Part II

Quantum Computing

Chapter 3

Foundations of Quantum Computing

Quantum computation investigates the computational power of hypothetical computing devices that make use of quantum-mechanical properties, as introduced and discussed in previous chapters. An important objective is to find quantum algorithms that are significantly faster than any classical algorithm solving the same problem. The field started in the early 1980s with suggestions for analog quantum computers by Yuri Manin, Richard Feynman, and Paul Benioff, and reached more digital ground when in 1985 David Deutsch defined the universal quantum Turing machine. The following years saw only sparse activity, notably the development of the first algorithms by Deutsch and Jozsa and by Simon, and the development of quantum complexity theory by Bernstein and Vazirani. Interest in the field then increased tremendously after Peter Shor's 1994 discovery of his famous quantum algorithms for factoring large integers and for computing discrete logarithms.

In this chapter, we introduce some of the early quantum algorithms, and we cover the theoretical foundations by discussing the quantum circuit model of computation. We end the chapter with Grover's algorithm for unstructured search, a quantum algorithm that is maybe not as impressive in terms of speed-up (like Shor's algorithms) but is important due to the generality of the computational problem it solves.

In this chapter, if not specified otherwise, we restrict $\mathcal{H}$ to be $\mathcal{H} = \mathbb{C}^2$, the state space of a qubit. For any positive integer n, $\mathcal{H}^{\otimes n}$ stands for the n-fold tensor product $\mathcal{H} \otimes \ldots \otimes \mathcal{H}$ of $\mathcal{H}$ with itself. Similarly, for $U \in \mathcal{U}(\mathcal{H})$, $U^{\otimes n}$ is the n-fold tensor product $U \otimes \ldots \otimes U \in \mathcal{U}(\mathcal{H}^{\otimes n})$. Throughout, we consider the computational basis $\{|0\rangle, |1\rangle\}$ of $\mathcal{H}$ as well as the computational basis $\{|x\rangle\}_{x \in \{0,1\}^n}$ of $\mathcal{H}^{\otimes n}$, where $|x\rangle = |x_1, \ldots, x_n\rangle = |x_1\rangle \cdots |x_n\rangle$ for $x = (x_1, \ldots, x_n) \in \{0,1\}^n$.

3.1 Warm-up: Deutsch's Algorithm

Consider a binary function $f : \{0,1\} \to \{0,1\}$. We imagine a situation where f is not given to us by its function table, but, say, in the form of a very complicated and extremely inefficient algorithm. As such, we can learn $f(x)$ for any $x \in \{0,1\}$ by *computing* $f(x)$ using the algorithm, but there is no "shortcut". One typically speaks of **black-box access** then: the only way to learn (anything about) the function value of any input x is by means of making a **query** to an "oracle", which then provides the correct function value $f(x)$.

The task here now is to find out if $f(0) = f(1)$ or not. Obviously, this can be done by computing f *twice*, i.e., by making two queries — one for input 0 and one for input 1. The question is whether one can do better. It is intuitively quite clear, and not too hard to prove once rigorously formalized, that it is impossible to do any better with a classical algorithm: any classical algorithm with black-box access to f that only makes *one* query to f cannot predict whether $f(0) = f(1)$ or not with probability larger than $\frac{1}{2}$, when f is chosen uniformly at random from all functions $\{0,1\} \to \{0,1\}$. On the other hand, somewhat surprisingly, a *quantum* algorithm with black-box *quantum access* to f can do better. The latter means that the algorithm can make queries to the unitary $U_f \in \mathcal{U}(\mathcal{H}\otimes\mathcal{H})$ given by $U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$ for $x, y \in \{0,1\}$, where $\oplus$ is the binary XOR, i.e., addition mod 2.

Proposition 3.1. *There exists a quantum algorithm so that when given black-box access to U_f it makes a single query to U_f and outputs the bit $f(0) \oplus f(1)$ with certainty.*

Note that we do not have a formal notion of a "quantum algorithm" yet, but it should be clear that what we outline in the proof counts as one. It is called **Deutsch's algorithm**, named after David Deutsch. Before describing the algorithm, let us first look into the naive approach. Given that the desired quantum algorithm has access to U_f, which means that he can apply f to a *superposition* of inputs, that seems to be the way to go: form the superposition $(|0\rangle + |1\rangle)/\sqrt{2}$ and apply U_f to obtain

$$U_f \, \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle = \frac{1}{\sqrt{2}}|0\rangle|f(0)\rangle + \frac{1}{\sqrt{2}}|1\rangle|f(1)\rangle \, .$$

This indeed gives us a state that depends on *both*, $f(0)$ and $f(1)$. However, it is unclear how to now *extract information* on both, respectively on $f(0) \oplus f(1)$. If we measure the first qubit in the computational bases, then the second collapses to $|f(0)\rangle$ or $|f(1)\rangle$, and all information on the other is lost, and so we would still need another call to f. Also measuring the second qubit in the computational basis will not provide any information on $f(0) \oplus f(1)$.

Proof. The algorithm starts with the 2-qubit state

$$|+\rangle|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}}|0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) + \frac{1}{\sqrt{2}}|1\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \, .$$

and applies U_f. This yields

$$\begin{aligned} &\frac{1}{2}|0\rangle \otimes \big(|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle\big) \; + \; \frac{1}{2}|1\rangle \otimes \big(|0 \oplus f(1)\rangle - |1 \oplus f(1)\rangle\big) \\ &\qquad = \frac{1}{2}(-1)^{f(0)} \, |0\rangle \otimes \big(|0\rangle - |1\rangle\big) \; + \; \frac{1}{2}(-1)^{f(1)} \, |1\rangle \otimes \big(|0\rangle - |1\rangle\big) \\ &\qquad = \frac{1}{\sqrt{2}}\Big(|0\rangle + (-1)^{f(0)\oplus f(1)}|1\rangle\Big) \otimes \frac{1}{\sqrt{2}}(-1)^{f(0)}\big(|0\rangle - |1\rangle\big) \, . \end{aligned}$$

which is still a product state, and it has the predicate we are looking for encoded into the phase of the first qubit. Concretely, the first qubit is in state $|+\rangle$ if $f(0) = f(1)$, and it is in state $|-\rangle$ if $f(0) \neq f(1)$. Hence, by now measuring the first qubit in the Hadamard basis, or, equivalently, by applying the Hadamard operator and measuring in the computational basis, we obtain the correct result. □

