Lectures on Quantum Computation, Quantum Error Correcting Codes and Information Theory

by

K. R. Parthasarathy
Indian Statistical Institute
(Delhi Centre)

Notes by
Amitava Bhattacharya
Tata Institute of Fundamental Research, Mumbai

Preface

These notes were prepared by Amitava Bhattacharya on a course of lectures I gave at the Tata Institute of Fundamental Research (Mumbai) in the months of April 2001 and February 2002. I am grateful to my colleagues at the TIFR, in general, and Professor Parimala Raman in particular, for providing me a receptive and enthusiastic audience and showering on me their warm hospitality. I thank Professor Jaikumar for his valuable criticism and insight and several fruitful conversations in enhancing my understanding of the subject. Finally, I express my warm appreciation of the tremendous effort put in by Amitava Bhattacharya for the preparation of these notes and their LaTeX files.

Financial support from the Indian National Science Academy (in the form of C. V. Raman Professorship), TIFR (Mumbai) and Indian Statistical Institute (Delhi Centre) is gratefully acknowledged.

K. R. Parthasarathy
Delhi
June 2001

Contents

Lecture 1

Quantum Probability

In the Mathematical Congress held at Berlin, Peter Shor presented a new algorithm for factoring numbers on a *quantum computer.* In this series of lectures, we shall study the areas of quantum computation (including Shor's algorithm), quantum error correcting codes and quantum information theory.

1.1 Classical Versus Quantum Probability Theory

We begin by comparing classical probability and quantum probability. In classical probability theory (since Kolmogorov's 1933 monograph [11]), we have a sample space, a set of events, a set of random variables, and distributions. In quantum probability (as formulated in von Neumann's 1932 book [14]), we have a state space (which is a Hilbert space) instead of a sample space; events, random variables and distributions are then represented as operators on this space. We now recall the definitions of these notions in classical probability and formally define the analogous concepts in quantum probability. In our discussion we will be concerned only with *finite* classical probability spaces, and their quantum analogues—finite dimensional Hilbert spaces.

Spaces	
1.1 The sample space Ω: This is a finite set, say $\{1, 2, \ldots, n\}$.	**1.2 The state space H:** It is a complex Hilbert space of dimension n.

Events	
1.3 The set of events $\mathcal{F}_\Omega$: This is the set of all subsets of Ω. $\mathcal{F}_\Omega$ is a Boolean algebra with the *union* ($\cup$) operation for 'or' and the *intersection* ($\cap$) operation for 'and'. In particular, we have $E \cap (F_1 \cup F_2) = (E \cap F_1) \cup (E \cap F_2)$.	**1.4 The set of events $\mathcal{P}(H)$:** This is the set of all orthogonal projections in $\mathcal{H}$. An element $E \in \mathcal{P}(H)$ is called an *event*. Here, instead of '$\cup$' we have the max ($\vee$) operation, and instead of '$\cap$' the min ($\wedge$) operation. Note, however, that $E \wedge (F_1 \vee F_2)$ is not always equal to $(E \wedge F_1) \vee (E \wedge F_2)$. (They are equal if E, F_1, F_2 commute with each other).

Random variables and observables	
1.5 The set of random variables $\mathcal{B}_\Omega$: This is the set of all complex valued functions on Ω. The elements of $\mathcal{B}_\Omega$ are called *random variables*. $\mathcal{B}_\Omega$ is an Abelian C^*-algebra under the operations $(\alpha f)(\omega) = \alpha f(\omega)$; $(f+g)(\omega) = f(\omega) + g(\omega)$; $(f \cdot g)(\omega) = f(\omega) g(\omega)$; $f^*(\omega) \stackrel{\Delta}{=} f^\dagger(\omega) = f(\bar{\omega})$. Here, $\alpha \in \mathbb{C}$, $f, g \in \mathcal{B}_\Omega$, and the 'bar' stands for complex conjugation. The random variable $\mathbf{1}$ (defined by $\mathbf{1}(\omega) \stackrel{\Delta}{=} 1$), is the unit in this algebra.	**1.6 The set of observables $\mathcal{B}(\mathcal{H})$:** This is the (non-Abelian) C^*-algebra of all operators on $\mathcal{H}$, with '$+$' and '$\cdot$' defined as usual, and X^* defined to be the adjoint of X. We will use $X^\dagger$ instead of X^*. The identity projection I is the unit in this algebra. We say that an observable is real-valued if $X^\dagger = X$, that is, if X is Hermitian. For such an observable, we define $\mathrm{Sp}(X)$ to be the set of eigen values of X. Since X is Hermitian, $\mathrm{Sp}(X) \subseteq \mathbb{R}$, and by the spectral theorem, we can write X as $X = \sum_{\lambda \in \mathrm{Sp}(X)} \lambda E_\lambda$,

With each event $E \in \mathcal{F}_\Omega$ we associate the indicator random variable $\mathbf{1}_E$ defined by

$$\mathbf{1}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E; \\ 0 & \text{otherwise.} \end{cases}$$

For a random variable f, let $\mathrm{Sp}(f) \stackrel{\Delta}{=} f(\Omega)$. Then, f can be written as the following linear combination of indicator random variables:

$$f = \sum_{\lambda \in \mathrm{Sp}(f)} \lambda \mathbf{1}_{f^{-1}(\{\lambda\})},$$

so that

$$\mathbf{1}_{f^{-1}(\{\lambda\})} \cdot \mathbf{1}_{f^{-1}(\{\lambda'\})} = \mathbf{0} \text{ for } \lambda \neq \lambda';$$
$$\sum_{\lambda \in \mathrm{Sp}(f)} \mathbf{1}_{f^{-1}(\{\lambda\})} = \mathbf{1}.$$

Similarly, we have

$$f^r = \sum_{\lambda \in \mathrm{Sp}(f)} \lambda^r \mathbf{1}_{f^{-1}(\{\lambda\})},$$

and, in general, for a function $\varphi : \mathbb{C} \to \mathbb{C}$, we have the random variable

$$\varphi(f) = \sum_{\lambda \in \mathrm{Sp}(f)} \varphi(\lambda) \mathbf{1}_{f^{-1}(\{\lambda\})}.$$

Later, we will be mainly interested in real-valued random variables, that is random variables f with $\mathrm{Sp}(f) \subseteq \mathbb{R}$ (or $f^\dagger = f$).

where E_λ is the projection on the subspace $\{u : Xu = \lambda u\}$ and

$$E_\lambda E_{\lambda'} = \mathbf{0}, \lambda, \lambda' \in \mathrm{Sp}(X), \lambda \neq \lambda';$$
$$\sum_{\lambda \in \mathrm{Sp}(X)} E_\lambda = I.$$

Similarly, we have

$$X^r = \sum_{\lambda \in \mathrm{Sp}(X)} \lambda^r E_\lambda,$$

and in general, for a function $\varphi : \mathbb{R} \to \mathbb{R}$, we have

$$\varphi(X) = \sum_{\lambda \in \mathrm{Sp}(X)} \varphi(\lambda) E_\lambda.$$

Distributions and states	
1.7 A distribution $\boldsymbol{p}$: This is a function from $\mathcal{F}_\Omega$ to $\mathbb{R}$, determined by n real numbers $p_1, p_2, \ldots, p_n$, satisfying: $$\begin{aligned} p_i &\geq 0; \\ \sum_{i=1}^{n} p_i &= 1. \end{aligned}$$ The probability of the event $E \in \mathcal{F}_\Omega$ (under the distribution $\boldsymbol{p}$) is $$\Pr(E; \boldsymbol{p}) \triangleq \sum_{i \in E} p_i.$$ When there is no confusion we write $\Pr(E)$ instead of $\Pr(E; \boldsymbol{p})$. We will identify $\boldsymbol{p}$ with the sequence $(p_1, p_2, \ldots, p_n)$. The probability that a random variable f takes the value $\lambda \in R$ is $$\Pr(f = \lambda) \triangleq \Pr(f^{-1}(\{\lambda\}));$$ thus, a real-valued random variable f has a distribution on the real line with mass $\Pr(f^{-1}(\{\lambda\}))$ at $\lambda \in \mathbb{R}$.	**1.8 A state ρ:** In quantum probability, we have a state ρ instead of the distribution $\boldsymbol{p}$. A state is a non-negative definite operator on $\mathcal{H}$ with $\operatorname{Tr}\rho = 1$. The *probability* of the event $E \in \mathcal{P}(H)$ in the state ρ is defined to be $\operatorname{Tr}\rho E$, and the probability that the real-valued observable X takes the value λ is $$\Pr(X = \lambda) = \begin{cases} \operatorname{Tr}\rho E_\lambda & \text{if } \lambda \in \operatorname{Sp}(X); \\ 0 & \text{otherwise.} \end{cases}$$ Thus, a real-valued observable X has a distribution on the real line with mass $\operatorname{Tr}\rho E_\lambda$ at $\lambda \in \mathbb{R}$.

Expectation, moments, variance	
The expectation of a random variable f is $$\mathop{\mathbb{E}}_{\boldsymbol{p}} f \triangleq \sum_{\omega \in \Omega} f(\omega) p_\omega.$$ The *r-th* moment of f is the expectation of f^r, that is	The expectation of an observable X in the state ρ is $$\mathop{\mathbb{E}}_{\rho} X \triangleq \operatorname{Tr}\rho X.$$ The map $X \mapsto \mathbb{E}_\rho X$ has the following properties:

$$\begin{aligned}\mathop{\mathbb{E}}_{\boldsymbol{p}} f^r &= \sum_{\omega\in\Omega}(f(\omega))^r p_\omega \\ &= \sum_{\lambda\in \mathrm{Sp}(f)} \lambda^r \Pr(f^{-1}(\lambda)),\end{aligned}$$

and the *characteristic function* of f is the expectation of the complex-valued random variable e^{itf}, that is,

$$\mathop{\mathbb{E}}_{\boldsymbol{p}} e^{itf} = \sum_{\lambda\in \mathrm{Sp}(f)} e^{it\lambda} \Pr(f^{-1}(\lambda)).$$

The variance of a real-valued random variable f is

$$\mathbf{var}(f) \stackrel{\Delta}{=} \mathop{\mathbb{E}}_{\boldsymbol{p}}(f - \mathop{\mathbb{E}}_{\boldsymbol{p}} f)^2 \geq 0.$$

Note that

$$\mathbf{var}(f) = \mathop{\mathbb{E}}_{\boldsymbol{p}} f^2 - (\mathop{\mathbb{E}}_{\boldsymbol{p}} f)^2;$$

also, $\mathbf{var}(f) = 0$ if and only if all the mass in the distribution of f is concentrated at $\mathbb{E}_{\boldsymbol{p}} f$.

(1) It is linear;

(2) $\mathbb{E}_\rho X^\dagger X \geq 0$, for all $X \in \mathcal{B}(\mathcal{H})$.

(3) $\mathbb{E}_\rho I = 1$.

The *r-th moment* of X is the expectation of X^r; if X is real-valued, then using the spectral decomposition, we can write

$$\mathop{\mathbb{E}}_{\rho} X^r = \sum_{\lambda\in \mathrm{Sp}(X)} \lambda^r \operatorname{Tr} \rho E_\lambda.$$

The *characteristic function* of the real-valued observable X is the expectation of the observable e^{itX}. The variance of a (real-valued) observable X is

$$\begin{aligned}\mathbf{var}(X) &\stackrel{\Delta}{=} \operatorname{Tr}\rho(X - \operatorname{Tr}\rho X)^2 \\ &= \operatorname{Tr}\rho X^2 - (\operatorname{Tr}\rho X)^2 \\ &\geq 0.\end{aligned}$$

The variance of X vanishes if and only if the distribution of X is concentrated at the point $\operatorname{Tr}\rho X$. This is equivalent to the property that the operator range of ρ is contained in the eigensubspace of X with eigenvalue $\operatorname{Tr}\rho X$.

Extreme points

1.9 The set of distributions: The set of all probability distributions on Ω is a compact convex set (Choquet simplex) with exactly n extreme points, δ_j $(j = 1, 2, \ldots, n)$, where δ_j is determined by

1.10 The set of states: The set of all states in $\mathcal{H}$ is a convex set. Let ρ be a state. Since ρ is non-negative definite, its eigen values are non-negative reals, and we can write

$$\delta_j(\{\omega\}) \stackrel{\Delta}{=} \begin{cases} 1 & \text{if } \omega = j; \\ 0 & \text{otherwise.} \end{cases}$$

If $P = \delta_j$, then every random variable has a degenerate distribution under P: the distribution of the random variable f is concentrated on the point $f(j)$.

$$\rho = \sum_{\lambda \in \mathrm{Sp}(\rho)} \lambda E_\lambda;$$

since $\operatorname{Tr} \rho = 1$, we have

$$\sum_{\lambda \in \mathrm{Sp}(\rho)} \lambda \times \dim(E_\lambda) = 1.$$

The projection E_λ can, in turn, be written as a sum of *one-dimensional* projections:

$$E_\lambda = \sum_{i=1}^{\dim(E_\lambda)} E_{\lambda,i}.$$

Then,

$$\rho = \sum_{\lambda \in \mathrm{Sp}(\rho)} \sum_{i=1}^{\dim(E_\lambda)} \lambda E_{\lambda,i}.$$

Proposition 1.1.1 *A one-dimensional projection cannot be written as a non-trivial convex combination of states.*

Thus, the extreme points of the convex set of states are precisely the one-dimensional projections. Let ρ be the extreme state corresponding to the one-dimensional projection on the ray $\mathbb{C}u$ (where $\|u\| = 1$). Then, the expectation m of the observable X is

$$m = \operatorname{Tr} uu^\dagger X = \operatorname{Tr} u^\dagger X u = \langle u, Xu \rangle,$$

and

$$\begin{aligned} \mathbf{var}(X) &= \operatorname{Tr} uu^\dagger (X-m)^2 \\ &= \operatorname{Tr} \|(X-m)u\|^2. \end{aligned}$$

	Thus, $\mathbf{var}(X) = 0$ if and only if u is an eigenvector of X. So, even for this extreme state, not all observables have degenerate distributions: *degeneracy of the state does not kill the uncertainty of the observables!*

The product	
1.11 Product spaces: If there are two statistical systems described by classical probability spaces $(\Omega_1, \boldsymbol{p}_1)$ and $(\Omega_2, \boldsymbol{p}_2)$ respectively, then the probability space $(\Omega_1 \times \Omega_2, \boldsymbol{p}_1 \times \boldsymbol{p}_2)$ determined by $\Pr(\{(i,j)\}; \boldsymbol{p}_1 \times \boldsymbol{p}_2) \triangleq \Pr(\{i\}; \boldsymbol{p}_1) \Pr(\{j\}; \boldsymbol{p}_2)$, describes the two independent systems as a single system.	**1.12 Product spaces:** If $(\mathcal{H}_1, \rho_1)$ and $(\mathcal{H}_2, \rho_2)$ are two quantum systems, then the quantum system with state space $\mathcal{H}_1 \otimes \mathcal{H}_2$ and state $\rho_1 \otimes \rho_2$ (which is a non-negative definite operator of unit trace on $\mathcal{H}_1 \otimes \mathcal{H}_2$) describes the two independent quantum systems as a single system.
Dynamics	
1.13 Reversible dynamics in Ω: This is determined by a bijective transformation $T : \Omega \to \Omega$. Then, $f \rightsquigarrow f \circ T$ (for random variables) $P \rightsquigarrow P \circ T^{-1}$ (for distributions)	**1.14 Reversible dynamics in $\mathcal{H}$:** This is determined by a unitary operator $U : \mathcal{H} \to \mathcal{H}$. Then, we have the dynamics of Heisenberg: $X \rightsquigarrow U^\dagger X U$ for $X \in \mathcal{B}(\mathcal{H})$; Schrödinger $\rho \rightsquigarrow U \rho U^\dagger$ for the state ρ.

1.2 Three Distinguishing Features

We now state the first distinguishing feature.

Proposition 1.2.1 *Let E and F be projections in $\mathcal{H}$ such that $EF \neq FE$. Then, $E \vee F \leq E + F$ is false.*

Proof Suppose $E \vee F \leq E + F$. Then, $E \vee F - E \leq F$. So,

$$F(E \vee F - E) = (E \vee F - E)F.$$

That is, $FE = EF$, a contradiction.

□

Corollary 1.2.2 *Suppose E and F are projections such that $EF \neq FE$. Then, for some state ρ, the inequality $\operatorname{Tr} \rho(E \vee F) \leq \operatorname{Tr} \rho E + \operatorname{Tr} \rho F$ is false.*

Proof By the above proposition, $E \vee F \leq E + F$ is false; that is, there exists a unit vector u such that

$$\langle u, (E \vee F)u \rangle \not\leq \langle u, Eu \rangle + \langle u, Fu \rangle .$$

Choose ρ to be the one dimensional projection on the ray $\mathbb{C}u$. Then,

$$\begin{aligned} \operatorname{Tr}(E \vee F)\rho &= \langle u, (E \vee F)u \rangle , \\ \operatorname{Tr} E\rho &= \langle u, Eu \rangle , \\ \operatorname{Tr} F\rho &= \langle u, Fu \rangle . \end{aligned}$$

□

The second distinguishing feature is:

Proposition 1.2.3 (Heisenberg's inequality) *Let X and Y be observables and let ρ be a state in $\mathcal{H}$. Assume $\operatorname{Tr} \rho X = \operatorname{Tr} \rho Y = 0$. Then,*

$$\begin{aligned} \operatorname*{\mathbf{var}}_{\rho}(X) \operatorname*{\mathbf{var}}_{\rho}(Y) &\geq \left(\operatorname{Tr} \rho \, \frac{1}{2} \{X, Y\} \right)^2 + \left(\operatorname{Tr} \rho \, \frac{1}{2} \, i \, [X, Y] \right)^2 \\ &\geq \frac{1}{4} (\operatorname{Tr} \rho \, i \, [X, Y])^2, \end{aligned}$$

where

$$\begin{aligned} \{X, Y\} &\triangleq XY + YX; \text{ and} \\ [X, Y] &\triangleq XY - YX. \end{aligned}$$

Proof For $z \in \mathbb{C}$, we have

$$\operatorname{Tr} \rho(X + zY)^\dagger (X + zY) \geq 0.$$

If $z = re^{i\theta}$,

$$r^2 \operatorname{Tr} \rho Y^2 + 2r\Re e^{-i\theta} \operatorname{Tr} \rho YX + \operatorname{Tr} \rho X^2 \geq 0.$$

The left hand side is a degree-two polynomial in the variable r. Since, it is always non-negative, it can have at most one root. Thus, for all θ,

$$\begin{aligned}(\operatorname{Tr} \rho X^2)(\operatorname{Tr} \rho Y^2) &\geq (\Re e^{-i\theta} \operatorname{Tr} \rho YX)^2 \\ &\geq \left(\cos\theta \operatorname{Tr} \rho \frac{XY+YX}{2} + \sin\theta \operatorname{Tr} \rho i \frac{XY-YX}{2}\right)^2 \\ &= (x\cos\theta + y\sin\theta)^2,\end{aligned}$$

where $x \stackrel{\Delta}{=} \operatorname{Tr} \rho \ \frac{1}{2}\{X, Y\}$ and $y \stackrel{\Delta}{=} \operatorname{Tr} \rho \ \frac{i}{2}[X, Y]$. Note that the right hand side is maximum when $\cos\theta = \frac{x}{\sqrt{x^2+y^2}}$ and $\sin\theta = \frac{y}{\sqrt{x^2+y^2}}$ and the proposition follows.

□

Now we state the third distinguishing feature:

Extremal states (one-dimensional projections) are called *pure states.* The set of all pure states in an n-dimensional complex Hilbert space is a manifold of dimension $2n - 2$. (The set of all extremal probability distributions on a sample space of n points has cardinality n).

1.3 Measurements: von Neumann's Collapse Postulate

Suppose X is an observable (i.e. a Hermitian operator) with spectral decomposition

$$X = \sum_{\lambda \in \mathrm{Sp}(X)} \lambda E_\lambda.$$

Then, the measurement of X in the quantum state ρ yields the value λ with probability $\operatorname{Tr} \rho E_\lambda$. If the observed value is λ, then the state collapses to

$$\tilde{\rho}_\lambda = \frac{E_\lambda \rho E_\lambda}{\operatorname{Tr} \rho E_\lambda}.$$

The collapsed state $\tilde{\rho}_\lambda$ has its support in the subspace $E_\lambda(\mathcal{H})$.

1.4 Dirac Notation

Elements of the Hilbert space $\mathcal{H}$ are called *ket vectors* and denoted by $|u\rangle$. Elements of the dual space $\mathcal{H}^*$ are called *bra vectors* and denoted by $\langle u|$. The bra $\langle u|$ evaluated on the ket $|v\rangle$ is the bracket $\langle u \mid v\rangle$, the scalar product between u, v as elements of $\mathcal{H}$.

The operator $|u\rangle\langle v|$ is defined by

$$|u\rangle\langle v|(|w\rangle) \triangleq \langle v \mid w\rangle\, |u\rangle.$$

It is a rank one operator when u and v are non-zero.

$$\begin{aligned} \operatorname{Tr} |u\rangle\langle v| &= \langle v \mid u\rangle \\ (|u\rangle\langle v|)^\dagger &= |v\rangle\langle u| \\ |u_1\rangle\langle v_1||u_2\rangle\langle v_2|\cdots|u_n\rangle\langle v_n| &= (\langle v_1 \mid u_2\rangle\, \langle v_2 \mid u_3\rangle \cdots \langle v_{n-1} \mid u_n\rangle)|u_1\rangle\langle v_n|. \end{aligned}$$

The scalar product $\langle u \mid v\rangle$ is anti-linear (conjugate-linear) in the first variable and linear in the second variable.

1.4.1 Qubits

The Hilbert space $\mathbf{h} \triangleq \mathbb{C}^2$, with scalar product $\left\langle \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right], \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] \right\rangle = \bar{a}c + \bar{b}d$, is called a 1-*qubit Hilbert space*. Let

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then,

$$\begin{bmatrix} a \\ b \end{bmatrix} = a|0\rangle + b|1\rangle,$$

and the ket vectors $|0\rangle$ and $|1\rangle$ form an orthonormal basis for $\mathbf{h}$.

The Hilbert space $\mathbf{h}^{\otimes n} = (\mathbb{C}^2)^{\otimes n}$ is called the n-*qubit* Hilbert space. If $x_1x_2\cdots x_n$ is an n-length word from the binary alphabet $\{0, 1\}$, we let

$$\begin{aligned} |x_1x_2\cdots x_n\rangle &\triangleq |x_1\rangle|x_2\rangle\cdots|x_n\rangle \\ &\triangleq |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_n\rangle \\ &\triangleq |\mathbf{x}\rangle, \end{aligned}$$

where $\mathbf{x} = x_1 \times 2^{n-1} + x_2 \times 2^{n-2} + \cdots + x_{n-1} \times 2 + x_n$ (that is, as $x_1x_2\ldots x_n$ varies over all n-length words, the integer $\mathbf{x}$ varies in the range $\{0, 1, \ldots, 2^n - 1\}$).

Lecture 2

Quantum Gates and Circuits

2.1 Gates in n-qubit Hilbert Spaces

In ordinary (classical) computers, information is passed through a classical channel. Logic gates (like AND, OR, NOT) operate on these channels. Likewise, in a quantum computer, information is passed through a quantum channel and it is operated upon by quantum gates. A *quantum gate* is a unitary operator U in a (finite dimensional) Hilbert Space $\mathcal{H}$.

Not all the classical gates are reversible (for example if a AND $b = 0$, there are three possible values for the ordered pair (a, b)). On the contrary, all quantum gates are reversible.

If a gate U acts on an n-qubit Hilbert space $\mathcal{H}$ we depict it as in Figure 2.1. If U acts on a single qubit it is represented pictorially as shown in Figure 2.2.

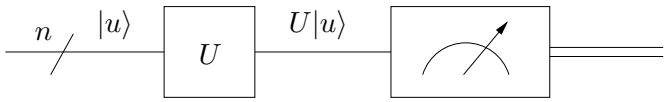

Figure 2.1: *A quantum circuit.*

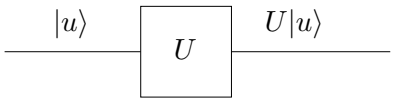

Figure 2.2: *A gate U acting on a single qubit.*

If the input is $|u\rangle$ and it passes through the gate U, then the output is written as $U|u\rangle$.

Any unitary operator U which acts on a single qubit can be written as

$$U = e^{i\alpha} \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix},$$

where $|a|^2 + |b|^2 = 1$ in the computational basis consisting of $|0\rangle$ and $|1\rangle$.

The action of the unitary operator U on the basis states can be computed as shown below.

$$U|0\rangle = e^{i\alpha} \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = e^{i\alpha}\{a|0\rangle - \overline{b}|1\rangle\}.$$

Similarly, $U|1\rangle = e^{i\alpha}\{b|0\rangle + \overline{a}|1\rangle\}$. By *measurement* on the n-qubit register of a quantum computer we usually mean measuring the observable

$$X = \sum_{j=0}^{2^n-1} j|j\rangle\langle j|,$$

and it is indicated in circuits by the ammeter symbol, as in Figure 2.1. Since by measuring we get *two* quantities, namely a classical value and a (collapsed) quantum state, pictorially it is indicated by a double line, as in Figure 2.1. The output consists of a value of X in the range $\{0, 1, \ldots, 2^n - 1\}$, where the probability of the event $\{X = j\}$ is $|\langle j|U|u\rangle|^2$, and a collapsed basis state $|j\rangle$, where j is the observed value.

As an example, let us simulate a Markov chain using a quantum circuit. Consider the circuit in Figure 2.3.

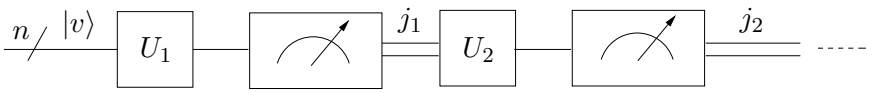

Figure 2.3: *A quantum circuit to simulate a Markov Chain.*

After each measurement, the observed classical parts $j_1, j_2, \ldots$ take

values in the space $\{0, 1, 2, \ldots, 2^n - 1\}$ with the following properties:

$$\begin{aligned}
\Pr(\{j_1\}) &= |\langle j_1|U_1|v\rangle|^2 && 0 \le j_1 \le 2^n - 1 \\
\Pr(\{j_2 \mid j_1\}) &= |\langle j_2|U_2|j_1\rangle|^2 && 0 \le j_2 \le 2^n - 1 \\
&\vdots && \vdots \\
\Pr(\{j_k \mid j_{k-1}j_{k-2}, \ldots, j_1\}) &= |\langle j_k|U_k|j_{k-1}\rangle|^2 && 0 \le j_k \le 2^n - 1 \\
&\vdots && \vdots
\end{aligned}$$

Thus, we have simulated a classical Markov chain with state space $\{0, 1, 2, \ldots, 2^n - 1\}$. The drawback here is that we need a separate unitary operator for each of the 2^n possible outcomes of the measurement.

Problem Given a doubly stochastic matrix M of size $n \times n$, does there exist a unitary matrix U such that, $|u_{ij}|^2 = p_{ij}$ for all $i, j \in \{0, 1, 2, \ldots, n\}$?

Existence of such a matrix will result in simplification of the quantum circuit for simulating a Markov chain.

2.2 Quantum Gates

2.2.1 One qubit gates

In classical computing, the only interesting one-bit gate is the NOT gate. In the quantum world, we have many 1-qubit gates. Some of them are given below.

1. Pauli gates: There are three such gates and they are denoted by X, Y, Z. The unitary matrices of X, Y, Z in the computational basis are given by

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

   The unitary matrix X is also called the *not gate* because $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$.

   These gates are called *Pauli gates* because the unitary matrices corresponding to these operators are the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$

of quantum mechanics. Pauli matrices are the basic spin observables taking values ± 1. X, Y, Z are hermitian, $X^2 = Y^2 = Z^2 = 1$ and X, Y, Z anticommute with each other i.e. $XY + YX = 0$.

2. Hadamard gate: The unitary matrix corresponding to the *Hadamard gate is* $H = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$. In this case, $H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $H|1\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$. Its n-fold tensor product $H^{\otimes^n}$ is the Hadamard gate on n-qubits satisfying

$$H^{\otimes^n}|00\ldots0\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_{x\in\{0,1\}^n} |x\rangle$$

and more generally

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

where $x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$.

3. Phase gate: The unitary matrix for this gate is $S = \left[\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right]$. This gate changes the phase of the ket vector $|1\rangle$ by i so that $|1\rangle$ becomes $i|1\rangle$, and leaves the ket vector $|0\rangle$ fixed.

4. $\frac{\pi}{8}$ gate: The unitary matrix for this gate is

$$T = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix} = e^{i\frac{\pi}{8}} \begin{bmatrix} e^{\frac{-i\pi}{8}} & 0 \\ 0 & e^{\frac{i\pi}{8}} \end{bmatrix}.$$

This gate changes the phase of $|1\rangle$ by $e^{i\frac{\pi}{4}}$.

2.2.2 Two qubit gates

1. Controlled NOT: This gate (Figure 2.4) acts as a NOT gate on the second qubit (target qubit) if the first qubit (control qubit) is in the computational basis state $|1\rangle$. So the vectors $|01\rangle$ and $|00\rangle$ are unaltered, while the vector $|10\rangle$ gets modified into $|11\rangle$ and vice versa.

   The unitary matrix for this gate is

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Figure 2.4: *Two qubit gates. A CNOT gate and a SWAP gate.*

The gate could also negate the content of the first qubit depending on the second qubit. Such a gate will have a different unitary matrix. The essential point is that a qubit can get negated depending on a control qubit. The control qubit will always be denoted by a solid dot in pictures.

2. Swap gate:

   This gate (Figure 2.4) swaps the contents of the two qubits. Because the vectors $|00\rangle$ and $|11\rangle$ are symmetric, they are unaltered, while the vector $|01\rangle$ gets mapped to $|10\rangle$ and vice versa.

   The unitary matrix for this gate is

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercise 2.2.1 Prove that the two circuits given in Figure 2.5 are equivalent.

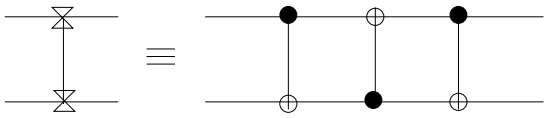

Figure 2.5: *Swap gate as a composition of three CNOT gates.*

Solution To check the equivalence of the circuits on the left hand side and right hand side we compute how the circuit on the right hand side acts on the basis state $|a, b\rangle$.

$$|a, b\rangle \to |a, a \oplus b\rangle \to |a \oplus (a \oplus b), a \oplus b\rangle = |b, a \oplus b\rangle \to |b, (a \oplus b) \oplus b\rangle = |b, a\rangle.$$

3. Controlled unitary: This is just like the controlled NOT, but instead of negating the target qubit, we perform the unitary transform prescribed by the matrix U (only if the control qubit is in state $|1\rangle$). It is represented schematically as shown in the first diagram of Figure 2.6.

2.2.3 Three qubit gates

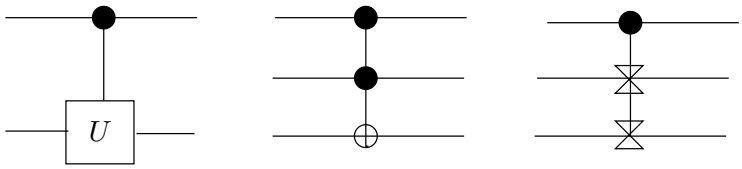

Figure 2.6: *A controlled unitary gate, Toffoli gate and a Fredkin gate.*

1. Toffoli gate: This (as in second diagram of Figure 2.6) is a double controlled NOT gate. The only computational basis vectors which get changed are $|110\rangle$ and $|111\rangle$. The corresponding unitary matrix is

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

2. Fredkin gate: This is a controlled swap gate (last diagram of Figure 2.6). The corresponding unitary matrix is

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

2.2.4 Basic rotations

We describe in this part, some basic rotation gates, each acting on a single qubit.

The basic rotation operators, which induce rotation by an angle θ about the x, y and z axis respectively, are denoted by $R_x(\theta), R_y(\theta)$ and $R_z(\theta)$ and they are defined by the following equations.

$$R_x(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} = e^{-\frac{i\theta X}{2}} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}X;$$

$$R_y(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} = e^{-\frac{i\theta Y}{2}} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Y;$$

$$R_z(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix} = e^{-\frac{i\theta Z}{2}} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Z.$$

More generally $R_{\hat{n}}(\theta) = (\cos\frac{\theta}{2})I - (i\sin\frac{\theta}{2})(\hat{n}_x X + \hat{n}_y Y + \hat{n}_z Z)$ is the matrix corresponding to rotation by an angle θ about the axis with direction cosines $(\hat{n}_x, \hat{n}_y, \hat{n}_z)$.

Theorem 2.2.2 (Euler) *Every 2×2 unitary matrix U can be expressed as*

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

□

Corollary 2.2.3 *Every 2×2 matrix U can be expressed as*

$$U = e^{i\alpha}AXBXC,$$

where A, B and C are 2×2 unitary operators and $ABC = I$.

Proof By Theorem 2.2.2 we can write

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

Set

$$A = R_z(\beta)R_y\left(\frac{\gamma}{2}\right), B = R_y\left(-\frac{\gamma}{2}\right)R_z\left(-\frac{\beta+\delta}{2}\right) \text{ and } C = R_z\left(\frac{\delta-\beta}{2}\right).$$

It is easy to check that A, B and C satisfy the required conditions.

