Modern Matrix Mechanics

Printed from: Quantum-Book/pde2mtx.tex on April 15, 2013

Robert Gilmore
Physics Department, Drexel University, Philadelphia, PA 19104

April 15, 2013

Abstract

Several convenient methods exist for transforming the Schrödinger eigenvalue equation into a matrix eigenvalue equation. These methods are based on introducing a set of points, called vertices, throughout the domain of interest, introducing a set of basis functions that are highly localized around each vertex, and then using these basis functions to compute matrix elements of Schrödinger's quadratic form. The eigenvectors of this matrix are computed and used to approximate the wavefunctions that are eigenfunctions of the Schrödinger hamiltonian. Common choices of basis functions include gaussians, pyramids, and Kronecker delta functions. These choices are at the heart of the three methods reviewed here: Gaussian methods, Finite Element Methods, and Finite Difference methods.

1 Background

Schrödinger's first formulation of the new quantum theory was as a variational principle. He proposed that the solutions $\psi(x)$ that made a certain action functional stationary held physical significance. The action functional proposed by Schrödinger is

$$I = \int \left[\Big((\mathbf{p} - \frac{q}{c}\mathbf{A})\psi(x) \Big)^* \cdot \Big((\mathbf{p} - \frac{q}{c}\mathbf{A})\psi(x) \Big) + V(x)\psi^*(x)\psi(x) \right] \, dV \quad (1)$$

where $\mathbf{p} = (\hbar/i)\boldsymbol{\nabla}$. He looked for functions $\psi(x)$ for which this action integral was stationary, subject to the condition that the wavefunctions were not identically zero: $\int \psi^*(x)\psi(x) \, dV = 1$.

When the normalization condition on the wavefunctions is imposed using Lagrange multipliers, and the action integral is integrated by parts, Schrödinger's well-known second order partial differential equation results.

Schrödinger also pointed out, at the bottom of the first page of his first paper on Wave Mechanics, that "... equation (1) can always be transformed so as to become a quadratic form (of ψ and its first derivatives) equated to zero." This statement makes a connection between the brand new Wave Mechanics and the Rayleigh-Ritz procedure, then well known from frequent use with partial differential equations. The general idea is to represent the (wave)function as a linear superposition of some set of known basis functions $f_j(x)$ with unknown coefficients c_j: $\psi(x) = \sum_j c_j f_j(x)$. In principle the sum should extend to "infinity" (countable or uncountable, depending on the potential); in practice we will always truncate the sum at some (large) finite value and hope (or test) for convergence. The approximations to the wavefunctions $\psi(x)$ depend *linearly* on the unknown expansion coefficients c_j. The action integral depends quadratically on the wavefunctions $\psi(x)$ (or linearly on ψ and also on $\psi^*(x)$). As a result the action integral depends quadratically (or bilinearly) on $c_j^* c_k$. Minimization of this quadratic form (in reality, the search for its stationary solutions), amounts to solving a matrix eigenvalue equation subject to constraints.

The coefficients of these terms can be constructed by opening up the parentheses in the action integral. We find the following result

$$I \quad \stackrel{\psi \to c_j f_j(x)}{\longrightarrow} \quad \sum_{j,k} c_j^* \left\{ \mathcal{K} + \mathcal{M} + \mathcal{V} - \lambda \mathcal{O} \right\}_{j,k} c_k \tag{2}$$

The constraint condition $\int \psi^*(x)\psi(x)\, dV = 1$ is enforced by the Lagrange multiplier λ, which will have an interpretation as an energy eigenvalue. The four matrices that appear within the brackets $\{\}$ are the kinetic energy matrix $\mathcal{K}$, the magnetic matrix $\mathcal{M}$, the potential energy matrix $\mathcal{V}$, and the overlap matrix $\mathcal{O}$. These matrices are defined by

$$
\begin{aligned}
\mathcal{K}_{j,k} &= \frac{\hbar^2}{2m}\int \boldsymbol{\nabla} f_j^*(x)\cdot\boldsymbol{\nabla} f_k(x)\; dV \\
\mathcal{M}_{j,k} &= \frac{q\hbar/i}{2mc}\int \left[-\boldsymbol{\nabla} f_j^*(x)\cdot\mathbf{A} f_k(x) + f_j^*(x)\mathbf{A}\cdot\boldsymbol{\nabla} f_k(x)\right]\; dV \\
\mathcal{V}_{j,k} &= \int f_j^*(x)\left[V(x) + \frac{q^2}{2mc^2}\mathbf{A}\cdot\mathbf{A}\right] f_k(x)\; dV \\
\mathcal{O}_{j,k} &= \int f_j^*(x) f_k(x)\; dV
\end{aligned}
\tag{3}
$$

If the magnetic term is absent (e.g., $q = 0$ or $\mathbf{A} = \mathbf{0}$) and only bound states are sought, all coefficients c_j can be assumed to be real. If the magnetic term is nonzero, it is convenient to lump the diamagnetic term ($\frac{q^2}{2mc^2}\mathbf{A}\cdot\mathbf{A}$) with the potential $V(x)$.

The quadratic form in Eq. (2) is made stationary by varying the coefficients c_j^*, c_k. This leads directly to the matrix eigenvalue equation:

$$
\delta I = 0 \Rightarrow \sum_k \left\{\mathcal{K} + \mathcal{M} + \mathcal{V} - \lambda\mathcal{O}\right\}_{j,k} = 0 \tag{4}
$$

The column vectors $c_k(\alpha)$ with energy eigenvalues E_α are orthonormal with respect to the overlap matrix:

$$
\sum_{j,k} c_j^*\mathcal{O}_{jk}c_k(\beta) = \delta(\alpha,\beta) \tag{5}
$$

It is clear from the expressions above that the devil is in the details of the set of functions $f_j(x)$ used to approximate the wavefunctions $\psi(x)$.

Remark: If a magnetic field is present, one must first compute the vector potential that represents the field. This is done through well-known integral relations (Helmholtz Theorem). If $\mathbf{B}(\mathbf{x}) = \boldsymbol{\nabla}\times\mathbf{A}$ is the magnetic field, then its vector potential $\mathbf{A(x)}$ is obtained through the volume and surface integrals

$$
\mathbf{A}(\mathbf{x}) = \frac{1}{4\pi}\boldsymbol{\nabla}\times\int_{\mathrm{V}}\frac{\mathbf{B}(\mathbf{y})}{||\mathbf{x}-\mathbf{y}||}\; dV + \frac{1}{4\pi}\int_{\partial\mathrm{V}}\frac{\mathbf{B}(\mathbf{y})}{||\mathbf{x}-\mathbf{y}||}\times\; d\mathbf{S} \tag{6}
$$

$\mathbf{A}$ is nonunique uip to the additin of the gradient of a scalar function: $\mathbf{A}\rightarrow\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi$.