Clearly, Deutsch's algorithm does not seem very relevant nor impressive from a practical perspective. Still, it nicely shows, in a simple way, how it is still possible to exploit the possibility of applying a function f to a quantum superposition of inputs, despite the observation that the naive approach does not give you anything. The idea here is to bring the function values from the basis vectors into the amplitudes, so that one gets *constructive* or *distructive interference*, depending on the function values. Here, this is achieved by applying U_f to $|x\rangle|-\rangle$ for $x \in \{0,1\}$, rather than to $|x\rangle|0\rangle$, so that we get

$$U_f|x\rangle|-\rangle = \frac{1}{\sqrt{2}}\big(|x\rangle|f(x)\rangle - |x\rangle|1 \oplus f(x)\rangle\big) = (-1)^{f(x)}|x\rangle|-\rangle \, .$$

This technique is typically referred to as **phase kickback**. Put differently, we use the fact that U_f has eigenvector $|x\rangle|-\rangle$ with eigenvalue $(-1)^{f(x)}$. Effectively, this gives access to a unitary $V_f \in \mathcal{U}(H)$ with $V_f|x\rangle = (-1)^{f(x)}|x\rangle$, and we now easily see that if $f(0) = f(1)$ then $V_f|+\rangle \equiv |+\rangle$ and if $f(0) \neq f(1)$ then $V_f|+\rangle \equiv |-\rangle$, and so we can distinguish the two cases with one call to V_f.

3.2 More Examples: Deutsch-Jozsa and Bernstein-Vazirani

Deutsch's algorithm is a special case of the **Deutsch-Jozsa algorithm**. The latter considers a function $f : \{0,1\}^n \to \{0,1\}$ with the promise that f is either constant or balanced, where the latter means that $|\{x \,|\, f(x) = 0\} = 2^{n-1}$, and the goal is to find out which of the two is the case. Classically, this requires $2^{n-1} + 1$ queries to f in the worst case, and k queries to get the right answer except with probability 2^{-k+1}. With a quantum algorithm, one query suffices.

Proposition 3.2. *For any positive integer $n \in \mathbb{N}$ there exists a quantum algorithm so that when given black-box access to $U_f \in \mathcal{U}(\mathcal{H}^{\otimes n} \otimes \mathcal{H})$ it makes a single query to U_f and predicts with certainty whether $f : \{0,1\}^n \to \{0,1\}$ is constant or balanced, given that it is one of the two.*

For the analysis of the algorithm, it will be convenient to have the following lemma at hand; we leave its proof as an exercise. $H^{\otimes n}$ is sometimes also called **Walsh-Hadamard transform**.

Lemma 3.3. *For any $x = (x_1, \ldots, x_n) \in \{0,1\}^n$,*

$$H^{\otimes n}|x\rangle = H|x_1\rangle \otimes \ldots \otimes H|x_n\rangle = \frac{1}{2^{n/2}} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y}|y\rangle$$

where $x \cdot y = x_1 y_1 \oplus \ldots \oplus x_n y_n \in \{0,1\}$.

Proof (of Proposition 3.2). The algorithm follows closely Deutsch's algorithm. It starts off with the $(n+1)$-qubit state $|+\rangle^{\otimes n} \otimes |-\rangle$, which equals

$$H^{\otimes n}|0,\ldots,0\rangle \otimes H|1\rangle = \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

by Lemma 3.3, and it applies U_f. This yields

$$\frac{1}{2^{n/2}} \sum_x |x\rangle \otimes \frac{1}{\sqrt{2}}\big(|f(x)\rangle - |1 \oplus f(x)\rangle\big) = \frac{1}{2^{n/2}} \sum_x (-1)^{f(x)}|x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,.$$

Again, we ignore the last qubit, which is in "product form", and apply $H^{\otimes n}$ to the first n qubits. This results in the n-qubit state

$$\frac{1}{2^{n/2}} \sum_x (-1)^{f(x)} \frac{1}{2^{n/2}} \sum_y (-1)^{x \cdot y}|y\rangle = \frac{1}{2^n} \sum_y \Bigg(\sum_x (-1)^{f(x)}(-1)^{x \cdot y} \Bigg)|y\rangle\,,$$

using again Lemma 3.3. Measuring this state in the computational basis $\{|y\rangle\}_{y \in \{0,1\}^n}$, we see that measurement outcome $y = (0,\ldots,0)$ is observed with probability

$$p_{(0,\ldots,0)} = \left| \frac{1}{2^n} \sum_x (-1)^{f(x)} \right|^2,$$

which is 1 if f is constant and 0 if f is balanced. □

The same algorithm can also be used for finding $s \in \{0,1\}^n$ when given black-box quantum access to the function $f_s : \{0,1\}^n \to \{0,1\}$, $x \mapsto s \cdot x$. In this context, it is then referred to as the **Bernstein-Vazirani algorithm**.

Proposition 3.4. *For any $n \in \mathbb{N}$ there is a quantum algorithm so that when given black-box access to $U_{f_s} \in \mathcal{U}(\mathcal{H}^{\otimes n} \otimes \mathcal{H})$ it makes a single query to U_{f_s} and outputs s with certainty.*

Proof. We consider the n-qubit state obtained by means of the Deutsch-Jozsa algorithm:

$$\frac{1}{2^n}\sum_{x,y}(-1)^{f_s(x)}(-1)^{x\cdot y}|y\rangle = \frac{1}{2^n}\sum_{y}\Bigg(\sum_{x}(-1)^{x\cdot(s\oplus y)}\Bigg)|y\rangle \,,$$

and we observe that the measurement outcome s is observed with probability 1. □

3.3 Quantum Algorithms and Complexity

We want to formalize the notion of a *quantum algorithm* and of the *complexity* of such an algorithm. Indeed, what is exciting about quantum computation is that, for certain computational problems, it allows for algorithms with (much) better complexity compared to classical models of computation. In the above examples, this was demonstrated for the notion of **query complexity**, i.e., when one considers algorithms that make black-box queries to some (partly) unknown "oracle" (typically a function), and we count the number of queries necessary to perform the desired computation. What is nice about this notion of complexity is that it allows for provable lower bounds, and thus for provable separation results between classical and quantum computation.

