

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

**22.51**  
**QUANTUM THEORY**  
of  
**RADIATION INTERACTIONS**

Fall 2012

**Prof. Paola Cappellaro**



# Contents

<b>1</b>	<b>Introduction to the course</b>	<b>7</b>
1.1	Why study Quantum Mechanics?	7
1.2	Striking Characteristics of QM	7
<b>2</b>	<b>Mathematical Formalism of Quantum Mechanics</b>	<b>9</b>
2.1	Linear vectors and Hilbert space	9
2.2	Operators	10
2.2.1	Hermitian operators	11
2.2.2	Operators and their properties	13
2.2.3	Functions of operators	14
<b>3</b>	<b>Axioms of Quantum Mechanics</b>	<b>15</b>
3.1	Introduction	15
3.2	The axioms of quantum mechanics	16
3.2.1	Observables and State Space	16
3.2.2	Quantum measurement	16
3.2.3	Law of motion	17
3.3	Strong measurements	18
3.3.1	Expectation values	18
3.3.2	Uncertainty relationships	18
3.3.3	Repeated measurements and Quantum Zeno Effect	19
<b>4</b>	<b>Two-level systems</b>	<b>21</b>
4.1	Generalities	21
4.2	Rotations and angular momentum	22
4.2.1	Classical rotations	22
4.2.2	QM angular momentum as generator of rotations	23
4.2.3	Example of Two-Level System: Neutron Interferometry	25
4.2.4	Spinor behavior	26
4.2.5	The SU(2) and SO(3) groups	26
<b>5</b>	<b>Time evolution</b>	<b>27</b>
5.1	The Schrödinger and Heisenberg pictures	27
5.2	Interaction Picture	28
5.2.1	Dyson Time-ordering operator	29
5.2.2	Some useful approximate formulas	30
5.3	Spin- $\frac{1}{2}$ precession	32
5.4	Examples: Resonance of a Two-Level System	33
5.4.1	Dressed states and AC Stark shift	34
5.5	The wave-function	35
5.5.1	Position representation	35
5.5.2	Momentum representation	35
5.5.3	Schrödinger equation for the wavefunction	37
5.6	Feynman's path-integral	37
<b>6</b>	<b>Composites systems and Entanglement</b>	<b>39</b>
6.1	Tensor product of Hilbert spaces	39
6.1.1	Product Operator Basis	40
6.2	Quantum Information Processing	40
6.3	Operators on two Qubits	41
6.4	No cloning Theorem	42
6.5	Entanglement and EPR paradox	44
6.5.1	Bell Inequalities	44
6.6	Teleportation (Bennet, Peres, Brassard)	47
6.7	Deutsch-Jozsa algorithm	48

<b>7</b>	<b>Mixed states</b>	<b>51</b>
7.1	Mixed States . . . . .	51
7.2	Dynamics of mixed states and operators . . . . .	52
7.2.1	Heisenberg picture . . . . .	53
7.2.2	Interaction picture . . . . .	53
7.3	Partial Trace . . . . .	54
7.3.1	Examples . . . . .	55
7.4	Entanglement measurement . . . . .	55
7.5	Mixed States and interpretation of the density matrix . . . . .	56
7.5.1	Classical Macro-states . . . . .	57
7.5.2	Quantum Macro-states . . . . .	58
7.5.3	Example: Spin- $\frac{1}{2}$ . . . . .	60
<b>8</b>	<b>Open Quantum Systems</b>	<b>61</b>
8.1	Combined evolution of system and bath . . . . .	61
8.2	Superoperators . . . . .	63
8.3	The Kraus Representation Theorem . . . . .	65
8.3.1	Amplitude-damping . . . . .	65
8.3.2	Phase-damping . . . . .	66
8.3.3	Depolarizing process . . . . .	67
8.4	The Master Equation . . . . .	69
8.4.1	Markov approximation . . . . .	69
8.4.2	Lindblad equation . . . . .	69
8.4.3	Redfield-Born theory of relaxation . . . . .	72
8.5	Other description of open quantum system dynamics . . . . .	74
8.5.1	Stochastic Liouville equation and cumulants . . . . .	74
8.5.2	Stochastic Wavefunctions . . . . .	76
<b>9</b>	<b>Harmonic Oscillator</b>	<b>79</b>
9.1	Harmonic Oscillator . . . . .	79
9.1.1	Classical harmonic oscillator and h.o. model . . . . .	79
9.1.2	Oscillator Hamiltonian: Position and momentum operators . . . . .	80
9.1.3	Position representation . . . . .	81
9.1.4	Heisenberg picture . . . . .	83
9.1.5	Schrödinger picture . . . . .	83
9.2	Uncertainty relationships . . . . .	83
9.3	Coherent States . . . . .	84
9.3.1	Expansion in terms of number states . . . . .	85
9.3.2	Non-Orthogonality . . . . .	86
9.3.3	Uncertainty relationships . . . . .	86
9.3.4	X-representation . . . . .	87
9.4	Phonons . . . . .	88
9.4.1	Harmonic oscillator model for a crystal . . . . .	88
9.4.2	Phonons as normal modes of the lattice vibration . . . . .	89
9.4.3	Thermal energy density and Specific Heat . . . . .	90
<b>10</b>	<b>The electromagnetic field</b>	<b>93</b>
10.1	Classical theory of the e.m. field . . . . .	93
10.2	Quantization of the e.m. field . . . . .	95
10.2.1	Zero-Point Energy and the Casimir Force . . . . .	96
10.3	Quantization of the e.m. field in the Coulomb gauge . . . . .	99
10.4	States of the e.m. field . . . . .	99
10.4.1	Photon number eigenstates . . . . .	99
10.4.2	Coherent states . . . . .	100
10.4.3	Measurement Statistics . . . . .	101
10.5	Atomic interactions with the quantized field . . . . .	102

<b>11 Perturbation Theory</b>	<b>107</b>
11.1 Time-independent perturbation theory . . . . .	107
11.1.1 Non-degenerate case . . . . .	107
11.1.2 Degenerate case . . . . .	111
11.1.3 The Stark effect . . . . .	115
11.2 Time-dependent perturbation theory . . . . .	119
11.2.1 Review of interaction picture . . . . .	119
11.2.2 Dyson series . . . . .	119
11.2.3 Fermi's Golden Rule . . . . .	120
<b>12 Interaction of Radiation with Matter</b>	<b>123</b>
12.1 Scattering Theory . . . . .	123
12.1.1 Cross Section . . . . .	125
12.1.2 Thermal Neutron Scattering . . . . .	127
12.2 Emission and Absorption . . . . .	133
12.2.1 Emission . . . . .	133
12.2.2 Absorption . . . . .	135
12.2.3 Blackbody Radiation . . . . .	135
12.3 Wigner-Weisskopf Theory . . . . .	136
12.3.1 Interaction of an atom with a single mode e.m. field . . . . .	136
12.3.2 Interaction with many modes of the e.m. field . . . . .	137
12.4 Scattering of photons by atoms . . . . .	139
12.4.1 Thomson Scattering by Free Electrons . . . . .	142
12.4.2 Rayleigh Scattering of X-rays . . . . .	145
12.4.3 Visible Light Scattering . . . . .	146
12.4.4 Photoelectric Effect . . . . .	147



# 1. Introduction to the course

## 1.1 Why study Quantum Mechanics?

### 1.2 Striking Characteristics of QM

## 1.1 Why study Quantum Mechanics?

Quantum mechanics (QM) is a fundamental and general theory that applies on a very wide range of scale, from subatomic systems to astrophysical objects.

It is nowadays also a widely applied subject, with real-life applications (from transistors to lasers) and an active area of research. For example, it is at the basis of research in nanotechnology, in materials science, in spintronics (where the electron spin replaces the charge as the fundamental unit for storing information in a computer), etc.

Also, there has been a resurgent interest in the fundamental theory of QM, due to the interest in quantum information as well as the availability of control on single quantum systems<sup>1</sup>. Although you might not be directly interested in quantum information per se, this discipline has shaped a new view of QM that is having impact also on practical applications. On the pedagogical side, it has pushed for a new way of presenting QM in college courses, that I will try to partially follow here, with more emphasis on discrete systems and applications instead of an historical presentation of QM theory. (You probably already have had that kind of historical introduction, both in undergrad and in 22.101).

## 1.2 Striking Characteristics of QM

QM is known for being *weird*, counterintuitive and difficult to understand<sup>2</sup>. The vast majority of physicists do not worry about the puzzling aspects of quantum mechanics, but simply use it as a tool without asking questions of principle. Nevertheless, the theoretical and, especially, experimental progress made over the past twenty years have led to a better grasp of certain aspects of the behavior of quantum objects.

In this class we will explore these counterintuitive phenomena, in particular we will emphasize three characteristics of QM:

### – Discreteness

This is the characteristic that gave the discipline its name of *quantum* mechanics<sup>3</sup>. While classical physics and the word around us seems to be continuous, in QM some quantities can only take a discrete set of values. Examples are the discrete energy levels of atoms or the amount of energy emitted in black body radiation (these are *historical* examples). In the first part of the course we will focus on discrete systems, in particular on a system that can assume only two states.

---

<sup>1</sup> This is exemplified by the work of the two Physics Nobel laureates in 2012, Dave Wineland and Serge Haroche.

<sup>2</sup> *For those who are not shocked when they first come across quantum theory cannot possibly have understood it.* Niels Bohr, quoted in Heisenberg, Werner (1971). *Physics and Beyond*. New York: Harper and Row. pp. 206. *I think I can safely say that nobody understands quantum mechanics.* Richard Feynman, in *The Character of Physical Law* (1965)

<sup>3</sup> Etymology: Latin, neuter of *quantus* how much. Plural *quanta*. 1 a : quantity, amount. b : a certain or an allotted amount : portion (from Merriam-Webster dictionary).

– **Interference**

Interference and diffraction are characteristics of waves, in particular light. In QM it was found that interference also applied to matter and it is actually a general phenomenon. E.g. diffraction observed with large objects such as fullerenes (C<sub>60</sub>). We will see how interference is linked to the possibility of finding a system in a superposition state and further explore even *weirder* phenomena such as entanglement.

– **Phase coherence**

The ability to observe interferences is linked to ability for a system to maintain a phase coherence among the different parts in a superposition state. Conversely, the loss of this phase coherence is linked to the disappearance of the QM properties of a system and the observation of classical physics behavior. This occurs e.g. when the system interacts with an environment and a good part of this course will focus on the study of these so-called *open* quantum systems.

## References

- A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki and H. Ezawa, *Demonstration of single-electron buildup of an interference pattern* [American Journal of Physics](#) **57**, 2 117-120 (1989)  
**Online Video**
- M. S. Chapman, T. D. Hammond, A. Lenef, J. Schmiedmayer, R. A. Rubenstein, E. Smith, and D. E. Pritchard, *Photon Scattering from Atoms in an Atom Interferometer: Coherence Lost and Regained* [Phys. Rev. Lett.](#) **75**, 3783 - 3787 (1995)
- M. Arndt, K. Hornberger, A. Zeilinger *Probing the limits of the quantum world* [Physics World](#) 35-40 (March 2005)

## 2. Mathematical Formalism of Quantum Mechanics

### 2.1 Linear vectors and Hilbert space

### 2.2 Operators

#### 2.2.1 Hermitian operators

#### 2.2.2 Operators and their properties

#### 2.2.3 Functions of operators

Quantum mechanics is a linear theory, and so it is natural that vector spaces play an important role in it. A physical state is represented mathematically by a vector in a Hilbert space (that is, vector spaces on which a positive-definite scalar product is defined); this is called the space of states. Physical properties like momentum, position, energy, and so on will be represented by operators acting in the space of states. We will introduce the essential properties of Hilbert spaces, mainly in the case of finite dimension, as the mathematical theory of Hilbert spaces of infinite dimension is much more complicated than that of spaces of finite dimension

### 2.1 Linear vectors and Hilbert space

**$\mathcal{D}$ : Linear vector space** A linear vector space is a set of elements, called vectors, which is closed under addition and multiplication by scalars.

Using Dirac notation, the vectors are denoted by *kets*:  $|k\rangle$ . We can associate to each ket a vector in the dual space called *bra*:  $\langle\psi|$ .

If two vectors  $|\psi\rangle$  and  $|\varphi\rangle$  are part of a vector space, then  $|\psi\rangle + |\varphi\rangle$  also belongs to the space. If a vector  $|\psi\rangle$  is in the space, then  $\alpha|\psi\rangle$  is also in the space (where  $\alpha$  is a complex scalar).

A set of *linearly independent* vectors  $\{|\varphi_i\rangle\}$  is such that  $\sum_k c_k |\varphi_k\rangle = 0$  if and only if  $c_k = 0 \forall k$  (no trivial combination of them sums to zero).

The *dimension* of the space  $N$  is the maximum number of linearly independent vectors (which is also the smallest number of vectors that span the space).

**$\mathcal{D}$ : Basis** A maximal set of linearly independent vectors in the space is called a basis. (e.g.  $\{|\phi_k\rangle\}$ ,  $k = 1, \dots, N$ ). Any vector in the space can be written as a linear superposition of the basis vectors:

$$|\psi\rangle = \sum_k a_k |\phi_k\rangle \quad (1)$$

To any vector we can thus associate a column vector of  $N$  complex numbers  $(a_1, a_2, \dots, a_n)^T$ . Here we are going to restrict ourselves to bounded, finite dimension spaces (even if many physical spaces are not: for example energy spaces can be unbounded and position has infinite dimension).

**$\mathcal{D}$ : Hilbert space** The Hilbert space is a linear vector space over complex numbers with an *inner product*.

**$\mathcal{D}$ : Inner product** An inner product is an ordered mapping from two vectors to a complex number (for a Hilbert space a mapping from a ket and a bra to a complex number  $c = \langle\psi|\varphi\rangle$ ) with the following properties:

- positivity:  $\langle\psi|\psi\rangle \geq 0$ . The equality holds only for the zero vector  $|\psi\rangle = 0$ .
- linearity in the second function:  $\langle\psi|(c_1\varphi_1 + c_2\varphi_2)\rangle = c_1\langle\psi|\varphi_1\rangle + c_2\langle\psi|\varphi_2\rangle$ .

- anti-linearity in the first function:  $(\langle c_1\varphi_1 + \langle c_2|\varphi_2)|\psi\rangle = c_1^*\langle\varphi_1|\psi\rangle + c_2^*\langle\varphi_2|\psi\rangle$ .
- skew symmetry:  $\langle\psi|\varphi\rangle = \langle\varphi|\psi\rangle^*$

**D: Norm** The norm of a vector is  $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$ .

Since the Hilbert space is characterized by its inner product, vectors are defined up to a global phase, that is,  $|\psi\rangle = e^{i\vartheta}|\psi\rangle$ . Relative phase is instead very important:  $|\psi\rangle + e^{i\vartheta}|\phi\rangle \neq |\psi\rangle + |\phi\rangle$ .

The inner product properties allow us to define two geometric inequalities:

- Schwartz inequality:  $|\langle\psi|\varphi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\varphi|\varphi\rangle$ .
- Triangular inequality:  $\|(\psi + \varphi)\| \leq \|\varphi\| + \|\psi\|$ .

The equality holds only if the two vectors are in the same direction:  $|\psi\rangle = c|\varphi\rangle$ .

There is also an antilinear correspondence between the dual vectors ket and bra:

$$c_1|\psi_1\rangle + c_2|\psi_2\rangle \rightarrow c_1^*\langle\psi_1| + c_2^*\langle\psi_2|$$

**D: Orthonormal set** A set of vectors  $\{|\varphi_k\rangle\}$  is orthonormal if for each pair the inner product  $\langle\varphi_k|\varphi_j\rangle = \delta_{k,j}$ .

## 2.2 Operators

We can define a set of operators that acting on the vectors return vectors:

**D: Operator** An operator  $A$  on a vector space is a mapping between two vectors in that space:  $A|\psi\rangle = |\phi\rangle$ .

A *linear* operator satisfies:

$$A(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1A|\psi_1\rangle + c_2A|\psi_2\rangle$$

To characterize and parametrize  $A$  we look at its action on each vector in the space. Because of linearity, it is however enough to characterize  $A$  with its action on the  $N$  basis vectors  $\{|\phi\rangle_k\}$ . In this way we can associate a matrix representation to any operator, in the same way we associated arrays of complex numbers with the vectors. In particular, given an orthonormal basis  $\{|v\rangle_k\}$ , the matrix representation of the operator  $A$  is an  $N \times N$  square matrix  $A$  whose elements are given by  $A_{k,j} = \langle v_k|A|v_j\rangle$ .

Let us consider an orthonormal basis  $\{v_i\}$ , then as seen any vector can be written as:  $|\psi\rangle = \sum_{i=1}^N a_i|v_i\rangle$ . The action of an operator  $A$  becomes:

$$A|\psi\rangle = |\varphi\rangle \rightarrow \sum_{i=1}^N Aa_i|v_i\rangle = \sum_{i=1}^N b_i|v_i\rangle$$

To extract one of the coefficients, say  $b_k$  we multiply by the bra  $\langle v_k|$ , obtaining:

$$\sum_{i=1}^N \langle v_k|Aa_i|v_i\rangle = b_k \rightarrow \sum_i A_{ki}a_i = b_k$$

The action of an operator can be thus seen as a matrix multiplication (again, here we are restricting to bounded, finite dimension spaces that support finite operators, hence this simple matrix representation).

**? Question:** Perform a simple matrix multiplication.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

This is equivalent to  $R_x \cdot \vec{v}_z = \vec{v}_y$ . □

The *domain* of an operator is the subspace on which it acts non-trivially (spanned by  $k \leq N$  vectors).

Two operators  $A$  and  $B$  are equal if their domains are the same and their action is equal  $\forall|\psi\rangle$  in their domains. The sum and product of operators are then defined as

$$(A + B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle$$

$$(AB)|\psi\rangle = A(B|\psi\rangle)$$

The operators are associative:

$$A(BC)|\psi\rangle = (AB)C|\psi\rangle$$

But they are not in general commutative:

$$AB|\psi\rangle \neq BA|\psi\rangle$$

**D: Commutator** . The commutator of two operators is  $[A, B] = AB - BA$ . Two operators commute/are commutable if  $[A, B] = 0$ .

### 2.2.1 Hermitian operators

An important class of operators are self adjoint operators, as observables are described by them.

**D: Adjoint** The adjoint of an operator  $A^\dagger$  is an operator acting on the dual space with the property:  $\langle(A^\dagger\psi)|\varphi\rangle = \langle\psi|(A\varphi)\rangle$ ,  $\forall\{|\psi\rangle, |\varphi\rangle\}$ . We can also have other notations. From  $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$  (where  $*$  indicates the complex conjugate) we have  $\langle(A^\dagger\psi)|\varphi\rangle = \langle\psi|(A\varphi)\rangle = \langle\varphi|A^\dagger\psi\rangle^*$ . Also, we can write the inner product as  $\langle\varphi|(A\psi)\rangle = \langle\varphi|A|\psi\rangle$  and  $\langle(A\varphi)|\psi\rangle = \langle\varphi|A^\dagger|\psi\rangle$ . In matrix representation, this means that the adjoint of an operator is the conjugate transpose of that operator:  $A_{k,j}^\dagger = \langle k|A^\dagger|j\rangle = \langle j|A|k\rangle^* = A_{j,k}^*$ .

**D: Self-adjoint** . A self adjoint operator is an operator such that  $A$  and  $A^\dagger$  operate on the same domain and with the property

$$\langle\psi|A|\varphi\rangle = \langle\varphi|A|\psi\rangle^*$$

or shortly,  $A^\dagger = A$ . In matrix representation we have then:  $A_{ki} = A_{ik}^*$ .

**? Question:** Prove that  $(cA)^\dagger = c^*A^\dagger$

We want to prove that  $(cA)^\dagger = c^*A^\dagger$ . We can take two strategies:

1) From the adjoint operator definition in the form:

$$\langle B^\dagger\phi|\psi\rangle = \langle\phi|B\psi\rangle,$$

with  $B = cA$  we obtain:

$$\langle(cA)^\dagger\phi|\psi\rangle = \langle\phi|cA\psi\rangle = c\langle\phi|A\psi\rangle = c\langle A^\dagger\phi|\psi\rangle = \langle c^*A^\dagger\phi|\psi\rangle$$

2) Alternatively, we can use the adjoint definition in Dirac's notation:

$$\langle\varphi|B^\dagger|\psi\rangle = \langle\psi|B|\varphi\rangle^*,$$

to get:

$$\langle\varphi|(cA)^\dagger|\psi\rangle = \langle\psi|cA|\varphi\rangle^* = c^*\langle\psi|A|\varphi\rangle^* = c^*\langle\varphi|A^\dagger|\psi\rangle = \langle\varphi|c^*A^\dagger|\psi\rangle$$

□

Note that we can write

$$\langle B^\dagger\phi|\psi\rangle = \langle\phi|B\psi\rangle = \langle\varphi|B|\psi\rangle = \langle\psi|B^\dagger|\varphi\rangle^*.$$

The second notation (based on Dirac's notation) could be seen as implying  $(|\varphi\rangle)^\dagger = \langle\varphi|$  (and thus  $(A|\varphi)\rangle^\dagger = \langle A^\dagger\phi|$ ). However, this applies the adjoint operation to a vector, while the adjoint is only properly defined for operators. For discrete dimensional spaces, which allow a matrix representation, there is no ambiguity since we have the equivalence of the adjoint with the complex-transpose of an operator (which can be defined also for vectors)<sup>4</sup>.

<sup>4</sup> See also [quant-ph/9907069](https://arxiv.org/abs/quant-ph/9907069) page 12, for a subtle difference between Hermitian and self-adjoint infinite-dimensional operators

---

**? Question:** Prove that  $(AB)^\dagger = B^\dagger A^\dagger$

$\forall |\psi\rangle$  we have  $|\varphi\rangle = (AB)|\psi\rangle \rightarrow \langle\phi| = \langle\psi|(AB)^\dagger$ . Define  $|\phi\rangle = B|\psi\rangle$ , then  $|\varphi\rangle = A|\phi\rangle$ ,  $\langle\varphi| = \langle\psi|B^\dagger$  and  $\langle\phi| = \langle\varphi|A^\dagger$ , so that  $\langle\phi| = \langle\psi|B^\dagger A^\dagger$ .  $\square$

---

A self-adjoint operator is also Hermitian in bounded, finite space, therefore we will use either term. Hermitian operators have some properties:

1. if  $A, B$  are both Hermitian, then  $A + B$  is Hermitian (but notice that  $AB$  is a priori not, unless the two operators commute, too.).
2. if  $A, B$  are both Hermitian but do not commute, then at least  $AB + BA$  is Hermitian.

---

**? Question:** Prove property # 2.

$(AB + BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB$ .  $\square$

---

Before describing other properties we need the following definition.

**$\mathcal{D}$ : Eigenvector** We define a right eigenvector as a column vector  $|\psi\rangle_R$  satisfying  $A|\psi\rangle_R = \lambda_R|\psi\rangle_R$ , so  $(A - \lambda_R\mathbb{1})|\psi\rangle_R = 0$ , which means the right eigenvalues  $\lambda_R$  must have zero determinant, i.e.,  $\det(A - \lambda_R\mathbb{1}) = 0$ . Similarly, a left eigenvector is such that  $\langle\psi|_L A = \lambda_L\langle\psi|_L$ .

The following properties will be very important in QM:

3. if  $A$  is Hermitian its eigenvalues are real (eigenvalues: scalar  $a$  such that  $A|\psi\rangle = a|\psi\rangle$ ). It is easy to show this properties from  $\langle\psi|A|\psi\rangle = a = a^*$ .
4. distinct eigenvectors of an Hermitian operator are orthogonal:  $A|\psi_1\rangle = a_1|\psi_1\rangle$ ,  $A|\psi_2\rangle = a_2|\psi_2\rangle \rightarrow \langle\psi_1|\psi_2\rangle = 0$  unless  $a_1 = a_2$ .
5. distinct eigenvalues correspond to orthogonal eigenvectors:

Given  $A|\psi_1\rangle = c_1|\psi_1\rangle$  and  $A|\psi_2\rangle = c_2|\psi_2\rangle$ , if  $c_1 \neq c_2 \rightarrow \langle\psi_1|\psi_2\rangle = 0$ .

As observables are given by Hermitian operators, the first properties will imply that the values that an observable can take on are only real values (as needed for the observable to have a physical meaning). On the domain of the operator, the eigenvectors form a complete orthogonal basis set.

**? Question:** Prove property # 5.

$\langle\psi_2|A\psi_1\rangle = \langle\psi_2|c_1\psi_1\rangle = \langle c_2^*\psi_2|\psi_1\rangle$ . For Hermitian operators then  $c_1 \langle\psi_2|\psi_1\rangle = c_2 \langle\psi_2|\psi_1\rangle$ , which is satisfied only if  $c_1 = c_2$  or if  $\langle\psi_1|\psi_2\rangle = 0$ . □

**? Question:** Prove property # 4.

Consider two eigenstates of  $A$   $|a_1\rangle$  and  $|a_2\rangle$ . We have  $\langle a_2|A|a_1\rangle = \langle a_1|A|a_2\rangle^*$  since  $A$  is Hermitian. Now  $\langle a_2|A|a_1\rangle = a_1\langle a_2|a_1\rangle$  and  $\langle a_1|A|a_2\rangle^* = (a_2\langle a_1|a_2\rangle)^* = a_2^*(\langle a_1|a_2\rangle)^*$  since  $a_2$  is real (being an eigenvector of  $A$ ). We thus have  $a_1\langle a_2|a_1\rangle = a_2\langle a_2|a_1\rangle$  which is satisfied iff  $a_1 = a_2$  (contrary to the hypothesis) or if  $\langle a_2|a_1\rangle = 0$ . □

## 2.2.2 Operators and their properties

**D: The Outer Product**  $|\psi\rangle\langle\varphi|$  is an operator, since acting on a vector returns a vector:  $(|\psi\rangle\langle\varphi|)|\phi\rangle = \langle\varphi|\phi\rangle|\psi\rangle$ .

It defines a projector operator  $P_i = |v_i\rangle\langle v_i|$ . The sum over all projectors on the space is the identity, therefore, for any basis set we have:  $\sum_i |v_i\rangle\langle v_i| = \mathbb{1}$  (closure relation). The product of two projectors is  $P_j P_k = \delta_{jk} P_j$ . Projectors derive their name from the property that they project out a vector component of the related basis vector: given  $P_j = |v_j\rangle\langle v_j|$ ,  $P_j|\psi\rangle = P_j \sum_k c_k |v_k\rangle = c_j |v_j\rangle$ .

**D: Trace** - The trace of an operator is the sum of the diagonal elements of an operator  $\text{Tr}\{A\} = \sum_{j=1}^N A_{jj} = \sum_j \langle v_j|A|v_j\rangle$ . It is independent of the choice of basis.

**D: Spectral Decomposition** - The spectral theorem states that given a self-adjoint operator  $A$  on a linear space  $\mathcal{H}$ , there exists an orthonormal basis of  $\mathcal{H}$  consisting of eigenvectors of  $A$ . Equivalently, we can state that  $A$  can be written as a linear combination of pairwise orthogonal projections (which are formed from its eigenvectors). This representation of  $A$  is called its spectral decomposition:  $A = \sum_j a_j |v_j\rangle\langle v_j|$ , where  $A|v_j\rangle = a_j |v_j\rangle$ . In this basis, the matrix representation of  $A$  is diagonal.

**■ Theorem:** If two hermitian operators commute, they share a common set of eigenvectors.

If  $[A, B] = 0$  then  $AB = BA$ . Given two eigenvectors of  $A$ , we have  $\langle a''|(AB - BA)|a''\rangle = a' \langle a'|B|a''\rangle - a'' \langle a'|B|a''\rangle$ . This is zero if  $a'' = a'$  (and  $\langle a'|B|a'\rangle$  is a diagonal term of  $B$  and it can be anything) or if  $\langle a'|B|a''\rangle = 0$  (off-diagonal, with  $a' \neq a''$ ). Thus  $B$  is diagonal in the basis of  $A$ 's eigenvectors, hence  $A$ 's eigenvectors are also eigenvectors of  $B$ . □

A simultaneous eigenvector of  $A$  and  $B$   $|a, b\rangle$  has the property:  $A|a, b\rangle = a|a, b\rangle$  and  $B|a, b\rangle = b|a, b\rangle$ . The notation  $|a, b\rangle$  is useful when the eigenvector is **degenerate**, that is, there exist more than one eigenvector with the same eigenvalue:  $A|a^{(i)}\rangle = a|a^{(i)}\rangle$ ,  $i = 1, \dots, n$ , where  $n$  is the degeneracy. Then the label  $b$  serves to distinguish different eigenvectors.

**D: Unitary operator** An operator fulfilling the conditions  $U^\dagger U = \mathbb{1}$  and  $UU^\dagger = \mathbb{1}$  is called unitary.

**■ Theorem:** Given two sets of basis kets  $\{|\psi\rangle_i\}$  and  $\{|\phi\rangle_i\}$  there exist a unitary operator such that  $|\phi\rangle_i = U|\psi\rangle_i$ ,  $\forall i$ . (The unitary operator is  $U = \sum_k |\varphi_k\rangle\langle\psi_k|$ ).

### 2.2.3 Functions of operators

Functions of operators are defined by the corresponding Taylor expansion of the function (if that exists). If  $f(x) = f(0) + f'(0)x + \dots + \frac{1}{n!}f^{(n)}(0)x^n + \dots$ , then  $f(A) = f(0)\mathbb{1} + f'(0)A + \dots + \frac{1}{n!}f^{(n)}(0)A^n + \dots$ , where the matrix power is defined recursively in terms of products  $A^n = A(A^{n-1})$ .

---

**? Question:** Show that given the spectral decomposition of the operator  $A = \sum_a \lambda_a |a\rangle \langle a|$  we have  $f(A) = \sum_a f(\lambda_a) |a\rangle \langle a|$ . We can first prove that  $A^2 = (\sum_a \lambda_a |a\rangle \langle a|)(\sum_a \lambda_a |a\rangle \langle a|) = \sum_{a,b} \lambda_a \lambda_b (|a\rangle \langle a|)(|b\rangle \langle b|) = \sum_a \lambda_a^2 |a\rangle \langle a|$ . Then show that if the theorem is valid for  $n - 1$  it is also valid for  $n$ . Finally, use the Taylor expansion to show it's true.  $\square$

**? Question:** Consider in particular the exponential function:  $\exp(\xi A) = \sum_n \frac{1}{n!}(\xi A)^n = \sum_k \exp(\xi a_k) |a_k\rangle \langle a_k|$ . Prove that  $f(ABA^{-1}) = Af(B)A^{-1}$ . It's easy to show that  $(ABA^{-1})^n = AB^nA^{-1}$  by expanding the product and using  $AA^{-1} = \mathbb{1}$ . In particular for unitary matrices  $U^{-1} = U^\dagger \rightarrow f(UAU^\dagger) = Uf(A)U^\dagger$ .  $\square$

---

## 3. Axioms of Quantum Mechanics

### 3.1 Introduction

### 3.2 The axioms of quantum mechanics

#### 3.2.1 Observables and State Space

#### 3.2.2 Quantum measurement

#### 3.2.3 Law of motion

### 3.3 Strong measurements

#### 3.3.1 Expectation values

#### 3.3.2 Uncertainty relationships

#### 3.3.3 Repeated measurements and Quantum Zeno Effect

### 3.1 Introduction

Every physical theory is formulated in terms of mathematical objects. It is thus necessary to establish a set of rules to map physical concepts and objects into mathematical objects that we use to represent them<sup>5</sup>. Sometimes this mapping is evident, as in classical mechanics, while for quantum mechanics the mathematical objects are not intuitive. In the same way as classical mechanics is founded on Newton's law or electrodynamics on the Maxwell-Boltzmann equations, quantum mechanics is also based on some fundamental laws, which are called the postulates or axioms of quantum mechanics. The axioms we are going to see apply to the dynamics of closed quantum systems. We want to develop a mathematical model for the dynamics of closed systems: therefore we are interested in defining states, observables, measurements and evolution. Some subtleties will arise since we are trying to define measurement in a closed system, when the measuring person is instead outside the system itself. We give below (and explain in the next few sections) one formulation of the QM axioms. Different presentations (for example starting from density operators instead of state vectors) are possible.

1. The properties of a quantum system are completely defined by specification of its state vector  $|\psi\rangle$ . The state vector is an element of a complex Hilbert space  $\mathcal{H}$  called the space of states.
2. With every physical property  $\mathcal{A}$  (energy, position, momentum, angular momentum, ...) there exists an associated linear, Hermitian operator  $A$  (usually called observable), which acts in the space of states  $\mathcal{H}$ . The eigenvalues of the operator are the possible values of the physical properties.
- 3.a If  $|\psi\rangle$  is the vector representing the state of a system and if  $|\varphi\rangle$  represents another physical state, there exists a probability  $p(|\psi\rangle, |\varphi\rangle)$  of finding  $|\psi\rangle$  in state  $|\varphi\rangle$ , which is given by the squared modulus of the scalar product on  $\mathcal{H}$ :  $p(|\psi\rangle, |\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$  (Born Rule).
- 3.b If  $A$  is an observable with eigenvalues  $a_k$  and eigenvectors  $|k\rangle$  ( $A|k\rangle = a_k|k\rangle$ ), given a system in the state  $|\psi\rangle$ , the probability of obtaining  $a_k$  as the outcome of the measurement of  $A$  is  $p(a_k) = |\langle k|\psi\rangle|^2$ . After the measurement the system is left in the state projected on the subspace of the eigenvalue  $a_k$  (Wave function collapse).
4. The evolution of a closed system is unitary. The state vector  $|\psi(t)\rangle$  at time  $t$  is derived from the state vector  $|\psi(t_0)\rangle$  at time  $t_0$  by applying a unitary operator  $U(t, t_0)$ , called the evolution operator:  $|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle$ .

---

<sup>5</sup> See: Leslie E. Ballentine, "Quantum Mechanics A Modern Development", World Scientific Publishing (1998). We follow his presentation in this section.

## 3.2 The axioms of quantum mechanics

### 3.2.1 Observables and State Space

A physical experiment can be divided into two steps: preparation and measurement. The first step determines the possible outcomes of the experiment, while the measurement retrieves the value of the outcome. In QM the situation is slightly different: the first step determines the *probabilities* of the various possible outcomes, while the measurement retrieve the value of a particular outcome, in a statistic manner. This separation of the experiment is reflected into the two types of mathematical objects we find in QM. The first step corresponds to the concept of a *state* of the system, while the second to observables.

The state gives a complete description of the set of probabilities for all observables, while these last ones are all dynamical variables that in principle can be measured. All the information is contained in the state, irrespectively on how I got the state, of its previous history. For the moment we will identify the state with the vectors of an Hilbert space  $|\psi\rangle$ . We will see later on that a more general definition exists in terms of state operators  $\rho$ .

All physical observables (defined by the prescription of experiment or measurement ) are represented by a linear operator that operates in linear inner product space (an Hilbert space in case of finite dimensional spaces). States of the system are represented by the direction/ray (not a vector) in the linear inner product space (again Hilbert space in the finite dimensional case).

### 3.2.2 Quantum measurement

The value of the measurement of an observable is one of the observable eigenvalues. The probability of obtaining one particular eigenvalue is given by the modulus square of the inner product of the state vector of the system with the corresponding eigenvector. The state of the system immediately after the measurement is the normalized projection of the state prior to the measurement onto the eigenvector subspace.

Let  $A$  be the observable with eigenvalues  $a_k$  and eigenvectors  $|k\rangle$ :  $A|k\rangle = a_k|k\rangle$ . Given a system in the state  $|\psi\rangle$ , the probability of obtaining  $a_k$  as the outcome of the measurement of  $A$  in this system is

$$p(a_k) = |\langle k|\psi\rangle|^2.$$

We can also write this in terms of the  $k^{th}$  eigenvector projector  $P_k = |k\rangle\langle k|$ :  $p(a_k) = \langle\psi|P_k|\psi\rangle$ . Since here we are considering strong, projective measurement, also called Von Neumann measurements, immediately after a measurement that gave us the result  $a_k$ , the state of the system is in the  $|k\rangle$  eigenstate. More precisely, the normalized output state after the measurement is

$$|\psi'\rangle = \frac{P_k|\psi\rangle}{\sqrt{|\langle\psi|P_k|\psi\rangle|}}.$$

If we repeat the experiment after the first measurement, we will obtain again the same result (with probability 1). If  $|\psi\rangle$  is an eigenstate of  $A$ ,  $A|\psi\rangle = a_\psi|\psi\rangle$ , then we will measure  $a_\psi$  with probability unity. This is the well-known collapse of the wavefunction.

The collapse of the wavefunction is of course a source of confusion and contradictions: as stated above it appears as an almost instantaneous evolution of the system from a given state to another one, an evolution which is not unitary (as evolution should be per axiom # 4). The source of contradiction stems from the fact that in this simple description of the measurement, the observer (or the measurement apparatus) are *external* to the system (thus the assumption of closed system is not respected) and might not even be quantum-mechanical. A more advanced theory of measurement attempts to solve these issues<sup>6</sup>.

On the other side, we note that operationally the wavefunction collapse is required to define a well-formulated theory. The collapse allows the experimenter to check the result of the measurement by repeating it (on the system just observed) thus giving confidence on the measurement apparatus and procedure. If this were not the case, no measurement could be ever believed to be the correct one, so no confirmation of the theory could be done.

#### Reference

M. Brune, E. Hagley, J. Dreyer, X. Matre, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche *Observing the Progressive Decoherence of the Meter in a Quantum Measurement*, [Phys. Rev. Lett.](#) **77**, 4887 - 4890 (1996)

---

<sup>6</sup> In addition to the “strong” or projective measurement presented here, generalized models for measurement exist, see for example POVM in Prof. Preskill’s online notes

### 3.2.3 Law of motion

We can define the time evolution operator  $U$ , such that

$$|\psi'\rangle = U|\psi\rangle, \text{ with } U^\dagger U = 1.$$

Since the state has all the information about the system at time  $t$ , the state of the system at the time  $t + dt$  depends only on the state at time  $t$  and on the evolution operator  $U(t, t + dt)$  (that thus depends only on the times  $t$  and  $t + dt$ , not on any previous times, otherwise it would bring extra information to the system).

The unitarity of the evolution is equivalent to the following statement regarding the evolution of the state vector. The dynamics of the system are generated by the system Hamiltonian  $\mathcal{H}$  (the observable corresponding to the total energy of the system), as described by Schrödinger equation:

$$i\hbar \frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

where  $\hbar$  is the reduced Planck's constant<sup>7</sup> ( $1.0545 \times 10^{-34}$  Js).

We would like to link this second statement (Schrödinger equation) to the previous statement regarding the unitarity of the evolution. To do so we first look at the evolution for an infinitesimal time  $dt$ .

For an infinitesimal evolution we have then:  $|\psi(t + dt)\rangle = |\psi\rangle - i dt \mathcal{H}|\psi\rangle$ . It follows that  $U(t, dt) = \mathbb{1} - i\mathcal{H}dt$ . Since the Hamiltonian is a self-adjoint operator, to the same order of approximation we retrieve the fact that  $U$  is unitary:  $UU^\dagger = (\mathbb{1} - i\mathcal{H}dt)(\mathbb{1} + i\mathcal{H}dt) \approx \mathbb{1} + o(dt^2)$ .

We can build the dynamics for any time duration in terms of infinitesimal evolutions,  $U(t, t') = U(t', t' - dt) \dots U(t + 2dt, t + dt)U(t + dt, t)$  since the propagator  $U$  depends only on the time  $t$ .

If the Hamiltonian is time independent (and setting  $t' = 0$ ), we obtain:  $|\psi(t)\rangle = U(0, t)|\psi(0)\rangle$ , where the evolution operator  $U$  is given by  $U = e^{-i\hat{\mathcal{H}}t}$ , i.e.  $U$  is an exponential operator.

**? Question:** Show from the infinitesimal time product and the Taylor expansion for the exponential that this is indeed the case. □

Equivalently, we can find a differential equation for the dynamics of the propagator: from  $U(t + dt, t_0) - U(t, t_0) = -\frac{i}{\hbar}\mathcal{H}U(t, t_0)$  we have the Schrödinger equation for the time evolution operator (propagator):

$$i\hbar \frac{\partial U}{\partial t} = \mathcal{H}U$$

This equation is valid also when the Hamiltonian is time-dependent (and we will see later on a formal solution to this equation).

As the Hamiltonian represents the energy of the system, its *spectral representation* is defined in terms of the energy eigenvalues  $\epsilon_k$ , with corresponding eigenvectors  $|k\rangle$ :  $\mathcal{H} = \sum_k \epsilon_k |k\rangle\langle k|$ . The evolution operator is then:  $U = \sum_k e^{-i\epsilon_k t} |k\rangle\langle k|$ . The eigenvalues of  $U$  are therefore simply  $e^{-i\epsilon_k t}$ , and it is common to talk in terms of eigenphases  $\epsilon_k t$ . If the Hamiltonian is time-independent we have also  $U^\dagger = U(-t)$ , it is possible to obtain an effective inversion of the time arrow.

**? Question:** What is the evolution of an energy eigenvector  $|k\rangle$ ?

First consider the infinitesimal evolution:  $|k(t + dt)\rangle = U(t + dt, t) |k(t)\rangle = (\mathbb{1} - i\mathcal{H}dt) |k(t)\rangle = (1 - i\epsilon_k dt) |k(t)\rangle$ . Thus we have the differential equation for the energy eigenket:  $\frac{d|k\rangle}{dt} = -i\epsilon_k |k\rangle$ , so that  $|k(t)\rangle = e^{-i\epsilon_k t} |k(0)\rangle$ . We can also use the spectral decomposition of  $U$ :  $|k(t)\rangle = U(t, 0) |k(0)\rangle = (\sum_h e^{-i\epsilon_h t} |h\rangle\langle h|) |k(0)\rangle = e^{-i\epsilon_k t} |k(0)\rangle$ . □

In conclusion, our picture of QM is a mathematical framework in which the system is completely described by its state, which undergoes a **deterministic** evolution (and invertible evolution). The measurement process, which connects the mathematical theory to the observed experiments, is probabilistic.

<sup>7</sup> We will quite often set  $\hbar = 1$ , that is, we will measure the energies in frequency units

### 3.3 Strong measurements

#### 3.3.1 Expectation values

Although the result of a single measurement is probabilistic, we are usually interested in the average outcome, which gives us more information about the system and observable. The average or *expectation value* of an observable for a system in state  $|\psi\rangle$  is given by

$$\langle A \rangle = \langle \psi | A | \psi \rangle$$

---

**? Question:** Prove that this can be simply derived from the usual definition of average

$\langle A \rangle = \sum_n p(a_k) a_k = \sum_n |\langle \psi | a_k \rangle|^2 a_k = \sum_n \langle \psi | n \rangle \langle n | \psi \rangle a_k = \langle \psi | (\sum_n a_k |k\rangle \langle k|) | \psi \rangle$ , and we get the desired result from  $A = \sum_n a_k |k\rangle \langle k|$ .  $\square$

---

#### 3.3.2 Uncertainty relationships

**D: Compatible Observables** Two observables  $A, B$  are said to be compatible if their corresponding operators commute  $[A, B] = 0$  and incompatible otherwise.

**D: Degeneracy** If there exist two (or more) eigenstates of an operator  $A$  with the same eigenvalues, they are called degenerate.

We have already seen how commuting operators have common eigenvectors and how a compatible observables can be used to distinguish between degenerate eigenvectors. We now look from a more physical point of view at the meaning of commuting (or compatible) observables. Suppose we first measure  $A$ , given a state  $|\psi\rangle$ . We retrieve e.g. the eigenvalue  $a$  and the state is now projected into the eigenstate  $|a\rangle$ . Allowing for the presence of degenerate eigenstates, we actually have a superposition state  $|\psi\rangle_{\text{Post-Meas}} = \sum_{i=1}^d c_i |a, b_i\rangle$ , where  $d$  is the degree of degeneracy of the eigenvalue  $a$ . We then measure  $B$  obtaining one of the  $b_i, \tilde{b}$ . The state is thus projected into  $|a, \tilde{b}\rangle$ . If we now measure again  $A$  we will retrieve the same result as before: the two measurements of commuting observables  $A$  and  $B$  do not interfere.

Consider now non-commuting observables. As  $AB|\psi\rangle \neq BA|\psi\rangle$  we cannot define a state  $|a, \tilde{b}\rangle$  which is described by the (separate) eigenvectors of the two observables. Also, if we repeat the same 3 successive measurements as above, we obtain a different result. In particular, the second measurement of  $A$  does not in general retrieve the same eigenvalue as the first one.

---

**? Question:** Show why measurement of non-commuting observables are not compatible.

Given a state  $|\psi\rangle$  we measure  $A$ , with result  $a$ . The state is now projected into the eigenstate  $|a\rangle$  as before (we neglect here degeneracy). Now we rewrite this state in the basis of the operator  $B$  (which is not the same as the basis for  $A$ , so  $|a\rangle \notin \{|b_i\rangle\}$ ):  $|a\rangle = \sum_i c_i(a) |b_i\rangle$ . When we measure  $B$  we will therefore obtain an eigenvalue  $b_i$  with probability  $|c_i(a)|^2$ , and the state is projected into:  $\frac{P_i|a\rangle}{\sqrt{\langle a | P_i | a \rangle}} = \frac{|b_i\rangle \langle b_i | a \rangle}{|\langle a | \langle b_i | \langle b_i | \langle a | \rangle|^{1/2}} = |b_i\rangle$ .

Again, this can be written as a non-trivial superposition of eigenstates of  $A$ :  $|b_i\rangle = \sum_j c_j(b_i) |a_j\rangle$  so that it is now possible to obtain a measurement  $a_j \neq a$  when we again measure  $A$ .  $\square$

---

**D: Variance of an operator.** We define an operator  $\Delta A = A - \langle A \rangle$  for any observable  $A$ . The expectation value of its square is the variance of  $A$ :  $\langle \Delta A^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2$ .

**■ Theorem: (Uncertainty relation).** For any two observables, we have

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2$$

From Schwartz inequality ( $|\langle\psi|\varphi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\varphi|\varphi\rangle$ ) we have  $\langle\Delta A^2\rangle\langle\Delta B^2\rangle \geq |\langle\Delta A\Delta B\rangle|^2$ . Now  $\Delta A\Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}$  (where we defined the anticommutator  $\{A, B\} = AB + BA$ ). Taking the expectation value (noting that  $\langle[\Delta A, \Delta B]\rangle = \langle[A, B]\rangle$ ) we have

$$\langle\Delta A\Delta B\rangle = \frac{1}{2}\langle[A, B]\rangle + \frac{1}{2}\langle\{\Delta A, \Delta B\}\rangle.$$

Now we can show that  $[A, B] = iC$  and  $\{A, B\} = D$  where  $C$  and  $D$  are hermitian operators. Then the first term in the RHS is purely imaginary and the second purely real. Thus we have:

$$\langle\Delta A^2\rangle\langle\Delta B^2\rangle \geq |\langle\Delta A\Delta B\rangle|^2 = \frac{1}{4}|\langle[A, B]\rangle|^2 + \frac{1}{4}|\langle\{\Delta A, \Delta B\}\rangle|^2 \geq \frac{1}{4}|\langle[A, B]\rangle|^2.$$

□

### 3.3.3 Repeated measurements and Quantum Zeno Effect

#### A. Photon Polarization

In the same way an electromagnetic wave can be polarized, also individual photons possess a polarization. In particular they can be in a state of linear or circular polarization (the most general case, is called elliptical polarization). We consider a photon polarizer. This can be thought as a filter that ensures photons coming out of it are only of the right polarization.

— In-class demonstration with polarizer filters —

The photon polarizer (a polarization filter) is very similar to a measurement process and indeed it can be described by a projector. Let's assume that light can be described as either being in the horizontal  $|h\rangle$  or vertical  $|v\rangle$  polarization. Then, for an horizontal polarizer, for example, we have  $P_H = |h\rangle\langle h|$ . If we send a photon in the state  $|\psi\rangle$  through this linear (horizontal) polarizer, its state after the polarizer will be  $|h\rangle$ . However the photon will emerge only with a probability  $|\langle h|\psi\rangle|^2$ . If we then send the photon to an orthogonal (vertical) polarizer  $P_V = |v\rangle\langle v|$ , the probability of a photon coming out is just zero. This situation is very similar to the case of repeated measurement. Thus the polarizer is a measurement process.

Now let's send an horizontally polarized photon ( $|h\rangle$ ) through a 45 degrees polarizer. This polarizer can be described by the projector operator  $P_{45} = |h+v\rangle\langle h+v|/2$ . The state after the polarizer is then  $(|h+v\rangle)/\sqrt{2}$ , and the probability of coming out is  $\frac{1}{2}$ . If now we send this photon through a  $|v\rangle\langle v|$  polarizer, we obtain as a final state  $|v\rangle$ , and the total probability is  $1/4$  (compare to zero before).

We can extend this even further. Assume we have a large number of polarizers each ensuring a polarization at a growing angle, each in a small step  $\vartheta$  with the horizontal (that is, the first polarizer's angle is  $\vartheta$ , the second  $2\vartheta$  etc.). The relevant projector is then

$$P_n(\vartheta) = (\cos(n\vartheta)|h\rangle + \sin(n\vartheta)|v\rangle)(\cos(n\vartheta)\langle h| + \sin(n\vartheta)\langle v|).$$

We start with a photon horizontally polarized  $|\psi\rangle_0 = |h\rangle$ . After the first polarizer, the photon emerges through in the state  $|\psi\rangle_1 = \cos\vartheta|h\rangle + \sin\vartheta|v\rangle$  with probability  $p_1(\vartheta) = |(\cos\vartheta\langle h| + \sin\vartheta\langle v|)|h\rangle|^2 = \cos^2\vartheta$ . Now passing through the second polarizer the photon will emerge again with probability  $\cos^2\vartheta$  and in the state  $|\psi\rangle_2 = \cos(2\vartheta)|h\rangle + \sin(2\vartheta)|v\rangle$ . After  $n$  polarizers, the state of the emerging photon is

$$|\psi\rangle_n = \cos(n\vartheta)|h\rangle + \sin(n\vartheta)|v\rangle.$$

Of course, we could get no photon at all, however the combined probability of getting a photon is  $\cos(\vartheta)^{2n} \approx 1$  if the angle  $\vartheta$  is small and the number of polarizer  $n$  is large. Thus we obtain an evolution of the system by using a measurement process.

#### B. Quantum Zeno effect

We consider a photon polarization rotator, whose action is to rotate the polarization about the propagation axis. By denoting  $\{h, v\}$  the horizontal and vertical polarization, respectively, the polarization rotator achieves the following transformation:

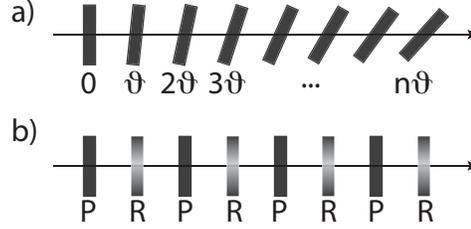


Fig. 1: a) Rotation by measurement. b) Quantum Zeno effect

$$\begin{aligned} |h\rangle &\rightarrow \cos(\vartheta)|h\rangle + \sin(\vartheta)|v\rangle \\ |v\rangle &\rightarrow \cos(\vartheta)|v\rangle - \sin(\vartheta)|h\rangle \end{aligned}$$

This corresponds to the evolution matrix  $U$ :

$$U = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

---

**? Question:** What are the eigenstates of  $U$ ?

By diagonalizing the matrix, we find the eigenvectors corresponding to right and left polarization:

$$\begin{aligned} R &= (|h\rangle + i|v\rangle)/\sqrt{2} \\ L &= (-i|h\rangle + |v\rangle)/\sqrt{2} \end{aligned}$$

With eigenvalues  $e^{i\vartheta}$  and  $e^{-i\vartheta}$  respectively. The evolution given by the polarization rotator is unitary. □

---

Now assume another experiment in which we alternate a polarizer rotator and an horizontal polarizer. First consider just a set of polarizer rotators, each described by the formula above:

$$U(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

After  $n$  of these rotators, the photon is rotated to  $U(\vartheta)^n |h\rangle = U(n\vartheta) |h\rangle = \cos(n\vartheta) |h\rangle + \sin(n\vartheta) |v\rangle$ . Now if we alternate with the horizontal polarizer, every time the photon is transmitted with probability  $\cos^2 \vartheta$  and rotate back to  $|h\rangle$ . Again for  $\vartheta$  small, the probability of a photon emerging tends to 1, and the final state of the photon is  $|h\rangle$ . This is a phenomenon called quantum Zeno effect<sup>8</sup> or we can call it a "watched milk never boils" phenomenon. The repeated measurements inhibit a (slow) evolution.

#### References

- B. Misra and E. C. G. Sudarshan, *The Zeno's paradox in quantum theory* *J. Math. Phys.* **18**, 756 (1977).
- W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, *Quantum Zeno effect*, *Phys. Rev. A* **41**, 2295 - 2300 (1990).
- Saverio Pascazio, Mikio Namiki, Gerald Badurek and Helmut Rauch *Quantum Zeno effect with neutron spin*, *Physics Letters A*, **179**, 155 - 160, (1993).

---

<sup>8</sup> Zeno's paradoxes are a set of problems (8 of which surviving) generally thought to have been devised by Zeno of Elea to support Parmenides's doctrine that "all is one" and that in particular, contrary to the evidence of our senses, motion is nothing but an illusion. The arrow paradox as related by Aristotle, (*Physics* VI:9, 239b5) states that "The third is ... that the flying arrow is at rest, which result follows from the assumption that time is composed of moments ... he says that if everything when it occupies an equal space is at rest, and if that which is in locomotion is always in a now, the flying arrow is therefore motionless." To make the argument more similar to the QM version, we can rephrase it as: If you look at an arrow in flight, at an instant in time, it appears the same as a motionless arrow. Then how do we see motion?

## 4. Two-level systems

### 4.1 Generalities

### 4.2 Rotations and angular momentum

#### 4.2.1 Classical rotations

#### 4.2.2 QM angular momentum as generator of rotations

#### 4.2.3 Example of Two-Level System: Neutron Interferometry

#### 4.2.4 Spinor behavior

#### 4.2.5 The SU(2) and SO(3) groups

### 4.1 Generalities

We have already seen some examples of systems described by two possible states. A neutron in an interferometer, taking either the upper or lower path. A photon linearly polarized either horizontally or vertically. A two level system (TLS) is the simplest system in quantum mechanics, but it already illustrates many characteristics of QM and it describes as well many physical systems. It is common to reduce or map quantum problems onto a TLS. We pick the most important states -the ones we care about- and then discard the remaining degrees of freedom, or incorporate them as a collection or continuum of other degrees of freedom termed a bath.

In a more abstract way, we can think of a TLS as carrying a binary information (the absence or presence of something, the information about a position, such as left or right, or up or down, etc.). Thus a TLS can be thought as containing a bit of information. By analogy with classical computers and information theory, TLS are thus called *qubits*. Their basis states are usually defined as  $|0\rangle$  and  $|1\rangle$  with a vector representation:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A general state is then  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ . If it is normalized we have  $|\alpha|^2 + |\beta|^2 = 1$ . Then, this state can also be written quite generally in terms of the two angles  $\vartheta$  and  $\varphi$ :

$$|\psi\rangle = \cos(\vartheta/2)|0\rangle + e^{i\varphi} \sin(\vartheta/2)|1\rangle$$

For this state, the probability of finding the system in the  $0[1]$  state is  $\cos^2(\vartheta/2)$  [ $\sin^2(\vartheta/2)$ ]. Notice that I could have written the state also as

$$|\psi\rangle \equiv |\phi\rangle = e^{-i\phi/2} \cos(\vartheta/2)|0\rangle + e^{i\phi/2} \sin(\vartheta/2)|1\rangle$$

The two states are in fact equivalent up to a *global* phase factor. While relative phase factors (in a superposition) are very important, global phases are irrelevant, since they yield the same results in a measurement outcome.

---

**? Question:** Show that a global phase factor does not change measurement outcomes and measurement statistics.

1.(Measurement outcome) As the possible measurement outcomes are the eigenvalues of the measurement operators, the first is trivially true.

2.(Statistics) Let's consider an observable  $A$  with eigenvectors  $|a\rangle = a_0|0\rangle + a_1|1\rangle$  corresponding to the eigenvalues  $a$ , then the probability of obtaining  $a$  from the measurement is  $p(a) = \|\langle\psi|a\rangle\|^2 = |a_0 \cos(\vartheta/2) + a_1 e^{i\phi} \sin(\vartheta/2)|^2 = \|\langle\phi|a\rangle\|^2 = |a_0 e^{-i\phi/2} \cos \vartheta/2 + a_1 e^{i\phi/2} \sin \vartheta/2|^2$ .  $\square$

---

Consider for example  $|a\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . Then we obtain  $p(a) = \frac{1}{2}(1 + \cos\varphi \sin\vartheta)$ . Thus the *relative* phase of  $|0\rangle$  w.r.t.  $|1\rangle$  is important. But a global phase multiplying the state is not.

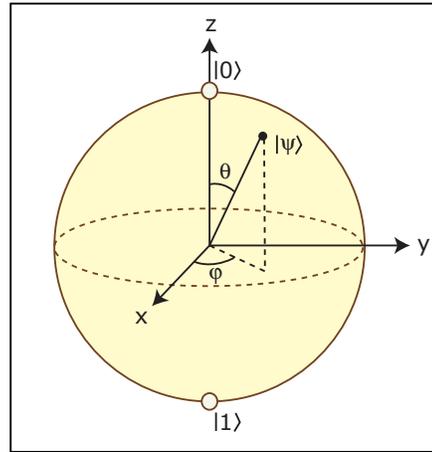


Image by MIT OpenCourseWare.

Fig. 2: Bloch sphere representation of a qubit (TLS) [From wikipedia]

Describing a TLS via the two angles  $\theta$  and  $\phi$  leads to a simple geometrical picture for the space occupied by this system. The angles define a point on a sphere of radius 1, which is called **Bloch sphere**. The TLS can then assume any of the points *inside* the sphere, as well as means to reach them). The unitary evolution for this particular system can then be described as *rotations* of the state vector in the sphere.

Using the example of a TLS we are thus going to introduce the concept of rotation and angular momentum, which can be generalized also to larger systems.

## 4.2 Rotations and angular momentum

### 4.2.1 Classical rotations

Let's review rotation in classical mechanics (geometry). The first property that we want to analyze is the fact that successive rotations about different axes do not commute. Consider for example to start with a vector aligned along the z axis and then effectuate two rotations, one about the y axis and one about the z axis. Depending on the order, we obtain a rotation or no rotation at all.

Rotation are represented in 3D by orthogonal  $3 \times 3$  matrices. (an *orthogonal* matrix is such that  $RR^T = R^T R = \mathbf{1}$ ). In particular, rotations about the 3 axes are as follow:

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{pmatrix}$$

$$R_y(\phi) = \begin{pmatrix} \cos(\phi) & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{pmatrix}$$

$$R_z(\phi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

It is easy then to show that  $R_\alpha(\vartheta)R_\beta(\varphi) \neq R_\beta(\varphi)R_\alpha(\vartheta)$  unless  $\alpha = \beta$ . What about if the rotation angles are very small? We might expect then that the order matters less. We thus consider infinitesimal rotations, where  $\phi = \epsilon \rightarrow 0$ :

$$\begin{aligned}
R_x(\epsilon) &\approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\epsilon^2}{2} & -\epsilon \\ 0 & \epsilon & 1 - \frac{\epsilon^2}{2} \end{pmatrix} \\
R_y(\epsilon) &= \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \frac{\epsilon^2}{2} \end{pmatrix} \\
R_z(\epsilon) &= \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & -\epsilon & 0 \\ \epsilon & 1 - \frac{\epsilon^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

If we then calculate for example  $R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon)$  we obtain:

$$\begin{aligned}
R_x(\epsilon)R_y(\epsilon) &= \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & 0 & \epsilon \\ \epsilon^2 & 1 - \frac{\epsilon^2}{2} & -\epsilon \left(1 - \frac{\epsilon^2}{2}\right) \\ -\epsilon \left(1 - \frac{\epsilon^2}{2}\right) & \epsilon & \left(1 - \frac{\epsilon^2}{2}\right)^2 \end{pmatrix} \\
R_y(\epsilon)R_x(\epsilon) &= \begin{pmatrix} 1 - \frac{\epsilon^2}{2} & \epsilon^2 & \epsilon - \frac{\epsilon^3}{2} \\ 0 & 1 - \frac{\epsilon^2}{2} & -\epsilon \\ -\epsilon & \epsilon - \frac{\epsilon^3}{2} & \left(1 - \frac{\epsilon^2}{2}\right)^2 \end{pmatrix}
\end{aligned}$$

and

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 0 & -\epsilon^2 & \frac{\epsilon^3}{2} \\ \epsilon^2 & 0 & \frac{\epsilon^3}{2} \\ \frac{\epsilon^3}{2} & \frac{\epsilon^3}{2} & 0 \end{pmatrix}$$

Thus we see that

1. If we ignored terms  $\propto \epsilon^2$  and higher, the rotations do commute.
2. At the second order in  $\epsilon$  we can write the result as  $R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = R_z(\epsilon^2) - \mathbb{1}$ . This result stands also for cyclic permutations of the subscripts. These commutation relationships are a guide in finding commutation relationships that the equivalent QM rotation operators should obey.

#### 4.2.2 QM angular momentum as generator of rotations

In QM we can as well define rotations, as we already did for classical mechanics. Although we will first study examples for TLS, rotations can be defined for any system (even higher dimensional systems). Generally, a rotation will be represented by an operator  $\mathcal{D}(R_\alpha(\phi))$  associated to a classical rotation  $R_\alpha(\phi)$ . We first define the action of an infinitesimal rotation. To do so we define the angular momentum operator  $\mathbf{J}$  in terms of the infinitesimal rotation :

$$\mathcal{D}(R_{\mathbf{n}}(\delta\phi)) = \mathbb{1} - i\delta\phi\vec{J} \cdot \vec{n}$$

where  $\vec{n}$  is a unit vector. A finite rotation can be found by repeating many infinitesimal rotations. For example, for a rotation about  $\mathbf{z}$ :

$$\mathcal{D}(R_z(\varphi)) = \lim_{N \rightarrow \infty} \left[ \mathbb{1} - iJ_z \frac{\varphi}{N} \right]^N = \mathbb{1} - iJ_z\varphi - \frac{1}{2}J_z^2\varphi^2 \dots = \exp(-iJ_z\varphi)$$

(Note that here again I took  $\hbar = 1$ .) The angular momentum can thus be considered as the generator of rotations.

### A. Rotations properties

- Identity:  $\exists \mathbb{1} : \mathbb{1}\mathcal{D} = \mathcal{D}\mathbb{1} = \mathcal{D}$
- Closure:  $\mathcal{D}_1\mathcal{D}_2$  is also a rotation  $\mathcal{D}_3$ .
- Inverse:  $\exists$  and inverse such that  $\mathcal{D}\mathcal{D}^{-1} = \mathbb{1}$
- Associativity  $(\mathcal{D}_1\mathcal{D}_2)\mathcal{D}_3 = \mathcal{D}_1(\mathcal{D}_2\mathcal{D}_3)$

### B. Commutation

In analogy with the classical case, we can write the commutation for the infinitesimal rotations:

$$\begin{aligned} & \mathcal{D}_x(\epsilon)\mathcal{D}_y(\epsilon) - \mathcal{D}_y(\epsilon)\mathcal{D}_x(\epsilon) \\ &= (\mathbb{1} - iJ_x\epsilon - \frac{1}{2}J_x^2\epsilon^2)(\mathbb{1} - iJ_y\epsilon - \frac{1}{2}J_y^2\epsilon^2) - (\mathbb{1} - iJ_y\epsilon - \frac{1}{2}J_y^2\epsilon^2)(\mathbb{1} - iJ_x\epsilon - \frac{1}{2}J_x^2\epsilon^2) = -(J_xJ_y - J_yJ_x)\epsilon^2 + O(\epsilon^3) \end{aligned}$$

and equate this to  $\mathcal{D}_z(\epsilon^2) - \mathbb{1} = -iJ_z\epsilon^2$ . With this analogy we justify the definition of angular momentum operators as operators that generate the rotations and obey the commutation relationships:

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k$$

### C. Spin- $\frac{1}{2}$

Although angular momentum operators have some classical analogy, they are more general, as they describe for example physical properties that have no classical counterparts, such as the spin. In particular, the lowest dimension in which the commutation relationships above hold is 2. The angular momentum  $\mathbf{S}$  for a TLS is represented by the operators:

$$\begin{aligned} S_x &= \frac{1}{2}\sigma_x = \frac{1}{2}(|0\rangle\langle 1| + |1\rangle\langle 0|) \\ S_y &= \frac{i}{2}\sigma_y = \frac{i}{2}(|0\rangle\langle 1| - |1\rangle\langle 0|) \\ S_z &= \frac{1}{2}\sigma_z = \frac{1}{2}(|0\rangle\langle 0| - |1\rangle\langle 1|) \end{aligned}$$

where  $\{\sigma_x, \sigma_y, \sigma_z\}$  are called Pauli operators or Pauli matrices. The Pauli matrices have the following properties:

1.  $\sigma_\alpha^2 = \mathbb{1}$
2.  $\sigma_i\sigma_j + \sigma_j\sigma_i = 0$ , that is, they anticommute.
3.  $\sigma_i\sigma_j = -\sigma_j\sigma_i = i\sigma_k$  (from the previous property)
4. Hermiticity:  $\sigma_i^\dagger = \sigma_i$
5. Zero trace:  $\text{Tr}\{\sigma_i\} = 0$
6. Determinant  $\det(\sigma_i) = -1$ .

