Vol. 3 – QUANTUM MECHANICS

# The Undergraduate Companion to Theoretical Physics

**Andrea Kouta Dagnino**[‡]

[‡]*Open University, Milton Keynes, UK.*

*E-mail:* k.y.dagnino@gmail.com

# Contents

*If quantum mechanics hasn't profoundly shocked you yet, then you haven't understood it yet.*

— Niels Bohr

# References

Several textbooks, online courses/resources were referenced heavily (to the extend of making this text completely unoriginal, yet hopefully helpful for revision) throughout the writing of these lecture notes. Using a typical bibliography (research paper style) would be a formidable task.Pinpointing exactly where each reference as been used is quite difficult for such a large and well-referenced subject, and would probably change the writing style to a far too formal one for lecture notes. Therefore we instead list the most relevant below giving a brief comment on which topics they were mostly used for:

- *OU textbooks*

  While probably not available online (and thus not very useful to the reader), I found these textbooks to be very helpful as a resource to understand the basics, though some parts of these notes definitely go beyond what is required in SM358. Many images that would be hard to draw by myself on Illustrator are also included from these textbooks.

- R. Shankar, *"Principles of Quantum Mechanics"*

  This is a fantastic pedagogical text, the perfect fit for a first introduction. It contains a short introduction to bra-ket notation and the linear algebra needed for QM which I found to be very useful. The only topics this text wasn't referenced for were those on quantum information, molecular/solid state physics, and the adiabatic approximation. Shankar was especially helpful while typesetting part I.

- D. Griffiths, *"Introduction to Quantum Mechanics"*

  Another great first introduction to the subject, although perhaps it leaves the algebra of QM hidden for too long. Consequently it mostly presents a wave-function approach to the subject, complementing Shankar's matrix algebra approach.

- L. Landau and E. Lifshitz, *"Quantum mechanics: non-relativistic theory"*

  A classic, advanced and comprehensive text written by one of the "greats". This text was mostly used for part IV of this text, especially atomic/molecular physics.

- John J. Sakurai and J. Napolitano, *"Modern Quantum Mechanics"*

  Much like Landau & Lifshitz. a classic, advanced and comprehensive text. It was especially helpful for approximation methods and the more advanced, group the-

oretical aspect of QM. Definitely not introductory, but I found it enlightening for a second course on QM.

- J. Binney *"The Physics of Quantum Mechanics"* Really nice textbooks. I enjoyed it especially for its alternative algebraic derivation of the Hydrogen energy levels. I used it mostly for the chapter on quantum informaion and part IV. Great for a first introduction.

- MIT OpenCourseWare 8.04-8.06 courses

  A great lecture series, following more or less the same structure as these notes. One who sits all three courses would (imo) have a sound undergraduate knowledge of QM. 8.04 and 8.05 cover parts I and II while 8.06 covers parts III and IV. The latter two specifically were in many parts written while listening to the 8.06 lectures. Most of chapter 2 was a copy of the first few lectures in 8.04 by Allan Adams.

- Course notes by D. Skinner on *"Part II Principles of Quantum Mechanics"*

  The lecture notes were especially clear in explaining angular momentum and covers parts II, III comprehensively. Good for a second introduction to QM.

- Course notes by B. Simons on *"Part II Advanced Quantum Mechanics"*

  I found it very helpful for part IV and V. It is one of few course notes I have found discussing atomic/molecular physics at the right level to be included in a more broader QM course, but that is not too superficial either. It has solutions to the problem sets which are quite rare, so I'd recommend it for practice problems too.

# Mathematical structure of Quantum Mechanics (in progress)

## Linear Spaces

We begin by defining a fundamental mathematical concept used in QM, the **Linear Space** which was introduced in the linear algebra course. Classically speaking, vectors are defined as objects with both a magnitude and direction. However, as we will see soon this definition is very limited, and breaking beyond the barrier of arrows with lengths and directions will enable us to create a mathematical structure for quantum mechanics.

**Definition 1.1**: Linear Space

A linear space $\mathcal{H}$ over a field $\mathbb{K}$ is a collection of **vectors** $|x_i\rangle$ over which the opeartions $+, \cdot$ are defined, such that $\forall\, |x_1\rangle, |x_2\rangle \in V$ and $\forall \alpha_1, \alpha_2 \in \mathbb{K}$ the following are satisfied:

(i) Closure under addition: $|x_1\rangle + |x_2\rangle \in V$
(ii) Closure under scalar multiplication: $\alpha_1 |x_1\rangle \in V$
(iii) Commutativity of addition: $|x_1\rangle + |x_2\rangle = |x_2\rangle + |x_1\rangle$
(iv) Associativity of addition: $|x_1\rangle + |x_2\rangle$
(vi) Associativity of addition: $|x_1\rangle + (|x_2\rangle + |x_3\rangle) = (|x_1\rangle + |x_2\rangle) + |x_3\rangle)$
(vi) Associativity of scalar multiplication: $\alpha_1(\alpha_2 |x_1\rangle) = \alpha_1\alpha_2 |x_1\rangle$
(vii) Right-distributivity: $(\alpha_1 + \alpha_2) |x_1\rangle = \alpha_1 |x_1\rangle + \alpha_2 |x_1\rangle$
(viii) Left-distributivity: $\alpha_1(|x_1\rangle + |x_2\rangle) = \alpha_1 |x_1\rangle + \alpha_1 |x_2\rangle$
(ix) Existence of **zero vector**: $\exists\, |0\rangle \in V$ such that $|x_1\rangle + |0\rangle = |x_1\rangle$
(x) Existence of inverse under addition: $\exists\, |-x_1\rangle \in V$ such that $|x_1\rangle + |-x_1\rangle = |0\rangle$

For example, despite not having a magnitude nor direction, the set of all $2 \times 2$ real matrices is a linear space over $\mathbb{R}$ (prove it)! Moreover, certain sets of functions may also be regarded as linear spaces.

**Example 1:** Function spaces are Linear Spaces

**The set** $\mathbb{R}_n[x] = \{f(x) = \sum_{i=0}^{n} a_i x^i : a_i \in \mathbb{R}, a_n \neq 0\}$ **is a linear space.**

*Proof.* □

## Linear (in)-dependence, bases and dimensions

**Definition 2:** LINEAR INDEPENDENCE

A set of vectors $|x_i\rangle \in V$ is linearly independent iff:

$$\sum_{i=1}^{n} \alpha_i |x_i\rangle = |0\rangle \implies \alpha_i = 0 \tag{0.0.1}$$

Otherwise, the set of vectors is linearly dependent.

We remark an important consequence of this definition: if a set of vectors is linearly independent, then none of its members can be expressed using the other members of the set and the standard $+, \cdot$ operations. Therefore, all of the vectors are individual and non-reproducible from the others linearly.

**Example 2: The vectors $|1\rangle = (1,1,1),\ |2\rangle = (\alpha,\beta,\gamma),\ |3\rangle = (\alpha^2,\beta^2,\gamma^2)$ for distinct $\alpha,\beta,\gamma \in \mathbb{R}^3$ are linearly independent.**

*Proof.* We consider $c_1|1\rangle + c_2|2\rangle + c_3|3\rangle = |0\rangle$, then:

$$c_1(1,1,1)+c_2(\alpha,\beta,\gamma)+c_3(\alpha^2,\beta^2,\gamma^2) = (0,0,0) \implies \begin{cases} c_1 + c_2\alpha + c_3\alpha^2 = 0 \\ c_1 + c_2\beta + c_3\beta^2 = 0 \\ c_1 + c_2\gamma + c_3\gamma^2 = 0 \end{cases} \tag{0.0.2}$$

Subtracting the second equation from the first gives:

$$c_2(\beta-\alpha) + c_3(\beta+\alpha)(\beta-\alpha) = 0 \implies c_2 + c_3(\beta+\alpha) = 0 \tag{0.0.3}$$

Similarly, for the other two combinations we get:

$$c_2 + c_3(\gamma+\alpha) = 0$$
$$c_2 + c_3(\beta+\gamma) = 0$$

from which we get that $c_3(\gamma+\alpha) = c_3(\gamma+\beta) = c_3(\beta+\alpha) = 0 \implies c_1 = c_2 = c_3 = 0$ as desired. □

## Inner products

## Subspaces, Direct sums and products

## Linear Operators

A linear operator on a space $\mathcal{H}$ transforms kets in V into other kets in $\mathcal{H}$ linearly.

**Definition : LINEAR OPERATORS**

Let $\hat{\Omega}$ be a linear operator acting on $\mathcal{H}$, and let $\alpha \in \mathbb{C}$. Then $\forall \psi_1, \psi_2 \in V$:

$$\hat{Q}(\alpha \left|\psi_1\right\rangle + \left|\psi_2\right\rangle = \alpha(\hat{Q} \left|\psi_1\right\rangle) + \hat{Q} \left|\psi_2\right\rangle \tag{0.0.4}$$

So, if for a basis $|i\rangle$ of $\mathcal{H}$ we have $\hat{\Omega} \left|j\right\rangle = \left|j'\right\rangle$ then, $\forall \left|\psi\right\rangle \in \mathcal{H}$:

$$\hat{\Omega} \left|\psi\right\rangle = \sum_j \hat{\Omega} a_j \left|j\right\rangle = \sum_j \hat{\Omega} a_j \left|j'\right\rangle \tag{0.0.5}$$

So the action of an operator on a vector preserves the same coordinates but over the transformed basis.

### Matrix representation of Linear Operators

It is a fundamental theorem in linear algebra that all linear operators can be represented by a matrix. Consider an operator $\hat{\Omega}$ acting on some basis vector $|i\rangle$ giving $|i'\rangle$. Then the projection of this transformed ket along $|j\rangle$ is:

$$\left\langle j \,\middle|\, i' \right\rangle = \left\langle j \middle| \hat{\Omega} \middle| i \right\rangle \equiv \Omega_{ji} \tag{0.0.6}$$

If we consider $\hat{\Omega} \left|\psi\right\rangle = \left|\psi'\right\rangle$ then:

$$\psi' = \left\langle j \,\middle|\, \psi' \right\rangle = \left\langle j \middle| \hat{\Omega} \middle| \psi \right\rangle = \langle j| \hat{\Omega} \Big( \sum_i \psi_i \left|i\right\rangle \Big) = \sum_i \psi_i \left\langle j \middle| \hat{\Omega} \middle| i \right\rangle = \sum_i \psi_i \Omega_{ji} \tag{0.0.7}$$

or in matrix form:

$$\begin{pmatrix} \psi_1' \\ \psi_2' \\ \vdots \\ \psi_n' \end{pmatrix} = \begin{pmatrix} \langle 1 \,|\, \Omega \,|\, 1 \rangle & \langle 1 \,|\, \Omega \,|\, 2 \rangle & \dots & \langle 1 \,|\, \Omega \,|\, n \rangle \\ \langle 2 \,|\, \Omega \,|\, 1 \rangle & \langle 2 \,|\, \Omega \,|\, 2 \rangle & \dots & \langle 2 \,|\, \Omega \,|\, n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle n \,|\, \Omega \,|\, 1 \rangle & \langle n \,|\, \Omega \,|\, 2 \rangle & \dots & \langle n \,|\, \Omega \,|\, n \rangle \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \tag{0.0.8}$$

**Definition: MATRIX REPRESENTATION OF A LINEAR OPERATOR**

For a linear operator $\Omega$ mapping $V_i$ to $V_j$ with bases $|i\rangle$ and $|j\rangle$ respectively:

$$\left\langle j \,\middle|\, i' \right\rangle = \left\langle j \middle| \hat{\Omega} \middle| i \right\rangle \equiv \Omega_{ji} \tag{0.0.9}$$

If $\hat{\Omega} \left|\psi\right\rangle = \left|\psi'\right\rangle$ with $\left|\psi\right\rangle = \sum_i \psi_i \left|i\right\rangle$ and $\left|\psi\right\rangle = \sum_j \psi_j' \left|j\right\rangle$ then:

$$\psi' = \sum_i \psi_i \Omega_{ji} \tag{0.0.10}$$

### Identity and Projection Operators

The projection operator can be defined as:

$$\hat{P}_i = |i\rangle \langle i| \tag{0.0.11}$$

Its action on some vector $|\psi\rangle$ is to provide its projection along $|i\rangle$. Applying two projections along two basis vectors gives:

$$\hat{P}_j \hat{P}_i = |j\rangle \langle j \,|\, i\rangle \langle i| = \hat{P}_i \delta ji \tag{0.0.12}$$

So if we first project some ket along $|i\rangle$, then applying $\hat{P}_j$ will give zero if $|i\rangle \neq |j\rangle$, and will do nothing if $|i\rangle = |j\rangle$. In other words:

$$\hat{P}_i \hat{P}_i = \hat{P}_i. \tag{0.0.13}$$

Let us then find the matrix elements of $\hat{P}_i$:

$$(\hat{P}_i)_{kl} = \left\langle k \middle| \hat{P}_i \middle| l \right\rangle = \langle k \,|\, i\rangle \langle i \,|\, l\rangle = \delta ki \delta il = \delta kl \tag{0.0.14}$$

The identity operator acting on a ket should yield the very same ket. We can define the operator as follows:

$$\hat{I} \to \sum_i |i\rangle \langle i| \,. \tag{0.0.15}$$

Indeed

$$\hat{I} \,|\psi\rangle = \sum_i |i\rangle \langle i \,|\, \psi\rangle = \sum_i |i\rangle \langle i| \left( \sum_j a_j \,|j\rangle \right) = \sum_i a_i \,|i\rangle = |\psi\rangle \tag{0.0.16}$$

Let us try to find the matrix representation of $\hat{I}$:

$$\left\langle j \middle| \hat{I} \middle| i \right\rangle = \langle j \,|\, i\rangle = \delta_{ji}{}^{1} \tag{0.0.17}$$

The attentive reader may have caught onto the relationship tying these two operators. Indeed:

$$\hat{I} \,|\psi\rangle = \hat{I} \sum_i \psi_i \,|i\rangle = \sum_i |i\rangle \langle i \,|\, \psi\rangle = \sum_i \hat{P}_i \,|\psi\rangle \implies \hat{I} \to \sum_i \hat{P}_i \tag{0.0.18}$$

[1] if one uses eq. 0.0.15 it is important to note that $|i\rangle$ in the definition is not equal to the $|i\rangle$ we're acting on, indeed only for one value of $i$ will they be equal.

### Products of Operators and Commutators

Taking the product of two operators is equivalent to applying one after the other to some ket. For example, consider the product of two linear operators $\hat{\Omega}$ and $\hat{\Lambda}$:

$$(\hat{\Omega}\hat{\Lambda})_{ji} = \left\langle j \middle| \hat{\Omega}\hat{\Lambda} \middle| i \right\rangle = \left\langle j \middle| \hat{\Omega}\hat{I}\hat{\Lambda} \middle| i \right\rangle \tag{0.0.19}$$

$$= \sum_k \left\langle j \middle| \hat{\Omega} \middle| k \right\rangle \left\langle k \middle| \hat{\Lambda} \middle| i \right\rangle \tag{0.0.20}$$

$$= \sum_k \Omega_{jk}\Lambda_{ki} \tag{0.0.21}$$

Also, operators, unlike scalars, generally do not commute, although some do. So, it would be useful to define some quantity that can tell us whether or not commutativity is present for two operators. The commutator is defined as:

$$[\hat{\Omega}, \hat{\Lambda}] = \hat{\Omega}\hat{\Lambda} - \hat{\Lambda}\hat{\Omega} \tag{0.0.22}$$

We summarize some important properties of commutators below:

**Properties:** COMMUTATORS

The commutator between linear operators has the following properties:

(i) $[\hat{\Omega}, \hat{\Lambda}] = \hat{\Omega}\hat{\Lambda} - \hat{\Lambda}\hat{\Omega}$
(ii) $[\hat{\Omega} + \hat{\Theta}, \hat{\Lambda}] = [\hat{\Omega}, \hat{\Lambda}] + [\hat{\Theta}, \hat{\Lambda}]$
(iii) $[\hat{\Omega}\hat{\Theta}, \hat{\Lambda}] = [\hat{\Omega}, \hat{\Lambda}]\hat{\Theta} + \hat{\Omega}[\hat{\Theta}, \hat{\Lambda}]$
(iv) $[\hat{\Omega}, \hat{\Theta}\hat{\Lambda}] = [\hat{\Omega}, \hat{\Theta}]\hat{\Lambda} + \hat{\Theta}[\hat{\Omega}, \hat{\Lambda}]$

### Hermitian adjoint operators

Recall that for a ket $\alpha |\psi\rangle$ corresponds a bra $\alpha * \langle\psi|$. Similarly, for $\hat{\Omega} |\psi\rangle$ corresponds a bra $\langle\psi| \hat{\Omega}^\dagger$, the operator $\hat{\Omega}^\dagger$ is called the **Hermitian adjoint**:

$$(\hat{\Omega}^\dagger)_{ij} = \left\langle i \middle| \hat{\Omega}^\dagger \middle| j \right\rangle = \left\langle \hat{\Omega} i \middle| j \right\rangle = \left\langle j \middle| \hat{\Omega} i \right\rangle * = (\hat{\Omega}*)_{ji} \tag{0.0.23}$$

from which it follows that the matrix representing the Hermitian adjoint of an operator is the conjugate transpose of the original operator's matrix.

**Properties:** HERMITIAN ADJOINT PROPERTIES

For two linear operators $\hat{\Omega},\ \hat{\Lambda}$:

$$(\hat{\Omega}\hat{\Lambda})^\dagger = \hat{\Lambda}^\dagger\hat{\Omega}^\dagger \tag{0.0.24}$$

*Proof.* Consider:

$$\langle\hat{\Omega}\hat{\Lambda}V| = \langle(\hat{\Omega}\hat{\Lambda})V| = \langle V|\,(\hat{\Omega}\hat{\Lambda})^\dagger \tag{0.0.25}$$

$$\langle\hat{\Omega}\hat{\Lambda}V| = \langle\hat{\Omega}(\hat{\Lambda}V)| = \langle\hat{\Lambda}V|\,\hat{\Omega}^\dagger = \langle V|\,\hat{\Lambda}^\dagger\hat{\Omega}^\dagger \tag{0.0.26}$$

□

**Definition:** HERMITIAN AND UNITARY OPERATORS

A linear operator $\hat{\Omega}$ is said to be:
(i) **Hermitian** iff $\hat{\Omega} = \hat{\Omega}^\dagger$
(ii) **anti-Hermitian** iff $\hat{\Omega} = -\hat{\Omega}^\dagger$
(i) **unitary** iff $\hat{\Omega}\hat{\Omega}^\dagger =$
.

**Theorem:** PROPERTIES OF UNITARY OPERATORS

A unitary operator $\hat{U}$ preserves the inner product of two vectors:

$$\langle\psi\,|\,\psi'\rangle = \left\langle\hat{U}\psi\,\middle|\,\hat{U}\psi'\right\rangle \tag{0.0.27}$$

Moreover the column kets (and the row bras) of a square matrix representing a unitary operator form an orthonormal basis.

*Proof.* Consider:

$$\left\langle\hat{U}\psi\,\middle|\,\hat{U}\psi'\right\rangle = \left\langle\psi\,\middle|\,\hat{U}^\dagger\hat{U}\,\middle|\,\psi'\right\rangle = \left\langle\psi\,\middle|\,\hat{I}\,\middle|\,\psi'\right\rangle = \langle\psi\,|\,\psi'\rangle \tag{0.0.28}$$

as desired. Also:

$$\hat{U}^\dagger\hat{U} = \hat{I} \implies \delta_{ij} = \left\langle i\,\middle|\,\hat{I}\,\middle|\,j\right\rangle = \left\langle i\,\middle|\,\hat{U}^\dagger\hat{U}\,\middle|\,j\right\rangle \tag{0.0.29}$$

$$= \sum_k \left\langle i\,\middle|\,\hat{U}^\dagger\,\middle|\,k\right\rangle\left\langle k\,\middle|\,\hat{U}\,\middle|\,j\right\rangle = \sum_k (\hat{U}^\dagger)_{ik}(\hat{U})_{kj} \tag{0.0.30}$$

$$= \sum_k (\hat{U})^*_{ki}(\hat{U})_{kj} = \langle U_i\,|\,U_j\rangle \tag{0.0.31}$$

so the columns kets are indeed orthonormal. Starting with $\hat{U}\hat{U}^\dagger = \hat{I}$ yields the same result for row bras. □

## Eigen-everything

### The eigenvalue problem

Consider a special ket $|\omega\rangle$ which, when acted upon by some operator $\hat{\Omega}$ is simply rescaled by a factor $\omega$. We can express this property as:

$$\hat{\Omega}\,|\omega\rangle = \omega\,|\omega\rangle \equiv (\hat{\Omega} - \omega\hat{I})\,|\omega\rangle = |0\rangle\,. \tag{0.0.32}$$

We call $\omega$ the **eigenvalue** and $|phi\rangle$ the corresponding **eigenket** of the operator. It can be shown that this is equivalent to asserting that $\det(\Omega - \omega I) = 0$. This is the **characteristic equation**.

**Theorem:** EIGEN-CHARACTERISTICS OF HERMITIAN AND UNITARY OPERATORS

(I) The eigenvalues of a Hermitian operator $\hat{\Omega}$ are real and there exists an eigenbasis formed by orthonormal eigenkets where the matrix representing $\hat{\Omega}$ is diagonal with corresponding eigenvalues in its diagonal entries.
(II) The eigenvalues of a unitary operator $\hat{U}$ are complex of unit modulus, and the corresponding eigenkets are mutually orthogonal.

*Proof.* (I) Consider:

$$\hat{\Omega}\,|\omega\rangle = \omega\,|\omega\rangle \implies \left\langle \omega \middle| \hat{\Omega} \middle| \omega \right\rangle = \langle \omega \,|\, \omega \,|\, \omega \rangle \tag{0.0.33}$$

$$\left\langle \omega \middle| \hat{\Omega}^\dagger \middle| \omega \right\rangle = \langle \omega \,|\, \omega^* \,|\, \omega \rangle \tag{0.0.34}$$

from which it follows due to the hermiticity of $\hat{\Omega}$ that $(\omega^* - \omega)\langle \omega \,|\, \omega \rangle = 0 \implies \omega^* = \omega \implies \omega \in \mathbb{R}$ since $|\omega\rangle \neq |0\rangle$ (which is a trivial eigenket).

Now consider an arbitrary eigenvalue $\omega_1$ for which there must exist $|\omega_1\rangle$, and the space $\mathcal{H}_\perp^{n-1}$ consisting of all kets orthogonal to $|\omega_1\rangle$. Choosing as our basis $\{|\omega_1\rangle, |\phi_\perp^1\rangle, ..., |\phi_\perp^{n-1}\rangle\}$ then

$$\Omega = \begin{pmatrix} \omega_1 & 0 \\ 0 & A_1 \end{pmatrix} \implies (\omega_1 - \omega)\det(A_1 - \omega I) = 0 \tag{0.0.35}$$

Again we choose $\omega_2$ such that $\det(A_1 - \omega_2 I) = 0$ so that:

$$\Omega = \begin{pmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & A_2 \end{pmatrix} \implies (\omega_1 - \omega)(\omega_2 - \omega)\det(A_2 - \omega I) = 0 \tag{0.0.36}$$

and repeat until we diagonalize completely $\Omega$.

(ii) Let $\hat{U}\,|u_i\rangle = u_i\,|u_i\rangle$ and $\hat{U}\,|u_j\rangle = u_j\,|u_j\rangle \implies \langle u_j|\,\hat{U}^\dagger = u_j^*\,\langle u_j|$. Hence:

$$\langle u_j|\,\hat{U}^\dagger\hat{U}\,|u_i\rangle = u_i u_j^*\,\langle u_j \,|\, u_i\rangle \implies (1 - u_i u_j^*)\,\langle u_i \,|\, u_j\rangle = 0 \tag{0.0.37}$$

It follows that if $i = j$ then $|u_i|^2 = 1$ and so $u_i$ are unit complex eigenvalues. It follows that if $i \neq j$ then $\langle u_i \,|\, u_j\rangle = |0\rangle$, so they eigenkets are mutually orthogonal.

□

Therefore, if $\hat{\Omega}$ is a Hermitian operator, then it is **unitarily diagonalizable** by transforming to its eigenbasis. Recall from your Linear Algebra course that to transform a matrix from

one basis (with representation $\Omega'$) to another (with representation $\Omega$), we must calculate:

$$\Omega = U^{-1}\Omega' U \tag{0.0.38}$$

In our case, since $U$ is the matrix containing the eigenkets, it is also unitary [2] and thus we find $\Omega = U^{\dagger}\Omega' U$.

### Degeneracy

### Simultaneous diagonalization

## Generalizing to infinite dimensions

### Functions of operators

### Infinite kets

Integration will be herein assumed to run over a suitable interval, generally $\mathbb{R}$ if not specified.

### The Dirac Delta Function

The analogue to the identity operator in the infinite limit is:

$$\hat{I} = \sum_i |i\rangle \langle i| \longrightarrow \hat{I} = \int_{-\infty}^{\infty} |x'\rangle \langle x'| \, dx' \tag{0.0.39}$$

which is known as the **completeness relation**. The analogue to the decomposition of a ket into its components in the infinite limit is:

$$|\psi\rangle = \sum_i \psi_i |i\rangle \longrightarrow |\psi\rangle = \int_{-\infty}^{\infty} \langle x'| \, dx' \tag{0.0.40}$$

so that by definition the components are given by:

$$\psi_i = \langle i \,|\, \psi\rangle \longrightarrow \langle x \,|\, \psi\rangle \, \psi(x) = \langle x \,|\, \psi\rangle \tag{0.0.41}$$

One can then see that to convert to the infinite dimensional limit it suffices to change all sums over $i$ with integrals in $x$, position. This is because unlike the discrete $i$ index, $x$ is continuous.

Another consequence of this definition is that:

$$\langle \psi \,|\, \phi\rangle = \int_{-\infty}^{\infty} \langle \psi \,|\, x\rangle \, \langle x \,|\, \phi\rangle \, dx = \int_{-\infty}^{\infty} \psi^*(x)\phi(x) dx \tag{0.0.42}$$

What is the analogue of the Kronecker delta? Consider this:

$$\left\langle x \,\middle|\, \hat{I} \,\middle|\, \psi \right\rangle = \psi(x) = \int_{-\infty}^{\infty} \langle x \,|\, x'\rangle \, \langle x' \,|\, \psi\rangle \, dx' = \int_{-\infty}^{\infty} \delta(x, x')\psi(x') dx' \tag{0.0.43}$$

[2] the eigenkets are mutually orthonormal, by theorem: properties of unitary operators

where we have defined $\delta(x, x') \equiv \langle x \,|\, x' \rangle$. It is then immediate that if $x \neq x'$ the orthogonality requirement gives:

$$\delta(x, x') = 0 \tag{0.0.44}$$

and if $x = x'$ then we integrate over an infinitesimally small interval $\mathcal{J} = [x - \epsilon, x + \epsilon]$ so that $\psi(x)$ may be regarded as constant:

$$\psi(x) = \psi(x) \int_{\mathcal{J}} \delta(x, x') dx' \implies \int_{\mathcal{J}} \delta(x, x') dx' = 0 \tag{0.0.45}$$

Noticing that $\delta(x, x')$ only depends on $x - x'$ we may then write $\delta(x - x')$ which is known as the Dirac delta function. Similarly to the Kronecker delta, it acts as a sampler/sieve:

$$\int_{-\infty}^{\infty} \delta(x - x') f(x') dx' = f(x) \tag{0.0.46}$$

**The Dirac delta function**

The Dirac delta function satisfies:

(i) $\delta x - x' = 0$ if $x \neq x'$
(ii) $\int_{-\infty}^{\infty} \delta(x - x') dx = 1$ if $x = x'$
(iii) $\int_{-\infty}^{\infty} \delta(x - x') f(x') dx' = f(x)$
(iv) $\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x'-x)}\, dk$

The latter way to define the Dirac delta function is through Fourier analysis. For a given continuous function $\psi$ we define its Fourier transform to be:

$$\psi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x)\, dx \tag{0.0.47}$$

and the inverse Fourier transform to be:

$$\psi(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx'} \psi(k)\, dk \tag{0.0.48}$$

Substituting one into the other gives:

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx'} \Big( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \psi(x)\, dx \Big)\, dk \tag{0.0.49}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{ik(x'-x)}\, dk \right) \psi(x) dx \tag{0.0.50}$$

which implies that

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x'-x)}\, dk = \mathcal{F}^{-1}\Big( \frac{1}{2\pi} e^{ikx} \Big) \tag{0.0.51}$$

where $\mathcal{F}$ is the inverse Fourier transform.

Finally let's investigate the derivative of the Dirac delta function.

$$\int_{-\infty}^{\infty} f(x')\frac{\partial}{\partial x}\delta(x-x')dx' = \frac{d}{dx}\int_{-\infty}^{\infty} f(x')\delta(x-x')dx' = \frac{df(x)}{dx} = \int_{-\infty}^{\infty}\delta(x-x')\frac{df}{dx'}dx' \tag{0.0.52}$$

so that:

$$\int_{-\infty}^{\infty} f(x')\frac{\partial}{\partial x}\delta(x-x')dx' = \int_{-\infty}^{\infty}\delta(x-x')\frac{df}{dx'}dx' \tag{0.0.53}$$

Setting $f(x') = 1$ then gives:

$$\delta(x-x') = x\delta'(x-x'). \tag{0.0.54}$$

## Operators in infinite dimensions

### Position operator $\hat{x}$

**Definition:** THE POSITION OPERATOR

The position operator is defined as:

$$\hat{x} = \int_{-\infty}^{\infty} x\,|x\rangle\,\langle x|\,dx \tag{0.0.55}$$

If we apply the position operator to some ket $|\phi\rangle$ and get $|\psi\rangle = \hat{x}\,|\phi\rangle$, then:

$$\psi(x') = \langle x'\,|\,\hat{x}\,|\,\psi\rangle = \int_{-\infty}^{\infty} x\,\langle x'\,|\,x\rangle\,\langle x\,|\,\phi\rangle\,dx = \int_{-\infty}^{\infty} x\delta(x-x')\phi(x)dx = x'\phi(x') \tag{0.0.56}$$

therefore we find that:

$$\hat{x}\,|x\rangle = x\,|x\rangle \tag{0.0.57}$$

which is the solution to the eigenvalue problem for $\hat{x}$. Similarly, for two states $|\phi\rangle, |\psi\rangle$ we find:

$$\langle\psi\,|\,\hat{x}\,|\,\phi\rangle = \int_{-\infty}^{\infty}\langle\psi\,|\,\hat{x}\,|\,x\rangle\,\langle x\,|\,\phi\rangle\,dx = \int_{-\infty}^{\infty}\psi^*(x)x\phi(x)dx \tag{0.0.58}$$

Clearly then, the position operator is Hermitian. Indeed:

$$\left\langle x_1\,\middle|\,\hat{x}^\dagger\,\middle|\,x_2\right\rangle = \langle x_x\,|\,\hat{x}\,|\,x_1\rangle^* = [x_1\delta(x_1-x_2)]^* = x_1\delta(x_1-x_2) = \langle x_1\,|\,\hat{x}\,|\,x_2\rangle \implies \hat{x} = \hat{x}^\dagger \tag{0.0.59}$$

### Momentum operator $\hat{p}$

Let us now investigate linear operators in infinite dimensional spaces. Consider $\hat{D} = \frac{d}{dx}$:

$$\left\langle x\,\middle|\,\hat{D}\,\middle|\,\psi\right\rangle = \left\langle x\,\middle|\,\frac{d\psi}{Dx}\right\rangle = \frac{d\psi(x)}{dx} \tag{0.0.60}$$

so

$$\frac{d\psi(x)}{dx} = \int_{-\infty}^{\infty} \left\langle x \left| \hat{D} \right| x' \right\rangle \langle x' \,|\, \psi \rangle \, dx' = \int_{-\infty}^{\infty} \delta'(x - x')\psi(x')dx' \tag{0.0.61}$$

$$\implies \left\langle x \left| \hat{D} \right| x' \right\rangle = \delta(x - x')\frac{d}{dx'} = \frac{d}{dx}(\delta'(x - x')) \tag{0.0.62}$$

where we used the derivative of the Dirac delta (eq 0.0.53). For sake of brevity we shall use $D_{xx'} = \left\langle x \left| \hat{D} \right| x' \right\rangle$

Let us now discuss the hermiticity of $\hat{D}$. Consider:

$$\hat{D} = \hat{D}^\dagger \implies D_{xx'} = D^*_{x'x} \tag{0.0.63}$$

However note that

$$D_{xx'} = \delta'(x - x') \text{ and } D^*_{x'x} = \delta'(x' - x)^* = \delta'(x' - x) = -\delta'(x - x') = -D_{xx'} \tag{0.0.64}$$

Hence $\hat{D}$ is anti-Hermitian and to make it Hermitian it suffices to multiply by some purely complex number. For dimensional reasons[3], we will multiply by $\frac{\hbar}{i} = -i\hbar$ then:

$$\hat{p} \equiv -i\hbar\frac{d}{dx} \implies p^*_{x'x} = i\delta'(x' - x) = -i\delta'(x - x') = p_{xx'}. \tag{0.0.65}$$

But is it really Hermitian? Consider the following:

$$\begin{aligned}
\langle \psi \,|\, \hat{p} \,|\, \phi \rangle &= -i\hbar \int_{-\infty}^{\infty} \phi^*(x)\frac{\partial \psi}{\partial x}dx \\
&= -i\hbar\left( [\phi^*\psi]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \psi(x)\frac{\partial \phi^*}{\partial x} \right) \\
&= -i\hbar[\phi^*\psi]_{-\infty}^{\infty} + (\langle \psi \,|\, \hat{p} \,|\, \phi \rangle)^*
\end{aligned}$$

so if the wave functions vanish at spatial infinity, then the term in the square brackets vanishes and the momentum operator is Hermitian. The space of all functions normalize to unity or the dirac delta function, that therefore vanish at infinity, is the **Physical Hilbert Space**. All wave functions we will treat in QM belong to this vector space.

**Definition:** MOMENTUM OPERATOR AND PHYSICAL HILBERT SPACE

We define the momentum operator as:

$$\hat{p} \equiv -i\hbar\frac{d}{dx} \tag{0.0.66}$$

which is Hermitian in the physical Hilbert space, the space of all functions satisfying the normalization condition.

[3] multiplying by $\hbar$ gives the right dimensions for momentum

Consider the eigenvalue problem for $\hat{p}$:

$$\hat{p}\,|p\rangle = p\,|p\rangle \implies \langle x\,|\,\hat{p}\,|\,p\rangle = p\,\langle x\,|\,p\rangle \tag{0.0.67}$$

$$\implies \int_{-\infty}^{\infty} \langle x\,|\,\hat{p}\,|\,x'\rangle\,\langle x'\,|\,p\rangle\,dx' = k\,\langle x\,|\,p\rangle \tag{0.0.68}$$

$$-i\hbar\frac{d\psi_p(x)}{dx} = k\psi_p(x) \tag{0.0.69}$$

where $\phi_p(x) \equiv \langle x\,|\,p\rangle$. The general solution to the latter ODE is clearly:

$$\psi_p(x) = Ae^{ipx/\hbar} \implies \langle x\,|\,p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \tag{0.0.70}$$

Here we normalised as follows:

$$\delta(p-p') = \langle p'\,|\,p\rangle = \int_{-\infty}^{\infty}\langle p'\,|\,x\rangle\,\langle x\,|\,p\rangle\,dx = |A|^2\int_{-\infty}^{\infty} e^{\frac{i(p-p')x}{\hbar}}\,dx = 2\pi\hbar|A|^2\delta(p-p') \tag{0.0.71}$$

where we used the Fourier definition of the dirac delta function.

It then follows that:

$$\psi(p) = \langle p\,|\,\psi\rangle = \int_{-\infty}^{\infty}\langle p\,|\,x\rangle\,\langle p\,|\,f\rangle\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ipx/\hbar}\psi(x)dx \tag{0.0.72}$$

$$\psi(x) = \langle x\,|\,\psi\rangle = \int_{-\infty}^{\infty}\langle x\,|\,p\rangle\,\langle p\,|\,x\rangle\,dp = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ipx/\hbar}\psi(p)dp \tag{0.0.73}$$

$$\tag{0.0.74}$$

which are the Fourier transform relations! This means that Fourier transforms are just passages from the position basis $|x\rangle$ to the momentum basis $|p\rangle$.

### Reciprocity between Position and Momentum representations

The Fourier relation between the position and momentum representations suggest a striking symmetry between these two bases in the Hilbert space. Indeed consider the following table summarizing the properties of the position and momentum operators:

Operators such as these are known as **reciprocal operators**.

| Property | Position $\hat{x}$ | Momentum $\hat{p}$ |
|---|---|---|
| Completeness | $\int_{-\infty}^{\infty} \lvert x\rangle \langle x\rvert \, dx = \hat{I}$ | $\int_{-\infty}^{\infty} \lvert p\rangle \langle p\rvert \, dp = \hat{I}$ |
| Orthogonality | $\langle x \,\vert\, x' \rangle = \delta(x - x')$ | $\langle p \,\vert\, p' \rangle = \delta(p - p')$ |
| Change of basis | $\langle x \,\vert\, p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ipx}{\hbar}}$ | $\langle p \,\vert\, x \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{-ipx}{\hbar}}$ |
| Diagonal representation | $\hat{x} = \int_{-\infty}^{\infty} x \lvert x\rangle \langle x\rvert \, dx$ | $\hat{p} = \int_{-\infty}^{\infty} p \lvert p\rangle \langle p\rvert \, dp$ |
| Matrix elements (diag.) | $\langle x \,\vert\, \hat{x} \,\vert\, x' \rangle = x\delta(x - x')$ | $\langle p \,\vert\, \hat{p} \,\vert\, p' \rangle = x\delta(p - p')$ |
| Cross representation | $\hat{x} = \int_{-\infty}^{\infty} \langle p \,\vert\, \hat{x} \,\vert\, p' \rangle \lvert p\rangle \langle p'\rvert \, dpdp'$ | $\hat{p} = \int_{-\infty}^{\infty} \langle x \,\vert\, \hat{p} \,\vert\, x' \rangle \lvert x\rangle \langle x'\rvert \, dxdx'$ |
| Matrix elements (cross) | $\langle x \,\vert\, \hat{p} \,\vert\, x' \rangle = -i\hbar \frac{\partial}{\partial x}(\delta(x - x'))$ | $\langle p \,\vert\, \hat{x} \,\vert\, p' \rangle = i\hbar \frac{\partial}{\partial p}(\delta(p - p'))$ |

# Part I

# Fundamental principles

# Reviewing Classical Mechanics

1

The goal of this chapter will be to revisit some classical results in Lagrangian and Hamiltonian mechanics. These will help us make the leap to Quantum Mechanics in an elegant manner, using the Hamiltonian formulation as our starting point (the passage using Lagrangian mechanics was also done, and it is known as the Feynman path integral formulation).

## 1.1 The first variation: deriving the Euler-Lagrange equation

Consider a system of $n$ particles with coordinates $(x_1, x_2...x_n)$ then we form a $n$-dimensional **configuration space** which configure the system unequivocally. Given a potential $V(x)$, and $x(t_i) = x_i, x(t_f) = x_f$ what makes the actual trajectory $x_{cl}$ in the configuration space special compared to the infinitely many others? This is the question that Lagrange addressed in his formulation.

The first step is to recall that for a given function $f(\boldsymbol{x})$ where $\boldsymbol{x}$ is the position of the particle in the configuration space, then for $\boldsymbol{x}^{min}$ to be a minimum, $\delta f^{(1)} = 0$ for any small variation $\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{\eta}$. Therefore, taylor expanding gives:

$$f(\boldsymbol{x}^0 + \boldsymbol{\eta}) = f(\boldsymbol{x}^0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\Big|_{\boldsymbol{x}^0} \eta_i + \mathcal{O}(\mu_i^2) \tag{1.1.1}$$

and so:

$$\delta f^{(1)} \equiv \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\Big|_{\boldsymbol{x}^{min}} \eta_i = 0 \implies \frac{\partial f}{\partial x_i}(\boldsymbol{x}^{min}) = 0, \ \ \forall i = 1...n \tag{1.1.2}$$

Let us apply this procedure to the following functional:

$$S[x] = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x}) dt \tag{1.1.3}$$

called the **action** which we hypothesize to be an extremum along the classical path $x_{cl}(t)$ and we consider a path in its vicinity $x_{cl}(t) + \eta(t)$ with the same start/end points. Then

$\eta(t_i) = \eta(t_f) = 0$ so that:

$$\begin{aligned} S[x_{cl}(t) + \eta(t)] &= \int_{t_i}^{t_f} \mathcal{L}(x_{cl}(t) + \eta(t), \dot{x}_{cl}(t) + \dot{\eta}(t))dt \\ &= \int_{t_i}^{t_f} \mathcal{L}(x_{cl}(t), \dot{x}_{cl}(t)) + \frac{\partial \mathcal{L}}{\partial x(t)}\Big|_{x_{cl}} \cdot \eta(t) + \frac{\partial \mathcal{L}}{\partial \dot{x}(t)}\Big|_{x_{cl}} \cdot \dot{\eta}(t) + \mathcal{O}(\eta^2, \dot{\eta}^2)dt \\ &= S[x_{cl}(t)] + \delta S^{(1)} + \mathcal{O}(\eta^2, \dot{\eta}^2) \end{aligned}$$

so for $S[x_{cl}(t)]$ to be an extremum, we need as was shown before that:

$$\delta S^{(1)} = \int_{t_i}^{t_f} \left[ \frac{\partial \mathcal{L}}{\partial x(t)}\Big|_{x_{cl}} \cdot \eta(t) + \frac{\partial \mathcal{L}}{\partial \dot{x}(t)}\Big|_{x_{cl}} \dot{\eta}(t) \right] dt = 0 \tag{1.1.4}$$

Integrating the second term by parts gives:

$$\int_{t_i}^{t_f} \left[ \frac{\partial \mathcal{L}}{\partial x(t)} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right]_{x_{cl}} \cdot \eta(t)dt = 0 \tag{1.1.5}$$

We now state two results that will help us simplify this result.

**Lemma 2.1** Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha < \beta$. Then $\exists \zeta \in \mathcal{C}^2(\mathbb{R})$ for which:

$$\begin{cases} \zeta > 0 \; \forall x \in (\alpha, \beta) \\ \zeta = 0 \; \forall x \in \mathbb{R} \setminus (\alpha, \beta) \end{cases} \tag{1.1.6}$$

This allows us to formulate the following theorem:

**Extremal values of inner products**

Assume $\langle \eta, g \rangle = 0, \forall \eta \in H$. If $g$ is continuous over $\mathcal{D}$ then $g = 0$ on $\mathcal{D} = [x_0, x_1]$.

*Proof.* Suppose $g > 0$ for some $c \in \mathcal{D}$, then due to Lemma 2.1 $\exists \alpha, \beta$ such that $x_0 < \alpha < c < \beta < x_1$, $g(x) > 0 \; \forall x \in (\alpha, \beta)$ and $g(x) = 0 \; \forall x \in \mathcal{D} \setminus (\alpha, \beta)$. Then:

$$\langle \eta, g \rangle = \int_{x_0}^{x_1} \eta(x)g(x)dx = \int_{\alpha}^{\beta} \eta(x)g(x)dx > 0 \tag{1.1.7}$$

but by assumption $\langle \eta, g \rangle = 0, \forall g \in H$. Therefore we must have that $g = 0$ everywhere on $\mathcal{D}$. □

An immediate consequence of this theorem comes from noting that (1.1.4) can be written as $\langle \eta, g \rangle = 0$ with $g = \left[ \frac{\partial \mathcal{L}}{\partial x(t)} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right]_{x_{cl}}$. Therefore we reach the famous **Euler-Lagrange** equation:

**Lagrangian Formulation**

Consider a particle in a configuration space moving from $(t_i, x_i)$ to $(t_f, x_f)$, and define the Lagrangian as a function $\mathcal{L} = T - V$ with $T$ as kinetic energy and $V$ as potential energy. Let us then define the functional $S[x(t)]$, called **action** as follows:

$$S[x(t)] = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x}) dt \tag{1.1.8}$$

Then the action is extremal (or sometimes erroneously called minimised) for some $x_{cl}(t)$ satisfying:

$$\left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right]_{x_{cl}} = 0 \tag{1.1.9}$$

which is known as the **Euler-Lagrange equation**.

Note that this formulation reduces down to Newton's Second Law assuming the potential is velocity independent. However, for the electromagnetic force is velocity dependent, and for this the Lagrangian will have to be modified so that the E-L equation reduces to the Lorentz force, the appropriate force law.

We also note that the Lagrangian, as opposed to the Newtonian formulation, is scalar in nature. Also, they are coordinate-independent, so any choice of independent coordinates will yield the same results.

**Definition 1**:Canonical Force and Momentum

For a Lagrangian $\mathcal{L}(q_1, \dot{q}_1, ..., q_n, \dot{q}_n)$, define the canonical momentum and the canonical force conjugate to $q_i$ as :

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \; F_i = \frac{\partial \mathcal{L}}{\partial q_i} \tag{1.1.10}$$

Note that these are not always the same as the forces and momenta encountered in basic classical mechanics. Indeed, any quantity such that

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} \tag{1.1.11}$$

has an associated conjugate force and momentum. For example, torque and angular momentum, in which case $q_i$ would be an angle.

**Definition 2** : Cyclic Coordinates and Conservation Laws

Suppose that $\mathcal{L}$ is independent of a coordinate $q_i$ but dependent on $\dot{q}_i$. Then $q_i$ is said to be a **cyclic coordinate** to which an associated conservation law reads:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0 \tag{1.1.12}$$

## 1.2 The Electromagnetic Lagrangian

The Lorentz force law gives the force acting on a charged particle moving through an electric field **E** and magnetic field **B**:

$$\mathbf{F} = q\Big(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\Big) \tag{1.2.1}$$

Because the magnetic force depends on the velocity, we cannot use the typical $\mathcal{L} = T - V$ mnemonic, we seek a new lagrangian.

**Electromagnetic Lagrangian**

The appropriate Lagrangian which reduces down to the correct force law turns out to be

$$\mathcal{L} = \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} - q\phi + \frac{q}{c}\mathbf{v}\cdot\mathbf{A} \tag{1.2.2}$$

where $\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$ and $\mathbf{B} = \nabla \times \mathbf{A}$.

Indeed, writing down the E-L equation:

$$\frac{d}{dt}(m\mathbf{v}) = -q\nabla\phi + \frac{q}{c}\Big[\nabla(\mathbf{v}\cdot\mathbf{A}) - \frac{d\mathbf{A}}{dt}\Big] \tag{1.2.3}$$

and by the chain rule: $\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{A} = \frac{\partial \mathbf{A}}{\partial t} + \nabla(\mathbf{v}\cdot\mathbf{A}) - \mathbf{v}\times(\nabla\times\mathbf{A})$ so:

$$\mathbf{F} = -q\nabla\phi - \frac{q}{c}\Big[\mathbf{v}\times\mathbf{B}\Big] = q\Big(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\Big) \tag{1.2.4}$$

Although it may seem very tempting, it is incorrect to interpret the second and third term of the LHS in 1.2.1 to be the potential energy of the electromagnetic field. It would be senseless to even define a potential $\Phi$, since a velocity-dependent force is non-conservative making $\Phi$ path dependent. We could solve this problem by finding $\Phi$ along the path of the particle, but to find this path we first need to solve the E-L equation, which involves $\Phi$. Overall, the mathematically correct way to interpret the second and third terms of LHS of the Lagrangian is to call them the **generalised potential**.

## 1.3 Hamiltonian Mechanics

The Hamiltonian formulation reverses the roles of $\dot{q}$and $p$ by introducing the Hamiltonian: $\mathcal{H}(q,p)$ such that:

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{1.3.1}$$

which is completely analogous to $\dot{p}_i = \frac{\partial L}{\partial q_i}$ The transformation $\dot{q}_i \longleftrightarrow p_i$ belongs to the family of **Legendre Transformations**.

**Definition 3:** THE LEGENDRE TRANSFORMATIONS

Consider the function $f(x)$ such that $\frac{df}{dx} = u(x)$ with $u(x)$ invertible. Let us then define:

$$g(u) = u \cdot x(u) - f(x(u)) \implies \frac{dg}{du} = x(u) \tag{1.3.2}$$

One calls $f$ and $g$ **Legendre transforms** of each other, and to transform from one to the other we simply exchange $x \longleftrightarrow u$.

It is then clear that if we define:

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \qquad\qquad p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \tag{1.3.3}$$

then $\mathcal{H}$ and $\mathcal{L}$ are Legendre transforms of each other with $u = p_i$ and $x = \dot{q}_i$. Using 1.3.1 we can then define:

$$\mathcal{H}(q,p) = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}(q,\dot{q}) \tag{1.3.4}$$

where $\dot{q}_i$ is a function of $q, p$.

We then get (note that in the second line we use the independence of $q$ and $p$ to eliminate terms of the form $:\frac{\partial \mathcal{L}}{\partial q_j}\frac{\partial q_j}{\partial p_i}$)

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial p_i} &= \frac{\partial}{\partial p_i}\Big(\sum_j p_j \dot{q}_j - \mathcal{L}\Big) \\
&= \dot{q}_i + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_i} \\
&= \dot{q}_i + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i} \\
&= \dot{q}_i
\end{aligned}$$

as required, as well as:

$$\frac{\partial \mathcal{H}}{\partial q_i} = \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_j} - \frac{\partial \mathcal{L}}{\partial q_j} - sum_j p_j \frac{\partial \dot{L}}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial q_i} = -frac\partial\mathcal{L}\partial q_j = -\dot{p}_i \tag{1.3.5}$$

which are the two canonical equations in Hamiltonian mechanics.

**Hamiltonian Formulation**

Consider a system with $n$ degrees of freedom in a $2n$-dimensional **phase space** with coordinates $(q,p)$. If we define the Hamiltonian to be:

$$\mathcal{H}(q,p) = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}(q,\dot{q}) \tag{1.3.6}$$

then the **Canonical Hamiltonian Equations** are:

$$\frac{\partial \mathcal{H}}{\partial q_i} = -\dot{p}_i, \qquad \frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i \tag{1.3.7}$$

## 1.4 The Electromagnetic Hamiltonian

Recall the Electromagnetic Lagrangian:

$$\mathcal{L} = \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} - q\phi + \frac{q}{c}\mathbf{v}\cdot\mathbf{A} \tag{1.4.1}$$

so that $\mathbf{p} = m\mathbf{v} + \frac{q\mathbf{A}}{c}$. Then using (1.3.3) we find that the Hamiltonian is:

$$\begin{aligned}
\mathcal{H} &= \mathbf{p}\cdot\mathbf{v} - \mathcal{L} \\
&= \frac{\mathbf{p}\cdot(\mathbf{p} - \frac{q\mathbf{A}}{c})}{m} - \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} + q\phi - \frac{q}{c}\mathbf{v}\cdot\mathbf{A} \\
&= \frac{\mathbf{p}\cdot(\mathbf{p} - \frac{q\mathbf{A}}{c})}{m} - \frac{1}{2m}(\mathbf{p} - \frac{q\mathbf{A}}{c})^2 + q\phi - \frac{q}{c}(\mathbf{p} - \frac{q\mathbf{A}}{mc})\cdot\mathbf{A} \\
&= \frac{1}{2m}(\mathbf{p} - \frac{q\mathbf{A}}{c})^2 + \frac{q\mathbf{A}}{mc} + q\phi - \frac{q}{c}(\mathbf{p} - \frac{q\mathbf{A}}{mc})\cdot\mathbf{A} \\
&= \frac{1}{2m}(\mathbf{p} - \frac{q\mathbf{A}}{c})^2 + q\phi
\end{aligned}$$

**Electromagnetic Hamiltonian**

The Hamiltonian describing electromagnetic interactions of a particle of charge $q$ in an electromagnetic field is:

$$\mathcal{H} = \frac{(\mathbf{p} - \frac{q\mathbf{A}}{c})^2}{2m} + q\phi \tag{1.4.2}$$

Now that we have reviewed the state of physics near the end of the 19th century, we are ready to delve into Quantum Mechanics.

# Experimental Motivation

2

The history of Quantum Mechanics can be dated back to the early 1900s, when theoretical physics seemed to be at a dead end. As Lord Kelvin put it: **There is nothing new to be discovered in physics now. All that remains is more and more precise measurement**. However, this could not be farther from the truth. indeed, it was about at this time when certain experiments had been performed whose results seemed to be inexplicable using classical physics. The most important were:

(i) Spectral lines

(ii) Blackbody radiation

(iii) the Photoelectric effect

(iv) Radioactive decay phenomena

## 2.1 Physics is in Chaos

### Spectral Lines

The first experiments related to the composition of chemical elements and the colors they emitted were performed in the 1860s by Gustav Kirchhoff and Robert Bunsen. They found that to each element was a designated spectral footprint or **spectrum**.

By 1885, Balmer had discovered empirically that the spectral lines of wavelength $\lambda$ could be written as:

$$\frac{1}{\lambda} = R\Big(\frac{1}{4} - \frac{1}{n^2}\Big) \tag{2.1.1}$$

where $R \approx 1.097 \times 10^5 cm^-1$ is the Rydberg constant, andd $n$ is an integer greater than 2. However a key question still remained unanswered: why where the spectra made of discrete lines, and not gradients? No classical theory could explain this result.

### Blackbody radiation

Several experiments by Rayleigh, Jeans and others had shown that the intensity $I$ of electromagnetic radiation emitted by a blackbody at temperature $T$ of wavelength $\lambda$ followed

a certain distribution shown below. No physical model for a blackbody could fit the data consistently.

On the one hand, there was the **Rayleigh-Jeans approximation**: $I(\lambda) = C\frac{T}{\lambda^4}$ which worked well for long wavelengths, but horrendously at short wavelengths such as UV.

On the other hand, we have **Wien's approximation**: $I(\lambda) = C\exp^{\frac{k}{\lambda T}}$ which instead worked well for short wavelengths, and blew up at longer wavelengths.

Neither however were able to consistently describe blackbody radiation over the entire spectrum, this was known as the **UV catastrophe**.

### The Photoelectric effect

Since 1888 another effect, called the photoelectric effect, had been under careful investigation. The effect consisted in the positive ionization of neutral metal zinc bars when irradiated with UV rays. Not only, experiments also ascertained that this emission of electrons was dependent on the frequency of incident radiation, which according to classical electromagnetism, would not be the case.

### Radioactivity

The final nail in the coffin has to do with radioactive decay, discovered by Ernest Rutherford in 1900. Despite being completely identical, atomic nuclei will not decay at identical times, but instead have a specific probability of decaying in a time interval $\delta t$. This probability is given by $\lambda\delta t$ with $\lambda$ defined as the **decay constant**, and varies across different isotopes.

Consider the following experiment: we take a box containing $N_0$ identical isotopes, and measure the number of remaining nuclei $N(t)$ at every instant $t$. On average:

$$\delta N = -\lambda N\delta t \implies \frac{dN}{dt} = -\lambda N \tag{2.1.2}$$

whose solution is:

$$N(t) = N(0)e^{-\lambda t} \tag{2.1.3}$$

One could know everything about the nucleus of a specific atom, and yet it would be impossible to determine when it would decay. This undeterministic answer, which goes entirely against classical physics, deeply troubled physicists at the time. These experiments seem to suggest two fundamental aspects of nature: **discreteness** and **indeterminism**.

## 2.2 Discreteness: Bohr and Planck's radical solutions

These problems would only be resolved with the radical idea of energy quantization, proposed by Planck in 1900. Indeed, to create a mathematical model that would fit the experimental data for blackbody radiation, he hypothesized that the energy in this radiation was quantized, and could only take discrete values given by:

$$E = h\nu \tag{2.2.1}$$

where $h \approx 6.626 \times 10^{-34} Js^{-1}$ is Planck's constant and $\nu$ is the frequency of the radiation. These **energy quanta** are now known as photons. More often, $\hbar = \frac{h}{2\pi}$ is used, in which case:

$$E = \hbar\omega \tag{2.2.2}$$

with $\omega$ defined as the angular frequency. It then follows that the momentum of a photon may be expressed as:

$$|\mathbf{p}| = \frac{E}{c} = \frac{h}{\lambda} = \hbar k \tag{2.2.3}$$

where $k = \frac{2\pi}{\lambda}$ is the wave-number. More generally, Sommerfield and Wilson, noticing the regularity in the quantization of quantities, proposed:

**The Quantization Condition**

For a system of momenta $p_i$ and corresponding coordinates $q_i$:

$$\oint p_i dq_i = n\hbar \tag{2.2.4}$$

where $n$ is an integer. Energy and momentum quantization then give:

$$E = \hbar\omega, \quad \mathbf{p} = \hbar\mathbf{k} \tag{2.2.5}$$

This quantization however did not only solve the UV catastrophe, but also explained the photoelectric effect, and the discreteness in atomic spectra.

### Bohr's Model

Consider the quantization condition for the coordinate $\phi$ and corresponding canonical angular momentum $l$. Then, if we define the $z$-axis to lie normal to the plane in which the nucleus and electron lie. Then we find:

$$\oint l_z d\phi = l \int_0^{2\pi} d\phi = nh \implies l = n\hbar \tag{2.2.6}$$

So, for a hydrogen-like atom of atomic number $Z$, we envision an electron moving in circular orbits around the nucleus, analogous to the solar system. These orbits however are not continuous. This is because as an electron orbits, it is accelerating and decelerating, emitting electromagnetic radiation and losing energy. This forces the electron to move closer and closer to the nucleus, eventually collapsing. To solve this issue, Bohr hypothesized that the energy levels were discrete, and corresponded to orbits where the electron would not collapse into the nucleus. We can thus define the potential to be:

$$V(r) = -\frac{Ze^2}{r} \tag{2.2.7}$$

with $r$, the radius from the nucleus, such that:

$$l = mvr = n\hbar$$

which is why the orbits are discrete. Using the centripetal force equation, one finds that:

$$\frac{1}{2}mv^2 = \frac{1}{2}\frac{Ze^2}{r^2} \implies E = -\frac{1}{2}mv^2 = -\frac{1}{2}\frac{Ze^2}{r^2} \tag{2.2.8}$$

multiplying by $mr^2$ gives the values of the orbit radii:

$$m^2v^2r^2 = mZe^2r \implies r = \frac{\hbar}{mZe^2}n^2 \tag{2.2.9}$$

Interestingly, the distance from the nucleus increases quadratically. We also remark that the **Bohr radius**, the radius of the first orbit in a hydrogen atom, is:

$$a_0 = \frac{\hbar}{me^2} \implies r = \frac{n^2}{Z}a_0 \tag{2.2.10}$$

We can then write the energy levels of the hydrogen-like atom as:

$$E = -\frac{m^2e^2}{2a_0}\frac{1}{n^2} \tag{2.2.11}$$

This means that if an electron jumps from an energy level $E_{n'}$ to a lower level $E_n$, then the energy of the emitted photon must be equal to the energy difference:

$$\frac{hc}{\lambda} = E_{n'} - E_n \tag{2.2.12}$$

The Bohr model can be summarized as follows:

**Bohr's atomic model**

The orbit radii of an electron for a hydrogen-like atom of atomic number $Z$ are given by:

$$r = \frac{\hbar}{mZe^2}n^2 = \frac{n^2}{Z}a_0 \tag{2.2.13}$$

and the corresponding energies are:

$$E = -\frac{m^2e^2}{2a_0}\frac{1}{n^2} \tag{2.2.14}$$

## 2.3 Indeterminacy: corpuscular and wave-like light

Although we have been using the terms **light** and **particle**, it is about time we actually define these two objects.

**Definition 1: Waves and Particles**

A **particle** is a localized bundle of energy and momentum, described at any instant by $(q, p)$ (or $(q, \dot{q})$). Given a set of initial conditions, the entire future of the system can be predicted with perfect accuracy using Lagrangian or Hamiltonian mechanics.

There is no interference between particle.
A **wave** is a disturbance spread over space, and is described by a wave function $\psi(\mathbf{r}, t)$. Solving the wave equation with appropriate boundary conditions we can find the wave at all future times. There is interference between waves.

We will be specially interested in waves that are periodic with respect to both space and time, called **plane waves**. They are expressible as:

$$\psi(x,t) = A\exp\left[2\pi i\Big(\frac{x}{\lambda} - \frac{t}{T}\Big)\right] \tag{2.3.1}$$

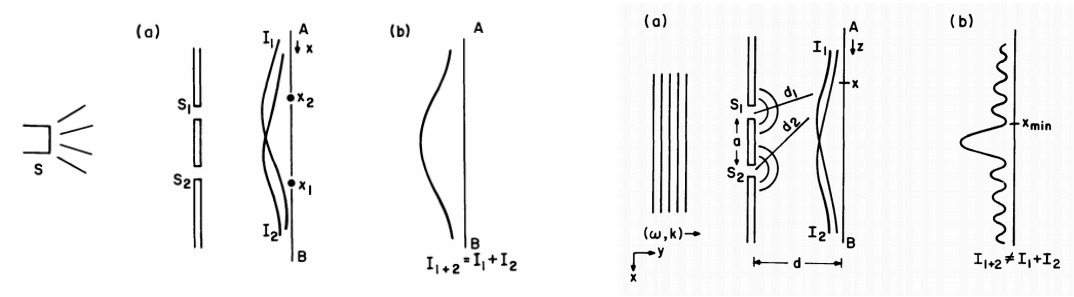
where $\lambda$ is the wavelength and $T$ its period. So the wave repeats itself every $T$ seconds in time, and every $\lambda$ meters in space. Alternatively, defining the wavenumber $k = \frac{2\pi}{\lambda}$ and angular frequency $\omega = \frac{2\pi}{T}$ we can then write:

$$\psi(\mathbf{r}, t) = Ae^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \tag{2.3.2}$$

We have seen how light behaves as a particle by looking at energy quanta, and phenomena such as spectral lines, the photoelectric effect and blackbody radiation. Let us now look at how light may behave as a wave.

### Light as a wave

Consider the following experiment, we take a wave $\psi = Ae^{i(ky-wt)}$ incident on a screen with two slits, a distance $a$ apart, with a row of detectors a distance d from the slits. First we




(a) The interference patterns for the double slit experiment using particles

(b) The interference patterns for the double slit experiment using waves

keep $S_1$ open, so that the wave will propagate and produce an intensity pattern $I_1 = |\psi_1|^2$. Similarly if we keep $S_2$ open we get $I_2 = |\psi_2|^2$. In both cases the wave is smooth, with a clear peak near the slit.

Now we keep both slits opened. One would expect the new pattern to be $I_{1+2} = I_1 + I_2$, but this is not the case. Instead $I_{1+2} = |\psi_1 + \psi_2|^2$. This effect is due to the fact that the waves have to travel different distances to arrive at some given point $x$ on the detector. The difference in phase between the two waves then produces an interference.

More specifically, for maximum intensity we need the difference in the distances travelled to be integer multiples of the wavelength, whereas for minimum intensity we need the difference in the distance travelled to be an odd multiple of the half wavelength:

$$|d_1 - d_2| = \begin{cases} n\lambda & \text{for constructive interference} \\ (2n+1)\frac{\lambda}{2} & \text{for destructive interference} \end{cases} \tag{2.3.3}$$

Let us now repeat the very same experiment using particles, with $I(x)$ defined as the number of particles detected per second. The pattern with just $S_1$ or $S_2$ open are similar, but when they are both open, classical mechanics predicts that $I_{1+2} = I_1 + I_2$ unlike waves. To avoid interactions between the particles themselves, they are sent one by one.

When we perform the double slit experiment with light, an interest result arises.

When light passes through the slits, in the case when only one slit is open, as well as when both slits are open, the interference pattern of a wave is produced.

Now we repeat the experiment, this time using very low intensity light. Starting with just $S_1$ open, the energy exchange will not be continuous anymore. Instead, the energy will arrive in bursts. After long enough, the envelope of $I_1$ is clearly visible. If we now open both slits however, we will not find $I_{1+2} = I_1 + I_2$ as is expected, but instead we get an interference pattern for a wave! This rules out the possibility for light bundles, photons, to move as particles, since if this were true, then a photon going through slit 1 would not care about whether or not slit 2 is open, unlike a wave which interferes with itself. In other words, consider a point of destructive interference, so that the number of photons arriving are higher when either $S_1$ or $S_2$ is open than when both are open. Then why would adding another pathway for photons by opening another slit **reduce** the photon count? Moreover, since the particles are coming out one at a time, we can't state that the photons through slit 1 are interfering with those through slit 2. It's as if the photons already knew where the others would distribute themselves!

So is light a wave, or is it a particle? This is the so-called $wave - particleduality$ of light.

## 2.4 Indeterminacy: corpuscular and wave-like electrons

The several experiments explained in the previous sections seem to support both a wave-like and particle-like nature of matter. This was further evinced when diffraction was achieved with electrons, suggesting that the wave-particle duality that had been already accepted for electromagnetic radiation is perhaps relevant to matter as well. This was what De Broglie conjectured in 1924, providing the De Broglie relations.

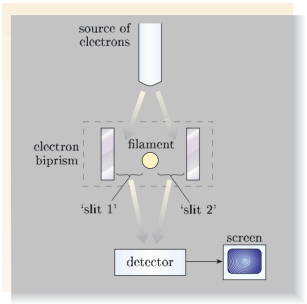
**Definition 2**: De Broglie relations

For any object we can associate a wavelength $\lambda_{DB}$, called the **De Broglie** wavelength such that:

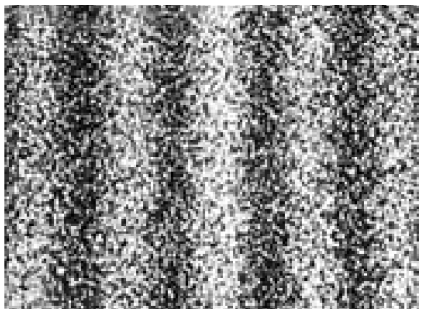
$$\lambda_{DB} = \frac{h}{p} \tag{2.4.1}$$

Further evidence for the wave-particle duality of electrons was collected by Tonomura in 1989.

A beam of electrons is shot against a **biprism**, which plays the role of the two slits. The electrons can pass on either side of the filament, and then detected. The electrons produced individual spots on the detector, but overall created an interference pattern.




(a) Experiment performed by Tonomura



(b) Interference pattern formed from the wave particle duality of electrons

So again we ask ourselves: is an electron a particle or a wave?

To answer this question, we will have to define the concept of superposition.

## 2.5 Understanding superposition

Consider two random properties of electrons:

(i) Colour: black or white

(ii) Hardness: hard or soft

which can be measured using the appropriate colour and hardness boxes. Obviously, if an electron comes out of a colour box as black, then is is black, and when it is again put in another colour box, there is $100\%$ probability of it going out of the black exit. The same goes for hardness. Colour and hardness are repeatable.

One natural question to ask is: are colour and hardness correlated? We will suppose that if an electron going through a colour box comes out white, then if it is then sent into a hardness box it has a $50\%$ chance of being soft and $50\%$ chance of being hard.

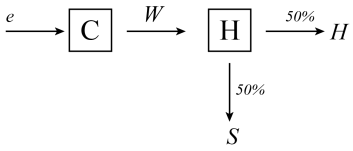



**Figure 2.3**. Hardness and colour are not correlated as shown

Now we perform the following experiment. A beam of electrons is sent to a colour box, and the white output is then sent to a hardness box, where the soft output is sent to another

colour box. Obviously, one expects the output to be $100\%$ white, but instead it is $50\%$ white and $50\%$ black!

This means that when we pass the electron beam through the colour box, and then the hardness box, we loose all information about the colour! Similarly, if we had sent the electron beam through the hardness box, and then the colour box, we would have lost all information on hardness. In general, we can't build a box giving both hardness and colour!

If we did try building such a box, took for example the soft-black output and sent it into a colour box, one would find that only $50\%$ are black!

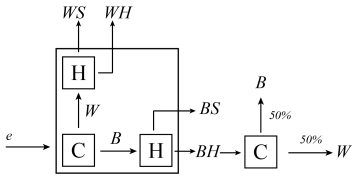



**Figure 2.4**. Colour hardness box and the inconsistency in observing hardness and colour simultaneously

It is therefore a fundamental aspect of nature, called the **uncertainty principle**, that it is impossible to measure certain couples of observables simultaneously, not because **we** can't measure them, but because it is meaningless to do so, just like the colour hardness box.

We now add mirrors to our apparatus, which will allow us to perform a set of experiments that will drill in the concept of superposition even further. These mirrors do not affect the observables in any way, if a beam of electrons is $100\%$ white before the mirror, it is still $100\%$ white after being reflected.

## Experiment 1

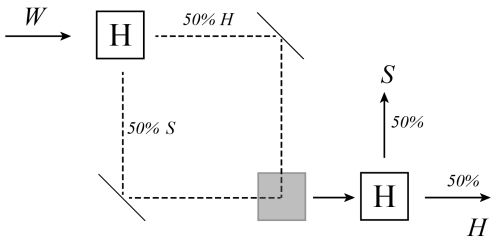



**Figure 2.5**. Experiment 1, we measure the hardness and softness of white electrons twice

Clearly, one would predict the electrons initially coming out hard will come out hard again, since colour measurement is repeatable. Same goes with white. Luckily, this prediction is true!

## Experiment 2

Again, we expect to find $50\%$ black and $50\%$ white. This corresponds with reality. W

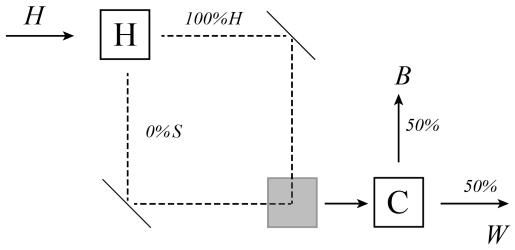



**Figure 2.6**. Experiment 2, we measure the colour of hard electrons

### Experiment 3

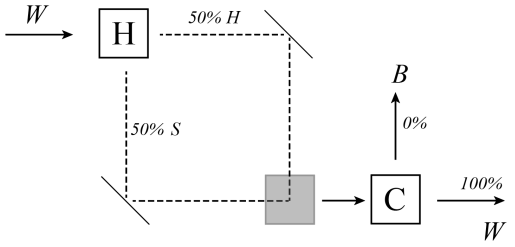



**Figure 2.7**. Experiment 3, we measure the colour of white electrons going through a hardness box first

Using our preceding arguments, one would expect $50\%$ black and $50\%$ white. However, this is wrong! All of the electrons come out white! No matter whether they are hard or soft, they are measured to be white.

### Experiment 4

We modify the apparatus slightly by inserting a small movable barrier, which absorbs all electrons incident upon it.

Now we run experiment 3 using the barriers.

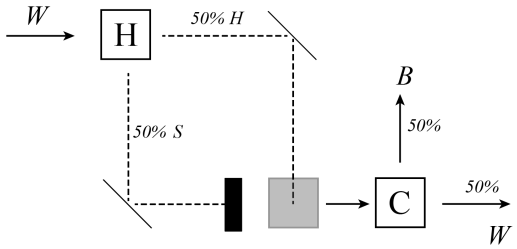



**Figure 2.8**. Experiment 4 is similar to experiment 3 but with a barrier

Classically, one would expect from locality that if we place a barrier in the soft path, the hard electrons won't care. So if all the hard electrons in experiment 3 came out white, then they should come out white again in this experiment.

However, in reality the output is $50\%$ black and $50\%$ white! How did the electrons through the hard path "know" that a barrier was placed in the soft path? This is a manifestation of **non-locality**, which we will see later on with the collapse of the wave function.

### Conclusions

So what path did the individual electron take in experiment 3? It can't have taken the hard path, because that was experiment 4 which had 50-50 ratio. However, it can't have taken the soft path either for the same reason. Could it have taken both paths? No, if we place a detector on both paths, it has been empirically verified that each electron always always takes one of the paths. You never see half an electron [1]. So it took neither? If we place a barrier in both paths however, nothing comes out, so it can't be this either.

**Definition 3** : SUPERPOSITION STATE

Electrons exist in this new mode of being, it is a **superposition**. For example, an initially white electron is in a superposition of $50\%$ hard and $50\%$ soft. It is neither hard, nor soft, nor both, nor either. So, having a definite colour means having an indefinite hardness, but being in a superposition of hard and soft.

## 2.6 The wave function

From all of these results, Born extrapolated one major idea, that of the wave function. He concluded that with each system is an associated complex wave $\psi(\mathbf{r}, t)$ which encoded all the information about this system.

**Born's Rule**

The probability of finding a particle at time $t$ in a volume $\delta V$ centered at $\mathbf{r}$ is given by:

$$P(\delta V) = |\psi(\mathbf{r}, t)|^2 \delta V \tag{2.6.1}$$

We must now revisit our definition of a particle.

**Definition 4:** QM particle

In quantum mechanics, a particle is an object which is found in only one place when its position is measured.

In the case of the electron diffraction experiment, this definition is self-evident, because the individual electrons appear as dots, despite travelling as a wave. What is truly happening is the detection of the electron wave results in the collapse of the wave, so that only one of the detector pixels is triggered, but leaving all the other pixels untouched. It's as if the responding pixel told the other pixels not to respond, this is **quantum non-locality**. This collapse is believed to be an instantaneous event [2].

[1] how do we know that this is not the wave function collapsing?

[2] see https://iopscience.iop.org/article/10.1088/1742-6596/410/1/012153/pdf

### The De Broglie wave function

The wave function for a free particle is called the **De Broglie wave function**. For a free particle moving in the positive x-direction with momentum $p$, this wave function must be a momentum eigenstate, since it is a state of definite momentum. We know from chapter 1 that in the position representation this wave function can be expressed as a plane wave:

$$\psi(x) = Ae^{\frac{ipx}{\hbar}} \tag{2.6.2}$$

As it turns out, to introduce time evolution, it suffices to multiply by $e^{\frac{iEt}{\hbar}}$ so that:

$$\psi_{dB}(x,t) = Ae^{i(px-Et)/\hbar} = Ae^{i(kx-wt)} \tag{2.6.3}$$

is the De Broglie wave function, with the standard relation $k = \frac{p}{\hbar}$ and $\omega = \frac{E}{\hbar}$, where $E$ in this case is only the kinetic energy of the particle.

An important consequence of this is that the particle has the same probability of being found in any small interval of length $\delta x$, since $P(\delta x) = |\psi(x,t)|^2\delta x = |A|^2\delta x$ which is constant along the entire x-axis.

### 2.6.1 Galilean transformations

It is interesting to see how the De Broglie relations hold for different Galilean observers. Consider two frames $S$ and $S'$ with aligned $x$-axes and with $S'$ moving to the right ($+x$ direction) relative to $S$ with constant non-relativistic velocity $v$. If the origins coincide at $t = 0$ then the frames are related by:

$$x' = x - vt,\ t' = t \tag{2.6.4}$$

Now a particle with momentum $p = m\tilde{v}$ in the $S$ frame has velocity:

$$\tilde{v}' = \tilde{v} - v \implies p' = p - mv \tag{2.6.5}$$

in the $S'$ frame. Therefore:

$$\lambda' = \frac{h}{p'} = \frac{h}{p - mv} \neq \lambda \tag{2.6.6}$$

For normal waves, Galilean observers will find discrepancies in the frequency but not the wavelength. Indeed, for waves like sound or water, the two observers need only to take a picture of the wave at the same instant to find the wavelength. Clearly then a wave that is not invariant under Galilean transformations must not be directly measurable.

## 2.7 Explaining electron interference

Let us denote the wave function from the two slits as $\psi(\mathbf{r},t)$, and the wave function when only slit 1 or slit 2 is open by $\psi_1(\mathbf{r},t),\ \psi_2(\mathbf{r},t)$ respectively. The probability of finding an electron in $\delta V$ is then:

$$P(\delta V) = |\psi_1 + \psi_2|^2\delta V = |\psi_1|^2 + |\psi_2|^2 + \underbrace{\psi_1^*\psi_2 + \psi_2^*\psi_1}_{\text{interference term}} \tag{2.7.1}$$

It follows that destructive interference occurs at points where
$phi_1 = \phi_2 + \pi$, and constructive interference occurs at points where $\phi_1 = \phi_2$. Here $\phi_{1,2}$ are the phases of the two wave functions.

One fundamental aspect about the electron interference pattern is that when we observe which slit the electron goes through, the interference pattern is destroyed, and we instead get either $\psi_1$ or $\psi_2$.

# The Four Postulates

3

The aim of this chapter will be to present the four fundamental postulates of Quantum Mechanics and discuss their implications.

## 3.1 The Postulates

**The Four Postulates of QM**

For a single-particle system in one dimension the postulates in quantum mechanics are:

I. The state of a particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space that is independent of time (this is called the **Schrödinger picture**).

II. The independent variables $x$ and $p$ of classical mechanics are represented by Hermitian operators $\hat{x}$ and $\hat{p}$ with the following matrix elements in the eigenbasis of $\hat{x}$:

$$\langle x|\,\hat{x}\,|x'\rangle = x\delta(x-x') \tag{3.1.1}$$
$$\langle x|\,\hat{p}\,|x'\rangle = -i\hbar\delta'(x-x') \tag{3.1.2}$$

III. If the particle is in a state $|\psi(x,t)\rangle$, the measurement of a variable represented by $\Omega$ will yield one of the eigenvalues $\omega$ with probability $\mathbb{P}(\omega) \propto |\,\langle\omega\,|\,\psi\rangle\,|^2$ and the state of the system will change to $|\omega\rangle$.

IV. The state vector $\psi(t)$ obeys the **Schrödinger equation**:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle \tag{3.1.3}$$

where $\hat{H}$ is the Hamiltonian operator in QM.

The first three postulates provide a description of how to interpret a system's state at a given time, like taking a snapshot of it. Instead, the last postulate tells us how the state evolves in time.

## 3.2 The First Postulate: what is a quantum state

The first postulate tells us that the quantum state of a system is described by a complex vector in a Hilbert space, the state vector $|\psi\rangle$. Because the system is described by a wave,

the wave function, the number of degrees of freedom is no longer finite like in classical mechanics, but infinite. This is highlighted by the fact that the state vector belongs to a Hilbert space.

It then follows from the definition of vector space that if$|\psi\rangle$ and $|\psi\rangle$ represent possible states, then any linear combination $\alpha\,|\psi\rangle + \beta\,|\psi\rangle$ is also a suitable combination. This is the **principle of superposition** we introduced in chapter 2.

**Superposition of states**

If$|\psi\rangle$ and $|\psi\rangle$ represent possible quantum states, then any linear combination $\alpha\,|\psi\rangle + \beta\,|\psi\rangle$ for $\alpha, \beta \in \mathbb{C}$ also represents a quantum state.

As was seen earlier, the probability that a measurement of the position of the particle yields a result in an interval $dx$ is given by $|\psi|^2 dx$. Therefore, if we integrate over $\mathbb{R}$ we must get 1, since the particle must be **somewhere**.

**Conditions for a wave function**

For a suitable state vector $|\psi(x,t)\rangle$:

$$||\psi||^2 = \int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = 1 \tag{3.2.1}$$

which implies :

(i) the square integrability of the wave function so that:

$$\lim_{x\to\infty} \psi(x,t) = 0, \tag{3.2.2}$$

(ii) the continuity of the wave function,

(iii) the boundedness of the wave function

(iv) the boundedness of the first derivative of the wave function at infinity:

$$\lim_{x\to\infty} \frac{\partial\psi(x,t)}{\partial x} < \infty, \tag{3.2.3}$$

(v) the single-valuedness of the wave function.

## 3.3 Second and Third Postulates: Measuring Observables

The second and third postulates provide information on observables and their measurement in QM.

Suppose we are given some dynamical variable $\omega(x,p)$, which we want to measure for a system in a state $|\psi\rangle$. The first step is to construct the quantum operator:

$$\hat{\Omega} = \omega(x \to \hat{x}, p \to \hat{p}) \tag{3.3.1}$$

The second step is to find the orthonormal eigenbasis $\{|\omega_i\rangle\}$ with eigenvalues $\omega_i$ of $\Omega$. The third step is to expand $|\psi\rangle$ in this basis and find the coordinate vector:

$$|\psi\rangle = \sum_i |\omega_i\rangle \langle\omega_i \,|\, \psi\rangle \tag{3.3.2}$$

Finally, the probability $\mathbb{P}(\omega)$ that the eigenvalue $\omega$ is measured is proportional to the coordinate/projection of $\psi$ along $\omega$. So:

$$\mathbb{P}(\omega) \propto |\langle\omega \,|\, \psi\rangle|^2 = |\langle\psi \,|\, \omega\rangle|^2 = \langle\psi \,|\, \omega\rangle \langle\omega \,|\, \psi\rangle = \langle\psi \,|\, \hat{p}_\omega \hat{p}_\omega \,|\, \psi\rangle = \langle\hat{p}_\omega\psi \,|\, \hat{p}_\omega\psi\rangle \tag{3.3.3}$$

since the projection operator along $\omega$ is $\hat{p}_\omega = |\omega\rangle\langle\omega|$.

Two important remarks must be made:

(i) If the quantum state is one of the eigenstates $|\omega_i\rangle$ then the measurement will definitely yield the measurement $\omega_i$.

(ii) Consider the superposition of two two eigenstates $|\omega_1\rangle$ and $|\omega_2\rangle$, such that the measurement of an observable $\omega$ yields the measurements $\omega_1, \omega_2$ respectively:

$$|\phi\rangle = \frac{\alpha\,|\omega_1\rangle + \beta\,|\omega_2\rangle}{(|\alpha|^2 + |\beta|^2)^{\frac{1}{2}}} \tag{3.3.4}$$

which is another suitable state vector (in general not an eigenstate). If we measure the same observable for $|\psi\rangle$ we can only find the values $\omega_1, \omega_2$ with respective probabilities [1]:

$$P_1 = \frac{|\alpha|^2}{|\alpha|^2 + |\beta|^2},\ P_2 = \frac{|\alpha|^2}{|\alpha|^2 + |\beta|^2} \tag{3.3.5}$$

No other measurements can be made. This is very different from our experience in classical mechanics, where any measurement could be made.

### How to construct the operator $\hat{\Omega}$

Because Hermitian operators generally do not commute, we must be careful when constructing them. For example, consider $\omega(x,p) = xp$, is the corresponding operator $\hat{\Omega} = \hat{x}\hat{p}$ or $\hat{\Omega} = \hat{p}\hat{x}$? The rule of thumb is use the sum:

$$\hat{\Omega} = \frac{\hat{x}\hat{p} + \hat{p}\hat{x}}{2} \tag{3.3.6}$$

### What if $\hat{\Omega}$ is degenerate?

Let $\omega_1 = \omega_2 = \omega$ be two eigenvalues of the operator $\hat{\Omega}$. We then select two orthonormal vectors $|\omega, 1\rangle$and $|\omega, 2\rangle$, so that:

$$\mathbb{P}(\omega) = |\langle\omega, 1 \,|\, \psi\rangle|^2 + |\langle\omega, 2 \,|\, \psi\rangle|^2 \tag{3.3.7}$$

[1] to find them simply dot to the left with $\langle\omega_1|$ anad $\langle\omega_2|$ and use the orthonormality of the eigenstates

Alternatively we can use the expression for the projection $\hat{p}$ along the eigenspace formed by $|\omega, 1\rangle$and $|\omega, 2\rangle$, and find $\mathbb{P}(\omega) \propto \langle\psi \,|\, \hat{p} \,|\, \psi\rangle$.

### Collapse of the State Vector

Part of postulate three asserts that if we measure the variable represented by $\hat{\Omega}$ of a system in a state $|\psi\rangle$ in a superposition of eigenstates:

$$|\psi\rangle = \sum_{\omega} |\omega\rangle \langle\omega \,|\, \psi\rangle \tag{3.3.8}$$

then it will collapse into one of the eigenstates $|\omega\rangle$ instantaneously. The acting operator is the projection operator $\hat{p}_\omega$ which projects the state vector in the corresponding eigenspace of eigenstate vectors.

### Expectation values

Consider a large sample of particles in a state $|\psi\rangle$, the average measurement obtained by operating $\hat{\hat{\Omega}}$ is the **expectation value** $\langle\hat{\Omega}\rangle$. Calculating:

$$\begin{aligned}\langle\hat{\Omega}\rangle &= \sum_i \mathbb{P}(\omega_i)\omega_i = \sum_i |\langle\omega_i \,|\, \psi\rangle|^2 \omega_i \\ &= \sum_i \langle\psi \,|\, \omega_i\rangle \langle\omega_i \,|\, \psi\rangle = \sum_i \left\langle\psi \middle|\, \hat{\Omega} \,\middle|\, \omega_i\right\rangle \langle\omega_i \,|\, \psi\rangle \\ &= \left\langle\psi \middle|\, \hat{\Omega} \,\middle|\, \psi\right\rangle\end{aligned}$$

since $I = \sum_i |\omega_i\rangle \langle\omega_i|$. For an eigenstate $\omega$, the expectation value is obviously $\omega$, as required.

**Expectation value**

The expectation value of an operator $\hat{\Omega}$ acting on a state vector $|\psi\rangle$ is given by:

$$\langle\hat{\Omega}\rangle = \left\langle\psi \middle|\, \hat{\Omega} \,\middle|\, \psi\right\rangle \tag{3.3.9}$$

## 3.4 The Fourth Postulate: Schrödinger's Equation

The fourth postulate tells us that the state vector obeys the Schrödinger's Equation (SE for short):

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \,. \tag{3.4.1}$$

The three parts to solving the SE are:

(i) Setting up the Hamiltonian

(ii) General approach to solution

(iii) Choosing a basis for the solution

### Setting up the Hamiltonian

We must make the substitution to construct the Hamiltonian operator:

$$\hat{H} \leftrightarrow \mathcal{H}(x \to \hat{x}, p \to \hat{p}). \tag{3.4.2}$$

For example, the harmonic oscillator has classical Hamiltonian has corresponding Hamiltonian operator:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \leftrightarrow \hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{3.4.3}$$

For a single particle, the Hamiltonian particle in the position representation:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{3.4.4}$$

where $\langle x \,|\, V(\hat{x}) \,|\, \psi\rangle = V(x)\psi(x)$. Does the Hamiltonian commute with the position operator? Consider:

$$[\hat{x}, \hat{H}] = [\hat{x}, \frac{\hat{p}^2}{2m}] + [\hat{x}, V(\hat{x})] = \frac{[\hat{x}, \hat{p}]\hat{p} + \hat{p}[\hat{x}, \hat{p}]}{2m} \tag{3.4.5}$$

we therefore need to find $[\hat{x}, \hat{p}]$. Then:

$$\langle x| \,[\hat{x}, \hat{p}]\, |\psi\rangle = \langle x| \,(\hat{x}\hat{p} - \hat{p}\hat{x}\,|\psi\rangle = -i\hbar\Big(x\frac{\partial\psi}{\partial x} - \frac{\partial(x\psi)}{\partial x}\Big) = i\hbar\,\langle x \,|\, \psi\rangle \tag{3.4.6}$$

from which we find the **canonical commutation relation**

$$[\hat{x}, \hat{p}] = i\hbar \tag{3.4.7}$$

**Canonical Commutator**

Two operators $\hat{A}, \hat{B}$ are said to be canonically conjugate if:

$$[\hat{A}, \hat{B}] = \pm i\hbar \tag{3.4.8}$$

One set of **canonically conjugate** operators are the position and momentum operators:

$$[\hat{x}, \hat{p}] = i\hbar \tag{3.4.9}$$

.

It follows that $[\hat{x}, \hat{H}] = \frac{i\hbar}{m}\hat{p}$.

### General Approach to the Solution

Initially consider a time-independent Hamiltonian:

$$i\hbar\,|\dot{\psi}\rangle = \hat{H}\,|\psi\rangle \tag{3.4.10}$$

which is analogous to the equation $|\ddot{x}\rangle = \hat{\Omega}\,|x\rangle$ encountered in the mathematical preliminaries. We can then write:

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle \tag{3.4.11}$$

using the propagator formulation. To construct it, we must find the normalized eigenkets $|E\rangle$ of $\hat{H}$ obeying the Time-independent SE (TISE).

**Time-Independent Schrödinger Equation**

The TISE equation is:

$$\hat{H}\,|E\rangle = E\,|E\rangle \tag{3.4.12}$$

Assume we have solved the TISE, and have found $|E\rangle$. Then:

$$|\psi(t)\rangle = \sum_E |E\rangle\,\langle E\,|\,\psi(t)\rangle \equiv \sum_E a_E(t)\,|E\rangle \tag{3.4.13}$$

where $a_E \equiv \langle E\,|\,\psi(x,t)\rangle$ is the probability of measuring the eigenstate $|E\rangle$. We then operate $(i\hbar\frac{\partial}{\partial t} - \hat{H})$ and find that:

$$0 = i\hbar\frac{\partial}{\partial t}\,|\psi\rangle - \hat{H}\,|\psi\rangle = \sum_E (i\hbar\dot{a}_E - Ea_E)\,|E\rangle \implies i\hbar\dot{a}_E = Ea_E \tag{3.4.14}$$

which gives the solution:

$$a_E(t) = \overbrace{\langle E\,|\,\psi(0)\rangle}^{a_E(0)} e^{-iEt/\hbar} \implies |\psi(t)\rangle = \sum_E |E\rangle\,\langle E\,|\,\psi(0)\rangle\,e^{-iEt/\hbar} \tag{3.4.15}$$

hence the propagator operator can be written as:

$$\hat{U}(t) = \sum_E |E\rangle\,\langle E|\,e^{-iEt/\hbar} = e^{-iHt/\hbar} \tag{3.4.16}$$

where the latter expression works provided the exponential series converges. The corresponding solutions are called stationary states, and calculating the probability distribution for a variable $\omega$:

$$\mathbb{P}(\omega,t) = |\,\langle\omega\,|\,\psi\rangle\,|^2 = |\left\langle\omega\,\middle|\,\psi(0)e^{-iHt/\hbar}\right\rangle|^2 = \langle\omega\,|\,\psi(0)\rangle = \mathbb{P}(\omega,0). \tag{3.4.17}$$

**Stationary State Solutions**

For a time-independent Hamiltonian the **stationary state solutions** are the energy eigenstates of $\hat{H}$, they are therefore the family of solutions:

$$|\psi(t)\rangle = |\psi(0)\rangle\,e^{-iHt/\hbar} = \sum_E a_E(0)e^{-iEt/\hbar}\,|E\rangle \tag{3.4.18}$$

which have time-independent probability densities. Since $\hat{H}$ is Hermitian, $\hat{U}(t)$ is

unitary and the propagator/time evolution of the state $|\psi\rangle$ is represented as a rotation in the Hilbert space.

Once we have found the propagator, we can then write:

$$\psi(x,t) = \langle x \,|\, \psi(t)\rangle = \left\langle x \,\middle|\, \hat{U}(t) \,\middle|\, \psi(0)\right\rangle \tag{3.4.19}$$

$$= \int_{-\infty}^{\infty} \left\langle x \,\middle|\, \hat{U}(t) \,\middle|\, x'\right\rangle \langle x' \,|\, \psi(0)\rangle \, dx' \tag{3.4.20}$$

$$= \int_{-\infty}^{\infty} U(t,x,x')\psi(x',0)dx' \tag{3.4.21}$$

so the matrix elements of $\hat{U}(t)$ act as a Green's function for the Schrödinger equation. Expanding $\hat{U}(t,x,x')$ we find:

$$\left\langle x \,\middle|\, \hat{U}(t) \,\middle|\, x'\right\rangle = \langle x| \sum_E \langle E \,|\, x'\rangle \, |E\rangle \, e^{-iEt/\hbar}| = \sum_E \langle x \,|\, E\rangle \, \langle E \,|\, x'\rangle \, e^{-iEt/\hbar} \tag{3.4.22}$$

so that

$$\psi(x,t) = \sum_E \left( \psi_E^*(x) e^{-iEt/\hbar} \int_{-\infty}^{\infty} \psi_E(x')\psi(x',0)dx' \right) \tag{3.4.23}$$

is the solution, where $\psi_E(x)$ is the wave function associated with the energy eigenstate $|E\rangle$.

Alternatively, if we had chosen a different initial time $t'$ then:

$$U(x,t,x',t') = \left\langle x \,\middle|\, \hat{U}(t-t') \,\middle|\, x'\right\rangle \tag{3.4.24}$$

### Choosing a Basis

The Schrödinger equation can of course be solved in any basis, but usually depending on the form of the Hamiltonian one will be more efficient than the other. Consider the one dimensional Hamiltonian:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{3.4.25}$$

If $V(\hat{x})$ is simple, it is simpler to use the $\hat{p}$ basis so that the matrix representation of the momentum operator is diagonal. For example, the particle in a constant force field $F$ has Hamiltonian:

$$\hat{H} = \frac{\hat{p}^2}{2m} - f\hat{x} \tag{3.4.26}$$

Clearly the momentum basis is much more advantageous, in it we get the equation:

$$\Big(\frac{p^2}{2m} - i\hbar f \frac{d}{dp}\Big)\psi(p) = E\psi(p) \tag{3.4.27}$$

whereas in the position basis it would have been a second order ODE.

## 3.5 The time-evolution operator

In this section we want to investigate time evolution in Quantum mechanics in general terms using an equivalent approach to Schrödinger's time-dependent equation. Suppose a system starts out in a state $|\alpha, t_0\rangle$ at time $t_0$, we ask ourselves what will the system's state $|\alpha, t_0; t\rangle$ be at time $t > t_0$? This question is answered by the **time-evolution operator** $\hat{U}(t, t_0)$

$$|\alpha, t_0; t\rangle = \hat{U}(t, t_0)\,|\alpha, t_0\rangle \tag{3.5.1}$$

Since probability must be conserved, we require that the time-evolution operator be unitary

$$\langle \alpha, t_0; t\,|\,\alpha, t_0; t\rangle = \langle \alpha, t_0\,|\,\alpha, t_0\rangle \implies U^\dagger(t, t_0)\hat{U}(t, t_0) = \mathbb{1} \tag{3.5.2}$$

It is clear that $U^\dagger(t, t_0)$ may be regarded as the time-devolution operator, it is the inverse of the time-evolution operator and takes a state back in time rather than forwards. We should also expect that time-evolution be composable, so that evolving a state from $t_0$ to $t_1$, then from $t_1$ to $t_2$ is equivalent to evolving it from $t_0$ to $t_2$ directly. This leads to

$$\hat{U}(t_2, t_1)\hat{U}(t_1, t_0) = \hat{U}(t_2, t_0),\ t_2 > t_1 > t_0 \tag{3.5.3}$$

Since time is a continuous variable we can also consider infinitesimal time-evolutions, where to first order in $dt$ we may write

$$\hat{U}(t_0 + dt, t_0) = \mathbb{1} - i\hat{\Omega}dt \tag{3.5.4}$$

Note that the unitarity of $U$ requires that $\Omega$ be hermitian:

$$(\mathbb{1} + i\hat{\Omega}^\dagger dt)(\mathbb{1} - i\hat{\Omega}dt) = \mathbb{1} + i(\hat{\Omega}^\dagger - \hat{\Omega})dt = \mathbb{1} \implies \hat{\Omega}^\dagger = \hat{\Omega} \tag{3.5.5}$$

The additivity property of $U$ is also satisfied

$$\hat{U}(t_0 + dt_1 + dt_2, t_0 + dt_1)\hat{U}(t_0 + dt_1, t_0) = (\mathbb{1} - i\hat{\Omega}dt_2)(\mathbb{1} - i\hat{\Omega}dt_1) \tag{3.5.6}$$

$$= \mathbb{1} - i\hat{\Omega}(dt_1 + dt_2) \tag{3.5.7}$$

$$= \hat{U}(t_0 + dt_1 + dt_2, t_0) \tag{3.5.8}$$

Recall from classical mechanics that the Hamiltonian is the generator of time-evolution. By the correspondence principle we should expect the same to occur in quantum mechanics. Therefore we may set $\hbar\hat{\Omega} = \hat{H}$, implying that

$$\hat{U}(t_0 + dt, t_0) = \mathbb{1} - \frac{i\hat{H}}{\hbar}dt \tag{3.5.9}$$

We can derive a fundamental equation for the time-evolution operator using this definition

$$\hat{U}(t+dt,t_0) = \hat{U}(t+dt,t)\hat{U}(t,t_0) = \left(\mathbb{1} - \frac{i\hat{H}}{\hbar}dt\right)\hat{U}(t,t_0) \tag{3.5.10}$$

$$\implies \hat{U}(t+dt,t_0) - \hat{U}(t,t_0) = -\frac{i\hat{H}}{\hbar}dt\hat{U}(t,t_0) \tag{3.5.11}$$

$$\implies \boxed{i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = \hat{H}\hat{U}(t,t_0)} \tag{3.5.12}$$

This equation is the **Schrödinger equation for the time-evolution operator**, and can be shown to be equivalent to the original Schroedinger equation. Indeed substituting

$$\hat{U}(t,t_0) = |\alpha,t_0;t\rangle\langle\alpha,t_0| \tag{3.5.13}$$

into (3.5.12) we find that

$$i\hbar\frac{\partial}{\partial t}(|\alpha,t_0;t\rangle\langle\alpha,t_0|) = H\langle\alpha,t_0;t|\,|\alpha,t_0\rangle \tag{3.5.14}$$

$$\implies i\hbar\frac{\partial}{\partial t}|\alpha,t_0;t\rangle = H|\alpha,t_0;t\rangle \tag{3.5.15}$$

as desired. We have thus derived the time dependent Schrödinger equation by simply assuming that the Hamiltonian operator is the generator of time-translations.

### Time-independent hamiltonian

If the Hamiltonian is time-independent, then a simple expression can be found for $\hat{U}(t,t_0)$. Indeed we can decompose the time evolution of a state from $t_0$ to $t$ into $N \to \infty$ infinitesimal steps of width $dt$ so that $t = t_0 + Ndt$, as shown below

$$t_0 \xrightarrow{\hat{U}(t_0+dt,t_0)} t_0+dt \xrightarrow{\hat{U}(t_0+2dt,t_0+dt)} t_0+2dt \longrightarrow \ldots \xrightarrow{\hat{U}(t,t-dt)} t$$

It follows that

$$\hat{U}(t,t_0) = \lim_{N\to\infty}\prod_{m=1}^{N}\hat{U}(t_0+mdt,t_0+(m-1)dt) \tag{3.5.16}$$

$$= \lim_{N\to\infty}\left(\mathbb{1} - \frac{i\hat{H}}{h}dt\right)^N \tag{3.5.17}$$

and since $dt = \frac{(t-t_0)}{N}$ we find that

$$\boxed{\hat{U}(t,t_0) = \exp\left(-\frac{i\hat{H}}{\hbar}(t-t_0)\right)} \tag{3.5.18}$$

### Commuting time-dependent hamiltonian

If instead the Hamiltonian is time-dependent, but the Hamiltonians at different times commute, then

$$\hat{U}(t,t_0) = \lim_{N\to\infty} \prod_{m=1}^{N} \hat{U}(t_0+mdt, t_0+(m-1)dt) \tag{3.5.19}$$

$$= \lim_{N\to\infty} \prod_{m=1}^{N} \exp\left(-\frac{i\hat{H}(t_0+mdt)}{\hbar}dt\right) \tag{3.5.20}$$

$$= \lim_{N\to\infty} \exp\left(-\frac{i}{\hbar}\sum_{m=1}^{N} \hat{H}(t_0+mdt)dt\right) \tag{3.5.21}$$

where in the last step we had to assume that the Hamiltonians at different times commute. Taking the $N\to\infty$ limit we get that

$$\hat{U}(t,t_0) = \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} \hat{H}(t')\,dt'\right) \tag{3.5.22}$$

### General time-independent hamiltonian

Let's integrate (3.5.12):

$$i\hbar\int_{t_0}^{t} dt' \frac{\partial}{\partial t'}\hat{U}(t',t_0) = \int_{t_0}^{t} dt' \hat{H}(t')\hat{U}(t',t_0) \tag{3.5.23}$$

$$\implies \hat{U}(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}(t')\hat{U}(t',t_0) \tag{3.5.24}$$

Substituting (3.5.24) back into (3.5.23) yields

$$\hat{U}(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}(t') + \left(\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H(t')H(t'')U(t'',t_0) \tag{3.5.25}$$

Reiterating this process ad infinitum we arrive at a power-series expansion of the type

$$\begin{aligned}\hat{U}(t,t_0) =& \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}(t') + \left(\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H(t')H(t'') \\ &+ \ldots + \left(\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\ldots H(t_n) + \ldots\end{aligned} \tag{3.5.26}$$

These integrals are all time-ordered, with the each integral running over $t$ smaller than whatever $t'$ contribution of the next outer integral is being evaluated. If we define a **time-ordering operator** $\mathcal{T}$ so that

$$\mathcal{T}[H(t_1)H(t_2)] = \begin{cases} H(t_1)H(t_2), & \text{for } t_1 \geq t_2 \\ H(t_2)H(t_1), & \text{for } t_1 < t_2 \end{cases} \tag{3.5.27}$$

Hence for example

$$\int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \mathcal{T}[H(t')H(t'')] = \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H(t')H(t'') + \int_{t_0}^{t} dt'' \int_{t_0}^{t''} dt' H(t'')H(t') \tag{3.5.28}$$

$$= 2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H(t')H(t'') \tag{3.5.29}$$

or more generally

$$\int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_n \mathcal{T}[H(t_1)...H(t_n)] = \frac{1}{n!}\left(\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_n \mathcal{T}[H(t_1)...H(t_n)] \tag{3.5.30}$$

Using this result we may write (3.5.26) then

$$\begin{aligned}\hat{U}(t,t_0) =& \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}(t') + \frac{1}{2}\left(\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t} dt'' \mathcal{T}[H(t')H(t'')] \\ &+ ... + \frac{1}{n!}\left(\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 ... \int_{t_0}^{t} dt_n \mathcal{T}[H(t_1)H(t_2)...H(t_n)] + ...\end{aligned} \tag{3.5.31}$$

or in even more compact notation

$$\hat{U}(t,t_0) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt' \hat{H}(t')\right)\right] \tag{3.5.32}$$

This is known as the **Dyson series**. In practice only the first few terms of (3.5.31) are calculated in the perturbative limit.

## 3.6 The pictures of quantum mechanics

In the previous section we have looked at states in the Hilbert space as varying with time, while the operators acting on them remained constant and time-independent (except when they had explicit time-dependence built into them). Geometrically this corresponds to viewing the states as vectors rotating under the action of $U(t,t_0)$. However, one may instead wish to view the states as fixed, with the operators acting on them as changing with time. It is well known that one can perform changes of basis on matrices, so if one wishes to maintain the basis fixed then it suffices to perform a time-dependent change of basis on the operators. Finally, in some instances it may be more favourable to let the state evolve under one portion of the Hamiltonian, and the operators evolve on another. These three correspond to different equivalent pictures in quantum mechanics, known as the Schrödinger, Heisenberg and interaction picture.

### Schrödinger's picture

This is the picture we are most familiar with. If we consider a state $|\psi(t_0)\rangle$ then it will evolve under the propagator as

$$|\psi(t)\rangle = \hat{U}(t,t_0)\,|\psi(t_0)\rangle \tag{3.6.1}$$

In the Schrödinger picture the operators $\hat{A}_S$ corresponding to physical observables are time-independent, so that the expectation values are

$$\langle A\rangle(t) \equiv \left\langle \psi(t)\,\middle|\,\hat{A}_S\,\middle|\,\psi(t)\right\rangle = \left\langle \psi(t_0)\,\middle|\,\hat{U}^\dagger(t,t_0)\hat{A}\_ShatU(t,t_0)\,\middle|\,\psi(t_0)\right\rangle \tag{3.6.2}$$

Assuming $|\psi(t_0)\rangle$ is an eigenstate $|E\rangle$ of the Hamiltonian then

$$\langle A\rangle(t) = \left\langle \psi(t_0)\,\middle|\,e^{iEt/\hbar}\hat{A}_S e^{-iEt/\hbar}\,\middle|\,\psi(t_0)\right\rangle = \langle A\rangle(t_0) \tag{3.6.3}$$

implying that the expectation value of any time-independent operator will be stationary. It is for this reason that energy eigenstates are often referred to as statioanry states, as mentioned previously.

### Heisenberg's picture

It is evident from (3.6.2) that if we wish to keep the states time-independent, then one should evolve the operators as

$$\hat{A}_H(t) = \hat{U}^\dagger(t,t_0)\hat{A}_S\hat{U}(t,t_0) \tag{3.6.4}$$

This change in perspective will yield a different set of equations of motion. It no longer makes sense to think of Schrödinger's equation as the states are time-independent, instead we should try looking at the time-evolution of the operators:

$$\frac{d\hat{A}_H(t)}{dt} = \frac{d}{dt}(\hat{U}^\dagger(t,t_0)\hat{A}_S\hat{U}(t,t_0)) \tag{3.6.5}$$

$$= \frac{\partial\hat{U}^\dagger(t,t_0)}{\partial t}\hat{A}_S\hat{U}(t,t_0) + \hat{U}^\dagger(t,t_0)\hat{A}_S\frac{\partial\hat{U}(t,t_0)}{\partial t} \tag{3.6.6}$$

$$= \frac{i}{\hbar}\Big(\hat{U}^\dagger(t,t_0)\hat{H}\hat{A}_S\hat{U}(t,t_0) - \hat{U}^\dagger(t,t_0)\hat{A}_S\hat{H}\hat{U}(t,t_0)\Big) \tag{3.6.7}$$

$$\implies \frac{d\hat{A}_H(t)}{dt} = \frac{i}{\hbar}[\hat{H}_H,\hat{A}_H] \tag{3.6.8}$$

where we defined $\hat{H}_H$ to be the Hamiltonian in the Heisenberg picture. In most practical applications where $\hat{H}$ is time-independent note that $[\hat{H},\hat{U}(t,t_0)] = 0$ so that $\hat{H}_H = \hat{H}$. Equation (3.6.8) is known as **Heisenberg's equation of motion**. Its equivalence to Schrödinger's equation is manifest from the fact that we used (3.5.12) in going from the second to third line.

### Interaction picture

In some instances it may be that the time-evolution is very simple for one part $H^0$ of the hamiltonian $H$ (for example if it is diagonalisable in a simple basis) but is complicated in another part $V$. Let's consider as an example the following Schrödinger picture hamiltonian

$$\hat{H}_S = \hat{H}_S^0 + \hat{V}_S(t) \tag{3.6.9}$$

where $H_S^0$ is time-independent and exactly solvable, while $V_S(t)$ is a more complicated, time-dependent term. Let us define the interaction picture state to be related to the Schrödinger picture states via

$$|\psi_I(t)\rangle = e^{i\hat{H}_S^0 t/\hbar} |\psi_S(t)\rangle \tag{3.6.10}$$

and similarly for the operators

$$\hat{A}_I(t) = e^{i\hat{H}_S^0 t/\hbar} \hat{A}_S(t) e^{-i\hat{H}_S^0 t/\hbar} \tag{3.6.11}$$

Most importantly, this means that

$$\hat{H}_I = \hat{H}_S^0 + e^{-i\hat{H}_S^0 t/\hbar} \hat{V}_S(t) e^{i\hat{H}_S^0 t/\hbar} \tag{3.6.12}$$

since $\hat{H}_S^0 = \hat{H}_I^0$. Lets verify what the resulting equations of motion are. For the states we find that

$$\frac{\partial}{\partial t}|\psi_I(t)\rangle = \frac{\partial}{\partial t}(e^{i\hat{H}_S^0 t/\hbar}|\psi_S(t)\rangle) = \frac{iH_S^0}{\hbar} e^{iH_S^0 t/\hbar}|\psi_S(t)\rangle + e^{iH_S^0 t/\hbar}\frac{\partial}{\partial t}|\psi_S(t)\rangle \tag{3.6.13}$$

$$= \frac{iH_S^0}{\hbar} e^{iH_S^0 t/\hbar}|\psi_S(t)\rangle - \frac{i}{\hbar} e^{iH_S^0 t/\hbar}(H_S^0 + V_S(t)) e^{-iH_S^0 t/\hbar}|\psi_I(t)\rangle \tag{3.6.14}$$

$$= \frac{1}{i\hbar}\hat{V}_I(t)|\psi_I(t)\rangle \implies \boxed{i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = \hat{V}_I(t)|\psi_I(t)\rangle} \tag{3.6.15}$$

For the operator, note that since $H_0^S$ is time-independent, $A_I(t) = A_H(t)$ and therefore

$$\boxed{\frac{d}{dt}A_I(t) = \frac{i}{\hbar}[\hat{H}_S^0, \hat{A}_I(t)]} \tag{3.6.16}$$

Equations (3.6.15) and (3.6.16) together are known as the **interaction picture equations of motion**. It is evident from the table below that the interaction picture is intermediate between the Schrödinger and Heisenberg pictures Note that we may define a time-evolution

| | Schrödinger picture | Interaction picture | Heisenberg picture |
|---|---|---|---|
| State | determined by $H$ | determined by $\hat{V}_I(t)$ | unchanged |
| Observable | unchanged | determined by $H_0$ | determined by $H$ |

operator $\hat{U}_I(t, t_0)$ in the interaction picture by

$$\hat{U}_I(t, t_0)|\psi_I(t_0)\rangle = |\psi_I(t)\rangle\langle\psi_I(t_0)| \tag{3.6.17}$$

Indeed dotting (3.6.15) to the left with $\langle\psi_I(t_0)|$ we see that

$$i\hbar\frac{\partial\hat{U}_I(t,t_0)}{\partial t} = \hat{V}_I(t)\hat{U}_I(t,t_0) \tag{3.6.18}$$

We already solved this problem in the previous section using the Dyson series, we quote the result below:

$$U_I(t,t_0) = \mathcal{T}\Big[\exp\Big(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_I(t')\Big)\Big] \tag{3.6.19}$$

# The Path-Integral formulation

4

This chapter contains a lot of advanced material and requires substantial information not yet explained so it should be skipped at a first reading. However, logically it should fit in this place in the textbook, hence why it is included so early in this volume.

## 4.1 Gaussian integrals

Before starting our discussion of path integrals we should prove a few results on Gaussian integrals that will be very useful. We start by evaluating

$$I(a) = \int_{-\infty}^{\infty} e^{-ax^2}\, dx, \quad \text{where } a > 0 \tag{4.1.1}$$

The trick for evaluating this Gaussian integral is noting that $x$ is a dummy variable so considering $I^2$ then one integral can be integrated over $x$ while the other over $y$

$$I^2 = \Big(\int_{-\infty}^{\infty} e^{-ax^2}\, dx\Big)^2 = \Big(\int_{-\infty}^{\infty} e^{-ax^2}\, dx\Big)\Big(\int_{-\infty}^{\infty} e^{-ay^2}\, dy\Big) \tag{4.1.2}$$

Consequently, changing variables from $(x, y)$ to $(r, \theta)$ then

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-a(x^2+y^2)}\, dxdy = \int_0^{\infty} e^{-ar^2} 2\pi r dr = \frac{2\pi}{a}\int_0^{\infty} \rho e^{-\rho^2}\, d\rho = \frac{\pi}{a} \tag{4.1.3}$$

so that

$$I(a) = \sqrt{\frac{\pi}{a}} \tag{4.1.4}$$

It follows immediately that

$$\int_{-\infty}^{\infty} e^{-a(x-c)^2}\, dx = \sqrt{\frac{\pi}{a}} \tag{4.1.5}$$

Geometrically, we can understand this result by noting that the integrand is a gaussian bell curve, so changing coordinates $x \mapsto x + c$ simply shifts the center of the bell curve, but not the area below it. We can generalise this result by computing

$$I(a, b) = \int_{-\infty}^{\infty} e^{-ax^2+bx+c}\, dx, \quad \text{where } a > 0 \tag{4.1.6}$$

We do so by completing the square of the exponent

$$ax^2 - bx - c = a\left(x - \frac{b}{2a}\right)^2 - \left(c + \frac{b^2}{2a}\right) \tag{4.1.7}$$

implying that

$$I(a,b) = e^{-c-b^2/2a}\int_{-\infty}^{\infty} e^{a(x-b/2a)^2}\,dx = \sqrt{\pi a}e^{b^2/2a+c} \tag{4.1.8}$$

These integrals are quite useful when considering 1D path integrals. In N-dimensions there is a much more general class of Gaussian integrals. The equivalent of $I(a)$ now becomes

$$I(\mathsf{A}) = \int_{\mathbb{R}^N} e^{-\mathbf{x}^T\mathsf{A}\mathbf{x}}\,d^N\mathbf{x} \tag{4.1.9}$$

where $\mathsf{A}$ is a symmetric positive-definite $N \times N$ matrix (this condition is equivalent to saying $a > 0$ in the 1D case). The trick this time is to diagonalise the matrix $\mathsf{A}$ (which can be diagonalised due to the spectral theorem) which will recast the matrix into the 1D form we have already solved. More specifically, let

$$\mathsf{A} = \mathsf{O}^T\mathsf{D}\mathsf{O} \tag{4.1.10}$$

where $\mathsf{O}$ is an orthogonal matrix and $\mathsf{D} = \text{diag}(a_1, a_2, ..., a_N)$. Performing a change of variables $\mathbf{x} \mapsto \mathbf{y} = \mathsf{O}\mathbf{x}$ with unit Jacobian (since $|\det \mathsf{O}| = 1$), we find that

$$I(\mathsf{A}) = \int_{\mathbb{R}^N} e^{-\mathbf{y}^T\mathsf{D}\mathbf{y}}\,d^N\mathbf{y} = \prod_{n=1}^{N}\left[\int_{\mathbb{R}} e^{-a_n y_n^2}dy_n\right] = \sqrt{\frac{\pi}{a_1...a_N}} = \sqrt{\frac{\pi^N}{\det \mathsf{A}}} \tag{4.1.11}$$

Again, we can generalise these results by considering

$$I(\mathsf{A},\mathbf{J}) = \int_{\mathbb{R}^N} e^{-\mathbf{x}^T\mathsf{A}\mathbf{x}+\mathbf{J}^T\mathbf{x}}\,d^N\mathbf{x} \tag{4.1.12}$$

This time we consider the change of coordinates $\mathbf{x} \mapsto \mathbf{y} = \mathbf{x} - \frac{1}{2}\mathsf{A}^{-1}\mathbf{J}$. Then

$$-\mathbf{x}^T\mathsf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} = -(\mathbf{y}^T + \frac{1}{2}\mathbf{J}^T\mathsf{A}^{-1})\mathsf{A}(\mathbf{y} + \frac{1}{2}\mathsf{A}^{-1}\mathbf{J}) + \mathbf{J}^T(\mathbf{y} + \frac{1}{2}\mathsf{A}^{-1}\mathbf{J}) \tag{4.1.13}$$

$$= -\mathbf{y}^T\mathsf{A}\mathbf{y} - \frac{1}{2}\mathbf{y}^T\mathbf{J} - \frac{1}{2}\mathbf{J}^T\mathbf{y} - \frac{1}{4}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J} + \mathbf{J}^T\mathbf{y} + \frac{1}{2}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J} \tag{4.1.14}$$

$$= -\mathbf{y}^T\mathsf{A}\mathbf{y} + \frac{1}{4}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J} \tag{4.1.15}$$

so

$$I(\mathsf{A},\mathbf{J}) = e^{\frac{1}{4}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J}}\int_{\mathbb{R}^N} e^{-\mathbf{y}^T\mathsf{A}\mathbf{y}}\,d^N\mathbf{y} \tag{4.1.16}$$

implying that

$$I(\mathsf{A},\mathbf{J}) = e^{\frac{1}{4}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J}}\sqrt{\frac{\pi^N}{\det \mathsf{A}}} \tag{4.1.17}$$

## 4.2 Path integral formulation

Our formulation of quantum mechanics thus far has been centered around the Hamiltonian. However, recall that in classical mechanics the Hamiltonian and Lagrangian formulations of the equations of motion are equivalent. Can the same be said of quantum mechanics, does the latter also have a Lagrangian formulation?

