Further Quantum Mechanics

University of Cambridge Part II Natural Sciences Tripos

Yue Wu

Yusuf Hamied Department of Chemistry
Lensfield Road,
Cambridge, CB2 1EW

yw628@cam.ac.uk

Acknowledgements

Nothing in these lecture notes is original. They are largely based on the notes by Dr. John Morgan, who lectured this course in 2025. Moreover, they are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Preface

This course focuses on quantum mechanics, and it is slightly more advanced than what you have learned in Part IB Chemistry A: *Introduction to Quantum Mechanics*. It will mostly focus on perturbation theory, including both the time independent and the time dependent cases, and it also covers topics that are removed from the A4 *Theoretical Techniques* course this year, namely normal modes. This course will avoid the rigorous mathematical formulation of quantum mechanics, and especially, it will not introduce concepts like projective Hilbert space or functional analysis. If you want a more mathematical approach to quantum mechanics, you can find my notes on Mathematical Tripos Part II: *Principles of Quantum Mechanics*.

Contents

1 Foundational Principles

We will start from a revision of the foundational principles of quantum mechanics that should be familiar from part IB Chemistry A.

1.1 Wavefunctions and Operators

In quantum mechanics, all physical information about a system is embodied in its wavefunction, denoted Ψ. The wavefunction is complex-valued, and we will use the *position representation* of the wavefunction, so it a function of the spatial coordinates. In the Born's interpretation, the probability density of finding a particle at $\mathbf{r}$ is

$$P(\mathbf{r}) \propto |\Psi(\mathbf{r})|^2 \,. \tag{1.1}$$

Wavefunctions should be single-valued and (at least) twice differentiable. Under such interpretation, we would often choose to normalise the wavefunction such that

$$\int \mathrm{d}\tau \, \Psi^*\Psi = 1 \,, \tag{1.2}$$

where $\mathrm{d}\tau$ is a shorthand notation for integrating over all spatial coordinates. This integral should converge for a proper wavefunction, so that it can be normalised.

A quantum mechanical system is defined by its Hamiltonian, H, the total energy operator. The Hamiltonian usually includes the kinetic and potential energies of the particles. A quantum mechanical system evolve according to the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi \,. \tag{1.3}$$

Often, the Hamiltonian operator is independent of time. For such time-independent system, the wavefunction satisfies the time-independent Schrödinger equation (which we often directly refers to as the Schrödinger equation)[1]

$$H\psi = E\psi \,, \tag{1.5}$$

where the constant E is the energy of the system[2] and $\psi(\mathbf{r})$ is now a wavefunction independent of time. This is a (partial) differential equation, so it is only analytically solvable in a few limited cases, most of which you have seen already. Approximations are generally needed to solve for more complex systems, and one of the most important approximation techniques, the perturbation theory, is exactly the main theme of this course.

In general, the Schrödinger equation will have multiple (usually a countably infinite number of) solutions, which can be indexed by a quantum number n such that

$$H\psi_n = E_n\psi_n \,. \tag{1.6}$$

The ψ_n are different states available to the system, and the state with the lowest energy E_n is known as the ground state. We usually arrange the states in sequence so that the ground state is labelled ψ_0 (or ψ_1 if you find it more convenient to start numbering from $n = 1$). We say two or more states are degenerate if they have the same energy. Any linear combination of degenerate wavefunctions is also a solution to the Schrödinger equation with the same energy.

---

[1]This is because if H is independent of t, then a special class of solutions of the time-dependent Schrödinger equation exists

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/h} \,, \tag{1.4}$$

where $\psi(\mathbf{x})$ is independent of time and satisfied the time-independent Schrödinger equation.

[2]This is essentially the conservation of energy — if a system has time-translational symmetry, then the energy of the system is conversed. This is an example of the Noether's theorem.

The Schrödinger equation is a specific example of the more general eigenvalue equation. In quantum mechanics, all physical observables A has a corresponding operator $\hat{A}$, if the wavefunction satisfies

$$\hat{A}\psi = a\psi\,, \tag{1.7}$$

then the value of A will always be measured to be a. If this equation is not satisfied, then measured values of A will be drawn from a probability distribution. The expectation and the uncertainty of A is

$$\langle A \rangle = \frac{\int \mathrm{d}\tau\,\psi^* \hat{A}\psi}{\int \mathrm{d}\tau\,\psi^*\psi} \tag{1.8}$$

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}\,. \tag{1.9}$$

To avoid writing integrations over and over again, we introduce the Dirac bra-ket notation, in which a bra is $\langle\psi| \equiv \psi^*$ and a ket is $|\psi\rangle \equiv \psi$, and a pair of a bra and a ket forms a bra-ket (or just bracket), which implies integrations[3]

$$\langle\psi|\psi\rangle := \int \mathrm{d}\tau\,\psi^*\psi\,. \tag{1.10}$$

An operator can be placed in middle of a bra-ket, so

$$\left\langle\psi\middle|\hat{A}\middle|\psi\right\rangle := \int \mathrm{d}\tau\,\psi^* \hat{A}\psi\,. \tag{1.11}$$

To know what the quantum operator $\hat{A}$ corresponding to a physical observable A is, we write the classical expression of A in terms of position and momentum, and we replace the position and momentum with their quantum operators[4]

$$\hat{x} \longrightarrow x \tag{1.12}$$

$$\hat{p}_x \longrightarrow -i\hbar\frac{\partial}{\partial x}\,. \tag{1.13}$$

In general, a wavefunction cannot simultaneously be an eigenfunction of two (or more) operators. However, this will happen if and only if the commutator of the two operators, defined as

$$[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A} \tag{1.14}$$

is zero. If this is the case, then we say the two operators commute. If the condition is not met, then the commutator places a lower bound on the product of the uncertainties

$$\Delta A \Delta B \geq \frac{1}{2}\left|\langle[A, B]\rangle\right|\,. \tag{1.15}$$

1.2 Hermitian Operators

Not all operators can correspond to physical observables. First of all, it has to be linear, so that

$$\hat{A}(a\,|\phi\rangle + b\,|\psi\rangle) = a\hat{A}\,|\phi\rangle + b\hat{A}\,|\psi\rangle\,, \tag{1.16}$$

---

[3] Actually more formally, $|\psi\rangle$ what truly fundamental, and is the quantum state of a system living in an abstract vector space called a (projective) Hilbert space $\mathcal{H}$. The wavefunction $\psi(\mathbf{x})$ is just the position-space representation of the vector $|\psi\rangle$. If a quantum state is a vector in the Hilbert space, $|\psi\rangle \in \mathcal{H}$, then an operator is an operator in $\mathcal{H}$, $\hat{A} : \mathcal{H} \to \mathcal{H}$, and a bra will be a dual vector living in the dual space of $\mathcal{H}$, $\langle\psi| \in \mathcal{H}^*$. What allows us to do this is the Risez representation theorem, which shows for any $|\psi\rangle \in \mathcal{H}$, there is always a corresponding $\langle\psi| \in \mathcal{H}^*$ and *vice versa*.

[4] This is the position-space representations of these operators, which we can act on wavefunctions that are written as a function of spatial coordinates. These operators have different representations in different basis.

where $a, b \in \mathbb{C}$ are constants. Moreover, operators that correspond to physical observable must be Hermitian. For any linear operator $\hat{A}$, we define its adjoint $\hat{A}^\dagger$ to be the operator such that

$$\left\langle \phi \middle| \hat{A} \middle| \psi \right\rangle = \left\langle \hat{A}^\dagger \phi \middle| \psi \right\rangle . \tag{1.17}$$

A operator is Hermitian if it is self-adjoint,[5] meaning $\hat{A} = \hat{A}^\dagger$, so that

$$\left\langle \phi \middle| \hat{A} \middle| \psi \right\rangle = \left\langle \hat{A}\phi \middle| \psi \right\rangle = \left\langle \psi \middle| \hat{A} \middle| \phi \right\rangle^* . \tag{1.18}$$

It is not difficult to show that taking the adjoint has the following properties:

$$(a\hat{A})^\dagger = a^* \hat{A}^\dagger \tag{1.19}$$

$$(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger \tag{1.20}$$

$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger \tag{1.21}$$

$$(\hat{A}\ket{\psi})^\dagger = \bra{\psi} \hat{A}^\dagger . \tag{1.22}$$

A Hermitian operator has the following nice properties:

(i) The eigenvalues are all real.

(ii) Eigenfunctions with different eigenvalue are orthogonal. Even if some eigenfunctions are degenerate, orthogonal eigenfunctions can always be constructed.

(iii) The eigenfunction form a complete basis set.

Two functions are orthogonal means that their inner product, defined as $\langle \phi | \psi \rangle$, is zero. Moreover, the completeness of the basis set means that any wavefunction ψ satisfying the boundary conditions can be expressed as a linear combination of the eigenfunctions,

$$\psi = \sum_i c_i \phi_i . \tag{1.23}$$

We can always make this basis set, which we denote as $\{\phi_i\}$, orthonormal by normalising the orthogonal eigenfunctions, so that $\langle \phi_i | \phi_j \rangle = \delta_{ij}$. Then taking the inner product of the above eigenfunction expansion with ϕ_j, we have

$$\begin{aligned} \langle \phi_j | \psi \rangle &= \left\langle \phi_j \middle| \sum_i c_i \phi_i \right\rangle \\ &= \sum_i c_i \langle \phi_j | \phi_i \rangle \\ &= \sum_i c_i \delta_{ij} = c_j , \end{aligned} \tag{1.24}$$

and so

$$\psi = \sum_i \langle \phi_i | \psi \rangle \ket{\phi_i} . \tag{1.25}$$

From now on, we will denote the eigenstates just by its quantum number in the bra-ket notation: $\ket{i} \equiv \ket{\phi_i}$, so that for example, we have

$$\ket{\psi} = \sum_i \langle i | \psi \rangle \ket{i} . \tag{1.26}$$

---

[5]There is actually some extremely subtle differences between self-adjoint and Hermitian operators in functional analysis, but they would only bother pure mathematicians.

