

Lecture 12. Entanglement and Bell's inequality violation.

1. The EPR paradox and entangled states

This last lecture explains why quantum optics and nonclassical light are important from most fundamental viewpoint. Namely, using the methods of quantum optics one can demonstrate that quantum mechanics is probabilistic. And this probabilistic behavior is typical not only for ensembles of quantum particles (which is also valid for classical particles) but also for a single quantum particle. This was the point of disagreement between Einstein and Bohr during the early years of quantum theory, about 80 years ago.

Continuous variables and an example from SPDC. In their paper of 1935, Einstein, Podolsky and Rosen (EPR), in the discussion with Bohr, considered an example that, from their viewpoint, showed that the quantum mechanical description was incomplete. Namely, they considered a pair of particles A,B created at some point at the same time moment, so that the conservation of momentum led to equations for their coordinates and momenta:

$$x_A - x_B = 0, \quad p_A + p_B = 0.$$

Then, after the particles have separated by a large distance, one can perform measurements on one of them, and hence bring it into a state with definite x or p. How then, asked EPR, the other particle is brought into a state with fixed x and p without any action on it? Does the reduction of its wavefunction occur non-locally? This led to the conclusion that the description with the wavefunction is incomplete. Moreover, one can measure the coordinate for one particle and the momentum for the other one and then violate the uncertainty relation. Such particles were called entangled (*verschränkt*); mathematically, they can be defined by the condition that the total wavefunction cannot be factorized:

$$|\Psi\rangle \neq |\Psi_A\rangle |\Psi_B\rangle.$$

Note that at the time of the argument, this example was purely a *Gedankenexperiment*. But nowadays, one can prepare such a state via SPDC (Fig.1). If the parametric gain is small, the state generated in the non-collinear regime is

$$|\Psi\rangle = |0\rangle_A |0\rangle_B + c |1\rangle_A |1\rangle_B, \quad |c| \ll 1.$$

This is definitely an entangled state since the wavefunction is not a product of the wavefunctions of particles A, B. But this is still a simplification: there are not just two plane-wave modes A, B but a whole spectrum of k-vectors and frequencies. For simplicity, let us omit the frequencies, then the two-photon part of the wavefunction has the form

$$|\Psi\rangle = \iint dk_A dk_B F(k_A, k_B) a^\dagger(k_A) a^\dagger(k_B) |0\rangle_A |0\rangle_B.$$

Here, k_A and k_B are transverse k-vectors (momenta) of the photons. If we look at the shape of the two-photon amplitude $F(k_A, k_B)$, we will see that it is stretched (Fig.2), i.e., the transverse wavevectors (momenta) of the two photons are anti-correlated. Similarly, the transverse coordinates of their exit from the crystal will be correlated. This represents the nowadays available example of a state entangled in continuous variables.

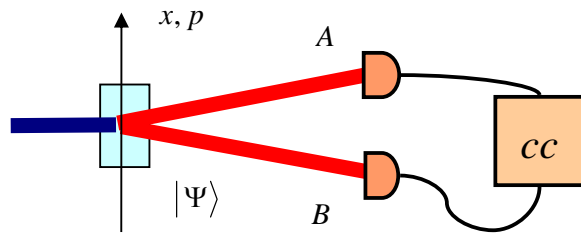


Fig.1

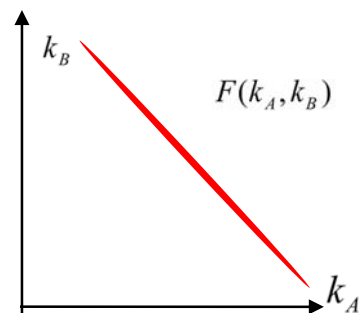


Fig.2

Dichotomic variables and an example from SPDC. For years, the EPR paradox remained just a subject of philosophic discussions. A breakthrough was made when Bohm proposed to formulate it in terms of dichotomic (binary) variables, such as the projection of the spin of a spin-1/2 particle on the direction of the magnetic field. This quantity is known to be measurable in the Stern-Gerlach experiment (Fig.3, top). One can then imagine two spin-1/2 particles born in such a way that each of them has the spin direction uncertain, but there is strict correlation between the spin directions of the two particles. The corresponding wavefunction can be written as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \{ |\uparrow\rangle_A |\downarrow\rangle_B + |\downarrow\rangle_A |\uparrow\rangle_B \}.$$

This state is also entangled.