The answer turns out to be yes, there is a Lagrangian-based approach to quantum mechanics known as the path-integral formulation. Recall that in the differential formulation of QM, one obtained the propagator by solving the differential equation

$$i\hbar\frac{d}{dt}\hat{U}(t) = \hat{H}\hat{U}(t) \tag{4.2.1}$$

Similarly, the path integral-formulation obtains the propagator by the use of an integral equation

$$U(\mathbf{x}_N, \mathbf{x}_0, t) = \int \mathcal{D}[\mathbf{x}]\; e^{iS[\mathbf{x}]/\hbar} \tag{4.2.2}$$

To understand where (4.2.2) comes from, let's start by considering a time-independent Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}) \tag{4.2.3}$$

Then we see that the propagator reads

$$U(\mathbf{x}', \mathbf{x}, t) = \left\langle ' \right| e^{-iHt/\hbar} \left| \mathbf{x} \right\rangle \left\langle \mathbf{x}' \right| \exp\Big( -\frac{it}{\hbar}\Big(\frac{\mathbf{p}^2}{2m} + V(\mathbf{x})\Big)\Big) \left| \mathbf{x} \right\rangle \tag{4.2.4}$$

Now consider subdividing $t$ into $N$ intervals of width $\varepsilon = \frac{t}{N}$. Then we see that

$$\exp\Big( -\frac{it}{\hbar}\Big(\frac{\mathbf{p}^2}{2m} + V(\mathbf{x})\Big)\Big) = \Big[\exp\Big( -\frac{i\epsilon}{\hbar}\Big(\frac{\mathbf{p}^2}{2m} + V(\mathbf{x})\Big)\Big)\Big]^N \tag{4.2.5}$$

We can use the Bakker-Hausdorff identity

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}+1/2[\hat{A},\hat{B}]+\ldots} \tag{4.2.6}$$

and working to first order in $\epsilon$ (which we will take to approach 0) then

$$\exp\Big( -\frac{i\epsilon}{\hbar}\Big(\frac{\mathbf{p}^2}{2m} + V(\mathbf{x})\Big) \approx \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar}V(\mathbf{x})\Big) \tag{4.2.7}$$

Substituting this into (4.2.4) then

$$U(\mathbf{x}', \mathbf{x}, t) = \left\langle \mathbf{x}' \right| \prod_{i=1}^{N} \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar}V(\mathbf{x})\Big) \left| \mathbf{x} \right\rangle \tag{4.2.8}$$

### 4.2.1 Configuration space integral

We can simplify this expression (surprisingly) by introducing an identity resolution

$$\mathbb{1} = \int d^3\mathbf{x}\, |\mathbf{x}\rangle\, \langle \mathbf{x}| \tag{4.2.9}$$

between each of the $N$ terms in the product, yielding

$$\left\langle \mathbf{x}' \middle| \int \prod_{i=1}^{N} d^3\mathbf{x}_i \Big[ \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar} V(\mathbf{x})\Big) \middle| \mathbf{x}_i \right\rangle \langle \mathbf{x}_i \Big] |\mathbf{x}\rangle \tag{4.2.10}$$

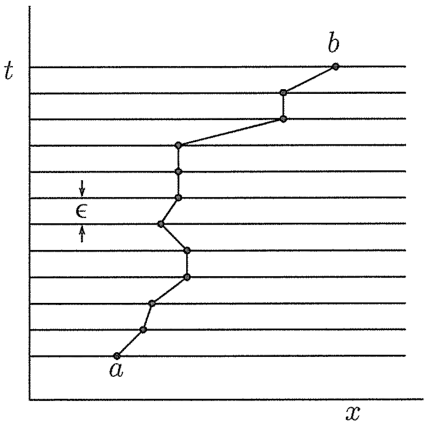
For example if $N = 3$ then we would find that

$$U(\mathbf{x}_3, \mathbf{x}_0, t) = \left\langle \mathbf{x}_3 \middle| \prod_{i=1}^{2} \int d^3\mathbf{x}_i \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar} V(\mathbf{x})\Big) \middle| \mathbf{x}_2 \right\rangle \tag{4.2.11}$$

$$\times \left\langle \mathbf{x}_2 \middle| \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar} V(\mathbf{x})\Big) \middle| \mathbf{x}_1 \right\rangle \tag{4.2.12}$$

$$\times \left\langle \mathbf{x}_1 \middle| \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar} V(\mathbf{x})\Big) \middle| \mathbf{x}_0 \right\rangle \tag{4.2.13}$$

There is a nice physical interpretation to this equation. Each of the brakets (or better, their modulus squared) evaluates the probability that the Hamiltonian takes a state from $|\mathbf{x}_{n-1}\rangle$ to $|\mathbf{x}_n\rangle$. Integrating over $\prod_{i=1}^{N} d^3\mathbf{x}_i$ is equivalent to integrating over all paths in phase space with fixed endpoints at $|\mathbf{x}_0\rangle$ and $|\mathbf{x}_N\rangle$. Consequently the sums of the probabilities along all paths taking $|\mathbf{x}_0\rangle$ at time $t = 0$ to $|\mathbf{x}_N\rangle$ at time $t$ gives the total probability of going from $|\mathbf{x}_0\rangle$ to $|\mathbf{x}_N\rangle$. In other words, we may view a quantum particle as taking all possible paths in configuration space from an initial state to a final state, each being weighed by the probability amplitude of this occurring. But what is this weight factor and does it have a physical significance?




Since $|\mathbf{x}\rangle$ are eigenstates of $V(\mathbf{x})$ we can write that the $n$th term will be

$$\left\langle \mathbf{x}_n \middle| \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \exp\Big( -\frac{i\epsilon}{\hbar} V(\mathbf{x})\Big) \middle| \mathbf{x}_{n-1} \right\rangle = \exp\Big(-\frac{i\epsilon}{\hbar} V(\mathbf{x}_{n-1})\Big) \left\langle \mathbf{x}_n \middle| \exp\Big( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\Big) \middle| \mathbf{x}_{n-1} \right\rangle \tag{4.2.14}$$

We can compute this matrix element by introducing yet another resolution of the identity

$$\mathbb{1} = \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\, |\mathbf{p}\rangle\, \langle \mathbf{p}| \tag{4.2.15}$$

where $\langle \mathbf{x} \,|\, \mathbf{p} \rangle = e^{i\mathbf{p}\cdot\mathbf{x}\hbar}$ We find that

$$\left\langle \mathbf{x}_n \left| \exp\left( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right) \right| \mathbf{x}_{n-1} \right\rangle = \left\langle \mathbf{x}_n \left| \exp\left( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right) \right| \mathbf{p} \right\rangle \mathbf{p}|\mathbf{x}_{n-1} \tag{4.2.16}$$

$$= \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \exp\left( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m} + \frac{i}{\hbar}\mathbf{p}\cdot(\mathbf{x}_n - \mathbf{x}_{n-1})\right) \tag{4.2.17}$$

Using the 3D Gaussian integral in (4.1.17) with $\mathbf{J} = \frac{i}{\hbar}(\mathbf{x}_n - \mathbf{x}_{n-1})$ and $\mathsf{A} = \frac{i\epsilon}{2m\hbar}\mathbb{1}$ we get that

$$\left\langle \mathbf{x}_n \left| \exp\left( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right) \right| \mathbf{x}_{n-1} \right\rangle = \frac{1}{(2\pi\hbar)^3}\left(\frac{\pi}{i\epsilon/2m\hbar}\right)^{3/2} \exp\left( -\frac{2m\hbar}{i\epsilon}\frac{1}{4}\frac{(\mathbf{x}_n - \mathbf{x}_{n-1})^2}{\hbar^2}\right) \tag{4.2.18}$$

and therefore

$$\left\langle \mathbf{x}_n \left| \exp\left( -\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right) \right| \mathbf{x}_{n-1} \right\rangle = \left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{3/2} \exp\left(\frac{im(\mathbf{x}_n - \mathbf{x}_{n-1})^2}{2\hbar\epsilon}\right) \tag{4.2.19}$$

To conclude we have that

$$U(\mathbf{x}_N, \mathbf{x}_0, t) = \left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{3N/2} \int \prod_{n=0}^{N-1} d^3\mathbf{x}_i \ \exp\left( \sum_{n=1}^{N} \frac{im(\mathbf{x}_n - \mathbf{x}_{n-1})^2}{2\hbar\epsilon} - \frac{i\varepsilon}{\hbar}V(\mathbf{x}_{n-1})\right) \tag{4.2.20}$$

To connect this expression to the classical analogue we define

$$\int \mathcal{D}[\mathbf{x}] \equiv \left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{3N/2} \int \prod_{n=0}^{N-1} d^3\mathbf{x}_i \tag{4.2.21}$$

to obtain

$$U(\mathbf{x}_N, \mathbf{x}_0, t) = \lim_{N\to\infty} \int \mathcal{D}[\mathbf{x}] \ \exp\left[\frac{i\epsilon}{\hbar}\left( \sum_{n=1}^{N} \frac{m(\mathbf{x}_n - \mathbf{x}_{n-1})^2}{2\epsilon^2} - V(\mathbf{x}_{n-1})\right)\right] \tag{4.2.22}$$

We recognise the classical action in a discretised form as

$$S[\mathbf{x}] = \int_0^t \left(\frac{m\dot{\mathbf{x}}^2}{2} - V(\mathbf{x})\right) dt = \lim_{N\to\infty} \sum_{n=1}^{N} \left(\frac{m(\mathbf{x}_n - \mathbf{x}_{n-1})^2}{2\epsilon^2} - V(\mathbf{x}_{n-1})\right)\epsilon \tag{4.2.23}$$

which allows us to write

$$U(\mathbf{x}_N, \mathbf{x}_0, t) = \int \mathcal{D}[\mathbf{x}] \ e^{iS[\mathbf{x}]/\hbar} \tag{4.2.24}$$

as desired. The weight for each path is therefore the total action along it.

**Path integral formulation**

The general procedure to evaluate $U(x,t,;x',t')$ is to:
(i) fix the endpoints $(x,t)$ and $(x',t')$ and consider all paths $x(t)$ connecting them.
(ii) find the action $S[x(t)]$ for each path.
(iii) Up to a normalization factor:

$$U(\mathbf{x}_N,\mathbf{x}_0,t,t_0)=\int \mathcal{D}[\mathbf{x}]\; e^{iS[\mathbf{x}]/\hbar} \tag{4.2.25}$$

### Configuration space integral

Instead of inserting an identity operator after each product in, let's instead resolve the identity between every exponential. It will prove useful to resolve the identity in two different ways using the real space and momentum space bases

$$\mathbb{1}=\int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3}\;|\mathbf{p}\rangle\langle\mathbf{p}|=\int d^3\mathbf{x}\;|\mathbf{x}\rangle\langle\mathbf{x}| \tag{4.2.26}$$

We then find that

$$\left\langle \mathbf{x}_N\left|\int\prod_{j=1}^{N}\prod_{i=1}^{N-1}\frac{d^3\mathbf{p}_i}{(2\pi\hbar)^3}\,d^3\mathbf{x}_i\left[\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right)\right|\mathbf{p}_i\right\rangle\left\langle\mathbf{p}_i\left|\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x})\right)\right|\mathbf{x}_{i-1}\right\rangle\langle\mathbf{x}_{i-1}\Big]|\mathbf{x}_0\right\rangle \tag{4.2.27}$$

For example if $N=3$ then we would find that

$$U(\mathbf{x}_3,\mathbf{x}_0,t)=\left\langle\mathbf{x}_3\left|\prod_{j=1}^{3}\prod_{i=1}^{2}\int\frac{d^3\mathbf{p}_j}{(2\pi\hbar)^3}\,d^3\mathbf{x}_i\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right)\right|\mathbf{p}_3\right\rangle\left\langle\mathbf{p}_3\left|\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x})\right)\right|\mathbf{x}_2\right\rangle \tag{4.2.28}$$

$$\times\left\langle\mathbf{x}_2\left|\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right)\right|\mathbf{p}_2\right\rangle\left\langle\mathbf{p}_2\left|\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x})\right)\right|\mathbf{x}_1\right\rangle \tag{4.2.29}$$

$$\times\left\langle\mathbf{x}_1\left|\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right)\right|\mathbf{p}_1\right\rangle\left\langle\mathbf{p}_1\left|\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x})\right)\right|\mathbf{x}_0\right\rangle \tag{4.2.30}$$

Using

$$\left\langle\mathbf{x}_n\left|\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}^2}{2m}\right)\right|\mathbf{p}_n\right\rangle=\exp\left(-\frac{i\epsilon}{\hbar}\frac{\mathbf{p}_n^2}{2m}+\frac{i}{\hbar}\mathbf{p}_n\cdot\mathbf{x}\right) \tag{4.2.31}$$

$$\left\langle\mathbf{p}_n\left|\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x})\right)\right|\mathbf{x}_{n-1}\right\rangle=\exp\left(-\frac{i\epsilon}{\hbar}V(\mathbf{x}_{n-1})-\frac{i}{\hbar}\mathbf{p}_n\cdot\mathbf{x}_{n-1}\right) \tag{4.2.32}$$

it follows that

$$U(\mathbf{x}_N,\mathbf{x}_0,t)=\int\prod_{j=1}^{N}\prod_{i=1}^{N-1}\frac{d^3\mathbf{p}_i}{(2\pi\hbar)^3}\,d^3\mathbf{x}_i\exp\left[\frac{i\epsilon}{\hbar}\sum_{j=1}^{N}\left(\frac{\mathbf{p}_n\cdot(\mathbf{x}_n-\mathbf{x}_{n-1})}{\epsilon}-\frac{\mathbf{p}_n^2}{2m}-V(\mathbf{x}_n)\right)\right] \tag{4.2.33}$$

It is easy to show that this will yield the same expression as the configuration space integral (by integrating over the $\mathbf{p}_i$ variables). Nevertheless the propagator in its current form is an integral over phase space rather than configuration space (thus explaining the absence of the weight factor $(m/2\pi i\hbar\epsilon)^{1/2}$), so it is called the **phase space integral**. Indeed by defining

$$\int \mathcal{D}[\mathbf{x}]\mathcal{D}[\mathbf{p}] \longleftrightarrow \lim_{N\to\infty} \int \prod_{j=1}^{N} \prod_{i=1}^{N-1} \frac{d^3\mathbf{p}_i}{(2\pi\hbar)^3}\, d^3\mathbf{x}_i, \qquad \lim_{N\to\infty} \epsilon \sum_{i=1}^{N} \to \int_0^t dt \tag{4.2.34}$$

then we see that

$$U(\mathbf{x}_N, \mathbf{x}_0, t) = \int \mathcal{D}[\mathbf{x}]\mathcal{D}[\mathbf{p}] \exp\left[\frac{i}{\hbar}\int_0^t dt(\mathbf{p}\cdot\dot{\mathbf{x}} - \mathcal{H})\right] \tag{4.2.35}$$

## 4.3 Retrieving the classical path: the stationary phase approximation

In the previous section we saw how one could obtain the time-evolution of a quantum system by simply letting it take all possible paths in phase/configuration space and then taking a sum over each of these paths. However it is difficult to see how one can retrieve the classical path in the classical limit ($\hbar \to 0$) if we are summing over all the possible paths a system can take.

Consider the classical path $x_{\text{cl}}(t)$ (the path which minimizes the action) starting at $(x', t')$ and ending at $(x, t)$. In the classical limit we should expect only this path to contribute and thus that

$$U(x, t; x', t') \sim e^{iS_{\text{cl}}/\hbar} \tag{4.3.1}$$

Let's now consider a small perturbation $\eta_1$ from the classical path which keep the endpoints fixed i.e. $\eta_1(t') = \eta_1(t) = 0$. We denote the deviated path as $x_1 = x_{\text{cl}} + \eta_1$. Suppose we repeat this process thus creating a grey region of paths in the neighborhood of the classical path.

Let us examine the contribution to the propagator due to the classical path and paths in its neighbourhood (grey region). Because the path is approximately stationary in the grey region, the individual contribution due to all paths close to the classical one will be coherent, since they will have the same phase/same action. Consequently, if we consider each small arrow in the phasor diagram for $U$ as the contribution due to some path close to or equal to $x_{cl}$, then they will add constructively.

It can be shown that for every path outside the neighbourhood of $x_{cl}$, we can find an associated path whose two phasors cancel out. Indeed, consider a path $x_i$. We can define a difference function as $\eta_i = x_{cl} - x_i$ in the interval $[t', t]$ with boundary conditions $\eta_i(t') = \eta_i(t) = 0$. Recall from the derivation of the Lagrange equation that:

$$e^{iS[x_i]/\hbar} = e^{iS[x_{cl}]/\hbar} \exp\left[\frac{i}{h}\int_t^{t'} \left(\frac{\partial\mathcal{L}}{\partial x_i(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{x}_i(t)}\right)\eta_i(t)\right] = e^{iS[x_{cl}]/\hbar} e^{i\delta S_1/\hbar} \tag{4.3.2}$$

It is evident that for paths close to $x_{cl}$ the action is unchanged $\delta S = 0$. However, as we

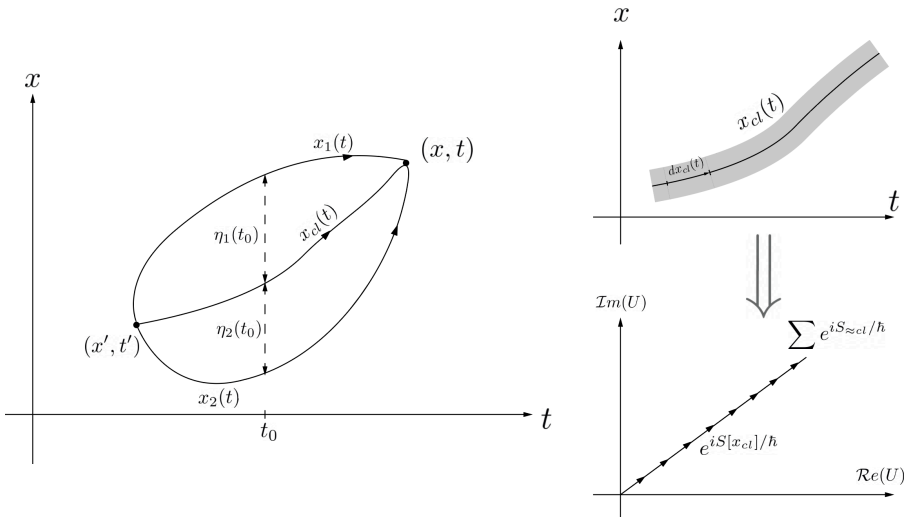



**Figure 4.1**. Paths from $(x', t')$ to $(x, t)$ near the classical path with stationary action, and their contribution to the propagator adding coherently.

get farther away from the classical path, we start to get an additional phase factor $e^{i\delta S[x_i]/\hbar}$ multiplied to the classical contribution $e^{iS[x_{cl}]/\hbar}$. These are very rapidly oscillating contributions for $\delta S \gg \hbar$ which will integrate to zero. Hence, as we get farther from the classical path, destructive interference starts to set in between the different paths and as the phase changes, the small arrows starts to add destructively, effectively cancelling each other out as shown below:

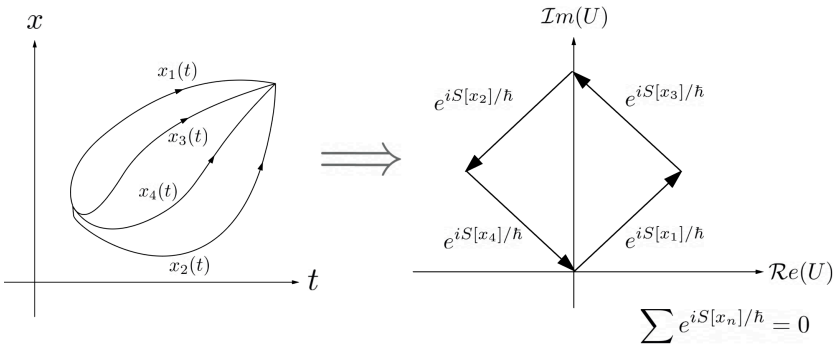



**Figure 4.2**. Phasor diagram showing decoherence in the contributions due to paths outside the classical region

Mathematically, we are trying to prove that if $f(x)$ has an extremum at $x_0$ then

$$\int_{-\infty}^{\infty} dx\, e^{i\alpha f(x)} \tag{4.3.3}$$

is dominated by values close to $x_0$ when $\alpha$ is very big. Indeed $f(x)$ oscillates very rapidly as $(x - x_0) \to \infty$ implying that the integral in the regions far from $x_0$ will sum to zero, as shown below

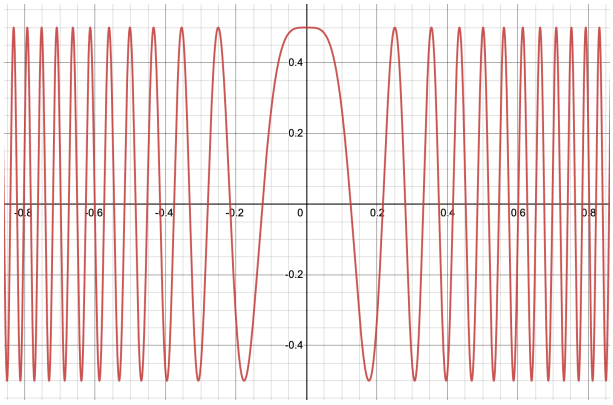



**Figure 4.3**. Plot of $\cos(\alpha x^2)$ for big $\alpha$.

Near $x_0$ on the other hand, we can Taylor expand $f$ as

$$f(x) \approx f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 ... \tag{4.3.4}$$

Substituting this into (4.3.3) we obtain

$$\int_{-\infty}^{\infty} dx\, e^{i\alpha f(x)} \approx e^{i\alpha f(x_0)} \int_{-\infty}^{\infty} d(\delta x)\, e^{i\alpha(\delta x^2 f''(x_0)/2 + ...)} \tag{4.3.5}$$

Since only the small deviations $\delta x$ will contribute to the integral then we may ignore all higher order $\delta x$ terms and end up with

$$\int_{-\infty}^{\infty} dx\, e^{i\alpha f(x)} \approx e^{i\alpha f(x_0)} \int_{-\epsilon}^{\epsilon} d(\delta x)\, e^{i\alpha(\delta x^2 f''(x_0)/2)} \tag{4.3.6}$$

where $[x_0 - \epsilon, x + \epsilon]$ is the neighborhood of $x_0$ for which the integral does not vanish. Next we can extend the limits of integration since high frequency terms will not contribute to the integral, so we end up with a Gaussian integral

$$\int_{-\infty}^{\infty} dx\, e^{i\alpha f(x)} \approx e^{i\alpha f(x_0)} \int_{-\infty}^{\infty} d(\delta x) e^{i\alpha(\delta x^2 f''(x_0)/2} = \sqrt{\frac{2\pi i}{\lambda f''(x_0)}} e^{i\alpha f(x_0)} \tag{4.3.7}$$

Extending this argument to the path integral we obtain

$$U(x, t; x', t') \sim e^{iS_{cl}/\hbar} \text{ in the classical limit } \hbar \to 0 \tag{4.3.8}$$

where $A$ measures the density of paths in the coherence region (grey region in Figure 4.1). One may quantify this coherence range as the set of paths for which $\delta S \ll \hbar$. For a classical path, since $S_{cl} \approx 10^{27}\hbar$, the grey region is very very narrow, allowing for virtually no deviation from the classical path. However, as we enter the realm of quantum phenomena, the action is much much smaller compared to $\hbar$ thus allowing for more deviation from the classical path and thus the counter-intuitive effects quantum mechanics is known for.

## 4.4 Potentials of form $V = a + bx + cx^2 + d\dot{x} + ex\dot{x}$

It turns out that we can use the stationary phase approximation $\hat{U}(t) = A(t)e^{iS_{cl}/\hbar}$ *exactly* for potentials of the form $V = a + bx + cx^2 + d\dot{x} + ex\dot{x}$. It's not hard to see why, indeed if $\mathcal{L}$ quadratic in $x, \dot{x}$ then cutting off the Taylor expansion at second order is no longer an approximation, it is exact. The result in (4.3.7) then becomes exact.

More concretely, let us compute the path integral

$$U(x_1, t_1; x_0) = \int_{x_0}^{x_1} e^{iS[x(t)]/\hbar} \mathcal{D}[x(t)] \tag{4.4.1}$$

We begin by writing all paths we are integrating over as $x(t) = x_{cl}(t) + \eta(t)$ where $\eta(x_0) = \eta(x_1) = 0$. The integration variables turn into: $\mathcal{D}[x(t)] = \mathcal{D}[\eta(t)]$ since $x_{cl}$ is constant. Therefore

$$\int_{x_0}^{x_1} \mathcal{D}[x(t)]\, e^{iS[x(t)]/\hbar} = \int_0^0 \mathcal{D}[\eta(t)]\, \exp\Big(\frac{i}{\hbar} S[x_{cl}(t) + \eta(t)]\Big) \tag{4.4.2}$$

where $\int_0^0$ means that we are integrating over path deviations $\eta$ satisfying the boundary condition $\eta(x_0) = \eta(x_1) = 0$. We may write $S[x_{cl}(t) + \eta(t)]$ using taylor series as:

$$S[x_{cl}(t) + \eta(t)] = \int_0^{t_1} \mathcal{L}(x_{cl} + \eta, \dot{x}_{cl} + \dot{\eta}) dt \tag{4.4.3}$$

$$\int_0^{t_1} \Bigg[ \mathcal{L}(x_{cl}, \dot{x}_{cl}) + \left(\frac{\partial \mathcal{L}}{\partial x}\right)_{x_{cl}} \eta + \left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right)_{x_{cl}} \dot{\eta} \tag{4.4.4}$$

$$+ \frac{1}{2}\left(\left(\frac{\partial \mathcal{L}}{\partial x^2}\right)_{x_{cl}} \eta^2 + 2\left(\frac{\partial^2 \mathcal{L}}{\partial x \partial \dot{x}}\right)_{x_{cl}} \eta\dot{\eta} + \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{x}^2}\right)_{x_{cl}} \dot{\eta}^2\right) \Bigg] dt \tag{4.4.5}$$

where we truncated the expansion here since we consider only quadratic Lagrangians. Since $\mathcal{L}$ is a polynomial this Taylor expansion is exact, we do not need to approximate $\eta$ to be small. The first term becomes $S[x_{\text{cl}}]$. We can integrate the second and third term by parts, as we did back when deriving the Euler-Lagrange equation and find

$$\int_0^{t_1} \left[\left(\frac{\partial \mathcal{L}}{\partial x}\right)_{x_{cl}} \eta + \left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right)_{x_{cl}} \dot{\eta}\right] dt = \int_0^{t_1} \left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right]_{x_{cl}} \eta(t) dt = 0 \tag{4.4.6}$$

Consequently we find that

$$S[x_{cl} + y] = S[x_{cl}] + \int_0^t (-c\eta^2 - e\eta\dot{\eta} + \frac{1}{2}m\dot{\eta}^2) dt \tag{4.4.7}$$

and therefore

$$U(x_1, t_1; x_0) = e^{iS_{cl}/\hbar} \overbrace{\int_0^0 \mathcal{D}[\eta(t)]\, \exp\left[\frac{i}{\hbar}\int_0^t \left(\frac{1}{2}m\dot{\eta}^2 - c\eta^2 - e\eta\dot{\eta}\right) dt\right]}^{A(t)} \tag{4.4.8}$$

$$\implies U(x_1, t_1; x_0) = A(t) e^{iS_{cl}/\hbar} \tag{4.4.9}$$

as desired.

## 4.5 Examples of path integrals

### Free particle

The classical path a free particle takes is simply a line through $(x_0, t_0)$ and $(x_1, t_1)$:

$$x_{cl}(t) = x_0 + \frac{x_1 - x_0}{t_1 - t_0}(t - t_1) \implies \mathcal{L} = \frac{1}{2}m\left(\frac{x_1 - x_0}{t_1 - t_0}\right)^2 \tag{4.5.1}$$

Therefore:

$$S_{cl} = \int_{t_0}^{t_1} \mathcal{L}dt = \frac{1}{2}m\frac{(x_1 - x_0)^2}{t_1 - t_0} \tag{4.5.2}$$

Consequently the propagator takes the form:

$$U(x_1, t_1; x_0, t_0) = A' \exp\left[\frac{im(x_1 - x_0)^2}{2\hbar(t_1 - t_0)}\right] \tag{4.5.3}$$

Note that as $(t_1 - t_0) \longrightarrow 0$, we must have $U \longrightarrow \delta(x_1 - x_0)$. Recall the definition of the delta function:

$$\delta(x_1 - x_0) = \lim_{\Delta \to 0} \sqrt{\frac{1}{\pi\Delta^2}} \exp\left[-\frac{(x_1 - x_0)^2}{\Delta^2}\right] \tag{4.5.4}$$

we find that the normalization factor becomes:

$$A' = \sqrt{\frac{m}{2\pi\hbar i(t_1 - t_0)}} \tag{4.5.5}$$

Setting $t_0 = 0$ we find that:

$$U(x, t; x_0) = \sqrt{\frac{m}{2\pi\hbar i t}} \exp\left[\frac{im(x - x_0)^2}{2\hbar t}\right] \tag{4.5.6}$$

We will see in the next chapter that this is the exact propagator we find when solving for the energy eigenvalues. However, we will do so in a much more tedious process than was presented here using the efficient path integral method.

### Harmonic oscillator

## 4.6 The Aharonov-Bohm and Aharonov-Casher effects

The path integral formulation of quantum mechanics is particularly useful in describing interference effects in double slit experiments. There are two special variants of the double slit experiment, namely the Aharonov-Bohm (AB) and Aharonov-Casher (AC) effects that are particularly interesting and exactly solvable using path integrals.

### Double slit interference

To understand the AB and AC effects one must first understand how double slit interference occurs with non-optical media such as electrons. We consider an electron source $P$ firing electrons against an impenetrable barrier with two narrow slits in it. The resulting interference pattern is observed at a point $Q$ on a panel behind the barrier.

The probability amplitude of the particle going from $P(0,0)$ at $t=0$ to $Q(x_f,L)$ at $t=T$ through the first slit at $(a,d)$ is given by multiplying the amplitude of the particle going from $P$ to the first slit, and the amplitude of the particle going from the first slit to $Q$:

$$U_{\text{slit 1}} = \int_{\text{slit 1}} \mathcal{D}[x]\, e^{iS[x]/\hbar} = \int_P^{\text{slit 1}} \mathcal{D}[x]\, e^{iS[x]/\hbar} \times \int_{\text{slit 1}}^{Q} \mathcal{D}[x]\, e^{iS[x]/\hbar} \tag{4.6.1}$$

$$= \left(\frac{m}{2\pi\hbar i}\right)^3 \int dt' \left(\frac{1}{t'(T-t')}\right)^{3/2} \exp\left[\frac{im}{2\hbar}\left(\frac{(x_f-a)^2+d^2}{t'} + \frac{a^2+L^2}{T-t'}\right)\right] \tag{4.6.2}$$

This integral over the time it takes for the particle to go through the slit can be evaluated analytically and mathematica yields

$$U_{\text{slit 1}} = \sqrt{\frac{\pi}{T^3}} \frac{\sqrt{A}+\sqrt{B_1}}{\sqrt{AB_1}} \exp\left[-\frac{(\sqrt{A}+\sqrt{B_1})^2}{T}\right] \tag{4.6.3}$$

where

$$A = -\frac{im(a^2+d^2)}{2\hbar},\ B_1 = -\frac{im[(x_f-a)^2+(L-d)^2]}{2\hbar} \tag{4.6.4}$$

Letting $L \to \infty$ then $B \approx -\frac{imL^2}{2}$ so that the prefactor can be taken to be constant with respect to $x_f$. Similarly we find that the propagator for the particle going in the second slit is

$$U_{\text{slit 2}} = \sqrt{\frac{\pi}{T^3}} \frac{\sqrt{A}+\sqrt{B_2}}{\sqrt{AB_2}} \exp\left[-\frac{(\sqrt{A}+\sqrt{B_2})^2}{T}\right] \tag{4.6.5}$$

where

$$B_2 = -\frac{im[(x_f+a)^2+(L-d)^2]}{2\hbar} \tag{4.6.6}$$

Thus the total propagator from $P$ to $Q$ is

$$U \sim \exp\left[-\frac{(\sqrt{A}+\sqrt{B_1})^2}{T}\right] + \exp\left[-\frac{(\sqrt{A}+\sqrt{B_2})^2}{T}\right] \tag{4.6.7}$$

The intensity of the interference pattern is given by the magnitude of the propagator squared, thus

$$I = |U|^2 \sim 2 + 2\cos\left[-\frac{(\sqrt{A}+\sqrt{B_1})^2}{iT} + \frac{(\sqrt{A}+\sqrt{B_2})^2}{iT}\right] \tag{4.6.8}$$

The argument of the cosine can be simplified to yield

$$-\frac{(\sqrt{A}+\sqrt{B_1})^2}{iT} + \frac{(\sqrt{A}+\sqrt{B_2})^2}{iT} = \frac{B_2-B_1+2\sqrt{A}(\sqrt{B_2}-\sqrt{B_1})}{iT} \tag{4.6.9}$$

$$\approx \frac{2mx_f a}{\hbar T} \tag{4.6.10}$$

and therefore

$$I \sim \cos^2\left(\frac{mx_f a}{\hbar T}\right) \tag{4.6.11}$$

Now note that the electron has de Broglie wavelength $\lambda_{\text{dB}} = \frac{h}{p} = \frac{hT}{mL}$. Letting $\theta$ be the angle from $P$ to $Q$ then $L\sin\theta \approx x_f$ and hence

$$\frac{mx_f a}{\hbar T} = \frac{\pi d \sin\theta}{\lambda} \tag{4.6.12}$$

where $d = 2a$. This implies that we obtain the same interference pattern as if we had used photons

$$I \sim \cos^2\left(\frac{\pi d \sin\theta}{\lambda_{\text{dB}}}\right) \tag{4.6.13}$$

### Aharonov-Bohm interference

Now consider the same set-up as before, only that we place a "fluxon", a line of magnetic moments/solenoid with flux $\Phi$ between the slits and the detector. The Lagrangian for this set-up is

$$L = \frac{\mathbf{p}^2}{2m} - \frac{e}{m}\mathbf{p}\cdot\mathbf{A} + e\phi \tag{4.6.14}$$

Suppose we perform a gauge transformation $\mathbf{A} \mapsto \mathbf{A} - \nabla\chi$ and $\phi \mapsto \phi + \frac{\partial\chi}{\partial t}$. Then we see that the Lagrangian transforms as

$$L \mapsto L + \frac{e}{m}\mathbf{p}\cdot\nabla\chi + e\frac{\partial\chi}{\partial t} \equiv L' + e\frac{d\chi}{dt} \tag{4.6.15}$$

and therefore the action acquires a global translation

$$S \mapsto S + e(\chi_f - \chi_i) \tag{4.6.16}$$

corresponding to a global phase shift of the propagator

$$U \mapsto \exp\left[\frac{ie}{\hbar}(\chi_f - \chi_i)\right] \tag{4.6.17}$$

In our case the most advantageous gauge transformation is that which cancels out $\mathbf{A}$, so that the result of the free-particle propagator may be applied. Since the electron moves in a region with $\mathbf{B} = 0$, we have $\nabla\times\mathbf{A} = 0$ implying that $\mathbf{A} = \nabla\chi$ where

$$\chi(\mathbf{r}) = \int_{\text{P}}^{\mathbf{r}} \mathbf{A}\cdot d\mathbf{l} \implies U(\mathbf{r}, t_f, P, t_i) \mapsto U\exp\left[\frac{ie}{\hbar}\int_P^{\mathbf{r}} \mathbf{A}\cdot d\mathbf{l}\right] \tag{4.6.18}$$

Applying this time-independent gauge transformation we see that $\mathbf{A} = 0$ and, since there is no electric field, $\phi = 0$. The Lagrangian in turns into the free-particle Lagrangian, whose propagator is well known. Thus, taking the gauge phase factor out of the path integral since it is path independent (because $\nabla\times\mathbf{A} = 0$), we find

$$U_{\text{slit 1}} = \exp\left(-\frac{ie}{\hbar}\int_{\text{slit 1}} \mathbf{A}\cdot d\mathbf{l}\right) U^{\text{free}}_{\text{slit 1}} \tag{4.6.19}$$

and

$$U_{\text{slit 2}} = \exp\Big(-\frac{ie}{\hbar}\int_{\text{slit 2}} \mathbf{A}\cdot d\mathbf{l}\Big) U^{\text{free}}_{\text{slit 2}} \tag{4.6.20}$$

Therefore the total propagator from $P$ to $Q$ is

$$U = \exp\Big(-\frac{ie}{\hbar}\int_{\text{slit 2}} \mathbf{A}\cdot d\mathbf{l}\Big)\Big[\exp\Big(-\frac{ie}{\hbar}\oint \mathbf{A}\cdot d\mathbf{l}\Big) U^{\text{free}}_{\text{slit 1}} + U^{\text{free}}_{\text{slit 2}}\Big] \tag{4.6.21}$$

We note that by Stokes' theorem, $\oint \mathbf{A}\cdot d\mathbf{l} = \Phi$ and

$$U \sim \exp\Big(-\frac{ie}{\hbar}\int_{\text{slit 2}} \mathbf{A}\cdot d\mathbf{l}\Big)\Big[e^{-ie\Phi/\hbar}\exp\Big(-\frac{(\sqrt{A}+\sqrt{B_1})^2}{T}\Big) + \exp\Big(-\frac{(\sqrt{A}+\sqrt{B_2})^2}{T}\Big)\Big] \tag{4.6.22}$$

So

$$I \sim 2 + 2\cos\Big[\frac{2mx_f a}{\hbar T} - \frac{e\Phi}{\hbar}\Big] \tag{4.6.23}$$

or alternatively

$$I \sim \cos^2\Big(\frac{mx_f a}{\hbar T} - \frac{e\Phi}{2\hbar}\Big) \tag{4.6.24}$$

Therefore there is a shift in the interference pattern by $\frac{e\Phi}{2\hbar}$. Note that $\frac{e\Phi}{2\hbar} \mapsto \frac{e\Phi}{2\hbar} + \pi$ so an increase in $\Phi$ by $\Phi_0 = \frac{2\pi\hbar}{e}$, the quantum of magnetic flux, does not alter the interference pattern.

### Aharonov-Casher interference

Let's reverse the situation of the Aharonov-Bohm experiment. We take a neutral particle, such as a neutron, with magnetic moment $\boldsymbol{\mu}$ shot at a double slit-set up with a line of electrons of charge density $\lambda$ between the slits and the detector.

The current distribution generating a moment $\boldsymbol{\mu}$ is

$$\mathbf{j} = \nabla\times\mathbf{M}, \;\; \text{where } \boldsymbol{\mu} = \int d^3\mathbf{r}\,\mathbf{M} \tag{4.6.25}$$

By Lorentz transforming the charge density vector we see that to first order in $V/c$

$$\rho = \frac{1}{c^2}\mathbf{J}\cdot\mathbf{V} \tag{4.6.26}$$

so that

$$L = \frac{1}{2}M\mathbf{V}^2 - \int d^3\mathbf{r}\,\varphi(\mathbf{r})\rho(\mathbf{r}) \tag{4.6.27}$$

$$= \frac{1}{2}M\mathbf{V}^2 - \frac{1}{c^2}\int d^3\mathbf{r}\,\varphi(\mathbf{r})\mathbf{V}\cdot(\nabla\times\mathbf{M}) \tag{4.6.28}$$

$$= \frac{1}{2}M\mathbf{V}^2 - \mathbf{V}\cdot(\mathbf{E}\times\boldsymbol{\mu}) \tag{4.6.29}$$

However, this is precisely the same hamiltonian as that of a particle with charge $e$ moving in a vector potential $\mathbf{A} = -\frac{1}{ec^2}(\mathbf{E}\times\boldsymbol{\mu})$. For example, if we have an electron producing a

radial electric field then

$$\mathbf{A} = \frac{\boldsymbol{\mu} \times (\mathbf{r} - \mathbf{R})}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{R}|^3} \tag{4.6.30}$$

which is precisely the vector potential of a magnetic dipole up to some fundamental constants. Going back to the general case, the neutral particle will acquire a phase

$$\varphi = \frac{1}{\hbar c^2} \int_P^{\mathbf{r}} (\mathbf{E} \times \boldsymbol{\mu}) \cdot d\mathbf{l} \tag{4.6.31}$$

We align our axes so that $\boldsymbol{\mu} = \mu\hat{\mathbf{z}}$ and assume that the line charge also lies along the $z$-axis so that $\mathbf{E} = \frac{\lambda}{2\pi\epsilon_0 r}\hat{\mathbf{r}}$. Then

$$\mathbf{A} = -\frac{1}{ec^2}\mathbf{E} \times \boldsymbol{\mu} = \frac{1}{ec^2}\frac{\lambda\mu}{2\pi\epsilon_0 r} = \frac{\Phi}{2\pi r} \text{ where } \Phi = \frac{\mu_0 \lambda \mu}{e} \tag{4.6.32}$$

Thus the interference pattern will be

$$I \sim \cos^2\left(\frac{m x_f a}{\hbar T} - \frac{\mu_0 \lambda \mu}{2\hbar}\right) \tag{4.6.33}$$

## 4.7 Statistical mechanics

### Imaginary time formalism

Consider a particle in 1D inside a potential $V(x)$ connected to a heat reservoir. Due to the entanglement of the particle with the reservoir, we no longer have a wave-function describing the state of the particle but rather a mixed thermal state, a density matrix $\rho$. This density matrix takes the form

$$\rho = \frac{1}{Z} e^{-\beta H}, \ \beta = \frac{1}{k_B T} \tag{4.7.1}$$

where $Z$, the partition function, is $\text{Tr}(e^{-\beta H})$, to ensure normalisation: $\text{Tr}(\rho) = 1$. Given this density matrix, one can compute the expectation value of some operator $A$ as

$$\langle A \rangle = \text{Tr}(\rho A) \tag{4.7.2}$$

Note the striking similarity between $e^{-\beta H}$ and $e^{-iHt/\hbar}$, could we perhaps describe the density matrix using $U(t)$ [1]? The answer is yes, if we define an **imaginary time** $t = -i\tau$ where

$$\tau = \hbar\beta = \frac{\hbar}{k_B T} \tag{4.7.3}$$

If we make this replacement then we can identify $e^{-\beta H} = U(-i\tau)$. However, since $H$ is only necessarily bounded below, $U(-i\tau)$ is well defined if and only if $\tau > 0$. We define the imaginary-time propagator as the **Euclidean propagator**

$$e^{-\beta H} = U(-i\tau) \equiv U_E(\tau) \tag{4.7.4}$$

[1] since we are working in equilibrium conditions, the Hamiltonian is time-independent so $U(t) = e^{-iHt/\hbar}$

Since the propagator formalism allows for imaginary times, we should be able to make an imaginary-time representation of a path integral. We then see that

$$U_E(x', x, \tau) = \int \mathcal{D}[x] e^{-S_E[x]/\hbar}, \text{ where } S_E[x] = \int_0^\tau \Bigg[ \frac{m}{2} \underbrace{\left(\frac{dx}{d\tau}\right)^2 + V(x)}_{\mathcal{L}_E} \Bigg] \tag{4.7.5}$$

where $\mathcal{S}_E$ is the **Euclidean action**, the integral of the Euclidean Lagrangian $\mathcal{L}_E$. The reason these quantities are dubbed "Euclidean" is because the Lorentz invariant $s = x^2 - c^2t^2$ becomes an Euclidean invariant $\sigma = x^2 + c^2\tau^2$ under the imaginary time transformation. Time and space are now treated on equal footing, as in Euclidean space-time. (4.7.5) can be seen by noting that the classical action is

$$iS[x] = i \int_0^t dt \left[ \frac{m}{2} \left( \frac{dx}{dt} \right)^2 - V(x) \right] \tag{4.7.6}$$

so by replacing $t \to -i\tau$ then we obtain

$$S[x] = \int_0^\tau (-i) d\tau \left[ \frac{m}{2} \left( i \frac{dx}{d\tau} \right)^2 - V(x) \right] \implies S_E[x] = iS[x] = \int_0^\tau \left[ \frac{m}{2} \left( \frac{dx}{d\tau} \right)^2 + V(x) \right] \tag{4.7.7}$$

Therefore, a particle moving in Euclidean time will experience an inverted potential. Another important consequence of time becoming imaginary is that $S_E$ must now be bounded below, or else the exponential factor will diverge.

This formalism allows us to calculate the matrix elements of $e^{-\beta H}$ as a path integral. However, we haven't calculated the partition function yet. We have that

$$Z = \int dx \ \langle x | e^{-\beta H} | x \rangle = \int dx \ U_E(x, x, \beta\hbar) = \int dx \int_{\text{periodic}} \mathcal{D}[x] e^{-S_E[x]/\hbar} \tag{4.7.8}$$

where the path integral is taken over periodic paths $x(0) = x(\beta\hbar)$. The larger the temperature, the smaller the period of these orbits (in Euclidean time of course).

The quantum fluctuations of a quantum system at finite temperature can be associated with the thermal fluctuations in the classical statistical ensemble of closed trajectories in euclidean time.

### Operator representation using path integrals

Let's try to calculate

$$\int_{x_i}^{x_f} \mathcal{D}[x] \ e^{iS[x]/\hbar} x(t_1) \tag{4.7.9}$$

We can view the path integral as the integral over $x(t_1) \equiv x_1$ product of path integral from $x_i$ to $x_1$ and from $x_1$ to $x_f$

$$\int_{x_i}^{x_f} \mathcal{D}[x]\, e^{iS[x]/\hbar} x(t_1) = \int dx_1 \Big[ \int_{x_i}^{x_1} \mathcal{D}[x]\, e^{iS[x]/\hbar} \times \int_{x_1}^{x_f} \mathcal{D}[x]\, e^{iS[x]/\hbar} \Big] x_1 \tag{4.7.10}$$

$$= \int dx_1 \, \langle x_f | U(t_f - t_1) | x_1 \rangle \, x_1 \, \langle x_1 | U(t_1 - t_i) | x_i \rangle \tag{4.7.11}$$

$$= \langle x_f | U(t_f - t_1) \hat{x} U(t_1 - t_i) | x_i \rangle \tag{4.7.12}$$

where we added a hat on $x_1$ in the last line to emphasize that it is now an operator. Now using the definition of $U(t)$ we get that

$$\langle x_f, t_1 | x | x_i, t_1 \rangle = \int_{x_i}^{x_f} \mathcal{D}[x]\, e^{iS[x]/\hbar} x(t_1) \tag{4.7.13}$$

in the Schrodinger picture and

$$\langle x_f | x(t_1) | x_i \rangle = \int_{x_i}^{x_f} \mathcal{D}[x]\, e^{iS[x]/\hbar} x(t_1) \tag{4.7.14}$$

in the Heisenberg picture.

# Quantum Dynamics in one dimension

5

In theory, this is all that there is to QM. However, there are several important applications of the postulates we stated in the previous chapter that help model the microscopic world. We will present the simplest of these models, in one dimension, in this chapter. Other applications of QM can be found in statistical mechanics, which constitutes volume 3 of this series, as well as most of the modern physics volume.

## 5.1 Parity

We end this chapter by proving a very useful theorem on the parity of solutions to the Schrödinger equation when the potential is symmetric. More precisely, suppose $V(-x) = V(x)$, then the 1D TISE in the coordinate basis is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \tag{5.1.1}$$

Suppose we substitute $x \mapsto -x$, then:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(-x)}{dx^2} + V(x)\psi(-x) = E\psi(-x) \tag{5.1.2}$$

Consequently $\psi(-x)$ is also a solution.

Now if $\psi(x)$ and $\psi(-x)$ are linearly independent, they form a fundamental set of solutions. So any solution can be expressed as a linear combination of $\psi(x)+\psi(-x)$ and $\psi(x)-\psi(-x)$ as well, which are even and odd respectively.

If instead $\psi(x)$ and $\psi(-x)$ are linearly dependent, then:

$$\psi(x) = c\psi(-x) \implies \psi(-x) = c\psi(x) \tag{5.1.3}$$

by re-substituting $x \mapsto -x$. Then:

$$\psi(x) = c^2\psi(x) \implies c = \pm 1 \tag{5.1.4}$$

So if $c = 1$ then $\psi(x) = \psi(-x)$ and we have an even solution. If $c = -1$ then $\psi(x) = -\psi(-x)$ so we have an odd solution.

Hence any solution to the TISE with symmetric potential can be written as superposition of even and odd solutions.

## 5.2 The Free Particle

For a free particle, TDSE turns into:

$$\hat{H}\left|\psi\right\rangle = i\hbar\frac{d}{dt}\left|\psi\right\rangle = \frac{\hat{p}^2}{2m}\left|\psi\right\rangle \tag{5.2.1}$$

and feeding the stationary state solutions $\left|\psi\right\rangle = \left|E\right\rangle e^{-iEt/\hbar}$ we find the TISE:

$$\hat{H}\left|E\right\rangle = E\left|E\right\rangle = \frac{\hat{p}^2}{2m}\left|E\right\rangle \tag{5.2.2}$$

We now note that if $\hat{p}\left|p\right\rangle = p\left|p\right\rangle$ then $\hat{p}^2\left|p\right\rangle = p^2\left|p\right\rangle$ so feeding one of the eigenstates $\left|p\right\rangle$ into the solution:

$$\left(\frac{p^2}{2m} - E\right)\left|p\right\rangle = \left|0\right\rangle \implies p = \pm\sqrt{2mE} \tag{5.2.3}$$

Therefore a particle with energy E can be described by a superposition of particles moving to the left and right with momentum $\pm\sqrt{2mE}$. The two energy orthogonal eigenstates can be defined as:

$$\left|E_+\right\rangle = \left|p = \sqrt{2mE}\right\rangle \left|E_-\right\rangle = \left|p = -\sqrt{2mE}\right\rangle \tag{5.2.4}$$

Interestingly, by the principle of superposition, if a particle is defined by:

$$\left|E\right\rangle = \alpha\left|E_+\right\rangle + \beta\left|E_-\right\rangle \tag{5.2.5}$$

then when measured it can be either moving to the left or right with momentum $\sqrt{2mE}$.

Now that we have found two orthogonal eigenstates we can construct the appropriate propagator (using $\left|p\right\rangle$ instead of $\left|E\right\rangle$ since we used it as a trial solution):

$$\hat{U}(t) = \int_{-\infty}^{\infty}\left|p\right\rangle\left\langle p\right| e^{-\frac{iE(p)t}{\hbar}}\,dp = \int_{-\infty}^{\infty}\left|p\right\rangle\left\langle p\right| e^{-\frac{ip^2t}{2m\hbar}}\,dp \tag{5.2.6}$$

Then we can evaluate the propagator in the position operator:

$$\left\langle x\left|\,\hat{U}(t)\,\right|x'\right\rangle = \int_{-\infty}^{\infty}\left\langle x\,|\,p\right\rangle\left\langle p\,|\,x'\right\rangle e^{\frac{-ip^2t}{2m\hbar}}\,dp = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{\frac{ip(x-x')}{\hbar}} e^{\frac{ip^2t}{2m\hbar}}\,dp = \sqrt{\frac{m}{2\pi\hbar it}} e^{\frac{im(x-x')^2}{2\hbar t}}$$

So, for any initial condition $\psi(x,0) = \psi_0(x)$ then:

$$\psi(x,t) = \int_{-\infty}^{\infty}\left\langle x\left|\,\hat{U}(t)\,\right|x'\right\rangle\psi_0(x')dx' = \sqrt{\frac{m}{2\pi\hbar it}}\int_{-\infty}^{\infty} e^{\frac{im(x-x')^2}{2\hbar t}}\psi_0(x')dx' \tag{5.2.7}$$

If we instead used $t = t'$ then we would have had to calculate $\left\langle x\left|\,\hat{U}(t)\,\right|x'\right\rangle$.

If the particle was initially localized at $x' = x'_0$, and so $\psi(x',t') = \delta(x' - x'_0)$ then we find

that:

$$\psi(x,t) = \left\langle x \left| \hat{U}(t-t') \right| x' \right\rangle \tag{5.2.8}$$

so the propagator in the position basis is the probability that a particle starting out at $(x_0', t')$ arrives at $(x,t)$. Hence equation 4.1.7 tells us that the probability amplitude of a particle ending up at $(x,t)$ is the sum of the contributions from all $x'$ with weight equal to the associated amplitude.

**Example (5.1.1 Sh)**
Show that the propagator for the free particle may be re-written as:

$$\hat{U}(t) = \sum_{\alpha=\pm} \int_0^\infty \left[ \frac{m}{\sqrt{2mE}} \right] |E,\alpha\rangle \langle E,\alpha| \, e^{-iEt/\hbar} dE \tag{5.2.9}$$

**Solution** We must perform a change of basis from the momentum representation to the energy representation. The best way to do this is to use the relation:

$$p = \pm\sqrt{2mE} \implies dp = \pm \frac{m}{\sqrt{2mE}} \, dE \tag{5.2.10}$$

Due to the degeneracy of the energy eigenstates, we then find that the integral is split into two, one for positive momentum states and one for negative momentum states (which represent a particle moving to the right and to the left respectively):

$$\hat{U}(t) = \int_0^\infty |p\rangle \langle p| \, e^{-\frac{iE(p)t}{\hbar}} \, dp + \int_{-\infty}^0 |p\rangle \langle p| \, e^{-\frac{iE(p)t}{\hbar}} \, dp \tag{5.2.11}$$

Note that the integration bounds for the first become $\int_0^\infty$ whereas for the second they are $\int_\infty^0$ from which we find that:

$$\hat{U}(t) = \int_0^\infty |p_+\rangle \langle p_+| \, e^{-iEt/\hbar} \frac{mdE}{2\sqrt{mE}} - \int_\infty^0 |p_-\rangle \langle p_-| \, e^{-iEt/\hbar} \frac{mdE}{2\sqrt{mE}} \tag{5.2.12}$$

$$= \frac{m}{\sqrt{2mE}} \left( \int_0^\infty |E,+\rangle \langle E,+| \, e^{-iEt/\hbar} dE + \int_0^\infty |E,-\rangle \langle E,-| \, e^{-iEt/\hbar} dE \right) \tag{5.2.13}$$

$$= \sum_{\alpha=\pm} \int_0^\infty \left[ \frac{m}{\sqrt{2mE}} \right] |E,\alpha\rangle \langle E,\alpha| \, e^{-iEt/\hbar} dE \tag{5.2.14}$$

as desired. ◀

**Example. (5.1.2 Sh)**
Solve the TISE for a free particle in the position representation.

**Solution** The TISE reads:

$$\frac{\hat{p}^2}{2m}\left|E\right\rangle = E\left|E\right\rangle \tag{5.2.15}$$

In the $x$-basis, we have that $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ and thus:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_E}{\partial x^2} = E\psi_E \implies \frac{\partial^2\psi_E}{\partial x^2} + \frac{2mE}{\hbar^2}\psi_E = 0 \tag{5.2.16}$$

The general solution can be written as:

$$\psi_E(x) = A\exp\left(\frac{i\sqrt{2mE}}{\hbar}x\right) + B\exp\left(\frac{i\sqrt{2mE}}{\hbar}x\right) \tag{5.2.17}$$

as found earlier. ◀

## 5.3 Wave-packets

we can construct the full stationary state solutions of the free particle by appending $e^{-\frac{iE}{\hbar}t} = e^{-\frac{i\hbar k^2}{2m}t}$ to (4.1.17):

$$\psi_k(x,t) = Ae^{i(kx-\frac{\hbar k^2}{2m}t)} + Be^{-i(kx+\frac{\hbar k^2}{2m}t)} \tag{5.3.1}$$

where we use $k$ as a subscript since for each energy level we have a unique wave-number.

For simplicity, we can let $k$ vary from $-\infty$ to $\infty$ and write:

$$\psi_k(x,t) = Ae^{i(kx-\frac{\hbar k^2}{2m}t)},\ k = \pm\frac{\sqrt{2mE}}{\hbar} \tag{5.3.2}$$

One immediately sees that the stationary states solutions are non-normalizable, they are de Broglie plane waves as postulated in chapter 1!

This means that a free particle can never exist in a stationary state, it is not a physically realizable system. It is therefore impossible to have a free particle of definite energy.

Consequently, we can only have a superposition of stationary states. We can indeed sum over all forms of $A(k)e^{i(kx-\frac{\hbar k^2}{2m}t)}$ with $A(k)$ varying for each wave-number/energy level.

We are taking several stationary states $\psi_k$, and superposing them with relative amplitude $A(k)$. Since $k$ is a continuous spectrum, we integrate:

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} A(k)e^{i(kx-\frac{\hbar k^2}{2m}t)}dk \tag{5.3.3}$$

where the factor $\frac{1}{\sqrt{2\pi}}$ is added for convenience as we shall soon see.

This wave-function is certainly normalizable, at the cost of losing determinism in the momentum representation. We no longer have a particle of definite momentum, but rather in a superposition of a range of $\hbar k$. This is called a **wave packet**.

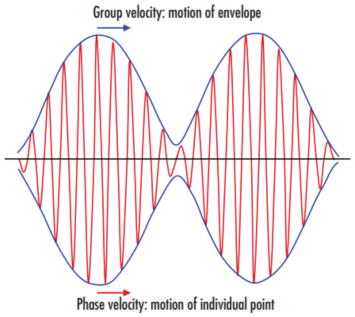
Suppose we are now given an initial wave-function $\psi(x,0)$, we can now use (4.2.3) to evolve the solution at later times. In order to have agreeing boundary conditions, we require:

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k) e^{ikx} dx \tag{5.3.4}$$

In other words, $A(k)$ is the inverse Fourier transform of $\psi(x)$:

$$A(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x,0) e^{-ikx} dx \tag{5.3.5}$$

A typical wave-packet is shown below: Since we superposed waves of different frequen-




cies, we expect them to travel at different velocities according to some dispersion relation.

We can define the phase velocity to be the velocity $v_k$ is the velocity at which the sinusoidal component of any one frequency travels:

$$v_k = \frac{\omega}{k} \tag{5.3.6}$$

The group velocity $v_g$ is instead defined as the velocity of the envelope, the signal as a whole:

$$v_g = \frac{d\omega}{dk} \tag{5.3.7}$$

We can see a similar phenomena when examining ripples in water. Suppose we toss a pebble in a pond, thus creating circular ripples. The group velocity is the speed at which these circular patterns expand. However, the phase velocity is the speed at which the ripples themselves move.

For the free particle, we find that:

$$v_{cl} = v_g = 2v_p \tag{5.3.8}$$

so the classical velocity of a particle agrees with the group velocity, and is half the phase velocity.

## 5.4 Gaussian packets

There is a special class of wave-packets, called **Gaussian packets**, which are of considerable interest when describing particles. It will be shown indeed in the next chapter that this class of wave-functions has minimum uncertainty, it saturates the Heisenberg uncertainty relation:

$$\Delta x \Delta p = \frac{\hbar}{2} \tag{5.4.1}$$

Gaussians therefore are as localized as a wave-function can get, and are a perfect candidate to model particles. Furthermore, they are also the ground state of the reputable harmonic oscillator.

Consider an initial wave function:

$$\psi(x', 0) = e^{ipx'/\hbar} \frac{e^{-x'^2/2\Delta^2}}{(\pi\Delta^2)^{1/4}} \tag{5.4.2}$$

which has mean position $\langle x \rangle = 0$ with uncertainty $\Delta x = \frac{\Delta}{\sqrt{(2)}}$, and mean momentum $\langle p \rangle = p_0$ with uncertainty $\frac{\hbar}{\sqrt{2}}$. We then get using the propagator for a free particle we found earlier:

$$\psi(x,t) = \sqrt{\frac{m}{2\pi\hbar i t}} \int_{-\infty}^{\infty} e^{\frac{im(x-x')^2}{2\hbar t}} e^{ipx'/\hbar} \frac{e^{-x'^2/2\Delta^2}}{(\pi\Delta^2)^{1/4}} dx' \tag{5.4.3}$$

$$= \left[\pi^{1/2}\Big(\Delta + \frac{i\hbar t}{m\Delta}\Big)\right]^{-1/2} \cdot \exp\left[\frac{-(x - p_0 t/m)^2}{2\Delta^2(1 + i\hbar t/m\Delta^2)}\right] \cdot \exp\left[\frac{ip_0}{\hbar}\Big(x - \frac{p_0 t}{2m}\Big)\right] \tag{5.4.4}$$

$$\implies \mathbb{P}(x,t) = \frac{1}{\sqrt{\pi(\Delta^2 + \hbar^2 t^2/m^2\Delta^2)}} \exp\left[\frac{-(x - p_0 t/m)^2}{\Delta^2 + \hbar^2 t^2/m^2\Delta^2)}\right] \tag{5.4.5}$$

It is important to note that:

(i) The mean position of the particle is:

$$\langle x \rangle = \frac{p_0 t}{m} \tag{5.4.6}$$

which is the classical relation we had in Newtonian mechanics.

(ii) The width of the packet grows at the same rate as $\Delta x(t)$:

$$\Delta x(t) = \frac{\Delta}{\sqrt{2}} \sqrt{1 + \frac{\hbar^2 t^2}{m^2 \Delta^4}} \tag{5.4.7}$$

Therefore the initial uncertainty in momentum (or better, velocity) is transferred into an increasing uncertainty in position. Indeed, $\Delta v = \frac{\hbar}{\sqrt{2}m}$ then for large times $\Delta x \approx \Delta v \cdot t = \frac{\hbar t}{\sqrt{2}m}$ which agrees with the above result.

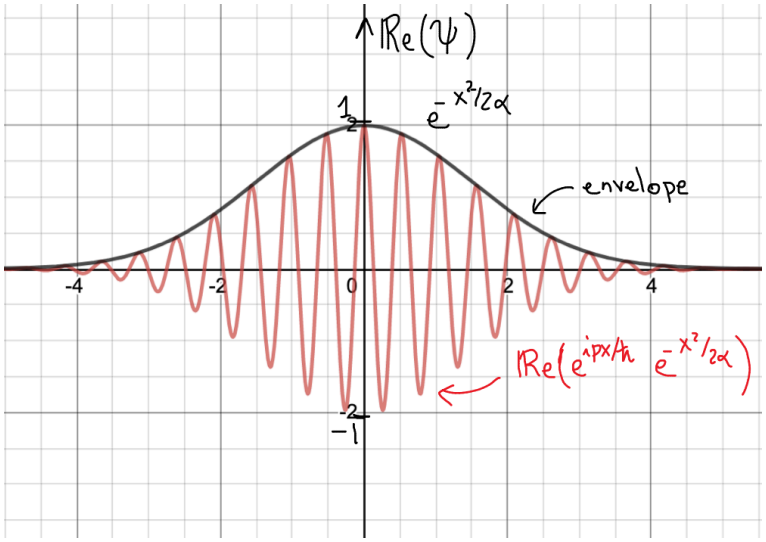



**Figure 5.1**. Plot of the real part of the Gaussian packet

**Example. (5.1.3 Sh)**

Use the alternative definition of the proapagator operator $\hat{U}(t) = e^{i\hat{H}t/\hbar}$ to re-do the gaussian problem, setting $p_0 = 0$ and $\Delta = 1$ so that:

$$\psi(x,0) = \frac{1}{\pi^{1/4}} e^{-x^2/2} \tag{5.4.8}$$

---

**Solution** For a free particle the Hamiltonian is:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} \tag{5.4.9}$$

so we may write the propagator as:

$$\hat{U}(t) = \exp\left[\frac{i\hbar t}{2m}\frac{\partial^{2n}}{\partial x^{2n}}\right] = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{i\hbar t}{2m}\right)^n \frac{\partial^{2n}}{\partial x^{2n}} \tag{5.4.10}$$

and therefore:

$$\psi(x,t) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{i\hbar t}{2m}\right)^n \frac{\partial^{2n}\psi(x,0)}{\partial x^{2n}} \tag{5.4.11}$$

Let us now remark that we may write $\psi(x,0)$ as a power series, which should facilitate

calculating the derivative in 4.2.10:

$$\psi(x,0) = \pi^{-1/4} \sum_{m=0}^{\infty} \Big( -\frac{1}{2} \Big)^m \frac{x^{2m}}{m!} \tag{5.4.12}$$

Then:

$$\frac{\partial^{2n}\psi(x,0)}{\partial x^{2n}} = \pi^{-1/4} \sum_{m=0}^{\infty} \Big( -\frac{1}{2} \Big)^m \frac{x^{2m}}{m!},\ n=0 \tag{5.4.13}$$

$$= \pi^{-1/4} \sum_{m=1}^{\infty} \Big( -\frac{1}{2} \Big)^m \frac{2m(2m-1)x^{2(m-1)}}{m!},\ n=1 \tag{5.4.14}$$

$$= \pi^{-1/4} \sum_{m=2}^{\infty} \Big( -\frac{1}{2} \Big)^m \frac{2m(2m-1)(2m-2)(2m-3)x^{2(m-2)}}{m!},\ n=2 \tag{5.4.15}$$

We can then write:

$$\psi(x,t) = \pi^{-1/4} \sum_{n=0}^{\infty} \frac{1}{n!} \Big( \frac{i\hbar t}{2m} \Big)^n \sum_{i=0}^{\infty} \Big( -\frac{1}{2} \Big)^i \frac{2i(2i-1)...(2i-2n+1)x^{2(i-n)}}{i!} \tag{5.4.16}$$

Looking at the first three terms of $n$:

$$\psi(x,t) = \pi^{-1/4} \Bigg[ \sum_{i=0}^{\infty} \Big( -\frac{1}{2} \Big)^i \frac{x^{2i}}{i!} + \frac{i\hbar t}{2m} \sum_{i=0}^{\infty} \Big( -\frac{1}{2} \Big)^i \frac{2i(2i-1)x^{2(i-1)}}{i!} \tag{5.4.17}$$

$$- \frac{\hbar^2 t^2}{4m^2} \sum_{i=0}^{\infty} \Big( -\frac{1}{2} \Big)^i \frac{2i(2i-1)(2i-2)(2i-3)x^{2(i-2)}}{i!} + ... \Bigg] \tag{5.4.18}$$

Collecting like terms in the same power of $x^k$ we find:

$$\text{term in } x^k = \pi^{-1/4} \Bigg[ \Big( -\frac{1}{2} \Big)^k \frac{x^{2k}}{k!} + \frac{i\hbar t}{2m} \Big( -\frac{1}{2} \Big)^{k+1} \frac{(2k+1)(2k+2)x^{2k}}{(k+1)!} \tag{5.4.19}$$

$$- \frac{\hbar^2 t^2}{4m^2} \Big( -\frac{1}{2} \Big)^{k+2} \frac{(2k+1)(2k+2)(2k+3)(2k+4)x^{2k}}{(k+2)!} + ... \Bigg] \tag{5.4.20}$$

$$= \pi^{-1/4} \Big( -\frac{1}{2} \Big)^k \frac{x^{2k}}{k!} \Bigg( 1 - \frac{i\hbar t}{2m} \frac{1}{2} \frac{(2k+2)(2k+1)}{k+1} + \Big( \frac{i\hbar t}{2m} \Big)^2 \tag{5.4.21}$$

$$\frac{1}{4} \frac{(2k+1)(2k+2)(2k+3)(2k+4)}{(k+1)(k+2)} \Bigg) \tag{5.4.22}$$

$$= \pi^{-1/4} \Big( -\frac{1}{2} \Big)^k \frac{x^{2k}}{k!} \Bigg( 1 + \frac{i\hbar t}{m} \Bigg)^{-k-1/2} \tag{5.4.23}$$

where in the latter step we used the expansion:

$$(1+x)^{-n-1/2} = 1 - (n+\frac{1}{2})x + \frac{(n+\frac{1}{2})(n+\frac{3}{2})}{2!}x^2 + ... \tag{5.4.24}$$

with $x = \frac{i\hbar t}{m}$. Finally we may write:

$$\psi(x,t) = \sum_{k=0}^{\infty} \pi^{-1/4}\Big(-\frac{1}{2}\Big)^k \frac{x^{2k}}{k!}\Big(1+\frac{i\hbar t}{m}\Big)^{-k-1/2} \tag{5.4.25}$$

$$= \Big[\pi^{1/2}\Big(1+\frac{i\hbar t}{m}\Big)\Big]^{-1/2} \sum_{k=0}^{\infty}\Big(-\frac{1}{2}\Big)^k \frac{1}{k!}\Big(\frac{x}{\sqrt{1+\frac{i\hbar t}{2m}}}\Big)^{2k} \tag{5.4.26}$$

$$= \Big[\pi^{1/2}\Big(1+\frac{i\hbar t}{m}\Big)\Big]^{-1/2} \exp\Big[\frac{x^2}{2\Big(1+\frac{i\hbar t}{2m}\Big)}\Big] \tag{5.4.27}$$

as was found earlier. ◀

## 5.5 Infinite potential well

Consider the following potential:

$$V(x) = \begin{cases} 0, \ |x| < \frac{L}{2} \\ \infty \ \text{otherwise} \end{cases} \tag{5.5.1}$$

which is called an **infinite well**. The TISE turns into:

$$\hat{H}\,|E\rangle = E\,|E\rangle \implies \frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E-V)\psi = 0 \tag{5.5.2}$$

In the region outside the box, initially set $V = V_0 > E$ so that:

$$\frac{d^2\psi}{dt^2} - \frac{2m}{\hbar^2}(V-E)\psi = 0 \tag{5.5.3}$$

which is the free particle equation, and has solution:

$$\psi = C_1 e^{-\kappa x} + C_2 e^{\kappa x}, \ \kappa = \sqrt{\frac{2m}{\hbar^2}(V-E)} \tag{5.5.4}$$

However, because the wave function must be normalizable we can't have $\psi$ blowing up as $x \to \infty$. Therefore, the only viable alternative, on physical grounds, is to set $C_1 = C_2 = 0$. Hence:

$$\psi(x) = 0, \ \text{ outside of box} \tag{5.5.5}$$

Let us instead consider the region inside the box. Here, $V = 0$ so we get the free particle solution:

$$\psi = C_1 e^{-i\kappa x} + C_2 e^{i\kappa x}, \ \kappa = \sqrt{\frac{2mE}{\hbar^2}} \tag{5.5.6}$$

so the energy spectrum is seemingly continuous. However, note that we require the wave function to be continuous, especially on the boundaries of the box at $x = \pm\frac{L}{2}$. We must therefore set the boundary conditions:

$$\psi\left(\pm\frac{L}{2}\right) = C_1 e^{\frac{-i\kappa L}{2}} + C_2 e^{\frac{i\kappa L}{2}} = 0 \implies \begin{pmatrix} e^{-i\kappa L/2} & e^{i\kappa L/2} \\ e^{i\kappa L/2} & e^{-i\kappa L/2} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.5.7}$$

whose non-trivial solutions are if the determinant vanishes so:

$$e^{-i\kappa L} - e^{i\kappa L} = 2i\sin(\kappa L) = 0 \implies \kappa = \frac{n\pi}{L},\ n \in \mathbb{Z} \tag{5.5.8}$$

which gives:

$$C_1 e^{-in\pi/2} + C_2 e^{in\pi/2} = 0 \implies C_1 = -\cos(n\pi)B \tag{5.5.9}$$

Finally, normalizing the solution we find that:

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right),\ n \text{ even} \\ \sqrt{\frac{2}{L}}\cos\left(\frac{n\pi x}{L}\right),\ n \text{ odd} \end{cases} \tag{5.5.10}$$

with a discrete spectrum of energy:

$$E_n = \frac{\hbar^2\pi^2 n^2}{2mL^2} \tag{5.5.11}$$

This quantum state in which the potential is greater than the energy $E$, is called a **bound state**, states where the particle is prevented from escaping to infinity.

Most importantly, the energy levels of bound states are always quantized and non-degenerate. Indeed, for a finite well, the conditions we impose on the wave function are:

(i) It must be continuous, especially at the boundaries of the box.

(ii) In the outside region, we need the rising exponential term to be zero.

which are four conditions, but we only have three parameters (the two falling exponential terms, the sine and cosine terms, seem to be four, but since we can scale them by any amount and not change the eigenvalue problem, we only have three all relative to an arbitrary proportionality constant). Therefore only special values for energy will yield a solution that satisfies these conditions.

More generally, for some potential which tends to $V_\pm$ as $|x| \to \infty$ binding a particle of energy $E < V_\pm$, we still have one more constraint than parameters. Indeed we divide space into tiny intervals of constant potential, and which therefore get longer and longer as $|x| \to \infty$ due to the stabilization of $V$. For each interval $\psi$ has two parameters, the coefficients of the growing/falling exponentials. The constraints however are the continuity of $\psi, \psi'$ at the interface between intervals. However, since adding one new interval every time means adding two new parameters and two new constraints. So going from the three interval case of a finite well to the infinite case of an arbitrary potential, we are still going to have the same mismatch of constraints and parameters as in the finite box, so one extra constraint.

One entirely new concept of the quantum mechanical approach to the particle in the box is that the lowest energy level, the ground state, is not zero, but $\frac{\hbar^2\pi^2}{2mL^2}$. This is due to the uncertainty principle. Consider:

$$\left\langle \bar{H} \right\rangle = \frac{\langle \hat{p}^2 \rangle}{2m} \tag{5.5.12}$$

In a bound state, $\langle \hat{p} \rangle = 0$. This is because bound states are stationary, and thus $\langle \hat{p} \rangle$ is time independent, so if it were non-zero then the particle would be capable of moving to infinity, which is not possible for a bound state. Therefore, using $\langle \hat{p}^2 \rangle = (\Delta p)^2 + (\langle \hat{p} \rangle)^2$:

$$\left\langle \bar{H} \right\rangle = \frac{(\Delta p)^2}{2m} \geq \frac{\hbar^2}{8m(\Delta x)^2} \tag{5.5.13}$$

Finally, since the particle is constrained inside the box, $\Delta x \leq \frac{L}{2}$ and so:

$$\left\langle \bar{H} \right\rangle \geq \frac{\hbar^2}{2mL^2} \tag{5.5.14}$$

In the ground state (which is an energy eigenstate), $\left\langle \bar{H} \right\rangle = E$ so that we reach:

$$E \geq \frac{\hbar^2}{2mL^2}. \tag{5.5.15}$$

which is the same order of magnitude as the actual ground state energy.

Finally let us construct the propagator as:

$$\hat{U}(t) = \sum_{n=1}^{\infty} |n\rangle \langle n| \exp\left[ -\frac{i}{\hbar}\Big(\frac{\hbar^2\pi^2 n^2}{2mL^2}\Big)t \right] \tag{5.5.16}$$

so that:

$$\left\langle x \,\middle|\, \hat{U}(t) \,\middle|\, x' \right\rangle = \sum_{n=1}^{\infty} \psi_n(x)\psi_n^*(x') \exp\left[ -\frac{i}{\hbar}\Big(\frac{\hbar^2\pi^2 n^2}{2mL^2}\Big)t \right] \tag{5.5.17}$$

and hence for some initial state $\psi(x', 0)$ we get:

$$\psi(x,t) = \sum_{n=1}^{\infty} \left( \int_{-\infty}^{\infty} \psi_n(x)\psi_n^*(x') \exp\left[ -\frac{i}{\hbar}\Big(\frac{\hbar^2\pi^2 n^2}{2mL^2}\Big)t \right] \psi(x',0)dx' \right) \tag{5.5.18}$$

$$= \sum_{n=1}^{\infty} c_n \psi_n(x) \exp\left[ -\frac{i}{\hbar}\Big(\frac{\hbar^2\pi^2 n^2}{2mL^2}\Big)t \right] \tag{5.5.19}$$

where

$$c_n = \int_{-\infty}^{\infty} \psi_n^*(x')\psi(x',0)dx' \tag{5.5.20}$$

is the probability that, upon an energy measurement, the wave-function collapses into the energy wave-function describing $|\psi_n\rangle$.

**Example (5.2.1 Sh)**
A particle is in the ground state of an infinite well of length $L$. Suddenly the box expands to twice its size, leaving the wave function undisturbed. What is the probability of finding the particle in the ground state of the new box?

**Solution** The particle is initially in the ground state of an infinite well of length $L$ so

$$\psi(x,0) = \sqrt{\frac{2}{L}} \cos\left(\frac{\pi x}{L}\right) \tag{5.5.21}$$

The ground state of the new well of length $2L$ is then:

$$\psi_{E_1}(x) = \sqrt{\frac{1}{L}} \cos\left(\frac{\pi x}{2L}\right) \tag{5.5.22}$$

Therefore, using the second postulate of quantum mechanics, the probability of finding the particle in the new ground state is:

$$P(E_1) = |\langle E_1 | \psi(0)\rangle|^2 = \left(\frac{\sqrt{8}}{3\pi} \int_{\frac{L}{2}}^{\frac{L}{2}} \cos\left(\frac{\pi x}{L}\right) \cos\left(\frac{\pi x}{2L}\right) dx\right)^2 = \frac{64}{9\pi^2} \tag{5.5.23}$$

which is around 0.72, so it is quite likely to make such a measurement. ◀

**Example. (5.2.2 Sh)**
Show that $\left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle \geq E_0$ where $E_0$ is the ground state. (b) Prove that every attractive potential has at least one bound state.
*Hint*: use the following mathematical ansatz

$$\psi(x) = \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2/2} \tag{5.5.24}$$

**Solution** Let us expand $|\psi\rangle$ in the energy eigenbasis so $|\psi\rangle = \sum_i \langle E_i | \psi\rangle |\psi\rangle$. Then, the expectation value for the hamiltonian operator is:

$$\langle H\rangle_\psi = \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle = \langle\psi| \sum_i \langle E_i | \psi\rangle \hat{H} |E_i\rangle \tag{5.5.25}$$

$$= \langle\psi| \sum_i \langle E_i | \psi\rangle E_i |E_i\rangle \tag{5.5.26}$$

$$= \sum_i \langle E_i | \psi\rangle E_i \langle\psi | E_i\rangle \tag{5.5.27}$$

$$= \sum_i |\langle E_i | \psi\rangle|^2 E_i = \sum_i c_i E_i \tag{5.5.28}$$

Clearly then, because $E_i \geq E_0$, we must have that:

$$\left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle \geq E_0 \tag{5.5.29}$$

with equality holding only when $c_i = 0$ for $i \neq 0$.

An attractive potential $V(x)$ is such that $V(x) = -|V(x)|$ and $V(\pm\infty) = 0$. Thus, for a bound state $E$ to exist we need its energy to be negative. Using the result in part a, one might try to find $\langle H \rangle_\psi$ and show that it can be made negative, so that the ground state has a negative energy.

We therefore need to show that $\langle H \rangle_\psi < 0$:

$$\left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle = \left\langle \psi \middle| -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} - \middle| V(x) \middle\| \psi \right\rangle = -\frac{\hbar^2}{2m}\left\langle \psi \middle| \frac{d^2\psi}{dx^2} \right\rangle - \langle \psi \,\|\, V(x) \,\|\, \psi \rangle \tag{5.5.30}$$

We can't proceed further without using a trial solution for $\psi_\alpha$.

$$\left\langle \hat{T} \right\rangle_\alpha = -\sqrt{\frac{\alpha}{\pi}}\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} \alpha(1-\alpha x^2)e^{-\alpha x^2}dx = \frac{\alpha\hbar^2}{4m} \tag{5.5.31}$$

Now to have a bound state we must therefore have that:

$$E_0 \leq \left\langle \hat{H} \right\rangle_\alpha \leq 0 \implies \left\langle \hat{T} \right\rangle_\alpha = \frac{\alpha\hbar^2}{4m} \leq \left\langle \middle| \hat{V} \middle| \right\rangle_\alpha \tag{5.5.32}$$

However, evaluating $\left\langle \middle| \hat{V} \middle| \right\rangle_\alpha$ is impossible to do explicitly, and we must therefore find some bounds for it. Let $|V(x_0) = 2n_0$ for some $x_0$ where the potential is piece wise continuous. Then, $\exists[x_0, x_1] \ni x_0$ such that $|V(x)| \geq n_0,\ \forall x \in [x_0, x_1]$. Therefore we find the following lower bound:

$$\sqrt{\frac{\alpha}{\pi}}\int_{-\infty}^{\infty} |V(x)|e^{-\alpha x^2}dx \geq \sqrt{\frac{\alpha}{\pi}}\int_{x_1}^{x_2} |V(x)|e^{-\alpha x^2}dx \geq \sqrt{\frac{\alpha}{\pi}}\int_{x_1}^{x_2} n_0 e^{-\alpha x^2}dx \tag{5.5.33}$$

where the first step is justified since the integrand is positive everywhere. We therefore need to find a value for $\alpha$ such that:

$$\left\langle \middle| \hat{V} \middle| \right\rangle_\alpha \geq \sqrt{\frac{\alpha}{\pi}}\int_{x_1}^{x_2} n_0 e^{-\alpha x^2}dx \geq \frac{\alpha\hbar^2}{4m} \tag{5.5.34}$$

Now because $\int_{x_1}^{x_2} n_0 e^{-\alpha x^2}dx \leq \int_{x_1}^{x_2} n_0 e^{-\alpha x_1}dx = n_0 e^{-\alpha x_1^2}(x_2 - x_1)$ then:

$$\left\langle \middle| \hat{V} \middle| \right\rangle_\alpha \geq \sqrt{\frac{\alpha}{\pi}} n_0 e^{-\alpha x_1^2}(x_2 - x_1) \geq \frac{\alpha\hbar^2}{4m} \tag{5.5.35}$$

If we choose $\alpha \leq \frac{1}{x_1^2} \implies e^{-\alpha x_1^2} \leq \frac{1}{e}$ then:

$$\alpha \leq \left(\frac{4mn_0(x_2 - x_1)}{e\hbar^2\sqrt{\pi}}\right)^2 \wedge \alpha \leq \frac{1}{x_1^2} \tag{5.5.36}$$

Which can always be solved since they're both positive upper bounds. Therefore the ground state is always a bound state in an attractive potential. ◀

**Example. (5.2.3 Sh)**
Find the bounded energy levels for a particle under a potential $V(x) = aV_0\delta(x)$.

**Solution** Clearly, for a bounded state we must have a negative energy, since $V(\pm\infty) = 0$.Let us consider the TISE:

$$\hat{H}\,|E\rangle = E\,|E\rangle \tag{5.5.37}$$

$$\implies \frac{\hat{p}^2}{2m}\,|E\rangle + V(\hat{x})\,|E\rangle = E\,|E\rangle \tag{5.5.38}$$

$$\implies -\frac{\hbar^2}{2m}\frac{\partial^2\psi_E(x)}{\partial x^2} + V(x)\psi_E(x) = E\psi_E(x) \tag{5.5.39}$$

$$\implies \frac{\partial^2\psi_E(x)}{\partial x^2} + \frac{2m}{\hbar^2}(E - V(x)) = 0 \tag{5.5.40}$$

We must now consider three different regions, corresponding to different values of $V(x)$. For $x < 0$ we have that $V(x) = 0$ and thus the solution is that of a free particle:

$$\psi_E(x) = Ae^{kx} + Be^{-kx},\ k = \frac{\sqrt{2m|E|}}{\hbar} \tag{5.5.41}$$

However, because the second term blows up at $-\infty$, we must set $B = 0$ so that:

$$\psi_E(x) = Ae^{kx} \tag{5.5.42}$$

Similarly, for $x < 0$ we find:

$$\psi_E(x) = Ce^{-kx} = Ae^{-kx} \tag{5.5.43}$$

where for the sake of continuity, we must have that $A = C$. Finally, for $x = 0$, the TISE turns into :

$$\frac{\partial^2\psi_E(x)}{\partial x^2} + \frac{2m}{\hbar^2}(E - aV_0\delta(x))\psi_E(x) = 0 \tag{5.5.44}$$

Integrating over an infinitesimal interval $[-\epsilon, \epsilon]$ we find:

$$\frac{d}{dx}\int_{-\epsilon}^{\epsilon}\frac{\partial\psi_E(x)}{\partial x}dx = \frac{2m}{\hbar^2}\int_{-\epsilon}^{\epsilon}(E - aV_0\delta(x))\psi_E(x) = -\frac{2maAV_0}{\hbar^2} \tag{5.5.45}$$

This is the change in the slope of $\psi_E(x)$ around the cusp at $x = 0$. Using the values of derivative of the wave function to the left and right of $x = 0$:

$$\frac{d\psi_E(0-)}{dx} = Ak \tag{5.5.46}$$

$$\frac{d\psi_E(0+)}{dx} = -Ak \tag{5.5.47}$$

so that the change in slope is equal to $2Ak$, then:

$$2Ak = -\frac{2maAV_0}{\hbar^2} \implies \frac{2m|E|}{\hbar^2} = \frac{m^2a^2V_0^2}{\hbar^4} \implies E = -|E| = -\frac{ma^2V_0^2}{2\hbar^2} \tag{5.5.48}$$

So the only bounded energy level (the only negative one) is $E = -\frac{ma^2V_0^2}{2\hbar^2}$. ◀

**Example. (5.2.4 Sh)**

A particle of mass $m$ is in the state $|n\rangle$ of a box of length $L$. Find the force encountered when the walls are quasistatically pushed in assuming the particle remains in the nth state of the box as it changes, in the classical case and quantum case.

---

**Solution** In the quantum mechanical case:

$$F = -\frac{\partial \mathcal{H}}{\partial L} = -\frac{\partial E_n}{\partial L} = -\frac{\partial}{\partial L}\Big(\frac{\hbar^2\pi^2n^2}{2mL^2}\Big) = \frac{\hbar^2\pi^2n^2}{mL^3} \tag{5.5.49}$$

For an electron in the ground state of a well of length $L = 10^{-10}m$, the force is about $F \approx 1.20 \times 10^{-7}N$.

Now let us work out the classical case. Here:

$$E_n = \frac{\hbar^2\pi^2n^2}{2mL^2} = \frac{p^2}{2m} \implies p = \frac{\hbar\pi n}{L} \implies v = \frac{\hbar\pi n}{mL} \tag{5.5.50}$$

Therefore, the time taken by the particle to hit the same wall twice is:

$$\Delta t = \frac{2L}{v} = \frac{2mL^2}{\hbar\pi n} \tag{5.5.51}$$

and so the frequency of collision on one wall is:

$$\nu = \frac{\hbar\pi n}{2mL^2} \tag{5.5.52}$$

Finally, since the change in momentum is $\Delta p = \frac{2\hbar\pi n}{L}$ then:

$$F = \nu\Delta p = \frac{\hbar^2\pi^2n^2}{mL^3} \tag{5.5.53}$$

which coincides with the quantum prediction. Furthermore, we can write:

$$F = \frac{\partial p}{\partial t} = \frac{\partial p}{\partial L}\frac{\partial L}{\partial t} \implies \frac{\hbar^2\pi^2n^2}{mL^3} = -\frac{\hbar\pi n}{L^2}\frac{\partial L}{\partial t} \tag{5.5.54}$$

so that the rate at which the box shrinks is given by:

$$\frac{\partial L}{\partial t} = -\frac{\hbar \pi n}{mL} \tag{5.5.55}$$

The quantization in the energy levels of the particle leads to a quantization of the shrinking speed of the box. ◀

## 5.6 Finite potential well

Let us now consider the problem of a symmetric potential well, with potential:

$$V(x) = \begin{cases} V_0 & |x| \geq L \text{ region 1} \\ 0 & |x| \leq L \text{ region 2} \end{cases} \tag{5.6.1}$$

with $E < V_0$. Then the TISE has solution in region 1:

$$\psi(x) = c_1 e^{\kappa x} + c_2 e^{-\kappa x},\ \kappa = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \tag{5.6.2}$$

However because the wave function must be normalisable, we need it to vanish at infinity and so:

$$\psi(x) = \begin{cases} c_1 e^{\kappa x}\ x \leq -L \\ c_2 e^{-\kappa x}\ x \geq L \end{cases} \tag{5.6.3}$$

Similarly, in region 2 the solution is akin to the free particle:

$$\psi(x) = b_1 \cos(kx) + b_2 \sin(kx),\ k = \sqrt{\frac{2mE}{\hbar^2}} \tag{5.6.4}$$

Because the potential is symmetric, it can be proven that the bound states are either even or odd. We can now impose the continuity of the wave function and its first derivative at the boundaries of the well:

(i) $\psi(-L^-) = \psi(-L^+)$ then: $b_1 \cos(kL) - b_2 \sin(kL) = c_1 e^{-\kappa L}$

(ii) $\psi(L^-) = \psi(L^+)$ then: $b_1 \cos(kL) + b_2 \sin(kL) = c_2 e^{-\kappa L}$

(iii) $\psi'(-L^-) = \psi'(-L^+)$ then: $k(b_1 \sin(kL) + b_2 \cos(kl)) = c_1 \kappa e^{-\kappa L}$

(iv) $\psi'(L^-) = \psi'(L^+)$ then: $k(-b_1 \sin(kL) + b_2 \cos(kL)) = -c_2 \kappa e^{-\kappa L}$

### Odd solutions

Because we are looking for even solutions, we must set $b_1 = 0$. The first and third conditions give:

$$\begin{cases} -b_2 \sin(kL) = c_1 e^{-\kappa L} \\ k b_2 \cos(kL) = c_1 \kappa e^{-\kappa L} \end{cases} \implies -k \cot(kL) = \kappa,\ k \neq (n - \frac{1}{2})\pi \tag{5.6.5}$$

### Even solutions

Because we are looking for odd solutions, we must $b_2 = 0$. The second and fourth conditions give:

$$\begin{cases} b_1 \cos(kL) = c_2 e^{-\kappa L} \\ -kb_1 \sin(kL) = -c_2 \kappa e^{-\kappa L} \end{cases} \implies k\tan(kL) = \kappa,\ k \neq \frac{n\pi}{2} \tag{5.6.6}$$

Note that irrelevant of the nature of the solutions, we have that:

$$k^2 + \kappa^2 = \frac{2mV_0}{\hbar^2} \tag{5.6.7}$$

Therefore, to find the bounded state solutions, we need to look at all the intersections of 4.4.5 or 4.4.6 with 4.4.7 , and number them with $n = 1, 2....$ The corresponding energy level is then given by

$$E = \frac{k^2\hbar^2}{2m} \tag{5.6.8}$$

for each $k$ where an intersection occurs.

**Example. (5.2.6 Sh)**
Show that there are always even solutions to the particle in the finite well, and only odd solutions when $V_0 \geq \frac{\hbar^2\pi^2}{8mL^2}$.

---

**Solution** Let us first prove that even solutions always exist. This can be seen graphically since $k\tan(kL)$ always passes by the origin, and so must intersect the circle at some points (see https://www.desmos.com/calculator/pgangnya2y?lang=it). Mathematically, we want to solve:

$$k^2(1 + \tan^2(kL)) = k^2\sec^2(kL) = \frac{2mV_0}{\hbar^2} \tag{5.6.9}$$

The only cases where an intersection may not occur is when the secant function is discontinuous. This would mean that $(n - \frac{1}{2})\pi$, which is a contradiction with 4.4.5 Therefore we always have a solution.

In the case of an odd solution, from the diagram we see that the first positive solution occurs only for some values of $V_0$. Indeed, we need the radius of the circle to be greater than the x-intercept of $-k\cot(kL)$. Therefore, for the x-intercept $\kappa = 0 \implies k = (n - \frac{1}{2})\pi$, and setting $n = 1$ because we are looking for the first solution, we get $k = \frac{\pi}{2}$. So:

$$\sqrt{\frac{2mV_0}{\hbar^2}} \geq \frac{\pi}{2} \implies V_0 \geq \frac{\hbar^2\pi^2}{8mL^2} \tag{5.6.10}$$

When $V_0 = \frac{\hbar^2\pi^2}{8mL^2}$, we have that the intersection occurs at $k = \frac{\pi}{2}$ so $E = \frac{\pi^2\hbar^2}{8m}$ is the ground state. ◀

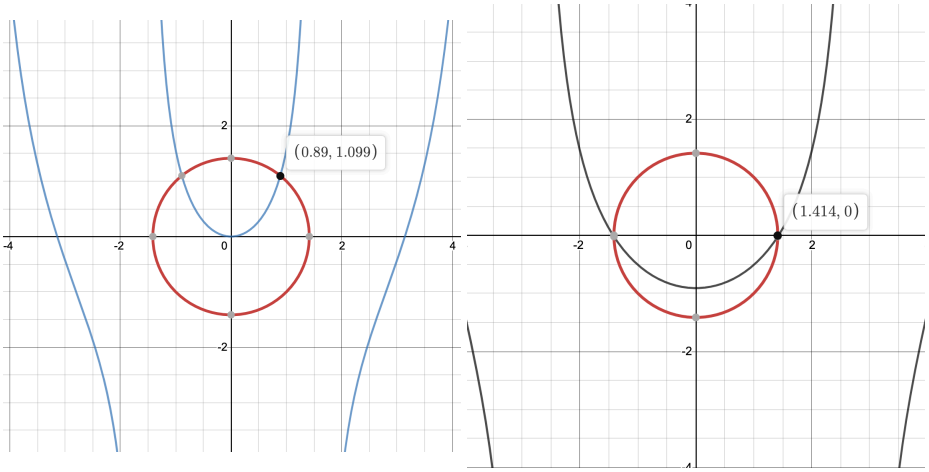



**Figure 5.2**. Plots of the even (left) and odd (right) solutions to the finite square well problem

## 5.7 Conservation of Probability: the Continuity equation

In electromagnetism, global conservation of charge states that the total charge in the universe is constant, and local conservation of charge states that any change in charge in a volume $\mathcal{V}$ is accounted by some flow of current through it:

$$\frac{\partial \rho(\mathbf{r},t)}{\partial t} = -\nabla \cdot \mathbf{J} \iff \frac{d}{dt}\int_{\mathcal{V}} \rho(\mathbf{r},t) d\mathbf{r}^3 = -\int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{S} \tag{5.7.1}$$

where $\rho$, $\mathbf{J}$ are the charge and current densities. Local conservation forbids phenomena such as charge suddenly appearing/disappearing in two different regions.

Similarly, in QM the globally conserved quantity is the total probability of finding a particle anywhere in the universe. Indeed, because the time-evolution of the wave function, the propagator operator, is unitary, we must have that:

$$\langle \psi(t) \,|\, \psi(t) \rangle = \left\langle \psi(0) \,\middle|\, \hat{U}^\dagger \hat{U} \,\middle|\, \psi(0) \right\rangle = \langle \psi(0) \,|\, \psi(0) \rangle \tag{5.7.2}$$

therefore we find the analogue to the global conservation of charge:

$$\langle \psi(t) \,|\, \psi(t) \rangle = \iiint \langle \psi(t) \,|\, \mathbf{r} \rangle \langle \mathbf{r} \,|\, \psi(t) \rangle \, d^3\mathbf{r} = \iiint \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) d^3\mathbf{r} = \iiint \mathbb{P}(\mathbf{r},t) d^3\mathbf{r} = \text{constant} \tag{5.7.3}$$

where we define the probability density in three dimensions:

$$\mathbb{P}(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2 \tag{5.7.4}$$

To derive the local conservation law in QM, consider the TDSE and its complex conjugate:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \psi + V\psi \tag{5.7.5}$$

$$-i\hbar \frac{\partial \psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \psi^* + V\psi^* \tag{5.7.6}$$

Multiplying the first by $\psi^*$ and the second by $\psi$, and then taking the difference:

$$i\hbar\frac{\partial}{\partial t}(\psi^*\psi) = -\frac{\hbar^2}{2m}(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*) \tag{5.7.7}$$

$$\implies \frac{\partial \mathbb{P}}{\partial t} = -\frac{\hbar}{2mi}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*) \tag{5.7.8}$$

$$\implies \frac{\partial \mathbb{P}}{\partial t} = -\nabla\cdot\mathbf{J} \tag{5.7.9}$$

where we define the **probability current density** to be:

$$\mathbf{J} = \frac{\hbar}{2mi}(\psi^*\nabla\psi - \psi\nabla\psi^*) = \frac{\hbar}{m}\mathrm{Im}\Big(\psi^*\nabla\psi\Big) \tag{5.7.10}$$

An immediate yet powerful consequence of the local conservation of probability density is that if the wave function is normalised at some $t'$, then it will be normalised for all subsequent $t > t'$. Indeed, consider a normalised wave function $\psi(\mathbf{r}, t')$ so that:

$$\mathcal{N}(t') \equiv \iiint \mathbb{P}(\mathbf{r}, t')d^3\mathbf{r} = 1 \tag{5.7.11}$$

Then we find that:

$$\frac{d\mathcal{N}(t')}{dt} = \frac{d}{dt}\iiint \mathbb{P}(\mathbf{r}, t')d^3\mathbf{r} = \iiint \nabla\cdot\mathbf{J}d^3\mathbf{r} = \iint \mathbf{J}d^2\mathbf{r} \tag{5.7.12}$$

Note that since the wave function must vanish at infinity (and its first derivative must be bounded), the latter surface integral must be null, so that:

$$\frac{d\mathcal{N}(t')}{dt} = 0 \tag{5.7.13}$$

and normalisation is preserved.

**Example. (5.3.1 Sh)**
Consider the case where $V(\mathbf{r} = V_r(\mathbf{r} - iV_i$, and $V_i$ is constant. Describe the time evolution of the probability of finding the particle.

**Solution** We can clearly see that $V$ is not Hermitian due to the imaginary part $V_i$. Therefore we can't use the result that $\frac{d\mathcal{N}(t')}{dt} = 0$. Therefore consider the TDSE and its complex conjugate:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V_r(\mathbf{r})\psi - iV_i\psi \tag{5.7.14}$$

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + V_r(\mathbf{r})\psi^* + iV_i\psi^* \tag{5.7.15}$$

Multiplying the first by $\psi^*$ and the second by $\psi$, and then taking the difference:

$$i\hbar\frac{\partial}{\partial t}(\psi^*\psi) = -\frac{\hbar^2}{2m}(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*) - 2iV_i\psi^*\psi \tag{5.7.16}$$

$$\implies \frac{\partial \mathbb{P}}{\partial t} = -\frac{\hbar}{2mi}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*) - \frac{2V_i\mathbb{P}}{\hbar} \tag{5.7.17}$$

$$\implies \frac{\partial \mathbb{P}}{\partial t} + \frac{2V_i}{\hbar}\mathbb{P} = -\nabla\cdot\mathbf{J} \tag{5.7.18}$$

We then find that the normalisation constant $\mathcal{N}$ evolves as:

$$\frac{d\mathcal{N}(t')}{dt} = \frac{d}{dt}\iiint \mathbb{P}(\mathbf{r},t')d^3\mathbf{r} = \iiint \nabla\cdot\mathbf{J}d^3\mathbf{r} - \frac{2V_i}{\hbar}\iint \mathbb{P}(\mathbf{r},t')d^3\mathbf{r} \tag{5.7.19}$$

$$= \iint \mathbf{J}d^2\mathbf{r} - \frac{2V_i}{\hbar}\iint \mathbb{P}(\mathbf{r},t')d^3\mathbf{r} \tag{5.7.20}$$

Note that since the wave function must vanish at infinity (and its first derivative must be bounded), if we evaluate the latter surface integral over a sphere of infinite radius, it must become null, so that:

$$\frac{d\mathcal{N}(t')}{dt} = -\frac{2V_i}{\hbar}\iint \mathbb{P}(\mathbf{r},t')d^3\mathbf{r} = -\frac{2V_i}{\hbar}\mathcal{N} \tag{5.7.21}$$

This differential equation can be easily solved, and we find that:

$$\mathcal{N} \propto e^{-2V_it/\hbar} \tag{5.7.22}$$

The probability of finding the particle therefore decreases exponentially, with a half life of $\frac{\hbar}{2V_i}\ln(2)$. This potential is very useful when modelling processes where a particle is absorbed. Here, the term $-\frac{2V_i}{\hbar}\mathcal{N}$ is called the sink term, because it is negative. Had it been positive (so with a potential $V(\mathbf{r}) = V_r(\mathbf{r}) + iV_i$) then it would be a source term. ◀

**Example. (5.3.3 Sh)**
Find the probability current density and probability density for:

$$\psi_{\mathbf{p}} = \left(\frac{1}{2\pi\hbar}\right)^{3/2} e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar} \tag{5.7.23}$$

**Solution** Let us firstly find the probability current density:

$$\mathbf{J} = \frac{\hbar}{2mi}\left(\left(\frac{1}{2\pi\hbar}\right)^3 e^{-i(\mathbf{p}\cdot\mathbf{r})/\hbar}\,\nabla e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar} - \left(\frac{1}{2\pi\hbar}\right)^3 e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar}\,\nabla e^{-i(\mathbf{p}\cdot\mathbf{r})/\hbar}\right) \tag{5.7.24}$$

$$= \frac{\hbar}{2mi(2\pi\hbar)^3}\left(\frac{i\mathbf{p}}{h} + \frac{i\mathbf{p}}{h}\right) \tag{5.7.25}$$

$$= \frac{\mathbf{p}}{(2\pi\hbar)^3 m} \tag{5.7.26}$$

where the gradient was calculated as follows:

$$\nabla e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar} = \nabla e^{i(xp_x+yp_x+zp_z)/\hbar} \tag{5.7.27}$$

$$= \frac{ip_x}{\hbar}e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar}\hat{x} + \frac{ip_y}{\hbar}e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar}\hat{y} + \frac{ip_z}{\hbar}e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar}\hat{z} \tag{5.7.28}$$

$$= \frac{i\mathbf{p}}{\hbar}e^{i(\mathbf{p}\cdot\mathbf{r})/\hbar}\hat{x} \tag{5.7.29}$$

Finding the probability density is much less tedious, and we find:

$$\mathbb{P}(\mathbf{r},t) = \psi^*\psi = \frac{1}{(2\pi\hbar)^3} \tag{5.7.30}$$

Notice that then:

$$\mathbf{J} = \mathbb{P}\cdot\frac{\mathbf{p}}{m} \tag{5.7.31}$$

which resembles the equation $\mathbf{J} = \rho\mathbf{v}$ in electrodynamics. We also find that:

$$\frac{\partial\mathbb{P}}{\partial t} = 0 = -\nabla\cdot\mathbf{J} \implies \nabla\cdot\mathbf{p} = 0 \tag{5.7.32}$$

so the momentum is divergenceless. ◀

# The Heisenberg Uncertainty Principle and Ehrenfest Theorem

6

## 6.1 Uncertainty relations

In classical mechanics, the state of a particle could be completely specified by a set of two variables $x_0, p_0$. Given these two observations one could know everything about the behaviour of the system in the future.

Quantum mechanics completely detaches itself from this picture, and states that the most one can know about a system is given by the state vector $|\psi\rangle$. One can therefore only find the probabilities of the outcomes of a measurement.

For an observable $\Omega$ recall that the expectation value $\langle\Omega\rangle$ is defined as:

$$\langle\Omega\rangle = \left\langle \psi \middle| \hat{\Omega} \middle| \psi \right\rangle \tag{6.1.1}$$

and the corresponding uncertainty about it is:

$$\Delta\Omega = \sqrt{\left\langle \psi \middle| (\hat{\Omega} - \left\langle \hat{\Omega} \right\rangle)^2 \middle| \psi \right\rangle} \tag{6.1.2}$$

Let $\hat{\Omega}$ and $\hat{\Lambda}$ be two Hermitian operators such that $[\hat{\Omega}, \hat{\Lambda}] = i\hat{\Gamma}$ where $\Gamma$ is Hermitian. Then:

$$(\Delta\Omega)^2(\Delta\hat{\Omega})^2 = \left\langle \psi \middle| (\hat{\Omega} - \left\langle \hat{\Omega} \right\rangle)^2 \middle| \psi \right\rangle \left\langle \psi \middle| (\hat{\Lambda} - \langle\Lambda\rangle)^2 \middle| \psi \right\rangle \tag{6.1.3}$$

Now let us define:

$$\Omega' = \Omega - \left\langle \hat{\Omega} \right\rangle \tag{6.1.4}$$

$$\Lambda' = \Lambda - \left\langle \hat{\Lambda} \right\rangle \tag{6.1.5}$$

where we clearly see that $[\hat{\Omega}, \hat{\Lambda}] = [\hat{\Omega}', \hat{\Lambda}']$. Now since $\hat{\Omega}'^\dagger = \hat{\Omega}^\dagger - \left\langle \hat{\Omega} \right\rangle^* = \hat{\Omega} - \left\langle \hat{\Omega} \right\rangle = \hat{\Omega}'$ we can conclude that $\hat{\Omega}'$ and $\hat{\Lambda}'$ are Hermitian. Therefore $\hat{\Omega}'^2 = \hat{\Omega}^\dagger\hat{\Omega}^\dagger$ and thus:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 = \left\langle \psi \middle| \hat{\Omega}'^2 \middle| \psi \right\rangle \left\langle \psi \middle| \hat{\Lambda}'^2 \middle| \psi \right\rangle = \left\langle \hat{\Omega}'\psi \middle| \hat{\Omega}'\psi \right\rangle \left\langle \hat{\Lambda}'\psi \middle| \hat{\Lambda}'\psi \right\rangle \tag{6.1.6}$$

We can now apply the Cauchy-Schwartz inequality $|V_1|^2|V_2|^2 \geq |\langle V_1 | V_2 \rangle|^2$ to $|\hat{\Omega}\psi\rangle$ and

$|\hat{\Lambda}\psi\rangle$ to find:

$$|\,|\hat{\Omega}'\psi\rangle\,|^2|\,|\hat{\Lambda}'\psi\rangle\,|^2 = \langle\Omega'\psi\,|\,\Omega'\psi\rangle\,\langle\Lambda'\psi\,|\,\Lambda'\psi\rangle = (\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq |\left\langle\hat{\Omega}'\psi\,\middle|\,\hat{\Lambda}'\psi\right\rangle|^2 \tag{6.1.7}$$

and since $\left\langle\hat{\Omega}'\psi\,\middle|\,\hat{\Lambda}'\psi\right\rangle = \left\langle\psi\,\middle|\,\hat{\Omega}'\hat{\Lambda}'\,\middle|\,\psi\right\rangle$ we find:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq \left\langle\psi\,\middle|\,\hat{\Omega}'\hat{\Lambda}'\,\middle|\,\psi\right\rangle \tag{6.1.8}$$

Now we can define the anti-commutator $\{\hat{A},\hat{B}\} = \hat{A}\hat{B} + \hat{A}\hat{B}$ and write:

$$\hat{\Omega}\hat{\Lambda} = \frac{1}{2}([\hat{\Omega}',\hat{\Lambda}'] + \{\hat{\Omega}',\hat{\Lambda}\}) \tag{6.1.9}$$

and so:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq |\left\langle\psi\,\middle|\,\frac{1}{2}([\hat{\Omega}',\hat{\Lambda}'] + \{\hat{\Omega}',\hat{\Lambda}\})\,\middle|\,\psi\right\rangle|^2 \tag{6.1.10}$$

which is proportional to the square norm of the sum of the expectation value of the commutator and the expectation value of the anti-commutator.

Since $[\hat{\Omega}',\hat{\Lambda}'] = i\Gamma$, the expectation value of the commutator is purely imaginary. Similarly, $\{\hat{\Omega}',\hat{\Lambda}'\}$ is Hermitian, and therefore real, hence the expectation value of the anti-commutator is also real.

Therefore, we have that:

$$|\left\langle\psi\,\middle|\,\frac{1}{2}([\hat{\Omega}',\hat{\Lambda}'] + \{\hat{\Omega}',\hat{\Lambda}\})\,\middle|\,\psi\right\rangle|^2 = \frac{1}{4}\left\langle\psi\,\middle|\,\{\hat{\Omega}',\hat{\Lambda}'\}\,\middle|\,\psi\right\rangle^2 + \frac{1}{4}\langle\psi\,|\,\Gamma\,|\,\psi\rangle^2 \tag{6.1.11}$$

therefore:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq \frac{1}{4}\left\langle\psi\,\middle|\,\{\hat{\Omega}',\hat{\Lambda}'\}\,\middle|\,\psi\right\rangle^2 + \frac{1}{4}\langle\psi\,|\,\Gamma\,|\,\psi\rangle^2 \tag{6.1.12}$$

Finally, in the case of **canonically conjugate operators**, $\Gamma = \hbar$:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq \frac{1}{4}\left\langle\psi\,\middle|\,\{\hat{\Omega}',\hat{\Lambda}'\}\,\middle|\,\psi\right\rangle^2 + \frac{\hbar^2}{4} \tag{6.1.13}$$

and finally since the first term is positive:

$$\Delta\hat{\Omega}'\cdot\Delta\hat{\Lambda}' \geq \frac{\hbar}{2} \tag{6.1.14}$$

**Uncertainty Relations**

For two Hermitian operators $\Omega$ and $\Lambda$ such that $[\Omega,\Lambda] = i\Gamma$ then:

$$(\Delta\hat{\Omega}')^2(\Delta\hat{\Lambda}')^2 \geq \frac{1}{4}\left\langle\psi\,\middle|\,\{\hat{\Omega}',\hat{\Lambda}'\}\,\middle|\,\psi\right\rangle^2 + \frac{1}{4}\langle\psi\,|\,\Gamma\,|\,\psi\rangle^2 \tag{6.1.15}$$

where $\Omega' = \Omega - \left\langle\hat{\Omega}\right\rangle$ and $\Lambda' = \Lambda - \left\langle\hat{\Lambda}\right\rangle$.

Note that the conditions for equality in 5.1.13 are:

$$\hat{\Omega}\left|\psi\right\rangle = c\Lambda\left|\psi\right\rangle \tag{6.1.16}$$

$$\left\langle\psi\right|\left\{\hat{\Omega}', \hat{\Lambda}'\right\}\left|\psi\right\rangle = 0 \tag{6.1.17}$$

where the first comes from the Cauchy-Schwartz inequality.

## 6.2 Minimum Uncertainty Packet

Suppose we want a packet with minimum uncertainty in $\hat{x}$ and $\hat{p}$. Then we need a state such that:

$$(\hat{p} - \langle\hat{p}\rangle)\left|\psi\right\rangle = c(\hat{x} - \langle\hat{x}\rangle)\left|\psi\right\rangle \tag{6.2.1}$$

and

$$\langle\psi\left|\right.(\hat{p} - \langle\hat{p}\rangle)(\hat{x} - \langle\hat{x}\rangle) + (\hat{x} - \langle\hat{x}\rangle)(\hat{p} - \langle\hat{p}\rangle)\left.\right|\psi\rangle = 0 \tag{6.2.2}$$

Solving the latter equation in the position basis:

$$\frac{d\psi}{dx} = \frac{i}{\hbar}(\langle\hat{p}\rangle + c(\hat{x} - \cancelto{0}{\langle\hat{x}\rangle}))\psi \tag{6.2.3}$$

where we can always translate our origin to $\langle\hat{x}\rangle$ so that in the new frame $\langle\hat{x}\rangle = 0$. Hence:

$$\psi = \psi(0)e^{i\langle p\rangle x/\hbar}e^{icx^2/2\hbar} \tag{6.2.4}$$

Now note that:

$$\langle\psi\left|\right.(\hat{p} - \langle\hat{p}\rangle)\hat{x}) + \hat{x}(\hat{p} - \langle\hat{p}\rangle)\left.\right|\psi\rangle = 0 \tag{6.2.5}$$

and since $\hat{p} - \langle\hat{p}\rangle = c\hat{x}$ we find:

$$\left\langle\psi\right|c^*\hat{x}^2 + c\hat{x}^2\left|\psi\right\rangle = 0 \implies (c + c^*)\left\langle\hat{x}^2\right\rangle = 0 \tag{6.2.6}$$

It follows immediately that $c$ is purely imaginary, and thus:

$$\psi = \psi(0)e^{i\langle p\rangle x/\hbar}e^{-|c|x^2/2\hbar} \tag{6.2.7}$$

which is an arbitrary Gaussian packet.

**Minimum uncertainty packet**

The family wave-function satisfying the equality in Heisenberg's uncertainty relation is:

$$\psi(x) = \psi(0)e^{i\langle p\rangle x/\hbar}e^{-|c|x^2/2\hbar} \tag{6.2.8}$$

## 6.3 Ground state energy of a hydrogen atom

The Hamiltonian for a hydrogen atom, where the proton is assumed to be stationary, whose only interaction with the electron is through a Coulomb potential, is:

$$\hat{H} = \frac{\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2}{2m} - \frac{e^2}{(\hat{x}^2 + \hat{y}^2 + \hat{z}^2)^{\frac{1}{2}}} \tag{6.3.1}$$

so that:

$$\left\langle \hat{H} \right\rangle = \frac{\langle \hat{p}_x^2 \rangle + \left\langle \hat{p}_y^2 \right\rangle + \langle \hat{p}_z^2 \rangle}{2m} - e^2 \left\langle \frac{1}{(\hat{x}^2 + \hat{y}^2 + \hat{z}^2)^{\frac{1}{2}}} \right\rangle \tag{6.3.2}$$

If we choose $\langle \hat{p}_i \rangle = 0$ then $\left\langle \hat{p}_i{}^2 \right\rangle = (\Delta p_i)^2 + \langle \hat{p}_i \rangle^2$ and so:

$$\left\langle \hat{H} \right\rangle = \frac{(\Delta \hat{p}_x)^2 + (\Delta \hat{p}_y)^2 + (\Delta \hat{p}_z)^2}{2m} - e^2 \left\langle \frac{1}{(\hat{x}^2 + \hat{y}^2 + \hat{z}^2)^{\frac{1}{2}}} \right\rangle \tag{6.3.3}$$

We can then argue that:

$$\left\langle \frac{1}{(\hat{x}^2 + \hat{y}^2 + \hat{z}^2)^{\frac{1}{2}}} \right\rangle \approx \frac{1}{\left\langle (\hat{x}^2 + \hat{y}^2 + \hat{z}^2)^{\frac{1}{2}} \right\rangle} \approx \frac{1}{(\langle \hat{x}^2 \rangle + \langle \hat{y}^2 \rangle + \langle z \rangle^2)^{\frac{1}{2}}} \tag{6.3.4}$$

For the ground state, we want to minimize $\left\langle \hat{H} \right\rangle$ so we need $\langle \hat{x}^2 \rangle = (\Delta \hat{x})^2 + \langle \hat{x} \rangle^2$ to be minimized. In other words, we need $\langle \hat{x} \rangle^2 = 0$. Then:

$$\left\langle \hat{H} \right\rangle \approx \frac{(\Delta \hat{p}_x)^2 + (\Delta \hat{p}_y)^2 + (\Delta \hat{p}_z)^2}{2m} - e^2 \frac{1}{((\Delta x)^2 + (\Delta \hat{y})^2 + (\Delta \hat{z})^2)^{\frac{1}{2}}} \tag{6.3.5}$$

Now due to the spherical symmetry, $\Delta \hat{x} = \Delta \hat{y} = \Delta \hat{z}$ and $\Delta \hat{p}_x = \Delta \hat{p}_y = \Delta \hat{p}_z$ and consequently:

$$\left\langle \hat{H} \right\rangle \approx \frac{3(\Delta \hat{p}_x)^2}{2m} - \frac{e^2}{\sqrt{3}\Delta \hat{x}} \tag{6.3.6}$$

We now use $\Delta \hat{x} \cdot \Delta \hat{p}_x \geq \frac{\hbar}{2}$ to find that:

$$\left\langle \hat{H} \right\rangle \geq \frac{3\hbar^2}{8m(\Delta \hat{x})^2} - \frac{e^2}{\sqrt{3}\Delta \hat{x}} \tag{6.3.7}$$

We differentiate with respect to $\Delta \hat{x}$ to find the minimum of $\left\langle \hat{H} \right\rangle$:

$$\frac{-6\hbar^2}{8m(\Delta \hat{x})^3} + \frac{e^2}{\sqrt{3}(\Delta \hat{x})^2} \tag{6.3.8}$$

so that $\Delta \hat{x} \approx 1.3 \frac{\hbar^2}{me^2}$. From here we find that:

$$\left\langle \hat{H} \right\rangle \geq -\frac{2me^4}{9\hbar^2} \tag{6.3.9}$$

It turns out that the real result is about twice as big:

$$E_g = -\frac{me^4}{2\hbar^2} \tag{6.3.10}$$

which is not too far off.