---

**? Question:** Show that  $S$  satisfies the commutation relationship.

1. Show it by multiplying the operators.
  2. Write down the matrix form and perform matrix multiplications. □
- 

We can now check what is the action of the spin operators on the TLS state vector  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ :

$$\begin{aligned} \sigma_x|\psi\rangle &= \alpha|1\rangle + \beta|0\rangle \\ \sigma_y|\psi\rangle &= i\alpha|1\rangle - i\beta|0\rangle \\ \sigma_z|\psi\rangle &= \alpha|0\rangle - \beta|1\rangle \end{aligned}$$

in particular,  $\sigma_x$  swap the two components (spin flip) and  $\sigma_z$  invert the sign of the  $|1\rangle$  component (phase shift), while  $\sigma_y$  does both.

## D. Spin- $\frac{1}{2}$ rotations

We can now look at rotations of spin- $\frac{1}{2}$ . In particular we want to calculate  $\mathcal{D}_\alpha(\varphi) = e^{-iS_\alpha\varphi}$ . For this we remember the property:  $\sigma_\alpha^2 = \mathbb{1}$ . With this, and using a Taylor expansion it is easy to show that we have

$$e^{-iS_\alpha\varphi} = \cos(\varphi/2)\mathbb{1} - i \sin(\varphi/2)S_\alpha$$

---

**? Question:** Calculate the exponential.

From  $(\sigma \cdot n)^2 = (\sigma_x n_x + \sigma_y n_y + \sigma_z n_z)^2 = \sigma_x^2 n_x^2 + n_x n_y (\sigma_x \sigma_y + \sigma_y \sigma_x) + \dots + n_y^2 \mathbb{1} + n_z^2 \mathbb{1} = \mathbb{1}$  and the Taylor expansion we obtain  $e^{-i\varphi\sigma \cdot n} = \mathbb{1} \sum_{n \text{ even}} (-i\varphi)^n/n! + \sigma \cdot n \sum_{n \text{ odd}} (-i\varphi)^n/n! = \mathbb{1} \cos \varphi + \sigma \cdot n \sin \varphi$ .  $\square$

---

### 4.2.3 Example of Two-Level System: Neutron Interferometry

Now we can revisit the TLS examples we have seen earlier. In particular we notice that the polarization rotator is represented by rotation operators, in particular rotations around the x-axis  $e^{-i\theta S_x}$ . Consider another very simple system, a neutron interferometer, such as the Mach-Zehnder interferometer.

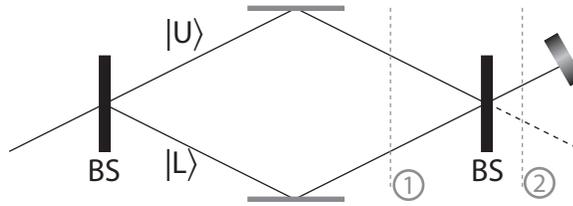


Fig. 3: Neutron Interferometer

We send in a beam of neutrons. The first beamsplitter divides the neutron flux into two parts, that will go into the upper arm or the lower arm. Thus the state of the system is at this point in time

$$|\psi\rangle_1 = \alpha |U\rangle + \beta |L\rangle, \quad \alpha^2 + \beta^2 = 1$$

We assume that the flux of neutrons is so low (neutrons can be very slow) so that only one neutron is present at any time inside the interferometer. The lower and upper beams are then reflected at the mirrors and recombined at the second beam splitter, after which the neutron flux is measured at one arm. If we assume that both beamsplitter works in the same way, delivering an equal flux to each arm (that is, the transmission and reflection are the same), then we have  $|\psi\rangle_1 = (|U\rangle + |L\rangle)/\sqrt{2}$  and  $|\psi\rangle_2 = |U\rangle$ .

---

**? Question:** What is the propagator describing the action of the Beamsplitter?

$U_{BS}|U\rangle = (|U\rangle + |L\rangle)/\sqrt{2}$  and we also know that  $U_{BS}(|U\rangle + |L\rangle)/\sqrt{2} = |U\rangle$ . We can verify that

$$U_{BS} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

performs as we want. In particular, notice that  $U_{BS}U_{BS} = \mathbb{1}$ .

Thus, if our observable is the number of neutron in the upper arm, the measurement always returns 1 with certainty (probability =1).

Let's now consider the case in which we want to measure at point 1 how many neutrons are in the upper arm. The observable is just the projector onto the upper arm  $Ob = P_U = |U\rangle\langle U|$  and we will detect one neutron (or zero neutrons) with probability  $\frac{1}{2}$ . In fact  $p(U) = \|\langle\psi|U\rangle\|^2 = \frac{1}{2}$ . Also, the average value of the number of neutron in the upper arm is  $1/2$  as well, since

$\langle Ob \rangle = |\langle \psi | P_U | \psi \rangle| = \frac{1}{2}$  After the measurement, the state is projected onto the upper arm, if we did detect a photon, or the lower arm, otherwise. We assume that the neutron is free to continue on its path after the measurement and we perform a second measurement after the second beamsplitter.  $\square$

**? Question:** What is the probability of measuring 1 neutron in the upper part in this case?

Now the state at 2 is  $|\psi\rangle_2 = (|U\rangle \pm |L\rangle)/\sqrt{2}$ , hence  $p(U) = \langle \psi | U \rangle = \frac{1}{2}$ .  $\square$

#### 4.2.4 Spinor behavior

By calculating  $e^{-iS_\alpha\varphi} S_\beta e^{iS_\alpha\varphi}$  we see that the rotations of the operator give the following result:

$$S_x \xrightarrow{S_z} S_x \cos(\varphi) - S_y \sin(\varphi),$$

$$S_y \xrightarrow{S_z} S_y \cos(\varphi) + S_x \sin(\varphi),$$

$$S_z \xrightarrow{S_z} S_z$$

These are the same rotation rules we would have expected classically. In particular, taking the expectation values, we see that they correspond exactly to the rotations in 3D of a vector, with a periodicity of  $2\pi$ . Things are a bit different (and more surprising) if we consider instead the state rotation. Consider the rotation of the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  with respect to  $S_z$ :

$$e^{-iJ_z\varphi}|\psi\rangle = e^{-i\varphi/2}\alpha|0\rangle + e^{i\varphi/2}\beta|1\rangle,$$

now the angle of rotation seems to be  $\varphi/2$ . This has an interesting consequence: if we rotate by  $\varphi = 2\pi$  instead of returning to the initial state, as we would have expected, we obtain  $e^{-iJ_z2\pi}|\psi\rangle = -|\psi\rangle$ . This is the so-called **spinor** behavior. Notice that from a simple measurement this minus sign (which is equivalent to a *global* phase) is irrelevant, hence we obtain the same expectation values for the angular momenta as before. We will see in P-Set on that experiments can be devised to show the spinor behavior (but they need to use more than one spin).

#### 4.2.5 The SU(2) and SO(3) groups

A group  $G$  is a finite or infinite set of elements together with a binary operation (called the group operation) that together satisfy the four fundamental properties of closure, associativity, the identity property, and the inverse property. A rotation group is a group in which the elements are orthogonal matrices with determinant 1. In the case of three-dimensional space, the rotation group is known as the special orthogonal group or  $SO(3)$ <sup>9</sup>. The special unitary group  $SU(2)$  is the set of 2 by 2 unitary matrices with determinant +1 [it is a subgroup of the unitary group  $U(2)$ ]. The two groups  $SO(3)$  and  $SU(2)$  both represent rotations, however there is a one-to-two correspondence for a given  $R \in SO(3)$  there are 2  $U \in SU(2)$ . This is because a  $2\pi$  and  $4\pi$  rotations are the same in  $SO(3)$  but they are  $\mathbb{1}$  and  $-\mathbb{1}$  in  $SU(2)$ .

<sup>9</sup> For a more rigorous and extensive explanation see J.J. Sakurai “Modern Quantum Mechanics”, Addison-Wesley (1994), page 168

## 5. Time evolution

### 5.1 The Schrödinger and Heisenberg pictures

### 5.2 Interaction Picture

5.2.1 Dyson Time-ordering operator

5.2.2 Some useful approximate formulas

### 5.3 Spin- $\frac{1}{2}$ precession

### 5.4 Examples: Resonance of a Two-Level System

5.4.1 Dressed states and AC Stark shift

### 5.5 The wave-function

5.5.1 Position representation

5.5.2 Momentum representation

5.5.3 Schrödinger equation for the wavefunction

### 5.6 Feynman's path-integral

In a previous lecture we characterized the time evolution of closed quantum systems as *unitary*,  $|\psi(t)\rangle = U(t, 0) |\psi(0)\rangle$  and the state evolution as given by Schrödinger equation:

$$i\hbar \frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

Equivalently, we can find a differential equation for the dynamics of the propagator:

$$i\hbar \frac{\partial U}{\partial t} = \mathcal{H}U$$

This equation is valid also when the Hamiltonian is time-dependent.

As the Hamiltonian represents the energy of the system, its *spectral representation* is defined in terms of the energy eigenvalues  $\epsilon_k$ , with corresponding eigenvectors  $|k\rangle$ :  $\mathcal{H} = \sum_k \epsilon_k |k\rangle\langle k|$ . The evolution operator is then:  $U = \sum_k e^{-i\epsilon_k t} |k\rangle\langle k|$ . The eigenvalues of  $U$  are therefore simply  $e^{-i\epsilon_k t}$ , and it is common to talk in terms of eigenphases  $\varphi_k(t) = \epsilon_k t$ . If the Hamiltonian is time-independent we have also  $U^\dagger = U(-t)$ , it is possible to obtain an effective inversion of the time arrow.

---

**? Question:** What is the evolution of an energy eigenvector  $|k\rangle$ ?

First consider the infinitesimal evolution:  $|k(t+dt)\rangle = U(t+dt, t) |k(t)\rangle = (\mathbb{1} - i\mathcal{H}dt) |k(t)\rangle = (1 - i\epsilon_k dt) |k(t)\rangle$ . Thus we have the differential equation for the energy eigenket:  $\frac{d|k\rangle}{dt} = -i\epsilon_k |k\rangle$ , so that  $|k(t)\rangle = e^{-i\epsilon_k t} |k(0)\rangle$ .

We can also use the spectral decomposition of  $U$ :  $|k(t)\rangle = U(t, 0) |k(0)\rangle = (\sum_h e^{-i\epsilon_h t} |h\rangle\langle h|) |k(0)\rangle = e^{-i\epsilon_k t} |k(0)\rangle$ . □

---

Notice that if a system is in a state given by an eigenvector of the Hamiltonian, then the system does not evolve. This is because the state will only acquire a global phase that, as seen, does not change its properties. Of course, superposition of energy eigenkets do evolve.

### 5.1 The Schrödinger and Heisenberg pictures

Until now we described the dynamics of quantum mechanics by looking at the time evolution of the state vectors. This approach to quantum dynamics is called the Schrödinger picture. We can easily see that the evolution of the

state vector leads to an evolution for the expectation values of the observables (which are the relevant physical quantities we are interested in and have access to).

From the evolution law for a state,  $|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle$ , we obtain the following relation, when expressing the state in the Hamiltonian eigenbasis:

$$|\psi\rangle = \sum_k c_k |\epsilon_k\rangle \rightarrow |\psi'\rangle = e^{-i\mathcal{H}t}|\psi\rangle = \sum_k c_k e^{-i\epsilon_k t} |\epsilon_k\rangle$$

Then the expectation value of an observable  $A$  evolves as:

$$\langle A \rangle = \sum_{k,j} c_k^* c_j \langle \epsilon_k | A | \epsilon_j \rangle \rightarrow \sum_{k,j} c_k^* c_j \langle \epsilon_k | A | \epsilon_j \rangle e^{-i(\epsilon_j - \epsilon_k)t}$$

Quite generally, we can also write  $\langle A(t) \rangle = \langle \psi(t) | A | \psi(t) \rangle = \langle (U\psi) | A | U\psi \rangle$ . By the associative property we then write  $\langle A(t) \rangle = \langle \psi | (U^\dagger A U) | \psi \rangle$ .

It would then seem natural to define an "evolved" observable  $A(t) = U^\dagger A U$ , from which we can obtain expectation values considering states that are fixed in time,  $|\psi\rangle$ . This is an approach known as **Heisenberg picture**.

Observables in the Heisenberg picture are defined in terms of observables in the Schrödinger picture as

$$A^H(t) = U^\dagger(t) A^S U(t), \quad A^H(0) = A^S$$

The state kets coincide at  $t = 0$ :  $|\psi\rangle_H = |\psi(t=0)\rangle_S$  and they remain independent of time. Analogously to the Schrödinger equation we can define the Heisenberg equation of motion for the observables:

$$\frac{dA^H}{dt} = -i[A^H, \mathcal{H}]$$

---

**? Question:** Derive the Heisenberg equation from the Schrödinger equation.

$\frac{dA^H}{dt} = \frac{d(U^\dagger A^S U)}{dt} = \frac{\partial U^\dagger}{\partial t} A^S U + U^\dagger A^S \frac{\partial U}{\partial t} = i(U^\dagger \mathcal{H}) A^S U + U^\dagger A^S (-i\mathcal{H}U)$ . Inserting the identity  $\mathbb{1} = U U^\dagger$  we have  $= i(U^\dagger \mathcal{H} U U^\dagger A^S U - U^\dagger A^S U U^\dagger \mathcal{H} U)$ . We define  $\mathcal{H}^H = U^\dagger \mathcal{H} U$ . Then we obtain  $\frac{dA^H}{dt} = -i[A^H, \mathcal{H}^H]$ .  $U$  and  $\mathcal{H}$  always commute for time-independent  $\mathcal{H}$ , thus  $\mathcal{H}^H = \mathcal{H}$ . □

---

## 5.2 Interaction Picture

We now consider yet another "picture" that simplifies the description of the system evolution in some special cases. In particular, we consider a system with an Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + V$$

where  $\mathcal{H}_0$  is a "solvable" Hamiltonian (of which we already know the eigen-decomposition, so that it is easy to calculate e.g.  $U_0 = e^{-i\mathcal{H}_0 t}$ ) and  $V$  is a perturbation that drives an interesting (although unknown) dynamics. In the so-called **interaction picture** the state is represented by

$$|\psi\rangle_I = U_0(t)^\dagger |\psi\rangle_S = e^{i\mathcal{H}_0 t} |\psi\rangle_S$$

where the subscript  $I, S$  indicate the interaction and Schrödinger picture respectively. For the observable operators we can define the corresponding interaction picture operators as:

$$A_I(t) = U_0^\dagger A_S U_0 \rightarrow V_I(t) = U_0^\dagger V U_0$$

We can now derive the differential equation governing the evolution of the state in the interaction picture (we now drop the subscript  $S$  for the usual Schrödinger picture):

$$i \frac{\partial |\psi\rangle_I}{\partial t} = i \frac{\partial (U_0^\dagger |\psi\rangle)}{\partial t} = i \left( \frac{\partial U_0^\dagger}{\partial t} |\psi\rangle + U_0^\dagger \frac{\partial |\psi\rangle}{\partial t} \right) = -U_0^\dagger \mathcal{H}_0 |\psi\rangle + U_0^\dagger (\mathcal{H}_0 + V) |\psi\rangle = U_0^\dagger V |\psi\rangle.$$

Inserting the identity  $\mathbb{1} = U_0 U_0^\dagger$ , we obtain

$$i \frac{\partial |\psi\rangle_I}{\partial t} = U_0^\dagger V U_0 U_0^\dagger |\psi\rangle = V_I |\psi\rangle_I.$$

This is a Schrödinger-like equation for the vector in the interaction picture, evolving under the action of the operator  $V_I$  only. However, in contrast to the usual Schrödinger picture, even the observables in the interaction picture evolve in time. From their definition  $A_I(t) = U_0^\dagger A_S U_0$ , we have the differential equation  $\frac{dA_I}{dt} = i[\mathcal{H}_0, A_I]$ , which is an Heisenberg-like equation for the observable, with the total Hamiltonian replaced by  $\mathcal{H}_0$ . The interaction picture is thus an intermediate picture between the two other pictures.

	<b>S</b>	<b>H</b>	<b>I</b>
$ \psi\rangle$	✓	×	✓
$A$	×	✓	✓

**Table 1:** Time dependence of states and operators in the three pictures

### 5.2.1 Dyson Time-ordering operator

If we now want to solve the state-vector differential equation in terms of a propagator  $|\psi(t)\rangle_I = U_I(t) |\psi\rangle_I$ , we encounter the problem that the operator  $V_I$  is usually time-dependent since  $V_I(t) = U_0^\dagger V U_0$ , thus in general  $U_I \neq e^{-iV_I t}$ . We can still write an equation for the propagator in the interaction picture

$$i \frac{dU_I}{dt} = V_I(t) U_I$$

with initial condition  $U_I(0) = \mathbb{1}$ . When  $V_I$  is time dependent and  $V_I(t)$  does not commute at different time, it is no longer possible to find a simple explicit expression for  $U_I(t)$ . Indeed we could be tempted to write  $U_I(t) = e^{-i \int_0^t V_I(t') dt'}$ . However in general

$$e^A e^B \neq e^{A+B} \quad \text{if } [A, B] \neq 0,$$

thus for example, although we know that  $U_I(t)$  can be written as  $U_I(t, 0) = U_I(t, t^*) U_I(t^*, 0)$  ( $\forall 0 < t^* < t$ ) we have that  $e^{-i \int_0^{t^*} V_I(t') dt'} e^{-i \int_{t^*}^t V_I(t') dt'} \neq e^{-i \int_0^t V_I(t') dt'}$ . Thus we cannot find an explicit solution in terms of an integral.

We can however find approximate solutions or formal solution to the evolution.

The differential equation is equivalent to the integral equation

$$U_I(t) = \mathbb{1} - i \int_0^t V_I(t') U_I(t') dt'$$

By iterating, we can find a formal solution to this equation :

$$\begin{aligned} U_I(t) = & \mathbb{1} - i \int_0^t dt' V_I(t') + (-i)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') + \dots \\ & + (-i)^n \int_0^t dt' \dots \int_0^{t^{(n-1)}} dt^{(n)} V_I(t') \dots V_I(t^{(n)}) + \dots \end{aligned}$$

This series is called the *Dyson* series.

Note that in the expansion the operators are time-ordered, so that in the product the operators at earlier times are at the left of operators at later times. We then define an operator  $\mathcal{T}$  such that when applied to a product of two operators it will return their time-ordered product:

$$\mathcal{T}(A(t)B(t')) = \begin{cases} A(t)B(t'), & \text{if } t < t' \\ B(t')A(t), & \text{if } t' < t \end{cases}$$

Now we can rewrite the expression above in a more compact way. We replace the limits of each intervals so that they span the whole duration  $\{0, t\}$  and we divide by  $n!$  to take into account that we integrate over a larger interval. Then we can write the products of integrals as powers and use the time-ordering operator to take this change into account. We then have:

$$U_I(t) = \mathcal{T} \left\{ \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \left( \int_0^t dt' V_I(t') \right)^n \right\}$$

where we recognize the expression for an exponential

$$U_I(t) = \mathcal{T} \left\{ \exp \left( -i \int_0^t dt' V_I(t') \right) \right\}$$

Note that the time-ordering operator is essential for this expression to be correct.

**? Question:** Prove that  $\int_0^t dt' \dots \int_0^{t^{(n-1)}} dt^{(n)} V_I(t') \dots V_I(t^{(n)}) = \frac{1}{n!} \mathcal{T} \left\{ \left( \int_0^t dt' V_I(t') \right)^n \right\}$  for  $n = 2$ .

### 5.2.2 Some useful approximate formulas

Besides the formal solution found above and the Dyson series formula, there are other approximate formulas that can help in calculating approximations to the time evolution propagator.

#### A. Baker-Campbell-Hausdorff formula

The Baker-Campbell-Hausdorff formula gives an expression for  $C = \log(e^A e^B)$ , when  $A, B$  do not commute. That is, we want  $C$  such that  $e^C = e^A e^B$ . We have<sup>10</sup>

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] - [B, [A, B]]) - \frac{1}{24}[B, [A, [A, B]]] \dots$$

The **Hadamard** series is the solution to  $f(s) = e^{sA} B e^{-sA}$ . To find this, differentiate the equation:

$$f'(s) = e^{sA} A B e^{-sA} - e^{sA} B A e^{-sA} = e^{sA} [A, B] e^{-sA}$$

$$f''(s) = e^{sA} A [A, B] e^{-sA} - e^{sA} [A, B] A e^{-sA} = e^{sA} [A, [A, B]] e^{-sA}$$

$$f'''(s) = e^{sA} [A, [A, [A, B]]] e^{-sA}$$

etc. and then construct the Taylor series for  $f(s)$ :

$$f(s) = f(0) + s f'(0) + \frac{1}{2} s^2 f''(0) + \frac{1}{3!} s^3 f'''(0) + \dots$$

to obtain

$$e^{sA} B e^{-sA} = B + [A, B] s + \frac{1}{2} [A, [A, B]] s^2 + \frac{1}{3!} [A, [A, [A, B]]] s^3 + \dots$$

With  $s = it$  and  $A = \mathcal{H}$ , this formula can be useful in calculating the evolution of an operator (either in the Heisenberg or interaction representation or for the density operator).

<sup>10</sup> See e.g. wikipedia for more terms and [mathworld](#) for calculating the series.

## B. Suzuki-Trotter expansion

Another useful approximation is the Suzuki-Trotter expansion<sup>11</sup>. To first order this reads:

$$e^{A+B} = \lim_{n \rightarrow \infty} (e^{A/n} e^{B/n})^n$$

Suzuki-Trotter expansion of the second order:

$$e^{A+B} = \lim_{n \rightarrow \infty} (e^{A/(2n)} e^{B/n} e^{A/(2n)})^n$$

In general we can approximate the evolution under a time-varying Hamiltonian by a piecewise constant Hamiltonian in small enough time intervals:

$$U(t, t_0) = U(t, t_{n-1}) \dots U(t_2, t_1) U(t_1, t_0), \quad t_0 < t_1 < t_2 < \dots < t_{n-1} < t,$$

where we usually take  $t_k - t_{k-1} = \delta t$  and consider the Hamiltonian  $\mathcal{H}$  to be constant during each of the small time interval  $\delta t$ .

## C. Magnus expansion

The Magnus expansion is a perturbative solution to the exponential of a time-varying operator (for example the propagator of a time-varying Hamiltonian). The idea is to define an effective time-independent Hamiltonian by taking:  $U = \mathcal{T} e^{-i \int_0^t dt' \mathcal{H}(t')} \equiv e^{-it\bar{H}}$ . The effective Hamiltonian is then expanded in a series of terms of increasing order in time  $\bar{H} = \bar{H}^{(0)} + \bar{H}^{(1)} + \bar{H}^{(2)} + \dots$ , so that

$$U = \exp\{-it[\bar{H}^{(0)} + \bar{H}^{(1)} + \bar{H}^{(2)} + \dots]\}$$

where the terms can be found by expanding  $\mathcal{T} e^{-i \int_0^t dt' \mathcal{H}(t')}$  and equating terms of the same time power. In order to keep the time order, commutators are then introduced. The lowest order terms are

$$\begin{aligned} \bar{H}^{(0)} &= \frac{1}{t} \int_0^t H(t') dt' \\ \bar{H}^{(1)} &= -\frac{i}{2t} \int_0^t dt' \int_0^{t'} dt'' [H(t'), H(t'')] \\ \bar{H}^{(2)} &= \frac{1}{6t} \int_0^t dt' \int_0^{t'} dt'' \int_0^{t''} dt''' \{ [[H(t'), H(t'')], H(t''')] + [[H(t'''), H(t'')], H(t')] \} \end{aligned}$$

The convergence of the expansion is ensured only if  $\|\mathcal{H}\|t \ll 1$ .

---

<sup>11</sup> See: M. Suzuki, *Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*, *Comm. Math. Phys.* **51**, 183-190 (1976)

### 5.3 Spin- $\frac{1}{2}$ precession

We consider the semi-classical problem of a spin-1/2 particle in a classical magnetic field. To each spin with spin angular momentum  $\mathbf{J}$  is associated a magnetic moment  $\boldsymbol{\mu} = \gamma\mathbf{S}$  where  $\gamma$  is called the gyromagnetic ratio, a property of each spin-carrying particle (nucleus, electron, etc.). The energy of the system in an external magnetic field is (classically) given by  $\boldsymbol{\mu} \cdot \mathbf{B}$ , where  $B$  is of course the field. Thus, the system Hamiltonian is simply  $\mathcal{H} = \gamma B_z S_z = \omega S_z$ , where we take the  $z$  axis to point along the external field for simplicity and we defined the Larmor frequency for the given system.

If the spin is initially in the state  $|0\rangle$ , the system does not evolve (as it is an eigenstate of the Hamiltonian). If instead it is prepared in a superposition state, it will undergo an evolution.

$$|\psi_0\rangle = \alpha_0|0\rangle + \beta_0|1\rangle \rightarrow |\psi(t)\rangle = \alpha(t)|0\rangle + \beta(t)|1\rangle$$

? **Question:** What are the functions  $\alpha(t)$ ,  $\beta(t)$ ?

- As  $|0\rangle$ ,  $|1\rangle$  are eigenstates of the Hamiltonian with eigenvalues  $\pm\omega/2$ , we know that their evolution is just a phase  $e^{\pm i\omega t/2}$ , so that  $\alpha(t) = \alpha_0 e^{-i\omega t/2}$  and  $\beta(t) = \beta_0 e^{+i\omega t/2}$ .
- $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ , with  $U = e^{-i\mathcal{H}t} = e^{-i\omega S_z t} = \mathbf{1} \cos(\omega t/2) - i \sin(\omega t/2) 2S_z$ . Then  $U(t)|0\rangle = (\cos \omega t/2 - i \sin \omega t/2)|0\rangle = e^{-i\omega t/2}|0\rangle$  and we find the same result.  $\square$

? **Question:** What is the probability of finding the spin back to its initial state?

Let's write the initial state as  $|\psi\rangle_0 = \cos(\vartheta/2)|0\rangle + e^{i\varphi/2} \sin(\vartheta/2)|1\rangle$ . Then the evolution is  $e^{i\omega t/2} \cos(\vartheta/2)|0\rangle + e^{i(\omega t+\varphi)/2} \sin(\vartheta/2)|1\rangle$  and the probability  $p = \cos^2(\omega t/2) + \cos^2 \vartheta^2 \sin^2(\omega t/2)$ . In particular, for  $\vartheta = \pi/4$  we have  $\cos^2(\omega t/2)$  (notice that this is an eigenstate of the  $S_x$  operator).  $\square$

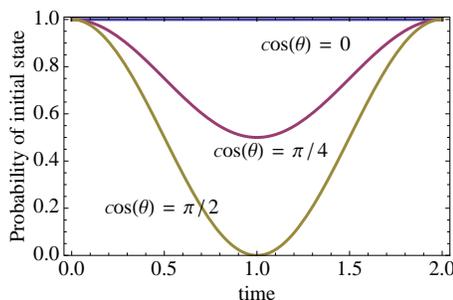


Fig. 4: Spin precession: probability of being in the initial state

? **Question:** What is the evolution of the magnetization in the x direction?

We want to calculate  $\langle S_x(t) \rangle$ . We can use the Heisenberg picture, and calculate  $U^\dagger S_x U = S_x \cos(\omega t) - S_y \sin(\omega t)$ . Thus we see that the periodicity is  $T = \frac{2\pi}{\omega}$  while it was  $\frac{4\pi}{\omega}$  for the spin state (spinor behavior). Then we know that  $\langle S_x \rangle = \cos(\varphi/2) \sin(\vartheta)$  and  $\langle S_y \rangle = \sin(\varphi/2) \sin(\vartheta)$  from which we find  $\langle S_x(t) \rangle = \cos(\varphi/2 + \omega t) \sin(\vartheta)$   $\square$

### Nuclear Magnetic Resonance

The evolution of the magnetization is what is usually detected in NMR. The precession of the spin in a large static magnetic field creates an oscillating magnetic field that in turns generate a current/voltage in a pickup coil. Fourier-transform of the signal gives spectroscopic information regarding the Larmor frequency; local modification of the magnetic field (due e.g. to electronic field) induces a variation on the Larmor frequency of each nuclear spin in a molecule, thus providing a way to investigate the structure of the molecule itself. Before we can have a full vision of a (simple) NMR experiment, we still need to answer the question on how we first prepare the spin in a superposition state (e.g. in a  $S_x$  eigenstate). We will be able to answer this question in a minute.

## 5.4 Examples: Resonance of a Two-Level System

We have looked at the precession of the spin at the Larmor frequency, which happens if the spin is initially in a superposition state. However, the question remained on how we rotate initially the spin away from its equilibrium state pointing along the large external magnetic field. Consider then a more general problem in which we add a (small) time-dependent magnetic field along the transverse direction (e.g. x-axis):

$$\vec{B}(t) = B_z \hat{z} + 2B_1 \cos(\omega t) \hat{x} = B_z \hat{z} + B_1 [(\cos(\omega t) \hat{x} + \sin(\omega t) \hat{y}) + (\cos(\omega t) \hat{x} - \sin(\omega t) \hat{y})],$$

where  $B_1$  is the strength of the radio-frequency (for nuclei) or microwave (for electron) field. The Hamiltonian of the system  $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1(t) + \mathcal{H}'_1(t)$  is then:

$$\mathcal{H} = \frac{\omega_0}{2} \sigma_z + \frac{\omega_1}{2} [\cos(\omega t) \sigma_x + \sin(\omega t) \sigma_y] + \frac{\omega_1}{2} [\cos(\omega t) \sigma_x - \sin(\omega t) \sigma_y],$$

where we defined the rf frequency  $\omega_1$ . We already know the eigenstates of  $\mathcal{H}_0$  ( $|0\rangle$  and  $|1\rangle$ ). Thus we use the interaction picture to simplify the Hamiltonian, with  $U_0 = e^{-i\omega\sigma_z/2}$  defining a frame rotating about the  $z$ -axis at a frequency  $\omega$ : this is the so-called *rotating frame*. Remembering that  $U_0 \sigma_x U_0^\dagger = \cos(\omega t) \sigma_x + \sin(\omega t) \sigma_y$ , it's easy to see that the perturbation Hamiltonian in the interaction frame is  $\mathcal{H}_{1I} = U_0^\dagger \mathcal{H}_1 U_0 = \frac{\omega_1}{2} \sigma_x$ . We also have  $\mathcal{H}'_{1I} = U_0^\dagger \mathcal{H}'_1 U_0 = \frac{\omega_1}{2} (\cos(2\omega t) \sigma_x - \sin(2\omega t) \sigma_y)$ . Under the assumptions that  $\omega_1 \ll \omega$ , this is a small, fast oscillating term, that quickly averages out during the evolution of the system and thus can be neglected. This approximation is called the *rotating wave approximation* (RWA). Under the RWA, the Hamiltonian in the rotating frame simplifies to

$$\mathcal{H}_I = \frac{\Delta\omega}{2} \sigma_z + \frac{\omega_1}{2} \sigma_x$$

where  $\Delta\omega = \omega_0 - \omega$ . Notice that if  $\Delta\omega$  is large ( $\gg \omega_1$ ), we expect that the eigenstates of the systems are still going to be close to the eigenstates of  $\mathcal{H}_0$  and the small perturbation has almost no effect. Only when  $\omega \approx \omega_0$  we will see a change: this is the resonance condition. In particular, for  $\Delta\omega = 0$  the new Hamiltonian  $\sim \sigma_x$  will cause a spin initially in, say,  $|0\rangle$  to rotate away from the  $z$  axis and toward the  $y$  axis. This is how a "pulse" is generated e.g. in NMR or ESR pulsed spectroscopy. For example, if the  $B_1$  field is turned on for a time  $t_{\pi/2} = \pi/2\omega_1$  we prepare the state  $|\psi\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$  that will then precess at the Larmor frequency, giving a spectroscopic signature in the recorded signal.

We want to study the Hamiltonian in the general case. Given the matrix representation

$$\mathcal{H}_I = \frac{1}{2} \begin{pmatrix} \Delta\omega & \omega_1 \\ \omega_1 & -\Delta\omega \end{pmatrix}$$

we can find the eigenvalues:

$$\omega_I = \pm \frac{\Delta\omega}{2} \sqrt{1 + (\omega_1/\Delta\omega)^2}.$$

There are two interesting limits, on resonance ( $\Delta\omega = 0$ ) where  $\omega_I = \omega_1$  and far off resonance ( $\Delta\omega \gg \omega_1$ ) where  $\omega_I \approx \Delta\omega \sim \omega_0$ . The eigenstates are found (e.g. via a rotation of the Hamiltonian) to be

$$\begin{aligned} |+\rangle_I &= \cos\vartheta |0\rangle + \sin\vartheta |1\rangle \\ |-\rangle_I &= \cos\vartheta |1\rangle - \sin\vartheta |0\rangle, \end{aligned}$$

with

$$\sin\vartheta = \sqrt{\frac{\omega_I - \Delta\omega}{2\omega_I}}, \quad \cos\vartheta = \sqrt{\frac{\omega_I + \Delta\omega}{2\omega_I}}$$

Consider the evolution of the state  $|0\rangle$  under the rotating frame Hamiltonian. At time  $t = 0$  the two frame coincide, so  $|\psi\rangle_I = |\psi\rangle = |0\rangle$ . The state then evolves as

$$|\psi(t)\rangle_I = \left[ \cos\left(\frac{\Omega t}{2}\right) - i\frac{\Delta\omega}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \right] |0\rangle - i\frac{\omega_1}{\Omega} \sin\left(\frac{\Omega t}{2}\right) |1\rangle$$

where we defined  $\Omega = \sqrt{\Delta\omega^2 + \omega_1^2}$ . The probability of *flipping* the spin (that is, of finding the spin in the  $|1\rangle$  state) is then  $p(1) = \frac{\omega_1^2}{\Delta\omega^2 + \omega_1^2} \sin^2\left(\frac{\Omega t}{2}\right)$ . Notice that only if  $\Delta\omega = 0$  we can have perfect inversion (i.e.  $p(1) = 1$  for  $t = \pi/\omega_1$ ). Notice that we have defined all the evolutions as in the rotating frame.

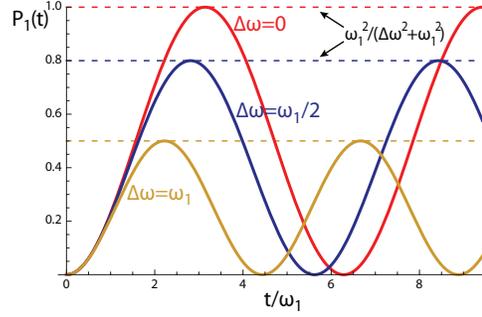


Fig. 5: Rabi oscillation. Probability of being in the  $|1\rangle$  state for different values of the ratio  $\omega_1/\Delta\omega$

### 5.4.1 Dressed states and AC Stark shift

This Hamiltonian is also used in Atomic physics to describe the ground and (one) excited levels coupled by an external e.m. field (for example in the visible spectrum). The evolution of an atom in an e.m. field (here we are considering a classical e.m. field, but we will see that we can also consider the quantized version) is usually described with the *dressed atom* picture. This picture (due to Cohen-Tannoudji) describes the atom as *dressed* by a cloud of virtual photons, with which it interacts.

This atomic TLS has (unperturbed) eigenstates  $|e\rangle = |0\rangle$  and  $|g\rangle = |1\rangle$  with energies  $E_0 - E_1 = \Delta\omega$ , which are coupled through an interaction  $\omega_1/2$ . When we consider the optical transition of an atom we usually call  $\omega_1$  the Rabi frequency.

The coupling mixes these states, giving two new eigenstates as seen before with energies  $\pm\omega_I = \pm\frac{\Delta\omega}{2}\sqrt{1 + (\omega_1/\Delta\omega)^2}$ , which is called the effective Rabi frequency.

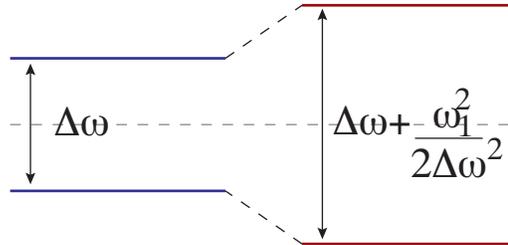


Fig. 6: Energy shift for small coupling perturbation

If the coupling is small, we can treat it as a perturbation, and the energies are just shifted by an amount  $\delta E = \frac{\omega_1^2}{4\Delta\omega}$ .

That is, the new energies are  $E'_0 = \frac{\Delta\omega}{2}(1 + \frac{\omega_1^2}{2\Delta\omega^2})$ . This shift in the context of a two-level atom dressed by the e.m. field is called the *AC Stark shift*. It is a quadratic effect that can be seen also as arising (in a more general context) from second order perturbation theory.

The perturbed energies are shown in the following diagram. Here we explore the range of the eigenvalues  $\pm\omega_I =$  found before, given a fixed value of the coupling  $\omega_1$  and a varying splitting  $\Delta\omega$  between the two levels. In red are the two perturbed energies, while the dashed lines follow the unperturbed energies. For  $\Delta\omega = 0$ , in the absence of a coupling term, the two eigenstate are degenerate. The perturbation lifts this degeneracy, giving rise to an *avoided crossing*. The eigenstates are a complete mix of the unperturbed states, yet remain split in energy by the strength of interaction  $\omega_1$ .

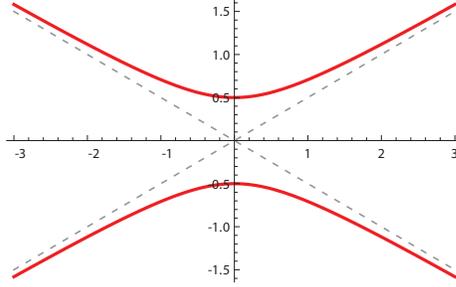


Fig. 7: Dressed atom energies as a function of the splitting  $\Delta\omega$  showing the avoided crossing

## 5.5 The wave-function

We have so far considered systems associated to observables with a discrete spectrum. That is, the system can assume only a discrete number of states (for example 2, for the TLS) and the possible outcomes of an experiment are a discrete set of values. Although for the first part of the class this is all that we'll need, it's important to introduce as well systems with a continuous set of states, as they lead to the concept of a particle's wave function<sup>12</sup>. This is an essential concept in non-relativistic QM that you might have seen before (and probably as one of the very first topics in QM).

### 5.5.1 Position representation

The state  $|\psi\rangle$  of a point-like particle is naturally expanded onto the basis made of the eigenstates of the particle's position vector operator  $\mathbf{R}$ . Of course the position of a point particle is a continuous variable (more precisely a vector whose components are the three commuting coordinate operators X, Y and Z). The rigorous mathematical definition of these continuous basis states is somewhat complex, so we will skip some of the details to instead obtain a practical description of the wave function. The basis states  $|\mathbf{r}\rangle$  satisfy the relations generalizing the orthonormality conditions:

$$\langle \mathbf{r} | \mathbf{r}' \rangle = \delta(\mathbf{r} - \mathbf{r}'), \quad \int d^3\mathbf{r} |\mathbf{r}\rangle \langle \mathbf{r}| = \mathbb{1}$$

where  $\delta(\mathbf{r} - \mathbf{r}')$  is the three-dimensional Dirac function. Developing  $|\psi\rangle$  in the  $|\mathbf{r}\rangle$  basis yields:

$$|\psi\rangle = \int d^3\mathbf{r} |\mathbf{r}\rangle \langle \mathbf{r} | \psi \rangle$$

where we define the wave function (in the position representation)

$$\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle$$

The shape of the wave function depends upon the physical situation under consideration. We may say that the wave function describes the state of the particle suspended, before measurement, in a continuous superposition of an infinite number of possible positions. Upon measurement of  $\mathbf{R}$  performed with a linear precision  $\delta r$ , this superposition collapses into a small wave packet of volume  $(\delta r)^3$  around a random position  $\mathbf{r}$ , with the probability  $p(\mathbf{r}) = |\langle \mathbf{r} | \psi \rangle|^2 (\delta r)^3$ .

### 5.5.2 Momentum representation

The position representation is suited for measurements of the particle's position. If one is interested in the particle momentum  $\mathbf{P}$  or velocity  $\mathbf{V} = \mathbf{P}/m$  (where  $m$  is the particle mass) it is appropriate to choose the momentum

<sup>12</sup> For a nice introduction to these concepts, see S. Haroche, J.-M. Raimond, *Exploring the quantum: atoms, cavities and photons*, Oxford University Press (2006). In this section we follow their presentation closely.

representation and to expand  $|\psi\rangle$  over the continuous basis of the momentum eigenstates  $|\mathbf{p}\rangle$ :

$$|\psi\rangle = \int d^3\mathbf{p} |\mathbf{p}\rangle \langle \mathbf{p} | \psi \rangle$$

where we define the wave function (in the position representation)

$$\tilde{\psi}(\mathbf{p}) = \langle \mathbf{p} | \psi \rangle$$

A simple system could be describing a single particle with a well defined momentum. The state is then  $|\psi\rangle = |\mathbf{p}\rangle$ . In the momentum representation, we obtain the wave function  $\tilde{\psi}(\mathbf{p}) = \delta(\mathbf{p})$ . We can as well describe this state in the position representation,  $|\mathbf{p}\rangle = \int d^3\mathbf{r} |\mathbf{r}\rangle \langle \mathbf{r} | \mathbf{p} \rangle$ . Following de Broglie's hypothesis which associates to a particle of momentum  $p$  a plane wave of wavelength  $\lambda = h/p$ , the momentum eigenstates are plane waves in the position representation

$$\psi_p(r) = \langle \mathbf{r} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}.$$

We can take this as the definition itself of the momentum eigenstates; from this definition the well-known commutation relationship between position and momentum follow. Otherwise one could state the commutation relationship as an axiom and derive the form of the momentum eigenstates in the position representation.

**? Question:** Show how  $[r_i, p_j] = i\hbar\delta_{ij} \Leftrightarrow \psi_p(r) = \frac{e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}}{(2\pi\hbar)^{3/2}}$

i) Hint: Show that the momentum generates translations in  $x$  and consider an infinitesimal translation.

ii) Hint: Show that  $[P_x, f(x)] = -i\hbar\partial_x f(x)$ .

1) We start from  $\langle p_x | x \rangle = \frac{e^{-ip_x x/\hbar}}{(2\pi\hbar)^{1/2}}$ . Then we have for any translation  $a$

$$\langle p_x | x + a \rangle \propto e^{-ip_x(x+a)/\hbar} = e^{-ip_x a/\hbar} \langle p_x | x \rangle$$

We thus recognized  $p$  as the generator of translation and the corresponding propagator  $U(a) = e^{-ip_x a/\hbar}$ . In the Heisenberg picture, we can thus show  $U(a)^\dagger x U(a) = x + a\mathbb{1}$ , since  $\forall |\psi\rangle$  we have

$$\langle \psi | U^\dagger(a) x U(a) | \psi \rangle = \langle \psi + a | x | \psi + a \rangle = \langle x \rangle + a.$$

Now we consider an infinitesimal translation  $\delta a$ . The propagator then becomes  $U(\delta a) \approx \mathbb{1} - ip_x \delta a/\hbar$ . Calculating again  $U(\delta a)^\dagger x U(\delta a) = x + \delta a\mathbb{1}$ , we obtain:

$$x + \delta a\mathbb{1} = (\mathbb{1} + ip_x \delta a/\hbar)x(\mathbb{1} - ip_x \delta a/\hbar) = x + \frac{i\delta a}{\hbar}(px - xp) + \frac{\delta a^2 p^2}{\hbar^2} = x - \frac{i\delta a}{\hbar}[x, p] + O(\delta a^2)$$

Neglecting terms in  $\delta a^2$  we thus proved the commutation relationship  $[x, p] = i\hbar\mathbb{1}$ . □

2) Now we start from the commutation relationship  $[x, p] = i\hbar$  and we calculate  $[x^n, p]$ . We start from the lower powers:

$$[x^2, p] = x[x, p] + [x, p]x = 2i\hbar x; \quad [x^3, p] = x[x^2, p] + [x, p]x^2 = 3i\hbar x^2; \quad [x^n, p] = ni\hbar x^{n-1}$$

Let's now consider any function of  $x$  and its commutator with  $p$ . Since by definition we can expand the function in a power series, it is easy to calculate the commutator:

$$[f(x), p] = \sum_n f^{(n)}(0)/n! [x^n, p] = \sum_n f^{(n)}(0) \frac{n}{n!} i\hbar x^{n-1} = i\hbar \frac{\partial f(x)}{\partial x}$$

Notice that this is also true for the wave function:  $[\hat{p}_x, \psi_p(x)] = -i\hbar\partial_x \psi_p(x) = \hat{p}(x|\mathbf{p}) - \langle x|\mathbf{p}|\hat{p} = p\psi_p(x)$  from which, solving the differential equation,  $\langle p_x | x \rangle = \frac{e^{-ip_x x/\hbar}}{(2\pi\hbar)^{1/2}}$  (where the denominator is chosen to have a normalized function). □

### 5.5.3 Schrödinger equation for the wavefunction

We have studied already the law governing the evolution of a quantum system. We saw that the dynamics of the system is generated by the system Hamiltonian  $\mathcal{H}$  (the observable corresponding to the total energy of the system), as described by Schrödinger equation:

$$\frac{i\hbar d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

We can express this same equation in the position representation. We want to describe the evolution of a point particle in a potential  $V(\mathbf{r})$  and with kinetic energy  $T = \frac{P^2}{2m}$ . The Hamiltonian of this simple system is thus given by  $\mathcal{H} = V(\mathbf{r}) + \frac{P^2}{2m}$ . By multiplying the Schrödinger equation with the bra  $\langle r|$  we obtain:

$$\frac{i\hbar d\langle r|\psi\rangle}{dt} = i\hbar\partial_t\psi(r) = \langle r|\mathcal{H}|\psi\rangle = \langle r|V(r)|\psi\rangle + \langle r|\frac{P^2}{2m}|\psi\rangle$$

Using the relationship

$$\langle x|P_x^2|\psi\rangle = (P_x^2\psi)(x, t) = (-i\hbar\partial_x)^2\psi(x, t) = -\hbar^2\frac{\partial^2\psi(x, t)}{\partial x^2},$$

we obtain

$$i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}, t) + V(\mathbf{r}, t)\psi(\mathbf{r}, t)$$

(where  $\Delta$  is the Laplacian operator in 3D).

## 5.6 Feynman's path-integral

The formal solution of the Schrödinger equation above can be written as  $|\psi(t)\rangle = U(t, 0)|\psi(0)\rangle$ . Using the position representation and the closure relation  $\int d^3\mathbf{r}|\mathbf{r}\rangle\langle\mathbf{r}| = \mathbb{1}$  we can write

$$\psi(\mathbf{r}, t) = \int d^3\mathbf{r}'\langle\mathbf{r}|U(t, 0)|\mathbf{r}'\rangle\psi(\mathbf{r}', 0),$$

where  $U(t, 0) = e^{-i\mathcal{H}t/\hbar}$  and the matrix element  $\langle\mathbf{r}|U(t, 0)|\mathbf{r}'\rangle$  is the Green function describing how a localized wave packet centered at  $\mathbf{r}'$  at time  $t = 0$  propagates at a later time in the potential  $V(\mathbf{r})$ . This equation represents the wave function of a single particle  $\psi(r, t)$  as a sum of partial waves propagating from  $r'$  at time 0 to  $r$  at time  $t$ ; it is thus the superposition of many different paths taken by the particle during its evolution. The probability of finding the particle at  $r$  results from a multiple-path interference process.

This picture of the wavefunction propagation can be used to give a qualitative introduction of Feynman's path-integral approach to quantum physics. We do not aim here for a rigorous derivation of that theory, only the main concepts will be presented<sup>13</sup>.

We start by expressing the probability amplitude that a particle, initially prepared at point  $x_i$ , will pass a time  $t$  later at point  $x_f$  as the matrix element between the initial and the final state of the system's evolution operator:  $\langle x_f|U(t, 0)|x_i\rangle$ . We expand this expression by slicing the time interval  $t$  into infinitesimal intervals  $\delta t$  and by introducing at each of these times a closure relationship on the position eigenstates:

$$\begin{aligned}\langle x_f|U(t, 0)|x_i\rangle &= \langle x_f|U(\delta t)^n|x_i\rangle = \\ &= \int dx_n\dots dx_k\dots dx_1\langle x_f|U(t, t-\delta t)|x_n\rangle\langle x_n|\dots U(\delta t)|x_k\rangle\langle x_k|\dots U(\delta t)|x_1\rangle\langle x_1|U(\delta t, 0)|x_i\rangle \\ &= \int dx_n\dots dx_1\langle x_k|U(\delta t)|x_{k-1}\rangle\dots\end{aligned}$$

<sup>13</sup> In this section we again closely follow the presentation in S. Haroche, J.-M. Raimond, *Exploring the quantum: atoms, cavities and photons*, Oxford University Press (2006)

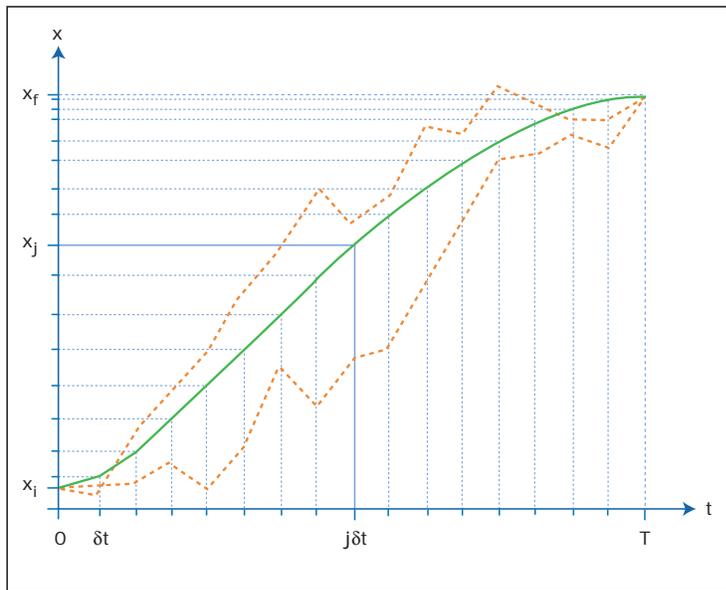


Image by MIT OpenCourseWare.

Fig. 8: Spacetime diagram of the propagation of a particle between two events. Taken from “Exploring the quantum”, S. Haroche, J-M. Raimond.,

We then evaluate the amplitude  $\langle x_k | U(\delta t) | x_{k-1} \rangle$  in the case  $U(t) = e^{-it(p^2/2m+V)/\hbar}$ . As  $\delta t$  is small, we can approximate it by the product of the two terms:

$U(t) = e^{-i\delta t(p^2/2m+V)/\hbar} \approx e^{-i\delta tV/\hbar} e^{-i\delta t p^2/2m\hbar} = e^{-i\delta tV/\hbar} e^{-i\delta t p^2/2m\hbar} (\int |p\rangle \langle p| dp)$  (where we introduced the closure expression for the momentum  $p$ ). We thus obtain the integral

$$\langle x_k | U(\delta t) | x_{k-1} \rangle \approx e^{-i/\hbar V(x_k)\delta t} \int dp e^{i/\hbar p(x_k - x_{k-1})} e^{-i/\hbar (p^2/2m)\delta t},$$

where we used the fact  $\langle x_k | p \rangle \propto e^{i/\hbar p x_k}$ . The integral over  $p$  is just the Fourier transform of a Gaussian, yielding a Gaussian function of  $x_k - x_{k-1}$ . The probability amplitude is then

$$\begin{aligned} \langle x_f | U(t, 0) | x_i \rangle &\propto \int dx_1 dx_2 \dots dx_n e^{i/\hbar \delta t [\frac{1}{2} m (x_f - x_n)^2 / \delta t^2 - V(x_n)]} \dots \\ &= \int dx_1 dx_2 \dots dx_n e^{i/\hbar \delta t [m v_n^2 / 2 - V(x_n)]} \dots e^{i/\hbar \delta t [m v_i^2 / 2 - V(x_i)]} \end{aligned}$$

where we introduced the velocity  $v_k = (x_k - x_{k-1})/\delta t$ . The probability amplitude for the system to go from  $x_i$  to  $x_f$  in time  $t$  is thus a sum of amplitudes - one for each possible classical path - whose phase is the system's action  $S = \int \mathcal{L} dt$  along the trajectory, where  $\mathcal{L} = \frac{1}{2} m v^2 - V(x) = m v^2 - \mathcal{H}$  is the Lagrangian of the system. This phase is expressed in units of  $\hbar$ .

We have derived this important result by admitting the Schrödinger equation formalism of quantum mechanics. Feynman proceeded the other way around, postulating that a quantum system follows all the classical trajectories with amplitudes having a phase given by the classical action and has derived from there Schrödinger's equation. At the classical limit  $S/\hbar \gg 1$ , the phase along a trajectory evolves very fast when the path is slightly modified, by changing for instance one of the  $x_j$ . The amplitudes of various neighboring paths thus interfere destructively, leaving only the contributions of the trajectories for which the phase, hence the action, is stationary. If the particles action in units of  $\hbar$  is much larger than 1, the particle follows a classical ray. Suppressing the contributions to the amplitude coming from trajectories far from the classical one does not appreciably affect this amplitude.

## 6. Composites systems and Entanglement

- 6.1 Tensor product of Hilbert spaces
  - 6.1.1 Product Operator Basis
- 6.2 Quantum Information Processing
- 6.3 Operators on two Qubits
- 6.4 No cloning Theorem
- 6.5 Entanglement and EPR paradox
  - 6.5.1 Bell Inequalities
- 6.6 Teleportation (Bennet, Peres, Brassard)
- 6.7 Deutsch-Jozsa algorithm

### 6.1 Tensor product of Hilbert spaces

Until now we have been concerned with the description and evolution of a single TLS. Although we have seen some examples of how it describes some real physical systems, of course many systems are more complicated and cannot be described by that formalism. We could of course start studying higher dimensional systems, such as more general angular momentum with dimension  $N$ . Here we focus instead on systems with dimension  $N = 2^n$  (with  $n$  integer) because we are interested in studying composite (or multipartite) systems, where two or more TLS systems interact. Let's consider 2 two-level Hilbert spaces  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , each spanned by the vectors:  $|0\rangle_A, |1\rangle_A$  and  $|0\rangle_B, |1\rangle_B$ . For each space we can define the Pauli Matrices and the identity on the space. They are two distinguishable Hilbert space (we will deal with indistinguishable particles later on). The action of a Pauli matrix on the vector of its own Hilbert space is as usual (e.g.  $\sigma_x^A |0\rangle_A = |1\rangle_A$ ). But operators of the  $A$  Hilbert space do not act on the vectors of the other Hilbert space, they leave them unchanged:  $\sigma_x^A |0\rangle_B = |0\rangle_B$ .

We can define the joint space  $\mathcal{H}_{AB}$  by a tensor product  $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ , which has dimensions  $N = 2^2 \times 2^2 = 2^4 = 16$ . When we consider a matrix representation of the Hilbert space, this corresponds to a kronecker product. For example, the kronecker product of two matrices (operators)  $A$  and  $B$  is given by:

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{22}B_{11} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & \dots \\ A_{21}B_{11} & A_{21}B_{12} & \dots & \dots \\ A_{21}B_{12} & & & \end{pmatrix}$$

that is, a  $4 \times 4$  matrix. In the same way, the vector states of the joint Hilbert space are defined by the kronecker products of the basis states of the two spaces. For example:

$$|0\rangle_A \otimes |1\rangle_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

A basis set for a two-qubit system (two TLS) is given by the four states:  $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ . Notation-wise, we normally do not write the identity:  $\sigma_x^A \otimes 1_B = \sigma_x^A$ .

If one spin space is spanned by 4 matrices, the joint domain A and B is spanned by 16 operators, which are the combinations of operators from the two spaces:  $\{1, \sigma_x^A, \sigma_x^B, \dots, \sigma_x^A \sigma_y^B, \dots, \sigma_z^A \sigma_z^B\}$ .

The joint space is still an Hilbert space. If  $|a\rangle$  is a vector in the  $\mathcal{H}_A$  space and  $|b\rangle$  in the  $\mathcal{H}_B$  space, taking a vector in the joint space  $|a\rangle \otimes |b\rangle$  it has the properties of a linear vector:

$$(|a_1\rangle + |a_2\rangle) \otimes |b\rangle = |a_1\rangle \otimes |b\rangle + |a_2\rangle \otimes |b\rangle$$

and

$$c(|a\rangle \otimes |b\rangle) = c|a\rangle \otimes |b\rangle = |a\rangle \otimes c|b\rangle$$

(notice that the scalar can be pushed trough as desired).

If  $A$  is an operator in  $\mathcal{H}_A$  and  $B$  in  $\mathcal{H}_B$ , each operator acts on its own domain:  $AB(|a\rangle \otimes |b\rangle) = (A|a\rangle) \otimes (B|b\rangle)$ . If  $\mathcal{H}_C = \mathcal{H}_{AB}$  is the joint Hilbert space, any operator in it can be written as a linear combination of operators in the two spaces:  $C = \sum_{i,j} c_{i,j} A_i B_j$ , where  $i$  and  $j$  run on the two domains and  $\{A_i\}$ ,  $\{B_j\}$  form complete sets (a basis for the operator spaces).

The inner product of vectors in the joint space are

$$(\langle b_1| \otimes \langle a_1|)(|a_2\rangle \otimes |a_2\rangle) = \langle a_1|a_2\rangle \langle b_1|b_2\rangle.$$

A ket of a joint space can also be written as  $|a, b\rangle$ , that is, a ket can be specified by as many quantum number as required to fully characterize the state.

### 6.1.1 Product Operator Basis

We can generalize these considerations to more than two TLS (or qubits or spin- $\frac{1}{2}$ ). We thus define a composite Hilbert space of dimension  $N = 2^n$ , where  $n$  is the number of qubits, as the tensor product of the Hilbert space for each qubit:  $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$ . A basis for this operator space is the product operator basis (also called generalized Pauli operators). Elements of this basis are defined as

$$P_l = \bigotimes_{j=1}^n P_l^{(j)},$$

where each  $P_l^{(j)}$  is either a Pauli matrix  $\{\sigma_x, \sigma_y, \sigma_z\}$  or the identity  $\mathbb{1}$  in the space of the qubit  $j$ . Notice that  $P_l^\dagger = P_l$  (hermitian) and  $\text{Tr}\{P_l P_{l'}\} = N \delta_{l,l'}$  (that is, the basis is orthogonal, but nor normalized).

## 6.2 Quantum Information Processing

Quantum information processing is the study of information processing tasks that can be accomplished (only) using quantum mechanical systems. What do we mean by *only*? What we refers to are tasks that can be possible only if the law of quantum mechanics apply to the system used for processing the information or that are accomplished in a more efficient way if performed by a quantum system (in terms of time or material resources). For example, Peter Shor showed in 1994 that it is possible to find the prime factors of a number using a quantum computer in an exponentially shorter time than in a classical computer. The scaling refers to the fact that if we want to factorize a number represented by  $n$  bits of information (e.g. in its binary representation the string is n-character long) it will take a time  $T_{cl} \propto 2^n$  for a classical computer to perform the computation, while only a time  $T_{qu} \propto n$  to a quantum computer. Although factoring the number 15 is easy<sup>14</sup>, factoring large numbers is a very time-consuming task, so much that encryption is based on number factoring (as the reverse operation, finding the product of two numbers, is instead an easy task).

---

<sup>14</sup> Why do I mention here 15? Because that is the number that has been possible to factorize until now by a quantum computer:

L.M.K Vandersypen, M. Steffen, G. Breyta, C.S. Yannoni, M.H. Sherwood and I.L. Chuang, *Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance*, *Nature* **414** 883-887 (2001)

While it is still a debate of where the power of quantum computation comes from, two main ingredients seems to have a preeminent role. Quantum superposition (in the form of parallelism that allows to compute all the possible solutions of a problem at once) and interference (that leads to algorithms that select a constructive interference for the correct solution, so that we obtain the right answer with high probability once we measure the quantum system and collapse the superposition state). As it is implied in this last statement, not all the tasks can be made more efficient on a quantum computer and in fact it has proven quite hard to find quantum algorithms (although the known ones are quite powerful).

Quantum information processing has ramifications well beyond quantum computation. Very active areas of research - and of practical results - are quantum communications, simulations, sensors, and of course on the theory side, quantum control, quantum complexity, entanglement theory, decoherence, etc.

Here we will adopt some of the language and tools of quantum information to explore ideas that connect to the very foundation of quantum theory. We will start by describing operations that can be performed on a quantum computer. As at its heart a quantum computer is just a QM physical system, these operations simply describe the evolution of the system itself.

In the same way as classical computer are physical systems, circuits made of wires and gates, a quantum computer is also composed of wires and quantum gates. The wires are used to carry the information around, while the gates perform operations, manipulate the information. Quantum gates however have the properties of being linear and *invertible*, as they represent the unitary evolution of a quantum system (a collection of TLS or qubits). This is different than usual classical gates, although invertible classical gates were already known.

### 6.3 Operators on two Qubits

There are several operators which are normally used in quantum computation and that describe the possible evolution of the system.

- Not<sup>A</sup> =  $\sigma_A^x \otimes 1_B$ ; Not<sup>B</sup> =  $1_A \otimes \sigma_B^x$ .

- Hadamard gate:  $H = (\sigma_x + \sigma_z)/\sqrt{2}$ .

- Controlled Not: rotate B conditionally on the state in the A subspace. Introducing the idempotents<sup>15</sup> (or projectors)  $E^+ = |0\rangle\langle 0|$  and  $E^- = |1\rangle\langle 1|$ , the CNOT is  $C^A NOT^B = E_A^+ + E_A^- \sigma_B^x$ :

$$C^A NOT^B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

If somebody has taken some computer science classes, you can realize that the truth table of the cnot gate is quite similar to that of a XOR gate. We can also just have general single qubit gates,  $U$ , that describe any general rotation

A	B	AB
0	0	00
0	1	01
1	0	11
1	1	10

Table 2: Truth table of the CNOT (1st qubit controller, 2nd qubit target)

on a single qubit. If we combine this single qubit rotations with the CNOT gates on any pair of qubit, we are able to build any possible algorithm (or computation) on the system. That also means that we are able to enforce any possible evolution of the system, by letting it evolve under these two types of gates. We says that they are universal gates.

---

<sup>15</sup> Idempotents since they square to themselves

### A. Measurement in the $\sigma_x$ basis

At the end of a circuit, the qubits are measured. While usually it is implicit that the qubits are measured in their computational basis ( $|0\rangle, |1\rangle$ ), which corresponds to the eigenvalues of the operator  $\sigma_z$ , this does not always have to be the case. The eigenvectors of  $\sigma_x$  form an equivalently good basis. We could have expressed a state vector as:  $|\psi\rangle_z = a|0\rangle + b|1\rangle \rightarrow |\psi\rangle_x = c|0\rangle_x + d|1\rangle_x = c|+\rangle + d|-\rangle$  (the last expression is a notation encountered often). The coefficients  $c$  and  $d$  can be calculated with a change of basis. First, notice that the eigenvectors of  $\sigma_x$  in the  $z$ -basis are given by the eigenvectors of the matrix

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

that is:

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \quad |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$$

The operator that performs the change of frame is therefore the Hadamard matrix:  $|\psi\rangle_x = H|\psi\rangle_z$ . This is also the reason why, instead of measuring in the  $x$ -basis, we can perform an Hadamard operation to bring back the qubit to the  $z$ -basis, and measure in this more usual basis.

