

Approximation Methods

1. Why approximation methods are needed

Exact solutions of the Schrödinger equation are exceptional. The harmonic oscillator, the hydrogen atom, the infinite square well, and a small number of other models are valuable not only because they can be solved exactly, but also because they provide reference problems for more realistic systems.

For a general Hamiltonian, the practical question is therefore not simply

Can the Schrödinger equation be solved exactly?

but rather

Can the problem be related to something we already understand well enough to construct a useful approximation and to assess when that approximation can be trusted?

There are several different ways of doing this, and they should be distinguished because the approximations are controlled in different ways.

1. **Stationary perturbation theory.** Start from an exactly solvable Hamiltonian and add a weak time-independent perturbation. The perturbation is weak only in comparison with the relevant energy separations of the unperturbed problem.
2. **Optimised reference Hamiltonian.** When the original Hamiltonian contains no obvious small parameter, introduce a solvable reference problem containing auxiliary parameters and choose them so that the resulting approximation is as accurate as possible.
3. **Variational method.** Choose a physically motivated family of trial states and minimise the expectation value of the Hamiltonian. For the ground state, the variational principle provides a rigorous upper bound to the exact energy.
4. **Time-dependent perturbation theory.** Start from the known stationary states of a reference Hamiltonian and calculate how a weak time-dependent interaction produces transitions between them.
5. **Sudden approximation.** Consider the opposite type of approximation, controlled not by the weakness of the change but by its short duration. If the Hamiltonian changes on a time scale much shorter than the relevant intrinsic quantum time scales, the state has essentially no time to evolve during the change.

These methods fall into two broad classes.

In some cases there is an identifiable **small dimensionless parameter**. For example, in stationary perturbation theory the relevant quantity is typically

$$\frac{|V_{mn}|}{|E_n^{(0)} - E_m^{(0)}|}$$

while in the sudden approximation the relevant ratio is

$$\frac{|\Delta E|\tau}{\hbar}$$

(see below for the detailed definitions). Such quantities tell us directly what physical limit makes the approximation possible.

In other methods, such as the variational method or an optimised reference Hamiltonian, there need not be an explicit small parameter. Their reliability is assessed differently: through variational bounds, optimisation with respect to auxiliary parameters, comparison of successive approximations, or comparison with independently known limits.

Thus there is no single criterion for all approximation methods. What is common to them is that the approximation must come with a clear statement of **what is being approximated, why the chosen procedure is expected to work, and how its reliability can be assessed.**

2. Stationary perturbation theory: non-degenerate case

2.1 Reference problem and perturbation

Suppose the Hamiltonian can be written as

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}.$$

The operator $\hat{H}_0$ defines the **reference problem**. It is chosen so that its eigenvalues and eigenstates are known:

$$\hat{H}_0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle,$$

with

$$\langle m^{(0)} | n^{(0)} \rangle = \delta_{mn}.$$

The operator $\hat{V}$ describes the difference between the physical Hamiltonian and the reference Hamiltonian.

The parameter λ is introduced only to organise the expansion. At $\lambda = 0$ we have the exactly solvable reference problem. At $\lambda = 1$ we recover the physical Hamiltonian. We therefore solve the problem formally as a power series in λ and set $\lambda = 1$ at the end.

This device makes the order of each contribution unambiguous. It does **not** mean that λ is itself a new physical parameter.

We want the eigenvalues and eigenstates of

$$(\hat{H}_0 + \lambda \hat{V}) |n\rangle = E_n |n\rangle.$$

Because the eigenstates of $\hat{H}_0$ form a complete basis, the exact perturbed state can be expanded as

$$|n\rangle = \sum_m C_{mn} |m^{(0)}\rangle.$$

Substituting this expansion into the Schrödinger equation gives

$$\sum_m C_{mn} (\hat{H}_0 + \lambda \hat{V}) |m^{(0)}\rangle = E_n \sum_m C_{mn} |m^{(0)}\rangle.$$

Now project onto an arbitrary basis state $\langle n'^{(0)} |$. Using

$$\langle n'^{(0)} | \hat{H}_0 | m^{(0)} \rangle = E_m^{(0)} \delta_{n'm},$$

we obtain

$$(E_n - E_{n'}^{(0)}) C_{n'n} = \lambda \sum_m V_{n'm} C_{mn},$$

where

$$V_{n'm} = \langle n'^{(0)} | \hat{V} | m^{(0)} \rangle.$$

This coefficient equation is exact. No approximation has yet been made. The approximation begins when we assume that the unknown energy and coefficients can be expanded in powers of λ .

2.2 Perturbative expansions

For a non-degenerate unperturbed level n we write

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \lambda^3 E_n^{(3)} + \dots,$$

and

$$C_{mn} = C_{mn}^{(0)} + \lambda C_{mn}^{(1)} + \lambda^2 C_{mn}^{(2)} + \dots.$$

At $\lambda = 0$, the exact state must reduce to the unperturbed state $|n^{(0)}\rangle$. Therefore

$$C_{mn}^{(0)} = \delta_{mn}.$$

Substitute both Taylor series into the exact coefficient equation. We are assuming that these series represent the solution in a neighbourhood of $\lambda = 0$. Two Taylor series that are equal for all sufficiently small λ must have equal coefficients at each power of λ . We can therefore equate the coefficients of $\lambda^0, \lambda^1, \lambda^2, \dots$ separately.

This is the basic mechanism of perturbation theory: instead of solving one difficult equation, we solve a sequence of simpler equations order by order.

2.3 First-order energy correction

At first order in λ ,

$$(E_n^{(0)} - E_{n'}^{(0)}) C_{n'n}^{(1)} + E_n^{(1)} \delta_{n'n} = V_{n'n}.$$

To obtain the energy correction, set $n' = n$. Then

$$E_n^{(0)} - E_n^{(0)} = 0, \quad \delta_{nn} = 1,$$

so the equation reduces immediately to

$$\boxed{E_n^{(1)} = V_{nn} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle}.$$

The first-order correction to the energy is therefore the expectation value of the perturbation in the unperturbed state.

The result should not be interpreted as saying that the state remains unchanged to first order. The state **does** acquire a first-order correction, derived in the next subsection. However, the contribution of that first-order state correction to the energy starts only at second order. The cross terms involving $\hat{H}_0$ are proportional to the zeroth-order eigenvalue equation and cancel against the corresponding first-order normalisation term. Consequently, the only first-order contribution to the energy is the diagonal expectation value of $\hat{V}$ in the unperturbed state.

A useful warning follows immediately: a vanishing first-order energy shift does **not** imply that the perturbation has no first-order effect. The wavefunction may still change at first order through off-diagonal matrix elements.

2.4 First-order correction to the state

For $n' \neq n$, the Kronecker delta vanishes and

$$(E_n^{(0)} - E_{n'}^{(0)}) C_{n'n}^{(1)} = V_{n'n}.$$

Hence

$$C_{n'n}^{(1)} = \frac{V_{n'n}}{E_n^{(0)} - E_{n'}^{(0)}} \quad (n' \neq n).$$

We still need the diagonal coefficient $C_{nn}^{(1)}$. This point is often passed over too quickly.

Write the state through first order as

$$|n\rangle = |n^{(0)}\rangle + \lambda \left[C_{nn}^{(1)} |n^{(0)}\rangle + \sum_{m \neq n} C_{mn}^{(1)} |m^{(0)}\rangle \right] + O(\lambda^2).$$

Normalisation requires

$$\langle n|n\rangle = 1.$$

Keeping only terms through first order gives

$$1 + \lambda \left(C_{nn}^{(1)} + \overline{C_{nn}^{(1)}} \right) = 1,$$

so

$$2 \operatorname{Re} C_{nn}^{(1)} = 0.$$

Thus normalisation fixes the real part of $C_{nn}^{(1)}$ to zero. There remains freedom to choose a constant phase convention for each eigenvector. We choose this convention so that the coefficient of the unperturbed state is real to the required order, such that

$$C_{nn}^{(1)} = 0.$$

This is simply a convenient choice of overall phase, made after imposing normalisation. It removes a coefficient that carries no independent physical information at this order.

The perturbed state is therefore

$$|n\rangle = |n^{(0)}\rangle + \lambda \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle + O(\lambda^2).$$

At the end we set $\lambda = 1$.

The perturbation mixes the original state with other unperturbed states. The amount of mixing is controlled by two quantities:

- the coupling matrix element V_{mn} ;
- the energy separation $|E_n^{(0)} - E_m^{(0)}|$.

A small matrix element can still produce a substantial admixture if the two unperturbed levels are very close. Conversely, a large energy separation suppresses mixing.

This is the first place where the denominator of perturbation theory acquires a direct physical meaning: the system is easier to mix with nearby levels than with remote ones.

2.5 Second-order energy correction

At second order,

$$(E_n^{(0)} - E_{n'}^{(0)}) C_{n'n}^{(2)} + E_n^{(1)} C_{n'n}^{(1)} + E_n^{(2)} \delta_{n'n} = \sum_m C_{mn}^{(1)} V_{n'm}.$$

Set $n' = n$. Because $C_{nn}^{(1)} = 0$,

$$E_n^{(2)} = \sum_m C_{mn}^{(1)} V_{nm}.$$

The term $m = n$ vanishes, so

$$E_n^{(2)} = \sum_{m \neq n} \frac{V_{mn} V_{nm}}{E_n^{(0)} - E_m^{(0)}}.$$

The perturbation operator is Hermitian, hence

$$V_{nm} = \overline{V_{mn}}.$$

Therefore

$$E_n^{(2)} = \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}.$$

Through second order,

$$E_n \simeq E_n^{(0)} + V_{nn} + \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}.$$

Sign of the ground-state correction

For the ground state,

$$E_m^{(0)} > E_0^{(0)} \quad (m \neq 0).$$

Every denominator in

$$E_0^{(2)} = \sum_{m \neq 0} \frac{|V_{m0}|^2}{E_0^{(0)} - E_m^{(0)}}$$

is therefore negative. Since the numerator is non-negative,

$$E_0^{(2)} < 0.$$

The second-order coupling to excited states always lowers the ground-state energy in non-degenerate perturbation theory. This negative second-order correction has direct physical consequences. A standard example is the van der Waals interaction between neutral atoms, for which the second-order correction gives an attractive potential $U(d) \propto -\frac{1}{d^6}$, where d is the center of mass separation. The same sign structure also appears in the effective attraction between electrons that leads to Cooper pairing in a superconductor.

2.6 When is non-degenerate perturbation theory valid?

The first-order admixture of the unperturbed state $|m^{(0)}\rangle$ into $|n^{(0)}\rangle$ is

$$\frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}}.$$

For perturbation theory to be controlled, this dimensionless ratio must remain small for the states that couple appreciably to $|n^{(0)}\rangle$:

$$\left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1.$$

Equivalently,

$$|V_{mn}| \ll |E_n^{(0)} - E_m^{(0)}|.$$

This criterion is more informative than saying only that “ $\hat{V}$ is small”. The perturbation is small **relative to the spectral separation of the states that it couples**.

The pairwise condition is a useful first check, but the first-order correction contains contributions from all coupled states. Its squared norm is

$$\|n^{(1)}\|^2 = \sum_{m \neq n} \left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right|^2.$$

A stronger global requirement for the state correction to remain small is therefore

$$\sum_{m \neq n} \left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right|^2 \ll 1.$$

Degenerate perturbation theory: general idea

Consider one particular r -fold degenerate eigenvalue $E^{(0)}$ of the unperturbed Hamiltonian. We follow how this one level changes when the perturbation is switched on.

