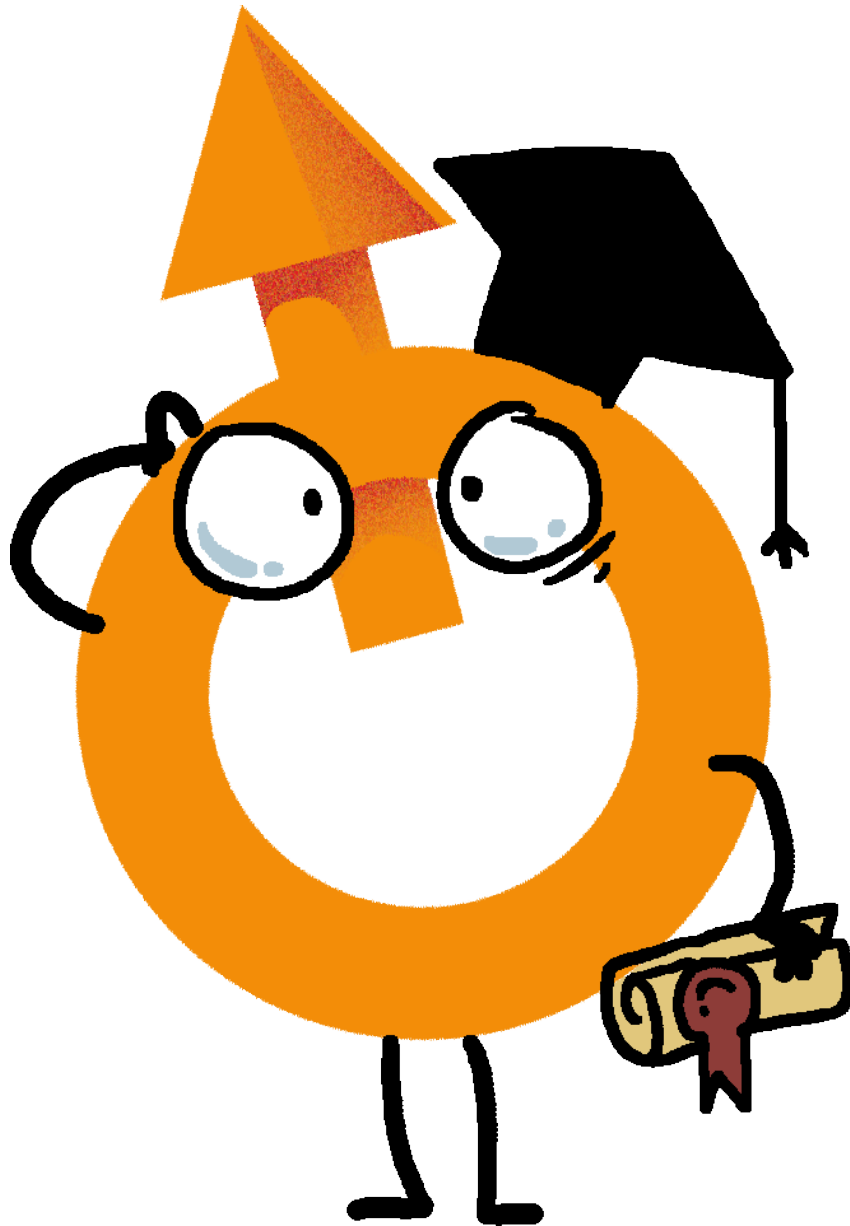


QUANTUM COMMUNICATION

Winter Semester 2021 / 2022

Falk Eilenberger, Institute of Applied Physics, Friedrich Schiller University, Jena

Fabian Steinlechner, Fraunhofer-Institute for Applied Optics and Precision Engineering IOF, Jena



The QuBitzies* are illustrated by Johannes Kretschmar of your most favourite [Lichtwerkstatt](#).

*The QuBitzies roam the vast spaces of the Bloch-Sphere. The arrow-organ indicates their location and provides orientation, in their very own and very peculiar habitat.

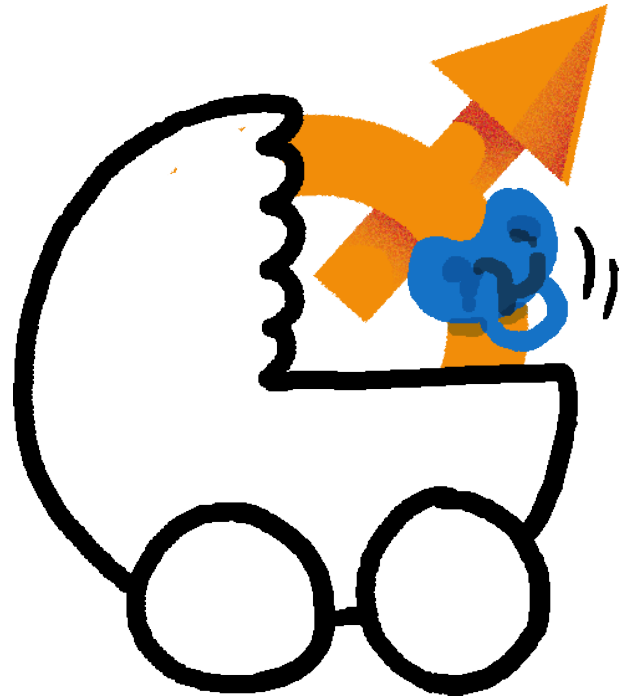
| | | |
|----------|---|-----------|
| 1 | Recap: Quantum Optics | 4 |
| 1.1 | Classical Modes and the Electromagnetic fields | 5 |
| 1.2 | The Quantization of the Fields and Modes | 12 |
| 1.3 | Time Evolution..... | 24 |
| 2 | Fundamentals..... | 26 |
| 2.1 | The principles of quantum theory | 26 |
| 2.2 | On quantum optics and the Nature of Photons | 30 |
| 2.3 | Matrix representations | 31 |
| 2.4 | Mixed States and the density matrix | 32 |
| 2.5 | Time evolution of quantum systems..... | 35 |
| 3 | Photonic Qubits | 37 |
| 3.1 | The Qubit | 37 |
| 3.2 | The Bloch Sphere | 39 |
| 3.3 | Observables and the Pauli-Matrices | 41 |
| 3.4 | Single Photon Operations, Gates, Elements and Hamiltonians..... | 44 |
| 3.5 | Mixed Single-Qubit States..... | 50 |
| 4 | Multiple Qubits and Entanglement..... | 51 |
| 4.1 | Product States and Non-Correlation..... | 52 |
| 4.2 | Non-Product States, Correlation, and Entanglement..... | 53 |
| 4.3 | The No-Cloning Theorem..... | 56 |
| 5 | Creating and Characterizing Photons | 58 |
| 5.1 | Single Photon Sources..... | 58 |
| 5.2 | Characterizing SPS: the Hanbury-Brown-Twiss Experiment..... | 61 |
| 5.3 | Photon Pair Sources based on SPDC..... | 63 |
| 5.4 | Characterizing PPS: The Hong-Ou-Mandel-Effect..... | 67 |
| 5.5 | Measuring Photonic Bell-States | 70 |
| 6 | EPR and the Bell-Inequalities | 74 |
| 6.1 | EPR'S Arguments on the Nature of Nature | 75 |
| 6.2 | Bell's Inequalities | 76 |
| 6.3 | A generalization: CHSH-Inequalities | 78 |
| 6.4 | Experimental Validation and some notes on loopholes | 85 |
| 7 | Quantum Key Distribution..... | 87 |
| 7.1 | Fundamentals of Cryptography | 88 |
| 7.2 | Physical Security Fundamentals..... | 91 |
| 7.3 | QKD with Single Qubits / BB84 | 92 |
| 7.4 | QKD with Entangled Qubits / Eckert 91..... | 96 |
| 7.5 | Overview over other security issues and mitigation strategies | 98 |
| 7.6 | Transmission rate and limits on transmission distance..... | 100 |

All notes subject to change, no guarantee to correctness, corrections welcome.

| | | |
|------------|---|------------|
| 8 | Advanced Quantum Communication Schemes..... | 105 |
| 8.1 | Quantum Teleportation..... | 105 |
| 8.2 | Entanglement Swapping and The Quantum Repeater..... | 107 |
| 8.3 | Superdense Coding..... | 109 |
| A 1 | Theoretical Description of Photon Detection..... | 112 |
| A 1.1 | Photon Detection | 112 |
| A 1.1.1 | Glauber’s quantum model for photodetection | 112 |
| A 1.2 | Threshold (“bucket/click”) detectors | 115 |
| A 1.3 | Correlation functions and coherence | 116 |
| A 1.4 | Quantum Interference and the Hong-Ou-Mandel-Effect..... | 120 |
| A 1.5 | Applications of HBT and HOM..... | 123 |
| A 2 | Single Photon Resolving Detectors: an Overview..... | 125 |
| A 2.1.1 | Single Channel Detectors (Bucket Type)..... | 126 |
| A 2.1.2 | Pixelated Detectors..... | 129 |
| A 3 | Three-Photon Processes and SPDC..... | 133 |
| A 3.1 | Fundamentals of Three Photon Processes | 133 |
| A 3.2 | Coupled Wave Equations | 136 |
| A 3.3 | Two-photon state produced in SPDC..... | 138 |
| A 3.4 | A note on the Connection to Joint-Probability-Densities and Correlation Properties of Stochastic Ensembles..... | 141 |

1 Recap: Quantum Optics

This first chapter is basically a recap of quantum optics, or, if you so like, the quantized version of Maxwell's equations. We shall glance over it in the lecture very briefly but, of course, it contains some of the basic physical concepts that we are making use of in quantum communication. The reason is simple: just as electromagnetic waves are the go-to-solution for the transmission of classic information so are their non-classic counterparts, namely photons, the go-to-solutions for the transmission of quantum information (which is a bold claim, since we have neither introduced the concept of photons nor of quantum information here). The reason is simple, however: being charge neutral and bosonic photons do not interact easily with each other. This means, they retain their quantum states, even when travelling over the vast, empty spaces of the universe but also when travelling through a lot of ubiquitous matter, such as glass or air. Moreover, we have learned to manipulate them with quite a lot of precision, so...photons it is.



But before digressing, let's get back to the recap of quantum electrodynamics. There is basically three ways on how you can deal with this chapter:

- 1) You could be a good student and go through the chapters. If you have already taken a class in wave electrodynamics, you will find that the first subchapter is nothing really new to you. If you have taken a course in quantum optics chapters 2 and 3 will also be a mere repetition. This will help you a lot in understanding the deeper connection between modes, field operators, and photons and in will make you a better person in general.
- 2) You could just ignore this chapter altogether and start directly on the next one. This is not a bad solution either, as this will introduce they laws by which photons abide as axioms, that need no further proofs but have been derived from experimental observation (and which require constant re-evaluation!). Which is just as well, because it's the way it is. If anyone of the students that went through the first chapter show's off as a superior student, always keep in mind that chapter 1 does exactly the same thing, it just introduced the axioms on the level of the fields and not the photons directly. Nevertheless: had you taken the time to go through chapter one, you would have a better understanding on the connection of quantum and classical light.
- 3) Is the most practical approach. You stick with the following summary:
 - a. Quantum Electrodynamics relies on exactly the same mathematical apparatus to introduce exactly the same modes as classical electrodynamics (e.g. plane wave with wavevector \mathbf{k} and a polarization index λ if you are in vacuum), including the same dispersion relation and scalar products.

- b. However, the expansion coefficients of any given mode $\mathbf{u}_\lambda(\mathbf{k})$ are no longer numbers $a_\lambda(\mathbf{k})$ and $a_\lambda^*(\mathbf{k})$ but operators $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}_\lambda^\dagger(\mathbf{k})$, which commute according to the commutations relations:

$$\begin{aligned} [\hat{a}_\lambda(\mathbf{k}), \hat{a}_{\lambda'}^\dagger(\mathbf{k}')] &= \delta_{\lambda\lambda'} \delta(\mathbf{k} - \mathbf{k}') \\ [\hat{a}_\lambda^\dagger(\mathbf{k}), \hat{a}_{\lambda'}^\dagger(\mathbf{k}')] &= 0 \\ [\hat{a}_\lambda(\mathbf{k}), \hat{a}_{\lambda'}(\mathbf{k}')] &= 0 \end{aligned} \quad (1)$$

- c. As a consequence, the state of any classical mode $\mathbf{u}_\lambda(\mathbf{k})$ can be described as a superposition $|\psi_\lambda(\mathbf{k})\rangle = \sum_{n=1}^{\infty} \alpha_n(\mathbf{k}) |n_\lambda(\mathbf{k})\rangle$ of mode specific number states $|n_\lambda(\mathbf{k})\rangle$. If a mode is in a number state $|n_\lambda(\mathbf{k})\rangle$ then we say, the mode is populated by n photons. The mode's energy is then $E_\lambda(\mathbf{k}) = n\hbar\omega(\mathbf{k})$, where $\omega(\mathbf{k})$ is the mode's frequency according to the specific dispersion relation of the system. Note that we have ignored the zero point energy here, because it does not generally play a role in quantum communication. The energy is measured with the help of the photon number operator $\hat{n}_\lambda(\mathbf{k}) = \hat{a}_\lambda^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k})$ by calculating its expectation value $\langle \hat{n}_\lambda(\mathbf{k}) \rangle = \langle \psi_\lambda(\mathbf{k}) | \hat{n}_\lambda(\mathbf{k}) | \psi_\lambda(\mathbf{k}) \rangle$.
- d. The number of photons can be increased by the application of the creation operator $\hat{a}_\lambda^\dagger(\mathbf{k})$, e.g. $\hat{a}_\lambda^\dagger(\mathbf{k})|n_\lambda(\mathbf{k})\rangle \sim |(n+1)_\lambda(\mathbf{k})\rangle$ and it can be decreased by the destruction operator $\hat{a}_\lambda(\mathbf{k})|n_\lambda(\mathbf{k})\rangle \sim |(n-1)_\lambda(\mathbf{k})\rangle$. The so-constructed ladder of photon number states begins at the quantum vacuum state $|0\rangle$.
- e. A common light source, such as a laser, is in a so-called coherent state, e.g. a Poisson-distributed superposition of number states. It's energy is thus ill-defined but it's behaviour, when subject to optical operations is just as we classical expect.
- f. Optical elements and the passage of time are described by the Heisenberg equation's of motions, i.e. by unitary operators, or equivalently by their Hamiltonian operators. If the number of modes are not mixed (e.g. the action of a beam splitter on an $n = 1$ number state) then the unitary operators may wittle down to simple matrix equations, just as we know and expect them from classical optics.

Each of the photon number modes in any spatial mode behaves according to the laws of quantum physics, as is outlined in the next chapter. The same is true for any superposition of modes, e.g. any change of basis.

You are now free to roam this chapter or skip to chapter 2.

1.1 Classical Modes and the Electromagnetic fields

In this chapter we will reiterate the fundamentals of quantum-electrodynamics (QED), i.e. the generalization Maxwell's Equations, which naturally lead to the concept of the modes and then to photons, which populate these individual modes. The notion of photons as quantum physical entities is at the centre of quantum optics in general and its properties are the fundamentals upon which quantum communication is built upon.

This chapter is only a brief overview over the most central concepts of the quantization of the field. For a more detailed analysis, see e.g. the lecture by Frank Setzpfandt on the "introduction to quantum optics".

This process of quantization is often termed the "construction of the laws of QED". This term is somewhat misleading; in reality, the process is really educated guesswork, which combines three trains of thought:

- **compatibility:** the classical electrodynamic equations (i.e. Maxwell's Equations) must retain their validity as an approximation to the new governing equations of QET,
- **construction:** we follow the same approach, that links classical with quantum mechanics; namely we first cast Maxwell's Equations into a canonical form, including a classic Hamiltonian and canonical position and momenta. These are then treated as operators. These classical quantities are constructed in a way, which leads to certain exchange rules, termed "Poisson Brackets {}", which carry over to the operator regime as commutation equations. This process automatically ensures the compatibility requirement.
- **Validation:** Following the same approach as has been successful for quantum mechanics does by no means guarantee that we end up with theory, which describes reality. It's validity has to be proven in experiments (or more precisely, it must withstand any attempt at falsification!). Spoiler alert: so far all experiments have validated this approach. Even to the point that we (as a scientific community) had to alter our understanding of the very nature of reality itself. This will be treated in later chapter.

1.1.1 Maxwell's Equation in Canonical Formulation

Maxwell's Equations can be written as the evolution equation to the Lagrangian density:

$$\mathcal{L}(\phi, \dot{\phi}, \mathbf{A}, \dot{\mathbf{A}}) = \frac{\epsilon_0}{2} \mathbf{E}^2(\mathbf{r}, t) - \frac{1}{2\mu} \mathbf{B}^2(\mathbf{r}, t) \quad (2)$$

where we have assumed free space propagation, i.e.

$$\mathbf{j} = 0 \quad \rho = 0 \quad (3)$$

and we have written the Lagrangian density in terms of the scalar potential ϕ and the vector potential \mathbf{A} . For the sake of simplicity, we adopt Coulomb (or radiation gauge)

$$\nabla \cdot \mathbf{A} = 0 \quad \dot{\phi} = 0 \quad (4)$$

Then the relation between the potentials \mathbf{A} and ϕ and the field \mathbf{E} and \mathbf{B} take the simple form

$$\mathbf{E} = -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} \quad \mathbf{B} = \nabla \times \mathbf{A}(\mathbf{r}, t) \quad (5)$$

Maxwell's Equations can be obtained from the Lagrangian density by application of the Euler-Lagrange-Equations

$$\frac{d}{dt} \frac{\delta \mathcal{L}}{\delta \dot{\phi}} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \quad \frac{d}{dt} \frac{\delta \mathcal{L}}{\delta \dot{A}_j} - \frac{\delta \mathcal{L}}{\delta A_j} = 0 \quad (6)$$

where the derivatives in front of the time derivatives are defined as the canonical momenta Π_ϕ and $\Pi_{\mathbf{A}}$ of the fields ϕ and \mathbf{A}

$$\Pi_\phi = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} = 0 \quad \Pi_{\mathbf{A}} = \frac{\delta \mathcal{L}}{\delta \dot{\mathbf{A}}} = \epsilon_0 \dot{\mathbf{A}} \quad (7)$$

One interesting side note is, that, due to a smart choice in the gauge freedom we get $\Pi_\phi = 0$ and $\dot{\phi} = 0$, this we basically only need to worry about \mathbf{A} and its momentum. This again has a physical interpretation: in free space (and in fact in all non-magnetic material), we only need worry about either the electric field \mathbf{E} or the magnetic field \mathbf{B} ; the other one is connected by a simple transformation.

By this construction, the canonical coordinates and canonical momenta automatically fulfil the classical commutation relations, termed Poisson-Brackets:

$$\begin{aligned} \{A_k(\mathbf{r}, t), A_j(\mathbf{r}', t)\} &= 0 \\ \{\Pi_A^k(\mathbf{r}, t), \Pi_A^j(\mathbf{r}', t)\} &= 0 \\ \{A_k(\mathbf{r}, t), \Pi_A^j(\mathbf{r}', t)\} &= \Delta_{kj}(\mathbf{r} - \mathbf{r}') \end{aligned} \quad (8)$$

We can now construct the classical Hamiltonian Density by executing a Legendre transformation with respect to the dynamical variables $\frac{\partial \Phi}{\partial t}$ and $\frac{\partial A}{\partial t}$. We arrive at:

$$\begin{aligned} \mathcal{H} &= \Pi_\phi \dot{\phi} + \Pi_A \dot{A} - \mathcal{L} \\ &= \Pi_A \dot{A} - \mathcal{L} \\ &= \frac{\epsilon_0}{2} \dot{A}^2 + \frac{1}{2\mu_0} (\nabla \times \mathbf{A})^2 \\ &= \frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2 \end{aligned} \quad (9)$$

Which is (somewhat unsurprisingly) the energy density of the electromagnetic field, which we could have guessed right away. But, we would have not gotten the definition of the canonical momenta and positions from just guessing the Hamiltonian density. This is however an important ingredient in the quantization process, as they are crucial in the definition of observables to the system.

1.1.2 Plane Waves as Classical Eigenmodes

The resulting Maxwell-Equation can be reformulated as the wave equation

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0 \quad \epsilon_0 \mu_0 = c^{-2} \quad (10)$$

Each solution to this equation (i.e. each EM-field) can then be written as a superposition of plane waves $u(\mathbf{k})$

$$\begin{aligned} \mathbf{A}(\mathbf{r}, t) &= \sum_\lambda \int_{\mathbb{R}^3} dk_x dk_y dk_z a_\lambda(\mathbf{k}) \frac{1}{\sqrt{(2\pi)^3 2\omega_{\mathbf{k}}}} \boldsymbol{\epsilon}_\lambda(\mathbf{k}) e^{i(\mathbf{k}\mathbf{r} - \omega_{\mathbf{k}}t)} + c.c \\ &= \sum_\lambda \int_{\mathbb{R}^3} dk_x dk_y dk_z a_\lambda(\mathbf{k}) \boldsymbol{\epsilon}_\lambda(\mathbf{k}) u(\mathbf{k}) + c.c \\ \mathbf{u}_\lambda(\mathbf{k}) &= \boldsymbol{\epsilon}_\lambda(\mathbf{k}) \frac{e^{i(\mathbf{k}\mathbf{r} - \omega_{\mathbf{k}}t)}}{\sqrt{(2\pi)^3 2\omega_{\mathbf{k}}}} = \boldsymbol{\epsilon}_\lambda(\mathbf{k}) u(\mathbf{k}) \end{aligned} \quad (11)$$

With the dispersion relation

$$\frac{\omega_{\mathbf{k}}^2}{c^2} = k_x^2 + k_y^2 + k_z^2 \quad \omega_{\mathbf{k}} = \pm \sqrt{k_x^2 + k_y^2 + k_z^2} \quad (12)$$

and a polarization state $\boldsymbol{\epsilon}_\lambda(\mathbf{k})$ with $|\boldsymbol{\epsilon}_\lambda(\mathbf{k})|^2 = 1$, which is constant over space for a given mode and orthogonal to \mathbf{k} , such that $\mathbf{k} \cdot \boldsymbol{\epsilon}_\lambda(\mathbf{k}) = 0$. Thus $\boldsymbol{\epsilon}_\lambda(\mathbf{k})$ spans a two-dimensional vector space with the basis vectors $\boldsymbol{\epsilon}_1(\mathbf{k})$ and $\boldsymbol{\epsilon}_2(\mathbf{k})$, which are mutually orthogonal $\boldsymbol{\epsilon}_1(\mathbf{k}) \cdot \boldsymbol{\epsilon}_2(\mathbf{k}) = 0$.

An important note here are the expansion coefficients $a_\lambda(\mathbf{k})$ and $a_\lambda^*(\mathbf{k})$, which are a set of complex numbers, that give amplitude and phase of the electric field in each mode $\mathbf{u}_\lambda(\mathbf{k})$. Herein lies the most profound difference to quantum optics.

1.1.3 The Scalar Product

To calculate the expansion coefficients $a(\mathbf{k})$ and $a_\lambda^*(\mathbf{k})$ from a given field distribution $\mathbf{A}(\mathbf{r}, t)$ we require a scalar product. This has the form

$$\begin{aligned} (\boldsymbol{\phi}(\mathbf{r}, t), \boldsymbol{\psi}(\mathbf{r}, t)) &= i \int_{\mathbb{R}^3} d\mathbf{r} \left(\boldsymbol{\phi}(\mathbf{r}, t)^* \cdot \partial_t \boldsymbol{\psi}(\mathbf{r}, t) - (\partial_t \boldsymbol{\phi}(\mathbf{r}, t))^* \cdot \boldsymbol{\psi}(\mathbf{r}, t) \right) \\ (\boldsymbol{\phi}(\mathbf{r}, t; \omega_1), \boldsymbol{\psi}(\mathbf{r}, t; \omega_2)) &= (\omega_1 + \omega_2) \int_{\mathbb{R}^3} d\mathbf{r} \boldsymbol{\phi}(\mathbf{r})^* \boldsymbol{\psi}(\mathbf{r}) \end{aligned} \quad (13)$$

Where the second line holds only for time-harmonic fields such that $\boldsymbol{\phi}(\mathbf{r}, t; \omega_1) = \boldsymbol{\phi}(\mathbf{r}) \exp(-i\omega_1 t)$ and $\boldsymbol{\psi}(\mathbf{r}, t; \omega_2) = \boldsymbol{\psi}(\mathbf{r}) \exp(-i\omega_2 t)$. Also note that by virtue of their nature as a basis set the set of plane waves are mutually orthogonal:

$$(\mathbf{u}_\lambda(\mathbf{k}), \mathbf{u}_{\lambda'}(\mathbf{k}')) = \delta_{\lambda\lambda'} \delta^3(\mathbf{k} - \mathbf{k}'). \quad (14)$$

Now we can calculate the expansion coefficients of the field $\mathbf{A}(\mathbf{r}, t)$ (or any other function) according to:

$$\begin{aligned} a_\lambda(\mathbf{k}) &= (\mathbf{u}_\lambda(\mathbf{k}), \mathbf{A}(\mathbf{r}, t)) \\ a_\lambda^*(\mathbf{k}) &= -(\mathbf{u}_\lambda^*(\mathbf{k}), \mathbf{A}(\mathbf{r}, t)) \end{aligned} \quad (15)$$

Note that plane waves are modes of the unstructured media and free space. In a structured medium, e.g. in a photonic crystal or a waveguide, the wave equation takes a different form and thus we get a different dispersion relation, a different set of eigenmodes and a different scalar product. The overall role of the modes and the nature of the scalar product¹ nevertheless remains totally unchanged. The same is true for the quantization, as to basically “stick a hat on the expansion coefficients”. Which we will get to later.

The dispersion relation states that there is an infinite number of plane waves $\mathbf{u}_\lambda^{(0)}(\mathbf{k})$, which belong to the same frequency $\omega(\mathbf{k}) = \omega_0$. Thus any linear combination of such modes $\mathbf{u}_\lambda^{(0)}(\mathbf{k})$ is also a mode $\mathbf{v}_\lambda^{(0)}(\mathbf{k})$ of the system. In fact, all such eigenmodes $\mathbf{u}_\lambda^{(0)}(\mathbf{k})$ form a vector space in which any number of bases may be constructed from superpositions of plane waves. Some examples are cylindrical waves, Gauss-Laguerre-Waves, Legendre-Waves, Bessel- and Matthieu-waves, etc.

1.1.4 Non-Plane-Wave Fields

In this chapter we will see how the decomposition into plane waves can then help us to decompose the field into other basis set of modes. These are useful in many theoretical and experimental scenarios. From an experimental point of view, they may match the symmetries of the system (e.g. circular optics), the nature of the available light sources, and the non-infinite size of the setup. From a theoretical point of view, they are oftentimes, nice as they have a more benign mathematical properties than plane waves. They are not infinitely extended in space and time and thus more therefore easier to make calculations with.

Note that for notational brevity we will ignore the vectorial nature of plane waves, these can be easily integrated, if required.

Any non-plane wave basis set $\mathbf{v}_\mu(\boldsymbol{\kappa}, \mathbf{r}, t)$ can be constructed from a superposition of plane wave modes $\mathbf{u}(\mathbf{k})$.

¹ An in-depth treatment can be found in Saleh Teich “Fundamentals of Photonics” and Synder/Love “Optical Waveguide Theory”

$$\mathbf{v}_\mu(\boldsymbol{\kappa}; \mathbf{r}, t) = \int d\mathbf{k} V_\mu^\lambda(\mathbf{k}, \boldsymbol{\kappa}) \mathbf{u}_\lambda(\mathbf{k}; \mathbf{r}, t) \quad (16)$$

Note that $\boldsymbol{\kappa}$ is now any set of indices, which enumerates the new basis set and $V(\mathbf{k}, \boldsymbol{\kappa})$ is a unitary matrix, i.e. $V^* = V^{-1}$. The unitarity of V dictates, that the orthogonality of the plane wave modes carries over to the orthogonality of the new modes, i.e. $(\mathbf{v}_\mu(\boldsymbol{\kappa}), \mathbf{v}_{\mu'}(\boldsymbol{\kappa}')) = \delta_{\mu\mu'} \delta^3(\boldsymbol{\kappa} - \boldsymbol{\kappa}')$.

Using these definitions the field under description will then have the form:

$$\mathbf{A}(\mathbf{r}, t) = \sum_\mu \int_{\mathbb{R}^3} d^3\kappa b(\boldsymbol{\kappa}) \mathbf{v}_\mu(\boldsymbol{\kappa}) + c.c$$

And the modal expansion coefficients for the basis set $\mathbf{v}(\mathbf{k})$ may be derived from the field as

$$\begin{aligned} b_\mu(\boldsymbol{\kappa}) &= (\mathbf{v}_\mu(\boldsymbol{\kappa}), \mathbf{A}(\mathbf{r}, t)) \\ b_\mu^*(\boldsymbol{\kappa}) &= -(\mathbf{v}_\mu^*(\boldsymbol{\kappa}), \mathbf{A}(\mathbf{r}, t)). \end{aligned} \quad (17)$$

One also can often approximate the light to belong exclusively to a certain range of harmonics with $\omega \in [\omega_0 - \Delta\omega, \omega_0 + \Delta\omega]$ and $\omega_0 \gg \Delta\omega$. If modal dispersion is also neglected, i.e. if the spatial modes essentially look the same for all of those ω we end up in the regime of the a slowly-varying envelope (SVEA). This makes the introduction of pulsed beams much simpler.

1.1.5 Example 1: Gaussian Modes

For the introduction of Gaussian modes, we will also assume paraxiality, i.e. the beam diameter is much larger than the wavelength of light. We will also assume that its propagation direction is centred along the z-axis. Thus

$$k_z \approx k \left(1 - \frac{k_x^2 + k_y^2}{k^2} \right) \quad k(\omega) = \frac{\omega}{c} \quad \omega_0 \gg \Delta\omega. \quad (18)$$

We also assume a slowly varying envelope (SVEA). We shall also assume that all modes are only excited with one type of polarization and thus ignore the vectorial nature of the fields and the μ subscripts. Under these assumptions, the Gaussian field $A(r, t)$ takes the form:

$$\begin{aligned} A(r, t) &= \int d\omega (b(\omega) v^{\text{Ga}}(\omega; \mathbf{r}) + c.c) \\ v^{\text{Ga}}(\omega; \mathbf{r}, t) &= \frac{4\pi}{s^2(z)} \exp\left(ik(\omega)z - \frac{x^2 + y^2}{s^2(z)} \right) e^{i\omega t} \\ s^2(z) &= w_0^2 + \frac{2iz}{k} \end{aligned} \quad (19)$$

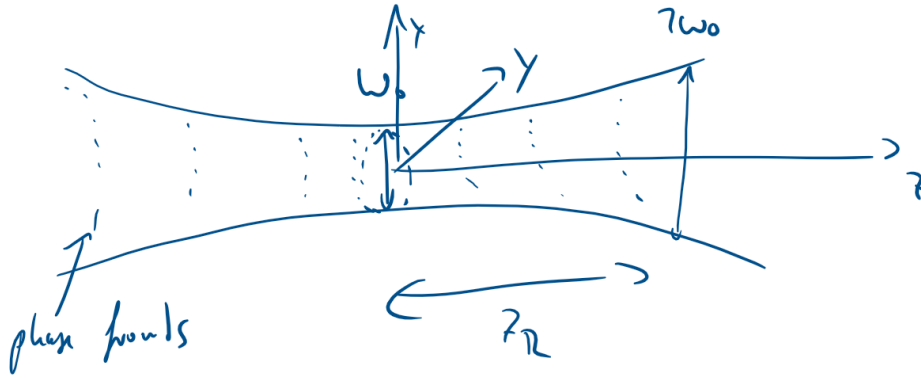


Fig. 1: Sketch of the diffraction properties of a Gaussian beam.

A Gaussian mode is defined for any single frequency ω by one parameter, its waist diameter w_0 and there is a set of fancy relations

$$z_R = \frac{\pi w_0^2}{\lambda} \quad w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_R}\right)^2} \quad NA = \frac{w_0}{z_R} = \frac{\lambda}{\pi w_0}, \quad (20)$$

which link this waist diameter to the way the beam diverges.

One can further show, that these transverse modal fields are minimum uncertainty localized transverse modes, i.e. they are the modes which for a given diameter have the least possible divergence

$$\Delta k \Delta x = \frac{1}{4\pi} \int d\mathbf{k} |f(\mathbf{k})|^2 \rightarrow w_0 NA = \frac{\lambda}{\pi}. \quad (21)$$

They are therefore well suited for long-range communication, as they require the smallest telescopes. Moreover, most lasers and optical fibers operate on modes, which are typically very close to Gaussian modes.

Note that the Gaussian modes, as presented here, is not a complete set of Eigenmodes. A possible completion will be given in the following example.

1.1.6 Example 2: Gauss-Laguerre Modes

We can extend the Gaussian Modes onto a complete set of Eigenmodes with rotational symmetry, allowing them to describe any kind of transversal field distribution. We here focus on Gauss-Laguerre-Modes, because they are experimentally most relevant as rationally symmetric modes and carriers of orbital angular momentum. We introduce

$$x = r \cos \varphi \quad y = r \sin \varphi \quad \sigma = x + iy. \quad (22)$$

The field $A(r, t)$ can be composed from the Gauss-Laguerre with the relation

$$A(r, t) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \int dk b_{lm}(\mathbf{k}) v_{lm}^{\text{LG}}(\omega; \mathbf{r}, t) + c. c \quad (23)$$

$$v_{lm}^{\text{LG}}(\omega; \mathbf{r}, t) = \frac{4\pi(-1)^{l+m} l!}{s^{2(l+m+1)}(z)} r^{|m|} e^{im\varphi} L_l^m\left(\frac{r^2}{s^2(z)}\right) e^{ik(\omega)z - \frac{r^2}{s^2(z)}} e^{i\omega t} + c. c$$

Note that $s(z)$ was defined above. The φ -dependency is in the phase term $\sim \exp(\sim im\varphi)$. Thus one can easily see that these modes are eigenfunctions to the operator, which measures the z-coordinate of the angular orbital momentum $\widehat{L}_z = \widehat{x} \widehat{p}_y - \widehat{y} \widehat{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial \varphi}$ with the Eigenvalue $\hbar m$:

$$\hat{L}_z v_{lm}^{\text{LG}}(\omega; \mathbf{r}, t) = m \hbar v_{lm}^{\text{LG}}(\omega; \mathbf{r}, t). \quad (24)$$

These beams thus carry a quantized and measurable orbital angular momentum. As this is a discrete quantity it can be used to conveniently transport information. Also note that this information transfer is quite robust: the angular momentum is a compatible measurable to both the direction \mathbf{k} of the beam, its frequency ω as well as its overall impulse l . Propagation through air typically induces perturbations along \mathbf{k} and l but very little on m . Information encoded in these modes is thus also robust.

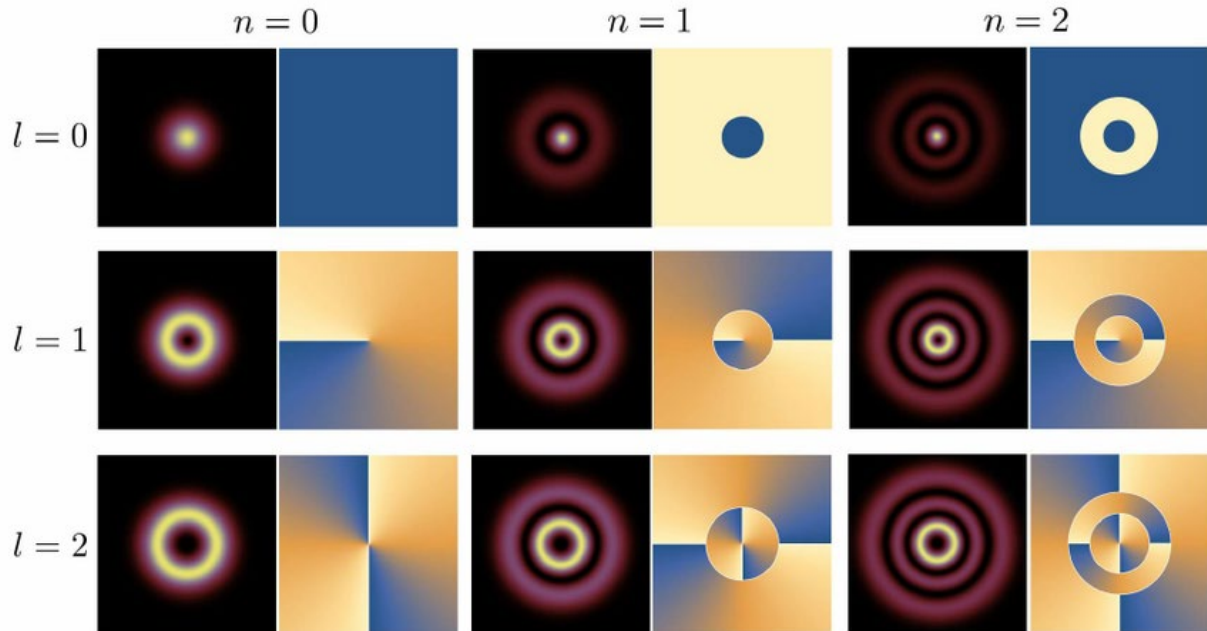


Fig. 2: Image of Gauss-Laguerre-Modes (in our notation l =OAM, n =radial Number)

1.1.7 Temporally localized wave packets

In the last chapters we have introduced two particular sets of non-plane wave modes and have taken this opportunity to briefly introduce the SVEA-approximation, which allows us to describe pulsed, i.e. temporally varying waves. In fact, it is often the case, that the spatial distribution of light is fixed to a certain number of well-known modes, whereas the temporal structure is where “the physics is happening”. For example, in or after a single mode fiber a laser pulse will always propagate in the mode dictated by the fiber geometry. As another example we can think of an atom or quantum dot emitting light: the light will always be fixed to a certain radiation mode, most likely a dipole mode.

All of these wavepackets do have a typical temporal structure, some common ones are noted here:

| Type | Typ. Emitter | Temporal Structure |
|------------|--------------------|--|
| Lorentzian | Atom / Quantum Dot | $b(\omega) = \frac{1}{\sqrt{\pi}\gamma} \frac{\sqrt{\gamma}}{i(\omega_k - \omega_0)}$ |
| Gaussian | Laser Pulse | $b(\omega) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\omega_k - \omega_0)^2}{4\sigma^2}\right\}$ |
| Rect. | Time Bin Encoder | $b(\omega) = \frac{1}{\sqrt{\Delta\omega}} \text{sinc}\left\{\frac{\omega_k - \omega_0}{\Delta\omega}\right\}^{1/2}$ |

These wavepackets can also be used for time-bin encoding in Quantum Communication.

1.1.8 Polarization modes

So far, we have pretty much ignored the polarization aspects of the modes. We shall now have a closer look at these. As with classical EM-theory these can be represented with Jones Vectors

$$\boldsymbol{\epsilon}(\mathbf{k}) = \begin{bmatrix} \epsilon_1(\mathbf{k}) \\ \epsilon_2(\mathbf{k}) \\ 0 \end{bmatrix} \quad (25)$$

Where we have assumed, without loss of generality, that $\mathbf{k} = k_z \mathbf{e}_z$. Then we can find a few single basis-vector systems, in which we can describe the polarization state of light:

$$\begin{aligned} \text{linear HV: } \epsilon &= h \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} 0 \\ 1 \end{bmatrix} = h|h\rangle + v|v\rangle \\ \text{linear diagonal: } \epsilon &= u \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} + d \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = u|u\rangle + d|d\rangle \\ \text{linear: } \epsilon &= l_1 \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} + l_2 \begin{bmatrix} -\sin \varphi \\ \cos \varphi \end{bmatrix} = l_1|l_{1\varphi}\rangle + l_2|l_{2\varphi}\rangle \\ \text{circular: } \epsilon &= l \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} + r \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} = l|l\rangle + r|r\rangle \end{aligned} \quad (26)$$

equally well. We are once getting ahead of ourselves and we are adapting bracket-notation instead of vectors – this will come in handy later on. In many of the classic implementations of QKD these basis sets are used heavily and thus they are important.

1.1.9 Summary

- Maxwell' equations can be cast into a classical Hamiltonian form, where the energy density is the Hamiltonian and both the vector field and its time derivative are conjugate coordinates and momenta will fulfil classical versions of the commutation relations.
- Plane waves are the classical eigenmodes of the electromagnetic field in free space. They are complete, meaning that every field can be constructed from and decomposed into a superposition of plane waves. Plane waves are thus a complete basis set; they are orthogonal with respect to a scalar product. There is a dispersion relation, which links the plane wave's spatial structure to their temporal frequency.
- The expansion coefficients of the plane wave modes are complex numbers of unit $Vs^{1/2}m^2$ (in more general terms: square root of generalized² energy per unit frequency), whereas the field has a unit of Vs/m (in more general terms: square root of generalized energy per unit frequency and unit wavenumber volume). The modes functions themselves are of unit $s^{1/2}$.
- There is an infinite set of non-plane wave modes, into which any field can be decomposed, as well. The properties of the scalar products does not change with the specific choice of the basis set. Basis functions of different basis sets are related by a unitaritan transformation.
- The non-plane-wave basis sets can be used to introduce basis sets, in which typical processes can be described particularly easy. These may capture particular spatial (e.g. Gaussian), polarization (e.g. linear or circular pol.) and temporal properties (Gaussian, Lorentzian) of light.

1.2 The Quantization of the Fields and Modes

In this chapter we will carry out the second quantization and see what kind of effects it has on the eigenmodes of the system. We will see that the main difference is the replacement of the complex

² Generalized means that a proper energy density is achieved by multiplication with ϵ_0 and division by a time-squared.

expansion coefficients with operator-valued quantities and that these operators each and by themselves behave exactly like a quantum harmonic oscillator. We will then see, how this finding gives rise to the concepts of photons, i.e. quantized excitations of the modes and how the excitations states of these modes are actually related to observable states of light.

1.2.1 Field Quantization in Space

We quantize the electromagnetic field by adding a "hat" to the vectorial fields and their momenta. In other words: we promote them from scalar (or vectorial) fields to operator fields. As their classical counterparts obey Poisson-Bracket relations, we postulate that these operator fields obey certain commutation relations:

$$\begin{aligned} [\hat{\mathbf{A}}^j(\mathbf{r},t), \hat{\mathbf{A}}^k(\mathbf{r}',t)] &= 0 \\ [\hat{\mathbf{\Pi}}_A^j(\mathbf{r},t), \hat{\mathbf{\Pi}}_A^k(\mathbf{r}',t)] &= 0. \end{aligned} \quad (27)$$

These relations have a physical meaning beyond a pure postulate of mathematics. They state that at a certain fixed point in time t one can measure the state of the quantum field $\hat{\mathbf{A}}$ at any two different points in space \mathbf{r} and \mathbf{r}' , without mutual influence. The same is true for the field's momentum $\hat{\mathbf{\Pi}}$. In terms of physical interpretations this means that Physics allows one to measure the instantaneous \mathbf{E} -field in all of space. The same is true for the \mathbf{B} -field.

However, nothing is stated here about what exactly one can learn at each point in space, i.e. if one can measure phase and amplitude at the certain point without mutual influence (which one cannot) and nothing is also stated about if one can measure \mathbf{E} - and \mathbf{B} -field at the same time. These questions are answered by the mixed commutation relation (also carried over from the classical Poisson-Brackets):

$$\begin{aligned} [\hat{\mathbf{A}}_i(\mathbf{r},t), \hat{\mathbf{\Pi}}_A^j(\mathbf{r}',t)] &= i\hbar\Delta_{ij}(\mathbf{r}-\mathbf{r}') \\ \hat{\mathbf{A}}_i &= -\hat{\mathbf{A}}^i \\ \Delta_{ij}(\mathbf{r}) &= \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\mathbf{r}} \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) \end{aligned} \quad (28)$$

Where the second equations comes into play, due to the relativistic nature of the fields and the third term is basically an ordinary δ -function, which is corrected for the divergence-free nature of the EM-field (i.e. that we have only two-polarizations for three spatial degrees of freedom).

This means that we cannot measure the same components of the $\hat{\mathbf{A}}$ -field and its momentum $\hat{\mathbf{\Pi}}$ independently at the same point in space and time. If you measure both, its result will depend on the order of the measurement. This carries over to the \mathbf{E} -field and \mathbf{B} -field being mutually dependent. Again, this is not a mere postulate but can be verified experimentally.

Also note that the commutator relations are scaled differently from their classical counterpart in that they have an \hbar . This also means that the units of the classical field \mathbf{A} and operator field $\hat{\mathbf{A}}$ are not quite the same anymore, they are no longer square roots of energy density but of action density.

1.2.2 Introduction of Quantum Plane Waves Modes

We can, of course decompose each quantum field into quantum plane waves, because the wave equation still holds. We will later see that due to the construction of the quantum fields, these quantum plane waves are eigenstates of the Hamiltonian-Operator of the system and thus remain shape invariant (except for a phase term) under the evolution of time. Note that it is convenient to integrate the change of units into the definition of the plane wave decomposition, so this now changes to.

$$\mathbf{A}(\mathbf{r}, t) = \sqrt{\frac{\hbar}{\epsilon_0}} \sum_{\lambda} \int d\mathbf{k} \{ \hat{a}_{\lambda}(\mathbf{k}) \mathbf{u}_{\lambda}(\mathbf{k}; \mathbf{r}, t) + c. c. \} \quad (29)$$

Neither the shape of the eigenmodes $\mathbf{u}_{\lambda}(\mathbf{k}; \mathbf{r}, t)$, nor their dispersion relation, nor the nature and result of their scalar product has changed in the slightest way. More specifically this means we can carry over from classical wave physics any wave property of the mode.

As we will later mostly just look into plane waves, it now makes sense to derive commutation relations for their operators. These can be found by plugging in the last equation into the commutations relations above. The calculation is tedious but straightforward. They read as:

$$\begin{aligned} [\hat{a}_{\lambda}(\mathbf{k}), \hat{a}_{\lambda'}^{\dagger}(\mathbf{k}')] &= \delta_{\lambda\lambda'} \delta(\mathbf{k} - \mathbf{k}') \\ [\hat{a}_{\lambda}^{\dagger}(\mathbf{k}), \hat{a}_{\lambda'}^{\dagger}(\mathbf{k}')] &= 0 \\ [\hat{a}_{\lambda}(\mathbf{k}), \hat{a}_{\lambda'}(\mathbf{k}')] &= 0 \end{aligned} \quad (30)$$

Which means that the state of any plane wave can be determined independently from the state of any other plane wave, except for the state of a plane wave and its own conjugate. To move ahead somewhat: you cannot determine the state of a mode and its phase, or in other words the electric and the magnetic field of one mode (think about measuring currents and voltages in Electronics, where both measurements necessarily influence each other). From a mathematical point of view this means that $\hat{a}_{\lambda}(\mathbf{k})$ and $\hat{a}_{\lambda}^{\dagger}(\mathbf{k})$ take the role of canonical conjugate variable and thus mathematically play the role of conjugate positions $\hat{q}_{\lambda}(\mathbf{k})$ and momenta $\hat{p}_{\lambda}(\mathbf{k})$.

Using the modal scalar product (f, g) we can invert the equation between $\hat{\mathbf{A}}$ and $\hat{a}_{\lambda}(\mathbf{k})$, namely:

$$\hat{a}_{\lambda}(\mathbf{k}) = \sqrt{\frac{\epsilon_0}{\hbar}} (\mathbf{u}_{\lambda}(\mathbf{k}; \mathbf{r}, t), \hat{\mathbf{A}}(\mathbf{r}, t)) \quad (31)$$

This equation gives us a recipe on how we can decompose any field into plane wave modes. Note that this is the exact same relation as the classical counterpart expect for a different scaling and the fact that the field itself is an operator-valued function.

As the nature of the scalar product has not changed, there is also no change in the introduction of non-plane-wave modes, i.e. the three equations above hold for $\hat{b}_{\mu}(\boldsymbol{\kappa})$, which are the quantum states of any arbitrary different set of modes $\mathbf{v}_{\mu}(\boldsymbol{\kappa}, \mathbf{r}, t)$, related to the plane wave modes $\mathbf{u}_{\lambda}(\mathbf{k}; \mathbf{r}, t)$ via a unitary transformation matrix $V_{\mu}^{\lambda}(\mathbf{k}, \boldsymbol{\kappa})$.

1.2.3 The Quantum Eigenmode Hamiltonian

Now that we have introduced the quantum modal operators $\hat{a}_{\lambda}(\mathbf{k})$, we can also derive the structure of the Hamiltonian operator $\hat{\mathcal{H}}$ in terms of the modal operators. This is achieved by the replacement of the expressions for \mathbf{A} with $\hat{\mathbf{A}}$ in the definition of the Hamilton-Operator in chapter 1.1.1 and by the subsequent expansion of $\hat{\mathbf{A}}$ into quantum plane wave modes as defined in chapter 1.2.2. The ensuing differential operator act only on the structure of the modes and after some tedious albeit straightforward calculation we can show that the QED analogue of the Hamilton-Operator is:

$$\begin{aligned} \hat{\mathcal{H}} &= \sum_{\lambda} \int d\mathbf{k} \frac{\hbar\omega(\mathbf{k})}{2} (\hat{a}_{\lambda}^{\dagger}(\mathbf{k}) \hat{a}_{\lambda}(\mathbf{k}) + \hat{a}_{\lambda}(\mathbf{k}) \hat{a}_{\lambda}^{\dagger}(\mathbf{k})) \\ &= \sum_{\lambda} \int d\mathbf{k} \hat{\mathcal{H}}_{\lambda}(\mathbf{k}) \end{aligned} \quad (32)$$

This result in and by itself is quite noteworthy. The Hamiltonian of the quantized fields is nothing but the sum of individual Hamiltonian's contributed from each mode. The individual Hamiltonian's for each mode are formally equivalent to harmonic oscillator Hamiltonian's, with the eigenfrequency identical to the frequency of the related mode.

The Hamiltonian has the following commutation relations with the modal operators:

$$\begin{aligned} [\mathcal{H}, \hat{a}_\lambda(\mathbf{k})] &= -\hbar\omega\hat{a}_\lambda(\mathbf{k}) \\ [\mathcal{H}, \hat{a}_\lambda^\dagger(\mathbf{k})] &= \hbar\omega\hat{a}_\lambda^\dagger(\mathbf{k}) \end{aligned} \quad (33)$$

This means, that you cannot measure the state of a single mode, without interfering with the energy of this state and vice versa. Moreover, you cannot measure the energy of the total radiation field, without messing with all the modes.

At this point we have not yet discussed the role of the Hamilton-Operator. Some of its aspects will (hopefully) become clearer in later chapter of this script but one aspect can be instantly carried over from classical mechanics. The Hamiltonian \mathcal{H} completely determines the equations of motion of any system (in the Heisenberg picture) it describes, if the canonical coordinates and momenta are known (which are $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}_\lambda^\dagger(\mathbf{k})$, see above). The system in question here is the state of the $\hat{\mathbf{A}}$ field in free space and the way it evolves.

In classical electrodynamics the equations of motion are derived, by application of the Poisson-brackets. In quantum electrodynamics we have to use the Commutator with a $-i/\hbar$ scaling to get the equations of motions in the Heisenberg-picture. A detailed derivation and some more insights into the manifold consequences of which are discussed in chapters 1.3 and later. Here we shall just use this equation to derive the equations of motion for the quantum modal operators:

$$\begin{aligned} \frac{\partial \hat{a}_\lambda(\mathbf{k})}{\partial t} &= -\frac{i}{\hbar} [\hat{a}_\lambda(\mathbf{k}), \mathcal{H}] \\ &= \frac{i}{\hbar} \hbar\omega\hat{a}_\lambda(\mathbf{k}) \\ &= i\omega(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) \\ \hat{a}_\lambda(\mathbf{k};t) &= \hat{a}_\lambda(\mathbf{k};t=0) \exp(i\omega(\mathbf{k})t) \end{aligned} \quad (34)$$

Which is the expected result: the state of any mode evolves with an $\exp(i\omega)$ phase term, just as we are used to in classical electrodynamics. In other words: the quantum modes interact with free space by the acquisition of a phase, which is proportional to the mode's frequency and the interaction duration.

1.2.4 Photons as Eigenstates of the Quantum Hamiltonian

So we have replaced the expansion coefficient with modal operators but have not yet made much progress in the understanding of its meaning or behaviour. As an operator is a highly abstract concept, it is always helpful to investigate its eigenstates and eigenvalues. Let's thus assume that we have found such an eigenstate $|\psi_n\rangle$ for the entire Hamiltonian:

$$\mathcal{H}|\psi_n\rangle = E_n|\psi_n\rangle \quad (35)$$

Now we can take this eigenstate of the Hamiltonian and let the modal expansion operators $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}_\lambda^\dagger(\mathbf{k})$ act on them. Using the commutation equations from the last chapter we find:

$$\begin{aligned}\mathcal{H} \hat{a}_\lambda(\mathbf{k})|\psi_n\rangle &= \hat{a}_\lambda(\mathbf{k})\mathcal{H}|\psi_n\rangle - \hbar\omega\hat{a}_\lambda(\mathbf{k})|\psi_n\rangle \\ &= (E_n - \hbar\omega)\hat{a}_\lambda(\mathbf{k})|\psi_n\rangle \\ \mathcal{H} \hat{a}_\lambda^\dagger(\mathbf{k})|\psi_n\rangle &= (E_n + \hbar\omega)\hat{a}_\lambda^\dagger(\mathbf{k})|\psi_n\rangle\end{aligned}\quad (36)$$

This means that as $\hat{a}_\lambda(\mathbf{k})$ acts on the Hamiltonian's eigenstate $|\psi_n\rangle$ it produces a new state $\hat{a}_\lambda(\mathbf{k})|\psi_n\rangle$. This new state is still an eigenstate to the Hamilton-operator, albeit with a by $\hbar\omega$ reduced eigenvalue (i.e. energy). The same is true for $\hat{a}_\lambda^\dagger(\mathbf{k})$ just that it increases the eigenvalue (i.e. energy). The same is, of course, true for consecutive applications $\hat{a}_\lambda(\mathbf{k})$ or applications of $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}_\lambda^\dagger(\mathbf{k}')$. We can thus use the modal expansion operators to generate entire arrays of eigenstates of the Hamiltonian if a single eigenstate is known.

Without loss of generality we can assume that the Eigenvalues of \mathcal{H} must be bound from below (it's an energy after all and negative energy is kind of hard to come by!). Thus, there should be a ground state $|\psi_0\rangle$ for which

$$\hat{a}_\lambda(\mathbf{k})|\psi_0\rangle = 0 \quad \forall \lambda, \mathbf{k} \quad (37)$$

This is called the quantum-vacuum state and is will be denoted as $|0\rangle$. However, if one calculates its energy one gets:

$$\begin{aligned}\mathcal{H}|0\rangle &= \sum_\lambda \frac{\int dk \hbar\omega(\mathbf{k})}{2} (\hat{a}_\lambda^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) + \hat{a}_\lambda(\mathbf{k})\hat{a}_\lambda^\dagger(\mathbf{k}))|0\rangle \\ &= \left(\int dk \hbar\omega(\mathbf{k}) \right) |0\rangle = \mathcal{E}|0\rangle\end{aligned}\quad (38)$$

This term is the quantum vacuum energy \mathcal{E} . It diverges and must be removed for all practical calculations of the energy. It's however not entirely unphysical. It leads e.g. to the Lamb-Shift, the Casimir-Force, and the Quantum-Unruh-Effect (dynamical Casimir Effect). For many cases, when we only investigate effects, which occur in a finite set of modes \mathcal{E} is finite anyway and can simply be ignored.

As $\hat{a}_\lambda^\dagger(\mathbf{k})$ and $\hat{a}_\lambda(\mathbf{k})$ can be used to move us up and down the ladder of Fock-States, we thus call them ladder-operators or creation and annihilation operators for the mode denoted by the index λ and \mathbf{k} .

Using the modal expansion operators, acting from the universal ground state $|0\rangle$ we can now introduce a particular array of eigenstates, denoted the (plane-wave) Fock-States $|n_{\mathbf{k},\lambda}\rangle$ for the mode denoted by \mathbf{k} and λ by applying $\hat{a}_\lambda^\dagger(\mathbf{k})$ n times to $|0\rangle$, such that

$$|n_{\mathbf{k},\lambda}\rangle \sim \left(\prod_{m=1}^n \hat{a}_\lambda^\dagger(\mathbf{k}) \right) |0\rangle \quad (39)$$

These then have the relative energy:

$$E_{n,\lambda}(\mathbf{k}) = \hbar\omega(\mathbf{k})n \quad (40)$$

Please note the proportionality sign in the above equations. The fact that one state fulfils an eigenvalue equation is not yet sufficient for it be a basis vector. It must also be normalized. Without loss of generality we can assume that the vacuum state is normalized, i.e. $\langle 0|0\rangle = 1$ but this is not necessarily true for any other state. We will address this issue now, but first have to deal with some technical problem. In the way we have defined Fock-States, they belong to infinitely extended plane waves and are thus not normalizable at all:

$$\begin{aligned}
 \langle 1_{k\lambda} | 1_{k'\lambda} \rangle &= \langle 0_{k\lambda} | \hat{a}_{\lambda'}(\mathbf{k}) \hat{a}_{\lambda}^{\dagger}(\mathbf{k}) | 0_{k'\lambda} \rangle \\
 &= \langle 0_{k\lambda} | \hat{a}_{\lambda}^{\dagger}(\mathbf{k}) \hat{a}_{\lambda'}(\mathbf{k}) | 0_{k'\lambda} \rangle + \delta_{ij} \delta(\mathbf{k} - \mathbf{k}') \\
 &= \delta_{ij} \delta(\mathbf{k} - \mathbf{k}')
 \end{aligned} \tag{41}$$

We'll now simply make a Basis transformation into a set of modes $\mathbf{v}_{j,\lambda}$, which are centered around a particular wave-vector \mathbf{k}_j and which themselves form an orthonormal basis. We can then decompose the quantum field $\hat{\mathbf{A}}$ into these modes using a Bogolioubov transformation

$$\begin{aligned}
 \hat{b}_{j\lambda} &= \sqrt{\frac{\epsilon_0}{\hbar}} (\mathbf{v}_{\lambda j}, \hat{\mathbf{A}}) \\
 &= \sum_{\lambda'} \int d\mathbf{k} \left(\alpha_{j\lambda\lambda'}(\mathbf{k}) \hat{a}_{\lambda}(\mathbf{k}) + \beta_{j\lambda\lambda'}(\mathbf{k}) \hat{a}_{\lambda'}^{\dagger}(\mathbf{k}) \right)
 \end{aligned} \tag{42}$$

By construction the new modal operators $\hat{b}_{j\lambda}$ (i.e. the new creation and annihilation operators), fulfil the commutation relations

$$\begin{aligned}
 [\hat{b}_{j\lambda}, \hat{b}_{j'\lambda'}^{\dagger}] &= \delta_{\lambda\lambda'} \delta_{jj'} \\
 [\hat{b}_{j\lambda}, \hat{b}_{j'\lambda'}] &= [\hat{b}_{j\lambda}^{\dagger}, \hat{b}_{j'\lambda'}^{\dagger}] = 0
 \end{aligned} \tag{43}$$

Which can be checked by just plugging the definition of the $\hat{b}_{j\lambda}$ into the commutators. For the sake of simplicity we can in most cases construct a set of modes such that $\beta_j(\mathbf{k}) = 0$ and of course we know $\sum_{\lambda'} \int d\mathbf{k} |\alpha_{j\lambda\lambda'}(\mathbf{k})|^2 = 1$. Thus we can now construct the first Fock-Modes $|1_{j\lambda}\rangle$, which belongs to the mode $\mathbf{v}_{\lambda j}$ by applying its associated creation operator $\hat{b}_{j\lambda}^{\dagger}$ onto the quantum vacuum state:

$$\begin{aligned}
 |1_{j\lambda}\rangle &= \hat{b}_{j\lambda}^{\dagger} |0\rangle \\
 &= \sum_{\lambda'} \int d\mathbf{k} \alpha_{j\lambda\lambda'}^*(\mathbf{k}) \hat{a}_{j'}^{\dagger}(\mathbf{k}) |0\rangle
 \end{aligned} \tag{44}$$

We can now test this state of the quantum field for its normalizability

$$\begin{aligned}
 \langle 1_{j\lambda} | 1_{j\lambda} \rangle &= \sum_{\mu\mu'} \int d\mathbf{k} d\mathbf{k}' \alpha_{j\lambda\mu}(\mathbf{k}) \alpha_{j\lambda\mu'}^*(\mathbf{k}') \langle 0 | \hat{a}_{\mu}(\mathbf{k}) \hat{a}_{\mu'}^{\dagger}(\mathbf{k}') | 0 \rangle \\
 &= \sum_{\mu\mu'} \int d\mathbf{k} d\mathbf{k}' \alpha_{j\lambda\mu}(\mathbf{k}) \alpha_{j\lambda\mu'}^*(\mathbf{k}') \langle 0 | \hat{a}_{\mu'}^{\dagger}(\mathbf{k}') \hat{a}_{\mu}(\mathbf{k}) + \delta_{\mu\mu'} \delta(\mathbf{k} - \mathbf{k}') | 0 \rangle \\
 &= \sum_{\mu\mu'} \int d\mathbf{k} d\mathbf{k}' \alpha_{j\lambda\mu}(\mathbf{k}) \alpha_{j\lambda\mu'}^*(\mathbf{k}') \delta_{\mu\mu'} \delta(\mathbf{k} - \mathbf{k}') \langle 0 | 0 \rangle \\
 &= \sum_{\mu} \int d\mathbf{k} \alpha_{j\lambda\mu}(\mathbf{k}) \alpha_{j\lambda\mu}^*(\mathbf{k}) \\
 &= 1
 \end{aligned} \tag{45}$$

This is now well-behaved! Keep in mind that the function $\alpha_j(\mathbf{k})$ may be very localized, such that from an experimental point of view here is very little difference to a plane wave here. We'll therefore in the future often forget the difference between $\hat{b}_{j\lambda}$ and $\hat{a}_{\lambda}(\mathbf{k})$. We will later do the same for the temporal structure of the mode; let's call this "modal doublethink". If you are really worried about this, then you are a good mathematician. Good on you.

