Solutions of problems of Quantum computation and quantum information by M.A. Nielsen and I.L. Chuang

Árpád Lukács

September 30, 2024

Abstract

See also the github repository for the LaTeXsource files and computer algebra (Reduce) calculations.

Part I
Fundamental concepts

1 Introduction and overview

Ex. 1.1. Let us generate k uniformly distributed random values of x and evaluate $f(x)$. If all the values are the same, we assume that the function is constant, otherwise, that it is balanced. If it is indeed constant, there is no possibility of making and error and claiming it to be balanced. If it is balanced, the probability of all values being of one type is $\varepsilon = 2 \times (1/2^k)$m which tells us that the number of necessary attempts is $k = 1 - \log_2 \varepsilon$.

Ex. 1.2. Let us create the operator D acting as follows: $D|\psi\rangle = |0\rangle$ and $D|\phi\rangle = |1\rangle$, the matrix of which is solved from the equations

$$\begin{aligned} D_{00}\psi_0 + D_{01}\psi_1 = 1\,, \qquad & D_{10}\psi_0 + D_{11}\psi_1 = 0\,, \\ D_{00}\varphi_0 + D_{01}\varphi_1 = 0\,, \qquad & D_{10}\varphi_0 + D_{11}\varphi_1 = 1\,, \end{aligned}$$

or in matrix form

$$D \begin{pmatrix} \psi_0 & \varphi_0 \\ \psi_1 & \varphi_1 \end{pmatrix} = I\,,$$

yielding

$$D = \frac{1}{\psi_0\varphi_1 - \varphi_0\psi_1} \begin{pmatrix} \varphi_1 & -\varphi_0 \\ -\psi_1 & \psi_0 \end{pmatrix}.$$

Let us now apply $(D^{-1} \otimes D^{-1})CNOT(D \otimes I)$ to $|\psi\rangle \otimes |0\rangle$ or $|\varphi\rangle \otimes |0\rangle$. The result is a clone. Where does this fail? D is not unitary, unless the states are orthogonal.

Conversely, if we could clone states, then making many copies and measuring in a basis orthogonal to, say, ψ, and then measure half of them in a basis $\Psi = (\psi, \psi')$ where $\langle\psi, \psi'\rangle = 0$, and half of them in the analogous basis to φ. If we made enough copies, only for either ψ or ψ' will the result never be the primed one, and that is the input state.

2 Introduction to quantum mechanics

2.1 Linear algebra

Ex. 2.1.

$$\begin{pmatrix}1\\-1\end{pmatrix}+\begin{pmatrix}1\\2\end{pmatrix}-\begin{pmatrix}2\\1\end{pmatrix}=0\,.$$

Ex. 2.2. In the basis $|0\rangle, |1\rangle$, the matrix representation is

$$(A_{ij})=\begin{pmatrix}0&1\\1&0\end{pmatrix}\,.$$

If we chose different input or output bases, the matrix would change, e.g., choosing the same input bases, but the output basis $|1\rangle, |0\rangle$ would yield

$$A'=\begin{pmatrix}1&0\\0&1\end{pmatrix}\,.$$

Ex. 2.3. If $A : V \to W$ is an operator, then

$$A\sum_k v_k|k\rangle=\sum_{jk}A_{jk}v_k|j\rangle\,,$$

and similarly for $B : W \to X$,

$$B\sum_j w_j|j\rangle=\sum_{ij}B_{ij}|i\rangle\,.$$

Let now $w = Av$, so $w_j = \sum_k A_{jk}v_k$, yielding

$$BAv=\sum_{ijk}B_{ij}A_{jk}v_k|i\rangle\,,$$

where we can recognise the matrix product

$$(BA)_{ik}=\sum_j B_{ij}A_{jk}\,.$$

Ex. 2.4. The definition of the identity operator is that it takes all vectors into themselves, including those of a given basis,

$$I|j\rangle=|j\rangle\,.$$

We may write the left hand side as $\sum_i I_{ij}|i\rangle$, and the right hand side as $\sum_i \delta_{ij}|j\rangle$, from which we may read off the matrix elements of I, $I_{ij} = \delta_{ij}$, which are the elements of a matrix with 1 in the diagonal and 0 elsewhere.

Ex. 2.5. The mapping $(\cdot,\cdot) : \mathbb{C}^n \to \mathbb{C}$ is an inner product:

(1) $(v, \sum \lambda_i w_i) = \sum_j v_j^* \sum_i \lambda_i (w_i)_j = \sum_i \lambda_i \sum_j v_j^* (w_i)_j = \sum_i \lambda_i (v, w_j)$,

(2) $(v, w) = \sum_i v_i^* w_i = (\sum_i v_i w_i^*)^* = (w, v)^*$,

(3) $(v, v) = \sum_i v_i^* v_i = \sum_i |v_i|^2 \geq 0$, and there is equality only if all $v_i = 0$, i.e., if $v = 0$.

Ex. 2.6. $(\sum_i \lambda_i v_i, w) = (w, \sum_i \lambda_i v_i)^* = [\sum_i \lambda_i (w, v_i)]^* = \sum_i \lambda_i^* (w, v_i)^* = \sum_i \lambda_i^* (v_i, w)$.

Ex. 2.7. $\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = (1,1) \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$. To calculate norms, we use the formula derived in the previous exercise, showing that both vectors have a norm of $\sqrt{2}$, and so they can both be normalised by multiplying them with $1/\sqrt{2}$.

Ex. 2.8. The vectors obtained using the Gram-Schmidt procedure are all normalised. Orthogonality is shown as follows:

$$\langle v_j | v_k \rangle = \frac{1}{N_k} \left\langle v_j \left| w_k - \sum_{i=1}^{k-1} \langle v_i | w_k \rangle v_i \right. \right\rangle = \frac{1}{N_k} \left(\langle v_j | w_k \rangle - \sum_{i=1}^{k-1} \langle v_i | w_k \rangle \langle v_j | v_i \rangle \right) ,$$

where N_k denotes the denominator used when normalising v_k [see eq. (2.17) in the book].

To proceed further, we shall assume without loss of generality that $j < k$, and prove by induction, first for $k = j + 1$, in which case we obtain

$$\langle v_j | v_{j+1} \rangle = \frac{1}{N_k} \left(\langle v_j | w_{j+1} \rangle - \sum_{i=1}^{j} \langle v_i | w_{j+1} \rangle \langle v_j | v_i \rangle \right) = \frac{1}{N_k} (\langle v_j | w_{j+1} \rangle - \langle v_j | w_{j+1} \rangle) = 0 \,,$$

where we have used that for $i < j + 1$ $\langle v_j | v_i \rangle = \delta_{i,j}$. Second, the induction step, using the same intermediate:

$$\langle v_j | v_{k+1} \rangle = \frac{1}{N_{k+1}} \left(\langle v_j | w_{k+1} \rangle - \sum_{i=1}^{k} \langle v_i | w_k \rangle \langle v_j | v_i \rangle \right) = \frac{1}{N_{k+1}} \left(\langle v_j | w_{k+1} - \langle v_j | w_k \rangle \right) = 0 \,.$$

Ex. 2.9. The outer product representation of any operator may be obtained using eq. (2.25) of the book,

$$\sigma_0 = I = |0\rangle\langle 0| + |1\rangle\langle 1| \,, \qquad \sigma_1 = |0\rangle\langle 1| + |1\rangle\langle 0| \,,$$
$$\sigma_2 = -\mathrm{i}|0\rangle\langle 1| + \mathrm{i}|1\rangle\langle 0| \,, \qquad \sigma_3 = |0\rangle\langle 0| - |1\rangle\langle 1| \,.$$

Ex. 2.10. Let $A = |v_j\rangle\langle v_k|$. The matrix representation of this operator is

$$A_{i\ell} = \langle v_i | A | v_\ell \rangle = \langle v_i | v_j \rangle \langle v_k | v_\ell \rangle = \delta_{ij} \delta_{k\ell} \,,$$

where in the last step we used the orthonormality of the basis. The result is a matrix with a 1 in the jk element (jth row kth column) and 0 everywhere else.

Ex. 2.11. For the 0 and 3 Pauli matrices, see the previous one. For all Pauli matrices, the characteristic polynomial is $c(\lambda) = \lambda^2 - 1$, with solutions $\lambda = \pm 1$. The corresponding eigenvectors are easily read off:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_0 - 1 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} , \quad \sigma_0 + 1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} ,$$

from where

$$|1\rangle_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \,, \qquad |-1\rangle_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \,.$$

Similarly,

$$\sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_0 - 1 = \begin{pmatrix} -1 & -\mathrm{i} \\ \mathrm{i} & -1 \end{pmatrix}, \quad \sigma_0 + 1 = \begin{pmatrix} 1 & -\mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix},$$

leading to

$$|1\rangle_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle + \mathrm{i}|1\rangle), \quad |-1\rangle_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle - \mathrm{i}|1\rangle).$$

Ex. 2.12. The characteristic polynomial of the matrix is

$$c(\lambda) = \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} - \lambda = x^2 - 2x + 1 = (x-1)^2,$$

i.e., the (degenerate) eigenvalue is 1. The matrix which should annihilate the eigenvectors is

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

and as this has the sole normalised solution $(0,1)^T$, the corresponding eigenspace is 1 dimensional, and the eigenspaces do not span $\mathbb{C}^2$.

Ex. 2.13. The adjoint of a dyad $|v\rangle\langle w|$:

$$\langle x \,|(|v\rangle\langle w|)\, y\rangle = \left\langle (|v\rangle\langle w|)^\dagger \, x \middle| \, y \right\rangle,$$

and evaluating the left hand side,

$$\langle x \,|(|v\rangle\langle w|)\, y\rangle = \langle x|v\rangle\langle w|y\rangle = \langle\, (|w\rangle\langle v|)\, x|\, y\rangle,$$

which, compared with the right hand side of the first equation yields that

$$(|v\rangle\langle w|)^\dagger = |w\rangle\langle v|.$$

Ex. 2.14.

$$\left\langle \left(\sum_i a_i A_i\right)^\dagger v \middle| \, w \right\rangle = \left\langle v \middle| \left(\sum_i a_i A_i\right) w \right\rangle = \sum_i a_i \langle v|A_i w\rangle = \sum_i a_i \left\langle A_i^\dagger v \middle| \, w \right\rangle$$
$$= \left\langle \left(\sum_i a_i^* A_i^\dagger\right) v \middle| \, w \right\rangle,$$

which tells us that

$$\left(\sum_i a_i A_i\right)^\dagger = \sum_i a_i^* A_i^\dagger.$$

Ex. 2.15.

$$\left\langle \left(A^\dagger\right)^\dagger v \middle| \, w \right\rangle = \langle v \,|A^\dagger w\,\rangle = \left\langle A^\dagger w \middle| \, v \right\rangle^* = \langle w \,|Av\rangle^* = \langle\, Av|\, w\rangle\,.$$

Ex. 2.16. Using the orthonormality of the basis, $\langle i|j\rangle = \delta_{ij}$,

$$P^2 = \left(\sum_{i=1}^k |i\rangle\langle i|\right)\left(\sum_{j=1}^k |j\rangle\langle j|\right) = \sum_{i,j=1}^k |i\rangle\langle i||j\rangle\langle j| = \sum_{i,j=1}^k \langle i|j\rangle |i\rangle\langle j| = \sum_{i=1}^k |i\rangle\langle i| = P.$$

Ex. 2.17. Using the spectral decomposition of the normal matrix, and the adjoint of a dyad,

$$\left(\sum_i \lambda_i |i\rangle\langle i|\right)^\dagger = \sum_i \lambda^* |i\rangle\langle i|\,,$$

which agrees with the original operator iff $\lambda_i = \lambda_i^*$ for all i.

Ex. 2.18. Again, use the spectral decomposition,

$$\left(\sum_i \lambda_i |i\rangle\langle i|\right)^\dagger \sum_j \lambda_j |j\rangle\langle j| = \sum_i \lambda_i^* \lambda_i |i\rangle\langle i|\,,$$

and that the identity operator is $\sum_i |i\rangle\langle i|$, from which we may read off that unitarity is equivalent to $|\lambda_i|^2 = \lambda_i^*\lambda_i = 1$ for all i, or $\lambda_i = \exp i\alpha_i$, α_i real.

Ex. 2.19. Direct calculation,

$$\sigma_0^\dagger = I^\dagger = I = \sigma_0\,, \qquad \sigma_1^\dagger = a\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}^\dagger = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}^T = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \sigma_1\,,$$

$$\sigma_2^\dagger = \begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}^\dagger = \begin{pmatrix}0 & \mathrm{i}\\ -\mathrm{i} & 0\end{pmatrix}^T = \begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix} = \sigma_2\,, \quad \sigma_3^\dagger = \begin{pmatrix}1 & \\ & -1\end{pmatrix}^\dagger = \begin{pmatrix}1 & \\ 0 & -1\end{pmatrix} = \sigma_3\,.$$

To verify unitarity, again, calculate directly

$$\sigma_i^\dagger \sigma_i = (\sigma_i)^2 = I\,.$$

Ex. 2.20. The matrix elements of an operator are obtained as $A'_{ij} = \langle v_i|A|v_j\rangle$, so, using the completeness relation,

$$A = \sum_{ij} |v_i\rangle\langle v_i|A|v_j\rangle\langle v_j| = \sum_{ij} A'_{ij}|v_i\rangle\langle v_j|\,.$$

Calculating the matrix elements in another basis yields

$$A''_{ij} = \langle w_i|A|w_j\rangle = \sum_{k\ell}\langle w_i||v_k\rangle\langle v_k|A|v_\ell\rangle\langle v_\ell||w_j\rangle = \sum_{k\ell}\langle w_i|v_k\rangle\langle v_k|A|v_\ell\rangle\langle v_\ell|w_j\rangle = \sum_{k\ell} U^*_{ki}A'_{k\ell}U_{\ell j}\,,$$

or in matrix form, $A'' = U^\dagger A' U$, where U is a matrix whose elements are $U_{ij} = \langle v_i|w_j\rangle$, unitary due to both bases being orthonormal.

Ex. 2.21. The following simplifications of the proof are possible: 1) $QMP = (PMQ)^\dagger = 0$. 2) The self-adjointness of QMQ is a direct consequence of $(QMQ)^\dagger = Q^\dagger M^\dagger Q^\dagger = QMQ$.

Ex. 2.22. Let $M = M^\dagger$, and $M|1\rangle = \lambda_1|1\rangle$ and $M|2\rangle = \lambda_2|2\rangle$, $\lambda_1 \neq \lambda_2$. In this case,

$$\langle 1|M|2\rangle = \langle 1\,|(M|2\rangle)\rangle = \langle 1|\lambda_2|2\rangle = \lambda_2\langle 1|2\rangle\,,$$

and similarly

$$\langle 1|M|2\rangle = \langle M^\dagger 1|2\rangle = \lambda_1\langle 1|2\rangle\,,$$

where we have used the fact that the eigenvalues of a self-adjoint operator are real, and so

$$(\lambda_1 - \lambda_2)\langle 1|2\rangle = 0\,.$$

Ex. 2.23. In the eigenbasis of the operator, $P = \sum_i |i\rangle\langle i|$, and also

$$P^2 = \left(\sum_i \lambda_i |i\rangle\langle i|\right)^2 = \sum_i \lambda_i^2 |i\rangle\langle i|\,,$$

where we have used the fact that the basis is orthonormal. All eigenvalues are real and satisfy $\lambda_i^2 = \lambda_i$, so $\lambda_i \in \{0, 1\}$.

Ex. 2.24. Let A a positive operator, and $B = (A + A^\dagger)/2$ and $C = (A - A^\dagger)/(2\mathrm{i})$. This way, $A = B + \mathrm{i}C$. For any vector,

$$\langle x|A|x\rangle = \langle x|B|x\rangle + \mathrm{i}\langle x|C|x\rangle\,,$$

and this is a real, positive number. As

$$\langle x|B|x\rangle = \langle x|B^\dagger|x\rangle^* = \langle x|B|x\rangle^*\,,$$

$\langle x|B|x\rangle$ is always real, and so is $\langle x|C|x\rangle$. As $\langle x|A|x\rangle$ is also real, this is only possible if for all vectors of x, $\langle x|C|x\rangle = 0$.

If for an operator C $\langle x|C|x\rangle = 0$ for all x, $C = 0$. Let us consider two vectors, x, y, and for such an operator

$$0 = \langle x + \alpha y|C|x + \alpha y\rangle = \langle x|C|\alpha y\rangle + \langle \alpha y|C|x\rangle = \alpha\langle x|C|y\rangle + \alpha^*\langle y|C|x\rangle\,,$$

and now adding this equation with $\alpha = 1$ to $-\mathrm{i}$ times this equation with $\alpha = \mathrm{i}$ yields $\langle x|C|y\rangle = 0$, i.e., all matrix elements of C vanish, therefore so does C. As a result, $A = B$ and $B = B^\dagger$, so A is self-adjoint.

Ex. 2.25. Take an arbitrary vector x,

$$\langle x|A^\dagger A|x\rangle = \langle Ax|Ax\rangle = \|Ax\|^2 \geq 0\,.$$

Ex. 2.26. Let $\psi = (|0\rangle + |1\rangle)/\sqrt{2}$. Then

$$\begin{aligned}
\psi^{\otimes 2} &= \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)\,,\\
\psi^{\otimes 3} &= \psi^{\otimes 2} \otimes \psi = \frac{1}{2^{3/2}}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \otimes (|0\rangle + |1\rangle)\\
&= \frac{1}{2^{3/2}}(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle)\,.
\end{aligned}$$

To do the same in terms of the Kronecker product, we shall use $\psi = (1, 1)^T/\sqrt{2}$, and so

$$\psi^{\otimes 2} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}\,,$$

$$\psi^{\otimes 3} = \frac{1}{2^{3/2}}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{pmatrix}\,.$$

Ex. 2.27. Tensor products of Pauli operators:

$$X \otimes Z = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \otimes \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = \begin{pmatrix} 0 & Z \\ Z & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

and

$$I \otimes X = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \otimes \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} X & \\ & X \end{pmatrix} = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{pmatrix},$$

and similarly

$$X \otimes I = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \otimes \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = \begin{pmatrix} I & \\ & I \end{pmatrix} = \begin{pmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & 1 & & \end{pmatrix}.$$

As we can see, $I \otimes X \neq X \otimes I$, the tensor product is non-commutative.

Ex. 2.28. As the matrix elements of the tensor product are given by eq. (2.50) in the book,

$$(A \otimes B)^* = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \dots & \dots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix}^* = \begin{pmatrix} A^*_{11}B^* & A^*_{12}B^* & \dots & A^*_{1n}B^* \\ A^*_{21}B^* & A^*_{22}B^* & \dots & A^*_{2n}B^* \\ \dots & \dots & \ddots & \vdots \\ A^*_{m1}B^* & A^*_{m2}B^* & \dots & A^*_{mn}B^* \end{pmatrix} = A^* \otimes B^*,$$

Similarly,

$$(A \otimes B)^T = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \dots & \dots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix}^T = \begin{pmatrix} A_{11}B^T & A_{21}B^T & \dots & A_{m1}B^T \\ A_{12}B^T & A_{22}B^T & \dots & A_{m2}B^T \\ \dots & \dots & \ddots & \vdots \\ A_{1n}B^T & A_{2n}B^T & \dots & A_{nm}B^T \end{pmatrix} = A^T \otimes B^T,$$

and $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$ already follows from $X^\dagger = (X^T)^*$.

Ex. 2.29. If $U^\dagger U = I = V^\dagger V$, then

$$(U \otimes V)^\dagger (U \otimes V) = (U^\dagger \otimes V^\dagger)(U \otimes V) = U^\dagger U \otimes V^\dagger V = I \otimes I = I\,,$$

and similarly with taking the adjoint of the second operator in stead of the first.

Ex. 2.30. If $H_i^\dagger = H_i$, $i = 1, 2$, then

$$(H_1 \otimes H_2)^\dagger = H_1^\dagger \otimes H_2^\dagger = H_1 \otimes H_2\,.$$

Ex. 2.31. Let us choose on the two spaces the eigenbasis of the operators A, B, then, using eq. (2.50) from the book,

$$A \otimes B = \begin{pmatrix} \lambda_1\mu_1 & 0 & \dots & 0 \\ 0 & \lambda_2\mu_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \lambda_n\mu_m \end{pmatrix},$$

where λ_i and μ_j are the eigenvalues of A and B, respectively. This is a positive matrix as all its eigenvalues are real and positive.

Ex. 2.32. Let P, Q be projectors, $P^2 = P$, $Q^2 = Q$, then

$$(P \otimes Q)^2 = P^2 \otimes Q^2 = P \otimes Q\,.$$

Ex. 2.33. The formula clearly holds for $n = 1$. Then

$$\begin{aligned} H^{\otimes n} &= \frac{1}{2^n} \sum_{x,y} (-1)^{x\cdot y} |x\rangle\langle y| \otimes \left[(|0\rangle + |1\rangle)\langle 0| + (|0\rangle - |1\rangle)\langle 1| \right] \\ &= \frac{1}{2^n} \sum_{x,y} (-1)^{x\cdot y} \left[(|x0\rangle + |x1\rangle)\langle y0| \ + (|x0\rangle - |x1\rangle)\langle y1| \right] \\ &= \frac{1}{2^n} \sum_{x'y'} = (-1)^{x'\cdot y'} |x'\rangle\langle y'|\,. \end{aligned}$$

For the $4 \otimes 4$ case, we need

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

yielding, using eq. (2.50)

$$H^{\otimes 2} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Ex. 2.34. The characteristic polynomial of the matrix is $p(\lambda) = (4-\lambda)^2 - 9 = (7-\lambda)(1-\lambda)$, so the eigenvalues are 7 and 1, so

$$A = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1, -1) + \frac{7}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1, 1)\,.$$

Using this, the square root is

$$\sqrt{A} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1, -1) + \frac{\sqrt{7}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1, 1) = \frac{1}{2} \begin{pmatrix} \sqrt{7}+1 & \sqrt{7}-1 \\ \sqrt{7}-1 & \sqrt{7}+1 \end{pmatrix},$$

and the logarithm,

$$\log A = \frac{\log 7}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1, 1) = \frac{\log 7}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Ex. 2.35. For all the Pauli matrices, $A^2 = I$ holds. Elementary algebra shows that $\sigma_i \sigma_j = \delta_{ij} + \mathrm{i}\varepsilon_{ijk}\sigma_k$, so for any unit vector $\vec{v}$, $(\vec{v}\vec{\sigma})^2 = I$ holds, and splitting the power series of the exponential for $n = 2k$ and $n = 2k+1$,

$$\exp(i\theta\vec{v}\vec{\sigma}) = \sum_n \frac{i^n \theta^n (\vec{v}\vec{\sigma})^n}{n!} = \sum_k \frac{\theta^{2k}}{(2k)!} I + \sum_k \frac{\mathrm{i}(-1)^k \theta^{2k+1}}{(2k+1)!} \vec{v}\vec{\sigma} = \cos\theta I + \mathrm{i}\sin\theta(\vec{v}\vec{\sigma})\,.$$

Ex. 2.36. The trace of the Pauli matrices, $\operatorname{Tr} A = \sum_i A_{ii}$, so

$$\operatorname{Tr}\sigma_0 = \operatorname{Tr} I = 2\,, \quad \operatorname{Tr}\sigma_1 = \operatorname{Tr}\begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = 0\,,$$
$$\operatorname{Tr}\sigma_2 = \operatorname{Tr}\begin{pmatrix} & -\mathrm{i} \\ \mathrm{i} & \end{pmatrix} = 0\,, \quad \operatorname{Tr}\sigma_3 = \operatorname{Tr}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = 0\,.$$

Ex. 2.37. For any matrices A, B,

$$\operatorname{Tr}(AB) = \sum_{ij} A_{ij}B_{ji} = \sum_{ij} B_{ji}A_{ij} = \operatorname{Tr}(BA)\,.$$

Ex. 2.38. For any matrices A, B,

$$\operatorname{Tr}(A+B) = \sum_i (A+B)_{ii} = \sum_i (A_i i + B_i i) = \operatorname{Tr} A + \operatorname{Tr} B\,,$$

and for any number z,

$$\operatorname{Tr}(zA) = \sum_i (zA)_i i = \sum_i zA_{ii} = z\sum_i A_{ii} = z\operatorname{Tr} A\,.$$

Ex. 2.39. Proving the scalar product nature of the Hilbert-Schmidt product requires linearity in the second argument,

$$(A, B+C) = \operatorname{Tr} A^\dagger(B+C) = \operatorname{Tr} A^\dagger B + \operatorname{Tr} A^\dagger C = (A,B) + (A,C)\,,$$

and

$$\operatorname{Tr}(A, zB) = \operatorname{Tr} A^\dagger zB = z\operatorname{Tr} A^\dagger B = z(A,B)\,,$$

and the exchange property,

$$(B,A) = \operatorname{Tr} B^\dagger A = \sum_{ij} B^*_{ji}A_{ij} = \left(\sum_{ij} A^*_{ji}B_{ij}\right)^* = (A,B)^*\,,$$

and positivity,

$$(A,A) = \sum_{ij} A^*_{ji}A_{ji} = \sum_{ij} |A_{ij}|^2\,.$$

A nice orthogonal basis is $E^{(ij)}$ having matrix elements $(E^{(ij)}_{k\ell} = \delta_{ik}\delta_{k\ell}$ as clearly any matrix has $A = \sum_{ij} A_{ij}E^{(ij)}$, and $(E^{(ij)}E^{(k\ell)} = \sum_{no} E^{(ij)}_{on}E^{(k\ell)}_{on} = \sum_{on} \delta_{io}\delta_{jn}\delta_{ko}\delta_{\ell n} = \delta_{ik}\delta_{j\ell}$.

Ex. 2.40. Perform matrix multiplications (see `ex.2.40.red` for computer algebra). A concise result is

$$\sigma_i\sigma_j = \delta_{ij}I + \mathrm{i}\epsilon_{ijk}\sigma_k\,,$$

where there is an implicit summation over the repeated index k.

Ex. 2.41. See previous one.

Ex. 2.42.

$$\frac{[A,B]+\{A,B\}}{2} = \frac{AB-BA+AB+BA}{2} = AB\,.$$

Ex. 2.43. See Ex. 2.40.

Ex. 2.44. If both commutators vanish, $AB = [A, B] + \{A, B\} = 0$, which, multiplied by A^{-1} yields $B = A^{-1}AB = A^{-1}([A, B] + \{A, B\}) = A^{-1}0 = 0$.

Ex. 2.45. $[A, B]^\dagger = (AB - BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = [B^\dagger, A^\dagger]$.

Ex. 2.46. $[A, B] = AB - BA$ and $[B, A] = BA - AB$, so $[A, B] = -[B, A]$.

Ex. 2.47. If $A = A^\dagger$ and $B = B^\dagger$, then $(\mathrm{i}[A, B])^\dagger = -\mathrm{i}[A, B]^\dagger = -\mathrm{i}[B^\dagger, A^\dagger] = \mathrm{i}[A^\dagger, B^\dagger] = \mathrm{i}[A, B]$.