But maybe more relevant from a practical perspective is the notion of **computational complexity**, which counts the number of "elementary steps" that the algorithm applies to the input. Here, an "elementary step" would be an "elementary unitary operation", which in this context is then referred to as a **gate**.

In order to formalize the above, let $\mathcal{G}$ be a non-empty set of such gates, i.e., of unitary operators, such that for every $G \in \mathcal{G}$ we have that $G \in \mathcal{U}(\mathcal{H}^{\otimes k})$ for some $k \leq n$, called the **arity** of G. For these gates to be "elementary" we will later require the arity k for every $G \in \mathcal{G}$ to be small, like at most 2.

In line with previous observations, a gate $G \in \mathcal{U}(\mathcal{H}^{\otimes k})$ with arity k can act on an n-qubit state (vector); but it then needs to be specified on which k of the n qubits, and which component of G acts on which of the designated k qubits. By default, we label the n qubits by "1","2",…,"n", and then write, say, $G_{7,2,8,5}$ to specify that the first component of $G \in \mathcal{U}(\mathcal{H}^{\otimes 4})$ acts on qubit "7", the second on qubit "2", etc. Similarly, $CNOT_{4,1}$ then refers to the CNOT gate that is controlled by qubit "4" and has target qubit "1". However, we will make little use of this notation; instead, we will mainly use pictorial descriptions (see e.g. Figure 3.1).

We can now define the following model of quantum computation, which is a notion of a quantum algorithm with *quantum input* and *quantum output.*

Definition 3.1. *An n-qubit* **quantum circuit** *with gate set $\mathcal{G}$ consists of a finite sequence $[U_1, \dots, U_t]$ of unitaries $U_i \in \mathcal{U}(\mathcal{H}^{\otimes n})$, where, for every $i \in \{1, \dots, t\}$, U_i is of the form $U_i = G_W$ with $G \in \mathcal{G}$ and $W = (w_1, \dots, w_k)$ a sequence of k distinct numbers in $\{1, \dots, k\}$, where k is the arity of G. The* **computational complexity** *of such a quantum circuit is given by t.*

Remark 3.1. Sometimes, we consider a relaxation of the the above definition where $\mathcal{H}^{\otimes n}$ is replaced by the n-fold tensor product $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ of possibly non-qubit state spaces.

A quantum circuit can be nicely depictured with "wires" and with "gates" that act on the "wires", as we have already been doing, e.g. in Section 2.6. Figure 3.1 shows such a pictorial representation of an example quantum circuit with gate set $\mathcal{G} = \{H, \mathit{CNOT}\}$.

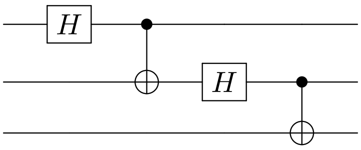

Figure 3.1: Pictorial representation of the quantum circuit $H_1, \mathit{CNOT}_{1,2}, H_2, \mathit{CNOT}_{2,3}$.

Definition 3.2. *An n-qubit quantum circuit $U_1, \dots, U_t$* **computes** *a unitary $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ if $U = U_t \cdots U_1$; it ε-***approximately computes** *U for $0 \leq \varepsilon \leq 1$ if $\|U - U_t \cdots U_1\| \leq \varepsilon$.*
More generally, for $n' \geq n$, an n'-qubit quantum circuit $U_1, \dots, U_t$ is said to **(approximately) compute** *an n-qubit unitary $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ if $U_t \cdots U_1|\varphi\rangle|0\rangle^{\otimes(n'-n)}$ equals (or approximates) $U|\varphi\rangle \otimes |0\rangle^{\otimes(n'-n)}$ for any $|\varphi\rangle \in \mathcal{S}(\mathcal{H}^n)$.*

The choice of the norm for measuring the distance between the unitaries is not so important; for concreteness, we take the operator norm[1] here. Similarly for the second part of the definition, where we take the norm induced by the inner product, but the fidelity is also a common choice.

Remark 3.2. Repeating Remark 2.1, we stress that the convention is to apply such a quantum circuit as depictured in Figure 3.1 *from the left to the right*, i.e., first H is applied and then *CNOT* etc., whereas the composition $U = U_t \cdots U_1$ is applied to a state $|\varphi\rangle$ *from right to left.*

Looking back to the decomposition of (multi-)control unitaries, the right hand side of Figure 2.4 can now be formally understood as (a pictorial representation of) a quantum circuit with gate set $\mathcal{G} = \{V, V^\dagger, \mathit{CNOT}\}$ that computes $C^2(U)$, and the right hand side of Figure 2.5 as a quantum circuit, with $\mathcal{G}$ consisting of *CNOT* and the Toffoli gate, that computes $C^n(U)$.

Not so surprising, quantum computing is at least as powerful than classical computing; in particular, once everything is formalized, it is not hard to see that the following holds.

Theorem 3.5. *For any set $\mathcal{G}$ of logic gates, like $\{\wedge, \neg\}$, let $f : \{0,1\}^{m_{in}} \to \{0,1\}^{m_{out}}$ be a function that can be classically computed by a classical circuit with gate set $\mathcal{G}$ and computational complexity c. Then, there exists an n-qubit quantum circuit that computes U_f with gate set $\{U_g \,|\, g \in \mathcal{G}\}$ and computational complexity t, where $t \leq 2c$ and $n \leq m_{in} + m_{out} + c$.*

The factor 2 blow-up in complexity comes from the fact that one has to undo all gates that produce an intermediary (qu)bit, in order to revert the state back to $|0\rangle$.

