

Notes for my course in Quantum Electronics

Mark Shtauf

School of Electrical Engineering, Dept. of Physical Electron. Tel-Aviv University,
Tel-Aviv, Israel shtauf@eng.tau.ac.il

Part I - Introduction and mathematical foundations

1. Refreshing known concepts, and discovering a few new ones

Elementary knowledge of quantum physics has been provided to you in the course Physics 3, that you should have all taken as undergraduates. There, you have learned that the quantum description of a particle is given by the quantum state (or wave) function denoted as $\psi(\vec{r}, t)$. You have also learned that the square modulus of the state function $|\psi(\vec{r}, t)|^2$ is a probability density function. Meaning that the probability of finding the particle in a volume ΔV in space is given by

$$P(\text{the particle is in } \Delta V \text{ at time } t) = \int_{\Delta V} |\psi(\vec{r}, t)|^2 d^3r. \quad (1)$$

Another thing that you have learned is that the state function satisfies the so-called Schrödinger equation

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

where ∇^2 is the Laplacian and where $V(\vec{r}, t)$ is the potential in which the particle is present. Solution of (2) with the proper initial conditions gives $\psi(\vec{r}, t)$ at all times. Equation (1) gives a physical meaning to the state-function and implies that the integration of $|\psi(\vec{r}, t)|^2$ over all space must be unity¹. The

¹ As we will see later, there are singular cases where the wave function cannot be properly normalized, such as the case of a free particle. But these cases are not

Schrödinger equation can be conveniently rewritten as

$$i\hbar \frac{\partial}{\partial t} \psi = \mathbf{H}\psi \quad (3)$$

where \mathbf{H} is an operator given by

$$\mathbf{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}, t). \quad (4)$$

The name that is used to describe this operator in quantum physics is *Hamiltonian*. Like the Hamiltonian, all operators that we will encounter in what follows, will be denoted by bold-face symbols in order to stress the difference with respect to regular variables. The Hamiltonian that we have in Eq. (4) corresponds to a particle of mass m that is affected by a potential $V(\vec{r}, t)$. The notion of a Hamiltonian was not invented only in the context of quantum mechanics and in fact, it is rigorously defined for a classical system following Newtonian dynamics. Those of you who have attended a course on analytical mechanics have surely learned that definition of the Hamiltonian. In classical mechanics, in the majority of cases, the Hamiltonian is equal to the total energy, expressed in terms of the coordinates of the system and its momenta. For a single point-size particle of mass m , subjected to a scalar potential as we have here, the classical Hamiltonian would be

$$H = \frac{p^2}{2m} + V(\vec{r}, t) \quad (5)$$

where \vec{p} is the momentum of the particle and $V(\vec{r}, t)$ is its potential energy when it happens to be at a position \vec{r} at time t . The process of quantization takes the classical Hamiltonian and turns it into a quantum operator by replacing the momentum p with a momentum operator

$$\vec{\mathbf{p}} = -i\hbar \vec{\nabla} \quad (6)$$

physical strictly speaking and we should not be too worried about them. It is easy to settle this difficulty by relating to these cases as to a limit of physical situations in which the state function is normalizable.

while leaving the coordinate vector \vec{r} as is², so that one obtains the Hamiltonian operator of Eq. (4). One may ask why is this the procedure that turns a classical system into a quantum one? The only answer that I can give you at this stage is “*because it works!*” Hopefully, this point will become more plausible to you at later stages when you get used to it and after we discuss it further.

The case of a time independent Hamiltonian

In many cases, the potential energy of the particle does not depend on time $V(\vec{r}, t) = V(\vec{r})$ and so the Hamiltonian is time independent. In that case we may solve the Schrödinger Eq. by separation of variables. In other words, we search for a solution of the form

$$\psi(\vec{r}, t) = u(\vec{r})\phi(t). \quad (7)$$

Substituting this into Eq. (3) and dividing by $\psi(\vec{r}, t)$ we have

$$i\hbar \frac{1}{\phi(t)} \frac{\partial \phi}{\partial t} = \frac{\mathbf{H}u}{u(\vec{r})}. \quad (8)$$

As usual in these cases, since the two sides of the equations depend on different variables, they should both be equal to a constant, which we denote by E . Thus, (72) gives us two equations

$$i\hbar \frac{\partial \phi}{\partial t} = E\phi(t) \quad (9)$$

$$\mathbf{H}u(\vec{r}) = Eu(\vec{r}). \quad (10)$$

The solution of the first equation is trivial $\phi(t) = K \exp(-i\omega t)$ with $\omega = E/\hbar$ and with K being an integration constant. Since in the end, the entire wave function needs to be normalized so that its square modulus integrates to

² To be more accurate we should treat this as if the coordinate \vec{r} turns into an operator \vec{r} whose effect on the state function is to multiply it by \vec{r} . Such consistency is imperative for the generalized treatment that will be presented later.

unity, the value of the constant K is immaterial at this stage and it may be safely taken as 1. The second equation is an eigenvalue equation, where it is obvious that E is simply the eigenvalue of the operator \mathbf{H} , and $u(\vec{r})$ is its corresponding eigenfunction. Let us assume that \mathbf{H} has a countable number of eigenvalues which we denote by E_n and the same number of corresponding eigenfunctions denoted by $u_n(\vec{r})$. In that case, all functions of the form $u_n(\vec{r}) \exp(-i\omega_n t)$ with $\omega_n = E_n/\hbar$ are solutions of the Schrödinger equation (3). Since, the Schrödinger equation is linear, the superposition of these functions is also a solution of (3) and thus, the most general solution of the Schrödinger equation with a time independent Hamiltonian is

$$\psi(\vec{r}, t) = \sum_n C_n u_n(\vec{r}) e^{-i\omega_n t}, \quad (11)$$

where the terms C_n are free constants, and the only constraint on their values is that the integral of $|\psi(\vec{r}, t)|^2$ over all space should be equal to 1. If the Hamiltonian has a continuum of eigenvalues, the index n should be changed into a continuous parameter and the sum turns into integration.

So what is the physical meaning of the eigenvalues E_n ? Many of you know that this is the spectrum of energies. In other words the values of E_n represent all possible values that the energy of our particle may have. If one measures the energy of the particle, one will surely obtain one of these values. What is the probability of obtaining a particular E_n in a measurement of the energy? The answer is $|C_n|^2$ provided of course that the state function (and hence the coefficients C_n) have been properly normalized. Why is E_n the energy spectrum, and why is $|C_n|^2$ the probability of measuring that energy? This is a consequence of an important principle in quantum mechanics, which I will explain after a “not very short” mathematical detour, which will be taken after we finish with the following famous example.

Recalling a famous example - a particle in a potential well

Assume a one dimensional potential well where

$$V(x) = \begin{cases} 0, & \text{when } |x| < a \\ V_0, & \text{when } |x| > a \end{cases},$$

so that the stationary Schrödinger equation (10) (together with the Hamiltonian of Eq. (4)) assumes the form

$$\frac{\partial^2 u(x)}{\partial x^2} + \frac{2m}{\hbar^2}(E - V)u(x) = 0. \quad (12)$$

Its solution has a different form in the regions $|x| > a$ and $|x| < a$. For $|x| < a$, $V = 0$ and the solutions can be written as harmonic functions

$$u(x) = \alpha_0 e^{ikx} + \beta_0 e^{-ikx} \quad (13)$$

with $k = \sqrt{2mE}/\hbar$ (convince yourselves that E must be positive as there are no solutions corresponding to negative energy). In the region where $|x| > a$ we must divide the solutions to two types. When $E > V_0$ we have harmonic solutions

$$u(x) = \alpha_- e^{ik'x} + \beta_- e^{-ik'x}, \quad \text{when } x < -a \quad (14)$$

$$u(x) = \alpha_+ e^{ik'x} + \beta_+ e^{-ik'x}, \quad \text{when } x > a \quad (15)$$

with $k = \sqrt{2m(E - V_0)}/\hbar$. For $E < V_0$ the solutions at $|x| > a$ have the form of decaying exponentials

$$u(x) = \alpha_- e^{\rho x}, \quad \text{when } x < -a \quad (16)$$

$$u(x) = \alpha_+ e^{-\rho x}, \quad \text{when } x > a, \quad (17)$$

where $\rho = \sqrt{2m(V_0 - E)}/\hbar$. We must now require that the boundary conditions at $|x| = a$ are continuous both for $u(x)$ and for its derivative $u'(x)$. This analysis is very similar to the analysis of optical, or electromagnetic waveguides, the solutions can be classified in terms of "modes" (like in optics), which can be divided into symmetric modes and antisymmetric modes. The fundamental (lowest energy) mode is a cosine function whose period is

approximately equal to twice the width of the well (i.e. $\sim 4a$). The tails decay exponentially at $|x| > a$. In the range where $E < V_0$, the energy cannot attain any arbitrary value. Only a set of discrete values can be allowed for the energy E , which are the eigenvalues of the Hamiltonian, or the energy spectrum of the problem. If one tries other values of E , the boundary conditions cannot be satisfied. When $E > V_0$ it may have any value, and the spectrum outside the well is said to be continuous.

The problem reduces to a very simple form when V_0 is taken to infinity. Then the decay is very quick and one may say that the boundary conditions for the wave-function are zero at $|x| = a$. The solutions then are of the form (13) with $k = n\pi/(2a)$ with n being a nonnegative integer. This implies that the allowed energy levels are

$$E_n = \frac{\hbar^2 \pi^2}{8ma^2} n^2 \quad (18)$$

The solutions corresponding to the lowest few energy levels are

$$\begin{aligned} u_1(x) &= \sqrt{\frac{1}{a}} \cos\left(\frac{\pi x}{2a}\right) \\ u_2(x) &= \sqrt{\frac{1}{a}} \sin\left(\frac{\pi x}{a}\right) \\ u_3(x) &= \sqrt{\frac{1}{a}} \cos\left(\frac{3\pi x}{2a}\right) \end{aligned} \quad (19)$$

etc. It can be readily concluded that the n -th order solution assumes the form $u_n(x) = a^{-1/2} \sin\left[n\pi\left(\frac{x}{2a} - \frac{1}{2}\right)\right]$.

The functions $u_n(x)$ are stationary solutions of the Hamiltonian, whereas the full solution has the form of $\psi(x, t) = \sum_{n=1}^{\infty} C_n u_n(x) \exp(i\omega_n t)$, where $\omega_n = E_n/\hbar$. For example, a particle can be in the state $\psi(x, t) = 2^{-1/2} u_1(x) \exp(-i\omega_1 t) + 2^{-1/2} u_2(x) \exp(-i\omega_2 t)$. Try to think what its energy is in this case.

2. Mathematical foundations³

*The vectors spaces \mathcal{E} and \mathcal{E}^**

We assume a generic linear inner product space (of the kind that you have studied in an introductory course in linear algebra) denoted by \mathcal{E} , whose elements (vectors) are called "Kets" (plural of "Ket") and denoted by placing a distinctive letter between a column "|" and a right angled bracket ">", as in $|\psi\rangle$, or $|\phi\rangle$. The name Ket and the notation using the right angled bracket in order to represent it are due to Paul Dirac, and the logic behind these choices will become clear imminently. Another linear space that we have to introduce is the space of all linear functionals that operate on vectors in \mathcal{E} . A functional is something that, when applied to a vector, it produces a complex number (a scalar). The space of functionals will be denoted by \mathcal{E}^* and it is dual to the space \mathcal{E} (because the functionals are defined with respect to the vectors in \mathcal{E}). The elements in \mathcal{E}^* will be called "Bras" and they are denoted by placing a distinctive letter between the symbols "<" and "|", as in $\langle\chi|$. The result of applying the Bra $\langle\chi|$ to a Ket $|\psi\rangle$ produces a complex number $\langle\chi|\psi\rangle$, called "BraKet," which sounds like "bracket," and hence the motivation for the choice of names. Since \mathcal{E}^* is a space of linear functionals, the relation $\langle\chi|(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha\langle\chi|\psi\rangle + \beta\langle\chi|\phi\rangle$, is satisfied for any two vectors $|\psi\rangle, |\phi\rangle$ and for any two complex numbers α and β . Similarly $(\alpha\langle\chi| + \beta\langle\chi'|)|\psi\rangle = \alpha\langle\chi|\psi\rangle + \beta\langle\chi'|\psi\rangle$ for any two Bras $\langle\chi|$ and $\langle\chi'|$ and for any two complex numbers α and β .

An important notion is that for any Ket $|\psi\rangle$ in \mathcal{E} , one can find a corresponding Bra in \mathcal{E}^* , which will be denoted by $\langle\psi|$, such that the result of applying this functional on some arbitrary vector $|\phi\rangle$ is the inner product between the vectors $|\psi\rangle$ and $|\phi\rangle$. We will call the Bra $\langle\psi|$ and the Ket $|\psi\rangle$ Hermitian conjugates of one another,

$$|\psi\rangle^\dagger = \langle\psi| \quad \text{and} \quad \langle\psi|^\dagger = |\psi\rangle.$$

³ This part is a summary of the corresponding discussion in Chapter 2 of [1].

Finally, as follows from the definitions of inner products,

$$\langle \psi | \phi \rangle = \langle \phi | \psi \rangle^*.$$

It should be noted that while for every Ket there is a corresponding Bra, the opposite statement is not always rigorous. In other words there may be a functional $\langle \chi |$ that corresponds to no legitimate vector $|\chi\rangle$. In order to see this, consider for example the space of all square integrable functions, with the usual definition of the inner product⁴, so that $|f\rangle$ represents some square integrable function $f(x)$. Assume a functional, represented by the Bra $\langle x_0 |$ whose effect on $|f\rangle$ is the following $\langle x_0 | f \rangle = f(x_0)$. Namely, it produces the value of $f(x)$ at the point x_0 . The Bra $\langle x_0 |$ is a perfectly legitimate linear functional, but if we were to look for the Ket $|x_0\rangle$ that represents its Hermitian conjugate, we would have found that it corresponds to Dirac's delta function $\delta(x - x_0)$, which is not a legitimate member in the space \mathcal{E} .⁵ Such peculiarities are unavoidable in the algebraic foundations of quantum physics, and occasionally their interpretation appears troublesome. Nonetheless, we will treat Kets such as $|x_0\rangle$ as semi-legitimate elements of \mathcal{E} , while introducing the appropriate caveats when necessary.

As is always the case with linear spaces, there is an infinite number of orthonormal bases that can span them. If we choose one particular such basis $\{|u_j\rangle\}_{j=1}^N$ then any vector $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{n=1}^N C_n |u_n\rangle \quad (20)$$

where C_n are complex coefficients. It should be noted that the number of basis elements N may be infinite. At this stage we consider cases in which the basis elements are countable (the index n is a natural number), but more

⁴ The inner product between two complex functions $f(x)$ and $g(x)$ is usually defined as $\int f^*(x)g(x)dx$ with the limits of the integration being from $-\infty$ to ∞ .

⁵ Dirac's Dela $\delta(x - x_0)$ is not a function in a strict mathematical sense, let alone square integrable.

general scenarios where the basis elements form a continuum (i.e. the index n is replaced by a real number) exist and will be discussed subsequently. The orthogonality of the basis vectors $\langle u_k | u_n \rangle = \delta_{k,n}$ implies that

$$C_n = \langle u_n | \psi \rangle \quad (21)$$

and hence we may write

$$|\psi\rangle = \sum_{n=1}^N \langle u_n | \psi \rangle |u_n\rangle = \left[\sum_{n=1}^N |u_n\rangle \langle u_n| \right] |\psi\rangle, \quad (22)$$

which implies that

$$\sum_{n=1}^N |u_n\rangle \langle u_n| = \mathbf{1} \quad (23)$$

is the Identity operator. Equation (23) is known as the *closure* relation, which is satisfied by all orthonormal bases. Note that the coefficients C_n uniquely define the vector $|\psi\rangle$ in the $\{|u_j\rangle\}_{j=1}^N$ basis. If we stack them on top of one another we obtain a column vector representation of $|\psi\rangle$. This column vector is meaningful only when we announce the basis in which it is written. The Bra $\langle\psi|$ is spanned by the bras $\{\langle u_j|\}_{j=1}^N$ in the form

$$\langle\psi| = \sum_{n=1}^N \langle\psi|u_n\rangle \langle u_n| = \sum_{n=1}^N C_n^* \langle u_n|, \quad (24)$$

and it can be uniquely represented (in the $\{|u_j\rangle\}_{j=1}^N$ basis) by a row vector whose elements are C_n^* ($n = 1 \dots N$). In what follows I will use $\vec{\psi}$ to denote the column vector representing the Ket $|\psi\rangle$ so that the Bra $\langle\psi|$ is correspondingly represented by its transpose conjugate row vector, denoted consistently by $\vec{\psi}^\dagger$. While the numeric values of the vector components depend on the choice of representation (i.e the basis), the scalar products between vectors $\vec{\phi}^\dagger \vec{\psi} = \langle\psi|\phi\rangle$ are independent of representation and hence we may expect that

a universal physical meaning can be assigned to them.⁶ The same is true for the vectors' norms, as the square of the norm is simply the scalar product of the vector with itself.

*Linear operators in \mathcal{E} and \mathcal{E}^**

An operator \mathbf{A} in \mathcal{E} is something that, when applied to a Ket $\mathbf{A}|\psi\rangle$ in \mathcal{E} produces another Ket in \mathcal{E} . It is called a linear operator if for any two Kets $|\psi\rangle$ and $|\phi\rangle$ it satisfies the relation $\mathbf{A}(|\psi\rangle + |\phi\rangle) = \mathbf{A}|\psi\rangle + \mathbf{A}|\phi\rangle$. We have already seen the unity operator $\mathbf{1}$, which is an example of a linear operator. The operators $|u_j\rangle\langle u_j|$ that were used to construct the identity are also linear operators. They are called projection operators. In general, for any normalized Ket $|\phi\rangle$, the operator

$$\mathbf{P}_{|\phi\rangle} = |\phi\rangle\langle\phi| \quad (26)$$

⁶ In order to see that, assume two vectors $|\psi\rangle$ and $|\phi\rangle$ whose representation in some basis $\{|u_j\rangle\}_{j=1}^N$ are

$$|\psi\rangle \rightarrow \vec{\phi} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \cdot \\ \cdot \\ c_N \end{pmatrix} \quad \text{and} \quad |\phi\rangle \rightarrow \vec{\psi} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \cdot \\ \cdot \\ d_N \end{pmatrix},$$

where $c_j = \langle u_j|\psi\rangle$ and $d_j = \langle u_j|\phi\rangle$. If we perform a scalar product $\vec{\phi}^\dagger\vec{\psi}$, we obtain

$$\vec{\phi}^\dagger\vec{\psi} = \sum_{j=1}^N c_j^*d_j = \sum_{j=1}^N \langle\phi|u_j\rangle\langle u_j|\psi\rangle = \langle\phi|\left[\sum_{j=1}^N |u_j\rangle\langle u_j|\right]|\psi\rangle = \langle\phi|\psi\rangle. \quad (25)$$

Namely, the scalar product performed with the column representations of the vectors is equal to the inner product between $|\phi\rangle$ and $|\psi\rangle$. While the components c_j and d_j depend on the choice of basis, the scalar product is independent of representation.

is called the projection operator corresponding to $|\phi\rangle$. Indeed when applied to an arbitrary vector $|\psi\rangle$ it produces

$$\mathbf{P}_{|\phi\rangle}|\psi\rangle = \langle\phi|\psi\rangle|\phi\rangle,$$

which is the projection of $|\psi\rangle$ on $|\phi\rangle$. As noted earlier, we will always follow the convention of denoting symbols representing operators with boldface letters. If we return to the example where $|\psi\rangle$ represents a square integrable function $\psi(x)$, then the derivative is an example for a linear operator. Namely $\mathbf{D}_x|\psi\rangle$ produces a new Ket $|\psi'\rangle$ which represents the function $\psi'(x) = d\psi/dx$. Another example for a linear operator is multiplication by x . Namely $\mathbf{X}|\psi\rangle = |\phi\rangle$ where $\phi(x) = x\psi(x)$. Since we are only going to be interested in linear operators, repetition of the word 'linear' seems superfluous and hence it will be dropped in what follows.

Operators can be combined arbitrarily with one another. For example the two operators \mathbf{A} and \mathbf{B} can be combined into \mathbf{AB} . This combination should be interpreted as follows

$$(\mathbf{AB})|\psi\rangle = \mathbf{A}[\mathbf{B}|\psi\rangle], \quad (27)$$

meaning that \mathbf{A} is applied to the result of applying \mathbf{B} to the vector. One must be very careful not to mix up the order of operators because in general $\mathbf{AB} \neq \mathbf{BA}$. It is customary to define the *commutator* between two operators as

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}, \quad (28)$$

which is an operator as well.⁷ The commutation relations between operators play an extremely important role in quantum physics. Some general properties of the commutator will be presented later in this chapter.

The action of an operator \mathbf{A} on a ket $|\psi\rangle$ can be written as

$$\mathbf{A}|\psi\rangle = \sum_{j,k} |u_j\rangle \langle u_j|\mathbf{A}|u_k\rangle \langle u_k|\psi\rangle = \sum_{j,k} \left(\langle u_j|\mathbf{A}|u_k\rangle \langle u_k|\psi\rangle \right) |u_j\rangle \quad (29)$$

where we have inserted the closure relation (23) twice in order to obtain the right-hand side. Equation (29) implies that the representation of the vector $\mathbf{A}|\psi\rangle$ in the $\{|u_j\rangle\}$ basis is obtained by multiplying the column vector $\vec{\psi}$ (which represents $|\psi\rangle$ in that basis) from the left by a matrix \underline{A} whose j, k -th element is $\langle u_j|\mathbf{A}|u_k\rangle$. This can be written as follows

$$\begin{pmatrix} \langle u_1|\mathbf{A}|\psi\rangle \\ \langle u_2|\mathbf{A}|\psi\rangle \\ \langle u_3|\mathbf{A}|\psi\rangle \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle u_1|\mathbf{A}|u_1\rangle & \langle u_1|\mathbf{A}|u_2\rangle & \langle u_1|\mathbf{A}|u_3\rangle & \cdots \\ \langle u_2|\mathbf{A}|u_1\rangle & \langle u_2|\mathbf{A}|u_2\rangle & \langle u_2|\mathbf{A}|u_3\rangle & \cdots \\ \langle u_3|\mathbf{A}|u_1\rangle & \langle u_3|\mathbf{A}|u_2\rangle & \langle u_3|\mathbf{A}|u_3\rangle & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix} \begin{pmatrix} \langle u_1|\psi\rangle \\ \langle u_2|\psi\rangle \\ \langle u_3|\psi\rangle \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}, \quad (30)$$

while keeping in mind that vectors and matrices may have infinite dimensions.

The eigenvalue equation of an operator \mathbf{A} is

$$\mathbf{A}|a_j\rangle = a_j|a_j\rangle \quad (31)$$

where a_j is the j -th eigenvalue of \mathbf{A} and $|a_j\rangle$ is the corresponding eigenvector.⁸ The set of all of the eigenvalues that an operator has is called the *spec-*

⁷ As an example, let us now consider the commutator $[\mathbf{x}, \mathbf{D}_x]$, in the space where Kets represent square integrable functions. We may write

$$[\mathbf{x}, \mathbf{D}_x]\psi(x) = x \frac{d\psi}{dx} - \frac{d[x\psi(x)]}{dx} = -\psi(x)$$

which shows us that $[\mathbf{x}, \mathbf{D}_x] = -1$. A general property of the commutator is that

$$[\mathbf{A}, \mathbf{B}] = -[\mathbf{B}, \mathbf{A}], \text{ and therefore in our particular example } [\mathbf{D}_x, \mathbf{x}] = 1.$$

⁸ Notice that not all of the operators in what follows will be represented by capital letters and hence expressions such as $\mathbf{a}|a\rangle = a|a\rangle$ will be occasionally seen.

trum of eigenvalues. In writing Eq. (31) we have assumed that the spectrum of the eigenvalues of \mathbf{A} is discrete, but in practice we will soon encounter cases where the spectrum of eigenvalues is continuous. In such cases the index j is dropped and the eigenvalue equation assumes the form $A|a\rangle = a|a\rangle$, where a can be any number within a continuous range of values. There are also cases in which the spectrum of eigenvalues is mixed. Namely, it is continuous in some range of values, but discrete in another range.⁹ Different eigenvalues will always have different eigenvectors (not necessarily vice versa).

It is common practice to refer to functions of operators. In principle, since we have already acquainted ourselves with operator multiplication, expressions such as \mathbf{A}^n for an integer n should not pose any problem for us. This means that we can also easily interpret functions of operators such as $\exp(\mathbf{A})$ as they can be expanded as a power series

$$e^{\mathbf{A}} = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{A}^n.$$

It is clear that the eigenvectors of an operator \mathbf{A} are also the eigenvectors of \mathbf{A}^n with the eigenvalue being a^n , and hence it can be trivially generalized that

$$e^{\mathbf{A}}|a\rangle = e^a|a\rangle,$$

implying in the general case that the eigenvalue of a function of an operator is the function of the operator's eigenvalue.

The Hermitian conjugate of an operator

The Hermitian conjugate of an operator \mathbf{A} is denoted by \mathbf{A}^\dagger and it can be defined by means of the relation $\langle\psi|\mathbf{A}^\dagger = (\mathbf{A}|\psi\rangle)^\dagger$. In other words, the operator \mathbf{A}^\dagger is the one that transforms the Bra $\langle\psi|$, which corresponds to the Ket $|\psi\rangle$, into the Bra that corresponds to the Ket $\mathbf{A}|\psi\rangle$ for any $|\psi\rangle$ in the

⁹ Recall the example of the Hamiltonian of a particle in a potential well that we have reviewed previously. It had a discrete spectrum of eigenvalues (energies) inside the well, and a continuous spectrum outside it.

space. Applying Hermitian conjugation to both sides of Eq. (29) we obtain

$$\langle \psi | \mathbf{A}^\dagger = \sum_{j,k} \langle \psi | u_j \rangle \langle u_j | \mathbf{A} | u_k \rangle \langle u_k |, \quad (32)$$

which implies that the row vector representing the Bra $\langle \psi | \mathbf{A}^\dagger$ in the $\{|u_j\rangle\}$ basis is given by $\vec{\psi}^\dagger \underline{\underline{A}}^\dagger$ where $\vec{\psi}^\dagger$ and $\underline{\underline{A}}^\dagger$ are the transpose conjugates of the vector $\vec{\psi}$ representing $|\psi\rangle$ and the matrix $\underline{\underline{A}}$ representing \mathbf{A} , respectively. Notice that the above definition also implies that for any two Kets $|\psi\rangle$ and $|\varphi\rangle$, $\langle \psi | \mathbf{A}^\dagger | \varphi \rangle = (\langle \varphi | \mathbf{A} | \psi \rangle)^*$, a relation that facilitates the search for the Hermitian conjugate operator in many cases.¹⁰

Hermitian operators

An operator \mathbf{A} is said to be Hermitian if

$$\mathbf{A}^\dagger = \mathbf{A}. \quad (33)$$

Such operators have a very important role in quantum physics. Notice that when we have a product of operators, Hermitian conjugation works as follows $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$, which means that a product of Hermitian operators is not hermitian, unless the two operators commute (i.e. unless $[\mathbf{A}, \mathbf{B}] = 0$). Since \mathbf{A} commutes with itself, \mathbf{A}^2 and consequently any power (or real function) of a Hermitian operator is Hermitian as well. One of the most important properties of Hermitian operators is that their *eigenvectors are necessarily orthogonal* to each other (as long as they correspond to non-identical eigenvalues). Thus, the eigenvectors of a Hermitian operator, once normalized, form an orthonormal basis for the space in which the operator acts. If some

¹⁰ As an exercise, let us consider the example where \mathcal{E} is the space of square integrable functions, and find the conjugate operator of $\mathbf{A} = \mathbf{D}_x = \partial/\partial x$. To do this, we use the relation $\int \psi^*(x) \mathbf{D}_x^\dagger \varphi(x) dx = \left(\int \varphi^*(x) \frac{\partial \psi}{\partial x} dx \right)^* = \int \psi^*(x) \left(-\frac{\partial \varphi}{\partial x} \right) dx$, where we have performed integration by parts and used the fact that square integrable functions vanish at $\pm\infty$. Thus, the result is $\mathbf{D}_x^\dagger = -\mathbf{D}_x$.

of the eigenvalues are degenerate, then any superposition of the corresponding eigenvectors is still an eigenvector with the same eigenvalue. In that case not every choice of linearly independent eigenvectors is orthogonal, but one can make a particular choice that is. A useful relation that we should stress for the sake of what follows is that a Hermitian operator \mathbf{A} can be expressed as a superposition of projectors onto its eigenvectors. In other words

$$\mathbf{A} = \sum_j a_j |u_j\rangle\langle u_j|. \quad (34)$$

Proving this relation is fairly straightforward and it is left as an exercise for you. Another very important property of Hermitian operators is that their eigenvalues are necessarily real. This can be readily seen by noting that Eq. (31) implies that $a_j = \langle a_j | \mathbf{A} | a_j \rangle$ and $a_j^* = \langle a_j | \mathbf{A}^\dagger | a_j \rangle$. Since in the case of Hermitian operators $\mathbf{A} = \mathbf{A}^\dagger$, one finds that $a_j = a_j^*$ indicating that the eigenvalues are real.

Unitary operator

An operator \mathbf{U} is said to be unitary if

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{U} \mathbf{U}^\dagger = \mathbf{1}. \quad (35)$$

An important property of unitary operators is that they do not change, neither the norm, nor the scalar product between vectors.¹¹ Additionally, since a unitary operator does not change vectors' norms, the modulus of its eigenvalues must be equal to 1. It is an excellent exercise for you to show that if \mathbf{A} is Hermitian then $[\exp(i\mathbf{A})]^\dagger = \exp(-i\mathbf{A})$ and that $\exp(i\mathbf{A}) \exp(-i\mathbf{A}) = \mathbf{1}$,¹² implying that $\exp(i\mathbf{A})$ is unitary.

¹¹ The scalar product of $\mathbf{U}|\psi\rangle$ and $\mathbf{U}|\phi\rangle$ is $\langle \psi | \mathbf{U}^\dagger \mathbf{U} | \phi \rangle = \langle \psi | \phi \rangle$.

¹² It can be easily done using the series expansion of the exponent and noting that since \mathbf{A} commutes with itself, the product of the exponents is the same as the exponent of the sum.

The trace of an operator

Similarly to matrices, the trace is defined as the sum of all the diagonal elements of the operator in some given representation, $\{|u_j\rangle\}$

$$\text{Tr}\{\mathbf{A}\} = \sum_j \langle u_j | \mathbf{A} | u_j \rangle. \quad (36)$$

Surprisingly, in spite of the fact that the above definition of the trace contains the vector elements of the basis that one chooses, the result is basis independent. To see that, we may choose another basis $\{|v_j\rangle\}$ and insert the unity operator $\sum_j |v_j\rangle\langle v_j|$ into Eq. (36), as follows

$$\begin{aligned} \text{Tr}\{\mathbf{A}\} &= \sum_{j,k} \langle u_j | v_k \rangle \langle v_k | \mathbf{A} | u_j \rangle = \sum_{j,k} \langle v_k | \mathbf{A} | u_j \rangle \langle u_j | v_k \rangle \\ &= \sum_k \langle v_k | \mathbf{A} | v_k \rangle. \end{aligned} \quad (37)$$

The independence of the trace on the choice of basis suggests that the trace of operators should have a universal meaning (like the length of vectors for example). When we return to quantum mechanics, we will see that this is indeed the case.

Commuting operators share the same eigenvectors

We will see that this property has major consequences in quantum physics. In order to demonstrate it, assume that $|u_j\rangle$ are eigenvectors of \mathbf{B} with eigenvalues b_j and that $[\mathbf{A}, \mathbf{B}] = 0$. In that case, we may write

$$\mathbf{A}\mathbf{B}|u_j\rangle = \mathbf{B}(\mathbf{A}|u_j\rangle) = b_j(\mathbf{A}|u_j\rangle)$$

which means that the vector $|v_j\rangle = \mathbf{A}|u_j\rangle$ is also an eigenvector of \mathbf{B} with the same eigenvalue b_j . Assuming that the eigenvalue b_j is not degenerate, this implies that $|v_j\rangle$ is proportional to $|u_j\rangle$, i.e. $\mathbf{A}|u_j\rangle = a_j|u_j\rangle$. In other words, $|u_j\rangle$ is an eigenvector of \mathbf{A} with an eigenvalue a_j . If the eigenvalue b_j of \mathbf{B} is degenerate then $\mathbf{A}|u_j\rangle$ is a superposition of the eigenvectors of \mathbf{B} that

have this particular eigenvalue. In this case, these degenerate eigenvectors of \mathbf{B} can be replaced by their appropriate superposition, such that the new set of degenerate vectors are also eigenvectors of \mathbf{A} . Alternatively, one could find the eigenvectors of \mathbf{A} and claim that they should also be the eigenvectors of \mathbf{B} . Assuming that the two operators are not degenerate with respect to the same groups of vectors, one can always find a unique set of vectors that are eigenvectors of both.

Important identities concerning the commutation relation

The following equations summarize the most useful properties of the commutator operation. Their proof is a homework exercise for you.

$$[\mathbf{A}, \mathbf{B}] = -[\mathbf{B}, \mathbf{A}] \quad (38)$$

$$[\mathbf{A}, \mathbf{B} + \mathbf{C}] = [\mathbf{A}, \mathbf{B}] + [\mathbf{A}, \mathbf{C}] \quad (39)$$

$$[\mathbf{A}, \mathbf{BC}] = [\mathbf{A}, \mathbf{B}]\mathbf{C} + \mathbf{B}[\mathbf{A}, \mathbf{C}] \quad (40)$$

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = 0 \quad (41)$$

$$[\mathbf{A}, \mathbf{B}]^\dagger = [\mathbf{B}^\dagger, \mathbf{A}^\dagger] \quad (42)$$

Operators that commute with their commutator

When two operators \mathbf{A} and \mathbf{B} commute with their commutator $[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] = [\mathbf{B}, [\mathbf{A}, \mathbf{B}]] = 0$, the following two relations hold:

$$[\mathbf{A}, F(\mathbf{B})] = [\mathbf{A}, \mathbf{B}]F'(\mathbf{B}), \quad (43)$$

and

$$e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}+\frac{1}{2}[\mathbf{A},\mathbf{B}]} = e^{\mathbf{A}+\mathbf{B}}e^{\frac{1}{2}[\mathbf{A},\mathbf{B}]} \quad (44)$$

The latter relation is known as Glauber's formula.¹³ We will find these two relations very useful in the analysis that follows. The proof of these two relations is part of your homework assignments.

