Quantum Computer Science

— course lecture notes HT 2008 (with updates 2010)*—

Bob Coecke
Oxford University Computing Laboratory

March 13, 2010

Contents

*Please report any error, typo or omission found in these notes both to the lectures and the class problems tutor, Jacob Biamonte, via email: Bob.Coecke@comlab.ox.ac.uk and Jacob.Biamonte@comlab.ox.ac.uk.

1 Historical and physical context

1.1 The birth of quantum mechanics

There is no agreed clear date attached to the 'birth of quantum theory', as is for example the case of Sir Isaac Newton's 1686 theory of (classical) mechanics and Albert Einstein's 1905/1917 theories of special/general relativity. Quantum Theory came about rather by several discoveries, insights and developments which ultimately lead to John von Neumann's Hilbert space quantum mechanical formalism which is currently still in use. Some of the most important of these discoveries, insights and developments are (e.g. [1]):

- In 1900 Max Planck noted that the physically observed

frequency dependence of so-called *black body radiation* requires energy to be quantized and come in the form of finite chunks.

- In 1905 Albert Einstein explained the photoelectric effect by postulating that light is also quantized and comes in packets which he called photons.
- In 1913 Niels Bohr explained the spectral lines of the hydrogen atom emission spetra by a new model for the atom which, a priori, involved discrete (i.e. quantized) energy levels.
- In 1924 Louis de Broglie suggested that, dually to the discrete 'particle' nature of light, matter should also be thought of as having 'wave'-like behavior.
- Around 1925 Werner Heisenberg constructed matrix mechanics and Erwin Schrödinger constructed wave mechanics including the Schrödinger equation.
- John von Neumann developed the mathematically rigorous Hilbert space formalism for quantum mechanics which is now still in use — first published in 1932 [2]. Also, Paul Dirac's bra-ket notation, which appeared in his 1930 book, remains in popular use today [5].

Before we turn to a formal development of quantum mechanics in the next chapter, we now informally discuss some crucial structural features of quantum mechanics. Once we have a mathematical formalism at our disposal, it will become much clearer how they inter-relate.

Superposition. When a system (e.g. a particle) admits some distinct properties e.g. 'being either here or there', 'being either 1, 2, 3, ..., or 111 years old', 'being either dead or alive', 'being either 0 or 1', (etc.) then a *superposition state* stands for a situation — where a kind of combination of these alternatives applies with is different than a probability distribution over the states. In the case of the so-called *Schrödinger's cat paradox* the cat is neither dead or alive, but somewhere in between, and in computer science terms a quantum bit (qubit) can take both the values 0 and 1 concurrently. In a sense, while the utterance 'quantum' indicates 'a passage from the continuous to the discrete', from an informatic perspective it is rather 'a passage from the discrete to the continuous'. But unfortunately, that continuous space is not evidently accessible, due to the nature of quantum measurement.

Uncertainty. A physical quantum system cannot admit both a sharp (= not in superposition) position and sharp momentum at the same time, a principle known as *Heisenberg's uncertainty principle*. An analogous principle arises often in quantum mechanics and applies to any pair of *non-compatible quantities*.

(non-local) Entanglement. There are states, e.g. *EPR-states and Bell-states*, which yield statistical correlations between systems separated by a large physical distance, and these correlations 'must travel faster than the speed of light', in fact, they are instantaneous. Surprisingly however, these correlations cannot be used to send information faster than light, and hence special relativity is not violated. These correlations have been experimentally observed many times during the past 30 years, and are typically referred to as *quantum entanglement*.

Intrinsic probabilities. When we perform a quantum measurement, i.e. verify some physical property of a superposition state, then the outcome will occur in a probabilistic manner. E.g. if we verify whether the cat is either dead or alive, or, whether the bit is either 0 or 1, then if we are in a superposition state — both outcomes can occur with some probability. There are mathematical theorems which state that by assigning additional statistical variables to the quantum system we cannot get rid of these probabilities [6, 7].

1.2 The status of quantum mechanics

1. Probably the most successful physical theory ever in terms of predictions e.g. *quantum electrodynamics* predicts correct results up to ± 10 digits!
2. It has many important applications such as:
   - The description of individual particles such as molecules, atoms, photons, electrons, protons and neutrons, and hence all the obvious applications in many fields ranging from chemistry, nuclear physics, and in the future possibly high-energy physics.
   - New technologies such as the laser and the electron microscope, and in particular, the *transistor* as a replacement for valves, enabling the scale at which micro-electronics (including computer hardware) can currently be built.
   - Important medical tools such as magnetic resonance imaging techniques.
   - Actual quantum informatic devices such as quantum communication systems and quantum processors?
3. Big conceptual and philosophical questions, initially raised by Einstein, remained unanswered. The biggest

of these is the so-called *measurement problem*: it is conceptually not clear at all what causes 'a measurement to take place'. But most of the physics community has moved forward — ignoring most conceptual problems and accepting quantum mechanics as a cook-book which provides 'weird' recipes on how to handle and interpret matter, and more recently information.

4. The formalism is still mathematically unsatisfactory for many reasons: it contains redundancies such as global phases, yields re-normalization problems in quantum field theory, lacks high-levelness etc. The formalism in fact hasn't changed since it's creation by John von Neumann, who actually denounced it three years after creating it! This lead to so-called quantum logic — a field of study launched by Birkhoff and von Neumann in 1936 [4], but there are serious doubts that this has given much insights either in quantum theory or in logic, and in no way did it have the capabilities to replace von Neumann's quantum mechanical formalism.

1.3 The birth of quantum informatics

Most of the scientific activity on 'pure quantum mechanics' which took place in the second half of the 20th century was either on its experimental confirmation, on its philosophical justification, or on generalizing/modifying its formalism. The passage to quantum informatics can be seen as a matter of change of attitude towards the so-called 'quantum-weirdness':

It's a feature, not a bug!

The first to mention quantum computing was Paul Benioff in 1980 who studied how particular kinds of quantum evolutions could simulate classical Turing machines. Richard Feynman on the other hand asked the dual question i.e. whether classical computers can simulate quantum evolution, and conjectured that such a simulation came with an exponential slow-down, while, in principle, quantum systems could simulate themselves without this exponential slow-down by simply relying on the natural quantum evolution they are already governed by. Hence, the first advantage of considering a quantum mechanical system as a computational device had been exposed. The key to this speed-up is that quantum evolution *physically computes* a function for several inputs at the same time, which are in *superposition*. But the nature of the quantum measurement process doesn't allow the state of the quantum system to be read without actually altering it, and converting this potential of *intrinsic parallelism* within quantum evolution into concrete examples of algorithmic speed-up of quantum computers as compared to classical computers turned out to be a highly non-trivial matter.

The first algorithm of that kind was the Deutch-Jozsa algorithm which exploits quantum parallelism—computing a function for several values at once, but uses this to solve a problem of little practical interest. What is often considered to be the start of quantum algorithmics was Peter Shor's 1994 factoring algorithm [13] which provided exponential speed-up as compared to all known classical factoring methods. Another well known quantum algorithm is Lov Grover's 1995 search algorithm [14] for unstructured data of size N which reduces the search-time from N to $\mathcal{O}(\sqrt{N})$.

But quantum informatics is not only about algorithmics. There are several intriguing quantum protocols which expose fascinating physical phenomena, some of which turn out to have applications which are most likely to be the first real-life incarnations of a quantum informatic revolution. Among these conceptually intriguing protocols are quantum teleportation and entanglement swapping. At the practical side there are the many variants of quantum key-distribution, within the field of protocol security. This is a nice example of how quantum informatics constitutes both a danger and provides the corresponding solution to the security of communication protocols: An actual quantum computer runing Shor's algorithm would provide a danger to many cryptographic protocols currently in use which typically rely on hardness of factoring. On the other hand, quantum key-distribution provides a solution to any such attack!

1.4 The status of quantum informatics

1. Many different experimental devices of a small number of qubits (< 10) are operational but *scalability* is still a major problem. This problem is due mainly to the *decoherence* of quantum data due to interaction with the environment, but (at least) theoretical solutions do exist.

2. The search for new kinds of algorithms and applications continues including resent efforts to use quantum processors to simulate chemical reactions at the quantum mechanical level!

3. It is commonly accepted that information security will likely be the first practicable application of quantum informatics, and quantum communication devices are available from commercial companies (MagiQ and ID Quantique). An actual quantum key distribution protocol has taken place between a Swiss bank and Geneva City Hall [21]. However, while the quantum component of the experiment worked perfectly, the authentication protocol failed to be secure due to flaws in the analysis of its classical component (e.g. [22]). The irony here is that a true danger to classical security protocols is posed by Shor's algorithm, i.e. by quantum informatics, while

it is again quantum informatics which provides the solution.

4. At a very fundamental level questions still remain e.g.:

   - What are the true origins of a quantum algorithmic computational speed-up?
   - What are the limits of quantum computation?
   - What is a model for general quantum computing?

   New computational models are being proposed e.g. the one-way model [19] which radically challenges all known methods related to the circuit or gate array model.

5. There is no real high-level quantum computer science. The current methods, from a programming perspective, are comparable with hacking with bits (but instead with complex numbers rather than with bits). There is also a convincing argument that current high-level computer science structures for distributed computing, hybrid computing, embedded computing and tools for verification might actually be extremely useful for the theoretical side of the quantum computational endeavor, both at the level of the quantum mechanical formalism and those posed by the new quantum computational paradigms. It seems that there is a true need for quantum computer science in the British/European sense.

6. The quantum computational activity has already provided some fresh insights into the domain of foundations of physics, providing new concepts and paradigms from informatics.

2 Qubits vs. bits

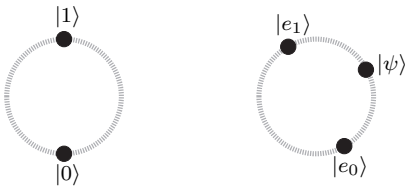
2.1 Acting on qubits

A **bit**:

- admits two distinct values 0 and 1,
- admits arbitrary transformations (can erase, copy etc. at ease).
- is freely readable (hence, the state of the bit is left unchanged if you measure it),

A **qubit**:

- a *sphere* of values (which in some particular manner is 'spanned' by two quantum states $|0\rangle$ and $|1\rangle$ or $|e_0\rangle$ and $|e_1\rangle$),

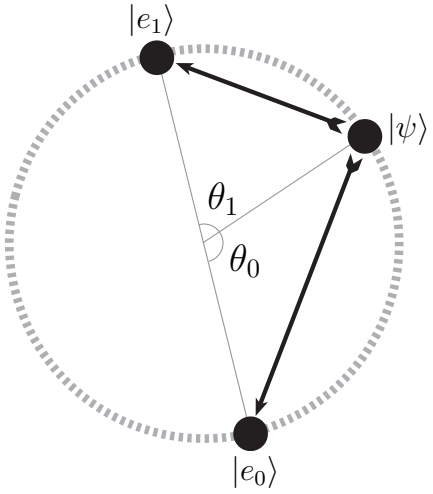


- only admits special transformations which preserve the angles (and hence opposites) on the sphere, and hence which are in particular *reversible*.
- only admits 'reading' through so-called *quantum measurements* $M(|e_0\rangle, |e_1\rangle)$ which
  - only have two possible outcomes $|e_0\rangle$ and $|e_1\rangle$,
  - change the initial state $|\psi\rangle$ to either $|e_0\rangle$ or $|e_1\rangle$,

  hence one could say that a measurement $M(|e_0\rangle, |e_1\rangle)$ does not tell us $|\psi\rangle = \alpha|e_0\rangle + \beta|e_1\rangle$ but destroys $|\psi\rangle$!

A metaphor: quantum measurement of Colors. Assume the points of the sphere, i.e. the possible states of the system, correspond to colors. We can for example ask if it is blue or red (= 2 colors), but not if it is either blue, red or green (= 3 colors). Assume now the system is green, and the measurement we perform asks if it is either blue or red. Then the outcome will either be blue or red, meaning that the system has indeed become blue or red respectively, but we will never get to know that the initial color was actually green.

Here's a more detailed look at quantum measurements:



The two transitions

$$\mathrm{P}_{e_0} :: |\psi\rangle \mapsto |e_0\rangle \qquad\qquad \mathrm{P}_{e_1} :: |\psi\rangle \mapsto |e_1\rangle$$

have respective chance $\mathsf{prob}(\theta_0)$ and $\mathsf{prob}(\theta_1)$ with

$$\mathsf{prob}(\theta_0) + \mathsf{prob}(\theta_1) = 1$$

since quantum theory dictates that for θ the angle on the sphere between the *initial state* and a possible *outcome state* (cf. the above picture) we have

$$\mathsf{prob}(\theta) = cos^2\frac{\theta}{2}\,,$$

and in particular do we have

$$\mathsf{prob}(0) = 1 \qquad \mathsf{prob}(90^o) = \frac{1}{2} \qquad \mathsf{prob}(180^o) = 0$$

and in general

$$0 < \mathsf{prob}(\theta) < 1 \qquad \text{for} \qquad 0 < \theta < 180^o\,.$$

Since there are impossible transitions (cf. $\mathsf{prob}(180^o) = 0$), we obtain two 'partial constant maps' on the sphere Q

$$\mathrm{P}_{e_0} : Q \setminus \{|\, e_1 \,\rangle\} \to Q :: |\psi\rangle \mapsto |\, e_0 \,\rangle.$$

$$\mathrm{P}_{e_1} : Q \setminus \{|\, e_0 \,\rangle\} \to Q :: |\psi\rangle \mapsto |\, e_1 \,\rangle$$

capturing the *dynamics of measurement*, which can be used as a *dynamic resource* when designing algorithms and protocols — as we shall see further, for so-called *degenerate measurements* these maps are not always constant. In fact, restricting to states and measurements which are such that measurements behave deterministically and don't change the state is equivalent to doing classical reversible computing! Hence:

- **bad:** quantum measurements destroy most data
- **good:** quantum measurements expose some data
- **good:** quantum measurements act on data

Conclusively, designing quantum algorithms and protocols boils down to exploiting the enlarged state space by acting on quantum data either with:

- a particular kind of reversible operations — which for example do not admit cloning as well as deleting, or,
- irreversible measurements, for which we have to perform acrobatics between 'the good' and 'the bad'.

2.2 Describing a qubit with complex numbers

Let $\mathbb{R}$ denote the *real numbers* and $\mathbb{C}$ denote the *complex numbers* i.e. numbers $z = x + i\,y$ where $x, y \in \mathbb{R}$ and i is implicitly defined within $i \cdot i = -1$ so i can be thought of as $\sqrt{-1}$. Hence for addition and multiplication of complex numbers $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2$ we have

$$z_1 + z_2 = (x_1 + y_1) + i\,(y_1 + y_2)$$

$$z_1 \cdot z_2 = (x_1 x_2 - y_1 y_2) + i\,(x_2 y_1 + y_2 x_1)\,.$$

The complex conjugate of $z = x + i\,y$ is $\bar{z} = x - i\,y$ hence

$$\bar{z} + z = 2x \qquad \text{and} \qquad \bar{z} \cdot z = x^2 + y^2\,.$$

The **state of a qubit** is described by a pair of complex numbers $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = z_1|0\rangle + z_2|1\rangle$ *up to a non-zero complex multiple*, which means that for any $z \in \mathbb{C}_0$ (= $\mathbb{C}$ without zero $\mathbb{C}/0$) the pairs

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad \text{and} \qquad z \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \begin{pmatrix} z \cdot z_1 \\ z \cdot z_2 \end{pmatrix}$$

both define the same state. Typically one writes these pairs as

$$|\psi\rangle := z_1 \cdot |0\rangle + z2 \cdot |1\rangle$$

to make a connection with bits. Ignoring the global redundancy of the non-zero complex number z, a qubit state is a *complex linear combination* of two reference states $|0\rangle$ and $|1\rangle$.

When representing the complex numbers in the 2D *complex plane*, passing from *cartesian* to *polar coordinates* yields the *amplitude* and *phase* of a complex number, respectively

$$r = \sqrt{x^2 + y^2} \qquad , \qquad \tan(\theta) = \left(\frac{y}{x}\right),$$

and conversely, the *real* and *complex parts* re-emerge as

$$x = r \cdot \cos(\theta) \qquad , \qquad y = r \cdot \sin(\theta)\,.$$

Hence a complex number can also be written as

$$z = r \cdot e^{i\theta} \quad \text{since} \quad e^{i\theta} := \cos(\theta) + i\,\sin(\theta)\,.$$

Hence, when representing a qubit by a pair of complex numbers (z_1, z_2) there is both a redundant *global phase* and *global amplitude*. Concerning the redundant global amplitude, one usually only considers *normalized* vectors i.e.

$$(\bar{z}_1\ \bar{z}_2) \circ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \bar{z}_1 \cdot z_1 + \bar{z}_2 \cdot z_2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 = 1\,,$$

since those are the ones which occur in the expressions for calculating the probabilities. Note here in particular that a pair of complex numbers has four 'real degrees of freedom', and hence that a pair of complex numbers up to a non-zero complex multiple has two 'real degrees of freedom', what indeed corresponds with points on a sphere. More generally, 'n-tuples of complex numbers up to a non-zero complex multiple' have $2n - 2$ 'real degrees of freedom'.

Exercise 2.1 If we take 'all pairs of real numbers up to a non-zero real multiple' to be the states of some system, which geometric object do we obtain as the state space? How would you define opposite states? Representing real number pairs in the XY-plane, when do two such pairs yield opposite states?

Special examples of states are $|0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ which constitute the so-called *computational basis* corresponding to the classical bit values 0 and 1. The states of the computational basis are indeed opposite states, which in terms of pairs of complex numbers $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ and $\begin{pmatrix} z_1' \\ z_2' \end{pmatrix}$ requires

$$\bar{z}_1 \cdot z_1' + \bar{z}_2 \cdot z_2' = 0\,,$$

or equivalently, in terms of an inner or *matrix product*,

$$(\bar{z}_1\ \bar{z}_2) \circ \begin{pmatrix} z_1' \\ z_2' \end{pmatrix} = 0\,.$$

In practice however, calculations can be performed within standard linear algebra, ignoring these redundancies. For example, quantum *logic gates* are 2×2-matrices of complex numbers

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

and induce a change of the state

$$|\psi\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto U(|\psi\rangle) = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} U_{11} \cdot z_1 + U_{12} \cdot z_2 \\ U_{21} \cdot z_1 + U_{22} \cdot z_2 \end{pmatrix},$$

which both preserves normalization and opposites i.e. the 'canonical opposites' $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ should stay opposite and preserve their global amplitude, resulting in both $\begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix}$ and $\begin{pmatrix} U_{12} \\ U_{22} \end{pmatrix}$ being normalized, i.e.

$$(\bar{U}_{11}\ \bar{U}_{21}) \circ \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} = \bar{U}_{11} \cdot U_{11} + \bar{U}_{21} \cdot U_{21} = 1$$

and

$$(\bar{U}_{12}\ \bar{U}_{22}) \circ \begin{pmatrix} U_{12} \\ U_{22} \end{pmatrix} = \bar{U}_{12} \cdot U_{12} + \bar{U}_{22} \cdot U_{22} = 1\,,$$

and

$$(\bar{U}_{11}\ \bar{U}_{21}) \circ \begin{pmatrix} U_{12} \\ U_{22} \end{pmatrix} = \bar{U}_{11} \cdot U_{12} + \bar{U}_{21} \cdot U_{22} = 0\,.$$

In other words, measurements are representable by particular families of *projectors*. The measurement with respect to the computational basis is described by the following pair of projectors

$$\mathrm{P}_0 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathrm{P}_1 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

which induce a change of state

$$|\psi\rangle \mapsto \mathrm{P}_0(|\psi\rangle) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ 0 \end{pmatrix} \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$|\psi\rangle \mapsto \mathrm{P}_1(|\psi\rangle) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ z_2 \end{pmatrix} \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

i.e. the possible outcome states indeed constitute the computational basis. All the other measurements on a qubit can be obtained by *rotations of the sphere* using the same transformations which characterize the logic gates, resulting in

$$U \circ \mathrm{P}_\beta \circ U^{-1}$$

i.e. using U^{-1} we first rotate 'backwards' to the computational basis, then preform the measurement in the computational basis, and then using U to rotate forward again. From the above requirements on U we obtain for

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \quad \text{and} \quad U^{-1} := \begin{pmatrix} \bar{U}_{11} & \bar{U}_{21} \\ \bar{U}_{12} & \bar{U}_{22} \end{pmatrix},$$

that they are indeed inverses i.e. both

$$\begin{pmatrix} \bar{U}_{11} & \bar{U}_{21} \\ \bar{U}_{12} & \bar{U}_{22} \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \text{ and } \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \bar{U}_{11} & \bar{U}_{21} \\ \bar{U}_{12} & \bar{U}_{22} \end{pmatrix}$$

yield the identity

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Exercise 2.2 In the lectures it was explained that we can represent *pairs of complex numbers up to a complex number* on a sphere, with $|0\rangle := (1, 0)$ as the sphere top and $|1\rangle := (0, 1)$ as the sphere bottom, and with the states $\frac{1}{\sqrt{2}}(1, e^{i\theta})$ on the 'equator', where we singled out $|+\rangle := \frac{1}{\sqrt{2}}(1, 1)$, $|-\rangle := \frac{1}{\sqrt{2}}(1, -1)$, $|y_+\rangle := \frac{1}{\sqrt{2}}(1, i)$ and $|y_-\rangle := \frac{1}{\sqrt{2}}(1, -i)$. What is the action of the following operations on these 6 special points (depict on the sphere):

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad S := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad T := \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$$

respectively called the Hadamard-, phase-, and $\pi/8$-gate (also called the T-gate), and

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

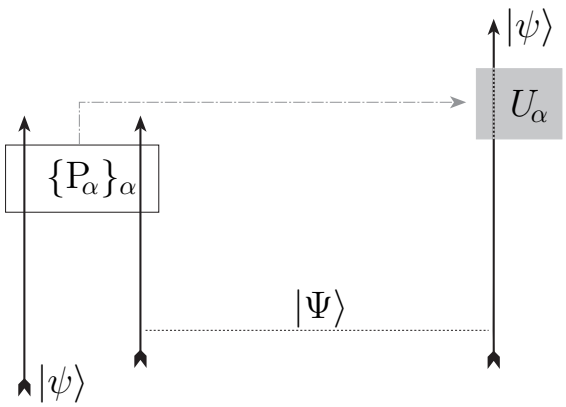
$$Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

typically called the Pauli X, Y and Z matrices. In particular, which of the 6 states mentioned above are either invariant or

permuted by these gates. Can you discover any special relations between these operations? (e.g. do some commute, are idempotent ($U \circ U = U$), involutive ($U \circ U = 1$), or do we have a relation like $U_1 \circ U_2 = U_3$ for certain triples? (Note: please spend as much time as needed on this important question.)

2.3 Describing two qubits

Since, for the case of qubits, the fact that $2+2 = 2\times 2$ might cause some confusion in the argument we wish to make, we will consider 'n-tuples of complex numbers up to a non-zero complex multiple', called 'qu*d*its' (d is for digit). Consider the following situation. We start with 3 qu*d*its, the first one being in an arbitrary state $|\psi\rangle$ and the other two being in a particular *joint state* $|\Psi\rangle$. Then we apply a particular *joint measurement* to the first two qu*d*its. Using quantum theory, it can then be shown that after performing a particular logic gate on the third qu*n*it it will be exactly in the same state $|\psi\rangle$ as the first qu*n*it initially was. This involves sending the *measurement outcome* which 'witnesses which projector P_α actually took place' to the third qu*n*it, such that the appropriate logic gate U_α can be applied.



All together something very weird has happened here:

- We were able to *teleport* the quantum state of the first qu*n*it to the third qu*n*it, that is, sending quantum data, but we only communicated finitary classical data, namely the measurement outcome! So what causes this magic?

The magic is hidden in the particular nature of the particular initial joint state describing the second and the third qu*n*it. Indeed, quantum theory tells us that a pair of qu*d*its is not described by assigning a state to each of them, but by assigning a $n\times n$-matrix (up to a complex multiple) to the pair of them, and hence, rather than $(n-1)+(n-1)$ complex degrees of freedom we obtain $(n\times n)-1$ complex degrees of freedom, that is, for n large enough, approximately

$$2n \mapsto n^2\,,$$

and it is the resulting additional degrees of freedom which enable communication. Actually, as we will see it what will come, it is not completely wrong to think of the $n\times n$-matrix representing a communication channel through which information can flow and be processed.

3 von Neumann's pure state formalism

We will only consider *finite-dimensional Hilbert spaces*, that is, in physical terms, all measurements have a finite number of outcomes. While for informatic purposes this suffices, infinite spectra do play an important role in general quantum mechanics e.g. position and momentum observables.

3.1 Hilbert space

Definition 3.1 A (finite-dimensional) *Hilbert space* is a vector space $\mathcal{H}$ over the complex number field $\mathbb{C}$ which also comes with an *inner-product,* i.e. a map

$$\langle - \mid - \rangle : \mathcal{H}\times\mathcal{H} \to \mathbb{C}\,,$$

satisfying

$$\langle\psi|c_1\cdot\psi_1 + c_2\cdot\psi_2\rangle = c_1\langle\psi|\psi_1\rangle + c_2\langle\psi|\psi_2\rangle$$

$$\langle c_1\cdot\psi_1 + c_2\cdot\psi_2|\psi\rangle = \bar{c}_1\langle\psi_1|\psi\rangle + \bar{c}_2\langle\psi_2|\psi\rangle$$

$$\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle} \qquad \langle\psi|\psi\rangle \in \mathbb{R}^+ \quad \langle\psi|\psi\rangle = 0 \Leftrightarrow \psi = \mathbf{0}$$

for all $c_1, c_2 \in \mathbb{C}$ and all $\phi, \psi, \psi_1, \psi_2 \in \mathcal{H}$.

A *linear operator* between Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is a map $f : \mathcal{H}_1 \to \mathcal{H}_2$ which satisfies

$$f(c_1\cdot\psi_1 + c_2\cdot\psi_2) = c_1\cdot f(\psi_1) + c_2\cdot f(\psi_2)\,,$$

for all $c_1, c_2 \in \mathbb{C}$ and all $\psi_1, \psi_2 \in \mathcal{H}_1$, hence the inner-product is linear in the second variable while being *anti-linear* in the first variable i.e. a so-called *sesquilinear* form. Two vectors $\psi, \phi \in \mathcal{H}$ are called *orthogonal* iff

$$\langle\psi\,|\,\phi\rangle = 0$$

and a vector $\psi \in \mathcal{H}$ is *normalized* iff

$$|\psi|^2 := \langle\psi\,|\,\psi\rangle = 1\,.$$

Exercise 3.2 Prove that $\mathbb{C}$ is itself a Hilbert space over $\mathbb{C}$ i.e. show that there exists an inner-product on $\mathbb{C}$.

