Part I

Introduction to Quantum Information Science

Chapter 1

The State-Vector Formalism

The goal of quantum mechanics is to provide a mathematical framework that provides the means to rigorous describe and predict the behavior of certain physical objects, typically particles like electrons, photons etc. In contrast to a typical physics textbook on quantum mechanics, here we are not interested in analyzing *specific* physical systems, like the hydrogen atom; instead, we want to understand the *general behavior* of "quantum-mechanical objects", and how this — often strange — behavior affects the concepts of *computation* and *information* as we know them, and which arose from abstracting the behavior of typical (and thus non-quantum) information-processing devices. This angle of quantum mechanics is called **quantum computing** or **quantum information-processing** if the focus lies on aspects related more to computing, or **quantum information theory** if the focus lies on aspects related to the behavior of information. In its union, it is referred to as **quantum information science**.

In this section, we introduce the so-called state vector formalism of quantum mechanics, which is one particular framework for describing quantum systems and their behavior. The state-vector formalism is somewhat limited in that there are certain aspects that can not be (well) captured; however, those are not very relevant in the context of quantum computing, which is what we focus on first.

1.1 State Spaces and State Vectors

Let $\mathcal{H}$ be an arbitrary Hilbert space.

Definition 1.1. *$\mathcal{S}(\mathcal{H})$ denotes the set of all norm-1 vectors in $\mathcal{H}$, i.e.,*

$$\mathcal{S}(\mathcal{H}) := \{|\varphi\rangle \in \mathcal{H} \,|\, \||\varphi\rangle\| = 1\}.$$

A vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ is called a **state vector***.*

The connection to physics is as follows. Any **(quantum) system**, i.e., any physical system that follows the laws of quantum mechanics (like the polarization of a photon, or the spin of an electron), is associated to a Hilbert space $\mathcal{H}$, called the **state space** of the system.[1] The **(quantum) state** of the system, which is meant to determine the future behavior of the system, can then be mathematically described by a state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$. In the context of quantum computing, such a quantum system is also referred to as a **register**. However, we tend to be a bit sloppy with the terminology and do not always distinguish well between the quantum

[1] Certain quantum systems have an infinite dimensional Hilbert space, like $L^2(\mathbb{R}^3)$, as state spaces; however, in these notes, we restrict to (systems with) finite dimensional state spaces.

system, the state of the system, and the description of the state by means of a state vector. Later on, when we consider multiple quantum systems, we will refer to them by A, B etc. in order to distinguish between them, and their respective state spaces are then by default denoted by $\mathcal{H}_A, \mathcal{H}_B$ etc.

In case of a 2-dimensional state space $\mathcal{H}$, which may then be assumed to be $\mathcal{H} = \mathbb{C}^2$, the quantum system, respectively the state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ describing its state, is typically called a **qubit**, and in case of dimension $d > 2$ it is sometimes referred to as a **qu*d*it**.

Let $\{|i\rangle\}_{i\in I}$ be some fixed orthonormal basis of $\mathcal{H}$. Then, any state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ can be written as a linear combination, we also say: as a **superposition**,

$$|\varphi\rangle = \sum_i \alpha_i |i\rangle$$

of the $|i\rangle$'s, where the α_i's, called **amplitudes**, satisfy

$$\sum_i |\alpha_i|^2 = \sum_{ij} \overline{\alpha}_i \alpha_j \langle i|j\rangle = \langle\varphi|\varphi\rangle = 1\,.$$

In case of a 2-dimensional state space $\mathcal{H}$, we consider a fixed orthonormal basis $\{|0\rangle, |1\rangle\}$ of $\mathcal{H}$, called the **computational basis** (or **Z-basis** or **rectilinear basis**). In case $\mathcal{H} = \mathbb{C}^2$, which we may well assume without loss of generality, the computional basis is given by the canonical basis

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \qquad \text{and} \qquad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix}.$$

A state vector $|\varphi\rangle \in \mathcal{S}(\mathbb{C}^2)$ can then be written as a superposition $|\varphi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ of $|0\rangle$ and $|1\rangle$, with $\alpha_0, \alpha_1 \in \mathbb{C}$ such that $|\alpha_0|^2 + |\alpha_1|^2 = 1$ (see Figure 1.1).

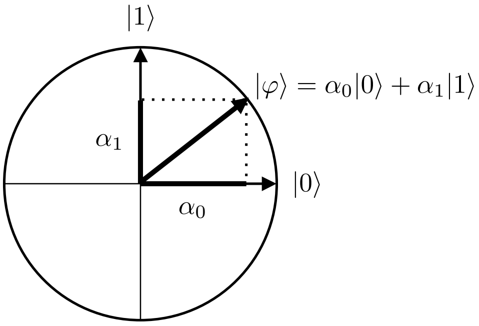

Figure 1.1: A qubit with (real) amplitudes α_0 and α_1.