□

Corollary 2.2.4 *In Figure 2.7, the circuit on the left hand side is equivalent to the circuit on the right hand side if* $AXBXC = e^{-i\alpha}U$, $ABC = I$ *and*

$$D = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix}.$$

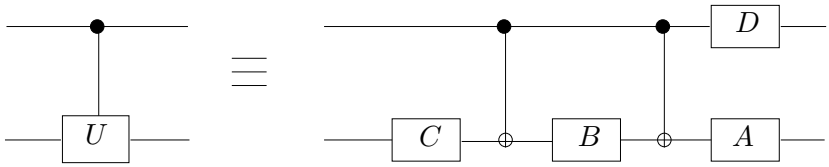

Figure 2.7: *Circuit implementing the controlled-U operation for single qubit* U. α, A, B *and* C *satisfy* $U = e^{i\alpha}AXBXC, ABC = I$.

Proof The equivalence of the circuits can be verified by checking how the computational basis states evolve.

$$\begin{aligned} |0\rangle|u\rangle &\to |0\rangle C|u\rangle \to |0\rangle BC|u\rangle \to |0\rangle ABC|u\rangle \to D|0\rangle ABC|u\rangle = |0\rangle|u\rangle. \\ |1\rangle|u\rangle &\to |1\rangle C|u\rangle \to |1\rangle XC|u\rangle \to |1\rangle BXC|u\rangle \to |1\rangle XBXC|u\rangle \\ &\to D|1\rangle AXBXC|u\rangle = e^{i\alpha}|1\rangle e^{-i\alpha}U|u\rangle = |1\rangle U|u\rangle. \end{aligned}$$

□

Corollary 2.2.5 *In Figure 2.8, the circuit on the left hand side is equivalent to the circuit on the right hand side if* $V^2 = U$.

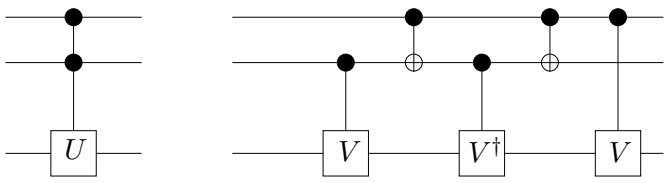

Figure 2.8: *Circuit for the* $C^2(U)$ *gate.* V *is any unitary operator satisfying* $V^2 = U$. *The special case* $V = (1-i)(I+iX)/2$ *corresponds to the Toffoli gate.*

Proof

$|00\rangle|u\rangle \rightarrow |00\rangle|u\rangle$.

$|01\rangle|u\rangle \rightarrow |01\rangle V|u\rangle \rightarrow |01\rangle V^\dagger V|u\rangle = |01\rangle I|u\rangle = |01\rangle|u\rangle$.

$|10\rangle|u\rangle \rightarrow |11\rangle|u\rangle \rightarrow |11\rangle V^\dagger|u\rangle \rightarrow |10\rangle V^\dagger|u\rangle \rightarrow |10\rangle VV^\dagger|u\rangle = |10\rangle|u\rangle$.

$|11\rangle|u\rangle \rightarrow |11\rangle V|u\rangle \rightarrow |10\rangle V|u\rangle \rightarrow |11\rangle V|u\rangle \rightarrow |11\rangle VV|u\rangle = |11\rangle U|u\rangle$.

□

Corollary 2.2.6 *A Toffoli gate can be expressed as a composition of controlled NOT's and 1-qubit gates.*

Proof Follows from the previous two corollaries.

□

Exercise 2.2.7 Derive and verify that the circuit on the right hand side of Figure 2.9 is a correct realization of the Toffoli gate using controlled NOT and single qubit gates.

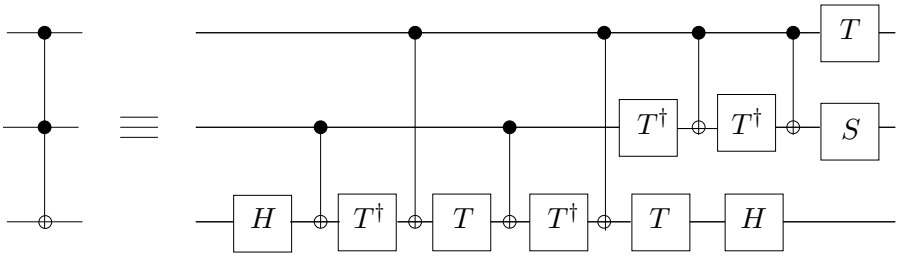

Figure 2.9: *Implementation of the Toffoli gate using Hadamard, phase, controlled NOT and* $\frac{\pi}{8}$ *gates.*

2.3 Some Simple Circuits

2.3.1 Quantum teleportation

In quantum teleportation, Alice (sender) can send a qubit to Bob (receiver) without using a quantum communication channel. In order to achieve this, Alice and Bob together generate an EPR pair (i.e. $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$) and share one qubit each.

Suppose Alice wants to send an unknown qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Then she cannot even measure it because she has only one copy of it. Even if Alice knows the state of the qubit $|\psi\rangle$ sending it to Bob through classical channel will not be possible at all. But by making use of the EPR pair Alice can send the qubit $|\psi\rangle$ to Bob just by sending two additional classical bits of information.

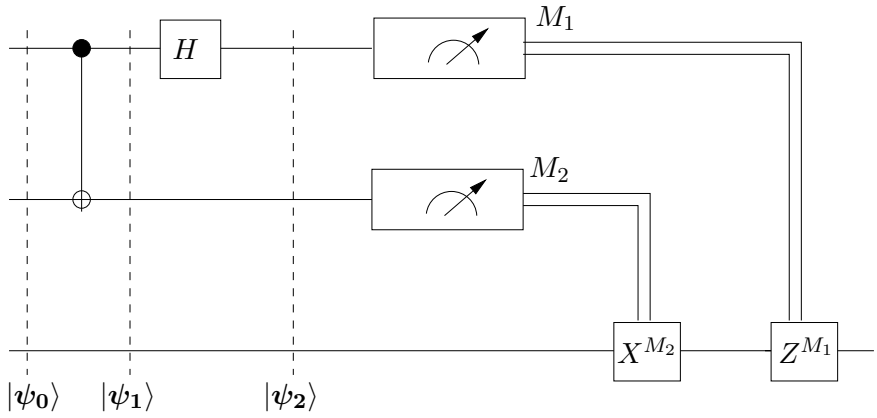

Figure 2.10: *Circuit used by Alice and Bob*

To accomplish the task Alice makes a circuit as shown in Figure 2.10. Alice has access to the top two qubits. So all operations Alice does involve only the top two qubits.

The initial state of the system is

$$|\psi_0\rangle = |\psi\rangle\frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)].$$

After the first CNOT gate the state of the system is

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)].$$

After she sends the first qubit through the Hadamard gate the state of the system is

$$|\psi_2\rangle = \frac{1}{2}[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)].$$

Collecting the first two qubits the state $|\psi_2\rangle$ can be re-written as

$$|\boldsymbol{\psi_2}\rangle = \frac{1}{2}[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)].$$

When Alice makes a measurement on the two qubits she can control, the state of Bob's qubit is completely determined by the results of Alice's measurement on her first two qubits. Hence if Alice sends the results of her measurement to Bob, he can apply appropriate gates on the qubit he can access and get the state $|\psi\rangle$. The action of Bob can be summarized as in the table below.

Alice measures	State of Bob's qubit	Gates needed to get $\|\psi\rangle$
00	$[\alpha\|0\rangle + \beta\|1\rangle]$	I
01	$[\alpha\|1\rangle + \beta\|0\rangle]$	X
10	$[\alpha\|0\rangle - \beta\|1\rangle]$	Z
11	$[\alpha\|1\rangle - \beta\|0\rangle]$	ZX

Thus, the state of the first qubit $|\psi\rangle$ is transferred to the third qubit which is with Bob. The above algorithm implies that one shared EPR pair and two classical bits of communication is a resource at least equal to one qubit of quantum communication.

2.3.2 Superdense coding: quantum communication through EPR pairs

If Alice and Bob initially share an EPR pair, Alice can send Bob two bits of classical information by passing a single qubit as follows. Alice makes a circuit as shown in Figure 2.11.

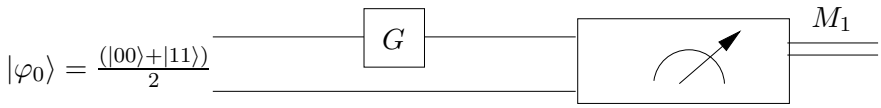

Figure 2.11: *Circuit used by Alice and Bob*

Alice selects the gate G according to the bits she wants to send. She selects a gate according to the table below and applies it to the qubit she possesses before transmitting it to Bob.

Bits to be sent	Gates to be used	Bob receives
00	I	$\frac{\|00\rangle+\|11\rangle}{\sqrt{2}}$
01	Z	$\frac{\|00\rangle-\|11\rangle}{\sqrt{2}}$
10	X	$\frac{\|10\rangle+\|01\rangle}{\sqrt{2}}$
11	iY	$\frac{\|01\rangle-\|10\rangle}{\sqrt{2}}$

The four possible states that Bob can receive are the so-called *Bell states* or *EPR pairs* which constitute the *Bell basis*. Since the Bell states form an orthogonal basis, they can be distinguished by measuring in the appropriate basis. Hence when Bob receives the qubit sent by Alice he has both the qubits. Then he does a measurement in the Bell basis and finds out the message she wanted to send. In classical computation it is impossible to send two bits of information by just passing a single bit. So a qubit can carry more than one bit of classical information.

2.3.3 A generalization of "communication through EPR states"

Let F be a finite abelian group of order n for example $(\mathbb{Z}/2\mathbb{Z})^k$ with $n = 2^k$. Let $\hat{F}$ denote its character group. Define the Hilbert space $\mathcal{H} \stackrel{\Delta}{=} L^2(F)$ to be the space of functions from F to $\mathbb{C}$ under the standard inner product. The characteristic functions of elements of the group F, $1_{\{x\}}$ where $x \in F$, form the standard orthonormal basis for $\mathcal{H}$. Define $|x\rangle \stackrel{\Delta}{=} 1_{\{x\}}$. Let $f \in \mathcal{H}$ and $x \in F$. For $a \in F$ and $\alpha \in \hat{F}$, define unitary operators U_a and V_α on $\mathcal{H}$ as

$$(U_a\, f)(x) \stackrel{\Delta}{=} f(x+a), \quad (V_\alpha\, f)(x) = \alpha(x)\, f(x).$$

U_a can be thought of as translation by the group element a and V_α can be thought of as multiplication by the character α. For $(a, \alpha) \in F \times \hat{F}$, define the *Weyl operator* $W_{a,\alpha} \stackrel{\Delta}{=} U_a V_\alpha$. It is a unitary operator.

Exercise 2.3.1 $W_{a,\alpha}W_{b,\beta} = \overline{\alpha(b)}W_{a+b,\alpha\beta}$. i.e. the $W_{a,\alpha}$ form a projective unitary representation of the group $F \times \hat{F}$. The term projective is used to refer to the fact that the unitary operators $W_{a,\alpha}$ form a representation of $F \times \hat{F}$ upto multiplication by a complex scalar (the number $\overline{\alpha(b)}$) of modulus unity.

Exercise 2.3.2 Show that the only linear operators which commute with $W_{a,\alpha}$ for all $(a,\alpha) \in F \times \hat{F}$, are the scalars. Hence, the $W_{a,\alpha}$'s form an irreducible projective representation of the group $F \times \hat{F}$, i.e. the only subspaces of $\mathcal{H}$ which are invariant under every $W_{a,\alpha}$ are the zero subspace and $\mathcal{H}$ itself.

Exercise 2.3.3 Show that the operators $\{W_{a,\alpha}\}_{(a,\alpha)\in F\times\hat{F}}$ are linearly independent. Thus, they span the space $\mathcal{B}(\mathcal{H})$ of (bounded) linear operators on $\mathcal{H}$.

Exercise 2.3.4 Show that $W_{a,\alpha}^\dagger = \alpha(a)W_{-a,\overline{\alpha}}$. Show also that $\operatorname{Tr} W_{a,\alpha} = n$ if $a = 0$ and α is the trivial character, where $n = |F|$; otherwise $\operatorname{Tr} W_{a,\alpha} = 0$. Hence, prove that $\operatorname{Tr} W_{a,\alpha}^\dagger W_{b,\beta} = n\delta_{(a,\alpha),(b,\beta)}$.

Exercise 2.3.5 Define

$$|\psi_0\rangle \stackrel{\Delta}{=} \frac{1}{\sqrt{n}} \sum_{x\in F} |x\rangle|x\rangle.$$

Also define $|(a,\alpha)\rangle \stackrel{\Delta}{=} (W_{a,\alpha} \otimes I)|\psi_0\rangle$, where I is the identity operator on $\mathcal{H}$. Then, $\{|(a,\alpha)\rangle\}_{(a,\alpha)\in F\times\hat{F}}$ is an orthonormal basis for $\mathcal{H} \otimes \mathcal{H}$.

Enumerate (a,α) as $f(a,\alpha) \in \{1,2,\ldots,n^2\}$, in some order. Define the Hermitian measurement operator

$$X \stackrel{\Delta}{=} \sum_{(a,\alpha)\in F\times\hat{F}} f(a,\alpha)\, |(a,\alpha)\rangle\langle(a,\alpha)| \,.$$

$|\psi_0\rangle$ is the entangled state which Alice and Bob share. Alice holds the first $\log n$ qubits of the state while Bob holds the other $\log n$ qubits. To send a message $m \in [n^2]$, Alice applies the unitary transformation $W_{a,\alpha}$, where $f(a,\alpha) = m$, on her qubits. She then sends her qubits to Bob, who then applies the measurement X on the $2\log n$ qubits which he now has. The outcome of the measurement is m, which is exactly what Alice intended to send. Thus Alice has communicated $2\log n$ classical bits of information using only $\log n$ qubits of quantum communication.

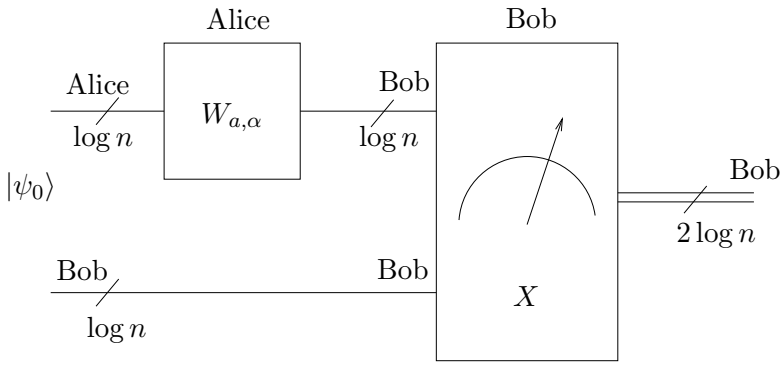

Figure 2.12: *Circuit used by Alice and Bob*

Exercise 2.3.6 In the case where $F = \mathbb{Z}/2\mathbb{Z}$, this reduces to communicating two classical bits at a time using one qubit, by the usual superdense coding technique!

2.3.4 Deutsche algorithm

This algorithm enables us to find out whether a function $f : \{0,1\} \to \{0,1\}$, is a constant function or not, by computing the function only once. In classical theory of computation we must evaluate the function twice before making such a conclusion.

Corresponding to the function f we consider the unitary operator U_f, where $U_f|xy\rangle = |x\rangle|y \oplus f(x)\rangle$, $x, y \in \{0, 1\}$. The circuit for implementing the algorithm is shown in Figure 2.13.

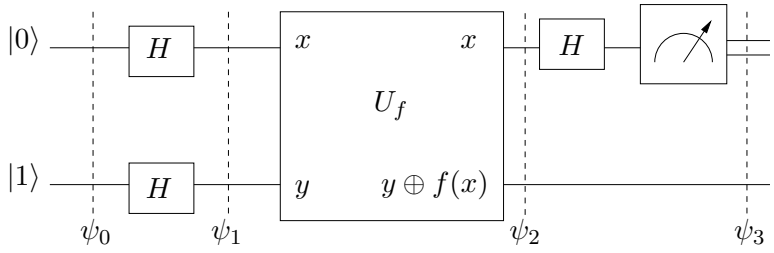

Figure 2.13: *Circuit for implementing Deutsche Algorithm.*

We follow the evolution of the circuit in Figure 2.13.

$$|\psi_0\rangle = |01\rangle$$
$$|\psi_1\rangle = \frac{1}{2}\left(|0\rangle + |1\rangle\right)\left(|0\rangle - |1\rangle\right).$$

Observe that $U_f|x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) = (-1)^{f(x)}|x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$.

$$|\psi_2\rangle = \begin{cases} \pm\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) & \text{if } f(0) = f(1); \\ \pm\frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) & \text{otherwise.} \end{cases}$$
$$|\psi_3\rangle = \begin{cases} \pm|0\rangle\frac{(|0\rangle-|1\rangle)}{\sqrt{2}} & \text{if } f(0) = f(1); \\ \pm|1\rangle\frac{(|0\rangle-|1\rangle)}{\sqrt{2}} & \text{otherwise.} \end{cases}$$
$$\implies |\psi_3\rangle = \pm|f(0) \oplus f(1)\rangle\frac{(|0\rangle - |1\rangle)}{\sqrt{2}}.$$

Thus, by measuring the first bit we get

$$\left\{\{f(0) \oplus f(1)\}, \pm|f(0) \oplus f(1)\rangle\frac{(|0\rangle - |1\rangle)}{\sqrt{2}}\right\}.$$

In this algorithm, both superposition and interference of quantum states are exploited.

2.3.5 Arithmetical operations on a quantum computer

We now see how addition may be performed on a quantum computer. Let x, y be two $n+1$ bit integers. Then we have

x	$=$		a_n	a_{n-1}	$\ldots$	a_0
y	$=$		b_n	b_{n-1}	$\ldots$	b_0
$x+y$	$=$	c_n	s_n	s_{n-1}	$\ldots$	s_0

and

$x' =$		a_{n-1}	a_{n-2}	$\ldots$	a_0
$y' =$		b_{n-1}	a_{n-2}	$\ldots$	b_0
$x'+y' =$	c_{n-1}	s_{n-1}	s_{n-2}	$\ldots$	s_0

Note that $s_0, s_1, \ldots s_{n-1}$ are same in both these additions. Also,

$$(c_n, s_n) = (a_n b_n \oplus c_{n-1}(a_n \oplus b_n), a_n \oplus b_n \oplus c_{n-1}).$$

Note that the Toffoli gate sends $|abc\rangle \rightarrow |ab\rangle|c \oplus ab\rangle$.

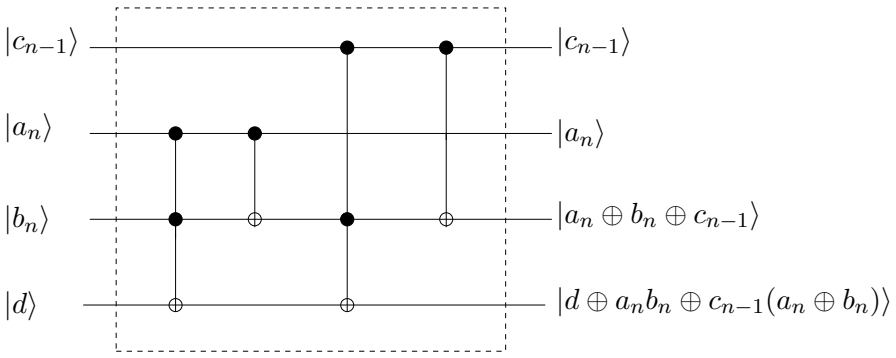

Figure 2.14: *Circuit for adding two single bit numbers with carry.*

Consider a subroutine for adding two single bit numbers with carry. The circuit for this subroutine is shown in Figure 2.14.

If we measure the last two qubits in the circuit in Figure 2.14, we get the outputs $\{s_n\}, \{c_n\}$ and the collapsed states $|s_n\rangle$, $|c_n\rangle$ provided $d = 0$. Hence, using this subroutine we can add two n-bit numbers.

Addition:

We would like to count the number of Toffoli and CNOT gates used by the circuit as a measure of complexity. Suppose α_n Toffoli and β_n CNOT gates are used for adding two n-bit numbers. Then

$$\begin{aligned} &\alpha_{n+1} = \alpha_n + 2, && \beta_{n+1} = \beta_n + 2 \\ \implies\ &\alpha_n = \alpha_1 + 2(n-1), && \beta_n = \beta_1 + 2(n-1). \end{aligned}$$

Consider the circuit in Figure 2.15.

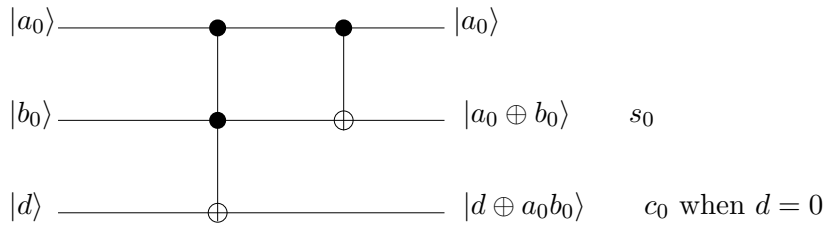

Figure 2.15: *Circuit for adding two single bit numbers without carry.*

Thus, $\alpha_1 = 1$ and $\beta_1 = 1$. This implies $\alpha_n = \beta_n = 2n - 1$. So by this method of adding two n bit numbers we need $2n - 1$ Toffoli and

$2n-1$ CNOT gates. The circuit for adding two n bit numbers is shown in Figure 2.16.

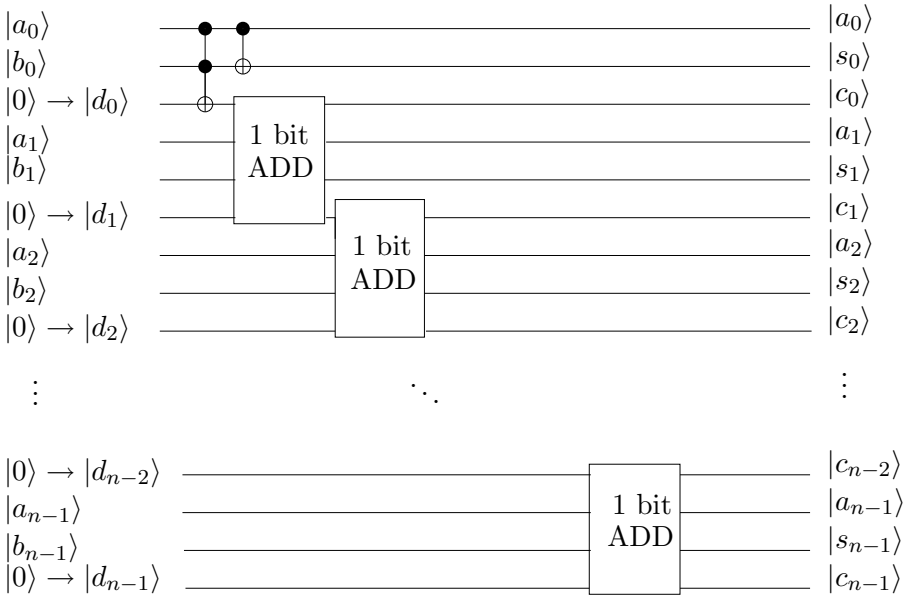

Figure 2.16: *Circuit for adding two n bit numbers without carry.*

Subtraction:

To evaluate $a-b$, where a, b are two n bit numbers, add a and $2^n - b$ to get $a + 2^n - b = e_n e_{n-1} \dots e_0$. Note that $2^n - b$ can be easily computed using only CNOT gates. If $e_n = 0$, then $a - b = -(1 \oplus e_{n-1})(1 \oplus e_{n-2}) \cdots (1 \oplus e_0)$. If $e_n = 1$, then $a - b = e_{n-1} e_{n-2} \dots e_0$.

Exercise 2.3.7 Count the number of gates required in the above subtraction algorithm.

Exercise 2.3.8 Device a circuit for addition (mod N), multiplication and division.

Lecture 3

Universal Quantum Gates

3.1 CNOT and Single Qubit Gates are Universal

In classical computation the AND, OR and NOT gates are universal which means that any boolean function can be realized using only these three gates. In this lecture, we prove the quantum analogue of this theorem. We show that any unitary transformation in an n-qubit Hilbert space can be approximated by compositions of Hadamard, CNOT, phase and $\pi/8$ gates to any desired degree of accuracy. We proceed by proving two propositions from which the theorem immediately follows.

Lemma 3.1.1 *Any $n \times n$ unitary matrix U can be expressed as a product of at most one phase factor and $\frac{n(n-1)}{2}$ unitary matrices, each of which acts on a 2-dimensional coordinate plane.*

Proof Let $U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \cdot & \cdot & \dots & \cdot \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{bmatrix}$.

If $u_{21} = 0$, do nothing. Otherwise, left multiply by a unitary matrix

$$U_1 = \left[\begin{array}{cc|c} \alpha & \beta & 0 \\ -\overline{\beta} & \overline{\alpha} & \\ \hline 0 & & I_{n-2} \end{array}\right]$$

such that $-\overline{\beta}u_{11} + \overline{\alpha}u_{21} = 0$ and $|\alpha|^2 + |\beta|^2 = 1$. Solving we get

$$\alpha = \frac{\overline{u}_{11}}{\sqrt{|u_{11}|^2 + |u_{21}|^2}} \text{ and } \beta = \frac{\overline{u}_{21}}{\sqrt{|u_{11}|^2 + |u_{21}|^2}}.$$

Now consider $M^1 = U_1 U$. The $M^1(2,1)$ entry is 0. If $M^1(3,1)$ is 0, we do nothing. Otherwise we left multiply by U_2 in the $(1,3)$ plane to make the entry $(3,1)$ in the resulting matrix 0. Continuing this way we get

$$U_{n-1}U_{n-2}\cdots U_1 U = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ 0 & v_{22} & \dots & v_{2n} \\ . & . & \dots & . \\ 0 & v_{n2} & \dots & v_{nn} \end{bmatrix}$$

where $|v_{11}| = 1$. Orthogonality between the 1^{st} and any other column shows that $v_{12} = v_{13} = \cdots = v_{1n} = 0$. Thus

$$v_{11}^{-1}U_{n-1}U_{n-2}\cdots U_1 U = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ 0 & & & \\ . & & \mathrm{W} & \\ . & & & \\ . & & & \\ 0 & & & \end{array}\right]$$

where W is an $n-1 \times n-1$ unitary matrix. The same procedure is repeated for the reduced matrix W. We repeat these operations till we get the identity matrix I. Pooling the phase factors we get $e^{i\alpha}U_m U_{m-1}\cdots U_1 U = I$ where $m \leq \binom{n}{2}$. It is to be noted that U_j is an element in $SU(2)$ acting in a two dimensional subspace. Transferring the U_j's to the right we get $U = e^{i\alpha}U_1^\dagger U_2^\dagger \cdots U_m^\dagger$.

□

Lemma 3.1.2 *Any matrix $U \in SU(2)$ acting in a 2-dimensional subspace can be realized using single qubit and r-controlled 1-qubit gates.*

Proof Consider $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ with computational basis $\{|x\rangle, x \in \{0,1\}^n\}$. Consider a pair x, y which differ in exactly one place, say i.

$$|x\rangle = |a\rangle|0\rangle|b\rangle,$$
$$|y\rangle = |a\rangle|1\rangle|b\rangle,$$

with a and b being words of length $i-1$ and $n-i$ respectively.

A unitary matrix U in the two dimensional plane spanned by $|x\rangle$ and $|y\rangle$ which leaves the other kets $|z\rangle$ fixed can be expressed as in Figure 3.1, with U replaced by $\tilde{U} = \begin{bmatrix} \alpha & \beta \\ -\beta & \overline{\alpha} \end{bmatrix}$ and $|\alpha|^2 + |\beta|^2 = 1$.

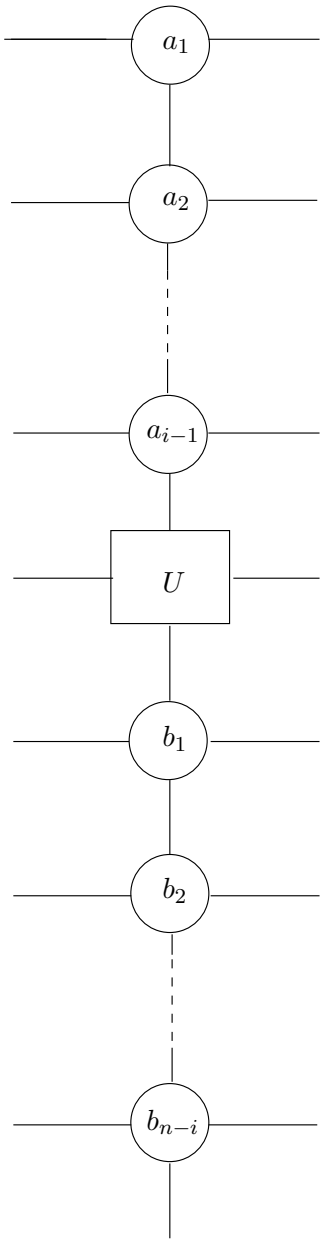

Figure 3.1: *A generalized controlled U operation on n-qubits.*

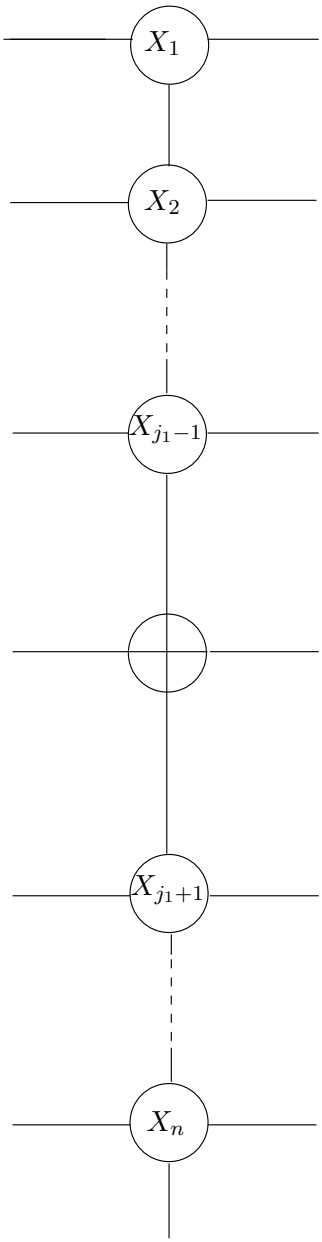

Figure 3.2: *A generalized controlled NOT operation on n-qubits.*

Suppose now x and y differ in r places. Then we can construct a sequence

$$x = x^{(0)}, x^{(1)}, x^{(2)}, \ldots, x^{(r-1)}, x^{(r)} = y$$

of n length words such that $x^{(i)}$ and $x^{(i+1)}$ differ exactly in one position for all $i = 0, 1, 2, \ldots, r-1$.

$$\begin{array}{lccl} \text{Let} & x, & x^{(1)} & \text{differ at position } j_1, \\ & x^{(1)}, & x^{(2)} & \text{differ at position } j_2, \\ & \vdots & \vdots & \vdots \\ \text{and} & x^{(r-1)}, & x^{(r)} & \text{differ at position } j_r. \end{array}$$

Now a controlled NOT gate (it is not the CNOT gate) is applied on x with the j_1 bit as target and the remaining $n-1$ bits as control bits. The NOT gate acts on the j_1 bit if the first bit is x_1, the second bit is x_2 and so on. This can be implemented with X (NOT) and CNOT gates as shown in the Figures 3.2 and 3.3.

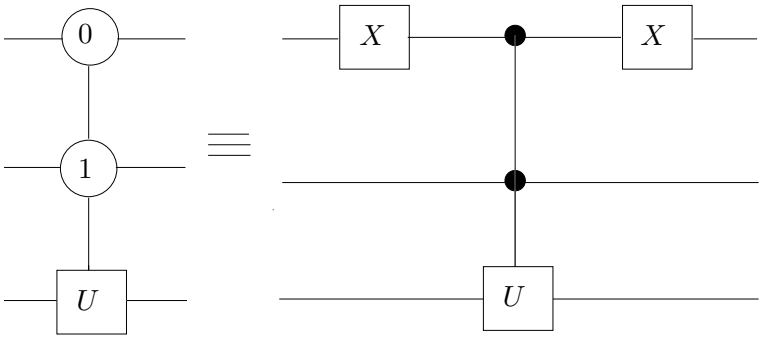

Figure 3.3: *Realizing a generalized controlled operation.*

We follow this by a controlled NOT on $x^{(1)}$ with j_2 as the target bit and the remaining $n-1$ as the control bits. After continuing this up to $x^{(r-1)}$, we apply $\tilde{U}$. Then we just do the reverse of the controlled NOT operations. This implements $\tilde{U}$ in the plane generated by $|x\rangle$ and $|y\rangle$ keeping all $|z\rangle$ fixed where z differs from both x and y.

Figure 3.3 shows how a generalized controlled 1-qubit gate can be realized using 1-qubit gates and r-controlled 1-qubit gate. This completes the proof.