Ex. 2.48. Let P be a positive matrix, then there exist Q, such that $P = Q^\dagger K$. Its polar decomposition is $P = UJ$ where $J = \sqrt{P^\dagger P} = \sqrt{Q^\dagger Q Q^\dagger Q} = Q^\dagger Q = P$. In this case, $U = I$. Similarly, $K = P$.
For a unitary matrix U, $J = \sqrt{U^\dagger U} = I$ and $K = \sqrt{UU^\dagger} = I$.
For a Hermitean matrix, H and $H^\dagger H$ share an eigensystem, so $U = I$ and $K = J = |H|$.

Ex. 2.49. Let N be normal, then $N = \sum_i \lambda_i |i\rangle\langle i|$ (by the spectral decomposition theorem). In this case, $N^\dagger N = \sum_{ij} \lambda_i^* \lambda_j |i\rangle\langle i|\, |j\rangle\langle j| = \sum_i |\lambda_i|^2 |i\rangle\langle i|$ and the resulting operator $J = \sum_i |\lambda_i| |i\rangle\langle i|$. The corresponding unitary operator is $U = \sum_i \lambda_i/|\lambda_i|\, |i\rangle\langle i|$ (and extended to zero eigenvalue subspaces). Due to $UJ = JU$, $K = J$.

Ex. 2.50. Let

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad A^\dagger A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad J = \sqrt{A^\dagger A} = \frac{1}{\sqrt{5}} \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}, \quad K = \sqrt{AA^\dagger} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix},$$

and so

$$U = AJ^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = K^{-1}A\,.$$

2.2 The postulates of quantum mechanics

Ex. 2.51. For the Hadamard gate H holds

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = H^\dagger = H^T\,,$$

so this exercise is equivalent to the next one.

Ex. 2.52. Using H from the previous exercise,

$$H^2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^2 = I\,.$$

Ex. 2.53. The characteristic polynomial is $p(\lambda) = \lambda^2 - 1$, so $\lambda = \pm 1$ with eigenvectors $(1, \sqrt{2} - 1)^T$ and $(-1, \sqrt{2} + 1)^T$.

Ex. 2.54. If A and B are commuting operators, we may diagonalise them on the same basis, $A = \sum_i a_i |i\rangle\langle i|$ and $B = \sum_i b_i |i\rangle\langle i|$. In this basis,

$$\exp A = \sum_i \mathrm{e}^{a_i} |i\rangle\langle i| \,, \quad \exp B = \sum_i \mathrm{e}^{b_i} |i\rangle\langle i| \,,$$

and

$$A + B = \sum_i (a_i + b_i) |i\rangle\langle i| \,, \quad \exp(A+B) = \sum_i \mathrm{e}^{a_i + b_i} |i\rangle\langle i| \,,$$

and

$$\exp A \exp B = \sum_i \mathrm{e}^{a_i} |i\rangle\langle i| \sum_j \mathrm{e}^{b_j} |j\rangle\langle j| = \sum_{ij} \mathrm{e}^{a_i + b_j} |i\rangle\langle i|j\rangle\langle j| = \sum_i \mathrm{e}^{a_i + b_i} |i\rangle\langle i| \,,$$

where we have used $\langle i|j\rangle = \delta_{ij}$.

Ex. 2.55. Let U be defined as in eq. (2.91). Then, using the power series of the exponential and that $H^\dagger = H$,

$$U(t_1, t_2)^\dagger = \exp \frac{\mathrm{i}H(t_2 - t_1)}{\hbar} \,,$$

and the solution of 2.54 to get

$$U(t_1, t_2)^\dagger U(t_1, t_2) = \exp \frac{\mathrm{i}H(t_2 - t_1)}{\hbar} \exp \frac{-\mathrm{i}H(t_2 - t_1)}{\hbar} = \exp 0 = I \,.$$

The product of the other order is done in the same way.

Ex. 2.56. Let U be unitary, then

$$U = \sum_i \mathrm{e}^{\mathrm{i}\kappa_i} |i\rangle\langle i|$$

with κ_i real, in some basis $|i\rangle$. The logarithm is computed as

$$\log U = \sum_i \log \mathrm{e}^{\mathrm{i}\kappa_i} |i\rangle\langle i| = \sum_i \mathrm{i}\kappa_i |i\rangle\langle i| \,,$$

so

$$K = -\mathrm{i} \log U = \sum_i \kappa_i |i\rangle\langle i|$$

is a normal (diagonalisable) operator with real eigenvalues, so it is self-adjoint (Hermitean).

Ex. 2.57. After the measurement described by the set of operators $\{L_\ell\}$, the system is in the state given by eq. (2.93) with a probability $p(\ell)$ given by eq. (2.92). A further measurement, described by the operators $\{M_m\}$ from this state brings the system to state

$$\frac{M_m L_\ell \psi}{\sqrt{\langle\psi| L_\ell^\dagger M_m^\dagger M_m L_\ell^\dagger |\psi\rangle} \sqrt{\langle\psi| L_\ell^\dagger L_\ell |\psi\rangle}} \,,$$

with a probability

$$p(m|\ell) p(\ell) = p_\ell(m) p(\ell) = \frac{\langle\psi| L_\ell^\dagger M_m^\dagger M_m L_\ell |\psi\rangle}{\langle\psi| L_\ell^\dagger L_\ell |\psi\rangle} \langle\psi| L_\ell^\dagger L_\ell |\psi\rangle = \langle\psi| L_\ell^\dagger M_m^\dagger M_m L_\ell |\psi\rangle \,,$$

where $p(m|\ell)$ denotes the conditional probability of the system ending up with the measured value m provided that the first measurement yielded value ℓ.

A measurement given by the operators $\{N_{\ell m}\}$, $N_{\ell m} = M_m L_\ell$ takes the system into state

$$\frac{N_{\ell m}\psi}{\sqrt{\langle\psi|N_{\ell m}^\dagger N_{\ell m}\psi\rangle}} = \frac{N_m L_\ell \psi}{\sqrt{\langle\psi|L_\ell^\dagger M_m^\dagger M_m L_\ell|\psi\rangle}}$$

with probability

$$p(\ell, m) = \langle\psi|N_{\ell m}^\dagger N_{\ell m}|\psi\rangle = \langle\psi|L_\ell^\dagger M_m^\dagger M_m L_\ell|\psi\rangle\,.$$

We shall now consider the consequences of the completeness relation,

$$\sum_{\ell m} N_{\ell m}^\dagger N_{\ell m} = \sum_{\ell m} L_\ell^\dagger M_m^\dagger M_m L_\ell = \sum_\ell L_\ell^\dagger \sum_m M_m^\dagger M_m L_\ell = \sum_\ell L_\ell^\dagger L_\ell = I\,.$$

As we can see, the two are physically equivalent, as they only differ in the normalisation of the output vector.

Ex. 2.58. Let $|m\rangle$ be an eigenvector of M, $M|m\rangle = m|m\rangle$, then the expectation value is

$$\mathbf{E}(M) = \langle M\rangle = \langle m|M|m\rangle = \langle m|m|m\rangle = m\langle m|m\rangle = m\,,$$

if $|m\rangle$ is normalised. The standard deviation is $\Delta^2(M) = \langle M^2\rangle - \langle M\rangle^2$, and the first one is

$$\langle M^2\rangle = \langle m|M^2|m\rangle = \langle m|m\cdot m|m\rangle = m^2\,,$$

where we have used the for the adjoint $\langle m|M = \langle m|M^\dagger = (M|m\rangle)^\dagger = (m|m\rangle)^\dagger = m\langle m|$ as for a self-adjoint observable m is real. This shows that $\Delta(M) = 0$ in an eigenstate.

Ex. 2.59. The expectation value is

$$\langle X\rangle = \langle 0|X|0\rangle = (1,0)\begin{pmatrix} & -\mathrm{i}\\ \mathrm{i} & \end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = 0\,,$$

and as $X^2 = I$, the standard deviation $\Delta^2(X) = \langle X^2\rangle - \langle X\rangle^2 = \langle X^2\rangle = 1$.

Ex. 2.60. Assuming $v^2 = 1$, the matrix $\vec{v}\cdot\vec{\sigma}$ satisfies $(\vec{v}\cdot\vec{\sigma})^2 = I$ using $\sigma_i\sigma_j = \delta_{ij}I + \mathrm{i}\epsilon_{ijk}\sigma_k$. Therefore, its eigenvalues must satisfy $\lambda^2 = 1$. As the Pauli matrices do not commute, $\vec{v}\cdot\vec{\sigma}$ cannot be $\pm I$, so its eigenvalues must be $\lambda_1 = 1$ and $\lambda_2 = -1$. The fact that the operators $P_\pm = (I \pm \vec{v}\cdot\vec{\sigma})/2$ act as the projectors on the eigenstates is easily verified in this basis.

Ex. 2.61. The probability of measuring $+1$ is the same as the expectation value of the projector, i.e.,

$$p(1) = \frac{1}{2}\langle 0|I + \vec{v}\cdot\vec{\sigma}|0\rangle = \frac{1}{2} + \langle 0|\vec{v}\cdot\vec{\sigma}|0\rangle = \frac{1+v_3}{2}\,,$$

as the 1, 1 component of all Pauli matrices except the third one vanishes. If the result 1 is obtained, the new state after the measurement is

$$\frac{P_+|0\rangle}{\sqrt{\langle 0|P_+|0\rangle}} = \frac{(I+\vec{v}\cdot\vec{\sigma})|0\rangle}{\sqrt{2(1+v_3)}} = \frac{1}{\sqrt{2}\sqrt{1+v_3}}\begin{pmatrix}1+v_3\\ \mathrm{i}v_1 + v_2\end{pmatrix}\,.$$

Ex. 2.62. If for a measurement $\{M_m\}$ the measurement operators and the POVM elements $E_m = M_m^\dagger M_m$ agree,

$$E_m = M_m^\dagger M_m = M_m \,,$$

it follows that the measurement operators are self-adjoint, $M_m^\dagger = E_m^\dagger = (M_m^\dagger M_m)^\dagger = M_m^\dagger M_m = E_m = M_m$, and so $M_m^2 = M_m^\dagger M_m = E_m = M_m$, they are projectors as well, and the completeness relation demands that $\sum_m M_m = I$,

$$I = I^2 = \left(\sum_m E_m\right)^2 = \sum_m E_m^2 + \sum_{m\neq n} E_m E_n = \sum_m E_m + \sum_{m\neq n} E_m E_n = I + \sum_{m\neq n} E_m E_n \,,$$

and all $M_m = E_m$ are positive operators, therefore so is $E_m E_n$, from which $E_m E_n = 0$ $(m \neq n)$ follows, i.e., they are mutually exclusive projectors, and so correspond to a projective operator.

Ex. 2.63. Let $\{M_m\}$ be a set of measurement operators. For each of these, $E_m = M_m^\dagger M_m$ is a positive self-adjoint operator, and therefore, its square root exists, and the polar decomposition of M_m is $M_m = U_m\sqrt{E_m}$. The completeness relation reads $\sum_m E_m = I$. Thus, E_m is a set of positive operators with the completeness property, i.e., a POVM.

Ex. 2.64. As the vectors ψ_k $(k = 1, \ldots, m)$ are linearly independent, for each vector, it is possible to find a normalised vector ϕ_k which is orthogonal to ψ_j $(j \neq k)$, and not orthogonal to ψ_k, and take

$$E_k = |\phi_k\rangle\langle\phi_k| \quad (k = 1, \ldots, m)\,, \text{ and } \quad E_{m+1} = 1 - \sum_{k=1}^{m} E_k \,.$$

It is simple to verify that E_k is positive, $\langle\psi_k|E_k|\psi_k\rangle \neq 0$, $\langle\psi_j|E_k|\psi_j\rangle = 0$ $(k \neq j)$ and $\sum_{k=1}^{m+1} E_k = I$, i.e., it is the POVM required.

To find this set of vectors, for each vector, move that one to the last position, apply the Gram-Schmidt procedure, and keep only the last vector. Equivalently, as the vectors are linearly independent, they span a subspace of the state space, and for each k, $\{\psi_j | j \neq k\}$ spans a different subspace. Let P_k denote the orthogonal projector to this subspace, $\phi_k' = \psi_k - P_k\psi_k$ and $\phi_k = \phi_k/\|\phi_k\|$.

Ex. 2.65. In the basis

$$|0'\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}\,, \quad |1'\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

the two vectors are $(1,0)^T$ and $(0,1)^T$, so not the same up to relative phases.

Ex. 2.66. Let us first use the action of the operators,

$$X_1 Z_2 \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{X_1|0\rangle Z_2|0\rangle + X_1|1\rangle Z_2|1\rangle}{\sqrt{2}} = \frac{|1\rangle|0\rangle + |0\rangle|1\rangle}{\sqrt{2}} = \frac{|10\rangle + |01\rangle}{\sqrt{2}}\,,$$

and the following scalar products,

$$\begin{aligned}
\langle 00|10\rangle &= \langle 0|1\rangle\langle 0|0\rangle = 0 \times 1 = 0\,, \quad \langle 11|10\rangle = \langle 1|1\rangle\langle 1|0\rangle = 1 \times 0 = 0\,,\\
\langle 00|01\rangle &= \langle 0|0\rangle\langle 0|1\rangle = 1 \times 0 = 0\,, \quad \rangle 11|01\rangle = \langle 1|0\rangle\langle 1|1\rangle = 0 \times 1 = 0\,,
\end{aligned}$$

which yields

$$\frac{(\langle 00| + \langle 11|)X_1 Z_2(|00\rangle + |11\rangle)}{2} = 0\,.$$

Ex. 2.67. Let us consider an orthonormal basis e_i in V such that $e_1, \ldots, e_{\dim W}$ is in W, and the rest is in its orthogonal complement. As $f_i = Ue_i$, $i = 1, \ldots, \dim W$ are also orthonormal, it is possible to extend these to an orthonormal basis f_i too, and define $Ue_i = f_i$ for $i > \dim W$. The operator such defined is defined on the whole space V and preserves scalar products, therefore it is unitary.

Ex. 2.68. Any state of a composite system can be represented as a matrix of its coefficient in the basis $|i,j\rangle$, i.e., $|\psi\rangle = \psi_{i,j}|i,j\rangle$. For the state $\psi = (|00\rangle + |11\rangle)/\sqrt{2}$, we have $\psi_{0,0} = \psi_{1,1} = 1/\sqrt{2}$. This is a rank-2 matrix. For any state $|a\rangle|b\rangle$, the resulting matrix is a rank-1 matrix, which can be shown by, e.g., choosing a basis in both spaces in which $|a\rangle$ and $|b\rangle$ are the first basis vectors, yielding $(|a\rangle|b\rangle)_{0,0} = 1$ and all other elements 0. As the rank of a matrix is invariant to basis transformations, this proves that ψ is not a product state.

Ex. 2.69. It is sufficient to show linear independence, as the space is four dimensional and there are four vectors. This follows from orthonormality, and that is shown as follows,

$$\begin{aligned}
\langle 00'|00'\rangle &= \frac{1}{2}(\langle 00|00\rangle + \langle 00|11\rangle + \langle 11|00\rangle + \langle 11|11\rangle) = \frac{1}{2}(1+0+0+1) = 1\,,\\
\langle 00'|01'\rangle &= \frac{1}{2}(\langle 00|00\rangle + \langle 11|00\rangle - \langle 00|11\rangle - \langle 11|11\rangle = \frac{1}{2}(1+0-0-1) = 0\,,\\
\langle 00'|10'\rangle &= \frac{1}{2}(\langle 00|10\rangle + \langle 00|01\rangle + \langle 11|10\rangle + \langle 11|01\rangle) = \frac{1}{2}(0+0+0+0) = 0\,,\\
\langle 00'|11'\rangle &= \frac{1}{2}(\langle 00|01\rangle - \langle 00|10\rangle + \langle 11|10\rangle - \langle 11|01\rangle) = \frac{1}{2}(0-0+0-0) = 0\,,\\
\langle 01'|01'\rangle &= \frac{1}{2}(\langle 00|00\rangle - \langle 00|11\rangle - \langle 11|00\rangle + \langle 11|11\rangle) = \frac{1}{2}(1-0-0+1) = 1\,,\\
\langle 01'|10'\rangle &= \frac{1}{2}(\langle 00|10\rangle + \langle 00|01\rangle - \langle 11|10\rangle - \langle 11|01\rangle) = \frac{1}{2}(0+0-0-0) = 0\,,\\
\langle 01'|11'\rangle &= \frac{1}{2}(\langle 00|01\rangle - \langle 00|10\rangle - \langle 11|01\rangle + \langle 11|01\rangle) = \frac{1}{2}(0-0-0+0) = 0\,,\\
\langle 10'|10'\rangle &= \frac{1}{2}(\langle 10|10\rangle + \langle 10|01\rangle + \langle 01|10\rangle + \langle 10|10\rangle) = \frac{1}{2}(1+0+0+1) = 1\,,\\
\langle 10'|11'\rangle &= \frac{1}{2}(\langle 01|01\rangle - \langle 01|10\rangle + \langle 10|01\rangle - \langle 10|10\rangle) = \frac{1}{2}(1-0+0-1) = 0\,,\\
\langle 11'|11'\rangle &= \frac{1}{2}(\langle 01|01\rangle - \langle 01|10\rangle - \langle 10|01\rangle + \langle 10|10\rangle) = \frac{1}{2}(1-0-0+1) = 1\,,
\end{aligned}$$

where we used the notation $|i,j'\rangle$ to denote the Bell basis elements corresponding to the two classical bits i, j as shown in eqs. (2.134-2.137).

Ex. 2.70. Bell states each have the form (see the notation in Soln. 2.69)

$$\psi = \frac{|ij\rangle \pm |\bar{i}\bar{j}\rangle}{\sqrt{2}}\,,$$

where we have used an overbar to denote the negation of the corresponding bit. As a result, evaluating an operator of the form $E \otimes I$ in such a state yields

$$(E \otimes I)\psi = \frac{1}{\sqrt{2}}(E_{ii}|ij\rangle + E_{\bar{i}i}|\bar{i}j\rangle \pm E_{i\bar{i}}|i\bar{j}\rangle \pm E_{\bar{i}\bar{i}}|\bar{i}\bar{j}\rangle)\,,$$

and when multiplying this vector with ψ, one obtains

$$\langle\psi|E\otimes I|\psi\rangle = \frac{1}{2}(\langle ij| \pm \langle \bar{i}\bar{j}|)(E_{ii}|ij\rangle + E_{\bar{i}i}|\bar{i}j\rangle \pm E_{i\bar{i}}|i\bar{j}\rangle \pm E_{\bar{i}\bar{i}}|\bar{i}\bar{j}\rangle) = E_{ii} + E_{\bar{i}\bar{i}}\,,$$

independently of i and j and the sign.

If Eve intercepts Alice's qubit, she can only perform local measurements, i.e., with operators of the form $\{M_m \otimes I\}$, and so the probabilities of the outcomes are

$$p(m) = \langle\psi|M_m^\dagger M_m \otimes I|\psi\rangle\,,$$

independent of which Bell state the system was in, therefore, she cannot distinguish the states.

Ex. 2.71. Let $\rho = \sum_j p_j|j\rangle\langle j|$ be the spectral decomposition of the density operator. As ρ is positive and $\operatorname{Tr}\rho = \sum_j p_j = 1$, $0 \le p_j \le 1$ for all j, so $p_j^2 \le p_j$, and either there is one $j = j_0$ for which $p_j = 1$ and all of them are 0, or $p_j < 1$ for all j. In the latter case $p_j^2 < p_j$. Then

$$\operatorname{Tr}\rho^2 = \operatorname{Tr}\sum_{j,k} p_j p_j |j\rangle\langle j||k\rangle\langle k| = \operatorname{Tr}\sum_j p_j^2|j\rangle\langle j| = \sum_j p_j^2 \le \sum_j p_j = 1\,,$$

and equality only holds when there is a $p_{j_0} = 1$, i.e., when the state is pure, $\rho = p_{j_0}|j_0\rangle\langle j_0|$.

Ex. 2.72. (1) Let ρ be a density matrix, i.e., it is a positive operator with $\operatorname{Tr}\rho = 1$. Then let $T = \rho - I/2$, so $\operatorname{Tr}T = \operatorname{Tr}\rho - \operatorname{Tr}I/2 = 0$. A basis on the space of traceless Hermitian 2×2 matrices is the set of the Pauli matrices, so it may be so expanded, $T = (1/2)\vec{r}\cdot\vec{\sigma}$, yielding

$$\rho = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}\,.$$

The eigenvalues of $\vec{r}\cdot\vec{\sigma}$ are $\pm|r|$ (as seen, e.g., by writing the characteristic polynomial of the 2×2 matrix A as $\lambda^2 - \lambda\operatorname{Tr}A + \det A$), so $|r| \le 1$ must hold in order that ρ is positive.

(2) To the state $\rho = I/2$ corresponds $\vec{r} = 0$, i.e., it is represented by the centre of the Bloch sphere.

(3) Let ρ correspond to a pure state, in which case $\operatorname{Tr}\rho^2 = 1$, and as

$$\left(\frac{I + \vec{r}\cdot\vec{\sigma}}{2}\right)^2 = \frac{I + 2\vec{r}\cdot\vec{\sigma} + (\vec{r}\cdot\vec{\sigma})^2}{4} = \frac{(1+r^2)I + 2\vec{r}\cdot\vec{\sigma}}{4}\,,$$

$\operatorname{Tr}\rho^2 = (1+r^2)/2$ which is unity iff $r = |\vec{r}| = 1$.

(4) In sec. 1.2, the pure states were parametrised in eq. (1.3). Let us extract the vector $\vec{v}$ by using the fact that $\operatorname{Tr}\sigma_i\sigma_j = 2\delta_{ij}$, so $v_i = \operatorname{Tr}(\sigma_i\rho)$, yielding

$$\begin{aligned} v_1 &= \operatorname{Tr}\sigma_1\rho = \langle\psi|\sigma_1|\psi\rangle = \sin\vartheta\cos\varphi\,,\\ v_2 &= \operatorname{Tr}\sigma_2\rho = \langle\psi|\sigma_2|\psi\rangle = \sin\vartheta\sin\varphi\,,\\ v_3 &= \operatorname{Tr}\sigma_3\rho = \langle\psi|\sigma_3|\psi\rangle = \cos\vartheta\,, \end{aligned}$$

which shows that the two descriptions of the Bloch vector agree.

Ex. 2.73. Let $\rho = \sum_i \lambda_i|i\rangle\langle i|$ and $\psi = \sum_i c_i|i\rangle$. In this case, the generalised inverse is

$$\rho^{-1} = \sum \lambda_i^{-1}|i\rangle\langle i|\,,$$

and so

$$\langle\psi|\rho^{-1}|\psi\rangle = \sum_i \langle\psi|i\rangle\lambda_i^{-1}\langle i|\psi\rangle = \sum_i |c_i|^2/\lambda_i\,.$$

According to theorem 2.6, a set $\{p_j, \psi_j\}$ is an ensemble for ρ iff there is a unitary matrix u with which

$$\sqrt{p_i}\psi_i = \sum_j u_{ij}\sqrt{\lambda_j}|j\rangle\,.$$

Let $\psi_1 = \psi$. In this case,

$$\sum_j \sqrt{p_1}c_j|j\rangle = \sqrt{p_1}\psi = \sum_j u_{1j}\sqrt{\lambda_j}|j\rangle\,,$$

yielding

$$u_{1j} = \frac{\sqrt{p_1}c_j}{\sqrt{\lambda_j}}\,,$$

and from the unitarity of u follows $uu^\dagger = I$, so

$$\frac{1}{p} = \sum_j \frac{|c_j|^2}{\lambda_j}\,, \quad p = \frac{1}{\langle\psi|\rho^{-1}|\psi\rangle}\,.$$

Constructing a minimal ensemble with $\psi_1 = \psi$ can be done as follows: first, we consider a matrix whose first row is $u_{1j} = c_j\sqrt{\lambda_j}/\sqrt{p_1}$, and extend this into a unitary $\operatorname{Rank}\rho \times \operatorname{Rank}\rho$ matrix (that this is possible can be shown using the vectors e_i and the Gram-Schmidt orthogonalisation procedure), and let the remaining vectors in the set be the vectors given by eq. (2.167).

Ex. 2.74. If the composite system is in a state $|a\rangle|b\rangle$, its density operator is $\rho = |a\rangle|b\rangle\langle a|\langle b$, and the reduced density operator is, using eq. (2.178),

$$\rho^A = \operatorname{Tr}_B\rho = |a\rangle\langle a|\operatorname{Tr}|b\rangle\langle b| = |a\rangle\langle a|\langle b|b\rangle = |a\rangle\langle a|\,,$$

which is a pure state, given by the state vector $|a\rangle$.

Ex. 2.75. For the Bell state $|00'\rangle = (|00\rangle + |11\rangle)/\sqrt{(2)}$ (see the notation in Soln. 2.69), the density operator is

$$\rho = |00'\rangle\langle 00'| = \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|) = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|)\,,$$

and so

$$\rho^1 = \operatorname{Tr}_2\rho = \frac{1}{2}(|0\rangle\langle 0| + 0 + 0 + |1\rangle\langle 1|) = \frac{I}{2}\,,$$

and, similarly, $\rho^2 = I/2$. For $|01'\rangle$,

$$\rho = |01'\rangle\langle 01'| = \frac{1}{2}(|00\rangle\langle 00| - |00\rangle\langle 11| - |11\rangle\langle 00| + |11\rangle\langle 11|)\,,$$

therefore $\rho^1 = I/2 = \rho^2$. For $|10'\rangle$,

$$\rho = |10'\rangle\langle 10'| = \frac{1}{2}(|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|)\,,$$

so

$$\rho_1 = \mathrm{Tr}_2 \rho = \frac{1}{2}(|0\rangle\langle 0| + 0 + 0 + |1\rangle\langle 1|) = \frac{I}{2}\,,$$

and

$$\rho_2 = \mathrm{Tr}_1 \rho = \frac{1}{2}(|1\rangle\langle 1| + 0 + 0 + |0\rangle\langle 0|) = \frac{I}{2}\,,$$

and we obtain the same result for $|11'\rangle$ by flipping the signs of terms 2 and 3 in the brackets.

Note, that the independence of the reduced density matrices on the Bell states also follows from Exercise 2.70, as the measurement statistics of all operators on one of the subsystems must be the same for all the Bell states. For the $|00\rangle$ state the result is given by eq. (2.191).

Ex. 2.76. Let us assume that the dimensions of the state spaces of the two subsystems are $\dim H_A = n$, $\dim H_B = m$, $n < m$. Let us consider a larger state space H'_A with $\dim H'_A = m$, and apply the theorem there.

The vectors $|j\rangle|k\rangle$ for $j > n$ do not appear in the expression of ψ, $a_{jk} = 0$ for $j > n$. As a result, $u_{ji} d_{ii} = 0$, and therefore in $\psi = \sum_i \lambda_i |i_A\rangle|i_B\rangle$, one can drop the zero eigenvalues, and for the non-zero eigenvalues have $|i_A\rangle = \sum_{j=1} n u_{ji} |j\rangle$, the sum run only over $1, \ldots, n$, i.e., the vector is in the subspace corresponding to H_A in H'_A.