The r states spanning its eigenspace D satisfy

$$\hat{H}_0 |i^{(0)}\rangle = E^{(0)} |i^{(0)}\rangle, \quad i = 1, \dots, r.$$

All other zeroth-order eigenstates are denoted by $|m^{(0)}\rangle$ and belong to different unperturbed energy levels:

$$\hat{H}_0 |m^{(0)}\rangle = E_m^{(0)} |m^{(0)}\rangle, \quad E_m^{(0)} \neq E^{(0)}.$$

Diagonalise the perturbation matrix restricted to D ,

$$V_{ij} = \langle i^{(0)} | \hat{V} | j^{(0)} \rangle.$$

Its eigenvectors define a new zeroth-order basis inside D :

$$|\alpha^{(0)}\rangle = \sum_{i=1}^r c_{i\alpha} |i^{(0)}\rangle, \quad \alpha = 1, \dots, r.$$

In this basis the perturbation operator is diagonal within D ,

$$\langle \beta^{(0)} | \hat{V} | \alpha^{(0)} \rangle = v_\alpha \delta_{\alpha\beta}.$$

The r branches originating from $E^{(0)}$ therefore begin as

$$E_\alpha = E^{(0)} + \lambda v_\alpha + \dots$$

For the branch that reduces to $|\alpha^{(0)}\rangle$ when $\lambda \rightarrow 0$, the first-order correction is expanded in the complete zeroth-order basis as

$$\sum_{\beta \neq \alpha} c_{\beta\alpha}^{(1)} |\beta^{(0)}\rangle + \sum_m c_{m\alpha}^{(1)} |m^{(0)}\rangle$$

The first sum lies inside D; the second contains states from the other unperturbed energy levels. The outside coefficients follow directly from the first-order equation:

$$c_{m\alpha}^{(1)} = \frac{\langle m^{(0)} | \hat{V} | \alpha^{(0)} \rangle}{E^{(0)} - E_m^{(0)}}.$$

The coefficients multiplying the other states $|\beta^{(0)}\rangle$ inside D are not fixed at first order; they are determined from the next-order equation.

If the first-order shifts v_α are distinct, the second-order energy correction is

$$E_\alpha^{(2)} = \sum_m \frac{(|\langle m^{(0)} | \hat{V} | \alpha^{(0)} \rangle|)^2}{E^{(0)} - E_m^{(0)}}.$$

What exactly is being expanded in λ ?

The parameter λ introduced in Section 2.1 is an **auxiliary switching parameter**. It defines the one-parameter family

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda \hat{V},$$

with the reference problem at $\lambda = 0$ and the physical Hamiltonian at $\lambda = 1$.

Rayleigh-Schrödinger perturbation theory assumes that the particular eigenvalue and eigenstate connected to $E_n^{(0)}$ and $|n^{(0)}\rangle$ can be followed smoothly from $\lambda = 0$ and expanded as a Taylor series there.

For the present non-degenerate discrete problem, this is the working assumption behind the order-by-order derivation. The important point is that λ is **not** automatically the physical coupling constant of the original problem. Statements about a physical parameter being non-analytic must therefore be kept separate from this formal construction.

The pure quartic oscillator in Section 3 illustrates that distinction. There the parameter α is a physical coefficient in the Hamiltonian. Taking $\alpha \rightarrow 0$ removes the confining potential and changes the spectral problem qualitatively: the bound discrete spectrum disappears and one obtains a free particle. The resulting scaling $E \propto \alpha^{1/3}$ is therefore a statement about the physical parameter α , not an example of the formal parameter λ suddenly acquiring a fractional power.

2.7 Example: quartic correction to the harmonic oscillator

Consider

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \alpha x^4.$$

Choose

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2, \quad \hat{V} = \alpha x^4.$$

The unperturbed energies are

$$E_n^{(0)} = \hbar\omega \left(n + \frac{1}{2} \right).$$

The first-order correction is

$$E_n^{(1)} = \alpha \langle n | x^4 | n \rangle.$$

We now derive the required matrix element rather than quoting it.

Introduce the standard oscillator length

$$x_0 = \sqrt{\frac{\hbar}{2m\omega}},$$

so that

$$x = x_0(a + a^\dagger).$$

Instead of expanding $(a + a^\dagger)^4$ directly, first act with x^2 on $|n\rangle$. Using

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle,$$

one finds

$$x^2|n\rangle = x_0^2 \left[\sqrt{n(n-1)}|n-2\rangle + (2n+1)|n\rangle + \sqrt{(n+1)(n+2)}|n+2\rangle \right].$$

Because x^2 is Hermitian,

$$\langle n|x^4|n\rangle = \|x^2|n\rangle\|^2.$$

The three states on the right-hand side are mutually orthogonal, so

$$\langle n|x^4|n\rangle = x_0^4 [n(n-1) + (2n+1)^2 + (n+1)(n+2)].$$

Simplifying,

$$\boxed{\langle n|x^4|n\rangle = 6x_0^4 \left(n^2 + n + \frac{1}{2} \right) = 6 \left(\frac{\hbar}{2m\omega} \right)^2 \left(n^2 + n + \frac{1}{2} \right)}.$$

Therefore

$$\boxed{E_n^{(1)} = 6\alpha \left(\frac{\hbar}{2m\omega} \right)^2 \left(n^2 + n + \frac{1}{2} \right)}.$$

For large n , the correction grows as n^2 , whereas the unperturbed level spacing remains $\hbar\omega$. The approximation therefore deteriorates for sufficiently highly excited states even if α is fixed and small.

This example illustrates a general lesson: the validity of perturbation theory depends on the state under consideration, not only on the coupling constant in the Hamiltonian.

3. When there is no obvious small parameter

3.1 The pure quartic oscillator: first scale the problem

Consider

$$\hat{H} = \frac{\hat{p}^2}{2m} + \alpha x^4.$$

There is no quadratic term, so there is no obvious harmonic-oscillator Hamiltonian to identify as the zeroth-order problem.

Before introducing an approximation, rescale the coordinate. Let

$$x = ay.$$

Then

$$\frac{d}{dx} = \frac{1}{a} \frac{d}{dy}, \quad \frac{d^2}{dx^2} = \frac{1}{a^2} \frac{d^2}{dy^2}.$$

The stationary Schrödinger equation becomes

$$\left[-\frac{\hbar^2}{2ma^2} \frac{d^2}{dy^2} + \alpha a^4 y^4 \right] \psi = E \psi.$$

Multiply by $2ma^2/\hbar^2$:

$$\left[-\frac{d^2}{dy^2} + \frac{2m\alpha a^6}{\hbar^2} y^4 \right] \psi = \frac{2ma^2 E}{\hbar^2} \psi.$$

Choose a so that the coefficient of y^4 is unity:

$$\frac{2m\alpha a^6}{\hbar^2} = 1.$$

Therefore

$$a = \left(\frac{\hbar^2}{2m\alpha} \right)^{1/6}.$$

The equation becomes parameter-free:

$$\left[-\frac{d^2}{dy^2} + y^4 \right] \psi = \varepsilon \psi,$$

where

$$\varepsilon = \frac{2ma^2 E}{\hbar^2}.$$

Hence

$$E = \frac{\hbar^2}{2ma^2} \varepsilon \propto \left(\frac{\alpha \hbar^4}{m^2} \right)^{1/3}.$$

The important result is the scaling

$$E \propto \alpha^{1/3}.$$

This dependence is non-analytic at $\alpha = 0$. It cannot be represented by a regular perturbation series

$$E = E^{(0)} + \alpha E^{(1)} + \alpha^2 E^{(2)} + \dots.$$

There is also a direct physical reason. At $\alpha = 0$ the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m},$$

so the confining potential disappears. The discrete bound spectrum for $\alpha > 0$ is therefore not a small deformation of a discrete bound spectrum at $\alpha = 0$; the reference problem itself changes character.

This is different from the auxiliary parameter λ in Section 2. There, λ was introduced only to organise the expansion of a fixed decomposition $\hat{H}_0 + \hat{V}$. Here α is a physical parameter whose zero limit is singular for the bound-state problem.

3.2 Introduce a solvable reference problem by adding and subtracting the same term

A useful reference Hamiltonian can nevertheless be introduced artificially. Add and subtract a harmonic term:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \alpha x^4 + \frac{1}{2}m\omega^2 x^2 - \frac{1}{2}m\omega^2 x^2.$$

Define

$$\hat{H}_0(\omega) = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2,$$

and

$$\hat{V}(\omega) = \alpha x^4 - \frac{1}{2}m\omega^2 x^2.$$

The original Hamiltonian has not been changed. We have only reorganised it.

This algebraic rewriting does **not** by itself make $\hat{V}(\omega)$ a small perturbation. There is no theorem saying that the residual interaction is weak for an arbitrary choice of ω . The purpose of the auxiliary parameter is instead to choose a reference problem that resembles the physical problem as closely as possible within a simple solvable family. The quality of the approximation must then be judged by the optimisation procedure and, where possible, by comparison with additional information.

The frequency ω is not a physical parameter of the original problem. It is an auxiliary parameter that controls the choice of reference system.

For the n th oscillator state,

$$E_n^{(1)} = \alpha \langle x^4 \rangle - \frac{1}{2}m\omega^2 \langle x^2 \rangle.$$

Using

$$\langle x^2 \rangle = \frac{\hbar}{m\omega} \left(n + \frac{1}{2} \right)$$

and the result derived in Section 2.7,

$$E_n^{(1)} = 6\alpha \left(\frac{\hbar}{2m\omega} \right)^2 \left(n^2 + n + \frac{1}{2} \right) - \frac{\hbar\omega}{2} \left(n + \frac{1}{2} \right).$$

One possible prescription is to choose ω so that

$$E_n^{(1)} = 0.$$

Then the first correction vanishes and the approximate energy is simply the reference oscillator energy evaluated at the optimised frequency.

This is an example of a broader strategy: when there is no obvious small parameter in the original problem, reorganise the Hamiltonian around a solvable reference family and choose its auxiliary parameters so that the first omitted corrections are reduced. This is an optimisation strategy rather than an ordinary perturbation expansion controlled by a known small physical parameter.

3.3 Optimisation by minimal energy

A more systematic choice is to minimise the approximate energy with respect to the auxiliary parameter ω .

For the pure quartic oscillator, the first-order estimate is

$$E_n(\omega) = \frac{\hbar\omega}{2} \left(n + \frac{1}{2} \right) + 6\alpha \left(\frac{\hbar}{2m\omega} \right)^2 \left(n^2 + n + \frac{1}{2} \right).$$

For the ground state,

$$E_0(\omega) = \frac{\hbar\omega}{4} + 3\alpha \left(\frac{\hbar}{2m\omega} \right)^2.$$

The stationary condition is

$$\frac{dE_0}{d\omega} = 0.$$

Differentiating,

$$\frac{\hbar}{4} - \frac{3\alpha\hbar^2}{2m^2\omega^3} = 0,$$

which gives

$$\omega_{\text{opt}} = \left(\frac{6\alpha\hbar}{m^2} \right)^{1/3}.$$

Substitution gives

$$E_0^{\text{opt}} = \frac{3}{8} \left(\frac{6\alpha\hbar^4}{m^2} \right)^{1/3}.$$

The scaling $E_0 \propto \alpha^{1/3}$ is now correct automatically.

The close relation to the variational method is not accidental. For the ground state, choosing a harmonic reference Hamiltonian and minimising the approximate energy with respect to ω is exactly equivalent to choosing the harmonic-oscillator ground state as a Gaussian trial family and minimising $\langle H \rangle$ with respect to its width. In Section 5 we will prove the variational theorem, which then tells us something stronger: this optimised value is a rigorous **upper bound** to the exact ground-state energy.

