PHY690G: Coherence and Quantum Entanglement

Anand Kumar Jha, IIT Kanpur, Semester I, 2018-19

Contents

LECTURE # 1

Classical mechanics overall is a very successful way of describing physical phenomena, in the sense that majority of modern engineering problems and situations do not require descriptions beyond what is provided by classical mechanics. However, since the beginning of the last century it started becoming clear that the actual physical reality goes beyond the classical description and requires a new way of for its description. This new way of describing physical phenomena is now known as quantum mechanics. Although for most part, it seems quantum mechanics is required to describe physical phenomena in the microscopic world while classical mechanics is sufficient for the macroscopic phenomena, modern experimental advances have now established that the differentiation in terms of length scale is not correct. In fact, it is now understood that the quantum description is *the* correct description irrespective of whether the physical system is macroscopic or microscopic. Nevertheless, in the macroscopic world the classical description is adequate for most physical phenomena and therefore in such situations one does not necessarily require quantum description.

Main differences between classical and quantum mechanics:

1. **Determinism:** Classical mechanics is in principle deterministic whereas quantum mechanics even in principle is indeterministic. If the complete information about the initial conditions of a classical system is known, it is in principle possible to precisely predict its future evolution. But this is not true for quantum systems. A quantum system remains indeterministic in general and its future evolution cannot be exactly predicted. In fact, this truly indeterministic nature of quantum mechanics is now being utilized in order to make true random number generators for various applications.

2. **Effect of measurement:** Within classical description, an act of measurement does not affect the system. However, in quantum description, one cannot make any measurement on the system without interacting with it. Therefore, an act of measurement changes the system in general. The only exception is when the system is in one of the eigen-states of the measurement operator corresponding to some physical observable. And in this case an act of measurement does not alter the system and one obtains the value of the physical observable with complete certainty.

3. **Equation of motion:** In classical mechanics, the equations of motion are in terms of real variables (position, moementum, etc.). Even when sometimes one writes the equation of motion in terms of complex variables, it is only for the sake of mathematical convenience. On the other hand, in quantum mechanics the equations of motion are written in terms of the wavefunction, which is intrinsically an abstract object. One obtains the real values corresponding to a physical observable only through an act of measurement on the wavefunction representing the quantum system.

4. **Duality**: Classical description is free from duality. A system is described either as particles using Newton's equations ($\frac{d\boldsymbol{P}}{dt} = m\frac{d^2\boldsymbol{r}}{dt^2}$) or as waves using wave equation ($\nabla^2\boldsymbol{E} = \mu_0\epsilon_0\frac{\partial^2\boldsymbol{E}}{\partial t^2}$). However, duality is the most essential feature of a quantum system. This means that a quantum system behaves simultaneously as particles and waves. This simultaneous existence of both the particle and wave properties is described using Schrödinger's equation ($\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t}$). The system is represented in terms of a wavefunction, which is an abstract quantity but y.

5. **Superposition Principle / Interference**: Wave-mechanics studies the dynamics of waves. The equations that describe this dynamics are called the wave equations. The wave equations are derived from the Maxwell's equations and are given as:

$$\nabla^2\boldsymbol{E}(\boldsymbol{r},t) = \mu_0\epsilon_0\frac{\partial^2\boldsymbol{E}(\boldsymbol{r},t)}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2\boldsymbol{E}(\boldsymbol{r},t)}{\partial t^2} \tag{1}$$

$$\nabla^2\boldsymbol{B}(\boldsymbol{r},t) = \mu_0\epsilon_0\frac{\partial^2\boldsymbol{B}(\boldsymbol{r},t)}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2\boldsymbol{B}(\boldsymbol{r},t)}{\partial t^2}, \tag{2}$$

where $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$ are the electric and magnetic fields, respectively. These equations describe the time evolution of electric and magnetic fields. One of the most important phenomena exhibited by fields is interference.

Classical mechanical description of interference is based on division of amplitude, which necessarily implies division of energy. On the other hand, the quantum mechanical description of interference in based on the division of

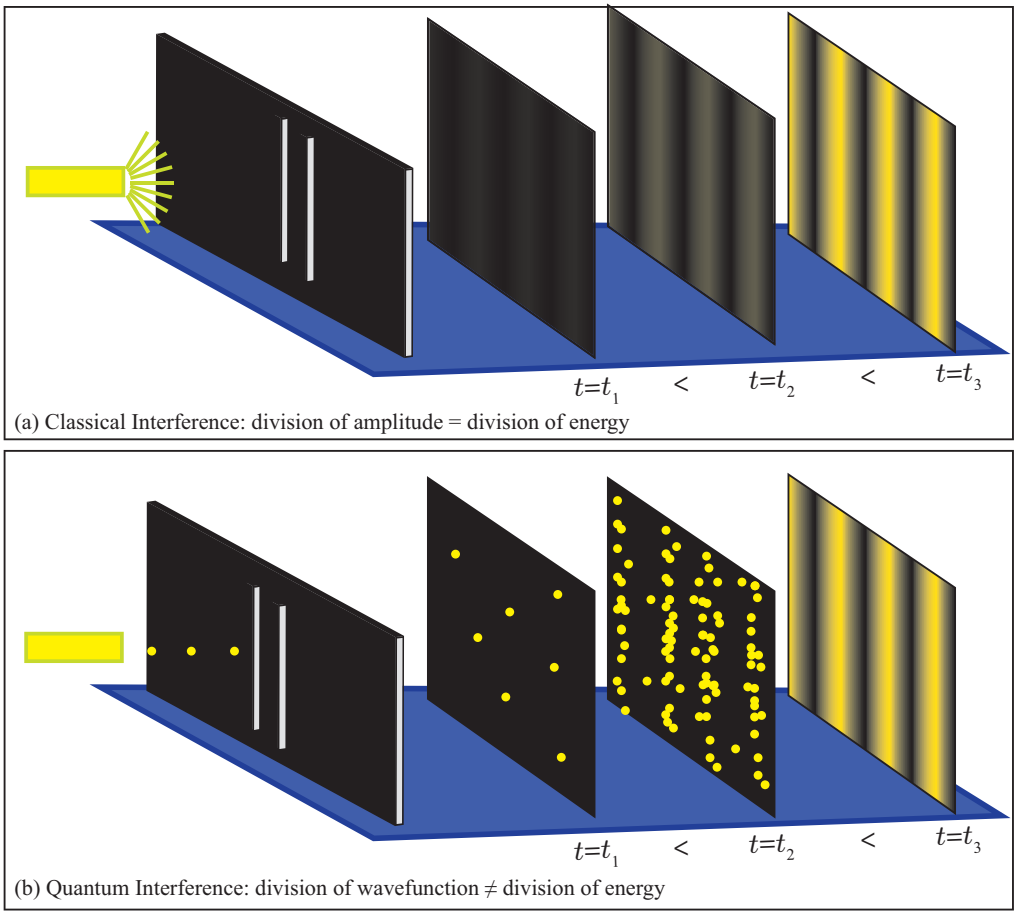


FIG. 1: Illustrating interference through classical and quantum mechanics.

wave-function which does not imply division of energy since a photon can not be divided (see Fig. 1). We now know from various experimental demonstrations that the classical description of interference works at high light levels but at very low light levels only quantum theory is able to explain the experimental results.

LECTURE # 2

Intro to coherence

If the field is deterministic, interference is perfect. However, the fields produced by real sources are more or less random. The reason for this is that a real, physical source consists of infinitely many emitters all of which emit in a statistically independent manner. As a result the total field emitted by the source becomes random. In the case of random fields, interference occurs to the extent that the fields at two different space-time points are mutually correlated. The nature and the degree of this correlation is described through what is known as a coherence function, and the equation that dictates the dynamics of a coherence function is known as the Wolf equation. Coherence as a requirement for interference remains the same both in classical and in quantum descriptions. This means that the two interfering field-amplitudes in the classical theory and the two interfering wave-functions in the quantum theory need to be mutually coherent for the interference to take place. Within quantum theory the concept of coherence can also be cast in terms of indistinguishability arguments which say that if a photon in two separate alternatives remain indistinguishable, that is, if there is no way of figuring out which particular interfering alternative a photon took then the two alternatives are mutually coherent and therefore interference will take place.

Coherence is the degree of order in a random field. Within classical description, randomness is in principle removable if complete information about the system becomes available. But it is not always possible to obtain the complete information about a system, and this is why one studies random fields through correlation functions. On the other hand, in quantum description, the randomness in not removable even in principle. So, the only concrete quantity that can be studied within quantum theory is anyway the correlation function. For some class of fields, classical and quantum descriptions of coherence yield the same average result, at least at the high light levels. However, there are many other classes of fields, which even at the high light levels can not be explained using classical coherence theory and for these fields one has to resort to the quantum theory of coherence. One such field is known as the entangled field. We will first study classical theory of coherence and then move onto the quantum theory of coherence.

Stochastic Processes: Some essential concepts

1. **Basics of probability theory**

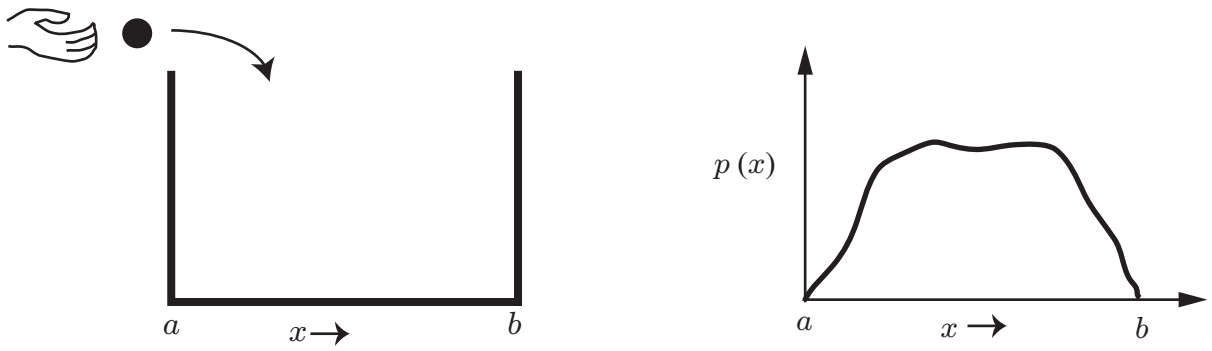


FIG. 2: Probability distribution of a random event.

First of all, let us review the basics of probability theory which will then be used to calculate different correlation function. Let us consider the situation shown in Fig.(2). A ball is thrown into a box of length $b-a$. Here, the position of the ball x is a random variable, the distribution of which depends on all kinds of driving forces acting on the box. Let $p(x)$ be the probability density that the ball after landing in the box is found at position x. Therefore, $p(x)dx$ is the probability that the position is between x and $x+dx$. Now, in the above example, the probability that the ball is found between a and b is unity. So, we have

$$\int_a^b p(x)dx = 1. \tag{3}$$

The probability that the ball is found between any other pair of points is given by

$$P(a_0 \leq x \leq b_0) = \int_{a_0}^{b_0} p(x)dx. \tag{4}$$

2. **The Ensemble Average or the Expectation Value**

For a random variable x, the collection of its all possible realizations constitute the ensemble of x. This is illustrated in Fig. 3. For a random process, individual realizations do not yield meaningful information. The most relevant

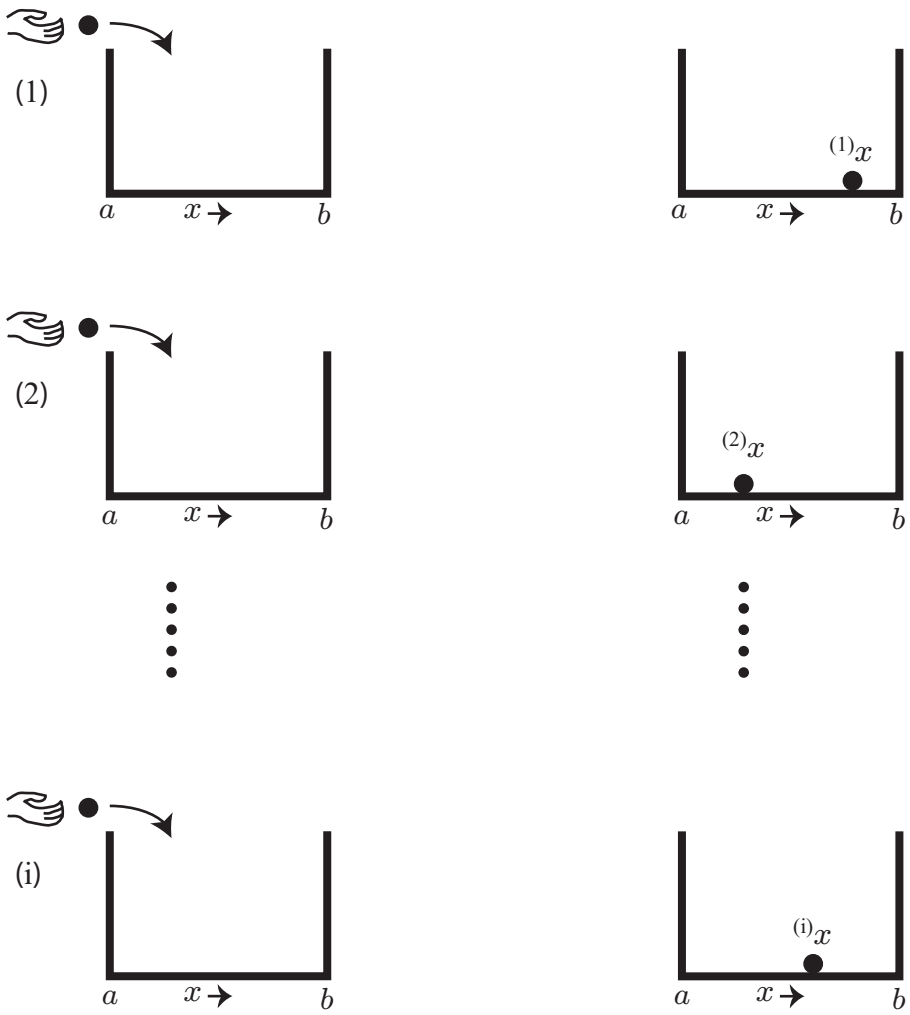


FIG. 3: Illustrating the realizations constituting an ensemble.

quantity is the ensemble average or the expectation value. The ensemble average or the expectation value is the average value of a random variable that one would obtain if the same experiment is repeated over a large number of times, and it is defined as

$$\langle x \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} {}^{(i)}x, \tag{5}$$

where ${}^{(i)}x$ is the ith realization in the ensemble. In situation in which the probability density is known the expectation value $\langle x \rangle$ can be written as

$$\langle x \rangle = \int xp(x)dx \tag{6}$$

Eq. (5) and Eq. (6) are equivalent definitions, and we will be using expectation value and ensemble average rather interchangeable in this course. Depending on whether the probability density of a random process is given or not, one uses one or the other definition.

The expectation value of x is not the only relevant quantity that one obtains in a random process. In fact, there can be a whole hierarchy of expectation values yielding different levels of information about a random process. For example, the expectation value

$$\langle x^r \rangle = \int x^r p(x) dx. \tag{7}$$

is the r^{th} order moment of the random process. The first moment is called the mean. In many situations, it is useful to define the higher order moments around mean, that is,

$$\langle (x - \langle x \rangle)^r \rangle = \int (x - \langle x \rangle)^r p(x) dx. \tag{8}$$

It is called the r^{th} central moment of x. The first central moment is always 0. The second central moment is called the variance, the square root of which is called the standard deviation $\sigma \equiv \sqrt{\langle (x - \langle x \rangle)^2 \rangle} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$. In most physical situations the mean and the standard deviation are the only relevant quantities. In fact, if the random process is Gaussian, that is, if the probability density is a Gaussian function, then all the higher order moments can be expressed in terms of the first two central moments. However, there are other physical situation where moments higher than the second order also become relevant.

3. **Joint probability**

What we have studied so far pertains to a single random variable. However, in many physical situations, there can be more than one random variables and then one is interested in the joint probability, which is the probability that the two random events happen together. Let x_1 and x_2 be two random variables and $p(x_1, x_2)$ be the joint probability density. The probability density for an individual random process can be obtained from the joint probability by integrating over the other random variable, that is, $p_1(x_1) = \int p(x_1, x_2) dx_2$ and $p_2(x_2) = \int p(x_1, x_2) dx_1$, where $p_1(x_1)$ and $p_2(x_2)$ are the probability densities corresponding to random variables x_1 and x_2, respectively. If x_1 and x_2 are completely uncorrelated then the joint probability density gets factorized, that is, $p(x_1, x_2) = p_1(x_1) p_2(x_2)$.

In the case of one random variable the quantities of interest are moments of different order as defined above. However, in the case of more than one random variables, one needs to generalize the concept of moments. The quantities of interest in this case are called correlation functions. For two random variables x_1 and x_2 the simplest correlation function that can be defined is:

$$\langle x_1 x_2 \rangle = \int x_1 x_2 p(x_1, x_2) dx_1 dx_2. \tag{9}$$

We note that if $x_1 = x_2$, it becomes the second moment $\langle x_2^2 \rangle$ of random variable x_2. Therefore it is clear that $\langle x_1 x_2 \rangle$ contains more information than just $\langle x_1^2 \rangle$ or $\langle x_2^2 \rangle$. In general a whole hierarchy of correlation functions can be defined for two random variables. For example,

$$\langle x_1^m x_2^n \rangle = \int x_1^m x_2^n p(x_1, x_2) dx_1 dx_2. \tag{10}$$

is the correlation function with m^{th} moment for x_1 and n^{th} moment for x_2. By adding time arguments the above definitions can be made even more general. For example, $\langle x_1(t_1) x_2(t_2) \rangle$ is the two-time correlation function of two random variables $x_1(t_1)$ and $x_2(t_2)$.

So far we have considered only the real random variables but in general a random variable can be complex. The complex random variables are very important. The use of complex random variables within classical theory has many advantages, including ease of calculations. In quantum theory the variables are intrinsically complex, so one cannot do without these complex random variables. Let $z_1(t_1)$ and $z_2(t_2)$ be two complex random variables. The two-time correlation function in this case is defined as $\langle z_1^*(t_1) z_2(t_2) \rangle$. We note that in this case one of the variables is taken to be the complex conjugate. This is because in physical situation only such kind of correlations turn out to be physically meaningful, even though one can define quantities such as $\langle z_1(t_1) z_2(t_2) \rangle$, $\langle z_1^*(t_1) z_2^*(t_2) \rangle$, etc. Again, a whole hierarchy of such correlations can be defined for the complex random variables in the same way as is done for real random variables. For example, one can have correlation such as

$\langle z_1^*(t_1)z_2^*(t_2)\cdots z_N^*(t_N)z_{N+1}(t_{N+1})z_{N+2}(t_{N+2})\cdots z_{2N}(t_{2N})\rangle$. We note that we have taken equal number of complex conjugate variables. This is only because correlation functions with unequal numbers of complex and complex-conjugate quantities are not very frequently encountered in physical situation and in this course we will only consider correlation functions with equal number of complex and complex-conjugate quantities.

As a special case, we note that the two-time correlation function $\langle z_1^*(t_1)z_2(t_2)\rangle$ is also called the mutual correlation function of random variables $z_1(t_1)$ and $z_2(t_2)$. In situation in which $z_1 = z_2 = z$, the correlation function $\langle z^*(t_1)z(t_2)\rangle$ is referred to as the auto-correlation function.

LECTURE # 3

4. **Complex Analytic Signal Representation**

We have seen that the random variables can be real or complex. In a physical theory one is concerned about the dynamics of real-valued quantities, such as, current, charge densities, etc. Within the classical description, the dynamical equations are given in terms of real physical quantities, whereas in quantum description the dynamical equations are written for wave functions, a necessarily complex quantity. However, as far as calculations are concerned, it turns out that dealing with complex random variables is much easier than dealing with real random variables. So, within classical description, a real random variable is first mapped onto a complex random variable. The dynamical evolution is studied with the complex random variable and at the end of a calculation the physical quantities are extracted by taking the real part of appropriate quantities. This mapping of real random variables to the complex one is called the complex analytic signal representation of real random variable. The analytic signals themselves are purely mathematical quantities and do not necessarily have any direct physical meaning but they make the calculations much easier within classical description.

Let $x(t)$ be a real function of a real variable t. Let us also assume that the Fourier transform of $x(t)$ exists such that

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} \tilde{x}(\omega) e^{-i\omega t} d\omega \\ &= \int_{-\infty}^{0} \tilde{x}(\omega) e^{-i\omega t} d\omega + \int_{0}^{\infty} \tilde{x}(\omega) e^{-i\omega t} d\omega \\ &= z^*(t) + z(t) \end{aligned}$$

Here, $z(t)$ is the complex analytic signal associated with the real random variable $x(t)$. We can write $z(t)$ as a Fourier transform:

$$z(t) = \int_{-\infty}^{\infty} \tilde{z}(\omega) e^{-i\omega t} d\omega$$

where

$$\begin{aligned} \tilde{z}(\omega) &= \tilde{x}(\omega) \qquad \text{when} \qquad \omega \geq 0 \\ &= 0 \qquad\quad \text{when} \qquad \omega < 0 \end{aligned}$$

We note that $z(t)$ is a complex analytic signal and therefore it is single valued and has continuous derivatives. The real random variable $x(t)$ can be obtained by taking the real part of $z(t)$, that is, $x(t) = z(t) + z^*(t) = 2\text{Re}z(t)$. Now since $x(t)$ is real, $x^*(t) = x(t)$. But $x^*(t)$ can be written as

$$\begin{aligned} x^*(t) &= \int_{-\infty}^{\infty} \tilde{x}^*(\omega) e^{+i\omega t} d\omega \\ &= \int_{\infty}^{-\infty} \tilde{x}^*(-\omega) e^{-i\omega t} d(-\omega) \qquad \text{(By substituting } \omega \text{ with } -\omega) \\ &= \int_{-\infty}^{\infty} \tilde{x}^*(-\omega) e^{-i\omega t} d\omega \end{aligned}$$

This simply implies that $\tilde{x}^*(-\omega) = \tilde{x}(\omega)$ and therefore that the positive frequency part of the signal contains as much information as the negative frequency part. It is known that for a complex analytic signal the real and imaginary parts are connected by a Hilbert transform. Here, $z(t)$ is a complex analytic signal. The real part of $z(t)$ is $\frac{1}{2}x(t)$. Let us denote its imaginary part by $\frac{1}{2}y(t)$, that is, $z(t) = \frac{1}{2}[x(t) + iy(t)]$. The real signals $x(t)$ and $y(t)$ then form the Hilbert transform pair.

$$\begin{aligned} y(t) &= \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(t')}{t' - t} dt' \\ x(t) &= -\frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{y(t')}{t' - t} dt', \end{aligned}$$

where the Cauchy's principle value is defined as

$$P\int_{-\infty}^{\infty}\frac{x(t')}{t'-t}dt' = \lim_{\delta\to 0}\left[\int_{-\infty}^{t-\delta}\frac{x(t')}{t'-t}dt' + \int_{t+\delta}^{\infty}\frac{x(t')}{t'-t}dt'\right].$$

The Joint probability function used in classical optics: The correlation functions

So far, we have seen how a real signal corresponding to a physical process can be represented in terms of a complex analytic signal for the ease of mathematical calculations. We have also studied how from a given random variable one can construct correlation function of different orders so as to extract relevant information up to various levels. We now consider the real classical electric field and study its correlation properties within the framework that we have just developed. Let us consider the classical electric field due to a chaotic light source. Let us take the electric field to be a real random signal $V^{(\mathrm{r})}(\boldsymbol{r},t)$. We represent this real signal in terms of the complex analytic signals

$$\begin{aligned} V^{(\mathrm{r})}(\boldsymbol{r},t) &= \int_{-\infty}^{\infty}\tilde{v}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\ &= \int_{-\infty}^{0}\tilde{v}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega + \int_{0}^{\infty}\tilde{v}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\ &= V^{*}(\boldsymbol{r},t) + V(\boldsymbol{r},t). \end{aligned}$$

Here, $V(\boldsymbol{r},t)$ is the complex analytic signal corresponding to the real random variable $V^{(\mathrm{r})}(\boldsymbol{r},t)$. Also, since $V^{(\mathrm{r})}(\boldsymbol{r},t)$ is real, we have $\tilde{v}^{*}(\boldsymbol{r},\omega) = \tilde{v}(\boldsymbol{r},-\omega)$. Let us look at the first few correlation functions that can be constructed using this analytic signal.

1. **Intensity:** $I(r,t) \equiv \langle V^{*}(\boldsymbol{r},t)V(\boldsymbol{r},t)\rangle$**:** This expectation value is the simplest second order correlation function corresponding to the analytic signal. The quantity $V^{*}(\boldsymbol{r},t)V(\boldsymbol{r},t)$ is the instantaneous intensity of the field at $(\boldsymbol{r},t)$, and the expectation value is the ensemble-averaged intensity of the field at $(\boldsymbol{r},t)$.

2. **Cross-correlation function:** $\Gamma(\boldsymbol{r_1},t,\boldsymbol{r_2},t_2) \equiv \langle V^{*}(\boldsymbol{r_1},t_1)V(\boldsymbol{r_2},t_2)\rangle$**:** The next expectation value we consider is the two-point correlation function. This is also called the cross-correlation function.

3. **Higher-order correlation function:** $\Gamma(\boldsymbol{r_1},t,\boldsymbol{r_2},t_2;\boldsymbol{r_3},t_3;\boldsymbol{r_4},t_4) \equiv \langle V^{*}(\boldsymbol{r_1},t_1)V^{*}(\boldsymbol{r_2},t_2)V(\boldsymbol{r_3},t_3)V(\boldsymbol{r_4},t_4)\rangle$**:** This is a four-point correlation function. In this course, we will not be studying it much within the classical theory but in the context of quantum theory of coherence, the four-points correlation functions will be studied in greater details.

LECTURE # 4

Second-order coherence theory (Temporal)

In this section, we will look at the temporal coherence effects and ask the question as to how field vibrations at a given space point $\boldsymbol{r_0}$ at time t_1 are correlated with field vibrations at the same point but at a different time t_2. To answer this question, we consider a Mach-Zehnder interferometer, as shown in Fig. 4. The field at position $\boldsymbol{r_0}$ is divided into two by the beam splitter, and the two fields reach the detector after having travelled for times t_1 and t_2. The electric field $V(\boldsymbol{r}, t)$ at the detector position $\boldsymbol{r}$ at time t is given by

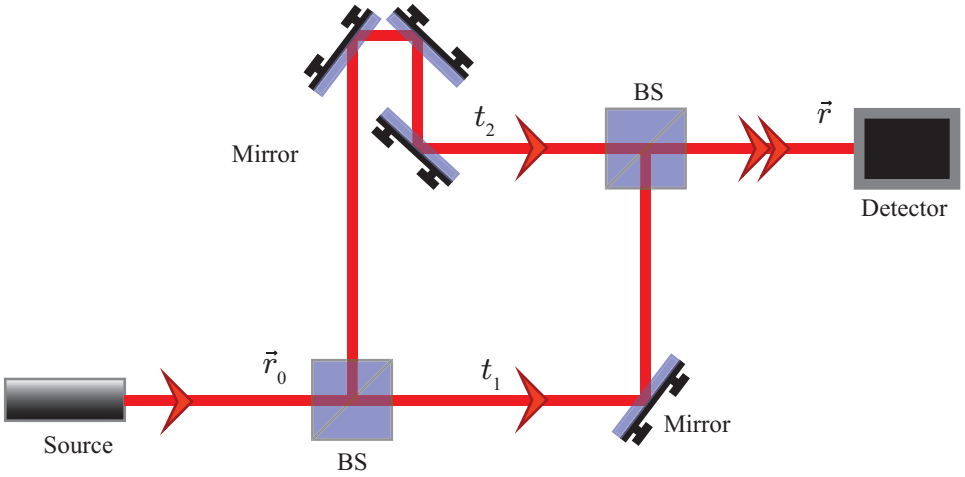


FIG. 4: Mach-Zehnder Interferometer.

$$V(\boldsymbol{r}, t) = k_1 V(\boldsymbol{r_0}, t - t_1) + k_2 V(\boldsymbol{r_0}, t - t_2).$$

Here $V(\boldsymbol{r_0}, t - t_1)$ and $V(\boldsymbol{r_0}, t - t_2)$ are the electric field vibrations at the source point $\boldsymbol{r_0}$ at time $t - t_1$ and $t - t_2$, respectively. Therefore, by studying the intensity at the detector one can study how the field vibrations that are separated by time $t_1 - t_2$ are correlated. But before we start calculating the expression for intensity at the detector, let us first go through some essential mathematical concepts.

Temporal correlations: some essential mathematical concepts

(1) **Stationarity (temporal):** A random process is said to be stationary when *all* the correlation functions are independent of the origin of time. For such processes, one has

$$\begin{aligned}\langle V^*(\boldsymbol{r_1}, t_1) \cdots V^*(\boldsymbol{r_n}, t_n) V(\boldsymbol{r_{n+1}}, t_{n+1}) \cdots V(\boldsymbol{r_{2n}}, t_{2n}) \rangle \\ = \langle V^*(\boldsymbol{r_1}, t_1 - t) \cdots V^*(\boldsymbol{r_n}, t_n - t) V(\boldsymbol{r_{n+1}}, t_{n+1} - t) \cdots V(\boldsymbol{r_{2n}}, t_{2n} - t) \rangle \end{aligned} \quad (11)$$

In most physical situation, however, it is very difficult to calculate all the correlation functions. In such cases, one looks for the stationarity-in-the-wide-sense. A field is said to be stationary in the wide sense if the mean and the two-time correlation function are independent of the origin of time,that is,

$$\begin{aligned}&\langle V(\boldsymbol{r}, t) \rangle = \langle V(\boldsymbol{r}, t - t) \rangle = \langle V(\boldsymbol{r}, 0) \rangle \\ &\implies \langle V^*(\boldsymbol{r}, t) V(\boldsymbol{r}, t) \rangle = \langle V^*(\boldsymbol{r}, 0) V(\boldsymbol{r}, 0) \rangle \implies I(\boldsymbol{r}, t) = I(\boldsymbol{r}, 0) = \text{constant} \end{aligned} \quad (12)$$

$$\begin{aligned}\text{and} \quad \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) &= \langle V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2) \rangle = \langle V^*(\boldsymbol{r_1}, t_1 - t_1 + t) V(\boldsymbol{r_2}, t_2 - t_1 + t) \rangle \\ &= \langle V^*(\boldsymbol{r_1}, t) V(\boldsymbol{r_2}, t + \tau) \rangle = \langle V^*(\boldsymbol{r_1}, 0) V(\boldsymbol{r_2}, \tau) \rangle = \Gamma(\boldsymbol{r_1}, \boldsymbol{r_2}, \tau), \end{aligned} \quad (13)$$

where $\tau = t_2 - t_1$. We find that if a process is stationary in the wide sense, the mean becomes independent of the origin of time and the two-time cross-correlation function depends only the difference of the two time arguments.

For notation simplifications, when $\boldsymbol{r_1} = \boldsymbol{r_2}$, we write $\Gamma(\boldsymbol{r_1}, \boldsymbol{r_1}, \tau) = \Gamma(\tau)$. The cross-correlation function $\Gamma(\tau)$ has the following properties:

$$\Gamma(0) \geq 0 \tag{14}$$
$$|\Gamma(\tau)| \leq \Gamma(0) \tag{15}$$
$$\Gamma(-\tau) = \Gamma^*(\tau) \tag{16}$$

(2) **Spectral representation of a time-domain random process**

Quite often, it is much more convenient to study the temporal correlations of a field in the frequency basis. In order to do this, we need the frequency or the spectral representation of random processes. If we assume that the Fourier transform of a random process $V(\boldsymbol{r}, t)$ exist and also that its Fourier integral can be inverted then we can write

$$V(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} \tilde{v}(\boldsymbol{r}, \omega) e^{-i\omega t} d\omega$$
$$\tilde{v}(\boldsymbol{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} V(\boldsymbol{r}, t) e^{i\omega t} dt$$

However, $V(\boldsymbol{r}, t)$ is a random process and its Fourier transform does not exist in the usual sense because $V(\boldsymbol{r}, t)$ is not absolutely integrable, that is, the requirement, $\int_{-\infty}^{\infty} |V(\boldsymbol{r}, t)| dt < \infty$ is not satisfied. Although this problem was overcome by Wiener by introducing what is known as the "generalized harmonic analysis," we will not be using the generalized harmonic functions in this course. Even though $V(\boldsymbol{r}, t)$ is random, we know that the moments of a stationary random process are not random. For example, the auto-correlation function of a stationary random process is absolutely integrable. Therefore, we will derive the spectrum in a way that is based on auto-correlation function being absolutely integrable. Let us for now use the above Fourier transform relations rather symbolically and calculate the ensemble average of the product $V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2)$.

$$\begin{aligned} \langle V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2) \rangle &= \left\langle \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{v}^*(\boldsymbol{r_1}, \omega_1) \tilde{v}(\boldsymbol{r_2}, \omega_2) e^{i\omega_1 t_1} e^{-i\omega_2 t_2} d\omega_1 d\omega_2 \right\rangle \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \tilde{v}^*(\boldsymbol{r_1}, \omega_1) \tilde{v}(\boldsymbol{r_2}, \omega_2) \rangle \, e^{i\omega_1 t_1} e^{-i\omega_2 t_2} d\omega_1 d\omega_2 \end{aligned} \tag{17}$$

The above equation can be written as

$$\Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\boldsymbol{r_1}, \omega_1; \boldsymbol{r_2}, \omega_2) e^{i\omega_1 t_1} e^{-i\omega_2 t_2} d\omega_1 d\omega_2, \tag{18}$$

where we have defined $W(\boldsymbol{r_1}, \omega_1; \boldsymbol{r_2}, \omega_2) \equiv \langle \tilde{v}^*(\boldsymbol{r_1}, \omega_1) \tilde{v}(\boldsymbol{r_2}, \omega_2) \rangle$ as the cross-spectral density function. This equation is the spectral representation of cross-correlation function. The cross-spectral density function quantifies the correlation between different frequency components of a random field. We find that in the most general situations, the cross-correlation function $\Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2)$ is given by the two-dimensional Fourier transform of the cross-spectral density function $W(\boldsymbol{r_1}, \omega_1; \boldsymbol{r_2}, \omega_2)$. The above equation can be inverted and be written as

$$W(\boldsymbol{r_1}, \omega_1; \boldsymbol{r_2}, \omega_2) = \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) e^{-i\omega_1 t_1} e^{+i\omega_2 t_2} dt_1 dt_2, \tag{19}$$

The pair of equations between the cross-correlation function and the cross-spectral density function is referred to as the generalized Wiener-Khintchine theorem. The two functions are quite often expressed in the following normalized:

$$\gamma(t_1, t_2) = \frac{\Gamma(t_1, t_2)}{\sqrt{\Gamma(t_1, t_1)\Gamma(t_2, t_2)}} = \frac{\Gamma(t_1, t_2)}{\sqrt{I(t_1) I(t_2)}} \tag{20}$$
$$\mu(\omega_1, \omega_2) = \frac{W(\omega_1, \omega_2)}{\sqrt{W(\omega_1, \omega_1) W(\omega_2, \omega_2)}} = \frac{W(\omega_1, \omega_2)}{\sqrt{S(\omega_1) S(\omega_2)}} \tag{21}$$

Here, $\gamma(t_1, t_2)$ is called the complex degree of temporal coherence and $\mu(\omega_1, \omega_2)$ is called the complex degree of spectral coherence. The magnitudes $|\gamma(t_1, t_2)|$ and $|\mu(\omega_1, \omega_2)|$ are called the degree of temporal coherence and the

degree of spectral coherence, respectively. Both these degrees range from 0 to 1, that is, $0 \leq |\gamma(t_1, t_2)| \leq 1$ and $0 \leq |\mu(\omega_1, \omega_2)| \leq 1$.

Now, for stationary random processes, the two-time cross-correlation function depends only on the difference of the time arguments, that is, $\Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) = \Gamma(t_1, t_2) \rightarrow \Gamma(\tau)$, where $\tau = t_2 - t_1$. In order that the right hand side of Eq. (18) also depends only on τ, we have to have

$$W(\omega_1, \omega_2) = \langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle = S(\omega_1)\delta(\omega_1 - \omega_2) \tag{22}$$

where $S(\omega)$ is called the spectral density and it represents the strength of individual monochromatic components in a random field.

The cross-correlation correlation can therefore be represented as

$$\Gamma(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{-i\omega\tau} d\omega \tag{23}$$

We note that the spectral density and the correlation function $\Gamma(\tau)$ do form a Fourier transform pair. The above integral can be inverted and we can write it as

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Gamma(\tau) e^{i\omega\tau} d\tau \tag{24}$$

Equations (23) and (24) representing the Fourier transform relations between the spectral density and the temporal correlation function are known as the Wiener-Khintchine theorem.[1]

(3) **Time Average:** In addition to the concept of ensemble average or expectation value, the other average used is the time average, which is a more useful concept since in realistic physical situations it is impossible to calculate the ensemble average because that requires having many realizations of the physical phenomenon. So, in realistic physical situations, in which one is left with a single realization of the phenomenon, the quantity of interest in the time average. The time-average of a quantity $f(\boldsymbol{r}, t)$ is defined as:

$$\langle f(\boldsymbol{r}, t)\rangle_{\mathrm{t}} = \lim_{T\rightarrow\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(\boldsymbol{r}, t) dt \tag{26}$$

We note that the time-average is defined for a particular realization while ensemble average is defined as the average over infinitely many different realizations. We note that although ensemble averages are always equal to the expectation values, time-averages need not always be equal to the expectation value. It is only for a particular class of random process known as the ergodic process that the time-average becomes equal to the ensemble average and thus becomes equal to the expectation value. So, only for ergodic processes, ensemble averages can be replaced by the time-averages.

(4) **Ergodicity:** A stationary random process is called ergodic if it satisfies the following condition:

$$\int_{-\infty}^{\infty} \left|\Gamma(\tau) - |\langle V(0)\rangle|^2\right| d\tau < \infty \qquad \text{or} \qquad \lim_{\tau\rightarrow\infty} \left|\Gamma(\tau) - |\langle V(0)\rangle|^2\right| \rightarrow 0 \tag{27}$$

Here the idea is that $\Gamma(\tau)$ should tend to $|\langle V(0)\rangle|^2$ rapidly such that a sufficiently long time-realization can be broken up into sections of shorter time-realization, each of which still remains longer than the time-period over

---

[1] Please note that it is quite tempting to define the spectral density as $S(\omega_1) \equiv \langle|\tilde{v}(\omega)|^2\rangle$ but this is an incorrect definition. In terms of the frequency correlations $\langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle$, the correct definition is

$$S(\omega_1) \equiv \lim_{\Delta\omega\rightarrow 0} \int_{\omega_1 - \frac{\Delta\omega}{2}}^{\omega_1 + \frac{\Delta\omega}{2}} \langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle d\omega_2, \tag{25}$$

which $\Gamma(\tau)$ tends to $|\langle V(0)\rangle|^2$. The average over this ensemble of sections then equals the time-average over one such section. Therefore, if the field represented by $V(\boldsymbol{r},t)$ is stationary and ergodic, the ensemble averages can be replaced by time averages. Therefore, the ensemble-average $\langle V^*(\boldsymbol{r},t)V(\boldsymbol{r},t)\rangle$ can be treated as time-averaged intensity. However, in classical electrodynamics the intensity is defined as the quantity $\langle [V^{(\mathrm{r})}(\boldsymbol{r},t)]^2\rangle_t$, which involves time-averaging over a few cycles. So, are these two definitions equivalent? In fact, it can be shown that the two definitions are indeed equivalent and are related as

$$\langle [V^{(\mathrm{r})}(\boldsymbol{r},t)]^2\rangle = \frac{1}{2}\langle V^*(\boldsymbol{r},t)V(\boldsymbol{r},t)\rangle$$

So, the complex analytic signal representation does yield a definition of intensity that is equivalent to the definition of intensity in classical electromagnetism. There are several other advantages of using the complex analytic signal representation which will become clear later in this course.

LECTURE # 5

Quantifying the temporal correlations

We are now prepared to answer the question as to how to quantify the correlation between the field vibrations at $t-t_1$ and $t-t_2$. We refer back to Fig. 4. The electric field $V(\boldsymbol{r},t)$ at the detector position $\boldsymbol{r}$ at time t is given by

$$V(\boldsymbol{r},t) = k_1 V(\boldsymbol{r_0}, t-t_1) + k_2 V(\boldsymbol{r_0}, t-t_2).$$

Here $V(\boldsymbol{r_0}, t-t_1)$ and $V(\boldsymbol{r_0}, t-t_2)$ are the electric field vibrations at the source point $\boldsymbol{r_0}$ at time $t-t_1$ and $t-t_2$, respectively. Therefore, the intensity at the detector is

$$I(\boldsymbol{r},t) = \langle V^*(\boldsymbol{r},t)V(\boldsymbol{r},t)\rangle \tag{28}$$

$$= |k_1|^2\langle V^*(\boldsymbol{r_0},t-t_1)V(\boldsymbol{r_0},t-t_1)\rangle + |k_2|^2\langle V^*(\boldsymbol{r_0},t-t_2)V(\boldsymbol{r_0},t-t_2)\rangle + k_1^*k_2\langle V^*(\boldsymbol{r_0},t-t_1)V(\boldsymbol{r_0},t-t_2)\rangle + \text{c.c.} \tag{29}$$

We see that the intensity at the detector contains information about correlations in the field. The above equation can be written as

$$I(\boldsymbol{r},t) = |k_1|^2 I(\boldsymbol{r_0},t-t_1) + |k_2|^2 I(\boldsymbol{r_0},t-t_2) + k_1^*k_2\Gamma(\boldsymbol{r_0},t-t_1;\boldsymbol{r_0},t-t_2) + \text{c.c}, \tag{30}$$

where we have

$$\langle V^*(\boldsymbol{r_0},t-t_1)V(\boldsymbol{r_0},t-t_1)\rangle = I(\boldsymbol{r_0},t-t_1) = I_0(t-t_1)$$
$$\langle V^*(\boldsymbol{r_0},t-t_2)V(\boldsymbol{r_0},t-t_2)\rangle = I(\boldsymbol{r_0},t-t_2) = I_0(t-t_2)$$
$$\langle V^*(\boldsymbol{r_0},t-t_1)V(\boldsymbol{r_0},t-t_2)\rangle = \Gamma(\boldsymbol{r_0},t-t_1;\boldsymbol{r_0},t-t_2) = \Gamma(t-t_1,t-t_2).$$

Suppressing the position variables, we write the above equation as

$$I_r(t) = |k_1|^2 I(t-t_1) + |k_2|^2 I(t-t_2) + k_1^*k_2\Gamma(t-t_1,t-t_2) + \text{c.c}, \tag{31}$$

where $I_r(t) = I(\boldsymbol{r},t)$. Equation (31) represents the general temporal interference law. It can describe interference due to fields of any type. We see that the detector intensity at time t depends on the source intensities at time $t-t_1$ and $t-t_2$ as well as the cross-correlation function of the source. Therefore, through the intensity measurements at the detector the temporal correlations of the source can in principle be extracted.

