Part III

Variational Methods

Chapter 8

Ritz Method in Quantum Mechanics

8.1 The Variational Principle

The variational method provides a powerful approach to approximate the ground state energy of a quantum system. The principle states that for any trial wavefunction Ψ, the expectation value of the Hamiltonian $\hat{H}$ will always be greater than or equal to the ground state energy E_0:

$$E_0 \leq \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{8.1.1}$$

This principle forms the foundation of the Ritz method and other variational techniques in quantum mechanics.

8.2 Application of the Ritz Method

In the Ritz method, we express the trial wavefunction Ψ as a linear combination of basis functions $\{\Psi_i\}$:

$$|\Psi\rangle = \sum_{i=1}^{N} c_i |\Psi_i\rangle \tag{8.2.1}$$

where c_i are variational parameters to be determined. Starting from the variational principle, we want to minimize the energy expectation value:

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{8.2.2}$$

Substituting our trial wavefunction (Eq. 8.2.1) into this expression:

$$E[c_i] = \frac{\left\langle \sum_j c_j^* \Psi_j \right| \hat{H} \left| \sum_i c_i \Psi_i \right\rangle}{\left\langle \sum_j c_j^* \Psi_j \middle| \sum_i c_i \Psi_i \right\rangle} = \frac{\sum_{i,j} c_i^* c_j \langle \Psi_j | \hat{H} | \Psi_i \rangle}{\sum_{i,j} c_i^* c_j \langle \Psi_j | \Psi_i \rangle} \tag{8.2.3}$$

Now, we introduce the Hamiltonian matrix $\mathbf{H}$ and the overlap matrix $\mathbf{S}$:

$$H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle, \quad S_{ij} = \langle \Psi_i | \Psi_j \rangle \tag{8.2.4}$$

With these definitions, we can rewrite the energy functional in matrix form:

$$E[c_i] = \frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{S} \mathbf{c}} \tag{8.2.5}$$

where $\mathbf{c}$ is the column vector of coefficients c_i.

8.2.1 Minimization Procedure

To find the best approximation to the ground state, we need to minimize $E[c_i]$ with respect to the coefficients c_i. We do this by setting the derivative of $E[c_i]$ with respect to c_k^* to zero:

$$\frac{\partial E}{\partial c_k^*} = 0 \tag{8.2.6}$$

Let's evaluate this derivative:

$$\frac{\partial E}{\partial c_k^*} = \frac{\partial}{\partial c_k^*} \left(\frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{S} \mathbf{c}} \right) = \frac{(\mathbf{H}\mathbf{c})_k (\mathbf{c}^\dagger \mathbf{S} \mathbf{c}) - (\mathbf{c}^\dagger \mathbf{H} \mathbf{c})(\mathbf{S}\mathbf{c})_k}{(\mathbf{c}^\dagger \mathbf{S} \mathbf{c})^2} \tag{8.2.7}$$

Setting this to zero and simplifying:

$$(\mathbf{H}\mathbf{c})_k (\mathbf{c}^\dagger \mathbf{S} \mathbf{c}) = (\mathbf{c}^\dagger \mathbf{H} \mathbf{c})(\mathbf{S}\mathbf{c})_k \tag{8.2.8}$$

This equation must hold for all k. We can rewrite it as:

$$\mathbf{H}\mathbf{c} = \lambda \mathbf{S}\mathbf{c} \tag{8.2.9}$$

where $\lambda = \frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{S} \mathbf{c}}$.

8.2.2 Generalized Eigenvalue Problem

Equation 8.2.9 is known as the generalized eigenvalue problem. To solve the generalized eigenvalue problem numerically, we can transform it into a standard eigenvalue problem. This is achieved through the following steps:

1. Since $\mathbf{S}$ is Hermitian positive definite, we can compute its matrix square root:

$$\mathbf{S}^{1/2} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^\dagger \tag{8.2.10}$$

   where $\mathbf{U}$ and $\mathbf{\Lambda}$ are obtained from the eigendecomposition of $\mathbf{S}$.

2. We can then transform the Hamiltonian to create a new Hermitian matrix:

$$\mathbf{H}' = \mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2} \tag{8.2.11}$$

3. This transforms our generalized eigenvalue problem into a standard eigenvalue problem:

$$\mathbf{H}'\mathbf{c}' = \lambda\mathbf{c}' \tag{8.2.12}$$

   where $\mathbf{c}' = \mathbf{S}^{1/2}\mathbf{c}$.

The transformed problem can now be solved using standard eigenvalue algorithms such as the Jacobi method.

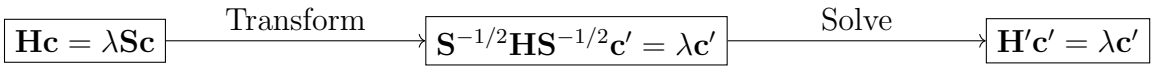

Figure 8.2.1: Transformation of the generalized eigenvalue problem to standard form.

