Constraint operator solution to quantum billiard problems

D. A. McGrew* and W. Bauer
National Superconducting Cyclotron Laboratory and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824-1321
(Received 7 December 1995)

We introduce an additional method to solve Schrödinger's equation for a free particle in an infinite well of arbitrary shape (the Helmholtz equation with Dirichlet boundary conditions), a problem of interest in the area of quantum chaos. We expand the wave function in a basis of products of sine functions, then use the constraint operator to contain the wave function to a region within the domain of the basis functions. In this manner, a quantum billiard problem of arbitrary shape can be solved. Several methods exist to solve problems of this sort, but as recent work reviewing these methods has shown, all have shortcomings. Our work represents a different direction in the solution of these problems. Our method is different in that it provides a means of computing an eigenbasis. It is also interesting from a physical standpoint in that it can represent the Hamiltonian of a classically chaotic system in the basis of a classically regular system. [S1063-651X(96)06607-X]

PACS number(s): 02.70.Rw, 05.45.+b, 03.65.Ge, 02.10.Sp

I. INTRODUCTION

A billiard system consists of a particle bouncing around in a rigid box of arbitrary shape. Billiard systems are useful in the study of chaos, as the chaoticity of the system is determined by the shape of the box. Circles and squares give rise to regular motion; in more complicated shapes, like stadia, both regular and chaotic motion is possible, depending on the initial conditions [1,2].

Quantum billiard systems are widely used in the study of quantum chaos. Quantum chaotic systems can be characterized by statistics. The distribution of normalized energy level spacings of a quantum system is one such characterization; chaotic systems have Wigner distributions and regular systems have Poisson distributions [3–7]. The wave functions of quantum chaotic systems qualitatively resemble a random superposition of plane waves, though ''scars'' in the quantum wave functions corresponding to classical periodic orbits can appear [8].

The decay of quantum billiard systems through small exit channels is of current interest. The chaoticity of the billiard system controls the decay of the system; regular billiard systems decay algebraically in time, while chaotic billiard systems decay exponentially in time [9]. Recent work [10,11] shows an even richer variety of behaviors. The quantum chaotic billiard decay problem is yet unsolved; existing methods of solving quantum billiards are unsuited for it.

Many methods exist for solving quantum billiard problems. The most used are the boundary integral method [12–17], the plane wave decomposition method [8,18,19], and the conformal mapping diagonalization method [20–23]. However, recent work reviewing the boundary integral method [24] and the plane wave decomposition method [25] demonstrates that both have weaknesses.

The boundary integral method solves billiard problem by deriving an integral equation for the normal derivative of the wave function using Green's theorem [12–17]. Discretization of the boundary integral results in a complex determinant nonlinear in the wave vector magnitude, the zeros of which correspond to solutions of the wave function equation. It is widely used, but has recently been shown to have problems when the box geometry is nonconvex [24].

The plane wave decomposition method assumes an expansion in plane waves with the same wave vector magnitude, then tries to force the wave function to be zero along the boundary of the box by proper selection of the plane wave components [8,18,19]. If it succeeds in making the wave function approximately zero on the boundary, then it has found an approximate eigenfunction of $\vec{\nabla}^2$. The procedure iterates over wave vector magnitudes, recording the eigenvalues that it finds. It is widely used to find quantum billiard wave functions, but cannot be relied upon for accurate spectra, as some eigenvalues can be stepped over in the iteration process. The wave functions it finds are not necessarily orthogonal to very good accuracy, as shown in [25].

The conformal mapping diagonalization method elegantly solves a billard problem by finding a conformal map from the shape of the box to the unit circle [20–23]. The problem is solved by the mapping, but this method is limited to two dimensional problems with boxes for which a conformal mapping to the unit disk can be found.

We present a method that solves the problem in a more ''quantum mechanical'' way. We find many eigenfunctions simultaneously by diagonalizing a Hamiltonian matrix. This results in a (truncated) complete set of eigenfunctions that are necessarily orthonormal (within the limitations of the diagonalization algorithm), unlike the methods mentioned above. The availability of a complete basis provides a straightforward approach to time-dependent problems like quantum chaotic decay. We also connect this method to some existing methods.

Random matrices that are band diagonal are of interest in quantum chaos [26]. Band diagonal Hamiltonians are ''natural'' in the sense that many systems have localized interactions and localized wave functions. Our method, introduced below, results in an approximately band diagonal Hamil-

*Electronic address: mcgrew@nscl.msu.edu, bauer@nscl.msu.edu

 © 1996 The American Physical Society

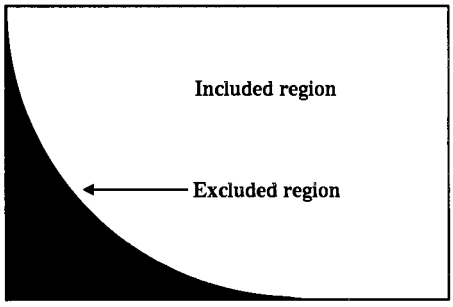

FIG. 1. The included and excluded regions shown for the desymmetrized stadium. The total region is the entire rectangle. The ratio of included to total regions μ here is 0.8570.

tonian. A connection between the two may prove revealing.

II. THE CONSTRAINT OPERATOR

Schrödinger's equation for the quantum billiard is

$$\vec{\nabla}^2\Psi(\vec{x})+\lambda\Psi(\vec{x})=0, \tag{1}$$

where λ is an energy eigenvalue of the system in some convenient units. The boundary condition, which characterizes the quantum billiard problem, is that $\Psi(\vec{x})=0$ on the boundary surface δI of an arbitrarily shaped region I, the included region.

We express the shape of I by starting with a larger region T where we can solve (1), then ''cutting away'' the unwanted parts of T to make I. We do this by constraining the wave function to be zero in the excluded region $E=T/I$. Figure 1 shows the regions I and E for the stadium billiard. (Most of our examples are two dimensional, but the method can be applied to three dimensions or higher.) In practice, we choose the region T so that it has a boundary with surfaces of constant coordinates in a coordinate system where $\vec{\nabla}^2$ is separable.

Define the constraint operator $\mathcal{C}$ for $\vec{x}$ in T, which multiplies functions on its right by the constraint function

$$c(\vec{x})=\begin{cases}1 & \text{if } \vec{x} \text{ is in } I\\ 0 & \text{if } \vec{x} \text{ is in } E.\end{cases} \tag{2}$$

The constraint operator is the projector for functions over the larger, simpler region T that are zero over E (it ''constrains'' functions to be zero in E.) Functions in the range of $\mathcal{C}$ are zero in E. Functions in the null-space of $\mathcal{C}$ are zero in the region I. $\mathcal{C}^2=\mathcal{C}$, so that $\mathcal{C}$ is idempotent and has eigenvalues 0 and 1. Therefore, it is the projector of its range, and $1-\mathcal{C}$ is the projector of its null-space. We define the included fraction $\mu=\int_I d\vec{x}/\int_T d\vec{x}$ as the ratio of the included volume to the total volume; we will use it below.

