1 Basic Concepts

1.1 Dirac's Bra-ket Notation

Let $\mathcal{H}$ be a complex Hilbert space. We use *Dirac's bra-ket* notation for the vectors in $\mathcal{H}$ and its dual space, which we briefly introduce here. Vectors in $\mathcal{H}$ are denoted as *ket*'s $|\cdot\rangle$, and for any $|\varphi\rangle \in \mathcal{H}$, the corresponding *bra*-vector is defined as the linear functional $\langle\varphi| : \mathcal{H} \to \mathbb{C}$ which maps $|\psi\rangle \in \mathcal{H}$ to the *inner product* of $|\varphi\rangle$ and $|\psi\rangle$, which is denoted as $\langle\varphi|\psi\rangle$; hence, by definition, $\langle\varphi||\psi\rangle = \langle\varphi|\psi\rangle$. Furthermore, for $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$, the *outer product* of $|\varphi\rangle$ and $|\psi\rangle$ is defined as the linear function $|\varphi\rangle\langle\psi| : \mathcal{H} \to \mathcal{H}$ which maps $|\eta\rangle \in \mathcal{H}$ to $|\varphi\rangle\langle\psi|\eta\rangle$; hence, by definition, $|\varphi\rangle\langle\psi||\eta\rangle = |\varphi\rangle\langle\psi|\eta\rangle$.

We only consider *finite-dimensional* Hilbert spaces, so that we always may assume that $\mathcal{H} = \mathbb{C}^d$ for some (finite) dimension d, and the set $\mathcal{E}nd(\mathcal{H})$ of linear maps $\mathcal{H} \to \mathcal{H}$ coincides with the set $\mathbb{C}^{d\times d}$ of d-dimensional square matrices. A vector $|\varphi\rangle \in \mathcal{H}$ should then be thought of as a column vector

$$|\varphi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} \in \mathbb{C}^d$$

and $\langle\varphi|$ as the corresponding transpose complex-conjugate row vector $|\varphi\rangle^\dagger = (\bar{a}_1, \ldots, \bar{a}_d)$. Then, as matrix product, $\langle\varphi||\psi\rangle$ indeed coincides with the standard inner product[1] of $|\varphi\rangle$ and $|\psi\rangle$, whereas the matrix product $|\varphi\rangle\langle\psi|$ results in the $d \times d$ matrix which indeed maps $|\eta\rangle$ to $|\varphi\rangle\langle\psi|\eta\rangle$.

An equality that is easy to see yet handy to be aware of is that $\mathrm{tr}(|\varphi\rangle\langle\psi|) = \langle\psi|\varphi\rangle$ for any $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$, and thus that $\mathrm{tr}(A|\varphi\rangle\langle\varphi|) = \langle\varphi|A|\varphi\rangle$ for any $A \in \mathcal{E}nd(\mathcal{H})$ and $|\varphi\rangle \in \mathcal{H}$.

Finally, for two (and similarly for more) vectors $|\varphi\rangle \in \mathcal{H}$ and $|\psi\rangle \in \mathcal{H}'$, we write $|\varphi\rangle|\psi\rangle$ as well as $|\varphi, \psi\rangle$ as a short hand for the vector $|\varphi\rangle \otimes |\psi\rangle \in \mathcal{H} \otimes \mathcal{H}'$.[2]

1.2 Postulates of Quantum Information Processing

Throughout the definitions in this section, $\mathcal{H}$ denotes an arbitrary (finite dimensional) complex Hilbert space.

Definition 1.1. *$\mathcal{D}(\mathcal{H})$ denotes the set of all linear operators (matrices) $\rho \in \mathcal{E}nd(\mathcal{H})$ which are positive semi-definite and have $\mathrm{tr}(\rho) = 1$. A matrix $\rho \in \mathcal{D}(\mathcal{H})$ is called a* **density matrix**.

Postulate 1 *Associated to a quantum system A is a complex Hilbert space $\mathcal{H}$, the state space. The behavior of A is determined by its state, which is described by a density matrix $\rho \in \mathcal{D}(\mathcal{H})$. A quantum system A with state space $\mathcal{H}$ can be prepared to be in any initial state $\rho \in \mathcal{D}(\mathcal{H})$.*

In quantum information processing, one considers the state of a quantum system to be *static*, i.e., to not change over time, *unless* it is actively influenced by the "experimenter"; see Postulates 3 and 4. This is in contrast, though not in contradiction, to the view typically taken in physics, where a state's *dynamical* behavior, described by the Schrödinger equation, is the main focus of interest.

When considering two or more quantum systems, A, B etc., by default we denote the corresponding state spaces as $\mathcal{H}_\mathsf{A}, \mathcal{H}_\mathsf{B}$ etc., and similarly we sometimes use A, B etc. as subscripts

[1] Note that we consider here the standard inner product which is conjugate-linear in the first argument and linear in the second argument; this is opposite to the convention typically used in mathematics.

[2] We try to avoid expressions like $|\varphi\rangle|\psi\rangle\langle\varphi|\langle\psi|$, for which it is syntactically not clear how to read them, and instead write $(|\varphi\rangle \otimes |\psi\rangle)(\langle\varphi| \otimes \langle\psi|)$, or use the latter convention $|\varphi, \psi\rangle\langle\varphi, \psi|$.

for the (density) matrices that act on the corresponding Hilbert spaces (and later also for the names of vectors that "live" in the corresponding Hilbert spaces).