Ex. 2.77. The question of a triple decomposition has been considered in Ref. [1]. The main line of the argument is that if the eigenvalues are unequal, the decomposition is unique. On the one hand, the tripartite state space can be split into subsystem 1 and subsystems 2 and 3, one may apply the Schmidt decomposition, and find a decomposition of the form

$$\psi = \sum_i \lambda_i |i_A\rangle|i_{BC}\rangle\,,$$

and on the other hand, if the tripartite decomposition is possible,

$$\psi = \sum_i \lambda_i |i_A\rangle|i_B\rangle|i_C\rangle\,,$$

and the two must agree, i.e., $|i_{BC}\rangle$ must agree with $|i_B\rangle|i_C\rangle$. We may therefore construct the counterexample "backwards", by starting with a 1-(23) decomposition in which the vectors $|_{BC}\rangle$ are entangled states, e.g.,

$$\psi = \alpha|0\rangle|01'\rangle + \beta|1\rangle|10'\rangle\,,$$

using Bell states (see Soln. 2.69 for the notation), and $\alpha \neq \beta$, $|\alpha|^2 + |\beta|^2 = 1$.

Ex. 2.78. Let $\psi = |\psi_A\rangle|\psi_B\rangle$. This is already a Schmidt decomposition with Schmidt number 1. In the other direction, a state with Schmidt number one is $\psi = \lambda_1 |1_A\rangle|1_B\rangle$ is also clearly a product state.

Also, if ψ is a product state, clearly ψ_A and ψ_B are pure states. For the other direction, let us assume that ψ has at least Schmidt number 2, i.e., in the sum

$$\psi = \sum_i \lambda_i |i_A\rangle|i_B\rangle$$

there are two or more nonzero λ_i's. In this case, the reduced states are

$$\rho_A = \mathrm{Tr}_B \sum_{i,j} \lambda_i \lambda_j |i_A\rangle|i_B\rangle\langle j_A|\langle j_B| = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|\,,$$

$$\rho_B = \mathrm{Tr}_A \sum_{i,j} \lambda_i \lambda_j |i_A\rangle|i_B\rangle\langle j_A|\langle j_B| = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|\,,$$

as $\mathrm{Tr}\,|i_B\rangle\langle j_B| = \langle j_B|i_B\rangle = \delta_{ij}$ and $\mathrm{Tr}\,|i_A\rangle\langle j_A| = \langle j_A|i_A\rangle = \delta_{ij}$. In order that the two ρ_A, ρ_B are pure states, both ρ_A and ρ_B must be rank-1 matrices, there can neither be two different $|i_A\rangle$ nor $|i_B\rangle$ vectors.

Ex. 2.79. The Schmidt decomposition is found by extracting from the states the matrix a, and then performing an SVD on these matrices. The matrices are as follows,

$$\psi_1 = \frac{|00\rangle + |11\rangle}{\sqrt{2}}\,, \quad a_1 = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$$

which is already in the Schmidt decompositon form (a is diagonal),

$$\psi_2 = \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2}\,, \quad a_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\,,$$
$$\psi_3 = \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}\,, \quad a_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\,.$$

The SVD of the matrices yields

$$a_2 = U_2 D_2 V_2\,, D_2 = \begin{pmatrix} 0 & \\ & 2 \end{pmatrix} U_2 = V_2^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\,,$$

yielding

$$|0_A\rangle = \sum_j (U_2)_{j0}|j\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,, \quad |1_A\rangle = \sum_j (U_2)_{j1}|j\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,,$$

and

$$|0_B\rangle = \sum_j (V_2)_{j0}|j\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,, \quad |1_B\rangle = \sum_j (V_2)_{j1}|j\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\,,$$

and similarly (by finding left and right eigenvalues)

$$a_3 = U_3 D_3 V_3\,, \quad D_3 = \begin{pmatrix} \frac{\sqrt{5}+1}{2} & \\ & \frac{-\sqrt{5}+1}{2} \end{pmatrix}\,,$$

and

$$U_3 = V_3^\dagger = \begin{pmatrix} \frac{2}{\sqrt{\frac{\sqrt{5}-5}{\sqrt{5}-3}(\sqrt{5}-1)}} & \frac{-\sqrt{2}}{\sqrt{\sqrt{5}+5}} \\ \sqrt{\frac{\sqrt{5}-3}{\sqrt{5}-5}} & \frac{\sqrt{5}+1}{\sqrt{2\sqrt{5}+10}}\,, \end{pmatrix}$$

and $|i_{A,B}\rangle$ $(i = 0, 1)$ as above.

Ex. 2.80. Te equations below eq. (2.204) defining the new bases on the two subsystem Hilbert spaces, $|i_A\rangle = \sum_j u_{ji}|j\rangle$ and $|i_B\rangle = \sum_k v_{ik}|k\rangle$ may be used to define unitary operators by

$$|i_A\rangle = U(\psi)|i\rangle\,, \quad |i_B\rangle = V(\psi)|i\rangle\,.$$

Let us apply the inverse to ψ,

$$\begin{aligned}(U(\psi)\otimes V(\psi))^\dagger|\psi\rangle &= \sum_i \lambda_i |i_A\rangle|i_B\rangle = \sum_i \lambda_i U(\psi)^\dagger|i_A\rangle \otimes V(\psi)^\dagger|i_B\rangle \\ &= \sum_i \lambda_i U(\psi)^\dagger U(\psi)|i\rangle \otimes V(\psi)^\dagger V(\psi)|i\rangle = \sum_j \lambda_j |i\rangle|i\rangle\,,\end{aligned}$$

which only depends on the Schmidt coefficients of the state ψ. This lets us define

$$U = U(\psi)U(\varphi)^\dagger\,, \quad V = V(\psi)V(\varphi)^\dagger\,,$$

so that

$$\begin{aligned}(U\otimes V)\varphi &= (U(\psi)\otimes V(\psi))(U(\varphi)^\dagger\otimes V(\varphi)^\dagger)\varphi\\ &= (U(\psi)\otimes V(\psi))\sum_j \lambda_j|i\rangle|i\rangle = \psi\,.\end{aligned}$$

To verify that $U(\psi)$ and $V(\psi)$ are both unitary, note that they take orthonormal bases into orthonormal bases.

Ex. 2.81. Let us use that basis in the subsystem A which diagonalises the density matrix ρ, i.e.,

$$\rho = \sum_j p_j|j^A\rangle\langle j^A|\,,$$

and write the purifications as

$$\psi = \sum_{jk}\psi_{j,k}|j^A\rangle|k^R\rangle\,, \quad \varphi = \sum_{jk}\varphi_{j,k}|j^A\rangle|k^R\rangle\,.$$

The condition for a purification is that

$$\rho = \mathrm{Tr}_2|\psi\rangle\langle\psi| = \mathrm{Tr}_2\sum_{j,k,\ell,m}\psi_{j,k}\psi^*_{\ell,m}|j^A\rangle|k^R\rangle\langle\ell^A|\langle m^R| = \sum_{j,k,\ell}\psi_{j,k}\psi^*_{\ell,k}|j^A\rangle\langle\ell^A|\,,$$

and similary for φ, yielding

$$\sum_k \psi_{j,k}\psi^*_{\ell,k} = p_j\delta_{j,\ell} = \sum_k \varphi_{j,k}\varphi_{\ell,k}\,.$$

Defining $u_{j,k} = p_j^{-1/2}\psi_{j,k}$, this yields $\sum_k u_{j,k}u^*_{\ell,k} = \delta_{j,\ell}$, i.e., u is a unitary matrix, and so is v defined as $v_{j,k} = p_j^{-1/2}\varphi_{j,k}$ and we may write

$$\begin{aligned}\psi &= \sum_{j,k}\sqrt{p_j}|j^A\rangle u_{j,k}|k^R\rangle = \sum_j\sqrt{p_j}|j^A\rangle U|j_{R_1}\rangle = (I\otimes U)\sum_j\sqrt{p_j}|j^A\rangle|j_{R_1}\rangle\,,\\ \varphi &= \sum_{j,k}\sqrt{p_j}|j^A\rangle v_{j,k}|k^R\rangle = \sum_{j,k}\sqrt{p_j}|j^A\rangle U|j_{R_2}\rangle = (I\otimes V)\sum_j\sqrt{p_j}|j^A\rangle|j_{R_2}\rangle\,,\end{aligned}$$

The operator defined by $W|j_{R_2}\rangle = |j_{R_1}$ is also unitary, as it maps the elements of an orthonormal basis into those of another, so

$$U_R = UWV^\dagger$$

has the desired property, $(I\otimes U_R)\varphi = \psi$.

Ex. 2.82. (1) The proof is basically eqs. (2.207–2.2.211).

(2) Measuring $|i\rangle$ in the R-basis corresponds to a projective measurement with measurement operators $I\otimes P_i$ with $P_i = |i\rangle\langle i|$, yielding probabilities

$$\begin{aligned}p(i) &= \sum_{k,\ell}\sqrt{p_kp_\ell}\langle k|\langle\psi_k|(I\otimes P_i)|\psi_\ell\rangle|\ell\rangle = \sum_{k,\ell}\sqrt{p_kp_\ell}\langle k|\langle\psi_k|(I\otimes|i\rangle\langle i|)|\psi_\ell\rangle|\ell\rangle\\ &= \sum_{k,\ell}\sqrt{p_kp_\ell}\delta_{k,\ell}\delta_{i,\ell} = p_i\,.\end{aligned}$$

The corresponding state of system A is

$$\frac{P_i \sum_k \sqrt{p_k}|\psi_k\rangle|k\rangle}{\sqrt{\sum_{k,l}\sqrt{p_k p_\ell}\langle k|\langle\psi_k|P_i|\psi_\ell\rangle|\ell\rangle}} = \frac{(I\otimes|i\rangle\langle i|)\sum_k\sqrt{p_k}|\psi_k\rangle|k\rangle}{\sqrt{\sum_{k,l}\sqrt{p_k p_\ell}\langle k|\langle\psi_k|(I\otimes|i\rangle\langle i|)|\psi_\ell\rangle|\ell\rangle}}$$
$$= \frac{\sum_k\sqrt{p_k}|\psi_k\rangle\delta_{ik}}{\sqrt{\sum_{k,\ell}\sqrt{p_k p_\ell}\delta_{k,\ell}\delta_{k,i}\delta_{i,\ell}}} = \psi_k\,.$$

(3) If $|AR\rangle$ is any purification, we know that it is unitary equivalent, with an operator $I\otimes U_R$, to the purification condidered in (1), (2) (see Excer. 2.81). The measurement operators are then

$$P_i' = (I\otimes U_R)P_i(I\otimes U_R^\dagger).$$

These are also rank-1 projectors, i.e., $P_i' = |i'\rangle\langle i|$ and the sought states are $|i'\rangle$.

Pr. 2.1. Let us expand the function f in a Taylor series,

$$f(x) = \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)x^k\,,$$

and note that $\vec{n}\cdot\vec{\sigma}$ has the property that $(\vec{n}\cdot\vec{\sigma})^2 = (n_i\sigma_i)^2 = n_i n_k\sigma_i\sigma_k = n_i n_k(\delta_{ik}+\mathrm{i}\epsilon_{ik\ell}\sigma_\ell) = 1$, so

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \sum_k\frac{1}{(2k)!}f^{(2k)}(0)\theta^{2k} + \theta\vec{n}\cdot\vec{\sigma}\sum_k\frac{1}{(2k+1)!}f^{(2k+1)}(0)\theta^{2k+1}\,,$$

and, similarly, $(-x)^{2k} = x^{2k}$, $(-x)^{2k+1} = -(-x)^{2k+1}$, so

$$\frac{f(\theta)+f(-\theta)}{2} = \sum_k\frac{1}{(2k)!}f^{(2k)}(0)\theta^{2k}\,,$$

and

$$\frac{f(\theta)-f(-\theta)}{2} = \sum_k\frac{1}{(2k+1)!}f^{(2k+1)}(0)\theta^{2k+1}\,,$$

therefore, comparison yields the desired result

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma}\,.$$

Pr. 2.2. (1) Let us work in the basis where ψ is transformed to its Schmidt form,

$$\psi = \sum_i\lambda_i|i_A\rangle|i_B\rangle\,,$$

and compute the partial trace

$$\rho_A = \mathrm{Tr}_2|\psi\rangle\langle\psi| = \sum_{i,j}\lambda_i\lambda_j|i_A\rangle\langle i_A|\langle i_B|j_B\rangle = \sum_i\lambda_i^2|i_A\rangle\langle i_A|\,,$$

where we used $\langle i_B|j_B\rangle = \delta_{ij}$. It is clear that the rank of ρ_A is the Schmidt number of ψ.

(2) By applying Gram-Schmidt orthogonalisation to the vectors α_j and to β_j we obtain an orthonormal basis in the space spanned by them. The resulting expansion is of the form

$$\psi = \sum_{i,j} a_{ij}|i\rangle|j\rangle\,,$$

and the Schmidt-decomposition is obtaine from the singular value decomposition of the matrix a, and the number of non-zero singular values is maximally that of the smaller of the number of rows and columns of a, both of which are the number of linearly independent vectors among α_i and β_i, repsectively.

(3) Again, writing all the state vectors in the form

$$\psi = \sum_{i,j} \psi_{i,j}|i\rangle|j\rangle\,, \quad \varphi = \sum_{i,j} \varphi_{i,j}|i\rangle|j\rangle\,, \quad \gamma = \sum_{i,j} \gamma_{i,j}|i\rangle|j\rangle\,,$$

the Schmidt number becomes the number of the non-zero singular values of the corresponding matrices ψ_{ij}, $\varphi_{i,j}$ and $\gamma_{i,j}$. The rank is the dimension of the space spanned by the columns of the matrices, it is clear that the rank of the matrix γ is $\mathrm{Rank}\,\gamma = \mathrm{Rank}\,\beta\gamma = \mathrm{Rank}(\psi - \alpha\varphi) \leq \mathrm{Rank}\,\psi + \mathrm{Rank}\,\varphi$, yielding $\mathrm{Sch}(\gamma) \leq \mathrm{Sch}(\psi) + \mathrm{Sch}\varphi$, or $\mathrm{Sch}(\psi) \geq \mathrm{Sch}(\gamma) - \mathrm{Sch}(\varphi)$. By exchanging the roles of φ and γ we obtain the desired inequality,

$$\mathrm{Sch}(\psi) \geq |\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)|\,.$$

Pr. 2.3. To calculate

$$(Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2\,,$$

note that any of these operators square to I, as for any $\vec{v}$, $(\vec{v}\cdot\vec{\sigma})^2 = v^2 I$, so the $4I$ equals the diagonal terms. To evaluate the cross term, note that

$$(A \otimes B)(C \otimes D) = AC \otimes BD\,,$$

and if the letter on one side of the tensor product agree, there an identity operator resutlts. These terms are, eq., $Q \otimes SR \otimes S = QR \otimes I$, and these drop out. The remaining terms are

$$4I + QR \otimes ST - RQ \otimes ST + RQ \otimes TS - QR \otimes TS\,,$$

which are the terms arising when expanding $4I + [Q,R] \otimes [S,T]$, thereby prooving eq. (2.33).

For any operator A, $\langle A\rangle \leq \lambda_{\max}(A)$. Applying this to the operator in the brackets in the expression we started with yields

$$\langle (Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2\rangle = \langle 4I + [Q,R] \otimes [S,T]\rangle \leq 8\,,$$

where we have used, e.g., that

$$[Q,R] = [\vec{q}\cdot\vec{\sigma}, \vec{r}\cdot\vec{\sigma}] = 2\mathrm{i}(\vec{q}\times\vec{r})\cdot\vec{\sigma}\,,$$

where $\vec{q}\times\vec{r}$ is the vector product of two unit vectors, therefore its length is at most 1, so the maximal eigenvalue of $[Q,R]$ is also at most 1.

For any operator A, $\sigma_A^2 = \langle A^2\rangle - \langle A\rangle^2$, yielding

$$\langle A\rangle \leq |\langle A\rangle|\,, \quad \langle A\rangle^2 = \langle A^2\rangle - \sigma_A^2 \leq \langle A^2\rangle\,,$$

Applying this inequality to the operator $Q \otimes S + R \otimes S + R \otimes T - Q \otimes T$, and plugging in our result for its square yields Tsirelson's inequality.

3 Introduction to computer science

Ex. 3.1. I do not think it is possible to verify if a natural process evaluates a function non-computable with a Turing machine. We can only verify what it evaluates for a finite value of inputs, and there definitely exists a Turing machine which gives the same results for thos (e.g., with a lot of states containing a lookup table). We might conjecture that a process evaluates a non-computable function if we try to model it with Turing machines and then every time we test it on a new input the model has to be changed.

Ex. 3.2. Turing machines can be numbered as follows: we put them into a table, where in the ith column are Turing machines wiht an alphabet of length i and in the jth row those of j states, and in each cell of the table, behind one another those with their programs alphabetically sorted.

Inputs can also be assigned a number. If in the kth cell of the input tape there is the a_kth element of the alphabet for a machine, we assign that input the number $p_1^{a_1} p_2^{a_2} \cdot \dots \cdot p_n^{a_n}$ where n is the length of the given input and p_j is the jth prime number. This gives a natural number, and as the prime factorisation is unique it is a 1:1 mapping of possible inputs to real numbers.

Ex. 3.3. Let us construct the Turing machine as follows: we take a two-tape Turing machine, with states r, ℓ for moving left, ℓ' for rewinding tape 2 and c for copying, and the following program:

$$\begin{aligned}
&\langle q_s, \triangleright, \triangleright, r, \triangleright, x, +1, +1\rangle\,, \\
&\langle r, x, y, r, x, y, +1, 0\rangle\,, \quad \forall x \neq b, y \\
&\langle r, b, y, \ell, b, y, -1, 0\rangle\,, \quad \forall y \\
&\langle \ell, x, y, \ell, b, x, -1, 1\rangle\,, \quad \forall x \neq \triangleright, y \\
&\langle \ell, \triangleright, \triangleright, \ell', \triangleright, \triangleright, 1, 1\rangle \\
&\langle \ell', x, y, \ell', x, y, 0, -1\rangle\,, \quad \forall x, y \neq \triangleright \\
&\langle \ell', x, \triangleright, c, x, \triangleright, 0, 1\rangle \\
&\langle c, x, y, c, y, b, +1, +1\rangle\,, \quad \forall x, y \neq b \\
&\langle c, x, b, h, x, b, 0, 0\rangle\,, \quad \forall x
\end{aligned}$$

The machine starts, it switches to state r, moves the first tape to the end of the input without moving the second tape, and when reaches a blank cell, switches to state ℓ. In that state, it moves first tape left, the second one right, and copies the number on tape 0 to tape 1 backwards until it reaches the start position (marked $\triangleright$), also erasing the input. Then it switches to state ℓ, and rewinds tape 2, until it reaches the start position on that one, and then switches to state c copies tape 2 onto 1 until it reaches a blank on tape 2, and finally halts.

Ex. 3.4. Let us use a three tape Turing machine, get the input on tapes 1 and 2, both with the least significant bit first, and leave the result on tape 3, also in the same order. (In the previous exercise we have seen that it is possible to reverse the bits, and to copy the result from tape 3 to tape 1). The program is

$$\begin{aligned}
&\langle q_s, \triangleright, \triangleright, \triangleright, q_a, \triangleright, \triangleright, \triangleright, 1, 1, 1\rangle\,, \\
&\langle q_a, x, y, z, q_a, x, y, x + y \bmod 2, 1, 1, 1\rangle\,, \quad \forall x \neq b, y \neq b \\
&\langle q_a, x, b, z, q_a, x, b, x, 1, 0, 1\rangle\,, \forall y \neq 0 \\
&\langle q_a, b, y, z, q_a, b, y, y, 0, 1, 1\rangle\,, \forall x \neq 0 \\
&\langle q_a, b, b, z, h, b, b, z, 0, 0, 0\rangle\,.
\end{aligned}$$

Ex. 3.5. This is known as the blank space halting problem. The proof is indirect, we shall assume that there is a solution to the blanks pace halting problem, i.e., a function $h_b(x)$ which is 0 if the Turing machine with Turing number x does not halt for an epty input, and 1 if it does. We shall show that if this was true then the function h in Box 3.2 would also be computable. To show this, to any Turing number x and input w we construct a new Turing machine with the program:

```
T_w(x):
  erase tape
  write w
  run T(x) on input w
```

It is clear that this is a Turing machine, and if h_b was computable, so was h.

Ex. 3.6. Let us construct the machine used in the construction in Box 3.2, replacing h with h_p. If for this machine $h_p(x)$ is true, then in more then $1/2$ of the runs, y is true, and the machine runs forever, contradicting $h_p(x)$ being true. If $h_p(x)$ is false, then in more than $1/2$ of the rund, y is false, and the machine stops. Again, this contradicts $h_p(x)$ being true.

Ex. 3.7. No. We may use the same proof as for the Turing machine, i.e., construct an oracle machine, and replace $h(x)$ with the function that computes that the oracle machine number x halts.

Ex. 3.8. To show that NAND can be used to simulate all other gates, let as proceed as follows: note that

$$x \,\text{NAND}\, x = \bar{x}\,,$$

and $x \wedge x = x$. This may be used to show that

$$(x \,\text{NAND}\, y) \,\text{NAND}\, (x \,\text{NAND}\, y) = x \wedge y\,,$$

constructing the AND gate. To produce the XOR gate, note, that

$$x \vee y = \overline{\bar{x} \wedge \bar{y}}$$

and

$$x \,\text{XOR}\, y = (x \vee y) \wedge \overline{(x \wedge y)}\,,$$

and thes only contained operations that we have already constructed (AND and NOT and OR in the latter case).

Ex. 3.9. 1. If $f(n)$ is $O(g(n))$, then there exist n_0 and $c > 0$ such that for $n > n_0$, $f(n) \leq c{\cdot}g(n)$. This, however, means that for the smae n_0, $g(n) \geq (1/c)f(n)$, i.e., $g(n)$ is $\Omega(f(n))$ with the constant $c' = 1/c$.

2. If, on the other hand, $g(n)$ is $\Omega(f(n))$, by definition, that means, that there are constants n_0 and $c > 0$ such that for $n > n_0$, $g(n) \geq cf(n)$. This, however, means that for the same n_0, $n > n_0$, $f(n) \leq (1/c)g(n)$, i.e., $f(n)$ os $O(g(n))$ with constant $1/c$.

Ex. 3.10. Let $g(n) = a_k n^k + \cdots + a_1 n + a_0$. What we need to show is that there exist c and n_0 that for $n > n_0$, $f(n) \leq cn^\ell$ for all $\ell > k$.

We shall first note that for any $m > 0$, and $n > 1$, $n^m > n^{m-1}$, so $a_m n^m + a_{m-1} n^{m-1} > (a_m + a_{m-1}) n^m$. This way, we may eliminate lower powers by induction, and see that $g(n) < (a_k + a_{k-1} + \cdots + a_0) n^k < (a_k + a_{k-1} + \cdots + a_0) n^\ell$ for $\ell > k$.

Ex. 3.11. It suffices to show that there is an n_0 that for $n > n_0$, $\log n < n$. This already holds for $n = 2$, and we shall show that $\log n - n$ is monotonously decreasing (for large enough n), by looking at its derivative, $(\log x - x)' = 1/(x \ln 2) - 1 < 0$ for $x > 2$.

Ex. 3.12. As for $n > 2^k$, $n^{\log n} > n^k$, n^k is $O(n^{\log n})$. On the other hand, assuming that $n^{\log n}$ is $O(n^k)$ would mean assuming the existence of n_0, k and c such that $n^{\log n} \leq cn^k$ for all $n > n_0$. On the other hand, for $n > 2^{k+1}$, $n^{\log n} > n^{k+1}$ and for large enough n, $n^{k+1} > cn^k$.

Ex. 3.13. As $n^{\log n} = 2^{(\log n)^2}$ and $c^n = 2^{n \log c}$, all we need to show is that $n \log c > (\log n)^2$ for n large enough, or, equivalently, $2^{\log n} \log c > (\log n)^2$, which is clear from the fact that the Taylor series of an exponential contains all powers with positive coefficients.

Ex. 3.14. For $n > n_0'$, $e(n) \leq c' f(n)$ and for $n > n_0''$, $g(n) \leq c'' h(n)$. It follows that for $n > n_0 = \max\{n_0', n_0''\}$, $e(n)g(n) \geq c'c'' f(n)h(n)$.

Ex. 3.15. At each compare-and-swap operation we exchange or do not exchange 2 elements in the list. Let us assume that for a given algorithm finishing in at most k steps, there exist such orderings of the list, that in each step, for one of them, the comparison results in a swap, and in the other one, no swap is done. This means, that running the steps backward, we started with the ordered list, and in each step, there is a branching, and therefore, in the first step, we obtain 2 new orderings, in the second, 4, and in the end, 2^k. This means, that some of the $k! < 2^k$ orderings were not visited, the algorithm could not have sorted those in k steps.

Ex. 3.16. The number of Boolean functions of n bits is 2^{2^n}, because the value of the most general function may be prescribed separately to each possible input, and this value may be either true or false.

Let us assume that we have a circuit with k gates and n wires. Then for each gate, we need to choose its input wires, either from one of the other gates or from the original n wires, yielding $\sim (n+k)^2$ possibilities, and the type of the gate, which is the number of 2-bit functions, 4. So we have $(4(n+k)^2)^k$ functions.

To be able to construct all the 2^{2^n} Boolean functions, we therefore need to satisfy

$$[4(n+k)^2]^k \geq 2^{2^n} ,$$

or taking a logarithm and neglecting constants,

$$k \geq \frac{2^n}{\log n} .$$

Ex. 3.17. If we could find the factors of a number in polynomial time, all we needed to do was compare the smallest factor with the input ℓ to see if the number has a factor smaller than ℓ.

If we could solve the factorisation decision problem in polynomial time, it would be possible to factorise the number in polynomial time by finding the smallest factor using a logarithmic search, dividing the number by that, and repeating this until all factors are found. The number of prime factors of a number are at most logarithmic (all factors are ≥ 2, so dividing them at least halves the number).

Ex. 3.18. If **P** agreed with **NP** than for any **NP** decision problem there was a polynomial time Turing machine that decided the problem. This machine can then be used to decide the relevant witness verification for both being in the language or for not being in the language by simply running it after a machine that just copies the input on a second tape and runs the machine on that one.

Ex. 3.19. A polynomial (actually, linear) time algorithm for the reachability problem can be constructed as follows: take one of the two vertices that need to be verified if they are reachable from each other. In each step, add one of their neighbours to a list, ad cross them off from the list of all vertices. When we ran out of neighbours, add the neighbours of the neighbours, and so on. If the second vertex is added, it is reachable. If at one point we run out of neighbours (i.e., all the neighbours of the vertices in the list of visited vertices are already crossed out from the list of all vertices) then the second point is not reachable. As in each step we add a vertex to the list of reachable vertices, the algorithm ends in at most as many steps as there are vertices in the graph.

Ex. 3.20. It is obvious that if the graph has an Euler cycle then all vertices are of even rank. Starting from one vertex along the cycle we leave that vertex, then move to another one, leave that one, move to another one, and so on. For each arrival at a vertex there is also a departure, each time adding two to the rank of the vertex.

If we have a connected graph with all vertices of even rank, we may construct the Euler cycle using Hierholzer's algorithm. We start from a vertex, and always move to a next one along an edge. It is not possible to get stuck as that would only be possible at a vertex with an odd rank. This is done as long as we reach the original vertex. After this has been done checks recursively, along this cycle, if there are edges left. Along those the algorithm is repeated, always inserting the new cycle into the old one. As in each step at least two edges are added, and the number of edges is $O(n^2)$, n denoting the number of vertices, the algorithm is polynomial time.