$\mathcal{C}$ is represented in the basis $\{\phi_i(\vec{x})\}$ of functions over T as the matrix C with elements

$$C_{ij}=\int_T d\vec{x}\,\phi_i(\vec{x})\mathcal{C}\phi_j(\vec{x})=\int_E d\vec{x}\,\phi_i(\vec{x})\,\phi_j(\vec{x}). \tag{3}$$

We will return to the matrix C later to discuss its properties and its computation.

The solutions to the problem we are interested in are the eigenvectors of $\vec{\nabla}^2$ that are in the range of $\mathcal{C}$. Intuitively, it seems that we can solve our problem by finding the eigenvectors of $\mathcal{C}\vec{\nabla}^2\mathcal{C}$. This is *almost* correct. To be completely correct, we will derive the solutions using a Green's function [27] then simplify the result using our knowledge of the constraint operator.

We need a Green's function $G(\vec{x},\vec{x}_0)$ that satisfies the equation

$$(\vec{\nabla}^2+\lambda)G(\vec{x},\vec{x}_0)=-4\pi\delta(\vec{x}-\vec{x}_0). \tag{4}$$

Multiplying (1) by $G(\vec{x},\vec{x}_0)$ and (4) by $\Psi(\vec{x})$, then subtracting the second result from the first gives

$$G(\vec{x},\vec{x}_0)\vec{\nabla}^2\Psi(\vec{x})-\Psi(\vec{x})\vec{\nabla}^2G(\vec{x},\vec{x}_0)=4\pi\Psi(\vec{x})\,\delta(\vec{x}-\vec{x}_0). \tag{5}$$

Integrating (5) over the included region I we obtain an expression for the wave function

$$\frac{1}{4\pi}\int_I d\vec{x}[G(\vec{x},\vec{x}_0)\vec{\nabla}^2\Psi(\vec{x})-\Psi(\vec{x})\vec{\nabla}^2G(\vec{x},\vec{x}_0)]=\begin{cases}\Psi(\vec{x}_0) & \text{if } \vec{x}_0 \text{ is in } I\\ 0 & \text{if } \vec{x}_0 \text{ is in } E.\end{cases} \tag{6}$$

To enforce the boundary condition that $\Psi=0$ on δI we use Green's theorem to change the volume integral into a surface integral, then apply the boundary condition, resulting in

$$\Psi(\vec{x}_0)=\frac{1}{4\pi}\int_S d\mathbf{A}\cdot[G(\vec{x},\vec{x}_0)\vec{\nabla}\Psi(\vec{x})] \tag{7}$$

for $\vec{x}_0$ in I. We then use Green's first identity to return to a volume integral,

$$\Psi(\vec{x}_0)=\frac{1}{4\pi}\int_I d\vec{x}[\vec{\nabla}G(\vec{x},\vec{x}_0)\cdot\vec{\nabla}\Psi(\vec{x})+G(\vec{x},\vec{x}_0)\vec{\nabla}^2\Psi(\vec{x})] \tag{8}$$

which is true for $\vec{x}_0$ in I.

We choose the basis $\{\phi_i(\vec{x})\}$ of eigenfunctions of $\vec{\nabla}^2$ over T. To express $G(\vec{x},\vec{x}_0)$ in this basis, we use (4) to determine the expansion coefficients, utilizing the eigenfunction expansion for $\delta(\vec{x}-\vec{x}_0)$. This results in

$$G(\vec{x},\vec{x}_0)=4\pi\sum_i\frac{\phi_i(\vec{x})\,\phi_i(\vec{x}_0)}{\lambda_i-\lambda}. \tag{9}$$

The wave function Ψ has the expansion

$$\Psi(\vec{x})=\sum_k\psi_k\phi_k(\vec{x}). \tag{10}$$

Now we put everything together. By substituting (9) and (10) into (8), we get the eigenvalue equation

$$\sum_k \psi_k \phi_k(\vec{x}_0) = \sum_{j,k} \frac{\psi_j}{\lambda_k - \lambda} \phi_k(\vec{x}_0) \int_I d\vec{x} [\vec{\nabla} \phi_k(\vec{x}) \cdot \vec{\nabla} \phi_j(\vec{x}) + \phi_k(\vec{x}) \vec{\nabla}^2 \phi_j(\vec{x})]. \quad (11)$$

This can be simplified by converting it to matrix form. To effect this change, we first multiply both sides by $\phi_i(\vec{x}_0)$, and use the orthonormality of the basis functions over the region T. Then

$$\psi_i = \sum_j \left[\int_I d\vec{x} \vec{\nabla} \phi_i \cdot \vec{\nabla} \phi_j - C_{ij} \lambda_j \right] \frac{1}{\lambda_i - \lambda} \psi_j = \sum_j [\delta_{i,j} \lambda_i - S_{ij} - C_{ij} \lambda_j] \frac{1}{\lambda_i - \lambda} \psi_j, \quad (12)$$

where

$$S_{ij} = \int_E d\vec{x} \nabla \phi_i \cdot \nabla \phi_j. \quad (13)$$

Here we have used the fact that

$$\int_I d\vec{x} \nabla \phi_i \cdot \nabla \phi_j = \int_T d\vec{x} \nabla \phi_i \cdot \nabla \phi_j - \int_E d\vec{x} \nabla \phi_i \cdot \nabla \phi_j = \lambda_i \delta_{i,j} - \int_E d\vec{x} \nabla \phi_i \cdot \nabla \phi_j. \quad (14)$$

The matrix C arises naturally in this derivation; the right-hand side of (6) reflects the fact that Ψ in the range of $\mathcal{C}$.

We now rearrange the equation into conventional matrix form,

$$(C\Lambda + S - \lambda)\psi = 0, \quad (15)$$

where Λ is the diagonal matrix containing the eigenvalues of the basis functions over T. Solving this eigensystem gives us a solution to (1) in the basis of eigenfunctions over T. Notice that the matrix in (15) is not Hermitian.

We now use our knowledge of C to make the eigensystem Hermitian. The wave function expansion (10) was equated to the right-hand side of (11), which is zero outside of I. Our wave function expansion actually represents a function that is equal to Ψ in I and equal to zero in E. Therefore the solutions to (15) are zero in E and are in the range of C, that is, $\mathcal{C}\Psi(\vec{x}) = \Psi(\vec{x})$ or $C\psi = \psi$. We can introduce this fact on the left-hand side of (8) to make the eigensystem Hermitian. The resulting generalized eigensystem is

$$(C\Lambda + \Lambda C - \Lambda + S - \lambda C)\psi = 0. \quad (16)$$

or equivalently,

$$[(C - \tfrac{1}{2}) \bigcirc H + S - \lambda C]\psi = 0, \quad (17)$$

where $H_{ij} = \lambda_i + \lambda_j$ and $\bigcirc$ represents the Hadamard or elementwise product defined by $(A \bigcirc B)_{ij} = A_{ij} B_{ij}$.