4. Variational method

4.1 Variational formulation of the stationary Schrödinger equation

Consider a quantum system that is characterized by a Hamiltonian $\hat{H}$. Let $\psi(q)$ be a normalised trial function,

$$\int dq |\psi(q)|^2 = 1.$$

Define the energy functional

$$\mathcal{E}[\psi] = \langle \psi | \hat{H} | \psi \rangle.$$

We seek stationary values of this functional subject to the normalisation constraint. Introduce a Lagrange multiplier E and define

$$\mathcal{F}[\psi, \bar{\psi}] = \langle \psi | \hat{H} | \psi \rangle - E(\langle \psi | \psi \rangle - 1).$$

Here, the overbar denotes complex conjugate. In the variational calculation, ψ and $\bar{\psi}$ are treated as independent variables. This is a standard mathematical device for varying a complex field; the final physical solution still has $\bar{\psi}$ equal to the complex conjugate of ψ .

The variation is

$$\delta\mathcal{F} = \int dq \delta\bar{\psi}(\hat{H}\psi - E\psi) + \int dq \delta\psi \overline{(\hat{H}\psi - E\psi)}.$$

Because $\delta\psi$ and $\delta\bar{\psi}$ are arbitrary,

$$\boxed{\hat{H}\psi = E\psi.}$$

The Lagrange multiplier is therefore the energy eigenvalue.

The precise statement is that exact energy eigenstates are **stationary points** of the energy functional under the normalisation constraint. The ground state is the absolute minimum. Excited eigenstates are stationary as well, but they need not be minima when arbitrary variations are allowed.

For practical ground-state calculations we rarely vary over all possible functions. We choose a restricted family of physically motivated trial functions and minimise the energy with respect to a small number of parameters.

4.2 Upper-bound theorem for the ground-state energy

Let the exact energy eigenstates be $|n\rangle$,

$$\hat{H}|n\rangle = E_n|n\rangle,$$

ordered so that

$$E_0 \leq E_1 \leq E_2 \leq \dots$$

Expand any normalised trial state as

$$|\psi\rangle = \sum_n c_n |n\rangle, \quad \sum_n |c_n|^2 = 1.$$

Its energy expectation value is

$$\langle H \rangle_\psi = \sum_n |c_n|^2 E_n.$$

Write

$$E_n = E_0 + (E_n - E_0).$$

Then

$$\langle H \rangle_\psi = E_0 \sum_n |c_n|^2 + \sum_n |c_n|^2 (E_n - E_0).$$

Using normalisation,

$$\langle H \rangle_\psi = E_0 + \sum_n |c_n|^2 (E_n - E_0).$$

Every term in the second sum is non-negative, so

$$\boxed{\langle H \rangle_\psi \geq E_0.}$$

This is the central variational theorem.

There is a second useful consequence. Suppose a normalised trial state is close to the exact ground state and write

$$|\psi\rangle = \sqrt{1 - \varepsilon^2} |0\rangle + \sum_{n>0} c_n |n\rangle, \quad \varepsilon^2 = \sum_{n>0} |c_n|^2 \ll 1.$$

The wavefunction error is of order ε , but the energy error is

$$\langle H \rangle_\psi - E_0 = \sum_{n>0} |c_n|^2 (E_n - E_0) = O(\varepsilon^2).$$

Thus the energy is **stationary** at the exact ground state: small first-order errors in the wavefunction produce only second-order errors in the variational energy. This is one reason a fairly simple trial function can sometimes give a surprisingly accurate ground-state energy.

It gives a clear practical prescription:

1. choose a normalised trial family $\psi(q; a_1, a_2, \dots)$;
2. calculate

$$E(a_1, a_2, \dots) = \langle \psi | \hat{H} | \psi \rangle;$$

3. minimise with respect to the free parameters;
4. the minimum is an upper bound to the exact ground-state energy.

The theorem also explains why physical insight matters. A trial function that respects the symmetry, boundary conditions, nodal structure and asymptotic decay of the exact state usually has a much larger overlap with the true ground state and gives a better bound.

4.3 Example: pure quartic oscillator with a Gaussian trial state

For

$$\hat{H} = \frac{\hat{p}^2}{2m} + \alpha x^4,$$

the ground state is even, smooth and has no nodes. A simple trial function with these properties is a Gaussian,

$$\psi_\omega(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right).$$

The parameter ω controls the width of the Gaussian. It is a variational parameter, not a physical oscillator frequency in the original Hamiltonian. This is the same Gaussian family that appeared in Section 3.3 when we optimised an auxiliary harmonic reference Hamiltonian. The variational theorem now tells us why that minimisation is especially useful for the ground state: its minimum is a rigorous upper bound to the exact energy.

Kinetic-energy expectation value

Differentiate the trial function:

$$\frac{d\psi}{dx} = -\frac{m\omega}{\hbar} x\psi,$$

and

$$\frac{d^2\psi}{dx^2} = \left[\left(\frac{m\omega}{\hbar}\right)^2 x^2 - \frac{m\omega}{\hbar}\right]\psi.$$

Since

$$\hat{p}^2 = -\hbar^2 \frac{d^2}{dx^2},$$

we obtain

$$\left\langle \frac{\hat{p}^2}{2m} \right\rangle = \frac{\hbar\omega}{2} - \frac{m\omega^2}{2} \langle x^2 \rangle.$$

The required Gaussian moments can be obtained without quoting a formula. Define

$$\beta = \frac{m\omega}{\hbar}$$

and

$$I_0(\beta) = \int_{-\infty}^{\infty} e^{-\beta x^2} dx = \sqrt{\frac{\pi}{\beta}}.$$

Differentiate with respect to β :

$$-\frac{dI_0}{d\beta} = \int_{-\infty}^{\infty} x^2 e^{-\beta x^2} dx.$$

After division by I_0 ,

$$\boxed{\langle x^2 \rangle = \frac{1}{2\beta} = \frac{\hbar}{2m\omega}}.$$

Hence

$$\boxed{\left\langle \frac{\hat{p}^2}{2m} \right\rangle = \frac{\hbar\omega}{4}}.$$

Quartic expectation value

Differentiate the Gaussian integral twice:

$$\frac{d^2 I_0}{d\beta^2} = \int_{-\infty}^{\infty} x^4 e^{-\beta x^2} dx.$$

Since

$$I_0(\beta) = \sqrt{\pi} \beta^{-1/2},$$

we have

$$\frac{d^2 I_0}{d\beta^2} = \frac{3\sqrt{\pi}}{4} \beta^{-5/2}.$$

Dividing by I_0 gives

$$\boxed{\langle x^4 \rangle = \frac{3}{4\beta^2} = 3 \left(\frac{\hbar}{2m\omega} \right)^2}.$$

The variational energy is

$$E(\omega) = \frac{\hbar\omega}{4} + 3\alpha \left(\frac{\hbar}{2m\omega} \right)^2.$$

Set

$$\frac{dE}{d\omega} = 0.$$

This gives

$$\omega_{\text{opt}} = \left(\frac{6\alpha\hbar}{m^2} \right)^{1/3},$$

and hence

$$E_0 \leq \frac{3}{8} \left(\frac{6\alpha\hbar^4}{m^2} \right)^{1/3}.$$

The inequality sign is essential. Unlike a generic approximation, the variational estimate comes with a rigorous direction of error for the ground-state energy.

4.4 Excited states

The variational idea can be extended to excited states, but the orthogonality constraints matter.

Suppose the exact ground state $|0\rangle$ were known. If we minimise $\langle H \rangle$ over all normalised states orthogonal to $|0\rangle$, the minimum is E_1 .

More generally, to obtain a rigorous upper bound to E_n , the trial state must be constrained to be orthogonal to the exact lower eigenstates $|0\rangle, \dots, |n-1\rangle$.

In practical calculations these exact lower states are usually not known. One often imposes orthogonality to previously obtained approximate states. This remains a useful computational strategy, but the rigorous upper-bound interpretation then requires additional care.

Symmetry can sometimes provide exact orthogonality automatically. For example, in a parity-symmetric one-dimensional potential the ground state is even and the first excited state is odd. Any odd trial state is then automatically orthogonal to the exact even ground state.

5. Time-dependent perturbation theory

5.1 Problem setup

When the Hamiltonian depends explicitly on time, energy is in general no longer conserved and the problem is no longer an eigenvalue problem. Instead, we ask how a state prepared at an initial time evolves and with what probability it is later found in one of the stationary states of a convenient reference Hamiltonian. We therefore separate the Hamiltonian into a time-independent part that is exactly solvable and a weak time-dependent perturbation:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t), \quad (5.1)$$

The stationary eigenvalue problem for the reference Hamiltonian is

$$\hat{H}_0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle, \quad (5.2)$$

In the absence of the perturbation, each of these states changes only by its dynamical phase,

$$|n^{(0)}; t\rangle = e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle, \quad (5.3)$$

Because the unperturbed eigenstates form a complete basis, the exact state of the perturbed system can be expanded in the same basis while allowing the expansion coefficients to vary with time:

$$|\Psi(t)\rangle = \sum_n C_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle, \quad (5.4)$$

This is the time-dependent analogue of variation of constants. The exponential factors contain the evolution that would occur under the reference Hamiltonian alone, while the coefficients $C_n(t)$ contain the change produced by the perturbation. For a normalised state, orthonormality of the unperturbed basis gives

$$\sum_n |C_n(t)|^2 = 1, \quad (5.5)$$

Thus, for a non-degenerate unperturbed level, $|C_n(t)|^2$ is the probability that a measurement of the unperturbed Hamiltonian at time t gives $E_n^{(0)}$. If an unperturbed energy is degenerate, the probability of that energy is obtained by summing $|C(t)|^2$ over all basis states in the corresponding eigenspace. The same decomposition also makes the interaction picture explicit. Removing the known evolution generated by the reference Hamiltonian gives

$$|\Psi_I(t)\rangle = e^{i\hat{H}_0 t/\hbar} |\Psi(t)\rangle, \quad (5.6)$$

The coefficients $C_n(t)$ are therefore precisely the components of the interaction-picture state in the unperturbed eigenbasis.

5.2 Exact equations for the amplitudes

The exact state satisfies the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \hat{H}(t) |\Psi(t)\rangle, \quad (5.7)$$

Substituting the expansion (5.4) into (5.7), the terms generated by the unperturbed Hamiltonian cancel against the time derivatives of the phase factors. Projecting the remaining equation onto an arbitrary unperturbed state $\langle m^{(0)}|$ gives an exact system of coupled equations for the amplitudes:

$$i\hbar \dot{C}_m(t) = \sum_n \tilde{V}_{mn}(t) C_n(t), \quad (5.8)$$

Here the Schrödinger-picture matrix elements of the perturbation are

$$V_{mn}(t) = \langle m^{(0)} | \hat{V}(t) | n^{(0)} \rangle, \quad (5.9)$$

and the transition frequencies associated with the unperturbed spectrum are

$$\omega_{mn} = \frac{E_m^{(0)} - E_n^{(0)}}{\hbar}, \quad (5.10)$$

The corresponding matrix elements in the interaction picture are therefore

$$\tilde{V}_{mn}(t) = e^{i\omega_{mn}t} V_{mn}(t), \quad (5.11)$$

Equivalently, at the operator level,

$$\tilde{V}(t) = e^{i\hat{H}_0 t/\hbar} \hat{V}(t) e^{-i\hat{H}_0 t/\hbar}, \quad (5.12)$$

Equation (5.8) is still exact. No perturbative approximation has yet been made. The approximation enters only when the amplitudes are expanded order by order in the weak interaction. This formulation is useful because the known evolution generated by the reference Hamiltonian has already been removed; the remaining time dependence of the amplitudes is generated by the interaction-picture perturbation.