2 Tessellation

The idea underlying the three procedures described below for mapping the wave equation to a matrix equation is to choose basis functions $f_j(x)$ that are highly localized around points (*vertices*) $\mathbf{a}$, so that the coefficients $c_j \to c_{\mathbf{a}}$ of these basis functions are good approximations to the wavefunction at a point. For each of the three methods described below, we will choose the basis functions $f_{\mathbf{a}}(x)$ to be highly localized around each vertex. The number of vertices is N_V.

We will develop these methods in two dimensions in such a way that their extension to three dimensions will not be difficult. The restriction to one dimension will be used to motivate the two-dimensional formulation.

The first step in the three computational methods we will describe is the placement of N_V vertices throughout the domain of interest. The same tessellation can be used for each of the three numerical computation methods described below.

The general idea is to use as many vertices as feasible. The more, the better the resolution but the slower the computation, so a balance has to be made. In general, the density of vertices is lower where the wavefunction is expected to vary more slowly, and higher where it is expected to vary more rapidly, with changes in position. For example, near corners or sharp changes in the value of the potential the wavefunction is expected to change quickly, so a higher density of vertices is called for. In the middle of a domain where the potential changes slowly it is usually possible to use a lower density of vertices.

If there are no walls confining the motion of the particle, the wavefunction for bound states is expected to approach zero as $r \to \infty$. If the computational domain is chosen large enough, the lower eigenfunctions (eigenfunctions with smaller energy eigenvalues) should have components that approach zero toward the edge of the domain. Such eigenfunctions may (provisionally) be deemed acceptable. Others that do not go to zero for large r do not approximate bound states well, and should be rejected.

If the potential contains hard walls, then the amplitudes of the wavefunctions should be set to zero at these boundaries.

When magnetic fields *and* hard walls are present the probability current normal to the walls must vanish.

In one dimension the tessellation is straightforward. It consists of a series of points, not necessarily equally spaced, along a one dimensional curve that

is often, but not necessarily, a line. In two dimensions, constructing a tessellation that is "convex" can be done without too much difficulty. Convexity means that the union of the simplices (triangles) containing any vertex is a convex region in R^2: two points in this region can be connected by a straight line that lies entirely within this region. In three dimensions constructions guaranteeing convexity are complicated.

The rows and columns of the $N_V \times N_V$ matrices in Eq. (3) are labeled by the vertices $\mathbf{a}, \mathbf{b}, \mathbf{c}, \cdots$ of the tessellations.

3 Construction of the Wavefunctions

The eigenvectors $c_i(\alpha)$ of the matrix equation (4) with eigenvalue $\lambda = E_\alpha$ are used to construct wavefunctions on the space in a straightforward way:

$$\psi_\alpha(\mathbf{x}) = \sum_k c_k(\alpha) f_k(\mathbf{x}) \tag{7}$$

The orthonormality property of the wavefunctions is related to the orthonormality property Eq.(5) of the matrix eigenfunctions

$$\int \psi_\alpha(\mathbf{x})\psi_\beta(\mathbf{x})dV = c_j^*(\alpha)\mathcal{O}_{jk}c_k(\beta) = \delta(\alpha,\beta) \tag{8}$$

3.1 Basis Function Choices

The differences between the three methods we review lie in the choice of basis functions $f_{\mathbf{a}}(\mathbf{x})$. In each case they are highly localized around the vertices $\mathbf{a}, \mathbf{b}, \mathbf{c}, \cdots$. Gaussian methods are based on gaussian functions; Finite Element Methods are based on pyramidal functions; Finite Difference methods are based on Kronecker delta functions. Gaussians are differentiable everywhere. Pyramidal functions are defined everywhere and differentiable except over a subset of points with zero measure. Kronecker delta functions are defined only on a set of points, the vertices.

4 Gaussian Methods

4.1 Motivation

The motivation behind this ensemble of computational methods ("Gaussian") is to use basis functions that can be dealt with easily and analytically. Gaussian basis functions fit this description to a "t".

4.2 Basis Functions

A useful set of functions consists of the gaussian functions:

$$f(\mathbf{x};\mathbf{a},A) = e^{-(\mathbf{x}-\mathbf{a})^t A(\mathbf{x}-\mathbf{a})/2} \tag{9}$$

Each basis function is identified by its vertex $\mathbf{a}$ and its covariance matrix A, where A is positive-definite. Each basis function is normalized so that its value at the vertex is $+1$. Further, these functions should be sufficiently localized so that all integrals can be extended to infinity without substantial error (A has large eigenvalues). Though highly localized, they usually extend over several (many) neighboring vertices.

4.3 One-dimensional Integrals

Before proceeding, we consider the elementary one-dimensional integrals

$$\begin{array}{rcccl} I^{(0)}(A) & = & \int_{-\infty}^{+\infty} e^{-A(x-a)^2/2}dx & = & \sqrt{\frac{2\pi}{A}} \\ I^{(2)}(A) & = & \int_{-\infty}^{+\infty} (x-a)^2 e^{-A(x-a)^2/2}dx & = & A^{-1}I^0(A) \end{array} \tag{10}$$

The product of two gaussians is a gaussian and the integral is easily computed:

$$\begin{array}{rcl} e^{-A(x-a)^2/2} \times e^{-B(x-b)^2/2} & = & e^{-C(x-c)^2/2} e^{-\frac{1}{2}\text{stuff}} \\ A + B & = & C \\ Aa + Bb & = & Cc \\ A(c-a)^2 + B(c-b)^2 & = & \text{stuff} \\ \int e^{-A(x-a)^2/2} \times e^{-B(x-b)^2/2}\, dV & = & I^{(0)}(C) e^{-\frac{1}{2}\text{stuff}} \end{array} \tag{11}$$

The product of the derivatives of two gaussians is also a gaussian multiplied by quadratic terms, and its integral is also easily computed:

$$\frac{d}{dx}e^{-A(x-a)^2/2} \times \frac{d}{dx}e^{-B(x-b)^2/2} =$$
$$AB\left\{(x-c)^2 + (\cdots)(x-c) + (c-a)(c-b)\right\} e^{-C(x-c)^2/2} e^{-\frac{1}{2}\text{stuff}}$$
$$\int \frac{d}{dx}e^{-A(x-a)^2/2} \times \frac{d}{dx}e^{-B(x-b)^2/2}\, dV = AB\left\{C^{-1} + (c-a)(c-b)\right\} I^{(0)}(C) e^{-\text{stuff}/2} \tag{12}$$

4.4 Multi-dimensional Integrals

Before proceeding it is useful to identify two integrals, generalizations of those in Eq. (10), that will appear often below. These are integrals over all space R^n, so that the coordinates $\mathbf{x}$ and vertex locations $\mathbf{a}$ are dummy variables.