1.3 Unitary Operators

Apart from position space representations, there are many other basis we can represent the wavefunctions, such as the momentum space, or the basis spanned by eigenstates. An important class of linear operators is the *unitary operators* that are used for transformations of basis. To define unitary operators, we first need to define the identity operator. We define the identity operator to be an operator $\hat{I}$ such that

$$\hat{I}\ket{\psi} = \ket{\psi} \tag{1.27}$$

for any $\ket{\psi}$. An useful way of constructing the identity operator is the *resolution of identity*. For any complete orthonormal basis set $\{\phi_i\}$, we have

$$\hat{I} = \sum_i \ket{i}\bra{i} . \tag{1.28}$$

This is because if we act this operator on any $\ket{\psi}$, we get its eigenfunction expansion is that basis

$$\hat{I}\ket{\psi} = \sum_i \ket{i}\braket{i|\psi} = \sum_i \braket{i|\psi}\ket{i} = \ket{\psi} , \tag{1.29}$$

where we moved $\braket{i|\psi}$ to the front to make the expression more obvious — we are allowed to do this because $\braket{i|\psi}$ is just a number.

Having defined the identity operator, we can define the inverse of an operator $\hat{A}$, denoted $\hat{A}^{-1}$, to be the operator such that

$$\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{I} . \tag{1.30}$$

A unitary operator $\hat{U}$ is one for which the adjoint is equal to the inverse,

$$\hat{U}^{-1} = \hat{U}^\dagger , \tag{1.31}$$

so that

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I} . \tag{1.32}$$

As we claimed before, unitary operators are used to define transformations between different representations.[6] Therefore, if we have an operator $\hat{A}$, we would like to know what its representation has transformed into after the transformation. Suppose the operator $\hat{A}$ acts on a general wavefunction ψ to get

$$\hat{A}\psi = \phi , \tag{1.33}$$

and after the transformation, we have

$$\psi' = \hat{U}\psi , \qquad \phi' = \hat{U}\phi . \tag{1.34}$$

Then we would like to find $\hat{A}'$, which is the representation of $\hat{A}$ after the transformation, such that

$$\hat{A}'\psi' = \phi' . \tag{1.35}$$

We have

$$\hat{A}'\hat{U}\psi = \hat{U}\phi = \hat{U}\hat{A}\psi \implies \hat{A}'\hat{U} = \hat{U}\hat{A} . \tag{1.36}$$

Therefore,

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger \quad \text{or} \quad \hat{A} = \hat{U}^\dagger\hat{A}'\hat{U} . \tag{1.37}$$

By taking the adjoint of the above equation, we can see that a Hermitian operator after unitary transformation is still an Hermitian operator.

---

[6]Unitary operators are linked to basis transformations, just like unitary matrices are linked to basis transformations in vector spaces. We will show this later, when we transform a wavefunction from the position basis (wavefunction) to the momentum basis (the momentum representation of the wavefunction), and transform an orbital from the AO basis in to the SO basis.

Moreover, a transformed eigenfunction of an operator is still an eigenfunction of the transformed operator with the same eigenvalue. This means that if

$$\hat{A}\psi_n = a_n\psi_n\,, \tag{1.38}$$

then since we can insert an identity operator anywhere,

$$\hat{A}\hat{U}^\dagger\hat{U}\psi_n = a_n\hat{U}^\dagger\hat{U}\psi_n\,, \tag{1.39}$$

and so

$$\hat{U}\hat{A}\hat{U}^\dagger\hat{U}\psi_n = a_n\hat{U}\psi_n\,, \tag{1.40}$$

which implies

$$\hat{A}'\psi_n' = a_n\psi_n'\,. \tag{1.41}$$

We can also show that the expectation values are the same:

$$\begin{aligned}
\int \mathrm{d}\tau\,\psi_n^*\hat{A}\psi_n &= \int \mathrm{d}\tau\,\psi_n^*\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}\psi_n \\
&= \int \mathrm{d}\tau\,(\hat{U}\psi_n)^*\hat{U}\hat{A}\hat{U}^\dagger\hat{U}\psi_n \\
&= \int \mathrm{d}\tau'\,\psi_n'\hat{A}'\psi_n'\,.
\end{aligned} \tag{1.42}$$

If we let $\hat{A} = \hat{I}$, then

$$\int \mathrm{d}\tau'\,\psi_n'\psi_n = \int \mathrm{d}\tau\,\psi_n^*\psi_n\,, \tag{1.43}$$

so a normalised state remains normalised after transformation.[7]

1.4 The Momentum Space Representation

Consider the transformation, in one-dimensional space, defined by the unitary operator

$$\phi(p) = \hat{U}\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\,\psi(x)e^{-ipx}\,, \tag{1.44}$$

which you may recognise as the *Fourier transform*. This is a unitary transformation because its inverse transform is given by

$$\psi(x) = \hat{U}^\dagger\phi(p) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \phi(p)e^{ipx}\,, \tag{1.45}$$

which is exactly the adjoint of the forward transformation. This transforms our wavefunction $\phi(x)$ from the position basis to the momentum basis.

Before we proceed, we first introduce a useful result.

Proposition 1.1.

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}x\,e^{i(p-p')x} = \delta(p-p')\,, \tag{1.46}$$

where $\delta(x)$ is the *Dirac delta function*.[8]

---

[7]We are deliberately not using the bra-ket notation here because technically, ψ_n and ψ_n' corresponds to the same $|n\rangle$, just being represented in different basis. For example, if ψ_n is represented in the position (x) basis and ψ_n' is represented in the momentum (p) basis, then $\psi_n(x) = \langle x|n\rangle$ and $\psi_n'(p) = \langle p|n\rangle$, where $|x\rangle$ and $|p\rangle$ are the position and momentum eigenstates.

[8]Dirac delta function $\delta(x)$ is a weird 'function' defined such that $\delta(x) \neq 0\ \forall x \neq 0$ but for any $\epsilon > 0$, we have

$$\int_{-\epsilon}^{\epsilon} \mathrm{d}x\,\delta(x) = 0\,. \tag{1.47}$$

You can think of this as a normalised Gaussian in the limit of its variance $\to 0$.

Dirac delta has a nice property.

$$\int_{-\infty}^{\infty} dx \, f(x)\delta(x-x') = f(x') \,. \tag{1.48}$$

Now it's time to investigate the momentum space representation. Specifically, we would like to find out how the position and momentum operators look like in the momentum representation. We do this by rewriting the position-space expectation value of momentum into the momentum space, and we find

$$\begin{aligned}
\langle p \rangle &= \int_{-\infty}^{\infty} dx \, \psi^*(x) \left(-i\frac{\partial}{\partial x}\right) \psi(x) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \left[\int_{-\infty}^{\infty} dp \, \phi(p)^* e^{-ipx}\right] \left(-i\frac{\partial}{\partial x}\right) \left[\int_{-\infty}^{\infty} dp' \, \phi(p')e^{-ip'x}\right] \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dp' \, p' \phi(p)^* \phi(p') e^{i(p-p')x} \\
&= dx \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dp' \, p' \phi(p)^* \phi(p')\delta p' - p \\
&= \int_{-\infty}^{\infty} dp \, \phi(p)^* p\phi(p) \,.
\end{aligned} \tag{1.49}$$

This is exactly the momentum-space average of the operator p. We see that in the momentum representation, the momentum operator is unsurprisingly p itself — and if fact that is why we call this representation the momentum representation.

To work out the momentum representation of the position operator, we do the same trick again.

$$\begin{aligned}
\langle x \rangle &= \int_{-\infty}^{\infty} dx \, \psi(x)^* x\psi(x) \\
&= \int_{-\infty}^{\infty} dx \left[\int_{-\infty}^{\infty} dp \, \phi(p)^* e^{-ipx}\right] x \left[\int_{-\infty}^{\infty} dp' \, \phi(p')e^{-ip'x}\right] \,.
\end{aligned} \tag{1.50}$$

Since

$$-i\frac{\partial}{\partial p}e^{ipx} = xe^{ipx} \,, \tag{1.51}$$

we can rewrite it as

$$\langle x \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \, \phi(p')^* \phi(p) \left(-i\frac{\partial}{\partial p}e^{i(p-p')x}\right) \,. \tag{1.52}$$

We now use integration by part, and we get

$$\begin{aligned}
\langle x \rangle &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \, \phi(p')^* e^{i(p-p')x} \left(i\frac{\partial}{\partial x}\phi(p)\right) \\
&= \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dp \, \delta(p-p')\phi(p')^* \left(i\frac{\partial}{\partial p}\phi(p)\right) \\
&= \int_{-\infty}^{\infty} dp \, \phi(p)^* \left(i\frac{\partial}{\partial p}\right) \phi(p) \,.
\end{aligned} \tag{1.53}$$

This is the expectation integral of the operator $i\partial/\partial p$.

