Error estimates, error bounds, and adaptive refinement in finite-element quantum theory

Sergio Fanchiotti
Physics Department, New York University, New York, New York 10003

Mark A. Rubin
Physics Department, Rockefeller University, New York, New York 10021
(Received 28 December 1990)

We derive a rigorous *a posteriori* bound on the error in energy eigenvalues obtained using C^1 finite elements, as well as several *a posteriori* error estimates, and test them numerically with potentials of analytically-known eigenspectrum. We also obtain numerical solutions, with error bounds and estimates, for the octic oscillator potential, illustrating the ability of the finite-element method to resolve-nearly degenerate states with extremely narrow splitting. The incorporation of adaptive refinement is shown to reduce the number of degrees of freedom needed to achieve a given level of accuracy.

I. INTRODUCTION

Despite its spectacular success in many fields of engineering and physics, the finite-element method (FEM) for the numerical solution of partial differential equations[1] is only beginning to be used for quantum-mechanical[2,3] and quantum-field-theoretic[4] problems. Such problems are characterized by the large number of degrees of freedom involved. Particularly if one envisions doing numerical quantum field theory in the Schrödinger picture,[4a] the only viable approach would seem to be one based on the notion of *adaptive refinement.* That is, once the numerical algorithm has obtained a relatively crude approximate solution to a problem, it must be capable of determining the error in this solution in each region of the relevant configuration space. Degrees of freedom can then be added where the error is greatest (and, if desired, removed where it is least). By proceeding in such an iterative manner it is possible to concentrate computational effort on those degrees of freedom which are most important for the overall accuracy of the solution, even if, at the outset, one has been unable to pick out these degrees of freedom using, e.g., physical intuition or symmetry arguments.

Needless to say, we are interested in the accuracy of our computations even if we do not plan on adaptive refinement. In either case, the FEM is an ideal tool. Since it gives an approximate solution which is defined throughout the configuration space, rather than at isolated points, one can insert the approximate solution back into the differential equation which is to be solved. The resulting function, the degree to which, at a given point in configuration space, the differential equation fails to be satisfied, is termed the *residual.* From the residual, functions can be obtained which bound or estimate errors in global quantities, such as energy eigenvalues, or which determine where to perform adaptive refinement.

If, in fact, an adaptive scheme is implemented, the use of the FEM is advantageous in yet another regard, in that it can easily accommodate local additions or deletions of degrees of freedom. We also note that the FEM can with little difficulty incorporate nonstandard boundary conditions, e.g., open boundary conditions corresponding to systems which can lose energy to the external environment.

In the present paper, we focus on error bounds, error estimates, and adaptive refinement in the context of time-independent quantum mechanics in one space dimension, leaving for subsequent work the application to open, time-dependent, higher-dimensional quantum-mechanical and quantum-field-theoretic systems. After giving a brief summary of the FEM, we derive a rigorous *a posteriori* bound on the error in the energy eigenvalues, as well as nonrigorous, though sharper, estimates of the error. We then obtain numerical solutions to two benchmark one-dimensional potentials of known eigenspectrum, the harmonic oscillator and the Morse potential, and compare our computed error bound and estimates with the actual errors. As an illustration of the high degree of precision attainable by the FEM with relatively little computational effort, we solve an octic oscillator potential containing pairs of nearly degenerate eigenstates. Examples are also given of the use of adaptive refinement.

II. FINITE-ELEMENT ESSENTIALS

We will describe the FEM directly in the context of the time-independent one-dimensional Schrödinger equation

$$\hat{H}\psi(x)=E\psi(x)\ , \tag{1}$$

where

$$\hat{H}=-\frac{1}{2m}\frac{\partial^2}{\partial x^2}+V(x)\ . \tag{2}$$

Since we are dealing with the time-independent equation we can, without any loss of generality, choose $\psi(x)$ to be real. We approximate the exact solution $\psi(x)$ by a weighted sum of a finite number of *shape functions* $N_i(x)$:

$$\psi(x)\approx\bar{\psi}(x)=\sum_{i=1}^{m}N_i(x)\psi_i \ . \tag{3}$$

The approximate solution $\bar{\psi}(x)$ satisfies

$$(\hat{H}-\bar{E})\bar{\psi}(x)=r(x) \ , \tag{4}$$

where $\bar{E}$ is an approximation to the exact energy eigenvalue E.

In general, we cannot demand that the residual $r(x)$ be identically zero, since we only have at our disposal the m parameters ψ_i. Instead, we demand that m *weighted residuals,* integrals of the residual multiplied by *weight functions* $W_i(x)$, vanish:

$$\int dx\ W_i(x)r(x)=0, \quad i=1,\ldots,m \ . \tag{5}$$

For several reasons, some of which we will mention below, it is usually preferable to take the W_i's equal to the N_i's; this is referred to as the "Galerkin FEM." The weighted residual equations are then

$$\sum_{j=1}^{m}\left[\int dx\ N_i(x)\hat{H}N_j(x)\right]\psi_j=\bar{E}\left[\int dx\ N_i(x)N_j(x)\right]\psi_j \ . \tag{6}$$

In all the problems with which we will deal in the present paper, we will impose Dirichlet boundary conditions on the solution at the ends of a finite interval $x\in[x_L,x_R]$. The most straightforward way to enforce these conditions is to choose shape functions which satisfy them:

$$N_i(x_L)=N_i(x_R)=0, \quad i=1,\ldots,m \ . \tag{7}$$

We can then integrate by parts in (6) to obtain the generalized algebraic eigenvalue problem

$$\sum_{j=1}^{m}H_{ij}\psi_j=\bar{E}\sum_{j=1}^{m}C_{ij}\psi_j \ , \tag{8}$$

where

$$H_{ij}=\frac{1}{2m}K_{ij}+V_{ij} \ , \tag{9}$$

$$K_{ij}=\int_{x_L}^{x_R}dx\frac{dN_i(x)}{dx}\frac{dN_j(x)}{dx} \ , \tag{10}$$

$$V_{ij}=\int_{x_L}^{x_R}dx\ N_i(x)V(x)N_j(x) \tag{11}$$

and

$$C_{ij}=\int_{x_L}^{x_R}dx\ N_i(x)N_j(x) \ . \tag{12}$$

K_{ij} and C_{ij} are termed the "stiffness" and "mass" matrices, respectively. Once H_{ij} and C_{ij} are known, the ψ_i's and $\bar{E}$ can be computed by any of several standard techniques, such as the shifted inverse iteration which we employ in Sec. VI.

One computational advantage of the Galerkin approach is seen to be that the matrices which result are symmetric. We note also that (8) would have resulted had we taken a Rayleigh-Ritz variational approach. The connection with the variational formulation is another useful feature; for example, it tells us that, apart from roundoff error, the approximation $\bar{E}$ we obtain for the energy of an eigenstate must be *above* the exact energy. These are not the only advantages of an FEM which can be expressed variationally; see, e.g., Ref. 5.

The FEM now imposes the following key requirements on the shape functions.

(1) They are *narrow based.* That is, each $N_i(x)$ vanishes everywhere except in a small compact region of configuration space. For any choice of $N_i(x)$, if $\hat{O}(x)$ is a local operator, it will obviously be true that

$$\int_{x_L}^{x_R}N_i(x)\hat{O}(x)N_j(x)=0 \tag{13}$$

unless either $i=j$, or i and j correspond to different shape functions with overlapping support. Since a weighted sum of $N_i(x)$'s must be capable of approximating functions defined over the entire interval $x\in[x_L,x_R]$, the left-hand side (LHS) of Eq. (13) will be zero for most pairs i,j. So the matrices which appear in the eigenvalue problem (8) will have the computationally advantageous property of being sparse and, if the shape functions are suitably ordered, banded with relatively narrow bandwidth.