□

Lemma 3.1.3 *If $n \geq 2$, then an n-controlled 1-qubit gate can be realized by $(n-1)$-controlled 1-qubit gates.*

Proof Let $U = V^2$ where $U, V \in SU(2)$. Then we see that the two circuits in Figure 3.4 are equivalent.

□

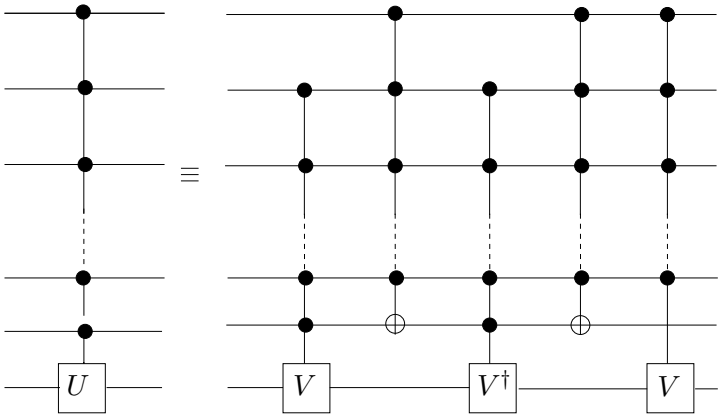

Figure 3.4: *n-controlled* 1*-qubit gate as a composition of five* $(n-1)$ *controlled* 1*-qubit gates.*

Exercise 3.1.4 Show that in Figure 3.4 the circuit on the left hand side is equivalent to the circuit on the right hand side.

Lemma 3.1.5 *A controlled 1-qubit gate can be realized using CNOT and single qubit gates.*

Proof Let $U = e^{i\alpha}AXBXC$, $ABC = I$, $D = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix}$. Then from Corollary 2.2.4 we know that the two circuits in Figure 3.5 are equivalent.

□

Proposition 3.1.6 *Any arbitary unitary matrix on an n-dimensional Hilbert space can be realized using phase, single qubit and CNOT gates.*

Proof The proof follows from Lemma 3.1.1, Lemma 3.1.2, Lemma 3.1.3 and Lemma 3.1.5.

□

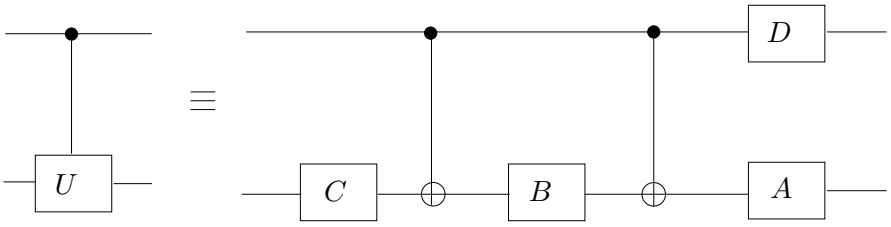

Figure 3.5: *Controlled* 1*-qubit gate as a composition of two CNOT and four* 1*-qubit gates.*

Proposition 3.1.7 *The group generated by* H *and* $e^{-i\frac{\pi}{8}Z}$ *is dense in* $SU(2)$.

Proof $H^2 = I$, $HZH = X$, $HYH = -Y$, $He^{-i\frac{\pi}{8}Z}H = e^{-i\frac{\pi}{8}X}$ and

$$\begin{aligned} e^{-i\frac{\pi}{8}Z}e^{-i\frac{\pi}{8}X} &= \cos^2\frac{\pi}{8}I - \Big(i\sin\frac{\pi}{8}\Big)\Big\{\Big(\cos\frac{\pi}{8}\Big)(X+Z) + \Big(\sin\frac{\pi}{8}\Big)Y\Big\} \\ &= R_{\vec{n}(\alpha)}, \end{aligned}$$

where $\cos\alpha = \cos^2\frac{\pi}{8}$, $\vec{n} = \frac{(\cos\frac{\pi}{8}, \sin\frac{\pi}{8}, \cos\frac{\pi}{8})}{\sqrt{1+\cos^2\frac{\pi}{8}}}$,

$$\begin{aligned} He^{-i\frac{\pi}{8}Z}e^{-i\frac{\pi}{8}X}H &= \cos^2\frac{\pi}{8}I - \Big(i\sin\frac{\pi}{8}\Big)\Big\{\Big(\cos\frac{\pi}{8}\Big)(X+Z) - \Big(\sin\frac{\pi}{8}\Big)Y\Big\} \\ &= R_{\vec{m}(\alpha)}, \end{aligned}$$

where, $\vec{m} = \frac{(\cos\frac{\pi}{8}, -\sin\frac{\pi}{8}, \cos\frac{\pi}{8})}{\sqrt{1+\cos^2\frac{\pi}{8}}}$. Now we need the following lemma.

Lemma 3.1.8 *If* $\cos\alpha = \cos^2\frac{\pi}{8}$, *then* α *is an irrational multiple of* π.

Proof See Appendix.

□

Any $R_{\vec{n}}(\theta)$ can be approximated as closely as we want by a suitable power of $R_{\vec{n}}(\alpha)$ because α is an irrational multiple of π. Similarly, any $R_{\vec{m}}(\phi)$ can be approximated by a suitable power of $R_{\vec{m}}(\alpha)$.

Since $\vec{n}$ and $\vec{m}$ are two linearly independent unit vectors, any $U \in SU(2)$ can be written as $U = e^{i\psi}R_{\vec{n}}(\theta_1)R_{\vec{m}}(\theta_2)R_{\vec{n}}(\theta_3)$. This is an immediate consequence of Euler's theorem (Theorem 2.2.2). This completes the proof.

□

Now we are ready for the main theorem.

Theorem 3.1.9 *The subgroup generated by the Hadamard gate H, phase gate, CNOT and the $\pi/8$ is dense in the unitary group $U(2)$.*

Proof Immediate from Proposition 3.1.6 and Proposition 3.1.7. □

3.2 Appendix

In this section we first give all the definitions and results needed to prove Lemma 3.1.8. The proofs which are routine are left out. The reader may refer to [3, 8] for a comprehensive treatment. We start with a few definitions.

A nonzero ring $\mathbf{R}$ with $1 \neq 0$ is called an *integral domain* if it has no zero divisors. In other words, it has the property that for $A, B \in \mathbf{R}$, if $AB = 0$, then $A = 0$ or $B = 0$.

An ideal is called *principal* if it is generated by a single element.

An integral domain in which every ideal is principal is called a *principal ideal domain.*

Exercise 3.2.1 Show that for any field k, $k[x]$ is a principal ideal domain.

An element P ($\neq \{0, 1\}$) of an integral domain $\mathbf{R}$ is called *prime* if the following is true: if P divides a product of two elements of $\mathbf{R}$, then it divides one of the factors.

A nonconstant polynomial $P \in \mathbb{F}[x]$ is called *irreducible* if it is written as a product of two polynomials $P_1, P_2 \in \mathbb{F}[x]$ then either P_1 or P_2 is a constant.

A polynomial is called *monic* if the coefficient of the leading term is 1.

A polynomial $a_0 + a_1 x + \cdots + a_n x^n$ in $\mathbb{Z}[x]$ is called *primitive* if g.c.d. $(|a_0|, \ldots, |a_n|) = 1$ and $a_n > 0$.

Remark 3.2.2 Every nonzero polynomial $P \in \mathbb{Q}[x]$ can be written as a product $P = cP_0$, where c is a rational number and P_0 is a primitive polynomial in $\mathbb{Z}[x]$. Note that this expression for P is unique and the polynomial P has integer coefficients if and only if c is an integer. In that case $|c|$ is the g.c.d. of the coefficients of P and c and the leading coefficient of P have the same sign.

The rational number c which appears in Remark 3.2.2 is called the *content* of P. If P has integer coefficients, then the content divides P in $\mathbb{Z}[x]$. Also, P is primitive if and only if its content is 1.

Lemma 3.2.3 *Let $\varphi : \mathbf{R} \longrightarrow \mathbf{R}'$ be a ring homomorphism. Then for any element $\alpha \in \mathbf{R}'$, there is a unique homomorphism $\Phi : \mathbf{R}[x] \longrightarrow \mathbf{R}'$ which agrees with the map φ on constant polynomials and sends $x \rightsquigarrow \alpha$.*

Let $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Lemma 3.2.3 gives us a homomorphism $\mathbb{Z}[x] \to \mathbb{F}_p$. This homomorphism sends a polynomial $P = a_m x^m + \cdots + a_0$ to its residue $\overline{P} = \overline{a_m} x^m + \cdots + \overline{a_0}$ modulo p.

Theorem 3.2.4 (Gauss's Lemma) *A product of primitive polynomials in $\mathbb{Z}[x]$ is primitive.*

Proof Let P and Q be two primitive polynomials in $\mathbb{Z}[x]$ and let R be their product. Obviously the leading coefficient of R is positive. To show that R is primitive, it is enough to show that no prime integer p divides all the coefficients of R. Consider the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$ defined above. Since P is primitive, its coefficients are not all divisible by p. So $\overline{P} \neq 0$. Similarly, $\overline{Q} \neq 0$. Since the polynomial ring $\mathbb{F}_p[x]$ is an integral domain, $\overline{R} = \overline{P}\,\overline{Q} \neq 0$. Therefore p does not divide one of the coefficients of R. This implies that R is primitive.

□

Proposition 3.2.5

1. *Let F, G be polynomials in $\mathbb{Q}[x]$, and let F_0, G_0 be the associated primitive polynomials in $\mathbb{Z}[x]$. If F divides G in $\mathbb{Q}[x]$, then F_0 divides G_0 in $\mathbb{Z}[x]$.*
2. *Let $F, G \in \mathbb{Z}[x]$ such that F is primitive and G is divisible by F in $\mathbb{Q}[x]$, say $G = FQ$, with $Q \in \mathbb{Q}[x]$. Then $Q \in \mathbb{Z}[x]$, and hence F divides G in $\mathbb{Z}[x]$.*
3. *Let F, G be polynomials in $\mathbb{Z}[x]$. If they have a common nonconstant factor in $\mathbb{Q}[x]$, then they have such a factor in $\mathbb{Z}[x]$ too.*

Proof To prove (1), we may clear denominators so that F and G become primitive. Then (1) is a consequence of (2). By Remark 3.2.2 we can write $Q = cQ_0$, where Q_0 is primitive and $c \in \mathbb{Q}$. By Gauss's Lemma, FQ_0 is primitive, and the equation $G = cFQ_0$ shows that it is the primitive polynomial Q_0 associated to Q. Therefore $Q = cQ_0$ is the expression for Q referred to in Lemma 3.2.2, and c is the content of Q.

Since c is the content of both G and Q, and $G \in \mathbb{Z}[x]$, it follows that $c \in \mathbb{Z}$, hence that $Q \in \mathbb{Z}[x]$. Now let us prove (3). Suppose that F, G have a common factor H in $\mathbb{Q}[x]$. We may assume that H is primitive, and then by (2) H divides both F and G in $\mathbb{Z}[x]$.

□

Corollary 3.2.6 *If a nonconstant polynomial F is irreducible in $\mathbb{Z}[x]$, then it is irreducible in $\mathbb{Q}[x]$.*

Proposition 3.2.7 *Let F be an integer polynomial with positive leading coefficient. Then F is irreducible in $\mathbb{Z}[x]$ if and only if either*

1. *F is a prime integer, or*
2. *F is a primitive polynomial which is irreducible in $\mathbb{Q}[x]$.*

Proof Suppose that F is irreducible. As in Remark 3.2.2, we may write $F = cF_0$, where F_0 is primitive. Since F is irreducible, this cannot be a proper factorization. So either c or F_0 is 1. If $F_0 = 1$, then F is constant, and to be irreducible a constant polynomial must be a prime integer. The converse is trivial.

□

Lemma 3.2.8 *In a principal ideal domain, an irreducible element is prime.*

Proof Let $\mathbf{R}$ be a principal ideal domain and F be an irreducible element in $\mathbf{R}$. Let $F|GH$, $G, H \in \mathbf{R}$. We assume that $F \nmid G$. Then the ideal generated by F and G is $\mathbf{R}$ (why?). Thus we may write $F_1F + G_1G = 1$ for some $F_1, G_1 \in \mathbf{R}$. This implies $F_1FH + G_1GH = H$. Hence $F|H$. This shows that F is prime.

□

Theorem 3.2.9 *Every irreducible element of $\mathbb{Z}[x]$ is a prime element.*

Proof Let F be irreducible, and suppose F divides GH, where G, $H \in \mathbb{Z}[x]$.

Case 1: $F = p$ is a prime integer. Write $G = cG_0$ and $H = dH_0$ as in Remark 3.2.2. Then G_0H_0 is primitive, and hence some coefficient a of G_0H_0 is not divisible by p. But since p divides GH, the corresponding coefficient, which is cda, is divisible by p. Hence p divides c or d, so p divides G or H.

Case 2: F is a primitive polynomial which is irreducible in $\mathbb{Q}[x]$. By Lemma 3.2.8, F is a prime element of $\mathbb{Q}[x]$. Hence F divides G or H in $\mathbb{Q}[x]$. By Proposition 3.2.5, F divides G or H in $\mathbb{Z}[x]$.

□

Lemma 3.2.10 *Let $F = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$ be an integer polynomial, and let p be a prime integer which does not divide a_n. If the residue $\overline{F}$ of F modulo p is irreducible, then F is irreducible in $\mathbb{Q}[x]$.*

Proof This follows from the natural homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$ (see Lemma 3.2.3). We may assume that F is primitive. Since p does not divide a_n, the degrees of $\overline{F}$ and F are equal. If F factors in $\mathbb{Q}[x]$, then it also factors in $\mathbb{Z}[x]$ by Corollary 3.2.6. Let $F = GH$ be a proper factorization in $\mathbb{Z}[x]$. Since F is primitive, G and H have positive degree. Since $\deg F = \deg \overline{F}$ and $\overline{F} = \overline{G}\,\overline{H}$, it follows that $\deg G = \deg \overline{G}$ and $\deg H = \deg \overline{H}$, hence that $\overline{F} = \overline{G}\,\overline{H}$ is a proper factorization, which shows that $\overline{F}$ is reducible.

□

Theorem 3.2.11 (Eisenstein criterion) *Let $F = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$ be an integer polynomial, and let p be a prime integer. Suppose that the coefficients of F satisfy the following conditions:*

1. *p does not divide a_n;*
2. *p divides other coefficients $a_{n-1}, \ldots a_0$;*
3. *p^2 does not divide a_0.*

Then F is irreducible in $\mathbb{Q}[x]$. If F is primitive, it is irreducible in $\mathbb{Z}[x]$.

Proof Assume F satisfies the hypothesis. Let $\overline{F}$ denote the residue modulo p. The conditions (1) and (2) imply that $\overline{F} = \overline{a}_n x^n$ and that $\overline{a}_n \neq 0$. If F is reducible in $\mathbb{Q}[x]$, then it will factor in $\mathbb{Z}[x]$ into factors of positive degree, say $F = GH$. Then $\overline{G}$ and $\overline{H}$ divide $\overline{a}_n x^n$, and hence each of these polynomials is a monomial. Therefore all coefficients of G and of H, except the highest ones are divisible by p. Let the constant coefficients of G, H be b_0, c_0. Then the constant coefficient of F is $a_0 = b_0 c_0$. Since p divides b_0 and c_0, it follows that p^2 divides a_0, which contradicts (3). This shows that F is irreducible. The last assertion follows from Proposition 3.2.7.

□

Corollary 3.2.12 *Let p be a prime. Then the polynomial $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ is irreducible in $\mathbb{Q}[x]$. (Such polynomials are called cyclotomic polynomials, and their roots are the p^{th} roots of unity.)*

Proof First note that $(x-1)f(x) = x^p - 1$. Now substituting $x = y+1$ into this equation we get

$$y\, f(y+1) = (y+1)^p - 1 = y^p + \binom{p}{1} y^{p-1} + \cdots + \binom{p}{p-1} y.$$

We have $\binom{p}{i} = p(p-1)\cdots(p-i+1)/i!$. If $i < p$, then the prime p is not a factor of $i!$, so $i!$ divides the product $(p-1)\cdots(p-i+1)$ This implies that $\binom{p}{i}$ is divisible by p. Dividing the expansion of $y\, f(y+1)$ by y shows that $f(y+1)$ satisfies the Eisenstein criterion and hence it is an irreducible polynomial. This implies that $f(x)$ is also irreducible. □

Theorem 3.2.13 *If $\cos\alpha = \cos^2\frac{\pi}{8}$, then α is an irrational multiple of π.*

Before proceeding to the proof of this theorem we shall establish a lemma.

Lemma 3.2.14 *Let $\lambda = \alpha/\pi$, where α is as in Theorem 3.2.13. Then $\beta = e^{2i\pi\lambda}$ is a root of the irreducible monic polynomial $m_\beta = x^4 + x^3 + \frac{1}{4}x^2 + x + 1$ (over $\mathbb{Q}[x]$).*

Proof Let m_β be the irreducible monic polynomial which has β as one of its roots. Note that $\sin 2\pi\lambda$ is not equal to zero. This means m_β has a complex root. Since its coefficients are rational it must also have the root $\overline{\beta}$. Thus, m_β must be divisible by $x^2 - 2Re\{\beta\} + 1$. Elementary computation shows that

$$2Re\{\beta\} = -\frac{1}{2} + \sqrt{2}.$$

So m_β is divisible by $p(x) = x^2 - (\sqrt{2} - \frac{1}{2})x + 1$. Since, $p(x)$ has irrational coefficients and m_β has rational coefficients, m_β must have another irrational root, say δ. This implies m_β has another quadratic factor with real coefficients. This means that $\deg(m_\beta) \geq 4$. Consider the polynomial $p'(x) = x^2 + (\sqrt{2} + \frac{1}{2})x + 1$. Multiplying $p(x)$ and $p'(x)$

we get $x^4 + x^3 + \frac{1}{4}x^2 + x + 1$. From the construction β is a root of the polynomial

$$m_\beta = x^4 + x^3 + \frac{1}{4}x^2 + x + 1,$$

which has no rational roots.

□

Proof of Theorem 3.2.13 Note that the polynomial $m_\beta(x)$ is not cyclotomic. Let us assume that λ is rational. Then $\beta = \frac{p}{q}$ is a root of the cyclotomic polymomial

$$\Phi_q(x) = x^{q-1} + x^{q-2} + \cdots + x + 1.$$

But $\Phi_q(x) = \prod_{p|q} \Phi_p(x)$, where p is prime. By Corollary 3.2.12 and Theorem 3.2.9 we know this is a prime factorization of $\Phi_q(x)$. Since, $m_\beta(x)$ is minimum irreducible polynomial and $\mathbb{Z}[x]$ is a unique factorization domain (follows from Theorem 3.2.9), $m_\beta(x)$ is prime. Thus, $m_\beta(x)$ must divide $\Phi_q(x)$. Hence, $m_\beta(x)$ must be a cyclotomic polynomial, a contradiction.

□

Lecture 4

The Fourier Transform and an Application

4.1 Quantum Fourier Transform

The quantum Fourier transform F on a finite dimensional Hilbert space $\mathcal{H}$ of dimension N is defined as a linear operator whose action on an orthonormal basis $|0\rangle, \ldots, |N-1\rangle$ is given by

$$F|j\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\frac{2\pi\, i\, j\, k}{N}} |k\rangle.$$

It can be easily verified that F defined as above is a unitary operator and the matrix of the transformation is $M(F) = [u_{jk}]$, where $u_{jk} = \frac{1}{\sqrt{N}} e^{\frac{2\pi\, i\, j\, k}{N}}$.

Theorem 4.1.1 *Let the dimension of the Hilbert space $\mathcal{H}$ be 2^n. Then the quantum Fourier transform F also has the following product representation. Let $j = j_1 2^{n-1} + j_2 2^{n-2} + \cdots + j_{n-1} 2 + j_n$. Then*

$$F\,|j\rangle = F\,|j_1 j_2 \ldots j_n\rangle = \frac{1}{2^{\frac{n}{2}}}(|0\rangle + e^{2\pi i 0 \cdot j_n}|1\rangle)(|0\rangle + e^{2\pi i 0 \cdot j_{n-1} j_n}|1\rangle)\cdots (|0\rangle + e^{2\pi i 0 \cdot j_1 j_2 \ldots j_n}|1\rangle).$$

Proof

$$\begin{aligned}
F|j\rangle &= \frac{1}{2^{\frac{n}{2}}} \sum_{k=0}^{2^n-1} e^{\frac{2\pi ijk}{2^n}} |k\rangle \\
&= \frac{1}{2^{\frac{n}{2}}} \sum_{k_1=0}^{1} \sum_{k_2=0}^{1} \cdots \sum_{k_n=0}^{1} e^{2\pi ij(\frac{k_1}{2^1}+\frac{k_2}{2^2}+\cdots+\frac{k_n}{2^n})} |k_1 k_2 \ldots k_n\rangle \\
&= \frac{1}{2^{\frac{n}{2}}} \sum_{k_1,k_2,\ldots,k_n} \otimes_{l=1}^{n} e^{\frac{2\pi ijk_l}{2^l}} |k_l\rangle \\
&= \frac{1}{2^{\frac{n}{2}}} \otimes_{l=1}^{n} (|0\rangle + e^{\frac{2\pi ij}{2^l}} |1\rangle)
\end{aligned}$$

since

$$\begin{aligned}
\frac{j}{2^l} &= \text{ integer } + \frac{j_{n-(l-1)}}{2} + \cdots + \frac{j_{n-1}}{2^{l-1}} + \frac{j_n}{2^l} \\
F|j\rangle &= \frac{1}{2^{\frac{n}{2}}} \otimes_{l=1}^{n} (|0\rangle + e^{2\pi i 0.j_{n-(l-1)} j_{n-(l-2)} \cdots j_n} |1\rangle) \\
&= \frac{1}{2^{\frac{n}{2}}} (|0\rangle + e^{2\pi i 0.j_n}|1\rangle)(|0\rangle + e^{2\pi i 0.j_{n-1} j_n}|1\rangle) \ldots \\
&\qquad (|0\rangle + e^{2\pi i 0.j_1 j_2 \ldots j_n}|1\rangle).
\end{aligned}$$

□

The circuit for implementing Fourier Transform on n-qubits is shown in Figure 4.1.

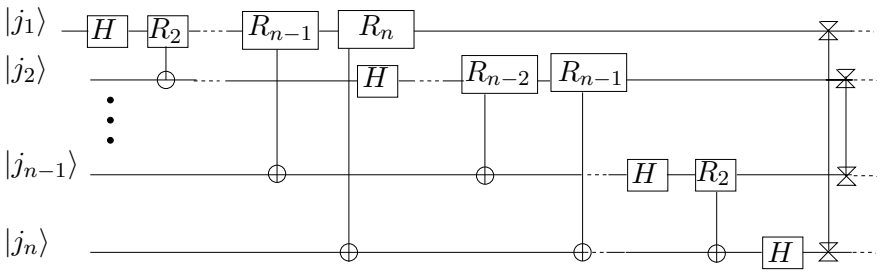

Figure 4.1: *Efficient circuit for quantum Fourier transform. The output on the k^{th} qubit from top is $\frac{1}{\sqrt{2}}(|0\rangle + e^{2\,\pi\, i\, 0.j_{n-k+1}\ldots j_n}|1\rangle)$. The correctness of the circuit follows from Theorem 4.1.1.*

In Figure 4.1, H represents the Hadamard gate and the unitary transform corresponding to the gate R_k is $\begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{2\pi i}{2^k}} \end{bmatrix}$. From the product representation it is easy to see that this circuit does compute the Fourier transform. To see how the circuit works we consider the input state $|j_1 j_2 \dots j_n\rangle$ and check how the system evolves. After the first Hadamard gate the state is

$$(H|j_1\rangle)|j_2 j_3 \dots j_n\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i \frac{j_1}{2}}|1\rangle)|j_2 j_3 \dots j_n\rangle.$$

After the controlled R_2 gate acting on the first qubit the state is

$$(R_2 H|j_1\rangle)|j_2 j_3 \dots j_n\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i(\frac{j_1}{2} + \frac{j_2}{2^2})}|1\rangle)|j_2 j_3 \dots j_n\rangle.$$

Hence, after the sequence of the controlled R_k's on the first qubit, the state is

$$\begin{aligned}(R_n R_{n-1} \dots R_2 H|j_1\rangle)|j_2 j_3 \dots j_n\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i(\frac{j_1}{2} + \frac{j_2}{2^2} + \dots \frac{j_n}{2^n})}|1\rangle)|j_2 j_3 \dots j_n\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0.j_1 j_2 \dots j_n}|1\rangle)|j_2 j_3 \dots j_n\rangle.\end{aligned}$$

Similarly, we can compute the action on the other qubits. The final state of the system is

$$\frac{1}{2^{\frac{n}{2}}}(|0\rangle + e^{2\pi i 0 \cdot j_1 j_2 \dots j_n}|1\rangle)(|0\rangle + e^{2\pi i 0 \cdot j_2 \dots j_n}|1\rangle) \dots (|0\rangle + e^{2\pi i 0 \cdot j_n}|1\rangle).$$

Now, if we perform the swap operation i.e. interchange the order of the qubits we get

$$\frac{1}{2^{\frac{n}{2}}}(|0\rangle + e^{2\pi i 0 \cdot j_n})(|0\rangle + e^{2\pi i 0 \cdot j_{n-1} j_n}) \dots (|0\rangle + e^{2\pi i 0 \cdot j_1 j_2 \dots j_n}),$$

which is exactly the quantum Fourier transform applied to $|j\rangle$. The number of Hadamard gates used is n and the number of controlled rotation gates used is $\frac{n(n-1)}{2}$. In the end at most $\lfloor \frac{n}{2} \rfloor$ swap gates are used. Therefore, this circuit uses $\Theta(n^2)$ gates. The best classical algorithm to compute Fourier transform on 2^n elements takes $\Theta(2^n(\log 2^n))$ gates. Thus to compute classical Fourier transform using classical gates takes exponentially more time to accomplish the task compared to computing quantum Fourier transform using a quantum computer.

Remark 4.1.2 This fact cannot be exploited very well because it is not possible to get access to the amplitudes in a quantum computer by measurements. Moreover, it is very difficult to obtain the initial state whose Fourier transform is to be computed. But quantum Fourier transform makes *phase estimation* "easy" which enables us to factor an integer efficiently in a quantum computer.

4.2 Phase Estimation

Let U be a unitary operator with eigenvector $|u\rangle$ and eigenvalue $e^{2\pi i\varphi}$ where $0 \leq \varphi \leq 1$. If $|u\rangle$ and controlled U^{2^j} are available then using Fourier transform one can efficiently estimate the phase φ. The circuit for the first stage of the phase estimation is shown below.

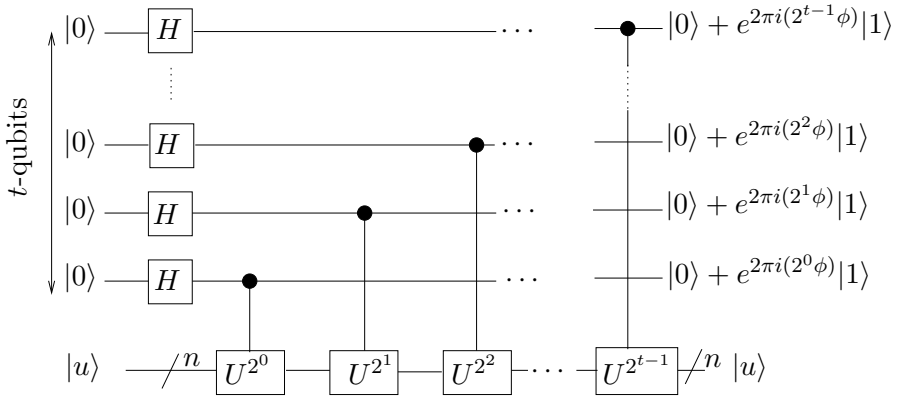

Figure 4.2: *First stage of the phase estimation circuit. Normalization factors of* $1/\sqrt{2}$ *have been omitted, on the right side.*

In the second stage of the phase estimation *inverse* Fourier transform is applied on some selected qubits and a measurement is done on those qubits in the computational basis. It will be shown that this yields a good estimate of the phase.

The first stage of the phase estimation uses two registers. The first register contains t qubits all in the state $|0\rangle$ and the second register contains n qubits in the state $|u\rangle$. The number of qubits t in the first register is chosen according to the accuracy and the probability of success required in the phase estimation procedure.

The final state after the first stage is

$$\frac{1}{2^{\frac{t}{2}}}(|0\rangle + e^{2\pi i 2^{t-1}\varphi}|1\rangle)(|0\rangle + e^{2\pi i 2^{t-2}\varphi}|1\rangle)\dots(|0\rangle + e^{2\pi i 2^{0}\varphi}|1\rangle)|u\rangle$$
$$= \frac{1}{2^{\frac{t}{2}}}\sum_{k=0}^{2^t-1} e^{2\pi i\varphi k}|k\rangle|u\rangle.$$

In the second stage *inverse* Fourier transform is applied on the first register (the first t qubits). This gives us a good estimate of φ. A schematic of the circuit is shown in Figure 4.3. To get a rough idea why this is true we consider the case when φ can be expressed exactly in t bits (in binary) $\varphi = 0\cdot\varphi_1\varphi_2\dots\varphi_t$. In this case the final state after stage one can be written as

$$\frac{1}{2^{\frac{t}{2}}}(|0\rangle + e^{2\pi i 0\cdot\varphi_t}|1\rangle)(|0\rangle + e^{2\pi i 0\cdot\varphi_{t-1}\varphi_t}|1\rangle)\dots(|0\rangle + e^{2\pi i 0\cdot\varphi_1\varphi_2\dots\varphi_t}|1\rangle)|u\rangle.$$

If we look at the product representation of the Fourier transform it is immediate that the above expression is the Fourier transform of the state $|\varphi_1\varphi_2\dots\varphi_t\rangle$. Hence measurement in the computational basis after the inverse Fourier transform will give the exact value of φ. If φ cannot be represented in t bits the observed value after measurement will be some $\tilde{\varphi}$. In the next section we analyze how good is $\tilde{\varphi}$ as an estimate of φ.

4.3 Analysis of the Phase Estimation Circuit

In this section we assume that $2^t\varphi$ is not an integer. We follow the evolution of the state $\underbrace{|0\rangle\dots|0\rangle}_{t}|u\rangle$ in the circuit depicted in Figure 4.3:

$$\begin{aligned}
\underbrace{|0\rangle\dots|0\rangle}_{t}|u\rangle &\to \frac{1}{2^{\frac{t}{2}}}\sum_{j=0}^{2^t-1}|j\rangle|u\rangle \\
&\to \frac{1}{2^{\frac{t}{2}}}\sum_{j}|j\rangle e^{2\pi ij\varphi}|u\rangle \\
&\to \frac{1}{2^t}\sum_{j,k} e^{\frac{-2\pi ijk}{2^t}+2\pi ij\varphi}|k\rangle|u\rangle \\
&= \frac{1}{2^t}\sum_{k}\frac{1-e^{2\pi i(\varphi-\frac{k}{2^t})2^t}}{1-e^{2\pi i(\varphi-\frac{k}{2^t})}}|k\rangle|u\rangle.
\end{aligned}$$

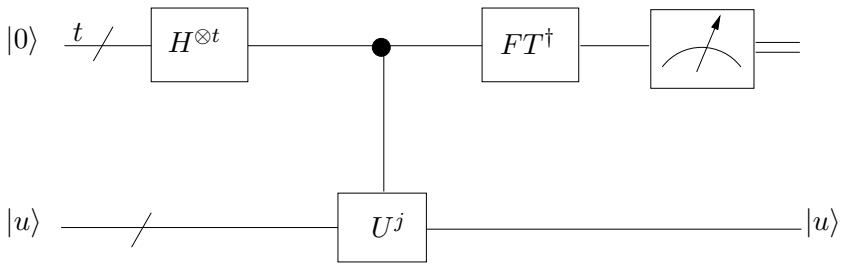

Figure 4.3: *The schematic for the overall phase estimation circuit.*

Hence measurement of the first register (the first t qubits) produces a random variable x with values in $X = \{0, 1, \ldots, 2^t - 1\}$ with

$$\begin{aligned}\Pr(x = k) &= \frac{1}{2^{2t}} \left| \frac{1 - e^{2\pi i(\varphi - \frac{k}{2^t})2^t}}{1 - e^{2\pi i(\varphi - \frac{k}{2^t})}} \right|^2 \\ &= \frac{1}{2^{2t}} \frac{\sin^2 \pi(\varphi - \frac{k}{2^t})2^t}{\sin^2 \pi(\varphi - \frac{k}{2^t})}.\end{aligned}$$

If the observed value is k, then $\frac{k}{2^t}$ is the desired estimate for φ. Let $a = \lfloor 2^t \varphi \rfloor$, $d = \varphi - \frac{a}{2^t}$ and δ be a positive integer $< 2^{t-1}$. We set $\frac{\delta}{2^t}$ to be the desired tolerance of error in the estimate of φ. In other words, the observed value of the random variable x should lie in

$$X_\delta = \{a - \delta + 1 (\text{mod } 2^t), a - \delta + 2 (\text{mod } 2^t), \ldots, a + \delta (\text{mod } 2^t)\}.$$

We now obtain a lower bound for the probability $\Pr(X_\delta)$.

We will need the following elementary fact which we leave as an exercise.