**Example.**
We wish to check that step (5.3.4) are justifiable. Consider the wave-function for the ground state of hydrogen (check normalization):

$$\psi(r,\theta,\phi) = \frac{e^{-r/a_0}}{\sqrt{\pi a_0^3}} \tag{6.3.11}$$

and prove that $\Delta x = \frac{\hbar^2}{me^2}$, and that $\frac{1}{\langle\sqrt{\hat{r}}\rangle} = \left\langle\frac{1}{\hat{r}}\right\rangle$

**Solution** We firstly check that the normalization condition is satisfied:

$$\begin{aligned}
\|\psi\|^2 &= \langle\psi\,|\,\psi\rangle && (6.3.12)\\
&= \int_{\mathbb{R}^3} \frac{1}{\pi a_0^3} e^{-2r/a_0} d\tau && (6.3.13)\\
&= \int_0^\infty \int_0^{2\pi} \int_0^\pi \frac{e^{-2r/a_0}}{\pi a_0^3} d\theta d\phi dr && (6.3.14)\\
&= \frac{4}{a_0^3} \int_0^\infty r^2 e^{-2r/a_0} dr = 1 && (6.3.15)
\end{aligned}$$

where we used the property of the gamma function:

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}} \tag{6.3.16}$$

so the wave-function is indeed normalized.

Now let us exploit the spherical symmetry of the system (ground state orbital is a sphere) to write that:

$$\left\langle\hat{x}^2\right\rangle = \left\langle\hat{y}^2\right\rangle = \left\langle\hat{z}^2\right\rangle \tag{6.3.17}$$

and thus:

◀

## 6.4 Ground state of a harmonic oscillator

Suppose we wish to find the form of the zero-energy state, that is, a state with energy $\hbar\omega$ which is as close as possible to the classical state $x = p = 0$. It must therefore saturate the uncertainty principle $\Delta x \Delta p = \frac{\hbar}{2}$.

We need to find a $\psi$ that minimizes $\langle H \rangle$, that is:

$$\langle H \rangle = \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle = \frac{\langle \hat{p}^2 \rangle}{2m} + \frac{1}{2} m\omega^2 \left\langle \hat{x}^2 \right\rangle \tag{6.4.1}$$

$$= \frac{(\Delta p)^2 + \langle p \rangle^2}{2m} + \frac{1}{2} m\omega^2 [(\Delta x)^2 + \langle x \rangle^2] \tag{6.4.2}$$

We can always set $\langle \hat{x} \rangle = 0$ through a change of reference frame. We will also restrict ourselves to states with $\langle \hat{p} \rangle = 0$ to further minimize $\langle H \rangle$. Then:

$$\langle H \rangle = \frac{(\Delta p)^2}{2m} + \frac{1}{2} m\omega^2 (\Delta x)^2 \tag{6.4.3}$$

and using $\Delta x \cdot \Delta p = \frac{\hbar}{2}$ we find:

$$\langle H \rangle = \frac{\hbar^2}{8m(\Delta x)^2} + \frac{1}{2} m\omega^2 (\Delta x)^2 \tag{6.4.4}$$

We can now minimize with respect to $(\Delta x)^2$:

$$\frac{\partial \langle H \rangle}{\partial (\Delta x)^2} = -\frac{\hbar - 2}{8m(\Delta x)^4} + \frac{1}{2} m\omega^2 = 0 \implies (\Delta x)^2 = \frac{\hbar}{2m\omega} \tag{6.4.5}$$

so that:

$$\langle H \rangle_{min} = \frac{\hbar}{\omega} 2 \tag{6.4.6}$$

This is the lowest mean energy an oscillator can have.

## 6.5 The Classical Limit

Of course, our quantum theory would not be acceptable if it didn't reduce to the classical framework in the domain of macroscopic systems.

Consider an operator $\hat{\Omega}$ with no explicit time-dependence:

$$\frac{d}{dt} \left\langle \hat{\Omega} \right\rangle = \frac{d}{dt} \left\langle \psi \middle| \hat{\Omega} \middle| \psi \right\rangle = \left\langle \dot{\psi} \middle| \hat{\Omega} \middle| \psi \right\rangle + \left\langle \psi \middle| \hat{\Omega} \middle| \dot{\psi} \right\rangle \tag{6.5.1}$$

Now we may use the Schrödinger equation:

$$|\dot{\psi}\rangle = -\frac{i}{\hbar} \hat{H} |\psi\rangle \tag{6.5.2}$$

$$\langle \dot{\psi}| = \frac{i}{\hbar} \hat{H} \langle \psi| \tag{6.5.3}$$

and find that:

$$\frac{d}{dt} \left\langle \hat{\Omega} \right\rangle = -\frac{i}{\hbar} ( \left\langle \psi \middle| \hat{\Omega} \hat{H} \middle| \psi \right\rangle - \left\langle \psi \middle| \hat{H} \hat{\Omega} \middle| \psi \right\rangle ) = -\frac{i}{\hbar} \left\langle \psi \middle| [\hat{\Omega}, \hat{H}] \middle| \psi \right\rangle \tag{6.5.4}$$

known as **Ehrenfest's theorem**.

**Ehrenfest's theorem**

The time-evolution of the expected value of an operator $\Omega$ with no explicit time dependence is:

$$\frac{d}{dt}\left\langle\hat{\Omega}\right\rangle = -\frac{i}{\hbar}\left\langle[\hat{\Omega},\hat{H}]\right\rangle \tag{6.5.5}$$

Recall that in classical Hamiltonian mechanics:

$$\frac{d\omega}{dt} = \{\omega,\mathcal{H}\} \tag{6.5.6}$$

is the time evolution of some dynamical variable $\omega$. The structural similarity with Ehrenfest's theorem is remarkable, an encouraging sign that our present quantum theory works.

Let us apply Ehrenfest's theorem to the position operator:

$$\frac{d}{dt}\langle\hat{x}\rangle = -\frac{i}{\hbar}\left\langle[\hat{x},\hat{H}]\right\rangle \tag{6.5.7}$$

If we assume that the Hamiltonian takes the form:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \tag{6.5.8}$$

then:

$$[\hat{x},\hat{H}] = \frac{1}{2m}[\hat{x},\hat{p}^2] \tag{6.5.9}$$

$$= \frac{1}{2m}([\hat{x},\hat{p}]\hat{p} + \hat{p}[\hat{x},\hat{p}]) \tag{6.5.10}$$

$$= \frac{2i\hbar}{2m}\hat{p} = \frac{i\hbar}{m}\hat{p} \tag{6.5.11}$$

so that:

$$\frac{d}{dt}\langle\hat{x}\rangle = \frac{\langle\hat{p}\rangle}{m} \tag{6.5.12}$$

This greatly resembles the relation $\frac{dx}{dt} = \frac{p}{m}$ in classical mehcanics. The only difference is that the variables $p$, $x$ have been replaced by the expectation values of their associated operators.

Let us repeat this process for $\hat{p}$.

$$\frac{d}{dt}\langle\hat{p}\rangle = -\frac{i}{\hbar}\langle[\hat{p},V(\hat{x})]\rangle \tag{6.5.13}$$

In the position basis:

$$[\hat{p},V(\hat{x})]\psi(x) = -i\hbar\frac{d}{dx}(V(x)\psi(x)) + i\hbar V(x)\frac{d\psi(x)}{dx} \tag{6.5.14}$$

$$= -i\hbar\psi(x)\frac{dV(x)}{dx} \tag{6.5.15}$$

which implies that:

$$[\hat{p}, V(\hat{x})] = -i\hbar \frac{dV(x)}{dx} \tag{6.5.16}$$

and thus:

$$\frac{d}{dt}\langle p \rangle = \left\langle \frac{dV(x)}{dx} \right\rangle = -\left\langle \frac{\partial H}{\partial x} \right\rangle \tag{6.5.17}$$

Thus, we have seen that in the classical limit, we regain Hamilton's equations:

$$\dot{x} \to \frac{d}{dt}\langle \hat{x} \rangle = \left\langle \frac{\partial H}{\partial p} \right\rangle \xrightarrow{\text{class. lim.}} \frac{\partial \mathcal{H}}{\partial p} \tag{6.5.18}$$

$$\dot{p} \to \frac{d}{dt}\langle \hat{p} \rangle = -\left\langle \frac{\partial H}{\partial x} \right\rangle \xrightarrow{\text{class. lim.}} -\frac{\partial \mathcal{H}}{\partial x} \tag{6.5.19}$$

# The Harmonic Oscillator

7

## 7.1 Quantization of the Oscillator in position representation

We consider a quantum oscillator, that is, a system whose Hamiltonian is of the form:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{7.1.1}$$

To solve the problem of the quantum oscillator, we must find the propagator $\hat{U}(t)$ by solving the TISE. We will do so in two ways, the first in the position representation and the second in the energy representation.

Firstly note that any energy eigenvalue must be positive, since:

$$\langle H\rangle = \frac{1}{2m}\left\langle \psi \middle| \hat{p}^2 \middle| \psi \right\rangle + \frac{1}{2}m\omega^2 \left\langle \psi \middle| \hat{x}^2 \middle| \psi \right\rangle \tag{7.1.2}$$

$$= \frac{1}{2m}\langle \hat{p}\psi \,|\, \hat{p}\psi\rangle + \frac{1}{2}m\omega^2 \langle \hat{x}\psi \,|\, \hat{x}\psi\rangle \geq 0 \tag{7.1.3}$$

since the norms of $|\hat{p}\psi\rangle$ and $|\hat{x}^2\psi\rangle$ must be positive. Therefore, setting $|\psi\rangle = |E\rangle$, for all energy eigenstates $|E\rangle$ we have that:

$$\langle H\rangle = E \geq 0 \tag{7.1.4}$$

as required.

We can now tackle the TISE equation for the oscillator:

$$\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\right)|E\rangle = E\,|E\rangle \tag{7.1.5}$$

which in the position representation becomes:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2\right)\psi = E\psi \tag{7.1.6}$$

The equation can be rearranged into the form $\frac{d^2\psi}{dx^2} + \frac{2mV}{\hbar^2}\psi = 0$:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}\left(E - \frac{1}{2}m\omega^2x^2\right)\psi = 0 \tag{7.1.7}$$

We begin by nondimensionalizing the equation, that is, to perform substitutions so that all quantities become dimensionless.

The substitutions we shall use are $x = by$, with $b = \sqrt{\hbar m\omega}$ and $\epsilon = \frac{E}{\hbar\omega}$ for reasons that will become clear soon.

Eq 6.1.7 then turns into:

$$\frac{d^2\psi}{dy^2} + (2\epsilon - y^2)\psi = 0 \tag{7.1.8}$$

In the limit $y \to \infty$, we get $\psi'' - y^{"}\psi = 0$ whose solution in the same limit is[1]:

$$\psi = Ay^m e^{\pm y^2/2} \tag{7.1.9}$$

but we can discard the positive exponential solution since it blows up and does not belong to the physical Hilbert space we are interested in.

In the limit as $y \to 0$ we get $\psi'' + 2\epsilon\psi = 0$ which has solution in the same limit as:

$$\psi = A\cos\left(\sqrt{2\epsilon}y\right) + B\sin\left(\sqrt{2\epsilon}y\right) \longrightarrow A + \frac{B}{\sqrt{2\epsilon}}y + \mathcal{O}(y^2) \tag{7.1.10}$$

So we can introduce the ansatz:

$$\psi(y) = u(y)e^{-y^2/2} \tag{7.1.11}$$

where $u(y)$ approaches $A + \frac{B}{\sqrt{2\epsilon}}y$ for small $y$ and $y^m$ for large $y$. We then get:

$$\frac{d^2\psi}{dy^2} = u''e^{-y^2/2} - 2yu'e^{-y^2/2} + (y^2 - 1)ue^{-y^2/2} = -(2\epsilon - y^2)ue^{-y^2/2} \tag{7.1.12}$$

which implies:

$$u'' - 2yu' + (2\epsilon - 1)u = 0 \tag{7.1.13}$$

From the mathematical methods volume, we can clearly see that this differential equation could be a candidate for a power series solution. Indeed, if we set:

$$u(y) = \sum_{n=0}^{\infty} c_n y^n \tag{7.1.14}$$

[1] the second derivative can be written as $\psi'' = Ay^{m+2}e^{y^2/2}\left(1 + \mathcal{O}(\frac{1}{y^2})\right) \approx Ay^{m+2}e^{y^2/2} = y^2\psi$

into eq. 6.1.13 we find that:

$$\sum_{n=1}^{\infty} C_n \big[ n(n-1)y^{n-2} - 2ny^n + (2\epsilon - 1)y^n \big] = 0 \tag{7.1.15}$$

Note that the first term can be rewritten as:

$$\sum_{n=2}^{\infty} C_n n(n-1)y^{n-2} = \sum_{n=0}^{\infty} C_{n+2}(n+2)(n+1)y^n \tag{7.1.16}$$

so that:

$$\sum_{n=0}^{\infty} y^n \big[ C_{n+2}(n+2)(n+1) + C_n(2\epsilon - 2n - 1) \big] = 0 \tag{7.1.17}$$

Therefore, since the above expression must vanish identically for all $y$, the coefficient in brackets must be zero for all $n$:

$$C_{n+1} = C_n \frac{(2n + 1 - 2\epsilon}{(n+2)(n+1)} \tag{7.1.18}$$

We can however immediately see a problem, there has been no restriction set on $\epsilon$, despite the fact that the energy eigenvalues must be positive. Moreover, $u(y)$ does not behave asymptotically as $y^m$, but has an infinite number of powers. This is an immediate consequence of the *infinite* power expansion that we used.

The only way to solve this problem is to make sure that $C_{m+2} = 0$ so that all subsequent coefficients are also null. This can only happen if $\epsilon$ has one of the special values:

$$\epsilon_n = \left( n + \frac{1}{2} \right), \;\; n = 0, 1, 2... \tag{7.1.19}$$

with corresponding energy quantization:

$$E_n = \left( n + \frac{1}{2} \right) \hbar\omega \tag{7.1.20}$$

For each $n$m we have a corresponding Hermite polynomial:

$$H_{n+1}(y) = 2yH_n - 2nH_{n-1}, \; H_0(y) = 1, \; H_1(y) = 2y \tag{7.1.21}$$

so that $u(y) = H_n(y)$. We can finally substitute back the initial variables $x, m, \omega$ to find the final normalized solution to the TISE:

$$\psi(x) = \left( \frac{m\omega}{\pi\hbar 2^{2n}(n!)^2} \right)^{1/4} \exp\left( -\frac{m\omega x^2}{2\hbar} \right) H_n\left( \sqrt{\frac{m\omega}{\hbar}} x \right) \tag{7.1.22}$$

## 7.2 Quantization of the oscillator in energy representation

Let us introduce the operator :

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{1}{\sqrt{2m\omega\hbar}}\hat{p} \tag{7.2.1}$$

and its conjugate $a^\dagger$ called the **ladder operators** introduced by the great P.A.M. Dirac. Notice that

$$a^\dagger a = \frac{m\omega}{2\hbar}\hat{x}^2 + \frac{1}{2m\omega\hbar}\hat{p}^2 + \frac{i}{2\hbar}[\hat{x},\hat{p}] = \frac{\hat{H}}{\hbar\omega} - \frac{1}{2} \tag{7.2.2}$$

so that:

$$\hat{H} = (a^\dagger a + \frac{1}{2})\hbar\omega \tag{7.2.3}$$

This is a factorization of the Hamiltonian $\hat{H} \sim (\hat{x}^2 + \hat{p}^2)$ as the product of $a \sim (\hat{x} + i\hat{p})$ and $a^\dagger \sim (\hat{x} - i\hat{p})$.

We may therefore derive the following commutation relation:

$$[a, a^\dagger] = aa^\dagger - a^\dagger a = 1 \tag{7.2.4}$$

If we define $\hat{H}' = (a^\dagger a + \frac{1}{2})$ (again to nondimensionalize) then we wish to solve the eigenvalue equation:

$$\hat{H}' \left|\epsilon\right\rangle = \epsilon \left|\epsilon\right\rangle \tag{7.2.5}$$

Note that:

$$\hat{H}' a \left|\epsilon\right\rangle = (a\hat{H}' - [a, \hat{H}']) \left|\epsilon\right\rangle \tag{7.2.6}$$

and since:

$$[a, \hat{H}'] = [a, a^\dagger + \frac{1}{2}] = [a, a^\dagger a] = [a, a^\dagger]a = a \tag{7.2.7}$$

we find that:

$$\hat{H}' a \left|\epsilon\right\rangle = (a\hat{H}' - a) \left|\epsilon\right\rangle = (\epsilon - 1)a \left|e\right\rangle \tag{7.2.8}$$

Therefore $a \left|\epsilon\right\rangle$ is an energy eigenstate with eigenvalue $\epsilon - 1$. Similarly, $a^\dagger \left|\epsilon\right\rangle$ is an energy eigenstate with eigenvalue $\epsilon + 1$. To summarize:

$$a \left|\epsilon\right\rangle = c_\epsilon \left|\epsilon - 1\right\rangle \tag{7.2.9}$$

$$a^\dagger \left|\epsilon\right\rangle = c_\epsilon \left|\epsilon + 1\right\rangle \tag{7.2.10}$$

Since the two operators lower and raise the energy $\epsilon$, they are referred to as lowering and raising ladder operators.

We now use the fact that the energy eigenvalues must be positive to find the ground state $\left|\epsilon_0\right\rangle$ which can no longer be lowered:

$$a \left|\epsilon_0\right\rangle = 0 \implies a^\dagger a \left|\epsilon_0\right\rangle \tag{7.2.11}$$

Using the definition of $a^\dagger a = \hat{H}' - \frac{1}{2}$ then:

$$\hat{H}' \left|\epsilon_0\right\rangle = \frac{1}{2} \left|\epsilon_0\right\rangle \implies \epsilon_0 = \frac{1}{2} \tag{7.2.12}$$

Again, we find the same energy quantization as before:

$$E_n = (n + \frac{1}{2})\hbar\omega,\ n = 0, 1, 2... \tag{7.2.13}$$

There cannot be other energy levels, since there is no degeneracy in one-dimension. Therefore, the ground states must coincide, and consequently subsequent operation of ladder operators shows that the above quantization is the only set of energy levels.

Since $\epsilon$ are solely dependent on $n$, we will denote them as $|n\rangle$:

$$a \left|n\right\rangle = c_n \left|n-1\right\rangle \tag{7.2.14}$$

$$\implies \left\langle n \middle| a^\dagger a \middle| n \right\rangle = |n-1|n-1\rangle\, c_n^* c_n \tag{7.2.15}$$

$$\implies \left\langle n \middle| \hat{H}' - \frac{1}{2} \middle| n \right\rangle = |c_n|^2 \tag{7.2.16}$$

$$\implies \left\langle n \middle| n \right\rangle = |c_n|^2 \tag{7.2.17}$$

$$\implies c_n = \sqrt{n} e^{i\phi} \tag{7.2.18}$$

where the phase $\phi$ is arbitrary, and conventionally set to zero. So:

$$a \left|n\right\rangle = \sqrt{n} \left|n-1\right\rangle \implies \boxed{\left\langle n' \middle| a \middle| n \right\rangle = \sqrt{n}\delta_{n',n-1}} \tag{7.2.19}$$

Similarly:

$$a^\dagger \left|n\right\rangle = \sqrt{n+1} \left|n+1\right\rangle \implies \boxed{\left\langle n' \middle| a^\dagger \middle| n \right\rangle = \sqrt{n+1}\delta_{n',n+1}} \tag{7.2.20}$$

We can therefore write:

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) \tag{7.2.21a}$$

$$\hat{p} = i\sqrt{\frac{m\omega\hbar}{2}}(a^\dagger - a) \tag{7.2.21b}$$

and use 6.2.19 and 6.2.20 to find their representing matrices:

$$\hat{x} \leftrightarrow \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 1 & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7.2.22}$$

and similarly:

$$\hat{p} \leftrightarrow i\sqrt{\frac{m\omega\hbar}{2}} \begin{pmatrix} 0 & i & 0 & 0 & \dots \\ 1 & 0 & i\sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & i\sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7.2.23}$$

For example, suppose we want to calculate $\langle 4 \,|\, \hat{x}^2 \,|\, 2\rangle$. In the position representation one would have to calculate a very long and tedious integral. In the energy representation however:

$$\left\langle 4 \middle| \hat{x}^2 \middle| 2 \right\rangle = \frac{\hbar}{2m\omega} \left\langle 4 \middle| (a + a^\dagger)^2 \middle| 2 \right\rangle \tag{7.2.24}$$

$$= \frac{\hbar}{2m\omega} \left\langle 3 \middle| a^2 + (a^\dagger)^2 + aa^\dagger + a^\dagger a \middle| 2 \right\rangle \tag{7.2.25}$$

Now notice that the terms $aa^\dagger |2\rangle$ and $|a^\dagger a|2\rangle$ are both null, since the effect of $a^\dagger$ is to raise $|2\rangle \to |3\rangle$ and the effect of $a$ is to lower $|3\rangle \to |2\rangle$. So the only contributions are from:

$$a^2 |2\rangle = \sqrt{2}a |1\rangle = |0\rangle \tag{7.2.26}$$

$$(a^\dagger)^2 |2\rangle = \sqrt{3}a^\dagger |3\rangle = 2\sqrt{3} |4\rangle \tag{7.2.27}$$

Due to the orthonormality of energy eigenstates, $\langle 3 \,|\, 0\rangle = 0$ and thus:

$$\left\langle 4 \middle| \hat{x}^2 \middle| 2 \right\rangle = \frac{\sqrt{3}\hbar}{m\omega} \tag{7.2.28}$$

We also then see that we can express $|n\rangle$ by repeatedly operating $a^\dagger$ on $|0\rangle$:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle \tag{7.2.29}$$

Finally, there is a final operator that is often useful, especially when counting energy quanta of the oscillator. This operator, $\hat{N}$ is the **number operator**, and may be expressed as:

$$\hat{N} = a^\dagger a \implies \hat{H} = \hbar\omega(\hat{N} + \frac{1}{2}) \tag{7.2.30}$$

For example, let us find $\hat{N} |n\rangle$:

$$\hat{N} |n\rangle = a^\dagger a |n\rangle = \sqrt{n}a^\dagger |n-1\rangle = n |n\rangle \tag{7.2.31}$$

as expected.

## 7.3 Revised second postulate

We saw from the previous section that there is no need to express the $\hat{x}, \hat{p}$ operators in any basis. It suffices to use the commutation relation $[\hat{x}, \hat{p}] = i\hbar$, which is basis-independent.

This suggests a reformulation of the second postulate, which up until now has always been expressed in the position representation.

**Reformulated Postulate II**

The variables $x$, $p$ in classical mechanics become operators $\hat{x}$, $\hat{p}$ defined by the commutation relation $[\hat{x}, \hat{p}] = i\hbar$.

This is a much more general postulate. Indeed, going back to the $x$-basis, it allows us to write:

$$\hat{x} \to x, \ \hat{p} \to -i\hbar \frac{d}{dx} + f(x) \tag{7.3.1}$$

where we could add a factor $f(x)$ without altering the commutation relation. This may seem very unphysical, since surely we cannot obtain the same solutions if we add this arbitrary function. However, note that we cannot measure the wave function directly, so we are allowed to perform this trick. It turns out that what we can measure, that is, probabilities, squares of matrix elements and eigenvalue spectra are all invariant by such a translation in the Hilbert space.

## 7.4 Properties of Harmonic oscillator

### Energy levels are quantized

Why does the classical oscillator have a seemingly continuous energy spectrum? Consider for example a mass of $1kg$ oscillating at a frequency $1$ rad/s with amplitude $20cm$. Then it has energy:

$$E = 10^5 \text{ erg} \tag{7.4.1}$$

Instead, the gap between allowed energies is:

$$\Delta E \approx 10^{-27} \text{ erg} \tag{7.4.2}$$

We immediately see that the relative size of the energy gap is $\frac{\Delta E}{E} = 10^{-33}$, a minuscule quantity that at a macroscopic level is not observable.

### Energy levels uniformly spaced

The energy gap $\Delta E = \hbar\omega$ remains constant. We may therefore associate a fictitious quanta to oscillators of frequency $\omega$, each with energy $\hbar\omega$, called **phonons**. Therefore, by acting $a$ on the wave-function, we are destroying phonons, whereas acting $a^\dagger$ creates phonons.

### Ground state has non-zero energy

The lowest allowed energy for an oscillator is $\frac{\hbar\omega}{2}$, and not zero since it would imply that $x = p = 0$ exactly, thus violating Heisenberg's uncertainty relation.

### 7.4.1 Comparison with Classical oscillator

The classical oscillator has solution $x = x_0 \cos(\omega t + \phi)$, so that the its velocity is given by:

$$\dot{x} = \sqrt{\frac{2E}{m} + \omega^2 x^2} = \omega\sqrt{x_0^2 - x^2} \tag{7.4.3}$$

So, the classical turning point is given by $\pm x_0$, the position of the particle cannot exceed these values.

Although in the Newtonian framework, the particle has a definite position, we can still assign a probability distribution to it. Indeed, consider the following experiment: we walk inside a room with a classical oscillator. Obviously, we expect the points where the particle has lowest speed. So, the probability distribution is equal to the time the particle spends at each point divided by the period of oscillation thus:

$$P_{cl}(x) = \frac{\omega}{2 * \pi\dot{x}} = \frac{1}{2\pi\sqrt{x_0^2 - x^2}} \tag{7.4.4}$$

For a classical oscillator with energy $E_n = (n + \frac{1}{2})\hbar\omega$, we see that the classical turning point $x_0$ must satisfy:

$$\frac{1}{2}m\omega^2 x_0^2 = (n + \frac{1}{2})\hbar\omega \implies x_0 = \pm\sqrt{\frac{(2n+1)\hbar}{m\omega}} \tag{7.4.5}$$

No matter what, the classical oscillator cannot leap over this turning point. However, we see clearly in the quantum oscillator that this is not the case, since the wave-function extends well over this point.

Furthermore, in the classical case we readily see that the most probable position measurement of the system is at the turning points. However, in the quantized oscillator, it's most likely to be found at or near the center.

Finally, we know from the correspondence principle that as we increase $n$, the quantum and classical pictures should start to overlap. This is true, since for large $n$ the wave-function starts oscillating too rapidly, so that any measurement can really only detect the averaged value of the wave-function over appreciable intervals. This averaged out wave-function is close to $P_{cl}$ as expected.

## 7.5 Relating the energy and position representation

We have seen that the energy representation is extremely useful to evaluate matrix elements representing operators. However, the position representation is most useful when trying to evaluate probability amplitudes.

So is there a way to find $\psi_n(x)$ other than just calculating $\langle x \,|\, n\rangle$?

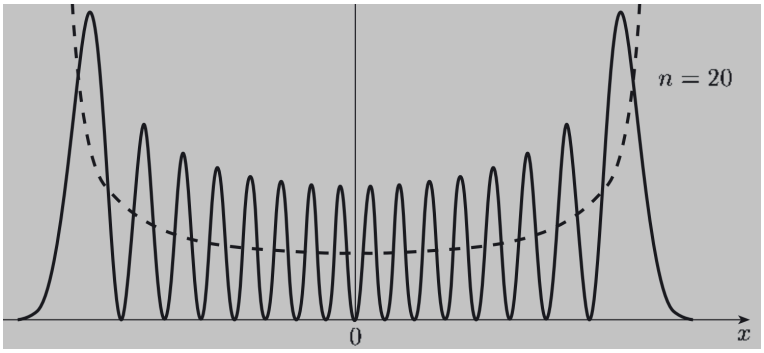



**Figure 7.1**. Plot of $|\psi_{20}|^2$ (solid line) and $P_{cl}$ (broken line).

Let us begin by projecting $a$ onto the $x$-basis:

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + \frac{i}{\sqrt{2m\omega}}\frac{d}{dx} = \frac{1}{\sqrt{2}}\left(y + \frac{d}{dy}\right) \tag{7.5.1}$$

as well as:

$$a^\dagger = \frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right) \tag{7.5.2}$$

So, in the $x$-basis $a\,|0\rangle = 0$ becomes:

$$\left(y + \frac{d}{dy}\right)\psi_0(y) = 0 \tag{7.5.3}$$

then we get that:

$$\psi(y) = A_0 e^{-y^2/2} \implies \psi(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left[-\frac{m\omega x^2}{2\hbar}\right] \tag{7.5.4}$$

upon normalization. We can also project the equation:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle \tag{7.5.5}$$

onto the $x$-basis to find that:

$$\langle x\,|\,n\rangle = \frac{1}{\sqrt{n!}}\left[\frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right)\right]^n \underbrace{\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-y^2/2}}_{\psi_0(x)} \tag{7.5.6}$$

Comparison with:

$$\psi(x) = \left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2}\right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \tag{7.5.7}$$

tells us that:

$$H_n(y) = e^{y^2/2}\left(y - \frac{d}{dy}\right)^n e^{-y^2/2} \tag{7.5.8}$$

## 7.6 Path integral derivation

The lagrangian for a harmonic oscillator may be expressed as:

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2 \tag{7.6.1}$$

Therefore, because it is of the form $\frac{1}{2}m\dot{x}^2 - a - bx - cx^2 - d\dot{x} - ex\dot{x}$, it is appropriate to ignore the actions due to the non-classical paths and write the propagator of the oscillator as:

$$U(x,t;x') = A(t)e^{iS_{cl}/\hbar} \tag{7.6.2}$$

Now:

$$S_{cl} = \int_0^t \mathcal{L}(x,\dot{x})dt = \int_0^t \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2 dt \tag{7.6.3}$$

The equation of motion for the classical harmonic oscillator is:

$$x_{cl}(t) = A\cos\omega t + B\sin\omega t \tag{7.6.4}$$
$$\dot{x}_{cl}(t) = \omega(-A\sin\omega t + B\cos\omega t) \tag{7.6.5}$$

So:

$$\frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\omega^2(A^2\sin^2\omega t + B^2\cos^2\omega t - AB\sin\omega t) \tag{7.6.6}$$
$$\frac{1}{2}m\omega^2 x^2 = \frac{1}{2}m\omega^2(A^2\cos^2\omega t + B^2\sin^2\omega t + AB\sin 2\omega t) \tag{7.6.7}$$

hence

$$\mathcal{L} = \frac{1}{2}m\omega^2\left[(B^2 - A^2)\cos 2\omega t - 2AB\sin 2\omega t\right] \tag{7.6.8}$$

Integrating from $0$ to $t$ we express the action as:

$$S_{cl} = \int_0^t \frac{1}{2}m\omega^2\left[(B^2 - A^2)\cos 2\omega t - 2AB\sin 2\omega t\right]dt \tag{7.6.9}$$
$$= \frac{1}{4}m\omega\Big((B^2 - A^2)\left[\sin 2\omega t\right]_0^t + 2AB\left[\cos 2\omega t\right]_0^t\Big) \tag{7.6.10}$$
$$= \frac{1}{4}m\omega(B^2 - A^2)\big(\sin 2\omega t - 4AB\sin^2\omega t\big) \tag{7.6.11}$$

where we used the half angle formula $\cos 2\omega t = 1 - 2\sin^2\omega t$.

Now we express $B$ and $A$ in terms of $x(0) = x'$ and $x(t) = x$:

$$A = x' \tag{7.6.12}$$

$$B = \frac{x - x' \cos \omega t}{\sin \omega t} \tag{7.6.13}$$

$$AB = x_0 \frac{x - x' \cos \omega t}{\sin \omega t} \tag{7.6.14}$$

and

$$B^2 - A^2 = \frac{x^2 - 2xx' \cos \omega t + x' \cos^2 \omega t + x' \cos \omega t - x'^2 \sin \omega t}{\sin^2 \omega t} \tag{7.6.15}$$

Finally we get after a hefty amount of algebraic simplification:

$$S_{cl} = \frac{1}{2} m \left( \omega \frac{(x^2 + x'^2) \cos \omega t - 2xx'}{\sin \omega t} \right) \tag{7.6.16}$$

so:

$$U(x, t; x') = A(t) \exp \left[ \frac{im\omega}{2\hbar \sin \omega t} \left( (x^2 + x'^2) \cos \omega t - 2xx' \right) \right] \tag{7.6.17}$$

which, up to a normalization factor of $A(t)$ corresponds with what we found previously.

# Scattering and tunnelling in 1D

8

## 8.1 Introduction

Scattering is of utmost importance in physics. The scattering of atoms off a gold foil is what led Rutherford to hypothesize the nuclear model of an atom. Furthermore, it is the scattering of light off surfaces that allows us to see,

Scattering in general is the phenomenon in which incident particles interact with some target, changing its speed, direction of motion etc...

The phenomena that we will discuss in this chapter however, which can be classified as tunneling phenomena, have no counterpart in classical physics. They involve the penetration of barriers, classically forbidden regions where the energy of the scattered particle is lower than the potential barrier it is trying to climb.

There are several applications of scattering, such as the development of STM (scanning electron microscopy) and nuclear processes.

## 8.2 Wave-packet approach (long method)

We consider potentials of the form:

$$V(x) = \begin{cases} 0, & \text{for } x < 0 \\ V_0, & \text{for } x > 0 \end{cases} \tag{8.2.1}$$

Imagine we shoot a particle with energy $E$ towards the barrier. Of course, one would expect that for $E > V_0$, the particle will simply hop over the barrier and proceed, whereas for $E < V_0$ the particle will bounce off the barrier and be reflected backwards.

This would be true if the particle were localised. However, we know from our previous discussions that the wave function for a particle extends out to infinity even in a finite potential well, and consequently there is a fraction of the wave-function that will extend over the barrier, representing the probability that the particle climbs the barrier.

How do we calculate this coefficient, the **reflection coefficient**, expressing the probability of the particle overcoming the potential barrier? We assume the particle initially has a wave-function:

$$\psi_i(x,0) = Ae^{ik_0(x+a)}e^{-(x+a)^2/2\Delta^2} \tag{8.2.2}$$

(i) Find the eigenfunction $\psi_E$ of the potential Hamiltonian

(ii) find the projection $a = \langle \psi_E \,|\, \psi_I \rangle$ where $\psi_I$ is the particle's wavefunction

(iii) express $\psi(x,t) = \int_{-\infty}^{\infty} a(E) e^{-iEt/\hbar} \psi_E(x) dE$

(iv) identify $\psi_R$ and $\psi_T$ and evaluate the reflection coefficient:

$$R = \int_{-\infty}^{\infty} |\psi_R|^2 \, dx \tag{8.2.3}$$

and the transmission coefficient:

$$T = \int_{-\infty}^{\infty} |\psi_T|^2 \, dx \tag{8.2.4}$$

**Step I**

We have already found the solutions of the eigenfunctions previously:

$$\psi_E(x) = \begin{cases} Ae^{ik_1 x} + Be^{-ik_1 x}, \; k_1 = \sqrt{\frac{2mE}{\hbar^2}} \\ Ce^{ik_2 x} + De^{-ik_2 x}, \; k_2 = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \end{cases} \tag{8.2.5}$$

We only considered eigenfunctions with $E > V_0$, since if $E < V_0$ the corresponding eigenfunctions will be shown to be orthogonal to $\psi_I$.

Since we only seek solutions with the transmitted wave moving to the right on top of the barrier, we set $D = 0$.

We now impose the continuity of $\psi$ and $\psi'$ at $x = 0$ to find:

$$\begin{cases} A + B = C \\ ik_1(A - B) = ik_2 C \end{cases} \implies \begin{cases} B = \frac{k_1 - k_2}{k_1 + k_2} A \\ C = \frac{2k_1}{k_1 + k_2} A \end{cases} \tag{8.2.6}$$

We can therefore use the Heaviside function:

$$\Theta(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x < 0 \end{cases} \tag{8.2.7}$$

to write:

$$\psi_{k_1}(x) = A\left[\left(e^{ik_1 x} + \frac{B}{A} e^{-ik_1 x}\right)\Theta(-x) + \frac{C}{A} e^{ik_2 x}\Theta(x)\right] \tag{8.2.8}$$

where we labelled the wavefunction with $k_1$ since there is a unique wave-number for each energy level $E$.

### Step 2

Consider:

$$a(k_1) = \langle \psi_{k_1} \,|\, \psi_I \rangle \tag{8.2.9}$$

$$= A\Bigg[ \int_{-\infty}^{\infty} \left( e^{-ik_1 x} + \frac{B^*}{A^*} e^{ik_1 x} \right) \Theta(-x) \psi_I(x) dx \tag{8.2.10}$$

$$+ \int_{-\infty}^{\infty} \frac{C^*}{A^*} e^{-ik_2 x} \Theta(x) \psi_I(x) dx \Bigg] \tag{8.2.11}$$

Now, since $\psi_I(x)$ is non-zero (to a good approximation) only for $x < 0$ and $\Theta(x)$ is non-vanishing only for $x > 0$, we can ignore the second integral.

Also, the second term in the first integral vanishes since $\psi_I(x)$ is orthogonal to $e^{ik_1 x}$. Hence:

$$a(k_1) = A \int_{-\infty}^{\infty} e^{-ik_1 x} \psi_I(x) dx = \left( \frac{\Delta^2}{\pi} \right) e^{-(k_1 - k_0)^3 \Delta^2 / 2} e^{ik_1 a} \tag{8.2.12}$$

is just the Fourier transform of $\psi_I$ (up to some constant, $A$, which we will show to be equal to $\frac{1}{\sqrt{2\pi}}$). Also, note that $a(k_1)$ has a peak at $k_1 = k_0$ for large $\Delta$.

### Step 3

We can express the wave-function for $t \geq 0$ as:

$$\psi(x,t) = \int_{-\infty}^{\infty} a(k_1) e^{-iE(k_1)t/\hbar} \psi_{k_1}(x) dk_1 \tag{8.2.13}$$

Then $\psi(x,t)$ evaluates to:

$$\psi(x,t) = \left( \frac{\Delta^2}{4\pi^3} \right)^{\frac{1}{4}} \int_{-\infty}^{\infty} \exp\left( \frac{-i\hbar k_1^2 t}{2m} \right) \exp\left[ -\frac{(k_1 - k_0)^2 \Delta^2}{2} \right] e^{ik_1 a} \tag{8.2.14}$$

$$\times \left[ e^{ik_1 x} \Theta(-x) + \frac{B}{A} e^{-ik_1 x} \Theta(-x) + \frac{C}{A} \exp\left[ i \underbrace{\sqrt{k_1^2 - \frac{2mV_0}{\hbar^2}}}_{k_2} x \right] \Theta(x) \right] dk_1 \tag{8.2.15}$$

Setting $t = 0$ then we find that $A = \frac{1}{\sqrt{2\pi}}$.

### Step 4

The first term can be expressed as $\Theta(-x) G(-a, k_0, t)$ where $G(x_0, k_0, t)$ is the Gaussian centered at $x + \frac{\hbar k_0 t}{m}$. As $t \longrightarrow \infty$, the gaussian is centered more and more at $x \to \frac{\hbar k_0 t}{m}$, whereas $\Theta(-x)$ vanishes at $x > 0$. So for large $t$ the first term, the original wave-packet, disappears as is expected.

The second term therefore represents the reflected wave function $\psi_R$ (since it is travelling with momentum $\hbar k_1$ to the left), and the third term represents the transmitted wave function $\psi_T$ (since it is travelling with momentum $\hbar k_2$ to the right). Hence, setting $k_1 = k_0$

since $a(k)$ has a peak there, and since we expect a completely elastic collision so that the energy of the reflected:

$$\psi_R = \Theta(-x)G(a, k_0, t)\left(\frac{B}{A}\right)_{k_1=k_0} \tag{8.2.16}$$

So:

$$R = \int_{-\infty}^{\infty} |\psi_R|^2 dx = \left(\frac{|B|}{|A|}\right)^2_{k_1=k_0} \tag{8.2.17}$$

and:

$$T = 1 - R \tag{8.2.18}$$

by the global conservation of probability.

## 8.3 Probability currents approach (shortcut)

Our final results for the coefficients only seem to depend on the relative sizes of $A$ and $B$ at $k_0 = k_1$. Therefore, would it be possible to reach the same expressions without resorting to the monstrosity that is (8.2.10) and (8.2.20)?

It turns out that we may use probability currents to relate the flux of particles transmitted through the step barrier and the flux of particles reflected back.

We can completely eliminate the complexity of the wave-packet approach by considering not an individual particle, but rather a beam of particles incident on the barrier. If we consider the beam as a time-independent steady current ($\nabla \cdot \mathbf{J} = 0$) so that there is no particle accumulation, then we can consider the entire system as a stationary state. There is no longer a need to use wave-packets, but rather a simply stationary state of energy equal to the average energy of the beam.

### Step potential with $E > V_0$

So we can consider the following eigenstate (assuming $k_0 = k_1$):

$$\psi(x) = A\left[\left(e^{ik_1 x} + \frac{B}{A}e^{-ik_1 x}\right)\Theta(-x) + \frac{C}{A}e^{ik_2 x}\Theta(x)\right] \tag{8.3.1}$$

Now the probability current of the incident beam is then:

$$J_{inc} = \frac{\hbar}{m}\mathrm{Im}(ik_1|A|^2|e^{ik_1 x}|^2) = \frac{\hbar k_1}{m}|A|^2 \tag{8.3.2}$$

Similarly, the probability current of the reflected and transmitted beams are respectively:

$$J_{ref} = \frac{\hbar k_1}{m}|B|^2,\ J_{tra} = \frac{\hbar k_2}{m}|C|^2 \tag{8.3.3}$$

(ignoring the negative term for the reflected current). Therefore:

$$R = \frac{J_{ref}}{J_{inc}} = \frac{|B|}{|A|} \tag{8.3.4}$$

$$T = \frac{J_{tra}}{J_{inc}} = \frac{k_2}{k_1}\frac{|C|}{|A|} \tag{8.3.5}$$

Finally, we can use the relations between $A, B, C$ in (8.2.6) to find that:

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 \tag{8.3.6a}$$

$$T = \frac{4k_1k_2}{(k_1 + k_2)^2} \tag{8.3.6b}$$

Notice that these equations are mass-independent, so how can the correspondence principle be satisfied? In other words, how can we recover the result that if we throw a ball against a wall, it will bounce back?

We resolve this issue by noting that no real, physical potential will be as abrupt as in the case we analyzed. In reality, the potential will increase over some region, forming a diffuse step.

It turns out that for (8.3.6) to work we need the wavelength of the incident particles to be much longer than the distance over which the potential increases. This condition is clearly not met for classical particles, which have an extremely short wavelength.

Note that if we interchange $k_1$ and $k_2$ the relations remain unchanged. This corresponds to the fact that a beam will scatter exactly the same way unconditional of the orientation of the barrier (upwards or downwards).

### Wave-packets

We can now use our results using stationary states to construct a physical wave-packet. Indeed, we may superpose different stationary states with amplitude modulated by $f(k_1)$.

Begin by writing down the full stationary state solution (setting $A = 1$, we shall worry about normalization later):

$$\psi(x,t) = \begin{cases} \left(e^{ik_1x} + \frac{k_1-k_2}{k_1+k_2}\right)e^{-iE(k_1)t/\hbar}, & \text{for } x < 0 \\ \frac{2k_1}{k_1+k_2}e^{ik_2x}e^{-iE(k_1)t/\hbar}, & \text{for } x > 0 \end{cases} \tag{8.3.7}$$

Superposing we find that:

$$\psi(x,t) = \begin{cases} \int_{k_0}^{\infty} f(k_1)\left(e^{ik_1x} + \frac{k_1-k_2}{k_1+k_2}\right)e^{-iE(k)t/\hbar}dk_1, & \text{for } x < 0 \\ \int_{k_0}^{\infty} f(k_1)\frac{2k_1}{k_1+k_2}e^{ik_2x}e^{-iE(k)t/\hbar}dk_1, & \text{for } x > 0 \end{cases} \tag{8.3.8}$$

We integrate from $k_0 = \sqrt{\frac{2mV_0}{\hbar}}$ to $\infty$, since we require $E > V_0 \implies k > k_0$. Note that in the limiting case as $V_0 \to 0$, we retrieve $k_0 = 0, k_1 = k_2$ and we retrieve the time evolution of a free wave packet travelling to the right:

$$\int_0^\infty f(k_1) e^{i(k_1 x - E(k)t/\hbar)} dk_1 \tag{8.3.9}$$

We may also write the incident, reflected and transmitted wave packets:

$$\psi_{inc}(x,t) = \int_{k_0}^\infty \left[ f(k_1) e^{ik_1 x} e^{-iE(k_1)t/\hbar} \right] \Theta(-x)\, dk_1 \tag{8.3.10a}$$

$$\psi_{ref}(x,t) = \int_{k_0}^\infty \left[ f(k_1) \Big( \frac{k_1 - k_2}{k_1 + k_2} \Big) e^{ik_1 x} e^{-iE(k_1)t/\hbar} \right] \Theta(-x)\, dk_1 \tag{8.3.10b}$$

$$\psi_{tra}(x,t) = \int_{k_0}^\infty \left[ f(k_1) \Big( \frac{2k_1}{k_1 + k_2} \Big) e^{ik_2 x} e^{-iE(k_1)t/\hbar} \right] \Theta(x)\, dk_1 \tag{8.3.10c}$$

It follows from Parseval's identity that $f(k)$ is the Fourier transform of the initial incident wavepacket $\psi(x, 0)$.

Now let $f(k)$ be sharply peaked at $k_0$. Then using the stationary phase condition, we can investigate how the peak of $\psi_{inc}$ moves by setting the phase in 7.3.10a to be stationary in the vicinity of $k_0$ where the only contribution to the intgral occurs:

$$\frac{d}{dk_1}\Big( kx - \frac{\hbar^2 k_1^2 t}{2m\hbar} \Big)|_{k_0} = 0 \implies x = \frac{\hbar k_0}{m} t \tag{8.3.11}$$

which describes a peak moving to the right at speed $\frac{\hbar k_0}{m}$. At negative times the peak is at $x < 0$. However, we must not have $x > 0$ for the incident packet, hence the peak must not exist for positive time $t > 0$.

This means that the main contribution to $\psi_{inc}$ becomes very very small as $t$ increases, and after sufficiently long it virtually disappears to give rise to the other two wave packets.

Similar treatment for the reflected wave gives:

$$x = -\frac{\hbar k_0}{m} t \tag{8.3.12}$$

Therefore, for $t > 0$, the peak is moving to the left with negative $x$. However, for $t < 0$ we have $x > 0$, which is not allowed for the reflected packet which only exists for $x < 0$. Hence, as in the case of the incident packet, $\psi_{ref}$ must disappear at $t < 0$.

Finally, for $\psi_{trans}$:

$$\frac{d}{dk_1}\Big( k_2 x - \frac{\hbar^2 k_1^2 t}{2m} \Big) k_0 \big|_{k_0} = 0 \tag{8.3.13}$$

It turns out after some algebra that $\frac{dk_2}{dk_1} = \frac{k_1}{k_2}$ thus:

$$\frac{k_1}{k_2}x - \frac{\hbar k_1}{m}t = 0 \implies x = \frac{\hbar k_2}{m}t \tag{8.3.14}$$

Things obviously become very complex for very very small $t$, where all three wave-packets co-exist.

### Step potential with $E < V_0$

Consider the case of a step potential where $E < V_0$. The eigenfunction may be expressed as:

$$\psi(x) = A\left[\left(e^{ik_1x} + \frac{B}{A}e^{-ik_2x}\right)\Theta(-x) + \frac{C}{A}e^{-k_2x}\Theta(x)\right] \tag{8.3.15}$$

where $k_1 = \frac{\sqrt{2mE}}{\hbar}$ and $k_2 = \frac{\sqrt{2m(V_0-E}}{\hbar}$ We then retrieve the boundary conditions:

$$C = \frac{2ik_1}{ik_1 - k_2}A \tag{8.3.16}$$

$$B = \frac{ik_1 + k_2}{ik_1 - k_2}A \tag{8.3.17}$$

Hence:

$$R = \frac{|B|^2}{|A|^2} = 1 \tag{8.3.18}$$

For a finite square step, any particle with insufficient energy will get reflected with $100\%$ certainty.

We may rewrite $\frac{B}{A}$ as:

$$\frac{B}{A} = -\frac{k_2 + ik_1}{k_2 - ik_1} = -e^{2i\delta(E)} \tag{8.3.19}$$

where

$$\delta(E) = \arctan\left(\frac{k_1}{k_2}\right)\arctan\left(\sqrt{\frac{E}{V_0 - E}}\right) \tag{8.3.20}$$

Thus the eigenfunctions may be written as (setting $A = 1$):

$$\psi(x,t) = \begin{cases} \left(e^{ik_1x} - e^{2i\delta(E)}e{-ik_1x}\right)e^{-iEt/\hbar}, & \text{for } x < 0 \\ e^{-k_2x}e^{-iEt/\hbar}, & \text{for } x > 0 \end{cases} \tag{8.3.21}$$

As before, we can use the wave-packet formalism to get a more physical solution:

$$\psi_{inc}(x,t) = \int_0^k f(k_1)\Big(e^{ik_1x} - e^{2i\delta(E)}e{-}ik_1x\Big)e^{-iEt/\hbar}\Theta(-x) \tag{8.3.22}$$

$$\psi_{ref}(x,t) = -\int_0^k f(k_1)\Big(e^{-ik_1x}e^{2i\delta(E)}\Big)e^{-iEt/\hbar}\Theta(-x) \tag{8.3.23}$$

$$\tag{8.3.24}$$

where we integrate from 0 to $k$ because the incoming wave packet cannot be composed of a stationary state with energy greater than $V_0$. As in the case for $E > V_0$ we can apply the stationary phase condition to examine the how the peak of $\psi$ behaves given a momenta distribution $f(k)$ sharply peaked at $k_0$.

For the reflected packet:

$$\frac{d}{dk_1}\Big(-k_1x + 2\delta(E) - \frac{Et}{\hbar}\Big)\Big|_{k_0} = 0 \tag{8.3.25}$$

yielding:

$$x = -\frac{\hbar k_0}{m}(t - \underbrace{2\hbar\delta'(E)}_{timedelay}) \tag{8.3.26}$$

It is interesting to note the time delay $2\hbar\delta'(E)$. One may differentiate 7.3.20 and find that:

$$\delta'(E) = \sqrt{\frac{1}{E(V_0 - E)}} \tag{8.3.27}$$

Without the time delay we would have a perfect reflection of the packet, as was discussed for the step potential $E > V_0$. However, we now have that the packet starts to move to the left only at $t = 2\hbar\delta'(E)$.

Note that the wave-function does not immediately drop to zero, but decays exponentially inside the barrier. This phenomenon is known as **barrier penetration**, and is responsible for quantum tunnelling.

## 8.4 Tunnelling

We now explore the aspect of scattering that has no counterpart in classical mechanics, tunnelling.

Indeed, if the barrier abruptly ended before the wave-function fully decayed (which happens at $x \to \infty$), then there would be a non-zero probability of the particle being detected in the classically forbidden region past the barrier.

Hence, let us consider a barrier of potential:

$$V(x) = \begin{cases} 0, & \text{for } x < 0 \text{ and } L < x \\ V_0, & \text{for} 0 < x < L \end{cases} \tag{8.4.1}$$

The wave-functions are then:

$$\psi(x) = \begin{cases} Ae^{ik_1 x} + Be{-}ik_1 x, & \text{for } x < 0 \\ Ce^{ik_2 x} + De^{-ik_2 x}, & \text{for} 0 < x < L \\ Fe^{ik_1 x}, & \text{for } L < x \end{cases} \tag{8.4.2}$$

with $k_1 = \frac{\sqrt{2mE_0}}{\hbar}$ and $k_2 = \frac{\sqrt{2m(V_0 - E_0)}}{\hbar}$. We can then impose the boundary conditions at $x = 0, L$ and find:

$$A + B = C + D \tag{8.4.3}$$
$$ik_1 A - ik_1 B = -k_2 C + k_2 D \tag{8.4.4}$$
$$Ce^{ik_2 L} + De^{-ik_2 L} = Fe^{ik_1 L} \tag{8.4.5}$$
$$k_2 Ce^{ik_2 L} - k_2 De^{-ik_2 L} = ik_1 Fe^{ik_1 L} \tag{8.4.6}$$

Hence, the transmission coefficient is:

$$T = \frac{|F|^2}{|A|^2} \tag{8.4.7}$$

The calculation of this quantity is quite laborious.

We begin by writing down the boundary conditions in a more suggestive fashion:

$$A + B = C + D \tag{8.4.8}$$
$$A - B = \frac{ik_2}{k_1}(D - C) \tag{8.4.9}$$
$$Ce^{ik_2 L} + De^{-ik_2 L} = Fe^{ik_1 L} \tag{8.4.10}$$
$$Ce^{ik_2 L} - De^{-ik_2 L} = \frac{ik_1}{k_2} Fe^{ik_1 L} \tag{8.4.11}$$
$$\tag{8.4.12}$$

We may then derive:

$$C = \frac{1}{2}\Big(1 + \frac{ik_1}{k_2}\Big) Fe^{ik_1 L} e^{-ik_2 L} \tag{8.4.13}$$
$$D = \frac{1}{2}\Big(1 - \frac{ik_1}{k_2}\Big) Fe^{ik_1 L} e^{ik_2 L} \tag{8.4.14}$$
$$\tag{8.4.15}$$

Hence:

$$A + B = \frac{1}{2}\Big(1 + \frac{ik_1}{k_2}\Big) Fe^{ik_1 L} e^{-ik_2 L} + \frac{1}{2}\Big(1 - \frac{ik_1}{k_2}\Big) Fe^{ik_1 L} e^{ik_2 L} \tag{8.4.16}$$
$$= Fe^{ik_1 L}\Big(\cosh(k_2 L) - \frac{ik_1}{k_2}\sinh(k_2 L)\Big) \tag{8.4.17}$$

and similarly:

$$A - B = -\frac{ik_2}{k_1}\frac{1}{2}\Big(1 + \frac{ik_1}{k_2}\Big)Fe^{ik_1L}e^{-ik_2L} - \frac{1}{2}\Big(1 - \frac{ik_1}{k_2}\Big)Fe^{ik_1L}e^{ik_2L} \tag{8.4.18}$$

$$= Fe^{ik_1L}\Big(\cosh(k_2L) + \frac{ik_2}{k_1}\sinh(k_2L)\Big) \tag{8.4.19}$$

Adding these two expressions we find:

$$2A = Fe^{ik_1L}\Big[2\cosh(k_2L) + i\Big(\frac{k_2}{k_1} - \frac{k_1}{k_2}\Big)\sinh(k_2L)\Big] \tag{8.4.20}$$

giving:

$$\frac{F}{A} = \frac{2e^{-ik_1L}}{2\cosh(k_2L) + i\Big(\frac{k_2}{k_1} - \frac{k_1}{k_2}\Big)\sinh(k_2L)} \tag{8.4.21}$$

Finally:

$$T = \frac{4}{4\cosh^2(k_2L) + \Big(\frac{k_2}{k_1} - \frac{k_1}{k_2}\Big)^2\sinh^2(k_2L)} \tag{8.4.22}$$

$$= \frac{1}{1 + \frac{1}{4}\Big(\frac{k_2}{k_1} + \frac{k_1}{k_2}\Big)^2\sinh^2(k_2L)} \tag{8.4.23}$$

where we used $\cosh^2 x = 1 + \sinh^2 x$. Since $k_1 = \frac{\sqrt{2mE}}{\hbar}$ and $k_2 = \frac{\sqrt{2m(V_0-E)}}{\hbar}$ then:

$$\frac{k_2}{k_1} + \frac{k_1}{k_2} = \sqrt{\frac{E}{V_0 - E_0}}\Big(1 + \frac{V_0 - E}{E}\Big) = \sqrt{\frac{V_0^2}{E(V_0 - E)}} \tag{8.4.24}$$

We can then simplify 8.4.27 as:

$$T = \frac{4E(V_0 - E)}{4E(V_0 - E) + V_0^2\sinh^2\Big[\frac{\sqrt{2mL^2(V_0-E)}}{\hbar}\Big]} \tag{8.4.25}$$

Whenever $k_2L >>> 1$, recall that:

$$\sinh x = \frac{e^x - e{-x}}{2} \approx \frac{1}{2}e^x \tag{8.4.26}$$

so that:

$$T \approx \frac{16E}{V_0}\Big(1 - \frac{E}{V_0}\Big)\exp\Big(-\frac{\sqrt{8mL^2(V_0 - E)}}{\hbar}\Big) \tag{8.4.27}$$

Note the exponential decay of the transmission coefficient as we increase the length of the barrier. This was expected, as we studied the limiting case $L \to \infty$ for a step potential previously.

Also, small note on terminology: $k_2$ is often called the attenuation constant.

## 8.5 Applications

An important type of radioactive decay in physics is alpha decay, consisting of an atomic nucleus emitting an energetic alpha particle. It was first discovered y Rutherford in 1898, who established that each nucleus emits an alpha particle at a specific energy $E_\alpha$. There is a relation between $E_\alpha$ and the decay constant $\lambda$:

$$\lambda = Ae^{-B/E_\alpha^{1/2}} \tag{8.5.1}$$

known as the Geiger-Nuttall law. It was however not understood how the emission of the alpha particle occurred. Indeed, modelling the alpha particle as confined within the nucleus did not explain some cases where the emission energies were lower than the energy required to surmount the nuclear potential.

This problem would be resolved using notions of quantum tunnelling. We may approximate the transmission coefficient as:

$$T \approx \exp\left( -2\int_{r_0}^{r_1} \frac{\sqrt{2m(V-E_\alpha)}}{\hbar} dr \right) \tag{8.5.2}$$

where $r_0$ and $r_1$ are the lower and upper bounds respectively of the classically forbidden region. For a coulomb barrier:

$$V(r) = \frac{2(Z-2)e^2}{4\pi\varepsilon_0 r} \tag{8.5.3}$$

where $2e$ is the charge of the alpha particle, and hence $(Z-2)e$ is the charge of the nucleus. Then after some calculations:

$$T \approx ae^{-b(Z-2)/\sqrt{E_a}} \tag{8.5.4}$$

## 8.6 Resonant transmission

We now consider the potential in 7.4.1 but with $E > V_0$.

The treatment is extremely similar to that of the previous section, and after some calculations one finds that:

$$T = \frac{1}{1 + \frac{1}{4}\frac{V_0^2}{E(E-V_0)} sin^2(k_2 L)} \tag{8.6.1}$$

Note that the transmission coefficient is exactly equal to 1 for some special values of $k_2$. More precisely, $T = 1$ whenever:

$$k_{2,n} L = n\pi,\ n \in \mathbb{Z} \tag{8.6.2}$$

Consequently, we find that:

$$E_n - V_0 = \frac{n^2\pi^2\hbar^2}{2mL^2} \tag{8.6.3}$$

Interestingly, the RHS, the energy of the particle relative to the barrier potential, belongs to the energy spectrum of an infinite well of width $L$ [1].

Particle beams incident on a barrier with this energy undergo **resonant transmission** (total transmission).

An important application of this resonance phenomenon is the Ramsauer-Townsend effect regarding the elastic scattering of electrons off noble gases.

These gases have a full valence shell, and are therefore unreactive. They therefore form a Coulomb potential which an electron could scatter against.

They observed that while at low energies, the scattering cross section was high, as they increased the energy it went down to zero. This effectively meant that all the incident electrons penetrate the electron cloud of the gas atoms.

[1] We may also write $k_2 = \frac{2\pi}{\lambda}$ so that:

$$\frac{L}{\lambda_{DB}} = \frac{n}{2} \tag{8.6.4}$$

Thus, resonance occurs when the de Broglie wavelength of the particle fits half integer times into the width of the well.

# Part II

# Quantum Mechanics and its Interpretation

# Symmetries in Quantum mechanics

9

## 9.1 Translations

In classical mechanics, symmetries of a system lead to conserved quantities. For example, if a system has translational symmetry so that $\mathcal{H} \mapsto \mathcal{H}$ under $x \mapsto x + \epsilon$ then we have a conservation law $\dot{p} = 0$.

How can we formulate similar arguments in quantum mechanics? We have seen by the correspondence principle that expectation values take the role of typical classical variables, so we could interpret translations in QM as:

$$\langle x \rangle \mapsto \langle x \rangle + \alpha \quad \text{or} \quad \langle p \rangle \mapsto \langle p \rangle \tag{9.1.1}$$

so that given a state $|\psi\rangle$, it will get translated to $|\psi_\alpha\rangle$ such that:

$$\langle \psi_\alpha \,|\, \hat{x} \,|\, \psi_\alpha \rangle = \langle \psi \,|\, \hat{x} \,|\, \psi \rangle + \alpha \tag{9.1.2}$$

If we define the translation operator $\hat{T}(\alpha)$ as:

$$\hat{T}(\alpha) \,|\psi\rangle = |\psi_\alpha\rangle \tag{9.1.3}$$

then we can write:

$$\left\langle \psi \,\middle|\, T^\dagger(\alpha) \hat{x} T(\alpha) \,\middle|\, \psi \right\rangle = \langle \psi \,|\, \hat{x} \,|\, \psi \rangle + \alpha \tag{9.1.4a}$$

$$\left\langle \psi \,\middle|\, T^\dagger(\alpha) \hat{p} T(\alpha) \,\middle|\, \psi \right\rangle = \langle \psi \,|\, \hat{p} \,|\, \psi \rangle \tag{9.1.4b}$$

This first picture, known as the **active transformation picture**, physically displaces the particle by $\epsilon$ in the positive $x$ direction.

It is clear from this interpretation that:

$$\hat{T}(\alpha) \,|x\rangle = e^{i\alpha g(x)/\hbar} \,|x + \alpha\rangle \tag{9.1.5}$$

where we must add the phase factor due to the considerations we made in section 6.3. Consequently:

$$\langle x \rangle \xrightarrow{\hat{T}(\alpha)} \langle x \rangle + \alpha \tag{9.1.6}$$

$$\langle p \rangle \xrightarrow{\hat{T}(\alpha)} \langle p \rangle + \alpha \langle f \rangle \tag{9.1.7}$$

where $f = g'$. To satisfy our definition of translation, we must therefore set $g$ constant, for sake of simplicity to zero. Then:

$$\hat{T}(\alpha)\,|x\rangle = |x+\alpha\rangle \tag{9.1.8}$$

Then:

$$|\psi_\alpha\rangle = \hat{T}(\alpha)\,|\psi\rangle = \hat{T}(\alpha)\int_{-\infty}^{\infty} |x\rangle\,\langle x\,|\,\psi\rangle\,dx \tag{9.1.9}$$

$$= \int_{-\infty}^{\infty} |x+\alpha\rangle\,\langle x\,|\,\psi\rangle\,dx \tag{9.1.10}$$

$$= \int_{-\infty}^{\infty} |x'\rangle\,\langle x'-\alpha\,|\,\psi\rangle\,dx \tag{9.1.11}$$

$$\implies \psi_\alpha(x) = \psi(x-\alpha) \tag{9.1.12}$$

where $x' = x + \alpha$. Consequently:

$$\langle\psi_\alpha\,|\,\hat{x}\,|\,\psi_\alpha\rangle = \int_{-\infty}^{\infty} \psi_\alpha^*(x)x\psi_\alpha(x)dx \tag{9.1.13}$$

$$= \int_{-\infty}^{\infty} x|\psi(x-\alpha)|^2dx \tag{9.1.14}$$

$$= \int_{-\infty}^{\infty} (x'+\alpha)|\psi(x')|^2dx' \tag{9.1.15}$$

$$= \langle\psi\,|\,\hat{x}\,|\,\psi\rangle + \alpha \tag{9.1.16}$$

as required by our definition of translation. Similarly:

$$\langle\psi_\alpha\,|\,\hat{p}\,|\,\psi_\alpha\rangle = \int_{-\infty}^{\infty} \psi_\alpha^*(x)\Big(-i\hbar\frac{d}{dx}\Big)\psi_\alpha(x)dx \tag{9.1.17}$$

$$= \int_{-\infty}^{\infty} \psi^*(x')\Big(-i\hbar\frac{d}{dx'}\Big)\psi(x')dx' \tag{9.1.18}$$

$$= \int_{-\infty}^{\infty} (x'-\alpha)|\psi(x')|^2dx' \tag{9.1.19}$$

$$= \langle\psi\,|\,\hat{p}\,|\,\psi\rangle \tag{9.1.20}$$

as required.

We can now define translational invariance through the requirement that:

$$\Big\langle\psi\,\Big|\,\hat{H}\,\Big|\,\psi\Big\rangle = \Big\langle\psi_\alpha\,\Big|\,\hat{H}\,\Big|\,\psi_\alpha\Big\rangle \tag{9.1.21}$$

but how do we get a more explicit conservation law?

Let us try to expand the translation operator as some exponential (we will see why soon):

$$\hat{T}(\alpha) = \sum_{k=0}^{\infty} \frac{1}{k!}\Big(\frac{-ia\hat{G}}{\hbar}\Big)^k \tag{9.1.22}$$

then substituting into 8.1.12 and taylor expanding then:

$$\left\langle x \left| \hat{T}(\alpha) \right| \psi \right\rangle = \psi(x-\alpha) \tag{9.1.23}$$

$$\implies \sum_{k=0}^{\infty} \frac{1}{k!}\Big(\frac{-ia}{\hbar}\Big)^k \left\langle x \left| \hat{G}^k \right| \psi \right\rangle = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \frac{d^{(k)}\psi}{dx^{(k)}} \tag{9.1.24}$$

so that:

$$\Big(-\frac{i\hat{G}}{\hbar}\Big)^k = \frac{d^{(k)}}{dx^{(k)}} \tag{9.1.25}$$

suggesting that $\hat{G} = \hat{p}$. Hence we may alternatively define $\hat{T}(\alpha)$ as:

$$\hat{T}(\alpha) = e^{-i\alpha\hat{p}/\hbar} \tag{9.1.26}$$

The conservation law for translational invariance then becomes:

$$\langle \psi \,|\, H \,|\, \psi \rangle = \langle \psi_\alpha \,|\, H \,|\, \psi_\alpha \rangle \tag{9.1.27}$$

$$= \left\langle \psi \left| e^{i\alpha\hat{p}/\hbar} \hat{H} e^{-i\alpha\hat{p}/\hbar} \right| \psi \right\rangle \tag{9.1.28}$$

$$= \left\langle \psi \left| \Big( \sum_{k=0}^{\infty} \frac{1}{k!}\Big(\frac{i a\hat{p}}{\hbar}\Big)^k \Big) \Big( \sum_{k=0}^{\infty} \frac{\hat{H}}{k!}\Big(\frac{-ia\hat{p}}{\hbar}\Big)^k \Big) \right| \psi \right\rangle \tag{9.1.29}$$

$$= \left\langle \psi \left| \hat{H} \right| \psi \right\rangle + \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{i\alpha}{\hbar}\right)^k \left\langle \psi \left| [{}^k\hat{p}, \hat{H}] \right| \psi \right\rangle = \langle H \rangle \tag{9.1.30}$$

$$\iff \left\langle \psi \left| [\hat{p}, \hat{H}] \right| \psi \right\rangle = 0 \tag{9.1.31}$$

where $[{}^k\hat{p}, \hat{H}] = [\hat{p}, [\hat{p}, \underbrace{\ldots}_{k-2} [\hat{p}, \hat{H}]] \underbrace{\ldots}_{k-2}]$ so we require that:

$$[\hat{p}, \hat{H}] = \mathbb{0} \tag{9.1.32}$$

We now invoke Ehrenfest' theorem to find finally that:

$$\frac{d\langle p \rangle}{dt} \tag{9.1.33}$$

as expected by the correspondence principle.

The second picture assumes that the state vectors are unaltered, and that the operators $\hat{x}$ and $\hat{p}$ are modified as follows:

$$T^\dagger(\alpha)\hat{x}\hat{T}(\alpha) = \hat{x} + \alpha\mathbb{I} \tag{9.1.34a}$$

$$T^\dagger(\alpha)\hat{p}\hat{T}(\alpha) = \hat{p} \tag{9.1.34b}$$

known as the **passive transformation picture**.

We can derive these relations from the active picture by noting that:

$$\left\langle \psi \middle| T^\dagger(\alpha)\hat{x}\hat{T}(\alpha) - \hat{x} - \alpha\mathbb{I} \middle| \psi \right\rangle = 0 \tag{9.1.35}$$

$$\left\langle \psi \middle| T^\dagger(\alpha)\hat{p}\hat{T}(\alpha) - \hat{p} \middle| \psi \right\rangle = 0 \tag{9.1.36}$$

Since $\psi$ is arbitrary, we can choose it to be any eigenvector of the sandwiched operator. Then, all these eigenvalues must be zero, implying that the sandwiched operators are null.

Since these two pictures are equivalent, all the results derived in the active picture also hold in the passive picture. In this picture however, since it is the operators that are altered, we define translational invariance by:

$$\hat{T}^\dagger(\alpha)\hat{H}\hat{T}(\alpha) = \hat{H} \tag{9.1.37}$$

Then, retrieve (8.1.33) as follows:

$$\hat{H} = e^{i\alpha\hat{p}/\hbar}\hat{H}e^{-i\alpha\hat{p}/\hbar} \tag{9.1.38}$$

$$= \left(\sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{ia\hat{p}}{\hbar}\right)^k\right)\left(\sum_{k=0}^{\infty} \frac{\hat{H}}{k!}\left(\frac{-ia\hat{p}}{\hbar}\right)^k\right) \tag{9.1.39}$$

$$= \hat{H} + \sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{i\alpha}{\hbar}\right)^k [{}^k\hat{p}, \hat{H}] = \langle H \rangle \tag{9.1.40}$$

$$\iff \left\langle \psi \middle| [{}^k\hat{p}, \hat{H}] \middle| \psi \right\rangle = 0 \tag{9.1.41}$$

so once again:

$$\boxed{\frac{d\langle p \rangle}{dt} = 0} \tag{9.1.42}$$

Consider for example a system of particles with translational invariance. In the passive picture we define this to be:

$$\hat{T}^\dagger(\alpha)\hat{H}\hat{T}(\alpha) = \hat{H} \implies \hat{H}\hat{T}(\alpha) = \hat{T}(\alpha)\hat{H} \implies [\hat{T}(\alpha), \hat{H}] = 0 \tag{9.1.43}$$

This implies that:

$$[\hat{T}(\alpha), \hat{H}^n] = 0 \tag{9.1.44}$$

for any natural number $n$. The case for $n = 1$ has already been proven. Suppose the above is true for $n - 1$ so that:

$$[\hat{T}(\alpha), \hat{H}^{n-1}] = 0 \tag{9.1.45}$$

Then:

$$[\hat{T}(\alpha), \hat{H}^n] = \hat{T}(\alpha)\hat{H}^n - \hat{H}^n\hat{T}(\alpha) \tag{9.1.46}$$

$$= \hat{T}(\alpha)\hat{H}^n - \hat{H}^{n-1}\hat{T}(\alpha)\hat{H}\hat{T}(\alpha) + \hat{H}^{n-1}\hat{T}(\alpha)\hat{H}\hat{T}(\alpha) - \hat{H}^n\hat{T}(\alpha) \tag{9.1.47}$$

$$= \hat{H}^{n-1}[\hat{T}(\alpha), \hat{H}] + [\hat{T}(\alpha), \hat{H}^{n-1}]\hat{H} \tag{9.1.48}$$

$$= 0 \tag{9.1.49}$$

as desired. So, since $\hat{U}(t) = e^{i\hat{H}t/\hbar}$ can be expanded in a taylor series involving powers of $\hat{H}$ we find that:

$$[\hat{T}(\alpha), \hat{U}(t)] = [\hat{T}(\alpha), \sum_{k=1}^{\infty} \frac{1}{k!}\Big(\frac{it}{\hbar}\Big)^k \hat{H}^k] \tag{9.1.50}$$

$$= \sum_{k=1}^{\infty} \frac{1}{k!}\Big(\frac{it}{\hbar}\Big)^k [\hat{T}(\alpha), \hat{H}^k] \tag{9.1.51}$$

so that:

$$[\hat{T}(\alpha), \hat{U}(t)] = 0 \tag{9.1.52}$$

This is a very important result, because it shows that if we repeat some experiment at two different places at the same instant, they will give the same result.

Indeed, suppose that at $t = 0$ two experimenters $A$ and $B$ prepare the same system at positions $x = 0$ and $x = a$. If we let $|\psi(0)\rangle$ be the state of the system prepared by $A$ then clearly $\hat{T}(\alpha)\,|\psi(0)\rangle$ is the state of the system prepared by $B$. Then, the two systems will evolve as $\hat{U}(t)\,|\psi(0)\rangle$ and $\hat{U}(t)\hat{T}(\alpha)\,|\psi(0)\rangle = \hat{T}(\alpha)\hat{U}\,|\psi(0)\rangle$. The latter is just the same state observed by $A$ at time $t$ but translated by $\alpha$. That means that according to the two experimenters, the two systems will look identical to them.

So, if two experiments are performed in two different parts of the world, then they must give the same results.

# Many-particle systems

10

## 10.1 Two-particle systems

We consider two particles with quantum operators $(\hat{x}_1, \hat{p}_1)$ and $(\hat{x}_2, \hat{p}_2)$. These must satisfy, as per postulate II, the canonical commutation relations:

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{10.1.1}$$
$$[\hat{x}_i, \hat{x}_j] = 0 \tag{10.1.2}$$
$$[\hat{p}_i, \hat{p}_j] = 0 \tag{10.1.3}$$

We can define the coordinate basis in this Hilbert space using the simultaneous eigenkets $|x_1, x_2\rangle$ of the position operators satisfying:

$$\hat{x}_1 |x_1, x_2\rangle = x_1 |x_1, x_2\rangle \tag{10.1.4}$$
$$\hat{x}_2 |x_1, x_2\rangle = x_2 |x_1, x_2\rangle \tag{10.1.5}$$

satisfying the normalization condition:

$$\langle x_1', x_2' \,|\, x_1, x_2\rangle = \delta(x_1' - x_1)\delta(x_2' - x_2) \tag{10.1.6}$$

In this basis:

$$\langle x_1, x_2 \,|\, \psi\rangle = \psi(x_1, x_2) \tag{10.1.7}$$
$$\hat{x_i} \to x_i \tag{10.1.8}$$
$$\hat{p_i} \to -i\hbar\frac{\partial}{\partial x_i} \tag{10.1.9}$$

Born's rule changes so that the probability density of observing particle 1 near $x_1$ and particle two near $x_2$ is:

$$\mathbb{P}(x_1, x_2) = |\langle x_1, x_2 \,|\, \psi\rangle|^2 \tag{10.1.10}$$

We may similarly define the momentum representation using simultaneous eigenkets of $\hat{p}_1$ and $\hat{p}_2$. More generally given any two commuting Hermitian operators $\Omega_1$ and $\Omega_2$, their simultaneous eigenkets provide a $\Omega$-basis. We define this two-particle Hilbert space spanned by any of these bases as $\mathcal{H}_{1\otimes 2}$.

Let us now consider the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ describing particles 1 and 2 respectively

and separately. We then have the commutation relations:

$$[\hat{x}_1^1, \hat{p}_1] = i\hbar\mathbb{I}_1 \tag{10.1.11}$$
$$[\hat{x}_1^2, \hat{p}_2] = i\hbar\mathbb{I}_2 \tag{10.1.12}$$

When a measurement is performed on the two particle system, we expect to obtain a pair of positions $(x_1, x_2)$ so we can denote the ket corresponding to particle 1 being measured at $x_1$ and particle 2 being measured at $x_2$ as:

$$|x_1\rangle \otimes |x_2\rangle \equiv |x_1, x_2\rangle \tag{10.1.13}$$

called the direct product. We then see that the set of vectors of the form $|x_i\rangle \otimes |x_j\rangle$ forms a position basis of the two-particle Hilbert space $\mathcal{H}_{1\otimes 2}$. This however is the exact definition of the tensor product space in Linear algebra:

$$\mathcal{H}_{1\otimes 2} = \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{10.1.14}$$

Hence any element in $\mathcal{H}_1 \otimes \mathcal{H}_2$ may be expressed in the position representation as:

$$|\psi\rangle = \sum_{i,j} c_{ij} \, |x_i\rangle \otimes |x_j\rangle = \sum_{i,j} c_{ij} \, |x_i, x_j\rangle \tag{10.1.15}$$

We define the inner product $\langle \cdot \,|\, \cdot \rangle$ over $\mathcal{H}_1 \otimes \mathcal{H}_2$ so as to satisfy the orthonormality condition (10.1.6). Suppose the inner product over $\mathcal{H}_1$ and $\mathcal{H}_2$ as $\langle \cdot \,|\, \cdot \rangle_1$ and $\langle \cdot \,|\, \cdot \rangle_2$ respectively. Then:

$$(\langle x_1'| \otimes \langle x_2'|)|(|x_1\rangle \otimes |x_2\rangle) = \langle x_1' \,|\, x_2\rangle_1 \langle x_2' \,|\, x_2\rangle_2 = \delta(x_1' - x_1)\delta(x_2' - x_2) \tag{10.1.16}$$

as required. Hence for two arbitrary vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$:

$$|\phi|\psi\rangle = \int |\phi, x_1, x_2\rangle \, x_1, x_2|\psi dx_1 dx_2 = \int \phi^*(x_1, x_2)\psi(x_1, x_2) dx_1 dx_2 \tag{10.1.17}$$

Note that sandwiching (10.1.15) with $\langle x_i| \otimes \langle x_j| = \langle x_i, x_j|$ we get an expression for $c_{ij}$:

$$c_{ij} = \langle x_i, x_j \,|\, \psi\rangle \tag{10.1.18}$$

Of course the results we have discussed in the coordinate basis may be easily expanded to any basis.

We can define the analogue of $\hat{x}_1$ (acting on $\mathcal{H}_1$) on $\mathcal{H}_1 \otimes \mathcal{H}_2$ as $\hat{x}_1^{1\otimes 2}$ so that:

$$\hat{x}_1^{1\otimes 2} \, |x_1\rangle \otimes |x_2\rangle = x_1 \, |x_1\rangle \otimes |x_2\rangle \tag{10.1.19}$$

and similarly

$$\hat{x}_2^{1\otimes 2} \, |x_1\rangle \otimes |x_2\rangle = x_2 \, |x_1\rangle \otimes |x_2\rangle \tag{10.1.20}$$

We can define the direct product of two operators $\Omega_1^{1\otimes 2}$ and $\Omega_2^{1\otimes 2}$ as:

$$(\Omega_1^{1\otimes 2} \otimes \Omega_2^{1\otimes 2}) \, |\omega_1\rangle \otimes |\omega_2\rangle = |\Omega_1^{1\otimes 2}\omega_1\rangle \otimes |\Omega_1^{1\otimes 2}\omega_2\rangle \tag{10.1.21}$$

Note that given two operator $\Omega_1, \Omega_2$ acting on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively:

$$[\Omega_1 \otimes \mathbb{I}_2, \mathbb{I}_1 \otimes \Omega_2] \left|\omega_1, \omega_2\right\rangle = (\Omega_1 \otimes \mathbb{I}_2)(\mathbb{I}_1 \otimes \Omega_2) \left|\omega_1, \omega_2\right\rangle - (\mathbb{I}_1 \otimes \Omega_2)(\Omega_1 \otimes \mathbb{I}_2) \left|\omega_1, \omega_2\right\rangle \tag{10.1.22}$$

$$= (\Omega_1 \otimes \mathbb{I}_2)(\omega_2 \left|\omega_1, \omega_2\right\rangle) - (\mathbb{I}_1 \otimes \Omega_2)(\omega_1 \left|\omega_1, \omega_2\right\rangle) \tag{10.1.23}$$

$$= \omega_1\omega_2 \left|\omega_1, \omega_2\right\rangle - \omega_2\omega_1 \left|\omega_1, \omega_2\right\rangle \tag{10.1.24}$$

$$= 0 \tag{10.1.25}$$

and since any ket $|\psi\rangle$ in this Hilbert space can be written as a linear superposition of these basis vectors, it follows that $[\Omega_1 \otimes \mathbb{I}_2, \mathbb{I}_1 \otimes \Omega_2] = 0$.

Another interesting property is:

$$(\Omega_1 \otimes \Gamma_2)(\Theta_1 \otimes \Lambda_2) \left|x_1, x_2\right\rangle = (\Omega_1 \otimes \Gamma_2) \left|\Theta_1 x_1\right\rangle \otimes \left|\Lambda_2 x_2\right\rangle \tag{10.1.26}$$

$$= \left|\Omega_1\Theta_1 x_1\right\rangle \otimes \left|\Gamma_2\Lambda_2 x_2\right\rangle \tag{10.1.27}$$

$$= (\Omega_1\Theta_1 \otimes \Gamma_2\Lambda_2) \left|x_1, x_2\right\rangle \tag{10.1.28}$$

so that $(\Omega_1 \otimes \Gamma_2)(\Theta_1 \otimes \Lambda_2) = (\Omega_1\Theta_1 \otimes \Gamma_2\Lambda_2)$.

Also:

$$[\Omega_1^{1\otimes 2}, \Lambda_1^{1\otimes 2}] \left|x_1, x_2\right\rangle = (\Omega_1^{1\otimes 2}\Lambda_1^{1\otimes 2} - \Lambda_1^{1\otimes 2}\Omega_1^{1\otimes 2}) \left|x_1, x_2\right\rangle \tag{10.1.29}$$

$$= \Omega_1^{1\otimes 2} \left|\Lambda_1 x_1, x_2\right\rangle - \left|\Lambda_1\Omega_1 x_1, x_2\right\rangle \tag{10.1.30}$$

$$= \left|\Omega_1\Lambda_1 x_1, x_2\right\rangle - \left|\Lambda_1\Omega_1 x_1, x_2\right\rangle \tag{10.1.31}$$

$$= ([\Omega_1, \Lambda_1] \otimes \mathbb{I}_2) \left|x_1, x_2\right\rangle \tag{10.1.32}$$

so that $[\Omega_1^{1\otimes 2}, \Lambda_1^{1\otimes 2}] = [\Omega_1, \Lambda_1] \otimes \mathbb{I}_2$.

Finally:

$$(\Omega_1^{1\otimes 2} + \Omega_2^{1\otimes 2})^2 = (\Omega_1^{1\otimes 2} + \Omega_2^{1\otimes 2})(\Omega_1^{1\otimes 2} + \Omega_2^{1\otimes 2}) \left|x_1, x_2\right\rangle \tag{10.1.33}$$

$$= (\Omega_1^{1\otimes 2} + \Omega_2^{1\otimes 2})(\left|\Omega_1 x_1, x_2\right\rangle + \left|x_1, \Omega_2 x_2\right\rangle) \tag{10.1.34}$$

$$= \left|\Omega_1^2 x_1, x_2\right\rangle + \left|x_1, \Omega_2^2 x_2\right\rangle + 2\left|\Omega_1 x_1, \Omega_2 x_2\right\rangle \tag{10.1.35}$$

$$= (\Omega_1^2 \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \Omega_2^2 + 2\Omega_1 \otimes \Omega_2) \left|x_1, x_2\right\rangle \tag{10.1.36}$$

so that $(\Omega_1^{1\otimes 2} + \Omega_2^{1\otimes 2})^2 = \Omega_1^2 \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \Omega_2^2 + 2\Omega_1 \otimes \Omega_2$.

Suppose that in a single-particle Hilbert space in two dimensions, we have two basis vectors $|+\rangle, |-\rangle$, we have the operators acting on $\mathcal{H}_1$ and $\mathcal{H}_2$:

$$\omega_1 = \begin{matrix} \\ + \\ - \end{matrix}\!\!\begin{matrix} \begin{matrix} + & - \end{matrix} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{matrix} \text{ and } \omega_2 = \begin{matrix} \\ + \\ - \end{matrix}\!\!\begin{matrix} \begin{matrix} + & - \end{matrix} \\ \begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{matrix} \tag{10.1.37}$$

so that for example $\langle + \left|\, \sigma_1 \,\right| + \rangle = a$, $\langle - \left|\, \sigma_1 \,\right| + \rangle = c$, $\langle + \left|\, \sigma_1 \,\right| - \rangle = b$ and $\langle - \left|\, \sigma_1 \,\right| - \rangle = d$.

Then the space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is spanned by $|+,+\rangle, |+,-\rangle, |-,+\rangle$ and $|-,-\rangle$. Then, we can

read off the elements of the operator $\sigma_1 \otimes \mathbb{I}_2$ as:

$$\sigma_1 \otimes \mathbb{I}_2 = \begin{array}{c} \\ ++ \\ +- \\ -+ \\ -- \end{array} \begin{array}{c} \begin{array}{cccc} ++ & +- & -+ & -- \end{array} \\ \begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix} \end{array} \tag{10.1.38}$$

For example, $\langle +,+ | \, \sigma_1 \otimes \mathbb{I}_2 \, | +,+ \rangle = \langle + | \, \sigma_1 \, | + \rangle = a$. All elements where the second state in $\mathcal{H}_2$ are not the same vanish Similarly:

$$\mathbb{I}_1 \otimes \sigma_2 = \begin{array}{c} \\ ++ \\ +- \\ -+ \\ -- \end{array} \begin{array}{c} \begin{array}{cccc} ++ & +- & -+ & -- \end{array} \\ \begin{pmatrix} e & f & 0 & 0 \\ g & h & 0 & 0 \\ 0 & 0 & e & f \\ 0 & 0 & g & h \end{pmatrix} \end{array} \tag{10.1.39}$$

Hence, we can use matrix multiplication to find that:

$$\sigma_1 \otimes \sigma_2 = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & dr & df \\ cg & ch & dg & dh \end{pmatrix} \tag{10.1.40}$$

## 10.2 Time evolution of multi-particle systems

Consider now a two-particle system described by a state vector in $\mathcal{H}_1 \otimes \mathcal{H}_2$. The Schrödinger equation:

$$i\hbar \, |\dot{\psi}\rangle = \left[ \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_1^2}{2m_2} + V(\hat{x}_1, \hat{x}_2) \right] |\psi\rangle \tag{10.2.1}$$

If the Hamiltonian is separable as:

$$\hat{H} = \hat{H}_1 + \hat{H}_2 \tag{10.2.2}$$

so that the two particles evolve independently of each other, then we may write the TISE as:

$$(\hat{H}_1 + \hat{H}_2) \, |E\rangle = E \, |E\rangle \tag{10.2.3}$$

Moreover, since $[\hat{H}_1 \otimes \mathbb{I}_2, \mathbb{I}_1 \otimes \hat{H}_2] = 0$, we may project this equation into a common energy eigenbasis. These eigenstates consist of states $|E_1\rangle \otimes |E_2\rangle = |E_1, E_2\rangle$ satisfying:

$$\hat{H}_1 \, |E_1\rangle = E_1 \, |E_1\rangle \, \hat{H}_2 \, |E_2\rangle = E_2 \, |E_2\rangle \tag{10.2.4}$$

so that particle 1 is in the energy eigenstate $|E_1\rangle$ and particle 2 is in the energy eigenstate $|E_2\rangle$. Then:

$$\hat{H} \, |E\rangle = (\hat{H}_1 + \hat{H}_2) \, |E_1\rangle \otimes |E_2\rangle = (E_1 + E_2) \, |E_1\rangle \otimes |E\rangle_2 = (E_1 + E_2) \, |E\rangle \tag{10.2.5}$$

so we see that the total energy eigenvalues are the sum of the energy eigenvalues of particle 1 and particle 2.

Feeding $|E\rangle = |E_1\rangle \otimes |E\rangle_2$ into the Schrödinger equation:

$$|\psi(t)\rangle = |E_1\rangle\, e^{-iE_1 t/\hbar} \otimes |E_2\rangle\, e^{-iE_2 t/\hbar} \tag{10.2.6}$$

Alternatively, we can work in the position basis and write the TISE as:

$$\Big[ -\frac{\hbar^2}{2m_1}\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big) + V_1(x_1) + V_2(x_2)\Big]\psi_E(x_1, x_2) = E\psi_E(x_1, x_2) \tag{10.2.7}$$

and use separation of variables by using the ansatz:

$$\psi_E(x_1, x_2) = \psi_{E_1}(x_1)\psi_{E_2}(x_2) \tag{10.2.8}$$

to find that:

$$\frac{1}{\psi_{E_1}(x_1)}\Big[ -\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2} + V_1(x_1)\Big]\psi_{E_1}(x_1) = E_1 \tag{10.2.9}$$

$$\frac{1}{\psi_{E_2}(x_2)}\Big[ -\frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2} + V_2(x_2)\Big]\psi_{E_2}(x_2) = E_2 \tag{10.2.10}$$

with $E = E_1 + E_2$. Once the solutions to the above are found:

$$\psi_E(x_1, x_2, t) = \psi_{E_1}(x_1)e^{-iE_1 t/\hbar}\psi_{E_2}(x_2)e^{-iE_2 t/\hbar} \tag{10.2.11}$$

The second case is that of two interacting particles where $V(\hat{x}_1, \hat{x}_2) \neq V(\hat{x}_1) + V(\hat{x}_2)$. In the special case where:

$$V(\hat{x}_1, \hat{x}_2) = V(\hat{x}_1 - \hat{x}_2) \tag{10.2.12}$$

which occurs very often (since most interactions are dependent on the relative displacement of the particles).

Notice that in classical mechanics, we can reduce this problem to two non-interacting fictitious particles by employing the relative coordinate $\hat{x} = \hat{x}_1 - \hat{x}_2$ and the CM coordinate:

$$\hat{x}_{CM} = \frac{m_1\hat{x}_1 + m_2\hat{x}_2}{m_1 + m_2} \tag{10.2.13}$$

so that:

$$\hat{x}_1 = \hat{x}_{CM} + \frac{m_2\hat{x}}{m_1 + m_2} \tag{10.2.14}$$

$$\hat{x}_2 = \hat{x}_{CM} - \frac{m_1\hat{x}}{m_1 + m_2} \tag{10.2.15}$$

and thus using the canonical commutation rules:

$$\hat{p}_{CM}^2 = \hat{p}_1^2 + \hat{p}_2^2 + 2\hat{p}_1\hat{p}_2 \tag{10.2.16}$$
$$\hat{p}^2 = \hat{p}_1^2 + \hat{p}_2^2 - 2\hat{p}_1\hat{p}_2 \tag{10.2.17}$$

The Hamiltonian turns into:

$$\hat{H} = \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} + V(\hat{x}_1, \hat{x}_2) \tag{10.2.18}$$
$$= \frac{\hat{p}_{CM}^2}{2M} + \frac{\hat{p}^2}{2\mu} + V(\hat{x}) \tag{10.2.19}$$

where $\mu = \frac{m_1 m_2}{m_1+m_2}$ is the reduced mass.

**Example (Sh. 10.1.3)**
Consider the classical Hamiltonian of the coupled system:

$$\mathcal{H} = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{1}{2}m\omega^2[x_1^2 + x_2^2 + (x_1 - x_2)^2] \tag{10.2.20}$$

We can use the normal coordinates:

$$x_I = \frac{1}{\sqrt{2}}(x_1 + x_2),\ x_{II} = \frac{1}{\sqrt{2}}(x_1 - x_2) \tag{10.2.21}$$

Then we can write:

$$x_1 = \frac{1}{\sqrt{2}}(x_I + x_{II}) \tag{10.2.22}$$
$$x_2 = \frac{1}{\sqrt{2}}(x_I - x_{II}) \tag{10.2.23}$$
$$p_1 = \frac{1}{\sqrt{2}}(p_I + p_{II}) \tag{10.2.24}$$
$$p_2 = \frac{1}{\sqrt{2}}(p_I - p_{II}) \tag{10.2.25}$$

and so:

$$\mathcal{H} = \frac{(p_I + p_{II})^2}{4m} + \frac{(p_I - p_{II})^2}{4m} + \frac{1}{4}m\omega^2[(x_I + x_{II})^2 + (x_I - x_{II})^2 + 4x_{II}^2] \tag{10.2.26}$$
$$= \frac{p_I^2}{2m} + \frac{p_{II}^2}{2m} + \frac{1}{4}m\omega^2(2x_I^2 + 6x_{II}^2) \tag{10.2.27}$$
$$= \frac{p_I^2}{2m} + \frac{p_{II}^2}{2m} + \frac{1}{2}m\omega^2(x_I^2 + 3x_{II}^2) \tag{10.2.28}$$

We can upgrade these variables to operators and find that the Hamiltonian operator is:

$$\hat{H} = \frac{\hat{p}_I^2}{2m} + \frac{\hat{p}_{II}^2}{2m} + \frac{1}{2}m\omega^2(\hat{x}_I^2 + 3\hat{x}_{II}^2) \tag{10.2.29}$$

The eigenvalue equation for $H$ can be written in the simultaneous eigenbasis of $\hat{x}_I$ and $\hat{x}_{II}$ is:

$$\Big[-\frac{\hbar^2}{2m}\Big(\frac{\partial^2}{\partial x_I^2}+\frac{\partial^2}{\partial x_{II}^2}\Big)+\frac{1}{2}m\omega^2(x_I^2+3x_{II}^2)\Big]\psi(x_I,x_{II})=E\psi(x_I,x_{II}) \tag{10.2.30}$$

We could have done this in the reverse order too, that is, promote to operators first and then use normal coordinates.

## 10.3 Higher dimensions

Although there is no real mathematical difference between a two-particle system of 1D particles and a one-particle system of 2D particles, we shall use different operator notation to differentiate the two.

For a particle in two dimensions, the two cartesian coordinates will be $\hat{x},\hat{y}$, with momenta $\hat{p}_x,\hat{p}_y$ etc... The position eigenket in three dimensions will be written as $|\mathbf{r}\rangle\equiv|x\rangle\otimes|y\rangle\otimes|z\rangle$, and the same goes for momentum.

**Example (Sh. 10.2.1)**
Find the energy eigenfunctions of a particle in a three dimensional cubic box of length $L$.

The Hamiltonian for a three dimensional box is in the position representation:

$$\hat{H}\to-\frac{\hbar^2}{2m}\nabla^2+V(x,y) \tag{10.3.1}$$

where:

$$V(x,y)=\begin{cases}0, & \text{in } [0,L]\times[0,L]\\ \infty, & \text{otherwise}\end{cases} \tag{10.3.2}$$

We see that inside the box, the energy eigenfunctions satisfy:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z)=E\psi(x,y,z) \tag{10.3.3}$$

We can use separation of variables by letting $\psi(x,y,z)=\psi_1(x)\psi_2(y)\psi_3(z)$ to find that:

$$-\frac{\hbar^2}{2m}\Big(\frac{1}{\psi_1(x)}\frac{\partial^2\psi_1(x)}{\partial x^2}+\frac{1}{\psi_2(y)}\frac{\partial^2\psi_2(y)}{\partial y^2}+\frac{1}{\psi_3(z)}\frac{\partial^2\psi_3(z)}{\partial z^2}\Big)=E \tag{10.3.4}$$

Since the LHS has three different independent terms, we can set each equal to some

constant to find three different ODEs:

$$\begin{cases} -\frac{\hbar^2}{2m}\frac{d^2\psi_1(x)}{dx^2} = E_1\psi_1(x) \\ -\frac{\hbar^2}{2m}\frac{d^2\psi_2(y)}{dy^2} = E_2\psi_2(y) \\ -\frac{\hbar^2}{2m}\frac{d^2\psi_3(z)}{dz^2} = E_{\check{c}}\psi_3(z) \end{cases} \tag{10.3.5}$$

with $E = E_1 + E_2 + E_3$. We've already found the solutions:

$$\begin{cases} \psi_1(x) = \sqrt{\frac{2}{L}}\sin\frac{n_x\pi x}{L} \\ \psi_2(y) = \sqrt{\frac{2}{L}}\sin\frac{n_y\pi y}{L} \\ \psi_3(z) = \sqrt{\frac{2}{L}}\sin\frac{n_z\pi z}{L} \end{cases} \tag{10.3.6}$$

where the single-particle energy eigenvalues are:

$$E_1 = \frac{\hbar^2\pi^2}{2mL^2}n_x^2,\ E_2 = \frac{\hbar^2\pi^2}{2mL^2}n_y^2,\ E_3 = \frac{\hbar^2\pi^2}{2mL^2}n_z^2 \tag{10.3.7}$$

so that the total energy eigenvalues are:

$$E = \frac{\hbar^2\pi^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2),\ \forall n_x, n_y, n_z \in \mathbb{N} \tag{10.3.8}$$

with corresponding eigenfunctions:

$$\psi(x,y,z) = \left(\frac{2}{L}\right)^{3/2}\sin\frac{n_x\pi x}{L}\sin\frac{n_y\pi y}{L}\sin\frac{n_z\pi z}{L} \tag{10.3.9}$$

Alternatively, since the Hamiltonian is separable $\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_3$, we can use the simultaneous energy eigenbasis $|E_1, E_2, E_3\rangle$ satisfying:

$$\hat{H}_1|E_1\rangle = E_1|E_1\rangle,\ \hat{H}_2|E_2\rangle = E_2|E_2\rangle,\ \hat{H}_3|E_3\rangle = E_3|E_3\rangle \tag{10.3.10}$$

so that:

$$\hat{H}|E_1,E_2,E_3\rangle = (\hat{H}_1 + \hat{H}_2 + \hat{H}_3)|E_1,E_2,E_3\rangle = (E_1+E_2+E_3)|E_1,E_2,E_3\rangle \tag{10.3.11}$$

$$\equiv E|E_1,E_2,E_3\rangle \tag{10.3.12}$$

with $E = E_1 + E_2 + E_3$.

**Example (Sh. 10.2.2)**
Find the energy eigenstates of the two-dimensional oscillator with hamiltonian:

$$\hat{H} = \frac{\hat{p}_x^2 + \hat{p}_y^2}{2m} + \frac{1}{2}m\omega_x^2\hat{x}^2 + \frac{1}{2}m\omega_y^2\hat{y}^2 \tag{10.3.13}$$

We need to solve the TDSE:

$$\left(\frac{\hat{p}_x^2+\hat{p}_y^2}{2m}+\frac{1}{2}m\omega_x^2\hat{x}^2+\frac{1}{2}m\omega_y^2\hat{y}^2\right)|E\rangle = E\,|E\rangle \tag{10.3.14}$$

Again, since the potential is separable, we can find the simultaneous eigenstates of $\hat{H}_1=\frac{\hat{p}_x^2}{2m}+\frac{1}{2}m\omega_x^2\hat{x}^2$ and $\hat{H}_2=\frac{\hat{p}_y^2}{2m}+\frac{1}{2}m\omega_y^2\hat{y}^2$, which we denote $|E_1,E_2\rangle$. These satisfy:

$$\left(\frac{\hat{p}_x^2}{2m}+\frac{1}{2}m\omega_x^2\hat{x}^2\right)|E_1\rangle = E_1\,|E_1\rangle \tag{10.3.15}$$

$$\left(\frac{\hat{p}_y^2}{2m}+\frac{1}{2}m\omega_y^2\hat{y}^2\right)|E_2\rangle = E_2\,|E_2\rangle \tag{10.3.16}$$

where $E_1=\frac{1}{2}\hbar\omega_x(n_x+1)$ and $E_2=\frac{1}{2}\hbar\omega_y(n_y+1)$.

Therefore

$$\left(\frac{\hat{p}_x^2+\hat{p}_y^2}{2m}+\frac{1}{2}m\omega_x^2\hat{x}^2+\frac{1}{2}m\omega_y^2\hat{y}^2\right)|E_1,E_2\rangle = (E_1+E_2)\,|E_1,E_2\rangle = E\,|E_1,E_2\rangle \tag{10.3.17}$$

implying that the energy eigenvalues are:

$$E=E_1+E_2=\frac{1}{2}\hbar\omega_x(n_x+1)+\frac{1}{2}\hbar\omega_y(n_y+1),\ \forall n_x,n_y=0,1,2... \tag{10.3.18}$$

Projecting onto the position representation we find the normalized eigenfunctions:

$$\psi_{n_x,n_y}(x,y)=\sqrt{\frac{m\sqrt{\omega_x\omega_y}}{\pi\hbar 2^{n_x+n_y}(n_x!)(n_y!)}}\exp\Big(-\frac{m}{2\hbar}(\omega_x x^2+\omega_y y^2)\Big) \tag{10.3.19}$$

$$\times H_{n_x}\left(\sqrt{\frac{m\omega}{\hbar}}x\right)H_{n_y}\left(\sqrt{\frac{m\omega}{\hbar}}y\right) \tag{10.3.20}$$

where we recall that:

$$H_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} \tag{10.3.21}$$

are the Hermite polynomials. Note that if we substitute $x\to -x$ and $y\to -y$, then:

$$\psi_{n_x,n_y}(-x,-y)=\sqrt{\frac{m\sqrt{\omega_x\omega_y}}{\pi\hbar 2^{n_x+n_y}(n_x!)(n_y!)}}\exp\Big(-\frac{m}{2\hbar}(\omega_x x^2+\omega_y y^2)\Big) \tag{10.3.22}$$

$$\times H_{n_x}\left(\sqrt{\frac{m\omega}{\hbar}}x\right)H_{n_y}\left(\sqrt{\frac{m\omega}{\hbar}}y\right) \tag{10.3.23}$$

Since $e^{-x^2}$ is an even function, its $n$th derivative is even if $n$ is even and odd if $n$ is odd. So, the hermite polynomial $H_n$ is odd if $n$ is odd and even if $n$ is even. Consequently,

it follows that the parity of $\psi_{n_x,n_y}(x,y)$ is determined by the parity of $n_x, n_y$:

$$\psi_{n_x,n_y} = \begin{cases} \text{even if } n_x, n_y \text{ have same parity} \\ \text{odd if } n_x, n_y \text{ have different parity} \end{cases} \tag{10.3.24}$$

In the isotropic case, where $\omega_x = \omega_y = \omega$, then:

$$\psi_{n_x,n_y}(x,y) = \sqrt{\frac{m\omega}{\pi\hbar 2^{n_x+n_y}(n_x!)(n_y!)}}\exp\left(-\frac{m\omega}{2\hbar}(x^2+y^2)\right)H_{n_x}\left(\sqrt{\frac{m\omega}{\hbar}}x\right)H_{n_y}\left(\sqrt{\frac{m\omega}{\hbar}}y\right) \tag{10.3.25}$$

So, the first three normalized eigenfunctions are:

$$\psi_{0,0}(x,y) = \sqrt{\frac{m\omega}{\pi\hbar}}\exp\left(-\frac{m\omega}{2\hbar}(x^2+y^2)\right) \tag{10.3.26}$$

$$\psi_{1,0}(x,y) = \sqrt{\frac{2}{\pi}}\frac{m\omega x}{\hbar}\exp\left(-\frac{m\omega}{2\hbar}(x^2+y^2)\right) \tag{10.3.27}$$

$$\psi_{0,1}(x,y) = \sqrt{\frac{2}{\pi}}\frac{m\omega y}{\hbar}\exp\left(-\frac{m\omega}{2\hbar}(x^2+y^2)\right) \tag{10.3.28}$$

Using polar coordinates, we substitute $x = r\cos\theta, y = r\sin\theta, x^2+y^2 = r^2$ and find that:

$$\psi_{0,0}(r,\theta) = \sqrt{\frac{m\omega}{\pi\hbar}}\exp\left(-\frac{m\omega r^2}{2\hbar}\right) \tag{10.3.29}$$

$$\psi_{1,0}(r,\theta) = \sqrt{\frac{2}{\pi}}\frac{m\omega r}{\hbar}\exp\left(-\frac{m\omega r^2}{2\hbar}\right)\cos\theta \tag{10.3.30}$$

$$\psi_{0,1}(r,\theta) = \sqrt{\frac{2}{\pi}}\frac{m\omega r}{\hbar}\exp\left(-\frac{m\omega r^2}{2\hbar}\right)\sin\theta \tag{10.3.31}$$

We see that generally, for a given $n = n_x + n_y$, $n_x$ can range from $0$ to $n$, each with a corresponding value of $n_y$. Consequently the energy level $E_n = \hbar\omega(n+1)$ is $n+1$-fold degenerate.

For the three-dimensional isotropic oscillator, it is easy to see how our results are generalized. The allowed energy levels are:

$$E_n = \hbar\omega\left(n+\frac{3}{2}\right) \tag{10.3.32}$$

where $n = n_x + n_y + n_z$.

Let us try to figure out the degeneracy of $E_n$. The first quantum number can take any value from $0$ to $n$, fixing the sum of the other two quantum numbers. So we find that if the first quantum number is $k$, then the sum of the other two must be $n-k$, and

there are $(n-k+1)$ ways to do so:

$$\sum_{k=0}^{n}(n-k+1) = (n+1)^2 - \frac{n(n+1)}{2} = \frac{(n+1)(n+2)}{2} \tag{10.3.33}$$

Hence $E_n$ is $\frac{(n+1)(n+2)}{2}$-fold degenerate.

## 10.4 Identical particles

Two particles are identical if it is impossible to detect any intrinsic difference in their properties. Electrons for example are believed to be identical.

In classical mechanics, identical particles are distinguishable, we can simply follow their trajectories, or we could "tag" the particles, give them a "name" so as to be able to follow them. Moreover, the classical hamiltonian satisfies the symmetry:

$$\mathcal{H}(r_1, p_1, r_2, p_2) = \mathcal{H}(r_2, p_2, r_1, p_1) \tag{10.4.1}$$

In quantum mechanics the concept of following a trajectory makes no sense, since the particles are not point-like, but rather a wave. When the waves mix, because the particles are indistinguishable there is no way to understand which "part" of the wave belongs to one particle and which belongs to the other. Hence identical particles are indistinguishable in quantum mechanics. Tagging the particles without affect their quantum states is also very hard.

So if we consider two identical particles crossing paths, in quantum mechanics it is impossible to determine where each particle goes after the exchange.

### (Requires knowledge of spin)

Suppose for example we have two electrons, one spin up and one spin down. How do we describe this two-electron system. For distinguishable particles we would just use the tensor product, but which one? There are two:

$$|\uparrow\rangle \otimes |\downarrow\rangle \equiv |\uparrow,\downarrow\rangle \quad \text{or} \quad |\downarrow\rangle \otimes |\uparrow\rangle \equiv |\downarrow,\uparrow\rangle \tag{10.4.2}$$

Unfortunately, these two states are not equivalent. Indeed, if they were equivalent, any linear superposition of these two states must also be equivalent, which is not the case. If we take some superposition:

$$|\psi_{\alpha,\beta}\rangle = \alpha\,|\uparrow,\downarrow\rangle + \beta\,|\downarrow,\uparrow\rangle \tag{10.4.3}$$

such that $|\alpha|^2 + |\beta|^2 = 1$, let us find the probability of measuring spin up along the $x$-axis for both particles, so in a state

$$|\psi_0\rangle = \frac{1}{2}(|\uparrow\rangle + |\downarrow\rangle) \otimes (|\uparrow\rangle + |\downarrow\rangle) \tag{10.4.4}$$

so that:

$$|\langle\psi_0|\psi_{\alpha,\beta}\rangle|^2 = |\frac{1}{2}(\alpha+\beta)|^2 = \frac{1}{2} + \mathbf{Re}(\alpha\beta^*) \tag{10.4.5}$$

Thus the state $|\psi_{\alpha,\beta}\rangle$ changes for different values of $\alpha$ and $\beta$. It is not true that all superpositions of (10.4.2) are physically equivalent.

We resolve this issue by noting that to make $|\uparrow,\downarrow\rangle$ and $|\downarrow,\uparrow\rangle$ equivalent, the two states must differ only by a phase factor:

$$|\uparrow,\downarrow\rangle = e^{i\phi}|\downarrow,\uparrow\rangle = e^{2i\phi}|\uparrow,\downarrow\rangle \implies e^{i\phi} = \pm 1 \tag{10.4.6}$$

so the only allowed state vectors are:

$$|\uparrow,\downarrow\rangle_S \equiv \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle) \tag{10.4.7}$$

called the **symmetric state vector** and:

$$|\uparrow,\downarrow\rangle_A \equiv \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle) \tag{10.4.8}$$

called the **antisymmetric state vector**.

Note also that categories of particles must be either symmetric or anti-symmetric. Indeed, if the Hilbert space of two identical particles contained symmetric and anti-symmetric state vectors, then we can construct another state in this same Hilbert space:

$$|\psi\rangle = \alpha|\uparrow,\downarrow\rangle_S + \beta|\uparrow,\downarrow\rangle_A \tag{10.4.9}$$

This state however is neither symmetric or anti-symmetric, so it cannot describe the two-particle system, and does not belong to the Hilbert space.

Particles always found in symmetric states are called **bosons** whereas particles always found in anti-symmetric states are called **fermions**.

### Pauli exclusion principle

Consider a two-fermion system in a state:

$$|\omega_1,\omega_1\rangle_A = |\omega_1,\omega_2\rangle - |\omega_2,\omega_1\rangle \tag{10.4.10}$$

in some basis $\omega_i$. If we set $\omega_1 = \omega_2 = \omega$ then

$$|\omega,\omega\rangle_A = 0 \tag{10.4.11}$$

so we cannot find two identical fermions in the same quantum state.

## 10.5 Bosonic and Fermionic Hilbert spaces

Consider the Hilbert space of symmetric boson states $\mathcal{H}_S$ and the Hilbert space of anti-symmetric fermionic states $\mathcal{H}_A$.

Consider an arbitrary vector $|\omega_1, \omega_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. We can construct one fermionic vector:

$$|\omega_1, \omega_2\rangle - |\omega_2, \omega_1\rangle \tag{10.5.1}$$

and one bosonic vector:

$$|\omega_1, \omega_2\rangle + |\omega_2, \omega_1\rangle \tag{10.5.2}$$

unless $\omega_1 = \omega_2$ in which case the Pauli exclusion principle allows us to construct only a bosonic vector.

So we can construct one bosonic Hilbert space and one fermionic Hilbert space from $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \mathcal{H}_A \oplus \mathcal{H}_S \tag{10.5.3}$$

Suppose we now want to evaluate the probability density of the measurements by an operator $\Omega$ acting on some state $|\psi\rangle \in \mathcal{H}_S$? Then we must evaluate:

$$\mathbb{P}_S(\omega_1, \omega_2) = |\, \langle \omega_1, \omega_2 \,|\, \psi \rangle_S \,|^2 \tag{10.5.4}$$

where $|\omega_1, \omega_2\rangle_S = \frac{1}{\sqrt{2}}(|\omega_1, \omega_2\rangle + |\omega_2, \omega_1\rangle)$. The normalization condition reads:

$$\frac{1}{2} \int |\, \langle \omega_1, \omega_2 \,|\, \psi \rangle_S \,|^2 d\omega_1 d\omega_2 = 1 \tag{10.5.5}$$

where the $\frac{1}{2}$ terms comes up because the states $|\omega_1, \omega_2\rangle$ and $|\omega_2, \omega_1\rangle$ which are physically equivalent are counted twice. It is therefore convenient to define:

$$\psi_S(\omega_1, \omega_2) = \frac{1}{\sqrt{2}} \langle \omega_1, \omega_2 \,|\, \psi \rangle_S \tag{10.5.6}$$

so that:

$$\int |\psi_S(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2 = 1 \tag{10.5.7}$$

and

$$\mathbb{P}_S(\omega_1, \omega_2) = 2|\psi_S(\omega_1, \omega_2)|^2 \tag{10.5.8}$$

Consider for example two non-interacting bosons in a one-dimensional box in quantum states $n = 3, n = 4$:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|3, 4\rangle + |4, 3\rangle) \tag{10.5.9}$$

Then, the wave-function is:

$$\psi_S(x_1,x_2) = \frac{1}{\sqrt{2}}\langle x_1,x_2\,|\,\psi\rangle_S \tag{10.5.10}$$
$$= \frac{1}{2}(\langle x_1,x_2\,|\,\psi\rangle + \langle x_2,x_1\,|\,\psi\rangle) \tag{10.5.11}$$
$$= \frac{1}{2^{3/2}}((\langle x_1,x_2\,|\,3,4\rangle + \langle x_1,x_2\,|\,4,3\rangle + \langle x_2,x_1\,|\,3,4\rangle + \langle x_2,x_1\,|\,4,3\rangle) \tag{10.5.12}$$
$$= \frac{1}{2^{3/2}}(\psi_3(x_1)\psi_4(x_2) + \psi_4(x_1)\psi_3(x_2) + \psi_3(x_2)\psi_4(x_1) + \psi_4(x_2)\psi_3(x_1)) \tag{10.5.13}$$
$$= \frac{1}{\sqrt{2}}(\psi_3(x_1)\psi_4(x_2) + \psi_4(x_1)\psi_3(x_2)) \tag{10.5.14}$$
$$= \langle x_1,x_2\,|\,\psi\rangle \tag{10.5.15}$$

where $\psi_n$ are the energy eigenfunctions of the one-dimensional box.

We may rewrite this in the form of a sign-less determinant:

$$\psi_S(x_1,x_2) = \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_3(x_1) & \psi_3(x_2)\\ \psi_4(x_1) & \psi_4(x_2)\end{vmatrix}_+ \tag{10.5.16}$$

Similarly, one would find in the fermionic case that:

$$\psi_A(x_1,x_2) = \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_3(x_1) & \psi_3(x_2)\\ \psi_4(x_1) & \psi_4(x_1)\end{vmatrix} \tag{10.5.17}$$

**Example (Sh 10.3.4)**
Two identical particles of mass $m$ are in a 1D box of length $L$. Energy measurement yields $E = \frac{\hbar^2\pi^2}{mL^2}$, what is the state vector of the system? What if the measured energy was $E = \frac{5\hbar^2\pi^2}{mL^2}$?

If we assume that the only degrees of freedom are orbital (so we don't have to worry about spin etc...), we find that the total energy of the system in the state $|n_1,n_2\rangle$ is given by:

$$E = \frac{\hbar^2\pi^2}{2mL^2}(n_1^2+n_2^2) \tag{10.5.18}$$

If $E = \frac{\hbar^2\pi^2}{mL^2}$, then since both $n_1$ and $n_2$ must be positive integers, it follows that $n_1 = n_2 = 1$. Therefore, the particles must be bosons by the Pauli exlusion principle, so we may write the symmetrized state of the system as:

$$|\psi\rangle = |1,1\rangle \tag{10.5.19}$$

In the position representation this becomes:

$$\psi(x_1, x_2) = \langle x_1, x_2 \,|\, 1, 1\rangle_S == \frac{2}{L} \sin\frac{\pi x_1}{L} \sin\frac{\pi x_2}{L} \tag{10.5.20}$$

If instead $E = \frac{5h^2\pi^2}{mL^2}$ then the only possibilities are $n_1 = 1, n_2 = 2$ and $n_1 = 2, n_2 = 1$. Here we cannot determine whether we are dealing with a fermion or a boson. In the former case:

$$|\psi_F\rangle = \frac{|1,2\rangle - |2,1\rangle}{\sqrt{2}} \tag{10.5.21}$$

whereas in the latter case:

$$|\psi_B\rangle = \frac{|1,2\rangle + |2,1\rangle}{\sqrt{2}} \tag{10.5.22}$$

These can be projected in the position basis:

$$\psi_F(x_1, x_2) = \frac{1}{\sqrt{2}}(\psi_1(x_1)\psi_2(x_2) - \psi_2(x_1)\psi_1(x_2)) \tag{10.5.23}$$

$$= \frac{\sqrt{2}}{L}\left(\sin\frac{\pi x_1}{L}\sin\frac{2\pi x_1}{L} - \sin\frac{2\pi x_1}{L}\sin\frac{\pi x_1}{L}\right) \tag{10.5.24}$$

and

$$\psi_B(x_1, x_2) = \frac{1}{\sqrt{2}}(\psi_1(x_1)\psi_2(x_2) + \psi_2(x_1)\psi_1(x_2)) \tag{10.5.25}$$

$$= \frac{\sqrt{2}}{L}\left(\sin\frac{\pi x_1}{L}\sin\frac{2\pi x_1}{L} + \sin\frac{2\pi x_1}{L}\sin\frac{\pi x_1}{L}\right) \tag{10.5.26}$$

## 10.6 $N$ identical particles

### Transposition operators

Consider a two-particle system, where the set of two particles resides in the product space $\mathcal{H} \otimes \mathcal{H}$, with basis $|u_i\rangle \otimes |u_j\rangle \equiv |u_i, u_j\rangle$. We will not assume anything about their distinguishability for now.