Let us now use the normalized mode operators to properly normalize the respective Fock states with $n > 1$, which we could not do previously. The result of this process is:

$$\begin{aligned}
 |n_{j\lambda}\rangle &= \frac{1}{\sqrt{n!}} (\hat{b}_{j\lambda}^\dagger)^n |0\rangle \\
 \hat{b}_{j\lambda} |n_{j\lambda}\rangle &= \sqrt{n} |n-1_{j\lambda}\rangle \\
 \hat{b}_{j\lambda}^\dagger |n_{j\lambda}\rangle &= \sqrt{n+1} |n+1_{j\lambda}\rangle
 \end{aligned} \tag{46}$$

Combining these equations, the reader can readily verify that the Fock states are Eigenstates to the *photon number operator*

$$\begin{aligned}
 \hat{n}_{j\lambda} &= \hat{b}_{j\lambda}^\dagger \hat{b}_{j\lambda} \\
 \hat{n}_{j\lambda} |n_{j\lambda}\rangle &= n |n_{j\lambda}\rangle
 \end{aligned}$$

That is, the position on the ladder (or more precisely, the number total number of creation operators n , which are required to create a certain state) can be determined by applying the number operator. Note that this kind of behaviour gives rise to the notion of the PHOTON, namely that the natural states of the modes are discrete excitations with a fixed energy defined by the frequency/wavelength of that mode, that can be created and annihilated in a certain way.³

So far, we have only considered states where all excitations are in a single mode of the field, i.e. only one mode was occupied with one or more photons. To describe field excitations across multiple modes we assign each mode $\hat{a}_\lambda(\mathbf{k})$ an independent Hilbert space (spanned by e.g. the Fock-States in this mode). This way, multi-mode number states of the quantum field may be denoted as the tensor product of the respective Fock States:

$$|n_{j_1\lambda_1}^{(1)}\rangle \otimes |n_{j_2\lambda_2}^{(2)}\rangle \dots \dots \otimes |n_{j_M\lambda_M}^{(M)}\rangle$$

Or in an equivalent shorthand notation:

$$|n_{j_1\lambda_1}^{(1)}, \dots, n_{j_M\lambda_M}^{(M)}\rangle$$

These states are now eigenstates to the *total photon number operator*

$$\hat{n} = \sum_\lambda \int d\mathbf{k} \hat{a}_\lambda^\dagger(\mathbf{k}) \hat{a}_\lambda(\mathbf{k}) \tag{47}$$

Which has a well-defined meaning for Fock-States in a particular mode $|n_{j\lambda}\rangle$ and also for multimode Fock-States across different modes. The multi-mode Fock states are now Eigenstates of this operator

$$\hat{n} |n_{j_1\lambda_1}^{(1)}, \dots, n_{j_M\lambda_M}^{(M)}\rangle = \left(\sum_{m=1}^M n_{j_m\lambda_m} \right) |n_{j_1\lambda_1}^{(1)}, \dots, n_{j_M\lambda_M}^{(M)}\rangle \tag{48}$$

with the eigenvalue $n = (\sum_{m=1}^M n_{k_m, \lambda_m})$, i.e. the total number of photons in all relevant field modes. Notice that, unlike in the single-mode case, the eigenvalue spectrum is now degenerate; that is, there are many possible multi-mode Fock states that correspond to the same eigenvalue.

As a simple example let us pick out three arbitrary field modes, with the mode excitations:

$$|1_{j_1\lambda_1}, 0_{j_2\lambda_2}, 2_{j_3\lambda_3}\rangle$$

³ Edwin: Trude, how can we create a photon? Trude: By applying a creation operator to a field state.

$$\begin{aligned} &|3_{j_1\lambda_1}, 0_{j_2\lambda_2}, 0_{j_3\lambda_3}\rangle \\ &|2_{j_1\lambda_1}, 1_{j_2\lambda_2}, 0_{j_3\lambda_3}\rangle \end{aligned}$$

We note that the states are now all Eigenstates of the total photon number operator, with Eigenvalue $n=3$. Hence any superposition of these states

$$|\Psi_n\rangle \propto |1_{j_1\lambda_1}, 0_{j_2\lambda_2}, 2_{j_3\lambda_3}\rangle + |1_{j_1\lambda_1}, 0_{j_2\lambda_2}, 2_{j_3\lambda_3}\rangle + |1_{j_1\lambda_1}, 0_{j_2\lambda_2}, 2_{j_3\lambda_3}\rangle$$

is also an Eigenstate of the total photon number operator $\hat{n} |\Psi_n\rangle = n |\Psi_n\rangle$. In general, the state $|\Psi_n\rangle$ is not an eigenstate of the Hamiltonian, due to the different Energy $\hbar\omega(\mathbf{k})$ associated with each mode excitation, i.e. $\hat{H} |\Psi_n\rangle \neq E_n |\Psi_n\rangle$.

These results now warrant a bit of interpretation, some of which is already hidden in the naming convention for the various operators, states, and eigenvalues. Let's try and summarize the findings in some straightforward bullet points:

- the Quantum Vector Potential, or Field Operator $\hat{\mathbf{A}}$
 - is composed of modal fields $\mathbf{v}_{\lambda j}$
 - and modal expansion coefficients $\hat{b}_{j\lambda}$, which are operators
 - defines the magnetic and electric quantum field operators $\hat{\mathbf{E}}, \hat{\mathbf{B}}$ via Maxwell Equn.
- the modal amplitude $\mathbf{v}_{\lambda j}$
 - is exactly the mode from classical electrodynamics
 - has a frequency ω_j
 - retains all classical properties related to scalar products, completeness, normalization, and modal transformations
- the modal expansion operators $\hat{b}_{j\lambda}$
 - fulfil bosonic commutation relations
 - each operator evolves according to $\exp(-i\hbar\omega_j)$
 - modal transformations mix the modal expansion operators, the new modal expansion operators $\hat{c}_{j\lambda}$ fulfil equal relations (i.e. there is not preferred set of modes)
- Fock-states
 - are a discrete and complete set of eigenstates to the Hamiltonian operator
 - can be numbered for each mode $\mathbf{v}_{\lambda j}$ with an index $|n_{j,\lambda}\rangle$; a mode is then said to be populated with n photons
 - each number contributes a discrete amount of energy $\hbar\omega_j$ to the total energy of the system
 - the $\hat{b}_{j,\lambda}^\dagger$ operator creates one photon in mode $\mathbf{v}_{\lambda j}$
 - the $\hat{b}_{j,\lambda}$ operator destroys one photon in mode $\mathbf{v}_{\lambda j}$

1.2.5 Coherent States

In the last chapter we have introduced Fock-States, which are eigenstates to both, the Hamilton-Operator (i.e. the energy of the system) as well as the photon number operator. We have also seen that they can be created from the quantum vacuum state $|0\rangle$ by repeated application of the photon creation operator \hat{a}^\dagger for any given mode (note that in this chapter we only consider a single mode and suppress the modal index).

Fock-states, are, however, fairly rare in nature (in fact, Fock-States with large numbers of photons in any given mode are extremely hard to produce!). The deeper reason being, that photons are typically produced in a random process, where a large number of emitters is each emitting a photon with a certain non-unity chance (the prime example is the amplification process in a laser). This randomness naturally leads to uncertainty in the photon number of a so-produced "coherent state of light" (one

All notes subject to change, no guarantee to correctness, corrections welcome.

can already kind of guess that the resulting state of the field should have a Poisson-distribution of the photon numbers).

We will nevertheless utilize the Fock-States, as we have seen that they are a complete set of eigenstates to the state of any given mode as a basis to construct new coherent states from. We will construct a new set of modes from a superposition of these Fock-States for a single mode, by application of a superposition of creation/annihilation-operators to the vacuum state:

$$\widehat{D}(\alpha) = \exp\{\alpha\hat{a}^\dagger - \alpha^*\hat{a}\} \quad (49)$$

Where α is a complex number and $\widehat{D}(\alpha)$ is a unitary operator (we will see in the next chapter, that this is a necessary requirement for such a generation operator). In this case $\widehat{D}(\alpha)$ is called the "Glauber displacement operator". Unitarity can be easily proven by checking the following relations:

$$\widehat{D}^\dagger(\alpha) = \widehat{D}^{-1}(\alpha) = \widehat{D}(-\alpha) \quad (50)$$

Let's now rewrite the operator, using the commutation relation $[\hat{a}, \hat{a}^\dagger]=1$:

$$\begin{aligned} \widehat{D}(\alpha) &= \exp\left\{\alpha\hat{a}^\dagger - \alpha^*\hat{a} - \frac{1}{2}[\alpha\hat{a}^\dagger, -\alpha^*\hat{a}] + \frac{1}{2}[\alpha\hat{a}^\dagger, -\alpha^*\hat{a}]\right\} \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\left\{\alpha\hat{a}^\dagger - \alpha^*\hat{a} - \frac{1}{2}[\alpha\hat{a}^\dagger, -\alpha^*\hat{a}]\right\} \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\left\{\hat{A} + \hat{B} + \frac{1}{2}[\hat{A}, \hat{B}]\right\} \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\{\hat{A}\} \exp\{\hat{B}\} \\ &\Leftrightarrow [[\hat{A}, \hat{B}], \hat{A}] = [[\hat{A}, \hat{B}], \hat{B}] = 0 \text{ with } \hat{A} = \alpha\hat{a}^\dagger, \hat{B} = -\alpha^*\hat{a} \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\{\alpha\hat{a}^\dagger\} \exp\{-\alpha^*\hat{a}\} \end{aligned} \quad (51)$$

Then we can apply this reformulated version of the Glauber-Operator $\widehat{D}(\alpha)$ on the Vacuum-State quite easily to get a better understanding on the state of the field.

$$\begin{aligned} |\alpha\rangle = \widehat{D}(\alpha)|0\rangle &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\{\alpha\hat{a}^\dagger\} \exp\{-\alpha^*\hat{a}\} |0\rangle \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \exp\{\alpha\hat{a}^\dagger\} |0\rangle \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \sum_n \frac{\alpha^n (\hat{a}^\dagger)^n}{n!} |0\rangle \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle \end{aligned} \quad (52)$$

This means, that the field is in a superposition of Fock-States and the probability $P(n)$ (amplitude square!!!) of finding the field in an $|n\rangle$ state is given by the Poisson distribution:

$$\begin{aligned} P(n) &= |\langle n|\widehat{D}(\alpha)|0\rangle|^2 \\ &= \frac{\exp\{-|\alpha|^2\} |\alpha|^{2n}}{n!} \\ &= P_{\text{Poisson}}(n, |\alpha|^2) \end{aligned} \quad (53)$$

From probability theory we know, that a series of Poisson-distributed events is maximally random, i.e. the occurrence of an event (i.e. the appearance of a photon) at any given point in time in a certain mode does by no means make the time of appearance of another photon more or less probable. In this respect, coherent states have no memory, photons are neither bunched, nor anti-bunched.

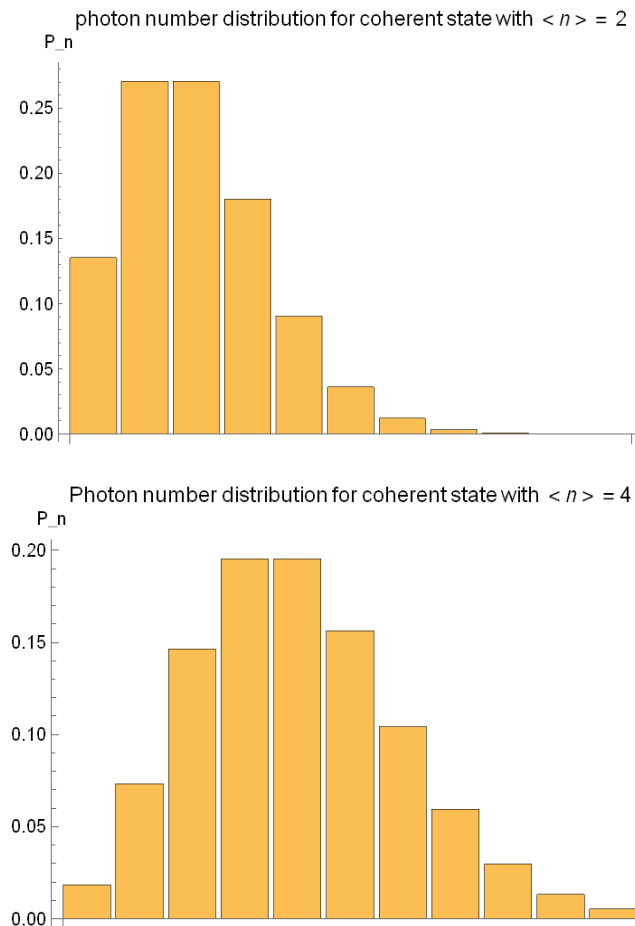


Fig. 3: Photon-Number Probabilities of two different coherent states.

We can quite easily find the expectation value and variance of the photon number operator:

$$\begin{aligned}
 \langle \hat{n} \rangle &= \langle \alpha | \hat{n} | \alpha \rangle \\
 &= \exp\{-|\alpha|^2\} \sum_{n,n'} \frac{\alpha^n}{\sqrt{n!}} \frac{\alpha^{*n'}}{\sqrt{n'!}} \langle n' | \hat{n} | n \rangle \\
 &= \exp\{-|\alpha|^2\} \sum_n \frac{|\alpha|^{2n}}{n!} n \\
 &= |\alpha|^2 \\
 (\Delta n)^2 &= \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2 \\
 &= |\alpha|^4 + |\alpha|^2 - |\alpha|^4 \\
 &= |\alpha|^2
 \end{aligned} \tag{54}$$

This has a few major ramifications. Ordinary light sources emit states of light, which do NOT have a defined number of photons. If you measure the energy you get so-called "shot-noise" even for a perfect detector, which limits the measurement accuracy. Examples:

- 10 μ W Signal on a 10 GHz Communication Channel $\rightarrow 10^{-15}$ J per time slot \rightarrow roughly 10^{-18} J per photon for light with a wavelength of 1000 nm \rightarrow 1000 Photons and a shot noise floor of

All notes subject to change, no guarantee to correctness, corrections welcome.

$\sqrt{1000} \approx 30$ photons \rightarrow SNR of roughly 30; no more than $\log_2 SNR \approx 5$ bits per time slot possible for fundamental information theoretical reasons

- low-Light image with roughly 10 Photons per pixel per frame \rightarrow 3 Photons Shot Noise \rightarrow 30 % Noise floor

Both Communication- as well as Imaging can profit from the usage of Fock-States. Particularly the latter one is a goal of Quantum-Imaging and a hot topic in research.

Let's now proceed to a few more properties of coherent states. First, they are robust against mixing (i.e. amplification and damping):

$$\hat{D}(\beta)|\alpha\rangle = |\alpha + \beta\rangle \quad (55)$$

Coherent states are also complete:

$$\int_{\mathbb{C}} \frac{d^2\alpha}{\pi} |\alpha\rangle\langle\alpha| = 1 \quad (56)$$

They are also eigenstates of the annihilation operator \hat{a}

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \quad (57)$$

We shall later see, that the time evolution of the any state is given by the application of the time evolution operator $\exp\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)$, in this case this yields:

$$\begin{aligned} \exp\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)|\alpha\rangle &= \exp\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)\exp\left\{-\frac{|\alpha|^2}{2}\right\}\sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\}\sum_n \frac{\alpha^n}{\sqrt{n!}}\exp\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)|n\rangle \\ &= \exp\left\{-\frac{|\alpha|^2}{2}\right\}\sum_n \frac{\alpha^n}{\sqrt{n!}}\exp\{-i\omega nt\}|n\rangle \text{ with } \hat{\mathcal{H}}|n\rangle = \hbar\omega n|n\rangle \\ &= \sum_n \exp\left\{-\frac{|\alpha|^2}{2}\right\}\frac{[\alpha \exp\{-i\omega t\}]^n}{\sqrt{n!}}|n\rangle \\ &= |\alpha \exp\{-i\omega t\}\rangle \end{aligned} \quad (58)$$

The coherent states do thus have a time evolution, which can be represented by a rotation in the α -plane, where the rate of rotation is only depended on the modes frequency.

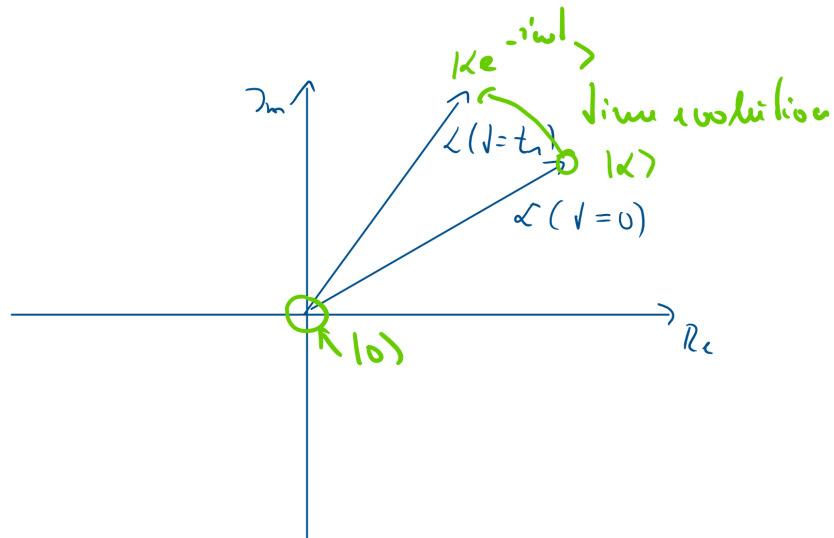


Fig. 4: Representation of coherent states and their evolution in the complex plane. Note that the axes correspond to conjugated variables and roughly to the amplitude of the field and its phase, i.e. the electric and magnetic fields.

1.3 Time Evolution

The field of science is called Quantum ElectroDYNAMICS. Let's thus take a closer look at the evolution of the field and find out, how this is related to the Hamiltonian-Operator, for as of now, we have only looked into static scenarios – with the exception that a part of the time evolution is already covered in the time-dependence of the modes of the electric fields.

We will then find, that this concept can indeed be expanded into arbitrary Interaction-Hamiltonians, which may describe real-world optical elements, such as beam-splitters, loss-elements and the like. As a last part we shall introduce time-bin operators in temporally localized fields.

1.3.1 Heisenberg Equations of Motion

In chapter 1.2.3 we had already – in a very brief manner – discussed the time evolution of operators. This shall be extended here. Note that we will operate in the Heisenberg-picture here, meaning that we treat the time-evolution of any field as an evolution of the operators \hat{a} and \hat{a}^\dagger . This approach is very natural for a quantum electrodynamics, as per the last chapter we have a thorough understanding of the relation of these operators with respect to the notion of photons of modes.

Just like in the aforementioned chapter 1.2.3, we start with the understanding that in classical mechanics the total time derivative of any not-explicitly time-dependent measurable $A(q_k, p_k)$ is given by its Poisson-Bracket with the system's Hamiltonian. By construction the same is true for quantum mechanics, with the difference that the we have to replace the Poisson-bracket with $-i/\hbar$ and the commutator and the measurable is now a Hermitian operator \hat{A}

$$\frac{d\hat{A}}{dt} = \frac{i}{\hbar} [\mathcal{H}, \hat{A}] \quad (59)$$

We can repeatedly apply this relation to get higher order derivatives of the operator \hat{A} , e.g.

$$\begin{aligned} \frac{d^2\hat{A}}{dt^2} &= \frac{d}{dt} \frac{d\hat{A}}{dt} = \left(\frac{i}{\hbar}\right)^2 [\mathcal{H}, [\mathcal{H}, \hat{A}]] \\ \frac{d^3\hat{A}}{dt^3} &= \frac{d}{dt} \frac{d^2\hat{A}}{dt^2} = \left(\frac{i}{\hbar}\right)^3 [\mathcal{H}, [\mathcal{H}, [\mathcal{H}, \hat{A}]]] \end{aligned} \quad (60)$$

From this we can reverse-engineer the explicit relation for the time-dependence of the operator, by writing it as a Taylor-series as a function of the time coordinate (we assume without loss of generality that the point of the Taylor-series expansion is a $t = 0$).

$$\begin{aligned} \hat{A}(t) &= \hat{A}(t=0) + t \left. \frac{d\hat{A}}{dt} \right|_{t=0} + t^2 \frac{1}{2!} \left. \frac{d^2\hat{A}}{dt^2} \right|_{t=0} + \dots \\ &= \hat{A} + \left(\frac{i}{\hbar}t\right) [\mathcal{H}, \hat{A}] + \frac{1}{2!} \left(\frac{i}{\hbar}t\right)^2 [\mathcal{H}, [\mathcal{H}, \hat{A}]] + \dots \end{aligned} \quad (61)$$

The last expression may seem complicated but it is the exact representation of the Baker-Campbell-Hausdorff-Theorem (which is, in fact, true for complex expansion parameters and non-hermitic operators. Using the BCH-theorem we get:

$$\begin{aligned} \hat{A}(t) &= e^{\frac{i}{\hbar}t\mathcal{H}} \hat{A}(t=0) e^{-\frac{i}{\hbar}t\mathcal{H}} \\ &= U(t) \hat{A} U^\dagger(t) \end{aligned} \quad (62)$$

The exponential (and thus unitarian) form of the Hamilton-operators is, due to this relation is called the “generator operator”.

In the Heisenberg picture an arbitrary Hermitian operator \hat{A} evolves in time under the influence of the time evolution operator $\hat{U}(t)$.

1.3.2 Temporal Wavepackets

We can now use these findings to introduce temporally localized wavepackets. In chapter 1.2.3 we had found that

$$\begin{aligned}\hat{a}_\lambda(\mathbf{k}; t) &= \hat{a}_\lambda(\mathbf{k}; t = 0) \exp(i \omega(\mathbf{k})t) \\ \hat{a}_\lambda^\dagger(\mathbf{k}; t) &= \hat{a}_\lambda^\dagger(\mathbf{k}; t = 0) \exp(-i \omega(\mathbf{k})t)\end{aligned}\quad (63)$$

We shall now generalize this to introduce the time-dependent annihilation operator, by integrating over all possible modes (which are no longer time dependent). We do so for a given polarization λ and at a fixed position $r = 0$. For plane waves in vacuum this is not a problem for all other types of waves it's not a big problem.

$$\hat{a}_\lambda(t, r = 0) = \int d\mathbf{k} \hat{a}_\lambda(\mathbf{k}; t) = \int d\mathbf{k} \hat{a}_\lambda(\mathbf{k}) \exp(i \omega(\mathbf{k})t) \quad (64)$$

The same can be done for the time-dependent number operator:

$$n_\lambda(t, r = 0) = \hat{a}_\lambda^\dagger(t) \hat{a}_\lambda(t) = \int d\mathbf{k} \int d\mathbf{k}' \hat{a}_\lambda(\mathbf{k}') \hat{a}_\lambda^\dagger(\mathbf{k}) \exp(i(\omega(\mathbf{k}) - \omega(\mathbf{k}'))t) \quad (65)$$

Assume that we have a temporal wave packet $|1_{j,\lambda}\rangle$, denoted with some wave packet-identification index j in a polarization λ , which is filled with exactly one photon, i.e:

$$|1_{j,\lambda}\rangle = \int d\mathbf{k} \alpha_j(\mathbf{k}) \hat{a}_\lambda^\dagger(\mathbf{k}) |vac\rangle \quad (66)$$

Obviously $\alpha_j(\mathbf{k})$ has to be defined in the very same way as was done in chapter 1.1.7, i.e. they have to be normalized appropriately for any mode function. For this we can simply take the wavepackets defined in 1.1.7. Then the expectation value of the time-dependent number operator is simply:

$$\langle n \rangle(t, r = 0) = \langle 1_{j,\lambda} | n_\lambda(t) | 1_{j,\lambda} \rangle = \left| \int d\mathbf{k} \alpha_j(\mathbf{k}) \exp(-i \omega(\mathbf{k})t) \right|^2. \quad (67)$$

This means that the absolute value of the number density operator expectation value of a single-photon wavepacket is given by the absolute value square of the Fourier-transform of its spectrum. Just as we would expect from classical wave-theory.

2 Fundamentals

What is *quantum* communication? Since every communication system is ultimately described by quantum physics, the answer to this question cannot be exclusively related to the inner workings of hardware alone. What makes a quantum communication system different from classical systems, is thus better defined at the operational resource level:

A quantum communication system uses distributed quantum information to perform specific information processing tasks in a way that would not be possible using classical resources alone.

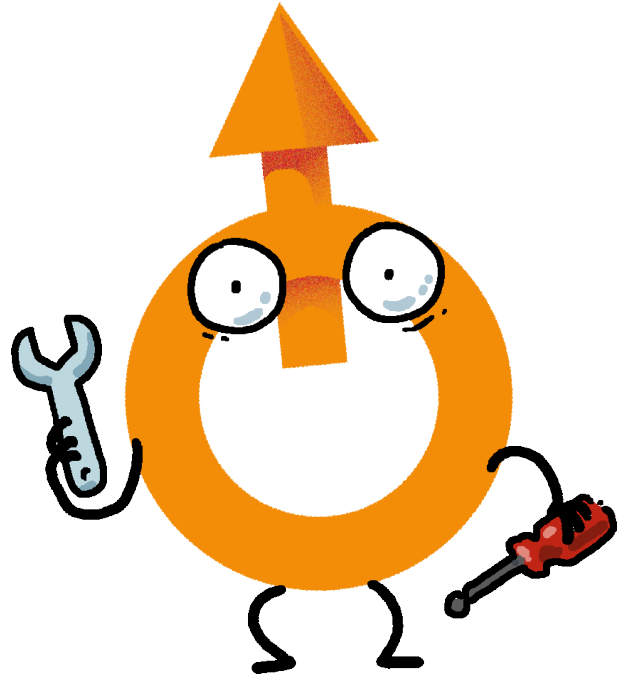
While state-of-the-art classical communication systems may well leverage quantum technology at the level of hardware (lasers, semiconductor technology, photodetection), they do not take advantage of quantum principles at the level of information distribution or processing itself. Not yet.

To understand what these key quantum principles are, let us cast some of the key concepts of quantum theory into a slightly different light – that of quantum information theory. We will assume some level of familiarity with linear algebra and probability theory extensively throughout. The reader is encouraged to consult the standard quantum theory textbooks for a review if deemed necessary.

2.1 The principles of quantum theory

Quantum theory provides a set of tools for calculating **probabilities for outcomes of measurements**⁴ applied to a certain state of the quantum system to be measured. A measurement corresponds to anything we may observe in a laboratory using a suitable measurement apparatus we may potentially build. Mathematically such an apparatus is represented by a so-called observable. Let's briefly state the principles and then look into them in a little more detail:

- States of a quantum physical system are completely described via a vector $|\psi\rangle$ in a linear vector space \mathcal{H} with a complex-valued inner product $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^* \in \mathbb{C}$
- Observable quantities are described by linear, hermitian operators \hat{O} with exclusively real eigenvalues. These eigenvalues $\{a_1, a_2, a_3, \dots\}$ denote the possible outcomes of a measurement (hence they can only be real).
- If the measurement outcome is guaranteed to be a_i , then the system is in the Eigenstate of $\hat{O}|i\rangle = a_i|i\rangle$.
- The Eigenstates $|i\rangle$ form a complete basis of the Hilbert Space. They are orthonormal $\langle i|j\rangle = \delta_{ij}$.
- Not all measurement outcomes need to be different, there may be a subset of eigenstates which are degenerate.



⁴ And nothing more. If you find that non-satisfactory, then deal with it. We shall later see that this is not a problem of the theory but the very essence of nature itself as can be tested in e.g. a Bell measurement.

- States that can be distinguished with certainty are orthogonal $\langle \phi | \psi \rangle = 0$. (the opposite is not always true, see above). Distinguishability here means, that they produce different results upon a specific measurement \hat{O} .
- The overlap of a state vector $|\psi\rangle$ with an eigenvector $|i\rangle$ determines the probability that the eigenvalue a_i will be observed, specifically: $P(a_i) = |\langle \psi | i \rangle|^2$.
- The time-evolution of an isolated quantum system conserves overlap between state vectors, i.e.: $\langle \phi(t=0) | \psi(t=0) \rangle = \langle \phi(t=T) | \psi(t=T) \rangle$, i.e. unperturbed quantum states retain their degree of (dis-)similarity.

2.1.1 Quantum States

The quantum state of a physical system is specified by a vector in Hilbert Space \mathcal{H} - that is – a complex vector space that is equipped with an inner product (a linear mapping from two vectors into complex numbers; $\mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$). For historical reasons we do not denote the state vector using more familiar notation, e.g. $\vec{\psi}$, but instead the Dirac “ket” notation:

$$|\psi\rangle \tag{68}$$

According to quantum theory, the state vector represents a state of *complete knowledge* about the preparation of the physical system – i.e. everything that we need to know, and everything that is principle knowable. Implicit in the structure of the linear vector space structure is the following statement: If $|\psi_1\rangle$ and $|\psi_2\rangle$ are possible quantum states, then so is any superposition state:

$$|\psi\rangle = \alpha_1 |\psi_1\rangle + \alpha_2 |\psi_2\rangle \tag{69}$$

with complex amplitudes α_1 and α_2 . While this may look trivial, it is arguably the most profound concept in quantum theory: the superposition principle is not only the culprit responsible for much quantum *weirdness* such a quantum nonlocality or the Heisenberg uncertainty principle, it is also the key feature in many quantum-enhancements such as exponential speedups in computing and tap-proof communication.

In the following we will see that experimentally accessible quantities, such as expectation values and probabilities are described by numbers and not the state vectors themselves. Or to put it more bluntly: you cannot measure the state $|\psi\rangle$ by any conceivable means. To arrive at these, we need a mapping from vectors to numbers, i.e. an inner product. Denoting the dual vector to $|\psi\rangle$ by the Dirac “bra”:

$$\langle \psi | = |\psi\rangle^\dagger = \alpha_1^* \langle \psi_1 | + \alpha_2^* \langle \psi_2 | \tag{70}$$

The inner product can be written conveniently as a „bra-ket“:

$$\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^* \tag{71}$$

In particular, the norm of a vector is a real number $\langle \psi | \psi \rangle > 0$.

2.1.2 Observables

Observable physical quantities, so-called *observables* or *measurables*, are described using linear operators. Possible measurement outcomes correspond to the eigenvalues $\{a_1, a_2, a_3, \dots\}$ of an operator \hat{A} . In accordance with our every-day lab experience, we can think of measurement outcomes as numbers on a read-out device. Extending this idea to the measurement of a quantum system, we may require measurement outcomes to be real-valued numbers $a_i \in \mathbb{R}$ (and not, e.g. complex numbers, as these can be described by two real numbers, i.e. give two measurement outcomes).

If we prepare a quantum system in a way that the measurement outcome for observable is a_i with certainty, then the system is in an Eigenstate of the operator:

$$\hat{A}|i\rangle = a_i|i\rangle \quad (72)$$

Moreover, if we obtain a result, say a_i , we expect a subsequent identical measurement to give the same result, and not some other result, say a_j . In other words, the measurement outcomes $a_i \neq a_j$ should correspond to unambiguously distinguishable quantum states. As stated initially, we require that states that can be distinguished with certainty are orthogonal; which implies that the eigenvectors corresponding to different eigenvalues (measurement outcomes) should be orthogonal, i.e. we have $\langle i|j\rangle = \delta_{ij}$. These requirements restrict the type of operator we can use to describe the measurement process; Operators that represent observables are so-called Hermitian operators:

$$\hat{A} = \hat{A}^\dagger \quad (73)$$

where the Hermitian conjugate \hat{A}^\dagger of operator \hat{A} is defined by the requirement

$$\langle \phi | \hat{A}^\dagger | \psi \rangle = \langle \psi | \hat{A} | \phi \rangle^* \quad \forall |\psi\rangle, |\phi\rangle \quad (74)$$

It is quite easy to show that Hermitian operators have orthogonal set of eigenvectors with real-valued eigenvalues; so let's do it. First, we show that the eigenvalues are real:

$$\langle i | \hat{A} | i \rangle = \langle i | \hat{A} | i \rangle^* \rightarrow a_i \langle i | i \rangle = a_i^* \langle i | i \rangle^* \rightarrow a_i = a_i^* \quad (75)$$

where we have used the fact that the norm is a real number. To prove that eigenvectors of a Hermitian operator with eigenvalues $a_i \neq a_j$ are orthogonal, we calculate conjugate of the Eigenvalue equation

$$\langle i | \hat{A}^\dagger = \langle i | \hat{A} = a_i^* \langle i | \quad (76)$$

And evaluate its inner product with an eigenvector $|j\rangle$,

$$\langle i | \hat{A}^\dagger | j \rangle = \langle i | \hat{A} | j \rangle = a_i \langle i | j \rangle \quad (77)$$

Similarly, taking the overlap of $\hat{A}|j\rangle = a_j|j\rangle$ with eigenvector $|i\rangle$ we have:

$$\langle i | \hat{A} | j \rangle = a_j \langle i | j \rangle \quad (78)$$

Subtraction of the two results shows that $\langle i | j \rangle = 0$. After normalizations can thus construct an orthonormal basis consisting of eigenvectors $\langle i | j \rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta symbol.

2.1.3 The Born rule

So far, we have discussed what happens when we measure an observable a physical system that is prepared in an Eigenstate, i.e. a situation that gives a definitive outcome. Now let us consider the more general situation in which state of the system prior to measurement is described by a normalized vector $|\psi\rangle$.

The *Born rule* states the following: For a system described by any normalized state vector $|\psi\rangle = \sum \alpha_i |i\rangle$ ($\sum |\alpha_n|^2 = 1$), the measurement of observable \hat{A} will yield an outcome a_j with a *probability* that is given by the modulus of the overlap with the eigenstate $|j\rangle$, i.e.:

$$P(a_j) = |\langle \psi | j \rangle|^2 = |\alpha_j|^2.$$

If we repeat many such experiments with identically prepared quantum systems, then we obtain the expected value for the observable:

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \sum_n P(a_n) a_n \quad (79)$$

Fun fact: this axiomatic rule was guessed by Max Born in one of his papers and stated there as a footnote only. Take this as a hint to by default not skip footnotes.

2.1.4 Projection operators

Hermitian operators can be decomposed into orthogonal projection operators $\hat{P}_i = |i\rangle\langle i|$

$$\hat{A} = \sum \hat{P}_i a_i \quad (80)$$

These projection operators fulfil the relations:

$$\begin{aligned} \hat{P}_i^2 &= \hat{P}_i \\ \hat{P}_i \hat{P}_j &= \delta_{ij} \hat{P}_i \end{aligned} \quad (81)$$

We can think of a projection operator as an elementary observable that essentially “asks” the quantum system the question: “will you give me result a_i ”. The operators’ eigenvalues (1 and 0) can be interpreted as the response (yes/no) to such a query:

$$\hat{P}_i |i\rangle = \sum_j \delta_{ij} |j\rangle \quad (82)$$

Phrasing the Born rule slightly differently, the probability of measuring a particular value a_j when we perform a projective measurement on a state prepared in a state $|\psi\rangle$ is the expected value of the corresponding projection operator:

$$P(a_j) = \langle \psi | \hat{P}_j | \psi \rangle = |\langle \psi | j \rangle|^2 \quad (83)$$

Whenever a measurement is made our knowledge about the state of the system also changes according to the outcome of the measurement. From the numerous potential outcomes, only one occurs in the measurement. Correspondingly the normalized post-measurement state becomes⁵:

$$|\psi\rangle \rightarrow \frac{\hat{P}_j |\psi\rangle}{\sqrt{P(a_j)}} = |j\rangle \quad (84)$$

So, when the quantum measurement has been performed, we must update the state vector according to the measurement outcome that was observed, i.e. the wave function *collapses* onto the corresponding eigenvector. **Whenever the state of the system is not an Eigenstate of the observable to be measured, the mere process of measurement will change the quantum state.**

⁵ Otherwise we could not guarantee that a repeated measurement would yield the same result, which would be contrary to what we observe in the real world.

2.1.5 Complementarity of Observables

Two observables \hat{A} , \hat{B} are said to be *complementary* (incompatible) when complete knowledge about the result of a measurement of the first means that we have absolutely no knowledge of the measurement outcome of the second. In quantum formalism, the complementarity of observables is captured using the commutator of the respective operators:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (85)$$

For any non-commuting observables $[\hat{A}, \hat{B}] \neq 0$, we can define an uncertainty relation for the expectation values of measurements:

$$\Delta\hat{A} \cdot \Delta\hat{B} \geq \frac{1}{2} |\langle [\hat{A}, \hat{B}] \rangle| \quad (86)$$

where $(\Delta\hat{A})^2 = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$. This means that a measurement in \hat{A} will randomly collapse the wavefunction into a basis, which is guaranteed to not produce a fixed result in \hat{B} . Or, from an experimental point of view: the more precisely you measure \hat{A} the more random the results of a consecutive measurement of \hat{B} will be.

2.2 On quantum optics and the Nature of Photons

The term “quantum information **distribution**” also implies that we are mostly concerned with the quantum properties of light, i.e. to the properties of photons. This is because photons are extremely robust and long-lived (after all we sometimes collect some them that are almost as old as the universe). They can also be easily created, manipulated, and detected and they are pretty damn fast. In other words: they are, in most cases, the best candidates of the distribution of quantum information.

To be able to properly speak of photons, this in principal requires a full introduction of the quantum theory of light, as is classically taught in lectures such as quantum optics. For the sake of brevity, we shall not do this here but just point to the most important results:

- The notion of modes carries over to quantum optics from classical electromagnetic theory in a 1:1 manner. All states of the (quantum) electromagnetic field can be decomposed into eigenmodes and the eigenmodes are calculated exactly the same way as they are calculated in classical theory (same scalar product, same eigenfrequency ω , etc...).
- The difference between electromagnetic theory and the quantum theory of light lies exclusively in the expansion coefficients of each mode. In classical theory this expansion coefficient is a complex number a (indicating amplitude and phase) and its conjugate a^* , whereas in quantum theory it's an operator \hat{a} and its conjugate \hat{a}^\dagger .
- The operator \hat{a} itself is not Hermitian, hence not an observable. This means that it cannot be determined by a measurement (in more physical terms: we cannot measure the amplitude of the electric field and its phase at the same point in space at the same time or in yet other terms you can either precisely measure the electric or the magnetic field at one point in space but not both).
- The absolute value of the operator $\hat{a}^\dagger \hat{a}$, however, is Hermitian and it's a measure of the energy contained in the mode. More specifically we find its eigenvalues are $E_n = \hbar\omega n$, where each value corresponds to an eigenstate of the quantum field $|n\rangle$. Each of these states is at-

tributed to the excitation of the (classical) mode with n photons. Each of these so-called number state $|n\rangle$ is orthogonal to all other number states of the (classical) mode $|n'\rangle$ such that $\langle n|n'\rangle = \delta_{nn'}$.

- The full state of the field in any one mode can be decomposed of the superpositon of these number states $|\psi\rangle = \sum_n a_n |n\rangle$, where a_n are complex numbers.

Or in summary for quantum optics each classical mode is by itself a superposition of (frequency)-degenerate photon number states; each of which can be considered as an individual mode in their own right.

Each state of the (global) quantum field $|\psi\rangle$ is thus a vector in some high-dimensional Hilbert space $|\psi\rangle \in \mathcal{H}$. It can be decomposed into the superposition of Basis states $|m\rangle$ with some arbitrary numbering scheme, such that $|\psi\rangle = \sum_m a_m |m\rangle$. The basis vectors $|m\rangle$ may be any superposition of frequency-photon-number eigenmodes. The only requirement is that the modes be orthogonal $\langle n|n'\rangle = \delta_{nn'}$ and the state itself must be normalized, e.g. $\langle\psi|\psi\rangle = 1$, i.e. $\sum_n |a_n|^2 = 1$.

More info on the general nature of Quantum Optics can be found in chapter 1 of the appendix.

2.3 Matrix representations

With the orthonormal eigenvectors we can write any state vector in terms of the orthonormal eigenvector basis, i.e.:

$$\begin{aligned} |\psi\rangle &= \sum_n \alpha_n |n\rangle \\ \langle\psi| &= \sum_n \alpha_n^* \langle n| \end{aligned} \quad (87)$$

where α_n are complex coefficients. If we group the ket coefficients into a column vector

$$|\psi\rangle \rightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \vdots \end{pmatrix} \quad (88)$$

and bra vectors into row vectors

$$\langle\psi| \rightarrow (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \dots)^* \quad (89)$$

We can express the action of any operator \hat{O} on a state vector as a simple matrix multiplication:

$$\hat{O}|\psi\rangle \rightarrow O_{ij} \alpha_j \quad (90)$$

with a matrix with elements $O_{ij} = \langle i|\hat{O}|j\rangle$

$$\hat{O} \rightarrow \begin{bmatrix} O_{11} & \dots & O_{1j} & \dots & \vdots \\ O_{21} & \dots & O_{2j} & \dots & \vdots \\ O_{31} & \dots & O_{3j} & \dots & \vdots \\ \vdots & \dots & \dots & \ddots & \vdots \\ \vdots & \dots & \dots & \dots & \ddots \end{bmatrix} \quad (91)$$

The matrix elements of the Hermitian conjugate operator are then given by transposition and complex conjugation $O'_{ij} = O_{ji}^*$. In the eigenvector basis of the observable \hat{A} , the matrix representation A_{ij} is diagonal matrix:

$$\hat{A} \rightarrow \begin{bmatrix} a_1 & 0 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & a_3 & 0 & 0 \\ 0 & 0 & 0 & a_4 & 0 \\ 0 & 0 & 0 & 0 & \ddots \end{bmatrix} \quad (92)$$

which is called the spectral decomposition of the observable. In the following we will mostly consider cases in which possible measurement outcomes are discrete and finite $\{a_1, a_2, \dots, a_n\}$, i.e. we will mostly deal with vectors of dimensionality N and matrices of dimensionality of $N \times N$. In the matrix formalism we can also better interpret functions of operators, simply

2.4 Mixed States and the density matrix

So far, we have looked into the state of a particular quantum system per-se. In reality, however, we will typically make experiments on a series of more-or-less identical copies of a system, for example to generate some kind of statistical data. In practice it may well be that any quantum system is in fact far from reproducible and will generate a different quantum state for each repetition. In a summary, we will get an ensemble of quantum states, with some degree of statistical distribution between the different pure quantum states.

In practice, many things can contribute to such effects: emitters may have multiple decay channels, dipole-vectors jitter in their orientation, various processes may lead to inhomogeneous broadening of spectroscopic lines, your helpful co-worker may occasionally change the temperature of some nonlinear crystal, just because he can. And he will. Your hands may shake slightly upon adjustment of some setup, due to a lack of Thorlabs sending lab snacks.

Such statistical ensembles of quantum states may be described with the help of the density operator

$$\hat{\rho} = \sum_i p_i \hat{\rho}_i = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (93)$$

where p_i is the probability that the quantum system is in state $|\psi_i\rangle$ and $\sum_i p_i = 1$ and $\hat{\rho}_i = |\psi_i\rangle\langle\psi_i|$ is the pure state density operator.

In reality, we are, however more interested in measurables than in the quantum state itself. Any measurable is, of course, defined by its measurement operator \hat{A} and can be characterized by expectation value $\langle\hat{A}\rangle$, which is defined as:

$$\langle\hat{A}\rangle = \sum_j p_j \langle\hat{A}\rangle_j = \sum_j p_j \text{Tr}(\hat{\rho}_j \hat{A}) = \text{Tr}\left(\sum_j p_j \hat{\rho}_j \hat{A}\right) = \text{Tr}(\hat{\rho} \hat{A}) \quad (94)$$

Where $\text{Tr}(\cdot)$ is the trace operator, i.e. the sum of the diagonal elements of the density matrix. We don't show this relation here, please look it up if you are interested. It is noteworthy that $\hat{\rho}$ (being a sum of obviously Hermitian $\hat{\rho}_i$ with real factors) can always be decomposed into eigenstates and appropriate eigenvalues, such that:

$$\hat{\rho} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \quad (95)$$

which is called the spectral decomposition of the density matrix. For example, a light source may emit 50% horizontally polarized photons and 50% diagonally upwards polarized photons, thus:

$$\begin{aligned}
 \hat{\rho} &= \frac{1}{2} |h\rangle\langle h| + \frac{1}{2} |u\rangle\langle u| \\
 &= \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
 &= \frac{1}{4} (2 + \sqrt{2}) \begin{bmatrix} \frac{1 + \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} & \frac{1}{\sqrt{4 + 2\sqrt{2}}} \\ \frac{1}{\sqrt{4 + 2\sqrt{2}}} & \frac{1}{\sqrt{4 + 2\sqrt{2}}} \end{bmatrix} \begin{bmatrix} \frac{1 + \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} \\ 1 \\ \frac{1}{\sqrt{4 + 2\sqrt{2}}} \end{bmatrix} \\
 &\quad + \frac{1}{4} (2 - \sqrt{2}) \begin{bmatrix} \frac{1 - \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} & \frac{1}{\sqrt{4 + 2\sqrt{2}}} \\ \frac{1}{\sqrt{4 + 2\sqrt{2}}} & \frac{1}{\sqrt{4 + 2\sqrt{2}}} \end{bmatrix} \begin{bmatrix} \frac{1 - \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} \\ 1 \\ \frac{1}{\sqrt{4 + 2\sqrt{2}}} \end{bmatrix} \\
 &= \frac{1}{4} (2 - \sqrt{2}) \left\{ \frac{1 - \sqrt{2}}{\sqrt{4 + 2\sqrt{2}}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{\sqrt{4 + 2\sqrt{2}}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}
 \end{aligned} \tag{96}$$

Which means that the spectrally decomposed version of this are again linear states of light. And this is surprisingly cumbersome.

Also note that:

$$\text{Tr}(\hat{\rho}) = 1 \tag{97}$$

And furthermore, for any quantum state $|\psi\rangle$, we get:

$$\langle \psi | \hat{\rho} | \psi \rangle \geq 0 \tag{98}$$

i.e. the density operator is always positive. For pure quantum state vectors the density matrix reduces to a projection operator $|\psi_i\rangle\langle\psi_i|$, for which the relation $\hat{\rho}^2 = \hat{\rho}$ is readily shown. This relation is useful as it allows us to quantify the “degree of mixedness”, i.e. the state purity:

$$\text{Purity}(\hat{\rho}) = \text{Tr}(\hat{\rho}^2) \tag{99}$$

The reader can readily verify that $\text{Purity}(|\psi_i\rangle\langle\psi_i|) = 1$ for a pure state and $\text{Purity}(\hat{\rho}_N) = \hat{1}/N$ for a completely mixed state of dimension N .

Note that the type of uncertainty here is a different one from the uncertainty introduced by the quantum measurement process. Each of these effects may in fact be fully quantified and measured, this may just be practically impossible or impractical to deal with. Also note that each of the effects, which contribute to some kind of statistical uncertainty are themselves subject to the laws of quantum physics (even your co-worker is!). They derive from a pure state and if the system is large enough, they are unaffected by external noise. Thus, any mixed state can be purified into a pure state of a larger system. We won't show the mathematical proof here.

2.4.1 Entropy in Quantum Physics

In classical physics there is an intricate relation between the notion of Entropy and Information in a System. If you are more interested in that please consult the seminal works by Landauer. We'll just summarize here: the more entropy a system has, the more information it contains. I typically think about the room of my kids: if there are toys lying around everywhere there is lots of information in the

room (e.g. to describe which toy is where takes a loooong time), whereas if the room is cleaned up you can describe it with a single piece of info: everything is, where it belongs⁶.

We would now like extend this concept to quantum physics and the idea that a pure quantum state is a minimum information/entropy state a little more formally. For a pure state, where we have complete information of the preparation procedure, we expect a measure describing disorder (if you're from a physics background) or information content (if you're a telecom engineering background) to be minimized. The von Neumann entropy is the extension of the concept of entropy from classical thermodynamics (Gibbs entropy) or information theory (Shannon entropy) to the quantum realm. It is defined as:

$$S(\hat{\rho}) = -\text{Tr}\{\hat{\rho} \text{Log}(\hat{\rho})\} \quad (100)$$

It is straightforward to verify that the von Neumann entropy⁷ of a physical system prepared in any pure quantum state $|\psi\rangle$ is zero:

$$S(|\psi\rangle\langle\psi|) = 0 \quad (101)$$

With the pure quantum states thus corresponding to minimum information. The state of maximum confusion, i.e. the opposite of a pure state, is the maximally mixed state in which each eigenstate of the system $|i\rangle$ appears with equal likelihood:

$$\hat{\rho}_M = \frac{1}{N} \sum_i |i\rangle\langle i| = \frac{\hat{1}}{N} \quad (102)$$

where $\hat{1}$ is the unit operator and N is the dimension of the state space. This is the state of maximum entropy in a Hilbert space of dimension N :

$$S(\hat{\rho}_M) \propto \log(N) \quad (103)$$

Hence you can see that the concept of the impurity of the state is closely related to the entropy of a quantum system. When you think about this for a while you can come to a few nifty conclusions on the relation of entropy, information and the nature of coincidences:

There are two *distinguishable* types of randomness in a quantum measurement: If you make measurements on a mixed state you have two contributions to the statistics of the measurement: the statistics of the quantum measurement process and the classical ensemble statistics that comes from the mixed'ness of the states. While the latter does contribute to the entropy the former does not. So, there is a conceptual difference between the two classes of randomness. Only classic-statistical randomness it attributed to entropy. The reason is: the quantum randomness can be reduced to zero by virtue of choosing a measurement operator, where the quantum state is an eigenstate. The selection of the (virtual) measurement operator, however, should not contribute to the entropy of a system.

Quantum states have a fixed entropy when not measured: A pure state does not have entropy. Any quantum operation that does not affect the purity of a state thus does not increase entropy. We shall later see that unitary operator leaves the purity of a quantum state unaffected and that all non-measurement operations on a quantum system belong to such unitary operators. In other words: unless you measure a quantum system, its entropy does not increase by its intrinsic evolution.

⁶ My colleagues tell me this example shows more than anything else, that I am German. Alas.

⁷ In the following the entropy is commonly defined in terms of the base-2 logarithm, so that a maximally mixed state of a two-level system corresponds to one bit of entropy.

A measurement operation *can* induce entropy and is thus irreversible if the measurement outcome was not yet known: As an example take a $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state, which is measured with the $S_x = |0\rangle\langle 0| - |1\rangle\langle 1|$ operator (we shall later see, that this e.g. corresponds to a diagonally polarized photon measured with an HV-polarization beamsplitter). The result is a $|0\rangle$ or a $|1\rangle$ state, each with 50% probability, thus a mixed state with $\hat{\rho} = 1/2(|0\rangle\langle 0| + |1\rangle\langle 1|)$ and an entropy of $S(\hat{\rho}) = \text{Tr}(\hat{\rho} \log \hat{\rho}) = \log(2) = 1$ (e.g. this is a maximum entropy state). As we have increased the entropy we have made an irreversible operation.

A measurement operation *does not need* to induce entropy and may thus be reversible if the measurement outcome was known to begin with: If the measurement had been in parallel with the state then we would have gotten one answer with certainty and retained a pure state. This operation is thus NOT irreversible.

The last two statements have profound impact on our ability to clone a quantum state. You can imagine that doing irreversible things to systems you'd like to clone is a big no-no.

If the measurement apparatus extracts information from the quantum system. It must thus increase the quantum system's entropy: If the entire system (measurement apparatus plus quantum system) is adiabatic⁸ then the overall entropy of the system cannot have changed by the measurement and the measurement must have reduced the entropy of the measurement apparatus. In other words: The measurement has transferred a certain degree of order from the quantum system to the measurement apparatus (its quantum information being measured leaves the measurement apparatus in a more well-defined state as before; e.g. it shows a specific reading and not just noise) but the apparatus must likewise transfer disorder to the quantum system. In this respect the measurement process in quantum physics may be a bit less mysterious: it's "simply" the random dephasing of a highly ordered state, when it gets in contact with a thermal bath of a large apparatus.

2.5 Time evolution of quantum systems

Let us consider an isolated quantum system that is prepared in a state $|\Psi(t=0)\rangle$. Depending on the dynamics of the system, the state of the system at some other time $t = T$ is determined by some operator acting on the initial state

$$|\Psi(T)\rangle = \hat{U}(T)|\Psi(0)\rangle \quad (104)$$

we call $\hat{U}(t)$ the time-evolution operator. Typically, we would derive this from the equation of motion of the system, however the specifics are not of interest at the moment. We are merely interested in the key features of the time evolution operator, as imposed by the laws of quantum mechanics. First, the operator should be linear. This does not come as a surprise, since aside from the measurement process, all quantum mechanics is linear. Second, we want the information content of the system to remain unchanged. We saw, in the previous section, that a pure quantum state has zero entropy and we expect a completely isolated quantum system to maintain its entropy irrespective of its inner workings. This immediately leads us to the following requirement: relationships between quantum states should be conserved. In particular, if two quantum states are unambiguously distinguishable by measurement at time $t=0$, then this distinguishability should be maintained. In other words, the inner product of vectors should be conserved:

$$\langle \Phi(T) | \Psi(T) \rangle = \langle \Phi(0) | \hat{U}^\dagger(T) \hat{U}(T) | \Psi(0) \rangle = \langle \Phi(0) | \Psi(0) \rangle \quad (105)$$

⁸ This can always be enforced if the apparatus is big enough or if you only consider sufficiently short times.

This will only be the case if the operator satisfies

$$\hat{U}^\dagger \hat{U} = 1 \quad (106)$$

We call operators that fulfil this condition *unitary*; Time-evolution, or more generally any transformation of an isolated quantum state is given by *unitary transformations*. This has several implications, as we shall see in the following.

As a final remark, we note that the quantities we access in an experiment are probabilities of certain measurement outcomes, or more generally, time-dependent observables:

$$\hat{A}(t) = \langle \psi | \hat{U}^\dagger(t) \hat{A} \hat{U}(t) | \psi \rangle \quad (107)$$

We can describe the time evolution of a quantum system in several ways: We can attribute the time-dependence entirely to the operators (this is called the Heisenberg picture), or we can keep the operators constant, and attribute the time-dependence to the quantum state (Schrödinger picture), or we do something in between, and attribute part of the time evolution to the state and the operators (for example in the Interaction picture).

Specifically, in the *Schrödinger picture* the states evolve in time under the influence of the time evolution operator $|\Psi(t)\rangle = \hat{U}(t)|\Psi(0)\rangle$, whereas in the *Heisenberg picture* the time-dependence is purely in the operators $\hat{O}(t) = \hat{U}^\dagger(t)\hat{O}\hat{U}(t)$. We will be using these pictures interchangeably- but more on this later.

In quantum optics (or quantum physics in general) the free (time-)evolution of any system is determined by its Hamiltonian⁹ \hat{H} . Using a few theorems and a bit nibbling around you indeed find that the system's Hamiltonian \hat{H} and the systems evolution operator $\hat{U}(t)$ according to:

$$\hat{U}(t) = e^{\frac{i}{\hbar}t\hat{H}} \quad (108)$$

The general terminology here is that the Hamiltonian \hat{H} generates the evolution operator \hat{U} ; it is thus called the "generator". As the Hamiltonian is the measurable for the system's energy it comes as no surprise to you that it is a Hermitian operator. The same way that an exponential of numbers (times i) relates purely real numbers to pure phases it relates Hermitian operators with Unitary operators. Thus, by construction we have guaranteed that \hat{U} is unitary, that its eigenvalues are phases and that its eigenmodes for a complete basis set.

While we have been discussing time-evolution here, the same kind of argumentation here applies to any kind of reversible (lossless¹⁰) and linear operation you can carry out on a quantum system. In the reality of quantum optics such an operation may be represented by any type of lossless optical system; e.g. phase shifters, beam splitters, polarization optics, holograms, fibers, gratings....you name them. They can all be represented by a unitary evolution operator \hat{U} , which connects the quantum system before $|\psi_0\rangle$ and after $|\psi_1\rangle$ the interaction

$$|\psi_1\rangle = \hat{U}|\psi_0\rangle \quad (109)$$

Moreover, each of these evolution operators \hat{U} are generated by an interaction Hamiltonian:

⁹ If you come from a classical world, then note that the Hamiltonian is basically a modified version of the evolution equation (e.g. Maxwell equations or the Schrödinger equation) with exactly the same information content

¹⁰ If your loss is not so bad as to make the modal structure of your system meaningless you can include loss here, too.

$$\hat{U} = e^{i\hat{H}} \quad (110)$$

Note that as opposed to the time-evolution Hamiltonian there is no time here and we have absorbed the \hbar into the definition of the interaction Hamiltonian for the sake of brevity.

In the following chapter we shall introduce the qubit as the simplest possible quantum system and then discuss specific optical elements and how their properties are related to their interaction Hamiltonians, their evolution operators and what that all means physically.

3 Photonic Qubits

As we now have a fundamental understanding of how the world works on a quantum level, we shall dive deeper into the realm of quantum information. We do so by dumbing down all the concepts from the last chapter until nothing is left but the most simple quantum system, that you can still rightfully call a quantum system. A quantum system which is composed of two modes and only two modes: the Qubit.



3.1 The Qubit

In the classical case we can encode information in any physical system that has at least two clearly distinguishable states – a bit. Such states may be a low or high voltage; a light being turned on or off or an apple having a bite taken out of it or not. In any case we can give these two specific states logical representations and call them:

$$|0\rangle, |1\rangle \quad (111)$$

Note that the formal similarity to quantum states is at this case purposefully selected but not yet obvious. Let's however call these the *computational basis states* (CBS). If these states are the basis states of an arbitrary quantum system, we have the possibility of introducing general superposition states, a *qubit state*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (112)$$

which is something, that one, of course, cannot do with a classical bit.

How can we physically realize Qubits in quantum photonics, which – as we recall – has the concept of modes and photons? The first option is to encode the qubit in the photon state of a single fixed mode with index i , which we shall call \hat{a}_i . This is known as the *single-rail qubit* representation. One possible implementation would be to differentiate between the single photon Fock-State and the two Photon Fock-State⁷ of the field in this particular mode:

$$\begin{aligned} |0\rangle &\equiv |n_i = 1\rangle = \hat{a}_i^\dagger |\text{vac}\rangle \\ |1\rangle &\equiv |n_i = 2\rangle = \hat{a}_i^\dagger \hat{a}_i^\dagger |\text{vac}\rangle \end{aligned} \quad (113)$$

Note that we have changed the notation of the number-States somewhat (they are now called $|n_i = 1\rangle$), to differentiate between them (and the vacuum-state) and the CBS. That is, the computational basis state $|1\rangle$ corresponds to a state of the field with a two photons in mode \hat{a}_i and the state $|0\rangle$ corresponding to a state with one photon. Keep in mind the specific numbers are chosen arbitrary.

All notes subject to change, no guarantee to correctness, corrections welcome.

Some problems with single-rail qubit encoding is that photon loss will affect the qubit state in the sense of that it changes its value. Moreover, it requires a handle on detectors and even more so on optical elements and sources that create and/or mix different Fock-states at will. This is indeed difficult, and the loss issue is rather unpractical if we want to transmit the state over a long distance. Moreover, if you want to implement operations which work differently, depending on the state of the qubit you'll have to resort to nonlinear optics and that's generally a bugger.

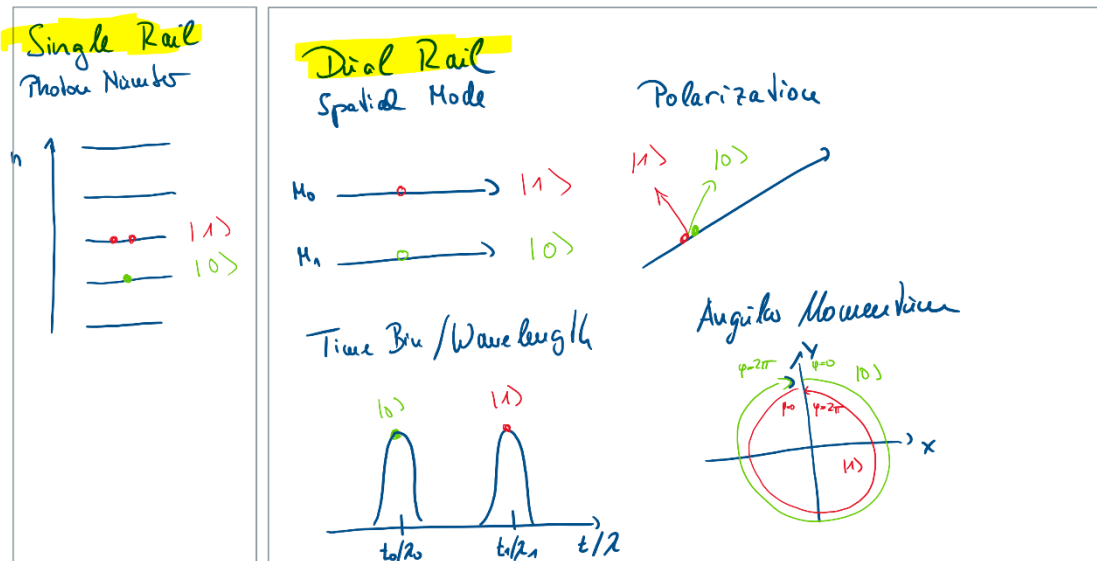


Fig. 5: Some important classes of photonic Qubits.