(2) The shape functions are piecewise polynomial. This guarantees convergence to the exact solution as the number of shape functions is increased.[5,6]

(3) The shape functions have sufficient continuity so that the entries in the matrices (10)–(12) are finite. The only restriction comes from the stiffness matrix; the $N_i(x)$'s must be at least C^0, i.e., continuous but with first derivative not necessarily continuous.

To give a concrete illustration: divide up the interval $x\in[x_L,x_R]$ into $(m-1)$ nonoverlapping *elements;* element e consists of all $x\in[x_e,x_{(e+1)}]$, where the *nodes* x_e, $e=1,\ldots,m$, satisfy

$$x_L=x_1<x_2<\cdots<x_{m-1}<x_m=x_R \ . \tag{14}$$

For the present example we will let the $N_i(x)$ be the so-called standard C^0 shape functions; i.e., they are zero at all but a single node:

$$N_i(x_j)=\delta_{ij}, \quad i,j=1,\ldots,m \ . \tag{15}$$

The expansion coefficients then have a straightforward interpretation:

$$\psi_i=\bar{\psi}(x_i) \ . \tag{16}$$

Finally, we implement the piecewise polynomial requirement by demanding that, within elements $(i-1)$ and i, $N_i(x)$ be a polynomial of a given degree in $(x-x_i)$—for the purpose of our present example, of degree one. Then the shape functions are the "roof functions"

$$N_i(x)=\begin{cases}\dfrac{x-x_{i-1}}{x_i-x_{i-1}}, & x\in[x_{i-1},x_i],\\ \dfrac{x_{i+1}-x}{x_{i+1}-x_i}, & x\in[x_i,x_{i+1}],\\ 0, & \text{in all other elements} \ .\end{cases} \tag{17}$$

Convergence, as the number of elements is increased, is faster the higher the polynomial order of the shape func-

tions. Higher-order shape functions can be constructed in several ways;[1,7] we will describe one such construction in Sec. IV.

III. ENERGY-ERROR BOUND

Given any candidate solution $\bar{\psi}(x)$ to the time-independent Schrödinger equation and its associated estimated energy eigenvalue $\bar{E}$, we can in a straightforward manner put an upper bound[8,8a] on the absolute value of the difference between $\bar{E}$ and the nearest exact eigenvalue E_I. Expanding $\bar{\psi}(x)$ in terms of the exact energy eigenfunctions

$$\bar{\psi}(x)=\sum_I a_I\psi_I(x) , \tag{18}$$

where

$$\hat{H}\psi_I(x)=E_I\psi_I , \tag{19}$$

Eq. (4) becomes

$$\sum_I (E_I-\bar{E})a_I\psi_I(x)=r(x) . \tag{20}$$

We can always choose the ψ_I to be an orthonormal set. Squaring both sides of (20) and integrating then gives

$$\sum_I (E_I-\bar{E})a_I^2=\int_{x_L}^{x_R} dx[r(x)]^2 . \tag{21}$$

But,

$$\sum_I (E_I-\bar{E})^2a_I^2\geq \min_I(E_I-\bar{E})^2\sum_I a_I^2 . \tag{22}$$

So, if $\bar{\psi}(x)$ is normalized, (21) and (22) imply

$$\min_I(E_I-\bar{E})^2\leq\int_{x_L}^{x_R} dx[r(x)]^2 . \tag{23}$$

Equation (23) is not simply an estimate but is, rather, a rigorous bound on the possible error in the energy eigenvalue. Furthermore, it is an *a posteriori* bound; that is, it is obtained directly from the approximate solution $\bar{\psi}(x)$, $\bar{E}$, and does not require the solution of an additional linear system (compare Ref. 9).

However, it is clear that (23) is likely to be vacuous if we insist on building $\bar{\psi}(x)$ out of shape functions with only C^0 continuity. The kinetic energy operator $\partial^2/\partial x^2$, acting on $\bar{\psi}(x)$ at any point which is not at least C^1 will yield, in $r(x)$, a δ-function singularity, making the right-hand side of (23) infinite. To obtain a meaningful bound, we must thus use shape functions which are C^1 or smoother.

IV. HIERARCHICAL C^1 SHAPE FUNCTIONS

The shape functions of lowest order which retain C^1 continuity at the interelement nodes are the so-called Hermite cubics. With each node x_i are associated two shape functions, ${}^0N_i(x)$ and ${}^1N_i(x)$, each with support only in the two elements adjacent to x_i, satisfying

$${}^0N_i(x_i)=\frac{d\,{}^1N_i(x_i)}{dx}=1 , \tag{24a}$$

$$\frac{d\,{}^0N_i(x_i)}{dx}={}^1N_i(x_i)=0 , \tag{24b}$$

$${}^0N_i(x_j)=\frac{d\,{}^0N_i(x_j)}{dx}$$

$$={}^1N_i(x_j)=\frac{d\,{}^1N_i(x_j)}{dx}=0 \quad \text{for } i\neq j . \tag{24c}$$

If $\bar{\psi}(x)$ is constructed out of shape functions satisfying (24a)–(24c),

$$\bar{\psi}(x)=\sum_{i=1}^{n}[{}^0N_i(x)\,{}^0\psi_i+{}^1N_i(x)\,{}^1\psi_i] , \tag{25}$$

the discretized degrees of freedom ${}^0\psi_i$, ${}^1\psi_i$ will again have a direct interpretation:

$${}^0\psi_i=\bar{\psi}(x_i) , \tag{26a}$$

$${}^1\psi_i=\frac{d\bar{\psi}(x_i)}{dx} . \tag{26b}$$

In detail: in element i,

$${}^0N_i(x)=2(y_i)^3-3(y_i)^2+1 , \tag{27a}$$

$${}^1N_i(x)=h_i[(y_i)^3-2(y_{i-1})^2+y_i] , \tag{27b}$$

$${}^0N_{i+1}(x)=-2(y_i)^3+3(y_i)^2 , \tag{27c}$$

$${}^1N_{i+1}(x)=h_i[(y_i)^3-(y_1)^2] , \tag{27d}$$

where

$$h_i=x_{i+1}-x_i \tag{28a}$$

and

$$y_i=\frac{x-x_i}{h_i} \tag{28b}$$

(see Fig. 1.).

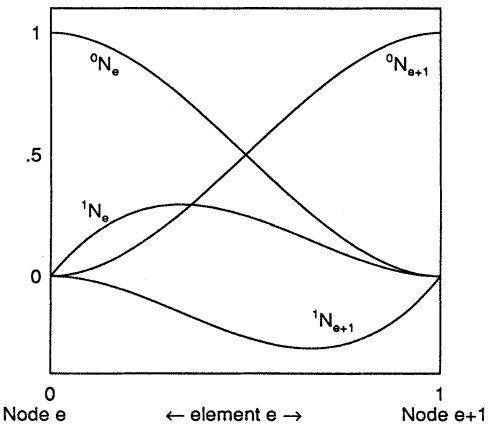

FIG. 1. C^1 standard shape functions in element e. The curves ${}^1N_e$ and ${}^1N_{e+1}$ are for $h_e=1$. (Note that ${}^0N_e$, ${}^1N_e$ also have support in element $e-1$, and ${}^0N_{e+1}$, ${}^1N_{e+1}$ in element $e+1$.)