¹³ More general versions of Glauber's formula exist, but are not necessary for our use.

When the vectors that form a basis for \mathcal{E} belong to a continuum

The bases that span \mathcal{E} that we have been considering so far, were of the form $\{|u_j\rangle\}_{j=1}^N$, where j was a natural number. There are situations in which an extension of this framework is necessary. One motivation for this extension comes up when recalling the existence of cases in which the eigenvalue spectrum of a Hermitian operator is continuous. In such cases, its eigenvectors, which are known to form a basis in \mathcal{E} , also belong to a continuum. While important examples for such situation will be presented in what follows, at this point we will assume the existence of a basis $\{|u\rangle\}$ that spans \mathcal{E} , with $u \in (-\infty, \infty)$, and explore its properties. By definition, any Ket $|\psi\rangle$ can be expressed as a superposition of the basis vectors $|u\rangle$,

$$|\psi\rangle = \int du C(u)|u\rangle \quad (45)$$

where the coefficients C_n of Eq. (20) are replaced with $C(u)$ and the sum is replaced by an integral. Extension of the ideas underpinning Eqs. (20) and (22) requires that the orthogonality relation is interpreted as follows

$$\langle u'|u\rangle = \delta(u - u'), \quad (46)$$

with Dirac's delta replacing Kronecker's delta that appeared in the discrete case. Application of $\langle u_0|$ from the left implies that

$$C(u) = \langle u|\psi\rangle, \quad (47)$$

and substitution of this expression back into Eq. (45) leads to the closure relation

$$\int du |u\rangle\langle u| = \mathbf{1}, \quad (48)$$

which is the continuous equivalent of Eq. (23). Notice that the expansion coefficients $C(u)$ are the elements of the column vector that represents the Ket

$|\psi\rangle$ in the $|u\rangle$ basis, except that this column vector does not contain discrete vector elements, as we had previously. Instead, it contains a continuum of values and therefore referring to it as a ‘column’ vector is somewhat non-intuitive. It is more customary to call $C(u)$ the function (or in a future context – wave function) representing $|\psi\rangle$ in the $|u\rangle$ basis. Finally, as we will find in what follows, it is a far more convenient choice of notation to replace $C(u)$ with $\psi(u)$, in which case Eq. (45) is rewritten as,

$$|\psi\rangle = \int du \psi(u) |u\rangle \quad (49)$$

with $\psi(u) = \langle u | \psi \rangle$.

Hermitian operators whose commutator is i

We shall see in what follows that Hermitian operators whose commutator is a constant proportional to i play a profoundly important role in the theory of quantum physics. For this reason, their properties are addressed here in some detail. We will denote these operators by \mathbf{q} and \mathbf{k} such that the relation

$$[\mathbf{q}, \mathbf{k}] = i \quad (50)$$

is satisfied. The most elegant approach for uncovering the special properties of such operators relies on some trickery as we will see now. We define an auxiliary operator

$$\mathbf{S}_\lambda(\mathbf{k}) = e^{-i\lambda\mathbf{k}}, \quad (51)$$

where λ is a scalar parameter. As we have seen previously, the Hermiticity of \mathbf{k} implies that $\mathbf{S}_\lambda(\mathbf{k})$ is unitary, and since \mathbf{q} and \mathbf{k} both commute with their commutator (which is a constant) we may use (43) to write

$$[\mathbf{q}, \mathbf{S}_\lambda(\mathbf{k})] = i(-i\lambda)\mathbf{S}_\lambda(\mathbf{k}) = \lambda\mathbf{S}_\lambda(\mathbf{k}). \quad (52)$$

which implies that $\mathbf{q}\mathbf{S}_\lambda(\mathbf{k}) = \mathbf{S}_\lambda(\mathbf{k})[\mathbf{q} + \lambda]$. Now assume that the operator \mathbf{q} has eigenvectors $|q\rangle$ and corresponding eigenvalues q . We see at once that

$$\mathbf{q}\mathbf{S}_\lambda(\mathbf{k})|q\rangle = (q + \lambda)\mathbf{S}_\lambda(\mathbf{k})|q\rangle. \quad (53)$$

In other words $\mathbf{S}_\lambda(\mathbf{k})|q\rangle = |q + \lambda\rangle$. If we fix the phase appropriately, we have

$$|q\rangle = \mathbf{S}_q(\mathbf{k})|0\rangle \quad (54)$$

This shows that the eigenvalues of the operator \mathbf{q} , whose only requirement is to be Hermitian and to have a commutator of i with another Hermitian operator \mathbf{k} , has a continuum of eigenvalues. For this reason, its eigenvectors $|q\rangle$ (which form a basis in \mathcal{E}) also belong to a continuum. Obviously, symmetry considerations imply that the same statements also apply to the eigenvalues and eigenvectors of \mathbf{k} .

Furthermore, since $\mathbf{S}_q^\dagger(\mathbf{k}) = \mathbf{S}_{-q}(\mathbf{k})$ we have $\langle q|\mathbf{S}_{-dq}(\mathbf{k}) = \langle q + dq|$. Denoting by $|k\rangle$ and k , the eigenvectors of \mathbf{k} and their respective eigenvalues. We may write

$$\langle q + dq|k\rangle = \langle q|\mathbf{S}_{-dq}(\mathbf{k})|k\rangle = \langle q|k\rangle + idqk\langle q|k\rangle + O(dq^2) \quad (55)$$

This implies that

$$\frac{d\langle q|k\rangle}{dq} = ik\langle q|k\rangle$$

so that

$$\langle q|k\rangle = \frac{1}{\sqrt{2\pi}}e^{iqk} \quad (56)$$

and therefore

$$\langle k|q\rangle = \frac{1}{\sqrt{2\pi}}e^{-iqk}, \quad (57)$$

which implies that $\tilde{\psi}(k) = \langle k|\psi\rangle$ is the fourier transform of $\psi(q) = \langle q|\psi\rangle$

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int dq \psi(q) e^{-iqk}.$$

From here it is obvious how the operator \mathbf{k} is represented in the $|q\rangle$ representation. What we are interested in is

$$\langle q|\mathbf{k}|\psi\rangle$$

for a generic state $|\psi\rangle$. Inserting the identity, the above can be rewritten as

$$\int dk \langle q|\mathbf{k}|k\rangle \langle k|\psi\rangle = \frac{1}{\sqrt{2\pi}} \int k \tilde{\psi}(k) e^{iqk} dk = \frac{1}{i} \frac{\partial \psi(q)}{\partial q}. \quad (58)$$

Similarly, in the opposite direction, the effect of \mathbf{q} on $\tilde{\psi}(k)$ is given by

$$\langle k|\mathbf{q}|\psi\rangle = i \frac{\partial \tilde{\psi}(k)}{\partial k}.$$

As we shall now see, these properties, that at this stage may appear to you as nothing more than esoteric mathematical peculiarities, are immensely useful and important in describing the works of nature.

Going back to quantum physics - observables

In the following we will briefly describe some postulates that quantum physics is based upon.

We will first describe the procedure for quantizing a system that is defined classically¹⁴ and focus initially on the case of a single particle. It starts

¹⁴ We are discussing the quantization of a system that has a classical description.

Quantities, such as the spin of particles, have no classical counterpart and are defined only in the realm of quantum physics. Also, by stating that the system is well defined classically, we mean that the coordinates and their conjugate momenta are defined, and that the classical Hamiltonian (in most cases just the total energy of the system) is expressed as a function of those parameters. If we consider a single particle for example, then there are only 3 coordinates represented by a vector \vec{r} and three corresponding (conjugate) momenta, that are represented by a vector \vec{p} .

from associating with the vectors of coordinates $\vec{r} = (r_1, r_2, r_3)^t$ and momenta $\vec{p} = (p_1, p_2, p_3)^t$, the respective vector operators $\vec{\mathbf{r}} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)^t$ and $\vec{\mathbf{p}} = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)^t$, whose components are scalar operators that satisfy the commutation relations

$$\begin{aligned} [\mathbf{r}_j, \mathbf{r}_i] &= [\mathbf{p}_j, \mathbf{p}_i] = 0 \\ [\mathbf{r}_j, \mathbf{p}_i] &= i\hbar\delta_{i,j}, \end{aligned} \quad (59)$$

with i and j being either 1,2, or 3.¹⁵ Then, the the system is completely described in terms of a state-vector $|\psi\rangle$ whose evolution in time is given by the Schrödinger equation

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \mathbf{H}|\psi\rangle. \quad (60)$$

Notice that this formulation is not related to any particular basis. It can be easily reduced to the familiar $|\vec{r}\rangle$ basis by applying the bra $\langle \vec{r}|$ to both sides.

Another postulate is the following. All measurable (i.e physically meaningful) quantities are represented by Hermitian operators. And the only possible outcomes in a measurement of that quantity are the eigenvalues of that operator. In other words, denoting some generic physical quantity by the Hermitian operator \mathbf{A} we have that the possible outcomes a_j in a measurement of that quantity satisfy the equation

$$\mathbf{A}|a_j\rangle = a_j|a_j\rangle. \quad (61)$$

Namely, the possible outcomes a_j are the eigenvalues of the operator \mathbf{A} and the state-vectors $|a_j\rangle$ are the corresponding eigenvectors¹⁶. The second part of this postulate is that the probability of attaining a particular value a_j in the measurement of an operator \mathbf{A} is

¹⁵ It should be noted that for the sake of notational convenience, we will occasionally use $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and $\mathbf{p}_x, \mathbf{p}_y, \mathbf{p}_z$ instead of $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ and $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, respectively, in what follows.

¹⁶ Note that since observables are represented by Hermitian operators the eigenvalues - i.e. the possible outcomes of measurements are *real*.

$$P(a_j) = |\langle a_j | \psi \rangle|^2, \quad (62)$$

where $|\psi\rangle$ is the state of the system. This postulate allows us to understand why the energies that can be attained by a particle in a potential well are quantized¹⁷ and equal only to the values that we have obtained in Eq. (18). These were just the eigenvalues of the energy operator in that particular problem, i.e the values for which $\mathbf{H}|\psi_E\rangle = E|\psi_E\rangle$. Similarly, the eigenvalues of $\vec{\mathbf{r}}$ and $\vec{\mathbf{p}}$ are the possible outcomes in a measurement of the particles position and momentum, respectively. Since we have encountered these operators in the previous section we know that their eigenvalues are continuous.

The quantization of classical quantities that are functions of \vec{p} and \vec{r} is usually quite straightforward, as they involve the replacement of these parameters by the corresponding operators. For example, the classical Hamiltonian of a particle in a scalar potential is $H = p^2/2m + V(\vec{r})$ and thus the quantum Hamiltonian becomes $\mathbf{H} = \mathbf{p}^2/2m + V(\vec{\mathbf{r}})$. The same is true for angular momentum $\vec{L} = \vec{r} \times \vec{p}$ which becomes $\vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$. But here we have to be more careful in general. For example, the classical angular momentum in the z direction is $L_z = xp_y - yp_x$, which turns into $\mathbf{L}_z = \mathbf{x}p_y - \mathbf{y}p_x$ which is a hermitian operator and everything is OK. On the other hand if we consider a quantity that involves terms of the form xp_x , we cannot quantize it as $\mathbf{x}p_x$ because the product is not Hermitian as the operators \mathbf{p}_x and \mathbf{x} do not commute. In this case quantization is performed with the help of the *symmetrization rule*. Namely

$$xp_x \longrightarrow \frac{1}{2} [\mathbf{x}p_x + p_x\mathbf{x}].$$

This is true for the quantization of all such combinations of \vec{p} and \vec{r} .

Since the probability of finding a result a_j in a measurement of an observable \mathbf{A} is $P(a_j) = |\langle a_j | \psi \rangle|^2$, the average value in a measurement of \mathbf{A}

¹⁷ Here the word *quantized* means *discrete*. This usage of the word should not be confused with its other, more important, meaning, which is "described within the realm of quantum physics".

is

$$\sum_j P(a_j)a_j = \langle \psi | \left[\sum_j a_j |a_j\rangle \langle a_j| \right] | \psi \rangle = \langle \psi | \mathbf{A} | \psi \rangle. \quad (63)$$

Similarly the mean square value in a measurement of \mathbf{A} can be shown to be $\langle \psi | \mathbf{A}^2 | \psi \rangle$. Therefore the variance in a measurement of that type is

$$\text{Var}\{\mathbf{A}\} = \langle \psi | \mathbf{A}^2 | \psi \rangle - \langle \psi | \mathbf{A} | \psi \rangle^2. \quad (64)$$

In what follows we will often encounter the notation $\langle \mathbf{A} \rangle$ that indicates the average of the physical quantity \mathbf{A} , taken with respect to the state of the system. This notation will also be used for higher moments. The notation $\Delta \mathbf{A}$ will be used to denote the standard deviation of an observable with respect to the state of a given system. Namely

$$\Delta \mathbf{A} = \sqrt{\langle \mathbf{A}^2 \rangle - \langle \mathbf{A} \rangle^2}. \quad (65)$$

A very brief description of a measurement and the uncertainty principle

As noted earlier earlier, for a system described by some generic state function, the result of a measurement is unknown deterministically. It can only be described in terms of probabilities. The Operator \mathbf{A} representing the observable, can be expressed in terms of its eigenvectors as

$$\mathbf{A} = \sum_j a_j |a_j\rangle \langle a_j|,$$

an expression that we have seen before. When this quantity \mathbf{A} is measured, the probability of obtaining a given result a_j is $|\langle a_j | \psi \rangle|^2$ as we have said earlier. Yet a physical fact is that if the system is measured again very briefly after a first measurement which gave the result a_{j_0} , the same result will follow again and again (provided of course that the measurement did not destroy

the measured particle, and that the state did not have time to evolve). This means that following a measurement, the system collapses to a particular eigenstate $|a_{j_0}\rangle$ of \mathbf{A} . If after performing an exact measurement of \mathbf{A} , one performs an exact measurement of a second observable \mathbf{B} , the result is going to be deterministic only if $|a_j\rangle$ is also an eigenstate of \mathbf{B} . In all other cases, the result in a measurement of \mathbf{B} is again unknown deterministically. When \mathbf{A} and \mathbf{B} commute they share the same eigenvectors and the measurements can be performed in either order, with the result of the second measure being known in advance.¹⁸ What this illustrates is a basic quantum mechanical property which states that simultaneous exact measurements of non-commuting observables are not possible. The picture of polarized light passing through a polarization beam-splitter may help visualize these things.

A special case of non-commuting observables is pairs of observables whose commutator is an imaginary constant. This is the property that the position and conjugate momentum operators have according to Eq. (59). For such observable pairs a very special uncertainty relation can be prescribed. We will derive this relation after the following example.

A simple numerical example

Here is a simple example. Consider three Hermitian operators \mathbf{A} , \mathbf{B} and \mathbf{C} defined in a two dimensional space that is spanned by the basis functions $|u_1\rangle$ and $|u_2\rangle$. Assume that the representations of these three operators in the basis of $|u_1\rangle$ and $|u_2\rangle$ are given by¹⁹

$$\mathbf{A} \rightarrow \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbf{B} \rightarrow \begin{pmatrix} 1.5 & 0.5i \\ -0.5i & 1.5 \end{pmatrix} \quad \text{and} \quad \mathbf{C} \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (66)$$

It is easily shown that

¹⁸ This is guaranteed to be so if that the measured eigenvalue is not degenerate.

¹⁹ Notice that there is no equality sign between the operators and the matrices that represent them. This is to stress that the matrices are representations in a particular basis, whereas the operators are a more general thing. The same is true for the distinction between the eigenvectors and their representation vectors later on.

$$[\mathbf{A}, \mathbf{B}] = 0, \quad [\mathbf{A}, \mathbf{C}] \rightarrow 2 \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \quad \text{and} \quad [\mathbf{B}, \mathbf{C}] \rightarrow \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \quad (67)$$

The eigenvalues of \mathbf{A} are $a_1 = 1$ and $a_2 = -1$ and their respective eigenvectors are

$$|a_1\rangle \rightarrow 2^{-1/2} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad |a_2\rangle \rightarrow 2^{-1/2} \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

The eigenvalues of \mathbf{B} are $b_1 = 1$ and $b_2 = 2$ and the respective eigenvectors are $|b_1\rangle = |a_1\rangle$ and $|b_2\rangle = |a_2\rangle$, as must follow from the fact that $[\mathbf{A}, \mathbf{B}] = 0$. The eigenvalues of \mathbf{C} are $c_1 = 1$ and $c_2 = -1$ (same as the eigenvalues of \mathbf{A}) and the respective eigenvectors are

$$|c_1\rangle \rightarrow 2^{-1/2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |c_2\rangle \rightarrow 2^{-1/2} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

which differ from the eigenvectors of \mathbf{A} and \mathbf{B} . It is easy to show that $\mathbf{A} = a_1|a_1\rangle\langle a_1| + a_2|a_2\rangle\langle a_2|$, with similar expressions holding for the other two operators. Assume that the system was initially prepared in the state $|\psi\rangle = |u_1\rangle$ whose representation in our basis is simply

$$|\psi_{in}\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Consider a measurement of \mathbf{A} followed by the measurement of \mathbf{B} followed by the measurement of \mathbf{C} . Since $|\psi_{in}\rangle = 2^{-1/2}|a_1\rangle - i2^{-1/2}|a_2\rangle$, the measurement of \mathbf{A} will yield either a_1 , or a_2 with probability of 0.5. For, example, if it yields a_1 , then the state of the system after measurement of \mathbf{A} is $|\psi_{in,B}\rangle = |a_1\rangle$. Since $|\psi_{in,B}\rangle = |b_1\rangle$, the measurement of B will yield $b_1 = 1$ with complete certainty and the state will remain unchanged (i.e. $|\psi_{in,C}\rangle = |\psi_{in,B}\rangle = |b_1\rangle$). Next we express $|\psi_{in,C}\rangle$ in terms of the eigenvectors of \mathbf{C} ,

$$|\psi_{in,C}\rangle = 0.5(1+i)|c_1\rangle + 0.5(1-i)|c_2\rangle.$$

Implying that either values c_1 , or c_2 will be obtained with probabilities $|0.5(1 \pm i)|^2 = 0.5$.

The uncertainty principle for position and conjugate momentum

The uncertainty principle states that the position \mathbf{r}_j and its conjugate momentum \mathbf{p}_j ($j = x, y, z$) satisfy the relation²⁰

$$\Delta \mathbf{r}_j \Delta \mathbf{p}_j \geq \frac{\hbar}{2}. \quad (68)$$

In order to prove this uncertainty we may drop the index j for simplicity. Assume a system in some state $|\psi\rangle$ and define a pair of centralized operators $\mathbf{r}_c = \mathbf{r} - \langle r \rangle$ and $\mathbf{p}_c = \mathbf{p} - \langle p \rangle$. The centralization of the operators did not change their commutation relations obviously. Then lets look at the norm of the vector $|\phi\rangle = (\mathbf{p}_c - i\lambda \mathbf{r}_c)|\psi\rangle$. It is given by

$$\begin{aligned} \langle \phi | \phi \rangle &= \lambda^2 \langle \mathbf{r}_c^2 \rangle + i\lambda \langle [\mathbf{r}_c, \mathbf{p}_c] \rangle + \langle \mathbf{p}_c^2 \rangle \\ &= \lambda^2 \langle \mathbf{r}_c^2 \rangle - \lambda \hbar + \langle \mathbf{p}_c^2 \rangle \geq 0 \end{aligned} \quad (69)$$

To guarantee that this quantity is greater than 0 for all $|\psi\rangle$, the discriminant of the above expression needs to be non-positive. In other words, we must have

$$\langle \mathbf{r}_c^2 \rangle \langle \mathbf{p}_c^2 \rangle \geq \frac{\hbar^2}{4}.$$

Since $\langle \mathbf{r}_c^2 \rangle = \Delta \mathbf{r}^2$ and $\langle \mathbf{p}_c^2 \rangle = \Delta \mathbf{p}^2$, we have the uncertainty relation (68) written above. Equivalently, we could have said that any two parameters that constitute a Fourier pair satisfy an uncertainty relation of this type. It can be also shown that the minimum amount of uncertainty characterizes the Gaussian case, when

$$\langle q | \psi \rangle = \psi(q) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-\frac{q^2}{4\sigma^2}}$$

²⁰ Since, as we have seen, the quantities \mathbf{r}_j and \mathbf{p}_j are defined solely in terms of their commutation relations, one may generalize that any two physical quantities that the commutation relation between them is given by $[\mathbf{a}, \mathbf{b}] = iC$ (with constant C) satisfy the relation $\Delta \mathbf{a} \Delta \mathbf{b} \geq \frac{C}{2}$.

$$\langle p|\psi\rangle = \tilde{\psi}(p) = \frac{\sqrt{\sigma/\hbar}}{\sqrt{2\pi}} e^{-\frac{p^2}{\hbar^2}\sigma^2}. \quad (70)$$

A little more about commuting observables

We have seen above that when two operators \mathbf{A} and \mathbf{B} do not commute, the properties that they represent cannot be determined with arbitrary accuracy. We have seen that measurement of \mathbf{A} collapses the state of the measured system to one of the eigenstates of \mathbf{A} , $|a_j\rangle$ and the result of the measurement a_j is just the corresponding eigenvalue. Thus, unless, $|a_j\rangle$ happens to be also an eigenstate of \mathbf{B} , the result in a subsequent measurement of \mathbf{B} is fundamentally uncertain. This is because any eigenvalue b_k of \mathbf{B} may be measured with probability $P(b_k) = |\langle b_k|a_j\rangle|^2$. The exact amount of uncertainty is quantified by the square root of the variance, as we have seen before, and its value depends on the particular operators whose simultaneous measurement is considered. A case of particular value and simplicity is that of the coordinate and momentum operators that we considered above, and where the product of uncertainties is bounded from below by $\hbar/2$. Let us now talk a little bit about operators whose commutator is 0. There is a particular situation that is worth some attention. It is when operators operate on separate subspaces. For example, we know that \mathbf{x} commutes with \mathbf{p}_y , implying that there is nothing that prevents the simultaneous knowledge of position along the x coordinate and the momentum along the y coordinate. Similarly the position of a particle along all three coordinates can be determined with arbitrary precision, because the operators \mathbf{X} , \mathbf{Y} and \mathbf{Z} commute. Based on what we have learned before, the operators \mathbf{X} , \mathbf{Y} and \mathbf{Z} should have the same eigenvectors. Similarly the operators \mathbf{X} and \mathbf{P}_y must have the same eigenvectors, but how can it be consistent with the fact that the eigenvectors of \mathbf{Y} and \mathbf{P}_y are different (since \mathbf{Y} and \mathbf{P}_y don't commute)?

The answer is quite simple, the operators \mathbf{X} and \mathbf{P}_y operate in different subspaces. The overall three-dimensional space in which the particle in our previous examples lives is spanned by the basis vectors

$$|\vec{r}\rangle = |x, y, z\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle$$

where the symbol \otimes indicates external (tensor) product²¹. Thus, the operator \mathbf{X} relates only to one property of the particle, but it is degenerate with respect to the other properties that pertain to the y and z coordinates. The word 'degenerate' means that for any given eigenvalue x_0 of \mathbf{X} , there are more than one eigenvectors. In fact the number of eigenvectors in our present case is infinite, because all the vectors $|x_0, y, z\rangle$ are eigenvectors of \mathbf{X} with the same eigenvalue x_0 .

Lets return to the particle in a box example, that we talked about in our first lesson. But now lets talk explicitly about the case of a particle in a three-dimensional cubic box of side a . The solution is obtained through separation of variables (along the x,y and z coordinates). The solution along each coordinate is identical to what we had before, i.e $u_n(x) = a^{-1/2} \sin \left[n\pi \left(\frac{x-a}{2a} \right) \right]$, and the full solution is

$$u_{n,m,l}(x, y, z) = a^{-3/2} \sin \left[n\pi \left(\frac{x-a}{2a} \right) \right] \sin \left[m\pi \left(\frac{y-a}{2a} \right) \right] \sin \left[l\pi \left(\frac{z-a}{2a} \right) \right]$$

The energy of the particle is given by

$$E_{n,m,l} = \frac{\hbar^2 \pi^2}{8ma^2} (n^2 + m^2 + l^2).$$

The state of the system is now determined, not by a single value like in the one-dimensional case, but rather by the three values associated with the 3 different coordinates. Specifying the energy of the particle is not sufficient in this case, because the energy can be degenerate. For example, if we know that the particle energy is $E = 6 \frac{\hbar^2 \pi^2}{8ma^2}$, the particle could be in any of the states²² (2, 1, 1), (1, 2, 1), or (1, 1, 2).

²¹ The meaning of a tensor product is discussed in Appendix I.

²² The ground-state energy on the other hand is not degenerate.

Time evolution

Given an initial state $|\psi(t_0)\rangle$ the solution of the (time independent) Schrödinger equation can be formally expressed as

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\mathbf{H}t}|\psi(0)\rangle, \quad (71)$$

an expression that can be easily verified by substitution into the Schrödinger equation. Expansion of $|\psi(t_0)\rangle$ in terms of the eigenstates of the Hamiltonian operator (the energy eigenstates) yields the expression that we have seen before

$$|\psi(t)\rangle = \sum_j C_j e^{-i\omega_j t} |E_j\rangle \quad (72)$$

where $|E_j\rangle$ are the energy eigenstates and where $\omega_j = E_j/\hbar$. Since the eigenstates are orthonormal, normalization implies that $\langle\psi(t)|\psi(t)\rangle = \sum_j |C_j|^2 = 1$ and the values of the individual coefficients C_j are determined by the initial conditions. If, for example, we know that at $t = 0$ the system started from $2^{-1/2}[|E_1\rangle + |E_2\rangle]$, then

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}[e^{-i\omega_1 t}|E_1\rangle + e^{-i\omega_2 t}|E_2\rangle] = \frac{e^{-i\omega_1 t}}{\sqrt{2}}[|E_1\rangle + e^{-i\Delta\omega_{2,1}t}|E_2\rangle] \quad (73)$$

where $\Delta\omega_{2,1} = \omega_2 - \omega_1$. The rightmost expression emphasizes the fact that the state oscillates at a frequency equal to the energy difference divided by \hbar . Notice that in the continuous case Eq. (72) changes into,

$$|\psi(t)\rangle = \int d\omega C(\omega) e^{-i\omega t} |\omega\rangle, \quad (74)$$

a relation that is reminiscent of an inverse Fourier transform and therefore implies an uncertainty relation between t and $\omega = E/\hbar$. Indeed if the wavefunction is well-localized in time, it is bound to contain a broad spectrum of energies. On the other hand, an eigen-state of the energy operator, has a well specified energy, but it is independent of time (see Eq. 72). That is why

energy states are called stationary. Nonetheless, although the uncertainty between t and E is similar to that between \mathbf{r}_j and \mathbf{p}_j , one should note that t is an evolution parameter and *not* an operator in quantum mechanics.

One may also consider the time evolution of the mean value of an observable $\langle\psi(t)|\mathbf{A}|\psi(t)\rangle$. This is a nice exercise in quantum algebra²³

$$\begin{aligned}\frac{d}{dt}\langle\psi(t)|\mathbf{A}|\psi(t)\rangle &= \langle\dot{\psi}|\mathbf{A}|\psi\rangle + \langle\psi|\mathbf{A}|\dot{\psi}\rangle = \langle\psi|\mathbf{A}|\dot{\psi}\rangle + \text{H.C.} \\ &= \frac{i}{\hbar}\langle\psi(t)|[\mathbf{H}, \mathbf{A}]|\psi(t)\rangle,\end{aligned}\tag{75}$$

where we use a dot to indicate a time derivative. You should try to see for yourselves what this gives when \mathbf{A} is a familiar operator (e.g. \mathbf{x} , or \mathbf{p}) and for the Hamiltonian of a particle in scalar potential that we have used before.

Schrödinger vs. Heisenberg pictures

We have mentioned before that there is an infinite number of possible representations, depending on the basis of choice. The transition between one representation to another is a unitary transformation. So far we considered possible representations of the space at a given time t . But in principle we can make this more general. Suppose that we use the following unitary transformation that transforms vectors from our current representation to an alternative one

$$|\psi\rangle_H = e^{\frac{i}{\hbar}\mathbf{H}t}|\psi(t)\rangle.\tag{76}$$

The vectors with the index H are the transformation of the vectors that have no index to a new representation. The nice thing about this particular transformation is that it removes the time dependence (recall Eq. (71) for $|\psi(t)\rangle$). Yet, when we change our representation of vectors, the representation of operators changes correspondingly, and thus the operator \mathbf{A} changes to²⁴

²³ In shortened notation this relation can be written as $d\langle A\rangle/dt = i/\hbar\langle[\mathbf{H}, \mathbf{A}]\rangle$.

²⁴ An easy way to see this is to notice that the Heisenberg representation of $\mathbf{A}|\psi\rangle$

is $e^{i\mathbf{H}t}\mathbf{A}e^{-i\mathbf{H}t}e^{i\mathbf{H}t}|\psi\rangle = e^{i\mathbf{H}t}\mathbf{A}e^{-i\mathbf{H}t}|\psi\rangle_H$.

$$\mathbf{A}_H(t) = e^{\frac{i}{\hbar}\mathbf{H}t} \mathbf{A} e^{-\frac{i}{\hbar}\mathbf{H}t}. \quad (77)$$

Thus in this representation, operators become time dependent, whereas state-vectors loose their time dependence. By taking a derivative of (77) one can readily show that

$$\frac{d\mathbf{A}_H}{dt} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{A}_H] + e^{\frac{i}{\hbar}\mathbf{H}t} \frac{\partial \mathbf{A}}{\partial t} e^{-\frac{i}{\hbar}\mathbf{H}t}, \quad (78)$$

where the rightmost term accounts for the possibility that the original operator had a direct time dependence of some sort. The original representation, where the state-vectors account for the time-dependence, is commonly referred to as the Schrödinger picture, whereas the latter representation, with time invariant state-vectors is named after Heisenberg. Notice that time independent operators in the Schrödinger picture, that commute with the Hamiltonian (and in particular the Hamiltonian itself) do not change between the Heisenberg and Schrödinger pictures. Such operators, and consequently the results of their measurement, are independent of time.

It is easy to show that physical quantities, such as the moments of an observable are the same in both pictures. An interesting exercise is to use (78) to find the time evolution of the \mathbf{x} and \mathbf{p} operators in the Heisenberg picture for the Hamiltonian of a particle in a scalar potential $\mathbf{H} = \mathbf{p}^2/2m + V(\vec{\mathbf{r}})$. The outcome is quite familiar,

$$\begin{aligned} \dot{\mathbf{x}}_H(t) &= \frac{\mathbf{p}_{x,H}}{m} \\ \dot{\mathbf{p}}_{x,H}(t) &= -\frac{\partial V(\vec{\mathbf{r}}_H)}{\partial x_H}. \end{aligned} \quad (79)$$

To obtain this result note that for any three operators \mathbf{A} , \mathbf{B} and \mathbf{C} that satisfy the relation $[\mathbf{A}, \mathbf{B}] = \mathbf{C}$ in the Schrödinger picture, one also obtains $[\mathbf{A}_H, \mathbf{B}_H] = \mathbf{C}_H$. Proving this relation will be left as an exercise.

A very important example: The simple harmonic oscillator

In this section we will see the quantum description of one of the most important systems in physics, the simple harmonic oscillator (SHO). The importance of this system stems from several things. One is that the SHO describes any system that is characterized by small fluctuations around an equilibrium point. The other, and perhaps more relevant reason in the context of this course, is that the SHO is the basis for the quantum description of electromagnetic fields. As we will see later, a mode of light can be rigorously viewed as a SHO. The concept of a *photon* will follow directly from this property. Vibrational modes in solids are also modeled as a SHO, and thus the concept of a *phonon* also follows from the SHO analysis.

If we return for a moment to the simplest description of a SHO in classical mechanics, the evolution of x is determined by Newton's equation

$$\ddot{x} + \frac{K}{m}x = 0, \quad (80)$$

whose solution is of course $x(t) = A \cos(\omega t + \theta)$, where $\omega = \sqrt{K/m}$ and where the amplitude A and the phase θ are constants defined by the initial conditions. The potential energy of a classical SHO is $Kx^2/2 = m\omega^2 x^2/2$ and hence the Hamiltonian operator is given by²⁵

$$\mathbf{H} = \frac{\mathbf{p}_x^2}{2m} + \frac{1}{2}m\omega^2 \mathbf{x}^2. \quad (81)$$

As the constants m , ω and \hbar overburden our notation we will get rid of them by referring to normalized quantities

$$\begin{aligned} \mathbf{q} &= \sqrt{\frac{m\omega}{\hbar}} \mathbf{x} \\ \mathbf{p} &= \frac{1}{\sqrt{m\hbar\omega}} \mathbf{p}_x, \end{aligned} \quad (82)$$

²⁵ It is easily seen that the total energy of the SHO in a classical treatment is given by $E = m\omega^2 A^2/2$. Of course, quantum mechanically, the allowed energy levels can only be eigenvalues of the Hamiltonian.

such that²⁶

$$\mathbf{H} = \frac{\hbar\omega}{2}[\mathbf{p}^2 + \mathbf{q}^2] \quad (83)$$

with

$$[\mathbf{q}, \mathbf{p}] = i, \quad (84)$$

a relation that should be familiar to us from Section 2. It is now customary to define the operator

$$\mathbf{a} = \frac{1}{\sqrt{2}}(\mathbf{q} + i\mathbf{p}) \quad (85)$$

whose Hermitian conjugate is

$$\mathbf{a}^\dagger = \frac{1}{\sqrt{2}}(\mathbf{q} - i\mathbf{p}), \quad (86)$$

such that the reverse relations are

$$\mathbf{q} = \frac{1}{\sqrt{2}}(\mathbf{a} + \mathbf{a}^\dagger) \quad (87)$$

$$\mathbf{p} = \frac{-i}{\sqrt{2}}(\mathbf{a} - \mathbf{a}^\dagger). \quad (88)$$

Using the commutation relation between \mathbf{q} and \mathbf{p} , the commutation relation between \mathbf{a} and \mathbf{a}^\dagger can be readily shown to be

$$[\mathbf{a}, \mathbf{a}^\dagger] = 1, \quad (89)$$

and the Hamiltonian can be expressed as

$$\mathbf{H} = \hbar\omega \left(\mathbf{a}^\dagger \mathbf{a} + \frac{1}{2} \right). \quad (90)$$

²⁶ The q representation of this Hamiltonian is $\mathbf{H} = \frac{\hbar\omega}{2} \left[-\frac{\partial^2}{\partial q^2} + q^2 \right]$ and the p_x representation is identical to it with q and p interchanged.

