

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	4
1.1	Wavefunctions and Operators	4
1.2	Hermitian Operators	5
1.3	Unitary Operators	7
1.4	The Momentum Space Representation	8
1.5	Vectors and Matrices Representations	10
1.6	Uncertainty Principle	11

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 is a function of the spatial coordinates. In the Born's interpretation, the probability density of finding a particle at \mathbf{r} is

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

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

$$\int d\tau \Psi^* \Psi = 1 , \quad (1.2)$$

where $d\tau$ 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 evolves according to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi = H \Psi . \quad (1.3)$$

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 refer to as the Schrödinger equation)¹

$$H\psi = E\psi , \quad (1.5)$$

where the constant E is the energy of the system² 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

$$H\psi_n = E_n\psi_n . \quad (1.6)$$

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.

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

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

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

²This is essentially the conservation of energy — if a system has time-translational symmetry, then the energy of the system is conserved. 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

$$\hat{A}\psi = a\psi, \quad (1.7)$$

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

$$\langle A \rangle = \frac{\int d\tau \psi^* \hat{A} \psi}{\int d\tau \psi^* \psi} \quad (1.8)$$

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \quad (1.9)$$

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 integrations³

$$\langle \psi | \psi \rangle := \int d\tau \psi^* \psi. \quad (1.10)$$

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

$$\langle \psi | \hat{A} | \psi \rangle := \int d\tau \psi^* \hat{A} \psi. \quad (1.11)$$

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 operators⁴

$$\hat{x} \longrightarrow x \quad (1.12)$$

$$\hat{p}_x \longrightarrow -i\hbar \frac{\partial}{\partial x}. \quad (1.13)$$

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

$$[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A} \quad (1.14)$$

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

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|. \quad (1.15)$$

1.2 Hermitian Operators

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

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

³Actually 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} , $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$, 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 Riesz representation theorem, which shows for any $|\psi\rangle \in \mathcal{H}$, there is always a corresponding $\langle \psi | \in \mathcal{H}^*$ and *vice versa*.

⁴This 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

$$\langle \phi | \hat{A} | \psi \rangle = \langle \hat{A}^\dagger \phi | \psi \rangle . \quad (1.17)$$

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

$$\langle \phi | \hat{A} | \psi \rangle = \langle \hat{A} \phi | \psi \rangle = \langle \psi | \hat{A} | \phi \rangle^* . \quad (1.18)$$

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

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

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

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

$$(\hat{A} | \psi \rangle)^\dagger = \langle \psi | \hat{A}^\dagger . \quad (1.22)$$

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,

$$\psi = \sum_i c_i \phi_i . \quad (1.23)$$

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

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

and so

$$\psi = \sum_i \langle \phi_i | \psi \rangle | \phi_i \rangle . \quad (1.25)$$

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

$$|\psi\rangle = \sum_i \langle i | \psi \rangle |i\rangle . \quad (1.26)$$

⁵There 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

$$\hat{I} |\psi\rangle = |\psi\rangle \quad (1.27)$$

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

$$\hat{I} = \sum_i |i\rangle \langle i| . \quad (1.28)$$

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

$$\hat{I} |\psi\rangle = \sum_i |i\rangle \langle i|\psi\rangle = \sum_i \langle i|\psi\rangle |i\rangle = |\psi\rangle , \quad (1.29)$$

where we moved $\langle i|\psi\rangle$ to the front to make the expression more obvious — we are allowed to do this because $\langle i|\psi\rangle$ 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

$$\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{I} . \quad (1.30)$$

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

$$\hat{U}^{-1} = \hat{U}^\dagger , \quad (1.31)$$

so that

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I} . \quad (1.32)$$

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

$$\hat{A}\psi = \phi , \quad (1.33)$$

and after the transformation, we have

$$\psi' = \hat{U}\psi , \quad \phi' = \hat{U}\phi . \quad (1.34)$$

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

$$\hat{A}'\psi' = \phi' . \quad (1.35)$$

We have

$$\hat{A}'\hat{U}\psi = \hat{U}\phi = \hat{U}\hat{A}\psi \implies \hat{A}'\hat{U} = \hat{U}\hat{A} . \quad (1.36)$$

Therefore,

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

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

⁶Unitary 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

$$\hat{A}\psi_n = a_n\psi_n, \quad (1.38)$$

then since we can insert an identity operator anywhere,

$$\hat{A}\hat{U}^\dagger\hat{U}\psi_n = a_n\hat{U}^\dagger\hat{U}\psi_n, \quad (1.39)$$

and so

$$\hat{U}\hat{A}\hat{U}^\dagger\hat{U}\psi_n = a_n\hat{U}\psi_n, \quad (1.40)$$

which implies

$$\hat{A}'\psi'_n = a_n\psi'_n. \quad (1.41)$$

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

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

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

$$\int d\tau' \psi'_n \psi_n = \int d\tau \psi_n^* \psi_n, \quad (1.43)$$

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

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

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

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

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.