To increase the polynomial order of the shape functions while retaining C^1 continuity at the interelement boundaries, we will use the "hierarchical" approach. If we want $\bar{\psi}(x)$ to be a piecewise polynomial of order $p>3$, we introduce, in each element e, shape functions ${}^hN_r^{(e)}(x)$, $r=4,\ldots,p$ which are polynomials of order r in $(x-x_e)$, vanish outside of element e and satisfy, at the boundaries of the element,

$$ {}^hN_r(x_i)=\frac{d\,{}^hN_r(x_i)}{dx}=0 \tag{29}$$

for $x_i=x_e$ or $x_i=x_{e+1}$. This condition, of course, does not determine uniquely the ${}^hN_r^{(e)}$'s. If we specify that the coefficient of $(x-x_e)^r$ is unity, then ${}^hN_4^{(e)}$ is fixed; for $r\geq 5$, ${}^hN_r^{(e)}$ is fixed by imposing, in addition, orthogonality in the derivative of the hierarchical shape functions:

$$\int_{x_e}^{x_{e+1}}dx\left[\frac{d\,{}^hN_r^{(e)}(x)}{dx}\right]\left[\frac{d\,{}^hN_{r'}^{(e)}(x)}{dx}\right]=0\ ,$$

$$4\leq r,r'\leq p,\quad r\neq r' \tag{30}$$

(see Fig. 2. for $r=4$ through 7). Alternatively, the ${}^hN_r^{(e)}$'s can be obtained as double integrals of Legendre polynomials.

The requirement (30) tends to reduce the condition number of the stiffness matrix, thereby contributing to a decrease in roundoff error. (When attempting an accurate solution, the condition number of the stiffness matrix is more significant in determining numerical stability than the condition number of the potential matrix, since the entries in K_{ij} will increase with decreasing h_i, while those of V_{ij} will decrease.) The approximate solution now has the form

$$\bar{\psi}(x)=\sum_{i=1}^{n}\left[{}^0N_i(x)\,{}^0\psi_i+{}^1N_i(x)\,{}^1\psi_i\right]+\sum_{e=1}^{n-1}\left[\sum_{r=4}^{p}{}^hN_r^{(e)}(x)\,{}^h\psi_r^{(e)}\right]\ . \tag{31}$$

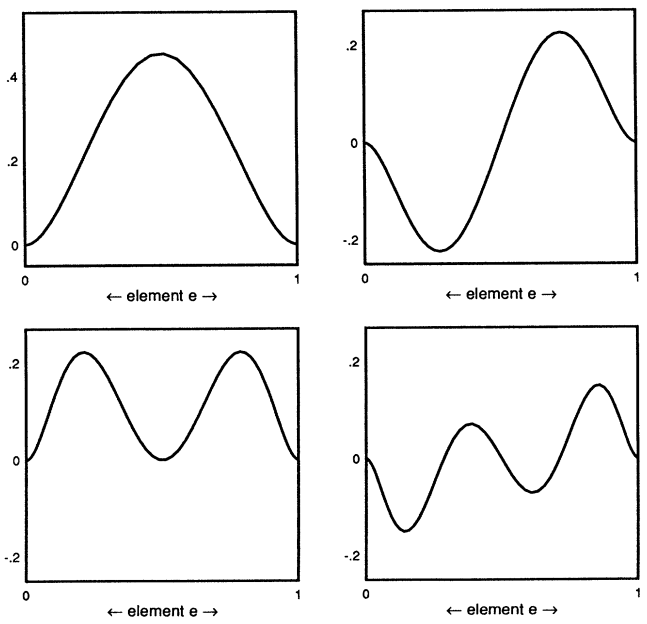

FIG. 2. C^1 hierarchical shape functions.

The fact that the hierarchical approach retains the interpretation (26a), (26b) of the coefficients of the standard shape functions ${}^0N_i(x)$, ${}^1N_i(x)$ is often a convenience but, of course, in no sense a necessity. More important is the utility of the hierarchical formulation for constructing local error estimators to be used in adaptive refinement, as we will see in the next section.

V. ENERGY-ERROR ESTIMATES

The error estimates we will present are extensions, to the case of eigenproblems, of an error estimate employed for boundary-value problems.[10,1,11] They are arrived at by considering a "perturbative" improvement of the FEM solution. Suppose $\bar{\psi}$ is an approximate finite-element solution to the equation

$$\hat{R}_E(x)=0\ , \tag{32a}$$

where

$$\hat{R}_E=\hat{H}-E\ . \tag{32b}$$

That is,

$$\bar{\psi}(x)=\sum_{i=1}^{m}N_i(x)\psi_i\ . \tag{33}$$

The sum in (33) is "shorthand" for an explicit expression such as (31); that is, $\bar{\psi}$ is a weighted sum of standard C^1 shape functions ${}^0N_i$, ${}^1N_i$, and hierarchical shape functions up to order p. The ψ_i's are determined by the weighted-residual conditions

$$\int_{x_L}^{x_R}dx\ N_i(x)r(x)=0,\quad i=1,\ldots,m\ , \tag{34}$$

with

$$r(x)=\hat{R}_{\bar{E}}\bar{\psi}(x)\ . \tag{35}$$

Following Ref. 1, we consider the improved solution $\tilde{\psi}(x)$ obtained by including a single additional hierarchical mode:

$${}^hN_k^{(e)},\quad k\geq p+1\ . \tag{36}$$

So,

$$\tilde{\psi}(x)=\sum_{i=1}^{m}N_i(x)\tilde{\psi}_i+{}^hN_k^{(e)}(x)\tilde{\psi}_k^{(e)}\ , \tag{37}$$

with

$$\int_{x_L}^{x_R}dx\ N_i(x)\tilde{r}(x)=0,\quad i=1,\ldots,m\ , \tag{38}$$

$$\int_{x_L}^{x_R}dx\ {}^hN_k^{(e)}(x)\tilde{r}(x)=0\ , \tag{39}$$

$$\tilde{r}(x)=\hat{R}_{\tilde{E}}\tilde{\psi}(x)=(\hat{H}-\tilde{E})\tilde{\psi}(x)\ . \tag{40}$$

If $\bar{\psi}(x)$ is already a good approximation, it is reasonable to assume that the coefficients of the first m shape functions are the same in $\tilde{\psi}(x)$ as is $\bar{\psi}(x)$,[1] so we may take

$$\tilde{\psi}(x)-\bar{\psi}(x)\approx{}^hN_k^{(e)}(x)\tilde{\psi}_k^{(e)}\ . \tag{41}$$

Using (38) and (39), the coefficient of the additional mode is found to be

$$\tilde{\psi}_k^e \approx -\left[\int_{x_L}^{x_R} dx\ {}^hN_k^{(e)}(x)\hat{R}_{\bar{E}}\ {}^hN_k^{(e)}(x)\right]^{-1}\int_{x_L}^{x_R} {}^hN_k^{(e)}(x)r(x)$$
$$\approx -\left[\int_{(e)} dx\ {}^hN_k^{(e)}(x)\hat{R}_{\bar{E}}\ {}^hN_k^{(e)}(x)\right]^{-1}\int_{(e)} {}^hN_k^{(e)}(x)r(x)\,, \tag{42}$$

where we can limit the integrals over x to the single element e in which the support of the hierarchical mode ${}^hN_k^{(e)}(x)$ lies.

In obtaining (42) we have, in addition to the approximation (41), taken

$$\tilde{E} \approx \bar{E}\ . \tag{43}$$

This is reasonable, since, for $\bar{\psi}(x)$ and $\tilde{\psi}(x)$ close to an exact solution, this difference is a *bilinear* functional of $\tilde{\psi}(x)-\bar{\psi}(x)$.

