

# Additional lecture notes in FYS4110

Joakim Bergli

October 11, 2023

## 1 Experimental tests of Bell's inequalities

### 1.1 Generation of entangled pairs

How do we generate entangled pairs of photons? There are as far as I know two ways that have been used in practice: Atomic cascades and parametric downconversion in nonlinear crystals. The first was used in the initial experiments, while the latter is used in all modern experiments.

**Atomic cascades** The principle of this is quite simple: An atom, ususally calcium, that has a very short lived excited state is excited to a state with higher energy than this state. The atom will then after some time spontaneously decay to the short lived state, emitting one photon. It will stay in the short lived state only for a short time, and then emit a second photon. Because the angular momentum of the electron changes by  $\pm 1$  in an electric dipole transition, and if one chooses the initial and final states to have the same angular momentum, the polarizations of the two photons must be correlated.

**Parametric downconversion** This is a process which takes place in nonlinear crystals. By this, we mean a material where the polarization is not a linear function of the applied electric field. In principle, this is the case for all crystals, but only some have strong nonlinearity so that the effect is not too small. Examples are beta barium borate (BBO) or potassium titanyl phosphate (KTP). In general, the polarization of a material in the presence of an applied electric field  $E$  is (for simplicity we write the equations for an isotropic material, even if the usual materials used in experiments are birefringent)

$$P = \epsilon_0 \chi^{(1)} E + \epsilon_0 \chi^{(2)} E^2 + \dots$$

where  $\chi^{(1)}$  is the usual linear electric susceptibility, while  $\chi^{(2)}$  describes the lowest order nonlinearity. This gives the displacement vector

$$D = \epsilon_0 E + P = \epsilon E + \epsilon_0 \chi^{(2)} E^2$$

where  $\epsilon = (1 + \chi^{(1)})\epsilon_0$ . The energy (Hamiltonian) is

$$H = \frac{1}{2} \int d^3r D \cdot E = \frac{\epsilon}{2} \int d^3r E^2 + \frac{\epsilon_0}{2} \int d^3r \chi^{(2)} E^3$$

The second term gives the nonlinear interaction in the crystal. We know that the electric field is given by the expression

$$E(r, t) = i \sum_{ka} \sqrt{\frac{\omega}{2\epsilon_0 V}} \epsilon_{ka} \left[ a_{ka} e^{i(kr - \omega t)} - a_{ka}^\dagger e^{-i(kr - \omega t)} \right].$$

This means that  $E^3$  will contain a term proportional to  $\hat{a}_{k_a} \hat{a}_{k'_a}^\dagger \hat{a}_{k''_a}^\dagger$  which means that one photon is annihilated and two are created. That is, one photon is split in two. A more detailed analysis is needed to determine the relations between the energies and polarizations of the two photons, and they will depend on the specific crystal and direction of propagation of the photons. In general, the two photons are entangled, and by choosing the proper direction and polarization of the incoming and outgoing beams one can control the exact state.

## 1.2 CHSH inequality

In the lecture notes, section 2.3.2, there is a discussion of one form of Bell inequality. Here we present a different form, the CHSH inequality.<sup>1</sup>

We consider two spin- $\frac{1}{2}$  particles propagating in opposite directions and reaching detectors A and B. Both detectors can be oriented to measure spin along any axis we choose, but we need only two different for each detector. We call them  $a$  and  $a'$  for A and  $b$  and  $b'$  for B. In a local realistic theory we would have two functions  $A(a, \lambda)$  and  $B(b, \lambda)$  both with value either  $+1$  or  $-1$  depending on whether the spin is parallel or antiparallel to the given axis. Here  $\lambda$  represents any set of hidden variables which is needed to fully specify the state in addition to the usual quantum state. These functions give in a fully deterministic way the outcome of the experiment with a given direction of the measurement axis and given values of the hidden variables. Consider now the quantity

$$S(\lambda) = A(a, \lambda)[B(b, \lambda) + B(b', \lambda)] + A(a', \lambda)[B(b, \lambda) - B(b', \lambda)].$$

For given values of  $b, b'$  and  $\lambda$ , we must have that  $B(b, \lambda)$  and  $B(b', \lambda)$  either have the same or opposite signs. This means that one of the square brackets is 0 and the other  $\pm 1$ .  $S(\lambda)$  can then only take the values  $\pm 2$ . The average

$$\langle S(\lambda) \rangle = \int d\lambda \rho(\lambda) S(\lambda)$$

must then satisfy

$$-2 \leq \langle S(\lambda) \rangle \leq 2.$$

We also need the maximal quantum value for  $S$ . Take the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

and the observables

$$A = \sigma_a = \mathbf{a} \cdot \boldsymbol{\sigma} = \begin{pmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{pmatrix}$$

where  $\mathbf{a} = (a_x, a_y, a_z)$  is the unit vector giving the direction of spin measurement, and we have similar expressions for the other observables. We then have

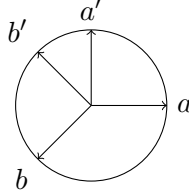
$$\begin{aligned} \langle \psi | \sigma_a \otimes \sigma_b | \psi \rangle &= \frac{1}{2} [\langle \uparrow\downarrow | \sigma_a \otimes \sigma_b | \uparrow\downarrow \rangle - \langle \uparrow\downarrow | \sigma_a \otimes \sigma_b | \downarrow\uparrow \rangle - \langle \downarrow\uparrow | \sigma_a \otimes \sigma_b | \uparrow\downarrow \rangle + \langle \downarrow\uparrow | \sigma_a \otimes \sigma_b | \downarrow\uparrow \rangle] \\ &= \frac{1}{2} [a_z(-b_z) - (a_x - ia_y)(b_x + ib_y) - (a_x + ia_y)(b_x - ib_y) + (-a_z)b_z] = -\mathbf{a} \cdot \mathbf{b}. \end{aligned}$$

<sup>1</sup>After J. F. Clauser, M. A. Horne, A. Shimony, and B. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

Therefore we get

$$\langle S \rangle = \langle \sigma_a \otimes \sigma_b + \sigma_a \otimes \sigma_{b'} + \sigma_{a'} \otimes \sigma_b - \sigma_{a'} \otimes \sigma_{b'} \rangle = -\mathbf{a} \cdot \mathbf{b} - \mathbf{a} \cdot \mathbf{b}' - \mathbf{a}' \cdot \mathbf{b} + \mathbf{a}' \cdot \mathbf{b}'.$$

If we choose the axes all to be in the same plane according to this scheme



we see that

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b}' = \mathbf{a}' \cdot \mathbf{b} = -\frac{1}{\sqrt{2}}, \quad \mathbf{a}' \cdot \mathbf{b}' = \frac{1}{\sqrt{2}}.$$

This gives

$$\langle S \rangle = 2\sqrt{2} > 2$$

One can show that this is the maximal value one can get for any state and observables.

### 1.3 First experiments

The first experiments that conclusively showed a violation of Bell's inequalities were by Aspect in 1981.<sup>2</sup> The source of entangled pairs of photons was an atomic cascade using calcium atoms, and the detectors were placed at a maximal distance of 6 m on opposite sides of the source.

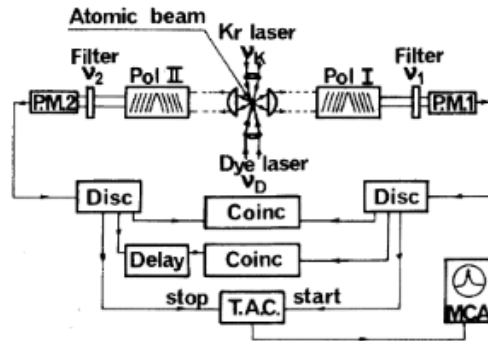


FIG. 2. Schematic diagram of apparatus and electronics. The laser beams are focused onto the atomic beam perpendicular to the figure. Feedback loops from the fluorescence signal control the krypton laser power and the dye-laser wavelength. The output of discriminators feed counters (not shown) and coincidence circuits. The multichannel analyzer (MCA) displays the time-delay spectrum.

<sup>2</sup>A. Aspect *et al.*, Phys. Rev. Lett. **47**, 460 (1981). Similar experiments were performed earlier (see Stuart J. Freedman and John F. Clauser, Phys. Rev. Lett. **28**, 938 (1972)) but with much smaller detection rates and therefore much less statistics.

The detectors are not able to measure the polarization of each photon. Instead, detectors that are insensitive to polarization are used, and polarizers are put in front of the detectors. This means that one can not directly measure correlation functions, since we only get coincidence detection if both photons are polarized along the direction specified by their polarizers. To account for this, one has to rewrite the CHSH inequality in terms of directly observable quantities, to get

$$-1 \leq [R(\mathbf{a}, \mathbf{b}) + R(\mathbf{a}, \mathbf{b}') + R(\mathbf{a}', \mathbf{b}) - R(\mathbf{a}', \mathbf{b}') - R_1(\mathbf{a}') - R_2(\mathbf{b})] / R_0 \leq 0 \quad (1)$$

where  $R(\mathbf{a}, \mathbf{b})$  is the rate of coincidences with polarizer A in orientation  $\mathbf{a}$ , and polarizer B in orientation  $\mathbf{b}$ ,  $R_1(\mathbf{a}')$  is the coincidence rate with polarizer B removed and polarizer A in orientation  $\mathbf{a}'$ , (and similarly for  $R_2(\mathbf{b})$ ), and  $R_0$  is the coincidence rate with the two polarizers removed. The result of the experiment was  $S_{exp} = 0.126 \pm 0.014$  violating inequality (1) by 9 standard deviations.

## 1.4 Loopholes

There were at least two serious concerns regarding the early experiments. These are usually expressed in terms of loopholes: Ways in which the experiments can show violation of Bell's inequalities while they still may not be really violated.

### Locality loophole

The locality requirement in the derivation of Bell's inequality [Lecture notes Section 2.3.2] includes the assumption that the probability distribution  $\rho(\lambda)$  for the hidden variable  $\lambda$  is independent of the settings of the two polarizers. But if somehow the process that creates the entangled pairs is affected by the settings of the polarizers that are going to select which measurement we will perform. Then the process creating the photon pair can make pairs with different polarizations depending on the settings of the polarizers, and one can violate Bell's inequalities even in a classical theory. The way to close this loophole is to have the direction of the polarizers undecided at the time of creation of the photons (or at least the creation event should be outside of the future lightcones of the events where the polarizers are set).

### Detection loophole

The photon detectors are far from perfect, detecting maybe 20% of the photons. If the efficiency of the detectors depends on the hidden variables, one can violate Bell's inequality in local hidden variable models. To close this loophole directly on photons, one would have to have better detectors, so that the probability of misdetection (not detecting a photon present or detecting one that is not present) is sufficiently small. Alternatively, one can create a pair of entangled particles of some other type than photons, where detectors are already good. It is this latter option that is chosen in practice, as we will discuss below.

Both of these may seem paranoid, in the sense that it seems very likely that they are not going to really affect the experiments. In particular, it seems difficult to understand how the settings of the polarizers can affect the process creating the entangled pairs, even if the polarizers are set before the photons are created. I have no idea which physical mechanism should be responsible for this process. But as matters of principle, they still are valid objections, and since one is trying to test the fundamental limitations of any classical theory, it is worth to try to close all loopholes.

## 1.5 Delayed choice experiments

Closing the locality loophole often goes under the name of delayed choice experiments, because one does not decide the settings of the polarizers until after the photons are created. Since the speed of light is large, one has to be very fast unless one can have a large distance between the two detectors.

The first delayed choice experiments were performed in 1982<sup>3</sup>, and were not very much more convincing than the original experiment. We include a discussion here only to show the ingenuity of experimentalists in designing the switching of the polarizers in the very short time available. The distance from the source to the detectors was about 6 m, giving a time of about 20 ns for a light signal to propagate from the polarizer in front of the detector to the source. There is no way to rotate a physical polarizer in this time, so instead they used two detectors with independent polarizers in front. This means that one can change the polarization by deciding which detector the light is sent to. To do this very fast, they had a water bath with transducers (a type of loudspeaker) on both ends. These created a standing ultrasound wave with a frequency of about 25 MHz in the water. This means that at a certain phase of the acoustic wave, the amplitude of density variations in the water is zero. The light is then passing the water unaffected. A quarter of a period later, the density variation is maximal, and this will scatter the light. Since the density variation is periodic in space, it acts as a grating, Bragg reflecting the incoming light if it is incident at the proper angle.

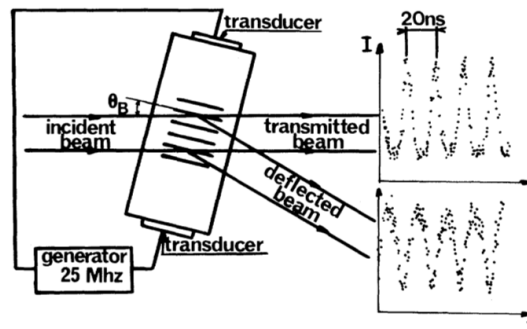


FIG. 3. Optical switch. The incident light is switched at a frequency around 50 MHz by diffraction at the Bragg angle on an ultrasonic standing wave. The intensities of the transmitted and deflected beams as a function of time have been measured with the actual source. The fraction of light directed towards other diffraction orders is negligible.

In this way, the detector that the light hits is switched at twice the acoustic frequency, about every 10 ns. The results were the same as without this elaborate switching device, and in full agreement with quantum mechanics, and violating the Bell inequality. However, in terms of closing the locality loophole, it is not fully convincing. As they write

The new feature of this experiment is that we change the settings of the polarizers, at a rate greater than  $c/L$ . The ideal scheme has not been completed since the change is not truly random, but rather quasiperiodic. Nevertheless, the two switches on the two sides are driven by different generators at different frequencies. It is then very natural to assume that they function in an uncorrelated way.

<sup>3</sup>A. Aspect, J. Dalibard and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

So in principle, both the process creating the entangled pairs and the detector on the other side of the setup could “know” the settings of both the polarizers on one side, as well as the state of the switch, since this can be predicted from the fact that it is periodic. It would be really strange if information, encoded in some hard to imagine physical signal, could be influencing the detectors and the photon source in just the right way.

Still, to be conclusive, the settings of the polarizers should be really random. How can this be achieved? Using a quantum random generator, of course! <sup>4</sup> This is implemented by a weak light source, emitting a stream of well-spaced photons. Each photon is directed to a 50/50 beamsplitter, and is either reflected or transmitted with equal probabilities. A detector is placed in each output of the beamsplitter, and a detection in one gives the random bit “0”, while a detection in the other gives “1”. Thus, the outcomes are random to the degree that a quantum system collapses randomly to one of the eigenstates in a measurement.

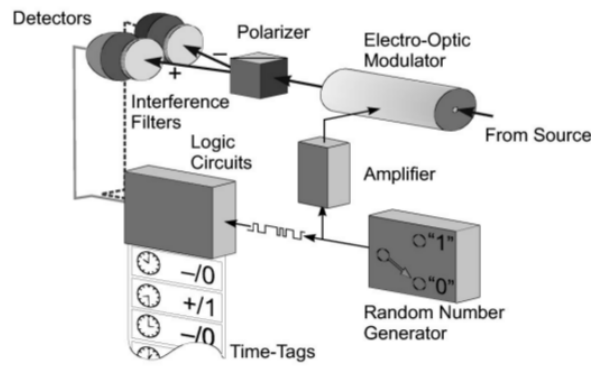


FIG. 2. One of the two observer stations. A random number generator is driving the electro-optic modulator. Silicon avalanche photodiodes are used as detectors. A “time tag” is stored for each detected photon together with the corresponding random number “0” or “1” and the code for the detector “+” or “-” corresponding to the two outputs of the polarizer.

