# Quantum physics II
# Prof. Zoë Holmes — EPFL

Notes by Joachim Favre

Quantum science and engineering master — Semester 1
Autumn 2024

I made this document for my own use, but I thought that typed notes might be of interest to others. There are mistakes, it is impossible not to make any. If you find some, please feel free to share them with me (grammatical and vocabulary errors are of course also welcome). You can contact me on Telegram:

`@JoachimFavre`

If you did not get this document through my GitHub repository, then you may be interested by the fact that I have one on which I put those typed notes and their LaTeX code. Here is the link (make sure to read the README to understand how to download the files you're interested in):

`https://github.com/JoachimFavre/UniversityNotes`

Please note that the content does not belong to me. I have made some structural changes, reworded some parts, and added some personal notes; but the wording and explanations come mainly from the Professor, and from the book on which they based their course.

I think it is worth mentioning that in order to get these notes typed up, I took my notes in LaTeX during the course, and then made some corrections. I do not think typing handwritten notes is doable in terms of the amount of work. To take notes in LaTeX, I took my inspiration from the following link, written by Gilles Castel. If you want more details, feel free to contact me on Telegram, as mentioned hereinabove.

`https://castel.dev/post/lecture-notes-1/`

I would also like to specify that the words "trivial" and "simple" do not have, in any course, the definition you find in a dictionary. We are at EPFL, nothing we do is trivial. Something trivial is something that a random person in the street would be able to do. In our context, understand these words as "simpler than the rest". Also, it is okay if you take a while to understand something that is said to be trivial (especially as I love using this word everywhere!).

Since you are reading this, I will give you a little advice. Sleep is a much more powerful tool than you may imagine, so do not neglect a good night of sleep in favour of studying (especially the night before an exam). I wish you to have fun during your exams.

*Version 2025–08–03–official*

*To Gilles Castel, whose work has*
*inspired me this note taking method.*

*Rest in peace, nobody*
*deserves to go so young.*

# Contents

# List of lectures

# Chapter 1

# Summary by lecture



- Definition of qubit.
- Explanation of the time evolution axiom.
- Definition of Pauli matrices, and explanation of their properties.
- Explanation of the measurement axiom.
- Definition of tensor product.
- Explanation of the composite system axiom.



- Explanation of global and local measurement.
- Explanation of the quantum eraser.
- Explanation of the no-signaling theorem.
- Explanation of the quantum psychics game.



- Explanation of the CHSH inequality.
- Explanation of Bell's inequalities.
- Definition of density operator, pure state, and mixed state.



- Proof of the properties of density operators.
- Definition and justification of reduced states.
- Definition of partial trace, and explanation of its properties.
- Proof of the no-signaling theorem.
- Explanation of the measurement problem, the collapse postulate and decoherence.



- Definition of Fermions and Bosons.
- Derivation of a basis for the wavefunction of identical Fermions and identical Bosons.
- Explanation of the second quantisation.
- Explanation of the Hong-Ou-Mandel effect.

**Lecture 13 : Ah yes, $\mathfrak{so}(3) = SO(3)$** — Monday 16th December 2024 *p. 96*

- Construction of irreps for $\mathfrak{so}(3)$.
- Explanation and examples of the Clebsch-Gordan decomposition.

# Chapter 2

# Basics

## 2.1 Qubits

**Definition: Qubit**

A qubit is a 2-level quantum system.
Its state can therefore always be written as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$.

*Examples* For instance, an electron's spin $\{|\uparrow\rangle, |\downarrow\rangle\}$, the polarisation of a photon $\{|H\rangle, |V\rangle\}$, the two enegy levels of any atom or molecule $\{|G\rangle, |E\rangle\}$ and a photon in different arms of a beamsplitter $\{|L\rangle, |R\rangle\}$ are examples of such 2-level systems.

**Theorem: Bloch sphere**

Let $|\psi\rangle$ be the state of some q-bit.
Then, there exists a $\theta \in [0, \pi]$ and a $\varphi \in [0, 2\pi[$ such that we can write:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle.$$

*Intuition* $(\theta, \varphi)$ can be interpreted as coordinates on a sphere. In other words, we can visualise is using the following unit vector:

$$\vec{\boldsymbol{v}} = \begin{pmatrix}\sin(\theta)\cos(\varphi) & \sin(\theta)\sin(\varphi) & \cos(\theta)\end{pmatrix}^T.$$

*Remark* When we will later look at density matrices, the fact that $\theta$ is divided by 2 will appear naturally.

*Proof* This directly comes from the fact that the general phase does not matter (i.e. $e^{i\delta}|\psi\rangle \equiv |\psi\rangle$ for all $\delta \in \mathbb{R}$), and that $|\alpha|^2 + |\beta|^2 = 1$.

## 2.2 Time evolution

**Axiom: Time evolution**

Let $H$ be the Hamiltonian of a system, and $|\psi_0\rangle$ be the initial state of the system.
We can get the state of a system at any time $t$, $|\psi(t)\rangle$, by solving the equation:

$$\frac{\partial|\psi(t)\rangle}{\partial t} = \frac{H}{i}|\psi_0\rangle.$$

This is also known as **Schrödinger's equation**.

*Remark* Note that, formally, $H = \frac{H'}{\hbar}$, and $H'$ is the real Hamiltonian of the system. Our convention allows to simplify notations. Note moreover that $H = H'$ in natural units *(but natural units are really awful)*.

**Definition: Unitary operator**

Let $U$ be an operator in some Hilbert space. It is said to be **unitary** if:

$$UU^\dagger = U^\dagger U = I.$$

*Properties*

- Unitary operators are reversible: given $|\psi'\rangle = U|\psi\rangle$, we can find $|\psi\rangle = U^\dagger|\psi'\rangle$.
- They preserved lengths: given $|\psi'\rangle = U|\psi\rangle$, we have

$$\langle\psi'|\psi'\rangle = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle = 1.$$

- As all operators, they are linear:

$$U(\alpha|\psi\rangle + \beta|\varphi\rangle) = \alpha U|\psi\rangle + \beta U|\varphi\rangle.$$

**Theorem: Unitary evolution**

There always exists a unitary $U(t)$ that solves Schrödinger's equation:

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle.$$

*Example*

For instance, if $H$ is time-independent:

$$U(t) = e^{-iHt}.$$

**Definition: Pauli Matrices**

The **Pauli matrices** are defined by:

$$\sigma_0 = I = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \sigma_1 = \sigma_X = X = NOT = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},$$

$$\sigma_2 = \sigma_Y = Y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad \sigma_3 = \sigma_Z = Z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

The different notation can be used interchangeably. Those matrices are both Hermitian and unitary.

*Pauli vector*

We can write the $(X, Y, Z)$ Pauli matrices as a pseudo-vector:

$$\vec{\boldsymbol{\sigma}} = \begin{pmatrix}\sigma_1 & \sigma_2 & \sigma_3\end{pmatrix}^T.$$

This is useful to use dot-products as a short-hand notation:

$$\vec{\boldsymbol{\sigma}} \bullet \begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix}^T = \alpha\sigma_1 + \beta\sigma_2 + \gamma\sigma_3.$$

*Properties*

Pauli matrices have properties that can save a lot of time.

1. $\mathrm{Tr}(I) = 2$ and $\mathrm{Tr}(X) = \mathrm{Tr}(Y) = \mathrm{Tr}(Z) = 0$.
2. Squaring a Pauli gives identity and multiplying two different Paulis gives the third one up to a constant factor. In other words, for $i, j \in \{1, 2, 3\}$:

$$\sigma_i\sigma_j = \delta_{ij}I + i\varepsilon_{ijk}\sigma_k,$$

   where $\varepsilon_{ijk}$ is the Levi-Civita symbol, and where we use Einstein's summation notation.
3. For $i, j \in \{1, 2, 3\}$:

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k.$$

4. For $i, j \in \{1, 2, 3\}$:

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}I.$$

5. Pauli matrices form an orthogonal basis. In other words, for $i, j \in \{0, 1, 2, 3\}$:

$$\mathrm{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}.$$

All properties directly come from the two first.

**Remark: Paulis are gates**

Since Paulis are unitary, they can be used as gates (i.e. discrete time evolution). For instance:

$$X|0\rangle = |1\rangle, \quad X|1\rangle = |0\rangle.$$

This justifies this gate can be called the $NOT$ gate.

**Theorem**

Let $H$ be an operator such that $H^2 = I$.
Then:

$$\exp(i\alpha H) = \cos(\alpha)I + i\sin(\alpha)H.$$

*Proof* This directly comes from the definition of the exponential of an operator. It will be proven in the first exercise series.

**Remark: Pauli are generators**

Since Paulis are Hermitian, they can be used as Hamiltonians. In fact, since they form a basis, any single-qubit Hamiltonian can be expressed as a sum of Paulis:

$$H = \sum_{i=1}^{3} \omega n_i \sigma_i = \omega(n_1 X + n_2 Y + n_3 Z) = \omega \vec{\boldsymbol{n}} \bullet \vec{\boldsymbol{\sigma}},$$

where $\omega \in \mathbb{R}$ is arbitrary, and $\begin{pmatrix} n_1 & n_2 & n_3 \end{pmatrix}^T \in \mathbb{R}^3$ is a unit vector.
Note that the sum ranges from 1 to 3, skipping $\sigma_0 = I$. Indeed, the corresponding unitary is $U(t) = \exp(-iHt)$, but once we take the exponential, any identity term would become a global phase:

$$\exp(i\theta I + i\varphi B) = \exp(i\theta I)\exp(i\varphi B) = \underbrace{\exp(i\theta)}_{\text{global}}\exp(i\varphi B).$$

Recall that global phases do not matter in quantum physics. Therefore, using our theorem above, any arbitrary qubit unitary can be written as:

$$U = \exp(-iHt) = \exp(-i\omega \vec{\boldsymbol{n}} \bullet \vec{\boldsymbol{\sigma}} t) = \cos(\omega t)I - i\sin(\omega t)\vec{\boldsymbol{n}} \bullet \vec{\boldsymbol{\sigma}}.$$

It can be proven that this induces a rotation on the Bloch sphere around the $\vec{\boldsymbol{n}}$ axis at a rate of $2\omega t$.

*Example* Let us consider:

$$\vec{\boldsymbol{n}} = \vec{\boldsymbol{n}}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \implies \vec{\boldsymbol{n}}_z \bullet \vec{\boldsymbol{\sigma}} = \sigma_z.$$

Then, we have:

$$\begin{aligned} e^{-i\omega \vec{\boldsymbol{n}} \bullet \vec{\boldsymbol{\sigma}} t}|\psi\rangle &= e^{-i\omega\sigma_z t}\left(\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle\right) \\ &= \cos\left(\frac{\theta}{2}\right)e^{-i\omega t}|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\varphi}e^{i\omega t} \\ &= e^{-i\omega t}\left(\cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)e^{i(\varphi+2\omega t)}|1\rangle\right). \end{aligned}$$

We then have $\theta(t) = \theta$ and $\varphi(t) = \varphi + 2\omega t$. Therefore, $U$ makes $|\psi\rangle$ rotate on the Bloch sphere around the $z$-axis at a rate of $2\omega t$, as expected.

## 2.3 Measurements

**Definition: Observables**

An **observable** is an Hermitian operator $M$.

*Property*

By using the spectral theorem, $M$ can always be orthodiagonalised with real eigenvalues:

$$M = \sum_k \lambda_k |\lambda_k\rangle\langle\lambda_k|, \quad \lambda_k \in \mathbb{R},$$

$$\langle\lambda_i|\lambda_j\rangle = \delta_{ij}.$$

**Definition: Expectation**

The **expectation** of an observable $M$ in state $|\psi\rangle = \sum_k \alpha_k |\lambda_k\rangle$ is:

$$\langle M\rangle = \sum_k |\alpha_k|^2 \lambda_k.$$

This is moreover equal to:

$$\langle M\rangle = \langle\psi|M|\psi\rangle$$

*Intuition*

This is indeed an expectation: the $|\alpha_k|^2$ are probabilities of different observation result, and the $\lambda_k$ are the value measured from those observations.

*Proof*

This is a direct equality:

$$\langle M\rangle = \langle\psi|M|\psi\rangle = \sum_{k,k',j} \alpha_k \alpha_{k'}^* \underbrace{\langle\lambda_k|\lambda_j\rangle}_{\delta_{kj}} \underbrace{\langle\lambda_j|\lambda_{k'}\rangle}_{\delta_{jk'}} \lambda_k = \sum_k |\alpha_k|^2 \lambda_k.$$

□

**Definition: Projector**

The **projector** on the $k^{\text{th}}$ outcome of an observable $M$ is defined by:

$$\Pi_k = |\lambda_k\rangle\langle\lambda_k|.$$

The probability of getting the outcome $k$ is given by:

$$p_k = \langle\psi|\Pi_k|\psi\rangle.$$

Moreover, in the case of an ideal measurement, the state after measurement is said to collapse to:

$$|\psi'\rangle = \frac{\Pi_k|\psi\rangle}{\sqrt{p_k}}.$$

*Remark*

The last claim has many issues. We will come back to it when we will see the measurement problem. The idea is that, after measuring a state, it may collapse to a lot of different states; although, in the end, it does not really matter.

**Thought experiment**

We consider the four following experiments, where we start with some state, let it evolve for some time (all experiments evolve under the same Hamiltonian), and measure it. All four experiments are possible to make in a lab.

1. We start with a system in state $|0\rangle$. We then wait half an hour and measure it, measuring $|0\rangle$ half of the time and $|1\rangle$ the other half.
2. We start with a system in $|1\rangle$. We then wait half an hour and measure it, measuring $|0\rangle$ half of the time and $|1\rangle$ the other half.
3. We start with a system in state $|0\rangle$. We then wait half an hour, and measure it. We wait another half hour and measure. For both measures, we get $|0\rangle$ half of the time, and $|1\rangle$ the other half.

4. We start with a system in state $|0\rangle$. We then wait a full hour, and measure. We always measure the result to be in state $|0\rangle$.

Let us explain those experiments mathematically.

1. Experiment 1 can be described as:

$$|0\rangle \xrightarrow{\text{wait } 30 \text{ min}} \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

   In other words:

$$U(30\,\text{min})|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

2. For experiment 2, we need to get a state that is orthogonal to $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ after waiting (because unitary transformations map orthogonal states to orthogonal states), so we can take:

$$|1\rangle \xrightarrow{\text{wait } 30 \text{ min}} \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

3. Since measuring a state collapses its state to $|0\rangle$ or $|1\rangle$, the third experiment is like chaining the first and second experiments.
4. Merging the idea of the first and second experiment, we have, for the fourth experiment:

$$|0\rangle \xrightarrow{\text{wait 30 minutes}} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{\text{wait 30 minutes}} \frac{1}{\sqrt{2}}\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = |0\rangle.$$

*Takeaway message* This shows the very unintuitiveness of quantum mechanics: the measure in experiment 3 had an impact.

# Chapter 3

# Composite systems and entanglement

## 3.1 Tensor product

**Definition: Tensor product**

Let $\mathcal{H}_A$ be a Hilbert space spanned by $\{|\mu_1\rangle, \dots, |\mu_{d_A}\rangle\}$ and $\mathcal{H}_B$ be a Hilbert space spanned by $\{|\nu_1\rangle, \dots, |\nu_{d_B}\rangle\}$.
The **tensor product** (also called Kronecker product) of $\mathcal{H}_A$ and $\mathcal{H}_B$, written $\mathcal{H}_A \otimes \mathcal{H}_B$, is a space spanned by:

$$|\lambda_{ij}\rangle = |\mu_i\rangle \otimes |\nu_j\rangle, \quad i \in \{1, \dots, d_A\}, j \in \{1, \dots, d_B\}.$$

It has dimension $d_{AB} = d_A d_B$. The inner product on basis elements of this Hilbert space is:

$$\langle \lambda_{ij} | \lambda_{k\ell} \rangle = \langle \mu_i \nu_j | \mu_i \nu_j \rangle = \langle \mu_i | \mu_k \rangle \langle \nu_j | \nu_\ell \rangle.$$

Moreover, the tensor product of vectors is defined in such a way that it follows the two following properties:

- *(Linearity)* $a(|\nu\rangle \otimes |\omega\rangle) = (a|\nu\rangle) \otimes |\omega\rangle = |\nu\rangle \otimes (a|\omega\rangle)$.
- *(Distributivity)* $(|\nu_1\rangle + |\nu_2\rangle) \otimes |\omega\rangle = |\nu_1\rangle \otimes |\omega\rangle + |\nu_2\rangle \otimes |\omega\rangle$.

*Properties*

Let $|\varphi\rangle = \sum_{ij} b_{ij} |\lambda_{ij}\rangle$ and $|\psi\rangle = \sum_{ij} c_{ij} |\lambda_{ij}\rangle$ be any arbitrary element of $\mathcal{H}_{AB}$. Then, their inner product is given by:

$$\langle \varphi | \psi \rangle = \sum_{ijk\ell} b_{ij}^* c_{k\ell} \delta_{ik} \delta_{j\ell} = \sum_{ij} \beta_{ij}^* c_{ij},$$

just like we would expect from any basis.
Similarly, the identity can simply be written as:

$$I_{AB} = I_A \otimes I_B = \sum_{ij} |\lambda_{ij}\rangle\langle\lambda_{ij}|.$$

*Remark*

We can chain tensor products to consider larger systems, such as $\mathcal{H}_{ABC} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. This has dimension $d_{ABC} = d_A d_B d_C$.

*Notation*

It is typical to drop the $\otimes$ when we write tensor products, or even merge them into the same ket:

$$|\mu_i\rangle \otimes |\nu_j\rangle = |\mu_i\rangle|\nu_j\rangle = |\mu_i \nu_j\rangle.$$

**Composite system**

Let $A, B$ be two systems living in Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.
The composite system $AB$ lives in a Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, where $\otimes$ is a tensor product (also called Kronecker product).

*Example*

Let us consider $A$ and $B$ to be single qubit systems, i.e. $AB$ is a two-qubit system. Both $A$ and $B$ are spanned by $\{|0\rangle, |1\rangle\}$, so $AB$

is spanned by:

$$|0\rangle = |0\rangle \otimes |0\rangle, \quad |1\rangle = |0\rangle \otimes |1\rangle, \quad |2\rangle = |1\rangle \otimes |0\rangle, \quad |3\rangle = |1\rangle \otimes |1\rangle.$$

Any state can then be expressed as:

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle + c_2|2\rangle + c_3|3\rangle.$$

We can for instance consider:

$$|\varphi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{|1\rangle - |2\rangle}{\sqrt{2}}.$$

It is indeed normalised since the sum of the modulus squared of its coefficient is 1. More formally:

$$\langle\varphi|\varphi\rangle = \left|\frac{1}{\sqrt{2}}\right|^2 + \left|-\frac{1}{\sqrt{2}}\right|^2 = 1.$$

**Operators in composite systems**

Let $T_A$ and $T_B$ be operators acting on their respective systems. The resulting joint operator on $AB$, written $T_A \otimes T_B$, is defined such that:

$$(T_A \otimes T_B)(|\varphi\rangle_A \otimes |\psi\rangle_B) = (T_A|\varphi\rangle_A) \otimes (T_B|\psi\rangle_B).$$

$T_A \otimes I_B$ and $I_A \otimes T_B$ are called **local operators**.

*Property* Local operators that act on different systems always commute:

$$\begin{aligned}(T_A \otimes I_B)(I_A \otimes T_B) &= (T_A I_A) \otimes (I_B T_B) \\ &= T_A \otimes T_B \\ &= (T_A \otimes I_B)(I_A \otimes T_B).\end{aligned}$$

*Notation* When we have indices that denote on which system an operator acts, we may drop identities. That way, we note:

$$T_A \otimes I_B \equiv T_A.$$

This moreover allows us to write $T_A \otimes T_B$ more easily:

$$T_A T_B = (T_A \otimes I_B)(I_A \otimes T_B) = T_A \otimes T_B.$$

**Explicit vector and matrix representation**

Up to an isomorphism, it is possible to show that, given any vector $|v\rangle$:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \otimes |v\rangle = \begin{pmatrix} a_1|v\rangle \\ \vdots \\ a_n|v\rangle \end{pmatrix}.$$

Similarly, given any matrix $A$:

$$\begin{pmatrix} a_1 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \otimes A = \begin{pmatrix} a_{11}A & \cdots & a_{1n}A \\ \vdots & \ddots & \vdots \\ a_{n1}A & \cdots & a_{nn}A \end{pmatrix}.$$

*Example* For instance:

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ b\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a\alpha \\ a\beta \\ b\alpha \\ b\beta \end{pmatrix}.$$

Monday 23rd September 2024 — **Lecture 2 : Star Trek goes brrr**

**Definition: Entangled state**

A state $|\Psi\rangle$ is said to be **entangled** if and only if it cannot be written as a product state:

$$|\Psi\rangle \neq |\varphi\rangle|\psi\rangle$$

*Example*

For instance, the following is an entangled state:

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

However, the following are two product states:

$$|0\rangle|1\rangle, \quad \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2} = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

**Axiom: Global measurement**

On an arbitrary composite system, a measurement is represented by a Hermitian operator $M_{AB} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$. This type of measurement is said to be **global**.

*Remark*

This is exactly what we had before for arbitrary quantum systems.

**Axiom: Local measurement**

A measurement on a single subsystem $A$ of a composite system $AB$ is represented by $M_A \otimes I$, where $M_A$ is a Hermitian operator. This type of measurements is said to be **local**.

*Remark*

Let us verify this makes sense for expected values. We decompose $M_A = \sum_i \mu_i |\mu_i\rangle\langle\mu_i|$ and $|\Psi\rangle = \sum_{i,j}^{\infty} \alpha_{ij} |\mu_i \nu_j\rangle$. Then:

$$\langle M_A \otimes I\rangle = \sum_{ij} |\alpha_{ij}|^2 \mu_i = \sum_i \left( \sum_j \mathbb{P}(\mu_i \nu_j) \right) \mu_i = \sum_i \mathbb{P}(\mu_i)\mu_i.$$

This is exactly what we want the expected value to be when we measure on the subsystem $A$.

## 3.2 The quantum eraser

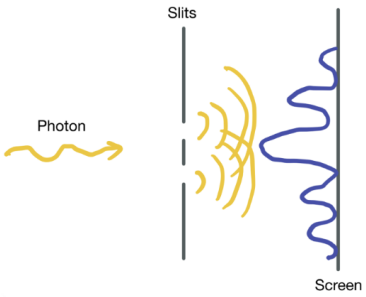
**Thought experiment: The quantum eraser**

We will consider four experiments, which are variations of the double-slit experiment. The idea is that the photon will behave as a wave (making interference pattern) if and only if we have no way of knowing through which slit it went through.

**Experiment 1**

We consider the regular double-slit experiment: we send a photon toward a wall that has two small holes, and with a screen behind. We moreover consider the photon to be horizontally polarised.

Then, we have interference patterns: photons behave as waves.




*Proof*

Let $|\Psi_1(x,t)\rangle$ be the wave function of the photon on that went through the first slit on the screen, and $|\Psi_2(x,t)\rangle$ be the one of the

second. The wave function at the screen is simply:

$$|\Psi(x,t)\rangle = \frac{|\Psi_1(x,t)\rangle + |\Psi_2(x,t)\rangle}{\sqrt{2}} \otimes |H\rangle.$$

The screen measures the position. In other words, our measurement is $|x\rangle\langle x|$, giving that the probability to the photons is measured on any given position $x$ on the screen is:

$$\begin{aligned}\mathbb{P}(x) &= \langle\Psi(x,t)|(|x\rangle\langle x| \otimes I)|\Psi(x,t)\rangle \\ &= \frac{1}{2}\langle\Psi_1(x,t) + \Psi_2(x,t)|x\rangle\langle x|\Psi_1(x,t) + \Psi_2(x,t)\rangle\langle H|H\rangle \\ &= \frac{|\Psi_1(x,t) + \Psi_2(x,t)|^2}{2}.\end{aligned}$$

This does rise to interference pattern. They are very high and destructive at places where $\Psi_1(x,t) \approx -\Psi_2(x,t)$, in which case:

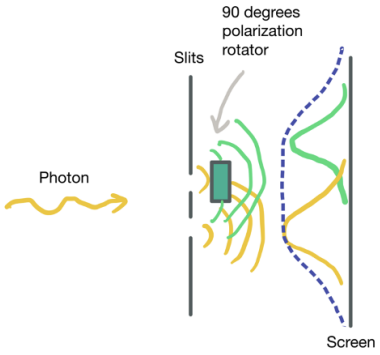
$$\mathbb{P}(x) \approx 0.$$

On the other hand, they are very high and constructive at places where $\Psi_1(x,t) \approx \Psi_2(x,t)$.

□

**Experiment 2**

We consider the same experiment setup as the first experiment, except that we add a 90° polarisation rotator (that turns the horizontally polarised photon to a vertically polarised one) behind one of the holes.
This time, we no longer have any interference. In other words, there are two regions: photons behave as particles.




*Proof*

We use the same notations as in the first experiment.
The first photon going through the first slit has a vertical polarisation, so:

$$|\Psi(x,t)\rangle = \frac{1}{\sqrt{2}}(|\Psi_1(x,t)\rangle|V\rangle + |\Psi_2(x,t)\rangle|H\rangle).$$

Then, the probability to measure the photon at any place on the screen, and writing $|\Psi(x,t)\rangle = |\Psi\rangle$ for a lighter notation:

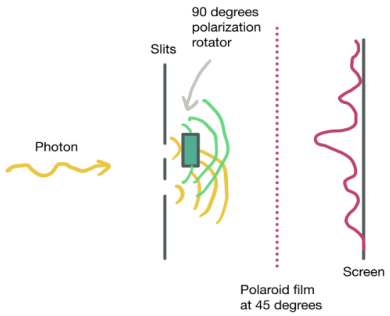
$$\begin{aligned}\mathbb{P}(x) &= \langle\Psi|(|x\rangle\langle x| \otimes I)|\Psi\rangle \\ &= \frac{1}{2}(\langle\Psi_1|\langle V| + \langle\Psi_2|\langle H|)(|x\rangle\langle x| \otimes I)(|\Psi_1\rangle|V\rangle + |\Psi_2\rangle|H\rangle) \\ &= \frac{1}{2}(\langle\Psi_1|x\rangle\langle x|\Psi_1\rangle + \langle\Psi_2|x\rangle\langle x|\Psi_2\rangle) \\ &= \frac{|\Psi_1(x,t)|^2 + |\Psi_2(x,t)|^2}{2}.\end{aligned}$$

This does not leave the room for any interference.

□

**Experiment 3: Quantum eraser**

We consider the same experiment setup as the second experiment, except that we add a 45° polarisation filter between the holes and the screen.
We now get some interference patterns; photons behave as waves. The interference patterns are however half the intensity as the ones of the first experiment.




*Remark* This experiment is named the quantum eraser since, by adding some more measurement *later* in the experiment, we managed to change the behaviour at the slit. In the second experiment it behaved as a particle, but adding some polarisation filter in this third experiment makes it believe again as a wave.

*Proof* The state after the polarisation rotator is the same as the second experiment:

$$\begin{aligned}|\Psi(x,t)\rangle &= \frac{|\Psi_1\rangle|V\rangle + |\Psi_2\rangle|H\rangle}{\sqrt{2}} \\ &= \frac{|\Psi_1\rangle(|\nearrow\rangle - |\swarrow\rangle)}{2} + \frac{|\Psi_2\rangle(|\nearrow\rangle + |\swarrow\rangle)}{2}.\end{aligned}$$

The polarisation filter only keeps the state $|\nearrow\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$. Therefore, the photon goes through the filter with probability $\frac{1}{2}$, in which case it collapses to:

$$\left|\widetilde{\Psi}(x,t)\right\rangle = \frac{|\Psi_1\rangle + |\Psi_2\rangle}{\sqrt{2}}\frac{|H\rangle + |V\rangle}{\sqrt{2}}.$$

When computing the probabilities, we only need to be careful about the fact that the polarisation filter lets through half of the photons. Therefore, we get the same result as in the first experiment, but with an extra $\frac{1}{2}$ factor:

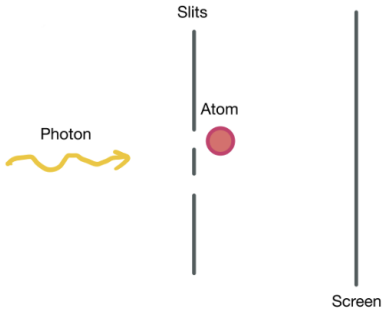
$$\mathbb{P}(x) = \frac{|\Psi_1 + \Psi_2|^2}{4}.$$

This therefore again gives rise to interference patterns.