If $f : \mathcal{H}_1 \to \mathcal{H}_2$ is a linear map then it always has a unique *adjoint* $f^\dagger : \mathcal{H}_2 \to \mathcal{H}_1$ which is implicitly defined within

$$\langle f^\dagger(\phi)|\psi\rangle = \langle \phi|f(\psi)\rangle$$

for all $\psi \in \mathcal{H}_1$ and all $\phi \in \mathcal{H}_2$. In Exercise 3.11.i we will construct this adjoint, hence or otherwise prove its existence, and uniqueness also follows in a straightforward manor.

Exercise 3.3 Show that $(g \circ f)^\dagger = f^\dagger \circ g^\dagger$.

A linear operator U is *unitary* if its inverse U^{-1} exists and, equivalently,

- $U^{-1} = U^\dagger$,
- U (and also $U^\dagger$) preserves the inner-product.

Exercise 3.4 Show that the two definitions of unitarity given above are indeed equivalent.

A subset of vectors $\mathcal{A} \subseteq \mathcal{H}$ is called a *subspace* of a vector space $\mathcal{H}$ if it is closed under linear combinations of the vectors it contains i.e.

$$\psi_1, \psi_2 \in \mathcal{A} \Rightarrow c_1 \cdot \psi_1 + c_2 \cdot \psi_2 \in \mathcal{A}\,.$$

Special types of subspaces are those formed by *rays* . Rays are the subspaces *spanned* (or generated) by a single vector i.e.

$$\mathsf{span}(\psi) = \{c \cdot \psi \mid c \in \mathbb{C}\}\,.$$

We are now ready to state a first postulate of von Neumann's formulation of quantum theory.

Postulate 3.5 [states and transformations] The state of a quantum system is described by a ray in a Hilbert space. Deterministic transformations of quantum systems are described by unitary operators acting on that Hilbert space.

Hence, from a computational perspective, the deterministic *logic gates* which we can apply to quantum data are exactly the unitary transformations. Besides unitary transformations, other linear *endo*-operators which play a special role in quantum theory are *self-adjoint operators* i.e.

$$\langle H(\phi)|\psi\rangle = \langle \phi|H(\psi)\rangle$$

that is for all $\psi \in \mathcal{H}_1$ and all $\phi \in \mathcal{H}_2$, $H^\dagger = H$. Self-adjoint endo-operators $\mathrm{P} : \mathcal{H} \to \mathcal{H}$ which are also *idempotent*, i.e. $\mathrm{P} \circ \mathrm{P} = \mathrm{P}$, are called *projectors*.

Exercise 3.6 i. If U is unitary and H self-adjoint show that $U^{-1} \circ H \circ U$ is also self-adjoint. **ii.** If U is unitary and P is a projector show that $U^{-1} \circ \mathrm{P} \circ U$ is a projector.

Special examples of projectors on $\mathcal{H}$ are the identity

$$1_{\mathcal{H}} : \mathcal{H} \to \mathcal{H} :: \psi \mapsto \psi$$

and the *zero-operator*

$$O_{\mathcal{H}} : \mathcal{H} \to \mathcal{H} :: \psi \mapsto \mathbf{0}\,.$$

Proposition 3.7 *Each self-adjoint operator $H : \mathcal{H} \to \mathcal{H}$ admits a so-called 'spectral decomposition'*

$$H = \sum_i a_i \cdot \mathrm{P}_i$$

where all $a_i \in \mathbb{R}$ and all $\mathrm{P}_i : \mathcal{H} \to \mathcal{H}$ are projectors which are 'mutually orthogonal' i.e. $\mathrm{P}_i \circ \mathrm{P}_j = O_{\mathcal{H}}, \forall i \neq j$.

The proof of this proposition can be performed relying on the matrix calculus and the fact that each self-adjoint operator admits a diagonal form (see Exercise 3.14 below).

Postulate 3.8 [measurements] A measurement on a quantum system is described by a self-adjoint operator. The set $\{a_i\}$ in the operator's spectral decomposition are the *measurement outcomes* while the set of projectors $\{\mathrm{P}_i\}$ describes the *change of the state* that takes place during a measurement. In particular, when a measurement takes place:

1. The initial state ψ undergoes one of the transitions

   $$\mathrm{P}_i :: \psi \mapsto \mathrm{P}_i(\psi)$$

   and the probability of the possible transitions is

   $$\mathsf{prob}(\mathrm{P}_i, \psi) = \langle \psi|\mathrm{P}_i(\psi)\rangle$$

   where ψ needs to be normalized.

2. The *observer* which performs the measurement receives the value a_i as a token-witness of that fact.

It should be clear that from a structural perspective the actual values of the measurement outcomes $\{a_i\}$ are of no significance, and in a sense the respective measurements represented by the self-adjoint operators

$$\sum_i a_i \cdot \mathrm{P}_i \qquad \text{and} \qquad \sum_i i \cdot \mathrm{P}_i$$

can be considered as equivalent, and in particular, the latter is completely determined by the set $\{\mathrm{P}_i\}_i$. On the other hand however, the measurement outcomes are typically physical quantities such as position, momentum and energy, which of course play an important quantitive role in physical theories.

Exercise 3.9 Show that, equivalently, we could have set

$$\mathsf{prob}(\mathrm{P}_i, \psi) = |\mathrm{P}_i(\psi)|^2$$

for the probability of each possible transition.

We have thus far only considered the description of individual quantum systems. A postulate on *compound systems* is still missing, but for this we need to introduce the *tensor product* of Hilbert spaces — we postpone this discussion until Subsection 3.3.

3.2 Matrices

A *basis* for a vector space $\mathcal{H}$ is a set of vectors $\{e_i\}_i$ which is such that each $\psi \in \mathcal{H}$ can, in a unique manner, be written as

$$\psi = \sum_i c_i \cdot e_i$$

for some set of complex numbers $\{c_i\}_i$, which we call the *coordinates* of ψ with respect to the basis $\{e_i\}_i$, and the number of basis vectors is the *dimension* of the vector space. Given a fixed basis $\{e_i\}_i$ of $\mathcal{H}_1$ any linear operator $f : \mathcal{H}_1 \to \mathcal{H}_2$ is completely determined by its action on the basis vectors since

$$f(\psi) = f\Big(\sum_i c_i \cdot e_i\Big) = \sum_i c_i \cdot f(e_i)\,.$$

Moreover, since given a basis $\{e'_i\}_i$ for $\mathcal{H}_2$ for each $f(e_i)$ there exists a unique set $\{m_{ij}\}_i$ such that

$$f(e_j) = \sum_i m_{ij} \cdot e'_i\,,$$

f is completely determined by its *matrix* $(m_{ij})_{ij}$ with respect to the basis $\{e_i\}_i$ and $\{e'_i\}_i$. By convention, the index i runs over *rows* and the index j runs over columns i.e.

$$\begin{pmatrix} m_{11} & \cdots & m_{1j} & \cdots & m_{1m} \\ \vdots & & \vdots & & \vdots \\ m_{i1} & \cdots & m_{ij} & \cdots & m_{im} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \cdots & m_{nj} & \cdots & m_{nm} \end{pmatrix}.$$

When applying f to a vector $\psi = \sum_j c_j \cdot e_j$ we have

$$\begin{aligned} f\Big(\sum_j c_j \cdot e_j\Big) &= \sum_j c_j \cdot f(e_j) \\ &= \sum_j c_j \cdot \Big(\sum_i m_{ij} \cdot e'_i\Big) \\ &= \sum_i \Big(\sum_j m_{ij} c_j\Big) \cdot e'_i \end{aligned}$$

so we obtain the usual formula for application of the matrix of f to the *column of vector coordinates* of ψ i.e.

$$\begin{pmatrix} \sum_j m_{1j} c_j \\ \vdots \\ \sum_j m_{ij} c_j \\ \vdots \\ \sum_j m_{nj} c_j \end{pmatrix} = \begin{pmatrix} m_{11} & \cdots & m_{1j} & \cdots & m_{1m} \\ \vdots & & \vdots & & \vdots \\ m_{i1} & \cdots & m_{ij} & \cdots & m_{im} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \cdots & m_{nj} & \cdots & m_{nm} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}.$$

On the other hand, for $(m'_{ij})_{ij}$ the matrix of another linear operator $g : \mathcal{H}_2 \to \mathcal{H}_3$ with respect to the basis $\{e''_i\}_i$ — we then apply the above result twice and obtain

$$\begin{aligned} (g \circ f)(e_k) &= g\Big(\sum_j m_{jk} \cdot e'_j\Big) \\ &= \sum_j m_{jk} \cdot g(e'_j) \\ &= \sum_j m_{jk} \cdot \Big(\sum_i m'_{ij} \cdot e''_i\Big) \\ &= \sum_i \Big(\sum_j m'_{ij} m_{jk}\Big) \cdot e''_i \end{aligned}$$

so we obtain the usual formula for post-composing the matrix of g with the matrix of f, i.e.

$$\begin{pmatrix} m'_{11} & \cdots & m'_{1j} & \cdots & m'_{1m} \\ \vdots & & \vdots & & \vdots \\ m'_{i1} & \cdots & m'_{ij} & \cdots & m'_{im} \\ \vdots & & \vdots & & \vdots \\ m'_{n1} & \cdots & m'_{nj} & \cdots & m'_{nm} \end{pmatrix} \begin{pmatrix} m_{11} & \cdots & m_{1k} & \cdots & m_{1\mu} \\ \vdots & & \vdots & & \vdots \\ m_{j1} & \cdots & m_{jk} & \cdots & m_{j\mu} \\ \vdots & & \vdots & & \vdots \\ m_{m1} & \cdots & m_{mk} & \cdots & m_{m\mu} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_j m'_{1j} m_{j1} & \cdots & \sum_j m'_{1j} m_{jk} & \cdots & \sum_j m'_{1j} m_{j\mu} \\ \vdots & & \vdots & & \vdots \\ \sum_j m'_{ij} m_{j1} & \cdots & \sum_j m'_{ij} m_{jk} & \cdots & \sum_j m'_{ij} m_{j\mu} \\ \vdots & & \vdots & & \vdots \\ \sum_j m'_{nj} m_{j1} & \cdots & \sum_j m'_{nj} m_{jk} & \cdots & \sum_j m'_{nj} m_{j\mu} \end{pmatrix}.$$

In fact, application of a linear operator to a vector can itself also be seen as a composition of functions, noting that there is a one-to-one correspondence between the vectors $\psi \in \mathcal{H}$ and the linear functions

$$\mathbb{C} \to \mathcal{H} :: 1 \mapsto \psi\,,$$

the matrix of the latter being the column of coordinates of ψ. Note here also that having a matrix calculus is a property also satisfied by relations — although not over a field but over a *semiring* i.e. a ring without inverses — cf. $\{0, 1\}$-valued matrices exactly encode all relations, hence in particular also all multi-valued ones. In fact, structurally, linear operators are mathematically much closer to being relations than to being functions — a fact which we will return to later in this course.

Having an inner-product on top of a vector space structure, i.e. having a Hilbert space as in Definition 3.1, turns out to be the same thing as fixing a basis in that vector space.

Exercise 3.10 For a vector space $\mathcal{H}$ with basis $\{e_i\}_i$ show that

$$\langle \phi \,|\, \psi \rangle := \begin{pmatrix} \overline{c'_1} & \cdots & \overline{c'_j} & \cdots & \overline{c'_n} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_j \\ \vdots \\ c_n \end{pmatrix}$$

with $\psi = \sum_i c_i \cdot e_i$ and $\phi = \sum_i c'_i \cdot e_i$ indeed defines an inner-product in the sense of Definition 3.1.

In a Hilbert space a basis is called *orthonormal* if

$$\langle e_i | e_j \rangle = \delta_{ij}$$

where $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ii} = 1$. In this case, since

$$\langle e_i | \psi \rangle = \Big\langle e_i \,\Big|\, \sum_j c_j \cdot e_j \Big\rangle = \sum_j c_j \langle e_i | e_j \rangle = c_i$$

we obtain the coordinates through the inner-product. Since

$$\langle \phi | \psi \rangle = \Big\langle \sum_i c'_i \cdot e_i \,\Big|\, \sum_j c_j \cdot e_j \Big\rangle = \sum_{ij} \overline{c}'_i c_j \langle e_i | e_j \rangle = \sum_i \overline{c}'_i c_i$$

in any Hilbert space the inner-product always coincides with the one defined in Exercise 3.10 whenever $\{e_i\}_i$ is an orthonormal basis of $\mathcal{H}$. It can be shown that each Hilbert space admits an orthonormal basis so what we can do with Hilbert spaces can be done in matrix calculus over $\mathbb{C}$ with the inner-product of Exercise 3.10.

Exercise 3.11 If $(m_{ij})_{ij}$ is the matrix of $f : \mathcal{H}_1 \to \mathcal{H}_2$ in some basis for $\mathcal{H}_1$ and some basis for $\mathcal{H}_2$ show that $(\overline{m}_{ji})_{ij}$ is the matrix of its adjoint $f^\dagger$ with respect to those basis.

So when we have agreed on a fixed basis for each Hilbert space, by Exercise 3.11.i the matrix corresponding to the adjoint of the linear operator with matrix

$$\begin{pmatrix} m_{11} & \cdots & m_{1j} & \cdots & m_{1m} \\ \vdots & & \vdots & & \vdots \\ m_{i1} & \cdots & m_{ij} & \cdots & m_{im} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \cdots & m_{nj} & \cdots & m_{nm} \end{pmatrix}$$

is its *conjugate transposed*

$$\begin{pmatrix} \overline{m}_{11} & \cdots & \overline{m}_{i1} & \cdots & \overline{m}_{n1} \\ \vdots & & \vdots & & \vdots \\ \overline{m}_{1j} & \cdots & \overline{m}_{ij} & \cdots & \overline{m}_{nj} \\ \vdots & & \vdots & & \vdots \\ \overline{m}_{1m} & \cdots & \overline{m}_{im} & \cdots & \overline{m}_{nm} \end{pmatrix},$$

and hence the matrix of a general self-adjoint operator is

$$\begin{pmatrix} r_{11} & \cdots & \overline{m}_{in} & \cdots & \overline{m}_{n1} \\ \vdots & & \vdots & & \vdots \\ m_{i1} & \cdots & r_{ii} & \cdots & \overline{m}_{in} \\ \vdots & & \vdots & & \vdots \\ m_{n1} & \cdots & m_{in} & \cdots & r_{nn} \end{pmatrix}$$

where $r_1, \ldots, r_i, \ldots, r_n \in \mathbb{R}$.

As is typical, we adopt the notation of writing a $\dagger$ in superscript after an operator (e.g. $A^\dagger$) to denote both the adjoint of a linear map and the conjugate transpose of a matrix. Since unitary operators preserve the inner-product, a unitary transformation sends an orthonormal basis $\{e_i\}_i$ to another orthonormal basis $\{U(e_i)\}_i$, and conversely, as is the case for any linear operator, a unitary transformation is completely determined by its action on an orthonormal basis. So when fixing an orthonormal basis, there is a bijective correspondence between unitary operators and the set of all orthonormal basis.

Exercise 3.12 i. Describe the matrix of a unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ with respect to the basis $\{e_i\}_i$ of $\mathcal{H}_1$ and the basis $\{U(e_i)\}_i$ of $\mathcal{H}_2$. **ii.** Describe the matrix of a unitary operator $U : \mathcal{H} \to \mathcal{H}$ with respect to the basis $\{e_i\}_i$ of $\mathcal{H}$ — hint: do this in terms of the vectors in $\{U(e_i)\}_i$. **iii.** For unitary operators of type $\mathcal{H}_1 \to \mathcal{H}_2$, explicitly describe the bijective correspondence with orthonormal basis of $\mathcal{H}_2$.

Let $(m_{jk})_{jk}$ be the matrix of $f : \mathcal{H}_1 \to \mathcal{H}_2$ for orthonormal basis $\{e_k\}_k$ of $\mathcal{H}_1$ and $\{e'_j\}_k$ of $\mathcal{H}_2$ i.e.

$$f(e_k) = \sum_j m_{jk} \cdot e'_j \,.$$

We would like to know what the matrix of f is for basis $\{U(e_k)\}_k$ of $\mathcal{H}_1$ and $\{U'(e'_j)\}_j$ of $\mathcal{H}_2$. Let

$$U(e_l) = \sum_k u_{kl} \cdot e_k \quad \text{and} \quad e'_j = \sum_i \overline{u}'_{ji} \cdot U'(e'_i)$$

where $(u_{kl})_{kl}$ is the matrix of U for the basis $\{e_i\}_i$ and $(u'_{ij})_{ij}$ is the matrix of U' for the basis $\{e'_i\}_i$, and hence $(\bar{u}_{ji})_{ij}$ the matrix of $U^{-1} = U^\dagger$ in that basis. We obtain

$$\begin{aligned} f(U(e_l)) &= f\Big(\sum_k u_{kl} \cdot e_k\Big) \\ &= \sum_k u_{kl} \cdot f(e_k) \\ &= \sum_k u_{kl} \cdot \Big(\sum_j m_{jk} \cdot e'_j\Big) \\ &= \sum_k u_{kl} \cdot \Big(\sum_j m_{jk} \cdot \Big(\sum_l \overline{u}'_{ji} \cdot U'(e'_i)\Big)\Big) \\ &= \sum_i \Big(\sum_{jk} \overline{u}'_{ji} m_{jk} u_{kl}\Big) \cdot U'(e'_i) \end{aligned}$$

so the resulting matrix for f in the basis $\{U(e_k)\}_k$ of $\mathcal{H}_1$ and $\{U'(e'_j)\}_k$ of $\mathcal{H}_2$ is the matrix product

$$(u'_{ij})^\dagger_{ij}\,(m_{jk})_{jk}\,(u_{kl})_{kl}\,.$$

When denoting the matrices of f, U and U' when expressed in the basis $\{e_i\}_i$ and $\{e'_i\}_i$ (slightly abusively) also as f, U and U' this expression simplifies to

$$U'^\dagger \circ f \circ U\,.$$

Proposition 3.13 *For each self-adjoint operator $H : \mathcal{H} \to \mathcal{H}$ there exists an orthonormal basis in which its matrix is 'diagonal ' i.e. all its non-diagonal elements become* 0*.*

Fixing an orhonormal basis $\{e_i\}_i$ in which we express all matrices let $\{U(e_i)\}_i$ be the basis in which the matrix of H is diagonal i.e., continuing our abuse of notation for the matrices of f and U in $\{e_i\}_i$,

$$U^\dagger \circ f \circ U\,.$$

is diagonal, so for the matrix of f expressed in $\{e_i\}_i$ we have

$$\begin{aligned} f &= (U \circ U^\dagger) \circ f \circ (U \circ U^\dagger) \\ &= U \circ (U^\dagger \circ f \circ U) \circ U^\dagger \\ &= U \circ M \circ U^\dagger \end{aligned}$$

where M is some diagonal matrix. Conversely, each matrix $N = U \circ M \circ U^\dagger$ with M diagonal defines a self-adjoint operator by Exercise 3.6.i, namely the one which has matrix N in the basis $\{e_i\}_i$, since we can interpret M as the matrix of a linear operator expressed in the basis $\{U(e_i)\}_i$. Note that this argument also provides a converse to Proposition 3.13.

Exercise 3.14 i. Is the orthonormal basis in which the matrix of a self-adjoint operator becomes diagonal unique? **ii.** Describe the matrix of general projectors in an orthonormal basis in which its matrix is diagonal. **iii.** Relying on Proposition 3.13 explicitly construct the spectral decomposition (cf. Proposition 3.7) of a self-adjoint operator $H : \mathcal{H} \to \mathcal{H}$.

Exercise 3.15 i. Which projectors P_+ and P_- have the states respectively described by $e_+ := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $e_- := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as their 'only outcome states' i.e. the range of the projector is the ray spanned by that vector. **ii.** Given $e_\theta := \begin{pmatrix} 1 \\ e^{i\theta} \end{pmatrix}$, pick a vector $e_\theta^\perp$ which is orthogonal e_θ and give the projectors P_θ and $\mathrm{P}_\theta^\perp$ which have the states described by these vectors as their only outcome states. **iii.** Give the matrices of the unitary operators U_+ and U_θ which are such that

$$\mathrm{P}_+ = U_+ \circ \mathrm{P}_0 \circ U_+^\dagger \qquad \text{and} \qquad \mathrm{P}_\theta = U_\theta \circ \mathrm{P}_0 \circ U_\theta^\dagger\,.$$

for $\mathrm{P}_0 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ **iv.** Give the probabilities for the input states e_-, e_+, e_θ and $e_\theta^\perp$ for the measurements

$$\{\mathrm{P}_+, \mathrm{P}_-^\perp\} \qquad \text{and} \qquad \{\mathrm{P}_\theta, \mathrm{P}_\theta^\perp\}\,.$$

Quantum mechanics in matrix terms. We provided both unitary operators and quantum measurements (as families of mutually orthogonal projectors arising from a self-adjoint operator through the spectral decomposition theorem) with an easy matrix representation, given a fixed basis $\{e_i\}_i$:

- Unitary operators are in one-to-one correspondence with ONBs, and represent in matrix terms as the ONB $\{U(e_i)\}_i$ written as a list of column vectors

$$\begin{pmatrix} U(e_1) & \cdots & U(e_n) \end{pmatrix}\,.$$

- *Non-degenerate quantum measurements* — i.e. quantum measurements for which the spectral decomposition only contains of projectors with rays as range, or equivalently, for which the number of mutually orthogonal projectors is equal to the dimension of the Hilbert space — are completely determined by a unitary operator, which itself is completely determined by a basis $\{U(e_i)\}_i$, with respect to which the quantum measurement is given by

$$\{U \circ \mathrm{P}_1 \circ U^\dagger, \ldots, U \circ \mathrm{P}_n \circ U^\dagger\}$$

where

$$\mathrm{P}_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \quad \cdots \quad \mathrm{P}_n = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

The outcome state for the projector $U \circ \mathrm{P}_i \circ U^\dagger$ is exactly the state described by the basis vector $U(e_i)$. This justifies the extremely handy slogan:

non-degenerate measurement = *orthonormal basis*!

- Degenerate quantum measurements require, in addition to a basis, specification of a partition

$$\{1, \ldots, n\} = I_1 \cup \ldots \cup I_k$$

which provides (as projectors) a family diagonal matrices which have 0s everywhere except for the ith diagonal elements for $i \in I_j$ where there are 1s i.e.

$$\left\{ \begin{pmatrix} r^j_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r^j_{nn} \end{pmatrix} \;\middle|\; \begin{array}{l} 1 \leq j \leq k\,,\, r^j_{ii} \in \{0,1\} \\ r^j_{ii} = 1 \Leftrightarrow i \in I_j \end{array} \right\}$$

- *Physically speaking*, basisd on the above we can make two choices for implementing a particular (non-degenerate) measurement $\{U \circ \mathrm{P}_- \circ U^\dagger\}_i$:

  1. We perform a measurement in the basis $\{U(e_i)\}_i$ i.e. we implement the projectors
  $$\{U \circ \mathrm{P}_1 \circ U^\dagger, \ldots, U \circ \mathrm{P}_n \circ U^\dagger\}\,;$$
  2. We first perform the unitary transformation $U^\dagger$, then the measurement in the basis $\{e_i\}_i$ i.e. we implement the projectors
  $$\{\mathrm{P}_1, \ldots, \mathrm{P}_n\}\,,$$
  and then (provided we are not just interested in the measurement outcome but also in the resulting state) we perform the unitary transformation U.

  The advantage of the second implementation is that we only need to rely on one particular quantum measurement, independent on which measurement we actually want to implement, namely the one in the fixed basis.

3.3 Tensor structure

We define the *direct sum* of Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ as

$$\mathcal{H}_1 \oplus \mathcal{H}_2 := \left\{(\psi_1, \psi_2) \;\middle|\; \psi_1 \in \mathcal{H}_1, \psi_2 \in \mathcal{H}_2\right\}$$

together with two linear *injections*

$$\iota_1 : \mathcal{H}_1 \to \mathcal{H}_1 \oplus \mathcal{H}_2 :: \psi \mapsto (\psi, \mathbf{0})$$
$$\iota_2 : \mathcal{H}_2 \to \mathcal{H}_1 \oplus \mathcal{H}_2 :: \psi \mapsto (\mathbf{0}, \psi)\,.$$

It is simple to extend the definition of the direct sum beyond the binary case — i.e. we can also consider $\bigoplus_i \mathcal{H}_i$ together with a family of linear injections

$$\{\iota_j : \mathcal{H}_1 \to \bigoplus_i \mathcal{H}_i\}_j\,.$$

We define the direct sum of two linear maps $f : \mathcal{H}_1 \to \mathcal{H}_1'$ and $g : \mathcal{H}_2 \to \mathcal{H}_2'$ component-wise i.e. as the linear map

$$f \oplus g : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1' \oplus \mathcal{H}_2' :: (\psi, \phi) \mapsto (f(\psi), g(\phi))\,.$$

Exercise 3.16 **i.** Show that $\mathcal{H}_1 \oplus \mathcal{H}_2$ is indeed a Hilbert space i.e. show that it comes with an inner-product — write this inner-product down without referring to a basis. **ii.** Describe a basis of $\mathcal{H}_1 \oplus \mathcal{H}_2$ in terms of basis for the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, and describe the coordinates of the elements of $\mathcal{H}_1 \oplus \mathcal{H}_2$ in that basis. **iii.** Show that

$$\sum_i \iota_i \circ \iota_i^\dagger = 1_{\oplus_i \mathcal{H}_i} \qquad \text{and} \qquad \iota_j^\dagger \circ \iota_i = \delta_{ij}\,.$$

where $\delta_{ij} = O : \mathcal{H}_i \to \mathcal{H}_j :: \psi \mapsto \mathbf{0}$ iff $i \neq j$ and $\delta_{ii} = 1_{\mathcal{H}_i}$.

A key mathematical application of the direct sum is that it provides Hilbert spaces which come with a *preferred basis*.

Exercise 3.17 Show that for each $n \in \mathbb{N}$ the direct sum provides a Hilbert space of dimension n together with a canonical choice for a basis — hint: recall that $\mathbb{C}$ is itself a Hilbert space which comes with a special element $1 \in \mathbb{C}$ (= the multiplicative inverse of the underlying field $\mathbb{C}$).

Physically, the direct sum describes *pairs of states*, and hence would constitute a likely candidate to describe compound systems, but quantum theory disagrees! This means we will need to introduce the *tensor product* of two Hilbert spaces.

Postulate 3.18 [compound systems] The joint state of a compound quantum system consisting of two subsystems is described by the tensor product of the Hilbert spaces which describe the two subsystems.