5.3 First-order transition amplitude

Suppose that the perturbation acts during the interval from t_0 to t_1 and vanishes outside it. Immediately before it is switched on, let the system be in the unperturbed state i . The initial condition for the interaction-picture amplitudes is then

$$C_n(t_0) = \delta_{ni}, \quad (5.13)$$

We now expand each amplitude in successive orders of the perturbation. As in stationary perturbation theory, an explicit switching parameter may be introduced to organise the powers, but it is convenient to suppress it in the notation:

$$C_n(t) = C_n^{(0)}(t) + C_n^{(1)}(t) + C_n^{(2)}(t) + \dots, \quad (5.14)$$

The initial conditions imply

$$C_n^{(0)}(t) = \delta_{ni}, \quad C_n^{(v)}(t_0) = 0 \quad (v \geq 1), \quad (5.15)$$

Substitution into the exact amplitude equation (5.8) and collection of the terms linear in the perturbation gives

$$i\hbar \dot{C}_m^{(1)}(t) = \tilde{V}_{mi}(t), \quad (5.16)$$

Integrating from the initial time to an arbitrary later time gives

$$C_m^{(1)}(t) = \frac{1}{i\hbar} \int_{t_0}^t \tilde{V}_{mi}(t') dt', \quad (5.17)$$

In particular, the first-order amplitude for a transition from the initial state i to a different final state f during the interval $[t_0, t_1]$ is

$$C_f^{(1)}(t_1) = \frac{1}{i\hbar} \int_{t_0}^{t_1} \tilde{V}_{fi}(t') dt', \quad f \neq i, \quad (5.18)$$

Using (5.11), the same result can be written directly in terms of the Schrödinger-picture matrix element of the perturbation:

$$C_f^{(1)}(t_1) = \frac{1}{i\hbar} \int_{t_0}^{t_1} V_{fi}(t') e^{i\omega_{fi}t'} dt', \quad (5.19)$$

The oscillating factor $\exp(i\omega_{fi}t)$ is essential. Contributions generated at different times are added as complex amplitudes; depending on their relative phases they may reinforce one another or cancel. The first-order amplitude is linear in the perturbation, whereas the leading non-vanishing transition probability is quadratic:

$$P_{i \rightarrow f} = |C_f^{(1)}(t_1)|^2, \quad f \neq i, \quad (5.20)$$

This distinction between the order of the amplitude and the order of the probability is important when perturbative results are compared with exact normalisation.

5.4 Higher orders, time ordering and the evolution operator

At second order the amplitude equation becomes

$$i\hbar \dot{C}_m^{(2)}(t) = \sum_n \tilde{V}_{mn}(t) C_n^{(1)}(t), \quad (5.21)$$

Inserting the first-order solution (5.17) and integrating once more gives, for the amplitude to reach a final state f ,

$$C_f^{(2)}(t_1) = \left(\frac{1}{i\hbar}\right)^2 \sum_n \int_{t_0}^{t_1} dt' \int_{t_0}^{t'} dt'' \tilde{V}_{fn}(t') \tilde{V}_{ni}(t''). \quad (5.22)$$

The nested integration limits express time ordering. The interaction at time t'' occurs first and can take the system from the initial state i to an intermediate state n ; the second interaction occurs at the later time t' and takes the system from n to the final state f . Higher orders repeat the same structure with successively ordered interaction times.

It is useful to collect the complete sequence of amplitudes into a finite-time evolution operator. In the interaction picture we denote this operator by $\hat{U}_I(t, t_0)$, defined by

$$|\Psi_I(t)\rangle = \hat{U}_I(t, t_0)|\Psi_I(t_0)\rangle, \quad (5.23)$$

Its equation of motion follows directly from the interaction-picture Schrödinger equation:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \hat{U}_I(t, t_0) &= \tilde{V}(t) \hat{U}_I(t, t_0), \\ \hat{U}_I(t_0, t_0) &= \hat{1}. \end{aligned} \quad (5.24)$$

Iterating this equation generates the ordered perturbation series

$$\hat{U}_I(t_1, t_0) = 1 + \frac{1}{i\hbar} \int_{t_0}^{t_1} dt \tilde{V}(t) + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t_1} dt' \int_{t_0}^{t'} dt'' \tilde{V}(t') \tilde{V}(t'') + \dots \quad (5.25)$$

The first two non-trivial terms in (5.25) reproduce the first- and second-order amplitudes derived above. For an initial unperturbed state i , the finite-time amplitude to be in state f at t_1 is simply the corresponding matrix element of the evolution operator:

$$C_f(t_1) = \langle f^{(0)} | \hat{U}_I(t_1, t_0) | i^{(0)} \rangle, \quad (5.26)$$

For a finite interval the natural object is therefore the evolution operator $\hat{U}_I(t_1, t_0)$. In a scattering problem one instead considers asymptotically free states before and after the interaction. The asymptotic limit of the interaction-picture evolution operator defines the scattering operator, or S -matrix:

$$\hat{S} \equiv \hat{U}_I(+\infty, -\infty), \quad S_{fi} = \langle f^{(0)} | \hat{S} | i^{(0)} \rangle. \quad (5.27)$$

The notation S comes from scattering. Its matrix elements are the transition amplitudes between asymptotic initial and final states. Thus, the same ordered perturbation series describes both finite-time quantum transitions and scattering; $\hat{U}_I$ refers to finite-time evolution, while $\hat{S}$ is reserved for the asymptotic scattering operator.

6. Transition rates and Fermi's golden rule

Section 5 gives the first-order transition amplitude from an initial stationary state to a specified final stationary state. For one isolated final level the natural quantity is the finite-time transition probability. A time-independent rate emerges only when many closely spaced final states are available: the finite interaction time selects a narrow energy interval, and the probabilities for the states inside that interval add to give a total probability proportional to the duration of the interaction.

We keep the unperturbed notation explicit: $E_i^{(0)}$ and $E_f^{(0)}$ are eigenvalues of the unperturbed Hamiltonian, and $|i^{(0)}\rangle$ and $|f^{(0)}\rangle$ are the corresponding stationary states.

6.1 A constant perturbation acting for a finite time

Suppose the system is in $|i^{(0)}\rangle$ at time t_0 . The perturbation is constant while it acts and vanishes outside the interval $[t_0, t_1]$:

$$\hat{V}(t) = \begin{cases} \hat{V}, & t_0 \leq t \leq t_1, \\ 0, & t < t_0 \text{ or } t > t_1, \end{cases} \quad (6.1)$$

The duration of the interaction is

$$\tau = t_1 - t_0. \quad (6.2)$$

For a specified final state $|f^{(0)}\rangle$, define the transition frequency and the matrix element of the perturbation by

$$\omega_{fi} = \frac{E_f^{(0)} - E_i^{(0)}}{\hbar}, \quad V_{fi} = \langle f^{(0)} | \hat{V} | i^{(0)} \rangle. \quad (6.3)$$

The first-order transition amplitude is

$$C_f^{(1)}(t_1) = \frac{V_{fi}}{i\hbar} \int_{t_0}^{t_1} dt' e^{i\omega_{fi}t'}, \quad f \neq i. \quad (6.4)$$

With $t_c = (t_0 + t_1)/2$, evaluation of the integral separates the physically irrelevant overall phase from the magnitude of the amplitude:

$$C_f^{(1)}(t_1) = -\frac{2iV_{fi}}{\hbar\omega_{fi}} e^{i\omega_{fi}t_c} \sin\left(\frac{\omega_{fi}\tau}{2}\right). \quad (6.5)$$

The exponential in Eq. (6.5) has unit modulus. The leading non-vanishing transition probability to this one final state is therefore

$$P_{i \rightarrow f}(\tau) = \frac{4|V_{fi}|^2}{\hbar^2} \frac{\sin^2(\omega_{fi}\tau/2)}{\omega_{fi}^2}. \quad (6.6)$$

At exact resonance, $\omega_{fi} = 0$, the limit of Eq. (6.6) is

$$P_{i \rightarrow f}(\tau) = \frac{|V_{fi}|^2 \tau^2}{\hbar^2}, \quad \omega_{fi} = 0. \quad (6.7)$$

The quadratic growth in Eq. (6.7) cannot continue indefinitely: once the transition probability is no longer small, first-order perturbation theory is no longer self-consistent. Away from resonance, Eq. (6.6) oscillates. A single isolated final level therefore does not give a constant long-time transition rate.

The finite interaction time also determines the energy resolution. Since $\Delta E_{fi} = E_f^{(0)} - E_i^{(0)} = \hbar\omega_{fi}$, the first zeros on either side of the central peak correspond to

$$|\Delta E_{fi}| = \frac{2\pi\hbar}{\tau}. \quad (6.8)$$

so the characteristic width of the energy-selection region scales as

$$\Delta E_{\text{res}} \sim \frac{\hbar}{\tau}. \quad (6.9)$$

The numerical factor depends on how the width is defined; the essential result is the inverse dependence on τ .

6.2 From many discrete final states to a continuum

Now suppose that many mutually orthogonal final states are available near the initial energy.

Imagine first a very large but finite system: the states are still discrete, but their level spacing is very small. Because the final states are orthogonal, their transition probabilities add.

Introduce the density of final states by

$$dN_f = \rho_f(E_f) dE_f. \quad (6.10)$$

where dN_f is the number of accessible final states in the interval dE_f . Additional quantum numbers can be included in ρ_f or summed explicitly. It is useful to define the finite-time energy-selection kernel

$$K_\tau(\varepsilon) = 4 \frac{\sin^2(\varepsilon\tau/2\hbar)}{\varepsilon^2}. \quad (6.11)$$

From here on, E denotes the final state energy. Define the slowly varying factor

$$F(E) = |V_{fi}(E)|^2 \rho_f(E). \quad (6.12)$$

The total probability transferred from the initial state into the accessible continuum is then

$$P_{\text{cont}}(\tau) = \int dE F(E) K_\tau(E - E_i^{(0)}). \quad (6.13)$$

This form separates the physics cleanly: $F(E)$ contains the transition strength and the number of available final states, while K_τ determines which final energies contribute during the finite interaction time.

6.3 Long-time energy selection and Fermi's golden rule

As τ increases, $K_\tau(\varepsilon)$ becomes higher and narrower around $\varepsilon=0$. Its height grows as τ^2 and its width decreases as $1/\tau$. The area under the peak is

$$\int_{-\infty}^{\infty} K_\tau(\varepsilon) d\varepsilon = \frac{2\pi\tau}{\hbar}. \quad (6.14)$$

Hence, under an energy integral, the family K_τ approaches a delta sequence:

$$K_\tau(\varepsilon) \rightarrow \frac{2\pi\tau}{\hbar} \delta(\varepsilon) \quad (\tau \rightarrow \infty). \quad (6.15)$$

This is a distributional statement, not a pointwise limit. If $F(E)$ changes little across the narrow interval of width of order $\hbar/\tau$ around $E_i^{(0)}$, it may be replaced by $F(E_i^{(0)})$ inside Eq. (6.13). One then obtains

$$P_{\text{cont}}(\tau) \simeq \frac{2\pi\tau}{\hbar} F(E_i^{(0)}). \quad (6.16)$$

The total probability is now proportional to τ . Dividing by the duration of the interaction gives the constant transition rate

$$w = \frac{P_{\text{cont}}(\tau)}{\tau} = \frac{2\pi}{\hbar} F(E_i^{(0)}). \quad (6.17)$$

This is Fermi's golden rule. Since

$$F(E_i^{(0)}) = |V_{fi}(E_i^{(0)})|^2 \rho_f(E_i^{(0)}),$$

the rule combines the strength of the perturbation with the number of final states available at the selected energy. Its dimensions are those of inverse time.