The pointwise error in $\bar{\psi}(x)$ is

$$\epsilon(x) \equiv \psi(x)-\bar{\psi}(x)\ . \tag{44}$$

Consistent with our assumption that $\bar{\psi}(x)$ is already close to $\psi(x)$, we take as an estimate for $\epsilon(x)$ the improvement in $\bar{\psi}(x)$ brought about by adding the single hierarchical mode ${}^hN_k^{(e)}(x)$. That is, we take

$$\tilde{\psi}(x) \approx \psi(x)\ . \tag{45}$$

Equations (41), (44), and (45) then give

$$\epsilon(x) \approx \psi(x)-\bar{\psi}(x) \approx {}^hN_k^{(e)}(x)\tilde{\psi}_k^{(e)}\ . \tag{46}$$

We can similarly consider the improvement brought about by adding several hierarchical modes; the same approximations as above result in

$$\epsilon(x) \approx \sum_e \sum_k {}^hN_k^{(e)}(x)\bar{\psi}_k^{(e)}\ . \tag{47}$$

However, if the error in the finite-element approximation to the wave function of the ith level is $\epsilon_i(x)$, then the magnitude of the difference between the approximate energy eigenvalue $\bar{E}_i$ and the corresponding exact energy eigenvalue E_i is bounded by[6]

$$|\bar{E}_i - E_i| \le \sum_{i' \le i} \|\epsilon_{i'}\|^2\ , \tag{48}$$

where

$$\|\epsilon\|^2 \equiv \int dx\ \epsilon(x)\hat{H}\epsilon(x)\ . \tag{49}$$

If the wave functions are reasonably accurate, (48) can usually be replaced by

$$|\bar{E}_i - E_i| \le \|\epsilon_i\|^2\ . \tag{50}$$

Using (32b), (42), (43), (47), (49), and (50), we obtain

$$|\bar{E}_i - E_i| \approx \sum_e \sum_{k,k'} \frac{\left[\int_{(e)} dx\ {}^hN_k^{(e)}r\right]\left[\int_{(e)} dx\ {}^hN_k^{(e)}\hat{H}\ {}^hN_{k'}^{(e)}\right]\left[\int_{(e)} dx\ {}^hN_{k'}^{(e)}r\right]}{\left[\int_{(e)} {}^hN_k^{(e)}\hat{R}_{E_i}\ {}^hN_k^{(e)}\right]\left[\int_{(e)} {}^hN_{k'}^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_{k'}^{(e)}\right]} \equiv \Delta_H\ . \tag{51}$$

An alternative estimate can be obtained using the formula for the lowest-order error in the eigenvalue induced by an error $\epsilon(x)$ in the eigenfunction,[6]

$$\bar{E}_i - E_i \approx \int dx\ \epsilon(x)(\hat{H}-E_i)\epsilon(x)\ . \tag{52}$$

Again using (32b), (42), and (43), we obtain, from (52),

$$\bar{E}_i - E_i \approx \sum_e \sum_{k,k'} \frac{\left[\int_{(e)} dx\ {}^hN_k^{(e)}r\right]\left[\int_{(e)} dx\ {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_{k'}^{(e)}\right]\left[\int_{(e)} dx\ {}^hN_{k'}^{(e)}r\right]}{\left[\int_{(e)} {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_k^{(e)}\right]\left[\int_{(e)} {}^hN_{k'}^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_{k'}^{(e)}\right]} \equiv \Delta_R\ . \tag{53}$$

Since, for small elements, the potential terms in $\int_{(e)} dx\ {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_{k'}^{(e)}$ and $\int_{(e)} dx\ {}^hN_k^{(e)}\hat{H}\ {}^hN_{k'}^{(e)}$ will be small, and since the ${}^hN_k^{(e)}$ have been chosen orthogonal in the derivatives, we may also hope to get good results by taking only the diagonal terms in (51) and (53):

$$\Delta_{H-D} \equiv \sum_e \sum_k \frac{\left[\int_{(e)} dx\ {}^hN_k^{(e)}r\right]^2\left[\int_{(e)} dx\ {}^hN_k^{(e)}\hat{H}\ {}^hN_k^{(e)}\right]}{\left[\int_{(e)} {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_k^{(e)}\right]^2} \tag{54}$$

$$\Delta_{R-D} \equiv \sum_e \sum_k \frac{\left[\int_{(e)} dx\ {}^hN_k^{(e)}r\right]^2}{\left[\int_{(e)} {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_k^{(e)}\right]}\ . \tag{55}$$

Finally, the Cauchy-Schwartz inequality tells us that Δ_{H-D} is bounded above by

$$\Delta_{H-D-\mathrm{CS}} \equiv \sum_e \sum_k \frac{\left[\int_{(e)} dx({}^hN_k^{(e)})^2\right]\left[\int_{(e)} dx\ {}^hN_k^{(e)}\hat{H}\ {}^hN_k^{(e)}\right]\left[\int_{(e)} dx\ r^2\right]}{\left[\int_{(e)} dx\ {}^hN_k^{(e)}\hat{R}_{\bar{E}_i}\ {}^hN_k^{(e)}\right]^2}\ . \tag{56}$$

VI. APPLICATIONS

As mentioned in Sec. II, we employ shifted inverse iteration[8] to obtain ψ_i and $\bar{E}$. Energy levels are computed successively, starting with the ground state and working upwards. At any given energy level, we start with an arbitrary $\psi_i^{(0)}$, and perform the iteration

$$\sum_j (H_{ij}-s_n C_{ij})\psi_j'^{(n+1)}=\sum_j C_{ij}\psi_j^{(n)} , \tag{57}$$

$$\psi_i^{(n+1)}=\sum_j P_{ij}\psi_j'^{(n+1)} , \tag{58}$$

where $\psi_i^{(n)}$ is the result for ψ_i after n iterations, P_{ij} is the projection operator which projects out the previously evaluated states, and the shift s_n is taken to be the expectation value of $\hat{H}$ at the nth iteration.[8a] (The eigenfunction is, of course, unchanged by an x-independent shift in the Hamiltonian.)

We begin by considering solutions using a fixed number of elements of equal size (that is, we do not yet implement adaptive refinement). Examining first the harmonic oscillator (Tables I–III), we find that $\Delta_{H-D-\mathrm{CS}}$ consistently overestimates the error, but that, for a certain number of the lowest-energy levels, the other four estimators are reasonably close to the actual error. The crossover from reasonable accuracy to significant inaccuracy occurs rather abruptly, near the point where Δ_R and Δ_{R-D} start to go negative. This type of behavior would be expected from the forms (51), (53), (54), (55) of the estimators. Each denominator contains factors of the form $\langle N_i|(\hat{H}-\bar{E})|N_j\rangle$ which will go through zero as $\bar{E}$ increases.

We can obtain a heuristic advance estimate for the maximum energy below which the estimators can be reliably used by looking for the smallest energy $\mathcal{E}$ such that

$$(\hat{H}-\mathcal{E})\,{}^hN_i^{(e)}=0 \tag{59}$$

for some hierarchial shape function ${}^hN_i^{(e)}$ which is used in the error estimator. Each ${}^hN_i^{(e)}$ vanishes outside of a single element and at the boundaries of that element. Its first derivative also vanishes at the boundaries of the element, and it is, of course, a polynomial; nevertheless, ${}^hN_i^{(e)}$ resembles the $\hat{i}$th eigenfunction $\phi_{\hat{i}}$ of the infinite square well of width h equal to the size of the element. If y is a local coordinate with the values 0, 1 at the ends of the element,

$${}^hN_i^{(e)}(y)\approx\phi_{\hat{i}}(y)=\sin(\hat{i}\pi y/h),\quad \hat{i}=i-3=1,2,\ldots . \tag{60}$$

If our solution was obtained using third-order polynomials, the lowest relevant state will be the square-well ground state; if fourth-order polynomials, the first excited state; and, in general,

$$\mathcal{E}=\frac{\pi^2(p-2)^2}{2mh^2} \tag{61}$$

if we are estimating the error to a solution obtained with pth-order polynomials.

This expression indicates the energy at which Δ_R and Δ_{R-D} become negative and the error estimates become essentially useless. It can be seen that the greatest accuracy is achieved at levels whose energy is less than

$$\mathcal{E}_0=\frac{\pi^2}{2mh^2} \tag{62}$$

no matter what order of polynomial is used in the solution. (This is presumably due to the increasing mismatch between the higher-level ${}^hN_i^{(e)}$'s and $\phi_{\hat{i}}$'s.)