So now we will define the tensor product of Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, which allows several formulations that are equivalent 'up to isomorphism'. Initially we define it as

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := \left\{ \begin{pmatrix} c_{11} & \ldots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \ldots & c_{nm} \end{pmatrix} \;\middle|\; \forall i, j : c_{ij} \in \mathbb{C} \right\},$$

where n is the dimension of $\mathcal{H}_1$ and m is the dimension of $\mathcal{H}_2$, together with the map

$$\xi : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 :: (\psi, \phi) \mapsto (c_i c_j')_{ij}$$

where $\psi = \sum_i c_i \cdot e_i$ and $\phi = \sum_j c_j' \cdot e_j'$ for orthonormal basis $\{e_i\}_i$ of $\mathcal{H}_1$ and $\{e_i'\}_i$ of $\mathcal{H}_2$, that is, using coordinates,

$$\xi :: \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \begin{pmatrix} c_1' \\ \vdots \\ c_m' \end{pmatrix} \right) \mapsto \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \begin{pmatrix} c_1' & \ldots & c_m' \end{pmatrix}.$$

Note that *the elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ are exactly the matrices of the linear maps of type $\mathcal{H}_2 \to \mathcal{H}_1$* (which by the taking the transpose or the adjoint are equivalent to the matrices of the linear maps of type $\mathcal{H}_1 \to \mathcal{H}_2$),[1] and when thinking of

[1] We could of course also have defined the tensor product such that it is exactly the matrices of the linear maps of type $\mathcal{H}_1 \to \mathcal{H}_2$, but in that case the presentation would have been slightly less clean due to the fact that matrix composition goes backward as compared to the (western) reading direction i.e. we would have

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := \left\{ \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \ldots & c_{mn} \end{pmatrix} \;\middle|\; \forall i, j : c_{ij} \in \mathbb{C} \right\},$$

together with

$$\xi : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 :: (\psi, \phi) \mapsto (c_j' c_i')_{ji}$$

vectors as linear functions of type $\mathbb{C} \to \mathcal{H}$ (cf. the discussion above), for the corresponding matrices we have

$$\xi :: (\psi, \phi) \mapsto \psi \circ \phi^T$$

where $(-)^T$ denotes the transpose. One easily verifies that $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a Hilbert space for the inner-product

$$\langle (c_{ij})_{ij} \mid (c'_{ij})_{ij} \rangle := \sum_{ij} \overline{c}_{ij} c'_{ij} \, .$$

A basis for the tensor product is provided by all $m \times n$-matrices which only contain a single 1 while all the other elements are 0. Notice that when applying ξ to a pair (e_i, e'_j) for basis $\{e_i\}_i$ of $\mathcal{H}_1$ and $\{e'_j\}_j$ of $\mathcal{H}_2$ we obtain the matrix with 1 in the jth row and ith column spot, and which is 0 everywhere else, in particular,

$$\xi :: (e_i, e'_j) \mapsto e_i \circ e'^T_j$$

for the matrices of the linear maps representing these basis vectors. So roughly speaking, while we obtain a basis for the direct sum by taking the '(disjoint) union' of the basis vectors of the underlying spaces, *we obtain a basis for the tensor product by taking the 'cartesian product' of the basis vectors of the underlying spaces*. In particular:

$$dim(\mathcal{H}_1 \oplus \mathcal{H}_2) = dim(\mathcal{H}_1) + dim(\mathcal{H}_2),$$

$$dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = dim(\mathcal{H}_1) \times dim(\mathcal{H}_2).$$

Finally, the above defined tensor product easily extends beyond the binary case by setting

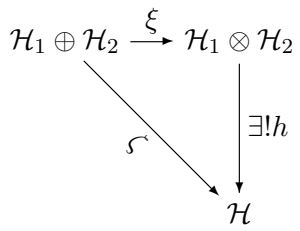
$$\bigotimes_{i=1}^{i=k} \mathcal{H}_i := \left\{ (c_{i_1 \ldots i_k})_{i_1 \ldots i_k} \;\middle|\; \forall i_1, \ldots, i_k : c_{i_1 \ldots i_k} \in \mathbb{C} \right\}$$

together with the map

$$\xi : \bigoplus_{i=1}^{i=k} \mathcal{H}_i \to \bigotimes_{i=1}^{i=k} \mathcal{H}_i :: (\psi_1, \ldots, \psi_k) \mapsto (c_{i_1} \ldots c_{i_k})_{i_1 \ldots i_k}$$

where $\psi_1 = \sum_{i_1} c_{i_1} \cdot e_{i_1}$, . . . , and $\psi_k = \sum_{i_k} c_{i_k} \cdot e_{i_k}$.

Exercise 3.19 i. Show that ξ is <u>not</u> linear. **ii.** Show that ξ is *bilinear* i.e. linear when conceived as a two-variable function where one variable takes its values in $\mathcal{H}_1$ and the other one in $\mathcal{H}_2$. **iii.** Show that for any other bilinear map $\zeta : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}$ there exists a <u>unique</u> linear map $h : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}$ such that $\zeta = h \circ \xi$ i.e. in a *commutative diagram*:



To stress that ξ is not 'globally' linear one rather writes its domain as the cartesian product $\mathcal{H}_1 \times \mathcal{H}_2$ than as the direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$. The so-called *universal property* of the tensor product expressed in Exercise 3.19.iii is in many cases taken as the definition of the tensor product, which defines it up to an 'isomorphism of vector spaces' — cf. [**?**] §8.7 & §8.8.

While our definition for the tensor product is very straightforward, it does depend on a choice of basis. We will work toward a basis-independent (but slightly less straightforward) construction. First replace the matrices in the above definition of the tensor product by 'formal linear combinations' with respect to a basis for the tensor product i.e.

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := \left\{ \sum_{ij} c_{ij} \cdot (e_i, e'_j) \;\middle|\; \forall i, j : c_{ij} \in \mathbb{C} \right\}$$

together with the map

$$\xi : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 :: (\psi, \phi) \mapsto (\psi, \phi)$$

where for $(\psi, \phi) \in \mathcal{H}_1 \otimes \mathcal{H}_2$ we set

$$\left(\sum_i c_i \cdot e_i \, , \sum_j c'_j \cdot e'_j \right) := \sum_{ij} c_i c'_j \cdot (e_i, e'_j) \, . \qquad (1)$$

This seems to allow us to think of 'pairs of states' as possible states of compound systems, but one still needs to be a bit cautious as is demonstrated by the following exercise.

Exercise 3.20 Show that ξ is <u>not</u> injective, and in particular, which $\psi \neq \phi$ are equalized under the action of ξ.

On the other hand, besides pairs of states there are many other ones too, the so-called *entangled states*, which are superpositions (= sums) of the pairs of states.

Exercise 3.21 Show that not all elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as (ψ, ϕ) for some $\psi \in \mathcal{H}_1$ and some $\phi \in \mathcal{H}_2$.

Usually the pairs $(\psi, \phi) \in \mathcal{H}_1 \otimes \mathcal{H}_2$ are denoted by $\psi \otimes \phi$ to indicate that we 'live in the tensor product' — and not in the directsum/cartesian product. Hence we obtain

$$\xi : \mathcal{H}_1 \oplus \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 :: (\psi, \phi) \mapsto \psi \otimes \phi,$$

which becomes

$$\xi :: \left(\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \begin{pmatrix} c'_1 \\ \vdots \\ c'_m \end{pmatrix} \right) \mapsto \begin{pmatrix} c'_1 \\ \vdots \\ c'_m \end{pmatrix} \begin{pmatrix} c_1 & \ldots & c_n \end{pmatrix} .$$

But the true 'natural' structural connection between the tensor product and linear maps is actually far more subtile than either of these two matricial candidates, as we will see in Section 11.2.

so $\xi = - \otimes -$, for which we moreover have

$$\Big(\sum_i c_i \cdot e_i\Big) \otimes \Big(\sum_i c'_i \cdot e'_i\Big) := \sum_{ij} c_i c'_j \cdot e_i \otimes e'_j \,.$$

We can actually think of $\psi \otimes \phi$ as a special case of the *tensor of two linear maps*, where we define the tensor of two linear maps $f : \mathcal{H}_1 \to \mathcal{H}'_1$ and $g : \mathcal{H}_2 \to \mathcal{H}'_2$ as the linear map

$$f \otimes g : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}'_1 \otimes \mathcal{H}'_2 :: \psi \otimes \phi \mapsto f(\psi) \otimes g(\phi) \,.$$

Hence you should be aware of the triple use of $- \otimes -$, namely

- as a map from pairs in the tensor product,
- as a connective on Hilbert spaces,
- as a tensor of two linear maps.

Exercise 3.22 i. First convince yourself that the above prescription for $f \otimes g$ is well-defined. Now rely on Exercise 3.19.iii to show (again) that the prescription of $f \otimes g$ indeed induces a unique linear map. **ii.** Show that

$$\langle \psi \otimes \psi' \mid \phi \otimes \phi' \rangle = \langle \psi | \phi \rangle \langle \psi' | \phi' \rangle \,.$$

indeed defines an inner-product on the tensor product of two Hilbert spaces. **iii.** Let $f, g : \mathcal{H} \to \mathcal{H}$ be linear maps with respective matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

in the basis $\{e_1, e_2\}$ of $\mathcal{H}$. Give the matrix of $f \oplus g$ in the basis

$$\{(e_1, \mathbf{0}), (e_2, \mathbf{0}), (\mathbf{0}, e_1), (\mathbf{0}, e_2)\}$$

and the matrix of $f \otimes g$ in the basis

$$\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\} \,.$$

iv. In these basis describe the 4×4 matrices for operations

$$\sigma_\oplus : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H} :: (\psi, \phi) \mapsto (\phi, \psi)$$

$$\sigma_\otimes : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H} :: \psi \otimes \phi \mapsto \phi \otimes \psi \,.$$

Are these operations unitary?

All the above results in a basis-independent construction of the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$. First we 'freely' introduce the following set of formal expressions

$$\Omega := \Big\{ \sum_i \alpha_i \cdot \psi_i \otimes \phi_i \;\Big|\; \forall i : \alpha_i \in \mathbb{C}, \psi_i \in \mathcal{H}_1, \phi_i \in \mathcal{H}_2 \Big\}$$

where all summations are finitary. Next we introduce a congruence (= equivalence relation) on Ω, namely

$$\sum_i \alpha_i \cdot \Big(\sum_j \beta_j \cdot \psi_{i,j}\Big) \otimes \phi_i \;\sim\; \sum_{ij} \alpha_i \beta_j \cdot \psi_{i,j} \otimes \phi_i$$

$$\sum_i \alpha_i \cdot \psi_i \otimes \Big(\sum_j \beta_j \cdot \phi_{i,j}\Big) \;\sim\; \sum_{ij} \alpha_i \beta_j \cdot \psi_i \otimes \phi_{i,j}$$

— which is clearly basis-independent. Expressing this congruence in respective basis for $\mathcal{H}_1$ and $\mathcal{H}_2$ yields

$$\sum_i \alpha_i \cdot \psi_i \otimes \phi_i \;\sim\; \sum_i \tilde{\alpha}_i \cdot \tilde{\psi}_i \otimes \tilde{\phi}_i$$

$$\Updownarrow$$

$$\forall j,k : \sum_i \alpha_i c^i_j d^i_k = \sum_i \tilde{\alpha}_i \tilde{c}^i_j \tilde{d}^i_k$$

where

$$\psi_i = \sum_j c^i_j \cdot e_j \quad , \quad \phi_i = \sum_k d^i_k \cdot e'_k \,,$$

$$\tilde{\psi}_i = \sum_j \tilde{c}^i_j \cdot e_j \quad , \quad \tilde{\phi}_i = \sum_k \tilde{d}^i_k \cdot e'_k \,,$$

which boils down to the matrices $(c_{jk})_{jk} := (\sum_i \alpha_i c^i_j d^i_k)_{jk}$ we started with. Conversely, each such matrix can be realised by setting $i := (j,k)$, $\psi_{jk} := e_j$, $\phi_{jk} := e'_k$ and $\alpha_{jk} := c_{jk}$,

The no-cloning theorem [8]. The passage from pairs of states to the tensor product, on which we only are allowed to act with unitary operations, comes with some drastic consequences. Assume we start with two quantum systems in states $\psi \otimes \phi_0 \in \mathcal{H}$ and we which, by means of some unitary operator $U : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$, to *copy* the state of the first one to the second one i.e. obtain $\psi \otimes \psi$. Assume we are able to do this both for ψ_1 and for ψ_2 i.e. we have

$$U(\psi_1 \otimes \phi_0) = \psi_1 \otimes \psi_1 \quad \text{and} \quad U(\psi_2 \otimes \phi_0) = \psi_2 \otimes \psi_2$$

— note here that it is crucial that in both cases U should be the same, since in general the state of the system is *unknown*, and measurement would alter it. Taking the inner-product of the above equalities yields

$$\langle U(\psi_1 \otimes \phi_0) | U(\psi_2 \otimes \phi_0) \rangle = \langle \psi_1 \otimes \psi_1 | \psi_2 \otimes \psi_2 \rangle \,,$$

that is, by $U^\dagger = U^{-1}$ and Exercise 3.22.ii,

$$\langle \psi_1 | \psi_2 \rangle \langle \psi_0 | \psi_0 \rangle = \langle \psi_1 | \psi_2 \rangle \langle \psi_1 | \psi_2 \rangle$$

and hence, assuming that all vectors are normalized,

$$\langle \psi_1 | \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle^2$$

forcing $\langle \psi_1 | \psi_2 \rangle = 0$ or $\langle \psi_1 | \psi_2 \rangle = 1$ i.e. ψ_1 and ψ_2 need to be either equal or orthogonal, so we cannot copy arbitrary states! There is also a corresponding no-deleting theorem [9], but that is slightly more subtle in its formulation.

Bell- and EPR-correlations. Two historically important examples of entangled states are the *Bell-state*

$$\mathsf{Bell} := e_1 \otimes e_1 + e_2 \otimes e_2$$

and the *EPR-state*

$$\mathsf{EPR} := e_1 \otimes e_2 - e_2 \otimes e_1$$

which respectively correspond to the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

i.e. the Bell-state corresponds to the identity, and the EPR-state seems hardly any more interesting. However, let's see what's happens when we measure them. First note that indeed there are no $a_1, a_2, a_3, a_4 \in \mathbb{C}$ such that either

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

so the Bell-state and the EPR-state are *truly entangled*, that is, cannot be written in the form $\psi \otimes \phi$. But the real magic starts when we measure them in the computational basis. If we measure the left system i.e. we apply the (degenerate) measurement

$$\{\mathrm{P}_1 \otimes \mathsf{id}, \mathrm{P}_2 \otimes \mathsf{id}\}$$

to the whole system we obtain

$$(\mathrm{P}_1 \otimes \mathsf{id})(\mathsf{Bell}) = e_1 \otimes e_1 \qquad (\mathrm{P}_1 \otimes \mathsf{id})(\mathsf{EPR}) = e_1 \otimes e_2$$

$$(\mathrm{P}_2 \otimes \mathsf{id})(\mathsf{Bell}) = e_2 \otimes e_2 \qquad (\mathrm{P}_2 \otimes \mathsf{id})(\mathsf{EPR}) = e_2 \otimes e_1$$

that is, we will now get a certain answer for a measurement on the second system i.e. we now apply

$$\{\mathsf{id} \otimes \mathrm{P}_1, \mathsf{id} \otimes \mathrm{P}_2\}$$

to the whole system, since in the case of the Bell-state we will always obtain the same outcome as we obtained when measuring the first system, while in the case of the EPR-state we will always obtain the opposite to what we obtained when measuring the first system (cf. identity). Typically these two systems are far apart so we witness a *non-local* effect. The experimental proof of this non-local effect requires making measurements in more than a single basis and the measurement outcomes must be shown to violate Bell's Inequality.

3.4 Dirac notation

A very popular notation in quantum mechanics and quantum informatics is the so-called *Dirac notation* or *bra-ket* notation [5]. Interestingly, while in most textbooks this is declared to be 'merely' a convenient notation, and sometimes even assumed as too informal and mathematically unsound, for us it will be a steping-stone to a compositional high-level formalism, and hence from the start we will formally justify it, and we will also need to slightly restrict it. Ultimately, we will extend it into a graphical notation in order to be able to cope with the intrinsic two-dimensional compositional structure of quantum mechanics cf. sequential composition $- \circ -$ and parallel composition $- \otimes -$. To justify this notation formally we want to think of vectors $\psi \in \mathcal{H}$ as linear maps

$$\mathbb{C} \to \mathcal{H} :: 1 \mapsto \psi\,,$$

which we (slightly abusively) also denote by ψ. The Dirac notation is formally justified by letting

- $|\psi\rangle := \psi$,
- $\langle\psi| := \psi^\dagger$,

in a table:

linear map	matrix	Dirac
$\psi : \mathbb{C} \to \mathcal{H}$	$\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$	$\lvert\psi\rangle$
$\psi^\dagger : \mathcal{H} \to \mathbb{C}$	$\begin{pmatrix} \bar{c}_1 & \dots & \bar{c}_m \end{pmatrix}$	$\langle\psi\rvert$

Hence in particular we have

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

We call $|\psi\rangle$ a *ket* and $\langle\psi|$ a *bra*. When writing one symbol after another we think of it as composition, either of linear functions or of the corresponding matrices. Hence an inner-product is a *bra-ket*(pronounced bracket):

linear map	matrix	Dirac notation
$\psi^\dagger \circ \phi$	$\begin{pmatrix} \bar{c}_1 & \dots & \bar{c}_m \end{pmatrix} \begin{pmatrix} c'_1 \\ \vdots \\ c'_m \end{pmatrix}$	$\langle\psi \mid \phi\rangle$

A projector on the ray spanned by ψ is a *ket-bra*:

linear map	matrix	Dirac
$\psi \circ \psi^\dagger$	$\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \begin{pmatrix} \bar{c}_1 & \dots & \bar{c}_m \end{pmatrix}$	$\mathrm{P}_\psi := \lvert\psi\rangle\langle\psi\rvert$

A projector on a ray indeed takes the shape $\psi \circ \psi^\dagger$. To see this, first note that with respect to a fixed basis we indeed have

$$\mathrm{P}_i = e_i \circ e_i^\dagger \,.$$

Hence, for a general projector

$$\mathrm{P}_\psi = U \circ \mathrm{P}_i \circ U^\dagger \qquad \text{with} \qquad \psi = U \circ e_i$$

— one verifies that each projector on a ray indeed admits such a representation using $\mathrm{P}_i \circ \phi = c \cdot e_i$ for some $c \in \mathbb{C}$ —

$$\begin{aligned} \mathrm{P}_\psi &= U \circ (e_i \circ e_i^\dagger) \circ U^\dagger \\ &= (U \circ e_i) \circ (e_i^\dagger \circ U^\dagger) \\ &= (U \circ e_i) \circ (U \circ e_i)^\dagger \\ &= \psi \circ \psi^\dagger. \end{aligned}$$

As an application of Dirac notation observe that

$$\mathrm{P}_\psi \circ \mathrm{P}_\phi = |\psi\rangle\underline{\langle\psi|\phi\rangle}\langle\phi| = O_{\mathcal{H}}$$

so two projectors on rays are orthogonal $\langle\psi|\phi\rangle = \mathbf{0}$ i.e. if and only if the rays ψ and ϕ on which they project are orthogonal — the underlined bra-ket is an inner-product, hence a scalar, so we obtain $c \cdot |\psi\rangle\langle\phi|$ for $c := \langle\psi|\phi\rangle$ which can indeed only be 0 if either c, ψ or ϕ would be $0/\mathbf{0}$. Here are some more examples of expressions in Dirac notation which illustrate the *compositional* nature of Dirac notation (here it is assumed that $f = f^\dagger$):

linear map	matrix	Dirac
$f \circ \psi$	$\begin{pmatrix} m_{11} & \dots & m_{1m} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{nm} \end{pmatrix} \begin{pmatrix} c'_1 \\ \vdots \\ c'_m \end{pmatrix}$	$f\|\psi\rangle$
$\phi^\dagger \circ f$	$\begin{pmatrix} \bar{c}_1 & \dots & \bar{c}_m \end{pmatrix} \begin{pmatrix} m_{11} & \dots & m_{1m} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{nm} \end{pmatrix}$	$\langle\phi\|f$
$\phi^\dagger \circ f \circ \psi$	$\dots = \sum_{ij} \bar{c}'_i m_{ij} c_j \in \mathbb{C}$	$\langle\phi\|f\|\psi\rangle$

As a special case we have probabilities for which we have

$$\langle\psi|\mathrm{P}_\phi|\psi\rangle = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\phi|\psi\rangle|^2$$

whenever the projector is of the form P_ϕ i.e. projects on a one-dimensional subspace. Hence we have again a very trivial computation which exposes an interesting feature: we can really think of quantum probabilities as some kind of *distance measure* between states. In Dirac notation one usually considers a privileged *computational basis* denoted by

$$\{\,|\,i\,\rangle \mid 0 \leq i \leq n-1\}$$

for each Hilbert space of dimension n. Mathematically speaking, such an n-dimensional Hilbert space which comes with a privileged basis can be produced as in Exercise 3.17, that is,

$$\mathcal{H} := \mathbb{C} \oplus \dots \oplus \mathbb{C}\,,$$

and in particular for a qubit we have

$$\mathcal{Q} := \mathbb{C} \oplus \mathbb{C}\,,$$

for which the computational basis vectors are $|\,0\,\rangle$ and $|\,1\,\rangle$. The vectors of the computational basis for $\mathcal{H} \otimes \mathcal{H}'$ are in the literature denoted in several ways:

$$|\,i\,\rangle \otimes |\,j\,\rangle \qquad |\,i\,\rangle|\,j\,\rangle \qquad |\,ij\,\rangle$$

but we will not be using the second which might cause confusion and even insinuates inconsistencies e.g. should

$$(|\psi\rangle\langle\psi|)(|\phi\rangle\langle\phi|)$$

be interpreted either as a composition or as a tensor i.e. as

$$|\psi\rangle\langle\psi|\phi\rangle\langle\phi| \qquad \text{or} \qquad |\psi \otimes \phi\rangle\langle\psi \otimes \phi|\,?$$

Exercise 3.23 Show that for projectors on rays we have

$$\mathrm{P}_\psi \otimes \mathrm{P}_\phi = \mathrm{P}_{\psi\otimes\phi}$$

by first showing that

$$(|\,\psi\rangle \otimes |\,\phi\rangle) \circ \lambda_{\mathbb{C}} = |\,\psi \otimes \phi\rangle\,,$$

where we used the unitary 'isomorphism' (check this!)

$$\lambda_{\mathbb{C}} : \mathbb{C} \to \mathbb{C} \otimes \mathbb{C} :: 1 \mapsto 1 \otimes 1\,,$$

by then showing that in general we have

$$(f_1 \otimes f_2) \circ (g_1 \otimes g_2) = (f_1 \circ g_1) \otimes (f_2 \circ g_2)\,,$$

and finally using these two facts to infer the above claim.

Conclusively, we have the following change of notation:

$$\begin{cases} e_1 \rightsquigarrow |0\rangle & e_{(i+1)} \otimes e_{(j+1)} \rightsquigarrow |ij\rangle \\ e_2 \rightsquigarrow |1\rangle & e_{(i_1+1)} \otimes \dots \otimes e_{(i_n+1)} \rightsquigarrow |i_1 \dots i_n\rangle \end{cases}$$

and hence an arbitrary state now takes the form

$$\Psi = \sum_i \alpha_i \,|\,i_1 \dots i_n\,\rangle\,.$$

If we assume that 'larger Hilbert spaces' $\mathcal{H}$ always arise as

$$\mathcal{H} := \mathcal{Q} \otimes \ldots \otimes \mathcal{Q}$$

then, using the notation

$$|i_1 \ldots i_n\rangle \qquad \text{with} \qquad i_1 \ldots i_n \in \{0,1\}$$

for the basis, this is actually nothing more then writing

$$|i\rangle \qquad \text{with} \qquad i \in \{0, \ldots, 2^n - 1\}$$

in binary rather than in decimal (or anything else). Note that this binary representation also allows for 'easy comparison' with classical computing with bit-strings. Important examples are the Bell-state and the EPR-state

$$Bell = |\,00\rangle + |\,11\rangle \qquad \text{and} \qquad EPR = |\,01\rangle - |\,10\rangle\,.$$

or, for three qubits, the *GHZ-state* and the *W-state*

$$GHZ := |\,000\rangle + |\,111\rangle \quad \text{and} \quad W = |\,100\rangle + |\,010\rangle + |\,001\rangle\,.$$

Usually one introduces a normalization constant resulting in

$$\frac{1}{\sqrt{2}}\left(|\,00\rangle + |\,11\rangle\right) \qquad \text{and} \qquad \frac{1}{\sqrt{2}}\left(|\,000\rangle + |\,111\rangle\right)$$

assuming that the basis vectors are normalized, but for both esthetic and ecological reasons we will drop these. The standard single qubit and two-qibit computational basis are

$$\{|0\rangle\,,\,|1\rangle\} \qquad \{|00\rangle\,,\,|01\rangle\,,\,|10\rangle\,,\,|11\rangle\}$$

with corresponding measurement projectors

$$\{|0\rangle\langle 0|\,,|1\rangle\langle 1|\} \quad \{|00\rangle\langle 00|\,,|01\rangle\langle 01|\,,|10\rangle\langle 10|\,,|11\rangle\langle 11|\}\,.$$

General self-adjoint operators take the form[2]

$$H = U\Big(\sum_{i=n}^{i=n} \alpha_i \cdot |\,i\rangle\langle i\,|\Big)U^\dagger = \sum_{i=n}^{i=n} \alpha_i \cdot U|\,i\rangle\langle i\,|U^\dagger$$

and for projectors we have $a_i \in \{0,1\}$, hence they are

$$\mathrm{P} = U\Big(\sum_{i\in I} |\,i\rangle\langle i\,|\Big)U^\dagger = \sum_{i\in I} U|\,i\rangle\langle i\,|U^\dagger$$

for $I \subseteq \{0,\ldots,n\}$, so quantum measurements take the shape

$$\Big\{\sum_{i\in I_1} U|\,i\rangle\langle i\,|U^\dagger\,,\;\ldots\;,\;\sum_{i\in I_k} U|\,i\rangle\langle i\,|U^\dagger\Big\}$$

for a partition $I_1 \cup \ldots \cup I_k = \{1,\ldots,n\}$.