□

**Experiment 4: Delayed quantum eraser**

We go back to something closer to the double-slit experiment. The photon is no longer polarised, but behind the first hole, we put an atom. If the photon goes through the slit behind which there is this atom, it will change the state of the atom from $|\uparrow\rangle$ to $|\downarrow\rangle$, without being absorbed. We observe *the atom* before the photon hits the screen, either in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ or in the basis $\{|\nearrow\rangle, |\swarrow\rangle\}$.




Depending on the choice of the basis and its outcome, we get different patterns:

- If we measure $|\uparrow\rangle$, we have a pattern $\mathbb{P}_{\uparrow}(x) = |\Psi_1(x)|^2$.
- If we measure $|\downarrow\rangle$, we have a pattern $\mathbb{P}_{\downarrow}(x) = |\Psi_2(x)|^2$.
- If we measure $|\nearrow\rangle$, we have a pattern $\mathbb{P}_{\nearrow}(x) = \frac{1}{2}|\Psi_1(x) + \Psi_2(x)|^2$.
- If we measure $|\swarrow\rangle$, we have a pattern $\mathbb{P}_{\swarrow}(x) = \frac{1}{2}|\Psi_1(x) - \Psi_2(x)|^2$.

In particular, in the first and second cases, the photon behaves a particle; in the third and fourth cases, the photon behaves as a wave, giving rise to interference patterns.

*Remark* This is named the delayed quantum eraser because we can choose very late to measure the atom in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ for the photon to behave as a particle, or in the basis $\{|\nearrow\rangle, |\swarrow\rangle\}$ for the photon to behave a wave.

**No-signaling theorem**

It is not possible to make a quantum protocol that allows to send information faster than the speed of light.

*Remark* This is very important for quantum physics to be coherent with special relativity.

*Proof* We will see a proof of this theorem when we consider density matrices later in the class.

**Theorem: Delayed quantum eraser signaling**

Let us consider the following faulty protocol. Bob sends to Alice a photon that when through the double-slits and atom of the fourth experiment (the delayed quantum eraser). Bob waits until just before Alice measures the photon with her screen. If Bob wants to send a 0, he measures the atom in $\{|\uparrow\rangle, |\downarrow\rangle\}$. If he wants to send a 1, he measure the atom in the basis $\{|\nearrow\rangle, |\swarrow\rangle\}$. Alice can then decode the information, using whether she sees interference patterns or not.

This protocol does not work.

*Proof* This is a direct consequence of the no-signaling theorem. However, let us get some physical intuition on why.

The idea is that, to see a pattern $\mathbb{P}(x)$, Alice needs to receive many photons on her screen and average the positions. Suppose that Bob wants to send a 0. In this cases, he measures the atom in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$, and has a $\frac{1}{2}$ probability to get either case. However, Alice does not know which was his result, so all she can measure is:

$$\mathbb{P}_0(x) = \frac{1}{2}\mathbb{P}_{\uparrow}(x) + \frac{1}{2}\mathbb{P}_{\downarrow}(x) = \frac{|\Psi_1(x)|^2 + |\Psi_2(x)|^2}{2}.$$

There is no interference pattern, as expected.

Similarly, if Bob wants to send a 1, all Alice can measure is the average of the two interference patterns $\mathbb{P}_{\nearrow}(x) = \frac{1}{2}|\Psi_1(x) + \Psi_2(x)|^2$ and $\mathbb{P}_{\swarrow}(x) = \frac{1}{2}|\Psi_1(x) - \Psi_2(x)|^2$:

$$\mathbb{P}_1(x) = \frac{1}{2}\left(\frac{|\Psi_1(x) + \Psi_2(x)|^2}{2} + \frac{|\Psi_1(x) - \Psi_2(x)|^2}{2}\right).$$

However, it is possible to show using analysis tools that $\frac{1}{2}|z_1+z_2|^2+\frac{1}{2}|z_1-z_2|^2=|z_1|^2+|z_2|^2$ for any $z_1,z_2\in\mathbb{C}$. This implies that $\mathbb{P}_1(x)=\mathbb{P}_0(x)$. In other words, when Bob wants to send 1, the interference patterns average out to give no interference pattern at all. Alice can therefore not distinguish the case Bob sends a 0 or a 1.

□

## 3.3 Bell inequalities

**Quantum psychics game** Alice and Bob pretend to be psychics, and Spock and Kirk try to find out if this is true. They make an experiment to decide this. They put Alice and Bob in separate boxes, preventing them from communicating. The only communication medium is that Alice can speak to Spock, and Bob can speak to Kirk.

Both Spock and Kirk toss a coin, and get either heads or tails, and tell the result to their participant (Spock to Alice, and Kirk to Bob). Alice and Bob must say 1 or $-1$, depending on the protocol and the result of their coins. The goal of Spock and Kirk is to make a protocol that allows to verify they are indeed psychics.

Alice and Bob are allowed to make a strategy before the game starts, but cannot communicate as soon as it started.

**Experiment 1** We consider the protocol where Alice and Bob must give the same answer if and only if the two coins gave the same result.

This is a bad protocol.

*Proof* Alice and Bob can just give 1 when they hear "Head", and $-1$ when they hear "Tail". This respects the protocol with probability 1, even though they do not need to be psychics.

□

**Experiment 2** We consider the protocol where Alice and Bob must give the same answer if and only if the two coins gave a different result.

This is again a bad protocol.

*Proof* Alice can do like in the first experiment, and Bob the other way around. This again, this respects the protocol with probability 1, even if they are not psychics.

□

**Experiment 3** We consider the protocol where Alice and Bob must give a different answer if and only if both coins give $H$.

Then:

1. If Alice and Bob only use classical physics, they have a $\frac{3}{4}$ probability of success and hence this is a good protocol.
2. However, if they share a Bell state $|\Phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$, they can get a higher probability of success (which is $\frac{2+\sqrt{2}}{4}$), and fool Kirk and Spock.

*Remark* Note that they are not transmitting any information, they just manage to get a higher correlation on their guesses. Moreover, this is what we call a "Bell inequality". We will formalise this right after, but the idea is that it shows hidden variables cannot explain quantum physics.

*Proof 1* We define $A_H\in\{1,-1\}$ to be what Alice says when she hears "Head", and similarly for $B_H,A_T,B_T$. Note that this means Alice and Bob have deterministic strategies. Using this deterministic case,

it is possible to extend this argument to the more general case where $(A_H, A_T)$ and $(B_H, B_T)$ are independent random variables, i.e. to the case where Alice and Bob have randomised strategies. The rules of the experiment ask for:

$$A_H B_H = -1, \quad A_T B_T = 1, \quad A_T B_H = 1, \quad A_T B_T = 1.$$

Multiplying all four equations together, we get:

$$A_H^2 B_H^2 A_T^2 B_T^2 = -1.$$

This is impossible, we always have $A_H^2 B_H^2 A_T^2 B_T^2 = 1$. Therefore, all equations cannot be fulfilled at the same time, and Alice and Bob cannot find a strategy that works all the time. The best they can do is fulfil three of those equations at any time. Since Kirk and Spock look at exactly one of those four equations with uniform probability, no matter their strategy, Alice and Bob cannot win more than $\frac{3}{4}$ of the time.

*Proof 2*

The idea is that quantum allows to make the random variables $(A_H, A_T)$ and $(B_H, B_T)$ dependent, without communicating. This can be considered to be some form of telepathy.

We consider the following protocol.

- If Alice gets told "Head", she measures in the $Z$-basis and says 1 on getting $|0\rangle$ and $-1$ on $|1\rangle$.
- If Alice gets told "Tail", she measures in the $X$-basis, and says 1 on getting $|+\rangle$ and $-1$ on $|-\rangle$.
- If Bob gets told "Head", he measures in the basis $\{|h\rangle, |\overline{h}\rangle\}$, and says 1 on getting $|h\rangle$ and $-1$ on $|\overline{h}\rangle$, where:

$$|h\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle + \cos\left(\frac{\pi}{8}\right)|1\rangle, \quad |\overline{h}\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle - \sin\left(\frac{\pi}{8}\right)|1\rangle.$$

- If Bob gets told "Tail", he measures in the basis $\{|t\rangle, |\overline{t}\rangle\}$, and says 1 on $|t\rangle$ and $-1$ on $|\overline{t}\rangle$, where:

$$|t\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle, \quad |\overline{t}\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle - \cos\left(\frac{\pi}{8}\right)|1\rangle.$$

We will check in the third exercise series that the probability they succeed is:

$$p = \cos\left(\frac{\pi}{8}\right)^2 = \frac{2+\sqrt{2}}{4} \approx 0.854,$$

hence beating the sceptic.

□

Monday 30th September 2024 — **Lecture 3 : Hidden variables are not a valid model**

**CHSH inequality**

We consider the following value, the total coalition coefficient:

$$C = \langle A_T B_T \rangle + \langle A_H B_T \rangle + \langle A_T B_H \rangle - \langle A_H B_H \rangle,$$

where, for instance, $\langle A_H B_H \rangle$ is the average value of the product of the two values Bob and Alice say if they both get head:

$$\begin{aligned}\langle A_H B_H \rangle = (-1)\mathbb{P}(A_H = 1, B_H = -1 \mid H, H) + (-1)\mathbb{P}(A_H = -1, B_H = 1 \mid H, H) \\ + \mathbb{P}(A_H = 1, B_H = 1 \mid H, H) + \mathbb{P}(A_H = -1, B_H = -1 \mid H, H).\end{aligned}$$

Then:

1. In the classical world, we always have $C \leq 2$.
2. In the quantum world, it is possible to reach $C = 2\sqrt{2}$.

*Remark* This is just a rephrasing of the third quantum psychics experiment above, in a more mathematical but less intuitive phrasing. There is a direct link between the probability of winning and the coalition coefficient $C$. Note however that this is typically phrased this way in literature, which we typically refer to as the CHSH inequality.

*Proof* Our goal is to link the probability of winning in the third experiment to the coalition coefficient. The inequalities will then come from our previous results.

We can now express their probability of winning, i.e. the probability that they give different answers if and only if both coins give head, is given by:

$$\begin{aligned}
&\mathbb{P}(\text{win}) \\
&= \frac{1}{4}\mathbb{P}(A_H = 1, B_H = -1 \mid H, H) + \frac{1}{4}\mathbb{P}(A_H = -1, B_H = 1 \mid H, H) \\
&\quad + \frac{1}{4}\mathbb{P}(A_H = 1, B_T = 1 \mid H, T) + \frac{1}{4}\mathbb{P}(A_H = -1, B_T = -1 \mid H, T) \\
&\quad + \frac{1}{4}\mathbb{P}(A_T = 1, B_H = 1 \mid T, H) + \frac{1}{4}\mathbb{P}(A_T = -1, B_H = -1 \mid T, H) \\
&\quad + \frac{1}{4}\mathbb{P}(A_T = 1, B_T = 1 \mid T, T) + \frac{1}{4}\mathbb{P}(A_T = -1, B_T = -1 \mid T, T).
\end{aligned}$$

Now note that, since the probabilities sum to one, we can write:

$$\begin{aligned}
\langle A_H B_H \rangle = 1 &- 2\mathbb{P}(A_H = 1, B_H = -1 \mid H, H) \\
&- 2\mathbb{P}(A_H = -1, B_H = 1 \mid H, H).
\end{aligned}$$

We can find something similar for all other terms. Doing some algebra, we can express the probability of winning as:

$$\begin{aligned}
\mathbb{P}(\text{win}) &= \frac{1}{2} + \frac{1}{8}(-\langle A_H B_H \rangle + \langle A_H B_T \rangle + \langle A_T B_H \rangle + \langle A_T B_T \rangle) \\
&= \frac{1}{2} + \frac{C}{8}.
\end{aligned}$$

In other words:

$$C = 8\mathbb{P}(\text{win}) - 4.$$

Let us now consider the classical and the quantum case.

1. In the classical case, we found that $\mathbb{P}(\text{win}) \leq \frac{3}{4}$. This gives that:
$$C \leq 8 \cdot \frac{3}{4} - 4 = 6 - 4 = 2,$$
as expected.
2. We found that it is possible to reach $\mathbb{P}(\text{win}) = \frac{2+\sqrt{2}}{4}$ in the quantum case, giving, in this case:
$$C = 8 \cdot \frac{2+\sqrt{2}}{4} - 4 = 4 + 2\sqrt{2} - 4 = 2\sqrt{2}.$$

□

**Bell's inequality** Let us be more formal and more general. We consider a very similar scenario, where we have some source that sets up the experience. Alice and Bob both have a measurement device that can measure in two different bases, $A, A'$ for Alice and $B, B'$ for Bob.

We take the following hypotheses, which intuition are explained after:

1. *(Setting independence)* We suppose that there exists some $\lambda$ so that:
$$\mathbb{P}(\ell \mid L, B, \lambda) = \mathbb{P}(\ell \mid L, B', \lambda).$$
where $L \in \{A, A'\}$ represents which measurement device was used by Alice.
2. *(Outcome independence)* We also suppose that the following is true:
$$\mathbb{P}(\ell \mid L, R, r, \lambda) = \mathbb{P}(\ell \mid L, R, r', \lambda).$$
where $L \in \{A, A'\}$ and $R \in \{B, B'\}$ represent which measurement device was used by Alice or Bob, respectively.
3. *(Single outcome assumption)* There is only one outcome obtained on each device.
4. *(No conspiracy assumption)* It is fair to calculate $C$ by averaging over many runs of the experiment. In other words, in some sense, there isn't a superior being tricking the probabilities giving us absurd statistical results.

We consider the following correlation coefficient:

$$C = |\langle LR\rangle - \langle LR'\rangle| + |\langle LR\rangle + \langle L'R\rangle|,$$

where, for instance:

$$\langle LR\rangle = \sum_{\ell,r\in\{-1,1\}} \ell r \mathbb{P}(\ell, r \mid L, R).$$

Then, $C \leq 2$. It is possible to make an experiment where $C = 2\sqrt{2}$, so one of the four hypotheses must be wrong.

*Hypotheses intuition*

Let us get some intuition on our hypotheses.

1. The first hypothesis states that Alice's result does not depend on Bob's choice of measurement device. She might get different probabilities depending on the results Bob gets, but it must depend on the result, not just the choice of the basis (similarly to what we discussed for using the delayed quantum eraser for signalling).
2. The second hypothesis is the most important one. It states that Alice's result does not depend on Bob's result, but only on some hidden variable $\lambda$, i.e. something that summarises all the shared history of the two qubits. In other words, we suppose that, when the source created the two qubits, it annotated some value $\lambda$ in it, that the qubits use on measurement. This supposes in some sense that the value is decided upon creation of the qubits, not on their measurement.
Another way to understand the first and second hypothesis is by noticing that that they together imply that:
$$\begin{aligned}\mathbb{P}(\ell, r \mid L, R, \lambda) &= \mathbb{P}(\ell \mid L, R, \lambda)\mathbb{P}(r \mid L, R, \lambda)\\ &= \mathbb{P}(\ell \mid L, \lambda)\mathbb{P}(r \mid R, \lambda),\end{aligned}$$
where we used first outcome independence and then setting independence. This property is named *factorisability*.
This property means that $\ell|L$ and $r|R$ are both *independent* random functions of the same $\lambda$. This is exactly the hypothesis that local hidden variable theories use: they suppose the behaviour of one qubit is not determined by the measurement of the other, but instead of some shared common $\lambda$.
Note that this does not mean $\ell|L$ and $r|R$ are not independent (just like the number of people in the subway in Lausanne and in Paris at any given time are very dependent random variables,

but one being big does not cause the other to be big, they are independent random functions of the time $t$).

3. The third hypothesis is quite unquestionable.
4. The fourth hypothesis basically states that we can estimate an expected value by averaging values obtained in a lab. Questioning this hypothesis would be a big issue for the scientific method.

*Implication* Since it is possible to make an experiment where $C = 2\sqrt{2}$, one of the four hypotheses must be wrong. We do not want to question the third or the fourth. Moreover, the first hypothesis would break the no-signaling theorem: Bob could choose $B$ or $B'$ depending on the value he wants to send to Alice.

This necessarily means that the second hypothesis, outcome independence, has to be violated. Note that this is more general than just quantum mechanics. No matter the physical model, the experiments we did that show $C = 2\sqrt{2}$ are valid. Therefore, outcome independence will never be possible, even in successors of quantum physics.

This notably shows that local hidden variables theory are not valid and will never be.

**Contextuality** Let us consider the following matrix of operators, named the Peres-Mermin square:

$$\begin{pmatrix} A & B & C \\ a & b & c \\ \alpha & \beta & \gamma \end{pmatrix}.$$

We moreover consider the following value:

$$\langle PM \rangle = \langle ABC \rangle + \langle abc \rangle + \langle \alpha\beta\gamma \rangle + \langle Aa\alpha \rangle + \langle Bb\beta \rangle - \langle Cc\gamma \rangle.$$

In the classical case, we have $|C| \leq 4$. With quantum, it is possible to choose the operators so that $C = 6$.

*Remark* This is a result which is very analogous to Bell's inequalities. The fact that $|C| \leq 4$ in the classical case means that the measure of some property does not depend on its context, i.e. it does not change because of the measure of other properties.

The fact that $C = 6$ is reachable with quantum means that this is not true for quantum physics.

*Proof* We only prove the quantum case, the classical case can be argued in a very similar way to the quantum psychic game. We consider the following Peres-Mermin square:

$$\begin{pmatrix} Z \otimes I & I \otimes Z & Z \otimes Z \\ I \otimes X & X \otimes I & X \otimes X \\ Z \otimes X & X \otimes Z & Y \otimes Y \end{pmatrix}.$$

Notice that all elements in a single row and all elements in a single column commute and are therefore jointly measurable. Computing one of the terms, we notice:

$$\langle ABC \rangle = \langle \psi | (Z \otimes I)(I \otimes Z)(Z \otimes Z) | \psi \rangle = \langle \psi | \psi \rangle = 1.$$

The product of every row and column gives identity, except for the last column which gives $-I$. This means that:

$$\langle PM \rangle = 1 + 1 + 1 + 1 + 1 - (-1) = 6.$$

□

# Chapter 4

# Density operators

## 4.1 Pure states

**Definition: Density operator**

Let $|\psi\rangle$ be some state.
Its corresponding **density operator** is $\rho = |\psi\rangle\langle\psi|$.

*Example*

For instance, if $|\psi\rangle = |1\rangle = \begin{pmatrix}0 & 1\end{pmatrix}^T$, then:

$$\rho = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

Similarly:

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\\0\end{pmatrix} \implies \rho = \frac{1}{2}\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 1 & -1 & 0\\ 0 & -1 & 1 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

Finally, for an arbitrary qubit $|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle$:

$$\rho = \begin{pmatrix}\cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{-i\varphi}\\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)e^{i\varphi} & \sin^2\left(\frac{\theta}{2}\right)\end{pmatrix}.$$

**Proposition: Expected value**

Let $\rho$ be a density operator, and $O$ be an observable.
Then, the expected value of $O$ under the state $\rho$ is defined by:

$$\langle O\rangle = \mathrm{Tr}(\rho O).$$

*Proof*

This is a direct proof, using properties of the trace:

$$\langle O\rangle = \mathrm{Tr}(\rho O) = \mathrm{Tr}(|\psi\rangle\langle\psi|O) = \mathrm{Tr}(\langle\psi|O|\psi\rangle) = \langle\psi|O|\psi\rangle.$$

□

**Theorem: Arbitrary single qubit state**

Let $|\psi\rangle$ be an arbitrary qubit, and $\rho$ be its corresponding density matrix. Moreover, let $\vec{\boldsymbol{v}}$ be the unit vector on the Bloch sphere that represents $|\psi\rangle$.
Then:

$$\vec{\boldsymbol{v}} = \begin{pmatrix}\langle X\rangle\\ \langle Y\rangle\\ \langle Z\rangle\end{pmatrix}, \quad \rho = \frac{1}{2}(I + \vec{\boldsymbol{\sigma}} \bullet \vec{\boldsymbol{v}}).$$

*Proof*

Since $\rho$ is a Hermitian matrix, we can represent it in the Pauli basis:

$$\rho = u_0 I + u_x X + u_y Y + u_z Z.$$

Since moreover $\rho = |\psi\rangle\langle\psi|$, we notice that $\mathrm{Tr}(\rho) = 1$. Therefore, we must have:

$$1 = \mathrm{Tr}(\rho) = u_0 \,\mathrm{Tr}(I) + u_x \,\mathrm{Tr}(X) + u_y \,\mathrm{Tr}(Y) + u_z \,\mathrm{Tr}(Z) = 2u_0$$
$$\iff u_0 = \frac{1}{2}.$$

We now need to get the other coefficients. Let us consider the following expectation value:

$$\langle X\rangle = \mathrm{Tr}(\rho X) = \mathrm{Tr}\left(\frac{1}{2}X + u_x X^2 + u_y YX + u_z ZX\right) = u_x \,\mathrm{Tr}(I).$$

Since $\mathrm{Tr}(I) = 2$, this means that $u_x = \frac{1}{2}\langle X\rangle$. We can do a similar reasoning for the other coefficients, getting:

$$\rho = \frac{1}{2}(I + \vec{\boldsymbol{\sigma}} \bullet \vec{\boldsymbol{v}}), \quad \text{where } \vec{\boldsymbol{v}} = \begin{pmatrix}\langle X\rangle \\ \langle Y\rangle \\ \langle Z\rangle\end{pmatrix}.$$

We will moreover verify in the fourth exercise sheet that:

$$\vec{\boldsymbol{v}} = \begin{pmatrix}\sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta)\end{pmatrix}.$$

So, $\vec{\boldsymbol{v}}$ is indeed the vector representing $|\psi\rangle$ and $\rho$ on the Bloch sphere .

□

## 4.2 Mixed states

**Classical randomness**

We want to add classical randomness in our constructions. Let's say that someone gives us a state, $|\psi\rangle$ with probability $p$ and $|\varphi\rangle$ with probability $1-p$. Our goal is to find a single mathematical object that allows us to easily get the expected value. The expected value is:

$$\begin{aligned}\langle O\rangle &= p\langle\psi|O|\psi\rangle + (1-p)\langle\varphi|O|\varphi\rangle \\ &= p\,\mathrm{Tr}(|\psi\rangle\langle\psi|O) + (1-p)\,\mathrm{Tr}(|\varphi\rangle\langle\varphi|O) \\ &= \mathrm{Tr}[(p|\psi\rangle\langle\psi| + (1-p)|\varphi\rangle\langle\varphi|)O].\end{aligned}$$

The mathematical entity that appears to have all the information about the expected value is therefore:

$$\rho = p|\psi\rangle\langle\psi| + (1-p)|\varphi\rangle\langle\varphi|.$$

This is also a density operator, but related to what we call a mixed state. This yields the following definition.

**Definition: Density operator**

Let's say we have a set of $m$ states $|\psi_k\rangle$, which appear with probability $p_k$. We can represent this using a **density operator**, defined by:

$$\rho = \sum_{k=1}^{m} p_k |\psi_k\rangle\langle\psi_k|.$$

If $m = 1$ (i.e. if $\rho = |\psi\rangle\langle\psi|$), $\rho$ is said to be a **pure state**. Otherwise, it is said to be a **mixed state**.

**Theorem: Bloch ball**

Let's say that we have states $|\psi\rangle$ and $|\varphi\rangle$ corresponding to Bloch sphere vectors $\vec{\boldsymbol{v}}_\psi$ and $\vec{\boldsymbol{v}}_\varphi$, respectively:

$$|\psi\rangle\langle\psi| = \frac{1}{2}(I + \vec{\boldsymbol{v}}_\psi \bullet \vec{\boldsymbol{\sigma}}), \quad |\varphi\rangle\langle\varphi| = \frac{1}{2}(I + \vec{\boldsymbol{v}}_\varphi \bullet \vec{\boldsymbol{\sigma}}).$$

Moreover, suppose that we have $|\psi\rangle$ with probability $p$ and $|\varphi\rangle$ with probability $1-p$. Then, $\rho = p|\psi\rangle\langle\psi| + (1-p)|\varphi\rangle\langle\varphi|$ has Block *ball* vector $\vec{\boldsymbol{u}}$, where:

$$\vec{\boldsymbol{u}} = p\vec{\boldsymbol{v}}_\psi + (1-p)\vec{\boldsymbol{v}}_\varphi.$$

*Intuition*

In other words, $\vec{\boldsymbol{u}}$ is a convex combination of $\vec{\boldsymbol{v}}_\psi$ and $\vec{\boldsymbol{v}}_\varphi$; it is on the line connecting those two vectors. Thus, whereas those two vectors are on the surface of the Bloch sphere, $\vec{\boldsymbol{u}}$ is inside of it. This is hence not a sphere, it is a ball.

This can be generalised $\rho$ being a linear combination of an arbitrary of pure states.

*Property*

Any pure state is on the Bloch sphere, $\|\vec{\boldsymbol{v}}\| = 1$. Any mixed state is inside the ball, $\|\vec{\boldsymbol{v}}\| < 1$.

*Proof*

We directly find that:

$$\begin{aligned}\rho &= p|\psi\rangle\langle\psi| + (1-p)|\varphi\rangle\langle\varphi| \\ &= \frac{1}{2}p(I + \vec{\boldsymbol{v}}_\psi \bullet \vec{\boldsymbol{\sigma}}) + \frac{1}{2}(1-p)(I + \vec{\boldsymbol{v}}_\varphi \bullet \vec{\boldsymbol{\sigma}}) \\ &= \frac{1}{2}[(p + (1-p))I + (p\vec{\boldsymbol{v}}_\psi + (1-p)\vec{\boldsymbol{v}}_\varphi) \bullet \vec{\boldsymbol{\sigma}}] \\ &= \frac{1}{2}(I + \vec{\boldsymbol{u}} \bullet \vec{\boldsymbol{\sigma}}).\end{aligned}$$

□

Monday 7th October 2024 — **Lecture 4 : Philosophy feat. Outer Wilds**

**Definition: Thermal state**

In thermodynamics, for a given Hamiltonian $H = \sum_i E_i|E_i\rangle\langle E_i|$, the thermal state $\gamma$ is defined by:

$$\gamma = \frac{e^{-\beta H}}{Z}, \quad Z = \sum_i e^{-\beta E_i},$$

where $Z$ is a normalisation term and $\beta$ is the inverse temperature (multiplied by some constant).

*Observation*

Note that $\gamma$ is a mixed state. Moreover, $Z$ is indeed a normalisation factor, since:

$$Z = \sum_i e^{-\beta E_i} = \mathrm{Tr}\big(e^{-\beta H}\big) \implies \mathrm{Tr}(\gamma) = 1.$$

*Intuition*

In a thermal distribution, we are in state $|E_i\rangle$ with probability $p_i = e^{-\beta E_i}/Z$. This lets us construct $\gamma$:

$$\gamma = \sum_i p_i|E_i\rangle\langle E_i| = \sum_i \frac{e^{-\beta E_i}}{Z}|E_i\rangle\langle E_i| = \frac{e^{-\beta H}}{Z},$$

since $\exp(A)$ has the same eigenvectors as $A$, but with eigenvalues $\exp(a)$ instead of $a$.

This is a good example of a physical system that requires both classical and quantum randomness, justifying an interest for mixed states.

**Properties of density operators**

Let $\rho$ be some operator.
$\rho$ is a density operator if and only if:

1. *(Normalisation constraint)* $\mathrm{Tr}(\rho) = 1$.
2. *(Positive semi-definite)* $\rho$ it has positive eigenvalues.
3. *(Hermitian)* $\rho^\dagger = \rho$.

*Proof* $\implies$

The first property directly comes from the fact that eigenvalues of a density operator represents probabilities. They must therefore sum to 1. The second property follows from a similar argument.
For the third one, we directly find that:

$$\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right)^\dagger = \sum_i p_i^* (|\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho$$

since, by a similar argument to the first properties, the eigenvalues are real.

□

*Proof* $\impliedby$

Let $\rho$ be an arbitrary operator respecting the three properties. Since it is Hermitian, we know it is orthodiagonalisable in a real basis:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad p_i \in \mathbb{R}.$$

Since it is positive semi-definite, $p_i \geq 0$ for all $i$. Since moreover it traces to 1:

$$\sum_i p_i = \mathrm{Tr}(\rho) = 1.$$

This means that the $p_i$ represent a valid probability distribution. This indeed yields that it is a valid density operator, where we have state $|\psi_i\rangle$ with probability $p_i$.

*Remark*

The fact that the Bloch vector has a norm less than 1 directly comes from those properties.