The dichotomic version of the EPR paradox is important for two reasons. First, it allows for an easy derivation of an inequality (Bell's inequality) that can be tested in experiment. Second, in an experiment one can use, instead of spin 1/2 particles, polarized single photons (see lecture 8). Polarization of a photon can be measured similarly to the spin direction (Fig.3, bottom), using a polarization beamsplitter instead of a large complicated setup with magnets.

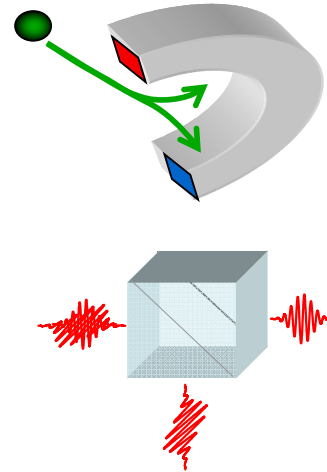


Fig.3

Polarization-entangled photons can be also produced via SPDC, but this requires a special type of phasematching. It is called type-II phasematching and it implies that the daughter photons are polarized orthogonally: one like an ordinary beam, the other one extraordinary. (The pump usually has extraordinary polarization). The situation is shown in Fig.4: o and e radiation is emitted into two cones displaced along the optic axis. The cones intersect along two lines (shown by arrows). In these directions, polarization can be either o (horizontal) or e (vertical), and the two-photon part of the state emitted via SPDC can be written as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \{ |H\rangle_A |V\rangle_B + |V\rangle_A |H\rangle_B \}.$$

This type of a state is called a Bell state, for the reasons that will be clear soon, and there are four of them:

$$\begin{aligned} |\Psi^{(+)}\rangle &= \frac{1}{\sqrt{2}} \{ |H\rangle_A |V\rangle_B + |V\rangle_A |H\rangle_B \}, & |\Psi^{(-)}\rangle &= \frac{1}{\sqrt{2}} \{ |H\rangle_A |V\rangle_B - |V\rangle_A |H\rangle_B \}, \\ |\Phi^{(+)}\rangle &= \frac{1}{\sqrt{2}} \{ |H\rangle_A |H\rangle_B + |V\rangle_A |V\rangle_B \}, & |\Phi^{(-)}\rangle &= \frac{1}{\sqrt{2}} \{ |H\rangle_A |H\rangle_B - |V\rangle_A |V\rangle_B \}. \end{aligned}$$

2. Bell's inequality

The discussion between EPR and Bohr became possible to test in experiment after Bell formulated in 1964 his famous theorem. This theorem, in the form of Clauser-Horne-Shimony-Holt inequality, can be derived using almost primary-school algebra, which we will do right now.

Let the state of particles A, B be described by dichotomic variables A, B, each taking values 1 or -1. We will imagine that they are measured in an experiment as shown in Fig.3: if upper

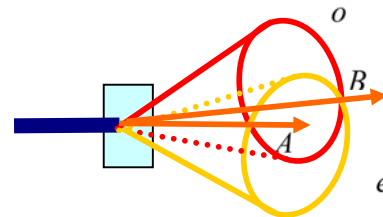


Fig.4

detector clicks, then $A=1$, and if lower detector clicks, then $A=-1$. For the measurement of B we have a different setup. Moreover, the same setup in a different configuration can measure other variables: A' for particle A, and B' for particle B. A' and B' are also dichotomic and take values $+1$ or -1 . One can imagine that for the measurement of A , the setups in Fig.3 look as they are; for the measurement of A' , they should be rotated. (How they should be rotated, we will see later.)