Postulate 2 *For two quantum systems A and B, the state space of the joint (or composite) system AB is given by $\mathcal{H}_{\mathsf{AB}} = \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$. If A and B are independently prepared to be in respective states $\rho_{\mathsf{A}} \in \mathcal{D}(\mathcal{H}_{\mathsf{A}})$ and $\rho_{\mathsf{B}} \in \mathcal{D}(\mathcal{H}_{\mathsf{B}})$, then the joint state of AB is $\rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}$.*

It is important, though, to realize that in general the state $\rho_{\mathsf{AB}} \in \mathcal{D}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ of a joint system is *not* of the form $\rho_{\mathsf{AB}} = \rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}$ with $\rho_{\mathsf{A}} \in \mathcal{D}(\mathcal{H}_{\mathsf{A}})$ and $\rho_{\mathsf{B}} \in \mathcal{D}(\mathcal{H}_{\mathsf{B}})$; we discuss this and its implications in more detail in Section 1.7.

Definition 1.2. *$\mathcal{U}(\mathcal{H})$ denotes the set of all unitary operators (matrices) $U \in \mathcal{E}nd(\mathcal{H})$.*

Postulate 3 *An isolated quantum system A can be operated on (only) by applying a unitary operator $U \in \mathcal{U}(\mathcal{H}_{\mathsf{A}})$; as a result, the state $\rho \in \mathcal{D}(\mathcal{H}_{\mathsf{A}})$ of A evolves to the new state $U\rho U^\dagger$. If A is part of a joint system AB, then applying U to A acts as applying $U \otimes \mathbb{I}_{\mathsf{B}}$ to AB.*

The only way for an experimenter to gain information on the state of a quantum system is by means of a **measurement**. A measurement is described by a (finite) collection $\mathbf{M} = \{M_i\}_{i \in I}$ of **measurement matrices** $M_i \in \mathcal{E}nd(\mathcal{H})$ which satisfy the condition

$$\sum_i M_i^\dagger M_i = \mathbb{I}\,.$$

Postulate 4 *For quantum system A in state $\rho \in \mathcal{D}(\mathcal{H}_{\mathsf{A}})$, measuring A with respect to the collection $\mathbf{M} = \{M_i\}_{i \in I}$ of measurement matrices has the following effect.*

1. *An* outcome *$i \in I$ is observed. The probability that a specific $i \in I$ is observed is given by*

$$p_i = \mathrm{tr}(M_i^\dagger M_i \rho) = \mathrm{tr}(M_i \rho M_i^\dagger)\,.$$

2. *After the measurement, the state ρ has* collapsed *to the post-measurement state*

$$\rho_i = \frac{1}{p_i} M_i \rho M_i^\dagger$$

where i is the outcome observed.

If A is part of a joint system AB, then measuring A with respect to $\mathbf{M} = \{M_i\}_{i \in I} \subset \mathcal{E}nd(\mathcal{H}_{\mathsf{A}})$ acts as measuring AB with respect to $\{M_i \otimes \mathbb{I}_{\mathsf{B}}\}_{i \in I}$.

Note that $p_i \geq 0$ for any $i \in I$ (because $\rho \geq 0$ and so is $M_i \rho M_i^\dagger$), and $\sum_i p_i = \sum_i \mathrm{tr}(M_i^\dagger M_i \rho) = \mathrm{tr}\big(\sum_i M_i^\dagger M_i \rho\big) = \mathrm{tr}(\mathbb{I}\rho) = 1$; hence, the p_i's indeed form a probability distribution.

We are sometimes a bit sloppy with the terminology and use the term "state" for a quantum system, its state, as well as the density matrix that describes it; however, this should not cause any confusion. For instance, we may say that we "measure a state ρ", where we actually mean that we measure a quantum system A whose state is described by the density matrix ρ.

The following is an immediate consequence of the fact that matrices acting on different subsystems commute.

Theorem 1.3. *The actions of acting (by means of a unitary operation or a measurement) on subsystem A and of acting on subsystem B of a joint system AB commute.*

1.3 Von Neumann Measurements and POVM's

We will mainly consider **Von Neumann** (also known as **projective**) **measurements**. These are measurements where the measurement matrices M_i are **orthogonal projections**, meaning that $M_i^\dagger = M_i$ and $M_i M_i = M_i$, and thus for which obviously the probability of observing i simplifies to $p_i = \mathrm{tr}(M_i\rho)$. In particular, we often consider **complete Von Neumann measurements**, where the M_i's are projections onto an orthogonal basis of $\mathcal{H}$: $\mathbf{M} = \{|i\rangle\langle i|\}_{i\in I}$, where $\{|i\rangle\}_{i\in I}$ forms an *orthonormal basis* of $\mathcal{H}$, i.e., the $|i\rangle$'s span $\mathcal{H}$ and $\langle i|j\rangle = 1$ if $i = j$ and else $= 0$ for any $i, j \in I$.[3] Note that the projections $M_i = |i\rangle\langle i|$ indeed satisfy the condition $\sum_i M_i^\dagger M_i = \mathbb{I}$, and $p_i = \mathrm{tr}(M_i\rho)$ simplifies to $p_i = \mathrm{tr}(|i\rangle\langle i|\rho) = \langle i|\rho|i\rangle$. In case of such a complete Von Neumann measurement, we also say that A is measured **in basis** $\{|i\rangle\}_{i\in I}$.[4]

In the special case where $\{|i\rangle\}_{i\in I}$ is the *canonical* basis of $\mathcal{H} = \mathbb{C}^d$, the p_i's are simply the diagonal elements of ρ; in case of an arbitrary orthonormal basis, the p_i's can be seen as the diagonal elements of ρ after a corresponding basis transformation. Also note that in case of a complete Von Neumann measurement, the collapsed state ρ_i equals[5] $\rho_i = |i\rangle\langle i|$, and as such is uniquely determined by the outcome i of the measurement; and thus we may just as well understand that the quantum system A actually *vanishes* as a result of the measurement, and if needed A can be "freshly" prepared in state $|i\rangle\langle i|$ (given the outcome i). Important to realize is that one cannot measure the system *in the original state* ρ a second time.

