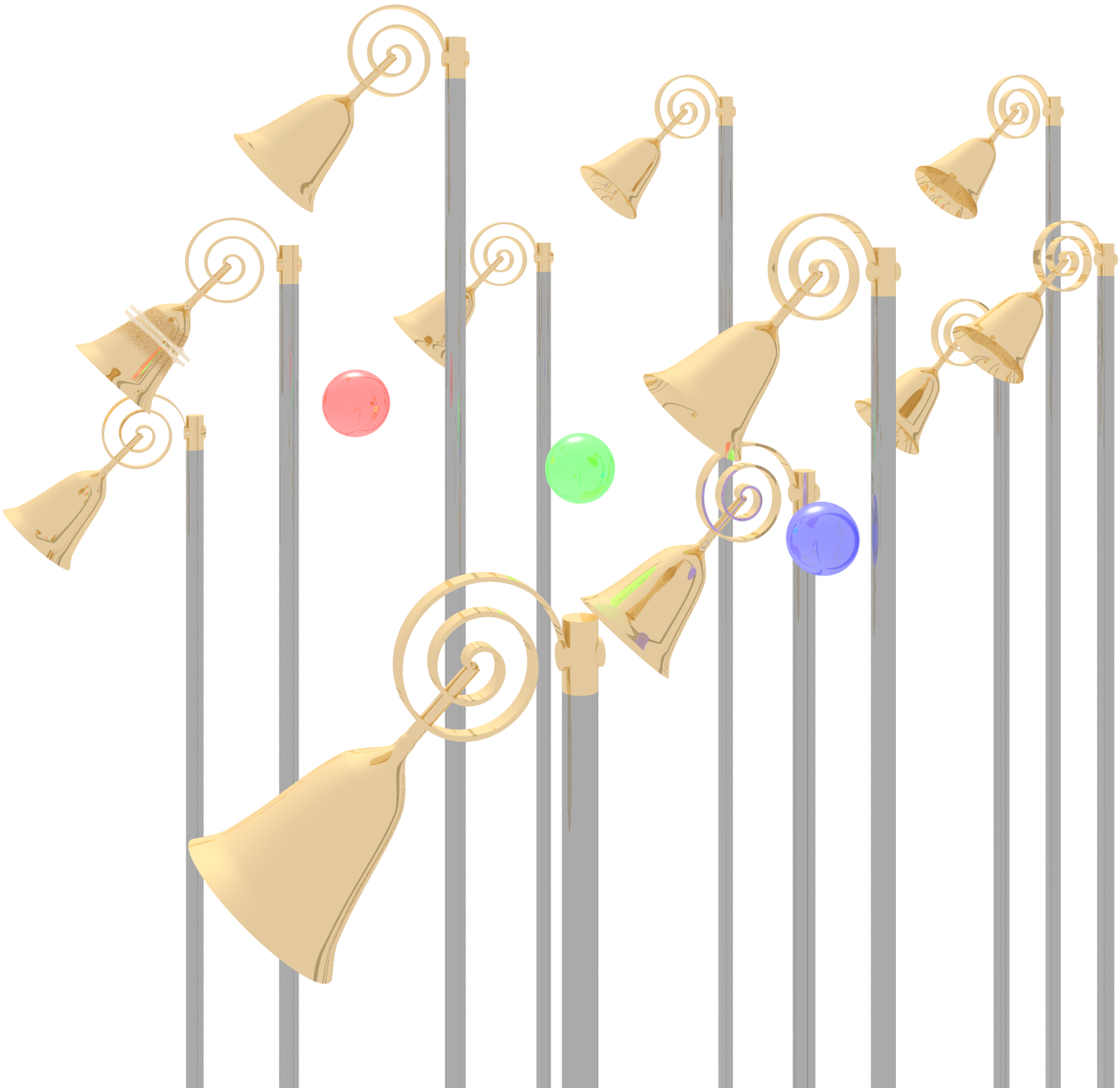
Electron's Destiny

A Very Short Advanced QM Book by

N. Booker



To my parents

Contents

1 Instead of a foreword **4**

1.1 Acknowledgements (4)

2 Formalism **5**

2.1 Vector spaces (5) 2.2 Operations in vector spaces (7) 2.3 Basis vectors (8) 2.4 Spectral decomposition (9) 2.5 Change of basis (11) 2.6 Operators (11) 2.7 Hermitian and unitary operators (13) 2.8 Projector (14) 2.9 Time evolution (15) 2.10 Measurement of discrete systems (15) 2.11 Density operator (16) 2.11.1 Symmetries (17)

3 Composite systems **19**

3.1 Tensor product (19) 3.2 Deletion of systems (21)

4 Time-dependent perturbation theory **23**

4.1 Perturbative expansion (23) 4.2 First and second orders (24) 4.3 (Almost) constant perturbation (25) 4.4 Harmonic perturbation (26)

5 Unitary time evolution **28**

5.1 Time evolution operator (28) 5.2 Suzuki-Trotter expansion (30) 5.3 Schrödinger and Heisenberg pictures (30) 5.4 Interaction picture (31)

6 Quantum light and atoms **32**

6.1 Quantum harmonic oscillator (32) 6.2 Light (32) 6.3 Light-atom interaction revisited (33)

7 Open quantum systems **35**

7.1 Superoperators (35) 7.2 Markovian evolution (36)

8 WKB method **39**

8.1 WKB wavefunction (39) 8.2 Uniform approximaion (41)

Chapter 1

Instead of a foreword

1.1 Acknowledgements

Electron's Destiny emerged from a series of lecture notes based on the advanced quantum mechanics course at University College London lectured by Prof. Tania Monteiro and Prof. Marzena Szymańska as well as Prof. Monteiro's notes for the course in the 2024-25 year. Other sources used for the book include:

- *Quantum Theory: Concepts and Methods* by Asher Peres (Technion - Israel Institute of Technology)
- *Advanced Quantum Mechanics* by Ben Simons (University of Cambridge)
- *Quantum Mechanics* by Eugen Merzbacher (University of North Carolina at Chapel Hill)

I want to extend my gratitude to Felix Halbwedl, Robert Schwarzl and Francisco Silva[1] for discussions and advice in improving this book and proofreading drafts of this book. I am also eternally indebted to Abhijeet Vats, under whose guidance I was able to develop my LaTeX skills to a satisfactory level. Without them, this book would undoubtedly not have been in its current form.
For any comments, suggestions or typos, please e-mail `zcapxix(at)ucl(dot)ac(dot)uk`.

[1] Who, importantly, proposed that postulates have their own `tcolorbox`.

Chapter 2

Formalism

2.1 Vector spaces

And now, after having avoided properly defined[1] fields for many years, it is time to confront them.

Definition 2.1 (Field) A *field* F is a set of elements a, b, c, $\cdots$ with two operations defined on it - the *addition* $+$ and the *multiplication* $\cdot$[a]. The operations observe the following axioms:

- **Associativity:**
$$a + (b + c) = (a + b) + c \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \tag{2.1}$$
- **Commutativity:**
$$a + b = b + a \quad a \cdot b = b \cdot a \tag{2.2}$$
- **Identity:**
$$\exists 0, 1 \in F \quad \text{such that} \quad a + 0 = a \quad \text{and} \quad a \cdot 1 = a \tag{2.3}$$
- **Inverse:**
$$\forall a \in F \quad \exists -a \in F \quad \text{such that} \quad -a + a = 0 \tag{2.4}$$
$$\forall a \neq 0 \in F \quad \exists -a^{-1} \in F \quad \text{such that} \quad a \cdot a^{-1} = 1 \tag{2.5}$$
- **Distributativity:**
$$a \cdot (b + c) = a \cdot b + a \cdot c \tag{2.6}$$

[a] Although these operations are not necessarily the conventional addition and multiplication in real numbers.

Definition 2.2 (Vector space) A *vector space* is a collection of objects that observe *vector addition* and scalar multiplication. While the two operations can be intuitively understood, we define them here for the sake of formality.

- **Vector addition:**
  - **Associativity:**
$$\vec{a} + (\vec{b} + \vec{w}) = \left(\vec{a} + \vec{b}\right) + \vec{w} \tag{2.7}$$
  - **Commutativity:**
$$\vec{a} + \vec{b} = \vec{b} + \vec{a} \tag{2.8}$$
  - **Identity:** There always exists a $\vec{0}$ for which
$$\vec{a} + \vec{0} = \vec{a} \tag{2.9}$$
    The zero vector is also called the *null vector*.

[1] By a physicist's standard.

  - **Inverse:** For every $\vec{a}$, there is always a $-\vec{a}$ for which

$$\vec{a} + -\vec{a} = 0 \tag{2.10}$$

- **Scalar multiplication:**
  - **Associativity:**

$$\lambda_1(\lambda_2\vec{a}) = (\lambda_1\lambda_2)\vec{a} \tag{2.11}$$

  - **Distributivity:**

$$\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b} \quad \text{and} \quad (\lambda_1 + \lambda_2)(\vec{a}) = \lambda_1\vec{a} + \lambda_2\vec{a} \tag{2.12}$$

  - **Negative:**

$$-1 \times \vec{a} = -\vec{a} \tag{2.13}$$

Remark 2.1 In mathematics, a *space* refers to a set of objects that follow specific rules and exhibit a particular structure. i.e. it is an 'environment' in which every object transforms in a certain way. Conversely, one can consider all possible objects that satisfy these specific rules to cover the entirety of the space they constitute.

> **Quote 2.1** Well, "theorem" refers to adult things that mathematicians work with, not kiddy things that physicists confuse themselves about.
>
> *Abhijeet Vats, 31 July 2024*

As kiddy physicists, we are satisfied by merely familiarising ourselves with the simplest of vector spaces: the *real vector space* $\mathbb{R}^n$ and the *complex vector space* $\mathbb{C}^n$.

Definition 2.3 (Real vector space) The real vector space $\mathbb{R}^n$ consists of all ordered n-tuples of real numbers. i.e. it is the set of all vectors of the form:

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, 1 \leq i \leq n\} \tag{2.14}$$

Each vector in $\mathbb{R}^n$ is an ordered list of n real numbers, and the operations in this vector space (addition and scalar multiplication) are defined using real numbers.

Definition 2.4 (Complex vector space) The complex vector space $\mathbb{C}^n$ consists of all ordered n-tuples of complex numbers. i.e. it is the set of all vectors of the form:

$$\mathbb{C}^n = \{(z_1, z_2, \ldots, z_n) \mid z_i \in \mathbb{C}, 1 \leq i \leq n\} \tag{2.15}$$

Each vector in $\mathbb{C}^n$ is an ordered list of n complex numbers, and the operations in this vector space (addition and scalar multiplication) are defined using complex numbers.

Definition 2.5 (Tuple) A *tuple* is an ordered list of objects. An n-tuple is simply a tuple that has n elements. 0-, 1- and 2-tuples are alternatively called the *empty tuple*, the *singleton* and the *ordered pair* respectively.

Remark 2.2 For vector spaces, the n-tuple consists of the components of the vector. Hence, n represents the number of dimensions of the vector space.

Definition 2.6 (Subspace) For a vector space V, a *subspace* $W \subset V$ satisfies all the conditions of a vector space but only contains some of the vectors that V contains. For example, if W has one less (but otherwise identical) basis vector than V, then $W \subset V$.

Definition 2.7 (Linear functional) A *functional* is a function that operates on a vector and returns a scalar. A *linear functional* (also known as a *one-form* or a *covector*) is a functional that satisfies

$$f(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2) = \alpha_1 f(\vec{v}_1) + \alpha_2 f(\vec{v}_2) \tag{2.16}$$

We look at two examples:

- The divergence $\nabla\cdot$ is *not* a functional, as it maps vector fields to scalar fields rather than mapping directly to scalars. Despite this, one can construct a functional by integrating the divergence of a vector field and applying the divergence theorem, after which one finds a scalar.
- The bra is a functional as it operates on a ket (i.e. vector) and returns a scalar.

From this, we can define a *dual space*:

Definition 2.8 (Dual space) The set of linear functionals V^* acting on a vector space V is a *dual space* (or the *dual* of V).

Remark 2.3 This makes more sense if we consider the inner product - bras (row vectors) operate on kets (column vectors), yielding scalars.
Remark 2.4 As column vectors are simply called vectors, row vectors are called *dual vectors*.

2.2 Operations in vector spaces

Again, from intuition, it is necessary to assign the notion of a 'length' to vectors. One can easily recognise this notion as the *norm* of a vector. Again, we define it more rigorously:

Definition 2.9 (Vector norm) The so-called vector norm observes the following properties:

- **Scaling behaviour:**
$$||\lambda\vec{a}|| = |\lambda||\vec{a}|| \tag{2.17}$$
- **Zero vector:**
$$||\vec{a}|| = 0 \tag{2.18}$$
is satisfied if and only if $\vec{a} = \vec{0}$.
- **Triangle inequality:**
$$||\vec{a} + \vec{b}|| \leq ||\vec{a}|| + ||\vec{b}|| \tag{2.19}$$

Remark 2.5 In 4D spacetime, if the norm of a vector is positive, the vector is *spacelike*. If the norm is negative, the vector is *timelike*. If the norm is zero, the vector is *lightlike/null*.
We can generalise the vector norm to a tensor of any rank using the index notation:

Definition 2.10 (Norm) The *norm* of a tensor generalises the concept of *magnitude* in vectors. For a rank-1 tensor:
$$||V|| = \sqrt{g_{ab}V^aV^b} \tag{2.20}$$
where g_{ab} is the metric tensor. For a rank-2 tensor:
$$||T|| = \sqrt{g_{ac}g_{bd}T^{ab}T^{cd}} \tag{2.21}$$
and so on.