Alternatively we can modify the eigenvalue equation (15) by inserting C before ψ and multiplying by C from the left

$$C(C\Lambda + S - \lambda 1)C\psi = 0,$$

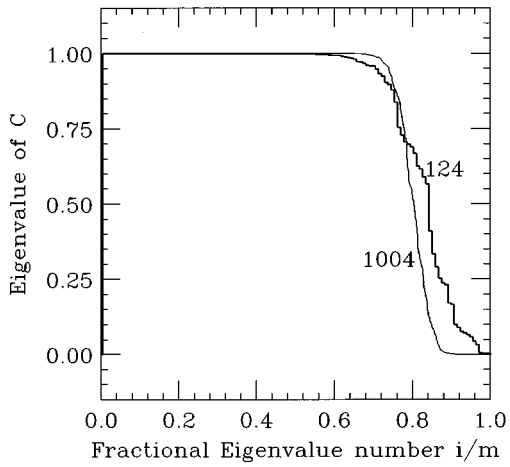

FIG. 2. The spectrum of C for the desymmetrized stadium, for $m = 124$ and $m = 1004$, where $r/m = \mu = 0.8570$.

$$(C\Lambda C + CSC - \lambda C)\psi = 0,$$

$$(C\Lambda C + CSC - \lambda)\psi = 0. \quad (18)$$

Here the idempotence of C makes the eigenproblem Hermitian.

The problem now looks ''quantum mechanical,'' as it has been reduced to finding the eigenvalues of a Hermitian matrix. We can express the problem as a perturbation of the original Hamiltonian Λ by using the complimentary projector $1 - C$, which is a ''small'' matrix in some problems. We will not pursue this here, as we are not interested in perturbative solutions.

Now we investigate the details involved in solving (18). As C is a projector, it can be decomposed as $C = PP^T$, where $P^T P = 1_{r \times r}$. P is an $m \times r$ matrix, where r is the rank of C. Then the eigensystem (18) is equivalent to

$$(P^T \Lambda P - P^T S P - \lambda)\psi' = 0, \quad (19)$$

where $\psi = P^T \psi'$ and ψ' is an r dimensional vector; there are r nontrivial solutions to (18).

As the constraint matrix for a region that is the union of two nonoverlapping regions must be the sum of the constraint matrices for the two regions from (3), we expect that the rank of a constraint matrix is proportional to the area in the range of the constraint matrix. When the included region is the total region the constraint matrix is the identity matrix, with rank m. Thus we expect that that the rank r of C is given by the closest integer to μm, which was confirmed numerically.

Before we solve (19), we need to find P; in other words, we need to find an orthonormal basis for the range of C. This matrix is only approximately a projector, due not to error in its elements but to the fact that it is representing $\mathcal{C}$ in a truncated basis; therefore its range is not well defined. We use the r eigenvectors of C with the largest eigenvalues to define the range of C. This step is justified in Appendices B and C. The spectrum of C proves to be close to that expected for it, so that the uncertainty in the range of C is not large. In Fig. 2, we plot the eigenvalues of C, in decreasing order, vs the fractional eigenvalue number, which is merely the eigen-

value number divided by the number of basis functions m. The spectrum is quite close to that of a projector with a fractional rank equal to the included fraction $\mu = \int_I d\vec{x} / \int_T d\vec{x}$.

III. NUMERICAL CALCULATION OF MATRIX ELEMENTS

Unless the boundary δI is along surfaces of constant coordinates we will need to use numerical integration to find the elements of C and S. This is because although the basis functions are analytically calculable, the integral of a basis function over an arbitrary region is not. We can integrate over E or I, as C and $1-C$ are trivially related. We have integrated over $1-C$ as the region is smaller, necessitating fewer grid points in the numerical integration. Thus we will find C and S by approximating (3) as a sum over all grid points in E of the Taylor expansion of the integrand.

To facilitate simple notation we will assume that $\vec{x}$ is two dimensional with components x_1, x_2. Consider a cell surrounding a grid point $\vec{x}_{\text{grid}}$ with dimensions h_1 and h_2. The integral of $\phi_i(\vec{x})\phi_j(\vec{x})$ over the cell is

$$\int_{-\frac{h_1}{2}}^{\frac{h_1}{2}} dx_1 \int_{-\frac{h_2}{2}}^{\frac{h_2}{2}} dx_2 \left[\phi_i(\vec{x})\phi_j(\vec{x}) \Big|_{\vec{x}_{\text{grid}}} + \sum_{k=1,2} \frac{\partial \phi_i(\vec{x})\phi_j(\vec{x})}{\partial x_k} \Big|_{\vec{x}_{\text{grid}}} x_k + \cdots \right]. \tag{20}$$

As we are using a regularly spaced grid, the integrals are over intervals centered around zero and odd powers of x_1 and x_2 integrate to zero and (20) becomes

$$h_1 h_2 \phi_i(\vec{x})\phi_j(\vec{x})|_{\vec{x}_{\text{grid}}}, \tag{21}$$

neglecting terms of order $h_1^3 h_2 + h_1 h_2^3$. Summing over all grid points,

$$C_{ij} = \sum_{k=1,}^{n} \phi_i(\vec{x}_k)\phi_j(\vec{x}_k) h_1^k h_2^k, \tag{22}$$

where $\vec{x}_k, k = 1,2, \ldots, n$ is a grid in the region E. h_1^k and h_2^k are the dimensions of the cell surrounding the kth grid point. If we define the $m \times n$ matrix $F_{ij} = \phi_i(\vec{x}_j)(h_1^k h_2^k)^{\frac{1}{2}}$, then $C_{\mathbf{E}} = FF^T$. Notice that FF^T is manifestly positive semidefinite, as expected. The matrices F and P should not be confused; even though $C = FF^T \simeq PP^T$, $P \neq F$. That the matrices cannot be equal is obviously true, as F is $m \times n$ and P is $m \times r$, and $n >> m \geq r$.

We observed that in some situations using grids with an odd number of grid points per side gave significantly better results than grids with an even number of grid points per side. We attribute this to the greater number of coincidental zeroes of the basis functions and grid points in the latter case, which degrades computation of the matrix elements. The number of coincidental zeroes and grid points can be lessened by using grids that begin at $h/2$, not 0. Our grids are evenly spaced, except at the boundary, where we put a point on the boundary and give it the appropriate weighting. The two dimensional numerical integration of the matrix elements merits no more consideration as the line integral computation of the matrix elements (discussed later) will be superior.