The second, and more practical way, is to fix the number-state and use a pair of orthogonal field modes M_i and M_j to encode the qubit, which may be orthogonal polarization modes, plane waves of different orientation, Gauss-Laguerre modes of different order or azimuthal phase, different beam paths, different modes of a single waveguide or modes of different waveguides, or different wavelengths or different time-bins or anything that you can imagine¹¹. This is called the *dual-rail qubit* representation. The Fock-state is typically fixed to a single photon state – everything else is complicated enough already:

$$\begin{aligned} |0\rangle &\equiv |n_i = 1, n_j = 0\rangle = \hat{a}_i^\dagger |\text{vac}\rangle \\ |1\rangle &\equiv |n_i = 0, n_j = 1\rangle = \hat{a}_j^\dagger |\text{vac}\rangle \end{aligned} \quad (114)$$

To make things less abstract, let's take these modes to be orthogonal polarization modes. Two particularly popular polarization modes are the linear horizontal $|H\rangle$ and linear vertical $|V\rangle$ polarization (typically in reference to an optical table or a polarizing beam splitter):

$$\begin{aligned} |0\rangle &\equiv |H\rangle = \hat{a}_H^\dagger |\text{vac}\rangle \\ |1\rangle &\equiv |V\rangle = \hat{a}_V^\dagger |\text{vac}\rangle \end{aligned} \quad (115)$$

But again, we will only use that to exemplify the physical meaning, of what we discuss here, and you can take any kind of qubit and apply the discussion to this qubit.

¹¹ In fact, we need not limit ourselves to two basis vectors but could take more. These states are then called qu-dit states.

3.2 The Bloch Sphere

The first thing we do is a bit of bookkeeping. We have introduced the expansion coefficient α and β which both are, of course complex numbers. However, this is in fact a bit overly complex and we can describe the entire state space with only two real numbers, which represent the latitude Θ (counted from the north pole to the south pole) and longitude ϕ of an imaginary sphere, according to:

$$|\psi\rangle = \alpha|H\rangle + \beta|V\rangle = \cos\frac{\Theta}{2}|H\rangle + e^{i\phi}\sin\frac{\Theta}{2}|V\rangle \quad (116)$$

Where we have used the fact that $\alpha^2 + \beta^2 = 1$ as a justification to introduce the polar angle Θ and the azimuthal angle ϕ (longitude) and have also utilized the fact that a cumulative phase is irrelevant. It thus becomes clear that the state of any polarization qubit and therefore ANY qubit state altogether can be represented as a point on the surface of the Sphere; the infamous Bloch Sphere according to the equation:

$$\begin{aligned} x &= r \sin\Theta \cos\phi \\ y &= r \sin\Theta \sin\phi \\ z &= r \cos\Theta \end{aligned} \quad (117)$$

Where $r = 1$ (we'll see later, that $r \neq 1$) also has a physical meaning.

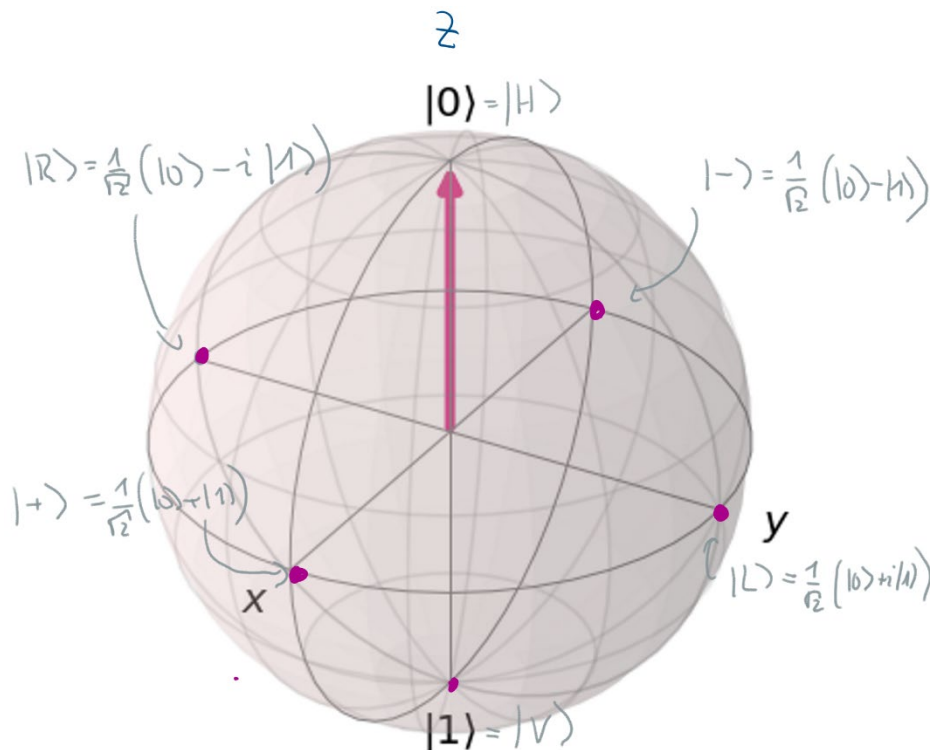


Fig. 6: Representation of a qubit on the Poincarè sphere. The Bloch-Sphere itself is created with IBM's QISKIT library.

On the Bloch sphere the state $|0\rangle = |H\rangle$ is represented by the north pole and $|1\rangle = |V\rangle$ is represented by the south pole, e.g. the CBS are exclusively along the z-axis of the Bloch sphere. The other axes have a profound meaning, too: The points on the x-axis, e.g. those on the equator facing the viewer or point straight away also belong to linear polarization, namely to the diagonal basis vectors $|+\rangle$ and $|-\rangle$, which can be constructed using the Hadamard operator \hat{H} :

All notes subject to change, no guarantee to correctness, corrections welcome.

$$\begin{bmatrix} |+\rangle \\ |-\rangle \end{bmatrix} = \hat{H} \begin{bmatrix} |0\rangle \\ |1\rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} |0\rangle \\ |1\rangle \end{bmatrix} \quad (118)$$

Here the Hadamard operator is given in its matrix representation (with the CBS as an expansion basis) as:

$$\hat{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (119)$$

In other words:

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \quad (120)$$

We will later see that find that such an action can be connected with the interaction Hamiltonian of a linear optical element. Of course, we know, that in this case of a polarization photon the element is a Half-Wave-Plate with its fast axis rotated 22.5 degrees with respect to the horizontal. In fact, we can connect any set of linearly polarized states to any other using a half-wave plate with the appropriate angle setting. For spatial modes the appropriate optical element is a 50/50-Beamsplitter.

Another set of special points on the Bloch sphere are those, where the sphere intersects the y-axis. This is where the left-handed and right-handed circular basis states $|L\rangle$ and $|R\rangle$ (sometimes also called $|\oslash\rangle$ and $|\otimes\rangle$) are located. They can also be constructed from $|H\rangle$ and $|V\rangle$ according to:

$$\begin{bmatrix} |L\rangle \\ |R\rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} |0\rangle \\ |1\rangle \end{bmatrix} = \hat{S}\hat{H} \begin{bmatrix} |0\rangle \\ |1\rangle \end{bmatrix} \quad (121)$$

In other words:

$$|L\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \quad |R\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \quad (122)$$

The procedure makes use of the now well-established Hadamard Gate \hat{H} and the phase gate \hat{S} :

$$\hat{S} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \quad (123)$$

Again, we can attribute this kind of transformation to an interaction Hamiltonian and a real kind of physical beam-splitter. In case of a polarization qubit \hat{S} is represented by a quarter wave plate with appropriate setting and in case of two spatial modes this is a mode-selective phase shifter.

To generalize this somewhat: a half-wave plate rotates the states, changes the longitude Θ and leaves the latitude ϕ fixed, i.e. it rotates a state perpendicular to the equator, whereas a quarter-waveplate changes the latitude ϕ and leaves the longitude Θ fixed, i.e. it rotates a state clockwise or anticlockwise parallel to the equator. Using two of these elements you can connect any two points of the sphere and change any polarization state into any other. That's also why you cannot buy 1/3-wave-plates or 0.93-wave plates; you can simply construct them from a half- and a quarter-wave plate.

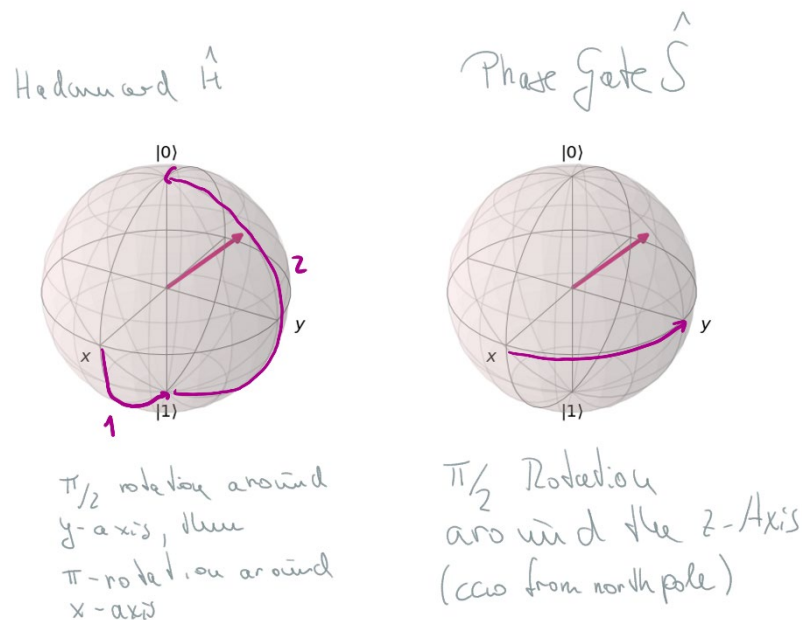


Figure 1: Actions of the \hat{H} -gates (Half-Wave-Plate) on a QuBit and \hat{S} -gate (Quarter Wave Plate) on a QuBit, interpreted as rotations on the Bloch-sphere.

An important thing to note about the three basis choices is that they are not orthogonal, in fact, they cannot be, because each of them is complete. One gets $|\langle H|U\rangle|^2 = |\langle H|D\rangle|^2 = |\langle H|R\rangle|^2 = |\langle H|L\rangle|^2 = \frac{1}{2} \neq 0$. When we measure a photon encoded as $|D\rangle$ in the basis (H/V) we are equally likely to get one result or the other, the (H/V) and (U/D) basis are therefore called *mutually unbiased*. The same applies to the (H/V) and (L/R) basis as well as the (U/D) and (L/R) basis. This is represented by the fact that these three sets span the state space's coordinate system as mutually orthogonal axes. I am leaving out the exact proof here, but from this you can imagine that the three types of bases introduced here are in and by themselves complete in the sense that you cannot find another basis set, which is mutually unbiased to the others.

From a standpoint of quantum information processing and of the three basis or in fact any other basis may be used to encode $|0\rangle$ and $|1\rangle$ states and we shall later see that some protocols actually only work if there is a certain degree of ambiguity as to this question.

3.3 Observables and the Pauli-Matrices

Now that we can describe the state of QuBits and understand the way that we can manipulate them, we must expand on the understanding of their measurement. For the sake of simplicity we shall identify the basis vectors as the eigenstates of the respective projection operators and construct measurement operators from the individual projectors, with measurement values 1, for the first basis vector and measurement value -1 for the second basis vector. The construction is particularly simple for the computation basis set (CBS) $|0\rangle$ and $|1\rangle$ (or $|H\rangle$ and $|V\rangle$):

$$\hat{\sigma}_3 = \hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (124)$$

Where the matrix representation is done in the CBS. The operator is termed the Pauli-z or third Pauli operator, and the alphabetic naming takes its name from the corresponding axis of the Bloch sphere.

Of course, we can construct similar measurement operators from the other two sets of basis vectors, namely:

All notes subject to change, no guarantee to correctness, corrections welcome.

$$\begin{aligned}
 \hat{\sigma}_1 = \hat{\sigma}_x &= |+\rangle\langle+| - |-\rangle\langle-| \\
 &= \frac{1}{2}[(|0\rangle + |1\rangle)(\langle 1| + \langle 0|) - (|0\rangle - |1\rangle)(-\langle 1| + \langle 0|)] \\
 &= \frac{1}{2}[|0\rangle\langle 1| + |0\rangle\langle 0| + |1\rangle\langle 1| + |1\rangle\langle 0| - |0\rangle\langle 0| - |1\rangle\langle 1| + |1\rangle\langle 0|] \\
 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
 \hat{\sigma}_2 = \hat{\sigma}_y &= |R\rangle\langle R| - |L\rangle\langle L| \\
 &= \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}
 \end{aligned} \tag{125}$$

Frequently there is a fourth Pauli-Operator $\hat{\sigma}_0 = \hat{1}$ introduced, which is the unit matrix. All four of these are obviously Hermitian, because they constitute measurements, e.g.:

$$\hat{\sigma}_i = \hat{\sigma}_i^\dagger \tag{126}$$

We also note that:

$$\hat{\sigma}_i \hat{\sigma}_j = \delta_{ij} \mathbb{I} + i \epsilon_{ijk} \hat{\sigma}_k \tag{127}$$

where ϵ_{ijk} is the Levi-Civita-Symbol or antisymmetric epsilon tensor.

Any linear operator \hat{M} on the qubit state space (e.g. any operator that acts on a two-dimensional Hilbert space and whose result still is in that space) can be constructed from a superposition of the Pauli-Operators:

$$\hat{A} = \sum_{i=0\dots 3} a_i \hat{\sigma}_i \tag{128}$$

If the expansion coefficients are real, then the resulting operator \hat{A} is Hermitian, i.e. it belongs to a measurement. In other words: any quantum measurement you can make on a qubit is a superposition of the Pauli measurement operators, or, from an optics point of view a polarization measurement. Conversely, if you can make Pauli-Measurements, you can make any possible measurement.

There is, in fact, an even stronger statement to this claim. The Pauli matrices $\hat{\sigma}_x$ and $\hat{\sigma}_y$ can be constructed from $\hat{\sigma}_z$ and the Hadamard operator \hat{H} (a 22.5° half wave plate) and a Hadamard plus Phase Shifter $\hat{S}\hat{H}$ (HWP plus QWP), according to the relations:

$$\hat{\sigma}_x = \hat{H} \hat{\sigma}_z \hat{H} \hat{\sigma}_y = \hat{S} \hat{H} \hat{\sigma}_z \hat{H} \hat{S}^\dagger \tag{129}$$

This means that you can make all three basis measurements with the help of a $\hat{\sigma}_z$ -measurement and a set of waveplates (beamsplitters). Experimentally this appears as somewhat of a no-brainer:

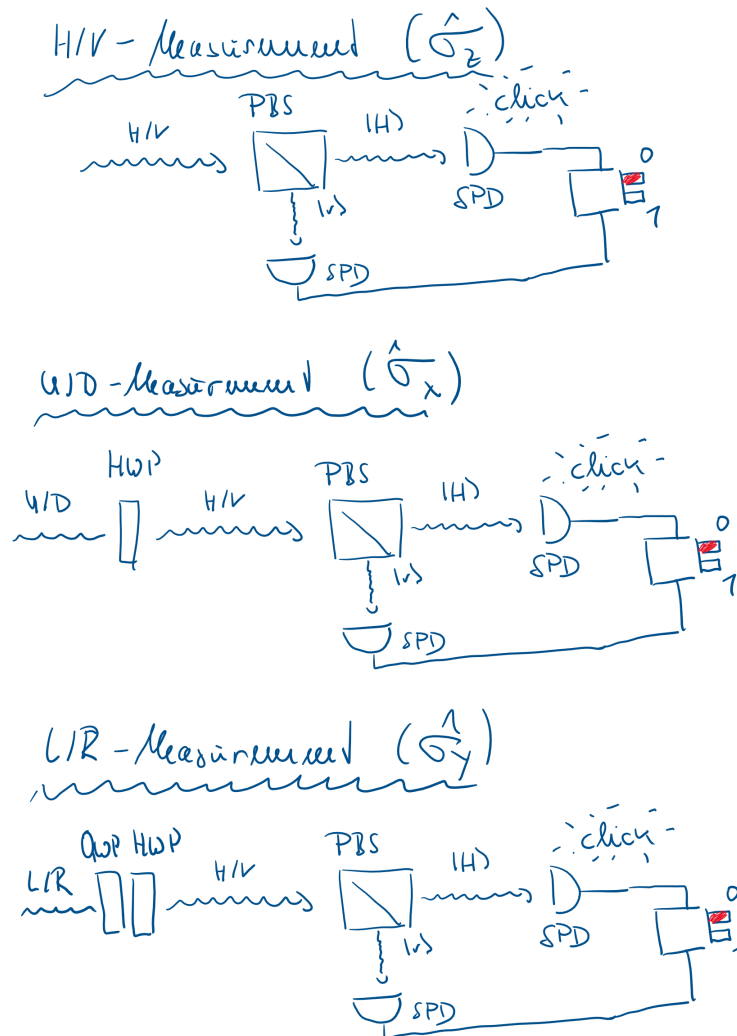


Figure 2: Experimental implementation of all three Basis measurements for Polarization Qubits using projection elements and a σ_z measurement.

The three types of basis state sets are mutually unbiased. You can see this relation by looking at the commutation relation of their observables $\hat{\sigma}_{1,2,3}$, for which the relation

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k \quad (130)$$

Hold. Again here ϵ_{ijk} is the Levi-Civita-Symbol or antisymmetric epsilon tensor. Thus the Pauli operators are mutually *complementary*, in the sense of that complete knowledge about the result of a measurement of the first means that we have absolutely no knowledge of the measurement outcome of the second.

In other words: if you decide to measure your Qubit $|\psi\rangle$ in the computational basis (we apply the $\hat{\sigma}_z$ operator) then there is absolutely no information left of the qubit, which you could measure in the other bases $|+/-\rangle$ or $|L/R\rangle$. Or, to put it in an even more blunt language: although the state of a qubit is characterized by two real numbers (e.g., the latitude and longitude on the Bloch-Sphere) you can only ever hope to extract a single bit of information from them. This is a profound finding, which cannot be stressed enough, because it limits the amount of classical information, which can be extracted from a quantum system to a point, were you can – upon a single measurement – never be quite certain as to the real state of the Qubit before the measurement. This potentially unintended loss of quantum information in the classical measurement process is the resource that quantum cryptography schemes

draw their strength from, it is also a source of challenges, if instead, complex quantum information is supposed to be transmitted and must be protected against destruction.

3.4 Single Photon Operations, Gates, Elements and Hamiltonians

In the last section we have used some gut feeling to map both unitary operators, such as \hat{H} and \hat{S} as well as Hermitian operators such as the Pauli Matrices onto optical elements. By doing so we have basically used our understanding of how polarization elements work on classical Jones vectors and then we have hoped and prayed that this description also holds in the realm of single photons, which we are discussing in the context of single rail photons.

Any optical element which works differently on two different classical modes M_0 and M_1 affects the entire quantum state of these modes \hat{a}_0 and \hat{a}_1 and not just the photon number states $|n_0 = 1\rangle$ and $|n_1 = 1\rangle$ that we have chosen to use for our dual-rail representation of the qubit. We therefore have to invoke the apparatus introduced in chapter 2.5 on time evolution to see how the optical element in question works on the quantum state. We shall then see that a large class of linear operators actually does not mix the number states, meaning that its action is solely and completely confined to the Hilbert-Subspace of the Qubit and that we can simply represent it by a 2x2-Jones-Type-matrix there, which, as discussed above, can be created by a complex superposition of the Pauli operators.

3.4.1 The Phase Shifter / Phase Gate

The first optical element is the antisymmetric phase-shifter. E.g. it adds a positive phase to the M_0 mode and a negative phase to the M_1 mode. Its generation Hamiltonian is expressed in terms of the mode operators, e.g. how they act of the field itself in the two modes that we have chosen for our two-rail representation of the qubit. Its Hamiltonian is (we shall see below that this actually matches our expectation, of what a phase shifter does in the context of a Qubit):

$$\mathcal{H}_1(\phi) = \hbar\phi(\hat{a}_0^\dagger\hat{a}_0 - \hat{a}_1^\dagger\hat{a}_1) \quad (131)$$

There is a deeper physical meaning in this formula. In quantum optics it is introduced that the \hat{a}_i^\dagger and \hat{a}_i operators correspond to the creation and destruction of a photon respectively. So, what the phase shifter does is: it destroys a photon in mode i and instantly recreates it, all the while adding (subtracting) an action of $\hbar\phi$ onto this photon, corresponding to the appropriate phase shift.

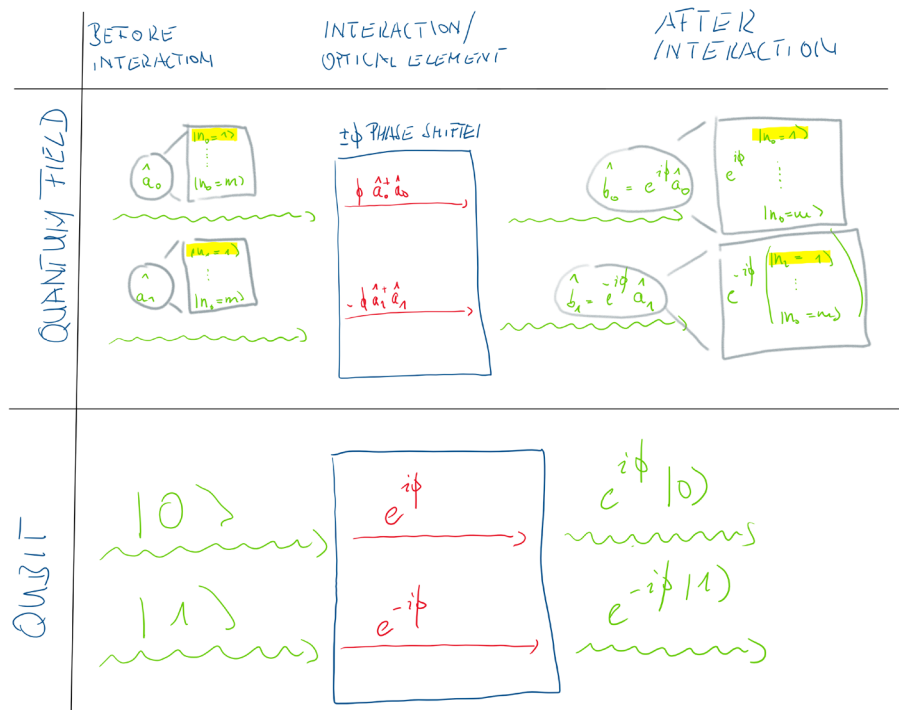


Fig. 7: Conceptual drawing of an antisymmetric phase shifter. Top row: the action of the antisymmetric phase shifter on the quantum fields of the two modes M_0 and M_1 . Displayed is the state of the quantum field before the interaction, the Feynman-Diagram of the interaction itself and the resulting state of the quantum field. The interaction is equal for all number states (i.e. there are m identical Feynman diagrams for every number state!). The number states used for our Qubit are marked in yellow. Bottom row: The action of the an antisymmetric phase shifter as experienced by the Qubit alone.

The Hamiltonian $\mathcal{H}_1(\phi)$ can of course be used to generate the evolution operator $\hat{U}_1(\phi)$:

$$\hat{U}_1(\phi) = \exp\left\{\frac{i}{\hbar} \mathcal{H}_1(\phi)\right\} = \exp\{i\phi(\hat{a}_0^\dagger \hat{a}_0 - \hat{a}_1^\dagger \hat{a}_1)\} \quad (132)$$

$\hat{U}_1(\phi)$ connects the quantum states of the field $\hat{a}_{0,1}$ in the modes corresponding to the qubit states $|0\rangle$ and $|1\rangle$ BEFORE the phase shifter with the quantum states $\hat{b}_{0,1}$ in the modes AFTER the phase shifter. We can simply calculate its effect on the quantum state of the field according to:

$$\begin{aligned} \hat{U}_1(\phi) \hat{a}_j \hat{U}_1^\dagger(\phi) &= \exp\left\{-\frac{i}{\hbar} \mathcal{H}_1(\phi)\right\} \hat{a}_j \exp\left\{\frac{i}{\hbar} \mathcal{H}_1(\phi)\right\} \\ &= \hat{a}_j + \left(-\frac{i}{\hbar}\right) [\mathcal{H}_1(\phi), \hat{a}_j] + \frac{1}{2} \left(-\frac{i}{\hbar}\right)^2 [\mathcal{H}_1(\phi), [\mathcal{H}_1(\phi), \hat{a}_j]] + \dots \\ &= \begin{cases} \hat{a}_j + (-i\phi) \hat{a}_j + \frac{1}{2} (-i\phi)^2 \hat{a}_j + \dots & \Leftrightarrow j = 0 \\ \hat{a}_j + (i\phi) \hat{a}_j + \frac{1}{2} (i\phi)^2 \hat{a}_j + \dots & \Leftrightarrow j = 1 \end{cases} \\ &= \begin{cases} \hat{a}_j \exp(-i\phi) & j = 0 \\ \hat{a}_j \exp(i\phi) & j = 1 \end{cases} \\ &\equiv \hat{b}_j \end{aligned} \quad (133)$$

Note that we have used a few fancy equations from quantum optics (e.g. the commutation relations for $[\mathcal{H}_1(\phi), \hat{a}_j]$) and the so-called Baker-Cambell-Hausdorff-theorem to get from one line to the next.

As this relation is purely linear in the operators, we can simply express it as a matrix:

$$\begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \end{bmatrix} = \begin{pmatrix} \exp(-i\phi) & 0 \\ 0 & \exp(i\phi) \end{pmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} \quad (134)$$

The states before and after the optical element are related in a linear manner (e.g. with numbers only in the matrix). This means that we don't mess around with the number state composition of the quantum states of both modes (except for that specific phase term). (e.g. any $|n_0 = 1\rangle$ state is transformed according to $|n_0 = 1\rangle \rightarrow \exp(-i\phi)|n_0 = 1\rangle$ but there is no intermixing with any other number state (e.g. the $|n_0 = 2\rangle$ state). Or in other words: the so-defined phase shifter operation does ONLY and ONLY act on our $|0\rangle$ and $|1\rangle$ states and we quite conveniently read off:

$$|\psi'\rangle = \hat{U}_1(\phi)|\psi\rangle = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} |\psi\rangle \quad (135)$$

Where $|\psi\rangle$ is of course the state of our qubit. As we had discussed above there is necessarily a relation with the Pauli operators, and in this case it's simply:

$$\hat{U}_1(\phi) = \exp(-i\phi\hat{\sigma}_1) = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} \quad (136)$$

Or in other words: the Pauli-Operator $\hat{\sigma}_1$ generates the antisymmetric phase shift gate which can be implemented by an antisymmetric beam splitter.

Two special cases of the phase shifter are the phase gate \hat{S} and the $\pi/8$ -gate \hat{T} , which are defined as:

$$\hat{S} = \exp\left(\frac{i\pi}{4}\right) \begin{bmatrix} e^{-i\frac{\pi}{4}} & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \quad (137)$$

and

$$\hat{T} = \exp\left(\frac{i\pi}{8}\right) \begin{bmatrix} e^{-i\frac{\pi}{8}} & 0 \\ 0 & e^{i\frac{\pi}{8}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix} \quad (138)$$

Obviously, we have $\hat{T}^2 = \hat{S}$,

Note that any type of phase shift on the Bloch sphere is a mere rotation about the polar axis. E.g. it does not change a $|0\rangle$ or a $|1\rangle$ state except for a global phase shift, which is of no consequence.

3.4.2 Beam Splitter

Now we may proceed to a more general optical element, the beam splitter. We can model the beam splitter in the exact same way. We guess what the Hamilton operator should look like, then derive the evolution operator, invoke the time evolution mechanism on that operator, see if it mixed the number states and then understand what that does to the qubit.

The beam splitter obviously has the effect of mixing the modes and we can create an interaction Hamiltonian by simply mixing them, thus

$$\hat{\mathcal{H}}(\theta, \varphi) = \hbar\theta \exp(i\varphi) \hat{a}_0^\dagger \hat{a}_1 + \hbar\theta \exp(-i\varphi) \hat{a}_1^\dagger \hat{a}_0 \quad (139)$$

You may wonder, why we chose this specific form. In fact, this is the most general bi-linear two-mode mixing Hermitian operator possible. I.e. we can only make the Hamiltonian Hermitian, if the creation and annihilation operators appear in pairs and if they also appear in sum with their Hermitian conjugates. In a more physical interpretation. The first term of this operator destroys photons in the $|1\rangle$ mode and creates them in the $|0\rangle$ mode with a probability, which is related to θ and a phase-shift

which is related to φ . The second term does exactly the same, but while destroying a photon in the $|1\rangle$ mode and creating one in the $|0\rangle$ mode.

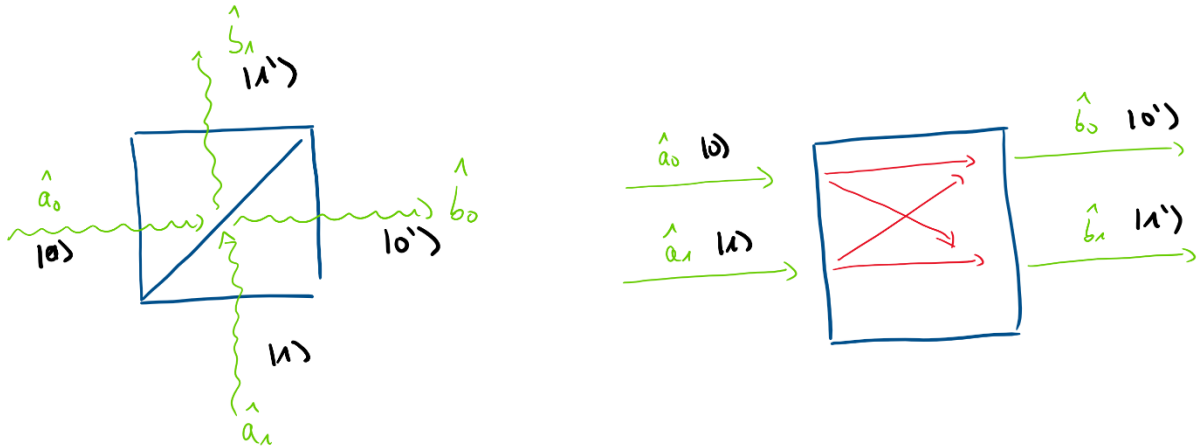


Fig. 8: A beam splitter together with its Feynman-Representation.

Applying the same steps as before we can then also find the appropriate transformation matrix, which connects the state of the modes:

$$\begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \end{bmatrix} = \begin{pmatrix} \cos \theta & -ie^{i\varphi} \sin \theta \\ -ie^{-i\varphi} \sin \theta & \cos \theta \end{pmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} \quad (140)$$

Again, same thing as before. The operation of the beam splitter does not mix up the different number states, e.g. its operation on the $n = 1$ qubit can be described with the same matrix, e.g.

$$|\psi'\rangle = \begin{bmatrix} \cos \theta & -ie^{i\varphi} \sin \theta \\ -ie^{-i\varphi} \sin \theta & \cos \theta \end{bmatrix} |\psi\rangle \quad (141)$$

This matrix can also be created from the superposition of Pauli operators, but we shall not go into details here.

We also combine the (symmetric) phase shifter with the beam splitter and obtain the most general case for a linear optical element (in the strict sense) with the Hamiltonian

$$\hat{\mathcal{H}}(\theta, \varphi) = \frac{\hbar\phi}{2} \hat{a}_1^\dagger \hat{a}_1 - \frac{\hbar\phi}{2} \hat{a}_2^\dagger \hat{a}_2 + \hbar\theta \exp(i\varphi) \hat{a}_1^\dagger \hat{a}_2 + \hbar\theta \exp(-i\varphi) \hat{a}_2^\dagger \hat{a}_1 \quad (142)$$

And the transformation matrix

$$\begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \end{bmatrix} = \begin{pmatrix} e^{\frac{i\phi}{2}} \cos \theta & -ie^{i\varphi} \sin \theta \\ -ie^{-i\varphi} \sin \theta & e^{-\frac{i\phi}{2}} \cos \theta \end{pmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix}. \quad (143)$$

Obviously, this also leave the number of photons unaffected and we may use the matrix directly on the qubit state:

$$|\psi'\rangle = \begin{pmatrix} e^{\frac{i\phi}{2}} \cos \theta & -ie^{i\varphi} \sin \theta \\ -ie^{-i\varphi} \sin \theta & e^{-\frac{i\phi}{2}} \cos \theta \end{pmatrix} |\psi\rangle \quad (144)$$

Which is the most general representation of any linear two mode mixing processes, represented by the most general unitarian 2x2 matrix possible. This means that any lossless, photon-number-conserving, linear mode transformation between pairs of modes can be written as such a matrix. This also

means that a combination of beam-splitters (waveplates) and a phase shifter can create ANY possible lossless, linear, photon-number-conserving interaction between two modes there is. This also means that ANY combination of optical elements, which mixes two modes can be replaced by a beam-splitter and a phase-shifter.

3.4.3 Mach-Zehnder-Interferometer

The action of the multiple interaction Hamiltonians can of course be "stacked" to form complex optical elements. Here is a look at a balanced Mach-Zehnder interferometer, i.e. a stacked 50/50- beam splitter and single mode phase shifter and a 50/50-beam splitter.

$$\begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \end{bmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} \quad (145)$$

As expected, we see that the total number of photons is preserved but there is a shift of contrast (we will later see, that the $\hat{b}_0^\dagger \hat{b}_0 - \hat{b}_1^\dagger \hat{b}_1$ -terms are indeed related to interference, which is enacted by the phase shift ϕ).

$$\begin{aligned} \text{Photon number conservation: } \hat{b}_0^\dagger \hat{b}_0 + \hat{b}_1^\dagger \hat{b}_1 &= \hat{a}_0^\dagger \hat{a}_0 + \hat{a}_1^\dagger \hat{a}_1 \\ \text{Interference: } \hat{b}_0^\dagger \hat{b}_0 - \hat{b}_1^\dagger \hat{b}_1 &= \cos \phi (\hat{a}_0^\dagger \hat{a}_0 + \hat{a}_1^\dagger \hat{a}_1) - i \sin \phi (\hat{a}_0^\dagger \hat{a}_1 + \hat{a}_1^\dagger \hat{a}_0) \end{aligned} \quad (146)$$

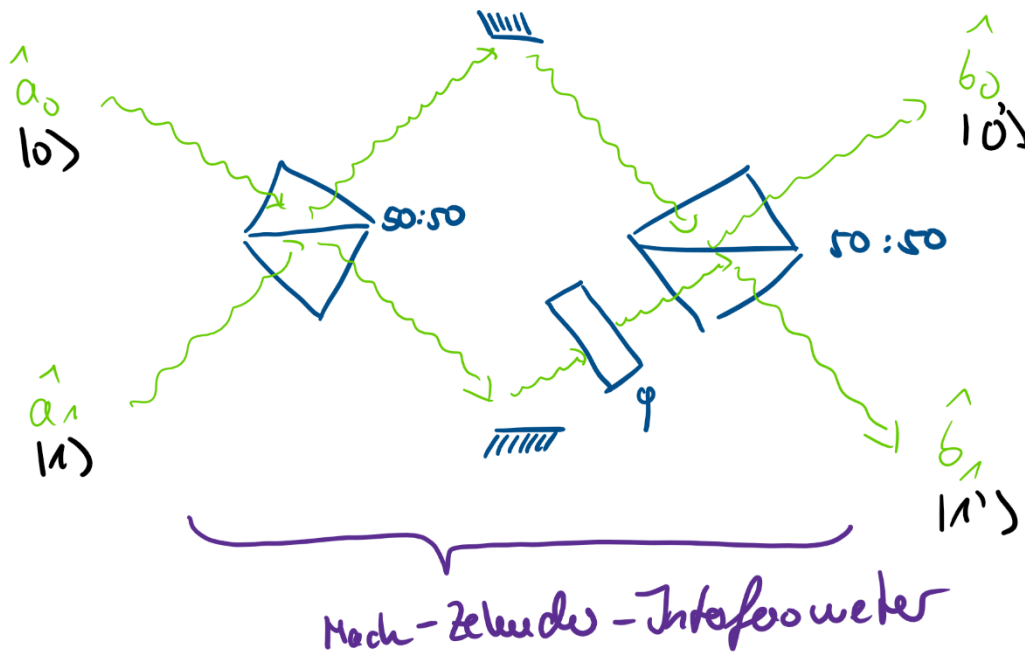


Fig. 9: A Mach-Zehnder Interferometer..

3.4.4 Squeezing Operators

This is, however, by no means a complete discussion of all possible interaction Hamiltonian and processes, which may act on a single photon. Note from above that we have also said "linear" optical element. This means "an optical element, which relies exclusively on two-wave mixing, e.g. linear optics". This is represented by $\hat{a}_1^\dagger \hat{a}_2$ pairs, where photon annihilation and photon creation are balanced.

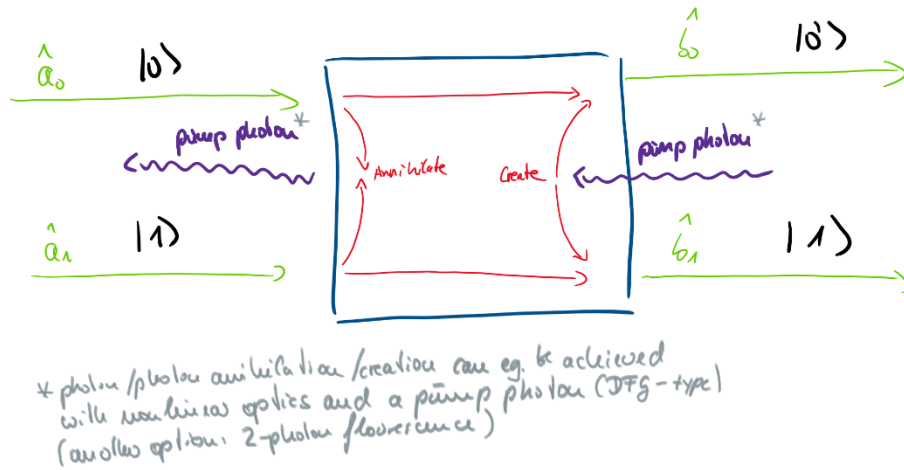


Fig. 10: Feynman representation of a mode squeezing operator. The green interaction paths are most commonly realized by a three photon interaction, where the resulting photon is discarded.

You can, however, retain mathematically linear operators by introducing elements such as $\hat{a}_1 \hat{a}_2$ into the Hamiltonian (simultaneous annihilation) or a similar pair of creation operators (simultaneous creation). These lead to **squeezed states of light** and do not preserve the number of photons. They thus leave the realm of the dual-rail qubit and don't belong to the canonical 2x2 matrix type of operators.

As optical elements can be implemented using nonlinear optics (e.g. sum frequency generation of a pump photon destroys a signal and an ideal photon at the same time). Hence, they are subject to nonlinear optics. In an often-applied approximation the pump photon (or pump photons) is assumed to be classical and the resulting interaction Hamiltonian is still linear. This concept will be discussed in more detail in the context of SPDC-sources for entangled photon pairs.

3.4.5 Some more notes

- Any ***N*-port interferometer** (i.e. an arbitrary $n \times n$ Unitary Matrix) can also be constructed from a series of 2-port beam-splitters and phase shifters \rightarrow any linear optical element for any number of modes can be thought of as a (possibly very complicated) set of beam-splitters and phase-shifters
- we **have only considered two-mode interaction**, i.e. Hamiltonians which consist of sums of bi-linear entries. This results in evolution operators \hat{U} whose action on the modes, i.e. $\hat{U} \hat{a}_j \hat{U}^\dagger$ can be **simplified to a Bogoliubov transformation** and an appropriate matrix $\alpha_{jl}(\hat{U})$. This is no longer possible, if the summand of the interaction Hamiltonian consists of more than two entries, e.g. of the type $\hat{a}_j \hat{a}_l \hat{a}_k$ as is the case for nonlinear optics, this has two profound consequences
 - the matrix $\alpha_{jl}(\hat{U})$ can be diagonalized, i.e. **there is a basis in which the action of the linear optical element is trivial**. This means that linear optical elements are no proper interactions in the strict sense at all.
 - **any non-trivial interaction of photons requires multi-photon, i.e. nonlinear, interactions**. Quantum photonics is thus deeply intertwined with nonlinear optics.

3.5 Mixed Single-Qubit States

In 2.4 we have introduced mixed states as a representation for the statistical uncertainty of a quantum field. Of course, such kind of uncertainty may also be attributed to the state of a (polarization) qubit. It may i.e. be produced by a light source, which does not always produce a fixed polarization, either by competing physical processes (e.g. a thermal emitter) or by statistical variations of the source over time (e.g. random/thermal reorientation of a dipole emitter or someone maliciously kicking your laser from time to time). Consider the mixed state:

$$\hat{\rho} = p|H\rangle\langle H| + (1-p)|V\rangle\langle V| \quad (147)$$

If the state was pure, e.g. $p = 1$, then the density matrix would correspond to the pure state $|H\rangle$ and its representative point on the equator of the Poincaré-Sphere. The same is true for $p = 0$. The mixed state above can thus be thought of as lying on the connection line between the $|H\rangle$ and the $|V\rangle$ point, with a fraction of p of the way from $|V\rangle$ to $|H\rangle$. Thus, mixed states lie inside the Poincaré sphere and the center of the sphere at $\hat{\rho}_{\text{Unpol}} = \frac{1}{2}|H\rangle\langle H| + \frac{1}{2}|V\rangle\langle V|$ is the maximally mixed state, i.e. completely unpolarized light.

It is also obvious that any point inside the Poincaré-Sphere may be reached with multiple mixtures. As one example, $\hat{\rho}_{\text{Unpol}} = \frac{1}{2}|R\rangle\langle R| + \frac{1}{2}|L\rangle\langle L| = \frac{1}{4}|R\rangle\langle R| + \frac{1}{4}|L\rangle\langle L| + \frac{1}{4}|U\rangle\langle U| + \frac{1}{4}|D\rangle\langle D|$ may be decomposed into mixtures of left- and right handed circular states or mixtures of left- and right handed and up. A density matrix decomposition of any point on inside the Poincaré-sphere is therefore never unique. It is, however, conceptually simple to use the three orthogonal axes to define the position of any point, which we have seen above are defined by the Pauli-Matrices. Thus, one can define any mixed polarization state (and thus any mixed Qubit state) according to:

$$\hat{\rho} = \frac{1}{2}(\mathbb{I} + \vec{s} \cdot \hat{\sigma}) \quad (148)$$

where \vec{s} ist the so-called *Stokes-Vector*, with each entry $s_i \in (-1,1)$. We immediately note that $\text{Tr}(\hat{\rho}) = 1$ is automatically fulfilled and the expectation value for a polarization measured along the axis i is given as

$$\text{Tr}(\hat{\rho}\hat{\sigma}_i) = \frac{1}{2}\text{Tr}(\hat{\sigma}_i + 2s_i) = s_i \quad (149)$$

Which is just, what we expected; i.e. if we measure any type of polarized light (pure or mixed) with a polarization beam splitter along the axis j , then we will get the value of the appropriate stokes vector entry as an average measurement result.

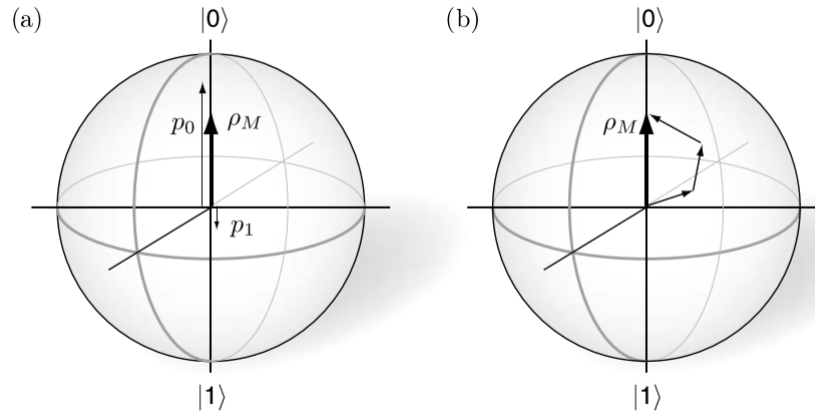
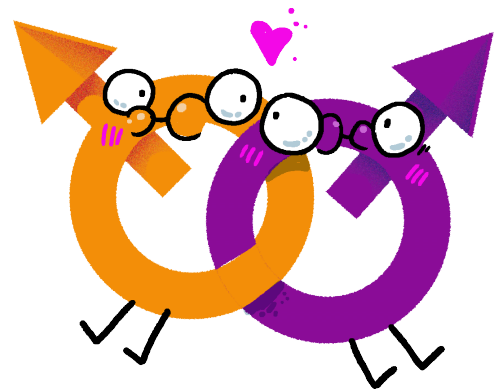


Fig. 11: A mixed state is represented by a Point inside the Poincaré-sphere. (a) Representation of the state as $\hat{\rho}_M = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ and (b) as an alternative but equally viable mixture. (Image Source Lovet & Kok, Introduction to Optical Quantum Information Processing)

4 Multiple Qubits and Entanglement

So far, we have only discussed individual Q-Bits. Many protocols in Quantum Communication and Quantum Information Processing in general rely explicitly on multiple Qubits. Imagine a physical system, which consists of multiple qubits, say for example multiple photons, which we shall number from 1 to N . In such a system each individual Qubit i must behave like an individual Qubit, the way we are used to dealing with. Thus, the state of this qubits must be given by

$$|\psi_i\rangle = \alpha_i|0_i\rangle + \beta_i|1_i\rangle \quad (150)$$



In other words: each Qubit's Basis thus spans its own two-dimensional Hilbert-Space \mathcal{H}_i which is independent of the Hilbert spaces of all other Qubits, because the basis state have a distinct and separable meaning (i.e. you can measure each state separately). A system of N Qubits must therefore span a Hilbert space \mathcal{H} :

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_N \quad (151)$$

Which means that the Hilbert space is spanned by the basis vectors composed of all possible combinations of individual computational basis vectors for the individual basis states $|b_1\rangle \otimes |b_2\rangle \otimes \dots \otimes |b_N\rangle$, where $|b_i\rangle \in \{0,1\}$. Thus, any possible state in the complete system is given by

$$\begin{aligned} |\psi\rangle &= \sum_{b_1=0}^1 \sum_{b_2=0}^1 \dots \sum_{b_N=0}^1 \alpha_{b_1 b_2 \dots b_N} |b_1\rangle \otimes |b_2\rangle \otimes \dots \otimes |b_N\rangle \\ &= \sum_{b_1=0}^1 \sum_{b_2=0}^1 \dots \sum_{b_N=0}^1 \alpha_{b_1 b_2 \dots b_N} |b_1 b_2 \dots b_N\rangle \end{aligned} \quad (152)$$

Where $\sum_{b_1=0}^1 \sum_{b_2=0}^1 \dots \sum_{b_N=0}^1 |\alpha_{b_1 b_2 \dots b_N}|^2 = 1$ must hold for reasons of normalization. The second line differs from the first in just the fact that the tensorial product of the basis vectors has been written in a shorthanded notation. To make this more obvious: $|b_1 b_2 \dots b_N\rangle$ is the state, where each Qubit i is in the state $|b_i\rangle$; e.g. $|000\rangle$ is a three qubit system in a state where all qubit have value zero, e.g. they are all horizontally polarized. These basis vectors $|b_1 b_2 \dots b_N\rangle$ are called the computational basis states.

If the composite system $|\psi\rangle$ is composed of mutually independent photons, the state of the complete system of N photons is simply:

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_N\rangle \quad (153)$$

then the relation of the quantum amplitudes is simply:

$$\alpha_{b_1 b_2 \dots b_N} = \alpha_{b_1}^{(1)} \cdot \alpha_{b_2}^{(2)} \cdot \dots \cdot \alpha_{b_N}^{(N)} \quad (154)$$

However, most states in the combined system cannot be rewritten in terms of individual product states, as defined above, which becomes immediately clear from simple combinatorial arguments. Assume that you have a N -Qubit system, then you require 2^N expansion coefficients $\alpha_{b_1 b_2 \dots b_N}$ to describe any possible state of that system. If you, however, have N individual states there are just $2N$ expansion coefficients α_i, β_i .

To make that more obvious: assume you have a three-qubit system. There are eight possible combinations of the individual qubit states $|b_i\rangle$ and thus eight possible basis states $|b_1 b_2 b_3\rangle$, ranging from $|000\rangle$ over $|001\rangle$ to $|111\rangle$, with a total of eight expansion coefficients $\alpha_{b_1 b_2 b_3}$. If the state was composed of individual states there were only six $\alpha_0^{(1)} \dots \alpha_0^{(3)}$ and $\alpha_1^{(1)} \dots \alpha_1^{(3)}$.

From this simple argument you immediately see that multi-qubit systems have a much larger complexity than all of their composite systems individually. Moreover, the difference scales exponentially. We shall later see that this difference is already apparent for $N = 2$. Any state that cannot be written as a product state shall be called entangled.

We shall see that entanglement entails correlations between quantum particles, which cannot be explained by classical physics. We will start to unravel the type of measurement correlations in two-photon systems, by discussing their differing behaviour under measurement.

4.1 Product States and Non-Correlation

For $N = 2$, the computational basis states are $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ and superpositions thereof to describe any state of the quantum system $|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$.

If the quantum system is composed of two otherwise independent photons it is a product state $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle$ of the individual photons states $|\psi_1\rangle$ and $|\psi_2\rangle$. Since the photons are – per construction – independent we expect, that any measurement (i.e. a polarization measurement) on the first Qubit does not affect the outcome of the measurement on the second Qubit, whatsoever. Nor does it produce any information on the state of the second Qubit.

To show how this maps out in the quantum formalism we assume an arbitrary measurement on Qubit one \hat{A}_1 , which shall be described by its two orthogonal projection operators and measurement results of ± 1 . The basis states of the projection operators shall be called $|a_1\rangle$ and $|a_2\rangle$ without loss of generality

$$\hat{A}_1 = |a_1\rangle\langle a_1| - |a_2\rangle\langle a_2| \quad (155)$$

We can decompose the state of the first qubit into the basis states of the first measurement operator, according to $|\psi_1\rangle = \cos \theta |a_1\rangle + \sin \theta \exp(i\phi) |a_2\rangle$:

$$|\psi_1\rangle|\psi_2\rangle = (\cos \theta |a_1\rangle + \sin \theta \exp(i\phi) |a_2\rangle) |\psi_2\rangle = \cos \theta |a_1\psi_2\rangle + \sin \theta \exp(i\phi) |a_2\psi_2\rangle \quad (156)$$

The measurement then collapses the first Qubit onto $|a_1\rangle$ with probability $\cos^2 \theta$ resulting in a joint state of $|a_1\rangle|\psi_2\rangle$ and onto $|a_2\rangle$ with probability $\sin^2 \theta$ resulting in a joint state of $|a_2\rangle|\psi_2\rangle$. The result

is a classically random ensemble and must therefore be treated in the mixed state formalism with a density matrix:

$$\begin{aligned}\hat{\rho} &= \cos^2 \theta |a_1\rangle\langle a_1| \otimes |\psi_2\rangle\langle \psi_2| + \sin^2 \theta |a_2\rangle\langle a_2| \otimes |\psi_2\rangle\langle \psi_2| \\ &= (\cos^2 \theta |a_1\rangle\langle a_1| + \sin^2 \theta |a_2\rangle\langle a_2|) \otimes |\psi_2\rangle\langle \psi_2| \\ &= \hat{\rho}_1 \otimes |\psi_2\rangle\langle \psi_2|\end{aligned}\quad (157)$$

From this result you can clearly see, that the measurement procedure has neither extracted any information from the second qubit, nor has it affected the second qubit in any tangible way or form. The measurement results are thus uncorrelated. Moreover, we note that the result has left Qubit 2 in a pure state. The same was true, if we had carried out the first measurement on the second Qubit.

Altogether this seems like a rather classical result: a measurement on Qubit 1 does not affect Qubit 2 and it also does not produce any prior information on Qubit 2. Or to put it in other terms: product states behave like classically independent systems, they are thus kind of boring.

4.2 Non-Product States, Correlation, and Entanglement

We shall now see that this classicality is not maintained for non-product states. For this we shall introduce a new basis set for the two-qubit system as an alternative to the CBS $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Among the many possible set of basis states, one, which stands out particularly, is the set of maximally entangled Bell-States $|\Psi/\Phi^\pm\rangle$:

$$\begin{aligned}|\Psi^\pm\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \\ |\Phi^\pm\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)\end{aligned}\quad (158)$$

These states cannot be expressed in terms of product states. I shall leave the proof thereof for you. Let's repeat our measurement experiment for any of these, say $|\psi\rangle = |\Phi^+\rangle$

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0_1 0_2\rangle + |1_1 1_2\rangle)\quad (159)$$

We measure the first Qubit in an arbitrary observable, which is again defined by its projection-based measurement operator. As a reminder this operator is

$$\hat{A}(\theta, \phi)_1 = |a_1\rangle\langle a_1| - |b_1\rangle\langle b_1|\quad (160)$$

The measurement corresponds to some arbitrary basis (not necessarily the CBS), which can be represented by a point on the Bloch sphere for $|a_1\rangle$ and a point on the opposite side for $|b_1\rangle$, which we can describe by the two angles θ and ϕ according to the equations:

$$|a_1\rangle = \cos \theta |0_1\rangle + \sin \theta \exp i\phi |1_1\rangle, |b_1\rangle = \sin \theta \exp(-i\phi) |0_1\rangle - \cos \theta |1_1\rangle\quad (161)$$

This simply means, that θ represents how far away on the Bloch-Sphere we are from the CBS. Here $\theta = 0$ and $\theta = \frac{\pi}{2}$ represent measurements in the CBS basis and $\theta = \pm \frac{\pi}{4}$ represent measurements on the equator of the Bloch-Sphere, e.g. the $|\pm\rangle$ or the $|L/R\rangle$ bases or superpositions thereof. The specific choice of factors also automatically ensures that $|a_1\rangle$ and $|b_1\rangle$ are orthonormal, i.e. they are a valid basis set.

As we must expand the CBS in which the initial state was defined into these states anyway it makes sense to expand the basis states into the eigenstates of the observable:

$$|0_1\rangle = \cos \theta |a_1\rangle + \sin \theta \exp i\phi |b_1\rangle |1_1\rangle = \sin \theta \exp(-i\phi) |a_1\rangle - \cos \theta |b_1\rangle \quad (162)$$

At any rate, we can now describe the first Qubit's CBS as a superposition of the measurement basis and we just plug that into the definition of $|\psi\rangle$:

$$\begin{aligned} |\psi\rangle &= \frac{1}{\sqrt{2}} [(\cos \theta |a_1\rangle + \sin \theta \exp(i\phi) |b_1\rangle) |0_2\rangle + (\sin \theta \exp(-i\phi) |a_1\rangle - \cos \theta |b_1\rangle) |1_2\rangle] \\ &= \frac{1}{\sqrt{2}} [(\cos \theta |0_2\rangle + \sin \theta \exp(-i\phi) |1_2\rangle) |a_1\rangle + (\sin \theta \exp(i\phi) |0_2\rangle - \cos \theta |1_2\rangle) |b_1\rangle] \end{aligned} \quad (163)$$

The measurement then collapses the first Qubit and each of the terms has a certain probability of being the resulting state after collapse. The probabilities are:

$$\begin{aligned} p(A_1 = +1) &= \langle \psi | \hat{P}_a | \psi \rangle \\ &= \langle \psi | a_1 \rangle \langle a_1 | \psi \rangle \\ &= \frac{1}{2} [(\cos \theta \langle 0_2 | + \sin \theta \exp(i\phi) \langle 1_2 |)] [(\cos \theta |0_2\rangle + \sin \theta \exp(-i\phi) |1_2\rangle)] \\ &= \frac{1}{2} [\cos^2 \theta + \sin^2 \theta] \\ &= \frac{1}{2} \\ p(A_1 = -1) &= \langle \psi | \hat{P}_b | \psi \rangle = \frac{1}{2} \end{aligned} \quad (164)$$

The states after the measurement are:

$$\begin{aligned} |\psi | A_1 = +1\rangle &= (\cos \theta |0_2\rangle + \sin \theta \exp(-i\phi) |1_2\rangle) |a_1\rangle \\ |\psi | A_1 = -1\rangle &= (\sin \theta \exp(i\phi) |0_2\rangle - \cos \theta |1_2\rangle) |b_1\rangle \end{aligned} \quad (165)$$

Here we note the first curious thing. The resulting probability distributions of Qubit 1 do not at all depend on the type of measurement applied. From the single particle picture, you would expect that a quantum particle must have one specific property (observable), where the result is fixed. Or to put it more bluntly: by now you have accepted that it may not be clear what property a Quantum Particle may have, but you would surely expect that it should have one property. Yet, ANY possible measurement, which you can apply on Qubit 1 gives the same result. It seems like Qubit 1 has become a particle without properties.

This also means that before the measurement there is no point on the Bloch Sphere, which describes the state of Qubit 1 by itself. Hence the initial state is NOT represented by two points in the Bloch sphere. In a sense Qubit 1 has ceased to exist as an independent particle. Instead, it is in a state, in which it does not make sense to think about the properties of Qubit 1 without resolving its connection with Qubit 2. Both Qubits have become ENTANGLED.

That said, let's explore the status of the joint system after the measurement on Qubit 1. As it is in a mixed state it must be described using the density matrix approach, where we can simply read off the entirety of the density operator from the table above

$$\hat{\rho} = \frac{1}{2} [|\psi | A_1 = +1\rangle \langle \psi | A_1 = +1| + |\psi | A_1 = -1\rangle \langle \psi | A_1 = -1|] \quad (166)$$

Which is clearly not factorizable in the same way, as the non-correlated state from above. Let's elaborate on this a bit more in-depth by explicitly calculating the state of the second Qubit. This is done by

calculating the partial trace over the first Qubit (e.g., a hypothetical measurement with the identity operator for Qubit 1).

$$\begin{aligned}
 \hat{\rho}_2 &= \text{Tr}_1 \hat{\rho} = \sum_i \langle a_i | \hat{\rho} | a_i \rangle \\
 &= \frac{1}{2} (\cos \theta |0_2\rangle + \sin \theta \exp(-i\phi) |1_2\rangle) (\cos \theta \langle 0_2| + \sin \theta \exp(i\phi) \langle 1_2|) \\
 &\quad + \frac{1}{2} (\sin \theta \exp(i\phi) |0_2\rangle - \cos \theta |1_2\rangle) (\sin \theta \exp(-i\phi) \langle 0_2| - \cos \theta \langle 1_2|) \\
 &= \frac{1}{2} (\cos^2 \theta + \sin^2 \theta) |0_2\rangle \langle 0_2| + \frac{1}{2} (\cos^2 \theta + \sin^2 \theta) |1_2\rangle \langle 1_2| \\
 &\quad + \frac{1}{2} (\cos \theta \sin \theta \exp(i\phi) - \cos \theta \sin \theta \exp(i\phi)) |0_2\rangle \langle 1_2| \\
 &\quad + \frac{1}{2} (\cos \theta \sin \theta \exp(-i\phi) - \cos \theta \sin \theta \exp(-i\phi)) |1_2\rangle \langle 0_2| \\
 &= \frac{1}{2} [|0_2\rangle \langle 0_2| + |1_2\rangle \langle 1_2|] \tag{167}
 \end{aligned}$$

This is not just any mixed state but a maximally mixed state according to the definition in chapter 2.4. This means that a measurement in Qubit 1 does not only increase the information content (entropy) of Qubit 1 it also increases the entropy of Qubit 2. Indeed, this is much weirder than you would initially expect. Let's set this aside for a second and use this finding as a means to define the entangledness of a quantum system:

The degree of Entanglement of a two-Qubit quantum system in a joined state $|\psi\rangle$ is measured by testing the purity of the state of Qubit 2 after a measurement A_1 is applied onto Qubit 1, i.e. let $\hat{\rho}$ be the state of the joint system after application of measurement. Then the entanglement E is calculated using $E = 2 \text{Tr}[(\text{Tr}_1 \hat{\rho})^2]$. $E \in [0,1]$ with $E = 0$ indicating non-entanglement and $E = 1$ indicating maximum entanglement. The specific kind of measurement of Qubit 1 does not matter. A generalization with larger systems is straightforward.

Let's return to the weirdness of entangled systems. Previously, we had seen that Quantum Systems are transferred from a pure into a mixed state by measurement only. But we have never even touched system Qubit 2. We have only measured Qubit 1. Still, in the process we have transformed Qubit 2 into a mixed state. This means we must have made implicitly some sort of measurement with Qubit 2. Let's find this out and do so by applying the observable $\hat{A}(\theta, -\phi)_2$, onto Qubit 2 (this is the same as for Qubit 1, with the only exception that the phase shift between the two measurement bases is reversed, e.g. the sense of the chirality is flipped).