The inequality is derived from the assumption that, as soon as particles A, B are created, each of them has certain parameters: particle A has parameters A, A' and particle B has parameters B, B' . For instance, for a spin-1/2 particle these can be projections of the spin on the horizontal and vertical axes. These parameters can be called *hidden variables*: their set $\{\lambda\} \equiv \{A, A', B, B'\}$ is assumed to exist, and its distribution is given by some probability $p(\{\lambda\}) \geq 0$, $\int d\lambda p(\{\lambda\}) = 1$. In Bohr's formulation (the so-called *Copenhagen picture*) of quantum mechanics, there are no a priori values of $\{\lambda\} \equiv \{A, A', B, B'\}$, i.e., there are no hidden variables. So let us see where the *hidden-variable* assumption leads us. Of course there is another assumption used here, namely, *locality*: there is no nonlocal influence of particle A on particle B, and vice versa.

So, let us introduce a new variable,

$$F \equiv \frac{1}{2} \{AB + AB' + A'B - A'B'\} = \frac{1}{2} \{A(B + B') + A'(B - B')\}.$$

Because $B, B' = \pm 1$, either $B = B'$ or $B = -B'$, and only one of the brackets is nonzero. Then, because $|A| = |A'| = 1$, F can only take values $+1$ or -1 . So F is also dichotomic. Then let us look at its mean value. If we try to average F , every measurement will bring us either $+1$ or -1 , and the absolute value of the average should be below 1 (Fig.5). Or, using the probability distribution,

$$\langle F \rangle = \int p(\{\lambda\}) F(\{\lambda\}) d\lambda;$$

$$|\langle F \rangle| \leq \int p(\{\lambda\}) |F(\{\lambda\})| d\lambda = 1.$$

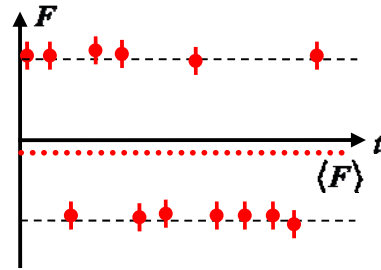


Fig.5

In other words, from the assumptions of (1) hidden variables and (2) locality it follows that $|\langle F \rangle| \leq 1$. This is the Bell inequality derived by Clauser, Horne, Shimony, Holt, so-called CHSH inequality. It is one of the numerous formulations of Bell's inequality. And, like others, it can be tested in experiment.

3. Violation of Bell's inequality for entangled photons from SPDC

Let us first show that a quantum mechanical calculation leads to the violation of the CHSH inequality. For this we have to choose the dichotomic variables. Let us consider a setup like shown in Fig.6: a nonlinear crystal, via SPDC, emits photon pairs in directions A, B. In each beam we have a setup with the polarizing prism performing the measurement of values A, A' (top) and B, B' (bottom). We will define the variables as follows.

1) Measurement of A. The upper polarizing prism is oriented in such a way that it transmits H polarization and reflects V polarization. Then, if the red detector (A_1) clicks, we

say that $A=1$; if the green one (A2) clicks, we say that $A=-1$. Altogether, it means that the value A is simply the first Stokes operator for photon A:

$\hat{A} \equiv \hat{S}_1^A = a_H^+ a_H - a_V^+ a_V$ (we are putting a hat because in quantum mechanics the measurement is described by operators).

2) Measurement of A' . The procedure is the same with the only difference that the prism is 45 degrees rotated. Note that we cannot do this measurement simultaneously with the first one. Then, the value A' is given by the second Stokes operator for photon A:

$$\hat{A}' \equiv \hat{S}_2^A = a_H^+ a_V + a_V^+ a_H.$$

3) Measurement of B. For the second setup, we would like to introduce more complicated measurements:

$$\hat{B} \equiv \frac{1}{\sqrt{2}} [\hat{S}_1^B + \hat{S}_2^B].$$

Such a measurement, as one can see by considering the Poincare sphere, can be performed by the setup in Fig.3 rotated by 22.5 degrees. Note that we choose such tricky orientations because we want to violate the inequality. And one can guess the orientation from certain physical requirements. Namely, we are choosing non-commuting variables corresponding to A and A' , to B and B' . From this, one can guess that for

4) measurement of B' , we introduce

$$\hat{B}' \equiv \frac{1}{\sqrt{2}} [\hat{S}_1^B - \hat{S}_2^B].$$

The operator \hat{F} will then take the form

$$\hat{F} = \frac{1}{2} \{ \hat{A}(\hat{B} + \hat{B}') + \hat{A}'(\hat{B} - \hat{B}') \} = \frac{1}{\sqrt{2}} \{ \hat{S}_1^A \hat{S}_1^B + \hat{S}_2^A \hat{S}_2^B \}.$$

Let us now calculate the mean value of \hat{F} for averaging over the singlet Bell state $|\Psi^{(-)}\rangle$.