In cases where one is only interested in the measurement outcome (and its distribution) but not in the post-measurement state, the general measurement formalism of a collection $\mathbf{M} = \{M_i\}_{i\in I}$ of measurement matrices can be simplified by considering the collection $\mathbf{E} = \{E_i\}_{i\in I}$ instead, where $E_i = M_i^\dagger M_i$, so that $p_i = \mathrm{tr}(M_i^\dagger M_i \rho)$ simplifies to $p_i = \mathrm{tr}(E_i\rho)$. Any collection $\mathbf{E} = \{E_i\}_{i\in I}$ of positive semi-definite matrices $E_i \geq 0$ with $\sum_i E_i = \mathbb{I}$ is called a **POVM** (which stands for a "Positive Operator-Valued Measure") and can be understood as being obtained from a collection of measurement matrices $\mathbf{M} = \{M_i\}_{i\in I}$ as above, but the M_i's — and thus the post-measurement states resulting from a POVM measurement — are not uniquely determined.

1.4 Pure States

Consider a system A with state space $\mathcal{H}$. We say that the state $\rho \in \mathcal{D}(\mathcal{H})$ of A is a *pure* state if ρ is of rank 1. The positivity condition then implies that it is of the form $\rho = |\varphi\rangle\langle\varphi|$ for some $|\varphi\rangle \in \mathcal{H}$, and the trace-condition on ρ implies that $|\varphi\rangle$ has Euclidean norm $\||\varphi\rangle\| = 1$; indeed, $\langle\varphi|\varphi\rangle = \mathrm{tr}(|\varphi\rangle\langle\varphi|) = 1$. In order to clearly distinguish a general state from a pure state, a general state is called a *mixed* state.[6] Obviously, in case of a pure state, ρ is uniquely determined by $|\varphi\rangle$ with $\||\varphi\rangle\| = 1$ and $\rho = |\varphi\rangle\langle\varphi|$, and thus we may just as well use such a vector $|\varphi\rangle$ as a description of the state of A.[7] And, vice versa, any norm-1 vector $|\varphi\rangle \in \mathcal{H}$ describes a pure state, given by the density matrix $\rho = |\varphi\rangle\langle\varphi| \in \mathcal{D}(\mathcal{H})$.

It is straightforward to verify that using this state-vector formalism for describing pure states, the composition of systems (Postulate 2) translates to: if A and B are independently

[3] Note that we are using the indices $i \in I$ as the "names" of the basis vectors; indeed we will often name basis vectors by numbers, like $\{|0\rangle, |1\rangle\}$, but the index set I may just as well consists of other "symbols".

[4] In the literature, one also finds the terminology that the **observable** M is measured of the system A, where M is a Hermitian matrix M acting on $\mathcal{H}$. This is to be understood in that A is measured in basis $\{|i\rangle\}_{i\in I}$, where $M = \sum_i \lambda_i |i\rangle\langle i|$ is the spectral decomposition (see Theorem A.6) of M. Vice versa, measuring in a basis may be phrased in terms of an observable M.

[5] Indeed, one can write ρ as $\rho = \sum_{j,k\in I} a_{j,k}|j\rangle\langle k|$ and then the claim is rather easy to verify using the orthogonality of $\{|i\rangle\}_{i\in I}$, but the claim will also become obvious from some later observations.

[6] The literature is somewhat inconsistent in whether a *mixed* state should be *any* state or any *non-pure* state. We adopt the former convention and use the term *mixed* merely to emphasize that the state we consider does not have to (but may) be pure.

[7] The vector $|\varphi\rangle$ is unique up to multiplication with $\omega \in \mathbb{C}$ with $|\omega| = 1$.

prepared to be in respective pure states $|\varphi_A\rangle$ and $|\varphi_B\rangle$, then the joint state of AB is $|\varphi_{AB}\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle$. Indeed, the tensor product $|\varphi_{AB}\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle$ acts component-wise on $\langle\varphi_{AB}| = \big(|\varphi_A\rangle \otimes |\varphi_B\rangle\big)^\dagger = \langle\varphi_A| \otimes \langle\varphi_B|$, so that

$$|\varphi_{AB}\rangle\langle\varphi_{AB}| = (|\varphi_A\rangle \otimes |\varphi_B\rangle)(\langle\varphi_A| \otimes \langle\varphi_B|) = |\varphi_A\rangle\langle\varphi_A| \otimes |\varphi_B\rangle\langle\varphi_B| \,.$$

Furthermore, the evolution $\rho \rightsquigarrow U\rho U^\dagger$ of a system (Postulate 3) translates to $|\varphi\rangle \rightsquigarrow U|\varphi\rangle$, and the "rule" for measuring a system (Postulate 4) translates as follows. If the state of A is given by $|\varphi\rangle \in \mathcal{H}$, then measuring A with respect to $\mathbf{M} = \{M_i\}_{i\in I}$ has the effect that i is observed with probability

$$p_i = \mathrm{tr}\big(M_i^\dagger M_i |\varphi\rangle\langle\varphi|\big) = \langle\varphi|M_i^\dagger M_i|\varphi\rangle \,,$$

and the state collapses to the post measurement state $\rho_i = \frac{1}{p_i} M_i|\varphi\rangle\langle\varphi|M_i^\dagger$, i.e. to $\frac{1}{\sqrt{p_i}} M_i|\varphi\rangle$ using the state-vector formalism.