We rewrite the state of the second Qubit system into two parts, according to the measurement outcome of A_2 (we could proceed with the complete $\hat{\rho}$ from above but then the equations get somewhat lengthy):

$$\begin{aligned}
 |\psi_2 |A_1 = +1\rangle &= \cos \theta |0_2\rangle + \sin \theta \exp(-i\phi) |1_2\rangle \\
 |\psi_2 |A_1 = -1\rangle &= \sin \theta \exp(i\phi) |0_2\rangle - \cos \theta |1_2\rangle \tag{168}
 \end{aligned}$$

Let's now apply the same measurement (let's call it A_2), which have applied to the first Qubit on the second qubit (e.g. we assume that both angles θ and ϕ are the same for Qubit 1 and Qubit 2. We now calculate the probabilities of A_2 by noting that $p(A_2 = a_q | A_1 = a_r) = \text{Tr}(\hat{\rho}_2(A_1 = a_r) |a_q\rangle \langle a_q|) = \sum_i \langle a_i | \hat{\rho}_2(A_1 = a_r) |a_q\rangle \langle a_q | a_i \rangle$. We read them off as:

$$\begin{aligned}
 p(A_2 = +1|A_1 = +1) &= |\langle a_1 | \psi_2 | A_1 = +1 \rangle|^2 \\
 &= |\langle a_1 | (\cos\theta |0_2\rangle + \sin\theta \exp(-i\phi) |1_2\rangle) \rangle|^2 \\
 &= |\cos^2\theta + \sin^2\theta|^2 \\
 &= 1 \\
 p(A_2 = -1|A_1 = +1) &= |\langle a_2 | \psi_2 | A_1 = +1 \rangle|^2 \\
 &= |\cos\theta \sin\theta \exp(-i\phi) - \cos\theta \sin\theta \exp(-i\phi)|^2 \\
 &= 0 \\
 p(A_2 = +1|A_1 = -1) &= 0 \\
 p(A_2 = -1|A_1 = -1) &= 1
 \end{aligned} \tag{169}$$

Note, that we have explicitly shown, how the first solution is obtained and then just given the result for the second to fourth. We now group the four cases into two classes. The situation $(A_2 = +1|A_1 = +1)$ and $(A_2 = -1|A_1 = -1)$ mean that the measurements on Qubit Number 2 will yield the SAME result as the measurement on Qubit Number 1 (correlation). The other two situations correspond to measurements with different results (anticorrelation). We find that both members in both of the classes are equal and they are 1 and 0 exclusively.

This result is profound: a measurement of Qubit 1 with observable A_1 will force Qubit 2 to instantly collapse into the same resulting state for observable A_2 irrespective of the measurement outcome. No matter what the results are, they perfectly correlated. Moreover, and this in as important point: the correlation is maintained irrespective of the measurement basis! The two Qubit always give the same results, irrespective of what you measure, as long as, you make the same measurement.

Or in other words, the observable in the A_1 measurement basis is perfectly correlated to the observable in the same basis, with a flipped phase as represented by the observable A_2 .

Here we have only discussed this relation for an initial two-Qubit system in the Φ^+ state but one can show that for the other three Bell-States there is a correlated Basis for Qubit 2 for any possible measurement of Qubit 1 (this is relation is just a slight bit more complicated than just a flip of the ϕ -phase). This leads us to an alternative definition of entanglement:

Two Qubits are completely entangled, if for any basis set for Qubit 1 there exists a corresponding basis set for Qubit 2, in which a measurement is guaranteed to yield the identical result. The degree of entanglement can be quantified by the maximum degree of correlation between a measurement in a basis set in Qubit 1 and the most correlated basis set in Qubit 2.

In other words: measurements in entangled systems produce correlated results, irrespective of the measurement!

4.3 The No-Cloning Theorem

After getting a better understanding on the thing that you can do with a Two-Qubit-System, we shall now look into something apparently simple, that you cannot do with a Two-Qubit System, not with any other quantum system in question. The question is very simple: can we copy the state of one qubit onto another qubit, without destroying the initial qubit.

Remember that the state of a (pure) qubit is represented by an arbitrary point on the Bloch-sphere, depending on the chosen basis vectors and the coefficients α and β or equally by the angles Θ and ϕ . If you attempt to measure its state, you must choose a certain basis in which to measure. This basis is represented by a specific Pauli-Operator or a superposition thereof. However, we have learned, that these operators are complementary, which in essence means, that you only ever get one chance of

measuring your polarization state (with a result of ± 1), without permanently and irrevocably destroying the specific state.

If you knew the specific basis in which the qubit was operated, then you'd be quite fine (in the sense of, that you'd only have to determine, on which side of the sphere your state is). In general, however, you end up in a situation, where you have absolutely no chance of measuring the complete state of your qubit, unless you have a lot of advance knowledge. Full stop.

To make it simple: a qubit may be any point on the Bloch-Sphere, i.e. it's defined by two real numbers, but you only ever get to measure on which side of the globe it (most likely) was. And as you cannot copy, what you cannot measure, you end up in a situation that in most of the cases you cannot clone a qubit.

This idea can be proven rigorously, with the two-Qubit notation¹². Suppose that we have a cloning operator \hat{U} , which operates on two combined qubits with states $|\phi\rangle$ and $|k\rangle$, such that it copies the state of $|\phi\rangle$ onto $|k\rangle$, i.e.:

$$\hat{U}(|\phi\rangle \otimes |k\rangle) = |\phi\rangle \otimes |\phi\rangle \quad (170)$$

As a cloning-operator \hat{U} must of course work in the same way for any other state $|\psi\rangle$, too, i.e.

$$\hat{U}(|\psi\rangle \otimes |k\rangle) = |\psi\rangle \otimes |\psi\rangle \quad (171)$$

Needless to say, that \hat{U} must be connected to a physical process and thus must be unitarian. Let's now compare the two results by taking their scalar product:

$$\begin{aligned} \langle \hat{U}(|\phi\rangle \otimes |k\rangle) | \hat{U}(|\psi\rangle \otimes |k\rangle) \rangle &= \langle \phi \otimes \phi | \psi \otimes \psi \rangle \\ \langle \hat{U}(|\phi\rangle \otimes |k\rangle) | \hat{U}(|\psi\rangle \otimes |k\rangle) \rangle &= \langle k \otimes \phi | \hat{U}^\dagger \hat{U} | \psi \otimes k \rangle = \langle k \otimes \phi | \psi \otimes k \rangle \end{aligned} \quad (172)$$

The first line is simply taken from the definition of the cloning operator, whereas the last line utilized the fact that \hat{U} is unitarian. Thus we find:

$$\langle \phi \otimes \phi | \psi \otimes \psi \rangle = \langle k \otimes \phi | \psi \otimes k \rangle \quad (173)$$

Because the tensor and the scalar product can be exchanged, we simplify both sides of the equation to:

$$\langle \phi | \psi \rangle \langle \phi | \psi \rangle = \langle \phi | \psi \rangle \langle k | k \rangle \quad (174)$$

Because $\langle k | k \rangle = 1$ we get:

$$\langle \phi | \psi \rangle^2 = \langle \phi | \psi \rangle \quad (175)$$

This result is crucial. It can either be fulfilled if $\langle \phi | \psi \rangle = 1$, which means that $|\phi\rangle = |\psi\rangle$, which is trivial or if $\langle \phi | \psi \rangle = 0$, which means that $|\phi\rangle$ is orthogonal to $|\psi\rangle$. In other words: if you have found a cloning operator that works on one specific quantum state (e.g. a Qubit), it can only work on orthogonal quantum states as well but it will not work for arbitrary quantum states. Thus, if you cannot find a cloning operator, i.e. any physical process, that copies quantum states, then you cannot copy a quantum state. As long as you have to stick to the laws of nature, that is.

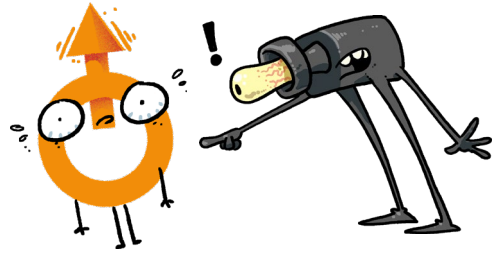
The central argument for the derivation of the no-cloning theorem is obviously the unitarity of \hat{U} . In terms of time evolution unitarity is equivalent to time-reversibility and thus to a constant entropy: In

¹² In fact, this works for any type of quantum system; qubit or not.

other words: quantum operations must not destroy information. The supposed cloning-operator, however, would just do that: it would destroy any information of the prior state $|k\rangle$ of the target system upon it being overwritten with $|\phi\rangle$. Thus cloning, from a thermodynamic point of view, is an irreversible process and quantum mechanics just does not provide any means to do that.¹³

5 Creating and Characterizing Photons

Now that we have understood everything about the nature of photonic qubits, we still have not discussed, how real single photon quantified and created. Due to timing reasons this chapter is a bit handwaving and a proper introduction can be found in the appendix.



For the characterization section we'll discuss two important experiments for there. First the Hanbury-Brown-Twiss Interferometer, which answers the question: does a source in fact emit single photons (and if yes, how frequently)? Then we shall deal with the Hong-Ou-Mandel interferometer, which tell us, if the two single photons are indistinguishable; e.g. if two consecutive photons emitted by a single photon source are exactly the same or if the source is fluctuating.

In the photon creation section we shall briefly introduce single photon emitters and photon pair emitters, which are based on the SPDC-process. We shall then briefly show, how the SPDC-process can be used to create Bell states; e.g. entangled photons. Also for this subchapter we shall remain somewhat handwaving and export a lot of the theory into the appendix.

The corresponding chapter of the Appendix are A1 for a description of photodetection and photon correlation on the framework of quantum fields, A2 for an overview over common single photon detector platforms, A3 for Spontaneous Emission and Single Photon Sources based thereupon and A4 for a basic description of the SPDC-process, which is most frequently used to create pairs of photons.

5.1 Single Photon Sources

We are now equipped with some of the basic tools required to formally describe the detection process, and coherence of quantum fields. In the following, it will be helpful to also have a physical model for typical quantum light sources. Here we briefly introduce some of the basic concepts – we will circle back to the physical implementation and key characteristics of various sources later on.

5.1.1 Spontaneous Emission Single Photon Sources

The process of Spontaneous emission does emit single photons, if a single emitter is considered. The simplest of such a emitters is a two-level system, which is initially in an excited state. The excited state has a lifetime of τ and an energy difference of ΔE to its ground state. The excited state will spontaneously decay within the lifetime and emit a single photon in the process. The photon will have a center frequency $\hbar\omega = \Delta E$, a duration of τ and a Lorentzian spectrum with $\Delta\omega = 1/\tau$.

¹³ Note that if you replace $|k\rangle$ with a many-body thermal bath, then you can “hide” the reversibility in the huge state-space and the fact that most of these states are in reality very hard to differentiate. Reversibility this thus practically impossible.

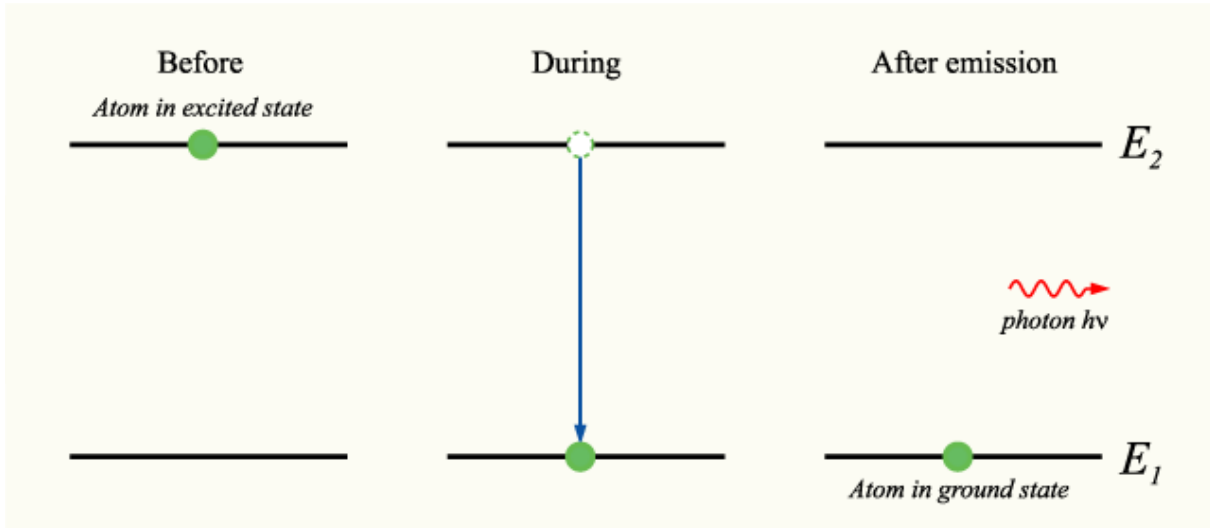


Fig. 12: In spontaneous emission an individual atom (or any other two-level system) begins in an excited state with a lifetime τ . It drops into the ground state and the energy difference ΔE and a single photon with $\hbar\omega = \Delta E$. The photons wavepacket has a duration of τ and a Lorentzian spectrum with a width $\Delta\omega = 1/\tau$. Imagesource: [Dialnet-ModelingABandwidthOfATwoLevelIndependentQuantumLas-6234592.pdf](https://www.researchgate.net/publication/3234592)

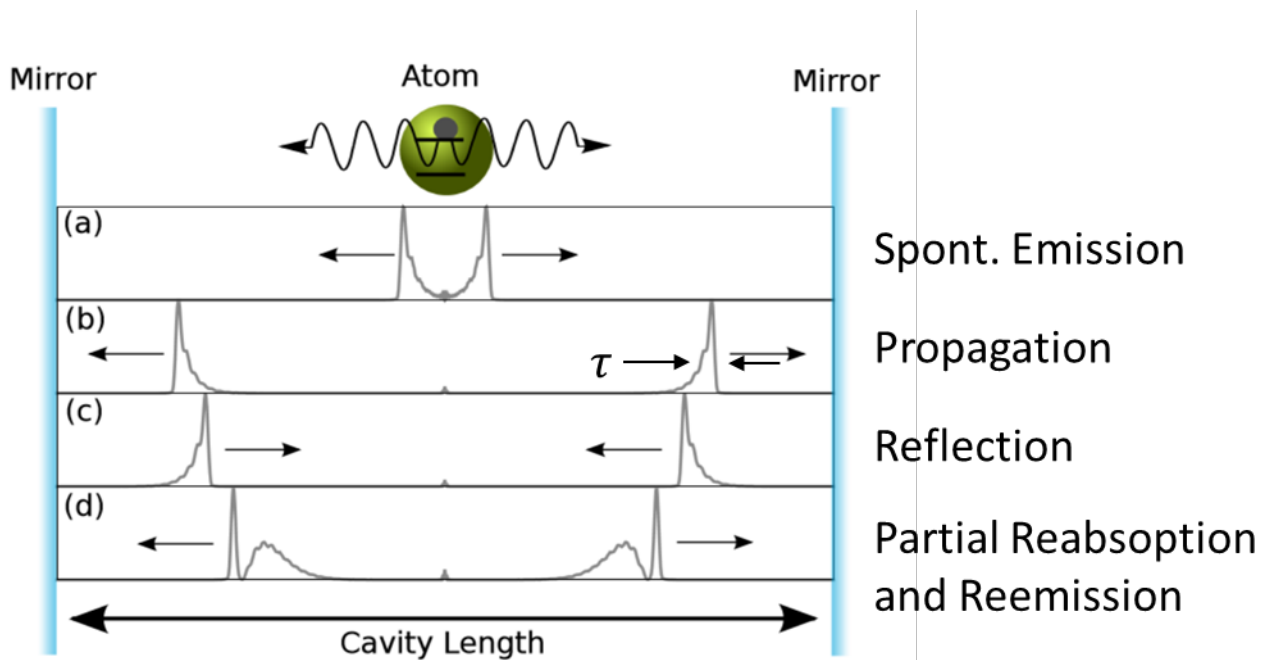


Fig. 13: Full Quantum-Mechanical simulation of the time-evolution of an initially excited two-level system in a one-dimensional cavity. The wavepackets can be understood as a single mode (with Lorentzian shape in spectrum) in which a single photon is being emitted. Reabsorption is not perfect, because the time-symmetry is wrong (strong-before-weak).

While this approach is conceptually extremely simple, it has a few requirements, which are not quite easy to fulfil in practice (at least not all of them at the same time):

- The system must be a near-perfect two-level system, without non-radiative decay (reduction of emission probability) or alternative decay channels (leads to mixed state emission and distinguishability). This is particularly difficult to achieve for emitters embedded in solid state systems, which can couple to a plethora of excitations such as phonons or hyperfine states. Many systems can fulfil this requirement only if they are operated at cryogenic temperatures

- The system must have a short lifetime to produce a high photon rate. This is not true for all emitters.
- The system must not switch its configuration upon excitation, such as many dyes do. This leads to blinking and bleaching, e.g. the temporary or permanent interruption of emission.
- The system must couple efficiently to a photonic state. No reabsorption must take place. Emitted photons must not be trapped in the emitter system, due to, e.g. total internal reflection.
- Add-on: one should be able to couple the emitter to a single spatial mode, e.g. a waveguide

5.1.2 Overview over Experimental Systems

Many such sources exist, however, all of them have one or more drawbacks:

- Single Atoms and ions are ideal single photons emitters, however, they don't stay in one place naturally. Therefore, one has to trap atoms and ions in a magnetic or optical trap and keep them in high vacuum.
- Dye molecules are cheap and easy to operate with. However, they suffer from blinking and bleaching and are thus unsuitable for long-term operation. Moreover, they typically have very wide emission band, making them unsuitable for indistinguishable photons. They also often rely on specific chemical environments, making them hard to integrate in optical systems.
- Quantum Dots are localized modifications in crystals with a bandgap. These defects act as local potential wells and exhibit discrete states. The defects can occur naturally or may be tailor-made. The host crystal may be a semiconductor (small gap) or a dielectric (large gap). Semiconductor QDs often have a small binding-energy (they are shallow in the gap) and must thus be operated cryogenically. However, they can be excited electrically. Dielectric QDs can be very deep in the band-gap and sometimes operated at room temperature. However, they cannot be pumped electrically and must be pumped optically.

Here a small table for comparison:

| | Trapped Atoms / Ions | Dye Molecules | Semicon. QDs | Dielectric QDs |
|--|--|----------------------|---|--------------------------------|
| Example | Rydberg Atoms, Rb-Vapor | Organic dyes | III-V QDs, CdSe-Particles | Diamond-NV, hBN-Defects |
| QE | high | high | Mid-high | Mid-high |
| Purity | high | low | Mid (Phonons) | Mid (Phonons) |
| Environment | High-Cav, Cryo | Spec. Chemical | Cryo | Room. Temp |
| Pumping | optical | optical | Optical / electrical | Optical |
| Integrability | nil | bad | good | exceptional |
| Emission Rate | Low (lifetime) | low (lifetime) | high | High |
| ToR-Loss / Single-Mode-Coupling | Possible, yet expensive (large optics) | ? | Possible, yet difficult (semicon-nano-optics) | Diamond: bad hBN: very easy |

A quantum-mechanical description of the emission process shall later be added in an appendix.

5.2 Characterizing SPS: the Hanbury-Brown-Twiss Experiment

The first-order correlation function quantifies the correlation of amplitudes and phases of two fields (i.e. the phase coherence). The second-order correlation, on the other hand, tells us about the correlation of intensities of two fields. Fig. 59 depicts a typical experimental setup.

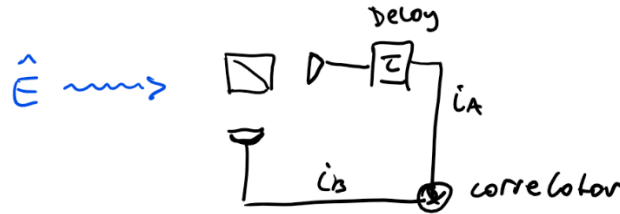


Fig. 14: Hanbury-Brown-Twiss setup is it was first used to measure the 2nd order intensity correlations of classical light in stellar interferometry and then for individual light sources. Note that the detectors are now interconnected and we thus are mostly interested in the order of clicks at the detectors.

The general expression for the second-order quantum correlation of fields A and B at times t_1 and t_2 reads

$$g_{AB}^{(2)}(t_1, t_2) = \frac{\langle \hat{E}_A^-(t_1) \hat{E}_B^-(t_2) \hat{E}_B^+(t_2) \hat{E}_A^+(t_1) \rangle}{\langle \hat{E}_A^-(t_1) \hat{E}_A^+(t_1) \rangle \langle \hat{E}_B^-(t_2) \hat{E}_B^+(t_2) \rangle} \quad (176)$$

For classical fields, this reduces to the Intensity-Intensity correlation function:

$$g_{\text{class}}^{(2)}(t_1, t_2) = \frac{\langle I_A(t_1) I_B(t_2) \rangle}{\langle I_A(t_1) \rangle \langle I_B(t_2) \rangle} \quad (177)$$

An important example is the intensity autocorrelation function of a stationary classical field

$$g_{\text{class}}^{(2)}(\tau) = \frac{\langle I(t+\tau) I(t) \rangle}{\langle I(t) \rangle^2} \quad (178)$$

It is straightforward to show that any classical light field must obey $g_{\text{class}}^{(2)}(\tau) \leq g_{\text{class}}^{(2)}(0)$ and $g_{\text{class}}^{(2)}(0) \geq 1$. This is not necessarily true for the autocorrelation of a quantum state of light $|\Psi\rangle$

$$g_{\text{QM}}^{(2)}(\tau) = \frac{\langle \hat{E}^-(t) \hat{E}^-(t+\tau) \hat{E}^+(t+\tau) \hat{E}^+(t) \rangle_{\Psi}}{\langle \hat{E}^-(t) \hat{E}^+(t) \rangle_{\Psi} \langle \hat{E}^-(t+\tau) \hat{E}^+(t+\tau) \rangle_{\Psi}} \quad (179)$$

While this expression is in general difficult to evaluate, we can get a good understanding of some general properties by evaluating it for $\tau = 0$ and a single-frequency mode. Here the correlation function becomes:

$$|g_{\text{QM}}^{(2)}(0)| = \frac{\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \rangle}{\langle \hat{a}^\dagger \hat{a} \rangle^2} = \frac{\langle \hat{n}(\hat{n}-1) \rangle}{\langle \hat{n} \rangle^2} \quad (180)$$

Evaluating this expression for a single photon Fock state $|\Psi\rangle = \hat{a}^\dagger |\text{vac}\rangle = |1\rangle$ we immediately see that

$$|g_{\text{QM}}^{(2)}(0)| = \langle 1 | \hat{n}(\hat{n}-1) | 1 \rangle = 0 \quad (181)$$

Single-photon states of light thus exhibit “anti-bunching”, a purely quantum phenomenon that cannot be described in classical coherence theory. Other examples that can be readily verified by the reader:

| STATE | $g^{(2)}(0)$ | COMMENT |
|-------|--------------|---------|
|-------|--------------|---------|

| | | |
|--|--|---|
| FOCK-STATE $n\rangle, n = 1$ | 0 | Perfect Anti-Bunching (one photon at a time) |
| FOCK-STATE $n\rangle, n > 1$ | $1 - 1/n$ | Anti-Bunching |
| COHERENT STATE $\alpha\rangle$ | 1 | Uncorrelated (a random stream of photons) |
| THERMAL STATE $\rho = \int f(\omega) \sum_n \frac{1 - \exp(-\frac{\hbar\omega}{k_B T})}{\exp(\frac{n\hbar\omega}{k_B T})} n(\omega)\rangle\langle n(\omega) d\omega$ | $1 + g^{(1)}(0) ^2$ | Bunching of photons of the same frequency (the more so the more narrowband) |
| SQUEEZED STATE (DEGENERATE PDC) | $g^{(2)}(0) = 3 + \frac{1}{\langle n \rangle}$ | Super-Bunched (photons always appear in correlated pairs) |

One can also, quite easily show, that $g_{QM}^{(2)}(\pm\infty) = 1$ and that the transition from the center value to the edge value is related to the bandwidth of the source in question, or more specifically its lifetime.

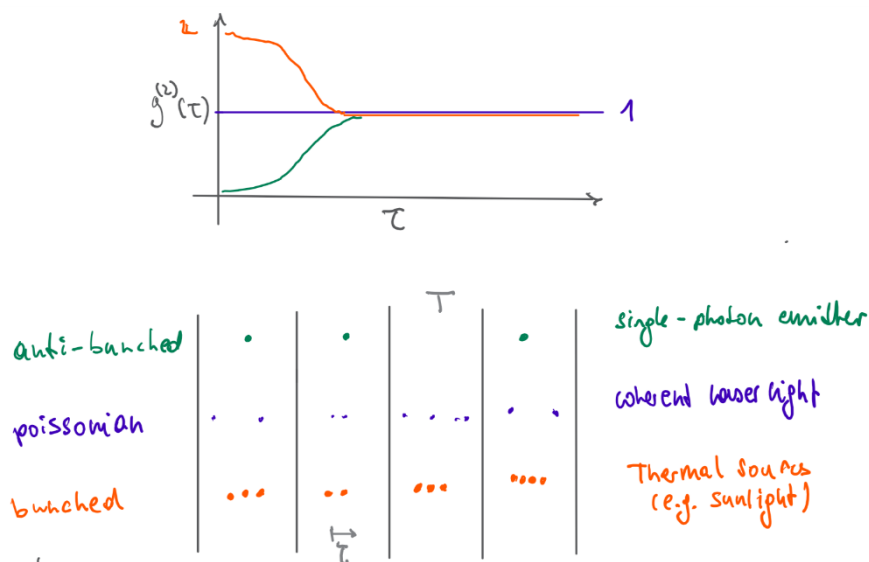


Fig. 15: Artists interpretation of photon arrival times in various sources, resulting in different $g^{(2)}$.

The Hanbury-Brown-Twiss experiment can thus be used to measure the “single-photon-ness” of a light source. It is the gold standard for this kind of characterization.

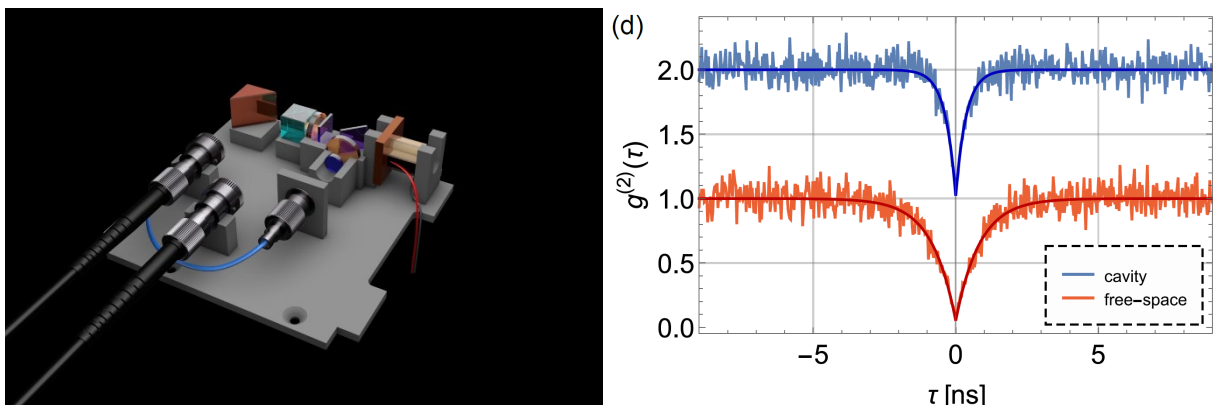


Fig. 16: (left) A real-life single-photon source with an integrated HBT-interferometer (fiber ports lead to SPDs and TDC). The single-photon source is an hBN-defect-state emitter localized in the brown plate. (right) $g^{(2)}$ -curves of the bare hBN emitter and the emitter located in a cavity. For the cavity case we measure $g^{(2)}(0) < 0.006$. The lifetime of the emitters is just below 1 nanosecond. Image source: ACS Photonics 2019, 6, 8, 1955–1962.

5.3 Photon Pair Sources based on SPDC

Now that we have talked about BSM, we should address the elephant in the room and discuss how Bell-States are created in the first place. In the last chapter we had seen, that the conceptually most convenient way is to use a CX and Hadamard-Gate, to convert a BS into CBS. If you also remember that quantum processes are unitary and thus reversible, we can just use the conversion circuit in reverse to convert a CBS into a BS using the operator

$$\hat{U} = \text{CX}_B \text{HAD}_A \quad (182)$$

I'll leave it up to you to prove this actually works. Nevertheless, a CX-Gate is not so easy to implement and thus experimentally this is probably not the way to go. The standard approach is rather to use nonlinear optics directly and create pairs of photons using the spontaneous parametric downconversion process (SPDC) in a $\chi^{(2)}$ -medium such as BBO or KTP or similar. We shall also see that we can also use the process to create heralded single photon sources.

In SPDC a strong pump field in mode p generates a nonlinear polarization response in a $\chi^{(2)}$ nonlinear crystal (non-centro-symmetric materials), that – with very low probability (typically on the order of $1E-9$ per pump photon) results in the emission of a pair of photons, called *signal* and *idler* photons. The range of possible frequencies and momenta of the signal and idler photons will be subject to energy and momentum conservation. Consider for a moment plane wave modes, then the signal s and idler i will obey the following conservation equations:

$$\begin{aligned} \mathbf{k}_p &= \mathbf{k}_s + \mathbf{k}_i \\ \omega_p &= \omega_s + \omega_i \end{aligned} \quad (183)$$

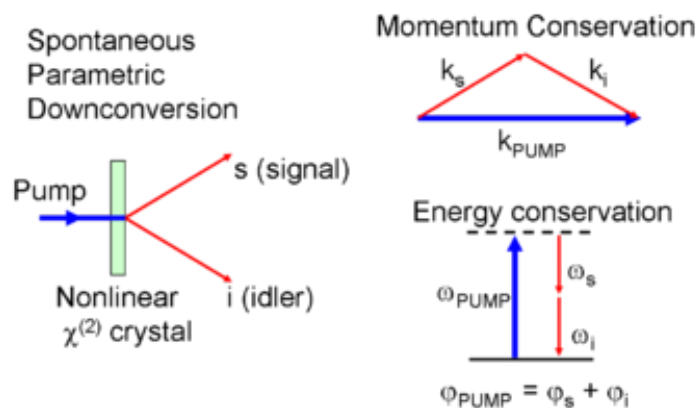


Fig. 17: Conservation Equations in thin-crystal SPDC.

Of course, this does not work any pair of modes but relies on phase-matching criteria, that in turn depend on the type of crystal, its size, orientation, temperature, structuring and the relative angles between the k -vectors. More specifically this process will also only work for a certain combination of polarization directions λ_s and λ_i for the signal and idler, given a certain polarization of the pump λ_p . For uniaxial nonlinear crystals we differentiate between type-(I) and type-(II) phase matching. In type-(I) the signal and idler mode are of the same polarization ($\lambda_i = \lambda_s$) and in type-(II) they are of opposite polarization ($\lambda_i \neq \lambda_s$).

The Hamiltonian governing the SPDC process then writes¹⁴:

$$\hat{\mathcal{H}}_{PDC} = q\chi^{(2)}L \hat{a}^\dagger_{k_s\lambda_s} \hat{a}^\dagger_{k_i\lambda_i} \hat{a}_{k_p\lambda_p} + h.c. \quad (184)$$

Where L is the length of the nonlinear crystal, $\chi^{(2)}$ is the nonlinear interaction strength and q is a proportionality constant.

We can describe the time evolution induced by this Hamiltonian either via the modes, or in the time-dependent quantum state. In the following, we choose the latter approach. The signal and idler modes are initially in their respective vacuum states and the pump is a coherent state with a large mean number of photons α :

$$|\Psi\rangle_{in} = |\alpha\rangle_{k_p} |\text{vac}\rangle_{k_s\lambda_s} |\text{vac}\rangle_{k_i\lambda_i} \quad (185)$$

The state of the field after the interaction is given by:

$$|\Psi\rangle = \exp\left(-\frac{iL}{q}\chi^{(2)} \left[\hat{a}^\dagger_{k_s\lambda_s} \hat{a}^\dagger_{k_i\lambda_i} \hat{a}_{k_p} + h.c. \right]\right) |\Psi\rangle_{in} \quad (186)$$

Since the pump is a coherent state, which is an eigenstate of the annihilation operator, we can replace its operator with the corresponding coherent state amplitude α .

$$|\Psi\rangle_{k_s\lambda_s, k_i\lambda_i} |\alpha\rangle_{k_p} = \exp(-iLq\chi^{(2)} [\alpha \hat{a}^\dagger_{k_s\lambda_s} \hat{a}^\dagger_{k_i\lambda_i} + h.c.]) |\text{vac}\rangle_{k_s\lambda_s} |\text{vac}\rangle_{k_i\lambda_i} |\alpha\rangle_{k_p} \quad (187)$$

Factoring out the state of the signal and idler photons, and noting that the h.c. term containing the annihilation operators will not contribute when it acts on vacuum, we have:

$$|\Psi\rangle_{SPDC} = |\Psi\rangle_{k_s\lambda_s, k_i\lambda_i} = \exp(-iLq\alpha\chi^{(2)} \hat{a}^\dagger_{k_s\lambda_s} \hat{a}^\dagger_{k_i\lambda_i}) |\text{vac}\rangle_{k_s\lambda_s} |\text{vac}\rangle_{k_i\lambda_i} \quad (188)$$

Application of the BCH-theorem, we obtain the SPDC state in the Fock state basis:

$$|\Psi\rangle_{SPDC} = \frac{1}{\cosh|\gamma|} \sum_{n=0}^{\infty} \frac{(-\gamma)^n}{|\gamma|^n} \tanh^n|\gamma| \cdot |n\rangle_{k_s\lambda_s} |n\rangle_{k_i\lambda_i} \quad (189)$$

Where we have introduced the overall gain $\gamma = iLq\alpha\chi^{(2)}$. Setting $\lambda = \tanh|\gamma|$ we can write the SPDC state in a more condensed form:

$$|\Psi\rangle_{SPDC} = \frac{1}{\sqrt{1-\lambda^2}} \sum_{n=0}^{\infty} e^{in\phi} \cdot \lambda^n \cdot |n\rangle_{k_s\lambda_s} |n\rangle_{k_i\lambda_i} \quad (190)$$

The state is also known as the two-mode squeezed vacuum state. It has the following Important characteristics:

- The number of photons in modes k_s and k_i are perfectly correlated: $\langle\Psi|(\hat{n}_s - \hat{n}_i)|\Psi\rangle = 0$.
- Average photon number : $\mu = \langle\Psi|\hat{n}|\Psi\rangle = \sinh^2|\gamma|$
- Probability of n photons: $P(n) = \langle\Psi|\hat{P}_n|\Psi\rangle = \frac{\tanh^{2n}|\gamma|}{\cosh^n|\gamma|} \propto \frac{\mu^n}{(1+\mu)^{n+1}}$

¹⁴ As dicussed in the previous chapter this literally means: the interaction process can annihilate a photon in mode k_p if it creates one in k_s and k_i at the same time and vice versa.

In the limit of low gain $\lambda \ll 1$, the SPDC state is dominated by a vacuum term, with a small photon pair contribution:

$$|\Psi\rangle_{SPDC} \approx \sqrt{1 - \lambda^2} |\text{vac}\rangle_{k_s \lambda_s} |\text{vac}\rangle_{k_i \lambda_i} + \lambda |1\rangle_{k_s \lambda_s} |1\rangle_{k_i \lambda_i} + O(\lambda^2) \quad (191)$$

SPDC in the low gain regime is a very simple and practical way of generating photon pairs and single photon states.

A much more detailed analysis of SPDC is presented in the Appendix (A 3).

5.3.1 Heralded Single Photon Sources

To produce single photon states we place a detector in the idler mode; when the detector fires (neglecting for the moment the practical issues such as detector noise, i.e. dark counts), we expect at least one photon to be present in the signal mode. Since $\lambda \ll 1$ the main contribution of this *heralded* signal photon state is the $n = 1$ Fock-state contribution.

However, the requirement to operate in the low-gain limit is quite strict, as the ration of photon-pairs $|1\rangle_{k_s} |1\rangle_{k_i}$ to pairs of pairs $|2\rangle_{k_s} |2\rangle_{k_i}$ scales with λ . As such, SPDC-based single photon sources are usually quite inefficient in terms of the maximal rate of photons, which can be extracted from a source, until the source properties degrade (e.g. HBT-dip changes).

If the photon lifetime $\tau = 1/\Delta\omega$, which in the case of SPDC is given by the phase-matching bandwidth, e.g. the range of frequencies $[\omega_p - \Delta\omega_p, \omega_p + \Delta\omega_p]$ for which $\Delta kL = |\mathbf{k}_p - \mathbf{k}_s - \mathbf{k}_i| \ll 2\pi$, has a given value then the rate of photons must be small compared to $\Delta\omega$ (with a factor λ).

5.3.2 Bell-State Entangled-Photon-Sources

So far, we have only created a photon-pair and nothing else. Without loss of generality, we can assume that we have type-(I)—phasematching and we also without loss of generality we can assume that the phase matching is in H-direction, this we have:

$$|1\rangle_{k_s H} |1\rangle_{k_i H} = |00\rangle \quad (192)$$

state. This is of course rather similar to a computational basis state. However, this relation will in fact hold for an entire set of k -vector pairs (typically a cone).

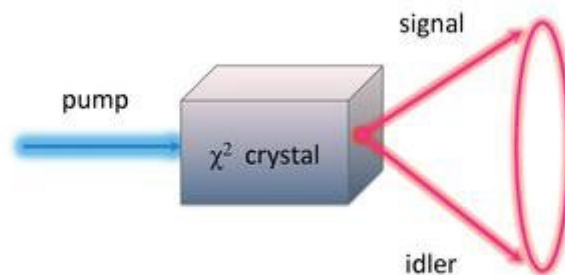


Fig. 18: SPDC in a thin crystal.

Assume now that the crystal is designed in such a way that it also supports SPDC for the mode of the type:

$$|1\rangle_{k_s V} |1\rangle_{k_i V} = |11\rangle \quad (193)$$

and that you have designed the input polarization of your pump-laser in such a way that both processes are equally likely. Due to the anisotropy of the crystal the H-polarized photons will experience a different dispersion relation and the phase-matching cone will be oriented differently in space. This means you have one cone, which emits $|00\rangle$ states and another cone, which emits $|11\rangle$ -states; but the two processes are distinguishable (e.g. by the propagation direction) and thus they cannot interfere into a joint quantum state.

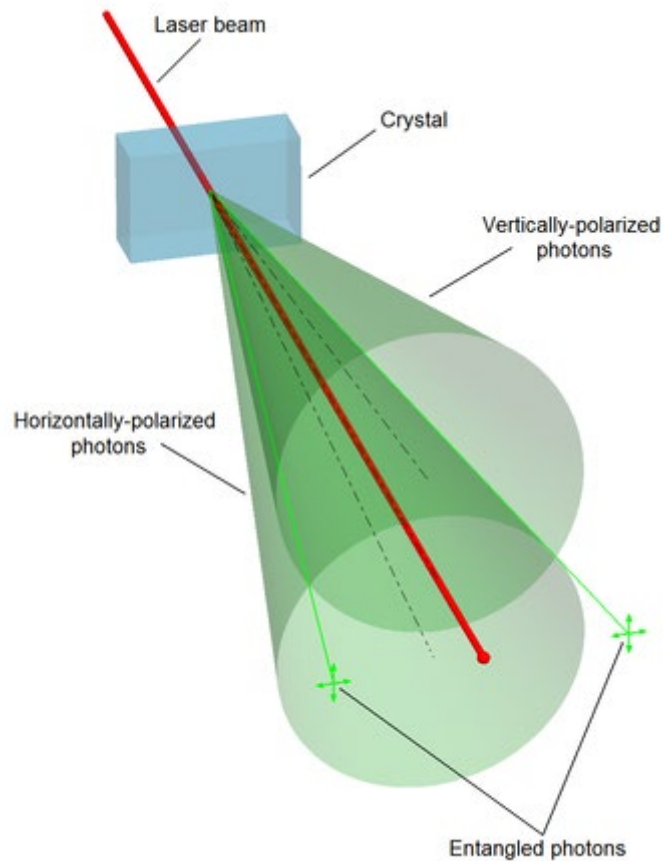


Fig. 19: SPDC in a nonlinear crystal; SPDC-cones and entanglement. Source: Wikipedia.

However, if the two cones intersect then both processes propagate in the same direction and hence a quantum superposition pair is formed, of the type:

$$\frac{1}{\sqrt{2}}(|1\rangle_{k_s H} |1\rangle_{k_i H} + |1\rangle_{k_s V} |1\rangle_{k_i V}) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\Phi^+\rangle \quad (194)$$

As this only happens for one particularly set of k-vectors you may now ignore the k-indices altogether. Now that we, in principle now how to use SPDC to generate entangled photon-pairs, I should make at least a few comments:

1. Other BS-states than $|\Phi^+\rangle$ be created by changing the dispersion and/or the type of phase-matching.
2. What the concept in essence means is: if you observe in the overlap direction and you get a photon pair, then you cannot know, from which cone the photon pair was coming from. Hence

All notes subject to change, no guarantee to correctness, corrections welcome.

you can't know if you get a HH or VV pair but since only these two processes are allowed, you know that you can get only those.

- The example here is about momentum-polarization entanglement. Any other two degrees of freedom of light can be entangled given an appropriate geometry and/or pump (time-energy; time-momentum; OAM-polarization; guided mode-wavelength). These may then require more complicated setups (e.g. structured crystals, waveguides) but may be interesting in their own right. Independent of the type of correlation it holds true, that they require at least two generations channels, which you cannot distinguish, to create an entangled pair.

5.3.3 Spontaneous Four Wave Mixing (SFWM)

SPDC-based processes require a material with non-vanishing $\chi^{(2)}$ coefficients. This is only the case for so-called non-centrosymmetric materials. For bulk-materials this is only the case for some special types of crystals, which rules out the SPDC-creation of photon-pairs in things such as glass-based optical fibers. However, we can instead utilize four-wave mixing, which occurs in all materials. This process is described by the $\chi^{(3)}$ coefficient.

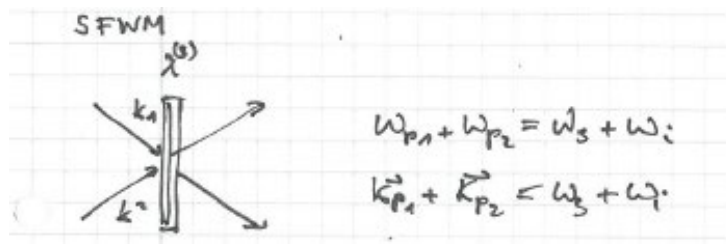


Fig. 20: SFWM in 3-rd order nonlinear materials.

Because a full description on the process is beyond the scope of this script, we shall just note a few important stepping stones here (note that p_1 and p_2 are pump photons and may in fact be the same mode)

- The phase matching conditions take the form $\omega_{p_1} + \omega_{p_2} = \omega_s + \omega_i$ and $\mathbf{k}_{p_1} + \mathbf{k}_{p_2} = \mathbf{k}_s + \mathbf{k}_i$
- The Hamiltonian now has a structure: $\hat{\mathcal{H}}_{FWM} \propto \chi^{(3)} \hat{a}_{k_s}^\dagger \hat{a}_{k_i}^\dagger \hat{a}_{k_{p1}} \hat{a}_{k_{p2}} + h.c.$, i.e. two pump photons are required in the process
- $\chi^{(3)}$ does occur in almost any material (even fibers)
- Involves two pump photons \rightarrow different scaling with input power

5.4 Characterizing PPS: The Hong-Ou-Mandel-Effect

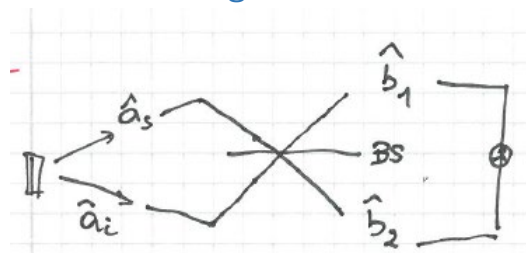


Fig. 21: HOM Interference on a Beam Splitter

Paul Dirac stated, that first-order interference effects can be thought of as each photon interfering with itself. Hong, Ou and Mandel experimentally showed that this is not the only form of interference

we can observe. Let's consider what happens when single photon states are incident from each of the two input ports of a 50:50 beam splitter. Their state is:

$$|\Psi\rangle_{s,i} = |1\rangle_s |1\rangle_i = \hat{a}^\dagger_s \hat{a}^\dagger_i |\text{vac}\rangle \quad (195)$$

If we then place a detector in each of the two output ports $\hat{b}^\dagger_1 \hat{b}^\dagger_2$, then there will be no simultaneous detections. To verify this, we can either calculate directly the correlation function for the input state $\langle \Psi | \hat{n}_{b_1} \hat{n}_{b_2} | \Psi \rangle$, or (faster) express the input state in terms of the detection modes, i.e. we replace

$$\begin{aligned} \hat{a}^\dagger_s &\rightarrow \frac{1}{\sqrt{2}} (\hat{b}^\dagger_1 + \hat{b}^\dagger_2) \\ \hat{a}^\dagger_i &\rightarrow \frac{1}{\sqrt{2}} (\hat{b}^\dagger_1 - \hat{b}^\dagger_2) \end{aligned} \quad (196)$$

Substituting these expressions into the input state, we find that the terms leading to a joint detection at detectors 1 and 2 will cancel

$$|1\rangle_s |1\rangle_i \rightarrow \hat{b}^{\dagger 2}_1 + \hat{b}^\dagger_1 \hat{b}^\dagger_2 - \hat{b}^\dagger_2 \hat{b}^\dagger_1 - \hat{b}^{\dagger 2}_2 |\text{vac}\rangle = |2\rangle_1 |0\rangle_2 - |0\rangle_1 |2\rangle_2 \quad (197)$$

Both photons will leave the beam splitter bunched into couples. As a result, there will be no coincident detection. This can be seen as destructive interference of transmitted and reflected photon pairs¹⁵, known as Hong-Ou-Mandel interference (see *PRL 1987, 59 2044*). HOM interference is a valuable tool in quantum information processing, and quantum optics – we will encounter it again at many instances.

After taking a closer look at the equations, we can also infer the HOM-effect from a handwaving explanation. Assume there are two photons, which are incident on a balanced beam splitter from its two-input port. There are in total four options, as each photon may or may not get reflected:

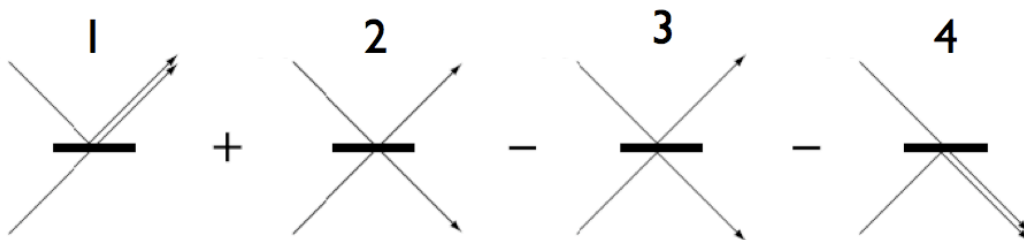


Fig. 22: Two-Photon Interference representation of the HOM-Experiment. If path 2 and 3 are indistinguishable, then they will interfere destructively and only photon pairs can be observed. Source: Wikipedia.

Note that the sign in front of the little pictures correspond to a 0 or π phaseshift. The phaseshift is particularly noteworthy for case 3, here they correspond to the phaseshift accumulated during reflection. A grossly simplified explanation is the reflection on a denser medium; nevertheless, the phase difference of π case 2 and 3 is universal, as it is related to the unitarity of the mixing operation.

Only indistinguishable photons show interference (that's why single photon interference is so easy to detect; one photon is necessarily indistinguishable from itself). In the HOM experiment this is only the case if the input photons are indistinguishable itself. As they interfere and have opposite sign, the modal contributions from option 2 and 3 thus cancel each other and the result of two-photon interference is such, that the two photons will either both go up or both go down (in the sense that they

¹⁵ Note that this is a consequence of the commutation operator relationships between the creation operators for bosons - what do you expect would happen if we replaced the photons with electrons ?

emerge in a superposition of 1 and 4; i.e. they go up and down simultaneously and their path is just decided up if you detect one photon). If you however detect one photon in, say, the upper branch then you know the other one is there as well and vice versa. Thus, if two indistinguishable photons meeting on a balanced beamsplitter they will leave the beamsplitter as a pair. If a pair is impinging on a beamsplitter they will go their separate ways. The beamsplitter can be considered something like the civil registry office for indistinguishable photon pairs.

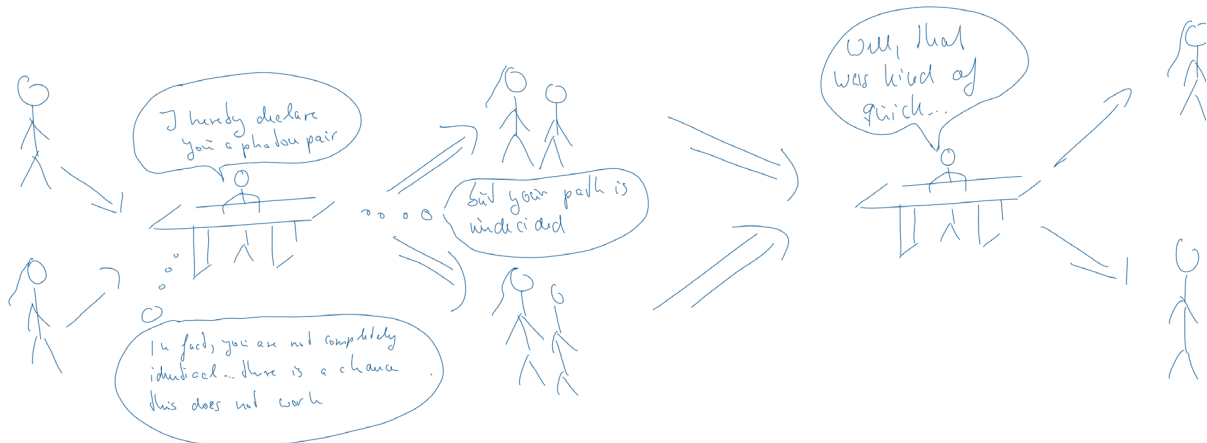


Fig. 23: The adventures of a beam splitter. Parts one and two. To pair and not to pair. Sorry for the silly joke. I could not resist.

The HOM-Interferometer is, however, not just a fancy quantum physical effect, but can be used to characterize properties of photon sources. One question you may ask is: are two consecutive photons of an SPS indistinguishable, i.e. does the source emit a PURE state? Or does it emit a mixed state, such that consecutive photons are this distinguishable? You can use the HOM to infer this question in such a way:

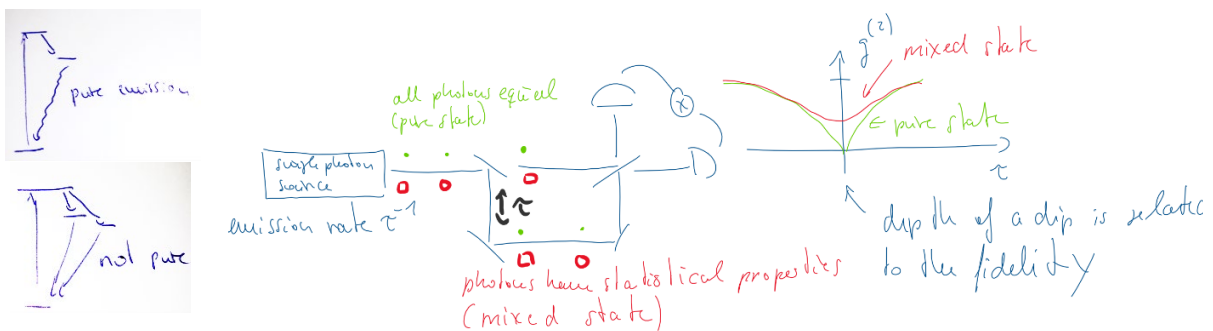


Fig. 24: (left) Examples of state diagrams of pure and mixed single photon emitters. (right) A HOM interferometer, which is used to test for the purity of a single-photon source. (green) A pure-state source, e.g. consecutive photons are indistinguishable and do thus interfere; (red) A mixed-state source, consecutive photons are distinguishable and thus do not interfere;

A HOM is also an important tool to characterize photon pair sources. Of such sources some come in the flavour of emitting indistinguishable photon pairs. Such sources are, e.g. important in Quantum Computing, more specifically in boson sampling. A HOM is here utilized in the following way:

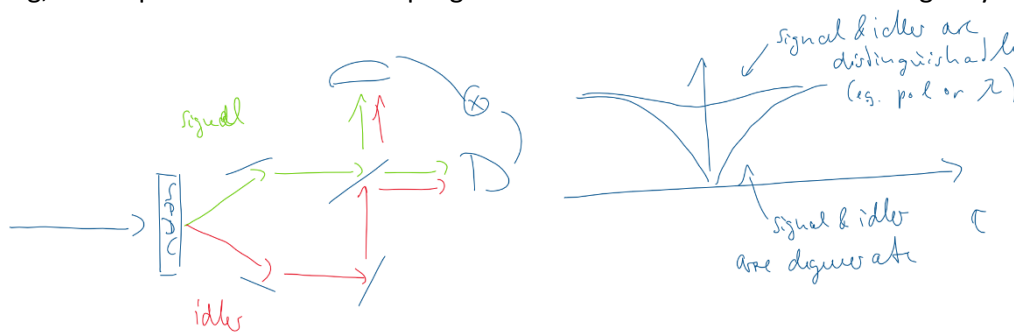


Fig. 25: A HOM interferometer, which is used to test for the indistinguishability of a photon pair source.

5.5 Measuring Photonic Bell-States

Many quantum protocols utilize Bell states as a measurement basis. This works, because Bell states are orthonormal and thus form a complete basis set of the 2-Qubit Hilbert space. As with many other things in Quantum Optics, such a measurement is conceptually simple, but in reality, difficult or next to impossible to implement. This chapter shall be devoted to discuss two of the most basic ideas, that we can use to implement an apparatus that is suitable to measure Bell states.



Fig. 26: Concept of a BS-Measuring device. Two Qubits enter the device and it tells you, which of the four BS the photons were in. If the photons have been in a superposition state of BSs then the resulting wavefunction will collapse into the one that corresponds to the measurement result.

The first approach is based on a very direct and brute-force mathematics inspired approach; which disentangles Bell states into CBS, which we can then easily measure. We shall see that this approach, however, requires the nonlinear interaction of Qubits. While this can be implemented in a Quantum Computer, it is generally not suitable for photons. For the case of photons, this approach is essentially the time reverse concept of SPDC and would thus require (efficient) sum-frequency generation of single photons; which is generally not feasible.

The second approach, which is the more feasible in quantum optics, is based purely on linear optical elements. It is a great simplification from an experimental point of view but this simplification comes at the expense of a reduced usability, that either materializes in a certain rate of error (the apparatus yields the wrong results) or a certain rate of failure (the apparatus yields no results) or a mixture thereof.

5.5.1 Measurement based on two-Qubit operations

We know that CBS and BS are both equally valid sets of basis vectors of the two-photon system. Thus, there should be a unitary operation, which converts between the two. It is therefore straightforward to attempt to use a CBS detector and have a unitary transformer \hat{U} before it. To do so, we must, however, introduce an important two-photon operation (A more in-depth discussion is frequently given in the context of Quantum-Computing, where these operators take the role of elementary quantum

gates, into which any quantum algorithm can be broken down into). The CX -gate (CNOT) applies a flip to Bob's Qubit if and only if Alice's is in state $|1\rangle$ ¹⁶

$$CX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (198)$$

The above-mentioned CBS to BS conversion can actually be written quite easily with these operations. It's composed by applying a CX -gate to Alice's photon with Bob's acting as the control photon and then a Hadamard-Gate (a simple beam splitter or a half-wave plate) to Alice's photon:

$$\begin{aligned} \hat{U} &= \text{HAD}_A \cdot CX_B \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \end{aligned}$$

I'll leave it as a homework problem to show, that this operation does indeed project the BS onto the CBS.

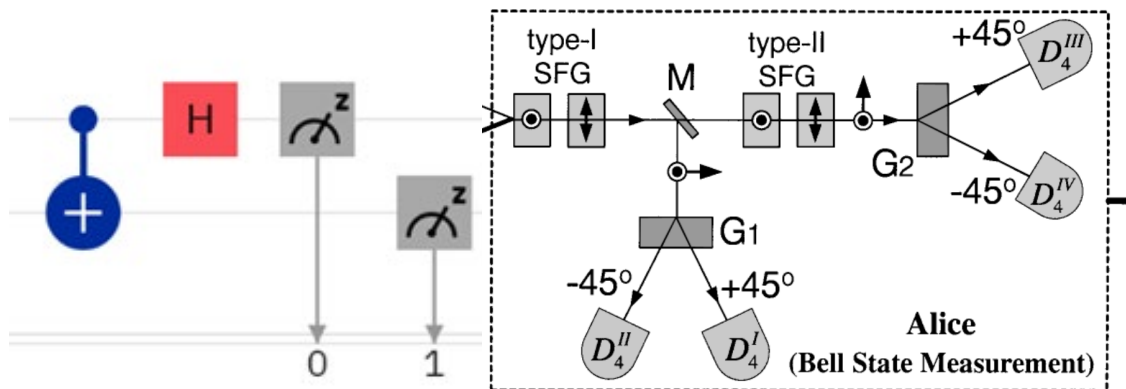


Fig. 27: (left) A CNOT-based BS measurement device in the quantum circuit representation. The single lines represent QuBits, the double lines classical bits. (right) A complete BS-measurement apparatus based on nonlinear optics from PRL 86, 1370 (2001). Note that most of the time the SFG-process does not occur and no measurement is recorded.

Two-photon gates, such as the CNOT, however, require photons to interact; hence nonlinear optics. Such gates typically operate at very low probabilities (success rates). so for practical applications we will have to resort to linear operations and skip the CNOT-approach. This will necessarily lead to an imperfect measurement.

5.5.2 Linear-optics based BSM

Thus, consider the following circuit, consisting of a 50:50 Beamsplitter, upon which both of the photons of the entangled pair are incident, two sets of polarization beam-splitters at the exit ports of the beamsplitter and 4 photon-number-counting detectors attached to a correlation setup.

¹⁶ The Matrices have to be multiplied with the Vector $[\alpha, \beta, \gamma, \delta]^\dagger$ to obtain the processed state vector $[\alpha', \beta', \gamma', \delta']^\dagger$

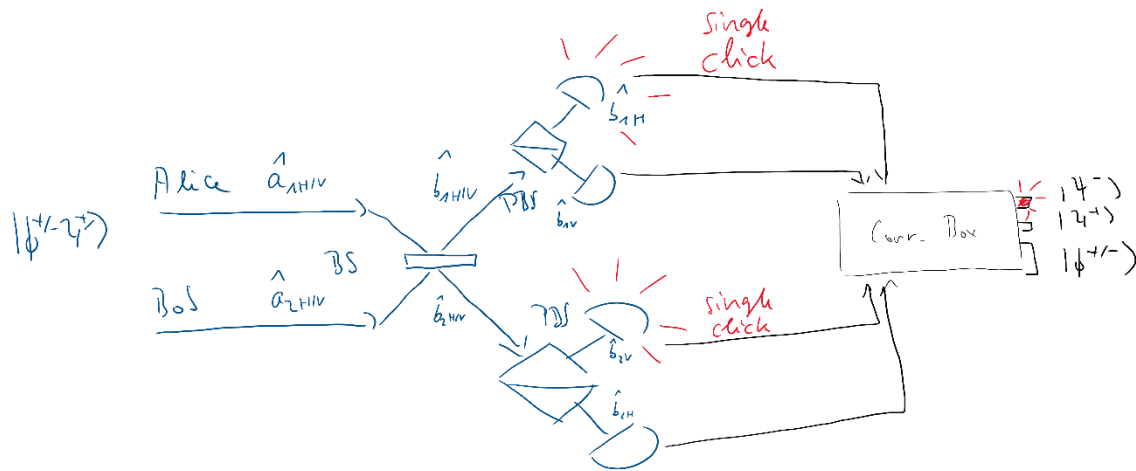


Fig. 28: A BS-PBS-type linear Bell State Measurement Scheme.

