

Assume for a moment that hidden variables exist. Then the measurement is fundamentally deterministic, but appears probabilistic because some degrees of freedom of the system (described by the hidden variables) are not known. To illustrate the idea, let us consider again the spin- $\frac{1}{2}$ particle. Suppose that the particle is prepared initially in state $|\uparrow_z\rangle$ and that a hidden variable λ determines the outcome of spin measurement on that particle along any axis \hat{n}_θ characterized by the rotation angle θ about the y axis. Since during the preparation stage we have no control over the hidden variable, it is reasonable to assume that λ is a random number uniformly distributed over the unit interval $0 \leq \lambda \leq 1$ with the probability distribution $\rho(\lambda) = \text{const}$ which is normalized as $\int_0^1 \rho(\lambda) d\lambda = 1$. Then the spin measurement along the axis \hat{n}_θ rotated by the angle θ yields the spin-up state $|\uparrow_\theta\rangle$ if $0 \leq \lambda < \cos^2 \frac{\theta}{2}$ and the spin-down state $|\downarrow_\theta\rangle$ if $\cos^2 \frac{\theta}{2} \leq \lambda \leq 1$. For example, in the case of $\theta = \pi/2$, we have $\hat{n}_{\pi/2} = \hat{x}$. Then the $|\uparrow_x\rangle$ state is obtained if $0 \leq \lambda < \frac{1}{2}$ while the $|\downarrow_x\rangle$ state is obtained if $\frac{1}{2} \leq \lambda \leq 1$.

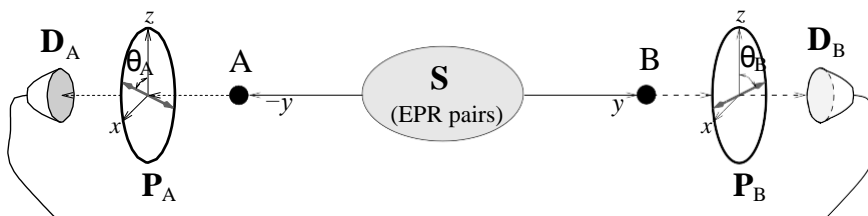


Fig. 8.11. Schematic representation of Bell's setup: S is the source of correlated (entangled) particles A and B, measured by corresponding apparatuses each consisting of analyzer P_i and detector D_i .

Considering the case of two spatially separated but correlated particles, Bell, however, has shown that certain statistical predictions of quantum mechanics are incompatible with hidden variable theories based on local realism. Let us outline his arguments following the treatment of Clauser and Horne. Suppose that EPR correlated (entangled) particles A and B are produced by a source of such particles, one pair at a time. The particles A and B fly in the opposite directions towards their respective analyzer-detector apparatuses. Each apparatus has an adjustable parameter characterized by variable θ , which may denote, e.g., the angle of the analyzer defining the measurement axis for spin- $\frac{1}{2}$ particles or for photon polarization, as shown schematically in Fig. 8.11. For fixed values of the parameters θ_A and θ_B , the probability $p_{AB}(\lambda, \theta_A, \theta_B)$ of simultaneous detection of both particles is a function of a hidden variable, or a set of such, collectively denoted by λ . Since the two apparatuses are assumed to be separated by sufficiently large distance, according to Einstein's locality constraint, the space-like separated detection events can not influence one another. Then the joint probability p_{AB} should factorize into

the product of individual detection probabilities p_A and p_B , according to

$$p_{AB}(\lambda, \theta_A, \theta_B) = p_A(\lambda, \theta_A) p_B(\lambda, \theta_B) . \quad (8.43)$$

At the system preparation and detection stages, we can neither control nor determine the hidden variables λ . Let us therefore characterize λ by some probability distribution $\rho(\lambda)$ normalized in the usual way $\int_{\Gamma} \rho(\lambda) d\lambda = 1$, where Γ encompasses the complete range of possible values of λ ($\lambda \in \Gamma$). Performing many identical measurements with fixed parameters of the system, the observed ensemble averaged probabilities are given by

$$P_A(\theta_A) = \int_{\Gamma} p_A(\lambda, \theta_A) \rho(\lambda) d\lambda , \quad (8.44a)$$

$$P_B(\theta_B) = \int_{\Gamma} p_B(\lambda, \theta_B) \rho(\lambda) d\lambda , \quad (8.44b)$$

$$P_{AB}(\theta_A, \theta_B) = \int_{\Gamma} p_{AB}(\lambda, \theta_A, \theta_B) \rho(\lambda) d\lambda . \quad (8.44c)$$

Below we use a theorem from number theory, stating that for any four numbers x_1, x_2, y_1, y_2 , such that $0 \leq x_{1,2}, y_{1,2} \leq 1$, the inequality

$$-1 \leq x_1 y_1 - x_1 y_2 + x_2 y_1 + x_2 y_2 - x_2 - y_1 \leq 0 \quad (8.45)$$

is always satisfied (Prob. 8.4). Let us denote by θ_A, θ_A^r and θ_B, θ_B^r four possible values of the parameters of apparatuses measuring particles A and B, respectively. Since for any θ_A, θ_B and λ , physically meaningful probabilities must satisfy $0 \leq p_A(\lambda, \theta_A), p_B(\lambda, \theta_B) \leq 1$, we can use (8.45) to write

$$\begin{aligned} -1 \leq & p_A(\lambda, \theta_A) p_B(\lambda, \theta_B) - p_A(\lambda, \theta_A) p_B(\lambda, \theta_B^r) \\ & + p_A(\lambda, \theta_A^r) p_B(\lambda, \theta_B) + p_A(\lambda, \theta_A^r) p_B(\lambda, \theta_B^r) \leq p_A(\lambda, \theta_A^r) + p_B(\lambda, \theta_B) . \end{aligned} \quad (8.46)$$

Multiplying all terms of this inequality by $\rho(\lambda)$ and integrating over λ taking into account (8.43), we obtain

$$\begin{aligned} -1 \leq & P_{AB}(\theta_A, \theta_B) - P_{AB}(\theta_A, \theta_B^r) \\ & + P_{AB}(\theta_A^r, \theta_B) + P_{AB}(\theta_A^r, \theta_B^r) \leq P_A(\theta_A^r) + P_B(\theta_B) . \end{aligned} \quad (8.47)$$

If, due to, e.g., rotational invariance, the probabilities $P_A(\theta_A)$ and $P_B(\theta_B)$ are constant and the joint probability $P_{AB}(\theta_A, \theta_B) = P_{AB}(\Delta\theta)$ is a function of only the angle difference $\Delta\theta = |\theta_A - \theta_B|$, by choosing the four angles so that

$$|\theta_A - \theta_B| = |\theta_A^r - \theta_B| = |\theta_A^r - \theta_B^r| = \frac{1}{3} |\theta_A - \theta_B^r| = \varphi ,$$

from (8.47) we obtain the so-called Bell's inequality

$$S(\varphi) = \frac{3P_{AB}(\varphi) - P_{AB}(3\varphi)}{P_A + P_B} \leq 1. \quad (8.48)$$

Below we show that in the cases of EPR correlated spin- $\frac{1}{2}$ particles and polarization-entangled single photons, such rotational invariance is indeed satisfied and that for a certain range of angles φ , the values of function $S(\varphi)$ exceed 1, violating Bell's inequality (8.48). This proves that any hidden variable theory based on Einstein's conviction of local realism is incompatible with certain predictions of quantum mechanics, which can be tested experimentally.