In case of a complete Von Neumann measurement, the formalism gets even simpler: measuring A in basis $\{|i\rangle\}_{i\in I}$ has the effect that i is observed with probability

$$p_i = \langle\varphi|(|i\rangle\langle i|)^\dagger |i\rangle\langle i||\varphi\rangle = \langle\varphi||i\rangle\langle i||\varphi\rangle = \langle\varphi|i\rangle\langle i|\varphi\rangle = |\langle i|\varphi\rangle|^2 \,,$$

and the state collapses to the post measurement state $\rho_i = \frac{1}{p_i}|i\rangle\langle i||\varphi\rangle\langle\varphi||i\rangle\langle i| = |i\rangle\langle i|$, i.e. to $|i\rangle$, using the state-vector formalism. Writing $|\varphi\rangle$ in the basis $\{|i\rangle\}_{i\in I}$:

$$|\varphi\rangle = \sum_j \alpha_j |j\rangle \,,$$

where the restriction of the norm implies that $\sum_j |\alpha_j|^2 = 1$, this amounts to

$$p_i = |\alpha_i|^2 \,.$$

1.5 Qubits

The "smallest" and thus simplest (non-degenerate) quantum system is the **qubit system**, which is given by the two-dimensional state space $\mathcal{H} = \mathbb{C}^2$. We denote the canonical orthonormal basis of $\mathbb{C}^2$ by $|0\rangle = \binom{1}{0}$ and $|1\rangle = \binom{0}{1}$. This is also called the **computational** (or the **rectilinear**) **basis**. A (pure) qubit state is thus given by a vector $|\varphi\rangle = \binom{\alpha_0}{\alpha_1} = \alpha_0|0\rangle + \alpha_1|1\rangle \in \mathbb{C}^2$ with $|\alpha_0|^2 + |\alpha_1|^2 = 1$ (see Figure 1), and measuring it in the computational basis has the effect that bit $i \in \{0, 1\}$ is observed with probability $p_i = |\alpha_i|^2$.

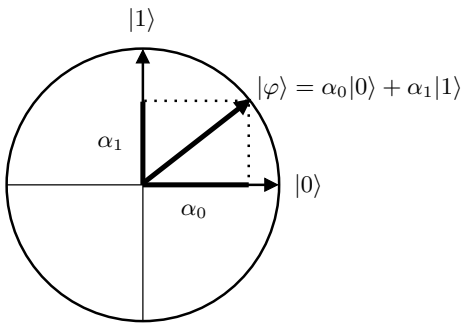

Fig. 1. Pure state of a qubit system.

The orthonormal basis $\{|+\rangle, |-\rangle\}$ of the qubit state space $\mathbb{C}^2$, consisting of the vectors

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad \text{and}$$
$$|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,,$$

is called the **Hadamard** (or the **diagonal**) **basis**, and the corresponding basis transformation, $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right) \in \mathcal{U}(\mathbb{C}^2)$, the **Hadamard transformation**. To compute the probabilities of observing "+" and "−" when measuring a pure qubit state $|\varphi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ in the Hadamard basis $\{|+\rangle, |-\rangle\}$, we re-write $|\varphi\rangle$ in terms of the Hadamard basis:

$$|\varphi\rangle = \frac{1}{\sqrt{2}}(\alpha_0 + \alpha_1)|+\rangle + \frac{1}{\sqrt{2}}(\alpha_0 - \alpha_1)|-\rangle\,,$$

which allows to read off the corresponding probabilities as $p_+ = \frac{1}{2}|\alpha_0+\alpha_1|^2$ and $p_- = \frac{1}{2}|\alpha_0-\alpha_1|^2$. By convention, we identify the outcome "+" with 0 and the outcome "−" with 1, and as such we from now on understand the outcome of measuring a qubit system in the Hadamard basis to be a bit $i \in \{0,1\}$.

For $x, \theta, \theta' \in \{0,1\}$, measuring the qubit $H^\theta|x\rangle$ in basis $\{H^{\theta'}|0\rangle, H^{\theta'}|1\rangle\}$ has the following effect. If $\theta = \theta'$ then the bit x is observed with certainty, but if $\theta \neq \theta'$ then a random (meaning: uniformly distributed) bit is observed.

An **n-qubit system** is the n-fold composition of a qubit system, and thus has state space $\mathcal{H} = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$, and its (pure) state can be written as

$$|\varphi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x|x\rangle = \sum_{x\in\{0,1\}^n} \alpha_x|x_1\rangle\cdots|x_n\rangle$$

with $\sum_x |\alpha_x|^2 = 1$ (and where we use the convention of omitting the $\otimes$-symbol).