We will mainly be interested in quantum algorithms with *classical input* and *classical output.* Such a notion can be easily obtained by modifying Definition 3.2 as follows.

Definition 3.3. *For any function $f : \{0,1\}^{m_{in}} \to \{0,1\}^{m_{out}}$ with $0 < m_{in}, m_{out} \leq n$, we say that an n-qubit quantum circuit ε-***approximately computes** *f if*

$$\left\| \big(|f(x)\rangle\langle f(x)| \otimes \mathbb{I}\big)\, U_t \cdots U_1 |x, 0\rangle \right\|^2 \geq 1 - \varepsilon$$

for all $x \in \{0,1\}^{m_{in}}$, where we understand that $|0\rangle \in \mathcal{S}\big(\mathcal{H}^{\otimes(n-m_{in})}\big)$ and $\mathbb{I} \in \mathcal{U}\big(\mathcal{H}^{\otimes(n-m_{out})}\big)$.

[1] $\|A\| = \max_{|\varphi\rangle} \|A|\varphi\rangle\|$ where the max is over all$|\varphi\rangle \in \mathcal{S}(\mathcal{H}^{\otimes n})$.

In other words, the algorithm proceeds by encoding the input into a quantum state, appending an ancilla register, running the quantum circuit, and then measuring (part of) the resulting state. The approximation parameter ε captures the probability of an incorrect outcome.

Sometimes, it will also be convenient to allow some classical "post-processing" of the measurement result obtained after the application of the quantum algorithm; again, this is without loss of generality since, by Theorem 3.5, such a classical "post-processing" could be incorporated into the quantum circuit. Note that as soon as $\varepsilon < 1/2$, one can amplify the success probability by repeating the algorithm.

Finally, we point out that we can easily extend the above to a notion of quantum algorithm *with black-box access* to a non-specified unitary $O \in \mathcal{U}(\mathcal{H}^{\otimes k})$ with a given arity k: we simply extend the set of gates $\mathcal{G}$ to $\mathcal{G} \cup \{O\}$, meaning that the U_i's may also be instructions to apply O to k of the qubits. The **query complexity** is then defined to be the number of i's for which U_i is an instruction to apply O. Unless specified differently, such a quantum algorithm with black-box access acts on the fixed input state $|0\rangle \in \mathcal{S}(\mathcal{H}^{\otimes n})$, and the classical output is obtained by measuring (a specified subset of) the resulting qubits. Once O is instantiated with a specific unitary, we can then make statements about the statistics of the measurement outcome. This finally puts the statements of Propositions 3.1, 3.2 and 3.4 on firm theoretical grounds. See Figure 3.2 below for the quantum circuit for Deutsch's algorithm, where we have "stacked" the two H gates that can be applied in parallel, and the last gate on the upper wire is a measurement in the computational basis.

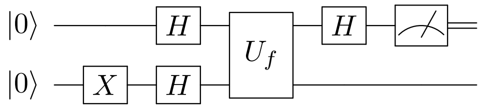

Figure 3.2: Quantum circuit for Deutsch's algorithm.

3.4 Universal Gate Sets

For the notion of quantum circuits to be *complete* as a model of computation, we need a gate set $\mathcal{G}$ that, in principle, enables to compute any unitary.

Definition 3.4. *A (possibly infinite) set of gates $\mathcal{G}$ is called* **perfectly universal** *if for any $n \in \mathbb{N}$ and any unitary $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ there exists an n-qubit quantum circuit with gate set $\mathcal{G}$ that computes $e^{i\alpha}U$ for some $\alpha \in \mathbb{R}$.*[2]
$\mathcal{G}$ is called **approximately universal** *if for any $n \in \mathbb{N}$, $\varepsilon > 0$ and $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ there exists an n-qubit quantum circuit with gate set $\mathcal{G}$ that ε-approximately computes $e^{i\alpha}U$ for some $\alpha \in \mathbb{R}$*

The freedom in the phase is motivated by the fact that such a global phase has no noticable effect. As a first step towards obtaining a universal gate set, we show that any unitary can be decomposed into *two-level* unitaries, defined as follows.

Definition 3.5. *For an arbitrary Hilbert space $\mathcal{H}$ with orthonormal basis $\{|i\rangle\}_{i \in I}$, a unitary $U \in \mathcal{U}(\mathcal{H})$ is said to be* **two-level** *(w.r.t. $\{|i\rangle\}_{i \in I}$) if $U|i\rangle = |i\rangle$ for* all but two *choices of $i \in I$.*

In other words, U acts non-trivially only on (at most) two basis vectors $|k\rangle$ and $|\ell\rangle$. It is then easy to see that $U|k\rangle = u_{11}|k\rangle + u_{21}|\ell\rangle$ and $U|\ell\rangle = u_{12}|k\rangle + u_{22}|\ell\rangle$, where $\tilde{U} := \left[\begin{smallmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{smallmatrix}\right] \in \mathcal{U}(\mathbb{C}^2)$, and we write $U = \mathbb{I} \oplus \tilde{U}$. It is also easy to see that $U^\dagger$ is two-level if U is, namely $U^\dagger = \mathbb{I} \oplus \tilde{U}^\dagger$.

[2] In the literature, the terminology *universal* is somewhat ambiguously used for different variations, like for *approximately* or *perfectly* universal, or when allowing the quantum circuit to be over more than n qubits.

Proposition 3.6. *For an arbitrary Hilbert space $\mathcal{H}$ and for any $U \in \mathcal{U}(\mathcal{H})$, there exists two-level unitaries $U_1, \dots, U_N \in \mathcal{U}(\mathcal{H})$ (w.r.t. any fixed basis) so that $U = U_1 \cdots U_N$.*

Proof. If $U = \mathbb{I}$ then there is nothing to prove, so we consider $U \neq \mathbb{I}$. Thus, there exists a maximal k (for any fixed ordering) so that $x := \langle k|U|k\rangle \neq 1$. We consider now two cases.