Proposition 1.2. The momentum representation of the position and momentum operators are

$$\hat{x}' = \hat{U}\hat{x}\hat{U}^\dagger = i\frac{\partial}{\partial p} \tag{1.54}$$

$$\hat{p}' = \hat{U}\hat{p}\hat{U}^\dagger = p \,. \tag{1.55}$$

9

1.4.1 The State Vector

You can see that the even though we transformed a wavefunction from the position representation to the momentum representation, it is still fundamentally the same thing — both representations describe the same underlying quantum state. Although we may define various unitary transformations and trying to represent a quantum state in different representations, there must be a invariant underlying object. Just as we may choose different basis to describe a vector, and the vector may have different coordinates in different basis, it is still the same vector. We can do the same thing in quantum mechanics.

The true underlying quantum state is called a *state vector*, denoted $|\psi\rangle$, and it lives in a special vector space called a *Hilbert space*. The position representation of a wavefunction is just writing out the coefficients of a state vector in the position basis $\{|x\rangle\}$, and so $\psi(x) = \langle x|\psi\rangle$. Similarly the momentum representation is $\psi(p)' = \langle p|\psi\rangle$. We can any other basis we want to represent a quantum state, but it should better be orthonormal and it should be complete, meaning that the basis should span the whole Hilbert space so that we can represent any quantum state we want in such a basis — the eigenstates of a Hermitian operator seems perfect for this purpose. Then a quantum states is defined by the countably infinite number of coefficients of each basis vector. This leads to the *vectors and matrices representations*.

1.5 Vectors and Matrices Representations

If we have a quantum state $|\psi\rangle$ and a complete, orthogonal basis set $\{|\phi_i\rangle\}$, then we can expand and write

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle \,, \tag{1.56}$$

where as we claimed before, if the basis functions $|\phi_i\rangle$ are orthonormal, then the coefficients c_i are given by

$$c_i = \langle \phi_i|\psi\rangle \,. \tag{1.57}$$

This expansion be rewritten in matrix notation as

$$|\psi\rangle = \begin{pmatrix} |\phi_1\rangle & |\phi_2\rangle & \cdots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} = \boldsymbol{\phi}^{\mathrm{T}} \mathbf{c} \,, \tag{1.58}$$

where $\boldsymbol{\phi}$ is the column vector of the basis set and $\mathbf{c}$ is the column vector of coefficients.

Now, let's do a basis transformation. We define a new basis set $|\theta_j\rangle = \sum_i |\phi_i\rangle U_{ij}$ for some set of coefficients $U_{ij} = \langle \phi_i|\theta_j\rangle$, which can be written in the matrix form as

$$\begin{pmatrix} |\theta_1\rangle & |\theta_2\rangle & \cdots \end{pmatrix} = \begin{pmatrix} |\phi_1\rangle & |\phi_2\rangle & \cdots \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & \cdots \\ U_{21} & U_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \,, \tag{1.59}$$

or

$$\boldsymbol{\theta}^{\mathrm{T}} = \boldsymbol{\phi}^{\mathrm{T}} \mathsf{U} \,. \tag{1.60}$$

If we want to keep the new basis orthonormal, then we must have

$$\begin{aligned} \delta_{ij} &= \langle \theta_i|\theta_j\rangle \\ &= \sum_{k,l} U_{ki}^* U_{lj} \langle \phi_k|\phi_l\rangle \\ &= \sum_k U_{ki}^* U_{kj} \,, \end{aligned} \tag{1.61}$$

which implies

$$U^\dagger U = I, \tag{1.62}$$

and so U is a unitary matrix. The transformation between orthonormal bases is given by a unitary matrix.

We hope to represent the same $|\psi\rangle$ in this new transformed basis $\{|\theta_i\rangle\}$

$$|\psi\rangle = \boldsymbol{\theta}^\mathrm{T}\mathbf{d}. \tag{1.63}$$

Then clearly, we should have

$$|\psi\rangle = \boldsymbol{\theta}^\mathrm{T}\mathbf{d} = \boldsymbol{\phi}^\mathrm{T}U\mathbf{d} = \boldsymbol{\phi}^\mathrm{T}\mathbf{c}, \tag{1.64}$$

so

$$\mathbf{c} = U\mathbf{d} \qquad \text{or} \qquad \mathbf{d} = U^\dagger\mathbf{c}. \tag{1.65}$$

The components transform in the opposite way as the basis.

The same thing goes for operators. If a Hermitian operator $\hat{A}$ is represented in the $|\phi_i\rangle$ basis with matrix elements $A_{ij} = \left\langle \phi_i \middle| \hat{A} \middle| \phi_j \right\rangle$, then in the new basis $|\theta_i\rangle$, the matrix element A' will be given by

$$\begin{aligned}
A'_{ij} &= \left\langle \theta_i \middle| \hat{A} \middle| \theta_j \right\rangle \\
&= \sum_{k,l} U^*_{ki} U_{lj} \left\langle \phi_k \middle| \hat{A} \middle| \phi_l \right\rangle \\
&= \sum_{k,l} U^*_{ki} A_{kl} U_{lj},
\end{aligned} \tag{1.66}$$

and so

$$A' = U^\dagger A U. \tag{1.67}$$

1.5.1 Symmetry Orbitals

The above transformation is exactly what we are doing when we transform from the atomic orbital basis (denoted $|\phi_i\rangle$) to symmetry orbitals (denoted θ_i) in Hückel theory. If we have figured out the symmetry orbitals to be $|\theta_j\rangle = \sum_i |\phi_i\rangle U_{ij}$ and the Hamiltonian in the atomic orbital basis to be $H_{ij} = \left\langle \phi_i \middle| \hat{H} \middle| \phi_j \right\rangle$, then the Hamiltonian in the symmetry orbital basis set is

$$H' = U^\dagger H U. \tag{1.68}$$

In practise we don't usually do this by matrix multiplication because we already know that a lot of the matrix elements will be zero, so we would rather work out H'_{ij} element by element. The goal of Hückel theory is usually to work out the Hückel molecular orbitals as well as their energies, which are the eigenvectors and eigenvalues of H. By doing a unitary basis transformation, the eigenvalues of H' should be the same as H, and the eigenvectors should be the same ones but represented in the symmetry orbital basis. The point is that the block diagonal form of H' makes them easier to find.

1.6 Uncertainty Principle

We have asserted that the uncertainties of two observables are related to the expectation value of their commutator. It is time to see where the uncertainty principle comes from.

In statistics, the variance of a random variable A is given by

$$\sigma_A^2 = \left\langle A^2 \right\rangle - \left\langle A \right\rangle^2 = \left\langle (A - \langle A \rangle)^2 \right\rangle . \tag{1.69}$$

This can be translated into the quantum mechanical expression of the uncertainty of a physical observable

$$\Delta A^2 = \left\langle \psi \left| \left(\hat{A} - \left\langle \hat{A} \right\rangle \right)^2 \right| \psi \right\rangle . \tag{1.70}$$

If $\hat{A}$ is a Hermitian operator, then we can write this as

$$\Delta A^2 = \left\langle \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \psi \left| \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \psi \right\rangle . \tag{1.71}$$

If we denote $\psi_A = \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \psi$, then ΔA^2 is exactly the norm of ψ_A:

$$\Delta A^2 = \langle \psi_A | \psi_A \rangle . \tag{1.72}$$

Suppose we have another observable with operator $\hat{B}$, then if $\psi_B = \left(\hat{B} - \left\langle \hat{B} \right\rangle \right) \psi$, we have

$$\Delta B^2 = \langle \psi_B | \psi_B \rangle \tag{1.73}$$

by exactly the same argument. The product of the two variances is

$$\Delta A^2 \Delta B^2 = \langle \psi_A | \psi_A \rangle \langle \psi_B | \psi_B \rangle . \tag{1.74}$$

Next, we will use Cauchy–Schwarz inequality.

Lemma 1.3 (Cauchy–Schwarz inequality). Let V be a vector space with an inner product $\cdot$, then for all $\mathbf{u}, \mathbf{v} \in V$,

$$|\mathbf{u} \cdot \mathbf{v}|^2 \leq (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) . \tag{1.75}$$

Therefore, we have

$$\Delta A^2 \Delta B^2 \geq |\langle \psi_A | \psi_B \rangle|^2 . \tag{1.76}$$

For any complex number $z \in \mathbb{C}$, we have

$$\left| z^2 \right| = \mathrm{Re}(z)^2 + \mathrm{Im}(z^2) = \left(\frac{z + z^*}{2} \right)^2 + \left(\frac{z - z^*}{2i} \right)^2 , \tag{1.77}$$

and since $\langle \psi_A | \psi_B \rangle = \langle \psi_B | \psi_A \rangle^*$, we have

$$\Delta A^2 \Delta B^2 \geq \left(\frac{\langle \psi_A | \psi_B \rangle + \langle \psi_B | \psi_A \rangle}{2} \right)^2 + \left(\frac{\langle \psi_A | \psi_B \rangle - \langle \psi_B | \psi_A \rangle}{2i} \right)^2 . \tag{1.78}$$

To simplify things up, we need to work out $\langle \psi_A | \psi_B \rangle$ and $\langle \psi_B | \psi_A \rangle$.

$$\begin{aligned}
\langle \psi_A | \psi_B \rangle &= \left\langle \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \psi \left| \left(\hat{B} - \left\langle \hat{B} \right\rangle \right) \psi \right\rangle \\
&= \left\langle \psi \left| \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \left(\hat{B} - \left\langle \hat{B} \right\rangle \right) \right| \psi \right\rangle \\
&= \left\langle \psi \left| \hat{A}\hat{B} - \hat{A} \left\langle \hat{B} \right\rangle - \hat{B} \left\langle \hat{A} \right\rangle + \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle \right| \psi \right\rangle \\
&= \left\langle \hat{A}\hat{B} \right\rangle - \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle .
\end{aligned} \tag{1.79}$$

Similarly

$$\langle \psi_B | \psi_A \rangle = \left\langle \hat{B}\hat{A} \right\rangle - \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle . \tag{1.80}$$

Defining the *anti-commutator*

$$\{\hat{A}, \hat{B}\} \coloneqq \hat{A}\hat{B} + \hat{B}\hat{A}, \tag{1.81}$$

we have

$$\langle \psi_A | \psi_B \rangle + \langle \psi_B | \psi_A \rangle = \left\langle \{\hat{A}, \hat{B}\} \right\rangle - 2 \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle, \tag{1.82}$$

$$\langle \psi_A | \psi_B \rangle + \langle \psi_B | \psi_A \rangle = \left\langle [\hat{A}, \hat{B}] \right\rangle. \tag{1.83}$$

This gives what is known as the Robertson–Schrödinger uncertainty principle.