We will now look for the allowed energy levels of the SHO, which are the eigenvalues of this Hamiltonian. Note that if we denote the eigenvalues of $\mathbf{a}^\dagger \mathbf{a}$ by the letter n , then the SHO's allowed energy levels will be equal to

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right). \quad (91)$$

The corresponding eigenstates are the same as the eigenstates of $\mathbf{a}^\dagger \mathbf{a}$ and we will denote them by $|n\rangle$. We will now see that the possible values of n are all the nonnegative integers. Thus the lowest possible energy of the SHO is not 0, as one might expect classically, but rather it is equal to $\frac{\hbar\omega}{2}$. This result has quite unusual consequences as it implies that the SHO oscillates even when it is not given any energy! But although this notion is highly nonintuitive, once we are aware of the uncertainty principle, it should not come as a total surprise. The ground-state oscillations of the SHO must exist in order to comply with the uncertainty principle, which implies that the coordinate and momentum of the SHO cannot be both simultaneously set to zero. The quantity $\hbar\omega$ is referred to as a "single *quantum* of energy" and so the total energy of the SHO is an integer number of quanta above the ground-state energy, which is $\frac{\hbar\omega}{2}$. We will now perform the calculation that shows that n is a nonnegative integer.

On the assumption that there exists a state $|n\rangle$ that is an eigenstate of the operator $\mathbf{a}^\dagger \mathbf{a}$ with eigenvalue n , one finds that $\mathbf{a}|n\rangle$ is also an eigenstate of $\mathbf{a}^\dagger \mathbf{a}$ with eigenvalue $n - 1$. To show this we write

$$\begin{aligned} (\mathbf{a}^\dagger \mathbf{a})(\mathbf{a}|n\rangle) &= (\mathbf{a}\mathbf{a}^\dagger - 1)(\mathbf{a}|n\rangle) \\ &= \mathbf{a}(\mathbf{a}^\dagger \mathbf{a} - 1)|n\rangle = (n - 1)\mathbf{a}|n\rangle. \end{aligned} \quad (92)$$

Thus we find that,

$$|n - 1\rangle = \frac{\mathbf{a}|n\rangle}{\sqrt{\langle n|\mathbf{a}^\dagger \mathbf{a}|n\rangle}} = \frac{1}{\sqrt{n}}\mathbf{a}|n\rangle, \quad (93)$$

a relation that can be readily generalized to

$$|n - m\rangle = \sqrt{\frac{(n - m)!}{n!}} \mathbf{a}^m |n\rangle. \quad (94)$$

The fact that $\mathbf{a}^\dagger \mathbf{a}$ is a nonnegative definite operator (i.e. its diagonal matrix elements are nonnegative in all representations), implies that the ladder of eigenvalues must stop at 0 so that all eigenvalues n must be integers.

In a similar manner, it can be shown that $\mathbf{a}^\dagger |n\rangle$ is also an eigenvalue of $\mathbf{a}^\dagger \mathbf{a}$ with eigenvalue $(n + 1)$, so that with proper normalization one finds that

$$\mathbf{a}^\dagger |n\rangle = \sqrt{n + 1} |n + 1\rangle. \quad (95)$$

and that

$$|n\rangle = \frac{(\mathbf{a}^\dagger)^n}{\sqrt{n!}} |0\rangle. \quad (96)$$

From now on, for obvious reasons, the operator $\mathbf{a}^\dagger \mathbf{a}$ will be referred to as the *number* operator, and the operators \mathbf{a} and \mathbf{a}^\dagger will be referred to as the *annihilation* and *creation* operators, respectively.

Having found the eigenvalues n to be integers, we will now look at the representation of the eigenstates $|n\rangle$ in the $|q\rangle$ and $|p\rangle$ bases. Using the fact that $\mathbf{a}|0\rangle = 0$ and writing \mathbf{a} explicitly, we have

$$\frac{1}{\sqrt{2}}(\mathbf{q} + i\mathbf{p})|0\rangle = 0. \quad (97)$$

Applying the bra $\langle q|$ from the left, yields

$$\frac{\partial \psi_0(q)}{\partial q} + q\psi_0(q) = 0 \quad (98)$$

an ordinary differential equation, whose (normalized) solution is

$$\psi_0(q) = \pi^{-1/4} e^{-q^2/2}. \quad (99)$$

The complementary function $\tilde{\psi}_0(p)$ is obtained by fourier transforming $\psi_0(q)$, or by using $\mathbf{q} = i\partial/\partial p$ in the k representation

$$\tilde{\psi}_0(p) = \pi^{-1/4} e^{-p^2/2}. \quad (100)$$

Recall that $|\psi_0(q)|^2$ and $|\tilde{\psi}_0(p)|^2$ are probability density functions. They are both Gaussian and they indicate an uncertainty of $1/\sqrt{2}$ in both k and q (as you have seen in your homework assignments), such that

$$\Delta q \Delta p = \frac{1}{2}. \quad (101)$$

Returning to the original parameters \mathbf{x} and \mathbf{p}_x , this relation translates into

$$\Delta \mathbf{x} \Delta \mathbf{p}_x = \frac{\hbar}{2}, \quad (102)$$

which, as we have seen in (68), is the smallest allowed uncertainty value. The ground-state of the SHO is thus said to be a minimum uncertainty state. The eigenstates representing higher energy levels of the the SHO can be obtained from the above results by using relation (95). For example,

$$\psi_1(q) = \frac{1}{\sqrt{2}} \left(q\psi_0(q) - \frac{\partial\psi_0(q)}{\partial q} \right) = \frac{\sqrt{2}}{\pi^{1/4}} q e^{-q^2/2}. \quad (103)$$

A general expression is obtained by successive application of \mathbf{a}^\dagger to the ladder of eigenstates,

$$\psi_n(q) = \pi^{-1/4} \frac{1}{\sqrt{2^n n!}} e^{-q^2/2} H_n(q), \quad (104)$$

where $H_n(q)$ is the n 'th order Hermite polynomial. These polynomials are given by the application²⁷ of the operation $(2q - d/dq)^n$ to 1. The first few polynomials are given by

²⁷ The Hermite polynomials can also be defined through the relation

$$H_n(q) = (-1)^n e^{q^2} d^n/dq^n e^{-q^2}.$$

$$\begin{aligned}
H_0(q) &= 1 \\
H_1(q) &= 2q \\
H_2(q) &= 4q^2 - 2 \\
H_3(q) &= 8q^3 - 12q.
\end{aligned} \tag{105}$$

They have various properties, the most notable of which is that they are of alternating parity and that the n 'th Hermite polynomial has exactly n real zeros. It is worth plotting the first few functions $\psi_n(q)$. This is the figure on page 502 of ref [1]. It is interesting to notice that at first sight the quan-

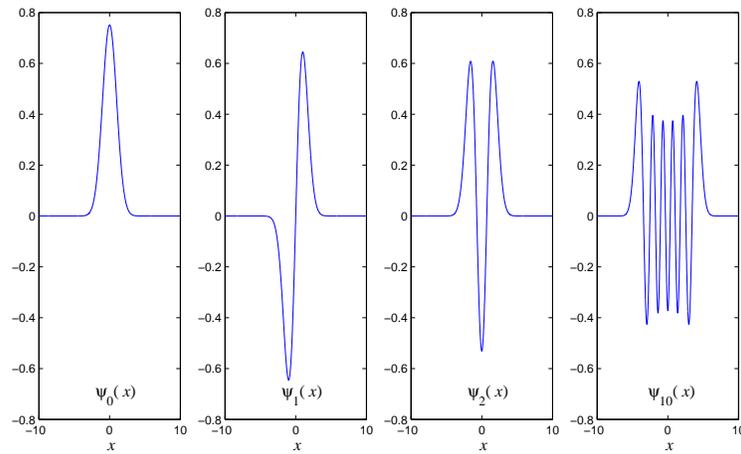


Fig. 1. Wave functions associated with the first three levels of the SHO. After [1] p. 502.

tum solution of the SHO bears no resemblance at all to the classical picture that we considered earlier. Let us consider the mean values of some measurable quantities. First, note that since the wavefunctions $\psi_n(q) = \langle q|n\rangle$ and $\tilde{\psi}_n(p) = \langle p|n\rangle$ have definite parity, the probability distributions are symmetric functions and thus both $\langle \mathbf{q} \rangle$ and $\langle \mathbf{p} \rangle$ are 0. Finding the standard deviation of \mathbf{q} and \mathbf{p} is a homework exercise. It is most conveniently done by using re-

lations (87) and (88)

$$\Delta\mathbf{q} = \Delta\mathbf{p} = \left(n + \frac{1}{2}\right)^{1/2} \quad (106)$$

and so

$$\Delta\mathbf{q}\Delta\mathbf{p} = n + \frac{1}{2}$$

giving minimum uncertainty only in the ground-state²⁸, i.e. when $n = 0$. It is instructive to return at this stage to the true coordinate and momentum $\mathbf{x} = \sqrt{\frac{\hbar}{m\omega}}\mathbf{q}$ and $\mathbf{p}_x = \sqrt{\hbar\omega m}\mathbf{p}$. Thus we find that

$$\Delta\mathbf{x} = \sqrt{\frac{E_n}{m\omega^2}} \quad \text{and} \quad \Delta\mathbf{p}_x = \sqrt{mE_n}.$$

Classically, the amplitude of oscillation A is related to the energy through $A = \sqrt{\frac{2E}{m\omega^2}}$ and the amplitude of the momentum $P_M = m\omega A = \sqrt{2mE}$, and so we find that the quantum uncertainties are

$$\Delta\mathbf{x} = \frac{1}{\sqrt{2}}A \quad \text{and} \quad \Delta\mathbf{p}_x = \frac{1}{\sqrt{2}}P_M.$$

Interestingly, the uncertainties do not scale with n and, against what perhaps could be naively expected, the relative uncertainty does not reduce when n is very large. This is only another manifestation of the fact that the energy states of the SHO are highly non-classical. As we will see in later stages, the SHOs that we usually see in nature are not in a fixed energy state. Their behavior resembles the behaviour of quantum SHOs in a so-called *coherent* state, which is a particular superposition of energy states. As we know, an energy state in quantum mechanics has to be stationary, implying in particular that the mean values of observables are independent of time. Therefore we

²⁸ Had we not normalized the coordinate and momentum to have a commutator of i , the product of uncertainties relation would be $\hbar\left(n + \frac{1}{2}\right)$.

should not be too surprised that the energy states do not resemble the classical solution which is clearly non-stationary. One could however propose a certain analogy that would help visualize the energy states in classical terms. Imagine the classical solution

$$X(t) = A \cos(\omega t + \theta)$$

$$P_x(t) = -P_M \sin(\omega t + \theta)$$

that would correspond to the case where θ is a random number that is uniformly distributed in the interval $[0, 2\pi]$. It is quite obvious that such distribution of θ would make the random processes $X(t)$ and $P_x(t)$ stationary. With this assumption the mean values of X and P_x are 0 and their standard deviations are exactly as they are in the quantum mechanical case when the SHO is in a well defined energy state.

Coherent states of the SHO

Let us now become familiar with the concept of coherent states. A coherent state is what we would often hear referred to as "classical" state, because as we shall see, a SHO that is in a coherent state behaves very much like what we are used to classically. Why do SHOs prefer to behave as if they are in a coherent state when they reach the classical limit is a question that will be addressed in your homework and I will refrain from addressing it here.²⁹

The formal definition of a coherent state is that it is an eigenstate of the annihilation operator \mathbf{a} . We will denote those states by $|\alpha\rangle$, where α designates the corresponding eigenvalue, as usual. Namely

$$\mathbf{a}|\alpha\rangle = \alpha|\alpha\rangle, \tag{107}$$

where, since \mathbf{a} is not a Hermitian operator, its eigenvalues α may be complex. We will now find out some properties of the coherent state. First we wish to expand the coherent state in terms of the energy states (or number-states)

²⁹ See Exam question 1 in 2014.

that we are already familiar with. Owing to the completeness of the energy state basis, we may write

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n(\alpha)|n\rangle.$$

Substituting this expression into (107) and using the fact that $\mathbf{a}|n\rangle = \sqrt{n}|n-1\rangle$ we find that

$$C_n = \frac{\alpha}{\sqrt{n}}C_{n-1}$$

an equation whose solution is

$$C_n = C_0 \frac{\alpha^n}{\sqrt{n!}}$$

The condition $\sum |C_n|^2 = 1$ leads to $C_0 = \exp(-|\alpha|^2/2)$ (to within an immaterial phase) and so we find that

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (108)$$

We may thus find the distribution of energy for a SHO that is in a coherent state. More relevantly, the probability that the measured energy is n quanta above the vacuum, is

$$P(n) = |C_n(\alpha)|^2 = e^{-|\alpha|^2} \frac{(|\alpha|^2)^n}{n!},$$

which is immediately recognized as Poisson's distribution. Knowing the properties of the Poisson distribution we may write

$$\langle n \rangle = |\alpha|^2 \quad (109)$$

$$\Delta n^2 = \langle n^2 \rangle - \langle n \rangle^2 = |\alpha|^2. \quad (110)$$

These are of course very important properties, and among other things they tell us that the relative energy uncertainty is

$$\frac{\Delta n}{\langle n \rangle} \longrightarrow 0$$

when $\langle n \rangle$ is large. This property is certainly in agreement with the classically observed reality.

Let us now consider the mean values of \mathbf{q} and \mathbf{p} for an SHO which is in a coherent state

$$\langle \mathbf{q} \rangle = \frac{1}{\sqrt{2}} \langle \alpha | \mathbf{a} + \mathbf{a}^\dagger | \alpha \rangle = \frac{\alpha + \alpha^*}{\sqrt{2}} \quad (111)$$

$$\langle \mathbf{p} \rangle = \frac{-i}{\sqrt{2}} \langle \alpha | \mathbf{a} - \mathbf{a}^\dagger | \alpha \rangle = \frac{\alpha - \alpha^*}{i\sqrt{2}} \quad (112)$$

If we express the complex number α as

$$\alpha = \frac{\bar{q} + i\bar{p}}{\sqrt{2}}$$

then \bar{q} and \bar{p} have the physical meaning of the mean coordinate and momentum. Let us now see how a coherent state $|\alpha\rangle$ evolves in time. The time evolution of all state-vectors in the case of a time independent Hamiltonian is

$$|\alpha(t)\rangle = e^{-\frac{i}{\hbar}\mathbf{H}t}|\alpha(0)\rangle.$$

This is most easily manipulated into a more explicit form by representing $|\alpha\rangle$ in terms of energy states

$$|\alpha(t)\rangle = e^{-i\omega t/2} e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{(\alpha e^{-i\omega t})^n}{\sqrt{n!}} |n\rangle = e^{-i\omega t/2} |e^{-i\omega t} \alpha(0)\rangle. \quad (113)$$

In other words, a coherent state remains a coherent state after evolution but its eigenvalue changes from α_0 to $\alpha_0 e^{-i\omega t}$. Evidently, this implies that

$$\begin{aligned} \langle q(t) \rangle &= |\alpha_0| \cos(\omega t + \theta) \\ \langle p(t) \rangle &= -|\alpha_0| \sin(\omega t + \theta), \end{aligned}$$

which is the same as the classical relations with θ being the phase of α_0 . Let us now consider the fluctuations Δq and Δp .

$$\begin{aligned}\langle \mathbf{q}^2 \rangle &= \frac{1}{2} \langle \alpha | (\mathbf{a} + \mathbf{a}^\dagger)^2 | \alpha \rangle = \frac{1}{2} \langle \alpha | \mathbf{a}^2 + (\mathbf{a}^\dagger)^2 + 2\mathbf{a}^\dagger \mathbf{a} + 1 | \alpha \rangle \\ &= \frac{|\alpha + \alpha^*|^2 + 1}{2}.\end{aligned}\quad (114)$$

Consequently,

$$\Delta q = \sqrt{\langle \mathbf{q}^2 \rangle - \langle \mathbf{q} \rangle^2} = \frac{1}{\sqrt{2}} \quad (115)$$

Similarly, and this will be left as a homework exercise, we find that

$$\Delta p = \frac{1}{\sqrt{2}}. \quad (116)$$

The product of uncertainties

$$\Delta q \Delta p = \frac{1}{2},$$

which is the smallest possible value permitted by the principle of uncertainty. In other words, we find that the coherent state of the SHO has the smallest possible position and momentum uncertainties, while the mean values of these quantities correspond exactly to what we observe in a classical SHO. This is the reason for why these states are often called "classical". We can now evaluate the wave functions in the q and p representations. The equation

$$\mathbf{a}|\alpha\rangle = \alpha|\alpha\rangle$$

can be written in the q representation by applying $\langle q|$ from the left and expressing \mathbf{a} in terms of the \mathbf{q} and \mathbf{p} operators

$$\frac{1}{\sqrt{2}} \left(q + \frac{\partial}{\partial q} \right) \psi_\alpha(q) = \frac{\bar{q} + i\bar{p}}{\sqrt{2}} \psi_\alpha(q),$$

where on the right-hand side I replaced α by $(\bar{q} + i\bar{p})/\sqrt{2}$ according to our earlier convention. This equation can be rewritten as

$$\frac{\partial \psi_\alpha}{\partial q_c} + (q_c - i\bar{p})\psi_\alpha = 0,$$

with $q_c \equiv q - \bar{q}$ being the distance between the coordinate q and the average q value (the index "c" stands for centered). The solution is quite straightforward³⁰

$$\psi_\alpha(q) = \pi^{-1/4} e^{-\frac{(q-\bar{q})^2}{2}} e^{i\bar{p}(q-\bar{q})} \quad (117)$$

and we find that the probability distribution density of q is Gaussian

$$|\psi_\alpha(q)|^2 = \frac{1}{\sqrt{\pi}} e^{-(q-\bar{q})^2}$$

with mean \bar{q} and standard deviation of $1/\sqrt{2}$. It can be readily shown (homework) that the exact same distribution is satisfied by the complementary wave function $\tilde{\psi}_\alpha(p)$. We may thus relate to a SHO in a coherent state as if it oscillates classically with a normalized uncertainty of $1/\sqrt{2}$. The uncertainties in real units would be $\Delta x = \sqrt{\frac{\hbar}{2m\omega}}$ and $\Delta p = \sqrt{\frac{m\hbar\omega}{2}}$ so that the uncertainty product is $\hbar/2$.

There are two more notions regarding coherent states that we should become familiar with at this stage. One is that, being eigenvectors of a non-Hermitian operator, different coherent states are not orthogonal to each other. We may show that by using the energy state expansion

$$\begin{aligned} \langle \alpha | \beta \rangle &= e^{-\frac{|\alpha|^2 + |\beta|^2}{2}} \sum_{n,m} \frac{(\alpha^*)^n \beta^m}{\sqrt{n!m!}} \langle n | m \rangle \\ &= e^{-\frac{|\alpha|^2 + |\beta|^2}{2}} \sum_n \frac{(\alpha^* \beta)^n}{n!} = e^{-\frac{|\alpha-\beta|^2}{2}} e^{\frac{\alpha^* \beta - \alpha \beta^*}{2}} \end{aligned} \quad (118)$$

so that $|\langle \alpha | \beta \rangle|^2 = e^{-|\alpha-\beta|^2}$. What this means is that although different coherent states are not generally orthogonal, they become nearly orthogonal when the eigenvalues are separated enough. Another useful property of coherent states is that they do form a basis for the space of functions that

³⁰ The term $\bar{p}(q - \bar{q})$ in the second exponent can be safely replaced with $\bar{p}q$, with the difference $\bar{p}\bar{q}$ only contributing to the phase, which is immaterial.

may describe the state of the SHO. In fact, the basis that they form is over complete (i.e. there is more than one way to represent every state-vector). To see that, we must show that the identity operator can be resolved into a sum of projectors onto coherent states. Indeed it can be shown that

$$\frac{1}{\pi} \int d^2\alpha |\alpha\rangle\langle\alpha| = \mathbf{1}, \quad (119)$$

Implying, as usual, that any $|\psi\rangle$ can be written as $|\psi\rangle = \frac{1}{\pi} \int d^2\alpha \langle\alpha|\psi\rangle |\alpha\rangle$. The integration is made with respect to $d^2\alpha = d\alpha_R d\alpha_I$ with α_R and α_I denoting the real and imaginary parts of α . When using polar coordinates, where $\alpha = r e^{i\theta}$, integration can be performed with respect to $r dr d\theta$ with r being between 0 and infinity and with θ being in the range of 0 to 2π .

The density operator

So far we have been saying that a quantum system is defined in terms of its quantum state-vector $|\psi\rangle$. While this remains true, in many cases we must account for situations in which the true quantum state of a system is not known to us. Instead, we only know the probability of the system to be in various states. For example, we may be interested in performing tests on a particle that comes from an ensemble. We know that in that ensemble, some particles are in quantum state $|\psi_1\rangle$, some are in $|\psi_2\rangle$ and some are in $|\psi_3\rangle$. Suppose that we also know that the probabilities that the particle is in each of these states (namely we know the fraction of particles in each state) are p_1 , p_2 and p_3 , respectively, where $p_1 + p_2 + p_3 = 1$. In this case one cannot prescribe a single state vector to the system, because one does not know what it is. The density matrix is something that was invented in order to resolve this issue in the most elegant way. Recall that the reason that one may want to know the state of the system is in order to predict the outcomes of measurements, or more accurately, in order to predict the probabilities of various outcomes. For example, we may want to know what is the probability of receiving an outcome b in a measurement of some operator \mathbf{B} . The only possible outcomes

are eigenvalues of \mathbf{B} and so, if we know that the state vector is either $|\psi_1\rangle$, $|\psi_2\rangle$, or $|\psi_3\rangle$, with probabilities p_1, p_2, p_3 , respectively, then the probability of measuring b is given by

$$p(b) = p_1 \langle b|\psi_1\rangle^2 + p_2 \langle b|\psi_2\rangle^2 + p_3 \langle b|\psi_3\rangle^2 = \langle b|\rho|b\rangle, \quad (120)$$

where $|b\rangle$ is the eigenstate of \mathbf{B} that corresponds to the eigenvalue b , and³¹

$$\rho = p_1 |\psi_1\rangle\langle\psi_1| + p_2 |\psi_2\rangle\langle\psi_2| + p_3 |\psi_3\rangle\langle\psi_3| \quad (121)$$

is what we call the *density matrix*. Equations (120) and (121) imply that ρ is a Hermitian semi-definite positive operator, as its diagonal elements in any representation are probabilities. When we do not limit our uncertainty to three states, the density matrix assumes the form³²

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|. \quad (122)$$

Equation (121) can also be written as

$$p(b) = \text{Trace}\{|b\rangle\langle b|\rho\}, \quad (123)$$

³¹ Notice that the vectors $|\psi_j\rangle$ need not be orthogonal to each other.

³² An important notion concerning the density matrix ρ is that, excluding the cases when it represents a pure state, its construction is not unique. Consider an example (exercise) in two-dimensions where $|\psi_1\rangle$ is represented by $2^{-1/2}(1, 1)^t$, $|\psi_2\rangle$ is represented by $(1, 0)^t$ and $|\psi_3\rangle$ is represented by $(0, 1)^t$. Assume also that the respective probabilities are $p_1 = 0.25$, $p_2 = 0.5$ and $p_3 = 0.25$. The density matrix $\rho = p_1 |\psi_1\rangle\langle\psi_1| + p_2 |\psi_2\rangle\langle\psi_2| + p_3 |\psi_3\rangle\langle\psi_3|$ is represented by

$$\frac{1}{8} \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}.$$

The same density matrix follows when there are two other states; $|\phi_1\rangle$ represented by $(0.3827, -0.9239)^t$ and $|\phi_2\rangle$ represented by $(0.9239, 0.3827)^t$ with the respective probabilities $p_1 = 0.3232$ and $p_2 = 0.6768$. There is no way to distinguish between the two mixtures.

which is the trace of ρ multiplied by the projection operator into state $|b\rangle$. Assume now that we want to calculate the mean value of some generic operator \mathbf{A} . Using the usual method, combined with the principles of stochastic analysis, we find that³³

$$\langle \mathbf{A} \rangle = \sum_j p_j \langle \psi_j | \mathbf{A} | \psi_j \rangle = \text{Trace}\{\mathbf{A}\rho\}, \quad (124)$$

which is also equal to $\text{Trace}\{\rho\mathbf{A}\}$ because of the invariance of the trace to cyclic permutations. Notice that Eq. (120) is a special case of Eq. (124), as can be seen when \mathbf{A} is replaced by $|b\rangle\langle b|$. If \mathbf{A} is set to be the identity operator we find another notable property of ρ stating that

$$\text{Trace}\{\mathbf{1}\rho\} = \text{Trace}\{\rho\} = 1. \quad (125)$$

If the system is in a pure state (i.e. its state-vector is known with probability 1), then the density operator becomes a simple projector

$$\rho = |\psi\rangle\langle\psi|. \quad (126)$$

In that case $\rho^2 = \rho$, a property that can be used in order to tell whether a given system is in a pure state, or not³⁴.

³³ To see that this is the case, note that

$$\begin{aligned} \text{Trace}\left\{\sum_j p_j |\psi_j\rangle\langle\psi_j| \mathbf{A}\right\} &= \sum_k \sum_j p_j \langle u_k | \psi_j \rangle \langle \psi_j | \mathbf{A} | u_k \rangle \\ &= \sum_j p_j \langle \psi_j | \mathbf{A} \left(\sum_k |u_k\rangle\langle u_k| \right) | \psi_j \rangle \\ &= \sum_j p_j \langle \psi_j | \mathbf{A} | \psi_j \rangle. \end{aligned}$$

³⁴ For example, if the density matrix is written as $\rho = 0.5|\phi_1\rangle\langle\phi_1| + 0.5|\phi_2\rangle\langle\phi_2| + 0.5|\phi_1\rangle\langle\phi_2| + 0.5|\phi_2\rangle\langle\phi_1|$, telling whether it represents a pure or, mixed state

Part B - The interaction of electromagnetic fields with matter, a semiclassical description

Consider an atom that is characterized by a Hamiltonian \mathbf{H}_A , whose eigenvalues (i.e. allowed energy levels) are E_n , with n being a nonnegative integer. The lowest level, E_0 is often referred to as the ground state. An eigenstate $|E_n\rangle$ corresponds to each of the energy levels. In case of degeneracy, there will be more than one state for each energy level, but for the time being we shall ignore this possibility. As we know, any possible state of the system can be spanned by the eigenfunctions $|E_n\rangle$ at any given instant. In the unperturbed case, we know that the state of the system is

$$|\psi(t)\rangle = \sum_n C_n e^{-i\omega_n t} |E_n\rangle \quad (127)$$

because the states $|E_n\rangle$ are the eigenstates of the system's Hamiltonian. But even if the Hamiltonian differs from \mathbf{H}_A , the state of the system at any given moment can still be spanned in terms of the eigenstates of \mathbf{H}_A , just because it is a Hermitian operator. In particular, when we consider an atom that is perturbed by an electromagnetic field, the state at any given instant can be expressed as

$$|\psi\rangle = \sum_n C_n(t) e^{-i\omega_n t} |E_n\rangle. \quad (128)$$

This is a general expression since the coefficients $C_n(t)$ are time dependent. Normalization of $|\psi\rangle$ implies that the sum of their squares at any given moment must equal unity. Our choice not to absorb the exponential factors into the definition of $C_n(t)$ is just a matter of future convenience.

The Hamiltonian in the presence of an electromagnetic field can be written as

is not immediately trivial without using the test $\rho^2 = \rho$. In reality, of course one could define $|\psi\rangle = 2^{-1/2}[|\phi_1\rangle + |\phi_2\rangle]$, in which case it becomes obvious that $\rho = |\psi\rangle\langle\psi|$ represents a pure state.

$$\mathbf{H} = \mathbf{H}_A + \mathbf{H}_I \quad (129)$$

where, if we assume an atom with a nucleus charge Ze and Z electrons of charge $-e$ that surround it,

$$\mathbf{H}_I = \vec{\mathbf{D}} \cdot \vec{E}(t) \quad (130)$$

is the contribution of the interaction to the Hamiltonian. The vector $\vec{E}(t)$ denotes the electric field at the position of the atom, and $\vec{\mathbf{D}}$ is the dipole moment vector

$$\vec{\mathbf{D}} = e \sum_{j=1}^Z \vec{\mathbf{r}}_j \quad (131)$$

with $\vec{\mathbf{r}}_j$ being the operator that corresponds to the position of the j 'th electron in the atom and Z is the total number of electrons. This form of the interaction Hamiltonian is known as the dipole approximation. The complete form of the interaction Hamiltonian is more complex, but under realistic conditions, for light whose wavelength is much longer than characteristic atomic radii (which are of the order of Angstroms), this approximation accurately applies. The state $|\psi\rangle$ satisfies the Schrödinger equation

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \mathbf{H}_A |\psi\rangle + \mathbf{H}_I |\psi\rangle. \quad (132)$$

If we substitute $|\psi\rangle$ from (128) and multiply by $\langle E_k|$ from the left, we obtain

$$\dot{C}_k = -\frac{i}{\hbar} \sum_n C_n(t) \langle E_k | \mathbf{H}_I | E_n \rangle e^{-i\omega_{nk}t}, \quad (133)$$

where we use a "dot" above the parameter in order to denote a time derivative and where $\omega_{nk} = \omega_n - \omega_k$. This is a set of coupled differential equations, which can be solved for a given initial condition to yield the population of all states. Unfortunately, a slight inconvenience follows from the fact that this set of equations cannot be solved analytically, even in the simplest case

where only two levels are involved (i.e. when ignoring all coefficients except C_1 and C_2). One can either approach the solution of (133) numerically, or resort to approximations. We start by solving the problem in the case where \mathbf{H}_I is small in a sense that will be defined shortly.

We will assume that at $t = 0$ the atom was in a particular energy state $|E_i\rangle$ (i stands for "initial") and we wish to find out the probability that we find the atom in state $|E_f\rangle$ (f stands for final) a short time t afterwards. At $t = 0$ we have $C_i = 1$ and C_f ($f \neq i$) are all 0. If t is small enough this situation cannot change by much and therefore we may approximate

$$C_f(t) = -\frac{i}{\hbar} \int_0^t dt' H_{fi}(t') e^{-i\omega_{fi}t'}. \quad (134)$$

where

$$H_{fi}(t) = \langle E_f | \mathbf{H}_I | E_i \rangle. \quad (135)$$

If we assume that the electric field is given by $E(t) = \vec{E}_0 \cos(\omega_0 t)$, then³⁵

$$H_{fi}(t) = \langle E_f | \vec{\mathbf{D}} \cdot \vec{E}_0 | E_i \rangle \cos(\omega_0 t) = (\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0) \cos(\omega_0 t) \quad (136)$$

with the coefficient $(\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0) = \langle E_f | \vec{\mathbf{D}} \cdot \vec{E}_0 | E_i \rangle$ defined through the above relation. Using this in (134) we obtain

$$C_f(t) = -\frac{i(\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0)}{2\hbar} t \left[e^{-\frac{i}{2}(\omega_{fi}-\omega_0)t} \text{sinc}\left(\frac{\omega_{fi}-\omega_0}{2}t\right) + e^{-\frac{i}{2}(\omega_{fi}+\omega_0)t} \text{sinc}\left(\frac{\omega_{fi}+\omega_0}{2}t\right) \right], \quad (137)$$

where $\text{sinc}(x) = \sin(x)/x$. Clearly, this approximation is valid only as long as the deviation of C_i from unity is negligible, implying that $|C_k|^2$ ($k \neq i$)

³⁵ Avoid confusing the terms \vec{E}_0 , E_0 and $|E_0\rangle$. The vector \vec{E}_0 is the amplitude of the electric field, whereas the scalar E_0 is the energy eigensatete corresponding to the eigenstate $|E_0\rangle$.

is much smaller than 1. A more rigorous procedure explaining the approximation will be reviewed in your homework. Notice that the contribution of the term in the square brackets is large only when $\omega_{fi} \simeq \pm\omega_0$ (we treat the field frequency ω_0 as a nonnegative number). This result suggests that a transition is likely only if the energy of the final state E_f is higher, or lower by $\hbar\omega_0$ from the initial energy E_i . This means that the electromagnetic field can either excite the atom to a higher level, or cause it to descend to a lower one. The first option relates to the absorption of radiation by the atom, the second correspond to the case of *stimulated emission*, which is the physical process that allows the operation of lasers and optical amplifiers. We will discuss this issue further at a later time. It is obvious that for any given final level E_f , either the first, or the second term dominates the coefficient and thus we may express the probability of transition to the final level $|E_f\rangle$ (be it through absorption, or emission) as

$$|C_f(t)|^2 = \frac{|\vec{\mathbf{D}}_{fi} \cdot \vec{\mathbf{E}}_0|^2}{4\hbar^2} t^2 \text{sinc}^2 \left(\frac{|\omega_{fi}| - \omega_0}{2} t \right). \quad (138)$$

When the light frequency is exactly at the atomic resonance (i.e. $\omega_0 = \omega_{fi}$), the probability of transition is proportional to t^2 . This is of course correct only for small values of t as long as $|C_f(t)|^2 \ll 1$. In order to find the evolution of $|C_f(t)|^2$ for larger t values one obviously needs to resort to a numerical solution of the set (133). For a fixed t , the dependence on the detuning from resonance is given by the square of the sinc function with zeros at $\omega_{fi} - \omega_0 = 2\pi m/t$ with m being integer.