Case 1: $\langle k|U|\ell\rangle = 0$ for all $\ell \neq k$. Here, we note that

$$1 = \langle k|k\rangle = \langle k|UU^\dagger|k\rangle = \sum_\ell \langle k|U|\ell\rangle\langle \ell|U^\dagger|k\rangle = \sum_\ell |\langle k|U|\ell\rangle|^2 = |\langle k|U|k\rangle|^2 = |x|^2 ,$$

and then consider the two-level (actually one-level) unitary V that acts non-trivially only on $|k\rangle$, and does so as $V|k\rangle = \bar{x}|k\rangle$. We then set $U' := UV$ and observe that $\langle k|U'|k\rangle = \langle k|U|k\rangle\bar{x} = x\bar{x} = 1$, while still $\langle k'|U'|k'\rangle = \langle k'|U|k'\rangle = 1$ for all $k' > k$.

Case 2: $\langle k|U|\ell\rangle \neq 0$ for some $\ell \neq k$. Let ℓ maximal and $y := \langle k|U|\ell\rangle$. Here,

$$1 = \langle \ell|\ell\rangle = \langle \ell|U^\dagger U|\ell\rangle = \sum_m \langle \ell|U^\dagger|m\rangle\langle m|U|\ell\rangle = \sum_m \langle m|U|\ell\rangle|^2 > |\langle \ell|U|\ell\rangle|^2$$

and thus $\ell < k$ by the maximality of k, and we then consider the two-level unitary V, called **Givens rotation**, that acts only on $|k\rangle$ and $|\ell\rangle$, and does so as

$$V|k\rangle = \bar{u}|k\rangle + \bar{v}|\ell\rangle \qquad \text{and} \qquad V|\ell\rangle = v|k\rangle - u|\ell\rangle$$

where

$$u = \frac{x}{\sqrt{|x|^2 + |y|^2}} \quad \text{and} \quad v = \frac{y}{\sqrt{|x|^2 + |y|^2}} .$$

It is easy to verify that V is unitary. Furthermore, setting $U' := UV$, we observe that

$$\langle k|U'|\ell\rangle = v\langle k|U|k\rangle - u\langle k|U|\ell\rangle = vx - uy = 0' ,$$

while still $\langle k|U'|\ell'\rangle = 0$ for all $\ell' > \ell$ with $\ell' \neq k$, and still $\langle k'|U'|k'\rangle = 1$ for all $k' > k$.

Thus, in either case, by a recursive application of the above to U', we obtain a finite sequence $V_1, \dots, V_N$ of two-level unitaries so that $UV_1 \cdots V_N = I$. By inverting the two-level unitaries, we obtain the claimed result. □

For the remainder, we again fix $\mathcal{H}$ to be $\mathcal{H} = \mathbb{C}^2$. The following shows that all single-qubit gates together with *CNOT* form a universal set of gates.

Theorem 3.7. *The gate set $\mathcal{G} = \{CNOT\} \cup \mathcal{U}(\mathcal{H})$ is perfectly universal.*

Proof. By Proposition 3.6 above, it is sufficient to show that any two-level unitary $U = \mathbb{I} \oplus \tilde{U}$ (w.r.t. the computational basis) can be computed with $\mathcal{G}$.[3] For this, we first consider the special case where the two basis vectors $|k\rangle$ and $|\ell\rangle$ on which U acts nontrivially are such that $k \in \{0,1\}^n$ and $\ell \in \{0,1\}^n$ differ in only one bit, say, $|k\rangle = |k_1\rangle \cdots |k_{n-1}\rangle|0\rangle$ and $|\ell\rangle = |k_1\rangle \cdots |k_{n-1}\rangle|1\rangle$.[4] In this case, we consider the multi-control unitary $C = C^{n-1}[k_1, \dots, k_{n-1}](\tilde{U})$ that is controlled by the first $n-1$ qubits to be in state $|k_1, \dots, k_{n-1}\rangle$ and acts on the last qubit as $\tilde{U}$, and we see that

$$C|k\rangle = |k_1, \dots, k_{n-1}\rangle \otimes \tilde{U}|0\rangle = |k_1, \dots, k_{n-1}\rangle \otimes (u_{11}|0\rangle + u_{21}|1\rangle) = u_{11}|k\rangle + u_{21}|\ell\rangle = U|k\rangle$$

[3] Note, however, that the computational complexity of computing general unitaries by means of two-level unitaries is quadratic in $d = 2^n$, and thus exponential in n. Thus, for *efficient* quantum computation, we need more clever ways to compute the unitaries of interest.

[4] That it is the last bit here makes the writing easier but is not crucial to the argument.

and similarly $C|\ell\rangle = U|\ell\rangle$, while $C|i\rangle = |i\rangle$ for any $i \neq k, \ell$. Thus, $C = U$, and by Corollay 2.8 such a multi-control unitary can be computed with $\mathcal{G}$.

In the more typical case where k and ℓ differ in more than one bit, we argue by induction on the number of bits they differ. Choose $k' \in \{0,1\}^n$ such that k' and k differ in one bit, and k' and ℓ in one bit *less* than k and ℓ. Consider the two-level Hermitian unitary V that maps $V|k\rangle = |k'\rangle$ and $V|k'\rangle = |k\rangle$. Then, $VUV|i\rangle = VU|i\rangle = V|i\rangle = |i\rangle$ for any $i \notin \{k, k', \ell\}$, but also $VUV|k\rangle = VU|k'\rangle = V|k'\rangle = |k\rangle$. Thus, the unitary $U' := VUV$ acts non-trivial only on $|k'\rangle$ and $|\ell\rangle$, and so we can apply the induction hypothesis to V and U'. Given that $U = VU'V$, this then proves the claim. □

In combination with Theorem 1.4, we immediately get the following "smaller" gate set.

Corollary 3.8. *The gate set $\mathcal{G} = \{CNOT\} \cup \bigcup_{0 \leq \theta < 4\pi} \{R_Y(\theta), R_Z(\theta)\}$ is perfectly universal.*

We conclude the discussion of universal gate sets with the following two fundamental results, which we state here without (full) proofs. The first result shows that we can replace the uncountable set of 1-qubit gates in the universal gate set $\mathcal{G}$ considered above by two particular 1-qubit gates and still get an *approximate* universal gate set, and the second result shows that these two single 1-qubit gates approximate any 1-qubit gate with low computational complexity.