Theorem 1.4 (Robertson–Schrödinger uncertainty principle). For two observables A and B,

$$\Delta A \Delta B \geq \sqrt{\left(\frac{1}{2}\left\langle \{\hat{A}, \hat{B}\} \right\rangle - \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle\right)^2 + \left(\frac{1}{2i}\left\langle [\hat{A}, \hat{B}] \right\rangle\right)^2}. \tag{1.84}$$

This version of the uncertainty principle is very general, but it simplifies in the special case of uncorrelated operators. The first parenthesised quantity in (1.84) is

$$\mathrm{cov}(\hat{A}, \hat{B}) \coloneqq \frac{1}{2}\left\langle \hat{A}\hat{B} - \hat{B}\hat{A} \right\rangle - \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle, \tag{1.85}$$

which is known as the *covariance* of the operators. The corresponding quantity in statistics is

$$\mathrm{cov}(X, Y) \coloneqq \langle XY \rangle - \langle X \rangle \langle Y \rangle, \tag{1.86}$$

but in quantum mechanics we have to be a little bit more careful because in general $\left\langle \hat{A}\hat{B} \right\rangle \neq \left\langle \hat{B}\hat{A} \right\rangle$. The covariance is a measure of the correlation between operators, and is 0 if the operators are uncorrelated. If this is the case, then (1.84) reduces to the more familiar Robertson uncertainty principle.

Theorem 1.5 (Robertson Uncertainty Principle). If A and B are two uncorrelated physical observables, then

$$\Delta A \Delta B \geq \frac{1}{2}\left| \left\langle [\hat{A}, \hat{B}] \right\rangle \right|. \tag{1.87}$$

Note that if the two operators are correlated, then the Robertson–Schrödinger lower bound is larger than the Robertson lower bond, in other words the Robertson–Schrödinger provides a stronger condition.

2 Normal Modes

In Part IB Chemistry A, we discussed the use of symmetry and group theory to make qualitative conclusions about normal modes.

Example. For example, consider the $C - H$ stretching modes in benzene, with point group D_{6h}.
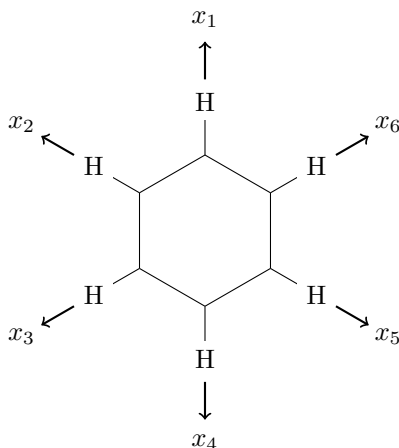

Figure 1: The basis set for representing the $C - H$ stretches in benzene.

Reducing the representation formed by these basis vectors, we get

$$\Gamma^{C-H} = A_{1g} \oplus E_{2g} \oplus B_{1u} \oplus E_{1u}. \tag{2.1}$$

Since there is only one combination for each irreducible representation, it should be easy to determine the symmetry adapted linear combinations of displacements along the $C - H$ bonds that transform according to the symmetries. Based on the Cartesian functions listed in the character table, intuition, or the projection operators, we can determine the normal modes to be

$$Q_{A_{1g}} = \frac{1}{\sqrt{6}}(x_1 + x_2 + x_3 + x_4 + x_5 + x_6) \tag{2.2}$$

$$Q_{E_{2g},1} = \frac{1}{2}(x_2 - x_3 + x_5 - x_6) \tag{2.3}$$

$$Q_{E_{2g},2} = \frac{1}{\sqrt{12}}(2x_1 - x_2 - x_3 + 2x_4 - x_5 - x_6) \tag{2.4}$$

$$Q_{E_{1u},1} = \frac{1}{\sqrt{12}}(2x_1 - x_2 - x_3 - 2x_4 + x_5 + x_6) \tag{2.5}$$

$$Q_{E_{1u},2} = \frac{1}{2}(x_2 + x_3 - x_5 - x_6) \tag{2.6}$$

$$Q_{B_{1u}} = \frac{1}{\sqrt{6}}(x_1 - x_2 + x_3 - x_4 + x_5 - x_6). \tag{2.7}$$

This method is nice and easy, especially for simple cases like this with restricted basis sets. But it does leave unanswered questions. What if we have more than one symmetry adapted combination transforming as a particular IR? These symmetry adapted combinations of displacement vectors are like symmetry orbitals, and some process analogous to combining symmetry orbitals to form molecular orbitals should be indicated. Most of the combinations we have here are only approximations (albeit useful ones) to normal modes, as we can see a lot of them would move the centre of mass. Also, it is impossible that only H atoms are moving during the vibrations — we also need the motion of C atoms as well as displacements in other directions. In such a complete basis set, how to correctly combine

the symmetry adapted displacement vectors to form the true normal modes would be a problem. Additionally, some indication of the frequencies would be useful.

To do this, we have to treat the mechanical problem of molecular vibration seriously.

2.1 Diatomic Vibration

We being by looking at the vibration of a diatomic molecule, and hopefully this can give us some insight on treating the general problem. Each atom in a diatomics can move in three independent directions (x, y and z), so there are in total six possible independent motions. We know that three of them are translations and rotations, leaving us with only one vibrational mode. It turns out that only the motion along the intermolecular axis is useful for vibration. The label the positions of the atoms relative to their equilibrium positions, we define two basis vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. Then the configuration of the atom can be labelled by a vector

$$\mathbf{v} = x_1 \mathbf{x}_1 + x_2 \mathbf{x}_2 \,, \tag{2.8}$$

where x_1 and x_2 are the components, meaning that atom 1 has displaced by length x_1 along the direction $\mathbf{x}_1$, and atom 2 has displaced by x_2 along $\mathbf{x}_2$. We can put the components into a single column vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.
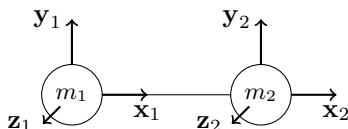

Figure 2: A basis set representing the motion of a diatomic molecule.

The simplest model for the internuclear potential is the harmonic oscillator, with the potential given by

$$V_{\text{HO}} = \frac{1}{2} k (x_1 - x_2)^2 \,, \tag{2.9}$$

where k is the force constant. We can write the Hamiltonian as[9]

$$\hat{H}(\mathbf{x}) = -\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2} + \frac{1}{2} k (x_1 - x_2)^2 \,, \tag{2.10}$$

where the first two terms are the kinetic energies of the two particles with masses of m_1 and m_2 respectively. Now this is a non-separable Hamiltonian — we have cross terms $x_1 x_2$ in the potential energy, so that we can not write the Hamiltonian as $H(x_1) + H(x_2)$. We would not like to solve this kind of partial differential equations directly.

To make progress, we want to try some coordinate transformation that would make the cross terms vanishes. The transformation

$$\begin{aligned} w_1 &= \frac{1}{\sqrt{2}} (x_1 - x_2) \\ w_2 &= \frac{1}{\sqrt{2}} (x_1 + x_2) \end{aligned} \tag{2.11}$$

of the components seems promising, since it eliminates the cross terms in the potential energy by making it into a single term kw_1^2. Moreover, these two coordinates have solid physical meaning — w_1 is the compression/extension along the internuclear axis, which is exactly the vibrational motion

---

[9]We treat this problem using Hamiltonian. If you do Part IB Mathematics, you will solve this type of problems using Lagrangian — you will find that these two methods are closely related.

we are trying to model, and w_2 describes the translation along the internuclear axis, one of the other motion of the molecules. This is described by the unitary transformation matrix

$$\mathsf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{2.12}$$

generated by putting the components of the transformed components in terms of the original components in rows,[10] so that the components transforms as desired:

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \mathsf{U} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{2.13}$$

or

$$\mathbf{w} = \mathsf{U}\mathbf{x}. \tag{2.14}$$

We can be satisfied that no physical properties will be altered by this unitary transformation.