Now, we take $k_1^*k_2 = |k_1||k_2|e^{i\phi_k}$ and write the cross correlation function as

$$\Gamma(t-t_1,t-t_2) = \sqrt{I(t-t_1)I(t-t_2)}\frac{\Gamma(t-t_1,t-t_2)}{\sqrt{I(t-t_1)I(t-t_2)}} = \sqrt{I(t-t_1)I(t-t_2)}|\gamma(t-t_1,t-t_2)|e^{i\arg\gamma}. \tag{32}$$

Therefore, the temporal interference law can now be written as

$$I_r(t) = |k_1|^2 I(t-t_1) + |k_2|^2 I(t-t_2) + 2|k_1||k_2|\sqrt{I(t-t_1)I(t-t_2)}|\gamma(t-t_1,t-t_2)|\cos(\arg\gamma + \phi_k) \tag{33}$$

We see that the intensity at the detector depends on $\arg\gamma$ and it changes in a sinusoidal manner as a function of $\arg\gamma$. This sinusoidal dependence is multiplied by the degree of coherence function $|\gamma(t-t_1,t-t_2)|$. The visibility of the interference pattern is defined in the following manner by

$$V = \frac{I_{\max}-I_{\min}}{I_{\max}+I_{\min}} = \frac{2|k_1||k_2|\sqrt{I(t-t_1)I(t-t_2)}}{|k_1|^2I(t-t_1)+|k_2|^2I(t-t_2)}|\gamma(t-t_1,t-t_2)|$$

We see that the visibility is proportional to the degree of temporal coherence. And for the special case when $|k_1|^2I(t-t_1) = |k_2|^2I(t-t_2)$, we have $V = |\gamma(t-t_1,t-t_2)|$, that is, the fringe visibility equals the degree of coherence.

Now, we know that the cross-correlation function is connected to the cross-spectral density function through the generalized Wiener-Khintchine theorem as

$$\Gamma(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(\omega_1,\omega_2)e^{-i(\omega_2t_2-\omega_1t_1)}d\omega_1d\omega_2, \tag{34}$$

$$W(\omega_1,\omega_2) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Gamma(t_1,t_2)e^{i(\omega_2t_2-\omega_1t_1)}dt_1t_2, \tag{35}$$

which says that the two-time temporal correlation function $\Gamma(t_1, t_2) = \langle V^*(t_1)V(t_2)\rangle$ is a two-dimensional Fourier transform of the cross-spectral density function $W(\omega_1, \omega_2) = \langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle$. Therefore, depending on the frequency correlations of the source, a source field can be divided into the following three categories:

1. **Case-I: No correlation between different frequency components**

   When the source has no correlations between its different frequency components, the cross-spectral density function is given by $W(\omega_1, \omega_2) = S(\omega_1)\delta(\omega_1 - \omega_2)$. The degree of spectral coherence in this case is $|\mu(\omega_1, \omega_2)| = \delta(\omega_1 - \omega_2)$. As we have seen before, having no frequency correlations implies that the source is stationary in time. This means that the cross-correlation function depends only on time-difference and the intensity of the source is independent of the origin of time and that the ensemble averages of Eq. (31) can be replaced by time-averages. We therefore have

   $$\begin{aligned} I(t - t_1) &= I \\ I(t - t_2) &= I \\ \Gamma(t - t_1, t - t_2) &= \Gamma(\tau). \end{aligned}$$

   We find that for sources having no frequency correlations, the intensity remains constant over time and the degree of temporal correlation depends only on the time difference. Such sources are also called continuous-wave (CW) fields. Many common lasers, such as He-Ne Lasers, Laser pointers have CW fields.

2. **Case-II: Perfect correlation between different frequency components**

   When the source has perfect correlations between its different frequency components, the cross-spectral density function is given by $W(\omega_1, \omega_2) = \langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle = \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)$. The temporal correlation function and the intensity of such a source are thus given by

   $$\begin{aligned} \Gamma(t_1, t_2) &= \int_{-\infty}^{\infty} \tilde{v}^*(\omega_1) e^{i\omega_1 t_1} d\omega_1 \int_{-\infty}^{\infty} \tilde{v}(\omega_2) e^{-i\omega_2 t_2} d\omega_2, \\ I(t) = \Gamma(t, t) &= \left| \int_{-\infty}^{\infty} \tilde{v}(\omega) e^{-i\omega t} d\omega \right|^2 \end{aligned}$$

   We find that in this case the temporal correlation function and the intensity both depend on time arguments. The Intensity has a temporal structure, the exact form of which depends on $\tilde{v}(\omega)$. The degree of spectral coherence in this case becomes $|\mu(\omega_1, \omega_2)| = 1$, indicating the perfect correlation between different frequencies. The degree of temporal coherence is also always equal to one, that is, $\gamma(t_1, t_2) = 1$. This type of correlations are found in pulsed laser. The pulsed structure in the intensity profile is due to the fact that all the frequencies $\tilde{v}(\omega)$ add up coherently.

   Measuring the temporal correlation function in this case is a little difficult as in this case, one cannot simply replace the ensemble average by the time-average. However, there are several techniques using which some or all of the information about these kinds of fields are extracted.

3. **Case-III: Partial correlation between different frequency components**

   Most realistic sources are effective partially correlated in their frequencies. The types of sources discussed above are the two extreme points. A source can be partially correlated in many different ways. However, we will study the most commonly used model, which is known as the Gaussian-Schell model source. The cross-spectral density of a Gaussian-Schell model source is given by

   $$W(\omega_1, \omega_2) = \langle \tilde{v}^*(\omega_1)\tilde{v}(\omega_2)\rangle = A \exp\left[-\frac{(\omega_1 - \omega_0)^2 + (\omega_2 - \omega_0)^2}{2\Omega^2}\right] \exp\left[-\frac{(\omega_1 - \omega_2)^2}{2\Omega_c^2}\right]$$

   In this case, the temporal correlation function and the intensity both depend on time arguments. The degree of spectral coherence in this case becomes $|\mu(\omega_1, \omega_2)| = \exp\left[-\frac{(\omega_1 - \omega_2)^2}{2\Omega_c^2}\right]$. Thus we find that in the limit $\Omega_c \to 0$, the above source essentially becomes a CW source. And in the limit $\Omega_c \to \infty$, $|\mu(\omega_1, \omega_2)| = 1$, and therefore the Gaussian Schell model source, in the limit $\Omega_c \to \infty$, represents a source having perfect frequency correlation.

The special case of a stationary quasi-monochromatic field with Gaussian spectrum

For the special case of a stationary field, that is, no correlation between different frequency components, we have $\Gamma(t-t_1, t-t_2) = \Gamma(\tau)$ and $I(t-t_1) = I(t-t_2) = I$. So, we can write Eq. (33) as

$$I_r = |k_1|^2 I + |k_2|^2 I + 2|k_1||k_2| I |\gamma(\tau)| \cos(\arg\gamma + \phi_k).$$

In many situations, the field of interest has a spectrum that is quasi-monochromatic. This means that the spectral width $\Delta\omega$ of the source is very small compared to the central frequency ω_0 of the source, that is, $\Delta\omega \ll \omega_0$. From our discussions in the earlier section, we know that the auto-correlation function of a stationary random field is the Fourier transform of the spectral density of the field, that is,

$$\Gamma(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{-i\omega\tau} d\omega.$$

Substituting $\omega_0 + \omega_2$ for ω and assuming quasi-monochromaticity $\Delta\omega \ll \omega_0$, we can write the auto-correlation function as

$$\Gamma(\tau) = \int_{-\infty-\omega_0}^{\infty-\omega_0} S(\omega_0+\omega_2) e^{-i(\omega_0+\omega_2)\tau} d\omega_2 \approx e^{-i\omega_0\tau} \int_{-\infty}^{\infty} S(\omega_0+\omega_2) e^{-i\omega_2\tau} d\omega_2 = e^{-i\omega_0\tau} \int_{-\infty}^{\infty} S(\omega_0+\omega) e^{-i\omega\tau} d\omega$$

We now consider the spectral density $S(\omega)$ to be a normalized Gaussian function given as

$$S(\omega) = \frac{1}{\sqrt{2\pi}\Delta\omega} \exp\left[-\frac{(\omega-\omega_0)^2}{2\Delta\omega^2}\right].$$

Here $\Delta\omega$ is the width of the spectral-density function and is called the frequency-bandwidth or the spectral-bandwidth of the source. So, for the above spectrum $I = \Gamma(0) = 1$. Now, using the Wiener-Khintchine theorem, we find the temporal correlation function for the above spectral density to be

$$\Gamma(\tau) = e^{-i\omega_0\tau} \exp\left[-\frac{\tau^2}{2\tau_c^2}\right],$$

where $\tau_c = 1/\Delta\omega$. The complex degree of coherence function can therefore be written as

$$\gamma(\tau) = \Gamma(\tau)/\Gamma(0) = \Gamma(\tau) = e^{-i\omega_0\tau} \exp\left[-\frac{\tau^2}{2\tau_c^2}\right],$$

Thus, we obtain the following expression for the intensity at the detector

$$I(\boldsymbol{r}) = |k_1|^2 + |k_2|^2 + 2|k_1||k_2| \exp\left[-\frac{\tau^2}{2\tau_c^2}\right] \cos(\omega_0\tau + \phi_k).$$

Coherence time

We will not attempt to define the coherence time τ_c in a formal manner. However, for Gaussian spectrum considered above it can be defined in terms of the temporal separation between two points after which the degree of correlation starts decreasing significantly, the standard deviation of a Gaussian distribution can therefore be taken as the coherence time. However, in situations in which the degree of coherence function is period, one has to be very careful about defining coherence time. So, for the above Gaussian spectrum, $\tau_c = 1/\Delta\omega$ is taken as the coherence time. The length travelled by light $c\tau_c$ during the coherence time is referred to as the coherence length. Below is a comparison table of the frequency-bandwidth and the coherence time for a few stationary random sources.

Typical source	frequency-bandwidth ($\Delta\omega$)	Coherence time (τ_c)	Coherence length ($= c\tau_c$)
very well-stabilized lasers	1 MHz (10^6 Hz)	10^{-6} s	300 m
diode lasers	1 GHz (10^9 Hz)	10^{-9} s	0.3 m
Sunlight	100 THz (10^{14} Hz)	10^{-14} s	3 μm

LECTURE # 6

Second-order coherence theory (Spatial)

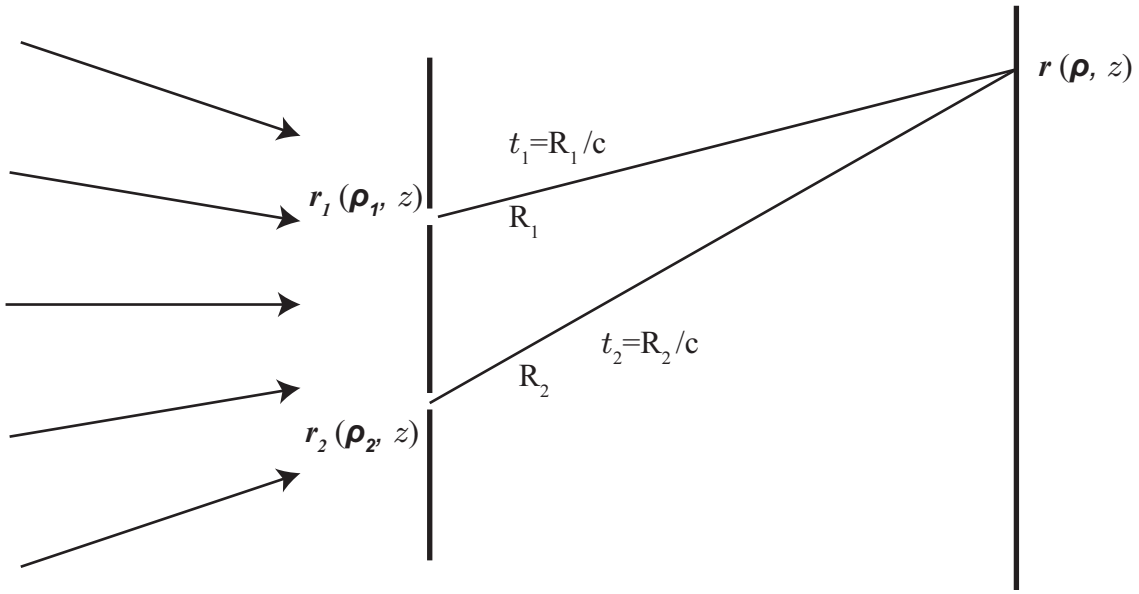


FIG. 5: Young's Interferometer.

In this section, we will look at the spatial coherence effects at a particular frequency ω. So, we represent a time-domain electric field $V(\boldsymbol{r},t)$ at position $\boldsymbol{r}$ at time t in terms of a Fourier integral as

$$V(\boldsymbol{r},t)=\int_{-\infty}^{\infty}\tilde{v}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega=\int_{-\infty}^{\infty}U(\boldsymbol{\rho};z)e^{-i\omega t}d\omega$$

Here $U(\boldsymbol{\rho};z)=\tilde{v}(\boldsymbol{r},\omega)$ is the field amplitude at frequency ω at the detector position. The field $U(\boldsymbol{\rho};z)=\tilde{v}(\boldsymbol{r},\omega)$ can in general be random and in such situations, the above Fourier transform is just symbolic and physical meanings will only be ascribed to correlations functions. We now ask the question as to how field vibrations at spatial location $(\boldsymbol{\rho_1},z=0)$ are correlated with field vibrations at spatial location $(\boldsymbol{\rho_2},z=0)$. This question seeking to quantify spatial coherence can be studied through the Young's double-slit interferometer, as shown in Fig. 5. In a Young's double-slit interferometer, one collects the fields emanating from spatial location $(\boldsymbol{\rho_1},z=0)$ and $(\boldsymbol{\rho_2},z=0)$ at the spatial location $(\boldsymbol{\rho},z)$ on a screen. The field $U(\boldsymbol{\rho};z)$ is written as

$$U(\boldsymbol{\rho};z)=k_1U(\boldsymbol{\rho}_1,0)e^{-i\omega t_1}+k_2U(\boldsymbol{\rho}_2,0)e^{-i\omega t_2} \tag{36}$$

Here $U(\boldsymbol{\rho}_1,0)e^{-i\omega t_1}$ and $U(\boldsymbol{\rho}_2,0)e^{-i\omega t_2}$ are the electric field vibrations at frequency ω at the source points $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$, at time t_1 and t_2, respectively. The constants k_1 and k_2 can in general be complex numbers. However, before we start calculating the expression for intensity at the detector, let us first go through some essential mathematical concepts.

Spatial correlations: some essential mathematical concepts

(1) **Stationarity (spatial):** A random process is said to be wide-sense spatially-stationary if the mean and the cross-spectral density function are independent of the space-origin, that is,

$$\langle U(\boldsymbol{\rho},z)\rangle=\langle U(0,z)\rangle,\implies\langle U^*(\boldsymbol{\rho},z)U(\boldsymbol{\rho},z)\rangle=\langle U^*(0,z)U(0,z)\rangle \tag{37}$$

and
$$\langle U^*(\boldsymbol{\rho_1},z)U(\boldsymbol{\rho_2},z)\rangle=\langle U^*(\boldsymbol{\rho_1}-\boldsymbol{\rho_1},z)U(\boldsymbol{\rho_2}-\boldsymbol{\rho_1},z)\rangle=\langle U^*(0,z)U(\boldsymbol{\Delta\rho},z)\rangle, \tag{38}$$

where $\boldsymbol{\Delta\rho}=\boldsymbol{\rho}_2-\boldsymbol{\rho}_1$. We find that if the process is spatially stationary in the wide sense, the correlation function depends only on the difference of the two spatial positions, just like in the case of a temporally stationary fields.

(2) **Angular-spectrum representation of a space-domain fields**

Just like a time-varying signal has a Fourier representation in the frequency domain, a space-varying signal has a Fourier representation in the spatial-frequency domain. In this section we will work out the Fourier representation of wave-fields going in the positive-z direction. Let us also assume that $U(\boldsymbol{\rho}, z)$ can be represented as a Fourier integral of the form:

$$U(\boldsymbol{\rho}, z) = \int_{-\infty}^{\infty} \tilde{U}(\boldsymbol{q}; z) e^{i\boldsymbol{q}\cdot\boldsymbol{\rho}} d^2\boldsymbol{q}$$
$$\text{or,} \qquad U(x, y; z) = \iint_{-\infty}^{\infty} \tilde{U}(q_x, q_y; z) e^{i(q_x x + q_y y)} dq_x dq_y$$

Here, we have used $\boldsymbol{\rho} = (x, y)$ and $\boldsymbol{q} = (q_x, q_y)$. We note that we are only taking a two-dimensional Fourier transform of the space-domain signal $U(\boldsymbol{\rho}, z)$. The variable z is kept at the propagation directly. Therefore, our formalism will be sufficient only for beam-like fields which propagate in the positive z direction. We know that $U(\boldsymbol{\rho}, z) \equiv \tilde{v}(\boldsymbol{r}, \omega)$ satisfies the Helmholtz equation:

$$(\nabla^2 + k^2) U(\boldsymbol{\rho}, z) = 0.$$

Therefore, putting the Fourier representation of $U(x, y; z)$ back in the above Helmholtz equation, we get

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2\right) \iint_{-\infty}^{\infty} \tilde{U}(q_x, q_y; z) e^{i(q_x x + q_x y)} dq_x dq_y = 0.$$
$$\text{or,} \quad \iint_{-\infty}^{\infty} \left[\nabla^2 \tilde{U}(q_x, q_y; z) + (-q_x^2 - q_y^2)\tilde{U}(q_x, q_y; z) + k^2 \tilde{U}(q_x, q_y; z)\right] e^{i(q_x x + q_y y)} dq_x dq_y = 0$$
$$\text{or,} \quad \iint_{-\infty}^{\infty} \left[\frac{\partial}{\partial z^2}\tilde{U}(q_x, q_y; z) + (k^2 - q_x^2 - q_y^2)\tilde{U}(q_x, q_y; z)\right] e^{i(q_x x + q_y y)} dq_x dq_y = 0$$
$$\text{or,} \quad \iint_{-\infty}^{\infty} \left[\frac{\partial}{\partial z^2}\tilde{U}(q_x, q_y; z) + w^2 \tilde{U}(q_x, q_y; z)\right] e^{i(q_x x + q_y y)} dq_x dq_y = 0$$
$$\text{or,} \quad \frac{\partial}{\partial z^2}\tilde{U}(q_x, q_y; z) + w^2 \tilde{U}(q_x, q_y; z) = 0,$$

where $w^2 = k^2 - q_x^2 - q_y^2$. We note that

$$w = +(k^2 - q_x^2 - q_y^2)^{\frac{1}{2}} \quad \text{when} \quad q_x^2 + q_y^2 \leq k^2$$
$$= +i(q_x^2 + q_y^2 - k^2)^{\frac{1}{2}} \quad \text{when} \quad q_x^2 + q_y^2 > k^2$$

When w is real $U(q_x, q_y; z)$ represents a propagating wave, which goes up to ∞ but when w is imaginary it represents an evanescent wave, whose amplitude decays exponentially as a function of z. We are here interested only in the propagating solutions. So, we write the general solution of the above equation as

$$\tilde{U}(q_x, q_y; z) = a(q_x, q_y) e^{iwz} + b(q_x, q_y) e^{-iwz}$$

Therefore, we obtain for $U(x, y; z)$

$$U(x, y; z) = \iint_{-\infty}^{\infty} \left[a(q_x, q_y) e^{iwz} + b(q_x, q_y) e^{-iwz}\right] e^{i(q_x x + q_y y)} dq_x dq_y$$
$$\text{or,} \quad U(x, y; z) = \iint_{-\infty}^{\infty} \left[a(q_x, q_y) e^{i(q_x x + q_y y + wz)} + b(q_x, q_y) e^{i(q_x x + q_y y - wz)}\right] dq_x dq_y$$

We know that $e^{i(wz - \omega t)}$ represents a wave going in the $+z$ direction whereas $e^{i(-wz - \omega t)}$ represents a wave going in the $-z$ direction. We are only interested in the waves going in the $+z$ direction. So, we take

$$U(x, y; z) = \iint_{-\infty}^{\infty} a(q_x, q_y) e^{i(q_x x + q_y y + wz)} dq_x dq_y. \tag{39}$$

We note that the above representation is not a Fourier transform relationship. However, when we write the field at $z = 0$, we get

$$U(x, y; 0) = \iint_{-\infty}^{\infty} a(q_x, q_y) e^{i(q_x x + q_y y)} dq_x dq_y,$$

which indeed is a Fourier transform relationship. We can invert the above Fourier transform integral to get

$$a(q_x, q_y) = \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} U(x, y; 0) e^{-i(q_x x + q_y y)} dx dy.$$

The amplitude $a(q_x, q_y)$ is known as the angular-spectrum representation of wave-field going in the positive-z direction. The angular spectrum does not depend on z and thus it does not change upon propagation. So, we see that knowing the angular spectrum $a(q_x, q_y)$, one can calculate the filed $U(x, y; z)$ at any z. Also, the field $U(x, y; z)$ at z can be obtained in terms of the field at $z = 0$ by substituting for $a(q_x, q_y)$ in Eq. (39) as

$$U(x, y; z) = \iint_{-\infty}^{\infty} \left[\frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} U(x', y'; 0) e^{-i(q_x x' + q_y y')} dx' dy' \right] e^{i(q_x x + q_y y + wz)} dq_x dq_y,$$

$$\text{or,} \quad U(x, y; z) = \frac{1}{(2\pi)^2} \iiint\!\!\iint_{-\infty}^{\infty} U(x', y'; 0) e^{i[q_x(x - x') + q_y(y - y') + wz]} dx' dy' dq_x dq_y,$$

where $U(x', y'; 0)$ is the filed at $z = 0$. We now have a representation for $\tilde{v}(\boldsymbol{r}, \omega) \equiv U(\boldsymbol{\rho}, z) = U(x, y; z)$ in its conjugate basis:

$$U(x, y; z) = \iint_{-\infty}^{\infty} a(q_x, q_y) e^{i(q_x x + q_y y + wz)} dq_x dq_y.$$

In the above equation, we have used the fact that $\boldsymbol{k} = (q_x, q_v, w)$. Let us rewrite it as $\boldsymbol{k} = (\boldsymbol{q}, w)$, where $\boldsymbol{q} = (u, v)$. Let us also assume that $q_x^2 + q_y^2 \ll k^2$ and thus write $w = \sqrt{k^2 - q_x^2 - q_y^2} \approx k - \frac{q^2}{2k}$, where $k = |\boldsymbol{k}|$ and $q = |\boldsymbol{q}| = \sqrt{q_x^2 + q_y^2}$. This is known as the paraxial approximation and is applicable for the field that propagate mostly towards the positive-z direction. We can now write the above equation as

$$U(\boldsymbol{\rho}; z) = e^{ikz} \iint_{-\infty}^{\infty} a(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{\rho}} e^{-i\frac{q^2}{2k}z} d^2\boldsymbol{q}.$$

Angular correlation function

Now, if $U(\boldsymbol{\rho}; z)$ is random the above decomposition is true only in a symbolic senses. In this case, we need to look at the cross-correlations of the type $\langle U^*(\boldsymbol{\rho}_1; z) U(\boldsymbol{\rho}_2; z) \rangle$, which quantifies how field vibrations with frequency ω at positions $(\boldsymbol{\rho}_1, z)$ and $(\boldsymbol{\rho}_2, z)$ are correlated. We have

$$\langle U^*(\boldsymbol{\rho}_1; z) U(\boldsymbol{\rho}_2; z) \rangle = \iint_{-\infty}^{\infty} \langle a^*(\boldsymbol{q_1}) a(\boldsymbol{q_2}) \rangle e^{-i\boldsymbol{q_1}\cdot\boldsymbol{\rho}_1 + i\boldsymbol{q_2}\cdot\boldsymbol{\rho}_2} e^{i\frac{q_1^2 - q_2^2}{2k}z} d^2\boldsymbol{q_1} d^2\boldsymbol{q_2}$$

We recognize $\langle U^*(\boldsymbol{\rho}_1; z) U(\boldsymbol{\rho}_2; z) \rangle = W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z)$ as the cross-spectral density function. The ensemble average inside the integral $\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) \equiv \langle a^*(\boldsymbol{q_1}) a(\boldsymbol{q_2}) \rangle$ is called the angular correlation function of the field at frequency ω. It quantifies how the field vibrations with frequency ω at spatial frequencies $(\boldsymbol{q_1})$ and $(\boldsymbol{q_2})$ are correlated. We write the above equation as

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z) = \iint_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) e^{-i\boldsymbol{q_1}\cdot\boldsymbol{\rho}_1 + i\boldsymbol{q_2}\cdot\boldsymbol{\rho}_2} e^{i\frac{q_1^2 - q_2^2}{2k}z} d^2\boldsymbol{q_1} d^2\boldsymbol{q_2}.$$

Again, we note that this is not a Fourier transform relation. However, if we consider the cross-spectral density function at $z = 0$, we find that the cross-spectral density function and the angular correlation function form a Fourier transform pair.

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z = 0) = W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \iint_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) e^{-i(\boldsymbol{q_1}\cdot\boldsymbol{\rho}_1 - \boldsymbol{q_2}\cdot\boldsymbol{\rho}_2)} d^2\boldsymbol{q_1} d^2\boldsymbol{q_2}. \tag{40}$$

This can be seen as the spatial analog of the generalized Wiener-Khintchine theorem. Just as the frequency correlations of a source is characterized by the cross-spectral density function, the spatial-frequency correlations of a source is characterized by the angular correlation function. The cross-spectral density function and the angular correlation functions can be in the following normalized form.

The degree of the angular correlation function can be defined as

$$\alpha(\boldsymbol{q_1}, \boldsymbol{q_2}) = \frac{\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2})}{\sqrt{\mathcal{S}(\boldsymbol{q_1})\mathcal{S}(\boldsymbol{q_2})}} \tag{41}$$

$$\mu(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \frac{W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)}{\sqrt{S(\boldsymbol{\rho}_1)S(\boldsymbol{\rho}_2)}}, \tag{42}$$

The magnitudes $|\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2})|$ and $|\mu(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)|$ are called the degree of angular coherence and the degree of spectral coherence, respectively. Both these degrees range from 0 to 1, that is, $0 \leq |\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2})| \leq 1$ and $0 \leq |\mu(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)| \leq 1$.

Now, for the spatially stationary random processes, the cross-spectral density function $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$ at $z = 0$ depends only on $\Delta\boldsymbol{\rho}$. Therefore, in order that the right hand side of Eq. (40) also depends on $\boldsymbol{\Delta}\rho$ only, we have to have

$$\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) = \mathcal{S}(\boldsymbol{q_1})\delta(\boldsymbol{q_1} - \boldsymbol{q_2}) \tag{43}$$

So, in this case Eq. (40) can be written as

$$W(\boldsymbol{\Delta}\rho) = \int_{-\infty}^{\infty} \mathcal{S}(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{\Delta}\rho} d^2\boldsymbol{q}. \tag{44}$$

The above equation is the spatial analog of the one-dimensional Wiener-Khintchine theorem.

(3) **Space Average:** Just as we defined time-averaging, one can define space averaging for a random signal. There are only a few interferometers in which space-averaging is involved. Most of the interferometers can make measurements at a particular location with good enough accuracy unlike time measurements.

(4) **Ergodicity (space):** It is now natural to expect an ergodic theorem connecting space averaging with ensemble averaging for spatially stationary fields, just like in the case of time-domain signal. However, since most current interferometers do not employ space-averaging, the concept of spatial ergodic theorem has not been explicitly studied as much.

LECTURE # 7

Quantifying the spatial correlations

We now go back to the field in Eq. (36). Using this field, the intensity at $(\boldsymbol{\rho}, z)$ on the screen can be written as

$$\langle U^*(\boldsymbol{\rho};z)U(\boldsymbol{\rho};z)\rangle = |k_1|^2\langle U^*(\boldsymbol{\rho}_1)U(\boldsymbol{\rho}_1)\rangle + |k_2|^2\langle U^*(\boldsymbol{\rho}_2)U(\boldsymbol{\rho}_2)\rangle + 2\mathrm{Re}k_1^*k_2\langle U^*(\boldsymbol{\rho}_1)U(\boldsymbol{\rho}_2)\rangle e^{-i\omega(t_2-t_1)}$$

In terms of the cross-spectral density it becomes:

$$W(\boldsymbol{\rho},\boldsymbol{\rho};z) = |k_1|^2W(\boldsymbol{\rho}_1,\boldsymbol{\rho}_1) + |k_2|^2W(\boldsymbol{\rho}_2,\boldsymbol{\rho}_2) + 2\mathrm{Re}k_1^*k_2W(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)e^{-i\omega(t_2-t_1)}$$

This can now be written as

$$S(\boldsymbol{\rho};z) = |k_1|^2S(\boldsymbol{\rho}_1) + |k_2|^2S(\boldsymbol{\rho}_2) + 2|k_1||k_2|\sqrt{S(\boldsymbol{\rho}_1)S(\boldsymbol{\rho}_2)}|\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|\cos[\omega(t_2-t_1)+\arg\mu+\phi_k],$$

where $k_1^*k_2 = |k_1||k_2|e^{-i\phi_k}$ and $\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = |\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|e^{-i\arg\mu}$. This equation represents the general spatial-interference law. We see that the intensity at the detector depends on $\omega(t_2-t_1)+\arg\mu$ and it changes in a sinusoidal manner as a function of $\omega(t_2-t_1)\arg\mu$. This sinusoidal dependence is multiplied by the degree of coherence function $|\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|$. The visibility of the interference pattern is defined in the following manner by

$$V = \frac{S(\boldsymbol{\rho};z)_{\max} - S(\boldsymbol{\rho};z)_{\min}}{S(\boldsymbol{\rho};z)_{\max} + S(\boldsymbol{\rho};z)_{\min}} = \frac{2|k_1||k_2|\sqrt{S(\boldsymbol{\rho}_1)S(\boldsymbol{\rho}_2)}}{|k_1|^2S(\boldsymbol{\rho}_1)+|k_2|^2S(\boldsymbol{\rho}_2)}|\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|$$

We see that the visibility is proportional to the degree of temporal coherence. And for the special case when $|k_1|^2S(\boldsymbol{\rho}_1) = |k_2|^2S(\boldsymbol{\rho}_2)$, we have $V = |\mu(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)|$, that is, the fringe visibility equals the degree of coherence.

The spectral coherence function quantifies the degree of spatial coherence of the source. This function depends on the spatial-frequency correlations of the source and can be calculated using the angular correlation function of the source. The width of the spectral coherence function is called the transverse coherence length. Let us consider the following types of angular correlation functions.

1. **Case-I: No correlation between different spatial-frequency components**

   The angular correlation function in this takes the following form.

   $$\begin{aligned}\mathcal{A}(\boldsymbol{q_1},\boldsymbol{q_2}) = \langle a^*(\boldsymbol{q_1})a(\boldsymbol{q_2})\rangle &= \mathcal{A}(\boldsymbol{q_1},\boldsymbol{q_1})\delta(\boldsymbol{q_1}-\boldsymbol{q_2})\\ &= \mathcal{S}(\boldsymbol{q_1})\delta(\boldsymbol{q_1}-\boldsymbol{q_2})\end{aligned}$$

   The degree of angular correlation in this case is $|\alpha(\boldsymbol{q_1},\boldsymbol{q_2})| \to \delta(\boldsymbol{q_1}-\boldsymbol{q_2})$. As evaluated in the last section, the cross-spectral density function can be written as

   $$W(\Delta\boldsymbol{\rho}) = \int_{-\infty}^{\infty}\mathcal{S}(\boldsymbol{q_1})e^{i\boldsymbol{q_1}\cdot\Delta\boldsymbol{\rho}}d^2\boldsymbol{q_1}.$$

   The spectral-density in this case is given as

   $$S(\boldsymbol{\rho}) = W(\boldsymbol{\rho},\boldsymbol{\rho}) = \int_{-\infty}^{\infty}\mathcal{S}(\boldsymbol{q_1})d^2\boldsymbol{q_1} = \text{constant}.$$

   We note that, analogous to the temporal case, when there is no correlation between different spatial-frequencies, the correlation function depends only on the difference of the spatial coordinates and the spectral density is independent of the spatial location. As discussed before, this field is also called the spatially stationary field. We further note that in this case the cross-spectral density function does not depend on z.

2. **Case-II: Perfect correlation between different spatial-frequency components**

   The other extreme situation is when the source has perfect correlations between different spatial-frequency components. The angular correlation function in this case can be factorized as

   $$\mathcal{A}(\boldsymbol{q_1},\boldsymbol{q_2}) = a^*(\boldsymbol{q_1})a(\boldsymbol{q_2})$$

The degree of angular correlation in this case is $|\alpha(\boldsymbol{q_1}, \boldsymbol{q_2})| = 1$. In this case, the angular spectrum representation is no longer symbolic. The transformations exist as they are written

$$U(\boldsymbol{\rho}; z) = e^{ikz} \iint_{-\infty}^{\infty} a(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{\rho}} e^{-i\frac{q^2}{2k}z} d^2\boldsymbol{q}.$$

The cross-spectral density $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z) = \langle U^*(\boldsymbol{\rho}_1; z) U(\boldsymbol{\rho}_2; z)\rangle$ and the spectral density both depend on $\boldsymbol{\rho}_1$ as well as on $\boldsymbol{\rho}_2$. As a result, the spectral density has a spatial profile and the cross-spectral density does not die-off for any value of $\Delta\boldsymbol{\rho} = \boldsymbol{\rho}_2 - \boldsymbol{\rho}_1$.

3. **Case-III: Partial correlation between different spatial-frequency components**

The above two cases are the two extreme cases of the most general angular correlation, in which different spatial-frequencies are only partially correlated. One model that captures such partial correlations is the Gaussian-Schell model and the angular correlation function takes the following form in this model:

$$\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}; \omega) = A \exp\left[-\frac{(q_1^2 + q_2^2)}{2\delta^2}\right] \exp\left[-\frac{(\boldsymbol{q_1} - \boldsymbol{q_2})^2}{2\delta_c^2}\right]$$

We note that in the limit $\delta_c \to 0$, the second Gaussian function becomes a Dirac-delta function and thus represents the case when field vibrations at different spatial frequency components are completely uncorrelated. The other extreme situation is depicted by the limit $\delta_c \to \infty$, in which case the second Gaussian becomes unity and the angular correlation function factorizes.

Propagation of Correlation

The time evolution of the electric and magnetic fields is governed by the wave equations [Eq. (1)]. What is the equation that governs the propagation of correlation functions? To answer this question, we start with the wave-equation for a real field $V^{(\mathrm{r})}(\boldsymbol{r}, t)$. We have

$$\nabla^2 V^{(\mathrm{r})}(\boldsymbol{r}, t) = \frac{1}{c^2} \frac{\partial^2 V^{(\mathrm{r})}(\boldsymbol{r}, t)}{\partial t^2}$$

It can be shown that the complex analytical signal $V(\boldsymbol{r}, t)$ associated with $V^{(\mathrm{r})}(\boldsymbol{r}, t)$ also follows the wave equation, that is,

$$\nabla^2 V(\boldsymbol{r}, t) = \frac{1}{c^2} \frac{\partial^2 V(\boldsymbol{r}, t)}{\partial t^2}.$$

Next, we take the complex conjugate of the above equation and replace $\boldsymbol{r}$ by $\boldsymbol{r_1}$ and t by t_1, we then get

$$\nabla_1^2 V^*(\boldsymbol{r_1}, t_1) = \frac{1}{c^2} \frac{\partial^2 V^*(\boldsymbol{r_1}, t_1)}{\partial t_1^2}.$$

Multiplying both sides of this equation with $V(\boldsymbol{r_2}, t_2)$ and taking the ensemble average, we obtain

$$\nabla_1^2 \langle V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2)\rangle = \frac{1}{c^2} \frac{\partial^2 \langle V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2)\rangle}{\partial t_1^2}.$$

Recognizing the ensemble average to be the cross-correlation function, we write the above equation as

$$\nabla_1^2 \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) = \frac{1}{c^2} \frac{\partial^2 \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2)}{\partial t_1^2}.$$

In a similar way one can obtain the equation

$$\nabla_2^2 \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2) = \frac{1}{c^2} \frac{\partial^2 \Gamma(\boldsymbol{r_1}, t_1; \boldsymbol{r_2}, t_2)}{\partial t_2^2}.$$

The set of two above equations are known as Wolf equations and they govern the propagation of the cross-correlation function. If the process is stationary and ergodic, the above two equations take the following form

$$\nabla_1^2\Gamma(\boldsymbol{r_1},\boldsymbol{r_2},\tau) = \frac{1}{c^2}\frac{\partial^2\Gamma(\boldsymbol{r_1},\boldsymbol{r_2},\tau)}{\partial\tau^2},$$
$$\nabla_2^2\Gamma(\boldsymbol{r_1},\boldsymbol{r_2},\tau) = \frac{1}{c^2}\frac{\partial^2\Gamma(\boldsymbol{r_1},\boldsymbol{r_2},\tau)}{\partial\tau^2}.$$

Also, the cross-correlation function is the Fourier transform of the cross-spectral density function. Therefore, taking the Fourier transform of the equations on both sides, we obtain

$$\nabla_1^2 W(\boldsymbol{r_1},\boldsymbol{r_2},\omega) + k^2 W(\boldsymbol{r_1},\boldsymbol{r_2},\omega) = 0,$$
$$\nabla_2^2 W(\boldsymbol{r_1},\boldsymbol{r_2},\omega) + k^2 W(\boldsymbol{r_1},\boldsymbol{r_2},\omega) = 0,$$

where $k = \omega/c$. The above equation are the propagation equation for the cross-spectral density function.

Propagation of correlation and the van Cittert-Zernike theorem

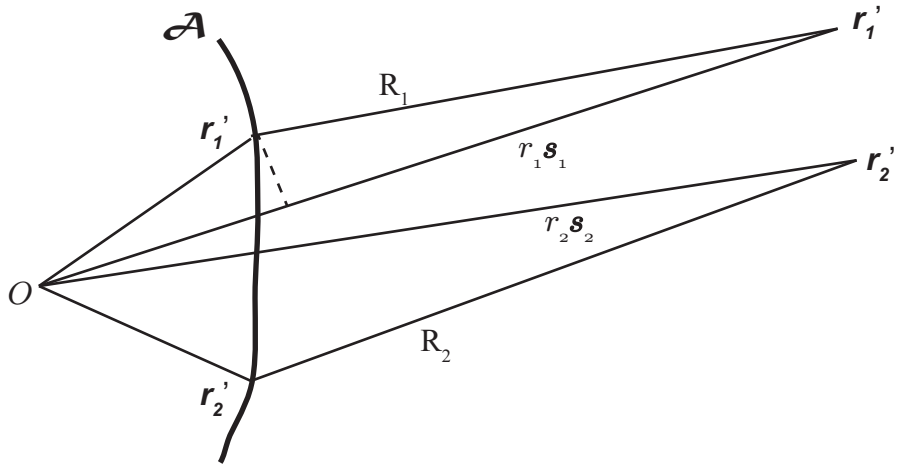


FIG. 6: propagation of coherence.

So far we have studied the coherence properties in one basis is decided by the field correlations in the conjugate basis. For, example the temporal coherence properties are decided by how different field vibrations are correlated in their frequencies. This is basically the statement of the generalized Wiener-Khintchine theorem.

$$\Gamma(\boldsymbol{r_1},\boldsymbol{r_2};t_1,t_2) = \iint_{-\infty}^{\infty} W(\boldsymbol{r_1},\boldsymbol{r_2};\omega_1,\omega_2)e^{i(\omega_1 t_1-\omega_2 t_2)}d\omega_1 d\omega_2$$

Similarly, the coherence properties between two spatial positions are decided by how the field vibrations are correlated in their spatial-frequencies, and correspondingly we get analogous forms for the Wiener Khintchine theorem.

$$W(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,\omega) = \iint_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{q_1},\boldsymbol{q_2};\omega)e^{-i(\boldsymbol{q_1}\cdot\boldsymbol{\rho}_1-\boldsymbol{q_2}\cdot\boldsymbol{\rho}_2)}d^2\boldsymbol{q_1}d^2\boldsymbol{q_2}.$$

In the previous section, we derived how propagation of coherence can be studied through Wolf equations, which are in the form of differential equations. In this section, we will be studying propagation of coherence through an integral equation for a primary source. We will be looking at how the spatial coherence properties of the field emanating from a primary source propagates to far field. We will start with the general formulation of the problem and then later focus onto the far-field form of the problem. Let us start with the field $\tilde{v}(\boldsymbol{r},\omega)$ that satisfies the Helmholtz equation, that is,

$$(\nabla^2 + k^2)\tilde{v}(\boldsymbol{r},\omega) = 0.$$

Now, let us consider a unit-amplitude diverging spherical wave-field emanating from point $\boldsymbol{r_1'}$. The resulting field at position $\boldsymbol{r_1}$ due to this diverging wave can be given by

$$\frac{e^{ikR_1}}{R_1}$$

where $R_1 = |\boldsymbol{r_1} - \boldsymbol{r_1'}|$. In fact, this wave satisfies the Helmholtz equation and so can be used as a basis for representing the propagation of fields. Now, if the amplitude of the field at $\boldsymbol{r_1'}$ is $\tilde{v}(\boldsymbol{r_1'}, \omega)$ then the resulting field at $\boldsymbol{r_1}$ is given by

$$\tilde{v}(\boldsymbol{r_1'}, \omega)\frac{e^{ikR_1}}{R_1}.$$

Let us now consider an infinitesimal area element $d^2\boldsymbol{r_1'}$ centered at $\boldsymbol{r_1'}$. The resulting field due to the area element thus becomes

$$\tilde{v}(\boldsymbol{r_1'}, \omega)\frac{e^{ikR_1}}{R_1}\Lambda_1(k)d^2\boldsymbol{r_1'},$$

where $\Lambda_1(k)$ depends on the orientation of the area element. The total field $\tilde{v}(\boldsymbol{r_1}, \omega)$ at $\boldsymbol{r_1}$ due to the source is calculated by integrating over the entire source area, that is,

$$\tilde{v}(\boldsymbol{r_1}, \omega) = \int_{\mathcal{A}} \tilde{v}(\boldsymbol{r_1'}, \omega)\frac{e^{ikR_1}}{R_1}\Lambda_1(k)d^2\boldsymbol{r_1'}.$$

Similarly, the total field $\tilde{v}(\boldsymbol{r_2}, \omega)$ at some other position $\boldsymbol{r_2}$ due to the same source can be written as

$$\tilde{v}(\boldsymbol{r_2}, \omega) = \int_{\mathcal{A}} \tilde{v}(\boldsymbol{r_2'}, \omega)\frac{e^{ikR_2}}{R_2}\Lambda_2(k)d^2\boldsymbol{r_2'},$$

where, as before, $R_2 = |\boldsymbol{r_2} - \boldsymbol{r_2'}|$, and $\Lambda_2(k)$ is the inclination factor corresponding to the area element $d^2\boldsymbol{r_2'}$. Using the field representations above, we obtain the following expression for the field correlation $\langle \tilde{v}^*(\boldsymbol{r_1}, \omega)\tilde{v}(\boldsymbol{r_2}, \omega)\rangle$:

$$\langle \tilde{v}^*(\boldsymbol{r_1}, \omega)\tilde{v}(\boldsymbol{r_2}, \omega)\rangle = \int_{\mathcal{A}}\int_{\mathcal{A}} \langle \tilde{v}^*(\boldsymbol{r_1'}, \omega)\tilde{v}(\boldsymbol{r_2'}, \omega)\rangle \frac{e^{ik(R_2-R_1)}}{R_1R_2}\Lambda_1^*(k)\Lambda_2(k)d^2\boldsymbol{r_1'}d^2\boldsymbol{r_2'}.$$

The quantities in the angular brackets are the cross-spectral densities function. So, we can rewrite the above equation as

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \int_{\mathcal{A}}\int_{\mathcal{A}} W(\boldsymbol{r_1'}, \boldsymbol{r_2'}, \omega)\frac{e^{ik(R_2-R_1)}}{R_1R_2}\Lambda_1^*(k)\Lambda_2(k)d^2\boldsymbol{r_1'}d^2\boldsymbol{r_2'}. \tag{45}$$

The above is the propagation equation for spatial coherence, which expresses the spatial correlations of the propagated field in terms of the spatial correlations at the source. It can be rigourously shown that $\Lambda_1(k) = \Lambda_2(k) = \frac{ik}{2\pi}$ but we will just use this result and will not derive it here. If the field is quasi-monochromatic, one can take the Fourier transform of the above equation and obtain the propagation equation for the cross-correlation function.

$$\Gamma(\boldsymbol{r_1}, \boldsymbol{r_2}, \tau) = \int_{\mathcal{A}}\int_{\mathcal{A}} \Gamma(\boldsymbol{r_1'}, \boldsymbol{r_2'}, \tau)\frac{e^{ik_0(R_2-R_1)}}{R_1R_2}\Lambda_1^*(k_0)\Lambda_2(k_0)d^2\boldsymbol{r_1'}d^2\boldsymbol{r_2'}. \tag{46}$$

We note that $k_0 = \omega_0/c$ is the mean wave-vector of the quasi-monochromatic field. We will next study this propagation equation for a spatially completely incoherent source.