Actually, the inequality will be violated for all four Bell states, but different operators \hat{F} have to be chosen. With the current one, the inequality will be violated for only $|\Psi^{(-)}\rangle$ and $|\Psi^{(+)}\rangle$.

As the first step, let us show that $|\Psi^{(-)}\rangle$ is an eigenstate of \hat{F} . Indeed,

$$\hat{F}|\Psi^{(-)}\rangle = \frac{1}{\sqrt{2}} \{ \hat{S}_1^A \hat{S}_1^B + \hat{S}_2^A \hat{S}_2^B \} \frac{1}{\sqrt{2}} \{ |H\rangle_A |V\rangle_B - |V\rangle_A |H\rangle_B \}.$$

We notice that the operators $\hat{S}_{1,2}^A$ act only on photon A, and the operators $\hat{S}_{1,2}^B$, only on photon B. Moreover, the states $|H\rangle, |V\rangle$ are eigenstates of \hat{S}_1 , with the eigenvalues 1,-1, while \hat{S}_2 acts on them as $\hat{S}_2|H\rangle = |V\rangle$, $\hat{S}_2|V\rangle = |H\rangle$. Then, will first act on the state by the B operators and see what happens:

$$\hat{F}|\Psi^{(-)}\rangle = \frac{1}{2} \{ \hat{S}_1^A [-|H\rangle_A |V\rangle_B - |V\rangle_A |H\rangle_B] + \hat{S}_2^A [|H\rangle_A |H\rangle_B - |V\rangle_A |V\rangle_B] \}$$

And now, we act on the state by the A operators:

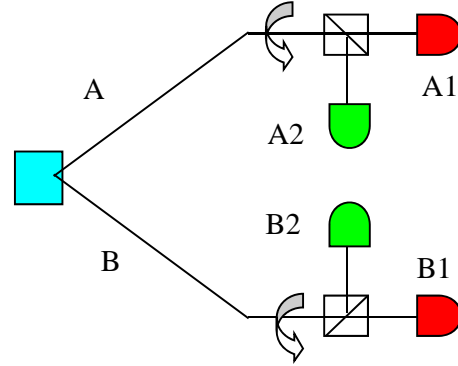


Fig.6

$$\hat{F}|\Psi^{(-)}\rangle = \frac{1}{2}\{-|H\rangle_A|V\rangle_B + |V\rangle_A|H\rangle_B + |V\rangle_A|H\rangle_B - |H\rangle_A|V\rangle_B\} = -\sqrt{2}|\Psi^{(-)}\rangle.$$

We see that $|\Psi^{(-)}\rangle$ is indeed an eigenstate of \hat{F} , $\hat{F}|\Psi^{(-)}\rangle = -\sqrt{2}|\Psi^{(-)}\rangle$, hence the mean value is easy to calculate:

$$\langle\hat{F}\rangle = \langle\Psi^{(-)}|\hat{F}|\Psi^{(-)}\rangle = -\sqrt{2}.$$

So we see that $|\langle\hat{F}\rangle| = \sqrt{2} > 1$, the CHSH inequality is violated (should be violated if the quantum mechanical calculation is true).

4. Experimental tests of Bell's inequality

In order to test the CHSH inequality in experiment, one has to perform four series of measurements:

- (1) $\langle\hat{S}_1^A(\hat{S}_1^B + \hat{S}_2^B)\rangle$: with polarizer A at 0 deg, polarizer B at 22.5 deg.
- (2) $\langle\hat{S}_2^A(\hat{S}_1^B + \hat{S}_2^B)\rangle$: with polarizer A at 45 deg, polarizer B at 22.5 deg.
- (3) $\langle\hat{S}_1^A(\hat{S}_1^B - \hat{S}_2^B)\rangle$: with polarizer A at 0 deg, polarizer B at 67.5 deg
- (4) $\langle\hat{S}_2^A(\hat{S}_1^B - \hat{S}_2^B)\rangle$: with polarizer A at 45 deg, polarizer B at 67.5 deg.