Another orthonormal basis of the qubit state space $\mathcal{H} = \mathbb{C}^2$ that is important to us is the so-called **Hadamard basis** (or **X-basis** or **diagonal basis**), given by the two basis vectors

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

and

$$|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,.$$

For a state space $\mathcal{H}$ with arbitrary dimension d, the **computational** basis is denoted by $|0\rangle, |1\rangle, \ldots, |d-1\rangle$.[2] We will later see possible generalizations of the Hadamard basis to higher dimensions.

[2] Sometimes, it is more convenient to write the computational basis as $|1\rangle, |2\rangle, \ldots, |d\rangle$ instead.

Strictly speaking, we can — and will — also consider $\mathcal{S}(\mathbb{C}) = \{\omega \in \mathbb{C} \,|\, |\omega| = 1\}$, but then we do not think of or refer to the elements as (state) vectors; instead, an element of $\mathcal{S}(\mathbb{C})$ is called a **phase**. We point out that, together with the multiplication in $\mathbb{C}$, $\mathcal{S}(\mathbb{C})$ forms a group.

Definition 1.2. *Two state vectors $|\varphi\rangle, |\varphi'\rangle \in \mathcal{S}(\mathcal{H})$ are* **equivalent***, denoted as $|\varphi\rangle \equiv |\varphi'\rangle$, if $|\varphi\rangle = \omega|\varphi'\rangle$ for some $\omega \in \mathcal{S}(\mathbb{C})$.*

We will see that two state vectors that are equivalent, or, as we also say, *equal up to a (global) phase*, behave identically under the physically-relevant operations, and thus they describe the same state.

1.2 Unitary Evolution

The natural operation U to apply to a state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$, so as to transform it into another state vector $|\varphi'\rangle = U|\varphi\rangle \in \mathcal{S}(\mathcal{H})$, is a *unitary* $U \in \mathcal{U}(\mathcal{H})$; indeed, unitarity ensures that the norm is preserverd. From the physics perspective, for any unitary $U \in \mathcal{U}(\mathcal{H})$ there exists (in principle) a way to manipulate a quantum system with state space $\mathcal{H}$ so that the (possibly unknown) state $|\varphi\rangle$ of the system evolves from $|\varphi\rangle$ to $|\varphi'\rangle = U|\varphi\rangle$. Vice versa, any physical manipulation of a given system *without causing it to interact with the environment* corresponds to a unitary $U \in \mathcal{U}(\mathcal{H})$.

A particular unitary that we have already (implicitly) encountered is the **Hadamard operator** $H \in \mathcal{U}(\mathbb{C}^2)$, which maps the computational basis into the Hadamard basis, i.e.,

$$H : |0\rangle \mapsto |+\rangle, |1\rangle \mapsto |-\rangle$$

and thus maps any $|\varphi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle \in \mathcal{S}(\mathbb{C}^2)$ to

$$H|\varphi\rangle = \alpha_0 H|0\rangle + \alpha_1 H|1\rangle = \frac{\alpha_0}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) + \frac{\alpha_1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) = \frac{\alpha_0 + \alpha_1}{\sqrt{2}}|0\rangle + \frac{\alpha_0 - \alpha_1}{\sqrt{2}}|1\rangle\,.$$

H is Hermitian, i.e. $H^\dagger = H$, and self-inverse, i.e., $H^2 = \mathbb{I}$, and so it also maps the Hadamard basis back into the computational basis.

Other important examples are the **Pauli operators** (or "gates") $X, Y, Z \in \mathcal{U}(\mathbb{C}^2)$, which act as

$$X : |0\rangle \mapsto |1\rangle, |1\rangle \mapsto |0\rangle\ , \quad Y : |0\rangle \mapsto i|1\rangle, |1\rangle \mapsto -i|0\rangle \quad \text{and} \quad Z : |0\rangle \mapsto |0\rangle, |1\rangle \mapsto -|1\rangle\,.$$

As matrices (with respect to the computational basis) they are

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\,.$$

Like H, they are Hermitian and self-inverse; as a matter of fact, they satisfy

$$X^2 = Y^2 = Z^2 = -iXYZ = iZYX = \mathbb{I}\,.$$

Furthermore,

$$XY = -YX\ , \qquad XZ = -ZX \qquad \text{and} \qquad YZ = -ZY\,.$$

In particular, $\{\pm 1, \pm i\} \cdot \{\mathbb{I}, X, Y, Z\}$ forms a group, the (1-qubit) **Pauli group**.