Exercise 4.3.1 Show that for any $\theta \in (0, \frac{\pi}{2}]$, $\frac{\sin\theta}{\theta} \geq \frac{2}{\pi}$.

Note that for $-2^{t-1} < k \leq 2^{t-1}$,

$$\begin{aligned}
\Pr(x = a + k(\text{mod } 2^t)) &= \frac{1}{2^{2t}} \frac{\sin^2 \pi(\varphi - \frac{a+k}{2^t})2^t}{\sin^2 \pi(\varphi - \frac{a+k}{2^t})} \\
&= \frac{1}{2^{2t}} \frac{\sin^2 \pi(2^t d - k)}{\sin^2 \pi(d - \frac{k}{2^t})} \\
&\leq \frac{1}{2^{2t}} \frac{1}{\sin^2 \pi(d - \frac{k}{2^t})} \\
&\leq \frac{1}{4(k - 2^t d)^2} \text{ (follows from Exercise 4.3.1).}
\end{aligned}$$

Now we are ready to give the bound.

$$\begin{aligned}
\Pr(X - X_\delta) &= \sum_{j=-2^{t-1}+1}^{-\delta} \Pr(x = a+j(\text{mod } 2^t)) + \sum_{j=\delta+1}^{2^{t-1}} \Pr(x = a+j(\text{mod } 2^t)) \\
&\leq \frac{1}{4}\Bigg(\sum_{j=-2^{t-1}+1}^{-\delta} \frac{1}{(j - 2^t d)^2} + \sum_{j=\delta+1}^{2^{t-1}} \frac{1}{(j - 2^t d)^2} \Bigg) \\
&< \frac{1}{4}\Bigg(\sum_{j=-2^{t-1}+1}^{-\delta} \frac{1}{j^2} + \sum_{j=\delta+1}^{2^{t-1}} \frac{1}{(j-1)^2} \Bigg) \text{ (since } 0 < 2^t d < 1) \\
&= \frac{1}{2} \sum_{j=\delta}^{2^{t-1}-1} \frac{1}{j^2} \\
&< \frac{1}{2} \int_{\delta-1}^{\infty} \frac{dy}{y^2} \\
&= \frac{1}{2(\delta - 1)}.
\end{aligned}$$

To approximate φ up to the first r bits (where $r < t$) in the binary expansion, we need to choose $\delta \leq 2^{t-r} - 1$. If we use $t = r + p$ qubits in the first register of the phase estimation circuit, the probability of obtaining an estimate of the phase within the desired error margin is at least $1 - \frac{1}{2(2^p-2)}$. Let $1 - \epsilon$ be the probability of obtaining an estimate within the desired tolerance of error. Then

$$\epsilon \geq \frac{1}{2(2^p - 2)} \Rightarrow p \geq \log\left(2 + \frac{1}{2\epsilon}\right).$$

Hence, if the desired accuracy is r and the required probability of successfully getting such an estimate is $1 - \epsilon$, then we need to choose

$$t \geq r + \log\left(2 + \frac{1}{2\epsilon}\right).$$

It is easy to see that the phase-estimation circuit uses polynomially many gates.

Lecture 5

Order Finding

5.1 The Order Finding Algorithm

For any two positive integers x, y denote their greatest common divisor (GCD) by (x, y). For any positive integer N let $\mathbb{Z}_N^*$ denote the set $\{x \mid x \in \mathbb{N},\ (x, N) = 1\}$. Under multiplication modulo N, $\mathbb{Z}_N^*$ is an abelian group. Let $\varphi(N)$ be the order of this group. Then $\varphi(\cdot)$ is called the Eulers's φ function. The *order* of an element $x \in \mathbb{Z}_N^*$ is defined to be the smallest positive integer r satisfying $x^r = 1 \pmod N$. In the classical model of computation finding the order of an element in $\mathbb{Z}_N^*$ is considered to be a *hard* problem. Using the phase estimation procedure of quantum computation we shall demonstrate how one can determine the order of an element with high probability using only a polynomial number of gates.

To solve the problem of order finding using a quantum computer we first translate the problem into a problem concerning unitary operators as follows.

Let N be an L bit number so that

$$N = 2^{j_0} + 2^{j_1} + 2^{j_2} + \cdots + 2^{j_{k-1}} \text{ where } 0 \le j_0 < j_1 < j_2 < \cdots < j_{k-1} < L.$$

Let the Hilbert space generated by L qubits be denoted by $\mathcal{H} = (\mathbb{C}^2)^{\otimes L}$. We define a unitary operator U in $\mathcal{H}$ by

$$U\,|y\rangle = \begin{cases} |x\,y \pmod N\rangle & \text{if } y < N,\ \ (y = 0, 1, 2, \ldots, N-1); \\ |y\rangle & \text{if } N \le y \le 2^L - 1. \end{cases}$$

It is to be noted that if $|x\,y_1 \pmod N\rangle = |x\,y_2 \pmod N\rangle$ for $0 \le y_1 < y_2 < N$ then we have $x\,(y_2 - y_1) \equiv 0 \pmod N$. But GCD of x and N is 1. So $N|(y_2 - y_1)$ which is impossible. This means U is a permutation matrix and hence unitary.

Let

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-2\pi i \frac{sk}{r}} |x^k \pmod N\rangle.$$

We observe that

$$\begin{aligned} U|u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-2\pi i \frac{sk}{r}} |x^{k+1} \pmod N\rangle \\ &= e^{2\pi i \frac{s}{r}} \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-2\pi i \frac{sk}{r}} |x^k \pmod N\rangle. \end{aligned}$$

Thus $|u_s\rangle$ is an eigenvector of the unitary matrix U with corresponding eigenvalue $e^{2\pi i \frac{s}{r}}$, for all $s \in \{0, 1, 2, \ldots, r-1\}$.

Now if we use the phase estimation algorithm we will get enough information to obtain the order r. But in order to be able to use the phase estimation we must be able to implement the controlled U^{2^j} operation efficiently. The other requirement is that we must be able to prepare the eigenvectors accurately.

The controlled U^{2^j} operations can be implemented using $O(L^3)$ gates as outlined in Appendix 2. But the second requirement seems impossible because we need to know r in order to prepare the eigen states. This problem can be solved by observing that

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |u_s\rangle = |1\rangle.$$

Thus in the phase estimation procedure if we set the number of qubits in the first register $t = 2L + 1 + \left\lceil 2 + \frac{1}{2\epsilon} \right\rceil$ and the L qubits in the second register in the state $|1\rangle$, then for each $s \in \{0, 1, \ldots, r-1\}$ we will get an estimate of the phase $\widetilde{\varphi} \approx \frac{s}{r}$ correct up to the first $2L+1$ bits with probability at least $\frac{1-\epsilon}{r}$. The circuit is shown in Figure 5.1.

It can be checked that if in the phase estimation circuit we feed in the superposition of eigen states

$$|u\rangle = \sum_{s=0}^{r-1} c_s |u_s\rangle, \qquad \text{where } \sum_{s=0}^{r-1} |c_s|^2 = 1$$

then the output state before measurement will be

$$\frac{1}{2^t} \sum_{s,k} c_s \left\{ \frac{1 - e^{2\pi i (\varphi_s - \frac{k}{2^t}) 2^t}}{1 - e^{2\pi i (\varphi_s - \frac{k}{2^t})}} \right\} |k\rangle |u_s\rangle.$$

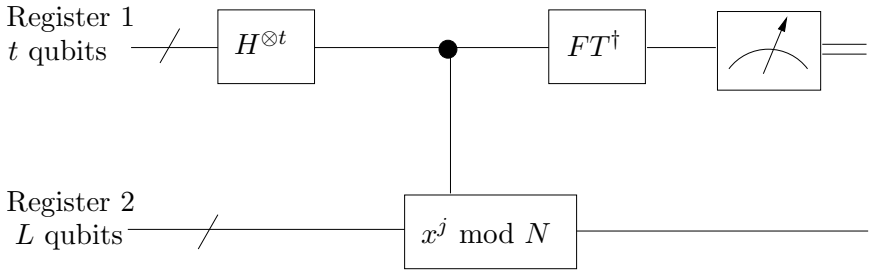

Figure 5.1: *Quantum circuit for order finding algorithm. The first register is initialized to state* $|0\rangle$ *and the second register is initialized to state* $|1\rangle$.

Hence on measuring the first t qubits we will get the value of the phase φ_s correct up to $2L+1$ bits with probability at least $|c_s|^2(1-\epsilon)$.

Now our job is to extract the value of r from the estimated phase. We know the phase $\widetilde{\varphi} \approx \frac{s}{r}$ correct up to $2L+1$ places. If this estimate is close enough, we should be able to get r because we know that $\widetilde{\varphi}$ is the ratio of two bounded integers. This task is accomplished efficiently using the following result from number theory.

Theorem 5.1.1 *If* $\frac{s}{r}$ *is a rational number such that* $\left|\frac{s}{r} - \widetilde{\varphi}\right| \leq \frac{1}{2r^2}$, *then* $\frac{s}{r}$ *is a convergent of the continued fraction for* $\widetilde{\varphi}$ *and hence can be efficiently computed using the continued fraction algorithm.*

Proof See Appendix 3.

□

We know that $|\frac{s}{r} - \tilde{\varphi}| \leq 2^{-(2L+1)} \leq \frac{1}{2r^2}$, since $r \leq N \leq 2^L$. So if we now use the continued fraction algorithm we will get the fraction $\frac{s'}{r'}$ which is equal to $\frac{s}{r}$ with $(r', s') = 1$. Thus if s and r are relatively prime then we get the order of the element x. We know that the number of positive integers relatively prime and less than r is at least $\frac{0{\cdot}1r \log\log r}{\log r}$ (see Appendix 4). The order finding algorithm fails if the phase estimation algorithm gives a bad estimate or if s divides r. The probability that the first case does not occur is at least $(1-\epsilon)$ and the second case does not occur is at least $\frac{0{\cdot}1 \log\log N}{\log N}$. Hence if we repeat the algorithm $O(L)$ times we will find the order with probability greater than $1-\delta$ for any fixed $\delta \in (0, 1]$. Note that the algorithm presented here can be implemented with $O(L^4)$ gates.

The algorithm can be summarized as follows
Inputs: Relatively prime integers N and x.
Output: Order of x.

Runtime: $O(L^4)$.

Procedure:

Initialize: Set "current smallest" equal to N.

1.	Prepare $U_{(x,N)}$	The equivalent sequence of controlled U^{2^j} operations.
2.	$\|0\rangle\|1\rangle$	Initial state.
3.	$\rightarrow \frac{1}{\sqrt{2^t}} \sum_{j=0}^{2^t-1} \|j\rangle\|1\rangle$	Create superposition.
4.	$\rightarrow \frac{1}{\sqrt{2^t}} \sum_{j=0}^{2^t-1} \|j\rangle\|x^j \pmod N\rangle$ $\approx$ $\frac{1}{\sqrt{r2^t}} \sum_{s=0}^{r-1} \sum_{j=0}^{2^t-1} e^{\frac{2\pi i s j}{r}} \|j\rangle\|u_s\rangle$	Apply $U_{(x,N)}$.
5.	$\rightarrow \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} \|\tilde{\varphi}\rangle\|u_s\rangle$	Apply inverse FT to first register.
6.	$\tilde{\varphi}$	Measure first register.
7.	Get denominator of all convergents of $\tilde{\varphi}$	Use Theorem 5.1.2 of Appendix 3.
8.	For all integers i obtained in Step 7, check if $x^i = 1$ and keep the smallest of them.	
9.	Update "current smallest"	
10.	Repeat steps 1 to 9 $O(\log N)$ times	
11.	Return "current smallest"	With a high probability. This is the order.

Appendix 1: Classical reversible computation

All quantum gates are reversible (i.e. from the output we can uniquely recover the input). But the classical gates like 'AND' and 'OR' are not reversible. So a quantum circuit cannot exist for any such gate. However, by adding a few extra wires we can obtain a gate which is reversible and the required function appears on specified wires. This is called a *reversible classical gate*. If the 'size' of the circuit is measured by

the number of 'wires' then this procedure uses only a constant multiple of the number of wires used in the earlier classical circuit. The latter gate can be implemented using a quantum gate. Reversible classical gates can be built using the Fredkin gate (see Figure 5.2).

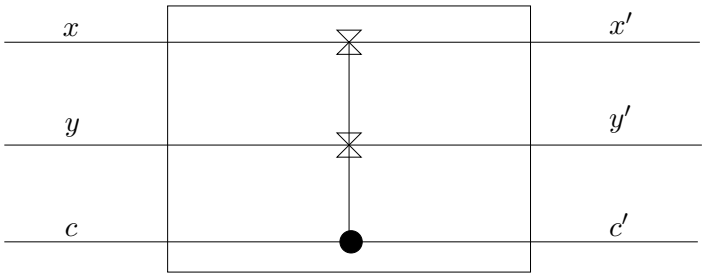

Figure 5.2: *Fredkin gate (controlled swap).*

If we set x to 0 then x' will be $y \wedge c$ which is the AND gate. If we set $x = 0$ and $y = 1$ then we get c on x' and $\neg c$ on y'. Thus we get both NOT and FANOUT gates. CNOT can also be used to copy classical bits. In the process of constructing functional equivalents of the classical gates using quantum gates some extra wires have been introduced. The outputs of these wires are called *junk*. But if the 'junk' is some arbitrary function of the input then the circuit may not behave as a quantum gate for the function $f(x)$. So instead of some junk output we would like to have some fixed output on the extra wires. This model is called *clean* computation. This can be done as shown in the Figures 5.3, 5.4 and 5.5.

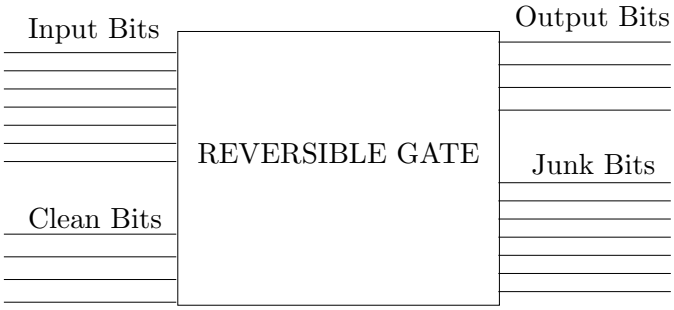

Figure 5.3: *Reversible gate*

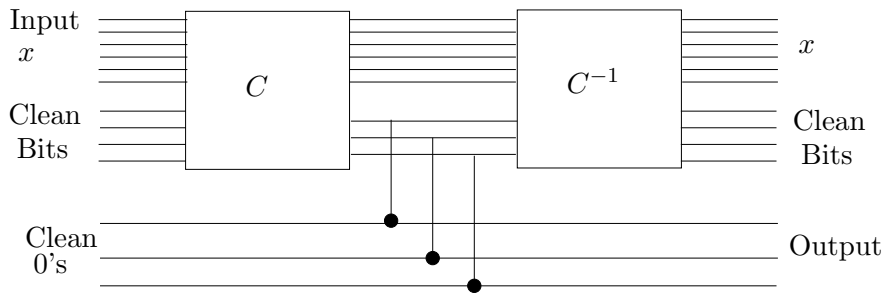

Figure 5.4: *Clean computation. Computing* $x \mapsto \langle x, f(x)\rangle$

Appendix 2: Efficient implementation of controlled U^{2^j} operation

To compute the sequence of controlled U^{2^j} operations we have to compute the transformation

$$\begin{aligned}|z\rangle|y\rangle &\to |z\rangle U^{z_t 2^{t-1}} \dots U^{z_1 2^0}|y\rangle \\ &= |z\rangle|x^{z_t 2^{t-1}} \times \dots \times x^{z_1 2^0} y \pmod N\rangle \\ &= |z\rangle|x^z y \pmod N\rangle.\end{aligned}$$

Thus the sequence of controlled U^{2^j} operations is equivalent to multiplying the content of the second register by the modular exponential $x^z \pmod N$, where z is the content of the first register. This can be computed using clean reversible computation (see Appendix 1).

This is achieved by first reversibly computing the function $x^z (\mathrm{mod}\,) N$ in a third register and then multiplying the contents of the third and the second register such that each qubit in the third register is in the state $|0\rangle$. The task is accomplished in two stages. In the first stage we compute x^{2^j} for all $j \in \{1, 2, \dots, t-1\}$ by successively squaring

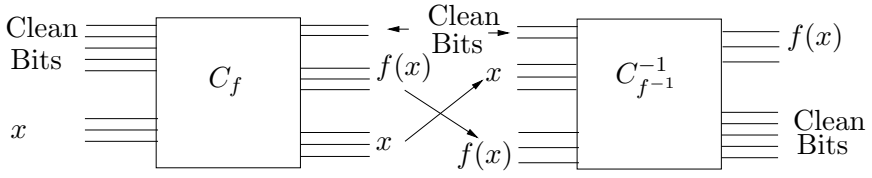

Figure 5.5: *Computing a bijective function* f

$x \pmod N$, where $t = 2L + 1 + \lceil \log 2 + \frac{1}{2\epsilon} \rceil = O(L)$. Each multiplication uses at most $O(L^2)$ gates (Indeed an $O(L \log L \log \log L)$ algorithm using FFT is known. See [2, 10].) and there are $t - 1$ such multiplications. Hence in this step at most $O(L^3)$ gates are used. In the second stage we compute $x^z \pmod N$ using the identity

$$x^z \pmod N = (x^{z_t 2^{t-1}} \pmod N)(x^{z_t 2^{t-2}} \pmod N) \cdots (x^{z_t 2^0} \pmod N).$$

Clearly this operation also uses at most $O(L^3)$ gates. Hence using $O(L^3)$ gates we compute the transformation $|z\rangle|y\rangle \rightarrow |z\rangle|x^z y \pmod N\rangle$.

Appendix 3: Continued fraction algorithm

A finite continued fraction of $n + 1$ variables is defined as

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots \cfrac{1}{a_n}}}}.$$

For convenience it is also written as $[a_0, a_1, \ldots, a_n]$. The nth convergent of a continued fraction $[a_0, a_1, \ldots a_N]$ is defined as $[a_0, a_1, \ldots, a_n]$ for $n \leq N$.

The nth convergent is easily computed by the following theorem.

Theorem 5.1.2 *If p_n and q_n are defined by*

$$p_0 = a_0, \;\; p_1 = a_1 a_0 + 1, \;\; p_n = a_n p_{n-1} + p_{n-2} \;\; \textit{for } 2 \leq n \leq N,$$
$$q_0 = 1, \;\; q_1 = a_1, \;\; q_n = a_n q_{n-1} + q_{n-2} \;\; \textit{for } 2 \leq n \leq N$$

then $[a_0, a_1, \ldots, a_n] = \frac{p_n}{q_n}$.

Proof We prove by induction. It is easy to check for the base cases $n = 1, 2$.
Induction Hypothesis: The conclusion holds for $1 \leq n \leq m$.

Induction step:

$$\begin{aligned}
[a_0, a_1, \ldots, a_m, a_{m+1}] &= \left[a_0, a_1, \ldots, a_{m-1}, a_m + \frac{1}{a_{m+1}}\right] \\
&= \frac{\left(a_m + \frac{1}{a_{m+1}}\right) p_{m-1} + p_{m-2}}{\left(a_m + \frac{1}{a_{m+1}}\right) q_{m-1} + q_{m-2}} \\
&= \frac{a_{m+1}(a_m p_{m-1} + p_{m-2}) + p_{m-1}}{a_{m+1}(a_m q_{m-1} + q_{m-2}) + q_{m-1}} \\
&= \frac{a_{m+1} p_m + p_{m-1}}{a_{m+1} q_m + q_{m-1}} \\
&= \frac{p_{m+1}}{q_{m+1}}.
\end{aligned}$$

□

Theorem 5.1.3 *The functions p_n and q_n satisfy the following relation*

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^n.$$

Proof We use induction. The result is true for the base cases $n = 1, 2$. Assume the result is true for any integer less than n.

$$\begin{aligned}
p_n q_{n-1} - p_{n-1} q_n &= (a_n p_{n-1} + p_{n-2}) q_{n-1} - p_{n-1}(a_n q_{n-1} + q_{n-2}) \\
&= -1(p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) \\
&= (-1)^n.
\end{aligned}$$

This completes the proof.

□

Let x be a real number. Then the system of equations

$$\begin{aligned}
x &= a_0 + \alpha_0 \quad && \text{with } a_0 \in \mathbb{Z} \text{ and } \alpha_0 \in [0, 1) \\
\frac{1}{\alpha_0} &= a_1 + \alpha_1 \quad && \text{with } a_1 \in \mathbb{Z} \text{ and } \alpha_1 \in [0, 1) \\
\frac{1}{\alpha_1} &= a_2 + \alpha_2 \quad && \text{with } a_2 \in \mathbb{Z} \text{ and } \alpha_2 \in [0, 1) \\
&\vdots
\end{aligned}$$

is called the *continued fraction algorithm*. The algorithm continues till $\alpha_n \neq 0$.

It is easy to see that if the algorithm terminates in $N+1$ steps then $x = [a_0, a_1, \ldots, a_N]$ and hence rational. But the converse of this is also true.

Theorem 5.1.4 *Any rational number can be represented by a finite continued fraction.*

Proof Let $x = \frac{h}{k}$. Then from the continued fraction algorithm we get the following set of equations.

$$\begin{aligned} h &= a_0 k + k_1 \quad (0 < k_1 < k) \\ k &= a_1 k_1 + k_2 \quad (0 < k_2 < k_1) \\ &\vdots \end{aligned}$$

We observe that $k > k_1 > k_2 \cdots$. Hence the algorithm must terminate. Also, this is exactly the Euclid's GCD algorithm. Hence its complexity is $O((\log(h+k))^3)$ (see [5, 10]).

□

Theorem 5.1.5 *If x is representable by a simple continued fraction with an odd (even) number of convergents, it is also representable by one with an even (odd) number of convergents.*

Proof Let $x = [a_0, \ldots, a_n]$. If $a_n \geq 2$, then $[a_0, \ldots, a_n] = [a_0, \ldots, a_n - 1, 1]$. If $a_n = 1$, then $[a_0, \ldots, a_{n-1}, 1] = [a_0, \ldots, a_{n-1} + 1]$.

□

Theorem 5.1.6 *Let x be a rational number and p and q two integers such that*

$$\left| \frac{p}{q} - x \right| \leq \frac{1}{2q^2}.$$

Then $\frac{p}{q}$ is a convergent of the continued fraction for x.

Proof Let $[a_0, \ldots, a_n]$ be the continued fraction expansion of $\frac{p}{q}$. From Theorem 5.1.5 it follows that without loss of generality we may assume n to be even. Let p_i and q_i be defined as in Theorem 5.1.2.

Let δ be defined by the equation

$$x = \frac{p_n}{q_n} + \frac{\delta}{2q_n^2}.$$

Then $|\delta| \leq 1$ and $\frac{p_n}{q_n} = \frac{p}{q}$ is the nth convergent. Let

$$\lambda = 2\left(\frac{q_n p_{n-1} - p_n q_{n-1}}{\delta}\right) - \frac{q_{n-1}}{q_n}.$$

The definition of λ ensures that the equation

$$x = \frac{\lambda p_n + p_{n-1}}{\lambda q_n + q_{n-1}}$$

is satisfied. Hence $x = [a_0, \ldots a_n, \lambda]$. By Theorem 5.1.3 we get

$$\begin{aligned}\lambda &= \frac{2}{\delta} - \frac{q_{n-1}}{q_n} \\ &> 2 - 1 \text{ since } q_i > q_{i-1} \\ &= 1.\end{aligned}$$

This implies that λ is a rational number greater than 1 and it has a finite continued fraction, say $[b_0, \ldots, b_m]$. Hence $x = [a_0, \ldots, a_n, b_0, \ldots, b_m]$. Thus $\frac{p}{q}$ is a convergent of x.

□

Appendix 4: Estimating $\frac{\varphi(r)}{r}$

Lemma 5.1.7 *The ratio $\frac{\varphi(r)}{r}$ is at least $\frac{\log\log r}{10\log r}$ for $r \geq 16$.*

Proof Let $r = \prod_{i=1}^{a} p_i^{\alpha_i} \prod_{j=1}^{b} q_j^{\beta_j}$, where

$$p_1 < p_2 < \cdots < p_a \leq \frac{2\log r}{\log\log r} < q_1 < q_2 < \cdots < q_b.$$

Then

$$\varphi(r) = \prod_{i=1}^{a}(p_i - 1)p_i^{\alpha_i - 1} \prod_{j=1}^{b}(q_i - 1)q_i^{\beta_j - 1}.$$

Note that $q_1^b \leq r$. This implies $b \leq \log_q r \leq \log r$. Since $q_1 > \frac{2\log r}{\log\log r}$, we have $b \leq \frac{\log r}{\log\log r - \log\log\log r + \log 2}$.

Hence,

$$\begin{aligned}
\frac{\varphi(r)}{r} &= \frac{\prod_{i=1}^{a}(p_i-1)p_i^{\alpha_i-1}\prod_{j=1}^{b}(q_j-1)q_j^{\beta_j-1}}{\prod_{i=1}^{a}p_i^{\alpha_i}\prod_{j=1}^{b}q_i^{\beta_j}} \\
&= \prod_{i=1}^{a}\left(\frac{p_i-1}{p_i}\right)\prod_{j=1}^{b}\left(1-\frac{1}{q_j}\right) \\
&> \prod_{i=2}^{\frac{2\log r}{\log\log r}}\left(\frac{i-1}{i}\right)\prod_{j=1}^{b}\left(1-\frac{1}{q_j}\right) \\
&= \frac{\log\log r}{2\log r}\prod_{j=1}^{b}\left(1-\frac{1}{q_j}\right) \\
&> \frac{\log\log r}{2\log r}\left(1-\frac{\log\log r}{2\log r}\right)^{b} \\
&> \frac{\log\log r}{2\log r}\left(1-\frac{\log\log r}{2\log r}b\right) \\
&= \frac{\log\log r}{2\log r}\left[1-\frac{\log\log r}{2\log r}\left(\frac{\log r}{\log\log r-\log\log\log r+\log 2}\right)\right] \\
&> \frac{\log\log r}{2\log r}\left[\frac{1-2E}{2(1-E)}\right] \text{ where } E=\frac{\log\log\log r-\log 2}{\log\log r} \\
&> \frac{\log\log r}{2\log r}\left(\frac{1-2E}{2}\right) \\
&> \frac{\log\log r}{10\log r} \text{ for } r\geq 16.
\end{aligned}$$

□

In fact the following theorem is true.

Theorem 5.1.8 $\underline{\lim}_{n\to\infty}\frac{\varphi(n)\log\log n}{n}=e^{-\gamma}$ *where* γ *is the Euler's constant.*

□

The interested reader may look up Hardy and Wright [6] for the proof.

Lecture 6

Shor's Algorithm

6.1 Factoring to Order Finding

Lemma 6.1.1 *Let N be an odd number with prime factorization $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}$, $m \geq 2$. Let*

$$A \stackrel{\Delta}{=} \{x \in \mathbb{Z}_N^* \mid (\mathrm{ord}(x) \text{ is odd}) \text{ or } (\mathrm{ord}(x) \text{ is even and } x^{\mathrm{ord}(x)/2} = -\mathbf{1})\},$$

where $\mathrm{ord}(x) = \min\{i \geq 1 \mid x^i = \mathbf{1}\}$. *If x is chosen at ramdom from $\mathbb{Z}_N^*$, then*

$$\Pr_{x \in \mathbb{Z}_N^*}(x \in A) \leq \frac{1}{2^{m-1}}.$$

Proof [1] Let $|\mathbb{Z}_N^*| = \varphi(N) = 2^\ell s$, where s is odd (note $\ell \geq 2$). Let V be the set of square-roots of $\mathbf{1}$ in $\mathbb{Z}_N^*$.

Lemma 6.1.2 (a) *If* $\mathrm{ord}(x)$ *is odd, then* $x^s = \mathbf{1}$.

(b) *If* $\mathrm{ord}(x)$ *is even, then* $x^{2^i s} \in V - \{\mathbf{1}\}$, *for some* $i \in \{0, 1, \dots, \ell - 1\}$.

(c) *If* $\mathrm{ord}(x)$ *is even and* $x^{\mathrm{ord}(x)/2} = -\mathbf{1}$, *then* $x^{2^i s} = -\mathbf{1}$ *for some* $i \in \{0, 1, \dots, \ell - 1\}$.

Proof (a) Since $x \in \mathbb{Z}_N^*$, we have $\mathrm{ord}(x) | \phi(N)$. Since $\mathrm{ord}(x)$ is odd, $\mathrm{ord}(x) | s$.

[1] Our proof is based on the proof of correctness of *Miller's primality test* in Kozen's book [12, page 206]. Nielsen and Chuang [4, Theorem A4.13, page 634] give a bound of 2^{-m}. Their bound is *not correct*: for $N = 21 = 3 \times 7$, we have $|\mathbb{Z}_N^*| = 12$ and $|A| = 6$. Then, $\frac{|A|}{|\mathbb{Z}_N^*|} \not\leq 2^{-2}$.

(b) and (c) Let $\mathrm{ord}(x) = 2^{\ell'}s'$ (where $\ell' \geq 1$ and s' is odd). Then, $\mathrm{ord}(x) | 2^{\ell'}s$, but $\mathrm{ord}(x) \nmid 2^{\ell'-1}s$. Hence, $x^{2^{\ell'-1}s} \in V - \{\mathbf{1}\}$. Now, if $x^{\mathrm{ord}(x)/2} = -\mathbf{1}$, then $x^{2^{\ell'-1}s'} = -1$. Hence, $x^{2^{\ell'-1}s} = -\mathbf{1}$.

□

For $i = 0, 1, \ldots, \ell - 1$, and $v \in V$, let $S_{i,v} \stackrel{\Delta}{=} \{x \in \mathbb{Z}_N^* : x^{2^i s} = v\}$. By Lemma 6.1.2, we have

$$A \subseteq S_{0,\mathbf{1}} \cup \bigcup_{i=0}^{\ell-1} S_{i,-\mathbf{1}}; \tag{6.1.3}$$

$$\text{and} \quad \mathbb{Z}_N^* = S_{0,\mathbf{1}} \cup \bigcup_{i=0}^{\ell-1} \bigcup_{v \in V - \{\mathbf{1}\}} S_{i,v}. \tag{6.1.4}$$

Lemma 6.1.5 *All the sets appearing on the right hand side of (6.1.4) are disjoint.*

Proof Consider two such sets $S_{i,v}$ and $S_{j,w}$ appearing above. If $i = j$ then $v \neq w$ and these sets are disjoint by defnition. Hence, suppose $i < j$; this implies that $w \neq \mathbf{1}$. But for each $x \in S_{i,v}$, we have $x^{2^{i+1}s} = v^2 = \mathbf{1}$. This implies that $x^{2^j s} = \mathbf{1} \neq w$, and therefore $x \notin S_{j,w}$.

□

To prove that $|A| \leq 2^{-m+1}|\mathbb{Z}_N^*|$, we will use the isomorphism

$$\begin{aligned} &\mathbb{Z}_N^* \to \mathbb{Z}_{p_1^{\alpha_1}}^* \times \mathbb{Z}_{p_2^{\alpha_2}}^* \times \cdots \times \mathbb{Z}_{p_m^{\alpha_m}}^*; \\ &\quad j \mapsto (j \pmod{p_1^{\alpha_1}}, j \pmod{p_2^{\alpha_2}}, \ldots, j \pmod{p_m^{\alpha_m}}), \end{aligned}$$

which follows from the *Chinese remainder theorem.*

Since p_i is odd, $\mathbf{1} \neq -\mathbf{1} \pmod{p_i^{\alpha_i}}$, for $i = 1, 2, \ldots, m$, and the 2^m elements in $W = \{+1, -1\}^m$ correspond to square roots of $\mathbf{1}$ in $\mathbb{Z}_N^*$; of these, the only trivial square roots are $\mathbf{1} = (1, 1, \ldots, 1)$ and $-\mathbf{1} = (-1, -1, \ldots, -1)$.

□

Lemma 6.1.6

$$|S_{0,\mathbf{1}}| = |S_{0,-\mathbf{1}}|; \tag{6.1.7}$$

$$|S_{j,-\mathbf{1}}| = |S_{j,w}|, \quad \textit{for } w \in W \textit{ and } j = 0, 1, \ldots, \ell - 1. \tag{6.1.8}$$

Proof To see (6.1.7), observe that $x \in S_{0,\mathbf{1}}$ if and only if $x^s = \mathbf{1}$, if and only if $(-x)^s = -\mathbf{1}$, if and only if $-x \in S_{0,-\mathbf{1}}$.

To prove (6.1.8), fix j and w. We first show that if $S_{j,-\mathbf{1}} \neq \emptyset$, then $S_{j,w} \neq \emptyset$. For, suppose $b = (b_1, b_2, \ldots, b_m) \in S_{j,-\mathbf{1}}$. Then, consider $c \in \mathbb{Z}^*_{p^{\alpha_1}} \times \mathbb{Z}^*_{p_2^{\alpha_2}} \times \cdots \times \mathbb{Z}^*_{p_m^{\alpha_m}}$, defined by

$$c_i = \begin{cases} 1 & \text{if } w_i = 1; \\ b_i & \text{if } w_i = -1. \end{cases}$$

Clearly, $c^{2^j s} = w$, so $S_{j,w} \neq \emptyset$. Furthermore, the map $x \mapsto cb^{-1}x$ is a bijection between $S_{j,-\mathbf{1}}$ and $S_{j,w}$. Hence, $|S_{j,-\mathbf{1}}| = |S_{j,w}|$.