$$I^{(0)}(A) = \int .. \int f(\mathbf{x};\mathbf{a},A) d^n x = \frac{(2\pi)^{n/2}}{\sqrt{\det(A)}} \tag{13}$$

$$I^{(2)}_{ij}(A) = \int .. \int (\mathbf{x}-\mathbf{a})_i (\mathbf{x}-\mathbf{a})_j f(\mathbf{x};\mathbf{a},A) d^n x = (A^{-1})_{ij} I^{(0)}(A) \tag{14}$$

4.5 Matrix Elements

Gaussians have the useful property that products of gaussians remain gaussians. We will repeatedly encounter integrals of functions multiplied by products of two gaussians. The product itself has the form

$$f(\mathbf{x};\mathbf{a},A) f(\mathbf{x};\mathbf{b},B) = f(\mathbf{x};\mathbf{c},C) e^{-\text{stuff}/2} \tag{15}$$

The parameters in this product can be computed by "completing the square". The (negative of the) sum of the terms in the exponents of the two gaussian functions is

$$(\mathbf{x}-\mathbf{a})^t A (\mathbf{x}-\mathbf{a}) + (\mathbf{x}-\mathbf{b})^t B (\mathbf{x}-\mathbf{b}) = (\mathbf{x}-\mathbf{c})^t C (\mathbf{x}-\mathbf{c}) + \text{stuff} \tag{16}$$

By comparing terms in Eq. (16) we find:

$$\begin{aligned} C &= A + B \\ C\mathbf{c} &= A\mathbf{a} + B\mathbf{b} \\ \text{stuff} &= (\mathbf{c}-\mathbf{a})^t A(\mathbf{c}-\mathbf{a}) + (\mathbf{c}-\mathbf{b})^t B(\mathbf{c}-\mathbf{b}) \\ &= \mathbf{a}^t A\mathbf{a} + \mathbf{b}^t B\mathbf{b} - \mathbf{c}^t C\mathbf{c} = (C\mathbf{c})^t (C^{-1})(C\mathbf{c}) \end{aligned} \tag{17}$$

When "stuff" > 20, the integral is zero for all practical purposes, and need not even be computed. In making this simplification we will create matrices that are "sparse". Fast matrix diagonalization techniques have been developed to handle sparse matrices.

The three integrals that we find useful are:

$$\begin{aligned} \int f(\mathbf{x};\mathbf{a},A) f(\mathbf{x};\mathbf{b},B) d^n x &= \mathcal{O}_{\mathbf{a},\mathbf{b}} = I^{(0)}(C) e^{-\text{stuff}} \\ \int (\mathbf{x}-\mathbf{r}_1)_i f(\mathbf{x};\mathbf{a},A) f(\mathbf{x};\mathbf{b},B) d^n x &= (\mathbf{c}-\mathbf{r}_1)_i \mathcal{O}_{\mathbf{a},\mathbf{b}} \\ \int (\mathbf{x}-\mathbf{r}_1)_i (\mathbf{x}-\mathbf{r}_2)_j f(\mathbf{x};\mathbf{a},A) f(\mathbf{x};\mathbf{b},B) d^n x &= \left((\mathbf{c}-\mathbf{r}_1)_i (\mathbf{c}-\mathbf{r}_2)_j + (C^{-1})_{ij}\right) \mathcal{O}_{\mathbf{a},\mathbf{b}} \end{aligned} \tag{18}$$

4.5.1 Overlap Matrix

The overlap quadratic form is

$$\int \sum_{\mathbf{a}} c^*_{\mathbf{a}} f(\mathbf{x};\mathbf{a},A) \sum_{\mathbf{b}} c_{\mathbf{b}} f(\mathbf{x};\mathbf{b},B) d^n x = \sum_{\mathbf{a},\mathbf{b}} c^*_{\mathbf{a}} \mathcal{O}_{\mathbf{a},\mathbf{b}} c_{\mathbf{b}} \tag{19}$$

The overlap matrix element $\mathcal{O}_{\mathbf{a},\mathbf{b}}$ is given in the first of Eqs. (18).

4.5.2 Potential Energy Matrix

The quadratic form describing the potential energy is obtained following the machinery developed to compute the overlap matrix in Eq. (19)

$$\int \sum_{\mathbf{a}} c^*_{\mathbf{a}} f(\mathbf{x};\mathbf{a},A)\ V(\mathbf{x})\ \sum_{\mathbf{b}} c_{\mathbf{b}} f(\mathbf{x};\mathbf{b},B) d^n x = \sum_{\mathbf{a},\mathbf{b}} c^*_{\mathbf{a}} \mathcal{V}_{\mathbf{a},\mathbf{b}} c_{\mathbf{b}} \tag{20}$$

The potential energy is expanded about $\mathbf{c}$ in a Taylor series to second order:

$$V(\mathbf{x}) \to V(\mathbf{c}) + V_i(\mathbf{c})(\mathbf{x}-\mathbf{c})_i + \frac{1}{2} V_{ij}(\mathbf{c})(\mathbf{x}-\mathbf{c})_i (\mathbf{x}-\mathbf{c})_j + \text{h.o.t.} \tag{21}$$

The integral over the constant term is proportional to the overlap matrix element: $V(\mathbf{c})\mathcal{O}_{\mathbf{a},\mathbf{b}}$. The integral over the linear term is zero "by symmetry". This leaves only the integral over the quadratic term, which is $\frac{1}{2}\mathrm{tr}(V_{*,*}C^{-1})\times \mathcal{O}_{\mathbf{a},\mathbf{b}}$. The net result is

$$\mathcal{V}_{\mathbf{a},\mathbf{b}} = \left(V(\mathbf{c}) + \frac{1}{2}\mathrm{tr}(V_{*,*}(\mathbf{c})C^{-1})\right) \times \mathcal{O}_{\mathbf{a},\mathbf{b}} \tag{22}$$

4.5.3 Kinetic Energy Matrix

The quadratic form describing the kinetic energy is obtained following the prescriptions above.

$$\frac{\hbar^2}{2m} \times \int \sum_{\mathbf{a}} c^*_{\mathbf{a}} \frac{\partial}{\partial x_i} f(\mathbf{x};\mathbf{a},A) \sum_{\mathbf{b}} c_{\mathbf{b}} \frac{\partial}{\partial x_i} f(\mathbf{x};\mathbf{b},B) d^n x = \sum_{\mathbf{a},\mathbf{b}} c^*_{\mathbf{a}} \mathcal{K}_{\mathbf{a},\mathbf{b}} c_{\mathbf{b}} \tag{23}$$

We find

$$\mathcal{K}_{\mathbf{a},\mathbf{b}} = \frac{\hbar^2}{2m} \times \int A_{ir}(\mathbf{x}-\mathbf{a})_r \times B_{is}(\mathbf{x}-\mathbf{b})_s \times f(\mathbf{x};\mathbf{a},A) f(\mathbf{x};\mathbf{b},B) d^n x \tag{24}$$