Another noteworthy example of a (parameterized) 1-qubit unitary is the **phase shift gate**

$$S_\theta : |0\rangle \mapsto |0\rangle, |1\rangle \mapsto e^{i\theta}|1\rangle$$

for an arbitrary $\theta \in \mathbb{R}$, which changes the state by means of a **local** phase. Important special cases of the phase shift gate are $S_\pi = Z$, $S_{\pi/2}$, which is called **phase gate** and denoted by S, and $S_{\pi/4}$, which is sometimes denoted by T and then referred to as ***T*-gate**.[3]

More generally, one may consider *isometries* V, which have the defining property $V^\dagger V = \mathbb{I}$ (but not necessarily $VV^\dagger = \mathbb{I}$). An isometry $V \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ maps a state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ into a state vector $|\varphi'\rangle = V|\varphi\rangle \in \mathcal{S}(\mathcal{H}')$ in a "bigger" Hilbert space, i.e., where $\dim(\mathcal{H}') \geq \dim(\mathcal{H})$. Thus, an isometry captures an operation that turns a quantum system into another one.

1.3 Measurements

Another (physically-relevant) way to act on a state is by means of a *measurement*, formally captured by means of the following definitions.

Definition 1.3. *Let I be an arbitrary non-empty finite set. A family $\mathbf{M} = \{M_i\}_{i\in I}$ of operators $M_i \in \mathcal{L}(\mathcal{H})$ with*

$$\sum_{i\in I} M_i^\dagger M_i = \mathbb{I}$$

is called a **family of measurement operators**, *or simply a* **measurement**. *The set of all such families for a given index set I is denoted $\mathcal{M}eas_I(\mathcal{H})$.*

Definition 1.4. *For any $\mathbf{M} \in \mathcal{M}eas_I(\mathcal{H})$ and any state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$, we define for every $i \in I$:*

$$p_i := \| M_i|\varphi\rangle\|^2 = \langle\varphi|M_i^\dagger M_i|\varphi\rangle \qquad \text{and} \qquad |\varphi^i\rangle := \frac{1}{\sqrt{p_i}} M_i|\varphi\rangle \in \mathcal{S}(\mathcal{H}) \,,$$

with $|\varphi^i\rangle$ undefined in case $p_i = 0$.

Note that $p_i \geq 0$ by definition, and $\sum_i p_i = 1$ by the defining property of a measurement and the fact that $|\varphi\rangle$ has norm 1. So, the p_i's form a *probability distribution*. Also, we see here that changing the phase of the state vector $|\varphi\rangle$ has no effect on the p_i's.

Here, the physical relevance is as follows, referred to as **Born's rule**. Any measurement device, which interacts with a quantum system and produces a measurement outcome (like a number on a monitor), is described by a measurement $\mathbf{M} \in \mathcal{M}eas_I(\mathcal{H})$, where $\mathcal{H}$ is the state space of the system. Using this device to measure a particular system with (possibly unknown) state $|\varphi\rangle$ then has the effect that outcome $i \in I$ is *observed* (i.e. displayed on the monitor) with probability p_i, and the state of the system *collapses* to the **post-measurement state** $|\varphi^i\rangle$.

Motivated by the above, for given state vector $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ and measurement $\mathbf{M} = \{M_i\}_{i\in I}$, we can — and will — speak of "the probability to observe (a particular outcome) i", which is then well defined to be $p_i = \langle\varphi|M_i^\dagger M_i|\varphi\rangle$, or "the probability that the measurement outcome lies in the set T", which is defined as $\sum_{i\in T} p_i$, or "satisfies some property Π", etc.

We can push this further and for instance considering yet another measurement $\mathbf{N} = \{N_j\}_{j\in J}$ that is applied to the post-measurement state that results from the first measurement; we may then speak of "the probability to observe i in the first measurement and j in the second", which is naturally defined to be

$$p_{ij} := p_i \, \langle\varphi^i|N_j^\dagger N_j|\varphi^i\rangle = \langle\varphi|M_i^\dagger N_j^\dagger N_j M_i|\varphi\rangle \,.$$

Here, $p_{j|i} := \langle\varphi^i|N_j^\dagger N_j|\varphi^i\rangle$ can be naturally understood as "the probability of the second measurement producing observation j *conditioned* on the first producing i".

[3] Confusingly, $S_{\pi/4}$ is sometimes also referred to as $\pi/8$ gate; the reason for this is that, up to an unimportant global phase $e^{i\pi/8}$, it is equal to the diagonal matrix with $e^{\pm i\pi/8}$ on its diagonal.