Consider some linear operator $\hat{P}_{21} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$, called the permutation operator. Its action on the basis vectors is:

$$\hat{P}_{21} |u_i, u_j\rangle = |u_j, u_i\rangle \tag{10.6.1}$$

so that:

$$\hat{P}_2 1^2 = \mathbb{I} \tag{10.6.2}$$

Note that the permutation operator is Hermitian, since:

$$\left\langle u_i, u_j \,\middle|\, \hat{P}_{21} \,\middle|\, u_k, u_l \right\rangle = \langle u_i, u_j \,|\, u_l, u_k\rangle = \delta il \delta_{jk} \tag{10.6.3}$$

and

$$\left\langle u_i, u_j \,\middle|\, \hat{P}_{21}^\dagger \,\middle|\, u_k, u_l \right\rangle = \langle u_j, u_i \,|\, u_k, u_l\rangle = \delta_{jk} \delta il \tag{10.6.4}$$

Combinining these two results, $\hat{P}_2 1 \hat{P}_{21} = \hat{P}_{21}^\dagger \hat{P}_{21} = \mathbb{I}$ so the permutation operator is also unitary, and thus preserves the inner product.

Let's define:

$$\hat{S} = \frac{1}{2}(1 + \hat{P}_{21}) \tag{10.6.5}$$

$$\hat{A} = \frac{1}{2}(1 - \hat{P}_{21}) \tag{10.6.6}$$

Now for a small algebraic digression. Consider a Hermitian operator $\hat{P}$ acting on $\mathcal{H}$ such that $\hat{P}^2 = \hat{P}$. Then, any vector $\mathbf{v} \in \mathcal{H}$ can be written as:

$$\mathbf{v} = \underbrace{\hat{P}\mathbf{v}}_{\in \mathrm{Im}(\hat{P})} + \underbrace{(1 - \hat{P})\mathbf{v}}_{\in \ker(\hat{P})} \tag{10.6.7}$$

Note however that $\hat{P}\mathbf{v} \in \mathrm{Im}(\hat{P})$ and:

$$\hat{P}(\mathbb{I} - \hat{P})\mathbf{v} = \hat{P}\mathbf{v} - \hat{P}\mathbf{v} = \mathbf{0} \implies (\mathbb{I} - \hat{P})\mathbf{v} \in \ker(\hat{P}) \tag{10.6.8}$$

It follows that:

$$\mathcal{H} = \mathrm{Im}(\hat{P}) \oplus \ker(\hat{P}) \tag{10.6.9}$$

Moreover:

$$\left\langle \hat{P}\mathbf{v}, (1 - \hat{P})\mathbf{u} \right\rangle = \left\langle \mathbf{v}, \hat{P}^\dagger (1 - \hat{P})\mathbf{u} \right\rangle = \left\langle \mathbf{v}, \hat{P}(1 - \hat{P})\mathbf{u} \right\rangle = 0 \tag{10.6.10}$$

implying that $\ker(\hat{P}) \perp \mathrm{Im}(\hat{P})$.

Operators satisfying these two properties:

(i) $\mathcal{H} = \mathrm{Im}(\hat{P}) \oplus \ker(\hat{P})$

(ii) $\ker(\hat{P}) \perp \mathrm{Im}(\hat{P})$

are called **orthogonal projection operators**. We have proven that all Hermitian operators $\hat{P}$ with $\hat{P}^2 = \hat{P}$ are orthogonal projection operators.

It turns out that $\hat{S}$ and $\hat{A}$ defined in (10.6.5) are orthogonal projection operators. Indeed, they are Hermitian. Moreover:

$$\hat{S}^2 = \frac{1}{4}(\mathbb{I} + 2\hat{P}_{21} + \mathbb{I}) = \frac{1}{2}(\mathbb{I} + \hat{P}_{21}) \tag{10.6.11}$$

so that $\hat{S}^2 = \hat{S}$ and similarly $\hat{A}^2 = \hat{A}$.

One final property of the permutation operator worth mentioning is the following. Consider some operator $\hat{B} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$. If we consider:

$$\hat{P}_{21} \hat{B}_1^{\mathcal{H}\otimes\mathcal{H}} \hat{P}_{21}^\dagger \left|u_1\right\rangle \otimes \left|u_2\right\rangle = \hat{P}_{21} \hat{B}_1^{\mathcal{H}\otimes\mathcal{H}} \left|u_2\right\rangle \otimes \left|u_1\right\rangle = \hat{P}_{21} \left|\hat{B}_1^{\mathcal{H}u_2}\right\rangle \otimes \left|u_1\right\rangle = \hat{B}_2^{\mathcal{H}\otimes\mathcal{H}} \left|u_1\right\rangle \otimes \left|u_2\right\rangle \tag{10.6.12}$$

so that:

$$\hat{B}_2^{\mathcal{H}\otimes\mathcal{H}} |u_1\rangle = \hat{P}_{21}\hat{B}_1^{\mathcal{H}\otimes\mathcal{H}}\hat{P}_{21}^\dagger \tag{10.6.13}$$

So, multiplying by $\hat{P}_{21}$ and $\hat{P}_{21}^\dagger$ to the left and right respectively we find:

$$\hat{B}_1^{\mathcal{H}\otimes\mathcal{H}} |u_1\rangle = \hat{P}_{21}\hat{B}_2^{\mathcal{H}\otimes\mathcal{H}}\hat{P}_{21}^\dagger \tag{10.6.14}$$

The permutation operator is very useful when trying to determine whether an operator is symmetric. Indeed, for some general operator $\hat{A}_{12}$:

$$\hat{P}_{21}\hat{A}_{12}hatP_{21}^\dagger = \hat{A}_{21} \tag{10.6.15}$$

so we see that symmetry is satisfied whenever.

$$\hat{P}_{21}\hat{A}_{12}hatP_{21}^\dagger = \hat{A}_{12} \implies \hat{P}_{21}\hat{A}_{12} = \hat{A}_{12}\hat{P}_{21} \implies [\hat{P}_{21}, \hat{A}] = 0 \tag{10.6.16}$$

### Permutation operator on $N$-particles

For $N$ particles, there are $N!$ ways of arranging them in order. The set of permutations of $N$ objects form the symmetric group $S_N$ (see Group theory in mathematical methods volume).

We define these permutations as:

$$\hat{P}_{i_1 i_2 \ldots i_N} |1\rangle |2\rangle \ldots |N\rangle = |i_1\rangle |i_2\rangle \ldots |i_N\rangle \tag{10.6.17}$$

so it moves the $i_n$th ket into the $n$th position. In the two-line notation used in the mathematical methods volume:

$$\hat{P}_{i_1 i_2 \ldots i_N} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 & \ldots & N \\ i_1 & i_2 & i_3 & \ldots & i_N \end{pmatrix} \tag{10.6.18}$$

Transpositions, like $\hat{P}_{21}$, are permutations on $N = 2$ objects. Hence, they are Hermitian and unitary.

Interestingly, it can be proven that any permutation can be written as a product of even and odd transpositions, and the number of transpositions in this product determines the parity of the permutation. The parity of permutations is well-defined, it is either even or odd.

It follows then that any permutation unitary (product of unitary operators is unitary, but product of Hermitian operators is not necessarily Hermitian).

### Symmetric and anti-symmetric states of $N$-particles

Our goal will be to find special states that are simultaneous eigenstates of all permutation operators in the $N$-permutation group $S_N$.

Suppose we have a $N$-particle system described by a state in $\bigotimes^N \mathcal{H}$. We postulate that there are symmetric states $|\psi_S\rangle \in \bigotimes^N \mathcal{H}$, satisfying:

$$\hat{P}_\sigma |\psi_S\rangle = |\psi_S\rangle \,,\ \forall \sigma \in S_N \tag{10.6.19}$$

that is, simultaneous eigenstates of all permutation operators with eigenvalue $+1$. Since permutation operators do not generally commute, we cannot find a simultaneous basis, so we know these symmetric states will not be form a basis for this $N$-particle Hilbert space.

If we follow the same logic we would find that:

$$\hat{P}_\sigma |\psi_A\rangle = -|\psi_A\rangle \,,\ \forall \sigma \in S_N \tag{10.6.20}$$

but this is nonsense, since $\mathbb{I}$ is a permutation, implying $|\psi_A\rangle = -|\psi_A\rangle$. We solve this issue by postulating instead that:

$$\hat{P}_\sigma |\psi_A\rangle = \text{sgn}(\sigma) |\psi_A\rangle \,,\ \forall \sigma \in S_N \tag{10.6.21}$$

The symmetric states form a vector subspace of $\bigotimes^N \mathcal{H}$:

$$\text{Sym}^N(\mathcal{H}) \subseteq \bigotimes^N \mathcal{H} \tag{10.6.22}$$

and similarly the anti-symmetric states form a vector subspace of $\bigotimes^N \mathcal{H}$:

$$\bigwedge^N \mathcal{H} \subseteq \bigotimes^N \mathcal{H} \tag{10.6.23}$$

We must now construct these symmetric and anti-symmetric states. To do so, we note that there should be a projector projecting from $\bigotimes^N \mathcal{H}$ to $\text{Sym}^N(\mathcal{H})$ and $\bigwedge^N \mathcal{H}$. We will prove that:

$$\hat{S} = \frac{1}{N!} \sum_{\sigma \in S_N} \hat{P}_\sigma \tag{10.6.24a}$$

$$\hat{A} = \frac{1}{N!} \sum_{\sigma \in S_N} \text{sgn}(\sigma) \hat{P}_\sigma \tag{10.6.24b}$$

called the **symmetrizer** and anti-symmetrizer operators, are orthogonal projectors that take us to the symmetric and anti-symmetric spaces.

Firstly, let us prove that $\hat{S}$ and $\hat{A}$ are Hermitian. To see why this is the case, consider the list of all permutations in $S_N$. If we apply Hermitian conjugation on them, we will get the

same list, just rearranged. In other words, we need to show that the map:

$$\dagger : S_N \to S_N \tag{10.6.25}$$

$$\hat{P}_\sigma \to \hat{P}_\sigma^\dagger \equiv \hat{P}_{\sigma'}^\dagger \tag{10.6.26}$$

is bijective. This is clearly true, since $\dagger$ is invertible (it is its own inverse). So:

$$\hat{S}^\dagger = \frac{1}{N!}\sum_{\sigma\in S_N}\hat{P}_\sigma^\dagger = \frac{1}{N!}\sum_{\sigma'\in S_N}\hat{P}_{\sigma'} = \hat{S} \tag{10.6.27}$$

The situation is slightly more complex for the anti-symmetrizer. Indeed, how do we know that when we evaluate the Hermitian conjugate of the permutations, their sign will still match up?

Luckily, this is not the case. Indeed, due to the unitarity of permutations:

$$\hat{P}^\dagger\hat{P} = \mathbb{I} \implies \text{sgn}(\hat{P}^\dagger\hat{P}) = 1 \tag{10.6.28}$$

If $\text{sgn}(\hat{P}) \neq \text{sgn}(\hat{P}^\dagger)$, so that one is even and the other odd, we would find that $\text{sgn}(\hat{P}^\dagger\hat{P}) = -1$, a contradiction. So we need that:

$$\text{sgn}(\hat{P}^\dagger) = \text{sgn}(\hat{P}) \tag{10.6.29}$$

and thus:

$$\hat{A}^\dagger = \frac{1}{N!}\sum_{\sigma\in S_N}\text{sgn}(\sigma)\hat{P}_\sigma^\dagger = \frac{1}{N!}\sum_{\sigma}\text{sgn}(\sigma)\hat{P}_{\sigma'}^\dagger = \frac{1}{N!}\sum_{\sigma'\in S_N}\text{sgn}(\sigma')\hat{P}_{\sigma'} = \hat{A} \tag{10.6.30}$$

Next we prove that, given a permutation $\hat{P}_{\sigma_0} \in S_N$, its action on $\hat{S}$ and $\hat{A}$ is simply to rearrange the order of sum of the permutation operators. Indeed

$$\hat{P}_{\sigma_0}\hat{S} = \frac{1}{N!}\sum_{\sigma\in S_N}\hat{P}_{\sigma_0}\hat{P}_\sigma = \frac{1}{N!}\sum_{\sigma'\in S_N}\hat{P}_{\sigma'} = S \tag{10.6.31}$$

and

$$\hat{P}_{\sigma_0}\hat{A} = \frac{1}{N!}\sum_{\sigma\in S_N}\text{sgn}(\sigma)\hat{P}_{\sigma_0}\hat{P}_\sigma \tag{10.6.32}$$

$$= \frac{1}{N!}\sum_{\sigma\in S_N}\text{sgn}(\sigma)\underbrace{\text{sgn}(\sigma_0)\text{sgn}(\sigma_0)}_{=1}\hat{P}_{\sigma_0}\hat{P}_\sigma \tag{10.6.33}$$

$$= \frac{1}{N!}\text{sgn}(\sigma_0)\sum_{\sigma\in S_N}\text{sgn}(\sigma)\text{sgn}(\sigma_0)\hat{P}_{\sigma_0}\hat{P}_\sigma \tag{10.6.34}$$

$$= \frac{1}{N!}\sum_{\sigma'\in S_N}\text{sgn}(\sigma')\hat{P}_{\sigma'} = \text{sgn}(\sigma_0)A \tag{10.6.35}$$

due to the closure property of $S_N$.

The final step in proving that $\hat{A}$ and $\hat{S}$ are orthogonal projectors, is to show that $\hat{S}^2 = \hat{S}$, $\hat{A}^2 = \hat{A}$ and $\hat{S}\hat{A} = \hat{A}\hat{S} = 0$. We will use the two properties proven above in the calculations:

$$\hat{S}^2 = \frac{1}{N!}\sum_{\sigma\in S_N}\hat{P}_\sigma\hat{S} = \frac{1}{N!}\sum_{\sigma\in S_N}\hat{P}_\sigma\hat{S} = \frac{1}{N!}\sum_{\sigma\in S_N}\hat{S} = \hat{S} \tag{10.6.36}$$

since there are $N!$ terms in the sum. Similarly:

$$\hat{A}^2 = \frac{1}{N!}\sum_{\sigma\in S_N}\mathrm{sgn}(\sigma)\hat{P}_\sigma\hat{A} = \frac{1}{N!}\sum_{\sigma\in S_N}\underbrace{\mathrm{sgn}(\sigma)\mathrm{sgn}(\sigma)}_{=1}\hat{A} = \hat{A} \tag{10.6.37}$$

Finally:

$$\hat{A}\hat{S} = \frac{1}{N!}\sum_{\sigma\in S_N}\mathrm{sgn}(\sigma)\hat{P}_\sigma\hat{S} = \frac{1}{N!}\sum_{\sigma}\mathrm{sgn}(\sigma)\hat{S} = \frac{1}{N!}\hat{S}\sum_{\sigma}\mathrm{sgn}(\sigma) = 0 \tag{10.6.38}$$

where the sum vanishes since the number of even permutations is equal to the number of odd permutations. Hence $\hat{S}, \hat{A}$ are indeed orthogonal projection operators. Furthermore, we easily note that for a $|\psi\rangle \in \bigotimes^N \mathcal{H}$:

$$\hat{P}_\sigma\hat{S}\,|\psi\rangle = \hat{S}\,|\psi\rangle \implies \hat{S}\,|\psi\rangle \in \mathrm{Sym}^N(\mathcal{H}) \tag{10.6.39}$$

$$\hat{P}_\sigma\hat{A}\,|\psi\rangle = \mathrm{sgn}(\sigma)\hat{A}\,|\psi\rangle \implies \hat{A}\,|\psi\rangle \in \bigwedge^N \mathcal{H}) \tag{10.6.40}$$

We therefore conclude that:

$$\hat{S} : \bigotimes^N \mathcal{H} \to \mathrm{Sym}^N(\mathcal{H}) \tag{10.6.41a}$$

$$\hat{A} : \bigotimes^N \mathcal{H} \to \bigwedge^N \mathcal{H} \tag{10.6.41b}$$

do indeed project the direct product space onto the symmetric state and anti-symmetric state spaces.

## 10.7 The hidden postulate: Symmetrization Postulate

One might wonder what role mixed states ( states that are neither symmetric or anti-symmetric) of identical particles play in quantum mechanics. Although they are mathematically possible, experiments have never shown them to be physically relevant. We may therefore create an additional postulate:

**Postulate V.** In a system of $N$ identical particles, physically realizable states are either **fermionic** (totally symmetric) or **bosonic** (totally anti-symmetric).

Quantum field theory gives us another way to define fermions and bosons according to their spin. Indeed, integer spin particles are bosons, whereas half-integer spin particles are fermions.

We can use these facts to examine the spin statistics of composite systems. For example, consider two hydrogen atoms. This system will be described by $\psi(e_1, p_1, e_2, p_2)$. Since electrons and protons are both fermions, we need:

$$\psi(e_1, p_1, e_2, p_2) = -\psi(e_2, p_1, e_1, p_2) \text{ and } \psi(e_1, p_1, e_2, p_2) = -\psi(e_1, p_2, e_2, p_1) \tag{10.7.1}$$

Therefore, exchanging both hydrogen atoms is equivalent to exchanging both protons and both electrons:

$$\psi(e_1, p_1, e_2, p_2) = \psi(e_2, p_2, e_1, p_1) \tag{10.7.2}$$

Therefore a hydrogen atom is a boson.

So how does all of this help us solve the exchange degeneracy problem? Given a mathematical state $|\psi\rangle \in \bigotimes^N \mathcal{H}$, there are up to $N!$ many possible representations due to exchange degeneracy. We know however that the only physical representations are the symmetric and anti-symmetric states, which are, as it turns out, unique (up to scaling).

Indeed, suppose there are two states $|\psi_1\rangle, |\psi_2\rangle \in \text{Sym}^N(\mathcal{H})$. Therefore, they can be written as:

$$|\psi_1\rangle = \sum_{\sigma \in S_N} c_\sigma^1 \hat{P}_\sigma |\psi\rangle, \ |\psi_2\rangle = \sum_{\sigma \in S_N} c_\sigma^2 \hat{P}_\sigma |\psi\rangle \tag{10.7.3}$$

and:

$$|\psi_1\rangle = \hat{S}|\psi_1\rangle = \hat{S} \sum_{\sigma \in S_N} c_\sigma^1 \hat{P}_\sigma |\psi\rangle = \sum_\sigma c_\sigma^1 \hat{S}\hat{P}_\sigma |\psi\rangle = \hat{S}|\psi\rangle \sum_\sigma c_\sigma^1 \tag{10.7.4}$$

Similarly:

$$|\psi_2\rangle = \hat{S}|\psi\rangle \sum_{\sigma \in S_N} c_\sigma^2 \implies |\psi_2\rangle \propto |\psi_1\rangle \tag{10.7.5}$$

so the two symmetric states must be proportional to each other. The same applies for anti-symmetric states.

So there is no ambiguity in what state representation to choose for $N$ identical particle-system, depending on the type of particles it is either the fermionic or bosonic state.

Finally, we can use the definition of determinants to simplify the expression for the fermionic state.

Let $|\psi\rangle \in \bigotimes^N \mathcal{H}$, so it may be expanded as:

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes ... \otimes |\psi_N\rangle \tag{10.7.6}$$

Then:

$$\hat{P}_\sigma |\psi\rangle = |\psi_{\sigma(1)}\rangle \otimes |\psi_{\sigma(2)}\rangle \otimes ... \otimes |\psi_{\sigma(N)}\rangle \tag{10.7.7}$$

so that:

$$\hat{A}|\psi\rangle = \frac{1}{N!} \sum_{\sigma \in S_N} \text{sgn}(\sigma) |\psi_{\sigma(1)}\rangle \otimes |\psi_{\sigma(2)}\rangle \otimes ... \otimes |\psi_{\sigma(N)}\rangle \tag{10.7.8}$$

Note that for a generic matrix $A$ with elements $A_{ij}$:

$$\det A = \sum_{\sigma \in S_N} \text{sgn}(\sigma) B_{\sigma(1),1} B_{\sigma(2),2} ... B_{\sigma(N),N} \tag{10.7.9}$$

so that:

$$\psi_A(\mathbf{r}_1, ..., \mathbf{r}_N) = \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \psi_1(\mathbf{r}_3) & \dots & \psi_1(\mathbf{r}_N) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \psi_2(\mathbf{r}_3) & \dots & \psi_2(\mathbf{r}_N) \\ \psi_3(\mathbf{r}_1) & \psi_3(\mathbf{r}_2) & \psi_3(\mathbf{r}_3) & \dots & \psi_3(\mathbf{r}_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1) & \psi_N(\mathbf{r}_2) & \psi_N(\mathbf{r}_3) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix} \tag{10.7.10}$$

Similarly, using a sign-less determinant called permanent:

$$\psi_S(\mathbf{r}_1, ..., \mathbf{r}_N) = \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \psi_1(\mathbf{r}_3) & \dots & \psi_1(\mathbf{r}_N) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \psi_2(\mathbf{r}_3) & \dots & \psi_2(\mathbf{r}_N) \\ \psi_3(\mathbf{r}_1) & \psi_3(\mathbf{r}_2) & \psi_3(\mathbf{r}_3) & \dots & \psi_3(\mathbf{r}_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1) & \psi_N(\mathbf{r}_2) & \psi_N(\mathbf{r}_3) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix}_+ \tag{10.7.11}$$

# Introduction to angular momentum

11

## 11.1 Rotations

We denote by $\hat{U}(\phi\mathbf{k})$ the operator associated with a rotation by $\phi$ about the z-axis. As in the case of translations, we may study this operator from two pictures, namely the active and passive pictures.

Let us start in the active picture, where the state itself gets rotated:

$$|\psi\rangle \mapsto |\psi_R\rangle = \hat{U}(\phi\mathbf{k})\,|\psi\rangle \tag{11.1.1}$$

The rotated state $|\psi_R\rangle$ must satisfy the classical relations:

$$\langle x\rangle_R = \langle x\rangle\cos\phi - \langle y\rangle\sin\phi \tag{11.1.2}$$
$$\langle y\rangle_R = \langle x\rangle\sin\phi + \langle y\rangle\cos\phi \tag{11.1.3}$$
$$\langle p_x\rangle_R = \langle p_x\rangle\cos\phi - \langle p\rangle_y\sin\phi \tag{11.1.4}$$
$$\langle p_y\rangle_R = \langle p_x\rangle\sin\phi + \langle p_y\rangle\cos\phi \tag{11.1.5}$$

where $\left\langle \hat{A}\right\rangle_R$ is the expectation value of $\hat{A}$ in the state $|\psi_R\rangle$. Similarly $\left\langle \hat{A}\right\rangle$ is the expectation value in the state $|\psi\rangle$.

Similarly to translations, we will define the action of $\hat{U}(\psi\mathbf{k})$ in the position representation as:

$$\hat{U}(\psi\mathbf{k})\,|x,y\rangle = |x\cos\phi - y\sin\phi, x\sin\phi + y\cos\phi\rangle \tag{11.1.6}$$

and proceed to construct an explicit form of $\hat{U}$. We begin by considering an infinitesimal rotation $\varepsilon\mathbf{k}$ which we can expand to first order:

$$\hat{U}(\varepsilon\mathbf{k}) = \mathbb{I} - \frac{i\varepsilon\hat{L}_z}{\hbar} \tag{11.1.7}$$

We need to determine $\hat{L}_z$, the **generator of infinitesimal rotations**. Then (11.1.6) becomes:

$$\hat{U}(\varepsilon\mathbf{k})\,|x,y\rangle = |x - y\varepsilon, x\varepsilon + y\rangle \tag{11.1.8}$$

**Infinitesimal rotations**

The infinitesimal rotation operator $\hat{U}(\epsilon\mathbf{k})$ has the action on the basis states:

$$\hat{U}(\varepsilon\mathbf{k})\,|x,y\rangle = |x-y\varepsilon, x\varepsilon+y\rangle$$

and may be expressed as

$$\hat{U}(\varepsilon\mathbf{k}) = \mathbb{1} - \frac{i\varepsilon\hat{L}_z}{\hbar}$$

Now that we have a definition of how the base states rotate infinitesimally, we can infinitesimally rotate any state $|\psi\rangle$ by expanding it in this basis:

$$\left\langle x,y\left|\hat{U}(\varepsilon\mathbf{k})\right|\psi\right\rangle = \iint \left\langle x,y\left|\hat{U}(\varepsilon\mathbf{k})\right|x',y'\right\rangle\langle x',y'\,|\,\psi\rangle\,dx'dy' \tag{11.1.9}$$

$$= \iint \langle x,y\,|\,x'-y'\varepsilon, x'\varepsilon+y'\rangle\,\psi(x',y')dx'dy' \tag{11.1.10}$$

$$= \iint \delta(x'=x+y'\varepsilon, y'=y-x'\varepsilon)\psi(x',y')dx'dy' \tag{11.1.11}$$

$$= \iint \delta(x'=x+y\varepsilon, y'=y-x\varepsilon)\psi(x',y')dx'dy' \tag{11.1.12}$$

$$= \psi(x+y\varepsilon, y-x\varepsilon) \tag{11.1.13}$$

where to go from the third line to fourth line, we solved:

$$y' = y - x'\varepsilon = y - (x+y'\varepsilon)\varepsilon = y - x\varepsilon + \mathcal{O}(\varepsilon^2) \tag{11.1.14}$$

$$x' = x + y'\varepsilon = x + (y-x\varepsilon)\varepsilon = x + y\varepsilon + \mathcal{O}(\varepsilon^2) \tag{11.1.15}$$

We can now taylor expand both sides to first order and use (11.1.7):

$$\langle x,y\,|\,\mathbb{1}\,|\,\psi\rangle - \frac{i\varepsilon}{\hbar}\left\langle x,y\left|\hat{L}_z\right|\psi\right\rangle = \psi(x,y) + \left(y\varepsilon\frac{\partial\psi}{\partial x} - x\varepsilon\frac{\partial\psi}{\partial y}\right) \tag{11.1.16}$$

implying that in the position basis:

$$\left\langle x,y\left|\hat{L}_z\right|\psi\right\rangle = \left(-i\hbar x\frac{\partial}{\partial y} + i\hbar y\frac{\partial}{\partial x}\right)\psi(x,y) \tag{11.1.17}$$

Now we may express (11.1.17) in polar coordinates in a more revealing way. Indeed:

$$\frac{\partial}{\partial\theta} = \frac{\partial x}{\partial\theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\theta}\frac{\partial}{\partial y} \tag{11.1.18}$$

$$= -r\sin\theta\frac{\partial}{\partial x} + r\cos\theta\frac{\partial}{\partial y} \tag{11.1.19}$$

$$= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \tag{11.1.20}$$

We may recognise the momentum and position operators in (11.1.17) so that:

**Generator of infinitesimal rotations**

The infinitesimal rotation operator $\hat{U}(\varepsilon \mathbf{z})$ is generated by the **angular momentum operator** $\hat{L}_z$ defined as:

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{11.1.21}$$

or

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \theta} \tag{11.1.22}$$

In the passive picture we start by looking at how the position and momenta operators are mapped:

$$\hat{U}^\dagger(\phi \mathbf{k})\hat{x}\hat{U}(\phi \mathbf{k}) = \hat{x} - \hat{y}\varepsilon \tag{11.1.23}$$
$$\hat{U}^\dagger(\phi \mathbf{k})\hat{y}\hat{U}(\phi \mathbf{k}) = \hat{y} + \hat{x}\varepsilon \tag{11.1.24}$$
$$\hat{U}^\dagger(\phi \mathbf{k})\hat{p}_x\hat{U}(\phi \mathbf{k}) = \hat{p}_x - \hat{p}_y\varepsilon \tag{11.1.25}$$
$$\hat{U}^\dagger(\phi \mathbf{k})\hat{p}_y\hat{U}(\phi \mathbf{k}) = \hat{p}_y + \hat{p}_x\varepsilon \tag{11.1.26}$$
$$\tag{11.1.27}$$

We can substitute $\hat{U}(\psi \mathbf{k}) = \mathbb{I} - \frac{i\varepsilon \hat{L}_z}{\hbar}$ into the first to find:

$$\left(\mathbb{I} + \frac{i\varepsilon \hat{L}_z}{\hbar}\right)\hat{x}\left(\mathbb{I} - \frac{i\varepsilon \hat{L}_z}{\hbar}\right) = \hat{x} - \frac{i\varepsilon}{\hbar}[\hat{x}, \hat{L}] = \hat{x} - \hat{y}\varepsilon \tag{11.1.28}$$
$$\implies [\hat{x}, \hat{L}_z] = -i\hbar\hat{y} \tag{11.1.29}$$

where we used $L_z^\dagger = L_z$ which follows directly from the fact that $\hat{U}^\dagger(\psi \mathbf{k})\hat{U}(\psi \mathbf{k}) = \mathbb{I}$. Similarly we also find the other relations:

**Commutation relations with $\hat{L}_z$**

We have that:

$$[\hat{x}, \hat{L}_z] = -i\hbar\hat{y} \qquad [\hat{y}, \hat{L}_z] = i\hbar\hat{x}$$
$$[\hat{p}_x, \hat{L}_z] = -i\hbar\hat{p}_y, \qquad [\hat{p}_y, \hat{L}_z] = i\hbar\hat{p}_x$$

These equations yield (**??**)

## 11.2 Derivation of $\hat{L}_z$ (passive picture)

We begin by considering $[\hat{x}_1, f(\hat{\boldsymbol{x}})]$ where $\hat{\boldsymbol{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)$ in multi-index notation (see tensor analysis in math methods volume). Then, assuming $f$ is well behaved, we can write it as the series expansion:

$$f(\hat{\boldsymbol{x}}) = \sum_{n=0}^{\infty} \frac{\partial^n f(\hat{\boldsymbol{x}}))}{n!}\hat{\boldsymbol{x}}^n = \sum_{n=0}^{\infty} k_n \hat{\boldsymbol{x}} \tag{11.2.1}$$

so that:

$$[\hat{x}_1, f(\hat{\boldsymbol{x}})] = \sum_{n=0}^{\infty} k_n [\hat{x}_1, \hat{\boldsymbol{x}}^n] \tag{11.2.2}$$

However:

$$[\hat{x}_1, \hat{\boldsymbol{x}}^n] = [\hat{x}_1, \hat{x}_1^{n_1} \hat{x}_2^{n_2} \hat{x}_3^{n_3} \hat{x}_4^{n_4}] \tag{11.2.3}$$

$$= \hat{x}_1^{n_1} [\hat{x}_1, \hat{x}_2^{n_2}] \hat{x}_3^{n_3} \hat{x}_4^{n_4} + \hat{x}_1^{n_1} \hat{x}_2^{n_2} [\hat{x}_1, \hat{x}_3^{n_3}] \hat{x}_4^{n_4} + \hat{x}_1^{n_1} \hat{x}_2^{n_2} \hat{x}_3^{n_3} [\hat{x}_1, \hat{x}_4^{n_4}] \tag{11.2.4}$$

since $[\hat{x}_1, \hat{x}_1^n] = 0$. Now we must let $\hat{x}_1$ be $\hat{x}, \hat{y}, \hat{p}_x$ or $\hat{p}_y$ so that accordingly:

$$\hat{x}_1 = \hat{x}, \hat{x}_2 = \hat{y}, \hat{x}_3 = \hat{p}_x, \hat{x}_4 = \hat{p}_y \implies [\hat{x}, \hat{\boldsymbol{x}}] = \hat{x}^{n_1} \hat{y}^{n_2} [\hat{x}, \hat{p}_x^{n_3}] \hat{p}_y^{n_4} \tag{11.2.5}$$

$$\hat{x}_1 = \hat{y}, \hat{x}_2 = \hat{x}, \hat{x}_3 = \hat{p}_x, \hat{x}_4 = \hat{p}_y \implies [\hat{y}, \hat{\boldsymbol{x}}] = \hat{y}^{n_1} \hat{x}^{n_2} \hat{p}_x^{n_3} [\hat{y}, \hat{p}_y^{n_4}] \tag{11.2.6}$$

$$\hat{x}_1 = \hat{p}_x, \hat{x}_2 = \hat{x}, \hat{x}_3 = \hat{y}, \hat{x}_4 = \hat{p}_y \implies [\hat{p}_x, \hat{\boldsymbol{x}}] = \hat{p}_x^{n_1} [\hat{p}_x, \hat{x}^{n_2}] \hat{y}^{n_3} \hat{p}_y^{n_4} \tag{11.2.7}$$

$$\hat{x}_1 = \hat{p}_y, \hat{x}_2 = \hat{x}, \hat{x}_3 = \hat{y}, \hat{x}_4 = \hat{p}_x \implies [\hat{p}_y, \hat{\boldsymbol{x}}] = \hat{p}_y^{n_1} \hat{x}^{n_2} [\hat{p}_y, \hat{y}^{n_3}] \hat{p}_y^{n_4} \tag{11.2.8}$$

Consequently we find that:

$$[\hat{x}, \hat{L}_z] = \sum_{n=0}^{\infty} k_n (\hat{x}^{n_1} \hat{y}^{n_2} [\hat{x}, \hat{p}_x^{n_3}] \hat{p}_y) \tag{11.2.9}$$

and using the relation:

$$[\hat{A}, \hat{B}^n] = \sum_{k=1}^{n} \hat{B}^{n-k} [\hat{A}, \hat{B}] \hat{B}^{k-1} \tag{11.2.10}$$

then:

$$[\hat{x}, \hat{p}_x^{n_3}] = \sum_{k=1}^{n_3} \hat{p}_x^{n_3-k} [\hat{x}, \hat{p}_x] \hat{p}_x^{k-1} = i\hbar n_3 \hat{p}^{n_3-1} \tag{11.2.11}$$

Hence:

$$[\hat{x}, \hat{L}_z] = \sum_{n=0}^{\infty} i\hbar k_n n_3 (\hat{x}^{n_1} \hat{y}^{n_2} \hat{p}_x^{n_3-1} \hat{p}_y^{n_4}) = i\hbar \frac{\partial \hat{L}_z}{\partial \hat{p}_x} \tag{11.2.12}$$

Similarly we also find that:

$$[\hat{y}, \hat{L}_z] = \sum_{n=0}^{\infty} i\hbar k_n n_4 (\hat{y}^{n_1} \hat{x}^{n_2} \hat{p}_x^{n_3} \hat{p}_y^{n_4-1}) = i\hbar \frac{\partial \hat{L}_z}{\partial \hat{p}_y} \tag{11.2.13}$$

and:

$$[\hat{p}_x, \hat{L}_z] = -\sum_{n=0}^{\infty} i\hbar k_n n_2 (\hat{p}_x^{n_1} \hat{x}^{n_2-1} \hat{p}_x^{n_3} \hat{p}_y^{n_4}) = -i\hbar \frac{\partial \hat{L}_z}{\partial \hat{x}} \tag{11.2.14}$$

$$[\hat{p}_y, \hat{L}_z] = -\sum_{n=0}^{\infty} i\hbar k_n n_3 (\hat{p}_y^{n_1} \hat{x}^{n_2} \hat{y}^{n_3-1} \hat{p}_y^{n_4}) = -i\hbar \frac{\partial \hat{L}_z}{\partial \hat{y}} \tag{11.2.15}$$

Equating these relations to 8.2.25 and 8.2.26 we get that

$$\begin{cases} \frac{\partial \hat{L}_z}{\partial \hat{p}_x} = -\hat{y} \\ \frac{\partial \hat{L}_z}{\partial \hat{p}_y} = \hat{x} \\ \frac{\partial \hat{L}_z}{\partial \hat{x}} = \hat{p}_y \\ \frac{\partial \hat{L}_z}{\partial \hat{y}} = -\hat{p}_x \end{cases} \implies \begin{cases} \hat{L}_z = g(\hat{x}, \hat{y}, \hat{p}_y) - \hat{y}\hat{p}_x \\ \hat{L}_z = g(\hat{x}, \hat{p}_x, \hat{p}_y) - \hat{y}\hat{p}_x \\ \hat{L}_z = f(\hat{y}, \hat{p}_x, \hat{p}_y) + \hat{x}\hat{p}_y \\ \hat{L}_z = f(\hat{y}, \hat{p}_x, \hat{p}_y) + \hat{x}\hat{p}_y \end{cases} \tag{11.2.16}$$

Comparing the first two we find that:

$$g(\hat{x}, \hat{y}, \hat{p}_x) = g(\hat{x}, \hat{p}_x, \hat{p}_y) \implies g(\hat{x}, \hat{p}_y) \tag{11.2.17}$$

and similarly comparing the latter two:

$$f(\hat{p}_x, \hat{y}, \hat{p}_y) = f(\hat{y}, \hat{p}_x, \hat{p}_y) \implies f(\hat{y}, \hat{p}_x) \tag{11.2.18}$$

Therefore:

$$f(\hat{y}, \hat{p}_x) - g(\hat{x}, \hat{p}_y) = -\hat{x}\hat{p}_y - \hat{y}\hat{p}_x \implies \begin{cases} f(\hat{y}, \hat{p}_x) = -\hat{y}\hat{p}_x + c\mathbb{I} \\ g(\hat{x}, \hat{p}_y) = \hat{x}\hat{p}_y + c\mathbb{I} \end{cases} \tag{11.2.19}$$

and hence we conclude that:

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x + c\mathbb{I} \tag{11.2.20}$$

## 11.3 Finite rotations

Let us now consider a finite rotation $\hat{U}(\phi\mathbf{k})$. We can split the rotation into several infinitesimal rotations by $\phi/N$, and compose them:

$$\hat{U}(\phi\mathbf{k}) = \lim_{N\to\infty} \left(\mathbb{I} - \frac{i\phi}{\hbar N}\hat{L}_z\right)^N = e^{-i\phi\hat{L}_z/\hbar} \tag{11.3.1}$$

hence we may define:

**Finite rotation operator**

The finite rotation operator $\hat{U}(\phi\mathbf{k})$ rotates the state by $\phi$ counterclockwise about the $z$-axis:

$$\hat{U}(\phi\mathbf{k}) = e^{-i\phi\hat{L}_z/\hbar} \tag{11.3.2}$$

It is not immediately clear that this operator rotates the state. If we change to polar coordinates however, we find:

$$\hat{L}_z \longrightarrow -i\hbar\frac{\partial}{\partial\theta} \tag{11.3.3}$$

so that:

$$\hat{U}(\phi\mathbf{k}) \to e^{-\phi\frac{\partial}{\partial\theta}} \tag{11.3.4}$$

If we act on some $\psi(r,\theta)$ then we retrieve a Taylor expansion:

$$e^{-\phi\frac{\partial}{\partial\theta}}\psi(r,\theta)=\sum_{k=0}^{\infty}(-\phi)^k\frac{\partial^k\psi}{\partial\theta^k}=\psi(r,\theta-\phi) \tag{11.3.5}$$

as desired.

We define $\hat{L}_z$ to be the angular momentum operator, since it is analogous to the classical angular momentum $l_z = xp_y - yp_x$ with the substitution of operators instead dynamical variables.

**Rotational invariance**

We define a system with Hamiltonian $\hat{H}$ to be **rotationally invariant** if:

$$\hat{U}^\dagger(R)\hat{H}\hat{U}(R)=\hat{H} \tag{11.3.6}$$

so that, by Ehrenfest's theorem:

$$[\hat{L}_z,\hat{H}]=0 \implies \frac{d\langle L_z\rangle}{dt}=0 \tag{11.3.7}$$

Another important application is that two operators that commutate share a common eigenbasis that diagonalizes them. So in this case the angular momentum operator and the hamiltonian operator share common eigenvalues in some basis, which we shall investigate next.

## 11.4 Eigenvalue problem of $\hat{L}_z$

Consider:

$$\hat{L}_z|l_z\rangle=l_z|l_z\rangle \implies -i\hbar\frac{\partial\psi(r,\theta)}{\partial\theta}=l_z\psi \tag{11.4.1}$$

The solution to this equation is:

$$\psi(r,\theta)=f(r)e^{il_z\theta/\hbar} \tag{11.4.2}$$

where $f(r)$ is some arbitrary normalizable radial function. It seems like there is no restriction on the eigenvalues $l_z$. However, we should also impose the condition of Hermiticity:

$$\left\langle\psi_1\middle|\hat{L}_z\middle|\psi_2\right\rangle=\left\langle\psi_2\middle|\hat{L}_z\middle|\psi_1\right\rangle^* \tag{11.4.3}$$

$$\iff \int_0^{2\pi}\int_0^{\infty}\psi_1^*\Big(-i\hbar\frac{\partial\psi_2}{\partial\theta}\Big)rdrd\theta=\left(\int_0^{\infty}\int_0^{2\pi}\psi_2^*\Big(-i\hbar\frac{\partial\psi_1}{\partial\theta}\Big)rdrd\theta\right)^* \tag{11.4.4}$$

$$\iff \int_0^{2\pi}\int_0^{\infty}\psi_1^*\Big(-i\hbar\frac{\partial\psi_2}{\partial\theta}\Big)rdrd\theta=-\left(\int_0^{\infty}\int_0^{2\pi}\psi_2^*\frac{\partial\psi_1}{\partial\theta}rdrd\theta\right)^* \tag{11.4.5}$$

We can integrate the LHS by parts:

$$\int_0^{2\pi}\int_0^{\infty}\psi_1^*\Big(-i\hbar\frac{\partial\psi_2}{\partial\theta}\Big)rdrd\theta = \int_0^{\infty}\Big([\psi_1^*\psi_2]_0^{2\pi} - \int_0^{2\pi}\psi_2\frac{\partial\psi_1^*}{\partial\theta}d\theta\Big)rdr \tag{11.4.6}$$

$$= \int_0^{\infty}[\psi_1^*\psi_2]_0^{2\pi}rdr - \int_0^{\infty}\int_0^{2\pi}\psi_2\frac{\partial\psi_1^*}{\partial\theta}d\theta rdr \tag{11.4.7}$$

$$= \int_0^{\infty}[\psi_1^*\psi_2]_0^{2\pi}rdr - \left(\int_0^{\infty}\int_0^{2\pi}\psi_2^*\frac{\partial\psi_1}{\partial\theta}d\theta rdr\right)^* \tag{11.4.8}$$

$$= -\left(\int_0^{\infty}\int_0^{2\pi}\psi_2^*\frac{\partial\psi_1}{\partial\theta}d\theta rdr\right)^* \tag{11.4.9}$$

$$\iff \int_0^{\infty}[\psi_1^*\psi_2]_0^{2\pi}rdr = 0 \tag{11.4.10}$$

This condition is satisfied whenever:

$$\psi(r,2\pi) = \psi(r,0) \tag{11.4.11}$$

In this case, we find that:

$$f(r)e^{il_z 2\pi/\hbar} = f(r) \implies l_z = m\hbar,\ \forall m\in\mathbb{Z} \tag{11.4.12}$$

**Angular momentum quantization**

The possible values for angular momentum eigenvalues are integer multiples:

$$l_z = m\hbar,\ \forall m\in\mathbb{Z} \tag{11.4.13}$$

where we call $m$ the **magnetic quantum number**.

Alternatively, we could have also derived this result by considering a superposition of two $l_z$ eigenstates:

$$\psi(r,\theta) = f_1(r)e^{i\theta l_z/\hbar} + f_2(r)e^{i\theta l_z'/\hbar} \tag{11.4.14}$$

then if we rotate by $2\pi$ we expect to get the same state:

$$e^{2\pi i\hat{L}_z/\hbar}\psi(r,\theta) = e^{2\pi i l_z/\hbar}f_1(r)e^{i\theta l_z/\hbar} + e^{2\pi i l_z'/\hbar}f_2(r)e^{i\theta l_z'/\hbar} = \psi(r,\theta) \tag{11.4.15}$$

$$\implies (e^{2\pi i l_z/\hbar}-1)f_1(r)e^{i\theta l_z/\hbar} + (e^{2\pi i l_z'/\hbar}-1)f_2(r)e^{i\theta l_z'/\hbar} = 0 \tag{11.4.16}$$

but the eigenfunctions from distinct eigenvalues are linearly independent, so that:

$$e^{2\pi i l_z/\hbar} = e^{2\pi i l_z'/\hbar} = 1 \implies l_z - l_z' = m\hbar \tag{11.4.17}$$

Now we argue by symmetry that if we can create an eigenstate of angular momentum $l_z$, then we can create another eigenstate of angular momentum $-l_z$ by flipping the system in the plane it resides in. Hence, the angular momentum eigenvalues must be symmetric

about zero so that:

$$l_z = \pm\frac{\hbar}{2}, \pm\frac{3\hbar}{2}... \text{ or } l_z = 0, \pm\hbar, \pm 2\hbar, ... \tag{11.4.18}$$

Notice that choosing a particular eigenvalue does not fix the associated eigenvector, due to the freedom in choosing $f(r)$. Let us define:

$$\Phi_m(\theta) = \frac{1}{\sqrt{2\pi}} e^{im\theta} \tag{11.4.19}$$

so that:

$$\langle \Phi_m \,|\, \Phi_n \rangle = \int_0^{2\pi} \frac{1}{2\pi} e^{i(n-m)\theta} d\theta = \delta_{mn} \tag{11.4.20}$$

giving orthonormality.

**$\hat{L}_z$ eigenfunctions**

We can express the angular momentum eigenfunctions as:

$$|\psi\rangle = |f_m\rangle \otimes |\Phi_m\rangle \iff \psi_m(r,\theta) = f_m(r)\Phi_m(\theta) \tag{11.4.21}$$

where

$$\langle \theta \,|\, \Phi_m \rangle = \Phi_m(\theta) = \frac{1}{\sqrt{2\pi}} e^{im\theta}, \ \langle \Phi_m \,|\, \Phi_n \rangle = \delta_{mn} \tag{11.4.22}$$

Luckily, the choice of $f_m(r)$ does not affect the results of an observation, as we will see in the next example. It can only affect the wave-function which is not directly observable.

**Example (Sh. 12.3.3)**
A particle is described by the wave function:

$$\psi(r,\theta) = Ae^{-r^2/2a^2} \cos^2\theta \tag{11.4.23}$$

What are the possible observable angular momenta, and the probability of measuring them upon observation?

**Solution** We begin by expressing $\psi$ as a superposition of angular momentum eigenfunctions:

$$\psi(r,\theta) = Ae^{-r^2/2a^2} \left(\frac{e^{i\theta} + e^{-i\theta}}{2}\right)^2 \tag{11.4.24}$$

$$= Ae^{-r^2/2a^2} \left(\frac{1}{4}e^{2i\theta} + \frac{1}{4}e^{-2i\theta} + \frac{1}{2}\right) \tag{11.4.25}$$

$$= A(r)\left(\frac{1}{4}\Phi_2 + \frac{1}{4}\Phi_{-2} + \frac{1}{2}\Phi_0\right) \implies |\psi\rangle = A(r)\left(\frac{1}{4}|\Phi_2\rangle + \frac{1}{4}|\Phi_{-2}\rangle \frac{1}{2}|\Phi_0\rangle\right) \tag{11.4.26}$$

We see that the only allowed values of the angular momentum for this particle are $0, \pm 2\hbar$. Their probabilities are:

$$P(l_z = 0) = \frac{|\langle \Phi_0 | \psi \rangle|^2}{\langle \psi | \psi \rangle} = \frac{|A(r)(\frac{1}{4}\langle \Phi_0 | \Phi_2 \rangle + \frac{1}{4}\langle \Phi_0 | \Phi_{-2} \rangle + \frac{1}{2}\langle \Phi_0 | \Phi_0 \rangle)|^2}{|A(r)|^2(\frac{1}{16}\langle \Phi_2 | \Phi_2 \rangle + \frac{1}{16}\langle \Phi_{-2} | \Phi_{-2} \rangle + \frac{1}{4}\langle \Phi_0 | \Phi_0 \rangle)} \tag{11.4.27}$$

$$= \frac{\frac{1}{4}}{\frac{1}{16} + \frac{1}{16} + \frac{1}{4}} = \frac{2}{3} \tag{11.4.28}$$

Similarly:

$$P(l_z = \pm 2\hbar) = \frac{|A(r)\frac{1}{4}\langle \Phi_{\pm 2} | \Phi_{\pm 2} \rangle|^2}{\frac{1}{16} + \frac{1}{16} + \frac{1}{4}} = \frac{1}{6} \tag{11.4.29}$$

◀

The eigenvalue equation for the Hamiltonian in a rotationally invariant problem (where $V$ is independent of $\theta$):

$$\left[ -\frac{\hbar^2}{2\mu}\left( \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \theta^2} \right) + V(r) \right] \psi_E(r, \theta) = E\psi_E(r, \theta) \tag{11.4.30}$$

but since $[\hat{L}_z, \hat{H}] = 0$, the two operators share a common eigenbasis, and we can substitute $\psi_{E,m}(r, \theta) = f_{E,m}(r)\Phi_m(\theta)$ into the above equation to find that:

$$\left[ -\frac{\hbar^2}{2\mu}\left( \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{m^2}{r^2} \right) + V(r) \right] f_{E,m}(r) = E f_{E,m}(r) \tag{11.4.31}$$

Notice the term $\frac{\hbar^2 m^2}{2\mu r^2} = \frac{l_z^2}{2\mu r^2}$, which is caused by substituting angular momentum eigenfunctions. If we take its gradient:

$$\mathbf{F} = -\nabla\left( \frac{\hbar^2 m^2}{2\mu r^2} \right) = \frac{\partial}{\partial r}\left( \frac{\hbar^2 m^2}{2\mu r^2} \right)\hat{\mathbf{r}} = \frac{\hbar^2 m^2}{\mu r^3}\hat{\mathbf{r}} = \frac{l_z^2}{\mu r^3}\hat{\mathbf{r}} \tag{11.4.32}$$

This is similar in form to the classical centrifugal force:

$$\mathbf{F}_{cent} = \frac{\mu v^2}{r}\hat{\mathbf{r}} = \frac{l_z^2}{\mu r^3}\hat{\mathbf{r}} \tag{11.4.33}$$

**Example (Sh. 12.3.6)**
Consider a particle of mass $\mu$ constrained to move on a circular path of radius $a$. Show that $\hat{H} = \frac{\hat{L_z}^2}{2\mu a^2}$ and solve the eigenvalue problem for the Hamiltonian.

**Solution** In the position representation, the Hamiltonian can be expressed as:

$$\hat{H} \to -\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial}{\partial \theta^2}\right) \tag{11.4.34}$$

Since the particle can only move in a circle of radius $a$, only the third tirm survives:

$$\hat{H} \to -\frac{\hbar^2}{2\mu a^2}\frac{\partial^2}{\partial \theta^2} = \frac{1}{2\mu a^2}\left(-i\hbar\frac{\partial}{\partial \theta}\right) \to \frac{\hat{L}_z^2}{2\mu a^2} \tag{11.4.35}$$

Therefore we need to solve the TISE equation:

$$\frac{\hat{L}_z^2}{2\mu a^2}|E\rangle = E|E\rangle \tag{11.4.36}$$

which in polar coordinates reads:

$$-\frac{\hbar^2}{2\mu a^2}\frac{\partial^2\psi_E}{\partial\theta^2} = E\psi_E \implies \frac{\partial^2\psi_E}{\partial\theta^2} = -\frac{2\mu a^2 E}{\hbar^2}\psi_E \tag{11.4.37}$$

Since this is a rotationally invariant problem (the potential $V(r)=0$ is symmetric), we can assert that $\hat{L}_z$ and $\hat{H}$ share a common eigenbasis. So, substituting $\psi_{E,m} = f_{E,m}(r)e^{im\theta}$ we find that:

$$-m^2 e^{im\theta} = -\frac{2\mu a^2 E}{\hbar^2}e^{im\theta} \implies E_m = \frac{m^2\hbar^2}{2\mu a^2} \tag{11.4.38}$$

This means that the energy eigenvalues are two-fold degenerate, for each angular momentum eigenvalue we may have two different energy eigenvalues. ◀

**Example (Sh. 12.3.7)**
Consider the Hamiltonian of the isotropic oscillator:

$$\hat{H} = \frac{\hat{p}_x^2+\hat{p}_y^2}{2\mu}+\frac{1}{2}\mu\omega^2(\hat{x}^2+\hat{y}^2) \tag{11.4.39}$$

Follow the derivation of the 1D harmonic oscillator to find the energy eigenfunctions.

**Solution** Physically, we see that this system is rotationally invariant, since the potential $V(r)=\frac{1}{2}\mu\omega^2 r^2$ only has a radial dependence.

Mathematically, we wish to show that $[\hat{H},\hat{L}_z]=0$. This is easy to verify using the commutation relations in 8.2.25 and 8.2.26 and the identity:

$$[\hat{A}^2,\hat{B}] = \hat{A}[\hat{A},\hat{B}]+[\hat{A},\hat{B}]\hat{A} \tag{11.4.40}$$

then

$$[\hat{p}_x^2, \hat{L}_z] = -i\hbar\hat{p}_x\hat{p}_y \tag{11.4.41}$$
$$[\hat{p}_y^2, \hat{L}_z] = i\hbar\hat{p}_x\hat{p}_y \tag{11.4.42}$$
$$[\hat{x}^2, \hat{L}_z] = -i\hbar\hat{x}\hat{y} \tag{11.4.43}$$
$$[\hat{y}^2, \hat{L}_z] = i\hbar\hat{x}\hat{y} \tag{11.4.44}$$

We find that:

$$[\hat{H}, \hat{L}_z] = 0 \tag{11.4.45}$$

as desired. Hence we can find a common eigenbasis for $\hat{H}$ and $\hat{L}_z$. Therefore, we can substitute angular momentum eigenfunctions $\psi_{E,m}(r,\theta) = f_{E,m}(r)\Phi_m(\theta)$ into the TISE (for sake of brevity we will simply write $f(r)$ instead of $f_{E,m}(r)$:

$$\left[-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{m^2}{r^2}\right) + \frac{1}{2}\mu\omega^2 r^2\right] f(r) = Ef(r) \tag{11.4.46}$$

Now let us examine the limiting case as $r \to 0$. Here we should examine solutions of the form $f(r) = r^k$

$$k(k-1)r^{k-2} + \frac{1}{r}kr^{k-1} - \frac{m^2}{r^2}r^k = -\frac{2\mu E}{\hbar^2}r^k \tag{11.4.47}$$

All terms on the LHS are of power $k-2$, whereas the RHS is of power $k$, and hence vanishes in the limit as $r \to 0$ so that:

$$k(k-1) + k - m^2 = 0 \implies k = |m| \tag{11.4.48}$$

since $f(r)$ should not diverge. Hence $f(r) \to r^{|m|}$ as $r \to 0$.

Now we consider the case $r \to \infty$. Here we retrieve:

$$\frac{\partial^2 f}{\partial r^2} = \frac{\mu^2\omega^2 r^2}{\hbar^2} f \tag{11.4.49}$$

which is identical to the case in the one dimensional oscillator:

$$f(r) \to r^m e^{-\mu\omega r^2/2\hbar} \tag{11.4.50}$$

so up to powers of $r$ we find that:

$$f(r) \to e^{-\mu\omega r^2/2\hbar} \tag{11.4.51}$$

Consequently we assume that $f(r) = r^{|m|}e^{-\mu\omega r^2/2\hbar}U(r)$. In dimensionless variables $\varepsilon = \frac{E}{\hbar\omega}$ and $y = \sqrt{\frac{\mu\omega}{\hbar}}r$ then $f(r) = y^{|m|}e^{-y^2/2}u(r)$. Hence we get:

$$\left(\frac{\partial^2}{\partial y^2} + \frac{1}{y}\frac{\partial}{\partial y} - \frac{m^2}{y^2} - y^2\right) f(y) = -2\varepsilon f(y) \tag{11.4.52}$$

Now we find that, letting $n = |m|$:

$$\frac{\partial f}{\partial y} = e^{-y^2/2}(ny^{n-1}u(y) - y^{n+1}u(y) + y^n u'(y)) \tag{11.4.53}$$

so that:

$$\frac{1}{y}\frac{\partial f}{\partial y} = y^n e^{-y^2/2}(n\frac{u(y)}{y^2} - u(y) + \frac{u'(y)}{y}) \tag{11.4.54}$$

Then

$$\frac{\partial^2 f}{\partial y^2} = y^n e^{-y^2/2}\Big[u(y)\Big(\frac{n(n-1)}{y^2} - 2n - 1 + y^2\Big) + u'(y)\Big(\frac{2n}{y} - 2y\Big) + u''(y)\Big] \tag{11.4.55}$$

and so 8.5.50 turns into:

$$u''(y)+u'(y)\Big(-2y+\frac{2n}{y}+\frac{1}{y}\Big)+u(y)\Big(\frac{n(n-1)}{y^2}-2n-1+y^2-1+\frac{n}{y^2}-\frac{n^2}{y^2}-y^2\Big) = -2\varepsilon u(y) \tag{11.4.56}$$

and upon simplification, and resubstituting $n = |m|$:

$$u''(y) + u'(y)\Big(\frac{2|m|+1}{y} - 2y\Big) + u(y)(2\varepsilon - 2|m| - 2) = 0 \tag{11.4.57}$$

We can now use the power series ansatz:

$$u(y) = \sum_{r=0}^{\infty} c_r y^r \tag{11.4.58}$$

we end up with:

$$\sum_{r=0}^{\infty} c_r\big(r(r+2|m|)y^{r-2} + y^r(2\varepsilon - 2|m| - 2r - 2)\big) = 0 \tag{11.4.59}$$

The first term is equal to:

$$\sum_{r=2}^{\infty} c_r r(r+2|m|)y^{r-2} = \sum_{r=0}^{\infty} c_{r+2}(r+2)(r+2+2|m|)y^r \tag{11.4.60}$$

Hence, since $y^k$ powers are linearly independent, we equate each coefficient to zero:

$$c_{k+2}(k+2)(k+2+2|m|) + c_k(2\varepsilon - 2|m| - 2k - 2) = 0 \tag{11.4.61}$$

and hence the two-term relation for $c_k$ is:

$$\frac{c_{k+2}}{c_k} = \frac{2(|m| + k + 1 - \varepsilon)}{(k+2)(k+2+2|m|)} \tag{11.4.62}$$

Hence, as $y \to \infty$ we find that:

$$\frac{c_{k+2}}{c_k} \to \frac{2}{k} \tag{11.4.63}$$

Observe that $e^{y^2}$ has a similar asymptotic behaviour:

$$e^{y^2} = \sum_{k=0}^{\infty} \frac{y^{2k}}{k!} \tag{11.4.64}$$

the coefficient of the power $y^{r+2}$ is $b_{r+2} = \frac{1}{(\frac{r}{2}+1)!}$, and the coefficient of $y^r$ is $b_r = \frac{1}{(\frac{r}{2})!}$. Their ratio is for large $r$ where $y^r$ dominates:

$$\frac{b_{r+2}}{b_r} = \frac{(\frac{r}{2})!}{(\frac{r}{2}+1)!} = \frac{1}{\frac{r}{2}+1} \cong \frac{2}{r} \tag{11.4.65}$$

In other words, our present solution is asymptotic to $e^{y^2}$, dominating limiting behaviour $e^{-y^2/2}$ inherent to the solution, and giving divergence. This means that the series must be truncated at some term $n$, so that:

$$|m| + n + 1 - \varepsilon = 0 \implies \varepsilon = |m| + n + 1 \tag{11.4.66}$$

This takes care of the behaviour as $y \to \infty$, but what about $y \to 0$? If we take $n$ to be odd, and assuming $c_1 \neq 0$

$$f(y) \cong y^{|m|} \sum_{r=0}^{n} c_r y^r \cong y^{|m|+1} \tag{11.4.67}$$

so we do not retrieve $y^{|m|}$. If we have $c_1 = 0$, then $c_{2k+1} = 0$ for all $k$, and thus the series actually terminates at an even $n$. So we can set $n = 2k$ and find finally that:

$$E_{m,n} = (2k + |m| + 1)\hbar\omega,\ n = 0, 1, 2... \tag{11.4.68}$$

◀

**Example (Sh. 12.3.8)**
Consider a particle of mass $\mu$ and charge $q$ in a vector potential $\mathbf{A} = \frac{B}{2}(-y\mathbf{i} + x\mathbf{j})$.

**Solution** The magnetic field is:

$$\mathbf{B} = \nabla \times \mathbf{A} = B\mathbf{k} \tag{11.4.69}$$

For a classical particle, the equation of motion would become:

$$\mu\ddot{\mathbf{r}} = \frac{q}{c}\dot{\mathbf{r}} \times \mathbf{B} \implies \begin{cases} \ddot{x} = \frac{qB}{\mu c}\dot{y} \\ \ddot{y} = -\frac{qB}{\mu c}\dot{x} \end{cases} \tag{11.4.70}$$

Differentiating the first equation and substituting the second equation:

$$\ddot{v_x} + \left(\frac{qB}{\mu c}\right)^2 v_x = 0 \implies v_x = A\cos(\omega_0 t + \phi_0) \tag{11.4.71}$$

where $\omega_0 = \frac{B}{\mu c}$. Therefore

$$v_y = -A\sin(\omega_0 t + \phi_0) \tag{11.4.72}$$

Hence:

$$x = \frac{A}{\omega_0}\sin(\omega_0 t + \phi) + c_1 \tag{11.4.73}$$

$$y = \frac{A}{\omega_0}\cos(\omega_0 t + \phi) + c_2 \tag{11.4.74}$$

The particle's trajectory is then:

$$(x - c_1)^2 + (y - c_2)^2 = \frac{A^2}{\omega_0^2} \tag{11.4.75}$$

where $c_1, c_2, \phi_0$ are determined by the initial conditions of the particle.

In quantum mechanics, we write the Hamiltonian of this system as:

$$\hat{H} = \frac{|\hat{\mathbf{p}} - q\mathbf{A}/c|^2}{2\mu} = \frac{(\hat{p}_x + \frac{q\hat{y}B}{2c})^2}{2\mu} + \frac{(\hat{p}_y - \frac{q\hat{x}B}{2c})^2}{2\mu} \tag{11.4.76}$$

Note that the operators $\hat{Q} = \frac{c\hat{p}_x + q\hat{y}B/2}{qB}$ and $\hat{P} = \hat{p}_y - q\hat{x}B/2c$ satisfy the canonical commutation relations:

$$[\hat{Q}, \hat{P}] = \left[\frac{(c\hat{p}_x + q\hat{y}B/2)}{qB}, \hat{p}_y - \frac{q\hat{x}B}{2c}\right] \tag{11.4.77}$$

$$= \left[\frac{(c\hat{p}_x + q\hat{y}B/2)}{qB}, \hat{p}_y\right] - \left[\frac{(c\hat{p}_x + q\hat{y}B/2)}{qB}, \frac{q\hat{x}B}{2c}\right] \tag{11.4.78}$$

$$= \frac{1}{2}[\hat{y}, \hat{p}_y] - \left[\frac{c}{qB}\hat{p}_x, \frac{qB}{2c}\hat{x}\right] \tag{11.4.79}$$

$$= \frac{1}{2}[\hat{y}, \hat{p}_y] + \frac{1}{2}[\hat{x}, \hat{p}_x] = i\hbar \tag{11.4.80}$$

We may then write the Hamiltonian (11.4.76) in terms of $\hat{P}$ and $\hat{Q}$ as:

$$\hat{H} = \left(\frac{qB\hat{Q}}{c}\right)^2 \frac{1}{2\mu} + \frac{\hat{P}^2}{2\mu} = \frac{1}{2}\omega_0^2 \mu \hat{Q}^2 + \frac{\hat{P}^2}{2\mu} \tag{11.4.81}$$

This is exactly the form of a harmonic oscillator with $\hat{x} \leftrightarrow \hat{Q}$ and $\hat{p_x} \leftrightarrow \hat{P}$. Since the algebraic solution of the harmonic oscillator relies only on the canonical commutation

relationship, we have that the allowed energy levels are:

$$E = \left(n + \frac{1}{2}\right)\hbar\omega_0,\ n = 0, 1, 2... \tag{11.4.82}$$

Alternatively, we may expand (11.4.76) as:

$$\hat{H} = \frac{\hat{p}_x^2}{2\mu} + \frac{\hat{p}_y^2}{2\mu} + \frac{q^2B^2}{8\mu c^2}(\hat{x}^2 + \hat{y}^2) + \frac{qB}{2\mu c}(\hat{y}\hat{p}_x - \hat{x}\hat{p}_y) \tag{11.4.83}$$

$$= \frac{\hat{p}_x^2}{2\mu} + \frac{\hat{p}_y^2}{2\mu} + \frac{1}{2}\mu\left(\frac{\omega}{2}\right)^2(\hat{x}^2 + \hat{y}^2) - \frac{\omega_0}{2}\hat{L}_z \tag{11.4.84}$$

$$= \hat{H}\left(\frac{\omega_0}{2}, \mu\right) - \frac{\omega_0}{2}\hat{L}_z \tag{11.4.85}$$

where $\hat{H}\left(\frac{\omega_0}{2}, \mu\right)$ is the Hamiltonian of an isotropic two-dimensional harmonic oscillator with mass $\mu$ and frequency $\frac{\omega_0}{2}$. Now consider the basis that diagonalizes $\hat{H}\left(\frac{\omega_0}{2}, \mu\right)$. Since our system has rotational symmetry about the $z$-axis, we can find a common diagonal basis for both $\hat{H}$ and $\hat{L}_z$. Consequently, we must have that $[\hat{H}, \hat{L}_z] = 0$ implying that:

$$[\hat{H}\left(\frac{\omega_0}{2}, \mu\right), \hat{L}_z] = 0 \tag{11.4.86}$$

Therefore the basis diagonalizing $\hat{H}\left(\frac{\omega_0}{2}, \mu\right)$ also diagonalizes $\hat{L}_z$. The eigenvalues of the isotropic oscillator were found to be:

$$(k + \frac{1}{2}|m| + \frac{1}{2})\hbar\omega_0,\ k = 0, 1, 2... \tag{11.4.87}$$

where $m\hbar$ are the associated angular momentum eigenvalues. So:

$$E = (k + \frac{1}{2}|m| - \frac{1}{2}m + \frac{1}{2})\hbar\omega_0 \tag{11.4.88}$$

are the allowed energy levels. If we define $n = k + \frac{1}{2}|m| - \frac{1}{2}m$, and note that $\frac{1}{2}|m| - \frac{1}{2}m$ is either zero or an integer $m$ then we retrieve back (11.4.82). ◀

## 11.5 Angular momentum in 3D

We may define the angular momentum operators as:

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \tag{11.5.1}$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \tag{11.5.2}$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{11.5.3}$$

Once again we consider the infinitesimal rotation operators first, and then generalize to

finite rotations. Here, we find that:

$$\hat{U}(\varepsilon_x \mathbf{i}) \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\varepsilon_x \\ 0 & \varepsilon_x & 1 \end{pmatrix} \tag{11.5.4}$$

$$\hat{U}(\varepsilon_y \mathbf{j}) \leftrightarrow \begin{pmatrix} 1 & 0 & \varepsilon_y \\ 0 & 1 & 0 \\ -\varepsilon_y & 0 & 1 \end{pmatrix} \tag{11.5.5}$$

$$\hat{U}(\varepsilon_z \mathbf{k}) \leftrightarrow \begin{pmatrix} 1 & -\varepsilon_z & 0 \\ \varepsilon_z & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{11.5.6}$$

Consequently, up to order $\varepsilon_x \varepsilon_y$

$$\hat{U}(-\varepsilon_y \mathbf{j})\hat{U}(-\varepsilon_x \mathbf{i})\hat{U}(\varepsilon_y \mathbf{j})\hat{U}(-\varepsilon_x \mathbf{i}) = \begin{pmatrix} 1 & \varepsilon_x\varepsilon_y & 0 \\ \varepsilon_x\varepsilon_y & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \hat{U}(-\varepsilon_x \varepsilon_y \mathbf{k}) \tag{11.5.7}$$

Substituting $\hat{U}(-\varepsilon_x \mathbf{i}) = \mathbb{I} - \frac{i\varepsilon_x \hat{L}_x}{\hbar}$, expanding terms giving $\varepsilon_x \varepsilon_y$ and matching their coefficients:

$$\left(\mathbb{I} + \frac{i\varepsilon_y \hat{L}_y}{\hbar}\right)\left(\mathbb{I} + \frac{i\varepsilon_x \hat{L}_x}{\hbar}\right)\left(\mathbb{I} - \frac{i\varepsilon_y \hat{L}_y}{\hbar}\right)\left(\mathbb{I} - \frac{i\varepsilon_x \hat{L}_x}{\hbar}\right) = \left(\mathbb{I} + \frac{i\varepsilon_x\varepsilon_y \hat{L}_z}{\hbar}\right) \tag{11.5.8}$$

we get that:

$$\frac{\varepsilon_x\varepsilon_y \hat{L}_x\hat{L}_y}{\hbar^2} - \frac{\varepsilon_x\varepsilon_y \hat{L}_y\hat{L}_x}{\hbar^2} + \frac{\varepsilon_x\varepsilon_y \hat{L}_y\hat{L}_x}{\hbar^2} - \frac{\varepsilon_x\varepsilon_y \hat{L}_y\hat{L}_x}{\hbar^2} = \frac{\varepsilon_x\varepsilon_y [\hat{L}_x, \hat{L}_y]}{\hbar^2} = \frac{i\varepsilon_x\varepsilon_y \hat{L}_z}{\hbar} \tag{11.5.9}$$

We can repeat this process with $\varepsilon_y \varepsilon_z$ and $\varepsilon_x \varepsilon_z$ to find that:

**Commutator algebra of angular momentum operators**

The angular momentum operators $\hat{L}_x, \hat{L}_y, \hat{L}_z$ satisfy the commutator algebra:

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \tag{11.5.10a}$$

$$[\hat{L}_y, \hat{L}_z] = i\hbar \hat{L}_x \tag{11.5.10b}$$

$$[\hat{L}_z, \hat{L}_x] = i\hbar \hat{L}_y \tag{11.5.10c}$$

They may be written more compactly as:

$$\hat{\mathbf{L}} \times \hat{\mathbf{L}} = i\hbar \hat{\mathbf{L}} \tag{11.5.11}$$

or

$$[\hat{L}_i, \hat{L}_j] = i\hbar \sum_{k=1}^{3} \epsilon_{ijk} \hat{L}_k \tag{11.5.12}$$

The total angular momentum operator is then defined similarly to vectors:

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{11.5.13}$$

Then, clearly:

$$\begin{aligned}
[\hat{L}^2, \hat{L}_i] = \sum_{j=1}^{3} [\hat{L}_j^2, \hat{L}_i] &= \sum_{j=1}^{3} (\hat{L}_j [\hat{L}_j, \hat{L}_i] + [\hat{L}_j, \hat{L}_i] \hat{L}_j) && (11.5.14)\\
&= i\hbar \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{jik} (\hat{L}_j \hat{L}_k + \hat{L}_k \hat{L}_j) && (11.5.15)\\
&= i\hbar \Big( \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{jik} \hat{L}_j \hat{L}_k - \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{kij} \hat{L}_k \hat{L}_j \Big) && (11.5.16)\\
&= i\hbar \Big( \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{jik} \hat{L}_j \hat{L}_k - \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{jik} \hat{L}_j \hat{L}_k \Big) && (11.5.17)\\
&= 0 && (11.5.18)
\end{aligned}$$

We may now construct the finite rotation operator in three dimensions. We consider a rotation by an angle $\theta$ in the direction $\hat{\theta}$, and let $\delta\boldsymbol{\theta}$ be an infinitesimal angle about $\hat{\theta}$. Then the vector $\mathbf{r}$ gets mapped by the rotation $\hat{U}(\delta\boldsymbol{\theta})$:

$$\mathbf{r} \to \mathbf{r} + \delta\boldsymbol{\theta} \times \mathbf{r} \tag{11.5.19}$$

Consequently:

$$\begin{aligned}
\hat{U}(\delta\boldsymbol{\theta}) |\psi\rangle &= \int_{-\infty}^{\infty} \hat{U}(\delta\boldsymbol{\theta}) |\mathbf{r}'\rangle \langle \mathbf{r}' | \psi\rangle \, d\mathbf{r}' && (11.5.20)\\
&= \int_{-\infty}^{\infty} |\mathbf{r}' + \delta\boldsymbol{\theta} \times \mathbf{r}'\rangle \langle \mathbf{r}' | \psi\rangle \, d\mathbf{r}' && (11.5.21)\\
\implies \Big\langle \mathbf{r} \Big| \hat{U}(\delta\boldsymbol{\theta}) \Big| \psi \Big\rangle &= \int_{-\infty}^{\infty} \langle \mathbf{r} | \mathbf{r}' + \delta\boldsymbol{\theta} \times \mathbf{r}'\rangle \langle \mathbf{r}' | \psi\rangle \, d\mathbf{r}' && (11.5.22)\\
&= \int_{-\infty}^{\infty} \delta(\mathbf{r} - (\mathbf{r}' + \delta\boldsymbol{\theta} \times \mathbf{r}')) \langle \mathbf{r}' | \psi\rangle \, d\mathbf{r}' && (11.5.23)\\
&= \psi(\mathbf{r} - \delta\boldsymbol{\theta} \times \mathbf{r}) \approx \psi(\mathbf{r}) - (\delta\boldsymbol{\theta} \times \mathbf{r}) \cdot \nabla\psi(\mathbf{r}) && (11.5.24)
\end{aligned}$$

where we used a taylor expansion to first order. Comparing this to:

$$\hat{U}(\delta\boldsymbol{\theta}) = \mathbb{I} - \frac{i\delta\theta \hat{L}_\theta}{\hbar} \tag{11.5.25}$$

we find that:

$$\hat{L}_\theta = (\hat{\theta} \times \mathbf{r}) \cdot (-i\hbar\nabla) = -\hat{\theta} \cdot ((-i\hbar\nabla) \times \mathbf{r}) = \hat{\theta} \cdot (\mathbf{r} \cdot (-i\hbar\nabla)) = \hat{\theta} \cdot \hat{\mathbf{L}} \tag{11.5.26}$$

so $\hat{L}_\theta = \hat{\theta} \cdot \hat{\mathbf{L}}$ is the generator of rotations about $\hat{\theta}$, and thus:

$$\hat{U}(\boldsymbol{\theta}) = \lim_{N\to\infty} \left( \mathbb{I} - \frac{i\boldsymbol{\theta} \cdot \hat{\mathbf{L}}}{N\hbar} \right)^N = e^{-i\boldsymbol{\theta}\cdot\hat{\mathbf{L}}/\hbar} \tag{11.5.27}$$

**Finite rotation operator**

The finite rotation operator $\hat{U}(\boldsymbol{\theta})$ rotates the state by $\theta$ counterclockwise about the $\hat{\theta}$-axis:

$$\hat{U}(\boldsymbol{\theta}) = e^{-i\boldsymbol{\theta}\cdot\hat{\mathbf{L}}/\hbar} \tag{11.5.28}$$

Once again, for rotational invariance to occur we require that:

$$\hat{U}(\boldsymbol{\theta})^\dagger \hat{H} \hat{U}(\boldsymbol{\theta}) = \hat{H} \tag{11.5.29}$$

for any $\boldsymbol{\theta}$. Choosing $\boldsymbol{\theta}$ to be about the $x, y, z$ axes we find that:

$$[\hat{H}, \hat{L}_i] = 0 \implies [\hat{H}, \hat{L}^2] = 0 \tag{11.5.30}$$

It follows that we can always find a common eigenbasis for $\hat{H}, \hat{L}^2$ and one of $\hat{L}_i$, usually $\hat{L}_z$. We cannot find a common eigebasis for all $\hat{L}_i$ since they do not commute with each other.

## 11.6 The eigen-problem for $\hat{\mathbf{L}}^2$ (operator)

Due to their commutativity we know that $\hat{L}^2$ and $\hat{L}_z$ share a common eigenbasis, say $|\alpha, \beta\rangle$:

$$\hat{L}^2 |\alpha, \beta\rangle = \alpha |\alpha, \beta\rangle \tag{11.6.1}$$

$$\hat{L}_z |\alpha, \beta\rangle = \beta |\alpha, \beta\rangle \tag{11.6.2}$$

**Lowering and raising operators**

We can now define the raising and lower operators:

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y \tag{11.6.3}$$

with

$$[\hat{L}_z, \hat{L}_\pm] = \pm\hbar\hat{L}_\pm \qquad [\hat{L}^2, \hat{L}_\pm] = 0 \tag{11.6.4}$$

These operators are called accordingly because the raise/lower the angular momentum

eigenvalue $l_z$ by $\hbar$ without altering the eigenvalue of $\hat{L}^2$. For example:

$$\hat{L}_z(\hat{L}_\pm |\alpha, \beta\rangle) = (\hat{L}_\pm \hat{L}_z \pm \hbar \hat{L}_\pm) |\alpha, \beta\rangle \tag{11.6.5}$$

$$= (\beta \pm \hbar)\hat{L}_\pm |\alpha, \beta\rangle \tag{11.6.6}$$

and:

$$\hat{L}^2(\hat{L}_\pm |\alpha, \beta\rangle) = \alpha \hat{L}_\pm |\alpha, \beta\rangle \tag{11.6.7}$$

From these it follows that $\hat{L}_\pm |\alpha, \beta\rangle$ is an eigenket of $\hat{L}_z$ with eigenvalue $\beta \pm \hbar$. Similarly, $\hat{L}_\pm |\alpha, \beta\rangle$ is also an eigenket of $\hat{L}^2$ with eigenvalue $\alpha$. Hence $\hat{L}_\pm |\alpha, \beta\rangle \propto |\alpha, \beta \pm \hbar\rangle$, and we can denote the constant of proportionality by $C_\pm(\alpha, \beta)$ so that:

$$\hat{L}_\pm |\alpha, \beta\rangle = C_\pm(\alpha, \beta) |\alpha, \beta \pm \hbar\rangle \tag{11.6.8}$$

Now observe that $\hat{L}_\pm$ produces infinitely many eigenstates $|\alpha, \beta + n\hbar\rangle$ of $\hat{L}_z$ for a fixed value of $\hat{L}^2$. This violates common sense, since classically we expect $l_z^2 \leq l^2$.

More rigorously, in our notation:

$$\left\langle \alpha, \beta \middle| \hat{L}^2 - \hat{L}_z^2 \middle| \alpha, \beta \right\rangle = \alpha - \beta^2 = \left\langle \alpha, \beta \middle| \hat{L}_x^2 + \hat{L}_y^2 \middle| \alpha, \beta \right\rangle \tag{11.6.9}$$

and since $\hat{L}_x^2 + \hat{L}_y^2$ is positive, this implies that $\alpha \geq \beta^2$. So, there must be a state $|\alpha, \beta_{max}\rangle$ which can no longer be raised:

$$\hat{L}_+ |\alpha, \beta_{max}\rangle = 0 \tag{11.6.10}$$

and a state $|\alpha, \beta_{min}\rangle$ which can no longer be lowered:

$$\hat{L}_- |\alpha, \beta_{min}\rangle = 0 \tag{11.6.11}$$

Now let us take the adjoint of (11.6.8):

$$\langle \alpha, \beta | \hat{L}_\mp = \langle \alpha, \beta \pm \hbar | C_\pm^*(\alpha, \beta) \tag{11.6.12}$$

and sandwiching (11.6.8) with (11.6.12) we find that:

$$\left\langle \alpha, \beta \middle| \hat{L}_\mp \hat{L}_\pm \middle| \alpha, \beta \right\rangle = |C_\pm(\alpha, \beta)|^2 \langle \alpha, \beta \pm \hbar | \alpha, \beta \pm \hbar \rangle = |C_\pm(\alpha, \beta)|^2 \tag{11.6.13}$$

Also:

$$\hat{L}_\mp \hat{L}_\pm = (\hat{L}_x \mp i\hat{L}_y)(\hat{L}_x \pm i\hat{L}_y) \tag{11.6.14}$$

$$= \hat{L}_x^2 + \hat{L}_y^2 \pm i(\hat{L}_x \hat{L}_y - \hat{L}_y \hat{L}_x) \tag{11.6.15}$$

$$= \hat{L}^2 - \hat{L}_z^2 \mp \hbar \hat{L}_z \tag{11.6.16}$$

Consequently we can write (11.6.13) as:

$$|C(\alpha, \beta)|^2 = |\hat{L}_\pm |\alpha, \beta\rangle|^2 = \alpha - \beta^2 \mp \hbar\beta \tag{11.6.17}$$

Therefore:

$$|\hat{L}_+ |\alpha, \beta_{max}\rangle|^2 = \alpha - \beta_{max}^2 - \hbar\beta_{max} = 0 \tag{11.6.18}$$
$$|\hat{L}_- |\alpha, \beta_{min}\rangle|^2 = \alpha - \beta_{min}^2 + \hbar\beta_{min} = 0 \tag{11.6.19}$$

It follows immediately that:

$$\alpha = \beta_{max}(\beta_{max} + \hbar),\ \beta_{max} = -\beta_{min} \tag{11.6.20}$$

Also note that to go from $|\alpha, \beta_{max}\rangle$ to $|\alpha, \beta_{min}\rangle$, we need to apply $\hat{L}_-$ $k$ times for some integer $k$, decreasing $\beta$ by $\hbar$ every time. Thus:

$$\beta_{max} - \beta_{min} = 2\beta_{max} = k\hbar \implies \beta_{\max} = \frac{\hbar k}{2} \tag{11.6.21}$$

so that given some integer $k$, we refer to $\frac{k}{2}$ as the angular momentum of the state:

$$\beta_{max} = \frac{\hbar k}{2} = -\beta_{min},\ \alpha = \beta_{max}(\beta_{max} + \hbar),\ k = 0, 1, 2... \tag{11.6.22}$$

In more conventional notation, we set $l = \beta_{max}$ and $m = \beta$, so the eigenvalues $m\hbar$ of $\hat{L}_z$ form a ladder with spacing $\hbar$. The maximum and minimum eigenvalues are given by $\pm l\hbar$ respectively, where $l$ is the maximum value of $m$. The corresponding eigenvalue for $\hat{L}^2$ is $\alpha = l(l+1)\hbar^2$.

**Eigenvalues of $\hat{L}^2$ and $\hat{L}_z$**

The eigenvalues of $\hat{L}^2$ and $\hat{L}_z$ are:

$$\hat{L}^2 |l, m\rangle = l(l+1)\hbar^2 |l, m\rangle,\ l = 0, \frac{1}{2}, 1, ... \tag{11.6.23a}$$
$$\hat{L}_z |l, m\rangle = m\hbar |l, m\rangle, \qquad |m| \leq l \tag{11.6.23b}$$

We call $l$ **the angular momentum quantum number**, and $m$ **the magnetic quantum number**.

Here we should have called $|l, m\rangle$ as $|l(l+1), m\rangle$, but for sake of brevity the former notation was adopted. We then rewrite the eigenvalue equation as:

$$\hat{L}_\pm |l, m\rangle = C_\pm(l, m) |l, m \pm 1\rangle \tag{11.6.24}$$

Now, using (11.6.23):

$$|C_\pm(l, m)|^2 = \left\langle l, m \middle| \hat{L}^2 - \hat{L}_z^2 \mp \hbar\hat{L}_z \middle| l, m \right\rangle \tag{11.6.25}$$
$$= l(l+1)\hbar^2 - m^2\hbar \pm m\hbar^2 \tag{11.6.26}$$
$$= \hbar^2(l \mp m)(l \pm m + 1) \tag{11.6.27}$$

so that:

**Ladder operator matrix elements**

The lowering and raising operators act on $|l,m\rangle$ as follows:

$$\hat{L}_\pm |l,m\rangle = \sqrt{(l \mp m)(l \pm m + 1)}\hbar\, |l, m \pm 1\rangle \tag{11.6.28}$$

We may therefore express the matrix components of $\hat{L}_x$ as:

$$\left\langle l', m' \middle| \hat{L}_x \middle| l, m \right\rangle = \left\langle l', m' \middle| \frac{\hat{L}_+ + \hat{L}_-}{2} \middle| l, m \right\rangle \tag{11.6.29}$$

$$= \frac{\hbar}{2}\delta_{l,l'}(\delta_{m',m+1}\sqrt{(l-m)(l+m+1)} + \delta_{m',m-1}\sqrt{(l+m)(l-m+1)}) \tag{11.6.30}$$

and similarly:

$$\left\langle l', m' \middle| \hat{L}_y \middle| l, m \right\rangle = \left\langle l', m' \middle| \frac{\hat{L}_+ - \hat{L}_-}{2} \middle| l, m \right\rangle \tag{11.6.31}$$

$$= \frac{\hbar}{2}\delta_{l,l'}(\delta_{m',m+1}\sqrt{(l-m)(l+m+1)} - \delta_{m',m-1}\sqrt{(l+m)(l-m+1)}) \tag{11.6.32}$$

Instead, $\hat{L}_z$ and $\hat{L}^2$ are diagonal in this basis:

$$\left\langle l', m' \middle| \hat{L}_y \middle| l, m \right\rangle = \delta_{l,l'}\delta_{m,m'} m\hbar \tag{11.6.33}$$

and

$$\left\langle l', m' \middle| \hat{L}^2 \middle| l, m \right\rangle = \delta_{l,l'}\delta_{m,m'} l(l+1)\hbar \tag{11.6.34}$$

This fully solves the angular momentum eigen-problem from an operator approach.

**Example (Sh. 12.5.3.)**
Consider a state $|l,m\rangle$, calculate $\langle L_x\rangle, \langle L_y\rangle$, $\Delta L_x$, $\Delta L_y$ and verify the uncertainty principle.

**Solution** In a state $|l,m\rangle$ we find that:

$$\left\langle l, m \middle| \hat{L}_x \middle| l, m \right\rangle = \frac{\hbar}{2}\Big(\sqrt{(l-m)(l+m+1)}\,\langle l, m \,|\, l, m+1\rangle \tag{11.6.35}$$

$$+ \sqrt{(l+m)(l-m+1)}\,\langle l, m \,|\, l, m-1\rangle\Big) = 0 \tag{11.6.36}$$

and similarly:

$$\Big\langle l,m\Big|\,\hat{L}_y\,\Big|\,l,m\Big\rangle = \frac{\hbar}{2i}\Big(\sqrt{(l-m)(l+m+1)}\,\langle l,m\,|\,l,m+1\rangle \tag{11.6.37}$$

$$-\sqrt{(l+m)(l-m+1)}\,\langle l,m\,|\,l,m-1\rangle\,\Big) = 0 \tag{11.6.38}$$

Instead:

$$\Big\langle l,m\Big|\,\hat{L}_x^2\,\Big|\,l,m\Big\rangle = \frac{1}{4}\Big\langle l,m\Big|\,\hat{L}_+^2+\hat{L}_-^2+\hat{L}_+\hat{L}_-+\hat{L}_-\hat{L}_+\,\Big|\,l,m\Big\rangle \tag{11.6.39}$$

$$= \frac{1}{4}\Big\langle l,m\Big|\,2\hat{L}_+\hat{L}_- - [\hat{L}_+,\hat{L}_-]\,\Big|\,l,m\Big\rangle \tag{11.6.40}$$

$$= \frac{1}{4}\Big\langle l,m\Big|\,2\hat{L}_+\hat{L}_- - 2\hbar\hat{L}_z\,\Big|\,l,m\Big\rangle \tag{11.6.41}$$

$$= \frac{1}{4}\Big(-2m\hbar^2 + 2\Big\langle l,m\Big|\,\hat{L}_+\hat{L}_-\,\Big|\,l,m\Big\rangle\Big) \tag{11.6.42}$$

$$= \frac{\hbar^2}{2}\Big(\sqrt{(l+m)(l-m+1)(l-(m-1))(l+(m-1)+1)}-m\Big) \tag{11.6.43}$$

$$= \frac{\hbar^2}{2}\big((l+m)(l-m+1)-m\big) \tag{11.6.44}$$

$$= \frac{\hbar^2}{2}\big(l(l+1)-m^2\big) \tag{11.6.45}$$

Similarly:

$$\Big\langle l,m\Big|\,\hat{L}_y^2\,\Big|\,l,m\Big\rangle = -\frac{1}{4}\Big\langle l,m\Big|\,\hat{L}_+^2+\hat{L}_-^2-\hat{L}_+\hat{L}_--\hat{L}_-\hat{L}_+\,\Big|\,l,m\Big\rangle \tag{11.6.46}$$

$$= \frac{1}{4}\Big\langle l,m\Big|\,\hat{L}_+\hat{L}_-+\hat{L}_-\hat{L}_+\,\Big|\,l,m\Big\rangle \tag{11.6.47}$$

$$= \frac{\hbar^2}{2}\big(l(l+1)-m^2\big) \tag{11.6.48}$$

So:

$$\Delta L_x = \Delta L_y = \sqrt{\frac{\hbar^2}{2}\big(l(l+1)-m^2\big)} \tag{11.6.49}$$

and thus:

$$\Delta L_x \Delta L_y = \frac{\hbar^2}{2}\big(l(l+1)-m^2\big) \tag{11.6.50}$$

Instead:

$$\Big\langle[\hat{L}_x,\hat{L}_y]\Big\rangle = i\hbar\,\langle L_z\rangle = m\hbar^2 \tag{11.6.51}$$

so the uncertainty principle requires:

$$\Delta L_x\Delta L_y = \frac{\hbar^2}{2}\big(l(l+1)-m^2\big) \geq \frac{m\hbar^2}{2} \tag{11.6.52}$$

$$\Longleftrightarrow\; l(l+1) \geq m(m+1) \tag{11.6.53}$$

which is clearly verified since $|m| \leq l$. The uncertainty principle is saturated when equality holds, that is, if $|m| = l$, or $m = \pm l$, so in a state $|l, \pm l\rangle$. ◀

## 11.7 The eigen-problem for $\hat{\mathbf{L}}^2$ (functional)

In spherical coordinates, we can write the gradient operator as:

$$\nabla = \hat{\mathbf{e}}_r \frac{\partial}{\partial r} + \hat{\mathbf{e}}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\mathbf{e}}_\phi \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi} \tag{11.7.1}$$

and as always

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \tag{11.7.2}$$

since $\theta \to \phi$ when transforming from polar to spherical coordinates.

Also, some lengthy algebra gives:

$$\hat{L}_\pm = \hbar e^{\pm i\phi} \Big( \pm \frac{\partial}{\partial \theta} + i \cot\theta \frac{\partial}{\partial \phi} \Big) \tag{11.7.3}$$

and after some even lengthier algebra:

$$\hat{L}^2 = -\hbar^2 \Big[ \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \Big( \sin\theta \frac{\partial}{\partial \theta} \Big) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial^2 \phi} \Big] \tag{11.7.4}$$

Let us now define the $(l, m)$-spherical harmonic as $Y_{lm} = \langle \theta, \phi \,|\, l, m \rangle$. Then the eigenvalue equation for $\hat{L}_z$ becomes:

$$-i\hbar \frac{\partial Y_{lm(\theta,\phi)}}{\partial \phi} = m\hbar Y_{lm}(\theta, \phi) \implies Y_{lm}(\theta, \phi) = F(\theta) e^{im\phi} \tag{11.7.5}$$

where $F(\theta)$ is normalizable. To determine the latter, one could use the eigenvalue equation for $\hat{L}^2$:

$$\Big[ \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \Big( \sin\theta \frac{\partial}{\partial \theta} \Big) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial^2 \phi} \Big] F(\theta) = l(l+1) F(\theta) \tag{11.7.6}$$

However, it is easier if we instead consider the state of maximum magnetic number $|l, l\rangle$:

$$\hat{L}_+ |l, l\rangle = 0 \tag{11.7.7}$$

For this state we must have that

$$0 = \Big\langle \theta, \phi \Big| \hat{L}_+ \Big| l, l \Big\rangle = \hbar e^{i\phi} \Big( \frac{\partial}{\partial \theta} + i \cot\theta \frac{\partial}{\partial \phi} \Big) Y_{ll}(\theta, \phi) \tag{11.7.8}$$

and substituting $Y_{ll}(\theta, \phi) = F(\theta) e^{il\phi}$ we get that:

$$\hbar e^{i(l+1)\phi} \Big( \frac{\partial}{\partial \theta} - l \cot\theta \Big) F(\theta) = 0 \implies \frac{\partial F(\theta)}{\partial \theta} = l \cot\theta F(\theta) \tag{11.7.9}$$

Its general solution up to a normalization constant $C$ is:

$$F(\theta) = C\sin^l(\theta) \tag{11.7.10}$$

We can find the forms of $\hat{Y}_{lm}(\theta,\phi)$ by repeatedly applying $\hat{L}_-$ to $|l,l\rangle$. This provides us with:

$$\hat{Y}_{lm}(\theta,\phi) = C(\hat{L}_-)^{l-m}\Big(\sin^l\theta e^{il\phi}\Big) \tag{11.7.11}$$

$$= C\Big(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\Big)^{l-m}\Big(\sin^l\theta e^{il\phi}\Big) \tag{11.7.12}$$

In the mathematical methods volume, we have shown that $Y_{lm}(\theta,\phi)$ are called **spherical harmonics** and can be expressed as:

$$Y_{lm}(\theta,\phi) = (-1)^{m+|m|}\left[\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}\right]^{\frac{1}{2}}P_l^{|m|}(\cos\theta)e^{im\phi} \tag{11.7.13}$$

where we used the associated Legendre polynomials:

$$P_l^m(x) = \frac{(1-x^2)^{m/2}}{2^l l!}\frac{d^{m+l}}{dx^{m+l}}(x^2-1)^l \tag{11.7.14}$$

Because $|l,m\rangle$ forms a complete basis of the Hilbert space under study, we can assert that:

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}|l,m\rangle\langle l,m| = \mathbb{I} \tag{11.7.15}$$

so that

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}Y_{lm}^*(\theta,\phi)Y_{lm}(\theta,\phi) = \delta(\theta-\theta')\delta(\phi-\phi') \tag{11.7.16}$$

The spherical harmonics are also orthonormal, because $|l,m\rangle$ are too orthonormal:

$$\langle l,m\,|\,l',m'\rangle = \int\langle l,m\,|\,\theta,\phi\rangle\langle\theta,\phi\,|\,l',m'\rangle\,d\Omega \tag{11.7.17}$$

$$= \int Y_{lm}^*(\theta,\phi)Y_{l'm'}^*(\theta,\phi)d\Omega = \delta_{ll'}\delta_{mm'} \tag{11.7.18}$$

**Spherical harmonics**

Any $\psi(r,\theta,\phi)$ may be expanded in the spherical harmonics basis $|l,m\rangle$:

$$\psi(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}C_{lm}(r)Y_{lm}(\theta,\phi) \tag{11.7.19}$$

with:

$$C_{lm} = \int Y^*_{lm}(\theta,\phi)\psi(r,\theta,\phi)d\Omega \tag{11.7.20}$$

**Example (Sh. 12.5.13)**
Consider a particle in a state $\psi = N(x+y+2z)e^{-\alpha r}$ where $N$ is a normalization constant. What are the possible observed angular momenta of this system, and their corresponding probability?

**Solution** We can rewrite the spherical harmonics as:

$$Y_1^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\theta} \tag{11.7.21}$$

$$= \mp\sqrt{\frac{3}{8\pi}}\frac{\sqrt{x^2+y^2}}{r}e^{\pm i\arctan(y/x)} \tag{11.7.22}$$

$$= \mp\sqrt{\frac{3}{8\pi}}\frac{\sqrt{x^2+y^2}}{r}\frac{x\pm iy}{\sqrt{x^2+y^2}} \tag{11.7.23}$$

$$= \mp\sqrt{\frac{3}{8\pi}}\frac{\sqrt{x\pm iy}}{r} \tag{11.7.24}$$

Also:

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\frac{z}{r} \tag{11.7.25}$$

Therefore:

$$Y_1^1 + Y_1^{-1} - iY_1^1 + iY_1^{-1} = \sqrt{\frac{3}{4\pi}}\cdot\sqrt{2}\cdot\frac{x+y}{r} \tag{11.7.26}$$

and so:

$$x+y+2z = \sqrt{\frac{4\pi}{3}}\frac{r}{\sqrt{2}}(Y_1^1 + Y_1^{-1} - iY_1^1 + iY_1^{-1}) + 2\sqrt{4\pi}3rY_1^0 \tag{11.7.27}$$

The wave-function may then be written as:

$$\psi = Nr\sqrt{\frac{4\pi}{3}}\left(\frac{Y_1^1 + Y_1^{-1} - iY_1^1 + iY_1^{-1}}{\sqrt{2}} + 2Y_1^0\right)e^{-\alpha r} \tag{11.7.28}$$

Hence:

$$|\psi\rangle = N(r)\left(\frac{(1-i)}{\sqrt{2}}|1,1\rangle + \frac{(1+i)}{\sqrt{2}}|1,-1\rangle + 2|1,0\rangle\right) \tag{11.7.29}$$

where we included the radial part into the normalization constant $N(r)$. Hence, the

possible observed values of $l_z$ and their respective probabilities are:

$$P(l_z = \hbar) = \frac{|\langle 1,1 \,|\, \psi\rangle|^2}{\langle \psi \,|\, \psi\rangle} = \frac{|N(r)|^2}{6|N(r)|^2} = \frac{1}{6} \tag{11.7.30}$$

$$P(l_z = -\hbar) = \frac{|\langle 1,-1 \,|\, \psi\rangle|^2}{\langle \psi \,|\, \psi\rangle} = \frac{|N(r)|^2}{6|N(r)|^2} = \frac{1}{6} \tag{11.7.31}$$

$$P(l_z = -\hbar) = \frac{|\langle 1,0 \,|\, \psi\rangle|^2}{\langle \psi \,|\, \psi\rangle} = \frac{4|N(r)|^2}{6|N(r)|^2} = \frac{2}{3} \tag{11.7.32}$$

◀

## 11.8 Rotationally invariant Hamiltonians in 3D

Now that we have fully solved the eigenvalue problem of the angular momentum operator, we are ready to tackle 3-dimensional problems with rotational invariance, that is, hamiltonians $\hat{H}$ which commute with $\hat{L}^2$ and $\hat{L}_i$.

For such problems we must have a central potential, that is, $V(r,\theta,\phi) = V(r)$, so that the TISE reads:

$$\Big[-\frac{\hbar^2}{2\mu}\Big(\overbrace{\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \underbrace{\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}}_{-\frac{\hat{L}^2}{\hbar^2 r^2}}}^{\nabla^2}\Big) + V(r)\Big]\psi_E(r,\theta,\phi) = E\psi_E(r,\theta,\phi) \tag{11.8.1}$$

Now since $\hat{H}$, $\hat{L}_z$ and $\hat{L}^2$ all commute with each other, they are simultaneously diagonalizable. We may substitute the most general form of the eigenfunctions of $\hat{L}^2$ and $\hat{L}_z$ and label them with the quantum numbers $E, l, m$ relating to the energy, angular momentum, and $z$-angular momentum eigenvalues:

$$\psi_{Elm}(r,\theta,\phi) = R_{Elm}(r)Y_l^m(\theta,\phi) \tag{11.8.2}$$

We then find that (recall $\hat{L}^2 Y_l^m(\theta,\phi) = \hbar^2 l(l+1)Y_l^m(\theta,\phi)$):

$$\Big[-\frac{\hbar^2}{2\mu}\Big(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{l(l+1)}{r^2}\Big) + V(r)\Big]R_{El} = ER_{El} \tag{11.8.3}$$

We may now introduce the function $U_{El} = rR_{El}$. Note that:

$$\Big(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\Big)R_{El} = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\Big(\frac{U_{El}}{r}\Big) \tag{11.8.4}$$

$$= \frac{1}{r^2}\frac{\partial}{\partial r}r^2\Big(\frac{1}{r}\frac{\partial U_{El}}{\partial r} - \frac{U_{El}}{r^2}\Big) \tag{11.8.5}$$

$$= \frac{1}{r^2}\Big(\frac{\partial}{\partial r}\Big(r\frac{\partial U_{El}}{\partial r}\Big) - \frac{\partial U_{El}}{\partial r}\Big) \tag{11.8.6}$$

$$= \frac{1}{r}\frac{\partial^2 U_{El}}{\partial r^2} \tag{11.8.7}$$

so that (11.8.3) turns into:

$$-\frac{\hbar^2}{2\mu}\left(\frac{1}{r}\frac{\partial^2 U_{El}}{\partial r^2}-\frac{l(l+1)}{r^3}U_{El}\right)+V(r)\frac{U_{El}}{r}=E\frac{U_{El}}{r} \tag{11.8.8}$$

$$\implies -\frac{\hbar^2}{2\mu}\left(\frac{\partial^2 U_{El}}{\partial r^2}-\frac{l(l+1)}{r^2}U_{El}\right)+V(r)U_{El}=EU_{El} \tag{11.8.9}$$

or more suggestively:

**Radial equation for rotationally invariant hamiltonians**

If we define $\psi(r,\theta,\phi)=\frac{U_{El}(r)}{r}Y_l^m(\theta,\phi)$ then the TISE becomes:

$$\left[\frac{d^2}{dr^2}+\frac{2\mu}{\hbar^2}\Big(E-\underbrace{V(r)-\frac{l(l+1)\hbar^2}{2\mu r^2}}_{V_{\text{eff}}(r)}\Big)\right]U_{El}=0 \tag{11.8.10}$$

This equation resembles the form of a one-dimensional potential problem, with potential barrier:

$$V_{eff}(r)=V(r)+\frac{l(l+1)\hbar^2}{2\mu r^2} \tag{11.8.11}$$

We see that the additional $\frac{l(l+1)\hbar^2}{2\mu r^2}$ term, known as the centrifugal potential barrier for obvious reasons (classically it would be $\frac{L^2}{2\mu r^2}$, taking its divergence would yield the centrifugal force) is repulsive, since the potential increases as the distance decreases.

## 11.9 Boundary conditions and limiting behaviour

The appropriate boundary conditions for central potentials are not however identical to those for a 1-dimensional potential. We must firstly require that if we define

$$\hat{D}_l\equiv\frac{d^2}{dr^2}+\frac{2\mu}{\hbar^2}\Big(E-\underbrace{V(r)-\frac{l(l+1)\hbar^2}{2\mu r^2}}_{V_{eff}(r)}\Big) \tag{11.9.1}$$

then (11.8.10) turns into

$$\hat{D}_l U_{El}=EU_{El} \tag{11.9.2}$$

We argue on physical grounds that the energy eigenvalues must be real, thus imposing the requirement that $\hat{D}_l$ be Hermitian:

$$\int_0^\infty U_1^*\hat{D}_l U_2 dr=\int_0^\infty(\hat{D}_l U_1)^* U_2 dr \tag{11.9.3}$$

$$\iff \int_0^\infty U_1^*\hat{D}_l U_2 dr=\int_0^\infty(\hat{D}_l U_1)^* U_2 dr \tag{11.9.4}$$

for any two $U_1, U_2 = U_{El}$ for some $E, l$. We then get that:

$$\int_0^\infty U_1^* \frac{d^2U_2}{dr} dr = \int_0^\infty \frac{d^2U_1^*}{dr} U_2 dr \tag{11.9.5}$$

$$\implies \left[U_1^* \frac{dU_2}{dr}\right]_0^\infty - \int_0^\infty \frac{dU_1^*}{dr}\frac{dU_2}{dr} = \left[U_2 \frac{dU_1^*}{dr}\right]_0^\infty - \int_0^\infty \frac{dU_1^*}{dr}\frac{dU_2}{dr} \tag{11.9.6}$$

$$\implies \left[U_1^* \frac{dU_2}{dr} - U_2 \frac{dU_1^*}{dr}\right]_0^\infty = 0 \tag{11.9.7}$$

after integrating by parts.

We clearly also require that $R_{El}$ be normalizable

$$\int_0^\infty |R_{El}|^2 r^2 dr = \int_0^\infty |U_{El}|^2 dr = 1 \tag{11.9.8}$$

For bounded states, this requires the exponential decay behaviour:

$$U_{El} \to 0, \; r \to \infty \tag{11.9.9}$$

whereas for unbounded states we can have sinusoidal behaviour:

$$U_{El} \to e^{ikr}, \; r \to \infty \tag{11.9.10}$$

In both cases however, we still retrieve:

$$\left[U_1^* \frac{dU_2}{dr} - U_2 \frac{dU_1^*}{dr}\right]_\infty = 0 \tag{11.9.11}$$

so that:

$$\left[U_1^* \frac{dU_2}{dr} - U_2 \frac{dU_1^*}{dr}\right]_0 = 0 \tag{11.9.12}$$

This is only possible for asymptotic behaviour $U_{El} \to c$ as $r \to 0$. This constant however must be set to zero, or else we would get $\psi \sim \frac{c}{r} Y_0^0$. This is not an appropriate wave-function, since:

$$\nabla^2 \psi \sim -4\pi c Y_0^0 \delta^3(\mathbf{r}) \tag{11.9.13}$$

which is incompatible with the TISE unless the potential has a delta term too.

So we have that $U_{El} \to 0$ as $r \to 0$.

### Behaviour as $r \to 0$

Further information can be found by assuming $V(r)$ has terms of order larger than $r^{-2}$. Then, in the limit as $r \to 0$ we have that (11.8.10) is dominated by the centrifugal potential so that:

$$\frac{d^2U_l}{dr^2} \sim \frac{l(l+1)}{r^2} U_l \tag{11.9.14}$$

Note that we dropped the $E$ subscript since the equation no longer depends on the energy eigenvalue.

Using the ansatz $U_{El} \sim r^\alpha$ for some $\alpha$ then:

$$\alpha(\alpha-1) = l(l+1) \implies \alpha = l \text{ and } \alpha = l-1 \tag{11.9.15}$$

which produces the fundamental set of solutions as long as $l \neq 0$:

$$U_l \sim \begin{cases} r^{l+1} \\ r^{-l} \end{cases} \tag{11.9.16}$$

Since this describes the behaviour as $r \to 0$, we must reject the first solution. Therefore:

$$U_l \sim Dr^{l+1} \tag{11.9.17}$$

### Behaviour as $r \to \infty$

Let us now reconsider the limit as $r \to$, where the potential $V(r)$ is not dominant. We make one further assumption, that $rV(r) \to 0$ as $r \to \infty$. Then we get that the TISE turns into:

$$\frac{d^2U_E}{dr^2} = -\frac{2\mu E}{\hbar^2}U_E \tag{11.9.18}$$

where we again dropped the $l$ subscript this time due to its redundancy.

For unbounded states we have $E > 0$ so that the particle exhibits free-particle behaviour far from the origin:

$$U_E = Ae^{ikr} + Be^{-ikr},\ k = \sqrt{\frac{2\mu E}{\hbar^2}} \tag{11.9.19}$$

For bounded states we have $E < 0$ so that the particle exhibits exponential decay far from the origin:

$$U_E = Ce^{-kr},\ k = \sqrt{\frac{2\mu|E|}{\hbar^2}} \tag{11.9.20}$$

**Asymptotic behaviour of $U_{El}$**

For unbounded states $E > 0$ one should try the ansatz

$$U_{El} = r^{l+1}(Ae^{ikr} + Be^{-ikr})f(r)$$

For bounded states $E < 0$ one should try the ansatz

$$U_{El} = r^{l+1}e^{-kr}f(r)$$

## 11.10 3D Isotropic oscillator

# Spin angular momentum

12

## 12.1 What is spin?

The astute reader may have noticed an inconsistency in our development of angular momentum eigenvalues. Indeed, it seems like for half integer values of $l$, $\hat{L}_z$ can have half-integer eigenvalues, whereas in the preceding sections it was shown that $\hat{L}_z$ only has integer eigenvalues.

It turns out that the extra half-integer contribution to the eigenvalue spectrum is due to the fact that we never specified the form of $\hat{L}$, but only specified its commutation relation $\hat{\mathbf{L}} \times \hat{\mathbf{L}} = i\hbar\hat{\mathbf{L}}$. Doing so would have yielded integer eigenvalues as would be expected.

Due to this freedom, the results that we uncovered without assuming a specific form of the angular momenta operators apply to any form of $\psi$ no matter its nature. Had we assumed that $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$ we would have uncovered results only applicable to scalar wave-functions.

Indeed, it turns out that for some particles, the wave-function assumes a more complicated aspect, such as a vector field $\boldsymbol{\Psi} = \psi_x\hat{\mathbf{x}} + \psi_y\hat{\mathbf{y}} + \psi_z\hat{\mathbf{z}}$. For such wave-functions, rotations do not simply assign a new rotated value to each point in space.

In the scalar case $\delta\boldsymbol{\theta}$ assigns to each $(x, y, z)$ a new rotated point $(x', y', z')$.

In the vector case, however, $\delta\boldsymbol{\theta}$ assigns to each $(x, y, z)$ a new rotated point $(x', y', z')$, but must also rotate the vector at $(x, y, z)$ accordingly. The former is done by the original $\hat{\mathbf{L}}$ **orbital angular momentum**, whereas the former is done by a new vector operator $\hat{\mathbf{S}}$ called **spin angular momentum**. We define their sum as the **total angular momentum**:

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} \tag{12.1.1}$$

Hence, to be more correct one should have substituted $\hat{\mathbf{L}} \to \hat{\mathbf{J}}$ in section 9.5, where we made no mention of the nature of the wave-function. Hence:

$$\hat{\mathbf{J}} \times \hat{\mathbf{J}} = i\hbar\hat{\mathbf{J}} \tag{12.1.2}$$

where the similar result with $\hat{\mathbf{L}}$ only holds for scalar wave-functions.

So, to answer our question, the eigenvalues we found in section 9.6 are not eigenvalues of $\hat{L}_z$ but eigenvalues of $\hat{J}_z = \hat{L}_z + \hat{S}_z$, with $\hat{S}_z$ contributing to the half-integer term.

## 12.2 Spin matrices

Let us explore how to generalize our definition of the generator of rotations for vector fields.

Consider a vector field $\mathbf{\Psi}(x,y) = \psi_x(x,y)\hat{\mathbf{x}} + \psi_y(x,y)\hat{\mathbf{y}}$. When acted from an infinitesimal rotation $\delta\theta_z\hat{\mathbf{k}}$:

$$\psi'_x(x,y) = \psi_x(x + y\delta\theta_z, y - x\delta\theta_z) - \psi_y(x + y\delta\theta_z, y - x\delta\theta_z)\delta\theta_z \tag{12.2.1}$$
$$\psi'_y(x,y) = \psi_x(x + y\delta\theta_z, y - x\delta\theta_z)\delta\theta_z + \psi_y(x + y\delta\theta_z, y - x\delta\theta_z) \tag{12.2.2}$$

We can rewrite (12.2.1) to order $\delta\theta_z$ as:

$$\psi'_x(x,y) = \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\psi_x(x,y) - \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\psi_y(x,y)\delta\theta_z \tag{12.2.3}$$
$$= \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\psi_x(x,y) - \psi_y(x,y)\delta\theta_z \tag{12.2.4}$$

and similarly for (12.2.2):

$$\psi'_y(x,y) = \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\delta\theta_z\psi_x(x,y) - \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\psi_y(x,y) \tag{12.2.5}$$
$$= \delta\theta_z\psi_x(x,y) - \left(\mathbb{I} - \frac{i\delta\theta_z\hat{L}_z}{\hbar}\right)\psi_y(x,y) \tag{12.2.6}$$

These can be expressed more compactly in vector form as:

$$\begin{pmatrix}\psi'_x\\ \psi'_y\end{pmatrix} = \left[\begin{pmatrix}1&0\\0&1\end{pmatrix} - \frac{i\delta\theta_z}{\hbar}\begin{pmatrix}\hat{L}_z&0\\0&\hat{L}_z\end{pmatrix} - \frac{i\delta\theta_z}{\hbar}\begin{pmatrix}0&-i\hbar\\ i\hbar&0\end{pmatrix}\right]\begin{pmatrix}\psi_x\\\psi_y\end{pmatrix} \tag{12.2.7}$$

Generalizing to higher dimensions:

$$\begin{pmatrix}\psi'_1\\ \vdots\\ \psi'_n\end{pmatrix} = \left[\begin{pmatrix}1&0&\dots&0\\0&1&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\dots&1\end{pmatrix} - \frac{i\delta\theta_z}{\hbar}\begin{pmatrix}\hat{L}_z&0&\dots&0\\0&\hat{L}_z&\dots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\dots&\hat{L}_z\end{pmatrix} - \frac{i\delta\theta_z}{\hbar}\hat{S}_z\right]\begin{pmatrix}\psi_1\\\vdots\\\psi_n\end{pmatrix} \tag{12.2.8}$$

or more abstractly as:

$$|\psi\rangle = \left(\mathbb{I} - \frac{i\delta\theta_z}{\hbar}(\hat{L}_z + \hat{S}_z)\right)|\psi\rangle = \left(\mathbb{I} - \frac{i\delta\theta_z}{\hbar}\hat{J}_z\right)|\psi\rangle \tag{12.2.9}$$

consequently $\hat{J}_z$ is generator of rotations about the z-axis.

Let us try to define $\hat{S}_z$. Since $\hat{J}_i$ are generators of a rotation

$$[\hat{J}_i, \hat{J}_j] = i\hbar \sum_j \varepsilon_{ijk} \hat{J}_k \tag{12.2.10}$$

$\hat{J}_i$ and $\hat{S}_i$ commute since they act on different parts on the wave-function, so it follows that:

$$[\hat{L}_i, \hat{L}_j] + [\hat{S}_i, \hat{S}_j] = i\hbar \Big( \sum_k \varepsilon_{ijk} L_k + \sum_k \varepsilon_{ijk} S_k \Big) \tag{12.2.11}$$

$$\implies [\hat{S}_i, \hat{S}_j] = i\hbar \sum_k \varepsilon_{ijk} S_k \tag{12.2.12}$$

States of definite spin $s$ are described by a complex space of dimension $2s+1$, since the magnetic quantum number $m \in -s, -s+1, ..., 0, s-1, s$ has finitely many values it can take.

Working in the eigenbasis of $\hat{S}_z$, consider states of $\frac{1}{2}$-spin. Since $s = \frac{1}{2}$, it follows that $s_z = m = \pm\frac{1}{2}$. For sake of brevity we shall let $|\uparrow\rangle = |\frac{1}{2}, \frac{1}{2}\rangle$ and $|\downarrow\rangle = |\frac{1}{2}, -\frac{1}{2}\rangle$. A spin half system can then be described by:

$$|\psi\rangle = \alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{12.2.13}$$

satisfying the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Wave-functions made out of two components are called **spinors**, and the operators acting on them must therefore by 2×2 matrices.

We now define the ladder operators:

$$\hat{S}\pm = \hat{S}_x + i\hat{S}_y \tag{12.2.14}$$

Note that the problem of determining the eigenvalues of $\hat{S}$ is mathematically equivalent to that we worked out in the previous chapter (one could indeed apply these results assuming a state of pure spin angular momentum, and no orbital angular momentum) so:

$$\hat{S}_\pm\,|s, m\rangle = \sqrt{(s \mp m)(s \pm m + 1)}\hbar\,|s, m \pm 1\rangle \tag{12.2.15}$$

Returning to spin-$\frac{1}{2}$ systems note that $\hat{S}_+\,|\uparrow\rangle = 0 = \hat{S}_-\,|\downarrow\rangle$ since $|\uparrow\rangle$ and $|\downarrow\rangle$ cannot be raised/lowered. Instead, $\hat{S}_+\,|\downarrow\rangle = \hbar\,|\uparrow\rangle$ and $\hat{S}_-\,|\uparrow\rangle = \hbar\,|\downarrow\rangle$.

Then (12.2.15) gives:

$$\hat{S}_x |\uparrow\rangle = \frac{1}{2}(\hat{S}_+ + \hat{S}_-) |\uparrow\rangle = \frac{1}{2}(0 + \hbar |\downarrow\rangle) = \frac{\hbar}{2} |\downarrow\rangle \tag{12.2.16}$$

$$\hat{S}_x |\downarrow\rangle = \frac{1}{2}(\hat{S}_+ + \hat{S}_-) |\downarrow\rangle = \frac{1}{2}(\hbar |\uparrow\rangle + 0) = \frac{\hbar}{2} |\uparrow\rangle \tag{12.2.17}$$

$$\hat{S}_y |\uparrow\rangle = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-) |\uparrow\rangle = \frac{1}{2i}(0 - \hbar |\downarrow\rangle) = +\frac{i\hbar}{2} |\downarrow\rangle \tag{12.2.18}$$

$$\hat{S}_y |\downarrow\rangle = \frac{1}{2i}(\hat{S}_+ - \hat{S}_-) |\downarrow\rangle = \frac{1}{2i}(\hbar |\uparrow\rangle - 0) = -\frac{i\hbar}{2} |\uparrow\rangle \tag{12.2.19}$$

We can use the above four equation to construct the spin-$\frac{1}{2}$ matrix representation in the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis:

$$\hat{S}_x = \begin{pmatrix} \langle\uparrow| \hat{S}_x |\uparrow\rangle & \langle\uparrow| \hat{S}_x |\downarrow\rangle \\ \langle\downarrow| \hat{S}_x |\uparrow\rangle & \langle\downarrow| \hat{S}_x |\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{12.2.20}$$

and similarly:

$$\hat{S}_y = \begin{pmatrix} \langle\uparrow| \hat{S}_y |\uparrow\rangle & \langle\uparrow| \hat{S}_y |\downarrow\rangle \\ \langle\downarrow| \hat{S}_y |\uparrow\rangle & \langle\downarrow| \hat{S}_y |\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{12.2.21}$$

and finally:

$$\hat{S}_z = \begin{pmatrix} \langle\uparrow| \hat{S}_z |\uparrow\rangle & \langle\uparrow| \hat{S}_z |\downarrow\rangle \\ \langle\downarrow| \hat{S}_z |\uparrow\rangle & \langle\downarrow| \hat{S}_z |\downarrow\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{12.2.22}$$

We can extend our results to spin along any axis $\hat{\mathbf{n}}$ (which is equivalent to simply rotating our coordinate basis). We will denote by $|\hat{\mathbf{n}}, +\rangle$ the eigenstate giving spin up along this new axis, and by $|\hat{\mathbf{n}}, -\rangle$ the eigenstate giving spin down. In this notation: $|\uparrow\rangle = |\hat{\mathbf{z}}, +\rangle$ and $|\downarrow\rangle = |\hat{\mathbf{z}}, -\rangle$.