## 4.3 Reduced states

**Observation**

So far, we have imagined that we have a state $|\psi\rangle_A$ in a system $A$ and state $|\varphi\rangle_B$ in system $B$. We then wrote down their joint state as $|\psi\rangle_A \otimes |\varphi\rangle_B$.
Now, we may want to go the other direction. We may have a state $|\Psi\rangle_{AB}$, and want to write down the state corresponding to the system $A$. Note that the sate may not even be a product state, we may for instance have $|\Psi\rangle_{AB} = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$. In the most general case, we may even have $\rho_{AB}$ to be a mixed state.
Given $\rho_{AB}$, our goal is therefore to find some mathematical entity $\rho_A$ that allows us to compute properties of the system $A$ only, i.e. so that:

$$\langle O_A \otimes I_B\rangle_{AB} = \mathrm{Tr}(\rho_{AB}(O_A \otimes I_B)) = \mathrm{Tr}(\rho_A O) = \langle O_A\rangle_A.$$

**Theorem: Reduced state**

Let $\rho_{AB}$ be a mixed state in system $\mathcal{H}_A \otimes \mathcal{H}_B$.
We define the state $\rho_A$, called a **reduced state**, as:

$$\rho_A = \sum_{k=1}^{d_B} (I \otimes \langle k|)\rho_{AB}(I \otimes |k\rangle).$$

Then, for all observable $O$ acting on $\mathcal{H}_A$:

$$\mathrm{Tr}(\rho_A O) = \langle O\rangle.$$

*Proof*

We make a constructive proof. In other words, we try to find $\rho_A$ from the result we wish to obtain.

Let us first recap what an expected value is. Expressing $O$ in its eigenbasis, $O = \sum_{j=1}^{d_A} \lambda_j |\lambda_j\rangle\langle\lambda_j|$, we have:

$$\langle O\rangle = \sum_{j=1}^{d_A} \lambda_j \mathbb{P}_A(\lambda_j),$$

where $\mathbb{P}_A(\lambda_j)$ is the probability to measure $\lambda_j$ on system $A$ in $\rho_{AB}$. We are going to express $\mathbb{P}_A(\lambda_j)$ in terms of $\rho_{AB}$. Let $\{|k\rangle_B\}$ be some arbitrary orthonormal basis of $\mathcal{H}_B$. We know that:

$$\mathbb{P}_{AB}(\lambda_j, k) = \mathrm{Tr}(|\lambda_j k\rangle\langle\lambda_j k|\rho_{AB}) = \langle\lambda_j k|\rho_{AB}|\lambda_j k\rangle.$$

This can be verified by the definition of mixed states. Another way to see this is that $\mathbb{P}_{AB}(\lambda_j, k)$ is the expected value of the indicator random variable that we measure $\lambda_j, k$. This can just be represented using the observable $|\lambda_j k\rangle\langle\lambda_j k|$.

But then, using the marginalisation of joint probabilities:

$$\mathbb{P}_A(\lambda_j) = \sum_{k=1}^{d_B} \mathbb{P}_{AB}(\lambda_j, k) = \sum_{k=1}^{d_B} \langle\lambda_j k|\rho_{AB}|\lambda_j k\rangle.$$

This gives us that:

$$\begin{aligned}
\langle O\rangle &= \sum_{j=1}^{d_A} \lambda_j \mathbb{P}_A(\lambda_j) \\
&= \sum_{j=1}^{d_A} \lambda_j \sum_{k=1}^{d_B} (\langle\lambda_j| \otimes \langle k|)\rho_{AB}(|\lambda_j\rangle \otimes |k\rangle) \\
&= \sum_{j=1}^{d_A} \lambda_j \langle\lambda_j| \left( \sum_{k=1}^{d_B} (I \otimes \langle k|)\rho_{AB}(I \otimes |k\rangle) \right) |\lambda_j\rangle,
\end{aligned}$$

We can now consider the $\rho_A$ defined in the hypotheses, by seeing that this directly gives us the result we swish for:

$$\langle O\rangle = \sum_{j=1}^{d_A} \lambda_j \langle\lambda_j|\rho_A|\lambda_j\rangle = \mathrm{Tr}\left( \sum_{j=1}^{d_A} \lambda_j |\lambda_j\rangle\langle\lambda_j|\rho_A \right) = \mathrm{Tr}(O\rho_A)$$

□

**Definition: Partial trace**

Let $\rho_{AB}$ be some mixed state in $\mathcal{H}_A \otimes \mathcal{H}_B$.

We define the **partial trace** over $\mathcal{H}_B$ as:

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_{k=1}^{d_B} (I \otimes \langle k|)\rho_{AB}(I \otimes |k\rangle).$$

*Remark*

Note that this is exactly the state we had in the theorem above. In other words, for all observable $O$:

$$\mathrm{Tr}(O\rho_A) = \mathrm{Tr}((O \otimes I_B)\rho_B).$$

*Intuition*

If we were doing the full trace, we would be computing:

$$\mathrm{Tr}(\rho_{AB}) = \sum_{jk} (\langle j| \otimes \langle k|)\rho_{AB}(|j\rangle \otimes |k\rangle).$$

In our case, we only trace on one of the systems (replacing the $|j\rangle$ by identities) getting an operator instead of a scalar.

**Properties**

The partial trace has the following properties:

1. $\mathrm{Tr}_B(O_A \otimes O_B) = O_A \,\mathrm{Tr}(O_B)$.
2. $\mathrm{Tr}_B(|ij\rangle\langle k\ell|) = |i\rangle\langle k| \,\mathrm{Tr}(|j\rangle\langle \ell|) = |i\rangle\langle k|\langle \ell|j\rangle$.
3. *(Cyclicity)* $\mathrm{Tr}_B(M_A N_A \otimes M_B N_B) = \mathrm{Tr}_B(M_A N_A \otimes N_B M_B)$.
4. *(Linearity)* $\mathrm{Tr}_B\left(\sum_i \lambda_i O_i\right) = \sum_i \lambda_i \,\mathrm{Tr}_B(O_i)$.

*Remark 1* It is not necessary to use the formal definition of the partial trace. Those properties are typically enough.

*Remark 2* For instance, the first property directly implies that the mixed reduced state of $\rho_A \otimes \rho_B$ would be:

$$\mathrm{Tr}_B(\rho_A \otimes \rho_B) = \rho_A \,\mathrm{Tr}(\rho_B) = \rho_A.$$

This is very intuitive: the mathematical entity that represents just system $A$ should indeed be $\rho_A$.

*Proof* Proving those properties is a good exercise.

**Example**

Let us consider a Bell state:

$$|\psi_{AB}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

Its corresponding mixed state is:

$$\rho_{AB} = |\Psi_{AB}\rangle\langle\Psi_{AB}| = \frac{|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|}{2}.$$

Then, the reduced state on system $A$ is:

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(\rho_{AB}) \\
&= \frac{\mathrm{Tr}_B(|00\rangle\langle 00|) + \mathrm{Tr}_B(|00\rangle\langle 11|) + \mathrm{Tr}_B(|11\rangle\langle 00|) + \mathrm{Tr}_B(|11\rangle\langle 11|)}{2} \\
&= \frac{|0\rangle\langle 0| \cdot \langle 0|0\rangle + |0\rangle\langle 1| \cdot \langle 0|1\rangle + |1\rangle\langle 0| \cdot \langle 1|0\rangle + |1\rangle\langle 1| \cdot \langle 1|1\rangle}{2} \\
&= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\
&= \frac{I}{2}.
\end{aligned}$$

This is a mixed state, with probability $\frac{1}{2}$ to have $|0\rangle$ and probability $\frac{1}{2}$ to have $|1\rangle$; this is just a classical coin toss. Here, we lost coherence: we have a classical mixture instead of quantum state in superposition.

*Remark 1* $\rho_A = \frac{I}{2}$ is named the maximally mixed state, it is at the centre of the Bloch sphere.

*Remark 2* We get the exact same result for of the four Bell states. Since multiple mixed state yield the same reduced state, the partial trace is not bijective.

*Remark 3* It is possible to find by symmetry that $\rho_B = \frac{I}{2}$ as well. Note however that $\rho_{AB} \neq \rho_A \otimes \rho_B$. In other words, when looking at only one of the two systems, this is just a classical coin toss; but, as soon as we

wish to analyse both systems at the same time, quantum physics adds some correlation.

**No-signaling theorem**

We suppose that Alice and Bob share many entangled qubit pairs.

There is no way for Bob to signal anything (i.e. transmit data) to Alice by applying operators and measuring (in any basis) its part of the pairs.

*Proof* The reduced mixed state is the unique mathematical object that determines the expected value of Alice's measure. It however does not depend on anything Bob does on its part of the qubit, giving our result.

□

# Chapter 5

# Measurement problem

**Axioms**

We will consider the following hypotheses:

1. A good measuring device is accurate.
2. Quantum mechanics is a universal (it applies to everything) and fundamental theory.
3. *(Weak physicalist postulate)* The description of the behaviour of large objects must be consistent with the laws governing the behaviour of the smaller objects of which they consist.

The two last hypotheses imply that we should be able to describe a measuring device using the rules of quantum mechanics.

**Measuring device**

A measuring device is a device which can extract information from a quantum system.

*Example*

For instance, we may have a system $|x\rangle \otimes |\psi\rangle$ where $|x\rangle \in \{|\text{ready}\rangle, |\text{up}\rangle, |\text{down}\rangle\}$ represents the state of the measuring device, and $|\psi\rangle$ represents the state of some electron. We want our measurement device to be such that:

$$|\text{ready}\rangle \otimes |\uparrow\rangle \to |\text{up}\rangle \otimes |\uparrow\rangle,$$

$$|\text{ready}\rangle \otimes |\downarrow\rangle \to |\text{down}\rangle \otimes |\downarrow\rangle,$$

**Problem**

The issue is that time evolution is linear, but doing a measure makes a projection on the state. Projections are not linear.

*Example*

For our example before, if we try to measure the state $|\psi\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$, our measurement device will be such that, by linearity:

$$|\text{ready}\rangle \otimes \frac{|\uparrow\rangle + |\downarrow\rangle}{\sqrt{2}} \to \frac{|\text{up}\rangle \otimes |\uparrow\rangle + |\text{down}\rangle \otimes |\downarrow\rangle}{\sqrt{2}}$$

This is however not how quantum mechanics predicts a measurement to happen.

*Implication*

Since we have a contradiction, we must drop one of our assumptions. Not all quantum mechanics formalism requires that a state collapse to some specific state after we measure it. This however does not fix everything, since it would still contradict our macroscopic world. Similarly, it seems to break the Born rule, which states we are able to measure something with a given probability.

**Measurement problem**

The measurement problem is thus to try and reconcile the concept of measurement and time evolution. This is a very philosophical question, to which there is no definite answer. Let us consider two potential solutions, and their criticisms.

**Solution: Collapse postulate**

A way to solve the problem is by considering that there are, in fact, two fundamental laws in quantum mechanics:

- When no measurement is going on, states evolve linearly.
- When a measurement is going on, states evolve according to the postulate of collapse.

*Issues*

The big issue is then the definition of measurement. A photon bouncing off an atom effectively measures an atom, and similarly seeing Schrödinger's cat dead is a measurement as well. So, we need to trace a line between what is a measurement and what is time evolution. This is not precise enough to be a fundamental role in the laws of physics.

*Personal remark*

Fun fact, this is the solution chosen by Outer Wilds. In this game, we define measurements relative to "conscious" beings; if a conscious being looks at a quantum object, then it is projected to one of its possibilities. Naturally, this raises many more questions than it solves (good luck defining consciousness physically; Conway's free will theorem for instance states that, under some axioms, we may suppose electrons have a free will), but this was just a good occasion for me to tell you to play this game.::)

**Solution: Decoherence**

Another way to solve this problem is is to add a term in the superposition, representing the environment. The environment then acts as some form of "sink", which absorbs information from the quantum system.

Let us consider the usual example to explain this, supposing we want to measure $|\psi\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$. We then postulate that the environment starts in some state $|E_0\rangle$, and that it then evolves differently whether the electron is in $|\uparrow\rangle$ or $|\downarrow\rangle$, i.e. that our measurement device is such that:

$$|E_0\rangle \otimes |\text{ready}\rangle \otimes \frac{|\uparrow\rangle + |\downarrow\rangle}{\sqrt{2}} \to \frac{|E_\uparrow\rangle \otimes |\text{up}\rangle \otimes |\uparrow\rangle + |E_\downarrow\rangle \otimes |\text{up}\rangle \otimes |\downarrow\rangle}{\sqrt{2}}.$$

This is a linear evolution, which can be explained with a physical Hamiltonian. Leaving $r = \langle E_\uparrow | E_\downarrow \rangle$, the reduced state corresponding to the measurement device and the electron (i.e. the partial trace of the density matrix corresponding to the evolved state) is given by:

$$\rho_{decoh} = \frac{|\text{up}\uparrow\rangle\langle\text{up}\uparrow| + |\text{down}\downarrow\rangle\langle\text{down}\downarrow| + r|\text{up}\uparrow\rangle\langle\text{down}\downarrow| + r^*|\text{down}\downarrow\rangle\langle\text{up}\uparrow|}{2}.$$

On the other hand, the Born rule tells us that, when measure $|\psi\rangle = \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}$, we should get $|\text{up}\uparrow\rangle$ with probability $\frac{1}{2}$, and $|\text{down}\downarrow\rangle$ with probability $\frac{1}{2}$, i.e. that measurement can be represented by the statistical mixture $\rho_{born}$, where:

$$\rho_{born} = \frac{|\text{up}\uparrow\rangle\langle\text{up}\uparrow| + |\text{down}\downarrow\rangle\langle\text{down}\downarrow|}{2}.$$

We see that $\rho_{decoh} \to \rho_{born}$ as $r \to 0$, which can again be done physically. This seems to make the Born rule coherent.

*Criticism*

Saying that $\rho_{decoh} \to \rho_{born}$ does not really make sense physically. $\rho_{decoh}$ is a reduced state, whereas $\rho_{born}$ is a statistical mixture. They are represented by the same mathematical object, and have a similar intuition, but they are fundamentally different.

More precisely, we define a mixed state $\rho$ that arises from ignorance of an underlying pure state (such as a statistical mixture) to be a **proper mixture**; and a mixed state that arises because we are considering a subsystem to be a **improper mixture**. Those two ob-

jects are mathematically the same, but physically and conceptually different.

*Remark* We will see in the fifth exercise series an exercise that puts more maths on all this reasoning.

# Chapter 6

# Identical multi-particle systems

## 6.1 Two identical particles

**Goal**

Let's say that we have two identical particles. Their joint wave-function is constrained in some way, and we are curious how.

*Remark* The following argument may seem fiddly, but textbooks do it in a similar way. Making a bulletproof argument would require a lot more explanations.

**Lemma**

Let $\psi(\vec{r}_1, \vec{r}_2)$ be the wavefunction of a system with two identical particles, of position $\vec{r}_1$ and $\vec{r}_2$ respectively.
Then, there is some $s \in \{1, -1\}$ so that, for all $\vec{r}_1, \vec{r}_2$:

$$\psi(\vec{r}_1, \vec{r}_2) = s\psi(\vec{r}_2, \vec{r}_1).$$

*Proof* Since particles are identical, the probability to the first at $\vec{r}_1$ and the second at $\vec{r}_2$ should be exactly the same as the probability to find the first at $\vec{r}_2$ and the second at $\vec{r}_1$, i.e.:

$$|\psi(\vec{r}_1, \vec{r}_2)|^2 = |\psi(\vec{r}_2, \vec{r}_1)|^2.$$

This implies that they are equal up to the phase, i.e. that there exists some $\varphi \in [0, 2\pi]$ such that:

$$\psi(\vec{r}_1, \vec{r}_2) = e^{i\varphi}\psi(\vec{r}_2, \vec{r}_1).$$

It is possible to argue that this $\varphi$ is the same for all $\vec{r}_1, \vec{r}_2$. The idea is that we suppose the universe is invariant under reflections (more generally, we suppose the universe is isotropic), so if $\psi(\vec{r}, -\vec{r}) = e^{i\varphi}\psi(-\vec{r}, \vec{r})$ for some $\vec{r}$, then we must also have $\psi(-\vec{r}, \vec{r}) = e^{i\varphi}\psi(\vec{r}, -\vec{r})$. By changing our coordinate system so that the origin is at the middle of $\vec{r}_1$ and $\vec{r}_2$, we can always express $(\vec{r}_1, \vec{r}_2)$ as $(\vec{r}, -\vec{r})$.
In particular, since this true for all $\vec{r}_1, \vec{r}_2$, this is also true for $\vec{r}_2, \vec{r}_1$:

$$\psi(\vec{r}_2, \vec{r}_1) = e^{i\varphi}\psi(\vec{r}_1, \vec{r}_2).$$

Combining the two:

$$\psi(\vec{r}_1, \vec{r}_2) = e^{2i\varphi}\psi(\vec{r}_1, \vec{r}_2) \implies e^{2i\varphi} = 1 \implies e^{i\varphi} = \pm 1.$$

This does yield our result:

$$\psi(\vec{r}_1, \vec{r}_2) = \pm\psi(\vec{r}_2, \vec{r}_1).$$

□

*Remark* This reasoning can be generalised so that it not only applies to positions, but also to any vector $\vec{\boldsymbol{x}}$ representing the variables associated with the particles (such as generalised position or momentum).

**Definition: Permutation operator**

We define the **permutation operator** $P_{12}$ so that:

$$P_{12}\psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) = \psi(\vec{\boldsymbol{x}}_2, \vec{\boldsymbol{x}}_1).$$

*Remark* We take the natural assumption that permuting the vectors twice takes us back where we started:

$$P_{12}^2\psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) = \psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2).$$

Note that this is a bit fiddly since, as we are working on the wave function, there could be some global phase. This is however natural, so we admit it. This implies that $P_{12}^2 = I$.

*Personal remark* This is strongly linked with the symmetric group, which we will use a lot. If the reader wants a refresher on this group, I invite them to take a look at my Algebra notes, available on my GitHub:

`https://github.com/JoachimFavre/UniversityNotes`

**Definition: Bosons and Fermions**

By our lemma, we have:

$$P_{12}\psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) = \psi(\vec{\boldsymbol{x}}_2, \vec{\boldsymbol{x}}_1) = \pm\psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2).$$

In other words, there are two different types of identical particles:

- Those that, when we swap, we get a $+1$. They are called **Bosons**.
- Those that, when we swap, we get a $-1$. They are called **Fermions**.

*Example* For instance, photons and gluons are Bosons; and electrons and quarks are Fermions. Experimentally, we notice that Bosons have integer spin, and Fermions have half-integer spin.

**Pauli exclusion principle**

We consider two identical Fermions.
They cannot be found at the same position.

*Proof* Since they are at the same place and have the same properties, their $\vec{\boldsymbol{x}}$ are exactly equal. So, we have:

$$\psi(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}) = -\psi(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}) \implies \psi(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}) = 0.$$

But then:

$$|\psi(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}})|^2 = 0.$$

So, the probability of finding them at the same place is indeed zero.

□

*Remark* If any of the properties differ between the two Fermions, then they are no longer identical and they can be found at the same position.

**Notations**

We will now write $x = \vec{\boldsymbol{x}}$, to simplify the notation.

**Proposition**

We consider an arbitrary two-particle wavefunction:

$$|\psi\rangle = \sum_{x,x'} a_{x,x'}|x, x'\rangle.$$

If the two particles are the same, and they are a Fermion, then:

$$a_{x,x'} = -a_{x',x}, \quad \forall x, x'.$$

*Example* Consider two Fermions that can be in state $|0\rangle$ or $|1\rangle$. Then, we necessarily have:

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

*Proof* We notice that:

$$|\psi\rangle = \sum_{x,x'} a_{x,x'}|x, x'\rangle = \sum_{x,x'} a_{x,x'}(-1)|x', x\rangle.$$

Now, in the sum, $x$ and $x'$ are just dummy variables. We are free to relabel them as $(x, x') \leftarrow (x', x)$. This yields:

$$|\psi\rangle = \sum_{x,x'} \underbrace{(-1)a_{x',x}}_{=a_{x,x'}}|x, x'\rangle.$$

By inspection, this means that:

$$a_{x,x'} = (-1)a_{x',x} = -a_{x',x}.$$

□

**Proposition** We consider an arbitrary two-particle wavefunction:

$$|\psi\rangle = \sum_{x,x'} a_{x,x'}|x, x'\rangle.$$

If the two particles are the same, and they are a Boson, then:

$$a_{x,x'} = a_{x',x}, \quad \forall x, x'.$$

*Example* Consider two Bosons that can be in state $|0\rangle$ or $|1\rangle$. Then, we can have any linear combination of the following states:

$$\frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |00\rangle, \quad |11\rangle.$$

In other words, their general state is of the form:

$$\frac{\alpha}{\sqrt{2}}(|01\rangle + |10\rangle) + \beta|00\rangle + \gamma|11\rangle.$$

*Proof* The proof is completely similar to the case of Fermions.

□

## 6.2 Multiple identical particles

**Generalisation** Let $\psi(x_1, \ldots, x_n)$ be the wavefunction of a system with $n$ identical particles. Then, there is some $s \in \{-1, 1\}$ so that, for all $j \neq k$:

$$\begin{aligned} P_{jk}\psi(x_1, \ldots, x_j, \ldots, x_k, \ldots, x_n) &= \psi(x_1, \ldots, x_k, \ldots, x_j, \ldots, x_n) \\ &= s\psi(x_1, \ldots, x_j, \ldots, x_k, \ldots, x_n). \end{aligned}$$

*Intuition* By permuting any two identical particles of an $n$ particle system, we get a $+1$ sign for Bosons and $-1$ sign for Fermions.

*Proof* We can use the exact same derivation as before to find that, for all $j, k$, there is some $s_{j,k} \in \{-1, 1\}$ (which may depend on $j$ on $k$) so that:

$$P_{jk}\psi(x_1, \dots, x_n) = s_{jk}\psi(x_1, \dots, x_n).$$

Our goal is thus to show that, in fact, $s_{jk}$ does not depend on $j, k$. To do so, we notice that:

$$P_{jk} = P_{1j}P_{2k}P_{12}P_{2k}P_{1j}.$$

Indeed, using the orbital notation of the permutation group:

$$(jk) = (1j)(2k)(12)(2k)(1j).$$

So, this means that, for all $j, k$:

$$\begin{aligned} s_{jk}\psi(x_1, \dots, x_n) &= P_{jk}\psi(x_1, \dots, x_n) \\ &= P_{1j}P_{2k}P_{12}P_{2k}P_{1j}\psi(x_1, \dots, x_n) \\ &= s_{1j}s_{2k}s_{12}s_{2k}s_{1j}\psi(x_1, \dots, x_n) \\ &= s_{12}\psi(x_1, \dots, x_n), \end{aligned}$$

since $s_{1j}^2 = s_{2k}^2 = 1$. By inspection, this means that $s_{jk} = s_{12}$ for all $j, k$, and thus indeed that $s_{jk}$ does not depend on $j, k$.

□

**Properties of the permutation operator**

Permutation operators have the following properties:

1. $P_{jk} = P_{kj}$. Indeed, swapping $j$ and $k$ is the same as swapping $k$ and $j$.
2. $P_{jk}^{-1} = P_{jk}$, since $P_{jk}^2 = I$.
3. They are Hermitian, i.e. $P_{jk}^\dagger = P_{jk}$.

**Property**

Let $|\psi_{12}\rangle$ be a state corresponding to two identical particles, and $O$ be an arbitrary observable.

Then:

$$\langle\psi_{12}|O|\psi_{12}\rangle = \langle\psi_{12}|P_{12}^\dagger O P_{12}|\psi_{12}\rangle.$$

*Proof* We can show this using the property of the permutation operator, but we will consider a more physical proof. Since the two particles are identical, swapping them should not change anything to the result. Therefore, those two expected value must be equal.

□

*Implication* In particular, this means that, for any $O$:

$$O = P_{12}^\dagger O P_{12} \iff P_{12}O = OP_{12} \iff [P_{12}, O] = 0,$$

where we used $P_{12}^\dagger = P_{12} = P_{12}^{-1}$.

In particular, this means that the permutation operator permutes with the Hamiltonian of identical particles. Commutation is another way to show symmetries of a system, and this will be formalised with representation theory presented later in the class.

**State of $n$ Bosons**

A basis for the state of $n$ Bosons is given by:

$$\left\{ \psi_{\vec{\boldsymbol{x}}} = \mathcal{N} \sum_{\mathbb{P} \in S_n} \mathbb{P}|x_1, \dots, x_n\rangle = \mathcal{N} \sum_{\mathbb{P} \in S_n} \left|x_{\mathbb{P}(1)}, \dots, x_{\mathbb{P}_n}\right\rangle \;\middle|\; \vec{\boldsymbol{x}} = (x_1, \dots, x_n)^T \right\},$$

where $S_n$ is the symmetric group over $n$ elements, and $\mathcal{N}$ is given by:

$$\mathcal{N} = \frac{1}{\sqrt{n!} \cdot \sqrt{\prod_k n_k!}},$$

where $n_k$ is the number of times $k$ is repeated in $\vec{\boldsymbol{x}}$.

*Intuition* The idea is that we take a state that is not properly symmetrised $|x_1, \dots, x_n\rangle$, and we symmetrise it using the symmetric group.

*Proof* We will prove the formula for $\mathcal{N}$ in the sixth exercise series.

*Example 1* Note that, for 2 Bosons, we had the three following possibilities:

$$|00\rangle, \quad |11\rangle, \quad \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

They can be indeed equivalently written as:

$$\psi_{\vec{\boldsymbol{x}}} \propto \sum_{\mathbb{P} \in S_2} \mathbb{P}|x_1, x_2\rangle = \sum_{\mathbb{P} \in S_2} \left|x_{\mathbb{P}(1)}, x_{\mathbb{P}(2)}\right\rangle,$$

where $\vec{\boldsymbol{x}} = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^T$ is some vector, and $S_2 = \{I, P_{12}\}$ is the permutation group with two elements.
For instance:

$$\sum_{\mathbb{P} \in S_2} |00\rangle = I|00\rangle + P_{12}|00\rangle = 2|00\rangle \propto |00\rangle,$$

$$\sum_{\mathbb{P} \in S_2} |01\rangle = |01\rangle + |10\rangle \propto \frac{|01\rangle + |10\rangle}{\sqrt{2}}.$$

*Example 2* Let us consider an unsymmetric state $|001\rangle$, of which we want to find a symmetric state. We notice that, in this case:

$$S_3 = \{I, P_{12}, P_{13}, P_{23}, P_{123}, P_{321}\},$$

where $P_{ijk} = P_{ij}P_{jk}$. Therefore:

$$\begin{aligned} &\sum_{\mathbb{P} \in S_3} \mathbb{P}|001\rangle \\ &= I|001\rangle + P_{12}|001\rangle + P_{23}|001\rangle + P_{13}|001\rangle + P_{123}|001\rangle + P_{321}|001\rangle \\ &= |001\rangle + |001\rangle + |010\rangle + |010\rangle + |100\rangle + |010\rangle \\ &\propto \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}. \end{aligned}$$

**State of $n$ Fermions** A basis for the state of $n$ Fermions is given by:

$$\left\{ |\psi_{\vec{\boldsymbol{x}}}\rangle = \frac{1}{\sqrt{n!}} \sum_{\mathbb{P} \in S_n} \operatorname{sign}(\mathbb{P})\mathbb{P}|x_1, \dots, x_n\rangle, \;\middle|\; \vec{\boldsymbol{x}} = (x_1, \dots, x_n)^T \right\},$$

where each $\vec{\boldsymbol{x}}$ does not have any repeated element (by the Pauli exclusion principle), and where $\operatorname{sign}(\mathbb{P})$ is the sign of the permutation, i.e. 1 if the number of swaps is even, -1 otherwise.

*Remark* The normalisation factor $\mathcal{N} = \frac{1}{\sqrt{n}}$ is simpler since we cannot have repeated terms in $\vec{\boldsymbol{x}}$.

## 6.3 Second quantisation

**Idea**

The second quantisation is a switch from listing the properties of each particle to counting how many particles have each property. Indeed, if we have two identical properties, it does not really make sense to list their properties separately; instead it is more meaningful to count how many there are that have those properties.
For instance:

$$|\uparrow\uparrow\rangle \leftrightarrow |2,0\rangle, \quad |\downarrow\downarrow\rangle \leftrightarrow |0,2\rangle, \quad \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \leftrightarrow |1,1\rangle.$$

Note that, in this case, we would have to remember that our particles are Bosons. For Fermions, the first quantisation equivalent to the second quantisation $|1,1\rangle$ would be $\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$.

**Creation and annihilation operators**

We introduce creation and annihilation operators to increase and decrease the particle numbers. It appears working with those in the second quantisation is simpler.