We can find the elements of S with a similar approach. From (13),

$$S_{ij} = h^2 \sum_{k,l} \left(\frac{\partial \phi_i}{\partial x_l} \frac{\partial \phi_j}{\partial x_l} \right)\Big|_{\vec{x}_k}, = \sum_{k,l} D_{ik}^l (D^l)_{kj}^T,$$

$$S = \sum_l D^l (D^l)^T, \tag{23}$$

where $D_{ik}^l = h(\partial \phi_i / \partial x_l)|_{\vec{x}_k}$. All of the matrices $D^l(D^l)^T$ are positive semidefinite.

For a function $\Psi(\vec{x})$ represented by a vector $\boldsymbol{\psi}$ to be in the null-space of FF^T the vector must satisfy

$$\sum_{j=1,}^{m} F_{kj}^T \boldsymbol{\psi}_j = 0, \quad (h_1 h_2)^{\frac{1}{2}} \sum_{j=1,}^{m} \phi_j(\vec{x}_k) \boldsymbol{\psi}_j = 0, \tag{24}$$

for $k = 1, \ldots, n$. Demanding Ψ to be in the range of $\mathcal{C}$ is thus equivalent to demanding that $\Psi(\vec{x}_k) = 0$ at the n grid points $\vec{x}_k$ used in the numerical integration. Only r of the equations (24) are linearly independent, giving rank r to the matrix C.

IV. COMPUTATIONAL ALGORITHM

Here we present a straightforward algorithm that solves (18); it uses a subroutine to find all of the eigenvalues and their associated eigenvectors. The only complication in this procedure is that finding all the elements of an $m \times m$ matrix that are equal to the product of three matrices (as is CSC) is an order m^4 process, which should be avoided. As the matrices will be large, the eigensystem solving routine should be one that uses only one array, such as Householder tridiagonalization followed by QR iteration. An uncomplicated algorithm that follows this advice and uses three storage arrays is as follows.

(a) In the first storage array we have the following.

(1) Compute the elements of C using numerical integration.

(2) Compute the eigenvalues and eigenvectors of C.

(3) Truncate the spectrum of C to find P such that $C = PP^T$ and $P^T P = 1$.

(b) In the second storage array we have the following.

(1) Compute the elements of S using numerical integration.

(2) Compute the eigenvalues and eigenvectors of S.

(3) Truncate the spectrum of S to find D such that $S = DD^T$.

(c) Compute $P^T D$ and write it into the third storage array.

(d) Compute $P^T \Lambda P$ (which is order m^3) and write the result in the second storage array.

(e) Compute $(P^T D)(P^T D)^T$ and add the result to $P^T \Lambda P$ in the second array.

(f) Compute the eigenvalues of $P^T \Lambda P + P^T S P$ which resides in the second array.

The algorithm solves three $m \times m$ eigensystems and three

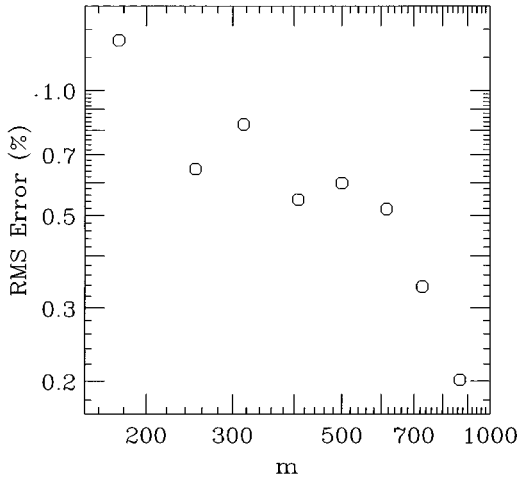

FIG. 3. The convergence of the rms fractional error in the energy eigenvalues as a function of the number of basis functions used for the disk problem. Here the number of grid points and the number of basis functions are equal.

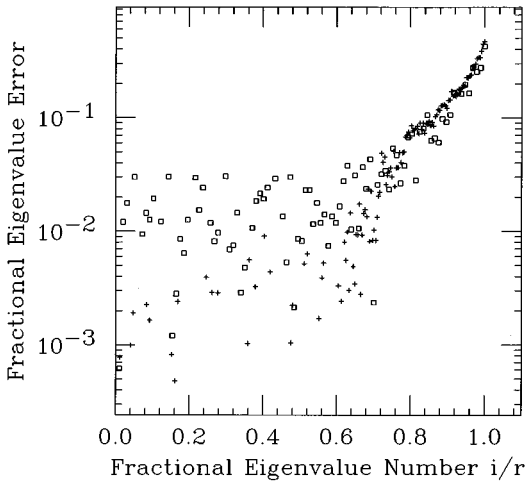

FIG. 4. The fractional eigenvalue error as a function of the fractional eigenvalue number for the unit disk with $m=124$ (boxes) and $m=316$ (pluses).

times computes the products of $m\times m$ matrices.

V. COMPARISON TO AN ANALYTIC CASE

To evaluate the accuracy of our method, we found the eigenfunctions of $\vec{\nabla}^2$ over a unit disk; we used a unit square for T and a unit disk for I. As the disk is ''separable,'' the eigenfunctions are known to be the product of sine functions and Bessel's functions. The eigenvalues are the zeroes of the integer order Bessel functions. We found the first 300 eigenvalues by finding the zeroes of the Bessel functions with a Newton-Raphson method for comparison.

Even though the circle and the square are both separable, their symmetries are quite different. Nevertheless the constraint operator solution works. With 868 basis functions, we obtained rms fractional eigenvalue error of 2.0138×10^{-3} for the first 300 eigenvalues (Fig. 3).

The errors of the eigenvalues depend on the eigenvalue number in a consistent way. To show this, we introduce the fractional eigenvalue number, which is the eigenvalue number i divided by the number of basis functions m. This enables us to put data for solutions with different m values on the same plot. Figure 4 shows the fractional error as a function of the fractional eigenvalue number. The dependence is nearly exponential for the higher eigenvalues, while the lower eigenvalues have fractional errors bounded by a constant.

We assume that $\Psi=0$ in E; we can easily find out how bad this assumption is. Define the volume fractional error ν to be the integral

$$\nu=\int_E d\vec{x}|\Psi|^2. \tag{25}$$

$0\leqslant\nu\leqslant1$, with only values near 0 being acceptable. The average volume fractional error for the first 300 eigenvalues was 7.113×10^{-3}. The volume fractional error also has a consistent distribution (Fig. 5.) The volume fractional error is roughly exponential in fractional eigenvalue number for low eigenvalue numbers.

We observed that the rms energy eigenvalue error converges to a finite value when n is increased but m is held fixed (Fig. 6). This is expected, as the matrix elements converge to their correct values with increasing n and there will always be a finite error when m is finite.