Let us study the propagation of cross-spectral density in some details and consider the situation when the source is spatially completely incoherent, that is, $W(\boldsymbol{r_1'}, \boldsymbol{r_2'}, \omega) = S(\boldsymbol{r_1'}, \omega)\delta(\boldsymbol{r_2'} - \boldsymbol{r_1'})$. The cross spectral density function at at $(\boldsymbol{r_1}, \boldsymbol{r_2})$ can be calculated using Eq. (45):

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi}\right)^2 \int_{\mathcal{A}}\int_{\mathcal{A}} S(\boldsymbol{r_1'}, \omega)\delta(\boldsymbol{r_2'} - \boldsymbol{r_1'})\frac{e^{ik(R_2-R_1)}}{R_1R_2}d^2\boldsymbol{r_1'}d^2\boldsymbol{r_2'},$$

$$\text{or,}\quad W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi}\right)^2 \int_{\mathcal{A}} S(\boldsymbol{r_1'}, \omega)\frac{e^{ik(R_2-R_1)}}{R_1R_2}d^2\boldsymbol{r_1'},$$

The last equation is called the van Cittert-Zernike theorem. This shows that the spectral density of the field at the source decides the cross-spectral density of the field at points away from the source. We now look at the far-zone behaviour of the van Cittert-Zernike theorem. In the far-zone, we can take $R_1 = r_1 - |r_1'|\cos\phi = r_1 - \boldsymbol{r_1'}.\boldsymbol{s_1}$. Similarly, we have $R_2 = r_2 - \boldsymbol{r_1'}.\boldsymbol{s_2}$ in the far zone. Also, we can safely assume that the factor $\frac{1}{R_1 R_2}$ does not vary much in the far-zone for the integration-range of interest and so we can take it out of the integral. The far-zone form of the above equation is then

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi}\right)^2 \frac{1}{R_1 R_2} \int_{\mathcal{A}} S(\boldsymbol{r_1'}, \omega) e^{ik(r_2 - r_1)} e^{-ik(\boldsymbol{s_2} - \boldsymbol{s_1}).\boldsymbol{r_1'}} d^2\boldsymbol{r_1'}, \tag{47}$$

$$\text{or,} \quad W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi}\right)^2 \frac{e^{ik(r_2 - r_1)}}{R_1 R_2} \int_{\mathcal{A}} S(\boldsymbol{r_1'}, \omega) e^{-ik(\boldsymbol{s_2} - \boldsymbol{s_1}).\boldsymbol{r_1'}} d^2\boldsymbol{r_1'} \tag{48}$$

We find that even an incoherent source produces finite spatial coherence upon propagation. We also note that in the far-zone the cross-spectral density of the field is the Fourier transform of the spectral density of the field at the source.

Now, let us look at the van Cittert-Zernike theorem for the case of a circular, spatially-incoherent source of radius a. We take the source to be uniformly illuminated, that is, $S(\boldsymbol{r_1'}, \omega) = I_0$, where I_0 is a constant. We also assume that the observation points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are far enough away from the source so that $r_1 \approx r_2 \approx z$. We now use the following transformations for evaluating the intergral

$$\begin{aligned}
&\boldsymbol{s_1} = \left(\frac{x_1}{r_1}, \frac{y_1}{r_1}, \frac{z}{r_1}\right) \approx \left(\frac{x_1}{z}, \frac{y_1}{z}, 1\right) \\
&\boldsymbol{s_2} = \left(\frac{x_2}{r_2}, \frac{y_2}{r_2}, \frac{z}{r_2}\right) \approx \left(\frac{x_2}{z}, \frac{y_2}{z}, 1\right) \\
&\boldsymbol{s_2} - \boldsymbol{s_1} = (w\cos\psi, w\sin\psi, 0); \qquad \boldsymbol{r'} = (\rho'\cos\theta', \rho'\sin\theta', 0) \\
&(\boldsymbol{s_2} - \boldsymbol{s_1}) \cdot \boldsymbol{r'} = \rho' w\cos(\theta' - \psi) = \left[\frac{(x_2 - x_1)x'}{z} + \frac{(y_2 - y_1)y'}{z}\right]
\end{aligned}$$

Using the above transformation, Eq. (48) for the source of interest becomes:

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi z}\right)^2 I_0 \int_0^{2\pi} \int_0^a e^{-ik\rho' w\cos(\theta' - \psi)} \rho' d\rho' d\theta' \tag{49}$$

The integral on the right hand side is a standard integral, using which we get

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega) = \left(\frac{k}{2\pi z}\right)^2 I_0 \pi a^2 \frac{J_1(v)}{v}. \tag{50}$$

Here $J_1(v)$ is the Bessel function of the first kind and first order, and

$$v = \frac{ka}{z}\left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right]^{1/2} = \frac{kad}{z}, \tag{51}$$

where d is the separation between the two observation points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. The first zero of $J_1(v)$ occurs at $v = 3.83$, this corresponds to

$$d = \frac{3.83z}{ka} = \frac{0.61\lambda z}{a}. \tag{52}$$

This distance can therefore be taken to be a measure of the transverse coherence length of the source in the far-field. Correspondingly, the coherence area is

$$\Delta A = \pi\left(\frac{d}{2}\right)^2 = \pi\left(\frac{0.61\lambda z}{2a}\right)^2 = \pi^2\left(\frac{0.61\lambda z}{2}\right)^2 \frac{1}{S}, \tag{53}$$

where $S = \pi a^2$ is the source area. We find that the far-zone coherence area of an incoherent source is inversely proportional to the area of the source. Smaller the source size, larger will be the coherence area in the far field. If the source is a point source, then effective coherence are would be infinity.

LECTURE # 8

Second order coherence theory (Angular)

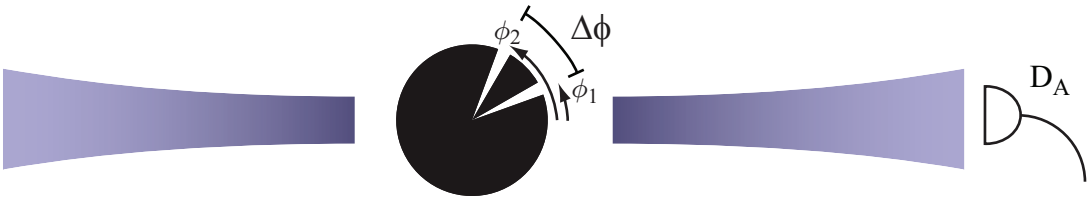


FIG. 7: Schematic representation of angular one-photon interference. A partially correlated field falls onto an angular aperture in the form of a double angular-slit. The angular separation between the slits is $\Delta\phi$.

In this section, we study angular coherence by asking the question as to how field vibrations at an angular position ϕ_1 are correlated with field vibrations at angular position ϕ_2. This question seeking to quantify angular coherence can be studied through the angular double-slit interferometer, as shown in Fig. 7. Just like the temporal and the spatial case, one can in general have situations in which different OAM modes are only partially correlated. The field $V(\phi)$ on the screen at angular position ϕ is given by

$$V(\phi) = k_1 V(\phi_1) + k_2 V(\phi_2), \tag{54}$$

where $V(\phi_1)$, and $V(\phi_2)$ are the fields at ϕ_1 and ϕ_2 respectively. The constants k_1 and k_2 can in general be complex numbers. Just like time and space domains, we can quantify the correlations in the angular domain by studying the intensity at (ϕ, t). However, before we start calculating the expression for intensity at the detector, let us first go through some essential mathematical concepts.

Angular correlations: some essential mathematical concepts

(1) **Stationarity (angular)** A random process can be called wide-sense agular-stationary if the mean and the two-point correlations functions are independent of the origin of angular positions, that is,

$$\begin{aligned} &\langle V(\phi,t)\rangle = \langle V(0,t)\rangle \implies \langle V^*(\phi)V(\phi)\rangle = \langle V^*(0)V(0)\rangle \\ &\langle V^*(\phi_1)V(\phi_2)\rangle = \langle V^*(\phi_1-\phi_1,t)V(\phi_2-\phi_1,t)\rangle = \langle V^*(0)V(\Delta\phi)\rangle, \end{aligned} \tag{55}$$

where $\Delta\phi = \phi_2 - \phi_1$. So, we see that if the correlation function is angular stationary in the wide sense, the correlation function depends only on the difference of the two angular positions, just like in the case of temporally and spatially stationary fields.

(2) **Representation of angular-domain signal in the conjugate basis: the azimuthal correlation function**

We have seen that a time-domain field has a Fourier representation in the frequency-domain. Correspondingly, the cross-correlation function and the cross-spectral density function form a Fourier pair.

$$\begin{aligned} V(t) &= \int_{-\infty}^{\infty} \tilde{v}(\omega) e^{-i\omega t} d\omega; \qquad \tilde{v}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} V(t) e^{i\omega t} dt \\ \langle V^*(\boldsymbol{r_1}, t_1) V(\boldsymbol{r_2}, t_2)\rangle &= \Gamma(\boldsymbol{r_1}, \boldsymbol{r_2}; t_1, t_2) = \iint_{-\infty}^{\infty} W(\boldsymbol{r_1}, \boldsymbol{r_2}; \omega_1, \omega_2) e^{i(\omega_1 t_1 - \omega_2 t_2)} d\omega_1 d\omega_2 \end{aligned}$$

Similarly, field-representations in the position-space and the momentum space form a Fourier pair, which is written as

$$V(\boldsymbol{r}) = \int_{-\infty}^{\infty} \tilde{v}(\boldsymbol{k}) e^{i\boldsymbol{k}.\boldsymbol{r}} d\boldsymbol{k}; \qquad \tilde{v}(\boldsymbol{k}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} V(\boldsymbol{r}) e^{-i\boldsymbol{k}.\boldsymbol{r}} d\boldsymbol{r}$$

However, in the context of this course, we are mostly interested in fields moving towards the positive-z direction and therefore in the Fourier relationship between transverse-position and transverse-momentum. We already know that the corresponding correlation functions, namely, the cross-spectral density function and the angular-correlation function also form a Fourier transform pair.

$$U(\boldsymbol{\rho}) = \iint_{-\infty}^{\infty} a(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{\rho}} d^2\boldsymbol{q}; \qquad a(\boldsymbol{q}) = \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} U(\boldsymbol{\rho}) e^{-i\boldsymbol{q}\cdot\boldsymbol{\rho}} d^2\boldsymbol{\rho}.$$

$$\langle U^*(\boldsymbol{\rho_1}, \omega) U(\boldsymbol{\rho_2}, \omega)\rangle = W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \omega) = \frac{1}{k^4} \iint_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}; \omega) e^{-i(\boldsymbol{q_1}\cdot\boldsymbol{\rho}_1 - \boldsymbol{q_2}\cdot\boldsymbol{\rho}_2)} d^2\boldsymbol{q_1} d^2\boldsymbol{q_2}.$$

Just like time and frequency, position and momentum, it is now very well established that field representations in angular-position and orbital angular momentum bases also form a Fourier pair. We can see this in the following way.

This Fourier relationship between angular-position and orbital angular momentum (OAM) is written as

$$V(\phi) = \sum_{l=-\infty}^{\infty} \alpha_l e^{il\phi}$$

$$\alpha_l = \frac{1}{2\pi} \int_{-\pi}^{\pi} V(\phi) e^{-il\phi} d\phi.$$

Here $V(\phi)$ is the analytic signal corresponding to the field amplitude at angular position ϕ, and α_l is the field amplitude in the Fourier basis represented by the orbital angular momentum (OAM) mode index l. The existence of a Fourier relationship between the angular position and OAM gives rise to interesting interference effects in the distribution of OAM-mode (or topological-charge) of a field when it passes through an angular aperture [9, 10].

Now, if $V(\phi)$ represents the amplitude of a random field, then the quantity of interest is the azimuthal correlation function $W(\phi_1, \phi_2) = \langle V^*(\phi_1) V(\phi_2)\rangle$.

$$\langle V^*(\phi_1) V(\phi_2)\rangle = \left\langle \sum_{l_1=-\infty}^{\infty} \alpha_{l_1}^* e^{-il_1\phi_1} \sum_{l_2=-\infty}^{\infty} \alpha_{l_2} e^{il_2\phi_2} \right\rangle$$

$$\text{or,} \quad \langle V^*(\phi_1) V(\phi_2)\rangle = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \left\langle \alpha_{l_1}^* \alpha_{l_2} \right\rangle e^{-il_1\phi_1 + il_2\phi_2},$$

$$\text{or,} \quad W(\phi_1, \phi_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} C_{l_1,l_2} e^{-il_1\phi_1 + il_2\phi_2},$$

The quantity $C_{l_1,l_2} = \langle \alpha_{l_1}^* \alpha_{l_2} \rangle$ can be termed as the OAM-correlation function. We see that the azimuthal correlation function and the OAM-correlation function also form a Fourier transform pair. Therefore, the last equation can be termed as the angular analog of the generalized Wiener-Khintchine theorem. Now, in situation in which the field in angular-stationary, we have $W(\phi_1, \phi_2) = W(\Delta\phi)$. Now, in order for the right hand side of the above equation to depend on $\Delta\phi$ only, we have to have $C_{l_1,l_2} = S_{l_1}\delta_{l_1,l_2}$, where S_l is the spectral density at mode index l and δ_{l_1,l_2} is the Kronecker-delta function. Therefore, the above equation can be written as

$$W(\Delta\phi) = \sum_{l=-\infty}^{\infty} S_l e^{-il\Delta\phi},$$

This equation can be referred to as the angular Wiener-Khintchine theorem.

(3) **Angle Average:** Just as we defined time-averaging and space-averaging, one can have a detection that involves averaging over an angule.

(4) **Ergodicity (angular)** It is now natural to expect an ergodic theorem connecting space averaging with ensemble averaging for spatially stationary fields, just like in the case of time-domain signal. However, since most current experiment do not employ angle-averaging, the concept of angular ergodic theorem has not been explicitly studied as much.

Quantifying the angular correlations

We now go back to Eq. (54). Using this field, the azimuthal intensity $S(\phi)$ at the screen is given by

$$\begin{aligned} & S(\phi) = \langle V^*(\phi)V(\phi)\rangle = |k_1|^2\langle V^*(\phi_1)V(\phi_1)\rangle + |k_2|^2\langle V^*(\phi_2)V(\phi_2)\rangle + k_1^*k_2\langle V^*(\phi_1)V(\phi_2)\rangle + \text{c.c.}, \\ \text{or,} \quad & S(\phi) = |k_1|^2W(\phi_1,\phi_1) + |k_2|^2W(\phi_2,\phi_2) + 2\text{Re}k_1^*k_2W(\phi_1,\phi_2), \\ \text{or,} \quad & S(\phi) = |k_1|^2S(\phi_1) + |k_2|^2S(\phi_2) + 2\text{Re}k_1^*k_2W(\phi_1,\phi_2), \end{aligned}$$

where $S(\phi_1) = W(\phi_1,\phi_1)$, $S(\phi_2) = W(\phi_2,\phi_2)$, and $\tau = t_2 - t_1$. This is the angular interference law. We rewrite the above equation as

$$S(\phi) = |k_1|^2S(\phi_1) + |k_2|^2S(\phi_2) + 2|k_1||k_2|\sqrt{S(\phi_1)S(\phi_2)}|\lambda(\phi_1,\phi_2)|\cos(\arg\lambda + \phi_k),$$

where $k_1^*k_2 = |k_1||k_2|e^{-i\phi_k}$, and $\lambda(\phi_1,\phi_2) = \frac{W(\phi_1,\phi_2)}{\sqrt{S(\phi_1)S(\phi_2)}} = |\lambda(\phi_1,\phi_2)|e^{i\arg\lambda}$ is defined as the complex degree of angular coherence. The visibility of the interference pattern in this case is given by

$$V = \frac{S_{\max} - S_{\max}}{S_{\max} + S_{\max}} = \frac{2|k_1||k_2|\sqrt{S(\phi_1)S(\phi_2)}}{|k_1|^2S(\phi_1) + |k_2|^2S(\phi_2)}|\lambda(\Delta\phi)|$$

For the special case when $|k_1|^2S(\phi_1) = |k_2|^2S(\phi_2)$, we have $V = |\lambda(\Delta\phi)|$, that is, the fringe visibility equals the degree of coherence. Let us consider the following types of angular OAM correlations function:

1. **No correlations between different OAM components**

   In this case the OAM correlation function can be written as $C_{l_1,l_2} = S_{l_1}\delta_{l_1,l_2}$, where δ_{l_1,l_2} is the Kronecker-delta function. Therefore, the azimuthal correlation function now becomes

   $$\begin{aligned} & W(\phi_1,\phi_2) = \sum_{l_1=-\infty}^{\infty}\sum_{l_2=-\infty}^{\infty} S_{l_1}\delta_{l_1,l_2}e^{-il_1\phi_1 + il_2\phi_2}, \\ \text{or,} \quad & W(\phi_1,\phi_2) = \sum_{l_1=-\infty}^{\infty} S_{l_1}e^{-il_1(\phi_1-\phi_2)}, \\ \text{or,} \quad & W(\Delta\phi) = \sum_{l=-\infty}^{\infty} S_l e^{-il\Delta\phi}. \end{aligned}$$

   We note that in this case the azimuthal correlation function depends only on the difference of the angular positions. This can called an angular stationary field. The equivalent situation in the temporal case is when different frequencies have no correlation in between them and also in the spatial case is when different spatial-frequencies have no correlation in between them.

2. **Perfect correlations between different OAM components**

   The OAM correlation function in this case factorizes and can be written as $C_{l_1,l_2} = \alpha_{l_1}^*\alpha_{l_2}$.

3. **Partial correlation between different OAM components**

   Such sources have not been studied as much as the temporally and spatially stationary fields.

Physical realization of the azimuthal-mode (or the OAM-mode) basis

We now study if the azimuthal modes of the type discussed above can be realized in a physical situation. First of all, we note that at a given frequency, the field $U(x,y,z)$ satisfies the Holmholtz equation, that is, $(\nabla^2 + k^2)U(x,y,z) = 0$. Suppose the field $U(x,y,z)$ has a rapid variation in the z-direction to the extent that the field can be written as $U(x,y,z) = U_0(x,y,z)e^{iwz}$, where $U_0(x,y,z)$ varies slowly with z and $w \approx k$. We can then write the Helmholtz equation as

$$\left[\frac{\partial^2U_0(x,y,z)}{\partial x^2} + \frac{\partial^2U_0(x,y,z)}{\partial y^2} + \frac{\partial^2U_0(x,y,z)}{\partial z^2} + 2ik\frac{\partial U_0(x,y,z}{\partial z}\right]e^{ikz} = 0 \tag{56}$$

Now, if we have $\left|\frac{\partial^2}{\partial z^2}U_0(x,y,x)\right| \ll \left|2iw\frac{\partial}{\partial z}U_0(x,y,z)\right|$, which is known as the paraxial approximation, the above equation can be written as

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+2ik\frac{\partial}{\partial z}\right)U_0(x,y,z)=0 \tag{57}$$

This equation is called the paraxial Helmholtz equation and is applicable only when the paraxial approximation is satisfied. In the cylindrical coordinate system, the above equation can be written as

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}+2ik\frac{\partial}{\partial z}\right]U_0(\rho,\phi,z)=0 \tag{58}$$

It turns out that one of the solutions to the above equation is the so called Laguerre-Gaussian mode $LG_p^l(\rho,\phi,z)$, that is, $U_0(\rho,\phi,z) = LG_p^l(\rho,\phi,z)$ is a solution to the above equation. The complete solution is then given by $LG_p^l(\rho,\phi,z)e^{ikz}$. The Laguerre-Gaussian modes form a family of orthogonal modes that have a well defined orbital angular momentum. The field amplitudes $LG_p^l(\rho,\phi,z)$ of normalized Laguerre-Gaussian modes is given by

$$\begin{aligned}LG_p^l(\rho,\phi,z) = \sqrt{\frac{2p!}{\pi(|l|+p)!}}\frac{1}{w(z)}\left[\frac{\sqrt{2}\rho}{w(z)}\right]^{|l|}L_p^l\left[\frac{2\rho^2}{w^2(z)}\right]\\ \times\exp\left[-\frac{\rho^2}{w^2(z)}\right]\exp\left[-\frac{ik^2\rho^2 z}{2(z^2+z_R^2)}\right]\exp\left[i(2p+|l|+1)\tan^{-1}\left(\frac{z}{z_R}\right)\right]e^{-il\phi},\end{aligned} \tag{59}$$

where k is the wave-vector magnitude of the field, z_R the Rayleigh range, and $w(z)$ the radius of the beam at z. L_p^l is the associated Laguerre polynomial given by

$$L_p^l(x)=\sum_{m=0}^{p}(-1)^m C_{p-m}^{p+l}\frac{x^m}{m!}, \tag{60}$$

with $C_k^n = \frac{n!}{k!(n-k)!}$. The amplitude of a Laguerre-Gaussian (LG) mode has an azimuthal phase dependence $e^{-il\phi}$, where l is called the azimuthal mode index. The first few Laguerre polynomials are given by

$$\begin{aligned}LG_0^l(x) &= 1\\ LG_1^l(x) &= -x+l+1\\ LG_2^l(x) &= \frac{1}{2}[x^2-2(l+2)x+(l+1)(l+2)]\end{aligned} \tag{61}$$

The Laguerre-Gaussian modes are the propagating solution to the wave equation. In 1992, Allen *et al.* showed that a light beam with a Laguerre-Gaussian amplitude distribution carries a well-defined orbital angular momentum [11]. The amplitude of a Laguerre-Gaussian (LG) mode has an azimuthal phase dependence $e^{-il\phi}$, where l is called the azimuthal mode index. Allen *et al.* showed that a Laguerre-Gaussian mode with index l possesses an orbital angular momentum of $l\hbar$ per photon. Thus, within quantum mechanics, l has the interpretation that a field with a Laguerre-Gaussian amplitude distribution carries an orbital angular momentum of $l\hbar$ per photon energy $\hbar\omega$ [11]. But within classical mechanics l can not be interpreted that way since there is no photon concept within classical mechanics. Here, l is interpreted as the topological charge of the field. Nevertheless, in this course, we will refer to l as the OAM-mode index even while dealing with the classical coherence theory. Although Laguerre-Guassian modes are, in general, characterized using two indices, l and p, in what follows we consider only those LG modes that have $p = 0$. So in the rest of this chapter, we shall drop the index p while writing LG modes and show the index l alone. Figure 8 shows the intensity and phase patterns of the first three LG modes: LG_0, LG_1 and LG_2. We note that ϕ is a 2π-periodic variable and l is discrete but infinite dimensional. Here, we are representing the orbital angular momentum basis with azimuthal phase given by $e^{-il\phi}$ [11, 12]. Light beams carrying orbital angular momentum are now routinely produced. There are many methods that are currently being employed to produce such beams. One method is based on using the so-called spiral-phase plates [12], and the other is based on using the computer

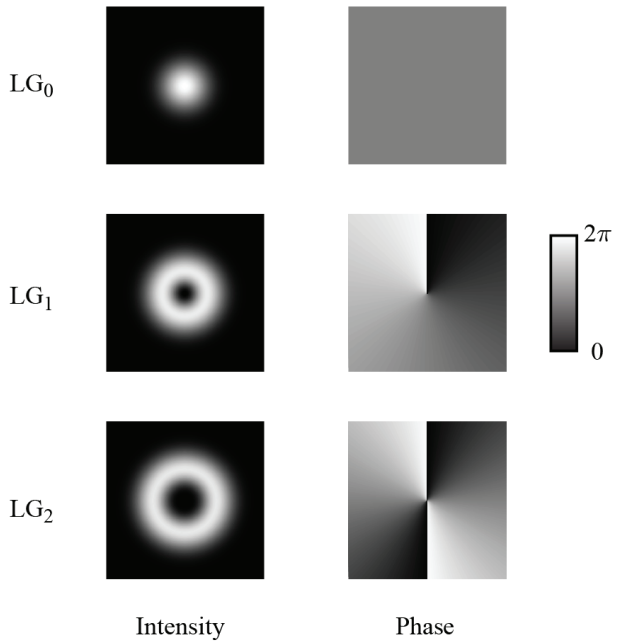


FIG. 8: The intensity and phase patterns of the first three Laguerre-Gaussian modes: LG_0, LG_1, and LG_2. The scales on the intensity plots are arbitraty, whereas the scales on the phase plots go from 0 to 2π, as shown.

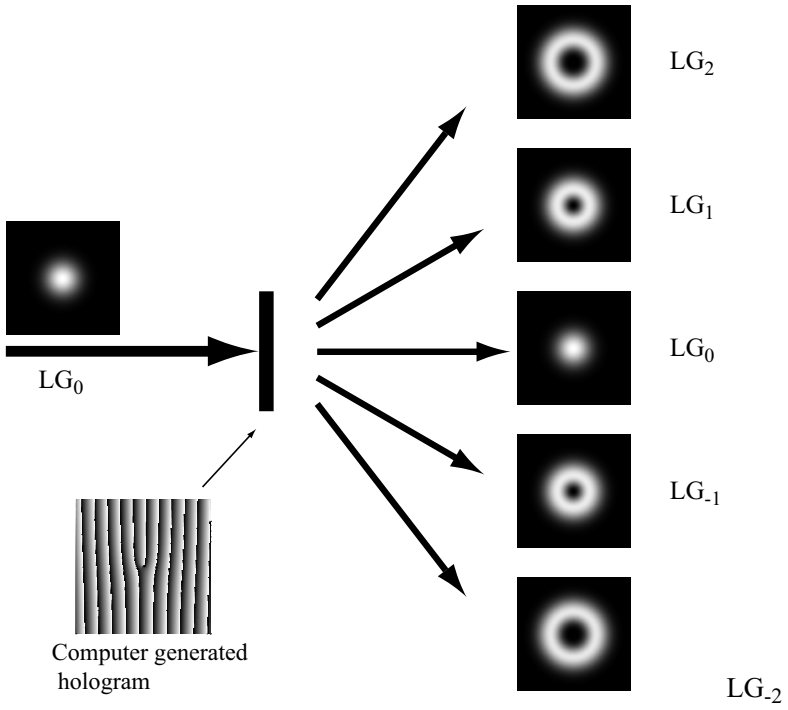


FIG. 9: Schematic of the method used to generate Laguerre-Gaussian (LG) beams using a computer generated hologram. The phase pattern of the computer generated hologram contains a phase grating and an azimuthal phase $e^{-il\phi}$, with $l = 1$. When a Gaussian beam (LG_0) falls on such a hologram, it gets diffracted into many different orders. In the 0th diffraction order, a beam with original phase and intensity structure is obtained. In the ± 1 diffraction order, $LG_{\pm 1}$ beam is obtained. Similarly in ± 2 diffraction order, $LG_{\pm 2}$ beam is obtained, and so on.

generated holograms [13]. Both these methods have their own advantages. Whereas the spiral-phase-plate method is more efficient, the computer generated holograms are easier to implement in experiments. One other method is using

stress birefringence [14]. Figure 9 depicts the schematic of a method used to generate LG beams using computer generated holograms.

LECTURE # 9

Second-order coherence theory (polarization)

So far, we have been studying only the scalar fields in which we were concerned with only the magnitude of a field. We are now going to study vector fields, which have both magnitude and direction. The direction in which an optical field vibrates is called its polarization, and in this section we are going to study the polarization properties of light. Our main aim is to quantify how electric field vibrations in one direction are correlated with the electric field vibrations in some other direction. We will be treating polarization just as we have treated time, space, and angular position. Polarization is a two-dimensional vector. An optical field with its electric field vector oscillating in the horizontal direction is called a horizontally-polarized field (or x-polarized field). Similarly a field with its electric field vector oscillating in the vertical direction is called a vertically-polarized field (or y-polarized field). The vectorial electric field $\boldsymbol{E}(t)$ at time t and at $z = 0$ can be written in the two-dimensional polarization basis as

$$\boldsymbol{E}(t) = E_x(t)\hat{\boldsymbol{x}} + E_y(t)\hat{\boldsymbol{y}}. \tag{62}$$

The field is taken to be monochromatic with frequency ω so that the time dependence of the field is of the form $e^{-i\omega t}$. The field at $z = 0$ is a superposition of fields vibrating in two different directions. The amplitudes $E_x(t)$ and $E_y(t)$ are the complex analytic signal corresponding to the real random electric field amplitudes. The unit vectors $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are along x- and y-directions respectively. In situation in which $E_x(t)$ and $E_y(t)$ are mutually perfectly correlated, Eq. (62) represents a perfectly coherence field. However, in situations in which $E_x(t)$ and $E_y(t)$ are not mutually perfectly correlated, the fields representation in Eq. (62) is only a symbolic representation and in these situations one needs to study relevant correlation functions.

Quantifying the polarization correlation

In order to study correlations in the polarization basis, first of all, we need to have a suitable interferometer. Figure. 13 represents one such scheme. We describe polarization interference using a scheme which is completely analogous to temporal interference using the Mach-Zehnder scheme.

We are interested in finding out how the electric field vibration in these two directions are correlated. To accomplish that we will have to interfere the two different types of vibrations. For that purpose, the field $\boldsymbol{E}(t)$ is first made to pass through a wave-plate, which introduces phases ϵ_1 and ϵ_2 in x- and y-polarization direction, respectively. The field at $z = z_1$, right after passing through the wave-plate is given by

$$\boldsymbol{E}(t, z_1) = E_x(t, z_1)\hat{\boldsymbol{x}} + E_y(t, z_1)\hat{\boldsymbol{y}} = E_x(t)e^{i\epsilon_1}e^{ikz_1}\hat{\boldsymbol{x}} + E_y(t)e^{i\epsilon_2}e^{ikz_1}\hat{\boldsymbol{y}}, \tag{63}$$

where kz_1 is the constant phase acquired due to propagation by fields in both the directions. Here we are assuming that ϵ_1 and ϵ_2 can be changed in a controllable manner. The two field are combined using a polarizer to study the interference effects. A polarizer lets only one kind of polarization pass through while blocking the other. We take the polarizer to be oriented along direction x' which is at an angle θ with respect to the x-direction. The total field $\boldsymbol{E}(t)$ at $z = z_2$ is therefore

$$\begin{aligned}
\boldsymbol{E}(t, z_2) &= E_x(t, z_2)\hat{\boldsymbol{x}} + E_y(t, z_2)\hat{\boldsymbol{y}} \\
&= E_{x'}(t, z_2)\hat{\boldsymbol{x'}} + E_{y'}(t, z_2)\hat{\boldsymbol{y'}} \\
&= \left[E_x(t)e^{i\epsilon_1}\cos\theta + E_y(t)e^{i\epsilon_2}\sin\theta\right]e^{ikz_2}\ \hat{\boldsymbol{x'}} + \left[-E_x(t)e^{i\epsilon_1}\sin\theta + E_y(t)e^{i\epsilon_2}\cos\theta\right]e^{ikz_2}\ \hat{\boldsymbol{y'}}.
\end{aligned}$$

Here $E_{x'}(t, z_2)$ is the part of the field that passes through the polarizer whereas $E_{y'}(t, z_2)$ is the part of the field that gets blocked by the polarizer. To study both parts of the field together, a polarizing beam splitter, instead of a polarizer, should be used, in which case the two fields come out through two separate ports.

(Aside: Let us look at the energy density to make sure that it is conserved. We have

$$\begin{aligned}
w = \langle \boldsymbol{E}^*(t, z_2)\cdot\boldsymbol{E}(t, z_2)\rangle &= \langle E^*_{x'}(t, z_2)E_{x'}(t, z_2)\rangle + \langle E^*_{y'}(t, z_2)E_{y'}(t, z_2)\rangle \\
&= \langle\left[E^*_x(t)e^{-i\epsilon_1}\cos\theta + E^*_y(t)e^{-i\epsilon_2}\sin\theta\right]\left[E_x(t)e^{i\epsilon_1}\cos\theta + E_y(t)e^{i\epsilon_2}\sin\theta\right]\rangle + \\
&\quad \langle\left[-E^*_x(t)e^{-i\epsilon_1}\sin\theta + E^*_y(t)e^{-i\epsilon_2}\cos\theta\right]\left[-E_x(t)e^{i\epsilon_1}\sin\theta + E_y(t)e^{i\epsilon_2}\cos\theta\right]\rangle \\
&= \langle E^*_x(t)E_x(t)\rangle + \langle E^*_y(t)E_y(t)\rangle.
\end{aligned}$$

Here we have used the fact that $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}} = 0$ and $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}} = \hat{\boldsymbol{y}} \cdot \hat{\boldsymbol{y}} = 1$. We find that the above energy density is same as the energy density of the field $\boldsymbol{E}(t)$ at $z = 0$ and thus that energy density is conserved as it should be.)

So, the part of the field that passes through the polarizer is

$$E_{x'}(t, z_2) = \left[E_x(t)e^{i\epsilon_1} \cos\theta + E_y(t)e^{i\epsilon_2} \sin\theta\right] e^{ikz_2}$$

We find the field that passes through the polarizer is a superposition of fields in the x and y directions. So, by studying the intensity or the energy density of the field passing through the polarizer, we should be able to study the polarization correlations. Now, the energy density $w_{x'}$ at the output of the polarizer is

$$\begin{aligned} w_{x'} &= \langle E^*_{x'}(t, z_2) E_{x'}(t, z_2)\rangle \\ &= \langle \left[E^*_x(t)e^{-i\epsilon_1} \cos\theta + E^*_y(t)e^{-i\epsilon_2} \sin\theta\right] \left[E_x(t)e^{i\epsilon_1} \cos\theta + E_y(t)e^{i\epsilon_2} \sin\theta\right]\rangle \\ &= \langle E^*_x(t)E_x(t)\rangle \cos^2\theta + \langle E^*_y(t)E_y(t)\rangle \sin^2\theta + \langle E^*_x(t)E_y(t)\rangle \cos\theta \sin\theta e^{i(\epsilon_2 - \epsilon_1)} + \text{c.c.} \\ &= J_{xx} \cos^2\theta + J_{yy} \sin^2\theta + J_{xy} \cos\theta \sin\theta e^{i\delta} + J_{yx} \cos\theta \sin\theta e^{-i\delta}, \end{aligned}$$

where $\delta = \epsilon_2 - \epsilon_1$. Let us write

$$J_{xy} = \sqrt{J_{xx}J_{yy}} \frac{J_{xy}}{\sqrt{J_{xx}J_{yy}}} = \sqrt{J_{xx}J_{yy}} j_{xy} = \sqrt{J_{xx}J_{yy}} |j_{xy}| e^{i\beta_{xy}}.$$

The quantity $|j_{xy}|$ is called the degree of coherence between the two polarization components. The energy density can now be written as

$$w_{x'} = J_{xx} \cos^2\theta + J_{yy} \sin^2\theta + 2\sqrt{J_{xx}J_{yy}} \cos\theta \sin\theta |j_{xy}| \cos(\delta + \beta_{xy}). \tag{64}$$

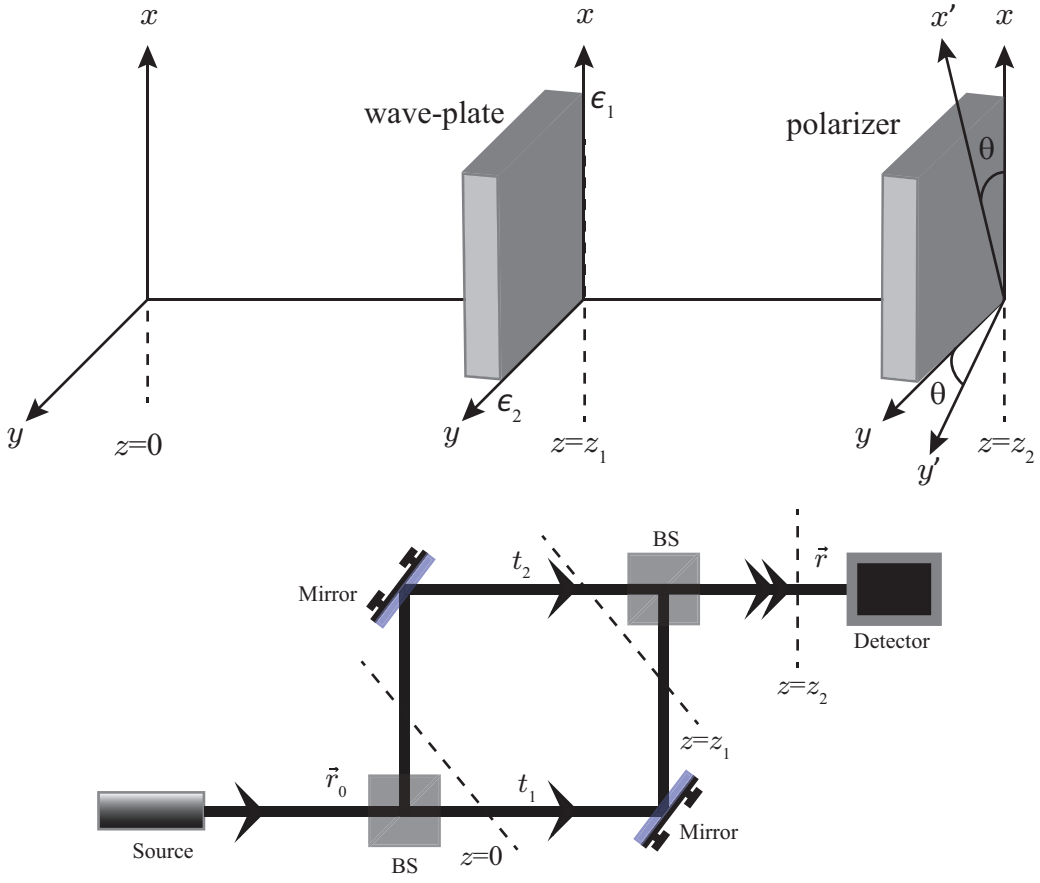


FIG. 10: Polarization Interference and its analogy with temporal interference.

This is the interference law for polarization. We note that the polarization interference law has the same generic structure as in the case of time, space, and angular position. If we take $k_1 = \cos\theta$ and $k_2 = \sin\theta$, we note the exact similarity. The expression for the energy density depends on two variables, δ and θ. Just like in the earlier situations, we define an expression for interference visibility that is based on optimizing the extrema over δ, while keeping θ fixed. The expression for visibility takes the following form.

$$V = \frac{\langle w_{x'}\rangle_{\max(\delta)} - \langle w_{x'}\rangle_{\min(\delta)}}{\langle w_{x'}\rangle_{\max(\delta)} + \langle w_{x'}\rangle_{\min(\delta)}} = \frac{2|k_1||k_2|\sqrt{J_{xx}J_{yy}}}{|k_1|^2 J_{xx} + |k_2|^2 J_{yy}}|j_{xy}| = \frac{2\cos\theta\sin\theta\sqrt{J_{xx}J_{yy}}}{\cos^2\theta J_{xx} + \sin^2\theta J_{yy}}|j_{xy}|.$$

We note that the expression for visibility has the same structure as in the case of temporal, spatial or angular coherence. Also, when $k_1\sqrt{J_{xx}} = k_2\sqrt{J_{yy}}$, that is $\cos\theta\sqrt{J_{xx}} = \sin\theta\sqrt{J_{yy}}$, $V = |j_{xy}|$, that is, the visibility becomes equal to the degree of coherence between two polarizations.

So far, we have seen how polarization interference is similar to the time-interference, space-interference or angle-interference. But there are two aspects of polarization basis for which we have not established the analogy. One is the Fourier space of polarization basis and also whether we have a concept like coherence-polarization, just like coherence-time, etc. The answer to the two questions are correlated. In fact, one can certainly have the Fourier space representation for polarization but since polarization is only a two-dimensional basis, the Fourier basis is also just two-dimensional. The Fourier representation of polarization basis can be written as

$$E_x(t) = E_{45}(t) + E_{-45}(t)$$
$$E_y(t) = E_{45}(t) - E_{-45}(t).$$

Here $E_{45}(t)$ and $E_{-45}(t)$ are the 45^0-polarized and -45^0-polarized fields, which are the Fourier components in this case. Also, we can write the above fields as

$$E_x(t) = E_{\rm rcp}(t) + iE_{\rm lcp}(t)$$
$$E_y(t) = E_{\rm rcp}(t) - iE_{\rm lcp}(t),$$

where $E_{\rm rcp}(t)$ and $E_{\rm lcp}(t)$ are the right-circularly polarized and the left-circularly polarized fields.

Now, regarding the second question, we note that in the context of polarization coherence, there is no concept that is analogous to the concept of coherence-time as in time-domain or transverse coherence area as in space-domain. The reason for this is that these other bases are infinite dimensional, and in an infinite dimensional basis, one can have a sense a width or a range. However, polarization is a two-dimensional basis that has only two basis vectors and so a width cannot be defined in this case.