The switching of the beam into one or the other detector is now achieved by an electro-optic modulator, which is a device that rotates the polarization of the photon if a voltage is applied to it. A polarizing beamsplitter then sends the photon to one or the other detector depending on the polarization of the photon. It can change its rate in a time of about 30 ns. To get sufficient time for this and the detectors to fire, the distance from the source to either detector was about 200 m, giving a time of about 1.3  $\mu$ s for light to cross from one detector to the other. The photons were sent through optical fibres, to minimize loss of photons (by absorption or scattering) on the way. As is seen in the following figure, they were well within the constraints of locality, thus conclusively closing the locality loophole.

<sup>4</sup>Gregor Weihs, Thomas Jennewein, Christoph Simon, Harald Weinfurter, and Anton Zeilinger, Phys. Rev. Lett. **81**, 5039 (1998).

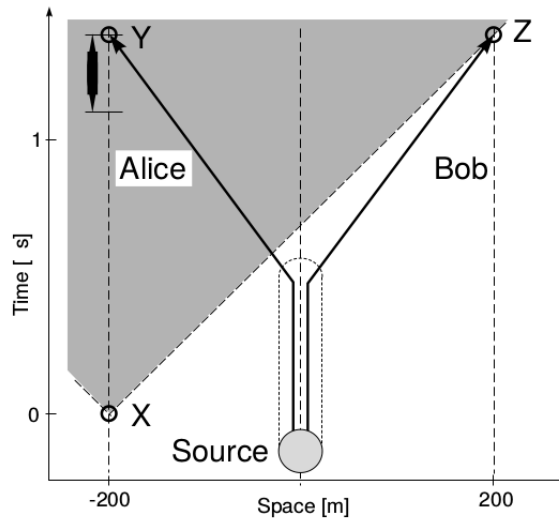


FIG. 1. Spacetime diagram of our Bell experiment. Selecting a random analyzer direction, setting the analyzer, and finally detecting a photon constitute the measurement process. This process on Alice's side must fully lie inside the shaded region which is invisible to Bob's during his own measurement. For our setup this means that the decision about the setting has to be made after point "X" if the corresponding photons are detected at spacetime points "Y" and "Z", respectively. In our experiment the measurement process (indicated by a short black bar) including the choice of a random number took less than only one-tenth of the maximum allowed time. The vertical parts of the kinked photon world lines emerging from the source represent the fiber coils at the source location, which are obviously irrelevant to the locality argument.

To further ensure that no unwanted communication was possible, they independently registered all events with timestamps provided by local clocks and only looked for coincidences (one photon detected at each end at the same time, thus being one entangled pair) afterwards. Previous experiments took the signals from each detector to a common coincidence detector, and one could worry that this allowed for some type of influence of one detector in the measurements of the other. The results were still perfectly in agreement with quantum mechanics, and violating Bell's inequality.

## 1.6 Entangled ions

To improve the efficiency of photon detectors sufficiently is not realistic at present. Therefore one must work with some other system than photons, where detection of the state is more certain. However, creating an entangled pair of anything else, and transporting the away from each other to sufficient distance so that one can measure their state faster than any signal can go from one to the other is equally difficult (or more). Decoherence from interactions with the environment will eat away all the entanglement logn before they are far enough apart. The trick is to use photons to transport the entanglement to two sufficiently separated places, transfer the entanglement to trapped ions (this process is called entanglement swapping, and we will discuss the principle of this in section 4, and then measure the ion state. The efficiency in the detection of the ionic state is quite good, and sufficient to close the detection loophole.

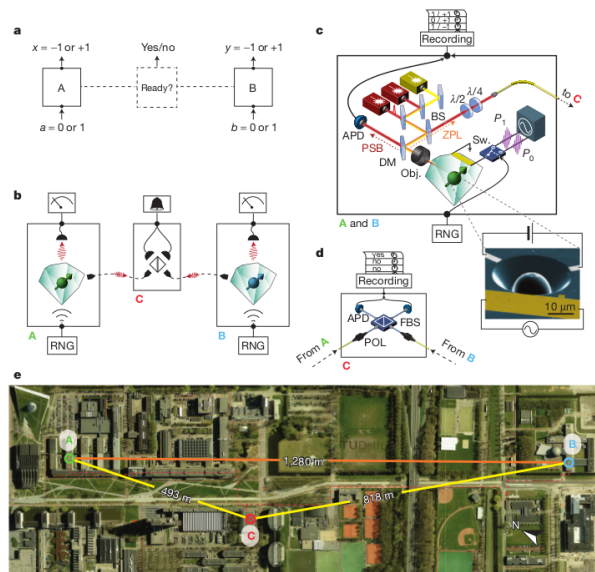
As far as I know, the first such experiments were by Rowe <sup>5</sup> using Beryllium ions. However, the

<sup>5</sup>M. A. Rowe et al., Nature **409**, 791 (2001)

two ions were still in the same trap, and not very far from each other. Better were the experiments<sup>6</sup> where two separate traps were used, but the distance was still only about 1 m, and while the detection loophole was closed, the locality loophole was definitely not.

## 1.7 Loophole free Bell test

The first experiments to close all loopholes were made in 2015<sup>7</sup>. This experiment is using a special defect, a Nitrogen-Vacancy (NV) centre in diamond instead of ions to be prepared in a Bell state. Photons are exchanged and entanglement is swapped to the NV centres. Then the state of each NV centre is measured, and the efficiency of the detectors is about 92%, sufficient to close the detection loophole. The distance between the two detectors was about 1.3 km, giving sufficient time to perform the basis selection and measurement in time to also close the locality loophole.



The result was that  $S = 2.42$ , for the first time confirming violation of Bell's inequality with all Loopholes closed.

## 1.8 Can vi use entanglement to send information?

The simplest quantum interpretation of the EPR experiment is that the measurement of one of the particles collapses the wavefunction for the pair, so that the other particle instantly changes its state. The appearance of the correlations violating Bell's inequality is then not surprising, as the outcome of the measurement on the first particle directly influences the state of the other particle before it is measured. This is often described as “quantum nonlocality”, but this notion should be used with caution because it is quite different from the classical nonlocality that would violate the assumptions behind Bell's inequality. If we believe that there is a real, physical change in the wavefunction of a distant particle when we make a measurement on the entangled partner, we could worry that there is a transfer of information faster than the speed of light, which would be backward in time in certain

<sup>6</sup>

<sup>7</sup>Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres B. Hensen *et al.*, Nature **526**, 682 (2015).



reference frames. However, we can show that the change in the wavefunction is of a special form that does not allow any information to be transmitted.

Consider the quantum state of two spatially separated particles, A and B. We know that the state can be written in the Schmidt-decomposed form

$$|\psi\rangle = \sum_n d_n |n_A\rangle \otimes |n_B\rangle$$

where  $|n_A\rangle$  and  $|n_B\rangle$  are orthonormal states in their respective Hilbert spaces. The density matrix of both particles is then

$$\rho = |\psi\rangle\langle\psi| = \sum_{n,m} d_m^* d_n |n_A n_B\rangle\langle n_A n_B|$$

and the reduced density matrix of particle A is

$$\rho_A = \text{Tr}_B \rho = \sum_n |d_n|^2 |n_A\rangle\langle n_A|$$

We could send information from B to A if we make an operation on B that changes the reduced density matrix of A. We will prove that this is impossible.

What can we do to particle B? Either, we apply a unitary transformation of the state of particle B, or we measure some observable on B.

## 1.9 Non-locality or non-realism?

What are we concluding by the experimental violation of Bell's inequality? All agree that it means that it is not possible to construct any local realistic theory in the form assumed by Bell that will explain the experiments. This in itself is rather remarkable: We can rule out a whole class of possible theories by this one experiment. However, Bell's argument is not concerned with quantum mechanics, and therefore we can logically not draw any conclusions about quantum mechanics from the experiments. This is probably why the opinions are divided about what these experiments imply for our understanding of quantum mechanics.

At present, I do not have any clear opinion, and we limit ourselves to a few quotes, illustrating different positions. As it is clear that one can not have a local realistic model, one can wonder if quantum mechanics is non-local or non-realistic, or both. Many find that the simplest is to accept that nature is non-local, as long as this non-locality does not imply the possibility of transmitting information

Let me finish on a personal note. For me, quantum nonlocality is an established fact. Yet, it remains mysterious: how does Nature organize its book-keeping to know which measurements should produce nonlocal correlations? Is Nature using an enormously-monstrously-vast Hilbert space to keep track of which physical systems (particles or modes) are entangled? For me this question, like the quantum measurement question, is a real physics question: some day Science will have something meaningful to say about both of these questions.<sup>8</sup>

---

<sup>8</sup>N. Gisin, *Foundations of Physics*, **42**, 80 (2012).

We are led, then, to distinguish two types of influence: the “causal” variety, which produce actual changes in some physical property of the receiver, detectable by measurements on that subsystem alone, and an “ethereal” kind, which do not transmit energy or information, and for which the only evidence is a correlation in the data taken on the two separate subsystems – a correlation which by its nature cannot be detected by examining either list alone. Causal influences *cannot* propagate faster than light, but there is no compelling reason why ethereal ones should not. The influences associated with the collapse of the wave function are of the latter type, and the fact that they “travel” faster than light may be surprising, but it is not, after all, catastrophic.<sup>9</sup>

The issue of realism is more involved than locality, and we quote this warning at the beginning of a review article on scientific realism <sup>10</sup>

It is perhaps only a slight exaggeration to say that scientific realism is characterized differently by every author who discusses it, and this presents a challenge to anyone hoping to learn what it is.

## 2 Interaction free measurements

In the lecture notes, section 3.1, there is a discussion of the interaction free measurements of Elitzur and Vaidman. The conclusion was that it is possible to detect the presence of an object without any photon ever interacting with it. Just the possibility of such interaction was sufficient. But it was not always successful. Here we find the success rate and see how to improve it in several ways and determine the optimal success rate.

### 2.1 The success rate in the original version

In the original version as discussed in the lecture notes, the photon is always detected in detector A if there is no object in the lower arm of the interferometer. If there is an object, we have that the state evolves according to

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}|0\rangle + i\frac{1}{\sqrt{2}}|1\rangle \rightarrow i\frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|lost\rangle \rightarrow -\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|lost\rangle \rightarrow -\frac{1}{2}|0\rangle - i\frac{1}{2}|1\rangle + \frac{1}{\sqrt{2}}|lost\rangle$$

where  $|lost\rangle$  indicates that the photon is absorbed by the object, and no detection occurs. This means that the probabilities of the different outcomes are

$$A: \frac{1}{4} \quad B: \frac{1}{4} \quad \text{No detection: } \frac{1}{2}$$

Only if we get the photon in detector B can we tell that the object is there and that the photon did not interact with it. So the probability that we succeed with detecting the object without interaction is only  $\frac{1}{4}$ .

<sup>9</sup>David J. Griffiths, Introduction to quantum mechanics

<sup>10</sup>Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/entries/scientific-realism/>

## 2.2 First improvement: Repeated trials

We can immediately improve on the probability of success by noting that if we get outcome A, the measurement is inconclusive, but no interaction took place. We can then repeat the experiment with a new photon and have another chance. If this still gives A, we repeat again until B or no detection. The total probability of success is

$$p = \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \dots = \frac{1}{3}$$

## 2.3 Second improvement: Optimized mirrors

Above, we have used a symmetric beam splitter that has the same probability for transmission and reflection. We can also consider splitters that are asymmetric, and optimize the reflection and transmission probabilities. Instead of Eq (3.1) from the lecture notes, we now use for the first beamsplitter

$$|0\rangle \rightarrow a|0\rangle + ib|1\rangle \quad |1\rangle \rightarrow a|1\rangle + ib|0\rangle \quad (2)$$

with  $a^2 + b^2 = 1$ . For the second beamsplitter we interchange the coefficients  $a$  and  $b$

$$|0\rangle \rightarrow b|0\rangle + ia|1\rangle \quad |1\rangle \rightarrow b|1\rangle + ia|0\rangle \quad (3)$$

We will see that the best result is obtained if  $b > a$ , so that the first beamsplitter is mostly reflective while the second is mostly transmissive. The fact that the beamsplitters are complementary to each other in the sense that  $a$  and  $b$  are interchanged ensures that we still have all photons detected by detector A if the box is empty:

$$|0\rangle \rightarrow a|0\rangle + ib|1\rangle \rightarrow ia|1\rangle - b|0\rangle \rightarrow iab|1\rangle - a^2|0\rangle - b^2|0\rangle - iab|1\rangle = -|0\rangle$$

If the box is full we have the evolution of the state

$$|0\rangle \rightarrow a|0\rangle + ib|1\rangle \rightarrow ib|1\rangle + a|lost\rangle \rightarrow -b|0\rangle + a|lost\rangle \rightarrow -b^2|0\rangle - iab|1\rangle + a|lost\rangle$$

This gives the probabilities of the different outcomes

$$\text{A: } b^4 \quad \text{B: } a^2b^2 \quad \text{No detection: } a^2$$

The probability of success in repeated trials is then

$$p = a^2b^2 + (b^4)a^2b^2 + (b^4)^2a^2b^2 + \dots = \frac{a^2b^2}{1 - b^4} = \frac{b^2}{1 + b^2}$$

In the limit  $b \rightarrow 1$  we have  $p \rightarrow \frac{1}{2}$ . Large  $b$  means that the first beamsplitter is mostly reflective, so that the photon most likely will not hit the object. It also means that the probability of getting the result A is close to one, which means that we have to repeat the experiment before eventually getting B and concluding that the object is there without interaction, or having no detection, which means that the photon reached the object.

## 2.4 Third improvement: Quantum repeated trials

One can do even better by exploiting quantum interference<sup>11</sup>. Instead of really detecting the photon and collapsing the wavefunction after each trial, one can send the outcome into a second identical interferometer, and then to another and so on. This is illustrated in the figure below, where the darkness of the lines indicates the probability amplitude at that point.



The number of beamsplitters is  $N$ , and they are identical with the action

$$|0\rangle \rightarrow a|0\rangle + ib|1\rangle \quad |1\rangle \rightarrow a|1\rangle + ib|0\rangle \quad (4)$$

where now  $|0\rangle$  represents a photon moving up and  $|1\rangle$  one moving down. We choose

$$a = \sin \frac{\pi}{2N} \quad b = \cos \frac{\pi}{2N}$$

and analyze first the case where there is no object. If we let the state after beamsplitter  $n$  be  $c_n|0\rangle + d_n|1\rangle$  we have that the action of one step in the interferometer is

$$c_n|0\rangle + d_n|1\rangle \xrightarrow{\text{mirrors}} id_n|0\rangle + ic_n|1\rangle \xrightarrow{\text{beamsplitter}} (-bc_n + iad_n)|0\rangle + (iac_n - bd_n)|1\rangle$$

which gives the recurrence relations

$$c_{n+1} = -bc_n + iad_n \quad d_{n+1} = iac_n - bd_n$$

With  $c_1 = a = \sin \frac{\pi}{2N}$  and  $d_n = ib = i \cos \frac{\pi}{2N}$  it is not difficult to see that the solution is

$$c_n = (-1)^{n+1} \sin \frac{\pi n}{2N} \quad d_n = (-1)^{n+1} i \cos \frac{\pi n}{2N}$$

After  $N$  beamsplitters we then have  $c_N = (-1)^{N+1}$  and  $d_N = 0$  so a photon injected into the interferometer will always exit up. To repeat the option of interacting with the object for each beamsplitter, we need to have not a single object, but one for each step, as illustrated in the following figure (this seems a bit unrealistic, but it is soon to be fixed).