The above also shows that the considered sequential application of the measurements $\mathbf{M}$ and $\mathbf{N}$ has "the same effect" as the one measurement $\{N_jM_i\}_{(i,j)\in I\times J}$, meaning that both induce the same (join) probabilities p_{ij} (and the same post-measurement states $|\varphi^{ij}\rangle$). We leave it as a simple exercise to verify that $\{N_jM_i\}_{(i,j)\in I\times J}$ satisfies Definition 1.3.

Similar results hold when composing a measurement with a unitary operator: we leave it as an exercise to show that a unitary followed by a measurement has the same effect as one suitably chosen measurement, and the same for a measurement followed by a unitary (that may then depend on the measurement outcome).

1.4 Projective Measurements

In this section, we introduce a special yet important class of measurements.

Definition 1.5. *$\mathbf{M} = \{M_i\}_{i\in I} \in \mathcal{M}eas_I(\mathcal{H})$ is called a* **projective** *(or* **Von Neumann***) measurement if M_i is a projection for every $i \in I$. Furthermore, $\mathbf{M}$ is called a* **rank-1 projective** *measurement if every M_i is of the form $M_i = |e_i\rangle\langle e_i|$ with $|e_i\rangle \in \mathcal{S}(\mathcal{H})$.*

The following in particular implies that a rank-1 projective measurement may be described by an orthonormal basis of $\mathcal{H}$.

Lemma 1.1. *If $\{P_i\}_{i\in I}$ is a projective measurement, then the projections P_i are pairwise mutually orthogonal: $P_i\,P_j = 0$ for $i \neq j$. In particular, if $\{P_i\}_{i\in I}$ is a rank-1 projective measurement, and thus $P_i = |e_i\rangle\langle e_i|$ for all $i \in I$, then $\{|e_i\rangle\}_{i\in I}$ is an orthonormal basis of $\mathcal{H}$.*

Proof. Given that $\sum_i P_i^\dagger P_i = \mathbb{I}$ and using the defining properties of projections, we see that for any $j \in I$ and $|\varphi\rangle \in \mathcal{H}$,

$$\langle\varphi|P_j|\varphi\rangle = \langle\varphi|P_j^\dagger P_j|\varphi\rangle = \sum_i \langle\varphi|P_j^\dagger P_i^\dagger P_i P_j|\varphi\rangle = \langle\varphi|P_j|\varphi\rangle + \sum_{i\neq j} \langle\varphi|P_j^\dagger P_i^\dagger P_i P_j|\varphi\rangle$$

and the claim follows from the observation that $\langle\varphi|P_j^\dagger P_i^\dagger P_i P_j|\varphi\rangle = \|P_i P_j|\varphi\rangle\|^2 \geq 0$. □

In the case of such a projective measurement $\mathbf{M} = \{P_i\}_{i\in I}$, Born's rule obviously simplifies to $p_i = \langle\varphi|P_i^\dagger P_i|\varphi\rangle = \langle\varphi|P_i|\varphi\rangle$. In case of a rank-1 projective measurement $\mathbf{M} = \{|i\rangle\langle i|\}_{i\in I}$, we say that "we measure the quantum system *in the basis* $\{|i\rangle\}_{i\in I}$"; here, Born's rule simplifies to

$$p_i = \langle\varphi||i\rangle\langle i||i\rangle\langle i||\varphi\rangle = \langle\varphi|i\rangle\langle i|\varphi\rangle = |\langle i|\varphi\rangle|^2$$

and

$$|\varphi^i\rangle = \frac{1}{\sqrt{p_i}}|i\rangle\langle i||\varphi\rangle = \frac{1}{\sqrt{p_i}}|i\rangle\langle i|\varphi\rangle \equiv |i\rangle$$

Therefore, when writing the state vector $|\varphi\rangle$ as a superposition

$$|\varphi\rangle = \sum_i \alpha_i|i\rangle$$

then the p_i's can easily be obtained from the amplitues as

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

For example, let us consider the rank-1 projective measurement $\mathbf{M} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ given by the computational basis $\{|0\rangle, |1\rangle\}$ of $\mathcal{H} = \mathbb{C}^2$. Then, for $|\varphi\rangle = |0\rangle$ we see immediately that

$p_0 = 1$ and $p_1 = 0$, i.e., the outcome 0 is observed with certainty. Similarly, for $|\varphi\rangle = |1\rangle$ we have $p_0 = 0$ and $p_1 = 1$. Whereas for $|\varphi\rangle = |+\rangle$ and $|\varphi\rangle = |-\rangle$, we get $p_0 = p_1 = \frac{1}{2}$. Correspondingly, when we consider the rank-1 projective measurement given by the Hadamard basis $\{|+\rangle, |-\rangle\}$ and the states $|+\rangle, |-\rangle, |0\rangle, |1\rangle$.