8.2.3 Example: Two-State System

To illustrate these concepts, let's consider a simple two-state system. Suppose we have a basis of two states $|\Psi_1\rangle, |\Psi_2\rangle$. Our trial wavefunction is:

$$|\Psi\rangle = c_1|\Psi_1\rangle + c_2|\Psi_2\rangle \tag{8.2.13}$$

The Hamiltonian and overlap matrices are 2x2:

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \tag{8.2.14}$$

The generalized eigenvalue problem becomes:

$$\begin{pmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} \\ H_{21} - \lambda S_{21} & H_{22} - \lambda S_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0 \tag{8.2.15}$$

The characteristic equation is:

$$\det(\mathbf{H} - \lambda \mathbf{S}) = 0 \tag{8.2.16}$$

Expanding this determinant gives a quadratic equation in λ, which can be solved to find the two energy eigenvalues. The corresponding eigenvectors give the coefficients c_1 and c_2 for each approximate eigenstate.

This two-state example, while simple, illustrates all the key features of the Ritz method: formulation of the trial wavefunction, construction of the Hamiltonian and overlap matrices, and solution of the generalized eigenvalue problem to obtain approximate energies and wavefunctions.

8.3 Choice of Basis Functions

The effectiveness of the Ritz method heavily depends on the choice of basis functions. Some common choices include:

- Harmonic oscillator eigenfunctions
- Hydrogen-like atomic orbitals
- Plane waves (for periodic systems)
- Gaussian functions

The choice is often guided by the symmetry of the problem and the behavior of the expected solution.

8.4 Example: Anharmonic Oscillator

Let's apply the Ritz method to find the ground state of a harmonic oscillator with a small anharmonic (x4) perturbation. This system is described by the Hamiltonian:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 + \lambda x^4 \tag{8.4.1}$$

where λ is a small parameter characterizing the strength of the anharmonic term. We'll use the eigenfunctions of the unperturbed harmonic oscillator as our basis set:

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \tag{8.4.2}$$

where $H_n(x)$ are the Hermite polynomials.

8.4.1 Matrix Elements

The Hamiltonian matrix elements are:

$$H_{mn} = \langle \psi_m | \hat{H} | \psi_n \rangle = H^{(0)}_{mn} + \lambda V_{mn} \tag{8.4.3}$$

where $H^{(0)}_{mn}$ are the matrix elements of the unperturbed harmonic oscillator Hamiltonian, and V_{mn} are the matrix elements of the x^4 term.

The unperturbed terms are diagonal by construction:

$$H^{(0)}_{mn} = \hbar\omega(n + \frac{1}{2})\delta_{mn} \tag{8.4.4}$$

For the x^4 term, we can use the properties of Hermite polynomials to calculate:

$$\begin{aligned} V_{mn} &= \langle \psi_m | x^4 | \psi_n \rangle \\ &= \frac{\hbar^2}{4m^2\omega^2} \Big[(2n^2 + 2n + 1)\delta_{mn} \\ &\quad + \sqrt{(n+1)(n+2)(n+3)(n+4)}\,\delta_{m,n+4} \\ &\quad + \sqrt{n(n-1)(n-2)(n-3)}\,\delta_{m,n-4} \\ &\quad + 6(n+1)\sqrt{n+2}\,\delta_{m,n+2} \\ &\quad + 6n\sqrt{n-1}\,\delta_{m,n-2} \Big] \end{aligned} \tag{8.4.5}$$

8.4.2 Numerical Implementation

Let's implement this in Python, using a truncated basis of the first few harmonic oscillator eigenstates:

```
import numpy as np
from scipy.linalg import eigh

# Constants
hbar = 1.0
m = 1.0
omega = 1.0
lambda_param = 0.01  # Strength of anharmonic term

def hamiltonian_element(m, n, lambda_param):
    if m == n:
        return hbar * omega * (n + 0.5) + lambda_param * (hbar**2 / (4 * m**2 *
    omega**2)) * (2*n**2 + 2*n + 1)
    elif m == n + 4 or n == m + 4:
        k = min(m, n)
```

```
        return lambda_param * (hbar**2 / (4 * m**2 * omega**2)) * np.sqrt((k+1)*(k+2)*(k+3)*(k+4))
    elif m == n + 2 or n == m + 2:
        k = min(m, n)
        return lambda_param * (hbar**2 / (4 * m**2 * omega**2)) * 6 * (k+1) * np.sqrt(k+2)
    else:
        return 0

# Number of basis states to use
N = 10

# Construct Hamiltonian matrix
H = np.zeros((N, N))
for i in range(N):
    for j in range(N):
        H[i, j] = hamiltonian_element(i, j, lambda_param)

# Solve eigenvalue problem
energies, states = eigh(H)

print(f"Ground state energy: {energies[0]}")
print(f"First excited state energy: {energies[1]}")

# Compare with unperturbed energies
print(f"Unperturbed ground state energy: {hbar * omega * 0.5}")
print(f"Unperturbed first excited state energy: {hbar * omega * 1.5}")
```

This script constructs the Hamiltonian matrix for the anharmonic oscillator and solves for its eigenvalues and eigenvectors.