Our goal is to expand $|\hat{\mathbf{n}}, \pm\rangle$ in the $\hat{S}_z$ eigenbasis, so in the form of spinors:

$$|\hat{\mathbf{n}}, \pm\rangle = \begin{pmatrix} \psi_1^{\pm} \\ \psi_2^{\pm} \end{pmatrix} \tag{12.2.23}$$

Let us assume that $\hat{\mathbf{n}}$ points along $(\theta, \phi)$ so that:

$$\hat{n}_z = \cos\theta \tag{12.2.24}$$

$$\hat{n}_x = \sin\theta\cos\phi \tag{12.2.25}$$

$$\hat{n}_y = \sin\theta\sin\phi \tag{12.2.26}$$

then in follows that $|\hat{\mathbf{n}}, \pm\rangle$ are eigenvectors of $\hat{\mathbf{n}} \cdot \hat{\mathbf{S}}$:

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{S}} = n_x \hat{S}_x + n_y \hat{S}_y + n_z \hat{S}_z \tag{12.2.27}$$

and hence:

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{S}} = \frac{\hbar}{2} \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} \tag{12.2.28}$$

Finding the eigenvalues and eigenvectors of the above matrix is straightforward. Some algebra gives the characteristic polynomial as:

$$\lambda^2 = \frac{\hbar^2}{4} \implies \lambda_1 = \frac{\hbar}{2},\ \lambda_2 = -\frac{\hbar}{2} \tag{12.2.29}$$

which is a reassuring sign. We proceed by finding the eigenspinors :

$$\begin{cases} \cos\theta\psi_1^{\pm} + \sin\theta e^{-i\phi}\psi_2^{\pm} = \pm\psi_1 \\ \sin\theta e^{i\phi}\psi_1^{\pm} - \cos\theta\psi_2^{\pm} = \pm\psi_2 \end{cases} \tag{12.2.30}$$

which gives after some algebraic manipulation:

$$\frac{\psi_1^+}{\psi_2^+} = \cot\left(\frac{\theta}{2}\right)e^{-i\phi} \tag{12.2.31}$$

$$\frac{\psi_1^-}{\psi_2^-} = -\tan\left(\frac{\theta}{2}\right)e^{-i\phi} \tag{12.2.32}$$

These give:

$$|\hat{\mathbf{n}}, +\rangle = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right)e^{-i\phi/2} \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi/2} \end{pmatrix} \qquad |\hat{\mathbf{n}}, -\rangle = \begin{pmatrix} -\sin\left(\frac{\theta}{2}\right)e^{-i\phi/2} \\ \cos\left(\frac{\theta}{2}\right)e^{i\phi/2} \end{pmatrix} \tag{12.2.33a}$$

Another way do derive these equations is to realize that the spin operator $\mathbf{S}$ is the generator of the SU(2) Lie algebra, and thus it can be used to "rotate" states in spin space. More specifically, note that $e^{-i\boldsymbol{\theta}\cdot\mathbf{S}/\hbar}$ rotates a spin about the $\hat{\boldsymbol{\theta}}$ axis by $\theta$ (counter-clockwise). Consequently we have that

$$|\hat{\mathbf{n}}, +\rangle = e^{-i\phi S_z/\hbar}e^{-i\theta S_y/\hbar}\,|\uparrow\rangle \tag{12.2.34}$$

To evaluate the matrix exponentials we use the fact that

$$e^{-i\boldsymbol{\theta}\cdot\mathbf{S}/\hbar} = \cos\frac{\theta}{2}\mathbb{1} - i(\hat{\boldsymbol{\theta}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2} \tag{12.2.35}$$

which will be proven later, and find that

$$e^{-i\theta S_y/\hbar} = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix}, \quad e^{-i\phi S_z/\hbar} = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix} \tag{12.2.36}$$

Consequently

$$|\hat{\mathbf{n}}, +\rangle = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix} \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta/2e^{-i\phi/2} \\ \sin\theta/2e^{i\phi/2} \end{pmatrix} \tag{12.2.37}$$

and

$$|\hat{\mathbf{n}}, -\rangle = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix} \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta/2 e^{-i\phi/2} \\ \cos\theta/2 e^{i\phi/2} \end{pmatrix} \tag{12.2.38}$$

as found earlier. The advantage of this approach is that it can be generalised to any spin-s state.

We can also go the other way, instead of finding the eigenstate given a spin direction, given a spinor, we can find the associated spin direction for which it is an eigenstate. Consider some spinor:

$$|\psi\rangle = \begin{pmatrix} c_1 e^{i\phi_1} \\ c_2 e^{i\phi_2} \end{pmatrix} \tag{12.2.39}$$

where $c_1, c_2$ are real. We wish to find the direction $\hat{n}$ along which it is an eigenstate with eigenvalue $\frac{\hbar}{2}$.

From the normalization condition, $c_1^2 + c_2^2 = 1$ it follows that we can find some angle $\theta$ such that $c_1 = \cos\frac{\theta}{2}$ and $c_1 = \sin\frac{\theta}{2}$. If we pull out a phase factor $e^{i(\phi_1+\phi_2)/2}$ then:

$$|\psi\rangle = \begin{pmatrix} \cos\frac{\theta}{2} e^{i(\phi_1-\phi_2)/2} \\ \sin\frac{\theta}{2} e^{-i(\phi_1-\phi_2)/2} \end{pmatrix} \equiv |\hat{\mathbf{n}}, +\rangle \tag{12.2.40}$$

So overall, $\hat{n}$ points along $(\theta, \phi)$ with:

$$\theta = 2(\arctan\frac{\rho_2}{\rho_1} + n\pi),\ n = 0, 1, 2... \tag{12.2.41}$$

$$\phi = \frac{\phi_1 - \phi_2}{2} \tag{12.2.42}$$

## 12.3 Pauli matrices

The matrices (up to the constant $\frac{\hbar}{2}$) representing $\frac{1}{2}$-spin operators are called **Pauli matrices**, and are denoted as:

$$\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z) \implies \boxed{\hat{\mathbf{S}} = \frac{\hbar}{2}\boldsymbol{\sigma}} \tag{12.3.1}$$

These matrices anti-commute with each other, so that:

$$\{\sigma_i, \sigma_j\} = \delta_{ij},\ \implies [\sigma_i, \sigma_j] = 2\sigma_i\sigma_j \tag{12.3.2}$$

which can be verified through direct calculation.

Moreover, the commutation relation for $\hat{\mathbf{S}}$ can be expressed as:

$$\frac{\hbar^2}{4}[\sigma_i, \sigma_j] = \frac{i\hbar}{2}\sum_k \varepsilon_{ijk}\sigma_k \implies \sigma_i\sigma_j = i\sum_k \varepsilon_{ijk}\sigma_k \tag{12.3.3}$$

so that:

$$\sigma_x\sigma_y = i\sigma_z,\ \sigma_y\sigma_z = i\sigma_x,\ \sigma_z\sigma_x = i\sigma_y \tag{12.3.4}$$

Hence, we may combine all these identities into the commutation relation:

$$[\sigma_i, \sigma_j] = 2i\sum_k \varepsilon_{ijk}\sigma_k \tag{12.3.5}$$

The Pauli matrices are also traceless. Finally:

$$(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})^2 = \mathbb{I} \tag{12.3.6}$$

To prove this, note that since $\hat{S}_z$ has eigenvalues $\pm\frac{\hbar}{2}$, we may write:

$$\left(\hat{S}_z + \frac{\mathbb{I}\hbar}{2}\right)\left(\hat{S}_z - \frac{\mathbb{I}\hbar}{2}\right) = \mathbb{0} \tag{12.3.7}$$

Indeed, given an arbitrary spinor in this Hilbert space, $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ then:

$$\left(\hat{S}_z + \frac{\mathbb{I}\hbar}{2}\right)\left(\hat{S}_z - \frac{\mathbb{I}\hbar}{2}\right)|\psi\rangle = \left(\hat{S}_z + \frac{\mathbb{I}\hbar}{2}\right)(-\beta\hbar|\downarrow\rangle) = 0 \tag{12.3.8}$$

Our choice of the z-axis is arbitrary, so it must be true that:

$$\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{S}} + \frac{\mathbb{I}\hbar}{2}\right)\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{S}} - \frac{\mathbb{I}\hbar}{2}\right) = \mathbb{0} \tag{12.3.9}$$

so that:

$$(\hat{\mathbf{n}}\cdot\hat{\mathbf{S}})^2 = \frac{\hbar^2}{4}\mathbb{I} \tag{12.3.10}$$

Finally, we also establish the identity:

$$\boxed{(\mathbf{A}\cdot\boldsymbol{\sigma})(\mathbf{B}\cdot\boldsymbol{\sigma}) = (\mathbf{A}\cdot\mathbf{B})\mathbb{I} + i(\mathbf{A}\times\mathbf{B})\cdot\boldsymbol{\sigma}} \tag{12.3.11}$$

where $\mathbf{A}$ and $\mathbf{B}$ are vector operators commuting with the Pauli matrices. To prove this result, note that (using summation notation for simplicity):

$$\sigma_i\sigma_j = \frac{1}{2}([\sigma_i,\sigma_j] + \{\sigma_i,\sigma_j\}) = \delta_{ij}\mathbb{I} + i\varepsilon_{ijk}\sigma_k \tag{12.3.12}$$

Thus:

$$\begin{aligned}(\mathbf{A}\cdot\boldsymbol{\sigma})(\mathbf{B}\cdot\boldsymbol{\sigma}) = A_iB_j\sigma_i\sigma_j &= A_iB_j(\delta_{ij}\mathbb{I} + i\varepsilon_{ijk}\sigma_k) &(12.3.13)\\ &= A_iB_j\delta_{ij}\mathbb{I} + i\varepsilon_{ijk}A_iB_j\sigma_k &(12.3.14)\\ &= A_iB_j\delta_{ij}\mathbb{I} + i\varepsilon_{kij}A_iB_j\sigma_k &(12.3.15)\\ &= (\mathbf{A}\cdot\mathbf{B})\mathbb{I} + i(\mathbf{A}\times\mathbf{B})\cdot\sigma &(12.3.16)\end{aligned}$$

as desired.

Now note that the Pauli matrices together with the identity matrix span the set of operators

acting on spinors. Indeed, consider a general matrix $T$:

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{12.3.17}$$

Then, we need to find $a, b, c, d$ not all equal to zero such that

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + c \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + d \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c+d & a-ib \\ a+ib & -c+d \end{pmatrix} \tag{12.3.18}$$

Therefore:

$$\begin{cases} d+c=\alpha \\ d-c=\delta \\ a-ib=\beta \\ a+ib=\gamma \end{cases} \implies \begin{cases} a=\frac{1}{2}(\gamma+\beta) \\ b=\frac{1}{2i}(\gamma-\beta) \\ c=\frac{1}{2}(\alpha-\delta) \\ d=\frac{1}{2}(\alpha+\delta) \end{cases} \tag{12.3.19}$$

The only case where these could all be equal to zero is if $\alpha = \beta = \gamma = \delta = 0$, showing that these matrices are linearly independent.

So, if we have some operator $\hat{T}$, we can express it as:

$$\hat{T} = a\mathbb{I} + \boldsymbol{\beta} \cdot \boldsymbol{\sigma} \tag{12.3.20}$$

where $\boldsymbol{\beta} = (b, c, d)$ we found earlier. Alternatively, using the spin operator:

$$\hat{T} = a\mathbb{I} + \frac{2\boldsymbol{\beta}}{\hbar} \cdot \hat{\mathbf{S}} \tag{12.3.21}$$

so we see that any operator on $\mathcal{H}_s$, the space of spinors, can be written as a linear combination of the identity operator and the spin operator.

## 12.4 Spin rotation operator

Consider the spin rotation operator:

$$\hat{U}(\boldsymbol{\theta}) = \exp(-i\boldsymbol{\theta} \cdot \mathbf{S}/\hbar) = \exp\Big[ -i\Big(\frac{\theta}{2}\Big)\hat{\theta} \cdot \boldsymbol{\sigma} \Big] = \sum_{n=0}^{\infty} \Big( -\frac{i\theta}{2} \Big)^n \frac{1}{n!} (\hat{\theta} \cdot \boldsymbol{\sigma})^n \tag{12.4.1}$$

Now note that:

$$(\hat{\theta} \cdot \boldsymbol{\sigma})^2 = \mathbb{I} \implies (\hat{\theta} \cdot \boldsymbol{\sigma})^{2n} = \mathbb{I} \tag{12.4.2}$$

and similarly:

$$(\hat{\theta} \cdot \boldsymbol{\sigma})^{2n+1} = \mathbb{I}(\hat{\theta} \cdot \boldsymbol{\sigma}) = \hat{\theta} \cdot \boldsymbol{\sigma} \tag{12.4.3}$$

Therefore:

$$\hat{U}(\boldsymbol{\theta}) = \sum_{n=0}^{\infty} \left( -\frac{i\theta}{2} \right)^n \frac{1}{n!} (\hat{\theta} \cdot \boldsymbol{\sigma})^n \tag{12.4.4}$$

$$= \sum_{n,even}^{\infty} \frac{(-1)^n}{n!} \left( \frac{\theta}{2} \right)^n \mathbb{I} + \sum_{n,odd}^{\infty} \frac{(-1)^n}{n!} (-i) \left( \frac{\theta}{2} \right)^n \hat{\theta} \cdot \boldsymbol{\sigma} \tag{12.4.5}$$

$$= \cos \frac{\theta}{2} \mathbb{I} - i \sin \frac{\theta}{2} (\hat{\theta} \cdot \boldsymbol{\sigma}) \tag{12.4.6}$$

We therefore have that the spin rotation operator by some angle $\boldsymbol{\theta}$ is:

$$\hat{U}(\boldsymbol{\theta}) = \cos \frac{\theta}{2} \mathbb{I} - i \sin \frac{\theta}{2} (\hat{\theta} \cdot \boldsymbol{\sigma}) \tag{12.4.7}$$

Note an interesting fact, a rotation of a spin by $\theta = 2\pi$ about an arbitrary axis $\hat{\mathbf{n}}$ returns

$$\hat{U}(\pi \hat{\mathbf{n}}) = -\mathbb{1} \tag{12.4.8}$$

while a rotation by $\theta = 4\pi$ returns

$$\hat{U}(\pi \hat{\mathbf{n}}) = \mathbb{1} \tag{12.4.9}$$

Consequently, spin-$\frac{1}{2}$ particles must be rotated twice around themselves in order to return to their original orientation. One way to visualise this weird behaviour is using **Dirac's belt trick**

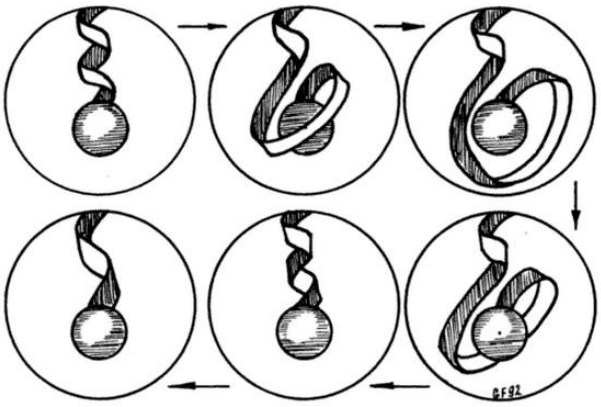


**Figure 12.1**. Dirac's belt trick shows how a belt with two rotations requires one to twist untwist itself, and thus two twists to go back to its original configuration.

## 12.5 Dynamics of spin

### Classical description

Recall from electromagnetism that the spin on a dipole moment $\boldsymbol{\mu}$ under the influence of a magnetic field $\mathbf{B}$ is given by:

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} \tag{12.5.1}$$

so that the interaction energy is:

$$\mathcal{H}_{dip} = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{12.5.2}$$

We may try to apply these results for a particle of mass $m$, charge $q$ in a circular orbit of radius $r$. Such a system would virtually look like a current:

$$I = \frac{q}{T} = \frac{qv}{2\pi r} \tag{12.5.3}$$

with magnetic moment (we use Gaussian units):

$$\mu = \frac{qv}{2\pi r} \cdot \frac{\pi r^2}{c} = \frac{qvr}{2c} = \frac{q}{2mc} mvr = \frac{ql}{2mc} \tag{12.5.4}$$

but $\boldsymbol{\mu}$ and $\mathbf{l}$ are parallel, and hence:

$$\boldsymbol{\mu} = \frac{q}{2mc}\mathbf{l} \tag{12.5.5}$$

We define the ratio of the magnetic moment to the angular momentum as the **gyromagnetic ratio** $\gamma$:

$$\gamma = \frac{q}{2mc} \tag{12.5.6}$$

Since the torque tends to align $\boldsymbol{\mu}$ along $\mathbf{B}$, the former will precess around the latter, just like a gyroscope under the influence of gravity. The angular equation of motion reads:

$$\boldsymbol{\tau} = \frac{d\mathbf{l}}{dt} = \boldsymbol{\mu} \times \mathbf{B} = \gamma(\mathbf{l} \times \mathbf{B}) \tag{12.5.7}$$

implying that in a small interval of time $\delta t$, the angular momentum vector moves by $\delta\mathbf{l}$ given by:

$$\delta\mathbf{l} = \gamma(\mathbf{l} \times \mathbf{B})\delta t \implies l \sin\theta \delta\phi = \delta l \tag{12.5.8}$$

so that the the tip of $\mathbf{l}$ precesses by an angle:

$$\delta\phi = -\gamma B \delta t \tag{12.5.9}$$

The precession frequency $\boldsymbol{\omega}_0$ is then:

$$\boxed{\boldsymbol{\omega}_0 = -\gamma \mathbf{B}} \tag{12.5.10}$$

### Quantum mechanical description

The electromagnetic hamiltonian for a particle of mass $m$, charge $q$ in a magnetic field is:

$$\hat{H} = \frac{(\mathbf{p} - q\mathbf{A}/c)^2}{2m} = \frac{|\mathbf{p}|^2}{2m} - \frac{q}{2mc}(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + \frac{q^2|\mathbf{A}|^2}{2mc^2} \tag{12.5.11}$$

If we consider a vector potential:

$$\mathbf{A} = \frac{B}{2}(-y\mathbf{i} + x\mathbf{j}) \implies \mathbf{B} = B\mathbf{k},\ B \text{ very small} \tag{12.5.12}$$

then the last term in (12.5.11) vanishes, being quadratic in $B$. Hence:

$$\langle \mathbf{r} \,|\, (\mathbf{p} \cdot \mathbf{A}) \,|\, \psi \rangle = -i\hbar \nabla \cdot (\mathbf{A}\psi) \tag{12.5.13}$$

$$= -i\hbar[(\nabla \cdot \mathbf{A})\psi + \mathbf{A} \cdot (\nabla \psi)] \tag{12.5.14}$$

but due to we can always apply gauge transformations for $\nabla \cdot \mathbf{A} = 0$ so that:

$$\langle \mathbf{r} \,|\, (\mathbf{p} \cdot \mathbf{A}) \,|\, \psi \rangle = -i\hbar \mathbf{A} \cdot (\nabla \psi) = \langle \mathbf{r} \,|\, \mathbf{A} \cdot \mathbf{p} \,|\, \psi \rangle \tag{12.5.15}$$

so that the middle term in (12.5.11) becomes:

$$\hat{H}_{int} = -\frac{q}{2mc}(2\mathbf{A} \cdot \mathbf{p}) = -\frac{qB}{2mc}(-\hat{y}\hat{p}_x + \hat{x}\hat{p}_y) = -\frac{q}{2mc}\mathbf{L} \cdot \mathbf{B} \tag{12.5.16}$$

We can define the quantum dipole moment operator $\boldsymbol{\mu}$

$$\boldsymbol{\mu} = \frac{q}{2mc}\mathbf{L} \implies \hat{H}_{int} = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{12.5.17}$$

where $\frac{q}{2mc}$ is known as the **gyromagnetic ratio**, it measures the ratio of the magnetic moment to the angular momentum of a particle. Alternatively, we can project this equation on the z-axis:

$$\mu_z = \frac{q\hbar}{2mc} \cdot n \;\text{ for } n = 0, 1, 2... \tag{12.5.18}$$

We define the quantity $\frac{q\hbar}{2mc}$ as the Bohr magneton, which for the electron has the value of:

$$\frac{e\hbar}{2mc} \approx 0.6 \times 10^{-8} eV/G \tag{12.5.19}$$

So we see that the orbital angular momentum is, quite expectedly, analogous to the standard angular momentum predicted by a classical description.

However, we have an additional contribution to the magnetic moment, called the spin magnetic moment, which has no analogue in classical mechanics and is purely quantum mechanical in nature.

Now note that the Pauli matrices together with the identity matrix span the space of 2×2 Hermitian matrices, so that any operator acting on spinors will be a linear combination of the identity operator and spin operator. In the case of the spin magnetic moment operator $\boldsymbol{\mu}_s$ (which is a vector operator), the identity gets washed out since it is a scalar operator, hence:

$$\boldsymbol{\mu}_s = \gamma \mathbf{S} \tag{12.5.20}$$

for some unknown constant $\gamma$. Let us write it as:

$$\boldsymbol{\mu}_s = -\frac{ge}{2mc}\mathbf{S} \tag{12.5.21}$$

where $g$ is the ratio of the spin and angular momentum magnetic moments. The fact that it is not equal to 1 shows that the spin of an electron cannot be attributes to just the rotation

of its mass about the $z$-axis, but is rather something more intrinsic. Now assume that:

$$\hat{H}_{int} = -\boldsymbol{\mu}_s \cdot \mathbf{B} = \frac{ge\hbar}{4mc}\boldsymbol{\sigma} \cdot \mathbf{B} \tag{12.5.22}$$

Experimentally, it was found that $g$ is very very close to 2. This is further supported by Quantum electrodynamics.

The propagator is then:

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} = e^{i\gamma t(\mathbf{S}\cdot\mathbf{B})t/\hbar} \tag{12.5.23}$$

which is surprisingly in the form of the rotation operator by an angle:

$$\boldsymbol{\theta}(t) = -\gamma \mathbf{B} t \tag{12.5.24}$$

giving a precession frequency:

$$\boldsymbol{\omega}_0 = -\gamma \mathbf{B} \tag{12.5.25}$$

For example, if we consider a magnetic field $\mathbf{B} = B\mathbf{k}$:

$$\hat{U}(t) = \exp(i\gamma t S_z B/\hbar) = \exp(i\omega_0 t\sigma_z/2) \tag{12.5.26}$$

so that:

$$\hat{U}(t) = \begin{pmatrix} e^{i\omega_0 t/2} & 0 \\ 0 & e^{-i\omega_0 t/2} \end{pmatrix} \tag{12.5.27}$$

For an electron in the state $|\hat{n}, +\rangle$ for example:

$$|\psi(0)\rangle = \begin{pmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{pmatrix} \tag{12.5.28}$$

so that after some time $t$:

$$|\psi(t)\rangle = \begin{pmatrix} \cos\frac{\theta}{2}e^{-i(\phi-\omega_0 t)/2} \\ \sin\frac{\theta}{2}e^{i(\phi-\omega_0 t)/2} \end{pmatrix} \tag{12.5.29}$$

the state has rotated in the Hilbert space with angular frequency $\omega_0$.

## 12.6 Paramagnetic resonance (long way)

Consider a spin-$\frac{1}{2}$ particle starting out in the state $|\psi(0)\rangle = |\uparrow\rangle$ interacting with the Hamiltonian:

$$\hat{H} = -\gamma \mathbf{B} \cdot \hat{\mathbf{S}} \tag{12.6.1}$$

where:

$$\mathbf{B} = B_0\mathbf{z} + B(\cos\omega t\mathbf{x} - \sin\omega t\mathbf{y}) \tag{12.6.2}$$

is the magnetic field, composed of a vertical component, and a clockwise rotating component.

We want to find an expression for the state of the particle at a later time $t$:

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle \tag{12.6.3}$$

Since the Hamiltonian is time-dependent, it is quite difficult to directly calculate the propagator $\hat{U}(t)$. However, we can investigate this system in a frame rotating (with respect to lab frame) with the same frequency $\boldsymbol{\omega} = -\omega\mathbf{z}$ as the magnetic field. Mathematically, this is equivalent as rotating the state of the system by $\boldsymbol{\omega} = \omega\mathbf{z}$. In this frame the state of the system is described by:

$$|\chi(t)\rangle = e^{-i\omega t\hat{S}_z/\hbar}\,|\psi(t)\rangle \tag{12.6.4}$$

The TDSE in the lab frame reads:

$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = -\gamma\mathbf{B}\cdot\hat{\mathbf{S}}\,|\psi(t)\rangle \tag{12.6.5}$$

$$\implies i\hbar\frac{\partial}{\partial t}(e^{i\omega t\hat{S}_z/\hbar}\,|\chi(t)\rangle\,) = -\gamma\mathbf{B}\cdot\hat{\mathbf{S}}(e^{i\omega t\hat{S}_z/\hbar}\,|\chi(t)\rangle\,) \tag{12.6.6}$$

$$\implies i\hbar\left(\frac{i\omega\hat{S}_z}{\hbar}e^{i\omega t\hat{S}_z/\hbar} + e^{i\omega t\hat{S}_z/\hbar}\frac{\partial}{\partial t}\right)|\chi(t)\rangle = -\gamma\mathbf{B}\cdot\hat{\mathbf{S}}(e^{i\omega t\hat{S}_z/\hbar}\,|\chi(t)\rangle\,) \tag{12.6.7}$$

$$\implies i\hbar e^{i\omega t\hat{S}_z/\hbar}\frac{\partial}{\partial t}\,|\chi(t)\rangle = (\omega\hat{S}_z e^{i\omega t\hat{S}_z/\hbar} - \gamma\mathbf{B}\cdot\hat{\mathbf{S}}e^{i\omega t\hat{S}_z/\hbar})\,|\chi(t)\rangle \tag{12.6.8}$$

We can now multiply by $e^{-i\omega t\hat{S}_z/\hbar}$ to the left so that:

$$i\hbar\frac{\partial}{\partial t}\,|\chi(t)\rangle = (\omega\hat{S}_z - e^{-i\omega t\hat{S}_z/\hbar}\gamma(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar})\,|\chi(t)\rangle \tag{12.6.9}$$

where we used the fact that $[\hat{A}, e^{ic\hat{A}}] = 0$ for any constant $c$.

We then see that the rotated ket $|\chi(t)\rangle$ also satisfies the TDSE with an effective Hamiltonian:

$$\hat{H}_{\text{eff}} = \omega\hat{S}_z - e^{-i\omega t\hat{S}_z/\hbar}\gamma(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar} \tag{12.6.10}$$

Let us try to simplify this expression. Firstly:

$$e^{-i\omega t\hat{S}_z/\hbar}(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar} = (\cos\frac{\omega t}{2}\mathbb{I} - i\sin\frac{\omega t}{2}\hat{\sigma}_z)(\mathbf{B}\cdot\hat{\mathbf{S}})(\cos\frac{\omega t}{2}\mathbb{I} + i\sin\frac{\omega t}{2}\hat{\sigma}_z) \tag{12.6.11}$$

$$= \cos^2\frac{\omega t}{2}(\mathbf{B}\cdot\hat{\mathbf{S}}) + i\sin\frac{\omega t}{2}\cos\frac{\omega t}{2}((\mathbf{B}\cdot\hat{\mathbf{S}})\hat{\sigma}_z - \hat{\sigma}_z(\mathbf{B}\cdot\hat{\mathbf{S}})) + \sin\frac{\omega t}{2}\hat{\sigma_z}(\mathbf{B}\cdot\hat{\mathbf{S}})\hat{\sigma_z} \tag{12.6.12}$$

Now:

$$((\mathbf{B}\cdot\hat{\mathbf{S}})\hat{\sigma}_z = \frac{\hbar}{2}(\mathbf{B}\cdot\boldsymbol{\sigma})(\mathbf{z}\cdot\boldsymbol{\sigma}) \tag{12.6.13}$$

$$= \frac{\hbar}{2}(B_z + i(\mathbf{B}\times\mathbf{z})\cdot\boldsymbol{\sigma}) \tag{12.6.14}$$

and similarly:

$$\hat{\sigma}_z((\mathbf{B}\cdot\hat{\mathbf{S}}) = \frac{\hbar}{2}(\mathbf{z}\cdot\boldsymbol{\sigma})(\mathbf{B}\cdot\boldsymbol{\sigma}) \tag{12.6.15}$$

$$= \frac{\hbar}{2}(B_z - i(\mathbf{B}\times\mathbf{z})\cdot\boldsymbol{\sigma}) \tag{12.6.16}$$

so that:

$$(\mathbf{B}\cdot\hat{\mathbf{S}})\hat{\sigma}_z - \hat{\sigma}_z(\mathbf{B}\cdot\hat{\mathbf{S}}) = 2i(B_y\hat{S}_x - \hat{B}_x\hat{S}_y) \tag{12.6.17}$$

Also, using the fact that the square of any Pauli matrix is zero:

$$\hat{\sigma}_z(\mathbf{B}\times\hat{\mathbf{S}})\hat{\sigma}_z = B_x\hat{\sigma}_z\hat{S}_x\hat{\sigma}_z + B_y\hat{\sigma}_z\hat{S}_y\hat{\sigma}_z + B_z\hat{\sigma}_z\hat{S}_z\hat{\sigma}_z \tag{12.6.18}$$

$$= B_z\hat{S}_z - iB_x\hat{\sigma}_z\hat{S}_y + iB_y\hat{\sigma}_z\hat{S}_x \tag{12.6.19}$$

$$= B_z\hat{S}_z - B_x\hat{S}_x - B_y\hat{S}_y \tag{12.6.20}$$

So 11.6.12 turns into:

$$\cos^2\frac{\omega t}{2}(\mathbf{B}\cdot\hat{\mathbf{S}}) - \sin\omega t(B_y\hat{S}_x - \hat{B}_x\hat{S}_y) + \sin^2\frac{\omega t}{2}(B_z\hat{S}_z - B_x\hat{S}_x - B_y\hat{S}_y) \tag{12.6.21}$$

Thus:

$$e^{-i\omega t\hat{S}_z/\hbar}(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar} = \hat{S}_x(B_x\cos^2\frac{\omega t}{2} - B_y\sin\omega t - \sin^2\frac{\omega t}{2}B_x) \tag{12.6.22}$$

$$+ \hat{S}_y(B_y\cos^2\frac{\omega t}{2} + B_x\sin\omega t - \sin^2\frac{\omega t}{2}B_y) \tag{12.6.23}$$

$$+ \hat{S}_z(B_z\cos^2\frac{\omega t}{2} + \sin^2\frac{\omega t}{2}B_z) \tag{12.6.24}$$

and hence:

$$e^{-i\omega t\hat{S}_z/\hbar}(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar} = \hat{S}_x(B_x\cos\omega t - B_y\sin\omega t) + \hat{S}_y(B_y\cos\omega t + B_x\sin\omega t) + B_z\hat{S}_z \tag{12.6.25}$$

Substituting back $B_x = B\cos\omega t$ and $B_y = -B\sin\omega t$ then:

$$e^{-i\omega t\hat{S}_z/\hbar}(\mathbf{B}\cdot\hat{\mathbf{S}})e^{i\omega t\hat{S}_z/\hbar} = B\hat{S}_x + B_0\hat{S}_z \tag{12.6.26}$$

Hence, finally, we find that the effective hamiltonian takes the form:

$$\hat{H}_{\text{eff}} = \omega\hat{S}_z - (B\hat{S}_x + B_0\hat{S}_z) \tag{12.6.27}$$

and thus:

$$\hat{H}_{\text{eff}} = -\gamma\underbrace{\Big(B\mathbf{x} + \Big(B_0 - \frac{\omega}{\gamma}\Big)\mathbf{z}\Big)}_{\mathbf{B}_{\text{eff}}}\cdot\hat{\mathbf{S}} \tag{12.6.28}$$

where the effective magnetic field $\mathbf{B}_{\text{eff}}$ is:

$$\mathbf{B}_{\text{eff}} = B\mathbf{x} + \Big(B_0 - \frac{\omega}{\gamma}\Big)\mathbf{z} \tag{12.6.29}$$

The effective propagator in the rotating frame is thus:

$$\hat{U}_{\text{eff}}(t) = e^{-i\hat{H}_{\text{eff}}t/\hbar} = \exp\Big(\frac{i\gamma(\mathbf{B}_{\text{eff}}\cdot\hat{\mathbf{S}})t}{\hbar}\Big) \tag{12.6.30}$$

This is simply a rotation by $\boldsymbol{\theta}(t) = -\gamma \mathbf{B}_{\text{eff}} t$, so using (12.4.7):

$$\hat{U}_{\text{eff}}(t) = \cos\frac{\gamma B_{\text{eff}} t}{2}\mathbb{1} + i\sin\frac{\gamma B_{\text{eff}} t}{2}(\hat{B}_{\text{eff}}\cdot\boldsymbol{\sigma}) \tag{12.6.31}$$

In the $\hat{S}_z$ basis, the we can use the expressions for the pauli matrices to write:

$$\hat{U}_{\text{eff}}(t) \leftrightarrow \cos\frac{\gamma B_{\text{eff}} t}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\frac{1}{B_{\text{eff}}}(\mathbf{B}_{\text{eff}}\cdot\boldsymbol{\sigma}) \tag{12.6.32}$$

Now:

$$\mathbf{B}_{\text{eff}}\cdot\boldsymbol{\sigma} = \begin{pmatrix}B_0 - \frac{\omega}{\gamma} & 0\\ 0 & -\left(B_0 - \frac{\omega}{\gamma}\right)\end{pmatrix} + \begin{pmatrix}0 & B\\ B & 0\end{pmatrix} \tag{12.6.33}$$

so:

$$\mathbf{B}_{\text{eff}}\cdot\boldsymbol{\sigma} = \begin{pmatrix}B_0 - \frac{\omega}{\gamma} & B\\ B & -\left(B_0 - \frac{\omega}{\gamma}\right)\end{pmatrix} \tag{12.6.34}$$

Therefore:

$$\hat{U}_{\text{eff}}(t) \leftrightarrow \begin{pmatrix}\cos\frac{\gamma B_{\text{eff}} t}{2} & 0\\ 0 & \cos\frac{\gamma B_{\text{eff}} t}{2}\end{pmatrix} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\begin{pmatrix}\frac{B_0-\omega/\gamma}{B_{\text{eff}}} & \frac{B}{B_{\text{eff}}}\\ \frac{B}{B_{\text{eff}}} & -\frac{B_0-\omega/\gamma}{B_{\text{eff}}}\end{pmatrix} \tag{12.6.35}$$

Now the state $|\uparrow\rangle$ in the rotating frame is given by:

$$|\chi(0)\rangle \leftrightarrow \begin{pmatrix}e^{-i\omega t/2}\\ 0\end{pmatrix} \tag{12.6.36}$$

so that:

$$|\chi(t)\rangle = \left[\begin{pmatrix}\cos\frac{\gamma B_{\text{eff}} t}{2} & 0\\ 0 & \cos\frac{\gamma B_{\text{eff}} t}{2}\end{pmatrix} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\begin{pmatrix}\frac{B_0-\omega/\gamma}{B_{\text{eff}}} & \frac{B}{B_{\text{eff}}}\\ \frac{B}{B_{\text{eff}}} & -\frac{B_0-\omega/\gamma}{B_{\text{eff}}}\end{pmatrix}\right]\begin{pmatrix}e^{-i\omega t/2}\\ 0\end{pmatrix} \tag{12.6.37}$$

$$= \begin{pmatrix}\cos\frac{\gamma B_{\text{eff}} t}{2}e^{-i\omega t/2}\\ 0\end{pmatrix} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\begin{pmatrix}\frac{B_0-\omega/\gamma}{B_{\text{eff}}}e^{-i\omega t/2}\\ \frac{B}{B_{\text{eff}}}e^{-i\omega t/2}\end{pmatrix} \tag{12.6.38}$$

$$= e^{-i\omega t/2}\begin{pmatrix}\cos\frac{\gamma B_{\text{eff}} t}{2} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\frac{B_0-\omega/\gamma}{B_{\text{eff}}}\\ \frac{iB}{B_{\text{eff}}}\sin\frac{\omega B_{\text{eff}}}{2}\end{pmatrix} \tag{12.6.39}$$

Finally, transforming back to the lab frame:

$$|\psi(t)\rangle = e^{-i\omega t/2}\begin{pmatrix}\left[\cos\frac{\gamma B_{\text{eff}} t}{2} + i\sin\frac{\gamma B_{\text{eff}} t}{2}\frac{B_0-\omega/\gamma}{B_{\text{eff}}}\right]e^{i\omega t/2}\\ \frac{iB}{B_{\text{eff}}}\sin\frac{\omega B_{\text{eff}}}{2}e^{-i\omega t/2}\end{pmatrix} \tag{12.6.40}$$

Note that for the interesting case where $B_0 = \frac{\omega}{\gamma}$, this state corresponds to $|\mathbf{n}(\theta(t),\phi(t)),+\rangle$

with:

$$\theta(t) = \gamma B_{\text{eff}} t, \quad \text{and} \quad \phi(t) = \gamma B_0 t \tag{12.6.41}$$

What about the spin magnetic moment? We can evaluate its expectation value:

$$\langle \hat{\mu}_z(t) \rangle = \gamma \left\langle \psi(t) \middle| \hat{S}_z \middle| \psi(t) \right\rangle \tag{12.6.42}$$

$$= \gamma \begin{pmatrix} \left[ \cos \frac{\gamma B_{\text{eff}} t}{2} - i \sin \frac{\gamma B_{\text{eff}} t}{2} \frac{B_0 - \omega/\gamma}{B_{\text{eff}}} \right] e^{-i\omega t/2} \\ -\frac{iB}{B_{\text{eff}}} \sin \frac{\omega B_{\text{eff}}}{2} e^{i\omega t/2} \end{pmatrix}^T \begin{pmatrix} \left[ \cos \frac{\gamma B_{\text{eff}} t}{2} + i \sin \frac{\gamma B_{\text{eff}} t}{2} \frac{B_0 - \omega/\gamma}{B_{\text{eff}}} \right] e^{i\omega t/2} \\ -\frac{iB}{B_{\text{eff}}} \sin \frac{\omega B_{\text{eff}}}{2} e^{-i\omega t/2} \end{pmatrix} \tag{12.6.43}$$

$$= \frac{\gamma \hbar^2}{4} \left( \cos^2 \frac{\gamma B_{\text{eff}} t}{2} + \sin^2 \frac{\gamma B_{\text{eff}} t}{2} \left( \frac{B_0 - \omega/\gamma}{B_{\text{eff}}} \right)^2 - \frac{B^2}{B_{\text{eff}}^2} \sin^2 \frac{\gamma B_{\text{eff}} t}{2} \right) \tag{12.6.44}$$

Resonance occurs when $B_0 = \frac{\omega}{\gamma}$, where $B_{\text{eff}} = B_0$ and:

$$\langle \hat{\mu}_z(t) \rangle = \cos \gamma B_0 t \tag{12.6.45}$$

so the magnetic spin moment precesses with frequency $\gamma B_0$.

We can prove this precession more generally. If $\hat{H} = -\gamma \mathbf{L} \cdot \mathbf{B}$, and $\mathbf{B}$ is homogeneous (in space), then using Ehrenfest's theorem:

$$\frac{d\left\langle \hat{\mathbf{L}} \right\rangle}{dt} = \frac{1}{i\hbar} \left\langle [\hat{\mathbf{L}}, \hat{H}] \right\rangle = \frac{1}{i\hbar} \left\langle [\hat{\mathbf{L}}, -\gamma \hat{\mathbf{L}} \cdot \mathbf{B}] \right\rangle = \left\langle \sum_i \frac{i\gamma B_i}{\hbar} ([\hat{\mathbf{L}}, \hat{L}_i]) \right. \tag{12.6.46}$$

$$= \left\langle \sum_i \frac{i\gamma B_i}{\hbar} \sum_j [\hat{L}_j, \hat{L}_i] \mathbf{e}_j \right\rangle = \left\langle \sum_i \frac{i\gamma B_i}{\hbar} \sum_{j,k} \varepsilon_{jik} i\hbar \hat{L}_k \mathbf{e}_j \right\rangle \tag{12.6.47}$$

$$= \left\langle \sum_i -\gamma B_i (\mathbf{e}_i \times \hat{\mathbf{L}}) \right\rangle = \left\langle (\gamma \hat{\mathbf{L}}) \times \mathbf{B} \right\rangle \tag{12.6.48}$$

$$\implies \boxed{\frac{d\left\langle \hat{\mathbf{L}} \right\rangle}{dt} = \langle \hat{\boldsymbol{\mu}} \rangle \times \mathbf{B}} = \langle \boldsymbol{\tau} \rangle \tag{12.6.49}$$

so we see that the expectation values do indeed follow the classical precession equation.

## 12.7 Paramagnetic resonance and Rabi oscillations (quick way)

Having sloshed through the horrendous algebra of the previous section, let's now try to work out the same problem in a more natural picture. Indeed in the previous section we moved to a frame which was rotating with frequency $\omega$, which resulting in transforming the Hamiltonian into a time-independent form. This change of basis coincides with the Interaction picture we discussed in Chapter 3, so let's try to figure out the equations of motion in this picture.

In the Schrödinger picture, the Hamiltonian can be written as

$$\hat{H} = H_0 + V(t) \tag{12.7.1}$$
$$H_0 = -\gamma(B_0 - B')\hat{S}_z, \tag{12.7.2}$$
$$V(t) = -\gamma B_\perp(\cos(\omega t)\hat{S}_x + \sin(\omega t)\hat{S}_y) - \gamma B'\hat{S}_z \tag{12.7.3}$$

where $B'$ is some unknown constant which we will figure out soon. In the Interaction picture, we find that

$$|\psi_I(t)\rangle = e^{iH_0 t}|\psi_S(t)\rangle = e^{-i\gamma(B_0-B')t\sigma_z/2},\ |\psi_I(t)\rangle = e^{-iV_I(t)t/\hbar}|\psi(0)\rangle \tag{12.7.4}$$

which corresponds to moving to a frame rotating with frequency $\gamma(B_0 - B')/2$ about the $z - axis$ and

$$V_I(t) = e^{i\gamma(B_0-B')t\sigma_z/2}V(t)e^{-i\gamma(B_0-B')t\sigma_z/2} \tag{12.7.5}$$

We then find that

$$e^{i\gamma(B_0-B')t\sigma_z/2} = \begin{pmatrix} e^{i\gamma(B_0-B')t/2} & 0 \\ 0 & e^{-i\gamma(B_0-B')t/2} \end{pmatrix},\ \text{and}\ V(t) = \begin{pmatrix} B' & B_\perp e^{-i\omega t} \\ B_\perp e^{i\omega t} & -B' \end{pmatrix} \tag{12.7.6}$$

which when substituted into (12.7.5) yields

$$V_I(t) = -\gamma\frac{\hbar}{2}\begin{pmatrix} e^{i\gamma(B_0-B')t/2} & 0 \\ 0 & e^{-i\gamma(B_0-B')t/2} \end{pmatrix}\begin{pmatrix} B' & B_\perp e^{-i\omega t} \\ B_\perp e^{i\omega t} & -B' \end{pmatrix}\begin{pmatrix} e^{-i\gamma(B_0-B')t/2} & 0 \\ 0 & e^{i\gamma(B_0-B')t/2} \end{pmatrix} \tag{12.7.7}$$

$$= -\gamma\frac{\hbar}{2}\begin{pmatrix} e^{i\gamma(B_0-B')t/2} & 0 \\ 0 & e^{-i\gamma(B_0-B')t/2} \end{pmatrix}\begin{pmatrix} B'e^{-i\gamma(B_0-B')t/2} & B_\perp e^{-i\omega t}e^{i\gamma(B_0-B')t/2} \\ B_\perp e^{i\omega t}e^{-i\omega(B_0-B')t/2} & -B'e^{i\gamma(B_0-B')t/2} \end{pmatrix} \tag{12.7.8}$$

$$= -\gamma\frac{\hbar}{2}\begin{pmatrix} B' & B_\perp e^{i\omega t}e^{i\gamma(B_0-B')t} \\ B_\perp e^{-i\omega t}e^{-i\gamma(B_0-B')t} & -B' \end{pmatrix} \tag{12.7.9}$$

If we let $B' = B_0 - \frac{\omega}{\gamma}$ then we see that

$$V_I(t) = -\gamma\left(\left(B_0 - \frac{\omega}{\gamma}\right)\hat{S}_z + B_\perp\hat{S}_x\right) \tag{12.7.10}$$

so the interaction picture interaction is time-independent. Letting the effective magnetic field be

$$\mathbf{B}_{\text{eff}} = B_\perp\mathbf{z} + \left(B_0 - \frac{\omega}{\gamma}\right)\mathbf{z} \implies V_I(t) = -\gamma\mathbf{B}_{\text{eff}}\cdot\hat{\mathbf{S}} \tag{12.7.11}$$

then it follows that the interaction picture propagator is

$$U_I(t) = e^{i\gamma\mathbf{B}_{\text{eff}}\cdot\boldsymbol{\sigma}t/2} = \cos\frac{\gamma B_{\text{eff}}t}{2}\mathbb{1} + i\frac{B_\perp\sigma_x + (B_0 - \omega/\gamma)\sigma_z}{B_{\text{eff}}}\sin\frac{\gamma B_{\text{eff}}t}{2} \tag{12.7.12}$$

This matches exactly the result (12.6.35) we got in the previous section (just with a lot less matrix algebra required). Hopefully this is a good showcase of how choosing the right

picture to work in can vastly simplify a problem, this will also happen when discussing time-dependent perturbation theory, scattering etc... Suppose we start out in a the state $|\uparrow\rangle$. Then we see that the state at time $t$ in the interaction picture is given by

$$|\psi_I(t)\rangle = \cos\frac{\gamma B_{\text{eff}} t}{2}|\uparrow\rangle + i\frac{B_\perp|\downarrow\rangle + (B_0 - \omega/\gamma)|\uparrow\rangle}{B_{\text{eff}}}\sin\frac{\gamma B_{\text{eff}} t}{2} \tag{12.7.13}$$

If we are interested in the transition amplitude $P_{\uparrow,\downarrow}(t)$ then

$$P_{\uparrow,\downarrow}(t) = \frac{B_\perp^2}{B_{\text{eff}}^2}\sin^2\left(\frac{\gamma B_{\text{eff}} t}{2}\right),\ B_{\text{eff}} = \sqrt{\left(B_0 - \frac{\omega}{\gamma}\right)^2 + B_\perp^2} \tag{12.7.14}$$

The oscillations are known as **Rabi oscillations** which have frequency $2\Omega$ where

$$\Omega = \gamma\sqrt{\left(B_0 - \frac{\omega}{\gamma}\right)^2 + B_\perp^2} \tag{12.7.15}$$

is known as the **Rabi frequency**. Note that when $\omega = \gamma B_0$ then the transition amplitude peaks to $P = 1$ when $t = \frac{\pi}{\gamma B_\perp}$ and we achieve paramagnetic resonance. The transition is assured to happen periodically.

## 12.8 The Ammonia molecule

Although our discussion of Rabi oscillations originated from the dynamics of a spin in a rotating magnetic field, our discussion extends equally well to any two-state system. One important example is ammonia.

Ammonia $NH_3$ is a molecule which takes a trigonal pyramidal shape due to the presence of a lone electron pair on the nitrogen atom. Due to the symmetry of this configuration, the ammonia molecule can be seen as a two-state system depending on the position of the nitrogen atom below/above the hydrogen plane.

Ammonia is also a polar molecule, due to the electronegativity difference between nitrogen and hydrogen. The result is that nitrogen tends to attract the electron cloud towards itself more strongly, forming a permanent dipole moment pointing to the center of the molecule by symmetry. This dipole can interact with an external electric field, much like how an electron's spin couples with an external magnetic field. Suppose we apply a rotating magnetic field

This problem is then exactly the same as that of paramagnetic resonance. The result is that the probability of transition from the $|\uparrow\rangle$ to the $|\downarrow\rangle$ state is

# Addition of angular momenta

13

## 13.1 Total angular momentum operator

Suppose we have a set of operators $\hat{J}_i^{(1)}$ and $\hat{J}_i^{(2)}$ satisfying the algebra of angular momentum operators in their respective Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$:

$$[J_i^{(1)}, J_j^{(1)}] = i\hbar\epsilon_{ijk}J_k^{(1)} \tag{13.1.1}$$
$$[J_i^{(2)}, J_j^{(2)}] = i\hbar\epsilon_{ijk}J_k^{(2)} \tag{13.1.2}$$

Suppose we want the angular momentum in the combined space $\mathcal{H}_1 \otimes \mathcal{H}_2$. Clearly it must be defined as:

$$\mathcal{J}_i \equiv \hat{J}_i^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_i^{(2)} \tag{13.1.3}$$

Indeed, we can check that this operator satisfies the commutator algebra:

$$[J_i, J_j] = [\hat{J}_i^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_i^{(2)}, \hat{J}_j^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_j^{(2)}] \tag{13.1.4}$$
$$= [\hat{J}_i^{(1)} \otimes \mathbb{1}, \hat{J}_j^{(1)} \otimes \mathbb{1}] + [\mathbb{1} \otimes \hat{J}_i^{(2)}, \mathbb{1} \otimes \hat{J}_j^{(2)}] \tag{13.1.5}$$

since operators $\hat{J}_i^{(1)} \otimes \mathbb{1}$ and $\hat{J}_i^{(2)} \otimes \mathbb{1}$ must commute. Consequently:

$$[J_i, J_j] = [\hat{J}_i^{(1)}, \hat{J}_j^{(1)}] \otimes \mathbb{1} + \mathbb{1} \otimes [\hat{J}_i^{(2)}, \hat{J}_j^{(2)}] \tag{13.1.6}$$
$$= i\hbar\epsilon_{ijk}(J_k^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes J_k^{(2)}) \tag{13.1.7}$$
$$= i\hbar\epsilon_{ijk}J_k \tag{13.1.8}$$

as desired.

**Total angular momentum operator**

Consider two operators $\hat{J}_i^{(1)}$ and $\hat{J}_i^{(2)}$ satisfying the algebra of angular momentum operators in their respective Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$. Then, the **total angular momentum** operator on the combiend Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is:

$$\mathcal{J}_i \equiv \hat{J}_i^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_i^{(2)} \tag{13.1.9}$$

## 13.2 General problem

Now there are mainly two sets of CSCO (complete set of commuting operators). The first option is to diagonalize $\mathbf{J}_1^2, \mathbf{J}_2^2, \hat{J}_z^{(1)}, \hat{J}_z^{(2)}$ to construct the uncoupled direct product basis $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ satisfying:

$$\mathbf{J}_1^2 |j_1, m_1\rangle \otimes |j_2, m_2\rangle = j_1(j_1+1) |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{13.2.1}$$
$$\mathbf{J}_2^2 |j_1, m_1\rangle \otimes |j_2, m_2\rangle = j_2(j_2+1) |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{13.2.2}$$
$$\hat{J}_z^{(1)} |j_1, m_1\rangle \otimes |j_2, m_2\rangle = m_1 |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{13.2.3}$$
$$\hat{J}_z^{(2)} |j_1, m_1\rangle \otimes |j_2, m_2\rangle = m_2 |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{13.2.4}$$

This basis is particularly useful when we are interested in the state of the components of a system.

However what if we are interested in the overall angular momentum of the entire system? Then we would have to choose to diagonalize $\mathbf{J}^2, \hat{J}_z, \mathbf{J}_1^2, \mathbf{J}_2^2$ [1] to construct the basis states $|j, m, j_1, j_2\rangle$ so that.

$$\mathbf{J}^2 |j, m, j_1, j_2\rangle = j(j+1) |j, m, j_1, j_2\rangle \tag{13.2.5}$$
$$\hat{J}_z |j, m, j_1, j_2\rangle = m |j, m, j_1, j_2\rangle \tag{13.2.6}$$
$$\mathbf{J}_1^2 |j, m, j_1, j_2\rangle = j_1(j_1+1) |j, m, j_1, j_2\rangle \tag{13.2.7}$$
$$\mathbf{J}_2^2 |j, m, j_1, j_2\rangle = j_2(j_2+1) |j, m, j_1, j_2\rangle \tag{13.2.8}$$

Our problem is to find how to transform between these two bases, how do we express $|j, m, j_1, j_2\rangle$ in the direct product (uncoupled) basis?

Let us expand the operator $\hat{J}^2$:

$$\hat{J}^2 = (\mathbf{J}_1 + \mathbf{J}_2) \cdot (\mathbf{J}_1 + \mathbf{J}_2) \tag{13.2.9}$$
$$= \hat{J}_1^2 + \hat{J}_2^2 + 2\mathbf{J}_1 \cdot \mathbf{J}_2 \tag{13.2.10}$$

since $[\mathbf{J}_1, \mathbf{J}_2] = 0$. Unfortunately, we do not quite know what $\mathbf{J}_1 \cdot \mathbf{J}_2$ looks like, we must therefore find an expression for this new operator. Consider:

$$\hat{J}_+^{(1)} \hat{J}_-^{(2)} = (\hat{J}_x^{(1)} + i\hat{J}_y^{(1)})(\hat{J}_x^{(2)} - i\hat{J}_y^{(2)}) \tag{13.2.11}$$
$$= \hat{J}_x^{(1)} \hat{J}_x^{(2)} + \hat{J}_y^{(1)} \hat{J}_y^{(2)} + i(\hat{J}_y^{(1)} \hat{J}_x^{(2)} - \hat{J}_x^{(1)} \hat{J}_y^{(2)}) \tag{13.2.12}$$

and similarly:

$$\hat{J}_-^{(1)} \hat{J}_+^{(2)} = (\hat{J}_x^{(1)} - i\hat{J}_y^{(1)})(\hat{J}_x^{(2)} + i\hat{J}_y^{(2)}) \tag{13.2.13}$$
$$= \hat{J}_x^{(1)} \hat{J}_x^{(2)} + \hat{J}_y^{(1)} \hat{J}_y^{(2)} - i(\hat{J}_y^{(1)} \hat{J}_x^{(2)} - \hat{J}_x^{(1)} \hat{J}_y^{(2)}) \tag{13.2.14}$$

[1] note that these all commute with each other by writing $\mathbf{J}^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + 2J_z^{(1)} J_z^{(2)} + J_+^{(1)} J_-^{(2)} + J_-^{(1)} J_+^{(2)}$ and using the standard commutation relations of $\mathbf{J}^2$ algebra

Adding (13.2.12) and (13.2.14) we find that:

$$\hat{J}_+^{(1)}\hat{J}_-^{(2)} + \hat{J}_-^{(1)}\hat{J}_+^{(2)} = 2(\mathbf{J}_1 \cdot \mathbf{J}_1 - \hat{J}_z^{(1)}\hat{J}_z^{(2)}) \tag{13.2.15}$$

Consequently, we can rearrange the above expression to $\hat{J}_+^{(1)}\hat{J}_-^{(2)} + \hat{J}_-^{(1)}\hat{J}_+^{(2)} + 2\hat{J}_z^{(1)}\hat{J}_z^{(2)} = 2\mathbf{J}_1 \cdot \mathbf{J}_2$ [2] and find that:

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_z^{(1)}\hat{J}_z^{(2)} + \hat{J}_+^{(1)}\hat{J}_-^{(2)} + \hat{J}_-^{(1)}\hat{J}_+^{(2)} \tag{13.2.19}$$

By physical arguments, we expect that:

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.20}$$

where we suppressed the $j_1, j_2$ quantum numbers in $|j_1 + j_2, j_1 + j_2\rangle$ (which should have been $|j_1 + j_2, j_1 + j_2, j_1, j_2\rangle$) since they are implicitly understood. We shall use this notation throughout the chapter unless when stated otherwise.

To check this, we need to show that$|j_1, j_1\rangle \otimes |j_2, j_2\rangle$ is an eigenstate of $\hat{L}^2$ and $\hat{L}_z$ with associated eigenvalues $(j_1 + j_2)(j_1 + j_2 + 1)$ and $j_1 + j_2$ respectively.

Therefore, let us apply the operator $\hat{L}_z$ on $|j_1, j_1\rangle \otimes |j_2, j_2\rangle$:

$$\hat{J}_z(|j_1, j_1\rangle \otimes |j_2, j_2\rangle) = (\hat{J}_z^{(1)} + \hat{J}_z^{(2)}) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.21}$$

$$= j_1 |j_1, j_1\rangle \otimes |j_2, j_2\rangle + j_2 |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.22}$$

$$= (j_1 + j_2) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.23}$$

so we see that $|j_1, j_1\rangle \otimes |j_2, j_2\rangle$ is an eigenvector of $\hat{J}_z$ with eigenvalue $j_1 + j_2$ as expected.

Let us apply $\hat{J}^2$ in a similar fashion:

$$\hat{J}^2(|j_1, j_1\rangle \otimes |j_2, j_2\rangle) = (\hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_z^{(1)}\hat{J}_z^{(2)} + \hat{J}_+^{(1)}\hat{J}_-^{(2)} + \hat{J}_-^{(1)}\hat{J}_+^{(2)}) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.24}$$

$$= (j_1(j_1 + 1) + j_2(j_2 + 1) + 2j_1 j_2) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.25}$$

where for example $\hat{J}_+^{(1)}\hat{J}_-^{(2)} |j_1, j_1\rangle |j_2, j_2\rangle = 0$ since the state $|j_1, j_1\rangle$ cannot be raised. We simplify the expression further by factoring the term $2j_1 j_2$ in the first two products:

$$\hat{J}^2 |j_1 + j_2, j_1 + j_2\rangle = (j_1(j_1 + j_2 + 1) + j_2(j_1 + j_2 + 1)) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.26}$$

$$= j(j + 1) |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.27}$$

[2] alternatively we can work in reverse:

$$\mathbf{J}_1 \cdot \mathbf{J}_2 = \hat{J}_z^{(1)}\hat{J}_z^{(2)} + \hat{J}_x^{(1)}\hat{J}_x^{(2)} + \hat{J}_y^{(1)}\hat{J}_y^{(2)} \tag{13.2.16}$$

$$= \hat{J}_z^{(1)}\hat{J}_z^{(2)} + \frac{1}{4}(\hat{J}_+^{(1)} + \hat{J}_-^{(1)})(\hat{J}_+^{(2)} + \hat{J}_-^{(2)}) - \frac{1}{4}(\hat{J}_+^{(1)} - \hat{J}_-^{(1)})(\hat{J}_+^{(2)} - \hat{J}_-^{(2)}) \tag{13.2.17}$$

$$= \hat{J}_z^{(1)}\hat{J}_z^{(2)} + \frac{1}{2}(\hat{J}_+^{(1)}\hat{J}_-^{(2)} + \hat{J}_-^{(1)}\hat{J}_+^{(2)}) \tag{13.2.18}$$

where we defined $j = j_1 + j_2$. So, we see that indeed $|j_1, j_1\rangle \otimes |j_2, j_2\rangle$ is indeed an eigenvector of $\hat{J}^2$ with the appropriate eigenvalue $j(j+1)$. We may therefore conclude that:

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle \tag{13.2.28}$$

To find the other eigenstates (with lower $\hat{J}_z$ eigenvalues) we can simply operate $\hat{J}_-$ on $|j, j\rangle = |j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle$ successively. For example:

$$\hat{J}_- |j, j\rangle = \sqrt{j(j+1) - j(j-1)}\, |j, j-1\rangle = \sqrt{2j}\, |j, j-1\rangle \tag{13.2.29}$$

while on the other hand:

$$\hat{J}_-^{(1)} + \hat{J}_-^{(2)} |j_1, j_1\rangle |j_2, j_2\rangle = \sqrt{2j_1}\, |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle + \sqrt{2j_2}\, |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \tag{13.2.30}$$

so that:

$$|j, j-1\rangle = \sqrt{\frac{j_1}{j}}\, |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle + \sqrt{\frac{j_2}{j}}\, |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \tag{13.2.31}$$

The first term represents the state where the first subsystem has a reduced angular momentum along the $z$ axis, whereas the second term represents the state where the second subsystem has a reduced angular momentum along the $z$ axis. Moreover, the larger the angular momentum along the $z$ axis, the more likely it is for it to be reduced.

Interestingly, the two subsystems have been correlated to form this entangled state. This means that knowing the angular momentum of each subsystem is not enough to know the angular momentum of the composite system.

To find $|j, j-2\rangle$ we simply operate $\hat{J}_-$ again. Visually, we can see this as shown in Figure 12.1.

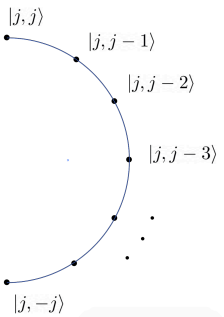



**Figure 13.1**. Diagram illustrating how to add angular momenta. Each state $|j, m\rangle$, which is represented by points. Operating $\hat{J}_-$ is equivalent to moving down the semi-circle.

The number of ways to orient the first subsystem is $2j_1 + 1$, and similarly the number of ways to orient the second subsystem is $2j_2 + 1$, hence the total number of angular momentum eigenstates of the combined system is $(2j_1+1)(2j_2+1)$. However, we have only found $2(j_1 + j_2) + 1$ eigenstates $|j, m\rangle$ in total, since

$$|j_1 - j_2| \leq m \leq j_1 + j_2 \tag{13.2.32}$$

where the lower bound for $m$ occurs when the two spins are perfectly anti-parallel.

So how do we reduce the total angular momentum rather than just the angular momentum along the $z$-axis? In other words, how do we calculate $|j-1, j-1\rangle$?

Firstly, we have that $|j, j-1\rangle$ and $|j-1, j-1\rangle$ are both eigenstates of $\hat{J}^2$ (which is Hermitian) with different eigenvalues. Consequently, they must be orthogonal. Moreover on physical grounds we argue that $|j-1, j-1\rangle$, having spin $j-1$ along the $z$-axis, must be a superposition of $|j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle$ and $|j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle$.

Indeed we have that:

$$\hat{J}_z |j-1, j-1\rangle = \hbar(j-1) |j-1, j-1\rangle \tag{13.2.33}$$

and expanding $|j-1, j-1\rangle$ into a linear superposition of $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$, then:

$$\hat{J}_z |j_1, m_1\rangle \otimes |j_2, m_2\rangle = \hbar(m_1 + m_2) |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{13.2.34}$$

For the superposition to be an eigenstate of $\hat{J}_z$ with eigenvalue $j-1$ we need $m_1 + m_2 = j - 1 = j_1 + j_2 - 1$ for each term. The only possibilities are then $m_1 = j_1, m_2 = j_2 - 1$ and $m_1 = j_1 - 1, m_2 = j_2$.

Hence we may write

$$|j-1, j-1\rangle = \alpha |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle + \beta |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \tag{13.2.35}$$

The only choice of $\alpha, \beta$ (up to an overall constant) which makes this state orthogonal to $|j, j-1\rangle$ is $\alpha = \sqrt{\frac{j_2}{j}}$ and $\beta = -\sqrt{\frac{j_1}{j}}$:

$$|j-1, j-1\rangle = \sqrt{\frac{j_2}{j}} |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle - \sqrt{\frac{j_1}{j}} |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \tag{13.2.36}$$

One may wonder how such a configuration is possible, why doesn't the total angular momentum add up like the angular momentum along $z$ does? Well, the two angular momenta don't necessarily have to be aligned, they may be at angles to each other. Consequently, they will not add up like scalars (such as the $z$-component of angular momenta) do.

We can represent the process that we have followed in the diagram below: We can summarize our results to write:

$$j_1 \otimes j_2 = (j_1 + j_2) \oplus (j_1 + j_2 - 1) \oplus ... \oplus (j_1 - j_2) \tag{13.2.37}$$

which means that the direct product of $j_1$-space and $j_2$-space gives the direct sum of $j_1 + j_2$-space, $j_1 + j_2 - 1$-space etc... until $j_1 - j_2$-space.

Consider for example the case where $j_1 = 1, j_2 = \frac{1}{2}$ so that $j = \frac{3}{2}$. Then the highest angular

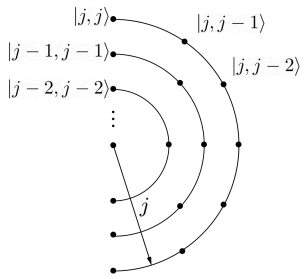



**Figure 13.2**. Diagram illustrating how to add angular momenta. Each semi-circle of radius $j$ represents the ensemble of states $|j,m\rangle$, which are each represented by points. Operating $\hat{J}_-$ is equivalent to moving down a semi-circle.

momentum state is $|\frac{3}{2},\frac{3}{2}\rangle$. Then:

$$|\frac{3}{2},\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}\,|1,0\rangle\otimes|\frac{1}{2},\frac{1}{2}\rangle + \sqrt{\frac{1}{3}}\,|1,1\rangle\otimes|\frac{1}{2},-\frac{1}{2}\rangle \tag{13.2.38}$$

We may then construct the following state:

$$|\frac{1}{2},\frac{1}{2}\rangle = \sqrt{\frac{1}{3}}\,|1,0\rangle\otimes|\frac{1}{2},\frac{1}{2}\rangle - \sqrt{\frac{2}{3}}\,|1,1\rangle\otimes|\frac{1}{2},-\frac{1}{2}\rangle \tag{13.2.39}$$

We define the constants:

$$C(j,m;j_1,j_2;m_1,m_2) \equiv \langle j,m\,|\,j_1,m_1\otimes j_2,m_2\rangle \tag{13.2.40}$$

as the **Clebsch-Gordan coefficients**. This allows us to easily transform between the uncoupled and coupled bases, fully solving the problem of adding angular momenta. Let us see some applications.

## 13.3 Spin-$\frac{1}{2}$ addition

Consider a system of two spin-$\frac{1}{2}$ particles, such as a hydrogen atom composed of one proton and one electron. Hence we have that $j_1 = j_2 = \frac{1}{2}$, and thus $j = 1$.

We also define the following shorthand notation:

$$|\uparrow\rangle_e = |\frac{1}{2},\frac{1}{2}\rangle_e \tag{13.3.1}$$

$$|\downarrow\rangle_e = |\frac{1}{2},-\frac{1}{2}\rangle_e \tag{13.3.2}$$

$$|\uparrow\rangle_p = |\frac{1}{2},\frac{1}{2}\rangle_p \tag{13.3.3}$$

$$|\downarrow\rangle_p = |\frac{1}{2},-\frac{1}{2}\rangle_p \tag{13.3.4}$$

Following the same procedure as in the previous section, we start with the maximal an-

gular momentum state:

$$|1,1\rangle = |\frac{1}{2},\frac{1}{2}\rangle_e \otimes |\frac{1}{2},\frac{1}{2}\rangle_p = |\uparrow\rangle_e \otimes |\uparrow\rangle_p \tag{13.3.5}$$

Operating $\hat{J}^-$ to this state:

$$|1,0\rangle = \frac{1}{\sqrt{2}}(|\frac{1}{2},-\frac{1}{2}\rangle_e \otimes |\frac{1}{2},\frac{1}{2}\rangle_p + |\frac{1}{2},\frac{1}{2}\rangle_e \otimes |\frac{1}{2},-\frac{1}{2}\rangle_p) \tag{13.3.6}$$

$$= \frac{1}{\sqrt{2}}(|\downarrow\rangle_e \otimes |\uparrow\rangle_p + |\uparrow\rangle_e \otimes |\downarrow\rangle_p) \tag{13.3.7}$$

Applying $\hat{J}^-$ again:

$$|1,-1\rangle = \frac{1}{\sqrt{2}}\Big(\frac{1}{\sqrt{2}}|\frac{1}{2},-\frac{1}{2}\rangle_e \otimes |\frac{1}{2},-\frac{1}{2}\rangle_p + \frac{1}{\sqrt{2}}|\frac{1}{2},-\frac{1}{2}\rangle_e \otimes |\frac{1}{2},-\frac{1}{2}\rangle_p\Big) \tag{13.3.8}$$

$$= |\frac{1}{2},-\frac{1}{2}\rangle_e \otimes |\frac{1}{2},-\frac{1}{2}\rangle_p = |\downarrow\rangle_e \otimes |\downarrow\rangle_p \tag{13.3.9}$$

Finally, we also have that:

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|\downarrow\rangle_e \otimes |\uparrow\rangle_p - |\uparrow\rangle_e \otimes |\downarrow\rangle_p) \tag{13.3.10}$$

It is important to note that the spin states $|1,1\rangle, |1,0\rangle, |1,-1\rangle$, called **triplet states**, are symmetric, whereas $|0,0\rangle$, known as the **singlet state**, is anti-symmetric.

Thus, if we consider an ensemble of two electrons, then we may construct its state as the direct product of a state vector in the spatial Hilbert space and a state vector in the spin Hilbert space. If the spatial component is symmetric, the spin state must be anti-symmetric, and vice versa.

The basis states of the two-electron Hilbert space is therefore:

$$|\Psi_1\rangle = \frac{1}{\sqrt{2}}(|\psi_1,\psi_2\rangle - |\psi_2,\psi_1\rangle) \otimes |1,1\rangle \tag{13.3.11}$$

$$|\Psi_2\rangle = \frac{1}{\sqrt{2}}(|\psi_1,\psi_2\rangle - |\psi_2,\psi_1\rangle) \otimes |1,0\rangle \tag{13.3.12}$$

$$|\Psi_3\rangle = \frac{1}{\sqrt{2}}(|\psi_1,\psi_2\rangle - |\psi_2,\psi_1\rangle) \otimes |1,-1\rangle \tag{13.3.13}$$

$$|\Psi_4\rangle = \frac{1}{\sqrt{2}}(|\psi_1,\psi_2\rangle + |\psi_2,\psi_1\rangle) \otimes |0,0\rangle \tag{13.3.14}$$

## 13.4 Adding orbital angular momentum and spin

Consider the addition of orbital angular momentum and spin-$\frac{1}{2}$: $\mathbf{J} = \mathbf{L} + \mathbf{S}$. We are looking for eigenstates of $\hat{J}^2, \hat{J}_z, \hat{L}^2, \hat{S}^2$, denoted by $|j, m_j, l\rangle$. We have that:

$$\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\mathbf{L}\cdot\mathbf{S} \tag{13.4.1}$$
$$= \hat{L}^2 + \hat{S}^2 + 2\hat{L}_z\hat{S}_z + \hat{L}_+\hat{S}_- + \hat{S}_+\hat{L}_- \tag{13.4.2}$$

Clearly the maximal angular momentum state for $j = l + \frac{1}{2}$ is:

$$|l+\frac{1}{2}, l+\frac{1}{2}\rangle = |l,l\rangle \otimes |\uparrow\rangle \tag{13.4.3}$$

Indeed:

$$\hat{J}_z\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle = \Big(l+\frac{1}{2}\Big)\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle \tag{13.4.4}$$

and

$$\hat{J}^2\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle = \hbar^2(l(l+1)+\frac{3}{4}+l)\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle \tag{13.4.5}$$
$$= \Big(l+\frac{1}{2}\Big)\Big(l+\frac{3}{2}\Big)\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle \tag{13.4.6}$$

as desired. Now, we may apply the lowering operator $\hat{J}_-$ to find that:

$$\hat{J}_-\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle = \sqrt{2l+1}\,|l+\frac{1}{2}, l-\frac{1}{2}\rangle \tag{13.4.7}$$

At the same time however, $\hat{J}_- = \hat{L}_- + \hat{S}_-$ so that:

$$\hat{J}_-\,|l+\frac{1}{2}, l+\frac{1}{2}\rangle = (\hat{L}_- + \hat{S}_-)\,|l,l\rangle \otimes |\uparrow\rangle \tag{13.4.8}$$
$$= \sqrt{l(l+1)-l(l-1)}\,|l,l-1\rangle \otimes |\uparrow\rangle + \sqrt{\frac{3}{4}+\frac{1}{4}}\,|l,l\rangle \otimes |\downarrow\rangle \tag{13.4.9}$$
$$= \sqrt{2l}\,|l,l-1\rangle \otimes |\uparrow\rangle + |l,l\rangle \otimes |\downarrow\rangle \tag{13.4.10}$$

Therefore:

$$|l+\frac{1}{2}, l-\frac{1}{2}\rangle = \sqrt{\frac{2l}{2l+1}}\,|l,l-1\rangle \otimes |\uparrow\rangle + \frac{1}{\sqrt{2l+1}}\,|l,l\rangle \otimes |\downarrow\rangle \tag{13.4.11}$$

More generally, the state $|l+\frac{1}{2}, m\rangle$ will be a linear superposition of $|l, m-\frac{1}{2}\rangle \otimes |\uparrow\rangle$ and $|l, m+\frac{1}{2}\rangle \otimes |\downarrow\rangle$. Hence:

$$|l+\frac{1}{2}, m\rangle = \alpha\,|l, m-\frac{1}{2}\rangle \otimes |\uparrow\rangle + \beta\,|l, m+\frac{1}{2}\rangle \otimes |\downarrow\rangle \tag{13.4.12}$$

We now impose the condition:

$$\begin{aligned}\hat{J}^2\,|l+\frac{1}{2},m\rangle =&\left(l+\frac{1}{2}\right)\left(l+\frac{3}{2}\right)|l+\frac{1}{2},m\rangle\\ =&\alpha\left[l(l+1)+\frac{3}{4}+m-\frac{1}{2}+\frac{\beta}{\alpha}\sqrt{l(l+1)-\left(m-\frac{1}{2}\right)\left(m+\frac{1}{2}\right)}\right]|l,m-\frac{1}{2}\rangle\otimes|\uparrow\rangle\\ &+\beta\left[l(l+1)+\frac{3}{4}-m-\frac{1}{2}+\frac{\alpha}{\beta}\sqrt{l(l+1)-\left(m-\frac{1}{2}\right)\left(m+\frac{1}{2}\right)}\right]|l,m+\frac{1}{2}\rangle\otimes|\downarrow\rangle\\ =&\alpha\left[l(l+1)+m+\frac{1}{4}+\frac{\beta}{\alpha}\sqrt{l(l+1)-m^2+\frac{1}{4}}\right]|l,m-\frac{1}{2}\rangle\otimes|\uparrow\rangle\\ &+\beta\left[l(l+1)-m+\frac{1}{4}+\frac{\alpha}{\beta}\sqrt{l(l+1)-m^2+\frac{1}{4}}\right]|l,m+\frac{1}{2}\rangle\otimes|\downarrow\rangle \quad (13.4.13)\end{aligned}$$

so comparing term by term:

$$l(l+1)+m+\frac{1}{4}+\frac{\beta}{\alpha}\sqrt{l(l+1)-m^2+\frac{1}{4}}=\left(l+\frac{1}{2}\right)\left(l+\frac{3}{2}\right) \quad (13.4.14)$$

$$\iff \frac{\beta}{\alpha}\sqrt{l(l+1)-m^2+\frac{1}{4}}=l-m+\frac{1}{2} \quad (13.4.15)$$

and similarly:

$$l(l+1)-m+\frac{1}{4}+\frac{\alpha}{\beta}\sqrt{l(l+1)-m^2+\frac{1}{4}}=\left(l+\frac{1}{2}\right)\left(l+\frac{3}{2}\right) \quad (13.4.16)$$

$$\iff \frac{\alpha}{\beta}\sqrt{l(l+1)-m^2+\frac{1}{4}}=l+m+\frac{1}{2} \quad (13.4.17)$$

Dividing (13.4.15) by (13.4.17) we find that

$$\frac{\beta}{\alpha}=\sqrt{\frac{l-m+\frac{1}{2}}{l+m+\frac{1}{2}}} \quad (13.4.18)$$

Combining this with normalization $\alpha^2+\beta^2=1$ we finally get that:

$$\alpha=\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}},\ \beta=\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} \quad (13.4.19)$$

Consequently:

$$|l+\frac{1}{2},m\rangle=\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}\,|l,m-\frac{1}{2}\rangle\otimes|\uparrow\rangle+\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}\,|l,m+\frac{1}{2}\rangle\otimes|\downarrow\rangle \quad (13.4.20)$$

It follows immediately that:

$$|l-\frac{1}{2},m\rangle = \sqrt{\frac{l-m+\frac{1}{2}}{2l+1}}\,|l,m-\frac{1}{2}\rangle\otimes|\uparrow\rangle - \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}\,|l,m+\frac{1}{2}\rangle\otimes|\downarrow\rangle \tag{13.4.21}$$

In summary one may write that:

$$l\otimes\frac{1}{2} = \left(l+\frac{1}{2}\right)\oplus\left(l-\frac{1}{2}\right) \tag{13.4.22}$$

# 14 Entanglement, the EPR paradox and Quantum information

## 14.1 Composite systems & Entanglement

Suppose we have two systems $A, B$ in states $i, j$ so that they are described by the state vectors $|i\rangle_A$ and $|j\rangle_B$, so that the state of both systems is described by:

$$|i,j\rangle_{AB} = |i\rangle_A \otimes |j\rangle_B \tag{14.1.1}$$

Surprisingly, not all states of the composite system $AB$ are in the form of a tensor product.

For example, consider two 2-state systems:

$$A : |u_1\rangle, |u_2\rangle \tag{14.1.2}$$
$$B : |v_1\rangle, |v_2\rangle \tag{14.1.3}$$

The general state of systems A,B are given by the superpositions:

$$|A\rangle = \alpha |u_1\rangle + \beta |u_2\rangle \tag{14.1.4}$$
$$|B\rangle = \gamma |v_1\rangle + \delta |v_2\rangle \tag{14.1.5}$$

Then the most general state of the composite system is:

$$|\Psi\rangle_{AB} = \alpha\gamma |u_1\rangle \otimes |v_1\rangle + \alpha\delta |u_1\rangle \otimes |v_2\rangle + \beta\gamma |u_2\rangle \otimes |v_1\rangle + \beta\delta |u_2\rangle \otimes |v_2\rangle \tag{14.1.6}$$

Using Baye's theorem, the probability of measuring system B in the state $|v_1\rangle$ given that system A is in the state $|u_1\rangle$ is:

$$P(v_1|u_1) = \frac{P(u_1 \cup v_1)}{P(u_1)} \tag{14.1.7}$$
$$= \frac{P(u_1 \cup v_1)}{P(u_1 \cup v_1) + P(u_1 \cup v_2)} \tag{14.1.8}$$
$$= \frac{1}{1 + P(u_1 \cup v_2)/P(u_1 \cup v_1)} \tag{14.1.9}$$
$$= \frac{1}{1 + \delta/\gamma} \tag{14.1.10}$$

So $P(v_1|u_1)$ is independent of the coefficients of $|u_1\rangle$(which is $\alpha$) and $|u_2\rangle$ (which is $\beta$). This implies that the systems $A$ and $B$ are independent/uncorrelated.

So systems described by product states $|\Psi\rangle_{AB} = |A\rangle\,|B\rangle$ must be uncorrelated. Correlated systems, such as a hydrogen atom, are not generally written as tensor products.

Such states that cannot be written as product of states are called **entangled states**.

**Entangled states**

Two particles are said to be entangled when they do not have their own quantum states, but rather they are described as a single state which cannot be expressed as a tensor product.

However, it turns out that the tensor product states do span the Hilbert space of all state vectors for the system (both correlated and uncorrelated) $AB$:

$$|\Psi\rangle_{AB} = \sum_{ij} c_{ij}\,|i\rangle_A\,|j\rangle_B \tag{14.1.11}$$

*Proof.* Consider some state $|\phi\rangle_{AB}$ that cannot be expressed as a superposition as shown in (14.1.11). We construct the state $|\chi\rangle_{AB}$ as:

$$|\chi\rangle_{AB} = |\phi\rangle_{AB} - \sum_{ij} c_{ij}\,|i\rangle_A\,|j\rangle_B\,,\ c_{ij} = \langle i,j\,|\,\phi\rangle_{AB} \tag{14.1.12}$$

Now the probability of system $A$ in the state $|\chi\rangle_{AB}$ to be in the state $|l\rangle_A$ is:

$$\sum_j |l,j|\chi\rangle_{AB} = |l,j|\phi\rangle - \sum_{ij} c_{ij}\,\langle l,j\,|\,i,j\rangle = 0 \tag{14.1.13}$$

implying that $|\chi\rangle_{AB}$ is the zero vector. Consequently:

$$|\phi\rangle_{AB} = \sum_{ij} c_{ij}\,|i\rangle_A\,|j\rangle_B \tag{14.1.14}$$

as desired. □

Consider a hydrogen atom composed of a proton and an electron. In the position representation:

$$\phi_i(\mathbf{x}_e) = \langle \mathbf{x}_e\,|\,\phi_i\rangle\,,\ \varphi_i(\mathbf{x}_p) = \langle \mathbf{x}_p\,|\,\varphi_j\rangle \tag{14.1.15}$$

The state of the hydrogen atom modelled as an uncorrelated system (very unrealistic!) would then be:

$$\langle \mathbf{x}_e, \mathbf{x}_p\,|\,\psi_{ij}\rangle = \langle \mathbf{x}_e\,|\,\phi_i\rangle\,\langle \mathbf{x}_p\,|\,\varphi_i\rangle = \phi_i(\mathbf{x}_e)\varphi_i(\mathbf{x}_p) \tag{14.1.16}$$

As a correlated system, we may write:

$$\langle \mathbf{x}_e, \mathbf{x}_p\,|\,\psi\rangle = \sum_{ij} c_{ij}\,\langle \mathbf{x}_e\,|\,\phi_j\rangle\,\langle \mathbf{x}_p\,|\,\varphi_i\rangle = \sum_{ij} \phi_i(\mathbf{x}_e)\varphi_j(\mathbf{x}_p) \tag{14.1.17}$$

This is very very strange. In classical probability theory, we cannot generally write correlated events' probabilities using uncorrelated event's probabilities. In quantum mechanics we can do this.

An interesting property of entangled systems lies in the collapse of their wave-function.

Consider for example the state:

$$|\psi\rangle = |u_1, v_1\rangle + |u_2\rangle \left(\alpha |v_1\rangle + \beta |v_2\rangle\right) \tag{14.1.18}$$

If we perform a measurement on system $A$, and find $|u_1\rangle_A$, then the state of the composite system collapses to:

$$|\psi'\rangle = |u_1, v_1\rangle \tag{14.1.19}$$

Therefore we also collapse the state of system $B$ to $|v_1\rangle_B$ by performing a measurement on system $A$. This is what we mean by correlated/entangled systems.

If instead we find $|u_2\rangle_A$ then the state of the composite system collapses to:

$$|\psi'\rangle = |u_2\rangle \left(\alpha |v_1\rangle + \beta |v_2\rangle\right) \tag{14.1.20}$$

so the outcome of a measurement on system $B$ is no longer uniquely determined.

Another interesting property of entangled systems lies in their time evolution.

For two uncorrelated subsystems, the total Hamiltonian is separable as:

$$\hat{H} = \hat{H}_A + \hat{H}_B \tag{14.1.21}$$

from which it follows that:

$$[\hat{H}_{AB}, \hat{H}_A] = [\hat{H}_A + \hat{H}_B, \hat{H}_A] = 0 \tag{14.1.22}$$

Therefore, we can measure the total energy of the subsystem A (or B) and of the composite system simultaneously. We can find a simultaneous diagonalized basis for $\hat{H}_{AB}$, $\hat{H}_A$ and $\hat{H}_B$ as was discussed in the chapter on multi-particle systems.

Now let's consider the evolution of entangled states. Each subsystem will evolve according to its TDSE:

$$i\hbar \frac{\partial}{\partial t} |i\rangle_A = \hat{H}_A |i\rangle_A \tag{14.1.23}$$

$$i\hbar \frac{\partial}{\partial t} |j\rangle_B = \hat{H}_B |i\rangle_B \tag{14.1.24}$$

Since we may write the general state of the composite system as:

$$|i\rangle_{AB} = \sum_{ij} c_{ij} |i\rangle_A |j\rangle_B \tag{14.1.25}$$

and the total Hamiltonian is:

$$\hat{H}_{AB} = \hat{H}_A + \hat{H}_B + \hat{H}_{int} \tag{14.1.26}$$

Note that now the hamiltonian is no longer separable:

$$[\hat{H}_{AB}, \hat{H}_A] \neq 0 \tag{14.1.27}$$

so it is impossible to measure the energy of the total system and the energy of a subsystem simultaneously, the two operators don't commute.

Consequently the LHS of the TDSE reads:

$$i\hbar\frac{\partial}{\partial t} = i\hbar\sum_{ij}\left[\frac{dc_{ij}}{dt}\left|i\right\rangle_A\left|j\right\rangle_B + c_{ij}\Big(\frac{\partial\left|i\right\rangle_A}{\partial t}\left|j\right\rangle_B + \left|i\right\rangle_A\frac{\partial\left|j\right\rangle_B}{\partial t}\Big)\right] \tag{14.1.28}$$

whereas the RHS of the TDSE reads:

$$\hat{H}_{AB}\left|i\right\rangle_A\left|j\right\rangle_B = \sum_{ij} c_{ij}\hat{H}_A\left|i\right\rangle_A\left|j\right\rangle_B + \hat{H}_B\left|i\right\rangle_A\left|j\right\rangle_B + \hat{H}_{int} \tag{14.1.29}$$

Equating (14.1.28) with (14.1.29) gives:

$$i\hbar\sum_{ij}\frac{dc_{ij}}{dt}\left|i\right\rangle_A\left|j\right\rangle_B = \sum_{ij} c_{ij}\hat{H}_{int}\left|i\right\rangle_A\left|j\right\rangle_B \tag{14.1.30}$$

Therefore the superposition coefficients evolve according solely to the interaction term in the Hamiltonian, that is, the term which couples subsystem A with subsystem B.

Sandwhiching (14.1.30) with $\left\langle k\right|_A\left\langle l\right|_B$ we finally find that:

$$i\hbar\frac{dc_{kl}}{dt} = \sum_{ij} c_{ij}\left\langle k,l\right|\hat{H}_{int}\left|i,j\right\rangle_{AB} \tag{14.1.31}$$

Consider for example an interaction hamiltonian such that $c_{ij}\left\langle k,l\right|\hat{H}_{int}\left|i,j\right\rangle_{AB}$ doesn't vanish only for $i = k = k_0$ and $j = l = l_0$, so that it starts out unentangled. Only $c_{k_0l_0}$ will not eventually vanish with time, but this also means that the systems will eventually get entangled.

## 14.2 Examples of entanglement

Consider the apparatus of a Mach-Zehnder interferometer, consisting of a source of photons incident on a half-silvered mirror, and two photon detectors at the sides of the beam splitter.

We stated that each photon actually travels through both paths, it will be both transmitted and reflected. Let's put this in quantitative terms.

Let's denote the state $|0\rangle$ as the vacuum-state, which contains no photons, and $|1\rangle$ as the state containing one photon. After a photon hits the beam splitter, its state will therefore be:

$$|BS\rangle = \alpha\,|0\rangle_R\,|1\rangle_T + \beta\,|1\rangle_R\,|0\rangle_T \tag{14.2.1}$$

satisfying $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$. This is yet again an entangled state

We also see entanglement in $\alpha$-decay. When a uranium-238 nucleus decays, it emits an $\alpha$-particle leaving thorium-234. Since the nucleus had no initial momentum, conservation laws force the momentum of the $\alpha$-particle and thorium isotope to be in opposite directions. However, because the $\alpha$-particle can be emitted in any direction, we cannot state a priori what the momentum of the two decay products must be.

We resolve this issue by postulating that the $\alpha$-particle and the thorium isotope are entangled, so that whenever the momentum of one is measured definitely, the state of the other must also collapse with opposite momentum.

If the $\alpha$-particle collided with an asteroid for example, no matter the distance with the thorium isotope, the latter will also assume a definite momentum.

## 14.3 EPR Argument

Suppose we have a nucleus that decays into an electron and positron. Alice measures the component of spin of the electron in some direction **a**. Similarly, bob measures the component of the spin of the positron in some direction **b**.

Nuclear physics states that the nucleus starts out with no net spin, so by conservation of angular momentum, the measured spin of the electron in some direction must be opposite to that of the positron in that same direction.

Consequently, if Alice measures the spin first, and finds that $S_a = \pm\frac{1}{2}$. Then, it is guaranteed that Bob measures the spin of the positron as $S_b = \mp\frac{1}{2}$.

For simplicity, we can choose the z-axis to be along **a** without loss of generality. Then we can construct the state vector using addition of angular momenta:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|z,\uparrow\rangle_e\,|z,\downarrow\rangle_p - |z,\downarrow\rangle_p\,|z,\uparrow\rangle_p) \tag{14.3.1}$$

After Alice has made the measurement that the electron is spin-up, the state collapses into the first term:

$$|\psi\rangle = |z,\uparrow\rangle_e\,|z,\downarrow_p\rangle \tag{14.3.2}$$

Now what will Bob measure? The state has collapsed into a spin-down state relative to the z-axis. Since Bob is measuring along some other direction $\mathbf{b} = (\theta, \phi)$, we find that the eigenstate pointing up along **b** is:

$$|\mathbf{b},\uparrow\rangle_p = \cos\frac{\theta}{2}e^{-i\phi/2}\,|z,\uparrow\rangle_p + \sin\frac{\theta}{2}e^{i\phi/2}\,|z,\downarrow\rangle_p \tag{14.3.3}$$

Therefore, the probability that Bob observes spin up along his axis of measurement is:

$$P(\mathbf{b},\uparrow) = |\langle \mathbf{b},\uparrow | z,\downarrow\rangle_p|^2 = \left|\sin\frac{\theta}{2}e^{-i\phi/2}\right|^2 = \sin^2\frac{\theta}{2} \tag{14.3.4}$$

and thus:

$$P(\mathbf{b},\uparrow) = \cos^2\frac{\theta}{2} \tag{14.3.5}$$

Note that if $\theta = 0$, then Bob is guaranteed to measure spin down, as would be expected from spin conservation.

Interestingly, the angle $\theta$ is dependent on the angle that Alice chooses. So why does Bob's measurement depend on Alice's measurement?

Perhaps more terrifyingly, suppose Bob and Alice are light years apart, and perform the measurement simultaneously. It seems like Bob's measurement will instantaneously provide information on what Alice's measurement was, thus violating the postulates of Special relativity!

Einstein, Podolski and Rosen postulated that the result of Alice's measurement was predetermined by a so-called **hidden variable**, so no information was really ever travelling between Alice and Bob.

## 14.4 Bell's inequality

Let $\sigma_A$ and $\sigma_B$ be the values measured by Alice and Bob respectively.

Therefore, $(\sigma_A, \sigma_B)$ can take four different values, namely

$$\left(\frac{1}{2},\frac{1}{2}\right),\left(-\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},-\frac{1}{2}\right),\left(-\frac{1}{2},-\frac{1}{2}\right) \tag{14.4.1}$$

The expectation value of $\sigma_A\sigma_B$ is:

$$\langle\sigma_A\sigma_B\rangle = \frac{1}{4}(P(A_\uparrow)P(B_\uparrow|A_\uparrow) + P(A_\downarrow)P(B_\downarrow|A_\downarrow) - P(A_\uparrow)P(B_\downarrow|A_\uparrow) - P(A_\downarrow)P(B_\uparrow|A_\downarrow)) \tag{14.4.2}$$

Also, $P(A_\uparrow) = P(A_\downarrow) = \frac{1}{2}$ and we have found that:

$$P(B_\uparrow|A_\uparrow) = P(B_\downarrow|A_\downarrow) = \sin^2\frac{\theta}{2},\ P(A_\downarrow|A_\uparrow) = P(A_\uparrow|A_\downarrow) = \cos^2\frac{\theta}{2} \tag{14.4.3}$$

Therefore:

$$\langle\sigma_A\sigma_B\rangle = \frac{1}{4}\left(\sin^2\frac{\theta}{2} - \cos^2\frac{\theta}{2}\right) = -\frac{1}{4}\cos\theta = -\frac{1}{4}\mathbf{a}\cdot\mathbf{b} \tag{14.4.4}$$

In the hidden variable theory postulated by EPR, there is some function $\sigma_e(\mathbf{v},\mathbf{a})$ where $\mathbf{v}$ is an $n$-vector containing the hidden variables, gives the probability of measuring $|\mathbf{a},+\rangle_e$. Similarly there must also be some function $\sigma_p(\mathbf{v},\mathbf{b})$. Both must take values $\pm\frac{1}{2}$ stochasti-

cally. By conservation of angular momentum we do know that:

$$\sigma_e(\mathbf{v},\mathbf{a}) + \sigma_p(\mathbf{v},\mathbf{b}) = 0 \tag{14.4.5}$$

Then:

$$\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle = \int d^n\mathbf{v}\rho(\mathbf{v})\sigma_e(\mathbf{a})\sigma_p(\mathbf{b}) \tag{14.4.6}$$

$$= -\int d^n\mathbf{v}\rho(\mathbf{v})\sigma_e(\mathbf{a})\sigma_e(\mathbf{b}) \tag{14.4.7}$$

Suppose we now choose another vector $\mathbf{c}$ instead of $\mathbf{b}$. Then:

$$\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle - \langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{c})\rangle = -\int d^n\mathbf{v}\rho\sigma_e(\mathbf{va})(\sigma_e(\mathbf{v},\mathbf{b}) - \sigma_e(\mathbf{v},\mathbf{c})) \tag{14.4.8}$$

We also know that $\sigma_e^2(\mathbf{v},\mathbf{b}) = \frac{1}{4}$ so that:

$$\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle - \langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{c})\rangle = -\int d^n\mathbf{v}\rho\sigma_e(\mathbf{v},\mathbf{a})(\sigma_e(\mathbf{v},\mathbf{b}) - \sigma_e(\mathbf{v},\mathbf{c}))\cdot 4\sigma_e^2(\mathbf{v},\mathbf{b}) \tag{14.4.9}$$

$$= -\int d^n\mathbf{v}\rho\sigma_e(\mathbf{v},\mathbf{a})\sigma_e(\mathbf{v},\mathbf{b})(1 - 4\underbrace{\sigma_e(\mathbf{v},\mathbf{b})\sigma_e(\mathbf{v},\mathbf{c})}_{\pm\frac{1}{4}}) \tag{14.4.10}$$

It follows that the term in the bracket is either $0$ or $2$. We can overestimate the above integral by taking $\sigma_e(\mathbf{v},\mathbf{a})\sigma_e(\mathbf{v},\mathbf{b}) = \frac{1}{4}$. Then:

$$|\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle - \langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{c})\rangle| \leq \frac{1}{4}\int d^n\mathbf{v}\rho(1 - 4\sigma_e(\mathbf{v},\mathbf{b})\sigma_e(\mathbf{v},\mathbf{c})) \tag{14.4.11}$$

$$= \frac{1}{4}\Big(1 + 4\langle \sigma_e(\mathbf{b})\sigma_p(\mathbf{c})\rangle\Big) \tag{14.4.12}$$

where the $\mathbf{v}$ was removed in the last line in the process of averaging $\sigma_e$ and $\sigma_p$. This is known as Bell's inequality:

$$\boxed{|\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle - \langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{c})\rangle| \leq \frac{1}{4}\Big(1 + 4\langle \sigma_e(\mathbf{b})\sigma_p(\mathbf{c})\rangle\Big)} \tag{14.4.13}$$

It turns out that the predictions of quantum mechanics do not satisfy the above inequality. Indeed plugging in (14.4.4), the LHS reads:

$$|\langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{b})\rangle - \langle \sigma_e(\mathbf{a})\sigma_p(\mathbf{c})\rangle| = |\frac{1}{4}\mathbf{a}\cdot(\mathbf{c} - \mathbf{b})| \tag{14.4.14}$$

whereas the RHS reads:

$$\frac{1}{4}\Big(1 + \mathbf{b}\cdot\mathbf{c}\Big) \tag{14.4.15}$$

Choose $\mathbf{a}, \mathbf{b}$ orthogonal $\mathbf{a} \cdot \mathbf{b} = 0$ and $\mathbf{c} = \mathbf{a} \sin\varphi + \mathbf{b} \cos\varphi$. Then Bell's inequality becomes:

$$|\frac{1}{4} \sin\varphi| \leq |\frac{1}{4}(1 - \cos\varphi)| \tag{14.4.16}$$

which is not satisfied!

Quantum mechanics is therefore inconsistent with the theory of local hidden variables.

There are several implications of Bell's inequality. Firstly, it shows that it simply does not make sense to ask the question "What spin would Alice measure?" beforehand, since the electron itself does not have a definite spin due to a hidden variable until a measurement is made.

## 14.5 Realism and locality

There are two fundamental ideas at the foundation of science, realism and locality.

**Realism and Locality**

**Realism** requires observables to have values independent of any measurement occurring. In other words measurements do not disturb the state of a system.
**Locality** requires that the effect of an event happening at some location at some other location takes at least the time it takes for light to travel between these two locations.

The attractive aspect of local hidden variable theories is that the salvage realism and locality. Bell's inequality however supports quantum mechanics, throwing these two concepts out of the window.

Consider for example the entangled state which we discussed in $\alpha$-decay of uranium-238. Suppose the $\alpha$-particle is in a laboratory set up, where a team of nuclear physicists are studying its properties. Knowing what is happening within the laboratory is is not enough to predict the behaviour of the particle. Indeed, one also needs knowledge of the entangled thorium isotope.

## 14.6 Polarization

To actually test the predictions of quantum mechanics and determine the disagreement with Bell's inequalities, several experiments have been performed. Most of these however haven't used the decaying nuclei but rather the polarization of light.

Despite being spin-1 particles, quantum field theory and several experiments have shown that photons only have two possible eigenvalues for spin, $\pm\hbar$.

Consider a photon moving in the $y$-direciton. We can define the two photon states as $|H\rangle$ and $|V\rangle$ corresponding to linear vertical polarization along the $z$-axis and linear horizontal polarization along the $x$-axis respectively.

Any photon propagating in the $y$-direction therefore is in some superposition:

$$|V_\theta\rangle = \cos\theta\,|V\rangle + \sin\theta\,|H\rangle \tag{14.6.1}$$

where $\theta$ is the angle to the $z$-axis at which the polarizer must be held for the photon to be transmitted 100% of the time. In other words, the photon in $|V_\theta\rangle$ is vertically polarized at an angle $\theta$.

Note that:

$$|V_{\theta+\pi/2}\rangle = |H_\theta\rangle = -\sin\theta\,|V\rangle + \cos\theta\,|H\rangle \tag{14.6.2}$$

Let us now turn to the experiments that verified Bell's inequality, known as the **Aspect experiments**.

Two entangled photons are released from the decay of a calcium atom with zero angular momentum. They are entangled in a state:

$$|P\rangle = \frac{1}{\sqrt{2}}(|V,V\rangle + |H,H\rangle) \tag{14.6.3}$$

Notice that:

$$\begin{aligned}\frac{1}{\sqrt{2}}(|V_\theta,V_\theta\rangle + |H_\theta,H_\theta\rangle) &= \frac{1}{\sqrt{2}}\big[(\cos^2\theta\,|V,V\rangle + \sin^2\theta\,|H,H\rangle + \sin\theta\cos\theta(|H,V\rangle + |V,H\rangle)) \\ &\quad + (\sin^2\theta\,|V,V\rangle + \cos^2\theta\,|H,H\rangle - \sin\theta\cos\theta(|H,V\rangle + |V,H\rangle))\big] \qquad (14.6.4)\\ &= \frac{1}{\sqrt{2}}(|V,V\rangle + |H,H\rangle) \qquad (14.6.5)\end{aligned}$$

so that if one photon is vertically (horizontally) polarized, so is the other photon. Suppose one observer measures the polarization of one photon at some angle $\theta_1$, and similarly we observe the polarization at some angle $\theta_2$.

Now we find that:

$$\langle\sigma_{\theta_1}\sigma_{\theta_2}\rangle = P(V_{\theta_1})P(V_{\theta_2}|V_{\theta_1}) + P(H_{\theta_1})P(H_{\theta_2}|H_{\theta_1}) - P(V_{\theta_1})P(H_{\theta_2}|V_{\theta_1}) - P(H_{\theta_1})P(V_{\theta_2}|H_{\theta_1}) \tag{14.6.6}$$

Since nothing is known about the spin of the photon before the measurement, $P(V_{\theta_1}) = P(H_{\theta_1}) = \frac{1}{2}$.

Just like in the case of entangled electron-positron systems, we find that the probability of measuring vertical polarization at some angle $\theta$ is:

$$P(V_{\theta_2}|V_{\theta_1}) = P(H_{\theta_2}|H_{\theta_1}) = |\,\langle V_{\theta_2}\,|\,V_{\theta_1}\rangle\,|^2 = \cos^2(\theta_2 - \theta_1) \tag{14.6.7}$$

and similarly:

$$P(H_{\theta_2}|V_{\theta_1}) = P(V_{\theta_2}|H_{\theta_1}) = |\,\langle H_{\theta_2}\,|\,V_{\theta_1}\rangle\,|^2 = \sin^2(\theta_2 - \theta_1) \tag{14.6.8}$$

Therefore we find that:

$$\langle \sigma_{\theta_1} \sigma_{\theta_2} \rangle = \cos 2(\theta_2 - \theta_1) \tag{14.6.9}$$

However, Bell's inequality assuming local hidden variables in this experiment would take the form:

$$| \langle \sigma_{\theta_1} \sigma_{\theta_2} \rangle - \left\langle \sigma_{\theta_1} \sigma_{\theta_2'} \right\rangle | \leq |1 + \left\langle \sigma_{\theta_1} \sigma_{\theta_2'} \right\rangle | \tag{14.6.10}$$

Substituting the predictions from quantum mechanics we get:

$$|\cos 2(\theta_2 - \theta_1) - \cos 2(\theta_2' - \theta_1)| \leq |1 + \cos 2(\theta_2' - \theta_1)| \tag{14.6.11}$$

Now taking $\theta_2' = \theta_2 + \theta_1$ we get the inequality:

$$|\cos 2(\theta_2 - \theta_1) - \cos 2\theta_2| \leq |1 + \cos 2\theta_2| \tag{14.6.12}$$

which is clearly not satisfied. Take for example $\theta_2 = \frac{7\pi}{12}$ and $\theta_1 = \frac{\pi}{6}$, so that $\theta_2' = \frac{11\pi}{12}$. Then the LHS reads $\frac{\sqrt{3}}{2}$ wheras the RHS reads $1 - \frac{\sqrt{3}}{2}$.

The Aspect experiments were the first to find what quantum mechanics predicted, thus showing that local hidden variable theories are incompatible with experiment.

## 14.7 Faster than light travel

We still have not resolved the issue of using entanglement to send information at speeds faster than light. However, it turns out that any sort of information sent through entanglement will be unreadable.

Consider for example the following experiment. Alice and Bob entangle two ensembles of 100 electrons each, $\{a_1, a_2, ..., a_{100}\}$ and $\{b_1, b_2, ..., b_{100}\}$. The entanglement is such that if $a_1$ is measured spin-up in some direction **n**, then $a_2$ must be measured spin-down in the same direction.

Bob stays on earth with his electrons, whereas Alice and her electrons move to a planet one light year away. On January 1, 2021 a soccer match between Zimbabwe and Yemen. Before separating, Alice and Bob decided that if Zimbabwe wins, Bob would measure the spin of his electrons in the $z$-direction. Since he cannot control the spin state of the electrons, they will be 50% spin up and 50% spin down. This will collapse the state of Alice's electrons instantaneously into either spin up or spin down in the z-direction. Hence, when Alice measures the spin of the electrons in the $z$ direction, they will be 50% spin up and 50% spin down. If instead Yemen wins, Bob will measure the spin of his electrons in the $y$-direction, collapsing the state of Alice's electrons. If Alice now measures the spin in the $z$-direction, she will still get 50% spin up and 50% spin down. Consequently, the two ensembles of atoms in case Zimbabwe wins and Yemen wins are indistinguishable for Alice's measurements.

One could however argue that Alice and Bob could compare their different measurements. If there was a correlation between Alice and Bob's measurement, so that if one measures spin up the other measures spin down, and vice versa, then Alice will know that she measured the spin in the same direction as Bob, so she will know if Zimbabwe or Yemen won.

If instead there is no correlation, she will know that she measured spin in a direction orthogonal to Bob, and hence she will again infer what direction Bob made his measurement in.

Unfortunately there is yet another flaw in this scheme. To compare their measurements, Alice and Bob would need to communicate, and this requires 2 light years...

Consequently, even though the electrons have collapsed in such a way as to suggest that information has been sent from Bob to Alice, the latter cannot read this message unless she has previous knowledge of the state of Bob's electrons before January 1, 2021. This however cannot be done, since measuring the state of the electrons breaks the entanglement.

## 14.8 Classical bits

A **bit** is any piece of information which can take two possible values. Suppose we flip a coin, and decide to associate the value $0$ with the outcome "heads", and the value $1$ with the outcome "tails". Flipping the coin will then give either a $0$ or a $1$ with certainty, there are two possible outcomes. The result of a coin flip therefore contains 1 bit of information.

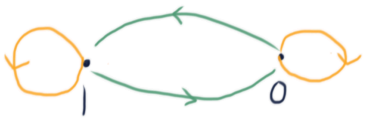
Bits are all **quantum bits**, or **qubits**, but in many scenarios the quantumness of these bits does not manifest itself. Such bits are modelled as **classical bits**.

Adopting dirac notation, we will designate the two states of a classic bit as $|0\rangle$ or $|1\rangle$.

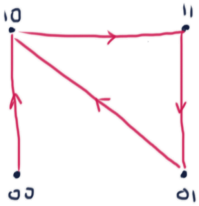
A possible state for a six coin toss could be $|010011\rangle$. There are $2^6$ possible states in this vector space.

Suppose we ask what is the temperature in the lecture room? If we are only interested in an estimate of the number to some degree of precision, then we can write the temperature in base 2, and find the number of bits contained in this information.

We can draw the time evolution of these bits using loop diagrams. For example:




(a) Two possible loop diagrams.




(b) Impossible loop diagram doesn't satisfy unitarity

The yellow loop diagram is representative of a classical coin toss. A coin that is tossed into the state "1" will remain in that state.

The green loop diagram is representative of a system switching states. For example, the stop sign has two possible states, red and green, and alternates between the two.

The red loop diagram for a 2-bit system is not a physically sensible one. Indeed, if the system finds itself in the state $|10\rangle$, it has lost track of where it comes from. This time

evolution therefore does not satisfy unitarity, we cannot construct a classical "propagator" for this system which is unitary.

## 14.9 Quantum cryptography

Encryption consists in rendering a message, originally in binary code known as **plaintext**, into a **cryptogram** through the use of a **cryptographic key**.

For the cryptogram to be uncrackable by an eavesdropper who is not in possession of the key, it was shown by Claude Shannon that the key must be random and as long as the message itself. Keys satisfying these properties are known as **one time pads**.

One problem, known as the **key distribution problem**, quickly arises however: how does the sender send the key to the receiver without revealing it to the eavesdropper?

Indeed, suppose the sender and receiever use a communication channel to transmit information about the key. The eavesdropper, with sufficient computational power, could break through the communication channel and get a copy of the key, thus breaking the encryption of the message. Any method of communication relying on classical physics will be vulnerable, since the eavesdropper can "measure" the key without disturbing it.

However we know that in the realm of quantum mechanics reality is not realizable, in other words the eavesdropper cannot measure the key without disturbing it. Furthermore, we also know that entangled photons violate non-locality, and may also be used in the distribution of keys, since measuring the key would immediately break the entanglement of these photons.

## 14.10 BB84 Protocol

The BB84 is one the most popular protocols which exploit quantum mechanics to solve the key distribution problem.

Suppose Alice and Bob wish to create a cryptographic key through a quantum communication channel. They also have a classical (insecure) communication channel to their disposal.

(i) Alice chooses either the $H/V$ basis or the $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ basis to transmit information over the quantum communication channel.

(ii) Alice sends one bit of information (either a 1 or a 0) in the chosen basis.

(iii) Bob chooses either the $H/V$ basis or the $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ basis to receive information.

(iv) Repeat (i)-(iii) until sufficient bits of information have been transmitted.

(v) In the classical channel, Alice and Bob tell each other what bases they used for each bit of information. Any bit where they didn't use the same basis is discarded.

(vi) To test for eavesdroppers, Alice and Bob choose a subset of the transmitted bits, and if they match, use the remaining bits to read the cryptographic key.

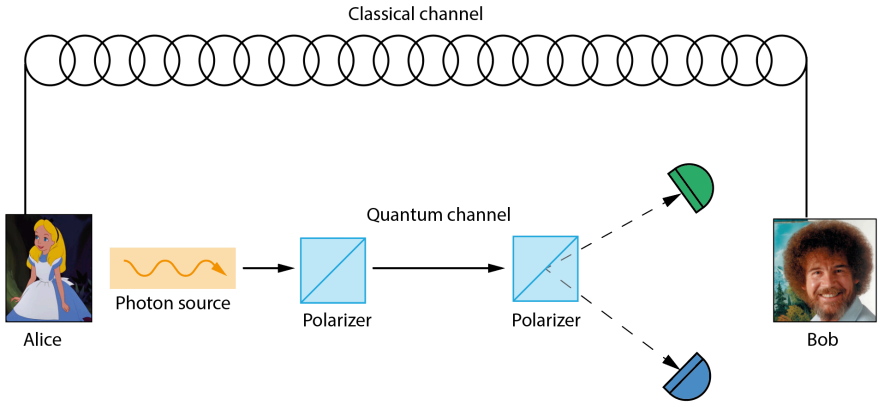



**Figure 14.2**. Apparatus for the BB84 protocol between Alice in the wonderland and Bob Ross.

Let's examine the steps carefully.

Firstly, Alice chooses one of the two available bases for the photon polarization. In the $H/V$ basis for example, if Alice wants to send a 1 then she will have to send a photon in the state $|V\rangle$. Bob must similarly choose a basis for detection. Here are the possible outcomes of Bob's measurements:

| A basis | $H/V$ | $H/V$ | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ |
|---|---|---|---|---|---|---|---|---|
| A sent bit | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| B basis | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ |
| B received bit | 1 | 0 | 1/0 | 1/0 | 1/0 | 1/0 | 1 | 0 |

where $1/0$ means that there is an equal probability of measuring bit 1 or bit 0. Next, Alice and Bob share what basis they used for each photon, and since they can only see a 100% correlation when they both use the same basis, they discard photons where different bases were employed.

In the absence of the eavesdropper, Alice and Bob would then use the remaining bits, which are necessarily copies of each other, may then be used as cryptographic keys.

But what if an eavesdropper Eve were to intercept the photons and measure their polarization in either the $H/V$ or the $|H_{\frac{\pi}{4}}\rangle$ bases? Since Eve does not yet know what bases Alice and Bob have used for each photon, there is a chance that her basis will not coincide with Alice's base. Consequently, when Eve sends back a photon with the measured polarization to Bob (in order to mitigate her measurement), there will be a chance for this photon's polarization to not coincide with the polarization of the photon sent by Alice. Hence, even if Bob and Alice use the same basis, due to the intervention of Eve, their bits may not coincide!

For example, suppose both Alice and Bob use the $H/V$ basis, and the eavesdropper uses

the $|H_{\frac{\pi}{4}}\rangle$ basis. If Alice sends 1, so a photon in state $|V\rangle$, then there is a possibility (50%) that the Eve will measure $|H_{\frac{\pi}{4}}\rangle$, and send a photon in this state to Bob. Bob will therefore have a probability of measuring either $|V\rangle$ or $|H\rangle$, and in the latter case this will not coincide with what Alice sent initially. When Alice and Bob compare their results, they will notice that Eve intervened.

| A basis | $H/V$ | $H/V$ | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ |
|---|---|---|---|---|---|---|---|---|
| A sent bit | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| E basis | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ |
| E received bit | 1 | 0 | 1/0 | 1/0 | 1/0 | 1/0 | 1 | 0 |
| B basis | $H/V$ | $H/V$ | $H/V$ | $H/V$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ | $H_{\frac{\pi}{4}}/V_{\frac{\pi}{4}}$ |
| B received bit | 1 | 0 | 1/0 | 1/0 | 1/0 | 1/0 | 1 | 0 |

Hence, if Alice sends 100 photons for each bit in each basis, so that she sends 800 photons in total, then 600 of these photons will coincide, and 200 will not coincide. Hence there is a 75% chance that even though Eve is eavesdropping, the photon polarization that Alice sends will eventually be measured by Bob.

## 14.11 Eckert Protocol

Again, Alice and Bob need to establish a cryptographic key. They have a source of entangled pairs of photons in the state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|V_\theta\rangle_A |H_\theta\rangle_B - |H_\theta\rangle_A |V_\theta\rangle_B) \tag{14.11.1}$$

(i) Alice chooses randomly a basis from $\theta = 0, \frac{\pi}{8}, \frac{\pi}{4}$. Similarly, Bob chooses randomly a basis from $\theta = -\frac{\pi}{8}, \frac{\pi}{8}, \frac{\pi}{4}$.

(ii) Alice and Bob make measurements on the polarization of their photons in their respective bases.

(iii) Alice and Bob reveal which basis they used for each photon over the classical communication channel.

(iv) Alice and Bob classify their measurement into group 1, where they employed different bases, and group 2, where they employed the same basis.

(v) Using the group 1 measurements, Alice and Bob determine whether or not quantum non-local correlations have been exhibited. If so, they use group 2 measurements to create a cryptographic key.

Clearly, if Alice and Bob use the same value of $\theta$, their measurements will be anti-correlated.

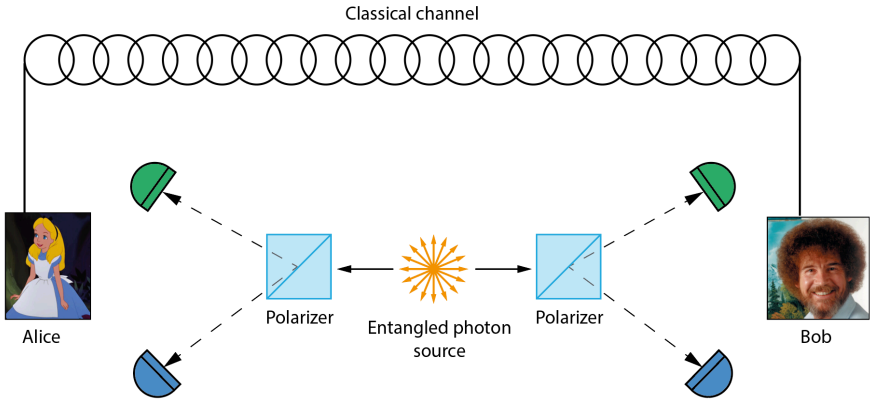



**Figure 14.3**. Apparatus for the Eckert protocol between Alice in the wonderland and Bob Ross

More generally, if Alice chooses $\alpha$, and Bob chooses $\beta$ as their polarizers' axes orientation, then we find that:

$$P_{++}(\alpha,\beta) = |\langle V_\alpha, V_\beta | \psi\rangle|^2 = \frac{1}{2}|\cos\alpha\sin\beta - \sin\alpha\cos\beta|^2 = \frac{1}{2}\sin^2(\alpha-\beta) \qquad (14.11.2)$$

Similarly:

$$P_{--}(\alpha,\beta) = |\langle H_\alpha, H_\beta | \psi\rangle|^2 = \frac{1}{2}|\sin\alpha\cos\beta - \cos\alpha\sin\beta|^2 = \frac{1}{2}\sin^2(\alpha-\beta) \qquad (14.11.3)$$

Also:

$$P_{+-}(\alpha,\beta) = P_{-+}(\alpha,\beta) = \frac{1}{2}\cos^2(\alpha-\beta) \qquad (14.11.4)$$

It follows that if $\alpha = \beta$, then $P_{++} = P_{--} = 0$ and $P_{+-} = P_{-+} = 1$, giving perfect anti-correlation. Hence, in the absence of an eavesdropper, if Alice and Bob share what basis they used, they know that the measurements in group 2 (where they employed the same basis) will be perfect copies of each other with $1 \leftrightarrow 0$.