8.8.2 Violations of Bell's Inequality

We outline now two physical schemes for testing Bell's inequality (8.48) against experimentally confirmed predictions of quantum mechanics.

Entangled spin- $\frac{1}{2}$ particles

Consider first two spin- $\frac{1}{2}$ particles A and B in the entangled state

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle_A |\uparrow_z\rangle_B + |\downarrow_z\rangle_A |\downarrow_z\rangle_B). \quad (8.49)$$

In a real experiment, these particles could be, e.g., Hg atoms. Then the entangled state (8.49) could be realized in two steps: First, a Hg_2 molecule is photodissociated into the singlet state $\frac{1}{\sqrt{2}}(|\uparrow_z\rangle_A |\uparrow_z\rangle_B + |\downarrow_z\rangle_A |\downarrow_z\rangle_B)$, whose two constituent particles (atoms) A and B fly in opposite directions, parallel and anti-parallel to the y axis, respectively. Next, a longitudinal magnetic field is applied to one of the particles to rotate its spin around the y axis by angle π , realizing thereby the σ_y transformation that results in state (8.49), to within the trivial overall phase factor $e^{i\pi/2}$ which can be omitted.

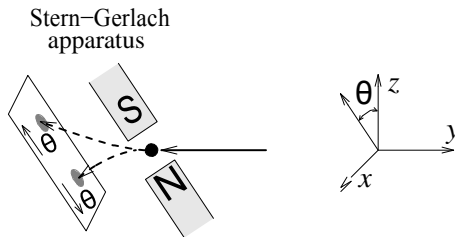


Fig. 8.12. Spin- $\frac{1}{2}$ particle passing through a Stern-Gerlach apparatus rotated by angle θ with respect to the z axis.

Let us first determine the probability of detecting a spin- $\frac{1}{2}$ particle in the spin-up state $|\uparrow_\theta\rangle$ along the axis rotated by angle θ with respect to the z

axis. Such a measurement can be realized by letting the particle pass a Stern-Gerlach apparatus and detecting its upward deflected component, as sketched in Fig. 8.12. This amounts to projecting the state of the particle onto the state $|\uparrow_\theta\rangle$, which can be obtained by applying the rotation operator $R_y(\theta)$ to state $|\uparrow_z\rangle$ (see Sect. 8.4),

$$|\uparrow_\theta\rangle = R_y(\theta) |\uparrow_z\rangle = e^{-i\theta\sigma_y/2} |\uparrow_z\rangle = \cos \frac{\theta}{2} |\uparrow_z\rangle + \sin \frac{\theta}{2} |\downarrow_z\rangle.$$

The probability $P(\theta)$ of detecting the particle in state $|\uparrow_\theta\rangle$ is then given by the expectation value of the projection operator $\Pi_\theta = |\uparrow_\theta\rangle\langle\uparrow_\theta|$, $P(\theta) = \langle\Pi_\theta\rangle$. If we now have a system of two spin- $\frac{1}{2}$ particles A and B, each analyzed by its own Stern-Gerlach apparatus rotated by the corresponding angle $\theta_{A,B}$, the joint detection probability $P_{AB}(\theta_A, \theta_B)$ is given by the expectation value of the product of two projection operators $\Pi_{\theta_A}^A = |\uparrow_{\theta_A}\rangle\langle\uparrow_{\theta_A}|$ and $\Pi_{\theta_B}^B = |\uparrow_{\theta_B}\rangle\langle\uparrow_{\theta_B}|$, $P_{AB}(\theta_A, \theta_B) = \langle\Pi_{\theta_A}^A \Pi_{\theta_B}^B\rangle$.

We can now easily calculate all detection probabilities for a pair of particles in the entangled state (8.49). For P_A and P_B we have

$$P_A(\theta_A) = P_B(\theta_B) = \frac{1}{2} \quad (8.50)$$

for any θ_A and θ_B , while for the joint detection probability P_{AB} we obtain

$$P_{AB}(\theta_A, \theta_B) = \frac{1}{2} \cos^2 \frac{\theta_A - \theta_B}{2}. \quad (8.51)$$

Equations (8.50) and (8.51) show that the rotational invariance assumed in the derivation of function $S(\varphi)$ is indeed satisfied in this case. Choosing the four detection angles as $\theta_A = 0$, $\theta_B = \pi/4$, $\theta'_A = \pi/2$ and $\theta'_B = 3\pi/4$ yields $\varphi = \pi/4$. From (8.48) we then obtain

$$S(\varphi) = \frac{3}{2} \cos^2 \frac{\varphi}{2} - \frac{1}{2} \cos^2 \frac{3\varphi}{2} = 1.2 \not\leq 1, \quad (8.52)$$

which clearly violates Bell's inequality. Hence, the predictions of quantum mechanics, which have been verified in many experiments, contradict and thereby invalidate the hidden variable theories based on Einstein's local realism. This leads to the inescapable conclusion that quantum mechanics is a nonlocal theory.

Entangled photons

We now describe an optical scheme for testing Bell's inequality (8.48) using pairs of photons in the entangled state

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|\rightarrow\rangle_A |\rightarrow\rangle_B + |\leftrightarrow\rangle_A |\leftrightarrow\rangle_B). \quad (8.53)$$

In fact, most of the experimental tests of Bell's inequalities have been performed using polarization-entangled photon pairs. These include a series of pioneering experiments by Aspect and coworkers, using atomic radiative cascade, as well as a number of experiments by several teams using nonlinear crystals to realize spontaneous parametric down-conversion. In the former experiments, a high-efficiency source of pairs of photons, at frequencies $\omega_A = 2\pi \times 7.102 \times 10^{14}$ rad/s and $\omega_B = 2\pi \times 5.44 \times 10^{14}$ rad/s, was obtained by two-photon excitation of state $4p^2\ ^1S_0$, via the intermediate state $3d4p\ ^1P_1$, of the $4p^2\ ^1S_0 \xrightarrow{\omega_A} 4s4p\ ^1P_1 \xrightarrow{\omega_B} 4s^2\ ^1S_0$ radiative cascade in calcium. The polarization entanglement of the photons comes about because of angular momentum conservation. Since relative to photons, the atoms are massive objects, their recoil during photon emission is negligible. Therefore, in addition to the polarization entanglement, the propagation directions of the two photons are also strongly correlated, due to energy and momentum conservation. In the experiments with nonlinear crystals possessing the $\chi^{(2)}$ nonlinearity, a pump photon at frequency ω_p is converted into a pair of photons (called signal and idler) with the frequencies $\omega_s \equiv \omega_A$ and $\omega_i \equiv \omega_B$ such that $\omega_s + \omega_i = \omega_p$. Here again, angular momentum conservation imposes polarization entanglement between the photons, while the phase-matching conditions result in a finite angle between the propagation directions of the photons, which makes it possible to redirect each photon to its own measuring apparatus. The measurements in different bases are realized by detecting photons that go through the usual optical polarizers rotated by the corresponding angle θ . This amounts to projecting the state of each photon onto the corresponding state $|\theta\rangle$, obtained by rotating the vertical polarization state $|\uparrow\rangle$ with the rotation operator $R(\theta)$ of (8.29),

$$|\theta\rangle = R(\theta) |\uparrow\rangle = \cos\theta |\uparrow\rangle + \sin\theta |\leftrightarrow\rangle.$$