8.4.3 Results and Analysis

Let's visualize how the ground state energy changes with the strength of the anharmonic term over a wider range, comparing it with perturbation theory:

Figure 8.4.1 shows how the ground state energy increases with the strength of the anharmonic term over a range of λ from 0 to 1. We can observe that for larger λ, the energy increase becomes non-linear, significantly deviating from the prediction of first-order perturbation theory.

8.4.4 Convergence Analysis

To assess the accuracy of our Ritz method approximation for a strong perturbation, we study how the ground state energy converges as we increase the number of basis functions:

Figure 8.4.2 demonstrates the convergence of the ground state energy as we increase the number of basis functions for $\lambda = 1$. We observe that significantly more basis functions are required to

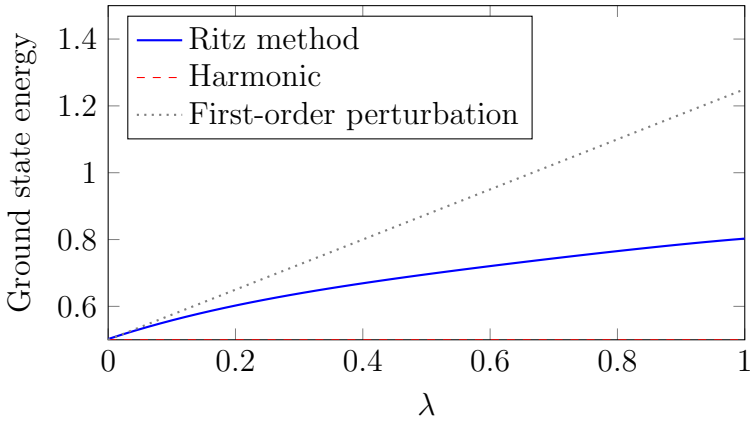

Figure 8.4.1: Ground state energy of the anharmonic oscillator vs. perturbation strength ($0 \leq \lambda \leq 1$)

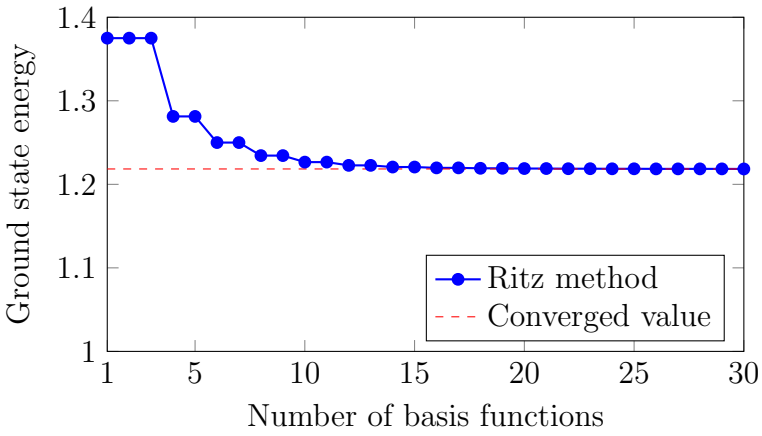

Figure 8.4.2: Convergence of ground state energy with increasing basis size for $\lambda = 1$

achieve convergence compared to the case with smaller λ, reflecting the strong perturbation to the harmonic oscillator.

8.4.5 Discussion

These results illustrate several important aspects of the Ritz method applied to the anharmonic oscillator:

1. Non-linear response: As seen in Figure 8.4.1, the ground state energy increases non-linearly with λ, significantly deviating from first-order perturbation theory for larger λ values. This highlights the limitations of perturbation theory and the advantage of the Ritz method in handling stronger perturbations.

2. Perturbation theory comparison: The first-order perturbation theory provides a good approximation for small λ but quickly becomes inaccurate as λ increases. The Ritz method, on the other hand, can provide accurate results even for large λ, provided we use enough basis functions.

3. Convergence behavior: Figure 8.4.2 shows that for a strong perturbation ($\lambda = 1$), many more basis functions are needed to achieve convergence compared to weaker perturbations. This demonstrates the Ritz method's ability to handle regimes far beyond the applicability of perturbation theory.

4. Basis choice: The harmonic oscillator eigenfunctions remain a viable choice of basis even for strong perturbations, as evidenced by the eventual convergence. However, the number of basis functions required increases significantly with the strength of the perturbation.