LECTURE # 10

Visibility of Polarization Interference

Now, we look at one feature of coherence in the polarization basis that is not a part of the current theory of coherence in either time, space or angular position. This concept is called the degree of polarization which is a unique, basis-independent way of quantifying how much polarization-correlation a field has. We note that Eq. (64) has two variables, δ and θ. For calculating the visibility above, we found the extremum of the energy density by optimizing over only the variable δ. We now define the visibility that is based on the extremum of the energy density optimized over both the variables δ and θ. The expression for visibility in this case becomes

$$V = \frac{\langle w_{x'}\rangle_{\max(\delta,\theta)} - \langle w_{x'}\rangle_{\min(\delta,\theta)}}{\langle w_{x'}\rangle_{\max(\delta,\theta)} + \langle w_{x'}\rangle_{\min(\delta,\theta)}} \tag{65}$$

For calculating the extremum of energy density $w_{x'}$, we first substitute $\cos^2\theta = \frac{1+\cos 2\theta}{2}$ and $\sin^2\theta = \frac{1-\cos 2\theta}{2}$, and rewrite Eq. (64) as

$$w_{x'} = \frac{1}{2}\left[J_{xx} + J_{yy} + (J_{xx} - J_{yy})\cos 2\theta + 2\sqrt{J_{xx}J_{yy}}\sin 2\theta|j_{xy}|\cos(\delta + \beta_{xy})\right].$$

We keep θ fixed and find the extremum of $w_{x'}$ by varying δ. We obtain

$$\langle w_{x'}\rangle_{\max(\delta)} = \frac{1}{2}\left[J_{xx} + J_{yy} + (J_{xx} - J_{yy})\cos 2\theta + 2\sqrt{J_{xx}J_{yy}}\sin 2\theta|j_{xy}|\right]; \quad \text{when} \quad \delta = -\beta_{xy} + 2m\pi$$
$$\langle w_{x'}\rangle_{\min(\delta)} = \frac{1}{2}\left[J_{xx} + J_{yy} + (J_{xx} - J_{yy})\cos 2\theta - 2\sqrt{J_{xx}J_{yy}}\sin 2\theta|j_{xy}|\right]; \quad \text{when} \quad \delta = -\beta_{xy} + (2m+1)\pi,$$

where $m = 0, 1, 2, \cdots$. Now we need to find the value of θ for which $\langle w_{x'}\rangle_{\max(\delta)}$ becomes maximum and also the value of θ for which $\langle w_{x'}\rangle_{\min(\delta)}$ becomes minimum. Let us first calculate $\langle w_{x'}\rangle_{\max(\delta,\theta)}$. By taking the derivative of the first equation with respect to θ, we find the condition for the maximum to be

$$(J_{xx} - J_{yy})\sin 2\theta = 2\sqrt{J_{xx}J_{yy}}\cos 2\theta|j_{xy}| = 2\sqrt{J_{xy}J_{yx}}\cos 2\theta,$$

which implies that we need to have

$$\sin 2\theta = \frac{2\sqrt{J_{xy}J_{yx}}}{\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}},$$
$$\text{and,} \qquad \cos 2\theta = \frac{J_{xx} - J_{yy}}{\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}}.$$

Substituting for the values of $\sin 2\theta$ and $\cos 2\theta$, we find that the maximum energy density $\langle w_{x'}\rangle_{\max(\delta,\theta)}$ is

$$\begin{aligned}
\langle w_{x'}\rangle_{\max(\delta,\theta)} &= \frac{1}{2}\left[J_{xx} + J_{yy} + \frac{(J_{xx} - J_{yy})(J_{xx} - J_{yy}) + 2\sqrt{J_{xx}J_{yy}}2\sqrt{J_{xx}J_{yy}}}{\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}}\right] \\
&= \frac{1}{2}\left[J_{xx} + J_{yy} + \frac{(J_{xx} - J_{yy})^2 + 4J_{xx}J_{yy}}{\sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}}\right] \\
&= \frac{1}{2}\left[J_{xx} + J_{yy} + \sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}\right]
\end{aligned}$$

Similarly, we can show that the minimum energy density $\langle w_{x'}\rangle_{\min(\delta,\theta)}$, optimized over both the variables is

$$\langle w_{x'}\rangle_{\min(\delta,\theta)} = \frac{1}{2}\left[J_{xx} + J_{yy} - \sqrt{(J_{xx} - J_{yy})^2 + 4J_{xy}J_{yx}}\right]$$

The visibility of the polarization interference pattern can now be calculated by substituting for $\langle w_{x'}\rangle_{\max(\delta,\theta)}$ and $\langle w_{x'}\rangle_{\min(\delta,\theta)}$ in Eq. (65). That is

$$\begin{aligned} V = \frac{\langle w_{x'}\rangle_{\max(\delta,\theta)} - \langle w_{x'}\rangle_{\min(\delta,\theta)}}{\langle w_{x'}\rangle_{\max(\delta,\theta)} + \langle w_{x'}\rangle_{\min(\delta,\theta)}} &= \frac{2\sqrt{(J_{xx}-J_{yy})^2 + 4J_{xy}J_{yx}}}{2(J_{xx}+J_{yy})} \\ &= \left[\frac{(J_{xx}-J_{yy})^2 + 4J_{xy}J_{yx}}{(J_{xx}+J_{yy})^2}\right]^{1/2} \\ &= \left[\frac{(J_{xx}+J_{yy})^2 - 4J_{xx}J_{yy} + 4J_{xy}J_{yx}}{(J_{xx}+J_{yy})^2}\right]^{1/2} \\ &= \left[1 - \frac{4(J_{xx}J_{yy} - J_{xy}J_{yx})}{(J_{xx}+J_{yy})^2}\right]^{1/2} \end{aligned} \tag{66}$$

Before, we proceed further, recall that our electric field at $z = 0$ can be written in terms of a row vector as

$$\boldsymbol{E} = E_x(t)\hat{\boldsymbol{x}} + E_y(t)\hat{\boldsymbol{y}} = [E_x(t) \quad E_y(t)] \tag{67}$$

In situation in which the field components in the two directions are not mutually completely coherent, we can represent the state in terms of a 2×2 density matrix J, which is called the polarization matrix. In the x-y coordinate system it takes the form

$$\mathrm{J} = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix} = \begin{pmatrix} \langle E_x^*(t)E_x(t)\rangle & \langle E_x^*(t)E_y(t)\rangle \\ \langle E_y^*(t)E_x(t)\rangle & \langle E_y^*(t)E_y(t)\rangle \end{pmatrix}.$$

J is a density matrix representing a physical system. So, we have tr J $= J_{xx} + J_{yy}$ as the trace of the matrix J and det J $= J_{xx}J_{yy} - J_{xy}J_{yx}$ as its determinant. And in terms of these quantities, the visibility can be written as

$$V = \left[1 - \frac{4\text{det J}}{(\text{tr J})^2}\right]^{1/2}, \tag{68}$$

We find that the visibility V that is optimized over both δ and θ, is a basis-independent quantity, that is, it does not depend of the choice of δ or θ. Now, we analyze the two extreme cases of visibility. The first case is when $V = 0$. From our discussion in the previous section, we understand that for a completely unpolarized field the correlations between two orthogonal polarization components should be zero and as a result the visibility V should be equal to zero. This means that

$$\begin{aligned} & (\text{tr J})^2 - 4\text{det J} = 0 \\ \text{or,} \quad & (J_{xx}+J_{yy})^2 - 4(J_{xx}J_{yy} - J_{xy}J_{yx}) = 0 \\ \text{or,} \quad & (J_{xx}-J_{yy})^2 + 4J_{xy}J_{yx} = 0. \\ \text{or,} \quad & (J_{xx}-J_{yy})^2 + 4|J_{xy}|^2 = 0. \end{aligned}$$

Both the terms in the summation on the left hand side are positive. So, in order that the left hand side equals to 0, we have to have $J_{xx} = J_{yy} = A$, and $J_{xy} = J_{yx} = 0$. These conditions dictate that the form of the polarization matrix for a completely unpolarized field (or $V = 0$) is

$$\mathrm{J}^{\text{unpol}} = A\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

Where $A \geq 0$. Next, we look at the case of a completely polarized field, that is, $V = 1$. Imposing this constraint on the visibility, we obtain the condition that det J $= 0$, that is, $|j_{xy}| = 1$. As a result, the polarization matrix in this case takes the following form

$$\mathrm{J}^{\text{pol}} = \begin{pmatrix} B & D \\ D^* & C \end{pmatrix},$$

with $BC - DD^* = 0$, where $B \geq 0$, $C \geq 0$, and D is a complex number.

Degree of Polarization

In this section, we will try to find a concrete physical interpretation for our basis-independent visibility derived in the previous section. For this purpose, we use the fact that a polarization matrix J can be uniquely represented as a sum of two matrices, one of which is completely polarized $\mathrm{J}^{\mathrm{pol}}$ and the other one completely unpolarized $\mathrm{J}^{\mathrm{unpol}}$, that is,

$$\mathrm{J} = \mathrm{J}^{\mathrm{pol}} + \mathrm{J}^{\mathrm{unpol}} = \begin{pmatrix} B & D \\ D^* & C \end{pmatrix} + A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{69}$$

Here $A \geq 0$, $B \geq 0$ and $C \geq 0$. Also $BC - DD^* = 0$. Let us first prove that the above decomposition is unique. Equating the left and the right hand side of the above equation, we obtain the conditions that

$$\begin{aligned} A + B &= J_{xx}, \\ A + C &= J_{yy}, \\ D &= J_{xy}, \\ D^* &= J_{yx}. \end{aligned}$$

Substituting for B and C from the above conditions in $BC - DD^* = 0$, we get

$$(J_{xx} - A)(J_{yy} - A) - J_{xy}J_{yx} = 0.$$

This is also the eigenvalue equation of matrix J, solving which, we obtain two solutions for A

$$\begin{aligned} A_1 &= \frac{\mathrm{tr\ J} + \sqrt{(\mathrm{tr\ J})^2 - 4\mathrm{det\ J}}}{2} \\ A_2 &= \frac{\mathrm{tr\ J} - \sqrt{(\mathrm{tr\ J})^2 - 4\mathrm{det\ J}}}{2} \end{aligned}$$

The two solutions of A are the eigenvalues of matrix J. Let us take the second solution $A = A_2$. We find that in this case,

$$B_2 = J_{xx} - A_2 = \frac{2J_{xx} - \mathrm{tr\ J} + \sqrt{(\mathrm{tr\ J})^2 - 4\mathrm{det\ J}}}{2} = \frac{(J_{xx} - J_{yy}) + \sqrt{(J_{xx} - J_{yy})^2 + 4|J_{xy}|^2}}{2} \geq 0.$$

In a similar manner it can be shown that for $A = A_2$, $C_2 \geq 0$. Next, it can similarly be shown that for the existence of the solution A_1, the terms B and C need to be negative, which makes the above decomposition unphysical. Thus, it is proved that the decomposition shown in Eq. (69) is unique.

The degree of polarization P can be defined as the fraction of the total energy density that is completely polarized, that is,

$$P = \frac{w^{\mathrm{pol}}}{w} = \frac{\mathrm{trJ}^{\mathrm{pol}}}{\mathrm{trJ}^{\mathrm{pol}} + \mathrm{trJ}^{\mathrm{unpol}}} = \frac{\mathrm{trJ}^{\mathrm{pol}}}{\mathrm{trJ}} = \frac{B + C}{2A + B + C}, \tag{70}$$

where w is the total energy density and w^{pol} is the energy density of the part of the field which is completely polarized. Using the values of B and C from above and using $A = A_2$, we find that the degree of polarization

$$P = \frac{\sqrt{(J_{xx} - J_{yy})^2 + 4|J_{xy}|^2}}{J_{xx} + J_{yy}} = \frac{\sqrt{(\mathrm{tr\ J})^2 - 4\mathrm{det\ J}}}{\mathrm{tr\ J}} = \left[1 - \frac{4\mathrm{det\ J}}{(\mathrm{tr\ J})^2}\right]^{1/2} = V$$

We see that the degree of polarization is equal to the basis-independent visibility of polarization interference, as evaluated in Eq. 66. Therefore, it becomes clear that the basis-independent visibility measurement is the unique way of quantifying the degree of polarization correlations in a field. Now, if the matrix J has the trace equal to unity, then $J_{xx} + J_{yy} = 2A + B + C = 1$. This means that $P = B + C$, and therefore, $A = (1 - P)/2$. Using the values of A and $B + C$ in terms of P, Eq. (69) can be written as

$$\mathrm{J} = \mathrm{J}^{\mathrm{pol}} + \mathrm{J}^{\mathrm{unpol}} = P \begin{pmatrix} B/P & D/P \\ D^*/P & C/P \end{pmatrix} + (1 - P) \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = P\mathrm{J}^{\mathrm{pol}}_{\mathrm{unit}} + (1 - P)\mathrm{J}^{\mathrm{unpol}}_{\mathrm{unit}}. \tag{71}$$

This means that if we write the polarization matrices corresponding to the polarized part and the unpolarized part as matrices with unit trace then the degree of polarization is simply the weightage of the polarized part.

LECTURE # 11

Coherent Mode Representation of Optical Fields

A coherent mode represents a completely coherent field. Coherent mode representation is a way of representing a partially coherent field as a mixture of coherent modes. If the sum contains only one term then it is already a coherent field. The motivation for finding the coherent-mode representation is that it is relatively much easier to deal with coherent modes for propagation problems. So, if a partially coherent field can be represented as a mixture of a bunch of coherent modes then one does not have to propagate the entire field. All one has to do is to propagate the individual coherent modes and after propagation add all the modes together. This makes the propagation problem much easier to handle. The discussion regarding the coherent versus incoherent mode is the same as the discussion pertaining to pure state versus mixed state in quantum mechanics. So, whatever general formalism we work out here will be applicable to the mixed states in quantum coherence theory as well.

1. finite dimensional basis (Polarization Basis)

A coherent mode represents a completely coherent field. Coherent mode representation is a way of representing a partially coherent field as a mixture of coherent modes. If the sum contains only one term then it is already a completely coherent field. We also note that the coherent mode representation in the case of a finite dimensional basis essentially means the diagolanlization of of the matrix representing the state of the system. First, let us look at the conditions that are required for the diagonalization of a matrix J. We know that, if

(i) tr J =1;

(ii) J is Hermitian, that is, J = J*; and

(iii) J is positive definite matrix, that is, $\sum_{i=1}^{2}\sum_{j=1}^{2} a_i^* a_j J_{ij} \geq 0$. For a Hermitian matrix J, this means that J has only positive eigenvalues.

then the matrix J can be diagonalized, with at least one non-zero eigenvalue. We find that our polarization matrix satisfies all the criteria needed for the diagonalization and hence it is proved that it can be diagonalized. Now, let us assume that there exist an eigen-basis ($\hat{\boldsymbol{x}}'$, $\hat{\boldsymbol{y}}'$) in which the matrix can be diagonalized with eigenvalues A_1 and A_2, with $A_1 + A_2 = 1$. Therefore, we can write a general polarization matrix in the eigen-basis ($\hat{\boldsymbol{x}}'$, $\hat{\boldsymbol{y}}'$) as

$$\mathrm{J} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = A_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + A_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

We note that in the eigen basis, the matrix J has been written as an incoherent sum of two completely-polarized fields. If $A_1 = 0$, then it represent a coherent state polarized along $\hat{\boldsymbol{y}}'$. Also, if $A_1 = A_2 = 1/2$, then the state above represents an equal mixture of two completely-polarized field. In fact, this is always true if the eigenvalues are degenerate. And in this case it can be verified that eigen-basis in not unique and that any orthonormal polarization basis can be the eigen-basis. We note that if the state is coherent, then one can always find a basis in which the polarization matrix has only one diagonal term. This is the generic structure of a completely polarized state.

We recall that by solving the eigenvalue equation of J, we obtain two eigenvalues A_1 and A_2, which are both non-negative. In terms of these eigenvalues, we can write the density matrix J in the eigen-basis as

$$\mathrm{J} = \mathrm{A}_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \mathrm{A}_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

which is the coherent mode representation. It can now be shown that tr J = A$_1$+A$_2$ and (tr J)^2−4det J = |A$_1$−A$_2$|, using which, we obtain the connection between P and the two eigenvalues, which is

$$P = \left[1 - \frac{4\det \mathrm{J}}{(\mathrm{tr}\ \mathrm{J})^2}\right]^{1/2} = \frac{|A_1 - A_2|}{A_1 + A_2}.$$

2. Infinite dimensional basis

Let us now move to the infinite dimensional basis, which is the case when we are working in time, space, or angular-position basis. A fully coherent field in the position basis is represented using the electric field $E(\boldsymbol{\rho})$. However, as

we have studied earlier, a partially coherent field cannot be written in terms of the electric field and in that case one has to represent the field in terms of the cross-spectral density function $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \langle E^*(\boldsymbol{\rho}_1)E(\boldsymbol{\rho}_2)\rangle$. If the field is fully coherent, the cross-spectral density takes the form: $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = E^*(\boldsymbol{\rho}_1)E(\boldsymbol{\rho}_2)$. However, in general, a field is not fully coherent and in that case the question is whether one could represent such a field as an incoherent mixture of coherent modes. The conditions for coherent-mode representation for a correlation function are stated as follows.

If

(i) $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$ is square integrable, that is, $\iint_D |W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)|^2 d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2 < \infty$;
(ii) $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$ is Hermitian, that is, $W^*(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = W(\boldsymbol{\rho}_2, \boldsymbol{\rho}_1)$; and
(iii) $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$ is non-negative definite function, that is, $\iint_D f^*(\boldsymbol{\rho}_1)f(\boldsymbol{\rho}_2)W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2 \geq 0$,

then $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$ is said to be a Hilbert-Schmidt Kernel and that it has a coherent-mode representation of the form:

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \sum_n \alpha_n \psi_n^*(\boldsymbol{\rho}_1)\psi_n(\boldsymbol{\rho}_2) \tag{72}$$

The function $\psi_n(\boldsymbol{\rho})$ are the eigenfunctions and the co-efficients α_n are the eigenvalues of the integral equation

$$\int W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)\psi_n(\boldsymbol{\rho}_1)d^2\boldsymbol{\rho}_1 = \alpha_n\psi_n(\boldsymbol{\rho}_2) \tag{73}$$

The first condition is just the requirement that the total energy should be finite. This is equivalent to the statement that the trace is finite. The second condition is to ensure that the associated spectral densities are real. And the last condition is to ensure that the eigenvalues are all non-negative. We see that all these conditions are readily fulfilled by the cross-spectral density function. The Hermiticity and the non-negative definiteness conditions assures that the integral equation has at least one nonzero eigenvalues and all the eigenvalues are real and non-negative, that is, $\alpha_n \geq 0$. The above equation can be written as

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \sum_n \alpha_n W^{(n)}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2), \tag{74}$$

where $W^{(n)}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \psi_n^*(\boldsymbol{\rho}_1)\psi_n(\boldsymbol{\rho}_2)$. We note that here there is no ensemble averaging since $\psi_n(\boldsymbol{\rho})$ represents a coherent mode. The above decomposition implies that for every partially coherent field, there exists at least one basis in which the cross-spectral density function can be represented as a mixture of modes that are completely coherent in the space-frequency domain. We note that the eigenfunctions form an orthonormal set, that is, $\iint \psi_n^*(\boldsymbol{\rho}_1)\psi_n(\boldsymbol{\rho}_2)d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2 = \delta_{mn}$. So, we see that the coherent mode representation is all about solving the integral equation to find the appropriate eigenfunctions and eigenvalues. In some situations, one may have an infinite number of terms in the coherent mode representation, in which case it can be written as

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \int g(\boldsymbol{k})W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{k})d^2\boldsymbol{k}, \tag{75}$$

We now look at a few examples:

(i) Fields with no correlation between different frequencies, that is, $W(\omega_1, \omega_2) = S(\omega_1)\delta(\omega_1 - \omega_2)$.
We know that the cross-correlation function can be written as

$$\Gamma(t_1, t_2) = \iint W(\omega_1, \omega_2)e^{i(\omega_1 t_1 - \omega_2 t_2)}d\omega_1 d\omega_2. \tag{76}$$

Substituting $W(\omega_1, \omega_2) = S(\omega_1)\delta(\omega_1 - \omega_2)$, we can write the cross-correlation function as

$$\Gamma(t_1, t_2) = \int S(\omega_1)e^{-i\omega_1(t_2 - t_1)}d\omega_1 \tag{77}$$

$$= \int S(\omega)e^{-i\omega(t_2 - t_1)}d\omega \tag{78}$$

$$= \int S(\omega)\Gamma((t_1, t_2, \omega)d\omega. \tag{79}$$

The last equation is the coherent mode representation of $\Gamma(t_1, t_2)$. Here $e^{-i\omega t}$ is the eigen function and $\Gamma((t_1, t_2, \omega) = e^{-i\omega(t_2 - t_1)}$ are the coherent modes.

(ii) Fields with no correlation between different spatial frequencies: $\mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) = S(\boldsymbol{q_1})\delta(\boldsymbol{q_1} - \boldsymbol{q_2})$
We have for the cross-spectral density function

$$W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \iint_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{q_1}, \boldsymbol{q_2}) e^{-i\boldsymbol{q_1}.\boldsymbol{\rho}_1 + i\boldsymbol{q_2}.\boldsymbol{\rho}_2} d^2\boldsymbol{q_1} d^2\boldsymbol{q_2}.$$

Substituting for the angular correlation function we obtain

$$\begin{aligned} W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) &= \int_{-\infty}^{\infty} S(\boldsymbol{q_1}) e^{i\boldsymbol{q_1}.(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)} d^2\boldsymbol{q} \\ &= \int_{-\infty}^{\infty} S(\boldsymbol{q}) e^{i\boldsymbol{q}.(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)} d^2\boldsymbol{q} \\ &= \int_{-\infty}^{\infty} S(\boldsymbol{q}) W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{q}) d^2\boldsymbol{q}. \end{aligned}$$

This is a coherent-mode representation of $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2)$. Here $e^{i\boldsymbol{q}.\boldsymbol{\rho}}$ is the eigen function and $W(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{q}) = e^{i\boldsymbol{q}.(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)}$ are the coherent modes.

(iii) Fields with no correlation between OAM modes: $C_{l_1,l_2} = C_{l_1}\delta_{l_1,l_2}$
In this case we have

$$W(\phi_1, \phi_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} C_{l_1,l_2} e^{-il_1\phi_1 + il_2\phi_2}$$

Substituting for the above correlation function, we get

$$\begin{aligned} W(\phi_1, \phi_2) &= \sum_{l_1=-\infty}^{\infty} C_{l_1} e^{il_1(\phi_2 - \phi_1)} \\ &= \sum_{l=-\infty}^{\infty} C_l e^{il(\phi_2 - \phi_1)} \\ &= \sum_{l=-\infty}^{\infty} C_l W_l(\phi_1, \phi_2), \end{aligned}$$

which is the coherent-mode representation of the field. Here, $e^{il\phi}$ is the eigen function and $W_l(\phi_1, \phi_2)$ are the coherent modes.

LECTURE # 12

Review of Quantum Mechanics

In classical mechanics, one straightaway talks about a field but in quantum mechanics things are described using the dual language of the state and the operator. The very basic ingredients of quantum mechanics are as follows:

1. **State:** A system in quantum mechanics is described using a state vector $|\psi\rangle$. The system can be finite dimensional or even infinite dimensional. As long as a system can be completely described using a state vector, the system is said to be in a pure state. However, when a quantum system cannot be described as consisting of a single pure-state ensemble, it is said to be in a mixed state and any pure state representation in this case is just symbolic. So, just as in classical wave theory where we have coherent and partially coherent states, in quantum mechanics we have pure and mixed states. The connections between coherent and partially coherent states are the as those between the pure and mixed states. We now look at the concepts of pure and mixed states in some detail.

   (a) **Pure State:** In situations in which the state of a quantum system can be completely described using a single state vector $|\psi\rangle$, then the quantum system is said to be in a pure state. In other words, if a system can be represented by a single ensemble of realizations, then the system is said to be in a pure state. A pure state can be represented as a column vector, the dimensionality of which is the dimensionality of the Hilbert space in which the system lives. A pure state $|\psi\rangle$ in an N-dimensional space can be written as

$$|\psi\rangle = \sum_{i=1}^{N} d_i |\phi_i\rangle = d_1 \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix} + d_2 \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ 0 \end{pmatrix} + \cdots + d_N \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ d_N \end{pmatrix} \tag{80}$$

   with $\sum_{i=1}^{N} |d_i|^2 = 1$. Here $|\phi_i\rangle$ are the basis states and $|d_i|^2$ is the probability that the state is found to be in the basis state $|\phi_i\rangle$. In situation, in which the state space is continuous the summation gets replaced by an integral.

   (b) **Mixed State:** When the state of a quantum system cannot be described as a single ensemble of realizations and has to be described as a mixture of ensembles of realization then the quantum system is said to be in a mixed state. For a mixed state the representation given by Eq. (80) is just symbolic, and just as in the case of classical field, the state of the quantum system in this case needs to be represented in terms of an ensemble averaged quantity. This is done by representing the quantum state by a density matrix ρ defined as

$$\begin{aligned} \rho &= \langle |\psi\rangle\langle\psi| \rangle_e \\ &= \sum_{ij} \langle d_i d_j^* \rangle_e |\phi_i\rangle\langle\phi_j| \end{aligned} \tag{81}$$

   The density matrix elements, which in a sense represent correlation between different basis states, should be compared with the cross-correlation functions of the second-order coherence theory, which also quantify correlations between two basis states. Now, if the density matrix satisfies the conditions similar to the ones for coherent-mode-representation then one can find a new set of basis vectors $|\psi_i\rangle$ in which the density matrix takes the diagonal form and be written as

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \sum_i p_i \rho_i, \tag{82}$$

   where p_i is a positive real number. Equation (82) should be compared with Eq. (74), which is the coherent mode representation for the correlation function. Equation (82) shows that a density matrix ρ can be represented a sum of the density matrices ρ_i with p_i as their weightage in the sum. In other words, when the quantum state is not pure, it can no longer be represented as a single ensemble of realizations but can

be represented as a sum of different pure-state ensemble of realizations, with the pure states given by $|\psi_i\rangle$ and their weightage in the ensemble being given by p_i. Thus, if the above sum contains only one term the state ρ represents a pure state while when it contains more than one term, the state ρ represents a mixed state. The degree of mixedness of a quantum state in a given basis is directly related to the partial coherence of the state in that basis. In the subsequent lectures, we will see how to calculate the degree of mixedness of a quantum state.

We note that when the quantum state of a system is represented using a matrix the associated matrix elements take physical meaning. For example, in the density matrix of Eq. (81), the diagonal element $|d_i|^2$ represents the probability that the quantum system is found in the state ϕ_i; the off-diagonal element $\langle d_i d_j^* \rangle_e$ represents the correlation between basis states $|\phi_i\rangle$ and $|\phi_j\rangle$; the eigenfunctions $|\psi_i\rangle$ and the eigenvalues p_i of the density matrix represent pure states and their weightage, respectively, with which the quantum state ρ can be represented. Furthermore, the matrix properties and operations imply physical conditions and operations on the quantum system. For example, the trace of a density matrix represents the total energy or the total probability of the system; the determinant of a density matrix is a pure-state quantifier in the sense that if the determinant is zero the state is pure otherwise it is mixed; diagonalizing a density matrix is equivalent to finding its pure states ensembles; multiplying a density matrix with a unitary matrix implies rotating the quantum state in its state space; multiplying the density matrix with a projection operator is equivalent to making a measurement on the quantum system, and so on.

2. **Operator:** An operator $\hat{A}$ is represented by an $n \times n$ matrix, where n is the dimensionality of the state-space. There can be several different types of operators in quantum mechanics. An observable in quantum mechanics is represented by a Hermitian operator; the rotation of a quantum state in the Hilbert space is represented by a Unitary operator; the act of measurement is represented by a projection operator, and so no.

3. **Measurement:** The action of an operator on a quantum states constitutes measurement in quantum mechanics. Every time a measurement is performed on a quantum system, the quantum state collapses into one of the eigenstates. Thus, in the context of measurement in quantum mechanics, the quantity that is physically most relevant is the expectation value of an operator. The expectation value of an operator $\hat{A}$ when it acts on a pure state state $|\psi\rangle$ is $\langle\psi|\hat{A}|\psi\rangle$. However, if the system is in a mixed state, the expectation value is given by the sum of the expectation values over all pure state ensembles, that is, $\langle\langle\psi|\hat{A}|\psi\rangle\rangle_e = \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle$, where $\sum_i p_i = 1$. We express the ensemble average in terms of the density matrix in the following manner:

$$\begin{aligned}
\langle\langle\psi|\hat{A}|\psi\rangle\rangle_e &= \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle \\
&= \sum_i \sum_j p_i \langle\psi_i|\hat{A}|\psi_j\rangle\langle\psi_j|\psi_i\rangle \\
&= \sum_i \sum_j p_i \langle\psi_j|\psi_i\rangle\langle\psi_i|\hat{A}|\psi_j\rangle \\
&= \sum_j \langle\psi_j| \sum_i p_i |\psi_i\rangle\langle\psi_i|\hat{A}|\psi_j\rangle \\
&= \sum_j \langle\psi_j|\rho\hat{A}|\psi_j\rangle \\
&= \mathrm{tr}(\rho\hat{A})
\end{aligned}$$

4. **Dynamics:** The dynamics of a pure state is governed by the Schrödinger Equation: $\hat{H}|\psi\rangle = i\hbar\dfrac{\partial|\psi\rangle}{\partial t}$. The dyanamics of a mixed state is governed by what is known as Liouville-von Neumann Equation: $[\hat{H}, \rho] = i\hbar\dfrac{\partial\rho}{\partial t}$, where [ ] represents the commutator.

Quantum Mechanical Correlation Functions

1. two-point correlation function for one-photon field:

An electromagnetic field in quantum mechanics is represented by a state, which, in addition to having all the labelling in the classical case, also has the labelling of the number of photons in the field. It specifies how many quanta of energy $n\hbar\omega$ does a particular state of the field has. This is unlike classical mechanics where the field can have any amount of energy. The electric field in quantum mechanics is represented by an operator $\hat{E}(\boldsymbol{r},t)$. If the field is real the operator has to be real, in the sense that the operator representing the field has to be Hermitian so that the eigenvalues are real. In situation in which the real electric field $\hat{E}(\boldsymbol{r},t)$ has a Fourier representation, we can write it through the quantization procedure as:

$$\hat{E}(\boldsymbol{r},t) = \int_{-\infty}^{\infty} \hat{\mathcal{E}}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega,$$

where $\hat{\mathcal{E}}(\boldsymbol{r},\omega)$ is the amplitude of the electric field in the frequency domain. The fact that $\hat{E}(\boldsymbol{r},t)$ is real ensures that $\hat{\mathcal{E}}^{\dagger}(\boldsymbol{r},\omega) = \hat{\mathcal{E}}(\boldsymbol{r},-\omega)$. The above field can be written as a sum of the positive-frequency and negative-frequency parts, which are directly related to the creation and annihilation operators

$$\begin{aligned}
\hat{E}(\boldsymbol{r},t) &= \int_{-\infty}^{0} \hat{\mathcal{E}}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega + \int_{0}^{\infty} \hat{\mathcal{E}}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\
\hat{E}(\boldsymbol{r},t) &= \int_{\infty}^{0} \hat{\mathcal{E}}(\boldsymbol{r},-\omega)e^{i\omega t}d(-\omega) + \int_{0}^{\infty} \hat{\mathcal{E}}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\
\hat{E}(\boldsymbol{r},t) &= \int_{0}^{\infty} \hat{\mathcal{E}}^{\dagger}(\boldsymbol{r},\omega)e^{i\omega t}d\omega + \int_{0}^{\infty} \hat{\mathcal{E}}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\
\hat{E}(\boldsymbol{r},t) &= \hat{E}^{(-)}(\boldsymbol{r},t) + \hat{E}^{(+)}(\boldsymbol{r},t)
\end{aligned}$$

We see that the field is a sum of its positive and negative frequency parts. We note that the two parts are individually not the Hermitian operators, only the sum is a Hermitian operator. The positive and negative frequency parts are connected through Hilbert transform and are mutually adjoint, that is,

$$\hat{E}^{(-)}(\boldsymbol{r},t) = \hat{E}^{(+)\dagger}(\boldsymbol{r},t).$$

It can be shown that the act of an annihilation operator on a state with n photons produces a state with $n-1$ photons. This means that if an annihilation operator acts on a state with no photons in it, it produces zero. On the other hand, the act of an creation operator on a state with n photons produces a state with $n+1$ photons. Thus it is clear that the annihilation operator is associated with the absorption of a photon whereas the creation operator is associated with the emission of a photon. In the above representation of the field operator, we note that the annihilation operator $\hat{E}^{(+)}(\boldsymbol{r},t)$ plays the same role in quantum formalism that an analytic signal does in the classical formalism. Next, we will see that the structure of the correlation functions defined in quantum theory using the creation and annihilation operators is also the same as the structure of correlation functions defined in classical theory using analytical signals.

In classical mechanics, one asks the question as to what is the intensity of a field at $(\boldsymbol{r},t)$, and for a coherent field the intensity is defined as $I(\boldsymbol{r},t) = V^*(\boldsymbol{r},t)V(\boldsymbol{r},t)$. In quantum mechanics, one cannot put the question like this. What one can ask is what is the probability that a photon from the field is absorbed at $(\boldsymbol{r},t)$. Let us now see how to calculate this quantity. Suppose $|i\rangle$ is the initial state of the field and $|f\rangle$ is the final state or a set of final states of the field in which the field goes after a photon has been absorbed. Then, the probability that the field makes this transition from the initial state to the final state is given by

$$\begin{aligned}
\sum_f |\langle f|\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle|^2 &= \sum_f \langle f|\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle^{\dagger}\langle f|\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle \\
&= \sum_f \langle i|\hat{E}^{(-)}(\boldsymbol{r},t)|f\rangle\langle f|\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle \\
&= \langle i|\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle = P(\boldsymbol{r},t)
\end{aligned}$$

This is the probability that a photon is absorbed at $(\boldsymbol{r},t)$. We see that the probability is just the expectation value of the operator $\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)$. Therefore, the operator $\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)$ can be considered as the instantaneous photon intensity operator. Also, if the state of the field is not pure then the above probability is given by an ensemble average, that is

$$\begin{aligned}
P(\boldsymbol{r},t) &= \langle\langle i|\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)|i\rangle\rangle_e \\
&= \mathrm{tr}(\rho\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)),
\end{aligned}$$

where $\rho = \{|i\rangle\langle i|\}_e$. This expression should be compared with the ensemble-averaged intensity expression that we have in classical theory, which is $\langle I(\boldsymbol{r},t)\rangle_e = \langle V^*(\boldsymbol{r},t)V(\boldsymbol{r},t)\rangle_e$. Thus, it is very clear that that the positive-frequency part of the electric field operator in quantum formalism plays the same role as does the complex analytic signal in classical formalism.

In classical formalism intensity is described in terms of energy per unit time per unit area, whereas in quantum formalism intensity is described in terms of the number of photons per unit time per unit area. At high light levels, both the descriptions are equivalent but at low light levels there are some fundamental differences which we will study in detail during the remainder of this course. Now that we know how to define the expectation value for intensity in quantum mechanics, we define the expectation value $G^{(1)}(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2,)$ for the correlation between fields at $(\boldsymbol{r_1},t_1)$ and $(\boldsymbol{r_2},t_2)$ as

$$\begin{aligned} G^{(1)}(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2,) &= \langle\langle i|\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle\rangle_e \\ &= \mathrm{tr}(\rho\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)). \end{aligned}$$

The correlation function $G^{(1)}(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2,)$ should be compared with its classical analogue, which is the cross-correlation function $\Gamma(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2,) = \langle V^*(\boldsymbol{r_1},t)V(\boldsymbol{r_2},t)\rangle_e$.

An example: one-photon interference in Michelson interferometer

In the quantum domain, we have seen that one-photon interference is described using Dirac's famous statement, "...a single photon interferes with itself." Qualitatively, interference is also described in terms of the indistinguishability argument, according to which interference takes place only if the photons in the two interference alternatives are completely indistinguishable. In terms of coherence, one says that interference takes place only if the two alternatives are coherent with each other. Let us review the quantum treatment of temporal one-photon interference using a Michaelson interferometer as shown in the figure below.

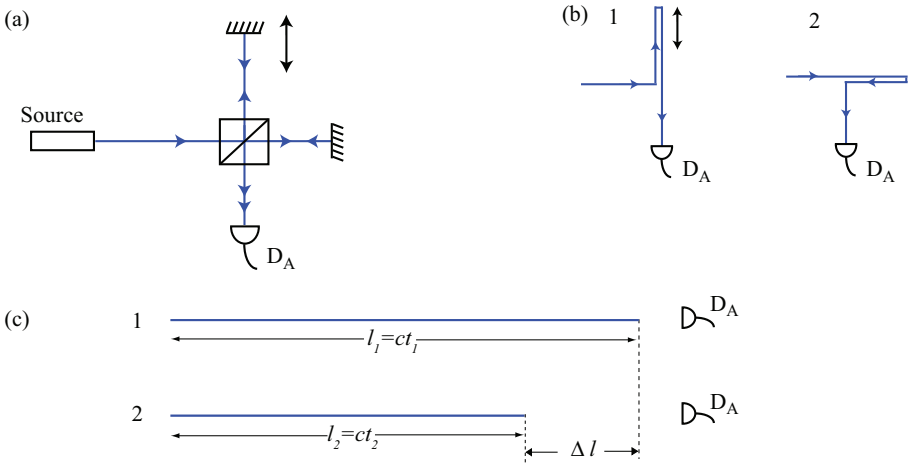


FIG. 11: Temporal one-photon interference.

Just as in the optical-field case we worked with the analytic signals, here we work with the positive frequency part of the field $\hat{E}^{(+)}(\boldsymbol{r},t)$. So, the Electric field operator $\hat{E}^{(+)}(\boldsymbol{r},t)$ at the detector position $\boldsymbol{r}$ at time t is given by

$$\hat{E}^{(+)}(\boldsymbol{r},t) = k_1\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_1) + k_2\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2).$$

Here k_1 and k_2 are constants that depend on the splitting ratio of the beam splitter and $\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_1)$ and $\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2)$ are the electric field operators right before the beam splitter at point $\boldsymbol{r}_0$ at times $t-t_1$ and $t-t_2$, respectively. Now, if the state of the field right before the beam splitter is given by $|\psi\rangle$ then the probability per

unit time that a photon is detected at $\boldsymbol{r}$ at time t can be written as

$$\begin{aligned}
P(\boldsymbol{r},t) &= \langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r},t)\hat{E}^{(+)}(\boldsymbol{r},t)|\psi\rangle\rangle \\
&= |k_1|^2\langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_1)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_1)|\psi\rangle\rangle + |k_2|^2\langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_2)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2)|\psi\rangle\rangle \\
&\qquad + 2\mathrm{Re}k_1^*k_2\langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_1)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2)|\psi\rangle\rangle \\
&= |k_1|^2P(\boldsymbol{r}_0,t-t_1) + |k_2|^2P(\boldsymbol{r}_0,t-t_2) + 2|k_1||k_2|G^{(1)}(\boldsymbol{r}_0,t-t_1;\boldsymbol{r}_0,t-t_2). \qquad (83)
\end{aligned}$$

where we have

$$\begin{aligned}
P(\boldsymbol{r}_0,t-t_1) &= \langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_1)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_1)|\psi\rangle\rangle \\
P(\boldsymbol{r}_0,t-t_2) &= \langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_2)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2)|\psi\rangle\rangle \\
G^{(1)}(\boldsymbol{r}_0,t-t_1;\boldsymbol{r}_0,t-t_2) &= \langle\langle\psi|\hat{E}^{(-)}(\boldsymbol{r}_0,t-t_1)\hat{E}^{(+)}(\boldsymbol{r}_0,t-t_2)|\psi\rangle\rangle
\end{aligned}$$

Equation (83) is the one-photon interference law. We find that it is very analogous to the interference law for optical fields as expressed in Eq. (31). If the one-photon field is stationary, in the sense that $P(\boldsymbol{r}_0,t-t_1)$ is independent of t_1 and $G^{(1)}(\boldsymbol{r}_0,t-t_1;\boldsymbol{r}_0,t-t_2)$ only depends on $\tau = t_1 - t_2$, then we see that as a function of τ the probability $P(\boldsymbol{r},t)$ shows interference and as soon as $\tau \gg \tau_c$, where τ_c is some coherence time, the interference fringes gets washed out. So,the condition for interference is that $\tau \ll \tau_c$. We find that the temporal interference effects produced by the one-photon field is quite analogous to the temporal interference effects produced by the wave-fields as studied in the previous lectures. Similar analogy can be established for interference effects in space, angle, and polarization degrees of freedom, and the entire optical coherence theory effects studied so far can be worked out in terms of the correlation functions for one-photon fields.

2. four-point correlation function for two-photon field:

Next, suppose we have two separate detectors and we ask the question: what is the probability per unit (time)^2 that one photon is detected at $\boldsymbol{r_1}$ at time t_1 and the other photon is detected at $\boldsymbol{r_2}$ at time t_2. Calculating in a similar manner as above, we can show that this probability in the case of a pure state is given by

$$\begin{aligned}
\sum_f |\langle f|\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle|^2 &= \sum_f \langle f|\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle^\dagger\langle f|\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle \\
&= \sum_f \langle i|\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)|f\rangle\langle f|\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle \\
&= \langle i|\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle \equiv P(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2)
\end{aligned}$$

The above probability is called the coincident probability $P(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2)$. In situations in which the field is in a mixed state, the probability needs to be ensemble averaged and thus becomes

$$\begin{aligned}
\langle P(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2)\rangle_e &= \langle\langle i|\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)|i\rangle\rangle_e \\
&= \mathrm{tr}\left(\rho\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)\hat{E}^{(+)}(\boldsymbol{r_1},t_1)\hat{E}^{(+)}(\boldsymbol{r_2},t_2)\right)
\end{aligned}$$

Now that we have photons being detected at two different space-time points, one can study correlations that involves four space-time points, which is defined in the following way:

$$\begin{aligned}
G^{(2)}(\boldsymbol{r_1},t_1;\boldsymbol{r_2},t_2;\boldsymbol{r_3},t_3;\boldsymbol{r_4},t_4) &= \langle\langle i|\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)\hat{E}^{(+)}(\boldsymbol{r_3},t_3)\hat{E}^{(+)}(\boldsymbol{r_4},t_4)|i\rangle\rangle_e \\
&= \mathrm{tr}\left[\rho\hat{E}^{(-)}(\boldsymbol{r_1},t_1)\hat{E}^{(-)}(\boldsymbol{r_2},t_2)\hat{E}^{(+)}(\boldsymbol{r_3},t_3)\hat{E}^{(+)}(\boldsymbol{r_4},t_4)\right].
\end{aligned}$$

In general, one-photon quantum treatment has exact analog in the classical theory and all the experimental results can be reproduced using the classical approach as well, except for some interpretations issues. However, in studying two-photon fields and above, we find that the classical theory is simply inadequate and cannot describe all the effects that are described by quantum theory and which are observed in real experiments. One such effect is known as quantum entanglement.

LECTURE # 13

Intro to quantum entanglement

In 1935 A. Einstein, B. Podolsky, and N. Rosen recognized a spooky feature in the quantum description of physical reality [1]. This feature, now known as entanglement, was originally called by Schrödinger "Verschränkung," which implies the existence of global states of a composite system that cannot be described as a product of the states of the individual subsystems [2]. For an extensive review of the subject of quantum entanglement, see the article by Horodecki *et al.* [3]. Einstein, Podolsky and Rosen were explicitly considering a system of two entangled particles, which had once interacted in the past but had no interactions thereafter. They showed that in such a system, by measuring either the position or the momentum of one of the particles, either the position or the momentum of the other particle can be predicted with complete certainty without in any way disturbing the other particle. But, since position and momentum are non-commuting observables, they cannot be measured simultaneously and thus cannot be regarded as simultaneous elements of reality of a system. So they concluded that in entangled two-particles systems, the reality of position and momentum of the second system depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. On noting this apparent non-locality in the quantum theory, Einstein, Podolsky and Rosen argued that the quantum-mechanical description of physical reality is not complete and that it should be supplemented by postulating the existence of "hidden variables," the specifications of which will predetermine the result of measuring any observable of the system. EPR's inherent belief was that the predictions of quantum mechanics are correct but only as a consequence of the statistical distribution of hidden variables. Until mid 1990s, studies related to entanglement were mostly centered at probing the foundations of quantum mechanics. However, in recent years entanglement is beginning to be seen as a resource that could also be used for many useful applications. Some of these proposed applications include quantum cryptography [4], quantum dense coding [5] quantum teleportation [6], entanglement swapping [7] and quantum lithography [8]. We are now going to study in detail one process that produces two-photon entangled states as conceived by EPR. The process is called parametric down-conversion. It is a non-linear optical process. So we will start with reviewing the very basics on nonlinear optics and study the down-conversion process in detail.