1.6 Mixed vs Pure States

The vector formalism for pure states is typically handier to work with (and probably also easier to understand) than the density-matrix formalism; we now show that actually also non-pure states can be understood in terms of the vector formalism. First, we show that any mixed state can be written as a convex linear combination of pure states (and vice versa that any convex linear combination of pure states is a mixed state).

Theorem 1.4. *Let $\mathcal{H}$ be a complex Hilbert space, and let ρ be a linear operator on $\mathcal{H}$. Then, $\rho \in \mathcal{D}(\mathcal{H})$ if and only if ρ can be written as*

$$\rho = \sum_{\ell=1}^{L} \varepsilon_\ell |\varphi_\ell\rangle\langle\varphi_\ell|\,,$$

where $\varepsilon_\ell \geq 0$ for any ℓ and $\sum_\ell \varepsilon_\ell = 1$, and $|\varphi_\ell\rangle \in \mathcal{H}$ with Euclidean norm $\||\varphi_\ell\rangle\| = 1$ for any ℓ.

Proof. For the backward implication, note that

$$\langle\psi|\rho|\psi\rangle = \sum_\ell \varepsilon_\ell \langle\psi||\varphi_\ell\rangle\langle\varphi_\ell||\psi\rangle = \sum_\ell \varepsilon_\ell |\langle\psi|\varphi_\ell\rangle|^2 \geq 0$$

for any vector $|\psi\rangle$. Furthermore,

$$\mathrm{tr}(\rho) = \sum_\ell \varepsilon_\ell \mathrm{tr}\big(|\varphi_\ell\rangle\langle\varphi_\ell|\big) = \sum_\ell \varepsilon_\ell \langle\varphi_\ell|\varphi_\ell\rangle = \sum_\ell \varepsilon_\ell = 1\,.$$

The forward implication follows from the spectral decomposition (Theorem A.6), which implies that ρ can be written as $\rho = \sum_i \lambda_i |i\rangle\langle i|$ for an orthonormal basis $\{|i\rangle\}_{i\in I}$, where by the positive semi-definiteness and the assumption on the trace, the λ_i's are non-negative and add up to 1. □

In Proposition 1.5 below, we show that a *randomized* system, i.e., a system whose state is given by density matrix ρ_ℓ with probability ε_ℓ, can be equivalently described by the fixed deterministic state given by the corresponding convex linear combination $\rho = \sum_\ell \varepsilon_\ell \rho_\ell$.[8] This in particular implies that in case of *pure* states, where with probability ε_ℓ the state is given by $|\varphi_\ell\rangle$, the system can be equivalently described by the density matrix

$$\rho = \sum_\ell \varepsilon_\ell |\varphi_\ell\rangle\langle\varphi_\ell| \,.$$

And, vice versa, it implies that any mixed state $\rho \in \mathcal{D}(\mathcal{H})$, which by Theorem 1.4 can be decomposed into $\rho = \sum_\ell \varepsilon_\ell |\varphi_\ell\rangle\langle\varphi_\ell|$, can be understood as a *randomized pure state*, where with probability ε_ℓ the state is given by $|\varphi_\ell\rangle$. Here, it is important to realize that in general the decomposition of ρ into pure states is *not* unique: there exist *distinct* distributions, also called **ensembles**, $\{(\varepsilon_\ell, |\varphi_\ell\rangle)\}_{\ell=1\ldots L}$ and $\{(\varepsilon'_\ell, |\varphi'_\ell\rangle)\}_{\ell=1\ldots L'}$ with *the same* density matrix $\sum_\ell \varepsilon_\ell |\varphi_\ell\rangle\langle\varphi_\ell| = \sum_\ell \varepsilon'_\ell |\varphi'_\ell\rangle\langle\varphi'_\ell|$. This means, although their respective descriptions as randomized states are different, the two states are actually identical. We will see a simple example below.

Proposition 1.5. *Let A be a system with a randomized state: A is in state $\rho_\ell \in \mathcal{D}(\mathcal{H})$ with probability ε_ℓ (where $\ell = 1, \ldots, L$), and let A' be the system with (fixed) state $\rho = \sum_\ell \varepsilon_\ell \rho_\ell$. Let p_i and p'_i be the respective probabilities to observe i when measuring A and A' with respect to some collection $\mathbf{M} = \{M_i\}_{i\in I}$ of measurement matrices. Then $p_i = p'_i$ for all $i \in I$.*

Proof. Follows immediately from the linearity of the trace: Let $p_{i|\ell} = \mathrm{tr}(M_i^\dagger M_i \rho_\ell)$ be the probability to observe i on measuring A, given that it is in state ρ_ℓ. Then, the probability p_i of observing i, not conditioned on the choice of ℓ, is given by

$$p_i = \sum_\ell \varepsilon_\ell p_{i|\ell} = \sum_\ell \varepsilon_\ell \mathrm{tr}(M_i^\dagger M_i \rho_\ell) = \mathrm{tr}\Big(M_i^\dagger M_i \sum_\ell \varepsilon_\ell \rho_\ell\Big) = \mathrm{tr}(M_i^\dagger M_i \rho) = p'_i \,.$$