The representations of gates, qubits and wires is usually done via diagrams like the following:

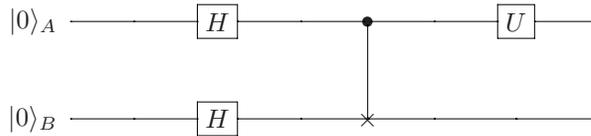


Fig. 9: Quantum circuit, showing Hadamard, CNot gate and a general gate  $U$

## 6.4 No cloning Theorem

We are going to study some properties of quantum states that distinguish them from classical states. One property that has been known for a long time, without stirring much interest before it was considered again in the optics of quantum computation is the impossibility of copying a quantum state. This impossibility seemed to doom quantum computation, because it seemed to forbid correction codes, but quantum resources offer other ways to perform error correction.

The so-called *No-cloning* theorem, states that:

■ **Theorem:** It is impossible to make a perfect copy of an unknown, pure state by an unitary operation.

Proof: I want to copy an arbitrary state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  on the blank initial state  $|i\rangle$  using a unitary operator  $U$ . The final state is therefore:

$$U|\psi\rangle \otimes |i\rangle \stackrel{?}{=} |\psi\rangle \otimes |\psi\rangle$$

for any state  $|\psi\rangle$  in the domain of the first system. If I assume to be able to copy any arbitrary state, I can assume that I can copy at least another state  $|\varphi\rangle$ , which is not the state  $|\psi\rangle$  and not orthogonal to it. For this second state we have:

$$U|\varphi\rangle \otimes |i\rangle = |\varphi\rangle \otimes |\varphi\rangle$$

Equating the inner products of the RHS and LHS of the two equations above, we obtain:

$$\begin{aligned} \langle \varphi, i | U^\dagger U | \psi, i \rangle &= \langle \varphi, \varphi | \psi, \psi \rangle = \\ \langle \varphi | \psi \rangle \langle i | i \rangle &= \langle \varphi | \psi \rangle \langle \varphi | \psi \rangle \\ \langle \varphi | \psi \rangle &= \langle \varphi | \psi \rangle^2 \end{aligned}$$

The last equation is verified iff  $\langle \varphi | \psi \rangle = 1$  or  $\langle \varphi | \psi \rangle = 0$ . In the first case, the two states are in effect the same state (up to a normalization factor or a global phase, which are not important). In the second case the two states are orthogonal, in contradiction with the hypothesis.  $\square$

A unitary operator cannot copy an arbitrary state. If we find an operator that can clone one state, it can only copy that state and states which are orthogonal to it, but it cannot clone all the other states. In a Hilbert space it is therefore possible to define an operator that clones the basis states, but not an arbitrary superposition of them.

### Example of "Cloning"

Consider the action of the CNOT gate on the state  $|\psi\rangle|0\rangle$ , where  $|\psi\rangle$  is the state we would like to clone and  $|0\rangle$  is the blank bit we want to copy on. If  $|\psi\rangle = |0\rangle$ , the CNOT gives us the state  $|00\rangle$ , if it is  $|1\rangle$  we obtain the state:  $|11\rangle$ . So it seems that it is possible to copy the state of the first qubit on the second qubit. But notice that for the moment we have only verified that we can copy two orthogonal state. If we have a more general state:  $|\psi\rangle = a|0\rangle + b|1\rangle$ , the action of the CNOT will give us:

$$\begin{aligned} CNOT|\psi, 0\rangle &= CNOT(a|00\rangle + b|10\rangle) = a|00\rangle + b|11\rangle \\ &\neq |\psi, \psi\rangle = (a|0\rangle + b|1\rangle)(a|0\rangle + b|1\rangle). \end{aligned}$$

Notice that approximate cloning is possible<sup>16</sup> (that is, it is possible to obtain an approximate copy of an arbitrary state up to an error  $\epsilon$ . The error is usually measured as the deviation from unity of the inner product of the original and "cloned" state:  $\epsilon = 1 - |\langle \psi | \varphi \rangle|$ ).

---

<sup>16</sup> Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin and Antonio Acin, *Quantum cloning*, *Rev. Mod. Phys.* **77**, 1225 - 1256 (2005) *Abstract* - The impossibility of perfectly copying (or cloning) an unknown quantum state is one of the basic rules governing the physics of quantum systems. The processes that perform the optimal approximate cloning have been found in many cases. These "quantum cloning machines" are important tools for studying a wide variety of tasks, e.g., state estimation and eavesdropping on quantum cryptography. This paper provides a comprehensive review of quantum cloning machines both for discrete-dimensional and for continuous-variable quantum systems. In addition, it presents the role of cloning in quantum cryptography, the link between optimal cloning and light amplification via stimulated emission, and the experimental demonstrations of optimal quantum cloning.

## 6.5 Entanglement and EPR paradox

It is nearly 70 years ago that Schrödinger gave the name *Verschraenkung* to a correlation of quantum nature. This term was then rather loosely translated to *entanglement*. Over the decades the meaning of the word *entanglement* has changed its flavor, going from a negative statement by Einstein and coworkers “*An entangled wavefunction does not describe the physical reality in a complete way*”, to more quantitative definitions (Bell, “*A correlation that is stronger than any classical correlation*”) to more practical ones (C. Bennett: “*A resource that enables quantum teleportation*”, P. Shor: “*that allows for faster algorithms*”).

A simple definition of entanglement is possible for pure, bipartite systems (i.e. composed of two subsystems).

**D: Entanglement** A pure state  $|\psi\rangle$  is called separable iff it can be written as  $|\psi\rangle = |\varphi\rangle_1 \otimes |\varphi\rangle_2$ , otherwise it is entangled. An example for a pure separable state is  $|00\rangle$ ; examples for pure entangled states are the Bell states

$$|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$$

$$|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$$

We will see some measure of entanglement and also some difficulties arising for example in defining and measuring entanglement for more complex systems.

Why is entanglement a difficult property to quantify and more importantly, to grasp its meaning?

We will review the so-called EPR paradox which makes it manifest some of the weirdness of QM as associated to entanglement.

In 1935 Einstein published a paper with some coworkers that asked :

*Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*<sup>17</sup>

We will rephrase their result in a slightly different way (due to Bohm) and following the presentation in Ballentine's book<sup>18</sup>.

### 6.5.1 Bell Inequalities

Let us suppose that we are capable of making a state  $|\psi\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$  of two identical spin-1/2 particles, with the two particles traveling with equal momenta in opposite directions. For example, they could originate in the decay of an unstable particle of zero spin and zero momentum, in which case momentum conservation implies that the particles move in opposite directions and have spin with zero sum.

Two experimentalists, conventionally named Alice and Bob (A,B), measure the spin component of each particle along a certain axis when the particles are very far apart compared with the range of any force of mutual interaction and when they have not interacted with each other for a long time.

Alice measures the spin component on the  $\hat{a}$  axis for the particle traveling to the left, particle  $a$ , while Bob measures the component along the  $\hat{b}$  axis of the particle traveling to the right, particle  $b$ . Let us first study the case where Alice and Bob both use the  $z$ -axis,  $\hat{a} = \hat{b} = \hat{z}$ . For the moment we can just think of the spins as a property of the particles, as it could be e.g. the color of a ball.

Alice measures the  $z$  component of the spin of particle  $a$ ,  $S_z^a$ , with the result  $\pm\frac{1}{2}$ , and Bob measures  $S_z^b$ . They obtain a series of random results, when they repeat the experiment. After the series of measurements has been completed, Alice and Bob meet and compare their results. They conclude that the results for each pair exhibit a perfect (anti-)correlation. When Alice has measured  $+1/2$  for particle  $a$ , Bob has measured  $-1/2$  for particle  $b$  and vice versa.

Upon reflection, this result is not very surprising. It can occur also for classical particles (or travelers!). Two travelers  $a$  and  $b$ , each carrying a suitcase, depart in opposite directions from the origin and eventually are checked by two customs inspectors Alice and Bob. One of the suitcases contains a red ball and the other a green ball, but the travelers have picked up their closed suitcases at random and do not know what color the ball inside is. If Alice checks the suitcase of traveler  $A$ , she has a 50% chance of finding a green ball. But if in fact she finds a green ball, clearly Bob

<sup>17</sup> A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 - 780 (1935)

<sup>18</sup> L. E. Ballentine, *Quantum Mechanics A Modern Development*, World Scientific Publishing (1998)

will find a red ball with 100% probability. Correlations between the two suitcases were introduced at the time of departure, and these correlations reappear as a correlation between the results of Alice and Bob.

However, the situation becomes weirder if Alice and Bob decide to use the  $x$  axis instead of the  $z$  axis for another series of measurements. In the classical case, this would correspond to the fact the ball hidden in the suitcases possess another property (for example they are shiny or matte). Again, one would obtain the same result of perfect anti-correlation of the results (try e.g. expressing the Bell state in the  $x$  basis).

In the usual classical picture of the world, one would assume –as stated by Einstein– the hypothesis of **LOCALITY** and **REALISM** (both of these hypothesis should be true at the same time). What do these hypotheses entail?

1. Realism: At preparation, particles  $a$  and  $b$  possess both the properties (color and gloss for the classical balls and  $\sigma_x, \sigma_z$ , with  $\sigma_{x,z} = x(z) = \pm 1$  for the quantum particles).
2. Locality: When I measure particle  $a$  I cannot modify instantaneously the result of measuring particle  $b$ , since  $b$  already had its own properties and there is no action at distance (faster than light).

In the EPR paper, the authors argue that since QM does not give a description coherent with these hypotheses, there must be a more complete theory able to fully describe reality while respecting these hypotheses.

The search for a theory of *hidden variables* is still open, but it has been shown already that local realism is in conflict with experiment.

The Bell inequalities want to show that these two hypotheses cannot be true together for quantum mechanics. They describe a more general experiment to what done until now.

**I** - Assume that A measure her particles along the axis  $\vec{a} = \vec{z}$  while B along the axis  $\vec{b}$  such that  $\vec{b} \cdot \vec{z} = \cos\vartheta$ . The results of the measurements are  $\langle \sigma_z^A \rangle = a$  and  $\langle \sigma_b^B \rangle = b$  and we are interested in the correlation  $\langle ab \rangle$ . This is given by

$$\begin{aligned} \langle \sigma_z^A \sigma_b^B \rangle &= \frac{1}{2} (\langle 01 | \sigma_z^A \sigma_b^B | 01 \rangle + \langle 01 | \sigma_z^A \sigma_b^B | 10 \rangle + \langle 10 | \sigma_z^A \sigma_b^B | 01 \rangle + \langle 10 | \sigma_z^A \sigma_b^B | 10 \rangle) \\ &= \frac{1}{2} (\langle 0 | \sigma_z^A | 0 \rangle \langle 1 | \sigma_b^B | 1 \rangle + \langle 0 | \sigma_z^A | 1 \rangle \langle 1 | \sigma_b^B | 0 \rangle + \langle 1 | \sigma_z^A | 1 \rangle \langle 0 | \sigma_b^B | 0 \rangle + \langle 1 | \sigma_z^A | 0 \rangle \langle 0 | \sigma_b^B | 1 \rangle) \\ &= \frac{1}{2} (\langle 1 | \sigma_b^B | 1 \rangle - \langle 0 | \sigma_b^B | 0 \rangle) = -\cos\vartheta \end{aligned}$$

where the last equation comes from the fact that  $\sigma_b^B = \cos\vartheta\sigma_z^B + \sin\vartheta\sigma_\perp^B$ .

**II** - Now we choose two other directions  $\vec{a}'$  and  $\vec{b}'$  each rotated by some angle  $\varphi$  with respect to the original directions. Then what we have done is a collective rotation of the coordinate frame, but we have seen already that the Bell state is unchanged by such a rotation. Thus by repeating the same analysis we will find that  $\langle a'b' \rangle = \langle ab \rangle = -\cos\vartheta$ .

**III** - Consider then the following experiment:

A can measure either  $\vec{a}$  or  $\vec{a}'$

B can measure either  $\vec{b}$  or  $\vec{b}'$

and we want to look at the correlation of the outcomes  $\langle ab \rangle$ ,  $\langle a'b \rangle$ ,  $\langle ab' \rangle$  and  $\langle a'b' \rangle$ . The quantity we are interested in is actually  $\langle S \rangle = \langle ab \rangle + \langle a'b \rangle + \langle ab' \rangle - \langle a'b' \rangle$ . There are two possible strategies:

a) One can measure each correlation in separate experiments (i.e. we measure separately  $\langle ab \rangle$  etc.). We then expect the results  $\langle ab \rangle = -\cos\vartheta_{ab}$ ,  $\langle a'b \rangle = -\cos\vartheta_{a'b}$  etc. and

$$\langle S \rangle = -(\cos\vartheta_{ab} + \cos\vartheta_{a'b'} + \cos\vartheta_{ab'} - \cos\vartheta_{a'b})$$

b) One can look at the outcome of the quantity  $S_k = (\sigma_a^A \sigma_b^B)_k + (\sigma_{a'}^A \sigma_b^B)_k + (\sigma_a^A \sigma_{b'}^B)_k - (\sigma_{a'}^A \sigma_{b'}^B)_k$  at each  $k^{th}$  experiment. Then the expectation value is  $\langle S \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k S_k$ . Notice that this definition of the quantity  $S_k$  implies that even in experiments were we measure e.g. along  $\vec{a}$  (i.e. we measure  $\sigma_a^A$  and not  $\vec{a}'$ ,  $\sigma_{z'}^A$  still has a well-defined value (realism). We can rewrite  $S_k$  as

$$S_k = \sigma_a^A (\sigma_b^B + \sigma_{b'}^B)_k - \sigma_{a'}^A (\sigma_b^B - \sigma_{b'}^B)_k$$

In each measurement, the possible results for  $\sigma_b^B$  are  $\pm 1$  (and the same for  $\sigma_{b'}^B$ ) so that the possible outcomes for  $\sigma_b^B + \sigma_{b'}^B$  are  $\{0, +2, -2\}$  and the same for the difference. Whenever  $\sigma_b^B + \sigma_{b'}^B = \pm 2$  we have however that  $\sigma_b^B - \sigma_{b'}^B = 0$  and vice-versa. Thus the possible outcomes for  $S_k$  are  $\pm 2\sigma_a$  or  $\pm 2\sigma_{a'}$  or finally  $S_k = \pm 2$  (since outcomes for  $\sigma_a$  are  $\pm 1$  and we assume that the act of measuring B does not change the outcome of A). Then, the expectation value for any possible choice of the axis direction is bounded by

$$-2 < \langle S \rangle < +2$$

If we now go back to the first strategy *a*) and choose as the measurement axes

$$\vec{a} = \vec{z}, \quad \vec{a}' = \vec{x}, \quad \vec{b} = \vec{z} - \vec{x}, \quad \vec{b}' = \vec{z} + \vec{x}$$

we find :

$$\begin{aligned} \langle ab \rangle &= -\cos \vartheta_{ab} = -1/\sqrt{2} & \langle a'b' \rangle &= -\cos \vartheta_{a'b'} = -1/\sqrt{2} \\ \langle ab' \rangle &= -\cos \vartheta_{ab'} = -1/\sqrt{2} & \langle a'b \rangle &= -\cos \vartheta_{a'b} = 1/\sqrt{2} \end{aligned}$$

which yields

$$\langle S \rangle = \langle ab \rangle + \langle a'b' \rangle + \langle ab' \rangle - \langle a'b \rangle = -\frac{4}{\sqrt{2}} = -2\sqrt{2} < -2$$

Thus the two hypothesis that we assumed in *b*) to arrive at the conclusion  $-2 < \langle S \rangle < +2$  must be wrong (or at least one of them: which one?)

## References

- A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* **Phys. Rev.** **47**, 777 - 780 (1935)

*Abstract* – In a complete theory there is an element corresponding to each element of reality. A sufficient condition for the reality of a physical quantity is the possibility of predicting it with certainty, without disturbing the system. In quantum mechanics in the case of two physical quantities described by non-commuting operators, the knowledge of one precludes the knowledge of the other. Then either (1) the description of reality given by the wave function in quantum mechanics is not complete or (2) these two quantities cannot have simultaneous reality. Consideration of the problem of making predictions concerning a system on the basis of measurements made on another system that had previously interacted with it leads to the result that if (1) is false then (2) is also false. One is thus led to conclude that the description of reality as given by a wave function is not complete.

- J. S. Bell, *On the Einstein Podolsky Rosen Paradox*, **Physics** **1**, 195-200 (1964)

- N.D. Mermin *Bringing home the atomic world: Quantum mysteries for anybody* **American Journal of Physics**, **49** (10), 940-943 (1981)

*Abstract* – A simple device is described, based on a version of Bell's inequality, whose operation directly demonstrates some of the most peculiar behavior to be found in the atomic world. To understand the design of the device one has to know some physics, but the extraordinary implications of its behavior should be evident to anyone. Except for a preface and appendix for physicists, the paper is addressed to the general reader.

- Alain Aspect, Philippe Grangier, and Gerard Roger, *Experimental Tests of Realistic Local Theories via Bell's Theorem*, **Phys. Rev. Lett.** **47**, 460 - 463 (1981)

*Abstract* – We have measured the linear polarization correlation of the photons emitted in a radiative atomic cascade of calcium. A high-efficiency source provided an improved statistical accuracy and an ability to perform new tests. Our results, in excellent agreement with the quantum mechanical predictions, strongly violate the generalized Bell's inequalities, and rule out the whole class of realistic local theories. No significant change in results was observed with source-polarizer separations of up to 6.5 m.

- Alain Aspect, Philippe Grangier, and Gerard Roger, *Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities*, **Phys. Rev. Lett.** **49**, 91 - 94 (1982)

*Abstract* – The linear-polarization correlation of pairs of photons emitted in a radiative cascade of calcium has been measured. The new experimental scheme, using two-channel polarizers (i.e., optical analogs of Stern-Gerlach filters), is a straightforward transposition of Einstein-Podolsky-Rosen-Bohm gedankenexperiment. The present results, in excellent agreement with the quantum mechanical predictions, lead to the greatest violation of generalized Bell's inequalities ever achieved.

## 6.6 Teleportation (Bennet, Peres, Brassard)

Two parties -Alice and Bob- want to transfer an unknown quantum state. They share a resource prior to the transfer, a pair of qubit in one of the Bell States, let say  $|\Phi^+\rangle = (|00\rangle + |11\rangle)/2$ . Alice possesses also another qubit in an unknown pure state  $|\psi\rangle = a|0\rangle + b|1\rangle$ , that she wishes to send to Bob. The circuit below shows the steps in the teleportation algorithm, starting with the gates that create the Bell State on the ancilla qubits.

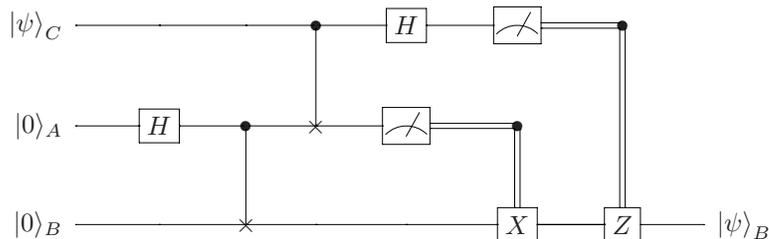


Fig. 10: Circuit for teleportation: the qubit  $|\psi\rangle_C$  (initially in Alice's hands) is teleported to Bob ( $|\psi\rangle_B$ ) by using two qubits in a Bell pair  $|\Phi^+\rangle_{AB}$ .

Alice then transforms her unknown qubit and her part of the shared pair to the Bell State basis by a cnot and a hadamard gate. She then measures them in this new basis and via a classical communication channel, tells the result of the measurement to Bob. Bob performs then an operation on his qubit (the second half of the entangled pair) based on whatever the measurement result was:

- if  $|00\rangle \rightarrow$  do nothing
- if  $|01\rangle \rightarrow \sigma_x$
- if  $|10\rangle \rightarrow \sigma_z$
- if  $|11\rangle \rightarrow \sigma_x\sigma_z$

This operation leaves Bob's qubit in the same state of the one initially owned by Alice. Notice that no superluminal speed of information transmission is proven by quantum teleportation, since classical communication is needed. Also, no cloning of an unknown, arbitrary state is happening (which is forbidden by quantum mechanics), since the original state is destroyed in the process.

The state of the 3 qubits at each step is as follows:

1.  $|\psi00\rangle \xrightarrow{H_A} (|\psi00\rangle + |\psi10\rangle)/\sqrt{2}$  (with  $|\psi\rangle = a|0\rangle + b|1\rangle$ )
2.  $\xrightarrow{C_{ANQT_B}} (|\psi00\rangle + |\psi11\rangle)/\sqrt{2} = |\psi\rangle |\Phi^+\rangle$
3.  $\xrightarrow{C_{BNQT_A}} (a|000\rangle + b|110\rangle + a|011\rangle + b|101\rangle)/\sqrt{2}$
4.  $\xrightarrow{H_C} [|00\rangle (a|0\rangle + b|1\rangle) + |01\rangle (a|1\rangle + b|0\rangle) + |10\rangle (a|0\rangle - b|1\rangle) + |11\rangle (a|1\rangle - b|0\rangle)]/2$
5.  $\xrightarrow{Meas. \dagger U_C} |\psi\rangle_B = a|0\rangle + b|1\rangle$

### References

- A. Fedrizzi, R. Ursin, T. Herbst, M. Nespoli, R. Prevedel, T. Scheidl, F. Tiefenbacher, T. Jennewein & A. Zeilinger, *High-fidelity transmission of entanglement over a high-loss free-space channel*, Nature Physics **5**, 389 - 392 (2009)  
Abstract – Quantum entanglement enables tasks not possible in classical physics. Many quantum communication protocols require the distribution of entangled states between distant parties. Here, we experimentally demonstrate the successful transmission of an entangled photon pair over a 144 km free-space link. The received entangled states have excellent, noise-limited fidelity, even though they are exposed to extreme attenuation dominated by turbulent atmospheric effects. The total channel loss of 64 dB corresponds to the estimated attenuation regime for a two-photon satellite communication scenario. We confirm that the received two-photon states are still highly entangled

by violating the Clauser-Horne-Shimony-Holt inequality by more than five standard deviations. From a fundamental point of view, our results show that the photons are subject to virtually no decoherence during their 0.5-ms-long flight through air, which is encouraging for future worldwide quantum communication scenarios.

Building on work done in:

- R. Ursin, et al (& A. Zeilinger), *Entanglement-based quantum communication over 144 km*, Nature Physics **3**, 481 - 486 (2007)

*Abstract* – Quantum entanglement is the main resource to endow the field of quantum information processing with powers that exceed those of classical communication and computation. In view of applications such as quantum cryptography or quantum teleportation, extension of quantum-entanglement-based protocols to global distances is of considerable practical interest. Here we experimentally demonstrate entanglement-based quantum key distribution over 144 km. One photon is measured locally at the Canary Island of La Palma, whereas the other is sent over an optical free-space link to Tenerife, where the Optical Ground Station of the European Space Agency acts as the receiver. This exceeds previous free-space experiments by more than an order of magnitude in distance, and is an essential step towards future satellite-based quantum communication and experimental tests on quantum physics in space.

- T. Schmitt-Manderbach, et al *Experimental Demonstration of Free-Space Decoy-State Quantum Key Distribution over 144 km* Phys. Rev. Lett. **98**, 010504 (2007)

## 6.7 Deutsch-Jozsa algorithm

To illustrate the power of quantum computation, we present one of the simplest quantum algorithm, the Deutsch-Jozsa algorithm. The algorithm's goal is to decide whether a given function  $f(x)$  is *constant* for all values of  $x$  or *balanced*, that is, equal to 1 for half of the values of  $x$  and to 0 for the other half. The goal is to make this decision with the minimum possible number of evaluations of the function value on trial  $x$  and with a given probability of arriving at the correct answer.

If the function  $f$  is defined on a space of dimension  $2^n$  (i.e.  $x$  can be stored in a  $n$ -bit string), the classical algorithm can decide the function with at least  $\frac{2^n}{2} + 1$  queries, while the quantum one only needs one query. The steps of the algorithm are illustrated in the following picture, where  $\mathbf{H}$  is the Hadamard gate and  $\mathbf{U}_f$  is a unitary gate which transform the state  $|x, y\rangle$  to  $\mathbf{U}_f|x, y\rangle = |x, y \oplus f(x)\rangle$  ( $\oplus$  indicates the addition modulo 2).

In the case where  $f$  is a function from 1 bit to 1 bit, there are only 4 possible  $f$ , two constant and two balanced

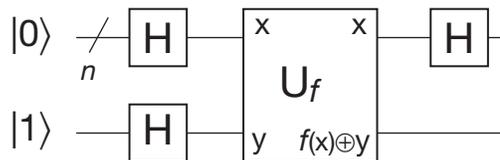


Fig. 11: Circuit implementing the Deutsch-Jozsa algorithm.

( $f_1(x) = 1$ ,  $f_2(x) = 0$ ,  $f_3(x) = x$ ,  $f_4(x) = \bar{x} = NOTx$ ). Since  $\mathbf{U}_f$  gives the sum  $y \oplus x$ , these functions correspond to the following  $\mathbf{U}_i$ :

$$\begin{aligned}
 f_1 &\rightarrow \mathbf{U}_1 = \mathbf{1}_x \otimes \mathbf{U}_{Not,y} \\
 f_2 &\rightarrow \mathbf{U}_2 = \mathbf{1}_x \otimes \mathbf{1}_y \\
 f_3 &\rightarrow \mathbf{U}_3 = \mathbf{U}_{CNot} \\
 f_4 &\rightarrow \mathbf{U}_4 = \mathbf{U}_{CNot} \mathbf{U}_{Not,y}
 \end{aligned} \tag{2}$$

Deutsch's algorithm is a perfect illustration of all that is miraculous, subtle, and disappointing about quantum computers. It calculates a solution to a problem faster than any classical computer *ever* can. It illustrates the subtle interaction of superposition, phase-kick back, and interference. Finally, unfortunately, it solves a completely pointless problem.

We begin by illustrating how superposition of quantum state creates *quantum parallelism* or the ability to compute on many states simultaneously.

Given a function  $f(x) : \{0, 1\} \rightarrow \{0, 1\}$  using a quantum computer, use two qubits  $|x, y\rangle$  and transform them into  $|x, y \oplus f(x)\rangle$  (where  $\oplus$  represents addition modular two). We use two qubits since we wish to leave the input  $x$  or the query register, “un-changed”. The second qubit,  $y$ , acts as a result register. Let  $U_f$  be the unitary transform that implements this.

Suppose we wish to calculate  $f(0)$ , then we could input  $x$  as  $|0\rangle$ , and  $y$ , our output register, as  $|0\rangle$  and apply the  $U_f$  transform, to obtain  $|0\rangle \otimes |0\rangle = |0, 0\rangle \rightarrow |0, 0 \oplus f(0)\rangle$ . If instead we want to calculate  $f(1)$ , then we could input  $x$  as  $|1\rangle$ , yielding the transformation :  $|1\rangle \otimes |0\rangle = |1, 0\rangle \rightarrow |1, 0 \oplus f(1)\rangle$ . In a quantum computer we can actually query the results of 0 and 1 simultaneously using quantum parallelism. For this, let  $x$  equal  $(|0\rangle + |1\rangle)/\sqrt{2}$  and  $y$  equal 0. From the input  $|\psi_1\rangle = \frac{|0,0\rangle+|1,0\rangle}{\sqrt{2}}$  we obtain the output  $|\psi_2\rangle = \frac{|0,f(0)\rangle+|1,f(1)\rangle}{\sqrt{2}}$ .  $U_f$  is applied to  $|0\rangle$  and  $|1\rangle$  simultaneously. This is known as quantum parallelism but there is still a problem since measurement produces either  $|0, f(0)\rangle$  or  $|1, f(1)\rangle$ . Hence we need to be clever about what type of question we ask, and how we go about extracting the answer. For this we use the circuit in the figure, which exploits another quantum mechanical property: *interference*.

The initial state is  $|\psi_0\rangle = |0, 1\rangle$ . We then apply the  $H$  gate to the query and result registers to obtain:  $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

Now, let's examine  $y \oplus f(x)$ :

$$\text{Suppose } f(x) = 0. \text{ Then } y \oplus f(x) = y \oplus 0 = \frac{1}{\sqrt{2}}(|0 \oplus 0\rangle - |1 \oplus 0\rangle) = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$\text{Suppose } f(x) = 1. \text{ Then } y \oplus f(x) = y \oplus 1 = \frac{1}{\sqrt{2}}(|0 \oplus 1\rangle - |1 \oplus 1\rangle) = \frac{1}{\sqrt{2}}(-|0\rangle + |1\rangle)$$

We can compactly describe this behavior as  $y \oplus f(x) = (-1)^{f(x)} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ .

Thus,  $U_f$  transforms  $|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  into:

$$(-1)^{f(x)} |x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Or we can say that:

$$U_f \left[ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right] = \frac{1}{2} \left[ (-1)^{f(0)} |0\rangle (|0\rangle - |1\rangle) + (-1)^{f(1)} |1\rangle (|0\rangle - |1\rangle) \right]$$

Suppose  $f$  is constant, that is  $f(0) = f(1)$ , then:

$$U_f \left[ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right] = \pm \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Suppose instead that  $f$  is balanced, that is  $f(0) \neq f(1)$ , then:

$$U_f \left[ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right] = \pm \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Now apply the Hadamard gate to the first qubit. Just before the measurement the system is in the state

$$|\psi_f\rangle = \begin{cases} \pm \frac{1}{\sqrt{2}} |0\rangle (|0\rangle - |1\rangle) & \text{if } f(0) = f(1) \\ \pm \frac{1}{\sqrt{2}} |1\rangle (|0\rangle - |1\rangle) & \text{if } f(0) \neq f(1) \end{cases}$$

Since in our case  $f(0) \oplus f(1) = 0 \Leftrightarrow f(0) = f(1)$  we can write this as  $|\psi_f\rangle = \pm |f(0) \oplus f(1)\rangle \left[ \frac{|0\rangle+|1\rangle}{\sqrt{2}} \right]$  Hence it is possible to measure the first qubit to find  $f(0) \oplus f(1)$ .

The Deutsch-Jozsa algorithm is a generalization of Deutsch's algorithm to a function  $f(x) : \{2^n\} \rightarrow \{0, 1\}$  that  $f$  is either constant or balanced. The algorithm just generalize to a larger number of qubits.



## 7. Mixed states

### 7.1 Mixed States

### 7.2 Dynamics of mixed states and operators

7.2.1 Heisenberg picture

7.2.2 Interaction picture

### 7.3 Partial Trace

7.3.1 Examples

### 7.4 Entanglement measurement

### 7.5 Mixed States and interpretation of the density matrix

7.5.1 Classical Macro-states

7.5.2 Quantum Macro-states

7.5.3 Example: Spin- $\frac{1}{2}$

## 7.1 Mixed States

Until now we have considered systems whose state was unequivocally described by a state vector. Although the result of an observable measurement on the state is probabilistic, until now the state of the system was well defined and evolved in a deterministic way. When we presented the fundamental concepts of QM we defined the state as a complete description of the set of probabilities for all observables. In particular, we put this into the context of the preparation step of an experiment. Since in order to obtain information about a system, the experiment has to be repeated many times, often we deal with an ensemble of systems (either an ensemble of *copies* of the same systems, or an ensemble *in time* of the same system). In many cases, when we repeat in experiment, it might be difficult to prepare the system in exactly the same state (or prepare perfectly identical copies), thus there is some uncertainty on the initial state.

To describe this situation in more abstract terms, we are thus interested in the case where our information regarding the system is not complete. Thus we will associate the concept of state of a system with an ensemble of similarly prepared systems. By this, we mean an ensemble of systems that could have been prepared in principle, we do not need to refer to a concrete set of systems that coexist in space.

The first postulate now reads: to each state corresponds a unique state operator  $\rho$ . The dynamical variable  $X$  over the ensemble represented by the state operator  $\rho$  has expectation value given by:  $\langle X \rangle = \text{Tr} \{ \rho X \} / \text{Tr} \{ \rho \} = \sum_i \langle i | \rho X | i \rangle$  (Notice that here the summation is done over some basis, but any basis is equivalent as it gives the same result). If we impose to  $\rho$  to have trace 1, the expectation value of  $X$  is just  $\langle X \rangle = \text{Tr} \{ \rho X \}$ . We impose further constraints on  $\rho$ :

- $\text{Tr} \{ \rho \} = 1$  as said.
- $\rho$  is self-adjoint  $\rho^\dagger = \rho$ , so that  $\langle X \rangle$  is real.
- $\rho$  is non-negative  $\langle u | \rho | u \rangle \geq 0$ .

These properties will allow us to associate a probability meaning to  $\rho$ . The state operator  $\rho$  can be expressed as the sum of projectors:  $\rho = \sum_{n=1}^N \rho_n |u_n\rangle \langle u_n|$ , where  $N$  is the dimension of the space (that is,  $\rho$  has a spectral representation in terms of projectors). With the properties established above, we have:  $\sum_n \rho_n = 1$ ,  $\rho_n = \rho_n^*$ , that is, the coefficients are real:  $0 \leq \rho_n \leq 1$ .

If the system can also be described by a state vector  $|\psi\rangle$ , the state operator is given by:  $\rho = |\psi\rangle \langle \psi|$ . A state that can be written in this way is called **pure state**.

Since the state operator for a pure state is a projector, it is an idempotent:  $\rho^2 = \rho$  (Proof:  $(|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|$ ). Therefore, the eigenvalues of  $\rho$  and  $\rho^2$  are the same, or  $\rho_n^2 = \rho_n$  and they must be either 0 or one. Since we know that the sum of the eigenvalues, which is equal to the trace, must be one, we can deduce that the state operator for a pure state has just one eigenvalue equal one and all the other are zero. This is the definition of a pure state, a state with only one non-zero eigenvalue (and equal to 1). An equivalent formulation is to say that  $\text{Tr}\{\rho^2\} = 1$ .

A more general state operator can be written as a convex sum of pure states. To define a convex sum, let's consider a set of state operators  $\{\rho_i\}$  and the operator  $\rho = \sum a_i \rho_i$ . If  $0 \leq a_i \leq 1 \forall i$  and  $\sum a_i = 1$ , the sum is said to be convex and  $\rho$  is a good state operator.

**? Question:** Show that the representation as a convex sum of pure states is not unique.

Consider  $\rho = a|\psi\rangle\langle\psi| + (1-a)|\varphi\rangle\langle\varphi|$  with  $0 \leq a \leq 1$ . Now define

$$|x\rangle = \sqrt{a}|\psi\rangle + \sqrt{1-a}|\varphi\rangle$$

$$|y\rangle = \sqrt{a}|\psi\rangle - \sqrt{1-a}|\varphi\rangle$$

By substitution,  $\rho = \frac{1}{2}|x\rangle\langle x| + \frac{1}{2}|y\rangle\langle y|$ . □

There is actually an infinite number of ways of representing  $\rho$ . A state operator that is not pure, is called **mixed state**. The properties of a mixed state are that  $\text{Tr}\{\rho^2\} < 1$  and it cannot be expressed in terms of one pure state only.

As said, the state operator for a pure state is the outer product of the pure state vector and its dual:  $\rho = |\psi\rangle\langle\psi|$ . The expectation value of an observable is therefore  $\langle X \rangle = \text{Tr}\{|\psi\rangle\langle\psi| X\} = \text{Tr}\{\langle\psi| X |\psi\rangle\}$  since the trace is invariant under permutation. We find the known result:  $\langle X \rangle = \langle\psi| X |\psi\rangle$ .

Imagine we have two state operators in the same Hilbert space. We have:

$$0 \leq \text{Tr}\{\rho_1 \rho_2\} \leq 1$$

the equality  $\text{Tr}\{\rho_1 \rho_2\} = 1$  is reached only if the two state operator are equal and pure.

## 7.2 Dynamics of mixed states and operators

For a pure state, the evolution is dictated by the Schrödinger equation:

$$i \frac{d|\psi\rangle}{dt} = \mathcal{H}|\psi\rangle$$

which has formal solution:  $|\psi(t)\rangle = U(t, 0)|\psi(0)\rangle$ . The unitary operator  $U$  (the propagator) that gives the evolution is the solution of the equation:

$$i \frac{dU}{dt} = \mathcal{H}U(t, 0)$$

If the Hamiltonian is time-independent, the propagator has the form:  $U(t, 0) = e^{-i\mathcal{H}t}$ . The dynamics of a pure state in state operator form ( $\rho = |\psi\rangle\langle\psi|$ ) is simply given by:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| = U(t, 0)|\psi(0)\rangle\langle\psi(0)|U^\dagger(0) = U(t, 0)\rho(0)U^\dagger(t, 0)$$

The equivalent of the Schrödinger equation for the state operators is the Liouville equation:

$$\frac{d\rho}{dt} = -i[\mathcal{H}, \rho],$$

which can be easily derived from the evolution of vector states described by Schrödinger equation.

---

**? Question:** Derive the Liouville equation.

Given the definition of density matrix as a convex sum of pure states:

$$\rho = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}|$$

where each vector state obeys Schrödinger equation:

$$i\hbar|\dot{\psi}\rangle = \mathcal{H}|\psi\rangle$$

we obtain, by taking the derivative of the first equation and inserting the second one:

$$\begin{aligned} i\hbar\dot{\rho} &= i\hbar \sum_{\alpha} p_{\alpha} (|\dot{\psi}_{\alpha}\rangle\langle\psi_{\alpha}| + |\psi_{\alpha}\rangle\langle\dot{\psi}_{\alpha}|) \\ &= \sum_{\alpha} p_{\alpha} (\mathcal{H}|\psi_{\alpha}\rangle\langle\psi_{\alpha}| + |\psi_{\alpha}\rangle\langle\psi_{\alpha}|\mathcal{H}) = [\mathcal{H}, \rho] \end{aligned}$$

□

---

The solution of the Liouville equation is:

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

### 7.2.1 Heisenberg picture

As the Liouville equation is more general than the Schrödinger equation, we would like to reformulate the QM dynamics starting from it. We are thus interested in obtaining the evolution of the observables in the Heisenberg picture starting from the Liouville equation.

The expectation value of an observable  $O$  at time  $t$  is given by the trace:  $\langle O(t) \rangle = \text{Tr} \{ \rho(t) O \} = \text{Tr} \{ U(t, 0) \rho(0) U^{\dagger}(t) O \} = \text{Tr} \{ \rho(0) U^{\dagger}(t) O U(t) \}$ . Notice that using the invariance of the trace under cyclic permutation it is possible to assign the time dependence either to the state operator (Schrödinger picture) or to the operator (Heisenberg picture). In the first one, the state evolves forward in time while the observable operator is time-independent. In the Heisenberg picture instead, the observable evolves "backward" (since as we saw  $U^{\dagger} = U(-t)$ , at least for time-independent hamiltonian) and the state operator is fixed. With this last picture we can follow the evolution of the observable without having to establish a state operator, that is, we can generalize this evolution to a class of state operators.

The operator in the Heisenberg picture at time  $t$  is given by:  $O_H(t) = U^{\dagger}(t, 0) O U(t, 0)$  and it evolves following the equation:

$$\frac{dO_H}{dt} = i [\mathcal{H}(t), O_H(t)] + \left( \frac{\partial O}{\partial t} \right)_H$$

The observable expectation value must be the same in the two pictures:

$$\frac{d\langle O(t) \rangle}{dt} = \text{Tr} \left\{ \frac{d\rho}{dt} O + \rho \frac{\partial O}{\partial t} \right\} = \text{Tr} \left\{ i\rho(t) [\mathcal{H}, O] + \rho(t) \frac{\partial O}{\partial t} \right\}$$

and:

$$\frac{d\langle O(t) \rangle}{dt} = \text{Tr} \left\{ \rho(0) \frac{dO_H}{dt} \right\} = \text{Tr} \left\{ i\rho(0) [\mathcal{H}, O_H] + \rho(0) \left( \frac{\partial O}{\partial t} \right)_H \right\}$$

### 7.2.2 Interaction picture

We revisit the interaction picture also in the context of the Liouville equation. Assume that the overall Hamiltonian of the system can be written as  $\mathcal{H} = \mathcal{H}_0 + V$  (where we separate the known, trivial part  $\mathcal{H}_0$  from the interesting one,  $V$ ). The transformation to the interaction picture is operated by the propagator  $U_I(t) = e^{-i\mathcal{H}_0 t}$ , such that  $|\psi\rangle_I = U_I^{\dagger} |\psi\rangle$  and  $A_I = U_I^{\dagger} A U_I$ .

The evolution of the density matrix in the interaction picture  $\rho_I = U_I^{\dagger} \rho U_I$ , is then:

$$i\dot{\rho}_I = i\dot{U}_I^{\dagger} \rho(t) U_I + iU_I^{\dagger} \dot{\rho} U_I + iU_I^{\dagger} \rho(t) \dot{U}_I$$

with

$$i\dot{U}^\dagger_I = -\mathcal{H}_0 U^\dagger_I(t) \quad \text{and} \quad i\dot{U}_I = U_I(t)\mathcal{H}_0$$

We obtain therefore:

$$\begin{aligned} & -\mathcal{H}_0 U^\dagger \rho(t) U + U^\dagger [\mathcal{H}, \rho(t)] U + U^\dagger \rho(t) U \mathcal{H}_0 \\ &= -[\mathcal{H}_0, \rho_I(t)] + [U^\dagger(t)\mathcal{H}(t)U(t), \rho_I(t)] \\ &= [\mathcal{H}_I, \rho_I(t)] \end{aligned}$$

where  $\mathcal{H}_I = U^\dagger_I(t) V U_I(t)$ .

#### A. Example: rf Hamiltonian in the rotating wave approximation

The interaction picture is particularly useful when the Hamiltonian is composed by a large part time independent ( $\mathcal{H}_0$ ) and a small, time-dependent part  $\mathcal{H}_1(t)$ . The interaction picture is defined by the operator  $U(t)_I = e^{-i\mathcal{H}_0 t}$ , which would give the evolution of the state operator if  $\mathcal{H}_1$  were zero. The interaction picture allows to make more evident the effect of the perturbation on the system, by isolating it and often by simplifying the calculations. Let for example consider the following Hamiltonian acting on a two level system <sup>19</sup>:

$$\mathcal{H} = \underbrace{\omega_0 \sigma_z}_{\mathcal{H}_0} + \underbrace{\omega_1 e^{-i\omega_0 t \sigma_z} \sigma_x e^{i\omega_0 t \sigma_z}}_{\mathcal{H}_1}, \quad \omega_0 \gg \omega_1, \quad \rho(0) = (\mathbb{1} + \epsilon \sigma_z)/2$$

Since  $[\mathcal{H}_0, \sigma_z] = 0$ , in the absence of the perturbation the system does not evolve, it is a constant of the motion. Let us define an unitary operator  $R = e^{-i\omega_0 t \sigma_z}$  that operates the transformation to the interaction picture. We can rewrite the Hamiltonian as:  $\mathcal{H} = \omega_0 \sigma_z + R \omega_1 \sigma_x R^\dagger$ .

The state operator in the interaction picture is given by:  $\rho(t)_I = R^\dagger(t) \rho(t) R(t)$ . Its evolution is therefore:

$$\frac{d\rho_I}{dt} = \frac{dR^\dagger}{dt} \rho R(t) + R^\dagger(t) \frac{d\rho}{dt} R + R^\dagger(t) \rho \frac{dR}{dt}$$

Notice that  $\frac{dR}{dt} = -i\omega_0 \sigma_z R$  and  $\frac{dR^\dagger}{dt} = i\omega_0 \sigma_z R^\dagger$ . We obtain:

$$\frac{d\rho_I}{dt} = i[\omega_0 \sigma_z, \rho] + R^\dagger(t) \frac{d\rho}{dt} R(t)$$

and using Liouville equation we have:

$$\frac{d\rho_I}{dt} = i[\omega_0 \sigma_z, \rho] - iR^\dagger [\mathcal{H}, \rho] R = -iR^\dagger [\mathcal{H}_1, \rho] R = \omega_1 [R^\dagger (R \sigma_x R^\dagger) R, R^\dagger \rho R] \Rightarrow \frac{d\rho_I}{dt} = -i[\omega_1 \sigma_x, \rho_I(t)]$$

Notice that this is true in general:

$$\frac{d\rho_I}{dt} = -i[\tilde{\mathcal{H}}_1, \rho_I], \quad \tilde{\mathcal{H}} = U^\dagger_I(t) \mathcal{H}_1(t) U_I(t)$$

### 7.3 Partial Trace

We define the partial trace of a bipartite system on  $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$  as a linear map  $\text{Tr}_B \{ \cdot \}$  from  $\mathcal{H}_{AB} \rightarrow \mathcal{H}_A$  (or  $\mathcal{H}_B$ ) that is determined by the equation

$$\text{Tr}_B \{ A \otimes B \} = A \text{Tr} \{ B \}$$

(where  $A, B$  are operators on  $\mathcal{H}_A, \mathcal{H}_B$  respectively). This can be extended to more general composite (multipartite) systems. As for the trace, the partial trace is independent of the basis.

*Why do we define the partial trace?* Consider a composite system composed of two parts, A and B, and an observable of the first system only  $O_A$ . The expectation value of the observable on the system A alone is given by:

<sup>19</sup> It could be a nuclear spin in a magnetic field under the action of a weaker rf field

$\langle O_A \rangle = \text{Tr} \{O_A \rho_A\}$  and on the composite system:  $\langle O_A \rangle = \text{Tr} \{(O_A \otimes \mathbb{1}_B) \rho_{AB}\}$ . We can rewrite this last equation as  $\langle O_A \rangle = \text{Tr}_A \{O_A \text{Tr}_B \{\rho_{AB}\}\}$  where  $\text{Tr}_B$  denote the partial trace on the B system. Thus, to obtain information about observables of a subsystem we can first take the partial trace of the state density operator and then use that to calculate the expectation value.

We use also the partial trace to reduce the dimensionality of the system:  $\rho_A = \text{Tr}_B \{\rho_{AB}\}$ .

To calculate the partial trace, write  $\rho$  as a sum of tensor products  $\rho = \sum_{ijkl} m_{ijkl} |a_i\rangle \langle a_j| \otimes |b_k\rangle \langle b_l|$ <sup>20</sup> and for each term we have:  $\text{Tr}_B \{|a_i\rangle \langle a_j| \otimes |b_k\rangle \langle b_l|\} = |a_i\rangle \langle a_j| \text{Tr} \{|b_k\rangle \langle b_l|\}$ .

We are often interested in describing a particular system inside a larger space and we would like to just describe the state of this system  $\rho_S$  without having to describe or know the overall system. The larger system containing the subsystem which we are interested in, can be the environment, a cavity, a field. By doing a partial trace over the environment degrees of freedom we discard the knowledge about them. In general we will obtain a state operator that describes a mixed state (that as we saw, describe some lack of knowledge on the system). The state operator can thus be seen as resulting from the reduction of a larger system to a smaller one, via the partial trace. If the initial multipartite system was entangled, the reduced system is left in a mixed state, since some information was lost. The partial trace reveals the level of entanglement of a state.

### 7.3.1 Examples

1) Pure product state (separable):  $\rho_{AB} = \rho_A \otimes \rho_B$ . The reduced density matrix is therefore:  $\rho_r^A = \text{Tr}_B \{\rho_{AB}\} = \rho_A$ . No information is lost about the A state.

2) Pure entangled state: Bell State.  $\rho = (|00\rangle + |11\rangle) \otimes (|00\rangle + |11\rangle)/2 = (|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|)/2$ . The partial trace over B picks up only the diagonal terms and it gives the reduced matrix:  $\rho_r^A = \text{Tr}_B \{\rho\} = (|0\rangle \langle 0| + |1\rangle \langle 1|)/2$ . All the information about the system A is now lost, since it is now in the maximally mixed state (the identity).

## 7.4 Entanglement measurement

We have seen examples of entangled states, but we haven't given a formal definition of entanglement yet. This is because it is not easy to give such a definition in the most general case. It is however possible to do so in the simplest case of bipartite pure systems. In that case we say that a state is entangled if it cannot be written as  $|\psi\rangle = |a\rangle \otimes |b\rangle$ . If such a decomposition exists, the state is called a *separable* or *product* state. The **Schmidt** decomposition can be used to check if the state is separable.

■ **Theorem:** For any vector  $v$  on the tensor product  $\mathcal{H}_1 \otimes \mathcal{H}_2$  of two Hilbert spaces, there exist orthonormal sets on each space  $\{u_i^1\}, \{u_i^2\}$  such that  $v$  can be written as  $v = \sum_{i=1}^m a_i u_i^1 \otimes u_i^2$  with  $a_i$  non-negative.

The proof is obtained from the singular value decomposition<sup>21</sup>.

The number  $m$  of the vectors needed for the decomposition is called the *Schmidt rank* and the  $a_i$  are the Schmidt coefficients. If the Schmidt rank of a vector is one, the associate state is separable. Note that  $a_i^2$  are the eigenvalues of the reduced density matrix obtained by taking the partial trace over the other system. As such, the rank is easily calculated by taking the partial trace.

The Schmidt rank is sometimes used to quantify entanglement for pure, bipartite systems. There exists many other measure of entanglement, however they coincide at least for this simplest case. For more complex cases, multi-partite, mixed states, the measures are not equivalent and sometimes ill-defined.

### A. Concurrence

One of the most used metrics for pure bipartite states is the concurrence. It can be operatively defined as:  $C = 2|\alpha\delta - \beta\gamma|$ , where the 4 coefficients are defined as:  $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$ . This metric has the following properties:

<sup>20</sup> Notice that by the Schmidt theorem (see [later](#)) we can always find such decomposition.

<sup>21</sup> The proof is presented in M. Nielsen & I. L. Chuang, *Quantum computation and quantum information* Cambridge University Press (2000).

1. The concurrence is bounded by 0 and 1:  $0 \leq C \leq 1$ .
2.  $C = 0$  iff the state is separable.
3.  $C = 1$  for any maximally entangled state.

The four Bell States are maximally entangled states. They correspond to the triplet and singlet manifolds:

$$\begin{aligned} |\varphi^+\rangle &= (|00\rangle + |11\rangle)/2 & |\varphi^-\rangle &= (|00\rangle - |11\rangle)/2 \\ |\psi^+\rangle &= (|01\rangle + |10\rangle)/2 & |\psi^-\rangle &= (|01\rangle - |10\rangle)/2 \end{aligned}$$

We can go from one of the Bell State to another with simple local operations (e.g.  $\sigma_x^1 |\varphi^+\rangle = |\psi^+\rangle$ ), but local operations (that is, operations on single qubit) cannot change the degree of entanglement.

The concurrence can be used to calculate the entanglement even for a mixed state of two qubits. For mixed qubit, an equivalent (more general) definition is given by

$$C(\rho) \equiv \max(0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4})$$

in which  $\lambda_1, \dots, \lambda_4$  are the eigenvalues of

$$A = \rho(\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)$$

in decreasing order ( $\rho^*$  is the complex conjugate of the density matrix).

## B. Entropy

The von Neumann entropy is defined as

$$S(\rho) = -\text{Tr} \{\rho \log \rho\}$$

The entropy of the reduced density matrix is a good measure of entanglement:

$$\mathcal{E} \rightarrow S(\rho_A) = -\text{Tr} \{\rho_A \log \rho_A\}$$

where  $\rho_A = \text{Tr}_B \{\rho\}$ . We can prove that this quantity is the same independently of which subsystem we trace over first.

## C. Purity

We can also consider the purity of the reduced state as a measure of entanglement

$$\mathcal{E} \rightarrow \text{Pur}(\rho_A) = -\text{Tr} \{\rho_A^2\}.$$

## Reference

Dagmar Bruss, *Characterizing entanglement*, Journal of Mathematical Physics, **43**, 9 (2002)

## 7.5 Mixed States and interpretation of the density matrix

We have seen how a mixed state emerged naturally from tracing over one part of a composite system, when the two parts were entangled. Now we can also introduce a density operator as a probabilistic description of a system, instead of the reduced system of a larger one. We consider an ensemble of systems: this ensemble can arise either because there are many copies of the same system (as for example in NMR, where there are  $10^{18}$  molecules in the sample) or because we are making many experiments on the same system (for example in a photon counting experiment from the same molecule). In this last case we have an ensemble over the time. The requirements on the ensemble are

1. that the elements of the ensemble do not interact with each other (first type of ensemble), and
2. that the system does not have memory (ensemble over time).

With these requirements, the physical ensembles we are considering are equivalent to a more abstract concept of ensemble, as seen at the beginning of the chapter.

### 7.5.1 Classical Macro-states

In classical statistical mechanics, equilibrium properties of macroscopic bodies are phenomenologically described by the laws of thermodynamics<sup>22</sup>. The macro-state  $M$  depends on a relatively small number of thermodynamic coordinates. To provide a more fundamental derivation of these thermodynamic properties, we can examine the dynamics of the many degrees of freedom  $N$ , comprising a macroscopic body. Description of each micro-state  $\mu$ , requires an enormous amount of information, and the corresponding time evolution is usually quite complicated. Rather than following the evolution of an individual (pure) micro-state, statistical mechanics examines an ensemble of micro-states corresponding to a given (mixed) macro-state. It aims at providing the probabilities  $p_M(\mu)$ , for the equilibrium ensemble.

#### A. Microcanonical ensemble

Our starting point in thermodynamics is a mechanically and adiabatically isolated system. In the absence of heat or work input to the system, the internal energy  $E$ , and the generalized coordinates  $x$ , are fixed, specifying a macro-state  $M = (E, x)$ . The corresponding set of individual micro-states form the **microcanonical** ensemble. All micro-states are confined to the surface  $H(\mu) = E$  in phase space. The probability distribution function for a microstate  $\mu$  of Hamiltonian  $\mathcal{H}$  is thus just given by the number of accessible states  $\Omega(E)$  at the fixed energy  $E$ :

$$p_E(\mu) = \frac{1}{\Omega(E, x)} \delta(\mathcal{H}(\mu) - E)$$

#### B. Canonical ensemble

Instead of fixing the energy of the system, we can consider an ensemble in which the temperature of the system is specified and its internal energy is then deduced. This is achieved in the canonical ensemble where the macro-states, specified by  $M = (T, x)$ , allow the input of heat into the system, but no external work. The system  $\mathcal{S}$  is maintained at a constant temperature through contact with a reservoir  $\mathcal{R}$ . The reservoir is another macroscopic system that is sufficiently large so that its temperature is not changed due to interactions with  $\mathcal{S}$ . The probability distribution function (p.d.f.) for a microstate  $\mu$  of Hamiltonian  $\mathcal{H}$  in the canonical ensemble is

$$p_T(\mu) = \frac{e^{-\beta\mathcal{H}(\mu)}}{Z(T, x)},$$

where the normalization  $Z(T, x) = \sum_{\{\mu\}} e^{-\beta\mathcal{H}(\mu)}$  is the partition function and  $\beta = 1/k_b T$  (with  $k_b$  the Boltzmann factor). Unlike in a microcanonical ensemble, the energy of a system exchanging heat with a reservoir is a random variable, and it is e.g. possible to define a probability distribution for the energy itself (by changing variables from  $\mu$  to  $\mathcal{H}(\mu)$  in the p.d.f. above.)

#### C. Gibbs and Grand-canonical ensemble

A generalization of the canonical ensemble is to allow the energy to vary by both the addition of heat and work. The Gibbs canonical ensemble describes a system where (mechanical) work is done (which changes the internal variables  $x$ ). In the Grand-canonical ensemble instead chemical work is performed (which varies the number of particles). Thus the chemical potential  $\mu_c$  is fixed and  $N$  can vary.

[The chemical potential of a thermodynamic system is the amount by which the energy of the system would change if an additional particle were introduced, with the entropy and volume held fixed. The chemical potential is a fundamental parameter in thermodynamics and it is conjugate to the particle number.]

---

<sup>22</sup> (Note: this section and the next one is taken from Prof. Kardar 8.333 “Statistical Mechanics I” notes as available on OCW, in some points with only small changes).

## D. Entropy

Given a probability distribution, we can define the entropy  $S$  as

$$S = -k_b \sum_a p_a \log(p_a)$$

(with the convention that  $x \log(x) \rightarrow 0$  for  $x \rightarrow 0$ ) where  $p_a$  describe the probability distribution ( $0 \leq p_a \leq 1$ ,  $\sum_a p_a = 1$ ). It is a measure of our knowledge about the state of the system.

For example, if  $p_j = 1$ ,  $p_i = 0$ ,  $\forall i \neq j$ ,  $S = 0$  (minimum entropy, maximum knowledge). If instead we have a uniform distribution  $p_i = 1/N$ ,  $\forall i$ ,  $S$  is maximum:

$$S = -k_b \frac{1}{N} \sum_i \log\left(\frac{1}{N}\right) = k_b \log(N).$$

In the ensemble interpretation of the density matrix, the entropy  $S(\rho) = -k_b \text{Tr}\{\rho \log \rho\}$  can be seen to have the same meaning as in classical statistics, since we give a probabilistic meaning to the density matrix. Given the decomposition into pure states:  $\rho = \sum p_i |\psi_i\rangle \langle \psi_i|$  we obtain that  $S(\rho) = -k_b \sum_i p_i \log(p_i)$ . In particular the entropy is maximized for the identity state.

The entropy  $S$  describes the lack of knowledge in the system and it can also be used to quantify subjective estimates of probabilities. In the absence of any information, the best unbiased estimate is that all  $N$  outcomes are equally likely. This is the distribution of maximum entropy. If additional information is available, the unbiased estimate is obtained by *maximizing* the entropy subject to the constraints imposed by this information. The entropy maximization method corresponds to finding the best unbiased estimate by minimizing the amount of information that we introduce in the estimate (given what we know about the distribution).

For example, in the canonical ensemble, we maximize the entropy given a fixed average energy. The canonical ensemble can in fact exchange energy with a large heat bath, so that the system is thermalized and the energy kept fixed. The microcanonical ensemble instead describes an isolated system, where the possible states of the system have the same energy and the probability for the system to be in any given state is the same.

### 7.5.2 Quantum Macro-states

We can as well formulate a statistical theory for QM. In QM we have seen already that micro-states are described by vectors in Hilbert spaces, evolving unitarily according to the Schrödinger equation. Unlike in classical mechanics, the value of an operator  $O$  is not uniquely determined for a particular micro-state. It is instead a random variable, whose average in a state  $|\psi\rangle$  is given by  $\langle O \rangle = \langle \psi | O | \psi \rangle$ .

As in the classical case, we can define quantum macro-states describing ensembles of micro-states. Macro-states of the system depend on only a few thermodynamic functions. We can form an ensemble of a large number  $N$  of micro-states  $\mu_a$  corresponding to a given macrostate. The different micro-states occur with probabilities  $p_a$ . (For example  $p_a = 1/N$  in the absence of any other information.) When we no longer have exact knowledge of the microstate, it is said to be in a mixed state.

A mixed quantum state is obtained from a set of possible states  $\{|\psi_a\rangle\}$ , with probabilities  $\{p_a\}$ . The ensemble average of the quantum mechanical expectation value of an observable  $O$  is thus

$$\overline{\langle O \rangle} = \sum_a p_a \langle \psi_a | O | \psi_a \rangle = \sum_{m,n,a} p_a \langle \psi_a | n \rangle \langle n | O | m \rangle \langle m | \psi_a \rangle = \sum_{m,n,a} \langle m | \psi_a \rangle p_a \langle \psi_a | n \rangle \langle n | O | m \rangle = \text{Tr}\{\rho O\}$$

where we defined the density matrix:

$$\rho = \sum_a p_a |\psi_a\rangle \langle \psi_a|$$

with the properties seen above (trace normalization to 1, hermiticity, positivity). We have also already seen that the density matrix obeys the Liouville equation:

$$i\hbar \frac{d\rho}{dt} = [\mathcal{H}, \rho]$$

Equilibrium requires time independent averages, and suggests  $\frac{d\rho}{dt} = 0$ . This condition is satisfied by choosing  $\rho = \rho(\mathcal{H})$ , so that  $[\rho(\mathcal{H}), \mathcal{H}] = 0$ .  $\rho$  may also depend on conserved quantities such that  $[\mathcal{H}, L] = 0$ . Various equilibrium quantum density matrices can now be constructed in analogy to classical statistical mechanics. For example, it is possible to use this minimization of the entropy to calculate the density matrix describing a mixed state.

#### A. Microcanonical ensemble:

As the internal energy has a fixed value  $E$ , a density matrix that includes this constraint is

$$\rho(E) = \frac{\delta(\mathcal{H} - E)}{\Omega(E)}$$

In the matrix representation this can be written as

$$\rho_{n,m} = \langle n | \rho | m \rangle = \sum_a p_a \langle m | \psi_a \rangle \langle \psi_a | n \rangle = \frac{1}{\Omega} \delta(E_n - E) \delta_{n,m},$$

where  $\mathcal{H} |n\rangle = E_n |n\rangle$ . Thus, only eigenstates of the correct energy can appear in the quantum wave-function and (for  $p_a = 1/N$ ) such states on average have the same amplitude,  $|\langle n | \psi_a \rangle|^2 = 1/\Omega$ . This is equivalent to the classical postulate of equal a priori equilibrium probabilities. The  $\Omega$  eigenstates of energy  $E$  are combined in a typical microstate with independent random phases. Note that the normalization condition  $\text{Tr} \{\rho\} = 1$ , implies that  $\Omega(E) = \sum_n \delta(E - E_n)$  is the number of eigenstates of  $\mathcal{H}$  with energy  $E$ .

Notice that we can also obtain the same result by using the maximization of the entropy method. For a microcanonical ensemble, we have no other knowledge on the system than the normalization constraint ( $\text{Tr} \{\rho\} = 1$ ). We thus want to find an unbiased estimate that reflects this minimum knowledge by maximizing the entropy. We thus calculate the density matrix by posing:

$$\begin{cases} \max(S) \\ \text{Tr} \{\rho\} = 1 \end{cases}$$

We can use the Lagrangian multiplier method to solve this problem. Define a function  $\mathcal{L} = S - \lambda [\text{Tr} \{\rho\} - 1]$ , where  $\lambda$  is a coefficient that multiply the constraint condition. The constrained maximum is found at the maximum of the function  $\mathcal{L}$ :

$$\begin{cases} \frac{d\mathcal{L}}{d\rho} = 0 & \rightarrow -k_b \text{Tr} \{\log_2 \rho + \mathbb{1}\} - \lambda \text{Tr} \{\mathbb{1}\} = 0 \\ \frac{d\mathcal{L}}{d\lambda} = 0 & \rightarrow \text{Tr} \{\rho\} = 1 \end{cases}$$

We therefore find  $\rho \propto \mathbb{1}$ , since  $\log(\rho) \propto \mathbb{1}$  from the first equation. From the normalization condition we obtain:  $\rho_{ii} = 1/N$ , where  $N$  is the dimension of the Hilbert space. This expresses the same condition as above (although for a discrete system).

#### B. Canonical ensemble:

A canonical ensemble describes a system with a fixed temperature. A fixed temperature  $T = 1/(k_B \beta)$  can be achieved by putting the system in contact with a reservoir. The canonical density matrix is then obtained by maximizing the system entropy under the constrain of a given average energy.

If the average energy is fixed we have another condition,  $\langle E \rangle = \text{Tr} \{\mathcal{H}\rho\}$  in addition to normalization. Therefore:

$$\mathcal{L} = -k_B \text{Tr} \{\rho \log_2 \rho\} - \lambda_1 [\text{Tr} \{\rho \mathcal{H}\} - \langle E \rangle] - \lambda_2 [\text{Tr} \{\rho\} - 1]$$

We can now calculate the maximum of  $\mathcal{L}$ :

$$k_b \text{Tr} \{\log_2 \rho + 1\} - \lambda_1 \text{Tr} \{\mathcal{H}\} - \lambda_2 \text{Tr} \{\mathbb{1}\} = 0 \rightarrow \log_2 \rho = -\lambda_1 \mathcal{H} + K \mathbb{1}$$

The density matrix is therefore an exponential:  $\rho = e^{-\beta \mathcal{H}} / Z$ , where  $\beta = 1/(k_B T)$  and  $Z$  is the partition function, determined by the normalization condition:

$$Z = \text{Tr} \{e^{-\beta \mathcal{H}}\} = \sum_n e^{-\beta E_n}$$

(where the last expression is calculated in the energy eigenbasis).  
We can calculate the average energy and the entropy:

$$\langle E \rangle = \text{Tr} \{ \mathcal{H} e^{-\beta \mathcal{H}} / Z \} = - \frac{\partial}{\partial \beta} (\ln Z)$$

$$S = -k_B \text{Tr} \{ \rho \log_2 \rho \} = k_B \beta \langle E \rangle + k_B \ln Z$$

In general, any macroscopic observable can be calculated from the partition function.