Ex. 3.21. L_1 being reducible to L_2 means that there is a mapping R_1 such that for any string x in the alphabet of L_1, $R_1(x)$ is in L_2 iff x is in L_1, and R_1 is calculable with a Turing machine in polynomial time. In this case, the length of the string $R_1(x)$ must also be polynomial, otherwise printing it would take longer than polynomial time. The reducibility of L_2 to L_3 means the existence of a similar mapping R_2. Therefore, $R := R_2 \circ R_1$ is a mapping reducing L_1 to L_3, and it is computable in polynomial time, as $R_1(x)$ is, and its length is polynomial, therefore $R_2(R_1(x))$ is computable in polynomial time in the length of $R_1(x)$, and a polynomial of a polynomial is polynomial, therefore, it is computable in polynomial time in the length of x as well.

Ex. 3.22. Using the results in the previous exercise, if L is complete, that means that for any L_1, there is a mapping R_1 reducing it to L. The reducibility of L to L' means that there is a mapping R' reducing L to L'. In that case, for the language L_1 is reduced to L' using the mapping $R'_1 := R' \circ R_1$.

Ex. 3.23. 1. To show that SAT is **NP**, we need to show that a witness to it can be verified in polynomial time by a Turing machine. To do this, we construct a Turing machine that first evaluates the inner brackets, writes it on the tape next to the input, then the next brackets, and so on. As there are a finite number of operations in the formula, this ends in a polynomial time in the number of operations (number of such steps is linear, and there is bookkeeping to know where to get the results from).

2. To show that SAT is **NP**-complete, we need to reduce another **NP**-complete problem to it. A Boolean circuit can be represented by a formula: each intermediate gate corresponds to a sub-expression in an inner bracket, and the last one is the result. This way, the equivalence of SAT and CSAT is demonstrated.

Ex. 3.24. A k-variable formula in the 2-conjunctive normal form is

$$(y_1 \vee y_2) \wedge \cdots \wedge (y_{2n-1} \vee y_{2m}) \,,$$

where all the y's may be taken from $x_1, \ldots, x_k$ and $\bar{x}_1, \ldots, \bar{x}_k$. The directed graph constructed encodes the relation "y_i being 0 means y_j has to be 1 for the formula to be true". If there is a path connecting x_i to $\bar{x}_i$, that means the formula cannot be satisfied.

For a directed graph, reachability can be decided in polynomial time. We list the vertices $v_1, \ldots, v_N$. We start with the starting vertex v_m, and in the first step, list the vertices that are connected to it. We cross them out from the list. In the next steps, we list vertices connected to all vertices in our list of reachable vertices. The procedure ends, if either the vertex whose reachability is the question is reached or there are no vertices to be added. As in each step either at least one vertex is checked or the procedure ends, this ends in polynomial time.

Ex. 3.25. Let us first consider how an algorithm can take a given amount of time. If the algorithm does stop, it has to take account of its progress, so the number of possible data stored in the space used multiplied by the number internal states must be at least as much as the number of steps.

If an algorithm is in **PSPACE** then for an input of length n it uses at most $p(n)$ bits of space, wher p is some polynomial. The number of possible data values stored in k bits of storage is 2^k, so in this case, the number of possible data values is $2^{p(n)}$, and with ℓ internal states, the total number of (internal and storage) states is $\ell 2^{p(n)}$ which is exponential, if the polynomial is of order N, then the algorithm uses time $O((2^N)^n)$.

Ex. 3.26. Let us assume that the Turing machine has ℓ internal states. Another part of the full state of the machine is position of the input tape, this is n states, and we have $k \log n$ bits on the second tape, which amounts to $2^{k \log n} = n^k$ states, so the total number of states is $\ell n n^k = \ell n^{k+1}$. The machine has to keep track of which step it is in, so the number of steps in the execution must be less than the number of total states. This way we have shown that $\mathbf{L} \subseteq \mathbf{P}$.

Ex. 3.27. It is clear that the algorithm constructs a vertex cover, as only such edges are removed from E' that at least one of whose endpoints are in the cover. It also contains at most double the size of a minimal vertex cover as a vertex cover must include at least one endpoint of each edge, and this includes for some edges both.

Ex. 3.28. If correctly accepting or rejecting a word x has probability $1/2 < k < 1$, then it is possible to repeat the test ℓ times and accept the majority as the result. This is only a constant multiplier. According to Condorcet's jury theorem the probability of the repeated runs yielding correct results is above $3/4$ for ℓ large enough, and what is large enough only depends on k, not the length of the input.

Ex. 3.29. The Fredkin gate can be expressed with the formulae

$$\begin{aligned} a' &= (\bar{c} \wedge a) \vee (c \wedge b) \,, \\ b' &= (\bar{c} \wedge b) \vee (c \wedge a) \,, \\ c' &= c \,. \end{aligned}$$

Applying another Fredkin gate with inputs a', b', c' and outputs a'', b'', c'' we obtain

$$c'' = c' = c \,,$$

so

$$a'' = [\bar{c} \wedge [(\bar{c} \wedge a) \vee (c \wedge b)]] \vee [c \wedge [(\bar{c} \wedge b) \vee (c \wedge a)]]$$
$$= \bar{c} \wedge \bar{c} \wedge a \vee \bar{c} \wedge c \wedge b \vee c \wedge \bar{c} \wedge b \vee c \wedge c \wedge a = \bar{c} \wedge a \vee c \wedge a = (\bar{c} \vee c) \wedge a = a\,,$$

and similarly

$$b'' = [\bar{c} \wedge [(\bar{c} \wedge b) \vee (c \wedge a)]] \vee [c \wedge [(\bar{c} \wedge a) \vee (c \wedge b)]]$$
$$= \bar{c} \wedge \bar{c} \wedge b \vee \bar{c} \wedge c \wedge a \vee c \wedge \bar{c} \wedge a \vee c \wedge c \wedge b = \bar{c} \wedge b \vee c \wedge b = (\bar{c} \vee c) \wedge b = b\,.$$

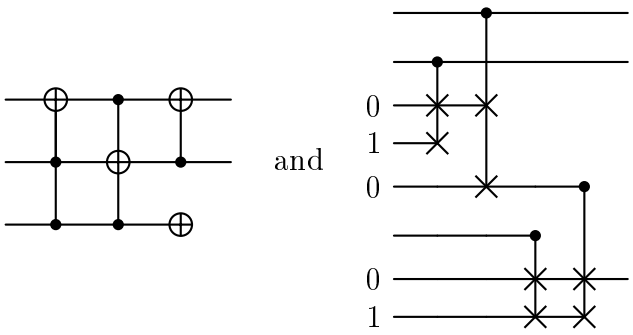
Ex. 3.30. Draw the possibilities, e.g., if $a = 1$ and all others 0, it moves u,d,u,u,u,d,d,d, yielding $a' = 1$, and all others 0 (as there is one ball).

If $a = 1$, $b = 0$ and $c = 1$, the movements of the a ball are u,d,u and the c ball are u, d, d, d and then a collision happens, and they swap direction, a ball moves d, c moves u, d, another collision happens, a moves d, u, c moves u, d, another collision, a moves d, u, c moves u, d, another collision, a moves d, u, d, u, u and exits at b', c moves u, u, u, and exits at c'.

The other possibilities may be analysed similarly.

Ex. 3.31. A (not necessarily optimal) solution is to express the outputs of the half-adder az $x \oplus y = (x \vee y) \wedge \overline{x \wedge y}$ and $c = x \wedge y$ and express the OR gates as in Fig. 3.16, and the OR and NOT gates using $x \vee y = \overline{\bar{x} \wedge \bar{y}}$. We then apply CNOTs to extend it to add it to a fourth register, and then add the inverse of the half-adder part expressed with Fredkin gates. The inverse is easily constructed using the Fredkin gates in reverse order, as the reverse of the Fredkin is its inverse.

Ex. 3.32. The following two figures show the simulation of a Fredkin gate with 3 Toffoli gates, and that of a Toffoli gate with four Fredkins. Control bits are marked with dots, the exchange/cnot bits are marked with crosses/opluses.



Pr. 3.1. (1) Let us assume that $f(n)$ is a computable function, i.e., there is such a Turing machine that for the input n on its tape it produced $f(n)$ on its tape when it halts. We wish to show that it is possible to evaluate this function on a Minsky machine.

The Turing machine has some internal states. We may enumerate the internal states, and store the state in one register. The interesting part is storing the values stored on an infinite amount of tape (but a finite alphabet) in a finite number of registers (capable of storing arbitrarily) large numbers. Let the alphabet consist of k symbols 0, 1, ..., $k-1$, then a finite word $a_1 a_2 \dots a_\ell$ on the tape can be stored in a register as $a_1 + a_2 k + \dots + a_\ell k^{\ell-1}$. The position of the head may be stored in another register. For simulating the Turing machine what remains to be done is to implement as a Minsky machine the change of state and the writing on the tape as a change according to tape position.

(2) Representing the Minsky machine on a Turing machine: if we have k registers, we choose a Turing machine with k tapes and write the numbers on the tape in binary. What is needed is the program to increment/decrement.

Pr. 3.2. To simulate a Minsky machine with a vector game, in addition to the values of the registers we need to keep track of the state of the machine to enforce the right vector being used. To do this, we add two additional components to the vector. The "increment" instruction going from state m to n is encoded in the vector $(0,\ldots,0,1,0\ldots,0,-m,n)$ where the 1 is in the component corresponding to the register to be incremented, and the decrement operation is encoded in two vectors, $(0,\ldots,0,-1,0,\ldots,0,-m,n)$ and $(0,\ldots,0,-m,p)$. The vectors are then listed in the order of decreasing m. To the beginning of the list we add $(0,\ldots,0,1,-1)$. See Ref. [2, 3].

Pr. 3.3. For a FRACTRAN program, we may construct a vector game. The input 2^n corresponds to a vector $(n,0,\ldots,0)$, and a multiplication with a rational q_i to the vector $(k_1,k_2,\ldots,0)$ where k_i are the exponents of the prime factors of the rational q_i. This way any vector machine can be represented as a FRACTRAN program [2]

Pr. 3.4. The above problems show that an algorithm that would decide if a FRACTRAN program reaches 1 would be equivalent to the decision problem.

Pr. 3.5. For any 2-bit reversible gate, the output part of a truth table is a permutation of the input part, and so, in any column, there are two zeros and two ones. This way, $\binom{4}{2}=6$ binary functions can be computed, and these may be listed: x, y, xXORy and their negations, i.e., the gates that one can construct are also constructable from the NOT and XOR gates.

The NOT and XOR gates do not form an universal set. By induction over the number of gates, one may show that the output of a network containing only these gates either does not depend on each input bit or becomes negated when that bit is negated. The logical or does not have this property, so that cannot be implemented with reversible two-bit gates.

The Toffoli gate is universal, therefore, if that could be implemented with reversible two-bit gates, then everything could be implemented with these too, and we have already seen that this is not the case.

Pr. 3.6. If the graph $G(V,E)$ has a Hamiltonian cycle, the length of that is $|V|$. A TSP optimal solution in the graph then has also length $|V|$. The approximator must return a TSP solution which is at most r times worse then the optimal, and just one edge chosen not in the original graph would yield a TSP solution at least of cost $\lceil r\rceil|V|+1$. So if the approximator is in **P**, it solves the **NP**-complete HC in polynomial time, which would prove $\mathbf{P}=\mathbf{NP}$.

Pr. 3.7. See Ref. [4]. A three-tape Turing machine is constructed as the reversible extension of a single-tape machine. The bounds for time and space are $4t(x)+4o(x)+4$ and $s(x)$, $t(x)+1$ and $o(x)+2$ on the three tapes, where $o(x)\leq s(x)$ is the size of the output. The main idea is to keep a history on the second tape, and a copy of the output on the third.

Part II
Quantum computation

4 Quantum circuits

4.1 Quantum algorithms

4.2 Single qubit operations

Ex. 4.1. We have calculated the normalised eigenvectors of the Pauli matrices in ex. 2.11. We shall now calculate the corresponding points on the Bloch sphere.

The case of $\sigma_0 = I$ is trivial, all vectors (all points) all eigenvectors. We may choose the two eigenvectors of σ_3, for example.

The case of σ_3 is as follows: the eigenvector corresponding to eigenvalue 1 is $|0\rangle$, which corresponds to $\cos(\theta/2) = 1$, so $\theta = 0$ and ϕ arbitrary. The other eigenvector is $|1\rangle$, similarly corresponding to $\theta = \pi$, and, as the coefficient of $|0\rangle$ vanishes, the relative phase becomes an overall phase, ϕ is arbitrary again.

For σ_1 we have for the eigenvalue 1 $\cos(\theta/2) = \sin(\theta/2) = 1/\sqrt{2}$, i.e., $\theta = \pi/2$ and the relative phase is 1, $\phi = 0$ and for the eigenvalue -1, we have $\cos(\theta/2) = 1/\sqrt{2}$, $\sin(\theta/2) = -1/\sqrt{2}$, yielding $\theta = -\pi/2$ and again $\phi = 0$.

For σ_2 we have for the eigenvalue 1 $\cos(\theta/2) = \sin(\theta/2) = 1/\sqrt{2}$, i.e., $\theta = \pi/2$ and the relative phase is i, $\phi = \pi/2$ and for the eigenvalue -1, we have $\cos(\theta/2) = 1/\sqrt{2}$, $\sin(\theta/2) = -1/\sqrt{2}$, yielding $\theta = -\pi/2$ and again $\phi = \pi/2$.

Ex. 4.2. Let A be such a matrix that $A^2 = I$, then for x real

$$\exp(\mathrm{i}Ax) = \sum_{n=0}^{\infty} \frac{(\mathrm{i}Ax)^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} + \mathrm{i}A \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = I\cos(x) + \mathrm{i}A\sin(x)\,.$$

As for all three Pauli matrices $X^2 = Y^2 = Z^2 = 1$ holds, this may be used with $A = X, Y, Z$ and $x = \theta/2$ to verify eqs. (4.4)-(4.6).

Ex. 4.3. For the Pauli matrix Z and the $\pi/8$ gate:

$$R_z(\pi/4) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\pi/8} & \\ & \mathrm{e}^{-\mathrm{i}\pi/8} \end{pmatrix} = \mathrm{e}^{-\mathrm{i}\pi/8} \begin{pmatrix} 1 & \\ & \mathrm{e}^{-\mathrm{i}\pi/4} \end{pmatrix} = \mathrm{e}^{-\mathrm{i}\pi/8} T\,,$$

i.e.,

$$T = \mathrm{e}^{\mathrm{i}\pi/8} T\,.$$

Ex. 4.4. We could try if two matrices suffice. In that case, depending on the order, either $R_z H$ or $R_x H$ should be such that it can be made proportional to R_x or to R_z, respectively. It is easily seen that this is not the case, as either the 11 and 12 elements cannot be turned into a cosine and a sine or the off-diagonal ones cannot vanish. Next, one would consider the form $H = \mathrm{e}^{\mathrm{i}\alpha} R_x(\theta) R_z(\theta') R_x(\theta)$, where examining $R_x(-\theta) H R_x(-\theta)$, and solving for θ such that this is diagonal. The result is

$$H = \mathrm{i} R_x(\pi/2) R_z(\pi/2) R_x(\pi/2)\,.$$

Ex. 4.5.

$$(\vec{n}\cdot\vec{\sigma})^2 = (n_i\sigma_i)(n_j\sigma_j) = n_in_j\sigma_i\sigma_j = n_in_j(\delta_{ij}I + \mathrm{i}\epsilon_{ijk}\sigma_k) = n^2I = I\,.$$

Ex. 4.6. The state corresponding to the Bloch vector $\vec{\lambda}$ is given by the density matrix

$$\rho = \frac{1+\vec{\lambda}\cdot\vec{\sigma}}{2}\,,$$

which is transformed as

$$\rho' = R_{\vec{n}}(\theta)\rho R_{\vec{n}}(\theta)^\dagger$$

where

$$R_{\vec{n}}(\theta) = \cos(\theta/2)I + \mathrm{i}\sin(\theta/2)(\vec{n}\cdot\vec{\sigma})\,,$$

therefore what we need is to show that

$$R_{\vec{n}}(\theta)\vec{\sigma}R_{\vec{n}}(\theta)^\dagger = R^{(3)}_{\vec{n}}(\theta)\cdot\vec{\sigma}\,,$$

which is easily done by computer algebra, using the products of the Pauli matrices.

Ex. 4.7. Using the product rule of Pauli matrices,

$$XYX = \mathrm{i}ZX = -Y\,,$$

and using this and eq. (4.5) yields

$$XR_y(\theta)X = X\left[\cos(\theta/2)I + \mathrm{i}\sin(\theta/2)Y\right] = \cos(\theta/2)I - \mathrm{i}\sin(\theta/2)Y = R_y(-\theta)\,.$$

Ex. 4.8. Any 2×2 matrix may be written as

$$U = (\alpha+\mathrm{i}\beta)I + \mathrm{i}(\vec{u}+\mathrm{i}\vec{v})\cdot\vec{\sigma}\,,$$

and for U to be unitary $U^\dagger U = I$ must hold with

$$U^\dagger = (\alpha-\mathrm{i}\beta)I - \mathrm{i}(\vec{u}-\mathrm{i}\vec{v})\cdot\vec{\sigma}\,,$$

so

$$U^\dagger U = (\alpha^2+\beta^2+u^2+v^2)I + 2(\beta\vec{u}-\alpha\vec{v}-\vec{u}\times\vec{v})\cdot\vec{\sigma} = I\,.$$

Unless $\vec{u}$ and $\vec{v}$ are collinear, together with $\vec{u}\times\vec{v}$ they form a basis, so the coefficient of $\vec{\sigma}$ can only vanish if all coefficients vanish. This is not a good solution, so they must be collinear. In this case the cross products vanish, and we may write $\vec{u} = \alpha\vec{w}$, $\vec{v} = \beta\vec{w}$, making the coefficients of the Pauli matrices vanish, and the coefficient of I is then

$$(\alpha^2+\beta^2)(1+w^2)\,,$$

which can be set to 1 if we choose

$$\alpha = \cos\alpha\cos(\theta/2)\,,\quad \beta = \sin\alpha\cos(\theta/2)\,,\quad w^2 = \tan^2(\theta/2)\,,$$

e.g., $\vec{w} = \tan\theta/2\vec{n}$, yielding

$$u = \cos\alpha\sin(\theta/2)\vec{n}\,,\quad v = \sin\alpha\sin(\theta/2)\vec{n}\,,$$

where $\vec{n}$ is a unit vector, and then

$$U = \exp(\mathrm{i}\alpha) R_{\vec{n}}(\theta)\,.$$

this completes the proof for (a).

(b) For the Hadamard gate, let us first separate the trace, $2\cos(\theta/2) = \operatorname{Tr} H = 0$ yielding $\theta/2 = \pm\pi/2$, $\theta = \pm\pi$. The decomposition may be obtained using $\operatorname{Tr}\sigma_a\sigma_b = 2\delta_{ij}$, so the coefficients are obtained as traces

$$\begin{aligned}
\mathrm{i}\mathrm{e}^{\mathrm{i}\alpha}\sin(\theta/2)n_1 &= \frac{1}{2}\operatorname{Tr}\sigma_1 H = 1/\sqrt{2}\,,\\
\mathrm{i}\mathrm{e}^{\mathrm{i}\alpha}\sin(\theta/2)n_2 &= \frac{1}{2}\operatorname{Tr}\sigma_2 H = 0\,,\\
\mathrm{i}\mathrm{e}^{\mathrm{i}\alpha}\sin(\theta/2)n_3 &= \frac{1}{2}\operatorname{Tr}\sigma_3 H = 1/\sqrt{2}\,,
\end{aligned}$$

yielding, e.g.,

$$\theta = \pi\,, \quad n_1 = n_3 = 1/\sqrt{2}\,, \quad n_2 = 0\,, \quad \alpha = -\pi/2\,.$$

(c) We may read off $\alpha = \pi/4$, $\theta = -\pi/2$, $n_1 = n_2 = 0$ and $n_3 = 1$.

Ex. 4.9. To understand why a single qubit unitary operator can be written in the form (4.12), note, that the columns of a unitary operator must be unit vectors, and a 2 element complex unit vector may be parametrised by an overall phase (δ), a relative phase (β) and an angle giving the magnitude of the two components (θ). This also determines the second column up to a relative phase (δ becomes the relative phase), and then there is a new overall phase (α).

Ex. 4.10. If we multiply the matrices, we get $U = \exp(\mathrm{i}\alpha) R_x(\beta) R_y(\gamma) R_x(\delta)$, or in matrix form

$$U = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\alpha-(\beta+\delta)/2)}\cos\frac{\gamma}{2} & -\mathrm{e}^{\mathrm{i}(\alpha-(\beta-\delta)/2)}\sin\frac{\gamma}{2} \\ \mathrm{e}^{\mathrm{i}(\alpha+(\beta-\delta)/2)}\sin\frac{\gamma}{2} & \mathrm{e}^{\mathrm{i}(\alpha+(\beta+\delta)/2)}\cos\frac{\gamma}{2} \end{pmatrix}.$$

The argument that this covers all unitary matrices is similar to that in the previous exercise: the rows of a unitary matrix have to be normalised and unitary; a single 2-component unit vector has the following parameters: magnitude angle (γ), common phase (β) and relative phase (δ), and this also determines the other row, the angle β becomes the relative phase of the two rows, and there is a new common phase α.

Ex. 4.11. First: see erratum to the exercise in the book! There is an infinite sequence of operators.

We may rephrase the Thm. 4.1 in the form that roations around two orthogonal axes generate all unitaries up to a phase. What we need is that we have two axes, and for that, we show via computation that

$$R_{\vec{n}}(\pi) R_{\vec{m}}(\gamma) R_{\vec{n}}(\pi) R_{\vec{m}}(-\gamma)$$

is a rotation with some angle around an axis $\vec{n}'$ such that $\vec{n}\cdot\vec{n}' = 0$.

To show that arbitrary axes are OK in stead of z, y, we just need a unitary V that takes $\vec{n}$ into z and $\vec{n}'$ into x and apply the decomposition to $V^\dagger U V$.

Ex. 4.12. Compare first the Hadamard operator,

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

with the form in eq. (4.12),

$$H = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\alpha-\beta/2-\delta/2)} \cos\frac{\gamma}{2} & -\mathrm{e}^{\mathrm{i}(\alpha-\beta/2+\delta/2)} \sin\frac{\gamma}{2} \\ \mathrm{e}^{\mathrm{i}(\alpha+\beta/2-\delta/2)} \sin\frac{\gamma}{2} & \mathrm{e}^{\mathrm{i}(\alpha+\beta/2+\delta/2)} \cos\frac{\gamma}{2} \end{pmatrix}$$

and read off the parameters

$$\alpha = \pi/2\,, \quad \beta = 0\,, \quad \gamma = \pi/2\,, \quad \delta = \pi\,.$$

Now, proceed as in the proof of Corollary 4.2, and set

$$A = R_z(\beta)R_y(\gamma/2) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\beta/2} \cos\frac{\gamma}{2} & -\mathrm{e}^{-\mathrm{i}\beta/2} \sin\frac{\gamma}{2} \\ \mathrm{e}^{\mathrm{i}\beta/2} \sin\frac{\gamma}{2} & \mathrm{e}^{\mathrm{i}\beta/2} \cos\frac{\gamma}{2} \end{pmatrix},$$

$$B = R_y(-\gamma/2)R_z(-(\delta+\beta)/2) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\beta+\delta)/4} \cos\frac{\gamma}{2} & \mathrm{e}^{-\mathrm{i}(\beta+\delta)/4} \sin\frac{\gamma}{2} \\ -\mathrm{e}^{\mathrm{i}(\beta+\delta)/4} \sin\frac{\gamma}{2} & \mathrm{e}^{-\mathrm{i}(\beta+\delta)/4} \cos\frac{\gamma}{2} \end{pmatrix},$$

and

$$C = R_z((\delta-\beta)/2) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\beta-\delta)/4} & \\ & \mathrm{e}^{-\mathrm{i}(\beta-\delta)/4} \end{pmatrix}.$$

It is easy to verify that $ABC = I$ and with the values of the angles α, β, γ, and δ obtained above, $\exp(\mathrm{i}\alpha)AXBXC = H$.

Ex. 4.13. Simple computation (see computer algebra code).

Ex. 4.14. Simple computation (see computer algebra code). With the global phase included

$$HTH = \mathrm{e}^{\mathrm{i}\pi/8} R_x(\pi/4)\,.$$

Ex. 4.15. See book errata!
(1) Direct calculation. See computer algebra code.
(2) Substitution.

4.3 Controlled operations

Ex. 4.16. Remember the numbering of states (page xxx): $0\ldots00$, $0\ldots01$ to $1\ldots11$. With this, the matrix of the Hadamard gate on the x_2 (upper) wire yields

$$I \otimes H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

In the case of putting the Hadamard gate on the x_1 (lower) wire, we obtain

$$H \otimes I = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

Note that the first one is putting Hadamard matrices in the place of elements of the unit matrix, and the second one is placing unit matrices on the elements of the Hadamard gate's matrix.

Ex. 4.17.

$$\mathrm{CNOT} = (I \otimes H)\mathrm{CZ}(I \otimes H)\,.$$

The control bit is, for both the CNOT and the CZ, the upper (x_2) one.

Ex. 4.18. The controlled Z-gate with the control bit on the upper (x_2) qubit has the matrix

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

which can be obtained by filling it columnwise, and keeping in mind that both bits are unchanged and there is a sign if bit 1 is 1 and the control bit is one. Similarly, putting the control bit on the lower (x_1) bit yields

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

The two matrixes agree.

Ex. 4.19. Let us write the elements of the density matrix as

$$\rho = \begin{pmatrix} \rho_{00,00} & \rho_{00,01} & \rho_{00,10} & \rho_{00,11} \\ \rho_{01,00} & \rho_{01,01} & \rho_{01,10} & \rho_{01,11} \\ \rho_{10,00} & \rho_{10,01} & \rho_{10,10} & \rho_{10,11} \\ \rho_{11,00} & \rho_{11,01} & \rho_{11,10} & \rho_{11,11} \end{pmatrix}\,.$$

With ρ above, the action of the CNOT gate is

$$\mathrm{CNOT}\rho\mathrm{CNOT} = \begin{pmatrix} \rho_{00,00} & \rho_{00,01} & \rho_{00,11} & \rho_{00,10} \\ \rho_{01,00} & \rho_{01,01} & \rho_{01,11} & \rho_{01,10} \\ \rho_{11,00} & \rho_{11,01} & \rho_{11,11} & \rho_{11,10} \\ \rho_{10,00} & \rho_{10,01} & \rho_{10,11} & \rho_{10,10} \end{pmatrix}\,.$$

Ex. 4.20. It can be shown by multiplying the matrices, that

$$(H \otimes H)\mathrm{CNOT}(H \otimes H) = \mathrm{CNOT}'\,,$$

where in the case of the CNOT the control bit is the upper one, and for CNOT$'$ it is the lower one.

This means, that CNOT has the same matrix in the $\pm$ basis as CNOT$'$ in the original one, keeping $|+\rangle|+\rangle$, $|-\rangle|+\rangle$, and exchangging $|+\rangle|-\rangle$ and $|-\rangle|-\rangle$. The latter can also be verified by direct calculation.