Basics of Nonlinear Optics

When we are studying electromagnetic fields in a medium, the electric field by itself cannot describe the entire dynamics. For the full description one has to include the concept of polarization as well. The reason is as follows: When an atom is placed in an external electric field, the electrons in the atom experience a force due to this field. The electrons, which are bound in the Coulomb field of the nucleus, also experience a restoring force by the nucleus. As a result of these two competing forces, the electrons get displaced from their equilibrium positions and the atom acquires a net dipole moment. The dipole moment per unit volume which is called the polarization $\boldsymbol{P}$ depends on the strength of the electric field. The Maxwell's equation in a medium take the following form:

$$\begin{aligned}
&\nabla \cdot \boldsymbol{D} = \rho; \\
&\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \\
&\nabla \cdot \boldsymbol{B} = 0; \\
&\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}.
\end{aligned}$$

Here $\boldsymbol{E}$ is the electric field and $\boldsymbol{D}$ is the electrical displacement vector and is given by $\boldsymbol{D} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P}$, where $\boldsymbol{P}$ is called the polarization, which is the dipole moment per unit volume. In the case of vacuum, when there is no medium, we have

$$\begin{aligned}
&\boldsymbol{P} = 0 \\
\text{and} \qquad &\boldsymbol{D} = \epsilon_0 \boldsymbol{E}.
\end{aligned}$$

In the case of linear optics, that is, when the applied field strength is small, the restoring force exerted on the electron is proportional to its displacement from the equilibrium position, and the atom is modelled as a harmonic oscillator.

The induced polarization $\boldsymbol{P}(\boldsymbol{r},t)$ depends linearly on the field strength $\boldsymbol{E}$ and is given by [15]

$$\begin{aligned} &\boldsymbol{P} = \epsilon_0 \chi^{(1)} \boldsymbol{E} \\ \text{and} \qquad &\boldsymbol{D} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P} = \epsilon_0 \boldsymbol{E} + \epsilon_0 \chi^{(1)} \boldsymbol{E} \end{aligned}$$

where $\chi^{(1)}$ is the linear susceptibility and is in general a tensor, and ϵ_0 is the permittivity of the free space. In the case of nonlinear optics, that is, when the applied field strength is strong, the restoring force exerted on the electron is no longer proportional to its displacement from the equilibrium position. As a result, the simple linear dependence of the induced polarization on the applied electric field is no longer valid [16]. The induced polarization $\boldsymbol{P}$ gets contributions that are not only linear in electric field strength, but also bilinear, trilinear, etc, and is given by [15]

$$\boldsymbol{P} = \epsilon_0 \chi^{(1)} \boldsymbol{E} + \epsilon_0 \chi^{(2)} \boldsymbol{E}^2 + \epsilon_0 \chi^{(3)} \boldsymbol{E}^3 + \cdots \tag{84}$$

$$\text{and} \qquad \boldsymbol{D} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P} = \epsilon_0 \boldsymbol{E} + \epsilon_0 \chi^{(1)} \boldsymbol{E} + \epsilon_0 \chi^{(2)} \boldsymbol{E}^2 + \epsilon_0 \chi^{(3)} \boldsymbol{E}^3 + \cdots \tag{85}$$

where $\chi^{(2)}$ and $\chi^{(3)}$ are known as the second-order and third-order nonlinear optical susceptibilities, respectively. $\boldsymbol{P}^{(2)} = \epsilon_0 \chi^{(2)} \boldsymbol{E}^2$ is the second-order nonlinear polarization, etc. The first-order susceptibility $\chi^{(1)}$ is usually of the order of unity. The second-order nonlinear susceptibility $\chi^{(2)}$ is usually of the order of 10^{-12} m/V. The third-order nonlinear susceptibility $\chi^{(3)}$ is usually of the order of 10^{-24} $\mathrm{m}^2/\mathrm{V}^2$. We now derive the wave equation inside the medium, assuming that $\rho = 0$, $\boldsymbol{J} = 0$, and that the material is magnetically isotropic, that is, $\boldsymbol{B} = \mu_0 \boldsymbol{H}$. The Maxwell's equation in the medium can now be written as

$$\begin{aligned} &\nabla \cdot \boldsymbol{D} = 0, \\ &\nabla \times \boldsymbol{E} = -\mu_0 \frac{\partial \boldsymbol{H}}{\partial t}, \\ &\nabla \cdot \boldsymbol{H} = 0, \\ &\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t}. \end{aligned}$$

We can now write

$$\begin{aligned} \nabla \times \nabla \times \boldsymbol{E} &= -\mu_0 \frac{\partial}{\partial t} (\nabla \times \boldsymbol{H}), \\ \nabla \times \nabla \times \boldsymbol{E} &= -\mu_0 \frac{\partial^2}{\partial t^2} \boldsymbol{D}, \\ \nabla \times \nabla \times \boldsymbol{E} &= -\mu_0 \frac{\partial^2}{\partial t^2} (\epsilon_0 \boldsymbol{E} + \boldsymbol{P}), \\ \nabla (\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} + \mu_0 \epsilon_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} &= -\mu_0 \frac{\partial^2 \boldsymbol{P}}{\partial t^2}, \end{aligned}$$

In linear optics $\nabla \cdot \boldsymbol{E} = 0$ and so the first term of the last equation also vanishes. However, this is the not true for the nonlinear optics. However, it can be shown that even for nonlinear optics the contribution due to the first term of the last equation is negligibly small and therefore we can write the last equation as

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \boldsymbol{P}}{\partial t^2}.$$

This is the wave-equation for the electric field in a nonlinear optical medium. We see that in vacuum $\boldsymbol{P} = 0$ and we retrieve the usual wave equation in the vacuum. It is often convenient to write polarization as $\boldsymbol{P} = \boldsymbol{P}^{(1)} + \boldsymbol{P}^{\mathrm{NL}}$. The

wave equation in this case takes the following form

$$\begin{aligned}
\nabla^2\boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2} &= \mu_0\frac{\partial^2(\boldsymbol{P}^{(1)}+\boldsymbol{P}^{\mathrm{NL}})}{\partial t^2},\\
\nabla^2\boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2(\boldsymbol{E}+\boldsymbol{P}^{(1)}/\epsilon_0)}{\partial t^2} &= \mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}\\
\nabla^2\boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2(\boldsymbol{E}+\chi^{(1)}\boldsymbol{E})}{\partial t^2} &= \mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}\\
\nabla^2\boldsymbol{E} - \frac{(1+\chi^{(1)})}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2} &= \mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}\\
\nabla^2\boldsymbol{E} - \frac{\epsilon^{(1)}}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2} &= \mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2}\\
\nabla^2\boldsymbol{E} - \frac{n^2}{c^2}\frac{\partial^2\boldsymbol{E}}{\partial t^2} &= \mu_0\frac{\partial^2\boldsymbol{P}^{\mathrm{NL}}}{\partial t^2},
\end{aligned}$$

where $n = \sqrt{\epsilon^{(1)}} = \sqrt{1+\chi^{(1)}}$ is called the refractive index of the medium. The last equation can also be seen as a homogeneous wave equation with an inhomogeneous driving term.

Second-Order Nonlinear Optical Effects

Now, let us look at some of the very common second-order nonlinear optical effects. We will do this more for conceptual clarity and therefore will assume that both the electric field and the polarization are scalar quantities

1. **Second Harmonic Generation (SHG)**: Let us take the electric field to be a real, monochromatic field such that we can write it as

   $$E(t) = E_0e^{-i\omega t} + E_0^*e^{i\omega t}$$

   The second-order polarization then becomes

   $$P^{(2)}(t) = \epsilon_0\chi^{(2)}E^2(t) = \epsilon_0\chi^{(2)}(E_0e^{-i\omega t} + E_0^*e^{i\omega t})^2 = \epsilon_0\chi^{(2)}\left[|E_0|^2 + |E_0|^2 + 2|E_0|^2(e^{-2i\omega t} + e^{2i\omega t})\right].$$

   We see that although the applied field has only a single frequency ω the resultant polarization consists of a static field $2|E_0|^2$ around the nonlinear medium and a real field at frequency 2ω. This frequency was not there but has been produced through the nonlinear interaction. This generation of frequency at double the applied frequency is called the second-harmonic generation. The exact dynamics of different fields can be studied by using the nonlinear wave equation as derived above.

2. **Sum and Difference Frequency Generation**: Now, let us take a real electric field that consists of two different frequencies, that is,

   $$E(t) = E_1e^{-i\omega_1 t} + E_2e^{-i\omega_2 t} + \text{c.c.}$$

   The second-order polarization in this case is

   $$\begin{aligned}
   P^{(2)}(t) = \epsilon_0\chi^{(2)}E^2(t) &= 2\epsilon_0\chi^{(2)}\left(E_1E_2^* + E_1^*E_2\right)\\
   &\quad + \epsilon_0\chi^{(2)}\left[|E_1|^2e^{-2i\omega_1 t} + |E_2|^2e^{-2i\omega_2 t} + 2E_1E_2e^{-i(\omega_1+\omega_2)t} + 2E_1E_2^*e^{-i(\omega_1-\omega_2)t} + \text{c.c.}\right].
   \end{aligned}$$

   In this case, we can clearly see the generations of the sum frequency $\omega_1+\omega_2$, the difference frequency $\omega_1-\omega_2$, as well as the second harmonic frequencies, $2\omega_1$ and $2\omega_2$.

   In order to have the difference frequency generation at $\omega_3 = \omega_2 - \omega_1$ from the field at frequency ω_1, the field at ω_2 need not be present. If the field at frequency ω_2 is present then the difference frequency generation is stimulated by the field at frequency ω_2. The process has a very high efficiency and is called the optical parametric amplification/oscillation. However, even in the absence of ω_2, one can have difference frequency generation. The reason is that vacuum mode at frequency ω_2 is always present. However, since this process is stimulated by the

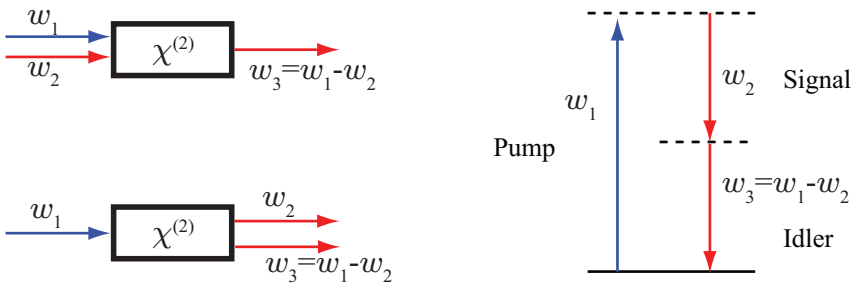


FIG. 12: Difference Frequency Generation.

vacuum mode, the efficiency of the process is very low and is called parametric down-conversion. Parametric refers to the fact that the energy remains conserved during the process. Non-parametric processess are non-conserving process and we will not be considering such processess. The explanation based on vacuum modes can only be proved using quantum formalism. The existence of vacuum modes have already been demonstrated through the casimir-force experiment. We see that for a medium to show the second-order nonlinear effects it has to have a non-zero value for the second-order nonlinear susceptibility. In fact, the second-order nonliear polarization can occur only in crystals that are non-centrosymmetric, that is, in crystals that have no inversion symmetry. To prove this suppose a medium that is centro-symmetric has a finite value of $\chi^{(2)}$, then we have to have

$$P = \epsilon_0 \chi^{(2)} E^2(t)$$
$$\text{and} \qquad -P = \epsilon_0 \chi^{(2)} (-E)^2$$

The above two equations imply that $\chi^{(2)} = 0$. Therefore, we have that only a medium that is non-centrosymmetric can show second order nonlinear effects including parametric down-conversions.

Parametric down-conversion

As outlined above, one studies the third and higher order nonlinear optical effects in a similar manner but we will not go in there. We will now concentrate on parametric down-conversion since it is one process that produces entangled two-photon field. Let us go back to the expression for polarization, we have

$$\boldsymbol{P}(\boldsymbol{r},t) = \epsilon_0 \chi^{(1)} \boldsymbol{E}(\boldsymbol{r},t) + \epsilon_0 \chi^{(2)} \boldsymbol{E}^2(\boldsymbol{r},t) + \epsilon_0 \chi^{(3)} \boldsymbol{E}^3(\boldsymbol{r},t) + \cdots, \tag{86}$$
$$\text{and} \qquad \boldsymbol{D}(\boldsymbol{r},t) = \epsilon_0 \boldsymbol{E}(\boldsymbol{r},t) + \boldsymbol{P} \tag{87}$$

Here, we are explicitly writing the space and time coordinates. Now, the total (electrical plus magnetic) energy density of the electromagnetic field is given by

$$\boldsymbol{D}(\boldsymbol{r},t) \cdot \boldsymbol{E}(\boldsymbol{r},t) = \Big[\epsilon_0 \boldsymbol{E}(\boldsymbol{r},t) + \epsilon_0 \chi^{(1)} \boldsymbol{E}(\boldsymbol{r},t)\Big] \cdot \boldsymbol{E}(\boldsymbol{r},t) + \Big[\epsilon_0 \chi^{(2)} \boldsymbol{E}^2(\boldsymbol{r},t)\Big] \cdot \boldsymbol{E}(\boldsymbol{r},t) + \cdots \tag{88}$$

The first term is the contribution to the electromagnetic energy due to the linear term. The second term is the contribution due to the second-order nonlinear term, etc. Since we are interested in second-order nonlinear optical effects, we will look at the second-order contribution to the energy which is

$$H(t) = \int_{\mathcal{V}} \boldsymbol{P}^{(2)}(\boldsymbol{r},t) \cdot \boldsymbol{E}(\boldsymbol{r},t) d^3\boldsymbol{r} \tag{89}$$
$$= \epsilon_0 \int_{\mathcal{V}} \chi^{(2)} \boldsymbol{E}^2(\boldsymbol{r},t) \cdot \boldsymbol{E}(\boldsymbol{r},t) d^3\boldsymbol{r} \tag{90}$$

where the integration extends over the volume $\mathcal{V}$ of the nonlinear medium. This is the general expression for the contribution to the total energy due to second-order nonlinear optical effects. However, if we are interested only in

parametric down-conversion for the scalar field, the above expression for the total energy takes the following form.

$$H(t) = \int_{\mathcal{V}} P^{(2)}(\boldsymbol{r},t)E(\boldsymbol{r},t)d^3\boldsymbol{r} \tag{91}$$

$$= \epsilon_0 \int_{\mathcal{V}} \chi^{(2)} E_p(\boldsymbol{r},t)E_s(\boldsymbol{r},t)E_i(\boldsymbol{r},t)d^3\boldsymbol{r} \tag{92}$$

In the quantum formalism the energy takes the form of the interaction Hamiltonian operator and is given by

$$\hat{H}(t) = \epsilon_0 \int_{\mathcal{V}} \chi^{(2)} \hat{E}_p(\boldsymbol{r},t)\hat{E}_s(\boldsymbol{r},t)\hat{E}_i(\boldsymbol{r},t)d^3\boldsymbol{r}. \tag{93}$$

where p, s and i stand for the pump, signal and idler, respectively, and where $\hat{E}_p(\boldsymbol{r},t)$ is the quantized electric field operator corresponding to the pump field, etc. It can be written as a sum of negative and positive frequency terms as

$$\hat{E}_p(\boldsymbol{r},t) = \hat{E}_p^{(+)}(\boldsymbol{r},t) + \hat{E}_p^{(-)}(\boldsymbol{r},t), \tag{94}$$

etc. The interaction Hamiltonian can thus be written as

$$\hat{H}(t) = \epsilon_0 \int_{\mathcal{V}} \chi^{(2)} \left[\hat{E}_p^{(+)}(\boldsymbol{r},t) + \hat{E}_p^{(-)}(\boldsymbol{r},t)\right] \left[\hat{E}_s^{(+)}(\boldsymbol{r},t) + \hat{E}_s^{(-)}(\boldsymbol{r},t)\right] \left[\hat{E}_i^{(+)}(\boldsymbol{r},t) + \hat{E}_i^{(-)}(\boldsymbol{r},t)\right] d^3\boldsymbol{r}. \tag{95}$$

The resulting expression for the interaction Hamiltonian $\hat{H}(t)$ is a sum of eight different terms with all possible combinations of the three fields. However, there are only two terms, $E_p^{(+)}E_s^{(-)}E_i^{(-)}$ and $E_p^{(-)}E_s^{(+)}E_i^{(+)}$, that lead to energy conserving processes and thus contribute appreciably to the down-conversion process. The contributions due to the other six terms, such as $E_p^{(-)}E_s^{(-)}E_i^{(-)}$ and $E_p^{(+)}E_s^{(+)}E_i^{(+)}$, get averaged out when the interaction Hamiltonian $\hat{H}(t)$ is integrated over time. Therefore, we neglect the contributions due to these other terms; neglecting these contributions is equivalent to making the rotating-wave approximation as in the case of treating atomic absorption and emission processes (see Ref. [17], Section 2.3). We note that these approximations hold only for second-order processes and that for the higher-order processes the non-energy-conserving terms may lead to important contributions. The effective interaction Hamiltonian for the process of parametric down-conversion can then be given by the following simplified form:

$$\hat{H}(t) = \epsilon_0 \int_{\mathcal{V}} d^3\boldsymbol{r}\chi^{(2)} \hat{E}_p^{\ (+)}(\boldsymbol{r},t)\hat{E}_s^{\ (-)}(\boldsymbol{r},t)\hat{E}_i^{\ (-)}(\boldsymbol{r},t) + \text{H.c.}, \tag{96}$$

where a quantized field is expressed as

$$\hat{E}^{(+)}(\boldsymbol{r},t) = \int A_k \hat{a}_{\boldsymbol{k}}(t) e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} d^3\boldsymbol{k}. \tag{97}$$

The constraints of energy and momentum conservation in down-conversion require that the sum of the energies of the signal and idler photons be equal to the energy of the pump photon and that the sum of the momenta of the signal and idler photons be equal to the momentum of the pump photon. These constraints due to conservation laws render the two photons entangled in their time-energy, position-momentum, and angular-position–OAM degrees of freedom. The phase-matching conditions in PDC can be adjusted so that both photons come out in a direction collinear with the pump, in which case it is known as the collinear phase-matching. Alternatively, the two photons come out in two separate directions, in which case it is known as the non-collinear phase-matching. Phase-matching conditions can also be adjusted so that the polarizations of the two photons are the same (type-I down-conversion) or orthogonal to each other (type-II down-conversion).

LECTURE # 14

Two-Photon State Produced by Parametric Down-Conversion: Temporal

In the present section, we use a simplified form of the interaction Hamiltonian of Eq. (96). We assume the transverse area of the interacting part of the nonlinear crystal to be very large. We also assume that the emission directions of the signal and idler fields are fixed such that the transverse wave-vectors of the pump field ($\boldsymbol{q}_p$) is equal to the sum of the transverse wave-vectors of the signal ($\boldsymbol{q}_s$) and idler ($\boldsymbol{q}_i$) fields, that is, $\boldsymbol{q}_p = \boldsymbol{q}_s + \boldsymbol{q}_i$, where $\boldsymbol{k}_p \equiv (\boldsymbol{q}_p, k_{pz})$, $\boldsymbol{k}_s \equiv (\boldsymbol{q}_s, k_{sz})$ and $\boldsymbol{k}_i \equiv (\boldsymbol{q}_i, k_{iz})$ are the wave-vectors of the pump, signal and idler fields, respectively. With the above assumptions, the interaction Hamiltonian $\hat{H}(t')$ for parametric down-conversion takes the following form:

$$\hat{H}(t) = \epsilon_0 \int_{-L}^{0} dz \chi^{(2)} \hat{E}_p^{\ (+)}(z,t) \hat{E}_s^{\ (-)}(z,t) \hat{E}_i^{\ (-)}(z,t) + \text{H.c.} \tag{98}$$

where L is the thickness of the nonlinear crystal. The three electric fields are given as

$$\hat{E}_p^{\ (+)}(z,t) = \int_0^\infty A_p d\omega_p V(\omega_p) e^{i[k_{pz}(\omega_p)z - \omega_p t]} e^{i(\omega_p \tau_p + \phi_p)}, \tag{99}$$

$$\hat{E}_s^{\ (-)}(z,t) = \int_0^\infty A_s^* d\omega_s \hat{a}_s^\dagger(\omega_s) e^{i[\omega_s t - k_{sz}(\omega_s)z]}, \tag{100}$$

$$\hat{E}_i^{\ (-)}(z,t) = \int_0^\infty A_i^* d\omega_i \hat{a}_i^\dagger(\omega_i) e^{i[\omega_i t - k_{iz}(\omega_i)z]}, \tag{101}$$

where k_{jz} is the z-component of wave-vector $\boldsymbol{k}_j$, with $j = p, s, i$. A_j is a frequency dependent quantity. However, it varies very slowly within the frequency range of interest for most down-conversion experiments; and therefore, it can be taken outside the integral. The pump field has been assumed to be very strong and will be treated classically. The strength of the pump field at frequency ω_p is represented by $V(\omega_p)$. τ_p represents the time taken by the pump photon in travelling the optical path length l_p between the laser and the nonlinear crystal, while ϕ_p is the phase other than the dynamical one accumulated during this travel. Using Eqs. (99), (100) and (101), we write Eq. (98) as

$$\hat{H}(t) = A_p A_s^* A_i^* \epsilon_0 \chi^{(2)} \int_{-L}^{0} dz \iiint_0^\infty d\omega_p d\omega_s d\omega_i \hat{a}_s^\dagger(\omega_s) \hat{a}_i^\dagger(\omega_i) V(\omega_p) \times e^{i[k_{pz}(\omega_p) - k_{sz}(\omega_s) - k_{iz}(\omega_i)]z} e^{i(\omega_s + \omega_i - \omega_p)t} e^{i(\omega_p \tau_p + \phi_p)} + \text{H.c.} \tag{102}$$

The second-order nonlinear susceptibility $\chi^{(2)}$ has been assumed to be independent of frequency over the range of interest. The state of the down-converted field at time $t = -t_{\text{int}}$ is given by $|\psi(-t_{\text{int}})\rangle = |\text{vac}\rangle_s |\text{vac}\rangle_i$, which is a

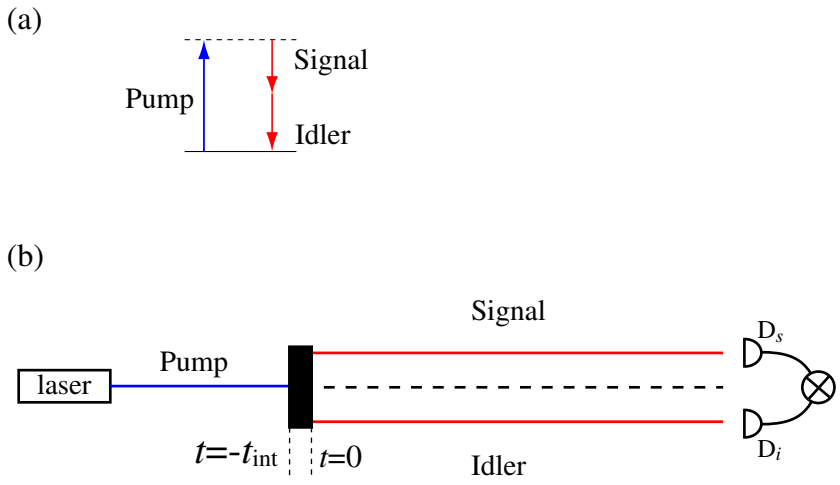


FIG. 13: Spontaneous Parametric Down-conversion.

vacuum state with no photons in either the signal or the idler mode. The state $|\psi(0)\rangle$ of the two-photon field at $t = 0$ is then calculated using the Schrödinger equation to be

$$|\psi(0)\rangle = \exp\left[\frac{1}{i\hbar}\int_{-t_{\rm int}}^{0} dt\hat{H}(t)\right]|\psi(-t_{\rm int})\rangle. \tag{103}$$

The parametric interaction is assumed to be very weak so that the state in Eq. (103) can be approximated by the first two terms of a perturbative expansion. The first term is simply the initial vacuum state. The second term $|\psi_{\rm tp}\rangle$ is calculated by substituting from Eq. (102) into Eq. (103) to be

$$|\psi_{\rm tp}\rangle = \frac{A_p A_s^* A_i^* \epsilon_0 \chi^{(2)}}{i\hbar}\int_{-t_{\rm int}}^{0} dt\int_{-L}^{0} dz\iiint_{0}^{\infty} d\omega_p d\omega_s d\omega_i V(\omega_p) \times e^{i[k_{pz}(\omega_p)-k_{sz}(\omega_s)-k_{iz}(\omega_i)]z} e^{i(\omega_s+\omega_i-\omega_p)t} e^{i(\omega_p\tau_p+\phi_p)}\hat{a}_s^\dagger(\omega_s)\hat{a}_i^\dagger(\omega_i)|{\rm vac}\rangle_s|{\rm vac}\rangle_i. \tag{104}$$

We note that although the interaction Hamiltonian $\hat{H}(t)$ in Eq. (102) contains two separate terms, including the Hermitian conjugate, the two-photon state $|\psi_{\rm tp}\rangle$ in Eq. (104) contains only one term. This is due to the fact that the operator $\hat{a}_s(\omega_s)\hat{a}_i(\omega_i)$ in the Hermitian conjugate term adds no contribution to the generated two-photon state when it acts on the vacuum state $|{\rm vac}\rangle_s|{\rm vac}\rangle_i$.

The interaction time $t_{\rm int}$ is taken to be much longer than the time scale over which down-conversion takes place. Therefore, both limits of the time integration in Eq. (104) is extended to infinity [18, 19]. Carrying out the time integration then yields

$$|\psi_{\rm tp}\rangle = A\int_{-L}^{0} dz\iiint_{0}^{\infty} d\omega_p d\omega_s d\omega_i V(\omega_p)\delta(\omega_s+\omega_i-\omega_p)e^{i[k_{pz}(\omega_p)-k_{sz}(\omega_s)-k_{iz}(\omega_i)]z}e^{i(\omega_p\tau_p+\phi_p)}|\omega_s\rangle_s|\omega_i\rangle_i. \tag{105}$$

Here, we have absorbed all the constant factors into A. Next, we evaluate the ω_p-integral and obtain

$$|\psi_{\rm tp}\rangle = A\int_{-L}^{0} dz\iint_{0}^{\infty} d\omega_s d\omega_i V(\omega_s+\omega_i)e^{i[k_{pz}(\omega_p)-k_{sz}(\omega_s)-k_{iz}(\omega_p-\omega s)]z}e^{i[(\omega_s+\omega_i)\tau_p+\phi_p]}|\omega_s\rangle_s|\omega_i\rangle_i. \tag{106}$$

By rearranging the above equation, we obtain the following expression for the two-photon state $|\psi_{\rm tp}\rangle$ at the exit face of the crystal:

$$|\psi_{\rm tp}\rangle = A\iint_{0}^{\infty} d\omega_s d\omega_i V(\omega_s+\omega_i)\Phi(\omega_s,\omega_i)e^{i[(\omega_s+\omega_i)\tau_p+\phi_p]}|\omega_s\rangle_s|\omega_i\rangle_i, \tag{107}$$

where

$$\Phi(\omega_s,\omega_i) = \int_{-L}^{0} dz e^{i[k_{pz}(\omega_p)-k_{sz}(\omega_s)-k_{iz}(\omega_i)]z} \tag{108}$$

is called the phase-matching function. The exact form of the phase-matching function $\Phi(\omega_s, \omega_p - \omega_s)$ depends on the nonlinear crystal parameters and the type of down-conversion (type-I or type-II). Let's make the change of variables: $\omega_p = \omega_s + \omega_i$ and $\omega_d = (\omega_s - \omega_i)/2$ such that $d\omega_s d\omega_i \to d\omega_p d\omega_d$. We can now write the above state as

$$|\psi_{\rm tp}\rangle = A\iint_{0}^{\infty} d\omega_p d\omega_d V(\omega_p)\Phi\left(\frac{\omega_p}{2}+\omega_d, \frac{\omega_p}{2}-\omega_d\right)e^{i(\omega_p\tau_p+\phi_p)}\left|\frac{\omega_p}{2}+\omega_d\right\rangle_s\left|\frac{\omega_p}{2}-\omega_d\right\rangle_i, \tag{109}$$

In the special case in which the incoming pump is a monochromatic field, that is, $V(\omega_p) = V_0\delta(\omega_0 - \omega_p)$, we get

$$|\psi_{\rm tp}\rangle = AV_0 e^{i(\omega_0\tau_p+\phi_p)}\int_{0}^{\infty} d\omega_d\Phi(\omega_d)\left|\frac{\omega_0}{2}+\omega_d\right\rangle_s\left|\frac{\omega_0}{2}-\omega_d\right\rangle_i, \tag{110}$$

We note that the above two-photon wavefunction cannot be written as a product of the wavefunctions of the signal and idler photon. This is referred to as the inseparability. The inseparability of a two-particle wavefunction in what is meant by two-particle entanglement. We find that the above two-particle state is inseparable in the frequency basis, therefore the above state is a two-particle entangled state in time-energy degree of freedom.

Temporal two-photon interference

Now that we know how to produce an entangled state and also what are the relevant quantities for detection. We will now attempt to describe two-photon interference and by that try to explore two-photon coherence. As in the case of one-photon, let us start with the two-photon interference in the temporal domain. We will try to describe two-photon interference experiment using the same intuition as that of one-photon interference. Two-photon interference effects are observed with the entangled two-photon fields and they can be described in terms of "a two-photon interfering with itself." In terms of coherence, one can say that two-photon interference occurs as long as the two two-photon alternatives are coherent with each other.

In a two-photon interference experiment, the two-photon field is made to go through two alternative pathways. The fields in the two alternative pathways are then combined, and the interference is observed in the coincidence count rate of two detectors. We begin by representing a general two-photon two-alternative interference experiment by the two-photon path diagrams of Fig. 14. We consider only the polarization-independent, temporal two-photon interference effects, assuming perfect spatial coherence. We also assume that the frequency bandwidth of the pump field is much smaller than that of the signal and idler fields. In Fig. 14, alternatives 1 and 2 are the two pathways by which a pump photon is down-converted and the down-converted signal and idler photons are detected in coincidence at single-photon detectors D_s and D_i. Two-photon interference is observed in the coincidence count rate of detectors D_s and D_i as long as the two alternatives are coherent, i.e., indistinguishable from each other. We adopt the convention that a signal photon is the one that reaches detector D_s and that an idler photon is the one that reaches detector D_i. In a two-photon interference experiment, these alternative pathways can be introduced by using beam splitters [20, 21], by passing the pump beam twice through a crystal [22], or even by using two different crystals [23]. In Fig. 14, l denotes the optical path length traveled by a photon and ϕ stands for phases other than the dynamic phase, such as the phase acquired due to reflections, geometric phase [24, 25], etc. Thus l_{s1} denotes the path length traveled by the signal photon in alternative 1, etc. For every optical path length traveled, the corresponding time elapsed is denoted by $\tau = l/c$. Thus τ_{s1} represents the time taken in traveling the distance l_{s1}. The various path-lengths and phases are used to define two length parameters and one phase parameter as follows:

$$\begin{aligned}
\Delta L &\equiv l_1 - l_2 = \left(\frac{l_{s1} + l_{i1}}{2} + l_{p1}\right) - \left(\frac{l_{s2} + l_{i2}}{2} + l_{p2}\right), \\
\Delta L' &\equiv l'_1 - l'_2 = (l_{s1} - l_{i1}) - (l_{s2} - l_{i2}), \\
\Delta\phi &\equiv \phi_1 - \phi_2 = (\phi_{s1} + \phi_{i1} + \phi_{p1}) - (\phi_{s2} + \phi_{i2} + \phi_{p2}).
\end{aligned} \tag{111}$$

Here $l_{1(2)}$ and $l'_{1(2)}$ are the two-photon path length and the two-photon path asymmetry-length for alternative 1(2); they are also referred to as the biphoton path length and the biphoton path-asymmetry length, respectively [26]. In a particular alternative, the two-photon path length is defined to be the mean of the optical path lengths traveled by the signal and idler photons added to the optical path length traveled by the pump photon. The two-photon path-

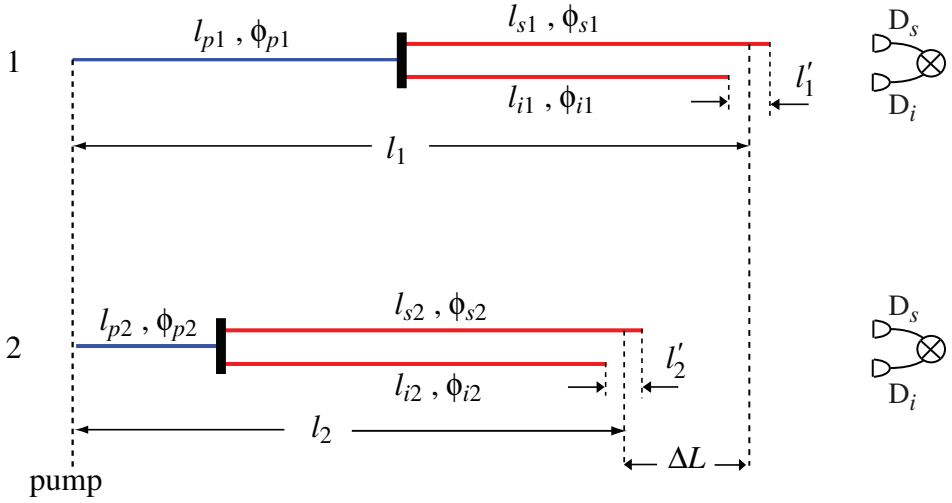


FIG. 14: Schematic representation of temporal two-photon interference using two-photon path diagrams.

asymmetry length is defined to be the difference of the optical path lengths traveled by the signal and idler photons. ΔL is the difference of the two-photon path lengths l_1 and l_2 whereas $\Delta L'$ is the difference of the two-photon path-asymmetry lengths l_1' and l_2'. The corresponding times are represented by $\Delta\tau$ and $\Delta\tau'$ respectively. Notice that if either ΔL or $\Delta L'$ is too large, alternatives 1 and 2 will become distinguishable and will no longer interfere.

The two-photon state $|\psi_{\rm tp}\rangle$ produced by down-conversion is given by Eq. (109). In a two-photon interference experiment, the complete two-photon state $|\psi\rangle$ is given by the coherent superposition of the two-photon states in alternatives 1 and 2, that is,

$$|\psi\rangle = |\psi_{\rm tp1}\rangle + |\psi_{\rm tp2}\rangle. \tag{112}$$

By substituting in Eq. (109) the relevant parameters corresponding to alternatives 1 and 2, we evaluate the two-photon states at the exit face of the crystal in the two alternatives. The complete two-photon state $|\psi\rangle$ is then given by

$$\begin{aligned}|\psi\rangle = A\iint_0^\infty d\omega_p d\omega_d V_1(\omega_p)\Phi_1\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right)e^{i(\omega_p\tau_{p1}+\phi_{p1})}\left|\frac{\omega_p}{2}+\omega_d\right\rangle_{s1}\left|\frac{\omega_p}{2}-\omega_d\right\rangle_{i1}\\ +A\iint_0^\infty d\omega_p d\omega_d V_2(\omega_p)\Phi_2\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right)e^{i(\omega_p\tau_{p2}+\phi_{p2})}\left|\frac{\omega_p}{2}+\omega_d\right\rangle_{s2}\left|\frac{\omega_p}{2}-\omega_d\right\rangle_{i2}.\end{aligned} \tag{113}$$

$V_{1(2)}(\omega_p)$ is the strength of the pump field at frequency ω_p in alternatives 1(2) and $\Phi_{1(2)}(\omega_s,\omega_d)$ is the phase-matching function in alternative 1(2). We denote the positive-frequency parts of the electric fields at detectors D_s and D_i by $\hat{E}_s^{(+)}(t)$ and $\hat{E}_i^{(+)}(t)$, respectively. The field at a detector is equal to the sum of the fields arriving at that detector by alternatives 1 and 2. Thus,

$$\begin{aligned}\hat{E}_s^{(+)}(t) &= \hat{E}_{s1}^{(+)}(t-\tau_{s1}) + \hat{E}_{s2}^{(+)}(t-\tau_{s2})\\ &= c_{s1}e^{i\phi_{s1}}\int_0^\infty d\omega_s f_s(\omega_s-\omega_{s0})e^{-i\omega_s(t-\tau_{s1})}\hat{a}_{s1}(\omega_s) + c_{s2}e^{i\phi_{s2}}\int_0^\infty d\omega_s f_s(\omega_s-\omega_{s0})e^{-i\omega_s(t-\tau_{s2})}\hat{a}_{s2}(\omega_s)\end{aligned} \tag{114}$$

and

$$\begin{aligned}\hat{E}_i^{(+)}(t) &= \hat{E}_{i1}^{(+)}(t-\tau_{i1}) + \hat{E}_{i2}^{(+)}(t-\tau_{i2})\\ &= c_{i1}e^{i\phi_{i1}}\int_0^\infty d\omega_i f_i(\omega_i-\omega_{i0})e^{-i\omega_i(t-\tau_{i1})}\hat{a}_{i1}(\omega_i) + c_{i2}e^{i\phi_{i2}}\int_0^\infty d\omega_i f_i(\omega_i-\omega_{i0})e^{-i\omega_i(t-\tau_{i2})}\hat{a}_{i2}(\omega_i)\end{aligned} \tag{115}$$

Here $\hat{E}_{s1}^{(+)}(t-\tau_{s1})$ is the positive-frequency part of the field arriving at detector D_s in alternative 1, etc; $f_s(\omega_s-\omega_{s0})$ and $f_i(\omega_i-\omega_{i0})$ are the amplitude transmission functions of the filters placed before detectors D_s and D_i, respectively. The two filters are centered at frequencies ω_{s0} and ω_{i0}. It is assumed that perfect phase-matching is satisfied at frequencies ω_0, ω_{s0} and ω_{i0} with $\omega_0=\omega_{s0}+\omega_{i0}$. c_{s1} is the probability amplitude that the signal photon is detected in alternative 1, etc. Since all the phase information is already contained in ϕ_{s1}, etc., we take the probability amplitudes to be real quantities. The probability amplitudes are normalized such that $c_{s1}^2+c_{s2}^2=c_{i1}^2+c_{i2}^2=1$

The coincidence count rate $R_{si}(t,t+\tau)$, which is the probability per (unit time)^2 that a photon is detected at D_s at time t and another at D_i at time $t+\tau$, is given by: $R_{si}(t,t+\tau)=\alpha_s\alpha_i\langle\langle\psi|\hat{E}_s^{(-)}(t)\hat{E}_i^{(-)}(t+\tau)\hat{E}_i^{(+)}(t+\tau)\hat{E}_s^{(+)}(t)|\psi\rangle\rangle$ [27], where α_s and α_i are the quantum efficiencies of detectors D_s and D_i, respectively. By substituting from Eqs. (113), (114) and (115), the coincidence count rate $R_{si}(t,t+\tau)$ can be shown to be

$$R_{si}(t,t+\tau) = \alpha_s\alpha_i\langle\langle\psi|[\hat{E}_{s1}^{(-)}(t-\tau_{s1})+\hat{E}_{s2}^{(-)}(t-\tau_{s2})][\hat{E}_{i1}^{(-)}(t+\tau-\tau_{i1})+\hat{E}_{i2}^{(-)}(t+\tau-\tau_{i2})] \tag{116}$$

$$\times[\hat{E}_{i1}^{(+)}(t+\tau-\tau_{i1})+\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})][\hat{E}_{s1}^{(+)}(t-\tau_{s1})+\hat{E}_{s2}^{(+)}(t-\tau_{s2})]|\psi\rangle\rangle \tag{117}$$

Out of the sixteen terms in the above summation, only four survives, and thus the expression for coincidence counts becomes

$$R_{si}(t,t+\tau) = \alpha_s\alpha_i\langle\langle\psi_{\rm tp1}|\hat{E}_{s1}^{(-)}(t-\tau_{s1})\hat{E}_{i1}^{(-)}(t+\tau-\tau_{i1})\hat{E}_{i1}^{(+)}(t+\tau-\tau_{i1})\hat{E}_{s1}^{(+)}(t-\tau_{s1})|\psi_{\rm tp1}\rangle\rangle \tag{118}$$

$$+\alpha_s\alpha_i\langle\langle\psi_{\rm tp2}|\hat{E}_{s2}^{(-)}(t-\tau_{s2})\hat{E}_{i2}^{(-)}(t+\tau-\tau_{i2})\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s1})|\psi_{\rm tp2}\rangle\rangle \tag{119}$$

$$+\alpha_s\alpha_i\langle\langle\psi_{\rm tp1}|\hat{E}_{s1}^{(-)}(t-\tau_{s1})\hat{E}_{i1}^{(-)}(t+\tau-\tau_{i1})\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle\rangle + {\rm H.c.} \tag{120}$$

$$\equiv R_{si}^{11}(t,t+\tau)+R_{si}^{22}(t,t+\tau)+R_{si}^{12}(t,t+\tau)+{\rm H.c.} \tag{121}$$

This is the two-photon interference law in the temporal domain. We see that there are four terms. First term is the coincidence count rate due to alternative 1 only. The second term is the coincidence count rate due to alternative 2 only. The other two are the cross terms.