### C. Grand Canonical ensemble

In the Grand Canonical ensemble the number of particles  $N$ , is no longer fixed. Quantum micro-states with indefinite particle number span a space called *Fock space* (we will come back to this concept when studying the e.m. field). The density matrix can be obtained as before, where we maximize the entropy, subjected now to conditions on the energy and the particle number. It can be shown (although we only mention it here) that

$$\rho(\beta, \mu) = \frac{e^{-\beta \mathcal{H} + \beta \mu N}}{\mathcal{Q}},$$

where the normalization is:

$$\mathcal{Q}(\beta, \mu) = \text{Tr} \{ e^{-\beta \mathcal{H} + \beta \mu N} \} = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_N(\beta)$$

#### 7.5.3 Example: Spin- $\frac{1}{2}$

Consider a spin- $\frac{1}{2}$  system in a magnetic field along z. The Hamiltonian is then  $\mathcal{H} = \frac{\hbar}{2} \gamma B \sigma_z = \hbar \omega \sigma_z$ . At thermal equilibrium, the density matrix is

$$\rho = \frac{e^{-\beta \hbar \omega \sigma_z / 2}}{Z}, \quad Z = \text{Tr} \{ e^{-\beta \hbar \omega \sigma_z / 2} \}$$

We find  $Z = e^{-\beta \hbar \omega / 2} + e^{\beta \hbar \omega / 2}$  and the expectation values:

$$\langle S_x \rangle = \langle S_y \rangle = 0. \quad \langle S_z \rangle = - \frac{\hbar}{2} \tanh \left( \frac{\beta \hbar \omega}{2} \right)$$

In the high temperature approximation, we can expand the exponential to find  $\rho = \frac{1}{2} + \frac{\beta \hbar \omega}{2} \sigma_z$ . This is the expression that is used for example in NMR.

# 8. Open Quantum Systems

- 8.1 Combined evolution of system and bath
- 8.2 Superoperators
- 8.3 The Kraus Representation Theorem
  - 8.3.1 Amplitude-damping
  - 8.3.2 Phase-damping
  - 8.3.3 Depolarizing process
- 8.4 The Master Equation
  - 8.4.1 Markov approximation
  - 8.4.2 Lindblad equation
  - 8.4.3 Redfield-Born theory of relaxation
- 8.5 Other description of open quantum system dynamics
  - 8.5.1 Stochastic Liouville equation and cumulants
  - 8.5.2 Stochastic Wavefunctions

We now proceed to the next step of our program of understanding the behavior of one part of a bipartite quantum system. We have seen that a pure state of the bipartite system may behave like a mixed state when we observe subsystem A alone. What if we want to know the dynamics of A only? Can we describe its evolution even if we don't have full knowledge of B? (the bath) We assume that the state of the bipartite system undergoes unitary evolution: how do we describe the evolution of A alone?

## 8.1 Combined evolution of system and bath

We will first start introducing the evolution of an open quantum system by considering it as a part of a larger (closed) system undergoing the usual unitary evolution. The total Hilbert space is thus  $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B$  and we assume the initial state is represented by the separable density matrix  $\rho = \rho_S \otimes |0\rangle\langle 0|_B$ <sup>23</sup>. The evolution of the total system is then

$$\rho(t) = U_{SB}(\rho_S \otimes |0\rangle\langle 0|_B)U_{SB}^\dagger$$

If we are only interested in the evolution of the system  $S$  we can at this point perform a partial trace on  $B$

$$\rho_S(t) = \text{Tr}_B \{\rho(t)\} = \sum_k \langle k|U_{SB}(\rho_S \otimes |0\rangle\langle 0|_B)U_{SB}^\dagger|k\rangle = \sum_k \langle k|U_{SB}|0\rangle\rho_S(0)\langle 0|U_{SB}^\dagger|k\rangle$$

where  $\{|k\rangle\}$  is an orthonormal basis for  $\mathcal{H}_B$ . As the result of  $\langle k|U_{SB}|0\rangle = \text{Tr}_B \{|0\rangle\langle k|U_{SB}\}$  we obtain an operator  $M_k$  that acts only on the  $S$  Hilbert space. For example, in a matrix representation the elements of  $M_k$  are simply  $M_k^{i,j} = \langle i|M_k|j\rangle$  (with  $|i\rangle, |j\rangle$  defined on  $\mathcal{H}_S$ ); that is, we have  $M_k^{i,j} = \text{Tr} \{|j, 0\rangle\langle i, k|U_{SB}\} = \langle i, k|U_{SB}|j, 0\rangle$ . Now we can write the evolution of the system only density matrix as

$$\rho_S(t) = \mathcal{M}(\rho_S(t)) = \sum_k M_k \rho_S(0) M_k^\dagger$$

---

<sup>23</sup> Here we only assume that the system  $B$  is in a pure state that we indicate as  $|0\rangle$ , we are not assuming that  $B$  is a TLS.

Since the propagator  $U_{SB}$  is unitary, we have that

$$\sum_k M_k^\dagger M_k = \mathbb{1}_S$$

**? Question:** Prove the above.

Inserting the definition for  $M_k$  we have  $\sum_k M_k^\dagger M_k = \sum_k \langle 0|U_{SB}^\dagger |k\rangle \langle k|U_{SB}|0\rangle = \langle 0|U_{SB}^\dagger (\sum_k |k\rangle \langle k|) U_{SB}|0\rangle = \mathbb{1}_S$ .  $\square$

The properties of the system density matrix are preserved by the map:

1.  $\rho_S(t)$  is hermitian:  $\rho_A(t)^\dagger = (\sum_k M_k \rho_S(0) M_k^\dagger)^\dagger = \sum_k M_k \rho_S(0)^\dagger M_k^\dagger = \rho_S(t)$ .
2.  $\rho_S(t)$  has unit trace. (since  $\text{Tr} \left\{ \sum_k M_k \rho_S M_k^\dagger \right\} = \text{Tr} \left\{ \rho_S \sum_k M_k M_k^\dagger \right\} = \text{Tr} \{ \rho_S \mathbb{1} \}$ )
3.  $\rho_S(t)$  is positive.

In the special case where there is only one term in the sum, we revert to the unitary evolution of the density matrix. In that case, a pure state, for example, would remain pure. If that is not the case, that is, the evolution is not unitary, it means that in the course of the evolution the system  $S$  and bath  $B$  became *entangled*, so that  $\rho_A$  is in a mixed state after partial trace. Because of the loss of unitarity, superoperators are in general not invertible and thus there is a specific arrow of time.

## A. Ancillary Bath

In many cases it is not possible to fully calculate the evolution of the total system ( $S + B$ ) as either it is too large or we have imperfect knowledge of the bath. However, if we have a description of the system dynamics in terms of the operator sum, it is possible to always *augment* the system and find a larger, composite system that evolves unitarily and yields the operator sum upon partial trace. The ancillary system might however not have all the characteristic of the (unknown) physical bath. What we are looking for is in fact a minimal description for the bath.

We choose as ancillary Hilbert space  $\mathcal{H}_B$  a space of dimensions at least equal to the number of terms in the operator sum. This space will have then a set of orthonormal vectors  $\{|k\rangle\}$ , and we can define a normalized state  $|0\rangle_B$  on  $\mathcal{H}_B$ . Then the unitary evolution operator of the combined system is defined by imposing the relationship:

$$U_{SB}(|\psi\rangle_S \otimes |0\rangle_B) = \sum_k (M_k \otimes \mathbb{1})(|\psi\rangle_S \otimes |k\rangle_B), \quad \forall |\psi\rangle_S \in \mathcal{H}_S$$

This ensures that the evolution of the reduced system is given by the Kraus map. The total system evolution is:

$$U_{SB}(\rho_S \otimes |0\rangle\langle 0|_B)U_{SB}^\dagger = \sum_{k,h} (M_k \otimes \mathbb{1}) |\psi_S, k\rangle \langle \psi_S, h| (M_h^\dagger \otimes \mathbb{1})$$

and upon taking the partial trace:

$$\begin{aligned} \rho_S(t) = \text{Tr}_B \{ \rho(t) \} &= \sum_j \langle j| \left( \sum_{k,h} (M_k \otimes \mathbb{1}) |\psi_S, k\rangle \langle \psi_S, h| (M_h^\dagger \otimes \mathbb{1}) \right) |j\rangle \\ &= \sum_j \langle j|k\rangle \langle h|j\rangle (M_k^\dagger |\psi\rangle \langle \psi| M_h) \\ \rho_S(t) &= \sum_j M_j^\dagger |\psi\rangle \langle \psi| M_j \end{aligned}$$

Although this relationship doesn't fully define the operator on the full Hilbert space, we can extend the operator as desired. In particular we want it to be unitary (and this imposes added constraints). As the operator  $U_{SB}$  as defined above preserves the inner product on the full Hilbert space, a unitary extension of it to the full space does indeed exist. Furthermore, we can check that upon taking a partial trace on  $B$  we retrieve the operator sum as desired, for an initial pure state on  $S$ . But any density matrix can be expressed as an ensemble of pure states, hence this property is true for any general state on  $S$ .  $\square$

## B. Non-uniqueness of the sum representation

The operator sum is of course not unique, since the choice of the set  $\{|k\rangle\}$  was quite arbitrary and not unique. If we had chosen another set  $\{|h\rangle\}$  we would have arrived to a different sum

$$\rho_S(t) = \sum_h N_h \rho_S(0) N_h^\dagger$$

---

**? Question:** What is the relationship between the operators  $M$  and  $N$ ?

They are related by the simple unitary transformation that connects the two sets of orthonormal vectors  $N_h = U_{hk} M_k$  with  $|h\rangle = \sum_k U_{hk} |k\rangle$ .  $\square$

---

## 8.2 Superoperators

We want to describe the quantum evolution of systems in the most general case, when the system evolves non-unitarily due to the presence of an environment<sup>24</sup>. As we have seen, the states need to be described by density operators. Therefore, the evolution is to be represented by a map connecting the initial density matrix to the evolved one  $\rho(t) = \mathcal{M}[\rho(0)]$ . The most general characteristics of this map will be determined by the fact that the properties of the density matrix should be in general maintained (such as unit trace). As the map  $\mathcal{M}$  is an operator acting on operators, it is called a superoperator.

Most generally, we can define a quantum operator describing the time evolution law for density matrices as a map  $\mathcal{M} : \rho \rightarrow \rho'$  with the following properties

1. Linear
2. Trace preserving
3. Hermiticity preserving
4. Positive
- 4' (Completely positive)

### A. Linearity

Although a non-linear map could also always map a density matrix to another density matrix, if we impose linearity we arrive at results that are more physical. Specifically, the linearity property retains the ensemble interpretation of the density matrix. What we mean is the following. Suppose we can write a density operator as a linear superposition of two densities,  $\rho = a\rho_1 + (1-a)\rho_2$ . The meaning of this expression is that with probability  $a$  we have a system described by  $\rho_1$  and with probability  $1-a$  by  $\rho_2$ . If the map describing the time evolution law is linear, this probabilistic interpretation is valid also for the evolved state. Assume now that the map is not linear, for example it depends on the trace of the density matrix:  $\mathcal{M}(\rho) = e^{iA\text{Tr}\{\rho M\}} \rho e^{-iA\text{Tr}\{\rho M\}}$ , where  $M$  is an operator in the Hilbert space of  $\rho$  and  $A$  an Hermitian operator. We now consider a density operator  $\rho_1$  such that  $\text{Tr}\{\rho_1 M\} = 0$ . We assume that we do not know exactly how we prepared the system, but with 50% probability is in  $\rho_1$ . Assume then the density matrix  $\rho = \frac{1}{2}(\rho_1 + \rho_\perp)$ , such that  $\text{Tr}\{\rho_\perp M\} = 0$ . Then,  $\mathcal{M}(\rho) = \rho$  as the traces are zero. If we now instead consider the initial density matrix  $\rho = \frac{1}{2}(\rho_1 + \rho_\parallel)$  (that is, still a 50% probability of being in  $\rho_1$ ), where  $\text{Tr}\{M\rho_\parallel\} > 0$  we obtain an evolution for  $\rho_1$ . That means, that in the two scenarios, the system behaves differently, even if we had prepared it in the state  $\rho_1$  (remember the probabilistic interpretation), so that the evolution of a potential state of a system  $\rho_1$  depends on another potential state ( $\rho_\perp$  or  $\rho_\parallel$ ), even if this second state never occurred.

---

<sup>24</sup> This presentation in this section and the following examples are taken from J. Preskill's notes at <http://www.theory.caltech.edu/people/preskill/ph229/>

## B. The superoperator preserves trace and hermiticity

Since the density matrix trace has the property to describe the sum of the probabilities of all possible states in the ensemble, it is important that the trace be preserved. An exception can be made for operators that describe measurement (and not time evolution). In that case  $0 \leq \text{Tr}\{\rho\} \leq 1$ . In this case,  $\text{Tr}\{\rho\}$  represent the probability that the measurement outcome described by the map  $\mathcal{M}$  has occurred and the normalized final state is  $\rho/\text{Tr}\{\rho\}$ . As more than one outcome of the measurement is possible, the probability of obtaining  $\rho$  might be less than one.

The superoperator preserves the hermiticity of the density matrix:  $[\mathcal{M}(\rho)]^\dagger = \mathcal{M}(\rho)$  if  $\rho^\dagger = \rho$

## C. Positivity and complete positivity

The property of positivity means that the map is such that  $\mathcal{M}(\rho)$  is non-negative if  $\rho$  is. Although this condition is enough to obtain a valid density matrix, it leads to a contradiction when we consider composite systems. Let's take a valid map  $\mathcal{M}_1$  on system 1. Then, if we consider a bipartite system and we apply the map  $\mathcal{M}_1 \otimes \mathbb{1}$  we would like to still obtain a density matrix on the composite system. Unfortunately, if the map is simply positive, this is not always the case. Thus, we require it to be *completely* positive. A map is completely positive if  $\mathcal{M}_1 \otimes \mathbb{1}_2$  is positive for any extension  $\mathcal{H}_2$  of the Hilbert space  $\mathcal{H}_1$ .

### 8.3 The Kraus Representation Theorem

We have seen in the preceding sections two different ways of describing the evolution of an open system.

The first description started from the evolution of a composite system (including the system of interest and a bath) and by tracing over arrived at a description of the open evolution via the operator sum.

The second description was instead quite abstract, and only defined the properties of the linear map describing the evolution in order to arrive at an acceptable (physical) evolved state (that still possess the characteristics of a density operator). The Kraus representation theorem reconciles these two description, by stating that they are equivalent.

■ **Theorem:** Any operator  $\rho \rightarrow S(\rho)$  in a space of dimensions  $N_S^2$  that obeys the properties 1-3,4' (Linearity, Trace preservation, Hermiticity preservation, complete positivity) can be written in the form:

$$S(\rho) = \sum_{k=1}^K M_k \rho M_k^\dagger, \quad \text{with} \quad \sum_{k=1}^K M_k^\dagger M_k = \mathbb{1}$$

where  $K \leq N_S^2$  is the Kraus number (with  $N_S$  the dimension of the system). As seen above, the Kraus representation is not unique<sup>25</sup>.

We consider three important examples of open quantum system evolution that can be described by the Kraus operators. To simplify the description we consider just a TLS that is coupled to a bath.

#### 8.3.1 Amplitude-damping

The amplitude-damping channel is a schematic model of the decay of an excited state of a (two-level) atom due to spontaneous emission of a photon. By detecting the emitted photon (“observing the environment”) we can get information about the initial preparation of the atom.

We denote the atomic ground state by  $|0\rangle_A$  and the excited state of interest by  $|1\rangle_A$ . The “environment” is the electromagnetic field, assumed initially to be in its vacuum state  $|0\rangle_E$ . After we wait a while, there is a probability  $p$  that the excited state has decayed to the ground state and a photon has been emitted, so that the environment has made a transition from the state  $|0\rangle_E$  (“no photon”) to the state  $|1\rangle_E$  (“one photon”). This evolution is described by a unitary transformation that acts on atom and environment according to

$$\begin{aligned} |0\rangle_S |0\rangle_E &\rightarrow |0\rangle_S |0\rangle_E \\ |1\rangle_S |0\rangle_E &\rightarrow \sqrt{1-p} |1\rangle_S |0\rangle_E + \sqrt{p} |0\rangle_S |1\rangle_E \end{aligned}$$

(Of course, if the atom starts out in its ground state, and the environment is at zero temperature, then there is no transition.)

By evaluating the partial trace over the environment, we find the Kraus operators

$$M_0 = \langle 0|U_{SE}|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad M_1 = \langle 1|U_{SE}|0\rangle = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}$$

The operator  $M_1$  induces a “quantum jump”, the decay from  $|1\rangle_A$  to  $|0\rangle_A$ , and  $M_0$  describes how the state evolves if no jump occurs. The density matrix evolves as

$$\begin{aligned} S(\rho) &= M_0 \rho M_0^\dagger + M_1 \rho M_1^\dagger = \\ &= \begin{pmatrix} \rho_{00} & \sqrt{1-p} \rho_{01} \\ \sqrt{1-p} \rho_{10} & (1-p) \rho_{11} \end{pmatrix} + \begin{pmatrix} p \rho_{11} & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \rho_{00} + p \rho_{11} & \sqrt{1-p} \rho_{01} \\ \sqrt{1-p} \rho_{10} & (1-p) \rho_{11} \end{pmatrix}. \end{aligned}$$

<sup>25</sup> The proof can be found in Prof. Preskill [online notes](#).

If we apply the channel  $n$  times in succession, the  $\rho_{11}$  matrix element decays as  $\rho_{11} \rightarrow \rho_{11}(1-p)^n$  so if the probability of a transition in time interval  $\delta t$  is  $\Gamma\delta t$ , then the probability that the excited state persists for time  $t$  is  $(1-\Gamma\delta t)^{t/\delta t} \approx e^{-\Gamma t}$ , the expected exponential decay law. Also we have  $\rho_{12} \rightarrow \rho_{12}(1-p)^{n/2} \approx \rho_{12}e^{-\Gamma t/2}$ . As  $t \rightarrow \infty$ , the decay probability approaches unity, so

$$S(\rho) = \begin{pmatrix} \rho_{00} + \rho_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

The atom always winds up in its ground state. This example shows that it is sometimes possible for a superoperator to take a mixed initial state to a pure state.

In the case of the decay of an excited atomic state via photon emission, it may not be impractical to monitor the environment with a photon detector. The measurement of the environment prepares a pure state of the atom, and so in effect prevents the atom from decohering. Returning to the unitary representation of the amplitude-damping channel, we see that a coherent superposition of the atomic ground and excited states evolves as

$$(a|0\rangle_S + b|1\rangle_S) |0\rangle_E \rightarrow (a|0\rangle_S + b\sqrt{1-p}|1\rangle_S) |0\rangle_E + \sqrt{p}|0\rangle_A |1\rangle_E$$

If we detect the photon (and so project out the state  $|1\rangle_E$  of the environment), then we have prepared the state  $|0\rangle_A$  of the atom. In fact, we have prepared a state in which we know with certainty that the initial atomic state was the excited state  $|1\rangle_A$  as the ground state could not have decayed. On the other hand, if we detect no photon, and our photon detector has perfect efficiency, then we have projected out the state  $|0\rangle_E$  of the environment, and so have prepared the atomic state

$$a|0\rangle_S + b\sqrt{1-p}|1\rangle_S$$

(or more precisely, if we normalize it:  $(a|0\rangle_S + b\sqrt{1-p}|1\rangle_S)/\sqrt{1-pb^2}$ ). Then  $p(0) = |a|^2 \rightarrow \frac{|a|^2}{1-p|b|^2} > |a|^2$ .

The atomic state has evolved due to our failure to detect a photon, it has become more likely that the initial atomic state was the ground state!

### 8.3.2 Phase-damping

Phase damping describes a process where the system interacts with a large environment composed of many small subsystems. The interaction of the system with each of the environment subsystems is weak (compared to the system energy, but strong compared to the subsystem energy). Therefore the system is unchanged, while the environment subsystem is changed. Since there will be many of these interactions with the environment subsystem, their combined action does have an effect on the system, however it will not be enough to change its energy.

An example is the interaction of a dust particle with photons. Collision of the particle with one photon is not going to change the particle state. However, if the particle was in the ground or excited state, the photon will acquire more or less energy in the collision, thus being excited to its first or second excited state. We now formalize this model.

When looking at the unitary evolution of this process, only the environment changes:

$$|0\rangle_S |0\rangle_E \rightarrow \sqrt{1-p}|0\rangle_S |0\rangle_E + \sqrt{p}|0\rangle_S |1\rangle_E = |0\rangle_S (\sqrt{1-p}|0\rangle_E + \sqrt{p}|1\rangle_E)$$

$$|1\rangle_S |0\rangle_E \rightarrow \sqrt{1-p}|1\rangle_S |0\rangle_E + \sqrt{p}|1\rangle_S |2\rangle_E = |1\rangle_S (\sqrt{1-p}|0\rangle_E + \sqrt{p}|2\rangle_E)$$

Thus a possible unitary is

$$U = \begin{pmatrix} \sqrt{1-p} & \sqrt{p} & 0 & 0 & 0 & 0 \\ \sqrt{p} & \sqrt{1-p} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{1-p} & 0 & \sqrt{p} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{p} & 0 & \sqrt{1-p} \end{pmatrix}$$

The Kraus operator are found by operating the partial trace of the operator above:

$$M_0 = \langle 0|U|0\rangle = \sqrt{1-p}\mathbb{1} \quad M_1 = \langle 1|U|0\rangle = \sqrt{p}|0\rangle\langle 0| \quad M_2 = \langle 2|U|0\rangle = \sqrt{p}|1\rangle\langle 1|$$

The state evolution is then

$$S(\rho) = \sum_{k=1}^3 M_k \rho M_k^\dagger = (1-p)\rho + p|0\rangle\langle 0|\rho|0\rangle\langle 0| + p|1\rangle\langle 1|\rho|1\rangle\langle 1|$$

In matrix form:

$$S(\rho) = \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}$$

Considering the Bloch vector:  $[n_x, n_y, n_z] \rightarrow [(1-p)n_x, (1-p)n_y, n_z]$  (that is, the transvers component are reduced. For  $p = 1$  the state becomes diagonal). Assume  $p = p(\Delta t) = \Gamma \Delta t$  is the probability of one such scatter events during the time  $\Delta t$ . Then if we have  $n$  such events in a time  $t = n\Delta t$  the off-diagonal terms become  $\propto (1-p)^n = (1 - \Gamma \Delta t)^{t/\Delta t} \approx e^{-\Gamma t}$ :

$$S(\rho, t) = \begin{pmatrix} \rho_{00} & e^{-\Gamma t} \rho_{01} \\ e^{-\Gamma t} \rho_{10} & \rho_{11} \end{pmatrix}$$

Consider for example an initial pure state  $\alpha|0\rangle + \beta|1\rangle$ . At long times, this state reduces to:

$$S(\rho, t) = \begin{pmatrix} |\alpha|^2 & e^{-\Gamma t} \alpha \beta^* \\ e^{-\Gamma t} \alpha^* \beta & |\beta|^2 \end{pmatrix} \xrightarrow{t \rightarrow \infty} \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}$$

thus any phase coherence is lost and the state reduces to a classical, incoherent superposition of populations. Because in this process phase coherence is lost (but the energy/population is conserved) the process is called dephasing and the time constant  $1/\Gamma$  is usually denoted by  $T_2$ . Then we have a representation of the superoperator, by expressing  $\rho$  as a linear vector:  $S(\rho, t) = S(t)\rho$ , where  $S = \text{diag}([1, e^{-\Gamma t}, e^{-\Gamma t}, 1])$ .

### 8.3.3 Depolarizing process

The depolarizing channel is a model of a decohering qubit that has particularly nice symmetry properties. We can describe it by saying that, with probability  $1 - p$  the qubit remains intact, while with probability  $p$  an “error” occurs. The error can be of any one of three types, where each type of error is equally likely. If  $\{|0\rangle|1\rangle\}$  is an orthonormal basis for the qubit, the three types of errors can be characterized as:

1. Bit-flip error:  $|\psi\rangle \rightarrow \sigma_x |\psi\rangle$  or  $|0\rangle \rightarrow |1\rangle$  &  $|1\rangle \rightarrow |0\rangle$ .
2. Phase-flip error:  $|\psi\rangle \rightarrow \sigma_z |\psi\rangle$  or  $|0\rangle \rightarrow |0\rangle$  &  $|1\rangle \rightarrow -|1\rangle$ .
3. Both errors:  $|\psi\rangle \rightarrow \sigma_y |\psi\rangle$  or  $|0\rangle \rightarrow i|1\rangle$  &  $|1\rangle \rightarrow -i|0\rangle$ .

If an error occurs, then  $|\psi\rangle$  evolves to an ensemble of the three states  $\sigma_x |\psi\rangle, \sigma_y |\psi\rangle, \sigma_z |\psi\rangle$ .

The depolarizing channel can be represented by a unitary operator acting on  $\mathcal{H}_{SE} = \mathcal{H}_S \otimes \mathcal{H}_E$ , where  $\mathcal{H}_E$  has dimension 4. The unitary operator  $U_{SE}$  acts as

$$U_{SE} |\psi\rangle_S \otimes |0\rangle_E \rightarrow \sqrt{1-p} |\psi\rangle_S \otimes |0\rangle_E + \sqrt{\frac{p}{3}} [\sigma_x |\psi\rangle_S \otimes |1\rangle_E + \sigma_y |\psi\rangle_S \otimes |2\rangle_E + \sigma_z |\psi\rangle_S \otimes |3\rangle_E]$$

The environment evolves to one of four mutually orthogonal states that “keep a record” of what transpired; if we could only measure the environment in the basis  $\{|\mu\rangle, \mu = 0, 1, 2, 3\}$ , we would know what kind of error had occurred (and we would be able to intervene and reverse the error).

**Kraus representation:** To obtain an operator-sum representation of the channel, we evaluate the partial trace over the environment in the  $\{|\mu\rangle_E\}$  basis. Then

$$M_\mu = \langle \mu | U_{SE} | 0 \rangle_E$$

$$M_0 = \sqrt{1-p} \mathbb{1}, \quad M_1 = \sqrt{\frac{p}{3}} \sigma_x, \quad M_2 = \sqrt{\frac{p}{3}} \sigma_y, \quad M_3 = \sqrt{\frac{p}{3}} \sigma_z$$

A general initial density matrix  $\rho_S$  of the qubit evolves as

$$\rho \rightarrow \rho' = (1-p)\rho + \frac{p}{3}(\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z)$$

It is also instructive to see how the depolarizing channel acts on the Bloch sphere. An arbitrary density matrix for a single qubit can be written as  $\rho = \frac{1}{2}(\mathbb{1} + \vec{n} \cdot \vec{\sigma})$ , where  $\vec{n}$  is the Bloch vector (with  $P = |\vec{n}|$  the polarization of the spin). Suppose we rotate our axes so that  $\vec{n} = \vec{z}$  and  $\rho = \frac{1}{2}(\mathbb{1} + P_z \sigma_z)$ . Then since  $\sigma_z \sigma_z \sigma_z = \sigma_z$  and  $\sigma_x \sigma_z \sigma_x = \sigma_y \sigma_z \sigma_y = -\sigma_z$ , we find

$$\rho' = \left(1 - p + \frac{p}{3}\right) \frac{1}{2}(\mathbb{1} - P_z \sigma_z) + \frac{2p}{3} \frac{1}{2}(\mathbb{1} - P_z \sigma_z) = \frac{1}{2} \left[ \mathbb{1} + \left(1 - \frac{4}{3}p\right) P_z \sigma_z \right]$$

or  $P'_z = \left(1 - \frac{4}{3}p\right) P_z$ . From the rotational symmetry, we see that  $P' = \left(1 - \frac{4}{3}p\right)$  irrespective of the direction in which  $P$  points. Hence, the Bloch sphere contracts uniformly under the action of the channel; the spin polarization is reduced by the factor  $\left(1 - \frac{4}{3}p\right)$  (which is why we call it the depolarizing process). This result was to be expected in view of the observation above that the spin is totally “randomized” with probability  $\frac{4}{3}p$ .

Why do we say that the superoperator is not invertible? Evidently we can reverse a uniform contraction of the sphere with a uniform inflation. But the trouble is that the inflation of the Bloch sphere is not a superoperator, because it is not positive. Inflation will take values of  $P \leq 1$  to values  $P > 1$ , and so will take a density operator to an operator with a negative eigenvalue. Decoherence can shrink the ball, but no physical process can blow it up again! A superoperator running backwards in time is not a superoperator.

## 8.4 The Master Equation

### 8.4.1 Markov approximation

In the case of coherent evolution, we find it very convenient to characterize the dynamics of a quantum system with a Hamiltonian, which describes the evolution over an infinitesimal time interval. The dynamics is then described by a differential equation, the Schrödinger equation, and we may calculate the evolution over a finite time interval by integrating the equation, that is, by piecing together the evolution over many infinitesimal intervals. It is often possible to describe the (not necessarily coherent) evolution of a density matrix, at least to a good approximation, by a differential equation. This equation, the master equation, will be our next topic. In fact, it is not at all obvious that there need be a differential equation that describes decoherence. Such a description will be possible only if the evolution of the quantum system is “Markovian,” or in other words, local in time. If the evolution of the density operator  $\rho(t)$  is governed by a (first-order) differential equation in  $t$ , then that means that  $\rho(t + dt)$  is completely determined by  $\rho(t)$ .

In general the density operator  $\rho_A(t + dt)$  can depend not only on  $\rho_A(t)$ , but also on  $\rho_A$  at earlier times, because the environment (reservoir) retains a memory of this information for a while, and can transfer it back to system. An open system (whether classical or quantum) is dissipative because information can flow from the system to the reservoir. But that means that information can also flow back from reservoir to system, resulting in non-Markovian fluctuations of the system.

Still, in many contexts, a Markovian description is a very good approximation. The key idea is that there may be a clean separation between the typical correlation time of the fluctuations and the time scale of the evolution that we want to follow. Crudely speaking, we may denote by  $\delta t_E$  the time it takes for the reservoir to “forget” information that it acquired from the system. After time  $\delta t_E$  we can regard that information as forever lost, and neglect the possibility that the information may feed back again to influence the subsequent evolution of the system. Our description of the evolution of the system will incorporate “coarse-graining” in time; we perceive the dynamics through a filter that screens out the high frequency components of the motion, with  $\omega \gg 1/\delta t_{coarse}$ . An approximately Markovian description should be possible, then, if  $\delta t_E \ll \delta t_{coarse}$ ; we can neglect the memory of the reservoir, because we are unable to resolve its effects. This “Markovian approximation” will be useful if the time scale of the dynamics that we want to observe is long compared to  $\delta t_{coarse}$ , e.g., if the damping time scale  $\delta t_{damp}$  satisfies

$$\delta t_{damp} \gg \delta t_{coarse} \gg \delta t_E$$

### 8.4.2 Lindblad equation

Our goal is to generalize the Liouville equation  $\dot{\rho} = -i[\mathcal{H}, \rho]$  to the case of Markovian but non-unitary evolution, for which we will have  $\dot{\rho} = \mathcal{L}[\rho]$ . The linear operator  $\mathcal{L}$ , which generates a finite superoperator in the same sense that a Hamiltonian  $\mathcal{H}$  generates unitary time evolution, will be called the Lindbladian.

We can derive the Lindblad equation from an infinitesimal evolution described by the Kraus sum representation, with the following steps:

1. From the Kraus sum we can write the evolution of  $\rho$  from  $t$  to  $t + \delta t$  as:  $\rho(t + \delta t) = \sum_k M_k(\delta t)\rho(t)M_k^\dagger(\delta t)$ .
2. We now take the limit of infinitesimal time,  $\delta t \rightarrow 0$ . We only keep terms up to first order in  $\delta t$ ,  $\rho(t + \delta t) = \rho(t) + \delta t \delta\rho$ . This implies that the Kraus operator should be expanded as  $M_k = M_k^{(0)} + \sqrt{\delta t}M_k^{(1)} + \delta tM_k^{(2)} + \dots$ . Then there is one Kraus operator such that  $M_0 = \mathbb{1} + \delta t(-i\mathcal{H} + K) + O(\delta t^2)$  with  $K$  hermitian (so that  $\rho(t + \delta t)$  is hermitian), while all others have the form:  $M_k = \sqrt{\delta t}L_k + O(\delta t)$ , so that we ensure  $\rho(t + \delta t) = \rho(t) + \delta\rho\delta t$ :

$$\begin{aligned} \rho(t + \delta t) &= M_0\rho(t)M_0^\dagger + \sum_{k>0} M_k\rho M_k^\dagger \\ &= [\mathbb{1} + \delta t(-i\mathcal{H} + K)]\rho[\mathbb{1} + \delta t(i\mathcal{H} + K)] + \delta t \sum_k L_k\rho L_k^\dagger \\ &= \rho - i\delta t[\mathcal{H}, \rho] + \delta t(K\rho + \rho K) + \delta t \sum_k L_k\rho L_k^\dagger \end{aligned}$$

3.  $K$  and the other operators  $L_k$  are related to each other, since they have to respect the Kraus sum normalization condition,

$$K = -\frac{1}{2} \sum_{k>0} L_k^\dagger L_k.$$

4. Finally we substitute  $K$  in the equation above and take the limit  $\delta \rightarrow 0$ :  $\rho(t + dt) = \rho(t) + dt\dot{\rho}$ . We thus obtain the Lindblad equation

$$\dot{\rho}(t) = \mathcal{L}[\rho] = -i[\mathcal{H}, \rho(t)] + \sum_{k=1}^M \left( L_k \rho(t) L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho(t) - \frac{1}{2} \rho(t) L_k^\dagger L_k \right)$$

The first term in  $\mathcal{L}[\rho]$  is the usual Schrödinger term that generates unitary evolution (thus we identify the hermitian operator  $\mathcal{H}$  with the usual Hamiltonian). The other terms describe the possible transitions that the system may undergo due to interactions with the reservoir. The operators  $L_k$  are called Lindblad operators or quantum jump operators. Each  $L_k \rho L_k^\dagger$  term induces one of the possible quantum jumps, while the  $-\frac{1}{2} L_k^\dagger L_k \rho - \frac{1}{2} \rho L_k^\dagger L_k$  terms are needed to normalize properly the case in which no jumps occur.

If we recall the connection between the Kraus representation and the unitary representation of a superoperator, we clarify the interpretation of the master equation. We may imagine that we are continuously monitoring the reservoir, projecting it in each instant of time onto the  $|\mu\rangle_E$  basis. With probability  $1 - O(\delta t)$ , the reservoir remains in the state  $|0\rangle_E$ , but with probability of order  $\delta t$ , the reservoir makes a quantum jump to one of the states  $|\mu\rangle_E$ . When we say that the reservoir has “forgotten” the information it acquired from the system (so that the Markovian approximation applies), we mean that these transitions occur with probabilities that increase linearly with time.

This is equation is also called the Kossakowski-Lindblad equation<sup>26</sup>.

The Lindblad equation above is expressed in the Schrödinger picture. It is possible to derive the Heisenberg picture Lindblad equation, which has the form:

$$\frac{dx}{dt} = i[\mathcal{H}, x] + \sum_k L_k^\dagger x L_k - \frac{1}{2} L_k^\dagger L_k x + x L_k^\dagger L_k \quad ,$$

where  $x$  is the observable under study.

Another way to express the Lindblad equation is for a “vectorized” density matrix:  $\dot{\rho} = (\mathbf{H} + \mathcal{G})\rho$ , with the generator  $\mathcal{G}$ :

$$\mathcal{G} = \sum_{m=0}^M \bar{L}_m \otimes L_m - \frac{1}{2} \mathbb{1} \otimes (L_m^\dagger L_m) - \frac{1}{2} (\bar{L}_m^\dagger \bar{L}_m) \otimes \mathbb{1}$$

and the Hamiltonian part will be given by  $\mathbf{H} = -i(\mathcal{H} \otimes \mathbb{1} - \mathbb{1} \otimes \mathcal{H})$ . In this form, the Lindblad equation becomes a linear equation (a matrix multiplying a vector, if we are considering e.g. discrete systems). Thus it is “easy” to solve the differential equation, finding:

$$\rho(t) = \exp [(\mathbf{H} + \mathcal{G})t] \rho(0),$$

where we identify the superoperator  $S = \exp [(\mathbf{H} + \mathcal{G})t]$ . More details on how to convert from Kraus sum, to Lindblad to superoperator description of the open quantum system dynamics can be found in T. F. Havel, *Robust procedures for converting among Lindblad, Kraus and matrix representations of quantum dynamical semigroups*, *J. Math. Phys.* **44**, 534 (2003).

#### A. Example: spin-1/2 dephasing

Dephasing, or transverse relaxation, is the phenomenon associated with the decay of the coherence terms (off-diagonals) in the density matrix. In NMR, since the signal is due to the ensemble of spins, a coherence term which lasts forever would require all the same spins of the different molecules to precess about the magnetic field at exactly the same rate. As previously mentioned, the frequency of a single spin depends on the local magnetic field, which depends on the external field, and on the field created by the surrounding spins. Due to rapid tumbling, the average

<sup>26</sup> Andrzej Kossakowski *On quantum statistical mechanics of non-Hamiltonian systems*, Rep. Math. Phys. **3** 247 (1972)  
Göran Lindblad *On the generators of quantum dynamical semigroups*, Commun. Math. Phys. **48** 119 (1976)).

field over time is the same, but does vary across the sample at a particular given time. This instantaneous variation causes the identical spins of all the molecules to slowly desynchronize and therefore lose coherence across the sample. Another example of dephasing was already presented when we described phase-damping for a dust particle interacting with many photons.

The dephasing noise can be thought as arising from random  $z$  rotation across the sample, so that the state of the system can be described by a statistical average over a distribution of rotation angles  $q(\vartheta)$ :

$$S_d(\rho) = \int d\vartheta q(\vartheta) e^{-i\vartheta/2\sigma_z} \rho(0) e^{i\vartheta/2\sigma_z}$$

Consider an initial density operator

$$\rho(0) = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}$$

The evolution under a propagator  $U_\vartheta = e^{-i\vartheta/2\sigma_z}$  gives

$$\rho(\vartheta) = \begin{pmatrix} \rho_{00} & \rho_{01} e^{-i\vartheta} \\ \rho_{10} e^{i\vartheta} & \rho_{11} \end{pmatrix}$$

Taking the integral over the angle distribution we find

$$\rho(\vartheta) = \begin{pmatrix} \rho_{00} & \rho_{01} \Gamma \\ \rho_{10} \Gamma^* & \rho_{11} \end{pmatrix},$$

where  $\Gamma = \langle e^{-i\vartheta} \rangle = \int q(\vartheta) e^{-i\vartheta} d\vartheta$ . If  $q(\vartheta) = q(-\vartheta)$  (as given by an isotropic environment) we obtain  $\langle e^{-i\vartheta} \rangle = \langle \cos \vartheta \rangle$ . For a non-Markovian environment where memory effects are present, we can describe the distribution  $q(\vartheta)$  as a Gaussian stochastic process, so that  $\Gamma = \langle \cos \vartheta \rangle \approx e^{-\langle \vartheta^2 \rangle / 2} = e^{-t^2 / T_2^2}$ . For a Markovian process instead we have an exponential decay  $\Gamma = e^{-t/T_2}$ .

We can also explicitly evaluate  $S_d$ :

$$\begin{aligned} S_d(\rho) &= \int d\vartheta q(\vartheta) [\cos(\vartheta/2)\mathbb{1} - i \sin(\vartheta/2)\sigma_z] \rho(0) [\cos(\vartheta/2)\mathbb{1} + i \sin(\vartheta/2)\sigma_z] = \\ &= \int d\vartheta q(\vartheta) [\cos^2(\vartheta/2)\rho(0) + \sin^2(\vartheta/2)\sigma_z \rho(0) \sigma_z - i \cos(\vartheta/2) \sin(\vartheta/2) (\sigma_z \rho(0) - \rho(0) \sigma_z)] \end{aligned}$$

By evaluating the integral, and assuming again a symmetric distribution, we have:

$$S_d(\rho) = (1 - p)\rho(0) + p\sigma_z \rho(0) \sigma_z$$

where  $p = \int d\vartheta q(\vartheta) \sin^2(\vartheta/2)$ . By comparison with the previous result we find  $p = \frac{1-\Gamma}{2}$ . From the superoperator, we can find the corresponding Kraus sum decomposition:

$$M_0 = \sqrt{1-p}\mathbb{1}, \quad M_1 = \sqrt{p}\sigma_z$$

We want now to describe this same evolution under a dephasing environment by a Lindblad equation. Notice that this is going to be possible only if we have a Markovian environment,  $\Gamma = e^{-t/T_2}$ .

Consider the action of the superoperator  $S_d(\rho) = \frac{1+\Gamma}{2}\rho(0) + \frac{1-\Gamma}{2}\sigma_z \rho(0) \sigma_z$ . If we consider a small time  $\Gamma = e^{-\delta t/T_2} \approx 1 - \delta t/T_2$  and we obtain:

$$S_d(\rho, \delta t) = \rho - \frac{\gamma \delta t}{2} \rho + \frac{\gamma \delta t}{2} \sigma_z \rho \sigma_z$$

where  $\gamma = 1/T_2$ . Then, taking the difference  $\rho(\delta t) - \rho(0)$  in the limit  $\delta t \rightarrow 0$  we have

$$\frac{\partial \rho}{\partial t} = \frac{\gamma}{2} (\sigma_z \rho \sigma_z - \rho) = \frac{\gamma}{2} (\sigma_z \rho \sigma_z - \frac{1}{2} \{\sigma_z \sigma_z, \rho\})$$

where we used the fact  $\sigma_z^2 = \mathbb{1}$ . Thus  $\sqrt{\frac{\gamma}{2}} \sigma_z$  is the Lindblad operator for dephasing.

Assume now that we had considered a non-Markovian environment, for which  $\Gamma = e^{-(t/T_2)^2}$ . Then if we tried to find the infinitesimal time evolution, we cannot define a differential equation, since  $\rho(\delta t) - \rho(0)$  is not  $\propto \delta t$ . For this type of environment, the Lindblad equation cannot be defined.

### 8.4.3 Redfield-Born theory of relaxation

Consider a system  $\mathcal{S}$  coupled to an environment  $\mathcal{E}$  (the heat bath) such that

$$\mathcal{H} = \mathcal{H}_0 + V = \mathcal{H}_S + \mathcal{H}_E + V,$$

and  $V$  describes the interaction between the system and the environment. Most generally it will take the form  $V = \sum_k A_k \otimes B_k(t)$ , with  $A$  acting on the system and  $B$  on the environment (and we have even allowed for a time-dependence of the random environment field). In the Schrödinger picture, the time evolution of the density matrix is given by the Liouville equation,  $i\hbar \frac{d\rho(t)}{dt} = [\mathcal{H}, \rho(t)]$ . Define the interaction picture density matrix

$$\rho_I(t) \equiv e^{\frac{i}{\hbar}(\mathcal{H}_S + \mathcal{H}_E)t} \rho(t) e^{-\frac{i}{\hbar}(\mathcal{H}_S + \mathcal{H}_E)t},$$

and similarly the interaction-picture system-environment interaction

$$V_I(t) \equiv e^{\frac{i}{\hbar}(\mathcal{H}_S + \mathcal{H}_E)t} V e^{-\frac{i}{\hbar}(\mathcal{H}_S + \mathcal{H}_E)t}.$$

Then the evolution in the interaction picture is given by

$$i\hbar \frac{d\rho_I(t)}{dt} = e^{\frac{i}{\hbar}\mathcal{H}_0 t} ([\mathcal{H}, \rho(t)] - [\mathcal{H}_0, \rho(t)]) e^{-\frac{i}{\hbar}\mathcal{H}_0 t} = e^{\frac{i}{\hbar}\mathcal{H}_0 t} [V, \rho(t)] e^{-\frac{i}{\hbar}\mathcal{H}_0 t} = [V_I(t), \rho_I(t)].$$

This has the formal solution

$$\rho_I(t) = \rho_I(0) + \frac{1}{i\hbar} \int_0^t dt_1 [V_I(t_1), \rho_I(t_1)]$$

(Note that this is the same equation as above, except in integral form).

Expanding once (by inserting the same equation at the place of  $\rho_I(t)$ ) we obtain,

$$\rho_I(t) = \rho_I(0) + \frac{1}{i\hbar} \int_0^t dt_1 [V_I(t_1), \rho_I(0)] + \frac{1}{(i\hbar)^2} \int_0^t \int_0^{t_1} dt_1 dt_2 [V_I(t_1), [V_I(t_2), \rho_I(t_2)]]$$

We could repeat this process to obtain an infinite series (the Dyson series we already saw).

Let us concentrate instead on the evolution of the (interaction picture) reduced density matrix  $\rho_S = \text{Tr}_{\mathcal{E}} \{\rho_I\}$ , obtained by tracing over the environment. To obtain the average density operator, we also need to take an ensemble average over the random fluctuating environment:

$$\rho_S(t) = \rho_S(0) + \frac{1}{i\hbar} \int_0^t dt_1 \langle \text{Tr}_{\mathcal{E}} \{ [V_I(t_1), \rho_I(0)] \} \rangle + \frac{1}{(i\hbar)^2} \int_0^t \int_0^{t_1} dt_1 dt_2 \langle \text{Tr}_{\mathcal{E}} \{ [V_I(t_1), [V_I(t_2), \rho_I(t_2)]] \} \rangle.$$

We want to find an explicit expression for the system evolution only (in the form of a differential equation). To do this, we will make a number of approximations.

#### A. Simplification: Separability and energy shift

We first assume that at time  $t = 0$  the system and environment are in a separable state:

$$\rho(0) = \rho_S(0) \otimes \rho_E(0).$$

(this can always be obtained by choosing  $t = 0$  appropriately).

This condition helps simplifying the second term in the LHS of the expression above. We have

$$\text{Tr}_{\mathcal{E}} \{ [V_I(t_1), \rho_I(0)] \} = \sum_k [A_I(t_1), \rho_S(0)] \text{Tr}_{\mathcal{E}} \{ B_k(t) \rho_E \},$$

that is, we consider the expectation value of the operators  $B_k$ . In general we will also need to take an ensemble average over the random fluctuating field  $\langle B_k(t) \rangle$ , as we look at expectation values for the density operator.

We can now make the assumption that  $\langle B_{\mathcal{E}} \rangle_{\mathcal{E}} = 0$ , which implies  $\langle V_I(t) \rangle_{\mathcal{E}} \sim \text{Tr}_{\mathcal{E}} \{V_I(t) \rho_{\mathcal{E}}(0)\} = 0$ . This is not restrictive, since, if  $V$  is of the form  $V = A_S \otimes B_{\mathcal{E}}$  with  $\langle B_{\mathcal{E}} \rangle_{\mathcal{E}} \neq 0$  we can replace  $V$  with  $V = A_S \otimes (B_{\mathcal{E}} - \langle B_{\mathcal{E}} \rangle_{\mathcal{E}})$ , and simultaneously add  $A_S \langle B_{\mathcal{E}} \rangle_{\mathcal{E}}$  to  $\mathcal{H}_S$ . With this condition,  $\langle V \rangle_{\mathcal{E}} = 0$  and since  $\rho_{\mathcal{E}}(0)$  has the same form in both Schrödinger and interaction pictures, the result holds in the interaction picture also. The same argument can be made if  $V = \sum_{\alpha} A_{S,\alpha} \otimes B_{\mathcal{E},\alpha}$ . Then the second term in the equation above vanishes and we have

$$\rho_S(t) = \rho_S(0) + \frac{1}{(i\hbar)^2} \int_0^t \int_0^{t_1} dt_1 dt_2 \langle \text{Tr}_{\mathcal{E}} \{ [V_I(t_1), [V_I(t_2), \rho_I(t_2)]] \} \rangle.$$

### B. Assumption 1: Born approximation

We can always write (in any picture)  $\rho(t) = \rho_S(t) \otimes \rho_{\mathcal{E}}(t) + \rho_{\text{correlation}}(t)$ , which may be taken as a definition of  $\rho_{\text{correlation}}$ . Let us assume (as done in the previous section) that the interaction is turned on at time  $t = 0$ , and that prior to that the system and environment are not correlated ( $\rho_{\text{correlation}}(0) = 0$ ). This assumption is not very restrictive, since we can always find a time prior to which the system and environment did not interact. Now however we make a stronger assumption.

We will assume that the coupling between the system and the environment is weak, so that  $\rho(t) \approx \rho_S(t) \otimes \rho_{\mathcal{E}}(t)$ , for timescales over which perturbation theory remains valid. Furthermore, we will assume that the correlation time  $\tau_{\mathcal{E}}$  (and thus the relaxation time) of the environment is sufficiently small that  $\rho_{\mathcal{E}}(t) \approx \rho_{\mathcal{E}}(0)$  if  $t \gg \tau_{\mathcal{E}}$ .

Note that since we assume that the environment is in a thermal equilibrium, it has a thermal density matrix

$$\rho_{\mathcal{E}}(0) = \frac{1}{Z_{\mathcal{E}}} \sum_n e^{-\frac{E_n}{k_B T}} |n_{\mathcal{E}}\rangle \langle n_{\mathcal{E}}|,$$

which is a stationary state, i.e.,  $[\rho_{\mathcal{E}}(0), \mathcal{H}_{\mathcal{E}}] = 0$ , so that  $\rho_{\mathcal{E}}(0)$  has the same form in both the interaction picture and Schrödinger picture. Then

$$\rho_S(t) = \rho_S(0) + \frac{1}{(i\hbar)^2} \int_0^t \int_0^{t_1} dt_1 dt_2 \langle \text{Tr}_{\mathcal{E}} \{ [V_I(t_1), [V_I(t_2), \rho_S(t_2) \otimes \rho_{\mathcal{E}}(0)]] \} \rangle.$$

We can also go further and explicitly write the partial trace:

$$\langle \text{Tr}_{\mathcal{E}} \{ [V_I(t_1), [V_I(t_2), \rho_S(t_2) \otimes \rho_{\mathcal{E}}(0)]] \} \rangle = \sum_{k,h} \langle B_k(t_1) B_h(t_2) \rangle [A_k^I(t_1), [A_h^I(t_2), \rho_S(t_2)]]$$

where  $\langle B_k(t_1) B_h(t_2) \rangle = G_{k,h}(t_1, t_2)$  is the correlation function for the environment.

Differentiating, we get

$$\frac{d}{dt} \rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^t ds \langle \text{Tr}_{\mathcal{E}} \{ [V_I(t), [V_I(s), \rho_S(s) \otimes \rho_{\mathcal{E}}(0)]] \} \rangle.$$

or

$$\frac{d}{dt} \rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^t ds \sum_{k,h} \langle B_k(t) B_h(s) \rangle [A_k^I(t), [A_h^I(s), \rho_S(s)]].$$

This should properly be considered a difference equation, since we have assumed that  $t \gg \tau_{\mathcal{E}}$ .

### C. Assumption 2: Markov approximation

We will also assume that we are working over timescales that are shorter than the gross timescale over which the system evolves, so that  $\rho_S(s) \approx \rho_S(t)$ . Thus we can replace  $\rho_S(s)$  in the integral with  $\rho_S(t)$ . We finally get the Redfield equation:

$$\frac{d}{dt} \rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^t ds \text{Tr}_{\mathcal{E}} \{ [V_I(t), [V_I(s), \rho_S(t) \otimes \rho_{\mathcal{E}}(0)]] \}$$

or

$$\frac{d}{dt}\rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^t ds \sum_{k,h} \langle B_k(t)B_h(s) \rangle [A_k^I(t), [A_h^I(s), \rho_S(t)]] .$$

We can change variables from  $s \rightarrow s' = t - s$  (so that we change the integrals as:  $\int_0^t ds \rightarrow \int_t^0 d(t - s') = -\int_t^0 ds' = \int_0^t ds'$ ). Then

$$\frac{d}{dt}\rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^t ds \sum_{k,h} \langle B_k(t)B_h(t-s) \rangle [A_k^I(t), [A_h^I(t-s), \rho_S(t)]] .$$

The correlation time of the thermal bath  $\mathcal{E}$  is assumed to be very short, so that the correlation function  $\langle B_k(t_1 - t_2)B_h(0) \rangle_{\mathcal{E}}$  differs only significantly from zero when  $t_1 \approx t_2$ . We can therefore extend the limit of integration to  $\infty$  (and call  $t - s = \tau$ ):

$$\frac{d}{dt}\rho_S(t) = \frac{1}{(i\hbar)^2} \int_0^\infty d\tau \sum_{k,h} \langle B_k(t)B_h(\tau) \rangle [A_k^I(t), [A_h^I(\tau), \rho_S(t)]] .$$

## D. Spectral densities

The next step in the simplification program is to take the expectation values with respect to the eigenstates of the system and then Fourier transform. We will write  $A_k(t) = \sum_p A_k^p e^{-i\omega_p t}$ :

$$\frac{d}{dt}\rho_S(t) = \frac{1}{(i\hbar)^2} \sum_{k,h} \sum_{p,q} \int_0^\infty d\tau G_{kh}(\tau) [A_k^p e^{-i\omega_p t}, [A_h^q e^{-i\omega_q(t-\tau)}, \rho_S(t)]] .$$

Here we used the fact that  $G(t, \tau)$  is stationary, and thus depend only on the difference  $t - \tau$ ,  $G(t, \tau) = G(t - \tau)$ . We then changed variables from  $\tau \rightarrow t - \tau$ . We can rewrite the equation as

$$\frac{d}{dt}\rho_S(t) = \frac{1}{(i\hbar)^2} \sum_{k,h} \sum_{p,q} [A_k^p, [A_h^q, \rho_S(t)]] e^{-i(\omega_p + \omega_q)t} \int_0^\infty d\tau G_{kh}(\tau) e^{i\omega_q \tau} .$$

Thus we have the integral  $\int_0^\infty e^{i\omega\tau} G(\tau) = J(\omega)$ , where the Fourier transform of the correlation function  $G(\tau)$  is the the spectral function  $J(\omega)$ . With some simplifications (due to statistical properties of the bath operators and to the fact that we only take terms resulting in an Hermitian operator), we finally arrive at the master equation:

$$\frac{d}{dt}\rho_S(t) = - \sum_k \sum_p J_k(\omega_p) [A_{-k}^p, [A_k^p, \rho_S(t)]]$$

We can also write the master equation as the Redfield equation (subscripts indicate matrix elements):

$$\frac{d}{dt}\rho_{a,a'} = \sum_{b,b'/b-b'=a-a'} R_{aa',bb'} \rho_{b,b'}$$

## 8.5 Other description of open quantum system dynamics

### 8.5.1 Stochastic Liouville equation and cumulants

Stochastic Liouville theory is based on a semiclassical model of decoherence, in which the Hamiltonian at any instant in time consists of a deterministic and a stochastic part, which represents the effects of a random noise. In the simplest case of NMR  $T_2$  relaxation, this typically takes the form

$$\mathcal{H}_{tot}(t) = \mathcal{H}_{det}(t) + \mathcal{H}_{st}(t) = \mathcal{H}_{det}(t) + \omega(t)\mathcal{H}_N,$$

where  $\mathcal{H}_{det}$  is the static deterministic Hamiltonian, and we separated the stochastic time dependence described by the coefficient  $\omega(t)$  from the noise generator  $\mathcal{H}_N$ .  $\omega(t)$  is a random variable due to stochastic, time-dependent fluctuating fields and  $\mathcal{H}_N$  is an operator which describes how these classical fields are coupled to the quantum system. We now introduce a superoperator  $\mathcal{L}(t)$  defined on Liouville (operator) space via

$$\mathcal{L}(t) = \mathcal{H}_{tot}^*(t) \otimes \mathbb{1} - \mathbb{1} \otimes \mathcal{H}_{tot}(t) = \mathcal{L}_{det}(t) + \omega(t)\mathcal{L}_N$$

where  $\mathcal{L}_N = \mathcal{H}_N^* \otimes \mathbb{1} - \mathbb{1} \otimes \mathcal{H}_N$ . This superoperator is the generator of motion for density operator  $\hat{\rho}$ , meaning

$$\rho(t) = \mathcal{U}\hat{\rho}(0) = \mathcal{T} \exp\left(-i \int_0^t dt' \mathcal{L}(t')\right) \hat{\rho}(0)$$

where  $\mathcal{T}$  is the usual Dyson time ordering operator. Since what is actually observed in an experiment is the statistical average over the microscopic trajectories of the system  $\langle \hat{\rho}(t) \rangle$ , we have to take the ensemble average superpropagator to obtain  $\langle \hat{\rho}(t) \rangle = \langle \mathcal{U} \rangle \hat{\rho}(0)$ . The problem of calculating the average of the exponential of a stochastic operator has been solved by Kubo<sup>27</sup> using the cumulant expansion.

First, expand the time-ordered average exponential  $S = \langle \mathcal{T} \exp(-i \int_0^t dt' \mathcal{H}(t')) \rangle$  via the Dyson series:

$$\begin{aligned} S &= \mathbb{1} - i \int_0^t dt' \langle \mathcal{H}(t') \rangle + \frac{(-i)^2}{2!} \mathcal{T} \int_0^t dt_1 \int_0^t dt_2 \langle \mathcal{H}(t_1) \mathcal{H}(t_2) \rangle + \dots \\ &\quad + \frac{(-i)^n}{n!} \mathcal{T} \int_0^t dt_1 \dots \int_0^t dt_n \langle \mathcal{H}(t_1) \dots \mathcal{H}(t_n) \rangle + \dots \end{aligned}$$

The term  $\langle \mathcal{H}(t_1) \dots \mathcal{H}(t_n) \rangle$  is called the  $n$ -th moment of the distribution. We want now to express this same propagator in terms of the cumulant function  $K(t)$ , defined by:

$$S = e^{K(t)}$$

The cumulant function itself can most generally be expressed as a power series in time:

$$K(t) = \sum_{n=1}^{\infty} \frac{(-it)^n}{n!} K_n = -itK_1 + \frac{(-it)^2}{2!} K_2 + \dots$$

Expanding now the exponential using the expression above we have:

$$S = \mathbb{1} + K(t) + \frac{1}{2!} (K(t))^2 + \dots = \mathbb{1} - itK_1 + \frac{(-it)^2}{2!} (K_2 + K_1^2) + \dots$$

By equating terms of the same order in the two expansions we obtain the cumulants  $K_n$  in terms of the moments of order at most  $n$ . For example:

$$\begin{aligned} K_1 &= \frac{1}{t} \int_0^t dt' \langle \mathcal{H}(t') \rangle \\ K_2 &= \frac{1}{t^2} \mathcal{T} \int_0^t dt_1 \int_0^t dt_2 \langle \mathcal{H}(t_1) \mathcal{H}(t_2) \rangle - K_1^2 \end{aligned}$$

The propagator can therefore be expressed in terms of the cumulant averages:

$$\begin{aligned} \langle \mathcal{H}(t') \rangle_c &= \langle \mathcal{H}(t') \rangle \\ \langle \mathcal{H}(t_1) \mathcal{H}(t_2) \rangle_c &= \mathcal{T} \langle \mathcal{H}(t_1) \mathcal{H}(t_2) \rangle - \langle \mathcal{H}(t_1) \rangle \langle \mathcal{H}(t_2) \rangle \end{aligned}$$

The propagator can therefore be written as:

$$S = \exp\left(-i \int_0^t dt' \langle \mathcal{H}(t') \rangle_c - \int_0^t dt_1 \int_0^t dt_2 \langle \mathcal{H}(t_1) \mathcal{H}(t_2) \rangle_c + \dots\right)$$

---

<sup>27</sup> R. Kubo, *Generalized Cumulant Expansion Method*, [Journal of the Physical Society of Japan](#), **17**, 1100-1120 (1962)

Note that if  $\mathcal{H}$  is a deterministic function of time, the ensemble averages can be dropped and  $\langle \mathcal{H}(t) \rangle_c = \int_0^t dt' \mathcal{H}(t')$  becomes the time-average Hamiltonian, which is the first term in the Magnus expansion. The second term in the cumulant expansion, on the other hand, becomes

$$\begin{aligned} & \mathcal{T} \int_0^t dt_1 \int_0^{t_1} dt_2 \mathcal{H}(t_1) \mathcal{H}(t_2) - \left( \int_0^t dt' \mathcal{H}(t') \right)^2 \\ &= 2 \int_0^t dt_1 \int_0^{t_1} dt_2 \mathcal{H}(t_1) \mathcal{H}(t_2) - \int_0^t dt_1 \int_0^t dt_2 \mathcal{H}(t_1) \mathcal{H}(t_2) \\ &= \int_0^t dt_1 \int_0^{t_1} dt_2 \mathcal{H}(t_1) \mathcal{H}(t_2) - \int_0^t dt_1 \int_{t_1}^t dt_2 \mathcal{H}(t_1) \mathcal{H}(t_2) \\ &= \int_0^t dt_1 \int_0^{t_1} dt_2 [\mathcal{H}(t_1), \mathcal{H}(t_2)], \end{aligned}$$

where we have used the fact that the time-ordering operator  $\mathcal{T}$  symmetrizes its argument with respect to permutation of the time points. This is the second term in the Magnus expansion for the “average” (effective) Hamiltonian. Proceeding in this fashion one can in principle derive average Hamiltonian theory<sup>28</sup> from the Dyson and cumulant expansions.

In terms of the so-called “cumulant averages”  $\langle \dots \rangle_c$ , the superpropagator is given by:

$$\langle \mathcal{U} \rangle = \exp \left( -i \int_0^t dt' \langle \mathcal{L}(t') \rangle_c - \frac{1}{2} \mathcal{T} \int_0^t dt_1 \int_0^t dt_2 \langle \mathcal{L}(t_1) \mathcal{L}(t_2) \rangle_c + \dots \right)$$

Provided  $\| \int_0^t dt' \mathcal{L}(t') \| \ll 1$  for all  $t > 0$ , we can safely neglect high order terms in the exponential’s argument.

## 8.5.2 Stochastic Wavefunctions

The Monte Carlo wavefunction was derived simultaneously in the 1990s by two groups interested in very different questions. A group of scientists in France, Dalibard, Castin, and Mølmer, wanted to simulate laser cooling of atoms quantum mechanically in three dimensions. Their numerical solution required discretizing space into a grid of  $40 \times 40 \times 40$  positions; to implement the master equation on such a space would have required a density matrix with  $O(40^6) \sim 10^9$  entries such calculations are beyond the scope of even modern computers. However, simulating a wavefunction with  $O(40^3)$  entries is quite feasible. Consequently the group sought to convert the master equation to something more like the Schrödinger equation<sup>29</sup>.

At the same time, Carmichael was interested in the effects that continuous monitoring would have on a system<sup>30</sup>. For example, a two-level atom prepared in an equal superposition of states can decay by emitting a photon; if that photon is detected, the experimenter knows with certainty that the atom is in its ground state. But what happens 50% of the time when a photon is not detected? Certainly, after a long time has passed, the atom must be in its ground state, but how does that happen? To study these and similar questions, Carmichael wanted to incorporate the effects of continuous monitoring, and understand how a measurement can cause the system state to suddenly jump into a different state.

The description on which both groups converged begins with the most general form of the master equation,

$$\frac{d\rho}{dt} = -i[\mathcal{H}, \rho] + \mathcal{L}(\rho),$$

with the Lindbladian

$$\mathcal{L}(\rho) = - \sum_k \frac{\gamma_k}{2} (L_k^\dagger L_k \rho + \rho L_k^\dagger L_k - 2L_k \rho L_k^\dagger).$$

<sup>28</sup> See for example Haeberlen, *High Resolution NMR in Solids: Selective Averaging*, Academic Press Inc., New York (1976)

<sup>29</sup> Jean Dalibard, Yvan Castin and Klaus Mlmer *Wave-function approach to dissipative processes in quantum optics*, *Phys. Rev. Lett.* **68**, 580583 (1992)

<sup>30</sup> H. J. Carmichael *Quantum trajectory theory for cascaded open systems* *Phys. Rev. Lett.* **70**, 22732276 (1993)

Using this explicit expression and rearranging the terms we have

$$\begin{aligned}\frac{d\rho}{dt} &= -i \left[ (\mathcal{H} - i \sum_k \frac{\gamma_k}{2} L_k^\dagger L_k) \rho - \rho (\mathcal{H} + \sum_k \frac{\gamma_k}{2} L_k^\dagger L_k) \right] + \sum_k \gamma_k L_k \rho L_k^\dagger \\ &= -i \mathcal{H}_{\text{eff}} \rho - \rho \mathcal{H}_{\text{eff}}^\dagger + \sum_k \gamma_k L_k \rho L_k^\dagger,\end{aligned}$$

where we have defined an effective *Hamiltonian*

$$\mathcal{H}_{\text{eff}} = \mathcal{H} - i \sum_k \frac{\gamma_k}{2} L_k^\dagger L_k$$

(notice that this is not a valid Hamiltonian in the usual sense, since it is not Hermitian, so its eigenvalues are not the energy, since they could be imaginary numbers).