But what if the eavesdropper Eve were to intercept the photons and measure their polarization? We can define the correlation function as the probability of correlation minus the probability of anti-correlation:

$$C(\alpha_i,\beta_j) = \sin^2(\alpha_i-\beta_j) - \cos^2(\alpha_i-\beta_j) = -\cos 2(\alpha_i-\beta_j) \qquad (14.11.5)$$

This is where group 1, the group of measurements where different bases were used, comes into play. Indeed, using group 1 measurements Alice and Bob can evaluate the following expression:

$$S = C(\alpha_1,\beta_1) + C(\alpha_1,\beta_3) + C(\alpha_3,\beta_1) - C(\alpha_3,\beta_3) \qquad (14.11.6)$$

Setting $\alpha_1 = 0, \alpha_3 = \frac{\pi}{4}, \beta_1 = \frac{\pi}{8}, \beta_3 = -\frac{\pi}{8}$, we find that:

$$S = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = -2\sqrt{2} \qquad (14.11.7)$$

If Eve eavesdrops and performs a polarization measurement, then she could create a new entangled pair. However, Alice and Bob's measurements for the new pair, which will be used to create the key, will not necessarily coincide with what Eve measured for the original pair.

So, Eve could try creating a pair of photons with polarization coinciding with her measurement, so for example $|V\rangle_A |H\rangle_B$. However, this has collapsed the state vector and broken the entanglement. The non-local correlations we derived earlier will no longer be present. The measurements performed by Alice will in no way depend or be correlated on what Bob measures.

Indeed, we find that:

$$P_{++}(\alpha,\beta) = |\langle V_\alpha, V_\beta | V, H\rangle|^2 = |\cos\alpha \sin\beta|^2 = \cos^2\alpha \sin^2\beta \tag{14.11.8}$$

Similarly:

$$P_{--}(\alpha,\beta) = |\langle H_\alpha, H_\beta | V, H\rangle|^2 = |-\sin\alpha \cos\beta|^2 = \cos^2\beta \sin^2\alpha \tag{14.11.9}$$

Also:

$$P_{+-}(\alpha,\beta) = \cos^2\alpha \cos^2\beta \tag{14.11.10}$$
$$P_{-+}(\alpha,\beta) = \sin^2\alpha \sin^2\beta \tag{14.11.11}$$

so that:

$$C'(\alpha_i,\beta_j) = \cos^2\alpha \sin^2\beta + \cos^2\beta \sin^2\alpha - \cos^2\alpha \cos^2\beta - \sin^2\alpha \sin^2\beta \tag{14.11.12}$$
$$= -\cos(2\alpha_i)\cos(2\beta_j) \tag{14.11.13}$$

Again, substituting $\alpha_1 = 0, \alpha_3 = \frac{\pi}{4}, \beta_1 = \frac{\pi}{8}, \beta_3 = -\frac{\pi}{8}$ we find that:

$$S' = -\sqrt{2} < S \tag{14.11.14}$$

So, we see that in the case where Eve has interfered, the non-local behaviour is disrupted.

## 14.12 Quantum teleportation

A loop immediately surfaces from our considerations of quantum cryptography protocols: what if the eavesdropper makes an exact copy of the intercepted photons *before* a measurement is made.

Is it possible to create an exact copy, a **clone** of an unknown quantum state? It turns out that this is impossible, as we shall now prove.

**No-cloning theorem**

It is impossible to clone quantum states.

*Proof.* Consider two quantum systems $A, B$ sharing the same Hilbert space $\mathcal{H}$, and suppose we wish to transform a system beginning in the state $|\psi\rangle_A |\phi\rangle_B$ to the state $|\psi\rangle_a |\psi\rangle_B$.

To do so, we need to create a propagator $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$ such that:

$$\hat{U} |\psi\rangle_A |\phi\rangle_B = e^{i\alpha} |\psi\rangle_A |\psi\rangle_B \tag{14.12.1}$$

where the phase factor does not change the "physicality" of the state.

Let $|\varphi\rangle_A$ be another state in $\mathcal{H}$. Then:

$$\begin{aligned}
\langle\psi \,|\, \varphi\rangle \langle\phi \,|\, \phi\rangle &= \langle\psi|_A \langle\phi|_B |\varphi\rangle_A |\phi\rangle_B && (14.12.2)\\
&= (\hat{U} |\varphi\rangle_B |\psi\rangle_A)^\dagger \hat{U} |\varphi\rangle_A |\phi\rangle_B && (14.12.3)\\
&= e^{-i\beta} \langle\psi|_A \langle\psi|_B |\varphi\rangle_A |\varphi\rangle_B && (14.12.4)\\
&= e^{-i\beta} \langle\psi \,|\, \varphi\rangle^2 && (14.12.5)\\
& && (14.12.6)
\end{aligned}$$

Since $|\phi\rangle$ is normalized, we find that:

$$\langle\psi \,|\, \varphi\rangle = e^{-i\beta} \langle\psi \,|\, \varphi\rangle^2 \tag{14.12.7}$$

implying that either $\langle\psi \,|\, \varphi\rangle = 0$ or $\langle\psi \,|\, \varphi\rangle = e^{i\beta}$. However, this cannot be possible true for any arbitrary state $\varphi$. It follows that we cannot create a propagator cloning a quantum state. □

Although it is not possible to clone a state, it turns out that we can teleport a state.

Suppose Alice wishes to teleport an unknown state $\psi$ of photon 1, to Bob, so that he may recreate the photon with this information. In general:

$$|\psi\rangle_1 = \alpha |V\rangle_1 + \beta |H\rangle_1 \tag{14.12.8}$$

Suppose we also have a source of entangled photons, producing photons 2 and 3 in the state:

$$|\phi\rangle_{23} = \frac{1}{\sqrt{2}}(|V\rangle_2 |H\rangle_3 - |H\rangle_2 |V\rangle_3) \tag{14.12.9}$$

The entangled photons are emitted so that photon 2 is sent to Alice and photon 3 is sent to Bob.

Since photons 2 and 3 are entangled, if Alice were able to couple photon 1 and 2, then Bob would be able to use quantum non-locality to receive information about $|\psi\rangle_1$.

But how does Alice accomplish this? Firstly note that the state of the entire system consisting of the three photons is:

$$|\Psi\rangle_{123} = \frac{\alpha}{\sqrt{2}}(|V\rangle_1 |V\rangle_2 |H\rangle_3 - |V\rangle_1 |H\rangle_2 |V\rangle_3) + \frac{\beta}{\sqrt{2}}(|H\rangle_1 |V\rangle_2 |H\rangle_3 - |H\rangle_1 |H\rangle_2 |V\rangle_3) \tag{14.12.10}$$

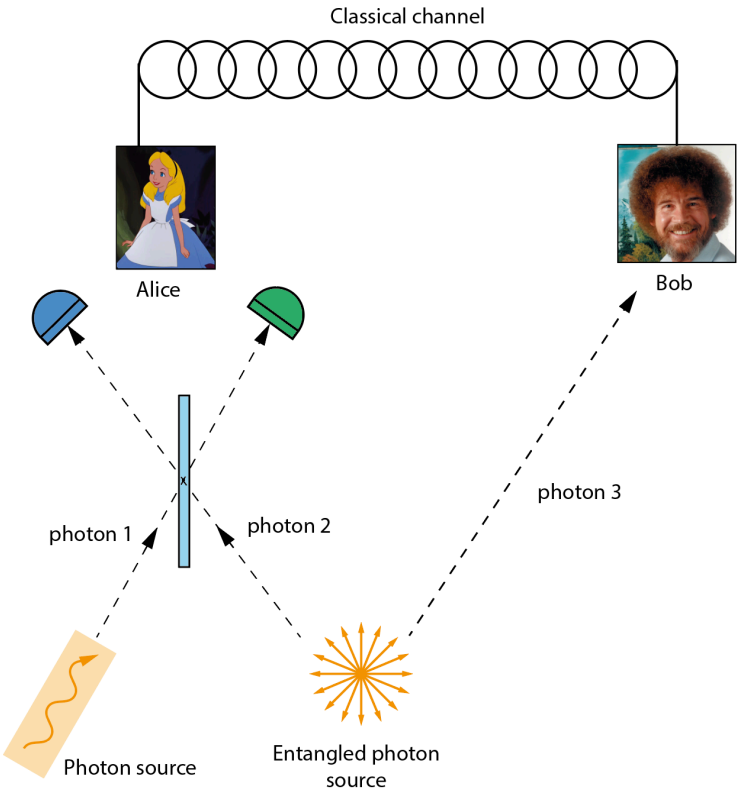



**Figure 14.4**. Apparatus for quantum teleportation of a photon from Alice in the wonderland to Bob Ross

Let us define the following entangled states for photons 1,2:

$$|\Psi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|V\rangle_1 |H\rangle_1 \pm |V\rangle_2 |H\rangle_2) \tag{14.12.11}$$

$$|\Phi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|V\rangle_1 |V\rangle_1 \pm |H\rangle_2 |H\rangle_2) \tag{14.12.12}$$

We may then write $|\Psi\rangle_{123}$ in terms of the above states, known as **Bell states**:

$$\begin{aligned} |\Psi\rangle_{123} = \frac{1}{2}\Big( & |\Psi^-\rangle_{12}(-\alpha|V\rangle_3 - \beta|H\rangle_3) + |\Psi^+\rangle_{12}(-\alpha|V\rangle_3 + \beta|H\rangle_3) \\ & + |\Phi^-\rangle_{12}(\beta|V\rangle_3 \alpha|H\rangle_3) + |\Phi^+\rangle_{12}(-\beta|V\rangle_3 + \alpha|H\rangle_3)\Big) \end{aligned} \tag{14.12.13}$$

In other words, the state of the combined system is a linear combination of the Bell states for photons 1 and 2, with coefficients given by the state of photon 3. These coefficients are all different, so if Alice were to perform a **Bell measurement**, that is a measurement that would collapse $|\Psi\rangle_{123}$ into one of the Bell states, then this would also collapse the state of photon 3 into one of the four coefficients.

The table below shows how the Bell measurements performed by Alice on photons 1 and

2 affect the state of photon 3 measured by Bob:

| Measurement | $\|\phi\rangle_3$ |
|---|---|
| $\|\Psi^-\rangle_{12}$ | $-\alpha\|V\rangle_3 - \beta\|H\rangle_3$ |
| $\|\Psi^+\rangle_{12}$ | $-\alpha\|V\rangle_3 + \beta\|H\rangle_3$ |
| $\|\Phi^-\rangle_{12}$ | $\beta\|V\rangle_3\,\alpha\|H\rangle_3$ |
| $\|\Phi^+\rangle_{12}$ | $-\beta\|V\rangle_3 + \alpha\|H\rangle_3$ |

So, if Alice tells Bob which of the Bell states she measured over a classical communication channel (this requires only 2 bits of information, since there are 4 possible Bell states), then Bob knows the shape of the state of photon 3. By performing polarization measurement, he may then determine $\alpha$ and $\beta$ by applying appropriate transformations, and hence reconstruct the original state of photon 1.

| $\|\phi\rangle_3$ | Transformation |
|---|---|
| $-\alpha\|V\rangle_3 - \beta\|H\rangle_3$ | $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ |
| $-\alpha\|V\rangle_3 + \beta\|H\rangle_3$ | $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ |
| $\beta\|V\rangle_3\,\alpha\|H\rangle_3$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ |
| $-\beta\|V\rangle_3 + \alpha\|H\rangle_3$ | $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ |

One can easily verify that all these matrices are unitary, and therefore preserve the normalization of $|\phi\rangle_3$.

For example, if Alice measures $|\Phi^-\rangle$, then Bob would then perform a measurement on photon 3. He then operates the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on the photon state, creating an identical copy of $|\psi\rangle_1$. In the process of doing so, Alice has not learnt anything about $|\psi\rangle_1$, since her measurement collapsed the state of photons 1,2. Consequently, the no-cloning theorem was not violated, since the original photon state was destroyed.

Bob however may be compelled to send information about $|\psi_1\rangle$ to Alice, thus violating the no-cloning theorem. However, to do so, Bob would have to repeat this protocol of quantum teleportation, and thus destroys his own copy of photon 1.

## 14.13 Performing Bell measurements

There remains one ambiguity in this procedure, how exactly does Alice perform Bell measurements on photons 1 and 2?

Recall that a beam splitter splits a beam of light by reflecting one half and reflecting the other half. Since the two photons are all incident on the beam splitter in two distinct directions, we may define the possible input spacial states of the photons on the beam syplitter as $|a\rangle$ and $|b\rangle$ as shown below.

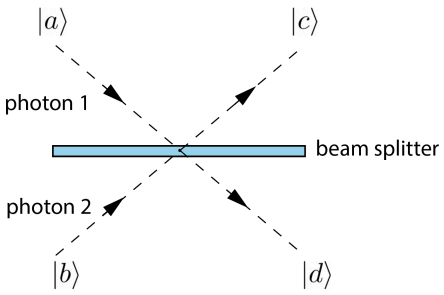



**Figure 14.5**. Input states $|a\rangle$ and $|b\rangle$, and output states $|c\rangle$ and $|d\rangle$ of a 50-50 beam splitter

The beam splitter will have the following effect:

$$|a\rangle \longrightarrow \frac{1}{\sqrt{2}}(i\,|c\rangle + |d\rangle) \tag{14.13.1}$$

$$|b\rangle \longrightarrow \frac{1}{\sqrt{2}}(|c\rangle + i\,|d\rangle) \tag{14.13.2}$$

where $i$ is due to the phase shift in light when reflected.

So, if a photon in the state $\frac{1}{\sqrt{2}}(|V\rangle - |H\rangle) \otimes |b\rangle$ is incident on the beam splitter, it will transform into $\frac{1}{2}(|V\rangle - |H\rangle) \otimes (|c\rangle + i\,|d\rangle)$.

In the apparatus of quantum teleportation, we see that photon 1 is arranged to be in the input $|a\rangle$ and photon 2 is arranged to be in the input $|b\rangle$. Then we may write the incoming state of photons 1 and 2 as:

$$|\Phi\rangle_{in} = (\alpha\,|V\rangle_1 + \beta\,|H\rangle_1)\,|a\rangle_1\,(\gamma\,|V\rangle_2 + \delta\,|H\rangle_2)\,|b\rangle_2 \tag{14.13.3}$$

After hitting the beam splitter, we find that:

$$|\Phi\rangle_{12} = \frac{1}{2}(\alpha\,|V\rangle_1 + \beta\,|H\rangle_1)(i\,|c\rangle_1 + |d\rangle_1)(\gamma\,|V\rangle_2 + \delta\,|H\rangle_2)(|c\rangle_2 + i\,|d\rangle_2) \tag{14.13.4}$$

However, note that due to exchange degeneracy, the photons leaving the beam splitter are indistinguishable. We must therefore construct the bosonic state:

$$|\Phi\rangle_{out} = \frac{1}{\sqrt{2}}(|\Phi\rangle_{12} + |\Phi\rangle_{21}) \tag{14.13.5}$$

Some lengthy algebra "simplifies" the above expression to:

$$
\begin{aligned}
|\Phi\rangle_{out} =& i\frac{\alpha\gamma+\beta\delta}{\sqrt{2}}\frac{1}{\sqrt{2}}(|c\rangle_1|c\rangle_2+|d\rangle_1|d\rangle_2)\frac{1}{\sqrt{2}}(|V\rangle_1|V\rangle_2+|H\rangle_1|H\rangle_2) \qquad (14.13.6)\\
&+i\frac{\alpha\gamma-\beta\delta}{\sqrt{2}}\frac{1}{\sqrt{2}}(|c\rangle_1|c\rangle_2+|d\rangle_1|d\rangle_2)\frac{1}{\sqrt{2}}(|V\rangle_1|V\rangle_2-|H\rangle_1|H\rangle_2)\\
&+i\frac{\alpha\delta+\beta\gamma}{\sqrt{2}}\frac{1}{\sqrt{2}}(|c\rangle_1|c\rangle_2+|d\rangle_1|d\rangle_2)\frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2+|H\rangle_1|V\rangle_2)\\
&+i\frac{\alpha\delta-\beta\gamma}{\sqrt{2}}\frac{1}{\sqrt{2}}(|c\rangle_1|d\rangle_2+|c\rangle_1|d\rangle_2)\frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2-|H\rangle_1|V\rangle_2)
\end{aligned}
\qquad (14.13.7)
$$

This is simply the superposition of Bell states for both the position and polarization part of the state vector for photons 1 and 2. For example, the first three lines correspond to the photons being detected on the same side, either of the beam splitter, either $|c\rangle$ or $|d\rangle$. The last line instead correponds to the photons being detected on opposite sides.

Suppose Alice places two detectors on the output modes of the beam splitter. Whenether both detectors have been activated, this will correspond to the photons being detected on different output states, that is they will have collapsed in the state described by the fourth line in $|\Phi\rangle_{out}$:

$$
i\frac{\alpha\delta-\beta\gamma}{\sqrt{2}}\frac{1}{\sqrt{2}}(|c\rangle_1|d\rangle_2+|c\rangle_1|d\rangle_2)\frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2-|H\rangle_1|V\rangle_2) \qquad (14.13.8)
$$

The polarization part corresponds to $|\Phi^-\rangle_{12}$, so $|\Phi\rangle_{123}$ must have collapsed to the first term in (14.12.13). Bob will then apply the necessarily transformation on photon 3, effectively teleporting the photon.

Note that if only one of the detectors on the beam splitter is activated, then we cannot distinguish between the three possible Bell states. More complex apparati are then needed.

# Part III

# Approximation methods

# Time-independent Perturbation theory

15

## 15.1 The problem

We have learned to solve several exact systems, such as the harmonic oscillator, the step potential, and the hydrogen atom. Unfortunately, in real life these exact systems come up on extremely rare occasions. The systems studied in modern research are significantly more complex, but can usually be regarded as a perturbed hamiltonian, that is, an exactly solvable hamiltonian plus a smaller perturbation hamiltonian.

For example, most potentials may be approximated as harmonic oscillators at their minimum, with a small anharmonic perturbation.

Whenever we have a Hamiltonian:

$$\hat{H} = \hat{H}^{(0)} + \delta\hat{H} \tag{15.1.1}$$

where $\hat{H}^{(0)}$ is a well known hamiltonian, and $\delta\hat{H}$ is the perturbation we impose on the system, we may use perturbation theory. However, we can cleverly insert a parameter $\lambda$ in front of the perturbation term:

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\delta\hat{H},\ \lambda \in [0,1] \tag{15.1.2}$$

It appears like we're shooting ourselves in the foot as we're solving an even more general problem of diagonalizing a family of Hamiltonians, but the introduction of $\lambda$ will allow us to set up the perturbation equations more easily.

Let us plot the energy spectrum of the unperturbed and perturbed hamiltonians:

We plot the energy spectrum of $\hat{H}^{(0)}$ on the $y$-axis, and let $\lambda$ evolve on the $x$-axis. We see that non-degenerate states will behave differently from degenerate states. Indeed the first excited state, which is degenerate, may split due to the perturbation. $\delta\hat{H}$ may cause one state to have a higher energy than the other. One must therefore differentiate between degenerate and non-degenerate perturbation theory. The ground state in this example may be analyzed using the latter, whereas the first excited state will require the former.

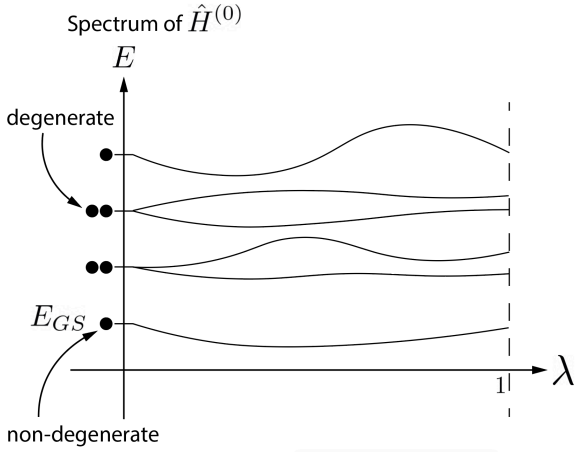



**Figure 15.1**. Evolution of energy spectra with $\lambda$

## 15.2 Non-degenerate perturbation theory

Suppose we have fully diagonalized $\hat{H}^0$, so we know the eigenstates $|k^{(0)}\rangle$ and the associated eigenvalues so:

$$\hat{H}^{(0)}\,|k^{(0)}\rangle = E_k^{(0)}\,|k^{(0)}\rangle \tag{15.2.1}$$

Suppose we also order the energy spectrum:

$$E_0^{(0)} \leq E_1^{(0)} \leq E_2^{(0)} \leq ... \tag{15.2.2}$$

If the eigenstate $|n^{(0)}\rangle$ is non-degenerate then we will not have equality so:

$$E_{n-1}^{(0)} < E_n^{(0)} < E_{n+1}^{(0)} \tag{15.2.3}$$

Let us label the perturbed hamiltonian's eigenstates as $|n\rangle_\lambda$ (we need the lambda since the spectra will evolve with $\lambda$ generally) so that the equation we wish to solve is:

$$\hat{H}(\lambda)\,|n\rangle_\lambda = E_n(\lambda)\,|n\rangle_\lambda \tag{15.2.4}$$

It is clear that $|n\rangle_0 = |n^{(0)}\rangle$ when $\lambda = 0$. We can Taylor expand about $\lambda$ to find:

$$|n\rangle_\lambda = |n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \lambda^2\,|n^{(2)}\rangle + ... + \lambda^k\,|n^{(k)}\rangle + ... \tag{15.2.5}$$

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + ... + \lambda^k E_n^{(k)} + ... \tag{15.2.6}$$

We can substitute this into the TISE:

$$(\hat{H}^{(0)} + \lambda\delta\hat{H} - E_n(\lambda))\,|n\rangle_\lambda = 0 \tag{15.2.7}$$

$$\iff ((\hat{H}^{(0)} - E_n^{(0)}) - \lambda(E_n^{(1)} - \delta\hat{H}) - \lambda^2 E_n^{(2)} - .... - \lambda^k E_n^{(k)}) \tag{15.2.8}$$

$$(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \lambda^2\,|n^{(2)}\rangle + ... + \lambda^k\,|n^{(k)}\rangle + ...) = 0 \tag{15.2.9}$$

It is important to note that(15.2.7) will be a polynomial equation in $\lambda$. Moreover, since it

must hold for all $\lambda$ in an interval range, we need the coefficients of all powers of $\lambda$ to vanish.

Hence:

$$\lambda^{(0)} : (\hat{H}^{(0)} - E_n^{(0)}) \, |n^{(0)}\rangle = 0 \tag{15.2.10}$$

$$\lambda^{(1)} : (\hat{H}^{(0)} - E_n^{(0)}) \, |n^{(1)}\rangle = (E_n^{(1)} - \delta\hat{H}) \, |n^{(0)}\rangle \tag{15.2.11}$$

$$\vdots$$

$$\lambda^{(k)} : (\hat{H}^{(0)} - E_n^{(0)}) \, |n^{(k)}\rangle = (E_n^{(1)} - \delta\hat{H}) \, |n^{(k-1)}\rangle + E_n^{(2)} \, |n^{(k-2)}\rangle + ... + E_n^{(k)} \, |n^{(0)}\rangle \tag{15.2.12}$$

$$\vdots$$

These can be solved recursively, starting from $|n^{(0)}\rangle$ we can find $|n^{(1)}\rangle$, from which we find $|n^{(2)}\rangle$ whose equation involves $|n^{(0)}\rangle$ and $|n^{(1)}\rangle$, and so forth...

Now suppose that $|n^{(1)}\rangle$ has a component along $|n^{(0)}\rangle$, say $\alpha \, |n^{(0)}\rangle$. Now observe that since $(\hat{H}^{(0)} - E_n^{(0)}) \, |n^{(0)}\rangle = 0$, we will have that:

$$(\hat{H}^{(0)} - E_n^{(0)})(|n^{(1)}\rangle + c \, |n^{(0)}\rangle) = (\hat{H}^{(0)} - E_n^{(0)}) \, |n^{(1)}\rangle = E_n^{(1)} - \delta\hat{H}) \, |n^{(0)}\rangle \tag{15.2.13}$$

so $|n^{(1)}\rangle + c \, |n^{(0)}\rangle$ will be a solution of the perturbative equation. Hence, we may set $c = -\alpha$ and remove the component along $|n^{(0)}\rangle$. We may do so for all states $|n^{(i)}\rangle$ so that

$$\left\langle n^{(0)} \middle| n^{(i)} \right\rangle = 0, \; 1 \leq i \tag{15.2.14}$$

Let's now take (15.2.10) and dot it with $\langle n^{(0)}|$:

$$\left\langle n^{(0)} \middle| \hat{H}^{(0)} - E_n^{(0)} \right\rangle = E_n^{(1)} - \left\langle n^{(0)} \middle| \delta\hat{H} \middle| n^{(0)} \right\rangle \tag{15.2.15}$$

Since $\left\langle n^{(0)} \middle| \hat{H}^{(0)} - E_n^{(0)} \right\rangle = E_n^{(0)} - E_n^{(0)} = 0$ we then find that:

$$\boxed{E_n^{(1)} = \left\langle n^{(0)} \middle| \delta\hat{H} \middle| n^{(0)} \right\rangle} \tag{15.2.16}$$

Let's redo this for (15.2.12) keeping in mind the orthogonality of $|n^{(k)}\rangle$ states with $|n^{(0)}\rangle$:

$$0 = -\left\langle n^{(0)} \middle| \delta\hat{H} \middle| n^{(k-1)} \right\rangle + E_n^{(k)} \implies \boxed{E_n^{(k)} = \left\langle n^{(0)} \middle| \delta\hat{H} \middle| n^{(k-1)} \right\rangle} \tag{15.2.17}$$

We have therefore found the first order energy correction. However, to find higher order terms we need to find the the associated eigenstates, starting with $|n^{(1)}\rangle$. We can dot (15.2.11) with $\langle k^{(0)}|$ where $k \neq n$. Then:

$$\left\langle k^{(0)} \middle| \hat{H}^{(0)} - \delta\hat{H} \middle| n^{(1)} \right\rangle = -\left\langle k^{(0)} \middle| \delta\hat{H} \middle| n^{(0)} \right\rangle \tag{15.2.18}$$

so that:

$$(E_k^{(0)} - E_n^{(0)}) \left\langle k^{(0)} \middle| n^{(1)} \right\rangle = -\left\langle k^{(0)} \middle| \delta\hat{H} \middle| n^{(0)} \right\rangle \tag{15.2.19}$$

For sake of brevity let $\delta H_{kn} = \left\langle k^{(0)} \middle| \delta\hat{H} \middle| n^{(0)} \right\rangle$:

$$\left\langle k^{(0)} \middle| n^{(1)} \right\rangle = -\frac{\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.20}$$

Then using the resolution of the identity:

$$|n^{(1)}\rangle = \sum_k |k^{(0)}\rangle \left\langle k^{(0)} \middle| n^{(1)} \right\rangle \tag{15.2.21}$$

but since we established that $\left\langle n^{(0)} \middle| n^{(1)} \right\rangle = 0$ we get:

$$|n^{(1)}\rangle = -\sum_{k\neq n} |k^{(0)}\rangle \frac{\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.22}$$

This is the first order correction to the eigenstates. We can similarly find higher order terms. This allows us to find the second order energy correction:

$$E_n^{(2)} = \left\langle n^{(0)} \middle| \delta\hat{H} \middle| n^{(1)} \right\rangle = -\sum_{k\neq n} \left\langle n^{(0)} \middle| \delta\hat{H} \middle| k^{(0)} \right\rangle \frac{\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.23}$$

$$= -\sum_{k\neq n} \frac{\delta H_{nk}\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.24}$$

Since the Hamiltonian must be Hermitian, the perturbation must also be Hermitian so that $\delta H_{kn} = (\delta H_{nk})^*$ and thus:

$$E_n^{(2)} = -\sum_{k\neq n} \frac{|\delta H_{nk}|^2}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.25}$$

In conclusion:

$$|n\rangle_\lambda = |n^{(0)}\rangle - \lambda \sum_{k\neq n} \frac{\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} |k^{(0)}\rangle + o(\lambda^2) \tag{15.2.26a}$$

$$E_n(\lambda) = E_n^{(0)} - \lambda\delta H_{nn} - \lambda^2 \sum_{k\neq n} \frac{|\delta H_{nk}|^2}{E_k^{(0)} - E_n^{(0)}} + o(\lambda^3) \tag{15.2.26b}$$

Interestingly, (15.2.26b) always overestimates the real ground state energy $E_0(\lambda)$. Indeed when $n = 0$ (remember we ordered the energy spectrum so that the lowest energy comes first) (15.2.26b) gives:

$$E_0^{(0)} \approx E_0^{(0)} + \lambda\delta H_{00} = \left\langle 0^{(0)} \middle| \hat{H}^{(0)} \middle| 0^{(0)} \right\rangle + \lambda \left\langle 0^{(0)} \middle| \delta\hat{H} \middle| 0^{(0)} \right\rangle \tag{15.2.27}$$

$$= \left\langle 0^{(0)} \middle| \hat{H}^{(0)} + \lambda\delta\hat{H} \middle| 0^{(0)} \right\rangle \tag{15.2.28}$$

$$= \left\langle 0^{(0)} \middle| \hat{H}(\lambda) \middle| 0^{(0)} \right\rangle \qquad \geq E_0(\lambda) \tag{15.2.29}$$

where we used the variational principle in the last line. Indeed for the ground state the

second order correction:

$$-\lambda^2 \sum_{k>0} \frac{|\delta H_{0k}|^2}{E_k^{(0)} - E_0^{(0)}} < 0 \tag{15.2.30}$$

since $E_k^{(0)} > E_0^{(0)}$.

More generally, for the $n$th energy state the second order correction may be split into two sums:

$$-\lambda^2 \sum_{k\neq n} \frac{|\delta H_{nk}|^2}{E_k^{(0)} - E_n^{(0)}} = \lambda^2 \sum_{k<n} \frac{|\delta H_{nk}|^2}{E_n^{(0)} - E_k^{(0)}} - \lambda^2 \sum_{k>n} \frac{|\delta H_{nk}|^2}{E_k^{(0)} - E_n^{(0)}} \tag{15.2.31}$$

Since the energy spectrum was ordered the first term will be positive, whereas the second will be negative due to the negative sign. Hence the states higher than $n$ will push the energy level up, whereas the states lower than $n$ will push the energy level down.

Suppose we have a hamiltonian:

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda \hat{V} = \begin{pmatrix} E_1^{(0)} & 0 \\ 0 & E_2^{(0)} \end{pmatrix} + \lambda \begin{pmatrix} 0 & V \\ V^* & 0 \end{pmatrix} \tag{15.2.32}$$

The eigenvalues of this matrix are:

$$E \pm (\lambda) = \frac{E_1^{(0)} + E_2^{(0)}}{2} \pm \frac{E_1^{(0)} - E_2^{(0)}}{2} \sqrt{1 + \frac{\lambda^2 |V|^2}{\left(\frac{E_1^{(0)} - E_2^{(0)}}{2}\right)^2}} \tag{15.2.33}$$

Now we may use the complex taylor expansion

$$f(\epsilon) = \sqrt{1+\epsilon^2} = 1 + \frac{\epsilon^2}{2} - \frac{\epsilon^4}{8} + \frac{\epsilon^6}{16} + o(\epsilon^8) \tag{15.2.34}$$

The radius of convergence of this expansion is 1 so that for fast convergence we require $|\epsilon| \ll 1$, that is:

$$|\lambda V| < \frac{|E_1^{(0)} - E_2^{(0)}|}{2} \tag{15.2.35}$$

So we need the perturbation to be smaller than the energy level differences, not only the energy levels themselves. More generally we require:

$$\left| \frac{\delta H_{kn}}{E_k^{(0)} - E_n^{(0)}} \right| \ll 1 \tag{15.2.36}$$

We see that for degenerate states this term blows up, and the inequality clearly cannot be satisfied.

## 15.3 Anharmonic oscillator

Let's try to apply time independent perturbation theory to the anharmonic oscillator, with exact hamiltonian:

$$\hat{H}^{(0)} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{15.3.1}$$

We adimensionalize the problem by setting a length scale $d^2 = \frac{\hbar}{m\omega}$. We impose a perturbation

$$\lambda\delta\hat{H} = \lambda\hbar\omega\frac{\hat{x}^4}{d^4} \tag{15.3.2}$$

Now we may use the typical ladder operator formalism:

$$\hat{x} = \frac{d}{\sqrt{2}}(a + a^\dagger) \tag{15.3.3}$$

$$\hat{p} = -i\frac{\hbar}{\sqrt{2}d}(a - a^\dagger) \tag{15.3.4}$$

so that:

$$\lambda\delta\hat{H} = \frac{\lambda\hbar\omega}{4}(a + a^\dagger)^4 \tag{15.3.5}$$

The first order energy correction to the ground state is then:

$$E_0^{(1)} = \left\langle 0 \middle| \delta\hat{H} \middle| 0 \right\rangle = \frac{3\hbar\omega}{4} \tag{15.3.6}$$

The second order correction is:

$$E_0^{(2)} = -\sum_{k\neq 0}\frac{|\delta H_{0k}|^2}{E_k^{(0)} - E_0^{(0)}} \tag{15.3.7}$$

Now it remains to evaluate the matrix elements $\delta H_{0k}$:

$$\delta H_{0k} = \frac{\hbar\omega}{4}\left\langle 0 \middle| (a + a^\dagger)^4 \middle| k \right\rangle \tag{15.3.8}$$

$$= \frac{\hbar\omega}{4}(4\langle 4 \,|\, k\rangle + 6\sqrt{2}\langle 2 \,|\, k\rangle + 3\langle 0 \,|\, k\rangle) \tag{15.3.9}$$

Thus:

$$\delta H_{02} = \frac{3\sqrt{2}\hbar\omega}{2} \tag{15.3.10}$$

$$\delta H_{04} = \frac{\sqrt{6}\hbar\omega}{2} \tag{15.3.11}$$

so that:

$$E_0^{(2)} = -\frac{|\delta H_{02}|^2}{E_2^{(0)} - \frac{\hbar\omega}{2}} - \frac{|\delta H_{04}|^2}{E_4^{(0)} - \frac{\hbar\omega}{2}} = -\frac{21}{8}\hbar\omega \tag{15.3.12}$$

Thus:

$$E_0(\lambda) = \frac{\hbar\omega}{2}(1 + \frac{3}{2}\lambda - \frac{21}{4}\lambda^2 + o(\lambda^4)) \tag{15.3.13}$$

## 15.4 Degenerate perturbation theory lifted at first order

### Setting up the equations

Consider an exact hamiltonian $\hat{H}^{(0)}$ with degenerate spectrum ordered as:

$$E_0^{(0)} < E_1^{(0)} < ... < E_n^{(0)} = E_{n+1}^{(0)} = ... = E_{n+N-1}^{(0)} < E_{n+N}^{(0)} < ... \tag{15.4.1}$$

composed of $N$ degenerate orthonormal states from $n$ to $n+N-1$ with the same energy. Let us label their eigenstates as:

$$|n^{(0)};1\rangle\,,\ |n^{(0)};2\rangle\,,\ ...\ |n^{(0)};N\rangle \tag{15.4.2}$$

$$\hat{H}^{(0)}\,|n^{(0)};i\rangle = E_n^{(0)}\,|n^{(0)};i\rangle\,,\ \left\langle n^{(0)};k \,\middle|\, n^{(0)};i\right\rangle = \delta_{ik},\ 1 \leq k,i \leq N \tag{15.4.3}$$

which span the subspace $\mathcal{H}_N$. Then it will be orthogonal and complementary to $\hat{\mathcal{H}}$ which contains all the other eigenstates with different energies (these can be degenerate or non-degenerate, doesn't matter a hoot). We see that the total Hilbert space will be the direct sum $\mathcal{H} = \mathcal{H}_N \oplus \hat{\mathcal{H}}$, so that $\mathcal{H}_N \perp \hat{\mathcal{H}}$ and thus:

$$\left\langle p^{(0)} \,\middle|\, n^{(0)};k\right\rangle = 0,\ \forall\,|p^{(0)}\rangle \in \hat{\mathcal{H}},\ \forall\,|n^{(0)};k\rangle \in \mathcal{H}_N \tag{15.4.4}$$

Then as always the degenerate state $|n^{(0)};k\rangle$ will evolve with $\lambda$ as a power series:

$$|n;k\rangle_\lambda = |n^{(0)};k\rangle + \lambda\,|n^{(1)};k\rangle + \lambda^2\,|n^{(2)};k\rangle + ... + \lambda^k\,|n^{(k)};k\rangle + ... \tag{15.4.5}$$

$$E_{n,k}(\lambda) = E_n^{(0)} + \lambda E_{n,k}^{(1)} + \lambda^2 E_{n,k}^{(2)} + ... + \lambda^k E_{n,k}^{(k)} + ... \tag{15.4.6}$$

Now the equation we wish to solve is:

$$\hat{H}(\lambda)\,|n;k\rangle_\lambda = E_{n,k}(\lambda)\,|n;k\rangle_\lambda \tag{15.4.7}$$

As before we will find that:

$$\lambda^{(0)} : (\hat{H}^{(0)} - E_n^{(0)})\,|n^{(0)};k\rangle = 0 \tag{15.4.8}$$

$$\lambda^{(1)} : (\hat{H}^{(0)} - E_n^{(0)})\,|n^{(1)};k\rangle = (E_{n,k}^{(1)} - \delta\hat{H})\,|n^{(0)};k\rangle \tag{15.4.9}$$

$$\vdots$$

$$\lambda^{(i)} : (\hat{H}^{(0)} - E_n^{(0)})\,|n^{(i)};k\rangle = (E_{n,k}^{(1)} - \delta\hat{H})\,|n^{(i-1)};k\rangle + E_{n,k}^{(2)}\,|n^{(i-2)};k\rangle + ... + E_{n,k}^{(i)}\,|n^{(0)};k\rangle \tag{15.4.10}$$

$$\vdots$$

We can assume that $\left\langle n^{(0)};k \,\middle|\, n^{(p)};k\right\rangle = 0$ for $p = 1,2,3...$ by methods identical to non-

degenerate perturbation theory.

### Finding $E^{(1)}_{n,k}$ by dotting (15.4.9) with $\langle n^{(0)}; l|$

Let's now dot (15.4.9) with some degenerate state $\langle n^{(0)}; l| \in \mathcal{H}_N$ so that:

$$0 = E^{(1)}_{n,k} \left\langle n^{(0)}; l \middle| n^{(0)}; k \right\rangle - \left\langle n^{(0)}; l \middle| \delta\hat{H} \middle| n^{(0)}; k \right\rangle \tag{15.4.11}$$

$$\iff E_{n,k}\delta_{kl} = \left\langle n^{(0)}; l \middle| \delta\hat{H} \middle| n^{(0)}; k \right\rangle \tag{15.4.12}$$

Therefore, we need $\delta\hat{H}$ to be diagonal in the block spanned by the basis of $\mathcal{H}_N$. Hence, when performing degenerate perturbation theory we need to:

Use a basis which makes the perturbation hamiltonian diagonal in the degenerate subspace

The elements of this diagonal block will be the first energy corrections:

$$E^{(1)}_{n,k} = \delta H_{kk} \equiv \left\langle n^{(0)}; k \middle| \delta\hat{H} \middle| n^{(0)}; k \right\rangle \tag{15.4.13}$$

Interestingly, this first order correction depends on $k$, so two different degenerate states in $\mathcal{H}_N$ will (generally) evolve differently with the perturbation. When this does occur:

$$E^{(1)}_{n,k} \neq E^{(1)}_{n,l},\ \forall k \neq l \tag{15.4.14}$$

then we say that the degeneracy is lifted at first order corrections. Visually, we see this as the splitting of the degenerate states in Figure 15.1.

### Finding a good basis

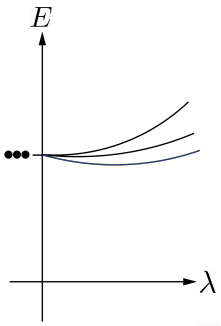
To check whether or not $\delta\hat{H}$ is diagonalized in a basis as required by (15.4.13), we check if these basis vectors are non-degenerate eigenstates of some other hermitian operator $\hat{A}$, and if $[\hat{A}, \delta\hat{H}] = 0$.

Indeed, suppose we have two states $|n^{(0)}; p\rangle$ and $|n^{(0)}; q\rangle$ are eigenvectors of $\hat{A}$ with associated eigenvalues $\lambda_p \neq \lambda q$ respectively. Then:

$$0 = \left\langle n^{(0)}; p \middle| [\hat{A}, \delta\hat{H}] \middle| n^{(0)}; q \right\rangle = \lambda_p \left\langle n^{(0)}; p \middle| \delta\hat{H} \middle| n^{(0)}; q \right\rangle - \lambda_q \left\langle n^{(0)}; p \middle| \delta\hat{H} \middle| n^{(0)}; q \right\rangle \tag{15.4.15}$$

$$\implies \left\langle n^{(0)}; p \middle| \delta\hat{H} \middle| n^{(0)}; q \right\rangle = 0,\ p \neq q \tag{15.4.16}$$

where we don't have to take complex conjugate of $\lambda_p$ since $\hat{A}$ is hermitian, thus proving that $\delta\hat{H}$ is diagonal (in the $\mathcal{H}_N$ subspace) in this basis.




**Finding $|n^{(1)};k\rangle\,|_{\hat{\mathcal{H}}}$ by dotting (15.4.9) with $|p^{(0)}\rangle$**

Now we may dot (15.4.9) with $|p^{(0)}\rangle \in \hat{\mathcal{H}}$ to find that:

$$(E_p^{(0)} - E_n^{(0)}) \left\langle p^{(0)} \middle| n^{(1)};k \right\rangle = \left\langle p^{(0)} \middle| \cancelto{0}{E_{n,k}^{(1)}} - \delta\hat{H} \middle| n^{(0)};k \right\rangle \tag{15.4.17}$$

$$\iff (E_p^{(0)} - E_n^{(0)}) \left\langle p^{(0)} \middle| n^{(1)};k \right\rangle = \left\langle p^{(0)} \middle| \delta\hat{H} \middle| n^{(0)};k \right\rangle \tag{15.4.18}$$

$$\iff \left\langle p^{(0)} \middle| n^{(1)};k \right\rangle = \frac{\left\langle p^{(0)} \middle| \delta\hat{H} \middle| n^{(0)};k \right\rangle}{E_p^{(0)} - E_n^{(0)}} \tag{15.4.19}$$

$$\iff \boxed{|n^{(1)};k\rangle\,|_{\hat{\mathcal{H}}} = -\sum_p \frac{\left\langle p^{(0)} \middle| \delta\hat{H} \middle| n^{(0)};k \right\rangle}{E_p^{(0)} - E_n^{(0)}} |p^{(0)}\rangle} \tag{15.4.20}$$

where $|n^{(1)};k\rangle\,|_{\hat{\mathcal{H}}}$ is the component of $|n^{(1)};k\rangle$ in $\hat{\mathcal{H}}$. Since we're only summing over $|p^{(0)}\rangle \in \hat{\mathcal{H}}$ we have only found the component in $\hat{\mathcal{H}}$.

**Finding $|n^{(1)};k\rangle\,|_{\mathcal{H}_N}$ by dotting (15.4.10) with $|n^{(0)};l\rangle$**

We still haven't found the component in the degenerate space $\mathcal{H}_N$, to do so we dot $|n^{(0)};l\rangle$ with (15.4.10) using $i = 2$:

$$\left\langle n^{(0)};l \middle| \hat{H}^{(0)} - E_n^{(0)} \middle| n^{(2)};k \right\rangle = \left\langle n^{(0)};l \middle| E_{n,k}^{(1)} - \delta\hat{H} \middle| n^{(1)};k \right\rangle + E_{n,k}^{(2)} \left\langle n^{(0)};l \middle| n^{(0)};k \right\rangle \tag{15.4.21}$$

We find that deconstructing $|n^{(1)};k\rangle$ into $|n^{(1)};k\rangle\,|_{\mathcal{H}_N}$ and $|n^{(1)};k\rangle\,|_{\hat{\mathcal{H}}}$:

$$0 = \left\langle n^{(0)};l \middle| E_{n,k}^{(1)} - \delta\hat{H} \middle| n^{(1)};k \right\rangle |_{\mathcal{H}_N} + \left\langle n^{(0)};l \middle| E_{n,k}^{(1)} - \delta\hat{H} \middle| n^{(1)};k \right\rangle |_{\hat{\mathcal{H}}} + E_{n,k}^{(2)}\delta_{lk} \tag{15.4.22}$$

Since $|n^{(0)};l\rangle$ and $|n^{(1)};k\rangle\,|_{\hat{\mathcal{H}}}$ are orthogonal (they belong to two orthogonal subspaces) we find that:

$$0 = \left\langle n^{(0)};l \middle| E_{n,k}^{(1)} - \delta\hat{H} \middle| n^{(1)};k \right\rangle |_{\mathcal{H}_N} - \left\langle n^{(0)};l \middle| \delta\hat{H} \middle| n^{(1)};k \right\rangle |_{\hat{\mathcal{H}}} + E_{n,k}^{(2)}\delta_{lk} \tag{15.4.23}$$

Let us now simplify:

$$\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\mathcal{H}_N}=\langle n^{(0)};l|\,\delta\hat{H}|\overbrace{\left(\sum_q|n^{(0)};q\rangle\langle n^{(0)};q|\right)}^{\text{projector onto }\mathcal{H}_N}|n^{(1)};k\rangle \tag{15.4.24}$$

$$=\sum_q\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(0)};q\right\rangle\left\langle n^{(0)};q\left|\,n^{(1)};k\right.\right\rangle \tag{15.4.25}$$

$$=E^{(1)}_{n,l}\left\langle n^{(0)};l\left|\,n^{(1)};k\right.\right\rangle \tag{15.4.26}$$

Since we're in both cases projecting onto $\mathcal{H}_N$, we may as well write:

$$\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\mathcal{H}_N}=E^{(1)}_{n,l}\left\langle n^{(0)};l\left|\,n^{(1)};k\right.\right\rangle|_{\mathcal{H}_N} \tag{15.4.27}$$

Hence, substituting back into (15.4.23):

$$-\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\hat{\mathcal{H}}}+(E^{(1)}_{n,k}-E^{(1)}_{n,l})\left\langle n^{(0)};l\left|\,n^{(1)};k\right.\right\rangle_{\mathcal{H}_N}+E^{(2)}_{n,k}\delta_{lk}=0 \tag{15.4.28}$$

Since in the second term we are already projecting into the degenerate subspace with the dot product, we may remove the $\mathcal{H}_N$ label. Consequently:

$$\boxed{-\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\hat{\mathcal{H}}}+(E^{(1)}_{n,k}-E^{(1)}_{n,l})\left\langle n^{(0)};l\left|\,n^{(1)};k\right.\right\rangle+E^{(2)}_{n,k}\delta_{lk}=0} \tag{15.4.29}$$

Setting $l=k$ we have that $\left\langle n^{(0)};k\left|\,n^{(1)};k\right.\right\rangle=0$ so:

$$\boxed{E^{(2)}_{n,k}=\left\langle n^{(0)};k\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\hat{\mathcal{H}}}=-\sum_p\frac{|\delta\hat{H}_{nk,p}|^2}{E^{(0)}_p-E^{(0)}_n}} \tag{15.4.30}$$

Setting $l\neq k$ we find that:

$$\left\langle n^{(0)};l\left|\,n^{(1)};k\right.\right\rangle=\frac{1}{E^{(1)}_{n,k}-E^{(1)}_{n,l}}\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\hat{\mathcal{H}}} \tag{15.4.31}$$

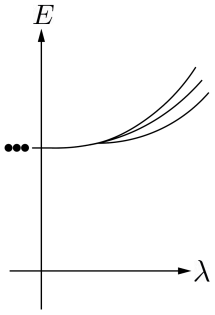
Notice how important it is for the degeneracy to be resolved to first order, or else the denominator $E^{(1)}_{n,k}-E^{(1)}_{n,l}$ would have blown up. Thus we find that:

$$\boxed{|n^{(1)};k\rangle_{\mathcal{H}_N}=\sum_{l\neq k}\frac{\left\langle n^{(0)};l\left|\,\delta\hat{H}\,\right|n^{(1)};k\right\rangle|_{\hat{\mathcal{H}}}}{E^{(1)}_{n,k}-E^{(1)}_{n,l}}|n^{(0)};l\rangle} \tag{15.4.32}$$

## 15.5 Degenerate perturbation theory lifted at second order

### What is a good basis?

Now suppose that degeneracy is not lifted to first order corrections, so that the situation looks like:




We will now have that:

$$\left\langle n^{(0)};l \,\middle|\, \delta\hat{H} \,\middle|\, n^{(0)};k \right\rangle = E_n^{(1)}\delta_{lk} \tag{15.5.1}$$

so the first order energy correction will not depend on $k, l$. Consequently dotting the first order perturbation equation with $|n^{(0)};l\rangle$ will not give us any new information, and this step will therefore be skipped.

Let us form the linear combinations:

$$|\psi^{(0)}\rangle = \sum_{k=1}^{N} |n^{(0)};k\rangle\, a_k^{(0)} \tag{15.5.2}$$

with unknown coefficients $a_k^{(0)}$. Since we are looking for $N$ such states due to the dimension of $\mathcal{H}_N$, we should label the desired basis by $|\psi_I^{(0)}\rangle$ with coefficients $a_{Ik}^{(0)}$ in the degenerate basis $\mathcal{H}_N$:

$$|\psi_I^{(0)}\rangle = \sum_{k=1}^{N} |n^{(0)};k\rangle\, a_{Ik}^{(0)},\ I = 1, 2...N \tag{15.5.3}$$

Note that for $|\psi_I^{(0)}\rangle$ to be a valid orthonormal basis we require:

$$\left\langle \psi_I^{(0)} \,\middle|\, \psi_J^{(0)} \right\rangle = \delta_{IJ} \iff \sum_{k=1}^{N} (a_{Ik}^{(0)})^* a_{Jk}^{(0)} = \delta_{IJ} \tag{15.5.4}$$

Our goal is to determine $a_{Ik}^{(0)}$ in order to construct a good basis from the degenerate eigenbasis.

### Setting up equations

$$|\psi_I\rangle = |\psi_I^{(0)}\rangle + \lambda\,|\psi_I^{(1)}\rangle + \lambda^2\,|\psi_I^{(2)}\rangle + o(\lambda^3) \tag{15.5.5}$$

$$E_{In}(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_{In}^{(2)} + o(\lambda^3) \tag{15.5.6}$$

satisfying:

$$\hat{H}\left|\psi_I\right\rangle = E_{In}(\lambda)\left|\psi_I\right\rangle \tag{15.5.7}$$

As always we get the perturbation equations:

$$\lambda^0 : (\hat{H}^{(0)} - E_n^{(0)})\left|\psi_I^{(0)}\right\rangle = 0 \tag{15.5.8}$$

$$\lambda : (\hat{H}^{(0)} - E_n^{(0)})\left|\psi_I^{(1)}\right\rangle = (E_n^{(1)} - \delta\hat{H})\left|\psi_I^{(0)}\right\rangle \tag{15.5.9}$$

$$\lambda^2 : (\hat{H}^{(0)} - E_n^{(0)})\left|\psi_I^{(2)}\right\rangle = (E_n^{(1)} - \delta\hat{H})\left|\psi_I^{(1)}\right\rangle + E_{In}^{(2)}\left|\psi_I^{(0)}\right\rangle \tag{15.5.10}$$

### Finding $\left|\psi_I^{(1)}\right\rangle$ component in $\hat{\mathcal{H}}$

We find the $\left|\psi_I^{(1)}\right\rangle$ component in $\hat{\mathcal{H}}$ by dotting (15.5.9) with $\langle p^{(0)}| \in \hat{\mathcal{H}}$:

$$\left\langle p^{(0)}\right|\hat{H}^{(0)} - E_n^{(0)}\left|\psi_I^{(1)}\right\rangle = \left\langle p^{(0)}\right|E_n^{(1)} - \delta\hat{H}\left|\psi_I^{(0)}\right\rangle \tag{15.5.11}$$

$$\implies (E_p^{(0)} - E_n^{(0)})\left\langle p^{(0)}\right|\left.\psi_I^{(1)}\right\rangle = -\sum_{k=1}^{N}\delta H_{p,nk}a_{Ik}^{(0)} \tag{15.5.12}$$

thus implying that:

$$\left|\psi_I^{(1)}\right\rangle|_{\hat{\mathcal{H}}} = -\sum_p\sum_{k=1}^{N}\frac{\delta H_{p,nk}a_{Ik}^{(0)}}{E_p^{(0)} - E_n^{(0)}}\left|p^{(0)}\right\rangle \tag{15.5.13}$$

### Finding $a_{Ik}^{(0)}$

We dot (15.5.10) with $\langle n^{(0)}; l|$:

$$\left\langle n^{(0)};l\right|E_n^{(1)} - \delta\hat{H}\left|\psi_I^{(1)}\right\rangle|_{\hat{\mathcal{H}}} + \left\langle n^{(0)};l\right|E_n^{(1)} - \delta\hat{H}\left|\psi_I^{(1)}\right\rangle|_{\mathcal{H}_N} + E_{In}^{(2)}a_{Il}^{(0)} = 0 \tag{15.5.14}$$

Now since degeneracy is not broken to first order we will find that:

$$\left\langle n^{(0)};l\right|\delta\hat{H}\left|\psi_I^{(1)}\right\rangle|_{\mathcal{H}_N} = \langle n^{(0)};l|\,\delta\hat{H}|\overbrace{\left(\sum_q|n^{(0)};q\rangle\langle n^{(0)};q|\right)}^{\text{projector onto }\mathcal{H}_N}\left|\psi_I^{(1)}\right\rangle \tag{15.5.15}$$

$$= \sum_q\left\langle n^{(0)};l\right|\delta\hat{H}\left|n^{(0)};q\right\rangle\left\langle n^{(0)};q\right|\left.\psi_I^{(1)}\right\rangle \tag{15.5.16}$$

$$= E_{n,l}^{(1)}\left\langle n^{(0)};l\right|\left.\psi_I^{(1)}\right\rangle \tag{15.5.17}$$

so that $\left\langle n^{(0)};l\right|E_n^{(1)} - \delta\hat{H}\left|\psi_I^{(1)}\right\rangle|_{\mathcal{H}_N} = 0$. Similarly, we also have that $\left\langle n^{(0)};l\right|E_n^{(1)}\left|\psi_I^{(1)}\right\rangle|_{\hat{\mathcal{H}}} = 0$ meaning that:

$$-\left\langle n^{(0)};l\right|\delta\hat{H}\left|\psi_I^{(1)}\right\rangle|_{\hat{\mathcal{H}}} + E_{In}^{(2)}a_{Il}^{(0)} = 0 \tag{15.5.18}$$

Substituting (15.5.13) into the above:

$$E^{(2)}_{In}a^{(0)}_{Il} = \left\langle n^{(0)};l \left| \delta\hat{H} \right| \psi^{(1)}_I \right\rangle |_{\hat{\mathcal{H}}} \tag{15.5.19}$$

$$= -\sum_p \sum_{k=1}^{N} \frac{\delta H_{p,nk} a^{(0)}_{Ik}}{E^{(0)}_p - E^{(0)}_n} \left\langle n^{(0)};l \left| \delta\hat{H} \right| p^{(0)} \right\rangle \tag{15.5.20}$$

$$= -\sum_p \sum_{k=1}^{N} \frac{\delta H_{p,nk}\delta H_{nl,p}}{E^{(0)}_p - E^{(0)}_n} a^{(0)}_{Ik} \tag{15.5.21}$$

This may be rearranged as:

$$\sum_{k=1}^{N} \Big( -\sum_p \frac{\delta H_{p,nk}\delta H_{nl,p}}{E^{(0)}_p - E^{(0)}_n} - E^{(2)}_{In}\delta_{kl} \Big) a^{(0)}_{Ik} = 0 \tag{15.5.22}$$

Let us introduce a matrix M with components:

$$M^{(2)}_{l,k} = -\sum_p \frac{\delta H_{p,nk}\delta H_{nl,p}}{E^{(0)}_p - E^{(0)}_n} \tag{15.5.23}$$

then one finds that:

$$\sum_k (M^{(2)}_{l,k} - E^{(2)}_{In}\delta_{lk}) a^{(0)}_{Ik} = 0 \tag{15.5.24}$$

or alternatively:

$$(\mathsf{M}^{(2)} - E^{(2)}_{In}\mathbb{1})\mathbf{a}_I = 0 \tag{15.5.25}$$

which is an eigenvalue equation for M! Hence the eigenvalues of M are the second order energy corrections, and the eigenvectors are $\mathbf{a}_I = (a^{(0)}_{I1}\ a^{(0)}_{I2}\ \ldots\ a^{(0)}_{IN})$, the coefficients to construct the good basis.

To find the component of $|\psi^{(1)}_I\rangle$ along $\mathcal{H}_n$ we must resort to the $\lambda^3$ order equation, but the calculations are completely analogous to before.

# The WKB approximation

16

We consider a particle with energy $E$ moving through a constant potential $V < E$. Recall then that the de Broglie wavelength for a plane wave $\psi(\mathbf{x}) = Ae^{\pm ip\mathbf{x}/\hbar}$ of momentum $p = \sqrt{2m(E-V)}$ is:

$$\lambda = \frac{h}{p} = \frac{2\pi\hbar}{p} \tag{16.0.1}$$

In the WKB approximation, we consider cases where $\lambda \ll L$, where $L$ is the length scale of the problem, that is, the order of magnitude of the lengths we are interested in.

For example, consider a particle of energy $E$ in a potential $V(\mathbf{x})$. In the classical framework, the position dependent momentum satisfies:

$$p^2(\mathbf{x}) = 2m(E - V(\mathbf{x})) \tag{16.0.2}$$

so we will thus define $p^2(\mathbf{x})$ as such in quantum mechanics as well.

Therefore, we may analogously define the position dependent or local de Broglie wavelength as:

$$\lambda(\mathbf{x}) = \frac{2\pi\hbar}{p(\mathbf{x})} \tag{16.0.3}$$

Then the TISE reads:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x}) = (E - V(\mathbf{x}))\psi(\mathbf{x}) \implies \hat{p}^2\psi(\mathbf{x}) = p^2(\mathbf{x})\psi(\mathbf{x}) \tag{16.0.4}$$

In the classically allowed region where $E \geq V$, we define the typical wavenumber $k(\mathbf{x})$, but now local as:

$$p^2(\mathbf{x}) = 2m(E - V(\mathbf{x})) = \hbar^2 k^2 \tag{16.0.5}$$

while in the classically forbidden region where $E < V$ we define the imaginary wavenumber $\kappa(x)$ as:

$$-p^2(\mathbf{x}) = 2m(V(\mathbf{x}) - E) = \hbar^2\kappa^2 \tag{16.0.6}$$

Looking at (16.0.4), it seems reasonable now to introduce an exponential wavefunction ansatz:

$$\psi(\mathbf{x}, t) = \sqrt{\rho(\mathbf{x}, t)}e^{iS(\mathbf{x},t)/\hbar} \tag{16.0.7}$$

where $\rho(x,t)$ is the probability density. The associated probability current is:

$$\mathbf{J} = \frac{\hbar}{m}\mathrm{Im}\left[\sqrt{\rho}e^{-iS/\hbar}\left(\frac{1}{2}\frac{\nabla\rho}{\sqrt{\rho}} + \frac{i}{\hbar}(\nabla S)\sqrt{\rho}\right)e^{iS/\hbar}\right] \tag{16.0.8}$$

$$= \frac{\hbar}{m}\mathrm{Im}\left(\frac{1}{2}\nabla\rho + \frac{i}{\hbar}(\nabla S)\rho\right) \tag{16.0.9}$$

$$\implies \mathbf{J} = \rho\frac{\nabla S}{m} \tag{16.0.10}$$

Compared to the fluid dynamics definition $\mathbf{J} = \rho\mathbf{v}$, we may identify the velocity as:

$$\mathbf{v} = \frac{\nabla S}{m} \implies \mathbf{p} = \nabla S \tag{16.0.11}$$

With this in mind, let's substitute the ansatz $\psi(x,t) = e^{iS(x,t)/\hbar}$ where $S(x,t) \in \mathbb{C}$ [1] into (16.0.4):

$$-\hbar^2\left(\frac{i}{\hbar}S'' - \frac{1}{\hbar^2}(S')^2\right)e^{iS/\hbar} = p^2(x)e^{iS/\hbar} \tag{16.0.12}$$

$$\iff (S')^2 - i\hbar S'' - p^2(x) = 0 \tag{16.0.13}$$

If $V(x)$ is slowly varying, so that $V(x) = V_0$ then $p(x) = p_0\sqrt{2m(E-V_0)}$, which we have already solved:

$$\psi(x) = e^{ip_0x/\hbar} \tag{16.0.14}$$

Comparing with our ansatz we identify $S(x) = p_0x$, implying that $S''(x) = 0$. Therefore, if we consider a slowly varying potential $V(x)$ which varies from a constant $V_0$ only infinitesimally, then we can't expect $S''(x)$ to increase considerably, so we may take it to be negligible.

To do so, we may take $\hbar$ to be a small parameter about which we may expand $S(x)$ as a power series:

$$S(x) = S_0(x) + \hbar S_1(x) + o(\hbar^2) \tag{16.0.15}$$

Substituting this into:

$$(S')^2 - i\hbar S'' - p^2(x) = 0 \tag{16.0.16}$$

and keeping terms of first order only we find that:

$$(S_0'(x) + \hbar S_1'(x) + o(\hbar^2))^2 - i\hbar(S_0''(x) + o(\hbar)) - p^2(x) = 0 \tag{16.0.17}$$

$$((S_0'(x))^2 - p^2(x)) + \hbar(2S_0'S_1' - iS_0'') + o(\hbar^2) = 0 \tag{16.0.18}$$

Since we're treating $\hbar$ as the small parameter we may equate the coefficients in parenthesis:

$$\begin{cases}(S_0'(x))^2 = p^2(x)\\ S_1'(x) = \frac{i}{2}\frac{S_0''(x)}{S_0'(x)}\end{cases} \tag{16.0.19}$$

[1] if we kept $S(x,t) \in \mathbb{R}$, then this would be an unphysical solution, since it is not normalizable

The first is easily solved:

$$S_0(x) = \pm \int_{x_0}^{x} p(x')dx' \tag{16.0.20}$$

The second instead is more difficult:

$$S_1' = \frac{i}{2}\frac{\pm p'(x)}{\pm p(x)} = \frac{i}{2}\frac{d}{dx}(\ln p(x))' \tag{16.0.21}$$

$$\implies S_1 = \frac{i}{2}\ln p(x) + c_0 \tag{16.0.22}$$

Therefore:

$$S = \pm \int_{x_0}^{x} p(x')dx' + \frac{i\hbar}{2}\ln p(x) + c_0 \tag{16.0.23}$$

and hence:

$$\psi(x) = e^{iS_0(x)/\hbar}\exp\left[\frac{i}{\hbar}\left(\frac{i\hbar}{2}\ln p(x) + c_0\right)\right] = Ae^{iS_0(x)/\hbar}\exp\left(-\ln\sqrt{p(x)}\right) \tag{16.0.24}$$

or more succintly:

$$\psi(x) = \frac{A}{\sqrt{p(x)}}\exp\left[\pm\frac{i}{\hbar}\int_{x_0}^{x} p(x')dx'\right] \tag{16.0.25}$$

In the notation of (16.0.7), we have that:

$$\rho(x) = \frac{C}{p(x)},\ S(x) = \int_{x_0}^{x} p(x')dx' \tag{16.0.26}$$

If $E > V$ then we defined $p(x) = \hbar k(x)$ so:

$$\psi(x) = \frac{A}{\sqrt{k(x)}}\exp\left[\pm i\int_{x_0}^{x} k(x')dx'\right] \tag{16.0.27}$$

while if $E < V$ then we defined $p(x) = i\hbar\kappa(x)$ so:

$$\frac{A}{\sqrt{\kappa(x)}}\exp\left[\pm\int_{x_0}^{x} \kappa(x')dx'\right] \tag{16.0.28}$$

Note that as the momentum, and hence wavenumber, becomes smaller, the probability density increases. Classically this makes sense, we are more likely to find the particle in regions where it spends more time, that is, where it has lower velocity. Indeed, the probability density is:

$$\rho = |\psi|^2 = \frac{|A|^2}{k(x)} = \frac{\hbar|A|^2}{p(x)} = \frac{\hbar|A|^2}{m}\frac{1}{v(x)} \tag{16.0.29}$$

Instead, the probability density is:

$$J = \frac{\rho}{m}\frac{\partial S}{\partial x} = \frac{\hbar|A|^2}{mk(x)}\hbar k(x) = \frac{\hbar^2|A|^2}{m} \tag{16.0.30}$$

which is constant. This was expected, as the continuity equation for a time-independent probability density gives a homogeneous probability current.

## 16.1 Validity of approximation

The WKB approximation is only valid when:

$$\hbar|S_0''| \ll |S_0'|^2 \implies \hbar\left|\frac{dp}{dx}\right| \ll p^2 \implies \left|\frac{d}{dx}\left(\frac{\hbar}{p(x)}\right)\right| \ll 1 \tag{16.1.1}$$

We recognize the local de Broglie wavelength $\lambda(x) = \frac{2\pi\hbar}{p(x)}$ and write:

$$\boxed{\left|\frac{d\lambda(x)}{dx}\right| \ll 2\pi} \tag{16.1.2}$$

so the wavelength must vary slowly for the WKB approximation to work. We can connect this with the variation of the potential:

$$p^2 = 2m(E - V(x)) \implies \left|\frac{dV}{dx}\right| = \frac{1}{m}\left|p\frac{dp}{dx}\right| \tag{16.1.3}$$

multiplying by $\lambda = \frac{\hbar}{p}$:

$$\lambda\left|\frac{dV}{dx}\right| = \frac{\hbar}{m}\left|\frac{dp}{dx}\right| \ll \frac{\hbar}{m}\frac{p^2}{\hbar} = \frac{p^2}{m} \tag{16.1.4}$$

where we used $\left|\frac{dp}{dx}\right| \ll \frac{p^2}{\hbar}$. We therefore conclude that:

$$\boxed{\lambda(x)\left|\frac{dV}{dx}\right| \ll \frac{p^2}{2m}} \tag{16.1.5}$$

Physically, this means that the variation of the potential over a de Broglie wavelength must be smaller than the kinetic energy of the particle.

For example, suppose we have some potential $V(x)$ acting on a particle with energy $E$. Suppose also that for $x$ close to $a$:

$$V(x) - E = b(x - a),\ g > 0 \tag{16.1.6}$$

so that the potential is approximately linear. Then, we see that:

$$p(x) = \sqrt{2m(E - V(x))} = \sqrt{2mb(x - a)} \implies \lambda = \frac{\hbar}{\sqrt{2mb(x - a)}} \tag{16.1.7}$$

Taking the derivative, we find that:

$$\frac{d\lambda}{dx} = \frac{\hbar}{\sqrt{2mb}}\frac{1}{2}\frac{1}{(a - x)^{3/2}} \tag{16.1.8}$$

which is clearly not small for $x \to a$. We therefore cannot use WKB near turning points. In these regions, we will have to solve the problem exactly, and try to connect them with the wkb approximate solution.

## 16.2 Connection formulae

We consider an infinite potential at $x = 0$ and a slowly varying unbound potential $V(x)$. A particle of energy $E$ is in this potential, such that $V(a) = E$ for some $a > 0$. Suppose using WKB we have found that for $x \ll a$

$$\psi(x) = \frac{2A}{\sqrt{k(x)}} \cos\left( \int_x^a k(x)dx' - \frac{\pi}{4} \right) - \frac{B}{\sqrt{k(x')}} \sin\left( \int_x^a k(x')dx' - \frac{\pi}{4} \right) \tag{16.2.1}$$

while for $x \gg a$:

$$\psi(x) = \frac{A}{\sqrt{k(x)}} \exp\left( - \int_a^x \kappa(x')dx' - \frac{\pi}{4} \right) - \frac{B}{\sqrt{k(x)}} \exp\left( \int_a^x \kappa(x')dx' - \frac{\pi}{4} \right) \tag{16.2.2}$$

Since the wave function must be normalizable, we must set $B = 0$, so $A \neq 0$:

$$\psi(x \ll a) = \frac{2A}{\sqrt{k(x)}} \cos\left( \int_x^a k(x')dx' - \frac{\pi}{4} \right) \tag{16.2.3}$$

$$\psi(x \gg a) = \frac{A}{\sqrt{k(x)}} \exp\left( - \int_a^x \kappa(x')dx' - \frac{\pi}{4} \right) \tag{16.2.4}$$

We can rewrite the first as:

$$\psi(x \ll a) = \frac{2A}{\sqrt{k(x)}} \cos\left( \int_x^a k(x')dx' - \frac{\pi}{4} \right) \tag{16.2.5}$$

$$= \frac{2A}{\sqrt{k(x)}} \cos\left( \int_0^a k(x')dx' - \underbrace{\int_0^x k(x')dx' - \frac{\pi}{4}}_{\delta} \right) \tag{16.2.6}$$

$$= \frac{2A}{\sqrt{k(x)}} \left[ \cos\left( \int_0^x k(x')dx' \right) \cos\delta + \sin\left( \int_0^x k(x')dx' \right) \sin\delta \right] \tag{16.2.7}$$

Now we impose that the wave-function vanish at the origin. This gives:

$$\cos\delta = 0 \implies \delta = \frac{(2n+1)\pi}{2},\ \forall n \in \mathbb{Z} \tag{16.2.8}$$

$$\implies \int_0^a k(x)dx = \left( n + \frac{3}{4}\pi \right),\ \forall n \in \mathbb{Z} \tag{16.2.9}$$

which is the Bohr-Sommerfield quantization. This integral can be eavluated, and gives which approximate energy levels are valid, only those for which the integrand is a specific multiple of $\pi$.

In this case:

$$\psi(x \ll a) = \frac{c}{\sqrt{k(x)}} \sin\left( \int_0^x k(x')dx' \right) \tag{16.2.10}$$

where we set $c = 2A \sin \delta$.

## 16.3 Deriving the connection formulae

We consider the Airy equation:

$$\frac{d^2\psi}{du^2} = u\psi \tag{16.3.1}$$

Its solution is:

$$\psi(u) = c \int_\Gamma e^{ik^3/3} e^{iku} \frac{dk}{2\pi} \tag{16.3.2}$$

where we integrate over some contour in the complex plane.

Since $k$ is a complex number, it will give an exponential decaying integrand if $\text{Im}(k^3) > 0$. Then letting $k = |k|e^{i\theta_k}$ we get $k^3 = |k|^3 e^{3i\theta_k}$. For its imaginary part to be positive we need $0 < 3\theta_k < \pi$. Note that if $\text{Im}(k^3) > 0$ for some $\theta_k$ then the same applies for $\theta_k + \frac{2\pi}{3}$. Consequently, in the Fourier space there will be three regions where $\text{Im}(k^3) > 0$:

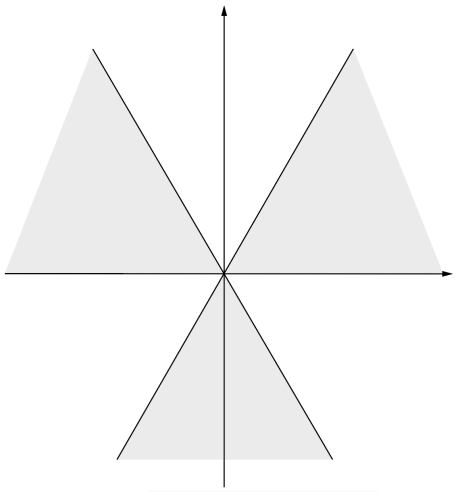


**Figure 16.1**. Regions in Fourier space with $\text{Im}(k^3) > 0$

Now from complex analysis we know that the contour must satisfy:

$$e^{ik^3/3} e^{iku} = 0 \text{ vanish at the ends of } \Gamma \tag{16.3.3}$$

If we use $\Gamma = \mathcal{C}_1$ and $c = 1$ we recover the Airy functions:

$$\psi(u) = \text{Ai}(u) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{k^3}{3} + ku\right) dk \tag{16.3.4}$$

Using a different contour $\Gamma = \mathcal{C}_2$, we find another solution:

$$\mathrm{Bi}(u) = -i \int_{\mathcal{C}_1} e^{ik^3/3} e^{iku} \frac{dk}{2\pi} + 2i \int_{\mathcal{C}_2} e^{ik^3/3} e^{iku} \frac{dk}{2\pi} \tag{16.3.5}$$

$$= \frac{1}{\pi} \int_0^\infty \left[ e^{-k^3/3+ku} + \sin\left(\frac{k^3}{3} + ku\right) \right] dk \tag{16.3.6}$$

# Time-dependent Perturbation theory

17

## 17.1 The Interaction picture

We consider time-dependent Hamiltonians of the form:

$$\hat{H}(t) = \hat{H}^{(0)} + \delta\hat{H}(t) \tag{17.1.1}$$

where the base hamiltonian $\hat{H}^{(0)}$ is time-independent while the perturbation $\delta\hat{H}(t)$ is time-dependent. Due to the time-dependence we can no longer consider states of definite energy, since the TISE equation which leads to the energy eigenequation was only derivable for time-independent hamiltonians. Because of this complication we need to further clarify what quantity we are looking for.

Suppose that we turn on the perturbation at some $t_i$, and keep it on until $t = t_f$ when it is turned off. Before the perturbation is activated the system is in some state, the perturbation then alters it, so that after $t_f$ the system is in some other state. What is this new state given the initial state?

To deal with the time evolution we introduce a new picture, known as the interaction picture.

Recall that in the Schrödinger picture, we time-evolve the states $|\psi_S(0)\rangle$ using the propagator $\hat{U}_S(t)$, keeping the operators unchanged:

$$|\psi_S(t)\rangle = \hat{U}_S(t)\,|\psi_S(0)\rangle \tag{17.1.2}$$

which when substituted into:

$$i\hbar\frac{d}{dt}\,|\psi_S(t)\rangle = \hat{H}(t)\,|\psi_S(t)\rangle \implies i\hbar\frac{d\hat{U}_S}{dt} = \hat{H}_S\hat{U}_S \tag{17.1.3}$$

In the Heisenberg picture, we time evolve the operators using similarity transformations in $\hat{U}(t)$, keeping the states unchanged. So, the states in this picture are:

$$|\psi_H(t)\rangle = \hat{U}_S^\dagger(t)\,|\psi_S(t)\rangle = |\psi_S(0)\rangle \tag{17.1.4}$$

We are rotating (not physically) the states in the Hilbert space to wash out the time-dependence of states in the Schrödinger picture. This unfortunately comes at the expense

of the operators gaining a time dependence through the similarity transformation:

$$\hat{A}_H = \hat{U}_S^\dagger(t)\hat{A}_S\hat{U}_S(t) \tag{17.1.5}$$

where $\hat{A}_S$ and $\hat{A}_H$ are arbitrary operators in the Schrödinger and Heisenberg picture respectively. Expectation values of observables therefore transform as:

$$\left\langle \psi_S(t) \middle| \hat{A} \middle| \psi_S(t) \right\rangle = \left\langle \psi_S(0) \middle| \hat{U}^\dagger(t)\hat{A}\hat{U}(t) \middle| \psi_S(0) \right\rangle \equiv \left\langle \psi_S(0) \middle| \hat{A}_H(t) \middle| \psi_S(0) \right\rangle \tag{17.1.6}$$

In our case $\hat{H} = \hat{H}^{(0)} + \delta\hat{H}(t)$. If $\delta\hat{H} = 0$ then the propagator would read:

$$\hat{U}_S(t) = e^{-i\hat{H}^{(0)}t/\hbar} \tag{17.1.7}$$

so given a state $\psi_S(t)$ in the Schrödinger picture we can view it in the rotating frame as:

$$|\psi_I(t)\rangle = e^{i\hat{H}^{(0)}t/\hbar}\,|\psi_S(t)\rangle \tag{17.1.8}$$

This removes the time evolution generated by $\hat{H}^{(0)}$, so its Schrödinger equation should hopefully be simpler. Also, note that if $\delta\hat{H} = 0$ then:

$$|\psi_I(t)\rangle = e^{i\hat{H}^{(0)}t/\hbar}\,|\psi(t)\rangle = |\psi_I(t)\rangle = e^{i\hat{H}^{(0)}t/\hbar}e^{-i\hat{H}^{(0)}t/\hbar}\,|\psi(0)\rangle = |\psi(0)\rangle \tag{17.1.9}$$

so the state would indeed be still, as expected by the Heisenberg picture. However, when the perturbation is turned on the state $|\psi_I(t)\rangle$ will evolve through time. This new state $|\psi_I(t)\rangle$ gives the **interaction picture**, which is a mix of the Schrödinger picture and the Heisenberg picture. When the hamiltonian is time independent ($\hat{H}^{(0)}$) it reduces to the Heisenberg picture, while for time-dependent hamiltonians ($\delta\hat{H}(t)$) it reduces to the Schrödinger picture as we shall now see.

To find $|\psi_I(t)\rangle$ we look at its Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\,|\psi_I(t)\rangle = -\hat{H}^{(0)}\,|\psi_I(t)\rangle + e^{i\hat{H^{(0)}}t/\hbar}(\hat{H}^{(0)} + \delta\hat{H})\,|\psi_S(t)\rangle \tag{17.1.10}$$

$$= -\hat{H}^{(0)}\,|\psi_I(t)\rangle + e^{i\hat{H^{(0)}}t/\hbar}(\hat{H}^{(0)} + \delta\hat{H})e^{i\hat{H^{(0)}}t/\hbar}\,|\psi_I(t)\rangle \tag{17.1.11}$$

Note that $\hat{H}^{(0)}$ commutes with $e^{i\hat{H^{(0)}}t/\hbar}$, so expanding the brackets the first two terms will cancel out giving:

$$i\hbar\frac{\partial}{\partial t}\,|\psi_I(t)\rangle = e^{i\hat{H^{(0)}}t/\hbar}\delta\hat{H}e^{-i\hat{H^{(0)}}t/\hbar}\,|\psi_I(t)\rangle \tag{17.1.12}$$

We now define $\delta\hat{H}_I = e^{i\hat{H^{(0)}}t/\hbar}\delta\hat{H}e^{-i\hat{H^{(0)}}t/\hbar}$. This is the perturbation hamiltonian in the rotating frame (interaction picture), and it gives the time evolution of the states $|\psi_I(t)\rangle$:

$$i\hbar\frac{\partial}{\partial t}\,|\psi_I(t)\rangle = \delta\hat{H}_I\,|\psi_I(t)\rangle \tag{17.1.13}$$

which follows exactly the evolution of states in the Schrödinger picture.

Consider an orthonormal eigenbasis:

$$\hat{H}^{(0)} |n\rangle = E_n |n\rangle \tag{17.1.14}$$

Then we postulate that since $|\psi\rangle$:

$$|\psi_I(t)\rangle = \sum_n c_n(t) |n\rangle \implies |\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n t/\hbar} |n\rangle \tag{17.1.15}$$

where the coefficients are now time-dependent. Substituting this ansatz into (17.1.13):

$$i\hbar \sum_m \dot{c}_m(t) |m\rangle = \sum_n c_n(t) \delta\hat{H}_I |n\rangle \tag{17.1.16}$$

$$= \sum_{n,m} c_n(t) |m\rangle \left\langle m \middle| \delta\hat{H}_I \middle| n \right\rangle \tag{17.1.17}$$

$$= \sum_{n,m} c_n(t) \delta\hat{H}^I_{mn} |m\rangle \tag{17.1.18}$$

where $\delta\hat{H}^I_{mn} = \left\langle m \middle| \delta\hat{H}_I \middle| n \right\rangle$. Equating components:

$$i\hbar \dot{c}_m(t) = \sum_n c_n(t) \delta\hat{H}^I_{mn} \tag{17.1.19}$$

We can convert the matrix elements to the Schrödinger picture as:

$$\delta\hat{H}^I_{mn} = \left\langle m \middle| \delta\hat{H}_I \middle| n \right\rangle = e^{i(E_m - E_n)t/\hbar} \left\langle m \middle| \delta\hat{H}_S \middle| n \right\rangle = e^{i(E_n - E_m)t/\hbar} \delta H^S_{mn} \tag{17.1.20}$$

Letting $\omega_{mn} = \frac{(E_m - E_n)}{\hbar}$ then we get:

$$\delta\hat{H}^I_{mn} = e^{i\omega_{mn}t} \delta H^S_{mn} \tag{17.1.21}$$

and hence:

$$i\hbar \dot{c}_m(t) = \sum_n c_n(t) e^{i\omega_{mn}t} \delta\hat{H}^S_{mn}(t) \tag{17.1.22}$$

## 17.2 Two-level system

Let's consider a two-level system with eigenstates $|a\rangle, |b\rangle$ and energies $E_a, E_b$ so that:

$$\omega_{ab} = \frac{E_a - E_b}{\hbar}, \ \hat{H} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \tag{17.2.1}$$

Suppose we also introduce a perturbation at $t = 0$:

$$\delta\hat{H}(t) = \begin{pmatrix} 0 & \alpha \\ \alpha^* & 0 \end{pmatrix} \delta(t) \equiv U\delta(t) \tag{17.2.2}$$

This has off diagonal components so it will shuffle the $|a\rangle$ and $|b\rangle$ states together. We ask, suppose the system starts out in the state $|a\rangle$ at $t = -\infty$, what is the probability that it is

in the state $|b\rangle$ at $t = \infty$?

Note that this question is the same as asking, suppose the state starts out in $|a\rangle$ at $t = 0^-$, what is the probability that it is then measured in $|b\rangle$ at $t = 0^+$? Indeed for $t \in (-\infty, 0)$:

$$|\psi(t)\rangle = e^{-iE_a t/\hbar} |a\rangle \tag{17.2.3}$$

so the state is only affected by a phase factor before the perturbation is on. Immediately after $t = 0$, the state will be in some state:

$$|\psi(0^+)\rangle = \gamma_a |a\rangle + \gamma_b |b\rangle \implies P_b(0^+) = |\gamma_b|^2 \tag{17.2.4}$$

so for $t \in (0, \infty)$ we find that:

$$|\psi(t)\rangle = \gamma_A e^{iE_a t/\hbar} |a\rangle + \gamma_B e^{iE_b t/\hbar} |b\rangle \implies P_b(t) = |\gamma_b|^2 \tag{17.2.5}$$

Our problem is thus to determine $\gamma_b$.

In the interaction picture the state of the system will be given by:

$$|\psi_I(t)\rangle = c_a(t) |a\rangle + c_B(t) |b\rangle \tag{17.2.6}$$

with $c_a(0^-) = 1$ and $c_b(0^-) = 0$. This state will evolve following (17.1.22), giving the following system of differential equations:

$$i\hbar \dot{c}_a(t) = e^{i\omega_{ab} t} \delta H_{ab}(t) c_b(t) \tag{17.2.7}$$

$$i\hbar \dot{c}_b(t) = e^{-i\omega_{ab} t} \delta H_{ba}(t) c_a(t) \tag{17.2.8}$$

Now the exponential function $e^{\pm i\omega_{ab} t}$, unlike $c_a(t)$, must be continuous at $t = 0$. Therefore since $\delta H_{ab} = \alpha\delta(t)$ and $\delta H_{ba} = \alpha^*\delta(t)$, we may substitute the exponentials with their values at $t = 0$:

$$i\hbar \dot{c}_a(t) = \alpha\delta(t) c_b(t) \tag{17.2.9}$$

$$i\hbar \dot{c}_b(t) = \alpha^*\delta(t) c_a(t) \tag{17.2.10}$$

Dealing with the delta function in physics is quite hard, no physical signal in the world is truly a delta function. We may approximate the perturbation instead using the regulation:

$$\delta(t) \approx \begin{cases} \frac{1}{t'}, \ 0 \le t \le t' \\ 0 \text{ otherwise} \end{cases} \tag{17.2.11}$$

So, the new initial conditions are $c_a(0) = 1, c_b(0) = 0$ (we no longer have to use the $0^\pm$ symbols since the perturbation is not a delta function). We are tasked with finding $c_a(t')$ and $c_b(t')$.

We have that for $t \in [0, t']$:

$$i\hbar\dot{c}_a(t) = \frac{\alpha}{t'}c_b(t) \tag{17.2.12}$$

$$i\hbar\dot{c}_b(t) = \frac{\alpha^*}{t'}c_a(t) \tag{17.2.13}$$

implying that:

$$i\hbar\ddot{c}_a(t) = \frac{\frac{\alpha}{t'}}{i\hbar}c_a(t) \implies \ddot{c}_a(t) = -\frac{|\alpha|}{\hbar t'}c_a(t) \tag{17.2.14}$$

The solution to our problem is therefore:

$$c_a(t) = \beta_0 \cos\left(\frac{|\alpha|t}{\hbar t'}\right) + \beta_1 \sin\left(\frac{|\alpha|t}{\hbar t'}\right), c_b(t) \propto \beta_0 \sin\left(\frac{|\alpha|t}{\hbar t'}\right) + \beta_1 \cos\left(\frac{|\alpha|t}{\hbar t'}\right) \tag{17.2.15}$$

Now we must set $c_b(0) = 0 \implies \beta_1 = 0$, as well as $c_a(0) = 1 \implies \beta_0 = 1$, giving:

$$c_a(t) = \cos\left(\frac{|\alpha|t}{\hbar t'}\right),\ c_b(t) = -\frac{i|\alpha|}{\alpha}\sin\left(\frac{|\alpha|t}{\hbar t'}\right) \tag{17.2.16}$$

For $t \geq t'$ we find that:

$$c_a(t \geq t') = \cos\left(\frac{|\alpha|}{\hbar}\right),\ c_b(t \geq t') = -\frac{i|\alpha|}{\alpha}\sin\left(\frac{|\alpha|}{\hbar}\right) \tag{17.2.17}$$

and hence:

$$|\psi(t > t')\rangle = \cos\left(\frac{|\alpha|}{\hbar}\right)e^{-iE_a t/\hbar}|a\rangle - \frac{i|\alpha|}{\alpha}\sin\left(\frac{|\alpha|}{\hbar}\right)e^{-iE_b t/\hbar}|b\rangle \tag{17.2.18}$$

yielding:

$$P_a(t) = \sin^2\left(\frac{|\alpha|}{\hbar}\right),\ P_b(t) = \cos^2\left(\frac{|\alpha|}{\hbar}\right) \tag{17.2.19}$$

## 17.3 Setting up the perturbation equations

We consider the time-dependent Hamiltonian:

$$\hat{H}(t) = \hat{H}^{(0)} + \lambda\delta\hat{H}_S(t) \tag{17.3.1}$$

where $\lambda$ is treated as a small parameter. We expand the state in the interaction picture:

$$|\psi_I(t)\rangle = |\psi_I^{(0)}\rangle + \lambda|\psi_I^{(1)}(t)\rangle + \lambda^2|\psi_I^{(2)}(t)\rangle + ... \tag{17.3.2}$$

which we substitute into the Schrödinger equation to find that:

$$i\hbar\frac{\partial}{\partial t}(|\psi_I^{(0)}\rangle + \lambda|\psi_I^{(1)}(t)\rangle + \lambda^2|\psi_I^{(2)}(t)\rangle + ...) = \lambda\delta\hat{H}_I(|\psi_I^{(0)}\rangle + \lambda|\psi_I^{(1)}(t)\rangle + \lambda^2|\psi_I^{(2)}(t)\rangle + ...) \tag{17.3.3}$$

Equating coefficients of $\lambda^k$ for $k = 0, 1, ...$ we find that:

$$i\hbar\frac{\partial}{\partial t}\left|\psi_I^{(0)}(t)\right\rangle = 0 \tag{17.3.4}$$

$$i\hbar\frac{\partial}{\partial t}\left|\psi_I^{(k)}(t)\right\rangle = \delta\hat{H}_I\left|\psi_I^{(k-1)}\right\rangle,\ k = 1, 2, ... \tag{17.3.5}$$

The first is to be expected, since $|\psi_I^{(0)}(t)\rangle$ is an eigenvalue of $\hat{H}^{(0)}$, a time-independent hamiltonian. Plugging in $t = 0$, and assuming that the system starts out in the state $|\psi^I(0)\rangle = |\psi(0)\rangle$ then:

$$|\psi_I(0)\rangle = |\psi(0)\rangle = |\psi_I^{(0)}(0)\rangle + \lambda\,|\psi_I^{(1)}(0)\rangle + \lambda^2\,|\psi_I^{(2)}(0)\rangle + ... \tag{17.3.6}$$

implying that

$$|\psi_I^{(0)}(0)\rangle = |\psi(0)\rangle \tag{17.3.7}$$

$$|\psi_I^{(k)}(0)\rangle = 0,\ k = 1, 2, ... \tag{17.3.8}$$

Note that the zeroth order equation with these initial conditions gives:

$$|\psi_I^{(0)}(t)\rangle = cnst. \implies \boxed{|\psi_I^{(0)}(t)\rangle = |\psi(0)\rangle} \tag{17.3.9}$$

Therefore, the first order equation reads.

$$i\hbar\frac{\partial}{\partial t}\left|\psi_I^{(1)}(t)\right\rangle = \delta\hat{H}_I(t)\,|\psi(0)\rangle \implies \boxed{|\psi_I^{(1)}(t)\rangle = \int_0^t \frac{\delta\hat{H}_I(t')}{i\hbar}\,|\psi(0)\rangle\,dt'} \tag{17.3.10}$$

where we are forced to use 0 as the lower limit of integration to satisfy $|\psi_I^{(1)}(0)\rangle = 0$.

Similarly, to second order we find that:

$$|\psi_I^{(2)}(t)\rangle = \int_0^t \frac{\delta\hat{H}_I(t')}{i\hbar}\,|\psi_I^{(1)}(t)\rangle\,dt' = \int_0^t \frac{\delta\hat{H}_I(t')}{i\hbar}\int_0^{t'}\frac{\delta\hat{H}_I(t'')}{i\hbar}\,|\psi(0)\rangle\,dt''dt' \tag{17.3.11}$$

Suppose we expand our state in the eigenbasis $\{|n\rangle\}$ of $\hat{H}^{(0)}$:

$$|\psi_I(t)\rangle = \sum_n c_n(t)\,|n\rangle \tag{17.3.12}$$

and similarly:

$$|\psi_I^{(k)}(t)\rangle = \sum_n c_n^{(k)}(t)\,|n\rangle \tag{17.3.13}$$

Then, substituting (17.3.13) into (17.3.9) we find that

$$|\psi_I(0)\rangle = \sum_n c_n(0)\,|n\rangle = |\psi_I^{(0)}(t)\rangle = \sum_n c_n^{(0)}(t)\,|n\rangle \implies \boxed{c_n^{(0)} = c_n(0)} \tag{17.3.14}$$

Similarly, substituting (17.3.13) into (17.3.10) we find that:

$$\sum_n c_n^{(1)}(t)\,|n\rangle = \int_0^t \frac{\delta\hat{H}_I(t')}{i\hbar}\sum_m c_m(0)\,|m\rangle \tag{17.3.15}$$

$$= \sum_{n,m}\int_0^t \frac{c_m(0)}{i\hbar}\,|n\rangle\left\langle n\left|\,\delta\hat{H}_I(t')\,\right|m\right\rangle \tag{17.3.16}$$

$$\implies c_n^{(1)}(t) = \sum_m \int_0^t \frac{c_m(0)}{i\hbar}\left\langle n\left|\,\delta\hat{H}_I(t')\,\right|m\right\rangle \tag{17.3.17}$$

Recalling that $\delta\hat{H}^I_{nm}(t') = e^{i\omega_{nm}t}\delta\hat{H}^S_{nm}(t')$ we finally find that:

$$c_n^{(1)}(t) = \sum_m \int_0^t e^{i\omega_{nm}t'}\frac{c_m(0)}{i\hbar}\delta\hat{H}^S_{nm}(t') \tag{17.3.18}$$

## 17.4 Constant perturbation

Let us consider a system starting out in an initial state $|i\rangle$, and acted upon by a constant perturbation $\delta\hat{H} \equiv \hat{V}$ which switches on at $t = 0$. We want to know the probability that the system ends up in some final eigenstate $|f\rangle$ upon measurement.

Therefore, we have that $c_n(0) = \delta_{ni}$, since the state starts out in the eigenstate $|i\rangle$. Consequently, letting $V_{fi} = \left\langle f\left|\,\hat{V}\,\right|i\right\rangle$

$$c_f^{(1)}(t_0) = \int_0^t e^{i\omega_{fi}t'}\frac{V_{fi}}{i\hbar}dt' \tag{17.4.1}$$

$$= \frac{1}{i\hbar}\frac{V_{fi}}{i\omega_{fi}}(e^{i\omega_{fi}t_0} - 1) \tag{17.4.2}$$

$$= -\frac{2i}{\hbar\omega_{fi}}e^{i\omega_{fi}t_0/2}\sin\frac{\omega_{fi}t_0}{2} \tag{17.4.3}$$

$$= -\frac{2iV_{fi}}{E_f - E_i}e^{i\omega_{fi}t_0/2}\sin\frac{\omega_{fi}t_0}{2} \tag{17.4.4}$$

The probability that the system transitions from $|i\rangle$ to $|f\rangle$ after time $t_0$ is:

$$P^{(1)}_{i\to f}(t_0) = |c_f^{(1)}(t_0)|^2 = \frac{4|V_{fi}|^2}{(E_f - E_i)^2}\sin^2\frac{\omega_{fi}t_0}{2} \tag{17.4.5}$$

The transition probability is periodic, and vanishes at

$$t_0 = \frac{2m\pi}{|\omega_{fi}|} = \frac{2m\hbar\pi}{|E_f - E_i|} \tag{17.4.6}$$

Furthermore, as $E_f \to E_i$:

$$\lim_{E_f \to E_i} P_{i\to f} = |V_{fi}|^2 \frac{t_0^2}{\hbar^2} \lim_{E_f \to E_i} \frac{\sin^2 \frac{\omega_{fi} t_0}{2}}{\left(\frac{\omega_{fi} t_0}{2}\right)^2} = \left(\frac{|V_{fi}| t_0}{\hbar}\right)^2 \tag{17.4.7}$$

This probability is unbounded, so we see that the first order perturbation is only valid for small enough $t_0$.

Now suppose that we have a continuous of energy levels $E_f$ to which transitions can occur. Our goal is to integrate over this continuum, and find the **transition rate**, the probability of some transition happening per unit time.

## 17.5 Continuum of states

The momentum of a free particle takes a continuum of values, so asking how many particles have momentum between two values is a senseless question. To solve this issue, we suppose we insert our free particle in some box of length $L$, hoping that in the end this parameter will get cancelled out, just like the broadening of the delta function did in the two-state system.

For a box of length $L$, we can use the momentum eigenstate $|\mathbf{k}\rangle$:

$$\psi_{\mathbf{k}}(x) = \frac{1}{\sqrt{L^3}} e^{ik_x x} e^{ik_y y} e^{ik_z z} \tag{17.5.1}$$

Applying periodic boundary conditions we get the proper momentum quantization:

$$k_x L = 2\pi n_x,\ k_y L = 2\pi n_y,\ k_z L = 2\pi n_z \tag{17.5.2}$$

and consequently for some small momentum interval cube $[k_x + dk_x, k_y + dk_y, k_z + dk_z]$:

$$dk_x L = 2\pi dn_x,\ dk_y L = 2\pi dn_y,\ dk_z L = 2\pi dn_z \tag{17.5.3}$$

The number of states in this cube is

$$dn = dn_x dn_y dn_z = \left(\frac{L}{2\pi}\right)^3 d^3\mathbf{k} \tag{17.5.4}$$

Suppose instead of looking at the number of states within some momentum interval, I want the number of states within some energy interval. Defining the density of states $\rho(E)$ as the number of states per unit energy, so that $\rho(E)dE$ gives the number of states with energy within an interval $dE$, then we require:

$$dn = \left(\frac{L}{2\pi}\right)^3 d^3\mathbf{k} = \rho(E) dE \tag{17.5.5}$$

Recall that:

$$E = \frac{\hbar^2 k^2}{2m} \implies dE = \frac{\hbar^2 k}{m} dk \tag{17.5.6}$$

In spherical coordinates, we may write $\mathbf{k} = (k, \theta, \phi)$ so that:

$$d^3\mathbf{k} = k^2 \sin\theta dk d\theta d\phi = \frac{m}{\hbar^2} k \sin\theta d\theta d\phi dE \tag{17.5.7}$$

Substituting this into (17.5.5) one finds that:

$$\left(\frac{L}{2\pi}\right)^3 \frac{m}{\hbar^2} k \sin\theta d\theta d\phi = \rho(E) \tag{17.5.8}$$

or defining the solid angle $d\Omega = \sin\theta d\theta d\phi$ then:

$$\rho(E) = \left(\frac{L}{2\pi}\right)^3 \frac{m}{\hbar^2} k d\Omega \tag{17.5.9}$$

## 17.6 Fermi's golden rule

As mentioned previously, we now integrate over the continuum of states to find the transition probability:

$$\sum_f P_{i\to f}(t_0) \equiv \int P^{(1)}_{i\to f}(t_0)\rho(E_f)dE_f \tag{17.6.1}$$

$$= \int \frac{4|V_{fi}|^2}{(E_f - E_i)^2} \sin^2\left(\frac{\omega_{fi}t_0}{2}\right)\rho(E_f)dE_f \tag{17.6.2}$$

Now we know that the probability of transition will be low for large $|E_f - E_i|$, so the contributions to the integral will come from a narrow band $\Delta E$ of energies near $E_i$. We may assume that both $|V_{fi}|^2$ and $\rho(E_f)$ are slowly varying, and bring them out of the integral:

$$\sum_f P_{i\to f}(t_0) = 4|V_{fi}|^2\rho(E_f) \int_{\Delta E} \sin^2\left(\frac{\omega_{fi}t_0}{2}\right) \frac{1}{\hbar^2\omega_{fi}^2} dE_f \tag{17.6.3}$$

$$= \frac{4|V_{fi}|^2\rho(E_f)}{\hbar} \int_{\Delta E} \sin^2\left(\frac{\omega_{fi}t_0}{2}\right) \frac{1}{\omega_{fi}^2} d\omega_{fi} \tag{17.6.4}$$

$$\tag{17.6.5}$$

Let's plot the integrand over $\omega_{fi}$ https://www.desmos.com/calculator/ndmxj0vog7:

We see that the majority (about 95%) of the integral will come from the first oscillation in the interval:

$$-\frac{2\pi}{t_0} \leq \omega_{fi} \leq \frac{2\pi}{t_0} \implies E_i - \frac{2\pi\hbar}{t_0} \leq E_f \leq E_i + \frac{2\pi\hbar}{t_0} \tag{17.6.6}$$

For this energy range to be narrow, we need $t_0$ to be large enough. This follows immediately from the uncertainty principle $t_0\Delta E \geq \hbar$. However, it should also be small enough so that first order perturbation is still approximately adequate. Therefore, we may integrate over infinity, since contributions outside the given energy bands are very small for large enough

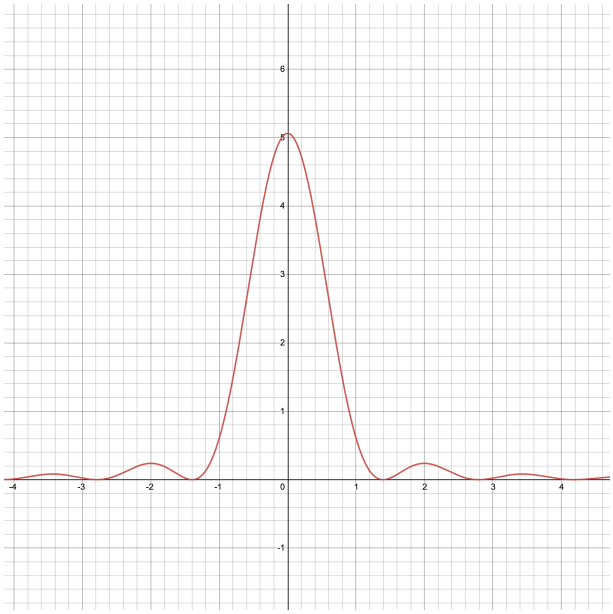


**Figure 17.1**. Plot of $\sin^2\left(\frac{\omega_{fi}t_0}{2}\right)\frac{1}{\omega_{fi}^2}$

$t_0$:

$$\int_{-\infty}^{\infty} \sin^2\left(\frac{\omega_{fi}t_0}{2}\right)\frac{1}{\omega_{fi}^2}d\omega_{fi} = \frac{t_0\pi}{2} \tag{17.6.7}$$

The total transition probability is then

$$\sum_f P_{i\to f}(t_0) = \frac{2\pi}{\hbar}|V_{fi}|^2\rho(E_f)t_0 \tag{17.6.8}$$

We see that the probability of transition is directly proportional to the time elapsed. The constant of proportionality is the transition rate:

$$\Gamma_{i\to[f]} = \frac{2\pi}{\hbar}|V_{fi}|^2\rho(E_f) \tag{17.6.9}$$

where $[f]$ means that we are considering transitions to a continuum of final states $|f\rangle$.

Indeed, one might wonder what the $f$ in this expression is. Because we assumed that the matrix elements and the probability density of states does not vary much, we can take any state in the contributing energy band as the final state.

### Validity of Fermi's golden rule

Fermi's golden rule is only valid for times $t_0$ which are small enough for first order perturbation to suffice, but large enough for us to bring out the density of states and perturbation matrix elements out of the integral in (17.7.13). It is important to ascertain that these two conditions are compatible with each other, and that it is possible to find such time intervals where Fermi's golden rule applies.

Firstly, we saw that the result from perturbation theory only holds when:

$$t_0 \ll \frac{\hbar}{|V_{fi}|} \tag{17.6.10}$$

However, from the energy-time uncertainty principle, we also require that $t_0$ be large enough for the $\frac{\sin x}{x}$ behaviour to be approximately a delta function, giving a narrow energy range $\Delta E$:

$$t_0 \Delta E = t_0 \hbar \omega_{fi} \geq \hbar \implies t_0 \geq \frac{1}{\omega_{fi}} \tag{17.6.11}$$

Combined, we find that Fermi's golden rule applies only in cases where:

$$\frac{1}{\omega_{fi}} \leq t_0 \ll \frac{\hbar}{|V_{fi}|} \tag{17.6.12}$$

## 17.7 Harmonic perturbation

We now consider harmonic perturbations of the form:

$$\delta\hat{H}(t) = \begin{cases} 2H' \cos\omega t, \ 0 < t < t_0 \\ 0, \ \text{otherwise} \end{cases} \tag{17.7.1}$$

where $H'$ is time independent. We ask again what the probability of a transition from some initial state $|i\rangle$ to some $|f\rangle$ is.

Therefore, we have that $c_n(0) = \delta_{ni}$, since the state starts out in the eigenstate $|i\rangle$. Consequently, letting $H'_{fi} = \left\langle f \middle| \hat{H}' \middle| i \right\rangle$

$$c_f^{(1)}(t_0) = \int_0^{t_0} e^{i\omega_{fi}t'} \frac{2H'_{fi}}{i\hbar} \cos\omega t' dt' \tag{17.7.2}$$

$$= \frac{H'_{fi}}{i\hbar} \int_0^{t_0} \left( e^{i(\omega_{fi}+\omega)t'} + e^{i(\omega_{fi}-\omega)t'} \right) dt' \tag{17.7.3}$$

$$= -\frac{H'_{fi}}{\hbar} \left[ \frac{e^{i(\omega_{fi}+\omega)t_0} - 1}{\omega_{fi} + \omega} + \frac{e^{i(\omega_{fi}-\omega)t'} - 1}{\omega_{fi} - \omega} \right] \tag{17.7.4}$$

### Absorption

In the case where $\omega_{fi} - \omega \approx 0$, then :

$$E_f - E_i - \hbar\omega \approx 0 \implies E_f \approx E_i + \hbar\omega \tag{17.7.5}$$

This is known as **stimulated absorption**, the system absorbs some energy $\hbar\omega$ by transitioning to a higher energy state.

In this case, we find that:

$$c_f^{(1)}(t_0) = -\frac{H'_{fi}}{\hbar}\frac{e^{i(\omega_{fi}-\omega)t_0}-1}{\omega_{fi}-\omega} \tag{17.7.6}$$

$$= -\frac{H'_{fi}}{\hbar}e^{i\omega_{fi}t_0/2}\frac{e^{i(\omega_{fi}-\omega)t_0/2}-e^{-i(\omega_{fi}-\omega)t_0/2}}{\omega_{fi}-\omega} \tag{17.7.7}$$

$$= -\frac{2iH'_{fi}}{\hbar(\omega_{fi}-\omega)}e^{i\omega_{fi}t_0/2}\sin\frac{(\omega_{fi}-\omega)t_0}{2} \tag{17.7.8}$$

$$\tag{17.7.9}$$

The probability that the system transitions from $|i\rangle$ to $|f\rangle$ after time $t_0$ is:

$$P_{i\to f}^{(1)}(t_0) = |c_f^{(1)}(t_0)|^2 = \frac{4|H'_{fi}|^2}{\hbar^2(\omega_{fi}-\omega)^2}\sin^2\frac{(\omega_{fi}-\omega)t_0}{2} \tag{17.7.10}$$

The transition probability is periodic, and vanishes at

$$t_0 = \frac{2m\pi}{|\omega_{fi}-\omega|} \tag{17.7.11}$$

Now suppose that we have a continuum of energy levels $E_f$ to which transitions can occur. The total probability of transitions is then given by summing over the continuum of states:

$$\sum_f P_{i\to f}(t_0) \equiv \int P_{i\to f}^{(1)}(t_0)\rho(E_f)dE_f \tag{17.7.12}$$

$$= \int \frac{4|H'_{fi}|^2}{\hbar^2(\omega_{fi}-\omega)^2}\sin^2\left(\frac{(\omega_{fi}-\omega)t_0}{2}\right)\rho(E_f)dE_f \tag{17.7.13}$$

Now we know that the probability of transition will occur for $E_f \approx E_i + \hbar\omega$, so the contributions to the integral will come from a narrow band $\Delta E$ of energies. We may assume that both $|V_{fi}|^2$ and $\rho(E_f)$ are slowly varying over this interval, and bring them out of the integral:

$$\sum_f P_{i\to f}(t_0) = 4|H'_{fi}|^2\rho(E_i+\hbar\omega)\int_{\Delta E}\frac{\sin^2\frac{(\omega_{fi}-\omega)t_0}{2}}{\hbar^2(\omega_{fi}-\omega)^2}dE_f \tag{17.7.14}$$

The calculation is exactly analogous as in the continuous perturbation, and the end result for the transition rate is:

$$\Gamma_{i\to[f]} = \frac{2\pi}{\hbar}|H'_{fi}|^2\rho(E_i+\hbar\omega) \tag{17.7.15}$$

### Emission

In the case where $\omega_{fi}+\omega \approx 0$, then :

$$E_f - E_i + \hbar\omega \approx 0 \implies E_f \approx E_i - \hbar\omega \tag{17.7.16}$$

This is known as **stimulated emission**, the system emits some energy $\hbar\omega$ by transitioning to a lower energy state.

Fermi's golden rule is identical to that for absorption, only that now we need $E_f = E_i - \hbar\omega$:

$$\Gamma_{i\to[f]} = \frac{2\pi}{\hbar}|H'_{fi}|^2\rho(E_i - \hbar\omega) \tag{17.7.17}$$

## 17.8 Dyson series derivation

We provide an alternative derivation for the Fermi golden rule using the Dyson series. This is just a reformulation of the derivations from the previous section but using slightly more advanced methods.

Suppose we have a Hamiltonian $H = H_0 + V(t)$ where $V(t)$ is a time-dependent perturbation that was adiabatically turned on at $t \to -\infty$. We consider two states, $|i\rangle$ and $|f\rangle$, that are eigenstates of the unperturbed hamiltonian $\hat{H}_0$ with energies $E_i$ and $E_f$ respectively. If the system starts out in the state $|i\rangle$ at time $t_0$, then after time $t$ we see that in the Schrodinger picture:

$$|i(t)\rangle\, 55 = e^{-iH_0t/\hbar}\hat{U}_I(t,t_0)e^{iH_0t/\hbar}\,|i\rangle \tag{17.8.1}$$

where $U_I(t,t_0)$ is the interaction-picture propagator, and given by the Dyson series

$$U_I(t,t_0) = \mathcal{T}\bigg[\exp\bigg(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\; V_I(t')\bigg)\bigg] \tag{17.8.2}$$

## 17.9 Hydrogen ionization

Suppose we have a photon of energy $E_\gamma$ incident on a hydrogen atom in its ground state, and ionizing it into the continuum of free states with momentum $k$. This means that:

$$E_f = \frac{\hbar^2k^2}{2m} = E_\gamma - \text{Ry} \geq 0 \tag{17.9.1}$$

where $\text{Ry} = \frac{\alpha\hbar c}{a_0} = 13.6$ eV as always.

To investigate this effect, we will consider harmonic perturbations by the photon. We therefore want the wavelength of the photon to be large enough for the perturbation to be constant over the atom, so $\lambda_\gamma \gg a_0$.

Noting that $\lambda_\gamma = \frac{2\pi c}{\omega}$ implying that:

$$\frac{\lambda}{a_0} = \frac{2\pi c}{\omega}\frac{2\text{Ry}}{\alpha\hbar c} = \frac{4\pi\text{Ry}}{\alpha\hbar\omega} \ll 1 \tag{17.9.2}$$

giving the following upper bound on the energy of the photon:

$$\frac{\hbar\omega}{\text{Ry}} \ll \frac{4\pi}{\alpha} \approx 1772 \tag{17.9.3}$$

Furthermore, we also want $\hbar\omega \gg \text{Ry}$ , because the free electron should not feel the coulomb attraction. If the photon has an energy in the order of magnitude of the Rydberg energy, the electron may be bound again by the proton through the Coulomb interaction.

We consider an electric field at an angle $\theta$ to the $z$-axis:

$$\mathbf{E} = 2\mathbf{E}_0 \cos\omega t \implies V(\mathbf{r}) = -2E_0 r\cos\theta\cos\theta'\cos\omega t \tag{17.9.4}$$

where $\theta'$ is the typical angular coordinate, while $\theta$ is the angle between the E field and the z-axis. This yields a perturbation

$$\delta\hat{H} = 2eE_0 r\cos\theta\cos\theta'\cos\omega = 2H'\cos\omega t \tag{17.9.5}$$

where $H' = eE_0 r\cos\theta\cos\theta'$. As desired, this is a harmonic perturbation, so our discussions from the previous section still hold.

Our initial state is the hydrogen ground state.

$$\psi_{GS}(\mathbf{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \tag{17.9.6}$$

while the final states are:

$$u_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{17.9.7}$$

Here $V$ is the volume over which we normalize the plane wave which we used in the discussion of continuum states.

We orient our axes so that the momentum of the electron $\mathbf{k}$ points along the $z$-axis.

To use Fermi's golden rule, we must first evaluate the matrix elements:

$$\left\langle f \left| \hat{H}' \right| i \right\rangle = \int \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{r}} (eE_0 r\cos\theta\cos\theta') \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} d^3\mathbf{r} \tag{17.9.8}$$

$$= \frac{eE_0}{\sqrt{\pi a_0^3 V}} \int e^{-ikr\cos\theta'} r\cos\theta\cos\theta' e^{-r/a_0} r^2 \sin\theta' dr d\theta' d\phi' \tag{17.9.9}$$

$$= \frac{eE_0}{\sqrt{\pi a_0^3}} \cos\theta 2\pi \int r^3 e^{-r/a_0} e^{-ikr\cos\theta'} \cos\theta' \sin\theta' dr d\theta' \tag{17.9.10}$$

$$= -\frac{eE_0}{\sqrt{\pi a_0^3 V}} \cos\theta 2\pi \int r^3 e^{-r/a_0} e^{-ikr\cos\theta'} \cos\theta' dr d(\cos\theta') \tag{17.9.11}$$

The radial integral is:

$$\int r^3 \exp\left[ -r\left( ik\cos\theta' + \frac{1}{a_0} \right) \right] dr = \frac{6}{\left( ik\cos\theta' + \frac{1}{a_0} \right)^4} \tag{17.9.12}$$

$$= \frac{6a_0^4}{\left( ia_0 k\cos\theta' + 1 \right)^4} \tag{17.9.13}$$

reducing the angular integral to:

$$\int_0^\pi \frac{6a_0^4\cos\theta'}{\left(ia_0k\cos\theta'+1\right)^4}d(\cos\theta') = -\int_{-1}^1 \frac{6ua_0^4}{(ia_0ku+1)^4} \tag{17.9.14}$$

$$= -\frac{16ia_0^5k}{(1+a_0^2k^2)^3} \tag{17.9.15}$$

and hence

$$\left\langle f\left|\hat{H}'\right|i\right\rangle = \frac{eE_0}{\sqrt{\pi a_0^3V}}\cos\theta 2\pi\frac{16ia_0^5k}{(1+a_0^2k^2)^3} \tag{17.9.16}$$

which simplifies to

$$\boxed{\left\langle f\left|\hat{H}'\right|i\right\rangle = 32eE_0\sqrt{\frac{\pi}{a_0^3V}}\frac{ia_0k}{(1+a_0^2k^2)^3}\cos\theta} \tag{17.9.17}$$

The differential transition rate is then:

$$d\Gamma = \frac{2\pi}{\hbar}\frac{V}{8\pi^3}\frac{m}{\hbar^2}k\left(32eE_0\sqrt{\frac{\pi}{a_0^3V}}\frac{ia_0k}{(1+a_0^2k^2)^3}\cos\theta\right)^2 d\Omega \tag{17.9.18}$$

and hence:

$$\frac{d\Gamma}{d\Omega} = \frac{256}{\pi}\frac{e^2E_0^2a_0^2}{\hbar}\frac{ma_0^2}{\hbar^2}\frac{k^3a_0^3}{(1+(ka_0)^2)^6}\cos^2\theta \tag{17.9.19}$$

$$= \frac{256}{\pi}\frac{(eE_0a_0)^2}{\hbar}\frac{ma_0^2}{\hbar}\frac{(ka_0)^3}{(1+(ka_0)^2)^6}\cos^2\theta \tag{17.9.20}$$

Since $\omega \gg \frac{\text{Ry}}{\hbar}$, we can state that $\omega \geq \frac{10\text{Ry}}{\hbar}$ and thus $ka_0 = \sqrt{\frac{\hbar\omega}{\text{Ry}}} \geq 2.5$. The 1 in the denominator can therefore be ignored, giving

$$\frac{d\Gamma}{d\Omega} = \frac{256}{\pi}\frac{(eE_0a_0)^2}{\hbar}\frac{ma_0^2}{\hbar}\frac{1}{(ka_0)^9}\cos^2\theta \tag{17.9.21}$$

Recalling that $\int\cos^2\theta d\Omega = \frac{4\pi}{3}$ we finally find that the total transition rate is:

$$\boxed{\Gamma = \frac{512}{3}\frac{(eE_0a_0)^2}{\hbar\text{Ry}}\frac{1}{(ka_0)^9}} \tag{17.9.22}$$

## 17.10 Einstein's argument

Let's consider a collection of atoms which we model as two-state systems with energy levels $E_a, E_b$, and corresponding eigenstates $|a\rangle, |b\rangle$.