The non-polarizing beam splitter induces the following mode transformations¹⁷, which should be obvious after the discussion on the HOM interferometer:

$$\hat{a}_{1\lambda} = \frac{1}{\sqrt{2}}(\hat{b}_{1\lambda} + \hat{b}_{2\lambda}) \quad \hat{a}_{2\lambda} = \frac{1}{\sqrt{2}}(\hat{b}_{1\lambda} - \hat{b}_{2\lambda}) \quad (199)$$

Let's first consider the Bell-State $|\Psi^-\rangle = 1/\sqrt{2}(|01\rangle - |10\rangle)$, which is in terms of creation operators:

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(\hat{a}_{1H}^\dagger \hat{a}_{2V}^\dagger - \hat{a}_{1V}^\dagger \hat{a}_{2H}^\dagger)|\text{vac}\rangle \quad (200)$$

After the beam splitter the Photon pair is now in the state:

$$\begin{aligned} |\Psi^-\rangle &= \frac{1}{2\sqrt{2}} [(\hat{b}_{1H}^\dagger + \hat{b}_{2H}^\dagger)(\hat{b}_{1V}^\dagger - \hat{b}_{2V}^\dagger) - (\hat{b}_{1V}^\dagger + \hat{b}_{2V}^\dagger)(\hat{b}_{1H}^\dagger - \hat{b}_{2H}^\dagger)]|\text{vac}\rangle \\ &= \frac{1}{2}(\hat{b}_{1V}^\dagger \hat{b}_{2H}^\dagger - \hat{b}_{1H}^\dagger \hat{b}_{2V}^\dagger)|\text{vac}\rangle \\ &= 1/\sqrt{2}(|01\rangle - |10\rangle) \\ &= |\Psi^-\rangle \end{aligned} \quad (201)$$

This means that the $|\Psi^-\rangle$ -state is fully invariant under the operation of the 50:50 BS. We can apply the same set of operations to the other three Bell-States and obtain the following results for the action of the initial 50:50 beamsplitter on a Bell-State

$$\begin{aligned} |\Psi^-\rangle &\rightarrow 1/\sqrt{2}(|H, V\rangle - |V, H\rangle) \\ |\Psi^+\rangle &\rightarrow \frac{1}{\sqrt{2}}(|HV, 0\rangle + |0, HV\rangle) \\ |\Phi^+\rangle &\rightarrow \frac{1}{2}(|2H, 0\rangle - |0, 2H\rangle + |2V, 0\rangle - |0, 2V\rangle) \\ |\Phi^-\rangle &\rightarrow \frac{1}{2}(|2H, 0\rangle - |0, 2H\rangle - |2V, 0\rangle + |0, 2V\rangle) \end{aligned} \quad (202)$$

Note, that we have resorted to the HV -notation to get a better understanding of the action of the polarizing beam-splitters to follow. Also note, that HV and $2V$ mean, that there are two photons of

¹⁷ Note that we must resort to state operator representation here, because the beam splitter is used with two photons and we cannot guarantee that on the output side there will only be superpositions of number states with $n = 0$ and $n = 1$. In fact we shall see this is not the case.

opposite or equal polarization in any arm of the output of the beamsplitter denoted by the modes \hat{b}_1^\dagger or \hat{b}_2^\dagger , respectively.

First note that $|\Psi^-\rangle$ and only $|\Psi^-\rangle$ does produce one photon in each arm of the output port of the beam-splitter. Thus, if any of $D_{1\lambda}$ and $D_{2\lambda}$ produce a correlated click, we certainly know that $|\Psi^-\rangle$ has been measured. $|\Psi^+\rangle$ is the only state in which both photons are always bunched in the same exit port of the beamsplitter but in different polarization states. Thus if D_{1H} and D_{1V} or D_{2H} and D_{2V} show correlated click, we know that $|\Psi^+\rangle$ has been detected.

$|\Phi^+\rangle$ and $|\Phi^-\rangle$ are different though. They are both marked by double-clicks (i.e. the simultaneous detection of two photons) of any of the single detectors. Although they are orthogonal, this orthogonality is entirely related to the mutually opposing phase of the vertically-polarized contributions. As phase and photon-number are mutually exclusive measureables (they do not commute), we cannot hope to distinguish them with our setup or any kind of advanced photon-counter.

| Bell-State | Detector Correlation Signature |
|--------------------------------------|--|
| $ \Psi^-\rangle$ | single photons at D_{1H} and D_{2V} or D_{1V} and D_{2H} |
| $ \Psi^+\rangle$ | single photons at D_{1H} and D_{1V} or D_{2H} and D_{2V} |
| $ \Phi^+\rangle$ or $ \Phi^-\rangle$ | two photons at D_{1H} or D_{1V} or D_{2H} or D_{2V} |

Fig. 29: Correlation scheme for a BS-PBS-Type Bell-State-Measurement Setup.

This is a profound limitation. With a BS-PBS-type scheme, we can only hope to discern three of the four possible Bell-States. Moreover, one can show, that this is not a problem with the specific setup but that it actually is a limiting case. Lütkenhaus et al. (1999) have in fact shown, that no better result can be achieved using only linear optical elements and classical communication (and perfect photon-counting detectors)¹⁸.

This is really bad news and seems to jeopardize all of the Quantum-Protocols introduced earlier in this chapter. However, there are a few strategies on how to deal with this issue:

1. The setup we have investigated is perfectly correct, i.e. it never returns wrong results, but does not produce distinguishable results. Trade-offs can be made here. The most extreme case would be to just guess the Bell State. Which would give perfect distinguishability of all four states but result in only 25% correct results. Intermediate solutions with more meaningful trade-offs are possible. Erroneous results must then be compensated at a high level of the experiment.
2. Post-Selection can be applied, such that only those photon-pairs are used in the quantum experiment, which are distinguishable, i.e. all results attributed to $|\Phi^+\rangle$ and $|\Phi^-\rangle$ could just be discarded. This results in additional loss of 50% (i.e. ~ 3 dB), which may not be so bad as compared to transmission and other losses.
3. Perfect Bell-State-Measurements can be applied if the LOCC-requirement is broken. This is particularly true for the "L". Nonlinear optical setups can be utilized to create perfects BSM-setups. This, however, require Bob's and Alice's photon to interact nonlinearly. As the nonlinear interaction rates of single photons are extremely weak, this typically leads to non-interaction and thus a failed measurement for the vast majority of the photon-pairs. The trade-off is effectively similar to the ones above: one sacrifices success-rate for distinguishability.

¹⁸ This requirement is commonly termed „LOCC“.

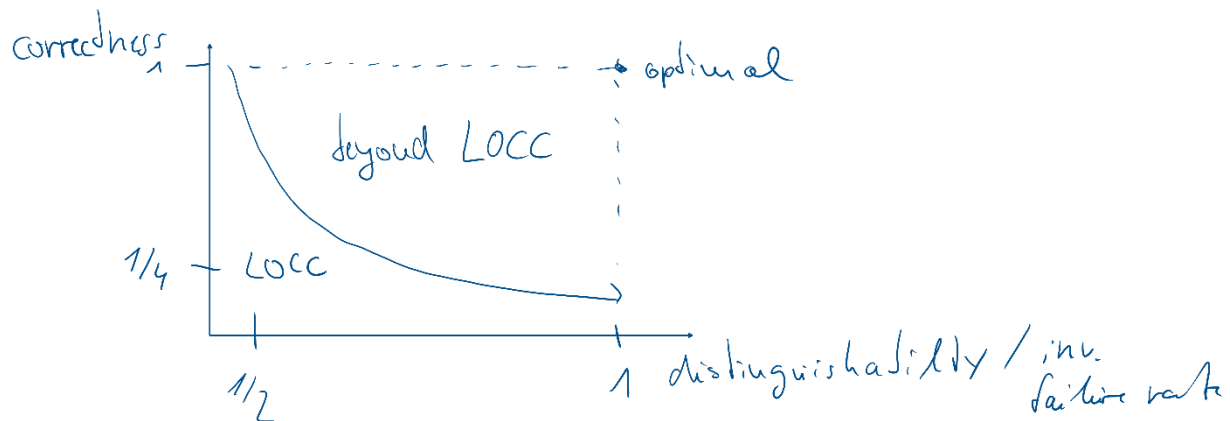


Fig. 30: The BSM-dilemma. Using LOCC you can either get correct or meaningful results. Never both. An optimal BSM requires nonlinear optics or non-classical communication (which by itself requires BSMs).

6 EPR and the Bell-Inequalities

Before we discuss quantum communication protocols, we shall first revisit Entanglement itself. We have seen that Entanglement is a natural property of multi-body quantum system, which occurs, if you apply the concept of superposition and wave-collapse onto them. As such, it is a direct result of the nature of quantum physics.

Its consequences are, however, subtle and have been a source for heated debated among the founders of the field: which culminated in the famous work by Albert Einstein, Boris Podolsky, and Nathan Rosen in 1935 and the reformulation of the Problem in terms of the Bell inequalities by John Bell in 1964 and later experimental demonstrations thereof, the first one in 1982 by Alain Aspect and then many later, which did close some existing loop-holes.



The core of the problem has already been introduced in Chapter 4.1. Some of these two-particle states actually behave in a way that Alice's measurement immediately determines the result for Bob's measurement (e.g. the Bell states). The argument, which now arose between Einstein and Bohr mainly did focus around the one question: if two particles are entangled, *are they still two individual particles or are they one unique quantum system?*

Think about it as you and your partner leaving the house in the morning, going to work in opposite directions, each grabbing one wallet in the dark. Yours is red and your partner's is green. Once you arrive at work, you can have a look at the wallet and see if it's yours or your partner's and you'll immediately know what your partner's measurements result is (the opposite of yours). The same applied to the coloring. But you'll also know that the wallets have been the same all along. For entangled quantum systems, we'll see that the latter is not the case and we'll see how this can be cast into measurable equations. More specifically even if you measure a superposition of color and owner, you'll still retain the connect between the results.

6.1 EPR'S Arguments on the Nature of Nature

The EPR argument basically claimed that Quantum Physics (at the least Copenhagen Interpretation) must be incomplete. They put forward the axiomatic claim, that any physical theory must be local and real. These words deserve some clarification:

Physical Realism: All properties of an object may be measured, such that the system may be fully determined by a the measurement conducted by the measurements of an (possibly infinitely skilful) observer. Any measurement does not affect the measured object, it merely reads off a property. The property itself is always determined and is fixed before and without the measurement.

Locality: Any object may only influence its immediate surrounding. Any action may propagate from the surrounding no faster than the speed of light. Thus, there can only be a causal relation between events in space-time that are separated space-like; not time-like. i.e. events that are within each other's light cone. Action at a distance (outside of the light cone) is impossible.

Keep in mind that we have in this lecture introduced plenty of concepts that are incompatible with these ideas. The former is violated by the concept of wave-function collapse and non-commuting measurements, whereas the latter is violated by entanglement. The EPR argument then basically suggest that because the Quantum Physics violates Realism and Locality it must therefore be **incomplete**.

Incompleteness means that there is something missing, there are some hidden variables, which we don't know (and may not have access to), but which predetermine the outcome of any quantum mechanical measurement and merely appear like quantum mechanical randomness to the (non-skilled) observer.

The entire argument can be boiled down to a simple thought experiment. Assume that we have a biphoton $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ state. Also assume that we take one of the photons of this state and send it to Alice and the other one to Bob and they both measure the polarization of the photon in the H/V -Basis. Also assume that their labs are sufficiently far away, such that there cannot be a causal relation between the two within the time-frame of the measurement. Both may either measure H or V . But they always measure the same polarization, no matter what. Assume Alice is the first to measure (which for time-like measurement events is actually not a meaningful statement) and assume she measures H , how does Bob's photon know it is supposed to also create an H -reading on Bob's detector?

The inventors of the Quantum Theory claimed that the laws of quantum physics create this correlation. Full stop. That's just the way it is, not matter if Einstein likes it or not. Action at a distance is very real and quantum physics (and therefore nature itself) thus cannot real AND local at the time time. The first measurement has in fact *determined (defined)* the state of both individual photons non-locally and that state simply did not exist before the measurement.

EPR argued that for locality and reality to hold, the must be hidden variables. Both Alice's and Bob's photon must have been imprinted at their common point of creation with the information that they will both eventually will create an H -result. They carried this information via an unknown (and possibly unknowable) hidden variable to both detectors. The detectors merely *measured (revealed)* the content of the hidden variables, its values had been fixed all along.

Note that because any property of any quantum system can be entangled, this argument basically expands to the full state of nature. We all know how the story went: the existence of proposed hidden variable was eventually disproven in experiment. Keep in mind however, that the EPR paper is still an

awesome work of science, because it strips down the conceptual differences between classical theories and quantum theories to two very basic concepts. It is a feat of trying to understand and disprove the arguments of the opposing side in the most honest way possible.

6.2 Bell's Inequalities

Funnily enough it took almost thirty years for John Bell to reformulate the arguments of EPR in a quantifiable manner. He was the first to realize that the existence of (unknowable) hidden variables is not a matter of taste, which you can believe in or not, but that their sheer existence leads one to predict different results for a certain type of correlation measurements on entangled system from those results predicted by quantum theory. This feat would later allow scientists to test in an experimental way, if EPR has indeed been right or if the world is in fact an unreal and/or non-causal place to live in.

The argument hinges on EPR's thought experiment but with a slightly modified setting. Namely, Alice and Bob now measure the polarization of each photon of the $|\phi^+\rangle$ -pair along three different angles off the H -basis ($\phi_a = 0^\circ$ and $\phi_b = \alpha$ and $\phi_c = \frac{\pi}{2} - \alpha$) where they select the measurement basis randomly and independently of each other for every individual photon. For any measurement they get one of two states, which we shall still call $|0_i\rangle$ and $|1_i\rangle$ with $i = a, b, c$ to denote that they are different states for each setting of the polarizer. See section 4.2. for the precise definition and the mathematical formulation of the measurement process.

The Bell-inequalities are now particularly concerned with the correlation probabilities, which can be determined after many repeated measurements. One example is $p(A = 0_a, B = 0_b)$, which is the fraction of all measurements were Alice has measured 0 in the a -Basis (with an α_a rotated polarizer) and Bob has measured 0 in the b -Basis (with an α_b rotated polarizer). Keep in mind that because Alice and Bob set their polarizers to randomly selected bases all the time, we may switch between probabilities p and measurement numbers N .

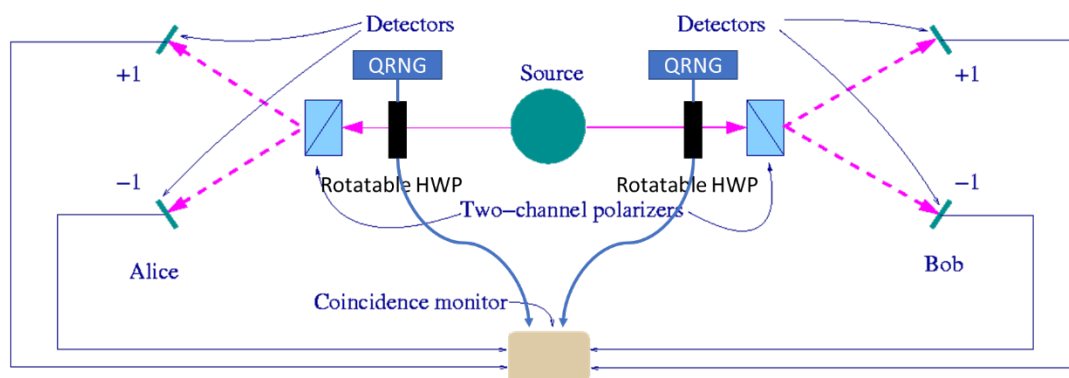


Fig. 31 Experimental Scheme of the Bell-Test, which is used to validate or invalidate Bell's inequality. Source: Wikipedia.

Let's stick with the German Wikipedia for a while and switch to code words. For the different measurement bases and results (0 or 1, respectively). The results in Basis a are termed 0_a =short, 1_a =tall, in basis b 0_b =blond, 1_b =dark, and in basis c 0_c =female, 1_c =male.

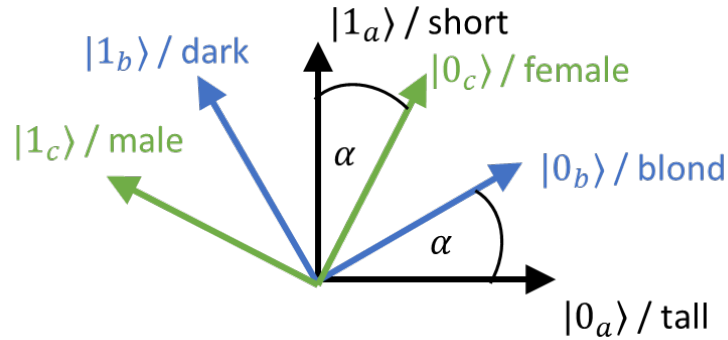


Fig. 32 Selected bases for the polarization measurement in Alice's and Bob's lab for a Bell test. The words next to the results are used as code-word for ease of interpretation.

In the hidden-variables explanation each photon of the entangled pair corresponds to a twin of the other partner, e.g. they may both be short, blond females or both be short, dark males. Alice and Bob may however only ask each twin one single question and they thus can measure two properties of the pair. E.g. size and color of hair.

For combinatorial reasons the number of tall, blond photons can be measured if Alice measured in basis a and Bob measures in basis b or vice versa. Its number is given by $N(A = 0_a, B = 0_b)$. It must be the same as the sum of tall, blond females plus the number of tall blond males:

$$N(A = 0_a, B = 0_b) = N(A = 0_a, B = 0_b | c = 0) + N(A = 0_a, B = 0_b | c = 1) \quad (203)$$

The first summand of the right-hand side will each not shrink, if we leave out the second condition ($B = 0_b$) for the left summand, e.g. we ignore or just not carry out the measurement on color of hair by Bob in the first summand. Because Bob is now basically unemployed, we can now use him to test for the gender instead.

$$N(A = 0_a, B = 0_b | c = 0) \leq N(A = 0_a | c = 0) = N(A = 0_a, B = 0_c) \quad (204)$$

The same can be said for the second summand if the Alice's measurement size is ignored and she measures gender instead. We can now also switch the roles of Alice and Bob, this will not change anything because the measured objects are identical twins:

$$N(A = 0_a, B = 0_b | c = 1) \leq N(B = 0_b | c = 1) = N(A = 1_c, B = 0_b) = N(A = 0_b, B = 1_c) \quad (205)$$

The entire formula thus reads as

$$N(A = 0_a, B = 0_b) \leq N(A = 0_a, B = 0_c) + N(A = 0_b, B = 1_c) \quad (206)$$

Thus, if hidden variables really exist (doesn't matter if they are measurable or not) this inequality must always hold.

Now we return to photons, for which we know quantum mechanics holds true and see if this relation is actually fulfilled. For a sufficiently large number of photons the terms in the last equations can be easily interpreted as probabilities and also as transmission ratios of particular filter settings. $N(A = 0_i, B = 0_j)$ describes an event, in which a photon passed a polarizer rotated by α_i in Alice's Lab and a polarizer with α_j rotation in Bob's lab. As the initial state was a $|\phi^+\rangle$ state we know that that

the probability for such an event is simply $\cos^2(\phi_i - \phi_j)$. The same is true for $N(A = 1_i, B = 1_j)$. The result for $N(A = 0_i, B = 1_j)$ and $N(A = 1_i, B = 0_j)$ is likewise $\sin^2(\phi_i - \phi_j)$.

The above inequality thus reads as:

$$\cos^2(\phi_a - \phi_b) \leq \cos^2(\phi_a - \phi_c) + \sin^2(\phi_b - \phi_c) \quad (207)$$

If we assume the situation as discussed above, namely $\phi_a = 0$, $\phi_b = \alpha$, and $\phi_c = \frac{\pi}{2} - \alpha$, this simplifies to

$$\cos^2(\alpha) \leq \sin^2 \alpha + \cos^2(2\alpha)$$

Which is obviously NOT true, unless $\alpha = 0$ or $\alpha = 45^\circ$. It is “most untrue” for $\alpha = 30^\circ$ with a difference of the LHS and RHS of 0.25.

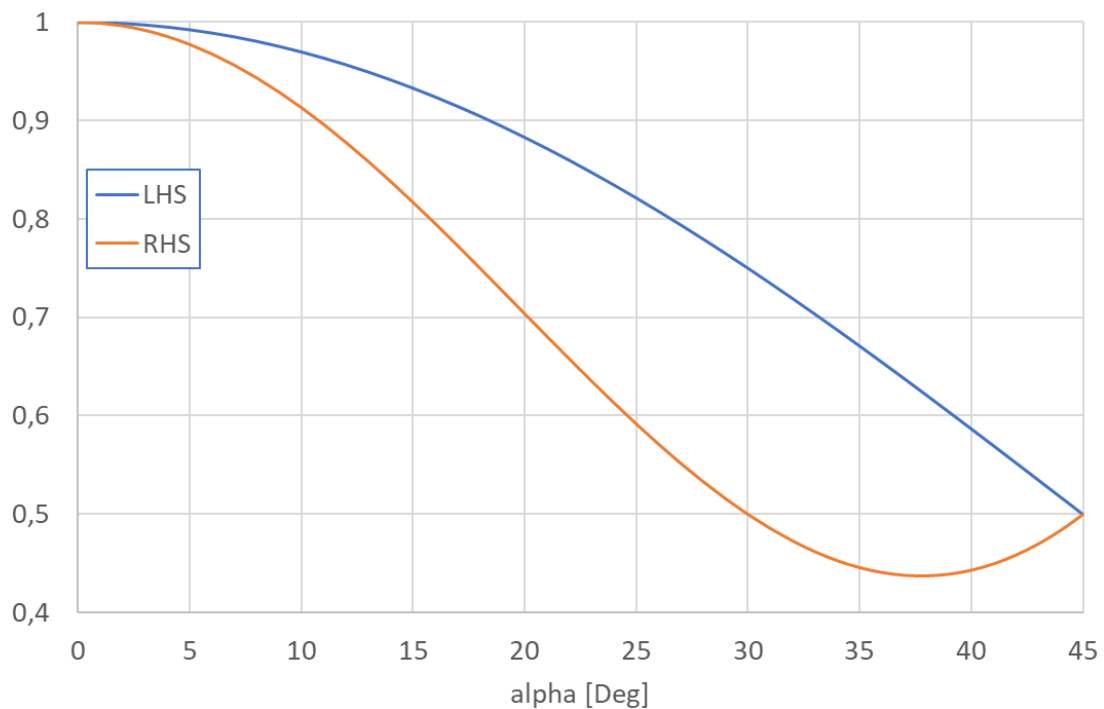


Fig. 33 Comparison of LHS and RHS of Bell inequalities for the quantum case. Classically the RHS should be equal or LARGER the LHS.

This means that we can guarantee (except for very specific angles) that the results we will measure in fact do violate the inequality, which we derived from the assumption of a hidden-variable theory. This is a deeply profound finding. There is a measurable difference between a theory with and without hidden variables and this difference can be tested for experimentally by evaluation of Bell’s inequalities (you can just MEASURE both sides of the inequality and if you find that your results violate it then, voilà). Should they hold then we know that Quantum Physics is indeed incomplete and hidden variables do exist. Should they not hold, then we can rule out hidden variables and we know that nature is not real and/or not causal.

6.3 A generalization: CHSH-Inequalities

Practical Bell-tests actually do not implement the scheme as proposed by Bell but they rely on a generalization of Bell’s scheme put forward by Clauser, Horne, Shimony, and Holt five years after Bell’s seminal breakthrough. The CHSH inequality is – as Bell’s original inequality – a constraint on the coincidences in a correlated set of measurements on a pair of entangled photons, which is smaller for a

classical (hidden-variables) geometry and larger for a quantum theory. To cut a long story short, the CHSH test also measures, if quantum particles are more strongly correlated than classical particles ever could be.

The CHSH inequality is again concerned with correlation measurements $E(i, j)$ of two different properties of the constituents of an entangled system, e.g. the results of an outcome of a polarization measurement where Alice measures along two different bases $i = a_1, a_2$ and Bob along $j = b_1, b_2$, each of which have eigenstates $|\pm_{i/j}\rangle$ and eigenvalues ± 1 or just \pm in shorthand notation. Also note that we briefly deviate from the established notation for the eigenvalues here, for ease of calculation of the number, i.e. $|+\rangle = |0\rangle$ and $|-\rangle = |1\rangle$.

Experimentally these correlations $E_{\text{Exp}}(i, j)$ are determined by long-term averages of the number of simultaneous clicks in Alice's and Bob's detectors for a fixed selection of bases i, j (of which there are four possible combinations). Here $N_{\pm\pm}^{i,j}$ denotes correlation, i.e. simultaneous clicks at both + for or both - detectors in Alice and Bob's labs. Likewise, $N_{\pm\mp}^{i,j}$ denotes anti-correlation, i.e. where one detector clicks at + and the other one at -, e.g. Alice and Bob record different values. The correlation of particle pairs for such a setting is then given by:

$$E_{\text{Exp}}(i, j) = \frac{N_{++}^{i,j} - N_{+-}^{i,j} - N_{-+}^{i,j} + N_{--}^{i,j}}{N_{++}^{i,j} + N_{+-}^{i,j} + N_{-+}^{i,j} + N_{--}^{i,j}} \quad (208)$$

A CHSH-type Bell test does then measure the quantity:

$$S = E(a_1, b_1) - E(a_1, b_2) + E(a_2, b_1) + E(a_2, b_2) \quad (209)$$

What we'll now see, is that we get different upper bounds for S if we calculate it according to the rules of Quantum mechanics and accordingly to the assumption of hidden variables.

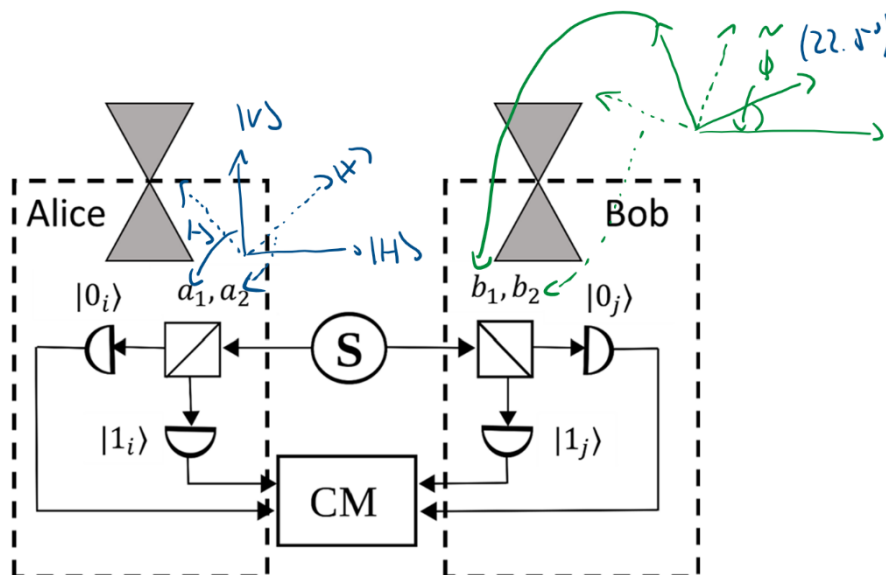


Fig. 34 Schematic of a two-channel Bell-test with polarization splitters using two time-like separated observer Alice and Bob. The test has to be re-run at least four times for all combinations of two different settings A_a, A_b of the left and B_c, B_d of the right polarizer. The strongest deviation for the classical prediction of $S = 2\sqrt{2}$ can be found at $a=0^\circ, b=45^\circ$ and $c=22.5^\circ, d=67.5^\circ$ (Tsirelson's bound)

According to the rules of quantum mechanics the correlation terms $E_Q(i, j)$ can be calculated from the system's biphoton state $|\psi\rangle$ using the expectation value

$$E_Q(i, j) = \langle \psi | \hat{\sigma}_i^a \otimes \hat{\sigma}_j^b | \psi \rangle = \langle \hat{\sigma}_i^a \hat{\sigma}_j^b \rangle \quad (210)$$

where $\hat{\sigma}_i^a$ denotes the Polarization measurement-operator for the appropriate measurement direction a_i acting on the photon in Alice's lab. $\hat{\sigma}_j^b$ is likewise b_j acting on the photon in Bob's lab. Here i and j denote the specific orientation of the polarization beams splitter. Without loss of generality, we assume that all measurements are in a linear polarization basis. Thus, the Polarization measurement operator can be obtained from the set of Pauli-matrices $\hat{\sigma}_{1,2,3}$ and the measurement direction vector $\vec{a}_i = [\cos 2\phi_i^a, \sin 2\phi_i^a, 0]$ denoted by the rotation angle ϕ_i for measurement basis a_i according to

$$\hat{\sigma}_i^a = [\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3] \cdot \vec{a}_i^\dagger \quad (211)$$

The same holds true for the $\hat{\sigma}_j^b$ operators. For example, if the measurement is oriented along the x -direction, e.g. we are measuring H/V-polarization we get $\hat{\sigma}_i^a = \hat{\sigma}_1$. Because S is composed of a sum of expectation values, we can introduce an appropriate measurement operator composed of the sum of four separate measurements $\hat{\sigma}_i^a \otimes \hat{\sigma}_j^b$, which will have the same expectation values, simply as

$$\hat{S} = \hat{\sigma}_1^a \hat{\sigma}_1^b - \hat{\sigma}_1^a \hat{\sigma}_2^b + \hat{\sigma}_2^a \hat{\sigma}_1^b + \hat{\sigma}_2^a \hat{\sigma}_2^b \quad (212)$$

We note, that the outcome of the measurement does not depend on the order that Alice and Bob make their measurement, because Alice's and Bob's measurement operators commute: $[\hat{\sigma}_i^a, \hat{\sigma}_j^b] = 0$, irrespective of their specific basis choice. Using this we can make use of the Khalfin-Tsirelson-Landau identity for the square of the CHSH-Operator \hat{S} .

$$\hat{S}^2 = 4\mathbb{I} - [\hat{\sigma}_1^a, \hat{\sigma}_2^a] [\hat{\sigma}_1^b, \hat{\sigma}_2^b] \quad (213)$$

The classical notion of hidden variables entirely hinges on the idea, that all quantities be determined and thus any classical theory must provide for $[\hat{\sigma}_1^a, \hat{\sigma}_2^a] = [\hat{\sigma}_1^b, \hat{\sigma}_2^b] = 0$. We thus get as the classical bound:

$$S_c \leq \sqrt{\langle \hat{S}^2 \rangle} = 2 \quad (214)$$

In quantum physics, however, the commutators do not vanish. We note, that $(\hat{\sigma}_i^a)^2 = (\hat{\sigma}_j^b)^2 = \mathbb{I}$, which can be easily seen from the fact that their eigenvalues are ± 1 . We may use this fact to derive an upper bound for the commutators, as $\|\hat{\sigma}_1^a, \hat{\sigma}_2^a\| \leq 2\|\hat{\sigma}_1^a\| \|\hat{\sigma}_2^a\| \leq 2$. The same is true for Bob's commutator. Thus, we get for the expectation value of their product: $-4 \leq \langle [\hat{\sigma}_1^a, \hat{\sigma}_2^a] [\hat{\sigma}_1^b, \hat{\sigma}_2^b] \rangle \leq 4$. The CHSH-measurement is maximized for the case of -4 and we get:

$$S_Q \leq \sqrt{\langle \hat{S}^2 \rangle} \leq \sqrt{4 - (-2)2} = 2\sqrt{2}. \quad (215)$$

Thus, any value of $2 < S \leq 2\sqrt{2}$ is a clear indicator of non-classical behaviour. However, these values can only be reached if $\langle [\hat{\sigma}_1^a, \hat{\sigma}_2^a] [\hat{\sigma}_1^b, \hat{\sigma}_2^b] \rangle \approx -4$. In practice this can be enforced by an appropriate choice initial states $|\psi\rangle$ and of the measurement bases a_i and b_j (which also means that not all choice of measurement bases allow for stronger than classical correlations!). For the sake of simplicity we restrict ourselves to $|\psi\rangle = |\phi^+\rangle$ and linear polarization states for the measurement. It is then further helpful to choose the respective measurement basis in each lab to be rotated by 45° . Thus, without loss of generality we chose for Alice to measure in linear polarization $|H/V\rangle$ and in diagonal polarization $|+/-\rangle$:

$$\phi_1^a = 0, \phi_2^a = \frac{\pi}{4} \quad (216)$$

The measurement bases for Bob are then simply rotated by a given angle $\tilde{\phi}$ with respect to Alice's bases, i.e.

$$\phi_1^b = 0 + \tilde{\phi}, \phi_2^b = \frac{\pi}{4} + \tilde{\phi} \quad (217)$$

It can be shown that the expectation value S_Q can then be maximized at $\tilde{\phi} = \frac{\pi}{8}$ and at this values we get exactly $S = 2\sqrt{2}$. This result is called Tsirelson's bound and the respective basis angles $0, 45^\circ$ and $22.5^\circ, 67.5^\circ$ are typically referred to as the Bell-Test angles.

6.3.1 CHSH on a Quantum Simulator

Since the Bell Test / CHSH experiment is such a milestone of Quantum Physics (keep in mind, it did not only settle a half-century old debate between titans of modern physics but it tell us something about the nature of the physical universe we live in, in general), we shall like to take the opportunity and experimentally verify its validity. We do so by quickly switching to the realm of quantum computation, but as we have discussed earlier; a Qubit is a Qubit. Keep in mind: the first demonstration of experimentally violating the CHSH inequality is merely forty years old and now we can repeat the experiment in a classroom.

We first construct a quantum circuit, that implements the CHSH-experiment with the appropriate setting for measurement angles:

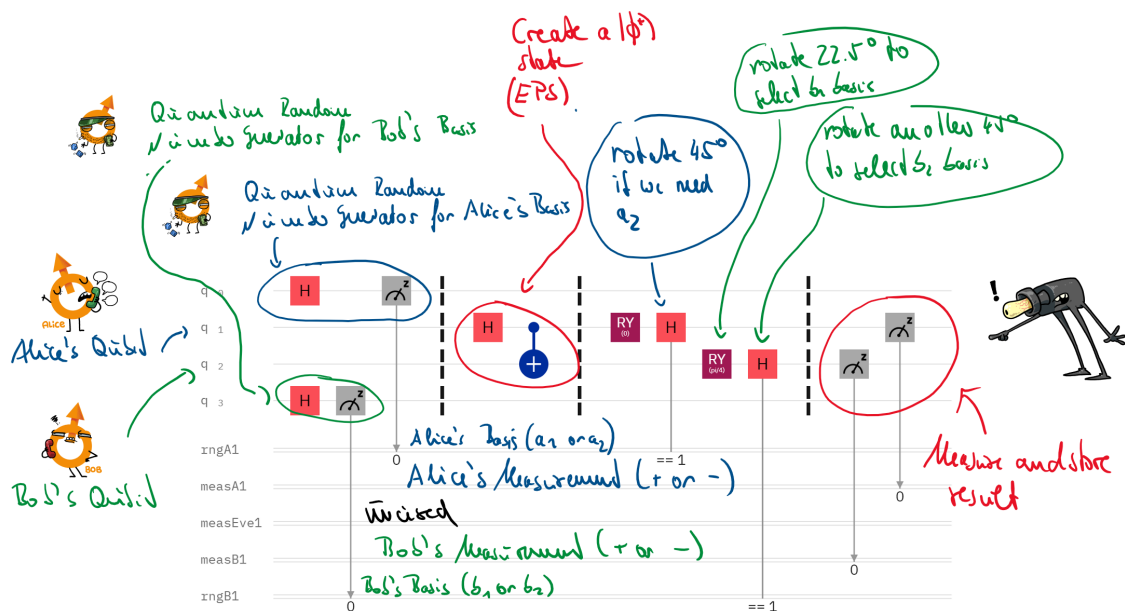


Fig. 35 A CHSH-Experiment with $\tilde{\phi} = 22.5^\circ$ implemented for a Quantum Computer. The basis selection is carried out via quantum random number generators for which two separate Qubits are consumed. Note that the measEve1 result line is unused in this experiment.

We then run the circuit (for the moment on a quantum simulator) and get the following results:

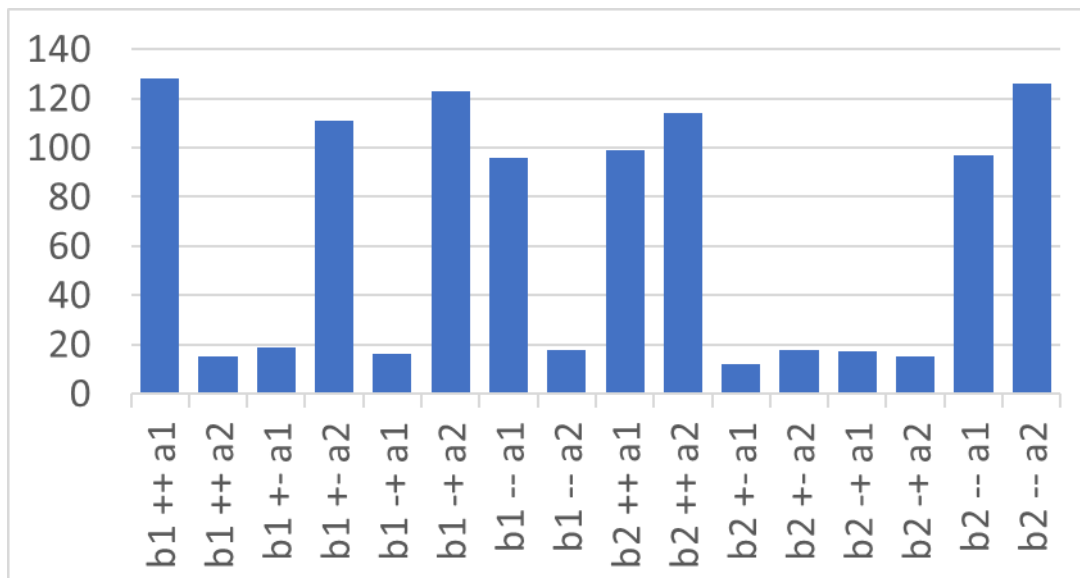


Fig. 36 Results for 1024 runs of the above displayed experiment. The x-axis labels are already modified from simple bitwise notation to the notation used in the formula above.

This representation is probably still a bit hard to really understand, therefore it makes sense to arrange the data in a table and use a bit of excel magic to mark up the results.

| Measuremer | Frequency | rngA | rngB | Weight | meas. | measB | Even |
|------------|-----------|------|------|--------|-------|-------|------|
| b1 ++ a1 | 128 | a1 | b1 | ● | + | + | ● |
| b1 ++ a2 | 15 | a2 | b1 | ● | + | + | ● |
| b1 +- a1 | 19 | a1 | b1 | ● | - | + | ● |
| b1 +- a2 | 111 | a2 | b1 | ● | - | + | ● |
| b1 -+ a1 | 16 | a1 | b1 | ● | + | - | ● |
| b1 -+ a2 | 123 | a2 | b1 | ● | + | - | ● |
| b1 -- a1 | 96 | a1 | b1 | ● | - | - | ● |
| b1 -- a2 | 18 | a2 | b1 | ● | - | - | ● |
| b2 ++ a1 | 99 | a1 | b2 | ● | + | + | ● |
| b2 ++ a2 | 114 | a2 | b2 | ● | + | + | ● |
| b2 +- a1 | 12 | a1 | b2 | ● | - | + | ● |
| b2 +- a2 | 18 | a2 | b2 | ● | - | + | ● |
| b2 -+ a1 | 17 | a1 | b2 | ● | + | - | ● |
| b2 -+ a2 | 15 | a2 | b2 | ● | + | - | ● |
| b2 -- a1 | 97 | a1 | b2 | ● | - | - | ● |
| b2 -- a2 | 126 | a2 | b2 | ● | - | - | ● |

Fig. 37 Same results as above but with a bit of excel-magic attached to visualize the data. The weight-column just shows, if this specific type of basis choice corresponds to the one E -term which has a negative sign. The even-column shows if the measurement was even (same detector clicked for Alice and Bob; either -- or ++) or odd (different detector clicked for Alice and Bob; either +- or -+).

The rest is really just book-keeping. We calculate each value for $E(i, j)$ by selecting the proper four rows from the table and make sure to add them up with the correct sign (according to the even-column!) in the denominator. We then add up all the $E(i, j)$'s and again make sure to use the correct sign (according to the weight-column). In this case we get an estimate of:

$$S = 2.98 > 2 \quad (218)$$

Don't you find this amazing? I do! We have found a clear violation of the CHSH inequality and thus have demonstrated that Quantum Physics is a reality, indeed!

Note that in this specific case we also get $S > 2\sqrt{2}$ but this is because we have used a measured number of occurrences to estimate expectation value $\langle S \rangle$. If you wanted to really get a serious fix on the violation of CHSH you'd have to rerun the experiment very frequently and then do proper statistics on the result. The distribution of estimated expectation values would then approach a Gauss shape with a center of gravity on $2\sqrt{2}$ and a width that is hopefully much smaller than $2\sqrt{2} - 2$.

6.3.2 CHSH for a Classical System/ Eavesdropping

Because I am feeling lucky (do you feel lucky, punk? Do you?) let's push the limits and see that we can't get beyond $S = 2$ for a classical system. To do so, we again use our quantum computer script from above albeit with a tiny modification. We will introduce an eavesdropper Eve, that attempts to make a copy of Alice's Qubit, like so:

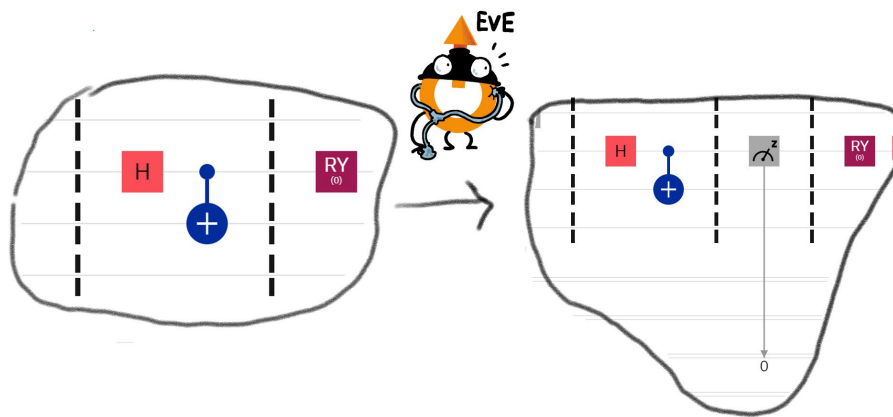


Fig. 38 We destroy the entanglement between Alice's and Bob's Qubit before their respective measurement (and basis selection) by introducing an eavesdropper "Eve". Eve conducts a measurement on Alice's Qubit and stores the result in the previously unused classical bit "measEve". Alice and Bob then proceed to measure in their respective bases and calculate $\langle S \rangle$.

Let's run the so-modified Quantum Circuit and we get:

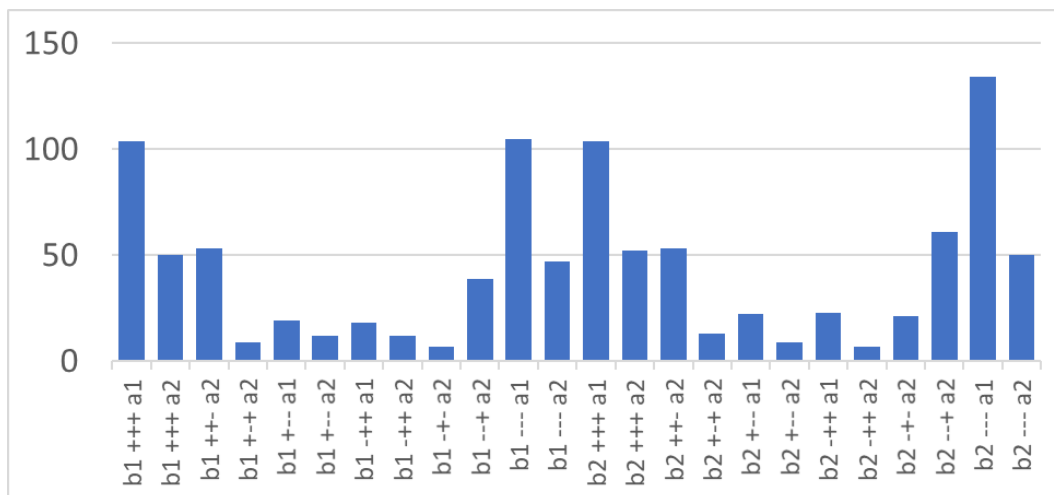


Fig. 39 Results for 1024 runs of the above displayed experiment. The x-axis labels are already modified from simple bitwise notation to the notation used in the formula above. The middle +/- corresponds to Eve's measurement result and is not used in the calculation of S .

The rest is again just book-keeping. We calculate each value for $E(i, j)$ by selecting the proper four rows from the table and make sure to add them up with the correct sign (according to the even-column!) in the denominator. We then add up all the $E(i, j)$'s and again make sure to use the correct sign (according to the weight-column). In this case we get an estimate of:

$$S = 1.42 < 2 \quad (219)$$

This means that we find a result that can be explained by classical correlations and hence Alice's and Bob's photons may not have been in a highly-correlated state of entanglement by the time they have been measured by Alice and Bob. This is an important train of thought later on for secret communication: if your communication is based on entangled photons, you can use a CHSH-test like this to check for non-classical correlation. If you find $S > 2$ then you are guaranteed that no one broke the entanglement and hence you can guarantee that no one has eavesdropped on your communication, therefore you are in a state of secrecy.

6.3.3 Beyond Tsirelson's Bound and a Real QC-Script

In the chapter above we had discussed that the "magic" Bell angles are the ones that achieve a maximum of $S = 2\sqrt{2}$. In this chapter I would like to discuss, what happens at different angles, i.e. we want to systematically vary $\tilde{\phi}$. Moreover in the last two chapters we have not run the measurement scripts on a real quantum computer but on a quantum simulator.

The simple reason for that is, that the QCs that I have access to do not support internal measurements and hence the idea of using a QRND to randomly select mutual measurement bases does not work. Instead we can just run, for every value of $\tilde{\phi}$ four difference script according to the combinations $a_1b_1, a_1b_2, a_2b_1,$ and a_2b_2 and determine respective $E(i, j)$ separately. This is in fact a little bit problematic because we run into a possible loophole (discussed below) but let's not worry about this too much for the moment.

I will not give you the script for this specific quantum experiment, because if could not do it better nor more elegantly as the solution found in the QISKIT book.

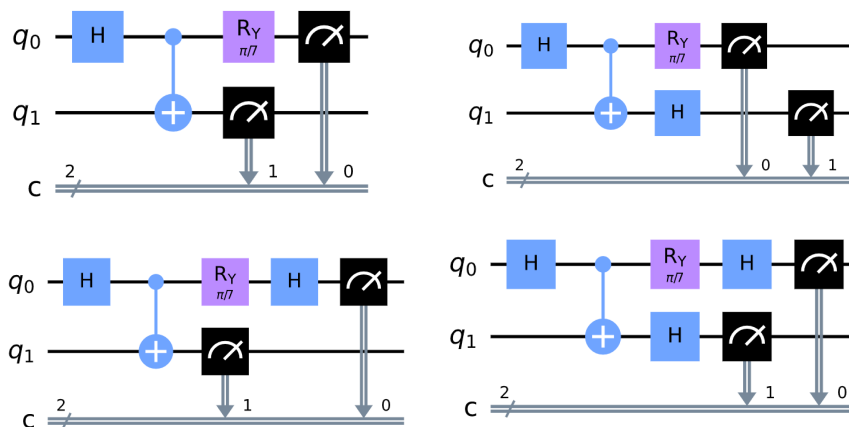


Fig. 40 A quadruplet of circuits for a CHSH-Test $\tilde{\phi} = \pi/14$. (From top left to bottom right): $a_1b_1, a_1b_2, a_2b_1,$ and a_2b_2 . These four circuit are run sequentially, one value of $E(i, j)$ is computed from each circuit. The four values are compiled into an estimate for $S(\tilde{\phi})$.

The resulting values for $S(\tilde{\phi})$ are found to be:

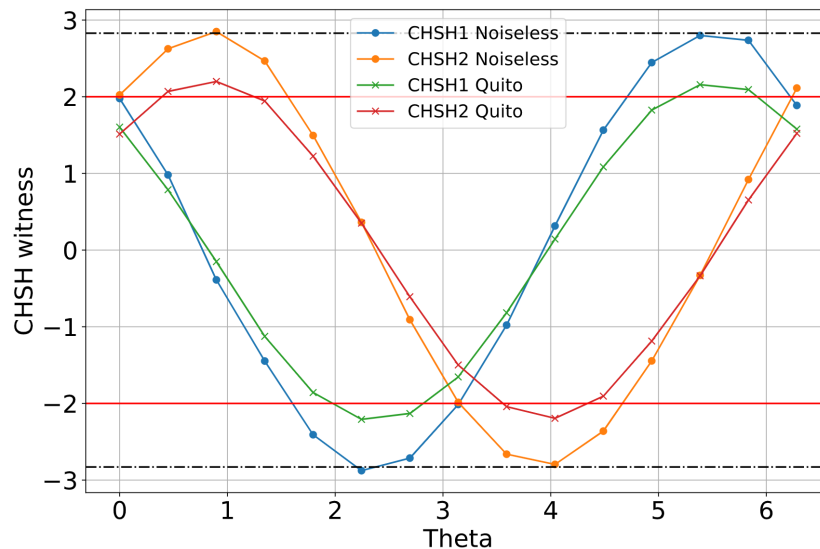


Fig. 41 $S(\vec{\phi})$ for a quantum simulation (noiseless) and measured on a real quantum computer (Quito). Note that the angle is called θ in this experiment. The red lines mark the classical limits. Everything beyond the red lines is quantum and disproved hidden variables. Note that real QCs are inherently noisy and thus do not reach far beyond $S = 2$. Also note that there is an ambiguity in the definition of the equation for S (the position of the “-“ sign). Here both cases are determined. Source: QISKIT.

This shows two things. First it shows that Tsirelson’s bound is really reached for 22.5° and also that we can REALLY, REALLY, REALLY see non-classical correlation in a REAL quantum experiment, although the intrinsic noise for such systems limits the correlation strength somewhat, however it is still well beyond $S = 2$.

6.4 Experimental Validation and some notes on loopholes

The first measurement of a non-classical value for the CHSH-measure S have been carried out by Aspect, Dalibard and Roger in 1982. Critics, however, did come up a set of so-called “loopholes”. These loopholes question hidden assumptions or claim imperfections in experiments to lead to higher measures for S than allowed classically, without requiring quantum physics for an explanation.

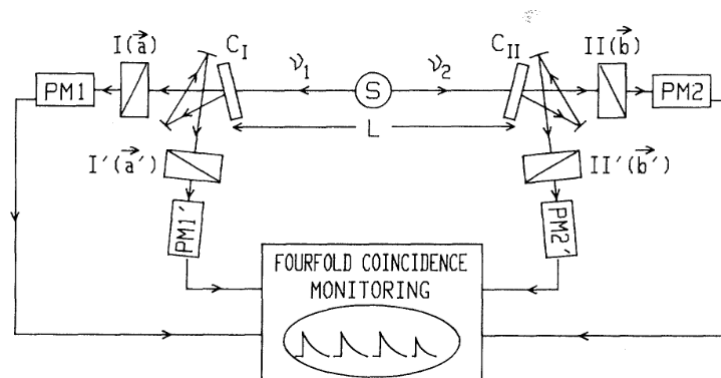


Fig. 42 Scheme of the Aspect-experiment, the first to successfully demonstrate $S > 2$. PRL **49**1804 (1982).

Here we shall only briefly discuss these loopholes and present some mitigation strategies, which may or may not have been used in experiments, so far. More specifically, some of the loopholes may be extended to the point of non-testability and we may have to resort to Okham’s razor to discount them.

Locality Loophole: Alice’s and Bob’s measurements have to be separated such that they cannot know of each other according to special relativity. While for the initial Aspect-paper this had been true for the measurement, this had not been true for the setting selection of the measurement basis, which

must be carried out late enough, such that a light-speed action may not propagate from Alice to Bob in time. Works by Zeilinger et al. closed this loophole.

Fair Sampling Loophole: In practice any detector has a quantum efficiency < 1 and thus only detects a fraction of the photons. This process must be “fair” in the sense that the detector must not detect correlated photon pairs with a higher probability than uncorrelated photon pairs. Initially this was indeed a big concern as detector did have QE in the range of 5% (modern ones can reach well beyond 50% or more, see chapter below). Rowe et al. did close this loophole by using a lasing medium that was triggered by the presence of an ion (no ion \rightarrow no lasing \rightarrow no light; ion \rightarrow lasing \rightarrow lots of light).

Freedom of Choice Loophole: This loophole disputes that independent and random selection of bases is possible. It can be shown, that if one assumes that a prior interaction of the random number generators (or their constituents) would induce correlations in their randomness mechanism a local hidden variable theory may be constructed. In practice one can use very old events to trigger the random number generator to push back the time of purported interaction. Zeilinger et al. right now hold the record, by using light from two very far away Quasars at opposing side of the sky to trigger the random number generators, back to 7.8 billion years. From a conceptual point of view this only makes “Non-freedom of Choice” interpretations less likely, as everything has interacted with everything else at the big bang. This interpretation would however imply that literally everything is predetermined (super-determinism) and there just is no freedom of choice in this universe. This seems equally queasy.

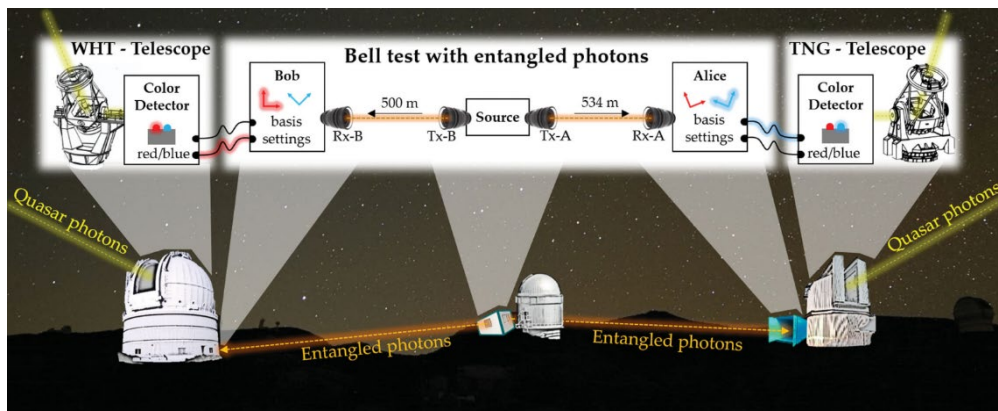


Fig. 43 Scheme of a Bell-Test with random number generators triggered by very old light (7.8 Gyr). PRL **121**080403 (2018).

7 Quantum Key Distribution

In the preceding chapters we introduced all of the basics required to discuss quantum physical communication protocols. Before we start, however, one needs to keep in mind that any kind of data-based communication architecture can usually be described in a layered approach. For any quantum physical communication this is not different and the same is true for communication architectures, that generate and distribute quantum keys, that allow two or more users to communicate securely. A layered model for this case could look like this:

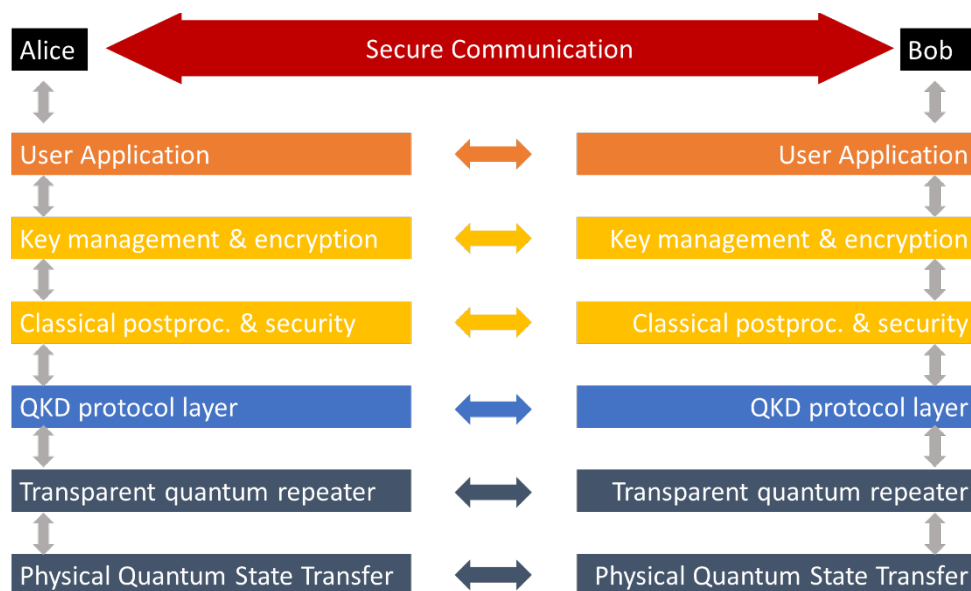
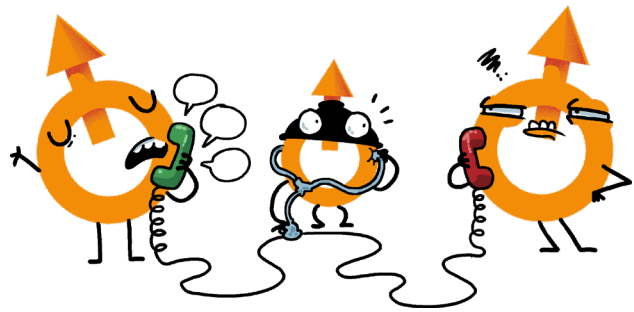


Fig. 44: A possible layer model for a secure quantum communication architecture. For a logical point of view each layer communicates with its counterpart, but uses the layer beneath to do so. In the end, the system serves the purpose of Alice communicating with Bob.

Note that the purpose of the communication architecture is for Alice to communicate with Bob. They do so with a User Application (for example a quantum secured messenger...let's call it QuatsApp). The two user applications, however, secure the data communication with quantum secure keys, which they somehow have to use and manage in some form. Their generation is done in a QKD-protocol. The QKD-protocol itself must use a specific physical implementation, consisting of devices to prepare, transfer and detect physical quantum states. They may also use quantum repeaters.

In this chapter we shall exclusively deal with the light blue QKD protocol layer, which takes an arbitrary physical implementation of Qubits for granted and utilizes these to generate a secure key, that only Alice and Bob do know, in order to use it for secure communication. We shall discuss the dark blue physical layer in later chapters. The yellow and orange layers will just be discussed very briefly here together with a more general introduction on cryptography and its role today.

7.1 Fundamentals of Cryptography

Until recently, cryptography (from greek *kryptos*: hidden, secret) was synonymous with the process of *encryption*, that is, the art and science of making messages un-readable to anyone but the intended receiver. Nowadays, cryptography covers everything from *authentication* and *digital signatures* (confirming identity and/or authorship of messages), *contract signature* and *commitment protocols* (allowing you to sign a contract without anyone knowing who signed it until a disclosure condition is met, e.g. a certain date passes), *private information retrieval* (reading data from a server without the server knowing what the query is), *asymmetric encryption protocols* (public key for encryption of messages and private key for reading the message), *symmetric key encryption protocols* (requiring identical shared secret keys), to the generation of random numbers. All of which playing together have given us online banking, cryptocurrencies, distributed ledger technologies, personal messengers, social networks or in short: the internet the way we know it today.

Without encryption an attacker can empty your bank account, read your messages, steal your identity, shop with your credit card. Without it, you could switch off power stations, crash airplanes, change mixtures for medicine, have cars built together the wrong way, in short: without it we'd be screwed pretty badly (that doesn't mean we aren't anyway, but that's a whole different story).

While the field of quantum cryptography relates to many of these applications¹⁹, we will focus on one of the most widely pursued applications in this lecture: **quantum key distribution** for symmetric encryption protocols.

Let us now focus on the task of secret communication between our protagonists: Alice and Bob. Alice has a precious manuscript that she wants to send to Bob, without anyone else – least of all a malicious eavesdropper (Eve) – being able to read it.

The very first step towards secret communication between Alice and Bob, is for them to prove to one another that: i) they are indeed who they claim to be and ii) Alice's message has not been tampered with; this process is called **authentication**. For example, when you read this text, you know that it was uploaded by someone with password access to the uni's moodle server. Another historic example is the wax seal: if the seal was unbroken, Bob could be sure that a letter is indeed from Alice, and that it had not been tampered with. So, without going into any detail on this, let's think of the authentication process as either a unique seal or a secret that only Alice and Bob know.

The second key component is **encryption**; even if the seal is broken, Alice wants to be sure that the message is meaningless or incoherent²⁰ to anyone but the intended recipient, Bob. Alice can do this by scrambling the message ("plaintext") into a cyphertext according to a certain algorithm. An example for such a procedure is a *transposition cypher*. The plaintext:

"Hello world"

becomes

"lfmmp xpsme"

when each letter is shifted by one letter in the alphabet. Bob can readily *decipher* the *cyphertext* (*lfmmp xpsme*) by the inverse procedure and recover the original *plaintext*. Of course, this example would be rather easy for Eve to break even without knowledge of the algorithm; for example, using

¹⁹ Incidentally, the first proposed application was tamperproof quantum money.

²⁰ (yes, quantum coherence will play a role in the following...)

knowledge of the English language, Eve would immediately infer that the sequence –mm- in the cyphertext is more likely to correspond to the letters “ss ee tt ff ll mm oo” than to “hh jj qq”. If the message were longer, Eve could also sort the symbols by the number of occurrences, and compare with the frequency at which we expect particular letters appear; for example in English, the letter “e” occurs more often than the letters “f, g, y, p, b, v, k, j, x, g, z” combined.

7.1.1 Symmetric Encryption

Ok, so how could one improve this? Alice and Bob can improve the security of this approach by combining the transposition algorithm with a *secret key*. To be specific, Alice and Bob could agree to shift each letter in the plaintext by a different amount, according to some sequence they agreed upon in advance. Since Alice and Bob need an identical key for encryption and decryption the message, such a scheme is referred to as a *symmetric key cryptosystem*.