Expanding the density matrix in terms of an ensemble of pure states,  $\rho = \sum_j p_j |\psi_j\rangle \langle \psi_j|$ , we can rewrite the master equation in a suggestive form:

$$\frac{d\rho}{dt} = \sum_j p_j \left[ -i(\mathcal{H}_{\text{eff}} |\psi_j\rangle \langle \psi_j| - |\psi_j\rangle \langle \psi_j| \mathcal{H}_{\text{eff}}^\dagger) + \sum_k \gamma_k L_k |\psi_j\rangle \langle \psi_j| L_k^\dagger \right]$$

Now we can interpret the first two terms of this equation as a Schrödinger evolution for each of the pure states in the density matrix expansion:

$$\frac{d}{dt} |\psi_j\rangle = -i \mathcal{H}_{\text{eff}} |\psi_j\rangle$$

while we interpret the last term as a *quantum jump* operator that changes  $|\psi_j\rangle$  into  $|\varphi_{j,k}\rangle = L_k |\psi_j\rangle$  with some probability.

We can then have a probabilistic picture of the pure state evolution. After an infinitesimal time, in the absence of *jumps*, the state will have evolved to

$$|\psi_j(t + \delta t)\rangle = (1 - i \mathcal{H}_{\text{eff}}) |\psi_j\rangle / \sqrt{1 - \delta p_j},$$

where we have introduced a normalization factor, which is needed since the Hamiltonian is not hermitian:

$$\delta p_j = \sum_k \delta p_{j,k} = \delta t \sum_k \gamma_k \langle \psi_j | L_k^\dagger L_k | \psi_j \rangle$$

If instead a jump has occurred, the state would have evolved to

$$|\varphi_{j,k}\rangle = \sqrt{\frac{\gamma_k \delta t}{\delta p_{j,k}}} L_k |\psi_j\rangle$$

Thus the evolution of the density matrix is given by

$$\rho(t + \delta t) = \sum_j p_j \left[ (1 - \delta p_j) |\psi_j(t + \delta t)\rangle \langle \psi_j(t + \delta t)| + \sum_k \delta p_{j,k} |\varphi_{j,k}\rangle \langle \varphi_{j,k}| \right]$$

This expression leads us to the following interpretation: the system undergoes a dynamics that yields two possible outcomes:

1. with probability  $1 - \delta p_j$  the system evolves to the state  $|\psi_j(t + \delta t)\rangle$ , according to the operator  $\mathcal{H}_{\text{eff}}$  with an appropriate normalization
2. with probability  $\delta p_j$  the system jumps to another state. There are many possible states the system can jump to, each one with a probability  $\delta p_{j,k}$ .

This probabilistic picture is of course a coarse graining of the continuous time evolution. However, by discretizing time it becomes easier to devise a simulation procedure to reproduce the desired dynamics, with a wavefunction Montecarlo procedure.



# 9. Harmonic Oscillator

## 9.1 Harmonic Oscillator

- 9.1.1 Classical harmonic oscillator and h.o. model
- 9.1.2 Oscillator Hamiltonian: Position and momentum operators
- 9.1.3 Position representation
- 9.1.4 Heisenberg picture
- 9.1.5 Schrödinger picture

## 9.2 Uncertainty relationships

## 9.3 Coherent States

- 9.3.1 Expansion in terms of number states
- 9.3.2 Non-Orthogonality
- 9.3.3 Uncertainty relationships
- 9.3.4 X-representation

## 9.4 Phonons

- 9.4.1 Harmonic oscillator model for a crystal
- 9.4.2 Phonons as normal modes of the lattice vibration
- 9.4.3 Thermal energy density and Specific Heat

## 9.1 Harmonic Oscillator

We have considered up to this moment only systems with a finite number of energy levels; we are now going to consider a system with an infinite number of energy levels: the quantum harmonic oscillator (h.o.).

The quantum h.o. is a model that describes systems with a characteristic energy spectrum, given by a ladder of evenly spaced energy levels. The energy difference between two consecutive levels is  $\Delta E$ . The number of levels is infinite, but there must exist a minimum energy, since the energy must always be positive. Given this spectrum, we expect the Hamiltonian will have the form

$$\mathcal{H} |n\rangle = \left( n + \frac{1}{2} \right) \hbar\omega |n\rangle,$$

where each level in the ladder is identified by a number  $n$ . The name of the model is due to the analogy with characteristics of classical h.o., which we will review first.

### 9.1.1 Classical harmonic oscillator and h.o. model

A classical h.o. is described by a potential energy  $V = \frac{1}{2}kx^2$ . If the system has a finite energy  $E$ , the motion is bound by two values  $\pm x_0$ , such that  $V(x_0) = E$ . The equation of motion is given by  $m \frac{d^2x}{dt^2} = -kx$  and the kinetic energy is of course  $T = \frac{1}{2}m\dot{x}^2 = \frac{p^2}{2m}$ . The energy is constant since it is a conservative system, with no dissipation. Most of the time the particle is in the position  $x_0$  since there the velocity is zero, while at  $x = 0$  the velocity is maximum.

The h.o. oscillator in QM is an important model that describes many different physical situations. We will study in depth a particular system described by the h.o., the electromagnetic field. Another system that can be described

by this model is solid-state crystals, where the oscillations of nuclei in the lattice can be described as a systems of coupled oscillators (phonons).

Notice that any potential with a local minimum can be locally described by an h.o.. Provided that the energy is low enough (or  $x$  close to  $x_0$ ), any potential can in fact be expanded in series, giving:  $V(x) \approx V(x_0) + b(x - x_0)^2 + \dots$  where  $b = \frac{d^2V}{dx^2}|_{x_0}$ .

### 9.1.2 Oscillator Hamiltonian: Position and momentum operators

We can define the operators associated with position and momentum. They are two observables ( $p, x$ ) with the commutation properties:  $[x, p] = i\hbar$ . With these two operators, the Hamiltonian of the quantum h.o. is written as:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{kx^2}{2} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2,$$

where we defined a parameter with units of frequency:  $\omega = \sqrt{k/m}$ .

We use the dimensionless variables,

$$P = \frac{p}{\sqrt{m\omega}}, \quad X = x\sqrt{m\omega}$$

and  $\hat{\mathcal{H}} = \mathcal{H}/\omega$ , to simplify the expression to  $\hat{\mathcal{H}} = \omega(X^2 + P^2)/2$  or  $\mathcal{H} = \frac{\omega}{2}(X^2 + P^2)$ .

We have said initially that we expect the hamiltonian to have the form  $\hat{\mathcal{H}} = (n + \frac{1}{2})|n\rangle\langle n|$ , if expressed in an appropriate basis. This simply corresponds to diagonalizing the Hamiltonian (thus the basis  $\{|n\rangle\}$  is the energy basis, or the basis formed by the eigenstates of the Hamiltonian). However the diagonalization is not as intuitive as for simple TLS (or n TLS) because we are considering a system with infinite dimensions.

The operator  $\hat{\mathcal{H}} = (n + \frac{1}{2})|n\rangle\langle n|$  is our guess for the diagonalized form of the Hamiltonian, which makes explicit the presence of energy levels, labeled by  $n$ . Correspondingly, there must be operators that act on the ladder of energy levels. For example, the fundamental operations possible are the raising or lowering of 1 quantum of energy, as well as an operator giving the number of energy quanta:  $N|n\rangle = n|n\rangle$ . The raising and lowering operators act as the following:  $a|n\rangle \propto |n-1\rangle$  and  $a^\dagger|n\rangle \propto |n+1\rangle$ . They are also called the annihilation and creation operators, as they destroy or create a quantum of energy.

Instead of deriving rigorously these operators, we guess their form in terms of the  $X$  and  $P$  operators:

$$\begin{aligned} a &= \frac{1}{\sqrt{2\hbar}}(X + iP) = \frac{1}{\sqrt{2\hbar}}(\sqrt{m\omega}x + \frac{i}{\sqrt{m\omega}}p) \\ a^\dagger &= \frac{1}{\sqrt{2\hbar}}(X - iP) = \frac{1}{\sqrt{2\hbar}}(\sqrt{m\omega}x - \frac{i}{\sqrt{m\omega}}p), \end{aligned}$$

and we will check a posteriori that indeed they act as annihilation and creation operators. Notice that  $a, a^\dagger$  are not hermitian, but they are one the hermitian conjugate of the other ( $a = (a^\dagger)^\dagger$ ). Also, we define the number operator as  $N = a^\dagger a$ . The commutation properties are:  $[a, a^\dagger] = 1$  and  $[N, a] = -a$ ,  $[N, a^\dagger] = a^\dagger$ .

Also we have:

$$\begin{aligned} x &= \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a) \\ p &= i\sqrt{\frac{m\omega\hbar}{2}}(a^\dagger - a) \end{aligned}$$

---

**? Question:** Prove the commutation relationships of the raising and lowering operators.

$$[a, a^\dagger] = \frac{1}{2\hbar}[X + iP, X - iP] = \frac{1}{2\hbar}([X, -iP] + [iP, X]) = -\frac{i}{\hbar}[X, P] = -\frac{i}{\hbar}[x, p] = 1$$

So we also have  $aa^\dagger = [a, a^\dagger] + a^\dagger a = 1 + a^\dagger a = 1 + N$ .

$$[N, a] = [a^\dagger a, a] = [a^\dagger, a]a = -a \quad \text{and} \quad [N, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger[a, a^\dagger] = a^\dagger$$

□

---

Notice that from now on we will take as usual  $\hbar = 1$ .

From the commutation relationships we have:

$$a|n\rangle = [a, N]|n\rangle = an|n\rangle - Na|n\rangle \rightarrow N(a|n\rangle) = (n-1)(a|n\rangle),$$

that is,  $a|n\rangle$  is also an eigenvector of the  $N$  operator, with eigenvalue  $(n-1)$ . Thus we confirm that this is the lowering operator:  $a|n\rangle = c_n|n-1\rangle$ . Similarly,  $a^\dagger|n\rangle$  is an eigenvector of  $N$  with eigenvalue  $n+1$ :

$$a^\dagger|n\rangle = [N, a^\dagger]|n\rangle = Na^\dagger|n\rangle - a^\dagger n|n\rangle \rightarrow N(a^\dagger|n\rangle) = (n+1)(a^\dagger|n\rangle).$$

We thus have  $a|n\rangle = c_n|n-1\rangle$  and  $a^\dagger|n\rangle = d_n|n+1\rangle$ . What are the coefficients  $c_n$ ,  $d_n$ ?  
Since

$$\langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = n$$

and

$$\langle n|a^\dagger a|n\rangle = (\langle an|)(a|n\rangle) = \langle n-1|n-1\rangle c_n^2,$$

we must have  $c_n = \sqrt{n}$ . Analogously, since  $aa^\dagger = N+1$ , as seen from the commutation relationship:

$$d_n^2 \langle n+1|n+1\rangle = \langle a^\dagger n|a^\dagger n\rangle = \langle n|aa^\dagger|n\rangle \langle n|(N+1)|n\rangle = n+1$$

So in the end we have :

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

All the  $n$  eigenvalues of  $N$  have to be non-negative since  $n = \langle n|N|n\rangle = \langle \psi_{n_1}|\psi_{n_1}\rangle \geq 0$  (this follows from the properties of the inner product and the fact that  $|\psi_{n_1}\rangle = a|n\rangle$  is just a regular state vector). However, if we apply over and over the  $a$  (lowering) operator, we could arrive at negative numbers  $n$ : we therefore require that  $a|0\rangle = 0$  to truncate this process. The action of the raising operator  $a^\dagger$  can then produce any eigenstate, starting from the 0 eigenstate:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle.$$

The matrix representation of these operator in the  $|n\rangle$  basis (with infinite-dimensional matrices) is particularly simple, since  $\langle n|a|n'\rangle = \delta_{n',n-1}\sqrt{n}$  and  $\langle n|a^\dagger|n'\rangle = \delta_{n',n+1}\sqrt{n+1}$ :

$$a = \begin{bmatrix} 0 & \sqrt{1} & 0 & \dots \\ 0 & 0 & \sqrt{2} & \dots \\ 0 & 0 & 0 & \dots \end{bmatrix} \quad a^\dagger = \begin{bmatrix} 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & \dots \end{bmatrix}$$

The Hamiltonian can be written in terms of these operators. We substitute  $a$ ,  $a^\dagger$  at the place of  $X$  and  $P$ , yielding  $\mathcal{H} = \omega(a^\dagger a + \frac{1}{2}) = \omega(N + \frac{1}{2})$  and the minimum energy  $\hbar\omega/2$  is called the zero point energy.

### 9.1.3 Position representation

We have now started from a (physical) description of the h.o. Hamiltonian and made a change of basis in order to arrive at a simple diagonal form of it. Now that we know its eigenkets, we would like to go back to a more intuitive picture of position and momentum. We thus want to express the eigenkets  $|n\rangle$  in terms of the position representation (see also section 5.5.1).

The position representation corresponds to expressing a state vector  $|\psi\rangle$  in the position basis:  $|\psi\rangle = \int dx \langle x|\psi\rangle |x\rangle = \int dx \psi(x) |x\rangle$  (where  $|x\rangle$  is the eigenstate of the position operator that is a continuous variable, hence the integral). This defines the wavefunction  $\psi(x) = \langle x|\psi\rangle$ .

The wave function description in the  $x$  representation of the quantum h.o. can be found by starting with the ground state wavefunction. Since  $a|0\rangle = 0$  we have  $\frac{1}{\sqrt{2}}(X + iP)|0\rangle = \frac{1}{\sqrt{2}}(\sqrt{m\omega}x + \frac{ip}{\sqrt{m\omega}})|0\rangle = 0$ . In the  $x$  representation, given  $\psi_0(x) = \langle x|0\rangle$

$$\frac{1}{\sqrt{2}} \langle x|(\sqrt{m\omega}x + \frac{ip}{\sqrt{m\omega}})|0\rangle = 0 \quad \rightarrow \quad (m\omega x + \frac{d}{dx})\psi_0(x) = 0 \quad \rightarrow \quad \psi_0(x) \propto e^{-m\omega x^2/2}$$

The other eigenstates are built using Hermite Polynomials  $H_n(x)$ , using the formula<sup>31</sup>  $|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle$  to derive differential equations:

$$\psi_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{n!2^n}} \left[ \sqrt{m\omega}x - \frac{1}{\sqrt{m\omega}} \frac{d}{dx} \right]^n \psi_0(x)$$

with solutions  $\psi_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{2^n n!}} H_n(x) \psi_0(x)$ .

The  $n = 2$  and  $n = 3$  wavefunctions are plotted in the following figure, while the second figure displays the probability distribution function. Notice the different parity for even and odd number and the number of zeros of these functions.

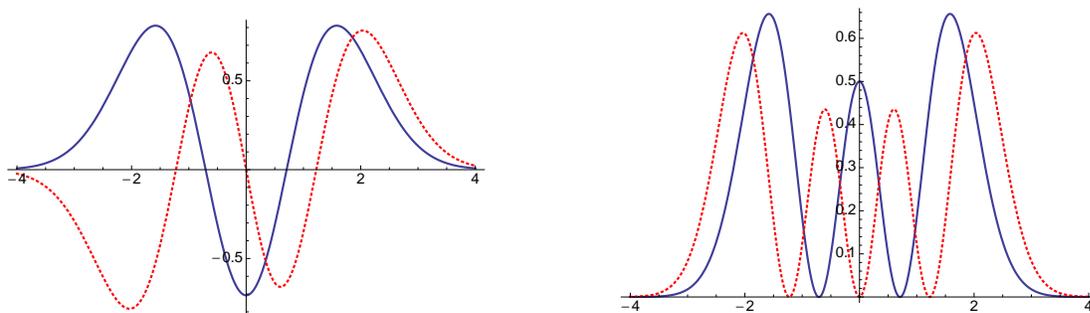


Fig. 12: Left: Harmonic oscillator wavefunction. Right: corresponding probability distribution function for  $n = 2$  (blue) and  $n = 3$  (Red, dotted).

Classically, the probability that the oscillating particle is at a given value of  $x$  is simply the fraction of time that it spends there, which is inversely proportional to its velocity  $v(x) = x_0\omega \sqrt{1 - \frac{x^2}{x_0^2}}$  at that position. For large  $n$ , the probability distribution becomes close to the classical one:

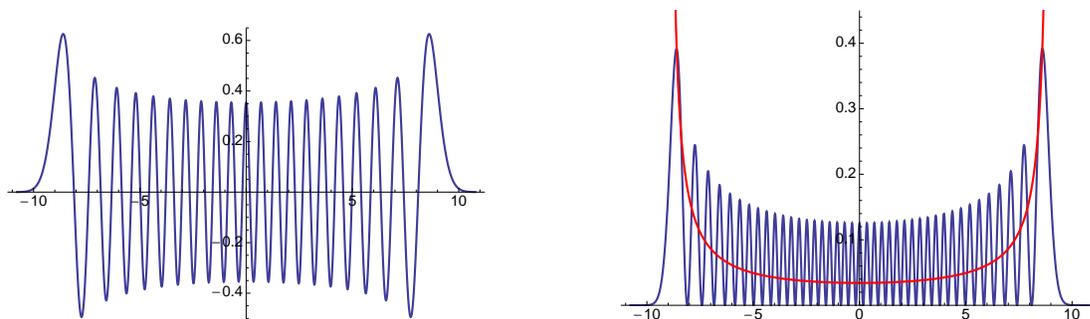


Fig. 13: Left: Harmonic oscillator wavefunction. Right: corresponding probability distribution function for  $n = 40$ . In Red, the classical probability.

<sup>31</sup> For more details on Hermite Polynomials and their generator function, look on Cohen-Tannoudji. Online information from: Eric W. Weisstein. [Hermite Polynomial](#). From MathWorld—A Wolfram Web Resource.

### 9.1.4 Heisenberg picture

We want now to study the time-evolution of the h.o. We first start with analyzing the evolution of the operators in the Heisenberg picture. We have

$$\frac{da}{dt} = i[\mathcal{H}, a] = i[\omega(a^\dagger a + \frac{1}{2}), a] = -i\omega a \quad \rightarrow \quad a(t) = a(0)e^{-i\omega t}$$

Similarly:

$$\frac{da^\dagger}{dt} = i[\mathcal{H}, a^\dagger] = i[\omega(a^\dagger a + \frac{1}{2}), a^\dagger] = i\omega a^\dagger \quad \rightarrow \quad a^\dagger(t) = a^\dagger(0)e^{i\omega t}$$

Notice that we could have found this last relationship just by taking the hermitian conjugate of the first one. Using these results, we can also find the time evolution of the position and momentum operators:

$$x(t) = x(0) \cos(\omega t) + \frac{p(0)}{m\omega} \sin(\omega t)$$

$$p(t) = p(0) \cos(\omega t) - m\omega x(0) \sin(\omega t)$$

and the corresponding expectation values, e.g.

$$\langle x(t) \rangle = \langle x(0) \rangle \cos(\omega t) + \frac{\langle p(0) \rangle}{m\omega} \sin(\omega t)$$

### 9.1.5 Schrödinger picture

An initial state can be expressed in terms of the number eigenvectors:  $|\psi\rangle = \sum_n c_n |n\rangle$ . Then its evolution is simply:  $|\psi(t)\rangle = \sum_n c_n e^{-in\omega t} |n\rangle$ . From this expression, one can calculate e.g.

$$\langle x(t) \rangle = \sum_{n,m} c_n c_m^* \langle m|x|n\rangle e^{-i\omega t(n-m)}.$$

Since  $x$  only connects states that differ by  $n - m = \pm 1$ , it's easy to see that the double sum simplifies and we retrieve the expression above, found in the Heisenberg picture.

## 9.2 Uncertainty relationships

The operators  $x$  and  $p$  for a quantum h.o. do not commute, so they do not share any eigenstate, nor they share eigenstates with the Hamiltonian. In particular the diagonal elements of  $x$  and  $p$  in the  $|n\rangle$ -basis representation are both zero, therefore the expectation values are also zero. In a series of measurements, it is possible to get a range of values; we associate this dispersion of values with the root mean square value of the eigenvalues:

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \tag{3}$$

$$\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} \tag{4}$$

Given the expression for  $x$  and  $p$  in terms of  $a$  and  $a^\dagger$  we can calculate  $\langle x^2 \rangle$ :

$$\langle x^2 \rangle = \frac{\hbar}{2m\omega} \langle n| aa + a^\dagger a^\dagger + a^\dagger a + aa^\dagger |n\rangle \tag{5}$$

$$= \frac{\hbar}{2m\omega} \langle n| a^\dagger a + aa^\dagger |n\rangle = \frac{\hbar}{2m\omega} (2n + 1) \tag{6}$$

and in the same way, we can calculate  $\langle p^2 \rangle$ :  $\langle p^2 \rangle = \frac{m\hbar\omega}{2} (2n + 1)$ . Since the expectation values are zero ( $\langle x \rangle = \langle p \rangle = 0$ ), the deviations are just:  $\Delta x = \sqrt{\langle x^2 \rangle}$  and  $\Delta p = \sqrt{\langle p^2 \rangle}$  and the uncertainty relation can be expressed by:

$$\Delta p \Delta x = \frac{\hbar}{2} (2n + 1) \tag{7}$$

We see that in general  $\Delta p \Delta x \geq \frac{\hbar}{2}$ , with equality for  $n = 0$ : the ground state of the harmonic oscillator is a state of minimum uncertainty. More generally, for any potential  $V(x)$ , the ground state of a local minimum is always a state of minimum uncertainty (since the potential can be always approximated by an harmonic potential).

We expect that higher energy states do not saturate the uncertainty bound. Classically, when a system has some finite energy, the particle is moving around so  $\Delta x = \sqrt{2}x_0$ . At the minimum energy (that classically is 0), the particle is at rest, localized ( $\Delta x = 0$ ). For the quantum h.o., even the minimum energy state is not localized, but rather it is a gaussian packet (as described by  $\psi_0(x)$ ) thus the state does have some uncertainty in its position. Still, as expected from the classical intuition, this uncertainty is the minimum possible.

From the expectation values  $\langle x^2 \rangle$  and  $\langle p^2 \rangle$  we can calculate the average kinetic and potential energy. We find that the average potential and kinetic energy are the same,  $\langle T \rangle = \langle V \rangle = \frac{\hbar\omega}{4} = \langle E \rangle / 2$ , as for classical conservative systems (virial theorem).

### 9.3 Coherent States

We now want to look at some connexion of the quantum h.o. with the classical one. We have seen that in the limit of vanishing energy, the classical and quantum oscillators are very different, since the minimum energy for the quantum h.o. is non-zero, while the classical h.o. is totally localized. On the opposite side, we saw that at high energy (high  $n$ ) the energy difference between two levels vanishes,  $\frac{\Delta E}{E} = \frac{\hbar\omega}{(n+1/2)\hbar\omega} \approx 0$ ; thus the energy becomes continuous, as it would be in the classical case. Still, to find a quantum-to-classical correspondence it is not enough to choose a stationary eigenstate of the Hamiltonian with a high energy (high  $n$ ): this state would still have zero expectation value for the momentum and position. In contrast, the position evolution in classical mechanics is  $x_{cl} = x_0 \cos \omega t$ : ideally we would like to find a state  $|\psi_{cl}\rangle$  such that  $\langle x(t) \rangle = \langle \psi_{cl}(t) | x | \psi_{cl}(t) \rangle = x_{cl}$ , as usually stated by Ehrenfest theorem. *Coherent states* achieve this result. For this reasons, these states are also called quasi-classical.

The coherent state was defined by Roy J. Glauber<sup>32</sup>. He was looking for a superposition of eigenstates that looked as classical as possible, without invoking any decoherence or the action of an external environment. The coherent states are pure quantum states, however when we look at expectation values with respect to these states, the limit of high energy we recover the classical results. For example, although the operator  $x$  and  $p$  do not commute and give rise to the known uncertainty relationships, when we consider the high energy limit of their expectation values the uncertainties become a vanishing contribution.

Glauber idea was to introduce a complex classical variable  $\alpha = \frac{1}{\sqrt{2}}(X + iP)$  (where  $X$  and  $P$  are the dimensionless variables defined previously). The classical equations of motion for  $x$  and  $p$  define the evolution of the variable  $\alpha$ :

$$\frac{dx}{dt} = \frac{p(t)}{m}, \quad \frac{dp}{dt} = -m\omega^2 x \quad \rightarrow \quad \frac{d\alpha}{dt} = -i\omega\alpha(t)$$

The evolution of  $\alpha$  is therefore just a rotation in its phase space:  $\alpha(t) = \alpha(0)e^{-i\omega t}$ . This is usual for a conservative system (in classical mechanics) or closed systems in QM. The initial conditions thus specify the overall evolution,  $\alpha_0 = \alpha(0)$  contains all the important information.

Since  $X = \sqrt{2}Re(\alpha)$  and  $P = \sqrt{2}Im(\alpha)$ , the expectation values for  $X$  and  $P$  oscillate, as usual in the classical case (again, here  $X$  and  $P$  are just normalized, classical variable).

$$\begin{aligned} \langle X \rangle &= \frac{1}{\sqrt{2}}(\alpha_0 e^{-i\omega t} + \alpha_0^* e^{i\omega t}) \\ \langle P \rangle &= \frac{-i}{\sqrt{2}}(\alpha_0 e^{-i\omega t} - \alpha_0^* e^{i\omega t}) \end{aligned}$$

The classical energy, given by  $\omega/2(X^2 + P^2) = \omega\alpha_0^2$ , is constant at all time.

Now consider the QM problem, where the variables are replaced by the corresponding operators:

$$X = (a + a^\dagger)/\sqrt{2}, \quad P = -i(a - a^\dagger)/\sqrt{2}, \quad \mathcal{H} = \omega(a^\dagger a + \frac{1}{2}),$$

and consider the evolution of the operator  $a$  in the Heisenberg picture. Its expectation value is given by

$$\frac{d\langle a \rangle}{dt} = -i\langle [a, \mathcal{H}] \rangle = -i\langle [a, \omega a^\dagger a] \rangle = -i\omega\langle a \rangle$$

<sup>32</sup> Roy J. Glauber, *Coherent and Incoherent States of the Radiation Field*, *Phys. Rev.* **131** 2766–2788 (1963).

Therefore the expectation value evolution is the same as for the  $\alpha$  variable:

$$\langle a(t) \rangle = \langle a(0) \rangle = e^{-i\omega t}, \quad \langle a(t)^\dagger \rangle = \langle a(0)^\dagger \rangle e^{i\omega t}$$

Inspired by this result, we consider a state that is an eigenstate of the annihilation operator  $a$ :  $a|\alpha\rangle = \alpha|\alpha\rangle$ . With respect to this state we have  $\langle X \rangle = \langle \alpha | (a + ia^\dagger) | \alpha \rangle / \sqrt{2} = (\alpha + \alpha^*) / \sqrt{2} = \text{Re}(\alpha) / \sqrt{2} \neq 0$ . The evolution of  $\langle X \rangle$  will then have the same oscillatory character as for its classical counterpart. This eigenstate of the annihilation operator has the desired property and we thus identify it with a coherent state.

The expectation values of position and momentum with respect to a coherent state give rise to the classical result. However, when considering the expectation value of the energy, there are still two contributions: the first one contributes to the classical energy ( $\omega a^\dagger a \rightarrow E = \omega |\alpha_0|^2$ ), while the second term is a purely QM contribution (zero point energy). The classical limit is reached at higher energy where the first contribution is much larger than the zero-point energy  $\hbar\omega/2$ .

### 9.3.1 Expansion in terms of number states

The coherent state can of course be expressed in terms of number eigenstates:  $|\alpha\rangle = \sum_n c_n |n\rangle$ . We want to derive the coefficients  $c_n$ . From

$$a|\alpha\rangle = \alpha|\alpha\rangle \rightarrow \sum_{n=0}^{\infty} c_n a|n\rangle = \sum_{n=1}^{\infty} c_n \sqrt{n} |n-1\rangle = \sum_{n=0}^{\infty} c_{n+1} \sqrt{n+1} |n\rangle$$

we obtain

$$\sum_{n=0}^{\infty} (\alpha c_n - c_{n+1} \sqrt{n+1}) |n\rangle = 0 \rightarrow \alpha c_n = c_{n+1} \sqrt{n+1}$$

We thus have a series of equations,

$$\begin{cases} c_1 = \alpha c_0 \\ c_2 = \frac{\alpha}{\sqrt{2}} c_1 = \frac{\alpha^2}{\sqrt{2}} c_0 \\ c_3 = \frac{\alpha}{\sqrt{3}} c_2 = \frac{\alpha^3}{\sqrt{6}} c_0 \end{cases}$$

So in general  $c_n = \frac{\alpha^n}{\sqrt{n!}} c_0$ . We finally obtain  $c_0$  from the normalization condition  $\langle \alpha | \alpha \rangle = 1$ :

$$|c_0|^2 = \left( \sum_{m,n} \frac{\alpha^n (\alpha^*)^m}{\sqrt{n!m!}} \langle m | n \rangle \right)^{-1} = \left( \sum_n \frac{|\alpha|^{2n}}{n!} \right)^{-1} = e^{-|\alpha|^2}$$

The coherent state can thus be expressed in terms of the number states as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

This also gives the probability for obtaining a particular energy level  $n$  when the system is in a quantum coherent state:

$$P_\alpha(n) = |\langle n | \alpha \rangle|^2 = e^{-\langle n \rangle} \frac{\langle n \rangle^n}{n!}$$

where we have used that the average number of photons is  $\langle n \rangle = \langle \alpha | a^\dagger a | \alpha \rangle = |\alpha|^2$ . Notice also that  $\Delta n^2 = |\alpha|^2$ . We thus see that the coherent states have a Poissonian distribution.

### 9.3.2 Non-Orthogonality

The coherent states  $|\alpha\rangle$  do not form a proper basis, since they are eigenvectors of a non-hermitian operator. In particular they are not orthogonal (even if they are normalized by the choice of  $c_0$ ):

$$\langle\alpha|\beta\rangle = \sum_{n,m} e^{-(|\alpha|^2+|\beta|^2)/2} (\alpha^*)^n \beta^m / \sqrt{n!m!} \langle n|m\rangle = e^{-(|\alpha|^2+|\beta|^2)/2} \sum_n (\alpha^*\beta)^n / n! = e^{-(|\alpha|^2+|\beta|^2-2\alpha^*\beta)/2}$$

Although not orthogonal, their superposition goes to zero as  $|\alpha - \beta| \rightarrow 0$ , since

$$|\langle\alpha|\beta\rangle|^2 = e^{-(|\alpha|^2+|\beta|^2-2\alpha^*\beta)/2} e^{-(|\alpha|^2+|\beta|^2-2\alpha\beta^*)/2} = e^{-|\alpha-\beta|^2}$$

Also, the set of coherent states is complete:

$$\int |\alpha\rangle \langle\alpha| d\alpha/\pi = \mathbb{1}$$

Because of this closure relation, any state can be written in terms of coherent state superposition, thus the coherent states form an overcomplete basis.

### 9.3.3 Uncertainty relationships

We have already seen that

$$\langle X \rangle = \sqrt{2} \operatorname{Re}[\alpha] = \frac{1}{\sqrt{2}}(\alpha + \alpha^*), \quad \langle P \rangle = -i\sqrt{2} \operatorname{Im}[\alpha] = \frac{i}{\sqrt{2}}(\alpha^* - \alpha)$$

Now consider the variance. We have:

$$\langle X^2 \rangle = \frac{1}{2} \langle \alpha | a^2 + (a^\dagger)^2 + aa^\dagger + a^\dagger a | \alpha \rangle = \frac{1}{2}(\alpha^2 + (\alpha^*)^2 + 2\alpha^*\alpha + 1)$$

and

$$\langle P^2 \rangle = \frac{1}{2} \langle \alpha | a^2 + (a^\dagger)^2 - aa^\dagger - a^\dagger a | \alpha \rangle = -\frac{1}{2}(\alpha^2 + (\alpha^*)^2 - 2\alpha^*\alpha - 1) = \frac{1}{2}[1 - (\alpha - \alpha^*)^2]$$

and for example:

$$\Delta X^2 = \frac{1}{2}[(\alpha^2 + (\alpha^*)^2 + 2\alpha^*\alpha + 1) - (\alpha + \alpha^*)^2] = \frac{1}{2}$$

We then have  $\Delta X^2 = \frac{1}{2}$  and  $\Delta P^2 = \frac{1}{2}$  so that the uncertainty relationship is saturated:

$$\Delta X \Delta P = \frac{1}{2}$$

The coherent state is thus a minimum uncertainty state (as the number states were).

**? Question:** What are the uncertainty relationship in terms of the variables  $x$  and  $p$ ?

$$\langle x \rangle = \sqrt{\frac{\hbar}{2m\omega}}(\alpha + \alpha^*), \quad \langle p \rangle = i\sqrt{\frac{\hbar m\omega}{2}}(\alpha - \alpha^*)$$

and

$$\begin{aligned} \langle x^2 \rangle &= \frac{\hbar}{2m\omega}(\alpha^2 + (\alpha^*)^2 + 2\alpha^*\alpha + 1) \\ \langle p^2 \rangle &= \frac{\hbar m\omega}{2}(\alpha^2 + (\alpha^*)^2 - 2\alpha^*\alpha - 1) \end{aligned}$$

We thus have the uncertainties for  $x$  and  $p$  and their uncertainty relationship

$$\Delta x^2 = \frac{\hbar}{2m\omega}, \quad \Delta p^2 = \frac{\hbar m\omega}{2} \rightarrow \Delta x \Delta p = \frac{\hbar}{2}$$

□

### 9.3.4 X-representation

We now want to obtain an expression for the wavefunction representing a coherent state, that is, we want to find the x-representation of the coherent state:  $\langle x | \alpha \rangle$ . For this, we start from the equation

$$\langle x | a | \alpha \rangle = \alpha \langle x | \alpha \rangle$$

as well as the explicit form of  $a$  in terms of  $x$  and  $p$ ,  $a = \frac{\sqrt{m\omega}}{\sqrt{2\hbar}}x + \frac{i}{\sqrt{2m\omega\hbar}}p$

$$\langle x | a | \alpha \rangle = \langle x | \left( \sqrt{\frac{m\omega}{2\hbar}}x + i\sqrt{\frac{1}{2\hbar m\omega}}p \right) | \alpha \rangle$$

Now we define the wavefunction in the x-representation  $\langle x | \alpha \rangle = \psi_\alpha(x)$  and we remember that

$$\langle x | p | \psi \rangle = -i\hbar\partial_x\psi(x)$$

and  $\langle x | p | x' \rangle = -i\hbar\partial_x\delta(x - x')$  to obtain:

$$\langle x | \left( \sqrt{\frac{m\omega}{2\hbar}}x + i\sqrt{\frac{1}{2\hbar m\omega}}p \right) | \alpha \rangle = \left( \sqrt{\frac{m\omega}{2\hbar}}x + \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial x} \right) \langle x | \alpha \rangle$$

Equating with the expression obtained before yields the differential equation:

$$\frac{\partial}{\partial x}\psi_\alpha(x) = \left( \sqrt{\frac{2m\omega}{\hbar}}\alpha - \frac{m\omega}{\hbar}x \right) \psi_\alpha(x)$$

with solution

$$\psi_\alpha(x) = Ae^{\sqrt{\frac{2m\omega}{\hbar}}\alpha x} e^{-\frac{m\omega}{2\hbar}x^2}$$

The constant A can be as usual obtained from the normalization condition:

$$\int_{-\infty}^{\infty} |\psi_\alpha(x)|^2 dx = 1 \rightarrow A = \left( \frac{m\omega}{2\pi\hbar} \right)^{1/4} e^{-\frac{\alpha^2}{2}}$$

The wavefunction representation is thus a Gaussian wavepacket:

$$\psi_\alpha(x) = \left( \frac{m\omega}{2\pi\hbar} \right)^{1/4} e^{-\frac{\alpha^2}{2}} e^{\sqrt{\frac{2m\omega}{\hbar}}\alpha x} e^{-\frac{m\omega}{2\hbar}x^2}$$

(not just a simple Gaussian, since  $\alpha$  can be complex).

## 9.4 Phonons

We have introduced the harmonic oscillator as an interesting model because of the energy level structure it gives rise to. A second reason for its utility is that it can model many different systems around their equilibrium point. Here we show how it can be used to describe vibrations in a crystal lattice and how the quantum-mechanical description can be used to derive some of the lattice properties, such as its specific heat.

### 9.4.1 Harmonic oscillator model for a crystal

We consider a crystal formed by ions of mass  $M$  in a lattice (for simplicity we will consider a monoatomic, one-dimensional lattice).

The ion equilibrium positions are  $R_n = nd$ , with  $d$  the lattice constant, but the actual position of the ions is  $r_n = R_n + x_n$ , where  $x_n$  is the displacement from the equilibrium.

The interaction potential among the ions is

$$U = \frac{1}{2} \sum_{n,m} \mathcal{U}(r_n - r_m) = \frac{1}{2} \sum_{n,m} \mathcal{U}(R_n - R_m + x_n - x_m)$$

Assuming the displacement  $x_n$  is small, we can expand the potential as:

$$U = \frac{1}{2} \sum_{n,m} \mathcal{U}(R_n - R_m) + \frac{1}{2} \sum_{n,m} (x_n - x_m) \partial \mathcal{U}(R_n - R_m) + \frac{1}{4} \sum_{n,m} (x_n - x_m)^2 \partial^2 \mathcal{U}(R_n - R_m)$$

The first term  $U_{eq} = \frac{1}{2} \sum_{n,m} \mathcal{U}(R_n - R_m)$  is the interaction potential at equilibrium, which is not of interest here. Consider the linear term:

$$\frac{1}{2} \sum_{n,m} (x_n - x_m) \partial \mathcal{U}(R_n - R_m) = \frac{1}{2} \sum_n x_n \sum_m [\partial \mathcal{U}(R_n - R_m) - \partial \mathcal{U}(R_m - R_n)] = \sum_n x_n \sum_m \partial \mathcal{U}(R_n - R_m)$$

The term  $\sum_m \partial \mathcal{U}(R_n - R_m) = F_n$  is the total force exerted on the atom  $n$  by all the other atoms. When all the atoms are at equilibrium, this force must be zero, since there can be no net force at equilibrium. We are then left with only the second order term:

$$U = \frac{1}{2} \sum_{n,m} (x_n - x_m)^2 \partial_x^2 \mathcal{U}$$

If we assume that only neighboring ions interact, to second order, we retrieve an harmonic potential:

$$U = \frac{1}{2} K \sum_n (x_n - x_{n+1})^2$$

Then, the Hamiltonian can be written as:

$$\mathcal{H} = \sum_n \frac{p_n^2}{2M} + \frac{1}{2} K \sum_n (x_n - x_{n+1})^2$$

while the equation of motion for each oscillator is:

$$M \ddot{x}_n = - \frac{\partial U}{\partial x_n} = -K [2x_n - x_{n-1} - x_{n+1}]$$

### 9.4.2 Phonons as normal modes of the lattice vibration

The classical oscillator model is solved by guessing a solution in terms of waves (normal modes of the oscillation):

$$x_k(t) \propto e^{-ikR_n} e^{-i\omega_k t}$$

This solution is motivated by the translation symmetry of the lattice. We can check that the solution we guessed is the correct one, by inserting it in the equation of motion

$$M\ddot{x}_n = -\frac{\partial U}{\partial x_n} \rightarrow -M\omega_k^2 x_n = -K[2x_n - x_{n-1} - x_{n+1}] = -Kx_n[2 - e^{ikd} - e^{-ikd}]$$

Thus if we set  $\omega_k^2 = \frac{K}{M}2[1 - \cos(kd)]$  the equation is verified. The relationship:

$$\omega(k) = 2\omega_0 \left| \sin\left(\frac{kd}{2}\right) \right|,$$

with  $\omega_0 = \sqrt{K/M}$ , is called the **dispersion** relation, which describes the frequency (energy) of the wave as a function of the wavelength.

This solution describes waves propagating in the chain with *phase* velocity  $c = \omega/k$  and *group* velocity  $v_g = \frac{\partial \omega}{\partial k}$  (the speed of sound in the given material). At small  $k$  the two velocities are equal, but for large  $k$  (small spacing between ions) we have  $v_g \rightarrow 0$ .

We can now turn to the corresponding quantum-mechanical model, by replacing the position and momentum coordinates in the Hamiltonian by the corresponding operators:

$$\mathcal{H} = \sum_n \frac{p_n^2}{2M} + \frac{1}{2}K \sum_n (x_n - x_{n+1})^2$$

Inspired by the classical solution, we also look for solutions (i.e. eigenvectors that diagonalize the Hamiltonian) in terms of waves. Then in this basis,  $x_n$  and  $p_n$  will be expressed as linear combinations of waves with different wavevectors, e.g.:

$$x_n = \frac{1}{\sqrt{N}} \sum_k X_k e^{ikR_n}$$

We then rewrite the operator  $X_k$  (and  $P_k$ ) in terms of the creation and annihilation operators  $a_k, a_k^\dagger$ :

$$x_n = \frac{1}{\sqrt{NM}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left( a_k e^{ikR_n} + a_k^\dagger e^{-ikR_n} \right)$$

and

$$p_n = -i\sqrt{\frac{M}{N}} \sum_k \sqrt{\frac{\omega_k}{2}} \left( a_k e^{-ikR_n} - a_k^\dagger e^{ikR_n} \right)$$

Similar to the solution for the simple h.o. we want to verify that the operators  $a_k, a_k^\dagger$  diagonalize the Hamiltonian. We thus first calculate the kinetic energy,  $T = \sum_n \frac{p_n^2}{2M}$ .

$$p_n^2 = -\frac{M}{N} \sum_{k,h} \sqrt{\frac{\omega_k \omega_h}{4}} \left[ a_k^\dagger a_h^\dagger e^{-i(k+h)R_n} - a_k^\dagger a_h e^{-i(k-h)R_n} + a_k a_h e^{i(k+h)R_n} - a_k a_h^\dagger e^{i(k-h)R_n} \right]$$

We then take the sum over  $n$ , remember that  $R_n = nd$  where  $n$  is an integer and  $d$  the distance between two ions, and invert the order of the sums:

$$T = -\frac{1}{4N} \sum_{k,h} \sqrt{\omega_k \omega_h} \left[ a_k^\dagger a_h^\dagger \sum_n e^{-i(k+h)R_n} - a_k^\dagger a_h \sum_n e^{-i(k-h)R_n} + \dots \right]$$

The sums  $\sum_n e^{-i(k+h)R_n} = \sum_n e^{-i(k+h)nd}$  are zero unless the exponent argument  $h + k$  is zero:

$$T = -\frac{1}{4N} \sum_{k,h} \sqrt{\omega_k \omega_h} \left[ a_k^\dagger a_h^\dagger \delta_{k,-h} - a_k^\dagger a_h \delta_{k,h} + \dots \right]$$

We thus obtain:

$$T = -\frac{1}{4} \sum_k \omega_k \left[ a_k^\dagger a_{-k}^\dagger - a_k^\dagger a_k + a_k a_{-k} - a_k a_k^\dagger \right]$$

We then calculate the potential energy. First we calculate  $x_n - x_{n+1}$ :

$$\begin{aligned} x_n - x_{n+1} &= \frac{1}{\sqrt{NM}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left( a_k e^{ikR_n} + a_k^\dagger e^{-ikR_n} - a_k e^{ikR_{n+1}} - a_k^\dagger e^{-ikR_{n+1}} \right) \\ &= \frac{1}{\sqrt{NM}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left( a_k e^{ikR_n} (1 - e^{ikd}) + a_k^\dagger e^{-ikR_n} (1 - e^{-ikd}) \right) \\ &= \frac{1}{\sqrt{NM}} \sum_k \frac{-2i}{\sqrt{2\omega_k}} \sin\left(\frac{kd}{2}\right) \left( a_k e^{ik(R_n+d/2)} + a_k^\dagger e^{-ik(R_n+d/2)} \right) \end{aligned}$$

The potential energy is then:

$$U = \frac{1}{2} K \sum_n (x_n - x_{n+1})^2 = -\frac{K}{2NM} \sum_k \frac{4 \sin^2\left(\frac{kd}{2}\right)}{2\omega_k} \left[ -a_k^\dagger a_{-k}^\dagger - a_k^\dagger a_k - a_k a_{-k} - a_k a_k^\dagger \right]$$

(where we used the same identities for the sum of exponential and the fact that  $\sin\left(\frac{kd}{2}\right) \sin\left(\frac{hd}{2}\right) \delta_{k,-h} = \sin^2\left(\frac{kd}{2}\right)$ ). By summing the potential and kinetic energy and imposing as before:

$$\omega_k = 2\omega_0 \left| \sin\left(\frac{kd}{2}\right) \right|$$

(with  $\omega_0^2 = \frac{K}{M}$ ) we can simplify the Hamiltonian to:

$$\mathcal{H} = \frac{1}{2} \sum_k \omega_k \left( a_k^\dagger a_k + a_k a_k^\dagger \right) = \sum_k \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right)$$

The operators  $a_k, a_k^\dagger$  do diagonalize the Hamiltonian. The number operator  $n_k = a_k^\dagger a_k$  describes the excitation number of a normal mode of the ion vibration. Instead of talking of excitations, we can introduce quasi particles, called **phonons**. The number of phonons then corresponds to the number of excitations. Thus the operators  $a_k, a_k^\dagger$  can create or annihilate a phonon of mode  $k$ .

### 9.4.3 Thermal energy density and Specific Heat

We want to first calculate the thermal energy density  $u = E/V$  and then the specific heat,  $c_V = \frac{\partial u}{\partial T}$  for a crystal at thermal equilibrium.

The thermal energy is given by the lattice vibration. Thus we want to calculate:

$$\langle E \rangle = \text{Tr} \{ \rho \mathcal{H} \}$$

with the Hamiltonian given above. The system in thermal equilibrium is described by the usual distribution:

$$\rho = \frac{e^{-\beta \mathcal{H}}}{Z}$$

Notice that  $\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta}$ . We thus need to calculate the partition function. Computing the trace in the number state basis, we have:

$$\begin{aligned} Z &= \text{Tr} \left\{ \exp \left[ -\beta \sum_k \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right) \right] \right\} = \text{Tr} \left\{ \prod_k \exp \left[ -\beta \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right) \right] \right\} \\ &= \prod_k e^{-\beta \omega_k / 2} \sum_n (e^{-\beta \omega_k})^n = \prod_k \frac{e^{-\beta \omega_k / 2}}{1 - e^{-\beta \omega_k}} \end{aligned}$$

Taking the derivative of the logarithm, we find:

$$u = -\frac{1}{V} \partial_\beta (\ln Z) = \frac{1}{V} \sum_k \frac{\omega_k}{2} \coth \left( \frac{\omega_k \beta}{2} \right)$$

This can also be written in terms of the average phonon number for the mode  $k$ ,

$$\begin{aligned} \langle n_k \rangle &= n(k) = \text{Tr} \left\{ a_k^\dagger a_k \rho \right\} = [e^{\beta \omega_k} - 1]^{-1} : \\ u &= \frac{1}{V} \sum_k \omega_k \left[ n(k) + \frac{1}{2} \right] \end{aligned}$$

The specific heat is then:

$$c_V = \frac{\partial u}{\partial T} = \sum_k \frac{\omega_k^2}{4V k_b T^2 \sinh^2 \left( \frac{\omega_k}{2k_b T} \right)}$$

Note that at high temperature (small  $\beta$ ) this is approximated by  $c_V \approx \sum_k \frac{k_b}{V} = N \frac{k_b}{V}$ , which is the classical Dulong-Petit law, stating that the specific heat is independent of the temperature and given by the density  $n = N/V$  and the system's dimension  $D$ ,  $c_V = D n k_b$ .



# 10. The electromagnetic field

## 10.1 Classical theory of the e.m. field

## 10.2 Quantization of the e.m. field

### 10.2.1 Zero-Point Energy and the Casimir Force

## 10.3 Quantization of the e.m. field in the Coulomb gauge

## 10.4 States of the e.m. field

### 10.4.1 Photon number eigenstates

### 10.4.2 Coherent states

### 10.4.3 Measurement Statistics

## 10.5 Atomic interactions with the quantized field

We will now provide a quantum-mechanical description of the electro-magnetic field. Our main interest will be in analyzing phenomena linked to atomic physics and quantum optics, in which atoms interact with radiation. Some processes can be analyzed with a classical description: for example we studied the precession and the manipulation of a spin by classical static and rf magnetic fields. Absorption and emission of light by an atom can also be described as the interaction with a classical field. Some other phenomena, such as spontaneous emission, can only arise from a QM description of both the atom and the field. There are various examples in which the importance of a quantum treatment of electromagnetism becomes evident:

- Casimir force
- Spontaneous emission, Lamb Shift
- Laser linewidth, photon statistics
- Squeezed photon states, states with subpoissonian distribution,
- Quantum beats, two photon interference, etc.

## 10.1 Classical theory of the e.m. field

Before introducing the quantization of the field, we want to review some basic (and relevant) concepts about e.m. fields.

Maxwell equations for the electric and magnetic fields,  $\mathbf{E}$  and  $\mathbf{B}$ , are:

Gauss's law	$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$
Gauss's law for magnetism	$\nabla \cdot \mathbf{B} = 0$
Maxwell-Faraday equation (Faraday's law of induction)	$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}$
Ampere's circuital law (with Maxwell's correction)	$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$

We will be interested to their solution in empty space (and setting  $c = 1/\sqrt{\mu_0 \epsilon_0}$ ):

Gauss's law	$\nabla \cdot \mathbf{E} = 0$
Gauss's law for magnetism	$\nabla \cdot \mathbf{B} = 0$
Maxwell-Faraday equation	$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}$
Ampere's circuital law	$\nabla \times \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$

Combining Maxwell equation in vacuum, we find the wave equations:

$$\nabla^2 E - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = 0 \quad \nabla^2 B - \frac{1}{c^2} \frac{\partial^2 B}{\partial t^2} = 0$$

**? Question:** Show how this equation is derived

We need to take the curl of Maxwell-Faraday equation and the time derivative of Ampere's law and use the vector identity  $\nabla \times (\nabla \times \vec{v}) = \nabla(\nabla \cdot \vec{v}) - \nabla^2 \vec{v}$  and Gauss law.  $\square$

A general solution for these equations can be written simply as  $E = E(\omega t - \vec{k} \cdot \vec{x})$ . By fixing the boundary conditions, we can find a solution in terms of an expansion in normal modes, where the time dependence and spatial dependence are separated. The solution of the wave equation can thus be facilitated by representing the electric field as a sum of normal mode functions:

$$\mathbf{E}(\vec{x}, t) = \sum_m f_m(t) \vec{u}_m(\vec{x}).$$

The normal modes  $u_m$  are the equivalent of eigenfunctions for the wave equation, so they do not evolve in time (i.e. they are function of position only). The  $u_m$  are orthonormal functions, called normal modes. The boundary conditions define the normal modes  $u_m$  for the field, satisfying:

$$\nabla^2 u_m = -k_m^2 u_m, \quad \nabla \cdot u_m = 0, \quad \vec{n} \times u_m = 0$$

(where  $n$  is a unit vector normal to a surface). This last condition is imposed because the tangential component of the electric field  $\mathbf{E}$  must vanish on a conducting surface. We can also choose the modes to satisfy the orthonormality condition (hence *normal* modes):

$$\int \vec{u}_m(x) \vec{u}_n(x) d^3x = \delta_{n,m}$$

Substituting the expression for the electric field in the wave equation, we find an equation for the coefficient  $f_m(t)$ :

$$\sum_m \frac{d^2 f_m}{dt^2} + c^2 k_m^2 f_m(t) = 0.$$

Since the mode functions are linearly independent, the coefficients of each mode must separately add up to zero in order to satisfy the wave equation, and we find :

$$\frac{d^2 f_m}{dt^2} + c^2 k_m^2 f_m(t) = 0.$$

As it can be seen from this equation, the dynamics of the normal modes, as described by their time-dependent coefficients, is the same as that of the h.o. with frequency  $\omega_m = ck_m$ . Hence the electric field is equivalent to an infinite number of (independent) harmonic oscillators. In order to find a quantum-mechanical description of the e.m. field we will need to turn this h.o. into quantum harmonic oscillators.

We want to express the magnetic field in terms of the same normal modes  $\vec{u}_m$  which are our basis. We assume for  $B$  the expansion:

$$\mathbf{B}(x, t) = \sum_n h_n(t) [\nabla \times u_n(x)],$$

From Maxwell-Faraday law:

$$\nabla \times E = \sum_n f_n(t) \nabla \times u_n = -\frac{1}{c} \partial_t B$$

we see that we need to impose  $h_n$  such that  $\frac{dh_n}{dt} = -cf_n$  so that we obtain

$$\sum_n -\frac{1}{c} \frac{dh_n}{dt} \nabla \times u_n = -\frac{1}{c} \partial_t B$$

which indeed corresponds to the desired expression for the magnetic field. We now want to find as well an equation for the coefficient  $h_n$  alone. From the expressions of the  $E$  and  $B$ -field in terms of normal modes, using Ampere's law,

$$\begin{aligned}\nabla \times B &= \frac{1}{c} \frac{\partial E}{\partial t} \quad \rightarrow \quad \sum_n h_n(t) \nabla \times (\nabla \times u_n) = \frac{1}{c} \sum_n \frac{df_n}{dt} u_n \\ &\quad \rightarrow \quad - \sum_n h_n \nabla^2 u_n = \frac{1}{c} \sum_n \frac{df_n}{dt} u_n\end{aligned}$$

(where we used the fact that  $\nabla \cdot u = 0$ ) we find

$$\frac{df_n(t)}{dt} = ck_n^2 h_n(t).$$

since we have  $\nabla^2 u_n = -k_n^2 u_n$ . Finally we have:

$$\frac{d^2}{dt^2} h_n(t) + c^2 k_n^2 h_n(t) = 0$$

The Hamiltonian of the system represent the total energy<sup>33</sup>:  $\mathcal{H} = \frac{1}{2} \frac{1}{4\pi} \int (E^2 + B^2) d^3x$ .

We can show that  $\mathcal{H} = \sum_n \frac{1}{8\pi} (f_n^2 + k_n^2 h_n^2)$ :

$$\begin{aligned}\mathcal{H} &= \frac{1}{8\pi} \int (E^2 + B^2) d^3x = \frac{1}{8\pi} \sum_{n,m} \left( f_n f_m \int u_n(x) u_m(x) d^3x + h_n h_m \int (\nabla \times u_n) \cdot (\nabla \times u_m) d^3x \right) \\ &= \sum_n \frac{1}{8\pi} (f_n^2 + k_n^2 h_n^2)\end{aligned}$$

where we used  $\int (\nabla \times u_n) \cdot (\nabla \times u_m) d^3x = k_n^2 \delta_{n,m}$ . We can then use the equation  $\frac{df_n(t)}{dt} = ck_n^2 h_n(t)$  to eliminate  $h_n$ . Then  $f_n$  can be associated with an equivalent position operator and  $h_n$  (being a derivative of the position) with the momentum operator.

Notice that the Hamiltonian for a set of harmonic oscillators, each having unit mass, is

$$\mathcal{H}_{h.o.} = \sum_n \frac{1}{2} (p_n^2 + \omega_n^2 q_n^2)$$

with  $q_n, p_n = \frac{dq_n}{dt}$  the position and momentum of each oscillator.

## 10.2 Quantization of the e.m. field

Given the Hamiltonian we found above, we can associate the energy  $\frac{1}{2}(p_n^2 + \omega_n^2 q_n^2)$  to each mode. We thus make the identification of  $f_n$  with an equivalent position:

$$Q_n = \frac{f_n}{2\omega_n \sqrt{\pi}}$$

and then proceed to quantize this effective position, associating an operator to the position  $Q_n$ :

$$\hat{Q}_n = \sqrt{\frac{\hbar}{2\omega_n}} (a_n^\dagger + a_n)$$

We can also associate an operator to the normal mode coefficients  $f_n$ :

$$\hat{f}_n = \sqrt{2\pi\omega_n \hbar} (a_n^\dagger + a_n)$$

---

<sup>33</sup> The factor  $4\pi$  is present because I am using cgs units, in SI units the energy density is  $\frac{\epsilon_0}{2} (E^2 + c^2 B^2)$ .

Notice that  $f_n(t)$  is a function of time, so also the operators  $a_n(t)$  are (Heisenberg picture). The electric field is the sum over this normal modes (notice that now the position is just a parameter, no longer an operator):

$$\mathbf{E}(x, t) = \sum_n \sqrt{2\hbar\pi\omega_n} [a_n^\dagger(t) + a_n(t)] \mathbf{u}_n(x)$$

Of course now the electric field is an **operator field**, that is, it is a QM operator that is defined at each space-time point  $(x, t)$ .

Notice that an equivalent formulation of the electric field in a finite volume  $V$  is given by defining in a slightly different way the  $u_n(x)$  normal modes and writing:

$$\mathbf{E}(x, t) = \sum_n \sqrt{\frac{2\hbar\pi\omega_n}{V}} [a_n^\dagger(t) + a_n(t)] \mathbf{u}_n(x).$$

We already have calculated the evolution of the operator  $a$  and  $a^\dagger$ . Each of the operator  $a_n$  evolves in the same way:  $a_n(t) = a_n(0)e^{-i\omega_n t}$ . This derives from the Heisenberg equation of motion  $\frac{da_n}{dt} = \frac{i}{\hbar}[\mathcal{H}, a_n(t)] = -i\omega_n a_n(t)$ . The magnetic field can also be expressed in terms of the operators  $a_n$ :

$$\mathbf{B}(x, t) = \sum_n ic_n \sqrt{\frac{2\pi\hbar}{\omega_n}} [a_n^\dagger - a_n] \nabla \times \mathbf{u}_n(x)$$

The strategy has been to use the known forms of the operators for a harmonic oscillator to deduce appropriate operators for the e.m. field. Notice that we could have used the equation  $\frac{dh_n(t)}{dt} = -cf_n(t)$  to eliminate  $f_n$  and write everything in terms of  $h_n$ . This would have corresponded to identifying  $h_n$  with position and  $f_n$  with momentum. Since the Hamiltonian is totally symmetric in terms of momentum and position, the results are unchanged and we can choose either formulations. In the case we chose, comparing the way in which the raising and lowering operators enter in the  $E$  and  $B$  expressions with the way they enter the expressions for position and momentum, we may say that, roughly speaking, the electric field is analogous to the position and the magnetic field is analogous to the momentum of an oscillator.

### 10.2.1 Zero-Point Energy and the Casimir Force

The Hamiltonian operator for the e.m. field has the form of a harmonic oscillator for each mode of the field<sup>34</sup>. As we saw in a previous lecture, the lowest energy of a h.o. is  $\frac{1}{2}\hbar\omega$ . Since there are infinitely many modes of arbitrarily high frequency in any finite volume, it follows that there should be an infinite zero-point energy in any volume of space. Needless to say, this conclusion is unsatisfactory. In order to gain some appreciation for the magnitude of the zero-point energy, we can calculate the zero-point energy in a rectangular cavity due to those field modes whose frequency is less than some cutoff  $\omega_c$ . The mode functions  $u_n(x)$ , solutions of the mode equation for a cavity of dimensions  $L_x \times L_y \times L_z$ , have the vector components

$$u_{n,\alpha} = A_\alpha \cos(k_{n,x}r_x) \sin(k_{n,y}r_y) \sin(k_{n,z}r_z)$$

for  $\{\alpha, \beta, \gamma\} = \{x, y, z\}$  and permutations thereof. The mode  $u_{n,\alpha}(x)$  are labeled by the wave-vector  $\vec{k}_n$  with components:

$$k_{n,\alpha} = \frac{n_\alpha\pi}{L_\alpha}, \quad n_\alpha \in \mathcal{N}$$

and the frequency of the mode is  $\omega_n = \sqrt{k_{n,x}^2 + k_{n,y}^2 + k_{n,z}^2}$ . At least two of the integers must be nonzero, otherwise the mode function would vanish identically.

The amplitudes of the three components  $A_\alpha$  are related by the divergence condition  $\nabla \cdot \vec{u}_n(x) = 0$ , which requires that  $\vec{A} \cdot \vec{k} = 0$ , from which it is clear that there are two linearly independent polarizations (directions of A) for each

<sup>34</sup> See: Leslie E. Ballentine, "Quantum Mechanics A Modern Development", World Scientific Publishing (1998). We follow his presentation in this section.

$k$ , and hence there are two independent modes for each set of positive integers  $(n_x, n_y, n_z)$ . If one of the integers is zero, two of the components of  $u(x)$  will vanish, so there is only one mode in this exceptional case. In this case the electric field can be written as:

$$E(x, t) = \sum_{\alpha=1,2} (\mathcal{E}_\alpha + \mathcal{E}_\alpha^\dagger) = \sum_{\alpha=1,2} \hat{e}_\alpha \sum_n \sqrt{\frac{\hbar\omega_n}{2\epsilon_0 V}} [a_n^\dagger e^{i(\vec{k}_n \cdot \vec{r} - \omega t)} + a_n e^{i(\vec{k}_n \cdot \vec{r} - \omega t)}]$$

Here  $V = L_x L_y L_z$  is the volume of the cavity. Notice that the electric field associated with a single photon of frequency  $\omega_n$  is

$$\mathcal{E}_n = \sqrt{\frac{2\pi\hbar\omega_n}{V}}$$

This energy is a figure of merit for any phenomena relying on atomic interactions with a vacuum field, for example, cavity quantum electrodynamics. In fact,  $\mathcal{E}_n$  may be estimated by equating the quantum mechanical energy of a photon  $\hbar\omega_n$  with its classical energy  $\frac{1}{2} \int dV (E^2 + B^2)$ .

Going back to the calculation of the energy density, if the dimensions of the cavity are large, the allowed values of  $k$  approximate a continuum, and the density of modes in the positive octant of  $k$  space is  $\rho(k) = 2V/\pi^3$  (the factor 2 comes from the two possible polarizations). The zero-point energy density for all modes of frequency less than  $\omega_c$  is then given by

$$E_0 = \frac{2}{V} \sum_{k=1}^{k_c} \frac{1}{2} \hbar\omega_k \approx \frac{2}{V} \frac{1}{8} \int d^3k \rho(k) \frac{1}{2} \hbar\omega(k)$$

where we sum over all positive  $k$  (in the first sum) and multiply by the number of possible polarizations (2). The sum is then approximated by an integral over the positive octant (hence the 1/8 factor). Using  $\omega(k) = kc$  (and  $d^3k = 4\pi k^2 dk$ ), we obtain

$$E_0 = \frac{2}{V} \frac{2V}{\pi^3} \frac{4\pi}{8} \int_{k=0}^{k_c} dk \frac{1}{2} \hbar k^3 c = \frac{c\hbar}{2\pi^2} \int_{k=0}^{k_c} dk k^3 = \frac{\hbar c k_c^4}{8\pi^2}$$

where we set the cutoff wave-vector  $k_c = \omega_c/c$ . The factor  $k_c^4$  indicates that this energy density is dominated by the high-frequency, short-wavelength modes. Taking a minimum wavelength of  $\lambda_c = 2\pi/k = 0.4 \times 10^{-6} \text{m}$ , so as to include the visible light spectrum, yields a zero-point energy density of  $23 \text{ J/m}^3$ . This may be compared with energy density produced by a 100 W light bulb at a distance of 1 m, which is  $2.7 \times 10^{-8} \text{ J/m}^3$ . Of course it is impossible to extract any of the zero-point energy, since it is the minimum possible energy of the field, and so our inability to perceive that large energy density is not incompatible with its existence. Indeed, since most experiments detect only energy differences, and not absolute energies, it is often suggested that the troublesome zero-point energy of the field should simply be omitted. One might even think that this energy is only a constant background to every experimental situation, and that, as such, it has no observable consequences. On the contrary, the vacuum energy has direct measurable consequences, among which the Casimir effect is the most prominent one.