This proves the claim. □

We now give a simple, yet important, example of the fact that the decomposition of a mixed state into a convex linear combination of pure states (Theorem 1.4) is not unique. We consider the so-called **fully mixed** state $\frac{1}{2}\mathbb{I} \in \mathcal{D}(\mathbb{C}^2)$, and we note that it can be decomposed into

$$\frac{1}{2}\mathbb{I} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

as well as into

$$\frac{1}{2}\mathbb{I} = \frac{1}{2}HH^\dagger = \frac{1}{2}H\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big)H^\dagger = \frac{1}{2}H|0\rangle\langle 0|H^\dagger + \frac{1}{2}H|1\rangle\langle 1|H^\dagger = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| \,.$$

Following Proposition 1.5 and its discussion, this means that the fully mixed state can be understood as being in state $|0\rangle$ with probability $\frac{1}{2}$ and in state $|1\rangle$ with probability $\frac{1}{2}$, but just as well it can be understood as being in state $|+\rangle$ with probability $\frac{1}{2}$ and in state $|-\rangle$ with probability $\frac{1}{2}$.[9] And, vice versa, this means that the above two (differently prepared) randomized systems are described by the same density matrix and thus they are in identical states (if the experimenter does not reveal his random choice).

[8] Actually, per se, Proposition 1.5 only shows equivalence when doing a single measurement; but we will see later (Section 1.8) that any information obtained from a quantum state by means of the operations allowed by the postulates can also be obtained by a suitable single measurement.

[9] Actually, the corresponding holds for *any* orthonormal basis of $\mathbb{C}^2$.

1.7 Composite States

For simplicity, and because it suffices for our purposes, we focus here on *pure* states. The state of a composite system AB is given (in case it is pure) by a length-1 vector $|\varphi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. If $|\varphi_{AB}\rangle$ can be written as $|\varphi_{AB}\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle$ with $|\varphi_A\rangle \in \mathcal{H}_A$ and $|\varphi_B\rangle \in \mathcal{H}_B$, then the quantum system AB is said to be in *product state*; we will see later on that this means that the two parts are in some sense independent. However, it is important to realize that in general $|\varphi_{AB}\rangle$ is *not* of the form $|\varphi_A\rangle \otimes |\varphi_B\rangle$; in this case, AB it is said to be *entangled*.

Recall from Postulate 4 that measuring subsystem A with respect to a collection $\{M_i\}_{i\in I}$ of measurement matrices acts as measuring AB with respect to $\{M_i \otimes \mathbb{I}_B\}_{i\in I}$. Restricting to complete Von Neumann measurements and to pure states, this amounts to the following. Let $|\varphi_{AB}\rangle \in \mathcal{H}_{AB}$ be the state of joint system AB, and let $|i\rangle_{i\in I}$ be an orthonormal basis of $\mathcal{H}_A$. We can write $|\varphi_{AB}\rangle$ as

$$|\varphi_{AB}\rangle = \sum_{j\in I} \alpha_j |j\rangle \otimes |\psi_j\rangle \,,$$

with $|\psi_j\rangle \in \mathcal{H}_B$ and $\||\psi_j\rangle\| = 1$, and where the restriction on the norm implies that $\sum_j |\alpha_j|^2 = 1$. Then, measuring subsystem A in basis $|i\rangle_{i\in I}$ has the effect that $i \in I$ is observed with probability

$$p_i = |\alpha_i|^2$$

and the state of A collapses to $|i\rangle \otimes |\psi_i\rangle$. Indeed, in case of a pure state $|\varphi_{AB}\rangle$, we can write

$$\begin{aligned} p_i &= \langle\varphi_{AB}|(|i\rangle\langle i| \otimes \mathbb{I}_B)^\dagger(|i\rangle\langle i| \otimes \mathbb{I}_B)|\varphi_{AB}\rangle \\ &= \sum_{j,k} \alpha_i \bar{\alpha}_k (\langle k| \otimes \langle\psi_k|)(|i\rangle\langle i| \otimes \mathbb{I}_B)(|i\rangle\langle i| \otimes \mathbb{I}_B)(|j\rangle \otimes |\psi_j\rangle) \\ &= \sum_{j,k} \alpha_i \bar{\alpha}_k \langle k||i\rangle\langle i||i\rangle\langle i||j\rangle \otimes \langle\psi_k||\psi_j\rangle = |\alpha_i|^2 \,, \end{aligned}$$

and

$$\frac{1}{\sqrt{p_i}}(|i\rangle\langle i| \otimes \mathbb{I}_B)|\varphi_{AB}\rangle = \frac{1}{\sqrt{p_i}} \sum_j \alpha_j |i\rangle\langle i||j\rangle \otimes |\psi_j\rangle = \frac{\alpha_i}{\sqrt{p_i}} |i\rangle \otimes |\psi_i\rangle \,,$$

which coincides with $|i\rangle \otimes |\psi_i\rangle$ as state vector, i.e., when comparing the corresponding density matrices.

Since the collapsed state, $|i\rangle \otimes |\psi_i\rangle$, is in product form, with the state of the measured subsystem uniquely determined by the outcome of the measurement, we may also understand that the measured subsystem A actually vanishes and the state of B collapses to $|\psi_i\rangle$.