□

Since $|W| = 2^m$, from (6.1.3), (6.1.4) and Lemma 6.1.6 we obtain

$$2^{m-1}|S_{0,\mathbf{1}} \cup S_{0,-\mathbf{1}}| = |\bigcup_{w \in W} S_{0,w}|,$$

$$\text{and for } i = 0, 1, 2, \ldots, \ell - 1, \quad (2^m - 1)|S_{i,-\mathbf{1}}| = |\bigcup_{w \in \{W - \{\mathbf{1}\}\}} S_{i,w}|,$$

which implies

$$\begin{aligned} 2^{m-1}|A| &\leq 2^{m-1}|S_{0,\mathbf{1}} \cup \bigcup_{i=0}^{\ell-1} S_{i,-\mathbf{1}}| \\ &\leq |S_{0,\mathbf{1}} \cup \bigcup_{i=0}^{\ell-1} \bigcup_{w \in \{W-\{\mathbf{1}\}\}} S_{i,w}| \\ &\leq |S_{0,\mathbf{1}} \cup \bigcup_{i=0}^{\ell-1} \bigcup_{v \in \{V-\{\mathbf{1}\}\}} S_{i,v}| \\ &= |\mathbb{Z}_N^*|\,. \end{aligned}$$

□

Lemma 6.1.1 is the main tool for analyzing the Shor's factoring algorithm. The crucial observation is that, if we can get a nontrivial square root of unity, then we can find a nontrivial factor of N using Euclid's G.C.D. algorithm. Lemma 6.1.1 tells us that if we randomly pick a number x, less than N and look at its order, with probability greater than $1 - \frac{1}{2^{m-1}}$ it is even and we can get a nontrivial square root of unity

by raising x to the power $\text{ord}(x)/2$. The lemma holds if N is odd and has at least two distinct prime factors. But a classical polynomial time algorithm exists for finding the prime number which divides N, if N is a prime power. So this gives us a polynomial time factoring algorithm. So far it is not known whether classical computers can factorize a number N in polynomial time, even if randomness is allowed. Below is the Shor's factoring algorithm.

Shor's factoring algorithm.

Input N

1) If N is even, return 2.

2) Use quantum order finding algorithm to find the order of 2. If $\text{ord}(2) = N - 1$, conclude N is prime and stop.

3) Check if N is of the form $p^\alpha,\ \alpha > 1$ by the subroutine Prime-power.

4) Pick an element $x \in N$.

5) If $x \mid N$, return x.

6) Use quantum order finding algorithm to find the order of x.

7) If $\text{ord}(x)$ is odd then abort.

8) If $x^{\frac{\text{ord}(x)}{2}} = -\mathbf{1} \pmod N$ then abort.

9) Get a nontrivial square root of $\mathbf{1} \pmod N$, by setting $y \leftarrow x^{\frac{\text{ord}(x)}{2}}$.

10) Use Euclid's G.C.D. algorithm to find the greatest common divisor of $(y-1, N)$ and $(y+1, N)$. Return the nontrivial numbers.

Output: With high probability it gives a divisor of N or tells if N is prime.

Subroutine: Prime-power

Input: Integer N.

1) Compute $y = \log_2 N$.

2) For all $i \in \{2, 3, \ldots, \log_2 N\}$ compute $x_i = \frac{y}{i}$.

3) Find $u_i < 2^{x_i} < u_i + 1$ for all $i \in \{2, 3, \ldots, \log_2 N\}$.

4) Check if $u_i \mid N$ or $u_i + 1 \mid N$ for all $i \in \{2, 3, \ldots, \log_2 N\}$. If any one of the numbers divide N, say u, then return u. Else fail.

Output: If N is a prime power of p, the subroutine "prime-power" returns p. If it is not a prime power it fails to produce any output. In $O((\log N)^3)$ steps it terminates. The most costly operation in the algorithm is the order finding algorithm. Since the order finding takes $O((\log N)^4)$ time, the time taken by this factoring algorithm is also $O((\log N)^4)$.

Remark 6.1.9 Step 1) just checks if the number N is divisible by 2. Step 2) checks if the number N is prime and Step 3) if N is a prime power. So after Step 3) Lemma 6.1.1 is applicable.

Probability of success in Shor's algorithm is greater than probability of success in order finding multiplied by the probability that the chosen element x is not in the set A, of Lemma 6.1.1. Running time of the algorithm is $O((\log N)^4)$. Thus, by running the algorithm only a constant number of times we can get probability of success greater than $1 - \epsilon$ for any fixed $\epsilon > 0$.

Exercise 6.1.10 Find a randomized polynomial time algorithm for factoring an integer N, if $\varphi(N)$ is known.

Lecture 7

Quantum Error Correcting Codes

7.1 Knill Laflamme Theorem

The mathematical theory of communication of messages through a quantum information channel is based on the following three basic principles.

1) Messages can be encoded as states and transmitted through quantum channels.
2) The output state may not be the same as the input state due to presence of noise in the channel.
3) There is a collection of "good" states which when transmitted through the noisy channel leads to output states from which the input state can be recovered with no error or with a small margin of error.

The aim is to identify the set of good states for a given model of the noisy channel and to give the decoding procedure.

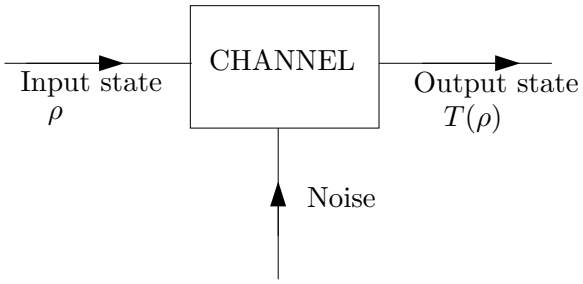

Figure 7.1: *A model of noisy quantum channel*

Let $\mathcal{H}$ be a finite dimentional complex Hilbert space. We assume that there is a linear space $\mathcal{E} \subset \mathcal{B}(\mathcal{H})$, called the error space such that

for any input state ρ on $\mathcal{H}$ the output state $T(\rho)$ has always the form

$$T(\rho) = \sum_j L_j \rho L_j^\dagger \tag{7.1.1}$$

where L_j belongs to $\mathcal{E}$ for every j (See Figure 7.1). If the same input state is transmitted again the operators L_j may be completely different. But they always come from the error space $\mathcal{E}$ and satisfy the equation

$$\mathrm{Tr}\left(\sum_{j=1}^{k} L_j^\dagger L_j\right)\rho = 1. \tag{7.1.2}$$

The L_j's may or may not depend on the density matrix ρ which is transmitted through the noisy channel.

Definition 7.1.3 A state ρ is said to have its support in a subspace $S \subset \mathcal{H}$ if $\mathrm{Tr}\,\rho E^S = 1$ where E^S is the orthogonal projection on S.

This means if we choose an orthonormal basis $e_1, \ldots e_k, e_{k+1}, \ldots, e_N$ for $\mathcal{H}$ such that $e_1, e_2, \ldots, e_k$ is an orthonormal basis for S then the matrix of ρ in this basis has the form $\left[\begin{array}{c|c} \tilde{\rho} & 0 \\ \hline 0 & 0 \end{array}\right]$ where $\tilde{\rho}$ is a $k \times k$ matrix. To recover the input state at the output of the channel we apply a recovery operator R of the form

$$R(T(\rho)) = \sum_j M_j T(\rho) M_j^\dagger,$$

$$\sum_j M_j^\dagger M_j = I.$$

It would be desirable to have $R(T(\rho)) = \rho$ for all ρ, whenever the L's are from $\mathcal{E}$ and they act on ρ as in (7.1.1). Of course this is too ambitious. We would like to achieve this pleasant situation at least for all ρ with support in some 'large' subspace $\mathcal{C} \subset \mathcal{H}$. Then we can encode messages in terms of states from $\mathcal{C}$ and recover them with the help of a decoding operation R. The idea is formalized in the following definition.

Definition 7.1.4 A subspace $\mathcal{C} \subset \mathcal{H}$ is called a $\mathcal{E}$-correcting quantum code, if there exist operators $M_1, M_2, \ldots, M_k$, such that for every ρ with support in $\mathcal{C}$ and any $L_1, L_2, \ldots, L_l \in \mathcal{E}$, with $\mathrm{Tr}(\sum_j L_j^\dagger L_j)\rho = 1$, one has $\sum_{i,j} M_i L_j \rho L_j^\dagger M_i^\dagger = \rho$.

Remark 7.1.5 Now consider $|u\rangle \in \mathcal{C}$. Then $|u\rangle\langle u|$ has support in $\mathcal{C}$. Consider the equations

$$\sum M_i L_j |u\rangle\langle u| L_j^\dagger M_i^\dagger = |u\rangle\langle u|$$

and

$$\langle u|(\sum_j L_j^\dagger L_j)|u\rangle = 1.$$

Choose any $|v\rangle \in \mathcal{H}$ such that $\langle u \mid v\rangle = 0$. Then we have

$$\sum_{i,j} |\langle v|M_i L_j|u\rangle|^2 = 0. \tag{7.1.6}$$

This is true if and only if $\langle v|M_i L_j|u\rangle = 0$ for all $|v\rangle \in \{|u\rangle\}^\perp$ and every i, j. Thus,

$$M_i L_j |u\rangle = c(u)|u\rangle \quad \text{for all } |u\rangle \in \mathcal{C}.$$

$M_i L_j$ is an operator and $\mathcal{C}$ is a subspace. Hence this can happen if and only if

$$M_i L \,\big|_{\mathcal{C}} = \lambda_i(L) I \,\big|_{\mathcal{C}} \text{ for all } L \in \mathcal{E}.$$

We state this as a proposition.

Proposition 7.1.7 *A subspace $\mathcal{C} \subset \mathcal{H}$ is an $\mathcal{E}$-correcting quantum code if and only if there exist operators $M_1, M_2, \ldots M_k$ in $\mathcal{H}$, such that, $\sum_i M_i^\dagger M_i = I$ and $M_i L \,\big|_{\mathcal{C}} = \lambda_i(L) I \,\big|_{\mathcal{C}}$ for all $L \in \mathcal{E}$.*

We would like to have a characterization of the quantum code $\mathcal{C}$ without involving the M_i's. That is, a condition entirely in terms of $\mathcal{C}$ and $\mathcal{E}$. This is achieved by the following remarkable criterion due to Knill and Laflamme.

Theorem 7.1.8 (Knill and Laflamme) *A subspace $\mathcal{C}$ with an orthonormal basis ψ_0, ψ_1, …, ψ_{k-1} is an $\mathcal{E}$-correcting quantum code if and only if*

1. $\langle \psi_i|L_1^\dagger L_2|\psi_j\rangle = 0$ *for all $i \neq j$, and all $L_1, L_2 \in \mathcal{E}$;*
2. $\langle \psi_i|L_1^\dagger L_2|\psi_i\rangle$ *is independent of $i = 0, 1, \ldots, k-1$.*

Proof *Necessity:* By the Proposition 7.1.7 we know that there must exist recovery operators $R_1, \ldots, R_l$ satisfying the equations $\sum_i R_i^\dagger R_i = I$ and $R_i L\psi = \lambda_i(L)\psi, \quad \psi \in \mathcal{C},\ L \in \mathcal{E}$. Let $L_1, L_2 \in \mathcal{E}$, then

$$\begin{aligned}\langle\psi_i|L_1^\dagger L_2|\psi_j\rangle &= \langle\psi_i|L_1^\dagger\Big(\sum_r R_r^\dagger R_r\Big)L_2|\psi_j\rangle \\ &= \sum_r \overline{\lambda_r(L_1)}\lambda_r(L_2)\langle\psi_i \mid \psi_j\rangle \\ &= \sum_r \overline{\lambda_r(L_1)}\lambda_r(L_2)\delta_{ij}.\end{aligned}$$

Sufficiency: Let the conditions (1) and (2) hold. Consider the subspaces $\mathcal{E}\psi_0, \mathcal{E}\psi_1, \ldots, \mathcal{E}\psi_{k-1}$. It can be verified that the correspondence $L\psi_i \to L\psi_j$, for all $L \in \mathcal{E}$ is a scalar product preserving map. So we can write the following table.

ψ_0	ψ_1	$\cdots$	ψ_j	$\cdots$	ψ_{k-1}
$\mathcal{E}\psi_0$	$\mathcal{E}\psi_1$	$\cdots$	$\mathcal{E}\psi_j$	$\cdots$	$\mathcal{E}\psi_{k-1}$

φ_0^0	φ_1^0	$\cdots$	φ_j^0	$\cdots$	φ_{k-1}^0
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
φ_0^{l-1}	φ_1^{l-1}	$\cdots$	φ_j^{l-1}	$\cdots$	φ_{k-1}^{l-1}

Here $\varphi_0^0, \varphi_0^1, \ldots, \varphi_0^{l-1}$ is an orthonormal basis for the subspace $\mathcal{E}\psi_0$. The map $L\psi_0 \to L\psi_j$, for any $L \in \mathcal{E}$, is a unitary isomorphism between the subspaces $\mathcal{E}\psi_0$ and $\mathcal{E}\psi_j$. So $\dim \mathcal{E}\psi_j = l$ for all $j \in \{0, 1, \ldots k-1\}$ and there exists a global unitary operator U_j, satisfying $U_j\varphi_0^i = \varphi_j^i$, $i = 0, 1, \ldots, l-1$. Since by the first condition $\langle L_1\psi_i \mid L_2\psi_j\rangle = 0$ for $L_1, L_2 \in \mathcal{E}$ and $i \neq j$, the subspaces $\mathcal{E}\psi_j \ \ j = 0, 1, \ldots k-1$ are mutually orthogonal. Let E_i be the projection on the span of the i^{th} row in the array $\{\varphi_j^i\}$. Now we define a unitary operator $V^{(i)}$ satisfying $V^{(i)}\varphi_j^i = \psi_j$ for $i = 0, 1, \ldots, l-1$.

Let $R_i = V^{(i)}E_i$ for $i = 0, 1, \ldots, l-1$ and $R_l = E_l$, the projection on $\{\varphi_j^i, 0 \le i \le l-1, 0 \le j \le k-1\}^\perp$. It can be verified that $\sum_{i=0}^{l} R_i^\dagger R_i = I$.

Now consider any $\psi = c_0\psi_0 + c_1\psi_1 + \cdots + c_{k-1}\psi_{k-1}$ in $\mathcal{C}$. Then

$$\begin{aligned}L\psi &= c_0 L\psi_0 + c_1 L\psi_1 + \cdots + c_{k-1}L\psi_{k-1}, \\ &= c_0 L\psi_0 + c_1 U_1 L\psi_0 + \cdots + c_{k-1}U_{k-1}L\psi_0.\end{aligned}$$

Let

$$L\psi_0 = \alpha_0(L)\varphi_0^0 + \alpha_1(L)\varphi_0^1 + \cdots + \alpha_{l-1}\varphi_0^{l-1}.$$

Then we have

$$\begin{aligned} U_j L\psi_0 &= \alpha_0(L)\varphi_j^0 + \alpha_1(L)\varphi_j^1 + \cdots + \alpha_{l-1}\varphi_j^{l-1} \\ \Rightarrow E_i U_j L\psi_0 &= \alpha_i(L)\varphi_j^i \\ \Rightarrow V^{(i)} E_i U_j L\psi_0 &= \alpha_i(L)\psi_j. \end{aligned}$$

That is,

$$\begin{aligned} R_i U_j L\psi_0 &= \alpha_i(L)\psi_j \text{ for } i = 0, 1, \ldots, l-1, \\ E_l U_j L\psi_0 &= 0 = R_l U_j L\psi_0. \end{aligned}$$

Thus we have,

$$\begin{aligned} R_i L\psi &= c_0\alpha_i(L)\psi_0 + c_1\alpha_i(L)\psi_1 + \cdots + c_{k-1}\alpha_i(L)\psi_{k-1} \\ &= \alpha_i(L)\psi \text{ for } i \in \{0, 1, \ldots, l-1\}, \text{ and} \\ R_l L\psi &= 0. \end{aligned}$$

That is, $R_i L \big|_{\mathcal{C}} = \alpha_i(L) I \big|_{\mathcal{C}}$, where $\alpha_l(L) = 0$.

□

Example 7.1.9 Let G be a finite group with identity element e and $\mathcal{H} = L^2(G)$, the Hilbert space of functions on G with

$$\langle f_1, f_2 \rangle = \sum_{x \in G} \overline{f_1(x)} f_2(x).$$

Let $E \subset G$ be called the error set and $C \subset G$ the code set. Let $\mathcal{E} = \text{lin}\{L_x \mid x \in E\}$, where $(L_a f)(x) = f(a^{-1}x)$, lin denotes linear span and $\mathcal{C} = \text{lin}\{1_{\{c\}} \mid c \in C\}$. It can be verified that $L_a 1_{\{b\}} = 1_{\{ab\}}$.

If $c_1 \neq c_2$, then

$$\begin{aligned} \left\langle 1_{\{c_1\}}, L_x^\dagger L_y 1_{\{c_2\}} \right\rangle &= \langle 1_{\{c_1\}}, 1_{\{x^{-1}yc_2\}} \rangle \\ &= 0 \text{ if } x^{-1}yc_2 \neq c_1 \\ &\qquad \text{or } x^{-1}y \neq c_1 c_2^{-1} \text{ or } E^{-1}E \cap CC^{-1} = \{e\}. \end{aligned}$$

Also,

$$\left\langle 1_{\{c\}}, L_x^\dagger L_y 1_{\{c\}} \right\rangle = \begin{cases} 1 & \text{if } x = y; \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\langle 1_{\{c\}}, L_x^\dagger L_y 1_{\{c\}} \rangle$ is independent of c. Hence by Knill–Laflamme theorem we see that $\mathcal{C}$ is an $\mathcal{E}$-correcting quantum code if $E^{-1}E \cap CC^{-1} = \{e\}$.

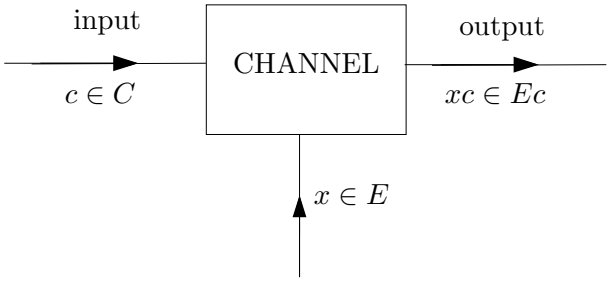

Figure 7.2: *A model of noisy classical channel.*

Consider the model of a noisy classical channel shown in Figure 7.2. If $E^{-1}E \cap CC^{-1} = \{e\}$ then for all distinct c_1, c_2, $Ec_1 \cap Ec_2 = \emptyset$. So C is an E–correcting classical code. If the output falls in the set Ec the message is decoded as c.

For example, set $G = \mathbb{Z}_2^3$, where $\mathbb{Z}_2 = \{0, 1\}$ with addition mod 2. Let the error set E be $\{100, 010, 001\}$ and the code set C be $\{000, 111\}$. Then $E-E = \{000, 110, 011, 101\}$ and $C-C = C = \{000, 111\}$ implying $(E - E) \cap (C - C) = \{000\}$.

In order to formulate our next proposition we introduce some notation. Let A be a finite abelian group with operation $+$, null element 0 and character group $\hat{A}$. In the Hilbert space $\mathcal{H} = L^2(A)$ of complex valued functions on A we define the unitary operators $U_a, a \in A, V_\alpha, \alpha \in \hat{A}$ by $(U_a f)(x) = f(x + a)$, $(V_\alpha f)(x) = \alpha(x)f(x)$. Then we have the Weyl commutation rules:

$$U_a U_b = U_{a+b},\ V_\alpha V_\beta = V_{\alpha\beta},\ U_a V_\alpha = \alpha(a) V_\alpha U_a.$$

Let $E \subset A, F \subset \hat{A}$ and let

$$\mathcal{E}(E, F) = \text{lin}\,\{U_a V_\alpha \mid a \in E, \alpha \in \hat{F}\}.$$

Our aim is to construct a quantum code which is $\mathcal{E}(E, F)$–correcting by using subgroups $C_1 \subset C_2 \subset A$. To this end, for any subgroup $C \subset A$, we define

$$C^\perp = \{\alpha \mid \alpha \in \hat{A},\ \alpha(x) = 1, \quad \text{for all } x \in C\}.$$

$C^{\perp}$ is called the *annihilator* of C. We have $C_1^{\perp} \supset C_2^{\perp}$. Clearly $C_1^{\perp}$, $C_2^{\perp}$ are subgroups of the character group $\hat{A}$ under multiplication. Suppose

$$\begin{aligned}(E-E)\cap C_2 &= \{0\}\\ F^{-1}F\cap C_1^{\perp} &\subseteq C_2^{\perp},\end{aligned}$$

and let S be the cross section for C_2/C_1 in the sense that $S \subset C_2$ and $C_2 = \cup_{a\in S} C_1 + a$ is a coset decomposition (or partition) of C_2 by C_1–cosets. Note that

$$S^{\perp} \stackrel{\Delta}{=} \{\alpha \mid \alpha \in \hat{A},\, \alpha(a) = 1 \text{ for all } a \in S\}$$

is a subgroup of $\hat{A}$. One may view C_2 as a classical E-correcting group code in A. Define

$$\psi_a(x) = (\#C_1)^{-\frac{1}{2}}\, 1_{C_1+a}(x),\ a \in S.$$

Theorem 7.1.10 $\lin\{\psi_a \mid a \in S\}$ *is an* $\mathcal{E}(E,F)$*-correcting quantum code of dimension* $\frac{\#C_2}{\#C_1}$.

Proof Note that $\langle \psi_{a_1} \mid \psi_{a_2}\rangle = \delta_{a_1 a_2},\ a_1, a_2 \in S$. It is enough to verify Knill-Laflamme conditions for

$$L_1 = U_{a_1}V_{\alpha_1},\ L_2 = U_{a_2}V_{\alpha_2},\ a_1, a_2 \in E,\ \alpha_1, \alpha_2 \in F.$$

Then by the Weyl commutation rules we have

$$L_1^{\dagger}L_2 = \alpha_1(a_2 - a_1)U_{a_2-a_1}V_{\alpha_1^{-1}\alpha_2},\ a_2 - a_1 \in E - E, \alpha_1^{-1}\alpha_2 \in F^{-1}F.$$

Let $a_1, a_2 \in S,\ a_1 \neq a_2$. We have for $a \in E - E, \alpha \in F^{-1}F$,

$$\langle \psi_{a_1} | U_a V_\alpha | \psi_{a_2}\rangle = (\#C_1)^{-1} \sum_{x\in A} 1_{C_1+a_1+a}(x)\alpha(x)1_{C_1+a_2}(x). \quad (7.1.11)$$

The x-th term in the summation on the right side of (7.1.11) is not equal to zero only if $x \in (C_1 + a_1 + a) \cap (C_1 + a_2)$, which implies the existence of $x_1, x_2 \in C_1$ such that

$$\begin{aligned}x_1 + a_1 + a &= x_2 + a_2\\ \implies a &= (x_2 - x_1) + a_2 - a_1.\end{aligned} \quad (7.1.12)$$

In (7.1.12) $a-(x_1-x_2)$ belongs to C_2 whereas a_2-a_1 belongs to $E-E$. By hypothesis $(E-E)\cap C_2 = \{0\}$. Thus the x-th term vanishes if $a \neq 0$.

Now consider the case $a = 0$. Then for $a_1, a_2 \in S$, $a_1 \neq a_2$, $C_1 + a_1$ and $C_1 + a_2$ are two disjoint cosets and therefore the right hand side of (7.1.11) vanishes once again. In other words

$$\langle \psi_{a_1} | U_a V_\alpha | \psi_{a_2} \rangle = 0 \quad \text{for all } a_1 \neq a_2,\ a \in E - E,\ \alpha \in F^{-1}F.$$

Now let us consider the case $a_1 = a_2 = b \in S$. Then the left hand side of (7.1.11) is equal to

$$(\#C_1)^{-1} \sum_{x \in A} 1_{C_1+b+a}(x)\, 1_{C_1+b}(x)\, \alpha(x). \tag{7.1.13}$$

The x-th term is not equal to zero only if

$$\begin{aligned} x \in (C_1 + b + a) \cap (C_1 + b) &\implies (C_1 + a) \cap C_1 \neq \emptyset \\ &\implies a \in C_1 \cap (E - E) \\ &\implies a = 0. \end{aligned}$$

Thus the expression (7.1.13) vanishes if $a \neq 0$. If $a = 0$ then (7.1.13) is equal to

$$(\#C_1)^{-1} \sum_{x \in A} 1_{C_1+b}(x)\, \alpha(x) = (\#C_1)^{-1} \alpha(b) \sum_{x \in C_1} \alpha(x).$$

If $\alpha \notin C_1^\perp$ then, α is a nontrivial character for C_1 and by Schur orthogonality the right hand side vanishes. If $\alpha \in C_1^\perp$, then

$$\alpha \in C_1^\perp \cap F^{-1}F \implies \alpha \in C_2^\perp \implies \alpha(b) = 1.$$

Thus the expression (7.1.13) is independent of b. In other words, Knill-Laflamme conditions are fulfilled for the orthonormal set $\{\psi_a \,|\, a \in S\}$.

□

Theorem 7.1.14 *Let $C_1 \subset C_2 \subset A$ be subgroups. Consider the subgroups $C_2^\perp \subset C_1^\perp \subset \hat{A}$ and the coset decomposition $C_1^\perp = \cup_{\alpha \in \tilde{S}} C_2^\perp \alpha$ with respect to the cross section $\tilde{S}$. Define*

$$\psi_\alpha = (\#C_2)^{-\frac{1}{2}} 1_{C_2} \alpha,\ \alpha \in \tilde{S}.$$

Let $E \subset A$, $F \subset \hat{A}$ be such that $(E\text{–}E) \cap C_2 = \{0\}$, $F^{-1}F \cap C_1^\perp \subset C_2^\perp$. Then $\mathrm{lin}\{\psi_\alpha \mid \alpha \in \tilde{S}\}$ is an $\mathcal{E}(E, F)$–correcting quantum code of dimension $(\#C_2)/(\#C_1)$.

Proof Let $b \in E - E$, $\beta \in F^{-1}F$, $\alpha_1, \alpha_2 \in \tilde{S}$. Then

$$\langle \psi_{\alpha_1} | U_b V_\beta | \psi_{\alpha_2} \rangle = (\#C_2)^{-1} \sum_x 1_{C_2+b}(x) \overline{\alpha_1(x)} \alpha_2(x) 1_{C_2}(x) \beta(x) \alpha_1(b). \quad (7.1.15)$$

If the x-th term in the right hand side of equation (7.1.15) is not equal to zero, then $C_2 + b \cap C_2 \neq \emptyset \implies b \in C_2 \cap (E - E) \implies b = 0$. Thus the right hand side of equation (7.1.15) vanishes whenever $b \neq 0$ for any α_1, α_2 in $\tilde{S}$. Let $b = 0$. Then the right hand side of equation (7.1.15) is

$$(\#C_2)^{-1} \sum_{x \in C_2} \overline{\alpha_1(x)} \alpha_2(x) \beta(x). \quad (7.1.16)$$

If $\alpha_1 = \alpha_2 = \alpha \in \tilde{S}$ this becomes $(\#C_2)^{-1} \sum_{x \in C_2} \beta(x)$ which is independent of $\alpha \in \tilde{S}$. So we consider the case $b = 0, \alpha_1 \neq \alpha_2$, $\alpha_1, \alpha_2 \in \tilde{S}$. Then the expression (7.1.16) is not equal to zero only if $\overline{\alpha_1}\alpha_2\beta \in C_2^\perp$. This implies $\beta \in C_1^\perp \cap F^{-1}F$. So by hypothesis β is in $C_2^\perp$. This implies $\overline{\alpha_1}\alpha_2 \in C_2^\perp$. i.e., α_1 and α_2 lie in the same coset of $C_2^\perp$ in $C_1^\perp$. This is impossible. So expression (7.1.16) must be equal to zero. In other words Knill-Laflamme conditions are fulfilled.

□

7.2 Some Definitions

7.2.1 Invariants

Let $\mathcal{C}$ be an $\mathcal{E}$ correcting quantum code with recovery operators $R_1, \ldots, R_l$. Suppose U is a unitary operator such that $U\mathcal{E}U^{-1} \subseteq \mathcal{E}$. Define, $S_j = UR_jU^{-1}$. We have $R_jL\psi = \lambda_j(L)\psi$, where $\psi \in \mathcal{C}$ and $L \in \mathcal{E}$. Since $\tilde{L} = U^{-1}LU$ is an element of $\mathcal{E}$ we have

$$\begin{aligned} S_jLU\psi &= UR_jU^{-1}LU\psi \\ &= UR_j\tilde{L}\psi \\ &= \lambda_j(\tilde{L})U\psi. \end{aligned}$$

In other words, if $\mathcal{C}$ is an error correcting quantum code with recovery operators $R_1, R_2, \ldots, R_l$ then for any unitary U, satisfying $U\mathcal{E}U^* \subseteq \mathcal{E}$, $U(\mathcal{C})$ is also $\mathcal{E}$–correcting with recovery operators $S_1, S_2, \ldots S_k$, where $S_j = UR_jU^{-1}$ for all j.

Definition 7.2.1 Two $\mathcal{E}$-correcting quantum codes $\mathcal{C}_1, \mathcal{C}_2$, are said to be equivalent if and only if there exists a unitary operator U, satisfying $U\mathcal{E}U^* \subseteq \mathcal{E}$, such that $U(\mathcal{C}_1) = \mathcal{C}_2$.

Remark 7.2.2 Finding invariants for the equivalence of $\mathcal{E}$-correcting quantum codes is an important problem in the development of the subject.

Let A be a finite set, called an *alphabet*, of cardinality N. An element $\boldsymbol{x}$ in A^n is called a *word* of length n. A word $\boldsymbol{x}$ is also written as $(x_1, x_2, \ldots, x_n)$. $C \subset A^n$ is called an $(n, M, d)_A$ code if, $\#C = M$ and

$$\min_{\boldsymbol{x},\boldsymbol{y}\in C, \boldsymbol{x}\neq\boldsymbol{y}} d(\boldsymbol{x}, \boldsymbol{y}) = d.$$

Here, $d(\boldsymbol{x}, \boldsymbol{y}) = \#\{i \mid x_i \neq y_i\}$. This is also known as the *Hamming distance* between $\boldsymbol{x}$ and $\boldsymbol{y}$.

If A is an abelian group with $+$ as its addition and 0 its null element then

$$w(\boldsymbol{x}) \stackrel{\Delta}{=} \#\{i \mid x_i \neq 0\}, \boldsymbol{x} = (x_1, x_2, \ldots, x_n)$$

is called the *weight* of $\boldsymbol{x}$. If $C \subset A^n$ is a subgroup with

$$d = \min_{\boldsymbol{x}\neq 0,\, \boldsymbol{x}\in C} w(\boldsymbol{x}), \#C = M,$$

then C is called an $(n, M, d)_A$ *group code*, and it is denoted by $\langle n, M, d\rangle_A$. If A is the additive group of a finite field $\mathbb{F}_q$ of q elements ($q = p^m$, for some prime p) and $C \subset \mathbb{F}_q^n$ is a linear subspace of the n-dimensional vector space $\mathbb{F}_q^n$ over $\mathbb{F}_q$ and $d = \min_{\boldsymbol{x}\neq\boldsymbol{0}} w(\boldsymbol{x})$, then C is called a *linear code* over $\mathbb{F}_q$ with *minimum distance* d and written as $[n, k, d]_q$ code, where $k = \dim C$. When $q = 2$, it is simply called an $[n, k, d]$ code (binary code).

An $\langle n, M, d\rangle_A$ code is *t-error correcting* when $t = \left\lfloor \frac{d-1}{2} \right\rfloor$.

7.2.2 What is a t–error correcting quantum code?

Let $\mathcal{G}$ be a Hilbert space of finite dimension and $\mathcal{H} = \mathcal{G}^{\otimes^n}$ its n-fold tensor product. A typical example is $\mathcal{G} = \mathbb{C}^2$, so that $\mathcal{H}$ is an n-qubit Hilbert space. Consider all operators in $\mathcal{H}$ of the form

$$X = X_1 \otimes X_2 \otimes \cdots \otimes X_n,$$

where $\#\{i \mid X_i \neq I\} \leq t$.

Denote by $\mathcal{E}_t$ the linear span of all such operators. An element $X \in \mathcal{E}_t$ is called an error operator of *strength* at most t. An $\mathcal{E}_t$-correcting quantum code $\mathcal{C} \subset \mathcal{H}$ is called a *t–error correcting quantum code*.

Remark 7.2.3 In an n–qubit quantum computer, if errors affect at most t wires among the n wires, they can be corrected by a t–error correcting quantum code.

7.2.3 A good basis for $\mathcal{E}_t$

We now construct a "good basis" for $\mathcal{E}_t \subset \mathcal{B}(\mathcal{H})$. Suppose $\dim \mathcal{G} = N$. Consider any abelian group A of cardinality N and identify $\mathcal{G}$ with $L^2(A)$. We define the unitary operators U_a, V_α and $W_{a,\alpha}$ as follows

$(U_a f)(x) = f(x+a)$ where $a \in A$,

$(V_\alpha f)(x) = \alpha(x) f(x)$ where $f \in L^2(A), \alpha \in \hat{A}$, and $W_{a,\alpha} = U_a V_\alpha$.