In 1948 H. B. G. Casimir showed that two electrically neutral, perfectly conducting plates, placed parallel in vacuum, modify the vacuum energy density with respect to the unperturbed vacuum. The energy density varies with the separation between the mirrors and thus constitutes a force between them, which scales with the inverse of the fourth power of the mirrors separation. The Casimir force is a small but well measurable quantity. It is a remarkable macroscopic manifestation of a quantum effect and it gives the main contribution to the forces between macroscopic bodies for distances beyond 100nm.

We consider a large cavity of dimensions  $V = L^3$  bounded by conducting walls (see figure). A conducting plate is inserted at a distance  $R$  from one of the  $yz$  faces ( $R \ll L$ ). The new boundary condition at  $x = R$  alters the energy (or frequency) of each field mode. Following Casimir, we shall calculate the energy shift as a function of  $R$ . Let  $W_X$  denote the electromagnetic energy within a cavity whose length in the  $x$  direction is  $X$ . The change in the energy due to the insertion of the plate at  $x = R$  will be

$$\Delta W = (W_R + W_{L-R}) - W_L$$

Each of these three terms is infinite, but the difference will turn out to be finite. Each mode has a zero-point energy of  $\frac{1}{2} \hbar k c$ . But while we can take the continuous

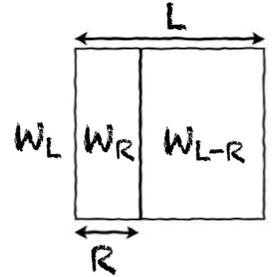


Fig. 14: Geometry of Casimir Effect

approximation in calculating  $W_L$  and  $W_{L-R}$ , for  $W_R$  we have to keep the discrete sum in the  $x$  direction (if  $R$  is small enough). With some calculations (see Ballentine) we find that the change in energy is

$$\Delta W = -\hbar c \frac{\pi^2}{720} \frac{L^2}{R^3}$$

When varying the position  $R$ , an attractive force (minus sign) is created between the conducting plates, equal to

$$F = -\frac{\partial \Delta W}{\partial R} = -\hbar c \frac{\pi^2}{240} \frac{L^2}{R^4}$$

The force per unit area (pressure) is then  $P = -\frac{\pi^2}{240} \frac{\hbar c}{R^4}$ . This is the so-called Casimir force. This force is very difficult to measure. The surfaces must be flat and clean, and free from any electrostatic charge. However, there has been measurements of the Casimir effect, since the experiment by by Sparnaay (1958).

The availability of experimental set-ups that allow accurate measurements of surface forces between macroscopic objects at submicron separations has recently stimulated a renewed interest in the Casimir effect and in its possible applications to micro- and nanotechnology. The Casimir force is highly versatile and changing materials and shape of the boundaries modifies its strength and even its sign. Modifying strength and even sign of the Casimir force has great potential in providing a means for indirect force transmission in nanoscale machines, which is at present not achievable without damaging the components. A contactless method would represent a breakthrough in the future development of nanomachines. More generally, a deeper knowledge of the Casimir force and Casimir torque could provide new insights and design alternatives in the fabrications of micro- and nanoelectromechanical-systems (MEMS and NEMS). Another strong motivation comes from the need to make advantage of the unique properties of Carbon Nanotubes in nanotechnology.

Measuring the Casimir force is also important from a fundamental standpoint as it probes the most fundamental physical system, that is, the quantum vacuum. Furthermore, it is a powerful experimental method for providing constraints on the parameters of a Yukawa-type modification to the gravitational interaction or on forces predicted by supergravity and string theory.

### 10.3 Quantization of the e.m. field in the Coulomb gauge

The quantization procedure and resulting interactions detailed above may appear quite general, but in fact we made an important assumption at the very beginning which will limit their applicability: we considered only the situation with no sources, so we implicitly treated only transverse fields where  $\nabla \cdot \vec{E} = 0$ . Longitudinal fields result from charge distributions  $\rho$  and they do not satisfy a wave equation. By considering only transverse fields, however, we have further avoided the issue of gauge. Since a transverse electric field  $E_T$  satisfies the wave equation, we were able to directly quantize it without intermediate recourse to the vector potential  $A$  and thus we never encountered a choice of gauge. In fact, the procedure can be viewed as corresponding to an implicit choice of gauge  $\phi = 0$ ,  $\nabla \cdot \vec{A} = 0$  that corresponds to a Lorentz gauge.

A more general approach may use the canonical Hamiltonian for a particle of mass  $m$  and charge  $q$  in an electromagnetic field. In this approach, the particle momentum  $p$  is replaced by the canonical momentum  $p - qA/c$ , so the Hamiltonian contains terms like  $\mathcal{H} \sim (p - qA/c)^2/2m$ . In this case, it is still possible to write a wave equation for the potentials. Then the potentials are quantized and for an appropriate choice of gauge we find again the same results. Specifically, for an appropriate choice of gauge, the  $p \cdot A$  terms imply the dipole interaction  $E \cdot d$  that we will use in the following.

Within the Coulomb gauge, the vector potential obeys the wave equation

$$\frac{\partial^2 A}{\partial t^2} - c^2 \nabla^2 A = 0$$

Taking furthermore periodic boundary conditions in a box of volume  $V = L^3$  the quantized electromagnetic field in the Heisenberg picture is:

$$\vec{A}(t, x) = \sum_{\alpha=1,2} \sum_k \sqrt{\frac{2\pi\hbar}{V\omega_k}} \left[ a_{k\alpha} e^{-i(\omega_k t - \vec{k} \cdot \vec{x})} + a_{k\alpha}^\dagger e^{i(\omega_k t - \vec{k} \cdot \vec{x})} \right] \hat{e}_\alpha(k)$$

The field can then be written in terms of the potential as  $E = -\frac{\partial A}{\partial t}$  and we find the similar result as before:

$$E(t, x) = \sum_{\alpha=1,2} \sum_k \sqrt{\frac{2\pi\hbar\omega_k}{V}} \left[ a_{k\alpha} e^{-i(\omega_k t - \vec{k} \cdot \vec{x})} - a_{k\alpha}^\dagger e^{i(\omega_k t - \vec{k} \cdot \vec{x})} \right] \hat{e}_\alpha(k)$$

and

$$B(t, x) = \sum_{\alpha=1,2} \sum_k \sqrt{\frac{2\pi\hbar\omega_k}{V}} \left[ a_{k\alpha} e^{-i(\omega_k t - \vec{k} \cdot \vec{x})} - a_{k\alpha}^\dagger e^{i(\omega_k t - \vec{k} \cdot \vec{x})} \right] (k \times \hat{e}_\alpha(k))$$

### 10.4 States of the e.m. field

Because of the analogies of the e.m. with a set of harmonic oscillators, we can apply the knowledge of the h.o. states to describe the states of the e.m. field. Specifically, we will investigate number states and coherent states.

#### 10.4.1 Photon number eigenstates

We can define number states for each mode of the e.m. field. The Hamiltonian for a single mode is given by  $\mathcal{H}_m = \hbar\omega_m (a_m^\dagger a_m + \frac{1}{2})$  with eigenvectors  $|n_m\rangle$ . The state representing many modes is then given by

$$|n_1, n_2, \dots\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \dots = |\vec{n}\rangle$$

Therefore the  $m^{\text{th}}$  mode of this state is described as containing  $n_m$  photons. These elementary excitations of the e.m. field behave in many respects like particles, carrying energy and momentum. However, the analogy is incomplete, and it is not possible to replace the e.m. field by a gas of photons.

In a state with definite photon numbers, the electric and magnetic fields are indefinite and fluctuating. The probability distributions for the electric and magnetic fields in such a state are analogous to the distributions for the position and momentum of an oscillator in an energy eigenstate. Thus we have for the expectation value of the electric field operator:

$$\langle E(x, t) \rangle = \langle \vec{n} | \sum_m \sqrt{2\hbar\pi\omega_m} [a_m^\dagger(t) + a_m(t)] u_m(x) | \vec{n} \rangle = 0$$

However, the second moment is non-zero:

$$\begin{aligned} \langle |E(x, t)|^2 \rangle &= 2\pi\hbar \sum_{p,m} \sqrt{\omega_p\omega_m} \langle [a_p^\dagger(t) + a_p(t)][a_m^\dagger(t) + a_m(t)] \rangle \vec{u}_p(x) \cdot \vec{u}_m(x) \\ &= 2\pi\hbar \sum_m \omega_m |u_m(x)|^2 \langle [a_m^\dagger(t) + a_m(t)]^2 \rangle = 2\pi\hbar \sum_m \omega_m |u_m(x)|^2 (2n_m + 1) \end{aligned}$$

The sum over all modes is infinite. This divergence problem can often be circumvented (but not solved) by recognizing that a particular experiment will effectively couple to the EM field only over some finite bandwidth, thus we can set cut-offs on the number of modes considered.

Notice that we can as well calculate  $\Delta B$  for the magnetic field, to find the same expression.

#### 10.4.2 Coherent states

A coherent state of the e.m. field is obtained by specifying a coherent state for each of the mode oscillators of the field. Thus the coherent state vector will have the form

$$|\vec{\alpha}\rangle = |\alpha_1 \alpha_2 \dots\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \dots$$

It is parameterized by a denumerably infinite sequence of complex numbers. We now want to calculate the evolution of the electric field for a coherent state. In the Heisenberg picture it is:

$$E(x, t) = \sum_m \sqrt{2\hbar\pi\omega_m} [a_m^\dagger e^{i\omega_m t} + a_m e^{-i\omega_m t}] u_m(x)$$

then, taking the expectation value we find:

$$\langle E(x, t) \rangle = \sum_m \sqrt{2\hbar\pi\omega_m} [\alpha_m^* e^{i\omega_m t} + \alpha_m e^{-i\omega_m t}] u_m(x)$$

This is exactly the same form as a normal mode expansion of a classical solution of Maxwell's equations, with the parameter  $\alpha_m$  representing the amplitude of a classical field mode. In spite of this similarity, a coherent state of the quantized EM field is not equivalent to a classical field, although it does give the closest possible quantum operator, in terms of its expectation value. Even if the average field is equivalent to the classical field, there are still the characteristic quantum fluctuations. A coherent state provides a good description of the e.m. field produced by a laser. Most ordinary light sources emit states of the e.m. field that are very close to a coherent state (lasers), or to a statistical mixture of coherent states (classical sources).

#### A. Fluctuations

We calculate the fluctuations for a single mode  $\Delta E_m$ . From  $\langle |E_m|^2 \rangle = \langle \alpha_m | E_m \cdot E_m | \alpha_m \rangle$  we obtain  $\Delta E_m^2 = 2\pi\hbar\omega_m |\vec{u}_m(x)|^2$ . Indeed, from  $\langle (a_m + a_m^\dagger)^2 \rangle$  we obtain:

$$\langle \alpha_m | (a_m^\dagger)^2 + a_m^2 + a_m^\dagger a_m + a_m a_m^\dagger | \alpha_m \rangle 2\pi\hbar\omega |u_m(x)|^2 = [1 + ((\alpha_m^*)^2 + \alpha_m^2 + 2\alpha_m^* \alpha_m)] 2\pi\hbar\omega |u_m(x)|^2$$

while we have  $\langle a_m + a_m^\dagger \rangle = (\alpha_m + \alpha_m^*) \sqrt{2\pi\hbar\omega} u_m(x)$ , so that we obtain

$$\langle (a_m + a_m^\dagger)^2 \rangle - \langle a_m + a_m^\dagger \rangle^2 = 2\pi\hbar\omega_m |\vec{u}_m(x)|^2$$

This is independent of  $\alpha_m$ , and is equal to the mean square fluctuation in the ground state. The Heisenberg inequality is therefore saturated when the field is in a coherent state,

## B. Photon statistics

The photon number distribution for each mode in a coherent state is obtained as for the h.o. The probability of finding a total of  $n$  photons in the field mode is governed by the Poisson distribution.

The probability of finding  $n$  photons in the mode  $m$  is  $P_\alpha(n_m) = |\langle n_m | \alpha \rangle|^2$ . Using the expansion of the coherent state in terms of the number states that we found for the h.o., we obtain  $P_\alpha(n_m) = \frac{|\alpha_m|^{2n}}{n!} e^{-|\alpha_m|^2}$ . This is a Poisson distribution, with parameter  $|\alpha_m|^2$ . Thus we have  $\langle n_m \rangle = |\alpha_m|^2$ , so that we can rewrite the pdf as  $P(n) = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}$ . From the known properties of the Poisson distribution, we also find  $\Delta n^2 = \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$ .

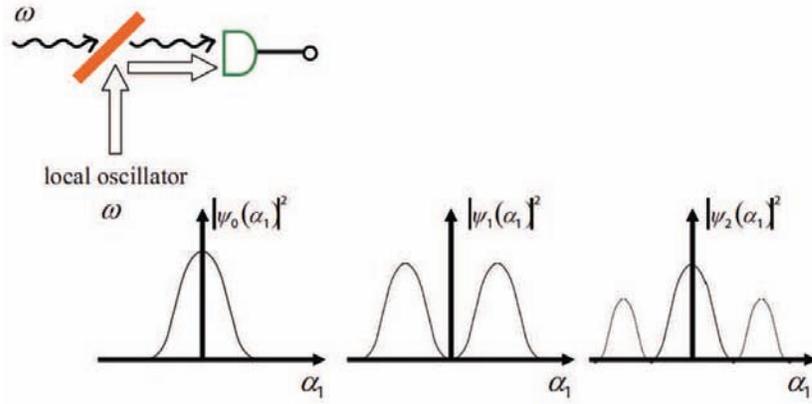
In particular we have the well-know shot-noise scaling  $\frac{\Delta n}{\langle n \rangle} = 1/\sqrt{\langle n \rangle}$  (i.e. the fluctuations go to zero when there are many photons.)

### 10.4.3 Measurement Statistics

We saw in the previous section the photon number distribution for a coherent state. This corresponds to the experimental situation in which we want to measure the number of photons in a field (such as laser light) which is well approximated by a classical field and thus can be represented by a coherent state.

This is not the only type of measurement of the e.m. field that we might want to do. Two other common measurement modalities are homodyne and heterodyne detection<sup>35</sup>.

Homodyne detection corresponds to the measurement of one quadrature amplitude. In practice, one mixes the e.m. field with a local oscillator at with a fixed frequency  $\omega$  (same as the field frequency) before collecting the signal with



© Yoshihisa Yamamoto. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

Fig. 15: Homodyne detection scheme and measurement statistics of the first three photon number eigenstates.

a photon counting detector.

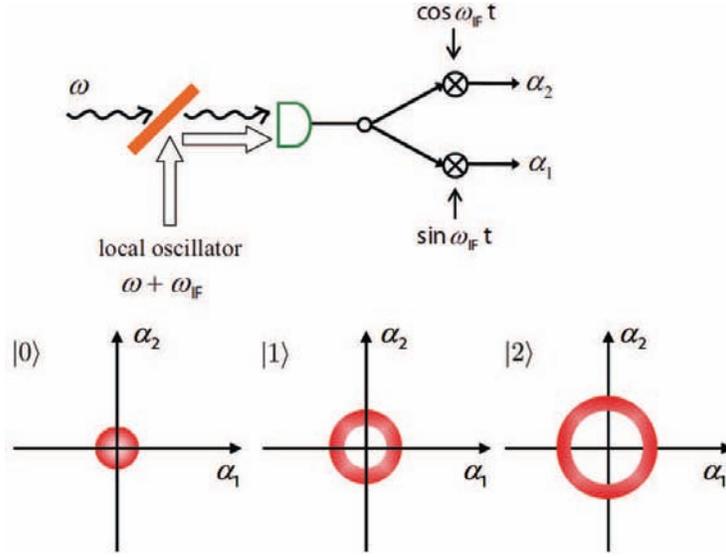
Thus the measurement corresponds to the observable  $O_{ho} = |\alpha_1\rangle \langle \alpha_1|$  (or  $O_{ho} = |\alpha_2\rangle \langle \alpha_2|$ , depending on the phase of the local oscillator), where  $|\alpha_{1,2}\rangle$  are the eigenstates of the quadrature operators  $a_1 = \frac{1}{2}(a + a^\dagger)$  and  $a_2 = \frac{i}{2}(a^\dagger - a)$ . The measurement statistics for a number state  $|m\rangle$  is thus:

$$P_m(\alpha_1) = |\langle \alpha_1 | m \rangle|^2 = \langle \alpha_1 | \frac{(a^\dagger)^n}{n!} | n \rangle = \sqrt{\frac{\omega}{\pi \hbar}} H_m^2(\alpha_1/2) e^{-\alpha_1^2/2}, \quad \langle O_{ho} \rangle = \langle m | O_{ho} | m \rangle = 0, \quad \langle \Delta O_{ho} \rangle = \frac{1}{2}$$

where  $H_n$  is the  $n^{th}$  Hermite polynomial and  $\langle \Delta O \rangle = \sqrt{\langle O^2 \rangle - \langle O \rangle^2}$ . We note that these results correspond to what we had found for the  $x$  operator in the case of the quantum harmonic oscillator.

Heterodyne detection corresponds to the simultaneous measurement of the two quadratures of a field. Operationally, one mixes the e.m. field with a local oscillator of frequency  $\omega$ , modulated at the Intermediate Frequency  $\omega_{IF}$ ; the

<sup>35</sup> We follow here the presentation in Prof. Yamamoto's Lectures



© Yoshihisa Yamamoto. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

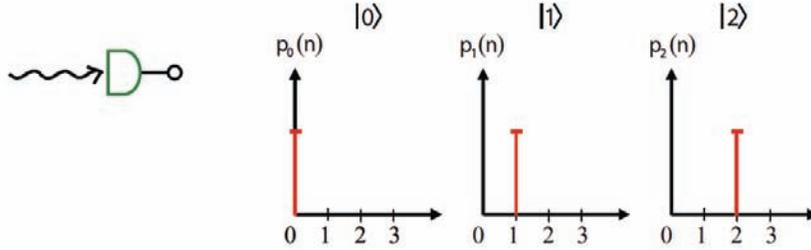
Fig. 16: Heterodyne detection scheme and measurement statistics of the first three photon number eigenstates.

signal, after collection, is demodulated by mixing it with  $\sin(\omega_{IF}t)$  and  $\cos(\omega_{IF}t)$ . Thus the measurement corresponds to the observable  $O_{he} = |\alpha\rangle\langle\alpha|$ . The measurement statistics for a number state  $|m\rangle$  is thus:

$$P_m(\alpha) = |\langle\alpha|m\rangle|^2 = \frac{e^{-|\alpha|^2}|\alpha|^{2m}}{m!}, \quad \langle O_{he} \rangle = \langle m|O_{he}|m\rangle = |\alpha|^2, \quad \langle \Delta O_{he} \rangle = |\alpha|^2$$

(note that of course this is equivalent to the case were we measured a number state for a coherent state). The measurement statistics for a coherent state  $|\beta\rangle$ , would be

$$P_\beta(\alpha) = |\langle\alpha|\beta\rangle|^2 = e^{-|\beta-\alpha|^2}$$



© Yoshihisa Yamamoto. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

Fig. 17: Photon counting detection scheme and measurement statistics of the first three photon number eigenstates.

For comparison, photon counting is of course the measurement of the observable  $O_n = |n\rangle\langle n|$ , with statistics for a number state  $|m\rangle$

$$P_m(n) = |\langle n|m\rangle|^2 = \delta_{m,n}, \quad \langle O_n \rangle = \langle m|O_n|m\rangle = \delta_{m,n}, \quad \langle \Delta O_n \rangle = 0$$

## 10.5 Atomic interactions with the quantized field

Let us consider the interaction of isolated neutral atoms with optical fields. Such atoms alone have no net charge and no permanent electric dipole moment. In an electric field  $\vec{E}$  associated, e.g., with an electromagnetic wave, the atoms

do develop an electric dipole moment  $\vec{d}$  which can then interact with the electric field with an interaction energy  $V$  given by

$$V = \vec{d} \cdot \vec{E}$$

We have already treated a similar case in a semiclassical way, although we were interested in the interaction with a magnetic field. The semi-classical treatment of this interaction, is quite similar: we treat the atom quantum mechanically and therefore consider  $\vec{d}$  as an operator, but treat the electromagnetic field classically and so consider  $E$  as a vector. We can write the dipole moment operator as

$$\vec{d} = e\vec{r} = \sum_{k,h} |k\rangle \langle k| \vec{d} |h\rangle \langle h|$$

where  $\{|k\rangle\}$  forms a complete basis. Transitions are only possible between states with different  $h$  and  $k$ :

$$\vec{d}_{h,k} = \langle h| \vec{d} |k\rangle \neq 0 \quad \text{iif} \quad k \neq h$$

and we will consider for simplicity a two-level atom:

$$\vec{d} = |0\rangle \langle 1| \vec{d}_{01} + |1\rangle \langle 0| \vec{d}_{10}$$

Let's first consider a single mode **classical** electromagnetic field, given by  $E = \vec{\mathcal{E}} e^{-i\omega t} + \vec{\mathcal{E}}^* e^{i\omega t}$ . The full *semi-classical* (sc) Hamiltonian is then:

$$\mathcal{H}_{sc} = \frac{1}{2} \hbar \omega_0 (|1\rangle \langle 1| - |0\rangle \langle 0|) - (|0\rangle \langle 1| \vec{d}_{01} + |1\rangle \langle 0| \vec{d}_{10}) \cdot (\vec{\mathcal{E}} e^{-i\omega t} + \vec{\mathcal{E}}^* e^{i\omega t})$$

If we assume  $\{\vec{d}_{10}, \mathcal{E}\} \in \mathbb{R}$ , we can rewrite this as

$$\mathcal{H}_{sc} = \frac{1}{2} \hbar \omega_0 \sigma_z - 2\sigma_x \vec{d}_{10} \cdot \vec{\mathcal{E}} \cos(\omega t)$$

Notice the correspondence with the spin Hamiltonian  $\mathcal{H}_{spin} = \Omega \sigma_z + B \cos(\omega t) \sigma_x$  describing the interaction of a spin with a time-varying, classical magnetic field.

We can now go into the interaction frame defined by the Hamiltonian  $\mathcal{H}_0 = \frac{1}{2} \omega_0 \sigma_z$ . Then we have:

$$\tilde{\mathcal{H}}_{sc} = -(|0\rangle \langle 1| \vec{d}_{01} e^{i\omega_0 t} + |1\rangle \langle 0| \vec{d}_{10} e^{-i\omega_0 t}) \cdot (\vec{\mathcal{E}} e^{-i\omega t} + \vec{\mathcal{E}}^* e^{i\omega t})$$

On resonance ( $\omega_0 = \omega$ ) we retain only time-independent contributions to the Hamiltonian (RWA), then

$$\tilde{\mathcal{H}}_{sc} \approx -(|1\rangle \langle 0| d \mathcal{E} + |0\rangle \langle 1| d^* \mathcal{E}^*)$$

Assuming for example that  $d \mathcal{E}$  is real, we obtain an Hamiltonian  $-d \mathcal{E} \sigma_x$ , in perfect analogy with the TLS already studied. (A more general choice of  $d \mathcal{E}$  just gives an Hamiltonian at some angle in the xy plane).

Now let us consider a full quantum-mechanical treatment of this problem. The interaction between an atom and a quantized field appears much the same as the semiclassical interaction. Starting with the dipole Hamiltonian for a two-level atom, we replace  $E$  by the corresponding operator, obtaining the interaction Hamiltonian

$$\begin{aligned} V &= -\vec{d} \cdot \vec{E} = -\sum_{\alpha} (\vec{\mathcal{E}}_{\alpha} + \vec{\mathcal{E}}_{\alpha}^{\dagger}) \cdot (\vec{d} |1\rangle \langle 0| + \vec{d}^* |0\rangle \langle 1|) \\ &= -\sum_{\alpha} \sum_m \sqrt{\frac{2\pi \hbar \omega_m}{V}} [a_m^{\dagger} e^{-i\vec{k}_m \cdot \vec{r}} + a_m e^{i\vec{k}_m \cdot \vec{r}}] (d_{\alpha} |1\rangle \langle 0| + d_{\alpha}^* |0\rangle \langle 1|) \end{aligned}$$

(where  $\alpha$  is the polarization and  $m$  the mode). As in the semiclassical analysis, the Hamiltonian contains four terms, which now have a clearer physical picture:

$a_m^\dagger 0\rangle\langle 1 $	Atom decays from $ 1\rangle \rightarrow  0\rangle$ and emits a photon (in the $m^{th}$ mode).
$a_m 1\rangle\langle 0 $	Atom is excited from $ 0\rangle \rightarrow  1\rangle$ and absorbs a photon (from the $m^{th}$ mode).
$a_m^\dagger 1\rangle\langle 0 $	Atom is excited from $ 0\rangle \rightarrow  1\rangle$ and emits a photon (in the $m^{th}$ mode).
$a_m 0\rangle\langle 1 $	Atom decays from $ 1\rangle \rightarrow  0\rangle$ and absorbs a photon (from the $m^{th}$ mode).

For photons near resonance with the atomic transition, the first two processes conserve energy; the second two processes do not conserve energy, and intuition suggests that they may be neglected. In fact, there is a direct correspondence between the RWA and energy conservation: the second two processes are precisely those fast-rotating terms we disregarded previously.

Consider the total Hamiltonian :

$$\mathcal{H} = \mathcal{H}_0 + V = \frac{\hbar}{2}\omega_0\sigma_z + \sum_m \hbar\omega_m \left( a_m^\dagger a_m + \frac{1}{2} \right) + V$$

If we go to the interaction frame defined by the Hamiltonian  $\mathcal{H}_0$ , each mode acquires a time dependence  $e^{\pm i\omega_m t}$  while the atom acquires a time dependence  $e^{\pm i\omega_0 t}$ :

$$\sum (\mathcal{E}_m a_m^\dagger e^{-i\vec{k}_m \cdot \vec{r}} e^{i\omega_m t} + \mathcal{E}_m^* a_m e^{i\vec{k}_m \cdot \vec{r}} e^{-i\omega_m t}) \cdot (e^{+i\omega_0 t} d_\alpha |1\rangle\langle 0| + d_\alpha^* e^{-i\omega_0 t} |0\rangle\langle 1|)$$

where  $\mathcal{E}_m = \sqrt{\frac{2\pi\hbar\omega_m}{V}}$ . Thus the time-dependent factors acquired are

$$\begin{aligned} a_m^\dagger|0\rangle\langle 1| &\rightarrow a_m^\dagger|0\rangle\langle 1|e^{+i(\omega_0-\omega_m)t} & a_m|1\rangle\langle 0| &\rightarrow a_m|1\rangle\langle 0|e^{-i(\omega_0-\omega_m)t} \\ a_m^\dagger|1\rangle\langle 0| &\rightarrow a_m^\dagger|1\rangle\langle 0|e^{i(\omega_0+\omega_m)t} & a_m|0\rangle\langle 1| &\rightarrow a_m|0\rangle\langle 1|e^{-i(\omega_0+\omega_m)t} \end{aligned}$$

For frequencies  $\omega_m$  near resonance  $\omega_m \approx \omega_0$ , we only retains the first two terms.

Then, defining the *single-photon Rabi frequency*,  $g_{m,\alpha} = -\frac{d_\alpha}{\hbar} \sqrt{\frac{2\pi\hbar\omega_m}{V}} e^{i\vec{k}_m \cdot \vec{r}}$ , the Hamiltonian in the interaction picture and in the RWA approximation is

$$\mathcal{H} = \sum_m \hbar (g_{m,\alpha} a_m |1\rangle\langle 0| + g_{m,\alpha}^* a_m^\dagger |0\rangle\langle 1|)$$

From now on we assume an e.m. with a single mode (or we assume that only one mode is on resonance). We can write a general state as  $|\psi\rangle = \sum_n \alpha_n(t) |1n\rangle + \beta_n(t) |0n\rangle$ , where  $|n\rangle = |n_m\rangle$  is a state of the given mode  $m$  we retain and here I will call the Rabi frequency for the mode of interest  $g$ . The evolution is given by:

$$\begin{aligned} i\hbar \sum_n \dot{\alpha}_n |1n\rangle + \dot{\beta}_n |0n\rangle &= \hbar \sum_n g [\alpha_n \sigma_- a^\dagger |1n\rangle + \beta_n \sigma_+ a |0n\rangle] \\ &= \hbar \sum_n g [\alpha_n \sqrt{n+1} |0, n+1\rangle + \beta_n \sqrt{n} |1, n-1\rangle] \end{aligned}$$

We then project these equations on  $\langle 1n|$  and  $\langle 0n|$ :

$$\begin{aligned} i\hbar \dot{\alpha}_n &= \hbar g \beta_{n+1}(t) \sqrt{n+1} \\ i\hbar \dot{\beta}_n &= \hbar g \alpha_{n-1}(t) \sqrt{n} \end{aligned}$$

to obtain a set of equations:

$$\begin{cases} \dot{\alpha}_n = -ig\sqrt{n+1}\beta_{n+1} \\ \dot{\beta}_{n+1} = -ig\sqrt{n+1}\alpha_n \end{cases}$$

This is a closed system of differential equations and we can solve for  $\alpha_n, \beta_{n+1}$ .

We consider a more general case, where the field-atom are not exactly on resonance. We define  $\Delta = \frac{1}{2}(\omega_0 - \omega)$ , where  $\omega = \omega_m$  for the mode considered. Then the Hamiltonian is:

$$\mathcal{H} = \hbar (g a |1\rangle\langle 0| + g^* a^\dagger |0\rangle\langle 1|) + \hbar\Delta (|1\rangle\langle 1| - |0\rangle\langle 0|)$$

We can assume that initially the atom is in the excited state  $|1\rangle$  (that is,  $\beta_n(0) = 0, \forall n$ ). Then we have:

$$\alpha_n(t) = \alpha_n(0)e^{i\Delta t/2} \left[ \cos\left(\frac{\Omega_n t}{2}\right) - \frac{i\Delta}{\Omega_n} \sin\left(\frac{\Omega_n t}{2}\right) \right]$$

$$\beta_n(t) = -\alpha_n(0)e^{-i\Delta t/2} \left[ \frac{2ig\sqrt{n+1}}{\Omega_n} \sin\left(\frac{\Omega_n t}{2}\right) \right]$$

with  $\Omega_n^2 = \Delta^2 + 4g^2(n+1)$ . If initially there is no field (i.e. the e.m. field is in the vacuum state) and the atom is in the excited state, then  $\alpha_0(0) = 1$ , while  $\alpha_n(0) = 0 \forall n \neq 0$ . Then there are only two components that are different than zero:

$$\alpha_0(t) = e^{i\Delta t/2} \left[ \cos\left(\frac{\Omega_0 t}{2}\right) - \frac{i\Delta}{\sqrt{\Delta^2 + 4g^2}} \sin\left(\frac{\Omega_0 t}{2}\right) \right]$$

$$\beta_0(t) = -e^{-i\Delta t/2} \frac{2ig}{\sqrt{\Delta^2 + 4g^2}} \sin\left(\frac{\Omega_0 t}{2}\right)$$

or on resonance ( $\Delta = 0$ )

$$\langle 1, n=0 | \psi(t) \rangle = \alpha_0(t) = \cos\left(\frac{gt}{2}\right)$$

$$\langle 0, n=0 | \psi(t) \rangle = \beta_0(t) = -i \sin\left(\frac{gt}{2}\right)$$

Thus, even in the absence of field, it is possible to make the transition from the ground to the excited state! In the semiclassical case (where the field is treated as classical) we would have no transition at all. These are called Rabi vacuum oscillations.



# 11. Perturbation Theory

## 11.1 Time-independent perturbation theory

11.1.1 Non-degenerate case

11.1.2 Degenerate case

11.1.3 The Stark effect

## 11.2 Time-dependent perturbation theory

11.2.1 Review of interaction picture

11.2.2 Dyson series

11.2.3 Fermi's Golden Rule

## 11.1 Time-independent perturbation theory

Because of the complexity of many physical problems, very few can be solved exactly (unless they involve only small Hilbert spaces). In particular, to analyze the interaction of radiation with matter we will need to develop approximation methods<sup>36</sup>.

### 11.1.1 Non-degenerate case

We have an Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \epsilon V$$

where we know the eigenvalue of the unperturbed Hamiltonian  $\mathcal{H}_0$  and we want to solve for the perturbed case  $\mathcal{H} = \mathcal{H}_0 + \epsilon V$ , in terms of an expansion in  $\epsilon$  (with  $\epsilon$  varying between 0 and 1). The solution for  $\epsilon \rightarrow 1$  is the desired solution.

We assume that we know exactly the energy eigenkets and eigenvalues of  $\mathcal{H}_0$ :

$$\mathcal{H}_0 |k\rangle = E_k^{(0)} |k\rangle$$

As  $\mathcal{H}_0$  is hermitian, its eigenkets form a complete basis  $\sum_k |k\rangle\langle k| = \mathbb{1}$ . We assume at first that the energy spectrum is not degenerate (that is, all the  $E_k^{(0)}$  are different, in the next section we will study the degenerate case). The eigensystem for the total hamiltonian is then

$$(\mathcal{H}_0 + \epsilon V) |\varphi_k\rangle_\epsilon = E_k(\epsilon) |\varphi_k\rangle_\epsilon$$

where  $\epsilon = 1$  is the case we are interested in, but we will solve for a general  $\epsilon$  as a perturbation in this parameter:

$$|\varphi_k\rangle = \left| \varphi_k^{(0)} \right\rangle + \epsilon \left| \varphi_k^{(1)} \right\rangle + \epsilon^2 \left| \varphi_k^{(2)} \right\rangle + \dots, \quad E_k = E_k^{(0)} + \epsilon E_k^{(1)} + \epsilon^2 E_k^{(2)} + \dots$$

---

<sup>36</sup> A very good treatment of perturbation theory is in Sakurai's book –J.J. Sakurai “Modern Quantum Mechanics”, Addison-Wesley (1994), which we follow here.

where of course  $|\varphi_k^{(0)}\rangle = |k\rangle$ . When  $\epsilon$  is small, we can in fact approximate the total energy  $E_k$  by  $E_k^{(0)}$ . The energy shift due to the perturbation is then only  $\Delta_k = E_k - E_k^{(0)}$  and we can write:

$$(\mathcal{H}_0 + \epsilon V) |\varphi_k\rangle = (E_k^{(0)} + \Delta_k) |\varphi_k\rangle_\epsilon \rightarrow (E_k^{(0)} - \mathcal{H}_0) |\varphi_k\rangle = (\epsilon V - \Delta_k) |\varphi_k\rangle$$

Then, we project onto  $\langle k|$ :

$$\langle k| (E_k^{(0)} - \mathcal{H}_0) |\varphi_k\rangle = \langle k| (\epsilon V - \Delta_k) |\varphi_k\rangle$$

The LHS is zero since  $\langle k| \mathcal{H}_0 |\varphi_k\rangle = \langle k| E_k^{(0)} |\varphi_k\rangle$ , and from the RHS  $\langle k| (\epsilon V - \Delta_k) |\varphi_k\rangle = 0$  we obtain:

$$\Delta_k = \epsilon \frac{\langle k| V |\varphi_k\rangle}{\langle k| \varphi_k\rangle} \rightarrow \Delta_k = \epsilon \langle k| V |\varphi_k\rangle$$

where we set  $\langle k| \varphi_k\rangle = 1$  (a non-canonical normalization, although, as we will see, it is approximately valid). Using the expansion above, we can replace  $\Delta_k$  by  $\epsilon E_k^1 + \epsilon^2 E_k^2 + \dots$  and  $|\varphi_k\rangle$  by its expansion:

$$\epsilon E_k^1 + \epsilon^2 E_k^2 + \dots = \epsilon \langle k| V (|k\rangle + \epsilon \varphi_k^{(1)}\rangle + \epsilon^2 \varphi_k^{(2)}\rangle + \dots)$$

and equating terms of the same order in  $\epsilon$  we obtain:

$$E_k^n = \langle k| V \varphi_k^{(n-1)}\rangle$$

This is a recipe to find the energy at all orders based only on the knowledge of the eigenstates of lower orders. However, the question still remains: how do we find  $\varphi_k^{(n-1)}\rangle$ ?

We could think of solving the equation:

$$(E_k^{(0)} - \mathcal{H}_0) |\varphi_k\rangle = (\epsilon V - \Delta_k) |\varphi_k\rangle \quad (*)$$

for  $|\varphi_k\rangle$ , by inverting the operator  $(E_k^{(0)} - \mathcal{H}_0)$  and again doing an expansion of  $|\varphi_k\rangle$  to equate terms of the same order:

$$|k\rangle + \epsilon \varphi_k^{(1)}\rangle + \dots = (E_k^{(0)} - \mathcal{H}_0)^{-1} (\epsilon V - \Delta_k) (|k\rangle + \epsilon \varphi_k^{(1)}\rangle + \dots)$$

Unfortunately this promising approach is not correct, since the operator  $(E_k^{(0)} - \mathcal{H}_0)^{-1}$  is not always well defined. Specifically, there is a singularity for  $(E_k^{(0)} - \mathcal{H}_0)^{-1} |k\rangle$ . What we need is to make sure that  $(E_k^{(0)} - \mathcal{H}_0)^{-1}$  is never applied to eigenstates of the unperturbed Hamiltonian, that is, we need  $|\psi_k\rangle = (\epsilon V - \Delta_k) |\varphi_k\rangle \neq |k\rangle$  for any  $|\varphi_k\rangle$ . We thus define the projector  $P_k = \mathbb{1} - |k\rangle \langle k| = \sum_{h \neq k} |h\rangle \langle h|$ . Then we can ensure that  $\forall |\psi\rangle$  the projected state  $|\psi\rangle' = P_k |\psi\rangle$  is such that  $\langle k| \psi\rangle' = 0$  since this is equal to

$$\langle k| P_k |\psi\rangle = \langle k| \psi\rangle - \langle k| k\rangle \langle k| \psi\rangle = 0$$

Now, using the projector,  $(E_k^{(0)} - \mathcal{H}_0)^{-1} P_k |\psi\rangle$  is well defined. We then take the equation  $(*)$  and multiply it by  $P_k$  from the left:

$$P_k (E_k^{(0)} - \mathcal{H}_0) |\varphi_k\rangle = P_k (\epsilon V - \Delta_k) |\varphi_k\rangle.$$

Since  $P_k$  commutes with  $\mathcal{H}_0$  (as  $|k\rangle$  is an eigenstate of  $\mathcal{H}_0$ ) we have  $P_k (E_k^{(0)} - \mathcal{H}_0) |\varphi_k\rangle = (E_k^{(0)} - \mathcal{H}_0) P_k |\varphi_k\rangle$  and we can rewrite the equation as

$$P_k |\varphi_k\rangle = (E_k^{(0)} - \mathcal{H}_0)^{-1} P_k (\epsilon V - \Delta_k) |\varphi_k\rangle$$

We can further simplify this expression, noting that  $P_k |\varphi_k\rangle = |\varphi_k\rangle - |k\rangle \langle k| \varphi_k\rangle = |\varphi_k\rangle - |k\rangle$  (since we adopted the normalization  $\langle k| \varphi_k\rangle = 1$ ). Finally we obtain:

$$|\varphi_k\rangle = |k\rangle + (E_k^{(0)} - \mathcal{H}_0)^{-1} P_k (\epsilon V - \Delta_k) |\varphi_k\rangle \quad (**)$$

This equation is now ready to be solved by using the perturbation expansion. To simplify the expression, we define the operator  $R_k$

$$R_k = (E_k^{(0)} - \mathcal{H}_0)^{-1} P_k = \sum_{h \neq k} \frac{|h\rangle \langle h|}{E_k^0 - E_h^0}$$

Now using the expansion

$$|k\rangle + \epsilon |\varphi_k^{(1)}\rangle + \dots = |k\rangle + R_k \epsilon (V - E_k^1 - \epsilon E_k^2 - \dots) (|k\rangle + \epsilon |\varphi_k^{(1)}\rangle + \dots)$$

we can solve term by term to obtain:

$$1^{st} \text{ order: } |\varphi_k^{(1)}\rangle = R_k (V - E_k^1) |k\rangle = R_k (V - \langle k|V|k\rangle) |k\rangle = R_k V |k\rangle$$

(where we used the expression for the first order energy and the fact that  $R_k |k\rangle = 0$  by definition). We can now calculate the second order energy, since we know the first order eigenstate:

$$E_k^2 = \langle k|V|\varphi_k^{(1)}\rangle = \langle k|V R_k V |k\rangle = \langle k|V \left( \sum_{h \neq k} \frac{|h\rangle \langle h|}{E_k^0 - E_h^0} \right) V |k\rangle$$

or explicitly

$$E_k^2 = \sum_{h \neq k} \frac{|V_{kh}|^2}{E_k^0 - E_h^0}$$

Then the second order eigenstate is

$$2^{nd} \text{ order: } \varphi_k^2\rangle = R_k V R_k V |k\rangle$$

### A. Formal Solution

We can also find a more formal expression that can yield the solution to all orders. We rewrite Eq. (\*\*) using  $R_k$  and obtain

$$|\varphi_k\rangle = |k\rangle + R_k (\epsilon V - \Delta_k) |\varphi_k\rangle = R_k H_1 |\varphi_k\rangle$$

where we defined  $H_1 = (\epsilon V - \Delta_k)$ . Then by iteration we can write:

$$|\varphi_k\rangle = |k\rangle + R_k H_1 (|k\rangle + R_k H_1 |\varphi_k\rangle) = |k\rangle + R_k H_1 |k\rangle + R_k H_1 R_k H_1 |\varphi_k\rangle$$

and in general:

$$|\varphi_k\rangle = |k\rangle + R_k H_1 |k\rangle + R_k H_1 R_k H_1 |k\rangle + \dots + (R_k H_1)^n |k\rangle + \dots$$

This is just a geometric series, with formal solution:

$$|\varphi_k\rangle = (\mathbb{1} - R_k H_1)^{-1} |k\rangle$$

### B. Normalization

In deriving the TIPT we introduced a non-canonical normalization  $\langle k|\varphi_k\rangle = 1$ , which implies that the perturbed state  $|\varphi_k\rangle$  is not normalized. We can then define a properly normalized state as

$$|\psi_k\rangle = \frac{|\varphi_k\rangle}{\sqrt{\langle \varphi_k|\varphi_k\rangle}}$$

so that  $\langle k|\psi_k\rangle = 1/\sqrt{\langle \varphi_k|\varphi_k\rangle}$ . We can calculate perturbatively the normalization factor  $\langle \varphi_k|\varphi_k\rangle$ :

$$\langle \varphi_k|\varphi_k\rangle = \langle k + \epsilon \varphi_k^1 + \dots | k + \epsilon \varphi_k^1 + \dots \rangle = 1 + \cancel{\epsilon \langle k|\varphi_k^1\rangle} + \dots + \epsilon^2 \langle \varphi_k^1|\varphi_k^1\rangle + \dots = 1 + \epsilon^2 \sum_{h \neq k} \frac{|V_{kh}|^2}{(E_h^0 - E_k^0)^2}$$

Notice that the state is correctly normalized up to the second order in  $\epsilon$ .

### C. Anti-crossing

Consider two levels,  $h$  and  $k$  with energies  $E_k^0$  and  $E_h^0$  and assume that we apply a perturbation  $V$  which connects only these two states (that is,  $V$  is such that  $\langle l|V|j\rangle = 0$  and it is different than zero only for the transition from  $h$  to  $k$ :  $\langle h|V|k\rangle \neq 0$ .)

If the perturbation is small, we can ask what are the perturbed state energies.

The first order is zero by the choice of  $V$ , then we can calculate the second order:

$$E_k^{(2)} = \sum_{j \neq k} \frac{|V_{kj}|^2}{E_k^0 - E_j^0} = \frac{|V_{kh}|^2}{E_k^0 - E_h^0}$$

and similarly

$$E_h^{(2)} = \sum_{j \neq h} \frac{|V_{hj}|^2}{E_h^0 - E_j^0} = \frac{|V_{kh}|^2}{E_h^0 - E_k^0} = -E_k^{(2)}.$$

This opposite energy shift will be more important (more noticeable) when the energies of the two levels  $E_k^0$  and  $E_h^0$  are close to each other. Indeed, in the absence of the perturbation, the two energy levels would “cross” when  $E_k^0 = E_h^0$ . If we add the perturbation, however, the two levels are repelled with opposite energy shifts. We describe what is happening as an “anti-crossing” of the levels: even as the levels become connected by an interaction, the levels never meet (never have the same energy) since each level gets shifted by the same amount in opposite directions.

### D. Example: TLS energy splitting from perturbation

Consider the Hamiltonian  $\mathcal{H} = \omega\sigma_z + \epsilon\Omega\sigma_x$ . For  $\epsilon = 0$  the eigenstates are  $|k\rangle = \{|0\rangle, |1\rangle\}$  and eigenvalues  $E_k^0 = \pm\omega$ . We also know how to solve exactly this simple problem by diagonalizing the entire matrix:

$$E_{1,2} = \pm\sqrt{\omega^2 + \epsilon^2\Omega^2},$$

$$|\varphi_1\rangle = \cos(\vartheta/2)|0\rangle + \sin(\vartheta/2)|1\rangle, \quad |\varphi_2\rangle = \cos(\vartheta/2)|1\rangle - \sin(\vartheta/2)|0\rangle \quad \text{with} \quad \vartheta = \arctan(\epsilon\Omega/\omega)$$

For  $\epsilon \ll 1$  we can expand in series these results to find:

$$E_{1,2} \approx \pm(\omega + \frac{\epsilon^2\Omega^2}{2\omega} + \dots)$$

$$|\varphi_1\rangle \approx |0\rangle + \frac{\vartheta}{2}|1\rangle = |0\rangle + \frac{\epsilon\Omega}{2\omega}|1\rangle \quad |\varphi_2\rangle \approx |1\rangle - \frac{\vartheta}{2}|0\rangle = |1\rangle - \frac{\epsilon\Omega}{2\omega}|0\rangle$$

As an exercise, we can find as well the results of TIPT. First we find that the first order energy shift is zero, since  $E_k^1 = \langle k|V|k\rangle = \langle 0|(\Omega\sigma_x)|0\rangle = 0$  (and same for  $\langle 1|(\Omega\sigma_x)|1\rangle$ ). Then we can calculate the first order eigenstate:

$$\varphi_1^1 = |0\rangle + (E_1^0 - \mathcal{H}_0)^{-1}P_1V|0\rangle = |0\rangle + [\omega(\mathbb{1} - \sigma_z)]^{-1}|1\rangle\langle 1|\epsilon\Omega\sigma_x|0\rangle = |0\rangle + \frac{1}{2\omega}\epsilon\Omega|1\rangle\langle 1|\sigma_x|0\rangle = |0\rangle + \epsilon\frac{\Omega}{2\omega}|1\rangle$$

similarly, we find  $\varphi_2^1 = |1\rangle - \epsilon\frac{\Omega}{2\omega}|0\rangle$ . Finally, the second order energy shift is  $E_1^2 = \frac{|V_{12}|^2}{E_1^0 - E_2^0} = \frac{(\epsilon\Omega)^2}{2\omega}$  in agreement with the result from the series expansion.

We can also look at the level anti-crossing: If we vary the energy  $\omega$  around zero, the two energy levels cross each other.

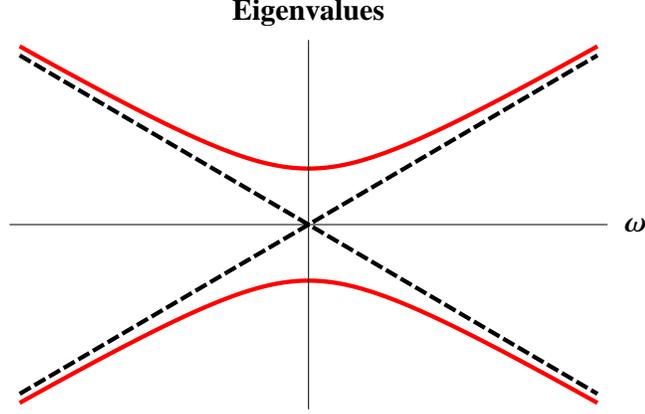


Fig. 18: Level anticrossing: Eigenvalues of the Hamiltonian  $\mathcal{H} = \omega\sigma_z + \epsilon\Omega\sigma_x$  as a function of  $\omega$ . Dashed lines:  $\Omega = 0$ . Red lines:  $\Omega \neq 0$  showing the anticrossing.

### 11.1.2 Degenerate case

If there are degenerate (or quasi-degenerate) eigenvalues of the unperturbed Hamiltonian  $\mathcal{H}_0$ , the expansion used above is no longer valid. There are two problems:

1. If  $|k'\rangle, |k''\rangle, \dots$  have the same eigenvalue, we can choose any combination of them as the unperturbed eigenket. But then, if we were to find the perturbed eigenket  $|\psi_k\rangle$ , to which state would this go to when  $\epsilon \rightarrow 0$ ?
2. The term  $R_k = \frac{P_k}{E_k^{(0)} - \mathcal{H}_0}$  can be singular for the degenerate eigenvalues.

Assume there is a  $d$ -fold degeneracy of the eigenvalue  $E_d$ , with the unperturbed eigenkets  $\{|k_i\rangle\}$  forming a subspace  $\mathbb{H}_d$ . We can then define the projectors  $Q_d = \sum_{k_i \in \mathbb{H}_d} |k_i\rangle\langle k_i|$  and  $P_d = \mathbb{1} - Q_d$ . These projectors also define subspaces of the total Hilbert space  $\mathbb{H}$  that we will call  $\mathbb{H}_d$  (spanned by  $Q_d$ ) and  $\mathbb{H}_{\bar{d}}$  (spanned by  $P_d$ ).

Notice that because of their nature of projectors, we have the following identities:

$$P_d^2 = P_d, \quad Q_d^2 = Q_d, \quad P_d Q_d = Q_d P_d = 0 \quad \text{and} \quad P_d + Q_d = \mathbb{1}.$$

We then rewrite the eigenvalue equation as:

$$\begin{aligned} (\mathcal{H}_0 + \epsilon V) |\varphi_k\rangle = E_k |\varphi_k\rangle &\rightarrow \mathcal{H}_0(Q_d + P_d) |\varphi_k\rangle + \epsilon V(Q_d + P_d) |\varphi_k\rangle = E_k(Q_d + P_d) |\varphi_k\rangle \\ &\rightarrow (Q_d + P_d)\mathcal{H}_0 |\varphi_k\rangle + \epsilon V(Q_d + P_d) |\varphi_k\rangle = E_k(Q_d + P_d) |\varphi_k\rangle \end{aligned}$$

where we used the fact that  $[\mathcal{H}_0, Q_d] = [\mathcal{H}_0, P_d] = 0$  since the projectors are diagonal in the Hamiltonian basis. We then multiply from the left by  $Q_d$  and  $P_d$ , obtaining 2 equations:

1.  $P_d \times [(Q_d + P_d)\mathcal{H}_0 |\varphi_k\rangle + \epsilon V(Q_d + P_d) |\varphi_k\rangle] = P_d \times (E_k(Q_d + P_d) |\varphi_k\rangle)$   
 $\rightarrow \mathcal{H}_0 P_d |\varphi_k\rangle + \epsilon P_d V(Q_d + P_d) |\varphi_k\rangle = E_k P_d |\varphi_k\rangle$
2.  $Q_d \times [(Q_d + P_d)\mathcal{H}_0 |\varphi_k\rangle + \epsilon V(Q_d + P_d) |\varphi_k\rangle] = Q_d \times (E_k(Q_d + P_d) |\varphi_k\rangle)$   
 $\rightarrow \mathcal{H}_0 Q_d |\varphi_k\rangle + \epsilon Q_d V(Q_d + P_d) |\varphi_k\rangle = E_k Q_d |\varphi_k\rangle$

and we simplify the notation by setting  $|\psi_k\rangle = P_d |\varphi_k\rangle$  and  $|\chi_k\rangle = Q_d |\varphi_k\rangle$

$$\mathcal{H}_0 |\psi_k\rangle + \epsilon P_d V(|\chi_k\rangle + |\psi_k\rangle) = E_k |\psi_k\rangle$$

$$\mathcal{H}_0 |\chi_k\rangle + \epsilon Q_d V(|\chi_k\rangle + |\psi_k\rangle) = E_k |\chi_k\rangle$$

which gives a set of coupled equations in  $|\psi_k\rangle$  and  $|\chi_k\rangle$ :

1.  $\epsilon P_d V |\chi_k\rangle = (E_k - \mathcal{H}_0 - \epsilon P_d V P_d) |\psi_k\rangle$
2.  $\epsilon Q_d V |\psi_k\rangle = (E_k - \mathcal{H}_0 - \epsilon Q_d V Q_d) |\chi_k\rangle$

Now  $(E_k - \mathcal{H}_0 - \epsilon P_d V P_d)^{-1}$  is finally well defined in the  $P_d$  subspace, so that we can solve for  $|\psi_k\rangle$  from (1.):

$$|\psi_k\rangle = \epsilon P_d (E_k - \mathcal{H}_0 - \epsilon P_d V P_d)^{-1} P_d V |\chi_k\rangle$$

and by inserting this in (2.) we find

$$(E_k - \mathcal{H}_0 - \epsilon Q_d V Q_d) |\chi_k\rangle = \epsilon^2 Q_d V P_d (E_k - \mathcal{H}_0 - \epsilon P_d V P_d)^{-1} P_d V |\chi_k\rangle.$$

If we keep only the first order in  $\epsilon$  in this equation we have:

$$[(E_k - E_d) - \epsilon Q_d V Q_d] |\chi_k\rangle = 0$$

which is an equation defined on the subspace  $\mathbb{H}_d$  only.

We now call  $U_d = Q_d V Q_d$  the perturbation Hamiltonian in the  $\mathbb{H}_d$  space and  $\Delta_k = (E_k - E_d)\mathbb{1}_d$ , to get:

$$(\Delta_k - \epsilon U_d) |\chi_k\rangle = 0$$

Often it is possible to just diagonalize  $U_d$  (if the degenerate subspace is small enough, for example for a simple double degeneracy) and notice that of course  $\Delta_k$  is already diagonal. Otherwise one can apply perturbation theory to this subspace. Then we will have found some (exact or approximate) eigenstates of  $U_d$ ,  $|k_i^{(0)}\rangle$ , s.t.  $U_d |k_i^{(0)}\rangle = u_i |k_i^{(0)}\rangle$  and  $\mathcal{H}_0 |k_i^{(0)}\rangle = E_d |k_i^{(0)}\rangle$ ,  $\forall i$ . Thus, this step sets what unperturbed eigenstates we should choose in the degenerate subspace, hence solving the first issue of degenerate perturbation theory.

We now want to look at terms  $\propto \epsilon^2$  in

$$(E_k - \mathcal{H}_0 - \epsilon U_d) |\chi_k\rangle = \epsilon^2 Q_d V P_d (E_k - \mathcal{H}_0 - \epsilon P_d V P_d)^{-1} P_d V |\chi_k\rangle$$

where we neglected terms higher than second order. Rearranging the terms, we have:

$$E_k |\chi_k\rangle = [\mathcal{H}_0 + \epsilon U_d + \epsilon^2 Q_d V P_d (E_k - \mathcal{H}_0)^{-1} P_d V] |\chi_k\rangle \quad \rightarrow \quad (\tilde{\mathcal{H}}_0 + \tilde{V}) |\chi_k\rangle = E_k |\chi_k\rangle$$

with

$$\tilde{\mathcal{H}}_0 = \mathcal{H}_0 + \epsilon U_d \quad \tilde{V} = \epsilon Q_d V P_d (E_k - \mathcal{H}_0)^{-1} P_d V Q_d$$

If there are no degeneracies left in  $\tilde{\mathcal{H}}_0$ , we can solve this problem by TIPT and find  $|\chi_k^{(n)}\rangle$ .

For example, to first order, we have

$$|\chi_{k,i}^{(1)}\rangle = \sum_{j \neq i} \frac{\langle k_j^{(0)} | \tilde{V} | k_i^{(0)} \rangle}{\epsilon(u_i - u_j)} |k_j^{(0)}\rangle$$

and using the explicit form of the matrix element  $\tilde{V}_{ij} = \langle k_j^{(0)} | \tilde{V} | k_i^{(0)} \rangle$ ,

$$\tilde{V}_{ij} = \langle k_j^{(0)} | \epsilon^2 V P_d (E_d^0 - \mathcal{H}_0)^{-1} P_d V | k_i^{(0)} \rangle = \epsilon^2 \sum_{h \notin \mathcal{H}_d} \frac{\langle k_j^{(0)} | V | h \rangle \langle h | V | k_i^{(0)} \rangle}{E_d^{(0)} - E_h^{(0)}}$$

we obtain:

$$|\chi_{k,i}^{(1)}\rangle = \epsilon \sum_{j \neq i} \frac{\langle k_j^{(0)} | V | h \rangle \langle h | V | k_i^{(0)} \rangle}{(u_i - u_j) E_d^{(0)} - E_h^{(0)}} |k_j^{(0)}\rangle$$

Finally, we need to add  $|\chi\rangle$  and  $|\psi\rangle$  to find the total vector:

$$\begin{aligned} |\varphi_k^{(1)}\rangle &= \sum_{h \notin \mathcal{H}_d} \left( \frac{\langle h|V|k_i\rangle}{E_d^0 - E_h^{(0)}} |h\rangle + \epsilon \sum_{j \neq i} \frac{\langle k_j^{(0)}|V|h\rangle}{(u_i - u_j)} |k_j^{(0)}\rangle \frac{\langle h|V|k_i^{(0)}\rangle}{E_d^{(0)} - E_h^{(0)}} \right) \\ |\varphi_k^{(1)}\rangle &= \sum_{h \notin \mathcal{H}_d} \frac{\langle h|V|k_i\rangle}{E_d^0 - E_h^{(0)}} \left( |h\rangle + \epsilon \sum_{j \neq i} \frac{\langle k_j^{(0)}|V|h\rangle}{(u_i - u_j)} |k_j^{(0)}\rangle \right) \end{aligned}$$

### Example: Degenerate TLS

Consider the Hamiltonian  $\mathcal{H} = \omega\sigma_z + \epsilon\Omega\sigma_x$ . We already solved this Hamiltonian, both directly and with TIPT. Now consider the case  $\omega \approx 0$  and a slightly modified Hamiltonian:

$$\mathcal{H} = (\omega_0 + \omega)|0\rangle\langle 0| + (\omega_0 - \omega)|1\rangle\langle 1| + \epsilon\Omega\sigma_x = \omega_0\mathbb{1} + \omega\sigma_z + \epsilon\Omega\sigma_x.$$

We could solve exactly the system for  $\omega = 0$ , simply finding  $E_{0,1} = \omega_0 \pm \epsilon\Omega$  and  $|\varphi\rangle_{0,1} = |\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ . We can also apply TIPT.

However the two eigenstates  $|0\rangle, |1\rangle$  are (quasi-)degenerate thus we need to apply degenerate perturbation theory. In particular, any basis arising from a rotation of these two basis states could be a priori a good basis, so we need first to obtain the *correct* zeroth order eigenvectors. In this very simple case we have  $\mathbb{H}_d = \mathbb{H}$  (the total Hilbert space) and  $\mathbb{H}_{\bar{d}} = 0$ , or in other words,  $Q_d = \mathbb{1}$ ,  $P_d = 0$ . We first need to define an equation in the degenerate subspace only:

$$(\Delta_k - \epsilon U_d) |\chi_k\rangle = 0$$

where  $U_d = Q_d V Q_d$ . Here we have:  $U_d = V = \Omega\sigma_x$ . Thus we obtain the *correct* zeroth order eigenvectors from diagonalizing this Hamiltonian. Not surprisingly, they are:

$$|\varphi_{0,1}^{(0)}\rangle = |\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle).$$

with eigenvalues:  $E_{0,1} = \omega_0 \pm \epsilon\Omega$ . We can now consider higher orders, from the equation:

$$(\tilde{\mathcal{H}}_0 + \tilde{V}) |\chi_k\rangle = E_k |\chi_k\rangle$$

with  $\tilde{\mathcal{H}}_0 = \omega_0\mathbb{1} + \epsilon\Omega\sigma_x$  and  $\tilde{V} = 0$ . Thus in this case, there are no higher orders and we solved the problem.

### Example: Spin-1 system

We consider a spin-1 system (that is, a spin system with  $S=1$  defined in a 3-dimensional Hilbert space). The matrix representation for the angular momentum operators  $S_x$  and  $S_z$  in this Hilbert space are:

$$S_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

The Hamiltonian of the system is  $\mathcal{H} = \mathcal{H}_0 + \epsilon V$  with

$$\mathcal{H}_0 = \Delta S_z^2; \quad V = S_x + S_z$$

Given that

$$S_z^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The matrix representation of the total Hamiltonian is :

$$\mathcal{H} = \begin{pmatrix} \Delta + \epsilon & \frac{\epsilon}{\sqrt{2}} & 0 \\ \frac{\epsilon}{\sqrt{2}} & 0 & \frac{\epsilon}{\sqrt{2}} \\ 0 & \frac{\epsilon}{\sqrt{2}} & \Delta - \epsilon \end{pmatrix}$$

Possible eigenstates of the unperturbed Hamiltonian are  $|+1\rangle, |0\rangle, |-1\rangle$ :

$$|+1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |-1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

with energies  $+\Delta, 0, +\Delta$  respectively. However, any combination of  $|+1\rangle$  and  $|-1\rangle$  is a valid eigenstate, for example we could have chosen:

$$|+1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad |-1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

This is the case because the two eigenstates are degenerate. So how do we choose which are the correct eigenstates to zeroth order<sup>37</sup>? We need to first consider the total Hamiltonian in the degenerate subspace.

The degenerate subspace is the subspace of the total Hilbert space  $H$  spanned by the basis  $|+1\rangle, |-1\rangle$ ; we can call this subspace  $H_Q$ . We can obtain the Hamiltonian in this subspace by using the projector operator  $Q$ :  $\mathcal{H}_Q = Q\mathcal{H}Q$ , with  $Q = |+1\rangle\langle+1| + |-1\rangle\langle-1| = S_z^2$ . Then:

$$\mathcal{H}_Q = Q(\Delta S_z^2 + \epsilon(S_z + S_x))Q = \Delta S_z^2 + \epsilon S_z$$

(Notice this can be obtained by direct matrix multiplication or multiplying the operators). In matrix form:

$$\mathcal{H}_Q = \begin{pmatrix} \Delta + \epsilon & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta - \epsilon \end{pmatrix} \rightarrow \mathcal{H}_Q = \begin{pmatrix} \Delta + \epsilon & 0 \\ 0 & \Delta - \epsilon \end{pmatrix}$$

where in the last line I represented the matrix in the 2-dimensional subspace  $H_Q$ . We can now easily see that the correct eigenvectors for the unperturbed Hamiltonian were the original  $|+1\rangle$  and  $|-1\rangle$  after all. From the Hamiltonian in the  $H_Q$  subspace we can also calculate the first order correction to the energy for the states in the degenerate subspace. These are just  $E_{+1}^{(1)} - E_{+1}^{(0)} = +\epsilon$  and  $E_{-1}^{(1)} - E_{-1}^{(0)} = -\epsilon$ .