Usually, one only measures the rate of coincidences between detectors A1 and B1. Indeed, consider the mean value (1). Note that the number of photons N_{A2} hitting detector A2 is equal to $N_A - N_{A1}$, where N_{A1} is the number of photons hitting A1 and N_A is the total number of photons in channel A (which can be measured by removing the polarizer). Then the expression (1) can be written as

$$\begin{aligned} \frac{1}{\sqrt{2}}\langle\hat{S}_1^A(\hat{S}_1^B + \hat{S}_2^B)\rangle &= \frac{\langle(N_{A1}(0^\circ) - N_{A2}(0^\circ))(N_{B1}(22.5^\circ) - N_{B2}(22.5^\circ))\rangle}{\langle N_A N_B \rangle} = \\ &= \frac{\langle(2N_{A1}(0^\circ) - N_A)(2N_{B1}(22.5^\circ) - N_B)\rangle}{\langle N_A N_B \rangle} = \\ &= \frac{4\langle N_{A1}(0^\circ)N_{B1}(22.5^\circ)\rangle - 2\langle N_A N_{B1}(22.5^\circ)\rangle - 2\langle N_B N_{A1}(0^\circ)\rangle + \langle N_A N_B \rangle}{\langle N_A N_B \rangle}. \end{aligned}$$

The whole combination F will be then given by

$$\begin{aligned} \frac{1}{2\langle N_A N_B \rangle} &[4\langle N_{A1}(0^\circ)N_{B1}(22.5^\circ)\rangle - 2\langle N_A N_{B1}(22.5^\circ)\rangle - 2\langle N_B N_{A1}(0^\circ)\rangle + \langle N_A N_B \rangle + \\ &+ 4\langle N_{A1}(45^\circ)N_{B1}(22.5^\circ)\rangle - 2\langle N_A N_{B1}(22.5^\circ)\rangle - 2\langle N_B N_{A1}(45^\circ)\rangle + \langle N_A N_B \rangle + \\ &+ 4\langle N_{A1}(0^\circ)N_{B1}(67.5^\circ)\rangle - 2\langle N_A N_{B1}(67.5^\circ)\rangle - 2\langle N_B N_{A1}(0^\circ)\rangle + \langle N_A N_B \rangle - \\ &- 4\langle N_{A1}(45^\circ)N_{B1}(67.5^\circ)\rangle + 2\langle N_A N_{B1}(67.5^\circ)\rangle + 2\langle N_B N_{A1}(45^\circ)\rangle - \langle N_A N_B \rangle]. \end{aligned}$$

After canceling similar terms, we get

$$\begin{aligned}
\langle F \rangle &= \frac{1}{\langle N_A N_B \rangle} [2\langle N_{A1}(0^\circ) N_{B1}(22.5^\circ) \rangle + 2\langle N_{A1}(45^\circ) N_{B1}(22.5^\circ) \rangle + 2\langle N_{A1}(0^\circ) N_{B1}(67.5^\circ) \rangle - \\
&- 2\langle N_{A1}(45^\circ) N_{B1}(67.5^\circ) \rangle - 2\langle N_A N_{B1}(22.5^\circ) \rangle - 2\langle N_B N_{A1}(0^\circ) \rangle + \langle N_A N_B \rangle] = \\
&= \frac{2}{\langle N_A N_B \rangle} [\langle N_{A1}(0^\circ) N_{B1}(22.5^\circ) \rangle + \langle N_{A1}(45^\circ) N_{B1}(22.5^\circ) \rangle + \langle N_{A1}(0^\circ) N_{B1}(67.5^\circ) \rangle - \\
&- \langle N_{A1}(45^\circ) N_{B1}(67.5^\circ) \rangle - \langle N_A N_{B1}(22.5^\circ) \rangle - \langle N_B N_{A1}(0^\circ) \rangle] + 1 \equiv 2S + 1
\end{aligned}$$