From now on, we will use the same notation for a linear operator and its matrix in the computational basis except when it is explicitly stated to be otherwise. Consider for example arbitrary *qubit measurement*with as set of projectors

$$\Big\{\,\mathrm{P}_0^U := U|\,0\rangle\langle 0\,|U^\dagger\,,\,\mathrm{P}_1^U := U|\,1\rangle\langle 1\,|U^\dagger\,\Big\}$$

physically either being

$$\Big\{\,\mathrm{P}_0^U := (U|\,0\rangle)(\langle 0\,|U^\dagger)\,,\,\mathrm{P}_1^U := (U|\,1\rangle)(\langle 1\,|U^\dagger)\,\Big\}\,,$$

or

$$\Big\{\,\mathrm{P}_0^U := U(|\,0\rangle\langle 0\,|)U^\dagger\,,\,\mathrm{P}_1^U := U(|\,1\rangle\langle 1\,|)U^\dagger\,\Big\}\,.$$

If we perform this measurement then two possible outcome states $U|\,0\rangle$ and $U|\,1\rangle$ can be obtained, 'up to normalization', by post-composing the input state ψ with the respective projectors i.e.

$$U|\,0\rangle\underline{\langle 0\,|U^\dagger|\psi\rangle} \qquad \text{and} \qquad U|\,1\rangle\underline{\langle 1\,|U^\dagger|\psi\rangle}$$

The underlined scalars yield the corresponding probabilities when multiplying them with their conjugate since

$$\langle\psi|\mathrm{P}_i^U|\psi\rangle = \langle\psi|U|\,i\rangle\langle i\,|U^\dagger|\psi\rangle = \overline{\langle i\,|U^\dagger|\psi\rangle}\langle i\,|U^\dagger|\psi\rangle\,.$$

Non-local correlations in Dirac Notation. To illustrate Dirac notation in action we now redo this calculation. Measurement in the computational basis of the first qubit of a qubit pair in the Bell-state either yields

$$\begin{aligned}
&(\mathrm{P}_0 \otimes 1_{\mathcal{Q}})(|\,00\rangle + |\,11\rangle)\\
&\quad = (|\,0\rangle\langle 0\,| \otimes 1_{\mathcal{Q}})(|\,00\rangle + |\,11\rangle)\\
&\quad = (|\,0\rangle\langle 0\,| \otimes 1_{\mathcal{Q}})|\,00\rangle + (|\,0\rangle\langle 0\,| \otimes 1_{\mathcal{Q}})|\,11\rangle\\
&\quad = (|\,0\rangle\langle 0\,|\,0\rangle) \otimes |\,0\rangle + (|\,0\rangle\langle 0\,|\,1\rangle) \otimes |\,1\rangle = |\,00\rangle
\end{aligned}$$

or

$$(\mathrm{P}_1 \otimes 1_{\mathcal{Q}})(|\,00\rangle + |\,11\rangle) = |\,11\rangle$$

which indeed yields a certain answer for a measurement on the second qubit, hence a non-local correlation.

4 Protocols from entanglement

With the machinery of tensor products and Dirac notation at hand we are now able to expose certain protocols, which surprisingly have only been recently discovered, and lift the 'weirdness' of non-local correlations one level higher — to outer space if you wish, where they meet the crew abord *Star Trek's USS Enterprise*.

[2]Note that, of course, in general

$$\Big|\sum_i \psi_i\Big\rangle\Big\langle\sum_i \psi_i\,\Big| \neq \sum_i |\psi_i\rangle\langle\psi_i|$$

e.g. for qubits $(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$ is a projector on the state $|0\rangle + |1\rangle$, while $|0\rangle\langle 0| + |1\rangle\langle 1|$ is the identity! Hence the sum seems to play two roles, and this will enable us to accomodate so-called *mixed states* to be introduced and studied in Section 11.4.

4.1 Bell-basis and Bell-matrices

While the standard 2-qubit quantum measurement is with respect to the computational basis

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle$$

a very important measurment basis is the *Bell-basis*:

$$|00\rangle + |11\rangle, \quad |00\rangle - |11\rangle, \quad |01\rangle + |10\rangle, \quad |01\rangle - |10\rangle\,.$$

This basis is obtained by respectively applying the unitaries

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

to the second qubit of the Bell-state, i.e. applying $1_{\mathcal{Q}} \otimes U$ to the whole system. For example, when we apply the fourth Bell matrix $U := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ to the qubit basis we obtain

$$|\,0\rangle \mapsto |\,1\rangle \qquad \text{and} \qquad |\,1\rangle \mapsto -|\,0\rangle$$

hence when we apply id $\otimes\, U$ we have

$$|\,00\rangle \mapsto |\,01\rangle \qquad \text{and} \qquad |\,11\rangle \mapsto -|\,10\rangle$$

so id $\otimes\, U$ applied to the Bell-state yields

$$|\,00\rangle + |\,11\rangle \mapsto |\,01\rangle - |\,10\rangle$$

i.e. the fourth Bell-basis vector. We call these matrices the *Bell-matrices*, and they are exactly (the transposed of) the matrices encoding the Bell-basis in our matricial definition of the tensor product, cf.

$$\sum_{ij} c_{ij}|\,ij\rangle \quad \stackrel{\simeq}{\longleftrightarrow} \quad (c_{ij})_{ij}$$

Alternatively, we can apply the transposed Bell-matrices to the first qubit of the Bell-state, i.e. applying $U^T \otimes 1_{\mathcal{Q}}$ to the whole system, and again we exactly obtain the Bell-basis. Hence, denoting the Bell-matrices by $\{U_i\}_i$ and the Bell-basis by $\{\Psi_i\}_i$ we have

$$(U_i^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle = \Psi_i = (1_{\mathcal{Q}} \otimes U_i)|\Psi_{Bell}\rangle$$

Note also that we produce a *Bell-basis measurement*

$$\Big\{(|00\rangle + |11\rangle)(\langle 00| + \langle 11|)\,,\\ (|00\rangle - |11\rangle)(\langle 00| - \langle 11|)\,,\\ (|01\rangle + |10\rangle)(\langle 01| + \langle 10|)\,,\\ (|01\rangle - |10\rangle)(\langle 01| - \langle 10|)\Big\}$$

from a measurement in the computational basis together with

$$U_{\textit{Bell-basis}} := \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$

for which we have (note the above matrix factors into a product $H \otimes I$ of and the CNOT gate)

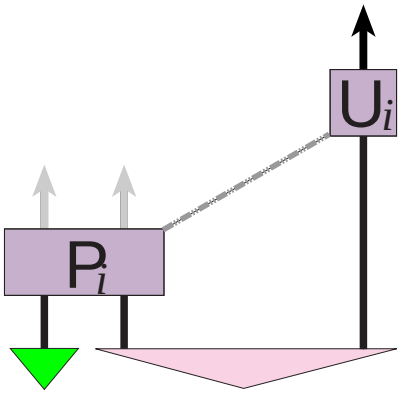
$$|\Psi_{\textit{Bell-basis}}\rangle = U_{\textit{Bell-basis}}|\Psi_{\textit{computational basis}}\rangle$$

and

$$\begin{aligned} \mathrm{P}_{Bell} &= |\Psi_{Bell}\rangle\langle\Psi_{Bell}| \\ &= |\Psi_{Bell}\rangle\,(|\Psi_{Bell}\rangle)^\dagger \\ &= U_{Bell}|\Psi_{comp.}\rangle\,(U_{Bell}|\Psi_{comp.}\rangle)^\dagger \\ &= U_{Bell}|\Psi_{comp.}\rangle\langle\Psi_{comp.}|U_{Bell}^\dagger \\ &= U_{Bell}\,\mathrm{P}_{comp.}\,U_{Bell}^\dagger\,. \end{aligned}$$

4.2 Teleportation and entanglement swapping

Quantum teleportation [11]. We are now ready to prove quantum teleportation, which we have already described above.



Denoting the state of the input qubit as

$$|\,\psi\rangle = c_0 \cdot |\,0\rangle + c_1 \cdot |\,1\rangle\,,$$

we assume that the second and third qubit are in the Bell-state. Then we perform a Bell-basis measurement on the first and the second qubit. If the outcome state of this measurement is the i-th Bell-basis vector, then we act with the transposed i-th Bell-matrix on the third qubit. Explicitly, we start with

$$\begin{aligned} &|\,\psi\rangle \otimes (|00\rangle + |11\rangle) \\ &= (c_0 \cdot |\,0\rangle + c_1 \cdot |\,1\rangle) \otimes (|00\rangle + |11\rangle) \\ &= c_0 \cdot (|000\rangle + |011\rangle) + c_1 \cdot (|100\rangle + |111\rangle) \\ &= |00\rangle \otimes (c_0 \cdot |0\rangle) + |01\rangle \otimes (c_0 \cdot |1\rangle) \\ &\quad + |10\rangle \otimes (c_1 \cdot |0\rangle) + |11\rangle \otimes (c_1 \cdot |1\rangle) \end{aligned}$$

There are four 'cases' corresponding to the possible outcome states of the Bell-basis measurement, so we need to concider

$$\Big((|\Psi\rangle\langle\Psi|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes (|00\rangle + |11\rangle)\Big)$$

for each Bell-basis vector $|\Psi\rangle$. E.g. for $|01\rangle - |10\rangle$ we obtain

$$(|01\rangle - |10\rangle) \otimes (c_0 \cdot |1\rangle - c_1 \cdot |0\rangle)$$

and applying $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ to this indeed yields

$$(|01\rangle - |10\rangle) \otimes |\psi\rangle \, .$$

The other three cases proceed analogously.

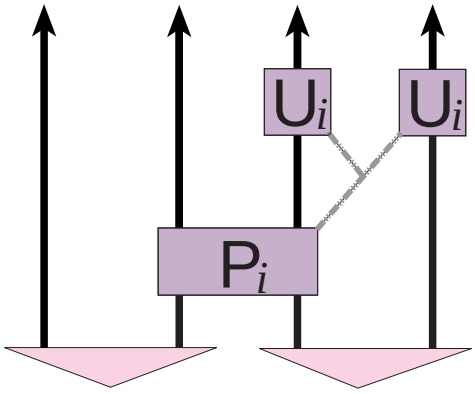
Exercise 4.1 i. Given that you only have the ability to perform measurements in the computational basis, Hadamard gates (see above), and CNOT-gates, i.e. gates with matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ,$$

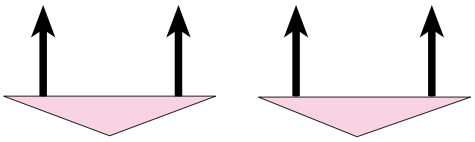
how would you perform a Bell-basis measurement? **ii.** When performing such a Bell-basis measurement involving a measurement in the computational basis, let boolean $b_1 \in \mathbb{B}$ be the outcome of measuring the first qubit, and let boolean $b_2 \in \mathbb{B}$ be the outcome of measuring the second qubit. Let $U_1 : \mathbb{B} \to \mathcal{U}$ and $U_2 : \mathbb{B} \to \mathcal{U}$ be functions with $\mathcal{U}$ the four element set consisting of the Pauli-matrices and the identity. Can you choose U_1 and U_2 such that $U_1(b_1) \circ U_2(b_2)$ provides the required correction of the third qubit in the teleportation protocol? (i.e. depending on the outcome b_2 of the measurement of the second qubit we perform unitary $U_2(b_2)$, and then, depending on the outcome b_1 of the measurement of the first qubit we perform unitary $U_1(b_1)$)

This implementation of the teleportation protocol is the one you'll find in most textbooks.

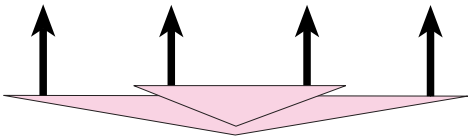
Entanglement swapping [12]. We start with four qubits, to which we refer as a, b, c, d, where a and b are in a Bell-state, and also c and d are in a Bell-state. Then we perform a Bell-basis measurement on c and d, and depending on the measurement outcome, analogously to what we did in the teleportation protocol, we apply the transpose of the corresponding Bell-matrix both to qubit c and d.



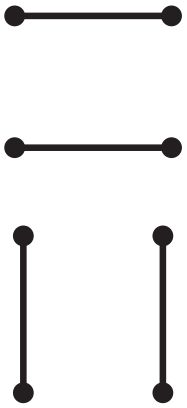
Now qubits a and d are in a Bell-state, and also qubits b and c are in a Bell-state. So, we 'swapped' the *entanglement* from the Bell-state entanglements:



to the different Bell-state entanglements:



Using a different geometry, we passed from



to

Exercise 4.2 i. Verify the entanglement swapping protocol. **ii.** If in the entanglement swapping protocol we start with two EPR-states rather than with two Bell-states, do we obtain the same result, i.e. do we still obtain two Bell-states, or, do we instead obtain two EPR-states, or something else? **iii.** Can you modify the measurement dependent 'unitary corrections' such that we do end up with two Bell-states, with two EPR-states, or instead with one EPR-state and one Bell-state?

When analyzing this protocol (and its proof) we can see that it includes two crucial components:

1. The measurement *destroys* correlations between pairs and *creates* new ones between other pairs.

2. The unitary 'corrections' using the Bell-matrices which guarantee that the resulting correlated pairs are indeed all in the Bell-state, and hence 'reverse' the non-deterministic differences due to the measurement.

A similar analysis also applies to the teleportation protocol.

We end this section with some additional comments on teleportation and entanglement swappping. While in the teleportation protocol we have

$$((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}) \, (|\, \psi\rangle \otimes |\Psi_{Bell}\rangle) = |\Psi_i\rangle \otimes U_i|\psi\rangle$$

(before applying $U_i^\dagger$ to the third qubit) we also have

$$(\langle\Psi_i| \otimes 1_{\mathcal{Q}}) \, (|\, \psi\rangle \otimes |\Psi_{Bell}\rangle) = U_i|\psi\rangle$$

so it seems that we actually only need the 'bra-part' of the projector $|\Psi\rangle\langle\Psi|$ to achieve teleportation. This seems to indicate that there are truly 'two components' to a projector, one, the ket-part, producing an outcome states, and one, the bra-part, being the 'action' which (for example) yields teleportation. Hence the Dirac-representation of a projector as a bra followed by a ket really reflects two distinct components in what such a projector actually does. Note that the 'unitary corrections' need not necessarily be the ones we used here, but our choice will again be motivated by conceptual analysis. For entanglement swapping we have

$$\begin{aligned}&(1_{\mathcal{Q}} \otimes (|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}) \left((|00\rangle + |11\rangle) \otimes (|00\rangle + |11\rangle)\right)\\ &= (1_{\mathcal{Q}\otimes\mathcal{Q}\otimes\mathcal{Q}} \otimes U_i)(|\,0\rangle\otimes|\Psi_i\rangle\otimes|\,0\rangle + |\,1\rangle\otimes|\Psi_i\rangle\otimes|\,1\rangle)\\ &= (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes U_i \otimes U_i)\\ &\qquad (|\,0\rangle\otimes|\Psi_{Bell}\rangle\otimes|\,0\rangle + |\,1\rangle\otimes|\Psi_{Bell}\rangle\otimes|\,1\rangle)\end{aligned}$$

Alternatively we can use the swap maps $\sigma : |ij\rangle \mapsto |ji\rangle$ to avoid explicit use of basis vectors yielding

$$\begin{aligned}&(1_{\mathcal{Q}} \otimes (|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}) \left(|\Psi_{Bell}\rangle \otimes |\Psi_{Bell}\rangle\right)\\ &\qquad = (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes U_i \otimes U_i)\, \mathcal{U}_{swap} \left(|\Psi_{Bell}\rangle \otimes |\Psi_{Bell}\rangle\right)\end{aligned}$$

where

$$\mathcal{U}_{swap} := (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes \sigma)(1_{\mathcal{Q}} \otimes \sigma \otimes 1_{\mathcal{Q}})\,.$$

— it should be clear that it becomes quite problematic to write things down nicely! When we consider only the 'bra-part' of the projector we obtain

$$(1_{\mathcal{Q}} \otimes \langle\Psi_i| \otimes 1_{\mathcal{Q}}) \left(|\Psi_{Bell}\rangle \otimes |\Psi_{Bell}\rangle\right) = (1_{\mathcal{Q}} \otimes U_i)|\Psi_i\rangle$$

i.e. the first and fourth qubit become entangled while no entanglement on the middle qubits have been created, any any entanglement on these qubits actually has been destroyed.

5 The structure of entanglement

We already saw that we can obtain the Bell-basis by acting with the Bell-matrices on the second qubit on a Bell-state. But actually we have much more.

5.1 Map-state duality and compositionality

Acting on the second qubit of a Bell-state with any linear operator f exactly yields the bipartite state encoded by the transposed to the matrix of that operator in our matricial definition of the tensor product, providing the linear map-bipartite state correspondence with a true (pseudo-)operational significance — we say pseudo-operational since in general f is not unitary and hence does not really correspond with a primitive physical operation. Indeed, when we apply $f^T := \begin{pmatrix} c_{00} & c_{10} \\ c_{01} & c_{11} \end{pmatrix}$ to the qubit basis we obtain

$$|\,0\rangle \mapsto c_{00}\cdot|\,0\rangle + c_{01}\cdot|\,1\rangle \quad \text{and} \quad |\,1\rangle \mapsto c_{10}\cdot|\,0\rangle + c_{11}\cdot|\,1\rangle$$

hence when we apply $1_{\mathcal{Q}} \otimes f^T$ we have

$$|\,00\rangle \mapsto c_{00}\cdot|\,00\rangle + c_{01}\cdot|\,01\rangle \;\text{ and }\; |\,11\rangle \mapsto c_{10}\cdot|\,10\rangle + c_{11}\cdot|\,11\rangle$$

so $1_{\mathcal{Q}} \otimes f^T$ applied to the Bell-state yields

$$|\,00\rangle + |\,11\rangle \mapsto c_{00}\cdot|\,00\rangle + c_{01}\cdot|\,01\rangle + c_{10}\cdot|\,10\rangle + c_{11}\cdot|\,11\rangle$$

that is

$$|\,00\rangle + |\,11\rangle \mapsto \sum_{ij} c_{ij}\cdot|\,ij\rangle\,.$$

Applying $f := \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}$ to the qubit basis yields

$$|\,0\rangle \mapsto c_{00}\cdot|\,0\rangle + c_{10}\cdot|\,1\rangle \quad \text{and} \quad |\,1\rangle \mapsto c_{01}\cdot|\,0\rangle + c_{11}\cdot|\,1\rangle$$

hence when we apply $f \otimes 1_{\mathcal{Q}}$ we have

$$|\,00\rangle \mapsto c_{00}\cdot|\,00\rangle + c_{10}\cdot|\,10\rangle \;\text{ and }\; |\,11\rangle \mapsto c_{01}\cdot|\,01\rangle + c_{11}\cdot|\,11\rangle$$

so $f \otimes 1_{\mathcal{Q}}$ applied to the Bell-state again yields

$$|\,00\rangle + |\,11\rangle \mapsto c_{00}\cdot|\,00\rangle + c_{01}\cdot|\,01\rangle + c_{10}\cdot|\,10\rangle + c_{11}\cdot|\,11\rangle.$$

The role which the Bell-state plays in this pseudo-operational correspondence is of course due to the fact that in the formal correspondence, the Bell-state corresponds to the identity. Hence a bipartite state Ψ_f represented by a matrix f^T can always be written down in two distinct pseudo-operational manners

$$(f^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle = |\Psi_f\rangle = (1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle.$$

We invite the reader to make a picture of this. Another useful property follows from

$$(f_1 \otimes f_2) \circ (g_1 \otimes g_2) = (f_1 \circ g_1) \otimes (f_2 \circ g_2)$$

as in Exercise 3.23. We have

$$\begin{aligned}(f \otimes \mathsf{id}) \circ (\mathsf{id} \otimes g) &= (f \circ \mathsf{id}) \otimes (\mathsf{id} \circ g)\\ &= (\mathsf{id} \circ f) \otimes (g \circ \mathsf{id})\\ &= (\mathsf{id} \otimes g) \circ (f \otimes \mathsf{id}).\end{aligned}$$

Again we invite the reader to make a picture of this.

Logic-gate teleportation [15]. So what happens if in the teleportation protocol we decide not to start with a Bell-state but with some other entangled state Ψ_f with matrix f^T? Explicitly, since for ordinary teleportation the input state was

$$|\,\psi\rangle \otimes |\Psi_{Bell}\rangle$$

now the input state is

$$\begin{aligned} |\,\psi\rangle \otimes \Psi_f &= |\,\psi\rangle \otimes ((1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle) \\ &= (1_{\mathcal{Q}}|\,\psi\rangle) \otimes ((1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle) \\ &= (1_{\mathcal{Q}} \otimes 1_{\mathcal{Q}} \otimes f)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big) \\ &= (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes f)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big) \end{aligned}$$

For ordinary teleportation the four 'cases' corresponding to the possible outcome states of the Bell-basis measurement we needed to concider were

$$\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big)$$

for each Bell-basis vector $|\Psi_i\rangle$, so now we have

$$\begin{aligned} &\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes f)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big) \\ &= (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes f)\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big). \end{aligned}$$

But we know from our study of teleportation that

$$(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes U_i^\dagger)\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big),$$

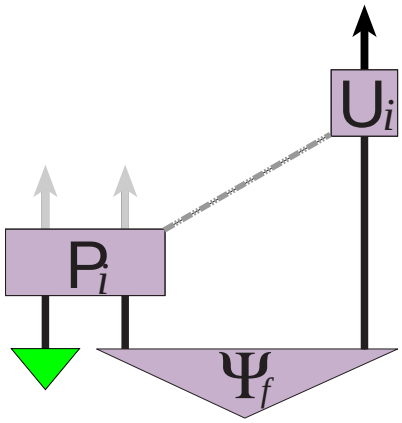
where U_i is the ith Bell-matrix, yields the state $|\psi\rangle$ for that third qubit, hence

$$\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big)$$

yields $U_i|\psi\rangle$ for the third qubit and hence

$$(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes f)\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big)$$

yields $fU_i|\psi\rangle$ for the third qubit. Assuming that f and U_i would commute, it then suffices to apply $U_i^\dagger$ to the last qubit to obtain $f|\psi\rangle$ i.e. we teleported the state and at the same time applied a (possibly unknown) operation f to this state $|\psi\rangle$.



This procedure is called *logic-gate teleportation* and turns out to be a universal quantum computational primitive. Of course, logic gates rarely commute, but in fact it suffices to find unitary operations $\{\tilde{U}_i\}_i$ such that

$$\tilde{U}_i \circ f = f \circ U_i$$

for each Bell-matrix, then applying $\tilde{U}_i^\dagger$ to the last qubit we do obtain $f|\psi\rangle$. In fact, with some appropriate acrobatics one can even go beyond the above considered situation [15].

Swapping from teleportation. In fact, entanglement swapping can be seen as a consequence of teleportation. For teleportation we have

$$\Big((|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,\psi\rangle \otimes |\Psi_{Bell}\rangle\Big) = |\Psi_i\rangle \otimes U_i|\psi\rangle$$

hence in particular, (i) considering $|\psi\rangle := |\,j\rangle$, (ii) applying $|\,j\rangle \otimes -$ to both sides of the equation, and (iii) using the $(\circ,\otimes)$-exchange properties, we obtain

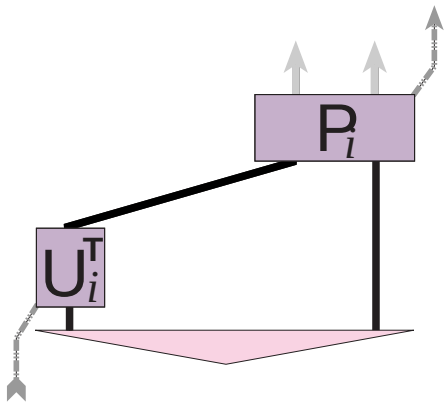
$$\begin{aligned} &\Big(1_{\mathcal{Q}} \otimes (|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,jj\rangle \otimes |\Psi_{Bell}\rangle\Big) \\ &\qquad = |\,j\rangle \otimes |\Psi_i\rangle \otimes U_i|\,j\rangle, \end{aligned}$$

hence adding for $j = 0, 1$ yields

$$\begin{aligned} &\sum_{j=0,1} \Big(1_{\mathcal{Q}} \otimes (|\Psi_i\rangle\langle\Psi_i|) \otimes 1_{\mathcal{Q}}\Big)\Big(|\,jj\rangle \otimes |\Psi_{Bell}\rangle\Big) \\ &\qquad = \sum_{j=0,1} |\,j\rangle \otimes |\Psi_i\rangle \otimes U_i|\,j\rangle \end{aligned}$$

and by linearity and $|\Psi_{Bell}\rangle = \sum_{j=0,1} |\,jj\rangle$ we indeed obtain entanglement swapping. Again, analyzing this calculation in a picture can be very instructive. *Compositionality* in computer science means breaking a big problem down in smaller, hopefully already known ones. Above we did this: we derived logic-gate teleportation and entanglement swapping from the teleportation protocol. On the other hand, the following protocol which is sometimes (wrongly?) referred to as some kind of converse to teleportation does have a nice conceptual derivation in terms of map-state duality.

Superdense coding [10]. It is our aim to use one quantum bit to communicate two classical bits i.e. in some way a kind of converse to quantum teleportation. We start with two qubits in a Bell-state, the two parties involved in the protocol each possessing one of the two qubits. Depending on which pair of classical bits we want to communicate one applies one of the four Bell-matrices to the first qubit which is then sent to the other party. The other party then performs a Bell-measurement on the pair of qubits and the outcome to that measurement reveals the encoded two bits.



To verify this protocol it suffices to recall that if we apply any of the Bell-matrices to the second qubit then we obtain the corresponding Bell-state, which we then will observe in the Bell-basis measurement which follows.