*Bosons*

For the Bosonic operators, it behaves just like the Quantum Harmonic oscillator:

$$c_i^\dagger |n_1, \ldots, n_i, \ldots\rangle = \sqrt{n_i + 1}|n_1, \ldots, n_i + 1, \ldots\rangle,$$

$$c_i |n_1, \ldots, n_i, \ldots\rangle = \sqrt{n_i}|n_1, \ldots, n_i - 1, \ldots\rangle.$$

It is a good exercise to check that:

$$[c_i, c_j] = \left[c_i^\dagger, c_j^\dagger\right] = 0, \quad \left[c_i, c_j^\dagger\right] = \delta_{ij}.$$

*Fermion*

On the other hand, the Fermionic case is much more messy and subtle. The creation operator is:

$$c_i^\dagger |n_1, \ldots, n_i, \ldots\rangle = (-1)^{n_1 + \ldots + n_{i-1}}(1 - n_i)|n_1, \ldots, n_i + 1, \ldots\rangle.$$

The factor $(1 - n_i)$ is here for the Pauli exclusion principle: if $n_i = 1$, then we cannot create a new particle with these properties. The $-1$ factor depending only on $n_1, \ldots, n_{i-1}$ (but not the $n_j$ with $j > i$) is here for the property $P_{12}|\psi\rangle = -|\psi\rangle$.
The annihilation operator is similarly defined by:

$$c_i |n_1, \ldots, n_i, \ldots\rangle = (-1)^{n_1 + \ldots + n_{i-1}} n_i |n_1, \ldots, n_i - 1, \ldots\rangle$$

It is again a good exercise to then check that:

$$\{c_i, c_j\} = \left\{c_i^\dagger, c_j^\dagger\right\} = 0, \quad \{c_i, c_j\}^\dagger = c_i c_j^\dagger + c_j c_i^\dagger = \delta_{ij}.$$

**Hong-Ou-Mandel effet**

We consider an experiment where we have a polarised photon coming from the left or right, hits a beamsplitter, and gets split into two. We note $L_H$ to be the annihilation operator for a photon on the left with a horizontal polarisation, and similarly for $L_V$ for vertical polarisation, and $R_H$ and $R_V$ for annihilation operators for photons on the right. We admit that the beamsplitter acts in the following way:

$$L_k^\dagger \to \frac{L_k^\dagger + R_k^\dagger}{\sqrt{2}}, \quad R_k^\dagger \to \frac{L_k^\dagger - R_k^\dagger}{\sqrt{2}},$$

where $k \in \{H, V\}$.
We consider two scenarios.

1. Let's say that we have a horizontally-polarised particle coming from the left and a vertically-polarised particle from the right. Mathematically, we start with:

$$|1\rangle_{L_H}|0\rangle_{L_V}|0\rangle_{R_H}|1\rangle_{R_V} = L_H^\dagger R_V^\dagger|0000\rangle.$$

Applying the effect of the beamsplitter:

$$\begin{aligned} L_H^\dagger R_V^\dagger|0000\rangle &\to \frac{1}{2}\Big(L_H^\dagger + R_H^\dagger\Big)\Big(L_V^\dagger - R_V^\dagger\Big)|0000\rangle \\ &= \frac{|1100\rangle + |0110\rangle - |1001\rangle - |0011\rangle}{2}. \end{aligned}$$

In other words, we end up with four different options: both are on the left, both are on the right, both go through or both bounce. This seems reasonable, since it is the four different possibilities for the two particles.


2. Let us now say that both particles are horizontally polarised, giving:

$$L_H^\dagger R_H^\dagger|0000\rangle \to \frac{1}{2}\Big(L_H^\dagger + R_H^\dagger\Big)\Big(L_H^\dagger - R_H^\dagger\Big)|0000\rangle = \frac{|2000\rangle - |0020\rangle}{\sqrt{2}},$$

where the extra $\sqrt{2}$ constant came from the creation operator, $L_H^\dagger|1000\rangle = \sqrt{2}|2000\rangle$.
In other words, either both photons end up on the left, or they both end up on the right. On the picture above, this is only the first two possibilities.

The idea to remember from this effect is that when Bosons are indistinguishable (like in the second case, where they both have the same polarisation), then they tend to end up in the same state. This is the opposite of identical Fermions, which are never in the same state.

*Remark* This effect is known as the Hong-ou-Mandel effect, or Bosonic Bunching.

# Chapter 7

# Perturbation theory

## 7.1 Non-degenerate and time-independent

**Goal**

Let's say that we have a Hamiltonian $H$, and that we want to find its eigenspectrum (i.e. eigenstates and eigenvalues). In many cases, we can split $\hat{H} = \hat{H}_0 + \lambda\hat{V}$ where $H_0$ is a part we can easily get its eigenspectrum, $\lambda \geq 0$ is some perturbation strength and $V$ is a perturbation.

The idea is therefore to compute results on $H$ from results on $H_0$.

**Theorem: Non-degenerate perturbation theory**

Let $H = H_0 + \lambda V$ be some Hamiltonian, with $\lambda \geq 0$. We moreover note $H_0|\varphi_n\rangle = \varepsilon_n|\varphi_n\rangle$ and $H|\psi_n\rangle = E_n|\psi_n\rangle$ to be their eigenvalues and eigenvectors. We suppose that $\varepsilon_n$ and $|\varphi_n\rangle$ are known, and our goal is to compute $E_n$ and $|\psi_n\rangle$.

We furthermore suppose that $|\varphi_n\rangle$ is approximately $\psi_n$, but with some correction terms $\left|\psi_n^{(i)}\right\rangle$ and $E^{(n)}$:

$$|\psi_n\rangle = |\varphi_n\rangle + \lambda\left|\psi_n^{(1)}\right\rangle + \lambda^2|\psi_n^2\rangle + \dots, \quad E_n = \varepsilon_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots$$

Finally, we take the global (non-physical) phase of $\left|\psi_n^{(1)}\right\rangle$ and $|\varphi_n\rangle$ so that $\left\langle\varphi_n\middle|\psi_n^{(1)}\right\rangle \in \mathbb{R}$.

If the eigenspectrum of $H_0$ is non-degenerate, i.e. if $\varepsilon_n \neq \varepsilon_m$ for all $n \neq m$, then:

1. $E_n^{(1)} = \langle\varphi_n|V|\varphi_n\rangle$,
2. $\left|\psi_n^{(1)}\right\rangle = \sum_{m\neq n} \frac{\langle\varphi_m|V|\varphi_n\rangle}{\varepsilon_n - \varepsilon_m}|\varphi_m\rangle$,
3. $E_n^{(2)} = \langle\varphi_n|V\left|\psi_n^{(1)}\right\rangle = \sum_{m\neq n} \frac{|\langle\varphi_m|V|\varphi_n\rangle|^2}{\varepsilon_n - \varepsilon_m}$.

*Remark 1*

It is possible to find these results in a more general way. However, this is rarely used past the second term and we will not do it in this class.

*Remark 2*

Note that we divide by $\varepsilon_n - \varepsilon_m$, hence the hypothesis that the eigensepctrum is non-degenerate. If it is degenerate, we have to use another method, called degenerate perturbation theory, which we will see later in the class.

*Proof* Let us feed our Ansatz on $H$, $|\psi_n\rangle$ and $E_n$ to the equation $H|\psi_n\rangle = E_n|\psi_n\rangle$:

$$\begin{aligned}&(H_0+\lambda V)\Big(|\varphi_n\rangle+\lambda\Big|\psi_n^{(1)}\Big\rangle+\lambda^2\Big|\psi_n^{(2)}\Big\rangle+\dots\Big)\\&=\Big(\varepsilon_n+\lambda E_n^{(1)}+\lambda^2E_n^{(2)}+\dots\Big)\Big(|\varphi_n\rangle+\lambda\Big|\psi_n^{(1)}\Big\rangle+\lambda^2\Big|\psi_n^{(2)}\Big\rangle+\dots\Big).\end{aligned}$$

We want this to be true for any $\lambda$, so we can consider each order separately. We will only need the three first orders:

- $H_0|\varphi_n\rangle = \varepsilon_n|\varphi_n\rangle$,
- $\lambda H_0\Big|\psi_n^{(1)}\Big\rangle + \lambda V|\varphi_n\rangle = \lambda\varepsilon_n\Big|\psi_n^{(1)}\Big\rangle + \lambda E_n^{(1)}|\varphi_n\rangle$,
- $\lambda^2 H_0\Big|\psi_n^{(2)}\Big\rangle + \lambda^2 V\Big|\psi_n^{(1)}\Big\rangle = \lambda^2\varepsilon\Big|\psi_n^{(2)}\Big\rangle + \lambda^2 E_n^{(1)}\Big|\psi_n^{(1)}\Big\rangle + \lambda^2 E_n^{(2)}|\varphi_n\rangle$.

Let us now look for our three results.

1. We want to find $E_n^{(1)}$. To do so, we can multiply by $\langle\varphi_n|$ on both sides of the equation of order $\lambda^1$, getting

$$\langle\varphi_n|H_0\Big|\psi_n^{(1)}\Big\rangle + \langle\varphi_n|V|\varphi_n\rangle = \varepsilon_n\Big\langle\varphi_n\Big|\psi_n^{(1)}\Big\rangle + E_n^{(1)}\langle\varphi_n|\varphi_n\rangle.$$

   We notice that $\langle\varphi_n|\varphi_n\rangle = 1$. Moreover, since $H_0|\varphi_n\rangle = \varepsilon|\varphi_n\rangle$ and $H_0$ is Hermitian, we know $\langle\varphi_n|H_0 = \varepsilon_n\langle\varphi_n|$. Therefore:

$$\begin{aligned}&\varepsilon_n\Big\langle\varphi_n\Big|\psi_n^{(1)}\Big\rangle + \langle\varphi_n|V|\varphi_n\rangle = \varepsilon_n\Big\langle\varphi_n\Big|\psi_n^{(1)}\Big\rangle + E_n^{(1)}\\\iff& E_n^{(1)} = \langle\varphi_n|V|\varphi_n\rangle.\end{aligned}$$

2. We now want to express $\Big|\psi_n^{(1)}\Big\rangle$ in the basis $|\varphi_m\rangle$,

$$\Big|\psi_n^{(1)}\Big\rangle = \sum_m\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle|\varphi_m\rangle.$$

   We therefore have to compute $\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle$ for all $m$. We first suppose $m \neq n$. Multiplying by $\langle\varphi_m|$ on both sides of our first order equation:

$$\begin{aligned}&\langle\varphi_m|H_0\Big|\psi_n^{(1)}\Big\rangle + \langle\varphi_m|V|\varphi_n\rangle = \varepsilon_n\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle + E_n^{(1)}\overbrace{\langle\varphi_m|\varphi_n\rangle}^{=0}\\\iff&\varepsilon_m\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle + \langle\varphi_m|V|\varphi_n\rangle = \varepsilon_n\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle\\\iff&\langle\varphi_m|V|\varphi_n\rangle = (\varepsilon_n-\varepsilon_m)\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle\\\iff&\Big\langle\varphi_m\Big|\psi_n^{(1)}\Big\rangle = \frac{\langle\varphi_m|V|\psi_n\rangle}{\varepsilon_n-\varepsilon_m}.\end{aligned}$$

   Let us now consider the case $m = n$. We know that $\langle\psi_n|\psi_n\rangle = 1$, so:

$$\begin{aligned}1&=\langle\psi_n|\psi_n\rangle\\&=\Big(\langle\varphi_n|+\lambda\Big\langle\psi_n^{(1)}\Big|+\lambda^2\Big\langle\psi_n^{(2)}\Big|+\dots\Big)\\&\quad\cdot\Big(|\varphi_n\rangle+\lambda\Big|\psi_n^{(1)}\Big\rangle+\lambda^2\Big|\psi_n^{(2)}\Big\rangle+\dots\Big).\end{aligned}$$

   Up to the first order, this reads:

$$1 = \langle\varphi_n|\varphi_n\rangle + \lambda\Big(\Big\langle\varphi_n\Big|\psi_n^{(1)}\Big\rangle + \Big\langle\varphi_n^{(1)}\Big|\varphi_n\Big\rangle\Big) + O(\lambda^2).$$

The first order on the left handside must be equal to the first order on the right handside, i.e.:

$$\begin{aligned} 0 &= \left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle + \left\langle \psi_n^{(1)} \middle| \varphi_n \right\rangle \\ \iff & \Re\Big(\left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle\Big) = 0 \\ \iff & \left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle = 0, \end{aligned}$$

where we used the hypothesis that $\left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle \in \mathbb{R}$.
Putting everything together, this does yield that:

$$\left| \psi_n^{(1)} \right\rangle = \sum_{m \neq n} \frac{\langle \varphi_m | V | \varphi_n \rangle}{\varepsilon_n - \varepsilon_m} |\varphi_m\rangle.$$

3. We can now consider the second order. Recall the equation of order $\lambda^2$:

$$H_0 \left| \psi_n^{(2)} \right\rangle + V \left| \psi_n^{(1)} \right\rangle = \varepsilon_n \left| \psi_n^{(2)} \right\rangle + E_n^{(1)} \left| \psi_n^{(1)} \right\rangle + E_n^{(2)} |\varphi_n\rangle.$$

Multiplying both sides by $\langle \varphi_n |$, this means, using the facts that $\langle \varphi_n | H_0 = \varepsilon_n \langle \varphi_n |$ and $\langle \varphi_n | \varphi_n \rangle = 1$:

$$\varepsilon_n \left\langle \varphi_n \middle| \psi_n^{(2)} \right\rangle + \langle \varphi_n | V \left| \psi_n^{(1)} \right\rangle = \varepsilon_n \left\langle \varphi_n \middle| \psi_n^{(2)} \right\rangle + E_n^{(1)} \left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle + E_n^{(2)}.$$

Notice that the $\varepsilon_n \left\langle \varphi_n \middle| \psi_n^{(2)} \right\rangle$ cancel out on both sides of the equation, and recall that $\left\langle \varphi_n \middle| \psi_n^{(1)} \right\rangle = 0$, so we are left with:

$$\begin{aligned} E_n^{(2)} &= \langle \varphi_n | V \left| \psi_n^{(1)} \right\rangle \\ &= \sum_{m \neq n} \frac{\langle \varphi_m | V | \varphi_n \rangle}{\varepsilon_n - \varepsilon_m} \langle \varphi_n | V | \varphi_m \rangle \\ &= \sum_{m \neq n} \frac{|\langle \varphi_m | V | \varphi_n \rangle|^2}{\varepsilon_n - \varepsilon_m}. \end{aligned}$$

□

**Theorem** Let $n$ be arbitrary. We define $\Delta$ to be the minimum energy difference to $\varepsilon_n$:

$$\Delta = \min_m |\varepsilon_m - \varepsilon_n|.$$

We suppose that at least one of the following conditions hold:

1. $\left| \dfrac{\langle \varphi_n | V^2 | \varphi_n \rangle}{\langle \varphi_n | V | \varphi_n \rangle} - \langle \varphi_n | V | \varphi_n \rangle \right| \ll \Delta$,
2. $|\langle \varphi_m | V | \varphi_n \rangle| \ll |\varepsilon_n - \varepsilon_m|$.

Then:

$$\left| E_n^{(2)} \right| \ll \left| E_n^{(1)} \right|.$$

*Remark 1* For perturbation theory to be applicable up to the first order, we need exactly $\left| E_n^{(2)} \right| \ll \left| E_n^{(1)} \right|$. This theorem therefore gives us two sufficient conditions for perturbation theory to be applicable, without having to compute $E_n^{(1)}$ and $E_n^{(2)}$.

*Remark 2* The second hypothesis is looser than the first one.

*Proof 1*

We first prove that the first hypothesis does indeed yield $\left|E^{(2)}\right| \ll \left|E^{(1)}\right|$.
We notice that, by the triangle inequality:

$$\left|E_n^{(2)}\right| = \left|\sum_{m\neq n} \frac{|\langle\varphi_m|V|\varphi_n\rangle|^2}{\varepsilon_n - \varepsilon_m}\right| \leq \sum_{m\neq n}\left|\frac{|\langle\varphi_m|V|\varphi_n\rangle|^2}{\varepsilon_n - \varepsilon_m}\right|,$$

Using the $\Delta = \min_m|\varepsilon_n - \varepsilon_m|$ we defined in the hypotheses, this gives us

$$\begin{aligned}\left|E_n^{(2)}\right| &\leq \frac{1}{\Delta}\sum_{m\neq n}|\langle\varphi_m|V|\varphi_n\rangle|^2 \\ &= \frac{1}{\Delta}\left(\sum_m|\langle\varphi_m|V|\varphi_n\rangle|^2 - |\langle\varphi_n|V|\varphi_n\rangle|^2\right) \\ &= \frac{1}{\Delta}\left(\sum_m\langle\varphi_n|V|\varphi_m\rangle\langle\varphi_m|V|\varphi_n\rangle - |\langle\varphi_n|V|\varphi_n\rangle|^2\right).\end{aligned}$$

We know $\sum_m|\varphi_m\rangle\langle\varphi_m| = I$ by the completeness relation, so:

$$\left|E_n^{(2)}\right| \leq \frac{1}{\Delta}\Big(\langle\varphi_n|V^2|\varphi_n\rangle - |\langle\varphi_n|V|\varphi_n\rangle|^2\Big).$$

Since $E_n^{(1)} = \langle\varphi_n|V|\varphi_n\rangle$, we thus want:

$$\begin{aligned}&\frac{1}{\Delta}\Big(\langle\varphi_n|V^2|\varphi_n\rangle - |\langle\varphi_n|V|\varphi_n\rangle|^2\Big) \ll |\langle\varphi_n|V|\varphi_n\rangle| \\ \iff &\frac{\langle\varphi_n|V^2|\varphi_n\rangle}{|\langle\varphi_n|V|\varphi_n\rangle|} - |\langle\varphi_n|V|\varphi_n\rangle| \ll \Delta \\ \impliedby &\left|\frac{\langle\varphi_n|V^2|\varphi_n\rangle}{\langle\varphi_n|V|\varphi_n\rangle} - \langle\varphi_n|V|\varphi_n\rangle\right| \ll \Delta,\end{aligned}$$

since $\langle\varphi_n|V^2|\varphi_n\rangle \geq 0$. This is indeed respected by the first hypothesis.

*Proof 2*

Remember that:

$$E_n^{(2)} = \sum_{m\neq n}\frac{|\langle\varphi_m|V|\varphi_n\rangle|^2}{\varepsilon_n - \varepsilon_m}.$$

Therefore, the second condition simply asks for all terms in this sum to decay sufficiently fast.

**Example**

We consider a perturbed quantum harmonic oscillator:

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

We have $H_0 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2$ and we can let $\lambda = 1$ to get $V = -qEx$. As seen in Quantum physics 1, the quantum oscillator has eigenstate $|n\rangle$ with eigenvalues $\varepsilon_n = \hbar\omega\big(n + \frac{1}{2}\big)$:

$$H_0|n\rangle = \varepsilon_n|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle.$$

We moreover exploit ladder operators when working with the quantum oscillator:

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

We can rewrite $V$ using those ladder operators:

$$V = -qEx = -qE\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right).$$

First order perturbation theory tells us that since, in our context, $|\varphi_n\rangle = |n\rangle$:

$$\begin{aligned} E_n^{(1)} &= \langle\varphi_n|V|\varphi_n\rangle \\ &= -qE\sqrt{\frac{\hbar}{2m\omega}}\langle n|\left(a + a^\dagger\right)|n\rangle \\ &= -qE\sqrt{\frac{\hbar}{2m\omega}}\langle n|\left(\sqrt{n}|n-1\rangle + \sqrt{n+1}|n+1\rangle\right) \\ &= 0. \end{aligned}$$

This makes sense since $|n\rangle$ are symmetric (they have an even probability function $|\langle x|n\rangle|^2$), so the expected value of $V \propto x$ should indeed be 0.
We can now compute the second correction term:

$$\begin{aligned} E_n^{(2)} &= \sum_{k\neq n}\frac{|\langle k|V|n\rangle|^2}{\varepsilon_k - \varepsilon_n} \\ &= \sum_{k\neq n}\left(-qE\sqrt{\frac{\hbar}{2m\omega}}\right)^2\frac{\left|\langle k|\left(a + a^\dagger\right)|n\rangle\right|^2}{\hbar\omega(n-k)} \\ &= \frac{q^2E}{2m\omega^2}\sum_{k\neq n}\frac{\left|\sqrt{n}\langle k|n-1\rangle\right|^2 + \left|\sqrt{n+1}\langle k|n+1\rangle\right|^2}{n-k} \\ &= \frac{q^2E^2}{2m\omega^2}\left(\frac{|\sqrt{n}|^2}{n-(n-1)} + \frac{\left|\sqrt{n+1}\right|^2}{n-(n+1)}\right) \\ &= \frac{q^2E^2}{2m\omega^2}(n-(n+1)) \\ &= \frac{-q^2E^2}{2m\omega^2}. \end{aligned}$$

This means that, up to the second order:

$$E_n \approx h\omega\left(n + \frac{1}{2}\right) - \frac{q^2E^2}{2m\omega^2}.$$

*Remark 1* This may seem like being a toy problem that was just designed to be an example. However, it does appear in real life: this was a problem Prof. Holmes used during their PhD.

*Remark 2* In this very specific case, it is possible to find the full perturbation by completing the square:

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 - qEx \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2\left(x - \frac{qE}{m\omega^2}\right)^2 - \left(\frac{qE}{m\omega^2}\right)^2 \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2(x-x_0)^2 - \underbrace{\frac{1}{2}m\omega^2x_0^2}_{E_n^{(2)}}. \end{aligned}$$

We thus notice that, still in this very specific case, the second order gave us the full correction: the perturbation reduces the energy by

$\frac{1}{2}m\omega^2 x_0^2 = \frac{q^2E^2}{2m\omega^2} = E_n^{(2)}$. In other words:

$$E_n = h\omega\left(n + \frac{1}{2}\right) - \frac{q^2E^2}{2m\omega^2}.$$

It is not true in the general case that the second order approximation gives us the exact result.

Monday 4th November 2024 — **Lecture 7 : Penpineappleapplepen**

## 7.2 Degenerate perturbation theory

**Theorem: Degenerate perturbation theory**

Let $H = H_0 + \lambda V$ be some Hamiltonian, with $\lambda \geq 0$. We moreover note $H_0|\varphi_n\rangle = \varepsilon_n|\varphi_n\rangle$ and $H|\psi_n\rangle = E_n|\psi_n\rangle$ to be their eigenvalues and eigenvectors. We suppose that $\varepsilon_n$ and $|\varphi_n\rangle$ are known, and our goal is to compute $E_n$ and $|\psi_n\rangle$.
We suppose that for the $n^{\text{th}}$ energy eigenstate of $H_0$, there is an $N$-fold degeneracy, i.e.:

$$H_0|\varphi_{n_i}\rangle = \varepsilon_n|\varphi_{n_i}\rangle, \quad i \in \{1, \dots, N\},$$

where $\varepsilon_n$ is the same for all $i$.
We furthermore suppose that $|\varphi_n\rangle$ is approximately $\psi_n$, but with some correction terms $\left|\psi_n^{(i)}\right\rangle$ and $E^{(n)}$:

$$|\psi_n\rangle = \sum_j c_j\left|\varphi_{n_j}\right\rangle + \lambda\left|\psi_n^{(1)}\right\rangle + \lambda^2|\psi_n^2\rangle + \dots, \quad E_n = \varepsilon_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots$$

We define a sub-matrix of $V$, which we call $V' \in \mathbb{C}^{N\times N}$, so that:

$$V'_{ij} = \langle\varphi_{n_i}|V'\left|\varphi_{n_j}\right\rangle = \langle\varphi_{n_i}|V\left|\varphi_{n_j}\right\rangle$$

Then, $E_n^{(1)}$ and $\left|\psi_n^{(0)}\right\rangle = \sum_j c_j\left|\varphi_{n_j}\right\rangle$ respect the following eigenvalue-eigenvector problem:

$$V'\vec{\boldsymbol{c}} = E_n^{(1)}\vec{\boldsymbol{c}}.$$

*Proof*

As mentioned in the hypotheses, our Ansatz is:

$$|\psi_n\rangle = \sum_j c_j\left|\varphi_{n_j}\right\rangle + \lambda\left|\psi_n^{(1)}\right\rangle + \lambda^2\left|\psi_n^{(2)}\right\rangle + \dots,$$

$$E_n = \varepsilon_n + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots.$$

It is very similar to the non-degenerate case, except that it also takes into account that all $\left|\varphi_{n_j}\right\rangle$ have the same energy $\varepsilon_n$. Just like for the non-degenerate pertubation theory, we take an eigenstate $(H_0 + \lambda V)|\psi_n\rangle = E_n|\psi_n\rangle$, and equate the orders. The first order of $\lambda$ yields:

$$\sum_j c_j V\left|\varphi_{n_j}\right\rangle + H_0\left|\psi_n^{(1)}\right\rangle = \varepsilon_n\left|\psi_n^{(1)}\right\rangle + E_n^{(1)}\sum_j c_j\left|\varphi_{n_j}\right\rangle.$$

Multiplying both sides by $\langle\varphi_{n_i}|$, we get, using the facts that $\langle\varphi_{n_i}|H_0 = \varepsilon_n H_0$ and $\left\langle\varphi_{n_i}\middle|\varphi_{n_j}\right\rangle = \delta_{ij}$:

$$\sum_j c_j\langle\varphi_{n_i}|V\left|\varphi_{n_j}\right\rangle + \varepsilon_n\left\langle\varphi_{n_i}\middle|\psi_n^{(1)}\right\rangle = \varepsilon_n\left\langle\varphi_{n_i}\middle|\psi_n^{(1)}\right\rangle + E_n^{(1)}\sum_j c_j\delta_{ij}.$$

Cancelling out the $\varepsilon_n \left\langle \varphi_{n_i} \middle| \psi_n^{(1)} \right\rangle$ on both sides, and simplifying the sum, this becomes:

$$\sum_j V_{ij} c_j = E_n^{(1)} c_i \implies V' \vec{\boldsymbol{c}} = E_n^{(1)} \vec{\boldsymbol{c}},$$

where we recognised our sub-matrix $V'$.
We identify an eigenvalue equation, which gives us the $0^{\text{th}}$ order correction to the eigenstate $\vec{\boldsymbol{c}}$, and the $1^{\text{st}}$ order correction to the energy $E_n^{(1)}$.

□

**Example: Stark effect**

We consider the stark effect, i.e. the splitting of the degeneracy of one-electron atoms (such as Hydrogen atoms) in a uniform electric field pointing in the $z$ direction. The Hamiltonian is given by:

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} - \frac{e^2}{4\pi\varepsilon_0 r} \underbrace{-e\mathcal{E}z}_{=V} = H_0 + V.$$

Let us consider the energy level $n = 2$. Notice that $H_0$ is a central potential Hamiltonian. As seen in Quantum physics 1, $H_0$ has eigenstate $|n, \ell, m\rangle$ where $n$ is the energy, $\ell$ is the total angular momentum, and $m$ is the angular moment in the $z$ direction. For $n = 2$, we have all the following states that have the same energy:

$$|200\rangle, \quad |2, 1, -1\rangle, \quad |210\rangle, \quad |211\rangle.$$

The $V$ we want to diagonalise is:

$$V = \begin{pmatrix} \langle 200|V|200\rangle & \langle 200|V|210\rangle & \langle 200|V|211\rangle & \langle 200|V|2,1,-1\rangle \\ \langle 210|V|200\rangle & \langle 210|V|210\rangle & \langle 210|V|211\rangle & \langle 210|V|2,1,-1\rangle \\ \langle 211|V|200\rangle & \langle 211|V|210\rangle & \langle 211|V|211\rangle & \langle 211|V|2,1,-1\rangle \\ \langle 2,1,-1|V|200\rangle & \langle 2,1,-1|V|210\rangle & \langle 2,1,-1|V|211\rangle & \langle 2,1,-1|V|2,1,-1\rangle \end{pmatrix}.$$

Let us compute a coefficient in this matrix:

$$V_{imi'm'} = \langle 2im|V|2i'm'\rangle = -e\mathcal{E} \iiint u_{im}^* \underbrace{r\cos(\theta)}_{=z} u_{i'm'} \underbrace{r^2 \sin(\theta)}_{=|\det J|} d\theta d\varphi dr.$$

Recalling the analysis from Quantum physics 1, we have, taking $a_0$ as the Bohr radius:

$$u_{00} \propto \left(1 - \frac{r}{2a_0}\right) \exp\left(-\frac{r}{2a_0}\right), \quad u_{10} \propto r\cos(\theta) \exp\left(-\frac{r}{2a_0}\right),$$

$$u_{11} \propto r\sin(\theta)\exp(i\varphi)\exp\left(-\frac{r}{2a_0}\right), \quad u_{1,-1} \propto r\sin(\theta)\exp(-i\varphi)\exp\left(-\frac{r}{2a_0}\right).$$

We notice that all the diagonal terms are 0, since we always have an odd function in $\theta$ (since all $u_{ij}$ are either odd or even in $\theta$, they are squared and multiplied by an odd function in $\theta$), and any term containing $u_{11}$ or $u_{1,-1}$ is 0 since $\exp(ik\varphi)$ integrates to 0. Therefore, $V$ is of the form:

$$V = \begin{pmatrix} 0 & \alpha & 0 & 0 \\ \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{where } \alpha = -3e\mathcal{E}a_0.$$

We directly notice that it is block diagonal, with the top-left corner being $\alpha \cdot \sigma_X$. Therefore, its eigenvalues are $\alpha$, $-\alpha$, 0 and 0, with eigenstates being respectively:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

This shows the Stark effect: states that all have the same energy level under $H_0$ split into states that have three different energy level.