Remark 2.6 The norm is significant in QM as it is closely related to probability.
While the norm is concerned with length, the *inner product* describes the general relation between two vectors, especially with respect to the angle between them. One can better understand this by recalling the most simple inner product $(\vec{a}, \vec{b})$:
$$(\vec{a}, \vec{b}) = \cos(\theta)||\vec{a}||\vec{b}|| \tag{2.22}$$
Again, it is convenient to define the inner product in index notation:

Definition 2.11 (Inner product) For rank-1 tensors:
$$\langle A, B\rangle = g_{ij}A^iB^j \tag{2.23}$$
For rank-2 tensors:
$$\langle A, B\rangle = g_{ac}g_{bd}A^{ab}B^{cd} \tag{2.24}$$
and so on.

Theorem 2.1 (Inner product properties) Unitary operators have the following properties:

- **Linearity in the *second argument*:**

$$(\vec{a}, \lambda\vec{b}) = \lambda(\vec{a}, \vec{b}) \tag{2.25}$$

- **Positive-definiteness:**

$$(\vec{a}, \vec{a}) \geq 0 \tag{2.26}$$

  where $(\vec{a}, \vec{a}) = 0$ only happens when $\vec{a}$ is the null vector.

- **Conjugate symmetry**

$$\left(\vec{a}, \vec{b}\right) = \left(\vec{b}, \vec{a}\right)^* \tag{2.27}$$

Definition 2.12 (Inner product space) A vector space with an inner product defined (i.e. all inner product properties are defined) is called an *inner product space*.

Definition 2.13 (Pseudo-Hilbert space) An inner product space which drops the positive-definiteness requirement (i.e. $(\vec{a}, \vec{a}) < 0$ is possible) is called a *indefinite inner product space* or a *pseudo-Hilbert space*.

Remark 2.7 The reasoning for the latter name will become clear when we discuss Hilbert spaces.

Definition 2.14 (Outer product) In index notation, the *outer product* is very simple:

$$C_{ij} = A_i B_j \tag{2.28}$$

> **Quote 2.2** There are many evil versions of outer products.
>
> *Tania Monteiro, 1 October 2024*

2.3 Basis vectors

Previously we briefly mentioned how all the elements of a vector space cover the entirety of said vector space. We now explore this more formally.

Definition 2.15 (Spanning set) The *spanning set* of a vector space is a set of vectors such that we can represent any point in the vector space with a combination of these vectors. Spanning sets of a vector space are said to *span* the vector space.

We note that the members of the spanning set do not have to be linearly independent, in which case redundancy arises. To eliminate this redundancy, we define a special category of spanning sets.

Definition 2.16 (Basis set) *Basis sets* are spanning sets whose members, the *basis vectors*, are all linearly independent from each other.

From a basis, we can recover the definition of dimensions:

Definition 2.17 (Dimension) The *dimension* of a vector space is the number of elements of its basis sets.

Derivation 2.1 (Spin-$\frac{1}{2}$ particles) The basis set of a spin-$\frac{1}{2}$ particle consists of its up- and down-states. Hence, the state space of a spin-$\frac{1}{2}$ particle is 2-dimensional. To generalise, one can prove that the state space of a spin-$\frac{n}{2}$ particle has $n + 1$ dimensions.

Exercise 2.1 Do it.

Derivation 2.2 (Polynomials) For a polynomial one has the basis $f_j(x) = x^k$. This spans into infinity, and the vector space of polynomials have infinite dimensions.

Now we still need something to convey the concept of perpendicularity. This manifests in *orthogonality*. Using the bra-ket notation:

Theorem 2.2 (Orthogonality) Two vectors are orthogonal when their inner product yields zero.

Remark 2.8 While orthogonality always implies linear independence, the same does not apply in reverse. For example, vectors $(1, 0)$ and $(1, 1)$ are linearly independent but *not* orthogonal.

Theorem 2.3 (Normalisation) A vector is *normalised* if

$$(\vec{a}, \vec{a}) = 1 \tag{2.29}$$

An eigenstate is always normalised.

Theorem 2.4 (Orthonormality) Combining orthogonality and normalisation, two states are *orthonormal* if

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \tag{2.30}$$

An *orthonormal basis* refers to a basis set that is both orthogonal and normalised.

Definition 2.18 (Hilbert space) A *Hilbert space* $\mathcal{H}$ is a complex inner product space that is complete with respect to the norm induced by the inner product. i.e. every Cauchy sequence in the space (with respect to the norm) converges to a limit that is also within the space.

Remark 2.9 We distinguish between finite-dimensional and infinite-dimensional Hilbert spaces:

- Finite-dimensional Hilbert spaces are sufficient for systems with a discrete set of states (commonly seen in quantum information). For example, a qubit is described by a two-dimensional Hilbert space, and a system of N qubits has a Hilbert space of dimension 2^N.
- Infinite-dimensional Hilbert spaces are used to describe systems with continuous variables, such as the position and momentum of particles in space. This corresponds well to wavefunctions, which are continuous.

Remark 2.10 For simplicity, we only investigate finite-dimensional Hilbert spaces in this book, and we will simply call them *Hilbert spaces*. In mathematics, the designation 'Hilbert space' by default refers to infinite-dimensional Hilbert spaces, while finite-dimensional Hilbert spaces are unimaginatively called *finite-dimensional Hilbert spaces*.

> **Quote 2.3** Quantum phenomena do not occur in a Hilbert space. They occur in a laboratory.
>
> *Asher Peres, in* Quantum Theory: Concepts and Methods*, 1993*

2.4 Spectral decomposition

We first recall the *bra-ket notation*:

Definition 2.19 (Bra-ket notation) In QM, the bra-ket notation or the *Dirac notation*. In it, vectors and covectors are denoted by *kets* $|\Psi\rangle$ and *bras* $\langle\Psi|$ respectively.

> **Quote 2.4** It was invented so that Dirac wouldn't have to learn more math than he absolutely needed. But also, I mean, it has given rise to a lot of nice mathematics. Rigged Hilbert spaces and nuclear spaces are good examples.
>
> *Abhijeet Vats, on the bra-ket notation, 4 May 2024*

In the bra-ket notation, wavefunctions become *state vectors*. We hence have the 1st postulate[2] of QM:

[2] Sometimes also known as 'axiom'.

Postulate 1 Every isolated physical system has associated with it a complex inner product space or Hilbert Space. This is known as its *state space*. A complete description of the system is given by its state vector, $|\Phi\rangle$, a unit-normalised vector in the state space.

> **Quote 2.5** Other thing, you announce the postulates as theorems, but could It be done like:
> Postulate 1 - ...
>
> *Francisco Silva, on a previous edition of the notes, 29 November 2024*

Definition 2.20 (Spectral representation) Any wavefunction $|\Psi\rangle$ can be 'broken down' into its *spectral representation* as follows:

$$|\Psi\rangle = \sum_j c_j |\psi_j\rangle \tag{2.31}$$

where c_j represents the probability amplitude of finding the system in the (eigen)state $|\psi_j\rangle$.
This process is then called *spectral decompositon*, the naming is due to it being closely related to spectral theory in mathematics.

Remark 2.11 The corresponding bra wavefunction's representation is then

$$\langle\Psi| = \sum_j c_j^* \langle\psi_j| \tag{2.32}$$

Note 2.1 While ϕ and ψ are the most common symbols for wavefunctions, the notation could in principle be any letter.

Derivation 2.3 (Spin-$\frac{1}{2}$ particles revisited) Previously we have seen that the state space of a spin-$\frac{1}{2}$ particle is 2-dimensional. We also recall that for such a particle $J = S = 1/2$, and the $|j, m_s\rangle$ basis states are

$$|\uparrow\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad |\downarrow\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{2.33}$$

As such, any arbitrary state can be written as

$$|\Psi\rangle = c_1|\uparrow\rangle + c_2|\downarrow\rangle = \begin{pmatrix}c_1\\c_2\end{pmatrix} \tag{2.34}$$

One can utilise the spectral representation to simplify the inner product. We migrate the inner product, previously seen in the wavefunction formalism

$$\underbrace{\int \Psi^*(x)\chi(x)dx}_{\text{wavefunction}} = \underbrace{\langle\Psi|\chi\rangle}_{\text{bra-ket}} \tag{2.35}$$

where x is some parameter. Recalling that $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, we reduce the inner product

$$\langle\Psi|\chi\rangle = \sum_{j,k} \left(c_j^*\langle\phi_j|\right)\left(b_k|\phi_k\rangle\right) = \sum_{j,k} c_j^* b_j = (\vec{c}, \vec{b}) = (\vec{c})^\dagger \cdot \vec{b} \tag{2.36}$$

where, for convenience, we have used the shorthand

$$\vec{c} = \begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix} \quad (\vec{c})^\dagger = (c_1, c_2, \ldots, c_n) \tag{2.37}$$

Remark 2.12 The norm can also be simplified by recalling that the norm of a vector $\vec{a}$ is merely the $\sqrt{(\vec{a}, \vec{a})}$.

2.5 Change of basis

Each matrix or operator is defined in some set of *basis states*. This is conceptually identical to a *coordinate system*. Often it would be useful to perform a *change of basis*, which corresponds to the concept of coordinate transformations. Note that this derivation is presented here mostly for completeness, and a more practical derivation can be found in Derivation **??**.

Derivation 2.4 (Change of basis) The same ket vector can be represented in different coordinate systems:

$$|\Psi\rangle = \sum_n c_n|\phi_n\rangle \rightarrow |\Psi\rangle = \sum_m d_m|\chi_m\rangle \tag{2.38}$$

Then, c and d in Dirac notation can be found:

$$c_n = \langle\phi_n|\Psi\rangle \quad d_m = \langle\chi_m|\Psi\rangle \tag{2.39}$$

Definition 2.21 (Similarity transform) We define the so-called *similarity transform* S, a matrix for which we have

$$S_{mn} = \langle\chi_m|\phi_n\rangle \tag{2.40}$$

Using closure relation:

$$d_m = \langle\chi_m|\phi_n\rangle\langle\phi_n|\Psi\rangle \tag{2.41}$$

which yields

$$d_m = S_{mn}c_n \rightarrow d = \hat{S}c \tag{2.42}$$

One can also perform a change of basis on a two-dimensional operator[a]:

$$\langle\chi|\hat{A}|\chi\rangle = \langle\chi|\phi\rangle\langle\phi|\hat{A}|\phi\rangle\langle\phi|\chi\rangle = S_{mn}\langle\phi|\hat{A}|\phi\rangle S_{nm} \rightarrow \hat{A}' = \hat{S}\hat{A}\hat{S}^\dagger \tag{2.43}$$

Note 2.2 Notice that in $S_{mn} = \langle\chi_m|\phi_n\rangle$, the *target* basis is the bra, while the *initial* basis is the ket!

[a] Which we will cover almost immediately.

Remark 2.13 Eigenvalues, determinants and traces are invariant under basis transformations.

2.6 Operators

Definition 2.22 (Linear operator) A linear operator $\hat{O}$ is a mapping $V \rightarrow W$ for vector spaces V and W satisfying the linearity condition

$$\hat{O}[\alpha|\phi\rangle \ + \beta|\psi\rangle] = \hat{O}\alpha|\phi\rangle \ + \hat{O}\beta|\psi\rangle \tag{2.44}$$

for any scalars $\alpha, \beta \in \mathbb{C}$ and vectors $|\phi\rangle, |\psi\rangle \in V$.

Remark 2.14 In this book, the notation is such that an operator $\hat{O}$ is denoted by a hat, while its matrix representation in a certain basis O_{mn} is denoted with indices and without a hat.

Definition 2.23 (Bounded and unbounded operators) In a Hilbert space $\mathcal{H}$, an operator $\hat{O}$ is...

- ...*bounded* if $\hat{O}$ is defined on the whole of $\mathcal{H}$.
- ...*unbounded* if $\hat{O}$ is defined on the whole of $\mathcal{H}$.

The objective is then to determine the matrix form of an operator in a certain basis. To do so, we migrate

the previously-known outer product to the bra-ket notation:

$$\underbrace{\int a(x)b^*(x)dx}_{\text{wavefunction}} = |\phi\rangle\langle\psi| = \underbrace{\begin{bmatrix} \phi_1\psi_1 & \phi_1\psi_2 & \dots & \phi_1\psi_n \\ \phi_2\psi_1 & \phi_2\psi_2 & \dots & \phi_2\psi_n \\ \vdots & \vdots & \ddots & \vdots \\ \phi_m\psi_1 & \phi_m\psi_2 & \dots & \phi_m\psi_n \end{bmatrix}}_{\text{bra-ket}} \tag{2.45}$$

where x is again some parameter.

Theorem 2.5 (Linearity of the outer product) The outer product has the properties of a linear operator. Acting on a ket vector, it returns a ket vector:

$$|\psi_1\rangle\langle\psi_2| \cdot |\psi_3\rangle = |\psi_1\rangle\langle\psi_2|\psi_3\rangle = \langle\psi_2|\psi_3\rangle|\psi_1\rangle \tag{2.46}$$

Theorem 2.6 (Closure relation) We have, for some wavefunction $|\phi_j\rangle$:

$$\hat{I} = \sum_j |\phi_j\rangle\langle\phi_j| \tag{2.47}$$

Retroactively from this, we can recover the definition of the *identity matrix*.

Exercise 2.2 Prove that

$$\left(\sum_j |\phi_j\rangle \langle\phi_j|\right) |\Psi\rangle = |\Psi\rangle \tag{2.48}$$

Remark 2.15 This is known as *resolution of the identity*.