6.4 Perturbative order and normalisation

The transition amplitude $C_f^{(1)}$ is first order in the perturbation, whereas the transition probability $|C_f^{(1)}|^2$ is second order. The exact amplitudes satisfy

$$\sum_n |C_n(t)|^2 = 1. \quad (6.18)$$

For a system initially in state i , write

$$C_i = 1 + C_i^{(1)} + C_i^{(2)} + \dots \quad (6.19)$$

$$C_{f \neq i} = C_f^{(1)} + C_f^{(2)} + \dots \quad (6.20)$$

At first order normalisation requires $\text{Re } C_i^{(1)} = 0$; a nonzero diagonal matrix element V_{ii} therefore changes only the phase of the initial amplitude at this order. At second order, normalisation gives

$$2 \operatorname{Re} C_i^{(2)} + |C_i^{(1)}|^2 + \sum_{f \neq i} |C_f^{(1)}|^2 = 0. \quad (6.21)$$

Thus, the probability accumulated in states $f \neq i$ is accompanied by the corresponding decrease of the survival probability. The golden-rule rate itself is obtained from the first-order amplitudes to $f \neq i$, so $C_i^{(2)}$ is not needed explicitly for the rate, but it is needed for a normalisation check through second order.

6.5 Connection with scattering

In a scattering problem the continuum of final states is the continuum of outgoing momentum states. The golden rule gives the transition rate into an element of final-state phase space. Dividing that rate by the incident particle flux gives the differential scattering cross section. This is the step that connects time-dependent perturbation theory with the scattering formulas used in the next application.

7. Elastic neutron scattering and the structure factor

Section 6 showed how Fermi's golden rule gives a transition rate into a continuum of final states and how, in a scattering problem, dividing that rate by the incident particle flux gives a differential cross section. We now apply that result to elastic neutron scattering from a condensed medium.

The medium contains N identical nuclei at positions $\mathbf{R}_1, \dots, \mathbf{R}_N$. During one scattering event we first treat these positions as fixed. The neutron is described before and after the collision by free plane waves. We then calculate the cross section for this fixed arrangement of nuclei and, only afterwards, average over the configurations sampled by the medium.

Two approximations are being used. First, in the **first Born approximation** the transition matrix element is evaluated between the unperturbed incident and outgoing plane waves. Second, in the **static approximation** the medium does not change its state during the scattering event, so no energy is transferred to it. The neutron therefore changes direction but not kinetic energy.

7.1 Initial and final neutron states

Let V be a large normalisation volume. The initial and final neutron wavefunctions are

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad \psi_f(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}' \cdot \mathbf{r}} \quad (7.1)$$

For a free neutron of mass m ,

$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}, \quad E_{\mathbf{k}'} = \frac{\hbar^2 k'^2}{2m} \quad (7.2)$$

In the static approximation the medium has the same energy before and after the event. Energy conservation therefore requires

$$E_{\mathbf{k}'} = E_{\mathbf{k}}, \quad k' = k \quad (7.3)$$

For a fixed set of nuclear positions $\{\mathbf{R}_j\}$, the neutron-medium interaction is taken to be the sum of the neutron-nucleus potentials,

$$\hat{V}(\mathbf{r}) = \sum_{j=1}^N \Phi(|\mathbf{r} - \mathbf{R}_j|) \quad (7.4)$$

where $\Phi(R)$ is the potential produced by one nucleus. Define the transferred wavevector

$$\mathbf{q} = \mathbf{k}' - \mathbf{k} \quad (7.5)$$

The neutron momentum therefore changes by $\hbar\mathbf{q}$. In elastic scattering $k' = k$, but $\mathbf{q}$ is generally non-zero because the direction of the neutron changes.

7.2 Transition matrix element for a fixed nuclear configuration

In the first Born approximation the matrix element entering Fermi's golden rule is

$$\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \frac{1}{V} \sum_{j=1}^N \int d^3r e^{-i\mathbf{q}\cdot\mathbf{r}} \Phi(|\mathbf{r} - \mathbf{R}_j|) \quad (7.6)$$

For the j th term introduce $\mathbf{R} = \mathbf{r} - \mathbf{R}_j$. Then $\mathbf{r} = \mathbf{R} + \mathbf{R}_j$, and

$$e^{-i\mathbf{k}'\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} = e^{-i\mathbf{q}\cdot\mathbf{R}_j} e^{-i\mathbf{q}\cdot\mathbf{R}}.$$

The dependence on the position $\mathbf{R}_j$ is therefore separated from the integral over the potential. Define its Fourier transform by

$$v_{\mathbf{q}} = \int d^3R e^{-i\mathbf{q}\cdot\mathbf{R}} \Phi(R) \quad (7.7)$$

For a central potential, $v_{\mathbf{q}}$ depends only on $q = |\mathbf{q}|$. Equation (7.6) becomes

$$\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \frac{v_{\mathbf{q}}}{V} \sum_{j=1}^N e^{-i\mathbf{q}\cdot\mathbf{R}_j} = \frac{\sqrt{N}}{V} v_{\mathbf{q}} \rho_{\mathbf{q}} \quad (7.8)$$

where

$$\rho_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{-i\mathbf{q}\cdot\mathbf{R}_j} \quad (7.9)$$

The meaning of $\rho_{\mathbf{q}}$ follows directly from the microscopic number density

$$n(\mathbf{r}) = \sum_{j=1}^N \delta(\mathbf{r} - \mathbf{R}_j).$$

Its Fourier component is

$$n_{\mathbf{q}} = \int d^3r e^{-i\mathbf{q}\cdot\mathbf{r}} n(\mathbf{r}) = \sum_{j=1}^N e^{-i\mathbf{q}\cdot\mathbf{R}_j} = \sqrt{N} \rho_{\mathbf{q}}.$$

For a homogeneous medium, the average density is spatially uniform and contributes only at $\mathbf{q} = 0$. Hence, for $\mathbf{q} \neq 0$, $\rho_{\mathbf{q}}$ is the normalised Fourier component of the fluctuation of the microscopic density about its uniform mean value. The case $\mathbf{q} = 0$ corresponds to the uniform density mode and is treated separately from the non-zero- q structure discussed below.

The transition probability contains the modulus squared of the matrix element. The part that depends on the nuclear positions is therefore

$$|\rho_{\mathbf{q}}|^2 = \frac{1}{N} \sum_{j,l=1}^N e^{-i\mathbf{q}\cdot(\mathbf{R}_j - \mathbf{R}_l)}.$$

Separating the terms with $j = l$ from those with $j \neq l$ gives

$$|\rho_{\mathbf{q}}|^2 = 1 + \frac{1}{N} \sum_{j \neq l} e^{-i\mathbf{q}\cdot(\mathbf{R}_j - \mathbf{R}_l)} \quad (7.10)$$

The first term in Eq. (7.10) comes from the N terms with $j = l$: for each of them $\mathbf{R}_j - \mathbf{R}_l = 0$, so the exponential is equal to one. The remaining terms depend on the separation vectors between different nuclei. Consequently, $|\rho_{\mathbf{q}}|^2$ changes when the spatial arrangement of the nuclei changes. This is the point at which information about the structure of the medium enters the scattering problem.

7.3 From Fermi's golden rule to the differential cross section

We now use the golden-rule result of Section 6. In a large normalisation volume, the allowed outgoing wavevectors are closely spaced. With periodic boundary conditions, one state occupies a volume $(2\pi)^3/V$ in $\mathbf{k}'$ -space. The number of final neutron states in $d^3k' = k'^2 dk' d\Omega$ is therefore

$$dN_f = \frac{V}{(2\pi)^3} k'^2 dk' d\Omega \quad (7.11)$$

Fermi's golden rule gives the transition rate to each final state. Therefore the rate into the set of final states counted by dN_f is

$$dw = \frac{2\pi}{\hbar} |\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) dN_f \quad (7.12)$$

The delta function imposes elastic energy conservation. Since $E_{\mathbf{k}'} = \hbar^2 k'^2 / 2m$,

$$\delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) = \frac{m}{\hbar^2 k} \delta(k' - k) \quad (7.13)$$

where only the positive root $k' = k$ is relevant. Substituting Eq. (7.11) into Eq. (7.12), integrating over k' , and using Eq. (7.8), the scattering rate per unit solid angle is

$$\frac{dw}{d\Omega} = \frac{Nm k}{4\pi^2 V \hbar^3} |v_q|^2 |\rho_{\mathbf{q}}|^2 \quad (7.14)$$

The incident plane wave contains one neutron in the normalisation volume V , so its particle density is $1/V$. Its speed is $v = \hbar k / m$. The incident particle flux is therefore

$$j_{\text{in}} = \frac{1}{V} \frac{\hbar k}{m} \quad (7.15)$$

By definition, the differential cross section is the scattering rate into $d\Omega$, divided by the incident flux and by $d\Omega$:

$$\frac{d\sigma}{d\Omega} = \frac{1}{j_{\text{in}}} \frac{dw}{d\Omega} \quad (7.16)$$

Substituting Eqs. (7.14) and (7.15), the normalisation volume cancels and we obtain, for the fixed nuclear configuration,

$$\left(\frac{d\sigma}{d\Omega} \right)_{\{\mathbf{R}_j\}} = N \left(\frac{mv_q}{2\pi\hbar^2} \right)^2 |\rho_{\mathbf{q}}|^2 \quad (7.17)$$

For one nucleus, $N = 1$ and $|\rho_{\mathbf{q}}|^2 = 1$. The Born cross section for one nucleus is therefore

$$\left(\frac{d\sigma}{d\Omega} \right)_1 = \left(\frac{mv_q}{2\pi\hbar^2} \right)^2 \quad (7.18)$$

Thus Eq. (7.17) can be written as

$$\left(\frac{d\sigma}{d\Omega} \right)_{\{\mathbf{R}_j\}} = N \left(\frac{d\sigma}{d\Omega} \right)_1 |\rho_{\mathbf{q}}|^2.$$

Everything in this expression has now been defined. The factor $(d\sigma/d\Omega)_1$ is the cross section of one nucleus in the Born approximation, while $|\rho_{\mathbf{q}}|^2$ contains the dependence on the positions of all nuclei in the particular configuration being considered.

7.4 Equilibrium average and the static structure factor

A macroscopic liquid or disordered solid is not measured in one specified microscopic configuration. During the measurement the system samples many configurations, and the measured elastic intensity corresponds to an equilibrium average of Eq. (7.17).

This motivates the definition of the static structure factor,

$$S(\mathbf{q}) = \langle |\rho_{\mathbf{q}}|^2 \rangle = \frac{1}{N} \left\langle \sum_{j,l=1}^N e^{-i\mathbf{q} \cdot (\mathbf{R}_j - \mathbf{R}_l)} \right\rangle \quad (7.19)$$

For $\mathbf{q} \neq 0$, $\rho_{\mathbf{q}}$ is a Fourier component of the density fluctuation. Equation (7.19) therefore says that $S(\mathbf{q})$ measures the mean-square amplitude of that density fluctuation at wavevector $\mathbf{q}$. Equivalently, because the double sum depends on the separation vectors $\mathbf{R}_j - \mathbf{R}_l$, $S(\mathbf{q})$ contains information about spatial correlations in the medium.

Averaging Eq. (7.17) gives

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = N \left(\frac{d\sigma}{d\Omega} \right)_1 S(\mathbf{q}) = N \left(\frac{mv_q}{2\pi\hbar^2} \right)^2 S(\mathbf{q}) \quad (7.20)$$

For a homogeneous isotropic medium, $S(\mathbf{q})$ depends only on $q = |\mathbf{q}|$, and we write it simply as $S(q)$. If the nuclear positions are uncorrelated, then for non-zero q the terms with $j \neq l$ average to zero and $S(q) = 1$. Any systematic departure of $S(q)$ from unity therefore reflects correlations in the positions of the nuclei.