Then the individual and joint detection probabilities $P_A(\theta_A) = \langle \Pi_{\theta_A}^A \rangle$, $P_B(\theta_B) = \langle \Pi_{\theta_B}^B \rangle$ and $P_{AB}(\theta_A, \theta_B) = \langle \Pi_{\theta_A}^A \Pi_{\theta_B}^B \rangle$ are given by the expectation values of the corresponding projection operators $\Pi_{\theta_A}^A = |\theta_A\rangle\langle\theta_A|$ and $\Pi_{\theta_B}^B = |\theta_B\rangle\langle\theta_B|$.

When the two photons are in the entangled state (8.53), similarly to the case of spin- $\frac{1}{2}$ particles, the quantum mechanical calculation of the detection probabilities yields

$$P_A(\theta_A) = P_B(\theta_B) = \frac{1}{2}, \quad (8.54)$$

$$P_{AB}(\theta_A, \theta_B) = \frac{1}{2} \cos^2(\theta_A - \theta_B). \quad (8.55)$$

Choosing the four detection angles as $\theta_A = 0$, $\theta_B = \pi/8$, $\theta_A^r = \pi/4$ and $\theta_B^r = 3\pi/8$, we obtain $\varphi = \pi/8$ and, correspondingly,

$$S(\varphi) = \frac{3}{2} \cos^2(\varphi) - \frac{1}{2} \cos^2(\varphi) \underset{C}{=} 1.2 \notin 1, \quad (8.56)$$

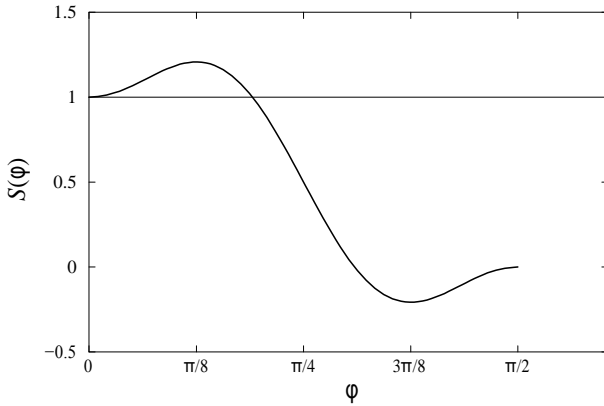


Fig. 8.13. Variation of the function $S(\varphi)$ with angle φ , as given by (8.56). The dashed line represents the upper bound of Bell's inequality (8.48).

which again violates Bell's inequality, refuting any hidden variable theory based on local realism. As shown in Fig. 8.13, where we plot the function $S(\varphi)$, inequality (8.48) is violated for the values of angle φ in the range $0 < \varphi < 3\pi/16$. The strongest violation, however, is attained in the vicinity of $\varphi \approx \pi/8$.

8.8.3 Greenberger–Horne–Zeilinger Equality

We have seen above that certain statistical predictions of quantum mechanics, which require averaging over a large number of measurements by repeating the experiment many times, violate Bell's inequality (8.48). In 1989 Greenberger, Horne and Zeilinger—GHZ—discovered a more powerful test of the existence of elements of reality which may be hidden from us due to our inability to control and detect them for whatever reason. In the experiment proposed by GHZ, such elements of reality, if existing, would reveal themselves in just a single measurement, and in complete violation of the predictions by quantum mechanics.

Following GHZ, instead of the EPR entangled state of two spin- $\frac{1}{2}$ particles, we consider the three particle entangled state

$$|\psi_3\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle^A |\uparrow_z\rangle^B |\uparrow_z\rangle^C - |\downarrow_z\rangle^A |\downarrow_z\rangle^B |\downarrow_z\rangle^C). \quad (8.57)$$

As before, let us assume that all three constituent particles A, B and C of this state are separated from each other by sufficiently large distances, so that the measurement performed at each particle site can not influence the other two. Consider three composite operators, $S_1 = \sigma^A \sigma^B \sigma^C$, $S_2 = \sigma^A \sigma^B \sigma^C$, and $S_3 = \sigma^A \sigma^B \sigma^C$, where the superscript of each Pauli operator indicates the

particle upon which that operator acts. Using (8.19), it is easy to see that all three operators S_i ($i = 1, 2, 3$) mutually commute,

$$[S_1, S_2] = [S_2, S_3] = [S_3, S_1] = 0.$$

In addition, the GHZ state (8.57) is a simultaneous eigenstate of these operators, with the eigenvalue +1,

$$S_1 |\psi_3\rangle = S_2 |\psi_3\rangle = S_3 |\psi_3\rangle = +1 |\psi_3\rangle. \quad (8.58)$$

This means that the product of the result of three spin measurements as given by the operators S_i (i.e., any two spins along the y axis and the third spin along the x axis), has to be +1. Consider, for example, operator S_1 : If the measurements on particles B and C result in the +1 eigenvalue of σ^B_y and -1 eigenvalue of σ^C_x , then measuring σ^A_y on particle A has to yield the eigenvalue -1, i.e., the state is $|\downarrow_x^A\rangle |\downarrow_y^B\rangle |\downarrow_y^C\rangle$. Other possible outcomes of the S_1 measurement are $|\uparrow_z^A\rangle |\uparrow_y^B\rangle |\uparrow_y^C\rangle$, $|\uparrow_z^A\rangle |\uparrow_y^B\rangle |\downarrow_y^C\rangle$, and $|\downarrow_x^A\rangle |\downarrow_y^B\rangle |\uparrow_y^C\rangle$. Equivalent reasoning applies to the operators S_2 and S_3 .