We observed that the rms energy eigenvalue error converges to zero as $m\rightarrow\infty$. The exponent of m was approximately -0.4 for the disk (Fig. 3).

The application of the constraint operator method to one dimensional problems is superfluous. For comparisons sake, we solved the one dimensional problem with a sine function basis. The rms fractional eigenvalue error decreased as m^{-1}. When the included region I is separable, the constraint operator method can be decomposed into two one dimensional problems [see Appendix C, Eqs. (C7) and (C8)]. Thus we expect that the rms fractional eigenvalue error in the two dimensional case can, at best, decrease as $m^{-1/2}$. The ob-

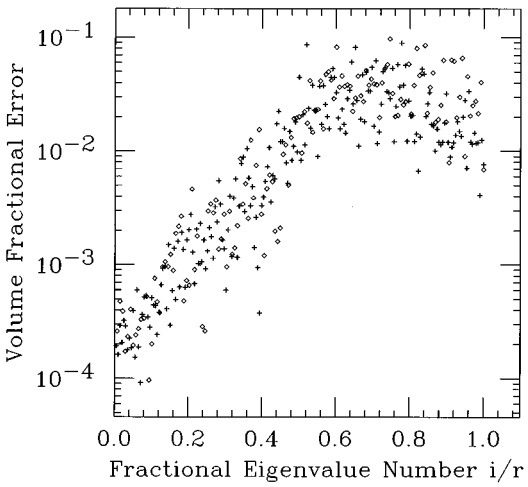

FIG. 5. The volume fractional error as a function of the fractional eigenvalue number for the unit disk with $m=124$ (boxes) and $m=316$ (pluses).

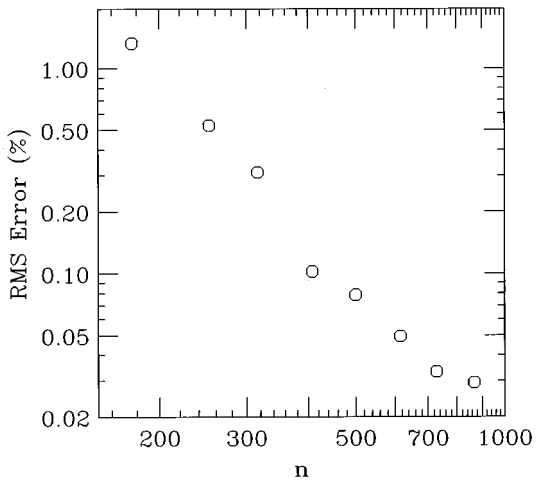

FIG. 6. The convergence of the rms fractional error in the energy eigenvalues as a function of the number of grid points used in the numerical integration for the disk problem, using 400 basis functions.

served exponent of m in the convergence of the eigenvalue error of the unit disk is consistent with this.

VI. COMPARISON WITH THE PLANE WAVE DECOMPOSITION METHOD

The plane wave decomposition method devised by Heller [18] uses a basis of plane waves to express the wave functions, then demands that the wave functions be zero on some points evenly spaced around the boundary surface. The eigenfunctions are found by an iterative process: an eigenvalue is assumed, the linear equations that set the wave functions equal to zero at the points are solved, then the error in the boundary condition is evaluated. The process is repeated until the error in the boundary condition vanishes; the wave function is then an eigenfunction of $\vec{\nabla}^2$ that satisfies the necessary boundary conditions.

Because of its iterative nature, it cannot be relied upon to find every eigenvalue [25], even though the error in the individual eigenvalues can be made arbitrarily small by varying the step size. The plane wave decomposition method finds eigenfunctions, but each eigenfunction is found in a different basis, as the wave vector magnitude is not the same for each eigenvector. This renders impracticable the task of manipulating the set of eigenvectors.

The constraint operator method has limited eigenvalue accuracy, but it finds a (truncated) complete set of eigenfunctions. The chore of knowing all m^2 matrix elements prevents the eigenvalue accuracy from being improved by using a large m. This is the price that is paid for finding a complete basis.

The relative merits of the plane wave decomposition method and the constraint operator method suggest that the methods could be used in conjunction. The plane wave decomposition method can be used to improve the eigenvalue accuracy of the constraint operator method, using the eigenvalues found by the latter as starting points. This would alleviate the plane wave decomposition method's problem of stepping over eigenvalues, and reduce the time spent searching for eigenvalues. The net result would be a complete set of eigenfunctions with accurate eigenvalues.

To check the method in a nonintegrable case, we also solved the desymmetrized stadium problem, where the included region is as shown in Fig. 1. No analytical solution is possible, so we compared our eigenvalues to those produced by the plane wave decomposition method for the same problem. With 378 basis functions, we obtained a rms fractional energy eigenvalue deviation of 0.002 for the first 100 eigenvalues of the two methods. The first four wave functions are plotted in Fig. 7, which shows the wave functions over the total region. We checked the distribution of normalized eigenvalue spacings in both of our test cases. The disk spacings closely follow a Poisson distribution, and the stadium spacings closely follow a Wigner distribution, as expected (Fig. 8). We regard this as evidence that the constraint operator solution preserves the essential details of a system.

VII. OTHER IDEAS AND FUTURE WORK

We are currently pursuing the approach of expanding the constraint function (2) in the eigenfunctions of $\vec{\nabla}^2$. The matrix elements of C and S are completely determined by the elements of this expansion. Fewer numerical integrations are required, and Green's theorem can be used to express the constraint function elements as integrals over the boundary surface δI. We expect better results from this approach as the problem of coincidental zeroes and grid points that can arise with the area integration will not be present when surface integrals are used.

This approach connects the constraint operator with the theory of Fourier expansions. Some of the properties of C mentioned above can be proven using this approach. The band diagonality of C and S is manifested, as a truncation of the constraint function expansion to m functions results in matrices with bandwidth 2 m.

The constraint operator method will work with arbitrary bases; a basis other than one that diagonalizes $\vec{\nabla}^2$ may result in a higher computational load but greater accuracy. In particular, we speculate that wavelets might do a better job of representing the constraint operator than the Fourier basis does.

We are planning to try our method on several quantum billiard problems that offer resistance to the standard methods, namely, the problems of scattering centers in a quantum well,triangular domains, and the problem of quantum chaotic decay.

APPENDIX A: OTHER OPERATORS, BASES, AND BOUNDARY CONDITIONS

Although we used a basis such that Λ was diagonal, this is not necessary. It may be possible to get better results by using other bases; for example, wavelets are more ''localized'' than plane waves and may better represent the constraint operator. Below we derive the correct eigensystem in a general basis. For complete generality, we assume that there is a quantum mechanical potential term $V(\vec{x})$.