Then we have

$$W_{(a,\alpha)} W_{(b,\beta)} = \overline{\alpha(b)} W_{a+b,\alpha\beta}$$
$$\text{and } \operatorname{Tr} W^\dagger_{(a,\alpha)} W_{(b,\beta)} = (\delta_{a,b}\delta_{\alpha,\beta}) N.$$

The family $\{W_{(a,\alpha)} \mid (a,\alpha) \in A \times \hat{A}\}$ is irreducible and the set

$$\left\{ \frac{1}{\sqrt{N}} W_{(a,\alpha)} \middle| (a,\alpha) \in A \times \hat{A} \right\}$$

is an orthonormal basis for the Hilbert space $\mathcal{B}(\mathcal{G})$ with scalar product $\langle X, Y \rangle = \operatorname{Tr} X^\dagger Y$, $X, Y \in \mathcal{B}(\mathcal{G})$. For $(\boldsymbol{a}, \boldsymbol{\alpha}) \in A^n \times \hat{A}^n (\cong (A \times \hat{A})^n)$ define $W_{(\boldsymbol{a},\boldsymbol{\alpha})} = W_{(a_1,\alpha_1)} \otimes W_{(a_2,\alpha_2)} \otimes \cdots \otimes W_{(a_n,\alpha_n)}$, so that

$$W_{(\boldsymbol{a},\boldsymbol{\alpha})} W_{(\boldsymbol{b},\boldsymbol{\beta})} = \prod_{i=1}^{n} \overline{\alpha_i(b_i)} W_{(\boldsymbol{a}+\boldsymbol{b},\boldsymbol{\alpha}\boldsymbol{\beta})}$$

and

$$\left\{ \frac{1}{N^{\frac{n}{2}}} W_{(\boldsymbol{a},\boldsymbol{\alpha})} \middle| (\boldsymbol{a},\boldsymbol{\alpha}) \in A^n \times \hat{A}^n \right\}$$

is an orthonormal basis for $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{G}^{\otimes^n})$. Define

$$w(\boldsymbol{a},\boldsymbol{\alpha}) = \#\{i \mid (a_i,\alpha_i) \neq (0,1)\}$$

the weight of $(\boldsymbol{a},\boldsymbol{\alpha})$ in the abelian group $(A \times \hat{A})^n$.

Then

$$\{W_{(\boldsymbol{a},\boldsymbol{\alpha})} \mid w(\boldsymbol{a},\boldsymbol{\alpha}) \leq t\}$$

is a linear basis for the subspace $\mathcal{E}_t$.

A subspace $\mathcal{C} \subset \mathcal{G}^{\otimes^n}$ is called a quantum code of *minimum distance* d, if $\mathcal{C}$ has an orthonormal basis $\psi_1, \psi_2, \ldots, \psi_k$ satisfying

1. $\langle\psi_i|W_{(\boldsymbol{a},\boldsymbol{\alpha})}|\psi_j\rangle = 0$, $i \neq j$, $w(\boldsymbol{a},\boldsymbol{\alpha}) \leq d$,
2. $\langle\psi_i|W_{(\boldsymbol{a},\boldsymbol{\alpha})}|\psi_i\rangle$ is independent of i whenever $w(\boldsymbol{a},\boldsymbol{\alpha}) \leq d$,
3. Either condition (1) or condition (2) is false, for some $(\boldsymbol{a},\boldsymbol{\alpha})$ with $w(\boldsymbol{a},\boldsymbol{\alpha}) = d+1$.

Such a quantum code is $\lfloor \frac{d-1}{2} \rfloor$-error correcting. We call it an $[\![n,k,d]\!]_A$ quantum code.

7.3 Examples

7.3.1 A generalized Shor code

We begin with a few definitions. Let A be a finite abelian group with binary operation $+$ and identity element 0. Let $\hat{A}$ denote its character group. Let $\mathcal{H}$ be the Hilbert space $L^2(A)^{\otimes^n}$. Let $U_{\boldsymbol{a}}$ and $V_{\boldsymbol{\alpha}}$ denote the Weyl operators. Let $C_n \subset A^n$ be a t-error correcting ($d(C_n) \geq 2t+1$) group code of length n with alphabet A. Let $D_{n,m} \subset (\hat{C_n})^m$ be a t-error correcting group code with alphabet $\hat{C_n}$ of length m.

An element in $D_{n,m}$ is denoted by $\boldsymbol{\chi}$. Sometimes we also denote by $\boldsymbol{\chi}$ the m-tuple $\chi_1, \chi_2, \ldots, \chi_m$, where each χ_i is in $\hat{C_n}$. Define

$$f_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \begin{cases} \#C_n^{-\frac{1}{2}}\boldsymbol{\alpha}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in C_n; \\ 0 & \text{otherwise.} \end{cases}$$

Let $F_{\boldsymbol{\chi}} = f_{\chi_1} \otimes f_{\chi_2} \otimes \cdots \otimes f_{\chi_m}$, where $\boldsymbol{\chi}$ is in $D_{n,m}$.

Theorem 7.3.1 *$\{F_{\boldsymbol{\chi}} \mid \boldsymbol{\chi} \in D_{n,m}\}$ is a t-error correcting quantum code in $L^2(A)^{\otimes^{mn}} \cong L^2(A^{mn})$.*

Proof Let $(\boldsymbol{a},\boldsymbol{\alpha}) \in A^{mn} \times \hat{A}^{mn}$ such that $w(\boldsymbol{a},\boldsymbol{\alpha}) \leq 2t$. We have

$$\langle F_{\boldsymbol{\beta}}|U_{\boldsymbol{a}}V_{\boldsymbol{\alpha}}|F_{\boldsymbol{\gamma}}\rangle = \sum_{\boldsymbol{x}\in A^{mn}} \prod_{j=1}^{m} \overline{f}_{\beta_j}(\boldsymbol{x}^{(j)} - \boldsymbol{a}^{(j)}) f_{\gamma_j}(\boldsymbol{x}^{(j)})\boldsymbol{\alpha}(\boldsymbol{x}). \quad (7.3.2)$$

Note that $w(\boldsymbol{a}) \leq 2t$ in A^{mn} and $w(\boldsymbol{\alpha}) \leq 2t$ in $\hat{A}^{mn}$.

Case 1: Let $\boldsymbol{a} \neq 0$. Then $\boldsymbol{a}^{(j)} \neq 0$ for some $j = j_0$, and $w(\boldsymbol{a}) \leq 2t$ implies $w(\boldsymbol{a}^{(j_0)}) \leq 2t$. Then $C_n + \boldsymbol{a}^{(j_0)} \cap C_n = \emptyset$. So every summand in the right hand side of equation (7.3.2) vanishes.

Case 2: Let $\boldsymbol{a} = 0$. Then the right hand side of equation (7.3.2) reduces to

$$\frac{1}{\#C_n} \sum_{\boldsymbol{x} \in C_n^m} \overline{\boldsymbol{\beta}}(\boldsymbol{x}) \boldsymbol{\gamma}(\boldsymbol{x}) \boldsymbol{\alpha}(\boldsymbol{x}).$$

Let $\boldsymbol{\beta} \neq \boldsymbol{\gamma}$, $\boldsymbol{\beta}, \boldsymbol{\gamma} \in D_{n,m}$. Then $\overline{\boldsymbol{\beta}}\boldsymbol{\gamma} \in D_{n,m}$ (a group code), and $w(\overline{\boldsymbol{\beta}}\boldsymbol{\gamma}) \geq 2t+1$. Since $w(\boldsymbol{\alpha}) \leq 2t$, $\boldsymbol{\alpha} \big|_{C_n^m}$ has weight $\leq 2t$. So $\overline{\boldsymbol{\beta}}\,\boldsymbol{\gamma}\,\boldsymbol{\alpha} \big|_{C_n^m}$ is nontrivial. By Schur orthogonality relations the right hand side of equation (7.3.2) is equal to 0.

We now consider the case when $\boldsymbol{\beta} = \boldsymbol{\gamma}$. Then the right hand side of equation (7.3.2) reduces to $\frac{1}{\#C_n} \sum_{\boldsymbol{x} \in C_n^m} \boldsymbol{\alpha}(\boldsymbol{x})$ which is independent of $\boldsymbol{\beta}$.

Thus the Knill-Laflamme conditions are fulfilled.

□

7.3.2 Specialization to $A = \{0,1\}, m = 3, n = 3$

Design of a 9-qubit, 1 error correcting, 2-dimensional code.

$$\begin{aligned}
C_3 &= \{000, 111\} \\
\hat{C}_3 &\text{ has two elements,} \\
\chi_1(000) &= \chi_1(111) = 1 \text{ (identity character) and} \\
\chi_2(000) &= +1, \chi_2(111) = -1. \\
f_{\chi_1} &= \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \\
f_{\chi_2} &= \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) \\
D_{3,3} &= \{(\chi_1, \chi_1, \chi_1), (\chi_2, \chi_2, \chi_2)\} \\
F_{\chi_1\chi_1\chi_1} &= f_{\chi_1}^{\otimes^3} \\
F_{\chi_2\chi_2\chi_2} &= f_{\chi_2}^{\otimes^3}.
\end{aligned}$$

Thus, we encode 0 as $F_{\chi_1\chi_1\chi_1}$ and 1 as $F_{\chi_2\chi_2\chi_2}$. The circuit for implementing the code is shown in Figure 7.3. This code is called the *Shor code*.

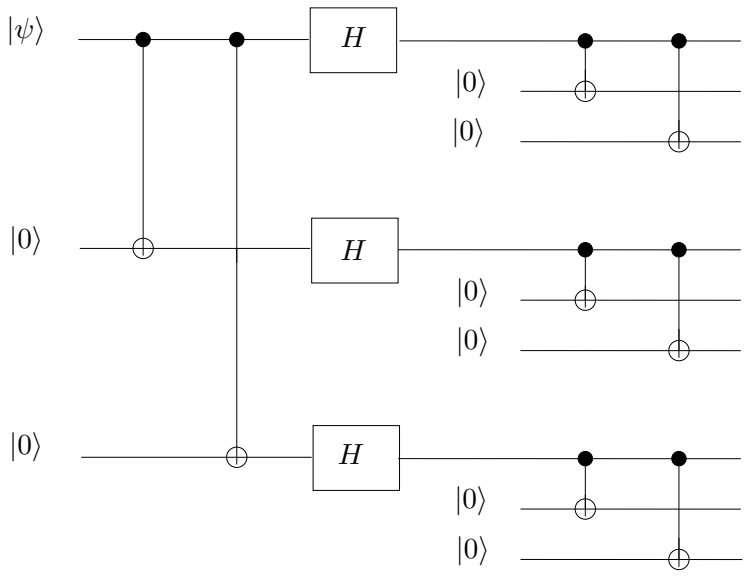

Figure 7.3: Circuit for encoding the Shor code.

7.3.3 Laflamme code

Laflamme found the following 5-qubit 1-error correcting quantum code.

$$\begin{aligned}
0 \mapsto |\psi_0\rangle = \frac{1}{4}\{&(|00000\rangle + |11000\rangle + |01100\rangle + |00110\rangle + |00011\rangle + |10001\rangle) \\
&- (|01010\rangle + |00101\rangle + |10010\rangle + |01001\rangle + |10100\rangle) \\
&- (|11110\rangle + |01111\rangle + |10111\rangle + |11011\rangle + |11101\rangle)\} \\
1 \mapsto |\psi_1\rangle = \frac{1}{4}\{&(|11111\rangle + |00111\rangle + |10011\rangle + |11001\rangle + |11100\rangle + |01110\rangle) \\
&- (|10101\rangle + |11010\rangle + |01101\rangle + |10110\rangle + |01011\rangle) \\
&- (|00001\rangle + |10000\rangle + |01000\rangle + |00100\rangle + |00010\rangle)\}.
\end{aligned}$$

The code can also be written in the following way. Let $a_0 = a_1 + a_2 + a_3 + a_4 + x \pmod 2$.

$$x \mapsto |\psi_x\rangle = \frac{1}{4} \sum_{a_1,a_2,a_3,a_4 \in \mathbb{Z}_2} (-1)^{(a_0a_2 + a_1a_3 + a_2a_4 + a_3a_0 + a_4a_1)} |a_0\rangle |a_1a_2a_3a_4\rangle.$$

This observation allows us to construct a simple circuit for implementing the Laflamme code. The circuit for the Laflamme code is shown in Figure 7.4.

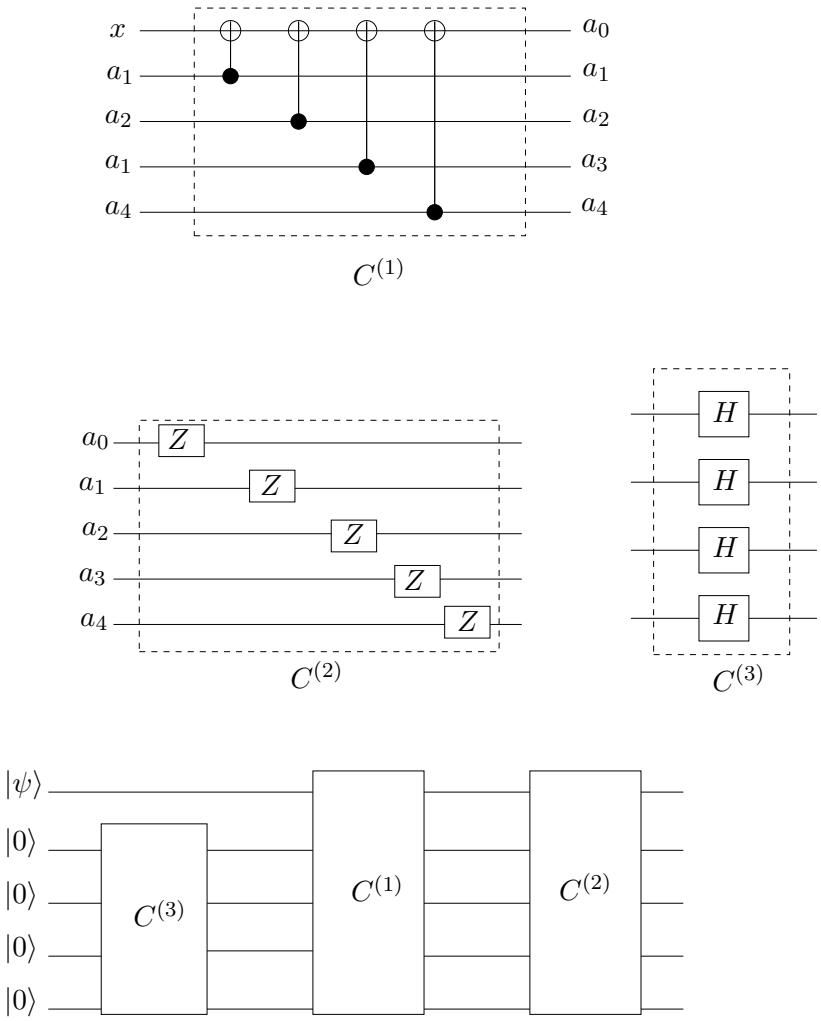

Figure 7.4: *Circuit for encoding the Laflamme code.*

7.3.4 Hadamard-Steane quantum code

Consider the Table 7.1. The ij^{th} entry, for $i, j > 1$, is the inner product of the i^{th} entry in the first row and j^{th} entry in the first column, computed over the field $\mathbb{F}_2$.

The portion inside the box is Hadamard $[7, 3, 4]$ simplex code. Let

	000	001	010	011	100	101	110	111
000	0	0	0	0	0	0	0	0
001	0	1	0	1	0	1	0	1
010	0	0	1	1	0	0	1	1
011	0	1	1	0	0	1	1	0
100	0	0	0	0	1	1	1	1
101	0	1	0	1	1	0	1	0
110	0	0	1	1	1	1	0	0
111	0	1	1	0	1	0	0	1

Table 7.1:

C be the set of all row vectors. Define

$$|\psi_0\rangle = \frac{1}{2\sqrt{2}} \sum_{\boldsymbol{x}\in C} |\boldsymbol{x}\rangle \text{ and } |\psi_1\rangle = \frac{1}{2\sqrt{2}} \sum_{\boldsymbol{x}\in C+(1,1,1,1,1,1,1,1)} |\boldsymbol{x}\rangle.$$

Then, $\text{lin}\{|\psi_0\rangle, |\psi_1\rangle\}$ is a 7-qubit single error correcting quantum code. Note that, $C \cup C + (1,1,1,1,1,1,1,1)$ is a group code of minimum distance 3.

Permute the columns to the order 4 6 7 1 2 3 5 in the table above. Then the enumerated rows can be expressed as

$$(x_1 \;\; x_2 \;\; x_3 \;\; x_1+x_2 \;\; x_1+x_3 \;\; x_2+x_3 \;\; x_1+x_2+x_3)$$

where x_1, x_2, x_3 vary in $\mathbb{F}_2$. In other words we have expressed the code as a parity check code with the first three positions for messages and the last four as parity checks. Then the Hadamard-Steane code can be expressed as

$$|\psi_a\rangle = \sum_{x_1,x_2,x_3} |x_1+a \;\; x_2+a \;\; x_3+a \;\; x_1+x_2+a \;\; x_1+x_3+a \;\; x_2+x_3+a \;\; x_1+x_2+x_3+a\rangle$$

where $a \in \{0,1\}$. Put $y_1 = x_1+x_3+a$, $y_2 = x_2+x_3+a$, $y_3 = x_1+x_2+x_3+a$. Then

$$|\psi_a\rangle = \sum_{y_1,y_2,y_3\in\{0,1\}} |y_2+y_3+a \; y_1+y_3+a \; y_1+y_2+y_3 \; y_1+y_2+a \; y_1 \;\; y_2 \;\; y_3\rangle.$$

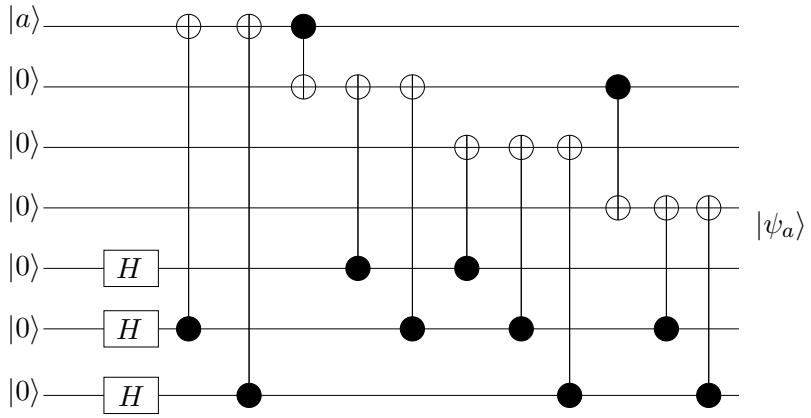

Figure 7.5: *Circuit implementing the Steane-Hadamard code.*

This shows that the code can be implemented by the circuit shown in Figure 7.5.

Exercise 7.3.3 Verify directly the Knill-Laflamme conditions for $\{|\psi_0\rangle, |\psi_1\rangle\}$, for single error correction.

7.3.5 Codes based on Bush matrices

Let $\mathbb{F}_q = \{a_1, a_2, \ldots, a_q\}$ be the field of $q = p^m$ elements, where p is prime.
Let $\mathbb{P}(t, q) = \{$ all polynomials of degree $\leq t$ with coefficients from $\mathbb{F}_q\}$, a linear space of dimension q^{t+1}.

We enumerate the elements of $\mathbb{P}(t-1, q)$, $t-1 \leq q$ as φ_0, φ_1, …, φ_{N-1} and construct the matrix B_t of order $q^t \times q$, $q^t = N$ as follows :

	a_1	a_2	$\cdots$	a_j	$\cdots$	a_q
$\varphi_0 = \mathbf{0}$	0	0	$\cdots$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
φ_i	$\varphi_i(a_1)$	$\varphi_i(a_2)$	$\cdots$	$\varphi_i(a_j)$	$\cdots$	$\varphi_i(a_q)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
φ_{N-1}	$\varphi_{N-1}(a_1)$	$\varphi_{N-1}(a_2)$	$\cdots$	$\varphi_{N-1}(a_j)$	$\cdots$	$\varphi_{N-1}(a_q)$

Denote the linear space of all the row vectors in B_t also by B_t.

Proposition 7.3.4 *B_t is a linear code of minimum distance $q-t+1$.*

Proof Consider the i-th row in B_t, $i \neq 0$. φ_i is a nonzero polynomial of degree $\leq t-1$. So φ_i has at most $t-1$ zeroes. Thus, the weight of this row $\geq q-t+1$. On the other hand consider the polynomial

$$\varphi(x) = (x-a_1)(x-a_2)\cdots(x-a_{t-1}).$$

Its zeros are exactly $a_1, a_2, \ldots a_{t-1}$. Thus, the weight of the corresponding row is $q-t+1$.

□

Corollary 7.3.5 *B_t is a $\lfloor \frac{q-t}{2} \rfloor$–error correcting group code.*

If E_t is the Hamming sphere of radius $\lfloor \frac{q-t}{2} \rfloor$ with $(0,\ldots,0)$ as center in $\mathbb{F}_q^{q^t}$ then $(E_t - E_t) \cap B_t = \{\mathbf{0}\}$.

Proposition 7.3.6 *Let $\boldsymbol{\alpha} \in B_t^{\perp} \subset (\hat{\mathbb{F}}_q)^q$. If $\boldsymbol{\alpha} \neq 1$, then $w(\boldsymbol{\alpha}) \geq t+1$. Thus $B_t^{\perp}$ is a $\lfloor \frac{t}{2} \rfloor$ error correcting group code. If F_t is the Hamming sphere of radius $\lfloor \frac{t}{2} \rfloor$ then $F_t^{-1} F_t \cap B_t^{\perp} = \{\mathbf{1}\}$.*

Proof Suppose $w(\boldsymbol{\alpha}) = r$, where $0 < r \leq t$. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_q)$, $\alpha_i \in \hat{\mathbb{F}}_q$, $\alpha_i \neq 1$ if and only if $i \in \{i_1 < i_2 < \cdots < i_r\}$. Write $b_j = a_{i_j}$, $j = 1, 2, \ldots, r$. For arbitrary $c_1, c_2, \ldots, c_r$ in $\mathbb{F}_q$ consider the Lagrange polynomial (for interpolation)

$$\varphi(x) = \sum c_j \frac{(x-b_1)(x-b_2)\cdots\widehat{(x-b_j)}\cdots(x-b_r)}{(b_j-b_1)(b_j-b_2)\cdots\widehat{(b_j-b_j)}\cdots(b_j-b_r)},$$

where "$\widehat{\ }$" indicates omission of that term. Then φ is a polynomial of degree $r-1$ $(\leq t-1)$ and $\varphi(b_j) = c_j$, $j = 1, 2, \ldots, r$. Corresponding to φ there is a row in B_t. Evaluating $\boldsymbol{\alpha}$ on this row we get

$$\boldsymbol{\alpha}(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_q)) = \prod_{j=1}^{r} \alpha_{i_j}(c_j) = 1,$$

since $\boldsymbol{\alpha} \in B_t^{\perp}$. Since c_j's are arbitrary, we have $\alpha_{i_j} = 1$ for all $j = 1, 2, \ldots, r$, a contradiction.

□

We can now use Theorem 7.1.10 and Theorem 7.1.14 to the case $C_1 \subset C_2 \subset A^q$, $A = \mathbb{F}_q$, as an additive group, $C_1 = B_{t'}$, $C_2 = B_t$, $0 < t' < t < q$. Then $B_t = B_{t'} \oplus S$, where S consists of all polynomials of the form

$$s(x) = x^{t'}(a_0 + a_1 x + \cdots + a_{t-t'-1}x^{t-t'-1}).$$

For any polynomial φ consider the state $|\varphi\rangle = |\varphi(a_1)\varphi(a_2)\ldots\varphi(a_q)\rangle$. For any $s \in S$ define

$$\psi_s = q^{-\frac{t'}{2}} \sum_{\varphi \in \mathbb{P}(t'-1,q)} |\varphi\rangle.$$

Then $\mathcal{C}_{t,t'} = \text{lin}\{\psi_s \mid s \in S\}$ is a quantum code with dim $\mathcal{C}_{t,t'} = q^{t-t'}$, which can correct $\left\lfloor \frac{q-t}{2} \right\rfloor \wedge \left\lfloor \frac{t'}{2} \right\rfloor$ errors.

Remark 7.3.7 Choose $t = \lfloor \theta q \rfloor$, $t' = \lfloor \theta' q \rfloor$, $0 < \theta' < \theta < 1$. Then, as $q \to \infty$, we have

$$\frac{\log \dim \mathcal{C}_{t,t'}}{\log \dim \mathcal{H}} = \frac{t - t'}{q} = \frac{\lfloor \theta q \rfloor - \lfloor \theta' q \rfloor}{q} \to (\theta - \theta').$$

Therefore,

$$\frac{\#\text{ errors corrected}}{\#\text{ qubits}} \geq \frac{\left\lfloor \frac{(1-\theta)}{2} \right\rfloor \wedge \left\lfloor \frac{\theta' q}{2} \right\rfloor}{q} \to \frac{1-\theta}{2} \wedge \frac{\theta'}{2}$$

as $q \to \infty$.

Then, for $\theta = \frac{3}{4}$, $\theta' = \frac{1}{4}$ we get, $\theta - \theta' = \frac{1}{2}$ and $\frac{1-\theta}{2} \wedge \frac{\theta'}{2} = \frac{1}{8}$. It means 50% of the qubits are used for sending the messages, 50% for error checking and up to $12\frac{1}{2}\%$ errors can be corrected.

7.3.6 Quantum codes from BCH codes

In this example we use the celebrated BCH (Bose-Chaudhuri-Hocquenhem) codes to construct a quantum code. We begin with a few facts from classical coding theory. Let $\mathbb{F}_q^n$ be a vector space over the finite field $\mathbb{F}_q$ with $q = p^m$, where p is a prime. Choose and fix a primitive element α of $\mathbb{F}_{q^n}$.

Let σ be a cyclic permutation defined by

$$\sigma(a_0, \ldots, a_{n-1}) \mapsto (a_{n-1}, a_0, \ldots, a_{n-2}).$$

Then a subspace $C \subset \mathbb{F}_q^n$ invariant under the cyclic permutation σ is called a *cyclic code* of length n. For every word $\boldsymbol{w} = (w_0, \ldots, w_{n-1}) \in \mathbb{F}_q^n$ we associate the *word polynomial* $w(x) = w_0 + w_1 x + \cdots + w_{n-1}x^{n-1}$. If w is in C it is called the *code word polynomial*. Let $\mathcal{R}_n = \mathbb{F}_q[x]/(x^n - 1)$. Then $\mathcal{R}_n$ can be viewed as a vector space over $\mathbb{F}_q$ and it is isomorphic to $\mathbb{F}_q^n$. Under the identification $w \rightsquigarrow w(x)$ the image C^* of a cyclic code C in $\mathcal{R}_n$ is an ideal with a single generator polynomial g_C. Without loss of generality we may assume g_C to be monic and therefore unique. It is known that g_C is a divisor of $x^n - 1$. If $\deg(g_C) = k$ then $\dim C = n - k$. If g_C has a string of successive powers $\alpha^a, \alpha^{a+1}, \ldots, \alpha^{a+b-2}$ as its roots and $0 \leq a < a + b - 2 \leq q^n - 2$, then $d(C) \geq b$ (where $d(C)$ is the minimum distance of C). For any cyclic code denote

$$C^{\perp} = \{\boldsymbol{x} \mid \boldsymbol{xy} = x_1 y_1 + \cdots + x_n y_n = 0, \quad \text{for all } \boldsymbol{y} \in C\}.$$

Then $C^{\perp}$ is also a cyclic code called the dual of C.

Conversely if g is a divisor of $x^n - 1$ then there exists a unique cyclic code C_g generated by g. Suppose $x^n - 1 = gh$ where $g(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k$, $h(x) = b_0 + b_1 x + \cdots + b_{n-k-1}x^{n-k-1} + x^{n-k}$ so that $a_0 b_0 = -1$. Define $\tilde{h} = b_0^{-1}(1 + b_{n-k-1}x + \cdots + b_0 x^{n-k})$. If h has a string of successive powers $\alpha^l, \alpha^{l+1}, \ldots, \alpha^{l+m+2}$ as its roots then so does the polynomial $\tilde{h}$ which can be written as

$$\tilde{h} = (-1)^{n-k}(\beta_1 \cdots \beta_{n-k})^{-1}(1 - \beta_1 x) \cdots (1 - \beta_{n-k}x)$$

where $\beta_1, \ldots, \beta_{n-k}$ are the roots of h in $\mathbb{F}_{q^n}$. It is known that $C^{\perp} = C_{\tilde{h}}$ and therefore it follows that $d(C^{\perp}) \geq m$. (For complete proofs we refer to [15, 7, 13]).

Let $x^n - 1 = g_1 g_2 g_3$, $d(C_{g_1}) = d_1$, $d(C_{g_3}) = d_3$. Note that $C_{g_1 g_2}^{\perp} = C_{\tilde{g_3}}$. By Theorem 7.1.10 we get a quantum code $\mathcal{C}$ of dimension $(\#C_{g_1})/(\#C_{g_1 g_2}) = q^{\deg(g_2)}$. If C_{g_1} and C_{g_3} are respectively t_1 and t_3 – error correcting codes then $\mathcal{C}$ can correct $\min(t_1, t_3)$ errors.

Lecture 8

Classical Information Theory

8.1 Entropy as information

8.1.1 What is information?

Let us consider a simple statistical experiment of observing a random variable X, which takes one of the values $x_1, x_2, \ldots, x_n$ with respective probabilities $p_1, \ldots, p_n$ ($p_i \geq 0$ for all i and $\sum_i p_i = 1$). When we observe X we gain some information because the uncertainty regarding its value is eliminated. So the information gained is the uncertainty eliminated. We wish to have a mathematical model which gives us a measure of this information gained. A function which measures this information gained or the uncertainty associated with a statistical experiment must depend only on the probabilities p_i and it should be symmetric. This is based on the intuition that changing the names of the outcomes does not change the uncertainty associated with the random variable X.

The desirable properties of a function H which measures the uncertainty associated with a statistical experiment are listed below.

1) For each fixed n, $H(p_1, p_2, \ldots, p_n; n)$ is a nonnegative symmetric function of p_1, p_2, ..., p_n.
2) $H(\frac{1}{2}, \frac{1}{2}; 2) = 1$. This is to fix the scale of the measurement. One can look at the information obtained by performing one of the simplest statistical experiments. That is, tossing an unbiased coin and observing the outcome. An outcome of this experiment is said to give one unit of information.
3) $H(p_1, p_2, \ldots, p_n; n) = 0$ if and only if one of the p_i's is 1. This corresponds to the case when there is no uncertainty in the outcome of the experiment.

4) Let X and Y be two independent statistical experiments. Let XY denote the experiment where the experiments X and Y are performed together and the output is the ordered pair of the outcomes of X and Y. Then $H(XY) = H(X) + H(Y)$.
5) $H(p_1, p_2, \ldots, p_n; n)$ attains its maximum when $p_i = \frac{1}{n}$, for all $i \in 1, 2, \ldots, n$. That is, we gain maximum information when all possible outcomes are equally likely.
6) $H(p_1, p_2, \ldots, p_n, 0; n+1) = H(p_1, p_2, \ldots, p_n; n)$.
7) $H(p_1, p_2, \ldots, p_n; n)$ is continuous in $p_1, \ldots p_n$. This is a natural condition because we would like to say that, if two statistical experiments have the same number of possible outcomes and their associated probabilities are *close*, then the information contained in each of them should also be *close*.

Let $H_0 = -\sum_{j=0}^{n} p_j \log_2 p_j$. This function is also known as the *entropy* function. It can be verified that this function satisfies all the above desired properties.

Let X, Y be two statistical experiments in which the outcomes of X and Y are $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ respectively. Suppose

$$\Pr(X = x_i) = p_i,\ \Pr(Y = y_j \mid X = x_i) = q_{ij},\ \Pr(Y = y_j) = q_j.$$

Then $\Pr(X = x_i, Y = y_j) = p_i q_{ij}$. Let $H(q_{i1}, \ldots, q_{im}) = H_i(Y)$. We define *conditional entropy* as $H(Y \mid X) = \sum_{i=1}^{n} p_i H_i(Y)$, i.e. the *entropy of Y on knowing X*.

Exercise 8.1.1 Verify that H_0 defined earlier satisfies the following equality.

$$H_0(XY) = H_0(X) + H_0(Y \mid X). \tag{8.1.2}$$

This can be interpreted as follows: The total information obtained by performing the experiments X and Y together is equal to the sum of the information obtained by performing X and the information left in Y after knowing the outcome of X.

This seems to be a reasonable property that the function H should have. Note that Property 4) is a special case of equation (8.1.2). If we replace Property 4) by the hypothesis, $H(XY) = H(X) + H(Y \mid X)$ then there is a unique function which satisfies all the above properties. Hence H_0 is the only candidate as a measure of entropy. From now onwards we use H to denote the measure of entropy and $H(\boldsymbol{p})$ to denote

$H(p_1, p_2, \ldots, p_n; n)$. If a random variable X has a probability distribution $\boldsymbol{p}$, we sometimes write $H(\boldsymbol{p})$ instead of $H(X)$.