5.2 The logic of bipartite entanglement

Recalling that a bipartite state Ψ_f represented by a matrix f^T can always be written down either as

$$(f^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle = |\Psi_f\rangle = (1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle.$$

Using the property

$$(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$$

which straightforwardly follows from the 'pointwise' definition of $- \otimes -$ on linear operators, we can now take the adjoint of these resulting in

$$\langle\Psi_f| = \Big((f^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle\Big)^\dagger = \langle\Psi_{Bell}|(\bar{f} \otimes 1_{\mathcal{Q}})$$

with $\bar{f}$ the conjugate of f and

$$\langle\Psi_f| = \Big((1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle)\Big)^\dagger = \langle\Psi_{Bell}|(1_{\mathcal{Q}} \otimes f^\dagger).$$

All this results in four alternative representations for a bipartite projector, namely

$$\begin{aligned} \mathrm{P}_f = |\Psi_f\rangle\langle\Psi_f| &= (1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle\langle\Psi_{Bell}|(\bar{f} \otimes 1_{\mathcal{Q}}) \\ &= (1_{\mathcal{Q}} \otimes f)|\Psi_{Bell}\rangle\langle\Psi_{Bell}|(1_{\mathcal{Q}} \otimes f^\dagger) \\ &= (f^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle\langle\Psi_{Bell}|(\bar{f} \otimes 1_{\mathcal{Q}}) \\ &= (f^T \otimes 1_{\mathcal{Q}})|\Psi_{Bell}\rangle\langle\Psi_{Bell}|(1_{\mathcal{Q}} \otimes f^\dagger) \end{aligned}$$

And again, to get a feel for what this actually stands for, you want to represent these formulae in a picture. In fact, we can now formulate the structural crux behind all of the above discussed protocols.

$$\begin{aligned} &(\langle\Psi_f| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_g\rangle) \\ &= (((\langle\Psi_{Bell}|(1_{\mathcal{Q}} \otimes f^\dagger)) \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes ((1_{\mathcal{Q}} \otimes g)|\Psi_{Bell}\rangle)) \\ &= (\langle\Psi_{Bell}|\otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes f^\dagger \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes g)(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \\ &= (\langle\Psi_{Bell}|\otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes g)(1_{\mathcal{Q}} \otimes f^\dagger \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \\ &= g(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes f^\dagger \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \\ &= g(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{\bar{f}}\rangle) \\ &= g(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes \bar{f})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \\ &= (g \circ \bar{f})(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \end{aligned}$$

— following this calculation is quasi impossible without the support of a picture. This result is quite intriguing: since

$$(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle)$$

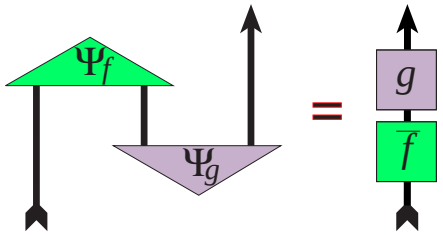
is just an instance of the teleportation protocol, namely conditioned on the fact that the measurement on the first two qubits yields the Bell-state, so

$$(\langle\Psi_{Bell}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) = 1_{\mathcal{Q}}\,,$$

hence we obtain

$$(\langle\Psi_f| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_g\rangle) = g \circ \bar{f}.$$

which can be represented in a picture as:

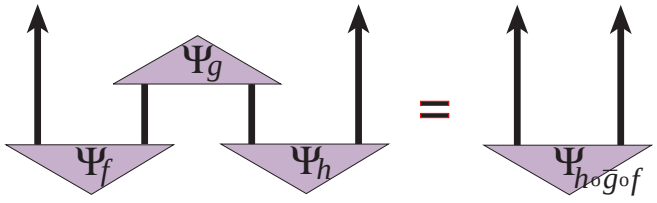


What is particularly interesting about this expression is the fact that while on the left, respecting compositional order, we first have an expression involving g and then one involving f, on the right we first have $\bar{f}$ and only then g. This seems as if there would be some weird reversal in the causal order!

Exercise 5.1 Show that we also have

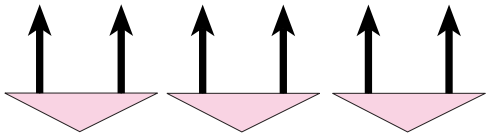
$$(1_{\mathcal{Q}} \otimes \langle\Psi_g| \otimes 1_{\mathcal{Q}})(|\Psi_f\rangle \otimes |\Psi_h\rangle) = \Psi_{h\circ\bar{g}\circ f}$$

which can be represented in a picture as:

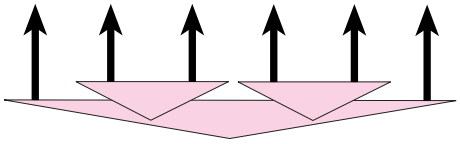


Exercise 5.2 [BA exam 2006] Four qubits are such that the first and second are in the joint state $a \cdot |00\rangle + b \cdot |11\rangle$ and the third and fourth are also in the joint state $a \cdot |00\rangle + b \cdot |11\rangle$. Then we perform a Bell-basis measurement on the second and third qubits. **i.** For each of the possible outcomes of the measurement, assuming $a, b \in]0, 1[$, what is the resulting state of the qubits? (Write these states in Dirac notation in the computational basis.) **ii.** Can you find values for a and b such that after applying some well-chosen unitary corrections the states of the qubits do not depend on the measurement outcome anymore? (Explicitly give these corrections.) **iii.** What is the probability for the second and the third qubit to end up in a Bell-state (when doing no corrections)?

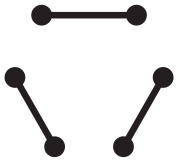
Exercise 5.3 We wish to design a slightly more sophisticated version of entanglement swapping involving six qubits, and yielding the passage from Bell-state entanglements



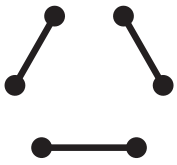
to Bell-state entanglements



that is, using a different geometry, from



to



Can you come up with a protocol which does that?

Correctness of the logic-gate teleportation protocol is in fact fully captured by the slightly elaborated variant[3]

$$\begin{aligned}
&(\mathrm{P}_f \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_g\rangle) \\
&\quad = (1_{\mathcal{Q}\otimes\mathcal{Q}} \otimes (g \circ \bar{f}))((|\Psi_f\rangle\langle\Psi_{Bell}|) \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_{Bell}\rangle) \\
&\quad = |\Psi_f\rangle \otimes (g \circ \bar{f})
\end{aligned}$$

which moreover provides a straightforward extension of the proof of logic-gate teleportation (and hence teleportation itself) beyond the case of qubits. We proceed in two steps:

1. First we show that we have

$$(\langle\Psi_{Bell}| \otimes 1_{\mathcal{H}})(1_{\mathcal{H}} \otimes |\Psi_{Bell}\rangle) = 1_{\mathcal{H}} \tag{2}$$

   where Ψ_{Bell} stands for the state corresponding to the identity matrix on $\mathcal{H}$ i.e.

$$\Psi_{Bell} := \sum_i |\, ii\rangle.$$

   for any dimension of the underlying Hilbert space.

2. Next, given a family of $n := \mathsf{dim}(\mathcal{H})$ unitary operators

$$\{U_i : \mathcal{H} \to \mathcal{H}\}_i$$

   which are such that the states

$$\{|\Psi_{U_i}\rangle : \mathbb{C} \to \mathcal{H}\}_i$$

   are mutually orthogonal, by the above we know that

$$(\langle\Psi_{U_i}| \otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}} \otimes |\Psi_f\rangle) = f \circ \bar{U}_i$$

   to which we can now apply the reasoning we made above for qubit-gate teleportation.

We can push all this a bit further within the domain of the absurd, as is illustrated by the following exercise.

Exercise 5.4 Assume we start with three qubits in a state

$$|\psi\rangle \otimes |\Psi_{f_1}\rangle\,.$$

To these we will apply five measurements, respectively including the projectors

$$\mathrm{P}_{f_2}\ \mathrm{P}_{f_3}\ \mathrm{P}_{f_4}\ \mathrm{P}_{f_5}\ \mathrm{P}_{f_6}$$

which we assume to have taken place in these five measurements, and

- first P_{f_2} takes place on the first two qubits,
- next P_{f_3} takes place on the last two qubits,
- next P_{f_4} takes place on the first two qubits,
- next P_{f_5} takes place on the last two qubits,
- finally P_{f_6} takes place on the first two qubits.

What is the state of the third qubit after all this?

[3] In course-notes v.7 there was a clumsy typo here.

5.3 Quantifying entanglement

Clearly, the Bell-state is more entangled than any (separable) tensor $|\psi\rangle\otimes|\phi\rangle$. But what exactly do we mean by 'being more entangled'? There are many different proposals for a precise conception of this, but consensus has yet to be reached in the research community. One way to attack this problem is to ask to which extent an entangled state enables typical quantum phenomena, such as non-local correlations and teleportation and how 'efficient' we can do that. For example, given a Bell-state we can do full teleportation, with as crux

$$(\langle\Psi_{Bell}|\otimes 1_{\mathcal{Q}})(1_{\mathcal{Q}}\otimes|\Psi_{Bell}\rangle) = 1_{\mathcal{Q}}\,.$$

On the other hand

$$\begin{aligned}(\langle\Psi_{Bell}|\otimes 1_{\mathcal{Q}})&(1_{\mathcal{Q}}\otimes|\psi\otimes\phi\rangle)\\ &= (\langle\Psi_{Bell}|(1_{\mathcal{Q}}\otimes|\psi\rangle))\otimes|\phi\rangle = |\phi\rangle\langle\bar{\psi}|\end{aligned}$$

i.e. the third qubit will end up in the state $|\phi\rangle$ independent of what the initial state of the first qubit was. In general we would end up somewhere between these two extremes of 'full teleportation' and 'no teleportation'. To understand this better we need some more linear algebra.

A linear operator $f:\mathcal{H}\to\mathcal{H}$ is *positive* iff

a. *it can be written as $f = g^\dagger\circ g$ for some other linear operator $g:\mathcal{H}\to\mathcal{H}'$.*

Each positive operator is obviously always self-adjoint, hence its eigenvalues will always be real. But we have more:

Proposition 5.5 *A linear operator $f:\mathcal{H}\to\mathcal{H}$ is positive if and only if, equivalently,*

b. *for all $\psi\in\mathcal{H}$ we have that $\langle\psi|f|\psi\rangle\in\mathbb{R}^+$;*

c. *f admits a self-adjoint square-root i.e. f can be decomposed as $f = g\circ g$ with g self-adjoint.*

Exercise 5.6 Prove Proposition 5.5 — you can proceed by respectively showing that $\mathbf{a}\Rightarrow\mathbf{b}$, $\mathbf{a}\&\mathbf{b}\Rightarrow\mathbf{c}$, $\mathbf{c}\Rightarrow\mathbf{a}$ where for proving $\mathbf{a}\&\mathbf{b}\Rightarrow\mathbf{c}$ you can use the fact that positive operators are self-adjoint and hence all diagonalize in some basis.

Next we show that each linear opertor factors into a unitary one and a positive one, something which is known as the *polar decomposition* of linear operators.

Theorem 5.7 *Each linear operator f can be written as*

$$f = U\circ g = g'\circ U$$

where U is unitary and g and g' are both positive, and in particular, g and g' are uniquely determined. Explicitly,

$$g = \sqrt{f^\dagger\circ f}\qquad\text{and}\qquad g' = \sqrt{f\circ f^\dagger}\,.$$

Since we know that bipartite states are in bijective correspondence with linear maps we can now use the above result to classify and quantify entanglement. For any bipartite state we have

$$|\Psi_f\rangle = (1_{\mathcal{H}}\otimes f)|\Psi_{Bell}\rangle = (1_{\mathcal{H}}\otimes(U\circ g))|\Psi_{Bell}\rangle.$$

Applying $U^\dagger$ to the last qubit we obtain

$$(1_{\mathcal{H}}\otimes U^\dagger)|\Psi_f\rangle = (1_{\mathcal{H}}\otimes g)|\Psi_{Bell}\rangle$$

i.e. we can always undo the effect of the unitary component, and we can do that in a reversible manner. Hence it is the (unique) positive component g which determines the 'degree of entanglement'. But since g is itself self-adjoint it admits diagonalization i.e. can be written as

$$g = U'\circ h\circ U'^\dagger$$

with h diagonal, so using Subsection 5.2 we obtain

$$(U'^T\otimes(U\circ U')^\dagger)|\Psi_f\rangle = (1_{\mathcal{H}}\otimes h)|\Psi_{Bell}\rangle$$

so we can again reversibly undo the effect of the unitaries, so the 'degree of entanglement' for bipartite states is now reduced to diagonal positive matrices i.e. a list of n positive reals. Again using map-state duality, but now in the converse direction, we straightforwardly obtain the following.

Proposition 5.8 *Each bipartite state $\Psi\in\mathcal{H}\otimes\mathcal{H}'$ admits a 'Schmidt decomposition' i.e. it can be written as*

$$\Psi = \sum_i r_i\cdot e_i\otimes e_i'\quad\text{with}\quad\{r_i\}_i\subseteq\mathbb{R}^+$$

for some well-chosen ONBs $\{e_i\}_i$ of $\mathcal{H}$ and $\{e_i'\}_i$ of $\mathcal{H}'$.

So for the qubit case the problem is reduced to comparing a pair of real numbers 'up to a real number' (cf. normalization) i.e. a one-dimensional problem. Extreme cases are $(1,0)$ (and equivalently $(0,1)$) which corresponds to pure tensors $\psi\otimes\phi$ while $(1,1)$ captures for example both the EPR-state and the Bell-state. In general, we will say that (a,b) for $a\geq b$ capture 'more entanglement' than (c,d) for $c\geq d$ iff

$$\frac{a}{b}\leq\frac{c}{d}.$$

For normalized coefficients $a^2+b^2 = c^2+d^2 = 1$ this boils down to $a\leq c$ or equivalently $d\leq b$, or again equivalently $a^2\leq c^2$ or $d^2\leq b^2$. So we can now assign a quantitative measure to bipartite qubit states by setting

$$\Psi\mapsto(a,b)\mapsto 2\cdot b^2$$

which is 0 whenever the state is disentangled and one whenever it is 'maximally entangled' e.g. in the Bell-state.

For higher dimensions the problem becomes more delicate, but a very natural candidate is the so-called *majorization order* [17] which is defined as follows. Let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ be decreasing lists of positive reals such that

$$\sum_i a_i = \sum_i b_i = 1\,.$$

We say that

$$(a_1, \ldots, a_n) \geq (b_1, \ldots, b_n)$$

if and only if

$$a_1 \leq b_1 \quad a_1{+}a_2 \leq b_1{+}b_2 \quad a_1{+}a_2{+}a_3 \leq b_1{+}b_2{+}b_3 \quad \text{etc.}$$

So given a bipartite state with $\{r_i\}_i$ we can consider the decreasing list $(r_{i_1}^2 \geq r_{i_2}^2 \geq \ldots \geq r_{i_n}^2)$, since for normalised states we have $\sum_i r_i^2 = 1$, and by comparing these lists we can compare their degree of entanglement. It can be shown that this order has an operational physical meaning [18]:

- One bipartite state Ψ can be converted into another bipartite state Φ by means of local operations and classical communication iff Ψ is above Φ w.r.t the above order i.e. iff Ψ is more entangled than Φ.

Exercise 5.9 **i.** Show that the majorization order is a partial order (= reflexive, anti-symmetric, transitive) on decreasing n-lists of positive reals, and hence induces a pre-ordering (= reflexive, transitive) on bipartite entangled states. **ii.** Draw a *Hasse-diagram* for the bipartite states

$$|00\rangle \qquad \frac{1}{\sqrt{2}}(|00\rangle + |22\rangle) \qquad \sqrt{\frac{3}{4}}|11\rangle + \sqrt{\frac{1}{4}}|22\rangle$$

$$\frac{1}{\sqrt{4}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle + \frac{1}{\sqrt{4}}|22\rangle \qquad \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)\,.$$

$$\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

That is, if a state is above another in the majorization order, then you depict it above it and connect them by a line, but if $a \leq b \leq c$ then you only draw a line between $a\&b$ and $b\&c$, but not between $a\&c$. **iii.** Characterise those states which are either maximal or minimal in the majorization order

---

All this is restricted to bipartite states. Understanding multipartite entanglement is still very much an open problem both with respect to quantitative and qualitative understanding.

5.4 Trace from Bell-states

For a linear operator $f : \mathcal{H} \to \mathcal{H}$ there exists a unique scalar

$$\mathrm{tr}(f) = \sum_i \langle i\,|f|\,i\rangle = \sum_i m_{ii}$$

which turns out to be independent of the choice of the basis. From Exercise 5.10i. it indeed folows that the basis-dependent expression can be rewritten as a basis-independent one.

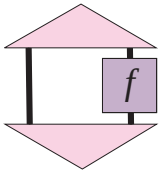
Exercise 5.10 **i.** Show for that

$$\mathrm{tr}(f) = \langle \Psi_{Bell}|(1_{\mathcal{H}} \otimes f)|\Psi_{Bell}\rangle$$

where we considered the generalized Bell-state

$$\Psi_{Bell} := \sum_i |\,ii\rangle \in \mathcal{H} \otimes \mathcal{H}\,,$$

(with norm $\sqrt{\mathsf{dim}(\mathcal{H})}$!), that is, in a picture,



ii. Show that

$$\langle \Psi_{Bell}|(1_{\mathcal{H}} \otimes (f \circ g))|\Psi_{Bell}\rangle = \langle \Psi_{Bell}|(1_{\mathcal{H}} \otimes (g \circ f))|\Psi_{Bell}\rangle\,.$$

Can you represent this equation in a picture? **iii.** For a linear operator $c : \mathbb{C} \to \mathbb{C}$ we clearly have that $\mathrm{tr}(c) = c$. Can you use part (ii) of this exercise to conclude from this that

$$\mathrm{tr}(\mathrm{P} \circ \rho_\psi) = \langle \psi|\mathrm{P}|\psi\rangle$$

where $\rho_\psi := |\psi\rangle\langle\psi|$?

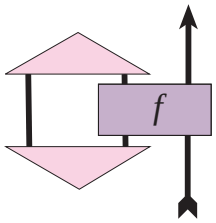
Exercise 5.10iii. exposes yet another variant on the rule for calculating probabilities. The advantage of this rule is that it extends to mixed states which we will discuss in the last section of this course. Another major advantage of the representation of the trace as in Exercise 5.10 is that it enables a straightforward definition of the *partial trace*, which, given a linear map $f : \mathcal{H} \otimes \mathcal{H}_1 \to \mathcal{H} \otimes \mathcal{H}_2$ is defined as a linear operator of type

$$\mathrm{tr}^{\mathcal{H}}_{\mathcal{H}_1,\mathcal{H}_2}(f) : \mathcal{H}_1 \to \mathcal{H}_2$$

with as prescription

$$\mathrm{tr}^{\mathcal{H}}_{\mathcal{H}_1,\mathcal{H}_2}(f) := (\langle \Psi_{Bell}| \otimes 1_{\mathcal{H}_2})(1_{\mathcal{H}} \otimes f)(|\Psi_{Bell}\rangle \otimes 1_{\mathcal{H}_1})$$

i.e., in a picture,



Exercise 5.11 Show that for

$$f : \mathcal{H} \to \mathcal{H}_1 \qquad \text{and} \qquad g : \mathcal{H} \to \mathcal{H}_2$$

we have

$$\mathrm{tr}^{\mathcal{H}}_{\mathcal{H}_1,\mathcal{H}_2}(|\Psi_g\rangle\langle\Psi_f|) = g \circ f^\dagger.$$

Can you represent this equation in a picture?

Exercise 5.12 [BA exam 2006] Let $\mathrm{P}_{GHZ} : \mathcal{Q}\otimes\mathcal{Q} \to \mathcal{Q}\otimes\mathcal{Q}$ be the projector which projects on the ray spanned by the state

$$\Psi_{GHZ} := |000\rangle + |111\rangle \in \mathcal{Q}\otimes\mathcal{Q}\otimes\mathcal{Q}$$

i. Compute

$$f := \mathsf{tr}^{\mathcal{Q}}_{\mathcal{Q}\otimes\mathcal{Q},\mathcal{Q}\otimes\mathcal{Q}}(\mathrm{P}_{GHZ}) : \mathcal{Q}\otimes\mathcal{Q} \to \mathcal{Q}\otimes\mathcal{Q}\,.$$

i.e. provide the matrix of f in the computational basis. **ii.** Assume now that in the above we substitute $\Psi_{GHZ} \in \mathcal{Q}\otimes\mathcal{Q}\otimes\mathcal{Q}$ by any arbitray state $\Psi \in \mathcal{Q}\otimes\mathcal{H}$. Prove positivity of

$$f := \mathsf{tr}^{\mathcal{Q}}_{\mathcal{H},\mathcal{H}}(\mathrm{P}) : \mathcal{H} \to \mathcal{H}$$

by showing that it always factors as $f = g^\dagger \circ g$ for some linear operator $g : \mathcal{H} \to \mathcal{Q}$.

6 Algorithms and gates

We now go over to the standard quantum computing stuff, a story about trying to design quantum algorithms by playing around with quantum logic gates. In order to compare the complexity of quantum algorithms with those of classical algorithms we need to have a particular model for quantum computing which straightforwardly compares to a model for classical computing. This model is the so-called *circuit model* or *gate-array model* basisd on the following three steps:

$$\text{preparation} \rightsquigarrow \text{logic gates} \rightsquigarrow \text{measurement}.$$

Comparing algorithms is now a matter of comparing the number of gates that need to be applied.

6.1 Special gates

The the 'controled not' or CNOT-gate is an important gate in the quantum computing literature is because of its entangling capabilities. Whenever the first qubit is in state $|\,0\rangle$ it doesn't alter the second qubit, but if the first qubit is in state $|\,0\rangle$ then the third Bell-matrix is applied to the second qubit, hence

$$|\,0i\,\rangle \mapsto |\,0i\,\rangle \qquad |\,1i\,\rangle \mapsto (|\,1\rangle \otimes U_2|\,i\,\rangle)$$

where U_2 is the third Bell-matrix or NOT-gate. The main application of this gate is preparation of Bell-states. We have

$$\mathsf{CNOT}\,((|\,0\rangle + |\,1\rangle)\otimes|\,0\rangle) = |\,00\rangle + |\,11\rangle$$

and more general we have

$$\mathsf{CNOT}\,((c_0\cdot|\,0\rangle + c_1\cdot|\,1\rangle)\otimes|\,0\rangle) = c_0\cdot|\,00\rangle + c_1\cdot|\,11\rangle\,.$$

Exercise 6.1 Define CNOT^σ as the gate obtained by exchanging the role played by the first and the second qubit. What is the effect of first applying CNOT, then CNOT^σ, and then again CNOT to a pair of qubits?

Exercise 6.2 Consider the following eight equations:

$$\mathsf{CNOT}\circ(U_i\otimes 1_{\mathcal{Q}}) = (\xi\otimes\xi')\circ\mathsf{CNOT}$$

$$\mathsf{CNOT}\circ(1_{\mathcal{Q}}\otimes U_i) = (\xi\otimes\xi')\circ\mathsf{CNOT}$$

where U_i can be any of the four Bell-matrices. Verify for each of these eight equations whether there exist operations ξ and ξ' such that it holds. What can you conclude from this for logic-gate teleportation of a CNOT-gate?

Proposition 6.3 *Each n-qubit gate can be obtained by compostion, tensor, 1-qubit gates and the* CNOT*-gate.*

Proof: See [17] pp.191–194. □

Other important gates are

$$H := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad S := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad T := \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$$

respectively called Hadamard-, phase-, and Tofoli-gate.

Proposition 6.4 *Each n-qubit gate can be approximated with arbitrary accuracy using compostion, tensor, the H-gate, the S-gate, the T-gate and the* CNOT*-gate — i.e. the difference in probabilities when performing measurements after applying the gates can be kept arbitrary small, explicitly*

$$\Big|\langle\Psi|U^\dagger\circ\mathrm{P}\circ U|\Psi\rangle - \langle\Psi|U^\dagger_{approx}\circ\mathrm{P}\circ U_{approx}|\Psi\rangle\Big|$$

can be kept arbitrary small for all Ψ and P.

Proof: See [17] pp.194–197. □

In the standard literature, rather than the Bell-matrices, the Pauli-matrices occur, that is

$$X := U_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$Y := i\,U_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$Z := U_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which of course do not have exactly same the 'nice' correspondence with the Bell-basis as the Bell-matrices have. But they are self-adjoint and give rise to a group, the so-called *Pauli group*. In the representation of a qubit as a sphere, they represent a 180° rotation respectively around the X-axis, the Y-axis and the Z-axis.

6.2 The Deutch-Jozsa algorithm

The aim of this 'pedagogical' example is to illustrate how easy it is try to exploit quantum parallelism, but how hard it is to actually succeed. Consider a Boolean function $f : \mathbb{B} \to \mathbb{B}$ where $\mathbb{B} := \{0, 1\}$. Since in general f is not injective, hence not reversible, we extend it in a way that after its execution we still 'remember' the argument, by setting

$$\tilde{f} : \mathbb{B} \times \mathbb{B} \to \mathbb{B} \times \mathbb{B} :: (i, j) \mapsto (i, j + f(i) \bmod 2)\,.$$

The function f can be recovered as $f(i) = \pi_2(\tilde{f}(i, 0))$ where $\pi_2 : \mathbb{B} \times \mathbb{B} \to \mathbb{B}$ is the second projection, while its argument can be retained as $i = \pi_1(\tilde{f}(i, 0))$. In fact, we produce pairs

$$\tilde{f}(-, 0) : \mathbb{B} \times \mathbb{B} \to \mathbb{B} :: i \mapsto (i, f(i))\,.$$

consiting of both the argument and the image. The second argument of $\tilde{f}$ enables it to be bijective: we can rewrite $\tilde{f}$ as

$$\tilde{f} :: \begin{cases} (i, j) \mapsto (i, j) & f(i) = 0 \\ (i, j) \mapsto (i, \sigma(j)) & f(i) = 1 \end{cases}$$

where $\sigma : \mathbb{B} \to \mathbb{B}$ permutes 0 and 1, from which bijectivity clearly follows either by considering cases, or by putting everything in one line as

$$\tilde{f} :: (i, j) \mapsto (i, \sigma^{f(i)}(j))\,.$$

But besides enabling bijectivity, the second input seems to be of no use at all, but as we will see further it will turn out to play a crucial role. Next we consider the by $\tilde{f}$ induced unitary permutation of basis vectors

$$U_f : \mathcal{Q} \otimes \mathcal{Q} \to \mathcal{Q} \otimes \mathcal{Q} :: |\,i\,j\rangle \mapsto |\,i\,(j + f(i) \bmod 2)\rangle\,.$$

We are now all set to exploit quantum parallelism:[4]

$$U_f(-, |\,0\rangle) :: |\,0\rangle + |\,1\rangle \mapsto |\,0\,f(0)\rangle + |\,1\,f(1)\rangle$$

i.e. with one execution of U_f we actually obtain the image under f both for 0 and 1. This idea moreover easily extends to functions $f : \mathbb{B}^n \to \mathbb{B}$, setting

$$U_f : \mathcal{Q}^{\otimes n} \otimes \mathcal{Q} \to \mathcal{Q}^{\otimes n} \otimes \mathcal{Q} :: |\,i\,j\rangle \mapsto |\,i\,(j + f(i) \bmod 2)\rangle\,,$$

and then considering

$$U_f(-, |\,0\rangle) :: \sum_{i \in \mathbb{B}^n} |\,i\rangle \mapsto \sum_{i \in \mathbb{B}^n} |\,i\,f(i)\rangle\,.$$

But there is a major problem! While our state encodes all possible argument/image pairs (or if you prefer I/O-pairs), this data is not accessible. Indeed, when we measure in the basis

$$\left\{|\,0\rangle\,, \ldots\,, |\,2^{n+1} - 1\rangle\right\}$$

then the outcome states with non-zero probability are

$$\left\{|\,i\,f(i)\rangle\right\}_{i \in \mathbb{B}^n}$$

but measurement

- will only expose one of these, and,
- destroy all the other components in the superposition.