We conclude by remarking that in the literature, one also finds the terminology that an **observable** M is measured of the system A, where M is a Hermitian matrix in $\mathcal{L}(\mathcal{H})$. In our terminology, this corresponds to the projective measurement $\mathbf{M} = \{M_i\}_{i \in I}$ where the M_i's are the orthogonal projections into the eigenspaces of M, and the i's are the corresponding (real) eigenvalues. Vice versa, every projective measurement $\mathbf{M} = \{M_i\}_{i \in I}$ (with $I \subset \mathbb{R}$) may be phrased in terms of an observable M. We do not make use of this terminology/formalism.

1.5 POVMs

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 family $\mathbf{M} = \{M_i\}_{i \in I}$ of measurement matrices can be simplified.

Definition 1.6. *Let $\mathbf{E} = \{E_i\}_{i \in I}$ be a non-empty finite family of matrices $E_i \in \mathcal{L}(\mathcal{H})$. $\mathbf{E}$ is called a* **POVM** *(which stands for a "Positive-Operator Valued Measure") if*

$$E_i \geq 0 \;\; \forall\, i \in I \quad \textit{and} \quad \sum_{i \in I} E_i = \mathbb{I}\,.$$

For a finite index set I, we let $\mathcal{POVM}_I(\mathcal{H})$ denote the set of POVM's $\mathbf{E} = \{E_i\}_{i \in I}$.

The following is a trivial observation. For every measurement $\mathbf{M} = \{M_i\}_{i \in I}$ in $\mathcal{M}eas_I(\mathcal{H})$, the family $\mathbf{E} = \{E_i\}_{i \in I}$ with $E_i = M_i^\dagger M_i$ is in $\mathcal{POVM}_I(\mathcal{H})$, and

$$p_i = \langle\varphi|M_i^\dagger M_i|\varphi\rangle = \langle\varphi|E_i|\varphi\rangle$$

for all $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ and $i \in I$. Hence, every measurement gives rise to a POVM, and the POVM is sufficient to compute p_i. Vice versa, every POVM arises from some measurement:

Lemma 1.2. *For every $\mathbf{E} = \{E_i\}_{i \in I} \in \mathcal{POVM}_I(\mathcal{H})$ there exists a measurement $\mathbf{M} = \{M_i\}_{i \in I}$ such that $E_i = M_i^\dagger M_i$ for every $i \in I$.*

The existence of a decomposition $E_i = M_i^\dagger M_i$ follows immediately from the spectral decomposition (Theorem 0.3) of E_i and the positivity of E_i. For instance, $M_i := \sqrt{E_i}$, defined according to Definition 0.1, does the job. However, we stress that the decomposition $E_i = M_i^\dagger M_i$ is *not* unique in general, and therefore the post-measurement state $|\varphi^i\rangle$ is not uniquely determined by $\mathbf{E}$ (when given $|\varphi\rangle$), but p_i is. As such, the POVM formalism is applicable if we are merely interested in the measurement statistics but not in the post-measurement state.

1.6 Perfect Distinguishability

We consider the following question. If an "experimenter" is given one or another state, how easy or hard is it for him to find out in which state of the two it is by means of performing an arbitrary measurement. This motivates the following terminology. Two state vectors $|\varphi\rangle, |\psi\rangle \in \mathcal{S}(\mathcal{H})$ are called **perfectly distinguishable** if there exists a POVM $\mathbf{E} = \{E_0, E_1\} \in \mathcal{POVM}_{\{0,1\}}(\mathcal{H})$ such that $\langle\varphi|E_0|\varphi\rangle = 1 = \langle\psi|E_1|\psi\rangle$.

Theorem 1.3. *Two state vectors $|\varphi\rangle \in \mathcal{S}(\mathcal{H})$ and $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ are perfectly distinguishable if and only if they are orthogonal, i.e., $\langle\varphi|\psi\rangle = 0$.*

Proof. That orthogonality is a sufficient condition for perfectly distinguishability is obvious: we simply take a measurement that is described by an orthonormal basis that contains $|\varphi\rangle$ and $|\psi\rangle$ as basis vectors, or, more formally, we set

$$E_0 = |\varphi\rangle\langle\varphi| \qquad \text{and} \qquad E_1 = \mathbb{I} - |\varphi\rangle\langle\varphi| \,.$$