Consider next the operator $S_4 = \sigma^A_x \sigma^B_x \sigma^C_x$. Using the equalities $\sigma^2 = I$ and $\sigma_y \sigma_x = -\sigma_x \sigma_y$, it is easy to show that $S_4 = -S_1 S_2 S_3$. Since all three operators S_1 , S_2 and S_3 have the eigenvalue +1, the eigenvalue of S_4 is -1,

$$S_4 |\psi_3\rangle = -1 |\psi_3\rangle. \quad (8.59)$$

Thus, all of the possible outcomes of the S_4 measurement on state (8.57) are $|\downarrow_x^A\rangle |\downarrow_x^B\rangle |\downarrow_x^C\rangle$, $|\downarrow_x^A\rangle |\uparrow_x^B\rangle |\uparrow_x^C\rangle$, $|\uparrow_x^A\rangle |\downarrow_x^B\rangle |\uparrow_x^C\rangle$ and $|\uparrow_x^A\rangle |\uparrow_x^B\rangle |\downarrow_x^C\rangle$. Let us now assign to the operators σ_j^i and σ_j^i ($j = A, B, C$) the corresponding

“elements of reality” m_x^i and m_y^i , each having the value +1 or -1 which is revealed by the relevant measurement. From (8.58) we have

$$m_x^A m_y^B m_y^C = 1, \quad m_y^A m_x^B m_y^C = 1, \quad m_y^A m_y^B m_x^C = 1. \quad (8.60)$$

Multiplying the three equalities and taking into account that $(m^i)^2 = 1$, we obtain

$$\begin{aligned} & (m^A m^B m^C)(m^A m^B m^C)(m^A m^B m^C) \\ &= m_x^A m_y^B m_y^C (m^A)^2 (m^B)^2 (m^C)^2 \\ &= m_x^A m_x^B m_x^C = 1. \end{aligned} \quad (8.61)$$

On the other hand, the quantum mechanical prediction from (8.59) is

$$m_x^A m_x^B m_x^C = -1, \quad (8.62)$$

which contradicts (8.61). Several recent experiments by Zeilinger and coworkers, using three- and four-photon GHZ states have clearly confirmed the quantum mechanical result, refuting the hypothesis of the existence of elements of reality.

8.9 Entropy and Information Theory

A pivotal theme of information theory—both classical and quantum—is the quantification of the information, produced by a source, through the smallest amount of memory needed to faithfully represent it, or the minimum amount of communication needed to reliably convey it. In classical information theory, this reduces to the compressibility of information characterized by a given probability distribution of its source, the measure of which is Shannon's entropy. In quantum information theory, it is the von Neumann entropy that plays the same role. A novel feature of quantum information theory, not having a classical counterpart, is quantum entanglement, which, as we have already seen, is a vital information resource; hence the necessity of verifying and quantifying the entanglement, which is a very active topic of current research. Our aim in this section is to outline certain aspects of information theory, whose detailed discussion is beyond the scope of this book.

The Shannon Entropy

Consider a message composed of a long string of letters $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ chosen from a binary alphabet, e.g., $x \in \{0, 1\}$. Assume that the letters in the message are statistically independent, with 0 appearing with probability $p_0 \equiv p$ and 1 with probability $p_1 = 1 - p$. When n is very large, a typical message would contain about np zeros and $n(1-p)$ ones. The number of such messages is given by the binomial coefficient $\binom{n}{np}$ which can be approximated as

$$\binom{n}{np} \approx 2^{nH_{\text{bin}}(p)},$$

where the function

$$H_{\text{bin}}(p) = -p \log_2 p - (1-p) \log_2 (1-p) \quad (8.63)$$

is referred to as the Shannon entropy of a binary source. Clearly $0 \leq H_{\text{bin}}(p) \leq 1$, with $H_{\text{bin}}(p) = 0$ when $p = 0$ or 1 , and $H_{\text{bin}}(p) = 1$ when $p = \frac{1}{2}$. Let us assign to each typical message a positive integer number, which can be represented by a binary string of length $nH_{\text{bin}}(p)$. Then, to convey a typical message between two parties, instead of sending the message itself, it is enough to send the binary string identifying that message. That string is shorter than the original message, since $H_{\text{bin}}(p) < 1$ for any $p \neq \frac{1}{2}$. This procedure thus yields data compression. The fact that for $p = \frac{1}{2}$, corresponding to maximum Shannon entropy, the data can not be compressed points to a physical interpretation of these concepts: $p = \frac{1}{2}$ means completely random distribution of zeros and ones in a typical message string in which the sequence of letters contains no pattern. As such, the message has to be transferred as a whole and no compression can achieve the same result. It then makes perfect sense that maximum entropy is associated with maximum information. And it is

no accidental coincidence that maximum entropy in statistical physics is also associated with complete randomness.

The above result can be generalized to the case of an alphabet containing $k \geq 2$ letters, i.e., $x \in \{x_1, x_2, \dots, x_k\}$. Assuming that each letter x_j appears with the corresponding probability p_j ($\sum_j p_j = 1$), a typical message of length n would have np_1 instances of x_1 , np_2 instances of x_2 , etc. The total number of permutations in such message strings is given by

$$\frac{n!}{\prod_j (np_j)!} \approx 2^{nH(X)},$$

where

$$H(X) \equiv H(p_1, p_2, \dots, p_k) = - \sum_j p_j \log_2 p_j \quad (8.64)$$

is the Shannon entropy of the ensemble $X = \{x_j, p_j\}$. Again, we can assign to each typical message a positive integer number, and send that number to the receiver using only $nH(p)$ bits. Importantly, as the length n of the message grows, the probability of having to deal with an atypical message, in which the statistical weights of various letters x_j deviate from the typical ones np_j , quickly approaches zero. Therefore, the above procedure achieves (asymptotically as $n \rightarrow \infty$) optimal data compression with the rate $H(X)$, which is Shannon's noiseless coding theorem.

The Von Neumann Entropy

Generalizing now the notion of entropy to quantum ensembles, for a quantum system characterized by the density operator ρ , the so-called von Neumann entropy $S(\rho)$ of ρ is defined as

$$S(\rho) = -\text{Tr}(\rho \log_2 \rho). \quad (8.65)$$

The von Neumann entropy is invariant under unitary basis transformations, $S(U \rho U^\dagger) = S(\rho)$, which follows from (1.48). Choosing an orthogonal basis $\{|\psi_i\rangle\}$ in which ρ is diagonal,

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|,$$

where λ_i are the eigenvalues of ρ , we can write

$$S(\rho) = - \sum_i \lambda_i \log_2 \lambda_i = H(Y), \quad (8.66)$$

which shows that the von Neumann entropy $S(\rho)$ reduces to the Shannon entropy $H(Y)$ for the ensemble $Y = \{|\psi_i\rangle, \lambda_i\}$. Clearly, if the system is in a pure state $\rho = |\Psi\rangle\langle\Psi|$, its entropy vanishes, $S(\rho) = 0$. Conversely, the entropy

attains the maximum $S(\rho) = \log_2 N$ in the case of a completely mixed state $\rho = \frac{1}{N} \sum_i |\psi_i\rangle\langle\psi_i| = \frac{1}{N}I$, where N is the dimension of the corresponding Hilbert space.