However, we will see that the problem of this basis transformation is that although the potential terms is nicely simplified, the kinetic term is screwed up. From chain rule, we have

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \frac{\partial w_1}{\partial x_1} + \frac{\partial}{\partial w_1} + \frac{\partial w_2}{\partial x_1} \frac{\partial}{\partial w_2} \\ &= \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial w_1} - \frac{\partial}{\partial w_2} \right), \end{aligned} \tag{2.15}$$

$$\begin{aligned} \frac{\partial}{\partial x_2} &= \frac{\partial w_1}{\partial x_2} + \frac{\partial}{\partial w_1} + \frac{\partial w_2}{\partial x_2} \frac{\partial}{\partial w_2} \\ &= \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial w_1} + \frac{\partial}{\partial w_2} \right). \end{aligned} \tag{2.16}$$

Therefore in this transformed basis, the Hamiltonian is

$$\hat{H}(\mathbf{w}) = -\frac{\hbar^2}{4m_1} \left(\frac{\partial}{\partial w_1} - \frac{\partial}{\partial w_2} \right)^2 - \frac{\hbar^2}{4m_2} \left(\frac{\partial}{\partial w_1} + \frac{\partial}{\partial w_2} \right)^2 + kw_1^2. \tag{2.17}$$

The kinetic part contains a mixed partial derivative

$$\frac{\hbar^2}{2m_1} \left(\frac{1}{m_1} - \frac{1}{m_2} \right) \frac{\partial^2}{\partial w_1 w_2^2}, \tag{2.18}$$

which is non-vanishing for $m_1 \neq m_2$.

It seems that by doing a unitary transformation, we can only separate one of the kinetic and the potential part. Whenever we separate one of the terms, the other will necessarily be screwed up.

However, the above trial did provide us some inspiration. The mixed partial derivative term do vanish if the masses of the two particles are equal. What if we do some scaling to the coordinates based on the masses? In turns out that for this particular question, the useful scaling is

$$Q_1 = \sqrt{\frac{m_1 m_2}{m_1 + m_2}} (x_1 - x_2) \tag{2.19}$$

$$Q_2 = \frac{1}{\sqrt{m_1 + m_2}} (m_1 x_1 + m_2 x_2) \tag{2.20}$$

We will explain how we obtained this later, but now let's try this out. Again, the potential is straightforward to rewrite in the **Q** basis, as

$$V_{\mathrm{HO}} = \frac{k(m_1 + m_2)}{2m_1 m_2} Q_1^2. \tag{2.21}$$

---

[10]Then the basis vector would transform as $\mathsf{U}^\dagger$. Remember that the basis vectors and the components transform in the inverse way.

For the kinetic energies, we again need the chain rule, and we obtain

$$\frac{\partial}{\partial x_1} = \frac{1}{\sqrt{m_1 + m_2}}\left(\sqrt{m_1 m_2}\frac{\partial}{\partial Q_1} + m_1\frac{\partial}{\partial Q_2}\right), \tag{2.22}$$

$$\frac{\partial}{\partial x_1} = \frac{1}{\sqrt{m_1 + m_2}}\left(-\sqrt{m_1 m_2}\frac{\partial}{\partial Q_1} + m_2\frac{\partial}{\partial Q_2}\right). \tag{2.23}$$

If we substitute this into the expression of the Hamiltonian, we will see that all the mixed derivatives magically cancels out and we get

$$\hat{H}(Q) = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_1^2} + \frac{k(m_1 + m_2)}{2m_1 m_2}Q_1^2 - \frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_2^2}. \tag{2.24}$$

This can be nicely separated into

$$\hat{H}(\mathbf{Q}) = \hat{H}(Q_1) + \hat{H}(Q_2), \tag{2.25}$$

where

$$\hat{H}(Q_1) = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_1^2} + \frac{k(m_1 + m_2)}{2m_1 m_2}Q_1^2, \tag{2.26}$$

$$\hat{H}(Q_2) = -\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_2^2}. \tag{2.27}$$

If we compare (2.26) with the Hamiltonian of a canonical harmonic oscillator, we see that this is an harmonic oscillator with a mass 1 and a modified force constant of

$$k' = \frac{k(m_1 + m_2)}{m_1 m_2} =: \frac{k}{\mu}, \tag{2.28}$$

where we defined the familiar reduced mass $\mu := m_1 m_2/(m_1 + m_2)$. This gives an angular frequency

$$\omega = \sqrt{\frac{k'}{1}} = \sqrt{\frac{k}{\mu}}. \tag{2.29}$$

This is a familiar result. The vibration of a diatomic molecule is the same as the vibration of a single molecule with reduced mass of the system. The second equation (2.27) has only a kinetic term. This is the translation of the whole molecule along the internuclear axis — it can also be thought of as a harmonic oscillator with zero force constant, and so a zero frequency.

2.1.1 Mass-Weighted Coordinates

Now let's consider what is going on. The transformation we proposed can be expressed by the matrix

$$\mathsf{A} = \frac{1}{\sqrt{m_1 + m_2}}\begin{pmatrix} \sqrt{m_1 m_2} & -\sqrt{m_1 m_2} \\ m_1 & m_2 \end{pmatrix}. \tag{2.30}$$

You may verify that this is not a unitary matrix — the easiest way to see this is by checking its determinant. Therefore, physical observable do not have to preserved by this transformation. For example, we can see that all masses has reduced to 1, and the actual masses has somehow been taken into the modified force constant. However, the observable we really care about, the (angular) frequency of the oscillator ω, is related to the ratio of these, and this is unchanged by our transformation.

To see what it really going on in our transformation, we can split our transformation into two phases. The first phase is to mass-weight the coordinates, defining $q_i = \sqrt{m_1}x_i$, which is a diagonal transform, but is not unitary. This is represented by the diagonal matrix

$$\mathsf{B} = \begin{pmatrix} \sqrt{m_1} & 0 \\ 0 & \sqrt{m_2} \end{pmatrix}. \tag{2.31}$$

The Hamiltonian in this mass-weighted basis set is

$$\hat{H}(\mathbf{q}) = -\frac{\hbar^2}{2}\left(\frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2}\right) + \frac{k}{2}\left(\frac{q_1}{\sqrt{m_1}} - \frac{q_2}{\sqrt{m_2}}\right)^2. \tag{2.32}$$

This transformation does not remove the cross term in the potential, but it has reduced all the kinetic terms to a nice and simple form where all the masses are unity. This is exactly the case that we wanted before — if the masses are equal, then when you do the unitary transformation to simplify the potential term, all the cross terms in the kinetic term after the transformation will automatically cancel out! We are then ready to do our second transformation defined by

$$Q_1 = \frac{1}{\sqrt{m_1 + m_2}}\left(\sqrt{m_2}q_1 - \sqrt{m_1}q_2\right), \tag{2.33}$$

$$Q_2 = \frac{1}{\sqrt{m_1 + m_2}}\left(\sqrt{m_1}q_1 + \sqrt{m_2}q_2\right), \tag{2.34}$$

which can be represented by the unitary matrix

$$\mathsf{C} = \frac{1}{\sqrt{m_1 + m_2}}\begin{pmatrix}\sqrt{m_2} & -\sqrt{m_1} \\ \sqrt{m_1} & \sqrt{m_2}\end{pmatrix}. \tag{2.35}$$

The combination of these two transformations is the one we claimed before:

$$\mathsf{A} = \mathsf{CB}. \tag{2.36}$$

2.2 Quadratic Form

Before investigating more complicated cases, you may find that it is extremely useful to formulate the above process in matrices. Hence, we may rewrite the Hamiltonian in the following matrix form

$$\hat{H}(\mathbf{x}) = \frac{\hbar^2}{2}\begin{pmatrix}\frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2}\end{pmatrix}\begin{pmatrix}\frac{1}{m_1} & 0 \\ 0 & \frac{1}{m_2}\end{pmatrix}\begin{pmatrix}\frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2}\end{pmatrix} + \frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}k & -k \\ -k & k\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix} \tag{2.37}$$

$$=: -\frac{\hbar^2}{2}\dot{\mathbf{x}}^{\mathrm{T}}\mathsf{T}_x\dot{\mathbf{x}} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathsf{H}_x\mathbf{x}. \tag{2.38}$$

The subscripts x here stands for the x basis, and H is confusingly the canonical notation for the Hessian matrix which unfortunately happens to share its notation with the Hamiltonian.

Now we can do the transformations. The first step is to mass-scale the coordinates by

$$\mathbf{q} = \mathsf{B}\mathbf{x}, \tag{2.39}$$

where

$$\mathsf{B} = \begin{pmatrix}\sqrt{m_1} & 0 \\ 0 & \sqrt{m_2}\end{pmatrix}. \tag{2.40}$$

Note that by chain rule, the derivatives transform in the opposite way as the coordinates, so

$$\dot{\mathbf{q}} = \mathsf{B}^{-1}\dot{\mathbf{x}}, \tag{2.41}$$

and so

$$\dot{\mathbf{x}}^{\mathrm{T}}\mathsf{T}_x\dot{\mathbf{x}} = \dot{\mathbf{q}}^{\mathrm{T}}\mathsf{B}^{\mathrm{T}}\mathsf{T}_x\mathsf{B}\dot{\mathbf{q}} =: \dot{\mathbf{q}}^{\mathrm{T}}\mathsf{T}_q\dot{\mathbf{q}}. \tag{2.42}$$

This transformation is designed to reduce the T matrix to the identity matrix in this basis:

$$\mathsf{T}_q := \mathsf{B}^{\mathrm{T}}\mathsf{T}_x\mathsf{B} = \mathsf{I}, \tag{2.43}$$

and so

$$\dot{\mathbf{x}}^{\mathrm{T}}\mathsf{T}_x\dot{\mathbf{x}} = \dot{\mathbf{q}}^{\mathrm{T}}\mathsf{I}\dot{\mathbf{q}} = \dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{q}}. \tag{2.44}$$

18

Similarly for the Hessian matrix,

$$\mathbf{x}^{\mathrm{T}}\mathsf{H}_x\mathbf{x} = \mathbf{q}^{\mathrm{T}}(\mathsf{B}^{-1})^{\mathrm{T}}\mathsf{H}_x\mathsf{B}^{-1}\mathbf{q} =: \mathbf{q}^{\mathrm{T}}\mathsf{K}\mathbf{q}\,, \tag{2.45}$$

where the Hessian matrix in this mass-scaled coordinate is called the *dynamical matrix* and is given the special notation K. We can calculate this as

$$\mathsf{K} := (\mathsf{B}^{-1})^{\mathrm{T}}\mathsf{H}_x\mathsf{B}^{-1} = \begin{pmatrix} \frac{k}{m_1} & -\frac{k}{\sqrt{m_1 m_2}} \\ -\frac{k}{\sqrt{m_1 m_2}} & \frac{k}{m_2} \end{pmatrix}. \tag{2.46}$$

Similarly, we then do the unitary transformation that diagonalises K as well, and we finally get

$$\mathsf{T}_Q = \mathsf{I}\,, \ \ \mathsf{H}_Q = \begin{pmatrix} 0 & 0 \\ 0 & \frac{k}{\mu} \end{pmatrix} \tag{2.47}$$

in the $\mathbf{Q}$ basis. Since both matrices are diagonal, we are then able to separate the Hamiltonian.