5. Computational considerations: As we increase the number of basis functions, the computational cost grows substantially. For very strong perturbations or higher precision requirements, one might consider alternative basis choices, such as a basis that incorporates some of the anharmonic character, or other computational techniques like the Numerov method or Rayleigh-Ritz with a different basis.

6. Excited states: While we focused on the ground state, the method simultaneously provides approximations to excited state energies, which would show similar convergence behavior but likely require even more basis functions for accurate results.

This anharmonic oscillator example demonstrates the power and versatility of the Ritz method in handling perturbations of various strengths. It provides a smooth transition between the regimes of validity for perturbation theory and numerical techniques for strongly perturbed systems, making it a valuable tool in computational quantum mechanics.

8.5 Conclusion

The Ritz method is a powerful variational technique in quantum mechanics that allows for systematic improvement of approximate solutions. By choosing an appropriate basis set and solving a generalized eigenvalue problem, we can obtain accurate estimates of energy levels and wavefunctions for a wide range of quantum systems.

The method's flexibility in choice of basis functions, its ability to provide upper bounds to the true ground state energy, and its straightforward numerical implementation make it a valuable tool in computational quantum mechanics. As we've seen in the harmonic oscillator example, even with a relatively small number of basis functions, the Ritz method can provide excellent approximations to exact solutions.

Understanding and applying the Ritz method not only provides practical computational tools but also deepens our insight into the nature of quantum mechanical systems and the power of variational approaches in physics.

Chapter 9

Galerkin Method for Time-Dependent Problems

In this chapter we present a comprehensive and rigorous derivation of the Galerkin method as applied to the time-dependent Schrödinger equation. Our approach begins with an ansatz for the wavefunction, followed by a derivation of the Galerkin conditions via the minimization of the L^2 norm of the residual. We then derive the resulting matrix system for the time evolution of the expansion coefficients and describe how to solve the time-dependent problem in the case of a non-orthogonal basis.

9.1 The Time-Dependent Schrödinger Equation and the Ansatz

We consider the time-dependent Schrödinger equation in one spatial dimension:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = \hat{H}\,\psi(x,t), \tag{9.1.1}$$

with the Hamiltonian operator given by

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x), \tag{9.1.2}$$

where $\psi(x,t)$ is the wavefunction, $\hbar$ is the reduced Planck constant, m is the particle mass, and $V(x)$ is a prescribed potential.

In the Galerkin method we approximate the true solution $\psi(x,t)$ by a finite-dimensional expansion (ansatz):

$$\psi_K(x,t) = \sum_{k=1}^{K} c_k(t)\,\phi_k(x), \tag{9.1.3}$$

where $\{\phi_k(x)\}_{k=1}^K$ is a set of (not necessarily orthonormal) basis functions spanning a subspace of the Hilbert space, and the coefficients $c_k(t)$ (complex in general) encode the time dependence.

9.2 Derivation of the Galerkin Conditions

To approximate solutions to the time-dependent Schrödinger equation (9.1.1), we begin with the finite-dimensional ansatz (9.1.3) for the wavefunction, written in bra–ket notation:

$$|\psi_K(t)\rangle = \sum_{k=1}^{K} c_k(t)|\phi_k\rangle.$$

Since this ansatz generally does not satisfy (9.1.1) exactly, we define the *residual*

$$|R(t)\rangle = \Big(i\hbar\frac{\partial}{\partial t} - \hat{H}\Big)|\psi_K(t)\rangle, \tag{9.2.1}$$

which quantifies the discrepancy between our approximate and the true solution. For an exact solution, the residual vanishes, i.e., $\langle R(t)|R(t)\rangle = 0$; for an approximate solution, it is nonzero.

We now seek the best possible evolution for the ansatz by minimizing the squared norm of the residual,

$$\|R(t)\|^2 = \langle R(t)|R(t)\rangle, \tag{9.2.2}$$

with respect to the time derivatives $\dot{c}_k(t)$ of the coefficients, while holding the values $c_k(t)$ fixed at time t. The idea is that if at time t our ansatz is exact, then the only source of error at time $t+\epsilon$ comes from the choice of the time derivatives $\dot{c}_k(t)$.

Substituting the ansatz into (9.2.1) yields

$$|R(t)\rangle = \sum_{k=1}^{K}\Big(i\hbar\,\dot{c}_k(t)|\phi_k\rangle - c_k(t)\hat{H}|\phi_k\rangle\Big). \tag{9.2.3}$$

We now compute the partial derivative of the squared norm $\langle R(t)|R(t)\rangle$ with respect to the complex conjugate of the time derivative, $\dot{c}_j^*(t)$. A straightforward calculation shows that

$$\frac{\partial\langle R(t)|R(t)\rangle}{\partial\dot{c}_j^*(t)} = i\hbar\langle\phi_j|R(t)\rangle. \tag{9.2.4}$$