## 7.3 Dyson series

**Definition: Time-ordering operator**

Let $H_1(t), \dots, H_n(t)$ be operators depending on time.
We define the **time ordering operator** $T$ so that:

$$T[H(t_1)H(t_2)\cdots H(t_n)] = H(t_{i_1})H(t_{i_2})\cdots H(t_{i_n})$$

where $t_{i_1} \geq t_{i_2} \geq \ldots \geq t_{i_n}$.

*Example* For instance, if $t_2 > t_1$:

$$T(H(t_2)H(t_1)) = T(H(t_1)H(t_2)) = H(t_2)H(t_1).$$

*Remark* This can be defined more properly, in which case it is possible to show it is linear. In other words:

$$T\left(\int H(t)dt\right) = \int T(H(t))dt.$$

**Theorem: Dyson series**

Let $H(t)$ be a Hamiltonian depending on time.
The unitary $U(t, t_0)$ that solves $|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$, where $|\psi(t)\rangle$ is subject to Schrödinger's equation (where we take $\hbar = 1$), is given by:

$$\begin{aligned} U(t, t_0) &= T\left(\exp\left(-i\int_{t_0}^{t} dt_1 H(t_1)\right)\right) \\ &= I + \sum_{n=1}^{\infty}(-1)^n \frac{1}{n!}\int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n T(H(t_1)\cdots H(t_n)). \end{aligned}$$

*Remark* When $H(t) = H$ is time-independent, this simply reads:

$$U(t, t_0) = T\left(\exp\left(-iH\int_{t_0}^{t} dt_1\right)\right) = \exp(-iH(t - t_0)),$$

as expected.

*Proof* Schrödinger's equation for time-dependent Hamitlonians gives, leaving $\hbar = 1$:

$$i\frac{\partial}{\partial t}|\varphi(t)\rangle = H(t)|\varphi(t)\rangle.$$

Let us feed the expression $|\varphi(t)\rangle = U(t, t_0)|\varphi(t_0)\rangle$ into Schrödinger's equation:

$$i\frac{\partial}{\partial t}U(t, t_0)|\varphi(t_0)\rangle = H(t)U(t, t_0)|\varphi(t_0)\rangle.$$

Since it holds for any states, we get an equation for $U(t, t_0)$:

$$i\frac{\partial}{\partial t}U(t, t_0) = H(t)U(t, t_0),$$

with the initial condition $U(t_0, t_0) = I$.

Integrating both sides from $t_0$ to $t$, this yields:

$$\begin{aligned} & i\int_{t_0}^{t} dt_1 \frac{\partial}{\partial t_1} U(t_1,t_0) = \int_{t_0}^{t} dt_1 H(t_1)U(t_1,t_0) \\ \iff & i(U(t,t_0)-I) = \int_{t_0}^{t} dt_1 H(t)U(t_1,t_0) \\ \iff & U(t,t_0) = I - i\int_{t_0}^{t} dt_1 H(t_1)U(t_1,t_0). \end{aligned}$$

We can now again plug this formulation of $U(t,t_0)$ in the integral:

$$\begin{aligned} U(t,t_0) &= I - i\int_{t_0}^{t} dt_1 H(t_1)\left(I - i\int_{t_0}^{t_1} dt_2 H(t_2)U(t_2,t_0)\right) \\ &= I - i\int_{t_0}^{t} dt_1 H(t_1) - i^2\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H(t_1)H(t_2)U(t_2,t_0). \end{aligned}$$

Continuing to do this iteratively, we get, for any $m$:

$$\begin{aligned} & U(t,t_0) \\ &= I + \sum_{n=1}^{m-1} (-i)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n) \\ &\quad + (-i)^m \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{m-1}} dt_m H(t_1)H(t_2)\cdots H(t_m)U(t_m,t_0). \end{aligned}$$

Notice that $U(t_m,t_0)$ only multiplies the last term in the sum. Since $t_0 \leq t_m \leq \ldots \leq t_1$ by definition of domain of integration, we notice that $\lim_{m\to\infty} t_m = t_0$. This means that, $\lim_{m\to\infty} U(t_m,t_0) = I$. So, taking $m \to \infty$, this gives us that:

$$\begin{aligned} & U(t,t_0) \\ &= I + \sum_{n=1}^{\infty} (-i)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n). \end{aligned}$$

Note that it is a pain how the limits of integration depend on the parameter we are integrating over. It would be much nicer if all limits where between $t_0$ and $t$. We consider one of the integrals in the sum by simplicity. When $n = 2$, we have:

$$J_2 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} H(t_1)H(t_2).$$

By definition of integrals, we notice $t_2 \in [t_0,t_1]$, and hence $t_2 \leq t_1$. Therefore:

$$J_2 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 T(H(t_1)H(t_2)).$$

We can also change the order of integration in order to integrate on $t_2$ first, while being careful on the integral bounds:

$$\begin{aligned} J_2 &= \int_{t_0}^{t} dt_2 \int_{t_2}^{t} dt_1 H(t_1)H(t_2) \\ &\overset{t_1 \leftrightarrow t_2}{=} \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 H(t_2)H(t_1) \\ &= \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 T(H(t_1)H(t_2)), \end{aligned}$$

where we switched the dummy variables $t_1$ and $t_2$.

To summarise, we found two equivalent expressions for $J_2$:

$$J_2 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 T[H(t_1)H(t_2)] = \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 T[H(t_1)H(t_2)].$$

We can thus average those two expressions to get:

$$J_2 = \frac{J_2 + J_2}{2} = \frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T(H(t_1)H(t_2)).$$

It is possible to do this very generally, yielding:

$$U(t, t_0) = I + \sum_{n=1}^{\infty} (-i)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T[H(t_1) \cdots H(t_n)].$$

We can finally use the fact that $T$ is linear, and the definition of the exponential:

$$\begin{aligned} U(t, t_0) &= T\left( I + \sum_{n=1}^{\infty} (-i)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n H(t_1) \cdots H(t_n) \right) \\ &= T\left( \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \left( \int_{t_0}^{t} dt_1 H(t_1) \right)^n \right) \\ &= T\left( \exp\left( -i \int_{t_0}^{t} dt_1 H(t_1) \right) \right). \end{aligned}$$

□

**Remark**

Alternatively to the Dyson series, when we want to find the unitary $U(t, t_0)$ corresponding to some time-dependent Hamiltonian $H(t)$, we can discretise it:

$$U(t, t_0) = \prod_j \exp(-iH(t_i)\delta t).$$

The error depends on the difference between the continuous Hamiltonian and the discretised one:

$$\left\| \int_{t_0}^{t} ds H(s) - \delta t \sum_{r=1}^{m_t} H(t_0 + r\delta t) \right\|^2,$$

for some operator norm.

So, this is usable in some scenario, if the Hamiltonian does not oscillate too quickly. We will not use it in this class, but it is good to know that this less intimidating and more intuitive approach also exists and is also sometimes used in practice.

## 7.4 Interaction picture and time-dependent perturbation theory

**Observation**

In Schrödinger's picture, the state evolves but the observable stays constant:

$$i\frac{\partial}{\partial t}|\varphi_S(t)\rangle = H(t)|\varphi_S(t)\rangle, \quad O_S(t) = O_S.$$

On the other hand, in Heisenberg's picture, the observable evolves but the state is constant:

$$O_H(t) = U_S(t, t_0)^\dagger O_S U_S(t, t_0), \quad |\varphi_H(t)\rangle = |\varphi_S(t_0)\rangle.$$

Those two pictures are equivalent, because we have equivalent expectation values:

$$\langle \varphi_H(t)|O_H(t)|\varphi_H(t)\rangle = \langle \varphi_H|U(t, t_0)^\dagger O_S \underbrace{U_S(t, t_0)|\varphi_H\rangle}_{=|\varphi_S(t)\rangle} = \langle \varphi_S(t)|O_S|\varphi_S(t)\rangle.$$

We are interested in another approach, called the interaction picture, that merges Schrödinger's and Heisenberg's picture.

**Theorem: Interaction picture**

We suppose $H(t) = H_0 + V(t)$, where $H_0$ is treated in Heisenberg's style and $V(t)$ is treated in Schrödinger's style. In other words, we define operators and states to evolve as:

$$O_I(t) = e^{iH_0(t-t_0)} O_S(t) e^{-iH_0(t-t_0)},$$
$$|\varphi_I(t)\rangle = e^{iH_0(t-t_0)}|\varphi_S(t)\rangle = e^{iH_0(t-t_0)} U_S(t,t_0)|\varphi_S(t_0)\rangle.$$

This is equivalent to the Schrödinger's and Heisenberg's picture . Moreover, let us define $U_I(t,t_0)$ to be the evolution unitary, i.e. the operator so that $U_I(t,t_0)|\varphi_I(t_0)\rangle = |\varphi_I(t)\rangle$. Then:

$$\begin{aligned} U_I(t,t_0) &= e^{iH_0(t-t_0)} U_S(t,t_0) \\ &= T\left(\exp\left(-i\int_{t_0}^{t} dt_1 V_I(t_1)\right)\right) \\ &= I + \sum_{i=1}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n T(V_I(t_1)\cdots V_I(t_n)). \end{aligned}$$

where $V_I(t)$ is $V(t)$ in the interaction picture, i.e. $V_I(t) = e^{iH_0(t-t_0)} V(t) e^{-iH_0(t-t_0)}$.

*Intuition* Intuitively, $O_I(t)$ evolves in the Heisenberg's picture under $H_0$, and $|\varphi_I(t)\rangle$ evolves in a way that is similar to the Schrödinger's picture under $V(t)$. The idea in the definition of $|\varphi_I(t)\rangle$ is that it we let it evolve in the Schrödinger picture under $H(t)$, and then cancel out the evolution under $H_0$.

*Remark* This allows to do time-independent perturbation theory (where the perturbation $V(t)$ depends on time) when we are able to compute the first few terms of this sum. For instance, up to the first order:

$$U_I(t,t_0) \approx I - i\int_{t_0}^{t} dt_1 V_I(t_1).$$

However, this also has other applications.

*Proof equivalence* We show that this is equivalent to Schrödinger's picture (which is equivalent to Heisenberg's picture) by showing that expected values are equal:

$$\begin{aligned} &\langle\varphi_I(t)|O_I(t)|\varphi_I(t)\rangle \\ &= \langle\varphi_S(t)|e^{-iH_0(t-t_0)} e^{iH_0(t-t_0)} O_S(t) e^{-iH_0(t-t_0)} e^{iH_0(t-t_0)} U_S(t,t_0)|\varphi_S(t_0)\rangle \\ &= \langle\varphi_S(t)|O_S(t)|\varphi_S(t)\rangle, \end{aligned}$$

since $\exp(A)\exp(B) = \exp(A+B)$ when $[A,B] = 0$.

*Proof equality* Recall that we define:

$$|\varphi_I(t)\rangle = e^{iH_0(t-t_0)} U_S(t,t_0)|\varphi_S(t_0)\rangle.$$

In particular, $\varphi_I(t_0) = |\varphi_S(t_0)\rangle$. Moreover, by definition, $|\varphi_I(t)\rangle = U_I(t,t_0)|\varphi_I(t)\rangle$. This tells us that:

$$U_I(t,t_0) = e^{iH_0(t-t_0)} U_S(t,t_0).$$

Therefore, by differentiating both sides:

$$\begin{aligned}\frac{\partial U_I}{\partial t} &= e^{iH_0(t-t_0)} iH_0 U_S(t,t_0) - ie^{iH_0(t-t_0)} H(t) U_S(t,t_0) \\ &= -ie^{iH_0(t-t_0)} \underbrace{(H(t)-H_0)}_{=V} U_S(t,t_0) \\ &= -ie^{iH_0(t-t_0)} V e^{-iH_0(t-t_0)} e^{iH_0(t-t_0)} U_S(t,t_0) \\ &= -iV_I(t) U_I(t,t_0).\end{aligned}$$

Recall however that Schrödinger's equation is:

$$i\frac{\partial}{\partial t} U(t,t_0) = H'(t) U(t,t_0).$$

We therefore recognise that our equation is Schrödinger's equation with $H'(t) = V_I(t)$. This makes sense since, in the interaction picture, the observable evolves because of $H_0$ and the state evolves because of $V(t)$; we do expect the equation that governs the way our states evolve to be Schrödinger's equation of Hamiltonian $V(t)$ (in the interaction picture).
We can therefore use the Dyson series:

$$\begin{aligned}U_I(t,t_0) &= T\left(\exp\left(-i\int_{t_0}^{t} dt_1 V_I(t_1)\right)\right) \\ &= I + \sum_{i=1}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n T(V_I(t_1)\cdots V_I(t_n)).\end{aligned}$$

□

Monday 11th November 2024 — **Lecture 8 : Dodgy bit of Physicist math**

**Example 1**

Let $V$ be a constant operator. We consider a Hamiltonian of the form $H(t) = H_0 + V(t)$, where:

$$V(t) = \begin{cases} 0, & t < t_0, \\ V, & t > t_0. \end{cases}$$

In other words, we turn on some constant potential at $t_0$. We consider $H_0|n\rangle = E_n|n\rangle$ to be the eigendecomposition of $H_0$. We moreover suppose that we start at an eigenstate $i$ in the Schrödinger picture, $|\varphi_S(t_0)\rangle = |i\rangle$. We are interested in the transition amplitude and transition probabilities:

$$C_n(t) = \langle n|\varphi_S(t)\rangle = \langle n|U_S(t,t_0)|i\rangle, \quad P_{i\to n}(t) = |\langle n|\varphi_S(t)\rangle|^2 = |C_n(t)|^2.$$

Using the interaction picture, and the fact $\langle n|H_0 = E_n\langle n|$:

$$C_n(t) = \langle n|U_S(t,t_0)|i\rangle = \langle n|e^{-iH_0(t-t_0)} U_I(t,t_0)|i\rangle = e^{-iE_n(t-t_0)}\langle n|U_I(t,t_0)|i\rangle.$$

Now, using the first order expansion of $U_I(t,t_0)$, and supposing $n \neq i$:

$$\begin{aligned}\langle n|U_I(t,t_0)|i\rangle &= \langle n|i\rangle - i\int_{t_0}^{t} dt_1 \langle n|V_I(t_1,t_0)|i\rangle \\ &= -i\int_{t_0}^{t} dt_1 \langle n|e^{iH_0(t-t_0)} V(t_1,t_0) e^{-iH_0(t-t_0)}|i\rangle \\ &= -i\int_{t_0}^{t} dt_1 e^{-i(E_n-E_i)(t-t_0)} \langle n|V(t_1,t_0)|i\rangle.\end{aligned}$$

We also directly get that:

$$\begin{aligned} P_{i\to n}(t) &= \left| e^{-iE_n(t-t_0)}\langle n|U_I(t,t_0)|i\rangle \right|^2 \\ &= |\langle n|U_I(t,t_0)|i\rangle|^2 \\ &= \left| -i\int_{t_0}^{t} dt_1 e^{i(E_n-E_i)(t_1-t_0)}\langle n|V(t_1,t_0)|i\rangle \right|^2. \end{aligned}$$

Note that we have only used $V(t) = 0$ for all $t < t_0$ for now, and hence this is a general expression for the transition probability of any potential up we turn on $t_0$, up to the first order. We can now exploit the fact that $V(t) = V$ is in fact constant for $t \geq t_0$, to get our transition probability to simplify to:

$$\begin{aligned} P_{i\to n}(t) &= \left| \langle n|V|i\rangle \int_{t_0}^{t} dt_1 e^{i(E_n-E_i)(t_1-t_0)} \right|^2 \\ &= \left| \langle n|V|i\rangle \frac{\exp(i(E_n-E_i)(t-t_0))-1}{E_n-E_i} \right|^2 \\ &= |\langle n|V|i\rangle|^2 \frac{4}{(E_n-E_i)^2}\sin^2\left(\frac{(E_n-E_i)(t-t_0)}{2}\right). \end{aligned}$$

We now exploit a "dodgy bit of Physicist maths":

$$\lim_{t\to\infty} \frac{\sin^2(xt)}{x^2} = \pi t\delta(x).$$

The fact that there is a $t$ on the right hand side is very dodgy, and not at all formal. Intuitively, we have:

$$\frac{\sin^2(xt)}{(xt)^2} \to \pi\delta(x) \iff \frac{\sin^2(xt)}{x^2 t} \to \pi\delta(x) \iff \frac{\sin^2(xt)}{x^2} \to \pi t\delta(x).$$

Another way to see this is that $\frac{\sin^2(xt)}{x^2}$ gets spikier and spikier as $t$ gets larger, so it gets well approximated by a constant times a Dirac delta. However, since $\frac{\sin^2(xt)}{x^2}$ integrates to $\pi|t|$ on $\mathbb{R}$, the constant must be $\pi t$.
So now, we get, supposing $t_0 = 0$ by simplicity:

$$\lim_{t\to\infty} P_{i\to n}(t) = 4|\langle n|V|i\rangle|^2 \lim_{t\to\infty} \frac{1}{\omega^2}\sin^2\left(\frac{\omega t}{2}\right) = 2\pi t|\langle n|V|i\rangle|^2\delta(E_n-E_i).$$

Note that if $n \neq i$, the probability of transition is 0. This makes a lot of sense, by conservation of energy a state cannot just magically gain energy. Note that this probability can grow larger than 1, but this is because we used only the first order. Insuring that $|E_1| \gg |E_2|$ requires $t$ to be sufficiently small for the distribution to be almost normalised. Considering higher orders would allow the approximation to be valid for a larger time $t$.
Moreover, it is sometimes easier to think as the transition probability rate is:

$$\frac{\partial P_{i\to n}(t)}{\partial t} = 2\pi|\langle n|V|i\rangle|^2\delta(E_n-E_i).$$

This shows that, as time passes, the probability to go to another state of same energy increases linearly. This is some form of *Fermi's golden rule*.

**Example 2** Let $V$ be a constant operator. We consider a Hamiltonian of the form $H(t) = H_0 + V(t)$, where $V(t)$ is an oscillatory potential:

$$V(t) = \begin{cases} 0, & \text{if } t \leq t_0, \\ Ve^{i\omega t} + V^\dagger e^{-i\omega t}, & \text{if } t > t_0. \end{cases}$$

As found in the first example, since $V(t) = 0$ for $t \leq t_0$, when $n \neq m$:

$$P_{i\to n}(t) = \left| -i \int_{t_0}^{t} dt_1 e^{i(E_n - E_1)t_1} \langle n|V(t)|i\rangle \right|.$$

Therefore, in our context, the transition probability is given by:

$$\begin{aligned} P_{i\to n}(t) &= \left| -i \int_{t_0}^{t} dt_1 e^{i(E_n - E_1)t_1} \left( \langle n|V|i\rangle e^{i\omega t_1} + \langle n|V^\dagger|i\rangle e^{-i\omega t_1} \right) \right| \\ &= \left| \frac{1 - \exp(i(E_n - E_i + \omega)t)}{E_n - E_i + \omega} \langle n|V|i\rangle + \frac{1 - \exp(i(E_n - E_i - \omega)t)}{E_n - E_i - \omega} \langle n|V^\dagger|i\rangle \right|^2. \end{aligned}$$

We can now do some algebra to get a squared sine term, and take the limit as $t \to \infty$:

$$\lim_{t\to\infty} P_{i\to n}(t) = 2\pi t \left| \langle n|\hat{V}|i\rangle \right|^2 \delta(E_n - E_i + \omega) + 2\pi t \left| \langle n|V^\dagger|i\rangle \right|^2 \delta(E_n - E_i - \omega).$$

This means that, for a large $t$, it will tend to transition to states of energy $E_n = E_i \pm \omega$.

**Example 3** Let $V$ be a constant operator. We consider a Hamiltonian of the form $H(t) = H_0 + V(t)$, where:

$$V(t) = Ve^{\varepsilon t}.$$

Note that $V(t_1)$ commutes with $V(t_2)$ for any $t_1, t_2$. We can therefore drop the time-ordering operator in the Dyson series. Therefore, up to the second order, we have:

$$U_I(t, -\infty) = I - i \int_{-\infty}^{t} dt_1 V_I(t_1) - \frac{1}{2} \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 V_I(t_1) V_I(t_2).$$

So, this means that, when $n \neq i$:

$$P_{i\to n}(t) = \left| -i \int_{-\infty}^{t} dt_1 \langle n|V_I(t_1)|i\rangle - \frac{1}{2} \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \langle n|V_I(t_1)V_I(t_2)|i\rangle \right|^2.$$

We have two integrals to compute. The first one gives:

$$\begin{aligned} I_1(t) = & \int_{-\infty}^{t} \langle n|V_I(t)|i\rangle \\ &= \langle n|V|i\rangle \int_{-\infty}^{t} dt_1 e^{i[(E_n - E_1)t_1 - i\varepsilon t_1]} \\ &= \langle n|V|i\rangle \frac{\exp[i((E_n - E_i)t_1 - i\varepsilon t_1)]}{i(E_n - E_i - i\varepsilon)} \Big|_{-\infty}^{t}. \end{aligned}$$

So, if we were to ignore the second integral, we could do some algebra to get:

$$P_{i\to n}(t) = |\langle n|V|i\rangle|^2 \frac{e^{2\varepsilon t}}{(E_n - E_i)^2 + \varepsilon^2},$$

$$\frac{\partial P_{i\to n}(t)}{\partial t} = |\langle n|V|i\rangle|^2 \frac{2\varepsilon e^{2\varepsilon t}}{(E_n - E_i)^2 + \varepsilon^2}.$$

For the second integral, we can compute:

$$\begin{aligned} I_2(t) &= \sum_m \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \langle n|V_I(t_1)|m\rangle\langle m|V_I(t_2)|i\rangle \\ &= \sum_m \langle n|V|m\rangle\langle m|V|i\rangle \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 e^{i(E_n-E_m-i\varepsilon)t_1} e^{i(E_m-E_i-i\varepsilon)t_2} \\ &= \sum_m \langle n|V|m\rangle\langle m|V|i\rangle \int_{-\infty}^{t} dt_1 e^{i(E_n-E_m-i\varepsilon)t_1} \frac{e^{i(E_m-E_i-i\varepsilon)t_2}}{E_m-E_i-i\varepsilon}\Big|_{-\infty}^{t_1} \\ &= -\sum_m \langle n|V|m\rangle\langle m|V|i\rangle \frac{e^{i(E_n-E_i-2i\varepsilon)t}}{(E_m-E_i-i\varepsilon)(E_n-E_i-2i\varepsilon)}. \end{aligned}$$

Considering that $\varepsilon \approx 2\varepsilon$ in the limit of small $\varepsilon$, we get:

$$P_{i\to n}(t) = \left| \langle n|V|i\rangle + \sum_m \frac{\langle n|V|m\rangle\langle m|V|i\rangle}{E_m-E_i-i\varepsilon} \right|^2 \frac{e^{-2\varepsilon t}}{(E_n-E_i)^2+\varepsilon}.$$

This is very similar to what we obtained when we considered a single order, except that we have some more sum in the absolute value. A way to understand this new terms is as some form of virtual transition: we go from $i$ to some other $m$, then from $m$ to $n$. Intuitively, this means that, for higher order terms, we will have more virtual energy levels.

# Chapter 8

# Variational principle

**Proposition: Variational principle**

Let $H$ be a Hamiltonian, of diagonalisation $H|\varphi_n\rangle = E_n|\varphi_n\rangle$, and ground state $|\varphi_0\rangle$ and ground state energy $E_0$. Then:

- For all normalised $|\psi\rangle$, $\langle\psi|H|\psi\rangle \geq E_0$.
- For all normalised $|\psi\rangle$ orthogonal to $|\varphi_0\rangle$, $\langle\psi|H|\psi\rangle \geq E_1$.

We might sometimes want to work with non-normalised states $|\psi'\rangle$, in which case the variational principle becomes:

- For all $|\psi'\rangle$, $\frac{\langle\psi'|H|\psi'\rangle}{\langle\psi'|\psi'\rangle} \geq E_0$.
- For all $|\psi'\rangle$ orthogonal to $|\varphi_0\rangle$, $\frac{\langle\psi'|H|\psi'\rangle}{\langle\psi'|\psi'\rangle} \geq E_1$.

*Remark 1*

This principle is conceptually very simple—it literally just states that the average energy of a state is always more than the energy of the ground state—but it might require having to solve very nasty integrals. In exams, it might be easier to start with questions that are conceptually harder, since, if we understand them well, they will be easier to solve than hard integrals.

*Remark 2*

This allows to get an estimate of the ground state energy, by taking some Ansatz and finding its minimum energy. To find an estimate of the first excited state energy, we can then take an Ansatz that is orthogonal to the first one: if the first one is a good approximation to the ground state, the second one cannot be the ground state and can therefore not yield the ground state energy.

Note that, in this class, we will always be given the Ansätze. In real life, finding good Ansätze is a big field of research.

*Proof 1*

We consider the case where $|\psi\rangle$ is normalised, the other case directly follows by leaving $|\psi\rangle = \frac{1}{\||\psi'\rangle\|}|\psi'\rangle$.

We know $H$ can be expressed as:

$$H = \sum_{n=0}^{\infty} E_n|\varphi_n\rangle\langle\varphi_n|.$$

We note $p_n(\psi) = |\langle\psi|\varphi_n\rangle|^2$ for simpler notations. Given any normalised $|\psi\rangle$, we thus have:

$$\langle\psi|H|\psi\rangle = \sum_{n=0}^{\infty} E_n p_n(\psi) \geq E_0 \sum_{n=0}^{\infty} p_n(\psi) = E_0,$$

since the probabilities sum to 1 for normalised states.

Moreover, given any normalised $|\psi\rangle$ so that $\langle\varphi_0|\psi\rangle = 0$, the reasoning is completely similar and just uses the fact $p_0(\psi) = 0$:

$$\langle\psi|H|\psi\rangle = \sum_{n=1}^{\infty} E_n p_n(\psi) \geq E_1 \sum_{n=1}^{\infty} p_n(\psi) = E_1.$$

□

**Example**

Let us consider the simple Harmonic oscillator:

$$H = T + V = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2.$$

We want to guess the ground state energy. To do so, we consider the following class of non-normalised quantum states:

$$\psi_a(x) = \frac{1}{x^2 + a}.$$

We need to compute the three following integrals:

$$\langle\psi_a|T|\psi_a\rangle = \frac{-\hbar^2}{2m}\int_{-\infty}^{\infty} dx \frac{1}{x^2+a}\frac{d^2}{dx^2}\frac{1}{x^2+a},$$

$$\langle\psi_a|V|\psi_a\rangle = \frac{1}{2}m\omega^2\int_{-\infty}^{\infty} dx \frac{x^2}{(x^2+a)^2},$$

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} dx \frac{1}{(x^2+a)^2}.$$

We will not compute those integrals in class (and, in exams, we would have some small hints on how to compute them), but solving them and doing some bit of algebra, we find:

$$E(a) = \frac{\langle\psi_a|H|\psi_a\rangle}{\langle\psi_a|\psi_a\rangle} = \frac{\langle\psi_a|T|\psi_a\rangle}{\langle\psi_a|\psi_a\rangle} + \frac{\langle\psi_a|V|\psi_a\rangle}{\langle\psi_a|\psi_a\rangle} = \frac{\hbar^2}{2m}\frac{1}{a} + \frac{1}{2}m\omega^2 a.$$

By the variational principle, $E(a) \geq E_0$ for all $a$. So, in particular, $E_0 \leq \min_a E(a)$. We now do some more algebra to minimise $E(a)$:

$$\frac{\partial E(a)}{\partial a} = -\frac{\hbar^2}{ma^2} + \frac{1}{2}m\omega^2 = 0 \implies a = \frac{\hbar}{m\omega\sqrt{2}} \implies E(a) = \frac{\hbar\omega}{2} \approx 0.72\hbar\omega.$$

The real ground state is $0.5\hbar\omega$, so this is not too bad of an approximation.

Let's say that we now want to compute an estimate on the first excited state energy. To do so, we pick an Ansatz for the first excited wave function that is orthogonal to the ground state. Indeed, assuming our first Ansatz gets close the ground state, the lowest energy of this new Ansatz cannot be the ground state and will thus be an estimate for the first excited state. Since $\psi_a(x)$ is even, we can take $\varphi_a(x)$ to be odd. The following works well:

$$\varphi_a(x) = \frac{x}{(x^2+a)^2}.$$

By doing a lot of algebra, we find that:

$$\min_a E_1(a) = \sqrt{3}\hbar\omega \approx 1.732\hbar\omega.$$

The correct energy is $\frac{3}{2}\hbar\omega$, so this is again a reasonable result.