In fact, this type of symmetric encryption can be made perfectly secure if the key is completely random and sufficiently long. This is known as a Vernam cypher, or a one-time pad encryption protocol. To understand the Vernam cypher it will be more convenient to consider Alice’s message as a binary bit sequence, say “1110 0111”. Alice and Bob share a secret key “1001 1001”. Alice uses the key sequence to transpose the bits of the plaintext (in the case of bits this corresponds to a bit flip). Mathematically, this operation is nothing but the bit-wise modulo-2 addition “ \oplus ” of the key and the plaintext.

$$\text{Plaintext} \oplus \text{Key} = \text{Cipher} \quad (220)$$

Bob can then recover the original message using the exact same procedure:

$$\text{Cipher} \oplus \text{Key} = \text{Plaintext} \quad (221)$$

This protocol is perfectly secure if:

- The key is perfectly random and known only to Alice and Bob
- The key length is as long as the message
- The key is only used once (one-time pad)

If these conditions are met, then the ciphertext is just as random as the secret key and the encryption protocol is unconditionally secure: the Vernam cipher contains no information about the message whatsoever, and no code-breaker may ever extract any meaning from the ciphertext, regardless of computational power and ingenuity. In practice the 2nd requirement may be relaxed and good estimations suggest that a very reasonable security can be obtained with state of the art symmetric ciphers if there is one secret key but for every 10⁶ data bit, which need communication.

Such symmetric encryption ciphers are, however, rarely used (at least not on their own), because they require that Alice and Bob have to exchange secret keys using a secure channel, before they could even start to communicate. A classic example for this are TAN-lists, that have been used for quite some time in online banking: you get a letter (as in paper) by the bank with a list of secret keys and for every transaction the bank will ask you to enter one specific key. In practice you tend to lose the list and it is quite cumbersome. *Quantum Key Distribution Systems solve just exactly this problem: they use the properties of Qubits and the laws of quantum physics to generate secret key, which are guaranteed to only be known to Alice and Bob.*

7.1.2 Asymmetric Encryption

Before moving on to the implementation of quantum key distribution, a short note on how this issue is solved in contemporary communication systems. The solution is based on so-called asymmetric cryptographic protocols. A prominent example is a public / private key encryption protocol: Alice, wants to

All notes subject to change, no guarantee to correctness, corrections welcome.

receive secret messages from Bob (and maybe even Chloe and Dave). She produces two types of keys: a public key, that she – as the name suggests – makes publicly available and a private key, known only to her. The public key allows Bob to encrypt a message; but the message can only be deciphered using Alice’s private key. To illustrate this, think of Alice sending out open combination padlocks to Bob, Chloe, and Dave; they can lock the padlock by simply closing and setting any number (which is a simple problem) – but only Alice can open the lock and make the message readable (for everybody else this is a hard problem, they would have to try all the numbers).

In practice Asymmetric Encryption schemes are rather slow (as is the case with QKD) and thus often used to exchange keys, that are then used with symmetric ciphers to encrypt larger sets of data. The most common scheme is termed Diffie-Hellman-Scheme. A brief explanation using the trapdoor-function “Color-mixing” is displayed in Fig. 45.

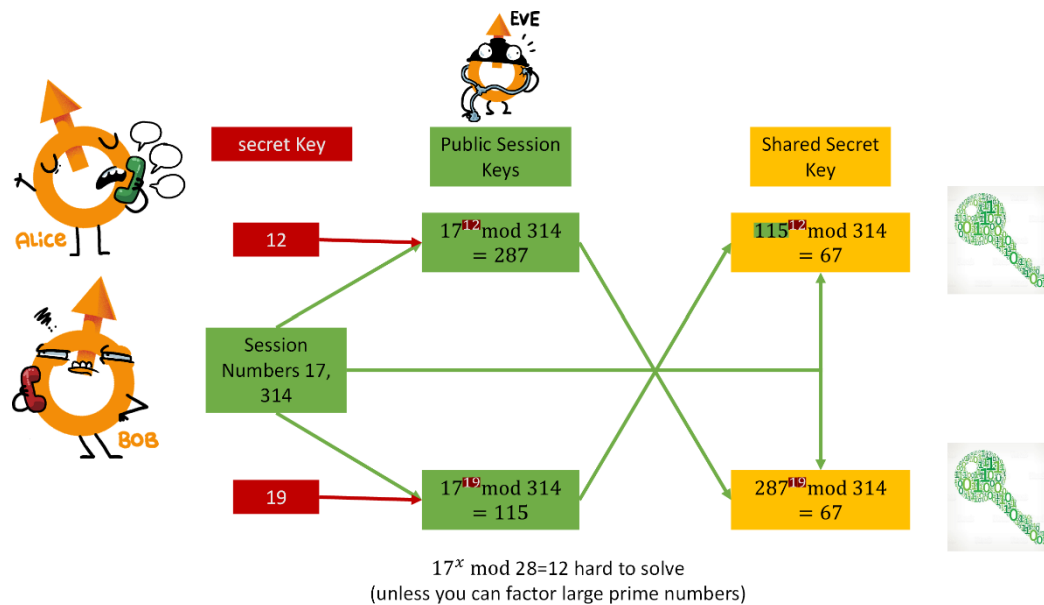


Fig. 45: The Diffie-Hellman-Scheme uses asymmetric encryption to established a shared secret, i.e. it distributes keys, using trap-door type functions. The most comm trap-door is the discrete log $l = a^b \bmod c$, where l easy to calculate if a, b, c are given but b is hard to calculate if a, c, l are known. (red) Secret Data, (green) public data, (yellow) shared secret.

In practice, these open padlocks can be coded using mathematical problems that are hard to solve, but easy to verify, so-called trap-door functions. In the example above this trapdoor function is the discrete log $l = a^b \bmod c$, where l easy to calculate if a, b, c are given but b is hard to calculate if a, c, l are known. This is used e.g. the RSA protocol. The usage of the words “easy” and “hard” implies a certain qualitative nature of the argument and it remains valid only if:

- the inverse problem is really hard, in the sense that no faster algorithm for the specific problem exists (not proven)
- the supposed attacker is not willing and capable to spend exceeding amounts of resources into brute-forcing the attack (the digital communication is valuable, the more an expensive attack is viable)
- the supposed attacker does not simply wait (and store the information) to let Moore’s Law turn the hard problem into an easy one.

In fact, all of the above security assumptions are flawed. For example, having a good algorithm for factoring large numbers: while such algorithms are not known for classical computers, there is an efficient way of doing this on a quantum computer. And these may become operational in a matter of

All notes subject to change, no guarantee to correctness, corrections welcome.

years and certainly within decades. This has in fact, inspired a new field of cryptography: post-quantum cryptography, that is, the development of crypto algorithms that are at least *assumed to be* hard to also solve on a quantum computer. However: “assumed to be” is the scary key word. Do you want paypal to rely on assumptions?

The second assumption is certainly not true for large nation states, which may spend virtually infinite resources in brute-forcing codes. As the amount of values, which we protect with such codes grow, there is also a growing incentive into investing large amounts of resources and money into breaking codes. So, the more ubiquitous digital infrastructures get, the more worthwhile expenses required for an attack may become and the more efficient they become from a cost-value point of view.

The third is also not true. We know that a large portion of the internet’s traffic was copied and stored for possible later decryption and there is no reason to believe this has changed. Moreover, many critical, cryptographically protected systems have exceedingly long deployment times and must be designed such that they remain secure for their entire life span (one prime example being navigation satellites: time from development to end of life > 20 years).

To cut a long story short: we need encryption, and we need more every day. Classical systems for encryption are, however, fundamentally flawed.

But – and that’s why we are here together – quantum theory also provides us with the means to solve this potential security issue: information-theoretically secure communication via one-time pad encryption with secret and random keys that are generated via quantum key distribution. With measurable, physical security for each message, which cannot be broken by anyone that has to obey to the laws of nature.

7.2 Physical Security Fundamentals

In this chapter we shall use single qubit to transmit data secretly from Alice to Bob. The security of this approach is entirely based upon the no-cloning theorem (see section 4.3). We shall reiterate its consequences here: Alice has a single photon that she wishes to transmit to Bob via a quantum channel. Let’s say she can prepare the photon in one of two states: $|\phi\rangle$ or $|\psi\rangle$.

What Alice doesn’t know, is that Eve is listening in to the transmission of the quantum state. Let’s assume Eve has a very sneaky quantum measurement device that can somehow infer the state of the photon without destroying it. The readout register of Eves devices is initialized in a state $|init\rangle$ and should switch to the state $|A\ sent\ \psi\rangle$ or $|A\ sent\ \phi\rangle$ depending on the state sent by Alice. Additionally, since Eve wants to go entirely undetected, the operation should leave the state sent by Alice unchanged.

From the no-cloning theorem we know, that

$$\langle\psi|\phi\rangle = \langle\psi|\phi\rangle\langle Alice\ sent\ \psi|Alice\ sent\ \phi\rangle \quad (222)$$

Which admits only two possible solutions; either $\langle\psi|\phi\rangle = 0$, which means that Alice sent orthogonal states; or $\langle Alice\ sent\ \psi|Alice\ sent\ \phi\rangle = 1$, that is, the readout is the same, irrespective of the state sent by Alice, which doesn’t convey any information about the transmitted state and would beat the purpose of the listening device.

With this knowledge, we can identify a crucial ingredient on secure communication: Alice must use non-orthogonal states in her communication. The first and most-well known implementation of a such a protocol of call BB84 and is discussed in the next section.

7.3 QKD with Single Qubits / BB84

The BB84²¹ Protocol allows Alice and Bob to establish a secret key that they can use to encrypt subsequent messages; hence the name quantum key distribution. Alice and Bob are connected via a Quantum Channel that conserves qubit states and a classical communication channel (public channel, e.g. internet). To implement the protocol Alice will need a single-photon source, a random number generator, and a polarization modulator. Bob needs is a single-photon detector, a random number generator, and a polarization detection module.

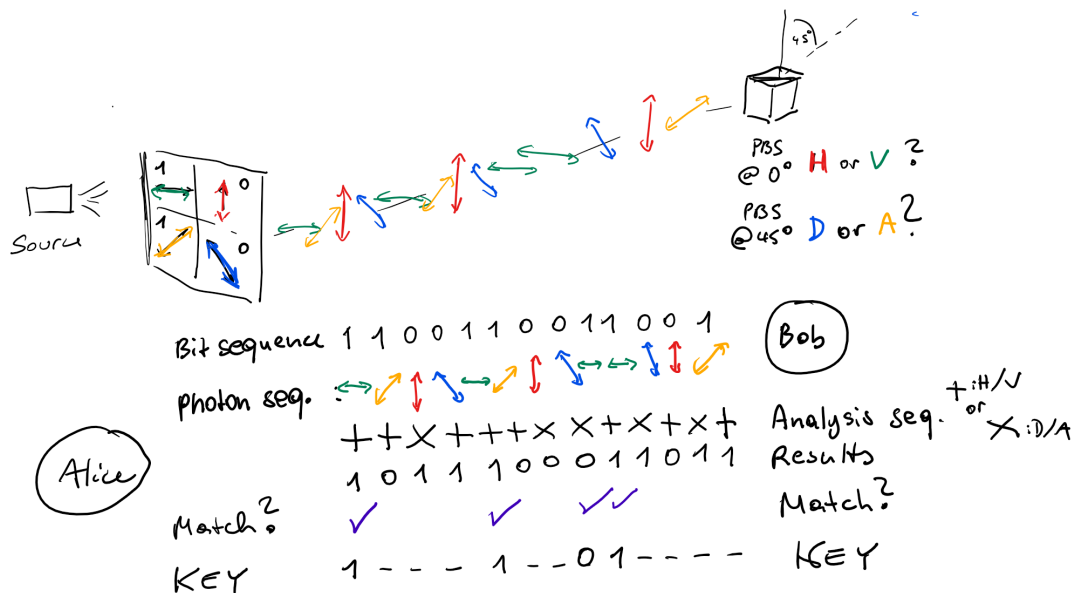


Fig. 46 BB84 with single-photon polarization states. Additional graphical support provided in the lecture slides.

The basic idea is to use two types of non-orthogonal quantum states to encode the bit values of the subsequent key. Alice chooses according to some random pattern how to encode the bit values of the key onto each photon, e.g. 1=H, 0=V and 1=D, 0=A. Bob analyses the photons, either in the D/A basis or the H/V basis. After the transmission of photons is complete, Alice and Bob publicly announce the encoding and analysis basis they used for each photon. They discard all bits that correspond to mismatching encoding and analysis basis. If no one interfered with the quantum state transmission, then the remaining bits should be strongly correlated. They verify that this is the case, by comparing a fraction of the – ideally perfectly – correlated bit sequence. To see what happens if an eavesdropper interferes, let's assume Eve intercepts a fraction of the transmitted photons, measures their polarization, and transmits a (different) photon, that she encodes according to the result she obtained. However, since Eve does not know which basis the photons are encoded in, she can only guess. When she guesses correctly, she will know the correct bit value and the attack will go undetected. However, if she guesses incorrectly, then she will introduce an Error with 50% probability. Alice and Bob can thus identify the attack when they compare parts of their key.

Alice & Bob agree on the two different encoding bases (i.e. H/V or U/D) and on an encoding pattern in these bases for the bit values (0=H or U, 1=V or D) via a classic communication channel.

Private randomness generation: Alice generates two random number sequences, one for the bit value (Alice's raw key) and one for the encoding basis

²¹ C. H. Bennett and G. Brassard, *Proc. Internat. Conf. Computer Systems and Signal Processing* (1984).
 All notes subject to change, no guarantee to correctness, corrections welcome.
 Version of 11.03.2022, Page 92

Transmission: Alice sends single photons with polarizations according her random number sequences over the quantum channel to bob.

Detection: Bob receives photons and measures polarization. For each photon he chooses randomly between the {H,V} and the {D,A} measurement basis. The detected bit sequence is Bob's raw key.

At this stage Alice and Bob each have a **raw key**, where some of the bit values should be identical (namely those where Alice and Bob use the same polarization basis to encode and detect, respectively)

Basis reconciliation: Alice and Bob announce measurement basis they chose for each photon sent/detected via a public communication channel, but keep the corresponding bit value secret. Alice and Bob keep only bits values where they chose the same polarization basis for encoding/detection. Hence, they discard approximately half of the raw key that is received by Bob (they will have chosen the same basis only approx. 50% of the time, and Bob's key will be shorter than Alice's due to photon losses in the transmission channel). If everything has gone well, then the resulting bit sequence should be nearly identical for Alice and Bob. The bit sequence at this stage of the protocol is called the **sifted Key**. The concept is summarized below, for the case of some loss, perfect detectors and no eavesdropping.

| Time Code | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Alice's Bits Value | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| Alice's Basis Choice | H/V | U/D | H/V | U/D | H/V | U/D | U/D | U/D | U/D | H/V | H/V | U/D |
| Polarization Sent | V | D | H | U | H | D | U | D | D | H | V | U |
| Bob's Basis Choice | U/D | H/V | H/V | H/V | U/D | U/D | H/V | U/D | H/V | H/V | U/D | U/D |
| Bob's Detection | U | V | H | H | D | D | - | D | - | H | D | U |
| Sift Filter | n | n | y | n | n | y | n | y | n | y | n | y |
| Sifted Key | | | 0 | | | 1 | | 1 | | 0 | | 0 |

Table 1: Transmission, detection, determination of raw keys and generation of a sifted key.

QBER estimation: Alice and Bob have to assume that any errors in the key are due to an eavesdropper. In the case of an "intercept and resend" eavesdropping attack, Eve will have to decide on a measurement basis beforehand, which will be correct in 50% of the cases. In this case, Eve gets to resend the qubit and does not influence Bob's measurement (we are in the case of orthogonality or equality in the no-cloning theorem). If Eve selects the wrong basis then she will randomly flip Bob's result. Bob will again get the proper result in 50% of these cases by pure coincidence. For so any qubit Eve tampers with, she'll have a probability of 25% to change the result between Alice and Bob. If Eve tampers with a fraction ϵ of all keys, she'll induce differences in a fraction of $\epsilon/4$ keys.

To identify any tampering of the key, Alice (or Bob) sends a **fraction of the key** across a public channel. Bob (or Alice) compares it with his (her) fraction of the sifted key to estimate the quantum bit error rate (QBER). As one example: assume Eve tampers with every Qubit and Bob and Alice invest only 72 bits of their sifted key for QBER detection. Assuming no measurement noise, they will detect that bits have been changed in

$$p = 1 - \left(\frac{3}{4}\right)^{72} = 1 - 10^{-9} \quad (223)$$

of all cases. In practice, detector noise, non-perfect modulation and background light will always introduce a certain QBER and one has to agree on an upper acceptable threshold after which the key is assumed to be compromised and has to be resent. The acceptable threshold level is determined by a security proof, based on the ability of the following steps to exclude erroneous bits (error correction/information consolation) and invalidate the value of Eve's residual knowledge on the key (privacy amplification).

All notes subject to change, no guarantee to correctness, corrections welcome.

| | | | | | | | | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Alice's Bit | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | |
| A&B Basis | H/V | U/D | H/V | U/D | H/V | U/D | U/D | U/D | U/D | H/V | H/V | U/D | |
| Pol. Sent | V | D | H | U | H | D | U | D | D | H | V | U | |
| Eve's Basis | U/D | H/V | H/V | H/V | U/D | U/D | H/V | U/D | H/V | H/V | H/V | U/D | 50% |
| Eve's Pol. | U | V | H | H | D | D | H | D | V | H | V | U | |
| Eve's Bit | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | |
| Eve is Correct? | n | y | y | y | n | y | y | y | y | y | n | y | 75% |
| Bob's Pol. | H | D | H | U | V | D | U | D | U | H | V | U | |
| Bob's Bit | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | |
| Error? | y | n | n | n | y | n | n | n | y | n | n | n | 25% |

Table 2: Consequences of a complete (e.g. $\epsilon = 1$) measure-and-resend attack on the part of sifted key, which is announced publicly and thus sacrificed to determine a QBER rate.

Error correction: If the QBER is lower than the threshold value for security, Alice and Bob proceed with the remaining bits of their respective sifted key. As they should have the same QBER as the publicly compared one, one can be quite sure that these bits are in fact not completely equal. In this step we should identify those bits and correct them (or throw them away), all the while not exposing too much information about the rest of the key. This is important because at this state Eve may in fact know quite a few bits of the sifted key and she knows certainly, which of these are correct.

For error correction Alice and Bob can for example use a low density parity check on blocks of bits (block-wise XOR of all the bits). Alice and Bob compare the parity of their blocks, which will and only expose one additional bit per block. The block's parity will only mismatch if the block contains an error, if the block size is chosen such that the probability of it containing more than one faulty bit value is small²².

If an error is detected, Alice and Bob can either discard the entire block (which would be wasteful) or they can break the block down into smaller chunks to pinpoint the faulty bit value. During this procedure Alice and Bob are leaking information about the key to Eve. This information leakage + the initial fraction $2 \cdot \text{QBER}$ must be taken into account in the next, and final step of the protocol, the privacy amplification. To cut a long story short. You'll get an estimate on the minimum number of bits Eve still cannot know of the corrected and sifted key under the least favourable scenario. If this number is large enough (e.g. 128 or 1024 bits) you can now construct a secret key, of the same (guaranteed) strength.

As an example. Assume the QBER was roughly $1/16$. You can now safely assume that Eve does not know more than $3 \cdot \text{QBER} = 3/16$. For this QBER you could select a good error correction block length of, say 8 bits. Which will mean that Eve will gain another $1/8 = 2/16$ information of the sifted bit. The actual numbers are a bit worse but in this case you'll know that Eve may not know more than a fraction of $5/16$ of all the bits of the key. If you want to create a secret of length 1024 bits then you'll have to have an original sifted key length of roughly 1490 bits.

Both contributions to the maximum number of bits that Eve may know scale with the QBER so there is a practical minimum limit on the QBER that can be tolerated. This is typically in the order of 11%.

Privacy amplification: At this stage Alice and Bob share an identical bit sequence, but the sequence is not completely private since Eve knows a fraction of the bits. To counteract this, Alice and Bob use a privacy amplification protocol. This procedure allows them to increase the secrecy, i.e. to "invalidate"

²² There are other schemes possible here. You can, e.g. select larger blocks (which may have multiple errors) and then reshuffle the key and repeat the process a few times.

the information Eve may have about the key. Of course, this procedure consumes some of the key, so that A and B end up with a shorter, but more private/secret/secure, bit sequence.

This is done by using the partially secret key as an input value to a universal hash function. A hash function is a function that reduces a long number into a short one in a way, that minor changes in the input lead to major changes of the output but in an unpredictable manner. A sought-after property for such a hash-function is the waterfall property: this means that any single bit-flip in the input will change any output bit with a probability of 50% in an uncorrelated way. If the hash-function (there are tons) behaves sufficiently unpredictable then any approximate knowledge of the input (i.e. Eve's info) will not give away info on the final result (i.e. resistance against differential key analysis).

Alice and Bob now share a secret key; they can set an information-theoretic upper bound on the privacy of this key. If the bound is too high for their liking, then they can continue amplifying its privacy at the expense of total key length until they get to the level of security they wish to ensure.

7.3.1 Implementation with laser sources / Decoy State Sources

The security of the BB84 protocols relies on single-photon qubits, which may be difficult to build and characterize. So, what if Alice were to implement the protocol using a weak laser pulse source instead of a true single-photon emitter?

We recall that a laser with an *average* photon number μ emits photons according to a Poissonian number distribution:

$$p(n|\mu) = \frac{\mu^n}{n!} e^{-\mu}$$

This means that every once in a while, the source will emit more than one photon. This opens the door to a so-called "photon number splitting attack" (PNS). The idea behind this attack is that Eve, using a quantum non-demolition measurement (i.e. one that does not absorb the photon state upon detection), can identify which of the pulses contain more than one photon. After identifying the photon number in each pulse, she selectively blocks out all of the pulses that contain only one photon. In the remaining multi-photon pulses she "splits" off one of the photons and stores it in a perfect quantum memory.

Eve now has a photon in her memory that is in the same state as the photon received by Bob; if her quantum memory is good enough, she can just sit and wait until unsuspecting Alice and Bob compare their measurement basis in the key sifting step of the BB84 protocol. Once Eve knows the basis, she performs the same measurement on her photon, and thus gets a perfect copy of the now-not-so-secret key (without introducing any errors). Admittedly, this scenario might seem somewhat academic: after all, Eve needs a quantum memory, a quantum non-demolition measurement, and a way of hiding the loss she introduces in the transmission channel. But if you boldly claim "unconditional security" you'd better find a way to solve this issue.

Fortunately, this type of attack can be identified by a slight modification of the protocol. For Eve to perform a PNS his attack, she will have to block many of the pulses (assuming that $\mu < 1$) and introduce an unnaturally high loss for the single photon pulses. The idea of the decoy state BB84 protocol, is to purposely introduce multi-photon pulses into the BB84 sequence and to detect the photon number statistics of both the signal pulses with average photon number μ_s and decoy pulses with a different average photon number μ_d . For each pulse Alice varies the average photon number according to

All notes subject to change, no guarantee to correctness, corrections welcome.

(another) random sequence. She now has two types of pulses (more typically 3). When Alice and Bob compare the transmission probability for the different pulse types, Eve's attack will lead to a higher-than-expected transmission probability for pulses with a higher average photon number and can be detected as well.²³

7.3.2 Extension to Entangled Qubits (BBM92)

Entangled photons may be used to extend the BB84-scheme, with the source now not being operated by Alice but located in the middle of the link, operated by Charlie. Alice and Bob use the same state encoding and steps as in the original BB84 with single photons. The only difference is that Alice is now a receiver and no longer a sender of photons.

If the source emits, without loss of generality, $|\phi^\pm\rangle$ -states, then Alice and Bob will share an identical key from all measurements, in which they measure in the same basis. The measurements in different bases are still discarded. The rest of the protocol remains unchanged. The protocol was invented in 1992 by Bennett, Brassard, and Mermin after the seminal work by Ekert, that first suggested the usage of entangled states for QKD. Ekert's protocol is discussed below in a separate section. Just as Ekert91 BBM92 has so-called unconditional security; e.g. the protocol itself is secure, if (and only if) properly implemented.

7.4 QKD with Entangled Qubits / Ekert 91

One of the key points in the previous chapters was the notion, that entangled photons are more strongly correlated than classically correlated pairs of photons. We had also seen in chapter 4.1, that a measurement on either partner of the entangled qubit pair destroys the entanglement and leads to a mixed states, which is classically correlated. Also remember that the most promising mode of attack for QKD-based communication is an intercept-and-resend attack, which requires a measurement to be made; thus leaving an entangled photon pair in a now classical state.

If Alice and Bob each receive a Qubit from an entangled source, they can now conduct a Bell-test. If the test ends up with $S \leq 2$, then they know, that have only received classically correlated photons and there is an eavesdropper present. This is one of the two key ideas of the Ekert91-protocoll.

The second idea is, that in the original implementation of the Bell-test (see chapter 6.2) all the events where Alice and Bob have measured along the same basis (e.g. both measure along a , meaning both measure in a H/V-basis) are discarded. However, in this situation we know that both Alice and Bob will always measure the same (the opposite) result because the initial state was a $|\phi^\pm\rangle$ -state ($|\psi^\pm\rangle$ -state). These previously used qubit can then be used to establish a secret key.

²³ For an example of a state-of-the-art polarization-based decoy state QKD source, refer to Jofre et al. Optics Express, Vol. 19, 2011. For an example of a state-of-the-art time-bin encoded decoy state QKD source refer to Baaron et al. 2018, <https://arxiv.org/pdf/1804.05426.pdf>

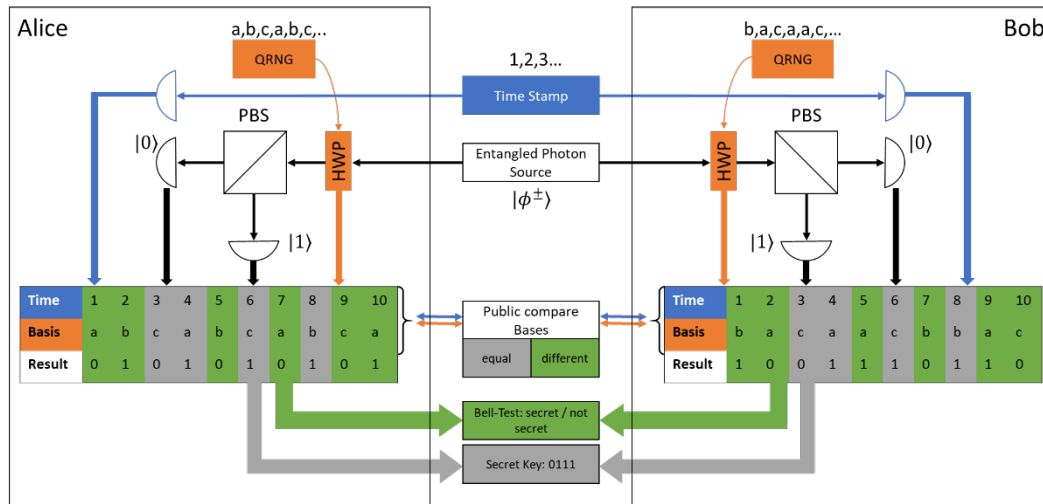


Fig. 47 Sketch of the Eckert91-Scheme. Entangled Photons are measured in independent bases (called a, b, c). Bases are the compared publicly. Different bases are used for Bell-testing. Same bases are used for secret key generation.

A sketch of the protocol can be seen in Fig. 47. First Alice and Bob each measure a series of polarization values for randomly selected bases for a proper set of three basis vectors, called a, b , and c . They then publish their basis selection. Different base measurements are used to establish S . If $S > 2$ then we know the photons are still entangled and can use the equal base measurements to establish a secret key.

| | | | | | | | | | | | | | |
|----------------|----|----|---|----|---|----|----|----|----|----|----|----|----|
| Time Stamp | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| Alice's Basis | a | b | c | b | a | c | a | a | c | b | b | c | a |
| Alice's Result | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Bob's Basis | b | c | c | a | a | B | c | b | a | c | b | b | a |
| Bob's Result | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| Equal | n | n | y | N | y | n | n | n | n | n | y | n | y |
| Used for S | -+ | +- | | -+ | | -+ | -- | ++ | +- | ++ | | +- | |
| Key Bit | | | 1 | | 0 | | | | | | 1 | | 1 |

Table 3: An example for the Eckert91 protocol.

A simple example for the operation of the scheme is given in Table 3. Note that this is the plain vanilla implementation. In a practical implementation one still has to deal with loss, errors and has to apply privacy amplification.

The role of loss is very similar to its role in the BB84-protocoll. Loss will induce time-stamps, where Alice or Bob simply did not get any click and these have to be discarded in the public announcement stage. As with BB84 loss does reduce the data transmission rate to the point, where the number of true events is smaller than the number of dark-count, leading to an initial continuous drop off in the key bit rate and then a sudden failure of the protocol.

Errors (both induced by the eavesdropper, as well as bad entanglement, detector noise, stray light, etc...) are all classical noise sources and induce (false) classical correlations in the measurement. They thus reduce S . If $S < 2$ then secrecy cannot be guaranteed and the protocol has failed (we may need to try again). For values of $2 < S < 2\sqrt{2}$ we may derive an upper boundary for the fraction of photons Eve may know and may use privacy amplification to deny her any useful knowledge on the secret key.

As with BB84 this amplification happens at the cost of secret bits per transmitted Qubit and thus reduces the data rate further.

One major advantage over BB84 is the independence of the secrecy with respect to the source. The secrecy hinges on the entanglement of the source, which is measured in every step and does not need to be claimed by the operator of the source. As such any manipulation of the source, which may lead to the operator of the source gaining knowledge on the key would reduce S the same way as any other attack. Thus, entangled photon QKD has the advantage that you need not even trust the operator of the infrastructure for his integrity. In fact, the source may be built, operated, or manipulated by the NSA, Huawei, or your overly nosy neighbour: they still cannot attain any knowledge about the secret key Alice has shared with Bob.

7.5 Overview over other security issues and mitigation strategies

After having discussed some of the theoretical concepts underlying quantum cryptography, let us recall the purpose of quantum key distribution, and equally important, the issues that it does not address as well as possible issues with the implementation.

In quantum key distribution, as the name suggests, allows Alice and Bob to share a secret key, that they can subsequently use to encrypt classical messages. If they do this using the one-time pad (Vernam Cypher), then the encryption level is said to be information-theoretically secure – that is – as secure as the key itself. QKD protocols such as BB84 provide a means of distributing such keys as secure as required. In technical terms: Alice and Bob can establish by information-theoretical means: an eavesdropper who attempts to interfere with the quantum state transmission involved in the protocol will leave a detectable trace, that Alice and Bob can measure and thus establish an upper bound of the amount of information such an attack may have revealed.

However, QKD does not address the issue of how this key is used in an encryption system²⁴ which is only as strong as its weakest component. For example, QKD cannot protect Alice and Bob from Eve “looking over their shoulder” and obtaining their secrets directly at the source e.g. via malware installed on their PC. Ensuring that the entire encryption chain is secure, that is, ensuring end-to-end security is an important issue that occupies many scientist and engineers in the IT security community. As QKD is only a part of a complete encryption system, this leaves room for attacker to focus on breaking, upstream of downstream components of this chain. Upstream components, not covered by QKD, involve e.g. attacks on the user authentication system. In other words: QKD does not provide Alice any type of guarantee that she is really talking to Bob and not some imposter. Downstream parts of the encryption chain may involve attacks on the symmetric cipher used for the data payload or attacks on the devices that Alice and Bob use to process data or Alice and Bob themselves.

However, even in the face of protocols with unconditional security, a new generation of “quantum hackers” have identified or at least proposed feasible ways to even break the QKD-process itself. These attacks rely on the notion that unconditional security does only apply to the protocol and not necessarily to the physical device that the protocol is implemented upon. Thus, even QKD-system are vul-

²⁴ – Strictly speaking, the information theoretical argument only holds if the Vernam cypher is used for encryption. This one-time-pad encryption requires a key of equal length as the message to be securely transmitted, and, in practice, state-of-the-art QKD systems do not yet provide the required key rates. QKD is thus often combined with other encryption standards such as the Advanced Encryption Standard (AES).

nerable to so-called side-channel attacks, that exploit or enforce information leakage out of the system. In this specific case quantum mechanics is also no help, as we know from the discussion on entropy and the measurement process we know that a macroscopic measurement system must leave traces of the measured quantum information in its thermodynamic state. Or to put it in other words: quantum mechanics guarantees that there are side channels in your system, no matter what. These must be protected within the implementation itself.

To give a concrete and very real example, the very first implementation of the BB84 protocol used electro-optic modulators that were driven by kV voltages that resulted in an audible signal each time the basis was changed – paraphrasing Bennet’s take on this in a recent conference: “we demonstrated QKD that was information-theoretically secure against a deaf eavesdropper”. Any attacker with a microphone could this have gained enough side-channel information to break the supposed unconditional security.

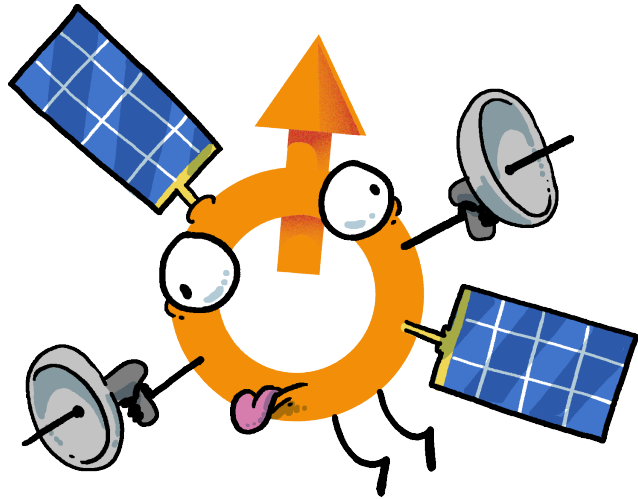
| SECURITY ISSUE | DESCRIPTION | COUNTERMEASURES |
|---|--|---|
| Trojan-horse attack | Eve probes the QKD equipment with light to gain information about the device settings | privacy amplification (PA), isolators, filters |
| Multi-photon emission | When more than one photon is emitted in a pulse, information is redundantly encoded on multiple photons | PA, characterisation, decoy states, SARG04 and other protocols |
| Imperfect encoding | Initial states do not conform to the protocol | PA, characterisation |
| Phase correlation between signal pulses | Non-phase-randomised pulses leak more info to Eve, decoy states fail | phase randomisation, PA |
| Bright-light attack | Eve manipulates the photon detectors by sending bright-light to them | active monitoring, measurement device independent QKD (MDI-QKD) |
| Efficiency mismatch and time-shift attack | Eve can control, at least partially, which detector is to click, gaining information on the encoded bit | MDI-QKD, detector symmetrisation |
| Back-flash attack | Eve can learn which detector clicked and hence knows the bit | isolators, MDI-QKD, detector symmetrisation |
| Manipulation of Local Oscillator reference | In continuous variable QKD (CV-QKD), the local oscillator (LO) can be tampered with by Eve if it is sent on a communications channel | Generate LO at the receiver. Phase reloading, i.e. only synchronise the phase of LO |

Fig. 48: Potential attacks due to imperfect implementation of QKD. Table taken from: https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp27_qkd_imp_sec_FINAL.pdf

Obviously, we can’t discuss all possible modes of attacks in this script. So, we have just summarized some of the most important attacks and mitigation strategies for them in Fig. 48. Also keep in mind that the promise of unconditional security of the protocol tends to embolden scientists and engineers alike to the point that they may start to become particularity sloppy in the implementation. So, if you ever develop a QKD system, be double wary of possible quantum hacks and other side channel vulnerabilities. At least initially QKD systems will protect high-value assets and will be a prime target for an attack. If the system is successful and scales into a mass market it may protect lower value targets but many, many more of them. It will also operate in a plethora of much less well-defined environments. In other words: your system will become an even more exposed target for attack. So take such issues seriously and don’t get fooled by the beauty of the protocol into a false sense of security.

7.6 Transmission rate and limits on transmission distance

So far, we have only looked at the protocol level of the communication infrastructure (at the key-distribution infrastructure, to be precise), not at the physical level. I.e. we have not yet dealt with the problem of how we shall actually get the Qubits (e.g. the photons) from Alice to Bob. While this, of course, depends somewhat on the specific implementation of the Qubit, we shall treat this rather generically here, as the key point is the modified role of loss in transmission systems.



Because of the no-cloning theorem, we cannot use signal amplifiers and must either live with the loss of any given transmission line or chose a non-lossy transmission line. In this chapter we shall discuss the two most commonly implemented schemes and see how loss in such schemes scales with distance. We shall also discuss the impact of unavoidable noise sources and their interaction with loss in the derivation of QBER; which determines how likely it is to extract a secure bit from a physical qubit.

7.6.1 Noise Sources / Dark Count Rate

The security of QKD-protocol evolves around the notion that no-cloning-theorem forces an eavesdropper to induce involuntary and random changes into the qubits exchanged between Alice and Bob. These changes lead to a measurable mismatch between Alice and Bob, which is recorded in the $QBER$. The nature of the protocols guarantees that Alice and Bob can distil a secure sequence of bits if $QBER < QBER_{MAX}$, e.g. is below a critical value. In practice you want to be well below this threshold because as the number of secure bits (e.g. the key rate) vs the number of received qubits (the transmission rate) drops towards zero, if the $QBER$ approaches $QBER_{MAX}$.

This critical value $QBER_{MAX}$ is calculated under the worst-case assumption that any contribution to $QBER$ is caused by an attempted eavesdropping. In reality there are many sources for quantum errors, which are totally unrelated to someone listening to your system, some of which cannot be avoided for the system in question. The most important noise sources are those that create a constant level of erroneous measurements per time. The most prominent among them is the dark-count rate ρ of your single-photon detectors: this means that any single photon detector will report the arrival of a photon every now and so often, even if there really was none. In general, this dark-count rate is exponentially dependent on the wavelength the detector is sensitive at; also exponentially dependent on its temperature and also on its material and measurement principle. As a rule of thumb: you want to have cold detectors and short wavelength. A typical rate for state-of-the-art silicon-based single photon avalanche detectors (SPADS) is in the range of $\rho = 10 \dots 1000s^{-1}$. Let's assume $\rho = 100s^{-1}$ for brevity.

As a rule of thumb, you want to have $QBER \ll 1$ and thus you need a signal rate $R \gg \rho$ or in other words: the brightness of your source and the loss of your transmission system must be designed such that a signal rate of R is received that is well above the detectors dark count rate ρ because otherwise you'll never get a secure key from the number of photons that you receive.

7.6.2 Fiber-Based Transmission

Some practical considerations using real numbers. Let's assume that Alice and Bob are connected via a single mode fiber (typical loss value: $\gamma = \frac{0.3 \text{ dB}}{\text{km}}$). Alice sends out single photons at a repetition rate $R=100 \text{ MHz}$ – i.e. one photon per 10ns. Due to loss in the transmission fiber, Bob will only receive a fraction of these photons. After L km of lossy fiber Bob receives photons at a rate of:



$$R_t(L) = R_0 \times 10^{-L[\text{km}] \times \frac{\gamma}{10}} \quad (224)$$

For a transmission distance of 100 km, for example, Bob only receives 100.000 photons per second, not yet taking into account losses in his receiver or the efficiency of single-photon detectors.

At the same time Bob's receiver will occasionally produce a signal, without there being a photon altogether (or Bob may detect a "faulty" photon for whatever reason). This will give him a wrong result. If the detectors dark-count rate is roughly 1000 s^{-1} then $QBER = 0.01$, which is quite acceptable.

If the length of the line however goes up to, say 166 km, then the count rate will go down to 1000 s^{-1} and $QBER = 1$ and you can't run QKD anymore because every key might be compromised. Thus, for a given transmission line the transmissible bit rate drops exponentially with distance and at some distance the protocol fails, because the QBER becomes too large. As fiber transmission losses are already close to the theoretical minimum these numbers are not bound to change dramatically to, say, planetary scale.

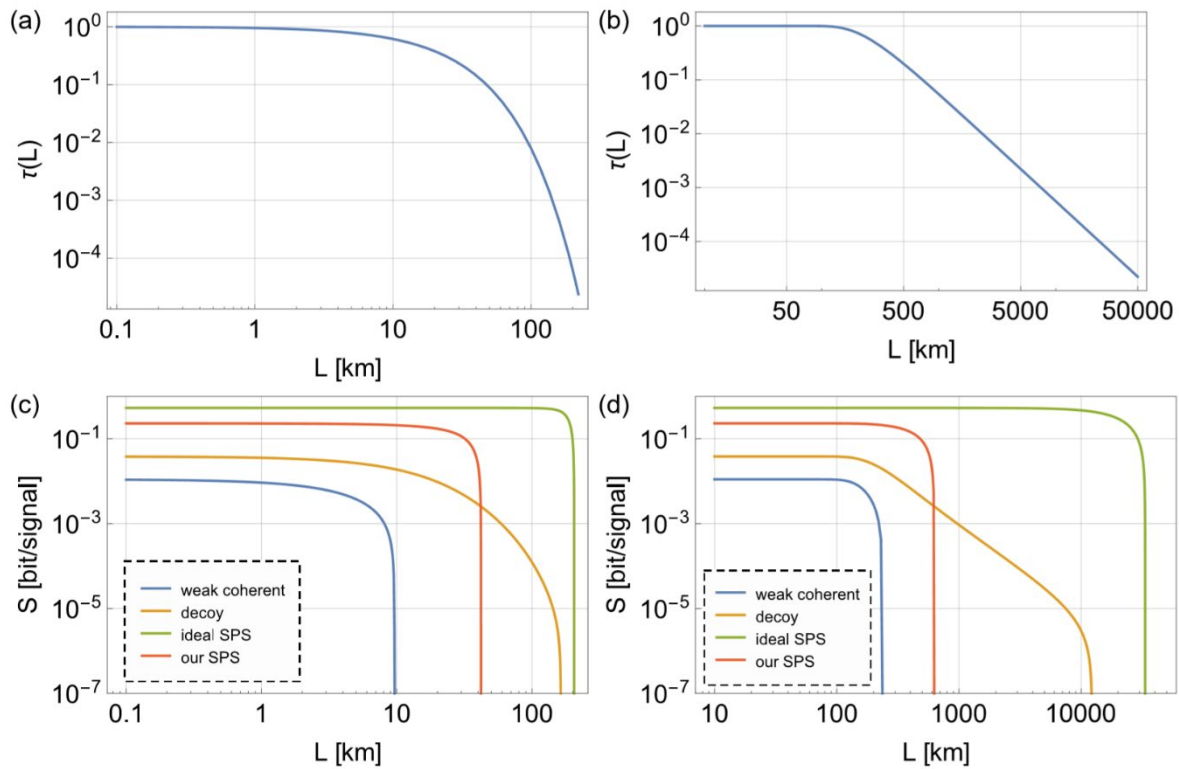


Fig. 49: Channel loss (top) and secret bit rate (bottom) as a function of distance for various physical implementation of the BB84 protocol (left) fiber-based, 0.21 dB/km loss, (right) free space communication, $d=5$ cm on satellite and $d=20$ cm on ground. (Source: T.Vogl et al. "Compact Cavity-Enhanced Single-Photon Generation with Hexagonal Boron Nitride", ACS Photonics 6 1955 (2019)).

7.6.3 Satellite-Based Transmission

While glass (as in a telecom-fiber) is probably the most transparent solid state material which we know, vacuum is still more transparent and diffraction losses have a much more favourable scaling than exponential. It is this a straightforward idea to use a satellite (or a constellation thereof) to exchange qubits between Alice and Bob.

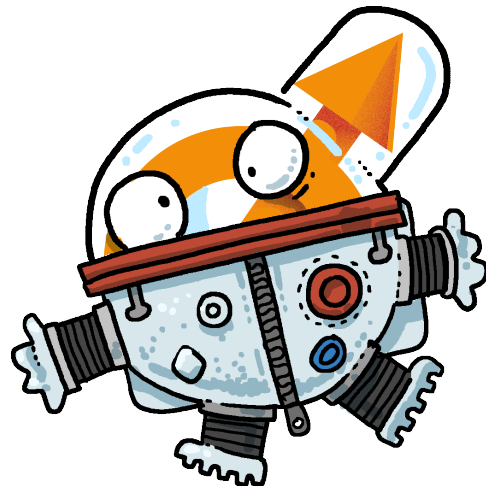




Fig. 50: Chinese Quantum Science Satellite (Micius) sends entangled photons to optical ground stations separated by 1200 km on ground © Jian-Wei Pan.

For practical reasons it is most convenient to place a Qubit source on a satellite and a similar order-of-magnitude-calculation for the loss can be given. We assume that the Alice’s photon source is mounted on a satellite and that both the satellite as well as the detector are equipped with a $d = 20$ cm telescope; for the sake of simplicity, we also assume that the beam is a Gaussian with $1/e^2$ diameter of d . At $\lambda = 800$ nm, we get a Rayleigh length of $z_0 \approx 40$ km. For a $d = 200$ cm telescope we get $z_0 \approx 4000$ km. After a certain distance the beam thus covers an area of

$$A(z) \sim A_0 \left(\frac{z}{z_0} \right)^{-2} \quad (225)$$

Division by the area of the detector (which we here assume to be of the same size A_0) will give you the relative loss. You also must account for the effects of the turbulence in the atmosphere L_T , which can be approximated as anywhere between 10 ... 20 dB for the downlink case and 30 ... 60 dB for an up-link case (because the angular divergence, which is introduced by mainly the lower part of the atmosphere can act on a longer path in the uplink-case), this is called the “shower curtain effect”.

At any rate, here are some losses, which you can expect:

| Orbit Type | Telescope Diameter (both) | Orbital Height z | Loss ($-L_T$) |
|-----------------|---------------------------|--------------------|-----------------|
| Low earth orbit | 20 cm | 300 km | 18 dB |
| | 200 cm | 300 km | 2 dB |
| Geostationary | 20 cm | 30.000 km | 56 dB |
| | 200 cm | 30.000 km | 18 dB |

Table 4: Diffractive signal loss estimation for a signal wavelength of $\lambda = 800$ nm, in various orbital heights. Does not include scattering/turbulence loss.

Because we know from the discussion above, that 30 dB loss is acceptable, we can immediately conclude, that a low-earth-orbit approach with a constellation of small and (comparatively) cheap satel-

lites is feasible as well as a geostationary solution with a VERY expensive single satellite and VERY expensive ground stations. We also see that direct reception on a mobile phone scale device will also not be an option (unless you have a very large mobile phone).

At any rate, the situation remains as above: for a given implementation you will experience a drop of data rate and a relative increase of QBER, which will rapidly kill off your protocol, if it approaches a certain critical threshold and thus impose a hard limit on the length of your communication line.

7.6.4 Alternative Schemes

Of course, there are alternative schemes on how to deal with the physical level of the communication hardware, which have been developed in particular with respect to the issue of loss and non-amplification. Some research directions are given below (without any claim for completeness), with a rough grouping into general strategies



Fig. 51: The largest trusted-node network at the time of writing of this script, running over 2000 km from Beijing to Shanghai. Note, that a trusted node network does not guarantee secrecy against the operator of the network nodes.

1. Modifications to the Single-Photon-Sources (see chapter on sources)
 - higher single photon rate
 - shorter photon lifetime (this allows you to synchronize the detector and effectively reduce the dark count rate)
 - sources at particular wavelengths (atmospheric windows, low-loss fibers, Fraunhofer-Lines)
2. Modification to the detectors (see chapter on detectors)
 - lower dark count rate
 - higher quantum efficiency
3. Modification to the Qubit and its physical implementation
 - non-polarization Qubits (orbital angular momentum, multi-mode in specialty fibers)
 - implementation as Qudits (Qubit with more basis states than just two)
4. Modifications to the protocol, with some loss of security
 - trusted node networks

- put the source in the middle between Alice and Bob (can be extended to use entangled photons)
5. Advanced Quantum Magic (see chapter on Quantum Teleportation)
- The quantum repeater

8 Advanced Quantum Communication Schemes

This chapter shall be devoted to non-cryptographic protocols in quantum communication. These have a slightly different take on the role and outcome of the quantum-ness. Whereas in Quantum Communication, you are specifically looking for violations of the Quantum State of light transmitted from Alice to Bob, we are in this case assuming that no perturbation is present (at least initially) and we will explore, what Alice and Bob can do, if they connect their mutual photons with local particles of their own.

If the connection is made on a Quantum level, then we will see that this leads to Quantum Teleportation, e.g. that a local destruction of the quantum state in Alice's lab and its reappearance in Bob's. If the connection is made on a classical level, this leads to the concept of dense coding, where many classical bits of information can be transmitted from Alice to Bob, using a single qubit.

8.1 Quantum Teleportation

We shall now see, how Entanglement can be transferred between previously unentangled systems. Suppose that Alice has a Qubit in quantum state $|\psi\rangle_A = \alpha|0\rangle_A + \beta|1\rangle_A$, with $|\alpha|^2 + |\beta|^2 = 1$, which may be unknown to Alice herself. The subscript A denotes the location of the photons as being in Alice's lab.



Also suppose that there is a photon-pair source, which is emitting biphotons in the maximally entangled Bell-State $|\psi^-\rangle_{AB} = 1/\sqrt{2}(|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B)$, where the A-part of the bi-photon is emitted towards Alice and the B-part is emitted towards Bob. Note, that the specific choice of the $|\psi^-\rangle$ -Bell-State is arbitrary; any Bell-State may, in fact, be chosen, as long as it is known.

The total state of the system of the three photons can be written as the Tensor-Product of the two states, namely

$$\begin{aligned}
 |\psi\rangle_A |\psi^-\rangle_{AB} &= (\alpha|0\rangle_A + \beta|1\rangle_A) \otimes \frac{|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B}{\sqrt{2}} \\
 &= \frac{\alpha|00\rangle_A|1\rangle_B - \alpha|01\rangle_A|0\rangle_B + \beta|10\rangle_A|1\rangle_B - \beta|11\rangle_A|0\rangle_B}{\sqrt{2}}
 \end{aligned} \tag{226}$$

In a next step we rewrite the Alice's Photon, which is written in the Computational Basis-State into Bell-States, using the relations:

$$\begin{aligned}
 |00\rangle &= \frac{|\Phi^+\rangle + |\Phi^-\rangle}{\sqrt{2}} & |11\rangle &= \frac{|\Phi^+\rangle - |\Phi^-\rangle}{\sqrt{2}} \\
 |01\rangle &= \frac{|\Psi^+\rangle + |\Psi^-\rangle}{\sqrt{2}} & |10\rangle &= \frac{|\Psi^+\rangle - |\Psi^-\rangle}{\sqrt{2}}
 \end{aligned} \tag{227}$$

Which we then substitute into the above relation, yielding:

$$\begin{aligned}
 |\psi\rangle_A |\psi^-\rangle_{AB} &= \frac{|\Phi^+\rangle_A}{\sqrt{2}} (\alpha|1\rangle_B - \beta|0\rangle_B) + \frac{|\Phi^-\rangle_A}{\sqrt{2}} (\alpha|1\rangle_B + \beta|0\rangle_B) \\
 &\quad - \frac{|\Psi^+\rangle_A}{\sqrt{2}} (\alpha|0\rangle_B - \beta|1\rangle_B) - \frac{|\Psi^-\rangle_A}{\sqrt{2}} (\alpha|0\rangle_B + \beta|1\rangle_B)
 \end{aligned}
 \tag{228}$$

In a next step Alice can make a Bell-State-Measurement, with an appropriate setup (this has some caveats in its own rights, we'll get to that later) and Bob's state will collapse into one of the terms in the parenthesis with equal probability. Suppose that Alice's measurement yields a $|\Psi^-\rangle$ -state. Then we know, that the state of Bob's photon is now $\alpha|0\rangle_B + \beta|1\rangle_B$, which is an exact copy of the initial state of the (unknown) photon Alice has. Of course, we cannot guarantee that Alice will measure this state (in fact this only happens in 25% of all cases) so we must add another step to complete the procedure.

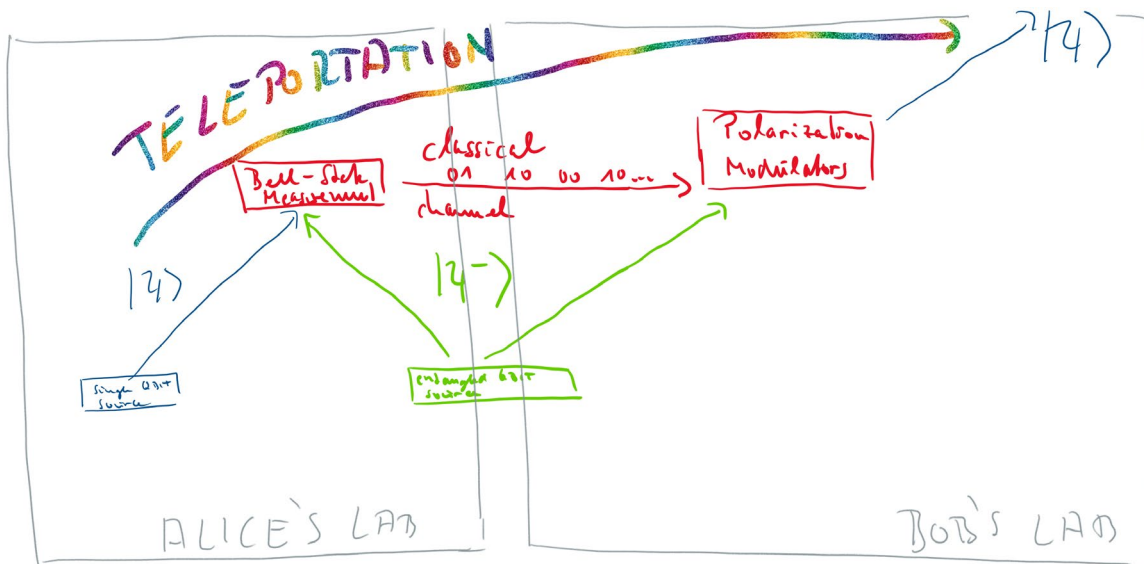


Fig. 52: Schematic of the Quantum-Teleportation Scheme. Alice and Bob share Qubits of an entangled photons of an entangled Qubit-source. Alice uses her photons together with an unknown photon to conduct Bell-State measurements. She transmits the results to Bob who now manipulates his photon according to the Bell-measurement result and obtains a perfect copy of the (still unknown and now destroyed) photon Alice once had. The outcome is a transfer of the quantum state of Alice to Bob.

First Alice measures the Bell-State of her system. The result can be any of four $|\Phi^\pm\rangle$ or $|\Psi^\pm\rangle$. Alice can now transmit the result of her measurement to Bob on a classical communication channel, using two ordinary bits. According to these two bits, Bob can make the following manipulations to his state:

| Alice measures | Bob needs to | Bob applies | Polarization Implementation |
|------------------|--|----------------------------|---|
| $ \Psi^+\rangle$ | nothing | nothing | nothing |
| $ \Psi^-\rangle$ | Phase shift $ 1\rangle$ by π | σ_Z | HWP with angle $\theta = 0$ wrt. $ H\rangle$ -axis |
| $ \Phi^-\rangle$ | bit flip | σ_X | HWP with angle $\theta = 45$ wrt. $ H\rangle$ -axis |
| $ \Phi^+\rangle$ | Phase shift $ 1\rangle$ by π then bit flip | σ_Z then σ_X | HWP with angle $\theta = 0$ wrt. $ H\rangle$ -axis then HWP with angle $\theta = 45$ wrt. $ H\rangle$ -axis |

After these modifications Bob always ends up with a Qubit in the same initial state as Alice used to have originally. The quantum state has been teleported from Alice to Bob.

A few remarks on this, however, need to be made

1. This protocol does not violate the **no-cloning-theorem**. Alice's state is destroyed in the Bell-State-Measurement-Process. During this measurement she learns nothing of her state. There is only ever the one copy of the initial state.
2. This protocol does not violate the **no-signalling-theorem**. While the entanglement is transferred instantaneously, Bob can only measure Alice's state with 25% probability before the information of the result of Alice's measurement has arrived. In fact, one may show, that due to the equal distribution of probabilities, Bob may infer no meaningful information at all (in the sense of better than pure guesswork) about Alice's state before the arrival of her result whatsoever.
3. The application of above protocol requires Bob to retain his Qubit until the information of Alice's result has arrived. For photons this is difficult. However, for a lot of situations Bob may proceed to use his state instantaneously, i.e. to make a measurement, but needs the results from Alice's measurement to a-posteriori make any sense of his measurement result.
4. Alice can only carry out her Bell-State measurement, if both of her photons are in the same set of modes $|0\rangle$ and $|1\rangle$. They must thus interfere. They must thus arrive within a time-frame defined by their mutual coherence-time. This makes such experiments very hard, if no photons are available "on demand". If they are only available in a heralded manner, one needs to resort to post-selection detection (more later).

8.2 Entanglement Swapping and The Quantum Repeater

We will now generalize the concept of Quantum Teleportation by replacing the Single Qubit Source, which is residing in Alice's lab by another entangled photon source. We'll also shuffle around the names of the lab's owners somewhat and label them according to cities. This will help you get in touch with the local geography and also bring you closer to a possible applications.

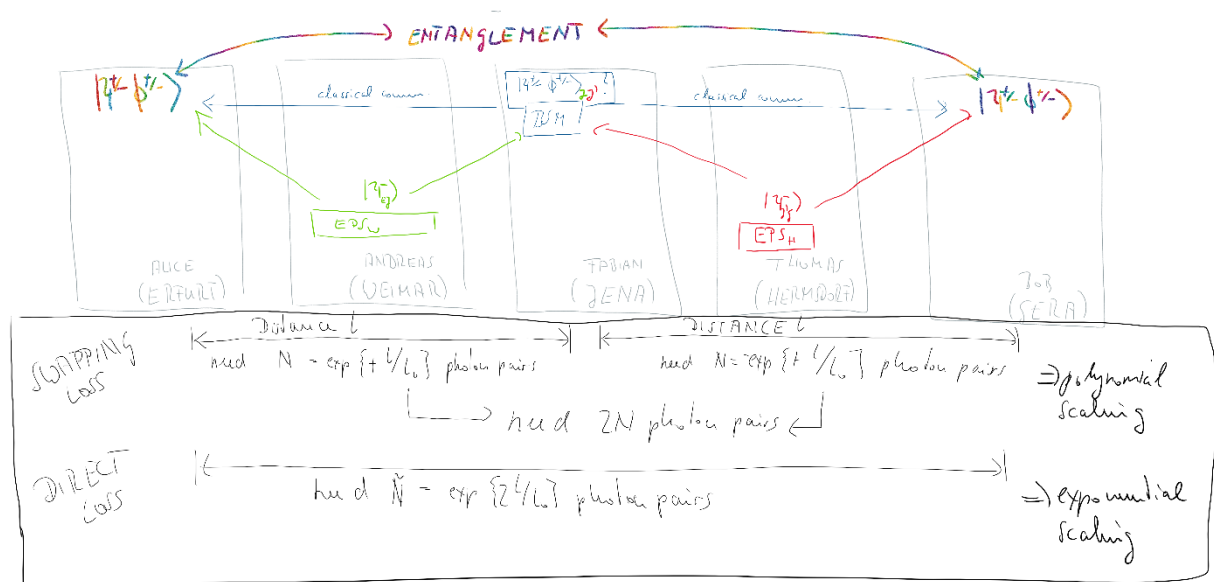


Fig. 53: Schematic of the Entanglement Swapping Scheme and its application as a Quantum Repeater. Andreas and Thomas, how reside in Weimar and Hermsdorf, respectively. Each are in command of an Entangle Photon Pair Source (EPS). Andreas sends on of his Photons to Erfurt, where Alice resides and his second photon to Jena into Fabian's lab. Thomas sends the first to Jena, too and the second to Bob, who lives in Gera. Fabian is conducting a Bell-State-Measurement in his lab. As a consequence of the measurement Alice and Bob now have an entangled photon pair in a Bell-State. If Fabian communicates them the result then Alice and Bob also now, which Bell-State their photons pair is in. As an added benefit, the loss in this scenario scales only polynomial with the loss of each section, whereas for direct sending, it would scale exponentially.

All notes subject to change, no guarantee to correctness, corrections welcome.

Assume that the two Entangled-Photons-Sources (EPS) are located in Jena's neighbouring towns of Weimar (to the West) and Hermsdorf (to the East). They are placed in such a way, that they both send one photon to Jena (denoted with J for the Weimar source and J' for the Hermsdorf source) into an auxiliary lab, run by Fabian. The other photons are going further west to Erfurt (for the Weimar-source) and further East to Gera for the Hermsdorf-source). They are denoted with E and G , respectively. Erfurt and Gera also happen to be the location of the labs of Alice and Bob respectively. Let's also assume that both sources are constructed in a way, that they emit in the $|\Psi^-\rangle$ state. The combined system is then in the following state:

$$\begin{aligned}
 |\psi\rangle &= |\Psi^-\rangle_{EJ} |\Psi^-\rangle_{J'G} \\
 &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)_{EJ} \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)_{J'G} \\
 &= \frac{1}{2} (|0101\rangle - |0110\rangle - |1001\rangle + |1010\rangle)_{EJJ'G} \\
 &= \frac{1}{2} (|\Psi^+\rangle_{JJ} |\Psi^+\rangle_{EG} - |\Psi^-\rangle_{JJ} |\Psi^-\rangle_{EG} - |\Phi^+\rangle_{JJ} |\Phi^+\rangle_{EG} + |\Phi^-\rangle_{JJ} |\Phi^-\rangle_{EG})
 \end{aligned} \tag{229}$$

Fabian in Jena now makes a Bell-State-Measurement on the photon pair JJ' the photon pair EG owned by Alice and Bob will also collapse into a Bell-State. If Fabian now communicates his result of the Bell-State-Measurement to Alice and Bob they also know, which Bell-State their System collapsed into. As a consequence Alice and Bob now share a maximally entangled photon pair, although each of these photons have never interacted directly, apart from the interaction mediated by the entanglement of their initial partner photons. The measurement of the Bell-State by Fabian has basically transferred (i.e. "swapped") the entanglement onto another pair of photons.