We can now outline the quantum analog of the classical data compression, which is known as Schumacher's quantum noiseless coding theorem. Consider a quantum source which produces messages composed of sequences of n qubits represented by, e.g., polarization states of photons. Assume that the possible states of the qubits are drawn from a set of distinct pure states $\{|\psi_i\rangle\}$, not necessarily orthogonal to each other, with each state occurring with the respective probability p_i ($\sum_i p_i = 1$). Thus, each qubit in a message is characterized by the density operator $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ and the whole message is described by the tensor product density operator $\rho^{\otimes n} = \rho \otimes \rho \otimes \dots \otimes \rho$ which spans an $N = 2^n$ dimensional Hilbert space $H^{(N)}$. We can diagonalize $\rho^{\otimes n}$ through an appropriate unitary transformation. The corresponding orthogonal basis states will be represented by the products of eigenstates $|x_0\rangle$ and $|x_1\rangle$ of individual qubits, while the eigenvalues of $\rho^{\otimes n}$ will be products of eigenvalues $p_0 = p$ and $p_1 = 1 - p$ of ρ . In this basis, the information content of $\rho^{\otimes n}$ is essentially that of a classical source producing message strings corresponding to $|x^{(1)}, x^{(2)}, \dots, x^{(n)}\rangle$ with probabilities $p^{(1)}p^{(2)} \dots p^{(n)}$. The von Neumann entropy $S(\rho)$ of ρ is obviously equal to the Shannon entropy $H_{\text{bin}}(p)$ of (8.63). This means that we have about $M = 2^{nH_{\text{bin}}(p)}$ orthogonal message strings, which can be encoded in a quantum system whose Hilbert space $H^{(M)}$ is M dimensional. This requires only $nS(\rho) = nH_{\text{bin}}(p)$ qubits, whose product space can therefore accommodate all the typical quantum messages with high fidelity. Since $H_{\text{bin}}(p) < 1$ for any $p \neq \frac{1}{2}$, quantum data compression is thereby achieved. Only for completely random qubits $\rho = \frac{1}{2}I$ ($p = \frac{1}{2}$), in which case

$S(\rho) = 1$ ($H_{\text{bin}}(p) = 1$), no compression is possible ($M = N$), in complete analogy with the classical case.

Entropy as a Measure of Entanglement

Consider a two-component (bipartite) quantum system $A+B$ in a pure state $|\Phi\rangle$. To test whether the subsystems are entangled or not, following the prescription of Sect. 1.3.3, we can perform the Schmidt decomposition of $|\Phi\rangle$. If this decomposition has more than one term, i.e., the Schmidt number is greater than one, $|\Phi\rangle$ is an entangled state, and the reduced density operator of one of the subsystems, say A , represents a mixed state, $\rho_{\text{mixed}}^A = \text{Tr}_B(|\Phi\rangle\langle\Phi|) = \sum_i p_i |\psi_i^A\rangle\langle\psi_i^A|$. On the other hand, for a factorisable state of the form $|\Phi\rangle = |\Psi^A\rangle \otimes |\Psi^B\rangle$, the Schmidt number is one, while the reduced density operator of A represents a pure state $\rho_{\text{pur}}^A = |\Psi^A\rangle\langle\Psi^A|$. Recall that the von Neumann entropy of a pure state ρ_{pur}^A is zero, $S(\rho_{\text{pur}}^A) = 0$, while it is maximized to $S(\rho_{\text{mixed}}^A) = \log_2 N$ for a mixed state ρ_{mixed}^A with all $p_i = 1/N$, which results from a maximally entangled state. Thus the von Neumann entropy is a monotonic function of entanglement between a pair of subsystems and is invariant under local unitary transformations, which do not

affect the entanglement. Therefore the von Neumann entropy of the density matrix for either subsystem of a bipartite system can serve as a convenient measure of entanglement, and is often referred to simply as the entropy of entanglement.

Considering now a two-component quantum system in a mixed state represented by the density operator ρ , one measure of entanglement is the so-called entanglement of formation $E(\rho)$, which is the minimum average entropy of entanglement of an ensemble of pure states $\{|\Phi\rangle\}$ that represents $\rho = \sum_{\Phi} P_{\Phi} |\Phi\rangle\langle\Phi|$. In general, for multistate subsystems the entanglement of formation is difficult to calculate, but in the simplest case of subsystems represented by two-state quantum systems (qubits), $E(\rho)$ can be expressed through the concurrence $C(\rho)$ defined as $C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$, where $\lambda_1, \dots, \lambda_4$ are the square roots of the eigenvalues of matrix $\rho\tilde{\rho}$, taken in decreasing order. Here $\tilde{\rho} \equiv \sigma_y^A \sigma_y^B \rho^* \sigma_y^A \sigma_y^B$, with ρ^* being the complex conjugate of ρ , and $\sigma_y^{A,B}$ the Pauli matrices acting on the corresponding qubits. Several other measures of entanglement of bipartite quantum systems have been suggested, including the distillable entanglement and logarithmic negativity. All of these measures reduce to the (von Neumann) entropy of entanglement in the case of pure states, and they share the important property of being the entanglement monotones, i.e., they do not increase under local operations and classical communication between the parties.

Quantifying entanglement of bipartite quantum systems is an important and difficult problem attracting at present much attention. In the general case of mixed states, the various measures of entanglement are nonequivalent, leading to much debate in the scientific community. The entanglement of three- and more-component systems represents an even more difficult problem. Usually one performs the pairwise decompositions of the compound system in all possible ways and then computes the measures of bipartite entanglements. This procedure, however, is capable of characterizing only certain aspects of multiparticle entanglement and for some states even fails to detect genuine entanglement. The characterization of multiparticle entanglement thus requires much further research.

Before closing this section, let us note that most of the experiments aimed at detecting entanglement between quantum systems employ the so-called entanglement witnesses. In general, entanglement witnesses are linear inequalities for expectation (mean) values of appropriate physical observables, which upon violating the inequalities verify the presence of entanglement in a bipartite system. The Bell inequalities described in the previous section are perhaps the most representative examples of entanglement witnesses.

Problems

8.1. Verify the truth table of Fig. 8.3.

8.2. Show that, if perfect cloning of quantum states were possible, then two distant parties, Alice and Bob, sharing a pair of entangled qubits in state $|B_{11}\rangle = \frac{1}{\sqrt{2}}(|0^A\rangle|1^B\rangle - |1^A\rangle|0^B\rangle)$, could communicate information with superluminal velocity. (Hint: Let Alice measure her qubit in either $\{|0\rangle, |1\rangle\}$ or $\{|+\rangle, |-\rangle\}$ basis. Now show that if Bob could perfectly clone his qubit, he would be able, with some probability, to find out Alice's basis.)

8.3. In the EPR protocol for quantum cryptography, in order to eavesdrop on the private key, Eve may employ the strategy of generating the three qubit entangled states $|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ and then sending the first qubit to Alice, the second one to Bob and keeping the third qubit to her own. Show that when Alice and Bob test the fidelity of their pairs by performing measurements in jointly determined random bases, on the average in one out of four measurements the correlation between the qubit states is violated. (Hint: Express the $|\text{GHZ}\rangle$ state in the $\{|+\rangle, |-\rangle\}$ basis.)

8.4. Prove that for any given four positive numbers x_1, x_2, y_1, y_2 , such that $0 \leq x_{1,2}, y_{1,2} \leq 1$, the inequality

$$-1 \leq x_1 y_1 - x_1 y_2 + x_2 y_1 + x_2 y_2 - x_2 - y_1 \leq 0$$

is always satisfied. (Hint: See J. F. Clauser and M. A. Horne, Phys. Rev. D 10, 526 (1974).)