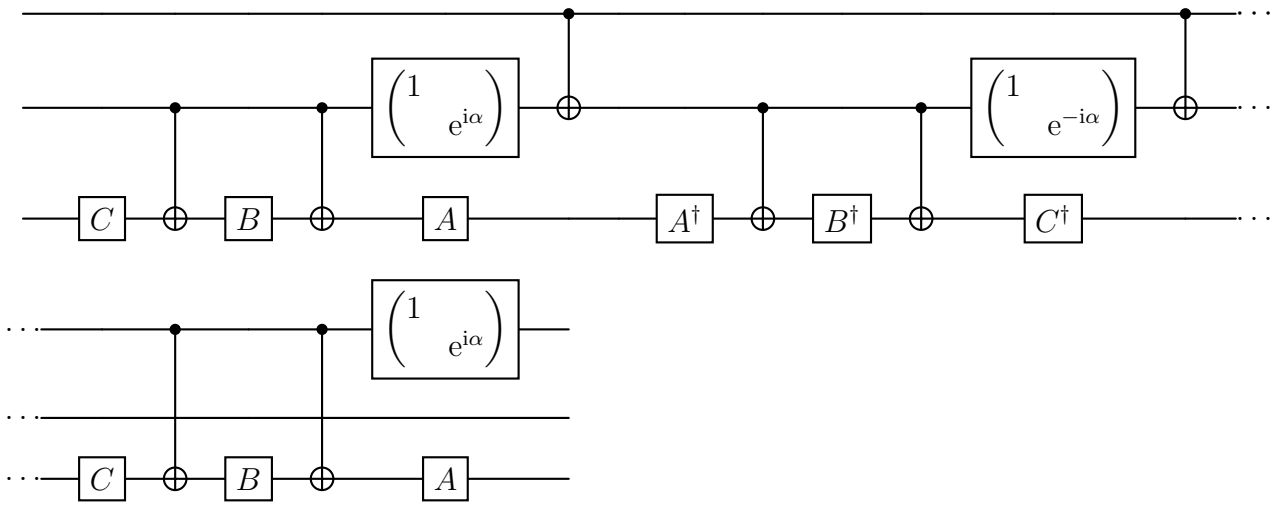
Ex. 4.21. Let us consider the case when the top two qubits are 0, then the leftmost controlled V does nothing, the left CNOT similarly does nothing, the controlled $V^\dagger$ does not act, neither does the second CNOT, and neither the last controlled V.

Now if the top qubits are 01, then none of the CNOTs act, and neither does the rightmost controlled V, so the action on the lowest qubit is $VV^\dagger = I$.

If the top qubits are 10, then the CNOTS both act, so between the two CNOTs the middle qubit becomes I too, and so the controlled $V^\dagger$ and the right V both act, and the effect on the lowest qubit is $V^\dagger V = I$.

If the two upper qubits are 11 then both the CNOTs act, and the middle qubit is 0 between the two, so the action on the last qubit is $V^2 = U$.

Ex. 4.22. We insert the CV and $CV^\dagger$ gates from fig. 4.6 into fig. 4.8. In this way, we construct C^2U using $3 \times 4 = 12$ single qubit gates and $3 \times 2 + 2 = 8$ CNOT gates. We need to reduce this to 8 one qubit gates and 6 CNOTs. To this, let us notice the following:



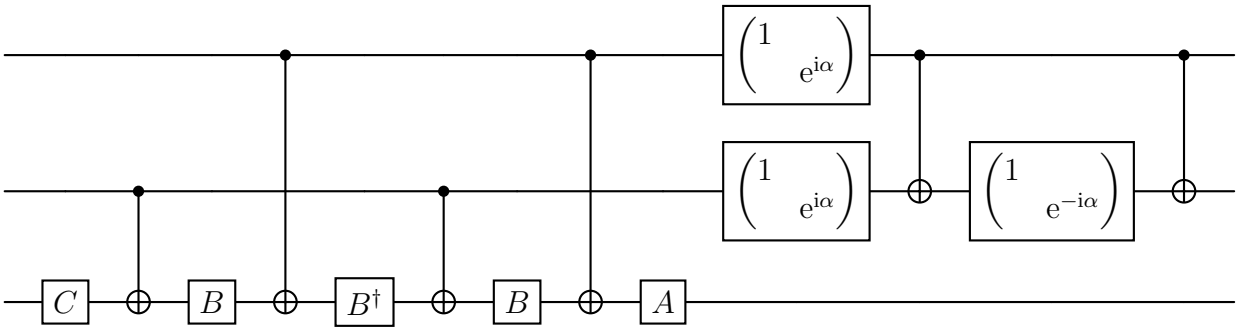
in the circuit, the $AA^\dagger = I = C^\dagger C$, so 4 single qubit operators drop out, we are down to 8. We still need to save 2 CNOTs.

Notice that the phase gates commute with all CNOTs that use their qubit as control, so they may be moved to be juts before and after the CNOT third from the right.

On the top two lines, the CNOTS 3rd and 6th from the left cancel each other. The difference is that CNOTs 4 and 5 are controlled by the result of CNOT 3, i.e., parity of the two control lines, so they may be replaced by two CNOTs, controlling the third line, one controlled from line 1 the other from 2, left of the CNOT originally 6th from the left.

The block consisting now of the phase shifts and the CNOTs that were originally 3rd and 6th from the left is diagonal, so it may be moved around freely.

Next we notice that $\mathrm{CNOT}^2 = I$, so there are 4 CNOTs that drop out. What remains is the following circuit:



which consists of 8 single qubit gates and 6 CNOTs. See also [5]

Ex. 4.23. As

$$R_x(\theta) = e^{i\alpha} R_z(\beta) R_y(\gamma) R_z(\delta)$$

with $\alpha = 0$, $\beta = -\pi/2$, $\gamma = \theta$, and $\delta = \pi/2$, the controlled $R_x(\theta)$ is given by fig. 4.6, with $A = R_z(\beta)R_y(\gamma/2)$, $B = R_y(-\gamma/2)R_z(-(\delta+\beta)/2)$, and $C = R_z((\delta-\beta)/2)$. This only includes 3 single qubit gates as $\alpha = 0$. I haven't been able to find a representation with fewer gates.

For $R_y(\theta)$, the decomposition is trivial, $\alpha = \beta = \delta = 0$ and $\gamma = \theta$, and so $A = R_y(\theta/2)$, $B = R_y(-\theta/2)$ an $C = I$, so one gate fewer is needed.

It is possible to show that we cannot do the same tric for $R_x(\theta)$, as that would require a decomposition $AB = I$, $AXBX = R_x(\theta)$. The first of these means $B = A^\dagger$, and $AXA^\dagger X = R_x(\theta)$, $AXA^\dagger = R_x(\theta)X$. The left hand side is a Hermitean matrix, and the right is not, unless $\theta = 0$.

Ex. 4.24. It is easy to verify what the circuit does on vectors of the form $|00\rangle|x\rangle$, $|01\rangle|x\rangle$, etc.

The simplest case is that of $|00\rangle|x\rangle$ in which case all the CNOTS do nothing, and so x is acted on by the operator $TT^\dagger TT^\dagger H = I$.

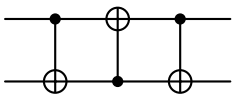
In the case of vectors of the form $|01\rangle|x\rangle$, the operator $HTT^\dagger XTT^\dagger XH = I$ acts on x.

In the case of vectors of the form $|10\rangle|x\rangle$ the middle qubit becomes 1 after the first CNOT acting on it, then collects a phase $\exp(-\mathrm{i}\pi/4)$ from the $T^\dagger$ gate second from the left on it, and then becomes $|0\rangle$ with this phase again at the next CNOT, the top qubit collects a phase $\exp(\mathrm{i}\pi/4)$ from the T gate on the top wire, the two cancel, so the operator $HTXT^\dagger TXT^\dagger H = I$ acts on x.

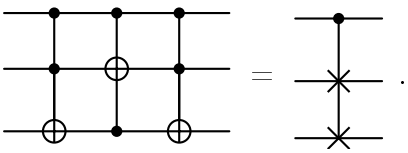
In the case of vectors of the form $|11\rangle|x\rangle$ similarly a phase $\exp(\mathrm{i}\pi/4)$ is collected by the top qubit, $\mathrm{i}\exp(-\mathrm{i}\pi/4)$ by the second, and so the operator $\mathrm{i}HTXT^\dagger XTXT^\dagger XH = X$ acts on x.

The above actions are identical to the action of the Toffoli gate.

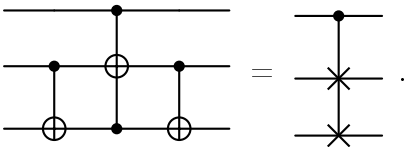
Ex. 4.25. (1) In fig. 1.7, the swap gate was constructed from 3 cnots as



and we know that all these gates can be controlled separately if we replace all cnots with Toffolis,



By simple calculation, we can show that



(2) If we insert $V = (1-\mathrm{i})(I+\mathrm{i}X)/2$ and $V^\dagger = (1+\mathrm{i})(I-\mathrm{i}X)/2$ into the circuit in fig. 4.8, we obtain an implementation of the Toffoli gate. Adding the two (leftmost and rightmost) CNOTs from the figure above, we obtain an implementation of the Fredkin gate. It contains he 2 CNOTs and 3 controlled V's of fig. 4.8 and two additional CNOTs, i.e., 7 two-qubit gates.

(3) In the resulting circuit, the same (middle) qubit controls both the Vs on the lowest qubit, i.e., we may save two additional two qubit gates if we replace the controlled V's with a controlled XV and a controlled VX.

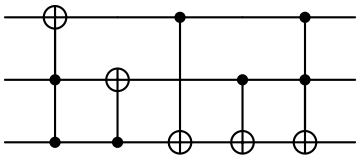
Ex. 4.26. By examining the result of the circuit on states of the form $|0\rangle|0\rangle|x\rangle, |0\rangle|1\rangle|x\rangle, \ldots$ the claim is verified, and the phases are

$$\theta(0,0,0) = \theta(0,0,1) = \cdots = \theta(1,1,1) = 0\,, \quad \theta(1,0,1) = \pi\,.$$

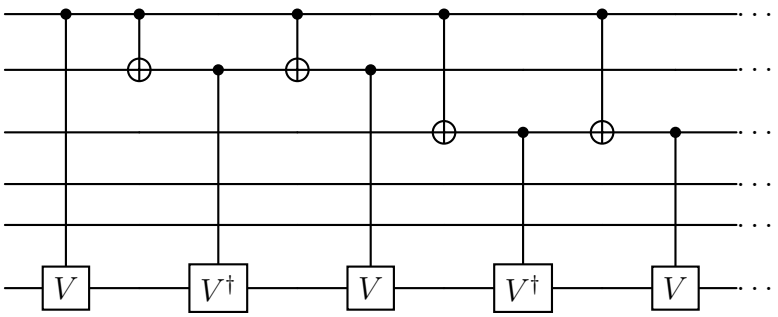
Ex. 4.27. There is a semi-systematic aproach based on the truth table of the circuit[6]. From the truth table, we may figure out the three output qubits as logical functions of the three input ones as

$$o_3 = x_3 \oplus (x_2 x_1)\,, \quad o_2 = x_2 \oplus x_1\,, \quad o_1 = x_2 \oplus x_2 \oplus x_3 x_1 \oplus x_2 x_1\,.$$

We recognise that o_3 is the result of a Toffoli gate controlled by the input qubits 1,2 and the target is the third, and o_2 is the output of a CNOT qubit 1 contolling 2. So we start the circuit with such a Toffoli and CNOT, and then note that the qubit 1 still has to be constructed. It contains x_2 and x_1, so we add two CNOTs, and as that still does not get the desired result, we add a Toffoli. Expanding the products shows that this yields the desired circuit:



Ex. 4.28. The C^5U gate is implemented using V, such that $V^4 = U$. The circuit is the following [7],

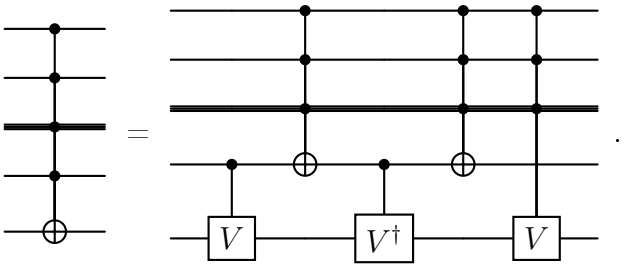


The idea behind it is that

$$x_5 - (x_5 \oplus x_4) + x_4 - (x_5 \oplus x_4) + x_3 - (x_5 \oplus x_3) + x_2 - \ldots \\ +(x_5 \oplus x_4 \oplus x_3) + \cdots = 4x_5x_4x_3x_2x_1\,,$$

where in the sum, we add 1 for a true and 0 for a false logical variable, and a positive power means an operator V and a negative one a $V^\dagger$.

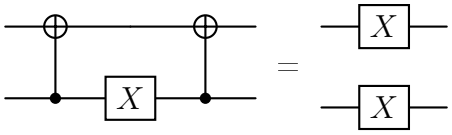
Ex. 4.29. Let us note that $V = (1-\mathrm{i})(I+\mathrm{i}X)/2$ is such that $V^2 = X$. Now the construction may be done recursively,



Let the cost for n control qubit be C_n. It is clear that $C_n = O(n)+C_{n-1}$, resulting in $C_n = O(n^2)$ [7].

Ex. 4.30. See 4.29 and Ref. [7].

Ex. 4.31. All these circuit identities can be verified by calculating the matrices (e.g., using computer algebra). In some cases it may be worthwile to do it more by checking on the elements of a basis, e.g., eq. (4.32),



It is easy to verify that the results for the basis vectors with labels 00, 01, 10, 11 are the basis vectors 11, 10, 01, and 00, respectively.

4.4 Measurement

Ex. 4.32. A projective measurement with measurement operators P_i gives the ith result with probability $p_i = \operatorname{Tr} P_i\rho$ and the resulting state is $\rho_i = P_i\rho P_i/\operatorname{Tr}(P_i\rho P_i) = P_i\rho P_i/p_i$. If the observer does not learn the outcome, i.e., an ensemble of systems is not separated according to the result, then after the measurement, the system is with probability p_i in state ρ_i, i.e.,

$$\rho' = \sum_i p_i \frac{P_i\rho P_i}{p_i} = \sum_i P_i\rho P_i\,,$$

which is the statement to be prove, eq. (4.40).

In the case of a composite system and a measurement on subsystem 2, in the above formula, all projection matrices are of the form $P_i = P_i \otimes I$, where P_i on the left acts on a product space and P_i on the right acts on the Hilbert space of ststem 2. In the present case, the latter are $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$.

The reduced density matrix of subsystem 1 before the measurement is $\rho_1 = \operatorname{Tr}_2\rho$. After the measurement it is

$$\rho_1' = \operatorname{Tr}_2\rho' = \sum_i \operatorname{Tr}_2(P_i\rho P_i)\,.$$

Expanding the density matrix in a product basis,

$$\rho = \sum_{ijk\ell} \rho_{ij,k\ell}|ij\rangle\langle k\ell|\,,$$

and calculating the traces, as in eq. (2.178), and using the orthonormality of the basis,

$$\rho_1 = \operatorname{Tr}_2\rho = \sum_{i,j,k} \rho_{ij,kj}|i\rangle\langle k|\,,$$

and similarly,

$$\rho_1' = \operatorname{Tr}_2\rho' = \sum_m \operatorname{Tr}_2 P_m\rho P_m\,,$$

where, using $P_m = |m\rangle\langle m|$,

$$P_m|ij\rangle\langle k\ell|P_m = (I \otimes |m\rangle\langle m|)|ij\rangle\langle k\ell|(I \otimes |m\rangle\langle m|) = \delta_{m,j}\delta_{\ell,m}|im\rangle\langle km|\,,$$

therefore

$$\rho_1' = \mathrm{Tr}_2 \sum_{ijk\ell}\sum_m \rho_{ij,k\ell}\delta_{m,j}\delta_{\ell,m}|im\rangle\langle km| = \sum_m \rho_{im,km}\,\mathrm{Tr}_2|im\rangle\langle km| = \sum_m \rho_{im,km}|i\rangle\langle m| = \rho_1\,,$$

so whe have demonstrated $\rho_1 = \rho_1'$, i.e., that a measurement on subsystem 2 without learning the outcomes does not affect the reduced density matrix of subsystem 1.

Ex. 4.33. Measurement of the Bell states corresponds to the measurement operators

$$B_{ij} = |\beta_{ij}\rangle\langle\beta_{ij}|$$

where β_{ij} denote the Bell states as given in eqs. (1.23)-(1.26). The circuit in the exercise first applies a unitary operator

$$U = (H \otimes I)\mathrm{CNOT}$$

and then performs the measurement using the projectors of the computational basis. Measurement of a pure state ψ in the computational basis gives probability amplitudes $\langle i,j|\psi\rangle$, measurement on $U\psi$ yields $\langle i,j|U\psi\rangle$, so the corresponding measurement operators are $U|\psi\rangle\langle\psi U^\dagger$, i.e., the measurement operators are now $UP_{ij}U^\dagger$, where $P_{ij} = |i,j\rangle\langle i,j|$ the measurement operators in the computational basis. What one needs to verify (matrix multiplication, may be done using computer algebra) is

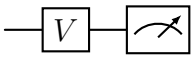
$$B_{ij} = UP_{ij}U^\dagger\,,$$

which completes the proof.

Ex. 4.34. If U has eigenvalues ± 1, that means it is of the form

$$U = \lambda_1|\lambda_1\rangle\langle\lambda_1| + \lambda_2|\lambda_2\rangle\langle\lambda_2|\,, \quad \lambda_i \in \{\pm 1\}\,.$$

One possibility for its measurement would require to find an operator V such that $V|0\rangle = \mathrm{e}^{\mathrm{i}\alpha_1}|\lambda_1\rangle$ and $V|1\rangle = \mathrm{e}^{\mathrm{i}\alpha_2}|\lambda_2\rangle$, and then use the circuit



We shall show that the corresponding measuring operators (see the previous solution) are the ones corresponding to U,

$$P_0' = V|0\rangle\langle 0|V^\dagger = |\lambda_1\rangle\langle\lambda_1| = P_{\lambda_1}\,,$$

and

$$P_1' = V|1\rangle\langle 1|V^\dagger = |\lambda_2\rangle\langle\lambda_2| = P_{\lambda_2}\,.$$

The operator V is unitary and

$$V = \mathrm{e}^{\mathrm{i}\alpha_1}|\lambda_1\rangle\langle 0| + \mathrm{e}^{\mathrm{i}\alpha_2}|\lambda_2\rangle\langle 1|\,.$$

However, this is a rather different operator from U. From U a controlled U may be constructed, and from a controlled U the circuit in the book. That circuit maps the incoming state into the following states,

$$|0\rangle|\psi\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|\psi\rangle \to \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + |1\rangle U|\psi\rangle \to \frac{1}{\sqrt{2}}(|0\rangle(I+U)|\psi\rangle + |1\rangle(I-U)|\psi\rangle\,,$$

and then in this final state is a measurement performed on the qubit 2. The corresponding measurement operators are $P_0 = |0\rangle\langle 0| \otimes I$ and $P_1 = |1\rangle\langle 1| \otimes I$. The probabilities are

$$p_i = \operatorname{Tr} P_i \frac{1}{2}(|0\rangle|(1+U)\psi\rangle + |1\rangle(1+U)|\psi\rangle)(\langle 0|\langle\psi|(1+U) + \langle 1|\langle\psi|(1-U)) = \left|\left\langle \psi \left| \frac{1\pm U}{\sqrt{2}} \right| \psi \right\rangle\right|^2 .$$

The resulting output state is

$$\psi_{\text{out}} = \frac{P_i \frac{1}{\sqrt{2}}(|0\rangle(I+U)|\psi\rangle + |1\rangle(I-U)|\psi\rangle}{\sqrt{p_i}} = \frac{\frac{I\pm U}{\sqrt{2}}|\psi\rangle}{\left|\left\langle \psi \left| \frac{1\pm U}{\sqrt{2}} \right| \psi \right\rangle\right|}\,,$$

which is both the same as in the case of measuring U.

Note, that we have assumed that the operator has one positive and one negative eigenvalue, otherwise it would be the trivial operator $\pm I$.

Ex. 4.35. The first circuit first maps an input state $|0\rangle \otimes \psi_1 + |1\rangle \otimes \psi_2$ to $|0\rangle \otimes \psi_1 + |1\rangle \otimes \psi_2$, and then performs the measurement on the upper qubit, resulting in the probabilities

$$p_0 = \|\psi_1\|^2\,, \quad p_1 = \|U\psi_2\|^2 = \|\psi_2\|^2\,,$$

and the output states

$$\frac{\psi_1}{\|\psi_1\|}\,, \quad \frac{U\psi_2}{\|U\psi_2\|} = \frac{U\psi_2}{\|\psi_2\|}\,.$$

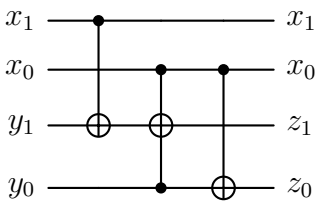
In the case of the second circuit, the measurement produces, with the same probabilities the post-measurement states

$$\frac{\psi_1}{\|\psi_1\|}\,, \quad \frac{\psi_2}{\|\psi_2\|}\,,$$

and then if the result was 1, the operator U is applied to the second state, yielding the same output states as in the first case. This completes the proof that the measurement commutes with control.

4.5 Universal quantum gates

Ex. 4.36. To construct the modulo 2 adder, $z = x + y$ mod4, let us consider the bits of z, $z_0 = x_0 \oplus y_0$, this is implemented using a CNOT, and $z_1 = x_1 \oplus y_1 \oplus x_0 y_0$, where we may construct the $x_2 \oplus z_2$ part using a CNOT and the rest using a Toffoli gate, yielding

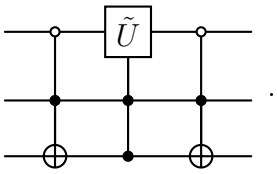


.

Ex. 4.37. We proceed as in the 3×3 case constructing U_1, then U_2, and use a similar formula (with the nontrivial rows being the columns 1, 4) to construct U_3. Then the same formula is used to construct U_4 (nontrivial columns 2, 3), U_5 (2, 4) and finally U_6 is constructed as U_3 in the 3×3 case. See computer algebra code for details.

Ex. 4.38. Let us consider the elements of the canonical basis of the d dimensinal space. Let us assign the vertices of a graph to the basis vectors, so we have d vertices. If we have a product of k two-level matrices, for each of these, we link the vertices corresponding to the basis elements mixed by the two-level matrix. In this way, for k matrices we get a graph of k edges. In a graph of d vertices, with less than $d-1$ edges all vertices cannot be linked. On the other hand, it is easy to construct a matrix that has no nonzero elements, i.e., in the image of any of the basis elements all others have a non-zero coefficient.

Ex. 4.39. The operator acts on basis vectors 2 (binary 010) and 7 (binary 111), so the Gray code is 010, 011, 111. We need to implement only one exchange, 010 to 011 with a C^2NOT acting on bit 1 (rightmost) controlled by bit 2 and the inverse of bit 3, then the C^2U controlled by 1 and 2 acting on 3 (the matrix looks the same but the element b is shifted one to the right, c one down and a one down and one to the right). The corresponding circuit is



.

Ex. 4.40. The error

$$E(U,V) = \max_{\psi:\|\psi\|=1} \|(U-V)\psi\|$$

is invariant to unitary transformations, as if T is unitary

$$E(TUT^\dagger, TVT^\dagger) = \max_{\psi:\|\psi\|=1} \|T(U-V)T^\dagger\psi\| = \max_{T\phi} \|T(U-V)T^\dagger T\phi\| = E(U,V)$$

where $\psi = T\phi$, and $\|T\phi\| = \|\phi\|$ and unitaries are bijective on the unit sphere, and for all vectors ξ, $\|T\xi\| = \|\xi\|$. We may therefore write $R_{\hat{n}}(\alpha) = UR_z(\alpha)U^\dagger$ where R is any such rotation that rotates $\hat{n}$ into the z axis.

In the case of the axis being the z axis,

$$R_z(\alpha) - R_z(\alpha+\beta) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\alpha} - \mathrm{e}^{-\mathrm{i}(\alpha+\beta)} & \\ & \mathrm{e}^{\mathrm{i}\alpha} - \mathrm{e}^{\mathrm{i}(\alpha+\beta)} \end{pmatrix} = R_z(\alpha) \begin{pmatrix} 1 - \mathrm{e}^{-\mathrm{i}\beta/2} & \\ & 1 - \mathrm{e}^{\mathrm{i}\beta/2} \end{pmatrix},$$

and as $R_z(\alpha)$ is unitary, we only need the norm of the rightmost matrix. For any vector $\psi = (\psi_0, \psi_1)^T$, this matrix maps it into the vector $((1-\mathrm{e}^{-\mathrm{i}\beta/2})\psi_0, (1-\mathrm{e}^{\mathrm{i}\beta/2})\psi_1)^T$ whose norm is $|1-\mathrm{e}^{-\mathrm{i}\beta/2}|\|\psi\|$.

Ex. 4.41. First, we calculate the state,

$$\psi = (H \otimes H \otimes I)\mathrm{Toffoli}(I \otimes I \otimes S)\mathrm{Toffoli}(H \otimes H \otimes I)|0,0\rangle \otimes \psi\,,$$

and then show that of we project using the operator

$$|0\rangle\langle 0| \otimes |0\rangle\langle 0| \otimes I\,,$$

The projected vector is

$$\sqrt{\frac{5}{8}}\frac{1+\mathrm{i}}{\sqrt{2}}R_z(\theta)\psi\,,$$

and the full vector is

$$|0\rangle|0\rangle\sqrt{\frac{5}{8}}\frac{1+\mathrm{i}}{\sqrt{2}}R_z(\theta)\psi + \frac{\mathrm{i}}{2\sqrt{2}}\frac{1+\mathrm{i}}{\sqrt{2}}(-|0\rangle|1\rangle - |1\rangle|0\rangle + |1\rangle|1\rangle)Z\psi$$

Consequently, 0, 0 is measured on qubits 3,2 with a probability $\sqrt{5/8}$, and in this case, the output state of the first (unmeasured) qubit is $R_z(\theta)\psi$ with a global phase.

Using $Z^2 = 1$, the procedure can be enhanced. If 0, 0 is measured, the output is used, if not, Z is applied, and the procedure is repeated. In all steps, there is a $\sqrt{5/8}$ chance of getting the result, so the probability of not yet stopping in step k is q^k where $q = 1 - \sqrt{5/8}$. $q^k \to 0$, so the probability of not yet having the result vanishes.

Ex. 4.42. Let us note first that $\sin^2\theta = 1 - \cos^2\theta = 1 - 9/25 = 16/25$, so $\sin\theta = \pm 4/5$. We shall assume the positive sign (the negative sign case is similar).