The linear terms are expanded about $\mathbf{c}$: $\mathbf{x}-\mathbf{b} = (\mathbf{x}-\mathbf{c}) + (\mathbf{c}-\mathbf{b})$ and the integrals are carried out using the third of Eqs. (18):

$$\mathcal{K}_{\mathbf{a},\mathbf{b}} = \frac{\hbar^2}{2m} \times \left((\mathbf{c}-\mathbf{a})_r (AB^t)_{rs} (\mathbf{c}-\mathbf{b})_s + \mathrm{tr}(AC^{-1}B^t)\right) \times \mathcal{O}_{\mathbf{a},\mathbf{b}} \tag{25}$$

4.5.4 Magnetic Matrix

The matrix elements of the magnetic matrix are constructed using the integral expressions given in Eq. (18) after the vector potential $\mathbf{A}(\mathbf{x})$ is expanded about $\mathbf{c}$ and the constant and linear terms are retained:

$$\mathcal{M}_{\mathbf{a},\mathbf{b}} = \frac{q\hbar/i}{2mc} \left\{ \mathbf{A}_r(\mathbf{c}) \left(B_{rs}(\mathbf{c}-\mathbf{b})_s - A_{rs}(\mathbf{c}-\mathbf{a})_s\right) + \frac{\partial \mathbf{A}_r}{\partial \mathbf{x}_s}\Big|_{\mathbf{c}} (A-B)_{rt} C^{-1}_{ts} \right\} \times \mathcal{O}_{\mathbf{a},\mathbf{b}} \tag{26}$$

4.6 Construction of the Global Matrices

The global kinetic, potential, magnetic, and overlap matrices $\mathcal{K}, \mathcal{V}, \mathcal{M}, \mathcal{O}$ are built up matrix element by matrix element. This involves a double sweep over the vertics $\mathbf{a}, \mathbf{b}$. Each of the other three $\mathbf{a}, \mathbf{b}$ matrix elements $\mathcal{K}_{\mathbf{a},\mathbf{b}}, \mathcal{V}_{\mathbf{a},\mathbf{b}}, \mathcal{M}_{\mathbf{a},\mathbf{b}}$ are proportional to the overlap matrix element $\mathcal{O}_{\mathbf{a},\mathbf{b}}$. If this is sufficiently small the other three matrix elements can also be zeroed out. Even so, the overlap matrix element $\mathcal{O}_{\mathbf{a},\mathbf{b}}$ must first be computed. Computation of these small matrix elements reduces the efficiency of this procedure.

4.7 Eigenfunctions

The wavefunctions $\psi_\alpha(\mathbf{x})$ associated with energy eigenvalue $\lambda = E_\alpha$ are constructed from the $N_V \times 1$ column vectors $c_{\mathbf{b}}(\alpha)$ and the basis functions $f(\mathbf{x}; \mathbf{b}, B)$:

$$\psi_\alpha(\mathbf{x}) = \sum_{\mathbf{b}} c_{\mathbf{b}}(\alpha) f(\mathbf{x}; \mathbf{b}, B) = \sum_{\mathbf{b}} c_{\mathbf{b}}(\alpha) e^{-(\mathbf{x}-\mathbf{b})^t B(\mathbf{x}-\mathbf{b})/2} \tag{27}$$

At any lattice site several basis function contribute to the value of the wavefunction:

$$\psi_\alpha(\mathbf{x} = \mathbf{a}) = \sum_{\mathbf{b}} c_{\mathbf{b}}(\alpha) f(\mathbf{a}; \mathbf{b}, B) = \sum_{\mathbf{b}} c_{\mathbf{b}}(\alpha) e^{-(\mathbf{a}-\mathbf{b})^t B(\mathbf{a}-\mathbf{b})/2} \tag{28}$$

The values of $e^{-(\mathbf{a}-\mathbf{b})^t B(\mathbf{a}-\mathbf{b})/2}$ can be considered as the elements of an $N_V \times N_V$ matrix $M_{\mathbf{a},\mathbf{b}}(B)$ that is not symmetric, so that the value of the eigenfunctions can finally be expressed as a matrix product:

$$\psi_\alpha(\mathbf{a}) = \sum_{\mathbf{b}} M_{\mathbf{a},\mathbf{b}}(B) c_{\mathbf{b}}(\alpha) \tag{29}$$

A minor problem with Gaussian methods is that the matrix eigenfunctions cannot immediately be interpreted as values of the eigenfunctions on the lattice sites.

5 Finite Element Methods

5.1 Motivation

One of the attractive features of finite element methods is the use of basis functions for which it is known beforehand which integrals need be evaluated and which are zero. This is done by a choice of basis functions that are highly localized around each vertex. In fact, these basis functions are nonzero only in elements that contain the vertex.

5.2 Basis Functions

The basis function "at" vertex $\mathbf{a}$ is $f(\mathbf{x};\mathbf{a})$. This is defined to be $+1$ at $\mathbf{x}=\mathbf{a}$ and drop off linearly as the boundary of the convex region around $\mathbf{a}$ is approached. If $\mathbf{a}$ is contained in k triangles (simplices) this function is nonzero in these k triangles. When the value of $f(\mathbf{x};\mathbf{a})$ is plotted over the domain it has the shape of a k-sided pyramid: hence the name "pyramidal function". This basis function is continuous everywhere and differentiable almost everywhere. It is not differentiable at $\mathbf{a}$, on the edges emanating from $\mathbf{a}$, and on the edges opposite $\mathbf{a}$. These edges delineate the boundary of the convex region surrounding $\mathbf{a}$. The non-differentiability over a measure zero set is usually not a problem. When it becomes a problem there are standard medicines to solve this problem.

5.3 One-dimensional Integrals

In one dimension consider adjacent points $a<b<c<d$ in a tessellation. The basis function $f_b(x)$ rises linearly from 0 at $x=a$ to $+1$ at $x=b$ and then decreases linearly to 0 at $x=c$. The basis function $f_c(x)$ is similar, being nonzero only in the range $b<x<d$. All basis functions are everywhere continuous and nondifferentiable at three vertices. Explicitly, these pyramids are

$$\begin{array}{lllll} f_b(x)= & \dfrac{x-a}{b-a} & a\le x\le b & \dfrac{c-x}{c-b} & b\le x\le c \\ f_c(x)= & \dfrac{x-b}{c-b} & b\le x\le b & \dfrac{d-x}{d-c} & c\le x\le d \end{array} \tag{30}$$

The functions are zero outside their indicated ranges.