With the closure relation seen in Equation 2.47, one can simplify Derivation 2.4 somewhat:

Derivation 2.5 (Change of basis using the identity) The same ket vector can be represented in different coordinate systems:

$$|\Psi\rangle = \sum_k c_k |\phi_k\rangle = \mathbb{I} \cdot \sum_k c_k |\phi_k\rangle = \sum_{j,k} c_k |\chi_j\rangle \langle\chi_j| \cdot |\phi_k\rangle = \sum_{j,k} \langle\chi_j \mid \phi_k\rangle c_k |\chi_j\rangle = \sum_j d_j |\chi_j\rangle \tag{2.49}$$

The same approach can also be used to perform a change of basis on a two-dimensional operator:

$$\langle\chi|\hat{A}|\chi\rangle = \langle\chi|\mathbb{I}\hat{A}\mathbb{I}|\chi\rangle = \langle\chi|\phi\rangle\langle\phi|\hat{A}|\phi\rangle\langle\phi|\chi\rangle = S_{mn}\langle\phi|\hat{A}|\phi\rangle S_{nm} \to \hat{A}' = \hat{S}\hat{A}\hat{S}^\dagger \tag{2.50}$$

Exercise 2.3 Suppose we introduce a unitary operator which affects the basis change for a single j-th member of the basis so $|\phi_j = \hat{S}|\chi_j\rangle$ What form would the operator take? Try out $\hat{S}\sum_n |\phi_n\rangle\langle\chi_n|$ and see if it has the correct behaviour. What about its matrix representation, does it yield the correct matrix elements?

Now we can finally represent an operator in matrix form. For a basis $|\phi\rangle$:

$$\hat{A} = \underbrace{\sum_{j,k} A_{ij}|\phi_j\rangle\langle\phi_k|}_{①} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2.51}$$

Corresponding to the spectral decomposition of a state vector, ① is the spectral decomposition of the operator.

Remark 2.16 Eigenvalues become eigenenergies and eigenvectors become eigenstates. The eigenstate with the smallest eigenenergy is the ground state; the 1st excited state is the eigenstate with the 2nd smallest eigenenergy; and so on.

2.7 Hermitian and unitary operators

We recall that the complex conjugate of some object $\phi = a + bi$ is $\phi^* = a - bi$. We then define the *Hermitian conjugate*:

Definition 2.24 (Hermitian conjugate) The Hermitian conjugate[a] or the *Hermitian adjoint*[b] is essentially the transpose of a complex conjugate.

$$\hat{A}^\dagger = \left(\hat{A}^*\right)^T \tag{2.52}$$

For a vector:

$$|\psi\rangle^\dagger = \langle\psi| \quad \langle\psi|^\dagger = |\psi\rangle \tag{2.53}$$

For an operator:

$$\left(\langle\psi_1|\hat{A}|\psi_2\rangle\right)^* = \langle\psi_2|\hat{A}^\dagger|\psi_1\rangle \tag{2.54}$$

[a] Used in the context of a matrix.
[b] This is used in more mathematical contexts - i.e. an operator.

Note 2.3 When you take the Hermitian conjugate, you *always* reverse the product order.

$$(AB)^\dagger = B^\dagger A^\dagger \tag{2.55}$$

Definition 2.25 (Hermitian operator) An operator is *Hermitian* or (more mathematically) *self-adjoint* if it is its own Hermitian conjugate:

$$\hat{A}^\dagger = \hat{A} \tag{2.56}$$

It observes

$$\langle m|\hat{A}|n\rangle = \langle \hat{A}m|n\rangle = a_m\langle m|n\rangle = a_n\langle m|n\rangle \tag{2.57}$$

Definition 2.26 (Unitary operator) An operator is *unitary* if its Hermitian conjugate is its inverse:

$$\hat{A}^\dagger = \hat{A}^{-1} \quad \text{or} \quad \hat{A}^\dagger\hat{A} = \mathbb{I} \tag{2.58}$$

Theorem 2.7 (Unitary operator properties) Unitary operators have the following properties:

- **Perservation of the norm:**
$$||\hat{U}|\psi\rangle|| = |||\psi\rangle|| \tag{2.59}$$
- **Preservation of the inner product between vector pairs:**
$$\left(\hat{U}\,|\phi\rangle\,, \hat{U}\,|\psi\rangle\right) = \langle\phi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\phi|\psi\rangle \tag{2.60}$$
- Unitary operators have eigenvalues of the form $e^{i\theta}$.

Theorem 2.8 (Antiunitary operator properties) An *antiunitary operator* $\hat{A}$ is defined by the following two properties:

- **Anti-linearity:** For complex numbers $a, b \in \mathbb{C}$ and states $|\psi\rangle, |\phi\rangle$:
$$\hat{A}\big(a|\psi\rangle + b|\phi\rangle\big) = a^*\hat{A}|\psi\rangle + b^*\hat{A}|\phi\rangle \tag{2.61}$$
where a^* and b^* are the complex conjugates of a and b.
- **Preservation of the inner product up to complex conjugation:** For states $|\psi\rangle$ and $|\phi\rangle$:
$$\langle\psi|\phi\rangle = \langle\hat{A}\phi|\hat{A}\psi\rangle^* \tag{2.62}$$

Remark 2.17 The most common example of an antiunitary operator is the time-reversal operator $\hat{T}$. It can then be shown that the Hermitian conjugate of a Hermitian operator is

$$\hat{A} = \sum_{j,k} A_{ij} |\phi_j\rangle\langle\phi_k| = \sum_j \lambda_j |\chi_j\rangle\langle\chi_j| \tag{2.63}$$

where λ_j are the eigenvalues of $\hat{A}$. For the Hamiltonian, this is intuitively

Postulate 2

$$\hat{H} = \sum_j E_j |\chi_j\rangle\langle\chi_j| \tag{2.64}$$

> **Quote 2.6** The Hamiltonian reigns supreme as the king of all operators (...) because it describes dynamics.
>
> *Tania Monteiro, 1 October 2024*

2.8 Projector

Definition 2.27 (Projector) Mathematically, a *projector* $\hat{P}$ is a linear operator that satisfies

$$\hat{P}^2 = \hat{P} \tag{2.65}$$

Functionally, it maps a quantity in a vector space V into a subspace $W \subset V$.

Derivation 2.6 (Projection from 3D space to 2D space) A simple example of a projector is $\mathrm{diag}(1,1,0)$, which performs the following projection

$$\mathrm{diag}(1,1,0)\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \tag{2.66}$$

from which we can cover its mathematical significance. The satisfaction of $\hat{P}^2 = \hat{P}$ is also observed.

Derivation 2.7 (Quantum mechanical projectors) In QM, the concept is identical. Noting that the Hilbert space is constructed by a series of orthonormal basis vectors $|\phi_j\rangle$, we can infer that projectors project quantities from the full Hilbert space with all orthonormal bases to a subspace with fewer orthonormal bases. Hence, QM projectors are generally constructed by the spectral decomposition

$$\hat{P} = \sum_j^r |\phi_j\rangle\langle\phi_j| \tag{2.67}$$

r is the so-called *projector rank*, which will be completely trivial for anyone who knows what a tensorial rank is.

> **Exercise 2.4** Using the spectral decomposition form of the projector and the orthonormality condition $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, verify that Equation 2.67 is a projector by proving $\hat{P}^2 = \hat{P}$.

Remark 2.18 When one applies the decomposed projector (as seen in Equation 2.67) to a vector $|v\rangle$, only the components of $|v\rangle$ that lie within the subspace spanned by $\{|\phi_j\rangle\}$ are preserved.
Using the projector, one can rewrite the spectral composition of a Hermitian operator is

$$\hat{A} = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j| = \sum_j \lambda_j \hat{P}_j \tag{2.68}$$

For a non-degenerate λ_j

$$\hat{P}_j = |\phi_j\rangle\langle\phi_j| \tag{2.69}$$

For a degenerate λ_j, the target subspace of the projector span over more than one basis:

$$\hat{P}_j = \sum_m |\phi_m\rangle\langle\phi_m| \tag{2.70}$$

for all ms corresponding to the eigenvalue λ_j.

2.9 Time evolution

Postulate 3 Wave functions evolve with time according to the time-dependent Schrödinger equation (TDSE)

$$i\hbar\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle \tag{2.71}$$

A general solution of the TDSE in wavefunction form is

$$\psi(x,t) = \sum_n c_n\phi_n(x)e^{-iE_nt/\hbar} \tag{2.72}$$

Definition 2.28 (Time evolution operator) The *time evolution operator* is a unitary operator defined as

$$\hat{U}_\chi = \sum_j e^{-iE_jt/\hbar}|\chi_j\rangle\langle\chi_j| \tag{2.73}$$

It describes the time evolution of state vectors. For example, for an arbitrary state $|\psi\rangle$:

$$|\psi(t_2)\rangle = \hat{U}(t_2,t_1)|\psi(t_1)\rangle \tag{2.74}$$

2.10 Measurement of discrete systems

Postulate 4 Associated with every observable quantity is a Hermitian operator $\hat{M}$ with a spectral decomposition $\hat{M} = \sum_j \lambda_j P_j$, The eigenvalues λ_j label the possible outcomes of the measurement. $P_j = |\psi_j\rangle\langle\psi_j|$, where $|\psi_j\rangle$ are eigenstates of $\hat{M}$. When the measurement is performed on a system in state $|\Psi\rangle$, the probability that eigenvalue λ_j is returned is $\langle\Psi|P_j|\Psi\rangle$. After the measurement, the state of the system becomes

$$|\Psi'\rangle = \frac{P_j|\Psi\rangle}{\sqrt{\langle\Psi|P_j|\Psi\rangle}} \tag{2.75}$$

We now look at two examples of this.

Definition 2.29 (Expetation value) For each obserable $\hat{X}$ in some quantum state Ψ, there is an *expectation value* given by

$$\langle\hat{X}\rangle = \langle\Psi|\hat{X}|\Psi\rangle \tag{2.76}$$

Using spectral decomposition, one finds

$$\langle\hat{X}\rangle = \sum_j \lambda_j|\langle\phi_j|\Psi\rangle|^2 \tag{2.77}$$

where we can recover the *probability amplitude* $p = |\langle\phi_j|\Psi\rangle|^2$.

Definition 2.30 (Global phase) Two states $|\Psi\rangle$ and $|\phi\rangle$ are said to differ only by a so-called *global phase* if

$$|\phi\rangle = e^{i\theta}|\Psi\rangle \tag{2.78}$$

Remark 2.19 There are no measurable differences between the states $|\Psi\rangle$ and $|\phi\rangle$. One can prove this by considering

$$\langle\phi|\hat{X}|\phi\rangle = e^{-i\theta}e^{i\theta}\langle\Psi|\hat{X}|\Psi\rangle = \langle\Psi|\hat{X}|\Psi\rangle \tag{2.79}$$

2.11 Density operator

We consider two possible states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ (which may or may not be orthogonal) with probabilities p_1 and $p_2 = 1 - p_1$. The expectation is

$$\langle \hat{O} \rangle = p_1 \langle \Psi_1 | \hat{O} | \Psi_1 \rangle + p_2 \langle \Psi_2 | \hat{O} | \Psi_2 \rangle \tag{2.80}$$

Definition 2.31 (Pure and mixed states) We distinguish pure and mixed states:

- *Pure states* can be described by a single state vector.
- *Mixed states* need to be described by more than one state vector.

Remark 2.20 Our motivating example is thus a mixed state. In real life, pure states are very uncommon due to noise.

Definition 2.32 (Maximally mixed state) A *maximally mixed state* is equally likely to be in any one of the states in its Hilbert space. Hence, it has only one probability $p_j = p$. The density matrix (as we will see almost immediately) of a N-dimensional maximally mixed state is

$$\rho = \frac{1}{N} \mathbb{I} \tag{2.81}$$

Remark 2.21 A maximally mixed state is so-called because it represents the most uncertain or random distribution over all possible quantum states, and can be interpreted as having no information about which specific state the system is in. The classical analogue of this is a uniform probability distribution.

> **Quote 2.7** Were it not the case you might well already be shopping for quantum computers in PC World.
>
> *Tania Monteiro, on the lack of pure states, in her lecture notes*

One can consider mixed states to be a combination of pure states with various probabilities. This can be better seen by introducing the so-called *density operator*.

Definition 2.33 (Density operator) For a mixed state, the so-called density operator is given by

$$\hat{\rho} = \sum_j p_j |\Psi_j\rangle\langle\Psi_j| \tag{2.82}$$

where $|\Psi_j\rangle$ are all possible states whose corresponding probabilities are p_j. For a pure state, this reduces to

$$\hat{\rho} = |\Psi\rangle\langle\Psi| \tag{2.83}$$

In this sense, each pure state component $|\Psi_j\rangle\langle\Psi_j|$ in a mixed state is the contribution of that pure state $|\Psi_j\rangle$ to the mixed state. The density operator represents the mixed state - more intuitively (and perhaps less carefully), one can also say that the density operator *is* the mixed state.

Theorem 2.9 (Density operator properties) The density operator has the following properties:

- $\hat{\rho}$ is Hermitian:
$$\hat{\rho} = \hat{\rho}^\dagger \tag{2.84}$$
- $\hat{\rho}$ has unit trace:
$$\mathrm{Tr}(\hat{\rho}) = 1 \tag{2.85}$$
- $\hat{\rho}$ has a positive expectation value:
$$\langle \hat{\rho} \rangle > 0 \tag{2.86}$$
This is because the probabilities always observe $p_j > 0$.

Definition 2.34 (Trace) As can be recalled, the *trace* is

$$\mathrm{Tr}(\hat{A}) = \sum_j A_{jj} = \sum_j \langle \phi_j | \hat{A} | \phi_j \rangle \tag{2.87}$$

Theorem 2.10 (Trace properties) The trace has the following properties:

- The trace is linear.
- The trace is cyclic invariant:
$$\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A}) \tag{2.88}$$
- The trace is basis independent:
$$\mathrm{Tr}_\psi(\hat{A}) = \mathrm{Tr}_\phi(\hat{A}) \tag{2.89}$$

As a pure state is a single projector, we know that the density operator of a pure state observes

$$\hat{\rho} = \hat{\rho}^2 \tag{2.90}$$

Definition 2.35 (Purity) We then want to know how pure a given mixed state is. For this, the *purity* $\mathcal{P}$ is defined as

$$\mathcal{P} = \mathrm{Tr}(\hat{\rho}^2) \tag{2.91}$$

For a pure state, $\mathcal{P}$ is expectedly 1.