(1) If θ given by $\cos\theta = 3/5$ is a rational multiple of π, i.e., $\theta = \pi p/q$, where $p, q \in \mathbb{N}$, then $2q\theta = 2\pi p$, so

$$1 = \mathrm{e}^{\mathrm{i}m\theta} = \left(\frac{3}{5} + \mathrm{i}\frac{4}{5}\right)^m\,,$$

where $m = 2q > 0$, which may be recast as

$$(3+4\mathrm{i})^m = 5^m\,.$$

(2) Let us proceed by induction. Assuming $(3+4\mathrm{i})^k = 3 + \mathrm{i}4 (\mathrm{mod}\ 5)$, we have

$$(3+5\mathrm{i})^{k+1} = (3+4\mathrm{i})^k(3+4\mathrm{i}) = (3+4\mathrm{i})(3+4\mathrm{i})\ (\mathrm{mod}\ 5) = (3+5\mathrm{i})^2 = (3+4\mathrm{i})\ (\mathrm{mod}\ 5)\,,$$

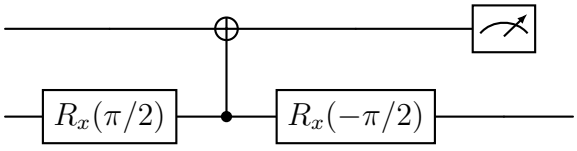
and the last step is verified by computation. As a result, the m obtained in the above subexcercise (1) cannot exist, as $5^m = 0\ (\mathrm{mod}\ 5)$.

Ex. 4.43. In exer. 4.41 we have shown how to obtain $R_z(\theta)$, $\cos\theta = 3/5$ using Hadamard, Toffoli and phase gates and measurements. As θ is an irrational multiple of π, any other angle may be approximated with powers (repeated application) of that circuit. In sec. 4.5.3, it was shown that Hadamard, phase, CNOT and $\pi/8$ are a universal set, so all that is to add is that the $\pi/8$ gate is approximated.

Ex. 4.44. If α is irrational, the $\mathrm{C}^2\mathrm{i}R_x(\pi\alpha)$ gate is universal, as it can approximate then any $R_x(\theta)$.

For $\theta = \pi$ we get the Toffoli. Setting one control bit to 1 yields the CNOT.

Setting $\theta = \pi/2$ and the control qubits to 1 yields a Pauli X. The circuit



provides a state $|-\rangle$ when the measurement result is 1 (up to a phase of i).

Inputting $|-\rangle$ on target bit, 1 on one control yields a Z rotation on the other control qubit.

We know that CNOTs and 2-level gates are universal, so we need to approximate 2-level gates. These can be obtained using $R_x(\theta)$ and $R_z(\theta)$ gates. (See Ref. [9].)

Ex. 4.45. The Hadamard gate has a matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

the matrix of the S gate is

$$S = \begin{pmatrix} 1 & \\ & \mathrm{i} \end{pmatrix},$$

the matrix of the CNOT is

$$\mathrm{CNOT} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & & 1 \\ & & 1 & \end{pmatrix},$$

and that of the Toffoli gate is

$$\mathrm{Toffoli} = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & & 1 \\ & & & & & & 1 & \end{pmatrix},$$

which are all matrices with either integer elements or integer times $1/\sqrt{2}$, so when multiplying such matrices, a matrix of the form $2^{-k/2}M$ results, where every Hadamard matrix contributes 1 to k.

The $\pi/8$ gate has a matrix

$$T = \begin{pmatrix} 1 & \\ & \mathrm{e}^{\mathrm{i}\pi/4} \end{pmatrix} = \begin{pmatrix} 1 & \\ & (1+\mathrm{i})/\sqrt{2} \end{pmatrix},$$

so in this case, M may also include integer times $\sqrt{2}$ elements.

4.6 Simulation of quantum systems

Ex. 4.46. The state vector of a single qubit has 2 complex components, for n qubits, a tensor product is taken, so the state vector has 2^n (complex) components. The density matrix (an operator acting on this space) is therefore a complex $2^n \times 2^n$ matrix. This matrix is self-adjoint, so the main diagonal must be real, and the elements above the diagonal are complex conjugates of the ones below it, and there is one constraint (total probability of 1) that the trace is 1, so we have $4^n - 1$ real parameters.

Ex. 4.47. Using the fact that the operators H_k commute, the terms in the series of the exponential may be reorganised as

$$\begin{aligned}
\mathrm{e}^{-\mathrm{i}Ht} &= \sum_{n=0}^{\infty} \frac{(-\mathrm{i}Ht)^n}{n!} = \sum_n \frac{(-\mathrm{i}t\sum_k H_k)^n}{n!} \\
&= \sum_{n,m_1,m_2,\ldots,m_{k-1}} \frac{(-\mathrm{i}t)^n}{n!} \frac{n!}{(n-m_1-m_2-\ldots)!m_1!m_2!\ldots} H_1^{m_1} H_2^{m_2} \ldots \\
&= \sum_{m_1,m_2,\ldots,m_k} \frac{(-\mathrm{i}t)^{\sum_i m_i}}{m_1!m_2!\ldots m_k!} H_1^{m_1} H_2^{m_2} \ldots H_k^{m_k} \\
&= \left(\sum_{m_1} \frac{(-\mathrm{i}H_1 t)^{m_1}}{m_1!}\right)\left(\sum_{m_2} \frac{(-\mathrm{i}H_2 t)^{m_2}}{m_2!}\right)\ldots\left(\sum_{m_k} \frac{(-\mathrm{i}H_k t)^{m_k}}{m_k!}\right) \\
&= \mathrm{e}^{-\mathrm{i}H_1 t}\mathrm{e}^{-\mathrm{i}H_2 t}\ldots\mathrm{e}^{-\mathrm{i}H_k t}\,.
\end{aligned}$$

Ex. 4.48. If H_k involves a maximum of c particles out of n then the different H_k terms are given as

$$\binom{n}{c} = \frac{n(n-1)\ldots(n-c)}{c!} = O(n^{c+1})\,.$$

Ex. 4.49. Use the power series of the exponential, e.g.,

$$\mathrm{e}^{A\Delta t} = 1 + A\Delta t + \frac{1}{2}(A\Delta t)^2 + O(\Delta t^3)\,,$$

so

$$\mathrm{e}^{(A+B)\Delta t} = I + A\Delta t + B\Delta t + \frac{1}{2}(A^2 + AB + BA + B^2)(\Delta t)^2 + O(\Delta t^3)\,,$$

and the other side of the equation is

$$\begin{aligned}
&e^{A\Delta t}e^{B\delta t}\mathrm{e}^{-\frac{1}{2}[A,B](\Delta t)^2} \\
&\quad = (I + A\Delta t + \frac{1}{2}A^2(\Delta t)^2)(I + B\Delta t + \frac{1}{2}B^2(\Delta t)^2)(I - \frac{1}{2}[A,B](\Delta t)^2) + O(\Delta t^3) \\
&\quad = I + (A+B)\Delta t + \left(AB + \frac{1}{2}(A^2 + B^2 - [A,B]\right)(\Delta t)^2 + O(\Delta t^3) \\
&\quad = I + (A+B)\Delta t + \frac{1}{2}\left(A^2 + AB + BA + B^2\right)(\Delta t)^2 + O(\Delta t^3)\,,
\end{aligned}$$

and this agrees with the expansion of $e^{(A+B)\Delta t}$ above.

Similarly, to prove eq. (4.103), on the left hand side we have

$$e^{\mathrm{i}(A+B)\Delta t} = I + \mathrm{i}(A+B)\Delta t + O(\Delta t^2)\,,$$

and on the right

$$e^{\mathrm{i}A\Delta t}\mathrm{e}^{\mathrm{i}B\Delta t} = (I + \mathrm{i}A\Delta t)(I + \mathrm{i}B\Delta t) + O(\Delta t^2) = I + \mathrm{i}(A+B)\Delta t + O(\Delta t^2)\,,$$

and the two agree.

In the case of eq. (4.104), the left hand side is

$$\mathrm{e}^{\mathrm{i}(A+B)\Delta t} = I + \mathrm{i}(A+B)\Delta t - \frac{1}{2}(A^2 + AB + BA + B^2)(\Delta t)^2 + O(\Delta t^3)\,,$$

and the right one is

$$
\begin{aligned}
&\mathrm{e}^{\mathrm{i}A\Delta t/2}\mathrm{e}^{\mathrm{i}B\Delta t}\mathrm{e}^{\mathrm{i}A\Delta t/2} \\
&\quad = (I + \frac{\mathrm{i}}{2}A\Delta t - \frac{1}{8}A^2(\Delta t)^2)(I + \mathrm{i}B\Delta t - \frac{1}{2}B^2(\Delta t)^2)(I + \frac{\mathrm{i}}{2}A\Delta t - \frac{1}{8}A^2(\Delta t)^2) + O(\Delta t^3) \\
&\quad = I + \mathrm{i}(A+B)\Delta t - \frac{1}{2}\left(A^2 + AB + BA + B^2\right)(\Delta t)^2 + O(\Delta t^3)\,.
\end{aligned}
$$

Ex. 4.50. (a) With $H = \sum_{k=1}^{L} H_k$,

$$
\begin{aligned}
&U_{\Delta t} \\
&= \left[\mathrm{e}^{-\mathrm{i}H_1\Delta t}\mathrm{e}^{-\mathrm{i}H_2\Delta t}\dots\mathrm{e}^{-\mathrm{i}H_L\Delta t}\right]\left[\mathrm{e}^{-\mathrm{i}H_L\Delta t}\mathrm{e}^{-\mathrm{i}H_{L-1}\Delta t}\dots\mathrm{e}^{-\mathrm{i}H_1\Delta t}\right] \\
&= \left[\left(I - \mathrm{i}H_1\Delta t - \frac{1}{2}H_1^2(\Delta t)^2\right)\left(I - \mathrm{i}H_2\Delta t - \frac{1}{2}H_2^2(\Delta t)^2\right)\dots\left(I - \mathrm{i}H_L\Delta t - \frac{1}{2}H_L^2(\Delta t)^2\right)\right] \\
&\left[\left(I - \mathrm{i}H_L\Delta t - \frac{1}{2}H_L^2(\Delta t)^2\right)\left(I - \mathrm{i}H_{L-1}\Delta t - \frac{1}{2}H_{L-1}^2(\Delta t)^2\right)\dots\left(I - \mathrm{i}H_1\Delta t - \frac{1}{2}H_1^2(\Delta t)^2\right)\right] \\
&+ O(\Delta t^2) \\
&= I - 2\mathrm{i}H\Delta t - \left(2\sum_k H_k^2 + \sum_{k\neq j}(H_kH_j + H_jH_k)\right) + O(\Delta t^3) \\
&= \mathrm{e}^{-2\mathrm{i}H\Delta t} + O(\Delta t^3)\,.
\end{aligned}
$$

(b) Let us note, that

$$E(U_{\Delta t}^m, \mathrm{e}^{-2\mathrm{i}mH\Delta t}) = \|U_{\Delta t}^m - (\mathrm{e}^{-2\mathrm{i}H\Delta t})^m\|\,,$$

and for any operators A, B with $C = A - B$, we have $A^m - B^m = (B+C)^m - B^m = B^m + CB^{m-1} + BCB^{m-2} + \dots - B^m$, and so $\|A^m - B^m\| = \|CB^{m-1} + BCB^{m-1} + \dots\| \leq \|CB^{m-1}\| + \|BCB^{m-2}\| + \dots \leq m\|B\|^{m-1}\|C\|$, so

$$E(U_{\Delta t}^m, \mathrm{e}^{-2\mathrm{i}mH\Delta t}) \leq m\|U_{\Delta t}^{m-1}\|E(U_{\Delta t}, \mathrm{e}^{-2\mathrm{i}H\Delta t})$$

and from (a) we know (the definition of O) that there is a number α' such that $\|U_{\Delta t} - \mathrm{e}^{-2\mathrm{i}H\Delta t}\| \leq \alpha'\Delta t^3$, so

$$E(U_{\Delta t}^m, \mathrm{e}^{-2\mathrm{i}mH\Delta t}) \leq m\|U_{\Delta t}\|^{m-1}\alpha'\Delta t^2 = m\alpha\Delta t^3$$

with $\alpha = \|U_{\Delta t}\|^{m-1}\alpha'$. As $U_{\Delta t}$ is unitary, $\|U_{\Delta t}\| = 1$.

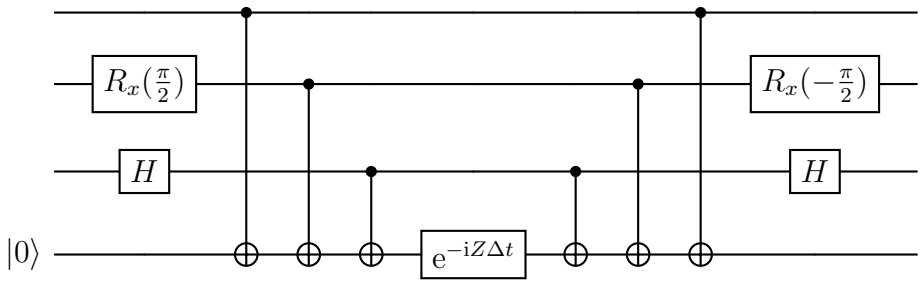
Ex. 4.51. We may express the Pauli matrices X, Y with Z and single-qubit gates as follows:

$$Y = R_x(-\pi/2)^\dagger Z R_x(-\pi/2)\,,$$

and

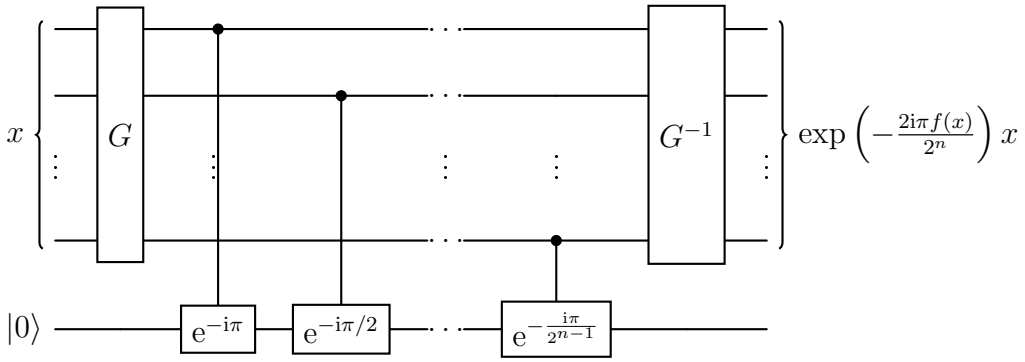
$$X = HZH\,,$$

and then we apply the techniques used for the Hamiltonian in eq. (4.113),



We have used the fact that pairs of the form $UU^\dagger$ cancel.

Pr. 4.1. Let G be the gate computing the function f reversibly using T Toffoli gates, then the circuit



implements the mapping using $2T + n$ Toffoli and controlled phase shift gates.

5 The quantum Fourier transform and its applications

5.1 The quantum Fourier transform

Ex. 5.1. The operator defined by eq. (5.3) has the matrix

$$F_{kj} = \frac{1}{\sqrt{N}} \mathrm{e}^{2\pi \mathrm{i} kj/N} ,$$

so what we need to show is $F^\dagger F = I$, and the matrix elements of $F^\dagger F$ are

$$\sum_k F^*_{k\ell} F_{kj} = \frac{1}{N} \sum_{k=0}^{N-1} \mathrm{e}^{2\pi \mathrm{i} k(j-\ell)/N} = \delta_{\ell,j} ,$$

which can be seen as follows: if $\ell \neq j$, the sum is a geometric series, and the explicit summation formula gives zero due to the periodicity of $\mathrm{e}^{2\pi \mathrm{i} k/N}$ with a period N, and at $\ell = j$, all summands are 1.

Ex. 5.2. The Fourier transform of the state $|00\ldots0\rangle$ is

$$\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle = \frac{1}{\sqrt{N}} \left[|00\ldots0\rangle + |0\ldots01\rangle + \ldots |11\ldots1\rangle \right] ,$$

where $N = 2^n$ and n is the number of qubits, which is obtained by setting $j = 0$ in eq. (5.2).

Ex. 5.3. Eq. (5.1), in the case of $N = 2^n$ requires the computation of 2^n sums, each sum containing N summands, so the total number of operations is $N^2 = 2^{2n}$.

The number of operations may be reduced using the trick of eq. (5.4). What is given there is the contribution of $x_{j_1 j_2 \ldots j_n}$ to $y_{k_1 k_2 \ldots k_n}$, and we may note that as $\exp(2\pi \mathrm{i} 0.j_n) = \pm 1$, so the contributions to $y_{0 k_2 \ldots k_n}$ and $y_{1 k_2 \ldots k_n}$ are either the same if $j_1 = 0$ or differ by a sign if $j_1 = 1$. This can be used to split the transform in a half, and compute the steps separately. This way, in each step, the number of bits of the output index are reduced by one. The input indices, over which a loop is still needed, still range over n bit numbers, so there will be 2^n steps, so the total number of steps is $n2^n$.

Ex. 5.4. We follow the procedure of Sec. 4.3. We first decompose the R-gate as $R = \mathrm{e}^{\mathrm{i}\alpha} AXBXCX$ where $ABC = I$, as follows,

$$R_k = \begin{pmatrix} 1 & \\ & \mathrm{e}^{2\pi\mathrm{i}/2^k} \end{pmatrix}$$

which when compared with eq. (4.12) yields $\gamma = 0$, and $\alpha = \pi/2^k$, $\beta = 2\pi/2^k$ and $\delta = 0$ (note that $\alpha - \beta/2 = 0$ and $\alpha + \beta/2 = 2\pi/2^k$), yielding $A = R_z(2\pi/2^k)$, $B = R_z(-\pi/2^k)$ and $C = R_z(-\pi/2^k)$. We may now use the circuit in Fig. 4.6.

Ex. 5.5. One possibility for the inverse-FT is to put the gates in reverse order, and replace them by their adjoints. The adjoint of a controlled gate is the controlled version of the adjoint, $R_k^\dagger$ just shifts with the negative phase, and $H^\dagger = H$.

Another possibility is based on the following property of the classical Fourier-transformation: the inverse is given by the same formula as the original transformation just with the phases reversed. So, the mapping

$$|k\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{\ell=0}^{N-1} \mathrm{e}^{-\mathrm{i}2\pi(k-\ell)/N} |\ell\rangle$$

is the inverse, as in this case, with this transformation applied after the direct one,

$$|j\rangle \mapsto \frac{1}{N} \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \mathrm{e}^{2\pi\mathrm{i}(j-\ell)k/N} |\ell\rangle = |j\rangle\,,$$

where we have used the fact that the same sum has been evaluated in the case of the classical Fourier transformation, see ex. 5.1

As the formula above is the same as for the direct Fourier transform, with the phases reversed, the circuit implementing it is also the same, just replacing the controlled S-gates with controlled $S^\dagger$-gates (phase shift gates with negative phase).

Ex. 5.6. According to the calculation in box 4.1, if operators in a product of operators are approximated, the errors add up linearly. The depth of the circuit implementing the quantum Fourier transform scales as n^2, where n is the number of qubits, so if the error of each operator scales as $1/p(n)$, the total error scales as $n^2/p(n)$.

5.2 Phase estimation

Ex. 5.7. Let us apply the operator implemented by the circuit in fig. 5.2 to a state of the form $|j\rangle \otimes \psi$, where $j = j_{t-1}2^{t-1} + \cdots + j_1 2 + j_0$. If $j_k = 1$, an operator U^{2^k} is applied to ψ, so the result is

$$|j\rangle \otimes \prod_{k:j_k=1} U^{2^k} \psi = |j\rangle \otimes U^{\sum_k j_k 2^k} \psi = |j\rangle \otimes U^j \psi\,.$$

Ex. 5.8. In the case of an initial state which is a superposition of the eigenstates of U, the derivation in sec. 5.2.1 may be repeated, but in eq. (5.23), the there is also a summation $\sum_u c_u$ in the front and in the exponent, ϕ is replaced by φ_u, u running over the eigenvalues of U.

The formula in eq. (5.27) can now be used to obtain the formula for $P(|m-b| > e|u)$, i.e., the conditional probability, and then the total probability of such a "wrong" result is

$$p(|m-b| > e) = \sum_{u'} P(|m-b| > r|u')P_{u'} \geq P_(|m-b| > e|u)P_u + \sum_{u' \neq u} P|u'|\,,$$

the rest of the derivation yields

$$P(|m-b|>e|u)\leq 1-\varepsilon\,,$$

and so

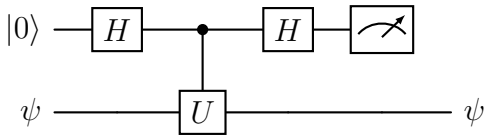
$$p(|m-b|>e)\leq |c_u|^2\varepsilon+\sum_{u'\neq u}|c_u|^2\,,$$

so the probability

$$p(|m-b|\leq e)\geq 1-|c_u|^2\varepsilon-\sum_{u'\neq u}|c_{u'}|^2=|c_u|^2(1-\varepsilon)\,,$$

where we have used $\sum_{u'}|c_{u'}|^2=1$.

Ex. 5.9. In this case, the phases to be measured, are 0 and π, corresponding to the eigenvalues 1 and -1, so they can be both exectly represented in the form $2\pi 0.b$ with $b=0,1$, as binary fractions. This means that in the circuit in fig. 5.2-5.3 we need one bit, $t=1$, no swap, and the quantum Fourier transform for 1 bit is the Hadamard operator, so the resulting circuit looks like



Note that the circuit is the same as the one considered in ex. 4.34

5.3 Applications: order-finding and factoring

Ex. 5.10. The powers of 5 mod 21 are: 5, 4, 20, 16, 17, 1, so the order of 5 is 6.

Ex. 5.11. According to Fermat's theorem, the order r of any x such that $\gcd(x,N)=1$ is a divisor of $\varphi(n)=\prod_j p_j^{\alpha_j-1}(p_j-1)$ where the prime factorisation of N is $N=\prod_j p_j^{\alpha_j}$, so on one hand, $r|\varphi(N)$, so $r\leq\varphi(N)$, and comparing the formula for $\varphi(N)$ with the factorisation of N, $\varphi(N)<N$.

Ex. 5.12. As $\gcd(x,N)=1$, the operator defined by eq. (5.36) is a permutation of the basis vectors (it maps basis vectors onto basis vectors and has an inverse), therefore it preserves scalar products.

Ex. 5.13. Substituting eq. (5.37) into eq. (5.44),

$$\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|u_s\rangle=\frac{1}{r}\sum_{k,s=0}^{r-1}\exp\left(-\frac{2\pi\mathrm{i}sk}{r}\right)|x^k\,(\mathrm{mod}\ N)\rangle\,,$$

and in the summation, when the summation over s is performed, the sum calculated is a geometric series. When $k\neq 0$, the result vanishes due to periodicity of $\mathrm{e}^{2\pi\mathrm{i}m}$, and when $k=0$, the exponent is 1, and the sum is r terms, each 1, canceling the $1/r$, and noting that $x^0=1$, the proof is complete.

The same argument proves eq. (5.45). In this case, the summation index shall be k', and the summation over s gives a $r\delta_{kk'}$.

Ex. 5.14. Let us remember the derivation in the solution of ex. 5.7. We apply $H^{\otimes t}$ to the register holding j, so the circuit implements the mappings

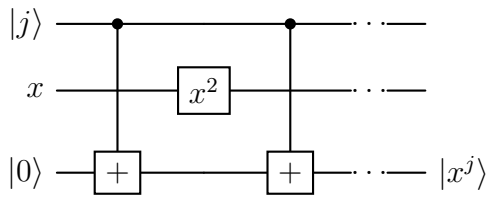
$$|0\rangle \otimes |0\rangle \xmapsto{H^{\otimes t}\otimes I} \frac{1}{2^{t/2}}(|0\rangle + |1\rangle)^{\otimes t} \otimes |0\rangle = \frac{1}{2^{t/2}} \sum_j |j\rangle \otimes |0\rangle$$

$$\xmapsto{V} \frac{1}{2^{t/2}} \sum_j |j\rangle |x^j \pmod N\rangle \,.$$

To implement this V usign L^3 gates, we first need a circuit that implements the mapping

$$|y\rangle \mapsto |y + x^{2^k} \pmod N\rangle$$

which we do as in box 5.2, using modular exponentiation, that requires $O(L)$ squaring operations (implemented as in reversible computing, $O(L^2)$ gates), and in each step we need a controlled addition, and there are L steps, so the total cost is $O(L^2)$. The circuit is as follows:



Ex. 5.15. Let us write the prime factorisation of the two integers as

$$x = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k} \,, \quad y = p_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k} \,.$$

The greatest common divisor of the two can be written as

$$\gcd(x,y) = \prod_i p_i^{\min\{\alpha_i,\beta_i\}} \,,$$

as this divides both, and any prime on any higher power does not divide the one with the lowest power. Similarly,

$$\mathrm{lcm}(x,y) = \prod_i p_i^{\max\{\alpha_i,\beta_i\}} \,,$$

and

$$xy = \prod_i p_i^{\alpha_i+\beta_i} \,.$$

Ex. 5.16. The integral is evaluated as follows,

$$I_x = \int_x^{x+1} \frac{1}{y^2} \mathrm{d}y = \left[-\frac{1}{y}\right]_x^{x+1} = \frac{1}{x} - \frac{1}{x+1} = \frac{1}{x(x+1)} \,,$$

and so

$$I_x - \frac{2}{3x^2} = \frac{x}{x^2(x+1)} - \frac{2(x+1)}{3x^2(x+1)} = \frac{x/3 - 2/3}{x^2(x+1)} \,,$$

which is positive is $x > 2/3$. (The denominator is positive, the numerator is an increasing function, and vanishes at $x = 2$).

Now we may write $1/x^2 \leq 3I_x/2$, and so

$$\sum_{q\,\text{prime}} \frac{1}{q^2} < \sum_{q=2}^{\infty} \frac{1}{q^2} \leq \frac{3}{2} \int_2^{\infty} \frac{1}{y^2} \mathrm{d}y = \frac{3}{2} \left[-\frac{1}{y}\right]_2^{\infty} = \frac{3}{4} \,.$$

Ex. 5.17. (1) If N is L bit long, then $0 \le N \le 2^L - 1$, so if $N = a^b$, and $a \ge 2$, then from $a^L > 2^L - 1$ follows that $b < L$ must hold.

(2) $y = \log_2 N$ is calculated with the following algorothm [12]: a, we divide N by 2 enough times so that the result is between 1 and 2. If it is one, we are done, in $O(L)$ steps. If not, precision is enhanced, by squaring y a couple of times till the result is between 2 and 4, and dividing that by 2, each time getting some additional bit of the fractional part of y, in the form $0 \ldots 01$ after the previous digits, the number of zeros is the number of squarings necessary.

The number $x = y/b$ is calculated using the usual method of division on paper. We take some bits from the beginning of y, until the number made up by those is above b, write that divided by b (know small multiples of b, $O(L)$ steps as $b < L$) to the beginning of x, and replace it at the beginning of y with the remainder. As the length of y is reduced in each step, the number of steps needed is $O(L)$ (length of N), each step required $O(L)$ operations, so the total is $O(L^2)$.