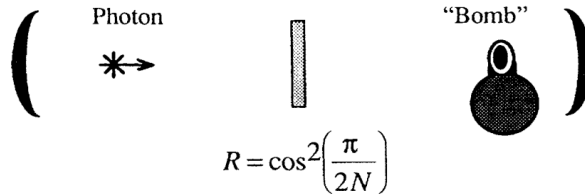


To get the outcome  $|1\rangle$  (photon exiting down) at the end of the chain, we need to have reflection on all the beamsplitters, so that we never get absorbed by any object. This would indicate the presence of the object since it is impossible if the object is not there. The probability for this is  $p = \cos^{2N} \frac{\pi}{2N}$ . For large  $N$  we can expand this to get

$$p = \cos^{2N} \frac{\pi}{2N} = 1 - \frac{\pi^2}{4N} + \dots$$

<sup>11</sup>This idea and the real experiment are discussed in Kwiat et al., Phys. Rev. Lett. **74**, 4763 (1995).

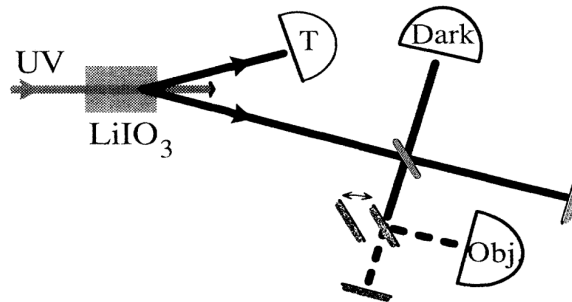
which approaches 1 in the limit  $N \rightarrow \infty$ . This means that we succeed in close to all cases if  $N$  is large. It is not easy to imagine successfully building this device with large  $N$  since it requires a long chain of optical elements which would be difficult to align, and also it is not so elegant to have  $N$  objects to be detected and not just one. Fortunately, the whole chain can be collapsed, and one beamsplitter used repeatedly.



We use two mirrors and one beamsplitter in the middle, creating two optical cavities coupled by the beamsplitter. A photon is inserted into the left cavity at time  $T = 0$ . For a beam splitter reflectivity of  $\cos^2 \frac{\pi}{2N}$ , and in the absence of any absorber, the photon will with certainty be located in the right cavity at time  $T_N = N \times$  (round-trip time), due to interference effects. Therefore, a detector inserted into the left cavity at time  $T_N$  would not fire. However, in the presence of an absorber or scatterer in the right cavity, the photon wave function is continually projected back onto the left cavity. Making the coupling weaker (i.e., increasing the reflectivity) and the number  $N$  greater, one can reduce the probability that the photon ever leaves the first cavity when the object is in the second. A detector inserted into the left cavity will then nearly always fire at time  $T_N$ . Again, the probability of an interaction-free measurement can be made arbitrarily close to 1.

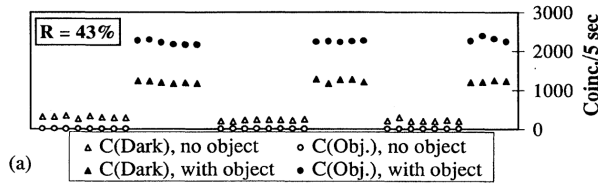
## 2.5 The real experiment

The real experiment was less ambitious, and used only a single pass through an interferometer. The setup is shown below



A laser beam is passing through a crystal of  $\text{LiIO}_3$ , which creates entangled pairs. One photon of the pair is sent to detector T to confirm that another photon enters the interferometer. The second photon is split by a beamsplitter and reflected back using two mirrors. The object whose presence is to be detected is implemented as a mirror that can be outside of the beam, corresponding to no object, or in the beam, deflecting the photon. To make an additional check, it is not absorbed but sent to a detector (marked “Obj”). In the absence of an object, the interferometer is such that all photons are sent back to the source, and no photons exit in the direction of the detector marked “Dark”, so this detector should never detect any photons in the absence of an object. In the presence of the object, there is no amplitude to return from the lower branch of the interferometer, the destructive interference

is removed, and photons can reach the “Dark” detector. This shows up as coincidences between the detectors “T” and “Dark” as was indeed observed.



### 3 Quantum cryptography

One of the ways in which one can exploit quantum mechanics in communication is in cryptography, in particular in distributing cryptographic keys in a secure way. Here we describe the principles of quantum key distribution and some examples of experiments.

#### 3.1 Secure classical communication: one time pad

One of the simplest ways to construct a classical code is to replace each letter in the alphabet with another, which is called a substitution cipher. This is easy to code and decode: To code you need a table of the type  $a \rightarrow r, b \rightarrow m, \dots$ , and to decode you need the reverse table  $r \rightarrow a, m \rightarrow b, \dots$ . These tables are referred to as the encoding and decoding keys, respectively. Unfortunately, it is also easy to break this code, at least if you have a reasonably long coded message to study and know the language it is written in. The point is that some letters are more common than others, so we can guess that a letter occurring frequently in the coded message is corresponding to one of the frequently occurring letters. To do even better, one can look at groups of letters or short words, and one will be able to reconstruct the decoding key, and break the code. It becomes much more difficult to decode if the same key is not used all the time. This was the principle for example of the famous Enigma system used during WW2. There, the key was implemented as a series of wheels with contacts for each letter and wires connecting them. After each letter was encrypted, the wheels were rotated, so that effectively each letter was encoded using a new key, although according to some pattern that provides some help in an attempt to break the code. However, we get the idea that if we could really use a new key for each letter, the code could never be broken.

In our digital world, we are happy only to send 0 and 1, so we replace the alphabet by just those two symbols. Then there are only two possible keys:

$$\begin{array}{ccc} 0 \rightarrow 0 & \text{or} & 0 \rightarrow 1 \\ 1 \rightarrow 1 & & 1 \rightarrow 0 \end{array}$$

These are conveniently represented by addition mod 2. Adding 0 gives the first key, adding 1 (mod 2) gives the second key. This means that if both the encoder and decoder have access to the same random sequence of bits (with equal probabilities for 0 and 1 and no correlations between any bits), they can do the following.

Original message	0	1	1	0	1	1	0	...
Random sequence	1	1	0	1	0	1	0	...
Coded message	1	0	1	1	1	0	0	...
Same random sequence	1	1	0	1	0	1	0	...
Decoded message	0	1	1	0	1	1	0	...

To encode the message, one adds the random sequence (mod 2). To decode, one adds the same random sequence again (which is the same as subtracting (mod 2)). As long as the random sequence is truly random, the encoded message will be equally random, and it is absolutely impossible to break the code. It is important that there are no patterns in the random sequence. For example is it tempting to have a sequence of finite length and then start from the beginning again once you reach the end. This would introduce a pattern that in principle could be discovered and used to break the code (although in practice it could be extremely difficult if the message is not many times the length of the random sequence). To make it principally unbreakable, each bit in the random sequence can be used only once, and it is therefore referred to as a one time pad. The disadvantage of this type of code is that it requires the sender and receiver to share the random sequence at some time before they can communicate, either by meeting or sending the information in some way they trust nobody can intercept.

What if I am in some remote place and run out of random sequence and have some really secret message for a friend back home? Or I find that I want to send a secret to someone I never met? If we can somehow generate the same random sequence in two places while at the same time being certain that nobody, even if they listen to all communication between us, will get any information about the sequence, we can use this as a one time pad to encode the message. This trick can be performed if we are able to send and measure particles in well defined quantum states (like electrons with spin or photons with polarization) as well as classical signals.

### 3.2 Quantum key distribution

We have two parties A and B that want to establish a common random sequence, which we will call the key, to be used as a one time pad. A first generates one random sequence of bits to be used as the key (or at least parts of it will be the key). A also generates an additional random sequence specifying the basis which she will use to encode the key in a two-level quantum system. If the bit specifying the basis is 0, she will use the basis of eigenstates of  $\sigma_z$  which we denote  $\{|\uparrow\rangle, |\downarrow\rangle\}$ , while if the basis bit is 1 she will use the basis of eigenstates of  $\sigma_x$  which we denote  $\{|\rightarrow\rangle, |\leftarrow\rangle\}$  and which are defined by

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle).$$

The encoding of the key bit is then

$$\begin{array}{ll} \text{If basis bit is 0:} & \begin{array}{l} 0 \rightarrow |\uparrow\rangle \\ 1 \rightarrow |\downarrow\rangle \end{array} & \text{If basis bit is 1:} & \begin{array}{l} 0 \rightarrow |\rightarrow\rangle \\ 1 \rightarrow |\leftarrow\rangle \end{array} \end{array}$$

For each bit A now prepares a two-level system in the corresponding state and sends it to B. B has no idea which basis was used to prepare the states, so he can do no better than randomly measuring according to one of the bases for each particle that arrives. When B uses the same basis for measurement as A used for preparation, the result of the measurement is perfectly determined, and agrees with the state that A prepared. In the cases where different bases are used, the result is random and there is no correlation between that prepared state and the result. Both A and B now send their random bases to each other using any type of classical communication. This can be done without any encryption, it does not matter if anyone gets to know this information. The cases where the bases are different are useless and discarded, but for those with a common basis we are guaranteed that B knows the

corresponding bit in A's random key. The process is exemplified in the following table:

A random key	1	0	1	1	0	1	0	...
A random basis	0	0	1	0	1	0	1	...
A state	$ \downarrow\rangle$	$ \uparrow\rangle$	$ \leftarrow\rangle$	$ \downarrow\rangle$	$ \rightarrow\rangle$	$ \downarrow\rangle$	$ \rightarrow\rangle$	...
B random basis	0	1	1	0	0	0	1	...
B measures	$\downarrow$	$\rightarrow$	$\leftarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\rightarrow$	...
Bits that we use	1	–	1	1	–	1	0	...

What happens if somebody tries to eavesdrop and listen to the communication, quantum or classical? It is clear that the information on the basis for A and B, communicated on the open classical channel does not provide any information on the key, and is therefore useless to intercept. What about intercepting the quantum channel? We will not provide a full proof of the security of the protocol<sup>12</sup>, but illustrate what happens in an example. One possible way to gain information from the quantum channel is if the eavesdropper (E) measures the particles. It is clear that the state is affected by the measurement, and we assume in this example that E will pass on the state after the measurement (which is now the eigenstate corresponding to the measurement outcome). Since experiments are (almost?) exclusively using photons, these are absorbed during detection, and E will instead resend the state corresponding to the measured result. This type of attempt to steal the information is therefore known as measure and resend attack. At the time of the exchange of quantum information, E has no idea about the basis that is used by A or B for that particle, this is only given by classical communication later. So E can do no better than randomly choosing a basis for the measurement. We only have to consider those cases where the basis of A and B are the same, since these are the only that are going to be used. Sometimes E will measure in same basis as the one used by A and B, in which case the state will not be modified, and E gets the information about the key at that bit. E can then decode that bit if she intercepts the coded message. If E uses a basis different from A, the state will be modified, and the result of the measurement will not be correlated with the actual value A has for that bit. The fact that the state is changed also means that B may not get the same value as A for the key at that point. In the example below, we have not specified what happens when A and B do not use the same basis, as these bits are discarded anyway. We have marked with circles those places where the keys for E or B are different from the one of A.

Bit number	1	2	3	4	5	6	7	...
A random key	1	0	1	1	0	1	0	...
A random basis	0	0	1	0	1	0	1	...
A state	$ \downarrow\rangle$	$ \uparrow\rangle$	$ \leftarrow\rangle$	$ \downarrow\rangle$	$ \rightarrow\rangle$	$ \downarrow\rangle$	$ \rightarrow\rangle$	...
E random basis	0	0	0	1	1	1	0	...
E measures	$\downarrow$	–	$\uparrow$	$\leftarrow$	–	$\leftarrow$	$\downarrow$	...
Key that E gets	1	–	0	1	–	1	1	...
E sends to B	$ \downarrow\rangle$	–	$ \uparrow\rangle$	$ \leftarrow\rangle$	–	$ \leftarrow\rangle$	$ \downarrow\rangle$	...
B random basis	0	1	1	0	0	0	1	...
B measures	$\downarrow$	–	$\rightarrow$	$\downarrow$	–	$\uparrow$	$\rightarrow$	...
Key that B gets	1	–	0	1	–	1	0	...

In this example we see several possible outcomes:

<sup>12</sup>See Nielsen and Chuang, Quantum Computation and Quantum Information



1. All can use the same basis, and then also get the same result (Bit 1).
2. E uses the wrong basis, and E and B get the wrong key (Bit 3).
3. E uses the wrong basis, but E and B still get the right key (Bit 4).
4. E uses the wrong basis, and E gets the right key while B gets the wrong key (Bit 6).
5. E uses the wrong basis, and E gets the wrong key while B gets the right key (Bit 7).

The point is that the interference by E results in a certain chance that B will not get the right key, and this can be used to detect the presence of E. Before using this key, A and B will validate by comparing a certain number of bits randomly chosen. The probability that a given bit will give a wrong key to B is  $\frac{1}{4}$  since there is a chance  $\frac{1}{2}$  that E uses the wrong basis and then a chance  $\frac{1}{2}$  that B will get the wrong result. The probability that comparing one specific bit at random will reveal E is then  $\frac{3}{4}$ . If they compare  $n$  bits, the probability that they detect E is  $p_n = 1 - \left(\frac{3}{4}\right)^n$  which becomes very small for not so large  $n$ . For example  $p_{100} = 3,2 \cdot 10^{-13}$ . As we will see below, one can experimentally generate 100 bits in less than a second even over long distances, so only a small fraction of the bits need to be tested in order to ensure virtual certainty about the non-presence of any eavesdropper. This picture is complicated by the fact that there will be naturally occurring noise or detector errors, which gives a certain background rate of erroneous key transfers.

## 4 Entanglement swapping

### 4.1 Principle

Entanglement is normally a consequence of interaction. If two particles, which initially are in a product state, interact, the final state will generally be entangled. But can we create entanglement between two particles which never have interacted with each other?

Imagine A and B are far from each other and has two TLS each, call them  $A_1, A_2, B_1$  and  $B_2$ . We want  $A_1$  and  $B_1$  to end in an entangled state. We let both parties entangle their pair of particles so that they are in the states

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{A_1}\downarrow_{A_2}\rangle - |\downarrow_{A_1}\uparrow_{A_2}\rangle)$$

$$|\psi_B\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{B_1}\downarrow_{B_2}\rangle - |\downarrow_{B_1}\uparrow_{B_2}\rangle)$$

The total state is

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle = \frac{1}{2} \left[ |\uparrow_{A_1}\uparrow_{B_1}\rangle \otimes |\downarrow_{A_2}\downarrow_{B_2}\rangle - |\uparrow_{A_1}\downarrow_{B_1}\rangle \otimes |\downarrow_{A_2}\uparrow_{B_2}\rangle - |\downarrow_{A_1}\uparrow_{B_1}\rangle \otimes |\uparrow_{A_2}\downarrow_{B_2}\rangle + |\downarrow_{A_1}\downarrow_{B_1}\rangle \otimes |\uparrow_{A_2}\uparrow_{B_2}\rangle \right]$$

Were we have just rearranged the grouping of the particles, but not changed the state in any way. We now rewrite this state in the Bell basis as defined in Eq. (3.12) in the lecture notes to get

$$|\psi\rangle = \frac{1}{2} \left[ |\phi^+\rangle_{A_1B_1} \otimes |\phi^+\rangle_{A_2B_2} - |\phi^-\rangle_{A_1B_1} \otimes |\phi^-\rangle_{A_2B_2} - |\psi^+\rangle_{A_1B_1} \otimes |\psi^+\rangle_{A_2B_2} + |\psi^-\rangle_{A_1B_1} \otimes |\psi^-\rangle_{A_2B_2} \right]$$

So far we just rewrote the state. Now, let A and B send their particles  $A_2$  and  $B_2$  to some common point where they are measured in the Bell basis. The state of  $A_2$  and  $B_2$  then collapses to one of the four Bell states and due to the structure of the state  $|\psi\rangle$ , the particles  $A_1$  and  $B_1$  are left in the same state. So we start from  $A_1$  and  $A_2$  entangled and  $B_1$  and  $B_2$  entangled. After the measurement,  $A_1$  and  $B_1$  are entangled and  $A_2$  and  $B_2$  are entangled.  $A_1$  and  $B_1$  end in an entangled state without ever having interacted with each other.