Expression (138) can be quite useful when the source of electro-magnetic field is a laser so that the linewidth of the radiation is very small. In many cases, the source of light is thermal in nature, meaning that it occupies a range of frequencies in the neighborhood of ω_{fi} . The different frequencies are not coherent with each other and so to find out the total effect, we integrate

the probability of transition³⁶ over ω_0 . If we assume that the bandwidth $\Delta\omega$ of the incident radiation is much greater than t^{-1} , then³⁷

$$t \operatorname{sinc}^2 \left(\frac{\omega_{fi} - \omega_0}{2} t \right) \rightarrow 2\pi \delta(\omega_{fi} - \omega_0)$$

and the total probability of transition to the final state $|E_f\rangle$ is given by the expression

$$\int d\omega_0 |C_f(t)|^2 = \frac{\pi |\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0(\omega_{fi})|^2}{2\hbar^2} t, \quad (139)$$

where I have explicitly written the dependence of E_0 on frequency. Thus the rate of transition is simply

$$W_{i \rightarrow f} = \frac{\pi |\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0(\omega_{fi})|^2}{2\hbar^2}. \quad (140)$$

A different scenario, but with the same result follows when monochromatic light leads to a transition of a system from a discrete quantum state $|E_i\rangle$ to a continuum of final states. In this case, ω_0 is constant, but ω_{fi} assumes a

³⁶ Notice that had we used a coherent source of radiation, we would have to integrate the probability amplitude $C_f(t)$ (as opposed to the probability $|C_f(t)|^2$) over ω_0 .

³⁷ The Dirac delta function can be approximated as a limit of many things. For example $\lim_{a \rightarrow \infty} [a \operatorname{sinc}(\omega a/2)] = 2\pi \delta(\omega)$. The meaning of this is very simple. Integration of $a \operatorname{sinc}(\omega a/2)$ with respect to ω from a value much smaller than $-a^{-1}$ to a value much greater than a^{-1} (recall that $2 \times 2\pi/a$ is the width of the sinc) is 2π . When a is large enough for the resolution of frequencies lower than a^{-1} to become irrelevant, the function $a \operatorname{sinc}(\omega a/2)$ can be replaced by a $2\pi \delta(\omega)$. Similarly the integral of $a^2 \operatorname{sinc}^2(\omega a/2)$ over ω is $2\pi a$ (as can be readily deduced from the fact that the FT of $a \operatorname{sinc}(\omega a/2)$ is a window of width a and a height of 2π and by applying Parseval's theorem). Therefore $a \operatorname{sinc}^2(\omega a/2)$ also becomes $2\pi \delta(\omega)$ when a is large enough.

continuum of values. The result is the same, although it is customary to add a summation over all final states for which $\omega_{fi} = \omega_0$

$$W_{i \rightarrow f} = \sum_f \frac{\pi |\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0(\omega_{fi})|^2}{2\hbar^2}. \quad (141)$$

This relation is famously known as Fermi's Golden rule³⁸. This whole thing is also related to the famous Einstein B coefficient through

$$B_{if} w(\omega_{fi}) = \frac{\pi \langle |\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0|^2 \rangle_\theta}{2\hbar^2}, \quad (142)$$

where this time $\langle \cdot \rangle_\theta$ indicates averaging with respect to the possible orientations of the Atoms dipole moments in space (assuming as in Einstein's theory a cloud of gas with atoms at random orientations) and where $w(\omega)$ is the power spectral density of the radiation. We should note that the Einstein "A" coefficient, which is related to spontaneous emission, cannot be extracted from a semiclassical model.

Solution of a two level system - Rabi oscillations

Let us consider the case of a two-level system. In the case of the atom this is what effectively happens when there are only two states whose energy difference divided by \hbar falls within the bandwidth of the electromagnetic

³⁸ This rule is often written as

$$W_{i \rightarrow f} = \int d\omega_{fi} \sum_f \frac{\pi |\vec{\mathbf{D}}_{fi} \cdot \vec{E}_0(\omega_0)|^2}{2\hbar^2} \delta(\omega_0 - \omega_{fi}),$$

and sometimes, mistakenly in my opinion, integration over ω_{fi} is not written explicitly (e.g. see Eq. 4.41 of [3]) In any case the FGR is meaningful only when t is large enough so that the bandwidth of final states is much greater than t^{-1} . In the same time t is small enough for the overall probability of transition to be small.

field. We will denote these levels by the indices 1 and 2 and assume that the energy of level 2 is higher than that of level one, so that $\omega_{21} > 0$. Equations (133) can then be written as³⁹

$$\dot{C}_1(t) = -\frac{i}{2\hbar}(\vec{D}_{12} \cdot \vec{E}_0)C_2(t)e^{i(\omega_0 - \omega_{21})t} \quad (143)$$

$$\dot{C}_2(t) = -\frac{i}{2\hbar}(\vec{D}_{21} \cdot \vec{E}_0)C_1(t)e^{-i(\omega_0 - \omega_{21})t}, \quad (144)$$

where I have used the fact that $(\vec{D}_{11} \cdot \vec{E}_0) = (\vec{D}_{22} \cdot \vec{E}_0) = 0$, as follows from the fact that the wave-functions of atomic bound-states have a well defined parity⁴⁰ (whereas the dipole moment is odd as indicated by Eq. (131)). Since the phase of \vec{D}_{12} will end up being immaterial to the final results⁴¹, I will just set it to zero in order to avoid the cluttering of the derivation, in which case it follows from the Hermiticity of $\vec{\mathbf{D}}$ that $\vec{D}_{12} = \vec{D}_{21}$. Equations (143)-(144) were written with the help of the rotating wave approximation that we have used before. Namely, the terms oscillating at $\omega_0 + \omega_{21}$ were neglected. These equations are exactly solvable, and subject to the initial conditions $C_1(0) = 1$ and $C_2(0) = 0$, their solution can be expressed as

³⁹ Notice that the time dependence of the field is written as $\vec{E}_0 \exp(i\omega_0 t)$ in Eq. (143) and as $\vec{E}_0 \exp(-i\omega_0 t)$ in Eq. (144). This is the manifestation of the rotating wave approximation that we have discussed earlier.

⁴⁰ For atoms, spherical symmetry implies that the Hamiltonian \mathbf{H}_A is symmetric with respect to sign inversion of the position operators $(\mathbf{x}, \mathbf{y}, \mathbf{z})$. This implies that \mathbf{H}_A and the inversion operator (that replaces $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ with $(-\mathbf{x}, -\mathbf{y}, -\mathbf{z})$) have joint eigenstates. Thus the wave functions $\psi_n(\vec{r}) = \langle \vec{r} | E_n \rangle$ are either symmetric or antisymmetric. A nonzero dipole moment exists only when one of the eigenstates is symmetric and one is anti-symmetric.

⁴¹ The Hermiticity of $\vec{\mathbf{D}}$ implies that $\vec{D}_{12} = \vec{D}_{21}^*$. We may represent $\vec{D}_{12} \cdot \vec{E}_0 = r \exp(i\phi)$ and $\vec{D}_{21} \cdot \vec{E}_0 = r \exp(-i\phi)$, where r and ϕ are real numbers. Examination of Eqs. (143-144) then immediately reveals that including a nonzero ϕ is equivalent to changing the time origin t by a little bit.

$$C_1(t) = \frac{1}{\Omega_R} \left[\Omega_R \cos\left(\frac{\Omega_R t}{2}\right) - i\Delta\omega \sin\left(\frac{\Omega_R t}{2}\right) \right] e^{i\Delta\omega t/2} \quad (145)$$

$$C_2(t) = -\frac{i}{\hbar\Omega_R} (\vec{D}_{21} \cdot \vec{E}_0) \sin\left(\frac{\Omega_R t}{2}\right) e^{-i\Delta\omega t/2} \quad (146)$$

where $\Delta\omega = \omega_0 - \omega_{21}$ is the detuning,

$$\Omega_R = \sqrt{\chi_R^2 + \Delta\omega^2}, \quad (147)$$

and

$$\chi_R = \frac{|\vec{D}_{21} \cdot \vec{E}_0|}{\hbar}. \quad (148)$$

It is common to refer to χ_R and to Ω_R as the *Rabi frequency* and the *generalized Rabi frequency*, respectively. It can be readily verified that eqs. (145)-(146) are the exact solutions of eqs. (143)-(144) and that $|C_1|^2 + |C_2|^2 = 1$ for all t . The probability of transition is given by

$$\begin{aligned} P_{1 \rightarrow 2} = |C_2|^2 &= \frac{|\vec{D}_{21} \cdot \vec{E}_0|^2}{4\hbar^2} t^2 \text{sinc}^2\left(\frac{\Omega_R t}{2}\right) \\ &= \frac{1}{4} \chi_R^2 t^2 \text{sinc}^2\left(\frac{\Omega_R t}{2}\right) \\ &= \frac{1}{2} \frac{\chi_R^2}{\Omega_R^2} [1 - \cos(\Omega_R t)] \end{aligned} \quad (149)$$

which is almost identical to (138). A notable difference is that the probability in (149) oscillates at the generalized Rabi frequency Ω_R and not at the detuning frequency $\Delta\omega$ that characterizes the oscillation of the approximate solution (138). This difference is particularly evident when the detuning is $\Delta\omega = 0$ and the exact probability of transition assumes the simple form

$$P_{1 \rightarrow 2} = |C_2|^2 = \frac{1}{2} [1 - \cos(\chi_R t)]. \quad (150)$$

Note that in this case, one obtains oscillations whose frequency Ω_R is linearly proportional to the amplitude of the electric field^{42, 43}.

⁴² It is interesting to note that the exact result obtained here reduces to the approximate result of Eq. (138) whenever $P_{1 \rightarrow 2} \ll 1$. At small detunings, i.e. when $\Delta\omega \ll \chi_R$, the results coincide for small values of t . In particular, when $\Delta\omega = 0$ both expressions behave like $\chi_R^2 t^2 / 4$ when $t \ll \chi_R^{-1}$. On the other hand, when $\Delta\omega \gg \chi_R$, we have $\Omega_R \simeq \Delta\omega$ and expressions (138) and (149) become identical.

In this regime, it is of course guaranteed that $P_{1 \rightarrow 2} \ll 1$.

⁴³ Another aspect of the above is the dependence on the optical field. If we express $C_2(t)$ of Eq. (146) to first order in χ_R (equivalent to first order in E_0) in the case where $\Delta\omega \neq 0$, then the transition probability is once again identical to the one obtained in the perturbational analysis. Thus we see that the true dependence on the electric field is not linear, which is what underlines the field of nonlinear optics. The nonlinear dependence on E_0 could only be observed when laser light became available.

Part C - Quantization of the electromagnetic field

Motivation - SHO in thermal equilibrium

My goal now is to introduce to you the problem of black-body radiation which puzzled the world of science for many years and eventually motivated the quantization of the electromagnetic field. But, before we approach this problem directly, let me quickly introduce one of the most central and basic results of statistical mechanics. This result is known as Boltzmann's distribution and it tells us the probability of any arbitrary system to be in a particular "state"

$$p(\text{state}) = \frac{1}{Z} e^{-\beta H(\text{state})}$$

where $\beta = (K_b T)^{-1}$ with K_b being Boltzmann's constant, where $H(\text{state})$ is the energy (Hamiltonian) of the system when it is in that particular state. The parameter Z is a normalization constant which equals the summation of $e^{-\beta H(\text{state})}$ over all possible states. In statistical mechanics it is known by the name "partition function", but we will not need to use it here. Depending on whether states are discrete, or continuous, the quantity $p(\text{state})$ represents probability, or probability density, respectively. The concept of a "state" requires some clarification. We have seen what this word means in the quantum mechanical context, but it can also be used classically. Knowing the state of a system classically means being able to answer every question concerning measurements that may be performed on it. For a single particle, for example, the state at a given instance in time is completely defined in terms of its position $\vec{r} = (x, y, z)$ and momentum $\vec{p} = (p_x, p_y, p_z)$. In fact, the position and momentum is all that is required for us to predict the particles evolution according to Newton's equations. Let us consider an example where a particle of mass m is in thermal equilibrium with a heat-bath at temperature T . The probability of finding it in a particular state is therefore given by

$$p(\vec{r}, \vec{p}) = \frac{1}{Z} e^{-\frac{p_x^2 + p_y^2 + p_z^2}{2mK_b T}}$$

which implies that \vec{p} is a Gaussian random vector with independent components, such that the mean square value of each component is $\bar{p}_i^2 = mK_bT$. The average energy is $\frac{3}{2}K_bT$, or $\frac{1}{2}K_bT$ for each dimension of motion. If there is a gas of N such particles, the overall energy of the gas is $\frac{3N}{2}K_bT$. Now, I want to consider another example. This time we will look at a one dimensional SHO (a mass, or two masses on a spring). In this case

$$p(x, p_x) = \frac{1}{Z} e^{-\frac{p_x^2}{2mK_bT}} e^{-m\omega^2 \frac{x^2}{2K_bT}},$$

such that x and p_x are independent Gaussian variables with mean values of $\bar{p}^2 = mK_bT$ and $\bar{x}^2 = K_bT/(m\omega^2)$. The mean energy of the SHO in thermal equilibrium is therefore equal to⁴⁴ K_bT . Notice that in both our examples, the result is universal and it does not depend on the details of the particle, or the SHO. Thus, the mass, or the spring constant ($K = m\omega^2$) are totally irrelevant to the final result. This is a manifestation of the so called equipartition theorem of statistical mechanics. Pausing on this result a little, it is not difficult to become convinced that this result is very troubling, as indeed it was for physicists in the era that preceded quantum mechanics. We have just found out that the gas of free particles has energy of $\frac{3}{2}NK_bT$, but what if each particle consists of smaller particles that may move relative to each other like two masses on a spring⁴⁵? Then in addition to the energy $\frac{3}{2}NK_bT$ associated with the motion of the center of mass of each two particles, there is quite a significant additional energy in the amount of NK_bT . What if the particles in the gas consist of more than two smaller particles? then the energy increases farther and this has very significant consequences (for example it affects the heat capacity which is easily measurable). The troubling part is that it makes absolutely no difference how big those masses are. So if

⁴⁴ Here the kinetic energy and potential energy receive $\frac{1}{2}K_bT$ each so that the total energy is K_bT

⁴⁵ Recall that any two masses with forces holding them at some distance from each other can be viewed as connected by a spring when the perturbation is small.

an atom consists of a nucleus and electrons, does each electron that oscillated around the nucleus take up $K_b T$ energy? That appears a bit strange and inconsistent with experimental observation.

The solution to the problem comes from quantum mechanics. The extension of Boltzmann's principle to the quantum domain can be phrased as follows. A system in thermal equilibrium will be in a mixed state represented by the density matrix

$$\rho_{\text{th}} = \frac{1}{Z} e^{-\beta \mathbf{H}} \quad (151)$$

where the subscript 'th' is shorthand for thermal equilibrium, \mathbf{H} is the quantum Hamiltonian – i.e. the operator representing what we classically related to as energy, and $Z = \text{Trace}\{\exp(-\beta \mathbf{H})\}$. The probability of finding the system in some quantum state $|\psi\rangle$ is then given as usual by $\langle \psi | \rho_{\text{th}} | \psi \rangle$. Since ρ_{th} is orthogonal in the energy representation, a particularly simple expression corresponds to the case where $|\psi\rangle$ is an energy state. In particular, in the case of the SHO,⁴⁶

$$\begin{aligned} p(|n\rangle) &= \frac{e^{-\beta \hbar \omega n}}{\sum_n e^{-\beta \hbar \omega n}} \\ &= (1 - e^{-\beta \hbar \omega}) e^{-\beta \hbar \omega n}. \end{aligned} \quad (152)$$

In what follows we may just refer to the above as $p(n)$, i.e. to the probability that the SHO is in an energy state that has an energy $\hbar \omega(n + 1/2)$. Now we can calculate the average excitation of the SHO in a state of thermal equilibrium

$$\langle n \rangle = \sum n p(n) = (1 - U) \sum_n n U^n = (1 - U) U \frac{\partial}{\partial U} \sum_n U^n = \frac{U}{1 - U}$$

where I used $U = \exp(-\hbar \omega \beta)$. The mean energy of the SHO in thermal equilibrium is thus equal to

⁴⁶ The vacuum contribution to the energy gets cancelled, as can be easily seen.

$$\bar{E} = \hbar\omega \left(\langle n \rangle + \frac{1}{2} \right) = \frac{\hbar\omega}{\exp(\hbar\omega/K_B T) - 1} + \frac{\hbar\omega}{2}. \quad (153)$$

The latter term represents the energy of the vacuum. It is conceptually troubling, but not easily observable in measurements. In what follows we will concentrate on the measurable energy, which is represented by the first term on the right-hand-side of Eq. (153). It is interesting to consider this result as a function of the temperature and the frequency. In the low temperature, or high frequency limit ($K_b T \ll \hbar\omega$) it reduces to $\bar{E} = \hbar\omega \exp(-\hbar\omega/K_b T)$. In the opposite (high temperature, or low frequency) limit $K_b T \gg \hbar\omega$, one obtains the classical result $\bar{E} = K_b T$. This solves the riddle that we raised earlier. In order for the SHO degrees of freedom in the molecules (or atoms) of the gas to have an effect on the gas, the temperature must be greater than $\hbar\omega/K_b$. Note that when the mass is close to 0, and for a given spring constant, the frequency $\omega = \sqrt{K/m}$ is very large and the SHO does not increase the mean energy of the system. In many cases, it is convenient to represent the distribution of n in thermal equilibrium in terms of the mean value $\bar{n} = \langle n \rangle$. The resulting expression is

$$p(n) = \frac{1}{\bar{n} + 1} \left(\frac{\bar{n}}{\bar{n} + 1} \right)^n. \quad (154)$$

The Boltzmann distribution gives us the probability that the SHO is in an energy state $|n\rangle$.

The electromagnetic field as a collection of SHOs

Let us consider the simplest possible problem, where we look at a cavity consisting of 2 infinitely large flat mirrors positioned at $z = 0$ and at $z = L$. We assume that there are no charges or currents inside the cavity. The electromagnetic field that satisfies the boundary conditions, and as you have studied in courses on electromagnetic theory, it will have the form of *standing waves*, which are waves in which the temporal dependence can be separated

from the spatial dependence. One can identify two types of fields that may exist in that cavity. The type where the electric field vector points in the x direction and the magnetic field points in the y direction, and the type where the electric field is in the y direction and the magnetic field is in the x direction. We denote the modes of the first type by the index $\lambda = 1$ and those of the second type by the index $\lambda = 2$. Thus, for the first type we have

$$\vec{E}_{k,1}(z, t) = \vec{1}_x \left(\frac{2k^2 c^2}{V \varepsilon_0} \right)^{1/2} q_{k,1}(t) \sin(kz) \quad (155)$$

$$\vec{B}_{k,1}(z, t) = -\vec{1}_y \left(\frac{2}{V \varepsilon_0 c^2} \right)^{1/2} p_{k,1}(t) \cos(kz) \quad (156)$$

and for the second type

$$\vec{E}_{k,2}(z, t) = \vec{1}_y \left(\frac{2k^2 c^2}{V \varepsilon_0} \right)^{1/2} q_{k,2}(t) \sin(kz) \quad (157)$$

$$\vec{B}_{k,2}(z, t) = \vec{1}_x \left(\frac{2}{V \varepsilon_0 c^2} \right)^{1/2} p_{k,2}(t) \cos(kz), \quad (158)$$

where V represents the "effective volume" of the cavity⁴⁷. In order for the field to satisfy the boundary conditions on the mirrors, the wave-number k may only assume discrete values that are integer multiples of π/L , i.e.

$$k = m \frac{\pi}{L},$$

with $m = 0, 1, 2, \dots$ being a non-negative integer. You may wonder why did I write the space dependent part explicitly, but wrote the time dependent part as $q_{k,j}(t)$ and $p_{k,j}(t)$. Since you have taken a course on electromagnetic theory you surely remember that these functions have a known sinusoidal dependence on time. Nonetheless, I want to derive the equations that these

⁴⁷ In our assumptions the volume is infinite, but the same sort of arguing could follow if we considered finite mirrors that are much closer to each other than their lateral dimensions.

quantities satisfy, so that I can later claim that they are equivalent to a "generalized" coordinate and momentum. In other words, I will later want to claim that when quantizing the electromagnetic field, the quantities $q_{k,j}$ and $p_{k,j}$ become operators whose commutation relations are the same as the commutation relations of the position and momentum operators. This is certainly not a rigorous procedure. If I wanted to be rigorous, I would have to first teach you some classical electrodynamics and show you how to define a classical Hamiltonian for the electromagnetic field, and how to identify the conjugate variables that play the role of position and momentum. Then the process of quantization would be as natural as it was in the case of mechanical systems. But this procedure is not part of our course and therefore I need to come up with some not very rigorous arguments that will hopefully be plausible enough to convince you that you can believe me when I say that $q_{k,j}$ and $p_{k,j}$ are equivalent to position and momentum. So after this very long disclaimer let us find the equations that $q_{k,j}$ and $p_{k,j}$ must satisfy. This is done by substituting Eqs. (155 -158) in the the following two Maxwell's equations

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \nabla \times \vec{B} = \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t},$$

and the result of this substitution yields

$$\frac{dq_{k,\lambda}}{dt} = p_{k,\lambda} \tag{159}$$

$$\frac{dp_{k,\lambda}}{dt} = -k^2 c^2 q_{k,\lambda}. \tag{160}$$

The above two equations are identical (to within the coefficients) to the equations satisfied by the coordinate and momentum of a one dimensional SHO, and one easily finds that $q_{k,\lambda}(t) = A \cos(\omega t + \phi)$ and $p_{k,\lambda} = -\omega A \sin(\omega t + \phi)$, where $\omega = kc$ and where A and ϕ are the amplitude and initial phase that are determined by initial conditions.

The total energy of the electro-magnetic field is given by

$$H = \frac{1}{2} \int dV \left[\epsilon_0 E^2(\vec{r}, t) + \frac{1}{\mu_0} B^2(\vec{r}, t) \right],$$

and by substituting Eqs. (155–158) into this expression we find that

$$H = \sum_{k,\lambda} \frac{1}{2} (p_{k,\lambda}^2 + \omega^2 q_{k,\lambda}^2). \quad (161)$$

What we have above is the energy expressed in terms of the quantities $q_{k,j}$ and $p_{k,j}$, which we identified (not rigorously, but we have had this discussion before ...) as equivalent to the conjugate coordinate and momentum. Therefore, it can be claimed that Eq. (161) is the classical Hamiltonian of the electromagnetic field in our one dimensional cavity. Quantization of this field should hence follow by treating $q_{k,j}$ and $p_{k,j}$ as operators, as we shall indeed do in a little while. At this point, it is important to note that the Hamiltonian in Eq. (161) is exactly the same as the Hamiltonian of an infinite array of independent SHOs, with each SHO identified by a particular pair of indices (k, λ) . The quantization of the field should therefore also be equivalent to the quantization of many independent SHOs.

Let us now consider a 3-dimensional cavity of dimensions L^3 . The situation is quite similar, except that now we have 3 wave-vector components k_x, k_y and k_z , each of which can receive integer multiples of⁴⁸ π/L . Apart

⁴⁸ In fact, in this case we may write the electric field as

$$E_x(\vec{r}, t) = E_x(t) \cos(k_x x) \sin(k_y y) \sin(k_z z)$$

$$E_y(\vec{r}, t) = E_y(t) \sin(k_x x) \cos(k_y y) \sin(k_z z)$$

$$E_z(\vec{r}, t) = E_z(t) \sin(k_x x) \sin(k_y y) \cos(k_z z)$$

and it is evident that at most one of the k vector components is allowed to be 0. If two components of k are 0, the electric field drops to 0. The magnetic field doesn't drop to 0, but the two are coupled whenever $\omega_k = |k|c > 0$. Thus the case of two components being 0 is not possible. Three components on the other hand can be 0, because then $\omega_k = 0$ and there is no coupling to the magnetic field.

from that, the principles remain the same as before, only the summation in (161) needs to be performed over the 3-dimensional grid of values that can be assumed by the wave-vector \vec{k} . For each legal value of \vec{k} and for each polarization⁴⁹ the electromagnetic field can be assigned a "coordinate" $q_{\vec{k},\lambda}$ and a momentum $p_{\vec{k},\lambda}$ and an energy equal to $\frac{1}{2} \left(p_{\vec{k},\lambda}^2 + \omega_k^2 q_{\vec{k},\lambda}^2 \right)$ in the classical picture, or $\hbar\omega(n + 1/2)$ in the quantum case. In other words, the field can be viewed as consisting of many discrete SHOs, one for every mode (\vec{k}, λ) .

Black-body radiation

We are now ready to consider the problem of black-body radiation (BBR), which was one of the most important problems in the second half of the nineteenth century and in the beginning of the twentieth century. The analysis of BBR was one of the first to suggest that light comes in discrete quanta of energy and eventually led to the modern description of light quantization.

People were trying to understand the emission spectrum from dark objects when heated to high temperatures. The reason for the great interest in this problem was the strange theoretical prediction that the spectral power density of the electromagnetic field would diverge at high frequencies, a phenomenon that had been referred to as the *ultraviolet catastrophe*. The formal definition of a black body is an object that absorbs all the radiation that is incident upon it. Since finding such a perfectly "black" substance was difficult, Kirchoff suggested that a good model for a black body would be that of a closed cavity in which there is a small hole. The idea was that radiation incident on the hole of the cavity, enters it and its probability of exiting upon reflection is negligibly small. This was therefore considered to be an almost

⁴⁹ In the 3-D case, the unit polarization vectors of the electric field are orthogonal to \vec{k} so that its divergence remains 0. We still have 2 orthogonal polarizations for each \vec{k} , but their absolute directions in the 3D space depend on the value of \vec{k} .

perfect black-body that absorbed all frequencies for which the wavelength was considerably smaller than the hole (the fact that long wavelength RF waves were not absorbed does not pose a particular problem). People were then experimenting with such black bodies, which they heated in various methods and observed the emitted spectrum of radiation. The observed spectrum defied the theoretical predictions of an ultraviolet catastrophe and thus formed one of the most important problems of those days. It was eventually solved by Plank who proposed the quantization of the field's energy.

Let us now do some calculations in order to understand the problem (of the ultraviolet catastrophe) and its solution. What we want to do is find out the power per solid angle per unit frequency that is emitted by a black body at temperature T . Having seen the structure of the free electromagnetic field in the 3 dimensional cavity in the previous section, we are ready to perform the calculation. We start by counting the number of modes (i.e the number of SHOs) in a frequency range between ω and $\omega + d\omega$. To do this we note that $\omega = kc$ and so the number of oscillators that we seek is the number of modes whose k vector has a modulus between $k = \omega/c$ and $k + dk$ where $dk = d\omega/c$. Finding the exact number is cumbersome because the points in k space (representing the possible modes) are discrete, but we can approximate by pointing out that the volume occupied by each point in k -space is one in $\frac{\pi^3}{L^3}$ where I assume that the box dimensions are identical in all directions⁵⁰. The volume enclosed between k and $k + dk$ is $4\pi k^2 dk$, but we need to divide it by 8 because we only consider positive k values. Then we need to divide this volume by the volume that each point (representing a mode) occupies in order to find the number of modes. This procedure yields

$$2 \times \frac{1}{8} \times \frac{V}{\pi^3} \times 4\pi k^2 dk = \frac{V}{c^3 \pi^2} \omega^2 d\omega, \quad (162)$$

where the factor of 2 is because there are two oscillators at each \vec{k} owing to the fact that there can be two independent polarizations of the electromag-

⁵⁰ This means that, the density of points is $\frac{V}{\pi^3}$.

netic field. Now, observation of the ultraviolet catastrophe is immediate. The average energy per unit volume per unit frequency $d\omega$ is then

$$\bar{W}_T(\omega) = \frac{K_b T}{c^3 \pi^2} \omega^2,$$

A value that diverges as ω^2 at high frequency. This is the famous catastrophe. The quantum result settles this issue very elegantly, since each SHO has a mean energy given by Eq. (153) and therefore

$$\bar{W}_T(\omega) = \hbar \omega \langle n \rangle \frac{1}{c^3 \pi^2} \omega^2 = \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar \omega / K_b T) - 1}. \quad (163)$$

This has a completely different shape. At low frequencies $\hbar \omega \ll K_b T$ it reduces to the classical result, whereas at high frequencies it reduces to 0 exponentially

$$\bar{W}_T(\omega) \simeq \frac{\hbar \omega^3}{\pi^2 c^3} \exp(-\hbar \omega / K_b T).$$

Calculating the entire energy integrated over all frequencies, one finds

$$\int \bar{W}_T(\omega) d\omega = \frac{\pi^2 K_b^4 T^4}{15 c^3 \hbar^3}$$

which has fourth power dependence on the temperature. This dependence is known as the Stefan Boltzmann radiation law.

The quantized electromagnetic field

Let us now represent a little more rigorously the quantization of the field. I will not derive the quantization once again for the more rigorous case, but I will try to quote more accurately the central results. First of all, the metallic cavity is not needed in the more general case of field quantization. In principle, the k vectors can be viewed as the vectors that represent the spatial fourier transform of the field, which is equivalent to the decomposition of the field into plain waves. In this case \vec{k} has a continuum of values. For

reasons of convenience (that have to do with tradition) people often prefer to consider cases in which the \vec{k} space consists of discrete points. They do that by imagining a fictitious cavity of some volume $V = L^3$, whose dimension L is considered to be very large. The cavity is not assumed to be metallic, but rather it is assumed to follow periodic boundary conditions so that the field at $x = L$ equals the field at $x = 0$, and similarly in the two other dimensions. Then the grid of k -vectors consists of all the points $(m_x, m_y, m_z)2\pi/L$, where (m_x, m_y, m_z) are integers that can be negative, positive, or 0. The reason for assuming periodic (rather than 0) boundary conditions is that they are less restrictive in terms of the fields that they can represent. Thus for example, one is not limited only to standing waves, and for $L \rightarrow \infty$ all physical fields can be properly represented.

Let me now write down the representation of the electric and magnetic fields in terms of their conjugate coordinate and momenta $q_{\vec{k},\lambda}(t)$ and $p_{\vec{k},\lambda}(t)$.

$$\vec{E}(\vec{r}, t) = \frac{1}{\sqrt{\varepsilon_0 V}} \sum_{\vec{k}, \lambda} \hat{e}_{\vec{k}, \lambda} \left[\omega_k q_{\vec{k}, \lambda}(t) \sin(\vec{k} \cdot \vec{r}) + p_{\vec{k}, \lambda}(t) \cos(\vec{k} \cdot \vec{r}) \right] \quad (164)$$

$$\vec{B}(\vec{r}, t) = \frac{1}{c\sqrt{\varepsilon_0 V}} \sum_{\vec{k}, \lambda} \left(\hat{k} \times \hat{e}_{\vec{k}, \lambda} \right) \cdot \left[\omega_k q_{\vec{k}, \lambda}(t) \sin(\vec{k} \cdot \vec{r}) + p_{\vec{k}, \lambda}(t) \cos(\vec{k} \cdot \vec{r}) \right]. \quad (165)$$

Obtaining these forms is again an exercise in classical electrodynamics. We will not attempt to demonstrate it here, but rather we take it for granted that these forms are correct and that $q_{\vec{k},\lambda}$ and $p_{\vec{k},\lambda}$ are indeed legitimate coordinate and momenta. The energy

$$H = \frac{1}{2} \int dV \left[\varepsilon_0 E^2(\vec{r}, t) + \frac{1}{\mu_0} B^2(\vec{r}, t) \right] = \sum_{\vec{k}, \lambda} \frac{1}{2} (p_{\vec{k}, \lambda}^2 + \omega_k^2 q_{\vec{k}, \lambda}^2), \quad (166)$$

as can be shown by direct substitution (HW exercise), reveals that each field mode functions just as a mechanical SHO whose mass is equal to unity.