We note that Δ_{R-D} seems to give slightly better results than the other estimators; and, of course, it and Δ_{H-D} require less computational effort than either Δ_R or Δ_H. (For each error estimator value presented, the number of hierarchical levels included was such that the fractional change due to the last level was less than 1%, up to a maximum order of 20. The highest order included in an estimator is listed under "*O*" in the tables.)

TABLE I. Harmonic oscillator, 30 elements of order 3, $X_L=-20.00$, $X_R=20.00$.

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	0.500034195335	0.304E-04	11	0.305E-04	11	0.314E-04	11	0.319E-04	11	0.524E-03	13	0.433E-01
1	1.500293454586	0.285E-03	9	0.272E-03	9	0.315E-03	9	0.318E-03	9	0.640E-02	11	0.134E+00
2	2.500886217776	0.902E-03	9	0.854E-03	9	0.107E-02	9	0.110E-02	9	0.225E-01	11	0.232E+00
3	3.504010617959	0.469E-02	7	0.441E-02	9	0.658E-02	7	0.664E-02	7	0.204E+00	9	0.544E+00
4	4.502471894751	0.260E-02	7	0.245E-02	7	0.374E-02	7	0.387E-02	7	0.117E+00	9	0.378E+00
5	5.521940842680	0.475E-01	7	0.334E-01	7	0.114E+00	5	0.113E+00	5	0.496E+01	7	0.134E+01
6	6.506225483588	0.160E-01	9	0.109E-01	9	0.787E-01	7	0.789E-01	7	0.992E+01	5	0.653E+00
7	7.546621920308	0.659E-01	9	0.499E-01	9	0.440E+01	5	0.440E+01	5	0.240E+03	5	0.201E+01
8	8.554183924713	-0.636E-01	9	-0.120E-02	20	0.899E+00	5	0.899E+00	5	0.384E+02	5	0.211E+01
9	9.564944720022	-0.689E+00	7	-0.423E+00	5	0.138E+03	5	0.138E+03	5	0.520E+04	5	0.233E+01
10	10.61250867836	-0.158E+00	11	-0.129E-02	20	0.311E+00	5	0.311E+00	5	0.233E+02	7	0.309E+01
11	11.66070850098	-0.482E-01	17	0.590E-01	17	0.130E+01	5	0.130E+01	5	0.798E+02	7	0.361E+01
12	12.69802938909	-0.164E+00	5	-0.505E-01	19	0.274E+01	5	0.274E+01	5	0.995E+02	7	0.391E+01
13	13.75435355007	-0.151E+00	15	0.568E-01	20	0.404E+00	11	0.395E+00	13	0.165E+02	9	0.440E+01
14	14.84085296377	-0.221E+00	15	0.642E-01	20	0.535E+00	13	0.478E+00	13	0.261E+02	9	0.504E+01
15	15.93572881946	-0.269E+00	5	-0.265E+00	5	0.130E+02	5	0.130E+02	5	0.114E+04	5	0.558E+01
16	17.02677456337	-0.316E+00	15	0.574E-01	20	0.148E+01	11	0.135E+01	10	0.529E+02	9	0.597E+01
17	18.12008919787	-0.184E+00	20	0.133E+00	18	0.838E+00	13	0.789E+00	12	0.392E+02	9	0.646E+01
18	19.24676407535	-0.163E+00	20	0.179E+00	18	0.108E+01	12	0.101E+01	12	0.526E+02	9	0.695E+01
19	20.34175444607	-0.388E-01	20	0.267E+00	16	0.171E+01	12	0.153E+01	10	0.133E+04	7	0.784E+01
20	21.53070663148	-0.119E+00	20	0.231E+00	16	0.223E+01	11	0.200E+01	10	0.131E+03	9	0.784E+01

TABLE II. Harmonic oscillator, 80 elements of order 3, $X_L=-20.00$, $X_R=20.00$.

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	0.500000179470	0.168D-06	6	0.168D-06	6	0.169D-06	6	0.169D-06	6	0.152E-05	13	0.629E-02
1	1.500001552852	0.149E-05	8	0.149E-05	8	0.150E-05	6	0.153E-05	8	0.141E-04	13	0.189E-01
2	2.500006766471	0.653E-05	8	0.651E-05	8	0.681E-05	8	0.682E-05	8	0.660E-04	13	0.403E-01
3	3.500020290313	0.197E-04	8	0.196E-04	8	0.208E-04	8	0.209E-04	8	0.214E-03	13	0.716E-01
4	4.500048034047	0.475E-04	10	0.464E-04	8	0.502E-04	8	0.503E-04	8	0.547E-03	13	0.113E+00
5	5.500096922081	0.966E-04	10	0.953E-04	10	0.105E-03	10	0.105E-03	10	0.119E-02	13	0.164E+00
6	6.500174607494	0.175E-03	10	0.172E-03	10	0.192E-03	10	0.193E-03	10	0.231E-02	13	0.225E+00
7	7.500289275735	0.295E-03	11	0.287E-03	10	0.325E-03	10	0.326E-03	10	0.411E-02	13	0.296E+00
8	8.500449508125	0.463E-03	11	0.447E-03	10	0.515E-03	10	0.517E-03	10	0.686E-02	13	0.377E+00
9	9.500664186240	0.691E-03	11	0.668E-03	11	0.776E-03	10	0.779E-03	10	0.107E-01	11	0.466E+00
10	10.50094242485	0.991E-03	11	0.953E-03	11	0.113E-02	11	0.113E-02	11	0.162E-01	11	0.565E+00
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
45	45.61348908668	0.770E+00	7	0.453E+00	7	0.676E+01	5	0.676E+01	5	0.339E+03	5	0.835E+01
46	46.68897880866	0.100E+01	7	0.600E+00	7	0.856E+01	5	0.856E+01	5	0.473E+03	5	0.105E+02
47	47.61770077767	0.268E+01	7	0.146E+01	7	0.133E+03	5	0.133E+03	5	0.659E+04	5	0.804E+01
48	48.68008492698	0.323E+01	5	0.323E+01	5	0.181E+04	5	0.181E+04	5	0.894E+05	5	0.107E+02
49	49.70467421388	0.248E+00	7	0.205E+00	7	0.587E+02	5	0.587E+02	5	0.467E+04	5	0.108E+02
50	50.65188433621	-0.146E+02	5	-0.146E+02	5	0.715E+05	5	0.715E+05	5	0.387E+07	5	0.941E+01
51	51.73005541314	-0.126E+01	5	-0.126E+01	5	0.386E+03	5	0.386E+03	5	0.187E+05	5	0.121E+02
52	52.73480963272	-0.235E+00	7	0.116E-01	17	0.575E+02	5	0.575E+02	5	0.390E+04	5	0.115E+02
53	53.69633929665	-0.415E+01	7	-0.219E+01	5	0.161E+04	5	0.161E+04	5	0.977E+05	5	0.108E+02
54	54.77379682297	0.224E+01	5	0.224E+01	5	0.121E+04	5	0.121E+04	5	0.587E+05	5	0.132E+02
55	55.78178124251	0.358E+00	9	0.298E+00	9	0.389E+02	5	0.389E+02	5	0.207E+04	5	0.127E+02

Table IV presents results for the Morse potential[12] with four bound states:

$$V_M(x)=16[\exp(-2x)-2\exp(-x)+1]\ . \tag{63}$$

The presence or absence of "edge effects" in this or other potentials can be checked by recomputing the problem on an array of elements coinciding with the original array, but extending beyond it.[2] (For greatest efficiency this should also be done adaptively, doing the computation on the second set of elements starting from the solution on the first set.) As an example of such a check, compare Tables IV(a) and IV(b). If we are only interested in levels 0 and 1 no further computations need be done, while the results for levels 2 and 3 should be checked on a still larger interval.