## 5 Quantum computing

### 5.1 Shor factorization algorithm

The discovery by Peter Shor (in 1994) of the quantum algorithm for factorizing integers was the first example of an interesting problem where an efficient quantum algorithm was found. We will not go into the full details of the algorithm, only sketch the main ideas and point out some of the properties.

Let  $N$  be the number that we want to factorize, and  $L$  the number of binary digits. If  $x$  is any number, we call the smallest number  $r$  such that  $x^r \equiv 1 \pmod{N}$  the order of  $x$ . We use two facts from number theory:

1. If  $x^2 \equiv 1 \pmod{N}$  and  $x \not\equiv \pm 1 \pmod{N}$  either  $\gcd(x-1, N)$  or  $\gcd(x+1, N)$  is a factor in  $N$ . There is also an efficient classical algorithm to find it (using of the order  $L^3$  operation).
2. If  $N = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$  with  $p_i$  different primes and  $x$  is chosen randomly between 1 and  $N-1$  and  $r$  is the order of  $x$  then the probability that  $r$  is even and  $x^{r/2} \not\equiv -1 \pmod{N}$  is greater than  $1 - 1/2^m$ . This gives a nontrivial solution  $y = x^{1/2}$  with  $y^2 \equiv 1 \pmod{N}$ .

#### 5.1.1 The quantum Fourier transformation

We have a vector of  $N$  elements  $(x_0, x_1, \dots, x_{N-1})$ . The discrete Fourier transformation of this is defined as

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{2\pi i j k / N}.$$

We can encode the initial vector in a quantum state of  $n$  qubits (we assume for simplicity that  $N = 2^n$  is a power of 2) as

$$|\psi\rangle = \sum_j x_j |j\rangle$$

where

$$|j\rangle = |j_1 j_2 \cdots j_n\rangle = |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle$$

with  $j_1 j_2 \cdots j_n$  the binary digits of  $j$ . We define the unitary transformation  $U_{QFT}$  for the quantum Fourier transformation by the action on each of the basis states

$$U_{QFT}|j\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i j k / N} |k\rangle.$$

Then we see that

$$U_{QFT}|\psi\rangle = \sum_j y_k |k\rangle.$$

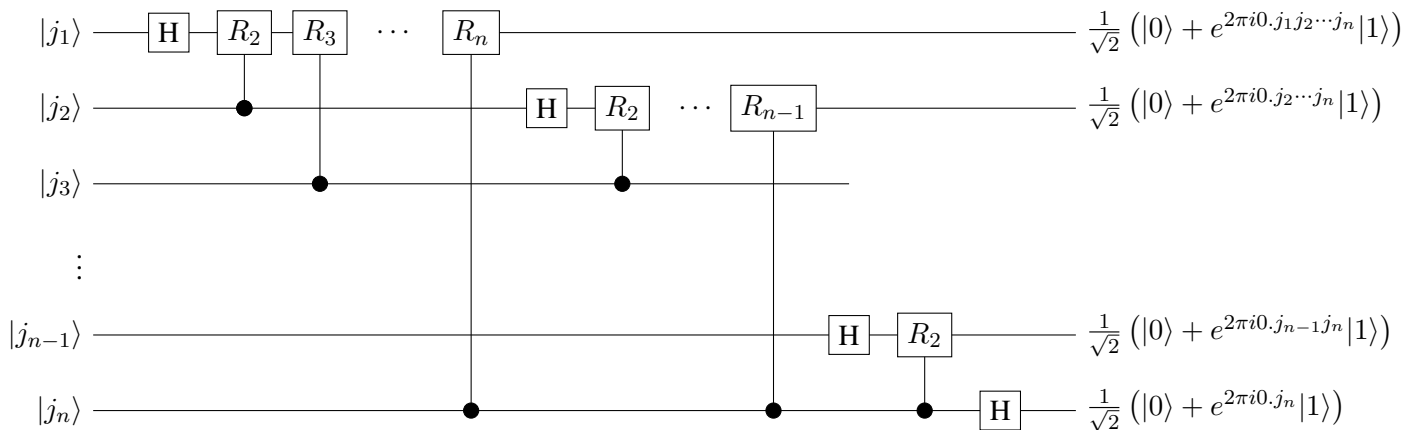
We can get better understanding if we examine the action of  $U_{QFT}$  on each of the qubits by rewriting ( $k = k_1 k_2 \cdots k_n = \sum_l k_l 2^{n-l}$ , where  $k = k_1 k_2 \cdots k_n$  are the binary digits of  $k$ )

$$\begin{aligned} U_{QFT}|j\rangle &= \frac{1}{2^{n/2}} \sum_{k=0}^{N-1} e^{2\pi i j k / 2^n} |k\rangle \\ &= \frac{1}{2^{n/2}} \sum_{k_1=0}^1 \cdots \sum_{k_n=0}^1 \bigotimes_{l=1}^n e^{2\pi i j k_l 2^{-l}} |k_l\rangle \\ &= \frac{1}{2^{n/2}} \bigotimes_{l=1}^n \left[ \sum_{k_l=0}^1 e^{2\pi i j k_l 2^{-l}} |k_l\rangle \right] \\ &= \frac{1}{2^{n/2}} \bigotimes_{l=1}^n \left[ |0\rangle + e^{2\pi i j 2^{-l}} |1\rangle \right]. \end{aligned}$$

Since  $j = j_1 j_2 \cdots j_n$  and dividing by  $2^l$  in binary notation is just a shift of the decimal point by  $l$  places, we have that in binary digits  $j 2^{-l} = j_1 j_2 \cdots j_{n-l} . j_{n-l+1} \cdots j_n$ . The integral part is irrelevant since  $e^{2\pi i j_1 j_2 \cdots j_{n-l}} = 1$ , which gives the QFT in the following form

$$U_{QFT}|j\rangle = \frac{1}{2^{n/2}} (|0\rangle + e^{2\pi i 0 . j_n} |1\rangle) (|0\rangle + e^{2\pi i 0 . j_{n-1} j_n} |1\rangle) \cdots (|0\rangle + e^{2\pi i 0 . j_1 \cdots j_n} |1\rangle). \quad (5)$$

We observe that the state is a product state, that is, it is not entangled. This gives hope that it can be realized by operations on single qubits, controlled by the state of the others (since the phases  $e^{2\pi i 0 . j_1 \cdots j_n}$  depend on the other qubits). We will show that the following quantum circuit will do the QFT



Here the single qubit operator

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i / 2^k} \end{pmatrix}$$

and these operations are controlled by the state of another qubit. This means that they are executed if the control qubit is in state  $|1\rangle$  while nothing is done if the state of the control qubit is in state  $|0\rangle$ . Let us follow the state of the first qubit

$$|j_1\rangle \xrightarrow{H} \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i 0 \cdot j_1} |1\rangle)$$

since  $e^{2\pi i 0 \cdot j_1} = e^{\pi i j_1}$  is 1 if  $j_1 = 0$  and  $-1$  if  $j_1 = 1$ . The controlled  $R_2$  gate gives the state

$$\frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i 0 \cdot j_1 j_2} |1\rangle).$$

Each successive  $R_k$  gate adds a similar phase factor, so that the final state is

$$\frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i 0 \cdot j_1 j_2 \cdots j_n} |1\rangle).$$

The action on the rest of the qubits is similar, and the final state is exactly the same as in Eq. (5), except that the order of the qubits is reversed. If we measure the state at this point, we can just read the result in the reverse order. If we are continuing the calculation, we can reverse the order by a number of SWAP gates (interchanging the state of two qubits).

The number of gates that is needed is  $1 + 2 + 3 + \cdots + n = \frac{1}{2}n(n+1)$ . That is, the number of gates increases as  $n^2$ . This is assuming that we can efficiently do the controlled  $R_k$  gates. We know that all two qubit gates can be decomposed in single qubit gates and CNOT. The Fast Fourier Transformation (FFT), which is the most efficient classical algorithm, needs of the order  $n2^n$  operations, so the quantum algorithm is much more efficient for large  $n$ . However, it is not generally useful in the same way as the classical FFT. Imagine for example that we have a measured time series with  $N = 2^n$  points and want to calculate the Fourier transform of this, which gives the frequency spectrum of the measured signal. We can encode this information in the amplitudes of  $n$  qubits, but to prepare this state needs in general at least  $2^n$  operations, as this is the number of measured values in our time series. Thus, the total number of operations can not scale better than  $2^n$ , similar to the FFT. In addition, we need to read out the final state, which at best will require  $2^n$  operations, as this is the number of values in the Fourier transform. In reality, it may take even more because of the random nature of the quantum measurement. The Shor factoring algorithm is efficient because it does not apply the QFT to a general state, but to a very specific state that has underlying periodicity and which we can generate efficiently. In addition, we are not interested in the full Fourier spectrum, but only in the period, which requires much less information. Because the signal we transform is periodic, the spectrum will have a very large amplitude at the corresponding period, and this is exactly the value we are going to measure.

## 5.2 Simulating quantum problems

The fact that the number of dimensions of the Hilbert space grows exponentially with the size of the system means that it is not possible to even represent the state of many quantum systems on a classical computer, and certainly not possible to compute any of its properties. As an illustration, consider a magnet consisting of a number  $n$  of atoms, each with a spin- $\frac{1}{2}$  and corresponding magnetic moment. The general state of this system is

$$|\psi\rangle = \sum_{j=0}^{2^n-1} c_j |j\rangle$$

with the  $2^n$  complex coefficients  $c_j$ . Imagine we have a  $4 \times 4 \times 4$  array of such atoms, hardly a large sample, and also certainly too small if we like to imagine we are approaching the thermodynamic limit and for example finding the critical temperature for the ferromagnet-paramagnet phase transition. Still, the number of coefficients is  $2^{64} = 1.8 \cdot 10^{18}$ , which would require of the order a million TB of computer memory, clearly not feasible. The solution to this is to use a quantum system to represent the state, that is, use a quantum computer. We have two alternatives:

1. We can use a universal quantum computer, consisting of a number of qubits with the possibility to realize a set of universal set of gates. The question is whether we can find an efficient algorithm to simulate a given system Hamiltonian on this computer. The advantage of it is that if such a general algorithm can be implemented, studying a new system just means reprogramming the computer. The disadvantage is that it likely requires a large number of good (low decoherence) qubits.
2. Use a quantum system that has Hamiltonian identical to the one that we are interested in simulating. This is called analog quantum simulation. The advantage of this approach is that it may require smaller resources in terms of the size of the system and the required control operations. The disadvantage is that it is specific to one special system, and if we want to study a different system we may have to build a new computer.

In both cases, an important question is how to get useful information out of the computer. In general, the full state has the exponentially large amount of information stored in the coefficient  $c_j$ , and we can not directly neither access this information nor process it in any way. As with all quantum algorithms, it is important that it can so be arranged that the measurement of the state will give the information we are interested in directly. We will give some examples of both approaches.

## 5.3 Digital quantum simulation

### 5.3.1 Time evolution

We would like to simulate the time evolution of a quantum system. In general, this is difficult and probably not efficient. But for a special class of Hamiltonians that are sums of terms involving a limited subspace each, it can be done in a way that does not scale exponentially with system size. We assume that the Hamiltonian can be written on the form

$$H = \sum_{k=1}^L H_k$$

where each term  $H_k$  acts on a subspace of at most  $m$  dimensions and  $m$  is independent of the size  $L$  of the system. Fortunately, this includes most realistic physical systems. Most solid-state systems fall in this class, for example the 1-dimensional Ising model for a magnet with the Hamiltonian

$$H = -B \sum_i \sigma_i^x - J \sum_i \sigma_i^z \sigma_{i+1}^z.$$

Here  $\sigma_i^x$  and  $\sigma_i^z$  are the usual Pauli matrices for spin  $i$ ,  $B$  is the external magnetic field and  $J$  is a coupling constant for neighbouring spins. Another example is the Hubbard model (in 1 dimension)

$$H = -t \sum_{i,\sigma} (a_{i,\sigma}^\dagger a_{i+1,\sigma} + a_{i,\sigma} a_{i+1,\sigma}^\dagger) + U \sum_i n_{i,\uparrow} n_{i,\downarrow}$$

where  $n_{i,\sigma} = a_{i,\sigma}^\dagger a_{i,\sigma}$  is the number operator for spin  $\sigma$  on site  $i$ .

A little more surprising is that also the motion of a single particle in a potential  $V(x)$  is of this type. The Hamiltonian is

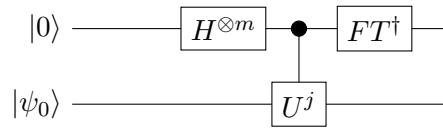
$$H = \frac{p^2}{2m} + V(x).$$

The second term is local in the position representation but not the first term.

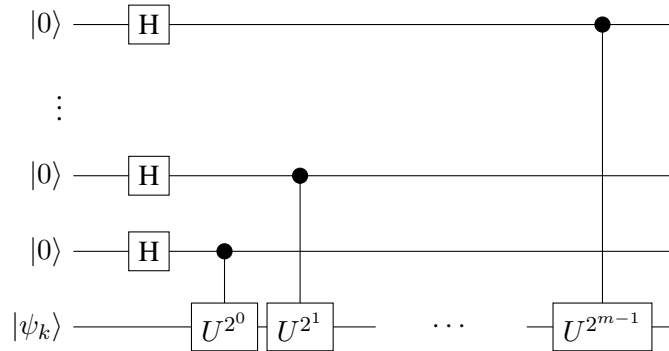
### 5.3.2 Finding eigenvalues

One of the main problems in quantum mechanics is determining the eigenvalues of the Hamiltonian. Here we describe how to do this on a quantum computer.

The general idea is described in the following quantum circuit



The upper line represents  $m$  qubits that we use to store the eigenvalue. The accuracy with which we can calculate it increases with increasing  $m$ . The lower line represents a number of qubits sufficient to store a state of the system that is studied.  $FT^\dagger$  represents the inverse Fourier transformation. The initial state  $|\psi_0\rangle$  of the lower line must have a non-zero inner product with the state we are calculating the eigenvalue for. We are assuming that we can efficiently prepare this as well as calculate the controlled  $U^j$  operation, controlled by the number  $j$  that is stored in the qubit in the upper line. We show how to do this, writing out the first part of the circuit in more detail by showing each of the  $m$  qubits corresponding to the upper line.