So the only thing we achieved so far is introducing randomness concerning which outcome we are actually calculating, which we could as well have done by flipping a coin. And in fact, this problem can not be overcome i.e. we will never be able to extract all the desired data from the superposition state. Does this mean our endeavor was a waste of time?

Actually, as David Deutch and (later) Richard Jozsa showed, while we are not able to extract more than one argument/image pair from the superposition state it turns out that alternatively we can extract a particular *bit of data* from it, encoding a certain property of f, which classically would require knowledge of many argument/image pairs, in the worst case $\frac{n}{2} + 1$ pairs. For a function f which is either

- *constant* i.e. for all arguments the image is the same,
- *balanced* i.e. the number of 0- and 1-images are equal,

we will be able to verify with certainty which of the two it is. Let us consider

$$U_f(-, |\,0\rangle - |\,1\rangle) : \mathcal{Q}^n \to \mathcal{Q}^n \otimes \mathcal{Q}$$

[4] Note that a matrix representation is not very useful in this case.

(so we exploit the 2nd input!) that is[5]

$$|\,i\rangle \mapsto |\,i\,(0 + f(i) \bmod 2)\rangle - |\,i\,(1 + f(i) \bmod 2)\rangle$$

and since

$$\begin{cases} 0 + f(i) \bmod 2 = f(i) \\ 1 + f(i) \bmod 2 = 1 - f(i) \end{cases}$$

we obtain

$$|\,i\rangle \mapsto |\,i\rangle \otimes \Big(|\,f(i)\rangle - |1 - f(i)\rangle\Big)$$

yielding, for $f(i) = 0$ and $f(i) = 1$ respectively,

$$|\,i\rangle \mapsto |\,i\rangle \otimes (|0\rangle - |1\rangle) \quad \text{and} \quad |\,i\rangle \mapsto |\,i\rangle \otimes (|1\rangle - |0\rangle)\,,$$

in short,

$$|\,i\rangle \mapsto (-1)^{f(i)}|\,i\rangle \otimes (|0\rangle - |1\rangle)\,.$$

For the simple $f : \mathbb{B} \to \mathbb{B}$ case this becomes

$$|\,0\rangle + |\,1\rangle \mapsto \Big((-1)^{f(0)}|\,0\rangle + (-1)^{f(1)}|\,1\rangle\Big) \otimes (|0\rangle - |1\rangle)\,.$$

In this case constant means $f(0) = f(1)$ yielding

$$U_f\Big((|\,0\rangle + |\,1\rangle) \otimes (|\,0\rangle - |\,1\rangle)\Big) = \pm(|\,0\rangle + |\,1\rangle) \otimes (|\,0\rangle - |\,1\rangle)$$

while balanced means $f(0) \neq f(1)$ yielding

$$U_f\Big(|\,0\rangle + |\,1\rangle) \otimes (|\,0\rangle - |\,1\rangle)\Big) = \pm(|\,0\rangle - |\,1\rangle) \otimes (|\,0\rangle - |\,1\rangle).$$

Hence it suffices to measure the first qubit in the basis

$$\{|\,0\rangle + |\,1\rangle\,,\, |\,0\rangle - |\,1\rangle\}$$

to achieve our goal. Crucial is the fact that we obtained two mutually orthogonal states $|\,0\rangle + |\,1\rangle$ and $|\,0\rangle - |\,1\rangle$, each representing one of the two alternatives we wish to distinguish. In the general case we obtain

$$\sum_i |\,i\rangle \mapsto \Big(\sum_i (-1)^{f(i)}|\,i\rangle\Big) \otimes (|0\rangle - |1\rangle)$$

that is, for f constant,

$$U_f\Big(\Big(\sum_i |\,i\rangle\Big) \otimes (|\,0\rangle - |\,1\rangle)\Big) = \pm\Big(\sum_i |\,i\rangle\Big) \otimes (|\,0\rangle - |\,1\rangle)\,.$$

For a state $\sum_j \epsilon_j|\,j\rangle$ with $\epsilon_j := (-1)^{f(j)} \in \{0, 1\}$ we have that it is orthogonal to $\sum_i |\,i\rangle$ if and only if

$$0 = \sum_{ij} \epsilon_j \langle i \mid j\rangle = \sum_j \epsilon_j\,,$$

that is, exactly when f is balanced. Hence the case of a constant f can be distinguished by measurement from that of a balanced f, when performing a measurement on the first n qubits which includes $\sum_i |\,i\rangle$ as an outcome state:

- If f is constant we obtain $\sum_i |\,i\rangle$ with certainty.
- If f is balanced we cannot obtain $\sum_i |\,i\rangle$.

Since the number of required evaluations of f (encodes as U_f) is classically proportional to 2^n, while in the quantum case a single evaluation suffices, we have a true example of substantial algorithmic speed-up. On the other hand, the striking artificiality of this example confirms the hardness of the quantum informatic endeavor and the need for better/high-level methods to study quantum algorithms.

The text book version of the facts. Typically e.g. [17] one extends the presentation by making also the preparation of states explicit, as well as the generation of arbitrary measurements using some 'standard' logic gates and measurement in the computational basis. For creating superpositions within the computational basis we apply the so-called *Hadamard gate*

$$H := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

to first basis vector yielding

$$|\,0\rangle \overset{H}{\longmapsto} |\,0\rangle + |\,1\rangle$$

which we can generalize to

$$|\,0\rangle^{\otimes n} \overset{H^{\otimes n}}{\longmapsto} (|\,0\rangle + |\,1\rangle)^{\otimes n} = \sum_{i \in \mathbb{B}^n} |\,i\rangle\,.$$

To get the input which induces inference we set

$$|\,1\rangle \overset{H}{\longmapsto} |\,0\rangle - |\,1\rangle\,.$$

Measuring in the basis

$$\{|\,0\rangle + |\,1\rangle\,,\, |\,0\rangle - |\,1\rangle\}$$

is achieved by first applying a Hadamard gate and then measuring in the computational basis, while a measurement including $\sum_{i \in \mathbb{B}^n} |\,i\rangle$ as an outcome state is achieved by applying $H^{\otimes n}$ before a measurement in the computational basis.

6.3 Grover's search algorithm

This algorithm searches an unsorted databasis of size N in $O(\sqrt{N})$ time while classically this takes $O(N)$ time. Grover's is provably the fastest quantum algorithm that does search. Consider a boolean function

$$f : \mathbb{B}^n \to \mathbb{B}$$

which assigns 1 to the searched entry and 0 else. Search boils down to evaluating such a function for different input values

[5] Note that $|\,i + j\rangle \neq |\,i\rangle + |\,j\rangle$ e.g. $|\,1\rangle = |\,0 + 1\rangle \neq |\,0\rangle + |\,1\rangle$.

until the right one is found. Referring to the original title of Lov Grover's paper [14] we may consider

$$haystack \to \{needle, straw\}\,.$$

We can also calculate the inverse to f and then $f^{-1}(1)$ is the desired value. This is what Grover's algorithm does. The intuition behind this algorithm is purely geometrical. Let $\omega := f^{-1}(1)$ and set

$$U_f : \mathcal{Q}^{\otimes n} \to \mathcal{Q}^{\otimes n} :: \begin{cases} +|\omega\rangle \;\mapsto\; -|\omega\rangle \\ +|\,i\,\rangle \;\mapsto\; +|\,i\,\rangle \quad (i \neq \omega) \end{cases}$$

Let $\epsilon := \frac{1}{\sqrt{2^n}}(|0\rangle + |1\rangle)^{\otimes n}$ and set

$$U_\epsilon = 2|\epsilon\rangle\langle\epsilon| - 1_{\mathcal{Q}^{\otimes n}}$$

Consider the plane spanned by $|\omega\rangle$ and $|\epsilon\rangle$ and let $|\omega\rangle^\perp$ be the orthocomplement to $|\omega\rangle$ in this plane. When restricting to this plane one easily verifies that:

- $U_f \sim$ refection against $ray(|\omega\rangle^\perp)$
- $U_\epsilon \sim$ refection against $ray(|\epsilon\rangle)$
- $U_\epsilon \circ U_f \sim 2\theta$ rotation towards $|\omega\rangle$ for

$$sin\theta = \langle\epsilon|\omega\rangle = \frac{1}{\sqrt{2^n}}$$

Hence for r rounds with

$$\frac{\pi}{2} - \theta = 2\theta r \quad i.e. \quad r = \frac{\frac{\pi}{\theta} - 2}{4}$$

the vectors $|\omega\rangle$ and $(U_\epsilon \circ U_f)^n(|\epsilon\rangle)$ become 'almost' aligned. Hence the protocol:

- The input state is ϵ.
- Apply $U_\epsilon \circ U_f$ "the closest integer to $\frac{\frac{\pi}{\theta}-2}{4}$" times.
- Measure in computational basis.

With very high probability we will obtain the answer. The probability arises from the fact that $\frac{\frac{\pi}{\theta}-2}{4}$ is not an integer. This probability increases with N since θ decreases. For $N = 2^n >> 1$ we have $\theta \simeq sin\theta = \frac{1}{\sqrt{N}}$ so

$$r = \frac{\frac{\pi}{\theta} - 2}{4} \sim O(\sqrt{N})\,.$$

6.4 Shor's factoring algorithm

This was the first quantum algorithm, which, if it could be efficiently implemented on a quantum computer, would have unavoidably an important impact.

6.4.1 Period finding

Given a function

$$f : \mathbb{B}^n \to \mathbb{B}^m$$

we intend to find its period. To do this we will first produce a big state which encodes all input-output pairs i.e.

$$\sum_{\alpha\in\mathbb{B}^n} |\,\alpha\,f(\alpha)\rangle \in \mathbb{B}^n \otimes \mathbb{B}^m$$

and as we know from the previous section, we can produce such a state by a single unitary operation. The next step is a so-called *discrete quantum Fourier transform*

$$|\,\alpha\rangle \mapsto \sum_{\beta\in\mathbb{B}^n} e^{i2\pi\alpha\beta/2^n}\,|\,\beta\rangle$$

performed on the first n qubits, which can shown both to admit an inverse and to be unitary. Hence our state now becomes

$$\sum_{\alpha\in\mathbb{B}^n}\sum_{\beta\in\mathbb{B}^n} e^{i2\pi\alpha\beta/2^n}\cdot|\,\beta\,f(\alpha)\rangle \in \mathbb{B}^n \otimes \mathbb{B}^m.$$

Remarkably, we are now done and it suffices to measure the state of the first n qubits to 'with a high probability' find the period of the function! Indeed, suppose that f has period ω so we have for all α that

$$f(\alpha + \omega) = f(\alpha)\,.$$

Now fix a value for β and we are interested in the probability of obtaining this outcome in a measurement of the first qubit, so we are interested in the weight of the term

$$\sum_{\alpha\in\mathbb{B}^n} e^{i2\pi\alpha\beta/2^n}\cdot|\,\beta\,f(\alpha)\rangle = |\,\beta\rangle \otimes \sum_{\alpha\in\mathbb{B}^n} e^{i2\pi\alpha\beta/2^n}\cdot|\,f(\alpha)\rangle$$

which, for $|\,\beta\rangle$ normalized, is determined by the length of

$$\sum_{\alpha\in\mathbb{B}^n} e^{i2\pi\alpha\beta/2^n}\cdot|\,f(\alpha)\rangle.$$

Since f is periodic, several values of α, namely

$$\alpha\,, \quad \alpha+\omega\,, \quad \alpha+2\omega\,, \quad \alpha+3\omega\,, \quad \ldots$$

will contribute to the coefficient of same term $|\,f(\alpha)\rangle$. Since sometimes these coifficients are positive and sometimes they are negative, averagely they will more or less annihilate each other giving that term a very small weight. However, if

$$\beta = k\cdot\frac{2^n}{\omega} \qquad \text{for} \qquad k \in \mathbb{N}$$

then the components contributing to the same term have

$$e^{i2k\pi\alpha/\omega},\; e^{i2k\pi}e^{i2k\pi\alpha/\omega},\; e^{i2\cdot 2k\pi}e^{i2k\pi\alpha/\omega},\; e^{i3\cdot 2k\pi}e^{i2k\pi\alpha/\omega},\; \ldots$$

as coefficients, which by $e^{il\cdot 2k\pi} = 1$ for $l \in \mathbb{N}$ are in fact all the same. That is, for $\beta = k \cdot \frac{2^n}{\omega}$ we get *constructive interference*, hence will have a very high probability, and in principle we could then compute ω if it wasn't for the presence of k. To find the actual period with high probability it turns out that that it suffices to repeat the above roughly $\log\log\frac{\omega}{n}$ times.

The bad thing is that we only obtain a probabilistic result. The good thing is that with a growing number of qubits (i.e. n) the number of computations (i.e. for each of the possible input values $\alpha \in 2^n$ of f) and the corresponding stored data grows exponentially, and it turns out that for a 'small computer' with only 270 qubits we would actually have performed more computations and stored more results than the estimated number of particles in the universe.

6.4.2 Factoring and code-breaking

To quantify 'how long' it takes a certain algorithm to do a task, it is natural to ask how the needed time increases with the size of the input e.g. if we want to factor a large number N, how does the required time grows when we let N grow. Alternatively, we can measure the input in the number of required (qu)bits, i.e. $n = \log_2 N$, enabling a direct comparison between classical and quantum algorithms. On conventional computers the best known factoring algorithm runs in

$$\mathcal{O}\left(\exp\left(\left(\frac{64}{9}\right)^{\frac{1}{3}}(\ln N)^{\frac{1}{3}}(\ln\ln N)^{\frac{2}{3}}\right)\right)$$

time, so the required time grows exponentially with the number of digits $n = \log_2 N$ of the number we wish to factor. As an example, in 1994 a 129 digit number was successfully factored on 1600 workstations in parallel in a period of 8 months. The same computer setup would however require 800.000 years to factor a 250 digit number, and significantly longer than the age of the universe to factor a 1000 digit number. Of course, computers do become faster, but everytime the speed doubles 'we can just add a digit' to maintain the above mentioned 'absurdly long' required computation times.

The hardness of factoring large numbers is crucial for public key crypto-systems, e.g. those used in banks, of which the secrecy typically relies on the assumed difficulty (not to say impossibility) to factor a number of approximately 250 digits. But in 1994 Peter Shor of AT&T proposed an algorithm for factoring which on a gate-array-type quantum computer would run in

$$\mathcal{O}\Big((\log_2 N)^3\Big)$$

time, which means that it is only polynomial in the number of qubits, and factoring a 250 bit number turns out to only take a few billion computational steps — e.g. Microsoft's Xbox 360 game console does about a hundred billion in a second. The 'quantum part' of his algorithm is the above discussed period finding algorithm, while the 'classical part' of his algorithm relates period-finding to factoring, more specifically, to finding a single factor.

So we wish to factor a number $N_0 := N$. First choose a number $N_1 < N_0$ and we use Euclid's algorithm to look for common factors i.e. we perform a number of divisions N_i/N_{i+1} yielding quotients $q_1, q_2, \ldots q_x$ and remainders $N_2, N_3, \ldots, N_{x-1}, 0$ i.e.

$$\begin{aligned} N_0 &= q_1 \cdot N_1 + N_2 \\ N_1 &= q_2 \cdot N_2 + N_3 \\ &\vdots \\ N_{x-2} &= q_{x-1} \cdot N_{x-1} + N_x \\ N_{x-1} &= q_x \cdot N_x (+0) \end{aligned}$$

with the greatest common divisor being N_x. To see that N_x is indeed a common divisor of N_0 and N_1 it suffices to substitute the last equation in the one just before yielding $N_{x-2} = (\ldots q \ldots) \cdot N_x$, next substituting the two last in the one before that yielding $N_{x-3} = (\ldots q \ldots) \cdot N_x$, etc. So ultimately we will both obtain $N_1 = (\ldots q \ldots) \cdot N_x$ and $N_0 = (\ldots q \ldots) \cdot N_x$. If this procedure yield a non-trivial common devisor, we have our desired factor of $N = N_0$. On the other hand, if the greatest common devisor is 1 we have established that N_0 and N_1 are co-prime.

Consider the sequence of N_1-powers modulo N_0 i.e.

$$f(0),\ f(1),\ f(2),\ \ldots \qquad \text{for} \qquad f(\alpha) = N_1^\alpha \bmod N_0\,.$$

For ω the smallest number such that $N_1^\omega \bmod N_0 = 1$ this series becomes

$$1,\ N_1,\ N_1^2,\ \ldots,\ N_1^{\omega-1},\ 1,\ N_1,\ N_1^2,\ \ldots,\ N_1^{\omega-1}, \ldots$$

Indeed, if

$$N_1^\omega \bmod N_0 = 1 \quad \text{then} \quad N_1^\omega - k \cdot N_0 = 1$$

for some $k \in \mathbb{N}$, so we have $N_1^\omega = 1 + k \cdot N_0$ and hence

$$N_1^{\omega+l} = N_1^l \cdot (1 + k \cdot N_0) = N_1^l + (k \cdot N_1^l) \cdot N_0 = N_1^l \bmod N_0\,.$$

Thus, the sequence which we obtain is periodic, and we will *use a quantum computer to find this period* ω in the way it was described in the previous subsection. If ω is even we proceed as discussed below. If ω is odd we need to start over again and pick another number N_1 — this only happens in 50% of the cases so not to many runs are needed in general. Rewriting $N_1^\omega \bmod N_0 = 1$ as

$$(N_1^{\frac{\omega}{2}})^2 - 1 = 0 \bmod N_0$$

we obtain

$$(N_1^{\frac{\omega}{2}} - 1)(N_1^{\frac{\omega}{2}} + 1) = k \cdot N_0$$

for some $k \in \mathbb{N}$ i.e. we obtain two factors $(N_1^{\frac{\omega}{2}} - 1)$ and $(N_1^{\frac{\omega}{2}} + 1)$ of which the product is equal to some multiple of N_0, so if N_0 is not prime either $(N_1^{\frac{\omega}{2}} - 1)$ or $(N_1^{\frac{\omega}{2}} + 1)$ should have a non-trivial factor in common with it. We can extract this common factor using Euclid's algorithm. So from two numbers N_0 and N_1 with no common factor we have build a pair N_1' and N_1'' of which at least one has a non-trivial common factor with N_0 whenever the latter is not prime, and this procedure involved period-finding for which we can use a quantum computer. If this non-trivial common factor ends up being N_0 itself we start over again with a different N_1 — but there are more optimal ways to deal with this issue.

6.5 Quantum key distribution

So bye-bye to all our money in the banks? In fact, for the positive-minded who beliefs that someday quantum computers will be fact, there already is a solution available to the above stated problem, a solution somewhat ironically also provided by quantum informatics as it was the case for the problem itself, and in fact, you can already buy it online either at MagiQ:

`http://www.magiqtech.com/`

or at the Swiss basisd ID quantique:

`http://www.idquantique.com/`

The BB84 quantum key distribution protocol is a simple protocol which goes as follows. There are two parties, namely *Alice* and *Bob*. Alice prepares her qubits in either of the states

$$|\,0\rangle \qquad |\,1\rangle \qquad |\,0\rangle + |\,1\rangle \qquad |\,0\rangle - |\,1\rangle$$

in a randomly distributed manner, and sends them one by one to Bob. Bob, in order to know their content has to choose between measuring them in either of the basis

$$\{|\,0\rangle\,,\,|\,1\rangle\} \qquad \text{or} \qquad \{|\,0\rangle + |\,1\rangle\,,\,|\,0\rangle - |\,1\rangle\}\,.$$

Choosing 'the wrong' basis for measuring a qubit would of course destroy its data and yield an outcome unrelated to its actual initial state, and there is no way for Bob to know (without Alice's help) whether an outcome reflects the true initial state or not. So Bob can't do anything else but measuring the qubits in a randomly picked basis of his choice and records for each of the qubits which basis he used and what the outcome was. After all this is done, Alice tells Bob publicly for all of the qubits she has send whether they either belonged to

$$\{|\,0\rangle\,,\,|\,1\rangle\} \qquad \text{or} \qquad \{|\,0\rangle + |\,1\rangle\,,\,|\,0\rangle - |\,1\rangle\}\,,$$

and Bob at its turn tells Alice in which basis he measured them. They only retain the outcome-digits

$$0 \sim |\,0\rangle \quad 1 \sim |\,1\rangle \quad 0 \sim |\,0\rangle + |\,1\rangle \quad 1 \sim |\,0\rangle - |\,1\rangle$$

for those states of which their basis match, and the resulting string of bits is a secret shared only by the two of them.

The safety of this protocol follows from the fact that an *eavesdropper* intercepting the states cannot measure them without in 50% of the cases altering them, causing a very high number of mismatches in the key shared by Alice and Bob, which can easily be detected by them if they compare a small number of their key-digits. So the crucial quantum-feature which guarantees secrecy in this protocol is the fact that measurements in general alter the states, in the case of the protocol being in 50% of the cases.

A mild variant of this scheme is the Ekert91 quantum key distribution protocol in which Alice and Bob now share Bell-states. Each measures their Bell-states in a basis of their choice, and after having measured all their qubits they again compare their basis, only retaining those outcome-digits for which their basis match. In fact, the difference between BB84 and Ekert91 essentially boils down to interpreting the 'identity' $1_{\mathcal{Q}} : \mathcal{Q}_{Alice} \to \mathcal{Q}_{Bob}$ either as:

- effectively sending a qubit;
- sharing a Bell-pair.

7 Mixed states

Thus far we defined a state to be a ray in a Hilbert space. It turs out to be useful to have a more general notion of state which will enable us to describe:

1. Situations where there is a (probabilistic) lack of complete knowledge on the actual state of a quantum system.
2. Large statistical ensembles of quantum systems.
3. Subsystems of a bigger (entangled) quantum system.
4. Non-isolated (=open) quantum systems; decoherence.

All of these can be represented by the same mathematical object, namely a *density operator*, also referred to as a *mixed state*. As compared to a *pure state* which is a ray in a Hilbert space $\mathcal{H}$, a *density operator* is defined to be a linear operator $\rho : \mathcal{H} \to \mathcal{H}$ which is:

- positive (and hence self-adjoint);
- has trace equal to one.

We should now reformulate the axioms of quantum mechanics with respect to this new generalised notion of state.

Postulate 7.1 [extension to mixed states] The state of a quantum system is a density operator $\rho : \mathcal{H} \to \mathcal{H}$. Deterministic transformations correspond to $\rho \mapsto U \circ \rho \circ U^\dagger$ where $U : \mathcal{H} \to \mathcal{H}$ is unitary. Pure measurements are described by a set of projectors $\{\mathrm{P}_i : \mathcal{H} \to \mathcal{H}\}_i$ with $\sum_i \mathrm{P}_i = 1_\mathcal{H}$ [6] and they cause a state transition

$$\rho \mapsto \frac{\mathrm{P}_i \circ \rho \circ \mathrm{P}_i}{\mathrm{Tr}(\mathrm{P}_i \circ \rho)}$$

and this transition happens with probability $\mathrm{Tr}(\mathrm{P}_i \circ \rho)$.

Exercise 7.2 Show that when setting $\rho := |\psi\rangle\langle\psi|$ these postulates boil down to those for *pure states*.

These postulates can in fact be derived from those for pure states given the above heuristics. Consider a probabilistic lack of knowledge on a set of states i.e. we have a family of pure states $\{\psi_i\}_i$ together with respective probabilistic weights $\{\omega_i\}_i$ of the system actually being in that state. The probability for a certain outcome in a measurement is the weighted sum of the individual probabilities i.e.

$$\begin{aligned}\sum_j \omega_j \langle\psi_j|\mathrm{P}_i|\psi_j\rangle &= \sum_j \omega_j \mathrm{Tr}\left(\mathrm{P}_i \circ |\psi_j\rangle\langle\psi_j|\right) \\ &= \mathrm{Tr}\Big(\mathrm{P}_i \circ \Big(\sum_j \omega_j |\psi_j\rangle\langle\psi_j|\Big)\Big) \\ &= \mathrm{Tr}(\mathrm{P}_i \circ \rho)\,.\end{aligned}$$

We claim that $\sum_j \omega_j |\psi_j\rangle\langle\psi_j|$ is indeed a density matrix:

- For all $|\phi\rangle$ we have that $\langle\phi|\psi_j\rangle\langle\psi_j|\phi\rangle = |\langle\phi|\psi_j\rangle|^2$ is positive and hence so is

  $$\sum_j \omega_j \langle\phi|\psi_j\rangle\langle\psi_j|\phi\rangle = \langle\phi|\Big(\sum_j \omega_j |\psi_j\rangle\langle\psi_j|\Big)|\phi\rangle$$

  which establishes positivity.

- We moreover have

  $$\begin{aligned}\mathrm{Tr}\Big(\sum_j \omega_j |\psi_j\rangle\langle\psi_j|\Big) &= \sum_j \omega_j \mathrm{Tr}\Big(|\psi_j\rangle\langle\psi_j|\Big) \\ &= \sum_j \omega_j = 1\end{aligned}$$

  what completes the claim.

[6] Recall that this is the same thing as saying that these projectors arise as the spectral decomposition of a self-adjoint operator.

Conversely, all mixed states clearly arise in this way. The transition under unitaries and measurement in Postulate 7.1 is also induced by this heuristics. Exactly the same argument holds for statistical ensembles of pure states.

Exercise 7.3 Does a density matrix always uniquely represent a particular set of pure states together with a corresponding set of probabilistic weights? If not, characterise all pairs of states (ϕ_1, ϕ_2) and weights (ω_1, ω_2) with

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

as density matrix. Can you find an example of (ϕ_1, ϕ_2, ϕ_3) with $\omega_1 = \omega_2 = \omega_3 = \frac{1}{3}$ which again is described by the same density matrix? Repeat this question but now for

$$\begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{3}{4} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Now consider the situation that we have $|\Phi\rangle \in \mathcal{K} \otimes \mathcal{H}$. A measurement of $\mathcal{H}$ 'alone' is formally realised by considering $\{1_\mathcal{K} \otimes \mathrm{P}_i\}_i$ where $\{\mathrm{P}_i : \mathcal{H} \to \mathcal{H}\}_i$ a measurement of $\mathcal{H}$. Hence the respective probabilities are given by

$$\begin{aligned}\langle\Phi|(1_\mathcal{K} \otimes \mathrm{P}_i)|\Phi\rangle &= \langle\Psi_{Bell}|(1_\mathcal{K} \otimes (f^\dagger \circ \mathrm{P}_i \circ f))|\Psi_{Bell}\rangle \\ &= \mathrm{Tr}(f^\dagger \circ \mathrm{P}_i \circ f) \\ &= \mathrm{Tr}(\mathrm{P}_i \circ f \circ f^\dagger) \\ &= \mathrm{Tr}(\mathrm{P}_i \circ \rho)\end{aligned}$$

— this can again easily be seen in a picture. We claim that $f \circ f^\dagger$ is indeed a density matrix:

- It is positive by Proposition 5.5.