LECTURE # 15

Deriving the two-photon interference law

Let us first evaluate $\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle$. Substituting from Eqs. (113), (114) and (115), we obtain

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_{i2}e^{i\phi_{i2}}c_{s2}e^{i\phi_{s2}} \times \int_0^\infty d\omega' f_i(\omega'-\omega_{i0})e^{-i\omega'(t+\tau-\tau_{i2})}\hat{a}_{i2}(\omega')\int_0^\infty d\omega f_s(\omega-\omega_{s0})e^{-i\omega(t-\tau_{s2})}\hat{a}_{s2}(\omega) \times \iint_0^\infty d\omega_p d\omega_d V_2(\omega_p)\Phi_2\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right)e^{i(\omega_p\tau_{p2}+\phi_{p2})}\left|\frac{\omega_p}{2}+\omega_d\right\rangle_{s2}\left|\frac{\omega_p}{2}-\omega_d\right\rangle_{i2}. \quad (122)$$

or, $$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_2e^{i\phi_2} \times \iiiint_0^\infty d\omega d\omega' d\omega_p d\omega_d V_2(\omega_p)\Phi_2\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right)f_s(\omega-\omega_{s0})f_i(\omega'-\omega_{i0}) \times e^{-i\omega'(t+\tau-\tau_{i2})}e^{-i\omega(t-\tau_{s2})}e^{i\omega_p\tau_{p2}}\hat{a}_{i2}(\omega')\hat{a}_{s2}(\omega)\left|\frac{\omega_p}{2}+\omega_d\right\rangle_{s2}\left|\frac{\omega_p}{2}-\omega_d\right\rangle_{i2}. \quad (123)$$

Here we have substituted $c_2 = c_{s2}c_{i2}$ and $\phi_2 = \phi s2+\phi i2+\phi p2$. Operating the annihilation operators and integrating over ω and ω', we obtain

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_2e^{i\phi_2}\iint_0^\infty d\omega_p d\omega_d V_2(\omega_p)\Phi_2\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right) \times f_s\left(\frac{\omega_p}{2}+\omega_d-\omega_{s0}\right)f_i\left(\frac{\omega_p}{2}-\omega_d-\omega_{i0}\right)e^{-i\left(\frac{\omega_p}{2}-\omega_d\right)(t+\tau-\tau_{i2})}e^{-i\left(\frac{\omega_p}{2}+\omega_d\right)(t-\tau_{s2})}e^{i\omega_p\tau_{p2}}|{\rm vac}\rangle_{s2}|{\rm vac}\rangle_{i2}. \quad (124)$$

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_2e^{i\phi_2}\iint_0^\infty d\omega_p d\omega_d V_2(\omega_p)\Phi_2\left(\frac{\omega_p}{2}+\omega_d,\frac{\omega_p}{2}-\omega_d\right) \times f_s\left(\frac{\omega_p}{2}+\omega_d-\omega_{s0}\right)f_i\left(\frac{\omega_p}{2}-\omega_d-\omega_{i0}\right)e^{-i\omega_p\left(t+\frac{\tau}{2}-\tau_2\right)}e^{i\omega_d\left(\tau+\tau_2'\right)}|{\rm vac}\rangle_{s2}|{\rm vac}\rangle_{i2}, \quad (125)$$

where we have used $\tau_2 = \tau_{p2}+(\tau_{s2}+\tau_{i2})/2$, and $\tau_2' = \tau_{s2}-\tau_{i2}$. We assume that the pump field is centered around frequency ω_0 and that the signal and idler fields are centered around frequencies ω_{s0} and ω_{i0}, respectively. Let us substitute $\omega_p = \omega_p'+\omega_0$ and $\omega_d = \omega_d'+\omega_{d0}$, where $\omega_{d0} = (\omega_{s0}-\omega_{i0})/2$, and assume that the frequency bandwidths of the pump, signal and idler fields are much smaller compared to their mean frequencies. Thus the lower limits of the integral can be extended to $-\infty$ and we get:

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_2e^{i\phi_2}e^{-i\omega_0\left(t+\frac{\tau}{2}-\tau_2\right)}e^{i\omega_{d0}\left(\tau+\tau_2'\right)}\iint_{-\infty}^\infty d\omega_p' d\omega_d' V_2(\omega_p'+\omega_0) \times \Phi_2\left(\frac{\omega_p'}{2}+\omega_d'+\omega_{s0},\frac{\omega_p'}{2}-\omega_d'+\omega_{i0}\right)f_s\left(\frac{\omega_p'}{2}+\omega_d'\right)f_i\left(\frac{\omega_p'}{2}-\omega_d'\right)e^{-i\omega_p'\left(t+\frac{\tau}{2}-\tau_2\right)}e^{i\omega_d'\left(\tau+\tau_2'\right)}|{\rm vac}\rangle_{s2}|{\rm vac}\rangle_{i2}, \quad (126)$$

Now we assume that the spectral width $\Delta\omega_p$ of the pump field is small enough such that the phase-matching function and the filter functions do not change appreciably in the frequency range $(\omega_0-\Delta\omega_p/2, \omega_0+\Delta\omega_p/2)$. The above assumption regarding the pump spectral width remains valid for most continuous wave and pulsed lasers. However, this assumption may not remain valid for very short pulsed pump [19, 28]. Thus the dependence of the phase matching function and the filter functions on ω_p' can be ignored and we can write the above integral as a product of two integrals:

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\rm tp2}\rangle = Ac_2e^{i\phi_2}e^{-i\omega_0\left(t+\frac{\tau}{2}-\tau_1\right)}e^{i\omega_{d0}\left(\tau+\tau_2'\right)}\int_{-\infty}^\infty d\omega_p' V_2(\omega_p'+\omega_0)e^{-i\omega_p'\left(t+\frac{\tau}{2}-\tau_2\right)} \times \int_{-\infty}^\infty d\omega_d'\Phi_2\left(\omega_d'+\omega_{s0},-\omega_d'+\omega_{i0}\right)f_s\left(\omega_d'\right)f_i\left(-\omega_d'\right)e^{i\omega_d'\left(\tau+\tau_2'\right)}|{\rm vac}\rangle_{s2}|{\rm vac}\rangle_{i2}, \quad (127)$$

Next we define

$$g_2^*(t) \equiv \int_{-\infty}^{\infty} d\omega_d' \Phi_2\left(\omega_d' + \omega_{s0}, -\omega_d' + \omega_{i0}\right) f_s\left(\omega_d'\right) f_i\left(-\omega_d'\right) e^{-i\omega_d' t} \quad \text{and} \tag{128}$$

$$v_2(t) \equiv \int_{-\infty}^{\infty} d\omega_p' V_2(\omega_p' + \omega_0) e^{-i\omega_p' t}, \tag{129}$$

and write the above state as

$$\hat{E}_{i2}^{(+)}(t+\tau-\tau_{i2})\hat{E}_{s2}^{(+)}(t-\tau_{s2})|\psi_{\text{tp2}}\rangle = Ac_2 e^{i\phi_2} e^{-i\omega_0\left(t+\frac{\tau}{2}-\tau_2\right)} e^{i\omega_{d0}\left(\tau+\tau_2'\right)} \times v_2\left(t+\frac{\tau}{2}-\tau_2\right) g_2^*\left(\tau+\tau_2'\right)|\text{vac}\rangle_{s2}|\text{vac}\rangle_{i2}, \tag{130}$$

Using the above expression, we evaluate the third term of Eq. (118) to be

$$R_{si}^{12}(t,t+\tau) = Kc_1^* c_2 \left\langle e^{-i(\omega_0\Delta\tau+\omega_{d0}\Delta\tau'+\Delta\phi)} v_1^*\left(t+\frac{\tau}{2}-\tau_1\right) v_2\left(t+\frac{\tau}{2}-\tau_2\right) g_1\left(\tau+\tau_1'\right) g_2^*\left(\tau+\tau_2'\right)\right\rangle, \tag{131}$$

where $k = |A|^2\alpha_s\alpha_i$ is a constant. The angular bracket $\langle\ \rangle$ represents the ensemble average. In most experiments, one does not measure the instantaneous coincidence count rate $R_{si}^{12}(t,t+\tau)$. Instead, one measures the coincidence count rate averaged with respect to t over the photon collection time T_{collect} and with respect to τ over the coincidence time-window T_{coinc}. Assuming that the fields are stationary, we can replace the ensemble average by time average taken with respect to both t and τ, that is,

$$\langle\cdots\rangle = \langle\cdots\rangle_{t,\tau} = \lim_{T_{\text{collect}}\to\infty}\lim_{T_{\text{coinc}}\to\infty} \frac{1}{T_{\text{collect}}}\frac{1}{T_{\text{coinc}}}\int_{-T_{\text{collect}}/2}^{T_{\text{collect}}/2}\int_{-T_{\text{coinc}}/2}^{T_{\text{coinc}}/2}\cdots dt d\tau, \tag{132}$$

and thus have

$$\begin{aligned} R_{si}^{12} &= \langle R_{si}^{12}(t,t+\tau)\rangle_{t,\tau} \\ &= Kc_1^* c_2 e^{-i(\omega_0\Delta\tau+\omega_{d0}\Delta\tau'+\Delta\phi)} \left\langle v_1^*\left(t+\frac{\tau}{2}-\tau_1\right) v_2\left(t+\frac{\tau}{2}-\tau_2\right)\right\rangle_t \left\langle g_1\left(\tau+\tau_1'\right) g_2^*\left(\tau+\tau_2'\right)\right\rangle_\tau, \end{aligned} \tag{133}$$

In a two-photon interference experiment, the photon collection time T_{collect} is usually a few seconds whereas the reciprocal of the pump spectral width $1/\Delta\omega_p$ for a typical laboratory laser source is usually a few microseconds or smaller. Thus T_{collect} is always much longer than $1/\Delta\omega_p$. Therefore, the correlation $\langle v_1^*\left(t+\frac{\tau}{2}-\tau_1\right) v_2\left(t+\frac{\tau}{2}-\tau_2\right)\rangle_t$ depends only on the time difference, that is,

$$\left\langle v_1^*\left(t+\frac{\tau}{2}-\tau_1\right) v_2\left(t+\frac{\tau}{2}-\tau_2\right)\right\rangle_t = \langle v_1^*(t)v_2(t+\Delta\tau)\rangle_t = \sqrt{|v_1|^2|v_2|^2}\frac{\langle v_1^*(t)v_2(t+\Delta\tau)\rangle_t}{\sqrt{|v_1|^2|v_2|^2}} = \sqrt{|v_1|^2|v_2|^2}\gamma(\Delta\tau), \tag{134}$$

where $|v_1|^2 = \langle v_1^*(t)v_1(t)\rangle_t$, etc. Also, the coincidence time window T_{coinc} in a two-photon interference experiment is usually a few nanoseconds whereas the reciprocal of the signal-idler frequency bandwidth, $1/\Delta\omega$, is typically a few picoseconds or smaller [20, 29]. Thus T_{coinc} is always much longer than $1/\Delta\omega$. Therefore, we can write

$$\langle g_1\left(\tau+\tau_1'\right) g_2^*\left(\tau+\tau_2'\right)\rangle_\tau = \langle g_1(\tau)g_2^*(\tau-\Delta\tau')\rangle_\tau = \sqrt{|g_1|^2|g_2|^2}\frac{\langle g_1(\tau)g_2^*(\tau-\Delta\tau')\rangle_\tau}{\sqrt{|g_1|^2|g_2|^2}} = \sqrt{|g_1|^2|g_2|^2}\gamma(\Delta\tau') \tag{135}$$

where $|g_1|^2 = \langle g_1(\tau)g_1^*(\tau)\rangle_\tau$, etc. Therefore, the coincidence count rate of Eq. (133) R_{si}^{12} can be written as

$$R_{si}^{12} = Kc_1^* c_2 e^{-i(\omega_0\Delta\tau+\omega_{d0}\Delta\tau'+\Delta\phi)}\sqrt{|v_1|^2|v_2|^2}\sqrt{|g_1|^2|g_2|^2}\gamma(\Delta\tau)\gamma(\Delta\tau') \tag{136}$$

Finally, the coincidence count rate R_{si} can be shown to be

$$R_{si} = R_1 + R_2 + 2\sqrt{R_1 R_2}\gamma(\Delta\tau)\gamma'(\Delta\tau')\cos(\omega_0\Delta\tau + \omega_{d0}\Delta\tau' + \Delta\phi) \tag{137}$$

where $R_1 = K|c_1 g_1 v_1|^2$ and $R_2 = K|c_2 g_2 v_2|^2$. Also, we can write the above equation in terms of ΔL and $\Delta L'$ as

$$R_{si} = R_1 + R_2 + 2\sqrt{R_1 R_2}\gamma(\Delta L)\gamma'(\Delta L')\cos(k_0\Delta L + k_{d0}\Delta L' + \Delta\phi) \tag{138}$$

where $k_0 = \omega_0/c$ and $k_d0 = \omega_{d0}/c$. The product $\gamma'(\Delta L')\,\gamma(\Delta L)$ is the time-averaged degree of two-photon correlation. In an interference experiment, two-photon fringes are observed only so long as the product $\gamma'(\Delta L')\,\gamma(\Delta L)$ is greater than zero. Two-photon interference gets washed out once either of the two correlation function becomes much smaller than unity. From the expression for the coincidence count rate R_{si}, it follows that two-photon interference effects in the temporal domain can be completely described in terms of the variations of the two length parameters, ΔL and $\Delta L'$. In the special case, when the down-converted fields are degenerate, that is $k_{d0} = 0$, and $R_1 = R_2 = C/2$, the expression for the coincidence count rate reduces to

$$R_{si} = C[1 + \gamma'(\Delta L')\,\gamma(\Delta L)\cos(k_0\Delta L + \Delta\phi)]. \tag{139}$$

In what follows we use this simplified expression for the coincidence count rate.

The time-averaged degree of correlations of the pump and the signal-idler fields depend on the spectrum of the pump and the signal-idler fields, respectively. For the special case when the pump is a stationary (continuous-wave) field having a Gaussian spectrum of rms frequency width $\Delta\omega_p$, the time-averaged degree of correlation of the pump field can be evaluated using the generalized Wiener-Kintchine theorem to be

$$\gamma(\Delta L) = \exp\left[-\frac{1}{2}\left(\frac{\Delta L}{l^p_{\rm coh}}\right)^2\right]. \tag{140}$$

Here $l^p_{\rm coh} = c/\Delta\omega_p$ is the coherence length of the pump field. One can evaluate the time-averaged degree of correlation $\gamma'(\Delta L')$ of the signal-idler field for various phase-matching conditions. In situations in which the signal-idler field has a Gaussian spectrum of width $\Delta\omega$, the time-averaged degree of correlation of the signal-idler field can be shown to be as

$$\gamma'(\Delta L') = \exp\left[-\frac{1}{2}\left(\frac{\Delta L'}{l_{\rm coh}}\right)^2\right], \tag{141}$$

where $l_{\rm coh} = c/\Delta\omega$ is the rms width of $\gamma'(\Delta L')$ as a function of $\Delta L'$ and is a measure of the reciprocal bandwidth of the signal-idler field. It is also taken quite often as a measure of the coherence length of the signal-idler field [20, 21, 26]. However, we note that this is strictly true only when the signal-idler field is completely stationary.

We now look at the effects of varying ΔL and $\Delta L'$ on the coincidence count rate R_{si} of Eq. (139) by considering two limiting cases.

Case I: For $\Delta L' = 0$ and $\Delta\phi = 0$,

$$R_{si} = C\,[\,1 + \gamma(\Delta L)\cos(k_0\Delta L)\,]. \tag{142}$$

Interference is observed in the coincidence count rate as ΔL is varied and gets washed out once ΔL exceeds the pump coherence length. Thus ΔL plays the same role in two-photon interference as does the optical path-length difference in one-photon interference. It is because of this analogy that we call ΔL the two-photon path-length difference. The coincidence fringes seen in Franson-type interferometers [30, 31] and in the double-pass setup [22] are examples of effects due to variations in ΔL.

Case II: For ΔL and $\Delta\phi$ fixed,

$$R_{si} = C\,[\,1 + K\gamma'(\Delta L')\,], \tag{143}$$

where $K = \gamma(\Delta L)\cos(k_0\Delta L + \Delta\phi)$ is constant. The coincidence count rate can show a dip when the two alternatives interfere destructively ($K < 0$), and a hump when the two alternatives interfere constructively ($K > 0$), as $\Delta L'$ is varied. These profiles, with widths equal to $l_{\rm coh}$, represent how the coherence between two two-photon alternatives changes with a variation in $\Delta L'$. $\Delta L'$ has no one-photon counterpart.

One Photon Interference Effects with Entangled photons

In many two-photon interference experiments, one-photon interference effects are also observed. These one-photon effects cannot be described by second-order (in the field) coherence theory [34] because the one-photon path-length differences involved in these experiments are much longer than the coherence lengths of the one-photon fields. Interference effects in one photon count rates have previously also been observed in many two-photon experiments including

induced coherence [23], frustrated two-photon creation [22] and interference experiment from separate pulses [35]. However, a simpler theory based on classical probability can be worked out to explain all these effects. Let us consider a complete set $\{Y_1, Y_2, \cdots Y_N\}$ of mutually exclusive random events. Since the set is complete, we have

$$\sum_{i=1}^{N} P(Y_i) = 1 \tag{144}$$

Now, let X be another random event and let $P(X, Y_i)$ be the joint probability that X and Y_i happen together. Then the probability that event X happens is given by

$$P(X) = \sum_{i=1}^{N} P(X, Y_i) \tag{145}$$

Use this simple result from classical probability theory, we can obtain the detection probability of one of the two photons of the entangled pair as follows. Suppose a photon of the entangled pair is detected at location X and its entangled partner has the probability of getting detected an N separate locations, given by $\{Y_1, Y_2, \cdots Y_N\}$. Then clearly, $\{Y_1, Y_2, \cdots Y_N\}$ is a mutually exclusive set of random events with $\sum_{i=1}^{N} P(Y_i) = 1$. Therefore, the one-photon count rate R_{X} at a detection position X is equal to the sum of the coincidence count rates R_{XY_i} between X and all the other positions Y_i where the entangled partner of the photon detected at X can go, i.e.,

$$R_{\mathrm{X}} = \sum_{i=1}^{N} R_{\mathrm{XY}_i}, \tag{146}$$

Summing over the detector positions R_{Y_i} in Eq. (146) is same as taking the partial trace over all the possible modes of the twin. The summation turns into an integral whenever the twin has finite probabilities of arriving at a continuous set of detection points.

LECTURE # 16

Some example of two-photon interference effects

1. Hong-Ou-Mandel (HOM) Effect

As illustrated in Fig. 15(a), in the Hong-Ou-Mandel (HOM) experiment [20], the signal and idler photons from PDC are mixed at a beam splitter and the two-photon interference effect is observed in the coincidence count rate of detectors D_s and D_i, as a function of the beam splitter position x. The experiment can be understood in terms of the two-photon path diagrams shown in Fig. 15(b). In alternative 1, both the signal and idler photons get reflected by the beam splitter, while in alternative 2, they both get transmitted by the beam splitter. Using the two-photon path diagrams, we find that $\Delta L = 0$, $\Delta L' = 4x\cos\theta$, and $\Delta\phi = \pi$. The coincidence count rate R_{si} is then calculated using Eq. (139) to be

$$R_{si} = C[1 - \gamma'(4x\cos\theta)] \tag{147}$$

At $x = 0$, that is, at the balanced position of the beam splitter, the coincidence count rate R_{si} is equal to zero. This implies that at the balanced position of the beam splitter, both photons always leave through the same output port of the beam splitter. As a result, a null is observed in the coincidence count rate at the balanced position, leading to a dip in the coincidence count rate as a function of the beam splitter position x. An intuitive explanation of this effect can be given in terms of the bunching of signal and idler photons at a beam splitter [32].

2. The postponed compensation effect

Hon-Ou-Mandel effect was described in terms of the bunching of entangled photons at a beam-splitter. However, the bunching interpretation is not adequate for the postponed compensation [21] and related experiments [33] in which HOM-like effects are observed, even when signal and idler photons do not simultaneously arrive at a beam splitter. Let us now look at one of the HOM-like experiment, called here as postponed compensation experiment [21]. Fig. 16 shows the experiment and the corresponding two-photon pat diagram. We find that in this case, we have $\Delta L = D$, $\Delta L' = 4x\cos\theta$, and $\Delta\phi = \pi$. The coincidence count rate R_{si} is then calculated using Eq. (139) to be

$$R_{si} = C[1 - \gamma(D)\gamma'(4x\cos\theta)\cos(k_0 D)] \tag{148}$$

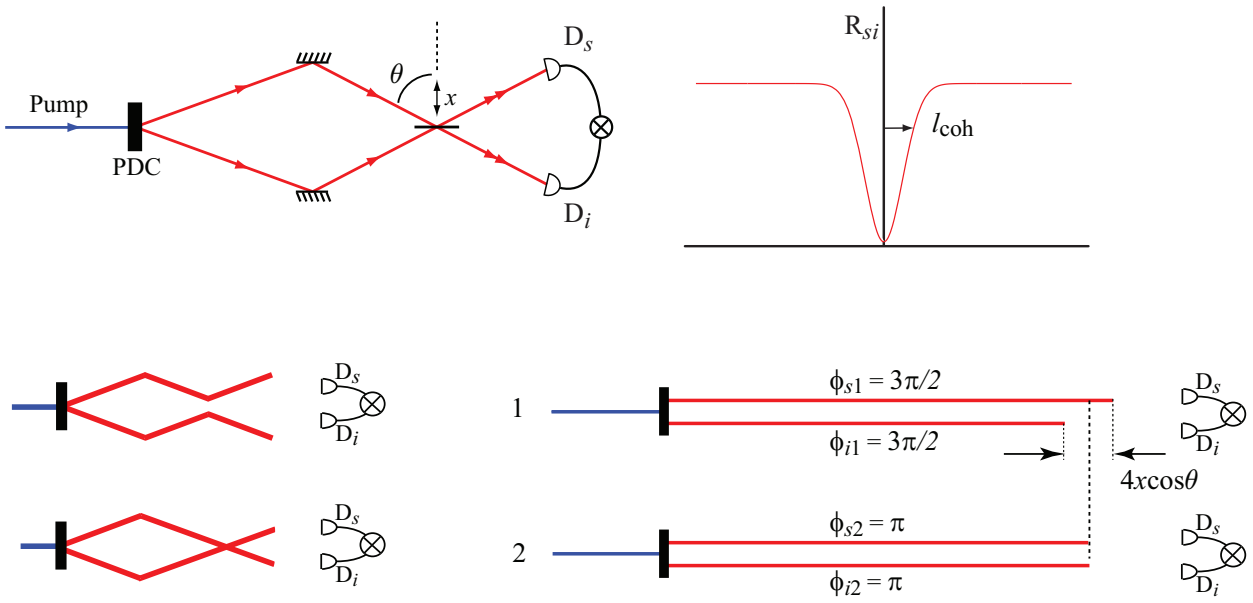


FIG. 15: (a) Schematic of the Hong-Ou-Mandel (HOM) experiment. (b) Two-photon path diagrams illustrating two-photon interference in a HOM experiment. In alternative 1, both the signal and idler photons get reflected by the beam splitter, while in alternative 2, they both get transmitted.

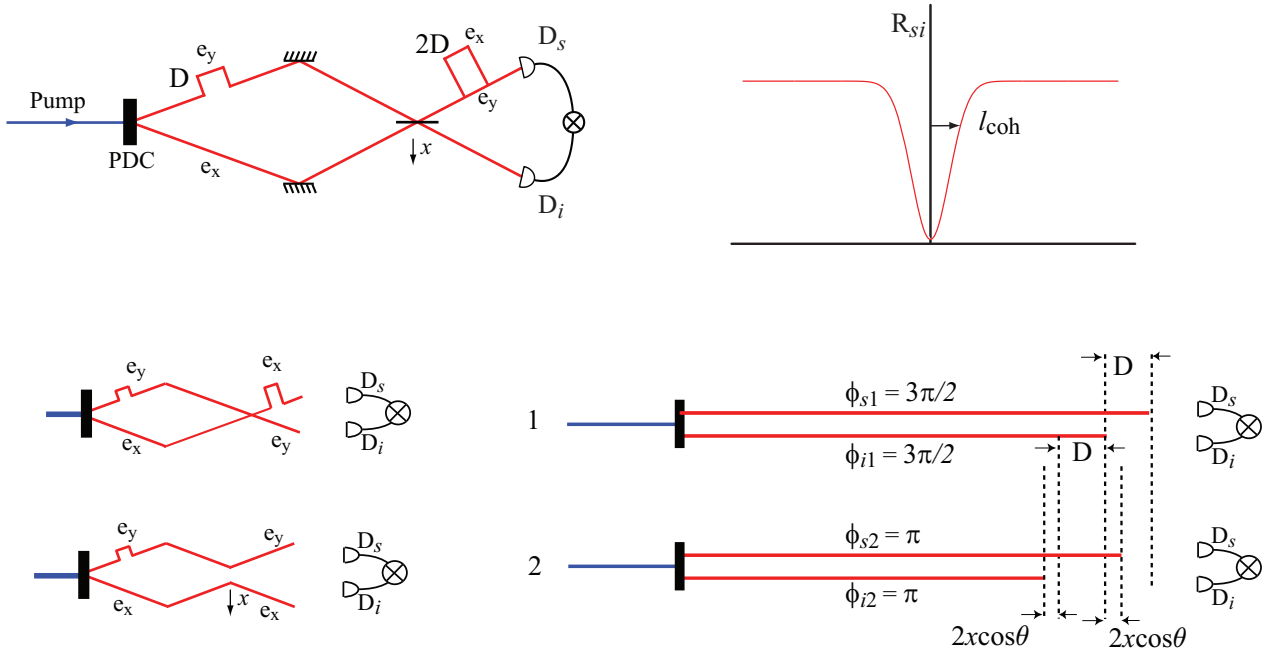


FIG. 16: Two-photon path diagram for the postponed compensation experiment.

Suppose we choose the value of D such that $\gamma(D) \approx 1$ and $\cos(k_0 D) = 1$, then R_{si} shows the same HOM like variation as a function of the beam splitter position x.

3. Induced coherence experiment

Let us now look at Mandel's famous "induced coherence experiment [23]," using the formalism developed in Chapter 2 for describing temporal two-photon interference effects. The schematic of the experiment is depicted in Fig. 17. In this experiment, two parametric down-converters (PDC1 and PDC2) are pumped coherently. When the paths of the idler photons (i_1 and i_2) are aligned, one-photon fringes are observed at detector D_A as the beam splitter position x is varied. To explain this effect, we first calculate the coincidence count rate R_{AB} of detectors D_A and D_B. From the two-photon path diagrams shown in Fig. 17(b), one finds that that $\Delta L = x\cos\theta, \Delta L' = 2x\cos\theta$ and $\Delta\phi = \pi$. Substituting these quantities into Eq. (139), we obtain

$$R_{AB} = C\left[1 - \gamma'(2x\cos\theta)\gamma\left(x\cos\theta\right)\cos\left(k_0 x\cos\theta\right)\right]. \tag{149}$$

Next, we calculate the coincidence count rate R_{CB} of detectors D_C and D_B. From the two-photon path diagrams shown in Fig. 17(c), we have, $\Delta L = -x\cos\theta, \Delta L' = -2x\cos\theta$ and $\Delta\phi = 0$. Substituting these quantities into Eq. (139), we obtain

$$R_{CB} = C\left[1 + \gamma'(2x\cos\theta)\gamma\left(x\cos\theta\right)\cos\left(k_0 x\cos\theta\right)\right]. \tag{150}$$

To calculate the one-photon count rates at detectors D_A and D_B, we note that the twin of a photon detected at D_A can go only to D_B while the twin of a photon detected at D_B can go to both D_A and D_C. Therefore, using Eqs. (146), we find that the one-photon count rates R_A and R_B at detectors D_A and D_B, respectively are given as:

$$R_A = R_{AB} \qquad \text{and} \tag{151}$$
$$R_B = R_{AB} + R_{CB}. \tag{152}$$

Using Eqs. (149) and (150), we then obtain

$$R_A = C\left[1 - \gamma'(2x\cos\theta)\gamma\left(x\cos\theta\right)\cos\left(k_0 x\cos\theta\right)\right], \tag{153}$$
$$R_B = 2C. \tag{154}$$

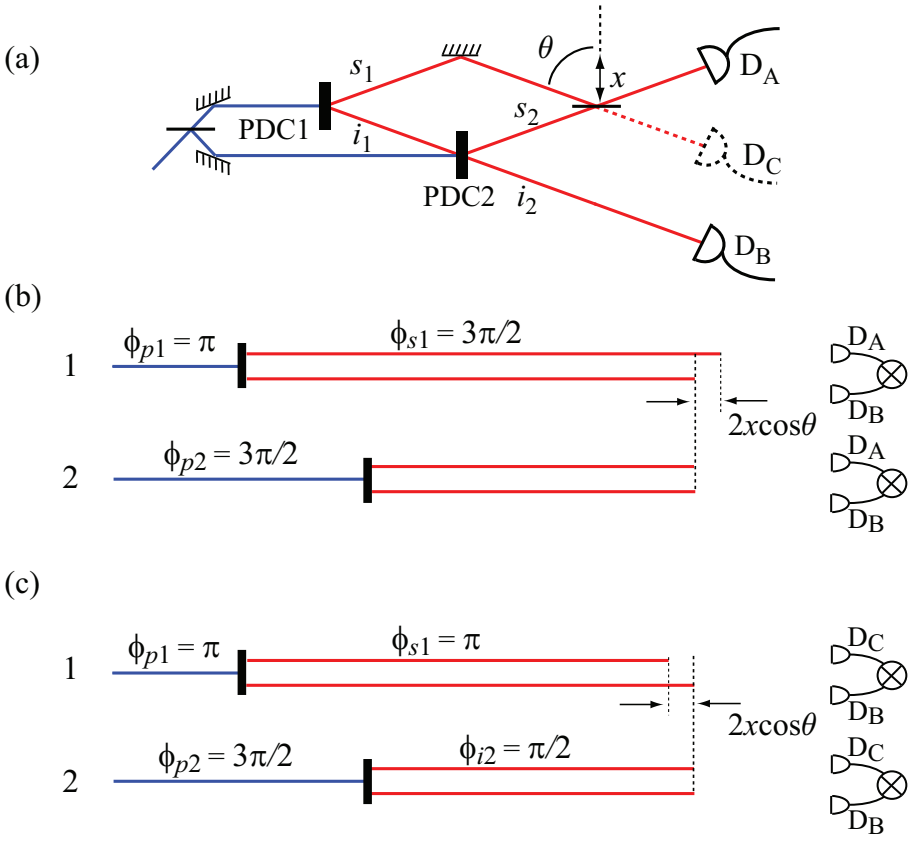


FIG. 17: (a) The schematic setup for the induced-coherence experiment [23]. In alternative 1, the pump photon gets down-converted in PDC1 while in alternative 2, it gets down-converted in PDC2. (b) Two-photon path diagrams representing the alternative pathways by which the signal and idler photons can get to detectors D_A and D_B. (c) Two-photon path diagrams representing the alternative pathways by which the signal and idler photons can get to detectors D_C and D_B.

The one-photon count rate R_A at detector D_A thus shows interference fringes as a function of x whereas the one-photon count rate R_B at detector D_B does not. These were the results reported in Ref. [23] and explained in terms of 'induced coherence'. Here we have shown that they can also be explained in terms of two-photon interference effects alone.

LECTURE # 17

Two-Photon State Produced by Parametric Down-Conversion: Spatial

The interaction Hamiltonian $\hat{H}(t')$ for parametric down-conversion is given by [see Eq. (96)]

$$\hat{H}(t) = \epsilon_0 \int_{\mathcal{V}} d^3\boldsymbol{r}\chi^{(2)}\hat{E}_p^{\ (+)}(\boldsymbol{r},t)\hat{E}_s^{\ (-)}(\boldsymbol{r},t)\hat{E}_i^{(-)}(\boldsymbol{r},t) + \text{H.c.}, \tag{155}$$

where $\mathcal{V}$ is the volume of the interacting part of the nonlinear crystal; $\chi^{(2)}$ is the second-order nonlinear susceptibility; $\hat{E}_j^{\ (+)}(\boldsymbol{r},t)$ and $\hat{E}_j^{\ (-)}(\boldsymbol{r},t)$ are the positive- and negative-frequency parts of the field, where $j = p, s, i$ stands for the pump, signal and idler, respectively. We assume that the nonlinear crystal is embedded in a passive linear medium of suitable refractive index and that $\chi^{(2)}$ is independent of frequency over the range of interest. [36, 37]. We have seen in the case of classical field that the electric field $V(\boldsymbol{r},t)$ at position $\boldsymbol{r}$ at time t can be represented in terms of a Fourier integral as

$$\begin{aligned} V(\boldsymbol{r},t) &= \int_{-\infty}^{\infty} \tilde{v}(\boldsymbol{r},\omega)e^{-i\omega t}d\omega \\ &= \int_{-\infty}^{\infty} U(\boldsymbol{\rho},z;\omega)e^{-i\omega t}d\omega \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} Aa(\boldsymbol{q},\omega)e^{i(\boldsymbol{q}.\boldsymbol{\rho}+k_z z)}d^2\boldsymbol{q}\right)e^{-i\omega t}d\omega \\ &= \iint_{-\infty}^{\infty} Aa(\boldsymbol{q},\omega)e^{i(\boldsymbol{q}.\boldsymbol{\rho}+k_z z)}e^{-i\omega t}d^2\boldsymbol{q}d\omega, \end{aligned}$$

where A is a constant and were we have used $\vec{\boldsymbol{k}} = \vec{\boldsymbol{q}}\hat{\boldsymbol{\rho}} + k_z\hat{z}$. The mode decomposition stays the same within quantum formalism. Therefore, we can write the three electric fields as

$$\hat{E}_p^{\ (+)}(\boldsymbol{r},t) = \iint_{-\infty}^{\infty} A_p V(\boldsymbol{q_p},\omega_p)e^{i(\boldsymbol{q_p}.\boldsymbol{\rho}+k_{pz}z)}e^{-i\omega_p t}d^2\boldsymbol{q_p}d\omega_p, \tag{156}$$

$$\hat{E}_s^{\ (-)}(\boldsymbol{r},t) = \iint_{-\infty}^{\infty} A_s^*\hat{a}_s^\dagger(\boldsymbol{q_s},\omega_s)e^{-i(\boldsymbol{q_s}.\boldsymbol{\rho}+k_{sz}z)}e^{i\omega_s t}d^2\boldsymbol{q_s}d\omega_s, \tag{157}$$

$$\hat{E}_i^{\ (-)}(\boldsymbol{r},t) = \iint_{-\infty}^{\infty} A_i^*\hat{a}_i^\dagger(\boldsymbol{q_i},\omega_i)e^{-i(\boldsymbol{q_i}.\boldsymbol{\rho}+k_{iz}z)}e^{i\omega_i t}d^2\boldsymbol{q_i}d\omega_i. \tag{158}$$

Here $\boldsymbol{r} = (\boldsymbol{\rho}, z)$. A_j is a frequency dependent quantity, and as it varies very slowly within the frequency range of interest for most realistic experiments, it is taken outside the integral. The pump field has been assumed to be very strong, and it is being treated classically. The strength of the pump field at $(\boldsymbol{q}_p, \omega_p)$ is represented by $V(\boldsymbol{q_p}, \omega_p)$. Using the expressions for the three fields above, we write Eq. (155) as

$$\begin{aligned} \hat{H}(t) = A_p A_s^* A_i^* \epsilon_0\chi^{(2)}\int_{\mathcal{V}} d^3\boldsymbol{r}\iiint_0^{\infty} d\omega_p d\omega_s d\omega_i \iiint d\boldsymbol{q}_p d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_p,\omega_p)\hat{a}_s^\dagger(\boldsymbol{q}_s,\omega_s)\hat{a}_i^\dagger(\boldsymbol{q}_i,\omega_i) \\ \times e^{i[(\boldsymbol{q}_p-\boldsymbol{q}_s-\boldsymbol{q}_i).\boldsymbol{\rho}+(k_{pz}-k_{sz}-k_{iz})z]}e^{i(\omega_s+\omega_i-\omega_p)t} + \text{H.c.} \end{aligned} \tag{159}$$

The state $|\psi(0)\rangle$ of the down-converted field at $t = 0$ is given by Eq. (103)

$$|\psi(0)\rangle = \exp\left[\frac{1}{i\hbar}\int_{-t_{\text{int}}}^{0} dt\hat{H}(t)\right]|\psi(-t_{\text{int}})\rangle, \tag{160}$$

Here $|\psi(-t_{\text{int}})\rangle = |\text{vac}\rangle_s|\text{vac}\rangle_i$ is the state of the down-converted field at $t = -t_{\text{int}}$ which is a vacuum state with no photons in either the signal or the idler mode. The parametric interaction is assumed to be very weak and the state in Eq. (160) is approximated by the first two terms of a perturbative expansion. The first term is the initial vacuum

state. The second term $|\psi\rangle_{\rm tp}$ is calculated by substituting from Eq. (159) into Eq. (160), which yields

$$|\psi_{\rm tp}\rangle = \frac{1}{i\hbar}\int_{-t_{\rm int}}^{0} dt\hat{H}(t)|\psi(-t_{\rm int})\rangle = \frac{A_p A_s^* A_i^* \epsilon_0 \chi^{(2)}}{i\hbar}\int_{-t_{\rm int}}^{0} dt \iint d^2\boldsymbol{\rho}dz \iiint_0^\infty d\omega_p d\omega_s d\omega_i$$
$$\times \iiint d\boldsymbol{q}_p d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_p,\omega_p) e^{i[(\boldsymbol{q}_p-\boldsymbol{q}_s-\boldsymbol{q}_i).\boldsymbol{\rho}+(k_{pz}-k_{sz}-k_{iz})z]}$$
$$\times e^{i(\omega_s+\omega_i-\omega_p)t}\hat{a}_s^\dagger(\boldsymbol{q}_s,\omega_s)\hat{a}_i^\dagger(\boldsymbol{q}_i,\omega_i,)|{\rm vac}\rangle_s|{\rm vac}\rangle_i, \quad (161)$$

where we have substituted $\int d^3\boldsymbol{r} \to \iint d^2\boldsymbol{\rho}dz$. The interaction time $t_{\rm int}$ is assumed to be much longer than the time scale over which down-conversion takes place and much smaller than the time interval between two consecutive down-conversion events. The limits of time integration in Eq. (161) are therefore extended to $\pm\infty$ [18, 19]. We evaluate the time integral, which yields the delta function, that is, $\int_{-\infty}^{\infty} e^{i(\omega_s+\omega_i-\omega_p)t} = \delta(\omega_p-\omega_s-\omega_i)$. We also assume that the effective interaction area is much smaller that the actual size of the nonlinear medium. Thus the limit of the $\boldsymbol{\rho}$ integral can be extended to $\pm\infty$ and the integral yields a delta function, that is, $\int d^2\boldsymbol{\rho}e^{i(\boldsymbol{q}_p-\boldsymbol{q}_s-\boldsymbol{q}_i).\boldsymbol{\rho}} = \delta(\boldsymbol{q}_p-\boldsymbol{q}_s-\boldsymbol{q}_i)$. Integrating over ω_p and $\boldsymbol{q}_p$ we obtain

$$|\psi_{\rm tp}\rangle = A\iint_0^\infty d\omega_s d\omega_i \iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i)\Phi(\omega_s,\omega_i)|\omega_s,\boldsymbol{q}_s\rangle_s|\omega_i,\boldsymbol{q}_i\rangle_i, \quad (162)$$

where

$$\Phi(\omega_s,\omega_i) = \int_{-L}^{0} e^{i(k_{pz}-k_{sz}-k_{iz})z}dz \quad (163)$$

is the phase-matching function, where L is the thickness of the crystal. All the constant factors have been absorbed into A. This a very general form for the state of the two-photon field produced by parametric down-conversion. Let us now look at the two special cases:

1. Case I: Signal and idler emission directions are fixed. This means that we fix $\boldsymbol{q}_s$ and $\boldsymbol{q}_i$. So, we have in this case: $V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i) = V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i)\delta(\boldsymbol{q}_s-\boldsymbol{q}_{s0})\delta(\boldsymbol{q}_i-\boldsymbol{q}_{i0})$. The two-photon wavefunction is therefore given by:

$$|\psi_{\rm tp}\rangle = A\iint_0^\infty d\omega_s d\omega_i \iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i)\delta(\boldsymbol{q}_s-\boldsymbol{q}_{s0})\delta(\boldsymbol{q}_i-\boldsymbol{q}_{i0})\Phi(\omega_s,\omega_i)|\omega_s,\boldsymbol{q}_s\rangle_s|\omega_i,\boldsymbol{q}_i\rangle_i \quad (164)$$
$$= A\iint_0^\infty d\omega_s d\omega_i V(\boldsymbol{q}_{s0}+\boldsymbol{q}_{i0},\omega_s+\omega_i)\Phi(\omega_s,\omega_i)|\omega_s,\boldsymbol{q}_{s0}\rangle_s|\omega_i,\boldsymbol{q}_{i0}\rangle_i \quad (165)$$
$$\equiv A\iint_0^\infty d\omega_s d\omega_i V(\omega_s+\omega_i)\Phi(\omega_s,\omega_i)|\omega_s\rangle_s|\omega_i\rangle_i, \quad (166)$$

   Thus we get the state of the two-photon as obtained when considering only the temporal effects.

2. Case II: Signal and idler fields are monochromatic. In this case we have fixed values for frequencies ω_s and ω_i. Therefore, we write $V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i) = V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i)\delta(\omega_s-\omega_{s0})\delta(\omega_i-\omega_{i0})$. The two-photon wavefunction is therefore given by:

$$|\psi_{\rm tp}\rangle = A\iint_0^\infty d\omega_s d\omega_i \iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s+\boldsymbol{q}_i,\omega_s+\omega_i)\delta(\omega_s-\omega_{s0})\delta(\omega_i-\omega_{i0})\Phi(\omega_s,\omega_i)|\omega_s,\boldsymbol{q}_s\rangle_s|\omega_i,\boldsymbol{q}_i\rangle_i \quad (167)$$
$$= A\iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s+\boldsymbol{q}_i)\Phi(\omega_{s0},\omega_{i0})|\omega_{s0},\boldsymbol{q}_s\rangle_s|\omega_{i0},\boldsymbol{q}_i\rangle_i \quad (168)$$
$$\equiv A\iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s+\boldsymbol{q}_i)\Phi(\omega_{s0},\omega_{i0})|\boldsymbol{q}_s\rangle_s|\boldsymbol{q}_i\rangle_i, \quad (169)$$

   This is the two-photon state in the momentum basis, when the signal and idler fields are monochromatic. In the special case of pump field being a plane wave, that is, $V(\boldsymbol{q}_s+\boldsymbol{q}_i) = V(\boldsymbol{q}_s+\boldsymbol{q}_i)\delta(\boldsymbol{q}_s+\boldsymbol{q}_i-\boldsymbol{q}_0)$, we get

$$|\psi_{\rm tp}\rangle = AV(\boldsymbol{q}_0)\int d\boldsymbol{q}_s\Phi(\omega_{s0},\omega_{i0})|\boldsymbol{q}_s\rangle_s|\boldsymbol{q}_0-\boldsymbol{q}_s\rangle_i, \quad (170)$$

In situations when $\Phi(\omega_{s0}, \omega_{i0}) = 1$ and $\boldsymbol{q}_0 = 0$, the two-photon state becomes

$$|\psi_{\rm tp}\rangle = AV(\boldsymbol{q}_0 = 0) \int d\boldsymbol{q}_s |\boldsymbol{q}_s\rangle_s |-\boldsymbol{q}_s\rangle_i, \tag{171}$$

We find that when signal photon has the transverse momentum $\boldsymbol{q}_s$, the idler photon has the transverse momentum $-\boldsymbol{q}_s$. This is the same state as was considered by EPR in the context of EPR paradox. And so, we see that such states can indeed by produced in a physical experiment. This is an entangled state of two-photon in the momentum basis. By taking the Fourier transform, the above state can be written the position basis and one can show similar correlation in the position basis as well.

Two-Photon Interference: Spatial

Figure 18 represents a generic situation for studying the spatial coherence properties of the two-photon field. The signal and idler photons produced by PDC go through a pair of double-holes located at plane z. They are detected in coincidence by detectors D_s and D_i located at positions $\boldsymbol{r}_s$ and $\boldsymbol{r}_i$, respectively. There are two alternative pathways, by which signal and idler photons can reach detectors D_s and D_i. In alternative 1, the signal and idler photons go through the pair of holes located at $\boldsymbol{r}_{s1}(\boldsymbol{\rho}_{s1}, z)$ and $\boldsymbol{r}_{i1}(\boldsymbol{\rho}_{i1}, z)$, and in alternative 2, they go through those located at $\boldsymbol{r}_{s2}(\boldsymbol{\rho}_{s2}, z)$ and $\boldsymbol{r}_{i2}(\boldsymbol{\rho}_{i2}, z)$. In principle, there are two more alternative pathways: one in which the signal and idler photons go through the pair of holes located at $\boldsymbol{r}_{s1}(\boldsymbol{\rho}_{s1}, z)$ and $\boldsymbol{r}_{i2}(\boldsymbol{\rho}_{i2}, z)$, and the second in which they go through those located at $\boldsymbol{r}_{s2}(\boldsymbol{\rho}_{s2}, z)$ and $\boldsymbol{r}_{i1}(\boldsymbol{\rho}_{i1}, z)$. In what follows we explicitly assume that the phase matching condition is such that the probability amplitudes of these two other alternatives are negligibly small.