To make the evolution of the state through the circuit more transparent, we have assumed that the system register (lower line) is initially in an exact eigenstate  $|\psi_k\rangle$  of the Hamiltonian. If we let  $U = e^{-\frac{i}{\hbar}Ht}$ , the state  $|\psi_k\rangle$  is also an eigenstate of  $U$ . If

$$H|\psi_k\rangle = E_k|\psi_k\rangle$$

then

$$U|\psi_k\rangle = e^{-\frac{i}{\hbar}E_k t}|\psi_k\rangle = e^{2\pi i\phi}|\psi_k\rangle$$

where the phase  $\phi$  can be assumed to be between 0 and 1 and have binary digits  $0.\phi_1\phi_2\cdots\phi_m$ . Let us follow the evolution of the one of the qubits from the upper line. Let it be number  $j + 1$  from the

bottom, so that it is executing the controlled  $U^{2^j}$  operation. In the following equation, we write the state of this qubit and the register for the state  $|\psi_k\rangle$ , since they are the only ones affected at that step

$$\begin{aligned} |0\rangle|\psi_k\rangle &\xrightarrow{H} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|\psi_k\rangle \xrightarrow{U^{2^j}} \frac{1}{\sqrt{2}}(|0\rangle|\psi_k\rangle + |1\rangle U^{2^j}|\psi_k\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 2^j \phi} |1\rangle)|\psi_k\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0.\phi_{j+1}\dots\phi_m} |1\rangle)|\psi_k\rangle. \end{aligned}$$

Note the curious fact that even though the qubit controls the  $U^{2^j}$  operation acting on the state  $|\psi_k\rangle$ , it is the qubit state that is changed while the state  $|\psi_k\rangle$  is not. This means that this state propagates from one controlled  $U^{2^j}$  operation to the next unchanged, and the action on each of the control qubits is similar, only  $j$  changes along the register. This means that the final state is

$$\frac{1}{2^{m/2}} \left( |0\rangle + e^{2\pi i 0.\phi_m} |1\rangle \right) \left( |0\rangle + e^{2\pi i 0.\phi_{m-1}\phi_m} |1\rangle \right) \dots \left( |0\rangle + e^{2\pi i 0.\phi_1\dots\phi_m} |1\rangle \right).$$

If we compare this to the result of the quantum Fourier transformation, Eq (5), we see that the expressions are identical, and applying the inverse Fourier transformation gives a final state that is  $|\phi_1\phi_2\dots\phi_m\rangle$ . Measuring this we get the phase  $\phi$  from which we find the energy  $E_k = -\frac{2\pi\hbar\phi}{t}$ .

The number  $m$  of qubits determines the accuracy with which the energy is found. If one has a limited number of qubits there is a trick to increase the accuracy by repeated operation. After measuring the energy once with the available qubits, one subtracts the energy found from the Hamiltonian.

### 5.3.3 Example: Finding the bond length of $\text{H}_2$

As an example of how to use a quantum computer to solve an eigenvalue problem in quantum mechanics, we explain how one can find the ground state energy of a hydrogen molecule. In general, any molecule has the Hamiltonian

$$H = -\sum_i \frac{\nabla_{R_i}}{2M_i} - \sum_i \frac{\nabla_{r_i}}{2} - \sum_{ij} \frac{Z_i}{|R_i - r_j|} + \sum_{ij} \frac{Z_i Z_j}{|R_i - R_j|} + \sum_{ij} \frac{1}{|r_i - r_j|}$$

where  $R_i$ ,  $M_i$  and  $Z_i$  are the nuclear positions, masses and charges,  $r_i$  are the positions of the electrons and the energy is measured in Hartree,  $\hbar^2/m_e e^2 a_0$ , with  $m_e$  the electron mass and  $a_0$  the Bohr radius. To go from this Hamiltonian to what is actually simulated on the quantum computer is quite involved, and we will only give an outline<sup>13</sup>

1. We make the Born-Oppenheimer approximation, which means that we assume that the electrons move on a timescale much shorter than that of the nuclei. This means that we can assume the positions of the nuclei to be fixed parameters, calculate the electronic energy, and then find the positions that minimizes the energy.
2. We choose a set of 1-particle wavefunctions (orbitals),  $\phi_i(r)$ . An  $n$ -particle wavefunction can then be expanded in a basis of functions of the type

<sup>13</sup>If you want to read more, see O'Malley et al., Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016) and Seeley et al., The Bravyi-Kitaev transformation for quantum computation of electronic structure, J. Chem. Phys. **137**, 224109 (2012). The first three steps are standard in quantum chemistry, see for example T. Helgaker et al., Molecular Electronic-Structure Theory.

$$\Phi_{\{i_1, \dots, i_n\}}(r_1, \dots, r_n) = \phi_{i_1}(r_1)\phi_{i_2}(r_2) \cdots \phi_{i_n}(r_n) \quad + \text{Antisymmetrization}$$

where the antisymmetrization must be included since the electrons are Fermions. To simplify the following equations, we write  $I = \{i_1, \dots, i_n\}$  for the set of indices.

3. We define the annihilation and creation operators for the 1-particle states, so that

$$a_i^\dagger |0\rangle = |\phi_i\rangle$$

For Fermions they have to satisfy the anticommutation relations

$$\{a_i, a_j^\dagger\} = \delta_{ij}$$

An  $n$ -particle state is then

$$|\Psi_I\rangle = a_{i_1}^\dagger a_{i_2}^\dagger \cdots a_{i_n}^\dagger |0\rangle$$

and the anticommutation of the operators takes care of the antisymmetrization.

As usual, we can represent the Hamiltonian in the given basis, and write (note that the sum is over all possible index sets)

$$H = \sum_{IJ} \langle \Psi_I | H | \Psi_J \rangle |\Psi_I\rangle \langle \Psi_J|.$$

For an 1-particle operator like a potential  $V(r_i)$  we have

$$\langle \Psi_I | V(r_i) | \Psi_J \rangle = \int dr_1 \cdots dr_n \phi_{i_1}^*(r_1) \cdots \phi_{i_n}^*(r_n) V(r_i) \phi_{j_1}(r_1) \cdots \phi_{j_n}(r_n)$$

If our 1-particle states are orthonormal,

$$\int dr \phi_i^*(r) \phi_j(r) = \delta_{ij},$$

the two basis sets  $I$  and  $J$  can differ on at most one index to get a nonzero matrix element. The general form of the 1-particle terms in the Hamiltonian is then

$$H^{(1)} = \sum_{pq} h_{pq} a_p^\dagger a_q$$

with

$$h_{pq} = \int dr \phi_p^*(r) \left( -\frac{\nabla_r}{2} - \sum_i i \frac{Z_i}{|R_i - r|} \right) \phi_q(r).$$

Similarly, we find that the 2-particle terms in the Hamiltonian can be written on the form

$$H^{(2)} = \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$$

with

$$h_{pqrs} = \int dr_1 dr_2 \frac{\phi_p^*(r_2) \phi_q^*(r_2) \phi_s(r_1) \phi_r(r_2)}{|r_1 r_2|}$$



We apply these general formulas to a standard set of orbitals used to describe the Hydrogen molecule. There are two spatial orbitals, each a combination of three Gaussian functions<sup>14</sup>. For each spatial orbital there are two spin orientations, resulting in 4 spin oprbitals. In terms of these, the Hamiltonian is

$$H^{(1)} = h_{00}a_0^\dagger a_0 + h_{11}a_1^\dagger a_1 + h_{22}a_2^\dagger a_2 + h_{33}a_3^\dagger a_3.$$

$$H^{(2)} = h_{0110}a_0^\dagger a_1^\dagger a_1 a_0 + h_{2332}a_2^\dagger a_3^\dagger a_3 a_2 + h_{0330}a_0^\dagger a_3^\dagger a_3 a_0 + h_{1221}a_1^\dagger a_2^\dagger a_2 a_1 + (h_{0220} - h_{0202})a_0^\dagger a_2^\dagger a_2 a_0 \\ + (h_{1331} - h_{1313})a_1^\dagger a_3^\dagger a_3 a_1 + h_{0132}(a_0^\dagger a_1^\dagger a_3 a_2 + a_2^\dagger a_3^\dagger a_1 a_0) + h_{0312}(a_0^\dagger a_3^\dagger a_1 a_2 + a_2^\dagger a_1^\dagger a_3 a_0).$$

with the  $h_{pq}$  and  $h_{pqrs}$  determined by the above integrals. They are functions of the bond length (distance between the two protons).

4. To simulate the Hamiltonian on the quantum computer, we must map the states on the qubits the Hamiltonian on a corresponding qubit operator. There are several schemes for this. In the experiments, the Bravyi-Kitaev transformation was used. We need to express the terms in the Hamiltonian in terms of qubit operator so that the usual commutators for the Pauli matrices follow from the anticommutators of the creation and annihilation operators. There is a systematic way to do this, (see Seeley et al., The Bravyi-Kitaev transformation for quantum computation of electronic structure, J. Chem. Phys. **137**, 224109 (2012)), the following formulas is one possibility

$$a_0^\dagger a_0 = \frac{1}{2}(\mathbb{1} - \sigma_0^z) \\ a_1^\dagger a_1 = \frac{1}{2}(\mathbb{1} - \sigma_1^z \sigma_0^z) \\ a_2^\dagger a_2 = \frac{1}{2}(\mathbb{1} - \sigma_0^z) \\ a_3^\dagger a_3 = \frac{1}{2}(\mathbb{1} - \sigma_3^z \sigma_2^z \sigma_1^z)$$

$$a_0^\dagger a_1^\dagger a_1 a_0 = \frac{1}{4}(\mathbb{1} - \sigma_0^z - \sigma_1^z \sigma_0^z + \sigma_1^z) \\ a_2^\dagger a_3^\dagger a_3 a_2 = \frac{1}{4}(\mathbb{1} - \sigma_2^z - \sigma_3^z \sigma_2^z \sigma_1^z + \sigma_3^z \sigma_1^z) \\ a_0^\dagger a_3^\dagger a_3 a_0 = \frac{1}{4}(\mathbb{1} - \sigma_0^z - \sigma_3^z \sigma_2^z \sigma_1^z + \sigma_3^z \sigma_2^z \sigma_1^z \sigma_0^z) \\ a_1^\dagger a_2^\dagger a_2 a_1 = \frac{1}{4}(\mathbb{1} - \sigma_2^z - \sigma_1^z \sigma_0^z + \sigma_2^z \sigma_1^z \sigma_0^z) \\ a_0^\dagger a_2^\dagger a_2 a_0 = \frac{1}{4}(\mathbb{1} - \sigma_2^z - \sigma_0^z + \sigma_2^z \sigma_0^z) \\ a_1^\dagger a_3^\dagger a_3 a_1 = \frac{1}{4}(\mathbb{1} - \sigma_3^z \sigma_2^z \sigma_1^z - \sigma_1^z \sigma_0^z + \sigma_3^z \sigma_2^z \sigma_0^z).$$

---

<sup>14</sup>See Whitfield et al., Simulation of Electronic Structure Hamiltonians Using Quantum Computers, Molecular Physics **109**, pages 735 (2011).

$$\begin{aligned}
a_0^\dagger a_3^\dagger a_1 a_2 + a_2^\dagger a_1^\dagger a_3 a_0 &= \frac{1}{8} (-\sigma_2^x \sigma_0^x + \sigma_2^x \sigma_1^z \sigma_0^x - \sigma_2^y \sigma_0^y + \sigma_2^y \sigma_1^z \sigma_0^y - \sigma_3^z \sigma_2^x \sigma_0^x \\
&\quad + \sigma_3^z \sigma_2^x \sigma_1^z \sigma_0^x - \sigma_3^z \sigma_2^y \sigma_0^y + \sigma_3^z \sigma_2^y \sigma_1^z \sigma_0^y) \\
a_0^\dagger a_1^\dagger a_3 a_2 + a_2^\dagger a_3^\dagger a_1 a_0 &= \frac{1}{8} (\sigma_2^x \sigma_0^x + \sigma_2^x \sigma_1^z \sigma_0^x + \sigma_2^y \sigma_0^y + \sigma_2^y \sigma_1^z \sigma_0^y + \sigma_3^z \sigma_2^x \sigma_0^x \\
&\quad + \sigma_3^z \sigma_2^x \sigma_1^z \sigma_0^x + \sigma_3^z \sigma_2^y \sigma_0^y + \sigma_3^z \sigma_2^y \sigma_1^z \sigma_0^y).
\end{aligned}$$

Using these, the Hamiltonian can be expressed as

$$H = f_0 \mathbb{1} + f_1 \sigma_0^z + f_2 \sigma_1^z + f_3 \sigma_2^z + f_1 \sigma_1^z \sigma_0^z + f_4 \sigma_2^z \sigma_0^z + f_5 \sigma_3^z \sigma_1^z + f_6 \sigma_2^x \sigma_1^z \sigma_0^x + f_6 \sigma_2^y \sigma_1^z \sigma_0^y \quad (6)$$

$$+ f_7 \sigma_2^z \sigma_1^z \sigma_0^z + f_4 \sigma_3^z \sigma_2^z \sigma_0^z + f_3 \sigma_3^z \sigma_2^z \sigma_1^z + f_6 \sigma_3^z \sigma_2^x \sigma_1^z \sigma_0^x + f_6 \sigma_3^z \sigma_2^y \sigma_1^z \sigma_0^y + f_7 \sigma_3^z \sigma_2^z \sigma_1^z \sigma_0^z. \quad (7)$$

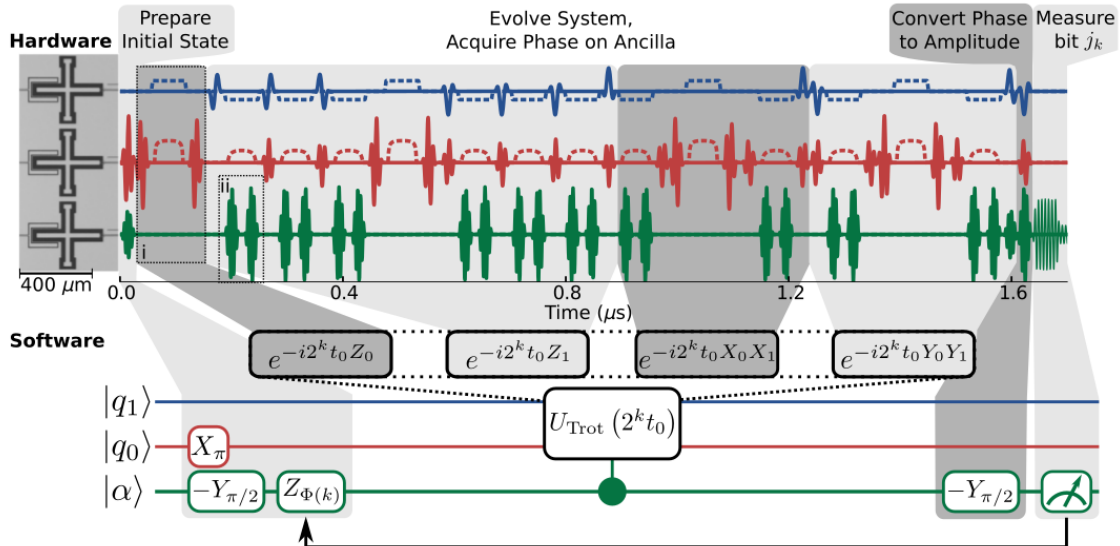
where  $f_i$  are functions of  $h_{pq}$  and  $h_{pqrs}$ .

We are going to use for the initial state  $\phi_0$  a product state (the Hartree-Fock solution) of basis vectors (that is, each qubit starts in an eigenstate of  $\sigma_i^z$ ). Looking at (6) we can see that for qubits 1 and 3 the Hamiltonian only has terms with  $\sigma_i^z$ . This means that qubits 1 and 3 are going to remain in the initial eigenstate throughout the whole calculation. We can then replace them by the initial value, and we do not need these qubits. Thus, we can replace the Hamiltonian by (renumbering qubit 2 to qubit 1)

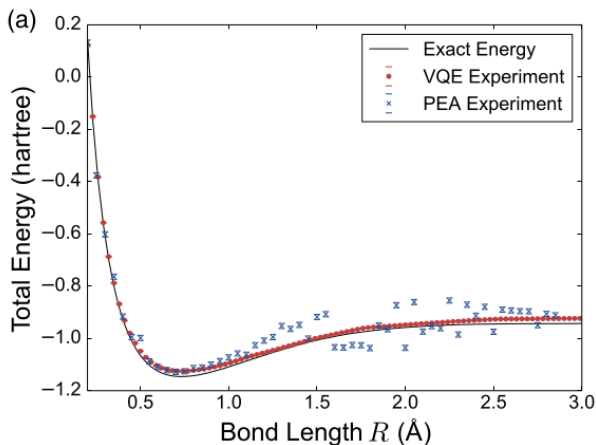
$$H = g_0 \mathbb{1} + g_1 \sigma_0^z + g_2 \sigma_1^z + g_3 \sigma_0^z \sigma_1^z + g_4 \sigma_0^x \sigma_1^x + g_3 \sigma_0^y \sigma_1^y,$$

where  $g_i$  can be expressed in terms of  $f_i$  and are functions of the bond length.