Now we want to calculate the first order correction to the eigenstates  $|\pm 1\rangle$ . This will have two contributions:  $|\psi\rangle_{\pm 1}^{(1)} = Q|\psi\rangle_{\pm 1}^{(1)} + P|\psi\rangle_{\pm 1}^{(1)}$  where  $P = \mathbb{1} - Q = |0\rangle\langle 0|$  is the complementary projector to  $Q$ . We first calculate the first term in the following way. We redefine an unperturbed Hamiltonian in the subspace  $H_Q$ :

$$\tilde{\mathcal{H}}_0 = \mathcal{H}_Q = Q\mathcal{H}Q = \Delta S_z^2 + \epsilon S_z$$

and the perturbation in the same subspace is (following Sakurai):

$$\tilde{V} = V_Q = \epsilon Q(VP(\Delta - \mathcal{H}_0)^{-1}PV)Q = \epsilon Q((S_z + S_x)|0\rangle\langle 0|(\Delta|0\rangle\langle 0|)^{-1}|0\rangle\langle 0|(S_z + S_x))Q = \frac{\epsilon}{\Delta}QS_xPS_xQ$$

In matrix form:

$$V_Q = \frac{\epsilon}{2\Delta} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \rightarrow V_Q = \frac{\epsilon}{2\Delta} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{\epsilon}{2\Delta}(\mathbb{1} + \sigma_x)$$

Now the perturbed eigenstates can be calculated as:

$$Q|\psi\rangle_k^{(1)} = |k\rangle + \epsilon \sum_{h \in H_Q=k} \frac{\langle h|V_Q|k\rangle}{E_k^{(1)} - E_h^{(1)}} |h\rangle$$

<sup>37</sup> Here by correct eigenstates I means the eigenstates to which the eigenstates of the total Hamiltonian will tend to when  $\epsilon \rightarrow 0$

In our case:

$$Q|\psi\rangle_{+1}^{(1)} = | +1 \rangle + \epsilon \frac{\langle -1 | V_Q | +1 \rangle}{E_{+1}^{(1)} - E_{-1}^{(1)}} | -1 \rangle = | +1 \rangle + \epsilon \frac{\epsilon}{2\Delta} \frac{\langle -1 | (\mathbb{1} + \sigma_x) | +1 \rangle}{2\epsilon} | -1 \rangle = | +1 \rangle + \frac{\epsilon}{4\Delta} | -1 \rangle,$$

$$Q\psi_{-1}^{(1)} = | -1 \rangle - \frac{\epsilon}{4\Delta} | 1 \rangle$$

In order to calculate  $P\psi_{\pm 1}^{(1)}$  we can just use the usual formula for non-degenerate perturbation theory, but summing only over the states outside  $H_Q$ . Here there's only one of them  $|0\rangle$ , so :

$$P|\psi\rangle_{\pm 1}^{(1)} = \epsilon \frac{\langle 0 | V | \pm 1 \rangle}{E_{\pm 1} - E_0} | 0 \rangle = \frac{\epsilon}{\sqrt{2}\Delta} | 0 \rangle$$

Finally, the eigenstates to first order are:

$$|\psi_{+1}\rangle^{(1)} = | +1 \rangle + \frac{\epsilon}{4\Delta} | -1 \rangle + \frac{\epsilon}{\sqrt{2}\Delta} | 0 \rangle$$

and

$$|\psi_{-1}\rangle^{(1)} = | -1 \rangle - \frac{\epsilon}{4\Delta} | 1 \rangle + \frac{\epsilon}{\sqrt{2}\Delta} | 0 \rangle$$

The energy shift to second order is calculated from  $\Delta_{\pm}^{(2)} = \sum_{h \notin H_Q} \frac{|\langle h | V | \pm 1 \rangle|^2}{\Delta - E_h^{(0)}}$ :

$$\Delta_{+1}^{(2)} = \frac{|\langle 0 | V | +1 \rangle|^2}{\Delta} = \frac{\epsilon^2}{2\Delta}$$

and

$$\Delta_{-1}^{(2)} = \frac{|\langle 0 | V | -1 \rangle|^2}{\Delta} = \frac{\epsilon^2}{2\Delta}$$

To calculate the perturbation expansion for  $|0\rangle$  and its energy, we use non-degenerate perturbation theory, to find:

$$\Delta_{+1}^{(1)} = \langle 0 | V | 0 \rangle = 0$$

$$|\psi_{+1}\rangle^{(1)} = | 0 \rangle + \epsilon \left( \frac{\langle +1 | V | 0 \rangle}{-\Delta} | +1 \rangle + \frac{\langle -1 | V | 0 \rangle}{-\Delta} | -1 \rangle \right) = -\frac{\epsilon}{\Delta} \frac{| +1 \rangle + | -1 \rangle}{\sqrt{2}}$$

and  $\Delta_0^{(2)} = -\frac{\epsilon^2}{\Delta}$ .

### 11.1.3 The Stark effect

We analyze the interaction of a hydrogen atom with a (classical) electric field, treated as a perturbation<sup>38</sup>. Depending on the hydrogen's state, we will need to use TIPT or degenerate TIPT, to find either a quadratic or linear (in the field) shift of the energy. The shift in energy is usually called Stark shift or Stark effect and it is the electric analogue of the Zeeman effect, where the energy level is split into several components due to the presence of a magnetic field. Measurements of the Stark effect under high field strengths confirmed the correctness of the quantum theory over the Bohr model.

Suppose that a hydrogen atom is subject to a uniform external electric field, of magnitude  $|\mathbf{E}|$ , directed along the  $z$ -axis. The Hamiltonian of the system can be split into two parts. Namely, the unperturbed Hamiltonian,

$$H_0 = \frac{p^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0 r},$$

---

<sup>38</sup> This section follows [Prof. Fitzpatrick](#) online lectures

and the perturbing Hamiltonian

$$H_1 = e |\mathbf{E}| z.$$

Note that the electron spin is irrelevant to this problem (since the spin operators all commute with  $H_1$ ), so we can ignore the spin degrees of freedom of the system. Hence, the energy eigenstates of the unperturbed Hamiltonian are characterized by three quantum numbers—the radial quantum number  $n$ , and the two angular quantum numbers  $l$  and  $m$ . Let us denote these states as the  $|nlm\rangle$ , and let their corresponding energy eigenvalues be the  $E_{nlm}$ . We use TIPT to calculate the energy shift to first and second order.

### A. The quadratic Stark effect

We first want to study the problem using non-degenerate perturbation theory, thus assuming that the unperturbed states are non-degenerate. According to TIPT, the change in energy of the eigenstate characterized by the quantum numbers  $n, l, m$  in the presence of a small electric field is given by

$$\Delta E_{nlm} = e |\mathbf{E}| \langle n, l, m | z | n, l, m \rangle + e^2 |\mathbf{E}|^2 \sum_{n', l', m' = n, l, m} \frac{|\langle n, l, m | z | n', l', m' \rangle|^2}{E_{nlm} - E_{n'l'm'}}.$$

This energy-shift is known as the Stark effect. The sum on the right-hand side of the above equation seems very complicated. However, it turns out that most of the terms in this sum are zero. This follows because the matrix elements  $\langle n, l, m | z | n', l', m' \rangle$  are zero for virtually all choices of the two sets of quantum number  $n, l, m$  and  $n', l', m'$ . Let us try to find a set of rules which determine when these matrix elements are non-zero. These rules are usually referred to as the selection rules for the problem in hand.

Now, since  $[L_z, z] = 0$ , it follows that

$$\langle n, l, m | [L_z, z] | n', l', m' \rangle = \langle n, l, m | L_z z - z L_z | n', l', m' \rangle = \hbar (m - m') \langle n, l, m | z | n', l', m' \rangle = 0.$$

Hence, one of the selection rules is that the matrix element  $\langle n, l, m | z | n', l', m' \rangle$  is zero unless

$$m' = m.$$

The selection rule for  $l$  can be similarly calculated from properties of the total angular momentum  $L^2$  and its commutator with  $z$ . We obtain that the matrix element is zero unless

$$l' = l \pm 1.$$

Application of these selection rules to the perturbation equation shows that the linear (first order) term is zero, while the second order term yields

$$\Delta E_{nlm} = e^2 |\mathbf{E}|^2 \sum_{n', l' = l \pm 1} \frac{|\langle n, l, m | z | n', l', m \rangle|^2}{E_{nlm} - E_{n'l'm}}.$$

Only those terms which vary quadratically with the field-strength have survived. Hence, this type of energy-shift of an atomic state in the presence of a small electric field is known as the **quadratic Stark effect**.

Now, the electric polarizability of an atom is defined in terms of the energy-shift of the atomic state as follows:

$$\Delta E = -\frac{1}{2} \alpha |\mathbf{E}|^2.$$

Hence, we can write

$$\alpha_{nlm} = 2 e^2 \sum_{n', l' = l \pm 1} \frac{|\langle n, l, m | z | n', l', m \rangle|^2}{E_{n'l'm} - E_{nlm}}.$$

Although written for a general state, the equations above assume there is no degeneracy of the unperturbed eigenvalues. However, the unperturbed eigenstates of a hydrogen atom have energies which only depend on the radial quantum number  $n$ , thus they have high (and increasing with  $n$ ) order of degeneracy. We can then only apply the above results to the  $n = 1$  eigenstate (since for  $n \geq 1$  there will be coupling to degenerate eigenstates with the same

value of  $n$  but different values of  $l$ ). Thus, according to non-degenerate perturbation theory, the polarizability of the ground-state (i.e.,  $n = 1$ ) of a hydrogen atom is given by

$$\alpha = 2e^2 \sum_{n>1} \frac{|\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2}{E_n - E_1}.$$

Here, we have made use of the fact that  $E_{n10} = E_{n00} = E_n$ .

The sum in the above expression can be evaluated approximately by noting that

$$E_n = -\frac{e^2}{8\pi\epsilon_0 a_0 n^2},$$

where  $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2}$  is the Bohr radius. Hence, we can write

$$E_n - E_1 \geq E_2 - E_1 = \frac{3}{4} \frac{e^2}{8\pi\epsilon_0 a_0},$$

which implies that the polarizability is

$$\alpha < \frac{16}{3} 4\pi\epsilon_0 a_0 \sum_{n>1} |\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2.$$

However, thanks to the selection rules we have,  $\sum_{n>1} |\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2 = \langle 1, 0, 0 | z^2 | 1, 0, 0 \rangle = \frac{1}{3} \langle 1, 0, 0 | r^2 | 1, 0, 0 \rangle$ , where we have made use of the fact that the ground-state of hydrogen is spherically symmetric. Finally, from  $\langle 1, 0, 0 | r^2 | 1, 0, 0 \rangle = 3a_0^2$  we conclude that

$$\alpha < \frac{16}{3} 4\pi\epsilon_0 a_0^3 \simeq 5.3 4\pi\epsilon_0 a_0^3.$$

The exact result (which can be obtained by solving Schrödinger's equation in parabolic coordinates) is

$$\alpha = \frac{9}{2} 4\pi\epsilon_0 a_0^3 = 4.5 4\pi\epsilon_0 a_0^3.$$

## B. The linear Stark effect

We now examine the effect of an electric field on the excited energy levels  $n \geq 1$  of a hydrogen atom. For instance, consider the  $n = 2$  states. There is a single  $l = 0$  state, usually referred to as 2s, and three  $l = 1$  states (with  $m = -1, 0, 1$ ), usually referred to as 2p. All of these states possess the same energy,  $E_2 = -e^2/(32\pi\epsilon_0 a_0)$ . Because of the degeneracy, the treatment above is no longer valid and in order to apply perturbation theory, we have to recur to degenerate perturbation theory.

We first need to  $U_d = Q_d V Q_d$ , where  $Q_d$  is the projector obtained from the degenerate 2s and 2p states (that is, the operator that project into the degenerate subspace). This operator is,

$$U_d = e |\mathbf{E}| \begin{pmatrix} 0 & \langle 2, 0, 0 | z | 2, 1, 0 \rangle & 0 & 0 \\ \langle 2, 1, 0 | z | 2, 0, 0 \rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & \langle 2, 0, 0 | z | 2, 1, 0 \rangle \\ \langle 2, 1, 0 | z | 2, 0, 0 \rangle & 0 \end{pmatrix},$$

where the rows and columns correspond to the  $|2, 0, 0\rangle$ ,  $|2, 1, 0\rangle$ ,  $|2, 1, 1\rangle$  and  $|2, 1, -1\rangle$  states, respectively and in the second step we reduce the operator to the degenerate subspace only. To simplify the matrix we used the selection rules, which tell us that the matrix element of between two hydrogen atom states is zero unless the states possess the same  $n$  quantum number, and  $l$  quantum numbers which differ by unity. It is easily demonstrated, from the exact forms of the 2s and 2p wave-functions, that

$$\langle 2, 0, 0 | z | 2, 1, 0 \rangle = \langle 2, 1, 0 | z | 2, 0, 0 \rangle = 3a_0.$$

It can be seen, by inspection, that the eigenvalues of  $U_d$  are  $u_1 = 3 e a_0 |\mathbf{E}|$ ,  $u_2 = -3 e a_0 |\mathbf{E}|$ , with corresponding eigenvectors

$$\begin{aligned} |k_1^{(0)}\rangle &= \frac{|2, 0, 0\rangle + |2, 1, 0\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ |k_2^{(0)}\rangle &= \frac{|2, 0, 0\rangle - |2, 1, 0\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

In the absence of an electric field, all of these states possess the same energy,  $E_2$ . The first-order energy shifts induced by an electric field are given by

$$\begin{aligned} \Delta E_1 &= +3 e a_0 |\mathbf{E}|, \\ \Delta E_2 &= -3 e a_0 |\mathbf{E}|, \end{aligned}$$

Thus, the energies of states 1 and 2 are shifted upwards and downwards, respectively, by an amount  $3 e a_0 |\mathbf{E}|$  in the presence of an electric field. States 1 and 2 are orthogonal linear combinations of the original 2s and 2p(m=0) states. Note that the energy shifts are linear in the electric field-strength, so this is a much larger effect than the quadratic effect described in the previous section.

The energies of states 2p(m=1) and 2p(m=-1) (which are outside the degenerate subspace) are not affected to first-order (as we already saw above for the non-degenerate case). Of course, to second-order the energies of these states are shifted by an amount which depends on the square of the electric field-strength, the quadratic shift found previously. Note that the linear Stark effect depends crucially on the degeneracy of the 2s and 2p states. This degeneracy is a special property of a pure Coulomb potential, and, therefore, only applies to a hydrogen atom. Thus, alkali metal atoms do not exhibit the linear Stark effect.

## 11.2 Time-dependent perturbation theory

### 11.2.1 Review of interaction picture

When first studying the time evolution of QM systems, one approach was to separate the Hamiltonian much in the same way we did above for TIPT. We wrote (see Section 5.2):

$$\mathcal{H} = \mathcal{H}_0 + V(t)$$

where  $\mathcal{H}_0$  is a "solvable" Hamiltonian of which we already know the eigen-decomposition,

$$\mathcal{H}_0|k\rangle = E_k^0|k\rangle,$$

(so that it is easy to calculate e.g.  $U_0 = e^{-i\mathcal{H}_0 t}$ ) and  $V(t)$  is a perturbation that drives an interesting (although unknown) dynamics. Here we even allow for the possibility that  $V$  is time-dependent. For any state  $|\psi\rangle = \sum_k c_k(0)|k\rangle$  the evolution can be written as  $|\psi\rangle = \sum_k c_k(t)e^{-iE_k^0 t}|k\rangle$ . This correspond to explicitly writing down the evolution due to the known Hamiltonian (if  $\mathcal{H} = \mathcal{H}_0$  then we would have  $c_k(t) = c_k(0)$  and the evolution would be given by only the phase factors). In other words, if we want to compare the state evolution with the initial eigenstates, by calculating the overlap  $|\langle k|\psi(t)\rangle|^2$ , we would be really interested only in the dynamics driven by  $V$  since  $|\langle k|\psi(t)\rangle|^2 = |c_k(t)|^2$  (while  $E_k^0$  do not play a role).

We define states in the interaction picture by

$$|\psi\rangle_I = U_0(t)^\dagger |\psi\rangle = e^{i\mathcal{H}_0 t} |\psi\rangle$$

Similarly we define the corresponding interaction picture operators as:

$$A_I(t) = U_0^\dagger A U_0 \rightarrow V_I(t) = U_0^\dagger V U_0$$

We can now derive the differential equation governing the evolution of the state in the interaction picture, starting from Schrödinger equation.

$$i \frac{\partial |\psi\rangle_I}{\partial t} = i \frac{\partial (U_0^\dagger |\psi\rangle)}{\partial t} = i \left( \frac{\partial U_0^\dagger}{\partial t} |\psi\rangle + U_0^\dagger \frac{\partial |\psi\rangle}{\partial t} \right)$$

Inserting  $\partial_t U_0 = i\mathcal{H}_0 U_0$  and  $i\partial_t |\psi\rangle = \mathcal{H}_0 |\psi\rangle$ , we obtain

$$i \frac{\partial |\psi\rangle_I}{\partial t} = U_0^\dagger \mathcal{H}_0 |\psi\rangle - U_0^\dagger (\mathcal{H}_0 + V) |\psi\rangle = U_0^\dagger V |\psi\rangle.$$

Inserting the identity  $\mathbb{1} = U_0 U_0^\dagger$ , we obtain  $= U_0^\dagger V U_0 U_0^\dagger |\psi\rangle = V_I |\psi\rangle_I$ :

$$\boxed{i \frac{\partial |\psi\rangle_I}{\partial t} = V_I(t) |\psi\rangle_I}$$

This is a Schrödinger-like equation for the vector in the interaction picture, evolving under the action of the operator  $V_I(t)$  only.

### 11.2.2 Dyson series

Besides expressing the Schrödinger equation in the interaction picture, we can also write the equation for the propagator that describes the evolution of the state:

$$\frac{dU_I}{dt} = -iV_I U_I, \quad |\psi(t)\rangle_I = U_I(t) |\psi_I(0)\rangle$$

Since  $V_I(t)$  is time-dependent, we can only write formal solutions for  $U_I$ . One expression is given by the Dyson series. The differential equation is equivalent to the integral equation

$$U_I(t) = \mathbb{1} - i \int_0^t V_I(t') U_I(t') dt'$$

By iterating, we can find a formal solution to this equation :

$$\begin{aligned} U_I(t) = & \mathbb{1} - i \int_0^t dt' V_I(t') + (-i)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') + \dots \\ & + (-i)^n \int_0^t dt' \dots \int_0^{t^{(n-1)}} dt^{(n)} V_I(t') \dots V_I(t^{(n)}) + \dots \end{aligned}$$

This is the Dyson series.

### 11.2.3 Fermi's Golden Rule

The problem that we try to solve via TDPT is to calculate the **transition probability** from an initial state to a final state. Consider an initial state  $|i\rangle$  which is an eigenstate of  $\mathcal{H}_0$  ( $\mathcal{H}_0 |i\rangle = E_i |i\rangle$ ). Then in the interaction picture we have the evolution

$$|i(t)\rangle_I = U_I(t) |i\rangle = \sum_k c_k(t) |k\rangle, \quad \text{with} \quad c_k(t) = \langle k | U_I(t) | i \rangle$$

We can insert the perturbation expansion for  $U_I(t)$  to obtain an expansion for  $c_k(t)$ :

$$c_k(t) = \langle k | \mathbb{1} - i \int_0^t V_I(t') U_I(t') dt' | i \rangle = \langle k | \left[ \mathbb{1} - i \int_0^t dt' V_I(t') + (-i)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') + \dots \right] | i \rangle$$

In the expansion we will obtain terms such as  $\langle k | V_I(t) | i \rangle$  that we can simplify since:

$$\langle k | V_I(t) | i \rangle = \langle k | (U_0^\dagger V(t) U_0) | i \rangle = \langle U_0 k | V(t) | U_0 i \rangle = \langle k | e^{i\omega_k t} V(t) e^{-i\omega_i t} | i \rangle = \langle k | V | i \rangle e^{i\omega_{ki} t} = V_{ki}(t) e^{i\omega_{ki} t}$$

where we defined  $\omega_j = E_j/\hbar$  and  $\omega_{ki} = \omega_k - \omega_i$ . Using these relationships and the series expansion we obtain:

$$\begin{aligned} c_k^{(0)}(t) &= \langle k | \mathbb{1} | i \rangle = \delta_{ki} \\ c_k^{(1)}(t) &= -i \int_0^t \langle k | V_I(t') | i \rangle dt' = -i \int_0^t V_{ki}(t') e^{i\omega_{ki} t'} dt' \\ c_k^{(2)}(t) &= - \int_0^t dt' \int_0^{t'} dt'' V_{kh}(t') V_{hi}(t'') e^{i\omega_{kh} t'} e^{i\omega_{hi} t''} \end{aligned}$$

From this expansion we can calculate the transition probability as  $P(i \rightarrow k) = |c_k(t)|^2$ .

We first consider the case where the perturbation  $V$  is time-independent and it is turned on at the time  $t = 0$ . Then we have

$$c_k^{(1)}(t) = -i V_{ki} \int_0^t e^{i\omega_{ki} t'} dt' = \frac{V_{ki}}{\omega_{ki}} (1 - e^{i\omega_{ki} t}) = -2i \frac{V_{ki}}{\omega_{ki}} e^{i\omega_{ki} t/2} \sin\left(\frac{\omega_{ki} t}{2}\right)$$

Then to first order perturbation, the transition probability is

$$P(i \rightarrow k) = \frac{4|V_{ki}|^2}{\omega_{ki}^2} \sin^2\left(\frac{\omega_{ki} t}{2}\right)$$

We can plot this transition probability as a function of the energy separation  $\omega_{ki}$  between the two states. We would expect that if the separation in energy is smaller, it will be easier to make the transition. This is indeed the case, since  $P$  has the shape of a sinc function square.

Notice that the peak height is proportional to  $t^2$ , while the zeros appear at  $2k\pi/t$ , that is, the peak width is proportional to  $1/t$  (the other peaks are quite small). This means that the probability is significantly different than

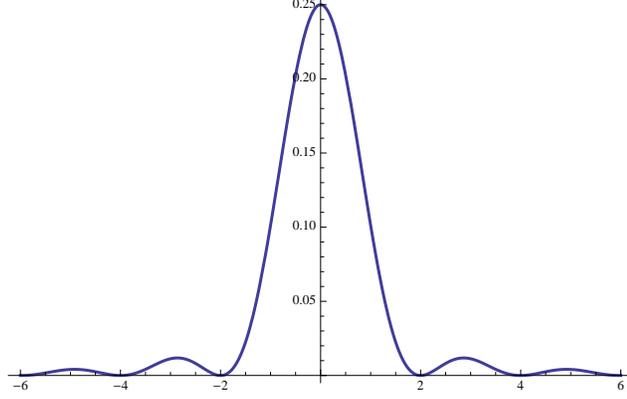


Fig. 19: Transition probability

zero only for  $\omega_{ki}t \leq 2\pi$ . In terms of energy, we have that  $\Delta t \Delta E \sim \hbar$  (where we defined  $\Delta t$  as the duration of the interaction), or in other words, we can have a change of energy in the system only at short times, while at long times we require quasi-conservation of energy. Consider the limit of the sinc function:

$$\lim_{t \rightarrow \infty} \frac{\sin(\omega t/2)}{\omega} = \pi \delta(\omega)$$

Then, from  $f(x)\delta(x) = f(0)$  and  $\text{sinc}(0) = 1$ , we obtain

$$\lim_{t \rightarrow \infty} \left( \frac{\sin(\omega t/2)}{\omega} \right)^2 = \frac{\sin(\omega t/2)}{\omega} \lim_{t \rightarrow \infty} \frac{\sin(\omega t/2)}{\omega} = \frac{\sin(\omega t/2)}{\omega} \pi \delta(\omega) = \left( \frac{\sin(\omega t/2)}{\omega t/2} \right) \frac{t}{2} \pi \delta(\omega) = \frac{\pi t}{2} \delta(\omega)$$

We have then found the transition probability at long time:

$$P(i \rightarrow k) \xrightarrow{t \rightarrow \infty} \frac{\pi t}{2} \delta(\omega) 4|V_{ki}|^2,$$

which confirms the fact that in the long-time limit we need to enforce energy conservation. A better defined quantity is the *rate* of transition:

$$W(i \rightarrow k) = 2\pi |V_{ki}|^2 \delta(\omega).$$

Notice that for  $\omega_{ki} = 0$ , from  $c_k^{(1)}(t) = -iV_{ki} \int_0^t e^{i\omega_{ki}t'} dt'$  we obtain  $c_k^{(1)}(t) = -iV_{ki}t$  and thus the probability  $|c_k(t)|^2 = |V_{ki}|^2 t^2$ . There is a quadratic dependence on time for a single final state.

Now we consider a continuum of final states, all with energy  $E_{k_f} \approx E_i$ . Then the probability of a transition to this continuum is given by the sum of the probability for each individual state:  $P_f = \sum_k |c_k|^2 \rightarrow \int dE_k \rho(E_k) |c_k|^2$ , where we defined the density of states  $\rho(E_k)$ , such that  $\rho(E_k)dE_k$  is the number of states with energy between  $E_k$  and  $E_k + dE_k$ . We can then rewrite the probability as

$$P_{i \rightarrow f} = 4 \int dE \rho(E) \sin^2 \left( \frac{(E - E_i)t}{2} \right) \frac{|V_{ki}|^2}{(E - E_i)^2}$$

Using the limit of the sinc function, we find

$$P_{i \rightarrow f} = 4 \int dE \rho(E) \delta(E - E_i) \frac{\pi t}{2} \frac{|V_{ki}|^2}{(E - E_i)^2}$$

Since all the states are in a neighborhood of the energy, we expect  $|V_{ki}|^2 \approx |\bar{V}_{ki}|^2$  over the range of energy of interest. Thus by evaluating the integral (with the delta function) we obtain the transition probability:

$$P_{i \rightarrow f} = 2|V_{ki}|^2 \pi t \rho(E_k)|_{E_k \approx E_i}$$

Similarly, we can calculate the transition rate to a continuum of states. From the expression for a single state,  $W_{i \rightarrow k} = \frac{2\pi}{\hbar} |V_{ki}|^2 \delta(E_k - E_i)$ , we integrate over all final energies,  $W_{i \rightarrow f} = \int W_{i \rightarrow k} \rho(E_k) dE_k$ , where  $f$  is the continuum of states  $k$  such that  $E_k \approx E_i$ . Then we obtain the transition rate:

$$W = \frac{2\pi}{\hbar} |V_{ki}|^2 \rho(E_k) |_{E_k \approx E_i}$$

This is Fermi's Golden Rule.

### Virtual Transitions

If the matrix element of the interaction connecting two given state is zero, we have seen from the expression above that no transition is possible, *to first order*.

However, consider  $c_k^{(2)}(t)$ . This is given by

$$c_k^{(2)}(t) = - \sum_h V_{kh} V_{hi} \int_0^t dt' \int_0^{t'} dt'' e^{i\omega_{kh}t'} e^{i\omega_{hi}t''} = i \sum_h \frac{V_{kh} V_{hi}}{\omega_{ki}} \int_0^t dt' (\underbrace{e^{i\omega_{ki}t'}}_{\text{same as before}} - \underbrace{e^{i\omega_{kh}t'}}_{\approx 0})$$

If  $E_h \neq E_k, E_i$ , the second term oscillates rapidly and goes to zero. Finally we have:

$$W_{i \rightarrow k} = \frac{2\pi}{\hbar} \left| V_{ki} + \sum_h \frac{V_{kh} V_{hi}}{\omega_{ki}} \right|^2 \delta(E_k - E_i)$$

or for a continuum

$$W_{i \rightarrow f} = \frac{2\pi}{\hbar} \left| V_{ki} + \sum_h \frac{V_{kh} V_{hi}}{\omega_{ki}} \right|^2 \rho(E_k) |_{E_k \approx E_i}$$

Notice that even if  $V_{ik} = 0$ , we can still have a transition to  $k$ , via *virtual* transitions to intermediate states, which are connected to the two relevant levels.

## 12. Interaction of Radiation with Matter

### 12.1 Scattering Theory

#### 12.1.1 Cross Section

#### 12.1.2 Thermal Neutron Scattering

### 12.2 Emission and Absorption

#### 12.2.1 Emission

#### 12.2.2 Absorption

#### 12.2.3 Blackbody Radiation

### 12.3 Wigner-Weisskopf Theory

#### 12.3.1 Interaction of an atom with a single mode e.m. field

#### 12.3.2 Interaction with many modes of the e.m. field

### 12.4 Scattering of photons by atoms

#### 12.4.1 Thomson Scattering by Free Electrons

#### 12.4.2 Rayleigh Scattering of X-rays

#### 12.4.3 Visible Light Scattering

#### 12.4.4 Photoelectric Effect

### 12.1 Scattering Theory

We want to describe the interaction of radiation with matter as a scattering process. Specifically, we are interested in calculating the rate of scattering (and then the cross section), which is nothing else than the transition rate from an initial state (initial state of the matter + incoming particle) and a final state (final state of the target + outgoing radiation)<sup>39</sup>.

This is a problem that can be solved by TDPT. Instead of considering a constant perturbation as done to derive Fermi's Golden rule, we analyze the case of a scattering potential, in its most general form. We describe a scattering

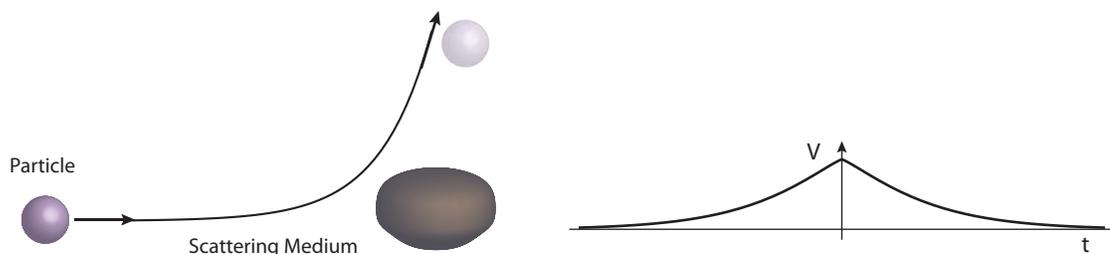


Fig. 20: Model for scattering: Left, particle trajectory, right time dependency of the potential.

event as a particle coming close to a target or a medium, interacting with it and then being deflected away. Thus, as a function of time, the interaction Hamiltonian  $V$  varies as in the figure 20.

<sup>39</sup> A very good resource for scattering theory is Chen, S.H.; Kotlarchyk, M., *Interactions of Photons and Neutrons with Matter*, (2007), which we follow closely in this chapter.

We want to calculate the probability of scattering from an initial state to a final state:

$$P_{scatt} = |\langle f | U_I(t) | i \rangle|^2 = |\langle f | (\mathbb{1} - i \int_{-\infty}^{\infty} V_I(t') dt' + \dots) | i \rangle|^2$$

Notice that we consider negative times as well. This corresponds to the so-called *adiabatic switching*, since the interaction is assumed to be turned on slowly from the beginning of time and to go down to zero again for long times.

### A. Scattering and Transition matrices

In scattering problems, the propagator  $U_I$  is usually called the **scattering matrix**  $S$ . To simplify the calculation, we can assume again that  $V$  is actually time-independent. Then from the first order TDPT we obtain:

$$\langle f | S^{(1)} | i \rangle = -i V_{fi} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} dt = -2\pi i \delta(\omega_f - \omega_i) V_{fi}$$

Now consider the second order contribution:

$$\langle f | S^{(2)} | i \rangle = -\langle f | \left( \sum_m V | m \rangle \langle m | V \right) | i \rangle \int_{-\infty}^{\infty} dt_1 e^{i\omega_{fm}t_1} \int_{-\infty}^{t_1} dt_2 e^{i\omega_{mi}t_2}$$

Notice that the last integral is not well defined for  $t \rightarrow -\infty$ . To solve it, we rewrite it as

$$\lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{t_1} dt_2 e^{i(\omega_{mi} - i\epsilon)t_2} = \lim_{\epsilon \rightarrow 0^+} -i \frac{e^{i\omega_{mi}t_1 + \epsilon t_1}}{\omega_{mi} - i\epsilon} \Big|_{-\infty}^{t_1}$$

Now when taking the limit  $t \rightarrow -\infty$  the exponential term  $e^{\epsilon t} \rightarrow 0$  (thus getting rid of the oscillations). Then we are left with only

$$\int_{-\infty}^{t_1} dt_2 e^{i\omega_{mi}t_2} = \lim_{\epsilon \rightarrow 0^+} -i \frac{e^{i(\omega_{mi} - i\epsilon)t_1}}{\omega_{mi} - i\epsilon}$$

and we obtain (setting now  $\epsilon = 0$ )

$$\langle f | S^{(2)} | i \rangle = i \sum_m V_{fm} V_{mi} \int_{-\infty}^{\infty} dt_1 \frac{e^{i(\omega_{fi} - i\epsilon)t_1}}{\omega_{mi} - i\epsilon} = -2\pi i \delta(\omega_f - \omega_i) \sum_m \frac{\langle f | V | m \rangle \langle m | V | i \rangle}{\omega_i - \omega_m}$$

Looking at the first and second order of the scattering matrix, we start seeing a pattern emerge. We can thus rewrite :

$$\boxed{\langle f | S | i \rangle = -2\pi i \delta(\omega_f - \omega_i) \langle f | T | i \rangle}$$

where  $T$  is called the **transition matrix**. Its expansion is given by:

$$\langle f | T | i \rangle = \langle f | V | i \rangle + \sum_m \frac{\langle f | V | m \rangle \langle m | V | i \rangle}{\omega_i - \omega_m} + \sum_{m,n} \frac{V_{fm} V_{mn} V_{ni}}{(\omega_i - \omega_m)(\omega_i - \omega_n)} + \dots$$

### B. Scattering Probability

We can now turn to calculate the scattering probability:  $P_S = |\langle f | S | i \rangle|^2$ . In order to obtain the total scattering probability, we will need to consider all possible final states. We found:

$$P_s = 4\pi^2 |\langle f | T | i \rangle|^2 \delta^2(\omega_f - \omega_i)$$

We calculate the square of the Dirac function from its definition based on the limit of the integral:

$$\delta^2(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{i\omega t} \delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \delta(\omega) = \lim_{t \rightarrow \infty} \frac{t}{\pi} \delta(\omega)$$

Then although the probability is not so well defined, since it contains a limit:

$$P_s = \lim_{t \rightarrow \infty} 4\pi t |\langle f | T | i \rangle|^2 \delta(\omega_f - \omega_i)$$

the *rate* of scattering is well defined, since it is  $W_S = P_S/(2t)$ :

$$W_S = 2\pi |\langle f | T | i \rangle|^2 \delta(\omega_f - \omega_i)$$

This is the rate for one isolated final state. If instead we have a continuum of final states, with density of states  $\rho(\omega_f)$  we need to sum over all possible final states:

$$W_S = 2\pi \int 2\pi |\langle f | T | i \rangle|^2 \delta(\omega_f - \omega_i) \rho(\omega_f) d\omega_f = 2\pi |\langle f | T | i \rangle|^2 \rho(\omega_i)$$

Notice that to first order, this is equivalent to the Fermi Golden rule.

### 12.1.1 Cross Section

We now use the tools developed in TDPT to calculate the scattering cross section. This is defined as the rate of scattering divided by the incoming flux of “particles”:

$$\boxed{\frac{d^2\sigma}{d\Omega dE} \propto \frac{W_S(\Omega, E)}{\Phi_{inc}}}$$

We consider a particle + medium system, where the particle is some radiation represented by a plane wave of momentum  $\vec{k}$ . In general, we will have to define also other degrees of freedom denoted by the index  $\lambda$ , e.g for photons we will have to define the polarization while for particles (e.g.e neutrons) the spin.

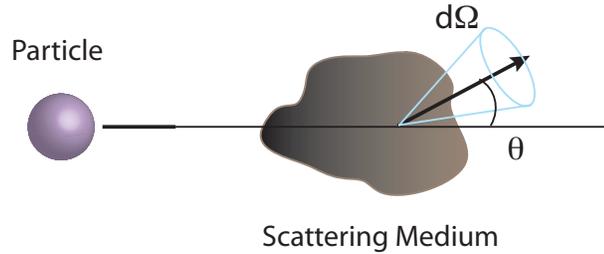
The unperturbed Hamiltonian is  $\mathcal{H}_0 = \mathcal{H}_R + \mathcal{H}_M$  (radiation and medium). We assume that for  $t \rightarrow \pm\infty$  the radiation and matter systems are independent, with (eigen)states:

$$|i\rangle = |k_i, m_i\rangle, \quad |f\rangle = |k_f, m_f\rangle$$

with energies:

$$\mathcal{H}_R |k_i\rangle = \hbar\omega_i |k_i\rangle, \quad \mathcal{H}_R |k_f\rangle = \hbar\omega_f |k_f\rangle, \quad \mathcal{H}_M |m_i\rangle = \epsilon_i |k_i\rangle, \quad \mathcal{H}_M |m_f\rangle = \epsilon_f |m_f\rangle$$

and total energies:  $E_i = \hbar\omega_i + \epsilon_i$  and  $E_f = \hbar\omega_f + \epsilon_f$ .



## Scattering Rate

The rate of scattering is given by the expression found earlier:

$$W_{fi} = \frac{2\pi}{\hbar} |\langle f|T|i\rangle|^2 \delta(E_f - E_i)$$

As usual, we want to replace, if possible, the delta-function with the final density of states. However, only the radiation will be left in a continuum of states, while the target will be left in one (of possibly many) definite state. To describe this distinction, we separate the final state into the two subsystems.

We first define the partial projection on radiation states only,  $T_{k_f, k_i} = \langle k_f|T|k_i\rangle$ . By writing the delta function as an integral we have:

$$W_{fi} = \frac{2\pi}{\hbar} \langle m_f|T_{k_f, k_i}|m_i\rangle \langle m_i|T_{k_f, k_i}^\dagger|m_f\rangle \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} e^{i(\omega_f - \omega_i)t} e^{i(\epsilon_f - \epsilon_i)t/\hbar}$$

Now, since  $e^{-i\mathcal{H}_R t/\hbar}|m_i\rangle = e^{-i\epsilon_i t/\hbar}|m_i\rangle$  (and similarly for  $|m_f\rangle$ ) we can rewrite

$$\langle m_f|T_{k_f, k_i}|m_i\rangle e^{i(\epsilon_f - \epsilon_i)t/\hbar} = \langle m_f|e^{i\mathcal{H}_R t/\hbar}T_{k_f, k_i}e^{-i\mathcal{H}_R t/\hbar}|m_i\rangle = \langle m_f|T_{k_f, k_i}(t)|m_i\rangle$$

and obtain a new expression for the rate as a correlation of “transition” events:

$$W_{fi} = \frac{1}{\hbar^2} \int_{-\infty}^{\infty} e^{i(\omega_f - \omega_i)t} \langle m_i|T_{k_f, k_i}^\dagger(0)|m_f\rangle \langle m_f|T_{k_f, k_i}(t)|m_i\rangle$$

## Final density of states

The final density of states describe the available states for the radiation. As we assumed that the radiation is represented by plane waves (and assuming for convenience they are contained in a cavity of edge  $L$ ), the final density of states is

$$\rho(k_f)d^3k_f = \left(\frac{L}{2\pi}\right)^3 k_f^2 dk_f d\Omega$$

We can express this in terms of the energy,  $\rho(k)d^3k = \rho(E)dEd\Omega$ . For example, for photons, which have  $k = E/\hbar c$  we have

$$\rho(E) = 2 \left(\frac{L}{2\pi}\right)^3 \frac{E^2}{\hbar^3 c^3} = 2 \left(\frac{L}{2\pi}\right)^3 \frac{\omega_k^2}{\hbar c^3}$$

where the factor 2 takes into account the possible polarizations. For neutrons (or other particles such that  $E = \frac{\hbar^2 k^2}{2m}$ ):

$$\rho(E) = \left(\frac{L}{2\pi}\right)^3 \frac{k}{\hbar^2} = \left(\frac{L}{2\pi}\right)^3 \frac{\sqrt{2mE}}{\hbar^3}$$

If the material target can be left in more than one final state, we sum over these final states  $f$ . Then the average rate is given by  $\overline{W}_S = \sum_f W_{fi} \rho(E) dEd\Omega$  (assuming that  $W_{fi}$  does not change very much in  $d\Omega$  and  $dE$ ).

## Incoming Flux

The incoming flux is given by the number of scatterer per unit area and unit time,  $\Phi = \frac{\#}{A t}$ . In the cavity considered, we can express the time as  $t = L/v$ , thus the flux is  $\Phi = \frac{v}{L^3}$ . For photons, this is simply  $\Phi = c/L^3$ , while for massive particles (neutrons)  $v = \hbar k/m$ , yielding  $\Phi = \frac{\hbar k}{m L^3}$ .

Average over initial states

If the scatterer is at a finite temperature  $T$  it will be in a mixed state, thus we need to sum over all possible initial states:

$$\rho_i = \frac{e^{-\beta\mathcal{H}_M}}{Z} \quad \rightarrow \quad P_i = \frac{e^{-\epsilon_i/k_b T}}{\sum_i e^{-\epsilon_i/k_b T}}$$

We can finally write the total scattering rate as:

$$\begin{aligned} W_S(i \rightarrow \Omega + d\Omega, E + dE) &= \rho(E) \sum_i P_i \sum_f W_{fi} \\ &= \rho(E) \sum_{f,i} \frac{P_i}{\hbar^2} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \langle m_i | T_{k_f, k_i}^\dagger(0) | m_f \rangle \langle m_f | T_{k_f, k_i}(t) | m_i \rangle dE d\Omega = \frac{\rho(E)}{\hbar^2} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \langle T_{if}^\dagger(0) T_{fi}(t) \rangle \end{aligned}$$

where  $\langle \cdot \rangle$  indicates an ensemble average at the given temperature.

### 12.1.2 Thermal Neutron Scattering

Using the scattering rate above and the incoming flux and density of state expression, we can find the cross section for thermal neutrons. From

$$\rho(E)/\Phi = \left[ \left( \frac{L}{2\pi} \right)^3 \frac{mk_f}{\hbar^2} \right] / \left[ \frac{\hbar k_i}{mL^3} \right] = \frac{(mL^3)^2 k_f}{(2\pi\hbar)^3 k_i}$$

we obtain

$$\frac{d^2\sigma}{d\Omega d\omega} = \hbar \frac{W}{\Phi} = \hbar \frac{\rho(E)}{\Phi} \frac{1}{\hbar^2} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \langle T_{if}^\dagger(0) T_{fi}(t) \rangle = \frac{1}{2\pi} \left( \frac{mL^3}{2\pi\hbar^2} \right)^2 \frac{k_f}{k_i} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \langle T_{if}^\dagger(0) T_{fi}(t) \rangle$$

Now the eigenstates  $|k_{i,f}\rangle$  are plane waves,  $\langle r|k\rangle = \psi_k(r) = e^{ik\cdot r}/L^{3/2}$ . Then, defining  $Q = k_i - k_f$  the transition matrix element is

$$T_{fi}(t) = \langle k_f | T(t) | k_i \rangle = \int_{L^3} d^3r \psi_{k_f}(r)^* T(r, t) \psi_{k_i}(r) = \frac{1}{L^3} \int_{L^3} d^3r e^{iQ\cdot r} T(r, t)$$

and

$$T_{fi}(0)^\dagger = \frac{1}{L^3} \int_{L^3} d^3r e^{-iQ\cdot r} T(r, 0)^\dagger$$

### Fermi Potential

To first order, we can approximate  $T$  by  $V$ , the nuclear potential in the center of mass frame (of the neutron+nucleus). You might recall that the nuclear potential is a very strong ( $V_0 \sim 30\text{MeV}$ ) and narrow ( $r_0 \sim 2\text{fm}$ ) potential. These characteristics seem to preclude a perturbative approach, since the assumption of a weak interaction (compared to the unperturbed system energy) is not satisfied. Still, the fact that the potential is narrow means that the interaction only happens for a very short time. Thus, if we average over time, we expect a weak interaction. More precisely, the scattering interaction only depends on the so-called *scattering length*  $a$ , which is on the order  $a \sim V_0 r_0$ . If we keep  $a$  constant, different combinations of  $V$ ,  $r$  will give the same scattering behavior. We can thus replace the strong nuclear potential with a weaker, pseudo-potential  $\tilde{V}_0$ , provided this has a much longer range  $\tilde{r}_0$ , such that  $a \sim V_0 r_0 = \tilde{V}_0 \tilde{r}_0$ . We can choose  $\tilde{V}_0$ ,  $\tilde{r}_0$  so that the potential is weak (eV) but the range is still short compared to the wavelength of the incoming neutron,  $k\tilde{r}_0 \ll 1$ . Then, it is possible to replace the potential with a simple delta-function at the origin.

$$V(r) = \frac{2\pi\hbar^2}{\mu} a \delta(r)$$

We can also define the **bound scattering length**,  $b = \frac{\mu}{m_n}a \approx \frac{A+1}{A}$ , where  $m_n$  is the neutron's mass and  $A$  the nucleus mass number. Then the potential is

$$V(r) = \frac{2\pi\hbar^2}{m_n}a\delta(r)$$

Note that  $b$  (interaction length or bound scattering length) is a function of the potential strength and range, which depend on the isotope from which the neutron is scattered off.

Then to first order the transition matrix is  $T_{fi} = \frac{2\pi\hbar^2}{m_n}b$ , or more generally, if there are many scatterers, each at a position  $r_x(t)$ , we have:

$$T_{fi}(t) = \frac{2\pi\hbar^2}{m_n} \sum_x b_x e^{iQ \cdot r_x(t)}$$

The scattering cross section becomes

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{1}{2\pi} \frac{k_f}{k_i} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \left\langle \sum_{x,y} b_x b_y e^{-iQ \cdot r_x(0)} e^{iQ \cdot r_y(t)} \right\rangle$$

Notice that since the collisions are spin-dependent, we should average over isotopes and spin states and replace  $b_x b_y$  with  $\overline{b_x b_y}$ .

### Scattering Lengths

Notice that  $b$  does not depend explicitly on position, although the position determines which isotope/spin we should consider. What is  $\overline{b_x b_y}$ ? We have two contributions. For  $x = y$  this is  $\overline{b^2} \delta_{x,y}$ , while for  $x \neq y$ , it is  $\overline{b^2} (1 - \delta_{x,y})$ . We then write  $\overline{b_x b_y} = (\overline{b^2} - \overline{b^2}) \delta_{x,y} + \overline{b^2} = b_i^2 + b_c^2$  which defines the **coherent scattering length**  $b_c = \overline{b}$  and the **incoherent scattering length**  $b_i^2 = \overline{b^2} - \overline{b^2}$ . If there are  $N$  scatterers, we have  $\sum b_x b_y = N(b_i^2 + b_c^2)$ .

### Structure Factors

Using these definition, we arrive at a simplified expression:

$$\frac{d^2\sigma}{d\Omega d\omega} = N \frac{k_f}{k_i} (b_i^2 S_S(Q, \omega) + b_c^2 S(Q, \omega))$$

where we used the **self-dynamic** structure factor

$$S_S(Q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \left\langle \frac{1}{N} \sum_x e^{-iQ \cdot r_x(0)} e^{iQ \cdot r_x(t)} \right\rangle$$

which simplifies to

$$S_S(Q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \left\langle e^{-iQ \cdot r(0)} e^{iQ \cdot r(t)} \right\rangle$$

if all nuclei are equivalent (same isotope), and the **full dynamic** structure factor

$$S(Q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega_{fi}t} \left\langle \frac{1}{N} \sum_{x,y} e^{-iQ \cdot r_x(0)} e^{iQ \cdot r_y(t)} \right\rangle$$

The structure factors depend only on the material properties. Thus they give information about the material when obtained from experiments.

## Intermediate Scattering Function

From the expressions above for the structure factors, it is clear that they can be obtained as the Fourier Transform (with respect to time) of the quantities:

$$F_S(Q, t) = \frac{1}{N} \left\langle \sum_x e^{-iQ \cdot r(0)} e^{iQ \cdot r_x(t)} \right\rangle$$

and

$$F(Q, t) = \frac{1}{N} \left\langle \sum_{x,y} e^{-iQ \cdot r_x(0)} e^{iQ \cdot r_y(t)} \right\rangle$$

These are called the intermediate scattering functions. Going even further, we can write even these function as a Fourier Transform (with respect to position). For example, for equivalent targets (no distribution in isotope nor spin), we have

$$F_S(Q, t) = \left\langle e^{-iQ \cdot r(0)} e^{iQ \cdot r(t)} \right\rangle$$

By defining a the position of a test particle,  $n(R, t) = \delta(R - r(t))$ , we can calculate the fourier transform  $n(Q, t)$ :

$$n(Q, t) = \int d^3r e^{iQ \cdot R} n(R, t) = e^{iQ \cdot r(t)}$$

Then we have  $F_S(Q, t) = \langle n(Q, t) n(-Q, 0) \rangle$ . We can as well define the van-Hove space-time self correlation function,

$$G_s(r, t) = \int d^3r' \langle n(r', 0) n(r + r', t) \rangle$$

which represents a correlation of the test particle in space-time. The intermediate scattering function is obtained from  $G_s$  as

$$F_S(Q, t) = \int d^3r e^{iQ \cdot r} G_s(r, t)$$

These final relationship makes it clear that  $F_S$  is the Fourier transform (with respect to space) of the time-dependent correlation of the test particle density,  $n(R, t)$ , which only depends on the target characteristics.

### Example I: Resting, free nucleus

We consider the scattering from one resting free nucleus. We need only consider the self dynamics factor and we have  $b_c = \bar{b} = b$ :

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{k_f}{k_i} b^2 S(Q, \omega) = \frac{\sigma_b}{2\pi\hbar} \frac{k_f}{k_i} S(Q, \omega)$$

where we introduced the bound cross section  $\sigma_b = 4\pi b^2$  (with units of an area). Since the nucleus is free, the intermediate function is very simple. From

$$F_S(Q, t) = \left\langle e^{-iQ \cdot r(0)} e^{iQ \cdot r(t)} \right\rangle$$

we can use the BCH formula to write

$$F_S(Q, t) = \left\langle e^{-iQ \cdot [r(0) - r(t)] + \frac{1}{2} [Q \cdot r(0), Q \cdot r(t)]} \right\rangle$$

Then we want to calculate  $[r(0), r(t)]$  in order to simplify the product of the two exponential. For a free particle,  $r(t) = r(0) + \frac{p}{m}t$  and  $[r(0), p] = i\hbar$ . Then we have

$$F_S(Q, t) = \left\langle e^{-iQ \cdot [r(0) - r(t)] + i\frac{\hbar}{2m} Q^2 t} \right\rangle = \left\langle e^{-iQ \cdot p/m} \right\rangle e^{+i\frac{\hbar}{2m} Q^2 t}$$

and for a nucleus at rest ( $p = 0$ ) we have

$$F_S(Q, t) = e^{ihtQ^2/(2m)}$$

This gives the structure factor

$$S_s(Q, \omega) = \delta\left(\omega - \frac{\hbar Q^2}{2m}\right)$$

and the cross-section

$$\frac{d^2\sigma}{d\Omega dE} = \frac{\sigma_b}{2\pi\hbar} \frac{k_f}{k_i} \delta\left(\omega - \frac{\hbar Q^2}{2m}\right)$$

Since  $Q = k_f - k_i$ , we have  $Q^2 = k_i^2 + k_f^2 - 2k_i k_f \cos\vartheta$ . Also,  $\omega = E_f - E_i$  and  $k_a^2 = 2mE_a \approx 2AE_a$  where we substituted  $A$  for the mass of the nucleus.

We can then integrate the cross-section over the solid angle, to find  $\frac{d\sigma}{dE}$ :

$$\frac{d\sigma}{dE} = \int_0^\pi \frac{\sigma_b}{2\pi\hbar} \frac{k_f}{k_i} \delta\left(\omega - \frac{\hbar Q^2}{2m}\right) 2\pi \sin\vartheta d\vartheta = \frac{A\sigma_b}{4E_i} \int_{(A-1)^2/(A+1)^2 E_i}^{E_i} \delta(x) dx$$

Defining the free-atom cross section  $\sigma_f$

$$\sigma_f = \left(1 + \frac{1}{A}\right)^{-2} \sigma_b$$

we have

$$\frac{d\sigma}{dE} = \begin{cases} \sigma_f \frac{(A+1)^2}{4AE}, & \text{for } \left(\frac{A-1}{A+1}\right)^2 E < E_f < E \\ 0, & \text{otherwise} \end{cases}$$

This expression for the cross section can also be obtained more simply from an energy conservation argument.

## Example II: Scattering from a crystal lattice

We consider now the scattering of neutrons from a crystal. For simplicity, we will consider a one-dimensional crystal lattice modeled as a 1D quantum harmonic oscillator. The position  $r \rightarrow x$  (in 1D) of a nucleus in the lattice is then the position of an harmonic oscillator of mass  $M$  and frequency  $\omega_0$ ,

$$x = \sqrt{\frac{\hbar}{2M\omega_0}}(a + a^\dagger)$$

with evolution given by the Hamiltonian

$$\mathcal{H} = \frac{p^2}{2M} + \frac{M\omega_0^2}{2}x^2 = \hbar\omega_0\left(a^\dagger a + \frac{1}{2}\right)$$

If we consider no variation of isotope and spin for simplicity, we only need the self-intermediate structure function is

$$F_S(Q, t) = e^{-iQ \cdot x(0)} e^{iQ \cdot x(t)} = e^{-iQ \cdot [x(0) - x(t)]} e^{+\frac{1}{2}[Q \cdot x(0), Q \cdot x(t)]}$$

First remember that

$$x(t) = x(0) \cos(\omega_0 t) + \frac{p(0)}{M\omega_0} \sin(\omega_0 t)$$

for an harmonic oscillator. Then  $[x(0), x(t)] = [x(0), p(0)] \frac{1}{M\omega_0} \sin(\omega_0 t) = \frac{i\hbar}{M\omega_0} \sin(\omega_0 t)$ . Also we have

$$\Delta x(t) = x(t) - x(0) = x(0)[1 - \cos(\omega_0 t)] + \frac{p(0)}{M\omega_0} \sin(\omega_0 t) = \sqrt{\frac{\hbar}{2M\omega_0}}(ae^{-i\omega_0 t} + a^\dagger e^{i\omega_0 t})$$

We want to evaluate  $\langle e^{iQ\Delta x(t)} \rangle$ . Using again the BCH formula, we have

$$e^{iQ\sqrt{\frac{\hbar}{2M\omega_0}}(ae^{-i\omega_0 t} + a^\dagger e^{i\omega_0 t})} = e^{\alpha a - \alpha^* a^\dagger} = e^{-\alpha^* a^\dagger} e^{\alpha a} e^{-|\alpha|^2 [a, a^\dagger]/2}$$

with  $\alpha = iQ\sqrt{\frac{\hbar}{2M\omega_0}}e^{-i\omega_0 t}$ . Since  $[a, a^\dagger] = 1$ , we only need to evaluate the expectation value  $\langle e^{-\alpha^* a^\dagger} e^{\alpha a} \rangle$ , by expanding in series the exponentials:

$$\langle e^{-\alpha^* a^\dagger} e^{\alpha a} \rangle = \sum_{n,m} \langle a^{\dagger m} a^n \rangle \frac{\alpha^n (-\alpha^*)^m}{n!m!}$$

Only the terms with  $m = n$  survive (the other terms are not diagonal in the number basis)

$$\langle e^{-\alpha^* a^\dagger} e^{\alpha a} \rangle = \sum_n \langle (a^\dagger)^n a^n \rangle \frac{(-|\alpha|^2)^n}{(n!)^2}$$

Now  $\langle a^{\dagger n} a^n \rangle = n! \langle (a^\dagger a)^n \rangle$ , thus we finally have

$$\langle e^{-\alpha^* a^\dagger} e^{\alpha a} \rangle = \sum_n \langle (a^\dagger a)^n \rangle \frac{(-|\alpha|^2)^n}{n!} = e^{-|\alpha|^2 \langle a^\dagger a \rangle}$$

This result is a particular case of the Bloch identity,  $\langle e^A \rangle = e^{\langle A^2 \rangle/2}$  where  $A = \alpha a + \beta a^\dagger$  is any combination of the creation and annihilation operators. Finally, we obtained for the intermediate function:

$$F_S(Q, t) = e^{-iQ \cdot x(0)} e^{iQ \cdot x(t)} = e^{-\frac{Q^2 \hbar}{2M\omega_0} (\langle \hat{n} \rangle + \frac{1}{2})} e^{+\frac{1}{2} \frac{i\hbar Q^2}{M\omega_0} \sin(\omega_0 t)}$$

We can also rewrite this using the Bloch identity Using the Bloch identity,  $\langle e^A \rangle = e^{\langle A^2 \rangle/2}$  where  $A = \alpha a + \beta a^\dagger$  is any combination of the creation and annihilation operators, we can rewrite this as

$$F_S(Q, t) = \langle e^{-iQ \cdot x(0)} e^{iQ \cdot x(t)} \rangle = \langle e^{iQ\Delta x} \rangle e^{+\frac{1}{2}[Q \cdot x(0), Q \cdot x(t)]} = e^{-Q^2 \langle \Delta x^2 \rangle/2} e^{+\frac{1}{2}[Q \cdot x(0), Q \cdot x(t)]}$$

Now,

$$\langle \Delta x^2 \rangle = \langle x(0)^2 \rangle + \langle x(t)^2 \rangle + 2 \langle x(0)x(t) \rangle - \langle [x(0), x(t)] \rangle = 2 \langle x^2 \rangle + 2 \langle x(0)x(t) \rangle - \langle [x(0), x(t)] \rangle$$

from which we obtain

$$F_S(Q, t) = e^{-Q^2 \langle x^2 \rangle} e^{Q^2 \langle x(0)x(t) \rangle}$$

If the oscillator is in a number state  $|n\rangle$ , we have

$$\langle x^2 \rangle = \frac{\hbar}{2M\omega_0} (2n + 1), \quad \langle x(0)x(t) \rangle = \frac{\hbar}{2M\omega_0} [2n \cos(\omega_0 t) + e^{i\omega_0 t}]$$

If we consider an oscillator at thermal equilibrium, we need to replace  $n$  with  $\langle n \rangle_{th}$ . In the high temperature limit,  $\langle n \rangle \gg 1$  and we can simplify:

$$F_S(Q, t) = e^{-\frac{\hbar Q^2}{2M\omega_0} \langle n \rangle [1 - \cos(\omega_0 t)]} = e^{-Q^2 W_0/2} e^{Q^2 W(t)/2}$$

with  $W_0 = \frac{2\langle n \rangle \hbar}{M\omega_0}$  and  $W(t) = W_0 \cos(\omega_0 t)$ . This form of the intermediate function is the same expression one would obtain from a classical treatment and the term  $e^{-Q^2 W_0/2}$  is called the Debye-Waller factor.

The intermediate structure function is thus a Gaussian function, with a time-dependent width,  $W_0 - W(t)$ . If  $W_0 < 1$  we can make an expansion of the time-dependent term:

$$F_S(Q, t) = e^{-Q^2 W_0/2} e^{Q^2 W_0 \cos(\omega_0 t)/2} \approx e^{-Q^2 W_0/2} \left[ 1 + W_0 \cos(\omega_0 t) + \frac{1}{2} W_0^2 \cos^2(\omega_0 t) + \dots \right]$$

Then the structure factor, which is the Fourier transform of  $F_S$  will be a sum of Dirac functions at frequencies  $\omega = \pm n\omega_0$  corresponding to the n-phonon contribution to the scattering. Here the terms  $\delta(\omega - n\omega_0)$  correspond to scattering events where the energy has been transferred from the neutron to the oscillator, while terms  $\delta(\omega + n\omega_0)$  describe a transfer of energy from the lattice to the neutron. The constant term yields  $\delta(\omega)$  which describes no energy exchange or elastic scattering (zero-phonon term). Note that the expansion coefficient,  $W_0$  can be expressed in terms of the temperature, since in the high temperature limit,  $\langle n \rangle \approx \frac{k_b T}{\hbar\omega_0}$ , from which  $W_0 = \frac{2k_b T}{M\omega_0^2}$ .

In the low temperature limit,  $\langle n \rangle \rightarrow 0$ . Thus we have:

$$F_S(Q, t) = e^{-\frac{\hbar Q^2}{2M\omega_0}\{2\langle n \rangle[1-\cos(\omega_0 t)]+1-e^{i\omega_0 t}\}} \approx e^{-Q^2 \frac{\hbar Q^2}{2M\omega_0}} e^{\frac{\hbar Q^2}{2M\omega_0} e^{i\omega_0 t}}$$

Expanding in series the second term, we have

$$F_S(Q, t) \approx e^{-Q^2 \frac{\hbar Q^2}{2M\omega_0}} \left[ 1 + \frac{\hbar Q^2}{2M\omega_0} e^{i\omega_0 t} + \frac{1}{2} \left( \frac{\hbar Q^2}{2M\omega_0} \right)^2 e^{2i\omega_0 t} + \dots \right]$$

Even at low temperature, the structure factor (the Fourier transform of the expression above) is a sum of Dirac function, also called a phonon expansion. However in this case only terms  $\delta(\omega - n\omega_0)$  appear, since energy can only be given from the neutron to the lattice (which is initially in its ground state).

## 12.2 Emission and Absorption

Atoms and molecules can absorb photons and make a transition from their ground state to an excited level. From the excited state, they can emit photons (either in the presence or absence of a preexisting e.m. field) and transition to a lower level. Using TDPT and the quantization of the field we can calculate the transition rates.

### 12.2.1 Emission

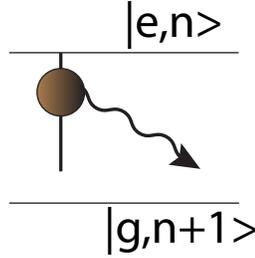


Fig. 21: Model for emission: the atom (molecule) makes a transition from the excited level ( $|e\rangle$ ) to the ground state ( $|g\rangle$ ) while the number of photons in the mode  $k, \lambda$  goes from  $n$  to  $n + 1$ .

The rate of emission is given simply by

$$W = \frac{2\pi}{\hbar} |\langle f | V | i \rangle|^2 \rho(E_f).$$

We separate the field and the atom (or molecule) levels:

$$|i\rangle = |n_{k\lambda}\rangle |e\rangle, \quad |f\rangle = |n_{k\lambda} + 1\rangle |g\rangle$$

As we are looking at atomic/optical processes the dipolar approximation is adequate and the interaction is given by:  $V = -\vec{d} \cdot \vec{E} = -e\vec{r} \cdot \vec{E}$ . Remember the expression for the electric field:

$$\vec{E} = \sum_{k,\lambda} \sqrt{\frac{2\pi\hbar\omega_k}{L^3}} \left( a_{k\lambda} e^{ikr} + a_{k\lambda}^\dagger e^{-ikr} \right) \vec{\epsilon}_{k\lambda}$$

The position of the electron which makes the transition can be written as  $\vec{r} = \vec{R} + \vec{\rho}$ , where  $\vec{R}$  is the nucleus position. Since the relative position of the electron with respect to the nucleus is  $\rho \ll \lambda$ , we can neglect it and substitute  $r$  with  $R$  in the exponential ( $\vec{\rho} \cdot \vec{k} \ll 1$ ). This simplifies the calculation, since  $R$  is not an operator acting on the electron state. Then, from the rate:

$$W = \frac{2\pi}{\hbar} |\langle g | \vec{d} | e \rangle \cdot \langle n_{k\lambda} + 1 | \vec{E} | n_{k\lambda} \rangle|^2 \rho(E_f)$$

we obtain

$$W = \frac{(2\pi e)^2}{L^3} \sum_{k',\lambda'} \omega_{k',\lambda'} \left| \langle n_{k\lambda} + 1 \left( a_{k',\lambda'} e^{ik'R} + a_{k',\lambda'}^\dagger e^{-ik'R} \right) | n_{k\lambda} \rangle \langle g | \vec{r} \cdot \vec{\epsilon}_{k',\lambda'} | e \rangle \right|^2 \rho(E_f)$$

Since we are *creating* a photon, only terms  $\propto a^\dagger$  survive and specifically the term with the correct wavevector and polarization:  $\langle n_{k\lambda} + 1 | a_{k\lambda}^\dagger | n_{k\lambda} \rangle = \sqrt{n_{k\lambda} + 1}$  (all other terms are zero). Then we have:

$$W = \frac{(2\pi e)^2}{L^3} \omega_{k,\lambda} (n_{k,\lambda} + 1) |\langle g | \vec{r} \cdot \vec{\epsilon}_{k,\lambda} | e \rangle|^2 \rho(E_f)$$

Since the atom is left in a specific final state, the density of states is defined by the e.m. field:

$$\rho(E_f) dE_f = \rho(\hbar\omega_k) \hbar d\omega_k$$

As  $\omega_k = ck$  and  $\rho(k)d^3k = \left(\frac{L}{2\pi}\right)^3 k^2 dk d\Omega = \left(\frac{L}{2\pi}\right)^3 \frac{\omega^2}{c^3} d\omega d\Omega$  we have:

$$\rho(E) = \left(\frac{L}{2\pi}\right)^3 \frac{\omega^2}{\hbar c^3} d\Omega$$

We define the dipole transition matrix element from the dipole operator  $\vec{d} = e\vec{r}$ ,  $d_{ge} = \langle g | d | e \rangle$ . The rate of emission is then:

$$W = \frac{\omega_k^3}{2\pi\hbar c^3} (n_{k\lambda} + 1) |\vec{\epsilon}_{k\lambda} \cdot \vec{d}_{ge}|^2 d\Omega$$

From this expression it is easy to see that there are two contributions to emission:

**Spontaneous** emission:

$$W = \frac{\omega_k^3}{2\pi\hbar c^3} |\vec{\epsilon}_{k\lambda} \cdot \vec{d}_{ge}|^2 d\Omega$$

which happens even in the vacuum e.m. and **stimulated** emission:

$$W = \frac{\omega_k^3}{2\pi\hbar c^3} n_{k\lambda} |\vec{\epsilon}_{k\lambda} \cdot \vec{d}_{ge}|^2 d\Omega$$

which happens only when there are already  $n$  photons of the correct mode.