The integrals of products of these functions and their derivatives are simple to carry out on the interval $b \leq x \leq c$. The results depend on the length $L = c - b$ of the interval:

	$f_b(x)$	$f_c(x)$	$df_b(x)/dx$	$df_c(x)/dx$
$f_b(x)$	$L/3$	$L/6$	$-1/2$	$1/2$
$f_c(x)$	$L/6$	$L/3$	$-1/2$	$1/2$
$df_b(x)/dx$	$-1/2$	$-1/2$	$1/L$	$-1/L$
df_c/dx	$1/2$	$1/2$	$-1/L$	$1/L$

(31)

Observe in passing: the integral over two basis functions is proportional to the measure of the domain; the integral over derivatives of the two basis functions is inversely proportional to the measure of the domain; and the integral of a basis function with a derivative of a basis function is independent of the measure of the domain.

5.4 Multi-dimensional Integrals

The action integral I of Eq. (1) can be partitioned into integrals over each of the triangular elements resulting from the tesselation. Over a given element with three vertices $\mathbf{a}, \mathbf{b}, \mathbf{c}$ only three basis functions are nonzero. These are $f(\mathbf{x}; \mathbf{a})$, $f(\mathbf{x}; \mathbf{b})$, and $f(\mathbf{x}; \mathbf{c})$. For convenience we relabel these three vertices $\mathbf{a} \to (x_1, y_2)$, $\mathbf{b} \to (x_2, y_2)$, and $\mathbf{c} \to (x_3, y_3)$. In the element with these three vertices we create the linear function $f_1(x, y)$ from two matrices as follows:

$$M_0 = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \qquad M_1 = \begin{bmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \qquad f_1(x, y) = \frac{\det(M_1)}{\det(M_0)} \tag{32}$$

The function $f_2(x, y)$ is obtained in the same way, using a matrix M_2 obtained from M_0 by replacing the second row $(x_2, y_2) \to (x, y)$. Similarly for $f_3(x, y)$. The function $f_1(x, y)$ is linear in the coordinates x and y, assumes value $+1$ at (x_1, y_1) and vanishes at the other two vertices in this triangle as well as the edge opposite (x_1, y_1) that connects (x_2, y_2) with (x_3, y_3). Similarly for $f_2(x, y)$ and $f_3(x, y)$.

In this element we construct a simple linear combination of these three functions to represent $\psi(x, y)$, $(x, y) \in \Delta(\mathbf{a}, \mathbf{b}, \mathbf{c})$

$$\begin{aligned} \psi(x,y) &= c_1 f_1(x,y) + c_2 f_2(x,y) + c_3 f_3(x,y) \\ \psi^*(x,y) &= c_1^* f_1(x,y) + c_2^* f_2(x,y) + c_3^* f_3(x,y) \end{aligned} \tag{33}$$

Since these functions are well-defined and linear, the four 3×3 matrices describing the quadratic forms for the kinetic, potential, magnetic, and overlap matrices involving the three amplitudes c_1, c_2, c_3 and their complex conjugates can be constructed.

5.5 Simplification of the Calculations

The simplest way to carry out these calculations is to introduce a new set of coordinates (u, v) and a linear transformation $(u, v) \rightarrow (x, y)$ as follows

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} (1-u-v) + \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{34}$$

The three corners $(u, v) = (0, 0), (1, 0), \text{and } (0, 1)$ of the right triangle in (u, v) space map to the vertices $(x_1, y_1), (x_2, y_2), \text{and } (x_3, y_3)$ of the element in (x, y) space. Using this linear transformation converts the integration over the element to integrations over a right triangle. Note that

$$\det \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{bmatrix} = \det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} = 2! \text{Area } \Delta(\mathbf{a}, \mathbf{b}, \mathbf{c}) \tag{35}$$

The measure of the triangle with vertices $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is proportional to the jacobian of the matrix M_0 in Eq. (32). The proportionality factor is $1/2!$. In n dimensions there are similar jacobians of $(n+1) \times (n+1)$ matrices and proportionality factors are $1/n!$.

5.6 Matrix Elements

Matrix elements: involving products of elements are proportional to M_0; involving products of derivatives are inversely proportional to M_0; and those involving one basis function and the derivative of a basis function are independent of M_0.

5.6.1 Overlap Matrix

The results for the four matrices of interest are:

$$\mathcal{O} = \frac{\det(M_0)}{24} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{36}$$

5.6.2 Potential Energy Matrix

The potential in this region is expressed as a linear combination of the three basis functions: $V(x,y) = V_1 f_1(x,y) + V_2 f_2(x,y) + V_3 f_3(x,y)$, where $V_1 = V(x_1, y_1)$. The integrals are easily carried out:

$$\mathcal{V} = \frac{\det(M_0)}{120} \begin{bmatrix} (6,2,2) & (2,2,1) & (2,1,2) \\ (2,2,1) & (2,6,2) & (1,2,2) \\ (2,1,2) & (1,2,2) & (2,2,6) \end{bmatrix} \tag{37}$$

The meaning of $(2,1,2)$ in the $(1,3)$ matrix element is

$$\begin{pmatrix} 2 & 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = 2V_1 + 1V_2 + 2V_3$$

5.6.3 Kinetic Energy Matrix

The matrix elements of the kinetic energy matrix (apart from the quantum factor $\frac{\hbar^2}{2m}$) are

$$\begin{array}{rcl} \mathcal{K}_{ii} & = & \frac{1}{2}\left((x_j - x_k)^2 + (y_j - y_k)^2\right)/\det(M_0) \\ \mathcal{K}_{jk} & = & -\frac{1}{2}\left((x_j - x_i)(x_k - x_i) + (y_j - y_i)(y_k - y_i)\right)/\det(M_0) \end{array} \tag{38}$$

In these expressions $1 \le i \ne j \ne k \le 3$.

5.6.4 Magnetic Matrix

The magnetic matrix is hermitian. It is the antisymmetric matrix given below, multiplied by the imaginary factor $\frac{1}{2m}\frac{q}{c}\frac{\hbar}{i}$:

$$\begin{array}{r} \mathcal{M}_{ij} = -\mathcal{M}_{ji} = \frac{1}{8}\left\{4\overline{A}_y\left((x_i - x_k) + (x_j - x_k)\right) + \mathbf{\Delta x}\cdot\mathbf{\Delta A_y}\right\} \\ -\frac{1}{8}\left\{4\overline{A}_x\left((y_i - y_k) + (y_j - y_k)\right) + \mathbf{\Delta y}\cdot\mathbf{\Delta A_x}\right\} \end{array} \tag{39}$$

In this expression (ijk) is a cyclic permutation of (123), $\overline{A}_x = A_{1x}+A_{2x}+A_{3x}$ and

$$\mathbf{\Delta x}\cdot\mathbf{\Delta A_y} = (x_1-x_2)(A_{1y}-A_{2y})+(x_2-x_3)(A_{2y}-A_{3y})+(x_3-x_1)(A_{3y}-A_{1y})$$

The other two expressions are similar.