Quantization is now performed in a perfectly routine manner by assigning operators $\mathbf{q}_{\vec{k},\lambda}$ and $\mathbf{p}_{\vec{k},\lambda}$ for the coordinates and momenta, such that their commutation relations are given by

$$[\mathbf{q}_{\vec{k},\lambda}, \mathbf{p}_{\vec{k}',\lambda'}] = i\hbar\delta_{\vec{k},\vec{k}'}\delta_{\lambda,\lambda'}.$$

With this substitution, the energy H turns into the quantum Hamiltonian (energy operator) \mathbf{H} and the electric and magnetic fields become operators. In what follows we will use the convenient normalization relations that we introduced in the context of the mechanical SHO⁵¹.

$$\sqrt{\frac{\omega}{\hbar}}\mathbf{q}_{\vec{k},\lambda} \rightarrow \mathbf{q}_{\vec{k},\lambda}, \text{ and } \sqrt{\frac{1}{\hbar\omega}}\mathbf{p}_{\vec{k},\lambda} \rightarrow \mathbf{p}_{\vec{k},\lambda},$$

and the commutation relations reduce to

$$[\mathbf{q}_{\vec{k},\lambda}, \mathbf{p}_{\vec{k}',\lambda'}] = i\delta_{\vec{k},\vec{k}'}\delta_{\lambda,\lambda'}.$$

Finally, with the use of the annihilation operator

$$\mathbf{a}_{\vec{k},\lambda} = \frac{1}{\sqrt{2}}(\mathbf{q}_{\vec{k},\lambda} + i\mathbf{p}_{\vec{k},\lambda}),$$

and the creation operator, its complex conjugate, $\mathbf{a}_{\vec{k},\lambda}^\dagger$, we may rewrite the Hamiltonian as

$$\mathbf{H} = \sum_{\vec{k},\lambda} \frac{\hbar\omega_k}{2} (\mathbf{p}_{\vec{k},\lambda}^2 + \mathbf{q}_{\vec{k},\lambda}^2) = \sum_{\vec{k},\lambda} \hbar\omega_k \left(\mathbf{a}_{\vec{k},\lambda}^\dagger \mathbf{a}_{\vec{k},\lambda} + \frac{1}{2} \right). \quad (167)$$

It is worthwhile now to consider the question of notation. How do we describe the state-vector of the overall field? Since the various SHOs constituting the electromagnetic field do not interact with each other, the overall state-vector

⁵¹ I use the same notation for the original and normalized quantities. In what follows, unless explicitly specified otherwise, only the normalized form will be used.

is a tensor product of all the individual ones⁵². Thus, for example, the overall field can be in a state

$$|\psi_A\rangle = |n_{\vec{k}_1, \lambda_1}^{(1)}, n_{\vec{k}_2, \lambda_2}^{(2)}, \dots\rangle = |n^{(1)}\rangle_{\vec{k}_1, \lambda_1} \otimes |n^{(2)}\rangle_{\vec{k}_2, \lambda_2} \otimes \dots$$

Which means that mode $\{\vec{k}_1, \lambda_1\}$ is excited to the level number $n^{(1)}$, mode $\{\vec{k}_2, \lambda_2\}$ is excited to energy level number $n^{(2)}$, etc. An equally legitimate state would be

$$|\psi_B\rangle = |\alpha_{\vec{k}_1, \lambda_1}^{(1)}, \alpha_{\vec{k}_2, \lambda_2}^{(2)}, \dots\rangle = |\alpha^{(1)}\rangle_{\vec{k}_1, \lambda_1} \otimes |\alpha^{(2)}\rangle_{\vec{k}_2, \lambda_2} \otimes \dots$$

which means that mode $\{\vec{k}_1, \lambda_1\}$ is excited to a coherent state characterized by the parameter $\alpha^{(1)}$, mode $\{\vec{k}_2, \lambda_2\}$ is in coherent state $\alpha^{(2)}$, etc. Moreover, we could have a state

$$|\psi_C\rangle = |n_{\vec{k}_1, \lambda_1}, \alpha_{\vec{k}_2, \lambda_2}, \dots\rangle,$$

where mode $\{\vec{k}_1, \lambda_1\}$ is excited to a number n and mode $\{\vec{k}_2, \lambda_2\}$ is in coherent state $\alpha^{(2)}$, and so on. This latter expression is perhaps less popular, but it is just as legitimate as the previous two. Moreover, any superpositions of $|\psi_A\rangle$, $|\psi_B\rangle$ and $|\psi_C\rangle$ are also legitimate state-vectors as we have seen before. We must always keep in mind that what we have is a lot of SHOs, one for each mode of the electromagnetic field. Since the number of modes is infinite, a practical notation ignores all modes that are not excited. Thus,

$$|n_{\vec{k}_1, \lambda_1}^{(1)}\rangle$$

indicates that of all modes only the one that corresponds to the wave-vector \vec{k} and polarization state λ is excited to a number state $n^{(1)}$, whereas all the other modes of the field are in their vacuum state. Similarly

⁵² It is assumed in this part that the meaning of the term tensor product is known to the reader. Please consult Appendix I for the definition and necessary properties of tensor products in the context of quantum physics.

$$|n_{\vec{k}_1, \lambda_1}, \alpha_{\vec{k}_2, \lambda_2}\rangle$$

implies that only modes $\{\vec{k}_1, \lambda_1\}$ and $\{\vec{k}_2, \lambda_2\}$ are excited with the first being excited to a number state n and with the latter being excited to a coherent state defined by the complex parameter α . All other modes are in the vacuum state and hence they need not be specified.

We are now ready to start introducing the concept of a photon. When the various modes of the field are in number states, as for example in the case of $|\psi\rangle = |n_{\vec{k}_1, \lambda_1}^{(1)}, n_{\vec{k}_2, \lambda_2}^{(2)}\rangle$, we say that the field has $n^{(1)}$ photons in mode $\{\vec{k}_1, \lambda_1\}$ and $n^{(2)}$ photons in mode $\{\vec{k}_2, \lambda_2\}$. We also say that the total number of photons is $n_{tot} = n^{(1)} + n^{(2)}$. A photon thus appears as a way of counting the excitation level of the various SHOs. What is then the meaning of a field containing a single photon. A straightforward interpretation of the above description is that a single photon field is what would exist if exactly one mode of our cavity was excited to the first level. So that the state-vector is $|1_{\vec{k}, \lambda}\rangle$, for some particular \vec{k} and λ . This conclusion is certainly correct, but it has some discouraging attributes. Most obvious is the fact that a single cavity mode is completely non-localized. After all, the concept of a cavity is fictitious as it was merely introduced as a mathematical trick in order to be able to relate to discrete modes of the field. Luckily, there is more depth in this story. Consider for example the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left[|1_{\vec{k}_1, \lambda_1}\rangle + |1_{\vec{k}_2, \lambda_2}\rangle \right]. \quad (168)$$

which is a superposition of two number states of two separate modes. If one performs a measurement of the number of photons in state $\{\vec{k}_1, \lambda_1\}$, he will find either 0, or 1 photons there. The same is true for $\{\vec{k}_2, \lambda_2\}$. This is a pure state in which neither the energy, nor the number of photons for each one of the two modes is defined. The total energy is also undefined, as the state $|\psi\rangle$ is not an eigenstate of the Hamiltonian (167). A measurement of the energy may yield either $\hbar\omega_{k_1}$, or $\hbar\omega_{k_2}$. On the other hand the total number

of photons in the field is certain and equal to 1. So that we have exactly one photon in the overall electromagnetic field. Let us make it a little more rigorous now. In the past we defined $\mathbf{N} = \mathbf{a}^\dagger \mathbf{a}$ as the *Number operator* of the SHO, but there was no clear value in defining it because it was equivalent to the Hamiltonian \mathbf{H} to within multiplication by $\hbar\omega$. Now the situation is slightly different as we define the total number operator as

$$\mathbf{N} = \sum_{\vec{k}, \lambda} \mathbf{N}_{\vec{k}, \lambda},$$

where $\mathbf{N}_{\vec{k}, \lambda} = \mathbf{a}_{\vec{k}, \lambda}^\dagger \mathbf{a}_{\vec{k}, \lambda}$ is the number operator of one individual mode of the electromagnetic field. The energy operator, on the other hand, is defined as

$$\mathbf{H} = \sum_{\vec{k}, \lambda} \hbar\omega_k \left(\mathbf{N}_{\vec{k}, \lambda} + \frac{1}{2} \right)$$

and the two are no longer proportional. Thus the state $|\psi\rangle$ that we defined in our example (168) is an eigenstate of \mathbf{N}_{tot} (with eigenvalue equal to 1), but not of \mathbf{H} and so the total number of photons is unity, whereas the energy is not uniquely defined. Let us take this reasoning further. Consider modes that represent fields propagating along a particular direction in space, to which we will refer as the z direction. Thus, we are referring to modes whose k -vectors are $k_j \hat{z}$, with $k_j = \frac{2\pi j}{L}$ and with j being a positive integer number. Let us also assume for simplicity that we are concerned only with modes of one given polarization λ (e.g. x -polarized modes only). Then the only mode index that we need to consider is a positive scalar number k . It is often more intuitive to index the mode in terms of its frequency ω , since it has a unique relation to the wave-number $k = \omega/c$. Let us now consider the state-vector

$$|\psi\rangle = \sum_{\omega} f(\omega) |1_\omega\rangle \quad (169)$$

with the coefficients $f(\omega)$ satisfying $\sum_{\omega} |f(\omega)|^2 = 1$, which is required for the state $|\psi\rangle$ to be normalized. By applying \mathbf{N}_{tot} to $|\psi\rangle$ we find that $|\psi\rangle$ is again

an eigenstate of \mathbf{N}_{tot} with eigenvalue 1, so that there is exactly one photon in this field. Here a group of modes is excited and as we shall see later, the photon can be somewhat localized both in time and in space. We will need to define the concept of detection in order to quantify the sense in which the photons are localized, and therefore we leave the further discussion of this problem for later.

Let us now represent the electric and magnetic fields in terms of the creation and annihilation operators as follows

$$\vec{\mathbf{E}}(\vec{r}) = i \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \hat{e}_{\vec{k}, \lambda} \left[\mathbf{a}_{\vec{k}, \lambda} e^{i\vec{k} \cdot \vec{r}} - \mathbf{a}_{\vec{k}, \lambda}^\dagger e^{-i\vec{k} \cdot \vec{r}} \right] \quad (170)$$

$$\vec{\mathbf{B}}(\vec{r}) = \frac{i}{c} \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \left(\hat{k} \times \hat{e}_{\vec{k}, \lambda} \right) \left[\mathbf{a}_{\vec{k}, \lambda} e^{i\vec{k} \cdot \vec{r}} - \mathbf{a}_{\vec{k}, \lambda}^\dagger e^{-i\vec{k} \cdot \vec{r}} \right], \quad (171)$$

where $\hat{e}_{\vec{k}, \lambda} \cdot \vec{k} = 0$ and where $\hat{k} = \vec{k}/k$. Expressions (170)-(171) are the quantum representations of the electric and magnetic fields. They are quantum operators represented in the Schrödinger picture. Notice that \vec{r} is NOT an operator in this description. It is just an indication of the point in 3D space that one is referring to. It is often more convenient to represent these operators in the Heisenberg picture where their time dependence is explicitly included and the form is closer to what we are familiar with in classical field theory. To represent the time dependence of the fields we find the time dependence of \mathbf{a} in the Heisenberg picture

$$\frac{d\mathbf{a}}{dt} = \frac{i}{\hbar} [\mathbf{H}, \mathbf{a}] = i\omega [\mathbf{a}^\dagger \mathbf{a}, \mathbf{a}] = -i\omega \mathbf{a}, \quad (172)$$

implying that

$$\mathbf{a}_{\vec{k}, \lambda}(t) = \mathbf{a}_{\vec{k}, \lambda}(0) e^{-i\omega_k t}.$$

Consequently

$$\vec{\mathbf{E}}(\vec{r}, t) = i \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \hat{e}_{\vec{k}, \lambda} \left[\mathbf{a}_{\vec{k}, \lambda} e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} - \mathbf{a}_{\vec{k}, \lambda}^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right], \quad (173)$$

and similarly for the magnetic field. The quadrature components of the electric field are quantized directly from (164)

$$\vec{\mathbf{E}}(\vec{r}, t) = - \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{\varepsilon_0 V} \right)^{\frac{1}{2}} \hat{e}_{\vec{k}, \lambda} \left[\mathbf{q}_{\vec{k}, \lambda} \sin(\vec{k} \cdot \vec{r} - \omega_k t) + \mathbf{p}_{\vec{k}, \lambda} \cos(\vec{k} \cdot \vec{r} - \omega_k t) \right]$$

so that the \mathbf{q} and \mathbf{p} operators represent the quadrature components of the electric field. Together with the notion that the energy operator is given by $\mathbf{H} = \frac{1}{2} \hbar \omega_k (q_{\vec{k}, \lambda}^2 + p_{\vec{k}, \lambda}^2)$ we may represent the electric field in terms of the phase space coordinates q and p . We will return to this description later.

Fully quantum description of the interaction between the field and a two-level atom

The Hamiltonian of the entire atom and field system can be written as

$$\mathbf{H} = \mathbf{H}_A + \mathbf{H}_f + \mathbf{H}_I \quad (174)$$

where the interaction hamiltonian is $\mathbf{H}_I = \vec{\mathbf{D}} \cdot \vec{\mathbf{E}}$, as we have seen before, where $\mathbf{H}_f = \sum_{\vec{k}, \lambda} \hbar \omega_k \left(\mathbf{a}_{\vec{k}, \lambda}^\dagger \mathbf{a}_{\vec{k}, \lambda} + \frac{1}{2} \right)$ and where \mathbf{H}_A can be expressed in the form

$$\mathbf{H}_A = \hbar \omega_g |g\rangle\langle g| + \hbar \omega_e |e\rangle\langle e|.$$

For simplicity, we have assumed a two level atom (meaning that only two levels participate in the interaction with the field) with $|g\rangle$ and $|e\rangle$ denoting "ground" and "excited" states and where $\hbar \omega_g$ and $\hbar \omega_e$ are their energies in the absence of interaction with the electromagnetic field. In what follows it is convenient to define $\omega_0 = \omega_e - \omega_g$, or equivalently to set the energy ladder such that the ground-state energy is 0. In this case, the atomic Hamiltonian is simply $\mathbf{H}_A = \hbar \omega_0 |e\rangle\langle e|$. The dipole moment operator is characterized as we had previously. Its diagonal matrix elements in the basis of $\{|g\rangle, |e\rangle\}$ vanish, and so

$$\mathbf{D}_i = \mu_i |e\rangle\langle g| + \mu_i^* |g\rangle\langle e| = \mu_i (|e\rangle\langle g| + |g\rangle\langle e|)$$

where $i = x, y, z$ is the cartesian coordinate of $\vec{\mathbf{D}}$ and where in the last term we have assumed for simplicity a symmetric scenario, for which the matrix coefficient is real. The time evolution of the state-vector is governed by the Schrödinger equation

$$\frac{\partial |\psi\rangle}{\partial t} = -\frac{i}{\hbar} (\mathbf{H}_A + \mathbf{H}_f) |\psi\rangle - \frac{i}{\hbar} \mathbf{H}_I |\psi\rangle$$

where the first two terms on the right-hand side represent free evolution, whereas the last term represents the effect of the interaction. The state-vector $|\psi\rangle$ represents the entire system, which consists of the atom and the

infinite number of (field modes) SHOs that surround it. In the past (i.e. in the semiclassical treatment) we approached this problem by expanding the state $|\psi(t)\rangle$ in terms of the energy states of the unperturbed atomic Hamiltonian \mathbf{H}_A , i.e. as $|\psi(t)\rangle = C_g(t)|g\rangle + C_e(t)e^{-i\omega_0 t}|e\rangle$, substituted it in the Schrödinger equation and found the equations for $C_{g,e}$. This definition of the system state is no longer sufficient. Now the state of the system must reflect the state of the atom and the field (all of its modes that can interact with the atom) together. We will try a more tractable approach which will be described in what follows. First, we perform a “change of coordinates,” so that instead of following the time evolution of $|\psi(t)\rangle$ in the Schrödinger picture, we will follow the evolution of

$$|\tilde{\psi}(t)\rangle = e^{\frac{i}{\hbar}(\mathbf{H}_A + \mathbf{H}_f)t} |\psi(t)\rangle.$$

so that the free evolution is factored out, as we shall now see. This representation is often called the interaction picture. Substitution into the Schrödinger equation reveals that

$$\frac{\partial |\tilde{\psi}\rangle}{\partial t} = -\frac{i}{\hbar} \tilde{\mathbf{H}}_I |\tilde{\psi}\rangle. \quad (175)$$

where

$$\tilde{\mathbf{H}}_I = e^{\frac{i}{\hbar}(\mathbf{H}_A + \mathbf{H}_f)t} \mathbf{H}_I e^{-\frac{i}{\hbar}(\mathbf{H}_A + \mathbf{H}_f)t}.$$

Using $\mathbf{H}_I = \vec{\mathbf{D}} \cdot \vec{\mathbf{E}}$ in the above expression, we find that we can write $\tilde{\mathbf{H}}_I = \vec{\tilde{\mathbf{D}}} \cdot \vec{\tilde{\mathbf{E}}}$, where

$$\vec{\tilde{\mathbf{D}}}(t) = e^{\frac{i}{\hbar}\mathbf{H}_A t} \vec{\mathbf{D}} e^{-\frac{i}{\hbar}\mathbf{H}_A t}$$

$$\vec{\tilde{\mathbf{E}}}(t) = e^{\frac{i}{\hbar}\mathbf{H}_f t} \vec{\mathbf{E}} e^{-\frac{i}{\hbar}\mathbf{H}_f t}$$

Notice that these are just the time dependent operators that correspond to the Heisenberg picture in the absence of interaction between the atom and the field. Thus, the electric field operator $\vec{\tilde{\mathbf{E}}}$ is just the expression for $\vec{E}(\vec{r}, t)$

written in (173). As to $\tilde{\mathbf{D}}(t)$, we obtain it by substituting the explicit form of $\vec{\mathbf{D}}$, which leads to

$$\begin{aligned}\tilde{\mathbf{D}} &= \vec{\mu} e^{\frac{i}{\hbar} \mathbf{H}_A t} (|g\rangle\langle e| + |e\rangle\langle g|) e^{-\frac{i}{\hbar} \mathbf{H}_A t} \\ &= \vec{\mu} \left(|g\rangle\langle e| e^{-i\omega_0 t} + |e\rangle\langle g| e^{i\omega_0 t} \right),\end{aligned}\quad (176)$$

where we made use of the fact that $|g\rangle$ and $|e\rangle$ are eigenstates of \mathbf{H}_A . With all of the above, we find that

$$\tilde{\mathbf{H}}_I = i \sum_{\vec{k}, \lambda} s_{\vec{k}, \lambda} \left\{ \mathbf{a}_{\vec{k}, \lambda} e^{-i\omega_k t + i\vec{k} \cdot \vec{r}} - \mathbf{a}_{\vec{k}, \lambda}^\dagger e^{i\omega_k t - i\vec{k} \cdot \vec{r}} \right\} \left(e^{i\omega_0 t} |e\rangle\langle g| + e^{-i\omega_0 t} |g\rangle\langle e| \right),$$

where $s_{\vec{k}, \lambda} = \left(\frac{\hbar \omega_k}{2\varepsilon_0 V} \right)^{1/2} (\hat{e}_{\vec{k}, \lambda} \cdot \vec{\mu})$.

Let us now return to our problem of finding out the probability of a transition from one energy state to another. Assume that initially the state of the system is $|\psi_i\rangle = |a_i, n_i\rangle$ at $t = 0$, where $a_i = e, g$ refers to the state of the atom and where a particular mode of the field is excited by n photons (we avoid writing the mode index explicitly in order to avoid cluttering the notation). We want to know the probability that at $t > 0$ the system is found in state $|\psi_f\rangle = |a_f, n_f\rangle$, where the initial and final states are orthogonal, i.e. $\langle a_f, n_f | a_i, n_i \rangle = 0$. As we have argued in the semiclassical case, the solution $|\tilde{\psi}(t)\rangle$ can be shown to satisfy the following integral equation, which is just a different form of (175)

$$|\tilde{\psi}(t)\rangle = |\tilde{\psi}(0)\rangle - \frac{i}{\hbar} \int_0^t \tilde{\mathbf{H}}_I(t') |\tilde{\psi}(t')\rangle dt'. \quad (177)$$

Since we will be assuming small t , $|\tilde{\psi}(t')\rangle$ does not differ much from $|\tilde{\psi}(0)\rangle$ and it can be replaced by it on the right-hand side, leading to

$$|\tilde{\psi}(t)\rangle = |\tilde{\psi}(0)\rangle - \frac{i}{\hbar} \int_0^t \tilde{\mathbf{H}}_I(t') |\tilde{\psi}(0)\rangle dt', \quad (178)$$

which can be used to arrive at the transition probabilities as follows

$$|\langle a_f, n_f | \tilde{\psi}(t) \rangle|^2 = \frac{s^2}{\hbar^2} \times \\ \times \left| \langle a_f, n_f | \left[\int_0^t dt' \left\{ \mathbf{a}_{\vec{k},\lambda} e^{-i\omega_k t'} - \mathbf{a}_{\vec{k},\lambda}^\dagger e^{i\omega_k t'} \right\} \left(e^{i\omega_0 t'} |e\rangle \langle g| + e^{-i\omega_0 t'} |g\rangle \langle e| \right) \right] | a_i, n_i \rangle \right|^2.$$

The result can be described as the sum of four types of terms;

$$\frac{s^2}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 - \omega_k)t}{2} \right) \left| \langle a_f, n_f | \mathbf{a}_{\vec{k},\lambda} | e \rangle \langle g | a_i, n_i \rangle \right|^2 \quad (179)$$

$$\frac{s^2}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 - \omega_k)t}{2} \right) \left| \langle a_f, n_f | \mathbf{a}_{\vec{k},\lambda}^\dagger | g \rangle \langle e | a_i, n_i \rangle \right|^2 \quad (180)$$

$$\frac{s^2}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 + \omega_k)t}{2} \right) \left| \langle a_f, n_f | \mathbf{a}_{\vec{k},\lambda} | g \rangle \langle e | a_i, n_i \rangle \right|^2 \quad (181)$$

$$\frac{s^2}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 + \omega_k)t}{2} \right) \left| \langle a_f, n_f | \mathbf{a}_{\vec{k},\lambda}^\dagger | e \rangle \langle g | a_i, n_i \rangle \right|^2. \quad (182)$$

The first term excites the atom, but annihilates a photon in the field. This is an absorption process, where a photon of light is absorbed in order to excite the atom. This term is nonzero only for initial and final states of the form $|g, n\rangle$ and $|e, n-1\rangle$, respectively. The transition probability is proportional to $|\langle e, n-1 | [\mathbf{a}_{\vec{k},\lambda} | e \rangle \langle g |] |g, n\rangle|^2 = n$. The second term reduces the atom to a lower energy state whereas a photon is generated. This term represents emission and it is non-zero only when the initial and final states are of the form $|e, n\rangle$ and $|g, n+1\rangle$, respectively. The square magnitude of this term is $|\langle g, n+1 | [\mathbf{a}_{\vec{k},\lambda}^\dagger | g \rangle \langle e |] |e, n\rangle|^2 = n+1$. Thus the absorption and emission coefficients end up being

$$\frac{s^2 n}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 - \omega_k)t}{2} \right) \quad (183)$$

$$\frac{s^2 (n+1)}{\hbar^2} t^2 \text{sinc}^2 \left(\frac{(\omega_0 - \omega_k)t}{2} \right). \quad (184)$$

The last two processes, involving the terms in (181) and (182), represent highly improbable events because of the very large argument inside of the sinc function. These processes have negligible probability and so they can be ignored in our treatment. Note that the expressions in (183) and (184) are practically identical to the semiclassical result (138). The only difference is that while semi-classically, the probabilities of absorption and emission are identical, here the ratio between them is

$$\frac{n}{n+1},$$

so that the probability of photon emission is higher. This difference is negligible when the number of photons in the field is large, but it is critical at small values of n . In particular, in the absence of a field, for $n = 0$, the atom has a probability of relaxing from the excited level $|e\rangle$ to the lower level $|g\rangle$ while producing a photon. This event is referred to as *spontaneous emission* and it is impossible in the semi-classical description, where the excited state of the atom is stationary. Thus, only the fully quantum description explains the process of spontaneous emission, which prior to field quantization, could only be treated empirically.

In principle, if one is able to generate a system where an atom interacts with a single mode of the electromagnetic field, or with an exceedingly small number of modes, the process of spontaneous emission is reversible. In such a system the atom excites the field in that single mode, but then the excitation is returned to the atom. The atom and the single mode field oscillate jointly such that the overall energy is conserved. These oscillations resemble the semiclassical Rabi oscillation phenomena, except that there is no externally excited electric field. A full analysis of this scenario is known by the name of the Jaynes Cummings model. In all common situations in nature (not including special and very well controlled experiments), this scenario is not very relevant as the number of electromagnetic field modes that interact with the atom is extremely large. Then, the process of spontaneous emission is irreversible and an atom, after relaxing into its ground state releases a

photon into one of an infinite number of modes. It then stays in its ground state forever, or until it is excited externally.

The intensity operator

In classical optics it is known that the photocurrent generated in a photo-detection device in response to light impinging upon it is proportional to the light intensity. The intensity is the magnitude of the pointing vector and for quasi-monochromatic waves that have a well defined direction of propagation, it is given by

$$I(\vec{r}, t) = 2\varepsilon_0 c \vec{E}^*(\vec{r}, t) \cdot \vec{E}(\vec{r}, t) = 2\varepsilon_0 c |E(\vec{r}, t)|^2 \quad (185)$$

where $\vec{E}(\vec{r}, t)$ is the "analytic signal" of the electric field. The analytic signal is defined as the positive frequency part of the signal. In other words, for a true physical signal $f(t)$, whose Fourier transform is

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt,$$

the analytic signal $f_a(t)$ is given by

$$f_a(t) = \frac{1}{2\pi} \int_0^{\infty} \tilde{f}(\omega) e^{-i\omega t} d\omega,$$

namely, it consists only of the positive frequency components of the original signal. By definition $f_a(t)$ is a complex quantity⁵³ and it is related to the true signal through

$$f(t) = 2\Re\{f_a(t)\}.$$

In the case of narrow-band signals, i.e. signals whose bandwidth is much narrower than their central frequency, the analytic signal is given by

$$f_a(t) = f_{c.e}(t) e^{-i\omega_0 t},$$

⁵³ Its real and imaginary parts are related through the Kramers Kroening relations.

where ω_0 is the central frequency and $f_{c.e}(t)$ is the complex envelope (base-band) signal whose spectrum is centered around 0. This point is elaborated on further in Appendix II.

The quantum equivalent of the analytic signal $\vec{E}(\vec{r}, t)$ appearing in Eq. (185), is the positive frequency part of the electric field operator

$$\vec{\mathbf{E}}^+(\vec{r}, t) = i \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{2\varepsilon_0 V} \right)^{\frac{1}{2}} \hat{e}_{\vec{k}, \lambda} \mathbf{a}_{\vec{k}, \lambda} e^{i(\vec{k} \cdot \vec{r} - \omega_k t)}, \quad (186)$$

whereas the negative frequency part is

$$\vec{\mathbf{E}}^-(\vec{r}, t) = -i \sum_{\vec{k}, \lambda} \left(\frac{\hbar \omega_k}{2\varepsilon_0 V} \right)^{\frac{1}{2}} \hat{e}_{\vec{k}, \lambda} \mathbf{a}_{\vec{k}, \lambda}^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)}, \quad (187)$$

and it is the Hermitian conjugate of $\vec{\mathbf{E}}^+$. The quantum equivalent of the classical definition of intensity then appears very naturally as

$$\mathbf{I}(\vec{r}, t) = 2\varepsilon_0 c \mathbf{E}^-(\vec{r}, t) \mathbf{E}^+(\vec{r}, t), \quad (188)$$

Note that the ordering of \mathbf{E}^- and \mathbf{E}^+ in Eq. (188) is critical (unlike in the classical case). Creation operators appear to the left of annihilation operators, a situation that is conventionally referred to as *normal ordering*. The quantum intensity operator $\mathbf{I}(\vec{r}, t)$ represents a measurable quantity and the photocurrent generated in a photo-detecting device is proportional to it⁵⁴. Expression (188) for the quantum intensity operator is in fact quite plausible when we recall that the process of light detection involves the absorption of photons by atoms. This is just what we have analyzed in the previous section, where we have seen that the probability amplitude that corresponds to a transition from an initial state $|g, n_k\rangle$ to a final state $|e, n_k - 1\rangle$ is proportional to the expectation value of the interaction Hamiltonian

⁵⁴ Expression (188) is justified only in an approximate form, justified in the case where the spatial bandwidth (the angular opening of the detected field) is small. This limitation is, however, identical to the classical case.

$$\langle e, n_k - 1 | \vec{\mathbf{D}} \cdot \vec{\mathbf{E}} | g, n_k \rangle = \langle e | \vec{\mathbf{D}} | g \rangle \cdot \langle n_k - 1 | \vec{\mathbf{E}}^+ | n_k \rangle$$

where we have used the fact that only $\vec{\mathbf{E}}^+$ which consists of annihilation operators has a nonzero matrix element of the needed kind. The term $\langle e | \vec{\mathbf{D}} | g \rangle$ is a property of the detector, whereas the term $\langle n_k - 1 | \vec{\mathbf{E}}^+ | n_k \rangle$ is totally independent of it. In the case of photo-detection the excited level should in fact be viewed as a continuum to which the electron of the atom is excited by the photon so that it contributes to the photocurrent. Once again, the details are of little concern to us in this course. The probability of transition is the square of the probability amplitude and it is proportional to

$$\langle n_k | \vec{\mathbf{E}}^- | n_k - 1 \rangle \cdot \langle n_k - 1 | \vec{\mathbf{E}}^+ | n_k \rangle = \langle n_k | \vec{\mathbf{E}}^- \cdot \vec{\mathbf{E}}^+ | n_k \rangle$$

where the last step can be verified directly, by showing that the same result follows with and without the $|n_k - 1\rangle\langle n_k - 1|$ term⁵⁵. In the more general case, the generated photocurrent will be proportional to

$$\langle \psi | \vec{\mathbf{E}}^- \cdot \vec{\mathbf{E}}^+ | \psi \rangle,$$

where $|\psi\rangle$ denotes the quantum state of the light beam.

Let us now return to the special case that we introduced in the previous section, where the quantum state of the field confined to propagate along the z axis was described by

$$|\psi\rangle = \sum_{\omega} \tilde{f}(\omega) |1_{\omega}\rangle = \sum_{\omega} \tilde{f}(\omega) \mathbf{a}^{\dagger}(\omega) |0\rangle$$

where the usefulness of the rightmost form will become clear imminently. We now wish to calculate the photo-current for the above given state of the field. We write down $\mathbf{E}^+(\vec{z}, t)$ as follows

⁵⁵ In fact, the simplest way of doing that is to replace $|n_k - 1\rangle\langle n_k - 1|$ by $\mathbf{1} = \sum_m |m\rangle\langle m|$. This will not affect the result, because all terms for which $m \neq n_k - 1$ do not contribute anything to the expression.

$$\vec{\mathbf{E}}^+(z, t) = i \sum_{\omega} \left(\frac{\hbar\omega}{2\varepsilon_0 V} \right)^{\frac{1}{2}} \mathbf{a}(\omega) e^{-i\omega(t-z/c)},$$

Assuming, as is almost always the case, that the band of frequencies is very narrow relative to the central frequency of the signal, we may approximate the above by

$$\vec{\mathbf{E}}^+(\vec{z}, t) = i \left(\frac{\hbar\omega_0}{2\varepsilon_0 V} \right)^{\frac{1}{2}} \sum_{\omega} \mathbf{a}(\omega) e^{-i\omega(t-z/c)},$$

with ω_0 being the central frequency of the signal. The expected photocurrent will be proportional to

$$c \frac{\hbar\omega_0}{V} \sum_{\omega, \omega', u, u'} \tilde{f}^*(\omega) \tilde{f}(\omega') e^{-i(u'-u)(t-z/c)} \langle 0 | \mathbf{a}(\omega) \mathbf{a}^\dagger(u) \mathbf{a}(u') \mathbf{a}^\dagger(\omega') | 0 \rangle.$$

We now use the commutation relations to write down

$$\begin{aligned} \left(\mathbf{a}(\omega) \mathbf{a}^\dagger(u) \right) \left(\mathbf{a}(u') \mathbf{a}^\dagger(\omega') \right) &= \left(\delta_{\omega, u} + \mathbf{a}^\dagger(u) \mathbf{a}(\omega) \right) \left(\delta_{\omega', u'} + \mathbf{a}^\dagger(\omega') \mathbf{a}(u') \right) \\ &= \delta_{\omega, u} \delta_{\omega', u'} + \delta_{\omega, u} \mathbf{a}^\dagger(\omega') \mathbf{a}(u') + \delta_{\omega', u'} \mathbf{a}^\dagger(\omega) \mathbf{a}(u) + \mathbf{a}^\dagger(u) \mathbf{a}(\omega) \mathbf{a}^\dagger(\omega') \mathbf{a}(u') \end{aligned}$$

The contribution of the latter 3 terms is 0, whereas the contribution of the first term comes down to

$$c \frac{\hbar\omega_0}{V} \left| \sum_{\omega} \tilde{f}(\omega) e^{i\omega(t-z/c)} \right|^2 = c \frac{\hbar\omega_0}{V} |f(t-z/c)|^2.$$

with $f(t)$ being the inverse Fourier transform (or the discrete inverse Fourier transform) of $\tilde{f}(\omega)$. We thus find that our single photon has a detectable intensity waveform that is proportional to $|f(t)|^2$.

Classical sources of light

As we have seen in the description of the mechanical SHO, not all possible states coincide with the behavior that one observes when looking at a classical, macroscopic SHO. We then showed that the pure state that behaves very

much like what we see in the classical picture is the coherent state, which we denote by $|\alpha\rangle$. The coherent state can be expressed in terms of the number states as

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (189)$$

We also showed that the coherent state $|\alpha\rangle$ is an eigenstate of the annihilation operator \mathbf{a} . You have a homework exercise to show that $|\alpha\rangle$ can also be expressed as⁵⁶

$$e^{(\alpha\mathbf{a}^\dagger - \alpha^*\mathbf{a})} |0\rangle = \mathbf{D}(\alpha) |0\rangle \quad (190)$$

where the operator, $\mathbf{D}(\alpha)$ is called the displacement operator because, as we have seen earlier, the coordinate and momenta of a coherent state α are distributed just as they are for the vacuum state ($\alpha = 0$) except that their mean values are displaced to the real and imaginary parts of α , respectively.

Now is the time when the importance of α can be appreciated. The claim is that a laser, which can be viewed as a classical source of deterministic light (as opposed to light that fluctuates in a noisy manner) generates light in a coherent state⁵⁷. In other words, what we observe at the output of a very good laser is the quantum state $|\alpha\rangle_{\omega_0} \otimes |0\rangle_{\omega \neq \omega_0}$, with ω_0 denoting the laser frequency. Here we consider a laser to be perfect in the sense that its linewidth is 0. In reality the laser has some linewidth, and so it emits into several frequency modes. If we discretise the modes as we usually do, then the output of a laser can be expressed as

$$|\alpha(\omega_1)_{\omega_1}, \alpha(\omega_2)_{\omega_2}, \dots\rangle,$$

⁵⁶ To do that, use the Baker Hausdorf formula to write $D(\alpha)$ as $e^{-|\alpha|^2} e^{i\alpha\mathbf{a}^\dagger} e^{i\alpha^*\mathbf{a}}$.

The rightmost term does not do anything to the vacuum, whereas the Taylor expansion of the middle term generates Eq. (189).