Also, it should be clear that it is a necessary condition when restricting to *projective* measurements. For general measurements, we can argue as follows. Using that $E_1 \geq 0$ can be written as $E_1 = M_1^\dagger M_1$, we see that

$$0 = \langle\varphi|E_1|\varphi\rangle = \langle\varphi|M_1^\dagger M_1|\varphi\rangle = \|M_1|\varphi\rangle\|^2$$

and thus $M_1|\varphi\rangle = 0$, and hence also $E_1|\varphi\rangle = 0$. Similarly, $\langle\psi|E_0|\psi\rangle = 0$ implies $E_0|\psi\rangle = 0$. It follows that

$$\langle\varphi|\psi\rangle = \langle\varphi|(E_0 + E_1)|\psi\rangle = \langle\varphi|(E_0 + E_1^\dagger)|\psi\rangle = \langle\varphi|E_0|\psi\rangle + \langle\varphi|E_1^\dagger|\psi\rangle = 0 \,.$$

Hence, $|\varphi\rangle$ and $|\psi\rangle$ must be orthogonal. □

We now see that two states, given by state vectors $|\varphi\rangle, |\psi\rangle \in \mathcal{S}(\mathcal{H})$, are perfectly distinguishable if and only if $\langle\varphi|\psi\rangle = 0$, and they are perfectly *in*distinguishable (in the obvious sense) if and only if $|\varphi\rangle$ and $|\psi\rangle$ are identical up to the phase, i.e., $|\varphi\rangle \equiv |\psi\rangle$, or, equivalenty, $|\langle\varphi|\psi\rangle| = 1$. Thus, we understand the extreme cases. For the cases in-between, the following seems to be a suitable measure for capturing how far away we are from one or the other extreme case.

Definition 1.7. *The* **fidelity** *of two state vectors $|\varphi\rangle, |\psi\rangle \in \mathcal{S}(\mathcal{H})$ is defined as*[4]

$$F(|\varphi\rangle, |\psi\rangle) := |\langle\varphi|\psi\rangle| \,.$$

Indeed, it turns out that the **distingushing advantage** of two states $|\varphi\rangle, |\psi\rangle \in \mathcal{S}(\mathcal{H})$, defined as

$$adv(|\varphi\rangle, |\psi\rangle) := \max_{0 \leq E_0 \leq \mathbb{I}} \big(\langle\varphi|E_0|\varphi\rangle - \langle\psi|E_0|\psi\rangle\big) = \max_{0 \leq E_0 \leq \mathbb{I}} \langle\varphi|E_0|\varphi\rangle + \langle\psi|(\mathbb{I} - E_0)|\psi\rangle - 1 \,,$$

is determined by the fidelity:

$$adv(|\varphi\rangle, |\psi\rangle) = \sqrt{1 - F(|\varphi\rangle, |\psi\rangle)^2} \,.$$

The fidelity should be thought of as a measure of distance, but obviously it is not a metric in the mathematical sense; in particular, small fidelity means that the states are far away, and large fidelity (i.e., a fidelity close to 1) means that the states are close to each other. The distingushing advantage, however, turns out to be a metric, the so called **trace distance**.

[4] Be aware: in some literature, the fidelity is defined as $|\langle\varphi|\psi\rangle|^2$.

1.7 The Bloch Sphere

The Bloch sphere, which we introduce here, offers a nice geometrical description of the space of qubits *modulo the (irrelevant) phase*. First, we observe that, for any Hilbert space $\mathcal{H}$, the map

$$\mathcal{S}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})\,,\, |\varphi\rangle \mapsto |\varphi\rangle\langle\varphi|$$

induces an *injection* on $\mathcal{S}(\mathcal{H})/{\equiv}$, i.e., the set of equivalence classes $\{\omega|\varphi\rangle \,|\, \omega \in \mathcal{S}(\mathbb{C})\}$.

Given that the global phase of a state vector is irrelevant, $\rho := |\varphi\rangle\langle\varphi|$ can thus be understood as a description of the state in terms of an operator, which, in contrast to the state-vector description, is *unique*; indeed, this is how states are described in the so-called **density operator** formalism. Note that by construction, ρ is positive-semidefinite (and thus Hermitian) and has trace $\mathrm{tr}(\rho) = \mathrm{tr}(|\varphi\rangle\langle\varphi|) = \langle\varphi|\varphi\rangle = 1$.

We now focus on $\mathcal{H} = \mathbb{C}^2$. It is not too hard to see that $\{\mathbb{I}, X, Y, Z\}$ forms an $\mathbb{R}$-basis of the Hermitian operators in $\mathcal{L}(\mathbb{C}^2)$, and therefore $\rho = |\varphi\rangle\langle\varphi|$ must be of the form

$$\rho = \frac{1}{2}\big(w\mathbb{I} + xX + yY + zZ\big)\ ,$$

for real-valued $w, x, y, z \in \mathbb{R}$, by applying the trace, which evaluates to 1 on the left and to w on the right hand side, we see that $w = 1$. Additionally, the fact that $\rho^2 = |\varphi\rangle\langle\varphi|\varphi\rangle\langle\varphi| = \rho$ implies that