5.7 Assembling the Matrix

The matrix form of the Schrödinger equation is an $N_V \times N_V$ matrix. Each of the matrices above is a 3×3 matrix. The nine matrix elements in each of these four matrices must be placed in the appropriate positions in the four large $N_V \times N_V$ matrices $\mathcal{K}, \mathcal{V}, \mathcal{M}, \mathcal{O}$. This is done element by element. For each triangular element the 3×3 matrices are constructed and the matrix elements *assembled* into the respective matrices. As the triangular elements are swept over, the large matrices are built up. After all the triangular elements are swept over and matrix elements put into their appropriate places the resulting generalized eigenvalue equation is solved.

5.8 Eigenfunctions

The $N_V \times 1$ matrix eigenvectors $c_{\mathbf{a}}(\alpha)$ with eigenvalue E_α are used to build up the wavefunctions as usual according to

$$\psi_\alpha(x,y) = \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha) f_{\mathbf{a}}(x,y) \tag{40}$$

Since each basis function is zero at all vertices except one, each component of the eigenvector is the value of the wavefunction in state α at vertex $\mathbf{a}$:

$$\psi_\alpha(\mathbf{b}) = \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha) f_{\mathbf{a}}(\mathbf{b}) \to \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha)\delta(\mathbf{a},\mathbf{b}) = c_{\mathbf{b}}(\alpha) \tag{41}$$

This means that the compoents $c_{\mathbf{a}}(\alpha)$ of the α eigenvector can be pumped directly to a visual display without further processing to exhibit the value of the wavefunction at each vertex $\mathbf{b}$.

6 Finite Difference Method

The finite difference method is the oldest of the methods used to convert partial differential equations to matrix equations. In this method derivatives are approximated by differences in the values of functions at neighboring points. The approximation gets better as the points get closer. There is a prejudice against using this method unless the domain under investigation is rectilinear. However, it can be used for an arbitrary-shaped domain.

6.1 Motivation

The method is straightforward and only the derivatives need be approximated by matrices. All other functions - the overlap integrals and the kinetic energy matrix, are represented by diagonal matrices.

6.2 Basis Functions

The basis functions $f(\mathbf{x};\mathbf{a})$ are defined *only* at the vertices. They are Kronecker delta functions:

$$f(\mathbf{x};\mathbf{a}) \to f(\mathbf{b};\mathbf{a}) = \delta(\mathbf{b},\mathbf{a}) \tag{42}$$

As such, the eigenvectors $c_{\mathbf{a}}(\alpha)$ are the values of the eigenfunctions $\psi_\alpha(\mathbf{a})$ at the N_V vertices $\mathbf{a}$.

6.3 Gradient Matrix

6.3.1 One-Dimensional Case

If b_i are neighbors of a so that $b_i - a$ is small and $(b_i - a)^2 \ll b_i - a$ (more correctly, $(b_i - a)^2 |d^2\psi/dx^2| \ll |(b_i - a)| d\psi/dx|$) then we can approximate a function $f(x)$ at points b_i by

$$f(b_i) \simeq f(a) + (b_i - a)\nabla_a f, \qquad \text{or} \qquad (f(b_i) - f(a)) \simeq (b_i - a)\nabla_a f \tag{43}$$

Including several nearby points yields

$$\sum_i (b_i - a)(f(b_i) - f(a)) \simeq \sum_i (b_i - a)^2 \nabla_a f \qquad \text{or} \qquad \nabla_a = \sum_i w_i(a, b_i) f(b_i) \tag{44}$$

The weights $w(a, b_i)$ are

$$w(a, b_i) = \frac{(b_i - a)}{\sum_j (b_j - a)^2} \tag{45}$$

The matrix elements for the differential operator $(D_x)_{a,b_i}$ for row a vanish except for the columns associated with the neighboring points b_i, for which they are $w(a, b_i)$. The diagonal matrix element in this row is chosen so that the sum of all the elements in row a is zero.

As an example, if a is surrounded by four vertices at distances $\pm\Delta$ and $\pm\frac{3}{2}\Delta$ from a, the nonzero matrix elements in this row are $(\cdots - \frac{3}{2}, -1, 0, +1, +\frac{3}{2}, \cdots)/\frac{13}{2}\Delta$.

6.3.2 Multi-dimensional Case

The arguments below are carried out on a two-dimensional grid.

Assume a vertex at $\mathbf{x}_0 = (x_0, y_0)$ is surrounded by vertices we label $1, 2, \ldots, n$ with coordinates $\mathbf{x_0} + \mathbf{\Delta x}_i = (x_0 + \Delta x_i, y_0 + \Delta y_i)$ and "surrounded" means that each vertex i is connected to the central vertex by one edge. The gradient is accessible through a truncated Taylor series expansion that becomes a better and better approximation as the tessellation becomes finer and finer:

$$\Delta f_i = f(\mathbf{x}_0 + \Delta \mathbf{x}_i) - f(\mathbf{x}_0) \simeq \boldsymbol{\nabla} f \cdot \mathbf{\Delta x}_i \tag{46}$$

Then in linear approximation we can write

$$\begin{bmatrix} \Delta x_1 & \Delta y_1 \\ \Delta x_2 & \Delta y_2 \\ \vdots & \vdots \\ \Delta x_n & \Delta y_n \end{bmatrix} \begin{bmatrix} (\boldsymbol{\nabla} f)_x \\ (\boldsymbol{\nabla} f)_y \end{bmatrix} \simeq \begin{bmatrix} \Delta f_1 \\ \Delta f_2 \\ \vdots \\ \Delta f_n \end{bmatrix} \tag{47}$$

For simplicity, call the $n \times 2$ matrix on the left T. Now multiply both sides by $T^\dagger = T^t$. On the left we find a 2×2 invertible matrix $T^t T$, so

$$\begin{bmatrix} (\boldsymbol{\nabla} f)_x \\ (\boldsymbol{\nabla} f)_y \end{bmatrix} \simeq (T^t T)^{-1} \begin{bmatrix} \Delta x_1 & \Delta x_2 & \cdots & \Delta x_n \\ \Delta y_1 & \Delta y_2 & \cdots & \Delta y_n \end{bmatrix} \begin{bmatrix} \Delta f_1 \\ \Delta f_2 \\ \vdots \\ \Delta f_n \end{bmatrix} \tag{48}$$

The matrix representative of ∇_x in the row for $\mathbf{x}_0$ has nonzero matrix elements wx_i in columns describing the contiguous vertices i and a compensating diagonal matrix element $-\sum_i wx_i$ whose origin is the difference term in $f(x_0 + \Delta x_i) - f(x_0)$. The weights wx_i, wy_i are

$$\begin{bmatrix} wx_i \\ wy_i \end{bmatrix} = (T^t T)^{-1} \begin{bmatrix} \Delta x_i \\ \Delta y_i \end{bmatrix} \tag{49}$$

The x-component $\mathbf{D}_x$ (a matrix) of the gradient operator is constructed by placing the x-weights for the vertex x_0 in the row labelled by x_0 and the appropriate columns. The y-component $\mathbf{D}_y$ of the gradient operator is constructed similarly. Note that the x weights wx_i generally depend on the x *and* y coordinates of the surrounding set $(x, y)_i$.