The rigorous bound, while applicable to all states and requiring less additional computation than any of the estimates, is seen to be conservative in the extreme. However, by increasing the polynomial order of the shape functions used in the solution and decreasing the element size, quite sharp bounds can be obtained. Table V presents results for the octic oscillator potential[13]

$$V_8(x)=(x^2-3)^4\ . \tag{64}$$

TABLE III. Harmonic oscillator, 40 elements of order 5, $X_L=-20.00$, $X_R=20.00$.

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	0.500000000338	0.313E-09	13	0.313E-09	13	0.314E-09	13	0.315E-09	13	0.734E-08	19	0.375E-03
1	1.500000004173	0.389E-08	13	0.387E-08	13	0.394E-08	13	0.395E-08	13	0.980E-07	19	0.135E-02
2	2.500000025397	0.238E-07	13	0.236E-07	13	0.244E-07	13	0.244E-07	13	0.640E-06	19	0.343E-02
3	3.500000109892	0.104E-06	13	0.102E-06	13	0.107E-06	13	0.107E-06	13	0.303E-05	19	0.737E-02
4	4.500000322519	0.293E-06	11	0.287E-06	11	0.306E-06	11	0.307E-06	11	0.930E-05	19	0.128E-01
5	5.500001061205	0.970E-06	11	0.937E-06	11	0.102E-05	11	0.102E-05	11	0.342E-04	19	0.242E-01
6	6.500001526689	0.140E-05	11	0.136E-05	11	0.149E-05	11	0.150E-05	11	0.481E-04	19	0.286E-01
7	7.500006834740	0.561E-05	9	0.537E-05	9	0.608E-05	9	0.609E-05	9	0.250E-03	19	0.639E-01
8	8.500002299094	0.199E-05	9	0.195E-05	9	0.217E-05	9	0.220E-05	9	0.698E-04	19	0.338E-01
9	9.500030331635	0.252E-04	9	0.238E-04	9	0.280E-04	9	0.281E-04	9	0.121E-02	19	0.137E+00
10	10.50000379459	0.371E-05	11	0.355E-05	11	0.402E-05	11	0.404E-05	11	0.125E-03	19	0.438E-01
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
45	45.65196588378	0.528E+00	8	0.523E+00	8	0.305E+01	7	0.305E+01	7	0.154E+02	9	0.497E+01
46	46.60156786799	0.576E+00	8	0.577E+00	8	0.510E+01	7	0.510E+01	7	0.244E+02	9	0.456E+01
47	47.70693241061	0.106E+01	8	0.106E+01	8	0.869E+01	7	0.869E+01	7	0.401E+02	9	0.580E+01
48	48.66498737680	0.960E+00	9	0.948E+00	8	0.115E+02	7	0.115E+02	7	0.543E+02	7	0.520E+01
49	49.67497864864	0.162E+01	9	0.161E+01	8	0.241E+02	7	0.241E+02	7	0.116E+03	7	0.564E+01
50	50.76045157281	0.265E+01	7	0.265E+01	7	0.738E+02	7	0.739E+02	7	0.348E+03	7	0.637E+01
51	51.70979331430	0.453E+01	7	0.453E+01	7	0.309E+03	7	0.309E+03	7	0.192E+04	7	0.583E+01
52	52.74904470906	-0.165E+01	9	-0.158E+01	8	0.622E+03	7	0.622E+03	7	0.679E+04	7	0.659E+01
53	53.82119567791	-0.650E+01	7	-0.652E+01	7	0.126E+04	7	0.126E+04	7	0.644E+04	7	0.706E+01
54	54.77707400744	0.538E+01	7	0.540E+01	7	0.135E+04	7	0.135E+04	7	0.600E+04	7	0.668E+01
55	55.82150180104	0.143E+00	9	0.281E+00	8	0.182E+03	7	0.182E+03	7	0.106E+04	7	0.750E+01

TABLE IV. (a) Morse potential, 80 elements of order 5, $X_L = -3.00$, $X_R = 20.00$. (Exact energies are 3.75, 9.75, 13.75, and 15.75.) (b) Morse potential, 100 elements of order 5, $X_L = -3.75$, $X_R = 25.00$.

(a)

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	3.750000000564	0.516E-09	13	0.519E-09	13	0.518E-09	13	0.522E-09	13	0.116E-07	19	0.165E-02
1	9.750000002323	0.215E-08	13	0.215E-08	13	0.217E-08	13	0.218E-08	13	0.489E-07	19	0.337E-02
2	13.75000000365	0.339E-08	13	0.339E-08	13	0.344E-08	13	0.345E-08	13	0.784E-07	19	0.425E-02
3	15.75000006785	0.176E-08	13	0.176E-08	13	0.179E-08	13	0.180E-08	13	0.410E-07	19	0.307E-02

(b)

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	3.750000000571	0.523E-09	13	0.526E-09	13	0.525E-09	13	0.528E-09	13	0.118E-07	19	0.166E-02
1	9.750000002354	0.218E-08	13	0.218E-08	13	0.219E-08	13	0.221E-08	13	0.500E-07	19	0.340E-02
2	13.75000000340	0.343E-08	13	0.343E-08	13	0.348E-08	13	0.349E-08	13	0.799E-07	19	0.429E-02
3	15.75000000236	0.178E-08	13	0.178E-08	13	0.181E-08	13	0.182E-08	13	0.417E-07	19	0.309E-02

TABLE V. Octic oscillator, 40 elements of order 10, $X_L = -4.00$, $X_R = 4.00$.

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	5.275264807242	0.905E-19	13	0.905E-19	13	0.905E-19	13	0.906E-19	13	0.190E-17	20	0.644E-07
1	5.275266882429	0.905E-19	13	0.905E-19	13	0.905E-19	13	0.906E-19	13	0.190E-17	20	0.644E-07
2	18.35624876859	0.778E-18	13	0.778E-18	13	0.779E-18	13	0.779E-18	13	0.161E-16	20	0.187E-06
3	18.35632536649	0.778E-18	13	0.778E-18	13	0.779E-18	13	0.779E-18	13	0.161E-16	20	0.187E-06
4	34.55668084695	0.384E-17	13	0.384E-17	13	0.385E-17	13	0.385E-17	13	0.786E-16	20	0.414E-06
5	34.55902456256	0.384E-17	13	0.384E-17	13	0.385E-17	13	0.385E-17	13	0.786E-16	20	0.414E-06
6	51.67287798620	0.131E-16	13	0.131E-16	13	0.132E-16	13	0.132E-16	13	0.267E-15	20	0.762E-06
7	51.72481220404	0.133E-16	13	0.133E-16	13	0.133E-16	13	0.133E-16	13	0.270E-15	20	0.766E-06
8	68.15033320527	0.285E-16	13	0.286E-16	13	0.287E-16	13	0.287E-16	13	0.584E-15	20	0.113E-05
9	68.91305247474	0.331E-16	13	0.331E-16	13	0.333E-16	13	0.333E-16	13	0.677E-15	20	0.121E-05
10	81.48791591123	0.394E-16	13	0.394E-16	13	0.396E-16	13	0.396E-16	13	0.818E-15	20	0.133E-05
11	86.14778659438	0.660E-16	13	0.660E-16	13	0.664E-16	13	0.664E-16	13	0.138E-14	20	0.173E-05
12	95.86981079236	0.849E-16	13	0.849E-16	13	0.855E-16	13	0.855E-16	13	0.183E-14	20	0.200E-05
13	104.8775504336	0.129E-15	13	0.129E-15	13	0.130E-15	13	0.130E-15	13	0.291E-14	20	0.251E-05
14	115.2467731573	0.192E-15	13	0.192E-15	13	0.194E-15	13	0.194E-15	13	0.461E-14	20	0.316E-05
15	126.1988707088	0.296E-15	14	0.296E-15	14	0.299E-15	14	0.299E-15	14	0.777E-14	20	0.410E-05
16	137.8634667294	0.607E-15	15	0.607E-15	15	0.613E-15	15	0.614E-15	15	0.164E-13	20	0.594E-05
17	150.1574216374	0.195E-14	13	0.195E-14	13	0.197E-14	13	0.197E-14	13	0.479E-13	20	0.101E-04
18	163.0652069075	0.575E-14	13	0.575E-14	13	0.582E-14	13	0.582E-14	13	0.130E-12	20	0.167E-04
19	176.5638703443	0.123E-13	13	0.123E-13	13	0.124E-13	13	0.124E-13	13	0.269E-12	20	0.239E-04
20	190.6372214796	0.255E-13	13	0.255E-13	13	0.258E-13	13	0.258E-13	13	0.533E-12	20	0.337E-04