2.3 The General Case

Now we are ready to generalise the above method for all molecules. Consider a molecule consisting of M atoms. The position of each atom is described by a displacement vector $\mathbf{x}_i = (x_i, y_i, z_i)$ $(1 \leq i \leq N)$ from its equilibrium position. We can compactly string them together as a $3N$-dimensional vector

$$\mathbf{x} = (x_1, y_1, z_1, \ldots, x_N, y_N, z_N)\,. \tag{2.48}$$

A general Hamiltonian can be written as

$$\hat{H} = -\frac{\hbar^2}{2}\sum_i^{3N}\frac{1}{m_i}\frac{\partial^2}{\partial x_i^2} + \frac{1}{2}\sum_i^{3N}\sum_{j>1} k_{ij}(x_i - x_j)^2\,. \tag{2.49}$$

Note that we sum over $i > j$ in the potential term to avoid double counting, and we allow a potential term to rise between any two coordinates, although in practise we can set most of them to be zero. We can write this into the matrix form introduced above

$$\hat{H} = -\frac{\hbar^2}{2}\dot{\mathbf{x}}^{\mathrm{T}}\mathsf{T}_x\dot{\mathbf{x}} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathsf{H}_x\mathbf{x}\,, \tag{2.50}$$

where T is a diagonal matrix whose elements are the reciprocal masses

$$T_{x,ii} = \frac{1}{m_i}\,, \tag{2.51}$$

and the Hessian matrix elements are

$$H_{x,ij} = \frac{\partial^2 V}{\partial x_i \partial x_j} = \begin{cases} \sum_n k_{in} & \text{if } i = j \\ -k_{ij} & \text{if } i \neq j \end{cases}. \tag{2.52}$$

We first scale the Hamiltonian into the mass-weighted basis so that

$$\hat{H}(\mathbf{q}) = -\frac{\hbar^2}{2}\dot{\mathbf{q}}^{\mathrm{T}}\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^{\mathrm{T}}\mathsf{K}\mathbf{q}\,. \tag{2.53}$$

The matrix T_q in the mass-weighted basis is the identity matrix so we have omitted it, and the dynamical matrix is simply

$$K_{ij} = \frac{1}{\sqrt{m_i m_j}}\mathsf{H}_{x,ij}\,. \tag{2.54}$$

19

We now need to transform the basis such that the K is made diagonal. We can always do this because K is a real symmetric matrix, which is always diagonalisable by a orthogonal matrix (a real unitary matrix). This transformation is guaranteed to keep T matrix still diagonalised because the identity matrix is always the identity matrix in any basis. This is done by the orthogonal matrix C such that

$$\mathsf{H}_Q = \mathsf{C}^\dagger \mathsf{K} \mathsf{C} \,, \tag{2.55}$$

where each column of C is the normalised eigenvector of K, and H_Q is a diagonal matrix whose elements are eigenvalues $\{\lambda_i\}$ of K in corresponding order. Finally in this basis, we have

$$\hat{H}(\mathbf{Q}) = -\frac{\hbar^2}{2}\dot{\mathbf{Q}}^{\mathrm{T}}\dot{\mathbf{Q}} + \frac{1}{2}\mathbf{Q}^{\mathrm{T}}\mathsf{H}_Q\mathbf{Q} \tag{2.56}$$

$$= \sum_i^{3N}\left(-\frac{\hbar^2}{2}\frac{\partial^2}{\partial Q_i^2} + \frac{1}{2}\lambda_1 Q_i^2\right) \tag{2.57}$$

separated into $3N$ one-dimensional harmonic oscillator with modified force constants $\{\lambda_i\}$. The angular frequencies of the modes are therefore

$$\omega_i = \sqrt{\lambda_i}\,. \tag{2.58}$$

The eigenvectors of K are the normal modes in terms of the $\mathbf{q}$ coordinates, which can be transformed into the $\mathbf{x}$ coordinates. Note that although the normal modes are orthogonal in the $\mathbf{q}$ coordinates, they are not necessarily so in the $\mathbf{x}$ coordinates since the mass scaling is not unitary. However, we may still often find them to be orthogonal, as they must be so if they transform according to different irreducible representations.

Note that since we used all $3N$ displacement vectors, there will be 6 (or 5) normal modes corresponding to translations or rotations. If we use a restricted basis set, there may be less of them. They can be easily identified as modes with frequencies zero.

2.4 Example: AAA Linear Molecule

As an simple example, let's look at the AAA linear molecule of three identical atoms connected by two bonds of equal strengths. An example of this would be the azide anion. For simplicity, we only consider the motions along the internuclear axis.
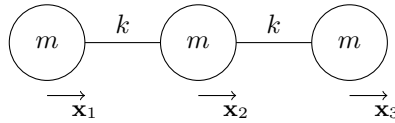

Figure 3: An AAA molecule.

The potential can be modelled as

$$V = \frac{k}{2}(x_1 - x_2)^2 + \frac{k}{2}(x_2 - x_3)^2\,. \tag{2.59}$$

The Hessian matrix can be found by either taking the partial derivative or by expanding the above function.

$$\mathsf{H}_x = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}\,. \tag{2.60}$$

The dynamical matrix is

$$\mathsf{K} = \begin{pmatrix} \frac{k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & \frac{2k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & \frac{k}{m} \end{pmatrix}\,. \tag{2.61}$$

We then simply need to diagonalise this matrix. The eigenvalues and the corresponding normal mode frequencies are

$$\lambda_1 = 0 \qquad\qquad \lambda_2 = \frac{k}{m} \qquad\qquad \lambda_3 = \frac{3k}{m} \qquad\qquad (2.62)$$

$$\omega_1 = 0 \qquad\qquad \omega_2 = \sqrt{\frac{k}{m}} \qquad\qquad \omega_3 = \sqrt{\frac{3k}{m}}. \qquad\qquad (2.63)$$

We can also find the normalised eigenvectors, which give the normal mode coordinates in the $\mathbf{q}$ basis (which is also the normal coordinates in the $\mathbf{x}$ basis since all atoms have equal masses)

$$Q_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \mathbf{q} \qquad\qquad Q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \mathbf{q} \qquad\qquad Q_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \mathbf{q}. \qquad (2.64)$$

In mode 1, all atoms move in the same direction with the same magnitude. This is the translation. Mode 2 is the symmetric stretch, in which both bonds are simultaneously stretched/compressed, and mode 3 is the antisymmetric stretch, in which when one bond is stretched, the other is compressed.
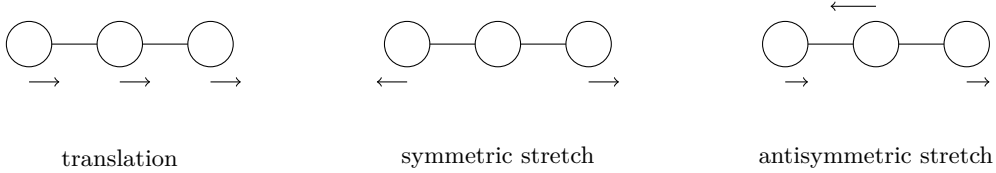

Figure 4: Normal modes of an AAA molecule.

2.5   The Use of Symmetry

The dynamical matrix in the above example is 3×3, which we can easily diagonalise directly. For larger molecules, we will encounter the same issue that we had when constructing MOs: any bigger matrix will be too tedious to diagonalise directly by have. The solution is the same too: we can use an intermediate symmetry-adapted basis set to transform the dynamical matrix to a block diagonal form.

We will summarise the process of constructing normal modes by symmetry as well as molecular orbitals in the table below.

2.6   Example: ABBA Linear Molecule

Consider the motions along the bond for a centrosymmetric ABBA linear molecule, for example acetylene. The masses of the two types of atoms are now different, and we also allow the strength of the A $-$ B bond to be different from the B $-$ B bond.
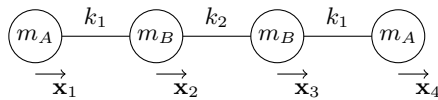

Figure 5: A centrosymmetric ABBA linear molecule.

MOs	Normal modes
1. Write down the Hamiltonian matrix in the AO basis.	1. Write down the dynamical matrix in the mass-weighted basis.
2. Determine the IRs spanned by the AO basis set.	2. Determine IRs spanned by the mass-weighted basis set.
3. Construct symmetry adapted combinations of the AOs.	3. Construct symmetry adapted combinations of the mass-weighted coordinates.
4. Transform the Hamiltonian matrix into the symmetry-adapted basis set.	4. Transform the dynamical matrix into the symmetry-adapted basis set.
5. The eigenvalues are the orbital energies and the eigenvectors are the MO coefficients in the symmetry-adapted basis set.	5. The eigenvalues are the frequencies squared and the eigenvectors are the normal mode coordinates in the symmetry-adapted basis set.
6. The eigenvectors can be transformed back to the AO basis set.	6. The eigenvectors can be transformed back to the coordinate displacement basis set.