Note 8.1.3 If Property 4) is not changed then there can be other functions which satisfy properties 1) to 7). See [1] for other measures of entropy.

The entropy function H has several important properties. Some of them are listed in the following exercises.

Exercise 8.1.4 Show that $H(XY) \geq H(X)$.

Mutual information $H(X : Y)$ of two statistical experiments is defined as $H(X : Y) = H(X) + H(Y) - H(XY) = H(X) - H(X \mid Y)$. It is the information about X gained by observing Y.

Exercise 8.1.5 Show that $H(Y : X) \geq 0$, where X and Y are two statistical experiments.

Exercise 8.1.6 Let X, Y, Z be three statistical experiments. Then show that the inequality $H(X \mid Y) \geq H(X \mid YZ)$ holds.

Exercise 8.1.7 (Subadditivity) Show that $H(XY) \leq H(X) + H(Y)$, where X and Y are two statistical experiments.

Exercise 8.1.8 (Strong subadditivity) Show that

$$H(XYZ) + H(Y) \leq H(XY) + H(YZ),$$

where X, Y and Z are three statistical experiments. Equality holds if and only if $\{Z, Y, X\}$ is a Markov chain.

The following identity is also very useful.

Theorem 8.1.9 (Chain rule for conditional entropy)

$$H(X_1, \ldots, X_n \mid Y) = \\ H(X_1 \mid Y) + H(X_2 \mid YX_1) + \cdots + H(X_n \mid YX_1 \cdots X_{n-1}).$$

Proof We prove by induction.

Base case: $n = 2$.

$$\begin{aligned} H(X_1X_2 \mid Y) &= H(X_1X_2Y) - H(Y) \\ &= H(X_1X_2Y) - H(X_1Y) + H(X_1Y) - H(Y) \\ &= H(X_2 \mid X_1Y) + H(X_1 \mid Y) \\ &= H(X_1 \mid Y) + H(X_2 \mid X_1Y). \end{aligned}$$

Induction hypothesis: For all $n \in \{2, 3, \ldots k\}$

$$\begin{aligned} &H(X_1, \ldots, X_n \mid Y) = \\ &\qquad H(X_1 \mid Y) + H(X_2 \mid YX_1) + \cdots + H(X_n \mid YX_1 \cdots X_{n-1}). \end{aligned}$$

Induction step:

$$\begin{aligned} H(X_1, \ldots, X_{k+1}|Y) &= H(X_1|Y) + H(X_2 \ldots X_{k+1} \mid YX_1) \text{ (by base case)} \\ &= H(X_1 \mid Y) + H(X_2 \mid YX_1) + \cdots + \\ &\quad H(X_{k+1} \mid YX_1 \ldots X_k) \text{ (by induction hypothesis).} \end{aligned}$$

□

Exercise 8.1.10 (Data processing inequality) Let $X \to Y \to Z$ be a Markov chain. Then $H(X) \geq H(X : Y) \geq H(X : Z)$.

Exercise 8.1.11 (Data pipeline inequality) Let $X \to Y \to Z$ be a Markov chain. Then $H(Z) \geq H(Z : Y) \geq H(Z : X)$.

8.2 A Theorem of Shannon

Let A be an alphabet of size N. Denote by $S(A)$ the free semigroup generated by A. Any element $W \in S(A)$ can be expressed as $W = a_{i_1}a_{i_2} \ldots a_{i_n}$, where $a_{i_j} \in A$ for each j. We say that W is a *word* of length n. Let B be another alphabet, say of size M. Any map $C : A \to S(B)$ is called a *code* and any word in the image of C is called a *codeword*. Extend C to a map $\tilde{C} : S(A) \to S(B)$ by putting $\tilde{C}(W) = \tilde{C}(a_{i_1}a_{i_2} \ldots a_{i_n}) = C(a_{i_1})C(a_{i_2}) \ldots C(a_{i_n})$. We say that C is *uniquely decipherable* if $\tilde{C}$ is injective (or one to one). C is called an *irreducible code* if no code word of C is an extension of another code word. An irreducible code is uniquely decipherable. Indeed, in such a case we can recover a word W in $S(A)$ from its image $\tilde{C}(W)$ by just reading $\tilde{C}(W)$ left to right.

Theorem 8.2.1 *Let $A = \{a_1, \ldots, a_N\}$ and $B = \{b_1, \ldots, b_M\}$ be two alphabets. Let $C : A \to S(B)$ be an irreducible code. Let the lengths of the words $C(a_1), C(a_2), \ldots, C(a_N)$, be $n_1, n_2, \ldots, n_N$, respectively. Then*

$$M^{-n_1} + M^{-n_2} + \cdots + M^{-n_N} \leq 1. \tag{8.2.2}$$

Conversely, if $n_1, n_2, \ldots, n_N$ are nonnegative integers satisfying this inequality then there exists an irreducible code $C : A \to S(B)$ such that $C(a_i)$ has length n_i for each $i = 1, 2, \ldots, N$.

Proof Let $C : A \to S(B)$ be an irreducible code with $L = \max_i n_i$. Denote by w_i the number of code words of length i.
Necessity: Since there can be at most M words of length 1 we have $w_1 \leq M$. Since C is irreducible, words of length 2 which are extensions of the code words of length 1 cannot appear in the image of C. This gives $w_2 \leq M^2 - w_1 M$.

Continuing this way we get $w_L \leq M^L - w_1 M^{L-1} - \cdots - w_{L-1} M$. The last inequality can be rewritten as

$$w_1 M^{-1} + w_2 M^{-2} + \cdots + w_L M^{-L} \leq 1. \tag{8.2.3}$$

Sufficiency: We pick any w_1 words of length 1. Then we pick any w_2 words of length 2 which are not extensions of the w_1 words of length 1 already picked. This is possible because inequality (8.2.3) is satisfied. This way we keep picking words of required lengths.

□

Suppose the letters a_i, $i = 1, 2, \ldots, N$ of the alphabet A are picked with probabilities p_i, $i = 1, 2, \ldots, N$ respectively. Then the expected length of the code is $\sum_{i=1}^{N} p_i n_i$, where n_i is the length of $C(a_i)$.

Let

$$q_j = \frac{M^{-n_j}}{\sum_{i=1}^{N} M^{n_i}} \quad \text{and} \quad \overline{\ell}(C) = \sum_{i=1}^{N} p_i n_i.$$

By using the inequality "arithmetic mean is greater than or equal to geometric mean" we get

$$\prod \left(\frac{q_j}{p_j}\right)^{p_j} \leq \sum_j p_j \left(\frac{q_j}{p_j}\right) = \sum_j q_j = 1.$$

Taking logarithm on both sides and using (8.2.3), we get

$$\overline{\ell}(C) \geq -\frac{\sum p_i \log_2 p_i}{\log_2 M}.$$

Hence the average length of an irreducible code must be at least $-\frac{\sum p_i \log_2 p_i}{\log_2 M}$.

Let n_j be an integer between $-\frac{\log_2 p_j}{\log_2 M}$ and $-\frac{\log_2 p_j}{\log_2 M} + 1$ for all $j \in \{1, 2, \ldots, N\}$. Then $\sum_j M^{-n_j} \leq \sum_j p_j \leq 1$. By the above discussion we know that an irreducible code C' exists with length of $C'(a_i)$ equal to n_i. The expected length of this code word $\overline{\ell}(C')$ $(= \sum_j n_j p_j)$ satisfies

$$\frac{\sum p_j \log_2 p_j}{\log_2 M} \leq \overline{\ell}(C') \leq -\frac{\sum p_j \log_2 p_j}{\log_2 M} + 1.$$

Theorem 8.2.4 (Sardinas-Patterson, 1953) *Let $A = \{a_1, \ldots, a_N\}$ and $B = \{b_1, \ldots, b_M\}$ be two alphabets. Let $C : A \to S(B)$ be a uniquely decipherable code. Let the lengths of the words $C(a_1)$, $C(a_2)$, …, $C(a_N)$ be n_1, n_2, …, n_N respectively. Then $\sum_{j=1}^{N} M^{-n_j} \leq 1$.*

Proof Let $w_j = \#\{i \mid n_i = j\}$. Then the desired inequality can be rewritten as

$$\sum_{j=1}^{L} w_j M^{-j} \leq 1 \text{ where } L = \max(n_1, n_2, \ldots, n_N).$$

Let $Q(x) = \sum_{j=1}^{L} w_j x^j$ and let $N(k)$ denote the number of B words of length k. Then we have the following recursive relation.

$$N(k) = w_1 N(k-1) + w_2 N(k-2) + \cdots + w_L N(k-L), \qquad (8.2.5)$$

where $N(0) = 1$ and $N(j) = 0$ if $j < 0$. Consider the formal power series $F(x) = \sum_{k=0}^{\infty} N(k) x^k$. We know that $N(k) \leq M^k$. Hence the formal series converges in the case $|x| < M^{-1}$. From (8.2.5) we have $F(x) - 1 = Q(x)F(x) \Rightarrow F(x) = \frac{1}{1-Q(x)}$. $F(x)$ is analytic in the disc $(|x| < M^{-1})$ and $1 - Q(x) > 0$ when $|x| < M^{-1}$. Therefore, by continuity we have, $Q(M^{-1}) \leq 1$. This is the required inequality.

□

Corollary 8.2.6 *Let A and B be as in Theorem 8.2.1. Suppose the letters $a_1, a_2, \ldots, a_N$ are picked with probabilities $p_1, p_2, \ldots, p_N$ respectively. Then for any uniquely decipherable code C from A to $S(B)$ one has $\overline{\ell}(C) \geq \frac{-\sum p_i \log_2 p_i}{\log_2 M}$.*

Thus, Theorem 8.2.4 implies that corresponding to any uniquely decipherable code $C : A \to S(B)$ with length of code words $n_1, n_2, \ldots, n_N$ there exists an irreducible code $C' : A \to S(B)$ with lengths of code words $n_1, n_2, \ldots, n_N$.

Remark 8.2.7 Suppose an i.i.d. sequence $X_1, X_2, \ldots$ of letters from A comes from a source with $\Pr(X_j = a_i) = p_i$. Then $\Pr((X_1 X_2 \ldots X_n) = (a_{i_1} a_{i_2} \ldots a_{i_n})) = p_{i_1} p_{i_2} \ldots p_{i_n}$ and $H(X_1 X_2 \ldots X_n) = nH(p_1, \ldots, p_N)$. Now consider blocks of length n. The new alphabet is A^n. Encode $C : \boldsymbol{a} \to C(\boldsymbol{a})$, where $\boldsymbol{a} = a_{i_1} a_{i_2} \ldots a_{i_n}$ and $C(\boldsymbol{a}) \in S(B)$, in a uniquely decipherable form, so that the following inequalities hold.

$$\frac{nH(p_1, p_2, \ldots, p_N)}{\log_2 M} \leq \sum_{\boldsymbol{a}} p(\boldsymbol{a}) \ell(C(\boldsymbol{a})) < \frac{nH(p_1, p_2, \ldots, p_N)}{\log_2 M} + 1.$$

This implies

$$\left| \frac{\sum_{\boldsymbol{a}} p(\boldsymbol{a}) \ell(C(\boldsymbol{a}))}{n} - \frac{H(p_1, p_2, \ldots, p_N)}{\log_2 M} \right| < \frac{1}{n}. \tag{8.2.8}$$

In this block encoding procedure, the expected length of an encoded block is

$$\overline{\ell}(C) = \sum_{\boldsymbol{a}} p(\boldsymbol{a}) \ell(C(\boldsymbol{a})).$$

The ratio of expected length of an encoded block and the size of the $\boldsymbol{a}$ block, namely $\frac{\sum_{\boldsymbol{a}} p(\boldsymbol{a}) \ell(C(\boldsymbol{a}))}{n}$, is called the *compression coefficient*. Equation (8.2.8) tells us that, as n increases the compression coefficient tends to $\frac{H(p_1, p_2, \ldots, p_N)}{\log_2 M}$.

8.3 Stationary Source

We consider a discrete information source $\mathcal{I}$ which outputs elements $x_n \in A$, $n = 0, \pm 1, \pm 2, \ldots$ where A is a finite alphabet. Thus a 'possible life history' of the output can be expressed as a bilateral sequence

$$\boldsymbol{x} = (\ldots, x_{-1}, x_0, x_1, x_2, \ldots), \quad x_n \in A. \tag{8.3.1}$$

Any set of the form

$$\left\{ \boldsymbol{x} \;\middle|\; \boldsymbol{x} \in A^{\mathbb{Z}}, x_{t_1} = a_1, \ldots, x_{t_n} = a_n \right\} = [a_1 \ldots a_n]_{t_1, t_2, \ldots, t_n}$$

is called cylinder with base $a_1, a_2, \ldots, a_n$ at times $t_1 < t_2 < \cdots < t_n$. Consider the smallest σ-algebra $\mathcal{F}_A$ containing such cylinders. Any probability measure $\boldsymbol{\mu}$ on the Borel space $(A^{\mathbb{Z}}, \mathcal{F}_A)$ is uniquely determined by the values of $\boldsymbol{\mu}$ on the cylinders. The probability space $(A^{\mathbb{Z}}, \mathcal{F}_A, \boldsymbol{\mu})$ is called a *discrete time random process*.

Consider the shift transformation $T : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ defined by $T\boldsymbol{x} = \boldsymbol{y}$ where $y_n = x_{n-1}$ for all $n \in \mathbb{Z}$. If the probability measure $\boldsymbol{\mu}$ is invariant under T we say that $(A^{\mathbb{Z}}, \mathcal{F}_A, \boldsymbol{\mu})$ is a *stationary information source* and we denote it by $[A, \boldsymbol{\mu}]$. (From now onwards when there is no chance of confusion we may use the notation $\mu(E)$ to denote $\Pr(E, \boldsymbol{\mu})$.) For such a source

$$\mu([a_1 a_2 \ldots a_n]_{t_1, t_2, \ldots, t_n}) = \mu([a_1 a_2 \ldots a_n]_{t_1+1, t_2+1, \ldots, t_n+1}).$$

The information emitted by such a source during the time period t, $t+1$, …, $t+n-1$ is also the information emitted during the period 0, 1, …, $n-1$ and is given by

$$H_n(\boldsymbol{\mu}) = -\sum_C \mu(C) \log \mu(C).$$

where the summation is over all cylinders based on a_0, a_1, …, a_{n-1} at times 0, 1, 2, …, $n-1$, a_j varying in A. We call $\frac{H_n(\boldsymbol{\mu})}{n}$ as the *rate* at which information is generated by the source during $[0, n-1]$. Our next result shows that this rate converges to a limit as $n \to \infty$.

Theorem 8.3.2 *For any stationary source $[A, \boldsymbol{\mu}]$ the sequence $\frac{H_n(\boldsymbol{\mu})}{n}$ monotonically decreases to a limit $H(\boldsymbol{\mu})$.*

Proof For any a_0, a_1, …, $a_{n-1} \in A$ we write

$$[a_0 a_1 \ldots a_{n-1}] = [a_0 a_1 \ldots a_{n-1}]_{0,1,2,\ldots,n-1}.$$

Consider the output during $[0, n-1]$ as a random variable. Then we can express

$$H_{n+1}(\boldsymbol{\mu}) = -\mathbb{E}(\log \mu[x_{-n}, x_{-(n-1)}, \ldots, x_0])$$
$$H_n(\boldsymbol{\mu}) = -\mathbb{E}(\log \mu[x_{-n}, x_{-(n-1)}, \ldots, x_{-1}])$$

where the expectation is with respect to $\boldsymbol{\mu}$. We now show that the sequence $H_{n+1}(\boldsymbol{\mu}) - H_n(\boldsymbol{\mu})$ is monotonic decreasing. Let A, B and C be schemes determined by the cylinders $[x_0]$, $[x_{-n}, x_{-(n-1)}, \ldots, x_{-1}]$ and

$[x_{-(n+1)}]$ respectively. Then the joint scheme BC is given by the cylinder $[x_{-(n+1)}, x_{-n}, \ldots, x_{-1}]$. Then we have

$$H(A \mid B) = H_{n+1}(\boldsymbol{\mu}) - H_n(\boldsymbol{\mu}) \text{ and}$$
$$H(A \mid BC) = H_{n+2}(\boldsymbol{\mu}) - H_{n+1}(\boldsymbol{\mu}).$$

By using the fact $H(A \mid BC) \leq H(A \mid B)$ we get

$$H_{n+2}(\boldsymbol{\mu}) - H_{n+1}(\boldsymbol{\mu}) \leq H_{n+1}(\boldsymbol{\mu}) - H_n(\boldsymbol{\mu}).$$

Also $H_2(\boldsymbol{\mu}) \leq 2H_1(\boldsymbol{\mu})$. Thus the sequence $H_1(\boldsymbol{\mu}), H_2(\boldsymbol{\mu}) - H_1(\boldsymbol{\mu}), \ldots, H_n(\boldsymbol{\mu}) - H_{n-1}(\boldsymbol{\mu}), \ldots$ is monotonic decreasing. Since

$$\frac{H_n(\boldsymbol{\mu})}{n} = \frac{H_1(\boldsymbol{\mu}) + (H_2(\boldsymbol{\mu}) - H_1(\boldsymbol{\mu})) + \cdots + (H_n(\boldsymbol{\mu}) - H_{n-1}(\boldsymbol{\mu}))}{n},$$

it follows that $\frac{H_n(\boldsymbol{\mu})}{n}$ is monotonic decreasing. But $\frac{H_n(\boldsymbol{\mu})}{n}$ is bounded from below. Hence $\lim_{n\to\infty} \frac{H_n(\boldsymbol{\mu})}{n}$ exists.

□

Lecture 9

Quantum Information Theory

9.1 von Neumann Entropy

Following the exposition of quantum probability in chapter 1 we now replace the classical sample space $\Omega = \{1, 2, \ldots, n\}$ by a complex Hilbert space $\mathcal{H}$ of dimension n and the probability distribution $p_1, p_2 \ldots p_n$ on Ω by a state ρ, i.e., a nonnegative definite operator ρ of unit trace. Following von Neumann we define the entropy of a quantum state ρ by the expression $S(\rho) = -\operatorname{Tr}(\rho \log \rho)$ where the logarithm is with respect to the base 2 and it is understood that the function $x \log x$ is defined to be 0 whenever $x = 0$. We call $S(\rho)$ the *von Neumann entropy* of ρ. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of ρ (inclusive of multiplicity) we have

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i. \tag{9.1.1}$$

If ρ is the diagonal matrix $\operatorname{diag}(p_1, \ldots, p_n)$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$ a probability distribution, then $S(\rho) = H(\boldsymbol{p}) = -\sum_i p_i \log p_i$.

9.2 Properties of von Neumann Entropy

Property 1) $0 \leq S(\rho) \leq \log_2 d$, where d is the dimension of the Hilbert space $\mathcal{H}$. $S(\rho) = 0$ if and only if ρ is pure, i.e., $\rho = |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$ in $\mathcal{H}$. $S(\rho) = \log_2 d$ if and only if $\rho = d^{-1}I$.

Property 2) For any unitary operator U, $S(U\rho U^\dagger) = S(\rho)$.

Property 3) For any pure state $|\psi\rangle$, $S(|\psi\rangle\langle\psi|) = 0$.

Note that Property 3) is already contained in Property 1).

Suppose $\mathcal{H}_A \otimes \mathcal{H}_B$ describes the Hilbert space of a composite quantum system whose constituents are systems A and B with their states

coming from the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. For any operator X on $\mathcal{H}$ we define two operators X^A and X^B on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively by

$$\langle u|X^A|v\rangle = \sum_j \langle u \otimes f_j|X|v \otimes f_j\rangle \tag{9.2.1}$$

$$\langle u'|X^B|v'\rangle = \sum_i \langle e_i \otimes u'|X|e_i \otimes v'\rangle \tag{9.2.2}$$

for all $u, v \in \mathcal{H}_A$, $u', v' \in \mathcal{H}_B$, $\{e_i\}$, $\{f_j\}$ being orthonormal bases in $\mathcal{H}_A$, $\mathcal{H}_B$ respectively. Note that the right side of (9.2.1) and (9.2.2) are sesquilinear forms on $\mathcal{H}_A$ and $\mathcal{H}_B$, and therefore the operators X^A and X^B are uniquely defined. A simple algebra shows that X^A and X^B are independent of the choice of orthonormal bases in $\mathcal{H}_A$ and $\mathcal{H}_B$. We write $X^A = \mathrm{Tr}_B X$, $X^B = \mathrm{Tr}_A X$. Tr_A and Tr_B are called the operators of *relative trace* on the operator variable X. Note that $\mathrm{Tr}\, X^A = \mathrm{Tr}\, X^B = \mathrm{Tr}\, X$. If X is nonnegative definite so are X^A and X^B. In particular, for any state ρ of the composite system ρ^A and ρ^B are states on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. We call them the *marginal states* of ρ.

Let $|i_A\rangle$, $|j_B\rangle$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$ be orthonormal bases for $\mathcal{H}_A$, $\mathcal{H}_B$ respectively. Then $\{|i_A\rangle|j_B\rangle, 1 \le i \le m, 1 \le j \le n\}$ is an orthonormal basis for $\mathcal{H} = \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and hence any joint pure state $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{i,j} a_{ij}|i_A\rangle|j_B\rangle. \tag{9.2.3}$$

The $m \times n$ matrix $A = [a_{ij}]$ can be expressed as

$$[a_{ij}] = U \left[\begin{array}{c|c} \mathrm{D} & 0 \\ \hline 0 & 0 \end{array}\right] V$$

where U is a unitary matrix of order $m \times m$, V is a unitary matrix of order $n \times n$ and $D = \mathrm{diag}(s_1, s_2, \ldots, s_r)$, $s_1 \ge s_2 \ge \cdots \ge s_r \ge 0$, r being the rank of $[a_{ij}]$. It follows that $s_1, s_2, \ldots, s_r$ are positive eigenvalues of the matrices $\sqrt{A^\dagger A}$ and $\sqrt{AA^\dagger}$, called the *singular values* of A.

Define the vectors

$$|\alpha_A^i\rangle = \sum_{k=1}^m u_{ki}|k_A\rangle, \quad 1 \le i \le m$$

$$|\beta_B^j\rangle = \sum_{l=1}^n v_{jl}|l_B\rangle, \quad 1 \le j \le n$$

where $U = [u_{ki}]$, $V = [v_{jl}]$. Then (9.2.3) becomes

$$|\psi\rangle = \sum_{i=1}^{r} s_i |\alpha_A^i\rangle |\beta_B^i\rangle. \tag{9.2.4}$$

Here $|\alpha_A^1\rangle, |\alpha_A^2\rangle, \ldots, |\alpha_A^r\rangle$ and $|\beta_B^1\rangle, |\beta_B^2\rangle, \ldots, |\beta_B^r\rangle$ are orthonormal sets in $\mathcal{H}_A$ and $\mathcal{H}_B$ of same cardinality and $s_1, s_2, \ldots, s_r$ are the singular values of A. The decomposition of $|\psi\rangle$ in the form (9.2.4) is called the *Schmidt decomposition* of $|\psi\rangle$.

Property 4) Let $|\Omega\rangle\langle\Omega|$ be a pure state for AB and let ρ^A and ρ^B be its marginal states. Then $S(\rho^A) = S(\rho^B)$.

Proof By Schmidt decomposition we know that if $|\psi\rangle$ is a pure state for the composite system, AB; then there exist orthonormal states $|i_A\rangle$ for system A and orthonormal states $|i_B\rangle$ for system B such that $|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle$, where λ_i's are nonnegative real numbers satisfying $\sum_i \lambda_i^2 = 1$. So we can write $|\Omega\rangle\langle\Omega| = \sum \lambda_i \lambda_j |i_A\rangle\langle j_A| \otimes |i_B\rangle\langle j_B|$. Thus $\rho^A = \sum \lambda_i^2 \, |i_A\rangle\langle i_A|$ and $\rho^B = \sum \lambda_i^2 \, |i_B\rangle\langle i_B|$. Hence the eigenvalues of ρ^A and ρ^B are same. Therefore by (9.1.1) we have $S(\rho^A) = S(\rho^B)$.

□

Property 5) Let $\rho_1, \rho_2, \ldots, \rho_n$ be states with mutually orthogonal support and let $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ be a probability distribution. Then

$$S\Big(\sum_i p_i \rho_i\Big) = H(\boldsymbol{p}) + \sum_i p_i S(\rho_i),$$

where $H(\boldsymbol{p}) = -\sum p_i \log p_i$.

Proof Let λ_i^j and $|e_i^j\rangle$ be the eigenvalues and corresponding eigenvectors of ρ_i. Then $\sum p_i \rho_i$ has eigenvalues $p_i \lambda_i^j$ with respective eigenvectors $|e_i^j\rangle$. Thus,

$$\begin{aligned} S\Big(\sum_i p_i \rho_i\Big) &= -\sum_{i,j} p_i \lambda_i^j \log p_i \lambda_i^j \\ &= -\sum_i p_i \log p_i - \sum_i p_i \sum_j \lambda_i^j \log \lambda_i^j \\ &= H(\boldsymbol{p}) + \sum_i p_i S(\rho_i). \end{aligned}$$

□

An immediate consequence of Property 5) is the following.

Corollary 9.2.5 (Joint entropy theorem) *Let* $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ *be a probability distribution,* $\{|i\rangle, i = 1, 2, \ldots, n\}$ *an orthonormal set of states in* $\mathcal{H}_A$ *and* $\{\rho_i, i = 1, 2, \ldots, n\}$ *a set of density operators in* $\mathcal{H}_B$. *Then*

$$S\Big(\sum_i p_i |i\rangle\langle i| \otimes \rho_i\Big) = H(\boldsymbol{p}) + \sum_i p_i S(\rho_i).$$

Property 6) The following theorem shows that the correspondence $\rho \to S(\rho)$ is continuous. For any matrix ρ, by $|\rho|$ we mean the positive square root of $\rho^\dagger\rho$. For two positive semidefinite matrices ρ and σ, we define the trace distance as $\mathrm{tr}|\rho - \sigma|$. Now we are ready to state Fannes' inequality.

Theorem 9.2.6 (Fannes' inequality) *Suppose* ρ *and* σ *are density matrices such that the trace distance between them satisfies* $\mathrm{Tr}\,|\rho - \sigma| < \frac{1}{e}$. *Then* $|S(\rho) - S(\sigma)| \leq \mathrm{Tr}\,|\rho - \sigma| \log d + \eta(\mathrm{Tr}\,|\rho - \sigma|)$, *where* d *is the dimension of the Hilbert space, and* $\eta(x) = -x \log x$.

Proof Let $r_1 \geq \cdots \geq r_d$ and $s_1 \geq \cdots \geq s_d$ be the eigenvalues of ρ and σ respectively. By the spectral decomposition we can write $\rho - \sigma = Q - R$, where Q and R are positive operators with orthogonal support, so $\mathrm{Tr}\,|\rho - \sigma| = \mathrm{Tr}(R) + \mathrm{Tr}(Q)$. Defining $V = R + \rho = Q + \sigma$, we get $\mathrm{Tr}\,|\rho - \sigma| = \mathrm{Tr}(R) + \mathrm{Tr}(Q) = \mathrm{Tr}(2V) - \mathrm{Tr}(\rho) - \mathrm{Tr}(\sigma)$. Let $t_1 \geq \cdots \geq t_d$ be the eigenvalues of V. By the variational principle for the i^{th} eigenvalue it follows that $t_i \geq \max(r_i, s_i)$. Hence $2t_i \geq r_i + s_i + |r_i - s_i|$ and

$$\mathrm{Tr}\,|\rho - \sigma| \geq \sum_i |r_i - s_i|\,. \tag{9.2.7}$$

When $|r - s| \leq \frac{1}{e}$, from mean value theorem it follows that $|\eta(r) - \eta(s)| \leq \eta(|r - s|)$. Since $|r_i - s_i| \leq \frac{1}{e}$ for all i, it follows that

$$|S(\rho) - S(\sigma)| = \Big|\sum_i (\eta(r_i) - \eta(s_i))\Big| \leq \sum_i \eta(|r_i - s_i|).$$

Setting $\Delta = \sum_i |r_i - s_i|$ and observing that

$$\eta(|r_i - s_i|) = \Delta\eta(|r_i - s_i|/\Delta) - |r_i - s_i| \log(\Delta),$$

we obtain

$$|S(\rho) - S(\sigma)| \leq \Delta \sum \eta(|r_i - s_i| / \Delta) + \eta(\Delta) \leq \Delta \log d + \eta(\Delta).$$

By (9.2.7) and monotonicity of $\eta(\cdot)$ on the interval $[0, 1/e]$, we get

$$|S(\rho) - S(\sigma)| \leq \mathrm{Tr}\,|\rho - \sigma| \log d + \eta(\mathrm{Tr}\,|\rho - \sigma|).$$

□

Property 7) For any two quantum states ρ, σ we define the *relative entropy* $S(\rho||\sigma)$ of ρ with respect to σ by

$$S(\rho||\sigma) = \begin{cases} \mathrm{Tr}\,\rho \log \rho - \mathrm{Tr}\,\rho \log \sigma & \text{if supp } \rho \subset \text{supp } \sigma; \\ \infty & \text{otherwise.} \end{cases} \tag{9.2.8}$$

Theorem 9.2.9 (Klein's inequality) *$S(\rho||\sigma) \geq 0$, where equality holds if and only if $\rho = \sigma$.*

Proof Let the eigen decompositions of the states ρ and σ be given by $\rho = \sum_i p_i\, |i\rangle\langle i|$, $\sigma = \sum_j q_j\, |j\rangle\langle j|$. Then we have

$$\begin{aligned} S(\rho||\sigma) &= \sum p_i \log p_i - \sum \langle i|\rho \log \sigma|i\rangle \\ &= \sum p_i \log p_i - \sum_{i,j} p_i\, |\langle i \mid j\rangle|^2 \log q_j. \end{aligned}$$

We may assume $S(\rho||\sigma)$ to be finite. Since $-\log x$ is a convex function in the interval $[0, 1]$ and $\sum_j |\langle i \mid j\rangle|^2 = 1$, we have

$$-\sum_j |\langle i \mid j\rangle|^2 \log q_j \geq -\log \sum_j |\langle i \mid j\rangle|^2 q_j.$$

Putting $r_i = \sum_j |\langle i \mid j\rangle|^2 q_j$ and observing that $\sum_i r_i = 1$, we have

$$S(\rho||\sigma) \geq -\sum_i p_i \log \frac{r_i}{p_i} \geq 0.$$

□

Property 8) Let ρ^{AB} be a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ with marginal states ρ^A and ρ^B. We denote by $S(A)$, $S(B)$ and $S(AB)$ the von Neumann entropy of ρ^A, ρ^B and ρ^{AB} respectively. The quantum mutual information of the systems A and B is defined as $S(A:B) = S(A) + S(B) - S(AB)$.

Theorem 9.2.10 $S(A:B) \geq 0$.

Proof Observe that

$$\begin{aligned} S(A) &= -\operatorname{Tr} \rho^A \log \rho^A \\ &= -\operatorname{Tr} \rho^{AB} \log(\rho^A \otimes I_B). \end{aligned}$$

Substituting in the expression for $S(A:B)$ we get

$$\begin{aligned} S(A:B) &= -\operatorname{Tr} \rho^{AB}(\log \rho^A \otimes I_B + \log I_A \otimes \rho^B) + \operatorname{Tr} \rho^{AB} \log \rho^{AB} \\ &= S(\rho^{AB} || \rho^A \otimes \rho^B) \\ &\geq 0. \end{aligned}$$

□

Let ρ^{AB} be a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ with marginal states ρ^A and ρ^B. The *conditional entropy* of the state ρ^A given the state ρ^B is defined as $S(A \mid B) = S(AB) - S(B)$. Note that the state ρ^{AB} may be a pure state and the state ρ^B an impure state. So $S(A \mid B)$ can be less than zero.

Property 9) Let A be a quantum system with Hilbert space $\mathcal{H}_A$. By a projective measurement we mean a family of projection operators $P_1, P_2, \ldots, P_n$ in $\mathcal{H}_A$ satisfying $\sum_{i=1}^n P_i = I$. When such a measurement is made in a state ρ the outcome of the measurement is j with probability $\operatorname{Tr} \rho P_j$. According to collapse postulate 1.3 if the outcome is j the state collapses to $\frac{P_j \rho P_j}{\operatorname{Tr} \rho P_j}$. Thus the post measurement state, ignoring the individual outcome, is

$$\sum_j (\operatorname{Tr} \rho P_j) \frac{P_j \rho P_j}{\operatorname{Tr} \rho P_j} = \sum_j P_j \rho P_j.$$

Theorem 9.2.11 *Let ρ be the state of a quantum system and $P_1, \ldots, P_n$ be a projective measurement and let $\rho' = \sum_j P_j \rho P_j$. Then $S(\rho') \geq S(\rho)$ and equality holds if and only if $\rho' = \rho$.*

Proof

$$\begin{aligned}
0 &\leq S(\rho||\rho') \\
&= \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\rho\log\rho' \\
&= \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\left(\sum_i P_i\rho\log\rho'\right) \\
&= \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\sum_j P_j\rho(\log\rho')P_j \\
&= \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\sum_j P_j\rho P_j(\log\rho') \\
&= S(\rho') - S(\rho).
\end{aligned}$$

□

By a *generalized measurement* we mean a set of operators $L_1, \ldots, L_n$ satisfying $\sum_{i=1}^n L_i^\dagger L_i = I$. If ρ is a state in which such a generalized measurement is made, the probability of the outcome i is $\operatorname{Tr}\rho L_i^\dagger L_i$ and the post measurement state is $\frac{L_i\rho L_i^\dagger}{\operatorname{Tr}\rho L_i^\dagger L_i}$. Thus the post measurement state, ignoring the individual outcome, is

$$\sum(\operatorname{Tr}\rho L_i^\dagger L_i)\frac{L_i\rho L_i^\dagger}{\operatorname{Tr}\rho L_i^\dagger L_i} = \sum_i L_i\rho L_i^\dagger.$$

Remark 9.2.12 A generalized measurement may decrease the entropy.