⁵⁷ Although, it is not very difficult to show that classical (large) currents generate coherent states, I will avoid doing that because it requires a detour in which certain relations and operators need to be defined.

where each mode is excited to a coherent state. This is the same as having many different zero-linewidth lasers, one at each frequency⁵⁸. A more convenient notation is just to write the above as⁵⁹ $|\alpha(\omega)\rangle$, which implies multiple excitations of all frequency modes ω , each is excited into a coherent state characterized by the parameter $\alpha(\omega)$. Note that these concepts and their various notations can be fairly confusing. It should be noted, for example, that the state $|\alpha(\omega_1)_{\omega_1}, 0_{\omega_2}\rangle + |0_{\omega_1}, \alpha(\omega_2)_{\omega_2}\rangle$ is not the same as $|\alpha(\omega_1)_{\omega_1}, \alpha(\omega_2)_{\omega_2}\rangle$. In fact the former state can by no means be considered classical, as it would generate very "un-classical" observations and it is very difficult to generate in the laboratory. A special property that characterizes coherent states is that they behave classically in all cases. Thus, when we generate a coherent state, we can predict all measurement results in terms of classical optics. Even quantum related questions can be answered classically for coherent states. An important example is that of the photo-detection process. Let us consider for example, the average intensity of light in a coherent state

$$\langle \alpha(\omega) | \mathbf{E}^- \mathbf{E}^+ | \alpha(\omega) \rangle = \left| \sum_{\omega} \alpha(\omega) e^{-i\omega(t-z/c)} \right|^2 = |\alpha(t - z/c)|^2$$

with $\alpha(t)$ representing the inverse Fourier transform of $\alpha(\omega)$ (we usually use a tilde to distinguish between functions in the time and frequency domains, but here the correct interpretation is clear from the context). Thus, if we consider a classically deterministic pulse whose slowly varying envelope is $\alpha(t)$, the intensity—measured as current in a photo-detecting device—will be just what we expect.

The definition of "classical states" as coherent states is a bit too narrow. Thermal light that is generated by the sun, or by a light-bulb is also quite

⁵⁸ Note that modes that are not excited, can also be viewed as coherent states corresponding to $\alpha = 0$.

⁵⁹ Try not to be confused by the difference with respect to the notation $|1_{\omega}\rangle$ that we used earlier to describe a single photon at only one mode indexed by the frequency ω .

classical, as the rules of classical optics apply to it very accurately. A single mode (single frequency component) of thermal light is described by a density matrix

$$\rho = \sum p(n) |n\rangle \langle n|,$$

with $p(n) = \frac{\langle n \rangle^n}{(1 + \langle n \rangle)^{n+1}}$, as we have seen earlier. We have also seen earlier (exercise set 4) that the density matrix of a thermal mode can be described as

$$\rho = \frac{1}{\pi N_0} \int d^2\alpha e^{-\frac{|\alpha|^2}{N_0}} |\alpha\rangle \langle \alpha|,$$

which is nothing but a statistical mixture of coherent states, each with a probability $p(\alpha) = (\pi N_0)^{-1} e^{-\frac{|\alpha|^2}{N_0}} d^2\alpha$.⁶⁰ Thus, generating thermal radiation is just the same as generating coherent radiation whose parameter is picked at random from a Gaussian distribution.

By analogy, we can now generalize by saying that an optical system is in a classical state whenever its density operator can be expressed as

$$\rho = \int d^2\alpha P(\alpha) |\alpha\rangle \langle \alpha|,$$

with $P(\alpha)$ being a nonnegative function of α . Such states can always be generated by properly randomizing the amplitude of a perfectly coherent signal. Light that is in a classical state can be described with perfect accuracy using classical optics. States, for which $P(\alpha)$ is non-positive, or for which it doesn't exist are not classical states and they contain features that cannot be explained by classical arguments. A trivial example of a non-classical state is the number state $|n\rangle$ for⁶¹ $n > 0$. If we write

⁶⁰ This implies, of course, that the real and imaginary parts of α are zero-mean Gaussian variables with mean square of $N_0/2$ each, so that the mean of $|\alpha|^2$ is N_0 .

⁶¹ Note that if $n = 0$ the 0 number state, which is the vacuum, is also a coherent state with 0 amplitude.

$$\rho = |n\rangle\langle n| = \int d^2\alpha P(\alpha) |\alpha\rangle\langle\alpha|,$$

we can immediately conclude that $P(\alpha)$ cannot be nonnegative. To see that we may take for example $\langle m|\rho|m\rangle$, which is the m, m matrix component of ρ and whose value is obviously zero when $m \neq n$. We thus have

$$0 = \langle m|\rho|m\rangle = \int d^2\alpha P(\alpha) e^{-|\alpha|^2} \frac{|\alpha|^m}{m!},$$

and since everything except $P(\alpha)$ in the integrand is positive, $P(\alpha)$ must have regions in which it is smaller than 0. We have seen earlier (in the context of the mechanical SHO) and we will see again in what follows that number states are extremely non-classical and the prediction of the results of experiments that use light in a number-state cannot be made based on classical optics considerations.

Examples of handling quantized optical fields

Some of the things that we studied regarding quantized light were pretty abstract. Yet those things are often addressed experimentally, using experimental set-ups that are very familiar to us from laboratories. In this section I want to relate the quantum description to experimental phenomena that are often observed in the lab. The most basic element in all experiments in optics is a beam splitter (BS). It is usually represented by a cube, but a fiber coupler performs the exact same function. Of course a more descriptive name for the BS would be beam splitter and combiner, as it can obviously perform either function. Nonetheless, the name that I will use here is just BS.

The functionality of a BS is easily represented in the classical domain.

$$\begin{pmatrix} E_3 \\ E_4 \end{pmatrix} = \begin{pmatrix} s & ir \\ ir & s \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}, \quad (191)$$

where $E_{1,2}$ are the complex envelopes (or the analytic signals) of the fields entering ports 1 and 2, and where $E_{3,4}$ are the complex envelopes of the fields

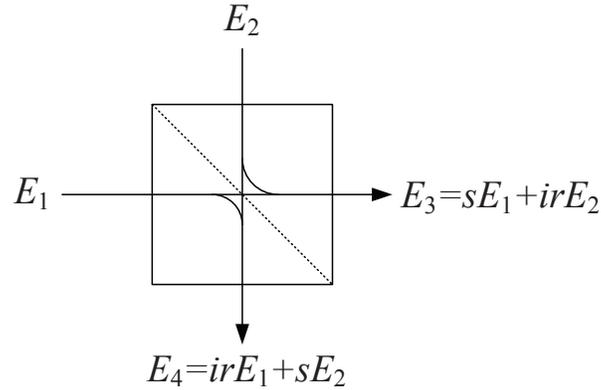


Fig. 2. Schematic of a beam-splitter/combiner.

emitted from ports 3 and 4. In the above and also in what follows –unless stated differently – we treat the fields as scalars. This is legitimate if the BS does not affect polarizations, in other words if the polarizations of the incident fields are maintained at the output, so that orthogonal polarization modes do not mix. The constants s and r are real and they represent the amplitude transmission and reflection coefficients, respectively. Usually the transmission coefficient is denoted by t , but we chose s , letting t represent time. In what follows, we will denote the transfer matrix of scalar fields, such as the one appearing in (191) by $T_M(\omega)$, where for generality, we include the possible dependence of the transfer matrix on frequency. Note that $T_M(\omega)$ is just the two-dimensional equivalent of the transfer function that we are familiar with from the study of linear systems. We assume that only a single mode \vec{k}_1 is incident into port 1 and a single mode \vec{k}_2 is incident into port 2 (in what follows we will identify the modes by the port number, dropping the wave-vector symbol \vec{k} from the index). The outputs correspond to modes \vec{k}_3 and \vec{k}_4 . We will assume that the length of the k vectors are identical, so that there is only one optical frequency involved. This picture can be easily generalized to particular narrow-band pulse-modes, where the pulse shape is fixed. The matrix that connects the inputs with the outputs must be symmetric (because

the BS is a symmetric device) and unitary, for conservation of energy, so that the coefficients s and r satisfy $r^2 + s^2 = 1$. In the quantum picture, the equivalents of the analytic signals are the positive frequency field operators \mathbf{E}^+ , which, for a single mode, are given by expressions of the form

$$\mathbf{E}_j^+ = \left(\frac{\hbar\omega}{2\varepsilon V_0} \right)^{1/2} \mathbf{a}_j e^{-i\omega(t-z/c)}$$

where $j = 1, 2, 3, 4$ correspond to the various BS ports. We are assuming that the frequencies of all ports are identical and equal to⁶² ω . It is quite clear that the quantum equivalent of the slowly varying amplitudes are simply the annihilation operators \mathbf{a}_j , where we ignore the constant coefficient $\left(\frac{\hbar\omega}{2\varepsilon V_0} \right)^{1/2}$, which should have no consequences to our arguments. We may now apply Eq. (191) to the annihilation operators instead of to the classical complex envelopes. It is useful to express the relations explicitly

$$\mathbf{a}_3 = s\mathbf{a}_1 + ir\mathbf{a}_2 \quad (192)$$

$$\mathbf{a}_4 = ir\mathbf{a}_1 + s\mathbf{a}_2 \quad (193)$$

and in the opposite direction

$$\mathbf{a}_1 = s\mathbf{a}_3 - ir\mathbf{a}_4 \quad (194)$$

$$\mathbf{a}_2 = -ir\mathbf{a}_3 + s\mathbf{a}_4. \quad (195)$$

It is important to realize (it is a homework exercise for you to do that) that in this representation, the operators \mathbf{a}_3 and \mathbf{a}_4 do indeed represent independent modes, i.e. such that $[\mathbf{a}_j, \mathbf{a}_{j'}^\dagger] = \delta_{j,j'}$, with j, j' equal to either 3 or 4. Now lets assume a number of specific scenarios. The first scenario that we consider is that where the beams incident upon ports 1 and 2 are in a coherent state.

⁶² If we have to deal with fields of different frequencies then each frequency can be studied separately, as the BS is linear and thus there is no interaction, or coupling between them.

In this case the entire system can be represented by the state-vector⁶³ $|\psi\rangle = |\alpha_1, \beta_2\rangle$. Let us now try to see what the quantum description of the output fields is. To do this most easily, I will need an identity that you were supposed to prove at home

$$|\alpha\rangle = \mathbf{D}(\alpha)|0\rangle.$$

where $\mathbf{D}(\alpha)$ is the *displacement* operator

$$\mathbf{D}(\alpha) = e^{\alpha \mathbf{a}^\dagger - \alpha^* \mathbf{a}}.$$

This unitary operator (showing that it is indeed unitary is an exercise for you) has various properties, the one that we will need most is

$$\mathbf{D}(\alpha)\mathbf{D}(\alpha') = e^{\alpha\alpha'^* - \alpha^*\alpha'} \mathbf{D}(\alpha + \alpha'),$$

where the term $\exp(\alpha\alpha'^* - \alpha^*\alpha')$ is immaterial and can be omitted, as it only affects the state's phase. Let us now represent the input state $|\psi\rangle$ as

$$|\psi\rangle = \mathbf{D}_1(\alpha)\mathbf{D}_2(\beta)|0\rangle \tag{196}$$

where $\mathbf{D}_j(\cdot)$ operates on the j 'th mode (or port) with $j = 1, 2$. We now write these operators explicitly using the relations that were introduced earlier

$$\mathbf{D}_1(\alpha) = e^{\alpha \mathbf{a}_1^\dagger - \alpha^* \mathbf{a}_1} = e^{(s\alpha) \mathbf{a}_3^\dagger - (s\alpha)^* \mathbf{a}_3} e^{(ir\alpha) \mathbf{a}_4^\dagger - (ir\alpha)^* \mathbf{a}_4} = \mathbf{D}_3(s\alpha)\mathbf{D}_4(ir\alpha).$$

Similarly,

$$\mathbf{D}_2(\beta) = \mathbf{D}_3(ir\beta)\mathbf{D}_4(s\beta).$$

⁶³ One could ask the question, how come the entire system is described in terms of the excitation of only modes 1 and 2. In what we have studied previously we said that the excitation of all independent modes in a system must be specified, unless if they are in the vacuum state. So why are we satisfied with the specification of modes 1 and 2 only? The answer is of course that modes 3 and 4 are not independent here. Rather they are related to modes 1 and 2 through Eqs. (193). Therefore, the state of the system can be represented either by specifying the excitation of modes 1 and 2, or the excitation of modes 3 and 4.

Substituting into (196) we obtain

$$\begin{aligned} |\psi\rangle &= \mathbf{D}_1(\alpha)\mathbf{D}_2(\beta)|0\rangle = \mathbf{D}_3(s\alpha + ir\beta)\mathbf{D}_4(ir\alpha + s\beta)|0\rangle \\ &= |(s\alpha + ir\beta)_3, (ir\alpha + s\beta)_4\rangle. \end{aligned} \quad (197)$$

Thus, what we found is that coherent states behave in perfect agreement with the predictions of classical optics. This is in fact the case in all situations, a fact that we will accept in all following examples without proving it for every individual case.

Let us now consider the next scenario, where we look at a single photon that is incident upon port number 1. How do we specify the state of the system in this case? The answer is the following

$$|\psi\rangle = |1_1, 0_2\rangle$$

implying a single photon in mode 1, and no photons in mode 2. We express the state-vector as

$$|\psi\rangle = |1_1, 0_2\rangle = \mathbf{a}_1^\dagger|0\rangle = (s\mathbf{a}_3^\dagger + ir\mathbf{a}_4^\dagger)|0\rangle = s|1_3, 0_4\rangle + ir|0_3, 1_4\rangle.$$

What we find is that from the standpoint of the BS output, the field is in a superposition state between having a photon in port 3 and having a photon in port 4. The probability of finding the photon in port 3 is proportional to the average of the intensity operator in that port $\mathbf{I}_3 = \mathbf{E}_3^- \mathbf{E}_3^+$, which is the same as $\mathbf{a}_3^\dagger \mathbf{a}_3$. Similarly, the probability of detecting the photon at port 4 is the mean of $\mathbf{a}_4^\dagger \mathbf{a}_4$. The probability of detecting two clicks (two photoelectrons) at ports 3 and 4 is the mean of $\mathbf{a}_3^\dagger \mathbf{a}_3 \mathbf{a}_4^\dagger \mathbf{a}_4$. If we calculate those means we find that the probability of observing the photon at port 3 is s^2 , at port 4 it is r^2 and for both ports simultaneously it is 0.

Thus we have found that the single photon state behaves like a particle that cannot be split. It either goes one way, or another, and it cannot be detected in both output arms. This notion, although quite profound, may not strike you as unusual. Most of you have become used to thinking of light

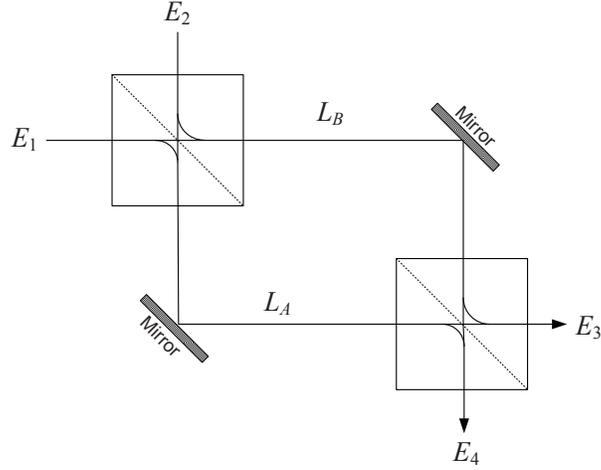


Fig. 3. Schematic of a Mach-Zehnder interferometer (MZI).

as of a stream of discrete particles and in this sense there is no innovation to be found in our last observation. Moreover, a plausible conclusion from our preceding analysis is that the classical intensity gives the probability for the photon to be found in one place, or another. Let us make this scenario a little more complicated by considering the case of a Mach-Zehnder interferometer (MZI) consisting of two beam splitters, as shown in Figure 3. In this case we assume equal symmetric beam splitters with $s = r = 2^{-1/2}$ so that the incoming light is split equally between the two output arms.

The classical description of this device is the following. Assuming that the optical lengths of the two MZI paths are L_A and L_B , the signals at the MZI outputs are given by

$$E_3 = \frac{1}{\sqrt{2}} \left(\frac{iE_1}{\sqrt{2}} + \frac{E_2}{\sqrt{2}} \right) e^{i\omega L_A/c} + \frac{i}{\sqrt{2}} \left(\frac{E_1}{\sqrt{2}} + \frac{iE_2}{\sqrt{2}} \right) e^{i\omega L_B/c},$$

$$E_4 = \frac{i}{\sqrt{2}} \left(\frac{iE_1}{\sqrt{2}} + \frac{E_2}{\sqrt{2}} \right) e^{i\omega L_A/c} + \frac{1}{\sqrt{2}} \left(\frac{E_1}{\sqrt{2}} + \frac{iE_2}{\sqrt{2}} \right) e^{i\omega L_B/c}$$

which can be rewritten in a more compact form as

$$\begin{pmatrix} E_3 \\ E_4 \end{pmatrix} = e^{i\bar{\Phi}} \begin{pmatrix} \cos(\Delta\phi) & \sin(\Delta\phi) \\ -\sin(\Delta\phi) & \cos(\Delta\phi) \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} \quad (198)$$

where $\Delta\Phi = \omega(L_A - L_B)/2c$ is the difference between phases accumulated in the two arms and where $\bar{\Phi} = \omega(L_A + L_B)/2c$ is the average phase delay. The average phase $\bar{\Phi}$ has no physical consequences, neither in the classical, nor in the quantum description, and it will therefore be omitted in what follows. In what follows we will refer to the 2×2 matrix in Eq. (198) as T_M . The MZI is a very basic and fundamental device in the optical sciences with numerous applications. For acquiring some insight consider the case in which only input port number 1 is excited (as is the case in many practical applications). In this case the outputs are $E_3 = E_1 \cos(\Delta\Phi)$ and $E_4 = E_1 \sin(\Delta\Phi)$. Clearly the sum of their squares is equal to $|E_1|^2$, demonstrating the conservation of energy (the transfer matrix is unitary and therefore this property is of course maintained in general, i.e. when both input ports are excited). The case of a balanced MZI, where $L_A = L_B$, such that $\Delta\Phi = 0$ is of particular relevance to our following discussion, since then $E_4 = 0$ and all the optical power comes out of port number 3.

We will now return to the quantum picture. If the inputs are excited by coherent states $|\psi\rangle = |\alpha_1, \alpha_2\rangle$, the output states will be $|\alpha_3, \alpha_4\rangle$, where the relations between α_3, α_4 and α_1, α_2 are given by equation (198) with α_j replacing E_j ($j = 1, 2, 3, 4$) in all places. In order to prove this, one could use the same procedure as the one that led to Eq. (197). More interesting is the case in which a single photon is incident in port number 1, i.e. $|\psi\rangle = |1_1, 0_2\rangle$. In this case we may be tempted to argue (based on our discussion of the BS) that following the first BS, the photon goes either through the top, or the bottom arm of the MZI with equal probability. Then the same happens when it enters the second BS so that the probability of exiting through either output arm is $1/2$ regardless of the path that the photon took in the MZI⁶⁴. As we will now see, this conclusion is drastically inconsistent with the actual reality

⁶⁴ If this were true, it would be in direct contradiction to the classical case. There, we have seen for example that when the MZI is balanced no light at all should come out of port number 4!

and in fact the photon will exit through ports 3 and 4 with probabilities that are proportional to the classical intensities, i.e. to $\cos^2(\Delta\Phi)$ and to $\sin^2(\Delta\Phi)$, respectively. This is rather reassuring on the one hand, because it implies that our classical intuition is still of some use⁶⁵, but on the other hand it is quite disturbing because the photon, whom we have earlier identified as an unsplittable particle, now seems to have split into two parts that interfered with each other!

To rigorously evaluate the probabilities that the photon comes out of ports 3 and 4 we write the operator \mathbf{a}_1^\dagger in terms of \mathbf{a}_3^\dagger and \mathbf{a}_4^\dagger , as we did earlier. Note that since T_M is a real and unitary matrix, its inverse is equal to its transpose. The needed relation between the quantum operators is

$$\mathbf{a}_1^\dagger = \cos(\Delta\phi)\mathbf{a}_3^\dagger - \sin(\Delta\phi)\mathbf{a}_4^\dagger,$$

which implies that

$$|1_1, 0_2\rangle = \mathbf{a}_1^\dagger|0\rangle = \cos(\Delta\phi)|1_3, 0_4\rangle - \sin(\Delta\phi)|0_3, 1_4\rangle,$$

which indeed confirms the fact that for $\Delta\phi = 0$, the photon will never exit through port 4.

This wave-particle duality of the photon is a concept that puzzled scientists quite a lot in the very early days of quantum mechanics. Of course it relates not only to photons, but to all particles in general, except that similar experiments with massive particles are usually much more difficult to perform. The attempt to ask particle-like questions about photons always leads to very strange and incorrect conclusions. The absence of "which path" information is at the heart of quantum physics and nature in general. The assertion that the photon, as a particle that cannot be split, must have gone through either the top path, or the bottom path, leads to many weird and incorrect conclusions. The reality is that the photon behaves as a wave and as a particle at exactly the same time. Imagine a situation where one of the two MZI paths is blocked. Then if the photon goes into the blocked arm it

⁶⁵ There will be other cases in which even this statement will no longer be true.

is absorbed and not observed at either output. One may correctly conclude that if it is observed at one of the outputs it must have gone through the un-blocked arm. But then the probability of observing it at either output is 50% as is the case in measurements. One can then ask the following question: if the photon chose to go through the unblocked arm then it never reached the obstacle that blocked that arm so how does it know that it is blocked? In other words one is able to tell that one arm of the MZI is blocked by an obstacle, without having the photon (or anything else) ever being in direct contact with that obstacle⁶⁶. To continue along the same line of thought, one may ask what would happen if one puts sensors that indicate with certainty whether the photon passed through a given arm, or not. The answer is that then interference will disappear altogether and the photon will behave as a particle. What if the sensors can detect the path of the photon with a limited probability? Meaning that an observer can tell which way the photon went, but only with a limited certainty. The answer is (and you should be able to show it yourselves) that the degree of interference will reduce proportionally to the certainty with which the “which path” information is known. The experimental manifestation of this will be reduction in the visibility of the fringes that are observed when the optical path difference $\Delta\Phi$ is changed^{67,68}.

It is interesting to consider an example that shows how one could obtain complete, or partial information on the path through which the photon went inside the MZI. Assume a MZI that a polarization rotator was introduced into one of its arms. Assuming that the incident single-photon field is in the

⁶⁶ This apparent paradox, with the obstacle being a bomb, has been discussed by Elitzur and Veidman.

⁶⁷ When the length of one of the arms is varied, $\Delta\Phi$ changes accordingly and the probability of observing the photon at a given output oscillates sinusoidally between a minimum value p_{min} and a maximum value p_{max} . The visibility is defined as $V = (p_{max} - p_{min}) / (p_{max} + p_{min})$. If there is complete destructive interference for certain values of $\Delta\Phi$ then $p_{min} = 0$ and $V = 1$. Otherwise, when interference is not complete, $p_{min} > 0$ and the value of V is reduced.

⁶⁸ This issue is famously discussed in the writings of Richard Feynman.

\vec{I}_x polarized mode, the rotator changes its polarization to $\cos(\theta)\vec{I}_x + \sin(\theta)\vec{I}_y$, with θ being a controlled angle. If θ is equal to $\pi/2$ the photon's polarization is completely flipped. Then, by observing the polarization at the output, one could tell without ambiguity whether the photon passed through the arm that contained the polarization rotator, or not. In this case interference is lost completely, and the photon emerges out of the two output ports with equal probabilities. In an intermediate case, when $0 < \theta < \pi/2$ one cannot tell for sure which way the photon went. That is because at the output its polarization may still turn out to be \vec{I}_x , even if it passed through the rotator. Then interference is not lost completely and the visibility is greater than 0, but smaller than 1.

We may conclude that single photons are seen to be dual entities that on some occasions behave as waves, whereas on others they behave as solid particles. We saw that they behave as waves whenever the path that they take towards their final destination (the measurement apparatus) is fundamentally unknown. When the path is known, they behave as particles. Yet, in all of the above examples, the probability of detecting the photon at a given output port is proportional to the intensity that would be detected in that port according to a classical calculation. Not all cases are like that, and in some situations, coming up with a classical analog is essentially impossible, as we will see in the following example.

The Hong, Ou and Mandel experiment

Consider the set-up shown in Fig. 4. It is the basis for the famous experiment⁶⁹, originally performed by Hong, Ou and Mandel (HOM) in 1987. Two identical photons are injected simultaneously into the two input arms of a BS at the output of which two detectors are placed. The quantum state at the input is $|\psi\rangle = |1_1, 1_2\rangle$, using the same machinery that we had above one can easily show that the output state is $2^{-1/2}[|2_3, 0_4\rangle + |0_3, 2_4\rangle]$. In other words, the two photons bunch and come out together either through port 3,

⁶⁹ C.K. Hong et. al., Phys. Rev. Lett., vol. 59, p. 2044, 1987.

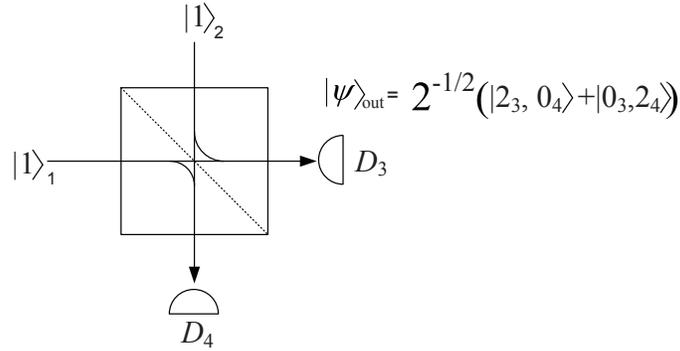


Fig. 4. Schematic of the HOM set-up.

or through port 4. There is never a situation where photons are simultaneously measured in both ports 3 and 4. They are always bunched and reach only one of these two ports. This situation is impossible classically, and indeed it does not happen if the states at the BS inputs are coherent⁷⁰ $|\alpha_1, \beta_2\rangle$. One may ask what would happen for example, if the photons incident upon ports 1 and 2 are delayed relative to each other. Or, how much can they be delayed before the HOM interference disappears? Answering this question was the goal of the original HOM experiment. They plotted the coincidence rate as a function of the delay between the two photons. At perfect overlap the coincidence rate fell to 0, whereas for large delay it stabilized on the incidental coincidence value of 1/2. The width of the dip can be interpreted as the temporal width of the photon in the experiment. It teaches on the properties of the source that was used to create the two identical photons.

Entanglement and The Einstein Podolsky and Rosen paradox

Perhaps the most famous paradox in quantum physics is the one that Einstein Podolsky and Rosen published in 1935. It has to do with the concept of

⁷⁰ In fact, if $\alpha = \pm i\beta$, an output signal will be registered only in one of the two BS outputs. The output into which light will come out is however deterministic and it depends on whether the phase difference between α and β is $\pi/2$, or $-\pi/2$. But with single photons, as in the HOM set-up, there is no phase to be set.

Entanglement which is unique to quantum physics and it implies an apparent violation of the principle of locality. In this section we will explain the concept of entanglement and discuss the issues raised by EPR. We will rely on the simplest example of entanglement that involves polarization entangled photons. This scheme of entanglement is also the one that was studied experimentally in the context of the EPR paradox.

Polarization of single photons

Before delving into the heart of the entanglement problem, we introduce the tools required for describing polarized photons. The specification of a mode of the electromagnetic field involves the specification of a wave vector⁷¹ \vec{k} and a polarization state, which we have denoted as λ . Using the lab frame of reference, we may expand all polarization modes in terms of the horizontal polarization mode H and a vertical polarization mode V . A horizontally, or vertically polarized single photon is denoted by $|1_{\vec{k},H}\rangle$, or $|1_{\vec{k},V}\rangle$, respectively. An alternative, and often more convenient, notation of the two photons would be $|H_{\vec{k}}\rangle$, or $|V_{\vec{k}}\rangle$, respectively. Similarly, a photon linearly⁷² polarized along some arbitrary angle θ from the H orientation in the H, V plane can be denoted as $|\theta_{\vec{k}}\rangle$. Since we are only dealing with one particular spatial mode, the specification of \vec{k} is superfluous and we may avoid it in what follows. Let us now define the relation between $|H\rangle$, $|V\rangle$ and $|\theta\rangle$. Recall that one may express either of these state-vectors as $\mathbf{a}_x^\dagger|0\rangle$ with $x = H, V$, or θ , depending on the photon that we wish to describe. Defining the desired relations between the state-vectors reduces to expressing the relations between the

⁷¹ We have seen that one can relax the definitions and refer to a packet of modes with close values of \vec{k} as an individual mode, that can be viewed as a quantum SHO whose frequency of oscillations is the central frequency of the packet.

⁷² Since knowledge of optics and of the math of polarizations is not a mandatory prerequisite, I am referring only to the simplest and most intuitive case of linearly polarized light. Those of you who have a background in optics can trivially generalize all our examples and conclusions to elliptic polarizations.

creation operators, which follow from the very definition of the modes of the electromagnetic field. As you may recall, for a given k -vector one defines two orthogonal polarization modes of the electromagnetic field, directed along \hat{e}_1 and \hat{e}_2 . The orientations of these two unit polarization vectors are chosen arbitrarily, with the only condition being orthogonality, i.e. $\hat{e}_1 \cdot \hat{e}_2 = 0$. Let us consider two particular bases $\{\hat{e}_H, \hat{e}_V\}$ and $\{\hat{e}_\theta, \hat{e}_{\theta'}\}$ where $\theta' = \theta + \pi/2$ such that $\hat{e}_\theta \cdot \hat{e}_{\theta'} = 0$. As usual, the electric fields (and hence the annihilation operators) satisfy the relations

$$\begin{aligned}\mathbf{a}_H &= \cos(\theta)\mathbf{a}_\theta - \sin(\theta)\mathbf{a}_{\theta'} \\ \mathbf{a}_V &= \sin(\theta)\mathbf{a}_\theta + \cos(\theta)\mathbf{a}_{\theta'},\end{aligned}$$

and in the opposite direction

$$\begin{aligned}\mathbf{a}_\theta &= \cos(\theta)\mathbf{a}_H + \sin(\theta)\mathbf{a}_V \\ \mathbf{a}_{\theta'} &= -\sin(\theta)\mathbf{a}_H + \cos(\theta)\mathbf{a}_V.\end{aligned}$$

Thus, the state $|H\rangle$ and $|V\rangle$ can be written as⁷³

$$\begin{aligned}|H\rangle &= \cos(\theta)|\theta\rangle - \sin(\theta)|\theta'\rangle \\ |V\rangle &= \sin(\theta)|\theta\rangle + \cos(\theta)|\theta'\rangle,\end{aligned}$$

or equivalently,

⁷³ Notice that if we rewrite the relation between $|\theta\rangle$ and $|H\rangle, |V\rangle$ in the original notation that we introduced, we have

$$|1_\theta, 0_{\theta'}\rangle = \cos(\theta)|1_H, 0_V\rangle + \sin(\theta)|0_H, 1_V\rangle,$$

where the right-hand-side cannot be written as a product state (as opposed to the left-hand-side $|1_\theta, 0_{\theta'}\rangle = |1\rangle_\theta \otimes |0\rangle_{\theta'}$). As we will see in what follows this property of not being able to express a state as a product is known as entanglement. What we have here is entanglement between modes (i.e between SHOs), i.e. it has to do with how our modes are defined. This is not entanglement between different photons, which is what we will encounter in what follows.

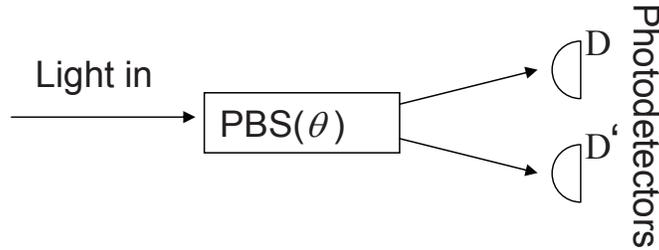


Fig. 5. Schematic of a polarization analyzer. The PBS is a polarization beam splitter that separates the incident beam into orthogonal polarization components which are detected by the two photo-detectors D and D' .

$$|\theta\rangle = \cos(\theta)|H\rangle + \sin(\theta)|V\rangle$$

$$|\theta'\rangle = -\sin(\theta)|H\rangle + \cos(\theta)|V\rangle.$$

An analogous, but more general relation allows consideration of all orthogonal polarization states (including elliptical polarizations) as legitimate modes of the electromagnetic field in the overall treatment.

Polarization measurement

A polarization analyzer consists of a polarization splitting device followed by two detectors as illustrated in Fig. 5. The polarization splitting device is usually a birefringent crystal (such as Calcite or YVO4) that spatially separates the incident optical signal into two orthogonal components that are defined by the principal axes of the crystal. For example, if the principle axes are aligned with the H and V orientations then the H and V components of the incident light are separated into two separate beams that are shined onto the two detectors D and D' . Since the birefringent crystal can be rotated at will⁷⁴, the polarization states of the beams at its output can be referred to as θ and $\theta' = \theta + \pi/2$, where we use our previously defined notation. If the field whose polarization is to be analyzed consists of a single photon, then the photon (being a non-dividable entity) collapses either onto the $|\theta\rangle$ state,

⁷⁴ Obviously, instead of rotating the birefringent crystal, one can equivalently introduce a polarization rotator in front of it.

or the $|\theta'\rangle$ state, generating a click in either detector D , or detector D' , respectively. Measurement of polarizations in the single photon case therefore comes down to choosing the angle of the polarization splitter and recording which of the two detectors clicked⁷⁵.

An example of entanglement between a pair of photons

The ideas of the previous paragraph are easily expanded to the case of two photons. The short-hand notation for two polarized photons will be for example $|\psi\rangle = |H, V\rangle$, or $|\psi\rangle = |\theta_1, \theta_2\rangle$, where $\theta_{1,2}$ define the polarization angles of the two photons relative to the reference direction H (which needs not be the same for the two photons). The state $|\psi\rangle$ that we have written above is just a simple example of a case in which each photon has a well defined polarization state. Yet, as we have seen earlier, superpositions of states are also legitimate quantum states and thus we may consider for example the state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}[|H, V\rangle - |V, H\rangle]. \quad (199)$$

This state is special, in the sense that it cannot be written as a product of two states each corresponding to one of the particles⁷⁶. Not all states that contain multiple terms are like that. For example the state $|\psi\rangle = 2^{-1}[|HH\rangle - |HV\rangle + |VH\rangle - |VV\rangle]$ can be written as $|\psi\rangle = 2^{-1}[|H\rangle_1 + |V\rangle_1] \otimes [|H\rangle_2 - |V\rangle_2]$, as can be readily checked. States that cannot be written as a product are said to be *entangled*. Entangled states have very interesting and unusual properties. Consider the state $|\psi^-\rangle$ that is written above and assume that photons 1 and

⁷⁵ Notice that the quantum polarization measurement only gives a “yes” or “no” answer to the question “Is the photon polarization θ ?”, where θ is the angle to which the measuring device has been set.