To ensure that the evolution of the ansatz minimizes the instantaneous growth of the residual, we require that the above derivative vanishes for each $j = 1,\dots,K$, i.e.,

$$\langle\phi_j|R(t)\rangle = 0, \quad j = 1,\dots,K.$$

This immediately implies the so-called *Galerkin condition*

$$\left\langle \phi_j \middle| \left(i\hbar \frac{\partial}{\partial t} - \hat{H} \right) \psi_K(t) \right\rangle = 0, \quad j = 1, \ldots, K, \tag{9.2.5}$$

which can be equivalently stated saying that we are imposing that the residual (error) vector must be orthogonal to the directions that span the sub-space included in our ansatz. This is rather intuitive, in the sense that we are imposing that within the chosen subspace the error is vanishing, whereas in general it will be finite along directions that are outside the chosen subspace.

9.3 Matrix Formulation

Let us define the overlap and Hamiltonian matrices as follows:

$$S_{jk} = \langle \phi_j \mid \phi_k \rangle, \quad H_{jk} = \langle \phi_j \mid \hat{H} \mid \phi_k \rangle. \tag{9.3.1}$$

Note that when the basis $\{\phi_k\}$ is non-orthogonal, the matrix $\mathbf{S}$ is not the identity but is Hermitian and positive definite.

Substitute the ansatz (9.1.3) into the Galerkin condition:

$$\begin{aligned} \langle \phi_j \mid (i\hbar\partial_t - \hat{H})\, \psi_K \rangle &= i\hbar \sum_{k=1}^{K} \dot{c}_k(t) \langle \phi_j \mid \phi_k \rangle - \sum_{k=1}^{K} c_k(t) \langle \phi_j \mid \hat{H} \mid \phi_k \rangle \\ &= i\hbar \sum_{k=1}^{K} S_{jk} \dot{c}_k(t) - \sum_{k=1}^{K} H_{jk} c_k(t) = 0, \quad j = 1, \ldots, K. \end{aligned} \tag{9.3.2}$$

In compact matrix notation, letting

$$|c(t)\rangle \doteq \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_K(t) \end{pmatrix},$$

we obtain the system

$$i\hbar\, \mathbf{S}\, \frac{d|c(t)\rangle}{dt} = \mathbf{H}|c(t)\rangle. \tag{9.3.3}$$

9.4 Time Evolution in a Non-Orthogonal Basis

The time evolution of the coefficients is governed by the differential equation (9.3.3). In order to solve it, we use a method similar to what discussed in the previous lecture. First of all, since the overlap matrix $\mathbf{S}$ is Hermitian positive definite, it admits an eigendecomposition:

$$\mathbf{S} = \mathbf{U}\, \mathbf{\Lambda}\, \mathbf{U}^\dagger, \tag{9.4.1}$$

where $\mathbf{U}$ is unitary and $\mathbf{\Lambda}$ is a diagonal matrix with positive entries. We define the square root and its inverse by

$$\mathbf{S}^{1/2} = \mathbf{U}\,\mathbf{\Lambda}^{1/2}\,\mathbf{U}^\dagger, \quad \mathbf{S}^{-1/2} = \mathbf{U}\,\mathbf{\Lambda}^{-1/2}\,\mathbf{U}^\dagger. \tag{9.4.2}$$

Introduce the transformed Hamiltonian

$$\mathbf{H}' = \mathbf{S}^{-1/2}\,\mathbf{H}\,\mathbf{S}^{-1/2}, \tag{9.4.3}$$

and define the transformed coefficient vector

$$|c'\rangle = \mathbf{S}^{1/2}|c\rangle. \tag{9.4.4}$$

Substituting $|c\rangle = \mathbf{S}^{-1/2}|c'\rangle$ into (9.3.3) gives

$$i\hbar\,\frac{d|c'(t)\rangle}{dt} = \mathbf{H}'|c'(t)\rangle,$$

thus we are left with a standard Schroedinger equation, for the modified Hamiltonian $\mathbf{H}'$.

9.4.1 Time Propagation

The solution of this equation is then found as usual, for example diagonalizing the modified hamiltonian, and then expressing the initial state in terms of its eigenvectors. Specifically, consider the eigenbasis of $\mathbf{H}'$, namely the states $\mathbf{H}'|d_l\rangle = E'_l|d_l\rangle$, then

$$\begin{aligned} |c'(t)\rangle &= e^{-\frac{i}{\hbar}\mathbf{H}'t}|c'(0)\rangle \\ &= \sum_l e^{-\frac{i}{\hbar}E'_l t}|d_l\rangle\langle d_l|c'(0)\rangle. \end{aligned}$$

Thus, in terms of the original coefficients we have:

$$\begin{aligned} \mathbf{S}^{1/2}|c(t)\rangle &= \sum_l e^{-\frac{i}{\hbar}E'_l t}|d_l\rangle\langle d_l|\mathbf{S}^{1/2}|c(0)\rangle \\ |c(t)\rangle &= \sum_l e^{-\frac{i}{\hbar}E'_l t}\mathbf{S}^{-1/2}|d_l\rangle\langle d_l|\mathbf{S}^{1/2}|c(0)\rangle. \end{aligned}$$

9.5 Example: Particle in a Double-Well Potential

Consider a particle in a one-dimensional double-well potential:

$$V(x) = V_0(x^2 - a^2)^2 \tag{9.5.1}$$

where V_0 determines the height of the barrier between wells, and $\pm a$ are the locations of the minima.