TABLE VI. Octic oscillator, 45 elements of order 3, $X_L = -4.00$, $X_R = 4.00$.

n	energy	Δ_R	O	Δ_{R-D}	O	Δ_H	O	Δ_{H-D}	O	Δ_{H-D-CS}	O	bound
0	5.275279373898	0.127E-04	10	0.129E-04	10	0.128E-04	10	0.130E-04	10	0.157E-03	13	0.186E+00
1	5.275281444392	0.127E-04	10	0.129E-04	10	0.128E-04	10	0.130E-04	10	0.157E-03	13	0.186E+00
2	18.35640327136	0.142E-03	10	0.143E-03	10	0.147E-03	10	0.148E-03	10	0.171E-02	13	0.592E+00
3	18.35647982088	0.142E-03	10	0.143E-03	10	0.147E-03	10	0.148E-03	10	0.171E-02	13	0.592E+00
4	34.55747350507	0.757E-03	10	0.752E-03	10	0.802E-03	10	0.808E-03	10	0.987E-02	13	0.137E+01
5	34.55981755211	0.757E-03	10	0.752E-03	10	0.803E-03	10	0.808E-03	10	0.987E-02	13	0.137E+01
6	51.67540368502	0.249E-02	11	0.245E-02	11	0.270E-02	10	0.271E-02	10	0.361E-01	13	0.253E+01
7	51.72736902275	0.253E-02	11	0.248E-02	11	0.273E-02	10	0.275E-02	10	0.366E-01	13	0.254E+01
8	68.15551187789	0.525E-02	11	0.510E-02	11	0.583E-02	11	0.586E-02	11	0.846E-01	13	0.374E+01
9	68.91904926467	0.609E-02	11	0.591E-02	11	0.676E-02	11	0.680E-02	11	0.986E-01	13	0.403E+01
10	81.49514804937	0.750E-02	11	0.721E-02	11	0.845E-02	11	0.850E-02	11	0.129E+00	11	0.453E+01
11	86.16017460453	0.130E-01	11	0.124E-01	11	0.147E-01	11	0.148E-01	11	0.230E+00	11	0.597E+01
12	95.88663337244	0.174E-01	9	0.170E-01	11	0.201E-01	9	0.202E-01	9	0.338E+00	11	0.707E+01
13	104.9035465940	0.273E-01	9	0.265E-01	11	0.319E-01	9	0.321E-01	9	0.561E+00	11	0.891E+01
14	115.2840729369	0.400E-01	9	0.385E-01	11	0.473E-01	9	0.476E-01	9	0.875E+00	11	0.108E+02
15	126.2522516492	0.585E-01	9	0.541E-01	9	0.702E-01	9	0.706E-01	9	0.137E+01	11	0.132E+02
16	137.9382813679	0.841E-01	9	0.768E-01	9	0.102E+00	9	0.103E+00	9	0.212E+01	11	0.158E+02
17	150.2605679104	0.119E+00	9	0.107E+00	9	0.148E+00	9	0.149E+00	9	0.324E+01	11	0.188E+02
18	163.2049733358	0.167E+00	9	0.148E+00	9	0.211E+00	9	0.213E+00	9	0.493E+01	11	0.222E+02
19	176.7498384025	0.231E+00	9	0.201E+00	9	0.298E+00	9	0.300E+00	9	0.727E+01	9	0.260E+02
20	190.8791158656	0.281E+00	7	0.267E+00	9	0.393E+00	7	0.396E+00	7	0.107E+02	9	0.300E+02

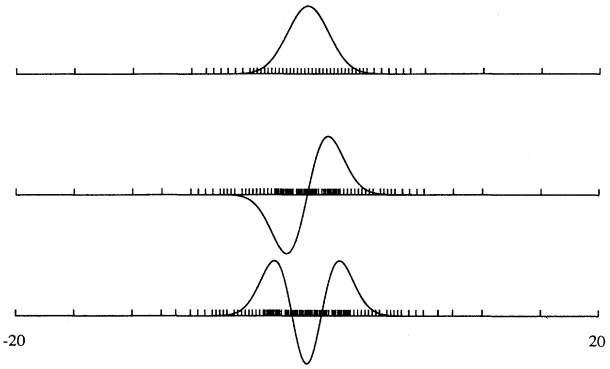

FIG. 3. Harmonic-oscillator wave functions computed adaptively, showing final distributions of elements; see Table VII.

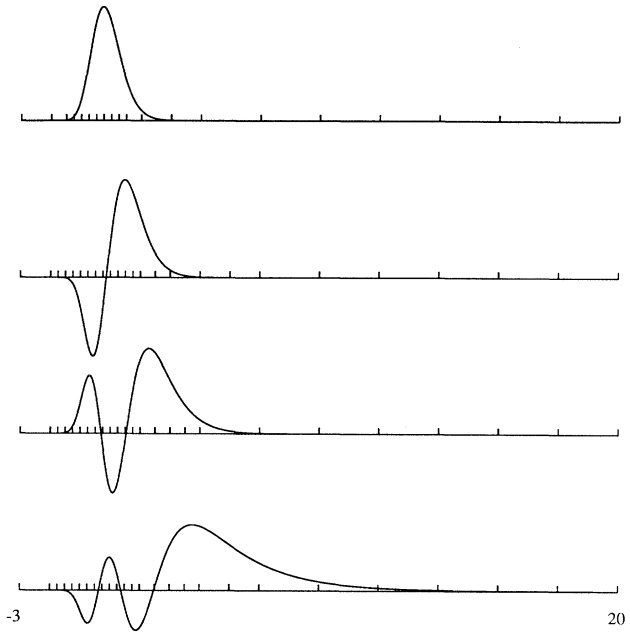

FIG. 4. Morse potential wave functions computed adaptively, showing final distributions of elements; see Table VIII.

The separation in energy between the two lowest-lying states, of order 10^{-6}, is clearly seen; the error in each of these states is bounded at a level of 10^{-8}. The values of the *estimators* in Table V should be ignored, as they are comparable to the roundoff error. (The roundoff error can usually be taken to be equal to the limit set by the number of bits used to represent the mantissa of a double-precision number; on the machines employed for the present work, this gives a maximum accuracy of about a part in 10^{15}. If required, a more careful estimate of the roundoff error could be made by using *forward* iteration to compute the *largest* eigenvalues of the matrices H and C, thus obtaining their condition numbers.[6,8]) The energy levels of the octic oscillator are not analytically known, but the results of the high-precision computation in Table V can be taken as exact in comparison with the less-precise results in Table VI. The error estimators in Table VI are again seen to indicate well the actual error in the energy.