⁷⁶ In a later section we will learn about the Schmidt decomposition, that implies that any state of the two polarized photons can be expressed as $q_a|\theta_1, \theta_2\rangle + q_b|\theta'_1, \theta'_2\rangle$, with $|q_a|^2 + |q_b|^2 = 1$. If a state can be expressed as a product, then only one of the coefficients q_a and q_b will be non-zero.

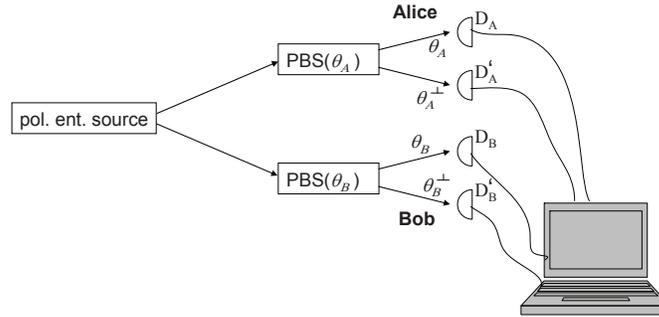


Fig. 6. Schematic of the EPR measurement set-up with polarization entangled photons.

2 are well separated in space. Typically, one talks about one photon being sent to one user named Alice, and the other photon sent to another user whose name is Bob. Alice and Bob each have a polarization analyzer which they use in order to perform a polarization measurement as depicted in Fig. 6, where we assume that Alice and Bob's polarization analyzers are aligned with the angles θ_A and θ_B , respectively. We will now assign the variables $A(\theta_A)$ and $B(\theta_B)$ to the outcomes of Alice and Bob's measurements according to the following procedure: If detector D_A in Alice's receiver clicks, we say that $A(\theta_A) = 1$, otherwise, if detector D'_A clicks we say that $A(\theta_A) = -1$. Similarly the variable $B(\theta_B)$ receives the values 1 and -1 , depending on which of the detectors D_B , or D'_B clicks, respectively.

Note that $A(\theta_A)$ and $B(\theta_B)$ are random variables in which θ_A and θ_B are merely parameters indicating the alignment angle of the measurement apparatus. We may ask ourselves what would be the probability that the two top detectors click for a given photon-pair. This is the probability that both $A(\theta_A) = 1$ and $B(\theta_B) = 1$, which we denote by⁷⁷

$$P_{\theta_A, \theta_B}(1_A, 1_B) = |\langle \theta_A, \theta_B | \psi \rangle|^2$$

⁷⁷ Note that this is not the probability of having one photon in mode "A" and one photon in mode "B", which would have a similar notation.

Noting that

$$\begin{aligned} |\theta_A, \theta_B\rangle &= \cos(\theta_A) \cos(\theta_B) |HH\rangle + \sin(\theta_A) \sin(\theta_B) |VV\rangle \\ &+ \cos(\theta_A) \sin(\theta_B) |HV\rangle + \sin(\theta_A) \cos(\theta_B) |VH\rangle, \end{aligned}$$

we find that

$$\begin{aligned} P_{\theta_A, \theta_B}(1_A, 1_B) &= \frac{1}{2} |\cos(\theta_A) \sin(\theta_B) - \sin(\theta_A) \cos(\theta_B)|^2 \\ &= \frac{1}{2} \sin^2(\theta_B - \theta_A). \end{aligned} \quad (200)$$

From the above result, it is evident that if Alice and Bob's polarization analyzers are aligned identically $\theta_A = \theta_B$ (doesn't matter at what angle with respect to H and V) the detectors D_A and D_B will never click simultaneously. To find the probability $P_{\theta_A, \theta_B}(-1_A, -1_B)$ it is sufficient to replace θ_A and θ_B in (200) with $\theta_A + \pi/2$ and $\theta_B + \pi/2$, respectively. As one would expect, this procedure indicates that $P_{\theta_A, \theta_B}(-1_A, -1_B) = P_{\theta_A, \theta_B}(1_A, 1_B)$ so that when the polarization analyzers are aligned ($\theta_A = \theta_B$) simultaneous clicks of detectors D'_A and D'_B are also impossible. The probability $P_{\theta_A, \theta_B}(1_A, -1_B)$ is obtained by replacing θ_B with $\theta_B + \pi/2$ and it is equal to $P_{\theta_A, \theta_B}(-1_A, 1_B)$, which is obtained when θ_A in (200) is replaced with $\theta_A + \pi/2$. The probability that D_A and D'_B click together is thus the same as the probability that D'_A and D_B click together, and they are equal to

$$P_{\theta_A, \theta_B}(1_A, -1_B) = P_{\theta_A, \theta_B}(-1_A, 1_B) = \frac{1}{2} \cos^2(\theta_B - \theta_A). \quad (201)$$

The probability of a "cross-click" (of any type) is therefore given by $\cos^2(\theta_B - \theta_A)$ and when the two polarization analyzers are aligned (i.e. when $\theta_A = \theta_B$), it equals unity. We may conclude that when Alice and Bob's analyzers are aligned in the same way (no matter at what angle with respect to the directions H and V) detectors D_A and D_B , or D'_A and D'_B never click simultaneously. One will always record clicks either in D_A and D'_B , or in D'_A and D_B .

Let us now consider the situation from the standpoint of a single photon. It is quite obvious that in the above scheme we have no preference to whether detector D_A , or D'_A clicks in Alice's receiver, or whether D_B , or D'_B clicks in Bob's receiver. All we know is that the clicks of the two photons will be correlated, but we do not have any knowledge on the polarizations that they are going to collapse to. Thus the probability that D_A clicks is simply

$$\begin{aligned} P_{\theta_A}(1_A) &= P_{\theta_A, \theta_B}(1_A, 1_B) + P_{\theta_A, \theta_B}(1_A, -1_B) \\ &= \frac{1}{2}[\sin^2(\theta_B - \theta_A) + \cos^2(\theta_B - \theta_A)] = \frac{1}{2}, \end{aligned} \quad (202)$$

and it is obviously equal to the complementary probability⁷⁸ $P_{\theta_A}(-1_A)$.

The Einstein-Podolsky-Rosen paradox

Assume that the two entangled photons are transmitted towards our two observers, Alice and Bob, that are well separated in space. The fact that the results of their measurements are correlated independently of the distance between them creates certain conceptual difficulties that were originally presented by Einstein Podolsky and Rosen. Assume that Alice decides to align her analyzer along an angle θ . Her measurement will determine whether the photon that reaches her is polarized along θ , or along $\theta' = \theta + \pi/2$ (depending on whether $A(\theta)$ is 1, or -1). Note that the result is unknown initially, because Alice's photon is unpolarized. The a priori probability of obtaining any result is $1/2$. The same is true for Bob. Yet, once Alice performs her measurement, and assuming that the detector D_A clicked, the wave-function on Bob's side collapses and it is absolutely certain that if Bob aligns his analyzer in the same way as Alice did, detector D'_B will click. In other words, it appears that Alice's measurement collapsed the entire wave-function into the state $|\theta, \theta'\rangle$. The problem is that Alice and Bob may be a very large distance

⁷⁸ In a later section I will show that marginal probability distributions (i.e. those concerning only one of the two photons) are obtained from a *reduced density matrix*, which is obtained by tracing the two-photon density matrix over the states of the photon that one is not considering.

away from each other. The time between their measurements can be arbitrarily small, and thus, it appears as if Alice's measurement caused the collapse of the wave-function at Bob's location instantaneously, apparently violating the principle that no signal can be transmitted faster than the speed of light. EPR used this scenario in their 1935 paper to question the completeness of quantum theory. A famous notion that was developed as a result of this paper was the notion of *hidden variables*. The hidden variable concept was meant to indicate that there is more to reality than what is apparent in quantum theory. In other words, since it is impossible for the measurement performed by Alice to cause the collapse of a wave function on Bob's side while respecting causality, the outcomes of Bob's and Alice's measurement could not have been completely unknown in advance. What EPR have suggested is the possibility that there is a variable ξ that is external to quantum theory and therefore "hidden" from it, that determines *in advance* what detectors will click for each alignment of the analyzers on the two sides. Searching for a theory that addresses the issue of locality and hidden variables was a very active area of research between 1935 and 1964, which is when Bell wrote his famous paper showing that the existence of hidden variables is not possible.

Bell's Inequalities

Let us now review very briefly what Bell's inequalities are. More precisely we will look at a later version of Bell's inequalities that was introduced in 1969, a few years after Bell's original paper, by Clauser, Horne, Shimony and Holt and which is known as the CHSH inequality. The key element in this theory is the notion that for any set of four numbers X_1, X_2, X_3 and X_4 whose values are allowed to be either 1, or -1 , the following equality holds

$$\begin{aligned} & X_1X_2 + X_1X_4 + X_3X_2 - X_3X_4 \\ &= X_1(X_2 + X_4) + X_3(X_2 - X_4) = \pm 2, \end{aligned} \tag{203}$$

as can be verified in a straightforward manner. The hidden variable suggestion made in the EPR article is that there is a hidden variable ξ given which the

value of $A(\theta_A)$ and $B(\theta_B)$ is determined for any value of θ_A and θ_B . We may express this idea formally by stating that the quantities $A(\theta_A, \xi)$ and $B(\theta_B, \xi)$ are perfectly deterministic functions⁷⁹ of θ_A, θ_B and ξ . With the above notions established, let us define X_1, X_2, X_3, X_4 as follows

$$\begin{aligned} X_1 &= A(\theta_A^{(1)}, \xi) \\ X_2 &= B(\theta_B^{(1)}, \xi) \\ X_3 &= A(\theta_A^{(2)}, \xi) \\ X_4 &= B(\theta_B^{(2)}, \xi), \end{aligned}$$

where $\theta_A^{(1)}, \theta_A^{(2)}$ and $\theta_B^{(1)}, \theta_B^{(2)}$ are two possible alignment angles for Alice and Bob's analyzers, respectively. Their numeric values will be specified in what follows. We may now substitute the above specified values of X_1, X_2, X_3, X_4 into relation (203) to obtain⁸⁰

$$\begin{aligned} &A(\theta_A^{(1)}, \xi)B(\theta_B^{(1)}, \xi) + A(\theta_A^{(1)}, \xi)B(\theta_B^{(2)}, \xi) \\ &+ A(\theta_A^{(2)}, \xi)B(\theta_B^{(1)}, \xi) - A(\theta_A^{(2)}, \xi)B(\theta_B^{(2)}, \xi) = \pm 2. \end{aligned} \quad (204)$$

Of course, the quantity in (204) cannot be checked directly in a measurement. That is because in every measurement Alice and Bob each need to decide

⁷⁹ Note that we wrote for Alice's experimental outcome the expression $A(\theta_A, \xi)$ and not $A(\theta_A, \theta_B, \xi)$. That is because the entire point of this exercise is that something that Bob chooses should not affect the result in Alice's experiment which can be very distant from it. The same is true for Bob's variable $B(\theta_B, \xi)$.

⁸⁰ It is important to understand that writing the expression in Eq. (204) is only meaningful under the assumption that the quantities $A(\theta_A^{(1)}, \xi)$, $A(\theta_A^{(2)}, \xi)$, $B(\theta_B^{(1)}, \xi)$, and $B(\theta_B^{(2)}, \xi)$ coexist within the studied physical reality, in spite of the fact that a way of determining them all simultaneously is unknown. As we have seen, quantum theory does not allow these quantities to coexist, but this is exactly what brought up the EPR paradox. Without assuming that these quantities coexist, as the hidden variable theory does, there is no meaning to saying that the sum of terms in Eq. (204) should be ± 2 .

whether they want to align their analyzer at an angle $\theta_{A,B}^{(1)}$, or $\theta_{A,B}^{(2)}$. Once they chose a certain alignment (e.g. $\theta_A^{(1)}$ for Alice and $\theta_B^{(2)}$ for Bob), they cannot tell what the result would have been had they chosen a different alignment⁸¹. Instead what they can do is perform a very large number of measurements for each combination of alignment angles $\theta_{A,B}^{(1,2)}$ and then average the measured products of the form $A(\theta_A^{(j)})B(\theta_B^{(k)})$ with j, k being either 1, or 2. If the hidden variables theory is correct, then the measured average of $A(\theta_A^{(j)})B(\theta_B^{(k)})$ is the same as the average of the expression $A(\theta_A^{(j)}, \xi)B(\theta_B^{(k)}, \xi)$ with respect to the distribution density of ξ . Although the distribution of ξ is unknown, it is clear that when expression (204) is averaged with respect to ξ (whatever its distribution), the result must be contained in the range between -2 and 2 . Denoting the ensemble average of $A(\theta_A)B(\theta_B)$ by $C(\theta_A, \theta_B)$ we obtain the CHSH version of Bell's inequality

$$-2 \leq C(\theta_A^{(1)}, \theta_B^{(1)}) + C(\theta_A^{(1)}, \theta_B^{(2)}) + C(\theta_A^{(2)}, \theta_B^{(1)}) - C(\theta_A^{(2)}, \theta_B^{(2)}) \leq 2. \quad (205)$$

Having concluded the above, we may now predict the results of the same experiment by using the tools devised within the framework of quantum physics. Using the relation

$$\begin{aligned} C(\theta_A, \theta_B) &= 1 \times P_{\theta_A, \theta_B}(1_A, 1_B) + 1 \times P_{\theta_A, \theta_B}(-1_A, -1_B) \\ &\quad + (-1) \times P_{\theta_A, \theta_B}(-1_A, 1_B) + (-1) \times P_{\theta_A, \theta_B}(1_A, -1_B), \end{aligned}$$

and taking advantage of Eqs. (200) and (201) we have

$$C(\theta_A, \theta_B) = -\cos[2(\theta_A - \theta_B)]. \quad (206)$$

It is not difficult to find a set of angles for which the inequality (205) is not satisfied. For example, we may consider the angles $\theta_A^{(1)} = 0$, $\theta_A^{(2)} = \pi/4$,

⁸¹ They would have been able to verify this relation directly only if they could fix ξ to be the same in a number of separate measurements where they change the analyzers alignments. But with ξ being "hidden" that would be difficult to do.

$\theta_B^{(1)} = \pi/8$ and $\theta_B^{(2)} = -\pi/8$, for which

$$C(\theta_A^{(1)}, \theta_B^{(1)}) + C(\theta_A^{(1)}, \theta_B^{(2)}) + C(\theta_A^{(2)}, \theta_B^{(1)}) - C(\theta_A^{(2)}, \theta_B^{(2)}) = -2\sqrt{2}$$

thereby violating the CHSH inequality. This result clearly indicates that quantum theory cannot coexist with the hidden variables theory, as the two produce contradicting results. One of the two theories must therefore be incorrect. This conundrum has been finally resolved when violation of the CHSH inequality was measured experimentally⁸². By today, the violation of Bell's inequalities has been proven experimentally beyond reasonable doubt, thereby contradicting the hidden variable theory. Nonetheless, in spite of the agreement among the majority of scientists, some scientists are still pointing out loopholes in Bell's arguments as well as in the experiments that have been performed.

Estimating the degree of entanglement

So far, the entanglement between a pair of particles has been defined as a situation in which the state that represents the two particles cannot be expressed as a tensor product. We have also seen that when two particles are entangled weird things happen. Yet, it is quite intuitive that not all entangled states are equally entangled. Consider for example, the state

$$|\psi\rangle = \cos(\theta)|HV\rangle + \sin(\theta)|VH\rangle. \quad (207)$$

When $\theta = -\pi/4$ this becomes the state $|\psi^-\rangle$ with which we are already acquainted. But what happens if θ is a very small number? For example, as-

⁸² The process of experimentally demonstrating the violation of Bell's inequalities was quite interesting. It started with Freedman and Clauser in 1972 and ended with Aspect in 1981. All experiments used polarization entangled photons. The first experiments addressed a variation of the CHSH inequality, rather than the form described here. Once the process of generating entangled photons through spontaneous parametric down-conversion was established and mastered in the mid eighties, Bell violation measurements became routine. In most cases today the CHSH inequality that we described is measured.

sume that $\theta = 10^{-9}$. Our state is still entangled in the sense that it cannot be expressed as a tensor product, but at the same time it is almost identical to $|HV\rangle$ which contains no entanglement at all. This similarity to (or practically the indistinguishability from) $|HV\rangle$ must be reflected in measurements. This means that the state $|\psi\rangle$ with $\theta = 10^{-9}$ is less entangled than $|\psi^-\rangle$. What is then a good measure for the amount of entanglement that is contained in a system? This question is more complex than it may initially seem, and it is difficult to present a measure that is always simple, intuitive, and valid (for example, for all pure and mixed state possibilities). But some measures are intuitive and useful in some of the cases, and we will describe them here. Two concepts that we will need in order to proceed with this discussion are the state's *entropy*, and the *reduced density matrix* of an entangled state.

The entropy and the 'uncertainty' contained in quantum system

We must be careful when defining the concept of uncertainty. As we have seen, uncertainty is intrinsic to quantum theory, as the results of most measurements can only be characterized probabilistically. Nonetheless, we can distinguish between a system that is in a pure-state, for which there are some particular measurements whose outcome is deterministic, and systems in a mixed state, for which there are no nontrivial⁸³ measurements whose outcome is known in advance. Consider for example the density matrix

$$\rho = \cos^2(\alpha)|H\rangle\langle H| + \sin^2(\alpha)|V\rangle\langle V|. \quad (208)$$

Let us consider the measurement of this state in a polarization analyzer of the kind illustrated in Fig. 5. When $\alpha = m\pi/2$ (with m being an integer) ρ represents a pure state and provided that the axes of the polarization analyzer are aligned with H and V, one can predict with perfect certainty which one of the detectors will click. When α is not an integer multiple of

⁸³ Of course one can always think of measurements whose outcome is deterministic. The unity operator is Hermitian and hence it corresponds to a measurable quantity, whose outcome is always 1.

$\pi/2$, ρ represents a mixed state and there is no setting of the polarization analyzer's axes for which the measurement outcome is certain. Yet, for some values of α (e.g. $\alpha = \pi/4$) there is more uncertainty, in the sense that the two possible outcomes of the measurement are equally likely for any alignment of the analyzer, whereas for other values of α (e.g. $\alpha = 10^{-10}$) the state is almost pure and there is an alignment of the analyzer for which the measurement outcome is almost certain. It seems that the question of how much uncertainty there is in a state, coincides with the question of how mixed the state is.

The entropy of a quantum system is a rigorous tool for quantifying the above ideas, and its definition is the following

$$S(\rho) = -\text{Trace}(\rho \log(\rho)). \quad (209)$$

Since we know that the expectation value (average) of a general operator \mathbf{A} with respect to a given quantum system is $\langle \mathbf{A} \rangle = \text{Trace}(\mathbf{A}\rho)$, we may also write the entropy as $S(\rho) = -\langle \log(\rho) \rangle$. The fact that we have a logarithm of an operator should not scare us at this stage, after having seen functions of operators before. The entropy is evaluated by using a representation basis that consists of the eigenvectors of ρ , where

$$\rho = \sum_i p_i |u_i\rangle\langle u_i|, \quad (210)$$

and where the (nonnegative) eigenvalues p_i have the meaning of probabilities. In this case expression (209) reduces to

$$S(\rho) = -\sum_i p_i \log(p_i), \quad (211)$$

which is the famous expression for entropy in statistical mechanics and in classical information theory. If the logarithm inside the sum uses the basis of 2, the units of entropy are called "bits" (otherwise, for natural logarithms the units of entropy are called "nats"). When a system is in a well prescribed state $|\psi\rangle$ so that its density operator is $\rho = |\psi\rangle\langle\psi|$, there is one eigenvalue

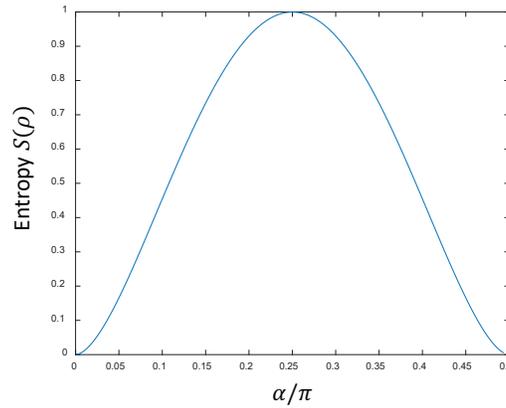


Fig. 7. The entropy of the system described by the density matrix of Eq. (208).

that is equal to 1 whereas the other eigenvalues are equal to 0. In this case substitution in Eq. (211) yields $S(\rho) = 0$. On the other hand, if a system is equally likely to be in any one of M orthogonal states (assuming an M dimensional space), the entropy assumes its highest possible value of $S(\rho) = \log(M)$. Returning to the example of Eq. (208), the entropy is given by

$$S(\rho) = -\cos^2(\alpha) \log_2(\cos^2(\alpha)) - \sin^2(\alpha) \log_2(\sin^2(\alpha)), \quad (212)$$

which is plotted in Fig. 7 for α between 0 and $\pi/2$. The entropy is maximal for maximally mixed states and it is zero when the state of the system is pure. We are now ready to return to the case of entangled systems, where entropy will be used to determine the degree of entanglement.

The reduced density matrix

Consider the case of two particles A and B that are described by a joint state-vector $|\psi\rangle$. In order to deal with a concrete example, let us refer to the spatial wave function $\psi(x_a, x_b) = (\langle x_a| \otimes \langle x_b|)|\psi\rangle$, where I assumed for simplicity that each particle lives in a single dimension. The function $P_{A,B}(x_a, x_b) = |\psi(x_a, x_b)|^2$ is the joint probability density function of the two particles. If we are interested only in particle A and not in particle B ,

we can easily find the marginal probability density function

$$P_A(x_a) = \int P_{A,B}(x_a, x_b) dx_b,$$

but what would be the quantum state-vector that would describe the state of particle A while ignoring particle B ? If the system is in a product state $|\psi\rangle = |u\rangle_A \otimes |v\rangle_B$ then it is quite obvious that $\psi_A(x_a) = \langle x_a | u \rangle_a$ and $P_A(x_a) = |\psi_A(x_a)|^2$. In this case, it is also obvious that $\psi(x_a, x_b) = \psi_A(x_a)\psi_B(x_b)$ and the joint probability density is a product of the marginal probability densities. Yet in the more general case, the process of attaching a wave-function to an individual particle does not seem obvious. As we shall now see, it is not only not obvious, but also impossible in cases where the two particles are entangled so that their joint state is not a tensor product of the individual states of the particles.

Recall that all physically observable quantities are represented by hermitian operators. Denote by \mathbf{G}_A a generic operator that acts only on the subspace of particle A . The mean outcome in the measurement of this operator is $\text{Trace}(\mathbf{G}_A \otimes \mathbf{1}_B \rho)$, where $\mathbf{G}_A \otimes \mathbf{1}_B$ is the extension of \mathbf{G}_A to the external product space. I will now denote by $|a_j\rangle_A$ and $|b_k\rangle_B$ the members of orthonormal bases that span the individual subspaces of particles A and B . We may then write

$$\begin{aligned} \text{Trace}(\mathbf{G}_A \otimes \mathbf{1}_B \rho) &= \sum_j \sum_k \langle a_j | \langle b_k | \mathbf{G}_A \mathbf{1}_B \rho | a_j \rangle | b_k \rangle \\ &= \sum_j \langle a_j | \mathbf{G}_A \rho_A | a_j \rangle = \text{Trace}_A(\mathbf{G}_A \rho_A). \end{aligned} \quad (213)$$

where Trace_A denotes a trace with respect to the basis vectors of subspace A only. The operator ρ_A , which is defined as

$$\rho_A = \sum_k \langle b_k | \rho | b_k \rangle = \text{Trace}_B(\rho) \quad (214)$$

is clearly the density matrix that describes particle A while ignoring particle B and it is called the reduced density matrix of particle A . Similarly, the

reduced density matrix of particle B is $\rho_B = \text{Trace}_A(\rho)$. For observations that are performed separately on either particle A , or on particle B , the corresponding reduced density matrix provides adequate information on the system. The state of the individual particle is completely described by it.

First let us now consider the example in which the two particle system is in a product state $|\psi\rangle = |u\rangle_A|v\rangle_B$ as we considered earlier. The density matrix in this case is $\rho = |u\rangle_A\langle u| \otimes |v\rangle_B\langle v|$ and the partial trace with respect to B yields $\rho_A = \text{Trace}_B(\rho) = (|u\rangle\langle u|)_A \otimes \sum_k \langle b_k|v\rangle_B \langle v|b_k\rangle = (|u\rangle\langle u|)_A$. Thus we find that particle A is in a pure state $|u\rangle_A$, as expected.

As a second example let us consider the state $|\psi^-\rangle = \frac{1}{\sqrt{2}}[|H_A V_B\rangle - |V_A H_B\rangle]$ of the type that was introduced in the previous section. In this case

$$\rho = |\psi^-\rangle\langle\psi^-| = \frac{1}{2} \left[|H_A V_B\rangle\langle H_A V_B| + |V_A H_B\rangle\langle V_A H_B| - |H_A V_B\rangle\langle V_A H_B| - |V_A H_B\rangle\langle H_A V_B| \right]$$

and $\rho_A = {}_B\langle H|\rho|H\rangle_B + {}_B\langle V|\rho|V\rangle_B = \frac{1}{2}[|H\rangle\langle H|_A + |V\rangle\langle V|_A] = \frac{1}{2}\mathbf{1}_A$. Thus we found, consistently with our earlier discussion, that from the standpoint of local measurements, particle A is in a maximally mixed quantum state. An analogous procedure would show the same result for particle B , i.e. that $\rho_B = \frac{1}{2}\mathbf{1}_B$. In fact, as we show in the subsequent paragraph, in the case of entanglement between two particles, the reduced density matrices always have the same spectrum of nonzero eigenvalues. This implies that the entropies of the two reduced density matrices of any two-particle entangled system which is in pure state, will have the same value $S(\rho_A) = S(\rho_B)$. Notice also that in the case of a product state (no entanglement) $S(\rho_{A,B}) = 0$, whereas in a maximally entangled state $S(\rho_{A,B}) = \log(M)$, where M is the smaller of the dimensions of the two individual spaces. The entropy of the reduced matrices is therefore a very useful parameter for quantifying the *degree of entanglement*

and it is indeed the most widely used metric for the entanglement of pure states.

In order to demonstrate the fact that the matrices ρ_A and ρ_B share the same eigenvalues, we first introduce the **Schmidt decomposition** which states that any pure two-particle state $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{i=1}^M g_i |u_i\rangle_A \otimes |v_i\rangle_B, \quad (215)$$

where $|u_i\rangle$ and $|v_i\rangle$ are two sets of orthonormal vectors in the spaces of the two individual particles labeled by the subscripts A and B . The summation runs from 1 to M , where M is the dimension of the smaller of the two spaces. In order to prove this we recall that *any* two-particle state-vector can be written as $|\psi\rangle = \sum_{i,j} c_{ij} |a_i\rangle \otimes |b_j\rangle$ with any bases $|a_i\rangle_{i=1}^M$ and $|b_j\rangle_{j=1}^N$, where we assume without loss of generality that $M \leq N$. In matrix form this can be written as follows

$$(|a_1\rangle |a_2\rangle \dots |a_M\rangle) \mathbf{C}_{M \times N} \begin{pmatrix} |b_1\rangle \\ |b_2\rangle \\ \vdots \\ |b_N\rangle \end{pmatrix}, \quad (216)$$

where $\mathbf{C}_{M \times N}$ is an $M \times N$ matrix whose elements are $C_{i,j}$. According to the singular value decomposition theorem $\mathbf{C}_{M \times N} = \mathbf{U}_{M \times M}^\dagger \mathbf{\Lambda}_{M \times N} \mathbf{V}_{N \times N}$, where $\mathbf{U}_{M \times M}$ and $\mathbf{V}_{N \times N}^\dagger$ are unitary matrices, whereas $\mathbf{\Lambda}_{M \times N}$ is diagonal with at most M nonzero elements $\Lambda_{1,1}$ to $\Lambda_{M,M}$ on the diagonal (recall that $M \leq N$). Thus, we may define

$$\begin{pmatrix} |u_1\rangle \\ |u_2\rangle \\ \vdots \\ |u_M\rangle \end{pmatrix} = \mathbf{U}^* \begin{pmatrix} |a_1\rangle \\ |a_2\rangle \\ \vdots \\ |a_M\rangle \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} |v_1\rangle \\ |v_2\rangle \\ \vdots \\ |v_M\rangle \end{pmatrix} = \mathbf{V} \begin{pmatrix} |b_1\rangle \\ |b_2\rangle \\ \vdots \\ |b_N\rangle \end{pmatrix}, \quad (217)$$

which implies that Eq. (216) can be rewritten as,

$$(|u_1\rangle |u_2\rangle \dots |u_M\rangle) \mathbf{\Lambda}_{M \times N} \begin{pmatrix} |v_1\rangle \\ |v_2\rangle \\ \vdots \\ |v_N\rangle \end{pmatrix} = \sum_{i=1}^M \Lambda_{i,i} |u_i\rangle_A \otimes |v_i\rangle_B, \quad (218)$$

and proves the Schmidt decomposition (215) with $g_i = \Lambda_{i,i}$. From Eq. (215) it is trivial to show that $\rho_A = \sum_{i=1}^M |g_i|^2 |u_i\rangle\langle u_i|$ and $\rho_B = \sum_{i=1}^M |g_i|^2 |v_i\rangle\langle v_i|$, showing that the spectra of eigenvalues are indeed identical.

So far we have discussed the ways in which the degree of entanglement can be quantified in pure state entangled systems. The situation in the case of mixed states entanglement is far less trivial, with the difficulty being that mixed states can be formed in a variety of ways. Consider for example a mixed state in which there is an equal 0.5 probability of the system to be in either of the states $|\psi^\pm\rangle = \frac{1}{\sqrt{2}}[|HV\rangle \pm |VH\rangle]$, so that

$$\rho = \frac{1}{2} |\psi^+\rangle\langle\psi^+| + \frac{1}{2} |\psi^-\rangle\langle\psi^-|.$$

The form of this density matrix appears to imply that the system is formed from maximally entangled particles and hence the degree of entanglement should be maximal as well. Indeed if we calculate the marginal density matrices, it can easily be shown that $\rho_A = \rho_B = \frac{1}{2}\mathbf{1}$, and hence the entropy $S(\rho_A) = S(\rho_B) = 1$ bit, which is the maximal value. The assertion that our mixed state is maximally entangled can be trivially refuted if one notes that the same ensemble can be formed from an equal mixture of $|HV\rangle$ and $|VH\rangle$, since (as can be easily shown) the same density matrix can be expressed as

$$\rho = \frac{1}{2} [|HV\rangle\langle HV| + |VH\rangle\langle VH|].$$

In this case, the building blocks are not entangled at all and hence the mixed state should have no entanglement. A common metric for quantifying entanglement in mixed systems is called *entanglement of formation* and it is defined as

$$E(\rho) = \min \sum_i p_i E(|\psi_i\rangle\langle\psi_i|),$$

where $E(|\psi_i\rangle\langle\psi_i|)$ is the degree of entanglement of a pure state $|\psi_i\rangle$ and where minimization is performed with respect to the various possibilities of decomposing ρ into $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Interestingly, in the case of two-level spaces (such as the space of photon polarizations), there is a rigorous way of extracting the degree of entanglement from the joint density matrix. This, however, goes beyond the scope of our course.

The complex envelope of the electric field in the quantum description

As we have seen earlier, the complex envelope of the electric field⁸⁴ of a single mode maps to the quantum annihilation operator \mathbf{a} . Thus the quadratures of the complex field map to

$$\mathbf{X} = \frac{1}{2}(\mathbf{a} + \mathbf{a}^\dagger) \quad (219)$$

$$\mathbf{Y} = \frac{1}{2i}(\mathbf{a} - \mathbf{a}^\dagger) \quad (220)$$

which are the in-phase and in-quadrature components of the electric field⁸⁵, respectively. The quadratures are measurable physical quantities and consistently with that, the operators \mathbf{X} , \mathbf{Y} that represent them are Hermitian. Clearly, they are equivalent to the position and momentum operators that are associated with the mode, except that they are smaller than them by a factor

⁸⁴ For the definition of complex envelope consult Appendix II.

⁸⁵ An obvious generalization is the following: Define $\mathbf{a}_\theta = \exp(i\theta)\mathbf{a}$ and then

$$\mathbf{X}(\theta) = \frac{1}{2}(\mathbf{a}e^{i\theta} + \mathbf{a}^\dagger e^{-i\theta}) \quad (221)$$

$$\mathbf{Y}(\theta) = \frac{1}{2i}(\mathbf{a}e^{i\theta} - \mathbf{a}^\dagger e^{-i\theta}) \quad (222)$$

These are the two field quadratures with respect to a rotated reference frame.

of $\sqrt{2}$. For this reason they satisfy the commutation relation $[\mathbf{X}, \mathbf{Y}] = i/2$, as well as the uncertainty relation $\Delta X \Delta Y \geq 1/4$. Notice that this uncertainty pertains to the state itself, not to the uncertainty that is associated with the measurement of the quadratures, which as we shall see in what follows, is larger. Consistently with what we have studied previously, in the case of coherent states, the uncertainty in the quadratures is identical and minimal so that $\Delta X = \Delta Y = 1/2$. There is a more general class of states that also satisfy the minimum uncertainty condition, but with $\Delta X \neq \Delta Y$. These states are called quadrature squeezed states, and you have encountered them in a homework exercise.