- We have

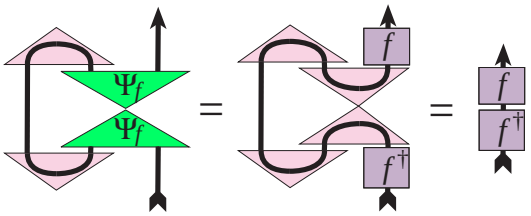
  $$\mathrm{Tr}(f \circ f^\dagger) = \langle\Phi|\Phi\rangle = 1$$

  whenever $|\Phi\rangle$ is normalised.

Again, all mixed states arise in this way by setting $f := \sqrt{\rho}$. The transition under unitaries and measurement in Postulate 7.1 is again induced by this heuristics. Exactly the same argument holds for open systems where $\mathcal{K}$ now describes the 'unknown' environment. Note also that we can extract ρ in a more direct manner from:

$$\rho = \mathrm{tr}^\mathcal{K}_{\mathcal{H},\mathcal{H}}(|\Phi\rangle\langle\Phi|)$$

i.e. we consider the density matrix of the pure state of the large system and trace out the component we are not interested in. In a picture we have:



The result of the converse derivation, i.e. finding $|\Phi\rangle$ given ρ, is called a *purification* of ρ.

Exercise 7.4 Are purifications always unique? If not, can you characterise all possible purifications of type $\mathcal{H}\otimes\mathcal{H}$ given a density matrix $\rho : \mathcal{H} \to \mathcal{H}$.

Comparing degrees of mixedness. We can use the majorisation order to compare the *mixedness* of mixed states. This can be done in two ways:

- Compare respective purifications in majorisation order.
- Diagonalise the density matrices and compare the lists of ordered diagonal elements in majorisation order.

These two are easily seen to be equivalent. Pure states are minimal elements in this order while there is a unique *top* element, namely the *maximally mixed state*

$$\perp_N := \frac{1}{N} \cdot 1_{\mathcal{H}} = \begin{pmatrix} \frac{1}{N} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{N} \end{pmatrix} ..$$

This state represents a situation where there is no information on the actual state of the system. Its classical counterpart is the uniform probability distribution since it behaves as such w.r.t. to any possible basis (cf. it's an identity up to a scalar).

Exercise 7.5 For a maximally mixed state $\perp_N : \mathcal{H} \to \mathcal{H}$ what are the possible purifications? What can you tell about their degree of entanglement?

Exercise 7.6 What is the mixed state describing a single qubit within the GHZ-state? What is the mixed state describing a pair of qubits within the GHZ-state?

Exercise 7.7 Assume we perform the teleportation protocol but Alice does not communicate the measurement outcome to Bob. Hence Bob cannot perform the unitary correction. What is the resulting state at Bob's end?

Decoherence. As a result of interaction between a quantum system and the environment the state of a system *decoheres*. Let $|\psi\rangle = \sum_i c_i|i\rangle$ be the state of the quantum system and let $|\epsilon\rangle$ be the state of the environment. We have

$$|\epsilon\rangle \otimes |\psi\rangle = |\epsilon\rangle \otimes \sum_i c_i|i\rangle \overset{decohere}{\rightsquigarrow} \sum_i c_i|\epsilon_i\rangle \otimes |i\rangle$$

$\Rightarrow$ the system becomes entangled with the environment

$\Rightarrow$ the system becomes part of a bigger system

$\Rightarrow$ the system's state becomes a mixed state

$\Rightarrow$ we loose lack of knowledge on the state

$\Rightarrow$ the state becomes less informative.

Ultimately the state could become the maximally mixed state i.e. the system could be in any state with equal probability. Obviously this is extremely bad for computational purposes. Decoherence is a major problem in the experimental realisation of quantum informatic devices. Avoiding decoherence requires *error-correction* [17].

8 Quantum logic and Gleason's theorem

In 1932 von Neumann published the current quantum mechanical formalism [2], but already in 1935 he wrote in a letter to G. Birkhoff [3]:

> "I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space no more." [von Neumann, 1935]

This resulted in a joint paper entitled 'The logic of quantum mechanics' [4], in which the order-theoretic structure which exists on the subspaces of a Hilbert space is taken to be the logical/structural feature which provides the key difference between classical and quantum behavior. The subspaces of a Hilbert space indeed come with order structure:

$$A \leq B \Leftrightarrow A \subseteq B$$

What makes this partial order special is the fact that not all subsets of $\mathcal{H}$ are subspaces, nor are the subsets of the set

$$\Sigma = \{|\,\psi\rangle \mid \psi \in \mathcal{H}\}$$

of all rays.

Birkhoff-von Neumann quantum logic. Classically, the algebra of (observable) properties that can be attributed to a physical or computational system consists of the subsets of the state space i.e. it is the powerset $\mathcal{P}(\Sigma)$. Any proposition on truth of an observable can be expressed in terms of such

a subset. Indeed, let $f : \Sigma \to \mathbb{R}$ be a physical observable e.g. energy values, color, location, speed, etc. Then

$$f^{-1}[E] \in \mathcal{P}(\Sigma)$$

expresses the property

"the value of f is within $E \subseteq \mathbb{R}$".

The basic idea in Birkhoff and von Neumann (1936) is the same thing. All statements of the form

"the value of H is within $E \subseteq \mathbb{R}$"

for a self-adjoint operator H can be represented by the projector P_E^H in the spectral decomposition, or the subspace A_E^H of its fixed points. Conversely, each subspace $A \subseteq \mathcal{H}$ is the eigenspace for some projector. Hence the algebra of observable properties of a quantum system seems to be

$$\mathcal{L}(\mathcal{H}) := \{A \subset \mathcal{H} \mid A \; is \; a \; subspace\}\,.$$

Since $\mathcal{L}(\mathcal{H}) \subseteq \mathcal{P}(\mathcal{H})$ set-theoretic inclusion provides a partial order and set-theoretic intersection provides greatest lower bounds. We also have least upper bounds, namely the linear span, which is definable as

$$\bigvee_i A_i = \bigcap \{A \in \mathcal{L}(\mathcal{H}) \mid \forall i : A_i \subseteq A\}\,,$$

and *orthocomlements*

$$A^\perp = \{\psi \in \mathcal{H} \mid \forall \phi \in A : \langle \psi | \phi \rangle = 0\}\,.$$

which satisfies

$$A \leq B \Rightarrow B^\perp \leq A^\perp \qquad A^{\perp\perp} = A$$

$$A \wedge A^\perp = \mathbf{0} \qquad A \vee A^\perp = \mathcal{H}$$

Hence one obtains a structure

$$\Big(\mathcal{L}(\mathcal{H})\,,\, \cap\,,\, \vee\,,\, {}^\perp\,,\, \mathcal{H}\,,\, \mathbf{0}\Big)$$

which very strongly resembles a classical Boolean algebra i.e. an ordinary propositional logic

$$(L\,,\, \wedge\,,\, \vee\,,\, \neg\,,\, tt\,,\, ff)$$

which comes with conjunction, disjunction, negation, true and false, and for which we can set

$$a \Rightarrow b \; := \neg a \vee b\,.$$

There is however a major difference since

$$A \; and \; (B \; or \; C) \;\neq\; (A \; and \; B) \; or \; (A \; and \; C)\,,$$

in $\mathcal{L}(\mathcal{H})$ e.g. in $\mathcal{Q}$ for $|+\rangle := |0\rangle + |1\rangle$ we have

$$|+\rangle \cap (|0\rangle \vee |1\rangle) = |+\rangle \neq \mathbf{0} = (|+\rangle \cap |0\rangle) \vee (|+\rangle \cap |1\rangle).$$

This example also shows that the suprema are *not disjunctive* due te fact that there are superposition states. This turns out to have dramatic consequences for its logical status e.g. there is notion of *deduction*, nor of *modus ponense* (etc.) i.e. we do not have

$$\frac{A \wedge B \vdash C}{A \vdash B \Rightarrow C} \qquad \frac{A \vdash B \;\wedge\; B \Rightarrow C}{A \vdash C}\,.$$

People have played around with with the Sasaki hook

$$A \Rightarrow B := A^\perp \vee (B \wedge A)$$

but this requires replacing $\wedge$ by the non-commutative and non-associative binary connective

$$\mathrm{P}_A(B) := A \wedge (B \vee A^\perp)\,,$$

which cannot be interpreted as a conjunction — an interpretation for this which does seem to make sense is one in terms of a Hoare-style weakest precondition/ strongest postcondition semantics. Therefore it makes more sense to call this an *algebra*. There are some important results of this setting.

Theorem 8.1 (Gleason 1957) *Let $dim(\mathcal{H}) \geq 3$. For each state $|\psi\rangle$ there exists exactly one function*

$$\omega_{|\psi\rangle} : \mathcal{L}(\mathcal{H}) \to [0,1]$$

such that

$$\omega_{|\psi\rangle}(A) = 1 \Leftrightarrow |\psi\rangle \subseteq A$$

$$\sum_i \omega_{|\psi\rangle}(A_i) = \omega_{|\psi\rangle}(\bigvee_i A_i)$$

where we assume all $\{A_i\}_i$ to be always mutually orthogonal.

Existence is not a surprise since we know such a map, namely

$$\omega_{|\psi\rangle} :: A \mapsto \langle \psi | \mathrm{P}_A | \psi \rangle\,.$$

The fact that this is the only one is quite astonishing. It means that in some manner the quantum probability structure is already encoded in the partial order. Another fascinating result is the Mackey-Piron-Solèr Theorem [25], which gives the exact assumptions one needs to impose on an order-theoretic structure for it to be the lattice of 'closed' subspaces of an infinite-dimensional Hilbert space — key to the proof is the fundamental theorem of projective geometry. Some of the axioms admit reasonable physical interpretations. The following is an import consequence of the previous result.

Theorem 8.2 (Gleason 1957) *Let $dim(\mathcal{H}) \geq 3$. The collection of all functions*

$$\omega : \mathcal{L}(\mathcal{H}) \to [0,1]$$

which are such that

$$\sum_i \omega(A_i) = 1$$

where $\{A_i\}_i$ are mutually orthogonal and $\bigvee_i A_i = \mathcal{H}$, are in bijective correspondence with density matrices. In particular

$$\omega_\rho :: A \mapsto \mathsf{Tr}(\rho\, \mathrm{P}_A)\,.$$

This result shows that all possible probabilistic behaviors, for whatever kind of system including both classical and quantum uncertainty, can always be described by a density operator.

However, this order-theoretic approach dramatically failed in capturing the Hilbert space tensor product at any algebraic or conceptual level. That is, any attempt to axiomatize it essentially requires the full-blown Hilbert space structure. So we have good reasons to abandon the order-theoretic approach, but is there any other candidate mathematical structure which would enable us to capture the tensor product at a higher level of abstraction/conceptualization.

9 Mixed operations

This is a very involved topic which has its roots in C^*-algebra [23]. We will mainly mention the key concepts and results. Similarly as density operators describe a more general notion of state there are more general notions of operation (contra unitaries) and of measurement.

Denote the set of all mixed states of type $\mathcal{H} \to \mathcal{H}$ as $\Sigma(\mathcal{H})$. The type of a generalised notion of operation is

$$\mathcal{F} : \Sigma(\mathcal{H}) \to \Sigma(\mathcal{H})$$

and from the heuristics of mixed states it follows that:

- It is a *convex-linear* map i.e.

$$\mathcal{F}(\sum_i p_i \rho_i) = \sum_i p_i \mathcal{F}(\rho_i)\,.$$

  Hence, thinking of density matrices as vectors in $\mathcal{H} \otimes \mathcal{H}$, these generalised operations have a matrix of type $\mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$. If $\mathcal{H}$ is n-dimensional then ρ is an $n \times n$-matrix and $\mathcal{F}$ has an $(n \times n) \times (n \times n)$ matrix i.e. 4×4 for a qubit and 16×16 for a pair of qubits i.e. 256 entries.

- It preserves positivity in a 'strong sense' i.e. whenever $\rho \in \Sigma(\mathcal{H} \otimes \mathcal{K})$ then $(\mathcal{F} \otimes 1_{\Sigma(\mathcal{K})})(\rho)$ has to be positive too. This condition is called *complete positivity*.

Complete positivity is essential. Consider the *transpose*

$$(-)^T : \mathcal{F} : \Sigma(\mathcal{Q}) \to \Sigma(\mathcal{Q}) :: \rho \mapsto \rho^T\,.$$

If $\rho \in \Sigma(\mathcal{Q})$ then indeed $\rho^T \in \Sigma(\mathcal{Q})$ but when applying $((-)^T \otimes 1_{\Sigma(\mathcal{Q})})$ to the Bell-state the result is not positive.

Exercise 9.1 Attempt to calculate this using matrices just to convince yourself how much fun this is.

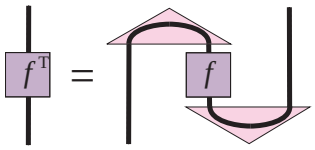
We will show this in a more structural manner. For

$$f = \sum_{ij} f_{ij} |i\rangle\langle j| : \mathcal{H}_a \to \mathcal{H}_b$$

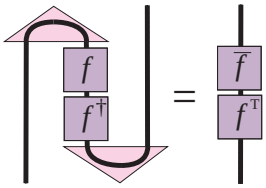
we have

$$\begin{aligned}\Big(\sum_i \langle ii| \otimes 1_{\mathcal{H}_a}\Big) \circ (1_{\mathcal{H}_b} \otimes f \otimes 1_{\mathcal{H}_a}) \circ \Big(1_{\mathcal{H}_b} \otimes \sum_j |jj\rangle\Big) \\ = \sum_{ij} \langle i|f|j\rangle |j\rangle\langle i| \\ = \sum_{ij} f_{ij} |j\rangle\langle i| = \sum_{ij} f_{ji} |i\rangle\langle j| = f^T\end{aligned}$$

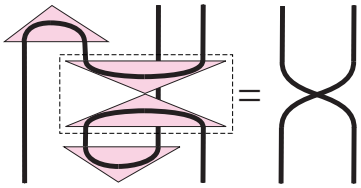
that is, in a picture



This can also be purely graphically derived by 'sliding the f-box'. Although we have $(f \circ f^\dagger)^T = \overline{f} \circ f^T$, graphically



i.e. the transposed of something positive is always positive, $((-)^T \otimes 1_{\Sigma(\mathcal{Q})})(|\Psi_{Bell}\rangle\langle\Psi_{Bell}|)$ is not positive:



Indeed, we have

$$\begin{aligned}&(\langle 01| - \langle 10|)\, \sigma_\otimes\, (|01\rangle - |10\rangle) \\ &= (\langle 10| - \langle 01|)\, (|01\rangle - |10\rangle) = -2 \leq 0\,.\end{aligned}$$

Exercise 9.2 What does the following expression stands for:

$$\Big(\sum_i \langle ii| \otimes 1_{\mathcal{H}_a}\Big) \circ (1_{\mathcal{H}_b} \otimes f^\dagger \otimes 1_{\mathcal{H}_a}) \circ \Big(1_{\mathcal{H}_b} \otimes \sum_j |jj\rangle\Big)$$

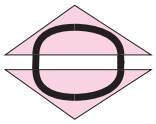
Exercise 9.3 What is the result of applying the operation

$$1_{\mathcal{H}\to\mathcal{H}} \otimes (-)^T : (\mathcal{H} \to \mathcal{H}) \to (\mathcal{H} \to \mathcal{H})$$

to the CNOT-gate?

The operation $1_{\mathcal{H}\to\mathcal{H}} \otimes (-)^T$ is called the *partial transpose*.

Exercise 9.4 How much is:



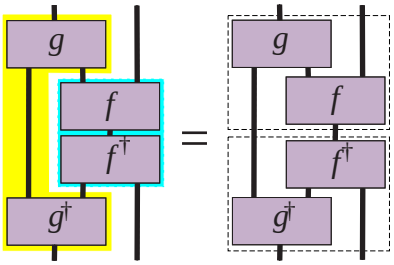
Examples of completely positive maps are:

$$\mathcal{F} : \Sigma(\mathcal{H}) \to \Sigma(\mathcal{H}) :: \rho \mapsto g \circ (1_{\mathcal{K}} \otimes \rho) \circ g^\dagger$$

where $g : \mathcal{K} \otimes \mathcal{H} \to \mathcal{H}$ is linear. Since

$$\begin{aligned}
&(g \otimes 1_{\mathcal{K}'}) \circ (1_{\mathcal{K}} \otimes (f \circ f^\dagger)) \circ (g \otimes 1_{\mathcal{K}'})^\dagger \\
&\quad = (g \otimes 1_{\mathcal{K}'}) \circ (1_{\mathcal{K}} \otimes f) \circ (1_{\mathcal{K}} \otimes f^\dagger) \circ (g \otimes 1_{\mathcal{K}'})^\dagger \\
&\quad = ((g \otimes 1_{\mathcal{K}'}) \circ (1_{\mathcal{K}} \otimes f)) \circ ((g \otimes 1_{\mathcal{K}'}) \circ (1_{\mathcal{K}} \otimes f))^\dagger
\end{aligned}$$

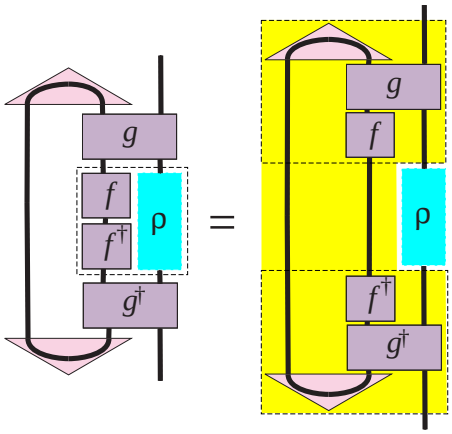
we have that $\mathcal{F}$ is indeed completely positive. This again can be immediately seen in a picture:



where the yellow part represent the completely positive map while the blue bit is the mixed state of the extended system in which it acts. By the following theorem it follows that each completely positive map arises like this.

Theorem 9.5 (Stinespring) *Each completely positive map can be realised by applying some linear operator to an extended system and then tracing this extended system out.*

This indeed assures that all completely positive maps admit the shape $g \circ (1_{\mathcal{K}} \otimes -) \circ g^\dagger$ since we have:



where ρ is the state of the system, $f \circ f^\dagger$ is the state of the ancilla, g is the linear map applied to system + ancilla, and the yellow part is the resulting completely positive map of the form $h \circ (1_{\mathcal{K}} \otimes \rho) \circ h^\dagger$ with h and $h^\dagger$ the dotted parts. Moreover, since it is known that each linear map can itself be realised by applying a unitary to a larger system it follows that each completely positive map can be realised by applying a unitary operator to an extended system. Hence it provides operational meaning for completely positive maps in terms of describing 'open system dynamics', or equivalently, 'being part of a bigger operation'.

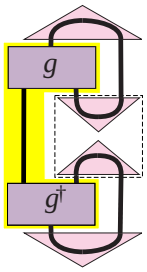
In *quantum information theory* one defines *channels* to be completely positive maps which are, depending on the author, either trace-preserving or trace-decreasing. The aim is to have an as 'clean' as possible channel i.e. with minimal noise. Noise is a result of interaction with the environment so we indeed need our generalised operations to model it. A quantity which measures the level of noise due to the environment is the *channel fidelity*:

$$\langle \Psi_{Bell}| \left(\mathcal{F} \otimes 1_{\mathcal{H}\to\mathcal{H}}\right) \left(|\Psi_{Bell}\rangle\langle\Psi_{Bell}|\right) |\Psi_{Bell}\rangle \in \mathbb{R}^+$$

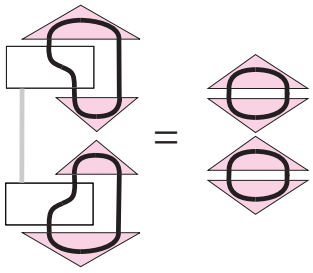
Note that this expression is quite hard to read.

Exercise 9.6 Show that channel fidelity is a positive number.

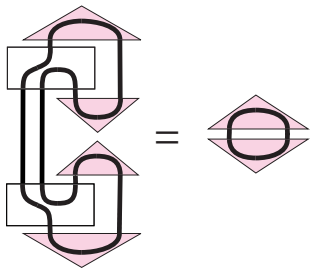
In a picture it becomes:



where the yellow part is $\mathcal{F} = g \circ (1_{\mathcal{K}} \otimes -) \circ g^\dagger$ and the dotted part the Bell-state. Via the trace we 'compare' the channel's input with its output. When the channel is perfectly clean, i.e. it is an identity, then we have:



If it is utterly dirty on the other hand, i.e. nothing is transmitted and it only produces noise, we have:



There are many other similar quantities.

Theorem 9.7 (Krauss) *For each completely positive map $\mathcal{F}$ there exist linear maps $\{E_k : \mathcal{H} \to \mathcal{H}\}_k$ such that*

$$\mathcal{F} : \Sigma(\mathcal{H}) \to \Sigma(\mathcal{H}) :: \rho \mapsto \sum_k E_k \circ \rho \circ E_k^\dagger \,.$$

A *generalised measurement* is described by a family of linear maps $\{E_k : \mathcal{H} \to \mathcal{H}\}_k$ and induces a transition

$$\rho \mapsto E_k \circ \rho \circ E_k^\dagger$$

with probability $\mathrm{Tr}(E_k \circ \rho \circ E_k^\dagger) = \mathrm{Tr}(E_k^\dagger \circ E_k \circ \rho)$. A *POVM* is described by a family of positive linear operators $\{M_k : \mathcal{H} \to \mathcal{H}\}_k$ and produces outcomes with probability $\mathrm{Tr}(M_k \circ \rho)$. No change of state is associated with a POVM.

Theorem 9.8 (Naimark) *Each POVM can be realised by applying some projective measurement to an extended system and then tracing this extended system out.*

10 More on tensors

We already proved the no-cloning theorem. The following exercises indicate how resource sensitivity is imposed by the tensor product, and how it is 'entangled' with entanglement.

Exercise 10.1 [BA exam 2006] **i.** We modify the setting in which we proved the no-cloning theorem by introducing a state of the "environment" i.e. we start with a system in state $\psi \otimes \phi_0 \otimes \Phi_0$ and we wish to find a unitary operator U_{clone} such that for any "unknown" state ψ we have

$$U_{clone}(\psi \otimes \phi_0 \otimes \Phi_0) = \psi \otimes \psi \otimes \Phi_\psi$$

where Φ_ψ is allowed to depend on ψ while ϕ_0 and Φ_0 are constants, but you are allowed to choose ϕ_0, Φ_0 and Φ_ψ. Does such a unitary operator U_{clone} exist? **ii.** Why is the operation

$$\delta : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{H} :: |\, ij\rangle \mapsto |\, ii\rangle$$

not a copying operation? More specifically, characterize the initial states for which δ does not map the input state $|\psi\rangle \otimes |\phi_0\rangle$ to the output state $|\psi\rangle \otimes |\psi\rangle$. **iii.** We reverse this cloning setting into some "pseudo-deleting" setting by starting with a system in state $\psi \otimes \psi \otimes \Phi_0$ while U_{delete} should be such that for any ψ we have

$$U_{delete}(\psi \otimes \psi \otimes \Phi_0) = \psi \otimes \phi_0 \otimes \Phi_\psi \,,$$

where again Φ_ψ can depend on ψ and you are allowed to choose ϕ_0, Φ_0 and Φ_ψ. Does such a unitary operator U_{delete} exist? Warning: this setting is not the setting of the so-called no-deleting theorem, since we could still be able to re-extract a copy of ψ out of Φ_ψ, something which is explicitly forbidden in the notion of deleting which gives rise to the no-deleting theorem [9].

The following exercise is closely related to ex.10.1 **ii** but puts it in a 'less decorated' formal context.

Exercise 10.2 [MSc mini-project 2006] We study maps between a Hilbert space $\mathcal{H}$ and $\mathcal{H} \otimes \mathcal{H}$. Is the map

$$\gamma : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H} :: |\, \psi\rangle \mapsto |\, \psi\rangle \otimes |\, \psi\rangle$$

linear, and for which arguments does the linear map

$$\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H} :: |\, i\rangle \mapsto |\, ii\rangle$$

coincide with γ?

The following exercise investigates the structure of the pure tensors within the whole tensor product of two qubits.

Exercise 10.3 [MSc mini-project 2006] **i.** Provide a state $\Phi \in \mathcal{Q} \otimes \mathcal{Q}$ which is orthogonal to the subspace of $\mathcal{Q} \otimes \mathcal{Q}$ spanned by the set

$$\mathsf{PT}_{\mathcal{Q}} := \left\{ |\psi\rangle \otimes |\psi\rangle \;\middle|\; |\psi\rangle \in \mathcal{Q} \right\}.$$

Next, find a basis $\mathcal{B}_{\mathcal{Q}}$ for the subspace of $\mathcal{Q} \otimes \mathcal{Q}$ spanned by PT, and prove that this is indeed a basis. (Hint: You can pick Φ and $\mathcal{B}_{\mathcal{Q}}$ from the union of the elements of the computational basis and the Bell-basis.) Is PT itself a subspace of $\mathcal{Q} \otimes \mathcal{Q}$? **ii.** We will now generalize the above to a Hilbert space $\mathcal{H}$ of arbitrary dimension. Provide a basis $\mathcal{B}_{\mathcal{H}}$ for the subspace of $\mathcal{H} \otimes \mathcal{H}$ spanned by the set

$$\mathsf{PT}_{\mathcal{H}} := \left\{ |\psi\rangle \otimes |\psi\rangle \;\middle|\; |\psi\rangle \in \mathcal{H} \right\}.$$

Next, find a set $\mathcal{A}$ of vectors which are all orthogonal to $\mathsf{PT}_{\mathcal{H}}$ and such that $\mathcal{A} \cup \mathcal{B}_{\mathcal{H}}$ is a basis of $\mathcal{H} \otimes \mathcal{H}$. (Hint: Are there some obvious ways in which you can embed each of the vectors included in the Bell-basis for $\mathcal{Q} \otimes \mathcal{Q}$ in the higher-dimensional space $\mathcal{H} \otimes \mathcal{H}$?)