7.5 Geometry of the transferred wavevector

Let θ be the angle between $\mathbf{k}$ and $\mathbf{k}'$. From $\mathbf{q} = \mathbf{k}' - \mathbf{k}$,

$$q^2 = k'^2 + k^2 - 2kk'\cos\theta \quad (7.21)$$

Elastic scattering gives $k' = k$, hence

$$q = 2k \sin \frac{\theta}{2} = \frac{4\pi}{\lambda} \sin \frac{\theta}{2} \quad (7.22)$$

where $\lambda = 2\pi/k$ is the neutron wavelength. Thus, for a fixed incident wavelength, changing the scattering angle changes the value of q at which the structure factor is measured. Small angles probe small q ; larger angles probe larger q .

7.6 What the elastic scattering experiment measures

Let $I(q)$ denote the measured elastic intensity at transferred wavevector q . For comparison, define $I_0(q)$ as the intensity corresponding to N independent nuclei with the same one-nucleus cross section,

$$I_0(q) \propto N \left(\frac{d\sigma}{d\Omega} \right)_1.$$

Equation (7.20) then gives

$$\frac{I(q)}{I_0(q)} = S(q) \quad (7.23)$$

This is the central result of the section. Once the one-nucleus scattering contribution is known, the angular dependence of the elastic intensity gives the static structure factor and therefore information about spatial density correlations in the medium.

The derivation above assumes elastic scattering. If the neutron changes its kinetic energy, the medium must gain or lose the corresponding energy. The scattering event must then be characterised by both the transferred wavevector $\mathbf{q}$ and the transferred energy $\hbar\omega = E_{\mathbf{k}} - E_{\mathbf{k}'}$. In that case the relevant quantity is the dynamic structure factor $S(\mathbf{q}, \omega)$ rather than the static $S(\mathbf{q})$.

8. Sudden perturbation

The approximation methods used so far were controlled mainly by the *strength* of an interaction. The sudden approximation is different. Here the essential question is how rapidly the Hamiltonian changes compared with the intrinsic time scales of the quantum system.

Suppose the Hamiltonian changes from $\hat{H}_i$ to $\hat{H}_f$ during a short interval of duration τ . We want to know whether the state has time to readjust while this change takes place.

8.1 Physical criterion

Let $\psi_i(q)$ be the initial normalised wavefunction, where q denotes the coordinates of the system. Initially, it is an eigenfunction of the Hamiltonian $\hat{H}_i$:

$$\hat{H}_i \psi_i(q) = E_i \psi_i(q).$$

During the interval $0 \leq t \leq \tau$, the Hamiltonian changes from $\hat{H}_i$ to $\hat{H}_f$. We write

$$\hat{H}(t) = \hat{H}_i + \hat{V}(t),$$

where $\hat{V}(t)$ describes the change during the switching interval. The question is whether the wavefunction changes appreciably before the switch is completed.

The time-dependent Schrödinger equation, with the initial condition $\psi(q, 0) = \psi_i(q)$, is

$$i\hbar \frac{\partial \psi(q, t)}{\partial t} = \hat{H}(t) \psi(q, t).$$

Integrating over $0 \leq t \leq \tau$ and keeping the leading term for a short switching time gives

$$\psi(q, \tau) \simeq \psi_i(q) - \frac{i}{\hbar} \int_0^\tau dt \hat{H}(t) \psi_i(q).$$

To determine whether the state has changed, consider a complete orthonormal set of eigenfunctions $\psi_n(q)$ of $\hat{H}_i$, with $\psi_i(q)$ among them. The amplitude along $\psi_n(q)$ at time t is

$$c_n(t) = \int dq \overline{\psi_n(q)} \psi(q, t).$$

For $n \neq i$ this amplitude is initially zero. Moreover,

$$\int dq \overline{\psi_n(q)} \hat{H}_i \psi_i(q) = E_i \int dq \overline{\psi_n(q)} \psi_i(q) = 0,$$

because eigenfunctions belonging to the orthonormal set are orthogonal. Define the perturbation matrix element

$$V_{ni}(t) = \int dq \overline{\psi_n(q)} \hat{V}(t) \psi_i(q).$$

Projecting the short-time wavefunction onto $\psi_n(q)$ therefore gives, for $n \neq i$,

$$c_n(\tau) \simeq -\frac{i}{\hbar} \int_0^\tau V_{ni}(t) dt. \quad (8.1)$$

Equation (8.1) shows how the changing Hamiltonian generates amplitudes in states orthogonal to the initial state. To estimate their combined magnitude, define the coupling-energy scale

$$\Delta E = \max_{0 \leq t \leq \tau} \left[\sum_{n \neq i} |V_{ni}(t)|^2 \right]^{1/2}.$$

The probability of occupying states orthogonal to the initial wavefunction is, to leading order, bounded by

$$P_{\text{change}} = \sum_{n \neq i} |c_n(\tau)|^2 \lesssim \left(\frac{\Delta E \tau}{\hbar} \right)^2.$$

The wavefunction therefore changes negligibly during the switch if

$$\tau \ll \frac{\hbar}{\Delta E}. \quad (8.2)$$

The time $\hbar/\Delta E$ is determined by those matrix elements of $\hat{V}(t)$ that can change the initial wavefunction. If the switch takes place much faster, $\psi(q, \tau)$ differs from $\psi_i(q)$ only by an overall phase, to the accuracy of the approximation. Although the wavefunction remains approximately unchanged, the Hamiltonian has become $\hat{H}_f$. The next subsection examines the wavefunction in terms of the eigenfunctions of this final Hamiltonian.

8.2 Transition probabilities as overlaps

Let $\psi_i(q)$ be the wavefunction immediately before the sudden change. Let $\psi_f(q)$ denote the normalised eigenfunctions of the new Hamiltonian,

$$\hat{H}_f \psi_f(q) = E_f \psi_f(q).$$

Because the wavefunction does not have time to readjust during the switch, immediately afterwards it is still $\psi_i(q)$. The functions $\psi_f(q)$ form a complete basis, so

$$\psi_i(q) = \sum_f C_{fi} \psi_f(q) \quad (8.3)$$

Multiplying by the complex conjugate $\overline{\psi_f(q)}$ and integrating over the coordinates gives

$$C_{fi} = \int dq \overline{\psi_f(q)} \psi_i(q) \quad (8.4)$$

The probability that a measurement of the new Hamiltonian gives the eigenstate f is therefore

$$W_{i \rightarrow f} = |C_{fi}|^2 = \left| \int dq \overline{\psi_f(q)} \psi_i(q) \right|^2 \quad (8.5)$$

Because the new eigenfunctions form a complete orthonormal set,

$$\sum_f W_{i \rightarrow f} = \sum_f |C_{fi}|^2 = 1.$$

Thus the sudden approximation reduces the transition problem to a calculation of overlaps between the wavefunction before the switch and the eigenfunctions of the Hamiltonian after the switch.

After the switch has ended, each component evolves with the energy of the new Hamiltonian,

$$\psi(q, t) = \sum_f C_{fi} e^{-iE_f(t-t_0)/\hbar} \psi_f(q),$$

where t_0 denotes the end of the rapid change.

8.3 Example: sudden change of oscillator parameters

Consider a one-dimensional harmonic oscillator whose mass and frequency change suddenly,

$$m, \omega \rightarrow m', \omega'.$$

Assume that before the change the oscillator is in its ground state. The initial ground-state wavefunction is

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) \quad (8.6)$$

The ground state of the new Hamiltonian is

$$\phi_0(x) = \left(\frac{m'\omega'}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m'\omega' x^2}{2\hbar}\right) \quad (8.7)$$

Immediately after the sudden change the wavefunction is still $\psi_0(x)$. The amplitude for finding the oscillator in the ground state of the new Hamiltonian is therefore

$$C_0 = \int_{-\infty}^{\infty} dx \phi_0(x) \psi_0(x).$$

Substituting Eqs. (8.6) and (8.7),

$$C_0 = \left(\frac{m\omega m'\omega'}{\pi^2 \hbar^2}\right)^{1/4} \int_{-\infty}^{\infty} dx \exp\left[-\frac{m\omega + m'\omega'}{2\hbar} x^2\right].$$

Using

$$\int_{-\infty}^{\infty} e^{-Ax^2} dx = \sqrt{\frac{\pi}{A}},$$

we obtain

$$C_0 = \frac{\sqrt{2} (m\omega m'\omega')^{1/4}}{\sqrt{m\omega + m'\omega'}}.$$

Hence

$$W_{0 \rightarrow 0'} = |C_0|^2 = \frac{2\sqrt{m\omega m'\omega'}}{m\omega + m'\omega'} \quad (8.8)$$

The result depends on the products $m\omega$ and $m'\omega'$, because these quantities determine the widths of the two Gaussian ground states. If

$$m\omega = m'\omega',$$

the two ground-state wavefunctions are identical and $W_{0 \rightarrow 0'} = 1$. If their widths are different, $W_{0 \rightarrow 0'} < 1$ and the remaining probability is distributed among excited states of the new oscillator.

8.4 Example: sudden change of nuclear charge in beta decay

A nuclear beta decay can change the nuclear charge on a time scale much shorter than a typical electronic orbital period. During the nuclear event, the electronic wavefunction then has essentially no time to readjust. Immediately afterwards, the same electronic wavefunction must be expanded in eigenfunctions of the new Coulomb Hamiltonian.

For a hydrogen-like 1s state, the initial ground-state wavefunction is

$$\psi_Z(r) = \left(\frac{Z^3}{\pi a_0^3} \right)^{1/2} e^{-Zr/a_0} \quad (8.9a)$$

After the change of nuclear charge, the corresponding 1s eigenfunction of the new Hamiltonian is

$$\phi_{Z'}(r) = \left(\frac{Z'^3}{\pi a_0^3} \right)^{1/2} e^{-Z'r/a_0} \quad (8.9b)$$

Suppose the charge changes suddenly from Z to Z' . The amplitude for the electron to be found in the new 1s state is

$$C_{1s} = \int d^3r \overline{\phi_{Z'}(r)} \psi_Z(r) = \frac{8(ZZ')^{3/2}}{(Z + Z')^3}.$$

Therefore

$$W_{1s \rightarrow 1s'} = |C_{1s}|^2 = \left[\frac{4ZZ'}{(Z + Z')^2} \right]^3 \quad (8.10)$$

For the change $Z = 1 \rightarrow Z' = 2$,

$$W_{1s \rightarrow 1s'} = \left(\frac{8}{9} \right)^3 \simeq 0.702.$$

The remaining probability corresponds to excited bound states and continuum states of the new Hamiltonian. The essential point is that no weak-coupling expansion is involved: the approximation is controlled by the short switching time, Eq. (8.2).

8.5 Example: K-shell ionisation accompanying beta decay

The previous example asked for the probability that an electron remains in a bound state after a sudden change of the nuclear charge. We now ask a different question: **what is the probability that a K-shell electron is transferred to the continuum and therefore leaves the atom?**

Consider a heavy atom of nuclear charge Z undergoing β^- decay. The nuclear charge changes as

$$Z \rightarrow Z + 1.$$

Assume that the emitted β particle crosses the region occupied by the K shell in a time much shorter than the orbital period of a K -shell electron. During this short interval the electronic wavefunction has no time to readjust. This is precisely the sudden approximation: immediately after the decay the electron is still described by the wavefunction it had before the decay.