These states interact with light containing photons of temperature $T$. Suppose these photons have frequency $\omega_{ba} = \frac{E_b-E_a}{\hbar}$. If the atom is initially in $|a\rangle$, then it will absorb a photon, transitioning to $|b\rangle$. If instead the atom is initially in $|b\rangle$, then it will emit another photon

by transitioning to $|b\rangle$.

The latter mechanism, which is stimulated emission, is used in laser technology. In lasers we have three levels, the electrons in the ground state are stimulated into transitioning to the third level, and then decay into the second level emitting photons of definite energy.

However, this is not the full picture, we also have a process known as spontaneous emission.

Let's suppose there are $N_a$ atoms in $|a\rangle$ and $N_b$ atoms in $|b\rangle$, as well as photons at temperature $T$. We work in the canonical ensemble, so that $\dot{N}_a = \dot{N}_b = 0$, and we assume we are in thermodynamic equilibrium, implying that:

$$\frac{N_b}{N_a} = \frac{e^{-\beta E_b}}{e^{-\beta E_a}} = e^{-\beta\hbar\omega_{ba}} \tag{17.10.1}$$

Finally, the spectral energy density (energy per volume per frequency range) of the photons is given by Planck's law:

$$u(\omega)d\omega = \frac{\hbar}{\pi^2 c^3}\frac{\omega^3 d\omega}{e^{\beta\hbar\omega} - 1} \tag{17.10.2}$$

In absorption, $|a\rangle \to |b\rangle$, the rate of transition should be proportional to $N_a$, since if more atoms occupy $|a\rangle$ then we would expect the transition to be more likely. Similarly, we also expect that the rate should be proportional to $u(\omega_{ba})$, the number of photons with sufficient energy to be absorbed. Let's then postulate that the transition rate per atom is:

$$\Gamma_{abs} = B_{ab}u(\omega_{ba})N_a \tag{17.10.3}$$

where $B_{ab}$ is a proportionality constant. Similarly, for stimulated emission:

$$\Gamma_{emi} = B_{ba}u(\omega_{ba})N_b \tag{17.10.4}$$

where $B_{ba}$ is a proportionality constant. To have equilibrium, we require that:

$$\dot{N}_b = \Gamma_{abs} - \Gamma_{emi} = B_{ab}u(\omega_{ba})N_a - B_{ba}u(\omega_{ba})N_b = 0 \tag{17.10.5}$$

$$\implies B_{ab}Na = B_{ba}N_b \implies \frac{B_{ab}}{B_{ba}} = e^{-\beta\hbar\omega_{ba}} \tag{17.10.6}$$

This is problematic, because while $\frac{B_{ab}}{B_{ba}}$ is temperature independent as we shall soon calculate, the RHS depends on temperature through $\beta$.

Let's introduce a new process, known as spontaneous emission. This will again be a transition $|b\rangle \to |a\rangle$, which now occurs independently of the number of photons. Its rate will be:

$$\Gamma_{spon} = AN_b \tag{17.10.7}$$

for some proportionality constant $A$. If we introduce this into (17.10.5) then:

$$\dot{N}_b = -AN_b - B_{ba}N_bU(\omega_{ba}) + B_{ab}U(\omega_{ba}N_a0 = 0 \tag{17.10.8}$$

$$\implies A = \left(B_{ab}\frac{N_a}{N_b} - B_{ba}\right)U(\omega_{ba}) \tag{17.10.9}$$

$$\implies U(\omega_{ba}) = \frac{A}{B_{ab}\left(\frac{N_a}{N_b} - \frac{B_{ba}}{B_{ab}}\right)} \tag{17.10.10}$$

$$\implies U(\omega_{ba}) = \frac{A}{B_{ab}}\frac{1}{\left(e^{\beta\hbar\omega_{ba}} - \frac{B_{ba}}{B_{ab}}\right)} \tag{17.10.11}$$

Comparing this with Planck's law, we see that:

$$\frac{B_{ba}}{B_{ab}} = 1,\ \frac{A}{B_{ab}} = \frac{\hbar\omega_{ba}^3}{\pi^2c^3} \tag{17.10.12}$$

The first result is familiar, indeed we saw that the transition rates for absorption and emission in the ionization of hydrogen are equal.

## 17.11 Light-atom interactions

We consider electromagnetic radiation with long wavelengths $\lambda \gg a_0$, so that we may approximate the associated electric field as homogeneous over an atom:

$$\mathbf{E}(t) = 2E_0\cos\omega t\mathbf{n} \tag{17.11.1}$$

implying that the perturbation Hamiltonian takes the form of

$$\delta\hat{H} \equiv V(\mathbf{r},t) = -q\mathbf{E}(t)\cdot\mathbf{r} = -2qE_0\cos\omega t(\mathbf{n}\cdot\mathbf{r}) \tag{17.11.2}$$

where $q$ is the charge of atom which is interacting with the radiation.

Let us define $\mathbf{d} = q\mathbf{r}$ as the dipole operator, giving:

$$\delta\hat{H} = -2E_0\cos\omega t(\mathbf{d}\cdot\mathbf{n}) \tag{17.11.3}$$

We have already solved this problem in the section on the Harmonic perturbation, with $\hat{H}' = -E_0\mathbf{d}\cdot\mathbf{n}$. We found that:

$$P_{b\to a}(t) = \frac{4|H'_{ab}|^2}{\hbar^2}\frac{\sin^2\left(\frac{\omega_{ba}-\omega}{2}t\right)}{(\omega_{ba}^2-\omega)^2} \tag{17.11.4}$$

$$= \frac{4|E_0|^2}{\hbar^2}|\langle a\,|\,\mathbf{d}\cdot\mathbf{n}\,|\,b\rangle\,|^2\frac{\sin^2\left(\frac{\omega_{ba}-\omega}{2}t\right)}{(\omega_{ba}^2-\omega)^2} \tag{17.11.5}$$

Now the electric field is a superposition of incoherent waves with different frequency $\omega_i$,

amplitude $E_0(\omega_i)$ and polarization $\mathbf{n}_i$. Our result therefore gets modified to:

$$P^i_{b\to a} = \frac{4|E_0(\omega_i)|^2}{\hbar^2} |\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle|^2 \frac{\sin^2\left(\frac{\omega_{ba}-\omega_i}{2}t\right)}{(\omega_{ba}^2-\omega_i)^2} \tag{17.11.6}$$

We will sum over all frequencies $\omega_i$ of the electromagnetic radiation near $\omega_{ba}$ (those far away don't contribute much, as we saw in the Fermi golden rule).

Recall that the electric field energy density is:

$$u_E = \frac{|E|^2}{8\pi} = \frac{4E_0^2\cos^2\omega t}{8\pi} \implies \langle u_E\rangle = \frac{1}{4\pi}E_0^2 \tag{17.11.7}$$

The average magnetic field energy density is identical for plane waves:

$$\langle u_B\rangle = \langle u_E\rangle = \frac{1}{4\pi}E_0^2 \tag{17.11.8}$$

giving an energy density of:

$$u(\omega_i) = \frac{1}{2\pi}(E_0(\omega_i))^2 \tag{17.11.9}$$

Substituting (17.11.9) into (17.11.6):

$$P^i_{b\to a} = \frac{8\pi u(\omega_i)}{\hbar^2} |\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle|^2 \frac{\sin^2\left(\frac{\omega_{ba}-\omega_i}{2}t\right)}{(\omega_{ba}^2-\omega_i)^2} \tag{17.11.10}$$

Integrating over frequencies $\omega$ of the radiation in the range near $\omega_{ab}$:

$$\sum_i P^i_{b\to a} = \frac{8\pi}{\hbar^2}\int u(\omega_i)|\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle|^2 \frac{\sin^2\left(\frac{\omega_{ba}-\omega_i}{2}t\right)}{(\omega_{ba}^2-\omega_i)^2}d\omega \tag{17.11.11}$$

$$= \frac{8\pi}{\hbar^2}u(\omega_{ba})\left\langle \left|\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle\right|^2\right\rangle_{\mathbf{n}_i}\int \frac{\sin^2\left(\frac{\omega_{ba}-\omega_i}{2}t\right)}{(\omega_{ba}^2-\omega_i)^2}d\omega \tag{17.11.12}$$

where we averaged the dipole operator matrix elements over all possible polarization directions $\mathbf{n}_i$. Here we went through the same logical reasoning of stating that for sufficiently large $t$ the $\frac{\sin x}{x}$ term contributes only near $\omega = \omega_{ab}$, so $u(\omega)$ may be taken to be constants over this integral. Instead, $|\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle|^2$ may vary even over a small frequency interval, since waves of the same frequency $\omega_{ba}$ can still have different polarizations. This justifies taking the average over all directions.

The integral was evaluated previously to be $\frac{1}{2}t\pi$ so the transition rate will be:

$$\Gamma_{b\to a} = \frac{4\pi^2}{\hbar^2}u(\omega_{ba})\left\langle \left|\langle a \,|\, \mathbf{d}\cdot\mathbf{n}_i \,|\, b\rangle\right|^2\right\rangle_{\mathbf{n}_i} \tag{17.11.13}$$

The matrix element average may be simplified further:

$$\left\langle \left| \langle a \,|\, \mathbf{d} \cdot \mathbf{n}_i \,|\, b \rangle \right|^2 \right\rangle_{\mathbf{n}_i} = \langle (\langle a \,|\, \mathbf{d} \,|\, b \rangle \cdot \mathbf{n}_i)^* (\langle a \,|\, \mathbf{d} \,|\, b \rangle \cdot \mathbf{n}_i) \rangle_{\mathbf{n}_i} \tag{17.11.14}$$

$$= \sum_{jk} \left\langle (\left\langle a \right| \hat{d}^j \left| b \right\rangle \mathbf{n}_i^j)^* (\left\langle a \right| \hat{d}^k \left| b \right\rangle \mathbf{n}_i^k) \right\rangle_{\mathbf{n}_i} \tag{17.11.15}$$

$$= \sum_{jk} (\left\langle a \right| \hat{d}^j \left| b \right\rangle)^* (\left\langle a \right| \hat{d}^k \left| b \right\rangle) \langle n_j n_k \rangle_{\mathbf{n}_i} \tag{17.11.16}$$

$$\tag{17.11.17}$$

Now by symmetry we have that $\langle n_j n_k \rangle_{\mathbf{n}_i} = \frac{1}{3}\delta_{jk}$, therefore:

$$\left\langle \left| \langle a \,|\, \mathbf{d} \cdot \mathbf{n}_i \,|\, b \rangle \right|^2 \right\rangle_{\mathbf{n}_i} = \frac{1}{3} |\langle a \,|\, \mathbf{d} \,|\, b \rangle|^2 \tag{17.11.18}$$

Our final result for the transition rate is then:

$$\Gamma_{b \to a} = \frac{4\pi^2}{3\hbar^2} u(\omega_{ba}) |\langle a \,|\, \mathbf{d} \,|\, b \rangle|^2 \tag{17.11.19}$$

which closely resembles the form of a Fermi Golden rule. It is important to note that this is a transition rate for a single atom, if we had $N$ such atoms the above result would scale accordingly.

Recall that in Einstein's argument we found that $\Gamma_{b \to a} = B_{ab} u(\omega_{ba}) N_b$, so comparison with (17.11.19) gives:

$$B_{ab} = B_{ba} = \frac{4\pi^2}{3\hbar^2} |\langle a \,|\, \mathbf{d} \,|\, b \rangle|^2 \tag{17.11.20}$$

and hence:

$$A = \frac{4\omega_{ba}^3}{3\hbar c^3} |\langle a \,|\, \mathbf{d} \,|\, b \rangle|^2 \tag{17.11.21}$$

From (17.10.7) we find that for a given population $N$ of atoms in state $|b\rangle$:

$$\frac{dN}{dt} = -AN \implies N(t) = N_0 e^{-At} \tag{17.11.22}$$

giving a decay lifetime of $\tau = \frac{1}{A}$.

## 17.12 Selection rules

When evaluating the rate of transitions in Hydrogen, one will notice a pattern in the transitions that are possible. There are a set of rules which state exactly which transitions are realizable, and which aren't.

# Part IV

# Quantum Mechanics of Matter

# The Hydrogen atom: gross structure

18

## 18.1 Setting up the Hamiltonian

We will now tackle the principal problem of quantum mechanics, the hydrogen atom.

Consider a quantum system made of an electron with charge $-e$, mass $m_e$ and a proton with charge $e$ and mass $m_p$. We denote the position and momentum of the proton by $\mathbf{x}_p$ and $\mathbf{p}_p$, and similarly for the electron. These satisfy the commutation relations:

$$[\hat{x}_p^i, \hat{p}_p^j] = i\hbar\delta_{ij} \tag{18.1.1}$$
$$[\hat{x}_e^i, \hat{p}_e^j] = i\hbar\delta_{ij} \tag{18.1.2}$$
$$[\hat{x}_e^i, \hat{p}_p^j] = 0 \tag{18.1.3}$$
$$\tag{18.1.4}$$

We may then write the hamiltonian as:

$$\hat{H} = \frac{\mathbf{p}_p}{2m_p} + \frac{\mathbf{p}_e^2}{2m_e} + V(|\mathbf{x}_e - \mathbf{x}_p|) \tag{18.1.5}$$

We can go to the CM frame of reference where:

$$\mathbf{X} = \frac{m_e\mathbf{x}_e + m_p\mathbf{x}_p}{m_e + m_p} \tag{18.1.6}$$
$$\mathbf{x} = \mathbf{x}_e - \mathbf{x}_p \tag{18.1.7}$$

and similarly:

$$\mathbf{P} = \mathbf{p}_p + \mathbf{p}_e \tag{18.1.8}$$
$$\mathbf{p} = \alpha\mathbf{p}_e - \beta\mathbf{p}_p \tag{18.1.9}$$

To find $\alpha, \beta$ we need to impose:

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \implies \alpha + \beta = 1 \tag{18.1.10}$$

and

$$[\hat{p}_i, \hat{X}_j] = 0 \implies \alpha m_e - \beta m_p = 0 \tag{18.1.11}$$

so that:

$$\alpha = \frac{m_e}{m_e + m_p},\ \beta = \frac{m_p}{m_e + m_p} \tag{18.1.12}$$

Let us introduce the reduced mass $\mu = \frac{m_e m_p}{m_e + m_p}$ and the total mass $M = m_e + m_p$, which gives:

$$\mathbf{p} = \mu\left(\frac{\mathbf{p}_e}{m_p} - \frac{\mathbf{p}_p}{m_e}\right) \tag{18.1.13}$$

Obviously, the commutation rules are still satisfied for the CM operators. Indeed:

$$[\hat{X}_i, \hat{P}_j] = \left[\frac{m\hat{x}_e^i + m_p\hat{x}_p^i}{m + m_p}, \hat{p}_p^j + \hat{p}_e^j\right] = i\hbar\left(\frac{m_p}{m + m_p} + \frac{m}{m + m_p}\right)\delta_{ij} = i\hbar\delta_{ij} \tag{18.1.14}$$

All other canonical commutation relations for this new set of operators may be verified similarly.

To rewrite the Hamiltonian with these new operators, it is useful to express the original operators as:

$$\mathbf{p}_p = \frac{m_p}{M}\mathbf{P} - \mathbf{p} \tag{18.1.15}$$

$$\mathbf{p}_e = \frac{m_e}{M}\mathbf{P} + \mathbf{p} \tag{18.1.16}$$

The kinetic energy operator then becomes:

$$\frac{\mathbf{p}_p^2}{2m_p} + \frac{\mathbf{p}_e^2}{2m_e} = \frac{1}{2m_p}\left(\frac{m_p}{M}\mathbf{P} - \mathbf{p}\right)^2 + \frac{1}{2m_e}\left(\frac{m_e}{M}\mathbf{P} + \mathbf{p}\right)^2 \tag{18.1.17}$$

$$= \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} \tag{18.1.18}$$

as expected, the CM motion and relative motion are independent.

Finally, we write $V(|\mathbf{x}_e - \mathbf{x}_p|) = V(|\mathbf{x}|)$ so the Hamiltonian separates into the center of mass motion and relative motion subsystems' Hamiltonians:

$$\hat{H} = \overbrace{\frac{\mathbf{P}^2}{2M}}^{\hat{H}_{CM}} + \underbrace{\frac{\mathbf{p}^2}{2\mu} + V(r)}_{\hat{H}_{rel}} \tag{18.1.19}$$

where we defined $r = |\mathbf{x}|$.

Since $[\hat{H}_{rel}, \hat{H}_{CM}] = 0$ we may consider wave functions that are of the form:

$$\psi(\mathbf{x}, \mathbf{X}) = \psi_{CM}(\mathbf{X})\psi_{rel}(\mathbf{x}) \tag{18.1.20}$$

Using separation of variables:

$$\psi_{rel}(\mathbf{x})\frac{\mathbf{P}^2}{2M}\psi_{CM}(\mathbf{X}) + \psi_{CM}(\mathbf{X})\left[\frac{\mathbf{p}^2}{2\mu} + V(r)\right]\psi_{rel}(\mathbf{x}) = E\psi_{CM}(\mathbf{X})\psi_{rel}(\mathbf{x}) \tag{18.1.21}$$

and dividing through by $\psi(\mathbf{x}, \mathbf{X})$ we get:

$$\frac{1}{\psi_{CM}(\mathbf{X})}\frac{\mathbf{P}^2}{2M}\psi_{CM}(\mathbf{X}) + \frac{1}{\psi_{rel}(\mathbf{x})}\left[\frac{\mathbf{p}^2}{2\mu} + V(r)\right]\psi_{rel}(\mathbf{x}) = E \tag{18.1.22}$$

As always, the first term is only **X**-dependent, whereas the second term is only **x**-dependent, so the two must be both equal to a constant such that their sum is $E$. We denote these two constants as $E_{CM}$ and $E_{rel}$ so that:

$$\frac{\mathbf{P}^2}{2M}\psi_{CM}(\mathbf{X}) = E_{CM}\psi_{CM}(\mathbf{X}) \tag{18.1.23}$$

$$\left[\frac{\mathbf{p}^2}{2\mu} + V(r)\right]\psi_{rel}(\mathbf{x}) = E_{rel}\psi_{rel}(\mathbf{x}) \tag{18.1.24}$$

$$E_{CM} + E_{rel} = E \tag{18.1.25}$$

Notice that the center of mass system moves as a free particle, and this must of course be the case since the hydrogen atom overall is not in any potential, it is free to roam around.

Finally, we can set the potential $V$ to be the Coulomb potential:

$$V(r) = -\frac{Ze^2}{r} \tag{18.1.26}$$

where $Z$ is the atomic number of the atom (for hydrogen $Z = 1$).

so

$$\left[\frac{\mathbf{p}^2}{2\mu} - \frac{Ze^2}{r}\right]\psi_{rel}(\mathbf{x}) = E_{rel}\psi_{rel}(\mathbf{x}) \tag{18.1.27}$$

## 18.2 Finding the energy levels

Using the results from the chapter on angular momentum:

$$-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2 U_{El}}{\partial r^2} - \frac{l(l+1)}{r^2}U_{El}\right) - \frac{Ze^2}{r}U_{El} = EU_{El} \tag{18.2.1}$$

or more suggestively:

$$\left[\frac{d^2}{dr^2} + \frac{2\mu}{\hbar^2}\left(E + \frac{Ze^2}{r} - \frac{l(l+1)\hbar^2}{2\mu r^2}\right)\right]U_{El} = 0 \tag{18.2.2}$$

Let us also undimensionalize this problem by setting $r = \frac{a_0}{2Z}x$ where $a_0 = \frac{\hbar^2}{\mu e^2}$ is known

as the Bohr radius (this choice simplifies the problem best):

$$\left[\frac{4Z^2}{a_0^2}\frac{d^2}{dx^2}+\frac{2\mu}{\hbar^2}\left(E+\frac{Ze^2}{x}\frac{2Z}{a_0}-\frac{l(l+1)\hbar^2}{2\mu x^2}\frac{4Z^2}{a_0^2}\right)\right]U_{El}=0 \tag{18.2.3}$$

$$\implies\left[\frac{d^2}{dx^2}+\frac{a_0^2}{4Z^2}\frac{2\mu}{\hbar^2}\left(E+\frac{Ze^2}{x}\frac{2Z}{a_0}-\frac{l(l+1)\hbar^2}{2\mu x^2}\frac{4Z^2}{a_0^2}\right)\right]U_{El}=0 \tag{18.2.4}$$

$$\implies\left[-\frac{d^2}{dx^2}+\frac{l(l+1)}{x^2}-\frac{1}{x}-\frac{a_0E}{2Ze^2}\right]U_{El}=0 \tag{18.2.5}$$

where we used $\frac{2\hbar^2Z^2}{\mu a_0^2}=\frac{2\hbar^2Z}{\mu a_0}\frac{me^2}{\hbar^2}=\frac{2Ze^2}{a_0}$ for the coefficient of $E$, and $\frac{a_0^2}{4Z^2}\frac{2\mu}{\hbar^2}\frac{2Z^2e^2}{a_0}=\frac{a_0\mu e^2}{\hbar^2}=1$ for the coefficient of $\frac{1}{x}$.

Let us now define yet another adimensional parameter $\kappa^2=-\frac{a_0E}{2Ze^2}$ so that:

$$\left[-\frac{d^2}{dx^2}+\frac{l(l+1)}{x^2}-\frac{1}{x}+\kappa^2\right]U_{El}=0 \tag{18.2.6}$$

Using a series solutions approach, we will get a three-term recursion relation, which is not quite desirable. We must simplify the equation further, or find a useful mathematical ansatz.

We will use the latter approach, by considering the limiting behaviour of the physical solutions we are interested in.

Clearly, in the limit as $x\to\infty$, the dominant term is $\kappa^2U_{El}$ so that the differential equation becomes:

$$\frac{d^2U_{El}}{dx^2}=\kappa^2u\implies U_{El}=e^{\pm\kappa x} \tag{18.2.7}$$

After all of this renaming-variable-bonanza, we may as well define $\rho=\kappa x$ which is again adimensional. Then (18.2.6) turns into:

$$\left[-\frac{d^2}{d\rho^2}+\frac{l(l+1)}{\rho^2}-\frac{1}{\kappa\rho}+1\right]U_{El}=0 \tag{18.2.8}$$

To check consistency, we see that as $\rho\to\infty$ the dominant term is $U_{El}$ so that

$$\frac{d^2U_{El}}{d\rho^2}=U_{El}\implies U_{El}=e^{\pm\rho} \tag{18.2.9}$$

as we found earlier. On physical grounds, we only want normalizable solutions at infinity so the positive exponent solution may be discarded.

If instead $\rho\to0$ then the dominant term is the centrifugal term, so that:

$$\frac{d^2U_{El}}{d\rho^2}=\frac{l(l+1)}{\rho^2}U_{El}\implies U_{El}\sim\rho^{l+1} \tag{18.2.10}$$

so we may substitute the ansatz $U_{El}=\rho^{l+1}e^{-\rho}f(\rho)$, and hope that this will yield some

other differential equation which can be solved for $f$ through series methods.

The algebra is lengthy but it yields:

$$\rho\frac{d^2f}{d\rho^2} + 2(l+1-\rho)\frac{df}{d\rho} + \left[\frac{1}{\kappa} - 2(l+1)\right]f = 0 \tag{18.2.11}$$

This may at first sight look horrendous, but note that the first term $\rho\frac{d^2f}{d\rho^2}$ lowers the exponent of a series solution by 1, while the second and third terms maintains the exponent, so we have a two-term recursion relation, just as we wanted!

Substituting the series ansatz $f = \sum_{k=0}^{\infty} a_k\rho^k$ we finally get:

$$a_{k+1} = \frac{2(k+l+1) - \frac{1}{\kappa}}{(k+1)(k+2l+2)}a_k \tag{18.2.12}$$

We now have a problem, as $k \to \infty$ we get that $\frac{a_{k+1}}{a_k} \to \frac{2}{k}$, just like in the case of the harmonic oscillator.

In the case of $e^{2\rho}$ we have that:

$$e^{2\rho} = \sum_{k=0}^{\infty} b_k\rho^k = \sum_{k=0}^{\infty}\frac{2^k}{k!}e^k \tag{18.2.13}$$

so that $b_k = \frac{2^k}{k!}$. But then:

$$\frac{b_{k+1}}{b_k} = \frac{2}{k+1} \leq \frac{2}{k} = \frac{a_{k+1}}{a_k} \tag{18.2.14}$$

So our series solution will give $f \sim e^{2\rho}$ giving an unnormalizable solution [1]. We solve this problem by truncating the series at some point.

We want $f$ to be a polynomial of degree $N$ so that $a_N \neq 0$ and $a_{N+1} = 0$. Then we require:

$$a_{N+1} = 0 \implies \frac{1}{\kappa} = 2(N+l+1) \tag{18.2.15}$$

which is the energy quantization condition we were looking for (recall $\kappa^2 \sim E$).

We define the principal quantum number as $n \equiv N + l + 1$ so that:

$$n = \frac{1}{2\kappa} = N + l + 1 \tag{18.2.16}$$

Keeping $n, l$ as independent quantum numbers (we could have chosen $N, l$ as well, but $N$ is not quite as physical as we would want) then:

$$E = -\frac{2Z^2e^2}{a_0}\kappa^2 = -\frac{2Z^2e^2}{a_0}\frac{a_0^2}{4} \tag{18.2.17}$$

[1] indeed using mathematica one can see that for a series with $\frac{a_{k+1}}{a_k}$ then $f(r) \sim 2xe^{2x}$

and thus we finally get:

$$E = -\frac{Z^2 e^2}{2a_0}\frac{1}{n^2} \tag{18.2.18}$$

It is important to note that $0 \leq l \leq n-1$ since $0 \leq N \leq n-1$.

Finally, the eigenfunctions may be written as:

$$\psi_{nlm}(r,\theta,\phi) \propto \left(\frac{r}{a_0}\right)^l f\left(\frac{r}{a_0}\right) e^{-Zr/na_0} Y_{ln}(\theta,\phi) \tag{18.2.19}$$

where $f$ is a polynomial of order $N = n - l - 1$, known as a Laguerre polynomial.

## 18.3 Degeneracy of energy levels

The energy spectrum of the hydrogen atom is highly degenerate. Let us construct the following diagram:

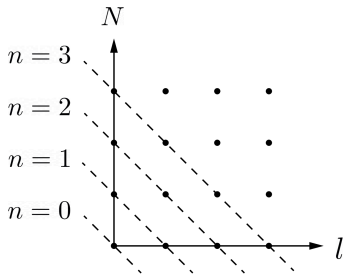



**Figure 18.1**. Degeneracy diagram of hydrogen energy levels

Now for a given $n$, $l$ can take $n$ different values, and so can $N$. We may therefore say that:

| $n$ | $l$ | number of states |
|---|---|---|
| $n=1$ | $l=0$ | 1 state |
| $n=2$ | $l=1,0$ | 4 states |
| $n=3$ | $l=2,1,0$ | 9 states |

More generally we have that for some given principal quantum number $n$, $l$ may take $n-1$ different values. For each given $l$, we have $2l+1$ different values of $m$ so the total degeneracy of the $n$-energy state is:

$$\sum_{l=0}^{n-1}(2l+1) = n(n-1) + n = n^2 \tag{18.3.1}$$

But where does this degeneracy come from? It is quite startling how the energy levels align perfectly, why should there be a degeneracy between solutions with different angular momenta?

Interestingly, one can get physical intuition behind the hydrogen spectrum degeneracy by tackling the problem of classical orbits.

Consider for example $n = 100$ so that $l = 0, 1, 2, ..., 99$. Now it turns out that $l = 0$ corresponds to the most elliptical orbit, whereas $l = 99$ corresponds to the most circular orbit.

Indeed, consider the case where $l = 0$. Then the effective potential is then simply the Coulomb potential, as shown below:

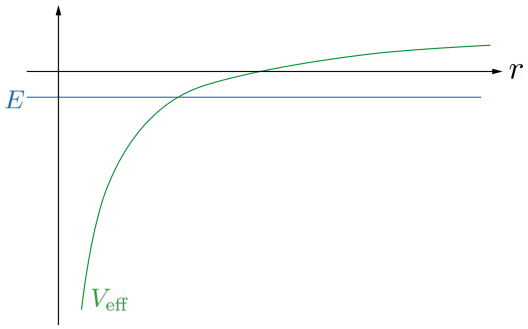



**Figure 18.2**. Dependence of effective potential $V_{eff}$ with $l = 0$ on $r$

We see that the particle can get very very close to the nucleus with extremely high kinetic energy, and then moves up along the potential by losing some kinetic energy, until it reaches the largest allowed radius. This is clearly the case of an elliptic orbit.

Now let's see what happens as we increase $l$, so that the centrifugal potential becomes more and more significant. Then:

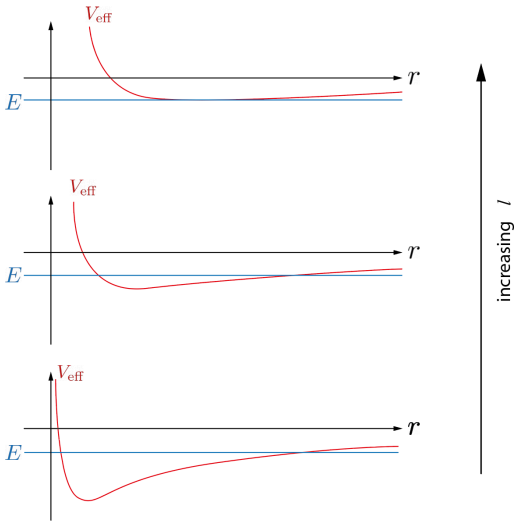



**Figure 18.3**. Dependence of effective potential $V_{eff}$ with increasing $l$ on $r$

We see that increasing the value of $l$ restricts the possibles values of the orbit's radius, until at some $l$ tangency is achieved and the orbit is perfectly circular. Any larger $l$ gives

unphysical solutions.

So we see that for a given $n$, the degeneracy is related to how eccentric the orbits of the electron may be.

## 18.4 Rydberg atoms

**Rydberg atoms**

**Rydberg atoms** are defined as atoms whose outermost electron has a very large principal quantum number $n$.

Now valence electron, the outermost electron, will only observe an effective nuclear charge $+e$. Indeed, even though the nucleus has charge $Ze$, it will experience a shielding effect from the other $Z-1$ electrons of total charge $-(Z-1)e$ whose clouds are closer to the nucleus.

Now the virial theorem applied for a coulomb potential gives:

$$\langle T\rangle = -\frac{1}{2}\langle V\rangle = \frac{e^2}{2}\left\langle\frac{1}{r}\right\rangle \tag{18.4.1}$$

Therefore in an energy $E$ eigenstate:

$$\langle H\rangle = E = \langle T\rangle + \langle V\rangle = \frac{1}{2}\langle V\rangle \tag{18.4.2}$$

so that:

$$\langle V\rangle = 2E = -2\frac{Ze^2}{2a_0}\frac{1}{n^2} \tag{18.4.3}$$

and thus:

$$\left\langle\frac{1}{r}\right\rangle = \frac{1}{a_0 n^2} \implies \langle r\rangle \approx a_0 n^2 \tag{18.4.4}$$

where we used the fact that $\left\langle\frac{1}{r}\right\rangle \approx \frac{1}{\langle r\rangle}$ to the same order of magnitude.

Indeed, by using some properties of Laguerre polynomials one can show that in the state $|nlm\rangle$:

$$\langle r\rangle_{nlm} = a_0 n^2\left[1+\frac{1}{2}\left(1-\frac{l(l+1)}{n^2}\right)\right] \tag{18.4.5}$$

It turns out that this is also the modal radius, the radius at which the radial probability density $\mathbb{P}(r)$ reaches a maximum value.

Inded, consider:

$$\psi_{nlm} = \mathcal{N} r^l f_{nl}(r) e^{-r/na_0} Y_l^m(\theta,\phi) \equiv R_{nl}(r) Y_l^m(\theta,\phi) \tag{18.4.6}$$

Then, the probability that the particle is in some spherical shell of thickness $dr$ is:

$$\mathbb{P}(r)dr = r^2 dr \int |\psi_{nlm}|^2 \sin\theta d\theta d\phi = r^2 R_{nl}^2(r)dr \tag{18.4.7}$$

where we used the orthonormality of the spherical harmonics. Now we have:

$$R_{nl}(r) \sim r^l(a_0 + ... + a_N r^N)e^{-r/na_0} \tag{18.4.8}$$

Now since we expect the radius to be quite large, we can only consider the $r^N$ dominant term

$$R_{nl}(r) \sim r^{n-1}e^{-r/na_0} \implies \mathbb{P}(r) \sim r^{2n}e^{-r/na_0} \tag{18.4.9}$$

This reaches a maximum when:

$$2nr^{2n-1} - \frac{2r^{2n}}{na_0} = 0 \implies r \sim a_0 n^2 \tag{18.4.10}$$

as we found earlier.

## 18.5 Ladder operator formalism

Quite like the harmonic oscillator and the isotropic oscillator, we can use the ladder operator formalism to tackle the problem algebraically.

Let $\hat{p}_r = i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)$ be the radial momentum operator, so that:

$$\hat{p}_r^2 = -\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) \tag{18.5.1}$$

One may ask why $\hat{\mathbf{r}} \cdot \mathbf{p}$ was not chosen instead. The answer is that this does not satisfy hermiticity:

$$(\hat{\mathbf{r}} \cdot \mathbf{p})^\dagger = \mathbf{p} \cdot \hat{\mathbf{r}} \tag{18.5.2}$$

We can get around this problem by applying the symmetrization procedure:

$$\hat{p}_r = \frac{1}{2}(\hat{\mathbf{r}} \cdot \mathbf{p} + \mathbf{p} \cdot \hat{\mathbf{r}}) \tag{18.5.3}$$

We then get that in the position basis:

$$\hat{p}_r = -\frac{i\hbar}{2}\left(\frac{1}{r}\mathbf{r} \cdot \nabla + \nabla \cdot \hat{\mathbf{r}}\right) \tag{18.5.4}$$

$$= -\frac{i\hbar}{2}\left(\frac{\partial}{\partial r} + \frac{3}{r} - \frac{r}{r^2} + \frac{\partial}{\partial r}\right) \tag{18.5.5}$$

$$= i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right) \tag{18.5.6}$$

as desired.

Note that $\hat{r}, \hat{p}_r$ are canonical conjugates:

$$[\hat{r}, \hat{p}_r] = i\hbar \tag{18.5.7}$$

We may then define:

$$\hat{A}_l = \frac{a_0}{\sqrt{2}}\left(\frac{i\hat{p}_r}{\hbar} - \frac{l+1}{r} + \frac{Z}{a_0(l+1)}\right) \tag{18.5.8}$$

where $a_0$ is the Bohr radius as always.

We then get that:

$$\hat{A}_l^\dagger \hat{A}_l = \frac{a_0^2}{2}\left(-\frac{i\hat{p}_r}{\hbar} - \frac{l+1}{r} + \frac{Z}{a_0(l+1)}\right)\left(\frac{i\hat{p}_r}{\hbar} - \frac{l+1}{r} + \frac{Z}{a_0(l+1)}\right) \tag{18.5.9}$$

$$= \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \left(\frac{Z}{(l+1)a_0} - \frac{(l+1)}{r}\right) + \frac{i(l+1)}{\hbar}\left[\hat{p}_r, \frac{1}{r}\right]\right) - \frac{2Z}{a_0 r} \tag{18.5.10}$$

We now use the important commutator property:

$$[\hat{A}, f(\hat{B})] = [\hat{A}, \hat{B}]\frac{df}{d\hat{B}} \tag{18.5.11}$$

which we have already used when deriving $\hat{L}_z$ in the passive picture. Then we find that:

$$\hat{A}_l = \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{(l+1)^2}{r^2} - \frac{2Z}{a_0 r} - \frac{i(l+1)}{\hbar}[\hat{p}_r, \hat{r}]\frac{1}{\hat{r}^2}\right) \tag{18.5.12}$$

$$= \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{(l+1)^2}{r^2} - \frac{2Z}{a_0 r} - \frac{l+1}{\hat{r}^2}\right) \tag{18.5.13}$$

$$= \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{l(l+1)}{r^2} - \frac{2Z}{a_0 r}\right) \tag{18.5.14}$$

$$= \frac{a_0^2 \mu}{\hbar^2}\hat{H}_l + \frac{Z^2}{2(l+1)^2} \tag{18.5.15}$$

so that the Hamiltonian may be written as:

$$\hat{H}_l = \frac{\hbar^2}{a_0^2 \mu}\left(\hat{A}_l^\dagger \hat{A}_l - \frac{Z^2}{2(l+1)^2}\right) \tag{18.5.16}$$

If instead we had evaluated $\hat{A}_l \hat{A}_l^\dagger$, then the only difference would have been in the sign of

the commutator $\left[\hat{p}_r, \frac{1}{r}\right]$, so we would get that:

$$\hat{A}_l = \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{(l+1)^2}{r^2} - \frac{2Z}{a_0 r} + \frac{i(l+1)}{\hbar}[\hat{p}_r, \hat{r}]\frac{1}{\hat{r}^2}\right) \tag{18.5.17}$$

$$= \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{(l+1)^2}{r^2} - \frac{2Z}{a_0 r} + \frac{l+1}{\hat{r}^2}\right) \tag{18.5.18}$$

$$= \frac{a_0^2}{2}\left(\frac{\hat{p}_r^2}{\hbar^2} + \frac{Z^2}{(l+1)^2 a_0^2} + \frac{(l+1)(l+2)}{r^2} - \frac{2Z}{a_0 r}\right) \tag{18.5.19}$$

$$= \frac{a_0^2 \mu}{\hbar^2}\hat{H}_{l+1} + \frac{Z^2}{2(l+1)^2} \tag{18.5.20}$$

where $\hat{H}_{l+1}$ is the hamiltonian with $l' = l+1$. Now let us see how $\hat{A}_l$ commutes with $\hat{H}_l$:

$$[\hat{A}_l, \hat{H}_l] = \frac{\hbar^2}{\mu a_0^2}[\hat{A}_l, \hat{A}_l^\dagger \hat{A}_l] = \frac{\hbar^2}{\mu a_0^2}[\hat{A}_l, \hat{A}_l^\dagger]\hat{A}_l \tag{18.5.21}$$

so we need to find $[\hat{A}_l, \hat{A}_l^\dagger]$:

$$[\hat{A}_l, \hat{A}_l^\dagger] = \frac{a_0^2 \mu}{\hbar^2}(\hat{H}_{l+1} - \hat{H}_l) \tag{18.5.22}$$

Consequently:

$$[\hat{A}_l, \hat{H}_l] = (\hat{H}_{l+1} - \hat{H}_l)\hat{A}_l \implies \boxed{\hat{A}_l \hat{H}_l = \hat{H}_{l+1}\hat{A}_l} \tag{18.5.23}$$

So clearly if we act $\hat{A}_l$ on some eigenstate $|E, l\rangle$, it will create another eigenstate with same energy $E$, but associated to an effective potential with $l' = l+1$.

Indeed:

$$\hat{A}_l \hat{H}_l |E, l\rangle = E\hat{A}_l |E, l\rangle = \hat{H}_{l+1}\hat{A}_l |E, l\rangle \tag{18.5.24}$$

so $\hat{A}_l |E, l\rangle$ is indeed an eigenvalue of $\hat{H}_{l+1}$.

This explains the alignment of the energy levels in the hydrogen spectrum, and thus its degeneracy. By applying $\hat{A}_l$ we can move from a state with quantum numbers $n, l$ to another state with quantum numbers $n, l+1$.[2].

However, we can't expect to be able to increase the angular momentum indefinitely. There comes a point at which the effective potential just barely touches the bound state energy. This, as we saw earlier, corresponds to the maximum angular momentum orbit, the particle is forbidden to be anywhere for larger values of $l$. Once we get to a circular orbit we cannot increase the angular momentum anymore without also increasing the energy. Hence, one might expect that in the quantum mechanical framework too there must be an upper limit to $l$, which we denote as $L$.

[2] the astute reader may have realized that $\hat{A}_l$ really only acts on the radial part of the wave-function. To raise the angular momentum for the entire wave-function one needs to construct the ladder operators using the Runge-Lenz operator

Consequently we must set:

$$\hat{A}_L |E, L\rangle = 0 \implies \left\langle E, L \middle| \hat{A}_L^\dagger \hat{A}_L \middle| E, L \right\rangle = 0 \tag{18.5.25}$$

and using (18.5.16) we find:

$$\left\langle E, L \middle| \frac{a_0^2 \mu}{\hbar^2} \hat{H}_L + \frac{Z^2}{2(L+1)^2} \middle| E, L \right\rangle = 0 \tag{18.5.26}$$

$$\iff E_L = -\frac{\hbar^2 Z^2}{2\mu(L+1)^2 a_0^2} \tag{18.5.27}$$

Defining $n \equiv L + 1$ as the principal quantum number we then get:

$$E_n = -\frac{Z^2 e^2}{2a_0} \frac{1}{n^2} \tag{18.5.28}$$

as found previously.

Defining the Rydberg constant to be:

$$\text{Ry} = \frac{e^2}{2a_0} \tag{18.5.29}$$

so that:

$$E_n = -\frac{Z^2 \text{Ry}}{n^2} \tag{18.5.30}$$

For hydrogen, we get that $\text{Ry} = 13.6$ eV, so that

$$E_n = -\frac{13.6 \text{ eV}}{n^2}, \quad \text{(hydrogen)} \tag{18.5.31}$$

For single-electron systems with other $Z$ we the find that:

$$E_n = -Z \frac{\mu}{\mu_H} \frac{13.6 \text{ eV}}{n^2}, \quad \text{(hydrogen-like)} \tag{18.5.32}$$

### 18.5.1 Muonic atoms

Muons ($\mu$) are elementary particles similar to electrons (same charge $-e$, so that $Z = 1$), but about 207 times more massive, so that $m_\mu = 207 m_e$. When captured by an proton, a muon may form a muonic atom. Recall that $m_p = 1840 m_e$ so that the reduced mass of a muonic atom becomes:

$$\mu = \frac{207 \cdot 1840 m_e^2}{2047 m_e} \approx 186 m_e \tag{18.5.33}$$

which is significantly greater than that for hydrogen, which we may approximate to $\mu_H \approx m_e$. The energy levels of a muonic atom are then:

$$E_n = -186 \frac{13.6 \text{ eV}}{n^2}, \tag{18.5.34}$$

so about two orders of magnitude grater than the hydrogen atom energy levels.

Consequently when photons are emitted due to transitions in muonic atom states, their energies will be in the $keV$ range rather than the $eV$ range, falling in the X-ray spectrum. These may be detected using solid-state detectors rather than optical spectroscopes.

Also, the Bohr radius $a_0 = \frac{\hbar^2}{\mu e^2} = \frac{1}{186}\frac{\hbar^2}{\mu_H e^2}$ is scaled by a factor of $\frac{1}{186}$. Consequently there will be a much higher change that the electron is found within the radius of the proton. This is problematic, because the electrostatic interaction between the proton and nucleus is different when they are separated versus when the first is inside the other. Here we will need to evaluate the electric potential energy of an electron inside the proton and use this as a perturbation.

### 18.5.2 Emission spectra

Suppose we have an electron transition from $n_1$ to $n_2$. By energy conservation a photon must be emitted with frequency given by Planck's law:

$$\nu = \frac{\Delta E}{h} = \frac{Z^2 \mathrm{Ry}}{h}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \tag{18.5.35}$$

For each $n_2$ we will therefore have a series of lines. The most important are for the hydrogen atom:

(i) Lyman series ($n_2 = 1$).

(ii) Balmer series ($n_2 = 2$).

(iii) Paschen series ($n_2 = 3$).

## 18.6 Eigenfunctions using ladder operators

### Circular orbits

Suppose we wish to express $|E, l, m\rangle$ in the position representation. To do so, we may firstly use the independence of radial and angular variables to write:

$$\langle \mathbf{r} \,|\, E, l, m\rangle = R_{El}(r) Y_l^m(\theta, \phi) \tag{18.6.1}$$

Now if we consider the maximum angular momentum state $|E, n-1\rangle$ (note that this is only the radial part, it does not include the spherical harmonics) then we expect:

$$\hat{A}_{n-1}\,|E, n-1\rangle = 0 \implies \left(\frac{i\hat{p}_r}{\hbar} - \frac{n}{r} + \frac{Z}{na_0}\right)|E, n-1\rangle = 0 \tag{18.6.2}$$

Remembering that in the position representation $\hat{p}_r = -i\hbar(\frac{\partial}{\partial r} + \frac{1}{r})$ we find that:

$$\left(\frac{\partial}{\partial r} - \frac{n-1}{r} + \frac{Z}{na_0}\right) R_{n,n-1} = 0 \tag{18.6.3}$$

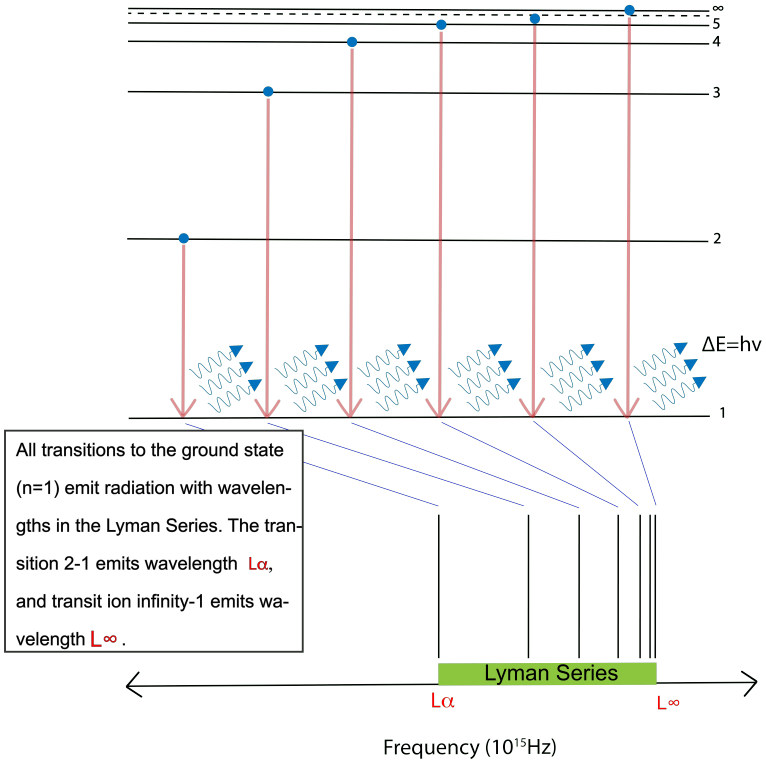



**Figure 18.4**. Lyman series of emission lines

which is just a first order linear ODE, and has integrating factor:

$$\Lambda = \exp\left[\int \frac{Z}{na_0} - \frac{n-1}{r}dr\right] = r^{1-n}e^{Zr/na_0} \tag{18.6.4}$$

so that:

$$\frac{d}{dx}(R_{n,n-1}(r)\Lambda = 0 \tag{18.6.5}$$

and thus

$$R_{n,n-1}(r) = \mathcal{N}r^{n-1}e^{-Zr/na_0} \tag{18.6.6}$$

where $\mathcal{N}$ is a normalization constant.

For Hydrogen, the ground state corresponds to $\psi_{GS}(r,\theta,\phi) = R_{1,0}(r)Y_{0,0}(\theta,\phi)$ and has the simple expression:

$$\psi_{GS}(r,\theta,\phi) = \mathcal{N}e^{-r/a_0} \tag{18.6.7}$$

The normalization constant is easily found by integration:

$$\mathcal{N}^2\int_0^\infty e^{-2r/a_0}4\pi r^2 dr = \mathcal{N}^2\frac{a_0^3}{4}4\pi = 1 \implies \mathcal{N} = \frac{1}{\sqrt{\pi a_0^3}} \tag{18.6.8}$$

so

$$\psi_{GS}(r,\theta,\phi) = \frac{2}{\sqrt{a_0^3}} e^{-r/a_0} \tag{18.6.9}$$

More generally, we find that for $R_{n,n-1}$:

$$\mathcal{N} = \frac{1}{\sqrt{2n!}} \left(\frac{2Z}{na_0}\right)^{3/2} \left(\frac{2Z}{na_0}\right)^{n-1} r^{n-1} e^{-Zr/na_0} \tag{18.6.10}$$

### Elliptic orbits

This all works for the most circular orbits, but what about the elliptic orbits? We suspect that just like $\hat{A}_{l-1}$ increases the angular momentum of the state while maintaining the energy, $\hat{A}^\dagger_{l-1}$ must instead lower the angular momentum.

To show this, we must prove that:

$$[\hat{H}_l, \hat{A}^\dagger_{l-1}] = (\hat{H}_l - \hat{H}_{l-1})\hat{A}^\dagger_{l-1} \implies \hat{A}^\dagger_{l-1}\hat{H}_l = \hat{H}_{l-1}\hat{A}^\dagger_{l-1} \tag{18.6.11}$$

so our goal will be to evaluate $[\hat{H}_l, \hat{A}^\dagger_{l-1}]$.

Recall:

$$[\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}] = \frac{a_0^2\mu}{\hbar^2}(\hat{H}_l - \hat{H}_{l-1}) \tag{18.6.12}$$

We then find that:

$$[[\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}], \hat{A}^\dagger_{l-1}] = \frac{a_0^2\mu}{\hbar^2}([\hat{H}_l, \hat{A}^\dagger_{l-1}] - [\hat{H}_{l-1}, \hat{A}^\dagger_{l-1}]) \tag{18.6.13}$$

$$[\hat{H}_l, \hat{A}^\dagger_{l-1}] = \frac{\hbar^2}{a_0^2\mu}[[\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}], \hat{A}^\dagger_{l-1}] + [\hat{H}_{l-1}, \hat{A}^\dagger_{l-1}] \tag{18.6.14}$$

Now we may expand the double commutator:

$$[[\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}], \hat{A}^\dagger_{l-1}] = [\hat{A}_{l-1}\hat{A}^\dagger_{l-1}\hat{A}^\dagger_{l-1}] - [\hat{A}^\dagger_{l-1}\hat{A}_{l-1}\hat{A}^\dagger_{l-1}] \tag{18.6.15}$$

and we may also substitute (18.5.16) into $[\hat{H}_{l-1}, \hat{A}^\dagger_{l-1}]$ to find:

$$[\hat{H}_{l-1}, \hat{A}^\dagger_{l-1}] = \frac{\hbar^2}{a_0^2\mu}[\hat{A}^\dagger_{l-1}\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}] \tag{18.6.16}$$

Hence we find that:

$$[\hat{H}_l, \hat{A}^\dagger_{l-1}] = \frac{\hbar^2}{a_0^2\mu}\Big([\hat{A}_{l-1}\hat{A}^\dagger_{l-1}\hat{A}^\dagger_{l-1}] - [\hat{A}^\dagger_{l-1}\hat{A}_{l-1}\hat{A}^\dagger_{l-1}] + [\hat{A}^\dagger_{l-1}\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}]\Big) \tag{18.6.17}$$

$$= \frac{\hbar^2}{a_0^2\mu}[\hat{A}_{l-1}\hat{A}^\dagger_{l-1}, \hat{A}^\dagger_{l-1}] \tag{18.6.18}$$

$$= \frac{\hbar^2}{a_0^2\mu}(\hat{A}_{l-1}\cancelto{0}{[\hat{A}^\dagger_{l-1}, \hat{A}^\dagger_{l-1}]} + [\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}]\hat{A}^\dagger_{l-1}) \tag{18.6.19}$$

$$= \frac{\hbar^2}{a_0^2\mu}[\hat{A}_{l-1}, \hat{A}^\dagger_{l-1}]\hat{A}^\dagger_{l-1}) \tag{18.6.20}$$

$$= (\hat{H}_l - \hat{H}_{l-1})\hat{A}^\dagger_{l-1} \tag{18.6.21}$$

After this whole commutator party, we may conclude that:

$$E\hat{A}^\dagger_{l-1}\,|E,l\rangle = \hat{A}^\dagger_l\hat{H}_l\,|E,l\rangle \tag{18.6.22}$$

$$= (\hat{H}_l\hat{A}^\dagger_{l-1} + [\hat{A}^\dagger_{l-1}, \hat{H}_l])\,|E,l\rangle \tag{18.6.23}$$

$$= (\hat{H}_l\hat{A}^\dagger_{l-1} - (\hat{H}_l - \hat{H}_{l-1})\hat{A}^\dagger_{l-1})\,|E,l\rangle \tag{18.6.24}$$

$$= \hat{H}_{l-1}\hat{A}^\dagger_{l-1}\,|E,l\rangle \tag{18.6.25}$$

so we see that $\hat{A}^\dagger_{l-1}$ produces a state with same energy eigenvalue, but lower angular momentum, as expected. We may therefore apply this operator on the most circular orbit $|E, n-1\rangle$ to produce elliptic orbits.

Now in the position representation we have that:

$$R_{n,n-2} = \hat{A}^\dagger_{n-2}\Big(-\frac{\partial}{\partial r} - \frac{n}{r} + \frac{Z}{(n-1)na_0}\Big)R_{n,n-1} \tag{18.6.26}$$

$$= (ar^{n-2} + br^{n-1})e^{-Zr/na_0} \tag{18.6.27}$$

To explain the last line, we note that the derivative operator will produce a term with $r^{n-1}$ when acting on the exponential, and another term $r^{n-2}$ when acting on the $r^{n-1}$ term. The $\frac{n}{r}$ will also reduce the power of $r$ to $r^{n-2}$. Finally, the $\frac{Z}{(n-1)na_0}$ term does nothing. Hence, overall we will have $r^{n-1}$ and $r^{n-2}$ terms multiplied by the exponential which survives differentiation. We may write this more neatly as:

$$R_{n,n-2} = (a + br)r^{n-2}e^{-Zr/na_0} \tag{18.6.28}$$

which has one node where $R(r) = 0$ due to the linear term.

Clearly, if we apply $\hat{A}^\dagger_{n-3}$ then we get:

$$R_{n,n-2} = (a + br + cr^2)r^{n-3}e^{-Zr/na_0} \tag{18.6.29}$$

which will have two nodes due to the quadratic polynomial. More generally we see that:

$$R_{n,l} = \begin{pmatrix} \text{polynomial of power} \\ n-l \end{pmatrix} r^{l-1} e^{-Zr/na_0} \tag{18.6.30}$$

which will have $n - l$ nodes. These will be spherical surfaces of fixed radii.

## 18.7 Spin degrees of freedom

Throughout this study of the gross structure of the hydrogen atom we have not yet included the spin degrees of freedom. Since the Hamiltonian is independent of spin, we need only to append an extra spin ket to the eigenstates $|n, l, m\rangle$. The resulting kets are $|n, l, m, m_s\rangle$, where $m_s$ is the spin of the electron along $z$.

The degeneracy of the state with principal quantum number $n$ is then doubled to $2n^2$. The energy levels however are unaffected.

Let us now consider a hydrogen atom in a magnetic field $\mathbf{B} = B\hat{\mathbf{k}}$. There will now be two additional terms due to the orbital angular momentum and spin angular momentum interactions with the magnetic field, which are $-\boldsymbol{\mu}_L \cdot \mathbf{B}$ and $-\boldsymbol{\mu}_S \cdot \mathbf{B}$, so that the total Hamiltonian becomes:

$$\hat{H} = \hat{H}_C - \Big( - \frac{eB}{2m_e c} \Big) \hat{L}_z - \Big( - \frac{eB}{m_e c} \Big) \hat{S}_z \tag{18.7.1}$$

We ignore the coupling due to the proton's angular momenta, since $m_p \ll m_e$ [3]. Since these interaction terms commute with the coulomb hamiltonian, they are simultaneously diagonalizable and hence:

$$\hat{H} \, |nlmm_s\rangle = \Big( - \frac{\text{Ry}}{n^2} + \frac{eB\hbar}{2m_e c}(m + 2m_s) \Big) \, |nlmm_s\rangle \tag{18.7.2}$$

For the example, the ground state splits into two levels:

$$E_1 = -\text{Ry} \pm \frac{e\hbar B}{2mc} \tag{18.7.3}$$

Instead, the first excited state splits into five levels:

$$E_2 = -\frac{\text{Ry}}{4} \pm \frac{eB\hbar}{2mc}(2, 1 \text{ or } 0) \tag{18.7.4}$$

More generally $E_n$ splits into $2n + 1$ states, instead of $2n^2$ without the magnetic field.

[3] the correction terms if we take into account the spin and orbital angular momenta of spin are of order $\frac{m}{M}$ and $\frac{m^2}{M^2}$ respectively.

## 18.8 Expectation values of $r^{-k}$

We have already found using the virial theorem that:

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{na_0^2} \tag{18.8.1}$$

Now from perturbation theory we know that adding a term $\alpha \hat{H}_1$ then:

$$\left. \frac{dE}{d\alpha} \right|_{\alpha=0} = \left\langle E \left| \hat{H}_1 \right| E \right\rangle \tag{18.8.2}$$

so if we let $\hat{H}_1 = -\frac{\hbar^2}{2\mu r^2}$. The hamiltonian now takes the form:

$$\hat{H}_{new} = \frac{\hbar^2}{2\mu}\left( \frac{\hat{p}_r^2}{\hbar^2} + \frac{l(l+1) - \alpha}{r^2} - \frac{2Z_0}{ra_0} \right) \tag{18.8.3}$$

If we set $l'(l'+1) = l(l+1) - \alpha$ then we have that the energy levels are identical to those of the normal hydrogenic atom:

$$E = -\frac{Z^2 e^2}{2a_0} \frac{1}{(l'+k+1)^2} \tag{18.8.4}$$

Hence:

$$-\frac{\hbar^2}{2\mu}\left\langle \frac{1}{r^2} \right\rangle = \frac{d}{d\alpha}\left( -\frac{Z^2 e^2}{2a_0} \frac{1}{(l'+k+1)^2} \right)_{\alpha=0} \tag{18.8.5}$$

$$= \frac{Z^2 e^2}{a_0} \frac{1}{(l'+k+1)^3} \left. \frac{dl'}{d\alpha} \right|_{\alpha=0} \tag{18.8.6}$$

$$= -\frac{Z^2 e^2}{a_0} \frac{1}{(l'+k+1)^3} \frac{1}{(2l'+1)} \Bigg|_{\alpha=0} \tag{18.8.7}$$

$$= -\frac{Z^2 e^2}{a_0} \frac{1}{n^3} \frac{1}{(2l+1)} \tag{18.8.8}$$

So that:

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2}{a_0^2 n^3 (l+\frac{1}{2})} \tag{18.8.9}$$

Finally, we may evaluate $\langle \frac{1}{r^3} \rangle$ by considering $[\hat{H}, \hat{p}_r]$. Its expectation value for a stationary state is null by Ehrenfest's theorem, hence:

$$\left\langle [\hat{H}, \hat{p}_r] \right\rangle = -\frac{l(l+1)\hbar^2}{2\mu} \left\langle [\hat{r}^{-2}, \hat{p}_r] \right\rangle - Ze^2 \left\langle [\hat{r}^{-1}, \hat{p}_r] \right\rangle \tag{18.8.10}$$

Now using the relation $[f(\hat{A}), \hat{B}] = [\hat{A}, \hat{B}]\frac{\partial f}{\partial \hat{A}}$ then:

$$\frac{l(l+1)\hbar^2}{2\mu}\left\langle [\hat{r}^{-2}, \hat{p}_r] \right\rangle = -\frac{2l(l+1)\hbar^2}{2\mu\hat{r}^3} i\hbar \tag{18.8.11}$$

$$-Ze^2\left\langle [\hat{r}^{-1}, \hat{p}_r] \right\rangle = i\hbar\frac{Ze^2}{\hat{r}^2} \tag{18.8.12}$$

so that:

$$\left\langle [\hat{H}, \hat{p}_r] \right\rangle = -i\hbar\frac{2l(l+1)\hbar^2}{2\mu}\left\langle \frac{1}{r^3} \right\rangle + i\hbar Ze^2\left\langle \frac{1}{r^2} \right\rangle \tag{18.8.13}$$

$$= -i\hbar\frac{l(l+1)\hbar^2}{\mu}\left\langle \frac{1}{r^3} \right\rangle + i\hbar Ze^2\frac{Z^2}{a_0^2 n^3 (l+\frac{1}{2})} = 0 \tag{18.8.14}$$

implying that:

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3 n^3 l(l+1)(l+\frac{1}{2})} \tag{18.8.15}$$

# The Hydrogen atom: fine structure

19

## 19.1 Scales

We recall from the previous chapter that the Bohr radius is a fundamental length scale of the hydrogen atom, found by setting the kinetic energy and coulomb potential energy to equal orders of magnitude:

$$a_0 = \frac{\hbar^2}{\mu e^2} \approx 53\text{pm} \tag{19.1.1}$$

where $\mu = \frac{m_e m_p}{m_e + m_p} \approx m_e$ is the reduced mass of the system. For $Z$ protons we will have to substitute $e^2 \to Ze^2$. The energy levels are:

$$E_n = -\frac{e^2}{2a_0}\frac{1}{n^2} \approx -\frac{13.6\text{ eV}}{n^2},\ n = 1, 2, ... \tag{19.1.2}$$

We define the fine structure constant as:

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \tag{19.1.3}$$

so that

$$\frac{e^2}{a_0} = \frac{me^4}{\hbar^2} = \alpha^2 mc^2 \implies E_n = -\frac{\alpha^2}{2}mc^2\frac{1}{n^2} \tag{19.1.4}$$

The momentum of the electron is approximately

$$p \approx \frac{\hbar}{a_0} = \alpha mc \implies v \approx \alpha c = \frac{c}{137} \tag{19.1.5}$$

## 19.2 Angular momentum bases

The electron spin may be specified by two quantum numbers $(s, m_s) = \left(\frac{1}{2}, \pm\frac{1}{2}\right)$. The electron's orbital angular momentum instead is specified by $(l, m)$ where $0 \leq l \leq n-1$ and $|m| \leq l$. Hence in the uncoupled basis we use the quantum numbers $(n, l, m, m_s)$ as was explained in Chapter 13.

However, we may also directly express the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ of the electron. So we transform from the uncoupled basis $(l, m, s, m_s)$ to the coupled basis $(s, l, j, m_j, l)$. We have traded knowledge of $m, m_s$ (state of individual components of an-

gular momentum) for knowledge of $j, m_j$ (state of total angular momentum). Since $s = \frac{1}{2}$ is assumed to be understood, we may omit it and simply write $(l, j, m_j)$.

Furthermore, recall that:

$$l \otimes \frac{1}{2} = \left(l + \frac{1}{2}\right) \oplus \left(l - \frac{1}{2}\right) \tag{19.2.1}$$

which states that the uncoupled basis states may be expressed as linear combinations of $l + \frac{1}{2}$-spin states and $l - \frac{1}{2}$-spin states. The most important relations for the hydrogen atom are:

$$0 \otimes \frac{1}{2} = \frac{1}{2},\ 1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2},\ 2 \otimes \frac{1}{2} = \frac{5}{2} \oplus \frac{3}{2} \tag{19.2.2}$$

We now introduce the spectroscopic notation for our uses:

(i) the orbital angular momentum number $l$ is associated to a letter $L$ as shown:

| $l$ | $L$ |
|---|---|
| 0 | $S$ |
| 1 | $P$ |
| 2 | $D$ |
| 3 | $F$ |

(ii) Add a right subscript to indicate the total angular momentum number $j$: $L_j$.

(iii) Add a left superscript to indicate the principal quantum number $n$: ${}^{n}L_j$.

This gives the following structure of hydrogen energy levels:

| | $S$<br>$\ell = 0$ | $P$<br>$\ell = 1$ | $D$<br>$\ell = 2$ |
|---|---|---|---|
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n = 4$ | ${}^{4}S_{\frac{1}{2}}$ | | |
| $n = 3$ | ${}^{3}S_{\frac{1}{2}}$<br>(2) | ${}^{3}P_{\frac{3}{2}}$<br>(6) ${}^{3}P_{\frac{1}{2}}$ | ${}^{3}D_{\frac{3}{2}}$<br>(10) ${}^{3}D_{\frac{5}{2}}$ |
| $n = 2$ | ${}^{2}S_{\frac{1}{2}}$<br>(2) | ${}^{2}P_{\frac{3}{2}}$<br>(6) ${}^{2}P_{\frac{1}{2}}$ | |
| $n = 1$ | ${}^{1}S_{\frac{1}{2}}$<br>(2) | | |

## 19.3 Constructing the fine structure hamiltonian

### Relativistic correction

Recall from special relativity the relation:

$$E^2 - \mathbf{p}^2c^2 = m^2c^4 \tag{19.3.1}$$

so promoting $E$ to $\hat{H}$ we get that:

$$\hat{H} = \sqrt{\mathbf{p}^2c^2 + m^2c^4} = mc^2\sqrt{1 + \frac{\mathbf{p}^2}{m^2c^2}} \tag{19.3.2}$$

$$\approx mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{(\mathbf{p}^2)^3}{8m^3c^2} \tag{19.3.3}$$

where we expanded for non-relativistic momenta. We have recovered the typical term $\frac{\mathbf{p}^2}{2m}$, but (ignoring the rest mass term) we now have an additional perturbation

$$\delta H_{rel} = -\frac{(\mathbf{p}^2)^3}{8m^3c^2} \tag{19.3.4}$$

### Spin-Orbit coupling

We have been ignoring an important problem, the electron does not only feel an electrostatic force due to the proton, because of its motion and intrinsic spin it will also experience a magnetic force. The electron orbiting a nucleus will experience a magnetic field due to the apparent motion of the proton in its frame of reference, thus forming a current. The hamiltonian for this interaction is given by:

$$\Delta\hat{H} = -\boldsymbol{\mu}_S \cdot \mathbf{B} = -\frac{e\mathbf{S}\cdot\mathbf{B}}{mc} \tag{19.3.5}$$

where we took $g = 2$. Mathematically, the magnetic field experienced by the electron moving at velocity $\mathbf{v}$ with respect to the lab frame which measures an electric field $\mathbf{E}$ is given by:

$$\mathbf{B}' \approx \frac{\mathbf{E}\times\mathbf{v}}{c} \tag{19.3.6}$$

where we took the Lorentz factor be approximately unity $\gamma \approx 1$. For the hydrogen atom:

$$U(r) = -\frac{e^2}{r} \implies \mathbf{E}(r) = \frac{1}{e}\frac{dU}{dr}\hat{\mathbf{r}} \tag{19.3.7}$$

so that:

$$\mathbf{B}' = \frac{1}{ec}\frac{1}{r}\frac{dU}{dr}\mathbf{r}\times\mathbf{v} = \frac{1}{ec}\frac{1}{r}\frac{dU}{dr}\mathbf{L} \tag{19.3.8}$$

Consequently the hamiltonian of the electron would then be:

$$\delta\hat{H}_{S-L} = \frac{1}{m_e^2c^2}\frac{1}{r}\frac{dU}{dr}(\mathbf{S}\cdot\mathbf{L}) \tag{19.3.9}$$

so we see that the spin and orbital angular momenta are coupled. The real value is actually half of this, due to relativistic corrections derived by Llewellyn Thomas. Hence we will take:

$$\delta\hat{H}_{S-L} = \frac{1}{2m_e^2c^2}\frac{1}{r}\frac{dU}{dr}(\mathbf{S}\cdot\mathbf{L}) \tag{19.3.10}$$

### Darwin correction

Solving the Dirac equation for the hydrogen atom we get an additional perturbation, known as the Darwin correction:

$$\delta H_D = \frac{\hbar^2}{8m^2c^2}\nabla^2 V \tag{19.3.11}$$

For the Coulomb potential $V = \frac{e^2}{r}$ then:

$$\delta H_D = -\frac{\pi e^2\hbar^2}{2m^2c^2}\delta(r) \tag{19.3.12}$$

Hence the fine structure hamiltonian becomes:

$$\hat{H} = \hat{H}^{(0)} - \frac{(\mathbf{p}^2)^2}{8m^3c^2} + \frac{e^2}{2m^2c^2}\frac{\mathbf{S}\cdot\mathbf{L}}{r^3} + \frac{\pi e^2\hbar^2}{2m^2c^2}\delta(r) \tag{19.3.13}$$

## 19.4 Evaluating the Darwin term

It is important to note that the Darwin term will only act on states that are non-zero at the origin $r = 0$. Looking at the structure of the wave-functions w see that only $l = 0$ states will be affected by the Darwin term. Hence:

$$E_{n00}^{(1),D} = \langle n00 \,|\, \delta H_D \,|\, n00\rangle \tag{19.4.1}$$

$$= \frac{\pi}{2}\frac{e^2\hbar^2}{m^2c^2}\int |\psi_{n00}|^2\delta(r)d^3\mathbf{r} \tag{19.4.2}$$

We now need to evaluate $|\psi_{n00}(0)|^2$, we state the result

$$|\psi_{n00}(0)|^2 = \frac{1}{\pi n^3 a_0^3} \tag{19.4.3}$$

so that

$$E_{n00}^{(1),D} = \frac{e^2\hbar^2}{2m^2c^2}\frac{1}{a_0^3n^3} = \alpha^4mc^2\frac{1}{2n^3},\ l = 0 \tag{19.4.4}$$

So where does the Darwin term come from? It is due to the fact that the electron is not a point particle in reality, but may be regarded as a cloud, a continuous charge distribution with density $\rho(\mathbf{r}') = -e\rho_0(\mathbf{r}')$. The potential energy becomes:

$$U(\mathbf{r}) = \int -e\rho_0(\mathbf{u})V_C(\mathbf{r}+\mathbf{U})d^3\mathbf{u} \tag{19.4.5}$$

where $V_C(r)$ is the electrostatic potential due to the proton. Expanding the potential about :

$$V(\mathbf{r}+\mathbf{u}) \approx V(\mathbf{r}) + \sum_i \frac{\partial V}{\partial x_i}\bigg|_{\mathbf{r}} u_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{\mathbf{r}} u_i u_j \tag{19.4.6}$$

Substituting this into (19.4.5) we find that:

$$U(\mathbf{r}) \approx V(\mathbf{r}) \int \rho(\mathbf{u}) d^3\mathbf{u} + \sum_i \frac{\partial V}{\partial x_i}\bigg|_{\mathbf{r}} \int u_i \rho(\mathbf{u}) d^3\mathbf{r} + \frac{1}{2}\sum_{i,j} \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{\mathbf{r}} \int u_i u_j \rho(\mathbf{u}) d^3\mathbf{u} \tag{19.4.7}$$

The first integral becomes:

$$V(\mathbf{r}) \int \rho(\mathbf{u}) d^3\mathbf{u} = eV(\mathbf{r}) \tag{19.4.8}$$

as would be expected by a point charge. Let us now assume that the charge distribution is spherically symmetric $\rho(\mathbf{u}) = \rho(u)$. Then we have that

$$\int u_i \rho(\mathbf{u}) d^3\mathbf{r} = 0 \tag{19.4.9}$$

Similarly, if $i \neq j$ then:

$$\int u_i u_j \rho(\mathbf{u}) d^3\mathbf{u} = 0 \tag{19.4.10}$$

If $i = j$ then the integral does not vanish, and since $\langle u_i^2 \rangle = \frac{1}{3}\langle u^2 \rangle$ and in this case:

$$\frac{1}{2}\sum_{i,j} \frac{\partial^2 V}{\partial x_i \partial x_j}\bigg|_{\mathbf{r}} \int u_i u_j \rho(\mathbf{u}) d^3\mathbf{u} = \frac{1}{6}\nabla^2 V \int u^2 \rho(\mathbf{u}) d^3\mathbf{u} \tag{19.4.11}$$

If the electron has radius $r_0 = \frac{\hbar}{mc}$ then we may set:

$$\rho(u) = \begin{cases} \frac{3e}{4\pi r_0^3}, & u < r_0 \\ 0, & u > r_0 \end{cases} \tag{19.4.12}$$

giving:

$$\frac{1}{6}\nabla^2 V \int u^2 \rho(\mathbf{u}) d^3\mathbf{u} = \frac{\hbar^2}{10 m^2 c^2} \tag{19.4.13}$$

and thus:

$$U(\mathbf{r}) = eV(\mathbf{r}) + \frac{\hbar^2}{10 m^2 c^2}\nabla^2 V \tag{19.4.14}$$

which is surprisingly close to the Darwin corrected potential:

$$U(\mathbf{r}) = eV(\mathbf{r}) + \underbrace{\frac{\hbar^2}{8 m^2 c^2}\nabla^2 V}_{\delta H_D} \tag{19.4.15}$$

## 19.5 Evaluating the relativistic correction

The first energy correction should be (in the uncoupled basis):

$$E^{(1),rel}_{nlm_lm_s} = -\frac{1}{8m^3c^2}\left\langle nlm_lm_s \left| \mathbf{p}^4 \right| nlm_lm_s \right\rangle \tag{19.5.1}$$

We can use the expression for non-degenerate perturbation theory because the relativistic perturbation is already diagonal in the uncoupled basis. To see why, note that $\mathbf{p}^2$ is rotationally invariant:

$$[\mathbf{p}^2, \mathbf{L}^2] = 0 \tag{19.5.2}$$
$$[\mathbf{p}^2, \hat{L}_z] = 0 \tag{19.5.3}$$
$$[\mathbf{p}^2, \hat{S}_z] = 0 \tag{19.5.4}$$

so $\mathbf{p}^2$ will be diagonal in the $(l, m_l, m_s)$ eigenbasis. The uncoupled basis is therefore a "good basis". Another way to do so is to use the fact that good basis vectors are non-degenerate eigenstates of some other hermitian operator $\hat{A}$ (in our case $\mathbf{L}^2, \hat{L}_z, \hat{S}_z$), and if $[\hat{A}, \delta\hat{H}] = 0$, as explained in section 14.4.

Now that we have justified (19.5.1) we may evaluate the correction (we omit the redundant $m_s$ label since the correction does not depend on it):

$$E^{(1),rel}_{nlm_l} = -\frac{1}{8m^3c^2}\left\langle \mathbf{p}^2\psi_{nlm_l} \left| \mathbf{p}^2\psi_{nlm_l} \right\rangle \right. \tag{19.5.5}$$
$$= -\frac{1}{8m^3c^2}\left\langle 2m(E_n - V(r))\psi_{nlm_l} \,|\, 2m(E_n - V(r))\psi_{nlm_l} \right\rangle \tag{19.5.6}$$
$$= -\frac{1}{2mc^2}\left(E_n^2 \left\langle nlm_l \,|\, nlm_l \right\rangle - 2E_n \left\langle V(r) \right\rangle_{nlm_l} + \left\langle V^2(r) \right\rangle_{nlm_l}\right) \tag{19.5.7}$$
$$= -\frac{1}{2mc^2}\left(E_n^2 + 2E_ne^2\left\langle \frac{1}{r} \right\rangle_{nlm_l} + e^4\left\langle \frac{1}{r^2} \right\rangle_{nlm_l}\right) \tag{19.5.8}$$

We have found the expectation values of $\frac{1}{r}$ and $\frac{1}{r^2}$ in the previous chapter:

$$\left\langle \frac{1}{r} \right\rangle = \frac{1}{a_0n^2}, \left\langle \frac{1}{r^2} \right\rangle = \frac{1}{a_0^2n^3(l+\frac{1}{2})} \tag{19.5.9}$$

giving:

$$E^{(1),rel}_{nlm_l} = -\frac{1}{2mc^2}\left(E_n^2 + \frac{2E_ne^2}{a_0n^2} + \frac{e^4}{a_0^2n^3(l+\frac{1}{2})}\right) \tag{19.5.10}$$
$$= -\frac{1}{2mc^2}\left(E_n^2 - 4E_n^2 + \frac{4nE_n}{l+\frac{1}{2}}\right) \tag{19.5.11}$$
$$= \frac{E_n^2}{2mc^2}\left(3 - \frac{4n}{l+\frac{1}{2}}\right) \tag{19.5.12}$$
$$= \frac{\alpha^4mc^2}{8n^4}\left(3 - \frac{4n}{l+\frac{1}{2}}\right) \tag{19.5.13}$$

Notice that the energy correction is independent of $m_l, m_s$. Consequently had we used the coupled basis we would have reached the same result.

## 19.6 Spin-Orbit coupling

We have that:

$$\delta H_{S-L} = \frac{e^2}{2m^2c^2}\frac{1}{r^3}\mathbf{S}\cdot\mathbf{L} = \frac{e^2}{4m^2c^2}\frac{1}{r^3}(\mathbf{J}^2 - \mathbf{S}^2 - \mathbf{L}^2) \tag{19.6.1}$$

so it is clear that we must work in the coupled basis. The first order energy correction is:

$$E^{(1),S-L}_{n,l,j,m_j} = \frac{e^2}{2m^2c^2}\frac{1}{r^3}\left\langle nljm_j \left| \mathbf{S}\cdot\mathbf{L} \right| nljm_j \right\rangle \tag{19.6.2}$$

The reason we can use the result from non-degenerate perturbation theory is because, as in the relativistic correction, the perturbation is already diagonalized in the $n, l, j, m_j$ basis. Indeed, we have that:

$$\left[\frac{\mathbf{S}\cdot\mathbf{L}}{r^3}, \mathbf{L}^2\right] = \left[\frac{\mathbf{S}\cdot\mathbf{L}}{r^3}, \mathbf{J}^2\right] = \left[\frac{\mathbf{S}\cdot\mathbf{L}}{r^3}, \hat{J}_z\right] = 0 \tag{19.6.3}$$

Therefore we find that:

$$E^{(1),S-L}_{n,l,j,m_j} = \frac{e^2\hbar^2}{4m^2c^2}\left(j(j+1) - l(l+1) - \frac{3}{4}\right)\left\langle\frac{1}{r^3}\right\rangle_{nljm_j} \tag{19.6.4}$$

and since

$$\left\langle\frac{1}{r^3}\right\rangle_{nljm_j} = \frac{1}{n^3a_0^3 l(l+1)(l+\frac{1}{2})} \tag{19.6.5}$$

we find that:

$$E^{(1),S-L}_{n,l,j,m_j} = \frac{e^2\hbar^2}{4m^2c^2}\frac{\left(j(j+1) - l(l+1) - \frac{3}{4}\right)}{n^3a_0^3 l(l+1)(l+\frac{1}{2})} \tag{19.6.6}$$

$$= \frac{e^4 a_0}{4mc^2}\frac{\left(j(j+1) - l(l+1) - \frac{3}{4}\right)}{n^3a_0^3 l(l+1)(l+\frac{1}{2})} \tag{19.6.7}$$

$$= \frac{4n}{4mc^2}\frac{e^4}{a_0^2n^4}\frac{\left(j(j+1) - l(l+1) - \frac{3}{4}\right)}{l(l+1)(l+\frac{1}{2})} \tag{19.6.8}$$

$$= \frac{E_n^2}{mc^2}\frac{n\left(j(j+1) - l(l+1) - \frac{3}{4}\right)}{l(l+1)(l+\frac{1}{2})} \tag{19.6.9}$$

## 19.7 Fine structure splitting

We see that as $l \to 0$ the spin-orbit energy correction is quite ambiguous. For such states we may take $j = l + \frac{1}{2}$, as $j = l - \frac{1}{2}$ would yield a negative value of $j$, and write:

$$\frac{E_n^2}{mc^2}\frac{n((l+\frac{1}{2})(l+\frac{3}{2}) - l(l+1) - \frac{3}{4})}{l(l+1)(l+\frac{1}{2})} \tag{19.7.1}$$

$$= \frac{E_n^2}{mc^2}\frac{nl}{l(l+1)(l+\frac{1}{2})} \tag{19.7.2}$$

$$= \frac{E_n^2}{mc^2}\frac{n}{(l+1)(l+\frac{1}{2})} \tag{19.7.3}$$

$$= \frac{2nE_n^2}{mc^2} = E_{n00}^{(1),D} \tag{19.7.4}$$

So we see that although physically the spin-orbit coupling disappears for $l = 0$ states, it gets replaced by the Darwin correction which is only relevant when $l = 0$. Consequently we can just sum the relativistic sum and the spin orbit terms (which is the same as the Darwin term):

$$E_{nljm_j}^{(1)} = \frac{(E_n^{(0)})^2}{2mc^2}\left\{3 - \frac{4n}{l+\frac{1}{2}} + 2n\left[\frac{j(j+1) - 3l(l+1) - \frac{3}{4}}{l(l+\frac{1}{2})(l+1)}\right]\right\} \tag{19.7.5}$$

$$= \frac{(E_n^{(0)})^2}{2mc^2}\left\{3 + 2n\left[\frac{j(j+1) - 3l(l+1) - \frac{3}{4}}{l(l+\frac{1}{2})(l+1)}\right]\right\} \tag{19.7.6}$$

Recall that $l \otimes \frac{1}{2} = (l+\frac{1}{2}) \oplus (l-\frac{1}{2})$ so for a fixed value of $j$ we will have only two possible values of $l$, either $l = j + \frac{1}{2}$ or $l = j - \frac{1}{2}$.

Setting $l = j + \frac{1}{2}$ we find:

$$E_{nljm_j}^{(1)} = \frac{(E_n^{(0)})^2}{2mc^2}\left\{3 + 2n\left[\frac{(j(j+1) - 3(j+\frac{1}{2})(j+\frac{3}{2}) - \frac{3}{4}}{(j+\frac{1}{2})(j+1)(j+\frac{3}{2})}\right]\right\} \tag{19.7.7}$$

$$= \frac{(E_n^{(0)})^2}{2mc^2}\left\{3 + 2n\left[\frac{-2j^2 - 5j - 3}{(j+\frac{1}{2})(j+1)(j+\frac{3}{2})}\right]\right\} \tag{19.7.8}$$

$$= \frac{(E_n^{(0)})^2}{2mc^2}\left\{3 + 2n\left[\frac{-2(j+1)(j+\frac{3}{2})}{(j+\frac{1}{2})(j+1)(j+\frac{3}{2})}\right]\right\} \tag{19.7.9}$$

$$= \frac{(E_n^{(0)})^2}{2mc^2}\left[3 - \frac{4n}{j+\frac{1}{2}}\right] \tag{19.7.10}$$

and similarly setting $l = j - \frac{1}{2}$ we find:

$$E^{(1)}_{nljm_j} = \frac{(E^{(0)}_n)^2}{2mc^2}\left\{3 + 2n\left[\frac{(j(j+1) - 3(j-\frac{1}{2})(j+\frac{1}{2}) - \frac{3}{4}}{j(j-\frac{1}{2})(j+\frac{1}{2})}\right]\right\} \tag{19.7.11}$$

$$= \frac{(E^{(0)}_n)^2}{2mc^2}\left\{3 + 2n\left[\frac{-2j^2 + j}{j(j-\frac{1}{2})(j+\frac{1}{2})}\right]\right\} \tag{19.7.12}$$

$$= \frac{(E^{(0)}_n)^2}{2mc^2}\left[3 - \frac{4n}{j+\frac{1}{2}}\right] \tag{19.7.13}$$

Therefore the energy corrections to first order have no real dependence on $l$, since given $j$ the two possible values of $l$ give the same result. Thus the total fine structure splitting of energy levels reads:

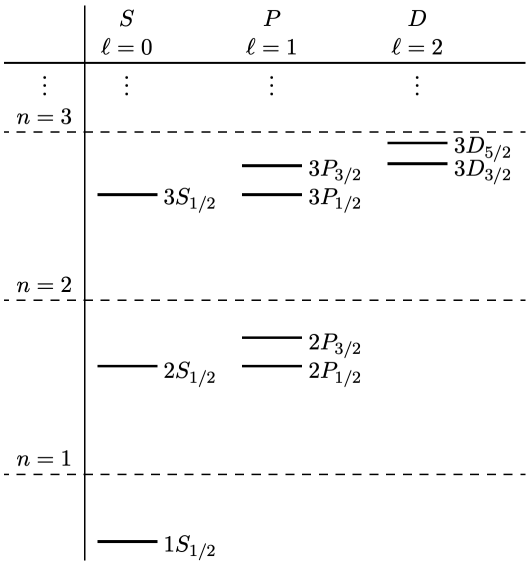
$$E^{(1)}_{nj} = \frac{(E^{(0)}_n)^2}{2mc^2}\left[3 - \frac{4n}{j+\frac{1}{2}}\right] = \alpha^4 mc^2 \frac{1}{2n^4}\left[\frac{3}{4} - \frac{n}{j+\frac{1}{2}}\right] \tag{19.7.14}$$

Note that this correction is always negative. Indeed the largest possible energy correction occurs when $\frac{n}{j+\frac{1}{2}}$ is minimized:

$$\min_{n,j} \frac{n}{j+\frac{1}{2}} - \frac{3}{4} = \min_{n,j} \frac{n}{l+1} - \frac{3}{4} = 1 - \frac{3}{4} = \frac{1}{4} \tag{19.7.15}$$

since $l_{max} = n - 1$. Consequently the lowest the energy correction can get is $-\frac{\alpha^4 mc^2}{8n^4}$. All energy levels are thus shifted down by some amount.

The hydrogen spectrum now looks like:




We have broken part of the degeneracy of $|n, l, j, m_j\rangle$ states. The energy is no longer dependent solely on $n$ (gross structure), but is also dependent on $j$, splitting all states with

$j \neq 0$ into two depending on the $j$ subscript in the spectroscopic notation ${}^nL_j$. However the multiplets are not split due to a lack of $m_j$ dependence. Hence states with the same $j$ are degenerate, which is why ${}^2S_{\frac{1}{2}}$ and ${}^2P_{\frac{1}{2}}$ are degenerate.

## 19.8 The Zeeman effect

Recall that an electron has two components of magnetic moment, an orbital component:

$$\boldsymbol{\mu}_l = -\frac{e}{2mc}\mathbf{L} \tag{19.8.1}$$

and an intrinsic spin component:

$$\boldsymbol{\mu}_l = -\frac{e}{mc}\mathbf{S} \tag{19.8.2}$$

Hence when placed in a magnetic field the Hamiltonian of the electron will read:

$$\delta\hat{H}_Z = \frac{e}{2mc}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} \tag{19.8.3}$$

If we orient our coordinate system so that $\mathbf{B} = B\hat{\mathbf{z}}$ then:

$$\delta\hat{H}_Z = \frac{eB}{2mc}(\hat{L}_z + 2\hat{S}_z) \tag{19.8.4}$$

The total hamiltonian is then:

$$\hat{H} = \hat{H}^{(0)} + \delta\hat{H}_{FS} + \delta\hat{H}_Z \tag{19.8.5}$$

For weak magnetic fields we may treat $\hat{H}^{(0)} + \delta\hat{H}_{FS}$ as our new $\hat{H}^{(0)}$ and $\delta\hat{H}_Z$ as our perturbation. But what do we mean by weak magnetic field? It means that the applied field $B$ must be smaller than the internal magnetic field $B_{in}$ which results in the spin-orbit coupling, the magnetic field experienced by the electron due to its relative motion around the proton.

## 19.9 Weak Zeeman effect

We will have to consider the full matrix

$$\left\langle n, l, j, m_j \middle| \delta\hat{H}_Z \middle| n, l', j, m_j' \right\rangle \tag{19.9.1}$$

owing to the fact that two degenerate states can have different $l$ or different $m_j'$ since they are not included in the fine structure energy splitting. Luckily, we have that:

$$[\mathbf{L}^2, \delta\hat{H}_Z] = \frac{eB}{2mc}([\mathbf{L}^2, \hat{L}_z] + 2[\mathbf{L}^2, \hat{S}_z]) = 0 \tag{19.9.2}$$

so the perturbation is already diagonal in $l$, we only need to worry about $m_j$:

$$\left\langle n, l, j, m_j \middle| \delta\hat{H}_Z \middle| n, l, j, m_j' \right\rangle \tag{19.9.3}$$

By a stroke of luck we also have that:

$$[\hat{J}_z, \delta\hat{H}_Z] = \frac{eB}{2mc}([\hat{J}_z, \hat{L}_z] + 2[\hat{J}_z, \hat{S}_z]) = 0 \tag{19.9.4}$$

so the perturbation is already diagonal in the coupled basis $|n, l, j, m_j\rangle$! We can therefore use the result from non-degenerate perturbation theory and simply compute the diagonal elements of $\delta\hat{H}_Z$.

We have:

$$E^{(1)}_{nljm_j} = \frac{eB}{2mc}\left\langle nljm_j \middle| \hat{L}_z + 2\hat{S}_z \middle| nljm_j \right\rangle \tag{19.9.5}$$

$$= \frac{eB}{2mc}\left\langle nljm_j \middle| \hat{J}_z + \hat{S}_z \middle| nljm_j \right\rangle \tag{19.9.6}$$

$$= \frac{eB}{2mc}(\hbar m_j + \left\langle nljm_j \middle| \hat{S}_z \middle| nljm_j \right\rangle) \tag{19.9.7}$$

Evaluating $\hat{S}_z$ is a bit tricky, the most immediate method is to use the Wigner-Eckart theorem. However here we present an alternative method.

### Vector operators and the Projection lemma

Recall that:

$$[\hat{J}_i, \hat{S}_i] = i\hbar\epsilon_{ijk}\hat{S}_k \tag{19.9.8}$$

by the standard angular momentum algebra. More generally, operators $\mathbf{V}$ whose components satisfy the commutator algebra:

$$[\hat{J}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}\hat{S}_k \tag{19.9.9}$$

are called vector operators. It is easy to show that $\mathbf{J}, \mathbf{L}, \mathbf{S}$ are all vector operators. Perhaps more surprisingly, we also have that $\mathbf{x}$ is a vector operator:

$$[\hat{J}_i, \hat{x}_j] = \epsilon_{ilk}[\hat{x}_l\hat{p}_k, \hat{x}_j] \qquad = \epsilon_{ilk}\hat{x}_l(-i\hbar\delta_{kj}) \tag{19.9.10}$$

$$= -i\hbar\epsilon_{ilj}\hat{x}_l = i\hbar\epsilon_{ijl}\hat{x}_l \tag{19.9.11}$$

as desired.

Furthermore, if $\mathbf{V}$ and $\mathbf{W}$ are vector operators then $\mathbf{V} \times \mathbf{W}$ is also a vector operator. Indeed:

$$[\hat{J}_i, (\mathbf{V} \times \mathbf{W})_j] = \epsilon_{jkl}[\hat{J}_i, \hat{V}_k\hat{W}_l] \tag{19.9.12}$$

$$= \epsilon_{jkl}([\hat{J}_i, \hat{V}_k]\hat{W}_l + \hat{V}_k[\hat{J}_i, \hat{W}_l]) \tag{19.9.13}$$

$$= i\hbar\epsilon_{jkl}(\epsilon_{ikm}\hat{V}_m\hat{W}_l + \hat{V}_k\epsilon_{iln}\hat{W}_n) \tag{19.9.14}$$

and since:

$$\epsilon_{jkl}\epsilon_{ikm}\hat{V}_m\hat{W}_l = \epsilon_{kjl}\epsilon_{kim}\hat{V}_m\hat{W}_l \tag{19.9.15}$$

$$= (\delta_{ij}\delta_{lm} - \delta_{jm}\delta_{il})\hat{V}_m\hat{W}_l \tag{19.9.16}$$

$$= \delta_{ij}\hat{V}_l\hat{W}_l - \hat{V}_j\hat{W}_i \tag{19.9.17}$$

and similarly:

$$\epsilon_{jkl}\epsilon_{iln}\hat{V}_k\hat{W}_n = \epsilon_{lkj}\epsilon_{lin}\hat{V}_k\hat{W}_n \tag{19.9.18}$$
$$= (\delta_{ik}\delta_{jn} - \delta_{ij}\delta_{kn})\hat{V}_k\hat{W}_n \tag{19.9.19}$$
$$= \hat{V}_i\hat{W}_j - \delta_{ij}\hat{V}_k\hat{W}_k \tag{19.9.20}$$

implying that

$$[\hat{J}_i, (\mathbf{V}\times\mathbf{W})_j] = i\hbar(\hat{V}_i\hat{W}_j - \hat{V}_j\hat{W}_i) \tag{19.9.21}$$
$$= i\hbar\epsilon_{ijk}(\mathbf{V}\times\mathbf{W})_k \tag{19.9.22}$$

Indeed:

$$\epsilon_{ijk}(\mathbf{V}\times\mathbf{W})_k = \epsilon_{ijk}\epsilon_{klm}\hat{V}_l\hat{W}_m \tag{19.9.23}$$
$$= -\epsilon_{kij}\epsilon_{klm}\hat{V}_l\hat{W}_m \tag{19.9.24}$$
$$= -(\delta_{jl}\delta_{im} - \delta_{jm}\delta_{il})\hat{V}_l\hat{W}_m \tag{19.9.25}$$
$$= \hat{V}_i\hat{W}_j - \hat{V}_j\hat{W}_i \tag{19.9.26}$$

as desired. Hence we have found that $\mathbf{V}\times\mathbf{W}$ is a vector operator with:

$$\boxed{[\hat{J}_i, (\mathbf{V}\times\mathbf{W})_j] = i\hbar\epsilon_{ijk}(\mathbf{V}\times\mathbf{W})_k} \tag{19.9.27}$$

Another important property is that:

$$([\mathbf{J}^2, \mathbf{W}])_j = [\hat{J}_i\hat{J}_i, \hat{W}_j] \tag{19.9.28}$$
$$= \hat{J}_i[\hat{J}_i, \hat{W}_j] + [\hat{J}_i, \hat{W}_j]\hat{J}_i \tag{19.9.29}$$
$$= i\hbar(\epsilon_{ijk}\hat{J}_i\hat{W}_k + \epsilon_{ijl}\hat{W}_l\hat{J}_i) \tag{19.9.30}$$
$$= i\hbar(\epsilon_{jki}\hat{W}_k\hat{J}_i + \epsilon_{jki}\hat{W}_k\hat{J}_i) + i\hbar\epsilon_{jki}[\hat{J}_i, \hat{W}_k] \tag{19.9.31}$$
$$= 2i\hbar(\mathbf{W}\times\mathbf{J})_j + i\hbar(i\hbar\epsilon_{ikl}\hat{W}_l)\epsilon_{jki} \tag{19.9.32}$$
$$= 2i\hbar(\mathbf{W}\times\mathbf{J})_j - i\hbar(i\hbar\epsilon_{ikj}\epsilon_{ikl}\hat{W}_l) \tag{19.9.33}$$
$$= 2i\hbar(\mathbf{W}\times\mathbf{J})_j - i\hbar(2i\hbar\delta_{jl}\hat{W}_l) \tag{19.9.34}$$
$$= 2i\hbar((\mathbf{W}\times\mathbf{J})_j - i\hbar\hat{W}_j) \tag{19.9.35}$$
$$\tag{19.9.36}$$

implying that:

$$\boxed{[\mathbf{J}^2, \mathbf{W}] = 2i\hbar(\mathbf{W}\times\mathbf{J} - i\hbar\mathbf{W})} \tag{19.9.37}$$

If we set $\mathbf{W} = \mathbf{J}$ then we get:

$$[\mathbf{J}^2, \mathbf{J}] = 2i\hbar(\mathbf{J}\times\mathbf{J} - i\hbar\mathbf{J}) = 0 \tag{19.9.38}$$

as expected.

Finally, we reach the important result of this (rather long) subsection:

$$[\mathbf{J}^2,[\mathbf{J}^2,\mathbf{V}]] = (2i\hbar)^2\left[(\mathbf{V}\cdot\mathbf{J})\mathbf{J} - \frac{1}{2}(\mathbf{J}^2\mathbf{V}+\mathbf{V}\mathbf{J}^2)\right] \tag{19.9.39}$$

To prove this:

$$\begin{aligned}
[\mathbf{J}^2,[\mathbf{J}^2,\mathbf{V}]] &= 2i\hbar[\mathbf{J}^2,\mathbf{W}\times\mathbf{J} - i\hbar\mathbf{W}] && (19.9.40)\\
&= 2i\hbar([\mathbf{J}^2,\mathbf{V}\times\mathbf{J}] - i\hbar[\mathbf{J}^2,\mathbf{V}]) && (19.9.41)\\
&= (2i\hbar)^2\{(\mathbf{V}\times\mathbf{J})\times\mathbf{J} - i\hbar(\mathbf{V}\times\mathbf{J}) - \frac{1}{2}[\mathbf{J}^2,\mathbf{V}]\} && (19.9.42)
\end{aligned}$$

Now we also have that:

$$\begin{aligned}
((\mathbf{V}\times\mathbf{J})\times\mathbf{J})_i &= \epsilon_{ijk}\epsilon_{jlm}\hat{V}_l\hat{J}_m\hat{J}_k && (19.9.43)\\
&= \epsilon_{jki}\epsilon_{jlm}\hat{V}_l\hat{J}_m\hat{J}_k && (19.9.44)\\
& && (19.9.45)\\
&= (\delta_{kl}\delta_{im} - \delta_{km}\delta_{il})\hat{V}_l\hat{J}_m\hat{J}_k && (19.9.46)\\
&= \hat{V}_k\hat{J}_i\hat{J}_k - \hat{V}_i\hat{J}_k\hat{J}_k && (19.9.47)
\end{aligned}$$

and since

$$\begin{aligned}
\hat{V}_k\hat{J}_i\hat{J}_k &= \hat{J}_i\hat{V}_k\hat{J}_k - [\hat{J}_i\hat{V}_k]\hat{J}_k && (19.9.48)\\
&= \hat{J}_i\mathbf{V}\cdot\mathbf{J} - i\hbar\epsilon_{ikl}\hat{V}_l\hat{J}_k && (19.9.49)\\
&= \hat{J}_i\mathbf{V}\cdot\mathbf{J} + i\hbar\epsilon_{ilk}\hat{V}_l\hat{J}_k && (19.9.50)\\
&= \hat{J}_i\mathbf{V}\cdot\mathbf{J} + i\hbar(\mathbf{V}\times\mathbf{J})_i && (19.9.51)\\
& && (19.9.52)
\end{aligned}$$

we get that

$$((\mathbf{V}\times\mathbf{J})\times\mathbf{J})_i = \hat{J}_i\mathbf{V}\cdot\mathbf{J} + i\hbar(\mathbf{V}\times\mathbf{J})_i - \hat{V}_i\mathbf{J}^2 \tag{19.9.53}$$

or more simply

$$(\mathbf{V}\times\mathbf{J})\times\mathbf{J} = \mathbf{J}\mathbf{V}\cdot\mathbf{J} + i\hbar(\mathbf{V}\times\mathbf{J}) - \mathbf{V}\mathbf{J}^2 \tag{19.9.54}$$

Then

$$[\mathbf{J}^2,[\mathbf{J}^2,\mathbf{V}]] = (2i\hbar)^2\{\mathbf{J}\mathbf{V}\cdot\mathbf{J} - \mathbf{V}\mathbf{J}^2 - \frac{1}{2}[\mathbf{J}^2,\mathbf{V}]\} \tag{19.9.55}$$

Note that:

$$\frac{1}{2}(\mathbf{J}^2\mathbf{V}+\mathbf{V}\mathbf{J}^2) = \mathbf{V}\mathbf{J}^2 + \frac{1}{2}[\mathbf{J}^2,\mathbf{V}] \tag{19.9.56}$$

giving:

$$[\mathbf{J}^2,[\mathbf{J}^2,\mathbf{V}]] = (2i\hbar)^2\left(\mathbf{V}\cdot\mathbf{J}\mathbf{J} - \frac{1}{2}(\mathbf{J}^2\mathbf{V}+\mathbf{V}\mathbf{J}^2)\right) \tag{19.9.57}$$

as desired (we used the fact that $\mathbf{J}$ commutes with $\mathbf{V}\cdot\mathbf{J}$ as can be easily verified).

Now we may take the expectation value of $|nljm_j\rangle$ of the above identity and use Ehrenfest's

theorem to get:

$$0 = \langle \mathbf{JV} \cdot \mathbf{J} \rangle - \frac{1}{2}(\left\langle \mathbf{J}^2\mathbf{V} \right\rangle + \left\langle \mathbf{V}\mathbf{J}^2 \right\rangle) \tag{19.9.58}$$

$$= \langle \mathbf{JV} \cdot \mathbf{J} \rangle - \hbar^2 j(j+1) \langle \mathbf{V} \rangle \tag{19.9.59}$$

giving us the projection lemma:

$$\langle k; jm_j \,|\, \mathbf{V} \,|\, k; jm_j \rangle = \frac{\langle k; jm_j \,|\, (\mathbf{V} \cdot \mathbf{J})\mathbf{J} \,|\, k; jm_j \rangle}{\hbar^2 j(j+1)} \tag{19.9.60}$$

The matrix elements of $\mathbf{V}$ in the total angular momentum basis is equal to the matrix elements of the typical projection of $\mathbf{V}$ onto This is reminiscent of the typical projection identity of normal vectors:

$$\text{Proj}_{\mathbf{b}}(\mathbf{a}) = \frac{(\mathbf{a} \cdot \mathbf{b})\mathbf{b}}{\mathbf{b}^2} \tag{19.9.61}$$

Applying this projection lemma to $\mathbf{S}$, and taking only the $z$-component we find the matrix elements of $\hat{S}_z$ in the coupled basis:

$$\left\langle nljm_j \,\middle|\, \hat{S}_z \,\middle|\, nljm_j \right\rangle = \frac{\left\langle k; jm_j \,\middle|\, (\mathbf{S} \cdot \mathbf{J})\hat{J}_z \,\middle|\, k; jm_j \right\rangle}{\hbar^2 j(j+1)} = \frac{\langle k; jm_j \,|\, (\mathbf{S} \cdot \mathbf{J}) \,|\, k; jm_j \rangle \, m_j}{\hbar j(j+1)} \tag{19.9.62}$$

Using the identity $\mathbf{S} \cdot \mathbf{J} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 + \mathbf{S}^2)$ then:

$$\left\langle nljm_j \,\middle|\, \hat{S}_z \,\middle|\, nljm_j \right\rangle = \frac{\hbar m_j \left( j(j+1) - l(l+1) + \frac{3}{4} \right)}{2j(j+1)} \tag{19.9.63}$$

The first order energy corrections due to the Zeeman perturbation will then be:

$$E^{(1)}_{nljm_j} = \frac{e\hbar B}{2mc} m_j \left( 1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \right) \tag{19.9.64}$$

The proportionality term in parenthesis is known as the **Lande g-factor** denoted $g(j,l)$:

$$g(j,l) = 1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \tag{19.9.65}$$

Unlike the fine structure splitting which was implicitly independent on $l$, in this case it can be shown that $g(l + \frac{1}{2}, l) \neq g(l - \frac{1}{2}, l)$, so that the $l$-dependence remains. Therefore, the energy correction is now dependent on $m_j, j, l$ so all degeneracies have been removed!

# Many electron atoms

20

## 20.1 The atomic hamiltonian

Until now we have only been considering systems with a single electron, and a nucleus consisting of $Z$ protons. In most cases however we have more than a single electron, consider for example a nucleus of $Z$ protons surrounded by $Z$ electrons. Then:

$$\hat{H} = \sum_{j=1}^{Z} \underbrace{\left( -\frac{\hbar^2}{2m}\nabla_j^2 - \frac{Ze^2}{r_j} \right)}_{\text{hydrogen-like hamiltonian}} + \underbrace{\frac{1}{2}\sum_{j\neq k}^{Z} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|}}_{\text{electron-electron interaction}} \tag{20.1.1}$$

where $\mathbf{r}_i$ is the vector displacement from the proton (at the origin) to the $i$th electron. The first term represents the typical hamiltonian which includes proton-electron interactions only, the second term also includes the electron-electron interactions, with the $\frac{1}{2}$ term to account for double counting. Due to the indistinguishability of electrons we will need to seek solutions of the form:

$$\psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_Z)\chi \tag{20.1.2}$$

which are totally antisymmetric under particle exchange. Although an exact solution to this problem is not yet known, we can apply perturbation theory and our knowledge of hydrogen-like systems to find approximate solutions.

## 20.2 Helium gross structure (perturbation)

For helium the hamiltonian reads:

$$\hat{H} = \left( -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{2e^2}{r_1} \right) + \left( -\frac{\hbar^2}{2m}\nabla_2^2 - \frac{2e^2}{r_2} \right) + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{20.2.1}$$

We will treat $\frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$ as a perturbation. Firstly we must find the solutions to

$$\left[ \left( -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{2e^2}{r_1} \right) + \left( -\frac{\hbar^2}{2m}\nabla_2^2 - \frac{2e^2}{r_2} \right) \right] \psi = E\psi \tag{20.2.2}$$

This is a separable Hamiltonian, corresponding to the case where the subsystems proton-electron 1 and proton-electron 2 do not interact with each other. We know that the solu-

tions will be of the form:

$$\Psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2)=\psi_{nlm}(\mathbf{r}_1)\psi_{n'l'm'}(\mathbf{r}_2)=R_{nl}(r_1)R_{n'l'}(r_2)Y_l^m(\theta_1,\phi_1)Y_{l'}^{m'}(\theta_2,\phi_2) \tag{20.2.3}$$

while the unperturbed energy levels are:

$$E^{(0)}_{n,n'}=-\left(\frac{1}{n^2}+\frac{1}{n'^2}\right)\text{Ry}=-4\left(\frac{1}{n^2}+\frac{1}{n'^2}\right)13.6\text{ eV}=-4\left(\frac{1}{n^2}+\frac{1}{n'^2}\right)\frac{e^2}{2a_{0,H}} \tag{20.2.4}$$

where $a_{0,H}$ is the hydrogen bohr radius. For example, the ground state will have energy :

$$E^{(0)}_{GS}=\frac{4e^2}{a_0}=108.8\text{ eV} \tag{20.2.5}$$

and the corresponding spatial wave-function will be:

$$\psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)=\frac{8}{\pi a_0^3}e^{-2(r_1+r_2)/a_0} \tag{20.2.6}$$

Note that the spatial wave-function is symmetric [1], so we require the spin component to be anti-symmetric. Only the singlet state is anti-symmetric so:

$$\Psi^{(0)}_{\text{GS}}=\frac{8}{\sqrt{2}\pi a_0^3}e^{-2(r_1+r_2)/a_0}(|\uparrow,\downarrow\rangle-|\downarrow,\uparrow\rangle) \tag{20.2.7}$$

Now let us evaluate the first order energy correction to the ground state. We need to calculate:

$$E^{(1)}_{\text{GS}}=e^2\,\langle\Psi^{(0)}_{GS}|\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}|\Psi^{(0)}_{GS}\rangle \tag{20.2.8}$$

$$=\left(\frac{8e}{\sqrt{2}\pi a_0^3}\right)^2\int\frac{e^{-4(r_1+r_2)/a_0}}{|\mathbf{r}_1-\mathbf{r}_2|}d^3\mathbf{r}_2d^3\mathbf{r}_1 \tag{20.2.9}$$

Now we may align our axes so that $\mathbf{r}_1$ lies on the $z$-axis, so that:

$$|\mathbf{r}_1-\mathbf{r}_2|=\sqrt{r_1^2+r_2^2-2r_1r_2\cos\theta_2} \tag{20.2.10}$$

and choose to integrate over $d^3\mathbf{r}_2$ first:

$$I=\int_0^{2\pi}\int_0^{\infty}\int_0^{\pi}\frac{e^{-4r_2/a_0}}{\sqrt{r_1^2+r_2^2-2r_1r_2\cos\theta_2}}r_2^2\sin\theta_2d\theta_2dr_2d\phi_2 \tag{20.2.11}$$

$$=2\pi\int_0^{\infty}\left[\frac{r_2}{r_1}e^{-4r_2/a_0}\sqrt{r_1^2+r_2^2-2r_1r_2\cos\theta_2}\right]_0^{\pi}dr_2 \tag{20.2.12}$$

$$=2\pi\int_0^{\infty}\frac{r_2}{r_1}e^{-4r_2/a_0}(|r_1+r_2|-|r_1-r_2|)dr_2 \tag{20.2.13}$$

[1] technically we should write $\frac{1}{\sqrt{2}}(\psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)+\psi_{100}(\mathbf{r}_2)\psi_{100}(\mathbf{r}_1))$ since the antisymmetric choice would vanish. However normalization gives $\frac{8}{\pi a_0^3}e^{-2(r_1+r_2)/a_0}$

Now in the range where $r_1 > r_2$ we get that $|r_1 + r_2| - |r_1 - r_2| = 2r_2$ and thus we get a contribution:

$$2\int_0^{r_1} \frac{r_2}{r_1} e^{-4r_2/a_0}(|r_1 + r_2| - |r_1 - r_2|)dr_2 \tag{20.2.14}$$
$$= \frac{4}{r_1}\int_0^{r_1} r_2^2 e^{-4r_2/a_0} dr_2 \tag{20.2.15}$$
$$= \frac{4}{r_1}\left[2\left(\frac{a_0}{4}\right)^3 - \frac{a_0}{4}\left(2\frac{a_0^2}{16} + 2\frac{a_0}{4}r_1 + r_1^2\right)e^{-4r_1/a_0}\right] \tag{20.2.16}$$
$$= \frac{4}{r_1}\left[\frac{a_0^2}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{2}r_1 + r_1^2\right)e^{-4r_1/a_0}\right] \tag{20.2.17}$$

In the range where $r_1 < r_2$ we get that $|r_1 + r_2| - |r_1 - r_2| = 2r_1$ and thus we get a contribution:

$$2\int_{r_1}^{\infty} \frac{r_2}{r_1} e^{-4r_2/a_0}(|r_1 + r_2| - |r_1 - r_2|)dr_2 \tag{20.2.18}$$
$$= 4\int_{r_1}^{\infty} r_2 e^{-4r_2/a_0} dr_2 \tag{20.2.19}$$
$$= 4\frac{a_0}{4}\left(\frac{a_0}{4} + r_1\right)e^{-4r_1/a_0} \tag{20.2.20}$$
$$= 4\frac{a_0}{4}\left(\frac{a_0}{4} + r_1\right)e^{-4r_1/a_0} \tag{20.2.21}$$

Taking their sum we get that:

$$\frac{I}{2\pi} = \frac{4}{r_1}\left[\frac{a_0^3}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{2}r_1 + r_1^2\right)e^{-4r_1/a_0}\right] + 4\frac{a_0}{4}\left(\frac{a_0}{4} + r_1\right)e^{-4r_1/a_0} \tag{20.2.22}$$
$$= \frac{4}{r_1}\left[\frac{a_0^3}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{2}r_1 + r_1^2 - \frac{a_0}{4}r_1 - r_1^2\right)e^{-4r_1/a_0}\right] \tag{20.2.23}$$
$$= \frac{4}{r_1}\left[\frac{a_0^3}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{4}r_1\right)e^{-4r_1/a_0}\right] \tag{20.2.24}$$

We now perform the $d^3\mathbf{r}_1$ integral:

$$\int \frac{8\pi}{r_1}\left[\frac{a_0^3}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{4}r_1\right)e^{-4r_1/a_0}\right]e^{-4r_1/a_0} r_1^2 \sin\theta_1 dr_1 d\theta_1 d\phi_1 \tag{20.2.25}$$
$$= 32\pi^2\int_0^{\infty} r_1 e^{-4r_1/a_0}\left[\frac{a_0^3}{32} - \frac{a_0}{4}\left(\frac{a_0^2}{8} + \frac{a_0}{4}r_1\right)e^{-4r_1/a_0}\right]dr_1 \tag{20.2.26}$$
$$= \pi^2\int_0^{\infty} r_1 e^{-4r_1/a_0}\left[a_0^3 - a_0\left(a_0^2 + 2a_0 r_1\right)e^{-4r_1/a_0}\right]dr_1 \tag{20.2.27}$$
$$= \pi^2\frac{5a_0^5}{128} \tag{20.2.28}$$

Consequently the first order energy correction is:

$$E_{\text{GS}}^{(1)} - \frac{32e^2}{\pi^2 a_0^6}\pi^2 \frac{5a_0^5}{128} = \frac{5e^2}{4a_0} \approx -\frac{5}{16}E_{\text{GS}}^{(0)} \tag{20.2.29}$$

The ground state of helium is approximately:

$$E_{\text{GS}} \approx \frac{11}{16}E_{\text{GS}}^{(0)} = -74.9 \text{ eV} \tag{20.2.30}$$

## 20.3 Helium excited states

The excited states of helium consist of one excited electron and an electron in the ground state. Indeed if both electrons are excited then one of them will transition to the ground state, releasing enough energy to ionize the atom and release the other electron.

Ignoring electron-electron repulsion, we may then construct a state with symmetric spatial wavefunction:

$$|\Psi\rangle_{para} = \frac{1}{\sqrt{2}}(|100\rangle \otimes |nlm\rangle + |nlm\rangle \otimes |100\rangle) \otimes |S = 0, S_z = 0\rangle \tag{20.3.1}$$

known as **parahelium**, and 3 states with anti-symmetric spatial wavefunction:

$$|\Psi\rangle_{para} = \frac{1}{\sqrt{2}}(|100\rangle \otimes |nlm\rangle - |nlm\rangle \otimes |100\rangle) \otimes |S = 1, S_z\rangle \tag{20.3.2}$$

known as **orthohelium**. We use upper-case $S$ and $S_z$ to denote the total spin and spin projection along $z$ quantum numbers. Let us evaluate the first order energy corrections for these states. We expect that parahelium, having a symmetric wave-function, will bunch up the electrons closer together than orthohelium. The electron repulsion will be stronger for parahelium, resulting in a higher energy.

We get that:

$$E_{nlm}^{(1)} = \frac{e^2}{2}(\langle 100| \otimes \langle nlm| \pm \langle nlm| \otimes \langle 100|)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}(|100\rangle \otimes |nlm\rangle \pm |nlm\rangle \otimes |100\rangle) \tag{20.3.3}$$

$$= e^2(\left\langle 100, nlm\right|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\left|100, nlm\right\rangle \pm \left\langle 100, nlm\right|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\left|nlm, 100\right\rangle) \tag{20.3.4}$$

Now we have that:

$$C_{nlm} = e^2\left\langle 100, nlm\right|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\left|100, nlm\right\rangle = \int \frac{e^2|\psi_{nlm}(\mathbf{r}_1)|^2|\psi_{\text{GS}}(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}d^3\mathbf{r}_1 d^3\mathbf{r}_2 \tag{20.3.5}$$

$$J_{nlm} = e^2\left\langle 100, nlm\right|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\left|100, nlm\right\rangle = \int \frac{e^2\psi_{nlm}(\mathbf{r}_1)\psi_{nlm}(\mathbf{r}_2)\psi_{\text{GS}}(\mathbf{r}_1)\psi_{\text{GS}}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d^3\mathbf{r}_1 d^3\mathbf{r}_2 \tag{20.3.6}$$

The first integral is known as the Coulomb integral. Since $|\psi(\mathbf{r})|^2$ is interpreted as the

probability density of the electron cloud, it follows that $e|\psi(\mathbf{r})|^2$ will be a charge density. Consequently the coulomb integral is just the potential energy associated with the two electron clouds.

The second integral is known as the exchange integral, and it is a consequence of the indistinguishability of electrons. For helium this is always positive, so as expected parahelium lies higher in energy than parahelium.

Thus the approximate energy to first order of an excited helium atom is:

$$E_{nlm} \approx E^{(0)}_{nlm} + C_{nlm} \pm J_{nlm} \tag{20.3.7}$$

where $+$ is used for singlet spin states, while $-$ is used for triplet spin states.

## 20.4 Variational principle

The perturbation theory calculations required to calculate the ground state of helium were quite long and complicated. It turns out that there is a simpler method that, in some instances, can be used to estimate ground state energies to a greater degree of accuracy.

Suppose we have a Hamiltonian $\hat{H}$ acting on a Hilbert space containing some arbitrary normalized state $|\psi\rangle$. Then we will have that

$$E_{\text{GS}} \leq \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle \tag{20.4.1}$$

Indeed:

$$\hat{H}\,|\psi\rangle = \hat{H}\Big(\sum_k c_k\,|E_k\rangle\Big) \tag{20.4.2}$$

$$= \sum_k c_k E_k\,|E_k\rangle \tag{20.4.3}$$

$$\leq \sum_k c_k E_{\text{GS}}\,|E_k\rangle \tag{20.4.4}$$

$$= E_{\text{GS}}\,|\psi\rangle \implies E_{\text{GS}} \leq \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle \tag{20.4.5}$$

as desired.