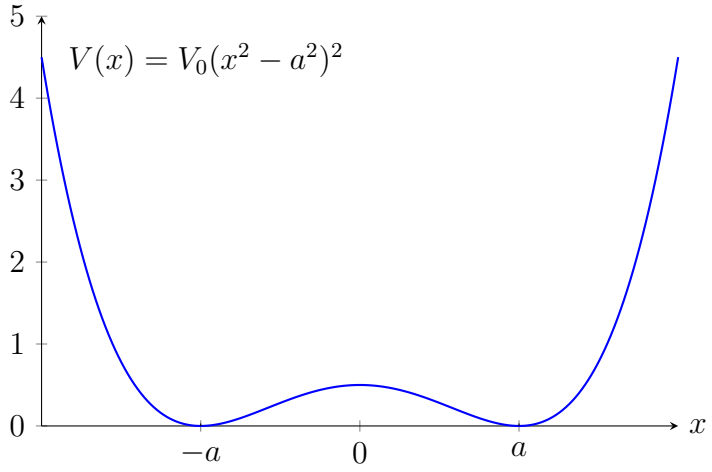

Figure 9.5.1: Sketch of the double-well potential

We use the eigenbasis of the harmonic oscillator centered at the midpoint between the two wells:

$$\phi_n(x) = \left(\frac{\alpha}{\pi}\right)^{1/4} \frac{1}{\sqrt{2^n n!}} H_n(\sqrt{\alpha}x) e^{-\alpha x^2/2} \tag{9.5.2}$$

where $H_n(x)$ are the Hermite polynomials and α is a parameter that can be optimized. The Hamiltonian matrix elements can be calculated analytically:

$$\begin{aligned}
(H_K)_{nm} = \hbar\alpha(n+\frac{1}{2})\delta_{nm} + V_0 \Bigg[& 3a^4\delta_{nm} + \frac{2n+1}{2\alpha}\delta_{nm} \\
& + \frac{\sqrt{(n+1)(n+2)}}{2\alpha}\delta_{m,n+2} + \frac{\sqrt{n(n-1)}}{2\alpha}\delta_{m,n-2}\Bigg] \\
& + V_0 \Bigg[\frac{3}{4\alpha^2}(2n^2+2n+1)\delta_{nm} + \frac{\sqrt{(n+1)(n+2)}}{2\alpha^2}(2n+3)\delta_{m,n+2} \\
& + \frac{\sqrt{n(n-1)}}{2\alpha^2}(2n-1)\delta_{m,n-2} + \frac{\sqrt{(n+1)(n+2)(n+3)(n+4)}}{4\alpha^2}\delta_{m,n+4} \\
& + \frac{\sqrt{n(n-1)(n-2)(n-3)}}{4\alpha^2}\delta_{m,n-4}\Bigg]
\end{aligned} \tag{9.5.3}$$

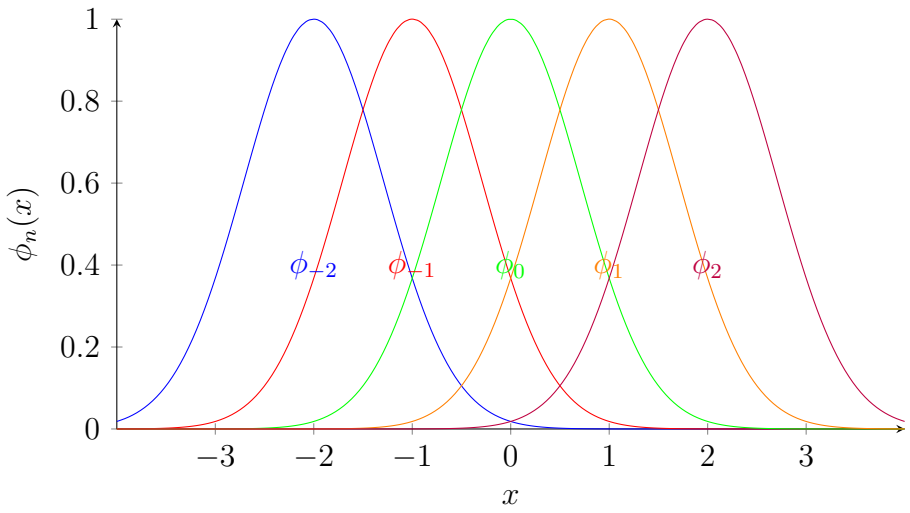

Figure 9.6.1: Sketch of localized Gaussian basis functions

9.6 Example: Localized Gaussian Basis Set

Another powerful choice for the basis functions in the Galerkin method is a set of localized Gaussian functions. These functions are particularly useful for problems where the wavefunction is expected to be localized in certain regions of space.