Finally, the density operator undergoes time evolution as one would expect of an operator:

$$\rho(t) = U(t)\rho(0)U^\dagger(t) \tag{2.92}$$

2.11.1 Symmetries

As should be well understood at this point, a *symmetry* is some transformation to a system that makes no change in its physical qualities. We will now look at three types of symmetries, from the more to the less general:

Derivation 2.8 (Wigner's theorem) The most general transformation is elucidated by the so-called *Wigner's theorem*[a]:

Theorem 2.11 (Wigner's theorem) For any transformation $U : |\Psi\rangle \to U|\Psi\rangle$ that preserves inner products, i.e., for some transformation U that observes:

$$\langle \Psi | A | \Psi \rangle \to \langle \Psi | U^\dagger U A U^\dagger U | \Psi \rangle = \langle \Psi | A | \Psi \rangle \tag{2.93}$$

where A is an observable, one can always represent U with a unitary or antiunitary operator, which then behaves as

$$A \to UAU^\dagger \tag{2.94}$$

[a] The actual theorem is far more complex, but we will only be concerned with its manifestation in quantum mechanics.

Derivation 2.9 (Symmetry of concrete systems) Symmetries of concrete[a] systems are associated with specific Hamiltonian operators.

Theorem 2.12 (Symmetry of concrete systems) n observable A is a symmetry of a system with Hamiltonian $\hat{H}$ if

$$[A, \hat{H}] = 0 \tag{2.95}$$

A is then said to be the generator of a corresponding unitary operator U_A[b]:

$$U_A(s) = e^{isA} \tag{2.96}$$

where s is a parameter that represents the extent or 'magnitude' of the transformation. It is analogous to the a in 2.100, but less well-defined.

Theorem 2.13 (Symmetry properties)

- In all dynamical pictures[a], we have[b]

$$\frac{dA_H}{dt} = \frac{1}{i\hbar}\left[A_H, \hat{H}\right] = 0 \tag{2.97}$$

- If $\Psi_s(t)$ is a solution to the Schrödinger equation, $U_A(s)\Psi_s(t)$ is also a solution.
- If $\Psi_s(0)$ is an eigenstate of A, $\Psi_s(t)$ is also an eigenstate.
- The linear combination of two symmetries A_1 and A_2 and their commutator are also symmetries.

[a] You will see what this means much later.
[b] This can be derived from Equation 5.27 and Equation 2.95.

[a] Here, 'concrete' simply means that the system is physical and specified with a well-defined Hamiltonian that governs its dynamics, not abstract or hypothetical.
[b] For a more detailed discussion, see *Spinors & Symmetries*.

As mentioned previously, the parameter s is poorly defined. Let us then look at a more specific example of a symmetry of a concrete system:

Derivation 2.10 (Coordinate transformations) For each coordinate transformation, which acts on the position x

$$F(x) = x' \quad x = F^{-1}(x') \tag{2.98}$$

one has a corresponding unitary operator U_F which acts on the wavefunction $\Psi(x)$.

$$U_F\Psi(x') = \Psi(x) \quad U_F\Psi(x) = \Psi(F^{-1}(x')) \tag{2.99}$$

In the case of spatial translations, we can derive, through some effort (we will not do this here), that the momentum operators are then the generators of the spatial translation operators

$$U_T(a) = \exp\left(-\frac{ia\hat{p}}{\hbar}\right) \tag{2.100}$$

for a corresponding spatial translation $T(x)$ by a

$$T(x) = x' = x + a \tag{2.101}$$

Chapter 3

Composite systems

3.1 Tensor product

Typical composite systems involve multiple particles or degrees of freedom. Suppose we have two systems represented by vector spaces A and B with orthonormal bases $|a_i\rangle$ and $|b_j\rangle$ and dimensions d_A and d_B. To satisfy Postulate 1, the combined space must also be a vector space. The states of the composite system's vector space C are therefore all the possible combinations of $|a_i\rangle$ and $|b_j\rangle$.
The mathematical operation used in generating the states of the composite system is known as the *tensor product*.

Derivation 3.1 (A recap) At this point, you have probably already seen composite systems in undergrad QM courses. Likewise, we have seen hidden tensor products already. For example, consider two spin-$\frac{1}{2}$ particles with the following triplet state:

$$|1,1\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle \left|\frac{1}{2},\frac{1}{2}\right\rangle \tag{3.1}$$

We note that there is no way we can multiply two ket products, so $\left|\frac{1}{2}\frac{1}{2}\right\rangle \left|\frac{1}{2}\frac{1}{2}\right\rangle$ is, in fact $\left|\frac{1}{2}\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\frac{1}{2}\right\rangle$. The $\otimes$ symbol is omitted as it is clear that this can only be a tensor product.
Remark 3.1 This example hides *a second* shorthand: the state $|1,1\rangle$ and $\left|\frac{1}{2},\frac{1}{2}\right\rangle$ themselves are tensor products. Their longhand forms are $|1\rangle \otimes |1\rangle$ and $\left|\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}\right\rangle$.

Note 3.1 Notably, $|jm\rangle$ and $|lm_l\rangle$ (or $|ls\rangle$) are *not* shorthands for tensor products.

Once again, the significance of the tensor product could be better seen in index notation we know and love from GR. In index notation, the outer product is

$$C_{ij} = A_i B_j \tag{3.2}$$

It then follows that

Definition 3.1 (Tensor product) The tensor product is the higher-ranking equivalent of the vector product

$$C_{ijkl} = A_{ij} \otimes B_{kl} \tag{3.3}$$

Remark 3.2 In practice, when the tensor product is obivous, which is almost always the case, the $\otimes$ symbol is often omitted, leaving $C_{ijkl} = A_{ij}B_{kl}$.
A numerical example is the tensor product of two matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & 0\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ 0\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & -1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{3.4}$$

Definition 3.2 (Composite vector space) The composite space C discussed earlier is then given by

$$C = A \otimes B \tag{3.5}$$

The dimension of C is $d_C = d_B d_A$. The same pattern applies in higher dimensions.

Remark 3.3 An implication is that one 2D harmonic oscillator $|n_3\rangle$ is mathematically equivalent to two 1D harmonic oscillators $|n_1\rangle \otimes |n_2\rangle$, even though they have profound physical differences.

Theorem 3.1 (Tensor product properties) The tensor product has the following properties:

- The tensor product is bilinear. Given $u_1, u_2 \in A$ and $v_1, v_2 \in B$:

$$(u_1 \otimes v_1)(u_2 \otimes v_2) = (u_1 u_2) \otimes (v_1 v_2) \tag{3.6}$$

$$(u_1 + u_2) \otimes v = u_1 \otimes v + u_2 \otimes v \tag{3.7}$$

$$u \otimes (v_1 + v_2) = u \otimes v_1 + u \otimes v_2 \tag{3.8}$$

$$(c \cdot u) \otimes v = u \otimes (c \cdot v) = c \cdot (u \otimes v) \tag{3.9}$$

  Only c is a scalar.

- The tensor product is associative:

$$(A \otimes B) \otimes C \cong A \otimes (B \otimes C) \tag{3.10}$$

  One can thus omit parentheses when writing tensor products of multiple spaces, i.e. $A \otimes B \otimes C$.

- The tensor product is distributive with respect to direct sums of vector spaces. Given vector spaces V_1, V_2, W_1 and W_2:

$$(V_1 \oplus V_2) \otimes (W_1 \oplus W_2) \cong (V_1 \otimes W_1) \oplus (V_1 \otimes W_2) \oplus (V_2 \otimes W_1) \oplus (V_2 \otimes W_2) \tag{3.11}$$

- The tensor product is not commutative. i.e. $A \otimes B$ is not isomorphic to $B \otimes A$.

Note 3.2 The so-called *direct sum* is typically not represented directly by the Einstein summation convention. This is because it corresponds to a 'concatenation' of index ranges, i.e. the direct sum $A_i \otimes B_j$ with $i = 0, \cdots, 3$ and $j = 0, \cdots, 3$ is C_k with $k = 0, \cdots, 7$, with $C_i = A_i$ and $C_{j+4} = B_j$. *'Concerning comment' courtesy of Felix Halbwedl, Technische Universität Graz.*

While this may seem tame at the moment, one should note that the practical use of the tensor product represents an increase in conceptual sophistication. We are now forced to migrate this definition to the bra-ket notation, and in doing so, we make a few comments about tensor products in QM.

Derivation 3.2 (Vector and operators) An arbitrary vector in C can be represented as

$$|\Psi\rangle = \sum_{i=1}^{d_a} \sum_{j=1}^{d_b} \gamma_{ij} |a_i\rangle \otimes |b_j\rangle \tag{3.12}$$

where γ_{ij} is the product of the (presumably known) coefficients α_i and β_j in vector spaces A and B. Here, $|\Psi\rangle$ is a *composite state*.
For composite operators in C, say, a composite of P in A and Q in B, the same idea follows:

$$P \otimes Q|\Psi\rangle = \sum_{j=1}^{d_a} \sum_{k=1}^{d_b} c_{jk} \left(P |a_j\rangle\right) \otimes \left(Q |b_k\rangle\right) \tag{3.13}$$

Note 3.3 Composite operators that effectively act as the identity on all but one subspace can be written with the following shorthand

$$P_A = P \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \cdots \quad P_B = \mathbb{I} \otimes P \otimes \mathbb{I} \otimes \mathbb{I} \cdots \quad \text{and so on.} \tag{3.14}$$

The subscript denotes the subspace which the operator actually acts on.

Definition 3.3 (Product state and entangled state) We distinguish between *product states* and *entangled states*:

- A composite state is a *product state* if it can be written as a tensor product of individual states. Consider two spin-$\frac{1}{2}$ particles A and B:

$$|\psi_1\rangle = |\uparrow_A\rangle\,|\uparrow_B\rangle \quad |\psi_2\rangle = |\downarrow_A\rangle\,|\uparrow_B\rangle \tag{3.15}$$

  A product state emerging from the two would be

$$|\psi_3\rangle = \frac{1}{2}\left(|\uparrow_A\rangle\,|\uparrow_B\rangle + |\uparrow_A\rangle\,|\downarrow_B\rangle + |\downarrow_A\rangle\,|\uparrow_B\rangle + |\downarrow_A\rangle\,|\downarrow_B\rangle\right) = \frac{1}{2}\left(|\uparrow_A\rangle + |\downarrow_A\rangle\right)\left(|\uparrow_B\rangle + |\downarrow_B\rangle\right) \tag{3.16}$$

- An *entangled state* is any composite state that is not a product state. An example is

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_A\rangle \otimes |\uparrow_B\rangle + |\downarrow_A\rangle \otimes |\downarrow_B\rangle\right) \tag{3.17}$$

  which is impossible to produce in a product state, under which there will inevitably be $|\uparrow_A\rangle \otimes |\downarrow_B\rangle$ and $|\uparrow_B\rangle \otimes |\downarrow_A\rangle$ terms.

Remark 3.4 Entangled states are so-called due to the correlation of measurement outcomes of different systems/subspaces. Consider the previous example. If particle A is spin-up/spin-down, particle B is invariably the same.

3.2 Deletion of systems

Now we consider the collapse of an entangled state under a measurement. Special relativity is preserved, and no information should be transmitted instantaneously. This is accomplished via the so-called *reduced state*.

Definition 3.4 (Density operator in composite systems) Consider again vector spaces A and B with orthonormal bases $|a_i\rangle$ and $|b_j\rangle$. The general density operator, as seen in Definition 2.33 is

$$\hat{\rho} = \sum_{i,j}\sum_{i',j'} \rho_{ij,i'j'} |a_i, b_j\rangle\langle a_{i'}, b_{j'}| \tag{3.18}$$

where each matrix element is

$$\rho_{ij,i'j'} = \langle a_i, b_j|\hat{\rho}|a_{i'}, b_{j'}\rangle \tag{3.19}$$

We apply the trace to a composite density matrix:

$$\mathrm{Tr}(\hat{\rho}) = \sum_{i,j} \rho_{ij,ij} \tag{3.20}$$

where one intuitively has the (diagonal) matrix elements

$$\rho_{ij,ij} = \langle a_i, b_j|\hat{\rho}|a_{i'}, b_{j'}\rangle \delta_{ii'}\delta_{jj'} \tag{3.21}$$

Now the concept of a *partial trace* is introduced, which allows the 'deletion' of a system[1] from the composite system, yielding the reduced state of the remaining system. We delete vector space A as an example:

[1] Perhaps because this system is inaccessible, among other reasons.

Definition 3.5 (Partial trace) The partial trace of some system A deletes A from the composite system C:

$$\mathrm{Tr}_A(\hat{\rho}) = \sum_i \rho_{ij,ij'} |b_j\rangle\langle b_{j'}| = \hat{\rho}_B \tag{3.22}$$

Summing over i removes the system A, which is said to be *traced out*.

> **Quote 3.1** Someone has stolen state A and taken it to the other room.
>
> *Tania Monteiro, 15 October 2024*

The expectation of some operator $\hat{O}_B$ (i.e. a measurement of B) then follows the standard form $\langle \hat{O}_B \rangle = \mathrm{Tr}(\hat{O}_B \hat{\rho}_B)$.

Remark 3.5 To delete B, we simply do the reverse:

$$\mathrm{Tr}_B(\hat{\rho}) = \sum_i \rho_{ij,i'j} |a_i\rangle\langle a_{i'}| = \hat{\rho}_A \tag{3.23}$$

Remark 3.6 The reduced density operator $\hat{\rho}_B$ expectedly describes the system B only. Importantly, $\hat{\rho}_B$ does not change when A is measured. i.e. no observable change of B that would indicate that A is measured happens. Any change in B is not observable without information about the measurement outcome of A. Hence special relativity is upheld.