We are now ready to execute the program on the quantum computer. To reduce the number of qubits needed, we will use the repeated determination of the energy, as discussed at the end of Sec. 5.3.2. We use two qubits to represent the state and one qubit to measure the phase (energy), so three qubits in total. We leave out the detailed description of how to find the exact control pulses to execute each gate, the figure gives an impression of the number of actual pulses.



The resulting energy as a function of the bond length is shown below



As we can see, the result of this calculation (labeled PEA) agrees with the exact (calculated by conventional numerics) but the accuracy is not that good. Also shown are the results of a quantum calculation using the Variational Quantum Eigensolver (VQE), which is a variational method based on a parametrized form of the eigenstate (in this case the unitary coupled cluster ansatz). This method is less demanding in terms of quantum resources, but also limited by the accuracy of the trial wavefunction.

## 6 Open quantum systems

In many situations, one has a quantum system of interest that one can control and measure but which is coupled (hopefully not too strongly) to some environment that one can not control or measure. Often we do not even know the details of the environment, such as its exact Hamiltonian or initial state. In general, the dynamics of the system will depend on all the details of the environment and the system-environment coupling. To describe all the details, it is then no other way than to specify the full model, including the environment degrees of freedom and Hamiltonian and study the full system+environment. However, in many cases one can get a very accurate description of the evolution of the system only in terms of a small number of parameters characterizing the environment. Here we will derive the Lindblad equation for the evolution of the system state in the presence of an environment and study the conditions and approximations that underly it. Then, we will study the solution of this equation for some example systems.

### 6.1 Derivation of the Lindblad equation

What we want to achieve can be illustrated by the following diagram

$$\begin{array}{ccc}
 \rho(0) = \rho_S(0) \otimes \rho_E(0) & \xrightarrow{\text{Unitary evolution}} & \rho(t) = U\rho(0)U^{-1} \\
 \downarrow \text{Tr}_E & & \downarrow \text{Tr}_E \\
 \rho_S(0) & \xrightarrow{\text{Dynamical map}} & \rho_S(t) = V(t)\rho_S(0)
 \end{array}$$

Here  $\rho_S$  is the system density matrix,  $\rho_E$  is the environment density matrix, and we assume that at  $t = 0$  they are uncorrelated so that  $\rho(0) = \rho_S(0) \otimes \rho_E(0)$ . The goal is to have an equation describing the dynamical map  $V(t)$  (in most cases approximately) that gives the evolution of the system density matrix without having to follow the evolution of the environment.

The total Hamiltonian has the form

$$H = H_S + H_E + H_I = H_0 + H_I$$

where  $H_S$  describes the system,  $H_E$  the environment and  $H_I$  is the interaction between the system and the environment which in general will have the form

$$H_I = \sum_{\alpha} S_{\alpha} \otimes E_{\alpha}$$

where  $S_{\alpha}$  are operators acting on the system and  $E_{\alpha}$  operators acting on the environment.

The equation for the time evolution of the total density matrix is

$$\frac{d\rho(t)}{dt} = -\frac{i}{\hbar}[H, \rho]$$

We pass to an interaction picture with the transformation

$$\rho^I = U_0^{-1} \rho U_0 \quad U_0 = e^{-\frac{i}{\hbar} H_0 t},$$

The density matrix has the time evolution given by

$$\frac{d\rho^I(t)}{dt} = -\frac{i}{\hbar}[H_I^I(t), \rho^I(t)] \quad (8)$$

with the interaction Hamiltonian in the interaction picture given by

$$H_I^I = U_0^{-1} H_I U_0 = \sum_{\alpha} e^{\frac{i}{\hbar} H_S t} S_{\alpha} e^{-\frac{i}{\hbar} H_S t} e^{\frac{i}{\hbar} H_E t} E_{\alpha} e^{-\frac{i}{\hbar} H_E t} = \sum_{\alpha} S_{\alpha}(t) E_{\alpha}(t).$$

We can formally integrate (8) to get

$$\rho^I(t) = \rho^I(0) - \frac{i}{\hbar} \int_0^t dt' [H_I^I(t'), \rho^I(t')].$$

Note that this is not the same as solving the differential equation, as the unknown function  $\rho^I(t)$  still appears on the right hand side. We can insert this expression back into (8) and find that

$$\frac{d\rho^I(t)}{dt} = -\frac{i}{\hbar}[H_I^I(t), \rho^I(0)] - \frac{1}{\hbar^2} \int_0^t dt' [H_I^I(t), [H_I^I(t'), \rho^I(t')]].$$

We are interested in having an equation for the system density matrix  $\rho_S^I = \text{Tr}_E \rho^I$  (still in the interaction picture), so we trace over the environment to get

$$\frac{d\rho_S^I(t)}{dt} = -\frac{i}{\hbar} \text{Tr}_E [H_I^I(t), \rho^I(0)] - \frac{1}{\hbar^2} \int_0^t dt' \text{Tr}_E [H_I^I(t), [H_I^I(t'), \rho^I(t')]]. \quad (9)$$

So far, this is exact and completely general without any assumptions or approximations. We now make two simplifying assumptions. The first is that at the initial time  $t = 0$ , the system and environment are uncorrelated, so that the density matrix is  $\rho(0) = \rho_S(0) \otimes \rho_E(0)$ . This can be because they did not

interact prior to  $t = 0$  or if correlations between them are decaying on timescales short compared to the system dynamics (we will return to this point below). The second assumption is that the initial state of the environment is in a stationary state with respect to the environment Hamiltonian  $H_E$ , which means that  $[H_E, \rho_E(0)] = 0$ . The typical situation is that the environment is in thermal equilibrium at  $t = 0$ , so that  $\rho_E(0) = e^{-\beta H_E} / \text{Tr} e^{-\beta H_E}$  with  $\beta = 1/k_B T$ . With these assumptions we can show that the first term on the right hand side of Eq. (9) can in general be taken to be zero. We have

$$\text{Tr}_E [H_I^I(t), \rho^I(0)] = \sum_{\alpha} S_{\alpha}(t) \rho_S^I(0) \text{Tr}_E [E_{\alpha}(t) \rho_E^I(0)] - \sum_{\alpha} \rho_S^I(0) S_{\alpha}(t) \text{Tr}_E [\rho_E^I(0) E_{\alpha}(t)]$$

and with a stationary  $\rho_E(0)$  the average

$$\langle E_{\alpha} \rangle = \text{Tr}_E [E_{\alpha}(t) \rho_E^I(0)] = \text{Tr}_E [E_{\alpha} \rho_E^I(0)]$$

is time independent. If this average is nonzero, we may define new operators  $E'_{\alpha} = E_{\alpha} - \langle E_{\alpha} \rangle$  and write the interaction as

$$H_I = \sum_{\alpha} S_{\alpha} \otimes E_{\alpha} = \sum_{\alpha} S_{\alpha} \otimes E'_{\alpha} + \sum_{\alpha} \langle E_{\alpha} \rangle S_{\alpha}.$$

The last term acts only on the system and can be included in  $H_S$  and clearly  $\langle E'_{\alpha} \rangle = 0$ .

To proceed, we make the Born approximation which is to assume that the environment is always in the stationary initial state and remains uncorrelated with the system. This is not true, as there will be correlations developing due to the interaction (in fact, it is precisely these correlations that cause the system state to be affected by the environment and which is what we want to use the Lindblad equation to study). But if the interaction is weak, which means that the correlations grow slowly, and the environment sufficiently complex so that it decays back to the thermal state quickly, it can be a good approximation. We can then write

$$\begin{aligned} \frac{d\rho_S^I(t)}{dt} &= -\frac{1}{\hbar^2} \int_0^t dt' \text{Tr}_E [H_I^I(t), [H_I^I(t'), \rho_S^I(t') \otimes \rho_E^I]] \\ &= -\frac{1}{\hbar^2} \int_0^t dt' \sum_{\alpha\beta} \text{Tr}_E [S_{\alpha}(t) \otimes E_{\alpha}(t), [S_{\beta}(t') \otimes E_{\beta}(t'), \rho_S^I(t') \otimes \rho_E^I]]. \end{aligned}$$

Note that this equation tells us that the change in  $\rho_S^I(t)$  depends on an integral over all the previous history,  $\rho_S^I(t')$  for all  $t' < t$ . We say that it is non-local in time. This is quite natural, since the action of the system on the environment at time  $t'$  will cause a change in the environment which acts back on the system at the later time  $t$ . With some approximation, we will now see that we can make it local in time. We define the correlation function of the environment as

$$C_{\alpha\beta}(t, t') = \text{Tr}_E [E_{\alpha}(t) E_{\beta}(t') \rho_E],$$

which for stationary  $\rho_E$  depends only on the time difference

$$C_{\alpha\beta}(t, t') = \text{Tr}_E [e^{iH_E t} E_{\alpha} e^{iH_E t'} e^{-iH_E t'} E_{\beta} e^{-iH_E t} \rho_E] = \text{Tr}_E [E_{\alpha}(t - t') E_{\beta}(0) \rho_E] = C_{\alpha\beta}(t - t').$$

In terms of this correlation function we can express

$$\begin{aligned} \frac{d\rho_S^I(t)}{dt} = & -\frac{1}{\hbar^2} \int_0^t dt' \sum_{\alpha\beta} \{ C_{\alpha\beta}(t-t') [S_\alpha(t)S_\beta(t')\rho_S^I(t') - S_\beta(t')\rho_S^I(t')S_\alpha(t)] \\ & + C_{\beta\alpha}(t'-t) [\rho_S^I(t')S_\beta(t')S_\alpha(t) - S_\alpha(t)\rho_S^I(t')S_\beta(t')] \}. \end{aligned}$$

The correlation function  $C_{\beta\alpha}(\tau)$  is large for small time differences  $\tau = t - t'$ , and decays fast for time differences larger than a characteristic time which we call the correlation time  $\tau_c$ . In the integral over  $t'$ , only  $t'$  with  $t - t' < \tau_c$  will give large contributions. We now make the Markov approximation, which is to replace  $\rho_S^I(t')$  by  $\rho_S^I(t)$  inside the integral. This is justified as long as the relaxation time of the environment is short compared to the timescale for the change of the state of the system. Changing the integration variable to  $\tau$  we get

$$\begin{aligned} \frac{d\rho_S^I(t)}{dt} = & -\frac{1}{\hbar^2} \int_0^t d\tau \sum_{\alpha\beta} \{ C_{\alpha\beta}(\tau) [S_\alpha(t)S_\beta(t-\tau)\rho_S^I(t) - S_\beta(t-\tau)\rho_S^I(t)S_\alpha(t)] \\ & + C_{\beta\alpha}(-\tau) [\rho_S^I(t)S_\beta(t-\tau)S_\alpha(t) - S_\alpha(t)\rho_S^I(t)S_\beta(t-\tau)] \}. \quad (10) \end{aligned}$$

Since the correlation function decays fast for  $\tau > \tau_c$  the main contribution to the integral comes from small  $\tau$ . For  $t \gg \tau_c$  it does not matter what the upper limit of the integral is, and we can replace it by infinity without changing the value of the integral significantly. The right hand side of the equation now depends on  $\rho_S^I(t)$  at the time  $t$ , so this equation is local in time. The integral over  $\tau$  could now be performed directly if we knew the time dependence in the operators  $S_\alpha(t)$ . We will now show how to do this with some manipulations and one additional approximation.

We denote the system eigenstates  $|\epsilon\rangle$ , so that

$$H_S|\epsilon\rangle = \epsilon|\epsilon\rangle$$

with projectors  $\Pi(\epsilon) = |\epsilon\rangle\langle\epsilon|$ . We define the operators

$$S_\beta(\omega) = \sum_{\epsilon-\epsilon'=\omega} \Pi(\epsilon)S_\beta\Pi(\epsilon')$$

and check that they satisfy the following relations

$$\begin{aligned} \sum_\omega S_\beta(\omega) &= \sum_\omega \sum_{\epsilon-\epsilon'=\omega} \Pi(\epsilon)S_\beta\Pi(\epsilon') = \sum_{\epsilon,\epsilon'} \Pi(\epsilon)S_\beta\Pi(\epsilon') = S_\beta, \\ [H_S, S_\beta(\omega)] &= -\omega S_\beta(\omega) \end{aligned}$$

and

$$[H_S, S_\beta^\dagger(\omega)] = \omega S_\beta^\dagger(\omega).$$

This implies that in the interaction picture we have

$$S_\beta(\omega, t) = e^{iH_S t} S_\beta(\omega) e^{-iH_S t} = e^{-i\omega t} S_\beta(\omega).$$

We therefore get

$$S_\alpha(t) = \sum_\omega S_\alpha(\omega, t) = \sum_\omega e^{-i\omega t} S_\alpha(\omega)$$

which we can insert into equation (10) to get

$$\frac{d\rho_S^I(t)}{dt} = - \sum_{\alpha\beta\omega\omega'} \left\{ e^{i(\omega' - \omega)\tau} \left[ S_\alpha(\omega), S_\beta^\dagger(\omega') \rho_S^I(t) \right] + h.c. \right\}. \quad (11)$$

where we have defined

$$\Gamma_{\alpha\beta}(\omega) = \frac{1}{\hbar^2} \int_0^\infty d\tau C_{\alpha\beta}(\tau) e^{-i\omega\tau}.$$

In Eq. (11), the terms with  $\omega' \neq \omega$  will be rapidly oscillating, and their effect will average out, giving small contributions to the time evolution of  $\rho_S^I(t)$ . We will ignore these terms, keeping only the resonant terms with  $\omega' = \omega$ . This is called the rotating wave approximation, and we used it previously in connection with the Rabi problem in a oscillating field.

We split  $\Gamma_{\alpha\beta}(\omega)$  in real and imaginary parts

$$\Gamma_{\alpha\beta}(\omega) = \frac{1}{2} \gamma_{\alpha\beta}(\omega) + i S_{\alpha\beta}(\omega)$$

with

$$\begin{aligned} \gamma_{\alpha\beta}(\omega) &= \Gamma_{\alpha\beta}(\omega) + \Gamma_{\alpha\beta}^*(\omega), \\ S_{\alpha\beta}(\omega) &= \frac{1}{2i} (\Gamma_{\alpha\beta}(\omega) - \Gamma_{\alpha\beta}^*(\omega)), \end{aligned}$$

and get

$$\frac{d\rho_S^I(t)}{dt} = -\frac{i}{\hbar} [H_{LS}, \rho_S^I(t)] - \frac{1}{2} \sum_{\alpha\beta\omega} \gamma_{\alpha\beta}(\omega) \left( S_\alpha(\omega) S_\beta^\dagger(\omega) \rho_S^I(t) + \rho_S^I(t) S_\alpha(\omega) S_\beta^\dagger(\omega) - 2 S_\beta^\dagger(\omega) \rho_S^I(t) S_\alpha(\omega) \right).$$

Here

$$H_{LS} = \hbar \sum_{\alpha\beta\omega} S_{\alpha\beta}(\omega) S_\beta^\dagger(\omega) S_\alpha(\omega)$$

is called the Lamb shift Hamiltonian, and it does not contribute to the decay of the state. It represents a renormalization of the system Hamiltonian due to the interaction with the environment.