9.6.0.1 Gaussian Basis Functions

We define our Gaussian basis functions as:

$$\phi_n(x) = \left(\frac{2\alpha}{\pi}\right)^{1/4} \exp\left(-\alpha(x - x_n)^2\right) \tag{9.6.1}$$

where α is a parameter controlling the width of the Gaussians, and x_n are the centers of the Gaussians.

9.6.1 Matrix Elements for Harmonic Oscillator

Let's consider a harmonic oscillator potential $V(x) = \frac{1}{2}m\omega^2 x^2$. We'll calculate the matrix elements of the Hamiltonian analytically using our Gaussian basis.

The kinetic energy matrix elements are:

$$\begin{aligned} T_{mn} &= -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty}\phi_m(x)\frac{d^2}{dx^2}\phi_n(x)dx \\ &= \frac{\hbar^2\alpha}{2m}S_{mn}\left[1-\alpha(x_m-x_n)^2\right] \end{aligned} \tag{9.6.2}$$

where S_{mn} is the overlap matrix element:

$$S_{mn} = \exp\left(-\frac{\alpha}{2}(x_m-x_n)^2\right) \tag{9.6.3}$$

The potential energy matrix elements for the harmonic oscillator are:

$$\begin{aligned} V_{mn} &= \frac{1}{2}m\omega^2\int_{-\infty}^{\infty}\phi_m(x)x^2\phi_n(x)dx \\ &= \frac{m\omega^2}{4\alpha}S_{mn}\left[1+2\alpha(x_m^2+x_n^2)-4\alpha^2(x_m-x_n)^2\right] \end{aligned} \tag{9.6.4}$$

The total Hamiltonian matrix elements are then $H_{mn} = T_{mn} + V_{mn}$.

Chapter 10

Variational Formulation for Classical Trajectories

In classical mechanics, the trajectory of a system between two fixed times can be determined by requiring that the action be stationary. In the continuous setting, the action is defined by

$$S[u] = \int_0^T L\big(u(t), \dot{u}(t)\big)dt, \tag{10.0.1}$$

where $u(t)$ is the generalized coordinate and L is the Lagrangian. Imposing $\delta S = 0$ under variations $\delta u(t)$ that vanish at the endpoints $u(0)$ and $u(T)$ leads to the Euler–Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{u}}\right) - \frac{\partial L}{\partial u} = 0. \tag{10.0.2}$$

This differential equation is the continuous condition for a trajectory to be physically admissible. The stationary action principle of classical mechanics is a variational principle that is extremely useful in cases when one seeks a solution to the equations of motion with constrained end points. In this Chapter we will see how we can solve this constrained problem numerically.

10.1 Discrete Action with Fixed Endpoints

To proceed further, we discretize the time interval $[0, T]$. Suppose we split the interval into $(n+1)$ subintervals by the nodes

$$t_0, t_1, \ldots, t_n, t_{n+1},$$

with uniform spacing $\Delta t = T/(n+1)$. Let x_k denote the approximation of $u(t_k)$. The endpoints x_0 and x_{n+1} are fixed by the boundary conditions, and the unknowns are the internal values

$$\mathbf{x} = (x_1, x_2, \ldots, x_n).$$

A simple finite-difference discretization of the kinetic energy and the potential energy leads to the discrete action

$$S_n(\mathbf{x}) = \sum_{k=0}^{n} \Delta t \left[\frac{m}{2} \left(\frac{x_{k+1} - x_k}{\Delta t} \right)^2 - V(x_k) \right], \tag{10.1.1}$$

where m is the mass and $V(x)$ is the potential energy. The goal is to determine the internal points $\mathbf{x}$ such that the discrete action is stationary.

10.2 Gradient and Hessian

The discretized action we have written above has now become a function (rather than a functional, as in the initial formulation) of the discrete trajectory points $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. In order to make the action stationary, we will need to use some numerical method to find a solution such that

$$\frac{\partial S_n}{\partial x_i} = 0, \quad i = 1, 2, \ldots, n. \tag{10.2.1}$$

In the following we will use Newton's method, for which it is essential to compute both the gradient of S_n and its Hessian (the Jacobian of the gradient). The Hessian tells us how the gradient changes with the variables and is crucial for constructing a quadratic approximation of the action in the neighborhood of a trial solution.