Spontaneous Emission

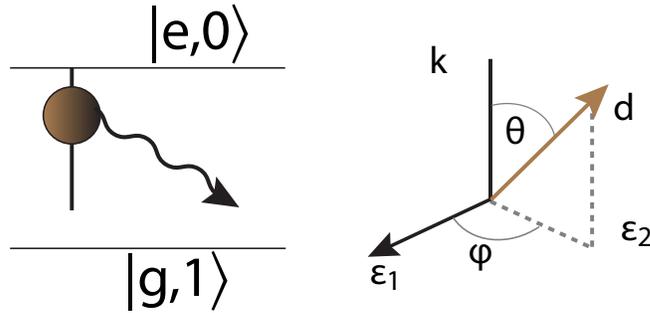


Fig. 22: Geometry of spontaneous emission

Since the photons emitted can have any polarization  $\epsilon$  and any wavevector  $\vec{k}$  direction, we have to sum over all possibilities. We assume that the dipole vector forms an angle  $\vartheta$  with respect to the wavevector  $k$ . Then the two possible polarization vectors are perpendicular to  $k$ , as in Fig. 22. The rate is the sum of the rates for each polarization  $W_{sp} = W_1 + W_2$ , each proportional to  $|d \cdot \epsilon_{k,1,2}|^2$ ,

$$d \cdot \epsilon_{k,1} = d \sin \vartheta \cos \varphi, \quad d \cdot \epsilon_{k,2} = d \sin \vartheta \sin \varphi$$

We thus obtain the typical  $\sin^2 \vartheta$  angular dependence of dipolar radiation (also seen for classical dipoles):

$$W_{sp} = \frac{\omega_k^3}{2\pi\hbar c^3} |d_{ge}|^2 \sin^2 \vartheta d\Omega$$

The total emission coefficient, or Einstein's emission coefficient, is obtained by integrating over the solid angle:

$$A_e = \int_{\Omega} W d\Omega = \frac{\omega_k^3}{2\pi\hbar c^3} |d_{ge}|^2 2\pi \int_{-1}^1 (1 - \mu^2) d\mu = \frac{4}{3} \frac{\omega_k^3}{\hbar c^3} d_{ge}^2$$

Given the rate, we can also calculate the power emitted, as rate times energy

$$P = \hbar\omega_k A_e = \frac{4}{3} \frac{\omega_k^4}{c^3} d_{ge}^2$$

Notice that this is very similar to the power emitted by a *classical* oscillating dipole (as if the e.m. field was emitted by orbiting electrons).

### Stimulated Emission

In the stimulated emission,  $W_{st}^{k\lambda} = n_{k\lambda} W_{sp}^{k\lambda}$ . Only photons with the same frequency ( $\vec{k}$ ) and polarization of the ones already in the field can be emitted. Then, as more photons in a particular mode are emitted, it becomes even more probable to produce photons in the same mode: we produce a beam of coherent photons (i.e. all with the same characteristics and phase coherent with each other). If the atoms can be kept in the excited (emitting) levels, we obtain a LASER (light amplification by stimulated emission of radiation). Of course, usually it is more probable to have the photons absorbed than to have it cause a stimulated emission, since at equilibrium we usually have many more atoms in the ground state than in the excited state,  $n_g \gg n_e$ . A mechanism capable of inverting the population of the atomic states (such as optical pumping) is then needed to support a laser.

### 12.2.2 Absorption

The rate of absorption is obtained in a way very similar to emission. The result is

$$W = \frac{2\pi}{\hbar} |\langle e | \vec{d} | g \rangle \cdot \langle n_{k\lambda} | \vec{E} | n_{k\lambda} + 1 \rangle|^2 \rho(E_f) = \frac{\omega_k^3}{2\pi\hbar c^3} n_{k\lambda} |\vec{\epsilon}_{k\lambda} \cdot \vec{d}_{eg}|^2 d\Omega$$

(as  $\langle n_{k\lambda} | a_{k\lambda} | n_{k\lambda} + 1 \rangle = \sqrt{n_{k\lambda}}$ ).

### 12.2.3 Blackbody Radiation

We consider a cavity with radiation in equilibrium with its wall. Then the polarization and  $\vec{k}$ -vector of the photons is random, and to obtain the total absorption rate we need to integrate over it, as done for the emission. We obtain

$$W_{ab} = \int_{\Omega} W_{ab}(\vartheta) d\Omega = n_k \frac{4}{3} \frac{\omega_k^3}{\hbar c^3} d_{ge}^2$$

for a given frequency (and wavevector length). Similarly, the total emission is obtained as the sum of spontaneous and stimulated emission:

$$W_e = W_{st} + W_{sp} = (n_k + 1) \frac{4}{3} \frac{\omega_k^3}{\hbar c^3} d_{ge}^2$$

In these expressions  $n_k$  is the number of photons in the mode  $k$ . Since we assumed to be at equilibrium,  $n_k$  depends only on the energy density at the associated frequency  $\omega_k$ . The energy density is given by the energy per volume, where the energy is given by the total number of photons times their energy,  $E = n_k \rho(\omega_k) \hbar \omega_k$ :

$$u(\omega_k) = \hbar \omega_k \rho(\omega_k) n_k / L^3$$

Then, from the density of states  $\rho(\omega_k) = 2 \left(\frac{L}{2\pi c}\right)^3 \omega^2 \int d\Omega = \frac{L^3}{c^3 \pi^2} \omega^2$ , we obtain

$$n_k = \frac{\pi^2 c^3}{\hbar \omega_k^3} u(\omega_k)$$

The rates can then be written in terms of the energy density and of Einstein's coefficients for absorption and emission:

$$B_{ab} = \frac{4}{3} \frac{\pi^2}{\hbar^2} d^2 \quad \rightarrow \quad W_{ab} = B_{ab} u(\omega_k)$$

$$B_{em} = B_{ab}, \quad A_e = \frac{4}{3} \frac{\omega_k^3}{\hbar c^3} d_{ge}^2 \quad \rightarrow \quad W_{em} = A_e + B_{em} u(\omega_k)$$

## Detailed Balancing

At equilibrium, we need to have the same number of photons absorbed and emitted (to preserve their total number). Then  $N_e W_{em}^k = N_g W_{ab}^k$ . Using Einstein's coefficient, we have  $N_e(A+uB) = N_g Bu$  which yields  $N_e A = uB(N_g - N_e)$ . This is the principle of detailed balancing.

We can solve for the energy density:  $u = \frac{A/B}{N_g/N_e - 1}$ . But from their explicit expressions we have  $A/B = \frac{\hbar\omega^3}{\pi^2 c^3}$  and from the condition that atoms are in thermal equilibrium, their population ratio is given by  $\frac{N_g}{N_e} = \frac{e^{-\beta E_g}}{e^{-\beta E_e}} = e^{-\beta(E_g - E_e)} = e^{\beta\hbar\omega_k}$  (since  $\hbar\omega_k$  is the exact energy needed for the transition from ground to excited state). Finally, we obtain the energy density spectrum for the black-body:

$$u(\omega_k, T) = \frac{\hbar\omega^3/\pi^2 c^3}{e^{\beta\hbar\omega_k} - 1}$$

## 12.3 Wigner-Weisskopf Theory

### 12.3.1 Interaction of an atom with a single mode e.m. field

Recall what we studied in Section 10.5. We consider again a two-level system (an atom) interacting with a single mode of the e.m. field. The Hamiltonian simplifies to  $\mathcal{H} = \mathcal{H}_0 + V$ , with

$$\mathcal{H}_0 = \hbar\nu a^\dagger a + \hbar\frac{\omega}{2}\sigma_z, \quad V = \hbar g(\sigma_+ a + \sigma_- a^\dagger)$$

where  $g = \frac{1}{2}\sqrt{\frac{\nu}{\hbar L^3}}d \cdot \epsilon$  is the dipole operator.

We move to the interaction frame defined by the  $\mathcal{H}_0$  Hamiltonian,  $U = e^{i\mathcal{H}_0 t}$ , then  $\mathcal{H}_I = UVU^\dagger$  or

$$\mathcal{H}_I = \hbar g e^{i\nu t a^\dagger a} e^{i\omega\sigma_z t/2} (\sigma_+ a + \sigma_- a^\dagger) e^{-i\nu t a^\dagger a} e^{-i\omega\sigma_z t/2} = \hbar g \left[ e^{i(\omega-\nu)t} \sigma_+ a + e^{-i(\omega-\nu)t} \sigma_- a^\dagger \right]$$

We will use the notation  $\Delta = \omega - \nu$ . We want now to study the evolution of a pure state in the interaction frame:  $i\hbar \dot{\psi} = \mathcal{H}_I |\psi\rangle$ . We can write a general state as  $|\psi\rangle = \sum_n \alpha_n(t) |e, n\rangle + \beta_n(t) |g, n\rangle$ . Notice that since we have a TLS,  $\sigma_+ |e\rangle = 0$  and  $\sigma_- |g\rangle = 0$ . The evolution is then given by:

$$\begin{aligned} i\hbar \sum_n \dot{\alpha}_n |e, n\rangle + \dot{\beta}_n |g, n\rangle &= \hbar g \sum_n [\alpha_n \sigma_- a^\dagger e^{-i\Delta t} |e, n\rangle + \beta_n \sigma_+ a e^{i\Delta t} |g, n\rangle] \\ &= \hbar g \sum_n [\alpha_n e^{-i\Delta t} \sqrt{n+1} |g, n+1\rangle + \beta_n e^{i\Delta t} \sqrt{n} |e, n-1\rangle] \end{aligned}$$

We then project these equations on  $\langle e, n|$  and  $\langle g, n|$ :

$$\begin{aligned} i\hbar \dot{\alpha}_n &= \hbar g \beta_{n+1}(t) e^{i\Delta t} \sqrt{n+1} \\ i\hbar \dot{\beta}_n &= \hbar g \alpha_{n-1}(t) e^{-i\Delta t} \sqrt{n} \end{aligned}$$

to obtain a set of equations:

$$\begin{cases} \dot{\alpha}_n = -ig\beta_{n+1} e^{i\Delta t} \sqrt{n+1} \\ \dot{\beta}_{n+1} = -ig\alpha_n e^{-i\Delta t} \sqrt{n+1} \end{cases}$$

This is a closed system of differential equations and we can solve for  $\alpha_n, \beta_{n+1}$ .

For example: we can assume that initially the atom is in the excited state  $|e\rangle$  and it decays to the ground state  $|g\rangle$  (that is,  $\beta_n(0) = 0, \forall n$ ). Then we have:

$$\alpha_n(t) = \alpha_n(0) e^{i\Delta t/2} \left[ \cos\left(\frac{\Omega_n t}{2}\right) - \frac{i\Delta}{\Omega_n} \sin\left(\frac{\Omega_n t}{2}\right) \right]$$

$$\beta_n(t) = -\alpha_n(0)e^{-i\Delta t/2} \left[ \frac{2ig\sqrt{n+1}}{\Omega_n} \sin\left(\frac{\Omega_n t}{2}\right) \right]$$

with  $\Omega_n^2 = \Delta^2 + 4g^2(n+1)$ . If initially there is no field (i.e. the e.m. field is in the vacuum state) then  $\alpha_0(0) = 1$ , while  $\alpha_n(0) = 0 \forall n \neq 0$ . Then there are only two components that are different than zero:

$$\alpha_0(t) = e^{i\Delta t/2} \left[ \cos\left(\frac{\Omega_0 t}{2}\right) - \frac{i\Delta}{\sqrt{\Delta^2 + 4g^2}} \sin\left(\frac{\Omega_0 t}{2}\right) \right]$$

$$\beta_1(t) = -e^{-i\Delta t/2} \frac{2ig}{\sqrt{\Delta^2 + 4g^2}} \sin\left(\frac{\Omega_0 t}{2}\right)$$

Thus, even in the absence of field, it is possible to make the transition from the ground to the excited state! In the semiclassical case (where the field is treated as classical) we would have no transition at all. The oscillations obtained in the quantum case are called the vacuum Rabi oscillations.

### 12.3.2 Interaction with many modes of the e.m. field

In analyzing the interaction of an atom with a *single* mode of radiation we found that transitions can occur only if energy is conserved. In the real world however we are always confronted with a finite linewidth of any transition. In order to find the linewidth we need to look at a multi-mode field.

Consider the same Hamiltonian as used in the previous section, but now we treat a field with many modes. The interaction Hamiltonian in the interaction frame is given by

$$V_I = \hbar \sum_k g_k^* a_k \sigma_+ e^{i(\omega - \nu_k)t} + g_k a_k^\dagger \sigma_- e^{-i(\omega - \nu_k)t}$$

We consider a case similar to the one consider at the end of the previous section, where initially the e.m. field is in the vacuum state and the atomic transition creates one photon. Now, however, this photon can be in one of many modes. The state vector is then:

$$|\psi(t)\rangle = \alpha(t) |e, 0\rangle + \sum_k \beta_k |g, 1_k\rangle$$

(now the index  $k$  in  $\beta_k$  label the mode and not the photon number) and the initial conditions are  $\alpha(0) = 1$ ,  $\beta_k(0) = 0$ ,  $\forall k$ . The system of equations for the coefficients are

$$\begin{cases} \dot{\alpha}(t) = -i \sum_k g_k^* e^{i(\omega - \nu_k)t} \beta_k(t) \\ \dot{\beta}_k(t) = -i g_k e^{-i(\omega - \nu_k)t} \alpha(t) \end{cases}$$

If we consider this transition as a decay process from the excited to the ground state,  $|\alpha(t)|^2$  gives the decay probability. To solve for  $\alpha(t)$  we first integrate  $\dot{\beta}$ :

$$\dot{\alpha} = -i \sum_k g_k^* e^{i(\omega - \nu_k)t} \left( \int_0^t -i g_k e^{-i(\omega - \nu_k)t'} \alpha(t') dt' \right)$$

We can rewrite the expression as:

$$\dot{\alpha} = - \sum_k |g_k|^2 \int_0^t dt' e^{-i(\omega - \nu_k)(t'-t)} \alpha(t')$$

#### Assumption 1)

We assume that the modes of the e.m. form a continuum, so that we can replace the sum by an integral  $\sum_k \rightarrow \int \rho(k) d^3k$ , with the density of states set by  $\nu_k = ck$  as usual:  $\rho(k) d^3k = 2 \left(\frac{L}{2\pi}\right)^3 k^2 dk d\varphi \sin\vartheta d\vartheta$ .

We then remember the explicit form of the interaction coupling in terms of the dipole operator:

$$|g_k|^2 = \frac{\nu_k}{4\hbar L^3} |d_{eg}|^2 \sin^2 \vartheta$$

and using again  $\nu_k = ck$  we obtain

$$\dot{\alpha} = -\frac{4|d_{eg}|^2}{(2\pi)^2 6\hbar c^3} \int_0^\infty \nu_k^3 d\nu_k \int_0^t dt' e^{-i(\omega-\nu_k)(t'-t)} \alpha(t')$$

**Assumption 2)**

In order for the transition to happen, we still need  $\nu_k \approx \omega$ .

This allows two simplifications: i) we can replace  $\nu_k^3$  with  $\omega^3$  in the integral, and ii) we can extend the lower integral limit to  $-\infty$  (since anyway we know that it will give contributions only for  $\nu_k \approx \omega$ ). By furthermore inverting the order of the integrals we obtain

$$\int_0^\infty \nu_k^3 d\nu_k \int_0^t dt' \dots \rightarrow \int_0^t dt' \alpha(t') \omega^3 \int_{-\infty}^\infty d\nu_k e^{-i(\omega-\nu_k)(t'-t)} = \int_0^t dt' \alpha(t') \omega^3 2\pi \delta(t-t') = 2\pi \alpha(t) \omega^3$$

Thus, the differential equation defining the evolution of  $\alpha(t)$  simplifies to

$$\dot{\alpha}(t) = -\frac{1}{2\pi} \frac{d_{eg}^2 \omega^3}{3\hbar c^3} \alpha(t) = -\frac{\Gamma}{2} \alpha(t)$$

Here we defined the rate of spontaneous emission

$$\Gamma = \frac{d_{eg}^2 \omega^3}{3\pi \hbar c^3}$$

Notice that the decay rate is related to Einstein's emission rate, as  $\Gamma = A_e/4\pi$  as we should expect, since it is related to the total emission (at any frequency) from the excited to the ground state.

Thus we have simply  $\alpha(t) = e^{-\Gamma t/2}$  and the decay probability  $P_d = e^{-\Gamma t}$ .

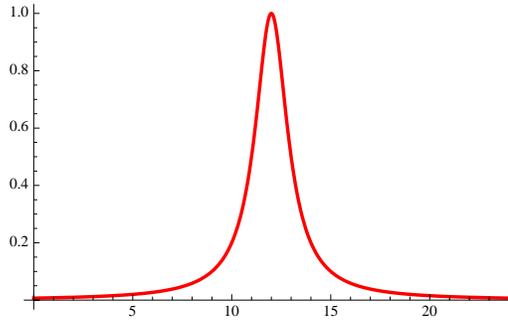


Fig. 23: Lorentzian lineshape, centered at  $\omega = 12$  and with a linewidth  $\Gamma = 2$

From the expression for  $\alpha(t)$  we can go back and calculate an explicit form for  $\beta_k(t)$ :

$$\beta_k(t) = -i \int_0^t dt' g_k e^{-i(\omega-\nu_k)t'} e^{-\Gamma t'/2} = g_k \frac{1 - e^{-i(\omega-\nu_k)t} e^{-\Gamma t/2}}{(\nu_k - \omega) + i\Gamma/2}$$

The frequency spectrum of the emitted radiation is given by  $P(\nu_k) = \rho(\nu_k) \sum_{\lambda=1,2} \int_{\Omega} d\Omega |\beta_k(t)|^2$  in the limit where  $t \rightarrow \infty$ .

$$P(\nu_k) \propto \lim_{t \rightarrow \infty} \frac{1 + e^{-\Gamma t} (1 - 2 \cos[(\omega - \nu_k)t])}{\frac{\Gamma^2}{4} + (\omega - \nu_k)^2} \sim 1 / \left( \frac{\Gamma^2}{4} + (\omega - \nu_k)^2 \right)$$

Thus the spectrum is a Lorentzian centered around  $\omega$  and with linewidth  $\Gamma$ .

## 12.4 Scattering of photons by atoms

In this section we want to study the scattering of photons by electrons (either free electrons or in an atom). We previously studied similar processes:

- Scattering theory (with an example for thermal neutrons)
- Emission and absorption of photons (in the dipole approximation)

Notice that these last processes only involved a single photon (either absorbed or emitted). Now we want to study the scattering of photons, meaning that there will be an incoming photon and an outgoing photon: this is a process that involves two photons.

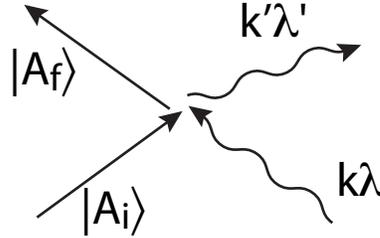


Fig. 24: Photon scattering cartoon

In order to study atom-photon interaction we need of course to start from the quantized e.m. field:

$$\mathcal{H} = \frac{\vec{p}^2}{2m} + \hbar\omega\left(n + \frac{1}{2}\right) = \frac{1}{2m} \left( p - \frac{eA}{c} \right)^2 + \hbar\omega\left(n + \frac{1}{2}\right)$$

We can separate the interaction Hamiltonian as:

$$\mathcal{H} = \mathcal{H}_0 + V = \underbrace{\frac{p^2}{2m} + \hbar\omega\left(n + \frac{1}{2}\right)}_{\mathcal{H}_0} + \underbrace{-\frac{e}{2mc}(p \cdot A + A \cdot p) + \frac{e^2}{2mc^2}A^2}_{V}$$

More generally, if there are many electrons, the interaction Hamiltonian is given by

$$V = \sum_i -\frac{e}{2mc} [p_i \cdot A(r_i) + A(r_i) \cdot p_i] + \frac{e^2}{2mc^2} A(r_i)^2$$

We already used the first term (in the dipole approximation  $\frac{e}{mc} p \cdot A \rightarrow d \cdot E$ ) to find emission and absorption processes. As stated, these processes only involve one photon. How do we obtain processes that involve two photons? Since from the term  $p \cdot A$  and in the first order perturbation theory we do not get them, we will need

- i) either terms  $\propto A^2$ , or
- ii) second order perturbation for the term  $\propto p \cdot A$ .

Notice that both these choices yield transitions that are  $\propto \alpha^2 = \left(\frac{e^2}{\hbar c}\right)^2$ , that is, that are second order in the fine structure constant.

Thus we want to calculate scattering transition rates given by  $W = \frac{2\pi}{\hbar} |K_1^{(2)} + K_2^{(1)}|^2 \rho(E_f)$ , where

- $K_1^{(2)}$  is the  $2^{nd}$  order contribution from  $V_1 = -\frac{e}{mc} \sum_i p_i \cdot A_i$  and
- $K_2^{(1)}$  is the  $1^{st}$  order contribution from  $V_2 = \frac{e^2}{2mc} \sum_i A_i^2$ .

$K_1^{(1)}$  is instead zero, since it only connects state that differ by one photon (thus it's not a scattering process) and we neglect higher orders than the second.

The initial and final eigenstates and eigenvalues are as follow (where  $\gamma$  indicate the photon):

	Initial	Final	In. Energy	Fin. Energy
$e^-$ :	$ A_i\rangle$	$ A_f\rangle$	$\epsilon_i$	$\epsilon_f$
$\gamma$ :	$ 1_{k,\lambda}, 0_{k',\lambda'}\rangle$	$ 0_{k,\lambda}, 1_{k',\lambda'}\rangle$	$\hbar\omega_k$	$\hbar\omega'_{k'}$
tot:	$ i\rangle$	$ f\rangle$	$E_i$	$E_f$

We first evaluate  $K_2^{(1)}$  for a single electron. We recall the expression for the vector potential (see Section 10.3):

$$\vec{A} = \sum_{k,\lambda} \sqrt{\frac{2\pi\hbar c^2}{L^3\omega_k}} \left( a_{k\lambda} e^{ik\cdot r} + a_{k\lambda}^\dagger e^{-ik\cdot r} \right) \vec{\epsilon}_{k\lambda}.$$

$K_2^{(1)}$  is proportional to  $A^2$ , but we only retain terms that link the correct modes ( $k, k'$ ) and that are responsible for the annihilation of a photon in mode  $k$  and the creation of a photon of mode  $k'$ . These are terms  $\propto a_{k'}^\dagger a_k$ . We find:

$$K_2^{(1)} = \langle f | V_2 | i \rangle = \frac{e^2}{2mc^2} \frac{2\pi\hbar c^2}{L^3 \sqrt{\omega_k \omega_{k'}}} \vec{\epsilon}_{k\lambda} \cdot \epsilon_{k'\lambda'} \\ \times \langle f | a_{k\lambda} a_{k'\lambda'}^\dagger e^{i(k-k')\cdot r} + a_{k\lambda}^\dagger a_{k'\lambda'} e^{-i(k-k')\cdot r} + a_{k\lambda}^\dagger a_{k'\lambda'}^\dagger e^{-i(k+k')\cdot r} + a_{k\lambda} a_{k'\lambda'} e^{i(k+k')\cdot r} | i \rangle$$

We now use the equality  $\omega_k = c|k|$  and  $\vec{k} - \vec{k}' = \vec{q} = \vec{p}/\hbar$  (the electron recoil momentum) to simplify the expression. Thus we obtain:

$$K_2^{(1)} = \frac{e^2}{2mc} \frac{2\pi\hbar c^2}{L^3 \sqrt{k k'}} \vec{\epsilon}_{k\lambda} \cdot \epsilon_{k'\lambda'} \langle A_f | e^{i\vec{q}\cdot\vec{r}} | A_i \rangle \langle 0_{k\lambda} 1_{k'\lambda'} | a_{k\lambda} a_{k'\lambda'}^\dagger | 1_{k\lambda} 0_{k'\lambda'} \rangle,$$

where the last inner product is just equal to 1. We can now extend  $K_2^{(1)}$  to many electrons:

$$K_2^{(1)} = \langle f | V_2 | i \rangle = \frac{e^2}{2m} \frac{2\pi\hbar c}{L^3 \sqrt{k k'}} \vec{\epsilon}_{k\lambda} \cdot \epsilon_{k'\lambda'} \langle A_f | \sum_i e^{i\vec{q}\cdot\vec{r}_i} | A_i \rangle.$$

This is the first contribution to the scattering matrix element, first order in perturbation theory from the quadratic term in the field potential.

We now want to calculate  $K_1^{(2)}$ , the second order contribution from the linear part  $V_1$  of the potential:

$$K_1^{(2)} = \sum_h \frac{\langle f | V_1 | h \rangle \langle h | V_1 | i \rangle}{E_i - E_h}$$

Note that this term describes virtual transitions to intermediate states since from first order transitions  $V_1$  can only create or annihilate one photon at a time. So there are two possible processes that contribute to  $K_1^{(2)}$ ,

- first absorption of one photon in the  $k\lambda$  mode followed by creation of one photon in the  $k'\lambda'$  mode: the intermediate state is zero photons in these two modes.
- first creation of one photon in the  $k'\lambda'$  mode followed by annihilation of the photon in mode  $k\lambda$ : the intermediate state is one photon in each mode.

Explicitly we have:

$$K_1^{(2)} = \sum_h \frac{\langle A_f | \langle 0_{k\lambda} 1_{k'\lambda'} | V_1 | 0_{k\lambda} 0_{k'\lambda'} \rangle | A_h \rangle \langle A_h | \langle 0_{k\lambda} 0_{k'\lambda'} | V_1 | 1_{k\lambda} 0_{k'\lambda'} \rangle | A_i \rangle}{\epsilon_i - \epsilon_h + \hbar\omega_k} \\ + \sum_h \frac{\langle A_f | \langle 0_{k\lambda} 1_{k'\lambda'} | V_1 | 1_{k\lambda} 1_{k'\lambda'} \rangle | A_h \rangle \langle A_h | \langle 1_{k\lambda} 1_{k'\lambda'} | V_1 | 1_{k\lambda} 0_{k'\lambda'} \rangle | A_i \rangle}{\epsilon_i + \hbar\omega_k - (\epsilon_h + \hbar\omega_k + \hbar\omega_{k'})}$$

Notice that  $K_1^{(2)}$  has an extra factor  $\propto \omega_k$  in the denominator with respect to  $K_2^{(1)}$ . Thus at higher energies of the incident photon (such as x-ray scattering) only  $K_1^{(2)}$  survives, while at lower energies (optical regime)  $K_2^{(1)}$  is more important.

## A. Types of Scattering

Depending on the energy  $\hbar\omega$  of the incident photon (with respect to the ionization energy  $E_I$  of the atom) and on the elastic or inelastic character of the scattering, the scattering process is designated with different names.

- **Rayleigh** scattering (Low energy, Elastic):  $\hbar\omega \ll E_I, |E_h - E_I|, E_f = E_I$ .  
The final state has the same energy as the initial one,  $E_f = E_i$  since the scattering is elastic. The scattering thus involve intermediate virtual levels, with energies  $E_h$ . We will find a cross section  $\sigma \propto \omega^4$ .
- **Raman** scattering (Low energy, Inelastic):  $\hbar\omega \ll E_I, E_f \neq E_I$ .  
Usually the final state is a different rovibrational state of the molecule, so the energy difference between initial and final state is small. If  $E_f > E_I$  the scattering process is called *Stokes*, otherwise if  $E_f < E_I$  the scattering process is called *anti-Stokes*.
- **Thomson** scattering (High energy, Elastic):  $\hbar\omega \gg E_I, E_f = E_I$ .  
This process is predominant for, e.g., soft x-ray scattering. This type of scattering can be interpreted in a semi-classical way, in the limit where the wavelength  $\lambda$  is larger than the atomic dimensions,  $\lambda \gg a_0$ . The cross section is then equivalent to what one would obtain for a free electron,  $\sigma = \frac{8}{3}\pi r_0^2$  with  $r_0$  the effective electron radius.
- **Compton** scattering (High energy, Inelastic):  $\hbar\omega \gg E_I, \lambda \ll a_0, E_f \neq E_I$ .  
For very high energy, the wavelength is small compared to the atom's size and the energy is much larger than the electron binding energy, so that the final state of the electron is an unbound state. Thus this scattering is very similar to Compton scattering (inelastic scattering) by a free electron.

Note that for x-ray scatterings the classification is slightly different than the one given above. There are two processes that competes with Coulomb scattering even at the x-ray energies:

- Electronic Raman scattering: an inelastic scattering process where the initial atomic state is the ground state and the final state an excited, discrete electronic state.
- Rayleigh scattering for x-rays: an elastic scattering process, where the final atomic state is the same as the initial state, since there is no atom excitation.

In addition to scattering processes, other processes involving the interaction of a photon with electrons are possible (besides absorption and emission of visible light that we already studied). In order of increasing photon energy, the interaction of matter with e.m. radiation can be classified as:

Rayleigh/Raman Scattering	Photoelectric Absorption	Thomson Scattering	Compton Scattering	Pair Production
$\hbar\omega < E_I$	$\hbar\omega \geq E_I$	$\hbar\omega \gg E_I$	$\hbar\omega \sim m_e c^2$	$\hbar\omega > 2m_e c^2$
$\sim \text{eV}$	$\sim \text{keV}$	$\sim \text{keV}$	$\sim \text{MeV}$	$\geq \text{MeV}$
Visible	X-rays	X-rays	$\gamma$ -rays	hard $\gamma$ -rays

## B. Semi-classical description of scattering

A classical picture is enough to give some scaling for the scattering cross section. We consider the effects of the interaction of the e.m. wave with an oscillating dipole (as created by an atomic electron).

The electron can be seen as being attached to the atom by a "spring", and oscillating around its rest position with frequency  $\omega_0$ . When the e.m. is incident on the electron, it exerts an additional force. The force acting on the electron is  $F = -eE(t)$ , with  $E(t) = E_0 \sin(\omega t)$  the oscillating electric field. This oscillating driving force is in addition to the attraction of the electron to the atom  $\sim -kx_e$ , where  $k$  (given by the Coulomb interaction strength and related to the binding energy  $E_I$ ) is linked to the electron's oscillating frequency by  $\omega_0^2 = k/m_e$ . The equation of motion for the electron is then

$$m_e \ddot{x}_e = -kx_e - eE(t) \quad \rightarrow \quad \ddot{x}_e + \omega_0^2 x_e = -\frac{e}{m_e} E(t)$$

We seek a solution of the form  $x_e(t) = A \sin(\omega t)$ , then we have the equation

$$(-\omega^2 + \omega_0^2)A = -\frac{e}{m_e} E_0 \quad \rightarrow \quad A = \frac{1}{\omega^2 - \omega_0^2} \frac{e}{m_e} E_0$$

An accelerated charge (or an oscillating dipole) radiates, with a power

$$P = \frac{2}{3} \frac{e^2}{c^3} a^2$$

where the acceleration  $a$  is here  $a = -\omega^2 A \sin(\omega t)$ , giving a mean square acceleration

$$\langle a^2 \rangle = \left( \frac{\omega^2}{\omega_0^2 - \omega^2} \frac{e}{m_e} E_0 \right)^2 \frac{1}{2}$$

The radiated power is then

$$P = \frac{1}{3} \left( \frac{e^2}{m_e c^2} \right)^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2} c E_0^2$$

The radiation intensity is given by  $I_0 = \frac{c E_0^2}{8\pi}$  (recall that the e.m. energy density is given by  $u = \frac{1}{2} E^2$  and the intensity, or power per unit area, is then  $I \sim cu$ ). Then we can express the radiated power as cross-section  $\times$  radiation intensity:

$$P = \sigma I_0$$

This yields the cross section for the interaction of e.m. radiation with atoms :

$$\sigma = \frac{8\pi}{3} \left( \frac{e^2}{m_e c^2} \right)^2 \left( \frac{\omega^2}{\omega_0^2 - \omega^2} \right)^2$$

or in SI units:

$$\sigma = \frac{8\pi}{3} \left( \frac{e^2}{4\pi\epsilon_0 m_e c^2} \right)^2 \left( \frac{\omega^2}{\omega_0^2 - \omega^2} \right)^2 = 4\pi r_e^2 \frac{2}{3} \left( \frac{\omega^2}{\omega_0^2 - \omega^2} \right)^2$$

where we used the classical electron radius<sup>40</sup>:

$$r_e = \frac{e^2}{m_e c^2}$$

which is about 2.8 fm ( $2.8 \times 10^{-15}$  m).

### 12.4.1 Thomson Scattering by Free Electrons

We consider first the Thomson scattering, which is well described by the scattering by free electrons. In this case we consider thus one single electron. Also in general, the photon should have energy high enough that the electron is seen as free even if in reality it is part of an atom (thus the photon energy must be larger than the atom's ionization energy,  $\hbar\omega \gg E_I$  or in other terms  $\lambda \gg$  than the atom's size). Note that in Thomson scattering the final electron is still a bound electron (elastic scattering) while in Compton scattering the electron is unbound (inelastic scattering). Still, since the binding energy is small compared to the other energy at play, the electron can be considered as a free electron, and many of the characteristics of Compton scattering still apply.

	Initial	Final		Initial	Final
$e^-$ :	$ A_i\rangle$	$ A_f\rangle$	En:	$mc^2 + \hbar ck$	$= \hbar ck' + \sqrt{p_e'^2 c^2 + m^2 c^4}$
$\varphi$ :	$ 1_{k,\lambda}, 0_{k',\lambda'}\rangle$	$ 0_{k,\lambda}, 1_{k',\lambda'}\rangle$	$p_x$ :	$\hbar k$	$= \hbar k' \cos \vartheta + p \cos \varphi$
tot:	$ i\rangle$	$ f\rangle$	$p_y$ :	0	$= \hbar k' \sin \vartheta - p \sin \varphi$

<sup>40</sup> The Bohr radius is a different quantity:  $r_B \sim \frac{\hbar^2}{m_e e^2}$  with some constants (depending on the units chosen) to give about  $r_B \sim 5 \times 10^{-11}$  m

The initial and final states, as well as energies and momentum are written above. They result from the conservation of energy and momentum for a relativistic electron which is initially at rest.

**? Question:** What is the ratio  $k/k'$ ? What is  $\Delta\lambda = \lambda' - \lambda$ ? (This is the usual Compton scattering formula).

From conservation of energy and momentum and with the geometry of figure 25, we can calculate the energy of the scattered photon.

$$E_\gamma + E_e = E'_\gamma + E'_e \quad \rightarrow \quad \hbar\omega + m_e c^2 = \hbar\omega' + \sqrt{|p|^2 c^2 + m^2 c^4}$$

$$\hbar\vec{k} = \hbar\vec{k}' + \vec{p} \quad \rightarrow \quad \begin{cases} \hbar k = \hbar k' \cos \vartheta + p \cos \varphi \\ \hbar k' \sin \vartheta = p \sin \varphi \end{cases}$$

From these equations we find  $p^2 = \frac{\hbar(\omega' - \omega)}{c^2} [\hbar(\omega' - \omega) - 2m_e c^2]$  and  $\cos \varphi = \sqrt{1 - \hbar^2 k'^2 \sin^2 \vartheta / p^2}$ . Solving for the change in the wavelength  $\lambda = \frac{2\pi}{k}$  we find (with  $\omega = kc$ ):

$$\Delta\lambda = \frac{2\pi\hbar}{m_e c} (1 - \cos \vartheta)$$

or for the frequency:

$$\hbar\omega' = \hbar\omega \left[ 1 + \frac{\hbar\omega}{m_e c^2} (1 - \cos \vartheta) \right]^{-1}$$

□

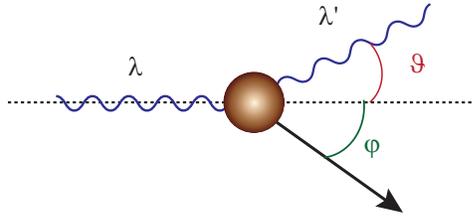


Fig. 25: Photon/Electron collision in Compton and Thomson scattering.

At these high energies,  $K_1^{(2)} \ll K_2^{(1)}$  thus we can consider only the  $K_2^{(1)}$  contribution, that we already calculated in the previous section.

To find the scattering rate and cross section we need the **density of states**:

$$\rho(E_f) dE_f = \left( \frac{L}{2\pi} \right)^3 k'^2 dk' d\Omega$$

where the final energy is  $E_f = \hbar ck' + \sqrt{p_e^2 c^2 + m^2 c^4} \approx \hbar ck' + \frac{p_e^2}{2m}$  (non-relativistic approximation). Thus we need to calculate  $\frac{dE_f}{dk'}$ . Noting that

$$p^2 / \hbar^2 = |k - k'|^2 = k^2 + k'^2 - 2kk' \cos \vartheta$$

we find

$$\frac{dE_f}{dk'} = \hbar c + \frac{\hbar^2}{2m} (2k' - 2k \cos \vartheta) = \hbar c \left[ 1 + \frac{\hbar k}{mc} \left( \frac{k'}{k} - \cos \vartheta \right) \right]$$

Solving the conservation of energy and momentum equations, we find

$$\frac{k'}{k} = \left[ 1 + \frac{\hbar k}{mc} (1 - \cos \vartheta) \right]^{-1}$$

Since  $\hbar k \ll mc$ , we can take only the first order term in  $1 + \frac{\hbar k}{mc} \left( \frac{k'}{k} - \cos \vartheta \right)$ . This is given by:  $1 + \frac{\hbar k}{mc} (1 - \cos \vartheta)$ . But this factor is just equal to  $k/k'$ . Thus we finally have:

$$\frac{dE_f}{dk'} = \hbar c \frac{k}{k'} \quad \rightarrow \quad \rho(E_f) = \left( \frac{L}{2\pi} \right)^3 \frac{k'^3}{k\hbar c} d\Omega$$

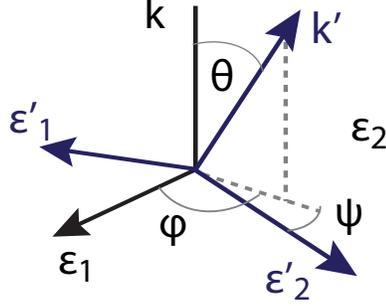


Fig. 26: Wave vectors and polarizations of scattering photons.  $\cos \gamma = \sin \vartheta \cos \psi$

Finally, to calculate the cross section, we recall the expression for the incoming flux of photons  $\Phi = c/L^3$ .

$$\frac{d\sigma}{d\Omega} = \frac{W_{fi}/d\Omega}{c/L^3} = \frac{2\pi}{\hbar} |K_2^{(1)}|^2 \frac{\rho(E_f)}{d\Omega} \frac{L^3}{c} = \left( \frac{e^2}{mc^2} \right)^2 \left( \frac{k'}{k} \right)^2 |\epsilon_{k\lambda} \cdot \epsilon_{k'\lambda'}|^2$$

With the angles defined in Fig. 26 we find:

$$\frac{d\sigma}{d\Omega} = r_e^2 \left( \frac{\omega'_k}{\omega_k} \right)^2 \sin^2 \gamma$$

where  $(\sin \gamma)^2 = 1 - \sin^2 \vartheta \cos^2(\varphi - \psi)$  and  $r_e$  is the classical electron radius. The average differential cross section (averaged over the polarization directions  $\psi$ ) is then given by

$$\left\langle \frac{d\sigma}{d\Omega} \right\rangle = r_e^2 \left( \frac{\omega'_k}{\omega_k} \right)^2 (1 - \sin^2 \vartheta / 2) = \frac{1}{2} r_e^2 \left( \frac{\omega'_k}{\omega_k} \right)^2 (1 + \cos^2 \vartheta)$$

### 12.4.2 Rayleigh Scattering of X-rays

Rayleigh scattering usually describes elastic scattering by low energy radiation. It describes for example visible light scattering from atoms: in that case, the predominant contribution comes from the term  $K_1^{(2)}$ . Rayleigh scattering also describes coherent, elastic scattering of x-rays from atoms (e.g. in a crystal) and is an important process in x-ray diffraction.

In the case of x-ray scattering, the photon energy is larger than the electronic excitation energy:  $\hbar\omega \gg E_b$ . Then we have, as stated above,  $K_1^{(2)} \ll K_2^{(1)}$  and we can neglect the  $K_1^{(2)}$  contribution. As we are considering now bound electrons, the recoil is zero, and  $\frac{dE_f}{dk'} = \hbar c$ . Then the density of states is simply  $\rho(E_f) = \frac{L}{2\pi}^3 \frac{k'^2}{\hbar c} d\Omega$ . The cross section is given by

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{2\pi}{\hbar} \frac{|K_2^{(1)}|^2}{c/L^3} \frac{\rho(E_f)}{d\Omega} = \frac{2\pi}{\hbar} \frac{c^2 r_e^2}{c/L^3} \left( \frac{2\pi\hbar}{L^3} \right)^2 \frac{1}{kk'} \left( \frac{L}{2\pi} \right)^3 \frac{k'^2}{\hbar c} |\epsilon_{k\lambda} \cdot \epsilon_{k'\lambda'}|^2 |\langle A_f | \sum_i e^{i\vec{q} \cdot \vec{r}_i} | A_i \rangle|^2 \\ &= r_e^2 \left( \frac{\omega'}{\omega} \right) |\epsilon_{k\lambda} \cdot \epsilon_{k'\lambda'}|^2 |\langle A_f | \sum_i e^{i\vec{q} \cdot \vec{r}_i} | A_i \rangle|^2 \end{aligned}$$

Consider an elastic scattering process (the inelastic scattering is called Raman scattering for x-rays). If the incoming x-ray is unpolarized, we have

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2} (1 + \cos^2 \vartheta) |\langle g | \sum_i e^{i\vec{q} \cdot \vec{r}_i} | g \rangle|^2$$

We define  $f(p) = \langle g | \sum_i e^{i\vec{q} \cdot \vec{r}_i} | g \rangle$  the atomic form factor.

- 1) Notice that for  $p \rightarrow 0$   $|\langle g | \sum_i 1 | g \rangle|^2 = Z^2$  (the atomic number squared). Thus in general we expect Rayleigh scattering to be weaker for lighter elements.
- 2) In general we can rewrite the sum as an integral  $\sum_i e^{i\vec{q} \cdot \vec{r}_i} \rightarrow \int e^{i\vec{q} \cdot \vec{r}} \tilde{\rho}(r) d^3r$  using the charge density  $\tilde{\rho}(r) = \sum_i \delta(r - r_i)$ . Then the atomic form factor takes the form:

$$f(p) = \langle g | \int e^{i\vec{q} \cdot \vec{r}} \tilde{\rho}(r) d^3r | g \rangle = \int e^{i\vec{q} \cdot \vec{r}} \bar{\rho}(r) d^3r$$

with  $\bar{\rho}(r) = \langle g | \tilde{\rho}(r) | g \rangle$ . Then the atomic form factor is the Fourier transform of the charge density.

#### Scattering from a crystal

In a crystal, we can rewrite the electron positions with the substitution  $r_i \rightarrow R_l + r_{li}$ , where  $R_l$  is the atom position (or the nucleus position or the atomic center of mass position). Then we need to sum over all atoms and all electrons in the atom. Then we have the structure factor:

$$G(q) = \langle g | \sum_{l,i} e^{i\vec{q} \cdot \vec{R}_l} e^{i\vec{q} \cdot \vec{r}_{li}} | g \rangle = \sum_l f_l(q) e^{i\vec{q} \cdot \vec{R}_l}$$

with  $f_l = \langle g | \sum_i e^{i\vec{q} \cdot \vec{r}_{li}} | g \rangle$ .

In a crystal we can rewrite the atom position as  $R_{lj} = \underbrace{l_1 a_1 + l_2 a_2 + l_3 a_3}_{\text{unit cell}} + \underbrace{r_j}_{\text{position in cell}}$ . Then

$$G(q) = \sum_{l_j} \underbrace{f_j(q) e^{i\vec{q} \cdot \vec{r}_j}}_{F(q)} e^{iq(l_1 a_1 + l_2 a_2 + l_3 a_3)} = \sum_{l_1, l_2, l_3} F(q) e^{iq(l_1 a_1 + l_2 a_2 + l_3 a_3)}$$

$F(q)$  is the form factor for the unit cell, which is tabulated for different crystals. The cross section can be written as:

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2} (1 + \cos^2 \vartheta) |F(q)|^2 \frac{\sin^2(N_1 q a_1 / 2)}{\sin^2(q a_1 / 2)} \frac{\sin^2(N_2 q a_2 / 2)}{\sin^2(q a_2 / 2)} \frac{\sin^2(N_3 q a_3 / 2)}{\sin^2(q a_3 / 2)}$$

Only when  $q a_n = 2\pi h$  the interferences terms do not vanish: this is Bragg's diffraction law.

### 12.4.3 Visible Light Scattering

When considering visible light, the wavelength is large compared to the atomic size. Then, instead of using the full interaction  $V_1 + V_2$  we can safely substitute it with the electric dipole Hamiltonian<sup>41</sup>,  $V = -\vec{d} \cdot \vec{E}$ . This Hamiltonian does not produce any two-photon process to first order, so in this case we need to consider the term  $K^{(2)}$ . This term involves virtual transitions. Since the duration of these transitions is very small, we do not have to worry about conservation of energy. Recall:

$$K^{(2)} = \sum_h \frac{\langle f|V|h\rangle \langle h|V_1|i\rangle}{E_i - E_h},$$

where  $V = -\vec{d} \cdot \vec{E}$ . The intermediate states are either  $|h\rangle = |A_h\rangle |0_{k\lambda}0_{k'\lambda'}\rangle$  or  $|h\rangle = |A_h\rangle |1_{k\lambda}1_{k'\lambda'}\rangle$ . It would be of course possible to derive the scattering cross section from the vector-potential/momentum Hamiltonian, and in that case both terms  $K_2^{(1)}$  and  $K_1^{(2)}$  should be included<sup>42</sup>.

The electric field in the Lorentz gauge is

$$E = \sum_{\ell,\xi} \sqrt{\frac{2\pi\hbar\omega_\ell}{L^3}} \left( a_{\ell\xi} e^{i\ell \cdot R} + a_{\ell\xi}^\dagger e^{-i\ell \cdot R} \right) \epsilon_{\ell\xi},$$

and thus we obtain for  $\langle h|V_1|i\rangle$  and  $\langle f|V|h\rangle$ :

$$\begin{aligned} & - \langle 0_{k\lambda}1_{k'\lambda'} | \left( a_{\ell\xi} e^{i\ell \cdot R} + a_{\ell\xi}^\dagger e^{-i\ell \cdot R} \right) | 0_{k\lambda}0_{k'\lambda'} \rangle = e^{-ik' \cdot R} \delta_{\ell,k'} \\ & - \langle 0_{k\lambda}0_{k'\lambda'} | \left( a_{\ell\xi} e^{i\ell \cdot R} + a_{\ell\xi}^\dagger e^{-i\ell \cdot R} \right) | 1_{k\lambda}0_{k'\lambda'} \rangle = e^{ik \cdot R} \delta_{\ell,k} \\ & - \langle 0_{k\lambda}1_{k'\lambda'} | \left( a_{\ell\xi} e^{i\ell \cdot R} + a_{\ell\xi}^\dagger e^{-i\ell \cdot R} \right) | 1_{k\lambda}1_{k'\lambda'} \rangle = e^{ik \cdot R} \delta_{\ell,k} \\ & - \langle 1_{k\lambda}1_{k'\lambda'} | \left( a_{\ell\xi} e^{i\ell \cdot R} + a_{\ell\xi}^\dagger e^{-i\ell \cdot R} \right) | 1_{k\lambda}0_{k'\lambda'} \rangle = e^{-ik' \cdot R} \delta_{\ell,k'} \end{aligned}$$

thus we have

$$K_1^{(2)} = \frac{2\pi\hbar}{L^3} \sqrt{\omega_k \omega_{k'}} e^{i(k-k')R} \sum_h \frac{\langle A_f | d \cdot \epsilon_{k'} | A_h \rangle \langle A_h | d \cdot \epsilon_k | A_i \rangle}{\epsilon_i - \epsilon_h + \hbar\omega_k} + \sum_h \frac{\langle A_f | d \cdot \epsilon_k | A_h \rangle \langle A_h | d \cdot \epsilon_{k'} | A_i \rangle}{\epsilon_i + \hbar\omega_k - (\epsilon_h + \hbar\omega_k + \hbar\omega_{k'})}$$

The scattering cross section is given as usual by  $\frac{d\sigma}{d\Omega} = \frac{W}{c/L^3}$ . and the density of state (assuming no recoil) is

$$\rho(E_f) = \left( \frac{L}{2\pi} \right)^3 \frac{k'^2}{\hbar c} d\Omega.$$

Finally the cross section is given by:

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{2\pi}{\hbar} \left( \frac{2\pi\hbar}{L^3} \right)^2 \omega_k \omega_{k'} \left( \frac{L}{2\pi} \right)^3 \frac{k'^2 L^3}{\hbar c} \left| \sum_h \frac{(d_{fh} \cdot \epsilon_{k'}) (d_{hi} \cdot \epsilon_k)}{\epsilon_i - \epsilon_h + \hbar\omega_k} + \frac{(d_{fh} \cdot \epsilon_k) (d_{hi} \cdot \epsilon_{k'})}{\epsilon_i - \epsilon_h - \hbar\omega_{k'}} \right|^2 \\ \frac{d\sigma}{d\Omega} &= k k'^3 \left| \sum_h \frac{(d_{fh} \cdot \epsilon_{k'}) (d_{hi} \cdot \epsilon_k)}{\epsilon_i - \epsilon_h + \hbar\omega_k} + \frac{(d_{fh} \cdot \epsilon_k) (d_{hi} \cdot \epsilon_{k'})}{\epsilon_i - \epsilon_h - \hbar\omega_{k'}} \right|^2 \end{aligned}$$

<sup>41</sup> A unitary transformation changes the Coulomb-gauge Hamiltonian into an expansion in terms of multipoles of the electromagnetic fields. For atomic interactions, only the electric dipole is kept, while higher multipoles, such as magnetic dipole and electric quadrupole, can be neglected. This unitary transformation is describe, e.g., in Cohen-Tannoudji's book, *Atom-Photons Interactions*

<sup>42</sup> This derivation can be found in Chen, S.H.; Kotlarchyk, M., *Interactions of Photons and Neutrons with Matter*, (2007)

## A. Rayleigh scattering

Rayleigh scattering describes elastic scattering, for which  $\omega_k = \omega_{k'}$  since  $|A_f\rangle = |A_i\rangle$ . Then we can simplify the cross section:

$$\frac{d\sigma}{d\Omega} = k^4 \left| \sum_h \frac{(d_{ih} \cdot \epsilon_k)(d_{hi} \cdot \epsilon_k)}{\epsilon_i - \epsilon_h + \hbar\omega_k} + \frac{(d_{ih} \cdot \epsilon_k)(d_{hi} \cdot \epsilon_k)}{\epsilon_i - \epsilon_h - \hbar\omega_k} \right|^2$$

At long wavelengths  $\hbar\omega_k \ll \epsilon_h - \epsilon_i$ , thus we can neglect  $\omega_k$  in the denominator. Then

$$\frac{d\sigma}{d\Omega} \propto \omega_k^4 \left| 2 \sum_h \frac{(d_{ih} \cdot \epsilon_k)(d_{hi} \cdot \epsilon_k)}{\epsilon_i - \epsilon_h} \right|^2$$

and simplifying we obtain that

$$\frac{d\sigma}{d\Omega} \propto \omega_k^4$$

This expression could have been found from the classical cross section we presented earlier, in the same limit  $\omega \ll \omega_0$ . The Rayleigh scattering has a very strong dependence on the wavelength of the e.m. wave. This is what gives the blue color to the sky (and the red color to the sunsets): more scattering occurs from higher frequencies photons (with shorter wavelength, toward the blue color).

As light moves through the atmosphere, most of the longer wavelengths pass straight through. Little of the red, orange and yellow light is affected by the air. However, much of the shorter wavelength light is scattered in different directions all around the sky. Whichever direction one looks, some of this scattered blue light reaches you. Since the blue light is seen from everywhere overhead, the sky looks blue. Closer to the horizon, the sky appears much paler in color, since the scattered blue light must pass through more air. Some of it gets scattered away again in other directions and the color of the sky near the horizon appears paler or white. As the sun begins to set, the light must travel farther through the atmosphere. More of the light is reflected and scattered and the sun appears less bright. The color of the sun itself appears to change, first to orange and then to red. This is because even more of the short wavelength blues and greens are now scattered and only the longer wavelengths are left in the direct beam that reaches the eyes. Finally, clouds appear white, since the water droplets that make up the cloud are much larger than the molecules of the air and the scattering from them is almost independent of wavelength in the visible range.

## B. Resonant Scattering

An interesting case arises when the incident photon energy matches the difference in energy between the atom's initial state and one of the intermediate levels. This phenomenon can occur both for elastic or inelastic scattering (Rayleigh or Raman). Assume that  $\hbar\omega_k = \epsilon_h - \epsilon_i$  for a particular  $h$  in the sum over all possible intermediate levels. Then, only first term important in  $K_1^{(2)}$  (describing first absorption and then emission) is important. In order to keep this term finite, we introduce a finite width of the level,  $\Gamma$ . The cross section then reduces to:

$$\frac{d\sigma}{d\Omega} = k k'^3 \left| \frac{(d_{fh} \cdot \epsilon_{k'})(d_{hi} \cdot \epsilon_k)}{\epsilon_h - \epsilon_i - \hbar\omega_k - i\hbar\Gamma/2} \right|_{\hbar\omega_k \approx \epsilon_h - \epsilon_i}^2 = k k'^3 \left[ \frac{|(d_{fh} \cdot \epsilon_{k'})(d_{hi} \cdot \epsilon_k)|^2}{(\epsilon_h - \epsilon_i - \hbar\omega_k)^2 + \hbar^2\Gamma^2/4} \right]_{\hbar\omega_k \approx \epsilon_h - \epsilon_i}$$

This cross section describes Raman resonance and, for  $k = k'$  resonance fluorescence.

### 12.4.4 Photoelectric Effect

In this section we want to use scattering theory of a photon from electron(s) in an atom to explain the photoelectric effect. We consider the case of an hydrogen-like atom with atomic number  $Z$  and we calculate the differential cross section

$$\frac{d\sigma}{d\omega} = \frac{W_{fi}}{\Phi_{inc}}$$

where  $W_{fi}$  is the transition rate for the scattering event and  $\Phi_{inc}$  is the incoming photon flux. The incoming photon flux can be calculated by assuming (for convenience) that the system is enclosed in a cavity of volume  $V = L^3$  (so that there's only one photon in that volume). The incoming flux of photons in the cavity is given by the number of photons per unit area and time:

$$\Phi = \frac{\#photons}{time \cdot Area} = \frac{1}{L/cL^2} = \frac{c}{L^3}$$

where the area is  $L^2$  and the time to cross the cavity is  $t = L/c$ . The transition rate  $W_{fi}$  is given by Fermi's Golden Rule, assuming an atom-photon interaction  $V$  and a density of final state  $\rho(E_f)$ :

$$W_{fi} = \frac{2\pi}{\hbar} |\langle f | V | i \rangle|^2 \rho(E_f)$$

Here the final density of states  $\rho(E_f)$  is expressed in terms of the momentum  $p$  of the scattered electron and the solid angle  $d\Omega$  where it is ejected. Indeed, as the photon is absorbed, the final density of states is only given by the free electron, again assumed to be enclosed in the volume  $V$ . The density of states for the electron is given by the density of momentum states in the cavity  $L^3$  assuming the electron propagates as a plane wave:

$$\rho(E_f) dE_f = \rho(\vec{p}) d^3\vec{p} = \left( \frac{L}{2\pi\hbar} \right)^3 p^2 dp d\Omega$$

with the (non-relativistic) energy for the electron given by  $E_f = p^2/(2m)$  giving  $dE_f = p dp/m$ . Finally

$$\rho(E_f) = \left( \frac{L}{2\pi\hbar} \right)^3 mp d\Omega$$

We next want to calculate the transition matrix element  $\langle f | V | i \rangle$ , where  $V = -\frac{e}{mc} \vec{A} \cdot \vec{p}$ . The relevant states are the photon states  $|1_{\vec{k}\lambda}\rangle$  and  $|0_{\vec{k}\lambda}\rangle$  and the electron momentum eigenstates, which in the position representation are  $\psi_i(\vec{r}) = \langle \vec{r} | i_e \rangle$  and  $\psi_f(\vec{r}) = \langle \vec{r} | f_e \rangle$ .

The matrix element between the relevant states is then:

$$\begin{aligned} V_{if} &= -\frac{e}{mc} \langle f_e | \langle 0_{\vec{k}\lambda} | \sum_{\vec{h}, \xi} \sqrt{\frac{2\pi\hbar c^2}{L^3\omega_h}} \left[ a_{\vec{h}\xi} e^{i\vec{h}\cdot\vec{r}} + a_{\vec{h}\xi}^\dagger e^{-i\vec{h}\cdot\vec{r}} \right] \epsilon_{\vec{h}, \xi} \cdot \vec{p} | 1_{\vec{k}\lambda} \rangle | i_e \rangle \\ &= -\sum_{\vec{h}, \xi} \frac{e}{m} \sqrt{\frac{2\pi\hbar}{L^3\omega_h}} \langle f_e | \left( \langle 0_{\vec{k}\lambda} | a_{\vec{h}\xi} | 1_{\vec{k}\lambda} \rangle e^{i\vec{h}\cdot\vec{r}} + \langle 0_{\vec{k}\lambda} | a_{\vec{h}\xi}^\dagger | 1_{\vec{k}\lambda} \rangle e^{-i\vec{h}\cdot\vec{r}} \right) \epsilon_{\vec{h}, \xi} \cdot \vec{p} | i_e \rangle \end{aligned}$$

The only surviving term is

$$V_{if} = -\frac{e}{m} \sqrt{\frac{2\pi\hbar}{L^3\omega_k}} \langle f_e | e^{i\vec{k}\cdot\vec{r}} \epsilon_{\vec{k}\lambda} \cdot \vec{p} | i_e \rangle$$

Then turning to the position representation of  $|i_e\rangle, |f_e\rangle$  and of the momentum operator, we can calculate an explicit expression. Using  $\psi_i(\vec{r}) = \langle \vec{r} | i_e \rangle$ ,  $\psi_f(\vec{r}) = \langle \vec{r} | f_e \rangle$  and  $\epsilon_{\vec{k}\lambda} \cdot \vec{p} = \epsilon_{\vec{k}\lambda} \cdot (-i\hbar\nabla)$ , we have:

$$\langle f_e | e^{i\vec{k}\cdot\vec{r}} \epsilon_{\vec{k}\lambda} \cdot \vec{p} | i_e \rangle = \int_V d^3\vec{r} \psi_f^*(\vec{r}) e^{i\vec{k}\cdot\vec{r}} \epsilon_{\vec{k}\lambda} \cdot (-i\hbar\nabla \psi_i(\vec{r}))$$

Finally

$$\langle f | V | i \rangle = -\frac{e}{m} \sqrt{\frac{2\pi\hbar}{L^3\omega_k}} \int_V d^3\vec{r} \psi_f^*(\vec{r}) e^{i\vec{k}\cdot\vec{r}} \epsilon_{\vec{k}\lambda} \cdot (-i\hbar\nabla \psi_i(\vec{r}))$$

The final wave function  $\psi_f$  is just a plane wave with momentum  $\vec{q} = \vec{p}/\hbar$  (in the volume  $L^3$ ). The initial wave function is instead a bound state. You should have seen that for an hydrogen-like atom the wave function is given by

$\psi_i(\vec{r}) = \frac{e^{-|\vec{r}|/a}}{\sqrt{\pi a^3}}$ , where  $a$  is the Bohr radius scaled by the atomic number  $Z$  ( $a = \hbar^2/(me^2Z)$ ). Replacing the explicit expressions for  $\psi_i$  and  $\psi_f$  in the previous result we obtain:

$$\langle f|V|i\rangle = -\frac{e}{m}\sqrt{\frac{2\pi\hbar}{L^3\omega_k}}\frac{1}{\sqrt{L^3}}\int_V d^3\vec{r}e^{i(\vec{k}-\vec{q})\cdot\vec{r}}\vec{\epsilon}_{\vec{k}\lambda}\cdot\left[-i\hbar\nabla\left(\frac{e^{-|\vec{r}|/a}}{\sqrt{\pi a^3}}\right)\right]$$

We now define  $\Delta\vec{k} = \vec{k} - \vec{q}$  and evaluate the integral:  $\int_V d^3\vec{r}e^{i\Delta\vec{k}\cdot\vec{r}}\vec{\epsilon}_{\vec{k}\lambda}\cdot\nabla\psi_i$  by parts:

$$\int_V d^3\vec{r}e^{i\Delta\vec{k}\cdot\vec{r}}\vec{\epsilon}_{\vec{k}\lambda}\cdot\nabla\psi_i = e^{i\Delta\vec{k}\cdot\vec{r}}\psi_i|_{L^3} - i\Delta\vec{k}\cdot\vec{\epsilon}_{\vec{k}\lambda}\int_V d^3\vec{r}e^{i\Delta\vec{k}\cdot\vec{r}}\psi_i(\vec{r})$$

Notice that the wavefunction vanishes at the boundaries, so the first term is zero. Also, by defining  $\vartheta$  the angle between  $\Delta\vec{k}$  and  $r$  we can rewrite the integral as:

$$-i2\pi\Delta\vec{k}\cdot\vec{\epsilon}_{\vec{k}\lambda}\int dr r^2\psi_i(r)\int_0^\pi e^{i\Delta kr\cos(\vartheta)}\sin(\vartheta)d\vartheta = -i\frac{\Delta\vec{k}\cdot\vec{\epsilon}_{\vec{k}\lambda}}{|\Delta\vec{k}|}\int dr\psi_i(r)r\sin(\Delta kr)$$

To evaluate this last integral, we can extend the interval of integration to infinity, under the assumption that  $L \gg a$ :

$$\langle f|V|i\rangle = -\frac{e}{mL^3}\sqrt{\frac{2\pi\hbar}{\omega_k}}(-i\hbar)\left(-i\frac{\Delta\vec{k}\cdot\vec{\epsilon}_{\vec{k}\lambda}}{|\Delta\vec{k}|}\right)\sqrt{\pi a^3}\int_0^\infty e^{-r/a}r\sin(\Delta kr)dr$$

and use the equivalence  $\int_0^\infty dre^{-r/a}r\sin(br) = \frac{2a^3b}{(1+a^2b^2)^2}$  to obtain:

$$-\frac{e2\pi\hbar}{mL^3}\sqrt{\frac{2\hbar a^3}{\omega_k}}\frac{\Delta\vec{k}\cdot\vec{\epsilon}a^3}{(1+a^2\Delta k^2)^2}$$

Notice that  $\Delta\vec{k}\cdot\vec{\epsilon}_k = \vec{k}\cdot\vec{\epsilon}_k - \vec{q}\cdot\vec{\epsilon}_k = -\vec{q}\cdot\vec{\epsilon}_k$  since  $\vec{k}$  and the polarization are always perpendicular.

Now considering the density of states and the incoming flux of photons  $\Phi_{inc} = c/L^3$  we obtain the scattering cross section:

$$\frac{d\sigma}{d\Omega} = \frac{32e^2a^3q(\vec{q}\cdot\vec{\epsilon}_k)^2}{m c \omega_k(1+a^2\Delta k^2)^4}$$

When the energy of the incoming photon is much higher than the electron binding energy, we have  $a\Delta k \gg 1$ . In this limit, we can rewrite the scattering cross section as

$$\frac{d\sigma}{d\Omega} = \frac{32e^2a^3q(\vec{q}\cdot\vec{\epsilon}_k)^2}{m c \omega_k(a^2\Delta k^2)^4} \propto \frac{a^3}{a^8} \propto a^{-5}$$

Now the constant  $a$  is the Bohr radius scaled by the atomic number  $Z$

$$a = \frac{\hbar^2}{me^2Z}$$

we thus find the well-known  $Z^5$  dependence of the photoelectric effect cross-section.

MIT OpenCourseWare  
<http://ocw.mit.edu>

22.51 Quantum Theory of Radiation Interactions  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.