$$\begin{aligned}\frac{1}{2}\big(\mathbb{I} + xX + yY + zZ\big) &= \frac{1}{4}\big(\mathbb{I} + xX + yY + zZ\big)^2 \\ &= \frac{1}{4}\big(\mathbb{I} + x^2\mathbb{I} + y^2\mathbb{I} + z^2\mathbb{I} + 2xX + 2yY + 2zZ + xy\{X,Y\} + xz\{X,Z\} + yz\{Y,Z\}\big) \\ &= \frac{1}{4}\big(1 + x^2 + y^2 + z^2\big)\mathbb{I} + \frac{1}{2}\big(xX + yY + zZ\big)\,,\end{aligned}$$

where we exploited that the **anticommutator** $\{X,Y\} := XY + YX$ of any two distinct Paulis vanishes. It follows that $x^2 + y^2 + z^2 = 1$. Thus, we obtain an injective map

$$\mathcal{S}(\mathbb{C}^2)/{\equiv}\ \to \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 = 1\}$$

into the real 2-sphere, which is then referred to as **Bloch sphere**. This map is also surjective; this follows by observing that the determinant of $\rho = \frac{1}{2}\big(\mathbb{I} + xX + yY + zZ\big)$ vanishes for points on the Bloch sphere, while its trace is 1, implying that ρ has eigenvalues 0 and 1 (since ρ is Hermitian, it has a spectral decomposition; see Section 0.3), which in turn implies that it is of the form $\rho := |\varphi\rangle\langle\varphi|$ for $|\varphi\rangle \in \mathcal{S}(\mathbb{C}^2)$ being an eigenvector for the eigenvalue 1. Thus, we can identify qubit states with points on the Bloch sphere and vice versa.

From the observation that

$$|0\rangle\langle 0| = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} = \frac{1}{2}(\mathbb{I} + Z) \qquad \text{and} \qquad |1\rangle\langle 1| = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = \frac{1}{2}(\mathbb{I} - Z)$$

we see that the computational basis vectors $|0\rangle$ and $|1\rangle$ correspond to the points $z = \pm 1$ on the Block sphere (see Figure 1.2). Similarly, the Hadamard basis vectors $|+\rangle$ and $|-\rangle$ correspond to $x = \pm 1$, while the points $y = \pm 1$ represent the vectors that form the **circular basis**, given by

$$|\circlearrowleft\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \qquad \text{and} \qquad |\circlearrowright\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\,.$$

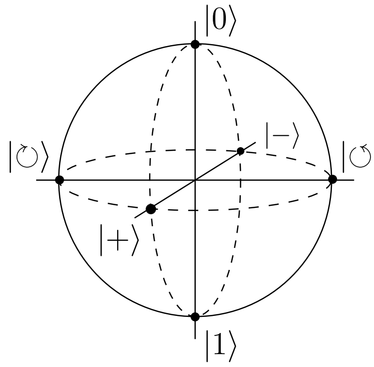

Figure 1.2: The Bloch sphere.

We can now view any $U \in \mathcal{U}(\mathbb{C}^2)$ as a function on the Bloch sphere; more formally, by "pulling back", applying U, and mapping to the Bloch sphere again, we obtain a group homomorphism from $\mathcal{U}(\mathbb{C}^2)$ to the functions on the Bloch sphere. For instance, from the basic properties of the Pauli operators (like $XZX = -Z$ etc.), we immediately see that

$$X|\varphi\rangle\langle\varphi|X = \frac{1}{2}X\big(\mathbb{I} + xX + yY + zZ\big)X = \frac{1}{2}\big(\mathbb{I} + xX - yY - zZ\big)\,,$$

which, e.g., shows that X maps $|0\rangle$ to $|1\rangle$ and vice versa (which of course we already knew), maps $|\circlearrowleft\rangle$ to $|\circlearrowright\rangle$ and vice versa, and leaves $|+\rangle$ and $|-\rangle$ untouched (modulo the phase!). Thus, as an action on the Bloch sphere, X is a rotation by 180° around the axis given by $|+\rangle$ and $|-\rangle$. Correspondingly for Y and Z.