Theorem 3.9. *For any $n \in \mathbb{N}$, $\varepsilon > 0$ and $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ there exists an n-qubit quantum circuit with gate set $\mathcal{G} = \{H, T, CNOT\}$ that ε-approximately computes U.*

Recall that $T = S_{\pi/4}$, sometimes also referred to as $\pi/8$ gate. Up to an irrelevant global phase, it coincides with $R_Z(\pi/4)$.

Proof (idea). Doing the calculation, one can show that $THTH$ performs a rotation of the Bloch sphere with angle θ defined by $\cos(\theta/2) = \cos^2(\pi/8)$. This θ can be shown to be an irrational multiple of 2π; as a consequence, a rotation with *any* angle can be approximated by a suitable number of repetitions $THTH$. The same holds for $HTHT$, but with respect to a different axis. With these two rotations, it is then possible to do *any* rotation, and thus in particular the rotations R_Z and R_Y, and the claim then follows from Corollary 3.8. □

Theorem 3.10 (Solovay-Kitaev). *Let $\mathcal{G}$ be a gate set that is closed under inversion, and let $n \in \mathbb{N}$ be a constant. Then, for any $U \in \mathcal{U}(\mathcal{H}^{\otimes n})$ that can be ε-approximately computed by a n-qubit quantum circuit with gate set $\mathcal{G}$ for any $\varepsilon > 0$, there exists an n-qubit quantum circuit with gate set $\mathcal{G}$ that ε-approximately computes U and has computational complexity $O\big(\log^4(1/\varepsilon)\big)$.*

In the remainder of these notes, when dealing with (efficient) quantum circuits, we typically leave the gate set $\mathcal{G}$ implicit, taking it as understood that the quantum circuits considered work with "simple" 1- and 2-qubit gates; this is well justified by the Solovay-Kitaev theorem.

3.5 Simon's Algorithm

As we have seen, the Deutsch-Jozsa algorithm performs exponentially better than any *deterministic* classical algorithm *in the worse case*, but only minorly better than a *randomized* classical algorithm with bounded-error. On the other hand, the Bernstein-Vazirani algorithm does clearly outperform any *randomized* classical algorithm, though only by a linear factor. Here, we present a problem where quantum algorithms are exponentially more efficient than randomized classical algorithms. We are still in the *query complexity* setting, where such a separation can be proven. Later, when discussing Shor's algorithm, we will see a super-polynomial separation in *computational complexity* between quantum and classical algorithms, but those come without proofs due to the lack of classical lower-bound proofs.

Here, the computational problem is the following. Given a function $f : \{0,1\}^n \to \{0,1\}^n$ with the promise that there exists a non-zero "period" $s \in \{0,1\}^n$ such that

$$f(x) = f(x') \iff x' \in \{x, x \oplus s\} \tag{3.1}$$

for all $x, x' \in \{0,1\}^n$, find s. For a classical algorithm with black-box access to f, in order to find s, it must query f on two inputs x and x' with $x' = x \oplus s$ (or else must have excluded all other choices for s by means of such a pair of queries). Thus, an algorithm that has made q queries, and so can check $q(q-1)$ differences, has a probability $O(q^2/2^n)$ of having found s. For this to be, say, a constant, q needs to be $\Omega(2^{n/2})$, i.e., exponential in n. We emphasize that probabilities here are over the random choice over all functions with the given constraint. **Simon's algorithm** shows that one can do exponentially better with a quantum algorithm.

Proposition 3.11. *For any integers $n, k \in \mathbb{N}$, there exists a quantum algorithm with black-box access to $U_f \in \mathcal{U}(\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{\otimes n})$ and query complexity $n+k-1$ that outputs s with property (3.1) with probability at least $1 - 2^{-k}$, assumed it exists.*

We remark that while the classical lower bound is meaningful only for a *randomly chosen* function f with the required property, the quantum upper bound holds for *any* such function.

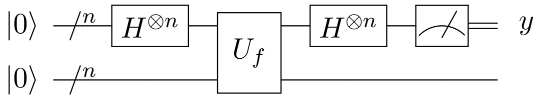

Figure 3.3: Quantum circuit for Simon's algorithm.

Proof. The algorithm is given in Figure 3.3. It starts off with $2n$ qubits in state $|0\rangle$ and applies H to the first n to obtain

$$\frac{1}{2^{n/2}} \sum_x |x\rangle|0\rangle \,,$$

where we understand that the sum is over $x \in \{0,1\}^n$. Applying U_f results in

$$\frac{1}{2^{n/2}} \sum_x |x\rangle|f(x)\rangle \,.$$

For the purpose of this analysis, let us assume that we now measure the second half of the state; whether this measurement takes place now or later (or not at all) makes no difference in the distribution of k; this follows from the fact that actions on different registers commute.

As a result of this measurement, we observe some value $z \in \{0,1\}^n$, and the state of the first n qubits collapses to

$$\frac{1}{\sqrt{2}}(|x\rangle + |x+s\rangle)$$

where x is such that $f(x) = z$. Following previous patterns, the algorithm applies H again to the first n qubits. This results in

$$\frac{1}{2^{n/2}} \sum_y \frac{1}{\sqrt{2}}\big((-1)^{x \cdot y}|y\rangle + (-1)^{(x+s)\cdot y}|y\rangle\big) = \frac{1}{2^{n/2}} \sum_y \frac{1}{\sqrt{2}}(-1)^{x\cdot y}\big(1 + (-1)^{s\cdot y}\big)|y\rangle \,.$$

Now we observe that the amplitude of $|y\rangle$ is 0 if $s \cdot y = 1$, and it is $\pm 1/\sqrt{2^{n-1}}$ otherwise. Thus, by measuring, we will observe a uniformly random $y \in \{0,1\}^n$ with the property that $s \cdot y = 0$. By repeating this procedure $n+k-1$ times, noting that $n+k-1$ random and independent vectors chosen from a vector space over $\mathbb{F}_2$ of dimension $n-1$ have full rank except with probability at most 2^{-k} (see below), s can then be found by means of basic linear algebra techniques. □

In the proof, we made use of the following technical lemma.