The integer part of 2^x may be calculated by multiplications by 2, as N has L bits, the integer part of x is $O(L)$. Now we need to enhance the result by calculating $2^{x'}$ where x' is the fractional part of x to the precision where $1/2^{[x]}$. This can be done using Taylor-series, the number of terms needed is $O(L)$ because the precision of the kth term is $1/k!$, $\log k!\ k \log k$, we need $2^L \sim k!$, so $k < O(L)$, and each term consist of L-bit multiplications.

(3) With repeated squaring we may compute u_1^b, where m is a power of 3 close to b, and then add the remining few multiplications. We multiply numbers of length L, and the number of squarings is $\log_2 b \sim \log_2 L$. Using, e.g., Karatsuba multiplication, this is below $O(L^2)$.

(4) As $b < L$, we loop over b, multiplying the number of steps by L.

Ex. 5.18. Factoring $N = 91$. The steps are as follows:

1. N is odd, proceed to step 2.
2. We need to check up to $b = \log_2 91 \approx 6.5$. See ex. 5.17.
3. We choose the "random" number x=4. A quick calculation using Euclid's algorithm yields $\gcd(4, 91) = 1$, so it is co-prime, we may proceed to order-finding.
4. The powers of 4 mod 91 are 1, 4, 16, 64, 74, 23, 1, so the order is $r = 6$. $x^3 = 64 \neq -1 \pmod{91}$, we may proceed to gcd.
5. $\gcd(63, 91) = 7$, a factor is found.

Ex. 5.19. The numbers below 15 are: 2 even, 3 prime, $4 = 2^2$ and even, 5 prime, 6 even, 7 prime, $8 = 2^3$ and even, $9 = 3^2$, 10 even, 11 prime, 12 even, 13 prime, 14 even.

5.4 General applications of the quantum Fourier transform

Ex. 5.20. Let us use the periodicity of f as follows:

$$\hat{f}(\ell) = \frac{1}{\sqrt{nr}} \sum_{x=0}^{N-1} \mathrm{e}^{-2\pi \mathrm{i} \ell x/N} f(x) = \frac{1}{\sqrt{N}} \sum_{m=0}^{n-1} \sum_{x'=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i} m \ell/n} \mathrm{e}^{-2\pi \mathrm{i} x' \ell/N} f(k)$$
$$= \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} \mathrm{e}^{-2\pi \mathrm{i} \ell m/n} \frac{1}{\sqrt{r}} \sum_{x'=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i} \ell x'/N} f(x') \,.$$

where $N = nr$ and we have written $x = mr + x'$. The first sum is the sum of a geometric series, and, unless $n|\ell m$, the sum formula shows that it vanishes. $n|\ell m$ for all m if $n|\ell$, i.e., ℓ is an integer multiple of $n = N/r$, and then all the summands are 1, and the first sum contributes $n = N/r$, and the result is

$$\hat{f}(\ell' n) = \frac{N}{r}\frac{1}{\sqrt{r}}\sum_{x'=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i}\ell' x'/r} f(x') = \frac{N}{r}\tilde{f}(\ell')\,,$$

where $\tilde{f}$ is the Fourier transform of f on the shorter interval $0 \le x \le r-1$, or we may write

$$\hat{f}(\ell) = \begin{cases} \dfrac{N}{r}\tilde{f}\left(\dfrac{\ell}{N/r}\right), \text{ if } N/r|\ell\,, \\ 0 \qquad\qquad \text{otherwise}\,. \end{cases}$$

In eq. (5.63) we use the inverse Fourier transform to express $f(x)$, using a single period of f. Note however that if x is outside the single period, $x = mr + x'$ then $f(x) = f(x')$, and substituting into eq. (5.64) yields

$$|f(x)\rangle = \frac{1}{\sqrt{r}}\sum_{\ell=0}^{r-1} \mathrm{e}^{2\pi \mathrm{i}\ell x/r}|\hat{f}(\ell)\rangle = \frac{1}{r}\sum_{\ell=0}^{r-1} \mathrm{e}^{2\pi \mathrm{i}\ell x'/r}|\hat{f}(\ell)\rangle = |f(x')\rangle\,,$$

as $\exp(2\pi \mathrm{i}\ell x/r) = \exp[2\pi \mathrm{i}(\ell x'/r + \ell m] = \exp(2\pi \mathrm{i}\ell x'/r)$, i.e., the periodicity is recovered.

Ex. 5.21. (1) Let us apply U_y to the state $|\hat{f}(\ell)\rangle$ defined in eq. (5.63),

$$\begin{aligned} U_y|\hat{f}(\ell)\rangle &= \frac{1}{\sqrt{r}}\sum_{x=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i}\ell x/r} U_y|f(x)\rangle = \frac{1}{\sqrt{r}}\sum_{x=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i}\ell x/r}|f(x+y)\rangle \\ &= \frac{1}{\sqrt{r}}\sum_{x'=y}^{y+r-1} \mathrm{e}^{-2\pi \mathrm{i}\ell(x'-y)/r}|f(x)\rangle = \mathrm{e}^{2\pi \mathrm{i}\ell y/r}\frac{1}{\sqrt{r}}\sum_{x=0}^{r-1} \mathrm{e}^{-2\pi \mathrm{i}\ell x/r}|f(x)\rangle \\ &= \mathrm{e}^{2\pi \mathrm{i}\ell y/r}|\hat{f}(\ell)\rangle\,. \end{aligned}$$

where we have shifted the summation variable, $x' = x + y$, used the fact that $\exp(2\pi \mathrm{i}\ell x/r)$ and $f(x)$ are both periodic, so the sum overlaps, and dropped the prime. The eigenvalue is $\lambda_\ell = \exp(2\pi \mathrm{i}\ell y/r)$.

(2) Let us express the given state $|f(x_0)\rangle$ with the Fourier transform as

$$|f(x_0)\rangle = \frac{1}{\sqrt{r}}\sum_{\ell=0}^{r-1} \mathrm{e}^{2\pi \mathrm{i}\ell x_0/r}|\hat{f}(\ell)\rangle\,,$$

and apply the black box part of the quantum phase estimation algorithm, implementing the mapping

$$|j\rangle \otimes \psi \mapsto |j\rangle \otimes U_y^j \psi\,.$$

At this point, the input to this mapping is

$$\frac{1}{\sqrt{N}}\sum_j |j\rangle|f(x_0)\rangle\,,$$

with $N = 2^t$, so the output is

$$\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle U_y^j|f(x_0)\rangle = \sum_{j=0}^{N-1}\sum_{\ell=0}^{r-1}\frac{1}{\sqrt{Nr}}\mathrm{e}^{2\pi\mathrm{i}\ell x_0/r}|j\rangle U_y^j|\hat{f}(\ell)\rangle = \frac{1}{\sqrt{Nr}}\sum_{j=0}^{N-1}\sum_{\ell=0}^{r-1}\mathrm{e}^{2\pi\mathrm{i}j\ell/r}\lambda_\ell^j|j\rangle|\hat{f}(\ell)\rangle\,,$$
$$= \frac{1}{\sqrt{Nr}}\sum_{j=0}^{N-1}\sum_{\ell=0}^{r-1}\mathrm{e}^{2\pi\mathrm{i}j\ell/r}\mathrm{e}^{2\pi\mathrm{i}\ell jx_0/r}|j\rangle|\hat{f}(\ell)\rangle = \frac{1}{\sqrt{r}}\sum_{\ell=0}^{r-1}\mathrm{e}^{2\pi\mathrm{i}j\ell/r}|\widetilde{N\ell x_0/r}\rangle|\hat{f}(\ell)\rangle\,.$$

Applying the inverse Fourier transform to the first register allows us to measure $N\ell x_0/r$, and using the continued fraction method, therefore, r.

Ex. 5.22. Using the double periodicity of f,

$$|\hat{f}(\ell_1,\ell_2)\rangle = \sum_{x_1=0}^{r-1}\sum_{x_2=0}^{r-1}\mathrm{e}^{-2\pi\mathrm{i}(\ell_1x_1+\ell_2x_2)/r}|f(x_1,x_2)\rangle$$
$$= \sum_{x_1=0}^{r-1}\sum_{x'=sx_1}^{sx_1+r-1}\mathrm{e}^{-2\pi\mathrm{i}[\ell_1x_1+\ell_2(x'-sx_1)]/r}|f(x_1,x'-\ell x_1)\rangle\,,$$

where we have introduced the new summation variable $x' = sx_1+x_2$. Note that $f(x_1,x'-\ell x_1) = f(0,x')$. Also, as $f(x_1,x_2) = b^{x_1}a^{x_2}$ and $a^r = 1\,(\mathrm{mod}\,r)$, and the exponential is also periodic with period r, the summation variable may run from 0 to $r-1$, yielding

$$|\hat{f}(\ell_1,\ell_2) = \sum_{x_1=0}^{r-1}\sum_{x_2=0}^{r-1}\mathrm{e}^{-2\pi\mathrm{i}[\ell_1x_1+\ell_2(x_2-sx_1)]/r}|f(0,x_2)\rangle\,,$$

and this sum vanishes unless $\ell_1 - s\ell_2$ is an integer multiple of r, as in the case where it is not, the sum over x_1 may be performed using the summation formula for a geometric progression, yielding zero. If $\ell_1 - s\ell_2$ is an integer multiple of r, the sum over x_1 yields r, and the result is

$$|\hat{f}(\ell_1,\ell_2)\rangle = r\sum_{j=0}^{r}\mathrm{e}^{-2\pi\mathrm{i}j\ell_2/r}|f(0,j)\rangle\,.$$

Ex. 5.23. Inserting formula (5.72) into (5.73), we need to evaluate four sums, of which the one over ℓ_1, collecting the terms only depending on ℓ_1 is, and replacing the summation variable x_1 with x_1' in eq. (5.72), is

$$\sum_{\ell_1=0}^{r-1}\mathrm{e}^{2\pi\mathrm{i}\ell_1(x_1-x_1')/r} = r\delta_{x_1,x_1'}\,,$$

and similarly the summation over ℓ_2 yields another Kronecker delta, so the resulting formula is

$$\frac{1}{r^2}\sum_{\ell_1=0}^{r-1}\sum_{\ell_2=0}^{r-1}\mathrm{e}^{2\pi\mathrm{i}(\ell_1x_1+\ell_2x_2)/r}|\hat{f}(\ell_1,\ell_2)\rangle = |f(x_1,x_2)\rangle\,.$$

Ex. 5.24. In order to obtain both $\ell_2 s/r$ and ℓ_2/r, and to apply thm. 5.1, one needs to choose the number of bits such that $2^{-2L+1} \leq 1/(2r^2)$. In that case, s is obtained by dividing the two.

Ex. 5.25. One possibility is to construct using reversible computation the functions, and then uncomputation. The necessary number of gates is, for L bits, $O(L)$ squarings to calculate one exponentiation, one squaring is $O(L^2)$, then $O(L)$ multiplications, again $O(L^2)$ cost, so a total of $O(L^3)$ is needed.

Ex. 5.26. Let $G = Z_{p_1} \times Z_{p_2} \times \dots Z_{p_N}$.

Any irreducible unitary representation of the cyclic group Z_p is 1 dimensional, and so on the generating element has order p, and its representation is a phase, $\rho_\ell(1) = \mathrm{e}^{\mathrm{i}\phi_{\ell,0}}$, such that $\rho_\ell(0) = \rho_\ell(1)^p = \mathrm{e}^{\mathrm{i}p\phi_{\ell,0}} = 1$, i.e., $\phi_{\ell,0} = 2\pi\ell/p$, so the representation of an arbitrary element is

$$\rho_\ell(g) = \mathrm{e}^{2\pi\mathrm{i}g\ell/p}\,,$$

$0 \le g < p$.

For G, a group element is given as $G \ni g = (g_1, g_2, \dots, g_N)$ where $0 \le g_i < p_i$. The representations are indexed by $\ell_1, \dots, \ell_N$, so

$$\rho_{\ell_1,\dots,\ell_N}(g_1, \dots, g_N) = \mathrm{e}^{(2\pi\mathrm{i}/|G|)\sum_i g_i\ell_i}\,,$$

where $|G| = p_1 \cdot \dots \cdot p_N$. The Fourier and inverse Fourier transform formulae are for such groups

$$\hat{f}(\ell) = \frac{1}{\sqrt{|G|}} \sum_{g_i=0}^{p_i-1} f(g)\mathrm{e}^{(2\pi\mathrm{i}/|G|)\sum_i g_i\ell_i}\,,$$

[see eq. (A2.9)] and

$$f(g) = \frac{1}{|G|} \sum_{\ell_i=0}^{p_i-1} \hat{f}(\ell_1, \dots, \ell_N)\mathrm{e}^{-(2\pi\mathrm{i}/|G|)\sum_i g_i\ell_i}\,,$$

[see eq. (A2.10)]. Comparing this with eq. (5.75), we can see that Nielsen and Chuang have used $g\ell$ as the short form of $\sum_i g_i\ell_i$ where $g = (g_1, \dots, g_N)$.

If K is a normal subgroup of G it is of the form $K = Z_{p_{i_1}} \times \dots \times Z_{p_{i_M}}$ where $M \le N$ and $1 \le i_k \le N$, and $\{p_{i_1}, p_{i_2}, \dots, p_{i_M}\} \subseteq \{p_1, p_2, \dots p_N\}$. In this case, the sum in eq. (5.76) may be rewritten as

$$|\hat{f}(\ell)\rangle = \frac{1}{|G|} \sum_{j\neq i_k, g_j=0}^{p_j-1} \mathrm{e}^{-(2\pi\mathrm{i}/|G|)\sum_j g_j\ell_j} |f(g_j, j \neq i_k)\rangle \sum_{g_{i_k}=0}^{p_k-1} \mathrm{e}^{-(2\pi\mathrm{i}/|G|)\sum_k g_{i_k}\ell_{i_k}}\,,$$

as f being constant on cosets of K is equivalent to f being independent of g_{i_k}, $k = 1, \dots, M$. The second sum in the last equation vanishes unless $\ell_{i_k} = 0$, where these are defined by that for these values these sums evaluate to $|K|$, as due to the orthogonality relation of characters and the fact that the character of the trivial representation is 1, so the result is

$$|\hat{f}(\ell : \ell_{i_k} = 0)\rangle = \frac{|K|}{|G|} \sum_{j\neq i_k, g_j=0}^{p_j-1} \mathrm{e}^{-(2\pi\mathrm{i}/|G|)\sum_j g_j\ell_j} |f(g_j, j \neq i_k)\rangle\,.$$

By measuring ℓ_k, the subgroup K is the one defined by i_k such that $\ell_{i_k} = 0$, $K = \prod_k Z_{p_{i_k}}$.

Ex. 5.27. This is Mosca's algorithm [13]. We shall start with a set of generators $\{a_1, \ldots, a_k\} \subset G$. The aim is to find new generators, $\{b_1, \ldots, b_\ell\} \subset G$, such that each of these generators alone generates a prime-order cyclic group.

The method is applying the known version of the hidden subgroup problem (ex. 5.26). To this end, first we replace the generators with ones that have a prime order. Let us assume that the order of one of the generators, a_i is pq, where $(p, q) = 1$ then $a^{pq} = 1$. In this case, according to the Euclidean algorithm, there are r, s such that $rp + sq = 1$, so $(a^p)^r (a^q)^s = a^{pq+rs} = a$. So, replacing a_i with a^p and a^q (and hence increasing the number of generators) does not change the generated group.

We shall now consider a mapping

$$g : \mathbb{Z}_q^k \to G\,, \quad (x_1, \ldots, x_k) \mapsto a_1^{x_1} \cdot \cdots \cdot a_k^{x_k}$$

where q is the maximum of the orders of the generators. The mapping g may be used as an input to the variant of the hidden subgroup problem where the group is already assumed to be a subgroup of a product of prime-order cyclic subgroups, yielding a subgroup K. A set of generators $y_1, \ldots, y_\ell \in \mathbb{Z}_q^k/K$ can also be computed, and so $\{g(y_i)|i = 1, \ldots, \ell\}$ is a set of generators of G with the desired property.

A Appendices

A.1 Notes on probability theory

Ex. A.1.1. The proof of Bayes' rule starts with the definition of conditional probability,

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{p(y|x)p(x)}{p(y)} = p(y|x)\frac{p(x)}{p(y)}\,.$$

Ex. A.1.2. For a set of mutually exclusive events x,

$$p(y) = \sum_x p(x,y) = \sum_x p(y|x)p(x)\,,$$

where we have used the definitions of the conditional probabilities, $p(y|x) = p(x,y)/p(x)$.

Ex. A.1.3. Let us assume the contrary, that for all values x of X such that $p(x) > 0$, $x < \mathbf{E}(X)$. Let x_0 the maximum of all these (as we are concerned with discrete variables taking one of a finite set of values, this exists), then

$$\mathbf{E}(X) = \sum_x xp(x) \leq \sum_x x_0 p(x) = x_0 < \mathbf{E}(x)\,,$$

which is a contradiction.

Ex. A.1.4. Linearity:

$$\mathbf{E}(\alpha X) = \sum_x \alpha x p(x) = \alpha \sum_x xp(x) = \alpha \mathbf{E}(X)\,,$$

and

$$\begin{aligned}\mathbf{E}(X+Y) &= \sum_{x,y}(x+y)p(x,y) = \sum_{x,y} xp(x,y) + \sum_{x,y} yp(x,y) = \sum_x xp(x) + \sum_y yp(y) \\ &= \mathbf{E}(X) + \mathbf{E}(Y)\,.\end{aligned}$$

Ex. A.1.5. For independent variables X and Y,

$$\mathbf{E}(XY) = \sum_{x,y} xyp(x,y) = \sum_{x,y} xyp(x)p(y) = \sum_{x} xp(x) \sum_{y} yp(y) = \mathbf{E}(X)\mathbf{E}(Y)\,.$$

Ex. A.1.6. Chebyshev's inequality is proven as follows:

$$\begin{aligned}
\Delta^2(X) &= \mathbf{E}(|X - \mathbf{E}(X)|^2) = \sum_x |x - \mathbf{E}(X)|^2 p(x) \\
&= \sum_{x:|x-\mathbf{E}(X)|<\lambda\Delta(X)} |x - \mathbf{E}(X)|^2 p(x) + \sum_{x:|x-\mathbf{E}(X)|\geq\lambda\Delta(X)} |x - \mathbf{E}(X)|^2 p(x) \\
&\geq \sum_{x:|x-\mathbf{E}(X)|<\lambda\Delta(X)} |x - \mathbf{E}(X)|^2 p(x) + \sum_{x:|x-\mathbf{E}(X)|\geq\lambda\Delta(X)} \lambda^2\Delta^2(X) p(x) \\
&\geq \sum_{x:|x-\mathbf{E}(X)|\geq\lambda\Delta(X)} \lambda^2\Delta^2(X) p(x) = \lambda^2\Delta^2(X) \sum_{x:|x-\mathbf{E}(X)|\geq\lambda\Delta(X)} p(x)\,,
\end{aligned}$$

which, upon division by $\lambda^2\Delta^2(X)$ yields the desired result.

A.2 Group theory

Ex. A.2.1. Let us assume that there is $g \in G$ such that $g^r \neq e$ for any non-zero integer r. In this case, all g^r must be distinct (otherwise, if $g^{r_1} = g^{r_2}$, $r_1 \neq r_2$, then $g^{r_2-r_1} = e$), and this is an infinite set, and subset of G, which is not possible, $|G| < \infty$.

Ex. A.2.2. Let $H \subseteq G$, and $g_{1,2} \in G$. Then either $g_1 H = \{g_1 h | h \in H\}$ and $g_2 H$ are disjoint or they are equal, as if $g_1 H \cup g_2 H \neq$, then there is h_1, h_2 such that $g_1 h_1 = g_2 h_2$, so $g_1 h_1 h_2^{-1} = g_2$, so $g_2 H = g_1 h_2 h_2^{-1} H = g_1 H$ as $h_2 \in H$. Also $g \in gH$ for all $g \in G$, so the cosets form a division of the total group.

All cosets have the same number of elements as H, as if it were not so, there would be $h_1, h_2 \in H$, $h_1 \neq h_2$, $g \in G$ such that $gh_1 = gh_2$, but that is not possible, as in this case $gh_1 h_2^{-1} = g$, multiplied by g^{-1} yielding $h_1 h_2^{-1} = e$, so $h_1 = h_2$.

As a result, the cosets have the same number of elements, $|gH| = |H|$, so $|H|$ is a divisor of $|G|$.

Ex. A.2.3. The elements $\{1, g, g^2, \dots\}$ form a (cyclic) subgroup of G. The order of this subgroup is the order of g.

Ex. A.2.4. If $y \in G_x$, then there is a $g \in G$ such that $y = g^{-1}xg$. In this case $G_y = \{h^{-1}yh | h \in G\} = \{h^{-1}g^{-1}xgh | h \in G\} = G_x$ as all elements of g may be written as gh.

Ex. A.2.5. In an Abelian group, for any $g, x \in G$, $g^{-1}xg = g^{-1}gx = x$, so $G_x\{g^{-1}xg | g \in G\} = \{x\}$.

Ex. A.2.6. For any element $g \in G$, $\langle g \rangle = \{1, g, g^2, \dots\} \leq G$ is a subgroup, so if r is the order of g then it is a divisor of $|G|$. If the latter is prime, they must agree.

Ex. A.2.7. Let $G = \langle g \rangle$ be cyclic subgroup and $H \leq G$ a subgroup. In this case, all elements of $h \in H$ are of the form of $h = g^k$, so $H = \{e, g^{k_1}, g^{k_2}, \dots\}$. If $H \neq \{1\}$, then there is a minimal non-zero k. H is then generated by g^k, and is therefore cyclic with order $|G|/k$. To

see this, note that taking any two elements $h_1, h_2 \in H$, and considering elements of the form $h_1^{n_1} h_2^{n_2}$ it is clear that what we get are all powers of the generator g with exponent $n_1 k_1 + n_2 k_2$, mod $r = |G|$. Using the Euclidean algorithm, the greatest common divisor d of k_1 and k_2 will be among the powers obtained, so $k_1 = k_1' d$, $k_2 = k_2' d$, and so $h_i = (g^d)^{k_i'}$.

Ex. A.2.8. Let us assume that $g^n = g^m$, and for simplicity sake, assume $m > n$, and multiply the equation with g^{-1} n times, yielding $1 = g^{m-n}$, which is only possible if $r|m-n$, and so $m = n(\mathrm{mod} r)$.

Ex. A.2.9. If $g_1, g_2 \in xH$, then $g_i = xh_i$, therefore $g_2 = xh_2$, $x = g_1 h_1^{-1}$, so $g_2 = g_1 h_1^{-1} h_2 = g_1 h$ where $h = h_1^{-1} h_2$.

Ex. A.2.10. As cosets define an equivalence relation on the group, each element belongs to a coset and only one. Each coset has the same number of elements as multiplication with a (sub)group element is a bijective mapping. Therefore the number of cosets is the number of all group elements divided by the elements of the subgroup whose cosets we are considering (and which is a divisor according to Lagrange's theorem), i.e., $G : H = |G|/|H|$ where the notation of $[G : H]$ is used for the number of cosets.

Ex. A.2.11. (1) As a representation is a homomorphism, $\rho(e) = I$, therefore $\chi(I) = \mathrm{Tr}\,\rho(e) = \mathrm{Tr}\, I = n$.

(2) If G is a finite group, then any element g has finite order, r, such that $g^r = e$. Also, diagonalising the matrix of $\rho(g)$ with a unitary transformation does not change the trace, therefore the character. In this basis, the matrix satisfies $\rho(g)^r = I$, so all its eigenvalues must satisfy $\lambda_i^r = 1$, therefore, they are complex numbers of unit magnitude, and the trace is their sum,

$$|\chi(g)| = |\,\mathrm{Tr}\,\rho(g)| = |\sum_k \lambda_k| \leq \sum_k |\lambda_k| = n\,.$$

(3) If $|\chi(g)| = n$, the direction of the eigenvalues of $\rho(g)$ must be the same in the complex plane [see part (2)], and they have unit magnitude, so all eigenvalues are the same. Therefore, the matrix is diagonal, and the eigenvalue is $\mathrm{e}^{\mathrm{i}\theta}$, where $\theta = 2\pi/r$, and so $\rho(g) = \mathrm{e}^{\mathrm{i}\theta} I$.

(4) $\chi(g^{-1}hg) = \mathrm{Tr}\,\rho(g)^{-1}\rho(h)\rho(g) = \mathrm{Tr}\,\rho(h) = \chi(h)$ using the cyclic property of the trace.

(5) As g and g^{-1} commute, the matrices $\rho(g)$ and $\rho(g^{-1})$ can be diagonalised with the same unitary transformation. In part (3) we have shown that the eigenvalues of the matrices are complex numbers of unit magnitude, and as $\rho(g^{-1})\rho(g) = I$, they are reciprocals, i.e., conjugates, $\lambda_k(g^{-1}) = \lambda_k(g)^*$, therefore, the same holds for the character, which is their sum, the trace of the representation matrices.

(6) All the eigenvalues are algebraic numbers, as they satisfy $\lambda_k^r - 1 = 0$, and the sum of algebraic numbers is algebraic.

Ex. A.2.12. Let us define another scalar product on the n-dimensional complex vector space by

$$\langle x, y\rangle_G := \sum_{g\in G} \langle \rho(g)x, \rho(g)y\rangle\,.$$

This scalar product is clearly invariant under the group action. Scalar products correspond to positive operators, i.e.,

$$\langle x, y\rangle_G = \langle x, A^\dagger A y\rangle = \langle Ax, Ay\rangle\,,$$

and the invariance reads

$$\langle A\rho(g)x, A\rho(g)y\rangle = \langle Ax, Ay\rangle\,,$$

writing $x' = Ax$, $y' = Ay$,

$$\langle A\rho(g)A^{-1}x', A\rho(g)A^{-1}y'\rangle = \langle x', y'\rangle\,,$$

demonstrating that $A\rho A^{-1}$ is unitary.

Ex. A.2.13. For a finite group, for all $g \in G$, $g^{|G|} = e$ (see excer. A2.6), so the same holds for the representation matrices, $\rho(g)^{|G|} = I$. Therefore, the representation consist of commuting diagonalisable matrices, so they can be diagonalised simultaneously. In the basis that diagonalised the representation, all the basis vectors correspond to 1-dimensional invariant subspaces, so for an irreducible representation there may be only one basis vector.

Note: finite order matrices, i.e., matrices A such that there is an integer k, $A^k = I$ are diagonalisable, as the minimal polynomial of the matrix is a divisor of $x^k - 1$, and this polynomial has simple roots.

Ex. A.2.14. For any element $g \in G$ let $C(g) = \{h^{-1}gh | g \in G\}$ denote the conjugacy class of the element. We shall also define the matrix $c(g) = \sum_{h\in C(g)} \rho(h)$.

We shall show that the matrix $c(g)$ commutes with the representation,

$$\rho(g')c(g) = \sum_{h^{-1}gh\in C(g)} \rho(g')\rho(h^{-1}gh) = \sum_h \rho(g'h^{-1}gh) = \sum_h \rho((hg'^{-1})^{-1}ghg'^{-1}g')$$

$$= \sum_h \rho(h'^{-1}gh'g') = \sum_h{}' \rho(h'^{-1}gh')\rho(g') = c(g)\rho(g').$$

According to Shur's lemma, therefore, $c(g)$ must be a constant times I, and using the fact that the trace of $\rho(g)$ is the character, constant on $C(g)$ (see excer. A2.11), $c(g) = c_g I$, $c_g = |C(g)|\chi(g)/d_\rho$.