An important (entangled) 2-qubit state is the so-called **EPR pair**, named after Einstein, Podolsky and Rosen, and also known as the first of the four **Bell states** (see Section 2):

$$|\Phi\rangle = |\Phi^+\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle + |1\rangle|1\rangle\big) \in \mathcal{H}_A \otimes \mathcal{H}_B \,,$$

with $\mathcal{H}_A = \mathbb{C}^2 = \mathcal{H}_B$. It has the following peculiar property. If subsystem A (i.e., the first of the two qubits) is measured in the computational basis $\{|0\rangle, |1\rangle\}$, then $i \in \{0, 1\}$ is observed with probability $p_i = \frac{1}{2}$ and the state of system B collapses to $|i\rangle$.[10] Thus, if subsequently B is measured in the computational basis, then *the same* $i \in \{0, 1\}$ is observed (with probability 1). The corresponding holds when first B and then A is measured. Thus, even when the two subsystems are geographically far apart, say at "opposite ends of the universe", as long as A is untouched,

[10] Indeed, $|\Phi\rangle$ is of the form $|\Phi\rangle = \sum_j \alpha_j |j\rangle \otimes |\psi_j\rangle$ with $\alpha_0 = \alpha_1 = \frac{1}{\sqrt{2}}$ and $|\psi_0\rangle = |0\rangle$ and $|\psi_1\rangle = |1\rangle$.

the outcome of measuring B is undetermined, 0 or 1 at random, but as soon as A is measured, then instantaneously the outcome of measuring B is determined (to be the same as for A). This is what Einstein called "spukhafte Fernwirkung", i.e., "spooky action at a distance". Even more interesting, since $|\Phi\rangle$ can also be written as

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\big(|+\rangle|+\rangle + |-\rangle|-\rangle\big) = \frac{1}{\sqrt{2}}\big(H|0\rangle \otimes H|0\rangle + H|1\rangle \otimes H|1\rangle\big) \in \mathcal{H}_A \otimes \mathcal{H}_B\,,$$

which is verified by a straightforward calculation, the corresponding also holds when measuring A and B in the Hadamard basis $\{|+\rangle, |-\rangle\}$.

We conclude this section on composite states with the famous No-Cloning Theorem.

Theorem 1.6 (No-Cloning). *For any complex Hilbert spaces $\mathcal{H}_A = \mathcal{H}_{A'}$ and $\mathcal{H}_B$, and for any state vector $|\psi_{A'B}\rangle \in \mathcal{H}_{A'} \otimes \mathcal{H}_B$, there exists no unitary $U \in \mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_{A'} \otimes \mathcal{H}_B)$ such that for every state vector $|\varphi_A\rangle \in \mathcal{H}_A$ it holds that*

$$U(|\varphi_A\rangle|\psi_{A'B}\rangle) = |\varphi_A\rangle|\varphi_A\rangle|\psi'_B\rangle$$

for some $|\psi'_B\rangle$ (that may depend on $|\varphi_A\rangle$).

Proof. Consider arbitrary complex Hilbert spaces $\mathcal{H}_A = \mathcal{H}_{A'}$ and $\mathcal{H}_B$. Let $|i\rangle$ and $|j\rangle$ be a pair of orthonormal vectors in $\mathcal{H}_A$. Assuming the existence of $|\psi_{A'B}\rangle$ and U as required, it holds that $U|i\rangle|\psi_{A'B}\rangle = |i\rangle|i\rangle|\psi'_B\rangle$ and $U|j\rangle|\psi_{A'B}\rangle = |j\rangle|j\rangle|\psi''_B\rangle$ and

$$\begin{aligned} U\frac{1}{\sqrt{2}}(|i\rangle + |j\rangle)|\psi_{A'B}\rangle &= \frac{1}{2}(|i\rangle + |j\rangle) \otimes (|i\rangle + |j\rangle) \otimes |\psi'''_B\rangle \\ &= \frac{1}{2}(|i\rangle|i\rangle + |i\rangle|j\rangle + |j\rangle|i\rangle + |j\rangle|j\rangle) \otimes |\psi'''_B\rangle. \end{aligned}$$

However, by linearity it also holds that

$$U\frac{1}{\sqrt{2}}(|i\rangle + |j\rangle)|\psi_{A'B}\rangle = \frac{1}{\sqrt{2}}U|i\rangle|\psi_{A'B}\rangle + \frac{1}{\sqrt{2}}U|j\rangle|\psi_{A'B}\rangle = \frac{1}{\sqrt{2}}|i\rangle|i\rangle|\psi'_B\rangle + \frac{1}{\sqrt{2}}|j\rangle|j\rangle|\psi''_B\rangle\,,$$

which is different from the above. This can e.g. be seen by applying $\langle i|\langle j|\langle\psi'''_B|$ to both expressions: with the former, it results in 1, with the latter in 0. □

1.8 General Versus Von Neumann Measurements

One reason why it is often sufficient to consider Von Neumann measurements is that any general measurement $\mathbf{M} = \{M_i\}_{i\in I}$ on a system A can be "simulated" by means of appending an independent system B, a so-called **ancilla**, to system A, applying a unitary transformation to the joint system AB, and then doing a Von Neumann measurement. We now show this in detail. Let $\{|j\rangle\}_{j\in J}$ be an orthonormal basis of $\mathcal{H}_A$, and let $\{|i\rangle\}_{i\in I}$ be an orthonormal basis of $\mathcal{H}_B$. Note that we are free to choose B, thus we may choose the dimension of $\mathcal{H}_B$ to be $|I|$. Let $i_\circ$ be some designated element of I. We define $U \in \mathcal{E}nd(\mathcal{H}_A \otimes \mathcal{H}_B)$ as follows. First, for any $j \in J$, we specify