# Chapter 9

# Groups and representation theory

## 9.1 Motivational examples

**Translational symmetry**

Let's say that we have a system with a translational symmetry, i.e. shifting it does not change it, $\psi(x+a) \equiv \psi(x)$ for all $a$. Intuitively, this is very restrictive. The case $\psi(x) =$ constant directly comes to mind, but this is not the only possibility. We can consider the following class of states, defined for any $p' \in \mathbb{R}$:

$$\psi(x) = e^{ip'x}.$$

This is indeed period, since shifting it only gives a non-physical global phase (it is global since it does not depend on $x$):

$$\psi(x+a) = e^{ip'x}e^{ip'a} = e^{ip'a}\psi(x) \equiv \psi(x).$$

Taking the Fourier transform, this yields:

$$\varphi(p) = \delta(p-p').$$

*Remark*

We notice two things from this example:

1. First, symmetries are associated with constants: for instance, the translational symmetry is associated with conservation of momentum. This can be generalised thanks to Noether's theorem.
2. Second, symmetries heavily constrain the allowed states. This is for instance useful for finding Ansätze for the variational principle. Another example of this is that we found in Quantum physics I that $V(x) = V(-x)$ necessarily implies the eigenstates are even or odd.

**The mystery of degeneracy**

Given a generic Hamiltonian $H$, we would expect the eigenvalues of $H$ to be distinct. However, in real cases, $H$ often has lots of degenerate eigenvalues. In early days of quantum mechanics, this was rather mysterious. Symmetries provide an explanation. Let us consider a unitary $U$ that leaves $H$ invariant:

$$U^\dagger HU = H \iff HU = UH.$$

Now, leaving $|\psi\rangle$ to be an eigenvalue of $H$, $H|\psi\rangle = E|\psi\rangle$, this yields:

$$H(U|\psi\rangle) = UH|\psi\rangle = U(E|\psi\rangle) = E(U|\psi\rangle).$$

In other words, $U|\psi\rangle$ is also an eigenstate of $H$, with the same energy $E$. In other words, a symmetry $U$ implies degeneracy.

## 9.2 Introduction to groups

**Definition: Symmetry** A symmetry describes some property of a system that is left unchanged by some operation. In quantum mechanics, this operation will be a unitary $U$.

**Definition: Group** A set $G$ equipped with some operation $* : G \times G \mapsto G$ is a **group** $(G, *)$, if:

- (Closed) $\forall a, b \in G,\ a * b \in G$.
- (Associativity) $\forall a, b, c \in G,\ (a * b) * c = a * (b * c)$.
- (Identity) $\exists e \in G, \forall g \in G,\ e * g = g * e = g$.
- (Inverse) $\forall a \in G, \exists b \in G,\ a * b = b * a = e$, and we note $b = a^{-1}$.

*Remark* We will typically consider a group over multiplication, such as $(\mathbb{R}^*, \cdot)$.

**Proposition** The set of all unitaries $U$ such that the map $\rho \mapsto U\rho U^\dagger$ leaves some property unchanged forms a group together with multiplication.

*Proof*

- Multiplying two unitaries gives a unitary matrix.
- $U = I$ is always such that $\rho \mapsto U\rho U^\dagger$ preserves all properties.
- If $\rho \mapsto U\rho U^\dagger$ preserves some property, then so does its inverse $\rho \mapsto U^\dagger \rho U$. Indeed, applying $U$ leaves a property unchanged, so undoing what $U$ did must also keep this property unchanged. In other words, we do have that if $U \in G$, then $U^{-1} = U^\dagger \in G$ .
- If applying $U$ and applying $V$ leave properties unchanged, applying $U$ and then $V$ also leaves property unchanged. In other words, if $U, V \in G$, then $UV \in G$.

□

**Definition: Finite group** Let $(G, *)$ be a group.
If $G$ is finite, $(G, *)$ is said to be a **finite group**. Its number of elements $|G|$ is named to be the **order** of the group.

**Definition: Cayley table** Let $(G, *)$ be a finite group.
Its **Cayley table** is a $|G| \times |G|$ table representing its operation. Each cell $(a, b)$ represents $a * b$.

**Intuition: Representation** Let $(G, *)$ be a finite group.
Informally, a representation $\{R(g)\}$ is a set of $d \times d$ matrices that obey the rules of the group:

$$R(g_1) \cdot R(g_2) = R(g_1 * g_2).$$

*Remark* We will build this definition formally. However, the idea is not too complicated, and we can already find some examples.

**Example 1** The only order 1 group is:

| $*$ | $e$ |
|---|---|
| $e$ | $e$ |

The following are three representations of this group:

$$e \to 1, \quad e \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad e \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Example 2** To find the only order 2 group, we notice that $e * a = a$ and, since $a$ needs an inverse, $a * a = e$. So:

| $*$ | $e$ | $a$ |
|---|---|---|
| $e$ | $e$ | $a$ |
| $a$ | $a$ | $e$ |

The following are five representations of this group:

$$\begin{cases} e \to 1, \\ a \to 1, \end{cases} \quad \begin{cases} e \to 1, \\ a \to -1, \end{cases} \quad \begin{cases} e \to I, \\ a \to \sigma_X, \end{cases} \quad \begin{cases} e \to I, \\ a \to \sigma_Z, \end{cases} \quad \begin{cases} e \to I, \\ a \to SWAP. \end{cases}$$

Indeed, all those matrices behave similarly to the group. For instance, for the fourth representation, $I \cdot I = I$, $I \cdot \sigma_Z = \sigma_Z \cdot I = \sigma_Z$ and $\sigma_Z \cdot \sigma_Z = I$, showing indeed $I$ behaves like $e$ and $\sigma_Z$ behaves like $a$.

**Example 3** It is possible to show that the following is the only order 3 group:

| $*$ | $e$ | $a$ | $b$ |
|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ |
| $a$ | $a$ | $b$ | $e$ |
| $b$ | $b$ | $e$ | $a$ |

This is the $\mathbb{Z}_3$ cyclic group. A possible representation is the third roots of unity:

$$e \to 1, \quad a \to e^{i\frac{2\pi}{3}}, \quad b \to e^{i\frac{4\pi}{3}}.$$

Note that this also represents the symmetrical rotations of an equilateral triangle.

**Example 4** Some order 4 group is:

| $*$ | $e$ | $a$ | $b$ | $c$ |
|---|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ | $c$ |
| $a$ | $a$ | $b$ | $c$ | $e$ |
| $b$ | $b$ | $c$ | $e$ | $a$ |
| $c$ | $c$ | $e$ | $a$ | $b$ |

This is the $\mathbb{Z}_4$ cyclic group. A representation is $e^{i2\pi\frac{j}{4}}$. We could construct another representation by recognising it is the group of some symmetries of a square.

**Example 5** The other order 4 group is:

| $*$ | $e$ | $a$ | $b$ | $c$ |
|---|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ | $c$ |
| $a$ | $a$ | $e$ | $c$ | $b$ |
| $b$ | $b$ | $c$ | $e$ | $a$ |
| $c$ | $c$ | $b$ | $a$ | $e$ |

A representation can be done by representing this as the symmetries of a rectangle (rotations of 0, rotation of $\pi$, vertical symmetry, horizontal symmetry).

**Definition: Cyclic group** The cyclic group $\mathbb{Z}_n$ of order $n$ is defined so that:

$$a_i a_j = a_{i+j \bmod n}, \quad e = a_0.$$

A representation is $a_j = e^{i2\pi\frac{j}{n}}$.

**Definition: Symmetric permutation group**

The symmetric permutation group $S_n$ is all the permutation of $n$ objects. Its order is

$$|S_n| = n!.$$

In physics, this is important since it is the symmetry of indistinguishable particles.

*Example*

For instance:

$$S_3 = \{I, SWAP_{12}, SWAP_{13}, SWAP_{23}, CYCLE_{123}, CYCLE_{321}\}.$$

Using the usual notation of the symmetric permutation group, this can also be written as:

$$S_3 = \big\{e, (1 \quad 2), (1 \quad 3), (2 \quad 3), (1 \quad 2 \quad 3), (3 \quad 2 \quad 1)\big\}$$

**Definition: Lie group**

Informally, a **Lie group** is a continuous group that depends analytically on some continuous parameter $\lambda$.

**Definition: Orthogonal matrix**

Let $O \in \mathbb{R}^{n\times n}$. It is said to be orthogonal if:

$$O^T O = O O^T = I.$$

*Remark 1*

Orthogonal matrices are the real equivalent to unitary matrices.

*Remark 2*

Note that:

$$1 = \det(I) = \det\big(O^T O\big) = \det(O)^2.$$

There are thus two possibilities, $\det(O) = 1$ and $\det(O) = -1$. In the first case, $O$ represents a rotation; in the second case it represents a reflection.

**Definition: Orthogonal group**

The **orthogonal group**, $O(d)$, is the group of orthogonal matrices:

$$O(d) = \big\{O \in \mathbb{R}^{n\times n} \;\big|\; O \text{ is orthogonal}\big\}.$$

This represents the group of matrices that preserve the norm of real vectors, i.e. matrices that do rotations and reflections. This is a lie group.

**Definition: Special orthogonal group**

The **special orthogonal group**, $SO(d)$, is the real rotation matrices on $d$ dimensional systems:

$$SO(d) = \big\{O \in \mathbb{R}^{n\times n} \;\big|\; O \text{ is orthogonal and } \det(O) = 1\big\}.$$

This is a lie group.

*Example*

For instance, an arbitrary element of $SO(2)$ is:

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

**Definition: Unitary group**

The **unitary group**, $U(d)$, is the group of unitary matrices:

$$U(d) = \big\{U \in \mathbb{C}^{n\times n} \;\big|\; U \text{ is unitary}\big\}.$$

This represents the group of matrices that preserves the norm of complex vectors. This is a lie group.

**Definition: Special unitary group**

The **special unitary group**, $SU(d)$, is the group of unitary matrices:

$$SU(d) = \big\{U \in \mathbb{C}^{n\times n} \;\big|\; U \text{ is unitary and } \det(U) = 1\big\}.$$

This is a lie group.

*Remark*

Note that:

$$1 = \det(I) = \det\big(U^\dagger U\big) = \det(U)^* \det(U) = |\det(U)|^2.$$

Therefore, there exists some $\varphi$ so that:

$$\det(U) = e^{i\varphi}.$$

Asking that $\det(U) = 1$ therefore simply fixes a choice on the global phase of $U$. In other words, this group represents the group of matrices that preserves the norm of complex vectors, quotiented by the global phase.

## 9.3 Basic definitions

**Definition: Abelian group**

Let $(G, *)$ be a group.
$(G, *)$ is said to be an **Abelian group** if $*$ is commutative, i.e. if:

$$g_1 * g_2 = g_2 * g_1, \quad \forall g_1, g_2 \in G.$$

**Definition: Subgroup**

Let $(G, *)$ be a group, and $H \subseteq G$ be a subset.
$(H, *)$ is said to be a **subgroup** of $G$, if it is a group.

**Definition: Proper subgroup**

Let $G$ be a group, and $H \subseteq G$ be a subgroup.
If $H \neq \{e\}$ and $H \neq G$, then $(H, *)$ is said to be a **proper subgroup**.

**Lagrange theorem**

Let $G$ be a finite group and $H$ be a subgroup.
Then, $|H|$ divides $|G|$.

*Implication* This implies that groups with a prime number of elements have no proper subgroup. Moreover, it is possible to show that $\mathbb{Z}_n$ for a prime $n$ is the unique group with no proper subgroup. This is an argument that allows to show that $\mathbb{Z}_3$ is unique.

**Definition: Group homomorphism**

Let $(G, *)$ and $(H, \star)$ be two groups.
A function $\varphi : G \mapsto H$ is said to be a **homomorphism** between $G$ and $H$ if:

$$\varphi(g_1 * g_2) = \varphi(g_1) \star \varphi(g_2), \quad \forall g_1, g_2 \in G.$$

*Intuition* In other words, a homomorphism is a function between groups that preserves group operations.

**Definition: Group isomorphism**

Let $(G, *)$ and $(H, \star)$ be two groups, and $\varphi : G \mapsto H$ be a homomorphism.
If $\varphi$ is also bijective, we say it is an **isomorphism**.
If there exists such an isomorphism, $G$ and $H$ are said to be **isomorphic**. Intuitively, this captures the idea that two groups are essentially the same.

*Intuition* In other words, an isomorphism is a function between groups that sets up a one-to-one correspondance between elements, in such a way that it preserves group operations.

**Definition: Group representation**

Let $(G, *)$ be a group, and $V$ be a vector space.
A **representation** $R$ of $G$ on $V$ is a group homomorphism from $G$ to a set of invertible matrices that act on $V$.

*Remark* We will only consider $V = \mathbb{C}^m$ in this class.

*Personal remark* The set of invertible matrices that act on $V$ is typically marked as $GL_n(V)$. So, $R : G \mapsto GL_n(V)$.

*Terminology* Representations might be referred to as "reps".

*Example* For instance, any group has a trivial representation:

$$R(g) = I, \quad \forall g \in G.$$

As mentioned earlier, the following are all representations of the parity group $Z_2 = \{e, a\}$:

$$\{1, -1\}, \quad \{1, 1\}, \quad \{I, X\}, \quad \{I, Z\}, \quad \{I, SWAP\}.$$

Lie groups over matrices have their fundamental representations which are "just the obvious matrices that define the group". For instance, $O(3)$ is the group of orthogonal matrices in three dimensions, and this group's fundamental representation is just those matrices:

$$R(O) = O, \quad \forall O \in O(3).$$

— Monday 25th November 2024 — **Lecture 10 : Diagonalising matrices, slowly but Schur-ly**

**Definition: Equivalent representation**

Let $G$ be a group, and $R_1, R_2$ be two representations on a space of matrices of the same dimension.
$R_1$ and $R_2$ are said to be **equivalent** if they are equal up to a change of basis, i.e. there exists some invertible matrix $S$ so that, for all $g \in G$:

$$R_1(g) = SR_2(g)S^{-1}.$$

**Example: Tensor product representation**

Let $G$ be a group, and $R_1, R_2$ be representations.
The two following are also representations on $G$:

$$R_1(g) \otimes R_2(g), \quad R_1(g)^{\otimes k}.$$

*Intuition*

This directly comes from the fact that $(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D)$. In other words, each representation just multiply independently.

**Definition: Direct sum**

Let $A, B$ be two matrices.
We define their **direct sum**, written $A \oplus B$, as:

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

**Example: Direct sum representation**

Let $G$ be a group, and $R_1, R_2$ be representations.
Their direct sum $R_1(g) \oplus R_2(g)$ is also a representation of $G$.

## 9.4 Irreducible representations

**Idea**

The goal is to reduce mess looking representations (often tensor product representations) into direct sums of simpler representations. In other words, this amounts to doing a block diagonalisation of the representation.

**Example**

Let us consider a tensor product representation of $SU(2)$:

$$R(U) = U \otimes U.$$

Our goal is to write write $R(U)$ as a direct sum:

$$R_1(g) \otimes R_2(g) = \begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}.$$

We notice that $SWAP$ commutes with all elements of the representation:

$$\begin{aligned} & SWAP(U \otimes U)SWAP^\dagger = U \otimes U \\ \iff & SWAP(U \otimes U) = (U \otimes U)SWAP \\ \iff & [U \otimes U, SWAP] = 0. \end{aligned}$$

We know that if two operators commute, then they are block-diagonalisable in the same basis. We notice that $SWAP$ has eigenvalues 1 and $-1$. Its $\lambda = 1$ eigenspace is spanned by:

$$|00\rangle, \quad |11\rangle, \quad \frac{|01\rangle + |10\rangle}{\sqrt{2}}.$$

Equivalently, using the Bell basis, the $\lambda = 1$ eigenspace is spanned by:

$$|\varphi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |\varphi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \quad |\psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}.$$

The $\lambda = -1$ eigenspace is spanned by the last Bell state:

$$|\psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

Therefore, in the Bell basis, there exists some matrices $U_1 \in \mathbb{C}^{3\times 3}$ and $U_2 \in \mathbb{C}$ so that:

$$U \otimes U = \begin{pmatrix} & & & 0 \\ & U_1 & & 0 \\ & & & 0 \\ 0 & 0 & 0 & U_2 \end{pmatrix} = U_1 \oplus U_2.$$

It is possible to convince ourselves that $U_1$ and $U_2$ are still unitaries, i.e. that $U_1 \in SU(3)$ and $U_2 \in SU(1) = \{1\}$. Thanks to this block-diagonal form, we can consider invariant subspaces (spaces that are left unchanged by any $U \otimes U$). In other words, looking at $U \otimes U = U_1 \oplus U_2$:

$$(U \otimes U)|\psi^-\rangle = |\psi^-\rangle,$$

$$(U \otimes U)\big(a|\varphi^+\rangle + b|\varphi^-\rangle + c|\psi^+\rangle\big) = d|\varphi^+\rangle + e|\varphi^-\rangle + f|\psi^+\rangle.$$

We notice that the two invariant subspaces are the symmetric and antisymmetric one. In other words, the first one consists of bosons and the second one of bosons and, physically, it makes sense that we cannot turn a photon to an electron.

*Takeaway*

The important things to take away from this example is that we can sometimes split a representation into a direct sum, in which case it shows invariant subspaces.

Note that, in some sense, this is what we did for addition of angular momentum in Quantum physics 1. We will later see that there is indeed a strong link, hence the importance.

**Definition: Reducibility**

Let $G$ be a group and $R : G \mapsto GL_n(V)$ be a representation.

$R$ is said to be **reducible** if there exists an invariant subspace. In other words, it is reducible if there exists some non-trivial subspace $W$ (i.e. $W \neq \{0\}$ and $W \neq V$) so

that, for each $|w\rangle \in W$:

$$R(g)|w\rangle \in W.$$

A representation is then said to be **irreducible** if it is not reducible. An irreducible representation is called an irrep.

**Definition: Complete reducibility** A representation $R(g)$ of a group $G$ is completely reducible if it splits into a direct sum of irreducible representation:

$$R(g) = \bigoplus_{j=1}^{k} R_j(g) = \begin{pmatrix} R_1(g) & & & \\ & R_2(g) & & \\ & & \ddots & \\ & & & R_k(g) \end{pmatrix}.$$

*Remark* We trivially notice that a completely reducible representation is reducible.

**Proposition** Let $R(g)$ be a unitary representation, i.e. $R(g) \in U(n)$ for all $g$.
If $R(g)$ is reducible, then it is completely reducible.

*Remark* In other words, a unitary representation is reducible if and only if it is completely reducible. We will only consider unitary representations in this class, we will thus consider the two interchangeably.

However, this is not necessarily true for arbitrary representations. For instance, the following is a representation of $(\mathbb{R}, +)$:

$$M(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

We do have $M(x)M(y) = M(x+y)$, so this is a valid representation. Moreover, it is reducible, since $M(x)|0\rangle = |0\rangle$. However, it is not completely reducible.

**Proposition** Let $G$ be a group. Then, the trivial representation $R(g) = 1$ for all $g \in G$, is completely reducible.

*Remark* If we are asked to find all the reducible representations of a group during an exam, then we can thus get some free points by citing the trivial representation.

*Proof* By definition, any one-dimensional representation is completely reducible.

□

**Schur's first lemma** Let $G$ be a group, and $R_1(g), R_2(g)$ be two non-equivalent irreducible representations.
If there is a matrix $A$ such that $AR_1(g) = R_2(g)A$ for all $g \in G$, then $A = 0$.

*Implication* So, if there exists some matrix $A$ so that $AR_1(g) = R_2A$ and $R_1, R_2$ are non-equivalent, then they are reducible.

**Schur's second lemma** Let $G$ be a group, and $R(g)$ be an irreducible representation.
If $AR(g) = R(g)A$ for all $g \in G$, then $A = \lambda I$ for some $\lambda \in \mathbb{C}$.

*Implication* So, if there is some non-identity matrix $A \neq \lambda I$ that commutes with everything $[A, R(g)] = 0$—for instance a symmetry—then the representation is reducible.

**Remark** Schur's lemma give us a means of indicating if a representation is reducible. As we will see right after, they also help us to block diagonalise operators.

**Property** Let $G$ be an Abelian group.
Then, its irreducible representations are one dimensional.

*Example*

For instance, let us consider the following representation of the $\mathbb{Z}_2$ group:

$$R(e) = I, \quad R(a) = X.$$

We notice that $A = X$ commutes with both $I$ and $X$. Since $X \neq \lambda I$, this means that $R$ is reducible. Since it is a unitary representation, it is also completely reducible. In other words, there is some basis so that:

$$R(g) = \begin{pmatrix} R_1(g) & \\ & R_2(g) \end{pmatrix}.$$

We recall that there are two 1D representations:

$$\{1, 1\}, \quad \{1, -1\}.$$

$R_1$ and $R_2$ must be either of those representations.
Moreover, using the $\{|+\rangle, |-\rangle\}$ basis (since we use the matrix $A = X$):

$$I = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

So, we clearly see that taking $R_1(g) = 1$ to be the trivial representation, and $R_2(I) = 1$ and $R_2(X) = -1$ to be the signed representation, we do get a valid representation, which is diagonal:

$$R(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R(X) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

*Proof*

Let $R$ be an arbitrary irreducible representation of $G$, and let $h \in G$ be arbitrary. By definition of representation, we notice that, for all $g \in G$:

$$R(g)R(h) = R(g * h) = R(h * g) = R(h)R(g).$$

However, this means that the matrix $R(h)$ commutes with all elements of the representation $R(g)$. By Schur's second lemma, we must have $R(h) = \lambda I$ for some $\lambda \in \mathbb{C}$. However, $R(h) = \lambda I$ is only irreducible if it is one-dimensional, so $R(h)$ is one-dimensional.

□

## 9.5 Block diagonalisation

**Observation**

In a quantum context, we are interested in studying systems (represented by some Hamiltonian $H$) with certain symmetries, i.e. we will have some Hamiltonian such that $[R(g), H] = 0$ for all $g \in G$, where $G$ is some symmetry group and $R(g)$ is a unitary transformation that represents that symmetry group on the Hilbert space of our system. Irreps tell us information on the structure of $H$.

**Theorem**

We state this theorem using an example, it can be generalised easily, but making it formal requires heavy notations, which would only decrease its pedagogical interest.
Let $H$ be Hamiltonian. Also, let $G$ be a group, and $R(g)$ be a representation so that:

$$[R(g), H] = 0, \quad \forall g \in G.$$

We moreover us suppose that $R$ can be decomposed into irreducible representations as follows:

$$R(g) = R_1(g) \oplus R_2(g) \oplus R_3(g) = \begin{pmatrix} R_1(g) & & \\ & R_2(g) & \\ & & R_3(g) \end{pmatrix}.$$

Furthermore, let us consider a representation of $H$ in this basis:

$$H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix},$$

where each $H_{ij}$ is a matrix, and $H_{ii}$ is a matrix of the same dimension as $R_i(g)$.
If $R(g) = R_1(g) \oplus R_2(g) \oplus R_3(g)$ is ordered so that equivalent representations are next to each others, and that $R_1(g)$ is equivalent to $R_2(g)$, but not to $R_3(g)$, then:

$$H = \begin{pmatrix} \lambda_1 I_1 & H_{12} & \\ H_{21} & \lambda_2 I_2 & \\ & & \lambda_3 I_3 \end{pmatrix}.$$

*Implication*

This implies that the dimension of the irreducible representations tells us the size of the degeneracies, when the irreps are non-equivalent. For instance, we have $d_3 = \dim I_3$ states with eigenvalue $\lambda_3$. In other words, the degeneracies of an Hamiltonian are determined by the dimension of its non-equivalent irreducible representations, and thus degeneracies of the Hamiltonian come from symmetries.

*Prof*

We notice that:

$$\begin{aligned} R(g)H &= \begin{pmatrix} R_1(g) & & \\ & R_2(g) & \\ & & R_3(g) \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \\ &= \begin{pmatrix} R_1(g)H_{11} & R_1(g)H_{12} & R_1(g)H_{13} \\ R_2(g)H_{21} & R_2(g)H_{22} & R_2(g)H_{23} \\ R_3(g)H_{31} & R_3(g)H_{32} & R_3(g)H_{33} \end{pmatrix}. \end{aligned}$$

Similarly:

$$\begin{aligned} HR(g) &= \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} R_1(g) & & \\ & R_2(g) & \\ & & R_3(g) \end{pmatrix} \\ &= \begin{pmatrix} H_{11}R_1(g) & H_{12}R_2(g) & H_{13}H_3(g) \\ H_{21}R_1(g) & H_{22}R_2(g) & H_{23}R_3(g) \\ H_{31}R_1(g) & H_{32}R_2(g) & H_{33}R_3(g) \end{pmatrix}. \end{aligned}$$

However, by hypothesis $R(g)H = R(g)H$ for all $g \in G$. This means that $R_j(g)H_{jk} = H_{jk}R_k(g)$ for any $j, k$. We consider the case $j = k$ (the diagonal terms) and the case $j \neq k$ (the non-diagonal terms) differently:

- We suppose $j = k$. We know that $R_k(g)H_{kk} = H_{kk}R_k(g)$. By Schur's second lemma, this means $H_{kk} = \lambda_k I$ for some $\lambda_k \in \mathbb{C}$.
- We suppose $j \neq k$, and that $R_j$ and $R_k$ are not equivalent relations. By Schur's first lemma, $H_{jk} = 0$.

This implies that $H$ is block diagonal (ordering $R_1 \oplus R_2 \oplus \dots$ in such a way that equivalent representations are next to each others). In our case, $R_1$ and $R_2$ are equivalent, and they are not equivalent

to $R_3$, so:

$$H = \begin{pmatrix} \lambda_1 I_1 & H_{12} & \\ H_{21} & \lambda_2 I_2 & \\ & & \lambda_3 I_3 \end{pmatrix}.$$

This reasoning can easily be generalised to any other scenario.

□

**Example**

Let us consider the Hamiltonian of two indistinguishable quits. It is easier to first notice the representation of the symmetry group:

$$I \otimes I, \quad SWAP.$$

The symmetry group that has this representation is the parity group $\mathbb{Z}_2$, and the representation we took is $R(e) = I$ and $R(a) = SWAP$. In the computation basis, we know that:

$$SWAP = \begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}.$$

As explained before, SWAP is diagonalised in the Bell basis $(|\psi^+\rangle, |\psi^-\rangle, |\varphi^+\rangle, |\varphi^-\rangle)$.

$$SWAP = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}.$$

We consider again $R_1(g) = 1$ to be the trivial representation, and $R_2(e) = 1$ and $R_2(a) = -1$ to be the sign representation. Therefore, in the Bell basis, we get:

$$R(g) = \begin{pmatrix} R_1(g) & & & \\ & R_1(g) & & \\ & & R_1(g) & \\ & & & R_2(g) \end{pmatrix} = R_1(g) \oplus R_1(g) \oplus R_1(g) \oplus R_2(g).$$

We notice that $R_1$ is equivalent to $R_1$ (trivially), but $R_1$ and $R_2$ are not equivalent. This yields:

$$H = \begin{pmatrix} \lambda_1 & H_{12} & H_{13} & 0 \\ H_{21} & \lambda_2 & H_{23} & 0 \\ H_{31} & H_{32} & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}.$$

*Remark* This is not a huge result as is, but if we were considering a more complex system this would have been a huge help.

Monday 2nd December 2024 — **Lecture 11 : Integrating over groups**

## 9.6 Finding the number of irreps

**Definition: Equivalent elements**

Let $G$ be a group, and $h, h' \in G$.
$h$ and $h'$ are said to be **equivalent**, written $h \sim h'$, if there exists some $g \in G$ so that:

$$h' = g^{-1}hg.$$

*Intuition*

Intuitively, some elements of a group do very different things, such as a $\frac{\pi}{10}$ rotation around the $z$-axis and a $\frac{\pi}{20}$ rotation about the $z$-axis, or $I$ and $SWAP$. However, some other elements are, in some sense, equivalent such as a $\frac{\pi}{10}$ rotation around the $x$-axis and a $\frac{\pi}{10}$ rotation about the $z$-axis, or permutations $\begin{pmatrix}1 & 2 & 3\end{pmatrix}$ and $\begin{pmatrix}1 & 3 & 2\end{pmatrix}$. In fact, we can do a change of basis to show they are indeed similar:

$$\begin{pmatrix}2 & 3\end{pmatrix}^{-1}\begin{pmatrix}1 & 2 & 3\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix} = \begin{pmatrix}1 & 3 & 2\end{pmatrix}.$$

This is of the form $g^{-1}hg = h'$, justifying the definition.

*Remark*

This is an equivalence relation, so we can consider the following definition.

**Definition: Conjugacy classes**

Let $G$ be a group.
Its **conjugacy classes** are the equivalence classes based on the relation defined above:

$$C_h = \left\{g^{-1}hg \;\middle|\; g \in G\right\}, \quad h \in G.$$

Its number of distinct conjugacy classes is written $N_c$.