In Fig. 18, subscripts p, s, and i stand for pump, signal, and idler, respectively. The distance travelled by a photon from the crystal to a hole is denoted by r. The distance travelled from a hole to the corresponding detector is denoted by d and the associated time elapsed by $t = d/c$. The transverse position vector of a photon is denoted by $\boldsymbol{\rho}$. Thus $\boldsymbol{\rho}_{s1}$ represents the transverse position vector of the signal photon in alternative 1, etc. We define two displacement parameters in terms of the transverse position vectors of the signal and idler photons in the two alternatives as:

$$\begin{aligned} \Delta\boldsymbol{\rho} &= \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2 \equiv \frac{\boldsymbol{\rho}_{s1} + \boldsymbol{\rho}_{i1}}{2} - \frac{\boldsymbol{\rho}_{s2} + \boldsymbol{\rho}_{i2}}{2}, \\ \Delta\boldsymbol{\rho}' &= \boldsymbol{\rho}'_1 - \boldsymbol{\rho}'_2 \equiv (\boldsymbol{\rho}_{s1} - \boldsymbol{\rho}_{i1}) - (\boldsymbol{\rho}_{s2} - \boldsymbol{\rho}_{i2}). \end{aligned} \tag{172}$$

Here $\boldsymbol{\rho}_{1(2)}$ and $\boldsymbol{\rho}'_{1(2)}$ are the two-photon transverse position vector and the two-photon position-asymmetry vector in alternative 1(2). For either alternative, the two-photon transverse position vector is defined to be the average of the transverse position vectors of the signal and idler photons, and the two-photon position-asymmetry vector is defined to be the difference of the transverse position vectors of the signal and idler photons.

The coincidence count rate $R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i)$, which is the probability per (unit time)^2 that a photon is detected at position $\boldsymbol{r}_s$ at time t and another at position $\boldsymbol{r}_i$ at time $t + \tau$, is given by $R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i) = \alpha_s \alpha_i \langle\langle \psi_{\rm tp} | \hat{E}_s^{(-)}(\boldsymbol{r}_s, t) \hat{E}_i^{(-)}(\boldsymbol{r}_i, t + \tau) \hat{E}_i^{(+)}(\boldsymbol{r}_i, t + \tau) \hat{E}_s^{(+)}(\boldsymbol{r}_s, t) | \psi_{\rm tp} \rangle\rangle$ [27], where α_s and α_i denote the quantum efficiencies of detectors D_s and D_i, respectively, and $\psi_{\rm tp}$ is the state of the two-photon field produced by PDC. The two-photon state was derived earlier [Eq. (167)] and is given as

$$|\psi_{\rm tp}\rangle = A \iint d\boldsymbol{q}_s d\boldsymbol{q}_i V(\boldsymbol{q}_s + \boldsymbol{q}_i) |\boldsymbol{q}_s\rangle_s |\boldsymbol{q}_i\rangle_i, \tag{173}$$

where we have taken $\Phi(\omega_{s0}, \omega_{i0}) = 1$. We denote the positive-frequency parts of the electric fields at detectors D_s and D_i by $\hat{E}_s^{(+)}(\boldsymbol{r}_s, t)$ and $\hat{E}_i^{(+)}(\boldsymbol{r}_i, t)$, respectively. $\hat{E}_s^{(+)}(\boldsymbol{r}_s, t)$ and $\hat{E}_i^{(+)}(\boldsymbol{r}_i, t)$ are equal to the sum of the signal and idler fields arriving at detectors D_s and D_i by alternatives 1 and 2, i.e.,

$$\hat{E}_s^{(+)}(\boldsymbol{r}_s, t) = k_{s1} \hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s1}, z) e^{-i\omega_s (t - t_{s1})} + k_{s2} \hat{E}_{s2}^{(+)}(\boldsymbol{\rho}_{s2}, z) e^{-i\omega_s (t - t_{s2})}, \tag{174}$$

$$\hat{E}_i^{(+)}(\boldsymbol{r}_i, t) = k_{i1} \hat{E}_{i1}^{(+)}(\boldsymbol{\rho}_{i1}, z) e^{-i\omega_i (t - t_{i1})} + k_{i2} \hat{E}_{i2}^{(+)}(\boldsymbol{\rho}_{i2}, z) e^{-i\omega_i (t - t_{i2})}. \tag{175}$$

Here $\hat{E}_{i1}^{(+)}(\boldsymbol{\rho}_{s1}, z)$ is the positive-frequency part of the signal field at position $(\boldsymbol{\rho}_{s1}, z)$, etc. The constant factor k_{s1} depends on the size of the hole at $\boldsymbol{\rho}_{s1}$ and the geometry of the arrangement, etc. By substituting from Eqs. (174)

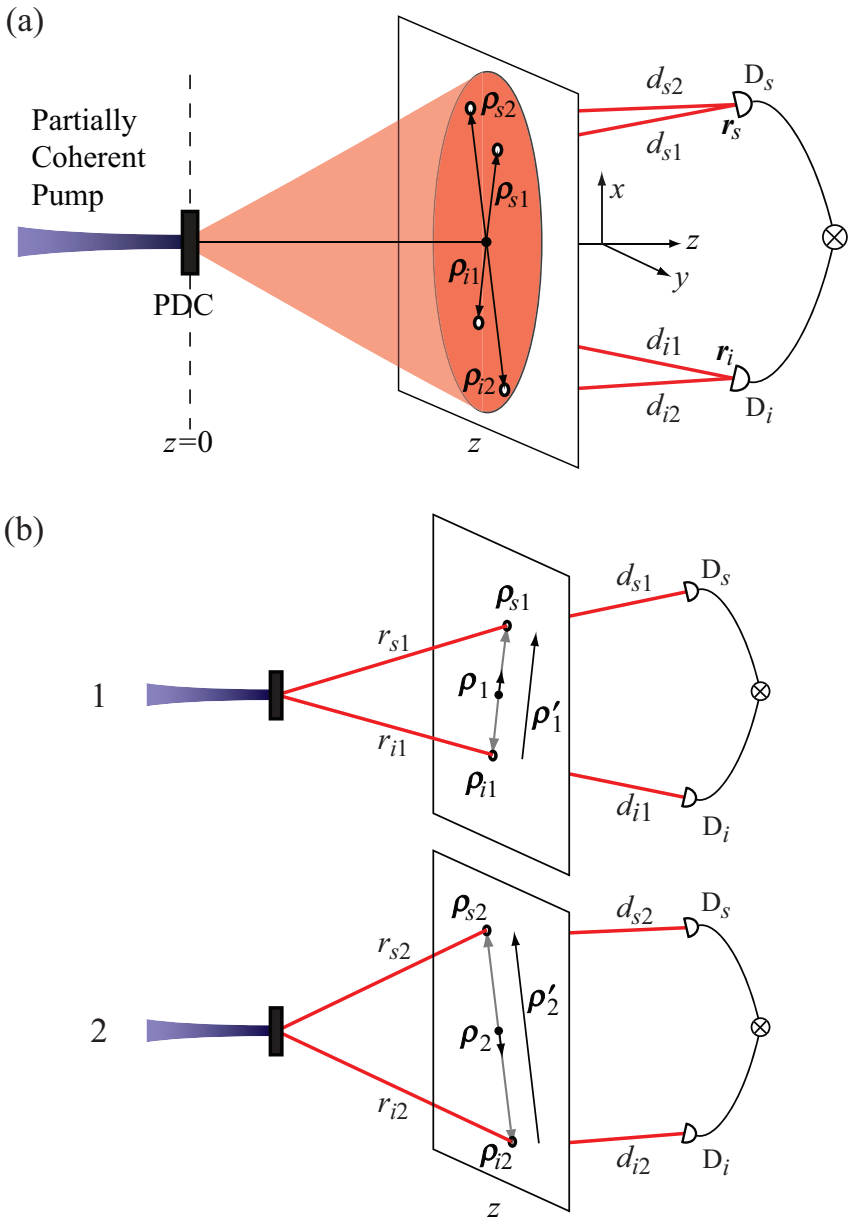


FIG. 18: (a) Schematic laboratory setup that could be used to study the spatial coherence properties of the two-photon field produced by PDC using a partially coherent pump beam. (b) 1 and 2 represent two alternative pathways by which the down-converted signal and idler photons can pass through the holes and get detected in coincidence at detectors D_s and D_i. In alternative 1, the signal and idler photons go through the pair of holes located at $\boldsymbol{r}_{s1}(\boldsymbol{\rho}_{s1}, z)$ and $\boldsymbol{r}_{i1}(\boldsymbol{\rho}_{i1}, z)$, and in alternative 2, they go through those located at $\boldsymbol{r}_{s2}(\boldsymbol{\rho}_{s2}, z)$ and $\boldsymbol{r}_{i2}(\boldsymbol{\rho}_{i2}, z)$.

and (175), we write the coincidence count rate $R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i)$ as

$$\begin{aligned} R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i) &= \alpha_s\alpha_i\langle\langle\psi_{\rm tp}|\hat{E}_s^{(-)}(\boldsymbol{r}_s)\hat{E}_i^{(-)}(\boldsymbol{r}_i)\hat{E}_i^{(+)}(\boldsymbol{r}_i)\hat{E}_s^{(+)}(\boldsymbol{r}_s)|\psi_{\rm tp}\rangle\rangle \\ &= \alpha_s\alpha_i|k_{s1}k_{i1}|^2\langle\langle\psi_{\rm tp}|\hat{E}_{s1}^{(-)}(\boldsymbol{\rho}_{s1}, z)\hat{E}_{i1}^{(-)}(\boldsymbol{\rho}_{i1}, z)\hat{E}_{i1}^{(+)}(\boldsymbol{\rho}_{i1}, z)\hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s1}, z)|\psi_{\rm tp}\rangle\rangle \\ &+ \alpha_s\alpha_i|k_{s2}k_{i2}|^2\langle\langle\psi_{\rm tp}|\hat{E}_{s2}^{(-)}(\boldsymbol{\rho}_{s2}, z)\hat{E}_{i2}^{(-)}(\boldsymbol{\rho}_{i2}, z)\hat{E}_{i2}^{(+)}(\boldsymbol{\rho}_{i2}, z)\hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s2}, z)|\psi_{\rm tp}\rangle\rangle \\ &+ \alpha_s\alpha_i k_{s1}k_{i1}k_{s2}^*k_{i2}^*\langle\langle\psi_{\rm tp}|\hat{E}_{s1}^{(-)}(\boldsymbol{\rho}_{s1}, z)\hat{E}_{i1}^{(-)}(\boldsymbol{\rho}_{i1}, z)\hat{E}_{i2}^{(+)}(\boldsymbol{\rho}_{i2}, z)\hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s2}, z)|\psi_{\rm tp}\rangle\rangle e^{i(k_0\Delta L + k_d\Delta L')} + \text{H.c.} \end{aligned} \tag{176}$$

The electric field operators $\hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s1}, z)$ and $\hat{E}_{i1}^{(+)}(\boldsymbol{\rho}_{i1}, z)$, within the paraxial approximation, can be written as (refer

to Lecture # 6):

$$\hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s1}, z) = e^{ik_0 z/2} \int d\boldsymbol{q} \hat{a}_s(\boldsymbol{q}) e^{i(\boldsymbol{q}\cdot\boldsymbol{\rho}_{s1} - q^2 z/k_0)}, \tag{177}$$

$$\hat{E}_{i1}^{(+)}(\boldsymbol{\rho}_{i1}, z) = e^{ik_0 z/2} \int d\boldsymbol{q'} \hat{a}_i(\boldsymbol{q'}) e^{i(\boldsymbol{q'}\cdot\boldsymbol{\rho}_{i1} - q'^2 z/k_0)}, \tag{178}$$

where $q^2 = |\boldsymbol{q}|^2$, $q'^2 = |\boldsymbol{q'}|^2$ and $k_0 = |\boldsymbol{k}_0(\omega_0)|$. By substituting from Eqs. (174) and (175), we write the coincidence count rate $R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i)$ as

$$R_{si}(\boldsymbol{r}_s, \boldsymbol{r}_i) = k_1^2 S^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, z) + k_2^2 S^{(2)}(\boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z) + k_1 k_2 W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z) e^{i(k_0 \Delta L + k_d \Delta L')} + \text{c.c.} \tag{179a}$$

where $k_1 = \sqrt{\alpha_s \alpha_i} k_{s1} k_{i1}$, $k_2 = \sqrt{\alpha_s \alpha_i} k_{s2} k_{i2}$,

$$W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z) = \langle\langle \psi_{\text{tp}} | \hat{E}_{s1}^{(-)}(\boldsymbol{\rho}_{s1}, z) \hat{E}_{i1}^{(-)}(\boldsymbol{\rho}_{i1}, z) \hat{E}_{i2}^{(+)}(\boldsymbol{\rho}_{i2}, z) \hat{E}_{s1}^{(+)}(\boldsymbol{\rho}_{s2}, z) | \psi_{\text{tp}} \rangle \rangle \tag{179b}$$

and

$$S^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, z) = W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, z). \tag{179c}$$

Equation (179a) is the interference law for the two-photon field. The first and second terms of Eq. (179a) are the coincidence count rates when coincidences are collected from only alternatives 1 and 2, respectively. These terms are recognized as the two-photon analogs of the spectral density functions of the second-order coherence theory. $S^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, z)$ will be referred to as the two-photon spectral density in alternative 1, etc. The interference term $W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z)$, which appears when coincidences are collected from both the alternatives, is a four-point fourth-order (in the field) correlation function. It satisfies four Wolf Equations [27, 34, 38] and is recognized as the two-photon analog of the cross-spectral density function. $W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z)$ will be referred to as the two-photon cross-spectral density function. To keep the notations simpler, we do not show the frequency arguments in the definitions of the two-photon spectral density and the two-photon cross-spectral density functions. The two-photon cross-spectral density can be written as

$$W^{(2)}(\boldsymbol{\rho}_{s1}, \boldsymbol{\rho}_{i1}, \boldsymbol{\rho}_{s2}, \boldsymbol{\rho}_{i2}, z) = e^{ik_0(r_1 - r_2)} \sqrt{S^{(2)}(\boldsymbol{\rho}_1, z) S^{(2)}(\boldsymbol{\rho}_2, z)} \mu^{(2)}(\Delta\boldsymbol{\rho}, z), \tag{180}$$

where $r_1 = (r_{s1} + r_{i1})/2$ and $r_2 = (r_{s2} + r_{i2})/2$. $\mu^{(2)}(\Delta\boldsymbol{\rho}, z)$ is the spatial two-photon coherence function, the magnitude of which is always between 0 and 1. Let us compare this with the one-photon interference law we had defined earlier:

$$S(\boldsymbol{\rho}; z; \omega) = |k_1|^2 S(\boldsymbol{\rho}_1, \omega) + |k_2|^2 S(\boldsymbol{\rho}_2, \omega) + 2|k_1||k_2| \sqrt{S(\boldsymbol{\rho}_1, \omega) S(\boldsymbol{\rho}_2, \omega)} \mu(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \omega) \cos[\omega(t_2 - t_1) + \phi],$$

We see that the structure of the two interference laws is exactly the same. It can be shown that, just like $\gamma(\Delta L)$, the functional form of $\mu^{(2)}(\Delta\boldsymbol{\rho}, z)$ is the same as that of the spatial coherence function of the pump field. Again, we see that the spatial two-photon coherence can be described in terms of the variations of just two transverse-vectors and that the coherence properties of the pump gets entirely transferred to the down-converted two-photon field. Also, if the degree spatial two-photon coherence $\mu^{(2)}(\Delta\boldsymbol{\rho}, z) = 0$ then there would be no two-photon interference.

LECTURE # 18

Lab Visit: Phase matching in Parametric down-conversion

LECTURE # 19

Can the quantum mechanical description of physical reality be considered complete?

This section is entirely based on the paper by A. Einstein, B. Podolsky and N. Rosen, Physical Review, 47, 777, (1935). In this paper, Einstein, Podolsky and Rosen (EPR) presented an argument to show that the quantum mechanical description of physical reality is not complete. In this paper EPR recognized a spooky feature in the quantum description of physical reality [1]. This feature, now known as entanglement, was originally called by Schrödinger "Verschränkung," which implies the existence of global states of a composite system that cannot be described as a product of the states of the individual subsystems [2]. For an extensive review of the subject of quantum entanglement, see the article by Horodecki *et al.* [3].

In their paper, EPR started out by defining physical reality. According to them: "If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity." Next they observed that: "In a complete theory there is an element corresponding to each element of reality. A sufficient condition for the reality of a physical quantity is the possibility of predicting it with certainty, without disturbing the system. In quantum mechanics in the case of two physical quantities described by non-commuting operators, the knowledge of one precludes the knowledge of the other. Then either (1) the description of reality given by the wave function in quantum mechanics is not complete or (2) these two quantities cannot have simultaneous reality. Consideration of the problem of making predictions concerning a system on the basis of measurements made on another system that had previously interacted with it leads to the result that if (1) is false then (2) is also false. One is thus led to conclude that the description of reality as given by a wave function is not complete."

We present EPR's argument by first reviewing one-particle quantum system. Suppose a one-particle system is in an eigenstate $|\psi\rangle$ of some measurement operator $\hat{A}$ corresponding a physically observable quantity. Then we know that the action of the measurement operator would yield a real value a for the physical quantity and the state after the measurement would remain the same $|\psi\rangle$, that is,

$$\hat{A}|\psi\rangle = a|\psi\rangle.$$

Further action of the operator would yield the same value a. In this sense we can say that for a particle in state $|\psi\rangle$, there is an element of physical reality corresponding to $\hat{A}$. As a concrete example, let us consider a one-particle system in one of the momentum eigenstates. The action of the momentum operator is,

$$\hat{K}|\psi_{k_0}\rangle = k_0|\psi_{k_0}\rangle,$$

where k_0 is a constant value of momentum. Let us rewrite this in terms of the wave-function. The momentum eigenfunction in the position basis is written as $\langle x|\psi_{k_0}\rangle = \psi_{k_0}(x) = e^{ik_0x}$. Also, we know that the position-basis representation of the momentum operator is $\hat{K} \to -i\frac{\partial}{\partial x}$. Therefore the above operator equation becomes

$$-i\frac{\partial}{\partial x}\psi_{k_0}(x) = k_0\psi_{k_0}(x)$$

We see that for this system momentum is real and has a fixed value. This means that there is an element of physical reality associated with momentum. But what about the position of the particle. The action of the position operator on the above state is:

$$\hat{X}|\psi_{k_0}\rangle \neq a|\psi_{k_0}\rangle.$$

That is, position measurements on the system do not yield a fixed value, since $|\psi_{k_0}\rangle$ is not an eigenstate of the position operator. So, the only question that can be asked in this case is, "what is the probability that the particle is found between x_1 and x_2?" This probability is

$$P(x_1, x_2) = \int_{x_1}^{x_2} |\psi_{k_0}(x)|^2 dx = \int_{x_1}^{x_2} dx = x_2 - x_1,$$

which means that the particle has equal probability of being anywhere. Therefore, one can have no certainty about the position of the particle. One can obtain a a value by direct measurement but a measurement disturbs the system in the sense that it will no longer remain in the state $|\psi_{k_0}\rangle$ after the measurement. Therefore, we conclude that there is no element of physical reality corresponding to position.

Next, suppose that the particle is in the eigenstate of the position operator. In this case we have

$$\hat{X}|\phi_{x_0}\rangle = x_0|\phi_{x_0}\rangle,$$

where x_0 is a fixed value of position. The position-basis representation of this eigenstate is $\langle x|\phi_{x_0}\rangle = \phi_{x_0}(x) = \delta(x-x_0)$. The positing operator in the position-basis is written as $\hat{X} = x$. Therefore the operator equation becomes

$$x\delta(x-x_0) = x_0\delta(x-x_0).$$

We see that for a one-particle system in a position eigenstate, position is real and has a fixed value x_0, that is, there is an element of physical reality associated with position. But what about the momentum in this case. The action of the momentum operator on the above state is:

$$\hat{K}|\phi_{x_0}\rangle \neq a|\phi_{x_0}\rangle.$$

This time we see that $|\phi_{x_0}\rangle$ is not an eigenstate of the momentum operator and so the only question that can be asked in this case is as to what is the probability that the particle is found to have momenta values between k_1 and k_2. The probability in this case is

$$P(k_1,k_2) = \int_{k_1}^{k_2} |\psi_{x_0}(k)|^2 dk,$$

where $\psi_{x_0}(k)$ is the momentum state wave-function which can be calculated from the position-space wave-function as

$$\psi_{x_0}(k) = \int_{-\infty}^{\infty} e^{-ikx}\phi_{x_0}(x)dx = \int_{-\infty}^{\infty} e^{-ikx}\delta(x-x_0)dx = e^{-ikx_0}.$$

The required probability, therefore, is

$$P(k_1,k_2) = \int_{k_1}^{k_2} |\psi_{x_0}(k)|^2 dk = \int_{k_1}^{k_2} dk = k_2 - k_1.$$

The particle has equal probability of having any momentum value. One can have no certainty about the momentum of the particle, and therefore, for the above state, there is no element of physical reality corresponding to momentum. Analyzing both the examples considered above, we conclude that a particle cannot simultaneously be in the eigenstates of both momentum and position. More generally, in quantum mechanics, if the operators $\hat{A}$ and $\hat{B}$ corresponding to two physical quantities do not commute, that is, $[A,B] \neq 0$, then both of them cannot have simultaneous physical reality, that is, a system cannot simultaneously be in the eigenstates of two non-commuting operators. This is the first principle of quantum mechanics and is quantified in terms of the Heisenberg uncertainty principle. If the quantum mechanical description of physical reality is to be considered complete, the first principle must remain intact for all quantum systems.

Next, we show an apparent contradiction to this first principle. Let us consider a system consisting of two particles. The particles have a common past and may have had interactions in the past, but they are now space-like separated. Let us assume that the wave-function corresponding to the two particles can be written as

$$\Psi(x_1,x_2) = \int_{-\infty}^{\infty} \psi_k^{(2)}(x_2)u_k^{(1)}(x_1)dk.$$

This wavefunction is a continuous coherent superposition of two-particle wavefunctions $\psi_k^{(2)}(x_2)u_k^{(1)}(x_1)$. This means that if we happen to find the first particle with wavefunction $u_k^{(1)}(x_1)$ then the second particle is guaranteed to be found with the wavefunction $\psi_k^{(2)}(x_2)$. Now, let us take a more specific situation in which we have,

$$\psi_k^{(2)}(x_2) = e^{-ik(x_2-x_0)}$$
$$\text{and} \qquad u_k^{(1)}(x_1) = e^{ikx_1},$$

$$\text{such that} \qquad \Psi(x_1,x_2) = \int_{-\infty}^{\infty} \psi_k^{(2)}(x_2)u_k^{(1)}(x_1)dk = \int_{-\infty}^{\infty} e^{ik(x_1-x_2+x_0)}dk.$$

We see that $u_k^{(1)}(x_1) = e^{ikx_1}$ is the eigenfunction of the momentum operator $\hat{K}_1 = -i\dfrac{\partial}{\partial x_1}$ of the first particle with the momentum eigenvalue k. Similarly, $\psi_k^{(2)}(x_2) = e^{-ik(x_2-x_0)}$ is the eigenfunction of the momentum operator $\hat{K}_2 = -i\dfrac{\partial}{\partial x_2}$ of the second particle with eigenvalue $-k$. So, if we measure the first particle and find it be in the momentum eigenstate $u_k^{(1)}(x_1)$ with momentum k then the second particle is guaranteed to be in the momentum eigenstate $\psi_k^{(2)}(x_2)$ with momentum $-k$. Next, let us write the above two-particle wavefunction as:

$$\begin{aligned}
\Psi(x_1, x_2) &= \int_{-\infty}^{\infty} \psi_k^{(2)}(x_2) u_k^{(1)}(x_1) dk \\
&= \int_{-\infty}^{\infty} e^{-ik(x_2-x_0)} e^{ikx_1} dk \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{-ikx} \delta(x - x_2 + x_0) dx \right] \left[\int_{-\infty}^{\infty} e^{ikx'} \delta(x_1 - x') dx' \right] dk \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{-ik(x-x')} dk \right] \delta(x - x_2 + x_0) \delta(x_1 - x') dx dx' \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - x') \delta(x - x_2 + x_0) \delta(x_1 - x') dx dx' \\
&= \int_{-\infty}^{\infty} \delta(x - x_2 + x_0) \delta(x_1 - x) dx \\
&= \int_{-\infty}^{\infty} \phi_x^{(2)}(x_2) v_x^{(1)}(x_1) dx,
\end{aligned}$$

where

$$\begin{aligned}
& \phi_x^{(2)}(x_2) = \delta(x - x_2 + x_0) \\
\text{and} \qquad & v_x^{(1)}(x_1) = \delta(x_1 - x)
\end{aligned}$$

We see that $v_x^{(1)}(x_1) = \delta(x_1 - x)$ is the eigenstate of position operator $\hat{X}_1$ of the first particle with eigenvalue x. Similarly $\phi_x^{(2)}(x_2) = \delta(x - x_2 + x_0)$ is the eigenstate of position operator $\hat{X}_2$ of the second particle with eigenvalue $x + x_0$. Thus, if we measure the first particle and find it to be in the position eigenstate $v_x^{(1)}(x_1)$ at x then the second particle is guaranteed to be in the position eigenstate $\phi_x^{(2)}(x_2)$ at $x + x_0$.

Comparing the above two set of observations for the two-particle system, we find that the second particle can either be in the momentum eigenstate or be in the position eigenstate, depending on the measurement performed on the first particle. But according to the first principle of quantum mechanics, a particle cannot simultaneously be in the eigenstatess of both position and momentum. This apparent contradiction with the first principle of quantum mechanics is referred to as the EPR paradox and this led EPR to conclude that the quantum description of physical reality is not complete and that it should be supplemented by postulating the existence of "hidden variables," the specifications of which will predetermine the result of measuring any observable of the system. EPR's inherent belief was that the predictions of quantum mechanics are correct but only as a consequence of the statistical distribution of hidden variables.

A "hidden variable" interpretation of quantum mechanics was formally constructed by David Bohm in 1950s [39, 40]. Bohm believed that hidden variable theories lead to precisely the same results for all physical processes as does the quantum theory, while providing a precise and continuous description of all physical processes. For quite some time, there was no conclusive way to test the validity of these two theories. Then in 1964 John S. Bell came up with his inequalities, an experimental violation of which ruled out any local hidden variable interpretations of quantum mechanics [41].

LECTURE # 20

Hidden Variable Interpretation of Quantum Mechanics

In the last section, we studied the entangled state that was conceived by EPR. The state was expressible as a continuous coherent sum of the product of eigenstates $\psi_k(x_2)u_k(x_1)$ in the momentum basis:

$$\Psi(x_1,x_2)=\int_{-\infty}^{\infty}\psi_k(x_2)u_k(x_1)dk=\int_{-\infty}^{\infty}e^{-ik(x_2-x_0)}e^{ikx_1}dk.$$

We note that the state $\Psi(x_1,x_2)$ cannot be represented as the product of the states corresponding to the individual particles, that is, $\Psi(x_1,x_2)\neq\Psi_1(x_1)\Psi_2(x_2)$. This is the general characteristic of an entangled state that it cannot be represented as the product of the states corresponding to individual particles. The above entangled state can also be expressed as a continuous coherent sum of the product of the eigenstates in the position basis, that is,

$$\Psi(x_1,x_2)=\int_{-\infty}^{\infty}\phi_x(x_2)v_x(x_1)dx=\int_{-\infty}^{\infty}\delta(x-x_2+x_0)\delta(x_1-x)dx,$$

We see that the general structure of an EPR entangled state is the same in both position and momentum bases. Just as it is formulated in terms of position and momentum, an EPR-type argument can be formulated in any pair of conjugate bases. For example, a similar EPR argument can be formulated for time and frequency. Also, position, momentum, time and frequency are all examples of infinite dimensional continuous bases. However, even if the basis is not continuous, an EPR-type argument can be formulated. For example, an EPR-type argument has been formulated for orbital angular momentum, which is an infinite-dimensional discrete basis. An entangled state in an infinite-dimensional discrete basis has the following generic structure

$$|\Psi\rangle=\sum_{n=-\infty}^{\infty}|\psi_n\rangle|u_n\rangle.$$

Here, the two-particle entangled state is represented as a discrete coherent sum of the product of the eigenstates. In fact, for the existence of an entangled state, the dimensionality of the discrete basis need not be infinity. As an example, let us consider an entangled state in a two-dimensional basis, which can be written as

$$|\Psi\rangle=|\psi_1\rangle|u_1\rangle+|\psi_2\rangle|u_2\rangle,$$

We note that the dimensionality has to be at least two for the entangled state to exist, otherwise the two-particle state becomes expressible as a product of the states corresponding to individual particles. An example of the two-dimensional basis is the polarization basis. Writing $|\psi_1\rangle=|H\rangle$, $|\psi_2\rangle=|V\rangle$, $|u_1\rangle=|H\rangle$, and $|u_2\rangle=|V\rangle$, we have a normalized, entangled state in the polarization basis

$$|\Psi\rangle=\frac{1}{\sqrt{2}}\left[|H\rangle|H\rangle+|V\rangle|V\rangle\right].$$

The first ket in a product refers to the first particle and second ket to the second particle. We find that if the first particle is detected to be horizontally polarized, the second particle is guaranteed to be horizontally polarized; and if the first particle is detected to be vertically polarized, the second particle is again guaranteed to be vertically polarized. The above state can be written in the $45^0/-45^0$-polarization basis as

$$|\Psi\rangle=\frac{1}{\sqrt{2}}\left[|45\rangle|45\rangle+|-45\rangle|-45\rangle\right].$$

$$\text{where}\qquad|45\rangle=\frac{1}{\sqrt{2}}\left[|H\rangle+|V\rangle\right]\quad\text{and}\quad|-45\rangle=\frac{1}{\sqrt{2}}\left[|H\rangle-|V\rangle\right].$$

Now, we find that if the first particle is detected to be 45^0-polarized, the second particle is guaranteed to be 45^0-polarized; and if the first particle is detected to be -45^0-polarized, the second particle is again guaranteed to be -45^0-polarized. This is the EPR paradox for a two-dimensional basis, which points out simultaneous correlations in two conjugate bases, H/V and $45^0/-45^0$.

Let us now go back to the conclusions of EPR which is that the quantum mechanical description is incomplete. In the conclusions, EPR also hinted that one can instead work out a theory in which it would be possible to get rid of

the apparent non-locality of quantum mechanics. This suggestion was later pushed forward by de-Broglie and David Bohm, and they worked out what is known as the hidden-variable theory as an alternative to quantum theory. A hidden variable theory can be understood as a set of rules that are imposed on the two particles such that the statistical predictions of quantum mechanics is reproduced without having to use the notion of entangled state. Figure. 19 is a cartoon illustrating the quantum-mechanical versus hidden-variable-model descriptions of entanglement. Let us now try to understand the hidden variable theory through various test cases:

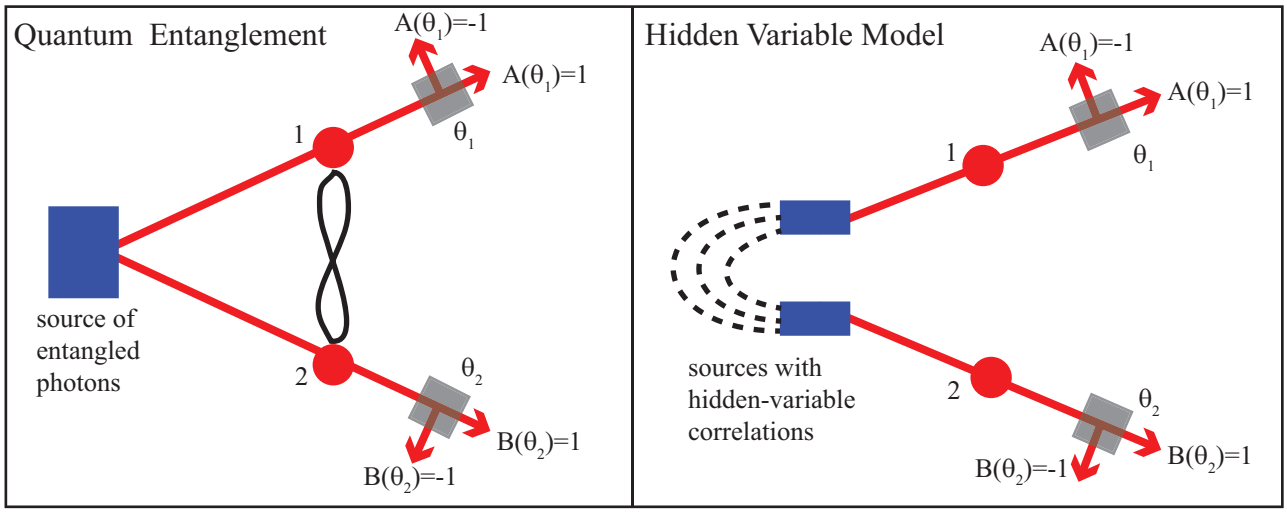


FIG. 19: Quantum-Mechanical versus Hidden-Variable-Model descriptions of entanglement.

(a) Measurement Outcome: If the first particle is horizontal then the second particle is also horizontal.

- **Quantum Mechanical wavefunction**: $|\Psi\rangle = |H\rangle|H\rangle$.
- **HVT model:**
  (i) Both sources emit simultaneously.
  (ii) Both sources emit photons with horizontal polarization.

(b) Measurement Outcome: If the first particle is 45^0 polarized then the second particle is also 45^0 polarized.

- **Quantum Mechanical wavefunction**: $|\Psi\rangle = |45\rangle|45\rangle$.
- **HVT model:**
  (i) Both sources emit simultaneously.
  (ii) Both sources emit photons with 45^0 polarization.

(c) Measurement Outcome: If the first particle is horizontal then the second particle is also horizontal, and if the first particle is vertical then the second particle is also vertical.

- **Quantum Mechanical wavefunction**: $|\Psi\rangle = \frac{1}{\sqrt{2}}\left[|H\rangle|H\rangle + |V\rangle|V\rangle\right] = \frac{1}{\sqrt{2}}\left[|45\rangle|45\rangle + |-45\rangle|-45\rangle\right]$.
- **HVT model:**
  (i) Both sources emit simultaneously.
  (ii) 50% of the time both sources emit horizontally polarized photons.
  (iii) 50% of the time both sources emit vertically polarized photons.

(d) Measurement Outcome: If the first particle is horizontal then the second particle is also horizontal, if the first particle is vertical then the second particle is also vertical, if the first particle is 45^0 polarized then the second particle is also 45^0 polarized, and if the first particle is -45^0 polarized then the second particle is also -45^0 polarized.

- **Quantum Mechanical wavefunction**: $|\Psi\rangle = \frac{1}{\sqrt{2}}\left[|H\rangle|H\rangle + |V\rangle|V\rangle\right] = \frac{1}{\sqrt{2}}\left[|45\rangle|45\rangle + |-45\rangle|-45\rangle\right]$.

- **HVT model:**
  (i) ?
  (ii) ?

The main question is: Can one construct a hidden variable model for every measurement outcome that can be explained using quantum mechanics. We see that for a lot of situations one is able to make a hidden variable model. But one needs to be sure if this can be done for any general measurement outcome. For quite some time there was no way of either proving or disproving this. And then in 1964 John S. Bell proposed a decisive test to check whether or not the local hidden variable interpretations of quantum mechanics were compatible with the statistical predictions of quantum mechanics [41]. Bell formalized EPR's idea of a deterministic world in terms of the local hidden variable models. He showed that for a single particle system local hidden variable models produce all the results correctly; however, in the case of a system consisting of two particles, the hidden variable models do not correctly predict all the results, which are predicted correctly by quantum mechanics. He then constructed an inequality and proved that the quantum-mechanical correlations could violate the inequality, but the correlations based on hidden variable models must satisfy it. Bell further showed that for the suggested hidden variable models [39, 40] to violate Bell's inequality, they have to be non-local and would have to involve faster than light propagation as well. In his own words, "in a hidden variable theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant [41]." A popular description of the concepts of non-locality and Bell's inequality has been presented by Kwiat and Hardy [42, 43].

LECTURE # 21

Bell Inequalities

We now derive the Bell inequality. We refer to Fig. 19. Let $A(\theta_1)$ and $B(\theta_2)$ be two dichotomic variables. Both $A(\theta_1)$ and $B(\theta_2)$ can have only two measurement outcomes: $+1$ and -1. $A(\theta_1) = 1$ represents emergence of the first particle through the polarizer kept at θ_1, whereas $A(\theta_1) = -1$ represents the absorption of the first particle by the polarized kept at θ_1, etc. We define

$$C(\theta_1, \theta_2) = \langle A(\theta_1) B(\theta_2) \rangle$$

to be the correlation between the two observables, ensemble-averaged over all possible realizations. According to the hidden variable theory, there is a hidden variable λ, with the probability density $\rho(\lambda)$ such that $\int \rho(\lambda) d\lambda = 1$, which decides the two outcomes. Therefore, within the hidden variable model, the correlation defined above can be written as

$$C(\theta_1, \theta_2) = \int A(\theta_1, \lambda) B(\theta_2, \lambda) \rho(\lambda) d\lambda.$$

We note that $A(\theta_1, \lambda)$ depends only on θ_1 and therefore cannot get influenced by a measurement performed on the second particle. Similarly $B(\theta_2, \lambda)$ depends only on θ_2 and is thus not influenced by a measurement performed on the first particle. Thus the locality assumption remain intact within the hidden variable description. Now, let us consider the quantity $C(\theta_1, \theta_2) - C(\theta_1, \theta_2')$. We have

$$\begin{aligned} |C(\theta_1, \theta_2) - C(\theta_1, \theta_2')| &= | \int A(\theta_1, \lambda)[B(\theta_2, \lambda) - B(\theta_2', \lambda)] \rho(\lambda) d\lambda| \\ &\leqslant \int |A(\theta_1, \lambda)[B(\theta_2, \lambda) - B(\theta_2', \lambda)]| \rho(\lambda) d\lambda \end{aligned}$$

Since $|A(\theta_1, \lambda)| = 1$, we have

$$|C(\theta_1, \theta_2) - C(\theta_1, \theta_2')| \leqslant \int |B(\theta_2, \lambda) - B(\theta_2', \lambda)| \rho(\lambda) d\lambda \tag{181}$$

Similarly,

$$|C(\theta_1', \theta_2) + C(\theta_1', \theta_2')| \leqslant \int |B(\theta_2, \lambda) + B(\theta_2', \lambda)| \rho(\lambda) d\lambda \tag{182}$$

Adding the above two equations, we get

$$|C(\theta_1, \theta_2) - C(\theta_1, \theta_2')| + |C(\theta_1', \theta_2) + C(\theta_1', \theta_2')| \leqslant \int [|B(\theta_2, \lambda) - B(\theta_2', \lambda)| + |B(\theta_2, \lambda) + B(\theta_2', \lambda)|] \rho(\lambda) d\lambda. \tag{183}$$

Now, since, B takes on only the values ± 1, we have

$$|B(\theta_2, \lambda) - B(\theta_2', \lambda)| + |B(\theta_2, \lambda) + B(\theta_2', \lambda)| = 2.$$

Using this and the normalization of $\rho(\lambda)$, that is $\int \rho(\lambda) d\lambda = 1$, we arrive at the Bell inequality

$$S \equiv |C(\theta_1, \theta_2) - C(\theta_1, \theta_2')| + |C(\theta_1', \theta_2) + C(\theta_1', \theta_2')| \leqslant 2,$$

where S is the so-called Bell parameter. We find that the Bell parameter remains less than or equal to 2 for any hidden variable model. We now calculate the bound on the Bell parameter for quantum mechanical description through wavefunctions. The correlation function in this case can be written as

$$\begin{aligned} C(\theta_1, \theta_2) &\equiv \langle A(\theta_1) B(\theta_2) \rangle \\ &= P(+, \theta_1, +, \theta_2) + P(-, \theta_1, -, \theta_2) - P(-, \theta_1, +, \theta_2) - P(+, \theta_1, -, \theta_2) \end{aligned}$$

where $P(+, \theta_1, -, \theta_2)$ is the probability that the first particle emerges out from the first polarizer kept at θ_1 and the second particle does not emerge out from the second polarizer kept at θ_2, etc. This can also be seen as the probability

that the first particle emerges our from the first polarizer kept at θ_1 and the second particle emerges from the second polarizer kept at $\theta_2 + \pi/2$. We denote this probability as $P(\theta_1, \theta_2 + \pi/2)$ and thus write the above correlation as

$$C(\theta_1, \theta_2) = P(\theta_1, \theta_2) + P(\theta_1 + \pi/2, \theta_2 + \pi/2) - P(\theta_1 + \pi/2, \theta_2) - P(\theta_1, \theta_2 + \pi/2)$$

The probability $P(\theta_1, \theta_2)$ is given by:

$$P(\theta_1, \theta_2) = \langle \Psi | \hat{a}^\dagger_{1\theta_1} \hat{a}^\dagger_{2\theta_2} \hat{a}_{2\theta_2} \hat{a}_{1\theta_1} | \Psi \rangle,$$

where $|\Psi\rangle = \frac{1}{\sqrt{2}} [|H\rangle|H\rangle + |V\rangle|V\rangle]$ and $a_{\hat{1\theta_1}}$ is the annihilation operator in the direction of the polarizer, which, in terms of the annihilation operators before the polarizer, is given by

$$\hat{a}_{1\theta_1} = \hat{a}_{1H} \cos\theta + \hat{a}_{1V} \sin\theta.$$

Similarly, we have

$$\hat{a}_{2\theta_2} = \hat{a}_{2H} \cos\theta + \hat{a}_{2V} \sin\theta.$$

And thus it can be shown that

$$P(\theta_1, \theta_2) = \frac{1}{2} \cos^2(\theta_1 - \theta_2).$$

The correlation function for quantum mechanics now becomes

$$\begin{aligned} C(\theta_1, \theta_2) &= P(\theta_1, \theta_2) + P(\theta_1 + \pi/2, \theta_2 + \pi/) - P(\theta_1 + \pi/, \theta_2) - P(\theta_1, \theta_2 + \pi/) \\ &= \frac{1}{2} \cos^2(\theta_1 - \theta_2) + \frac{1}{2} \cos^2(\theta_1 - \theta_2) - \frac{1}{2} \sin^2(\theta_1 - \theta_2) - \frac{1}{2} \sin^2(\theta_1 - \theta_2) \\ &= \cos 2(\theta_1 - \theta_2) \end{aligned}$$

Now, if we take the following set of values: $\theta_1 = 0$, $\theta_2 = 3\pi/8$, $\theta'_1 = -\pi/4$, and $\theta'_2 = \pi/8$, it can be shown that

$$S \equiv |C(\theta_1, \theta_2) - C(\theta_1, \theta'_2)| + |C(\theta'_1, \theta_2) + C(\theta'_1, \theta'_2)| = 2\sqrt{2}.$$

Thus we see that the quantum mechanical wavefunction does violate Bell inequality. And, therefore, a hidden variable description of quantum mechanics is not always possible. The first convincing test of the violations of Bell inequalities was performed by Aspect *et al.* [44, 45], using correlated photons produced in atomic cascade. Since then, using entangled pair of photons produced by parametric down-conversion, violations of the CHSH forms of Bell's inequality have been observed for various degrees of freedom including polarization [46, 47], phase and momentum [48], time and energy [30, 31, 49], spatial-parity [50], frequency [51] and OAM [52]. The results of these experiments have strongly confirmed the predictions of quantum mechanics and have provided very strong evidence against the local hidden variable interpretations of quantum mechanics.

LECTURE # 22

Entanglement Verification

Just as "bit" is a unit of classical information and refers to the state of a classical object that lives in a two dimensional space, "qubit" is the unit of information in quantum domain and refers to the state of a quantum object that lives in a two-dimensional Hilbert space. The essential feature that distinguishes a qubit from a classical bit is superposition. While a classical bit can have only two possible states, generally referred to as '0' and '1', a qubit can not only have the analogous states, $|0\rangle$ and $|1\rangle$, but also have the general superposition state, $\alpha|0\rangle + \beta|1\rangle$, where α and β are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$. In quantum information, one can have a more general "qudit" state which refers to the state of a quantum object that lives in a d-dimensional Hilbert space and can involve a general superposition state in a d-dimensional Hilbert space. As we already know, the other feature that distinguishes quantum information from classical information is entanglement. A large variety of quantum information protocols are based on entangled two-particle states, or in general two-qudit states. We will now study how one can make two-qubit states and verify whether or not a given quantum state is entangled.