This is quite an amazing result. We already know, that Entanglement is not bound to local action. This scheme, however, shows that Entanglement can also be created nonlocally. In a sense Photons-Pairs are not twins, which are created side-by-side in the same process, but pairs which may obtain their bond much later.

The process of entanglement swapping has two immediate applications. The first is related to Quantum Physics itself. While we use Photon pairs as model systems for entanglement, we may also create, e.g. entangled states of Photons and Spins. If two these are created and their photonic partners are subjected to a BSM, we are left with entangled Spins, which we may now exist at remote locations without having to transport either of the Spins, which may be very hard to achieve indeed.

The second may have even more profound practical applications. Assume that the communication lines from the EPS-Sources are subject to loss (fiber loss, scattering loss, diffraction loss). Let's also assume that each section has an equal length of l and loss which scales linearly with the length, such that the probability of a Photon-Pair to make the trip from the source to Alice/Fabian or Fabian/Bob is $p = \exp(-\frac{l}{l_0})$. Let's also assume that Fabian first waits for the photon from the Weimar source to arrive and is somehow able to detect its existence and its entangledness with the photon in Erfurt. For this to happen we have to invest in average $N = p^{-1}$ photon pairs. Let's then assume that Fabian can store this J -photon until the J' -photon from Hermsdorf arrives. This will cost another N photon pairs. The total cost of photon pairs to invest is thus $2N$ photons. If we were to place a EPS directly in the lab of Fabian and then attempt to directly share a photons pair directly between Alice and Bob, we'd require to have them survive a travel over twice the length $2l$ and thus, we'd thus need N^2 photons, which is of course much more expensive.

Let's further assume Alice and Bob were much further away and connected by a network of M identical nodes of length l . The cost is then NM as opposed to the N^M for a direct link. The latter cost of course prohibitive, whereas the multiple-entanglement-swapping scheme is, in principle, acceptable. Because the cost-reduction scaling and the overall topology of such a scheme is comparable to the repeater-scheme for classical communication this is called the Quantum Repeater Protocol and may be used for fiber-based secure communication in combination with Eckert91.

Before you get too carried away, a few words of caution, which are absolutely necessary because there is – as of yet – no practical Quantum Repeater and basically no hope of seeing one in the next few years, although it is conceptually quite elegant:

1. Fabian will have to store photons (in their quantum state) and release them on demand, to synchronize the (inherently unpredictable) arrival for the BSM. No such device exists. State-of-the-Art systems attempt to store the photon by converting it to Spin-Waves in Ultracold atomic gases. Hardly scalable.
2. Fabian must inspect the photons, which he received, for existence and entangledness. Particularly, for the latter no such device exists.
3. The scheme scales only well with Photon-Pair-On-Demand sources as all other sources need extensive synchronization techniques. None of the well-developed EPS work in an On-Demand mode.

8.3 Superdense Coding

So far we have used bits obtained by Bell-State-Measurement, being transmitted in a classical manner, to help us transfer or exchange entanglement, using Quantum Teleportation and Entanglement Swapping. We may also revert the scheme and use Entanglement to transmit information; and do so in a secure manner. While Eckert91 is the most straightforward approach to do so, it is by no means the end of the line. In fact, we may retain the safety of Eckert91 and transmit more than two classical bits of information encoded in a single Qubit.²⁵



²⁵ This is the limit for Qubits. However, photon pairs may be entangled in more than just a binary degree of freedom (called Qudits and Hyperentanglement). In this case we may transmit an arbitrary number of bits per photon pair. This has been shown and can be used to mitigate some of the issues with photonic BSMs as discussed in the next chapter. What, there is no next chapter? Bummer.

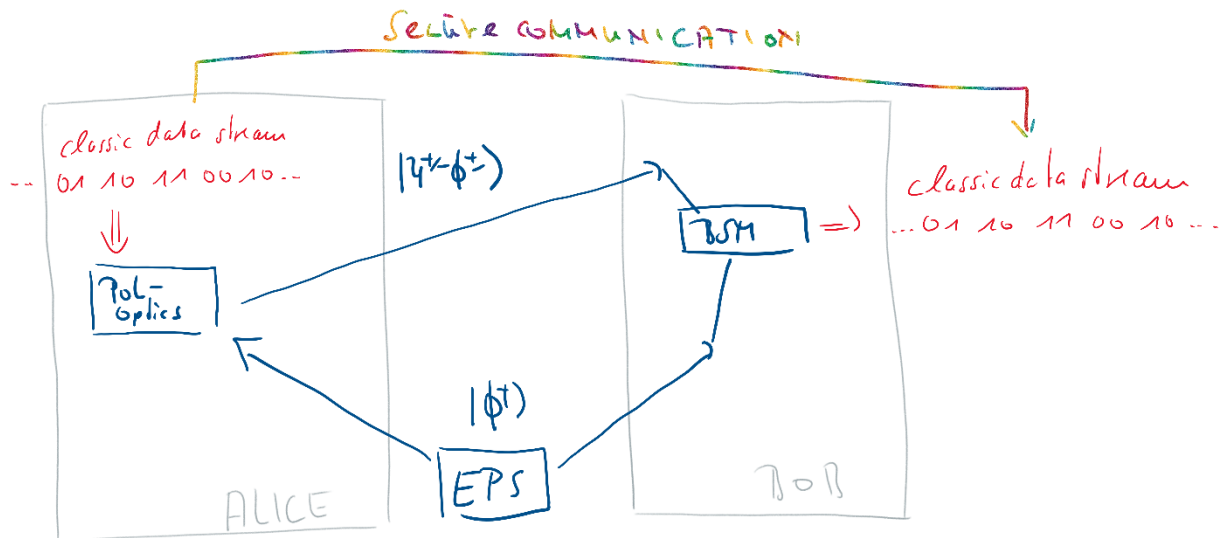


Fig. 54: The superdense coding scheme. Alice and Bob share an entangled photon pair. Alice manipulates her photons in four possible ways, switching between the Bell-States and sends her photon back to Bob. Bob performs a BSM and retains two bits of information from Alice from a single photon pair.

Let's assume that Alice and Bob share the two photons of the Bell-State $|\Phi^+\rangle$. Alice may now manipulate her photon in such a way that the photon-pair is in any of the four Bell-States. As she can create four possible Bell-States, her action can be interpreted as a message composed of two-classical bits. After doing so, she transmits her photon back to Bob, who makes a BSM with the re-united photon-pair. The measurement reveals Alice's action and thus the two bits of her message. The actions performed by Alice may be encoded like this:

| Alice's message | Bell-State | Alice Needs to | Alice applies | Polarization Implementation |
|-----------------|------------------|--|----------------------------|---|
| 00 | $ \Phi^+\rangle$ | Nothing | nothing | nothing |
| 01 | $ \Phi^-\rangle$ | Phase shift $ 1\rangle$ by π | σ_z | HWP with angle $\theta = 0$ wrt. $ H\rangle$ -axis |
| 10 | $ \Psi^+\rangle$ | bit flip | σ_x | HWP with angle $\theta = 45$ wrt. $ H\rangle$ -axis |
| 11 | $ \Psi^-\rangle$ | Phase shift $ 1\rangle$ by π then bit flip | σ_z then σ_x | HWP with angle $\theta = 0$ wrt. $ H\rangle$ -axis then HWP with angle $\theta = 45$ wrt. $ H\rangle$ -axis |

There are two particularly noteworthy features of this scheme. First, the action-table above is pretty much the same as the one for Quantum-Teleportation. In fact, as discussed above, superdense coding may be thought of as being reciprocal to Quantum-Teleportation.

The second noteworthy feature is the security. Any eavesdropper, who catches the photon sent from Alice back to Bob cannot make sense of its state; the retrieval of Alice's information requires command over the second photon, which only Bob has. If the eavesdropper, however, intercepts the photon on the way from Bob to Alice and replaces it with any photon of his own, the results of Bob's BSM will not reflect the data Alice has sent. If Alice occasionally sprinkles in Test-Data, which she then compares with Bob on an open channel, such a man-in-the-middle-attack would be noticed and communication could be cancelled.

As a side remark: the doubling of the channel capacity wrt. Eckert91 is dearly payed for. The photons has to travel twice the connection (square the losses) and Bob has to store his photon somehow (see Quantum Repeater).

A 1 Theoretical Description of Photon Detection

In the initial parts of the lecture we had introduced *modes* as excitable states in the quantum description of the electromagnetic field. When a field mode is excited to a certain discrete energy level we consider it as being occupied with a certain (or uncertain) number of Photons. This yields observations that cannot be explained by classical electromagnetic theory, such as the Bell-tests.

In order to describe such *non-classical* observations, we must first understand how to link between optical quantum states and experimentally accessible properties, such as classical voltages and currents; we must understand how the photo-detection process is described in quantum optics, both from a fundamental and a conceptual point of view. This shall be done with this and the following chapter.

A 1.1 Photon Detection

Photodetectors allow the experimenter to link optical fields she wishes to detect with electrical currents and voltages that are conveniently analysed using oscilloscopes, pulse counters, etc. and play an essential role in all of optics experimentation. Despite being equipped with a pair of quite remarkable photon receivers, obvious issues with practicality as well as limited time resolution, make them quite unsuitable for use in a controlled laboratory environment²⁶. The simplest practical method for detecting light is called “direct detection” (in contrast to *Homodyne* detection, which is not subject of this chapter). Direct photo detection is based on the absorption of photons, e.g. in a semiconductor diode, photocathode, bolometric detector...) whereby the deposited energy produces an electronic response that is proportional to the *intensity* of the incoming radiation. When the detector is sufficiently sensitive, a single photon already suffices to trigger a measurable electronic response with high probability (photo-detection efficiency, PDE). For example, in commercial Si-based single-photon Avalanche Diodes (SPAD, typical PDE >50%) a single absorbed photon results in the emission of an electron and an electron avalanche current (i.e. they operated in Geiger Mode), that, upon further amplification and pulse shaping yields an electronic pulse response or “click” (Fig. 55).

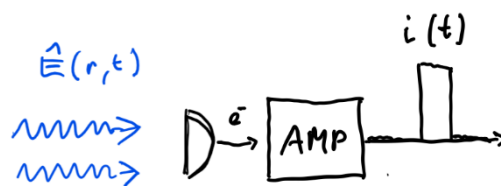


Fig. 55: photo detection Process: photons excite an electron, which goes through several amplification stages to give a detectable electrical signal “a click”.

A 1.1.1 Glauber’s quantum model for photodetection

The quantum theory of photo detection was established by Glauber in the 1960s. In the following we give a brief outline of his approach. Let’s consider an ideal detector consisting of a single atom (Fig. 56) that is localized at a position $r = 0$ and initially in its ground state $|g\rangle$. The interaction Hamiltonian that describes the coupling between the atom and the EM field is:

$$\hat{\mathcal{H}}_I = -\hat{p} \cdot \hat{E}(t) \quad (230)$$

²⁶ Notable recent exceptions include experiments on the single-photon responsivity of the human eye: Tinsley et al. “Detection of a single photon by humans”, *Nature Communications* **7**, 12172 (2016)

where $\hat{p} = d \cdot (|e\rangle\langle g| + |g\rangle\langle e|)$ is the operator for the transition from the atomic ground state $|g\rangle$ to the excited state $|e\rangle$, which are separated by an energy difference ΔE . d is the atom's dipole moment, describing the efficiency of the interaction and can in reality be tuned using resonant antennas in the vicinity of the atom in question, but this is a question for nanophotonics.

Note that $|e\rangle\langle g|$, which is the excitation/absorption operator, and $|g\rangle\langle e|$, which is the deexcitation/emission operator appear symmetrically, otherwise the Hamiltonian would not be symmetric and would thus violate time-reversal symmetry among others. In fact, you may know from laser physics that absorption and emission must be symmetric, because otherwise you may violate the 2nd law of thermodynamics (which is again all about violating time inversion symmetries).

The electric field operator is given by the superposition of all modal creation and annihilation operators at their specific frequencies (the spatial dependence is left for brevity; the atom samples the spatial behaviour at a very small point, this can be integrated into the dipole moment d). The electric field operator $\hat{E} \propto \hat{A}$ is thus:

$$\hat{E}(t) \propto \int (\hat{a}(\omega)e^{-i\omega t} + \hat{a}^\dagger(\omega)e^{i\omega t}) d\omega = \hat{E}^+(t) + \hat{E}^-(t) \quad (231)$$

Next, let us assume that the incident optical field is described by the quantum state $|\Psi_i\rangle$. During the interaction with this field, the atom may absorb a photon and transition from the atomic ground state $|g\rangle$ to the excited state $|e\rangle$, which we shall call a photo electron. This electron may, e.g. then be free to travel in the material of the detector and can itself be amplified and detected electronically. Since the optical field deposits energy in this process, it will transition to some lower-energy final state²⁷ $|\Psi_f\rangle$.

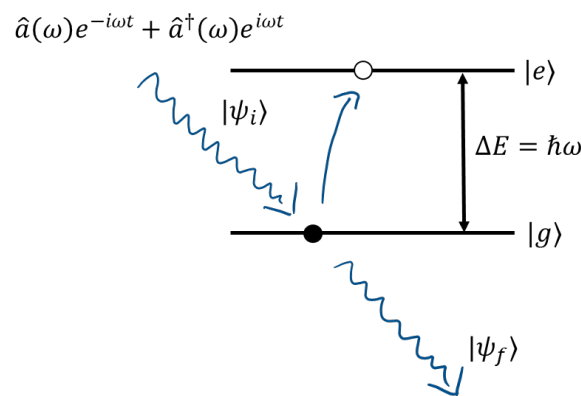


Fig. 56: Interaction of quantum field with two-level atom used in Glauber's model for the photo detection process.

To describe the resulting electrical photocurrent we must determine the rate at which electrons are excited via this interaction. Since we're interested in the absorption of photons, the relevant part of the electromagnetic field operator is that containing the annihilation operator (i.e. the positive frequency part of the field operator). The relevant matrix element of the interaction Hamiltonian is thus:

²⁷ Up to now, quantum states were constant in time and all time dependence was determined by the evolution of operators under unitary time evolution generated by the Hamiltonian $\hat{\mathcal{H}}_0$, i.e. we were operating in the Heisenberg picture. We're now entering the interaction picture, where the states can undergo transitions due to the interaction part $\hat{\mathcal{H}}_1$ of the total Hamiltonian $\hat{\mathcal{H}}_T = \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_0$.

$$\langle e | \langle \Psi_f | \mathcal{H}_I | \Psi_i \rangle | g \rangle = d \cdot \langle \Psi_f | \hat{E}^+(t) | \Psi_i \rangle \quad (232)$$

The calculation of this rate is a lengthy and non-trivial exercise in second-order time-dependent perturbation theory, which is beyond the scope of this lecture. The result of this calculation is known as Fermi's golden rule. It states that the transition rate is nonzero if and only if the energy difference between $\Delta E = \hbar\omega$ and then the time-dependent transition rate (i.e. the probability density for a transition event to occur) between the Eigenstates of the *free* Hamiltonian $\langle p_{g \rightarrow e, i \rightarrow f}(t) \rangle$ is proportional to:

$$\langle p_{g \rightarrow e, i \rightarrow f}(t) \rangle \propto |\langle \Psi_f | \hat{E}^+(t) | \Psi_i \rangle|^2 \quad (233)$$

This gives us the rate of photoelectrons emitted when the field makes a transition to a *particular* final state $|\Psi_f\rangle$. Since we are interested in the final state of the detector (i.e. the photoelectron), and not the final state of the field, we must take the sum over all possible final states for the field:

$$\langle p_{g \rightarrow e}(t) \rangle \propto \sum_f |\langle \Psi_f | \hat{E}^+(t) | \Psi_i \rangle|^2 = \sum_f \langle \Psi_i | \hat{E}^-(t) | \Psi_f \rangle \langle \Psi_f | \hat{E}^+(t) | \Psi_i \rangle \quad (234)$$

Assuming a complete basis $\sum_f |\Psi_f\rangle \langle \Psi_f| = \mathbb{I}$ for these states, we obtain Glauber's result for the rate of excitation of photoelectrons:

$$\langle p_{g \rightarrow e}(t) \rangle \propto \langle \Psi_i | \hat{E}^-(t) \hat{E}^+(t) | \Psi_i \rangle \quad (235)$$

The number of photoelectrons N_e promoted to the excited state $|e\rangle$ in a finite time interval T is then

$$N_e = \eta \int_T dt \langle \hat{E}^-(t) \hat{E}^+(t) \rangle_{\Psi_i} \quad (236)$$

Where we have grouped all relevant constants into a single coefficient η , that describes the quantum efficiency of the photo detector.

8.3.1 Coincidence detection

The expression above links the quantum state of the field to the electronic response in a single detector. In the description of quantum phenomena, in particular quantum entanglement, correlations of detection events play a central role. We thus must extend the theory of photo detection to two, or more, detectors. Glauber also considered this case, and showed that the joint probability density for the excitation of an electron at each of two detectors A and B at times t_A and t_B is given by:

$$\langle p_{g_A \rightarrow e_A, g_B \rightarrow e_B}(t_A, t_B) \rangle \propto \langle \Psi_i | \hat{E}_A^-(t_A) \hat{E}_B^-(t_B) \hat{E}_B^+(t_B) \hat{E}_A^+(t_A) | \Psi_i \rangle \quad (237)$$

Integration over a finite time intervals $[T_A^{(1)}, T_A^{(2)}]$ and $[T_B^{(1)}, T_B^{(2)}]$ gives us the expected number of electrons to be detected in A and B and the according time slots:

$$N_{A,B} = \eta_A \eta_B \int_{T_A^{(1)}}^{T_A^{(2)}} \int_{T_B^{(1)}}^{T_B^{(2)}} dt_1 dt_2 \langle \hat{E}_A^-(t_1) \hat{E}_B^-(t_2) \hat{E}_B^+(t_2) \hat{E}_A^+(t_1) \rangle_{\Psi_i} \quad (238)$$

where $\eta_A \eta_B$ denotes the efficiency of detectors A and B, respectively. Note that here, even more so, the quantum efficiency of the detectors play a crucial role, as they enter the equation quadratically.

Experimentally, the detection of coincident photoelectrons can be performed in a variety of ways. In a time-tagging configuration each detection event "click" at A and B is given a unique time tag $t_{A/B}^{(1)} \dots t_{A/B}^{(N)}$ which is then stored for post-processing. If the optical fields A and B are time correlated,

then the time tags will exhibit a pronounced peak around some delay $\tau = t_A^{(i)} - t_B^{(j)}$. This procedure allows any correlation to be analysed at a later point, but requires on $\mathcal{O}(N^2)$ delays to be measured, and large amounts of data to be stored.

More practically, the correlation of photo-detection events can also be evaluated directly using an electronic *AND* gate with a short *coincidence window* ΔT_c (typically on the order of somewhere between 1 ns and 1 ps), as depicted in Fig. 57. To put this approach into context with the expression above, we transform the absolute time coordinates as:

$$t_A, t_B \rightarrow t, t + \tau \quad (239)$$

For notational brevity, we define $p(t_A, t_B) = \eta_A \eta_B \langle \hat{E}_A^-(t_A) \hat{E}_B^-(t_B) \hat{E}_B^+(t_A) \hat{E}_A^+(t_B) \rangle$. Written this way, the total number of double-detection events $N_{A,B}$ in the interval $t \in [T^{(1)}, T^{(2)}]$

$$N_{A,B} = \int_{T^{(1)}}^{T^{(2)}} \int_{\Delta T_c} d\tau p(t + \tau, t) \stackrel{\text{def}}{=} \int_{T^{(1)}}^{T^{(2)}} dt R_c(t) \quad (240)$$

Where $R_c(t) = \int_{\Delta T_c} d\tau p(t + \tau, t)$ is the coincidence rate, that is, the probability of simultaneous detector clicks in A and B within the coincidence window ΔT_c .

Similar approaches can also be extended to higher-order coincidence detection events. Coincidence measurements are a practical and powerful tool that find application in a variety of quantum communication experiments.

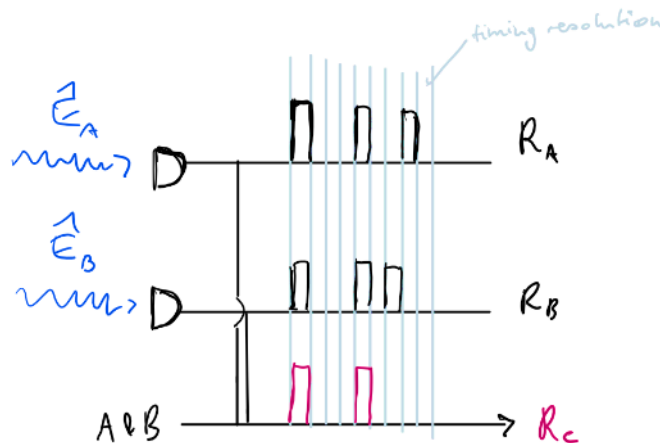


Fig. 57: Basic electronic coincidence counting circuit. The coincidence window approximately corresponds to the timing resolution.

A 1.2 Threshold (“bucket/click”) detectors

In the discussion so far, we have related the electronic detector response to expected values of products of Field operators. For a single detector, these are of the form $\langle \hat{E}^- \hat{E}^+ \rangle$. Restricting the following discussion to the case of a single detection mode, this reduces to the expected value of the photon number operator $\langle \hat{n} \rangle$, i.e. the detector response depends on the number of photons (using $\hat{E}^+ \hat{E}^- \sim \hat{a}^\dagger \hat{a}$, if there is only a single mode present).

Typical photon detectors, such as single-photon avalanche photo diodes (SPAD), however, do not have this “photon-number resolution” capability. They click upon the detection of a photon (or more) but

All notes subject to change, no guarantee to correctness, corrections welcome.

are blind to the number of photons in a field. To see how we model such “bucket” detectors, let us first formalize the photon-number resolving capability that such detectors lack, in order to later throw them away. Glauber’s theory of photo detection relates the number of photo electrons to the photon number operator

$$\hat{n} = \sum_{n=0}^{\infty} n |n\rangle\langle n| \quad (241)$$

Where $|n\rangle\langle n|$ denotes the projection operator onto a particular Fock state, i.e. the projection operator for an “n-photon detection event”. A detector that has photon number resolution is thus described via a set of projection operators:

$$\hat{P}_n = |n\rangle\langle n| \quad (242)$$

where n denotes the respective measurement result. The probability of such an n-photon detection event, when the field is prepared in a pure state $|\Psi\rangle$ is then:

$$p(\text{Detect } n \text{ Photons}) = \langle \Psi | \hat{P}_n | \Psi \rangle \quad (243)$$

Or in the case of a mixed input state $\hat{\rho}$

$$p(\text{Detect } n \text{ Photons}) = \text{Tr}(\hat{P}_n \hat{\rho}) \quad (244)$$

To model a detector that lacks the capability of discriminate between these detection outcomes, we must sum over all n-photon detection probabilities, that is:

$$p(\text{Detect at least 1 Photon}) = \sum_{n=1}^{\infty} \langle \Psi | \hat{P}_n | \Psi \rangle \quad (245)$$

Which naturally leads to the definition of the projection operator for a *bucket detector*:

$$\hat{P}_{\text{click}} = \sum_{n=1}^{\infty} \hat{P}_n \quad (246)$$

Likewise, the operator for calculating probabilities of *no detection* writes:

$$\hat{P}_{\text{no click}} = 1 - \hat{P}_{\text{click}} = |0\rangle\langle 0| \quad (247)$$

This projection operator notation will turn out to be more practical in some of the following chapters.

A coarse description of the practical implementation of contemporary single-photon-detectors is given in the Appendix to this script (A 2).

A 1.3 Correlation functions and coherence

So far, we have introduced correlation functions, because they can be described easily using the formalism of quantum photonics and they can also be measured quite easily (more details, see appendix on photodetectors). However, we shall now see, that such functions are not merely a mathematical toy and experimentally convenient measure but that they also have deep and fundamental meaning, which can help us understand the quantum nature of a light field and also allows us to redefine and expand our understanding of the vital concept of interference.

8.3.2 First-order correlation function

Returning for a moment to the case of a single detector, we note that we can interpret Glauber's result for the excitation rate of photoelectrons $p(t) \equiv G(t, t) = \langle \hat{E}^-(t) \hat{E}^+(t) \rangle$, as a correlation of the field operator with itself at time t , i.e. at delay $\tau = 0$. We can generalize this expression to arbitrary times t_1 and t_2 and to two distinct modes, denoted with A and B :

$$G_{AB}(t_1, t_2) = \langle \hat{E}_A^-(t_1) \hat{E}_B^+(t_2) \rangle \quad (248)$$

or its more commonly used normalized version:

$$g^{(1)}(t_1, t_2) = \frac{G_{AB}(t_1, t_2)}{\sqrt{G_{AB}(t_1, t_1) \cdot G_{AB}(t_2, t_2)}} \quad (249)$$

As we will see in the following, this quantity determines the maximum fringe visibility in an interference experiment. Consider the setup depicted in Fig. 58, where two (for the sake of illustration and notational brevity) classical fields originating from points A and B are superimposed on a balanced beam splitter (BS) and mixed into the output modes, denoted with A' and B' . To study the interference of these fields, it is convenient to apply an additional phase shift to field B , such that the field $E_{A'}$ and $E_{B'}$, in the output ports A' and B' of the beamsplitter is given by:

$$E_{A', B'}(t) = \frac{1}{\sqrt{2}} (E_A(t) \pm \exp(i\phi) E_B(t)) \quad (250)$$

The instantaneous count rate of photodetectors on the output modes is then:

$$p_{A', B'}(t) \propto |E^*(t)E(t)| \propto |E_A(t)|^2 + |E_B(t)|^2 \pm 2 \Re \{ \exp(i\phi) E_A(t) E_B(t) \} \quad (251)$$

Noting that typical detectors are rather slow (compared to the timescales of the source), we should take the average over the response time T , i.e.:

$$\langle f(t) \rangle_T = \frac{1}{T} \int_0^T dt f(t) \quad (252)$$

Which gives us the time average photocurrent $I(\phi)$:

$$\begin{aligned} I(\phi) &= \langle p_{A', B'}(t) \rangle_\phi \propto \langle |E_A(t)|^2 \rangle + \langle |E_B(t)|^2 \rangle \pm 2 \Re \{ \exp(i\phi) \langle E_A(t) E_B(t) \rangle \} \\ &= I_A + I_B \pm 2 \Re \{ \exp(i\phi) \langle E_A(t) E_B(t) \rangle \} \end{aligned} \quad (253)$$

Where the subscript ϕ denotes, that the time averaged photocurrent depends explicitly on the phase shift between the two interferometer arms. We can thus vary ϕ to find the minimum and maximum photocurrents observed:

$$I^{max/min} \propto I_A + I_B \pm 2\sqrt{I_A I_B} |g_{AB}^{(1)}| \quad (254)$$

Using the definition of the fringe visibility V :

$$V = \frac{I^{max} - I^{min}}{I^{max} + I^{min}} \propto \frac{2\sqrt{I_A I_B}}{I_A + I_B} |g_{AB}^{(1)}| \quad (255)$$

We can distinguish three scenarios, depending on the magnitude of $|g^{(1)}|$:

- complete coherence: $|g^{(1)}| = 1$
- partial coherence: $|g^{(1)}| < 1$

- incoherence: $|g^{(1)}| = 0$

For classical light, i.e. coherent states of light, we know that $g^{(1)}$ is mostly dependent on the spectral width of the light; e.g. perfectly monochromatic light will be completely coherent, whereas for any other light there is certain coherence length (i.e. a maximum path difference for A and B) inversely proportional to the bandwidth of the light in question.

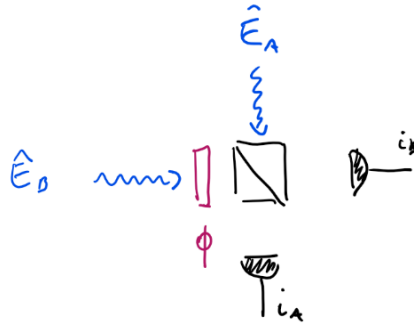


Fig. 58: First-order interference on a beam splitter. The Fringe Visibility observed when scanning the phase ϕ is determined by the First-order coherence function $g^{(1)}$. Note that the detector are not interconnected, they give individual readings and thus no recording of the order of events at the two detectors may be recorded.

So let's see if we get any different result for a monochromatic Fock-State $|n\rangle \sim \hat{a}^\dagger(\omega_0)^n |vac\rangle$, with frequency ω_0 . We substitute the one-dimensional field operators ($\hat{E}(t) \propto \int (\hat{a}(\omega)e^{-i\omega t} + \hat{a}^\dagger(\omega)e^{i\omega t}) d\omega'$) into

$$g^{(1)}(t_1, t_2) = \frac{\langle \hat{E}^-(t_1)\hat{E}^+(t_2) \rangle}{\sqrt{\langle \hat{E}^-(t_1)\hat{E}^+(t_1) \rangle \cdot \langle \hat{E}^-(t_2)\hat{E}^+(t_2) \rangle}} \quad (256)$$

After performing some delta-function acrobatics, we find:

$$g^1(t_1 - t_2) = \exp(t(t_1 - t_2)\omega_0) \quad (257)$$

That is, we observe perfect first-order coherence, i.e. $|g^{(1)}| = 1$. Unsurprisingly, a single frequency mode state is perfectly coherent for all time, irrespective of the photon number excitation. A number of interference experiments (Young's double slit experiment etc.) give the exact same result for Fock states, as we would expect in classical optics.

This also means, that first-order coherence does not allow us to distinguish between quantum states and classical states. This also gives another a-priori justification of the importance of double detector experiments as these measure second order coherence and it is only here that the differences between classical fields and quantum fields truly becomes apparent.

8.3.3 Second-order correlation and the Hanbury-Brown-Twiss Experiments

The first-order correlation function quantifies the correlation of amplitudes and phases of two fields (i.e. the phase coherence). The second-order correlation, on the other hand, tells us about the correlation of intensities of two fields. Fig. 59 depicts a typical experimental setup.

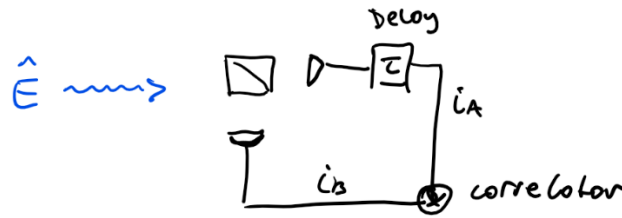


Fig. 59: Hanbury-Brown-Twiss setup is it was first used to measure the 2nd order intensity correlations of classical light in stellar interferometry and then for individual light sources. Note that the detectors are now interconnected and we thus are mostly interested in the order of clicks at the detectors.

The general expression for the second-order quantum correlation of fields A and B at times t_1 and t_2 reads

$$g_{AB}^{(2)}(t_1, t_2) = \frac{\langle \hat{E}_A^-(t_1) \hat{E}_B^-(t_2) \hat{E}_B^+(t_2) \hat{E}_A^+(t_1) \rangle}{\langle \hat{E}_A^-(t_1) \hat{E}_A^+(t_1) \rangle \langle \hat{E}_B^-(t_2) \hat{E}_B^+(t_2) \rangle} \quad (258)$$

For classical fields, this reduces to the Intensity-Intensity correlation function:

$$g_{\text{class}}^{(2)}(t_1, t_2) = \frac{\langle I_A(t_1) I_B(t_2) \rangle}{\langle I_A(t_1) \rangle \langle I_B(t_2) \rangle} \quad (259)$$

An important example is the intensity autocorrelation function of a stationary classical field

$$g_{\text{class}}^{(2)}(\tau) = \frac{\langle I(t+\tau) I(t) \rangle}{\langle I(t) \rangle^2} \quad (260)$$

It is straightforward to show that any classical light field must obey $g_{\text{class}}^{(2)}(\tau) \leq g_{\text{class}}^{(2)}(0)$ and $g_{\text{class}}^{(2)}(0) \geq 1$. This is not necessarily true for the autocorrelation of a quantum state of light $|\Psi\rangle$

$$g_{\text{QM}}^{(2)}(\tau) = \frac{\langle \hat{E}^-(t) \hat{E}^-(t+\tau) \hat{E}^+(t+\tau) \hat{E}^+(t) \rangle_{\Psi}}{\langle \hat{E}^-(t) \hat{E}^+(t) \rangle_{\Psi} \langle \hat{E}^-(t+\tau) \hat{E}^+(t+\tau) \rangle_{\Psi}} \quad (261)$$

While this expression is in general difficult to evaluate, we can get a good understanding of some general properties by evaluating it for $\tau = 0$ and a single-frequency mode. Here the correlation function becomes:

$$|g_{\text{QM}}^{(2)}(0)| = \frac{\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \rangle}{\langle \hat{a}^\dagger \hat{a} \rangle^2} = \frac{\langle \hat{n}(\hat{n}-1) \rangle}{\langle \hat{n} \rangle^2} \quad (262)$$

Evaluating this expression for a single photon Fock state $|\Psi\rangle = \hat{a}^\dagger |\text{vac}\rangle = |1\rangle$ we immediately see that

$$|g_{\text{QM}}^{(2)}(0)| = \langle 1 | \hat{n}(\hat{n}-1) | 1 \rangle = 0 \quad (263)$$

Single-photon states of light thus exhibit “anti-bunching”, a purely quantum phenomenon that cannot be described in classical coherence theory. Other examples that can be readily verified by the reader:

| STATE | $g^{(2)}(0)$ | COMMENT |
|---------------------------------|--------------|--|
| FOCK-STATE $ n\rangle, n = 1$ | 0 | Perfect Anti-Bunching (one photon at a time) |
| FOCK-STATE $ n\rangle, n > 1$ | $1 - 1/n$ | Anti-Bunching |
| COHERENT STATE $ \alpha\rangle$ | 1 | Uncorrelated (a random stream of photons) |

| | | |
|--|--|---|
| <p>THERMAL STATE</p> $\rho = \int f(\omega) \sum_n \frac{1 - \exp(-\frac{\hbar\omega}{k_B T})}{\exp(\frac{n\hbar\omega}{k_B T})} n(\omega)\rangle\langle n(\omega) d\omega$ | $1 + g^{(1)}(0) ^2$ | Bunching of photons of the same frequency (the more so the more narrowband) |
| <p>SQUEEZED STATE (DEGENERATE PDC)</p> | $g^{(2)}(0) = 3 + \frac{1}{\langle n \rangle}$ | Super-Bunched (photons always appear in correlated pairs) |

One can also, quite easily show, that $g_{QM}^{(2)}(\pm\infty) = 1$ and that the transition from the center value to the edge value is related to the bandwidth of the source in question, or more specifically its lifetime.

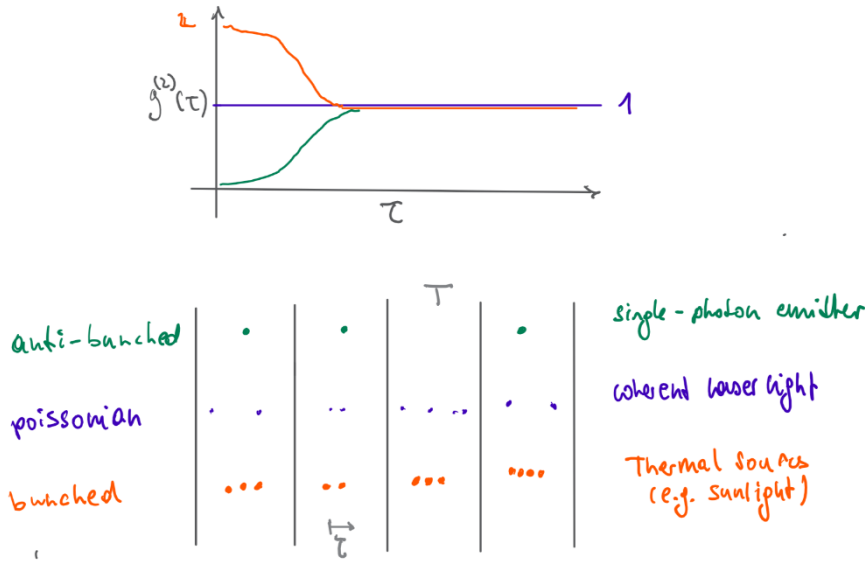


Fig. 60: Artists interpretation of photon arrival times in various sources, resulting in different $g^{(2)}$.

The Hanbury-Brown-Twiss experiment can thus be used to measure the “single-photon-ness” of a light source is the gold standard for this kind of characterization.

A 1.4 Quantum Interference and the Hong-Ou-Mandel-Effect

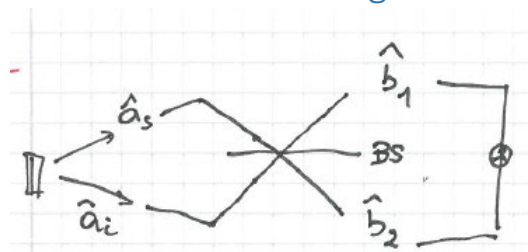


Fig. 61: HOM Interference on a Beam Splitter

Paul Dirac noted, that first-order interference effects can be thought of as each photon interfering with itself. Hong, Ou and Mandel experimentally showed that this is not the only form of interference we can observe. Let’s consider what happens when single photon states are incident from each of the two input ports of a 50:50 beam splitter. Their state is:

$$|\Psi\rangle_{s,i} = |1\rangle_s |1\rangle_i = \hat{a}_s^\dagger \hat{a}_i^\dagger |\text{vac}\rangle \quad (264)$$

Claim: If we place a detector in each of the two output ports $\hat{b}_1^\dagger \hat{b}_2^\dagger$, then there will be no simultaneous detections. To verify this we can either calculate directly the correlation function for the input state $\langle \Psi | \hat{n}_{b_1} \hat{n}_{b_2} | \Psi \rangle$, or (faster) express the input state in terms of the detection modes, i.e. we replace

$$\begin{aligned} \hat{a}_s^\dagger &\rightarrow \frac{1}{\sqrt{2}}(\hat{b}_1^\dagger + \hat{b}_2^\dagger) \\ \hat{a}_i^\dagger &\rightarrow \frac{1}{\sqrt{2}}(\hat{b}_1^\dagger - \hat{b}_2^\dagger) \end{aligned} \quad (265)$$

Substituting these expressions into the input state, we find that the terms leading to a joint detection at detectors 1 and 2 will cancel

$$|1\rangle_s |1\rangle_i \rightarrow \hat{b}_1^{\dagger 2} + \hat{b}_1^\dagger \hat{b}_2^\dagger - \hat{b}_2^\dagger \hat{b}_1^\dagger - \hat{b}_2^{\dagger 2} |\text{vac}\rangle = |2\rangle_1 |0\rangle_2 - |0\rangle_1 |2\rangle_2 \quad (266)$$

Both photons will leave the beam splitter bunched into couples. As a result, there will be no coincident detection. This can be seen as destructive interference of transmitted and reflected photon pairs²⁸, known as Hong-Ou-Mandel interference (see *PRL 1987, 59 2044*). HOM interference is a valuable tool in quantum information processing, and quantum optics – we will encounter it again at many instances.

8.3.4 HOM-interference for phase sensing

The output states produced in the HOM interferometer,

$$|2\rangle_1 |0\rangle_2 + |0\rangle_1 |2\rangle_2 \quad (267)$$

or more generally states of the form

$$|N\rangle_1 |0\rangle_2 + |0\rangle_1 |N\rangle_2 \quad (268)$$

are a useful tool in Metrology (so-called NOON states). To see why, let us feed such a NOON state into a MZ interferometer where one arm experiences an additional phase shift (caused e.g. by a small displacement of one of the mirrors in the interferometer). Using the phase shift operator defined in the previous Lectures, we note that this transforms the NOON state as:

$$|N\rangle_1 |0\rangle_2 + |0\rangle_1 |N\rangle_2 \rightarrow |N\rangle_1 |0\rangle_2 + e^{iN\phi} |0\rangle_1 |N\rangle_2 \quad (269)$$

Depending on the photon number, the phase factor is multiplied to $N\phi$. This feature is known as *super-resolution*. Compared to single-photon states and coherent states, NOON states allow measuring phase shifts with better precision; unfortunately, they are notoriously hard to produce for $N>2$ in experiment.

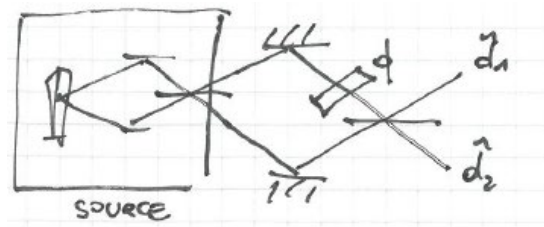


Fig. 62: Using HOM interference to produce a NOON state ($N=2$)

8.3.5 An interpretative note on HOM

After taking a closer look at the equations, we can also infer the HOM-effect from a hand waving explanation. Assume there are two photons, which are incident on a balanced beam splitter from its two-input port. There are in total four options, as each photon may or may not get reflected:

²⁸ Note that this is a consequence of the commutation operator relationships between the creation operators for bosons - what do you expect would happen if we replaced the photons with electrons?

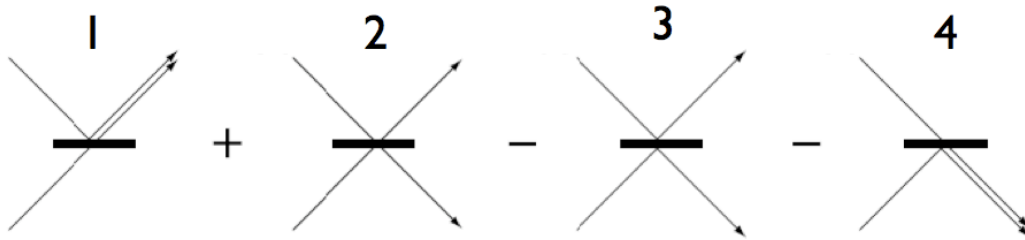


Fig. 63: Two-Photon Interference representation of the HOM-Experiment. If path 2 and 3 are indistinguishable, then they will interfere destructively and only photon pairs can be observed.

Note that the sign in front of the little pictures correspond to a 0 or π phaseshift. The phaseshift is particular noteworthy for case 3, here they correspond to the phaseshift accumulated during reflection. A grossly simplified explanation is the reflection on a denser medium; nevertheless the phase difference of π case 2 and 3 is universal, as it is related to the unitarity of the mixing operation.

Only indistinguishable photons show interference (that's why single photon interference is so easy to detect; one photons is necessarily indistinguishable from itself). In the HOM experiment this is only the case for paths 2 and 3 and only if the input photons are indistinguishable itself. As they interfere and have opposite sign, the modal contributions from option 2 and 3 thus cancel each other and the result of two-photon interference is such, that the two photons will either both go up or both go down (in the sense that they emerge in a superposition of 1 and 4; i.e. they go up and down simultaneously and their path is just decided up if you detect one photon). If you however detect one photon in, say, the upper branch then you know the other one is there as well and vice versa. This is example of entanglement.

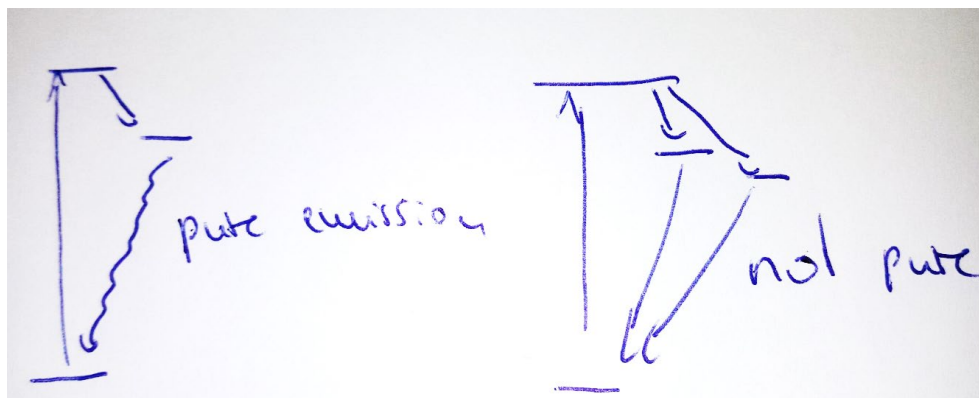


Fig. 64: Example of a level diagrams of single molecules or QDs, which emit pure and mixed states. A molecule or a QD is excite and decays non-radiatively into a single or two upper states, which emit fluorescence light. If only one path is possible the emitted state is pure (i.e. emission always in the same mode), if two paths are possible with a certain probability then the emitted states are mixed.

Summary: if two indistinguishable photons meeting on a balanced beamsplitter they will leave the beamsplitter as a pair. If a pair is impinging on a beamsplitter they will go their separate ways. The beamsplitter can be considered something like the civil registry office for photon pairs.

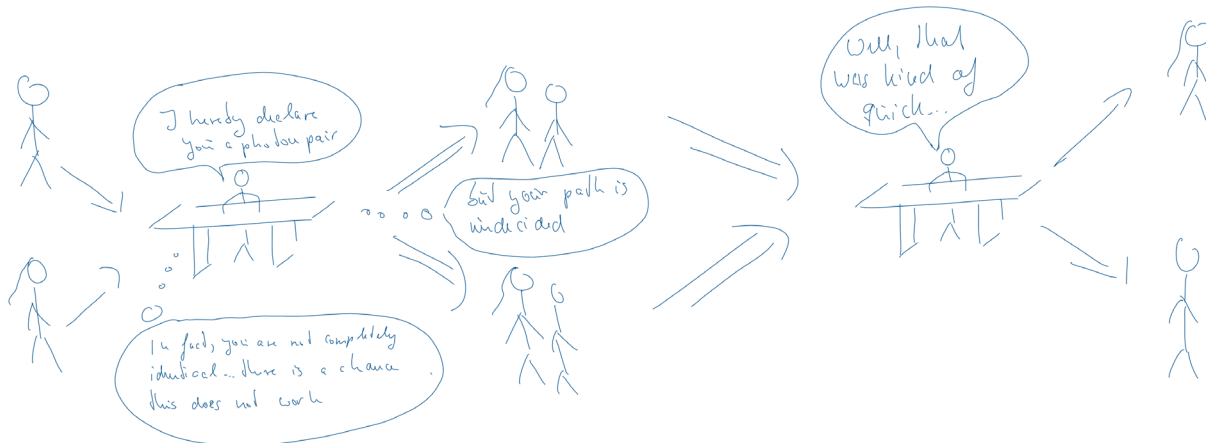


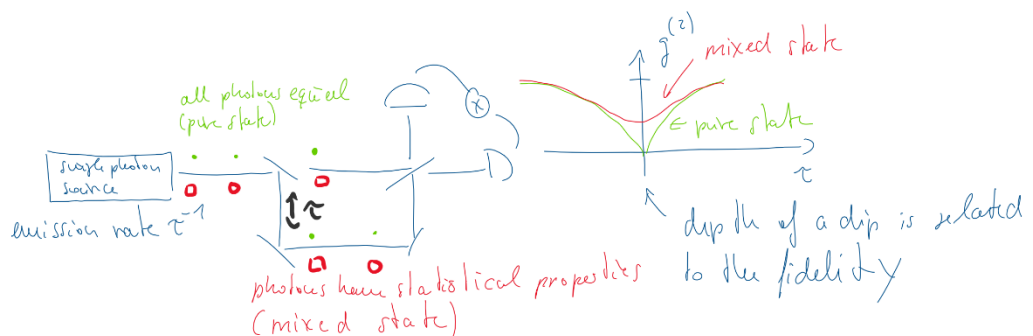
Fig. 65: The adventures of a beam splitter. Parts one and two. To pair and not to pair. Sorry for the silly joke. I could not resist.

A 1.5 Applications of HBT and HOM

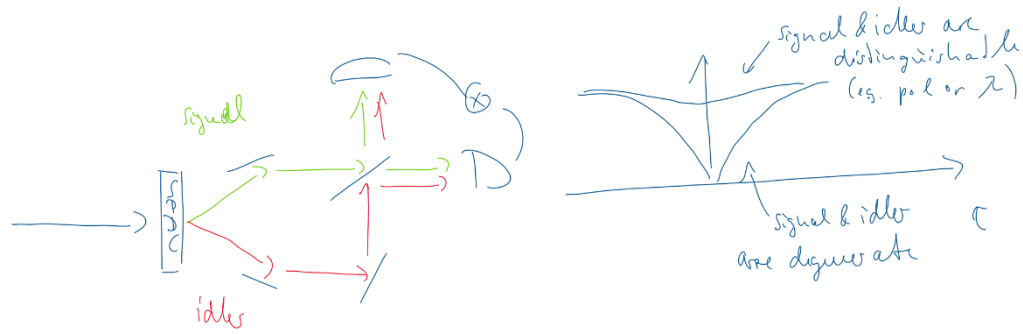
Both the HBT and the HOM-Effects are not just experiments, whose results can only be interpreted by a Quantum theory of light, as they can be conducted with any light source, quantum, or not. As such, their result gives a fairly comprehensive measure of the quality of any source as a single photons source. They can therefore be used to define the characteristics of a photon source, from a measurement point of view. They are the foremost experimental tool to characterize the quantum properties of light sources and their results are typically used directly to classify quantum light sources.

In this regard

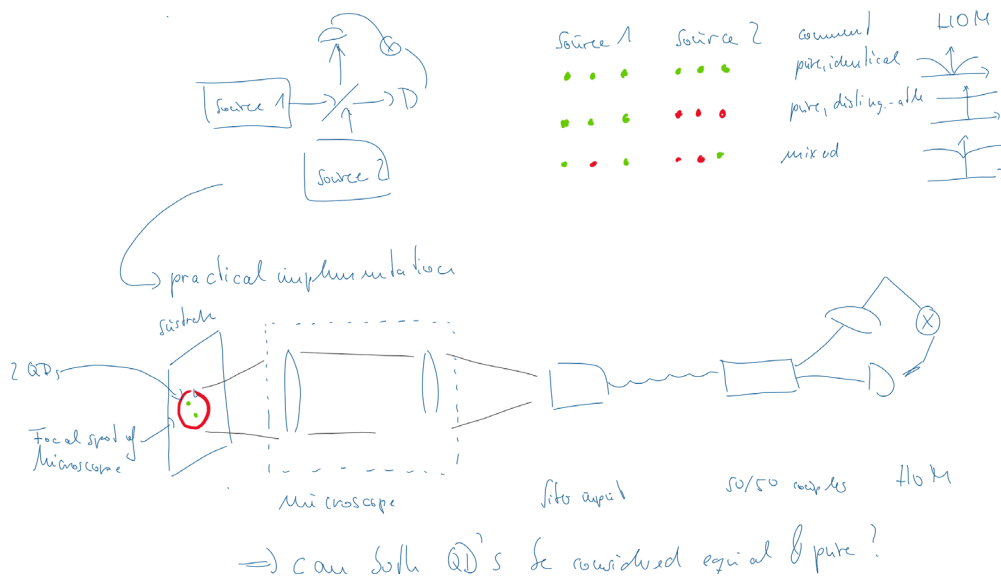
- HBT gives us information on the single photon character of a light source; i.e. the quantity $1 - g^{(2)}(\tau = 0)$ is an answer to the question, if a light source is emitting single photons and single photons only. HBT, is however, agnostic to the purity of the state of light. This means, in particular, that a light source may randomly emit in distinguishable modes k , as long as each of them is a $|1_k\rangle$ -mode you will see a perfect HBT-dip.
- The depth of the HOM-dip gives us information on the (in-)distinguishability of two light sources or two beam paths, which emanate from a the same light source.
 - o if HOM is used with a single photon source (HBT-checked) and two consecutive photons the depth of the HOM-dip will give you information of the purity of the emitted state



- if HOM is used with a two-photon source and one photon in each mode the depth of the dip tells you something about the similarity (indistinguishability of the photons)



- if HOM is used with two distinct light sources, it tells us, if these two light sources emit the same quantum state of light all the time. In particular this rules out the possibility of the emission of mixed states.



- The width of the HOM-dip is directly related to the coherence length of the light source, i.e. the inverse of its spectral bandwidth, i.e. the photon lifetime

While the implementation with a beam-splitter and two separate detectors is the canonical one, note that both HBT and HOM just require any kind of mixing process for two distinct modes and two coincidence detectors. In the simplest case the mixing may be simply achieved by diffraction and the detector may just be a single photon sensitive camera.

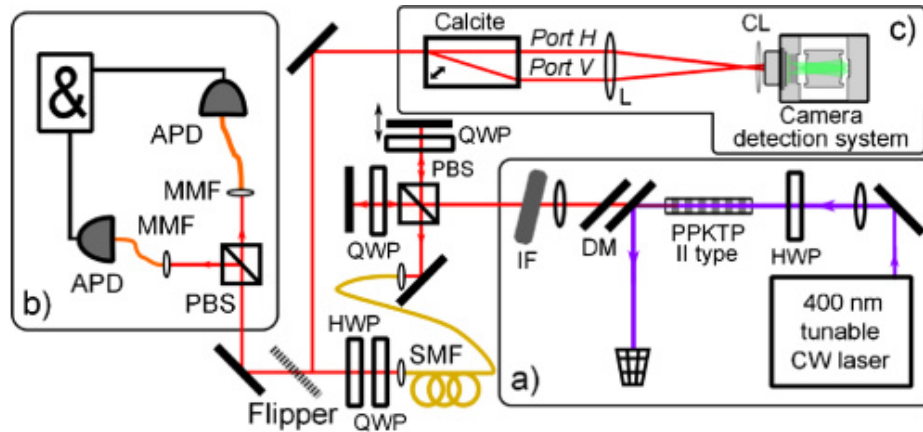


Fig. 66: Camera based HOM with two-photon source. Part a) generates two photon states with SPDC. The central part is used to define the polarization of the state (QWP), the delay τ (delay stage), and select a single spatial mode (fiber). Part b) is a classical HOM setup. Part c) is a camera based HOM. Jachua et al., *Opt. Lett.* 40, 1540 (2015)

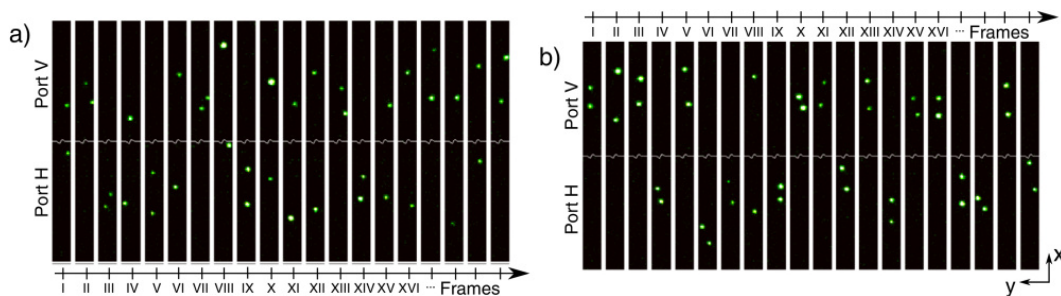


Fig. 67: Results: On the left the result for HH, VV and HV in the case of large τ (i.e. distinguishable beam paths and thus distinguishable modes) and HV in the case of $\tau = 0$ (i.e. polarization distinguishability) and on the right side for HH and VV for $\tau = 0$. Jachua et al., *Opt. Lett.* 40, 1540 (2015).

A 2 Single Photon Resolving Detectors: an Overview

We shall have a (non-exhaustive) overview over some types of single-photon detectors, which are used oftentimes in experiments in quantum photonics.

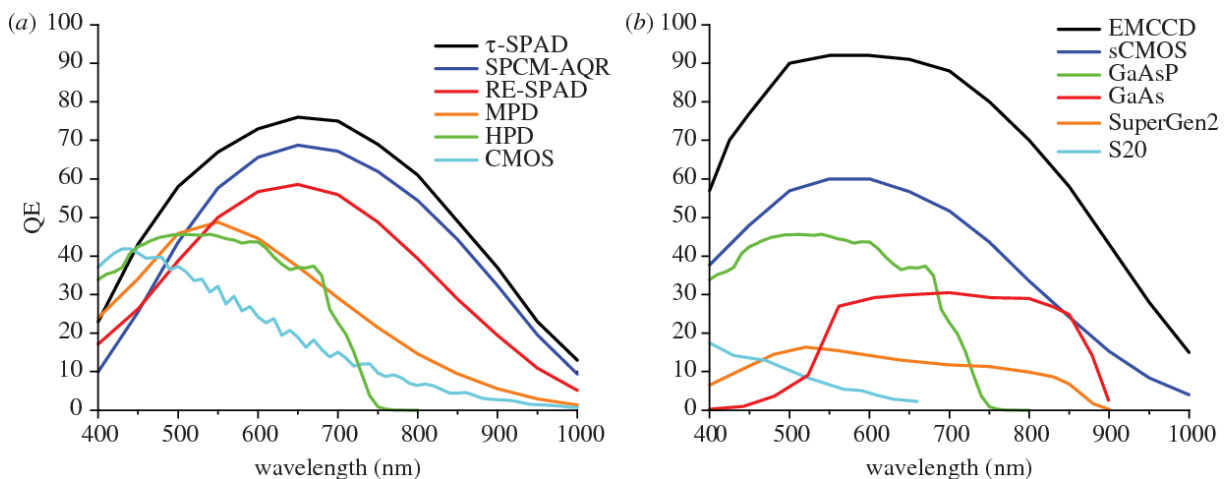


Fig. 68: Quantum efficiency (i.e. the probability of detection of a photon) of some common VIS platforms. (Left) Single-Channel detectors. (right) Pixelated Detectors. Source: Michalet et al., *Phil Trans R Soc B* 368, 20120035 (2012). Note the data is quite old, some of the values have improved drastically (e.g. BT-sCMOS now with up to 85% QE).

A 2.1.1 Single Channel Detectors (Bucket Type)

These are historically the first single-photon-capable detectors. They are typically stand-alone devices with superior quantum efficiency, count-rate and very low dark-count rates. They can be thought of as click-detectors; i.e. they produce a measure current or voltage spike upon the detection of a photon.

The spike is typically very short ($\sim 10 \dots 100$ ps), i.e. the time of detection of a photon can be measured very accurately and thus they are oftentimes combined with advanced correlation electronics or timestamp-electronics. However, the detection of an event is typically accompanied by a dead-time; i.e. a certain amount of time that is needed to restore the system in its single-photon sensitive state. During this time the system is generally incapable of measuring photons. This dead-time is typically much larger than the timing resolution ($\sim 1 \dots 100$ ns).

For all detectors temperature is an issue. The detectors are in a thermal equilibrium and thus subject to internal black-body radiation. They necessarily cannot discern between the detection of a photon from an external source or from the reabsorption of a thermal photon. This is of particular influence for detectors sensitive in the IR; these typically have to be cooled; not so much for detectors for the visible. Thermal dark-count rates typically scale exponentially with temperature (Boson-Distribution!). Thermal dark-count rates also typically scale with the quality of detector materials; they are necessarily the lowest for Si-based detectors.

8.3.6 Photomultiplier Tubes (PMTs)

In a photomultiplier tube the light is incident on a photocathode. There it will ionize an electron by supplying the work function to an atom (upper limit to wavelength!). The electron is then accelerated to a positively charged electrode (dynode). When it hits the electrode it has acquired so much kinetic energy, that it generates more than one secondary electron. The process is repeated with consecutively higher charges electrodes (dynodes) and a near-exponential increase of the number of electrons. The last electrode is the anode, here the electron cloud is measured as a current spike.

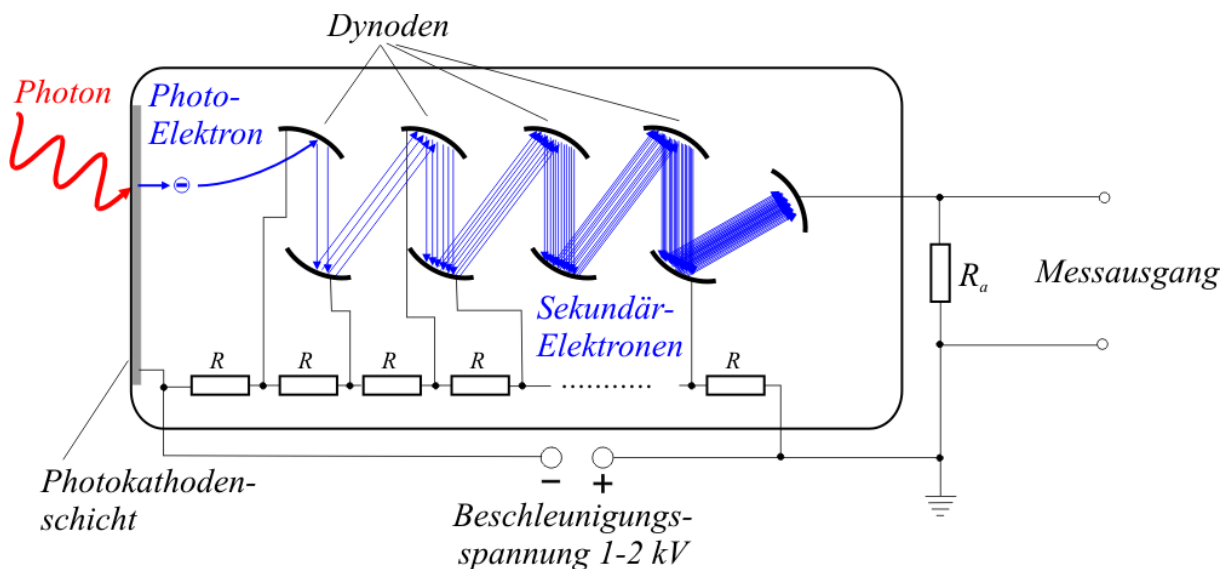


Fig. 69: Schematic of a PMT. Source: Wikipedia.