**Example**

We consider the permutation group over three elements again:

$$C3v = S_3 = \left\{e, \begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}, \begin{pmatrix}2 & 3\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\end{pmatrix}, \begin{pmatrix}1 & 3 & 2\end{pmatrix}\right\}.$$

An element is equivalent with any elements that have the same orbit structure. So, we have the three following conjugacy classes, showing $N_C = 3$:

$$\{e\}, \quad \left\{\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}, \begin{pmatrix}2 & 3\end{pmatrix}\right\}, \quad \left\{\begin{pmatrix}1 & 2 & 3\end{pmatrix}, \begin{pmatrix}1 & 3 & 2\end{pmatrix}\right\}.$$

**Property**

The identity will always be alone in its own conjugacy class:

$$C_e = \{e\}.$$

*Proof*

For any $g \in G$:

$$g^{-1}eg = g^{-1}g = e.$$

So, by definition, $C_e = \{e\}$.

□

**Theorem**

Let $G$ be a finite group, with $n$ non-equivalent irreps and $N_c$ distinct conjugacy classes.
Then, $n = N_c$.

*Proof*

We will later prove that $n \leq N_c$ using the Petit orthogonality theorem. Proving that $n \geq N_c$ can be done using the Grand orthogonality theorem.

**Burnside's lemma**

Let $G$ be a finite group with $n$ non-equivalent irreps, of dimensions $\ell_a$ for $a \in \{1, \ldots, n\}$ respectively.

Then:

$$\sum_{a=1}^{n} \ell_a^2 = |G|.$$

*Remark* While the previous theorem allows to know how many irreps we are missing, this lemma allows to make a guess on the dimension of those irreps.

**Corollary** Let $G$ be a finite group.
If $G$ is Abelian, $G$ has $|G|$ non-equivalent irreps.

**Example** For instance, let us consider again $\mathbb{Z}_2$. It has two irreps:

$$e \mapsto 1, a \mapsto -1, \quad e \mapsto 1, a \mapsto 1.$$

Both have dimension 1, so it does respect Burnside's lemma:

$$\ell_1^2 + \ell_2^2 = 1^2 + 1^2 = 2 = |Z_2|.$$

**Remark** Our first theorem allows us to know how many irreps we have, and Burnside's lemma puts strong constraints on their dimensions. Those two theorems thus help us know what irreps we are missing.
The orthogonality theorems, which come right after, are other tools to help on this idea.

**Observation** We can think of irreps as giving a vector of matrices:

$$(R(g))_{g\in G} = \begin{pmatrix} R(g_1) \\ R(g_2) \\ \vdots \end{pmatrix}.$$

**Grand orthogonality theorem** Let $G$ be a group be a finite group of order $N$. Also, let $R_a$ and $R_b$ be two non-equivalent unitary irreps of $G$, of dimension $n_a$ and $n_b$ respectively.
Then, for all $c, d \in \{a, b\}$:

$$\sum_{g\in G} \frac{n_c}{N} \left[R_c(g)^\dagger\right]_{jk} [R_d(g)]_{\ell m} = \delta_{cd}\delta_{jm}\delta_{\ell k}.$$

*Intuition* There are many proposition in this theorem. It states that:

- $\sum_{g\in G} \left[R_a(g)^\dagger\right]_{jk} [R_b(g)]_{\ell m} = 0.$
- If $j \neq m$ or $\ell \neq k$, then $\sum_{g\in G} \left[R_a(g)^\dagger\right]_{jk} [R_a(g)]_{\ell m} = 0.$
- $\sum_{g\in G} \left[R_a(g)^*\right]_{kj} [R_a(g)]_{kj} = \frac{N}{n_a}.$

Note that the $\frac{N}{n_a}$ is some form of normalisation coefficient.

*Personal intuition* Another way to understand this theorem, which justifies its name, is to notice that we can think of irreps as giving a vector of matrices:

$$(R(g))_{g\in G} = \begin{pmatrix} R(g_1) \\ R(g_2) \\ \vdots \end{pmatrix}.$$

In particular, we can only look at the $(j, k)$ component of each matrix, giving some vector:

$$\left(R(g)_{jk}\right)_{g\in G} = \begin{pmatrix} R(g_1)_{jk} \\ R(g_2)_{jk} \\ \vdots \end{pmatrix}.$$

What this theorem states is that the following set is an orthonormal set of vectors:

$$\left\{ \sqrt{\frac{n_c}{N}} \left( R_c(g)_{jk} \right)_{g\in G} \;\middle|\; c \in \{a,b\}, j, k \right\} = \left\{ \sqrt{\frac{n_c}{N}} \begin{pmatrix} R_c(g_1)_{jk} \\ R_c(g_2)_{jk} \\ \vdots \end{pmatrix} \;\middle|\; c, j, k \right\}.$$

*Remark* This theorem can be used to get missing irreps, as mentioned earlier.

**Example 1** We consider $\mathbb{Z}_2$ and its two irreps:

$$R_1(e) = 1, R_1(a) = 1, \quad R_2(e) = 1, R_2(a) = -1.$$

We want to verify the grand orthogonality theorem in this context. We notice that:

$$\sum_{g\in G} R_1(g)^\dagger R_2(g) = 1 \cdot 1 + 1 \cdot (-1) = 0.$$

Similarly, since $N = 2$ and $n_1 = n_2 = 1$:

$$\sum_{g\in G} R_1(g)^\dagger R_1(g) = 1^2 + (-1)^2 = 2 = \frac{2}{1} = \frac{N}{n_1},$$

$$\sum_{g\in G} R_2(g)^\dagger R_2(g) = 1^2 + 1^2 = 2 = \frac{N}{n_2}.$$

**Example 2** We consider $C3v$. The following is a 2D irrep for this group:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c_+ = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad c_- = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$\sigma = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma' = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \sigma'' = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

We notice that, considering that the $1,1$ element is the bottom-right corner of the matrix (using 0-indexing, like qubits $|0\rangle, |1\rangle$):

$$\sum_{g\in G} \left[R^\dagger(g)\right]_{11} R(g)_{11} = 1^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 = 3 = \frac{6}{2} = \frac{N}{n},$$

as expected by the grand orthogonality theorem.

## 9.7 Group Averaging

**Definition: Group averaging** Let $G$ be some finite group of cardinality $N$, $R(g) = U_g$ be some unitary $d$-dimensional representation of $G$, and $X$ be some $d \times d$ unitary.
We define the **group averaging** (or **group twirling**) of $X$ as:

$$\langle X \rangle_G = \frac{1}{N} \sum_{g\in G} R(g) X R(g)^\dagger = \frac{1}{N} \sum_{g\in G} U_g X U_g^\dagger.$$

**Proposition** Let $G$ be a finite group of cardinality $N$, $R$ be a unitary $d$-dimensional representation of $G$, and $X$ be a $d \times d$ unitary.
If $R$ is irreducible, we have:

$$\langle X \rangle_G = \frac{1}{d} \operatorname{Tr}(X) I.$$

*Remark* This generalises to compact groups $G$:

$$\langle X\rangle_G = \int_G d\mu(g) U(g) X U(g)^\dagger = \frac{1}{d}\operatorname{Tr}(X) I,$$

where $d\mu(g)$ is the continuous equivalent of $\frac{1}{N}$, generalising the idea of uniform distribution.
Note that $O(n)$, $SO(n)$, $U(n)$ and $SU(n)$ are compact groups.

*Proof* Expanding the matrix products using a sum, we get:

$$\begin{aligned}\langle X\rangle_G &= \frac{1}{N}\sum_{j,k,\ell,m}\sum_g [R(g)]_{\ell m} X_{mj}\left[R(g)^\dagger\right]_{jk}|\ell\rangle\langle k| \\ &= \frac{1}{N}\sum_{j,k,\ell,m}\left(\sum_g [R(g)]_{\ell m}\left[R(g)^\dagger\right]_{jk}\right) X_{mj}|\ell\rangle\langle k|.\end{aligned}$$

We can now use the Grand orthogonality theorem:

$$\langle X\rangle_G = \frac{1}{d}\sum_{j,k,\ell,m}\delta_{\ell k}\delta_{jm} X_{mj}|\ell\rangle\langle k| = \frac{1}{d}\sum_j X_{jj}\sum_k |k\rangle\langle k| = \frac{1}{d}\operatorname{Tr}(X) I.$$

□

**Example** We consider the single-qubit Pauli group:

$$G = \{\varphi P \mid \varphi \in \{\pm 1, \pm i\}, P \in \{I, \sigma_x, \sigma_y, \sigma_z\}\}.$$

Pauli matrices are dimension $d = 2$. So, given any state $\rho$:

$$\langle \rho\rangle_G = \frac{1}{2}\operatorname{Tr}(\rho) I = \frac{1}{2} I.$$

This is the maximally mixed state, which makes sense intuitively: letting $\rho$ evolve under an average Pauli, should give a state which representation on the Bloch sphere is not biased towards any direction.
Let us also compute this result using the definition of group averaging:

$$\langle \rho\rangle_G = \frac{1}{16}\sum_g U_g \rho U_g^\dagger = \frac{1}{16}\sum_{\varphi,P}\varphi P \rho \varphi^* P^\dagger.$$

We notice that $\varphi\varphi^* = |\varphi|^2 = 1$ for all $\varphi \in \{\pm 1, \pm i\}$. This yields that:

$$\langle \rho\rangle_G = \frac{4}{16}\sum_P P\rho P = \frac{1}{4}\sum_P P\rho P.$$

We know that we can express $\rho = \frac{1}{2}(I + r \bullet \sigma)$, and the identity $\sigma_i\sigma_j\sigma_i = -\sigma_j$ for $i \neq j$ *(the Professor left a note on the black board: USEFUL, REMEMBER ME)*, so:

$$\sigma_x\rho\sigma_x = \frac{1}{2}(I + r_x\sigma_x - r_y\sigma_y - r_z\sigma_z), \quad \sigma_y\rho\sigma_y = \frac{1}{2}(I - r_x\sigma_x + r_y\sigma_y - r_z\sigma_z),$$

$$\sigma_z\rho\sigma_z = \frac{1}{2}(I - r_x\sigma_x - r_y\sigma_y + r_z\sigma_z), \quad I\rho I = \frac{1}{2}(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z).$$

Adding everything, we notice that the coefficients in front of $I$ add to 2, and the coefficients in front of all other Paulis cancel out, yielding indeed:

$$\langle \rho\rangle_G = \frac{2I}{4} = \frac{1}{2} I.$$

**Proposition** Let $G$ be a finite group of cardinality $N$, $R(g) = U(g)$ be a unitary $d$-dimensional representation of $G$, and $X$ be a $d \times d$ unitary.

We moreover suppose that we can express $U(g)$ as a direct sum of irreducible representations:

$$U(g) = \bigoplus_x U_x(g) = \begin{pmatrix} U_1(g) & & \\ & 0 & \\ & & \ddots \end{pmatrix} + \begin{pmatrix} 0 & & \\ & U_2(g) & \\ & & \ddots \end{pmatrix} + \dots.$$

Finally, we note $\Pi_x$ to be the projector on the subspace represented by $U_x(g)$ (i.e., to be a projector on its corresponding block), $I_x$ to be the identity matrix on the subspace represented by $U_x(g)$, and $d_x$ be the dimension of $U_x(g)$ (i.e. $U_x(g) \in \mathbb{C}^{d_x \times d_x}$).

If all representations $U_x$ are non-equivalent, then:

$$\langle X \rangle_G = \sum_x \frac{1}{d_x} \operatorname{Tr}(X\Pi_x)\Pi_x = \bigoplus_x \frac{1}{d_x} \operatorname{Tr}(X\Pi_x) I_x.$$

*Remark 1* Note that, sometimes, we write:

$$\langle X \rangle_G = \bigoplus_x \frac{1}{d_x} \operatorname{Tr}(X\Pi_x)\Pi_x.$$

This is however only an abuse of notations.

*Remark 2* This proposition also generalises to compact groups.

*Remark 3* Note that this proposition is not true in general if there exists $x \neq x'$ so that $U_x$ and $U_{x'}$ are equivalent.

Consider for instance the trivial group $G = \{e\}$, with the representation:

$$U_g = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \oplus 1.$$

We use twice the trivial irrep, which are naturally equivalent. Considering $X = \sigma_X$ to be the Pauli matrix, we notice that:

$$\langle X \rangle_G = IXI^\dagger = X,$$

$$\frac{1}{1} \operatorname{Tr}(X|0\rangle\langle 0|)|0\rangle\langle 0| + \frac{1}{1} \operatorname{Tr}(X|1\rangle\langle 1|)|1\rangle\langle 1| = 0.$$

*Proof* Let us note $|x, i\rangle$ to be a basis for $U(g)$, where the $x$ indexes the representation and the $i$ indexes the elements of the representation. In other words:

$$\langle x, i|U(g)|x', j\rangle = \delta_{xx'}\langle i|U_x(g)|j\rangle.$$

We notice that:

$$\begin{aligned} \langle X \rangle_G &= \frac{1}{N} \sum_g U(g) X U(g)^\dagger \\ &= \frac{1}{N} \sum_{x,x'} \sum_{i,j,k,\ell} \sum_g (U_x(g))_{ij} \langle x, j|X|x', k\rangle \left(U_x(g)^\dagger\right)_{k\ell} |x, i\rangle\langle x', \ell|. \end{aligned}$$

We can apply the Grand orthogonality theorem since there is no representations $U_x$ and $U'_x$ with $x \neq x'$ that are equivalent. Our

equation therefore simplifies to:

$$\begin{aligned}\langle X\rangle_G &= \frac{1}{N}\sum_{x,x'}\sum_{i,j,k,\ell}\langle x,j|X|x',k\rangle\delta_{xx'}\delta_{i\ell}\delta_{jk}\frac{N}{d_x}|x,i\rangle\langle x',\ell| \\ &= \sum_x \frac{1}{d_x}\sum_j\langle x,j|X|x,j\rangle\sum_i|x,i\rangle\langle x,i| \\ &= \sum_x \frac{1}{d_x}\operatorname{Tr}(\Pi_x X)\Pi_x.\end{aligned}$$

□

**Example**

Let us consider $R_Z(\theta)=\exp(-i\theta\sigma_z)$. We want to compute:

$$\langle\rho\rangle_{R_Z(\theta)} = \frac{1}{2\pi}\int_0^{2\pi} d\theta R_Z(\theta)\rho R_Z(\theta)^\dagger.$$

$R_Z(\theta)$ is a representation of $U(1)$ (any element of $U(1)$ is on the unit circle, and can thus be interpreted as an angle). However, it is not an irreducible representation: $U(1)$ is Abelian, so all its irreducible representations have dimension 1. $U(1)$ has many irreps (infinitely many in fact), but it is possible to convince ourselves that the trivial one $U_1(\theta)=1$ and $U_2(\theta)=e^{i\theta}$ are examples of irreps, and that they are non-equivalent. We can decompose $R_Z(\theta)$ as a direct sum of irreps:

$$R_Z(\theta)=\exp(-i\theta\sigma_Z)=\begin{pmatrix}1 & 0\\ 0 & e^{i\theta}\end{pmatrix}=\begin{pmatrix}U_1(\theta) & \\ & U_2(\theta)\end{pmatrix}.$$

Note that here, we are in fact really considering $R_Z(\theta)'=\exp(-i\theta(\sigma_z+I))$ for slightly simpler computations. However, it only differs from $R_Z(\theta)$ by a global phase, so they are physically equivalent.
Both are our irreps are $1\times 1$ (i.e. $d_1=d_2=1$), so $\Pi_0=|0\rangle\langle 0|$ and $\Pi_1=|1\rangle\langle 1|$. This yields:

$$\begin{aligned}\langle\rho\rangle_\theta &= \frac{1}{2\pi}\int_0^{2\pi} R_Z(\theta)\rho R_Z(\theta)^\dagger d\theta \\ &= \bigoplus_x \frac{1}{d_x}\operatorname{Tr}(\rho\Pi_x)\Pi_x \\ &= \operatorname{Tr}(\rho|0\rangle\langle 0|)|0\rangle\langle 0| + \operatorname{Tr}(\rho|1\rangle\langle 1|)|1\rangle\langle 1| \\ &= \langle 0|\rho|0\rangle|0\rangle\langle 0| + \langle 1|\rho|1\rangle|1\rangle\langle 1|.\end{aligned}$$

This makes intuitive sense: we project on the $z$-axis. This is where the term twirling comes from: by averaging the position over a rotation along the $z$-axis, we loose any information on the $x$ and $y$ axes.

Monday 9th December 2024 — **Lecture 12 : The climax**

## 9.8 Petit orthogonality theorem

**Remark** We notice that the grand orthogonality theorem is an orthogonality relation between vectors of matrices, so it may be hard to exploit. We want to build the petit orthogonality theorem, which is an orthogonality relation between composant of traces of representations.

**Definition: Character** Let $G$ be a group, and $R$ be a representation.
We define the **character** of $R$ as the set $\{\chi_R(g) \mid g \in G\}$, where, for any $g$:

$$\chi_R(g) = \mathrm{Tr}(R(g)).$$

*Property* Note that, for instance:

$$\mathrm{Tr}\Big(R(g)^\dagger\Big) = \mathrm{Tr}(R(g))^* = \chi_R(g)^*.$$

**Property** The character of two equivalent representations are the same.

*Proof* This directly comes from the fact that a change of basis leaves the trace unchanged.
More formally, let $R_1, R_2$ be equivalent representations. By definition, this means that there exists some invertible matrix $P$ so that, for all $g \in G$:

$$R_1(g) = PR_2(g)P^{-1}.$$

But then, for all $g$:

$$\mathrm{Tr}(R_1(g)) = \mathrm{Tr}\big(PR_2(g)P^{-1}\big) = \mathrm{Tr}\big(P^{-1}PR_2(g)\big) = \mathrm{Tr}(R_2(g)).$$

□

**Proposition** Let $G$ be a group, and $R$ be a representation.
The representation of two elements of $G$ in the same conjugacy class have the same trace.

*Proof* Let $x, y$ be in the same conjugacy class, i.e.

$$x = u^{-1}yu.$$

We notice that, using first the homomorphism property and then the cyclicity of the trace:

$$\begin{aligned}
\mathrm{Tr}(R(x)) &= \mathrm{Tr}\big(R\big(u^{-1}yu\big)\big) \\
&= \mathrm{Tr}\big(R\big(u^{-1}\big)R(y)R(u)\big) \\
&= \mathrm{Tr}\big(R(u)R\big(u^{-1}\big)R(y)\big) \\
&= \mathrm{Tr}\big(R\big(u^{-1}u\big)R(y)\big) \\
&= \mathrm{Tr}(R(y)).
\end{aligned}$$

□

**Petit orthgonality theorem** Let $G$ be a finite group of order $N$. Also, let $R_a$ and $R_b$ be two non-equivalent unitary irreps of $G$, of dimension $n_a$ and $n_b$ respectively.
Then, for all $c, d \in \{a, b\}$:

$$\sum_{g \in G} \chi_a^*(g)\chi_b(g) = N\delta_{ab}.$$

Equivalently, noting $C_\mu$ to be equivalence classes, $n_\mu = |C_\mu|$, and $N_c$ to be the number of equivalence classes, then, for all $c, d \in \{a, b\}$:

$$\sum_{\mu=1}^{N_c} n_\mu \chi_a^*(C_\mu)\chi_b(C_\mu) = N\delta_{ab}.$$

*Rephrasing* Considering $N_c$ to be the number of conjugacy classes, and $n_\mu$ to be the number of elements in the $\mu^{\text{th}}$ conjugacy class, we can easily rephrase this theorem as:

$$\sum_{\mu=1}^{N_c} n_\mu \chi_a^*(C_\mu)\chi_b(C_\mu) = N\delta_{ab}.$$

This might be easier to use.

*Remark* This theorem states that the following set of vectors is orthonromal:

$$\left\{ \left[ \sqrt{\frac{n_\mu}{N}} \chi_a(C_\mu) \right]_{\mu=1}^{N_c} \;\middle|\; a \right\} = \left\{ \sqrt{\frac{1}{N}} \begin{pmatrix} n_1\chi_a(C_1) \\ n_2\chi_a(C_2) \\ \cdots \end{pmatrix} \;\middle|\; a \right\}$$

Note that each vector is $N_c$-dimensional. However, there can be at most $N_c$ orthonormal vectors of dimension $N_c$. Therefore, there are at most $N_c$ non-equivalent irreps. Recall that we claimed last lecture there are exactly $N_c$ non-equivalent irreps, this is thus one side of the argument.

*Proof* From the Grand orthogonality theorem, we notice that:

$$\begin{aligned} \sum_{g\in G} \frac{n_a}{N} \chi_a^*(g)\chi_b(g) &= \sum_{jk} \sum_{g\in G} \frac{n_a}{N} \left[ R_a(g)^\dagger \right]_{jj} [R_b(g)]_{kk} \\ &= \delta_{ab} \sum_{j=1}^{n_a} \sum_{k=1}^{n_a} \delta_{jk}\delta_{jk} \\ &= \delta_{ab} n_a, \end{aligned}$$

We get our result by multiplying by $\frac{N}{n_a}$ on both sides.

□

*Remark* We lost information when deriving this theorem. It might be easier to use than the Petit orthogonality theorem, but it is less powerful.

**Example** We want to verify our results for the following irrep of the $C3v$ group. Recall that the following is an irrep of this group:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c_+ = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad c_- = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$\sigma = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma' = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \sigma'' = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

We notice that, first, the trace of elements in the same conjugacy class are the same:

$$\chi(e) = 2, \quad \chi\big(c^+\big) = \chi\big(c^-\big) = -1, \quad \chi(\sigma) = \chi(\sigma') = \chi(\sigma'') = 0.$$

Moreover, the petit orthogonality theorem is indeed verified:

$$\sum_{\mu=1}^{N_C} n_\mu \chi^*(C_\mu)\chi(C_\mu) = 1 \cdot 2^2 + 3 \cdot 0^2 + 2 \cdot (-1)^2 = 6 = N.$$

**Theorem** Let $G$ be a group, and $R(g)$ be a representation.
We know we can write $R(g)$ as a direct sum:

$$R(g) = \bigoplus_a \bigoplus_{x=1}^{b_a} R_{a,x}(g),$$

where $a$ is an index for different non-equivalent irreps, and $x \in \{1, \dots, b_a\}$ is an index for the degeneracy of the irrep.
Then, noting $\chi_a(g) = \mathrm{Tr}(R_a(g))$ and $\chi_R(g) = \mathrm{Tr}(R(g))$:

$$b_a = \frac{1}{N} \sum_{\mu=1}^{N_c} n_\mu \chi_a^*(C_\mu) \chi_R(C_\mu).$$

*Intuition* This theorem allows to compute the degeneracy $b_a$ of an irreducible representation in the direct sum expansion of some representation $R(g)$.

*Proof* We notice that:

$$\begin{aligned}
\chi_R(g) &= \mathrm{Tr}(R(g)) \\
&= \mathrm{Tr} \begin{pmatrix} R_{1,1}(g) & & & & & & \\ & R_{1,2}(g) & & & & & \\ & & \ddots & & & & \\ & & & R_{1,b_1}(g) & & & \\ & & & & R_{2,1}(g) & & \\ & & & & & \ddots & \end{pmatrix} \\
&= \sum_{a,x} \mathrm{Tr}(R_a(g)) \\
&= \sum_a b_a \, \mathrm{Tr}(R_a(g)) \\
&= \sum_a b_a \chi_a(g).
\end{aligned}$$

Injecting this in the petit orthogonality theorem, this means:

$$\begin{aligned}
\sum_{\mu=1}^{N_c} n_\mu \chi_b^*(C_\mu) \chi_R(C_\mu) &= \sum_{\mu=1}^{N_c} n_\mu \sum_a b_a \chi_b^*(C_\mu) \chi_a(C_\mu) \\
&= \sum_a b_a N \delta_{ab} \\
&= N b_b.
\end{aligned}$$

This yields our result, by dividing both sides by $N$ (and replacing the pseudo-variable $b$ by $a$).

□

**Theorem** Let $G$ be a group of order $N$, and $R$ be a representation.
$R$ is an irrep if and only if:

$$\sum_{\mu=1}^{N_c} n_\mu |\chi_R(C_\mu)|^2 = N.$$

*Remark* This theorem finally allows us to know if a representation is irreducible.

*Proof* Just like the previous theorem, we decompose our representation $R$ into a direct sum of irreps:

$$R(g) = \bigoplus_a \bigoplus_{x=1}^{b_a} R_{a,x}(g),$$

Note that this is an irrep if and only if there is some $i$ so that $b_i = 1$ and for all $j \neq i$, $b_j = 0$. We found that:

$$\chi_R(g) = \sum_a b_a \chi_a(g).$$

Therefore, feeding it to the Petit orthogonality theorem:

$$\begin{aligned}\sum_{\mu=1}^{N_c} n_\mu |\chi_R(C_\mu)|^2 &= \sum_{i,j} b_i b_j \sum_{\mu=1}^{N_c} n_\mu \chi_i(C_\mu)^* \chi_j(C_\mu) \\ &= N \sum_{i,j} b_i b_j \delta_{ij} \\ &= N \sum_i b_i^2.\end{aligned}$$

As mentioned above, $R$ is an irrep if and only if a single $b_i$ is non-zero and it is equal to 1, i.e. if and only if $\sum_i b_i^2 = 1$. This gives our result.

□

**Example** We consider $C3v$, as usual. We can check that our 2D representation is indeed an irrep:

$$\sum_{\mu=1}^{N_c} n_\mu |\chi(C_\mu)|^2 = 1 \cdot 2^2 + 2 \cdot (-1)^2 + 3 \cdot 0 = 6 = N.$$

Since there are three conjugacy classes, there are three irreps, and we thus have two more irreps to find. As usual, we have the trivial representation:

$$g \mapsto 1, \quad \forall g \in C3v.$$

We can now use Burnside's lemma to identify the dimension of the remaining irrep:

$$\sum_i d_i^2 = N \implies 1^2 + 2^2 + d_3^2 = 6 \iff d_3 = 1.$$

Since it has dimension 1, the trace of an element is equal to the element. We can thus use the petit orthogonality theorem. The 2D irrep and trivial irrep have character, respectively:

$$\vec{\boldsymbol{v}} = (2, -1, -1, 0, 0, 0)^T, \quad \vec{\boldsymbol{u}} = (1, 1, 1, 1, 1, 1)^T.$$

The unknown irrep moreover has character:

$$\vec{\boldsymbol{w}} = (\chi_e, \chi_{c^-}, \chi_{c^+}, \chi_\sigma, \chi_{\sigma'}, \chi_{\sigma''}) = (\chi_e, \chi_c, \chi_c, \chi_\sigma, \chi_\sigma, \chi_\sigma)^T.$$

The petit orthogonality theorem tells us that:

$$\begin{cases} \vec{\boldsymbol{v}} \bullet \vec{\boldsymbol{w}} = 0 \iff 2\chi_e - 2\chi_c = 0, \\ \vec{\boldsymbol{u}} \bullet \vec{\boldsymbol{w}} = 0 \iff \chi_e + 2\chi_c + 3\chi_\sigma = 0. \end{cases}$$

However, this yields $\chi_e = \chi_c = -\chi_\sigma$. We know $R(e) = 1$, so:

$$R(e) = R(c_+) = R(c_-) = 1, \quad R(\sigma) = R(\sigma') = R(\sigma'') = -1.$$

We have used all the theory we built (mainly orthogonality relations) to successfully find a missing irrep.

# Chapter 10

# Lie algebras and angular momentum

## 10.1 Rotations

**Motivational example: 2D**

To study rotations, it suffices to study infinitesimal rotations. A rotation through an infinitesimal angle is almost no rotation (i.e. very close to identity), so we can write it as a perturbation from identity:

$$R(\theta) = I + A(\theta).$$

Now, we know that rotations are orthogonal matrices, i.e. they satisfy $R^T R = I$. Considering infinitesimal rotations, we can substitute in those two expressions, considering only the first order:

$$I = R^T R = \big(I + A^T\big)(I + A) = I + A^T + A + A^T A \approx I + A^T + A \iff A^T = -A.$$

This shows that $A$ is anti-symmetric. For instance, in two dimensions:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\theta = J\theta.$$

Note that here we made a choice on $J$: we could have chosen $J' = \frac{1}{2}J$, giving a similar result but with a dependence on $2\theta$ instead of $\theta$. This would be a correct choice as well, but choosing $J$ that way is simply more natural.
In other words, still for two dimensions, $R(\theta) = I + J\theta + O\big(\theta^2\big)$. This matrix $J$, the thing that determines the form of an infinitesimal rotation, is called the **generator** of rotations (i.e. of the group of rotations).
We recall that $e^x = \lim_{n\to\infty}\big(1 + \frac{x}{n}\big)^n$. We can use this to write an arbitrary rotation using a product of many small rotations:

$$R(\theta) = \lim_{n\to\infty} R\left(\frac{\theta}{n}\right)^n = \lim_{n\to\infty}\left(I + \frac{\theta J}{n} + O\left(\frac{\theta^2}{n^2}\right)\right)^n = e^{\theta J}.$$

We again see intuitively that $J$ generates the 2D rotations. As a small sanity check, we notice that, since $J^2 = I$ but skipping the details:

$$\begin{aligned} e^{\theta J} &= \sum_{n=0}^{\infty} \frac{\theta^n J^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{\theta^{2n} J^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{\theta^{2n+1} J^{2n+1}}{(2n+1)!} \\ &= \cos(\theta) I + \sin(\theta) J \\ &= \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}. \end{aligned}$$

This does give back our familiar representation of $SO(2)$, the group of 2D rotations.