Considering the sum $\sum_{g\in G} \chi(g^{-1})\rho(g)$, we may perform the sum by conjugacy class, and on each class insert the result from above, yielding

$$\sum_g \chi(g^{-1})\rho(g) = \frac{1}{d}\sum_i |C(g_i)|\chi(g^{-1})\chi(g)I\,,$$

where g_i is a set of representative elements from all characteristic classes. The trace of this is, on one hand, using the fact that the character is a class function,

$$\mathrm{Tr}\sum_g \chi(g^{-1})\rho(g) = \sum_g \chi(g)^2 = |G|\,,$$

therefore

$$\frac{n}{d_\rho} = \sum_i |C(g_i)|\chi(g^{-1})c_{g_i}\,.$$

The rest is number theory. It is shown that the RHS is algebraic integer and rational, therefore, integer. See Ref. [8]. For the algebraic integer, it is used, that characters are sums of roots of unity due to the fact that the order of an element of a finite group is finite.

Ex. A.2.15. Setting $i = j$ and $k = l$ in eq. (A2.3) yields

$$\sum_{g\in G} \chi^p(g^{-1})\chi^q(g) = |G|\delta_{pq}\,.$$

Also, according to excer. A2.11, characters are class function, and the sum over the group may be written as $\sum_{g\in G} = \sum_i \sum_{g\in C(g_i))}$ where g_i are representatives from each conjugacy class, on each class, the character is constant, $\chi(g) = \chi(g_i) = \chi_i$ is $g \in C(g_i)$, and als $\chi(g^{-1}) = \chi(g)^*$, yielding the relation

$$\sum_i r_i \chi_i^{p*}\chi_i^q = |G|\delta_{p,q}\,.$$

If we introduce the matrix $U_{iq} = \sqrt{r_i/|G|}\chi_q^p$, the above relation may be written as $U^\dagger U = I$, and then follows $UU^\dagger = I$, which has the matrix elements

$$\sqrt{\frac{r_i}{|G|}}\sqrt{\frac{r_j}{|G|}}\sum_p \chi_i^p\chi_j^{q*} = \delta_{ij}\,.$$

Ex. A.2.16. Let us enumerate the perturbations as p_0 to p_5 corresponding to 123, 231, 312, 213, 132, 321, respectively. The multiplication table of the group is as follows:

	p_0	p_1	p_2	p_3	p_4	p_5
p_0	p_0	p_1	p_2	p_3	p_4	p_5
p_1	p_1	p_2	p_0	p_5	p_3	p_4
p_2	p_2	p_0	p_1	p_4	p_5	p_3
p_3	p_3	p_4	p_5	p_0	p_1	p_2
p_4	p_4	p_5	p_3	p_2	p_0	p_1
p_5	p_5	p_3	p_4	p_1	p_2	p_0

The trivial representation, $\rho(g) = 1$ is clearly a representation, and it is easy to verify that $\rho = 1, 1, 1, -1, -1, -1$ is also a representation (sign, assigns +1 to cyclic and -1 to anti-cyclic permutations). The matrix representation is verified by computation, and its character (trace of the matrices is) 2, -1, -1, 0,0,0. Orthogonality is clearly satisfied.

Ex. A.2.17. If the regular representation wasn't faithful there would be an element $g \neq e$ also represented by the identity matrix. However, that would mean that it maps all element $h \in G$ to themselves, and the action is also the group action, i.e., $gh = h$ for all $h \in G$. This only holds for e in a group.

Ex. A.2.18. For the regular representation, all matrix elements are either 1 or 0. A 1 in the i, j element in the matrix $\rho(g)$ means that $gg_j = g_i$, so in the diagonal a 1 would mean that a group element leaves one invariant, which is only possible for the unit (if $gg' = g'$, this multiplied by g'^{-1} yields $g = e$). So, all representation matrices have 0 diagonals except that of the unit, and the representation is $|G|$ dimensiona, so $\chi(e) = |G|$ and the rest is $\chi(g \neq e) = 0$.

Ex. A.2.19. Inserting the character of the regular representation (see excer. A2.18) into eq. (A2.6) yields

$$c_p = \frac{1}{|G|}\sum_{i=1}^{r} r_i \chi_i^{\text{reg}*}\chi_i^\rho = \chi_e^\rho = d_\rho\,,$$

as only the conjugate class of e contributes, as $h^{-1}eh = e$ for all $h \in G$, this class has 1 element, $r_e = 1$, and on the unit element a character of a representation assumes the dimension of the representation (see excer. A2.11).

Ex. A.2.20. As the regular representation contains all irreducible representations d_ρ times, its matrices can be brought to a block-diagonal form, where each irrep is appears d_ρ times, yielding

$$\chi^{\mathrm{reg}} = \sum_{\rho\in\hat{G}} d_\rho \chi^\rho \,,$$

and we know that $\chi^{\mathrm{reg}}(g) = |G|\delta_{ge}$.

Ex. A.2.21. Evaluate the formula (A2.8) at $g = e$, and note that $\chi^\rho(e) = d_\rho$, yielding the desired result.

Ex. A.2.22. Substituting eq. (A2.10) into (A2.9) yields

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{|G|}} \sum_{g\in G} \rho(g) \frac{1}{\sqrt{|G|}} \sum_{\rho'\in\hat{G}} \sqrt{d_{\rho'}}\, \mathrm{Tr}\, \hat{f}(\rho')\rho'(g^{-1}) \,.$$

and exchanging summations, writing out indices yields

$$[\hat{f}(\rho)]_{ij} = \sum_{kl}\sum_{\rho'} \frac{\sqrt{d_\rho d_{\rho'}}}{|G|} [\hat{f}(\rho')]_{kl} \sum_{g\in G} [\rho(g)]_{ij}[\rho'(g)]_{lk}$$

and according to eq. (A2.3), the sum over $g \in G$ yields $(|G|/d_\rho)\delta_{ik}\delta_{jl}\delta_{\rho\rho'}$, and therefore the remaining sums yield $[\hat{f}(\rho)]_{ij}$.

Ex. A.2.23. The representations are labelled by the coefficient h in the exponent, and the value of the Fourier transform for the representation ρ_h is

$$\hat{f}(h) = \hat{f}(\rho_h) = \frac{1}{N}\sum_g f(g)\rho_h(g) = \frac{1}{N} f(g)\mathrm{e}^{-2\pi \mathrm{i} g h/N} \,,$$

as $d_{\rho_h} = 1$, $|G| = N$, and

$$f(g) = \frac{1}{\sqrt{N}} \sum_h \hat{f}(h)\mathrm{e}^{2\pi \mathrm{i} h g/N} \,,$$

as for a 1d representation, the trace is not needed.

The representations for $h = 0, 1, \ldots, N-1$ are inequivalent (as their character is orthogonal), and there are no more representations, as for $h = N$ we get the same matrices as for $h = 0$ due to the periodicity of the complex exponential. Also $\sum_h d_h^2 = \sum_h 1 = N$.

Note: the group is the addition with remainder.

Ex. A.2.24. The elements of the Fourier transform are

$$\hat{f}_0 = \frac{1}{\sqrt{6}} \sum_{i=0}^{5} f_i \,,$$

$$\hat{f}_1 = \frac{1}{\sqrt{6}} \sum_{i=0}^{5} f_i (-1)^{p_i} \,,$$

$$\hat{f}_2 = \begin{pmatrix} \frac{2f_0 - f_1 - f_2 - 2f_3 + f_4 + f_5}{2\sqrt{3}} & \frac{-f_1 + f_2 + f_4 - f_5}{2} \\ \frac{f_1 - f_2 + f_4 - f_5}{2} & \frac{2f_0 - f_1 - f_2 + 2f_3 - f_4 - f_5}{2\sqrt{3}} \end{pmatrix}$$

where by $(-1)^{p_i}$ we denote the sign of a permutation (the 1st representation in excer. A2.16) and $f_i = f(p_i)$. Listing the elements of the matrix $\hat{f}_2$ row-wise the matrix of the Fourier-transformation is

$$\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{6}} & -\frac{1}{2\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{6}} & \frac{1}{2\sqrt{6}} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} \end{pmatrix}$$

The matrix of the inverse transformation is the transpose of the above matrix, and

$$\begin{aligned} f_0 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1 + \sqrt{2}(\hat{f}_{2,11} + \hat{f}_{2,22}))\,, \\ f_1 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1) - \frac{1}{\sqrt{12}}(\hat{f}_{2,11} + \hat{f}_{2,22}) - \frac{1}{2}(\hat{f}_{2,12} - \hat{f}_{2,21})\,, \\ f_2 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1) - \frac{1}{\sqrt{12}}(\hat{f}_{2,11} + \hat{f}_{2,22}) + \frac{1}{2}(\hat{f}_{2,12} - \hat{f}_{2,21})\,, \\ f_3 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1) - \frac{1}{\sqrt{3}}(\hat{f}_{2,11} - \hat{f}_{2,22})\,, \\ f_4 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1) + \frac{1}{\sqrt{12}}(\hat{f}_{2,11} - \hat{f}_{2,22}) + \frac{1}{2}(\hat{f}_{2,12} + \hat{f}_{2,21})\,, \\ f_5 &= \frac{1}{\sqrt{6}}(\hat{f}_0 + \hat{f}_1) + \frac{1}{\sqrt{12}}(\hat{f}_{2,11} - \hat{f}_{2,22}) - \frac{1}{2}(\hat{f}_{2,12} + \hat{f}_{2,21})\,. \end{aligned}$$

A.3 The Solovay-Kitaev theorem

Ex. A.3.1. The newly defined distance function $D(U,V)$ is the sum of eigenvalues of $|U-V|$, whereas $E(U,V) = \|U-V\|$ is the absolute value of the maximal eigenvalue of $U-V$.

When calculating the eigenvalues of the matrix $M = R_{\hat{\mathbf{n}}}(\phi) - R_{\hat{\mathbf{m}}}(\theta)$, the eigenvalues are calculated as

$$\lambda_{1,2} = -t \pm \sqrt{t^2 - d}\,,$$

where $t = \mathrm{Tr}\,(R_{\hat{\mathbf{n}}}(\phi) - R_{\hat{\mathbf{m}}}(\theta))/2$ and $d = \det R_{\hat{\mathbf{n}}}(\phi) - R_{\hat{\mathbf{m}}}(\theta)$, and the term ($1/4$ of the discriminator) under the square root is

$$t^2 - d = \left(\cos\frac{\theta-\phi}{2} - \cos\frac{\theta+\phi}{2}\right)\hat{\mathbf{n}}\cdot\hat{\mathbf{m}} + \frac{\cos\theta + \cos\phi}{2} - 1\,,$$

which is negative, so the two eigenvalues are complex conjugates, their absolute values are equal, which completes the proof.

Ex. A.3.2. Using the series expansion of the exponential to second order,

$$\mathrm{e}^{-\mathrm{i}A} = I - \mathrm{i}A - \frac{1}{2}A^2 + O(\varepsilon^3)\,,$$

etc.,

$$\left[\mathrm{e}^{-\mathrm{i}A}, \mathrm{e}^{\mathrm{i}B}\right] = \left[I - \mathrm{i}A - \frac{1}{2}A^2, I - \mathrm{i}B - \frac{1}{2}B^2\right] + O(\varepsilon^3) = I - [A,B] + O(\varepsilon^3)\,,$$

and

$$\mathrm{e}^{-[A,B]} = I - [A,B] + O(\varepsilon^4)\,,$$

and eq. (A3.9) follows from here.

Ex. A.3.3. Let us note first, that the explicit form of the parametrisation follows from excer. 2.35,

$$u(\mathbf{a}) = \exp\left(-\frac{\mathrm{i}}{2}\mathbf{a}\cdot\vec{\sigma}\right) = \cos\frac{a}{2}I - \mathrm{i}\hat{\mathbf{a}}\cdot\vec{\sigma}\sin\frac{a}{2}\,,$$

where $a = |\mathbf{a}|$ and $\hat{\mathbf{a}} = \mathbf{a}/a$.

We now use the results of excer. A.3.1, and calculate $D(u(\mathbf{x}, u(\mathbf{y})) = 2E(u(\mathbf{x}, u(\mathbf{y}))$ as follows:

$$E(u(\mathbf{x}), u(\mathbf{y})) = \max_{\psi:\|\psi\|=1} \left\langle \psi \,\middle|\, \left(\mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\vec{\sigma}} - \mathrm{e}^{\mathrm{i}\mathbf{y}\cdot\vec{\sigma}}\right)\left(\mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\vec{\sigma}} - \mathrm{e}^{-\mathrm{i}\mathbf{y}\cdot\vec{\sigma}}\right) \,\middle|\, \psi \right\rangle^{1/2}$$

and simplify the operator whose expectation value is calculated as

$$\begin{aligned}\left(\mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\vec{\sigma}} - \mathrm{e}^{\mathrm{i}\mathbf{y}\cdot\vec{\sigma}}\right)\left(\mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\vec{\sigma}} - \mathrm{e}^{-\mathrm{i}\mathbf{y}\cdot\vec{\sigma}}\right) &= 2 - \left(\mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\vec{\sigma}}\mathrm{e}^{-\mathrm{i}\mathbf{y}\cdot\vec{\sigma}} + \mathrm{e}^{\mathrm{i}\mathbf{y}\cdot\vec{\sigma}}\mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\vec{\sigma}}\right)\,,\\ &= 2\left[1 - \cos\frac{x}{2}\cos\frac{y}{2} - \sin\frac{x}{2}\sin\frac{y}{2}\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}\right]I\,,\end{aligned}$$

where we have used the explicit form of the parametrisation and $\hat{\mathbf{x}}\cdot\vec{\sigma}\hat{\mathbf{y}}\cdot\vec{\sigma} + \hat{\mathbf{y}}\cdot\vec{\sigma}\hat{\mathbf{x}}\cdot\vec{\sigma} = 2\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}$, yielding the desired result [10].

Ex. A.3.4. In the above formula setting $y = 0$ and using $\cos(x/2) = 1 - 2\sin^2|x/4|$.

Ex. A.3.5. When the two vectors are small enough, we use the following replacements: $\cos\frac{x}{2} = 1 - x^2/8 + O(\varepsilon^4)$, $\sin x/2 = x/2 + O(\epsilon^3)$ yielding

$$D(u(\mathbf{x}), u(\mathbf{y})) = 2\sqrt{2}\left(\frac{x^2}{8} + \frac{y^2}{8} - \frac{1}{8}\mathbf{x}\cdot\mathbf{y}\right)^{1/2} + O(\epsilon^3)\,,$$

and

$$\|\mathbf{x} - \mathbf{y}\| = \left(x^2 + y^2 - 2\mathbf{x}\cdot\mathbf{y}\right)^{1/2}\,,$$

which agrees.

Ex. A.3.6. To approximate an element of $SU(2)$ with a sequence of elementary gates, i.e., words of a generating set $\mathcal{G}$, we use the following steps: if a zeroth approximation is needed, we choose the nearest element from $\mathcal{G}$. Assuming we have an nth approximation of the operator U to be approximated, we construct the next approximation by writing the error in the form $\Delta = UU_n^\dagger$, and then decompose this as $\Delta = VWV^\dagger W^\dagger$ in such a way that the norm of U, V is small enough, and approximate those to the nth level, and finally get $U_{n+1} = V_n W_n V_n^\dagger W_n^\dagger U_n$. For details, see Ref. [11].

A.4 Number theory

Ex. A.4.1. If $a|b$ and $b|c$, this means that $\exists n, m$ such that $b = na$ and $c = mb$, so $c = nma$, and noting that nm is an integer, this is the definition of $a|c$.

Ex. A.4.2. If $d|a$ and $d|b$, this means that $\exists n, m$ such that $a = nd$, $b = md$, so $ax + by = ndx + mdy = (nx + my)d$, therefore $d|ax + by$.

Ex. A.4.3. If $a|b$ then $\exists n$ such that $na = b$. If $a, b > 0$, then so is n, and, as it is an integer, $n \geq 1$, so $a \leq na = b$. Cosequently, if $a|b$ and $b|$, then $a \leq b$ and $b \leq a$, implying $a = b$.

Note that if the positivity condition is dropped, then $a = -b$ is also possible, but the above argument may be used for $|a|$ and $|b|$.

Ex. A.4.4. $697 = 17 \cdot 41$ and $36300 = 2^2 \cdot 3 \cdot 5^2 \cdot 11^2$.

Ex. A.4.5. For p prime, any integer x in the range $1, \ldots, p-1$ has $\gcd(x, p) = 1$ as all divisors d of x are $d \leq x$ and the divisors of p are only $p > x$ and 1.

In the case of $n = p^2$, all $x \in \{1, \ldots, n-1\}$ have $\gcd(x, n) = 1$ except $p, 2p, \ldots, p^2$, which have $\gcd(kp, p^2) = p$ for $p = 1, 2, \ldots, p$, so the classes that have multiplicative inverses are

$$\{1 \leq k \leq p^2 - 1 | \gcd(k, p^2) = 1\} = \{1 \leq k \leq p^2 - 1 | p \not| k\} .$$

Ex. A.4.6. Looping over the remainders from 1, ..., 24: the result is $17^{-1} = 17 \pmod{24}$.

Ex. A.4.7. Note that $(n+1)(n-1) = n^2 - 1 \equiv -1 \pmod{n^2}$, so $n^2 - (n-1) = n^2 - n + 1 = n(n-1)+1$ is the inverse of $n+1$. Verification: $(n+1)(n^2-n+1) = n^3+1 = nn^2+1 \equiv 1 \pmod{n}$.

Ex. A.4.8. If $ab = ab' = 1 \pmod{n}$ then $ab = kn+1$ and $ab' = kn'+1$ so $a(b-b') = (k-k')n$, and as a has a multiplicative inverse, its gcd with n is 1, so the only way $n|a(b-b')$ is if $n|(b-b')$.

Ex. A.4.9. Let us write the prime factorisations as $a = p_1^{j_1} p_2^{j_2} \ldots p_n^{j_n}$ and $b = p_1^{k_1} p_2^{k_2} \ldots p_n^{k_n}$ (we may write the same primes in both cases allowing 0 exponents). Then $\gcd(a, b) = p_1^{\min\{j_1,k_1\}} \cdot p_2^{\min\{j_2,k_2\}} \ldots p_n^{\min\{j_n,k_n\}}$.

In the example: $6825 = 3 \cdot 5^2 \cdot 7 \cdot 13$ and $1430 = 2 \cdot 5 \cdot 11 \cdot 13$, so $\gcd(6825, 1430) = 5 \cdot 13 = 65$.

Ex. A.4.10. As $187 = 11 \cdot 17$, the Euler-function assumes the value $\varphi(187) = \varphi(11)\varphi(17) = 10 \cdot 16 = 160$.

Ex. A.4.11. Let first $n = p^\alpha$, then the divisors of n are $p^{\alpha'}$, $\alpha' = 0, \ldots, \alpha - 1$. According to eq. (A4.23), $\varphi(p^\alpha) = p^{\alpha-1}(p-1)$, so the right side of eq. (A4.24) is

$$\sum_{d|n} \varphi(d) = 1 + \sum_{\alpha'=1}^{\alpha} p^{\alpha'-1}(p-1) = 1 + (p-1)\frac{p^\alpha - 1}{p-1} = p^\alpha .$$

For the general case, let us consider the prime factorisation of n and its divisors, $n = p_1^{\alpha_1} \ldots p_k^{\alpha_k}$. In this case, we have a multiple sum, in which we may use the multiplicative property, so

$$\sum_{d|n} \varphi(d) = \sum_{\beta_1=0}^{\alpha_1} \cdots \sum_{\beta_k=0}^{\alpha_k} \varphi(p_1^{\beta_1} \ldots p_k^{\beta_k}) = \sum_{\beta_1=0}^{\alpha_1} \varphi(p_1^{\beta_1}) \cdots \sum_{\beta_k=0}^{\alpha_k} \varphi(p_k^{\beta_k})$$

and use the case we have already shown to complete the proof.

Ex. A.4.12. The elements of $\mathbf{Z}_n^*$ are

$$\mathbb{Z}_n^* = \{k \in \{1, \ldots, n-1\} | \gcd(k, n) = 1\} ,$$

so it is clear that $|\mathbb{Z}_n^*| = \varphi(n)$. It is a group, (1) if $k, \ell \in \mathbb{Z}_n^*$ then $k\ell \in \mathbb{Z}_n^*$, as if $\gcd(k\ell, n) \neq 1$ then there is a prime $p|\gcd(k\ell, n)$, so $p|n$ and $p|k\ell$, so either $p|k$ or $p|\ell$ which would lead to $p|\gcd(k, n)$ or $p|\gcd(\ell, n)$, leading to a contradiction with both gcd's being 1. (2) All elements have inverses (by definition). (3) There is a unit, $1 \in \mathbb{Z}_n^*$.

Ex. A.4.13. (1) It is generally true that the powers of any group element form a subgroup. It is obviously closed to multiplication, all elements have an inverse due to Thm. A4.9, and the unit is included. (2) The size of the subgroup is the order of a, as higher powers simply repeat the sequence. (Note: this is a cyclic subgroup.)

Ex. A.4.14. If $g \in \mathbb{Z}_n^*$ is a generator, $\langle g \rangle = \{1, g, g^2, \dots\} = \mathbb{Z}_n^*$, then according to ex. A.4.13, the subgroup formed by the powers of g is the full group. The order of the subgroup is the order of g, and the order of the full group is $\varphi(n)$, so the two must agree.

Ex. A.4.15. The order r of any element $a \in \mathbb{Z}_n^*$ is the order of the subgroup formed by the powers of a according to ex. A.4.13. The order of the full group $\mathbb{Z}_n^*$ is $\phi(n)$, so according to Lagrange's theorem A2.1, $r|\varphi(n)$, i.e., $\exists k \in \mathbb{N}$ s.t. $rk = \varphi(n)$. In this case, $a^{\varphi(n)} = a^{rk} = (a^r)^k = 1^k = 1 \pmod n$, which is Lagrange's theorem A4.9.

Ex. A.4.16. According to thm. 4.9, $x^{\varphi(N)} = 1 \pmod N$. The definition of the order is that it is the minimal positive integer o such that $x^o = 1 \pmod N$. Consequently, for any multiple ko of o holds that $x^{ko} = (x^o)^k = 1 \pmod N$. If φN wasn't a multiple of o, then we could take the remainder of $\varphi(N)$ with o as $\varphi(N) = ko + r$, and so $x^{\varphi(N)} = x^{ko+r} = x^{ko}x^r = x^r \pmod N$ would hold, and so $x^r = 1 \pmod N$ as well, which contradicts the minimality of o.

(Note: the fact that the order divides $\varphi(N)$ follows also from Lagrange's theorem, see ex. A.4.15.)

Ex. A.4.17. Let us assume there is an efficient way to factor integers. Let $N = p_1^{\alpha_1} \dots p_n^{\alpha_n}$. The order of x is as follows: we know that the order r of x is a divisor of $\phi(N) = \prod_k p_k^{\alpha_k - 1}(p_k - 1)$. Knowing the factors of N and $\phi(N) = \prod_\ell q_\ell^{\beta_\ell}$, we only need to check its divisors, i.e., numbers with the same prime factors q_ℓ and exponents below or equal β_ℓ.

Ex. A.4.18. Following the split and invert technique

$$\frac{19}{17} = 1 + \frac{1}{8 + \frac{1}{2}}\,,$$

and

$$\frac{77}{65} = 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}}\,.$$

Ex. A.4.19. Let us verify the equality for $n = 1$, in this case $p_1 = 1 + a_0 a_1$, $q_1 = a_1$, $p_0 = a_0$, $q_0 = 1$, so

$$q_1 p_0 - p_1 q_0 = a_1 a_0 - (1 + a_0 a_1) \cdot 1 = -1\,.$$

Let us assume the equality has been proven for $n = 1, \dots, N-1$ and perform the induction step, by inserting the recursive formulae (A4.41,42) for p_N and q_N,

$$\begin{aligned} q_N p_{N-1} - p_N q_{N-1} &= (a_N q_{N-1} + q_{N-2}) - (a_N p_{N-1} - p_{N-1} q_{N-1}) \\ &= a_N (q_{N-1} p_{N-1} - p_{N-1} q_{N-1}) + q_{N-2} p_{N-1} - p_{N-2} q_{N-1} = -(-1)^{N-1}\,, \end{aligned}$$

which completes the proof.

A.5 Public key cryptography and the RSA cryptosystem

Ex. A.5.1. Let us encrypt the alphabet as follows: 'a'$\mapsto$ 2, 'z'$\mapsto$ 27, and store these numbers on 5 bits. Then "quantum" is encoded as 18, 22, 2, 15, 21, 22, 14. Now $p = 3$, $q = 11$, so $n = pq = 33$ and $\varphi(n = 33) = (p-1)(q-1) = 20$. We need to choose and integer e relative prime to this, such as 7. The multiplicative inverse is $d = 3$, as $7 \cdot 3 = 21 = 1 \pmod{20}$. The encryption function is $x \mapsto x^e \pmod{n}$ which yields for "quantum" 6, 22, 29, 27, 21, 22, 20.

Note: we started the encoding from 2 because 0 and 1 are mapped into themselves.

Ex. A.5.2. The order r is a divisor of $\varphi(n)$, and we have $ed = 1 \pmod{\varphi(n)}$, so $ed = k\varphi(n)+1 = k'r+1$, where $k' = k\varphi(n)/r$ is an integer. This means that d is also a multiplicative inverse of $e \bmod r$, so $d' = d \pmod{r}$ as the multiplicative inverse in a group is unique.

References

[1] A. Peres, "Higher order Schmidt decompositions", *Phys. Lett.* **A202**, 16–17 (1995), `arXiv:quant-ph/9504006`.

[2] J.H. Conway, "Unpredictable iterations", In: Number theory conference, University of Colorado Boulder, 1972, p. 49-52.

[3] R. Downey, "An invitation to structural complexity", *New Zealand J. Mathematics* **21**, 22-89 (1992).

[4] C.H. Bennett, "Logical reversibility of computation", *IBM J. Research Development* **17**, 525-532 (1975).

[5] `https://quantumcomputing.stackexchange.com/questions/7082/how-to-reduce-circuit-elements-of-a-decomposed-c2u-operation`.

[6] `https://quantumcomputing.stackexchange.com/questions/16115/partial-cyclic-permutation-with-only-toffoli-and-cnot-gates/23311#23311`.

[7] A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, "Elementary gates for quantum computation", *Phys. Rev.* **A 52**, 3457 (1995) `arXiv:quant-ph/9503016`.

[8] M. Hall, Jr., "The theory of groups" (AMS Chelsea Pub., Providence, RI, 1976).

[9] `https://quantumcomputing.stackexchange.com/questions/10216/why-is-deutschs-gate-universal`

[10] `https://enakai00.hatenablog.com/entry/2018/05/19/214521`

[11] C.A. Dawson, M.A. Nielsen, "The Solovay-Kitaev algorithm", `arXiv:0505030 [quant-ph]` (2005).

[12] J.C. Majithia and D. Levan, "A note on base-2 logarithm computation", *Proc. IEEE*, **61**1519–1520 (1973).

[13] K.K.H. Cheung and M. Mosca, "Decomposing finite Abelian groupa", `arXiv:cs/0101004`