8.5. Verify expressions (8.50) and (8.51) for the individual and joint detection probabilities $P_A(\theta_A)$, $P_B(\theta_B)$ and $P_{AB}(\theta_A, \theta_B)$ for a pair of particles in the entangled state (8.49). Also, calculate $P_A(\theta_A)$, $P_B(\theta_B)$ and $P_{AB}(\theta_A, \theta_B)$ for the two-particle singlet state $\frac{1}{\sqrt{2}}(|\uparrow_z^A\rangle|\downarrow_z^B\rangle - |\downarrow_z^A\rangle|\uparrow_z^B\rangle)$.

8.6. Given a pair of qubits in the pure entangled state $|\psi_2\rangle = \alpha|0^A\rangle|1^B\rangle + \beta|1^A\rangle|0^B\rangle$, calculate the concurrence $C(\rho)$ defined on the previous page. Show that for any two-qubit pure state $|\psi_2\rangle$, the concurrence is equal to

$$C(\psi_2) = |\langle\psi_2|\tilde{\psi}_2\rangle|, \quad \text{with} \quad |\tilde{\psi}_2\rangle \equiv \sigma_y^A \sigma_y^B |\psi_2^*\rangle. \quad (8.67)$$

(Hint: See W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998).)

Principles of Quantum Computation

In this chapter, after an outline of the general principles of operation of a quantum computer, we present several representative quantum algorithms for data processing and error correction. As we will see, these quantum algorithms can outperform their classical counterparts. The material here also serves to motivate the discussion in the next chapter pertaining to the physical implementations of quantum computation.

9.1 Operation of Quantum Computer

In the previous chapter we have studied the building blocks of a quantum information processing device. These are the qubits which constitute the register where quantum information is stored, and the quantum logic gates which manipulate that quantum information. We discuss now the Universal set of quantum gates and the principal steps involved in quantum computation.

9.1.1 Universal Gates for Quantum Computation

The Universal set of quantum gates, capable of realizing any unitary transformation on a multiqubit register, consists of the single-qubit rotational operations, such as $R_y(\theta)$ and $R_z(\theta)$ for spin- $\frac{1}{2}$ qubits or $R(\theta)$ and $T(\phi)$ for photon-polarization qubits, and the cnot two-qubit logic gate. This is to be contrasted with the reversible classical computation, where the universal logic gate—the Toffoli (or Fredkin) gate—involves three bits. As we show below, however, any three-qubit controlled-controlled-U operation (including the quantum Toffoli gate) can be decomposed into a sequence of two-qubit operations, which in turn can be implemented using the single-qubit rotations and cnot gate.

Consider a general unitary operation U on a single qubit,

$$U = \begin{pmatrix} e^{-i\beta} \cos \gamma & -e^{-i\delta} \sin \gamma \\ e^{i\delta} \sin \gamma & e^{i\beta} \cos \gamma \end{pmatrix}, \quad (9.1)$$

where, for simplicity, we disregard an overall phase factor $e^{i\alpha}$. From the definition of the rotation operators R_y and R_z in (8.25) and matrix multiplication, it follows that there exist real numbers θ_1 , θ_2 , and θ_3 , such that $U = R_z(\theta_1)R_y(\theta_2)R_z(\theta_3)$. Indeed, by choosing the three angles θ_i according to $\theta_1 + \theta_3 = 2\beta$, $\theta_1 - \theta_3 = 2\delta$, and $\theta_2 = 2\gamma$, we find that the product of the three rotations yields the right-hand-side of (9.1).

Next, we show that for any unitary matrix U , there exist matrices A , B and C such that $ABC = I$ and $AXBXC = U$. Indeed, by choosing $A = R_z(\theta_1)R_y(\theta_2/2)$, $B = R_y(-\theta_2/2)R_z(-(\theta_1 + \theta_3)/2)$, and $C = R_z((\theta_3 - \theta_1)/2)$, we have

$$\begin{aligned} ABC &= R_z(\theta_1)R_y\left[\frac{\theta_2}{2}\right]R_y\left[-\frac{\theta_2}{2}\right]R_z\left[-\frac{\theta_1 + \theta_3}{2}\right]R_z\left[\frac{\theta_3 - \theta_1}{2}\right] \\ &= R_z(\theta_1)R_z(-\theta_1) = I. \end{aligned} \quad (9.2)$$

Using the equalities $XX = I$, $XR_y(\theta)X = R_y(-\theta)$, and $XR_z(\theta)X = R_z(-\theta)$, we see that

$$\begin{aligned} AXBXC &= R_z(\theta_1)R_y\left[\frac{\theta_2}{2}\right]XR_y\left[-\frac{\theta_2}{2}\right]R_z\left[-\frac{\theta_1 + \theta_3}{2}\right]XR_z\left[\frac{\theta_3 - \theta_1}{2}\right] \\ &= R_z(\theta_1)R_y\left[\frac{\theta_2}{2}\right]XR_y\left[-\frac{\theta_2}{2}\right]XXR_z\left[-\frac{\theta_1 + \theta_3}{2}\right]XR_z\left[\frac{\theta_3 - \theta_1}{2}\right] \\ &= R_z(\theta_1)R_y(\theta_2)R_z(\theta_3) = U, \end{aligned} \quad (9.3)$$

which proves the statement.

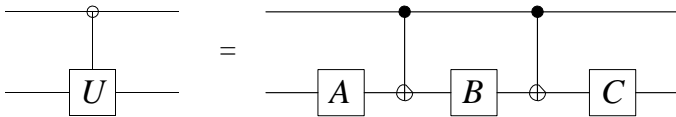


Fig. 9.1. Circuit generating the controlled-U gate.

Now it is easy to see that any two-qubit controlled-U operation can be simulated by the circuit of Fig. 9.1, which involves only the single-qubit operations A , B , and C , and the two-qubit **cnot** gates. Indeed, if the control (upper) qubit is in state $|0\rangle$, then the $ABC = I$ transformation is applied to the target (lower) qubit, while if the control qubit state is $|1\rangle$, the target qubit undergoes the transformation $AXBXC = U$.

It turns out that not only the **cnot** gate but almost any two-qubit logic gate is universal. In particular, the **cz** (controlled-Z) and **swap** (square-root of swap) gates are almost as good as the **cnot** gate, since we can easily construct circuits implementing the **cnot** transformation W_{cnot}^{AB} between qubits A and B through the W_{cz}^{AB} and W_{swap}^{AB} transformations,

$$W_{AB}^{\text{cnot}} = R_y^B(\pi/2) W_{BZ}^{AB} R_y^B(-\pi/2), \quad (9.4a)$$

$$W_{\text{cnot}}^{AB} = R_z^A(-\pi) R_z^A(\pi) W_{\text{swap}}^{AB} R_z^A(2\pi) W_{\text{swap}}^{AB}, \quad (9.4b)$$

where the $\sqrt{\text{swap}}$ gate, defined via $(W_{\text{swap}}^{AB})^2 = W_{\text{swap}}^{AB}$ has the following matrix representation,

$$W_{\text{swap}}^{AB} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1+i}{2} & \frac{1-i}{2} & 0 \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & e^{i\pi/4} & e^{-i\pi/4} & 0 \\ 0 & e^{-i\pi/4} & e^{i\pi/4} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{bmatrix}. \quad (9.5)$$

Controlled- Controlled- U

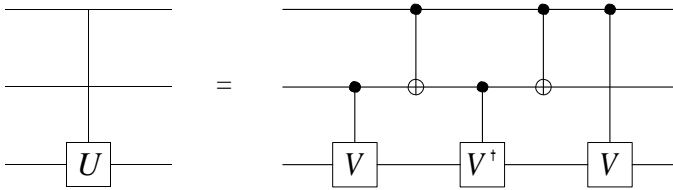


Fig. 9.2. Circuit generating the three-qubit controlled-controlled- U gate.