Now we need a Green's function $G(\vec{x},\vec{x}_0)$ that satisfies the equation

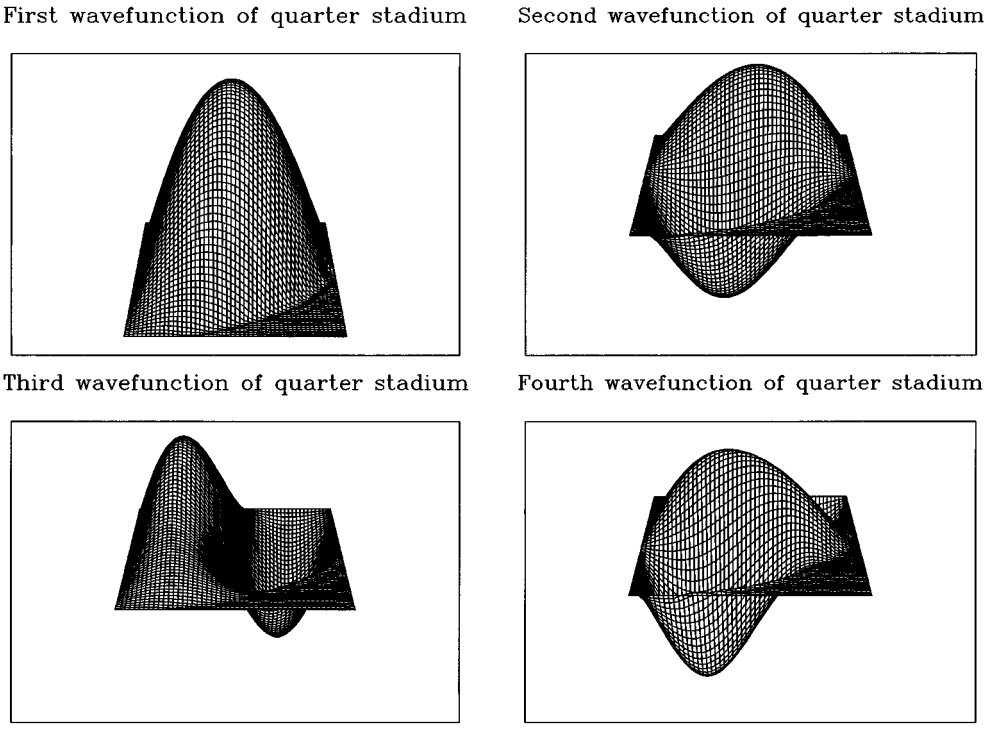

FIG. 7. $\Psi(\vec{x})$ for the first four eigenfunctions over the desymmetrized stadium. Something like the Gibbs phenomenon can be seen, with the maximum error occurring near, but not on, the boundary δI.

$$[\vec{\nabla}^2+V(x)+\lambda]G(\vec{x},\vec{x}_0)=-4\pi\delta(\vec{x}-\vec{x}_0). \quad \text{(A1)}$$

This leaves (5) through (7) unchanged. As our basis does not diagonalize $\vec{\nabla}^2$, this changes our expression for $G(\vec{x},\vec{x}_0)$ to

$$G(\vec{x},\vec{x}_0)=4\pi\sum_{ij}([T-V-\lambda]^{-1})_{ij}\phi_i(\vec{x})\phi_j(\vec{x}_0). \quad \text{(A2)}$$

The matrices T and V are given by

$$T_{ij}=-\int_T d\vec{x}\,\phi_i(\vec{x})\vec{\nabla}^2\phi_j(\vec{x}),$$

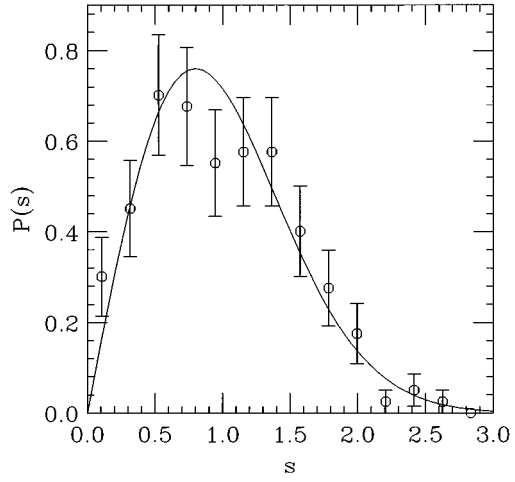

FIG. 8. The distributions of normalized energy spacings for the desymmetrized disk problem (eigenvalues 100–200) and the Wigner distribution.

$$V_{ij}=\int_T d\vec{x}\,\phi_i(\vec{x})V(x)\phi_j(\vec{x}). \quad \text{(A3)}$$

The expansion for Ψ is still

$$\Psi(\vec{x})=\sum_k \psi_k\phi_k(\vec{x}), \quad \text{(A4)}$$

but here the functions are not eigenfunctions of $\vec{\nabla}^2$. Substituting the expansions into (8),

$$\sum_k \psi_k\phi_k(\vec{x}_0)=\sum_{j,k,l}\psi_j([T-V-\lambda]^{-1})_{lk}\phi_k(\vec{x}_0)\times\int_S d\mathbf{A}\cdot[\phi_l\nabla\phi_j]. \quad \text{(A5)}$$

To convert this into matrix form, we multiply both sides by $\phi_i(\vec{x}_0)$ and then use the orthogonality of the eigenfunctions, giving

$$\psi_i=-\sum_{j,l}([T-V-\lambda]^{-1})_{li}A^T_{lj}\psi_j(T-V-A-\lambda 1)\psi=0, \quad \text{(A6)}$$

where $A_{ij}=\int_S d\mathbf{A}\cdot[\phi_j\vec{\nabla}\phi_i]$. A can be expressed in terms of C as

$$A_{ij}=\int_S d\mathbf{A}\cdot[\phi_i\vec{\nabla}\phi_j]=\int_I dv[\phi_i\vec{\nabla}^2\phi_j+\vec{\nabla}\phi_i\cdot\vec{\nabla}\phi_j]$$
$$=\int_T dv\left[\phi_i\mathcal{C}\vec{\nabla}^2\phi_j+\int_I dv\vec{\nabla}\phi_i\cdot\vec{\nabla}\phi_j\right]$$
$$=\sum_k C_{ik}(-T_{kj})+(S)_{ij}A=(1-C)T-S \tag{A7}$$

and the final form of the eigenvalue equation is

$$[C(T-V+S)C-\lambda]\phi=0. \tag{A8}$$

In the case of Neumann boundary conditions, $\vec{n}\cdot\vec{\nabla}\Psi(\vec{x})=0$ on the boundary surface, and (7) is replaced by

$$\Psi(\vec{x}_0)=\frac{1}{4\pi}-\int_S d\mathbf{A}\cdot[\Psi(\vec{x})\vec{\nabla}G(\vec{x},\vec{x}_0)], \tag{A9}$$

changing (A6) to

$$\psi_i=\sum_{j,l}([T-V-\lambda]^{-1})_{li}A^T_{lj}\psi_j(T-V+A-\lambda 1)\psi=0. \tag{A10}$$