Chapter 4

Time-dependent perturbation theory

4.1 Perturbative expansion

Derivation 4.1 (Simplest TDSE Hamiltonian) At this point, the time-dependent Schrödinger equation (TDSE) should be well-known:

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle \tag{4.1}$$

Suppose the Hamiltonian is time-invariant and has the known decomposition into

$$\hat{H} = \sum_k E_k|\phi_k\rangle\langle\phi_k| \tag{4.2}$$

Inserting this and the spectral decomposition $|\psi(t)\rangle = \sum_k c_k(t)|\phi_k\rangle$ into Equation 4.1 and integratting the result gives a first-order DE, which means that we have the general form of $c_k(t)$:

$$c_k(t) = c_k(0)\exp\left(\frac{-iE_kt}{\hbar}\right) \tag{4.3}$$

Substituting this into the decomposition of $|\psi(t)\rangle$, and we find

$$|\psi(t)\rangle = \sum_k \exp\left(\frac{-iE_kt}{\hbar}\right)c_k|\phi_k\rangle \tag{4.4}$$

which has an expectation value of

$$\langle\psi(t)|\hat{H}|\psi(t)\rangle = \sum_k E_k|c_k|^2 \tag{4.5}$$

In most cases, the Hamiltonian is time-dependent, has an unknown decomposition or satisfy both, and one must use perturbation theory to solve the problem. At this point, we have already seen *time-independent perturbation theory*. Now we investigate *time-dependent perturbation theory*.

Definition 4.1 (Perturbed Hamiltonian) The simplest form of the Hamiltonian used in perturbation theory is

$$\hat{H}(t) = \hat{H}_0 + \lambda\hat{V}(t) \tag{4.6}$$

where $\hat{H}_0$ is time-invariant and has a known decomposition. $\hat{V}(t)$ is the time-dependent perturbation term[a].

[a] Despite its notation, it has nothing to do with potential energy.

When $\hat{V}(t) = 0$, the scenario reduces to the previous case where the TDSE is directly integrable. When $\hat{V}(t)$ is *smol* but non-zero, one can use time-dependent perturbation theory.

Derivation 4.2 (Dirac variation-of-constants method) We start by giving ourselves a little leeway and allowing a time-variant $c_j \to c_j(t)$ in the eigenstates of $\hat{H}_0$:

$$|\psi(t)\rangle = \sum_j \exp\left(\frac{-iE_j t}{\hbar}\right) c_j(t)|\phi_j\rangle \tag{4.7}$$

Plugging this into the TDSE yields

$$i\hbar\partial_t \sum_j c_j(t) \exp\left(\frac{-iE_j t}{\hbar}\right)|\phi_j\rangle = (\hat{H}_0 + \lambda\hat{V}(t))|\psi(t)\rangle \tag{4.8}$$

Differentiating the LHS and noting that $\hat{H}_0|\phi_j\rangle = E_k|\phi_j\rangle$ gives

$$i\hbar \sum_j \dot{c}_j(t) \exp\left(\frac{-iE_j t}{\hbar}\right)|\phi_j\rangle = \lambda\hat{V}(t)\sum_j c_j(t)E_j \exp\left(\frac{-iE_j t}{\hbar}\right)|\phi_j\rangle \tag{4.9}$$

multiplying both sides by a $\langle\phi_m|$ and using $\langle\phi_m|\phi_k\rangle = \delta_{mk}$, we have

Theorem 4.1 (TDSE equivalent)

$$\dot{c}_m(t) = \frac{\lambda}{i\hbar}\sum_k c_k(t)\exp(i\omega_{mk}t)V_{mk}(t) \tag{4.10}$$

where:

- $\omega_{mk} = (E_m - E_k)/\hbar$ is the *Bohr angular frequency.*
- $V_{mk}(t) = \langle\phi_m|\hat{V}(t)|\phi_k\rangle$ is the matrix for the perturbation term.

As solving $c_m(t)$ yields $|\psi(t)\rangle$, this effectively solves the TDSE.

Previously in time-independent perturbation theory, we expanded eigenvalues E_j and eigenfunctions $|\psi_n\rangle$. Here, the idea is the same, but we expand the terms $c_m(t)$ instead:

$$c_m(t) = \sum_j \lambda^i c_m^{(j)}(t) \tag{4.11}$$

where each term is related to its previous term by

$$\dot{c}^{(n)}(t) = \frac{1}{i\hbar}\sum_k \exp(i\omega_{mk}t)V_{mk}(t)c_k^{(n-1)}(t) \tag{4.12}$$

Note 4.1 If λ is *smol*, then one only needs to calculate $c^{(1)}(t)$. However, this expansion converges as regardless of λ, the expanded terms are *smol* from the third order on. Usually calculating the first- and second-order terms provides very good approximations.

For convenience, we assume that the perturbation $\hat{V}(t)$ only emerges at $t = 0$. Therefore, the zeroth term is simply $c_m^{(0)} = \delta_{mj}$ which will inevitably be fixed to 1.

4.2 First and second orders

By plugging in the zeroth term to the differential equation, one finds that

Theorem 4.2 (First-order term)

$$\dot{c}_m^{(1)} = \frac{1}{i\hbar}\exp(i\omega_{mj}t)V_{mj}(t) \tag{4.13}$$

We integrate between $t = 0$ and some arbitrary $t = \tau$

- For $m = j$, this reduces to
$$c_j^{(1)}(\tau) = \frac{1}{i\hbar}\int_0^\tau V_{jj}(t)dt \tag{4.14}$$
- For $m \neq j$
$$c_m^{(1)}(\tau) = \frac{1}{i\hbar}\int_0^\tau V_{mj}(t)\exp(i\omega_{mj}t)dt \tag{4.15}$$

If $\dot{c}_m^{(1)}$ for $m \neq j$ is non-zero, it means that the system might be found instate $|\phi_m\rangle$, having transitioned from $|\phi_j\rangle$.

Theorem 4.3 (Rabi's formula) This *transition probability* and the probability of a system remaining in its initial state are, respectively:
$$P_{j\to m}^{(1)}(\tau) = |\dot{c}_m^{(1)}|^2 \quad P_{j\to j}^{(1)}(\tau)1 - |\dot{c}_j^{(1)}|^2 \tag{4.16}$$

The second-order term has a double integral

Theorem 4.4 (Second-order term)
$$c_m^{(2)} = -\frac{1}{\hbar^2}\int_0^\tau dt' \int_0^{t'} \sum_k \exp(i\omega_{mk}t')\exp(i\omega_{kj}t)V_{mk}(t')V_{kj}(t) \tag{4.17}$$

4.3 (Almost) constant perturbation

Still we want to consider simpler cases. One such case is a perturbation that is constant except for the instant it is switched on. i.e.
$$\hat{H} = \begin{cases} \hat{H}_0 & t < 0 \\ \hat{H}_0 + \hat{V} & 0 \leq t \leq \tau \end{cases} \tag{4.18}$$

Remark 4.1 This is distinct from a potential barrier we have seen before as there is *one* variation - when it is suddenly turned on.

Derivation 4.3 ((Almost) constant perturbation) By integrating $\dot{c}_m^{(1)}(\tau) = \frac{1}{i\hbar}\int_0^\tau V_{mj}(t)\exp(i\omega_{mj}t)dt$, we get the first-order term
$$c_m^{(1)}(\tau) = \frac{V_{mj}}{i\hbar}\int_0^\tau (1 - \exp(i\omega_{mj}t)) \tag{4.19}$$

Using trig identities, the transition probability is thus
$$P_{j\to k}^{(1)}(\tau) = \frac{4|V_{mj}|^2\sin^2(\omega_{kj}\tau/2)}{\hbar^2\omega_{kj}^2} = \frac{2|V_{mj}|^2}{\hbar^2}F(\tau,\omega) \tag{4.20}$$

Here, we have defined the (sadly unnamed) function
$$F(\tau,\omega) = \frac{\tau^2}{2}\operatorname{sinc}^2(\omega\tau/2) \tag{4.21}$$

where we have the so-called *sinc function*
$$\operatorname{sinc} x = \frac{\sin x}{x} \tag{4.22}$$

For long perturbations, we can consider the steady state where $\tau \to \infty$. In this case we have the convergence
$$\lim_{\tau\to\infty} F(\tau,\omega) = \pi\tau\delta(\omega) \tag{4.23}$$

and with it
$$P_{j\to k}^{(1)}(\tau) = \frac{2|V_{mj}|^2}{\hbar^2}\pi\tau\delta(\omega_{jk}) \tag{4.24}$$

which is zero unless $\omega_{jk} = (E_j - E_k)/\hbar$ is zero. Hence only transitions between states of identical energy ($E_k = E_j$) are allowed.
We can derive the transition rate by dividing both sides by τ. This yields the transition rate $\gamma^{(1)}_{j\to k}(\tau)$:

Theorem 4.5 (Fermi's golden rule)

$$\gamma^{(1)}_{j\to k}(\tau) = \frac{2\pi}{\hbar}|V_{mj}|^2\delta(E_j - E_k) \tag{4.25}$$

This formula for the transition rate is called *Fermi's golden rule*.

Fun fact 4.1 Fermi's golden rule was actually derived by Dirac 20 years before Fermi did so himself. It is so-named by Fermi (*golden rule No. 2* in full) due to its perceived importance. There also exists a *golden rule No. 1*, a similar formula involving the indirect transition rate.
Remark 4.2 It is interesting to consider the physical significance of $F(\tau, \omega)$. It is a Dirac delta-like function that 'regulates' transitions between states.

4.4 Harmonic perturbation

Definition 4.2 (Harmonic perturbation) In a *Harmonic perturbation*, the operator $\hat{V}(t)$ is a sinusoidal function with respect to time.

Derivation 4.4 (Dipole model) The interaction between atoms and light can be well-approximated as a Harmonic perturbation. In specific, we consider an atom as a dipole in an electric field. The potential energy[a] is

$$V = -\vec{E}\cdot\vec{D} \tag{4.26}$$

We will investigate this scenario in depth later on. But now, it will suffice to simply consider the quantised version of this scenario, where the electric field $\vec{E}$ and the dipole $\vec{D}$ are operators $\hat{E}$ and $\hat{D}$. Assuming that this light is a monochromatic laser, we have a perfectly sinusoidal electric field

$$\hat{E}(t) = \hat{E}_0\cos(\omega t) \tag{4.27}$$

Quote 4.1 Who needs *rockets* when we have *lasers*?

Laser Comanche in Command & Conquer: Generals

Figure 4.1: Laser Comanche

The perturbation Hamiltonian becomes

$$\hat{V} = \hat{H}_{\text{int}}\cos(\omega t) \tag{4.28}$$

where $\hat{H}_{\text{int}}$ is an *interaction Hamiltonian* that we will investigate later. As it turns out, this is a special case which we can generalise as

$$\hat{V} = \hat{A}e^{i\omega t} + \hat{A}^\dagger e^{-i\omega t} \to \hat{H} = \hat{H}_0 + \hat{A}e^{i\omega t} + \hat{A}^\dagger e^{-i\omega t} \tag{4.29}$$

Remark 4.3 This reduces to the laser case when $\hat{H}_{\text{int}} = 2\hat{A}$.
Substituting this into Equation 4.14 yields

$$c^{(1)}_m(\tau) = \frac{A_{mj}}{\hbar}(2i\tau e^{i(\omega_{mj}+\omega)\tau/2})\,\text{sinc}((\omega_{mj}+\omega)\tau/2) + \frac{A^*_{mj}}{\hbar}(2i\tau e^{i(\omega_{mj}-\omega)\tau/2})\,\text{sinc}((\omega_{mj}-\omega)\tau/2) \tag{4.30}$$

In the steady state $\tau \ll 0$, the terms become Dirac delta functions $\delta(\omega_{mj} - \omega)$ and $\delta(\omega_{mj} + \omega)$ which implies that

$$\omega_{mj} = \pm\omega \rightarrow E_m = E_j \pm \hbar\omega \tag{4.31}$$

Hence we recover the quantisation of photons.
We can also derive Fermi's golden rule for this scenario:

$$\Gamma^{(1)}_{k\rightarrow j}(\tau) = \frac{2\pi}{\hbar}|A_{kj}|^2\delta(E_k - E_j - \hbar\omega) \tag{4.32}$$

[a] Again distinct from the perturbation $\hat{V}$.

Chapter 5

Unitary time evolution

5.1 Time evolution operator

In this chapter, we study the solutions of the TDSE. We recall the equation to be

$$i\hbar\frac{\partial|\phi(t)\rangle}{\partial t} = \hat{H}|\phi(t)\rangle \tag{5.1}$$

where we have the so-called *time evolution operator* $U(t)$

$$|\phi(t)\rangle = U(t)|\phi(0)\rangle \tag{5.2}$$

When the Hamiltonian is time-independent, $U(t)$ is expressed by

$$U(t) = \exp\left(-\frac{i\hat{H}t}{\hbar}\right) \tag{5.3}$$

Theorem 5.1 (Unitary time evolution operator) The time evolution operator is *unitary*.

Here the exponent contains an operator. We can Taylor expand e^x as

$$e^x = \sum_j^\infty \frac{x^j}{j!} \tag{5.4}$$

For operators, we simply replace the number x with some operator $\hat{A}$.