6.4 Matrix Elements

In this representation the position operator $\mathbf{Q}_x$ for the x coordinate is diagonal. The diagonal matrix elements for this operator are the x components of the vertices. Similarly for the matrix $\mathbf{Q}_y$ representing the y coordinate.

6.4.1 Overlap Matrix

The overlap matrix is diagonal:

$$\mathcal{O}_{\mathbf{a},\mathbf{b}} = \sum_{\mathbf{c}} \delta(\mathbf{a}, \mathbf{c})\delta(\mathbf{b}, \mathbf{c}) = \delta(\mathbf{a}, \mathbf{b}) \tag{50}$$

6.4.2 Potential Energy Matrix

The matrix describing the potential is also diagonal: the matrix element at position $\mathbf{a}, \mathbf{a}$ is $V(\mathbf{a})$.

6.4.3 Kinetic Energy Matrix

Write $\mathbf{D}_x$ for the matrix that represents the x component of the gradient, constructed as described above. Then $\mathbf{D}_x^t$ is its transpose. Matrices $\mathbf{D}_y$ and $\mathbf{D}_y^t$ have similar meanings.

$$\mathcal{K} = \frac{\hbar^2}{2m} \left\{ \mathbf{D}_x^t \mathbf{D}_x + \mathbf{D}_y^t \mathbf{D}_y \right\} \tag{51}$$

6.4.4 Magnetic Matrix

The components of the vector potential $\mathbf{A}(\mathbf{x})$ are also represented by diagonal matrices. The magnetic matrix is

$$\mathcal{M} = -\frac{q\hbar/i}{2mc} \left\{ \left(\mathbf{A}_x \mathbf{D}_x - (\mathbf{A}_x \mathbf{D}_x)^t \right) + \left(\mathbf{A}_y \mathbf{D}_y - (\mathbf{A}_y \mathbf{D}_y)^t \right) \right\} \tag{52}$$

The real antisymmetric matrix within the curly brackets is multiplied by an imaginary number, so $\mathcal{M}$ is hermitian.

6.5 Eigenvectors

The eigenvectors of the matrix equation are the approximate eigenfunctions of the partial differential equation:

$$\psi_\alpha(\mathbf{x}) = \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha) f(\mathbf{x}; \mathbf{a}) \rightarrow \psi_\alpha(\mathbf{b}) = \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha) f(\mathbf{b}; \mathbf{a}) = \sum_{\mathbf{a}} c_{\mathbf{a}}(\alpha) \delta(\mathbf{b}, \mathbf{a}) = c_{\mathbf{b}}(\alpha) \tag{53}$$

7 Comparisons

The column vectors $c_{\mathbf{a}}(\alpha)$ are the values of the eigenfunctions $\psi_\alpha(\mathbf{x})$ at $\mathbf{x} = \mathbf{a}$ for finite element calculations (c.f., Eq. (41)) and finite difference calculations (c.f., Eq. (53)). However, this is not the case with gaussian methods (c.f., Eq. (29)).

The strategy for building up the matrices $\mathcal{K}, \mathcal{V}, \mathcal{M}, \mathcal{O}$ differs among the three methods.

For Gaussian methods a vertex $\mathbf{a}$ is fixed. This index identifies a row in each of the matrices. Then other vertices $\mathbf{b}$ (column indices) are scanned

over and those vertices that exhibit nonvanishingly small matrix elements are retained in the construction of matrix elements.

For Finite Element Methods the strategy is to partition the domain into simplices — triangles in R^2 — and to carry out integrations over each simplex. Each simplex is defined by three vertices so the integration over one simplex generates a quadratic form in the three coefficients $c_{\mathbf{a}}, c_{\mathbf{b}}, c_{\mathbf{c}}$ and their conjugates. The nine matrix elements of the quadratic forms for the four operators $\mathcal{K}, \mathcal{V}, \mathcal{M}, \mathcal{O}$ are placed in the appropriate place ("matrix assembly") in these four matrices, and the integration over remaining elements proceeds.

The Finite Difference strategy is intermediate. Simplices containing each vertex $\mathbf{a}$ are defined (as in Finite Element Methods) and the vertices $\mathbf{b}_i$ defining these (convex) domains are used to determine the matrix elements in the row defined by $\mathbf{a}$ in the gradient operator, as described in Sec. 6.3.2.

One important criterion for the numerical stability of the computations is the scaling properties of the four matrices.

For Gaussian methods the overlap and potential matrices are scale-independent, the kinetic energy matrix scales like the inverse square of the linear density, and the magnetic matrix scales like the inverse of the linear density.

For finite element methods the overlap and potential matrices scale like the mean *volume* of the tessellation, the kinetic energy matrix scales like the inverse of the mean volume, and the magnetic matrix is scale-independent.

For finite difference methods the overlap matrix is the identity and the potential energy matrix is scale independent. The gradient matrices, and therefore the magnetic matrix, scale inversely to the mean volume of the tesselation, and the kinetic energy matrix scales like the inverse square of the mean volume.

These back-of-the-envelop thoughts suggest that Gaussian methods may be more stable against finer meshing than the other two methods.

8 Useful Connections with Wave Mechanics

Once the wavefunctions are available at a discrete set of points it is possible to make many discrete connections with Wave Mechanics. We describe two below. The analysis is predicated on knowing the wavefunction at a discrete set of points. We use a discrete approximation for the components of the gradient operator. We also construct discrete analogs for the probability current operator and the canonical commutation relations.

8.1 Probability Current Operator

If the particle is charged, the electric current density is the probability current density multiplied by the particle charge.