The final form of the eigenvalue equation is then

$$[C(T-V-S)C-\lambda]\phi=0. \tag{A11}$$

APPENDIX B: APPROXIMATING A PROJECTOR

In (19) we choose the r eigenfunctions of $\mathcal{C}$ with the eigenvalues closest to unity to span the range of C. This is a sensible thing to do, but we would like a firmer justification for this step. Here we show that this step is equivalent to replacing C with the ''closest'' rank r projector C', namely, the one that minimizes the Euclidean matrix norm

$$\|C-C'\|_2. \tag{B1}$$

To show this, we write C and C' in terms of their unitary decompositions:

$$C=U^\dagger\Gamma U,\quad C'=V^\dagger\Gamma' V, \tag{B2}$$

where the diagonal matrix Γ_{ii} contains the eigenvalues of C, and $\Gamma'_{ii}=1$ if $i\leqslant r$ and is zero otherwise. Γ,Γ', and U are known to us; to solve the problem we want to find a unitary matrix V that minimizes

$$\|C-C'\|_2^2=\|\Gamma-UV^\dagger\Gamma' VU^\dagger\|$$
$$=\mathrm{Tr}[(\Gamma-UV^\dagger\Gamma' VU^\dagger)(\Gamma-UV^\dagger\Gamma' VU^\dagger)]$$
$$=\mathrm{Tr}[\Gamma^2+(\Gamma')^2-2\Gamma' UV^\dagger\Gamma VU^\dagger], \tag{B3}$$

which is equivalent to maximizing

$$\mathrm{Tr}[\Gamma' UV^\dagger\Gamma VU^\dagger]=\sum_{i=1,}^{m}\sum_{k=1,}^{m}\Gamma'_{ii}(UV^\dagger)_{ik}\Gamma_{kk}(VU^\dagger)_{ki}$$
$$=\sum_{i=1,}^{r}\sum_{k=1,}^{m}|(VU^\dagger)_{ik}|^2\Gamma_{kk}. \tag{B4}$$

Define the $m\times m$ matrix M by $M_{ij}=(VU^\dagger)_{ij}$, and note that it is orthostochastic (that is, $\Sigma_i M_{ij}=\Sigma_j M_{ij}=1$.) Define the vector γ by $\gamma_i=\Gamma_{ii}$. Then we want to maximize the sum

$$\sum_{i=1,}^{r}(M\gamma)_i. \tag{B5}$$

As M is orthostochastic, $M\gamma$ majorizes γ; the sum of the r largest elements of $M\gamma$ is less than or equal to the sum of the r largest elements of γ, with equality holding for $M=1$. We minimize the norm by taking $M=1$, which means that $U=V$. Thus C and C' are diagonal in the same basis, and replacing F with P is equivalent to replacing C with C'.

APPENDIX C: MATRIX STRUCTURE AND EIGENVALUE ERROR

In practice, the eigenvalue spectrum of C is observed to closely follow the function

$$\gamma(i)=\left[\exp\left(\frac{i'-\mu}{T}\right)+1\right]^{-\alpha}, \tag{C1}$$

where $i'=i/m$ is the ''fractional eigenvalue number'' and $\mu=r/m$ (Fig. 2). This function closely approximates the unit step function, which is the eigenvalue distribution expected for a projector. When $\alpha=1$, γ is the Fermi function; for $0<\alpha<1$, γ is greater than the Fermi function, and when $\alpha>1$, γ is less than the Fermi function; all cases are possible. In (Fig. 2) we see that $\gamma(\mu)$ is approximately the same for $m=124$ and $m=1004$, as expected, though the slope of γ at $i'=\mu$ is larger for larger m. The ''temperature'' T can be found using the slope $d\gamma/di'$ of (C1) at $i'=\mu$

$$T=\frac{-\alpha 2^{-\alpha-1}}{\left.\frac{d\gamma}{di'}\right|_\mu}=\frac{-\gamma(\mu)\log_2[\gamma(\mu)]}{\left.2\frac{d\gamma}{di'}\right|_\mu}. \tag{C2}$$

We observed that T is inversely proportional to m.

An essential fact about this method is that the matrices are approximately band diagonal in the basis of the eigenfunctions of $\vec{\nabla}^2$ arranged with their eigenvalues in increasing order. More precisely, they are well approximated by a matrix with all elements zero when $\|i-j\|>b$, where b is the bandwidth. This is important, as we need to represent infinite matrices with finite ones, and would like to understand exactly what we can safely neglect.

C can be expressed as the sum of a band diagonal matrix C_b with bandwidth b and a residual $\delta C=C-C_b$. The Weiland-Hoffman theorem [28] can now be used to bound the rms eigenvalue error of C_b

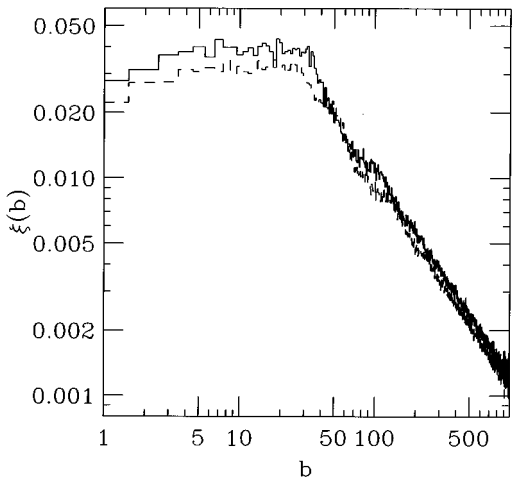

FIG. 9. The rms value ξ of the bth off-diagonal elements of C as a function of b, for the disk (upper curve) and stadium (lower curve), demonstrating the band diagonal nature of C.

$$\left[\frac{1}{m}\sum_{i=1,}^{m}[\lambda_i(C)-\lambda_i(C_b)]^2\right]^{1/2}\leqslant m^{-1/2}\|\delta C\|_2. \quad \text{(C3)}$$

Interestingly, $m^{-1/2}\|\delta C\|_2$ can be computed due to the idempotence of C

$$C_{ii}=\sum_j C_{ij}^2=\sum_{\|i-j\|\leqslant b} C_{ij}^2+\sum_{\|i-j\|>b} C_{ij}^2\|\delta C\|_2^2$$
$$=\sum_i\left[C_{ii}-\sum_{\|i-j\|\leqslant b} C_{ij}^2\right]. \quad \text{(C4)}$$

Here we have neglected the fact that m is finite, but this analysis works for large m.

It is easy to compute $m^{-1/2}\|\delta C\|_2$, and easy to update it if m is increased. A program can use this to determine the minimum value of m that has a certain maximum allowed rms error in the spectrum of C.