Note 5.1 (Diagonal matrices) We recall that the power of a diagonal matrix is easier to calculate. For power j

$$B = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix} \to B^j = \begin{pmatrix}\lambda_1^j & 0\\ 0 & \lambda_2^j\end{pmatrix} \tag{5.5}$$

where j is simply a summation index. We then have the exponential

$$e^B = \sum_j^\infty \begin{pmatrix}\lambda_1^j/j! & 0\\ 0 & \lambda_2^j/j!\end{pmatrix} = \begin{pmatrix}e^{\lambda_1} & 0\\ 0 & e^{\lambda_2}\end{pmatrix} \tag{5.6}$$

Theorem 5.2 (Distribution of exponentials) For numbers, we have

$$e^{a+b} = e^a e^b \tag{5.7}$$

The matrix analogy

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}} \tag{5.8}$$

is only true if $\hat{A}$ and $\hat{B}$ commute. i.e.

$$[\hat{A}\hat{B}] = 0 \tag{5.9}$$

We do, however, note that having operators on exponentials is usually not very helpful. Inspired by Note 5.1, we can derive the following cute trick:

Derivation 5.1 (A cute trick) As seen previously, one can always diagonalise a Hamiltonian in terms of its eigenvalues E_n and its eigenstates $|n\rangle$.

$$\hat{H} = \sum_n E_n |n\rangle\langle n| \tag{5.10}$$

Due to its diagonal nature, the Hamiltonian essentially acts as a scalar multiplication when applied to one of its eigenstates. The time evolution operator is then

$$U(t)|n\rangle = e^{-i\hat{H}t/\hbar}|n\rangle = e^{-iE_n t/\hbar}|n\rangle \tag{5.11}$$

We can then 'unpack' the full form of the evolution operator by adding a $|n\rangle\langle n|$

$$\hat{U}(t) = \sum_n |n\rangle\langle n| e^{-i\hat{H}t/\hbar} \tag{5.12}$$

which we can rewrite as

Theorem 5.3 (A cute trick) For a Hamiltonian $\hat{H} = \sum_n E_n|n\rangle\langle n|$, one can always represent its corresponding time evolution operator as

$$U(t) = e^{-iHt/\hbar} = \sum_n e^{-iE_n t/\hbar}|n\rangle\langle n| \tag{5.13}$$

Remark 5.1 If we start at a time t_1, one expectedly has

$$|\phi(t)\rangle = U(t - t_1)|\phi(t_1)\rangle \tag{5.14}$$

Derivation 5.2 (Step-wise changing Hamiltonian) A better way to put this is a Hamiltonian that is constant over various periods of time with discontinuous changes at their boundaries. We can consider the following Hamiltonian:

$$\hat{H} = \begin{cases} \hat{H}_1 & 0 \le t \le t_1 \\ \hat{H}_2 & t_1 \le t \le t_2 \end{cases} \tag{5.15}$$

One then yields the time evolution operators

$$U_1(t) = \exp\left(-\frac{i\hat{H}_1 t}{\hbar}\right) \quad U_2(t) = \exp\left(-\frac{i\hat{H}_2 t}{\hbar}\right) \tag{5.16}$$

The total (or effective) time evolution operator is this

$$U_{tot}(t) = \begin{cases} U_1(t) & 0 \le t \le t_1 \\ U_2(t - t_1)U_1(t_1) & t_1 \le t \le t_2 \end{cases} \tag{5.17}$$

Remark 5.2 Expectedly, the time evolution operators that come later in time are added to the left of those earlier in time[a].

[a] This comment is added at the advice of Felix Halbwedl (TU Graz).

Derivation 5.3 (Continuously changing Hamiltonian) The idea now is to approximate this Hamiltonian as a step-wise changing one. Only now the step is infinitely *smol*, almost like how we regarded integration as the sum of areas with infinitely *smol* x-intervals when we first studied it. Say we are at t. We take n points in time between 0 and t. Each interval is then t/n, with the Hamiltonian at that point being

$$H(jt/n) \quad j = 1, 2, \cdots, n \tag{5.18}$$

The resulting time evolution operator is then

$$U(t) \approx \prod_{j=1}^{n} \exp\left(-\frac{iH(jt/n)t}{\hbar}\right) \tag{5.19}$$

which we call a *Dyson series*. When $n \to \infty$ this becomes an integral:

Theorem 5.4 (Dyson's formula)

$$U(t) = \lim_{n\to\infty} \prod_{j=1}^{n} \exp\left(-\frac{iH(jt/n)t}{\hbar}\right) = \underbrace{T\left[\exp\left(-i\int_0^t dt' \frac{H(t')dt'}{\hbar}\right)\right]}_{①} \tag{5.20}$$

The result in ① form is called the *time-ordered exponential*, an alternative notation more commonly seen in QFT.

5.2 Suzuki-Trotter expansion

Sometimes we will encounter a two-term Hamiltonian whose terms $\hat{H}_1$ and $\hat{H}_2$ do not commute. In this case, calculating the time evolution operator will be immensely difficult. We can hence use the *Lie product formula* or the *Trotter product formula*:

Theorem 5.5 (Lie product formula) For any operators or square matrices $\hat{A}$ and $\hat{B}$, one has

$$e^{\hat{A}+\hat{B}} = \lim_{n\to\infty}\left(e^{\hat{A}/n}e^{\hat{B}/n}\right)^n = \lim_{n\to\infty}\left(e^{\hat{B}/n}e^{\hat{A}/n}\right)^n \tag{5.21}$$

where n is the so-called *Trotter number*.

For time evolution operators

$$U_1(t) = \exp\left(-\frac{i\hat{H}_1 t}{\hbar}\right) \quad U_2(t) = \exp\left(-\frac{i\hat{H}_2 t}{\hbar}\right) \tag{5.22}$$

this becomes

$$U(t) = e^{-\frac{it}{\hbar}(\hat{H}_1+\hat{H}_2)} = \lim_{n\to\infty}\left(U_1(t/n)U_2(t/n)\right)^n \tag{5.23}$$

When $n < \infty$, this is called the *Suzuki-Trotter expansion*.

Note 5.2 In general and for a given n, the Suzuki-Trotter expansion is valid to the n^{th} order in t and the error is of order t^{n+1}.

5.3 Schrödinger and Heisenberg pictures

Time evolution in quantum mechanics can be interpreted via more than one way, which we call *dynamical pictures*. We consider some operator $\hat{O}$ and its expectation as time evolves

$$\langle\hat{O}(t)\rangle = \langle\psi(0)|U^\dagger(t)\hat{O}U(t)|\psi(0)\rangle \tag{5.24}$$

There are two ways we can interpret this. The first is the *Schrödinger picture* which is the interpretation we have been familiar with up to this point.

Definition 5.1 (Schrödinger picture)

$$\langle\hat{O}(t)\rangle_S = \underbrace{{}_S\langle\psi(0)|U^\dagger(t)}_{\text{time evolution of state } \langle\psi|} \quad \hat{O}(0)_S \quad \underbrace{U(t)|\psi(0)\rangle_S}_{\text{time evolution of state } |\psi\rangle} \tag{5.25}$$

At $t = 0$, the *Heisenberg picture* is identical to the Schrödinger picture. It is

Definition 5.2 (Heisenberg picture)

$$\langle \hat{O}(t)\rangle_H = {}_H\langle \psi(0)| \underbrace{U^\dagger(t)\hat{O}(0)_H U(t)}_{\text{time evolution of operator } \hat{O}} |\psi(0)\rangle_H \tag{5.26}$$

Rather than the state, the operator is regarded to evolve with time. This interpretation is (barely) justified by two reasons:

- It can simplify calculations.
- It is more reminiscent of classical physics. This is more of a historical reason.

The equivalent of the Schrödinger equation in the Heisenberg picture is unimaginatively called the *Heisenberg equation*:

Theorem 5.6 (Heisenberg equation)

$$\partial_t \hat{O}(t)_H = \frac{i}{\hbar}\left[\hat{H}(t), \hat{O}(t)_H\right] \tag{5.27}$$

Remark 5.3 If $\left[\hat{H}(t), \hat{O}(t)_H\right] = 0$, $\hat{O}_H$ is constant in time. Hence the well-known result is recovered.

5.4 Interaction picture

In addition to the Heisenberg and Schrödinger pictures, there is an intermediate picture of time evolution called the *interaction picture*, which is used when the Hamiltonian is slightly perturbed:

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}} \tag{5.28}$$

where $\hat{H}_0$ is a solved Hamiltonian whose eigenstates and eigenenergies are known, and $\hat{H}_{\text{int}}$ represents extra terms arising from the interaction. We can define the evolution operators

$$U_0(t) = \exp\left(-\frac{i\hat{H}_0 t}{\hbar}\right) \quad U_{\text{int}}(t) = \exp\left(-\frac{i\hat{H}_{\text{int}} t}{\hbar}\right) \tag{5.29}$$

This results in the state evolution

$$|\psi(t)\rangle_I = U_{\text{int}}(t)|\psi(0)\rangle_I \tag{5.30}$$

To preserve the invariance of expectation values, the operator in the interaction picture is given by

$$\hat{O}_I(t) = U_0^\dagger(t)\hat{O}_I(0)U_0(t) \tag{5.31}$$

As such:

Definition 5.3 (Interaction picture)

$$\langle \hat{O}(t)\rangle_I = \underbrace{{}_I\langle\psi(0)|U_{\text{int}}^\dagger(t)}_{\text{time evolution of state } \langle\psi|} \underbrace{U_0^\dagger(t)\hat{O}(0)_I U_0(t)}_{\text{time evolution of operator } \hat{O}} \underbrace{U_{\text{int}}(t)|\psi(0)\rangle_I}_{\text{time evolution of state } |\psi\rangle} \tag{5.32}$$

Remark 5.4 Noting that $U(t) = U_0(t)U_{\text{int}}(t)$ and $U^\dagger(t) = U_{\text{int}}^\dagger(t)U_0^\dagger(t)$, we conclude that the interaction picture is equivalent to the Heisenberg and Schrödinger pictures.

Theorem 5.7 (Interaction picture Schrödinger equation) The interaction picture Schrödinger equation might look identical to the Schrödinger equation at first glance

$$i\hbar\partial_t|\psi(t)\rangle_I = \hat{H}_I|\psi(t)\rangle_I \tag{5.33}$$

but we have replaced the Hamiltonian $\hat{H}$ the so-called *interaction picture Hamiltonian*

$$\hat{H}_I = U_0^\dagger(t)VU_0(t) \tag{5.34}$$

Remark 5.5 In effect, we have eliminated the influence of $\hat{H}_0$ in the Schrödinger equation.

Chapter 6

Quantum light and atoms

6.1 Quantum harmonic oscillator

We now briefly review quantum harmonic oscillators (QHOs). A generic QHO has the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{6.1}$$

The energy eigenstates are $|n\rangle$ for an integer n. Each eigenstate's corresponding eigenenergy is then

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{6.2}$$

We can shift these energy eigenstates up and down using the so-called *ladder operators*:

Definition 6.1 (Ladder operators) We have the *lowering operator* or the *annihilation operator* $\hat{a}$ and the *raising operator* or the *creation operator* $\hat{a}^\dagger$

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} + i\hat{p}) \quad \hat{a}^\dagger = (\hat{a})^\dagger \tag{6.3}$$

which have the following effects:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad \hat{a}^\dagger = \sqrt{n+1}|n+1\rangle \tag{6.4}$$

Theorem 6.1 (Ladder operator commutation relations)

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{I} \tag{6.5}$$

Definition 6.2 (Number operator) The so-called *number operator* is then defined as

$$\hat{n} = \hat{a}\hat{a}^\dagger \quad \hat{n}|n\rangle = n|n\rangle \tag{6.6}$$

Derivation 6.1 (Hamiltonian) Using the ladder operators, one can represent the position and momentum operators as

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger + \hat{a}) \quad \hat{p} = i\sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger - \hat{a}) \tag{6.7}$$

Hence the Hamiltonian can be rewritten as

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{6.8}$$

6.2 Light

We recall Maxwell's equations:

- Source equations:
$$\nabla \cdot E = \frac{\rho}{\epsilon_0} \quad \nabla \times H = J_f + \partial_t D \tag{6.9}$$
where ρ is the charge density and $\mathbf{J}_f$ is the free current density.
- Structure equations:
$$\nabla \times E = -\partial_t B \nabla \cdot B = 0 \tag{6.10}$$

The generalised plane-wave E-field equation is hence

$$\left(\nabla^2 - \frac{1}{c^2}\partial_t^2\right) E = 0 \tag{6.11}$$

We note that E is a function of position and time $E(r,t)$. Separating variables by $E(r,t) = u(r)T(t)$ gives the two following equations:

Theorem 6.2 (Helmholtz equation)

$$\nabla^2 u(R) = -\frac{\omega^2}{c^2} u(r) \tag{6.12}$$

This is a vector equation.

Theorem 6.3 (Temporal equation)

$$\partial_t^2 T(t) = -\omega^2 T(t) \tag{6.13}$$

This equation admits solutions of the form $T(t) = e^{\pm i\omega t}$

In a QHO system we will have a range of frequencies ω_j with the associated solutions $u_j(r)$ and $T_j(t)$. Each solution, called a *mode* is then

$$E_j(r,t) = E_{0j}\alpha_j u_j(r) T_j(t) \tag{6.14}$$

where E_{0j} is a unit vector representing the polarisation of the mode (in effect, a *directional vector*) and α_j is a constant dependent on boundary conditions. The general solution then linearly sums over all the modes:

$$E(r,t) = \sum_j E_{0j}\alpha_j u_j(r) T_j(t) \tag{6.15}$$

This set of solutions can be quantised by replacing classical oscillators with quantum oscillators.

Definition 6.3 (Electric field operator)

$$\hat{E} = \sum_j u_j(r) E_{0j}(\hat{a}_j + \hat{a}_j^\dagger) \tag{6.16}$$

Definition 6.4 (Optical cavity) An *optical cavity* is a series of mirrors or other instruments that confines photons in a certain area (think lasers). In effect, it reduces the number of states to one: $|n\rangle$.