There is a second approximation, specific to a heavy atom. Before the decay the Coulomb potential of the nucleus is $-Z/r$ in atomic units, whereas the change caused by $Z \rightarrow Z + 1$ is only $-1/r$. Their ratio is therefore

$$\frac{|\Delta V|}{|V_Z|} = \frac{1}{Z} \ll 1.$$

Thus the change is sudden in time, but small relative to the original Coulomb field when $Z \gg 1$. The sudden approximation tells us **which state the electron has immediately after the decay**; first-order stationary perturbation theory then allows us to express the eigenstates of the new Hamiltonian in terms of those of the old one.

We use atomic units in this example,

$$\hbar = m_e = \frac{e^2}{4\pi\epsilon_0} = 1.$$

For one electron the Hamiltonian before and after the decay may be written as

$$\hat{H}_Z = -\frac{1}{2}\nabla^2 - \frac{Z}{r}, \quad \hat{H}_{Z+1} = \hat{H}_Z + \Delta\hat{V}, \quad \Delta V(r) = -\frac{1}{r}.$$

Let $\psi_0(\mathbf{r})$ denote the initial hydrogen-like 1s eigenfunction of $\hat{H}_Z$. Its energy is

$$E_0 = -\frac{Z^2}{2}.$$

Let $\psi_k(\mathbf{r})$ denote a continuum eigenfunction of the same reference Hamiltonian $\hat{H}_Z$, labelled by its asymptotic momentum k and normalised with respect to k . Its energy is

$$E_k = \frac{k^2}{2}.$$

After the decay, however, the continuum eigenfunctions belong to the new Hamiltonian $\hat{H}_{Z+1}$. Denote one of them by $\phi_k(\mathbf{r})$. The sudden approximation says that the electron itself is still in the old wavefunction ψ_0 immediately after the decay. Therefore, the one-electron probability of finding it in the final continuum interval $[k, k + dk]$ is

$$dW_1 = |C_{k0}|^2 dk, \quad C_{k0} = \int d^3r \overline{\phi_k(\mathbf{r})} \psi_0(\mathbf{r}).$$

We therefore need the overlap of the unchanged initial state with a continuum eigenfunction of the **new** Hamiltonian. To this end we construct the final continuum state perturbatively in the change of the Coulomb potential; formally this is the continuum, or scattering-state, version of stationary perturbation theory. This is where the small parameter $1/Z$ is used. Since $\hat{H}_{Z+1} = \hat{H}_Z + \Delta\hat{V}$, first-order stationary perturbation theory expresses ϕ_k in terms of eigenfunctions of $\hat{H}_Z$. For the present overlap we only need the component proportional to ψ_0 :

$$\phi_k(\mathbf{r}) = \psi_k(\mathbf{r}) + \psi_0(\mathbf{r}) \frac{V_{0k}}{E_k - E_0} + \{\text{terms orthogonal to } \psi_0\},$$

where

$$V_{k0} = \int d^3r \overline{\psi_k(\mathbf{r})} \Delta V(r) \psi_0(\mathbf{r}), \quad V_{0k} = \overline{V_{k0}}.$$

The unperturbed bound and continuum eigenfunctions are orthogonal, so $\int d^3r \overline{\psi_k} \psi_0 = 0$. Substituting the first-order expression for ϕ_k into the overlap therefore gives

$$C_{k0} \simeq \frac{V_{k0}}{E_k - E_0} \quad (8.11)$$

Equation (8.11) is thus the first-order expression for the projection of the unchanged initial K -shell wavefunction onto a continuum eigenfunction of the new Hamiltonian.

Because the continuum states are normalised with respect to k , $|C_{k0}|^2 dk$ is the one-electron probability for an outgoing momentum between k and $k + dk$. The K shell contains two electrons. To the present accuracy we neglect the much smaller probability that both are ionised in the same decay, so the probabilities for ejecting either electron add. Hence

$$dW = 2|C_{k0}|^2 dk = 2 \frac{|V_{k0}|^2}{(E_k - E_0)^2} dk.$$

Using

$$E_k - E_0 = \frac{k^2 + Z^2}{2},$$

we obtain

$$dW = \frac{8|V_{k0}|^2}{(k^2 + Z^2)^2} dk \quad (8.12)$$

Thus Eq. (8.12) is not an additional assumption. It follows directly from the sudden-overlap probability, Eq. (8.11), the continuum normalisation, the two electrons in the K shell, and the energies E_0 and E_k .

It remains to evaluate the matrix element V_{k0} . Since

$$\Delta V(r) = -\frac{1}{r}$$

is spherically symmetric, the angular integral is

$$\int d\Omega \overline{Y_{lm}(\Omega)} Y_{00}(\Omega) = \delta_{l0} \delta_{m0}.$$

The initial K -shell state has $l = 0$, so only continuum states with $l = 0$ contribute. The matrix element is therefore reduced to a radial integral,

$$V_{k0} = - \int_0^\infty dr r \overline{R_{k0}(r)} R_{10}(r).$$

The matrix element V_{k0} is determined mainly by distances $r \sim 1/Z$ (in atomic units), where the initial K -shell $1s$ wavefunction is concentrated. At these short distances, the nuclear Coulomb attraction dominates over the screening produced by the other atomic electrons. This allows us to approximate both the initial bound state and the relevant continuum s -wave by hydrogen-like radial functions corresponding to the Coulomb potential

$$V_C(r) = -\frac{Z}{r}.$$

The initial $1s$ radial wavefunction is

$$R_{10}(r) = 2Z^{3/2} e^{-Zr},$$

while $R_{k0}(r)$ is the regular continuum radial solution of the Schrödinger equation in the same Coulomb potential, with orbital angular momentum $l = 0$ and positive energy $E_k = k^2/2$. Its normalisation is chosen consistently with the continuum-state normalisation used in the ionisation probability. The actual potential experienced by an outgoing electron in a many-electron atom is screened and differs from $-Z/r$, particularly at large distances. The hydrogen-like approximation is justified here by the fact that the radial integral is dominated by the region close to the nucleus, where the nuclear Coulomb field is strongest. It provides an approximate evaluation of the matrix element, rather than an exact description of the outgoing electron.

Carrying out the radial integral in V_{k0} and substituting the result into Eq. (8.12) gives

$$dW = \frac{2^7}{Z^4(1 + k^2/Z^2)^4} f\left(\frac{k}{Z}\right) k dk \quad (8.13)$$

where the remaining dependence produced by the Coulomb continuum radial function is collected in the dimensionless function

$$f(\alpha) = \frac{1}{1 - e^{-2\pi/\alpha}} \exp\left[-\frac{4}{\alpha} \arctan \alpha\right] \quad (8.14)$$

Equation (8.13) is the probability for ionisation into the small momentum interval $[k, k + dk]$. The total ionisation probability is obtained by summing over the entire continuum, which here means integrating over all positive k :

$$W_{\text{ion}} = \int_0^\infty dW.$$

To expose the dependence on the nuclear charge, introduce

$$\alpha = \frac{k}{Z}, \quad k = Z\alpha, \quad dk = Z d\alpha.$$

Then $k dk = Z^2 \alpha d\alpha$, and Eq. (8.13) becomes

$$W_{\text{ion}} = \frac{2^7}{Z^2} \int_0^\infty d\alpha \frac{\alpha f(\alpha)}{(1 + \alpha^2)^4} \quad (8.15)$$

All dependence on Z is now explicit. The integral is dimensionless and has the numerical value

$$\int_0^\infty d\alpha \frac{\alpha f(\alpha)}{(1 + \alpha^2)^4} \simeq 5.10 \times 10^{-3}.$$

Therefore

$$W_{\text{ion}} \simeq \frac{0.65}{Z^2} \quad (8.16)$$

The result also gives a useful consistency check. The change of the Coulomb field is smaller than the original field by the factor $1/Z$. The transition amplitude generated by this change is correspondingly proportional to $1/Z$, while the probability is proportional to its square, giving the final Z^{-2} dependence.

9. How the methods fit together

The methods developed in these notes are controlled by different physical assumptions. Before choosing a calculation scheme, it is useful to identify exactly what is assumed to be small, what is varied, and what observable is being calculated.

Method	Physical assumption	Main result	Control or check
Stationary perturbation theory	weak mixing produced by a time-independent perturbation	energy and eigenstate corrections	$ V_{mn} / E_n^{(0)} - E_m^{(0)} \ll 1$ and small total admixture
Optimised reference Hamiltonian	a solvable reference family can be chosen to approximate a problem with no obvious small parameter	approximate spectrum	choose auxiliary parameters by cancellation or minimisation; check accuracy afterwards
Variational method	the ground-state energy is minimised over normalised trial states	upper bound to E_0	trial states satisfy known symmetry and boundary conditions
Time-dependent perturbation theory	weak time-dependent coupling	transition amplitudes and finite-time probabilities	leading amplitudes and transferred probability remain small
Fermi's golden rule	weak coupling and many closely spaced final states	transition rate	energy-selection width is narrow and depletion remains small
Static Born scattering	first Born approximation and negligible energy transfer to the target	differential cross section and $S(q)$	use plane-wave initial/final states and the static target approximation
Sudden approximation	Hamiltonian changes on a time scale much shorter than the relevant relative-phase time	overlap probabilities	$ \Delta E \tau/\hbar \ll 1$

Several connections between these methods are especially important.

Stationary and time-dependent perturbation theory are controlled by the *strength* of the coupling. The sudden approximation is controlled by a *time-scale ratio*. A large change of Hamiltonian can therefore still be treated as sudden if it occurs rapidly enough.

The route from time-dependent perturbation theory to scattering is:

transition amplitude → **finite-time transition probability** → **continuum of final states** → **transition rate** → **differential cross section**.

Section 6 supplies the passage from finite-time probabilities to Fermi's golden rule. Section 7 then divides the scattering rate by the incident particle flux and shows how, for many nuclei, the measured cross section contains the static structure factor $S(q)$.

The variational method and the optimised-reference method are useful in a different situation: the original Hamiltonian may contain no obvious small physical parameter. The approximation is then

organised by choosing a restricted family of states or reference Hamiltonians. For the variational ground-state calculation there is an additional rigorous statement: the calculated minimum is an upper bound to the exact ground-state energy.

10. Common checks in calculations

10.1 Identify the approximation before calculating

Do not begin with algebra. First state which physical limit is being used.

For non-degenerate stationary perturbation theory, the mixing between the unperturbed levels n and m is controlled by

$$\left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1.$$

For time-dependent perturbation theory, the leading transition amplitudes must remain small and the total probability transferred out of the initial state must remain much smaller than one.

For the golden-rule regime, the interaction time must be long enough to produce narrow energy selection, but not so long that the initial state is substantially depleted. In the notation of Section 6,

$$\frac{\hbar}{\tau} \ll \Delta E_F, \quad w\tau \ll 1,$$

where ΔE_F denotes the energy scale over which the factor $|V(E)|^2 \rho(E)$ changes appreciably.

For the sudden approximation,

$$\frac{|\Delta E| \tau}{\hbar} \ll 1.$$

The optimised-reference and variational methods need not be controlled by a small physical parameter. Their quality is assessed in a different way: by optimisation, by known bounds where available, and by comparison with additional information.

10.2 Check dimensions

Every energy correction must have dimensions of energy. Every transition rate must have dimensions of inverse time. Every probability must be dimensionless. The arguments of exponentials, sine and cosine functions must be dimensionless.

These checks immediately detect many common mistakes: a missing factor of $\hbar$, an incorrect power of momentum, or a missing density-of-states factor.

10.3 Inspect denominators

In stationary perturbation theory, the first-order coefficient describing the admixture of level m into level n contains

$$\frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}}.$$

If two unperturbed levels approach one another, the denominator becomes small and the non-degenerate expansion is no longer controlled. This is not an algebraic inconvenience; it signals strong mixing and the need for degenerate or nearly degenerate perturbation theory.