If we proceed without using symmetry, then we have a 4×4 dynamical matrix to diagonalise

$$\mathsf{K} = \begin{pmatrix} \frac{k_1}{m_A} & -\frac{k_1}{\sqrt{m_A m_B}} & 0 & 0 \\ -\frac{k_1}{\sqrt{m_A m_B}} & \frac{k_1+k_2}{m_B} & -\frac{k_2}{m_B} & 0 \\ 0 & -\frac{k_2}{m_B} & \frac{k_1+k_2}{m_B} & -\frac{k_1}{\sqrt{m_A m_B}} \\ 0 & 0 & -\frac{k_1}{\sqrt{m_A m_B}} & \frac{k_1}{m_A} \end{pmatrix}. \tag{2.65}$$

To simplify things up, we will create a $\tilde{\mathbf{q}}$ basis from symmetry adapted combinations of $\mathbf{q}$. The point group of the molecule is $D_{\infty h}$, and we can split the basis into two two-dimensional sets $\{\mathbf{q}_1, \mathbf{q}_4\}$ and $\{\mathbf{q}_2, \mathbf{q}_3\}$. The irreducible representations and the symmetry adapted combinations of basis vectors should be easy to find[11]

$$\tilde{\mathbf{q}}_1 = \frac{1}{\sqrt{2}}(\mathbf{q}_1 - \mathbf{q}_4) \qquad\qquad \Sigma_g^+ \tag{2.66}$$

$$\tilde{\mathbf{q}}_2 = \frac{1}{\sqrt{2}}(\mathbf{q}_2 - \mathbf{q}_3) \qquad\qquad \Sigma_g^+ \tag{2.67}$$

$$\tilde{\mathbf{q}}_3 = \frac{1}{\sqrt{2}}(\mathbf{q}_1 + \mathbf{q}_4) \qquad\qquad \Sigma_u^+ \tag{2.68}$$

$$\tilde{\mathbf{q}}_4 = \frac{1}{\sqrt{2}}(\mathbf{q}_2 + \mathbf{q}_3) \qquad\qquad \Sigma_u^+. \tag{2.69}$$

This is the basis transformation represented by the unitary matrix

$$\tilde{\mathsf{C}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix} \tag{2.70}$$

so that the components transforms as

$$\tilde{\mathbf{q}} = \tilde{\mathsf{C}}^\dagger \mathbf{q}. \tag{2.71}$$

---

[11]Recall when constructing a representation, we transform the basis, not the components.

We must have

$$\mathbf{q}^\dagger \mathsf{K} \mathbf{q} = \tilde{\mathbf{q}}^\dagger \tilde{\mathsf{K}} \tilde{\mathbf{q}} = \mathbf{q}^\dagger \tilde{\mathsf{C}} \tilde{\mathsf{K}} \tilde{\mathsf{C}}^\dagger \mathbf{q}\,, \tag{2.72}$$

and so

$$\tilde{\mathsf{K}} = \tilde{\mathsf{C}} \mathsf{K} \tilde{\mathsf{C}} \tag{2.73}$$

$$= \begin{pmatrix} \frac{k_1}{m_A} & -\frac{k_1}{\sqrt{m_A m_B}} & 0 & 0 \\ -\frac{k_1}{\sqrt{m_A m_B}} & \frac{k_1 + 2k_2}{m_B} & 0 & 0 \\ 0 & 0 & \frac{k_1}{m_A} & -\frac{k_1}{\sqrt{m_A m_B}} \\ 0 & 0 & -\frac{k_1}{\sqrt{m_A m_B}} & \frac{k_1}{m_A} \end{pmatrix}\,. \tag{2.74}$$

This is block diagonal as expected. For simplicity, we will only work out the Σ_u^+ (bottom right) block. The eigenvalues are

$$\lambda_3 = 0 \qquad\qquad \lambda_4 = \frac{k_1(m_A + m_B)}{m_A m_B}\,, \tag{2.75}$$

and the eigenvectors in the various basis sets are

$$Q_3 = \frac{1}{\sqrt{m_A + m_B}} \begin{pmatrix} \sqrt{m_A} \\ \sqrt{m_B} \end{pmatrix} \cdot \tilde{\mathbf{q}}_{\Sigma_u^+} \qquad\qquad Q_4 = \frac{1}{\sqrt{m_A + m_B}} \begin{pmatrix} \sqrt{m_B} \\ -\sqrt{m_A} \end{pmatrix} \cdot \tilde{\mathbf{q}}_{\Sigma_u^+} \tag{2.76}$$

$$= \frac{1}{\sqrt{2(m_A + m_B)}} \begin{pmatrix} \sqrt{m_A} \\ \sqrt{m_B} \\ \sqrt{m_B} \\ \sqrt{m_A} \end{pmatrix} \cdot \mathbf{q} \qquad\qquad = \frac{1}{\sqrt{2(m_A + m_B)}} \begin{pmatrix} \sqrt{m_B} \\ -\sqrt{m_A} \\ -\sqrt{m_A} \\ \sqrt{m_B} \end{pmatrix} \cdot \mathbf{q} \tag{2.77}$$

$$= \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \cdot \mathbf{x} \qquad\qquad = \frac{\sqrt{m_A m_B}}{\sqrt{2(m_A^2 + m_B^2)}} \begin{pmatrix} \sqrt{\frac{m_B}{m_A}} \\ -\sqrt{\frac{m_A}{m_B}} \\ -\sqrt{\frac{m_A}{m_B}} \\ \sqrt{\frac{m_B}{m_A}} \end{pmatrix} \cdot \mathbf{x}\,. \tag{2.78}$$

Mode three is the uniform translation, and mode four is the antisymmetric stretch in which when one $A - B$ bond is stretched, the other is compressed.

translation   antisymmetric stretch
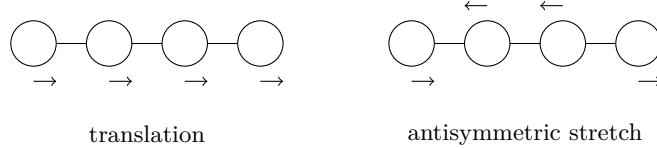
2.7 Beyond Harmonic Oscillators

Real molecules are not harmonic oscillators in general. We will have some complicated potential as a function of the nuclear coordinates, which defines the nuclear *potential energy surface* (PES). We can perform a Taylor expansion around an equilibrium position to get

$$V(\mathbf{R}) = V(\mathbf{R}_{\text{eq}}) + \sum_{i=1}^{3N} x_i \frac{\partial V}{\partial x_i}\bigg|_{\mathbf{R}_{\text{eq}}} + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{\mathbf{R}_{\text{eq}}} + O(x_i x_j x_k)\,. \tag{2.79}$$

$V(\mathbf{R}_{\text{eq}})$ is the energy at the bottom of the potential well, which we can set to 0, and the first derivatives the the equilibrium positions should also be 0. For sufficiently small oscillations, the cubic terms and above are not significant, so we can truncate out expression at the quadratic term, which is equivalent to making a harmonic approximation. We are left with

$$V(\mathbf{R}) \approx \frac{1}{2} \sum_{i,j=1}^{3N} x_i x_j \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{\mathbf{R}_{\text{eq}}} = \mathbf{x}^{\mathrm{T}} \mathsf{H}_x \mathbf{x} \tag{2.80}$$

23

with Hessian matrix elements

$$H_{x,ij} = \left.\frac{\partial^2 V}{\partial x_i \partial x_j}\right|_{\mathbf{R}_{eq}}.$$ (2.81)

For harmonic potentials, this is equivalent to finding the quadratic coefficients of V. If we want to include the effect of cubic terms and above, we can do this by perturbation theory, which is the topic of the rest of this course.

2.8 Transition States

The eigenvalues of the Hessian matrix represents the second derivatives of the potential energies along the principal directions. Recall that a positive second derivative corresponds to a minimum. Ignoring the six directions that do not corresponds to an internal change of the molecule, we can classify a stationary point on the potential energy surface according to the number of negative Hessian eigenvalues. If there are no negative Hessian eigenvalues, then this is a minimum on the PES. If all the eigenvalues are negative, the point is a maximum.

The Murrel–Laidler definition of a *transition state* is an index-1 saddle point, that is a stationary point with only one negative Hessian value. Therefore a transition state is a maximum along one direction and a minimum in all other orthogonal directions. The normal mode corresponding to the negative Hessian eigenvalue is the reaction coordinate at the transition state. Since the vibrational frequency is the square root of the eigenvalue, the frequency in such a direction is imaginary.[12]

3 Non-degenerate Perturbation Theory

3.1 Power Series Expansions

Often the Schrödinger equation of a Hamiltonian H will be too complicated to be solved directly:

$$\hat{H}\psi_n = E_n\psi_n,$$ (3.1)

where $n \in \mathbb{N}$ labelling the quantum states. Then we need to start to make approximations. The starting point of the perturbation theory is to find a simpler Hamiltonian, which we call $\hat{H}^{(0)}$, that somewhat resembles H but whose eigenfunctions $\psi_n^{(0)}$ and eigenvalues $E_n^{(0)}$ are known:

$$\hat{H}^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}.$$ (3.2)

At this state we will assume that the states are non-degenerate.