6.3 Light-atom interaction revisited

In quantum optics, we can approximate light-atom interaction with the *dipole approximation*, as previously seen in the hatmonic perturbation scenario. Classically, given the dipole $\vec{d} = q\vec{r}$ and the displacement $\vec{r}$ between the positive charge q and the negative charge $-q$, the classical interaction energy is

$$U = -\vec{D} \cdot \vec{E} \tag{6.17}$$

We have already quantised the electric field as $\hat{E}$, so we just need to quantised $\vec{D}$ as some dipole operator $\hat{D}$ and call it a day. For simplicity, we first assume the toy model of a *two-level atom*, in which the atom only have two states:

- The *ground state* $|g\rangle$.

- The sole *excted state* $|e\rangle$.

We therefore have the very simple Hamiltonian

Definition 6.5 (Two-level atom Hamiltonian)

$$\hat{H}_{2L} = 0|g\rangle\langle g| + \hbar\omega_a|e\rangle\langle e| = \hbar\omega_a|e\rangle\langle e| \tag{6.18}$$

where ω_a is the photon frequency during the transition of energy states.

Remark 6.1 The ground state term is zero as only energy differences are meaningful, and that we have chosen the ground state as is the reference point.

Derivation 6.2 (Dipole operator) We can then quantise the dipole moment and the displacement as operators $\hat{D}$ and $\hat{r}$. As we have only 2 states, the dimensionality is likewise 2. $\hat{D}$ therefore has 4 elements. The diagonal elements of $\hat{D}$ are intuitively

$$\langle g|\hat{D}|g\rangle = -e\langle g|\hat{r}|g\rangle \quad \langle e|\hat{D}|e\rangle = -e\langle e|\hat{r}|e\rangle \tag{6.19}$$

As the wavefunction is usually symmetric about the nucleus, the mean displacement is zero and both $\langle g|\hat{D}|g\rangle$ and $\langle e|\hat{D}|e\rangle$ vanish as a result.
Now we consider the off-diagonal elements. As there is diagonal symmetry we can say that

$$\langle g|\hat{D}|e\rangle = \langle e|\hat{D}|d\rangle = \vec{d} \tag{6.20}$$

where $\hat{d}$ is a real 3D vector dependent on the atomic states. Thus, the dipole operator is

$$\hat{D} = \vec{d}(|g\rangle\langle e| + |e\rangle\langle g|) \tag{6.21}$$

The Hamiltonian that describes the interaction between atom and light is known as the *dipole interaction Hamiltonian* or the *Jaynes-Cummings Hamiltonian*. We begin with the classical $U = -\vec{D}\cdot\vec{E}$ and quantise the three quantities. U is clearly the Hamiltonian, and we have previously quantised $\vec{D}$ and $\vec{E}$. Hence

Definition 6.6 (Jaynes-Cummings Hamiltonian)

$$\hat{H}_{JC} = \hbar g(r)(\hat{a} + \hat{a}^\dagger)(|g\rangle\langle e| + |e\rangle\langle g|) \tag{6.22}$$

where $\hbar g(r) = (-\vec{d}\cdot u(r)E_0)$ is the light-atom coupling strength.

The total Hamiltonian is simply the sum of the two-level atom and Jaynes-Cummings Hamiltonians:

Definition 6.7 (Total Hamiltonian)

$$\hat{H} = \hat{H}_{2L} + \hat{H}_{JC} = \hbar\omega_j(\hat{a}^\dagger\hat{a} + 1/2)\hbar\omega_a|e\rangle\langle e| + \hbar g(r)(\hat{a} + \hat{a}^\dagger)(|g\rangle\langle e| + |e\rangle\langle g|) \tag{6.23}$$

This can be simplified in the interaction picture, where one has $\hat{H}_0 = \hat{H}_{2L}$ and $\hat{H}_{\text{int}} = \hat{H}_{JC}$.

Chapter 7

Open quantum systems

7.1 Superoperators

It's a bird... it's a plane... it's superoperators!

Derivation 7.1 (Superoperator) As previously seen, unitary operators preserve purity. Thus, the evolution from a pure state to a mixed state must be non-unitary. Such a non-unitary transformation, however, can be constructed from a series of unitary transformations U_i with probabilities p_i. The following transforms the density operator of a pure state to that of a mixed state:

$$\hat{\rho}' = S[\hat{\rho}] = \sum_i p_i U_i \hat{\rho} U_i^\dagger \tag{7.1}$$

where S is the *superoperator*, a mathematical object that transforms one operator to another.
We know that such an evolution is non-unitary due to the presence of classical randomness. Indeed, we can forgo the probability amplitudes p_i altogether and absorb it into the unitary operators U_i. This yields the so-called *Kraus operators*:

Definition 7.1 (Kraus operator)

$$K_i = \sqrt{p_i} U_i \tag{7.2}$$

Although related to unitary operators via the previous relation, the Kraus operators themselves do not have to be unitary. In fact, the only other property they have to satisfy is the so-called *Kraus representation* or the *operator sum representation*:

Theorem 7.1 (Kraus representation) The Kraus operators satisfy

$$\sum_i K_i^\dagger K_i = \mathbb{I} \tag{7.3}$$

Definition 7.2 (Superoperator) In terms of the Kraus operators, the superoperator can hence be represented by

$$\hat{\rho}' = S[\hat{\rho}] = \sum_i K_i \hat{\rho} K_i^\dagger \tag{7.4}$$

Remark 7.1 This representation of $\hat{\rho}'$ in terms of K_i and $\hat{\rho}$ is known as the *Kraus decomposition*.

Theorem 7.2 (Preservation of density matrix properties) We know that ρ is Hermitian, has trace 1 and has non-negative eigenvalues (i.e. is positive semi-definite). The same properties also apply to $\hat{\rho}' = S[\hat{\rho}]$.

Theorem 7.3 (Complete positivity) For a composite system with 2 subsystems A and B. A superoperator where both $S[\hat{\rho}_A]$ and $S[\hat{\rho}_{AB}]$ positive semi-definite is said to be *completely positive*.

Remark 7.2 As surprising as it is, there are superoperators where $S[\hat{\rho}_A]$ is positive semi-definite, but $S[\hat{\rho}_{AB}]$ is not.
Remark 7.3 A superoperator S that is completely positive and preserves Hermicity and trace (i.e. all transformations in quantum evolution) can always be written in Kraus decomposition form.

7.2 Markovian evolution

As useful as it is, the Schrödinger equation restricts itself to isolated quantum systems. In nature, however, a quantum system usually instead interacts with other systems around it. We call such a system an *open quantum system*, and its surrounding systems its *environment*. Some comments can be made here:

- Even though an open quantum system includes the environment, our system of interest remains the small quantum system which the environment surrounds.
- It is possible to investigate the Schrödinger for the whole system, but as the environment is *large*, this is impractical.
- There is then a failure by the Schrödinger equation to describe open quantum systems.

What would then be convenient is to find an evolution equation, analogous to the Schrödinger equation, for a reduced state that covers purely our small system of interest $\hat{\rho}$. This would be of the form

$$\partial_t \hat{\rho} = S[\hat{\rho}(t)] \tag{7.5}$$

where $\partial_t \hat{\rho} = \lim_{\delta t \to 0} \frac{\hat{\rho}(t+\delta t))}{\delta t}$ and $\hat{\rho}(t+\delta t) = \hat{\rho}(t) + \partial_t \hat{\rho}(t)\delta t$. To get such an evolution equation, we see that the evolution has to be Markovian:

Definition 7.3 (Markovian evolution) The evolution of some operator $\hat{\rho}(t+\delta t)$ is Markovian if $\hat{\rho}(t+\delta t)$ must depend only on $\hat{\rho}(t)$, the state of the system immediately before. In other words:

- $\hat{\rho}(t+\delta t)$ must not depend on the environment.
- $\hat{\rho}(t+\delta t)$ must not depend on the $\hat{\rho}(t')$ for $t' < t$, the state of the system *any time earlier than* immediately before.

Remark 7.4 The second point is important in that information must flow from the system we study to the environment and not vice versa.
Remark 7.5 For example, 'Snakes and Ladders' is clearly Markovian. This is intuitive when one considers the game. At the same time, the weather is non-Markovian. The weather one year ago can continue to affect the weather tomorrow.

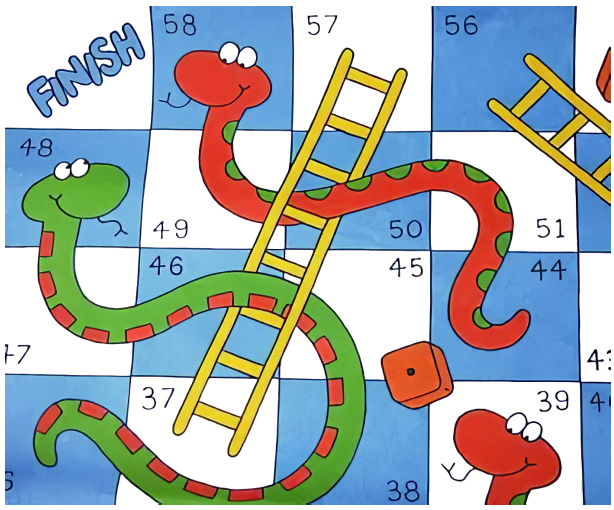


Figure 7.1: A game of 'Snakes and Ladders'.

Now we introduce the so-called *quantum master equation*.

> **Quote 7.1** But we think that everything, from the frantic dance of subatomic particles to the majestic swirl of galaxies, should be explained by just one grand physical principle, one master equation. If we can find that equation, how the universe really works at every time and place will at last be revealed.
>
> *Brian Greene, in* The Elegant Universe*, 2003*

Unfortunately, this is not *that* master equation. However, it is pretty cool too. The quantum master equation generalises the idea of a *master equation*[1] that describes the time evolution of a system.

Derivation 7.2 (GKSL equation) To start with, we assume that evolution must be Markovian. i.e. the quantum master equation must be of the form

$$\partial_t \hat{\rho} = S[\hat{\rho}(t)] \tag{7.6}$$

We recall that

$$\hat{\rho}(t+\delta t) = \hat{\rho}(t) + \delta t \partial_t \hat{\rho}(t) \tag{7.7}$$

and that $O(\delta t^n) = 0$ for $n \geq 1$. Now we try to derive a general form of $\partial_t \hat{\rho}(t)$. For simplicity, we consider $t = 0$. Thus one has $\hat{\rho}(t) \to \hat{\rho}(0)$, $\hat{\rho}(t+\delta t) \to \hat{\rho}(\delta t)$. The Kraus decomposition becomes

$$\hat{\rho}(\delta t) = \hat{\rho}(0) + \delta t \partial_t \hat{\rho}(0) = \sum_i K_i(\delta t)\hat{\rho}(0)K_i^\dagger(\delta t) \tag{7.8}$$

From this, we see that one of the Kraus operators (which we label K_0) must satisfy, for some linear operator A,

$$K_0 = \mathbb{I} + \delta t A \tag{7.9}$$

The remaining Kraus operators K_i satisfy, for linear operators L_i,

$$K_i = \sqrt{\delta t} L_i \tag{7.10}$$

One can verify this by inserting Equations 7.9 and 7.10 into Equation 7.7. This gives

$$\hat{\rho}(\delta t) = \hat{\rho}(0) + \delta t \left(A\hat{\rho}(0) + \hat{\rho}(0)A^\dagger + \sum_{j=1}^{\infty} L_j \hat{\rho}(0) L_j^\dagger \right) \tag{7.11}$$

where $\partial_t \rho(0) = A\hat{\rho}(0) + \hat{\rho}(0)A^\dagger + \sum_{j=1}^{\infty} L_j \hat{\rho}(0) L_j^\dagger$. By recalling the previous relabelling, we can rewrite this as the prototype of our master equation

$$\partial_t \rho(t) = A\hat{\rho}(t) + \hat{\rho}(t)A^\dagger + \sum_{j=1}^{\infty} L_j \hat{\rho}(t) L_j^\dagger \tag{7.12}$$

One can rewrite any linear operator as the sum of a real term and complex term. If we use a convenient scaling, A can be represnted as

$$A = \frac{i}{\hbar} H + M \tag{7.13}$$

where H is complex, M is real, and both are Hermitian. Recall that Kraus operators satisfiy the Kraus representation and plug K_0 into Equation 7.3, and we find that

$$M = -\frac{1}{2} \sum_{i=1}^{\infty} L_i^\dagger L_i \tag{7.14}$$

We can define a so-called *effective Hamiltonian*[a] in terms of H as seen previously

$$H_{\text{eff}} = H - \frac{i\hbar}{2} \sum_{i=1}^{\infty} L_i^\dagger L_i \tag{7.15}$$

[1] Which one can find in many field, not just physics.

This, along with Equation 7.14, one can simplify Equation 7.12 as

Theorem 7.4 (GKSL equation) The *Gorini-Kossakowski-Sudarshan-Lindblad equation*[a] or simply the *Lindblad equation* is a Markovian quantum master equation

$$\partial_t \hat{\rho}(t) = -\frac{i}{\hbar}\left(H_{\text{eff}}\hat{\rho}(t) - \hat{\rho}(t)H_{\text{eff}}^{\dagger}\right) + \sum_{i=1}^{\infty} L_i^{\dagger}\hat{\rho}(t)L_i \tag{7.16}$$

where L_i are the so-called *jump operators*.

[a] Named after Vittorio Gorini, Andrzej Kossakowski, George Sudarshan and Göran Lindblad.

[a] Which is not an actual Hamiltonian due to it not being Hermitian.

Remark 7.6 The Schrödinger and von Neumann equations are special cases of the GKSL equation.

Remark 7.7 The jump operator is so-called because it represents the quite literal 'jump' from one state to another. e.g. the jump from the excited state $|e\rangle$ to the ground state $|g\rangle$ when a photon is emitted.