There are two broad tests for verifying entanglement. One is through Bell-inequality violation and the other one is by doing EPR correlation measurements. The original Bell-inequality test was meant for verifying entanglement of two particles in situations in which each particle lives in a two-dimensional Hilbert space. However, more recently, Bell-inequalities have even been derived for higher dimensions. A Bell violation test involves two-photon interference measurement. The measurements need to be performed in more than one bases in order to calculate the Bell parameter S. If the Bell parameter $S > 2$, the state of the two particles is said to be entangled. The other method for verifying entanglement is by doing an EPR correlation measurements. This test is specifically suitable for entanglement of continuous variables as opposed to the Bell-inequality test which is suited only for the finite dimensional system. For EPR correlation measurements, one does not need to perform a two-photon interference experiment. What is required in this case are uncertainty measurements in two non-commuting bases. We will now look at entanglement in different degrees of freedom and see how the entanglement in a particular degree of freedom can be verified.

1. Polarization Entanglement: Polarization degree of freedom provides a two-dimensional basis. Therefore, polarization entanglement is only two dimensional, and what one can have using polarization entanglement is a two-qubit state only. For example, one can have the following two-qubit state in the polarization basis

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left[|H\rangle|H\rangle \pm |V\rangle|V\rangle\right]$$
$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left[|H\rangle|V\rangle \pm |V\rangle|H\rangle\right]$$

The above states are called the polarization Bell states. We know that entanglement guarantees simultaneous correlations in conjugate bases. For polarization entanglement, this means that $|\psi_+\rangle$ can also be written in the conjugate basis while still maintaining the two-qubit structure, that is, $|\psi_+\rangle = \frac{1}{\sqrt{2}}\left[|45\rangle|45\rangle + |-45\rangle|-45\rangle\right]$. The above two-qubit states can be experimentally generate using the process of PDC. The entanglement of such states can be proved by doing a Bell-inequality violation experiment. We note that we cannot perform an EPR-type experiment to verify polarization entanglement as it is just a two-dimensional basis.

2. Time-Energy Entanglement: This is an example of a continuous variable entanglement since both time and energy take up continuous range of values. Time-energy entangled states can be produced using the process of parametric down-conversion. The entanglement can be verified through an EPR type experiment. In the context of PDC, one needs to show that $\Delta t_s^{(i)} \Delta E_s^{(i)} < \frac{\hbar}{2}$, where $\Delta t_s^{(i)}$ and $\Delta E_s^{(i)}$ are conditional arrival-time and energy uncertainties of the signal photon given that the arrival-time and energy of the idler photon is already known. Although time-energy entanglement is intrinsically infinite dimensional, the degree of freedom can be used to make two-qubit states. Figure 20 shows how to make such a two-qubit state using Franson interferometer. The two-qubit state in this case can be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\tau_{s1}\rangle_s|\tau_{i1}\rangle_i + |\tau_{s2}\rangle_s|\tau_{i2}\rangle_i\right]$$

One can verify the entanglement of such two-qubit states by performing a Bell-inequality violation experiment. The above state involves superposition of two-photon states $|\tau_{s1}\rangle_s|\tau_{i1}\rangle_i$ and $|\tau_{s2}\rangle_s|\tau_{i2}\rangle_i$. We note that this can be ensured

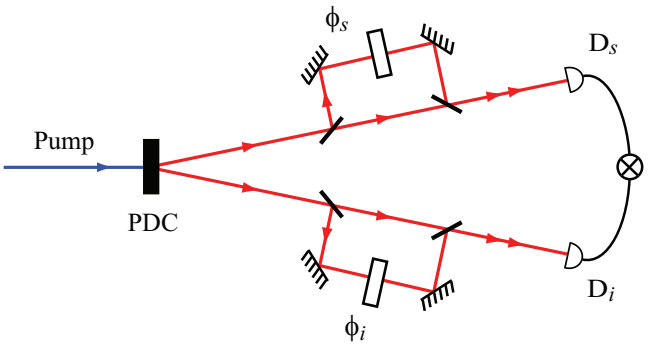


FIG. 20: Making two-qubit state using energy-time entanglement.

only if there is two-photon temporal coherence. We also note that, since time and energy form conjugate pairs, one could make a two-qubit state in the energy basis as well, that is, one can have

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\omega_1\rangle_s|\omega_2\rangle_i + |\omega_2\rangle_s|\omega_1\rangle_i\right],$$

where ω_1 and ω_2 are temporal frequencies.

3. Position-Momentum Entanglement: Just as time-energy, the position-momentum degree of freedom is also an example of continuous variable entanglement. Position-momentum entanglement can also be verified through an EPR-type experiment. For this, one has to calculate the conditional position and momentum uncertainties of the signal photon and prove that the uncertainty product violates Heisenberg uncertainty principle. Again, position-

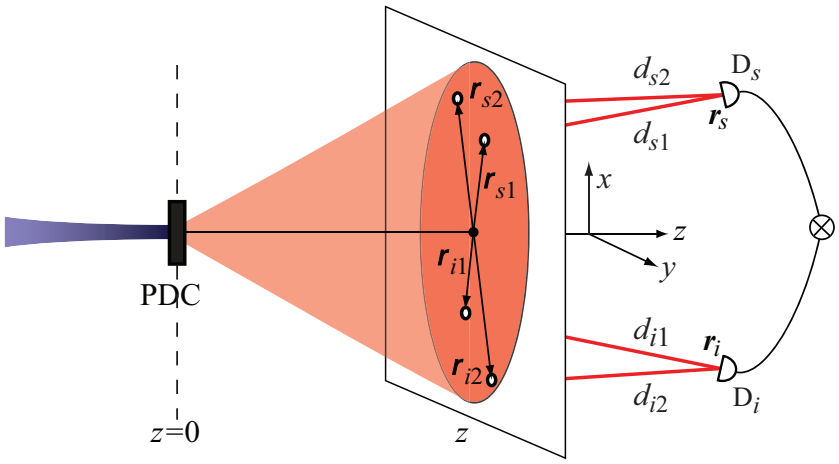


FIG. 21: Making two-qubit state using position-momentum entanglement.

momentum degree of freedom can also be used for defining a two-qubit state as shown in Figure 21. The two-qubit state can be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\boldsymbol{\rho}_{s1}\rangle_s|\boldsymbol{\rho}_{i1}\rangle_i + |\boldsymbol{\rho}_{s2}\rangle_s|\boldsymbol{\rho}_{i2}\rangle_i\right]$$

The production of the above state assumes two-photon spatial coherence. The entanglement of the state can be verified through the violation of a suitable Bell-inequality. Again, since position and momentum form conjugate pairs, one can alternatively make a two-qubit state in the momentum basis as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\boldsymbol{q}_1\rangle_s|\boldsymbol{q}_2\rangle_i + |\boldsymbol{q}_2\rangle_s|\boldsymbol{q}_1\rangle_i\right],$$

where $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are spatial frequencies.

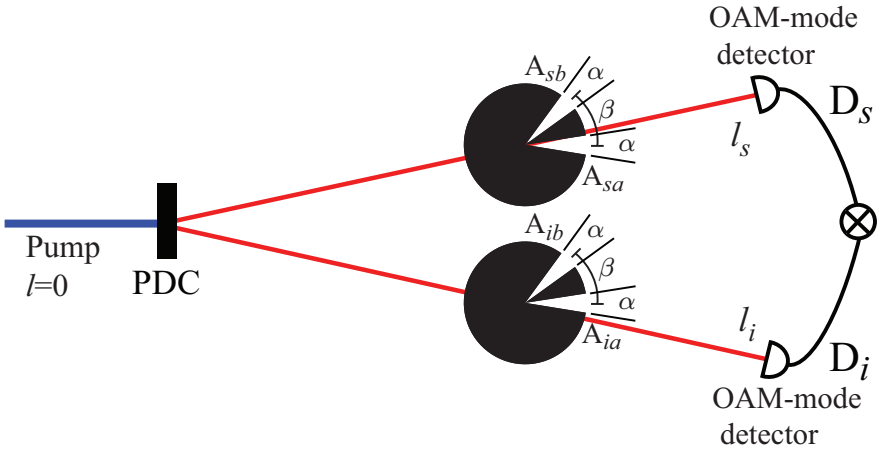


FIG. 22: Making two-qubit state using angular-position OAM entanglement.

4. Angular Position-Orbital Angular Momentum Entanglement: We now come to entanglement in angular position and OAM bases. This is also an example of a continuous variable entanglement. One can verify the entanglement in this case by doing an EPR-type measurement. One can also make a two-qubit state using the scheme shown in Figure 22. The state of the two qubits can be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|\theta_{s1}\rangle_s|\theta_{i1}\rangle_i + |\theta_{s2}\rangle_s|\theta_{i2}\rangle_i\right],$$

the entanglement of which can be verified through the violation of a Bell-inequality. Again, one can make a two-qubit state in the conjugate basis of orbital angular momentum as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[|l_1\rangle_s|l_2\rangle_i + |l_2\rangle_s|l_1\rangle_i\right],$$

where l_1 and l_2 are orbital angular momenta.

LECTURE # 23

Entanglement Quantification and connection between coherence and entanglement

In the last lectue, we saw how entanglement can be verified using either the Bell-inequality or EPR-correlations. We also summarized how two qubit states can be constructed for different degrees of freedom. Although Bell-inequality violation and EPR uncertainty measurement verify the presence of entanglement, they do not quantify the amount of entanglement present in a given two -qubit state. The field of entanglement quantification is a very active area of research and at this point one only knows how to uniquely quantify the entanglement of two-qubit states. In general, there is no unique way of quantifying multi-particle, high-dimensional entanglement. Let us first review the concepts of pure and mixed states in the context of one-particle systems.

Pure State (One Particle): In situations in which a one-particle system can be completely described using a single state vector $|\psi\rangle$, the one-particle system is said to be in a pure state. In other words, if a system can be represented by a single ensemble of realizations, then the system is said to be in a pure state. A pure state can be represented as a column vector, the dimensionality of which is the dimensionality of the Hilbert space in which the system lives, that is, $|\psi\rangle = \sum_i^N d_i|\phi_i\rangle$. An example of a pure state in two dimension is

$$|\psi\rangle = a|0\rangle + b|1\rangle = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix}$$

where $a^2 + b^2 = 1$.

Mixed State (One Particle): When a one-particle system cannot be described as a single ensemble of realizations and has to be described as a mixture of ensembles of realization then the one-particle state is said to be in a mixed state. Mixedness of a state in a given basis is directly related to the partial coherence of the state in that basis. The degree of mixedness or the degree of coherence is calculated by performing a coherent-mode representation of the state. A mixed state is represented by a density matrix ρ.

$$\rho = \langle|\psi\rangle\langle\psi|\rangle_e = \sum_{ij} d_i d_j^*|\phi_i\rangle\langle\phi_j|$$

The coherent mode representation of the state is obtained by diagonalizing ρ and can be represented as

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|,$$

If the above sum contains only one term then the state is pure.

Pure State (Two Particles): When a two-particle system can be described as a single ensemble of realizations then the state is said to be pure. A two-particle pure state can be written as

$$|\Psi\rangle = \sum_{ij} d_{ij}|\phi_i^{(a)}\rangle \otimes |\psi_j^{(b)}\rangle,$$

where $|\phi_i^{(a)}\rangle$ and $|\phi_j^{(b)}\rangle$ are the pure state-vectors corresponding to individual particles, and $\otimes$ represent the tensor product.

Mixed State (Two Particles): When a two-particle system can be described only as a mixture of ensembles of realizations, then the state is said to be mixed. The coherent mode representation or the diagonal representation of the state is given by

$$\rho = \sum_{ij}\sum_{mn} d_{ij}d_{mn}^*|\phi_i^{(a)}\rangle|\phi_j^{(b)}\rangle\langle\phi_m^{(a)}|\langle\phi_n^{(b)}| = \sum_i p_i|\Psi_i\rangle\langle\Psi_i| = \sum_i p_i\rho_i$$

If the above sum contains only one term then the two-particle state is a pure state. If the state contains more than one term then the state is mixed and hence partially coherent.

Now, the entanglement of two particles is not just about whether the two-particle state is pure or mixed. It is about whether or not a two-particle state is separable. If the joint state of the two particles cannot be written as a product of the states of the individual particles then the two-particle state is said to be inseparable. The degree of inseparability quantifies the amount of entanglement. However, if a two two-particle state is separable then it is not entangled. Next, we will see how this inseparability or entanglement is quantified. There are several methods for doing this quantification but we will only look at the ones that are more frequently used.

Entanglement of Formation: The inseparability of pure states can be uniquely quantified through Schmidt decomposition. A two-particle pure state can be represented as

$$|\Psi\rangle = \sum_{ij} d_{ij} |\phi_i^{(a)}\rangle |\psi_j^{(b)}\rangle,$$

where $|\phi_i^{(a)}\rangle$ and $|\psi_j^{(b)}\rangle$ represent the state-vector corresponding to the individual particles. The state $|\Psi\rangle$ can be represented in the Schmidt decomposed form as

$$|\Psi\rangle = \sum_{i} c_i |\tilde{\phi}_i^{(a)}\rangle |\tilde{\phi}_i^{(b)}\rangle,$$

with $\sum_i c_i^2 = 1$. Here $|\tilde{\phi}_i^{(a)}\rangle$ and $|\tilde{\psi}_j^{(b)}\rangle$ are the state vectors in the Schmidt basis. If the Schmidt decomposed form contains only one term then the state is separable but if it contains more than one term then the state is inseparable. As an example, let us take a two-particle pure state $|\Psi\rangle$ given by

$$|\Psi\rangle = \frac{|00\rangle + 2|01\rangle + |10\rangle + 2|11\rangle}{\sqrt{10}}.$$

This state can be written in the Schmidt decomposed form as

$$|\Psi\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{|0\rangle + 2|1\rangle}{\sqrt{5}}\right).$$

Since the Schmidt decomposed form contains only one term, we see that the state is separable and hence the entanglement of the state is zero.

Now, in case the state is not separable, the Schmidt coefficients c_i^2 should quantify the inseparability or the entanglement of the state. There can certainly be several definitions of entanglement based on the Schmidt coefficients but the one definition that has been adopted for the purpose is called the entanglement of formation and is defined as

$$E(\Psi) = -\sum_i c_i^2 \log_2 c_i^2.$$

We note that for a separable state we have $c_i^2 = 1$ and therefore $E(\Psi) = 0$, as expected.

It turns out that one does not always have to find the Schmidt decomposition as there are other ways to get the Schmidt coefficients. Let us take the partial trace of ρ over particle-b. We get

$$\begin{aligned}
\rho_a &= \mathrm{tr_b}\rho \\
&= \mathrm{tr_b} \sum_i \sum_j c_i c_j^* |\tilde{\phi}_i^{(a)}\rangle |\tilde{\phi}_i^{(b)}\rangle \langle\tilde{\phi}_j^{(a)}| \langle\tilde{\phi}_j^{(b)}| \\
&= \sum_i c_i^2 |\tilde{\phi}_i^{(a)}\rangle \langle\tilde{\phi}_i^{(a)}|.
\end{aligned}$$

We find that ρ_a is a diagonal matrix, with the diagonal terms being equal to the Schmidt coefficients of the pure two-particle state. So, in order to calculate the Schmidt coefficients, one only needs to calculate the reduced density matrix ρ_a (or ρ_b)of the system and take the diagonal entries. Now, since c_i^2 are the eigenvalues of the density matrix corresponding to particle-a, we find that the entanglement of formation is nothing but the von-Neumann entropy of particle-a. So, for a completely separable state the von Neumann entropy of the subsystems remain zero.

Now that we know how to define the entanglement of formation of a pure state, we move to the mixed state. We note that a two-photon mixed state ρ can be represented as a mixture of pure states. The diagonal representation of ρ,

$$\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| = \sum_i p_i \rho_i,$$

is one such decomposition, in which the individual pure states $|\Psi_i\rangle$ are orthogonal. However, in general, there can be several other decompositions of a mixed state, in terms of non-orthogonal pure states. The entanglement of formation of a mixed state is defined as the sum of the entanglement of formations of the pure state mixture that minimizes the sum, that is,

$$E(\rho) = \min \left\{ \sum_j p_j E(\Phi_j) \right\},$$

Here p_j are coefficients of the pure-state decompositions. In general, for a two-qudit state, it is very difficult to find out the decomposition that minimizes entanglement of formation. However, for a two-qubit state, it is possible to do so, which brings us to the second measure of entanglement known as "Concurrence."

Concurrence: Concurrence is another measure of entanglement which can be uniquely defined for two-qubit states [53]. Concurrence for a two-qubit mixed state is defined as follows: Take the two-qubit density matrix ρ and construct the density matrix $\tilde{\rho}$ using the spin flip operation

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y),$$

where $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is the Pauli's matrix. Construct a new matrix $\rho\tilde{\rho}$ and find its eigenvalues and arrange them in descending order $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. The concurrence of the two-qubit state is then given as

$$C(\rho) = \max\{0, \sqrt{\lambda}_1 - \sqrt{\lambda}_2 - \sqrt{\lambda}_3 - \sqrt{\lambda}_4\}.$$

LECTURE # 24

Quantum Cryptography

Cryptography is the art of combining a message with some additional information, called the key, such that the combined message becomes unintelligible to every person except to the intended receiver who has the key.

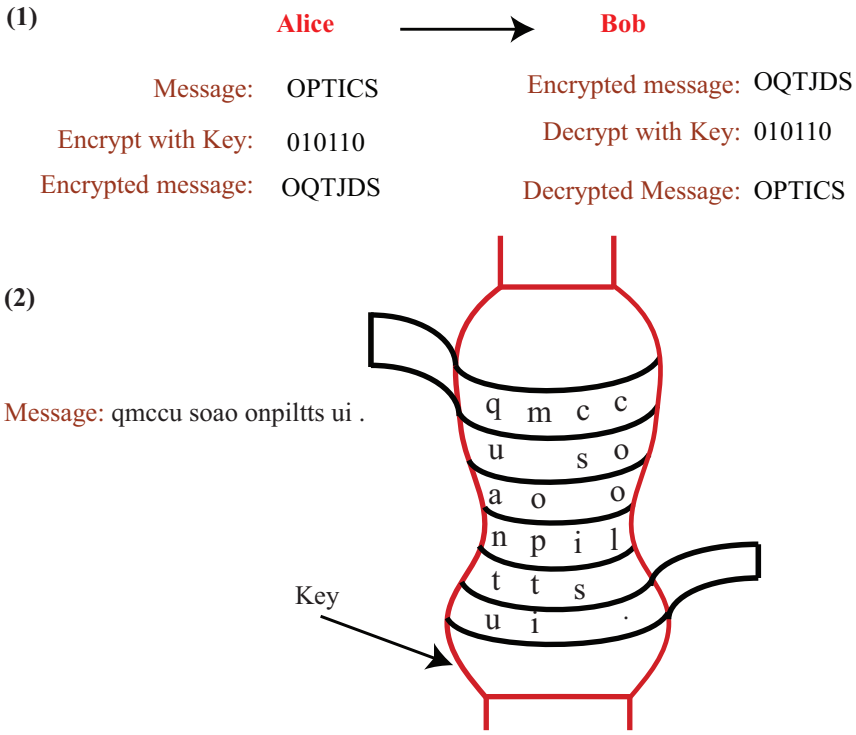


FIG. 23: Two examples of cryptography.

Cryptography is used for secure communication. A key is shared between the sender and the receiver; the sender encodes the message with the key and the receiver decodes the message with the same key. So, the main aim of cryptography is the secure distribution of the key between the sender and the receiver. In the quantum domain, the key distribution is achieved through the use of quantum principles and is therefore called quantum key distribution.

There are two generic protocols for doing quantum cryptography. One is called the BB84 protocol [54], which requires single photons and which utilizes the principle of superposition. The other protocol is called the Ekert protocol [4]; it requires a pair of photons that are entangled in a two-dimensional polarization basis. In the BB84 protocol, the sender (Alice) prepares a state in a randomly chosen basis and sends it to the receiver (Bob); Bob then measures the state sent by Alice in a separate, randomly-chosen basis. In Ekert protocol, Alice and Bob each receives one particle of the entangled pair. Alice then chooses a basis randomly and measures the state of her particle. After Alice's measurement, Bob also chooses a basis randomly and measures the state of his particle. The two protocols are very similar in the sense that Alice's "measuring a state" in the Ekert protocol is analogous to her "preparing a state" in the BB84 protocol. With this analogy in mind, we will keep the rest of our analysis and subsequent discussion centered on the BB84 protocol only.

Security in classical key distribution depends on the security of the communication channel. However, quantum key distribution is intrinsically secure because of the laws of quantum mechanics. If quantum key distribution succeeds, it is perfectly secure. So, what are the laws of quantum mechanics that makes quantum cryptography perfectly secure?

1. Measurement in an incompatible basis changes the quantum state. For example, If a horizontally polarized photon is measured in the H/V basis, the measurement will find with unit probability the state of the photon to be horizontally polarized and after the measurement the state of the photon will remain as horizontally polarized. On the other hand, if the photon is measured in the $45/-45$ basis, half the time the measurement will find the photon to be 45^0-polarized and the other half of the time the measurement will find the photon to be -45^0-polarized.

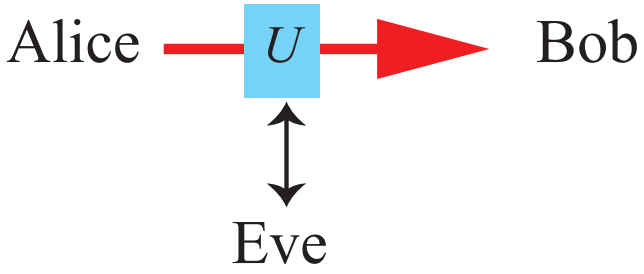


FIG. 24: Eavesdropping in BB84 Protocol.

2. No-cloning theorem: This theorem states that an arbitrary state of a photon cannot be cloned [55].
Proof: Suppose there is a unitary operator $\hat{U}$ that can clone a photon in the H/V basis, that is,

$$\hat{U}|H\rangle \rightarrow |HH\rangle, \tag{184}$$

$$\hat{U}|V\rangle \rightarrow |VV\rangle. \tag{185}$$

The same unitary operator, when applied to a 45^0-polarized photon, we get

$$\hat{U}(|H\rangle + |V\rangle) \rightarrow |HH\rangle + |VV\rangle \neq |H+V\rangle|H+V\rangle. \tag{186}$$

Thus, we see that an arbitrary quantum state cannot be cloned. So, even if the photon sent by Alice is captured by some eavesdropper, The eavesdropper cannot always make a clone of the photon and send the original photon to Bob. As a result any act of eavesdropping will result in an error in the state received by Bob and this will allow Alice and Bob to become aware of eavesdropping. This way quantum key distribution is perfectly secure.

Five steps of Quantum Cryptography

(1) Distribution of a " Raw key": The first step is to generate a key, called the "raw key". In the BB84 protocol, Alice prepares the state in a randomly chosen basis and sends the state to Bob, who then measures it in another randomly chosen basis. The sequence of states prepared by Alice and the sequence of states measured by Bob constitute the raw key. The raw key of course contains error and there is also a possibility that a potential eavesdropper may have some information on the key.

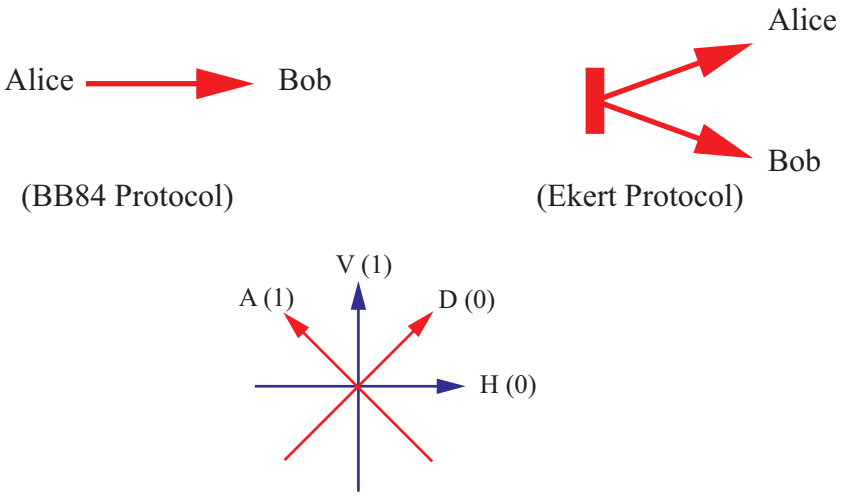
(2) Basis Reconciliation: The next step of quantum cryptography is called the basis-reconciliation. In this step both Alice and Bob publicly announce their bases for each set of prepared and measured states and after that keep only those measurement results for which both of them happen to use the same basis. The key obtained after the basis-reconciliation is called the "sifted key." In the absence of an eavesdropper, the sifted key contains no error because for each set both Alice and Bob use the same basis and so their measurement values are the same. However, the presence of an eavesdropper in the system introduces error in the sifted key, because a measurement act by an eavesdropper alters the state sent out by Alice and thus causes an error in Bob's measurement even when Bob is measuring in the same basis used by Alice to prepare the state. This fact is because of the no-cloning theorem of quantum mechanics [55], which dictates that an arbitrary quantum states cannot be copied. And in fact it is this feature of quantum mechanics that makes quantum cryptography 100% secure, at least in principle.

(3) Error Estimation: The third step is the error estimation, the goal of which is to estimate the error in the sifted key. In this step, both Alice and Bob select a random set of measurement results and then compare them publicly. If they find the error to be below the maximum allowed value, they proceed to steps four and five of the protocol, otherwise they simply abandon the protocol and start over again.

(4) Error correction: If the error obtained in the third step is below the allowed limit then the fourth step in quantum cryptography is the error correction. In this step Alice and Bob use different techniques to minimize the error in the sifted key. The length of the key obtained after error correction is invariably smaller than that of the sifted key. Although the key obtained after error correction is nearly error-free, there is still a possibility that an eavesdropper (Eve) has some information about the remaining key shared by Alice and Bob. In fact, most error-correcting codes end up increasing Eve's information on the remaining key.

(5) Privacy amplification: The fifth and final step in quantum cryptography is the privacy amplification, the aim of which is to minimize Eve's information on the remaining key.

The fourth and fifth steps are common to every cryptographic protocol, irrespective of whether it is quantum cryptography or the present-day, classical cryptography. Therefore, the main goal of quantum cryptography is to generate a sifted key with an error that is below the maximum allowed value. The hallmark of quantum cryptography is that once it is successfully completed, one can be absolutely sure that no eavesdropper has any information on the shared key.



Alice's Bases	DA	HV	DA	HV	HV	HV	DA	DA	HV	HV	DA
Alice's random bits	1	0	0	1	0	1	1	0	0	1	0
Sifted bits		0			0		1			1	

Bob's Bases	HV	HV	HV	DA	HV	DA	DA	HV	DA	HV	HV
Bob's random bits	0	0	1	1	0	0	1	0	1	1	1
Sifted bits		0			0		1			1	

FIG. 25: Quantum Cryptography: BB84 and Ekert Protocols.

The BB84 and the Ekert protocols are explained in Figure 25. In the BB84 protocol, Alice chooses a one-photon system that can exist in two-dimensional bases. Here, we are working in the polarization basis. Alice and Bob decide to work in two mutually-unbiased bases (MUB), H/V and D/A. Alice randomly chooses between the two mutually unbiased bases and sends either a '0' or a '1' to Bob. After receiving the photon, Bob then randomly chooses one of the two mutually unbiased bases and makes the measurement. If Alice and Bob were using Ekert protocol then they share an entangled state given by

$$\begin{aligned} |\psi\rangle &= |H\rangle|H\rangle + |V\rangle|V\rangle \\ &= |D\rangle|D\rangle + |A\rangle|A\rangle \end{aligned}$$

In both cases, for the example shown in the figure, the sifted key is '0011'. The original proposal by Bennett and Brassard [54] was based on using two-dimensional variable, for example, polarization. But more recently cryptography with higher-dimensional variables is gaining increased attention and the BB84 protocol for two-dimensional variables has now been extended to higher-dimensional variables [56]. In this context, we note that the orbital angular momentum states of a photon span a countably infinite-dimensional Hilbert space and are therefore a very suitable variable for high-dimensional quantum cryptography [**?**]. One of the main advantages of high-dimensional quantum cryptography is the increased value of the maximum allowed transmission error, which for the two-dimensional case is about only 17%.

LECTURE # 25

Quantum Teleportation

The no-cloning theorem prohibits generation of an exact copy of an arbitrary state of a particle [55]. However, using quantum entanglement it is possible to at least transfer an arbitrary state of a particle without actually transferring the particle itself. This is known as quantum teleportation [6] and works as follows(Figure 26). Alice and Bob share an entangled pair of photons, which are entangled in the polarization basis. The two particles can be in any of the four Bell-states given by

$$|\Phi^{\pm}\rangle_{\rm AB} = \frac{1}{\sqrt{2}}\left[|H\rangle_{\rm A}|H\rangle_{\rm B} \pm |V\rangle_{\rm A}|V\rangle_{\rm B}\right]$$
$$|\Psi^{\pm}\rangle_{\rm AB} = \frac{1}{\sqrt{2}}\left[|H\rangle_{\rm A}|V\rangle_{\rm B} \pm |V\rangle_{\rm A}|H\rangle_{\rm B}\right]$$

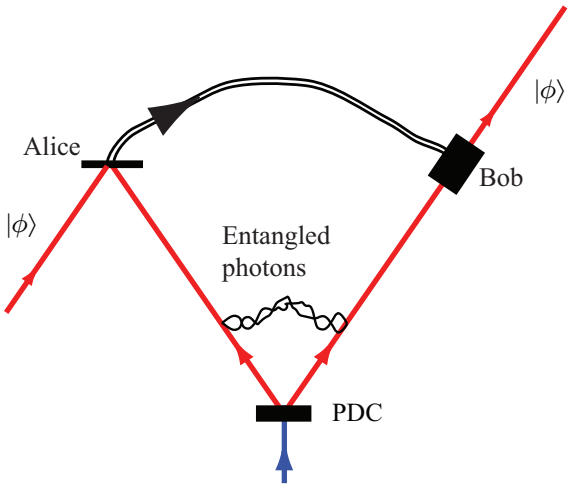


FIG. 26: Quantum Teleportation

Let us assume that the entangled particles are in following Bell state:

$$|\Psi^{-}\rangle_{\rm AB} = \frac{1}{\sqrt{2}}\left[|H\rangle_{\rm A}|V\rangle_{\rm B} - |V\rangle_{\rm A}|H\rangle_{\rm B}\right].$$

Now suppose that Alice has another particle with an unknown state $|\phi\rangle_{\rm C}$ that she wants to teleport to Bob. Alice does not know the quantum state of her particle, neither can she send the particle to Bob. But she is still able to send the unknown state to Bob, in the following way. We write the state of Alice's particle as

$$|\phi\rangle_{\rm C} = \alpha|H\rangle_{\rm C} + \beta|V\rangle_{\rm C} = \alpha\begin{pmatrix}1\\0\end{pmatrix} + \beta\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\alpha\\\beta\end{pmatrix},$$

with $|\alpha|^2 + |\beta|^2 = 1$. We note that the combined state of the two entangled particles and particle-C can be written as a pure tensor product of $|\Psi^{-}\rangle_{\rm AB}$ and $|\phi\rangle_{\rm C}$ and is given as

$$\begin{aligned}|\Psi\rangle_{\rm ABC} &= |\Psi^{-}\rangle_{\rm AB}|\phi\rangle_{\rm C}\\ &= \frac{1}{\sqrt{2}}\left(|H\rangle_{\rm A}|V\rangle_{\rm B} - |V\rangle_{\rm A}|H\rangle_{\rm B}\right)\left(\alpha|H\rangle_{\rm C} + \beta|V\rangle_{\rm C}\right)\\ &= \frac{\alpha}{\sqrt{2}}|H\rangle_{\rm A}|H\rangle_{\rm C}|V\rangle_{\rm B} - \frac{\alpha}{\sqrt{2}}|V\rangle_{\rm A}|H\rangle_{\rm C}|H\rangle_{\rm B} + \frac{\beta}{\sqrt{2}}|H\rangle_{\rm A}|V\rangle_{\rm C}|V\rangle_{\rm B} - \frac{\beta}{\sqrt{2}}|V\rangle_{\rm A}|V\rangle_{\rm C}|H\rangle_{\rm B}. \qquad (187)\end{aligned}$$

Alice has two particles, particle-A and particle-C. She performs a Bell state-measurement on these particle. This simply means that she wants to see which Bell state her two particles are in. The Bell-basis representing particle-A

and particle-C are given as

$$|\Phi^{\pm}\rangle_{\rm AC} = \frac{1}{\sqrt{2}}\left[|H\rangle_{\rm A}|H\rangle_{\rm C} \pm |V\rangle_{\rm A}|V\rangle_{\rm C}\right],$$
$$|\Psi^{\pm}\rangle_{\rm AC} = \frac{1}{\sqrt{2}}\left[|H\rangle_{\rm A}|V\rangle_{\rm C} \pm |V\rangle_{\rm A}|H\rangle_{\rm C}\right].$$

The above relations yield

$$|H\rangle_{\rm A}|H\rangle_{\rm C} = \frac{1}{\sqrt{2}}\left(|\Phi^{+}\rangle_{\rm AC} + |\Phi^{-}\rangle_{\rm AC}\right),$$
$$|V\rangle_{\rm A}|H\rangle_{\rm C} = \frac{1}{\sqrt{2}}\left(|\Psi^{+}\rangle_{\rm AC} - |\Psi^{-}\rangle_{\rm AC}\right),$$
$$|H\rangle_{\rm A}|V\rangle_{\rm C} = \frac{1}{\sqrt{2}}\left(|\Psi^{+}\rangle_{\rm AC} + |\Psi^{-}\rangle_{\rm AC}\right),$$
$$|V\rangle_{\rm A}|V\rangle_{\rm C} = \frac{1}{\sqrt{2}}\left(|\Phi^{+}\rangle_{\rm AC} - |\Phi^{-}\rangle_{\rm AC}\right).$$

Using the above relations Eq. (187) can be written as

$$|\Psi\rangle_{\rm ABC} = \frac{\alpha}{2}\left(|\Phi^{+}\rangle_{\rm AC} + |\Phi^{-}\rangle_{\rm AC}\right)|V\rangle_{\rm B} - \frac{\alpha}{2}\left(|\Psi^{+}\rangle_{\rm AC} - |\Psi^{-}\rangle_{\rm AC}\right)|H\rangle_{\rm B} + \frac{\beta}{2}\left(|\Psi^{+}\rangle_{\rm AC} + |\Psi^{-}\rangle_{\rm AC}\right)|V\rangle_{\rm B} - \frac{\beta}{2}\left(|\Phi^{+}\rangle_{\rm AC} - |\Phi^{-}\rangle_{\rm AC}\right)|H\rangle_{\rm B}. \quad (188)$$

The above equation can be rewritten as

$$|\Psi\rangle_{\rm ABC} = \frac{1}{2}|\Phi^{+}\rangle_{\rm AC}\left(\alpha|V\rangle_{\rm B} - \beta|H\rangle_{\rm B}\right) + \frac{1}{2}|\Phi^{-}\rangle_{\rm AC}\left(\alpha|V\rangle_{\rm B} + \beta|H\rangle_{\rm B}\right) + \frac{1}{2}|\Psi^{+}\rangle_{\rm AC}\left(-\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) + \frac{1}{2}|\Psi^{-}\rangle_{\rm AC}\left(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) \quad (189)$$

We find that particle-A and particle-C are equally likely to be found in any of the four Bell states and the probability is 1/4. So, after Alice has done the Bell-state analysis she sends that information to Bob and based on that information Bob deicides on a unitary transformation which guarantees that he has the original state. Here is how it is done:

(1) Alice's measurement gives her the Bell state $|\Phi^{+}\rangle_{\rm AC}$. The state of Bob's qubit will be $(\alpha|V\rangle_{\rm B} - \beta|H\rangle_{\rm B})$. So, Bob makes the unitary transformation given by $\hat{U}_{\phi^{+}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ such that the state of Bob's qubit becomes

$$U_{\phi^{+}}\left(\alpha|V\rangle_{\rm B} - \beta|H\rangle_{\rm B}\right) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -\beta \\ \alpha \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \left(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) \quad (190)$$

(2) Alice's measurement gives her the Bell state $|\Phi^{-}\rangle_{\rm AC}$. The state of Bob's qubit will be $(\alpha|V\rangle_{\rm B} + \beta|H\rangle_{\rm B})$. So, Bob makes the unitary transformation given by $\hat{U}_{\phi^{-}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ such that the state of Bob's qubit becomes

$$U_{\phi^{-}}\left(\alpha|V\rangle_{\rm B} + \beta|H\rangle_{\rm B}\right) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \left(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) \quad (191)$$

(3) Alice's measurement gives her the Bell state $|\Psi^{+}\rangle_{\rm AC}$. The state of Bob's qubit will be $(-\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B})$. So, Bob makes the unitary transformation given by $\hat{U}_{\psi^{+}} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ such that the state of Bob's qubit becomes

$$U_{\psi^{+}}\left(-\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -\alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \left(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}\right) \quad (192)$$

(4) Alice's measurement gives her the Bell state $|\Psi^-\rangle_{\rm AC}$. The state of Bob's qubit will be $(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B})$. So, Bob makes the unitary transformation given by $\hat{U}_{\psi^-} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ such that the state of Bob's qubit becomes

$$U_{\psi^-}\,(\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (\alpha|H\rangle_{\rm B} + \beta|V\rangle_{\rm B}) \tag{193}$$

Thus, we see that Bob gets the same state $|\psi\rangle$ that Alice intended to teleport. The quantum teleportation protocol that we have outlined is for a qubit, that is, for a photon in a two-dimensional Hilbert space. But in general, one can have a quantum teleportation protocol in an N-dimensional Hilbert space.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] E. Schrödinger, Naturewiss **23**, 807 (1935).
[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Reviews of Modern Physics **81**, 865 (pages 78) (2009), URL http://link.aps.org/abstract/RMP/v81/p865.
[4] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[5] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[7] B. Yurke and D. Stoler, Phys. Rev. Lett. **68**, 1251 (1992).
[8] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, Phys. Rev. Lett. **85**, 2733 (2000).
[9] B. Jack, M. Padgett, and S. Franke-Arnold, New Journal of Physics **10**, 103013 (2008).
[10] A. K. Jha, B. Jack, E. Yao, J. Leach, R. W. Boyd, G. S. Buller, S. M. Barnett, S. Franke-Arnold, and M. J. Padgett, Phys. Rev. A **78**, 043810 (pages 4) (2008), URL http://link.aps.org/abstract/PRA/v78/e043810.
[11] L. Allen, M. Beijersbergen, R. Spreeuw, and J. Woerdman, Phys. Rev. A **45**, 8185 (1992).
[12] M. W. Beijersbergen, R. P. C. Coerwinkel, M. Kristensen, and J. P. Woerdman, Opt. Comm. **112**, 321 (1994), ISSN 0030-4018, URL http://www.sciencedirect.com/science/article/B6TVF-46JGS4X-F/2/7668ad6e50d8b171195ddb0ef9e7e304.
[13] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, Nature **412**, 313 (2001).
[14] A. K. Spilman and T. G. Brown, Appl. Opt. **46**, 61 (2007), URL http://ao.osa.org/abstract.cfm?URI=ao-46-1-61.
[15] R. W. Boyd, *Nonlinear Optics* (Academic Press, New York, 2003), 2nd ed.
[16] N. Bloembergen, J. Opt. Soc. Am. **70**, 1429 (1980), URL http://www.opticsinfobase.org/abstract.cfm?URI=josa-70-12-1429.
[17] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, New York, USA, 2000), 3rd ed.
[18] Z. Y. Ou, L. J. Wang, and L. Mandel, Phys. Rev. A **40**, 1428 (1989).
[19] W. P. Grice and I. A. Walmsley, Phys. Rev. A **56**, 1627 (1997).
[20] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
[21] T. B. Pittman, D. V. Strekalov, A. Migdall, M. H. Rubin, A. V. Sergienko, and Y. H. Shih, Phys. Rev. Lett. **77**, 1917 (1996).
[22] T. J. Herzog, J. G. Rarity, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **72**, 629 (1994).
[23] X. Y. Zou, L. J. Wang, and L. Mandel, Phys. Rev. Lett. **67**, 318 (1991).
[24] J. Brendel, W. Dultz, and W. Martienssen, Phys. Rev. A **52**, 2551 (1995).
[25] R. Bhandari and J. Samuel, Phys. Rev. Lett. **60**, 1211 (1988).
[26] A. K. Jha, M. N. O'Sullivan, K. W. C. Chan, and R. W. Boyd, Phys. Rev. A **77**, 021801(R) (2008).
[27] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).
[28] T. E. Keller and M. H. Rubin, Phys. Rev. A **56**, 1534 (1997).
[29] M. H. Rubin, D. N. Klyshko, Y. H. Shih, and A. V. Sergienko, Phys. Rev. A **50**, 5122 (1994).
[30] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Phys. Rev. A **47**, R2472 (1993).
[31] J. Brendel, E. Mohler, and W. Martienssen, Phys. Rev. Lett. **66**, 1142 (1991).
[32] P. Grangier, Nature **419**, 577 (2002).
[33] D. V. Strekalov, T. B. Pittman, and Y. H. Shih, Phys. Rev. A **57**, 567 (1998).
[34] M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, 1999), 7th ed.
[35] Y.-H. Kim, M. V. Chekhova, S. P. Kulik, Y. Shih, and M. H. Rubin, Phys. Rev. A **61**, 051803 (2000).
[36] C. K. Hong and L. Mandel, Phys. Rev. A **31**, 2409 (1985).
[37] C. H. Monken, P. H. S. Ribeiro, and S. Pádua, Phys. Rev. A **57**, 3123 (1998).
[38] B. E. A. Saleh, M. C. Teich, and A. V. Sergienko, Phys. Rev. Lett. **94**, 223601 (2005).
[39] D. Bohm, Phys. Rev. **85**, 166 (1952).
[40] D. Bohm, Phys. Rev. **85**, 180 (1952).
[41] J. S. Bell, Physics **1**, 195 (1964).
[42] P. G. Kwiat and L. Hardy, American Journal of Physics **68**, 33 (2000), URL http://link.aip.org/link/?AJP/68/33/1.
[43] L. Hardy, Contemporary physics **39**, 419 (1998).
[44] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981).
[45] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).
[46] Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988).
[47] Y. H. Shih and C. O. Alley, Phys. Rev. Lett. **61**, 2921 (1988).
[48] J. G. Rarity and P. R. Tapster, Phys. Rev. Lett. **64**, 2495 (1990).
[49] J. D. Franson, Phys. Rev. Lett. **62**, 2205 (1989).
[50] T. Yarnall, A. Abouraddy, B. Saleh, and M. Teich, Phys. Rev. Lett. **99**, 170408 (2007).
[51] S. Ramelow, L. Ratschbacher, A. Fedrizzi, N. Langford, and A. Zeilinger, Arxiv preprint arXiv:0904.4916 (2009).
[52] J. Leach, B. Jack, J. Romero, M. Ritsch-Marte, R. W. Boyd, A. K. Jha, S. M. Barnett, S. Franke-Arnold, and M. J. Padgett, Opt. Express **17**, 8287 (2009), URL http://www.opticsexpress.org/abstract.cfm?URI=oe-17-10-8287.
[53] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[54] C. H. Bennett and G. Brassard (1984).

[55] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
[56] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, Phys. Rev. Lett. **88**, 127902 (2002).