I now wish to discuss a particularly relevant construction that is used for measuring the electric field. It is called the *balanced homodyne receiver*. This scheme is used both in the classical and in the quantum mechanical contexts, and it is convenient to start from the classical description of the way in which it operates. Consider the set-up illustrated in Fig. 8a, where an incident field whose complex envelope is denoted by E , is mixed with a local oscillator field with a complex envelope equal to β , using a standard symmetric beam-splitter. The signals emerging from the beam-splitter's output arms are photo-detected, producing electric currents that are proportional to the field intensities (i.e. to the square absolute values of the complex fields). Without any loss of generality, we may define the phase of β to be 0, such that it serves as a reference for all other phases in the experiment. In this case β is real-valued, and one can easily show that the subtraction between the two photocurrents produces a signal proportional to $2\beta\text{Im}(E)$, where $\text{Im}(\cdot)$ denotes the imaginary part. The imaginary part of E is the part that is said to be *in quadrature* with the local oscillator signal β . The real part of the electric field $\text{Re}(E)$ is often referred to as the *in-phase* part of the signal with respect to β and in order to measure it, the local oscillator is delayed by a quarter cycle so that $i\beta$ instead of β is used⁸⁶.

⁸⁶ It is easily shown that in this case the measured quantity is $2\beta\text{Re}(E)$.

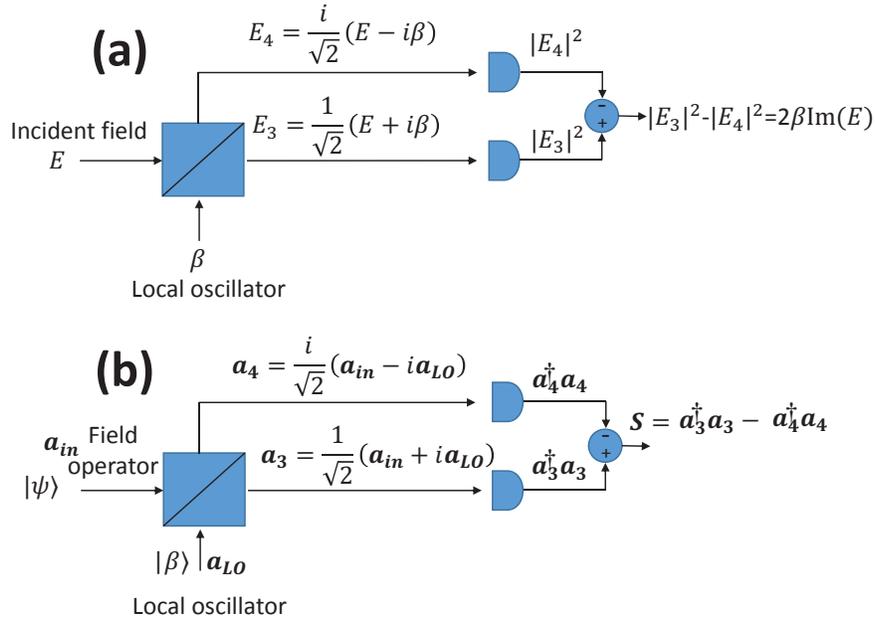


Fig. 8. Schematic of the Homodyne receiver designed to extract the in-quadrature component of the field. The value of β is set to be real. The in-phase component is obtained when β is replaced by $i\beta$. (a) Classical description (b) Quantum description.

We now shift to the quantum description of the balanced homodyne receiver, making use of the illustration in Fig. 8b. Now, the incident field is represented by the annihilation operator \mathbf{a}_{in} . The input signal is a quantum state $|\psi\rangle$ and the local oscillator is a coherent state $|\beta\rangle$. The electric field in the local-oscillator arm is \mathbf{a}_{LO} . Photo-detection in the two output ports (ports 3 and 4) is represented by the operators $\mathbf{a}_j^\dagger \mathbf{a}_j$, where j is either 3 or 4, and the output of the measurement apparatus is represented by the operator $\mathbf{S} = \mathbf{a}_3^\dagger \mathbf{a}_3 - \mathbf{a}_4^\dagger \mathbf{a}_4$. As in the classical case, measurement of the in-phase component requires that the local oscillator is delayed by a quarter cycle so that $|\beta\rangle$ is replaced by $|i\beta\rangle$. Rewriting the operator \mathbf{S} in terms of \mathbf{a}_{in} and \mathbf{a}_{LO} gives

$$\mathbf{S} = i \left(\mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{LO}} - \mathbf{a}_{\text{in}} \mathbf{a}_{\text{LO}}^\dagger \right), \quad (223)$$

and all the possible outcomes of the measurement are the eigenvalues of this Hermitian operator. Notice that \mathbf{S} acts in a space which is a tensor product of the spaces of the incident signal and of the local oscillator and it is not immediately obvious (at least to me) what its eigenvectors are. But, inspired by the classical measurement we can try to test how similar $\mathbf{S}/2\beta$ is to the in quadrature field component \mathbf{Y} given by Eq. (220). To that end we evaluate the second moment of the difference between the two operators, which turns out to be equal to⁸⁷

$$\langle \psi, \beta | \left(\frac{\mathbf{S}}{2\beta} - \mathbf{Y} \right)^2 | \psi, \beta \rangle = \frac{\langle \psi | \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} | \psi \rangle}{4\beta^2}, \quad (227)$$

⁸⁷ The derivation is a bit cumbersome, but straightforward.

$$\begin{aligned} \langle \psi, \beta | \left(\frac{\mathbf{S}}{2\beta} - \mathbf{Y} \right)^2 | \psi, \beta \rangle &= \\ \langle \psi, \beta | \frac{\mathbf{S}^2}{4\beta^2} + \mathbf{Y}^2 - \frac{1}{2\beta} \mathbf{YS} - \frac{1}{2\beta} \mathbf{SY} | \psi, \beta \rangle. \end{aligned} \quad (224)$$

First we write down the average of \mathbf{S}^2

$$\begin{aligned} \langle \psi, \beta | \mathbf{S}^2 | \psi, \beta \rangle &= \langle \psi, \beta | \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} \mathbf{a}_{\text{LO}} \mathbf{a}_{\text{LO}}^\dagger + \mathbf{a}_{\text{in}} \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{LO}}^\dagger \mathbf{a}_{\text{LO}} \\ &\quad - (\mathbf{a}_{\text{in}}^\dagger)^2 \mathbf{a}_{\text{LO}}^2 - \mathbf{a}_{\text{in}}^2 (\mathbf{a}_{\text{LO}}^\dagger)^2 | \psi, \beta \rangle \\ &= -\beta^2 \langle \psi | (\mathbf{a}_{\text{in}} - \mathbf{a}_{\text{in}}^\dagger)^2 | \psi \rangle + \langle \psi | \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} | \psi \rangle \\ &= 4\beta^2 \langle \psi | \mathbf{Y}^2 | \psi \rangle + \langle \psi | \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} | \psi \rangle, \end{aligned} \quad (225)$$

Then (leaving the detailed calculation for a homework exercise) we argue that

$$\langle \psi, \beta | \mathbf{YS} | \psi, \beta \rangle = \langle \psi, \beta | \mathbf{SY} | \psi, \beta \rangle = 2\beta \langle \psi | \mathbf{Y}^2 | \psi \rangle, \quad (226)$$

where the fact that β is real-valued was used once again. Substitution of (225) and (226) into (224) yields the result of Eq. 227.

which implies that although $\mathbf{S}/2\beta$ is not identical to \mathbf{Y} in general, it becomes indistinguishable from it in the limit where $|\beta|^2$, the intensity of the local-oscillator, is much larger than $\langle\psi|\mathbf{a}_{\text{in}}^\dagger\mathbf{a}_{\text{in}}|\psi\rangle$, the average number of photons present in the incident state. The idea behind this is that when a Hermitian operator \mathbf{O} satisfies $\langle\psi|\mathbf{O}^2|\psi\rangle = 0$, for all $|\psi\rangle$ then it is necessarily the zero-operator.⁸⁸ Since in the relevant limit the measured operator becomes identical to \mathbf{Y} , the homodyne setup constitutes an *exact* measurement of the in-quadrature part of the field, i.e. one that does not introduce any uncertainty, or noise. Nonetheless, owing to the inherent uncertainty that is present in the state $|\psi\rangle$ when it is not an eigenstate of the measured operator \mathbf{Y} , the signal-to-noise ratio (SNR) that characterizes this measurement is not infinite, and it is given by

$$\text{SNR} = \frac{\langle\psi|\mathbf{Y}|\psi\rangle^2}{\langle\psi|\Delta\mathbf{Y}^2|\psi\rangle}. \quad (228)$$

In the important special case where $|\psi\rangle$ is a coherent states, $\langle\psi|\Delta\mathbf{Y}^2|\psi\rangle = 1/4$ and the SNR is equal to $4\langle\psi|Y|\psi\rangle^2$.

Notice that so far we have only been concerned with the measurement of the in-quadrature field operator \mathbf{Y} . Measurement of \mathbf{X} is performed similarly by replacing $|\beta\rangle$ with $|i\beta\rangle$, as we have discussed before. But what happens if one wants to measure the entire field, i.e. the in-phase and in-quadrature components simultaneously? The idea then is to split the incoming signal into two identical replicas. One is used to measure the in-phase component and the other is used to measure the in-quadrature component in two separate simultaneous measurements, as illustrated in Fig. 9. Classically, the splitting and the simultaneous measurement of the two quadratures comes at absolutely no cost, provided of course that there is no noise that is produced by the photodetectors themselves, or by the subsequent electronic circuitry. Yet, quantum mechanically, the situation is quite different. The reason for this dif-

⁸⁸ One can think of it in terms of the outcomes of its measurement (the eigenvalues), which are characterized by 0 variance around 0, so they must all be equal to 0.

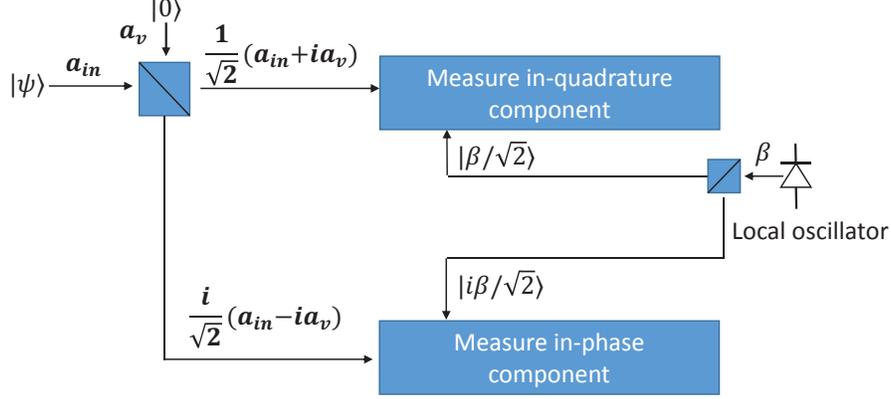


Fig. 9. The scheme for measuring both quadratures simultaneously.

ference is that instead of measuring $\mathbf{Y} = \frac{1}{2i}(\mathbf{a}_{in} - \mathbf{a}_{in}^\dagger)$ we are now measuring $\mathbf{Y}' = \frac{1}{\sqrt{2}}(\mathbf{Y} + \mathbf{Y}_v)$ where $\mathbf{Y}_v = \frac{1}{2i}(\mathbf{a}_v - \mathbf{a}_v^\dagger)$ is the in-quadrature component of the field entering through the empty port of the first beam-splitter (see Fig. 9).⁸⁹ Hence, the SNR of this measurement is

$$\text{SNR} = \frac{\langle \psi_{in}, 0_v | \mathbf{Y}' | \psi_{in}, 0_v \rangle^2}{\langle \psi_{in}, 0_v | \Delta \mathbf{Y}'^2 | \psi_{in}, 0_v \rangle}, \quad (229)$$

where $|\psi_{in}, 0_v\rangle = |\psi\rangle_{in} \otimes |0\rangle_v$, so that the averaging is performed both with respect to the state $|\psi\rangle$ which is present at the input port and with respect to the vacuum state $|0\rangle$ which is present in the empty port of the beam-splitter. It is easy to see that

$$\langle \psi_{in}, 0_v | \mathbf{Y}' | \psi_{in}, 0_v \rangle = \langle \psi | \mathbf{Y} | \psi \rangle + \langle 0 | \mathbf{Y}_v | 0 \rangle \quad (230)$$

$$\langle \psi_{in}, 0_v | \Delta \mathbf{Y}'^2 | \psi_{in}, 0_v \rangle = \langle \psi | \Delta \mathbf{Y}^2 | \psi \rangle + \langle 0 | \Delta \mathbf{Y}_v^2 | 0 \rangle, \quad (231)$$

and since for the vacuum state $\langle 0 | \mathbf{Y}_v | 0 \rangle = 0$ and $\langle 0 | \Delta \mathbf{Y}_v^2 | 0 \rangle = \frac{1}{4}$, the SNR of this measurement reduces to

⁸⁹ Of course, it is assumed that the intensity of the local oscillator is large to the extent that the difference between the measured operator and \mathbf{Y}' can be neglected.

$$\text{SNR} = \frac{\langle \psi | \mathbf{Y} | \psi \rangle^2}{\langle \psi | \Delta \mathbf{Y}^2 | \psi \rangle + \frac{1}{4}}. \quad (232)$$

When the measured state $|\psi\rangle$ is coherent, the uncertainty in its quadrature is $\langle \psi | \Delta \mathbf{Y}^2 | \psi \rangle = 1/4$, and the SNR is $2\langle \psi | \mathbf{Y} | \psi \rangle^2$ — exactly half of the SNR that characterized the measurement of \mathbf{Y} quadrature of a coherent state without splitting the signal. Namely, the greedy desire to simultaneously measure both quadratures of the electric field results in a penalty of a factor of 2 (3 dB) in the SNR of each quadrature. Remind me to tell you in class why this is also the reason for the fact that the smallest possible noise-figure that characterizes a large-gain optical amplifier is also equal to 3 dB.

Other topics to cover if time permits

The following should be covered if time permits. The order is immaterial.

- Dense coding
- Teleportation
- Quantum key distribution
- Time energy uncertainty relations
- No cloning theorem
- Amplification noise
- Phase-space representation of fields and the Wigner distribution
- Basic measurement theory, POVM

Appendix I - The Tensor Product

Assume that we have two independent and non interacting systems⁹⁰. Each system is described by a state-vector that belongs to a certain space. I will denote the state-vectors of the first system by $|\phi\rangle_1$ and the states of the second system by $|\psi\rangle_2$. Similarly, I will denote the spaces to which the two states belong as ε_1 and ε_2 , respectively. There is no requirement that the two systems, and hence the two spaces ε_1 and ε_2 be identical to each other. Suppose now that we want to expand our space in order to jointly relate to those systems. In other words, we want to prescribe a single state-vector $|\chi\rangle$ that simultaneously defines the states of the two individual systems. As we shall now see, the *tensor product*, also known as the *outer product* is what we need in order to perform this task. In the case of two completely independent systems, the tensor product offers nothing more than mathematical convenience. All questions regarding such systems can be answered easily by combining our knowledge about the description of the individual systems with simple probabilistic arguments. Yet, if we wish to relate to cases in which the two systems are allowed to interact, the use of the unified description provided by the tensor product formalism becomes inevitable. For convenience, most of the text that follows refers to two spaces of finite dimension. The expansion to more spaces, or to spaces of infinite dimensions is usually not difficult. The following discussion follows very closely the text in section F in chapter 2 of Ref. [1].

Definition

The vector space ε is called the tensor product of ε_1 and ε_2 ,

$$\varepsilon = \varepsilon_1 \otimes \varepsilon_2$$

if with each pair of vectors $|\phi\rangle_1$ and $|\psi\rangle_2$ belonging to ε_1 and ε_2 , respectively, there is an associated vector

⁹⁰ For example, in order to imagine something concrete, one can think of two separate potential wells with a particle in each one of them.

$$|\chi\rangle = |\phi\rangle_1 \otimes |\psi\rangle_2,$$

which is called the tensor product of $|\phi\rangle_1$ and $|\psi\rangle_2$. This vector needs to satisfy the following conditions:

(1) It is linear with respect to multiplication by complex numbers⁹¹

$$\begin{aligned} [\lambda|\phi\rangle_1] \otimes |\psi\rangle_2 &= \lambda[|\phi\rangle_1 \otimes |\psi\rangle_2] \\ |\phi\rangle_1 \otimes [\mu|\psi\rangle_2] &= \mu[|\phi\rangle_1 \otimes |\psi\rangle_2]. \end{aligned} \quad (233)$$

(2) It is distributive with respect to vector addition

$$\begin{aligned} |\phi\rangle_1 \otimes [|\psi_A\rangle_2 + |\psi_B\rangle_2] &= |\phi\rangle_1 \otimes |\psi_A\rangle_2 + |\phi\rangle_1 \otimes |\psi_B\rangle_2 \\ [|\phi_A\rangle_1 + |\phi_B\rangle_1] \otimes |\psi\rangle_2 &= |\phi_A\rangle_1 \otimes |\psi\rangle_2 + |\phi_B\rangle_1 \otimes |\psi\rangle_2. \end{aligned} \quad (234)$$

The scalar product in the tensor-product space

The scalar product between two tensor-product vectors $|\chi_A\rangle = |\phi_A\rangle_1 \otimes |\psi_A\rangle_2$ and $|\chi_B\rangle = |\phi_B\rangle_1 \otimes |\psi_B\rangle_2$ is defined as

$$\langle \chi_B | \chi_A \rangle = \langle \phi_B | \phi_A \rangle_1 \langle \psi_B | \psi_A \rangle_2. \quad (235)$$

This definition is quite natural and it can be easily seen to make sense. Recall that the absolute square value of the inner product between two vectors is the probability of measuring a system that is initially prepared in $|\chi_A\rangle$ in a state $|\chi_B\rangle$. This probability has to be equal to the product of probabilities that the first system is found in state $|\phi_B\rangle_1$ and the second system is found in a state $|\psi_B\rangle_2$, given that the two systems were prepared in states $|\phi_A\rangle_1$ and $|\psi_A\rangle_2$, respectively.

⁹¹ This property may appear confusing at first because it appears to imply that the relation between the individual vectors and their tensor product is not unique. But this should not be a real concern for us because the multiplication of a quantum state by a number has no physical significance. In other words, the state $|\phi\rangle$ and $\lambda|\phi\rangle$ represent exactly the same thing.

The tensor product space

The space $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$, is the space spanned by all vectors of the form $|\phi\rangle_1 \otimes |\psi\rangle_2$. A basis for ε can be constructed from all the possible tensor products of $|u_j\rangle_1 \otimes |v_k\rangle_2$ where $\{|u_j\rangle_1\}_{j=1}^{N_1}$ and $\{|v_k\rangle_2\}_{k=1}^{N_2}$ are orthonormal bases for ε_1 and ε_2 , respectively. The dimension of the product space is $N = N_1 N_2$. Clearly, all vectors of the form $|\chi\rangle = |\phi\rangle_1 \otimes |\psi\rangle_2$ belong to the space ε , but not all vectors in ε can be written as a tensor product of a state-vector in ε_1 and a state-vector in ε_2 . An obvious example is the state

$$\gamma \left(|\phi_A\rangle_1 \otimes |\psi_A\rangle_2 + |\phi_B\rangle_1 \otimes |\psi_B\rangle_2 \right),$$

where γ is a normalization factor. As we will see in later stages, states of this form are created when the two systems are allowed to interact with each other. They have a special and important role in the field of quantum physics.

Tensor products of operators

If an operator \mathbf{A}_1 that acts on the system 1 is applied, it changes the state-vector in system 1, but doesn't do anything to system 2. The same is true when an operator \mathbf{B}_2 that acts on system number 2 is applied. In other words

$$(\mathbf{A}_1 \otimes \mathbf{B}_2)(|\phi\rangle_1 \otimes |\psi\rangle_2) = (\mathbf{A}_1|\phi\rangle_1) \otimes (\mathbf{B}_2|\psi\rangle_2). \quad (236)$$

In the above we defined $\mathbf{A}_1 \otimes \mathbf{B}_2$ as an operator in ε that consists of a tensor product of individual operators, each acting on a separate sub-space. As a special case, the expansion of the operator \mathbf{A}_1 that acts only on the first sub-space is $\mathbf{A}_1 \otimes \mathbf{1}_B$. The same can be applied to an operator acting only on ε_2 . Just as with vectors, there exist operators in ε that cannot be written as a tensor product of operators in the two separate subsystems. Such operators introduce interactions between the systems. Notice that the extension of operators \mathbf{A}_1 and \mathbf{B}_2 that act on separate spaces commute with each other. Namely, $[(\mathbf{A}_1 \otimes \mathbf{1}_2), (\mathbf{1}_1 \otimes \mathbf{B}_2)] = 0$.

Tensor products of operators have various additional properties, two of which are listed below. Proving them is a homework exercise.

$$(i) (\mathbf{A}_1 \otimes \mathbf{B}_2)(\mathbf{C}_1 \otimes \mathbf{D}_2) = (\mathbf{A}_1 \mathbf{C}_1) \otimes (\mathbf{B}_2 \mathbf{D}_2).$$

$$(ii) \text{Trace}(\mathbf{A}_1 \otimes \mathbf{B}_2) = \text{Trace}(\mathbf{A}_1) \text{Trace}(\mathbf{B}_2).$$

Notation

As in most texts on quantum physics we will often use an abbreviated notation. Thus instead of writing $|\phi\rangle_1 \otimes |\psi\rangle_2$ I will often write $|\phi\rangle_1 |\psi\rangle_2$, or even $|\phi, \psi\rangle$, where the first position inside the combined ket corresponds to ε_1 , and the second position corresponds to ε_2 . Similarly, the $\mathbf{A}_1 \otimes \mathbf{B}_2$ will be replaced by $\mathbf{A}_1 \mathbf{B}_2$, with the meaning being clear from the context. The operator $\mathbf{A}_1 \otimes \mathbf{1}_2$ will be simply written as \mathbf{A}_1 leaving the expansion to the tensor product space implicit. This means for example that

$$(\mathbf{A}_1 + \mathbf{B}_2)|\phi\rangle_1 |\psi\rangle_2 = \mathbf{A}_1 |\phi\rangle_1 |\psi\rangle_2 + |\phi\rangle_1 \mathbf{B}_2 |\psi\rangle_2.$$

Eigenvalues and eigenvectors

Some relations concerning eigenvalues and eigenvectors are useful and important for what we will need. Assume a Hermitian operator \mathbf{A}_1 acting in the space ε_1 whose corresponding eigenvectors and eigenvalues in this space are $\{|a_n\rangle\}_{n=1}^{N_1}$ and $\{a_n\}_{n=1}^{N_1}$ respectively. All vectors in the product space ε that are of the form

$$|\chi\rangle = |a_n\rangle_1 |\psi\rangle_2$$

are eigenvectors of the extension of \mathbf{A}_1 into ε with the same eigenvalue a_n . Thus, a simple eigenvalue that is not degenerate in ε_1 becomes N_2 -fold degenerate in ε . Similarly, if $\{|b_k\rangle\}_{k=1}^{N_2}$ and $\{b_k\}_{k=1}^{N_2}$ are the eigenvectors and eigenvalues of an operator \mathbf{B}_2 acting on ε_2 , then vectors of the form $|\chi\rangle = |a_n\rangle_1 |b_k\rangle_2$ are eigenvectors of $\mathbf{A}_1 \mathbf{B}_2$ with the product eigenvalues $a_n b_k$.

An interesting situation is that of a sum of operators of the form $\mathbf{C} = \mathbf{A}_1 + \mathbf{B}_2$. The eigenvectors of \mathbf{C} are all vectors of the form $|a_n\rangle_1 |b_k\rangle_2$ and the

corresponding eigenvalues are $a_n + b_k$. Notice that since $[\mathbf{A}_1, \mathbf{B}_2] = [\mathbf{A}_1, \mathbf{C}] = [\mathbf{C}, \mathbf{B}_2] = 0$, one can always find a basis of eigenvectors of \mathbf{C} that are also eigenvectors of \mathbf{A}_1 and of \mathbf{B}_2 .

Representation in matrix form

In the finite dimensional case, one can use column vectors and matrices in order to represent states and operators. Any basis of the product state ε can be used for this representation. A particularly convenient choice is to use the basis vectors that consist of the individual bases of ε_1 and ε_2 . In order to illustrate this, let us assume the tensor product of two two-dimensional spaces. I will denote the basis vectors of ε_1 by $|\phi_a\rangle_1$ and $|\phi_b\rangle_1$ and the basis vectors of ε_2 by $|\psi_a\rangle_2$ and $|\psi_b\rangle_2$. The basis vectors of $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$ are

$$\begin{aligned} |\chi_1\rangle &= |\phi_a, \psi_a\rangle \\ |\chi_2\rangle &= |\phi_a, \psi_b\rangle \\ |\chi_3\rangle &= |\phi_b, \psi_a\rangle \\ |\chi_4\rangle &= |\phi_b, \psi_b\rangle. \end{aligned}$$

Let us now write down the representation of a product vector⁹²

$$\begin{aligned} &(x_1|\phi_a\rangle_1 + y_1|\phi_b\rangle_1) \otimes (x_2|\psi_a\rangle_2 + y_2|\psi_b\rangle_2) \\ &= x_1x_2|\chi_1\rangle + x_1y_2|\chi_2\rangle + y_1x_2|\chi_3\rangle + y_1y_2|\chi_4\rangle. \end{aligned}$$

The corresponding column vector in the $|\chi_j\rangle$ representation is $\begin{pmatrix} x_1x_2 \\ x_1y_2 \\ y_1x_2 \\ y_1y_2 \end{pmatrix}$.

Noting that the representations of the individual vectors in ε_1 and ε_2 are $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, respectively, we may write down the following relation.

⁹² Note that whenever there is a chance of ambiguity, or when it is more convenient, I will resort to the full notation of the tensor product without omitting the \otimes symbol.

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \otimes \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \\ y_1 \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} x_1 x_2 \\ x_1 y_2 \\ y_1 x_2 \\ y_1 y_2 \end{pmatrix}. \quad (237)$$

A similar relation can be shown to work for operators. Assume that \mathbf{A}_1 and \mathbf{B}_2 are operators acting in ε_1 and in ε_2 , respectively, such that their matrix representations in the above given bases are

$$\mathbf{A}_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \text{ and } \mathbf{B}_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}.$$

The combined operator can be shown to have the following representation as a matrix in the $|\chi_j\rangle$ basis

$$\mathbf{A}_1 \otimes \mathbf{B}_2 = \begin{pmatrix} a_1 \mathbf{B}_2 & b_1 \mathbf{B}_2 \\ c_1 \mathbf{B}_2 & d_1 \mathbf{B}_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 & a_1 b_2 & b_1 a_2 & b_1 b_2 \\ a_1 c_2 & a_1 d_2 & b_1 c_2 & b_1 d_2 \\ c_1 a_2 & c_1 b_2 & d_1 a_2 & d_1 b_2 \\ c_1 c_2 & c_1 d_2 & d_1 c_2 & d_1 d_2 \end{pmatrix}.$$

Some of you encountered this description, which is known as the Kronecker product. As an example, note that $\mathbf{A}_1 \otimes \mathbf{1}_2$, which is the extension of \mathbf{A}_1 to the tensor product basis is given by

$$\begin{pmatrix} a_1 & 0 & b_1 & 0 \\ 0 & a_1 & 0 & b_1 \\ c_1 & 0 & d_1 & 0 \\ 0 & c_1 & 0 & d_1 \end{pmatrix},$$

whereas the extension of \mathbf{B}_2 is represented by

$$\begin{pmatrix} a_2 & b_2 & 0 & 0 \\ c_2 & d_2 & 0 & 0 \\ 0 & 0 & a_2 & b_2 \\ 0 & 0 & c_2 & d_2 \end{pmatrix}.$$

However, this explicit representation is not particularly convenient in the general case and the more general notation, leaving the vectors and operators as products is often more useful.

Representation in an example of an infinite dimensional space

Assume for example the case of two independent particles that do not interact with each other. For simplicity we will assume that each particle exists in only one dimension. Denoting the two particles by the indices 1 and 2 and their corresponding spaces by ε_1 and ε_2 , we may describe the joint system as $|\chi\rangle = |\phi\rangle_1 \otimes |\psi\rangle_2$, implying that particle 1 is in state $|\phi\rangle_1$ and particle 2 is in state $|\psi\rangle_2$. In the abbreviated notation, this reads as $|\chi\rangle = |\phi_1, \psi_2\rangle$. We may denote the basis of subspaces by the set of state-vectors $|x\rangle_1$ and $|x'\rangle_2$, so that a basis for the tensor product space in the abbreviated notation is $|x_1, x'_2\rangle$. The x, x' representation of $|\chi\rangle$ is then

$$\langle x_1, x'_2 | \phi_1, \psi_2 \rangle = \phi(x)\psi(x'),$$

where $\phi(x)$ and $\psi(x')$ are the x representations of the wave functions of the individual particles. The x, x' representation of the tensor product is just the product of the two wave-functions, one with the argument x and the other with the argument x' .

A very similar situation exists in the case of particles that move in a three dimensional space, with no dependence between the three dimensions. A particle in a rectangular three dimensional box is an excellent example. If there is no coupling between the dimensions, then the three dimensional state-vector is a tensor product of the single dimensional vectors and the wave function is a product of three wave functions. Recalling the example of a particle in a box, which we have rehearsed in section 1, we find that in the case of a three dimensional box, the wave-function of the particle may, for example, assume the form

$$\psi(x, y, z) = u_1(x)u_2(y)u_2(z).$$

Indicating that the particle is in the ground state with respect to its motion along the x dimension and in the first excited state with respect to its motion in the y and z dimensions. Its total energy in this state is^{93, 94} $E = E_1 + 2E_2$.

⁹³ Note that the particle may also assume the state described by the wave function $u_1(x)u_2(y)u_2(z) + u_2(x)u_1(y)u_2(z)$, which is not a tensor product of three wave functions, but whose energy is still $E = E_1 + 2E_2$.

⁹⁴ As another, more general example, we may state that the space of functions in R^3 can be viewed as a tensor product of three spaces of functions in R^1 , although not all functions in R^3 are tensor products of three functions in R^1 .

Appendix II - The meaning of a complex envelope and analytic signals

Signals in nature are always real, but when they have a relatively narrow spectrum it is often convenient to treat them as complex. For example if we have a (deterministic) signal $\chi(t)$ whose spectrum is $\tilde{\chi}(\omega)$. It is clear that since $\chi(t)$ is real its fourier transform is symmetric, i.e. $\tilde{\chi}(\omega) = \tilde{\chi}^*(-\omega)$. Now what if $\chi(t)$ is narrowband in the vicinity of $\pm\omega_0$?. Then we may express its fourier transform as $\tilde{\chi}(\omega) = 1/2[\tilde{x}(\omega - \omega_0) + \tilde{x}^*(-\omega - \omega_0)]$ where $\tilde{x}(\omega)$ is a narrowband function (meaning that its width is very small relative to ω_0) and where the choice of defining this relationship with the 1/2 factor is for future convenience. The time dependence is given by

$$\begin{aligned}\chi(t) &= \frac{1}{2} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{x}(\omega - \omega_0) e^{-i\omega t} + \frac{1}{2} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{x}^*(-\omega - \omega_0) e^{-i\omega t} \\ &= \frac{1}{2} x(t) e^{-i\omega_0 t} + \frac{1}{2} x^*(t) e^{i\omega_0 t}\end{aligned}\quad (238)$$

where $x(t)$ is the inverse FT of $\tilde{x}(\omega)$ and it is called the slowly varying envelope, or the *complex envelope* of $x(t)$.

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \tilde{x}(\omega) e^{-i\omega t}\quad (239)$$

Notice that we may also write $\chi(t)$ as follows

$$\chi(t) = \text{Re} [x(t) e^{-i\omega_0 t}] = x_r(t) \cos(\omega_0 t) + x_i(t) \sin(\omega_0 t)\quad (240)$$

where $x_r(t)$ and $x_i(t)$ are the real and imaginary parts of $x(t)$ and they are called the real and imaginary quadratures (or sometimes the "in-phase" component and the "quadrature component"), respectively. In many cases we are interested in the power, or the intensity of the signal $\chi(t)$. These are proportional to $\chi^2(t) = x_r^2(t) \cos^2(\omega_0 t) + x_i^2(t) \sin^2(\omega_0 t) + 2x_r(t)x_i(t) \cos(\omega_0 t) \sin(\omega_0 t)$ and they consist of slowly varying terms as well as of terms that contain very

high frequencies of $2\omega_0$. In fact since all of our observation devices have a limited bandwidth that is usually much narrower than $2\omega_0$, but can be much broader than the bandwidth of $x(t)$, we do not observe the high frequency terms and end-up with the power being proportional to

$$(x_r^2(t) + x_i^2(t))/2 = (1/2)|x(t)|^2 \quad (241)$$

From here we may readily conclude that the two quadratures of the signal are orthogonal to each other within this approximation where oscillations at $2\omega_0$ are not detected. In other words if we want to know the power of the sum of two signals, one real and one imaginary, then it is equal to the sum of the two powers. Notice that expression (241) is proportional to the power, but it is not the power itself. Usually our $\chi(t)$ denotes the electric field so that we know that the optical power involves a proportionality coefficient related to the permeability of the vacuum.

An analytic signal is a related concept defined as follows. Assume that $\chi(t)$ is a general function (not necessarily narrow-band), describing a real physical process. Its analytic signal is defined as

$$\chi_a(t) = \int_0^\infty d\omega \tilde{\chi}(\omega) e^{-i\omega t}, \quad (242)$$

from which it is quite obvious that

$$\chi(t) = 2\Re\{\chi_a(t)\}.$$

If, in addition $\chi(t)$ is known to be a narrow-band signal that is centered around a frequency ω_0 . Then, the relation between its complex envelope $x(t)$ and its analytic signal $\chi_a(t)$ is given by

$$\chi_a(t) = \frac{1}{2}x(t)e^{-i\omega_0 t}. \quad (243)$$

If $\chi(t)$ denotes an optical field, then the field intensity is proportional to $|\chi_a(t)|^2$, or to $|x(t)|^2$.

References

1. C. Cohen-Tanoudji, B. Diu and F. Laloe *Quantum Mechanics*, vol. 1, Wiley, NY, 1977.
2. R. Loudon, *The quantum theory of light*, 3rd edition, Oxford Science Publications, 2000.
3. C. Gerry and P. Knight, *Introductory quantum optics*, Cambridge University Press, 2005.