Chapter 8

WKB method

8.1 WKB wavefunction

Interestingly, the *WKB method*[1] or the *WKB approximation* actually falls into the realm of mathematical physics. One can use it to approximate any differential equation whose highest-order derivative $\epsilon \frac{d^n y}{dx^n}$ has a *smol* parameter ϵ as a first-order DE.
In QM, this is often useful in the case of a one-dimensional potential that *slowly* varies with the spatial dimension x. Surprisingly, one example is quantum tunnelling. Up to this point, you have probably seen the simplified version, which regards the potential barrier as a simple step function. However, in real life, this supposed step function is almost always actually continuous. Instead of seeking a numerical solution, it is actually possible to derive an approximate analytical soluton through the WKB method.

Derivation 8.1 (WKB wavefunction) A constant potential has the wavefunction $\psi(x) = e^{ikx}$ for some parameter k. For a slowly varying potential, the wavefunction is parameterised:

$$\psi(x) = Ae^{\frac{i}{\hbar}S(x)} \tag{8.1}$$

where $S(x)$ is closely related to the acton[a]. To derive $S(x)$, we substitute this expression into the TISE:

$$-i\hbar 2mS''(x) + \frac{1}{2m}S'(x)^2 + V(x) - E = 0 \quad \text{or} \quad -i\hbar S''(x) + S'(x)^2 = p^2(x) \tag{8.2}$$

where the dashes represent x-derivatives and the momentum observes, much like plane wave solutions

$$p(x) \equiv \frac{h}{\lambda(x)} = \sqrt{2m(E - V(x))} \tag{8.3}$$

Admittedly, $\hbar$ is quite *smol*, and we can see that Equation 8.2 is of the type described in the premise of the WKB method. $\hbar$ corresponds to the *smol* parameter ϵ, and we can say that

$$O(\hbar^n) = 0 \quad \text{for} \quad n \geq 1 \tag{8.4}$$

It is convenient to pretend that $\hbar$ is a variable and regard $S(x)$ as a power series of $\hbar$:

$$S(x) = \sum_{i=0}^{\infty} \hbar^i S_i(x) \tag{8.5}$$

At first, this may seem a bit puzzling. But then we recall that $S(x)$ is related to the action. Hence

- $S_0(x)$ corresponds to the classical action that satisfies the Hamilton-Jacobi equation.
- All other $S_i(x)$ terms introduce quantum corrections (effectively perturbations) to the classical action $S_0(x)$.

[1] Named after Gregor Wentzel, Hendrik Anthony Kramers and Léon Brillouin. Rarely, Harold Jeffreys is included, whence the alternate short form *JWKB method* arises.

The inclusion of the classical $S_0(x)$ and the resulting failure to completely quantise the action means that this approximation is *semiclassical*. Inserting it into Equation 8.2 yields

$$-\frac{i\hbar}{2m}\sum_{i=0}^{\infty}\hbar^i S_i''(x)+\frac{1}{2m}\sum_{i=0}^{\infty}\hbar^i S_i'(x)^2+V(x)-E=0 \tag{8.6}$$

where the first term is eliminated due to all members of the sum having higher orders of $\hbar$.
Now we look at the rest of the equation. Significantly, by treating $\hbar$ as a variable, we have effectively turned Equation 8.2 into a series of equations. Each equation i has all their terms prefixed with $\hbar$ to the order of i - prefixes which cancel out. The first two of such equations, with coefficients $\hbar^0=1$ and $\hbar^1=\hbar$, are

$$\frac{1}{2m}S_0'(x)^2+V(x)-E=0 \tag{8.7}$$

$$-\frac{i\hbar}{2m}S_0''(x)+\frac{2}{2m}S_1'(x)S_0'(x)=0 \tag{8.8}$$

Equation 8.7 can be easily solved as

$$S_0(x)=\pm\int dx p(x)+c \tag{8.9}$$

Equation 8.8 is slightly trickier. One can first rewrite it as

$$\frac{S_0''(x)}{S_0'(x)}=-i2s_1'(x) \tag{8.10}$$

Integrating yields

$$\ln(S_0'(x))=-2iS_1(x)+c \tag{8.11}$$

Eliminating any terms with higher orders of $\hbar$ allows us to write

$$S(x)\approx S_0(x)\hbar S_1(x) \tag{8.12}$$

inserting this into Equation 8.1 yields

$$\psi(x)=\frac{Ae^{iC}}{\sqrt{\pm 1}}\frac{e^{\pm\frac{i}{\hbar}\int dx p(x)}}{\sqrt{p(x)}} \tag{8.13}$$

which we can simplify this as a superposition of positive and negative solutions

Theorem 8.1 (WKB wavefunction)

$$\psi(x)=\frac{A}{\sqrt{p(x)}}e^{i\int p(x)dx/\hbar}+\frac{B}{\sqrt{p(x)}}e^{-i\int p(x)dx/\hbar} \tag{8.14}$$

where A and B are arbitrary constants.

In classically forbidden regions ($E<V(x)$), it is convenient to rewrite Equation 8.14 in terms of a $q(x)=\sqrt{2m(V(x)-E)}$:

$$\psi(x)=\frac{C}{\sqrt{q(x)}}e^{\int q(x)dx/\hbar}+\frac{D}{\sqrt{q(x)}}e^{-\int q(x)dx/\hbar} \tag{8.15}$$

with C and D likewise being arbitrary constants.

[a] We will soon see the significance of this.

Remark 8.1 Consider a classical particle distributed at random, the probability (amplitude) of finding which is inversely proportional to its velocity. The probability amplitude is the square of the (normalised) wavefunction while the momentum is proportional to the velocity. Hence, the existence of the prefix $\frac{1}{\sqrt{p(x)}}$

in Equation 8.14 can be justified.

8.2 Uniform approximaion

We now consider the so-called *classical turning points* (CTPs) which are where transitions between allowed and forbidden regions take place. At the CTPs, $E = V(x)$ and we have zero momentum. The wavefunction seen in Equation 8.14 diverges to infinity, and the WKB method fails at the CTPs. To model the wavefunctions at the CTP, we use the *connection formulas*.

Derivation 8.2 (Connection formulas) We set a point $x = a$ to be very close to the CTP and assume $V(x)$ to be approximately linear[a], known as the *uniform approximation*:

$$V(x) - E = A(x - a) \tag{8.16}$$

where we have defined

$$A = \left.\frac{\partial V(x)}{\partial x}\right|_{x=a} \tag{8.17}$$

The TISE is thus

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} + A(x-a)\psi(x) = 0 \tag{8.18}$$

Remark 8.2 Note that A and $(x - a)$ are two separate terms!
If $A > 0$, then the potential is rising around $x = a$, and the barrier to the right. Likewise, if $A < 0$, the barrier is to the left.
We can perform a change of variable and set

$$z = \left(\frac{2mA}{\hbar^2}\right)^{\frac{1}{3}}(x - a) \tag{8.19}$$

which allows us to rewrite Equation 8.18 into the *Airy equation* or the *Stokes equation*:

Theorem 8.2 (Airy equation)

$$\frac{\partial^2\psi(z)}{\partial z^2} + z\psi(z) = 0 \tag{8.20}$$

The Airy equation, which has been studied since the 19$^{\text{th}}$ century, is so-called due to the *Airy function* being a solution to it:

Definition 8.1 (Airy functions) The Airy functions of the first and second kind, or alternatively the Airy and *Bairy* functions, are denoted as $\mathrm{Ai}(z)$ and $\mathrm{Bi}(z)$ respectively:

$$\mathrm{Ai}(z) = \frac{z^{1/2}}{3}\left[I_{-1/3}\left(\frac{2z^{3/2}}{3}\right) - I_{1/3}\left(\frac{2z^{3/2}}{3}\right)\right] \tag{8.21}$$

$$\mathrm{Bi}(z) = \left(\frac{z}{3}\right)^{1/2}\left[I_{-1/3}\left(\frac{2z^{3/2}}{3}\right) + I_{1/3}\left(\frac{2z^{3/2}}{3}\right)\right] \tag{8.22}$$

where $I_{-1/3}$ is the Bessel function of one-third order.

A general solution to the Airy equation is a linear combination of the two functions:

$$\psi(z) = \alpha\,\mathrm{Ai}(z) + \beta\,\mathrm{Bi}(z) \tag{8.23}$$

However, in the context of QM, we are not too concerned with the general form of the Airy functions, and instead utilise their simplified forms when z becomes *large* in either the positive or the negative direction.

Theorem 8.3 (Airy function approximations under large positive z)

$$\mathrm{Ai}(z) \approx \frac{1}{2\sqrt{\pi}} \frac{\exp\left[-\frac{2}{3}|z|^{3/2}\right]}{|z|^{1/4}} \tag{8.24}$$

$$\mathrm{Bi}(z) \approx \frac{1}{\sqrt{\pi}} \frac{\exp\left[+\frac{2}{3}|z|^{3/2}\right]}{|z|^{1/4}} \tag{8.25}$$

Theorem 8.4 (Airy function approximations under large negative z)

$$\mathrm{Ai}(z) \approx \frac{1}{\sqrt{\pi}} \frac{\cos\left[\frac{2}{3}|z|^{3/2} - \frac{\pi}{4}\right]}{|z|^{1/4}} \tag{8.26}$$

$$\mathrm{Bi}(z) \approx \frac{-1}{\sqrt{\pi}} \frac{\sin\left[\frac{2}{3}|z|^{3/2} - \frac{\pi}{4}\right]}{|z|^{1/4}} \tag{8.27}$$

While these reduced Airy functions are continuous across the CTP, they are not the full solution due to the assumption that $V(x)$ is linear instead of a smooth curve as is the case in the WKB method. To find the WKB Airy functions, we first perform the substitutions

$$z(x) = \begin{cases} \int_x^a p(x')dx'/\hbar & \text{when } z < 0 \text{ (allowed region)} \\ \int_a^x q(x')dx'/\hbar & \text{when } z > 0 \text{ (forbidden region)} \end{cases} \tag{8.28}$$

Note that the distinction between the two substitutions in the allowed and forbidden regions is purely one of formalism (recall $p(x)$ and $q(x)$).

Remark 8.3 By comparing, we can see that the post-substitution Airy function approximations are merely WKB functions as seen in Equation 8.14 and Equation 8.15.

Now consider some CTP *a to the right of which lies the forbidden area.* By adjusting their normalisation factors and considering the limit on either side of the CTP, we can recover the standard WKB form of the Airy functions in allowed and forbidden regions. Relabelling $\mathrm{Ai}(z)$ and $\mathrm{Bi}(z)$ as $\psi_A(x)$ and $\psi_B(x)$ respectively, we can write a rescaled general solution for the Airy equation

$$\psi(z) = \alpha\psi_A(x) + \beta\psi_B(x) \tag{8.29}$$

where α and β are likewise rescaled. The rescaled *WKB Airy functions* are

Theorem 8.5 (WKB Airy functions in the allowed region ($E > V(x)$))

$$\psi_A(x) = \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{p(x)}} \cos\left(\int_x^a p\left(x'\right) dx'/\hbar - \frac{\pi}{4}\right) \tag{8.30}$$

$$\psi_B(x) = \frac{-1}{\sqrt{\pi}} \frac{1}{\sqrt{p(x)}} \sin\left(\int_x^a p\left(x'\right) dx'/\hbar - \frac{\pi}{4}\right) \tag{8.31}$$

Theorem 8.6 (WKB Airy functions in the forbidden region ($E < V(x)$))

$$\psi_A(x) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{q(x)}} \exp\left(-\int_b^x q\left(x'\right) dx'/\hbar\right) \tag{8.32}$$

$$\psi_B(x) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{q(x)}} \exp\left(+\int_b^x q\left(x'\right) dx'/\hbar\right) \tag{8.33}$$

We are now in a position to discuss the connection formulas conceptually. So far, we have established the following:

- The WKB solutions in Equation 8.14 and Equation 8.15 fail at CTPs, where they blow up instead.

- Rather, the Airy equation is used to produce solutions - although with a catch. We now have two sets of solutions, in the allowed and forbidden regions (which are separated by the CTP a) respectively.
- The connection formulas are hence developed so that we can 'connect' the two sets of solutions on both sides of a.

Theorem 8.7 (Connection formulas for a)

$$\frac{1}{\sqrt{q(x)}}\exp\left(-\int_a^x q\left(x'\right)dx'/\hbar\right)\to\frac{2}{\sqrt{p(x)}}\cos\left(\int_x^a p\left(x'\right)dx'/\hbar-\frac{\pi}{4}\right) \tag{8.34}$$

$$-\frac{1}{\sqrt{p(x)}}\sin\left(\int_x^a p\left(x'\right)dx'/\hbar-\frac{\pi}{4}\right)\to\frac{1}{\sqrt{q(x)}}\exp\left(+\int_a^x q\left(x'\right)dx'/\hbar\right) \tag{8.35}$$

where the arrows denote the transition from the allowed to the forbidden region. Putting them together and accounting for arbitrary constants, one has

$$\begin{aligned}&\frac{A}{\sqrt{q(x)}}\exp\left(-\int_a^x q\left(x'\right)dx'/\hbar\right)-\frac{B}{\sqrt{p(x)}}\sin\left(\int_x^a p\left(x'\right)dx'/\hbar-\frac{\pi}{4}\right)\to\\&\frac{2A}{\sqrt{p(x)}}\cos\left(\int_x^a p\left(x'\right)dx'/\hbar-\frac{\pi}{4}\right)+\frac{B}{\sqrt{q(x)}}\exp\left(+\int_a^x q\left(x'\right)dx'/\hbar\right)\end{aligned} \tag{8.36}$$

Previously, we have established that the potential barrier lies to the right of a. Now we turn out gaze right and towards the tail end of the potential barrier, where we inevitably encounter another CTP b. Choosing b as our reference, the potential barrier is now to our left. The connection formulas for b has inverse arrow directions compared to their a counterparts, but are otherwise identical[b].

[a] This can be verified by performing a Taylor expansion around $V(a)$. The expansion is linear to the first order.

[b] This is because the arrows *still* denote the transition from the allowed to the forbidden region, which have now *reversed* their order along the x-direction.