The next thing we do to consider the whole one-parameter family of Hamiltonian

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$$ (3.3)

such that

$$\hat{H}(\lambda^*) = \hat{H}$$ (3.4)

at some specific value of $\lambda = \lambda^*$. $\hat{H}^{(1)}$ is the *perturbation* to the reference Hamiltonian. Usually we would just let

$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)}$$ (3.5)

so that $\hat{H}(\lambda)$ becomes the Hamiltonian $\hat{H}$ of interest at $\lambda = 1$. However, sometimes the perturbation to the reference Hamiltonian has a natural perturbation in it. For example, if we consider applying

---

[12]This statement should make you question the derivation of transition state theory presented in the Part IB MELT course. A rigorous treatment of reaction kinetics will be left to Part III.
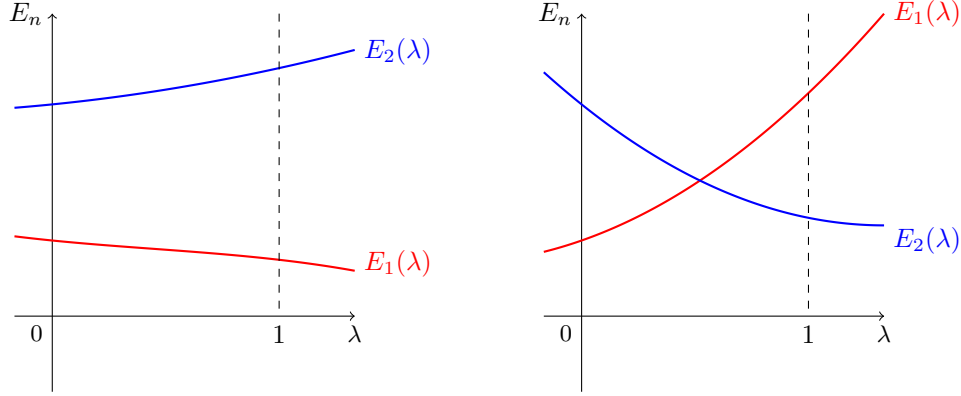
Figure 6: The perturbation theory is appropriate in the first case, in which the reference Hamiltonian is "close enough" to the true Hamiltonian such that the perturbation only leads to small changes in the energies and wavefunctions. The perturbation theory may not work in the second case where the changes in energy levels is comparable with the energy level separations. The reference Hamiltonian is too far away from the true Hamiltonian. Either we will then need too many terms in the series expansion to get a satisfactory result, or the series just may not converge.

a small electric field $\mathcal{E}\hat{\mathbf{z}}$ to a molecule with dipole $\boldsymbol{\mu}$, then the strength of the electric field $\mathcal{E}$ will be a natural perturbation parameter.

Then the fundamental assumption in perturbation theory is to assume that both the energies and the quantum states are analytic near $\lambda = 0$, so that if we gradually turn on the perturbation from the reference state where $\lambda = 0$ to some small non-zero λ, the energies and the wavefunctions will also change smoothly given by the power series

$$\psi_n(\lambda) = \psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \dots \tag{3.6}$$

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \tag{3.7}$$

In hope that these series do converge to the true values, we can then evaluate these power series expansions (usually the first few correction terms only) at the desired value of λ to get a approximation of the true E_n and ψ_n.

We will employ the conventional *intermediate normalisation*, where the unperturbed wavefunction will be normalised

$$\left\langle \psi_n^{(0)} \middle| \psi_n^{(0)} \right\rangle \equiv \langle n|n\rangle = 0\,. \tag{3.8}$$

This choice means that the perturbed wavefunctions $\psi_n(\lambda)$ will not be normalised. However, since $\psi_n^{(1)}$, $\psi_n^{(2)}$,...are correction to $\psi_n^{(0)}$, we may require that

$$\left\langle n \middle| \psi_n^{(i)} \right\rangle = 0 \tag{3.9}$$

for all $i > 0$, from which it follows that

$$\langle n|\psi_n\rangle = 1\,. \tag{3.10}$$

We can claim this because if $\psi_n^{(i)}$ has some non-zero $\psi_n^{(0)}$ component for $i > 0$, we can absorb it into $\psi_n^{(0)}$ and renormalise ψ_n to make the coefficient of $\psi_n^{(0)}$ to be 1 again.

3.2   The Perturbation Equations

To work out the coefficients $E_n^{(i)}$ and $\psi_n^{(i)}$ of the power series, we simply need to plug the expansion ansatz (3.7) and (3.6) into the Schrödinger equation (3.1) to get

$$(\hat{H}^{(0)} + \lambda\hat{H}^{(1)} - E_n^{(0)} - \lambda E_n^{(1)} - \lambda^2 E_n^{(2)} + \dots)(\psi_n^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi_n^{(2)} + \dots) = 0 \tag{3.11}$$

For this equation to be satisfied for all λ sufficiently small such that the series converges, the coefficients of all powers of λ must be zero. Gathering the terms in λ^0, λ^1, λ^2,..., we get a series of equations

$$(\hat{H}^{(0)} - E_n^{(0)})\left|\psi_n^{(0)}\right\rangle = 0 \tag{3.12}$$

$$(\hat{H}^{(0)} - E_n^{(0)})\left|\psi_n^{(1)}\right\rangle + (\hat{H}^{(1)} - E_n^{(1)})\left|\psi_n^{(0)}\right\rangle = 0 \tag{3.13}$$

$$(\hat{H}^{(0)} - E_n^{(0)})\left|\psi_n^{(2)}\right\rangle + (\hat{H}^{(1)} - E_n^{(1)})\left|\psi_n^{(1)}\right\rangle - E_n^{(2)}\left|\psi_n^{(0)}\right\rangle = 0 \tag{3.14}$$

$$\vdots$$

The zeroth-order equation (3.12) is exactly the unperturbed Schrödinger equation (3.2), which we should have already solved. The next equation of interest is the first-order equation (3.13). If we contract it with $\left\langle\psi_n^{(0)}\right| \equiv \langle n|$, we get

$$\left\langle n\left|H^{(0)} - E_n^{(0)}\right|\psi_n^{(1)}\right\rangle + \left\langle n\left|\hat{H}^{(1)} - E_n^{(1)}\right|n\right\rangle = 0\,. \tag{3.15}$$

Since $\hat{H}^{(0)}$ is Hermitian, the first term vanishes:

$$\left\langle n\left|H^{(0)} - E_n^{(0)}\right|\psi_n^{(1)}\right\rangle = \left\langle\psi_n^{(1)}\left|H^{(0)} - E_n^{(0)}\right|n\right\rangle^* = 0\,, \tag{3.16}$$

and we are left with

$$E_n^{(1)} = \left\langle n\left|\hat{H}^{(1)}\right|n\right\rangle\,. \tag{3.17}$$

This is the key result in perturbation theory. The first order correction in energy is simply the expectation value of the perturbation for the unperturbed state. This also implies that

$$E_n^{(0)} + \lambda E_n^{(1)} = \left\langle n\left|\hat{H}^{(0)} + \lambda\hat{H}^{(1)}\right|n\right\rangle = \left\langle n\left|\hat{H}\right|n\right\rangle\,. \tag{3.18}$$

The energy corrected to first order is the expectation value of the complete Hamiltonian with the unperturbed wavefunction.

A similar procedure yields the second-order energy. If we contract the second order perturbation equation (3.14) with $\langle n|$, we get

$$\left\langle n\left|\hat{H}^{(0)} - E_n^{(0)}\right|\psi_n^{(2)}\right\rangle + \left\langle n\left|\hat{H}^{(1)} - E_n^{(1)}\right|\psi_n^{(1)}\right\rangle - \left\langle n\left|E_n^{(2)}\right|n\right\rangle = 0\,. \tag{3.19}$$

The first term again vanishes since $\hat{H}^{(0)}$ is Hermitian. By the intermediate normalisation, we also have $E_n^{(1)}\left\langle n\middle|\psi_n^{(1)}\right\rangle$ in the second term vanishes and the third term $\left\langle n\left|E_n^{(2)}\right|n\right\rangle = E_n^{(2)}$. Therefore, the second order energy is

$$E_n^{(2)} = \left\langle n\left|\hat{H}^{(1)}\right|\psi_n^{(1)}\right\rangle\,. \tag{3.20}$$

To use this expression we will need to obtain the first order wavefunction $\psi_n^{(1)}$. This can be obtained by solving the complicated first-order equation (3.13), which we don't expect to have a simple closed form solution in general unless in very simple cases. We will introduce alternative methods to work this out later.

3.3 Higher-order Energies and Wigner's $2n + 1$ Theorem

By expanding higher order terms of (3.11), one can show that the nth order correction in energy is given by

$$E_n^{(i)} = \left\langle n\left|\hat{H}^{(1)}\right|\psi_n^{(i-1)}\right\rangle\,, \tag{3.21}$$

so it seems that we need to know the wavefunction to the order $i - 1$ to calculate the energy to order i. In fact, we can do better than this

Theorem 3.1 (Wigner's $2n+1$ rule). The $2n+1^{\text{th}}$ order correction of energy can be calculated from the first n^{th} order perturbation of wavefunctions.

This is easily seen in a different formulation of the perturbation theory called the *Wigner–Brillouin perturbation theory*, but it is a far more complicated topic. A brief proof of this is in section A, and here we will only show explicitly how to construct $E_n^{(3)}$ from $\psi_n^{(1)}$.

Appendices

A Wigner's $2n + 1$ Theorem