More generally, we consider the following unitaries, one for each Pauli operator and parameterized by $\theta \in \mathbb{R}$.

$$\begin{aligned}
R_X(\theta) &:= \cos\big(\tfrac{\theta}{2}\big)\mathbb{I} - i\sin\big(\tfrac{\theta}{2}\big)X = \begin{bmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2}) \\ -i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix} \\
R_Y(\theta) &:= \cos\big(\tfrac{\theta}{2}\big)\mathbb{I} - i\sin\big(\tfrac{\theta}{2}\big)Y = \begin{bmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix} \\
R_Z(\theta) &:= \cos\big(\tfrac{\theta}{2}\big)\mathbb{I} - i\sin\big(\tfrac{\theta}{2}\big)Z = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}\,.
\end{aligned}$$

It is easy to see that these are indeed unitaries with $R_{X/Y/Z}(\theta)^\dagger = R_{X/Y/Z}(-\theta)$ and

$$\begin{aligned}
R_Z(\theta)ZR_Z(\theta)^\dagger &= Z \quad , \\
R_Z(\theta)XR_Z(\theta)^\dagger &= \cos(\theta)X + \sin(\theta)Y \quad \text{and} \\
R_Z(\theta)YR_Z(\theta)^\dagger &= -\sin(\theta)X + \cos(\theta)Y\,.
\end{aligned}$$

Therefore, for any $|\varphi\rangle \in \mathcal{S}(\mathbb{C}^2)$ with Bloch-sphere coordinates (x, y, z), we have

$$\begin{aligned}
R_Z(\theta)|\varphi\rangle\langle\varphi|R_Z(\theta)^\dagger &= \frac{1}{2}\big(\mathbb{I} + xR_Z(\theta)XR_Z(\theta)^\dagger + yR_Z(\theta)YR_Z(\theta)^\dagger + zR_Z(\theta)ZR_Z(\theta)^\dagger\big) \\
&= \frac{1}{2}\big(\mathbb{I} + (x\cos(\theta) - y\sin(\theta))X + (x\sin(\theta) + y\cos(\theta))Y + zZ\big)\,,
\end{aligned}$$

and thus $R_Z(\theta)|\varphi\rangle$ has Bloch-sphere coordinates $(x\cos(\theta) - y\sin(\theta), x\sin(\theta) + y\cos(\theta), z)$. In other words, $R_Z(\theta)$ acts as a *rotation* around the z axis, and similarly for $R_X(\theta)$ and $R_Y(\theta)$. Because of this, these are called **rotation operators**. Also, it is easy to see that

$$ZR_Z(\theta)Z = R_Z(\theta) \quad , \qquad XR_Z(\theta)X = R_Z(-\theta) \quad \text{and} \qquad YR_Z(\theta)Y = R_Z(-\theta)$$

and similarly for $R_X(\theta)$ and $R_Y(\theta)$.

We conclude with the following characterization theorem for single-qubit unitaries.

Theorem 1.4 (Z-Y decomposition)**.** *For any $U \in \mathcal{U}(\mathbb{C}^2)$ there exist $\alpha, \beta, \gamma, \delta$ such that*

$$U = e^{i\alpha} R_Z(\beta) R_Y(\gamma) R_Z(\delta)\,.$$

This in particular implies that any single qubit unitary corresponds to a sequence of rotations of the Block sphere around the y- and z-axes, and thus to a single rotation around some axis. In mathematical terms, identifying U with its action on the Bloch sphere gives rise to a surjective homomorphism $SU(2) \to SO(3)$ from the special unitary group of degree 2 to the special orthogonal group of degree 3. The kernel of this homomorphism is $\{\pm\mathbb{I}\}$. This in turn gives rise to an isomorphism $SU(2)/\{\pm\mathbb{I}\} \leftrightarrow SO(3)$.

Proof. We note that

$$e^{i\theta/2} R_Z(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} \,,$$

which means that multiplying $e^{i\theta/2} R_Z(\theta)$ from the left to U has the (only) effect that the second row of U gets multiplied with the phase $e^{i\theta}$, while if it is multiplied from the right then the second column gets multiplied by the phase. Similarly for $e^{-i\theta/2} R_Z(\theta)$, but then for the first row/column, and with phase $e^{-i\theta}$.

Furthermore, by multiplying rows and columns of U with appropriate phases, and thus by appropriate choices of $\alpha, \beta, \delta \in \mathbb{R}$, we can obtain that U becomes

$$e^{-i\alpha} R_Z(-\delta)\, U R_Z(-\beta) = \begin{bmatrix} |u_{11}| & \omega |u_{12}| \\ |u_{21}| & |u_{22}| \end{bmatrix} =: U'$$

for some phase $\omega \in \mathcal{S}(\mathbb{C})$. Since U' is still unitary, and so its two columns are orthonormal, it holds that $\omega = -1$, or else either of u_{11} and u_{12} vanishes, in which case we may as well assume $\omega = -1$, by changing the phase of the first row in case $u_{11} = 0$. But then, with $\omega = -1$ and hence the two rows of U' being two orthonormal vectors in $\mathbb{R}^2$, $U' = R_Y(\gamma)$ for some $\gamma \in \mathbb{R}$. □