To perform adaptive refinement, we evaluate a measure of the error in each element: the contribution from a given element to one of the Δ's, or to the bound (i.e., $\int_{(e)} r^2$). After obtaining a solution with a given set of elements, we divide in two all elements in which the estimated error is greater than or equal to some fixed fraction of the largest error estimate in an element.[10,11] (In the examples below, we take this fraction to be 0.5.) The current approximation to the wave function is then mapped by point collocation[1] to the augmented set of degrees of freedom corresponding to the increased number of elements; this serves as a starting point for further shifted inverse iteration, leading to a new, more accurate solution. This process is repeated until the desired accuracy, as determined by one of the *global* error measures (not necessarily the same one used as a refinement indicator) is obtained.

The adaptively evaluated wave functions, and the final distributions of elements for the harmonic oscillator and Morse potential are shown, respectively, in Figs. 3 and 4. The corresponding energy eigenvalues, global error measures, and final numbers of elements employed are shown in Tables VII(a) and VIII. Both computations began with ten elements of equal size; refinement was controlled by the element contributions to Δ_{H-D}. All of the error measures produce similar results when used as adaptive indicators. Comparing Table VII(a) to Table VII(b), or

TABLE VII. (a) Harmonic oscillator, adaptive computation; see Fig. 3. Order=4, $X_L=-20.00$, $X_R=20.00$. (b) Harmonic oscillator, 80 elements of order 4, $X_L=-20.00$, $X_R=20.00$.

(a)

n	energy	final # of elements	Δ_R	Δ_{R-D}	Δ_H	Δ_{H-D}	Δ_{H-CS-D}	bound
0	0.500000000001	52	0.871E-12	0.872E-12	0.872E-12	0.873E-12	0.888E-11	0.623E-04
1	1.500000000001	78	0.594E-12	0.595E-12	0.596E-12	0.597E-12	0.499E-11	0.572E-04
2	2.500000000001	88	0.949E-12	0.949E-12	0.951E-12	0.952E-12	0.672E-11	0.874E-04

(b)

n	energy	Δ_R	Δ_{R-D}	Δ_H	Δ_{H-D}	Δ_{H-D-CS}	bound
0	0.500000000157	0.138E-09	0.138E-09	0.138E-09	0.138E-09	0.375E-08	0.451E-03
1	1.500000001759	0.155E-08	0.155E-08	0.156E-08	0.156E-08	0.413E-07	0.149E-02
2	2.500000009936	0.881E-08	0.880E-08	0.887E-08	0.888E-08	0.230E-06	0.351E-02

TABLE VIII. Morse potential, adaptive computation; see Fig. 4. Order=5, $X_L=-3.00$, $X_R=20.00$. (Exact energies are 3.75, 9.75, 13.75, and 15.75.)

n	energy	final # of elements	Δ_R	Δ_{R-D}	Δ_H	Δ_{H-D}	Δ_{H-CS-D}	bound
0	3.750000012048	20	0.774E-08	0.827E-08	0.783E-08	0.84E-08	0.106E-06	0.427E-02
1	9.750000002563	24	0.238E-08	0.238E-08	0.241E-08	0.242E-08	0.148E-07	0.340E-02
2	13.75000000772	24	0.704E-08	0.704E-08	0.772E-08	0.807E-08	0.282E-08	0.463E-02
3	15.75000007035	25	0.412E-08	0.412E-08	0.454E-08	0.469E-08	0.607E-07	0.353E-02

Table VIII to Table IV(a), the decrease in the number of degrees of freedom needed to achieve a given accuracy, when an adaptive procedure is used, is evident.

VII. CONCLUSIONS

When the "*R*-type" and "*H*-type" error estimators are in agreement, they are seen to track quite closely the error in the problems we have studied; it seems reasonable to expect that, when they agree, they may be taken with some degree of confidence to be good indicators of the actual energy error. With greater computational effort, errors can also be determined to high precision using the rigorous error bound. Any of these quantities also provides a good measure of the *local* error and may be employed effectively as a guide for adaptive refinement.

The above examples show clearly the ability of the adaptive approach to search out and find those regions of configuration space where more degrees of freedom are required for an accurate representation of the wave function. We anticipate that the increase in computational efficiency which is thus achieved will be more and more significant as the dimensionality of the problem increases, and will therefore be of the utmost importance for numerical quantum field theory.

ACKNOWLEDGMENTS

We thank Ronnie Manieri, Joel Runes, Alan Sokal, and Al Tino for helpful discussions during the course of this work. S. F. was partially supported by NSF Grant No. PHY-8715995. M. A. R. was partially supported by DOE Grant No. DE-AC02-87-ER40325-B_1.

[1]A particularly lucid exposition may be found in O. C. Zienkiewicz and K. Morgan, *Finite Elements and Approximation* (Wiley, New York, 1983).

[2]L. R. Ram-Mohan, S. Saigal, D. Dossa, and J. Shertzer, Comput. Phys. **4**, 50 (1989).

[3]C. D. Papageorgiou, A. D. Raptis, and T. E. Simos, J. Comput. Phys. **88**, 477 (1990); A. K. Commons, F. Ebrahimi, and S. T. Hafez, J. Phys. A **22**, 3229 (1989); J. Ftáčnik and I. Horváth, Phys. Rev. D **40**, 642 (1989).

[4]The FEM was first applied to Heisenberg-picture quantum theory in C. M. Bender and D. H. Sharp, Phys. Rev. Lett. **50**, 1535 (1983). Subsequent papers in this line of development include K. A. Milton and T. Grosse, Phys. Rev. D **41**, 1261 (1990); S. Hands, Phys. Lett. B **195**, 448 (1987); C. M. Bender, F. Cooper, K. A. Milton, S. S. Pinsky, and L. M. Simmons, Jr., Phys. Rev. D **35**, 3081 (1987); T. Matsuyama, Nucl. Phys. **B289**, 277 (1987); C. M. Bender and K. A. Milton, Phys. Rev. D **34**, 3149 (1986); L. Vázquez, *ibid.* **28**, 3253 (1983); V. Moncrief, *ibid.* **28**, 2485 (1983); D. Sen, Phys. Lett. **129B**, 239 (1983).

[4a]The authors of Ref. 4 apply the FEM to quantum field theory in the *Heisenberg* picture.

[5]G. Strang and G. J. Fix, *An Analysis of the Finite Element Method* (Prentice-Hall, Englewood Cliffs, NJ, 1973).

[6]I. Fried, *Numerical Solutions of Differential Equations* (Academic, New York, 1979).

[7]G. F. Carey and J. T. Oden, *Finite Elements: A Second Course* (Prentice-Hall, Englewood Cliffs, NJ, 1983).

[8]K.-J. Bathe, *Finite Element Procedures in Engineering Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1982).

[8a]The derivation of the bound presented here follows that of a bound on the *roundoff error* in numerically computed eigenvalues given in Ref. 8, p. 559.

[8b]When searching for the gound state, however, s_n is kept equal to zero for the first several iterations; when searching for higher states, it is initially kept equal to the energy of the previously computed state of lower energy. In our computation, shifting was initiated when a sufficiently large, fixed number of iterations had occurred, or a sufficiently small rate of change of the absolute value of the most slowly changing component of ψ_i had been reached. The implementation of shifting is crucial for achieving accurate results, especially in cases of near degeneracy of levels. There is no optimal algorithm to determine when to implement shifting; to verify that no level has been skipped, one can employ the *Sturm sequence check* in Ref. 8.

[9]P. Ladeveze and J. P. Pele, Int. J. Num. Meth. Eng. **28**, 1929 (1989), and references therein.

[10]D. W. Kelly, J. P. de S. R. Gago, O. C. Zienkiewicz, and I. Babuska, Int. J. Num. Meth. Eng. **19**, 1593 (1983).

[11]O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method,* 4th ed. (McGraw-Hill, London, 1989), Vol. 1.

[12]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory),* 3rd ed. (Pergamon, Oxford, 1977), p. 72.

[13]R. Blankenbecler, T. DeGrand, and R. L. Sugar, Phys. Rev. D **21**, 1055 (1980).