The final step is to diagonalize the matrix  $\gamma_{\alpha\beta}(\omega)$ , which means that we find a unitary matrix  $M(\omega)$  and diagonal matrix  $D(\omega)$  so that

$$\gamma_{\alpha\beta}(\omega) = M_{\alpha\gamma}(\omega) D_{\gamma\delta}(\omega) M_{\delta\beta}^\dagger(\omega).$$

Defining the operators

$$L_\gamma(\omega) = M_{\alpha\gamma}(\omega) S_\alpha^\dagger$$

and denoting the diagonal elements  $D_{\alpha\alpha} = \gamma_\alpha(\omega)$  we get the Lindblad equation

$$\frac{d\rho}{dt} = -\frac{i}{\hbar} [H, \rho] - \frac{1}{2} \sum_{\alpha\omega} \gamma_\alpha(\omega) (L_\alpha(\omega)^\dagger L_\alpha(\omega) \rho + \rho L_\alpha(\omega)^\dagger L_\alpha(\omega) - 2 L_\alpha(\omega) \rho L_\alpha(\omega)^\dagger).$$

Here  $H = H_S + H_{LS}$ .

## 6.2 Solving the Lindblad equation

The general form of the Lindblad equation is

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] - \frac{1}{2} \sum_i \gamma_i (L_i^\dagger L_i \rho + \rho L_i^\dagger L_i - 2L_i \rho L_i^\dagger)$$

with some set of Lindblad operators,  $L_i$ . In the following we will solve some illustrative examples.

### 6.2.1 Two level system at zero temperature

Let the Hamiltonian be

$$H = \frac{1}{2} \hbar \omega \sigma_z \quad (12)$$

We have the basis states

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

and we assume  $\omega > 0$  so that  $|0\rangle$  is the ground state. There is only one Lindblad operator,  $L = |0\rangle\langle 1| = \sigma_-$ , corresponding to the emission of one photon while the TLS goes from the excited to the ground state ( $\sigma_\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y)$ ). We write the density matrix in the form

$$\rho = \begin{pmatrix} p_1 & d \\ d^* & p_0 \end{pmatrix} \quad (13)$$

with the condition  $p_0 = 1 - p_1$  and get the Lindblad equation in matrix form

$$\frac{d}{dt} \begin{pmatrix} p_1 & d \\ d^* & p_0 \end{pmatrix} = -i\omega \begin{pmatrix} 0 & d \\ -d^* & 0 \end{pmatrix} - \frac{\gamma}{2} \begin{pmatrix} 2p_1 & d \\ d^* & -2p_1 \end{pmatrix}.$$

From this we get the equations

$$\begin{aligned} \frac{dp_1}{dt} &= -\gamma p_1 \\ \frac{dd}{dt} &= (-i\omega - \frac{\gamma}{2})d, \end{aligned}$$

with the solution

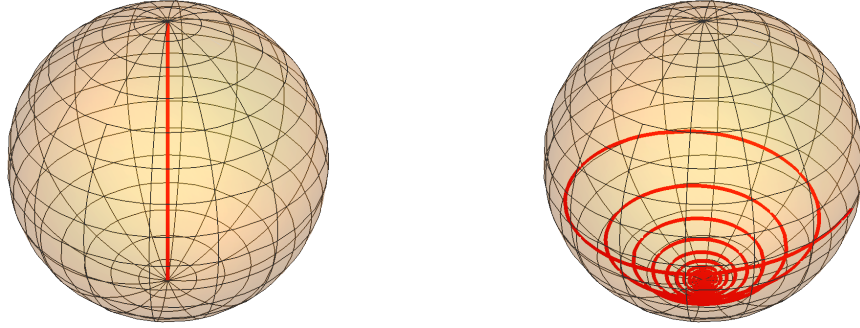
$$p_1(t) = p_1(0)e^{-\gamma t}, \quad d(t) = d(0)e^{i\omega t} e^{-\frac{\gamma}{2}t}.$$

To visualize the evolution of the state, we determine the Bloch-vector

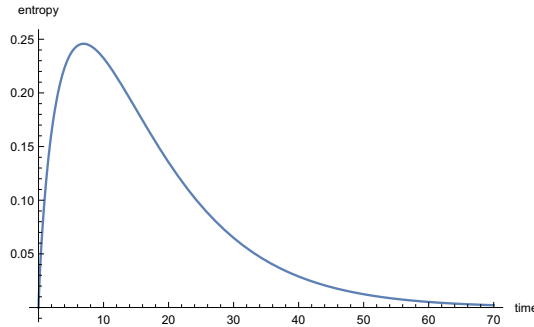
$$\begin{aligned} x &= 2 \operatorname{Re} d = (x(0) \cos \omega t - y(0) \sin \omega t) e^{-\frac{\gamma}{2}t} \\ y &= 2 \operatorname{Im} d = (x(0) \sin \omega t + y(0) \cos \omega t) e^{-\frac{\gamma}{2}t} \\ z &= (z(0) + 1)e^{-\gamma t} - 1. \end{aligned}$$

Two examples of the time evolution of the Bloch vector are shown below. On the left, the initial state is  $|1\rangle$ , corresponding to the Bloch vector  $(0, 0, 1)$ . On the right, the initial state has Bloch vector  $(1, 0, 0)$ . In both plots we have used  $\gamma = 0.1$  and  $\omega = 1$ .





We see that the evolution is a combination of decay to the ground state and precession in the  $xy$ -plane, resulting in a spiral trajectory. Note that the trajectory is not on the surface of the Bloch sphere, so the state is not a pure state, except for the initial state and the asymptotic state at  $t \rightarrow \infty$  which is the ground state  $|0\rangle$ . This we can quantify by showing the entropy as a function of time (for the initial Bloch vector  $(1, 0, 0)$  corresponding to the right figure above)



### 6.2.2 Two level system at finite temperature

At a nonzero temperature, the environment will have thermal excitations. To be concrete, we can think of the environment as the electromagnetic field and it means that we have a thermal radiation field surrounding the two-level system. Then we can have both emission and absorption processes, with the corresponding Lindblad operators  $L_1 = \sigma_-$  and  $L_2 = \sigma_+$ . The Lindblad equation is now

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] - \frac{\gamma_-}{2}(\sigma_+\sigma_-\rho + \rho\sigma_+\sigma_- - 2\sigma_-\rho\sigma_+) - \frac{\gamma_+}{2}(\sigma_-\sigma_+\rho + \rho\sigma_-\sigma_+ - 2\sigma_+\rho\sigma_-)$$

Here,  $\gamma_-$  is the emission rate and  $\gamma_+$  the absorption rate. These can in principle be determined from a detailed model of the interaction between the two-level system and the environment. However, we are satisfied to consider them as parameters of our theory. We can relate their ratio to the temperature,  $T$ , by the following argument. We know that

$$\gamma_+ \sim n \quad \text{and} \quad \gamma_- \sim n + 1$$

where  $n$  is the phonon number in one mode of the environment (which we consider to be described by harmonic oscillators) and the proportionality constants are the same in both expressions. We also know that in thermal equilibrium we have

$$n = \frac{1}{e^{\beta\hbar\omega} - 1}$$

where  $\beta = \frac{1}{k_B T}$  and  $\hbar\omega$  is the energy splitting of the two-level system, and therefore also the energy of the phonons emitted and absorbed. Then we have

$$\frac{\gamma_+}{\gamma_-} = \frac{n}{n+1} = e^{-\beta\hbar\omega}.$$

With the Hamiltonian (12) and the density matrix parametrized as (13) we get the Lindblad equation

$$\frac{d}{dt} \begin{pmatrix} p_1 & d \\ d^* & p_0 \end{pmatrix} = -i\omega \begin{pmatrix} 0 & d \\ -d^* & 0 \end{pmatrix} - \frac{\gamma_-}{2} \begin{pmatrix} 2p_1 & d \\ d^* & -2p_1 \end{pmatrix} - \frac{\gamma_+}{2} \begin{pmatrix} -2p_0 & d \\ d^* & 2p_0 \end{pmatrix}.$$

From this we get the equations

$$\begin{aligned} \frac{dp_1}{dt} &= -\gamma_- p_1 + \gamma_+ p_0 = -\gamma p_1 + \gamma_+ \\ \frac{dd}{dt} &= \left(-i\omega - \frac{\gamma}{2}\right)d, \end{aligned}$$

with  $\gamma = \gamma_+ + \gamma_-$ . The solution of these equations is

$$p_1(t) = \frac{\gamma_+}{\gamma} + (p_1(0) - \frac{\gamma_+}{\gamma})e^{-\gamma t}, \quad d(t) = d(0)e^{i\omega t} e^{-\frac{\gamma}{2}t}.$$

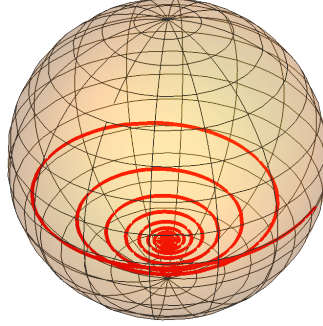
In the limit  $t \rightarrow \infty$  we approach a stationary state

$$d = 0, \quad p_0 = \frac{\gamma_-}{\gamma}, \quad \frac{\gamma_+}{\gamma}.$$

We can check that in this stationary state

$$\frac{p_1}{p_0} = \frac{\gamma_+}{\gamma_-} = e^{-\beta\hbar\omega}.$$

So the two level system will reach thermal equilibrium at the same temperature as the environment. The trajectory of the Bloch vector is similar to the zero temperature case, except that the final state now is the thermal equilibrium state (the figure corresponds to a temperature where  $\beta\hbar\omega = 1.6$ ).



### 6.2.3 Harmonic oscillator at zero temperature

We consider a harmonic oscillator with Hamiltonian

$$H = \hbar\omega\hat{a}^\dagger\hat{a}.$$

The eigenstates are  $|n\rangle$  with corresponding energies  $E_n = n\hbar\omega$ . As a simple model, we use only one Lindblad operator  $L = \hat{a}$ , the annihilation operator. This gives the Lindblad equation

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] - \frac{\gamma}{2}(\hat{a}^\dagger\hat{a}\rho + \rho\hat{a}^\dagger\hat{a} - 2\hat{a}\rho\hat{a}^\dagger)$$

Note that in general we could have one independent Lindblad operator for each transition  $|n\rangle \rightarrow |m\rangle$ , with independent constants  $\gamma_n$ . To see the implications of the fact that we use the same rate constant for all, we consider the equations for the matrix elements  $\rho_{mn} = \langle m|\rho|n\rangle$ . Taking the matrix elements of the Lindblad equation, we have that

$$\begin{aligned} \langle m|[H, \rho]|n\rangle &= \rho_{mn}(E_m - E_n) \\ \langle m|\hat{a}^\dagger\hat{a}\rho + \rho\hat{a}^\dagger\hat{a} - 2\hat{a}\rho\hat{a}^\dagger|n\rangle &= (m+n)\rho_{mn} - 2\sqrt{m+1}\sqrt{n+1}\rho_{m+1, n+1} \end{aligned}$$

We consider the diagonal elements  $p_n = \rho_{nn}$

$$\frac{dp_n}{dt} = -n\gamma p_n + (n+1)\gamma p_{n+1}.$$

The first term represents transitions  $|n\rangle \rightarrow |n-1\rangle$  with rate  $n\gamma$  and the second term represents transitions  $|n+1\rangle \rightarrow |n\rangle$  with rate  $(n+1)\gamma$ . We see that our assumption that  $\gamma$  is the same for all transitions means that the transition rate is proportional to  $n$ , the excitation state of the oscillator. To see that this can be realistic in a reasonable physical model, we recall the transition rate of dipole transitions, summed over final photon states, as given in the lecture notes, Eq. (4.101)

$$w_{BA} = \frac{4\alpha}{3c^2}\omega_{AB}^3|x_{BA}|^2$$

where we have assumed that the oscillator is oriented along the  $x$ -axis. The matrix element is

$$x_{BA} = \langle B|b|A \rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle n|\hat{a}^\dagger + \hat{a}|n+1 \rangle \sim \sqrt{n}$$

which gives  $w_{BA} = \gamma n$  as we have in the Lindblad equation.

We can find how the energy relaxes in this model. If  $n = \langle \hat{a}^\dagger \hat{a} \rangle$  is the average excitation state of the oscillator, we have

$$\begin{aligned} \frac{dn}{dt} &= \frac{d}{dt} \text{Tr}(\hat{a}^\dagger \hat{a} \rho) = \sum_n \langle n|\hat{a}^\dagger \hat{a} \frac{d\rho}{dt}|n \rangle = \sum_n n \frac{dp_n}{dt} = \gamma \sum_n n(-n\gamma p_n + (n+1)\gamma p_{n+1}) \\ &= -\gamma \sum_n (n^2 p_n - n(n+1)p_{n+1} - (n+1)p_{n+1} + (n+1)p_{n+1}) \\ &= -\gamma \sum_n (n^2 p_n - (n+1)^2 p_{n+1} + (n+1)p_{n+1}) = -\gamma \sum_n (n+1)p_{n+1} = -\gamma n \end{aligned}$$

which shows that the average excitation number, and therefore also the average energy decays exponentially,

$$n = n(0)e^{-\gamma t}.$$

We can also study more in detail how the state changes with time if we assume that the initial state is an energy eigenstate  $|n\rangle$ . The Lindblad equation for the matrix elements is

$$\frac{d\rho_{mn}}{dt} = -\frac{i}{\hbar} \rho_{mn}(E_m - E_n) - \frac{\gamma}{2}(m+n)\rho_{mn} + \gamma\sqrt{m+1}\sqrt{n+1}\rho_{m+1,n+1}$$

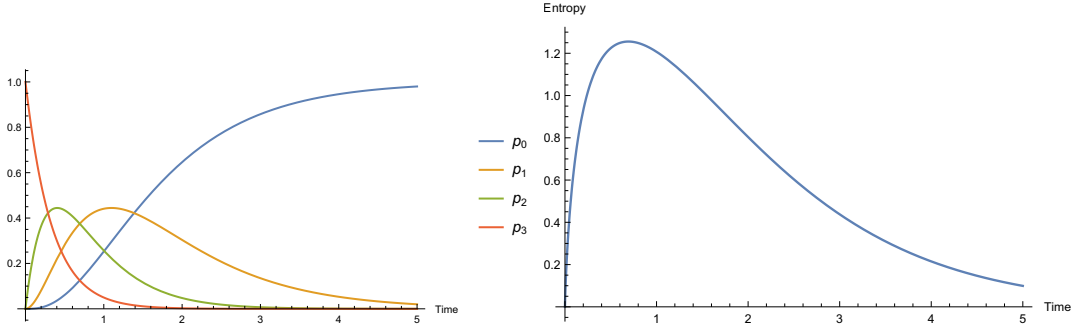
We see that the change in a given matrix element only depends on the same element, and the one diagonally adjacent. In particular, this means that if the density matrix is initially diagonal, it will remain so forever. So is we are initially in the state  $|n\rangle$  where  $p_n = \rho_{nn} = 1$  and all other matrix elements initially are zero, we can study the equations

$$\frac{dp_n}{dt} = -n\gamma p_n + (n+1)\gamma p_{n+1}.$$

I do not know any way to do this except systematically working through the equations for a given initial state. As an example, let us consider that we start from  $|3\rangle$ , so  $p_3(0) = 1$  and  $p_0(0) = p_1(0) = p_2(0) = 0$ . Then we have

$$\begin{aligned} \dot{p}_3 &= -3\gamma p_3 \implies p_3 = e^{-3\gamma t} \\ \dot{p}_2 &= -2\gamma p_2 + 3\gamma p_3 = -2\gamma p_2 + 3\gamma e^{-3\gamma t} \implies p_2 = 3(e^{-2\gamma t} - e^{-3\gamma t}) \\ \dot{p}_1 &= -\gamma p_1 + 2\gamma p_2 = -\gamma p_1 + 3\gamma(e^{-2\gamma t} - e^{-3\gamma t}) \implies p_1 = 3e^{-3\gamma t}(e^{\gamma t} - 1)^2 \\ \dot{p}_0 &= \gamma p_1 = 3\gamma e^{-3\gamma t}(e^{\gamma t} - 1)^2 \implies p_0 = e^{-3\gamma t}(e^{\gamma t} - 1)^3 \end{aligned}$$

A plot of the  $p_i$  as functions of time are shown below together with the entropy.