The electric current operator $\mathbf{j} = e\mathbf{v}$ is given by

$$\mathbf{j} = e\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)/m \tag{54}$$

In any state ψ the probability current is

$$\langle \mathbf{v} \rangle = \frac{1}{m}\ \mathrm{Re}\ \psi^* \left(\frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q}{c}\mathbf{A}\right)\psi = \frac{1}{m}\ \mathrm{Re}\ \psi^* \frac{\hbar}{i}\boldsymbol{\nabla}\psi - \frac{q}{mc}\psi^*\mathbf{A}\psi \tag{55}$$

At the vertex $\mathbf{x}_0$

$$\begin{bmatrix} \mathbf{v}_x \\ \mathbf{v}_y \end{bmatrix}_{\mathbf{x}_0} = \frac{\hbar}{m}\ \mathrm{Im}\ \psi_0^* \begin{bmatrix} wx_1 & wx_2 & \cdots & wx_n \\ wy_1 & wy_2 & \cdots & wy_n \end{bmatrix} \begin{bmatrix} \Delta\psi_1 \\ \Delta\psi_2 \\ \cdots \\ \Delta\psi_n \end{bmatrix} - \frac{q}{mc} \begin{bmatrix} \mathbf{A}_x \\ \mathbf{A}_y \end{bmatrix} \psi_0^*\psi_0 \tag{56}$$

8.2 Commutation Relations

To construct commutation relations between the nonlocal matrices representing the components of the gradient operator and the diagonal matrices representing the position operators, we introduce the set of commutators:

$$[\boldsymbol{\nabla}_i, x_j]\, f \rightarrow \begin{bmatrix} [\boldsymbol{\nabla}_x, x]\, f & [\boldsymbol{\nabla}_x, y]\, f \\ [\boldsymbol{\nabla}_y, x]\, f & [\boldsymbol{\nabla}_y, y]\, f \end{bmatrix} \tag{57}$$

and evaluate the commutators at the central point $\mathbf{x}_0$. To illustrate how this is done, we compute one of the elements of this 2×2 matrix:

$$[\boldsymbol{\nabla}_x, y]\, f = \boldsymbol{\nabla}_x(yf) - y(\boldsymbol{\nabla}_x f) \tag{58}$$

The first term in Eq. (58) is approximated by

$$\boldsymbol{\nabla}_x(yf) = \begin{bmatrix} wx_1 & wx_2 & \cdots & wx_n \end{bmatrix} \begin{bmatrix} (y_0 + \Delta y_1)(f_0 + \Delta f_1) \\ (y_0 + \Delta y_2)(f_0 + \Delta f_2) \\ \cdots \\ (y_0 + \Delta y_n)(f_0 + \Delta f_n) \end{bmatrix} \tag{59}$$

The second term in Eq. (58) is approximated by

$$y\boldsymbol{\nabla}_x(f) = y_0 \left[\begin{array}{cccc} wx_1 & wx_2 & \cdots & wx_n \end{array} \right] \left[\begin{array}{c} f_0 + \Delta f_1 \\ f_0 + \Delta f_2 \\ \cdots \\ f_0 + \Delta f_n \end{array} \right] \tag{60}$$

The first term (Eq (59)) has one contribution $y_0 f_0$ of zeroth order, two classes of contributions of first order, $y_0 \Delta f_i$ and $f_0 \Delta y_i$, and one class of contributions of second order: $\Delta y_i \Delta f_i$. The second term (Eq. (60)) has one zeroth order term, $y_0 f_0$, and one class of first order terms, $y_0 \Delta f_i$. The zero order contributions from the two terms cancel, as well as the first order terms of the form $y_0 \Delta f_i$. The second order terms can be neglected as unimportant for a fine-enough mesh, leaving

$$\boldsymbol{\nabla}_x(yf) - y(\boldsymbol{\nabla}_x f) = \left[\begin{array}{cccc} wx_1 & wx_2 & \cdots & wx_n \end{array} \right] \left[\begin{array}{c} \Delta y_1 \\ \Delta y_2 \\ \cdots \\ \Delta y_n \end{array} \right] f_0 \tag{61}$$

The remaining three commutation relations in Eq. (57) are treated similarly, with the result

$$\left[\begin{array}{cc} [\boldsymbol{\nabla}_x, x]\, f & [\boldsymbol{\nabla}_x, y]\, f \\ [\boldsymbol{\nabla}_y, x]\, f & [\boldsymbol{\nabla}_y, y]\, f \end{array} \right] = \left[\begin{array}{cccc} wx_1 & wx_2 & \cdots & wx_n \\ wy_1 & wy_2 & \cdots & wy_n \end{array} \right] \left[\begin{array}{cc} \Delta x_1 & \Delta y_1 \\ \Delta x_2 & \Delta y_2 \\ \cdots & \vdots \\ \Delta x_n & \Delta y_n \end{array} \right] f_0 =$$

$$(T^t T)^{-1}(T^t T) f_0 = \left[\begin{array}{cc} f & 0 \\ 0 & f \end{array} \right] \tag{62}$$

(recall Eq. (49)). This holds for any (reasonable) function f. Thus, to second order terms the commutation relations of $\boldsymbol{\nabla}$ and $\mathbf{x}$ are satisfied in the finite element approximation. The result is a diagonal matrix with $+1$ along the diagonal: the identity matrix I_2. Using $\boldsymbol{\nabla}$ to represent the momentum operator $\mathbf{p} = (\hbar/i)\boldsymbol{\nabla}$ we recover the finite element approximation statement of the canonical commutation relations: $[\mathbf{p}_i, \mathbf{x}_j] = (\hbar/i)\delta_{ij}$.

9 Summary

Convenient methods exist for transforming some partial differential equations to matrix equations. We have reviewed three effective methods for transforming the Schrödinger eigenvalue equation to a matrix eigenvalue equation. All three methods involve sowing a set of vertices throughout the domain of interest, introducing a set of basis functions that are highly localized around each vertex, and then introducing a prescription for computing matrices representing the operators of importance: the kinetic energy matrix, the potential energy matrix, the magnetic matrix, and an overlap matrix. The rows and columns of these $N_V \times N_V$ matrices are labeled by the N_V vertices. The resulting generalized eigenvalue equation is then treated by increasingly effective matrix diagonalization algorithms.

These methods are effective because, at the practical level, there is little difference between knowing the value of a wavefunction at *all* points in a continuous domain, as would result from solving a partial differential equation, and knowing the value of the wavefunction at only a finite but well-sampled set of points within the domain, as results from solving a matrix equation. In other words, even when you display a solution of a partial differential equation on a screen, only a finite set of point can be plotted. And integrations $\int\int \psi_\beta(\mathbf{x})^* \mathcal{T}(\mathbf{x}) \psi_\alpha(\mathbf{x}) d\mathbf{x}$ of an operator $\mathcal{T}(\mathbf{x})$ acting on a wavefunction, multiplied by the complex conjugate of a wavefunction, are usually carried out by sampling the integrand at a finite number of points rather than carrying out an analytic integral (usually never possible).

The three methods we have reviewed are based on different choices of localized basis functions: gaussian functions, pyramidal functions, and Kronecker delta functions. We have reviewed the simplest possible implementations of these powerful methods which use only one basis function at each vertex. As a rule of thumb, every worker who has devoted a good part of his/her life to using these methods has introduced a wrinkle of complexity in a search for increased power/efficiency, so there are very many extensions of gaussian- and fem-based computational methods currently available.