*3D case*

We may have wished to consider the 3D case as well. It is possible to convince ourselves that the following is a basis for anti-symmetric matrices $A$:

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In other words, we can write any matrix $A$ so that $A^T = -A$ as

$$A = \theta_x J_x + \theta_y J_y + \theta_z J_z$$

Note again that the choice of the $\{J_x, J_y, J_z\}$ basis is arbitrary. We chose this one because we know it will give us back our usual 3D rotation matrices. By a similar reasoning to the 2D case, this yields that an arbitrary 3D rotation can be written as:

$$R(\theta) = e^{\theta_x J_x + \theta_y J_y + \theta_z J_z}.$$

*General case*

More generally, we can always write a rotation as:

$$R(\theta) = e^{\sum_i \theta_i J_i},$$

where $\sum_i \theta_i J_i$ is an arbitrary anti-symmetric matrix.
We moreover notice that this looks a lot like the evolution operators $e^{iH}$, but without a $i$. So, in quantum physics, we tend to redefine our $J_k$ with an $i$, $\widetilde{J}_k = -iJ_k$, so that, instead:

$$R(\theta) = e^{i\sum_i \theta_k \widetilde{J}_k}.$$

**Motivational example: Structure constants**

An important property of rotations is the fact that they do not commute.
Let $R \approx I + A$ and $R' \approx I + B$ be two infinitesimal rotations. We know that:

$$I = RR^T \approx (I + A)(I + A)^T = (I + A)(I - A).$$

So, $R^{-1} \approx I - A$. This yields that:

$$RR'R^{-1} \approx (I + A)R'(I - A) = R' + AR' - R'A = I + B + AB - BA = R' + [A, B].$$

In other words:

$$RR' \approx R'R + [A, B]R \iff [R, R'] \approx [A, B]R.$$

This shows that the amount $R$ and $R'$ do not commute is captured by the commutator of the generators. Now, for generic rotations, $A$ and $B$ will be linear combinations of generators:

$$A = i\sum_j \theta_j \widetilde{J}_j, \quad B = i\sum_j \theta'_j \widetilde{J}_j.$$

We thus notice that:

$$[A, B] = i^2 \sum_{i,j} \theta_i \theta'_j \left[\widetilde{J}_i, \widetilde{J}_j\right].$$

This means that the commutator of the basis of the generators $\left[\widetilde{J}_i, \widetilde{J}_j\right]$ fully captures the non-commutation property of the group. We call this the structure constants of the group, and we will come back to it.

**Recipe to find and characterise the generator of a group**

Let us generalise our two motivational examples.
Let $G$ be some continuous group. To find and characterise the generators of $G$, we do:

1. Expand the group elements (such as a rotation) around the identity by considering the limit where continuous parameters go to zero (such as a small angle), $g \approx I + A$.
2. Find a basis in which any arbitrary $A$ for that group can be written, i.e. a set of $T_a$ such that $A = i \sum_a \theta_a T_a$.
3. For any two group elements near identity, $g_1 = I + A$ and $g_2 = I + B$, $g_1 g_2 g_1^{-1} = g_2 + [A, B]$. In particular, $[A, B]$ captures the essence of the group near identity.
4. $A$ and $B$ can be written as a linear combinations of the generators, and we know $g_1 g_2 g_1^{-1}$ is close to identity, so the commutation relation between generators give another generator. In particular, $[T_a, T_b] = i \sum_c f_{abc} T_c$. We sometimes use Einstein's notation to write $[T_a, T_b] = i f_{abc} J_c$.

This means in some sense that the commutation relation between generators capture the essence of the generator of the group.

## 10.2 Definitions

**Definition: Lie algebra**

Let $G$ be a lie group (i.e. some form of continuous group). A **lie algebra** $\mathfrak{g}$ is a linear space spanned by a linear combination $\sum_j \theta_j J_j$ of the generators associated with $G$, together with a lie bracket $[.,.]$.

In this class, the lie bracket will always be the commutator $[A, B] = AB - BA$; but this is an algebraic structure that is more general.

*Remark 1* As Lie groups are differentiable, it is always possible to write an element $g \in G$ as the exponential of an element $J \in \mathfrak{g}$ in the corresponding lie algebra, i.e.

$$\mathfrak{g} = \left\{ J \;\middle|\; e^{iJ} \in G \right\}.$$

*Remark 2* Since they are related by exponentiation, we sometimes write:

$$G = e^{i\mathfrak{g}}.$$

*Remark 3* It is typically a lot easier to analyse lie algebra than lie groups, since lie algebras is a linear spaces.

For instance, $\mathfrak{so}(n)$ is the vector space of the anti-symmetric matrices. Vector spaces behave very well, and it is for example very easy to represent an arbitrary anti-symmetric matrix. On the other hand, $SO(n)$ is a lot harder to manipulate.

**Definition: Structure constants**

Let $\mathfrak{g}$ be a lie algebra, and $J_k$ be generators (i.e. a basis of $\mathfrak{g}$).

We define the **structure constants** of $\mathfrak{g}$, noted $f_{abc}$, as the commutation relations of the generators:

$$[J_a, J_b] = i \sum_c f_{abc} J_c.$$

*Remark* Lie algebras are in one-to-one correspondance with structure constants. Therefore, if we find that two lie algebras have the same structure constants, then they are the same lie algebra (up to isomorphisms).

**Lie algebra representation**

Representing an algebra means that we can find a set of matrices such that the defining commutation relations are satisfied. For instance, we saw before that

$$J_x = -i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_y = -i \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_z = -i \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is moreover possible to show $[J_i, J_j] = i \cdot \varepsilon_{ijk} J_k$, where $\varepsilon_{ijk}$ is the Levi-Civita constant. As physicists, we started from the representations to get the structure constants. A more mathematical approach is to start from the structure constants

$f_{ijk} = \varepsilon_{ijk}$, and get some representation from it (such as the $J_x$, $J_y$, $J_z$ we found, but others are possible).

*Remark* Like for groups, it is easier as physicists to think of representation of lie algebras instead of the lie algebras directly. This is exactly what we did for rotations for instance; we saw the $J$ matrices, instead of their underlying group elements.

**Link between representations** Let $G$ be a Lie group associated to a Lie algebra $\mathfrak{g}$, $R$ be a representation of $G$. Then, we can find a representation of $\mathfrak{g}$, given by:

$$r(X) = \frac{d}{d\theta} R\big(e^{\theta X}\big)\Big|_{\theta=0}, \quad \forall X \in \mathfrak{g},$$

This is named the representation of $\mathfrak{g}$ induced by $R$.

*Remark* The converse is not true in general: exponentiating a representation of $\mathfrak{g}$ does not always give a representation of $G$. Now, this is true if $G$ is a simply connected group; although we will not define what this means in this class.

*Example* Consider the following element of the usual representation of $SO(3)$:

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Differentiating it, we do find an element of the representation of $\mathfrak{so}(3)$ (the vector space of anti-symmetric matrices) we presented:

$$\begin{aligned} \frac{d}{d\theta} R_x(\theta)\Big|_{\theta=0} &= \frac{d}{d\theta} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}\Big|_{\theta=0} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin(\theta) & -\cos(\theta) \\ 0 & \cos(\theta) & -\sin(\theta) \end{pmatrix}\Big|_{\theta=0} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ &= iJ_x. \end{aligned}$$

Monday 16th December 2024 — **Lecture 13 : Ah yes, $\mathfrak{so}(3) = SO(3)$**

## 10.3 Addition of angular momentum

**Personal remark** This lecture was not given by the Professor, we had to read the lecture notes instead. This is my best attempt to explain how I understood this subject, notably compiling some questions I asked to TAs. There are probably some inaccuracies, do not hesitate to point them out to me.

*Acknowledgement* A big thanks to Nathan Brunet, Jonas Daverio, Gabriel Pescia and Reyhaneh Saem for their answers to my questions and their patience.

**Definition: Angular momentum** Let $\mathcal{H}$ be a quantum system, with position operator $\hat{r} = (\hat{x}, \hat{y}, \hat{z})$ and momentum operator $\hat{p} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$.

We define its angular momentum operators as:

$$\begin{pmatrix}\hat{J}_x\\ \hat{J}_y\\ \hat{J}_z\end{pmatrix} = \begin{pmatrix}\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\\ \hat{z}\hat{p}_x - \hat{x}\hat{p}_z\\ \hat{x}\hat{p}_y - \hat{y}\hat{p}_x\end{pmatrix}.$$

*Intuition*

Recall that, in classical mechanics, we defined angular momentum as:

$$\vec{\boldsymbol{L}} = \vec{\boldsymbol{x}} \times \vec{\boldsymbol{p}} = \begin{pmatrix}yp_z - zp_y\\ zp_x - xp_z\\ xp_y - yp_x\end{pmatrix}.$$

We thus simply promote everything to an operator for quantum systems.

**Theorem**

Let $\mathcal{H}$ be a quantum system, of angular momentum operators $J_x, J_y, J_z$.
Then, $iJ_x, iJ_y, iJ_z$ are a basis for a representation of $\mathfrak{so}(3)$.

*Proof*

It is possible to prove, like we did in Quantum physics I, that:

$$[J_i, J_x] = i\hbar \sum_k \varepsilon_{ijk} J_k.$$

Leaving $\hbar = 1$ for simplicity, this shows that their structure constants are the same as the lie algebra corresponding to the orthogonal group $SO(3)$. Since lie algebras can be identified by their structure constants, this shows our result.

□

**Theorem**

Let $j \in \left\{0, \frac{1}{2}, 1, \dots\right\}$ be arbitrary.
We can construct a representation of $\mathfrak{so}(3)$, which we note $j$, given by some eigenvectors $|jm\rangle$ of $J_z$.
To do so, we define the following ladder operators:

$$J_- = J_x - iJ_y, \quad J_+ = J_x + iJ_y.$$

They are such that:

- $J_z|j,m\rangle = m$.
- $[J_z, J_\pm] = \pm J_\pm$.
- $[J_+, J_-] = 2J_z$.
- $J_+|j,m\rangle = \sqrt{j(j+1) - m(m+1)}|j, m+1\rangle$.
- $J_-|j,m\rangle = \sqrt{j(j+1) - m(m-1)}|j, m-1\rangle$.

All these properties allow to compute $\langle j, m'|J_z|j,m\rangle = m\delta_{m,m'}$, and the coordinates of $J_x = \frac{1}{2}(J_+ + J_-)$ and $J_y = \frac{1}{2i}(J_+ - J_y)$. These $J_x, J_y, J_z$ form a basis for a representation of $\mathfrak{so}(3)$.

*Example*

For instance, the $j = 0$ representation is the trivial irrep:

$$J_x = J_y = J_z = 0.$$

Moreover, the $j = \frac{1}{2}$ irrep is, in the $\left|\frac{1}{2}, \frac{1}{2}\right\rangle, \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ basis:

$$J_x = \frac{1}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad J_y = \frac{1}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad J_z = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

We recognise the Pauli matrices. As a last example, the $j = 1$ irrep is, in the $|1,1\rangle, |1,0\rangle, |1,-1\rangle$ basis:

$$J_x = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}, \quad J_y = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & -i & 0\\ i & 0 & -i\\ 0 & i & 0\end{pmatrix},$$

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

*Proof*

We make a constructive proof.

We want to construct a representation of $\mathfrak{so}(3)$, by diagonalising $J_z$. So, let $|m\rangle$ be an eigenstate associated to eigenvalue $m$:

$$J_z|m\rangle = m|m\rangle.$$

Since $[J_x, J_z] \neq 0$ and $[J_y, J_z] \neq 0$, this shows that we cannot block diagonalise them at the same time, and hence the $|m\rangle$ are not eigenstates of $J_x$ and $J_y$. Let us instead introduce some helper operators:

$$J_+ = J_x + iJ_y, \quad J_- = J_x - iJ_y.$$

Notice that they are such that, using the commutation properties of the angular momentum operators:

$$[J_z, J_\pm] = [J_z, J_x \pm iJ_y] = iJ_y \pm i \cdot (-iJ_x) = iJ_y \pm J_x = \pm J_\pm,$$

$$[J_+, J_-] = [J_x + iJ_y, J_x - iJ_y] = 2i[J_y, J_x] = 2i(-iJ_z) = 2J_z.$$

Moreover, we notice that $J_\pm^\dagger = J_\mp$ since $J_x$ and $J_y$ are Hermitian. We can thus evaluate:

$$\begin{aligned} J_z J_\pm|m\rangle &= J_\pm J_z|m\rangle + [J_z, J_\pm]|m\rangle \\ &= mJ_\pm|m\rangle \pm J_\pm|m\rangle \\ &= (m \pm 1)J_\pm|m\rangle. \end{aligned}$$

In other words, $J_+|m\rangle$ is an eigenket of $J_z$, with eigenvalue $m+1$; hence $J_+|m\rangle = c_{m+1}|m+1\rangle$ for some constant $c_{m+1}$. Similarly, $J_-|m\rangle = b_{m-1}|m-1\rangle$. This justifies the notation $J_+$ and $J_-$: they allow to move up or down on the ladder of eigenstates $\ldots, |m-2\rangle, |m-1\rangle, |m\rangle, |m+1\rangle, \ldots$.

Moreover, we can link $c_m$ and $b_m$ by using the fact that $J_+ = J_-^\dagger$:

$$b_m = \langle m|J_-|m+1\rangle = \langle m+1|J_-^\dagger|m\rangle^* = \langle m+1|J_+|m\rangle^* = c_{m+1}^*.$$

This shows that $J_+|m\rangle = c_{m+1}|m+1\rangle$ and $J_-|m\rangle = c_m^*|m-1\rangle$. We want to evaluate those values $c_m$ explicitly, by finding a recurrence relation. We notice that:

$$J_+J_-|m\rangle = c_m^* J_+|m-1\rangle = |c_m|^2|m\rangle,$$

$$J_-J_+|m\rangle = c_{m+1} J_-|m+1\rangle = |c_{m+1}||m\rangle.$$

This allows us to find a recurrence relation for our coefficients, using the fact $[J_+, J_-] = 2J_z$ as found above:

$$\begin{aligned} |c_m|^2 - |c_{m+1}|^2 &= \langle m|(J_+J_- - J_-J_+)|m\rangle \\ &= \langle m|[J_+, J_-]|m\rangle \\ &= \langle m|2J_z|m\rangle \\ &= 2m. \end{aligned}$$

We thus only need an initial condition in order to be able to solve our recurrence, i.e. we need to the value of some $c_{m_0}$. To do so, we notice that we aim to make a finite-dimensional representation of $\mathfrak{so}(3)$, so we cannot just have infinitely many such eigenvectors in our ladder. This means that there is some lower bound $\ell$ and

upper bound $j$ so that all our states are $|\ell\rangle, |\ell+1\rangle, \dots, |j-1\rangle, |j\rangle$. In particular, this asks for:

$$J_+|j\rangle = 0 \implies c_{j+1} = 0, \quad J_-|\ell\rangle = 0 \implies c_\ell = 0.$$

This allows us to find the initial condition for our recurrence relation:

$$\begin{aligned} 0 &= \langle j|J_-J_+|j\rangle \\ &= \langle j|J_+J_-|j\rangle + \langle j|[J_+, J_-]|j\rangle \\ &= |c_j|^2 + 2\langle j|J_z|j\rangle \\ &= |c_j|^2 + 2j. \end{aligned}$$

Therefore, $|c_j|^2 = 2j$. However this, together with the recurrence $|c_m|^2 = |c_{m+1}|^2 + 2m$ we found earlier, can be solved to find, for all $s \geq 0$:

$$\begin{aligned} |c_{j-s}|^2 &= |c_{j-s+1}|^2 + 2(j-s) \\ &= |c_{j-s+2}|^2 + 2(j-s+1) + 2(j-s) \\ &= \dots \\ &= |c_j|^2 + 2\sum_{k=j-s}^{j} k. \end{aligned}$$

Being careful on the fact that the sum may range between negative and positive numbers (we may have $j - s < 0$), we can finally simplify this to:

$$|c_{j-s}|^2 = (s+1)(2j-s).$$

In particular, we notice that this goes to zero if and only if $s = 2j$, i.e. $c_{-j} = 0$. This shows that $J_-|-j\rangle = c_{-j} = 0$. This means that the ladder terminates at this point, and hence that we had $\ell = -j$. Moreover, we know that we can apply $2j$ times the operator $J_-$ on $|j\rangle$ to get to $|-j\rangle$, so $2j$ must be an integer. This thus forces that $j$ is a half integer, $j \in \left\{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\right\}$. Since this is a fundamental parameter in our construction, we may write $|m\rangle \equiv |j, m\rangle$.
To finish with, we can express $|c_m|$:

$$\begin{aligned} &|c_{j-s}|^2 = (s+1)(2j-s) \\ \implies &|c_m|^2 = (j-m+1)(2j-j+m) = (j-m+1)(j+m) \\ \implies &|c_m| = \sqrt{(j-m+1)(j+m)}. \end{aligned}$$

We have a choice on the phase of the $c_m$, which makes sense since global phases do not matter in quantum physics. By simplicity, we can thus simply write:

$$c_m = \sqrt{(j-m+1)(j+m)} = \sqrt{j(j+1) - m(m-1)}.$$

□

**Theorem** Let $j \in \left\{0, \frac{1}{2}, 1, \dots\right\}$ be arbitrary.
The representation $j$ of $\mathfrak{so}(3)$ is an irrep.

*Remark 1* We did not define formally the notion of irreducible representation of lie algebras. Intuitively, this is just like for groups, a representation is reducible if all its elements can be block diagonalised. In particular, it suffices to block diagonalise a basis, $J_x, J_y, J_z$ in our case.

*Remark 2* Note that before in the course we were considering finite groups, and hence there was a finite number of irreps. Now that we are considering a continuous group, it is not surprising that we have infinitely many irreps.

**Clebsch-Gordan algorithm**

We consider the following algorithm, which allows to decompose the representation $j \otimes j'$ as a direct sum of irreps.

The idea is to decompose the $|J, M\rangle$ basis (associated to the angular momentum of the larger system $J_k = J_k^{(1)} \otimes I + I \otimes J_k^{(2)}$) into the basis $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ (associated to the angular momentum of the two sub-systems). To do so, we can use an inductive process with the following three rules:

1. Start with $|J = j_1 + j_2, M = j_1 + j_2\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle$.
2. Given $|J = j, M = m\rangle$, find $|J = j, M = m - 1\rangle$ using the fact it is normalised and:
$$\begin{aligned} &|J = j, M = m - 1\rangle \\ \propto\ & J_-|J = j, M = m\rangle \\ =\ & \Big(J_-^{(1)} \otimes I\Big)|J = j, M = m\rangle + \Big(I \otimes J_-^{(2)}\Big)|J = j, M = m\rangle. \end{aligned}$$
3. Given $|J = j + 1, M = j\rangle, \ldots, |J = j_1 + j_2, M = j\rangle$, find $|J = j, M = j\rangle$ using the fact it is orthogonal to all $|J = j + 1, M = j\rangle, \ldots, |J = j_1 + j_2, M = j\rangle$ and that it uses the same basis ket (i.e. it can be expressed in the basis $\{|j_1, m_1 = k\rangle \otimes |j_2, m_2 = j - k\rangle \mid k\}$).

We can moreover make the process faster, by using the following symmetry:

$$\begin{aligned} &(\langle j_1, -m_1| \otimes \langle j_2, -m_2|)|J = j, M = -m\rangle \\ =\ & (-1)^{j_1+j_2-j}(\langle j_1, m_1| \otimes \langle j_2, m_2|)|J = j, M = m\rangle. \end{aligned}$$

In other words, we get $|J = j, M = -m\rangle$ from $|J = j, M = m\rangle$ by replacing each $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ in its basis expansion by $(-1)^{j_1+j_2-j}|j_1, -m_1\rangle \otimes |j_2, -m_2\rangle$.

*Remark* Methodically, we may always start with $J = j_1 + j_2$ and $M = j_1 + j_2$. We can then find $|J = j_1 + j_2, M\rangle$ for each $M = j_1 + j_2, \ldots, -j_1 - j_2$. We can then decrease the value of $J$ by one, and start again; doing this until $J = j_1 - j_2$.

*Intuition* The idea is that we know $J_x^{(1)}, J_y^{(1)}, J_z^{(1)}$ form a basis for the $j_1$ irrep of $\mathfrak{so}(3)$, and similarly for $J_x^{(2)}, J_y^{(2)}, J_z^{(2)}$. We know that this implies $J_k^{(1)} \otimes J_k^{(2)}$ for $k \in \{x, y, z\}$ is also the basis for a representation of $\mathfrak{so}(3)$ (although not an irrep).

However, physically, $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ is a basis that shows the angular momentum of two subsystems. We can expect that there is some states $|J, M\rangle$ that should represent the total angular momentum of the full system. Since we want angular momentums to add, this means that the larger system has basis $J_k = J_k^{(1)} \otimes I + I \otimes J_k^{(2)}$. Indeed, for instance, we do have $J_z|j_1, m_1\rangle \otimes |j_2, m_2\rangle = m_1 + m_2$.

This construction thus allows to find how to change between bases $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ and $|J, M\rangle$. To do so, we use the three rules explained above iteratively. Let us explain the intuition behind them.

1. The first one states that the only way for the full system to reach total $z$-angular momentum $M = j_1 + j_2$ is that both sub-systems have a maximal $z$-angular momentum, $m_1 = j_1$ and $m_2 = j_2$.
2. The second rule simply uses the ladder operator on the full system, since we know how it will act on the subsystems.

3. The third rule uses orthogonality, and the fact that $|J = j, M = m\rangle$ uses the exact same basis as $|J = j', M = m\rangle$, $\{|j_1, m_1 = k\rangle \otimes |j_2, m_2 = m - k\rangle \mid k\}$. The latter fact can be understood intuitively by stating that, for the whole system to have $z$-angular momentum $M$, then the subsystems must be such that $m_1 + m_2 = M$.

Finally, we range $J \in \{|j_1 - j_2|, \dots, j_1 + j_2\}$. Indeed, the total angular momentum $J$ is maximal when the angular momentum of both subsystems go in the same direction, in which case it is $j_1 + j_2$; and it is minimal when the angular momentum of both subsystems go in the opposite direction, in which case it is $|j_1 - j_2|$. Note that this intuition relies heavily on the classical intuition, there is no real concept of "direction" for angular momentum in quantum.

**Example 1** We have:

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0.$$

*Proof*

- We start by finding all $|J, M\rangle$ for $J = 1$. We know:

$$|J = 1, M = 1\rangle = \left|j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\right\rangle \otimes \left|j_1 = \frac{1}{2}, m_1 = \frac{1}{2}\right\rangle.$$

For the sake of simplicity, we will write this equivalently as:

$$|1, 1\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle.$$

Since we know $J_-|1, 1\rangle = \sqrt{2}|1, 0\rangle$ and $J_-^{(1)}\left|\frac{1}{2}, \frac{1}{2}\right\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$:

$$\begin{aligned}
|1, 0\rangle &= \frac{1}{\sqrt{2}} J_-|1, 1\rangle \\
&= \frac{\left(J_-^{(1)} \otimes I\right)|1, 1\rangle + \left(I \otimes J_-^{(2)}\right)|1, 1\rangle}{\sqrt{2}} \\
&= \frac{J_-^{(1)}\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes J_-^{(2)}\left|\frac{1}{2}, \frac{1}{2}\right\rangle}{\sqrt{2}} \\
&= \frac{\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle}{\sqrt{2}}.
\end{aligned}$$

Finding $|1, -1\rangle$ is again very similar, since $J_-|1, 0\rangle = \sqrt{2}|1, -1\rangle$ and $J_-^{(1)}\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = 0$:

$$\begin{aligned}
&|1, -1\rangle \\
&= \frac{1}{\sqrt{2}} J_-|1, 0\rangle \\
&= J_- \frac{\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle}{2} \\
&= \frac{0 \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes 0}{2} \\
&= \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle.
\end{aligned}$$

- We now look for all $|J, M\rangle$ for $J = 0$. We know $|J = 0, M = 0\rangle$ is orthogonal to $|J = 1, M = 0\rangle$. Therefore, it must be:

$$|J = 0, M = 0\rangle = \frac{\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle - \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle}{\sqrt{2}}.$$

As a quick sanity check, one can verify that all those vectors are indeed normalised.

**Example 2**

We have:

$$1 \otimes 1 = 2 \oplus 1 \oplus 0.$$

*Proof*

- We start with $|J, M\rangle$ for $J = j_1 + j_2 = 2$. We know:

$$|J = 2, M = 2\rangle = |j_1 = 1, m_1 = 1\rangle \otimes |j_2 = 1, m_2 = 1\rangle.$$

Let us simplify the notation even more than before, since we know that $j_1 = 1$ and $j_2 = 1$ will always hold. We thus write this in the form:

$$|2, 2\rangle = |1\rangle \otimes |1\rangle.$$

*(We may even write this as $|2, 2\rangle = |1, 1\rangle$ if we really feel like we want to use horrendous notations as good physicists.)* Then, doing something completely similar to the previous example:

$$|2, 1\rangle = \frac{|1\rangle \otimes |0\rangle + |0\rangle \otimes |1\rangle}{\sqrt{2}}.$$

We can apply this again, using the fact $J_-|J = 2, M = 1\rangle = \sqrt{6}|J = 2, M = 0\rangle$:

$$\begin{aligned} |2, 0\rangle &= \frac{1}{\sqrt{6}} J_- |2, 1\rangle \\ &= \frac{\sqrt{2}|-1\rangle \otimes |1\rangle + \sqrt{2}|0\rangle \otimes |0\rangle + \sqrt{2}|-1\rangle \otimes |1\rangle + \sqrt{2}|0\rangle \otimes |0\rangle}{\sqrt{6} \cdot \sqrt{2}} \\ &= \frac{|-1\rangle \otimes |1\rangle + 2|0\rangle \otimes |0\rangle + |-1\rangle \otimes |1\rangle}{\sqrt{6}}. \end{aligned}$$

We can now exploit symmetry, to find that:

$$|2, -1\rangle = \frac{|-1\rangle \otimes |0\rangle + |0\rangle \otimes |-1\rangle}{\sqrt{2}},$$

$$|2, -2\rangle = |-1\rangle \otimes |-1\rangle.$$

- We now consider $J = 1$. We know that $|1, 1\rangle$ must be orthogonal with $|2, 1\rangle$, so we can take:

$$|1, 1\rangle = \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle}{\sqrt{2}}.$$

We then find:

$$|1, 0\rangle = \frac{|-1\rangle \otimes |1\rangle - |1\rangle \otimes |-1\rangle}{\sqrt{2}},$$

$$|1, -1\rangle = \frac{|-1\rangle \otimes |0\rangle - |0\rangle \otimes |-1\rangle}{\sqrt{2}}.$$

- We finally consider $J = 0$. We know $|0, 0\rangle$ must be orthogonal to both $|2, 0\rangle$ and $|1, 0\rangle$. It is possible to guess the answer, or solve a linear system, to find that:

$$|0, 0\rangle = \frac{|-1\rangle \otimes |1\rangle - |0\rangle \otimes |0\rangle + |1\rangle \otimes |-1\rangle}{\sqrt{3}}.$$

**$SO(3)$ representation**

Note that, all we have done so far only applies to representations of $\mathfrak{so}(3)$. Indeed, $SO(3)$ is not simply connected, so we cannot just get a representation of $SO(3)$ by exponentiating the representations of $\mathfrak{so}(3)$.

However, physically, this actually works. Indeed, exponentiating a representation of $SO(3)$, we get something called a "projective representation", where we allow global phases i.e. $R(g_1) \cdot R(g_2) = e^{i\varphi} R(g_1 * g_2)$. Since global phases do not matter in quantum physics, everything we did here actually also allows us to find representations $SO(3)$.

In practice, since there is this strong link between $\mathfrak{so}(3)$ and $SO(3)$, people may use the symbol $SO(3)$ when speaking of either. This is why, in exercise series, $|j, m\rangle$ is called a representation of $SO(3)$ (when, formally, it is really a representation of $\mathfrak{so}(3)$).

*Personal remark*

This actually took me so long to understand. A big thanks to Jonas Daverio for explaining me this.