Lemma 3.12. *Let $\mathbb{F}$ be a finite field, and let $n, k \in \mathbb{N}$. Then, the probability that a uniformly random matrix $A \in \mathbb{F}^{n\times(n+k)}$ does not have full rank n is at most $|\mathbb{F}|^{-k}$.*

Proof. Let p_i be the probability that the i-th row of A lies in the linear span of the first $i-1$ rows. Note that this linear span has cardinality at most $|\mathbb{F}|^{i-1}$. Thus,

$$p_i \leq \frac{|\mathbb{F}|^{i-1}}{|\mathbb{F}|^{n+k}} = \frac{1}{|\mathbb{F}|^{n+k-i+1}} .$$

Now, for A to not have full rank, there must exist a row that is in the linear span of the previous rows. Thus, by union bound, the probability of A not having full rank is at most

$$\sum_{i=1}^{n} p_i \leq \sum_{i=1}^{n} \frac{1}{|\mathbb{F}|^{n+k-i+1}} \leq \frac{1}{|\mathbb{F}|^{k}} \sum_{i=1}^{n} \frac{1}{|\mathbb{F}|^{n-i+1}} \leq \Big(\frac{1}{2^n} + \cdots + \frac{1}{2}\Big)\frac{1}{|\mathbb{F}|^{k}} \leq \frac{1}{|\mathbb{F}|^{k}} ,$$

as claimed. □

3.6 Grover's Algorithm for Unstructured Search

Grover's algorithm is again less impressive in terms of speed-up, but it applies to a very natural computational problem: given a function $f : \{0,1\}^n \to \{0,1\}$ for which there exist exactly M choices of $x \in \{0,1\}^n$ for which $f(x) = 1$, find such an x. Any classical algorithm with black-box access to f that succeeds with constant probability has query complexity $\Omega(2^n/M)$. With a quantum algorithm, we can gain a quadratic speed up.

Proposition 3.13. *For any positive integers n and $M \leq 2^n$, there exists a quantum algorithm with black-box access to $U_f \in \mathcal{U}(\mathcal{H}^{\otimes n} \otimes \mathcal{H})$ and query complexity $O\big(\sqrt{2^n/M}\big)$ that outputs x with $f(x) = 1$ with probability at least $1 - M/2^n$, given that there exist exactly M such x's.*

Note that the algorithm needs to know M, which is not very realistic in typical applications. As will become from the proof, the algorithm still works though, say with success probability at least $1/2$, if a sufficiently good approximation is known. As a matter of fact, by essentially running the algorithm with cleverly chosen guesses for M, an expected number of $O\big(\sqrt{2^n/M}\big)$ queries still suffice to find a solution in case M is unknown.

Proof. By the phase kickback technique from Sections 3.1 and 3.2, we may just as well assume black box access to $V_f \in \mathcal{U}(\mathcal{H}^{\otimes n})$ with $V_f|x\rangle = (-1)^{f(x)}|x\rangle$. The algorithm starts off with $|0\rangle \in \mathcal{S}(\mathcal{H}^{\otimes n})$ and applies $H^{\otimes n}$. This is followed by ℓ *Grover iterations* and, finally, the resulting state is measured. As illustrated in Figure 3.5 below, the Grover iteration consists of: applying V_f, applying $H^{\otimes n}$, applying a conditional phase shift $P = 2|0\rangle\langle 0| - \mathbb{I} \in \mathcal{U}(\mathcal{H}^{\otimes n})$, which is such that $P|0\rangle = |0\rangle$ and $P|x\rangle = -|x\rangle$ for $x \neq 0$, and applying $H^{\otimes n}$ once more.

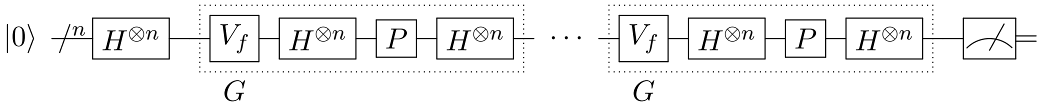

Figure 3.4: Quantum circuit for Grover's algorithm.

We now analyze the algorithm and determine ℓ, which then obviously determines the query complexity. Consider

$$|\psi\rangle = H^{\otimes n}|0\rangle = \frac{1}{2^{n/2}}\sum_x |x\rangle\,,$$

and note that $H^{\otimes n}PH^{\otimes n} = H^{\otimes n}(2|0\rangle\langle 0| - \mathbb{I})H^{\otimes n} = 2|\psi\rangle\langle\psi| - \mathbb{I}$, so that the Grover iteration may be written as

$$G = H^{\otimes n}PH^{\otimes n}V_f = (2|\psi\rangle\langle\psi| - \mathbb{I})V_f\,.$$

On order to understand the action of G, we set

$$|\gamma\rangle := \frac{1}{\sqrt{M}}\sum_{\substack{x \text{ s.t.}\\ f(x)=1}} |x\rangle \qquad \text{and} \qquad |\beta\rangle := \frac{1}{\sqrt{N-M}}\sum_{\substack{x \text{ s.t.}\\ f(x)=0}} |x\rangle\,,$$

where $N = 2^n$, and we observe that we can write

$$|\psi\rangle = \sqrt{\frac{N-M}{N}}\,|\beta\rangle + \sqrt{\frac{M}{N}}\,|\gamma\rangle = \cos(\theta_\circ)\,|\beta\rangle + \sin(\theta_\circ)\,|\gamma\rangle =: |\psi_{\theta_\circ}\rangle$$

for the proper choice of $\theta_\circ \in [0, \frac{\pi}{2}]$. We note that measuring $|\psi\rangle = |\psi_{\theta_\circ}\rangle$ (in the computational basis) produces a uniformly random $x \in \{0,1\}^n$, and thus $f(x) = 1$ will be satisfied with probability M/N only. What we will see is that the Grover iterations will bring $|\psi_{\theta_\circ}\rangle$ closer and closer to the "good" state $|\gamma\rangle$, which is such that measuring $|\gamma\rangle$ produces an x that satisfies $f(x) = 1$ with certainty.