8.3.7 (Silicon) Avalanche Photo Diode (SPAD)

An avalanche Photodiode is a slightly modified photodiode, which is operated in reverse bias voltage. The applied voltage is very high ($\sim 50 \dots 100$ V) and just short of the breakthrough voltage. The active area is composed of non-doped semiconductor; upon the absorption of a photon an electron-hole pair

is created. Due to the bias voltage the electron is drifting towards the p-n-junction exciting new electron-hole pairs along the way. This does particularly happen in the region of the p-n-junction due to the strong electric fields there.

As a result there is an avalanche of free electrons and holes which drift through the SPAD, every increasing the number of free electrons and holes. At the same time the avalanche leads to a current spike in the bias electronics. This current spike can be detected, counted and timed. The avalanche, however, needs to be actively terminated by quickly switching off the bias voltage and waiting for the recombination of all electron-hole pairs until it is switched back on.

The process works on all standard semiconductor platforms, e.g. Silicon (400-1100 nm), InGaAs (900-1700/1900 nm), or Ge (similar to InGaAs) or any other kind of exotic semiconductor. The resulting diodes are necessarily larger than, e.g. CCD or CMOS devices, both the control electronics (avalanche termination) as well as the high-voltage feed lines need quite a lot of space.

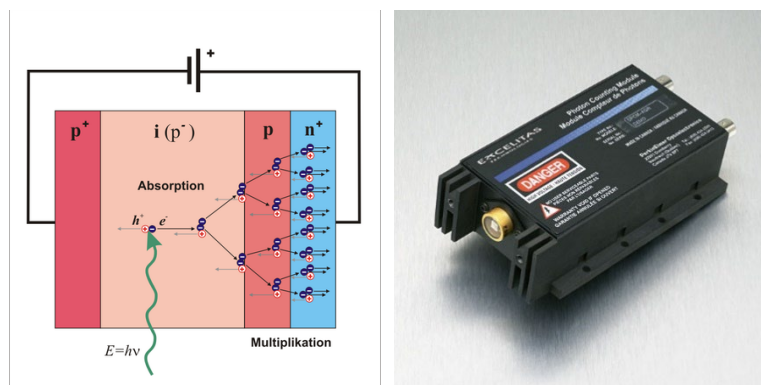


Fig. 70: (left) Schematic of a Silicon-APD. Source: wikipedia. (middle) A SPAD mounted on a photon counting module.

8.3.8 Superconducting nanowire detectors

In these detectors a superconducting nanowire is patterned to fill an area on the chip almost completely. The wire is cooled below the superconducting threshold and an external circuit drives a current through the superconductor, that is just short of the critical current at which superconductivity breaks down.

If a photon is absorbed by the superconducting wire, it creates a local hotspot with reduced critical current and the wire becomes ohmic. The change of voltage on the external circuit can be measured as a voltage spike. Due to the relative impedances of the hotspot and the amplifier the hotspot will typically cool by itself and the device automatically goes back into the superconducting state. Because the wire is very small (typically 50 nm) it cools back down into the ready-state in a matter of a few nanoseconds.

These devices are intrinsically broadband, as they rely on thermal absorption. They are very fast (~50 ps timing resolution) and have a comparatively short dead-time (cooling and inductive recycling). The operation is quite simple and many superconductors are suitable. They still require cryogenic operation but they work at 4 K, which is now attainable with closed-cycle cryostats, which as of 2018 have the size of a small refrigerator and can be run from an ordinary power plug. The QE can be tuned by application of an optical cavity to enhance the probability, that photons with certain wavelengths are indeed absorbed.

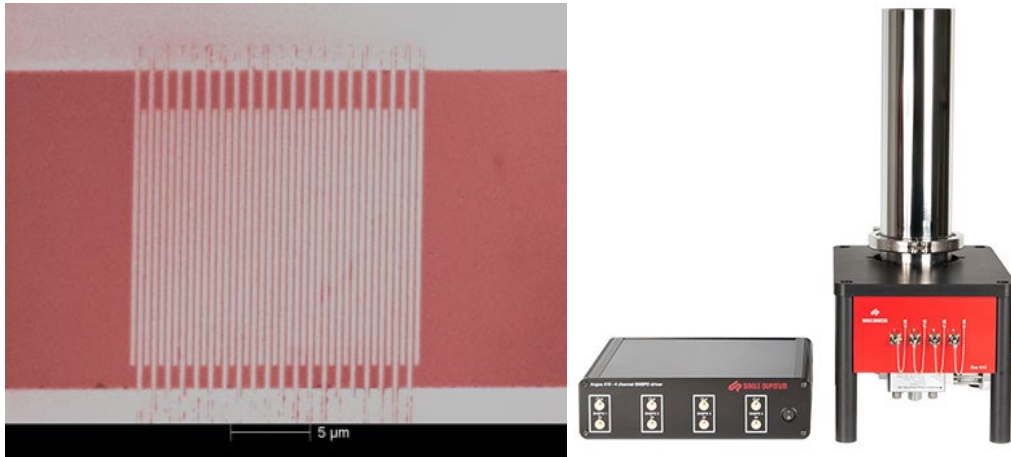


Fig. 71: (left) A superconducting nanowire from a superconducting nanowire detector (Wikipedia). (right) A contemporary 4K closed-cycle cryostat for multiple Superconducting nanowires complete with timing and correlation electronics. Source: Quantum Design

8.3.9 Transition edge sensors

Transition edge sensors can be thought of as highly sensitive thermometers based on the thermoresistive effect. They measure the change of conductivity of a metal as a photon is absorbed and it thus changes its temperature in a minuscule manner. The only way to reach enough sensitivity for a system which is still sufficiently small (which means you cannot make the thermal system arbitrarily small or arbitrarily well insulated), is to use the superconductor and stabilize it on the transition edge between the superconducting and conducting state. On this edge the resistivity grows by a few orders of magnitude over a temperature change in the region of milli- or microkelvins.

The system is operated with a current flowing through it and the change of resistivity is translated into a measurable change of voltage at the receiver. As the change of resistivity is (nearly) proportional to the amount of absorbed energy one can actually count photons or even (to a certain extent) measure their energy. The transition edge technique does, however, only work for low-temperature superconductors at roughly ~ 100 mK; they are thus extremely expensive and hard to maintain and operate. They are not very fast with dead-times in the range of tens of microseconds and very little timing resolution (order of a single microsecond) due to the larger size of the thermal bath.

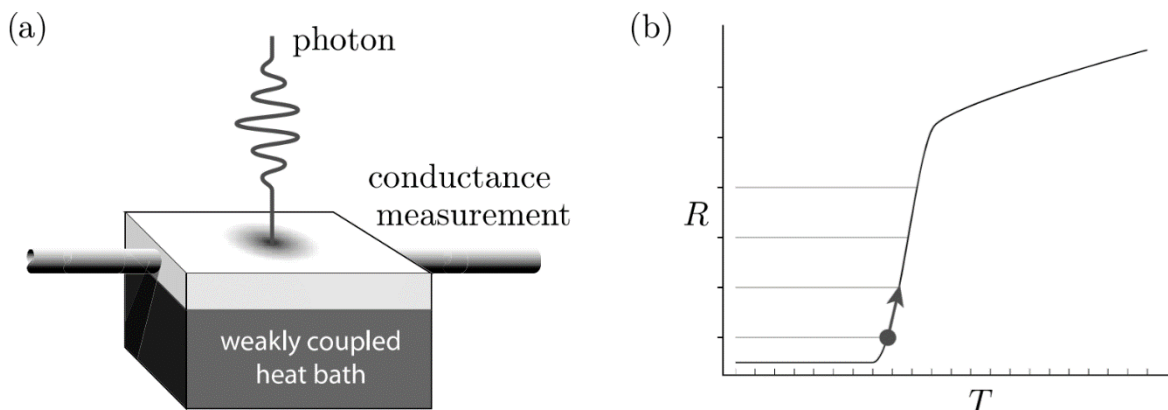


Fig. 72: (Left) Schematic of a TES. (right) Resistivity of a superconducting materials as a function of its temperature close to the transition edge. Horizontal bars denote changes of resistivity as individual photons are absorbed. (Lovett & Kok, Quantum Information Theory)

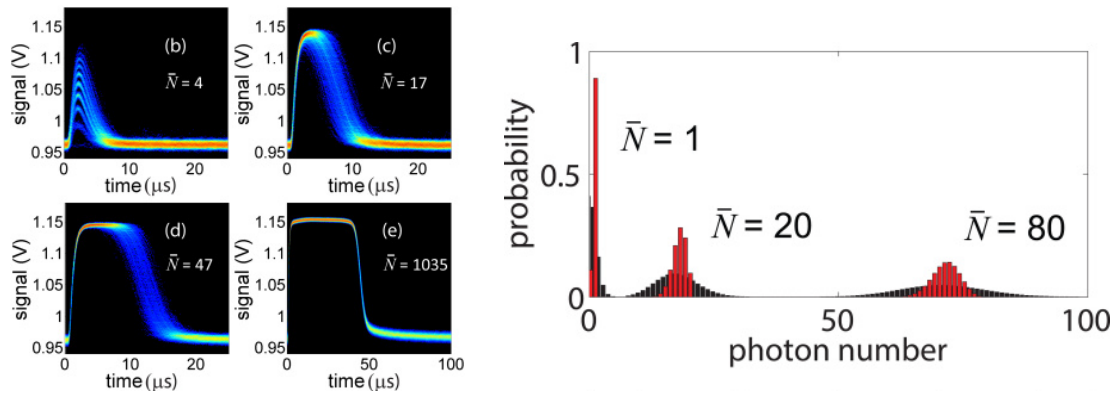




Fig. 73: (left) Response curves of a TES-sensor for illumination with coherent states of light with mean photon number \bar{N} . (right) Red: Readout probability of a TES, when illuminates with Fock-States. Black: Readout-Statistics, when illuminated with a coherent state. (Gerrits et al, Opt. Express 20, 23798 (2012))

8.3.10 Summary

| |  |  |
|------|---|---|
| PMT | Inexpensive (k€), room temp, few dark counts | low QE, large, |
| SPAD | inexpensive (10 k€), room temp, few dark counts (VIS), good timing | mediocre QE, high dark counts (IR), long dead-time |
| SND | Few dark counts, broadband, high QE, good timing, short dead-time | 4K-cryo, bulk, expensive (100 k€) |
| TES | Few dark counts, broadband, high QE, photon counting | bad timing, long dead-time, mK-Cryo, super-bulky, super-expensive (1 M€) |

A 2.1.2 Pixelated Detectors

Pixelated detectors (i.e. camera-like detectors) have seen the most dramatic development in the last decade or so. Driven by the development of consumer electronics and LIDAR-applications for autonomous mobility, they have improved in terms of (electronic) shutter speed, quantum efficiency, noise, pixel size, dynamic range, etc. As a matter of necessity these development are mostly limited to Silicon-based devices as here the small quantum-photonics market profits from the multi-billion dollar semiconductor industry. Consequentially this shall be discussed here. In general, there are extensions available in the NIR, based on, e.g. the InGaAs material system, and in the UV / X-ray range but they are typically much more expensive and much less developed.

While pixelated detectors in general do not have the timing resolution of bucket detectors with dedicated timing and counting electronics, these general make up this shortcoming by giving the ability to conduct measurement in many channels (up to millions) at once.

With these system you can not only measure if and when photons have arrived but you can extract more information from them, by coupling them with an optical system, such as an objective or a spectrometer. Thus you can differentiate between multiple modes of a system (e.g. k-vectors, positions, wavelengths, polarizations, etc.) and can therefore carry out experiment with multimode quantum

systems. The modal selection can then be done in a post-selection style and does not need to be implemented in the optical system (e.g. by using single mode fibres or small apertures). Moreover this scheme can be used to correlate photons in different modes (e.g. wavelength entangled photon pairs).

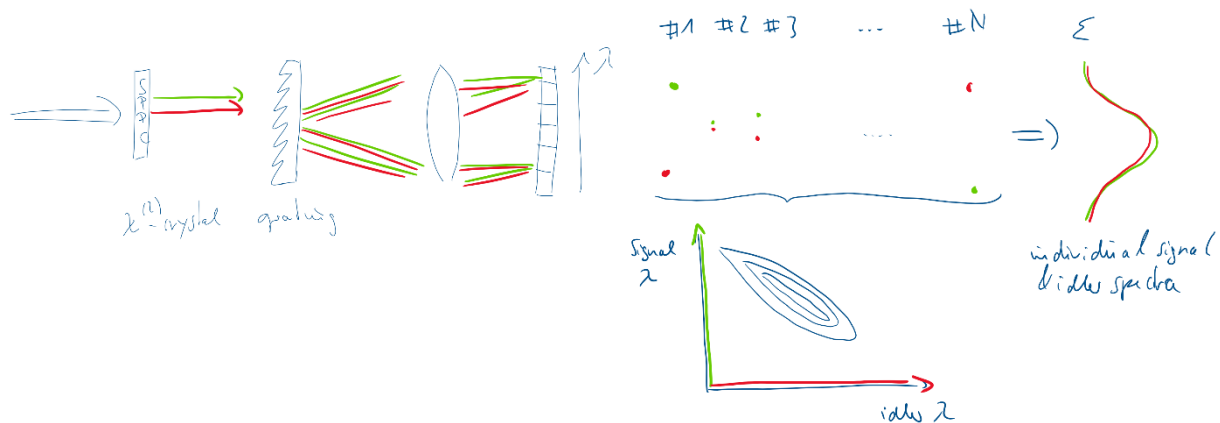


Fig. 74: Modal correlation imaging with a 1D-pixelated detector for a degenerated SPDC-based broadband light source. Both idler (red) and signal (green) are analysed with a spectrometer and always appear as frequency-entangled, anticorrelated pairs (due to energy conservation).

Moreover pixelated detectors can be used as a kind of poor-man's photon-counter. If a particular mode is imaged onto a set of pixels and each pixel has a certain probability of detecting the photon, then it is highly unlikely that two photons will hit the same detector but it will more likely hit two different detectors; the same is true for three, four, etc. Note that such a scheme of pixel counting does however require a very high QE, which is given for some modern systems, such as EMCCD or BT-sCMOS. I.e. the probability that a click-detector with N pixels will have n pixels clicking, if the detector is illuminated with a Fock-State $|n\rangle$ and each pixel has the same QE η is given by:

$$p(n|n) = \binom{N}{n} \frac{\eta^n n!}{N^n} \xrightarrow{N \rightarrow \infty} \eta^n \xrightarrow{\eta \rightarrow 1} 1.$$

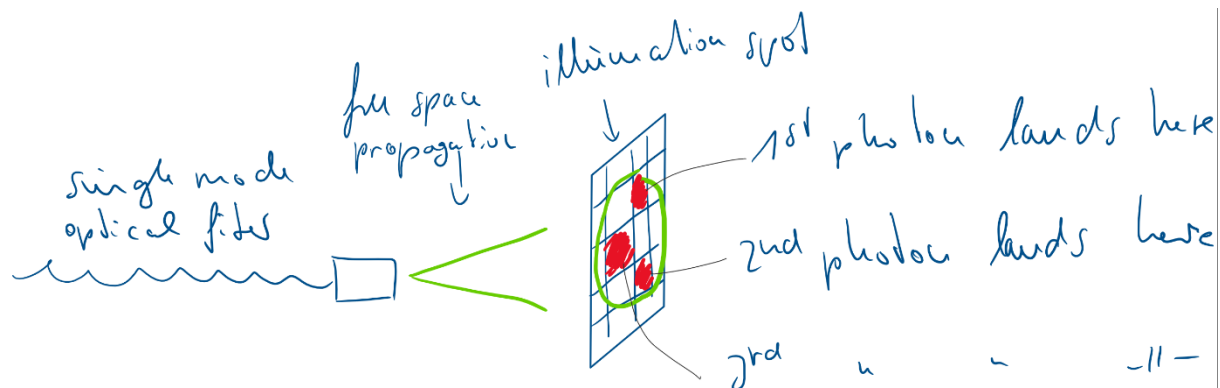


Fig. 75: Click-Detector-Multiplexing as a scheme for Photon-Counting.

8.3.11 Derived from CCD or CMOS cameras

This is the classic class of pixelated detectors. They have emerged from standard scientific cameras at the point, where their QE has started to become substantially higher than 30 %. There are various classes of single-pixel sensitive cameras in the market, which differ by pixel size, pixel number, QE-

spectrum, readout noise, area fill factor, readout speed, shutter speed, etc. Of course they also differ heavily in terms of price.

The most important classes are sCMOS, EMCCDs, and ICCDs. As all of these devices are improving in terms of their important parameters, basically every year, I shall not attempt to compare them here. In general, however, sCMOS tend to be the least sensitive, have the most pixel and be the cheapest (~20 k€), whereas EMCCDs and ICCDs tend to be more sensitive (+10% QE), have fewer pixels (~1/4) and be more expensive (~ 50 – 100 k€).

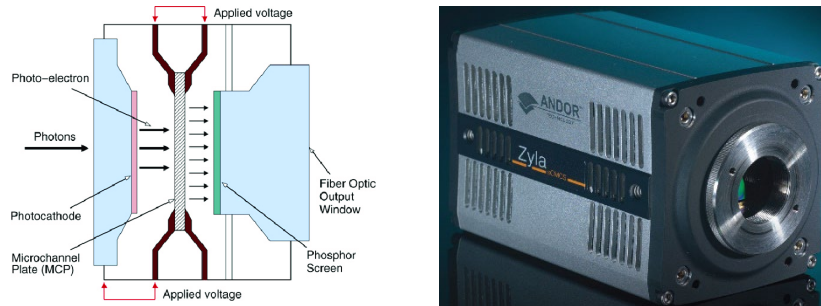


Fig. 76: (left) Operation scheme of an image intensifier, placed before the CCD of an ICCD. (right) Image of a cooled laboratory grade sCMOS (Andor).

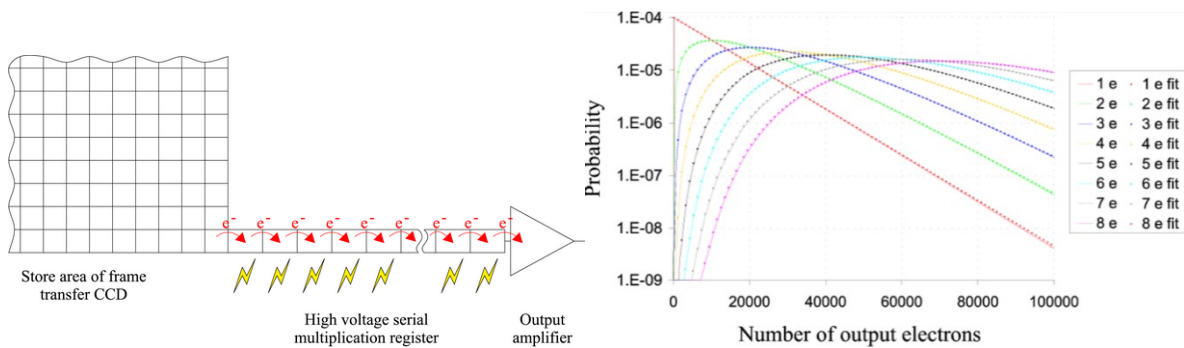


Fig. 77: Operation scheme of an EMCCD-Camera. Light creates free electrons in a CCD-register. Before readout each cell is passed through an series of multiplication registers (the semiconductor analogue of a PMT-Dynode) with a low multiplication probability. (right) The readout value for zero, a single and more photons can be clearly distinguished. Source: Wikipedia.

Each of these classes has their own approach to achieve a high QE. sCMOS systems mostly rely on the development of ever improved CMOS-detectors and processes derived from the consumer-driven semiconductor industry. ICCDs are derived from 2nd-generation night vision devices and use MCP-based image intensifiers to enhance the number of photons before detection. EMCCDs are based on later-generation of night-vision devices. They enhance the electronic sensitivity of the CCD-chip by placing an electron multiplication register between the CCD and the readout electronics. The approaches can in principle be combined to a certain degree.

Nevertheless each of these cameras have in common that they do not have timing resolution. I.e. they can be armed at an instance in time, left to their own devices for a certain exposure time and then read out. At which point in time of the exposure the photon was absorbed cannot be measured. Also the order in which multiple photons have been absorbed in one frame is not measurable. As their frame rate is limited (and transfer of frame to a data acquisition device is time consuming) one can only make a trade-off between timing accuracy (short exposure times, camera disarmed most of the time) and duty cycle (long exposure times, camera armed most of the time).

8.3.12 Derived from SPAD-devices

Newer developments attempt to mitigate this issue by combining the timing resolution of SPAD devices with the scalability of the CMOS process to create array SPAD (aSPAD) cameras. These devices are basically a lot of SPADs crammed on a single Si-chip, together with read-out circuitry, pixel based timing and photon-counting electronics.

Compared to camera-based devices these have a greatly reduced number of pixels and reduced size of the active area (typically in the range of 50x50 pixels as opposed to >5 MPixels for sCMOS, < 5% active area), due to the large size of the high-voltage, readout, and pixel-based timing-electronics. In the state of the art, they are also limited in their spectral response as the application of detection and electronics on the same chip, imposes a certain regime of spectral response of the QE. As these devices are under active development, it may be expected that the number of pixel will cross the 100x100 pixel range soon. The application of multi-wafer bonding techniques and micro-lenses will mitigate both the issues with the active area and the spectral response of the QE-curve.

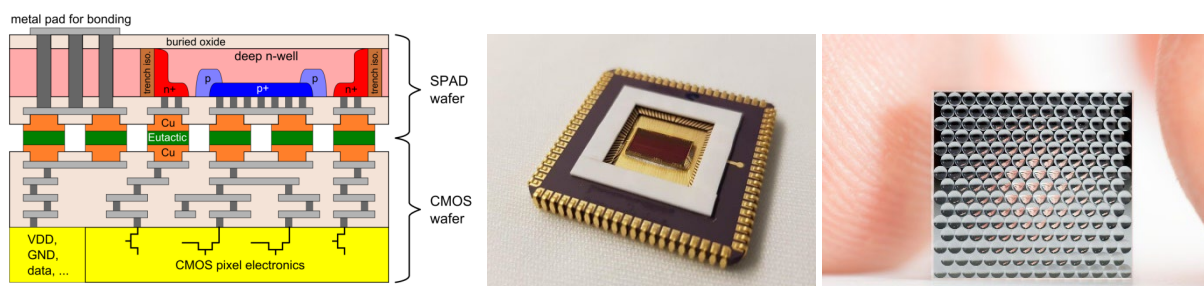


Fig. 78: (left) Schematic cross-section of a 3D-integrated SPAD-pixel with readout and timing electronics based on a CMOS process. (middle) A 2x192 aSPAD-array. (right) A microlens-array, which is used to trade numerical aperture for QE. All images: Fraunhofer-Leitprojekt QUILT.

aSPADs are, however, not just a new kind of camera. The availability of time-tags for each and every photon, which arrives at the camera opens perspectives for entirely new experiments of multimode correlation, high-order-correlation, and correlation based imaging (images are necessarily high-order-modal entities). They also, in principle, do not suffer from the timing accuracy / duty cycle trade-off as CCD and CMOS-based cameras do. They may thus help to establish hyper-entanglement measurement schemes, high-qudit based communication and sensing schemes, and quantum multimodal imaging. On-Chip correlation electronic can be used for enhanced data-compression and quantum compressed sensing schemes.

A 3 Three-Photon Processes and SPDC

In this chapter we will go through the Basis of nonlinear optical process of second order (i.e. three photon-processes) at a high pace and then focus on the special case of SPDC. To keep the discussion brief and simple we'll stick to classical Maxwell-Theory to end up at coupled-wave equations between pump, signal and idler, and then resort to the tried and tested method "just sticking a hat over all the development coefficients" to make the transition to the quantum realm. While this is conceptually total bullshit, it still gives the correct results. Alas, one former example of the fact, that you can clean dishes, even if you use dirty water.

A 3.1 Fundamentals of Three Photon Processes

In order to make the transition from the microscopic to the macroscopic Maxwell-Theory in optics one introduces the Polarization field \mathbf{P} , which summarizes the contributions from all induced dipoles of the matter a light field \mathbf{E} propagates through. In the process there is typically a position, where the assumption is made, that the individual dipoles are excited proportionally to \mathbf{E} , leading to a linear relationship between \mathbf{P} and \mathbf{E} . Microscopically this is connected to the fact, that the dipolar contribution of matter is mostly rooted in the motion of bound electrons, which are in state of an energetic minimum, forming a local harmonic oscillator with $E \sim \delta x^2$ relation of potential energy E with out-of-equilibrium position δx . If the microscopic oscillators are to be modelled quantum-mechanically, this means that their Hamiltonian is of the form $\hat{\mathcal{H}} = \omega_0 \hat{a}^\dagger \hat{a}$, with $\hat{a} = \hat{x} + i\hat{p}$.

If the incident E -Field is strong enough, δx will grow large enough, such that the local oscillators become (slightly) anharmonic. If the deviation is not too large²⁹ we can expand the Energy-Position relation into $E = a\delta x^2 + b\delta x^3 + \mathcal{O}(\delta x^3)$. If the microscopic oscillators are to be modelled quantum-mechanically, this means that their Hamiltonian is of the form $\hat{\mathcal{H}} = \omega_0 (\hat{a}^\dagger \hat{a} + \varepsilon \hat{a}^{2\dagger} \hat{a} + \varepsilon^* \hat{a}^\dagger \hat{a}^2)$. If you look carefully and consider an interaction Hamiltonian you find that it describes three-photon interaction, of the kind required to create entanglement and two-mode squeezing.

In classical Maxwell-Theory this changes the $P(E)$ relation to:

$$P_\mu(t) = \varepsilon_0 \int_{-\infty}^{\infty} d\tau R_{\mu\alpha}^{(1)}(t-\tau) \cdot E_\alpha(\tau) + \varepsilon_0 \iint_{-\infty}^{\infty} d\tau_1 d\tau_2 R_{\mu\alpha\beta}^{(2)}(t-\tau_1, t-\tau_2) \cdot E_\alpha(\tau_1) \cdot E_\beta(\tau_2) \quad (270)$$

Note that this relation is valid in time-domain and the linear and non-linear contribution of the matter is collected in the response-functions $R^{(1)}$ and $R^{(2)}$, which are of tensorial. In a hand waving manner, the response functions tell you :

- the time delay of the impact of the material on the P -Field after a photon has arrived
- the strength (i.e. interaction likelihood) of such a polarization event
- the maximum temporal separation (i.e. correlatedness), that multiple photons must arrive with, in order to lead to a polarization event
- the behaviour with respect to the polarization of the respective photons

In optics we typically operate with modes, which are defined in the frequency-domain as opposed to the time domain, in which the above relation is written in. Of course, we transform the equations into

²⁹ This is generally a valid assumption, because if it becomes too large, then the potential well typically merges with the continuum, the electrons can start to move freely, the matter ionizes and is destroyed.

frequency domain by expanding the electric fields $E_\alpha(t)$ into their frequency components $E_\alpha(t) = \int d\omega \tilde{E}_\alpha(\omega) \exp(-i\omega t)$. Thus:

$$P_\mu(t) = \varepsilon_0 \int_{-\infty}^{\infty} d\omega \chi_{\mu\alpha}^{(1)}(-\omega_\sigma; \omega) \cdot \tilde{E}_\alpha(\omega) \exp(-i\omega_\sigma t) \\ + \varepsilon_0 \iint_{-\infty}^{\infty} d\omega_1 d\omega_2 \chi_{\mu\alpha\beta}^{(2)}(-\omega_\sigma; \omega_1, \omega_2) \cdot \tilde{E}_\alpha(\omega_1) \cdot \tilde{E}_\beta(\omega_2) \exp(-i\omega_\sigma t) \quad (271)$$

With the susceptibility tensors:

$$\chi_{\mu\alpha}^{(1)}(-\omega_\sigma; \omega) = \int_{-\infty}^{\infty} d\tau R_{\mu\alpha}^{(1)}(\tau) \cdot \exp(i\omega\tau) \\ \chi_{\mu\alpha\beta}^{(2)}(-\omega_\sigma; \omega_1, \omega_2) = \iint_{-\infty}^{\infty} d\omega_1 d\omega_2 R_{\mu\alpha\beta}^{(2)}(\tau_1, \tau_2) \cdot \exp(i\omega_1\tau_1) \exp(i\omega_2\tau_2) \quad (272)$$

Obviously, there are a lot of symmetry conditions, which have to be imposed on the response functions $R^{(1)}$ and $R^{(2)}$. This carries over to the frequency representations, i.e. $\chi^{(1)}$ and $\chi^{(2)}$. Some of these symmetry conditions are of fundamental nature, others can be imposed for specific wavelength combinations (e.g. $\omega_1 = \omega_2$), crystal symmetry classes and spectral features, a few of which shall be discussed now.

$R^{(1)}$ and $R^{(2)}$ have to be both vanishing for $\tau < 0$ for causality reasons, thus $\chi^{(1)}$ and $\chi^{(2)}$ have to be analytic with all poles in the upper frequency plane. This leads to a connection between the real and imaginary parts of $\chi^{(1)}$ and $\chi^{(2)}$, which themselves account for polarization and loss. These relations are called Kramers-Kronig relations and they basically mean, that loss and polarization are two side of the same medal. Furthermore, $R^{(1)}$ and $R^{(2)}$ both have to be real, thus $\chi^{(1)}(-\omega_\sigma; \omega)^* = \chi^{(1)}(\omega_\sigma; -\omega^*)$ and $\chi^{(2)}(-\omega_\sigma; \omega_1, \omega_2)^* = \chi^{(2)}(\omega_\sigma; -\omega_1^*, -\omega_2^*)$. $R^{(2)}$ also has to be symmetric under the exchange of α and β , if the according time coordinates are also exchanged, thus $\chi_{\mu\alpha\beta}^{(2)}(-\omega_\sigma; \omega_1, \omega_2) = \chi_{\mu\beta\alpha}^{(2)}(-\omega_\sigma; \omega_2, \omega_1)$.

In a next step we make the experimentally relevant simplification³⁰ that we are just concerned with monochromatic waves and that we only have to be concerned with the wave amplitudes³¹ P_ω and E_ω of those monochromatic waves, such that

$$E_\alpha(\omega) = \frac{1}{2} \sum_{\omega' \geq 0} E_{\omega'; \alpha} \delta(\omega - \omega') + E_{-\omega'; \alpha} \delta(\omega + \omega') \\ P_\alpha(t) = \frac{1}{2} \sum_{\omega' \geq 0} P_{\omega'; \alpha} \exp(-i\omega t) + P_{-\omega'; \alpha} \exp(i\omega t) \quad (273)$$

We thus obtain:

³⁰ This is particularly relevant for Quantum Communications, as we often operate with cw-systems, here.

³¹ A wave amplitude is related to a discrete monochromatic wave, as opposed to the Fourier coefficient, which is an amplitude density. They carry optical intensity according to $I_\omega = 1/2c\varepsilon_0 n(\omega) |E_\omega|^2$, the unit is W/m^2 .

$$P_{\omega\sigma;\mu} = \varepsilon_0 \sum_{\omega} K(-\omega_{\sigma}; \omega_1) \chi_{\mu\alpha}^{(1)}(-\omega_{\sigma}; \omega_1) E_{\omega;\alpha} + \varepsilon_0 \sum_{\omega} K(-\omega_{\sigma}; \omega_1, \omega_2) \chi_{\mu\alpha\beta}^{(2)}(-\omega_{\sigma}; \omega_1, \omega_2) E_{\omega_1;\alpha} E_{\omega_2;\beta} \quad (274)$$

Where the sum over ω denotes summation over all electric field frequency modes, for which $\omega_{\sigma} = \sum_{l=1}^n \omega_l$, which is, of course the manifestation of photon energy conservation (which is kind of unexpected, as we are purely classical here and have never introduced the photons, as such). We have also introduced the numerical quantity K , which is used to keep track of the permutation density of the underlying process. Depending on the order of the process and the frequencies involved this takes the form:

| Process | Order | Frequencies $-\omega_{\sigma}; \omega_1, \dots, \omega_n$ | K |
|--|-------|---|---------------|
| linear absorption, refractive index | 1 | $-\omega; \omega$ | 1 |
| optical rectification | 2 | $0; \omega, -\omega$ | $\frac{1}{2}$ |
| Pockels effect | 2 | $-\omega; \omega, 0$ | 2 |
| SHG | 2 | $-2\omega; \omega, \omega$ | $\frac{1}{2}$ |
| SFG, DFG, SPDC | 2 | $-\omega_1 \mp \omega_2; \omega_1, \pm\omega_2$ | 1 |

For many processes we can make one further approximation: if all frequencies in question are far away from material resonances and there are no material resonances between the frequencies we can essentially drop out the frequency dependence of the $\chi^{(2)}$ -tensor altogether. This is a special application case of the Kramers-Kronig relations and called Kleinmann-symmetry.

There is a lot more to be said about the response-tensors, than can possibly be done here. Their calculation in and by itself are again problems of Quantum Mechanics and can often be understood in terms of perturbation calculations. However, as we know, the derivation of the electron orbitals itself is a hard problem, which cannot be tackled in an ab-initio manner. Often one has to determine these quantities experimentally.

The entries of the χ -tensors are heavily related to the symmetry of the materials in question. More specifically one can show, that

- inversion symmetric materials (and also amorphous materials) have $\chi^{(2)} = 0$.
- Non-inversion symmetric materials (e.g. crystals) can be grouped into several symmetry classes, which in turn determine the location and relative sign of non-zero entries of the $\chi^{(2)}$ -tensor.

More on this matter can be found in textbooks on nonlinear Optics, e.g. Boyd's "Nonlinear Optics" or Butcher's and Cotter's "The Elements of Nonlinear Optics".

We however include one further step, which is often done in the context of NLO and will allow you to access the literature in a more concise manner. We replace the $\chi^{(2)}$ -tensor with the d_{ijk} -tensor and then the contracted d_{il} -tensor:

$$d_{ijk} = \frac{1}{2} \chi_{ijk}^{(2)} \quad (275)$$

The d_{il} -tensor is constructed from the d_{ijk} -tensor by taking the following replacement rule:

| jk | 11 | 22 | 33 | 23,32 | 31,13 | 12,21 |
|------|----|----|----|-------|-------|-------|
| l | 1 | 2 | 3 | 4 | 5 | 6 |

The d_{il} -tensor is a 3x6-matrix:

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \quad (276)$$

The reduction of the number of elements from $3 \times 3 \times 3 = 27$ to $3 \times 6 = 18$ is related to the intrinsic permutation symmetry properties of the χ -tensor.

A 3.2 Coupled Wave Equations

In the last chapter we have derived the relation between an electric field and the (nonlinear) polarization, which it induces. At this point this is still an entirely local description of the action of the EM-field and does not allow us to describe, what happens if EM-radiation propagates through a medium. To achieve this, we, of course, have to plug in these relations into Maxwell's Equations.

First, however, we rewrite the Polarization equations for three-wave-mixing (SFG) in terms, if the reduced d_{il} -tensor (\mathbf{d} in matrix notation)

$$\begin{bmatrix} P_{\omega_3;x} \\ P_{\omega_3;y} \\ P_{\omega_3;z} \end{bmatrix} = 4\mathbf{d} \cdot \begin{bmatrix} E_{\omega_1;x}E_{\omega_2;x} \\ E_{\omega_1;y}E_{\omega_2;y} \\ E_{\omega_1;z}E_{\omega_2;z} \\ E_{\omega_1;y}E_{\omega_2;z} + E_{\omega_1;z}E_{\omega_2;y} \\ E_{\omega_1;x}E_{\omega_2;z} + E_{\omega_1;z}E_{\omega_2;x} \\ E_{\omega_1;x}E_{\omega_2;y} + E_{\omega_1;y}E_{\omega_2;x} \end{bmatrix} \quad (277)$$

If we choose a fixed set of Polarizations for the waves with the three frequencies we can compress the notation to:

$$P_{\omega_3} = 4d_{\text{eff}}E_{\omega_1}E_{\omega_2} \quad (278)$$

Where the values of d_{eff} is a superposition of the values of the d_{il} -tensor and depend on the selected polarization and symmetry conditions of the crystal. As an example, we take a negative uniaxial crystal of symmetry class $3m$ (such as BBO) we get:

$$\begin{aligned} d_{\text{eff}}^{(\text{I})} &= d_{32} \sin \theta - d_{22} \cos \theta \sin 3\phi \\ d_{\text{eff}}^{(\text{II})} &= d_{22} \cos^2 \theta \cos 3\phi \end{aligned} \quad (279)$$

Where type-I denotes the case where the lower-frequency photons have the same polarization and type-II is where they are in opposite polarization. θ is the angle between the propagation vector and the crystalline z-axis (optical axis). ϕ is the angle between the propagation vector and the xz-crystal plane. We also assume propagation in the z-direction only, to keep this discussion brief and simple (collinear interaction). Thus $\mathbf{k}_i = [0, 0, k_i]$, with $k_i = n_i \omega_i / c$ and $n_i = (\epsilon^{(1)}(\omega_i))^{1/2}$.

We then plug all of these expressions into Maxwell's Equations and end up with the so-called coupled wave equations.

$$\begin{aligned}
 \frac{dE_{\omega_3}}{dz} &= \frac{8\pi i d_{\text{eff}} \omega_3^2}{k_3 c^2} E_{\omega_1} E_{\omega_2} \exp(i\Delta k z) \\
 \frac{dE_{\omega_1}}{dz} &= \frac{8\pi i d_{\text{eff}} \omega_1^2}{k_2 c^2} E_{\omega_2}^* E_{\omega_3} \exp(-i\Delta k z) \\
 \frac{dE_{\omega_2}}{dz} &= \frac{8\pi i d_{\text{eff}} \omega_2^2}{k_2 c^2} E_{\omega_1}^* E_{\omega_3} \exp(-i\Delta k z)
 \end{aligned} \tag{280}$$

Here we have introduced the important quantity $\Delta k = k_1 + k_2 - k_3$, called the phase-mismatch. We will get to that in a second. For SPDC we typically have the case, that E_{ω_3} , e.g. the pump-wave, is very strong and is itself unaffected by the weak feedback from the few photons, which we create in the signal field ω_1 or the idler ω_2 . We can thus assume the so-called undepleted pump approximation, such that this field is constant and denote it as E_0 :

$$\begin{aligned}
 \frac{dE_{\omega_1}}{dz} &= \frac{8\pi i d_{\text{eff}} \omega_1^2}{k_2 c^2} E_{\omega_2}^* E_0 \exp(-i\Delta k z) \\
 \frac{dE_{\omega_2}}{dz} &= \frac{8\pi i d_{\text{eff}} \omega_2^2}{k_2 c^2} E_{\omega_1}^* E_0 \exp(-i\Delta k z)
 \end{aligned} \tag{281}$$

Let's, however, say a few more words about phase-matching:

- The value of Δk can be tuned by modification of the crystal orientation, propagation directions, crystal temperatures and by the introduction of a so-called quasi-phasematching. In the broader scheme of things, it can be tuned by the modification of the geometry and the resulting impact of the dispersive properties of the mode in question (which may not be plane waves but any kind of mode).
- If $\Delta k = 0$, then the intensities in the signal and idler grow exponentially with propagation length/crystal thickness (assume $E_{\omega_2}(z=0)$).

$$E_{\omega_1} = E_{\omega_1}(z=0) \cosh \frac{64\pi^2 \omega_1^2 \omega_2^2 d_{\text{eff}}^2}{k_1 k_2 c^4} |E_0^2| \tag{282}$$

- If $\Delta k \neq 0$, then the intensities in the signal and idler behave in a $\sin^2 L_{\text{PM}} z$ -manner with $L_{\text{PM}} = \Delta k^{-1}$.
- In the classic case this only happens if the process is seeded, i.e. if there is an initial intensity in the signal and/or idler mode. This case is called difference frequency generation and is the basis for OPAs and OPOs.
- In the quantum case the initial state is not zero but $|\text{vac}\rangle$ and thus there is always a seed. This case is then called SPDC.

Analytic solutions can also be found quite simply for $\Delta k \neq 0$ and also for non-collinear interaction. There is no new physics here, but it's quite an index-battle. So, let's skip that. The bottom line message is that:

- the entire SPDC-process can be described locally (coupled mode equations) but also globally (as the action of a device of finite thickness)
- the process is symmetric in signal and idler

- in both cases its action is to mix the modal excitations (i.e. photons) in the three modes, involved, we can thus derive a local and a global interaction Hamiltonian for the process, which describes the tree-mode-mixing process
- the interaction Hamiltonian acting on the quantum-vacuum in the signal- and idler-mode will create symmetric excitations (e.g. photons-pairs or pairs of pairs) in these modes.

A word on caution for those of you familiar with classical nonlinear processes, for which at least one ω_1 is fixed and ω_2 is then determined by this and the pump ω_p . This is not the case for SPDC, where only ω_p is fixed (or in case of a pulsed pump: an a certain range is possible). Here any combination of $\omega_1 + \omega_2 = \omega_p$ must be considered, each of which have individual phase mismatch conditions. In general phase matching will occur for a specific $\tilde{\omega}_1$ and a certain range around that will still be phase matched enough (this range depends on the crystal, phase-matching-type and crystal length), such that a broadband signal will be generated.

A 3.3 Two-photon state produced in SPDC

The previous section discussed how nonlinear polarization response in a second-order nonlinear crystal couples the pump, signal, and idler fields. In principle, we can use the coupled mode relations to calculate the modal distribution and photon number characteristics of PDC states to arbitrary order – from quantum all the way to classical nonlinear optics. In practice, the nonlinear interaction is weak and the PDC emission is dominated by the spontaneous emission of photon pairs. In order to derive an expression for the two-photon SPDC state, let us consider the following experimental situation, which accurately describes the majority of quantum optics experimentation: A strong pump laser which propagates along a principal axis of a nonlinear crystal of length L, which we choose as the z-axis of our coordinate system (Fig. 79).

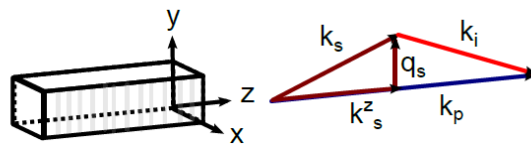


Fig. 79 Coordinate system used for calculation of SPDC state.

Assuming that the wave-vector distribution of the pump, signal, and idler fields are peaked around this preferential direction of propagation, it is prudent to separate the respective wave vectors into a longitudinal component k^z and a transverse component $\mathbf{q} = (k_x, k_y)$

$$\mathbf{k} = k^z(\omega, \mathbf{q})\mathbf{e}_z + \mathbf{q} \quad (283)$$

In this notation, the Hamilton operator for the nonlinear interaction reads:

$$\hat{H}_{PDC} \propto \int d^2\mathbf{r}_\perp dz \chi_{\lambda_p\lambda_s\lambda_i}^{(2)}(\mathbf{r}_\perp, z) \hat{E}_p^+(\mathbf{r}_\perp, z, t) \hat{E}_s^-(\mathbf{r}_\perp, z, t) \hat{E}_i^-(\mathbf{r}_\perp, z, t) + c. c \quad (284)$$

Where $\hat{E}_k^+(\mathbf{r}_\perp, z, t) = \hat{E}_k^-(\mathbf{r}_\perp, z, t)^\dagger$ are the transverse electromagnetic field operators ($k = p, s, i$) with transverse momentum decomposition³²:

³² $f(\omega) = \sqrt{\frac{\hbar\omega}{2\epsilon_0 c n(\omega)}}$ is a normalization constant. Typically, we pull this factor out of the integral and evaluate it at the center frequency of the respective field $f(\omega) = f(\omega_0)$.

$$\hat{E}_{\mathbf{k}}^-(\mathbf{r}_{\perp}, z, t) = \sum_{\lambda} \frac{\epsilon_{\lambda}}{(2\pi)^{\frac{3}{2}}} \int d\omega d\mathbf{q} \hat{a}_{\lambda}^{\dagger}(\omega, \mathbf{q}) f(\omega) \exp(ik^z(\omega, \mathbf{q})z + i\mathbf{q} \cdot \mathbf{r}_{\perp} - i\omega t) \quad (285)$$

In this decomposition, the field creation operators $\hat{a}_{\lambda,s(i)}^{\dagger}(\omega, \mathbf{q})$ create a signal (idler) photon in a plane-wave spatial mode with transverse wave vector \mathbf{q} , frequency ω and polarization vector $\epsilon_{\lambda,s(i)}$, whereas the pump annihilation operator $\hat{a}_{\lambda,p}(\omega, \mathbf{q})$ removes a photon from the pump field. Let's assume that the nonlinear material has only a single relevant non-zero nonlinear tensor coefficient³³, so that the susceptibility reduces to a scalar effective nonlinear coefficient

$$\chi^{(2)} = \chi_{\lambda_p \lambda_s \lambda_i}^{(2)} \cdot \epsilon_{\lambda_p} \cdot \epsilon_{\lambda_s} \cdot \epsilon_{\lambda_i} \quad (286)$$

Now, we only have to consider a single scalar operator for each field, i.e.:

$$\hat{E}_{\mathbf{k}}^-(\mathbf{r}_{\perp}, z, t) = \frac{f(\omega_0)}{(2\pi)^{\frac{3}{2}}} \int d\omega d\mathbf{q} \hat{a}_{\lambda}^{\dagger}(\omega, \mathbf{q}) \exp(ik^z(\omega, \mathbf{q})z + i\mathbf{q} \cdot \mathbf{r}_{\perp} - i\omega t) \quad (287)$$

We leave open the possibility of a spatial modulation of the nonlinear coefficient, such as periodic poling. In the case of an ideal quasi-phase matching, the spatial dependence of the nonlinear coefficient is of the form:

$$\chi^{(2)}(z) = \chi^{(2)} \exp\frac{i2\pi}{\Lambda} z \quad (288)$$

where Λ is the poling period. As discussed in the lecture, the poling period allows us to control the phase-matching condition in the crystal (but more on this later). With these definitions we're ready to calculate the multi-mode SPDC state. Because we're considering the interaction to be weak, we use the time-evolution operator in the interaction picture:

$$|\Psi_{\text{SPDC}}\rangle = \exp\left(-\frac{i}{\hbar} \int_T \hat{H}_{\text{PDC}} dt\right) |\Psi_{\text{initial}}\rangle \quad (289)$$

Some brief notes on the initial state: Since we're considering spontaneous parametric down-conversion (as opposed to stimulated), the signal and idler fields are initially in their vacuum states. The pump field, on the other hand, is a strong laser, which we describe via a multi-mode coherent state with a distribution of transverse momenta $E_p(q)$ and frequency $s(\omega)$. In total we assume to initial state to be of the form:

$$|\Psi_{\text{initial}}\rangle = |E_p(q), s(\omega)\rangle_p |\text{vac}\rangle_{s,i} \quad (290)$$

Now, remembering that coherent states are eigenstates of the photon annihilation operator, we can simply replace the pump mode operator with its classical field amplitudes and factor out the state of the pump field from the final state ($|\Psi_{\text{SPDC}}\rangle = |E_p(q), s(\omega)\rangle_p |\Psi_{\text{SPDC}}\rangle_{s,i}$). With all these restrictions, let's recap what the Hamiltonian we're considering now looks like:

³³ Depending on which nonlinear coefficient mediates the interaction, we can define three types of SPDC: type-0 (pump, signal, and idler co-polarized), type-I (signal and idler co-polarized, but orthogonal to pump), or type-II (signal and idler polarizations orthogonal).

$$\hat{H}_{PDC} \propto \int d^2 \mathbf{r}_\perp dz \chi^{(2)}(z) E_p(\mathbf{r}_\perp, z, t) \hat{E}_s^-(\mathbf{r}_\perp, z, t) \hat{E}_i^-(\mathbf{r}_\perp, z, t) + c. c \quad (291)$$

We now approximate the time evolution to the lowest order giving us a photon pair, that is:

$$|\Psi_{SPDC}\rangle_{s,i} = |vac\rangle_{s,i} - \frac{i}{\hbar} \underbrace{\int_{T,V} dt d\mathbf{r} \chi^{(2)}(\mathbf{r}) E_p(\mathbf{r}_\perp, z, t) \hat{E}_s^-(\mathbf{r}_\perp, z, t) \hat{E}_i^-(\mathbf{r}_\perp, z, t)}_{\text{two-photon contribution}} |vac\rangle_{s,i} + \dots \quad (292)$$

Dropping the vacuum term and normalizing, we get our final result for the two-photon state:

$$|\Psi_{SPDC}^{(2)}\rangle = \int d\omega_s d\mathbf{q}_s d\omega_i d\mathbf{q}_i \underbrace{\Phi(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i)}_{\text{"bi-photon mode function"}} |\omega_s, \mathbf{q}_s\rangle |\omega_i, \mathbf{q}_i\rangle \quad (293)$$

This looks quite neat, but still have a rather complicated expression for the bi-photon wave function in transverse momentum and frequency space:

$$\begin{aligned} & \Phi(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i) \\ &= \int_{T,V} dt d\mathbf{r}_\perp dz \chi^{(2)}(\mathbf{r}_\perp, z) \int d\omega_p d\mathbf{q}_p E_p(\mathbf{q}_p) s(\omega_p) e^{-i(\omega_p - \omega_s - \omega_i)t} e^{i(k_p^z - k_s^z - k_i^z)z} e^{i(\mathbf{q}_p - \mathbf{q}_s - \mathbf{q}_i) \cdot \mathbf{r}_\perp} \end{aligned}$$

So let's try to simplify this using a few reasonable assumptions. First, we assume that the nonlinear crystal is homogeneous in the transverse plane, and that its aperture is much larger than the transverse extension of the fields. With this we can extend the integration limits in x,y to and get:

$$\int d\mathbf{r}_\perp e^{i(\mathbf{q}_p - \mathbf{q}_s - \mathbf{q}_i) \cdot \mathbf{r}_\perp} \rightarrow \delta(\mathbf{q}_p - \mathbf{q}_s - \mathbf{q}_i) \quad (294)$$

Next, we extend the time integration to infinity:

$$\int dt e^{-i(\omega_p - \omega_s - \omega_i)t} \rightarrow \delta(\omega_p - \omega_s - \omega_i) \quad (295)$$

Finally, carrying out the integration over the crystal length L:

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} dz e^{i(k_p^z - k_s^z - k_i^z)z} \rightarrow \text{sinc}\left(\frac{L}{2}\Delta k^z\right) \quad (296)$$

and plugging in all this in above, we get:

$$\Phi(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i) = \sigma L E_p(\mathbf{q}_s + \mathbf{q}_i) s(\omega_s + \omega_i) \text{sinc}\left(\frac{L}{2}\Delta k^z\right) \quad (297)$$

This looks already a lot better. We now see that the modal distribution of the SPDC photons is governed by two main contributions:

1. The pump envelope $E_p(\mathbf{q}_s + \mathbf{q}_i) s(\omega_s + \omega_i)$ that determines the total momentum and the total energy of the photon pairs
2. The phase-matching term: $\text{sinc}\left(\frac{L}{2}\Delta k^z\right)$ that determines how energy and momentum is distributed among different modes of the signal and idler photon.

Playing with these two parameters, we can engineer the modal distribution of the photon pair emission to our liking. Some generic features to note:

- Due to energy and momentum conservation, the signal and idler photons are typically entangled: $\Phi(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i) \neq \phi_s(\omega_s, \mathbf{q}_s)\phi_i(\omega_i, \mathbf{q}_i)$
- Energy and momentum of one photon may also be correlated: $\Phi(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i) \neq \phi_\omega(\omega_s, \omega_i)\phi_q(\mathbf{q}_s, \mathbf{q}_i)$. This means that at different momenta (different emission angles) we may have a different spectrum (think of the famous SPDC emission cone rainbow image³⁴). Some people call this single-photon entanglement³⁵.
- Polarization-entanglement involves two mode-functions, e.g. $\Phi_{H_sV_i}(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i)$ and

$\Phi_{V_sH_i}(\omega_s, \mathbf{q}_s, \omega_i, \mathbf{q}_i)$. In order to get a high degree of polarization-entanglement, the polarization must be de-coupled from the mode function, i.e. the two mode functions must overlap.

Mode functions for typical experimental scenarios are presented in the lecture notes.

A 3.4 A note on the Connection to Joint-Probability-Densities and Correlation Properties of Stochastic Ensembles

Let's leave the realm of Qubits behind for a second and consider a systems of biphotons, which are entangled in a continuous variable. In this case the entries of the vector $\alpha, \beta, \gamma, \delta$, need to be generalized to a complex function, which depends on two parameters, i.e. the Joint Spectral Density $JSD(\omega_1, \omega_2)$.

One prominent example is the frequency for photon-pairs generated in the SPDC-process, for which $\omega_s + \omega_i \approx \omega_p$, where the approximate-sign is related to the bandwidth $\Delta\omega_p$ (i.e. the spectral width) of the pump-pulse. The specifics of the nonlinear environment, in which the conversion takes place will also dictate a certain signal frequency Ω_s , for which the process is phase matched and conversion is particularly efficient. The same parameters will also dictate a bandwidth of frequencies $\Delta\Omega_s$, out of which the conversion efficiency drops to zero. The JSD thus takes roughly the form as outlined in Fig. 80.

³⁴ https://en.wikipedia.org/wiki/Spontaneous_parametric_down-conversion

³⁵In this case, the entanglement is between the different degrees of freedom of one particle. We cannot separate the degrees of freedom of a single photon spatially and distribute one to Alice and one to Bob, so this type of entanglement is of limited use in applications in e.g. Quantum Teleportation. However, there may well be other applications where it can be useful. Needless to say, the notion of single-photon entanglement has been subject of many heated debates.

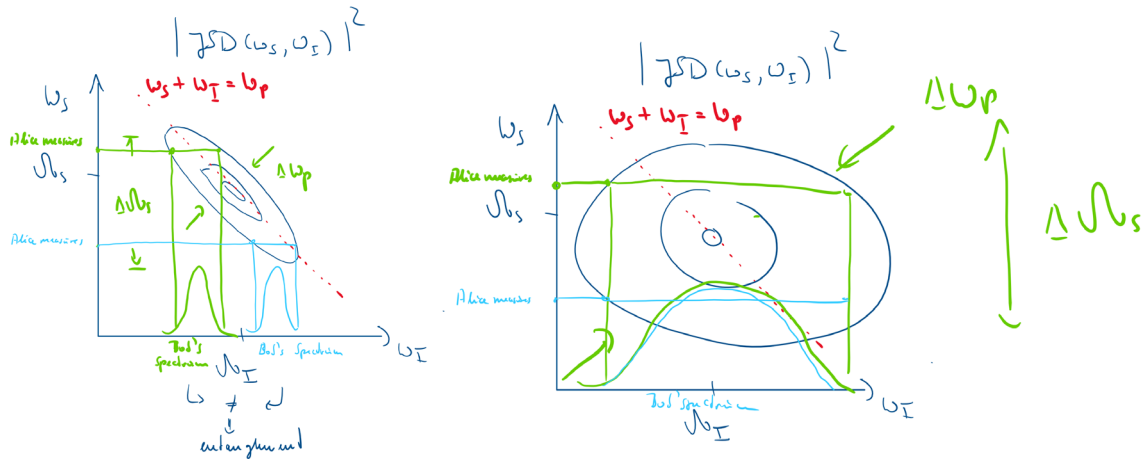


Fig. 80: Schematic JSD for an SPDC process with (left) a narrowband pump and broad phase matching and (right) a broadband pump with less phase matching.

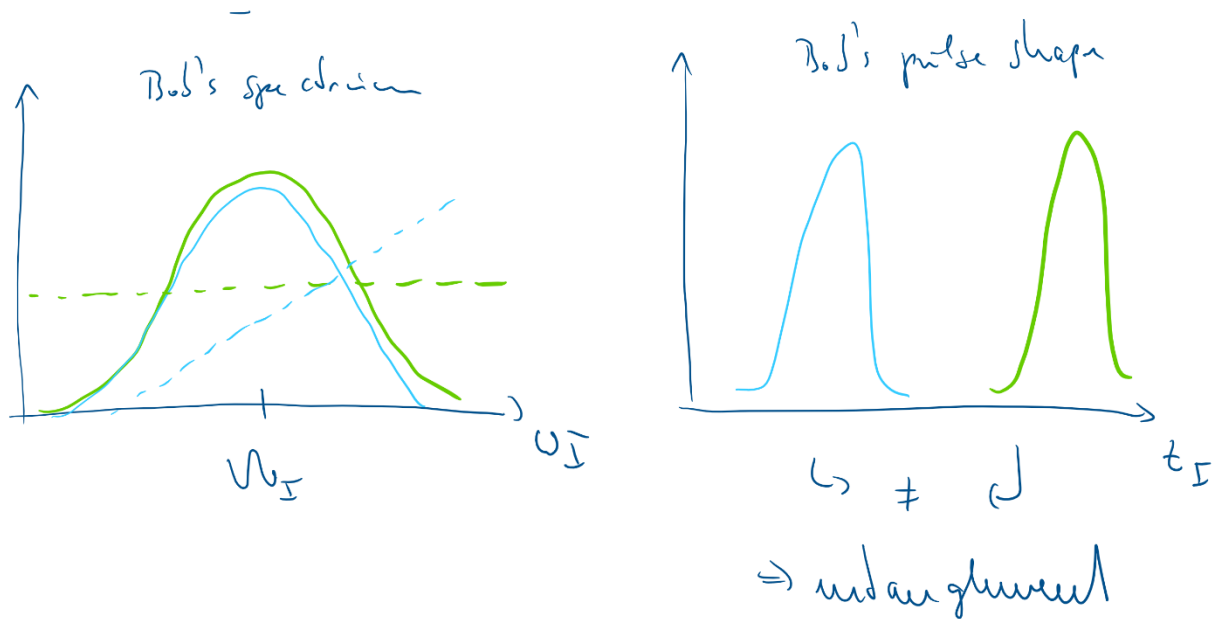


Fig. 81: Possible consequences of the complex-valuedness of the JSD as opposed to a JDF, which leads to subtle difference in (classical) correlation and (quantum) entanglement. (left) Spectral probability densities of Photons received by Bob after Alice has measured the wavelength of her photon in the right case of the above figure. The spectral probability densities are equal for both cases. However, their spectral phases may be different. (right) Temporal probability densities of the spectra to the left. Spectral phase difference manifests in different temporal structures (here, time of arrival), meaning that both photons are in fact distinguishable and thus belong to a mixed state.

Keep in mind, that the JSD describes the probability of observing a photon-pair at frequencies ω_s and ω_i . So, if the signal is sent to Alice and the idler to Bob and Alice measures the frequency of her photon (e.g. by using a spectrometer) this will automatically define the spectrum of the pulse for Bob. While the left case in Fig. 81 does clearly show an entangled situation, there may be non-entanglement for the right case, as the spectral intensity (!) of Bob's photon will not depend on Alice's measurement. There is, however, a catch that we will get into in a second.

This is, in fact, a situation that you may know from stochastics. In stochastics you may have two random variables, that are distributed in a correlated manner; e.g. you may randomly pick points on a map of the earth and record latitude x_1 and temperature x_2 . These two quantities should have a joint distribution function $JDF(x_1, x_2)$ and they should be correlated, as temperature clearly depends on latitude.

In stochastics you can test for correlation by checking, if the JDF can be recomposed from its marginals $M_{x_i} = \int JDF(x, y) dx_i$, such that $JDF(x_1, x_2) = M_{x_1}(x_2)M_{x_2}(x_1)$. In general, this is the case if $JDF(x_1, x_2)$ is of more or less elliptical shape with the main axes parallel to the coordinate axes.

On the ground of the conceptual similarity of correlation and entanglement one may be tempted to use this argument to dismiss the right-hand case of Fig. 80 as non-entangled (the shape is a circle after all). However, one must not forget that the state of a quantum system is not dictated by its probability intensities but by its amplitudes. The JSD is a complex quantity, thus although the measurement of Alice's frequency ω_s may result in the same probability spectrum $|\psi(\omega_i)|^2$ for Bob, it may have a totally different spectral phase $\arg(\psi(\omega_i))$. While this may not show up in a spectral measurement it will have a profound impact on the temporal structure of Bob's photon (e.g. different linear phases would lead to different mean times of arrival) and would then constitute distinguishability and thus break disentanglement.

Tracking back to the previous lectures, this is part of the reasoning behind Bell's inequalities: entanglement is indeed more than correlation; more complex one might be tempted to say. It is therefore also no miracle that entanglement can be used to construct richer and more extreme relations between systems, than correlation alone would be sufficient to explain.