For $-1 \leq \langle F \rangle \leq 1$, the variable $S \equiv \frac{1}{2} \{\langle F \rangle - 1\}$ should take values from -1 to 0. So, one has to check whether $0 \geq S \geq -1$, with

$$\begin{aligned}
S &= \frac{1}{\langle N_A N_B \rangle} [\langle N_{A1}(0^\circ) N_{B1}(22.5^\circ) \rangle + \langle N_{A1}(45^\circ) N_{B1}(22.5^\circ) \rangle + \langle N_{A1}(0^\circ) N_{B1}(67.5^\circ) \rangle - \\
&- \langle N_{A1}(45^\circ) N_{B1}(67.5^\circ) \rangle - \langle N_A N_{B1}(22.5^\circ) \rangle - \langle N_B N_{A1}(0^\circ) \rangle]
\end{aligned}$$

Quantum mechanics gives $\langle F \rangle = -\sqrt{2}$, hence $S \equiv -\frac{1}{2} \{\sqrt{2} + 1\} \approx -1.21$.

The corresponding coincidence counting rates are measured in experimental tests: 4 series with polarizers in both arms, 2 series with polarizers in 1 arm, 1 series with no polarizers.

The first experiments on Bell inequality violation were performed using photon pairs emitted in cascaded transitions in atoms:

- 1972 – Freedman, Clauser
- 1976 – Clauser
- 1976 – Fry and Thompson
- 1981,82 – Aspect, Grangier, Roger
- 1982 – Aspect, Dalibard, Roger

Further, these tests were performed with SPDC:

- 1988 – Ou, Mandel
- 1988 – Shih, Alley

After these first attempts, Bell tests were repeated with higher and higher accuracy, but some ‘loopholes’ for local realism still remained.

There were three main loopholes:

1. *Locality (communication) loophole.* To exclude the possibility of communication between particles A and B, measurements on A,B should be separated by a spacelike interval in the Minkowski space (Fig. 7, red rectangles).

Moreover, the choice of settings for the polarizers at A,B (green rectangles) should be also at spacelike interval to the measurement on the other side. This requires a relatively fast setting of polarizers and measurement and a relatively large separation between the measurement stations from the source.

2. *Freedom-of-choice loophole:* the settings of the polarizers should be chosen free or random. This implies using some really random choice – a random-number generator, for instance.

3. *Fair sampling (detection) loophole.* One can say that the detected photon pairs represent the whole ensemble if the detection efficiency is high enough. For the standard situation of

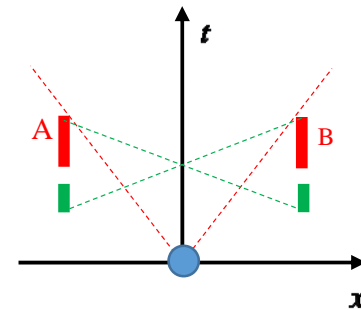


Fig. 7

state (1), about 82% detection efficiency is needed. However, as it was shown by Eberhard, if the state is non-maximally entangled,

$$|\Psi\rangle = \frac{1}{\sqrt{1+r^2}} \{ |H\rangle_A |V\rangle_B + r |V\rangle_A |H\rangle_B \},$$

Then the quantum efficiency can be reduced. For each value of the QE, a value of r exists needed to overcome the detection problem: for QE=2/3, $r = 0.001$ and for QE=70%, $r = 0.136$.

All three loopholes were overcome in papers [2]. The locality loophole was overcome by using large distances (~60 m, ~170 m and 1.3 km, respectively) with fast electronics. The freedom-of-choice loophole was eliminated by applying various random-number generators. The fair sampling loophole was closed by using the Eberhard state and detection with ~73% QE (nanowires), and in the last paper, spin-1/2 excitations (NV-centers in diamond) with perfect detection efficiency.

Literature:

1. D N Klyshko, Physics – Uspekhi, 41 (9) 885 - 922 (1998).
2. M. Giustina et al., Phys. Rev. Lett. **115**, 250401 (2015); L. K. Shalm et al., Phys. Rev. Lett. **115**, 250402 (2015); B. Hensen et al., Nature 526, 682 (2015).