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

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

⁷We 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.

⁸Dirac 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

$$\int_{-\epsilon}^{\epsilon} dx \delta(x) = 0. \quad (1.47)$$

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

Dirac delta has a nice property.

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

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

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

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 in 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.

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

Since

$$-i \frac{\partial}{\partial p} e^{ipx} = x e^{ipx}, \quad (1.51)$$

we can rewrite it as

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

We now use integration by part, and we get

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

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

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

$$\hat{p}' = \hat{U} \hat{p} \hat{U}^\dagger = p. \quad (1.55)$$

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

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle, \quad (1.56)$$

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

$$c_i = \langle \phi_i | \psi \rangle. \quad (1.57)$$

This expansion be rewritten in matrix notation as

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

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

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

or

$$\boldsymbol{\theta}^T = \boldsymbol{\phi}^T \mathbf{U}. \quad (1.60)$$

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

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

which implies

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}, \quad (1.62)$$

and so \mathbf{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 $\{|\theta_i\rangle\}$

$$|\psi\rangle = \boldsymbol{\theta}^T \mathbf{d}. \quad (1.63)$$

Then clearly, we should have

$$|\psi\rangle = \boldsymbol{\theta}^T \mathbf{d} = \boldsymbol{\phi}^T \mathbf{U} \mathbf{d} = \boldsymbol{\phi}^T \mathbf{c}, \quad (1.64)$$

so

$$\mathbf{c} = \mathbf{U} \mathbf{d} \quad \text{or} \quad \mathbf{d} = \mathbf{U}^\dagger \mathbf{c}. \quad (1.65)$$

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

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

and so

$$\mathbf{A}' = \mathbf{U}^\dagger \mathbf{A} \mathbf{U}. \quad (1.67)$$

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 $|\theta_i\rangle$) 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 $H_{ij} = \langle \phi_i | \hat{H} | \phi_j \rangle$, then the Hamiltonian in the symmetry orbital basis set is

$$\mathbf{H}' = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}. \quad (1.68)$$

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 \mathbf{H} . By doing a unitary basis transformation, the eigenvalues of \mathbf{H}' should be the same as \mathbf{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 \mathbf{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

$$\sigma_A^2 = \langle A^2 \rangle - \langle A \rangle^2 = \langle (A - \langle A \rangle)^2 \rangle. \quad (1.69)$$

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

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

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

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

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

$$\Delta A^2 = \langle \psi_A | \psi_A \rangle. \quad (1.72)$$

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

$$\Delta B^2 = \langle \psi_B | \psi_B \rangle \quad (1.73)$$

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

$$\Delta A^2 \Delta B^2 = \langle \psi_A | \psi_A \rangle \langle \psi_B | \psi_B \rangle. \quad (1.74)$$

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 $\mathbf{u}, \mathbf{v} \in V$,

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

Therefore, we have

$$\Delta A^2 \Delta B^2 \geq |\langle \psi_A | \psi_B \rangle|^2. \quad (1.76)$$

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

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

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

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

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

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

Similarly

$$\langle \psi_B | \psi_A \rangle = \langle \hat{B}\hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (1.80)$$

Defining the *anti-commutator*

$$\{\hat{A}, \hat{B}\} := \hat{A}\hat{B} + \hat{B}\hat{A}, \quad (1.81)$$

we have

$$\langle \psi_A | \psi_B \rangle + \langle \psi_B | \psi_A \rangle = \langle \{\hat{A}, \hat{B}\} \rangle - 2\langle \hat{A} \rangle \langle \hat{B} \rangle, \quad (1.82)$$

$$\langle \psi_A | \psi_B \rangle - \langle \psi_B | \psi_A \rangle = \langle [\hat{A}, \hat{B}] \rangle. \quad (1.83)$$

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

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

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

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

$$\text{cov}(\hat{A}, \hat{B}) := \frac{1}{2} \langle \hat{A}\hat{B} - \hat{B}\hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle, \quad (1.85)$$

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

$$\text{cov}(X, Y) := \langle XY \rangle - \langle X \rangle \langle Y \rangle, \quad (1.86)$$

but in quantum mechanics we have to be a little bit more careful because in general $\langle \hat{A}\hat{B} \rangle \neq \langle \hat{B}\hat{A} \rangle$. 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

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

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