For this purpose, let us first consider the action of V_f. It follows immediately by definition that

$$V_f|\beta\rangle = |\beta\rangle \qquad \text{and} \qquad V_f|\gamma\rangle = -|\gamma\rangle$$

Thus, within the space spanned by $|\gamma\rangle$ and $|\beta\rangle$ (over $\mathbb{R}$), the unitary V_f acts as a reflection across the axis spanned by $|\beta\rangle$. Similarly, for $P' := H^{\otimes n}PH^{\otimes n} = (2|\psi\rangle\langle\psi| - \mathbb{I})$, we have that

$$P'|\psi\rangle = |\psi\rangle \qquad \text{and} \qquad P'|\psi^\perp\rangle = -|\psi^\perp\rangle$$

for any $|\psi^\perp\rangle$ in the span of $|\gamma\rangle$ and $|\beta\rangle$ with $\langle\psi|\psi^\perp\rangle = 0$. Thus, within the space spanned by $|\gamma\rangle$ and $|\beta\rangle$ (over $\mathbb{R}$), the unitary P' acts as a reflection across the axis spanned by $|\psi\rangle$. The composition of the two is then a rotation by angle $2\theta_\circ$ towards $|\gamma\rangle$, see Fig. 3.5. Thus, the Grover iteration G maps $|\psi_{\theta_\circ}\rangle$ to $|\psi_{3\theta_\circ}\rangle$, and $|\psi_{3\theta_\circ}\rangle$ to $|\psi_{5\theta_\circ}\rangle$, etc.

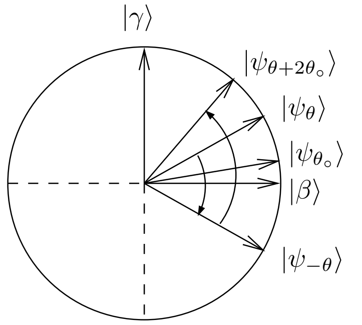

Figure 3.5: A Grover iteration, mapping $|\psi_\theta\rangle \overset{V_f}{\mapsto} |\psi_{-\theta}\rangle \overset{P'}{\mapsto} |\psi_{\theta+2\theta_\circ}\rangle$.

Alternatively, arguing purely trigonometrically, for any θ we have that

$$V_f|\psi_\theta\rangle = \cos(\theta)\,|\beta\rangle - \sin(\theta)\,|\gamma\rangle = |\psi_{-\theta}\rangle\,,$$

while working out $|\psi\rangle\langle\psi|$ and using basic trigonometric identities,

$$\begin{aligned}\big(2|\psi\rangle\langle\psi| - \mathbb{I}\big)|\beta\rangle &= (2\cos(\theta_\circ)^2 - 1)\,|\beta\rangle + 2\sin(\theta_\circ)\cos(\theta_\circ)\,|\gamma\rangle \\ &= \cos(2\theta_\circ)\,|\beta\rangle + \sin(2\theta_\circ)\,|\gamma\rangle\end{aligned}$$

and

$$\begin{aligned}\big(2|\psi\rangle\langle\psi| - \mathbb{I}\big)|\gamma\rangle &= 2\sin(\theta_\circ)\cos(\theta_\circ)\,|\beta\rangle + (2\sin(\theta_\circ)^2 - 1)\,|\gamma\rangle \\ &= \sin(2\theta_\circ)\,|\beta\rangle - \cos(2\theta_\circ)\,|\gamma\rangle\,,\end{aligned}$$

so that, putting all together, we obtain

$$\begin{aligned}G|\psi_\theta\rangle &= \big(2|\psi\rangle\langle\psi| - \mathbb{I}\big)V_f|\psi_\theta\rangle \\ &= \big(2|\psi\rangle\langle\psi| - \mathbb{I}\big)\big(\cos(\theta)\,|\beta\rangle - \sin(\theta)\,|\gamma\rangle\big) \\ &= \big(\cos(\theta)\cos(2\theta_\circ) - \sin(\theta)\sin(2\theta_\circ)\big)\,|\beta\rangle + \big(\cos(\theta)\sin(2\theta_\circ) + \sin(\theta)\cos(2\theta_\circ)\big)\,|\gamma\rangle \\ &= \cos(\theta + 2\theta_\circ)\,|\beta\rangle + \sin(\theta + 2\theta_\circ)\,|\gamma\rangle\,.\end{aligned}$$

Therefore, the state $|\psi\rangle = |\psi_{\theta_\circ}\rangle$ after the application of the inital $H^{\otimes n}$ evolves as $|\psi_{\theta_\circ}\rangle$, $|\psi_{3\theta_\circ}\rangle$, $|\psi_{5\theta_\circ}\rangle$ etc. Ideally, we want to choose ℓ such that $(2\ell+1)\theta_\circ = \frac{\pi}{2}$, so that the final state $|\psi_{(2\ell+1)\theta_\circ}\rangle$ that is measured equals $|\gamma\rangle$, and so we observe x with $f(x) = 1$ with certainty. Furthermore, using that $\sqrt{M/N} = \sin(\theta_\circ) \leq \theta_\circ$, we then have that ℓ, and thus the query complexity of the algorithm, is $O(1/\theta_\circ) = O(\sqrt{N/M})$. In general, this choice of ℓ will not be an integer, and then we have to round to the closest integer; the probability of observing x with $f(x) = 1$ is then still at least

$$\sin\Big(\frac{\pi}{2} \pm \theta_\circ\Big)^2 = 1 - \cos\Big(\frac{\pi}{2} \pm \theta_\circ\Big)^2 = 1 - \sin(\theta_\circ)^2 = 1 - M/N\,,$$

as claimed. □

Our focus here is on query complexity, but we do want to point out that by means of the techniques from Section 2.6, the P gate can be implemented using $O(n)$ elementary gates (plus $O(n)$ work qubits), and thus the computational complexity is larger by a factor $O(n)$ only.