10.4 Use symmetry before integration

Symmetry should be used before carrying out long calculations. Examples from these notes include:

- an odd operator has a vanishing diagonal matrix element in a parity eigenstate;
- in a parity-symmetric one-dimensional potential, even and odd states are exactly orthogonal;
- in the sudden oscillator example, the initial ground state is even, so its overlap with every odd eigenstate of the new oscillator vanishes.

A zero matrix element caused by symmetry should be recognised before evaluating an integral.

10.5 Distinguish an energy shift from a change of state

The first-order energy shift is determined by the diagonal matrix element

$$E_n^{(1)} = V_{nn},$$

whereas the first-order change of the wavefunction is determined by the off-diagonal matrix elements,

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)}.$$

Therefore

$$V_{nn} = 0$$

does **not** imply

$$\psi_n^{(1)} = 0.$$

A perturbation can leave the energy unchanged at first order while changing the wavefunction at first order.

10.6 Keep the time-dependent phase factor

In first-order time-dependent perturbation theory,

$$C_f^{(1)}(t_1) = \frac{1}{i\hbar} \int_{t_0}^{t_1} dt e^{i\omega_{fi}t} V_{fi}(t).$$

The factor $e^{i\omega_{fi}t}$ determines whether contributions generated at different times reinforce one another or cancel. It is the origin of resonance and energy selection and should not be removed before the time integral is evaluated.

10.7 Distinguish a finite-time probability from a transition rate

For one specified final state and a perturbation acting for a finite time τ , Section 6 gives

$$W_{i \rightarrow f}(\tau) = \frac{4|V_{fi}|^2}{\hbar^2} \frac{\sin^2(\omega_{fi}\tau/2)}{\omega_{fi}^2}.$$

This is a finite-time probability. It is not a constant rate.

A time-independent rate appears only after many closely spaced final states are available and the probabilities are summed over that dense set. Only under the resulting energy integral can the narrow finite-time peak be replaced by its delta-sequence limit.

10.8 Keep perturbative orders consistent with normalisation

When checking normalisation, it is useful to restore the auxiliary switching parameter λ that was introduced to organise perturbation theory. Write

$$\hat{H}(t) = \hat{H}_0 + \lambda \hat{V}(t).$$

For a system prepared initially in state i , the amplitudes are expanded as

$$C_i(t) = 1 + \lambda C_i^{(1)}(t) + \lambda^2 C_i^{(2)}(t) + \dots,$$

and, for $f \neq i$,

$$C_f(t) = \lambda C_f^{(1)}(t) + \lambda^2 C_f^{(2)}(t) + \dots.$$

The exact state is normalised, so

$$|C_i(t)|^2 + \sum_{f \neq i} |C_f(t)|^2 = 1.$$

Now expand this equation in powers of λ . At first order one obtains

$$2 \operatorname{Re} C_i^{(1)}(t) = 0.$$

Thus the first-order correction to the initial-state amplitude can change its phase, but it does not change the survival probability at first order.

At second order the normalisation condition gives

$$2 \operatorname{Re} C_i^{(2)}(t) + |C_i^{(1)}(t)|^2 + \sum_{f \neq i} |C_f^{(1)}(t)|^2 = 0.$$

This equation shows exactly how probability conservation works in perturbation theory. The transition probabilities obtained from the first-order amplitudes $C_f^{(1)}$ enter the normalisation condition at second order in the bookkeeping parameter λ . They are accompanied by a second-order decrease of the survival probability of the initial state.

Therefore, if normalisation is checked while retaining the probabilities $|C_f^{(1)}|^2$, the initial-state amplitude must also be retained through $C_i^{(2)}$. After the perturbative orders have been identified consistently, one sets $\lambda = 1$.

10.9 Check the continuum normalisation explicitly

For plane waves normalised in a large volume V ,

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3k.$$

This factor is essential in both Fermi's golden rule and the scattering calculation. In Section 7 the factors of the arbitrary normalisation volume cancel only after the scattering rate is divided by the incident flux.

10.10 Keep the two parts of the scattering result distinct

For the static Born scattering problem of Section 7, Eq. (7.8) can be written as

$$M_{\mathbf{k}'\mathbf{k}} = \frac{\sqrt{N}}{V} v_q \rho_{\mathbf{q}}.$$

Here v_q is the Fourier transform of the potential of one nucleus, whereas

$$\rho_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{-i\mathbf{q} \cdot \mathbf{R}_j}$$

contains the nuclear positions. After equilibrium averaging,

$$S(\mathbf{q}) = \langle |\rho_{\mathbf{q}}|^2 \rangle.$$

For a homogeneous medium and $\mathbf{q} \neq 0$, $\rho_{\mathbf{q}}$ is the normalised Fourier component of the density fluctuation. The static structure factor therefore describes spatial density correlations. If the neutron also transfers energy to the medium, the relevant quantity is instead the dynamic structure factor $S(\mathbf{q}, \omega)$.

10.11 In the variational method, build known physics into the trial state

Before introducing variational parameters, impose what is already known exactly: symmetry, boundary conditions, regularity, nodal structure where known, and asymptotic decay.

For a normalised trial function,

$$E[\psi] = \int dq \overline{\psi(q)} \hat{H} \psi(q), \quad \int dq |\psi(q)|^2 = 1,$$

and the variational theorem gives

$$E_0 \leq E[\psi].$$

The variational parameters should optimise what is unknown, rather than compensate for a trial function that violates exact constraints.

10.12 In the sudden approximation, compare time scales through energy differences

The sudden criterion is

$$\frac{|\Delta E| \tau}{\hbar} \ll 1.$$

Only energy differences enter because the zero of energy is arbitrary. During the rapid change the wavefunction is approximately unchanged. The transition probabilities are then obtained by expanding that unchanged wavefunction in the eigenfunctions of the new Hamiltonian.

10.13 Check limiting cases and physical meaning

A useful final test is to ask whether the formula behaves correctly in simple limits.

For example, when the perturbation is removed, perturbative corrections and transition amplitudes must vanish. In the sudden oscillator example, $m\omega = m'\omega'$ makes the two ground-state wavefunctions identical and gives $W_{0 \rightarrow 0} = 1$. In elastic scattering, $k' = k$ and therefore $q = 2k \sin(\theta/2)$. For a homogeneous uncorrelated arrangement and non-zero q , $S(q) = 1$.

The calculation should also have a clear physical interpretation. The negative second-order correction to a non-degenerate ground-state energy comes from coupling to higher unperturbed states. The golden-rule delta function expresses sharp energy selection produced by a long observation time. The structure factor in the elastic cross section contains information about spatial correlations. In the sudden approximation, the unchanged wavefunction is resolved into eigenfunctions of the new Hamiltonian.

11. Formula sheet and logical map

This section collects the main formulas together with the condition under which each of them is used.

11.1 Non-degenerate stationary perturbation theory

For

$$\hat{H} = \hat{H}_0 + \lambda \hat{V},$$

with a non-degenerate unperturbed eigenfunction $\psi_n^{(0)}$,

$$E_n^{(1)} = V_{nn}$$

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)}$$

$$E_n^{(2)} = \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}$$

A pairwise applicability check is

$$\left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1$$

and the total first-order admixture should satisfy

$$\sum_{m \neq n} \left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right|^2 \ll 1.$$

11.2 Degenerate perturbation theory

For an r -fold degenerate unperturbed level $E^{(0)}$, first restrict the perturbation to the degenerate subspace D and solve

$$\sum_{\beta \in D} V_{\alpha\beta} c_{\beta} = E^{(1)} c_{\alpha}.$$

The eigenvectors of this matrix give the correct zeroth-order linear combinations inside the degenerate subspace, and its eigenvalues give the first-order energy shifts.

11.3 Optimised reference Hamiltonian

Introduce a solvable family $\hat{H}_0(a)$ and write the exact Hamiltonian as

$$\hat{H} = \hat{H}_0(a) + \hat{V}(a).$$

The auxiliary parameter a may be chosen, for example, by cancelling the first correction,

$$E_n^{(1)}(a) = 0,$$

or by making the approximate energy stationary,

$$\frac{dE_{\text{approx}}(a)}{da} = 0.$$

The parameter a is not a new physical parameter of the original Hamiltonian; it specifies the chosen reference problem.

11.4 Variational method

For any normalised trial function,

$$E_0 \leq E[\psi], \quad E[\psi] = \int dq \overline{\psi(q)} \hat{H} \psi(q).$$

For a trial family $\psi(q; a_1, a_2, \dots)$, stationary values satisfy

$$\frac{\partial E}{\partial a_i} = 0,$$

and for a ground-state calculation one chooses the minimum.

11.5 Time-dependent perturbation theory

For $f \neq i$,

$$C_f^{(1)}(t_1) = \frac{1}{i\hbar} \int_{t_0}^{t_1} dt e^{i\omega_{fi}t} V_{fi}(t)$$

with

$$\omega_{fi} = \frac{E_f^{(0)} - E_i^{(0)}}{\hbar}.$$

If the auxiliary parameter λ is restored, then for $f \neq i$,

$$C_f(t) = \lambda C_f^{(1)}(t) + \lambda^2 C_f^{(2)}(t) + \dots$$

The leading non-vanishing transition probability is therefore

$$W_{i \rightarrow f} = |C_f^{(1)}|^2$$

This probability is second order in the bookkeeping parameter λ . After the perturbative order has been identified, λ is set equal to one.

11.6 Fermi's golden rule

For one specified final state under a constant perturbation acting for a time τ ,

$$W_{i \rightarrow f}(\tau) = \frac{4|V_{fi}|^2 \sin^2(\omega_{fi}\tau/2)}{\hbar^2 \omega_{fi}^2}.$$

For a dense continuum, with

$$dN_f = \rho_f(E_f) dE_f,$$

the golden-rule rate is

$$w = \frac{2\pi}{\hbar} |V_{fi}|_{E_f=E_i}^2 \rho_f(E_i).$$

The finite-time energy-selection peak must be narrow on the scale over which $|V(E)|^2 \rho(E)$ changes, while depletion remains small:

$$w\tau \ll 1.$$

11.7 Static Born scattering and the structure factor

The transferred wavevector is

$$\mathbf{q} = \mathbf{k}' - \mathbf{k}, \quad q = 2k \sin \frac{\theta}{2} \quad \text{for elastic scattering.}$$

For N identical nuclei,

$$\rho_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{-i\mathbf{q} \cdot \mathbf{R}_j}, \quad S(\mathbf{q}) = \langle |\rho_{\mathbf{q}}|^2 \rangle.$$

The static first-Born differential cross section is

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = N \left(\frac{mv_q}{2\pi\hbar^2} \right)^2 S(\mathbf{q}).$$

If the target is homogeneous and isotropic, write $S(\mathbf{q}) = S(q)$. If energy transfer to the target is resolved, the static description must be replaced by the dynamic structure factor $S(\mathbf{q}, \omega)$.

11.8 Sudden approximation

The criterion is

$$\frac{|\Delta E|\tau}{\hbar} \ll 1.$$

If $\psi_i(q)$ is the wavefunction immediately before the switch and $\psi_f(q)$ is a normalised eigenfunction of the new Hamiltonian,

$$W_{i \rightarrow f} = \left| \int dq \overline{\psi_f(q)} \psi_i(q) \right|^2.$$

11.9 Logical map

The common sequence is **physical limit** $\rightarrow$ **approximation** $\rightarrow$ **observable result**. The central discipline is the same throughout: identify the physical limit first, express its control condition clearly, and only then organise the calculation around it.