Compositionality with relations. We now present a surprising analogue to the compositionality result which enabled us to derive several quantum protocols. Let's move from the world of Hilbert space to sets, ...

$$Hilbert\ space \rightsquigarrow set$$

$$linear\ map \rightsquigarrow relation$$

$$tensor\ product \rightsquigarrow cartesian\ product$$

$$function\ composition \rightsquigarrow relational\ composition$$

Where do we end up in this way: something classical-like or something quantum-like? The obvious guess would be to think that we obtain something classical-like. However!

A relation 'from X to Y' is some $R \subseteq X \times Y$. Hence it is a family of pairs $(x, y) \in R$. When we write xRy for $(x, y) \in R$ we can write:

$$R = \{(x, y) \mid x \in X, y \in Y, xRy\}\,.$$

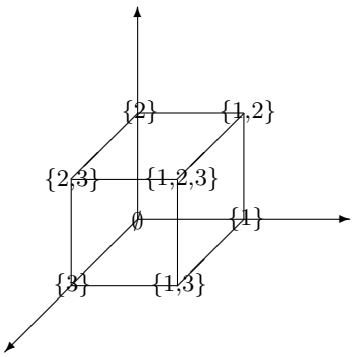
We have an analogue to scalars:

$$\mathcal{H} \otimes \mathbb{C} \simeq \mathcal{H} \ \rightsquigarrow \ X \times \{*\} \simeq X$$

and hence an analogue to state:

$$|\psi\rangle : \mathbb{C} \to \mathcal{H} \ \rightsquigarrow \ r : \{*\} \to X$$

so the 'relational states' are the subsets of X, i.e. elements of the powerset $\mathcal{P}(X)$. A notion of *Superposition* emerges:



Map state duality is a tautology:

$$\mathcal{P}(X \times Y) = \mathcal{P}(X \times Y)$$

where

- lefthandside are the states of the space $X \times Y$
- righthandside are all relations 'from X to Y'

For two relations $R_1 \subseteq X_1 \times Y_1$ and $R_2 \subseteq X_2 \times Y_2$ their *parallel composition* (cf. tensor) is defined as

$$(x_1, x_2)(R_1 \text{“}\times\text{”} R_2)(y_1, y_2) \ \Leftrightarrow \ x_1 R_1 y_1 \ \&\ x_2 R_2 y_2$$

what boils down to exactly being the cartesian product i.e. "pairs of pairs" $R_1 \times R_2 =$

$$\{((x_1, y_1), (x_2, y_2)) \in (X_1 \times Y_1) \times (X_2 \times Y_2) \mid (x_1, y_1) \in R_1\,,\ (x_2, y_2) \in R_2\}$$

For two relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$ their *sequential composition* is defined as

$$x(R; S)z'' \ \Leftrightarrow \ \exists y \in Y : xRy \ \&\ ySz$$

that is

$$R; S := \{(x, z) \in X \times Z \mid y \in Y : xRy\,,\ ySz\}\,.$$

We can now calculate what a ket-bra is. Since $r : \{*\} \to X$ is some $(\{*\} \times A) \subseteq \{*\} \times X$ we obtain

$$r^c; r = (A \times \{*\}); (\{*\} \times A) = A \times A$$

where r^c stands for the *converse* relation so

$$adjoint \rightsquigarrow converse.$$

By the trivial map state duality bipartite ket-bras are:

$$\mathrm{P}_R = R \times R \subseteq (X \times Y) \times (X \times Y)$$

in analogy with $\mathrm{P}_f := |\Psi_f\rangle\langle\Psi_f|$. We wish to study

$$(1_X \times \mathrm{P}_S); (\mathrm{P}_R \times 1_Z)$$

where $R \subseteq X \times Y$ and $S \subseteq Y \times Z$. We have

$$(x, y, z)(1_X \times \mathrm{P}_S)(x', y', z')$$

if and only if

$$x = x' \qquad y'Sz' \qquad (ySz)$$

and we have

$$(x', y', z')(\mathrm{P}_R \times 1_Z)(x'', y'', z'')$$

if and only if

$$z' = z'' \qquad x'Ry' \qquad (x''Ry'')$$

so

$$(x, y, z)(1_X \times \mathrm{P}_S))(\mathrm{P}_R \times 1_Z)(x'', y'', z'')$$

if there exists (x', y', z') such that $x = x'Ry'Sz' = z''$ i.e.

$$x(R; S)z$$

which exactly yields our well-known compositionality result for Hilbert spaces with corresponding 'seemingly backward in time information flow'.

Exercise 10.4 i. What is a Bell-state in the world of relations? **ii.** Show that in the world of relations there also exists an analogue to Exercise 5.1. **iii.** Can you come up with a notion of full and partial trace for the world of relations? **iv.** Using this notion of trace prove an analogue to Exercise 5.11 on partial traces in the world of relations.

There are two ways to pass from Hilbert space to the purely quantitative level of relations namely via

- matrix calculus where the Boolean semiring $\mathbb{B}$ (i.e. $1 + 1 = 1$ and not $\mathbb{Z}_2$) replaces the complex field $\mathbb{C}$.
- Dirac notation where the singleton set $\{*\}$ replaces $\mathbb{C}$ in the definition of *bra*s and *ket*s.

Can you figure out why in one case we have a two-element set while in the other case we have a singleton set — this is quite a hard one to solve!

What do sets, relations and cartesian product on-the-one-hand and Hilbert spaces, linear maps and tensor product on-the-other-hand have in common? Their *category-theoretic structure* are very similar! The fact that we picked sets wasn't important, but it was crucial to pick relations and not functions, and to pick cartesian product and not disjoint union.

Exercise 10.5 If we would have picked functions and cartesian product would there be some notion of superposition?

11 Semantics for quantum informatics

Goals of this semantics:

- We want a formal counterpart to the picture language which seems to help understanding entanglement.
- We want to do quantum theory in a more conceptual manner than just playing around with matrices.
- We hope to produce a better quantum formalism than the one around, a search which is by now a more that 70 year old lasting endeavor.
- Revealing a structure which lives on Hilbert spaces, but rather at the level of linear maps than at the level of the vectors in the Hilbert space since category theory reveals some structural connections which ordinary mathematical structures don't.

11.1 Symmetric monoidal categories

This Section is available as §1,2,3,4,5 in a paper at

`web.comlab.ox.ac.uk/internal/courses/materials05-06/qcs/Cats.pdf`

In the symmetric monoidal category $(\mathbf{Rel}, \times)$ with

- sets as objects;
- relations as morphisms;
- the cartesian product as its monoidal bifunctor with the singleton set $\{*\}$ as its unit;

and the symmetric monoidal category $(\mathbf{FdHilb}, \otimes)$ with

- finite dimensional Hilbert spaces as objects;
- linear maps as morphisms;
- the tensor product as its monoidal bifunctor with the one-dimensional Hilbert space $\mathbb{C}$ as its unit;

there is no natural diagonal i.e. no family of morphisms

$$\{\Delta_A : A \to A \otimes A\}_A$$

such that for all $f : A \to B$ we have commutation of

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle \Delta_A}\downarrow & & \downarrow{\scriptstyle \Delta_B} \\ A \otimes A & \xrightarrow[f \otimes f]{} & B \otimes B \end{array}$$

This while there are obvious candidates for such an operation. In $\mathbf{FdHilb}$ a first candidate would be

$$\mathcal{H} \to \mathcal{H} \times \mathcal{H} :: |\psi\rangle \mapsto |\psi\rangle \otimes |\psi\rangle$$

but this map fails to be linear. A second candidate would be

$$\Delta_{\{|\, i\rangle\}_i} : \mathcal{H} \to \mathcal{H} \times \mathcal{H} :: |\, i\rangle \mapsto |\, ii\rangle\,.$$

The obvious candidate for $\mathbf{Rel}$ is the 'function'

$$\Delta_X : X \to X \times X :: x \mapsto (x, x)$$

which written as a relation is

$$R_\Delta := \{(x, x) \mid x \in X\} \subseteq X \times X\,.$$

But it turns out that neither $\Delta_{\{|\, i\rangle\}_i}$ nor R_Δ make the required diagrams commute. Counterexamples can be found in §7 of the above mentioned reference. Analysis of these counterexamples for the **FdHilb**-case clearly show that the existence of superposition states is what causes the problem. In the case of relation the corresponding problem is that relations can be *multi-valued*, and this is why for the symmetric monoidal category $(\mathbf{Set}, \times)$ with

- sets as objects;
- functions as morphisms;
- the cartesian product as its monoidal bifunctor with the singleton set $\{*\}$ as its unit;

the function Δ_X does provide a natural diagonal. Note that this absence of a natural diagonal in **FdHilb** is strongly connected to the No-Cloning theorem. Moreover, while No-Cloning required unitarity in its proof, here we don't rely on unitarity whatsoever.

11.2 Naturality implies basis-independence

Generally speaking, given expression $\Lambda(-,\ldots,-)$ and $\Xi(-,\ldots,-)$ — where we only use brackets and the monoidal bifunctor — naturality of a family of morphisms

$$\{\xi_{A_1,\ldots,A_n} : \Lambda(A_1,\ldots,A_n) \to \Xi(A_1,\ldots,A_n)\}_{A_1,\ldots,A_n}$$

means that we have commutation of:

$$\begin{array}{ccc} \Lambda(A_1,\ldots,A_n) & \xrightarrow{\Lambda(f_1,\ldots,f_n)} & \Lambda(B_1,\ldots,B_n) \\ \downarrow{\scriptstyle \xi_{A_1,\ldots,A_n}} & & \downarrow{\scriptstyle \xi_{B_1,\ldots,B_n}} \\ \Xi(A_1,\ldots,A_n) & \xrightarrow[\Xi(f_1,\ldots,f_n)]{} & \Xi(B_1,\ldots,B_n) \end{array}$$

for all $f_i : A_i \to B_i$. Restricting this requirement to all the f_i being unitarity for the case of Hilbert spaces we obtain

$$\begin{array}{ccc} \Lambda(\mathcal{H}_1,\ldots,\mathcal{H}_n) & \xrightarrow{\Lambda(U_1,\ldots,U_n)} & \Lambda(\mathcal{H}'_1,\ldots,\mathcal{H}'_n) \\ \downarrow{\scriptstyle \xi_{\mathcal{H}_1,\ldots,\mathcal{H}_n}} & & \downarrow{\scriptstyle \xi_{\mathcal{H}'_1,\ldots,\mathcal{H}'_n}} \\ \Xi(\mathcal{H}_1,\ldots,\mathcal{H}_n) & \xrightarrow[\Xi(U_1,\ldots,U_n)]{} & \Xi(\mathcal{H}'_1,\ldots,\mathcal{H}'_n) \end{array}$$

We can take each $U_i : \mathcal{H}_i \to \mathcal{H}'_i$ to be a change of basis. Hence we obtain that naturality implies basis-independency. This then also immediately makes clear why $\Delta_{\{|\,i\rangle\}_i} : \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ couldn't have been natural.

Exercise 11.1 Show that $\Delta_{\{|\,i\rangle\}_i}$ indeed depends on the choice of basis i.e. there exists a basis $\{|\,e_i\rangle\}_i$ such that

$$\Delta_{\{|\,e_i\rangle\}_i} : \mathcal{H} \to \mathcal{H} \times \mathcal{H} :: |e_i\rangle \mapsto |e_i\rangle \otimes |e_i\rangle$$

does not coincide with $\Delta_{\{|\,i\rangle\}_i}$. Can you find the necessary condition on $\{|\,e_i\rangle\}_i$ such that $\Delta_{\{|\,i\rangle\}_i}$ and $\Delta_{\{|\,e_i\rangle\}_i}$ coincide?

Another important example of a map which depends on the choice of basis is

$$|\,e_i\rangle \mapsto \langle e_i\,|$$

since we have

$$c_i \cdot |\,e_i\rangle = |\,c_i \cdot e_i\rangle \mapsto \langle c_i \cdot e_i\,| = \bar{c}_i \cdot \langle e_i\,|\,.$$

That is, $\{|\,e_i\rangle \mapsto \langle e_i\,|\}_i$ and $\{|\,c_i \cdot e_i\rangle \mapsto \langle c_i \cdot e_i\,|\}_i$ define non-equal linear maps whenever one of the c_i has a non-trivial imaginary part. Hence it follows that also

$$|\,e_i\rangle \otimes |\,e_j\rangle \;\mapsto\; |\,e_j\rangle\langle e_i\,|$$

depends on the choice of basis since

$$c_i \cdot |e_i\rangle \otimes |e_j\rangle = |\,c_i \cdot e_i\rangle \otimes |e_j\rangle \mapsto |e_j\rangle\langle c_i \cdot e_i| = \bar{c}_i \cdot |e_j\rangle\langle e_i|.$$

Let $\mathcal{H} \multimap \mathcal{H}'$ be all linear maps $f : \mathcal{H} \to \mathcal{H}'$. The assignment

$$\mathcal{H} \otimes \mathcal{H}' \to \mathcal{H} \multimap \mathcal{H}' :: |\,i\,j\rangle \mapsto |\,i\rangle\langle j\,|$$

is exactly how we related the elements of the tensor product to linear maps i.e. *map-state duality*. Hence it follows that this correspondence is not basis-independent, hence not natural.

The solution to this problem consists of defining for each Hilbert space $\mathcal{H}$ the *conjugate Hilbert space* $\mathcal{H}^*$ which has the same vectors as $\mathcal{H}$ but in which each complex number is interpreted as the complex conjugate of a complex number for $\mathcal{H}$, that is, for $c \in \mathbb{C}$ and $\psi, \phi \in \mathcal{H}^*$ we have

$$c \bullet_{\mathcal{H}^*} \psi := \bar{c} \bullet_{\mathcal{H}} \psi \qquad \langle \psi\,|\,\phi\rangle_{\mathcal{H}^*} := \overline{\langle \psi\,|\,\phi\rangle}_{\mathcal{H}} = \langle \phi\,|\,\psi\rangle_{\mathcal{H}}\,.$$

It turns out that now we do have a natural isomorphism

$$\mathcal{H} \otimes \mathcal{H}' \simeq \mathcal{H}^* \multimap \mathcal{H}'\,.$$

Moreover

$$\mathcal{H}^{**} = \mathcal{H}\,.$$

This seems to indicate a logical interpretation where we take * to be *negation*, $\multimap$ to be *implication* and $\otimes$ to be 'coinciding' *conjunction-disjunction* since

$$(\mathcal{H} \otimes \mathcal{H}')^* \simeq \mathcal{H}^* \otimes \mathcal{H}'^*\,.$$

In fact, we have a 'degenerate' so-called Linear Logic in which conjunction and disjunction coincide.

11.3 †-compact categories

We define a †-*compact category* as a symmetric monoidal category which comes with the following additional data

- involution **dual** $A \mapsto A^*$;
- contravariant $\otimes$-involution[7] **adjoint** $f_{A\to B} \mapsto f^\dagger_{B\to A}$;

[7] I.e. the 'abstract' adjoint has to preserve the tensor-structure.

- **Bell-states** $\eta_A : \mathrm{I} \to A^* \otimes A$;

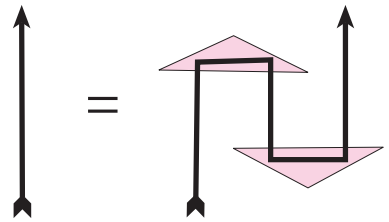
for which we have $\eta_{A^*} = \sigma_{A^*,A} \circ \eta_A$ and

$$\begin{array}{ccccc} A & \xleftarrow{\simeq} & \mathrm{I} \otimes A & \xleftarrow{\eta_{A^*}^\dagger \otimes 1_A} & (A \otimes A^*) \otimes A \\ {\scriptstyle 1_A}\uparrow & & & & \uparrow{\scriptstyle \simeq} \\ A & \xrightarrow[\simeq]{} & A \otimes \mathrm{I} & \xrightarrow[1_A \otimes \eta_A]{} & A \otimes (A^* \otimes A) \end{array}$$

As discussed in

`http://arxiv.org/abs/quant-ph/0510032`

such a category also admits a corresponding graphical calculus in which the above commuting diagram admits a *yanking* interpretation:



An example are Hilbert spaces, with their conjugates, adjoint of linear maps, and "natural" Bell-states. First define *conjugate kets* from kets by making the passage

$$|\,\psi\rangle : \mathbb{C} \to \mathcal{H} \quad \rightsquigarrow \quad |\,\psi\rangle_* : \mathbb{C} \to \mathcal{H}^*$$

for which we in particular have

$$|\,c \cdot \psi\rangle \otimes |\,\psi\rangle_* = c \cdot (|\,\psi\rangle \otimes |\,\psi\rangle_*) = |\,\psi\rangle \otimes |\,\bar{c} \cdot \psi\rangle_*$$

and the "natural" Bell-state are

$$\mathbb{C} \to \mathcal{H}^* \otimes \mathcal{H} :: 1 \mapsto \sum_i |\,i\rangle_* \otimes |\,i\rangle\,.$$

Note that this Bell-state arises through the "natural" map-state duality from the identity on $\mathcal{H}$. Also, each linear map $f : \mathcal{H} \to \mathcal{H}'$ induces a conjugate one

$$f_* : \mathcal{H}^* \to \mathcal{H}'^* :: |\,\psi\rangle_* \mapsto |\,f(\psi)\rangle_*$$

for which for

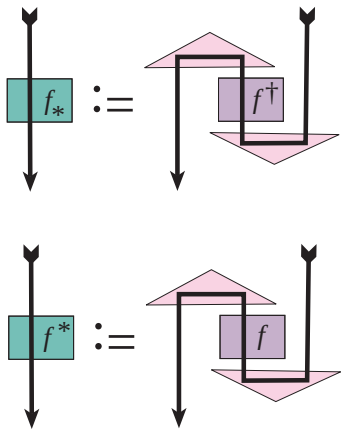
$$|\,\psi\rangle_* = \Big|\sum_i c_i |\,i\rangle\Big\rangle_* = \sum_i \bar{c}_i |\,i\rangle_*$$

we have

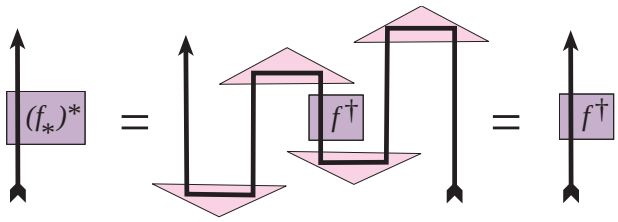
$$|\,\psi\rangle_* \mapsto |\,f(\psi)\rangle_* = \Big|\sum_{ij} c_i\, m_{ij} |\,j\rangle\Big\rangle_* = \sum_{ij} \bar{c}_i\, \overline{m}_{ij} |\,j\rangle_*$$

i.e. we obtain the conjugate matrix. Each such conjugate map admits an adjoint, which we will denote by $f_* : \mathcal{H}'^* \to \mathcal{H}^*$.

Both f_* and f^* admit a nice purely diagrammatic characterization, respectively:



We can abstractly show that these are factors of the adjoint:



and analogous we can prove that $(f^*)_* = f^\dagger$.

11.4 Classical uncertainty and open systems

But there is in fact an even more stunning presence of complex numbers at the level of abstract picture calculi of the kind we consider for which we refer to §4.d of the above mentioned reference. Exactly the same reasoning as done there holds for

$$f \mapsto f \otimes f_* \qquad \text{and} \qquad f \mapsto \ulcorner f \urcorner \otimes (\ulcorner f \urcorner)^\dagger$$

as for

$$f \mapsto f \otimes f^\dagger\,.$$

In fact, all of these represent the passage from pure states to *density matrices*, for kets respectively

$$|\psi\rangle \mapsto |\psi\rangle \otimes |\psi\rangle_* \qquad \text{and} \qquad |\psi\rangle \mapsto |\psi\rangle\langle\psi|$$

mixed states arise as $\mathbb{R}^+$-weighted sums of these

$$\sum_i r_i \cdot |\psi_i\rangle \otimes |\psi_i\rangle_* \qquad \text{and} \qquad \sum_i r_i \cdot |\psi_i\rangle\langle\psi_i|$$

they capture all probabilistic behaviors, including

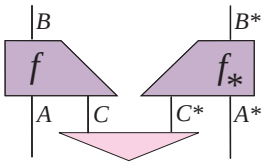
- 'lack of knowledge';
- 'statisical mixture';

- 'looking at part of a system'.

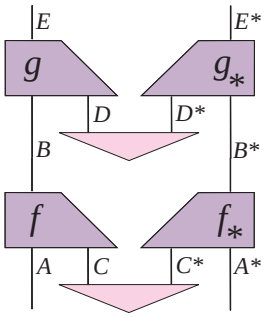
Exercise 11.2 Show that when we measure the second system of a compound system which is globally in state $|\Psi_f\rangle$ that the probabilities are given by $\mathsf{Tr}(\rho\,\mathrm{P})$ where $\rho := f \circ f^\dagger$.

Since $\rho := f \circ f^\dagger$ is positive, it is also self-adjoint, and hence admits a spectral decomposition so it can indeed be written as a density matrix of the above kind.

Besides *mixed states* there are of course also *mixed operations*. We end with a very recent result in which both mixed states and mixed operations are constructed at the abstract level of the picture calculus [**?**]:



and admit covariant composition:



References

[1] Kleppner, D. and Jackiw, R. (2000) *One Hundred Years of Quantum Physics*. Science **289**, 893–898. `arXiv:quant-ph/0008092`

[2] von Neumann, J. (1932) *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag. English translation (1955): *Mathematical Foundations of Quantum Mechanics*. Princeton University Press.

[3] Rédei, M. (1997) *Why John von Neumann did not like the Hilbert space formalism of quantum mechanics (and what he liked instead)*. Studies in History and Philosophy of Modern Physics **27**, 493–510.

[4] Birkhoff, G. and von Neumann, J. (1936) *The logic of quantum mechanics*. Annals of Mathematics **37**, 823–843.

[5] Dirac, P. A. M. (1947) *The Principles of Quantum Mechanics*, 3rd edition. Oxford University Press.

[6] Kochen, S. and Specker, E.P. (1967) *The Problem of Hidden Variables in Quantum Mechanics*. Journal Mathathematics and Mechanics **17**, 59–87.

[7] Belinfante, F. J. (1973) *A Survey of Hidden-Variables Theories*. Pergamon Press. (in Mathematical Institute's library)

[8] Wootters, W. and Zurek, W. (1982) A single quantum cannot be cloned. *Nature* **299**, 802–803.

[9] Pati, A. K. and Braunstein, S. L. (2000) Impossibility of deleting an unknown quantum state. *Nature* **404**, 164–165.

[10] Bennett, C. H. and Wiesner, S. J. (1992) *Communication via one- and two-particle operators on EPR states*. Physical Review Letters **69**, 2881–2884.

[11] Bennett, C. H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A. and Wooters, W. K. (1993) *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*. Physical Review Letters **70**, 1895–1899.

[12] Żukowski, M., Zeilinger, A., Horne, M. A. and Ekert, A. K. (1993) *'Event-ready-detectors' Bell experiment via entanglement swapping*. Physical Review Letters **71**, 4287–4290.

[13] Shor, P. W. (1994) *Algorithms for quantum computation: discrete logarithms and factoring*. Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Science Press.

[14] Grover, L. (1997) *Quantum mechanics helps in searching for a needle in a haystack*. Physical Review Letters **79**, 325–328. `arXiv:quant-ph/9706033`

[15] Gottesman, D. and Chuang, I. L. (1999) *Quantum teleportation is a universal computational primitive*. Nature **402**, 390–393.

[16] Gruska, J. (1999) *Quantum Computing*. McGraw-Hill.

[17] Nielsen, M. A. and Chuang, L. (2000) *Quantum Computation and Quantum Information*. Cambridge University Press.

[18] Nielsen, M. A. (1999) *Conditions for a class of entanglement transformations*. Physical Review Letters **83**, 436–439.

[19] Raussendorf, R. and Briegel, H.-J. (2001) A one-way quantum computer. *Physical Review Letters* **86**, 5188. Raussendorf, R., Browne, D. E. and Briegel, H.-J. (2003) Measurement-basisd quantum computation on cluster states. *Physical Review* A **68**, 022312. `arXiv:quant-ph/0301052`

[20] Kitaev, A. Yu., Shen, A. H. and Vyalyi, M. N. (2001) *Classical and Quantum Computing*. Garduate Studies in Mathematics **47**, American Mathematical Society.

[21] Quantum Technologies Group at ARC Seibersdorf research GmbH and the group Quantum Experiments and the Foundations of Physics of the University of Vienna (April 21, 2004) *World Premiere: Bank Transfer via Quantum Cryptography Based on Entangled Photons*. Press release. `www.quantenkryptographie.at/`

[22] Beth, Th., Mueller-Quade, J., and Steinwandt, R. (2004) Cryptanalysis of a practical quantum key distribution with polarization-entangled photons. To appear in *Quantum Information and Computation*. Available at `arXiv:quant-ph/0407130`

[23] Paulsen, V. (2002) *Completely Bounded Maps and Operator algebras*. Cambridge University Press.

[24] Mackey, G. W. (1963) *Mathematical Foundations of Quantum Mechanics*. W.A. Benjamin Inc.

[25] Piron, C. (1964) *Axiomatique quantique*. Helvetica Physica Acta **37**, 439–468. Piron, C. (1976) *Foundations of Quantum Physics*, W.A. Benjamin Inc. Solèr, M. P. (1995) *Characterization of Hilbert Spaces by Orthomodular Spaces'*. Communcations in Algebra **23**, 219–243.

[26] Coecke, B. (2003) The Logic of entanglement. An invitation. PRG-RR-03-12. web.comlab.ox.ac.uk/oucl/publications/tr/rr-03-12.html & arXiv:quant-ph/0402014

[27] Abramsky, S. and Coecke, B. (2004) *A categorical semantics of quantum protocols*. Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, IEEE Computer Science Press. arXiv:quant-ph/0402130 Coecke, B. (2005) *Quantum information-flow, concretely, and axiomatically.* In: *Proceedings of Quantum Informatics 2004*, pp. 15–29, Y. I. Ozhigov, Ed., Proceedings of SPIE Vol. 5833. arXiv:quant-ph/0506132 Coecke, B. (2005) *Kindergarten quantum mechanics — lecture notes*. In: *Quantum Theory: Reconsiderations of the Foundations* III. A. Khrennikov, Ed., American Institute of Physics Press. arXiv:quant-ph/0510032

[28] Abramsky, S. and Duncan, R. (2005) *A categorical quantum logic*. Mathematical Structures in Computer Science. arXiv:quant-ph/0512114