Therefore, if we choose an arbitrary state the energy expectation value in this state will be larger than or equal to the ground state. One may then get a good approximation to the ground state energy by minimizing this expectation value. The easiest way to do so is to choose a state $|\psi(\alpha)\rangle$ with a free variable $\alpha$ and then minimize $\left\langle \psi(\alpha) \middle| \hat{H} \middle| \psi(\alpha) \right\rangle$.

For example, let's consider the case of the harmonic oscillator:

$$\hat{H} = -\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{20.4.6}$$

whose eigenfunctions satisfy:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\frac{1}{2}m\omega^2x^2\psi=E\psi \tag{20.4.7}$$

Since the potential is symmetric the wavefunctions must be either even or odd. Furthermore the ground state cannot have nodes so it must be even. It is easy to see that $\psi(x,b)=Ae^{-bx^2}$ satisfies all these requirements. Let us firstly work out the normalization constant:

$$|A|^2\int_{-\infty}^{\infty}e^{2bx^2}dx=|A|^2\sqrt{\frac{\pi}{2b}}=1\implies A=\left(\frac{2b}{\pi}\right)^{1/4} \tag{20.4.8}$$

Now we get that:

$$\langle T\rangle=-\frac{\hbar^2}{2m}|A|^2\int_{-\infty}^{\infty}e^{-2bx^2}2b(2bx^2-1)dx=\frac{\hbar^2b}{2m} \tag{20.4.9}$$

while

$$\langle V\rangle=\frac{1}{2}m\omega^2|A|^2\int_{-\infty}^{\infty}e^{-2bx^2}x^2dx=\frac{m\omega^2}{8b} \tag{20.4.10}$$

Consequently we need to minimize:

$$\langle H\rangle=\frac{\hbar^2b}{2m}+\frac{m\omega^2}{8b} \tag{20.4.11}$$

$$\frac{d\langle H\rangle}{db}=\frac{\hbar^2}{2m}-\frac{m\omega^2}{8b^2}\implies b=\frac{m\omega}{2\hbar} \tag{20.4.12}$$

We therefore find that the minimum energy expectation value is $\frac{\hbar\omega}{2}$, which is unsurprisingly the ground state energy of the oscillator. Indeed, the type of wavefunction we chose (gaussian) matches the form of the true ground state wavefunction.

## 20.5 Helium gross structure (variational)

Let's apply the variational method to the Helium atom. Again we need to choose a trial wave-function. In perturbation theory we used the unperturbed ground state:

$$\psi_{\text{GS}}(\mathbf{r}_1,\mathbf{r}_2)=\frac{8}{\pi a_0^3}e^{-2(r_1+r_2)/a_0} \tag{20.5.1}$$

Here the $8$ and $2$ factors appear due to the fact that $Z=2$ in our system. However, what if we let $Z$ be a variable we minimize $\langle H\rangle$ with respect to? Physically, the electrons will experience a shielding effect due to each other, so the overall charge they observe will not be $2e$ but smaller, we define this new charge as "effective nuclear charge" $Z$. Then:

$$\psi_{\text{GS}}(\mathbf{r}_1,\mathbf{r}_2)=\frac{Z^3}{\pi a_0^3}e^{-Z(r_1+r_2)/a_0} \tag{20.5.2}$$

To be able to use the result from previous chapters, we express the Hamiltonian as:

$$\hat{H} = \left( -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{Ze^2}{r_1} \right) + \left( -\frac{\hbar^2}{2m}\nabla_2^2 - \frac{Ze^2}{r_2} \right) + \frac{(Z-2)e^2}{r_1} + \frac{(Z-2)e^2}{r_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (20.5.3)$$

Notice that the terms in the first two parenthesis are just the Hamiltonians for hydrogen-like atoms with atomic number $Z$. Their expectation values are known. So are the expectation values of $\frac{1}{r}$ and $\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$. Hence:

$$\langle H \rangle = -2\frac{Z^2 e^2}{2a_0} + 2e^2(Z-2)\left\langle \frac{1}{r} \right\rangle + \left\langle \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right\rangle \quad (20.5.4)$$

Now in the ground state:

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a_0} \quad (20.5.5)$$

instead:

$$\left\langle \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right\rangle = \frac{5e^2}{4b_0} = \frac{5Ze^2}{8a_0} \quad (20.5.6)$$

Consequently the expectation value is:

$$\langle H \rangle = -2\frac{Z^2 e^2}{2a_0} + 4e^2(Z-2)\frac{Z}{a_0} + \frac{5Ze^2}{8a_0} \quad (20.5.7)$$

which we minimize to:

$$\frac{d\langle H \rangle}{dZ} = -\frac{2Ze^2}{a_0} + \frac{2e^2(Z-1)}{a_0} + \frac{5e^2}{8a_0} = 0 \quad (20.5.8)$$

implying $Z = \frac{27}{16} < 2$ as expected. We then get that $\langle H \rangle = \frac{1}{2}\left(\frac{3}{2}\right)^6 \frac{e^2}{2a_0} \approx -77.5$ eV which is closer to $-78.9$ eV than the result from perturbation theory ($-74.8$ eV).

## 20.6 The central field approximation

It turns out that full shells (constant $n$) have a spherically symmetric geometry, since:

$$\sum_{m,l} |Y_l^m(\theta, \phi)| = \frac{4\pi}{2l+1} \quad (20.6.1)$$

This suggests that an electron in the outer shell (called **valence shell**) of an atom will experience a spherically symmetric charge distribution. From the POV of a valence electron, one would see the charge of the proton, $Ze$, screened by the other $(Z-1)e$ electrons which form a spherical charge distribution. Since the electric field outside such a configuration is the same as that of a point charge $e$ at the origin, it follows that we may approximate the potential of a valence electron as:

$$V(r) \approx \frac{-e^2}{r} \quad (20.6.2)$$

Consider instead an electron in the innermost electron. Since all electrons are outside and form an approximate spherical distribution, we have by Gauss' law that it will experience a potential solely due to the nucleus:

$$V(r) \approx \frac{-Ze^2}{r} \tag{20.6.3}$$

This suggests that we model the hamiltonian of the $i$th electron in an atom as an effective spherically symmetric hamiltonian

$$\hat{H}_i = \frac{\hat{p}_i^2}{2m} + V_i(r_i) \tag{20.6.4}$$

where $V \to \frac{-e^2}{r}$ for large $i$ and $V \to \frac{-Ze^2}{r}$ for small $i$.

If we choose a potential $V_i(r_i)$ that varies smoothly between these two limits, then we may solve Schrödinger's equation and find a set of solutions. If we anti-symmetrize these solutions, we find a corresponding electron charge density, which we may use to construct a better potential $V_i(r_i)$. Rinse and repeat, until we reach sufficiently accurate results.

Clearly for states of small $l$, electrons will be more likely to be found closer to the nucleus and hence the screening effect will be lower. This results in lower energy. As we increase $l$ we will get larger and larger screening, and thus higher energy levels. We must therefore expect that for a given shell $n$ the individual subshells $l$ will be ordered following:

$$s(l=0) < p(l=1) < d(l=2) < f(l=3)... \tag{20.6.5}$$

There are of course many exceptions due to the complex nature of the electron-electron interaction.

## 20.7 Spectroscopic notation and the L-S coupling scheme

For multi-electron atoms we define the operators:

In the presence of spin-orbit coupling, recall that $S, L, J, M_J$ are good quantum numbers.

We reintroduce the spectroscopic notation which denotes the set of states $|S, L, J, M_J = -J, ..., J\rangle$ as a term symbol ${}^{2S+1}L_J$ where $L$ is an uppercase letter determined by $L$:

| $L$ | Letter |
|---|---|
| 0 | $S$ |
| 1 | $P$ |
| 2 | $D$ |
| 3 | $F$ |

We easily see that the term symbol ${}^{2S+1}L_J$ has degeneracy $2J+1$.

Let's consider for example the electron configuration for Helium: $He : (1s)^2$. Clearly, since $s_1 = s_2 = \frac{1}{2}$ and $l_1 = l_2 = 0$ we have from the rules of addition of angular momenta that

$L = 0$ and $S = 1, 0$. Now since $S = 0$ is anti-symmetric, it must go with an even $L$ which has a symmetric spatial wavefunction. Alternatively, we may also see that if $S = 1$ then we may have $M_S = 1, 0, -1$. However, $M_S = \pm 1$ would require both electrons to be in the same state, violating the Pauli exclusion principle. Hence we may only have $S = 0, L = 0$ implying that $J = 0$. The spectroscopic notation is then ${}^1S_0$, denoting the singlet state $|J = 0, M_J = 0, S = 0, L = 0\rangle$.

It turns out that all closed subshells are singlets with $J = S = L = 0$.

We may also see this by noting that there is only one way to full a subshell. Consequently, the term symbol representing a full subshell must not be degenerate (singlet), and since the degeneracy of a term symbol is $2J+1$ we get that $J = 0$. This can only occur if $L = S = 0$, since for any other value of $L, S$ we would have multiplets $J = L+S, L+S-1, ..., |L-S|$. Yet another way would be to note that each orbital has 0 spin (singlet), implying that $S = 0$ by addition of angular momenta.

Let's take for example $(2p)^6$. Such a configuration has six electrons with $l = 1$ and $s = \frac{1}{2}$ each, and $m_l = \pm 1$, $m_s = \pm\frac{1}{2}$. It follows that $M_S = 0$ and $M_L = 0$. There is only one state that has only these values of $M_S$ and $M_L$, and that is $L = S = 0$ as desired.

When finding the possible term symbols of electron configurations, we may therefore ignore filled shells as they do not contribute to any angular momenta.

Let us now consider a more complex example, such as $(1s)^1(2s)^1$ which is the first excited state of helium. We have that $l_1 = l_2 = 0$ so that $L = 0$. Also $s_1 = s_2 = \frac{1}{2}$ implying $S = 1, 0$. Now since Pauli's exclusion principle has been automatically satisfied, we do not need to worry about (anti)-symmetry. If $S = 1, L = 0$ then $J = 1$ whereas if $S = L = 0$ then $J = 0$. Hence the possible term symbols are:

$$ {}^3S_1,\ {}^1S_0 \tag{20.7.1}$$

Similarly for $(1s)^1(2p)^1$ we find that $S = 1, 0$ and $L = 1$ so that the possible term symbols are

$$ {}^3P_2,\ {}^3P_1,\ {}^3P_0,\ {}^1P_1 \tag{20.7.2}$$

Let's now consider carbon with electron configuration $(1s)^2(2s)^2(2p)^2$, whose only unfilled shell is $(2p)^2$. We have that $l_1 = l_2 = 1$ and $s_1 = s_2 = \frac{1}{2}$ so that $L = 2, 1, 0$ and $S = 1, 0$. Now if $S = 0$ (anti-symmetric) we need either $L = 2$ or $L = 0$ (symmetric) [2], and thus $J = 2$ or $J = 0$. If instead $S = 1$ (symmetric) then we can only have $L = 1$ (anti-symmetric) and thus $J = 2, 1, 0$. Consequently the possible terms are:

$$ {}^1S_0,\ {}^1D_2,\ {}^3P_0,\ {}^3P_1,\ {}^3P_2 \tag{20.7.3}$$

Now we consider nitrogen with electron configuration $(1s)^2(2s)^3(2p)^3$. It follows from the rules of addition of angular momenta that $S = \frac{3}{2}, \frac{1}{2}$ while $L = 3, 2, 1, 0$. These total spin states have mixed symmetry, so we may not pair up $S - L$ without studying the problem

[2] recall that $|L, M_L\rangle$ has parity $(-1)^L$

further. The fastest method is to list all the possible electron arrangements consistent with the configuration given.

In this case there are 15 such arrangements, giving the pairs of $m_S$ and $m_L$ with degeneracies in the table below.

| $m_L$ | $m_S$ | Degeneracy |
|---|---|---|
| $\pm 2$ | $\pm\frac{1}{2}$ | 4 |
| $\pm 1$ | $\pm\frac{1}{2}$ | 8 |
| $\pm 0$ | $\pm\frac{3}{2}$ | 2 |
| $\pm 0$ | $\pm\frac{1}{2}$ | 6 |

To see why, note that to get $(m_L = \pm 2, m_S = \pm\frac{1}{2})$ we would need the following electron arrangements (where by convention the first orbital has $m_l = 0$, the second $m_l = -1$ and the third $m_l = 1$).

$$m_L = 2: \quad \underset{2p}{\uparrow\_\uparrow\downarrow} \quad \underset{2p}{\downarrow\_\uparrow\downarrow} \qquad m_L = -2: \quad \underset{2p}{\uparrow\uparrow\downarrow\_} \quad \underset{2p}{\downarrow\uparrow\downarrow\_}$$

Instead to get $(m_L = \pm 1, m_S = \pm\frac{1}{2})$:

$$m_L = 1: \quad \underset{2p}{\uparrow\downarrow\_\uparrow} \quad \underset{2p}{\uparrow\downarrow\_\downarrow} \quad \underset{2p}{\uparrow\_\uparrow\downarrow} \quad \underset{2p}{\downarrow\_\uparrow\downarrow}$$

$$m_L = -1: \quad \underset{2p}{\uparrow\downarrow\uparrow\_} \quad \underset{2p}{\uparrow\downarrow\downarrow\_} \quad \underset{2p}{\uparrow\uparrow\downarrow\_} \quad \underset{2p}{\downarrow\uparrow\downarrow\_}$$

To get $(m_L = 0, m_S = \pm\frac{3}{2}$ we need:

$$m_S = \frac{3}{2}: \quad \underset{2p}{\uparrow\uparrow\uparrow} \qquad m_S = -\frac{3}{2}: \quad \underset{2p}{\downarrow\downarrow\downarrow}$$

Finally to get $(m_L = 0, m_S = \pm\frac{1}{2}$ we need:

$$m_S = \frac{1}{2}: \quad \underset{2p}{\downarrow\uparrow\uparrow} \quad \underset{2p}{\uparrow\uparrow\downarrow} \quad \underset{2p}{\uparrow\downarrow\uparrow}$$

$$m_S = -\frac{1}{2}: \quad \underset{2p}{\uparrow\downarrow\downarrow} \quad \underset{2p}{\downarrow\downarrow\uparrow} \quad \underset{2p}{\downarrow\uparrow\downarrow}$$

The sum of degeneracies is 20, which must be the case. Indeed, if there are $t = 2(2l+1)$ possible electrons in a subshell, and we need to insert just $n$ electrons in this subshell, then there are precisely $\frac{t!}{l!(t-l)!}$ ways to do so. In our case we find $\frac{6!}{3!3!} = 20$ as desired. This is often useful to make sure the degeneracy counting was done correctly.

Now it is clear that if $L = 2$, then $S = \frac{3}{2}$ is not possible since $(2, \frac{3}{2})$ does not exist, hence we must have $S = \frac{1}{2}$. This would give the $(m_L, m_S)$ pairs $(2, \pm\frac{1}{2}), (1, \pm\frac{1}{2}), (0, \pm\frac{1}{2}), (-1, \pm\frac{1}{2}), (-2, \pm\frac{1}{2})$, which are 10.

If $L = 1$ then we may have $S = \frac{1}{2}$ by the same logic. We get the pairs $(1, \pm\frac{1}{2}), (0, \pm\frac{1}{2}), (-1, \pm\frac{1}{2})$ which are 6.

Finally if $L = 0$ then we may have either $S = \frac{3}{2}$ or $S = \frac{1}{2}$. If $S = \frac{3}{2}$ we get the pairs $(0, \frac{3}{2}), (0, \frac{1}{2}), (0, -\frac{1}{2}), (0, -\frac{3}{2})$ which are 4. We have exhausted all possible $(m_L, m_S)$ couples, meaning that $S = \frac{1}{2}$ is impossible. Indeed had we started our reasoning with $S = \frac{1}{2}$ we would not have been able to fill all $m_L, m_S$ combinations.

Hence the possible term symbols are:

$$^2D_{\frac{5}{2},\frac{3}{2}},\ ^2P_{\frac{3}{2},\frac{1}{2}},\ ^4S_{\frac{3}{2}} \tag{20.7.4}$$

Finally let's consider oxygen with configuration $(1s)^2(2s)^2(2p)^4$ which has valence shell $(2p)^4$. Now $m_S = 1, 0, -1$ and $m_L = 1, 0, -1$. Consequently we can have $S = 1, 0$ and thus $L = 2, 1, 0$. If $S = 1$ then we may only have $L = 1$ whereas if $S = 0$ then $L = 2, 0$. We therefore have the same term symbols as carbon, that is:

$$^1S_0,\ ^1D_2,\ ^3P_0,\ ^3P_1,\ ^3P_2 \tag{20.7.5}$$

It turns out due to symmetry that a subshell filled by $n$ electrons has the same term symbols as a subshell that is $n$ electrons from being full. This explains why $(2p)^2$ and $(2p)^4$ have equivalent term symbols. Similarly, $(2p)^5$ will have the same term symbols as $(2p)^1$, and $(2p)^6$, being a full subshell, will have term symbol $^1S_0$ as we proved earlier. For a given subshell we only need to work through half-filled states to solve the problem completely.

Intuitively this makes sense, as we may regard a subshell that is $n$ electrons from being full as a full subshell plus $n$ electron "holes". Since a full subshell has no angular momentum the problem is reduced to finding the angular momenta of $n$ electron holes, which is mathematically equivalent to finding the angular momenta of $n$ electrons in the subshell.

## 20.8 Hund's rules

We have seen that most electron configurations can have several values of $L, S, J$ corresponding to different ways to add the spin and orbital angular momenta of the electrons while satisfying the Pauli exclusion principle. Suppose we were interested in the ground state of carbon (which has 5 term symbols), how would we go around finding the correct state?

Hund's rules are a set of (mostly) empirical principles used to determine the energy ordering of different electron arrangements:

(i) The ground state has highest total spin number $S$.

(ii) The ground state has highest total orbital angular momentum number $L$.

(iii) If a subshell is half filled or less, the ground state will have total angular momentum number $J = |L - S|$. Otherwise, the ground state will have largest $J$.

Together with typical addition of angular momenta rules we can now determine the ground states of most elements. Let's do so for the first two rows of the periodic table.

Their physical justification is as follows:

(i) The ground state has highest total spin number $S$.

(ii) The ground state has highest total orbital angular momentum number $L$.

(iii) If a subshell is half filled or less, the ground state will have total angular momentum number $J = |L - S|$. Otherwise, the ground state will have largest $J$.

For carbon, it is easy to see that the term symbols with largest $S = 1$ are ${}^3P_{0,1,2}$, which have the same $L = 1$. Since the subshell is less than half-full we choose the state with $J = |L - S| = 0$, hence $(1s)^2(2s)^2(2p)^2\ {}^2P_{\frac{3}{2}}\,{}^3P_0$.

For oxygen we will have to choose $J = L + S = 2$ rather than $J = 0$ (as in the carbon case), hence $(1s)^2(2s)^2(2p)^2\ {}^3P_2$. For nitrogen instead we immediately have $(1s)^2(2s)^2(2p)^3\ {}^4S_{\frac{3}{2}}$.

Suppose we have not found all the possible term symbols, for example for fluorine which configuration $(1s)^2(2s)^2(2p)^5$. We will clearly have two paired electrons, and one unpaired. We begin by applying Hund's first rule, we must maximize $S$, and hence the allowed values of $m_S$, the following electron arrangements do the job:

$$\underset{2p}{\underline{\uparrow\downarrow}\,\underline{\uparrow}\,\underline{\uparrow\downarrow}} \qquad \underset{2p}{\underline{\uparrow\downarrow}\,\underline{\uparrow\downarrow}\,\underline{\uparrow}} \qquad \underset{2p}{\underline{\uparrow}\,\underline{\uparrow\downarrow}\,\underline{\uparrow\downarrow}} \tag{20.8.1}$$

with $m_S = \frac{1}{2}$ and hence $S = \frac{1}{2}$ (electrons fully aligned along $z$). These give $L = 1$ and hence $J = \frac{3}{2}, \frac{1}{2}$, from which we choose $J = \frac{3}{2}$ by Hund's third rule so the ground state of fluorine will be:

$$(1s)^2(2s)^2(2p)^5\ {}^2P_{\frac{3}{2}} \tag{20.8.2}$$

We have actually killed two birds with a stone, since we may also argue that Boron with configuration $(1s)^2(2s)^2(2p)^1$ will similarly have a ground state:

$$(1s)^2(2s)^2(2p)^5\ {}^2P_{\frac{1}{2}} \tag{20.8.3}$$

where we chose the smallest $J$ rather than the largest, by Hund's third rule.

Consider a more complex example. Dysprosium has ground state electron configuration $[Xe](4f)^10(6s)^2$, where $[Xe]$ denotes the ground state electron configuration of xenon, the closest noble gas. We find that Hund's first two rules imply:

$$\underset{4f}{\underline{\uparrow\downarrow}\,\underline{\uparrow\downarrow}\,\underline{\uparrow\downarrow}\,\underline{\uparrow}\,\underline{\uparrow}\,\underline{\uparrow}\,\underline{\uparrow\downarrow}} \tag{20.8.4}$$

with $m_L = 6 \implies L = 6$ and $m_S = 2 \implies S = 2$. Consequently $J = 8$ by Hund's third

rule, specifying the ground state (where $I$ corresponds to $L = 6$):

$$[\text{Xe}](4f)^{10}(6s)^{2\ 5}I_8 \tag{20.8.5}$$

# Diatomic molecules

21

## 21.1 The Born-Oppenheimer approximation

Molecules are much harder to solve exactly, and only a few cases can be solved analytically. This is because the potential due to more than one proton breaks the spherical symmetry we had, so all our tools from angular momenta have been rendered useless. The Hamiltonian operator for a two electron diatomic molecule, consisting of nuclei $A, B$ with atomic numbers $Z_A, Z_B$ and electrons $1, 2$, may be expressed as:

$$\hat{H} = \frac{\hat{\mathbf{p}}_A^2}{2M_A} + \frac{\hat{\mathbf{p}}_B^2}{2M_B} + \sum_{i=1}^{2}\left(\frac{\hat{\mathbf{p}}_i^2}{2m_e} - \frac{Z_A e^2}{|\mathbf{r}_i - \mathbf{R}_A|} - \frac{Z_B e^2}{|\mathbf{r}_i - \mathbf{R}_B|}\right) + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{Z_A Z_B e^2}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{21.1.1}$$

Transforming to the center of mass frame, with $\mu \approx \frac{M_A M_B}{M_A + M_B}$ neglecting the electron masses, then we may write:

$$\hat{H} = \frac{\hat{\mathbf{p}}_{AB}^2}{2\mu} + \sum_{i=1}^{2}\left(\frac{\hat{\mathbf{p}}_i^2}{2m_e} - \frac{Z_A e^2}{|\mathbf{r}_i - \mathbf{R}_A|} - \frac{Z_B e^2}{|\mathbf{r}_i - \mathbf{R}_B|}\right) + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{Z_A Z_B e^2}{R_{AB}} \tag{21.1.2}$$

where $\hat{\mathbf{p}}_{AB} = -i\hbar\nabla_{AB}$ is the momentum in the CM frame neglecting the overall translational motion, with $\mathbf{R}_{AB} = \mathbf{R}_A - \mathbf{R}_B$ as the relative distance.

Clearly this hamiltonian is not even remotely separable, due to the presence of all the electrostatic interactions. However, there is an approximation resulting from the smallness of $\frac{m_e}{m_p}$, known as the born-Oppenheimer approximation, which can be used to consider the electron and nuclei's behaviour separately.

**The Born-Oppenheimer approximation**

When studying the electrons in a molecule, we may model the nuclei as fixed (with no momenta), whereas to study the nuclei in a molecule, we may assume that a change in their position will lead to an immediate change in the electron's positions.

Hence we recover the electronic Schrödinger equation by removing the first term in (21.1.2), which is the kinetic energy of the nuclei, and the last term, which is just an additive con-

stant:

$$\sum_{i=1}^{2}\left(\frac{\hat{\mathbf{p}}_i^2}{2m_e}-\frac{Z_Ae^2}{|\mathbf{r}_i-\mathbf{R}_A|}-\frac{Z_Be^2}{|\mathbf{r}_i-\mathbf{R}_B|}\right)\psi_{el}(\mathbf{r}_1,\mathbf{r}_2)+\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}=E_{el}\psi_{el}(\mathbf{r}_1,\mathbf{r}_2) \tag{21.1.3}$$

and the nuclear Schrödinger equation

$$\left[\frac{\hat{\mathbf{p}}_{AB}^2}{2\mu}+E_{el}+\frac{Z_AZ_Be^2}{R_{AB}}\right]\psi_{nuc}(\mathbf{R}_{AB})=E\psi_{nuc}(\mathbf{R}_{AB}) \tag{21.1.4}$$

## 21.2 The hydrogen molecule ion

The hydrogen molecule ion $H_2^+$ is composed by two protons with position $\mathbf{R}_a$ and $\mathbf{R}_b$ and one electron at $\mathbf{r}$. The potential energy of this set up may be expressed as:

$$V=e^2\left(\frac{1}{|\mathbf{R}_a-\mathbf{R}_b|}+\frac{1}{|\mathbf{r}-\mathbf{R}_a|}+\frac{1}{|\mathbf{r}-\mathbf{R}_b|}\right) \tag{21.2.1}$$

Hence the electronic Schrödinger equation reads:

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2-\frac{e^2}{r_A}-\frac{e^2}{r_B}\right]\psi_{el}(\mathbf{r})=E_{el}\psi_{El}(\mathbf{r}) \tag{21.2.2}$$

To find the approximate solutions to (21.2.2) we will use the variational method, and choose trial wavefunctions that are linear combinations of atomic orbitals. This method is known as the LCAO method.

Let us consider the case where $R=|\mathbf{R}_a-\mathbf{R}_b|$ is very large. Then, the electronic ground state must consist of the electron being bound to one of the protons in the $1s$ orbital. It can be bound to either proton $A$ or proton $B$ hence:

$$\phi_{100}^A=\frac{1}{\sqrt{\pi a_0^3}}e^{-|\mathbf{r}-\mathbf{R}_A|/a_0} \tag{21.2.3}$$

$$\phi_{100}^B=\frac{1}{\sqrt{\pi a_0^3}}e^{-|\mathbf{r}-\mathbf{R}_B|/a_0} \tag{21.2.4}$$

$$\tag{21.2.5}$$

We should therefore take a superposition of these two 1s hydrogenic orbitals:

$$\psi_{el}=\alpha\phi_{100}^A+\beta\phi_{100}^B \tag{21.2.6}$$

The corresponding energy expectation value is:

$$\langle H\rangle = \frac{\left\langle \alpha\phi^A_{100}+\beta\phi^B_{100}\,\middle|\,\hat{H}\,\middle|\,\alpha\phi^A_{100}+\beta\phi^B_{100}\right\rangle}{\langle \alpha\phi^A_{100}+\beta\phi^B_{100}\,|\,\alpha\phi^A_{100}+\beta\phi^B_{100}\rangle} \tag{21.2.7}$$

$$= \frac{|\alpha|^2\langle E\rangle_A + |\beta|^2\langle E\rangle_B + \alpha^*\beta\left\langle\phi^A_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^B_{100}\right\rangle + \alpha\beta^*\left\langle\phi^B_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^A_{100}\right\rangle}{|\alpha|^2+|\beta|^2+\alpha^*\beta\langle\phi^A_{100}\,|\,\phi^B_{100}\rangle+\alpha\beta^*\langle\phi^B_{100}\,|\,\phi^A_{100}\rangle} \tag{21.2.8}$$

$$= \frac{|\alpha|^2H_{AA}+|\beta|^2H_{BB}+\alpha^*\beta H_{AB}+\alpha\beta^*H_{AB}}{|\alpha|^2+|\beta|^2+S(\alpha^*\beta+\alpha\beta^*)} \tag{21.2.9}$$

$$\tag{21.2.10}$$

where we defined:

$$H_{AA} \equiv \left\langle\phi^A_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^A_{100}\right\rangle = \left\langle\phi^B_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^B_{100}\right\rangle \tag{21.2.11}$$

$$H_{AB} \equiv \left\langle\phi^A_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^B_{100}\right\rangle = \left\langle\phi^B_{100}\,\middle|\,\hat{H}\,\middle|\,\phi^A_{100}\right\rangle \tag{21.2.12}$$

and the **interatomic overlap integral** for obvious reasons as:

$$S \equiv \left\langle\phi^A_{100}\,\middle|\,\phi^B_{100}\right\rangle = \left\langle\phi^B_{100}\,\middle|\,\phi^A_{100}\right\rangle \tag{21.2.13}$$

We further simplify our expressions by setting $\alpha,\beta$ to be real numbers:

$$\langle H\rangle = \frac{\alpha^2H_{AA}+\beta^2H_{AA}+2\alpha\beta H_{AB}}{\alpha^2+\beta^2+2\alpha\beta S} \tag{21.2.14}$$

We minimize $\langle H\rangle$ with respect to $\alpha$ and $\beta$ by simultaneously solving :

$$\begin{cases}\frac{\partial\langle H\rangle}{\partial\alpha}=0\\ \frac{\partial\langle H\rangle}{\partial\beta}=0\end{cases} \tag{21.2.15}$$

Now since $\langle H\rangle = \frac{N(\alpha,\beta)}{D(\alpha,\beta)}$ we have that:

$$\frac{\partial\langle H\rangle}{\partial\alpha} = \frac{\frac{\partial N}{\partial\alpha}D - N\frac{\partial D}{\partial\alpha}}{D^2} = 0 \implies \frac{\partial N}{\partial\alpha} - \langle H\rangle\frac{\partial D}{\partial\alpha} = 0 \tag{21.2.16}$$

and similarly for $\beta$. Applying this to (21.2.14) we find that:

$$2\alpha H_{AA}+2\beta H_{AB}-(2\alpha+2\beta S)\langle H\rangle = 0 \implies \alpha(H_{AA}-\langle H\rangle)+\beta(H_{AB}-S\langle H\rangle) = 0 \tag{21.2.17}$$

and

$$2\beta H_{BB}+2\alpha H_{AB}-(2\beta+2\alpha S)\langle H\rangle = 0 \implies \alpha(H_{AB}-S\langle H\rangle)+\beta(H_{AA}-\langle H\rangle) = 0 \tag{21.2.18}$$

For a non-trivial solution to exist, we need the **secular equation** to be satisfied:

$$\begin{vmatrix} H_{AA} - \langle H \rangle & H_{AB} - S\langle H \rangle \\ H_{AB} - S\langle H \rangle & H_{AA} - \langle H \rangle \end{vmatrix} = 0 \tag{21.2.19}$$

We therefore find that:

$$(H_{AA} - \langle H \rangle)^2 - (H_{AB} - S\langle H \rangle)^2 = 0 \tag{21.2.20}$$

$$\implies H_{AA} - \langle H \rangle = \pm(H_{AB} - S\langle H \rangle) \tag{21.2.21}$$

which has two solutions:

$$\langle H \rangle = \frac{H_{AA} \pm H_{AB}}{1 \pm S} \tag{21.2.22}$$

If we substitute $H_{AA} - \langle H \rangle = \pm(H_{AB} - S\langle H \rangle)$ into (21.2.17) we find that $\alpha = \pm\beta$, so the corresponding eigenfunctions will be:

$$\psi = \alpha(\phi^A_{100}(\mathbf{r}) \pm \phi^B_{100}(\mathbf{r})) \tag{21.2.23}$$

We can normalize this using the interatomic overlap integral:

$$\langle \psi \,|\, \psi \rangle = |\alpha|^2 2(1 \pm S) = 1 \implies \alpha = \frac{1}{\sqrt{2(1 \pm S)}} \tag{21.2.24}$$

so that:

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{2(1 \pm S)}}(\phi^A_{100}(\mathbf{r}) \pm \phi^B_{100}(\mathbf{r})) \tag{21.2.25}$$

The lowest of these energy levels gives an approximate ground state energy and wavefunction:

$$E^{el}_{\text{GS}} \approx \frac{H_{AA} + H_{AB}}{1 + S}, \; \psi_{\text{GS}}(\mathbf{r}) = \frac{1}{\sqrt{2(1+S)}}(\phi^A_{100}(\mathbf{r}) + \phi^B_{100}(\mathbf{r})) \tag{21.2.26}$$

whereas the highest of these energy levels gives the approximate first excited state energy:

$$E^{el}_{exc} \approx \frac{H_{AA} - H_{AB}}{1 - S}, \; \psi_{exc}(\mathbf{r}) = \frac{1}{\sqrt{2(1-S)}}(\phi^A_{100}(\mathbf{r}) - \phi^B_{100}(\mathbf{r})) \tag{21.2.27}$$

We used $el$ subscript to remind ourselves that this only gives the electronic ground state energy.

Now the probability density for the ground state is:

$$|\psi_{\text{GS}}(\mathbf{r})|^2 = \frac{1}{2(1+S)}((\phi^A_{100}(\mathbf{r}))^2 + (\phi^B_{100}(\mathbf{r}))^2 + 2\phi^A_{100}(\mathbf{r})\phi^B_{100}(\mathbf{r})) \tag{21.2.28}$$

The first two terms are just the probability density due to a hydrogen atom centered at different protons. The third term however represents an interaction between the two atomic orbitals, and is only of interest when both give a non-negligible contribution (so in the region between the protons). Since $\phi^A_{100}(\mathbf{r}), \phi^B_{100}(\mathbf{r})$ are both positive this interference will be constructive. There will be an enhanced probability of the electron being found between the two protons. It will help screen the two protons and reduce their electrostatic

repulsion, giving the ground state (lowest) energy. Furthermore, the increased probability between the two nuclei leads to a lower kinetic energy[1]. This is at the basis of chemical bonding, where the minimum energy is lower than the energies of either atomic orbitals.

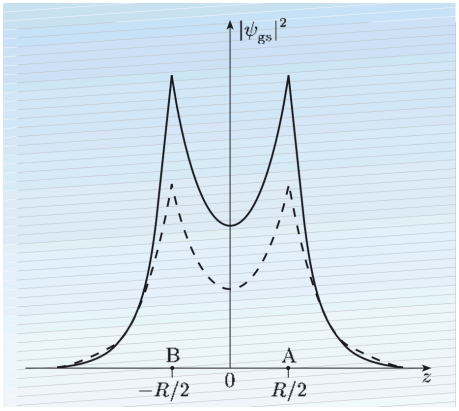



**Figure 21.1**. Probability density of the ground state for the hydrogen molecule ion

Instead, the probability density for the first excited state is:

$$|\psi_{exc}(\mathbf{r})|^2 = \frac{1}{2(1+S)}((\phi^A_{100}(\mathbf{r}))^2 + (\phi^B_{100}(\mathbf{r}))^2 - 2\phi^A_{100}(\mathbf{r})\phi^B_{100}(\mathbf{r})) \tag{21.2.29}$$

The first two terms are again just the probability density due to a hydrogen atom centered at different protons. The third term however represents a destructive interference between the two atomic orbitals. There will be a negligible probability of the electron being found between the two protons. Because the electron no longer screens the two protons, there will be a larger electrostatic repulsion between the two, leading to a higher energy in general. This is at the basis of chemical anti-bonding, where the energy is always higher than the energies of the contributing atomic orbitals.

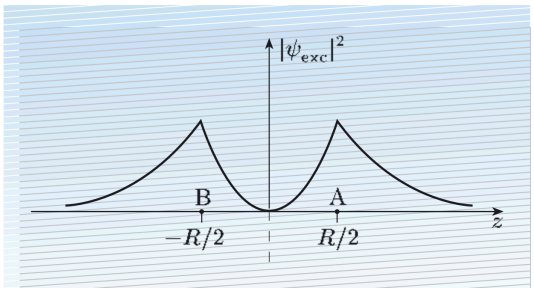



**Figure 21.2**. Probability density of the first excited state for the hydrogen molecule ion

We have forgotten to include the nuclear interaction term $E_{pp} = \frac{e^2}{R}$ to our energy, as well as their kinetic energy. If we only consider the static scenario then we get the total ground state static energy:

$$E_{stat} = \frac{e^2}{R} + E^{el}_{\text{GS}} \tag{21.2.30}$$

[1] classically, we see this as the fact that the electron is more likely to be found in regions of low momentum

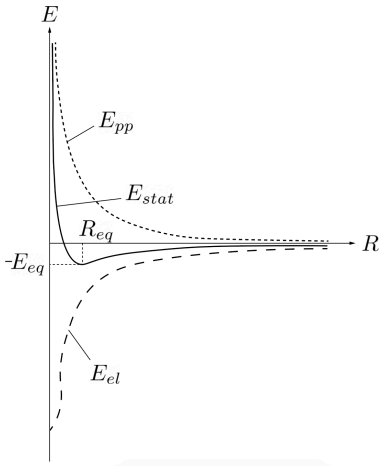



**Figure 21.3**. Energy curve of the static ground state energy

We set the zero energy point at $-$Ry, giving the following energy curves: Clearly at large $R$ the energy asymptotically tends to $-$Ry. Hence for large interatomic distances the hydrogen molecule ion is equivalent to a hydrogen atom in the $1s$ state plus a disassociated proton.

Furthermore, there is a radius $R_{eq} \approx 1.32 \times 10^{-10}$ m, known as the equilibrium separation, at which the static ground state energy is minimized to $E_{eq} \approx 1.76$ eV. The hydrogen molecule ion is therefore most stable in this configuration, the electron has managed to bond the two protons in a stable configuration statically, something that classical electrostatics cannot explain.

Our results coincide at least qualitatively with the exact solution as may be seen in the plots below.

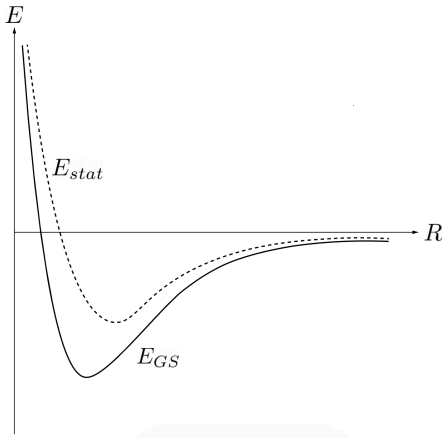



**Figure 21.4**. Comparison between variational method approximation (dotted line) and exact solution (solid line)

As expected the variational method gives an upper estimate to the exact energy. The exact equilibrium separation is $R_{eq} = 1.06 \times 10^{-10}$ m with energy $E_{eq} = 2.79$ eV. To make better

approximations one could for example set an effective nuclear charge parameter.

## 21.3 Molecular orbitals

To label a molecular orbital we use spectroscopic notation with slight adjustments. Since $\hat{H}$ no longer commutes with $\mathbf{L}$, we must choose a new set of quantum numbers. If we align the nuclei of the two protons along the $z$-axis, then the Hamiltonian will be symmetric under rotations about the $z$-axis, so it will commute with $\hat{L}_z$. We may therefore write the eigenfunctions as:

$$\psi_{el}(\mathbf{r}) = u(r,\theta)e^{im\phi} \tag{21.3.1}$$

Furthermore, note that the choice of the direction of the $z$-axis is completely arbitrary, the energy should not change if it points from proton $A$ to proton $B$ or viceversa. Flipping the $z$-axis corresponds to $\phi \to -\phi$ so it follows that $\psi_{el}(\mathbf{r}) = u(r,\theta)e^{im_l\phi}$ and $psi_{el}(\mathbf{r}) = u(r,\theta)e^{-im_l\phi}$ must have the same energy. In other words, states with $m_l$ and $-m_l$ must be degenerate.

Hence $|m_l|$ will be a good quantum number, and we will denote each value of it with a greek letter: $\sigma : |m_l| = 0$, $\pi : |m_l| = 1$, $\delta : |m_l| = 2$ and so forth. Furthermore we may include $g$ (gerade) or $u$ (ungerade) subscripts to denote the even or odd parity of the orbitals with respect to nuclear inversion. Finally, we add a number in front of the greek letters to denote the energy ordering (e.g. $1\sigma_u, 2\sigma_u, ...$).

The hydrogen molecule ion's ground state is therefore labelled by $1\sigma_g$, whereas the first excited state is $1\sigma_u$.

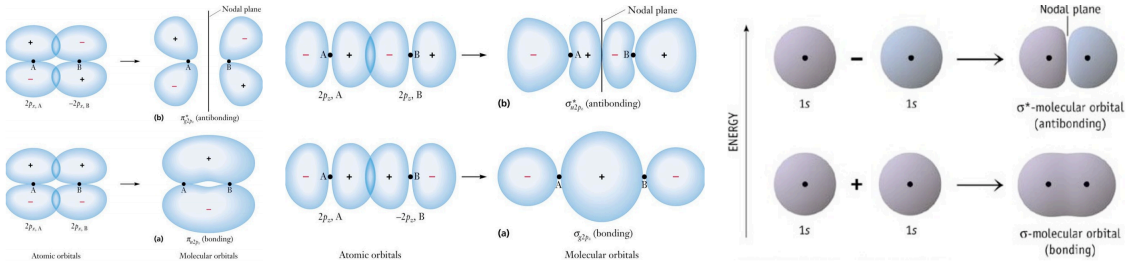



**Figure 21.5**. Labelling molecular orbitals and bonding/anti-bonding orbitals.

Note that a $\sigma$ molecular orbital only means that it is formed from the superposition of $m_l = 0$ orbitals. This imposes no restrictions on possible $l$ values.

The energy ordering of the molecular orbitals for the first two rows of the periodic table is given below.

Let's try for example to enumerate the bonding and anti-bonding orbitals formed by superposing atomic orbitals with $n = 3$.

If $l = 2$ then $|m_l| = 2, 1, 0$. For $m_l = 0$ we have a combination of $\phi^A_{3d_0}$ and $\phi^B_{3d_0}$ in (anti)-symmetric combination. Hence we have two $\sigma$ orbitals. These are $\sigma$ orbitals. For $|m_l| = 1$ we have combinations of $\phi^A_{3d_1}$ with $\phi^B_{3d_1}$ or $\phi^B_{3d_{-1}}$, and similarly $\phi^A_{3d_{-1}}$ with $\phi^B_{3d_1}$ or $\phi^B_{3d_{-1}}$

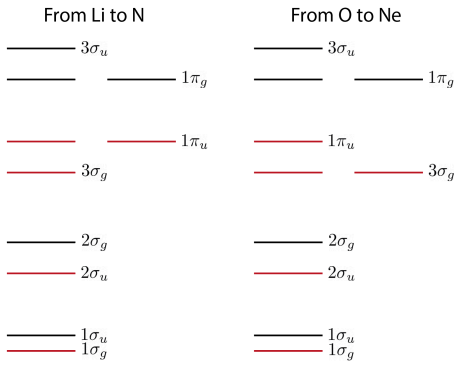



**Figure 21.6**. Energy ordering of molecular orbital energies, with bonding orbitals in red and anti-bonding orbitals in black.

giving four $\pi$ orbitals (two bonding and two anti-bonding) which are two-fold degenerate. Exactly the same reasoning for $|m_l| = 2$ which gives four $\delta$ orbitals.

If $l = 1$ then $|m_l| = 1, 0$. For $m_l = 0$ we have a combination of $\phi^A_{3p_0}$ and $\phi^B_{3p_0}$ in (anti)-symmetric combination. Hence we have two non-degenerate $\sigma$ orbitals, one bonding and one anti-bonding. For $|m_l| = 1$ we have combinations of $\phi^A_{3p_1}$ with $\phi^B_{3p_1}$ or $\phi^B_{3p_{-1}}$, and similarly $\phi^A_{3p_{-1}}$ with $\phi^B_{3p_1}$ or $\phi^B_{3p_{-1}}$ giving four $\pi$ orbitals (two bonding and two anti-bonding) which are two-fold degenerate.

Finally, if $l = 0$ then $|m_l| = 0$ so we have a combination of $\phi^A_{3s_0}$ and $\phi^B_{3s_0}$ in (anti)-symmetric combination. Hence we have two $\sigma$ orbitals, each non-degenerate.

So in total we have $1 + 3 + 5 = 9$ bonding orbitals, and 9 anti-bonding orbitals.

We now define the formal bond order as a measure of how strong the bonding in molecules may be:

$$B.O. = \frac{N_b - N_a}{2} \tag{21.3.2}$$

where $N_b$ is the number of electrons in bonding orbitals, and where $N_a$ is the number of electrons in anti-bonding orbitals. Note that generally within a row of the periodic table, the larger the bond order, the higher the disassociation energy $E_{eq}$ of the molecule. This is because as we increase the number of electrons in bonding orbitals (or decrease the number of electrons in anti-bonding orbitals) the screening between the protons increases, so we will need more energy to disassociate the two atoms.

This will help us see which diatomic molecules are stable. A B.O. of 1 corresponds to a single bond, a B.O. of 2 corresponds to a double bond etc...

Let's consider the hydrogen molecule $H_2$, which consists of two protons and two electrons. Since each molecular orbital can contain at most 2 electrons (except for degenerate orbitals where we may include more electrons) the ground state will be $1\sigma_g^2$. The helium molecule $He_2$ by similar reasoning should have ground state electronic configuration $1\sigma_g^2 1\sigma_u^2$. Note that the formal bond order for $He_2$ is 0, so this molecule should be very unstable. Indeed looking at the energy curve there is a shallow minimum at $E_{eq} = 9 \times 10^{-3}$ eV which is close

in order of magnitude to the lowest vibrational energy state of $He_2$. Detection therefore requires very low temperatures and low frequency light.

Next we have $Li_2$, which has configuration $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2$. The bond order is 1 so we have a single bond.

For the carbon molecule we will have the configuration $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4$. Since we have 8 electrons in bonding orbitals and 4 electrons in anti-bonding orbitals, the formal bond order will be 2, hence there will be a double bond.

Finally, for the nitrogen molecule, note that $1\pi_u$ is two-fold degenerate so it can hold 4 electrons in total. Hence the electronic configuration is $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4 3\sigma_g^2$. We have 10 electrons in bonding orbitals and 4 electrons in anti-bonding orbitals giving a formal bond order of 3. We therefore have a triple bond in the nitrogen molecule.

# Solid State Physics

22

## 22.1 The geometry of crystals

We define a crystal as an solid structure where a group of atoms is arranged in a periodic (hence infinite) pattern over a three-dimensional grid known as a lattice.

Lattice points are points on a lattice for which the surrounding environments look completely identical. In other words, translation which bring one lattice vector to another leave the crystal structure identical. With this in mind we define a lattice vector $\mathbf{R}$ as a vector joining any two lattice points, translations by such vectors are symmetries of the crystal.

We can investigate the structure of wave-functions on crystal lattices using the LCAO approach. Consider a line of six lithium atoms, where each valence electron is in the 2s atomic orbital. Then we may use the trial wave-function:

$$\psi(\mathbf{r}) = \sum_{k=1}^{6} c_k \phi_{2s}^k(\mathbf{r}) \tag{22.1.1}$$

where $\phi_{2s}^k(\mathbf{r})(\mathbf{r})$ is the wavefunction of a 2s orbital centered on the $k$th atom.

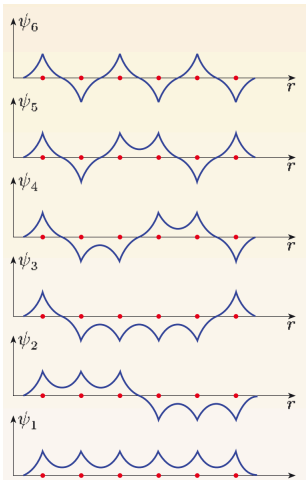
Applying the variational method, we will end up with a secular equation which involves the determinant of a $6 \times 6$ matrix. This will give six energy levels estimating the lowest energy levels of the system, the ground state having no nodes, and the $n$th excited level having $n$ nodes. They are shown below:

Each of these curves is a possible probability density for one electron.

There will be three orbitals $(\psi_1, \psi_2, \psi_3)$ which will have an energy lower than six non-interacting 2s orbitals. This is due to the screening of the Coulomb repulsion by the internuclear electron densities. Instead, the other three orbitals will have an energy higher than the six non-interacting 2s orbitals.

Each molecular orbital can hold two electrons, so the valence electrons in the six lithium atoms will occupy the three lowest energy bonding orbitals in the ground state, causing a chemical bond.

Increasing the number of atoms (say to $10^{23}$) we expect the number of molecular orbitals to increase dramatically. We can regard the resulting energy spectrum as a continuum forming an energy band.




Typically, energy bands for low lying states where the electrons are closely bound to their nuclei have narrow energy bands. Electrons that are farther from the nucleus (larger $l$) are less tightly bound, so there may be overlap between separate atomic orbitals.

The main issue with the LCAO approximation is that it does not accomodate for the translational symmetry crystals. Indeed, looking at the probability densities we see that the red dots corresponding to lattice points are not equivalent. There is a high density between some lattice points, and low density between others.

## 22.2 Bloch's theorem

A crystal has a discrete symmetry of translations by lattice vectors $\mathbf{R_n}$. Consequently the potential acting on an electron in the crystal must obey:

$$V(\mathbf{r}+\mathbf{R_n}) = V(\mathbf{r}) \tag{22.2.1}$$

Let's define $\hat{T}(\mathbf{n})$ as translation operator by $-\mathbf{R_n}$, so that:

$$\hat{T}(\mathbf{n})\psi(\mathbf{r}) = \psi(\mathbf{r}+\mathbf{R_n}) \tag{22.2.2}$$

Now because the Hamiltonian $\hat{H}$ of a crystal must be translationally invariant, we will look at the basis that diagonalizes both $\hat{H}$ and $\hat{T}(\mathbf{n})$.

The eigenfunctions of the translation operator must satisfy:

$$\hat{\mathbf{T}}(\mathbf{n})\psi(\mathbf{r}) = \lambda(\mathbf{n})\psi(\mathbf{r}) \tag{22.2.3}$$

Now note that:

$$\hat{T}(\mathbf{n}_1)\hat{T}(\mathbf{n}_1)\psi(\mathbf{r}) = \psi(\mathbf{r}+\mathbf{R}_{\mathbf{n}_1}+\mathbf{R}_{\mathbf{n}_2}) = \hat{T}(\mathbf{n}_1+\mathbf{n}_2)\psi(\mathbf{r}) \tag{22.2.4}$$

implying that the translation operator is additive:

$$\hat{T}(\mathbf{n}_1)\hat{T}(\mathbf{n}_1) = \hat{T}(\mathbf{n}_1 + \mathbf{n}_2)\psi(\mathbf{r}) \tag{22.2.5}$$

Thus:

$$\lambda(\mathbf{n}_1)\lambda(\mathbf{n}_2) = \lambda(\mathbf{n}_1 + \mathbf{n}_2) \tag{22.2.6}$$

which is only satisfied by exponential eigenvalues $e^{\mathbf{k}\cdot\mathbf{R_n}}$. This solution is only normalizable if $\mathbf{k}$ is complex, that is:

$$\lambda(\mathbf{n}) = e^{i\mathbf{k}\cdot\mathbf{R_n}} \tag{22.2.7}$$

Consequently the eigenvalue equation becomes:

$$\hat{\mathbf{T}}(\mathbf{n})\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R}_n) = e^{i\mathbf{k}\cdot\mathbf{R_n}}\psi(\mathbf{r}) \tag{22.2.8}$$

We can now write our eigenfunctions in the form:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) \tag{22.2.9}$$

which we can always do if we don't make any further assumptions about $u_{\mathbf{k}}(\mathbf{r})$. Then:

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{R_n}) = e^{i\mathbf{k}\cdot\mathbf{R_n}}u_{\mathbf{k}}(\mathbf{r}) \tag{22.2.10}$$

This is known as Bloch's theorem, it states that for electrons in an infinite crystal, the energy eigenfunctions can be expressed in the form of plane waves regulated by some periodic function $u_{\mathbf{k}}(\mathbf{r})$:

$$\psi(\mathbf{r} = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) \tag{22.2.11}$$

such that $u_{\mathbf{k}}(\mathbf{r} + \mathbf{R_n}) = u_{\mathbf{k}}(\mathbf{r})$ for any lattice vector $\mathbf{R_n}$.

Because of the translational symmetry we may impose periodic boundary conditions on our wavefunction:

$$\psi_{\mathbf{k}}(\mathbf{r} + L(\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z) = \psi_{\mathbf{k}}(\mathbf{r}) \tag{22.2.12}$$

This gives the following wave vector quantization:

$$k_x, k_y, k_z = \frac{2m\pi}{L},\ m \in \mathbb{Z} \tag{22.2.13}$$

## 22.3 The Tight-Binding model

In the tight binding model, we take linear combinations of atomic orbitals centered at each atom in the crystal:

$$\psi(\mathbf{r}) = \sum_i c_i\phi(\mathbf{r} - \mathbf{R}_i) \tag{22.3.1}$$

where $\mathbf{R}_i$ is the lattice vector to the lattice point $i$. Instead of using the variational principle, we instead use Bloch's theorem and force our solution to be consistent with it:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \mathcal{N}\sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i}\phi(\mathbf{r} - \mathbf{R}_i) \tag{22.3.2}$$

This ansatz is consistent with Bloch's theorem because for some lattice vector $\mathbf{R}_j$:

$$\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}_j) = \mathcal{N}\sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i}\phi(\mathbf{r}-(\mathbf{R}_i-\mathbf{R}_j)) \tag{22.3.3}$$

$$\mathcal{N}e^{i\mathbf{k}\cdot\mathbf{R}_j}\sum_i e^{i\mathbf{k}\cdot(\mathbf{R}_i+\mathbf{R}_j)}\phi(\mathbf{r}-(\mathbf{R}_i-\mathbf{R}_j)) \tag{22.3.4}$$

Let $\mathbf{R}_i' = \mathbf{R}_i - \mathbf{R}_j$, then this becomes:

$$\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}_j) = \mathcal{N}e^{i\mathbf{k}\cdot\mathbf{R}_j}\sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i'}\phi(\mathbf{r}-\mathbf{R}_i') \tag{22.3.5}$$

This is another sum over all lattice sites $i$ in the crystal, but with a different starting lattice vector. When summing over all the sites however the sums will be equal, and hence:

$$\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}_j) = e^{i\mathbf{k}\cdot\mathbf{R}_j}\psi_{\mathbf{k}}(\mathbf{r}) \tag{22.3.6}$$

as desired. To simplify notation, let $\langle\mathbf{r}\,|\,\phi_i\rangle = \phi(\mathbf{r}-\mathbf{R}_i)$, then:

$$|\psi_{\mathbf{k}}\rangle = \mathcal{N}\sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i}|\phi_i\rangle \tag{22.3.7}$$

Now suppose we treat each electron as subject to a single-particle Hamiltonian:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \sum_j V_j(\mathbf{r}) \tag{22.3.8}$$

Then the energy levels will be given by:

$$E(\mathbf{k}) = \frac{\left\langle\psi_{\mathbf{k}}\middle|\hat{H}\middle|\psi_{\mathbf{k}}\right\rangle}{\langle\psi_{\mathbf{k}}\,|\,\psi_{\mathbf{k}}\rangle} = \frac{|\mathcal{N}|^2\sum_{ij}e^{-i\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)}\left\langle\phi_i\middle|\hat{H}\middle|\phi_j\right\rangle}{\langle\psi_{\mathbf{k}}\,|\,\psi_{\mathbf{k}}\rangle} \tag{22.3.9}$$

Let's now asume that only terms due to nearest neighbor sites contribute to the energy, and ignore all matrix elements due to atomic intervals that are separated further.

We then find that:

$$\left\langle\psi_{\mathbf{k}}\middle|\hat{H}\middle|\psi_{\mathbf{k}}\right\rangle = |\mathcal{N}|^2\sum_i\left[\left\langle\phi_i\middle|\hat{H}\middle|\phi_i\right\rangle + \sum_{j\in\text{nn of }i}e^{-i\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)}\left\langle\phi_i\middle|\hat{H}\middle|\phi_j\right\rangle\right] \tag{22.3.10}$$

and similarly:

$$\langle\psi_{\mathbf{k}}\,|\,\psi_{\mathbf{k}}\rangle = |\mathcal{N}|^2\sum_i\left[1+\sum_{j\in\text{nn of }i}e^{-i\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)}\langle\phi_i\,|\,\phi_j\rangle\right] \tag{22.3.11}$$

Let us define $t_{ij} = \left\langle\phi_i\middle|\hat{H}\middle|\phi_j\right\rangle$ to be the $ij$ hopping matrix element, $S = \langle\phi_i\,|\,\phi_j\rangle$ to be the overlap integral, and let $E_0 = \left\langle\phi_i\middle|\hat{H}\middle|\phi_i\right\rangle$ be the energy of the atomic orbital. Note that due to the equivalence of lattice points, we should have that $t_{ij}$ takes the same value $t$ for

all neighbor sites $\langle ij \rangle$. Then:

$$E(\mathbf{k}) = \frac{E_0 + tf(\mathbf{k})}{1 + Sf(\mathbf{k})} \tag{22.3.12}$$

where

$$f(\mathbf{k}) = \sum_{j=\text{nn of } i} e^{-i\mathbf{k}\cdot(\mathbf{R}_i - \mathbf{R}_j)} \tag{22.3.13}$$

We further simplify our expression by noting that even for overlapping sites $|S| \ll 1$, so using a binomial expansion about $S$: $(1 + Sf(\mathbf{k}))^{-1} = 1 - Sf(\mathbf{k})$ and hence

$$E(\mathbf{k}) \approx E_0 - \beta f(\mathbf{k}),\ \beta = E_0 S - B \tag{22.3.14}$$

For a cubic lattice of length $a$, we find that:

$$f(\mathbf{k}) = \sum_{j=\text{nn of } i} e^{-i\mathbf{k}\cdot(\mathbf{R}_i - \mathbf{R}_j)} \tag{22.3.15}$$

$$= e^{-ik_x a} + e^{ik_x a} + e^{-ik_y a} + e^{ik_y a} + e^{-ik_z a} + e^{ik_z a} \tag{22.3.16}$$

$$= 2(\cos k_x a + \cos k_y a + \cos k_z a) \tag{22.3.17}$$

giving the dispersion relation:

$$E(\mathbf{k}) = E_0 - 2\beta(\cos k_x a + \cos k_y a + \cos k_z a) \tag{22.3.18}$$

Note that the minimum and maximum energy levels are $E_{min} = E_0 - 6\beta$ and $E_{max} = E_0 + 6\beta$ respectively, giving a band width of $12\beta$. Hence, orbitals with high overlap, and hence higher energy, will give wider bands than orbitals that barely overlap.

For atoms that are well separated, the energy bands will be very narrow, since the only possible energy levels will be approximated by those of isolated atoms. As we decrease the separation, the overlap increases, and hence the energy levels broaden to form bands.

Near the bottom of the energy bands, where $k_x, k_y, k_z$ are small, we can use the small angle approximation $\cos k_x a \approx 1 - \frac{1}{2}a^2 k_x^2$ and find that:

$$E(\mathbf{k}) = E_0 - 2\beta\left(3 - \frac{1}{2}a^2|\mathbf{k}|^2\right) = E_0' + \beta a^2 |\mathbf{k}|^2 \tag{22.3.19}$$

These energy levels resemble very much those of an electron in an infinite well of width $a$, where $E(\mathbf{k}) = \frac{\hbar^2|\mathbf{k}|^2}{2m}$. Hence we may identify an effective mass $m_{\text{eff}}$ in (22.3.19):

$$m_{\text{eff}} = \frac{\hbar^2}{2\beta a^2} \tag{22.3.20}$$

We see that in narrow bands, where the electron is tightly bound to the atom, the effective mass will be higher than in wider bands. We see this mathematically in the inverse proportionality $m_{eff} \propto \frac{1}{\beta}$.

## 22.4 Conductor and insulator bands

The theory of energy bands is closely related to the conductivity of materials.

We firstly note that electrons in a full band will not allow for a flow of current at absolute zero.

In the absence of an electric field, this result follows immediately from the fact that a full band will have equal numbers of electrons with momentum $\mathbf{k}$ and $-\mathbf{k}$, giving a net momentum of zero.

If we were to apply an electric field, then the electrons still would not be able to carry a current due to Pauli's exclusion principle. The flow of current implies a net momentum in some direction, implying that the full energy band will have more electrons in one range of momenta. But a full band by definition already has full momentum states, so one cannot do this without placing two or more electrons with the same spin and momentum.

Electrons in a partly full band can instead conduct electricity.

It follows that materials known as conductors will have a partially filled highest energy band, while materials known as insulators have completely full energy bands.

Suppose we now move away from absolute zero by increasing the temperature to room temperature, where the thermal excitations are now in the order of a couple eV.

Conductors will still conduct electricity in the same way. In insulators were the band gap is smaller than the energy of thermal excitations, electrons may be able to hop from the fully occupied lower band, known as **valence band**, to the nearest empty band, known as **conduction band**. Electrical current then flows from electrons in both the valence and conduction bands. Such materials are known as semi-conductors, and have band gaps of less than 2 eV usually.

Instead, electrons in insulators with larger band gaps will not be able to jump between bands, and hence no current will flow.

The Fermi energy, $E_F$, is the energy required to add an electron to a material. It therefore lies somewhere between the valence band and the conduction band for insulators.

## 22.5 Semiconductors

Suppose we have a semi-conductor in which electron $j$ from the valence band jumps from the valence band to the conduction band. The current density in the valence band is therefore:

$$\mathbf{J} = -\frac{e}{V}\sum_{i\neq j}\mathbf{v}_i = -\frac{e}{V}\sum_i \mathbf{v}_i + \frac{e}{V}\mathbf{v}_j = \frac{e}{V}\mathbf{v}_j \tag{22.5.1}$$

which we may view as the current due to a particle of charge $+e$, rather than $-e$. This virtual particle is known as an electron hole.

The energy of an electron hole increases with the depth of the hole in the energy band. This is because removing an electron from a lower energy band requires more energy.

We can increase the conductivity of semiconductors through a process known as doping.

Let's consider silicon, which is in group 14 of the periodic table and therefore has 4 electrons fully occupying the valence band.

Arsenic, on the other hand, has 5 electrons occupying its valence band. Suppose we substitute one silicon atom in a silicon crystal lattice with an arsenic atom. The latter will provide an extra atom which must occupy the conduction band, increasing the conductivity. Atoms which donate an electron to the conduction band are known as donor atoms, and create n-type semiconductors.

Boron, instead, has 3 electrons occupying its valence band. When we substitute a silicon atom with a boron atom, the latter will draw an electron from the valence band of the silicon crystal, creating an electron hole. Atoms which accept an electron from the valence band are known as acceptor atoms, and create p-type semiconductors.

# Particles in EM fields and the Adiabatic theorem

23

## 23.1 Gauge transformations

Recall that the electric and magnetic fields may be defined via the scalar and vector potentials as:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A} \tag{23.1.1}$$

Note that these definitions are invariant under transformations of the type:

$$\phi \mapsto \phi - \frac{1}{c}\frac{\partial \chi}{\partial t} \tag{23.1.2}$$

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\chi \tag{23.1.3}$$

known as Gauge transformations.

The TDSE for a particle of charge $q$ and mass $m$ in a 4-potential $A^\mu = (\psi/c, \mathbf{A})$ looks like:

$$i\hbar\frac{\partial \psi}{\partial t} = \left[\frac{1}{2m}\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)^2 + q\phi\right]\psi \tag{23.1.4}$$

where the TISE can be written explicitly as:

$$\hat{H}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + \frac{q^2}{2mc^2}\mathbf{A}^2\psi + \frac{i\hbar q}{2mc}(2\mathbf{A}\cdot\nabla\psi + \psi\nabla\cdot\mathbf{A}) + \phi\psi \tag{23.1.5}$$

Luckily, the TDSE is gauge invariant. This implies that when a gauge transformation $A^\mu \mapsto A^\mu + \Box\chi \equiv A'^\mu$ is applied to the Hamiltonian $\hat{H} \mapsto \hat{H}'$, the corresponding solution transforms as $\psi \mapsto \alpha\psi \equiv \psi'$ where $\alpha$ is some phase factor that may depend on $\mathbf{r}, t$. Consequently it may not be treated as some simple scalar factor, and most importantly it does not commute with the Hamiltonian in general. Then we find that:

$$i\hbar\frac{\partial \psi'}{\partial t} = i\hbar\frac{\partial \alpha}{\partial t} + \alpha\hat{H}\psi = \hat{H}'(\alpha\psi) \implies \frac{\partial \alpha}{\partial t} = -\frac{i}{\hbar}[\hat{H}'(\alpha\psi) - \alpha\hat{H}\psi] \tag{23.1.6}$$

Our goal is to find $\alpha$, and this can be done by solving the above differential equations. Now

suppose that by some lucky coincidence:

$$\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}'\Big)(\alpha\psi) = \alpha\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)\psi \tag{23.1.7}$$

Then it is quite easy to see that:

$$\hat{H}'(\alpha\psi) = \alpha\frac{1}{2m}\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)^2\psi + q\alpha\phi'\psi \tag{23.1.8}$$

$$\iff \hat{H}'(\alpha\psi) - \alpha\hat{H}\psi = -\frac{q}{c}\alpha\frac{\partial\chi}{\partial t}\psi \tag{23.1.9}$$

and thus:

$$\frac{\partial\alpha}{\partial t} = \frac{iq}{\hbar c}\frac{\partial\chi}{\partial t}\alpha \implies \boxed{\alpha = \exp\Big(\frac{iq\chi}{\hbar c}\Big)} \tag{23.1.10}$$

We must however check that this choice of $\alpha$ satisfies (23.1.7):

$$\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}'\Big)(\alpha\psi) = -i\hbar(\alpha\nabla\psi + \psi\nabla\alpha) - \frac{q}{c}(\mathbf{A} + \nabla\chi)\alpha\psi \tag{23.1.11}$$

$$= \alpha\Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)\psi - i\hbar\psi\nabla\alpha - \frac{q}{c}\alpha\psi\nabla\chi \tag{23.1.12}$$

$$= \Big(\mathbf{p} - \frac{q}{c}\mathbf{A}\Big)\psi \tag{23.1.13}$$

where we used the fact that:

$$\nabla\alpha = \frac{iq}{\hbar c}\alpha\nabla\chi \tag{23.1.14}$$

The gauge invariance of Schrödinger's equation will allows us to marry quantum mechanics with electromagnetism in a beautiful way.

We can also define gauge invariance for operators, an operator $O$ is said to be gauge invariant if for any $|\psi\rangle \in \mathcal{H}$:

$$\Big\langle \psi' \Big| \hat{O}' \Big| \psi' \Big\rangle = \Big\langle \psi \Big| \hat{O} \Big| \psi \Big\rangle \tag{23.1.15}$$

where $|\psi'\rangle = e^{iq\Lambda/\hbar c}|\psi\rangle$ and $\hat{O}' = \mathcal{G}\hat{O}\mathcal{G}$ are the gauge transformed state and operator respectively.

For example, $\hat{\mathbf{p}} \mapsto \hat{\mathbf{p}}$ under a gauge transformation. Consequently:

$$\langle \psi' | \hat{\mathbf{p}} | \psi' \rangle = \Big\langle \psi \Big| e^{-iq\Lambda/\hbar c}\hat{\mathbf{p}}e^{iq\Lambda/\hbar c} \Big| \psi \Big\rangle \neq \langle \psi | \hat{\mathbf{p}} | \psi \rangle \tag{23.1.16}$$

in general since $[e^{-iq\Lambda/\hbar c}, \hat{\mathbf{p}}] \neq 0$. It follows that $\mathbf{p}$ is no longer gauge invariant, it is not a physical quantity when electromagnetic fields are inserted. On the other hand, the generalized momentum $\mathbf{p} - \frac{q}{c}\mathbf{A}$ is gauge invariant since $[e^{-iq\Lambda/\hbar c}, \hat{\mathbf{p}} - q\mathbf{A}]$, so it will yield physical values for observables.

## 23.2 Quantization of magnetic fields on a torus

A torus can be formed by imposing periodic boundary conditions along $x$ and $y$ on a rectangle $\mathcal{R} = [0, L_x] \times [0, L_y]$. Let's embed a constant magnetic field $\mathbf{B}_0$ in the $z$-direction on

this torus.

We must choose a corresponding gauge **A**. We see that:

$$B_z = \partial_x A_y - \partial_y A_x = B_0 \implies A_y = B_0 x,\ A_x = 0 \tag{23.2.1}$$

is a possible gauge, but we must check that it is consistent with the boundary conditions on the torus. Clearly we have that $A_y(x, y + L_y) = A_x(x, y)$, but $A_y(x + L_x, y) \neq A_y(x, y)$. This is fine, since all we need is for $A_y(x + L_x, y)$ to be related by a gauge transformation to $A_y(x, y)$. We need to find this gauge transformation $\mathcal{G}$:

$$\mathcal{G}A_y(x, y) = A_y(x, y) + \partial_y \Lambda = A_y(x + L_x, y) \implies \partial_y \Lambda = B_0 L_x \implies \Lambda = B_0 L_x y \tag{23.2.2}$$

This gauge parameter seems to be ill-defined at $y = 0 = L_y$, it should take two different values. We can solve this issue by noting that $\Lambda$ can be ill-defined as long as $e^{iq\Lambda/\hbar c}$ is well-defined. We require that:

$$e^{iqB_0 L_x y/\hbar c} = e^{iqB_0 L_x (y+L_y)/\hbar c} \implies \frac{qB_0 L_x L_y}{\hbar c} = 2\pi n \tag{23.2.3}$$

so the magnetic field flux should be quantixzed:

$$\Phi = \frac{2\pi\hbar c}{q} n,\ n \in \mathbb{Z} \tag{23.2.4}$$

The flux quantum is $\Phi_0 = \frac{2\pi\hbar c}{q}$ so

$$\Phi = \Phi_0 n,\ n \in \mathbb{Z} \tag{23.2.5}$$

## 23.3 Landau levels

Consider a particle of mass $m$ and charge $q$ in a region of space with magnetic field **B**. In classical mechanics this particle can trace circular trajectories with cyclotron frequency $\omega_c = \frac{qB}{mc}$.

This time we choose the Landau gauge $\mathbf{A} = -By\mathbf{x}$. The Hamiltonian then reads:

$$\hat{H} = \frac{1}{2m}\left[\left(\hat{p}_x + \frac{qB}{c}\hat{y}\right)^2 + p_y^2\right] \tag{23.3.1}$$

where we neglected motion along $z$ which is completely decoupled from the $x-y$ motion. Firstly note that $[H, p_x] = 0$ so we have translational invariance along $x$. This also means that we should seek Bloch state solutions:

$$\psi(x, y) = \psi(y) e^{ik_x x} \tag{23.3.2}$$

Note that we are only allow to do this because of our clever gauge choice. Therefore:

$$\hat{H}\psi(x,y) = \frac{1}{2m}\left[\left(\hat{p}_x + \frac{qB}{c}\hbar k_x\right)^2 + p_y^2\right]\psi(x,y) \tag{23.3.3}$$

$$= \left[\frac{p_y^2}{2m} + \frac{1}{2}m\left(\frac{qB}{mc}\right)^2\left(y + \frac{\hbar k_x c}{qB}\right)\right]\psi(x,y) \tag{23.3.4}$$

$$= \left[\frac{p_y^2}{2m} + \frac{1}{2}m\omega_c^2\Big(y - y_0\Big)\right]\psi(x,y) \tag{23.3.5}$$

where $y_0 = -\frac{\hbar k_x c}{qB}$. This is the equation for a simple harmonic oscillator! The eigenstates are harmonic oscillator solutions travelling as plane waves along $x$:

$$\langle x, y \,|\, k_x, n\rangle = H_n(x)e^{ik_x x} \tag{23.3.6}$$

where $H_n(x)$ are the Harmonic oscillator solutions. The corresponding energy levels are known as Landau levels:

$$E_{k_x,n} = \hbar\omega_c\Big(n + \frac{1}{2}\Big) \tag{23.3.7}$$

Surprisingly the plane waves $e^{ik_x x}$ do not contribute to the total energy, each of these energy levels are infinitely degenerate in $k_x$. The characteristic length of an oscillator is $d = \sqrt{\frac{\hbar}{m\omega}}$ which for (23.3.5) reads:

$$d = \sqrt{\frac{\hbar mc}{qmB}} = \sqrt{\frac{\hbar c}{qB}} \equiv l_B \tag{23.3.8}$$

which we define as the magnetic length. This redefines $y_0 = -k_x l_B^2$, which is the height about which the solutions will oscillate.

Suppose that the region with the magnetic field is now bounded $\mathcal{R} = [0, L_x] \times [0, L_y]$. We take periodic boundary conditions along $x$ so that:

$$k_x = \frac{2\pi n_x}{L_x} \tag{23.3.9}$$

We need $y_0 = -k_x l_B^2 > 0$ for our solutions to lie within $\mathcal{R}$. Consequently $k_x < 0$ so we should only take negative $n_x$. However, $n_x$ cannot be too negative or else $y_0$ will get too large, larger than $L_y$. Thus the allowed $n_x$ values satisfy:

$$-\tilde{n}_x < n_x \leq 0 \tag{23.3.10}$$

where $\tilde{n}_x$ is the degeneracy for each $|n\rangle$ landau level in $\mathcal{R}$. It is easy to see that it satisfies

$$y_0 = \frac{2\pi\tilde{n}_x}{L_x}l_B^2 = L_y \implies \tilde{n}_x = \frac{L_x L_y}{2\pi\frac{\hbar c}{qB}} = \frac{\Phi}{\Phi_0} \tag{23.3.11}$$

Consequently we will have $\frac{\Phi}{\Phi_0}$ allowed $n_x$ values. This quantization is the starting point to understand the Quantum Hall effects, and it turns out that much of the material in the next

sections (the Berry connection in particular) will also be fundamental in understanding it. See my essay on this topic for a list of good references.

## 23.4 Classical Adiabatic invariants

Consider a classical pendulum whose pivot is periodically raised and lowered changing the oscillator length, and thus inducing a time-dependent oscillation frequency $\omega(t)$ in the pendulum. The hamiltonian reads:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega(t)^2 x^2 \tag{23.4.1}$$

The time-dependence of the Hamiltonian is expressed as:

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p}\frac{dp}{dt} + \frac{\partial H}{\partial t} \tag{23.4.2}$$

The Hamiltonian equations of motion are:

$$\frac{\partial H}{\partial p} = \dot{x}, \ \frac{\partial H}{\partial x} = -\dot{p} \tag{23.4.3}$$

and can be substituted into (23.4.2) to cancel out the first two terms, yielding:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = m\omega\frac{d\omega}{dt}x^2 \tag{23.4.4}$$

In an adiabatic change, the time scale $\tau$ of the change in the parameter must be larger than the time scale $T$ of an oscillation (period). More precisely the change in $\omega(t)$ over $T$ is much smaller than $\omega(t)$:

$$\frac{2\pi}{\omega^2}\left|\frac{d\omega}{dt}\right| \ll 1 \iff \boxed{\left|\frac{dT}{dt}\right| \ll 1} \tag{23.4.5}$$

where $T = \frac{2\pi}{\omega(t)}$. Note that this is completely analogous to the condition we require in the WKB approximation that the wavelength vary slowly over position: $|\frac{d\lambda}{dx}| \ll 1$. In the adiabatic approximation we instead impose that the period vary slowly over time.

We claim that the following quantity:

$$\boxed{I(t) = \frac{H(t)}{\omega(t)}} \tag{23.4.6}$$

known as the Adiabatic invariant, changes very slowly. Indeed, computing its time derivative:

$$\frac{dI}{dt} = \frac{\omega\dot{H} - \dot{\omega}H}{\omega^2} = \frac{m\omega^2\dot{\omega}x^2 - p^2\dot{\omega}/2m + 1/2m\omega^2\dot{\omega}x^2}{\omega^2} \tag{23.4.7}$$

$$= \frac{\dot{\omega}}{\omega^2}\left(\frac{1}{2}m\omega^2x^2 - \frac{p^2}{2m}\right) \tag{23.4.8}$$

we can identify the Lagrangian in the above expression:

$$\frac{dI}{dt} = -\frac{\dot{\omega}}{\omega}\mathcal{L} \tag{23.4.9}$$

Note that if we time average $\mathcal{L}$ over a period we will get roughly 0 (exactly 0 if $\omega$ was time independent), while $\frac{\dot{\omega}}{\omega}$ changes very slowly over a period. So, if average over a period:

$$I(t+T) - I(t) = -\int_t^{t+T} \frac{\dot{\omega}}{\omega}\mathcal{L}dt \approx\approx \frac{\dot{\omega}}{\omega}\int_t^{t+T} \mathcal{L}dt = 0 \tag{23.4.10}$$

Geometrically, in the phase space the energy surfaces are ellipses with axes $x_0 = \sqrt{\frac{2E}{m\omega^2}}$ and $y_0 = \sqrt{2mE}$. The area of the ellipse is:

$$\mathcal{A} = \oint pdx = \pi x_0 y_0 = 2\pi\frac{E}{\omega} = 2\pi I(t) \tag{23.4.11}$$

Consequently the adiabatic invariant tells us that the area of the ellipse in phase space varies very slowly. In classical mechanics we generally identify contour integrals of the type:

$$I(t) = \oint pdq \tag{23.4.12}$$

as adiabatic invariants, where $p$ and $q$ are generalized momenta. In quantum mechanics these adoanatic invariants give the Bohr-Sommerfield quantization:

$$\oint pdx = 2\pi\hbar(n + 1/2) \tag{23.4.13}$$

In quantum mechanics the adiabatic invariant can then be written as:

$$\frac{E}{\omega} = \hbar(n + 1/2) \tag{23.4.14}$$

Since this quantity is slowly varying, and $n$ can only be integer, then transitions, which give a fairly large change in $E/\omega$ will be heavily discouraged. States will therefore try not to undergo transitions in an oscillator.

## 23.5 Instantaneous eigenstates

Consider the TISE:

$$H(t)\left|\psi(t)\right\rangle = E(t)\left|\psi(t)\right\rangle \tag{23.5.1}$$

This equations doesn't look time independent as the name suggests, but our goal will be to find its solutions $|\psi(t)\rangle$ at every time $t$ (we are fixing time), known as instantaneous eigenstates. Also consider the TDSE:

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle = \hat{H}(t)\left|\Psi(t)\right\rangle \tag{23.5.2}$$

where $|\Psi(t)\rangle$ are the real solutions, not the instantaneous ones. We introduce ansatz

$$|\Psi(t)\rangle = c(t)\exp\left(-\frac{i}{\hbar}\int_0^t E(t')dt'\right)|\psi(t)\rangle \tag{23.5.3}$$

which is analogous to the ansatz $e^{-iEt/\hbar}|E\rangle$ used for a time-independent hamiltonian. (23.5.3) can be substituted into the TDSE, the LHS giving

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = i\hbar\dot{c}(t)\exp\left(-\frac{i}{\hbar}\int^t E(t')dt'\right)|\psi(t)\rangle + E(t)|\Psi(t)\rangle \tag{23.5.4}$$

$$+ i\hbar c(t)\exp\left(-\frac{i}{\hbar}\int^t E(t')dt'\right)|\dot{\psi}(t)\rangle \tag{23.5.5}$$

and the RHS giving:

$$\hat{H}(t)|\Psi(t)\rangle = E(t)|\Psi(t)\rangle \tag{23.5.6}$$

as a consequence of $|\psi(t)\rangle$ being an instantaneous eigenstate. After some simplification we find that:

$$\dot{c}(t)|\psi(t)\rangle + c(t)|\dot{\psi}(t)\rangle = 0 \tag{23.5.7}$$

Dotting to the left with $\langle\psi(t)|$ we find that:

$$\dot{c}(t) = -c(t)\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle \tag{23.5.8}$$

thus the coefficients $c(t)$ are:

$$c(t) = \exp\left(-\int_0^t \left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle dt'\right) \tag{23.5.9}$$

This is a bit worrying, since the coefficient looks decaying. However, the integrand $\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle$ is actually purely imaginary. Indeed:

$$\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle = \int \psi^*\frac{\partial\psi}{\partial t}d^3\mathbf{r} \tag{23.5.10}$$

$$= \int\left(\frac{\partial|\psi|^2}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right)d^3\mathbf{r} \tag{23.5.11}$$

$$= \frac{\partial}{\partial t}\left(\int|\psi|^2 d^3\mathbf{r}\right) - \left(\int\psi^*\frac{\partial\psi}{\partial t}d\mathbf{r}\right)^* \tag{23.5.12}$$

The first term clearly vanishes due to normalization, so we find that:

$$\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle = -\left(\int\psi^*\frac{\partial\psi}{\partial t}d^3\mathbf{r}\right)^* = -\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle^* \tag{23.5.13}$$

implying that $\left\langle\psi(t)\middle|\dot{\psi}(t)\right\rangle$ is purely imaginary, as required. We can finally write that:

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}\int_0^t E(t')dt'}e^{i\int i\langle\psi|\dot{\psi}\rangle dt'}|\psi(t)\rangle \tag{23.5.14}$$

There is one major mistake that we have glossed over. This result cannot be right because it would imply that a state always remains in its instantaneous eigenstate at all times. We have made no further assumptions about the system, and got an extremely strong (and dubious) result. This is because dotting to the left with just one state $|\psi(t)\rangle$ is not enough to specify $c(t)$ as it only gives one component of a vector equation. Wee should instead dot with all states in the Hilbert space. Luckily, we can still claim that in the adiabatic limit (23.5.14) is still approximately correct. This is known as the **Adiabatic theorem** which we will now prove.

## 23.6 Adiabatic theorem

Consider a set of instantaneous eigenstates $\{|\psi_n(t)\rangle\}$. Let's impose the initial condition that the system start out in one of these eigenstates $|\Psi(0)\rangle = |\psi_k(0)\rangle$. Let us also assume that the energy levels near $k$ are non-degenerate and can thus be ordered in the following way:

$$... \leq E_{k-1}(t) < E_k(t) < E_{k+1}(t) \leq ... \tag{23.6.1}$$

The Adiabatic theorem states that:

$$|\Psi(t)\rangle \approx e^{i\Theta_k(t)} e^{i\gamma_k(t)} |\psi_k(t)\rangle \tag{23.6.2}$$

where the phases are defined as

$$\Theta_k(t) = -\frac{i}{\hbar}\int_0^t E_k(t')dt' \tag{23.6.3}$$

$$\gamma_k(t) = \int_0^t i\left\langle \psi_k(t') \middle| \dot{\psi}_k(t') \right\rangle dt' \tag{23.6.4}$$

provided the Hamiltonian varies slowly.

How slowly? Let us expand $|\Psi(t)\rangle$ in the instantaneous eigenstate basis:

$$|\Psi(t)\rangle = \sum_n c_n(t) |\psi_n(t)\rangle \tag{23.6.5}$$

The TDSE reads:

$$i\hbar \sum_n (\dot{c}_n(t) |\psi_n(t)\rangle + c_n(t) |\dot{\psi}_n(t)\rangle) = \sum_n c_n(t) E_n(t) |\psi_n(t)\rangle \tag{23.6.6}$$

Dotting to the left with $|\psi_k(t)\rangle$, and using the orthonormality of $\{|\psi_n(t)\rangle\}$ then:

$$i\hbar \dot{c}_k(t) = c_k(t)E_k(t) - i\hbar \sum_n \left\langle \psi_k(t) \middle| \dot{\psi}_n(t) \right\rangle c_n(t) \tag{23.6.7}$$

$$\implies i\hbar \dot{c}_k(t) = (E_k(t) - i\hbar \left\langle \psi_k(t) \middle| \dot{\psi}_k(t) \right\rangle) c_k(t) + \sum_{n\neq k} \left\langle \psi_k(t) \middle| \dot{\psi}_n(t) \right\rangle c_n(t) \tag{23.6.8}$$

(this result also applies to other states). When the last term is negligible then we recover

(23.6.2). Note that the initial conditions are:

$$c_m(0) = \delta_{km} \tag{23.6.9}$$

so if the state is slowly varying then the last term should indeed be negligible. To verify this, consider:

$$i\hbar \dot{c}_m(0) = i\hbar \sum_{n \neq m} \left\langle \psi_m \middle| \dot{\psi}_n \right\rangle c_n(0) = i\hbar \left\langle \psi_m \middle| \dot{\psi}_k \right\rangle \neq 0, \ m \neq k \tag{23.6.10}$$

which is worrying, the other instantaneous eigenstates already start getting occupied at $t = 0$. To see how big $\left\langle \psi_m \middle| \dot{\psi}_k \right\rangle$ is consider:

$$H(t) \left| \psi_n(t) \right\rangle = E_n(t) \left| \psi_n(t) \right\rangle \tag{23.6.11}$$

Differentiating with respect to time we get:

$$\dot{H} \left| \psi_n \right\rangle + H \left| \dot{\psi}_n \right\rangle = \dot{E}_n \left| \psi_n \right\rangle + E_n \left| \dot{\psi}_n \right\rangle \tag{23.6.12}$$

Therefore:

$$\left\langle \psi_k \middle| \dot{H} \middle| \psi_n \right\rangle + E_k \left\langle \psi_k \middle| \dot{\psi}_n \right\rangle = E_n \left\langle \psi_k \middle| \dot{\psi}_n \right\rangle \implies \boxed{\left\langle \psi_k \middle| \dot{\psi}_n \right\rangle = \frac{\left\langle \psi_k \middle| \dot{H} \middle| \psi_n \right\rangle}{E_n - E_k}} \tag{23.6.13}$$

For adiabatic changes this term will be small so the coupling between the instantaneous eigenstates will be very slow.

Consider for example:

$$H(t) = \begin{cases} H_0 + \frac{t}{T} V, & 0 < t < T \\ H_0 + V, & t > T \end{cases} \tag{23.6.14}$$

Then we see that $\dot{H} = \frac{V}{T}$ so that:

$$\left\langle \psi_k \middle| \dot{H} \middle| \psi_n \right\rangle = \frac{1}{T} \frac{V_{kn}}{E_n - E_k} \tag{23.6.15}$$

In the adiabatic limit $T$ is very large, so the corrections in (23.6.13) will be very small.

## 23.7 Landau-Zener transitions

We consider a molecule modelled as two fixed nuclei with separation $R$ with an two-level electronic configuration specified by the wavefunctions $\psi_1(x, R)$ (ground state) and $\psi_2(x, R)$ (excited state) with energies $E_1(R)$ and $E_2(R)$:

$$\hat{H}(R)\psi_{1,2}(x, R) = E_{1,2}\psi_{1,2}(x, R) \tag{23.7.1}$$

We are interested in finding a special separation $R = R_0$ where $E_1(R_0) \approx E_2(R_0)$ below which the molecule is non-polar and above which it is polar. We can investigate this behaviour by making $R$ time-dependent, that is we vary the separation between the nu-

clei. The eigenstates of (23.7.1) for each value of $R = R(t')$ then give the instantaneous eigenstates at time $t'$.

Suppose for example that $R(t) = \alpha t$, then we see that at low times we have a non-polar molecule, while at large times we should have a polar molecule. However, could it be that the time-dependence of the system allows for a transition between $E_1(R_0)$ and $E_2(R_0)$?

### Toy example

We consider a two level system $\{|1\rangle, |2\rangle\}$ modelled by the hamiltonian $H(t) = \frac{\alpha t}{2}(|1\rangle\langle 1| - |2\rangle\langle 2|)$ with instantaneous energy levels $E_1 = \frac{\alpha t}{2}$ and $E_2 = -\frac{\alpha t}{2}$.

We see that:

$$|\psi_1(t)\rangle = \exp\left(-\frac{i}{\hbar}\int_0^t E_1(t')dt'\right)|1\rangle = e^{-i\alpha t2/4\hbar}|1\rangle \tag{23.7.2}$$

$$|\psi_2(t)\rangle = \exp\left(-\frac{i}{\hbar}\int_0^t E_2(t')dt'\right)|2\rangle = e^{i\alpha t2/4\hbar}|2\rangle \tag{23.7.3}$$

are both exact solutions of the TDSE. There is no coupling between the states $|1\rangle$ and $|2\rangle$ despite them crossing at $t = 0$.

We now complicate our model a bit by introducing off-diagonal elements:

$$\hat{H} = \begin{pmatrix} \alpha t/2 & H_{12} \\ H_{12}^* & -\alpha t/2 \end{pmatrix} \tag{23.7.4}$$

where $H_{12}$ is small. The energy levels are now:

$$E_\pm = \pm\sqrt{|H_{12}|^2 + \frac{\alpha^2 t^2}{4}} \tag{23.7.5}$$

We are interested in $t = 0$ where the energy levels $E_\pm = \pm H_{12}$ are quite close for small $H_{12}$. Here the system oscillates between $|1\rangle$ and $|2\rangle$ with Rabi frequency $\omega_{12} = \frac{|H_{12}|}{\hbar}$. We define $\tau_d = \frac{|H_{12}|}{\alpha}$ as the characteristic time scale in which the initial energy levels coinciding with $H_{12} = 0$ get deflected. For an adiabatic approximation we require:

$$\omega_{12}\tau_d \gg 1 \implies \frac{|H_{12}|^2}{\alpha\hbar} \tag{23.7.6}$$

## 23.8 Geometric phase

Consider a hamiltonian $H(\mathbf{R})$ where $\mathbf{R} \in \mathbb{R}^N$ is some vector in the configuration space containing the parameters of the system. Suppose these parameters change with time tracing a path $\Gamma(t)$ in $\mathbb{R}^N$ at time $t$. We study the phase in the adiabatic approximation:

$$\gamma_n(t) = i\int_0^t \left\langle \psi_n(R(t')) \middle| \frac{d}{dt'} \middle| \psi_n(R(t')) \right\rangle \tag{23.8.1}$$

known as the **geometric phase**. Using the chain rule the integrand of the phase simplifies to

$$i\left\langle \psi_n(\mathbf{R}(t'))\left|\frac{d}{dt'}\right|\psi_n(\mathbf{R}(t'))\right\rangle = i\left\langle \psi_n(\mathbf{R}(t'))\,|\,\nabla_{\mathbf{R}}\,|\,\psi_n(\mathbf{R}(t'))\right\rangle\cdot\frac{d\mathbf{R}}{dt'} \tag{23.8.2}$$

which leads to the integral in time $\int_0^t dt$ transforming into a path integral in configuration space $\int_\Gamma d\mathbf{R}$:

$$\gamma_n(t) = i\int_{\Gamma(t)}\langle\psi_n(\mathbf{R})\,|\,\nabla_{\mathbf{R}}\,|\,\psi_n(R)\rangle\cdot d\mathbf{R} \tag{23.8.3}$$

We see that unlike $\theta(t)$ which acts like a clock keeping track of $t$, the geometric phase does not really care about time but just the path $\Gamma(t)$ taken by the system in the configuration space. In this sense $\gamma$ is a geometric quantity.

We define the **Berry connection** as the integrand of (23.8.3)

$$A_n(\mathbf{R}) = i\left\langle\psi_n(R)\,|\,\nabla_{\mathbf{R}}\,|\,\psi_n(R)\right\rangle = \int_{\Gamma(t)} A_n(\mathbf{R})\cdot d\mathbf{R} \tag{23.8.4}$$

Note that the Berry connection is not invariant under gauge transformations, but it does transform in a rather special way. Let us define a new instantaneous eigenstate $|\psi_n'(\mathbf{R})\rangle = e^{-i\alpha(\mathbf{R})}\,|\psi_n(\mathbf{R})\rangle$. Then the Berry connection transforms as a vector potential:

$$A_n'(\mathbf{R}) = i\left\langle\psi_n(R)\left|e^{i\alpha(\mathbf{R})}\nabla_{\mathbf{R}}e^{-i\alpha(\mathbf{R})}\right|\psi_n(R)\right\rangle,\ \gamma_n(t) \tag{23.8.5}$$

$$= A_n(\mathbf{R}) + \nabla\alpha(\mathbf{R}) \tag{23.8.6}$$

The geometric phase now becomes:

$$\gamma_n'(t) = \gamma_n(t) + \int_{\Gamma(t)}\nabla_{\mathbf{R}}\alpha(\mathbf{R})\cdot d\mathbf{R} = \gamma_n(t) + \alpha(\mathbf{R}(t)) - \alpha(\mathbf{R}(0)) \tag{23.8.7}$$

We see that unless $\alpha(\mathbf{R}(t)) = \alpha(\mathbf{R}(0))$ the geometric phase is not gauge-invariant, and thus cannot be observed experimentally. However if the motion completes a loop then the geometric phase may become observable since it will be gauge invariant. A well known example of this is the Ahranov-Bohm effect.

It is important to note that if the instantaneous eigenstates can be chosen to be real then the geometric phase vanishes. Similarly, in 1D configuration space (only one parameter changes) then the Berry phase vanishes, since a closed loop in 1D just goes in one direction and back, cancelling out.

Finally, in 3D we can use Stoke's theorem to write:

$$\oint_\Gamma \mathbf{A}\cdot d\mathbf{R} = \iint_{\mathcal{S}_\Gamma}(\nabla\times\mathbf{A})\cdot d\mathbf{a} \tag{23.8.8}$$

where $\nabla\times\mathbf{A}$ is the **Berry curvature**. It is often useful in calculating the geometric phase.

## 23.9 Ahranov-Bohm effect

# Scattering in 3D

24

## 24.1 Initial assumptions

We will be interested in elastic scattering processes where there is no change in the identity of the particles involved in the process, their internal states are left unchanged.

To simplify calculations we will not consider spin degrees of freedom and work non-relativistically. In the next part on Relativistic QM we will consider the relativistic corrections, although a complete theory of scattering processes can only be given in the context of Quantum Field Theory. The interaction potentials between the colliding particles will also be assumed to be central and finite ranged up to some radius $a$.

## 24.2 Guessing the solution

Consider the TISE in the center of mass coordinates, and where $m$ is the reduced mass of the colliding particles:

$$\Big[ -\frac{\hbar^2}{2m}\nabla^2 + V(r)\Big]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{24.2.1}$$

Recall that when discussing scattering in 1D, we inserted plane wave ansatz for the reflected and transmitted waves, and used boundary conditions and the conservation of probability current to determine unknown coefficients. In a similar fashion we now insert as an ansatz the free particle energy $E = \frac{\hbar^2 k^2}{2m}$ to get:

$$\Big[ -\frac{\hbar^2}{2m}(\nabla^2 + k^2) + V(r)\Big]\psi(\mathbf{r}) = 0 \tag{24.2.2}$$

In the case where $V(r) = 0$ we get plane wave solutions of momentum $k = |\mathbf{k}|$. So far away from the potential, where it is zero (i.e. $r \gg a$) we may assume an incident plane wave solution $\phi(\mathbf{r}) = e^{ikz}$. We now ask ourselves what the ougoing waves could be? In 1D this was easy to answer, there were only a reflected and transmitted plane waves. In 3D we should now get spherical waves due to the spherical symmetry of the potential. Also, since the ingoing plane wave was travelling along $z$ we should also insert a dependence on the angular variables $\theta, \phi$. We thus take a scattered spherical wave solution:

$$\psi_S(\mathbf{r}) = f(\theta,\phi)\frac{e^{ikr}}{r} \tag{24.2.3}$$

where $f(\theta, \phi)$ is known as the scattering amplitude.Thus the full wave will be given by:

$$\psi(\mathbf{r}) = e^{ikz} + f(\theta, \phi)\frac{e^{ikr}}{r}, \; r \gg a \tag{24.2.4}$$

## 24.3 Differential Cross section

In a real scattering experiment we measure the so-called differential cross section. Suppose we place a detector which covers a solid angle $d\Omega$ with the impact position. We define the differential cross-section $d\sigma$ as:

$$d\sigma = \frac{\text{no. particles scattered per unit time into } d\Omega}{\text{no. particles incident per unit time per unit area}} = \frac{1}{J_{inc}}\frac{dn}{dt} \tag{24.3.1}$$

where $J_{inc}$ is the probability current of the incoming wave, and $dn$ is the number of particles scattered into $d\Omega$. This interpretation applies well when we consider a scattering process occuring several times giving several particle detections. If we instead only have one particle getting scattered then we replace "number of particles" with "probability of". The differential cross section can be calculated using (24.2.4). Firstly, recall that the incident flux of particles is given by the probability current:

$$\mathbf{J}_{inc} = \frac{\hbar}{m}\,\mathrm{Im}(\phi(\mathbf{r})\nabla\phi(\mathbf{r})) = \frac{\hbar k}{m}\hat{\mathbf{z}} \tag{24.3.2}$$

Now within a solid angle $d\Omega = \sin\theta d\theta d\phi$ there are $dn$ scattered particles by definition. We then have that:

$$dn = |\Psi_S(\mathbf{r})|^2 r^2 d\Omega dr = |f(\theta, \phi)|^2 d\Omega dr \tag{24.3.3}$$

Therefore, since $dt = dr\frac{m}{\hbar k}$ the number of scattered particles into $d\Omega$ per unit time is:

$$\frac{dn}{dt} = \frac{\hbar k}{m}|f(\theta, \phi)|^2 d\Omega \tag{24.3.4}$$

This finally gives us the differential cross section:

$$d\sigma = |f(\theta, \phi)|^2 d\Omega \implies \frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2 \tag{24.3.5}$$

## 24.4 Partial wave expansion

We now want to get to (24.2.4) more rigorously, and find an expression for the scattering amplitude. Since the potential is central we expect that $f(\theta, \phi) = f(\theta)$. We can then expand the full wave ansatz into the well-known solutions:

$$\psi_{E,l,m}(\mathbf{r}) = R_{E,l}(r)Y_{lm}(\theta, \phi)\frac{u_{E,l}(r)}{r}Y_{lm}(\theta, \phi) \tag{24.4.1}$$

satisfying the radial TISE:

$$\Big[ -\frac{\hbar^2}{2m}\Big(\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2}\Big) + V(r)\Big] u_{E,l}(r) = \frac{\hbar^2 k^2}{2m} u_{E,l} \tag{24.4.2}$$

$$\implies \Big( -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + \frac{2mV(r)}{\hbar^2 2}\Big) u_{E,l}(r) = k^2 u_{E,l} \tag{24.4.3}$$

Let's go to the limit where $r \gg a$ so that $V(r) = 0$. We assume that in this limit the centrifugal potential is still dominating (this does not hold for the Coulomb potential for example). Introducing the variable $\rho = kr$ reveals the Bessel equation:

$$\Big( -\frac{d^2}{d\rho^2} + \frac{l(l+1)}{\rho^2}\Big) u_{E,l} = u_{E,l} \tag{24.4.4}$$

The solutions to (24.4.4) are given by the spherical Bessel functions $J_l(\rho), B_l(\rho)$ (first and second type respectively):

$$u_{E,l} = A_l \rho J_l(\rho) + N_l \rho N_l(\rho) \tag{24.4.5}$$

where

$$\begin{cases} J_l(\rho) \approx \frac{1}{\rho}\sin\left(\rho - \frac{l\pi}{2}\right) \\ N_l(\rho) \approx -\frac{1}{\rho}\cos\left(\rho - \frac{l\pi}{2}\right), \end{cases} \quad , \ \rho \to \infty \tag{24.4.6}$$

Consequently we should find that:

$$\psi(\mathbf{r}) = \sum_{lm} (A_{l,m} J_l(kr) + B_{l,m} N_l(kr)) Y_{lm}(\theta, \phi) \tag{24.4.7}$$

Note however that due to the symmetry of the problem our solution should have azimuthal symmetry, forcing us to require $m = 0$ for which the spherical harmonics have no $\phi$-dependence. Our solution now reads:

$$\boxed{\psi(\mathbf{r}) = \sum_{l} (A_l J_l(kr) + B_l N_l(kr)) Y_{l,0}(\theta)} \tag{24.4.8}$$

The $e^{ikz}$ plane wave does not diverge at $r = 0$ while $N_l(kr)$ does. This allows us to set $B_{l,} = 0$ for the part of the solution giving us the incident wave:

$$e^{ikz} = e^{ikr\cos\theta} = \sum_l a_l J_l(kr) Y_{l,0}(\theta) \tag{24.4.9}$$

The $a_l$ can be determined as follows. Using the orthonormality of $Y_{l,m}$ we get:

$$\int e^{ikr\cos\theta} Y_{l,0}(\theta) d\Omega = a_l J_l(kr) \iff 2\pi\sqrt{\frac{2l+1}{4\pi}} \int_{-1}^{1} e^{ikrx} P_l(x) dx = a_l J_l(kr) \tag{24.4.10}$$

Luckily, the spherical Bessel functions of first type have a nice integral representation:

$$J_l(\rho) = \frac{1}{2}\frac{1}{(i)^l} \int_{-1}^{1} e^{i\rho x} P_l(x) dx \tag{24.4.11}$$

allowing us to write

$$a_l = \sqrt{4\pi}\sqrt{2l+1}(-1)^{l/2} \tag{24.4.12}$$

and giving us the following expansion

$$e^{ikz} = \sqrt{4\pi}\sum_l \sqrt{2l+1}(-1)^{l/2} J_l(kr) Y_{l,0}(\theta) \tag{24.4.13}$$

$$\approx \frac{\sqrt{4\pi}}{k}\sum_l \sqrt{2l+1}(-1)^{l/2}\frac{1}{2i}\left(\frac{e^{i(kr-l\pi/2)}}{r} - \frac{e^{-i(kr-l\pi/2)}}{r}\right),\ r \gg a \tag{24.4.14}$$

We have managed to express the plane wave as a superposition of ingoing and outgoing spherical waves, known as partial waves. Note also that this expansion includes an ingoing wave (incident on the potential) and an outgoing wave (reflected by the potential). Each one of these partial waves will scatter with the potential $V(r)$ and also produce $\psi_{sc}$.

Taking the $r \gg a$ limit of our ansatz in (24.4.8) we find:

$$\psi(\mathbf{r}) = \sum_l (A_l \rho J_l(kr) + B_l \rho N_l(kr)) Y_{l,0}(\theta) \tag{24.4.15}$$

$$= \sum_l A_l \left(\frac{\sin(kr - l\pi/2)}{kr} - \frac{B_l}{A_l}\frac{\cos(kr-l\pi/2)}{kr}\right) Y_{l,0}(\theta) \tag{24.4.16}$$

We let $A_l = C_l \cos\delta_l$ and $B_l = C_l \sin\delta_l$ which yields

$$\psi(\mathbf{r}) = \sum_l C_l \frac{\sin(kr - l\pi/2 + \delta_l)}{kr} Y_{l,0}(\theta) \tag{24.4.17}$$

$$= \sum_l \frac{e^{-i\delta_l}}{2i} C_l \left(\frac{e^{i(kr-l\pi/2)}e^{2i\delta_l}}{kr} - \frac{e^{i(kr-l\pi/2)}}{kr}\right) Y_{l,0}(\theta) \tag{24.4.18}$$

Equating this to (24.2.4) we find that:

$$f(\theta)\frac{e^{ikr}}{r} = \sum_l e^{-\delta_l} C_l \left(\frac{e^{kr-l\pi/2}e^{2\delta_l}}{kr} - \frac{e^{kr-l\pi/2}}{kr}\right) Y_{l,0}(\theta) \tag{24.4.19}$$

$$- \frac{\sqrt{4\pi}}{k}\sum_l \sqrt{2l+1}(-1)^{l/2}\frac{1}{2i}\left(\frac{e^{i(kr-l\pi/2)}}{r} - \frac{e^{-i(kr-l\pi/2)}}{r}\right) Y_{l,0}(\theta) \tag{24.4.20}$$

The LHS is an outgoing spherical wave, so we must have that the ingoing waves on the RHS cancel out. Thus:

$$C_l = \sqrt{4\pi}\sqrt{2l+1}(-1)^{l/2} e^{\delta_l} \tag{24.4.21}$$

The scattering amplitude can now be written as:

$$f(\theta)\frac{e^{ikr}}{r} = \frac{\sqrt{4\pi}}{k}\sum_l \sqrt{2l+1}(-1)^{l/2}\frac{1}{2i} e^{i(kr-l\pi/2)}(e^{2\delta_l} - 1) \tag{24.4.22}$$

We recognise $e^{2\delta_l} - 1 = 2ie^{\delta_l} \sin \delta_l$ and $e^{-l\pi/2} = (-1)^{l/2}$ which finally gives:

$$f(\theta) = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1} e^{\delta_l} \sin \delta_l \tag{24.4.23}$$

Note that by definition, we can find the phase shift from

$$\tan \delta_l = -\frac{B_l}{A_l} \tag{24.4.24}$$

Alternatively, we could have also argued that it would be nice to be able to express $f_k(\theta)\frac{e^{ikr}}{r}$ as an outgoing wave, and insert it into the above. Then we would get:

$$\psi(\mathbf{r}) = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1}(-1)^{l/2} \frac{1}{2i} \left( ?? - \frac{e^{-i(kr+l\pi/2)}}{r} \right), \; r \gg a \tag{24.4.25}$$

Due to the superposition principle each $l$ term will get scattered separately. Thus, the outgoing wave for each $l$ term must have the same amplitude as the ingoing wave, or else there would be an accumulation of probability density between these waves. Consequently, we can insert $?? = \frac{e^{-i(kr+l\pi/2)}}{r} e^{2i\delta_l}$ where $\delta_l$ is a phase shift. We get that:

$$\psi(\mathbf{r}) = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1}(-1)^{l/2} \frac{1}{2i} \left( \frac{e^{-i(kr+l\pi/2)+2i\delta_l}}{r} - \frac{e^{-i(kr+l\pi/2)}}{r} \right), \; r \gg a \tag{24.4.26}$$

just like as in 1D scattering. We can equate this to

$$\psi(\mathbf{r}) = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1}(-1)^{l/2} \frac{1}{2i} \left( \frac{e^{i(kr-l\pi/2)}}{r} - \frac{e^{-i(kr+l\pi/2)}}{r} \right) + f_k(\theta)\frac{e^{ikr}}{r}, \; r \gg a \tag{24.4.27}$$

This then gives:

$$f_k(\theta)\frac{e^{ikr}}{r} = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1}(-1)^{l/2} \frac{e^{i(kr-l\pi/2)}}{r} (e^{2i\delta_l} - 1) \tag{24.4.28}$$

so we find that:

$$f_k(\theta) = \frac{\sqrt{4\pi}}{k} \sum_l \sqrt{2l+1} \frac{1}{2i} e^{i\delta_l} \sin \delta_l \tag{24.4.29}$$

as found earlier. This together with the orthonormality of spherical harmonics can be used to find the total cross-section:

$$\sigma = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l \tag{24.4.30}$$

There is a remarkable result known as the Optical theorem that we can extract out of this

relation. Recall that $Y_{l,0}\theta = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta)$ so setting $\theta = 0$ we get:

$$f_k(0) = \frac{\sqrt{4\pi}}{k}\sum_l \frac{2l+1}{\sqrt{4\pi}}e^{i\delta_l}\sin\delta_l \tag{24.4.31}$$

Taking the imaginary part of the above:

$$\mathrm{Im}(f_k(0)) = \frac{1}{k}\sum_l (2l+1)\sin^2\delta_l = \frac{k}{4\pi}\sigma \tag{24.4.32}$$

In other words, the scattering amplitude can be found from the forward scattering amplitude.

## 24.5 Calculating phase shifts

Suppose we have been given the solution $R_{k,l}(r)$ to the radial equation for $r < a$, inside the potential sphere. This solution must be matched with the general radial solution in (24.4.8) at $r = a$. Hence:

$$R_{k,l}(a) = A_l J_l(ka) + B_l N_l(ka) \tag{24.5.1}$$
$$aR'_{k,l}(a) = ka(A_l J'_l(ka) + B_l N'_l(ka)) \tag{24.5.2}$$

Taking their ratio gives:

$$\frac{R'_{k,l}(a)}{R_{k,l}(a)} = k\frac{A_l J'_l(ka) + B_l N'_l(ka)}{A_l J_l(ka) + B_l N_l(ka)} \tag{24.5.3}$$
$$= k\frac{J'_l(ka) - \tan\delta_l N'_l(ka)}{J_l(ka) - \tan\delta_l N_l(ka)} \tag{24.5.4}$$

Inverting the above gives:

$$\tan\delta_l = \frac{J'_l(ka) - \frac{R'_l(a)}{kR_l(a)}J_l(ka)}{N'_l(ka) - \frac{R'_l(a)}{kR_l(a)}N_l(ka)} \tag{24.5.5}$$

### Hard-sphere example

Consider a hard-sphere potential:

$$V = \begin{cases}\infty, & r < a\\ 0, & r > a\end{cases} \tag{24.5.6}$$

The general solution for our wave-function is given by:

$$\psi(r,\theta) = \sum_l [A_l J_l(kr) + B_l N_l(kr)]P_l(\cos\theta) \tag{24.5.7}$$

However, since the potential is infinite inside the sphere we must have that the wave-function vanish on its surface. In other words:

$$\psi(a,\theta) = \sum_l (A_l J_l(ka) + B_l N_l(ka)) P_l(\cos\theta) = 0 \tag{24.5.8}$$

Luckily, the Legendre polynomials are orthonormal so we must have that:

$$A_l J_l(ka) + B_l N_l(ka) = 0 \implies \frac{B_l}{A_l} = -\frac{J_l(ka)}{N_l(ka)} \tag{24.5.9}$$

This allows us to calculate the phase shift:

$$\tan\delta_l = \frac{J_l(ka)}{N_l(ka)} \implies \sin^2\delta_l = \frac{J_l^2(ka)}{J_l^2(ka)+N_l^2(ka)} \tag{24.5.10}$$

giving us the total cross section:

$$\sigma = \frac{4\pi}{k^2}\sum_l (2l+1)\frac{J_l^2(ka)}{J_l^2(ka)+N_l^2(ka)} \tag{24.5.11}$$

## 24.6 Green's functions

Let's go back to the TISE:

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\Big]\psi(\mathbf{r}) = \frac{\hbar^2 k^2}{2m}\psi(\mathbf{r}) \tag{24.6.1}$$

Let $V(\mathbf{r}) = \frac{\hbar^2}{2m}U(\mathbf{r})$ giving us:

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = U(\mathbf{r})\psi(\mathbf{r}) \tag{24.6.2}$$

This equation is just begging for Green's functions if we view the potential $U(\mathbf{r})$ as some sort of source. We thus try to solve:

$$(\nabla^2 + k^2)G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{24.6.3}$$

Suppose we have found a solution $\psi_0(\mathbf{r})$ to the homogeneous problem:

$$(\nabla^2 + k^2)\psi_0(\mathbf{r}) = 0 \tag{24.6.4}$$

Then, just like one would solve an inhomogeneous ODE, we would have that:

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}'\psi(\mathbf{r}') \tag{24.6.5}$$

Indeed:

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = \int d\mathbf{r}'\delta(\mathbf{r}-\mathbf{r}')U(\mathbf{r}')\psi(\mathbf{r}') = U(\mathbf{r})\psi(\mathbf{r}) \tag{24.6.6}$$

as desired. The Green's function is known to be:

$$G_{\pm}(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\frac{e^{\pm ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \tag{24.6.7}$$

so the total wave-function is given by:

$$\psi(\mathbf{r}) = e^{ikz} + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')d\mathbf{r}' \tag{24.6.8}$$

Now for $r \gg a$ we have that $|\mathbf{r}-\mathbf{r}'| \approx r$ in (24.6). The same cannot be done for the exponential in since the wave-vector could be large. Taking $\mathbf{n}$ to be the unit vector along $\mathbf{r}$ then we can still write:

$$e^{\pm ik|\mathbf{r}-\mathbf{r}'|} \approx e^{\pm i\mathbf{k}\cdot\mathbf{r}}e^{\mp k\mathbf{n}\cdot\mathbf{r}'},\ r \gg a \tag{24.6.9}$$

For large $r$ the Green function can thus be approximated as:

$$G_{\pm}(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi r}e^{\pm ikr}e^{\mp ik\mathbf{n}\cdot\mathbf{r}'} \tag{24.6.10}$$

Inserting this into (24.6.8) and keeping only the $+$ Green function:

$$\psi(\mathbf{r}) = e^{ikz} - \frac{e^{ikr}}{r}\int \frac{1}{4\pi r}e^{-ik\mathbf{n}\cdot\mathbf{r}'}U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{24.6.11}$$

This allows us to identify the scattering amplitude:

$$f(\theta) = -\int \frac{1}{4\pi}e^{-ik\mathbf{n}\cdot\mathbf{r}'}U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{24.6.12}$$

## 24.7 Born approximation

The integral equation equation in (24.6.8) can be solved iteratively using the Born aproximation. Indeed substituting $\mathbf{r} \to \mathbf{r}'$ and $\mathbf{r}' \to \mathbf{r}''$ we find that:

$$\psi(\mathbf{r}') = e^{ikz'} + \int G(\mathbf{r}',\mathbf{r}'')U(\mathbf{r}'')d\mathbf{r}'' \tag{24.7.1}$$

which can be substituted into (24.6.8) to get:

$$\psi(\mathbf{r}) = e^{ikz} + \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')e^{ikz'}d\mathbf{r}' \tag{24.7.2}$$

$$+ \int G(\mathbf{r},\mathbf{r}')U(\mathbf{r}')\int G(\mathbf{r}',\mathbf{r}'')U(\mathbf{r}'')\psi(\mathbf{r}'')d\mathbf{r}''d\mathbf{r}' \tag{24.7.3}$$

We can iterate this process, and schematically get the Born series:

$$\psi(\mathbf{r}) = e^{ikz} + \int GUe^{ikz} + \int GU\int GUe^{ikz} + \int GU\int GU\int GUe^{ikz} + \ldots \tag{24.7.4}$$

This approximation works well when the free particle solution dominates over the interactions i.e. the incident wave has high energy relative to the scattering potential.

The first Born approximation only keeps the terms in (24.7.2) giving:

$$\psi(\mathbf{r}) = e^{ikz} + \int G(\mathbf{r}, \mathbf{r}')U(\mathbf{r}')e^{ikz'}d\mathbf{r}' \tag{24.7.5}$$

Inserting (24.6.10) yields

$$\psi(\mathbf{r}) = e^{ikz} - \frac{e^{ikr}}{r}\int \frac{1}{4\pi}e^{-ik\mathbf{n}\cdot\mathbf{r}'}U(\mathbf{r}')e^{ikz'}d\mathbf{r}' \tag{24.7.6}$$

Letting $\mathbf{K} = \mathbf{k}_s - \mathbf{k}_i = k(\mathbf{n} - k\mathbf{n}_i)$ be the difference in the wave-vector of the scattered and incident waves, then:

$$\psi(\mathbf{r}) = e^{ikz} - \frac{e^{ikr}}{r}\int \frac{1}{4\pi}e^{-i\mathbf{K}\cdot\mathbf{r}'}U(\mathbf{r}')d\mathbf{r}' \tag{24.7.7}$$

and hence the scattering amplitude takes the form:

$$\boxed{f(\theta) = -\frac{1}{4\pi}\int e^{-i\mathbf{K}\cdot\mathbf{r}'}U(\mathbf{r}')d\mathbf{r}'} \tag{24.7.8}$$

The scattering amplitude is just the fourier transform of the scattering potential at momentum $\mathbf{K}$. For a spherical potential we can perform some further simplifications:

$$f(\theta) = -\frac{1}{2}\int e^{-i\mathbf{K}\cdot\mathbf{r}'}U(r')r^2dr'\sin\theta d\theta \tag{24.7.9}$$

We can fix $\mathbf{K}$ and let $\theta'$ be its angle with $\mathbf{r}'$.Then:

$$f(\theta) = -\frac{1}{2}\int e^{-iKr'\cos\theta'}U(r')r^2dr'\sin\theta' d\theta' \tag{24.7.10}$$

$$= \frac{1}{2}\int \frac{1}{iKr'}\left(e^{-iKr'\cos\theta'}\right)_0^{\pi}U(r')r^2dr' \tag{24.7.11}$$

$$= -\frac{1}{K}\int r\sin(Kr')U(r')dr' \tag{24.7.12}$$

finally giving:

$$\boxed{f(\theta) = -\frac{2m}{\hbar^2 K}\int r\sin(Kr')V(r')dr'} \tag{24.7.13}$$

Also note that

$$K = 2k\sin(\theta/2) \tag{24.7.14}$$

where $\theta$ is the angle between $\mathbf{n}$ and $\mathbf{n}_i$.

# Part V

# Relativistic Quantum Mechanics

# The Klein Gordon equation

25

# The Dirac equation

26

# What next? Quantum field theory.

27

## Acknowledgments

This is the most common positions for acknowledgments. A macro is available to maintain the same layout and spelling of the heading.

**Note added.** This is also a good position for notes added after the paper has been written.