10.2.1 Gradient

We now derive the gradient of $S_n(\mathbf{x})$. We have seen above that the contribution of each interval to the discretized action is

$$\frac{m}{2\Delta t}(x_{k+1} - x_k)^2 - \Delta t\, V(x_k).$$

When we differentiate with respect to x_i, we notice that x_i appears in two consecutive terms, namely for $k = i-1$ and $k = i$. A compact way to express this is to use Kronecker delta functions. In particular, one can write

$$\frac{\partial S_n}{\partial x_i} = \sum_{k=0}^{n} \left[\frac{m}{\Delta t^2}(x_{k+1} - x_k)(\delta_{i,k+1} - \delta_{i,k}) - \frac{\partial V(x_k)}{\partial x_k}\, \delta_{i,k} \right] \Delta t.$$

Since the only nonzero contributions occur when $k = i$ or $k = i \pm 1$, the gradient simplifies to

$$\frac{\partial S_n}{\partial x_i} = \frac{m}{\Delta t}\Big(2x_i - x_{i-1} - x_{i+1}\Big) - \Delta t\, \frac{\partial V(x_i)}{\partial x_i}. \tag{10.2.2}$$

Here x_0 and x_{n+1} are known, so the derivatives are computed only for the free variables $x_1, \ldots, x_n$.

10.2.2 Hessian

Next, we differentiate the gradient with respect to x_j to obtain the Hessian. In Kronecker-delta notation, this reads

$$\frac{\partial^2 S_n}{\partial x_i \partial x_j} = \frac{m}{\Delta t}\Big(2\delta_{j,i} - \delta_{j,i-1} - \delta_{j,i+1}\Big) - \Delta t\, \frac{\partial^2 V(x_i)}{\partial x_i^2}\delta_{j,i}. \tag{10.2.3}$$

This shows that the Hessian is tridiagonal: the diagonal entries are given by

$$\frac{2m}{\Delta t} - \Delta t\, \frac{\partial^2 V(x_i)}{\partial x_i^2},$$

and the entries immediately above and below the diagonal are

$$-\frac{m}{\Delta t}.$$

This analytical expression for the Hessian is essential for the efficient application of Newton's method.

Derivation of the Newton Update

Derivation of Newton's Method Update:
For a nonlinear function $\mathbf{g}(\mathbf{x})$ in $\mathbb{R}^n$, Newton's method aims to find a zero of $\mathbf{g}$. Given a current approximation $\mathbf{x}^{(m)}$, we linearize $\mathbf{g}$ about $\mathbf{x}^{(m)}$ using its Jacobian $J(\mathbf{x}^{(m)})$:

$$\mathbf{g}(\mathbf{x}) \approx \mathbf{g}\big(\mathbf{x}^{(m)}\big) + J\big(\mathbf{x}^{(m)}\big)(\mathbf{x} - \mathbf{x}^{(m)}).$$

Setting $\mathbf{g}(\mathbf{x}) = 0$ and letting $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^{(m)}$, we obtain the linear system

$$J\big(\mathbf{x}^{(m)}\big)\,\delta\mathbf{x} = -\mathbf{g}\big(\mathbf{x}^{(m)}\big).$$

The updated solution is then given by

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \delta\mathbf{x}.$$

Often a damping factor α (with $0 < \alpha \leq 1$) is introduced to control the step size:

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \alpha\,\delta\mathbf{x}.$$

In our case, $\mathbf{g}(\mathbf{x})$ is the gradient $\nabla S_n(\mathbf{x})$ and $J(\mathbf{x})$ is the Hessian of S_n as given in (10.2.3).

10.3 Newton's Method for Stationarizing the Discrete Action

To find the stationary point of $S_n(\mathbf{x})$, we seek $\mathbf{x}$ such that $\nabla S_n(\mathbf{x}) = 0$. Newton's method iteratively solves

$$J\big(\mathbf{x}^{(m)}\big)\,\delta\mathbf{x} = -\nabla S_n\big(\mathbf{x}^{(m)}\big), \tag{10.3.1}$$

and updates

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \alpha\,\delta\mathbf{x}, \tag{10.3.2}$$

with $0 < \alpha \leq 1$ chosen to damp the step and avoid overshooting. The process is repeated until $\|\nabla S_n(\mathbf{x}^{(m)})\|$ is below a prescribed tolerance.

10.4 Specializing to a Given Potential

For example, if we specialize to the potential

$$V(x) = \frac{1}{2}kx^2 + \frac{1}{4}\lambda x^4,$$

then

$$\frac{\partial V(x)}{\partial x} = kx + \lambda x^3 \quad \text{and} \quad \frac{\partial^2 V(x)}{\partial x^2} = k + 3\lambda x^2.$$

Substituting these expressions into (10.2.2) and (10.2.3) yields explicit formulas for the gradient and Hessian. These can then be used in the Newton solver to compute the stationary trajectory that satisfies the fixed endpoints x_0 and x_{n+1}.