As we should expect, the state approaches  $|0\rangle$  at long times. The entropy is largest when all four levels are of the same order of magnitude, and decays as the state becomes more and more pure at long times. The fact that the entropy is nonzero indicates that the system is not in a pure state. This is because it has been entangled with the environment. The behaviour of this model is similar to what we observed for the two-level system at zero temperature in Sec. 6.2.1.

An interesting case to study, which shows a different behaviour, is if we assume the initial state to be a coherent state instead of an energy eigenstate. We define the following transformation of the density matrix

$$C(\lambda, \lambda^*, t) = \text{Tr} \left( \rho e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right)$$

We find the derivatives

$$\begin{aligned} \frac{\partial C}{\partial \lambda} &= \text{Tr} \left( \rho \hat{a}^\dagger e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) \\ \frac{\partial C}{\partial \lambda^*} &= -\text{Tr} \left( \hat{a} \rho e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) \end{aligned}$$

$$\frac{\partial C}{\partial t} = \text{Tr} \left( \frac{\partial \rho}{\partial t} e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) = -i\omega \text{Tr} \left[ (\hat{a}^\dagger \hat{a} \rho - \rho \hat{a}^\dagger \hat{a}) e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right] - \frac{\gamma}{2} \text{Tr} \left[ (\hat{a}^\dagger \hat{a} \rho + \rho \hat{a}^\dagger \hat{a} - 2\hat{a} \rho \hat{a}^\dagger) e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right] \quad (14)$$

From the general formula

$$e^B A e^{-B} = A + [B, A] + \dots$$

We derive the relations

$$\begin{aligned} e^{-\lambda \hat{a}^\dagger} \hat{a} e^{\lambda \hat{a}^\dagger} &= \hat{a} - \lambda [\hat{a}^\dagger, \hat{a}] = \hat{a} + \lambda \\ e^{\lambda \hat{a}^\dagger} \hat{a}^\dagger e^{-\lambda \hat{a}^\dagger} &= \hat{a}^\dagger + \lambda^* [\hat{a}, \hat{a}^\dagger] = \hat{a}^\dagger + \lambda^* \end{aligned}$$

which give

$$\hat{a} e^{\lambda \hat{a}^\dagger} = e^{\lambda \hat{a}^\dagger} (\hat{a} + \lambda) \quad \hat{a}^\dagger e^{-\lambda^* \hat{a}} = e^{-\lambda^* \hat{a}} (\hat{a}^\dagger + \lambda^*).$$

Using this we find

$$\begin{aligned} \text{Tr} \left( \hat{a}^\dagger \hat{a} \rho e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) &= \text{Tr} \left( \hat{a} \rho (\hat{a}^\dagger + \lambda^*) e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) = \text{Tr} \left( \hat{a} \rho \hat{a}^\dagger e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) + \lambda^* \frac{\partial C}{\partial \lambda^*} \\ \text{Tr} \left( \rho \hat{a}^\dagger \hat{a} e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) &= \text{Tr} \left( (\hat{a} + \lambda) \rho \hat{a}^\dagger e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) = \text{Tr} \left( \hat{a} \rho \hat{a}^\dagger e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) + \lambda \frac{\partial C}{\partial \lambda}. \end{aligned}$$

This means that we can rewrite Eq. (14) as

$$\frac{\partial C}{\partial t} = -\left(\frac{\gamma}{2} + i\omega\right)\lambda^* \frac{\partial C}{\partial \lambda^*} - \left(\frac{\gamma}{2} - i\omega\right)\lambda \frac{\partial C}{\partial \lambda}.$$

This equation can be solved by the method of characteristics to get

$$C(\lambda, \lambda^*, t) = C_0\left(\lambda e^{-(\frac{\gamma}{2}-i\omega)t}, \lambda^* e^{-(\frac{\gamma}{2}+i\omega)t}\right)$$

where  $C_0(\lambda, \lambda^*) = C(\lambda, \lambda^*, 0)$ .

We assume that at  $t = 0$  the state is the coherent state  $|z\rangle$  and calculate the function

$$C_0(\lambda, \lambda^*) = \text{Tr}\left(|z\rangle\langle z|e^{\lambda\hat{a}^\dagger}e^{-\lambda^*\hat{a}}\right) = \text{Tr}\left(e^{-\lambda^*\hat{a}}|z\rangle\langle z|e^{\lambda\hat{a}^\dagger}\right) = e^{\lambda z^* - \lambda^* z}.$$

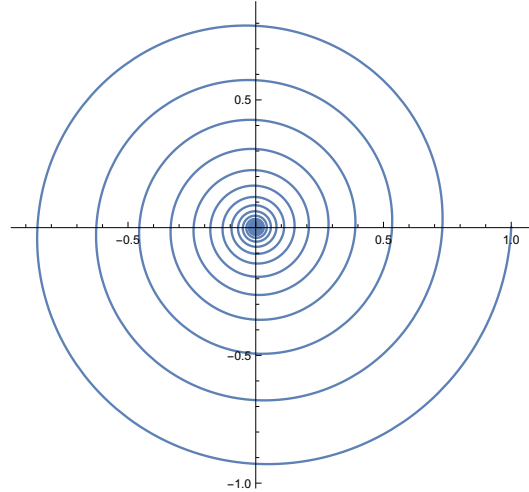
We then get that

$$C(\lambda, \lambda^*, t) = e^{\lambda z^*(t) - \lambda^* z(t)},$$

with

$$z(t) = z e^{-(\frac{\gamma}{2}+i\omega)t}.$$

If the transformation from  $\rho$  to  $C(\lambda, \lambda^*, t)$  is invertible (I believe that it is, but I have never seen this point discussed), we can conclude that the state at time  $t$  is  $|z(t)\rangle$ . That is, it is still a coherent state. Moreover, it is still a pure state, which means that the oscillator is not entangled with the environment despite the interaction that induces damping of the motion. The reason for this is that the coherent states are eigenstates of the Lindblad operator. The plot below shows  $z(t)$  for  $\gamma/\omega = 0.1$ .



The last case we want to analyze is the situation where the initial state is a superposition of two coherent states  $|\psi\rangle = N(|z_1\rangle + |z_2\rangle)$ . In this case, we have

$$\rho(0) = |\psi\rangle\langle\psi| = |N|^2(|z_1\rangle\langle z_1| + |z_1\rangle\langle z_2| + |z_2\rangle\langle z_1| + |z_2\rangle\langle z_2|)$$

which gives

$$C_0(\lambda, \lambda^*) = \text{Tr}\left(\rho(0)e^{\lambda\hat{a}^\dagger}e^{-\lambda^*\hat{a}}\right) = |N|^2\left(e^{\lambda z_1^* - \lambda^* z_1} + \langle z_1|z_2\rangle e^{\lambda z_1^* - \lambda^* z_2} + \langle z_2|z_1\rangle e^{\lambda z_2^* - \lambda^* z_1} + e^{\lambda z_2^* - \lambda^* z_2}\right)$$

where we used that

$$\text{Tr} \left( |z_1\rangle\langle z_2| e^{\lambda \hat{a}^\dagger} e^{-\lambda^* \hat{a}} \right) = \langle z_2|z_1\rangle e^{\lambda z_2^* - \lambda^* z_1}$$

because of the cyclic property of the trace and that the coherent states are eigenstates of  $\hat{a}$ . We then get

$$C(\lambda, \lambda^*, t) = |N^2| \left( e^{\lambda z_1^*(t) - \lambda^* z_1(t)} + \langle z_1|z_2\rangle e^{\lambda z_1^*(t) - \lambda^* z_2(t)} + \langle z_2|z_1\rangle e^{\lambda z_2^*(t) - \lambda^* z_1(t)} + e^{\lambda z_2^*(t) - \lambda^* z_2(t)} \right).$$

If we instead calculate the  $C(\lambda, \lambda^*, t)$  for the state  $|\psi(t)\rangle = N(|z_1(t)\rangle + |z_2(t)\rangle)$ , the crossterms would have the form

$$\langle z_2(t)|z_1(t)\rangle e^{\lambda z_2^*(t) - \lambda^* z_1(t)}$$

so the crossterms are reduced by the factor

$$\eta(t) = \frac{\langle z_2|z_1\rangle}{\langle z_2(t)|z_1(t)\rangle}$$

compared to the state  $|\psi(t)\rangle$ . Using that

$$|\langle z_2|z_1\rangle| = e^{-\frac{1}{2}|z_2 - z_1|^2}$$

we get

$$|\eta(t)| = \frac{e^{-\frac{1}{2}|z_2 - z_1|^2}}{e^{-\frac{1}{2}|z_2(t) - z_1(t)|^2}} = \frac{e^{-\frac{1}{2}|z_2 - z_1|^2}}{e^{-\frac{1}{2}e^{-\gamma t}|z_2 - z_1|^2}} = e^{-\frac{1}{2}|z_2 - z_1|^2(1 - e^{-\gamma t})}.$$

We consider times short, so that  $\gamma t \ll 1$  which means that  $1 - e^{-\gamma t} \approx \gamma t$  and get

$$|\eta(t)| \approx e^{-\frac{1}{2}\gamma|z_2 - z_1|^2 t} = e^{-\Gamma t}.$$

where we have defined the decay rate of the crossterms as

$$\Gamma = \frac{1}{2}\gamma|z_2 - z_1|^2.$$

When  $|z_2 - z_1| \gg 1$  we have that  $\Gamma \gg \gamma$  which means that in a short time  $\tau = 1/\Gamma$ , the state will decohere to a mixture of the two states  $|z_1(t)\rangle$  and  $|z_2(t)\rangle$ , while the amplitude of these two states has not decayed significantly. A particularly simple example of this phenomenon is seen if we consider  $z_1 = 0$  and  $z_2 = z$ . Then

$$\tau = \frac{2}{\gamma|z|^2}$$

Recalling that  $|z|^2 = n$  is the average excitation level of the oscillator, and that the rate of phonon emission is  $n\gamma$ , we conclude that  $\tau$  is of the order of the time to emit one single photon. That is, the emission of a single photon is sufficient to change the state from a superposition of two coherent states to a mixture of two coherent states. We can understand this from the point of view of our information about the state of the system. In the initial state  $\psi = N(|0\rangle|z\rangle)$  the system is in a superposition of the ground state and some coherent state. At the moment a photon is emitted, we could in principle detect this photon. The fact that the ground state can not emit any photon would mean that n detecting

an emitted photon, we collapse the wavefunction and conclude that the state was  $|z\rangle$ . The Lindblad equation is derived assuming that no measurements of the state of the environment is made, so it means that the information about whether a photon has been emitted or not does not reach us. Instead of collapsing the wavefunction in accordance with the measurement outcome, we have the situation that we know the collapse has happened (because sufficient time passed), but not the state to which the wavefunction collapsed. This means that we have classical uncertainty about the quantum state, so the state is a mixture of the two states.

#### 6.2.4 Two level system with external driving

Let the Hamiltonian be

$$H = \frac{1}{2}\omega_0\sigma_z + \frac{1}{2}\omega_1(\cos\omega t\sigma_x + \sin\omega t\sigma_y)$$

We transform to a rotating frame by the transformation

$$T = e^{-\frac{i}{2}\omega t\sigma_z}$$

The density matrix in the rotating frame is  $\rho' = T^\dagger\rho T$  and the Hamiltonian is

$$H' = T^\dagger H T + i\dot{T}^\dagger T = \frac{1}{2}\Delta\sigma_z + \frac{1}{2}\omega_1\sigma_x$$

with  $\Delta = \omega_0 - \omega$ . With the jump operator  $\alpha = |0\rangle\langle 1|$ , the Lindblad equation is

$$\dot{\rho}' = i[\rho', H'] - \frac{\gamma}{2} \left( \alpha^\dagger \alpha \rho' + \rho' \alpha^\dagger \alpha - 2\alpha \rho' \alpha^\dagger \right) \quad (15)$$

The stationary state is found by setting  $\dot{\rho}' = 0$  giving

$$i[\rho', H'] = \frac{\gamma}{2} \left( \alpha^\dagger \alpha \rho' + \rho' \alpha^\dagger \alpha - 2\alpha \rho' \alpha^\dagger \right)$$

We parametrize the density matrix using the Bloch vector, so that  $\rho' = \frac{1}{2}(1 + m_i\sigma_i)$ , and inserting this into Eq (15) we get

$$\begin{aligned} \Delta m_y &= -\frac{\gamma}{2} m_x \\ \Delta m_x - \omega_1 m_z &= \frac{\gamma}{2} m_y \\ \omega_1 m_y &= \gamma m_z + \gamma \end{aligned}$$

with solution

$$\begin{aligned} m_x &= -\frac{2\Delta\omega_1}{\omega_1^2 + 2\Delta^2 + \frac{1}{2}\gamma^2} \\ m_y &= \frac{\gamma\omega_1}{\omega_1^2 + 2\Delta^2 + \frac{1}{2}\gamma^2} \\ m_z &= -\frac{2\Delta^2 + \frac{1}{2}\gamma^2}{\omega_1^2 + 2\Delta^2 + \frac{1}{2}\gamma^2} \end{aligned}$$



To find the power absorbed from the driving field we consider the average energy of the system

$$E = \text{Tr } \rho H.$$

The time derivative of this is

$$\dot{E} = \text{Tr } \dot{\rho} H + \text{Tr } \rho \dot{H},$$

and using the Lindblad equation we get

$$\dot{E} = i \text{Tr}([\rho, H]H) - \frac{\gamma}{2} \text{Tr}(\alpha^\dagger \alpha \rho' + \rho' \alpha^\dagger \alpha - 2\alpha \rho' \alpha^\dagger) + \text{Tr } \rho \dot{H}.$$

The first term is always zero:

$$\text{Tr}([\rho, H]H) = \text{Tr}(\rho H^2 - H \rho H) = 0.$$

Therefore there are only two terms that contribute to the energy change

$$\dot{E} = -\frac{\gamma}{2} \text{Tr}(\alpha^\dagger \alpha \rho' + \rho' \alpha^\dagger \alpha - 2\alpha \rho' \alpha^\dagger) + \text{Tr } \rho \dot{H}.$$

The first term is proportional to  $\gamma$  and gives the energy loss due to emitted photons. The second term involves the time dependence of the Hamiltonian,  $\dot{H}$  and gives the energy gain due to the driving field. In the stationary state both terms should be equal. We want to express the last term in the rotating frame, and use

$$H = T H' T^\dagger - i T \dot{T}^\dagger.$$

We know that both the last term and  $H'$  are time-independent, and therefore we get

$$\dot{H} = \dot{T} H' T^\dagger + T H' \dot{T}^\dagger = \frac{i}{2} \omega T [H', \sigma_z] T^\dagger = \frac{\omega \omega_1}{2} T \sigma_y T^\dagger$$

Then we get

$$\dot{E} = \text{Tr } \rho \dot{H} = \frac{1}{2} \omega \omega_1 \text{Tr}(\rho' \sigma_y) = \frac{1}{2} \omega \omega_1 m_y.$$

Inserting the result for  $m_y$  in the stationary state we get the absorbed energy

$$\dot{E}_+ = \frac{1}{2} \frac{\gamma \omega \omega_1^2}{\omega_1^2 + 2\Delta^2 + \frac{1}{2}\gamma^2}.$$