$$U|j\rangle|i_\circ\rangle = \sum_{i\in I} M_i|j\rangle \otimes |i\rangle\,.$$

This defines the action of U on the subspace of $\mathcal{H}_A \otimes \mathcal{H}_B$ spanned by the orthonormal vectors $\{|j\rangle|i_\circ\rangle\}_{j\in J}$, but leaves the action of U on the remaining vectors undefined. Note that for $j, k \in J$, it holds that

$$\langle k|\langle i_\circ|U^\dagger U|j\rangle|i_\circ\rangle = \sum_{i,i'\in I} (\langle k|M_{i'}^\dagger \otimes \langle i'|)(M_i|j\rangle \otimes |i\rangle)$$

$$= \sum_{i,i' \in I} \langle k|M_{i'}^\dagger M_i|j\rangle\langle i'|i\rangle = \sum_{i \in I} \langle k|M_{i'}^\dagger M_i|j\rangle = \langle j|k\rangle\,.$$

Thus, U preserves the orthonormality of the vectors $\{|j\rangle|i_\circ\rangle\}_{j\in J}$. Therefore, the above action of U on the subspace can be extended to a unitary transformation $U \in \mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_B)$. The latter can easily be seen from the characterization of a unitary matrix as basis transformation between orthonormal bases.

Let the state of A be given by $|\varphi_A\rangle$; if the state of A is mixed, then the corresponding will follow from Theorem 1.4 and the linearity of the operations below (adding an ancilla, applying a unitary, and performing a measurement). Let us prepare system B to be in state $|\varphi_B\rangle = |i_\circ\rangle$, so that the joint system AB is in state $|\varphi_{AB}\rangle = |\varphi_A\rangle|i_\circ\rangle$. We now analyze the outcome when U is applied to AB, and then system B is measured in basis $\{|i\rangle\}_{i\in I}$ (meaning that AB is measured by means of the Von Neumann measurement $\{\mathbb{I} \otimes |i\rangle\langle i|\}_{i\in I}$). Applying U to AB has the effect that the state of AB evolves to

$$|\varphi'_{AB}\rangle = U|\varphi_A\rangle|i_\circ\rangle = \sum_{i\in I} M_i|\varphi_A\rangle \otimes |i\rangle$$

so that when (the evolved state of) B is measured in basis $\{|i\rangle\}_{i\in I}$, $i \in I$ is observed with probability

$$p_i = \sum_{i',i'' \in I} \left(\langle\varphi_A|M_{i''}^\dagger \otimes \langle i''|\right)\left(\mathbb{I} \otimes |i\rangle\langle i|\right)\left(M_{i'}|\varphi_A\rangle \otimes |i'\rangle\right) = \langle\varphi_A|M_i^\dagger M_i|\varphi_A\rangle$$

and the joint state collapses to

$$\frac{1}{\sqrt{p_i}}(\mathbb{I} \otimes |i\rangle\langle i|) \sum_{i'\in I}(M_{i'}|\varphi_A\rangle \otimes |i'\rangle) = \frac{1}{\sqrt{p_i}} M_i|\varphi_A\rangle \otimes |i\rangle\,,$$

i.e., the state of A collapses to $\frac{1}{\sqrt{p_i}}M_i|\varphi_A\rangle$. Thus, we have the very same behavior as when the original state of A is measured with respect to the general measurement $\mathbf{M}$.

If one is merely interested in the measurement outcome (and its distribution), but not in the post-measurement state, then also the converse holds: the action of appending an ancilla, applying a unitary transformation (to the joint system) and then performing a measurement, can be captured by means of a general measurement. Since we are not interested in the post-measurement state here, we use the POVM formalism as introduced in Section 1.3. Let B be the ancilla system, let $|0\rangle$ be its state, let $U \in \mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_B)$, and let $\mathbf{E} = \{E_i\}_{i\in I} \subset \mathcal{E}nd(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a POVM. Without loss of generality, we may assume that $|0\rangle = (1, 0, \ldots, 0)^\dagger \in \mathcal{H}_B$. For any state $|\varphi_A\rangle \in \mathcal{H}_A$, the action of adding the ancilla, applying U, and performing the measurement, leads to the measurement outcome $i \in I$ with probability

$$p_i = \langle\varphi_A|\langle 0|U^\dagger E_i U|\varphi_A\rangle|0\rangle = \langle\varphi_A|\langle 0|\tilde{E}_i|\varphi_A\rangle|0\rangle$$

for $\tilde{E}_i = U^\dagger E_i U$. From the operational interpretation of the tensor product of matrices, we obtain that $\langle\varphi_A|\langle 0|$ can be written as block-vector $\langle\varphi_A|\langle 0| = (\langle\varphi_A|, 0, \ldots, 0)$, where each 0 is the zero-vector in $\mathcal{H}_A$. It thus follows that

$$p_i = \langle\varphi_A|\tilde{E}_i^{11}|\varphi_A\rangle$$

where $\tilde{E}_i^{11} \in \mathcal{E}nd(\mathcal{H}_A)$ is the upper left block of $\tilde{E}_i$. It is obvious that if $E_i \geq 0$ then also $\tilde{E}_i = U^\dagger E_i U \geq 0$ as well as $\tilde{E}_i^{11} \geq 0$, and if the E_i's add up to 1 then so do the $\tilde{E}_i$'s as well as the $\tilde{E}_i^{11}$'s. Thus, $\{\tilde{E}_i^{11}\}_{i\in I}$ forms a POVM that produces the same outcome distribution when applied to the original state $|\varphi_A\rangle$.