Consider, finally, multi-qubit operations. An example of the three-qubit controlled-controlled- U gate and its decomposition into a sequence of two-qubit operations is shown in Fig. 9.2, where the operator V is defined via $VV = U$ (see Prob. 9.1). When U corresponds to the X or not gate, and therefore $V = \frac{1}{\sqrt{2}}(1 - i)(I + iX)$, this circuit realizes the universal quantum Toffoli or ccnot gate. This ccnot gate can further be used to implement multiqubit gates having any number of control and target qubits.¹ In analogy with the reversible classical computer discussed in Sect. 7.4, using an equivalent sequence of quantum logic gates, any function $f(x) : \{0, 1\}^k \rightarrow \{0, 1\}^l$ is computed via the transformation $|x, y\rangle \xrightarrow{U_f} |x, y \oplus f(x)\rangle$, as shown in Fig. 9.3. The unitary transformation U_f leaves the “data” register of k qubits containing the argument $x \in \{0, 1\}^k$ of the function unchanged, while the value of the function is written into the “target” register of l qubits $y \in \{0, 1\}^l$ via the addition modulo 2 operation, $|y \oplus f(x)\rangle$.

¹Of course, if a particular physical system is capable of explicitly realizing unitary transformations involving more than two qubits, implementing thereby multiqubit logic gates, it is an advantageous but not a fundamental factor, since single- and two-qubit gates suffice to construct any multiqubit transformation.

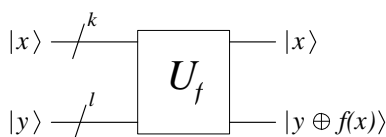


Fig. 9.3. Unitary transformation U_f for the evaluation of function $f(x)$. Symbol “/” is a short-hand notation for a set of n qubits.

9.1.2 Building Blocks of a Quantum Computer

In preparation for the discussion of the physical implementations of quantum information processing in the next chapter, let us list here the necessary ingredients of an envisioned quantum computer. The quantum computer is composed of (i) register containing K qubits; (ii) one- and two-qubit (and possibly $k(> 2)$ -qubit) logic gates applied to the register according to the particular algorithm; and (iii) measuring apparatus applied to the desired qubits at the end of (and, possibly, during) the program execution, which projects the qubit state onto the computational basis $\{|0\rangle, |1\rangle\}$.

The operation of the quantum computer consists of the following principal steps:

1. Initialization—Prepare all K qubits of the register in a well-defined initial state, such as, e.g., $|0\dots 000\rangle$.
2. Input—Load the input data using the logic gates.
3. Computation—Perform the desired unitary transformation by applying the sequence of logic gates according to the program.
4. Output—Measure the final state of the register in the computational basis.

In the following sections we consider several representative quantum algorithms for data processing and error correction.

9.2 Quantum Algorithms

Before we embark upon a detailed discussion, let us outline two general principles pertaining to most of the existing quantum algorithms. The first principle instructs us as to how to prepare the input state of the register, in order to make the best use of quantum parallelism. It may be formulated as follows:

- (i) Prepare the input state of the register in a superposition state of all possible “classical” inputs x

$$|\psi^{\text{in}}\rangle = \sum_x c_x |x\rangle, \quad \sum_x |c_x|^2 = 1. \quad (9.6)$$

Note that many problems are intractable on classical computers simply because there are too many possible inputs that should be processed and analyzed by a computer before the actual problem is solved. This usually requires either an enormous amount of CPU time, if a sequential processing of the input data is performed on a single processor, or enormous number of simultaneously working processors, to vastly parallelize the algorithm execution. On the other hand, owing to the quantum superposition principle, a single quantum register is capable of simultaneously storing and processing all of the classical inputs at once.

The second principle has to do with the measurement which should yield an intelligible result of computation, expressed in terms of classical quantities. It states:

- (ii) Design an algorithm in which all of the computational paths interfere with each other to yield with high probability the output state y

$$|\psi^{\text{out}}\rangle \subset c_y |y\rangle, \quad |c_y|^2 \subset 1, \quad \sum_{y' \neq y} |c_{y'}|^2 = 1. \quad (9.7)$$

We thus recognize that the required output state y , containing the sought after solution of the problem, is an eigenstate of the quantum register. It is thus a classical state, since no superposition is involved, and a single measurement reveals y with almost unit probability $|c_y|^2 \subset 1$. If the algorithm execution is not too costly in terms of the hardware involved, we may allow for several repetitions of the algorithm, followed by the measurements, and the condition on c_y can be somewhat relaxed to $|c_y|^2 > 1/2$. Then if we perform say N_r repetitions of the cycle (i)-(ii), all resulting in the same state y , the probability of obtaining an erroneous output $P_{\text{error}}(N_r) = (1 - |c_y|^2)^{N_r}$ rapidly goes to zero with increasing N_r .

Designing good quantum algorithms is a very difficult task requiring profound insight and ingenuity, particularly because we think largely in classical terms, with the quantum world often being rather counterintuitive.

9.2.1 Deutsch Algorithm

We begin with this simple algorithm, involving only two qubits, representing therefore a "toy problem", aimed at demonstrating the power of quantum parallelism employed in other quantum algorithms capable of performing useful tasks.

Suppose we are given some Boolean function $f(x) : \{0, 1\} \rightarrow \{0, 1\}$ of a single-bit argument x and the corresponding quantum "black-box" or "oracle" U_f evaluating that function via $|x, y\rangle \rightarrow |x, y \oplus f(x)\rangle$ (see. Prob. 9.2). Our aim is to determine a property of function f , i.e., whether it is constant $f(0) = f(1)$ or balanced $f(0) \neq f(1)$. On a classical computer, we would need to evaluate $f(x)$ twice, once for $x = 0$ and once for $x = 1$, and then

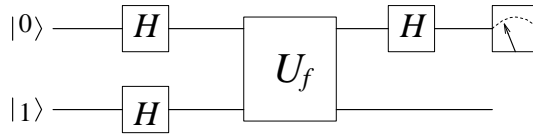


Fig. 9.4. Circuit implementing the Deutsch algorithm.

compare the results. A quantum computer, however, allows us to characterize the function by calling U_f only once. This is achieved by preparing the data qubit in an equally weighted superposition state of both classical inputs $x = 0$ and $x = 1$ and evaluating $f(x)$ for both x at the same time, followed by interfering the output of U_f and a single measurement that unambiguously reveals the function. The quantum circuit for doing this is shown in Fig. 9.4. Let us follow the states of the register through this circuit. The initial state of the two qubits is

$$|\psi_2^{(0)}\rangle = |0\rangle |1\rangle. \quad (9.8)$$

After applying the Hadamard gates to the data and target qubits, we have

$$|\psi_2^{(1)}\rangle = \frac{1}{2} (|0\rangle + |1\rangle)(|0\rangle - |1\rangle). \quad (9.9)$$