To investigate the rate of convergence of $m^{-1/2}\|\delta C\|$ to zero as b increases, we need a measure of the size of the kth off-diagonal. To account for the fact that different off-diagonals have different numbers of elements, we define ξ_k to be the rms value of the elements of the kth off-diagonal

$$\xi_k=\left[\frac{\Sigma_{|i-j|=k}C_{ij}^2}{N_k}\right]^{1/2}, \quad \text{(C5)}$$

where N_k is the number of elements in the kth off-diagonal. For large m, the bound on the rms eigenvalue error of C_b is

$$\frac{1}{m}\|\delta C\|_2^2=2\sum_{k=b,}^{\infty}\xi_k^2. \quad \text{(C6)}$$

We observed that $\xi_b\sim b^{-1}$ (Fig. 9), implying that the bound on the rms eigenvalue error of C_b is proportional to $b^{-1/2}$.

Whenever the wave function is separable, the matrices C and S must be separable as well, reducing the two dimensional problem to two one dimensional problems. When (1) can be solved by separation of variables, then the vector ψ will be decomposable into the Kroneker product $\psi=\xi\otimes\eta$. As there are separable eigenvectors $\psi=\xi\otimes\eta$, there are separable unitary matrices that diagonalize C and S. Therefore C can be written as the Kroneker product $C=C^x\otimes C^y$ and S can be written as the Kroneker sum $S=S^x\otimes 1+1\otimes S^y$. As we choose the total region T to correspond to a separable case, we can always write T as the Kroneker sum of two matrices: $T=T^x\otimes 1+1\otimes T^y$. Considering all of these separabilities, the matrix in (18) becomes

$$(C^x\otimes C^y(T^x\otimes 1+1\otimes T^y+S^x\otimes 1+1\otimes S^y-\lambda\otimes 1)C^x\otimes C^y$$
$$=[C^x(T^x+S^x-\lambda^x)C^x]\otimes C^y+C^x\otimes[C^y(T^y+S^y$$
$$-\lambda^y)C^y], \quad \text{(C7)}$$

where we have written $\lambda=\lambda^x\otimes 1+1\otimes\lambda^y$ and made repeated use of the mixed product identity. The solutions are given by the Kroneker product of the one dimensional solutions ξ and η, where

$$[C^x(T^x+S^x-\lambda^x)C^x]\xi=0,\quad [C^y(T^y+S^y-\lambda^y)C^y]\eta=0. \quad \text{(C8)}$$

There is a well-recognized correspondence between nonchaoticity and separability; this correspondence extends to the matrices C and S.

APPENDIX D: CONNECTION WITH THE FINITE DIFFERENCE METHOD

The constraint operator solution can be connected to the finite difference method. Consider the change of basis given by $F\psi^*=\psi$. From (10), the function

$$\Psi(\vec{x})=\sum_{i=1,}^{m}\phi_i(\vec{x})\sum_{j=1,}^{n}F_{ij}\psi_j^*,$$
$$=\sum_{j=1,}^{n}\psi_j\sum_{i=1,}^{m}\phi_i(\vec{x})\phi_i(\vec{x})(h_1h_2)^{1/2},$$
$$=\sum_{j=1,}^{n}\psi_j\delta_j^m(\vec{x}-\vec{x}_j), \quad \text{(D1)}$$

where $\delta_j^m(\vec{x}-\vec{x}_j)$ is the unit normalized δ function expanded in a size m basis.

This corresponds to changing to a basis of n δ functions, which is qualitatively similar to the finite difference method with an $n\times n$ array. n must be considerably larger than m in order for the numerical integration of the matrix elements to be accurate, and we have not neglected terms in representing $\vec{\nabla}^2$, demonstrating the relative merit of the constraint operator method.

[1] Ya. G. Sinai, Russ. Math. Surveys **25**, 137 (1970).
[2] L. A. Bunimovich, Commun. Math. Phys. **65**, 295 (1979).
[3] M. V. Berry and M. Tabor, Proc. R. Soc. London Ser. A, **356** (1977).
[4] S. W. McDonald and A. N. Kaufman, Phys. Rev. Lett. **42**, 1189 (1979).
[5] G. Casati, F. Valz-Gris, and I. Guarneri, Lett. Nuovo Cimento **28**, 279 (1980).
[6] M. V. Berry, Ann. Phys. (N.Y.) **131**, 163 (1981).
[7] O. Bohigas, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. **52**, 1 (1984).
[8] E. Heller, P. O'Connor and J. Gehlen, Phys. Scr. **40**, 354 (1989).
[9] W. Bauer and G. Bertsch, Phys. Rev. Lett. **65** , 2213 (1990).
[10] H. Alt, H. -D. Gräf, H. L. Harney, R. Hofferbert, H. Lengeler, A Richter, P. Schardt, and H. A. Weidenmüller, Phys. Rev. Lett. **74**, 62 (1995).
[11] H. Alt, H. -D. Gräf, H. L. Harney, R. Hofferbert, H. Rehfeld, A Richter, and P. Schardt, Phys. Rev. E (to be published).
[12] M. V. Berry and M. Wilkinson, Proc. R. Soc. London Ser. A **392**, 15 (1984).
[13] P. K. Banerjee *The Boundary Element Methods in Engineering* (McGraw-Hill, New York, 1994).
[14] G. De Mey, Int. J. Numer. Methods Eng. **10**, 59 (1976).
[15] Y. Niwa, S. Kobayashi, and M. Kitahara, *Developments in Boundary Element Methods 2* (Appl. Sci. Publications, London, 1980), pp. 143-176.
[16] R. J. Riddell, J. Comp. Phys. **31**, 21 (1979); **31**, 42 (1979).
[17] G. R. C. Tai and R. P. Shaw J. Acoust. Soc. Am. **56**, 796 (1974).
[18] E. Heller, Phys. Rev. Lett. **53**, 1515 (1984).
[19] E. Heller, *Chaos and Quantum Systems*, Proceedings of the Les Houches Summer School, edited by M.-J. Giannoni, A. Voros, and J. Zinn-Justin (Elsevier, Amsterdam, 1991).
[20] M. Robnik J. Phys. A **17**, 1049 (1984).
[21] M. V. Berry and M. Robnik, J. Phys. A **19**, 649 (1986).
[22] T. Prosen and M. Robnik, J. Phys. A **26**, 2371 (1993).
[23] T. Prosen and M. Robnik, J. Phys. A **27**, 8059 (1994).
[24] B. Li and M. Robnik, J. Phys. A **27**, 5509.
[25] B. Li and M. Robnik (unpublished).
[26] G. Casati, I. Guarneri, F. M. Izrailev, L. Molinari, K. Zyczkowski, Phys. Rev. Lett. **72**, 2697 (1994).
[27] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[28] R. H. Horn, and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, England, 1985).