Next, the U_f transformation leads to

$$|\psi_2^{(2)}\rangle = \begin{cases} \pm \frac{1}{2} (|0\rangle + |1\rangle)(|0\rangle - |1\rangle) & \text{if } f(0) = f(1) \\ \pm \frac{1}{2} (|0\rangle - |1\rangle)(|0\rangle - |1\rangle) & \text{if } f(0) \neq f(1). \end{cases}$$

Realizing that $|x\rangle(|0\rangle - |1\rangle)/\sqrt{2} \xrightarrow{U_f} (-1)^{f(x)} |x\rangle(|0\rangle - |1\rangle)/\sqrt{2}$, we can cast $|\psi_2^{(2)}\rangle$ in a more compact form

$$|\psi_2^{(2)}\rangle = \sum_{x=0,1} \frac{(-1)^{f(x)} |x\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (9.10)$$

Finally, after applying the Hadamard gate to the data qubit, we have

$$|\psi_2^{(3)}\rangle = \begin{cases} \pm |0\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) & \text{if } f(0) = f(1) \\ \pm |1\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) & \text{if } f(0) \neq f(1), \end{cases}$$

or, alternatively,

$$|\psi_2^{(3)}\rangle = \pm |f(0) \oplus f(1)\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \quad (9.11)$$

Thus, measurement on the data qubit yields state $|f(0) \oplus f(1)\rangle$, which is $|0\rangle$ if $f(0) = f(1)$ [$f(0) \oplus f(1) = 0$], and is $|1\rangle$ if $f(0) \neq f(1)$ [$f(0) \oplus f(1) = 1$]. This illustrates the power of superposition employed in quantum computation.

9.2.2 Deutsch–Jozsa Algorithm

The Deutsch–Jozsa algorithm is a generalization of the above algorithm to the multiqubit case. Namely, suppose that a Boolean function $f(x) : \{0, 1\}^k \rightarrow \{0, 1\}$ of a k -bit argument $x \equiv x_1 x_2 \dots x_k$ is computed by the quantum black-box U_f via $|x, y\rangle \rightarrow |x, y \oplus f(x)\rangle$. As before, our aim is to determine whether the function is constant $f(x) = \text{const}$, i.e., its value is the same for all $0 \leq x < 2^k$, or balanced, meaning that $f(x) = 0$ for exactly half of all possible x , and $f(x) = 1$ for the other half of the values of x . We are guaranteed that f is either constant or balanced. On a classical computer, we would need to evaluate $f(x)$ at least twice for two different arguments x and x^r , provided that the values of f for x and x^r are different, in which case we know that the function is balanced. If however, the values of f are the same, we need to call U_f with yet another argument $x^{rr} = x$, and compare the output with the previous values of f . Again, if these values are different, we learn that f is balanced, otherwise we have to test more x . Only after $2^{k/2} + 1$ queries of the function with different arguments yielding the same result we can be certain that the function is constant. The quantum circuit of Fig. 9.5, however, can give a definite answer to that problem after only a single evaluation of f for a superposition of all x .

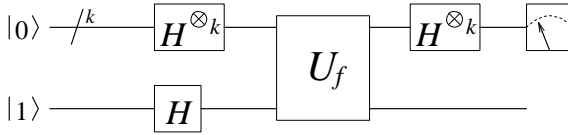


Fig. 9.5. Circuit implementing the Deutsch–Jozsa algorithm.

Let us follow the states of the register through this circuit. The data register is composed of k qubits, initially all in state $|0\rangle$, and the target is a single qubit, initially in state $|1\rangle$. Thus, the initial state of the system is

$$|\psi_{k+1}^{(0)}\rangle = |0\rangle^{\otimes k} |1\rangle. \quad (9.12)$$

Applying the Hadamard gates to the data and target qubits, we prepare the data register in the equally weighted superposition state of all possible x ,

$$|\psi_{k+1}^{(1)}\rangle = \frac{1}{\sqrt{2^k}} \sum_x |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (9.13)$$

Since the target qubit is prepared in state $(|0\rangle - |1\rangle)/\sqrt{2}$, the U_f transformation leads to

$$|\psi_{k+1}^{(2)}\rangle = \frac{1}{\sqrt{2^k}} \sum_x (-1)^{f(x)} |x\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (9.14)$$

Finally, the Hadamard transformation is applied to all k qubits of the data register, which yields the state

$$|\Psi_{k+1}^{(3)}\rangle = \sum_z \sum_x \frac{(-1)^{x \cdot z + f(x)} |z\rangle}{2^k} \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (9.15)$$

Here we have used the multiqubit generalization of (8.7),

$$H^{\otimes k} |x_1 \dots x_k\rangle = \sum_{z_1, \dots, z_k} \frac{(-1)^{x_1 z_1 + \dots + x_k z_k} |z_1 \dots z_k\rangle}{\sqrt{2^k}},$$

which, in a compact form, is

$$H^{\otimes k} |x\rangle = \sum_z \frac{(-1)^{x \cdot z} |z\rangle}{\sqrt{2^k}},$$

where $x \cdot z = x_1 z_1 + \dots + x_k z_k$ is the bitwise dot product of x and z .

The remarkable property of the final state (9.15) is that the amplitude c_0 of the state $|z = 0\rangle = |0\rangle^{\otimes k}$ of the data register is given by

$$c_0 = (\Psi_{k+1}^{(3)} |0\rangle)^{\otimes k} = \sum_x \frac{|0\rangle - |1\rangle}{\sqrt{2}} \frac{(-1)^{f(x)}}{2^k}.$$

Therefore, if the function f is constant, all 2^k terms enter this sum with the same sign ("+" for $f(x) = 0$ and "-" for $f(x) = 1$) and we have $c_0 = \pm 1$. On the other hand, if the function f is balanced, exactly half of the terms enter this sum with the "+" sign, and the other half enter with the "-" sign, which results in $c_0 = 0$. Thus, if the function f is constant, the measurement of the data register yields, with the probability $|c_0|^2 = 1$, the output state $|0\rangle^{\otimes k}$, i.e., all qubits are in state $|0\rangle$. If, on the other hand, the measurement yields $|c_0|^2 = 0$, i.e., at least one of qubits of the data register is found in state $|1\rangle$, the function f is balanced.

We have thus seen that, despite of its limited use for practical applications, the Deutsch–Jozsa algorithm unambiguously demonstrates the potential power of a quantum computer. In what follows, we describe two very useful quantum algorithms that can significantly speed up and efficiently solve two classes of important problems involving search/minimization and Fourier transform.

9.2.3 Grover Algorithm

The quantum search algorithm was invented by Grover, and it offers a quadratic speed-up over classical algorithms for searching an unsorted database. Suppose we have a list of $N = 2^k$ elements d_x stored in a computer memory. The index x , taking N different integer values $x = 0, 1, \dots, N - 1$, identifies