Example 9.2.13 Let $L_1 = |0\rangle\langle 0|$ and $L_2 = |0\rangle\langle 1|$. Note that $L_1^\dagger L_1 + L_2^\dagger L_2 = I$. Let $\rho = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$. Then

$$S(\rho) = -p\log p - (1-p)\log(1-p).$$

Let ρ be measured using the measurement operators L_1 and L_2. The resulting state is $\rho' = L_1\rho L_1^\dagger + L_2\rho L_2^\dagger = |0\rangle\langle 0|$. This implies $S(\rho') = 0$.

Property 10)

Theorem 9.2.14 *Let ρ^{AB} be a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ with marginal states ρ^A and ρ^B. Then the following inequalities hold.*

1) $S(AB) \leq S(A) + S(B)$,
2) $S(AB) \geq |S(A) - S(B)|$.

The first inequality is known as the sub-additivity inequality for the von Neumann entropy. The second is known as the triangle inequality or the Araki-Lieb inequality.

Proof The first inequality follows immediately from Klein's inequality (Theorem 9.2.9), $S(\rho) \leq -\operatorname{Tr}\rho\log\sigma$. Let $\rho = \rho^{AB}$ and $\sigma = \rho^A \otimes \rho^B$. Then

$$\begin{aligned} -\operatorname{Tr}(\rho\log\sigma) &= -\operatorname{Tr}(\rho^{AB}(\log\rho^A + \log\rho^B)) \\ &= -\operatorname{Tr}(\rho^A\log\rho^A) - \operatorname{Tr}(\rho^B\log\rho^B) \\ &= S(A) + S(B). \end{aligned}$$

Therefore we have $S(AB) \leq S(A) + S(B)$. From Klein's theorem it follows that equality holds if and only if $\rho^{AB} = \rho^A \otimes \rho^B$.

To prove the triangle inequality, we introduce a reference system R such that ρ^{ABR} is a pure state in $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_R$. Then by sub-additivity we have $S(R) + S(A) \geq S(AR)$. Since ρ^{ABR} is a pure state we have $S(AR) = S(B)$ and $S(R) = S(AB)$. Substituting we get $S(AB) \geq S(B) - S(A)$. By symmetry we get the second inequality.

□

Exercise 9.2.15 Let $\rho^{AB} = \sum_i \lambda_i |i\rangle\langle i|$ be the spectral decomposition for ρ^{AB}. Then, show that $S(AB) = S(B) - S(A)$ if and only if the operators $\rho_i^A = \operatorname{Tr}_B(|i\rangle\langle i|)$ have a common eigen basis, and the operators $\rho_i^B = \operatorname{Tr}_A(|i\rangle\langle i|)$ have orthogonal support.

Property 11) $S(\rho)$ is concave in ρ.

Theorem 9.2.16 *Let $\rho_1, \rho_2, \ldots, \rho_n$ be states and let $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ be a probability distribution. Then*

$$S\left(\sum_i p_i\rho_i\right) \geq \sum_i p_i S(\rho_i).$$

Proof Let ρ_i's be the states in $\mathcal{H}_A$. Consider an auxiliary Hilbert space $\mathcal{H}_B$, whose state space has an orthonormal basis $|i\rangle$ corresponding to the index i of the density operators ρ_i. Let a joint state on $\mathcal{H}_A \otimes \mathcal{H}_B$ be defined by

$$\rho^{AB} = \sum_i p_i\rho_i \otimes |i\rangle\langle i|.$$

Note that $S(AB) = H(\boldsymbol{p}) + \sum p_i S(\rho_i)$, by the joint entropy theorem (Corollary 9.2.5).

$$\rho^A = \sum p_i \rho_i \Rightarrow S(\rho^A) = S\Big(\sum p_i \rho_i\Big).$$
$$\rho^B = \sum p_i |i\rangle\langle i| \Rightarrow S(\rho^B) = H(\boldsymbol{p}).$$

By subadditivity we have, $S(\rho^A) + S(\rho^B) \geq S(\rho^{AB})$. Substituting we get $S(\sum p_i \rho_i) + H(\boldsymbol{p}) \geq H(\boldsymbol{p}) + \sum p_i S(\rho_i)$.

□

Property 12)

Theorem 9.2.17 $\sum p_i S(\rho_i) \leq S(\sum p_i \rho_i) \leq H(\boldsymbol{p}) + \sum p_i S(\rho_i)$.

Proof First let us consider the case when $\rho_i = |\psi_i\rangle\langle\psi_i|$ for all i. Let ρ_i's be the states in $\mathcal{H}_A$ and let $\mathcal{H}_B$ be an auxiliary Hilbert space with an orthonormal basis $|i\rangle$ corresponding to the index i of the probabilities p_i. Let $\rho^{AB} =| AB\rangle\langle AB |$ where $|AB\rangle = \sum \sqrt{p_i}|\psi_i\rangle|i\rangle$. In other words $\rho^{AB} = \sum_{i,j} \sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j| \otimes |i\rangle\langle j|$. Since ρ^{AB} is a pure state we have $S(A) = S(B) = S(\sum_i p_i |\psi_i\rangle\langle\psi_i|)$. After performing measurement on the state ρ^B in the $|i\rangle$ basis, the state of the system will be $\rho^{B'} = \sum_i p_i |i\rangle\langle i|$. But, projective measurements never decrease entropy and using the fact $S(\rho_i) = 0$ we get $S(A) \leq H(\boldsymbol{p}) + \sum_i p_i S(\rho_i)$. Note that the equality holds if and only if $\rho^B = \rho^{B'}$ and this occurs if and only if $|\psi_i\rangle's$ are orthogonal. Now we can prove the mixed state case.

Let $\rho_i = \sum_i p_{ij} |e^i_j\rangle\langle e^i_j|$ be an orthonormal decomposition for the state ρ_i. Let $\rho = \sum_{i,j} p_i p_{ij} |e^i_j\rangle\langle e^i_j|$. Applying the result for the pure state case and observing that $\sum_j p_{ij} = 1$ for all i, we get

$$\begin{aligned} S(\rho) &\leq -\sum_{i,j} p_i p_{ij} \log(p_i p_{ij}) \\ &= \sum_i p_i \log p_i - \sum_i p_i \sum_j p_{ij} \log p_{ij} \\ &= H(\boldsymbol{p}) + \sum_i p_i S(\rho_i). \end{aligned}$$

□

The sub-additivity and the triangle inequality for two quantum systems can be extended to three systems. This gives rise to a very important and useful result, known as the strong sub-additivity. The proof given here depends on a deep mathematical result known as Lieb's theorem.

Let A, B be bounded operator variables on a Hilbert space $\mathcal{H}$. Suppose the pair (A, B) varies in a convex set $\mathcal{C}$. A map $f : \mathcal{C} \to \mathbb{R}$ is said to be *jointly convex* if

$$f(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2) \leq \lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2).$$

for all $0 \leq \lambda \leq 1$, $(A_i, B_i) \in \mathcal{C}$, $i = 1, 2$.

Now we are ready to state the next property.

Property 13)

Theorem 9.2.18 *Relative entropy is jointly convex in its arguments.*

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two finite dimensional Hilbert spaces. Let α be a map from $\mathcal{B}(\mathcal{H}_1)$ to $\mathcal{B}(\mathcal{H}_2)$ which satisfies $\alpha(X^\dagger X) \geq \alpha(X)^\dagger \alpha(X)$. In our case α will be a star homomorphism. Let T_i, S_i, $i \in \{1, 2\}$ be positive operators in $\mathcal{H}_i$, $i = 1, 2$. The index i corresponds to the Hilbert space $\mathcal{H}_i$. To prove Theorem 9.2.18 we need the following lemma. This is also known as Lieb's inequality.

Lemma 9.2.19 *If* $\operatorname{Tr} XT_1 \geq \operatorname{Tr} \alpha(X)T_2$ *and* $\operatorname{Tr} XS_1 \geq \operatorname{Tr} \alpha(X)S_2$ *and* $T_i, i = 1, 2$ *are invertible then*

$$\operatorname{Tr} \alpha(X^\dagger) S_2^t \alpha(X) T_2^{1-t} \leq \operatorname{Tr} X^\dagger S_1^t X T_1^{1-t}. \tag{9.2.20}$$

Observe that (9.2.20) is true when the parameter t is equal to 1 or 0. We need to show that the conclusion of (9.2.20) holds even when t is a real number in the range $(0, 1)$. So Lieb's inequality is an interpolation inequality. To prove Lieb's inequality we need the following results.

Lemma 9.2.21 *The following equation is true.*

$$x^t = \frac{1}{\beta(t, 1-t)} \int_0^\infty \left[\lambda^{t-1} - \lambda^t (\lambda + x)^{-1}\right] d\lambda. \tag{9.2.22}$$

Proof We first perform the substitution $1 + \frac{\lambda}{x} = \frac{1}{u}$. Then,

$$\begin{aligned}
&\frac{1}{\beta(t,1-t)} \int_0^\infty \left[\lambda^{t-1} - \lambda^t(\lambda+x)^{-1}\right] d\lambda \\
&= \frac{1}{\beta(t,1-t)} \int_1^0 \left[x^{t-1} \left(\frac{1-u}{u}\right)^{t-1} - x^t \left(\frac{1-u}{u}\right)^t \left(\frac{u}{x}\right) \right] \left(-\frac{x}{u^2}\right) du \\
&= \frac{x^t}{\beta(t,1-t)} \int_0^1 (1-u)^{t-1} u^{-t} du \\
&= x^t.
\end{aligned}$$

□

Lemma 9.2.23 *Let $0 < t < 1$ and let A, B be two positive operators such that $A \leq B$. Then $A^t \leq B^t$.*

Proof

$$\begin{aligned}
A \leq B &\Rightarrow (\lambda + A)^{-1} \geq (\lambda + B)^{-1} \\
&\Rightarrow \lambda^t(\lambda + A)^{-1} \geq \lambda^t(\lambda + B)^{-1} \\
&\Rightarrow \lambda^{t-1} - \lambda^t(\lambda + A)^{-1} \leq \lambda^{t-1} - \lambda^t(\lambda + B)^{-1}.
\end{aligned}$$

Thus by spectral theorem and Lemma 9.2.21 we have $A^t \leq B^t$.

□

Lemma 9.2.24 *Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ be a strictly positive definite matrix where A_{11} and A_{22} are square matrices. Then A_{11} and A_{22} are also strictly positive definite and*

$$\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} \right)_{11} > A_{11}^{-1}.$$

Proof Note that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} =$$
$$\begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \\ -(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix}.$$

Therefore $(A^{-1})_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}$. Since $A_{12}A_{22}^{-1}A_{21}$ is a positive operator we have $(A^{-1})_{11} \geq A_{11}^{-1}$.

□

Lemma 9.2.25 *Let X be a positive operator in a finite dimensional Hilbert space $\mathcal{H}_0$ and let V be a contraction map. Then $(V^\dagger XV)^t \geq V^\dagger X^t V$.*

Proof Observe that the lemma is true when V is unitary. Let

$$U = \begin{bmatrix} V & \sqrt{1-VV^\dagger} \\ -\sqrt{1-V^\dagger V} & V^\dagger \end{bmatrix}.$$

Note that, since V is a contraction map, $\sqrt{1-VV^\dagger}$ and $\sqrt{1-V^\dagger V}$ are well defined and U is unitary.

Let P be the map $P : \mathcal{H}_0 \oplus \mathcal{H}_0 \to \mathcal{H}_0$ which is projection on the first co-ordinate. Then $V = PUP|_{\mathcal{H}_0}$. By Lemma 9.2.24 we have

$$\begin{aligned}
(\lambda I_{\mathcal{H}_0} + V^\dagger XV)^{-1} &= (\lambda I_{\mathcal{H}_0} + PU^\dagger PXPUP\,|_{\mathcal{H}_0})^{-1} \\
&\leq P(\lambda I + U^\dagger PXPU)^{-1}P\,|_{\mathcal{H}_0} \\
&= PU^\dagger(\lambda^{-1}P^\perp + P(\lambda+X)^{-1}P)UP\,|_{\mathcal{H}_0} \\
&= \lambda^{-1}PU^\dagger(I-P)UP\,|_{\mathcal{H}_0} + V^\dagger(\lambda+X)^{-1}V \\
&= \lambda^{-1}(I - V^\dagger V) + V^\dagger(\lambda+X)^{-1}V.
\end{aligned}$$

This implies

$$\begin{aligned}
&\frac{1}{\beta(1,1-t)}\int_0^\infty \lambda^{t-1} - \lambda^t(\lambda I + V^\dagger XV)^{-1}d\lambda \\
&\quad\geq \frac{1}{\beta(1,1-t)}\int_0^\infty \lambda^{t-1} - \lambda^t(\lambda^{-1}(I-V^\dagger V) + V^\dagger(\lambda+X)^{-1}V)d\lambda.
\end{aligned}$$

By applying Lemma 9.2.21 we get $(V^\dagger XV)^t \geq V^\dagger X^t V$. This completes the proof.

□

Remark 9.2.26 Lemma 9.2.25 holds even when the contraction V is from one Hilbert space $\mathcal{H}_1$ to another Hilbert space $\mathcal{H}_2$ and X is a positive operator in $\mathcal{H}_2$. In this case the operator U of the proof is from $\mathcal{H}_1 \oplus \mathcal{H}_2$ to $\mathcal{H}_2 \oplus \mathcal{H}_1$.

We look upon $\mathcal{B}(\mathcal{H}_1)$ and $\mathcal{B}(\mathcal{H}_2)$ as Hilbert spaces with the scalar product between two operators defined as $\langle X, Y\rangle = \operatorname{Tr} X^\dagger Y$. Define $V : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ by $V : XT_1^{\frac{1}{2}} = \alpha(X)T_2^{\frac{1}{2}}$.

Lemma 9.2.27 *V is a contraction map.*

Proof

$$\begin{aligned}||\alpha(X)T_2^{\frac{1}{2}}||^2 &= \operatorname{Tr} T_2^{\frac{1}{2}}\alpha(X)^\dagger\alpha(X)T_2^{\frac{1}{2}} \\ &\leq \operatorname{Tr}\alpha(X^\dagger X)T_2 \leq \operatorname{Tr} X^\dagger X T_1 \\ &= \operatorname{Tr} T_1^{\frac{1}{2}}X^\dagger X T_1^{\frac{1}{2}} \\ &= ||XT_1^{\frac{1}{2}}||^2.\end{aligned}$$

Hence the assertion is true.

□

Assume that T_1 and T_2 are invertible and put $\Delta_t X = S_1^t X T_1^{-t}$ and $D_t Y = S_2^t Y T_2^{-t}$. Note that $\Delta_t\Delta_s = \Delta_{t+s}$ and $D_t D_s = D_{s+t}$ for $s, t \geq 0$. Furthermore,

$$\begin{aligned}\langle XT_1^{\frac{1}{2}} \mid \Delta_t \mid XT_1^{\frac{1}{2}}\rangle &= \operatorname{Tr} T_1^{\frac{1}{2}}X^\dagger S_1^t X T_1^{\frac{1}{2}-t} \\ &= \operatorname{Tr}(X^\dagger S_1^t X)T_1^{1-t} \\ &\geq 0,\end{aligned}$$

and similarly $\langle YT_2^{\frac{1}{2}} \mid D_t \mid YT_2^{\frac{1}{2}}\rangle \geq 0$.

Hence Δ_t and D_t are positive operator semigroups and in particular $\Delta_t = \Delta_1^t$ and $D_t = D_1^t$.

Lemma 9.2.28 $\langle XT_1^{\frac{1}{2}} \mid \Delta_1 \mid XT_1^{\frac{1}{2}}\rangle \geq \langle XT_1^{\frac{1}{2}} \mid V^\dagger D_1 V \mid XT_1^{\frac{1}{2}}\rangle$.

Proof

$$\begin{aligned}\langle XT_1^{\frac{1}{2}} \mid \Delta_1 \mid XT_1^{\frac{1}{2}}\rangle &= \operatorname{Tr} T_1^{\frac{1}{2}}X^\dagger S_1 X T_1^{-\frac{1}{2}} \\ &= \operatorname{Tr} X^\dagger S_1 X \\ &= \operatorname{Tr} X X^\dagger S_1 \\ &\geq \operatorname{Tr}\alpha(XX^\dagger)S_2 \\ &\geq \operatorname{Tr}\alpha(X)\alpha(X^\dagger)S_2 \\ &= \operatorname{Tr} T_2^{\frac{1}{2}}\alpha(X)^\dagger S_2\alpha(X)T_2^{-\frac{1}{2}} \\ &= \langle XT_1^{\frac{1}{2}} \mid V^\dagger D_1 V \mid XT_1^{\frac{1}{2}}\rangle.\end{aligned}$$

□

Proof of Lemma 9.2.19 From Lemmas 9.2.28, 9.2.23 and 9.2.25 it follows that

$$\begin{aligned}
&\Delta_1 \geq V^\dagger D_1 V \\
\Rightarrow\ &\Delta_t \geq (V^\dagger D_1 V)^t \\
&\quad \geq V^\dagger D_1^t V \quad \text{(true since } V \text{ is a contraction map)} \\
&\quad = V^\dagger D_t V.
\end{aligned}$$

By expanding one can verify that the inequality

$$\langle X T_1^{\frac{1}{2}} \mid \Delta_t \mid X T_1^{\frac{1}{2}} \rangle \geq \langle \alpha(X) T_1^{\frac{1}{2}} \mid D_t \mid \alpha(X) T_1^{\frac{1}{2}} \rangle$$

is same as (9.2.20).

□

Proof of Theorem 9.2.18 Let $\mathcal{H}_2 = \mathcal{H} \otimes \mathcal{H}$ and $\alpha(X) = \left[\begin{smallmatrix} X & 0 \\ 0 & X \end{smallmatrix}\right]$.

For $0 < \lambda < 1$ define S_1, T_1, S_2 and T_2 as follows: $S_1 = \lambda\rho_1 + (1-\lambda)\rho_2$, $T_1 = \lambda\sigma_1 + (1-\lambda)\sigma_2$,

$$S_2 = \begin{bmatrix} \lambda\rho_1 & 0 \\ 0 & (1-\lambda)\rho_2 \end{bmatrix} \text{ and } T_2 = \begin{bmatrix} \lambda\sigma_1 & 0 \\ 0 & (1-\lambda)\sigma_2 \end{bmatrix},$$

where σ_1 and σ_2 are invertible. Then

$$\begin{aligned}
\operatorname{Tr} \alpha(X) S_2 &= \lambda \operatorname{Tr} \rho_1 X + (1-\lambda) \operatorname{Tr} \rho_2 X \\
&= \operatorname{Tr} S_1 X \text{ and} \\
\operatorname{Tr} \alpha(X) T_2 &= \lambda \operatorname{Tr} \sigma_1 X + (1-\lambda) \operatorname{Tr} \sigma_2 X \\
&= \operatorname{Tr} T_1 X.
\end{aligned}$$

Applying (9.2.20) with $X = I$ we get,

$$\begin{aligned}
\operatorname{Tr} S_2^t T_2^{1-t} &\leq \operatorname{Tr} S_1^t T_1^{1-t} \\
\lim_{t\to 1} \frac{1 - \operatorname{Tr} S_2^t T_2^{1-t}}{1-t} &\geq \lim_{t\to 1} \frac{1 - \operatorname{Tr} S_1^t T_1^{1-t}}{1-t} \\
\frac{d}{dt} \operatorname{Tr} S_2^t T_2^{1-t} \mid_{t=1} &\geq \frac{d}{dt} \operatorname{Tr} S_1^t T_1^{1-t} \mid_{t=1} \\
\operatorname{Tr} S_2 \log S_2 - \operatorname{Tr} S_2 \log T_2 &\geq \operatorname{Tr} S_1 \log S_1 - \operatorname{Tr} S_1 \log T_1.
\end{aligned}$$

That is,

$$\begin{aligned}
\operatorname{Tr} \lambda\rho_1 \log \lambda\rho_1 + (1-\lambda)\rho_2 \log(1-\lambda)\rho_2 - \lambda\rho_1 \log \lambda\sigma_1 - (1-\lambda)\rho_2 \log(1-\lambda)\sigma_2 \\
\geq S(\lambda\rho_1 + (1-\lambda)\rho_2 || \lambda\sigma_1 + (1-\lambda)\sigma_2).
\end{aligned}$$

Thus $\lambda S(\rho_1||\sigma_1)+(1-\lambda)S(\rho_2||\sigma_2) \geq S(\lambda\rho_1+(1-\lambda)\rho_2||\lambda\sigma_1+(1-\lambda)\sigma_2)$.

□

Property 14) Let ρ^{AB} be a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ with marginal states ρ^A and ρ^B. Then the conditional entropy is concave in the state ρ^{AB} of $\mathcal{H}_A \otimes \mathcal{H}_B$.

Proof Let d be the dimension of $\mathcal{H}_A$. Then

$$\begin{aligned} S\left(\rho^{AB}||\frac{I}{d}\otimes\rho^B\right) &= -S(AB) - \mathrm{Tr}\left(\rho^{AB}\log\left(\frac{I}{d}\otimes\rho^B\right)\right) \\ &= -S(AB) - \mathrm{Tr}(\rho^B\log\rho^B) + \log d \\ &= -S(A\mid B) + \log d. \end{aligned}$$

Therefore concavity of $S(A \mid B)$ follows from convexity of the relative entropy.

□

Property 15)

Theorem 9.2.29 (Strong subadditivity) *For any three quantum systems, A, B, C, the following inequalities hold.*

1) $S(A) + S(B) \leq S(AC) + S(BC)$.
2) $S(ABC) + S(B) \leq S(AB) + S(BC)$.

Proof To prove 1), we define a function $T(\rho^{ABC})$ as follows:

$$T(\rho^{ABC}) = S(A) + S(B) - S(AC) - S(BC) = -S(C \mid A) - S(C \mid B).$$

Let $\rho^{ABC} = \sum_i p_i \,|i\rangle\langle i|$ be a spectral decomposition of ρ^{ABC}. From the concavity of the conditional entropy we see that $T(\rho^{ABC})$ is a convex function of ρ^{ABC}. From the convexity of T we have

$$T(\rho^{ABC}) \leq \sum_i p_i T(|i\rangle\langle i|).$$

But $T(|i\rangle\langle i|) = 0$, as for a pure state $S(AC) = S(B)$ and $S(BC) = S(A)$. This implies $T(\rho^{ABC}) \leq 0$. Thus

$$S(A) + S(B) - S(AC) - S(BC) \leq 0.$$

To prove 2) we introduce an auxiliary system R purifying the system ABC so that the joint state ρ^{ABCR} is pure. Then using 1) we get

$$S(R) + S(B) \leq S(RC) + S(BC).$$

Since $ABCR$ is a pure state, we have, $S(R) = S(ABC)$ and $S(RC) = S(AB)$. Substituting we get

$$S(ABC) + S(B) \leq S(AB) + S(BC).$$

□

Property 16) $S(A : BC) \geq S(A : B)$

Proof Using the second part of Property 15) we have

$$\begin{aligned} S(A : BC) - S(A : B) &= S(A) + S(BC) - S(ABC) - \\ &\quad [S(A) + S(B) - S(AB)] \\ &= S(BC) + S(AB) - S(ABC) - S(B) \\ &\geq 0. \end{aligned}$$

□

Let $\mathcal{H}$ be the Hilbert space of a finite level quantum system. Recall that by a generalized measurement we mean a finite collection of operators $\{L_1, L_2, \ldots, L_k\}$ satisfying the relation $\sum_i L_i^\dagger L_i = I$. The set $\{1, 2, \ldots, k\}$ is the collection of the possible outcomes of the measurement and if the state of the system at the time of measurement is ρ then the probability p_i of the outcome i is given by $p_i = \operatorname{Tr} L_i \rho L_i^\dagger = \operatorname{Tr} \rho L_i L_i^\dagger$. If the outcome of the measurement is i, then the state of the system collapses to

$$\rho_i = \frac{L_i \rho L_i^\dagger}{p_i}.$$

Thus the post measurement state is expected to be $\sum_i p_i \rho_i = \sum_i L_i \rho L_i^\dagger$. The map $\mathcal{E}$ defined by

$$\mathcal{E}(\rho) = \sum_i L_i \rho L_i^\dagger \tag{9.2.30}$$

on the set of states is called a *quantum operation*.

If we choose and fix an orthonormal basis in $\mathcal{H}$ and express the operators L_i as matrices in this basis the condition that $\sum_i L_i^\dagger L_i = I$ can be interpreted as the property that the columns of the matrix

$$\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_k \end{bmatrix}$$

constitute an orthonormal set of vectors. The length of the column vector is kd where d is the dimension of the Hilbert space $\mathcal{H}$. Extend this set of orthonormal vectors into an orthonormal basis for $\mathcal{H} \otimes \mathbb{C}^k$ and construct a unitary matrix of order $kd \times kd$ of the form

$$U = \begin{bmatrix} L_1 & \cdots \\ L_2 & \cdots \\ \vdots & \vdots \\ L_k & \cdots \end{bmatrix}.$$

We can view this as a block matrix where each block is a $d \times d$ matrix. Define

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

so that for any state ρ in $\mathcal{H}$ we have

$$M = \rho \otimes |0\rangle\langle 0| = \begin{bmatrix} \rho & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

as states in $\mathcal{H} \otimes \mathbb{C}^k$. Then

$$UMU^\dagger = \begin{bmatrix} L_1\rho L_1^\dagger & L_1\rho L_2^\dagger & \cdots & L_1\rho L_k^\dagger \\ L_2\rho L_1^\dagger & L_2\rho L_2^\dagger & \cdots & L_2\rho L_k^\dagger \\ \vdots & \vdots & \vdots & \vdots \\ L_k\rho L_1^\dagger & L_k\rho L_2^\dagger & \cdots & L_k\rho L_k^\dagger \end{bmatrix}.$$

Thus we have $\mathrm{Tr}_{\mathbb{C}^k} U(\rho \otimes |0\rangle\langle 0|)U^\dagger = \sum_{i=1}^k L_i \rho L_i^\dagger = \mathcal{E}(\rho)$, where $\mathcal{E}(\rho)$ is defined as in (9.2.30). We summarize our discussion in the form of a lemma.

Lemma 9.2.31 *Let $\mathcal{E}$ be a quantum operation on the states of a quantum system with Hilbert space $\mathcal{H}$ determined by a generalized measurement $\{L_i, 1 \leq i \leq k\}$. Then there exists a pure state $|0\rangle$ of an auxiliary system with a Hilbert space $\mathcal{K}$ of dimension k and a unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ satisfying the property $\mathcal{E}(\rho) = \mathrm{Tr}_{\mathcal{K}} U(\rho \otimes |0\rangle\langle 0|)U^\dagger$ for every state ρ in $\mathcal{H}$.*

Property 17) Let AB be a composite system with Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and let $\mathcal{E}$ be a quantum operation on B determined by the generalized measurement $\{L_i,\ 1 \leq i \leq k\}$ in $\mathcal{H}_B$. Then $\mathrm{id} \otimes \mathcal{E}$ is a quantum operation on AB determined by the generalized measurement $\{I_A \otimes L_i, 1 \leq i \leq k\}$. If ρ^{AB} is any state in $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and $\rho^{A'B'} = \mathrm{id} \otimes \mathcal{E}(\rho^{AB})$, then, $S(A' : B') \leq S(A : B)$.

Proof Following Lemma 9.2.31, we construct an auxiliary system C with Hilbert space $\mathcal{H}_C$, a pure state $|0\rangle$ in $\mathcal{H}_C$ and a unitary operator U on $\mathcal{H}_B \otimes \mathcal{H}_C$ so that

$$\mathcal{E}(\rho^B) = \sum_i L_i \rho^B L_i^\dagger = \mathrm{Tr}_C\, U(\rho^B \otimes |0\rangle\langle 0|)U^\dagger.$$

Define $\tilde{U} = I_A \otimes U$. Let $\rho^{ABC} = \rho \otimes |0\rangle\langle 0|$ and $\rho^{A'B'C'} = \tilde{U}\rho^{ABC}\tilde{U}^\dagger$. Then for the marginal states we have $\rho^{A'} = \rho^A, \quad \rho^{B'C'} = U\rho^{BC}U^\dagger$ and therefore $S(A') = S(A), S(B'C') = S(BC)$. Thus using Property 16), we get

$$\begin{aligned} S(A : B) &= S(A) + S(B) - S(AB) \\ &= S(A) + S(BC) - S(ABC) \\ &= S(A') + S(B'C') - S(A'B'C') \\ &= S(A' : B'C') \\ &\geq S(A' : B'). \end{aligned}$$

□

Property 18) Holevo Bound

Consider an information source in which messages x from a finite set X come with probability $p(x)$. We denote this probability distribution by $\boldsymbol{p}$. The information obtained from such a source is given by

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x).$$

Now suppose the message x is encoded as a quantum state ρ_x in a Hilbert space $\mathcal{H}$. In order to decode the message make a generalized measurement $\{L_y, Y \in Y\}$ where $\sum_{y \in Y} L_y^\dagger L_y = I$. Given that the message x

came from the source, or equivalently, the state of the quantum system is the encoded state ρ_x the probability for the measurement value y is given by $\Pr(y \mid x) = \operatorname{Tr} L_y \rho_x L_y^\dagger$. Thus the joint probability $\Pr(x, y)$, that x is the message and y is the measurement outcome, is given by

$$\Pr(x, y) = p(x) \Pr(y \mid x) = p(x) \operatorname{Tr} \rho_x L_y^\dagger L_y.$$

Thus we obtain a classical joint system XY described by this probability distribution in the space $X \times Y$. The information gained from the generalized measurement about the source X is measured by the quantity $H(X) + H(Y) - H(XY)$ (see [9]). Our next result puts an upper bound on the information thus gained.

Theorem 9.2.32 (Holevo, 1973)

$$H(X) + H(Y) - H(XY) \leq S(\sum_x p(x)\rho_x) - \sum_x p(x) S(\rho_x).$$

Proof Let $\{|x\rangle, x \in X\}$, $\{|y\rangle, y \in Y\}$ be orthonormal bases in Hilbert spaces $\mathcal{H}_X$, $\mathcal{H}_Y$ of dimension $\#X$, $\#Y$ respectively. Denote by $\mathcal{H}_Z$ the Hilbert space of the encoded states $\{\rho_x,\ x \in X\}$. Consider the Hilbert space $\mathcal{H}_{XZY} = \mathcal{H}_X \otimes \mathcal{H}_Z \otimes \mathcal{H}_Y$. Choose and fix an element 0 in Y and define the joint state

$$\rho^{XZY} = \sum_x p(x)\ |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|\ .$$

In the Hilbert space $\mathcal{H}_{ZY}$ consider the generalized measurement determined by $\{\sqrt{E_y} \otimes U_y,\ y \in Y\}$ where $E_y = L_y^\dagger L_y$ and U_y is any unitary operator in $\mathcal{H}_Y$ satisfying $U_y|0\rangle = |y\rangle$. Such a measurement gives an operation $\mathcal{E}$ on the states of the system ZY and the operation $\mathrm{id} \otimes \mathcal{E}$ satisfies

$$(\mathrm{id} \otimes \mathcal{E})(\rho^{XZY}) = \sum_{x \in X,\ y \in Y} p(x)\ |x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y| = \rho^{X'Z'Y'},$$

say. By Property 17) we have $S(X : Z) = S(X : ZY) \geq S(X' : Z'Y')$. By Property 16)

$$S(X : Z) \geq S(X' : Y'). \tag{9.2.33}$$

Since $\rho^{XZ} = \sum p(x)\,|x\rangle\langle x| \otimes \rho_x$ we have from the joint entropy theorem $S(XZ) = H(\boldsymbol{p}) + \sum p(x)S(\rho_x)$. Furthermore

$$\begin{aligned} \rho^X &= \sum p(x)\,|x\rangle\langle x|, \quad S(X) = H(\boldsymbol{p}) = H(X) \\ \rho^Z &= \sum p(x)\rho_x, \quad S(Z) = S(\rho^Z) \\ S(X:Z) &= S(\textstyle\sum p(x)\rho_x) - \sum p(x)S(\rho_x). \end{aligned} \tag{9.2.34}$$

On the other hand

$$\begin{aligned} \rho^{X'Z'Y'} &= \sum_{x,y} p(x)\,|x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y| \\ \rho^{X'} &= \sum_x p(x)\,|x\rangle\langle x| \\ \rho^{Y'} &= \sum_{x,y} p(x)\,\mathrm{Tr}\,\rho_x E_y\,|y\rangle\langle y| \\ \rho^{X'Y'} &= \sum_{x,y} p(x)\,\mathrm{Tr}\,\rho_x E_y\,|x\rangle\langle x| \otimes |y\rangle\langle y|\,. \end{aligned}$$

Thus,

$$S(X':Y') = H(X) + H(Y) - H(XY). \tag{9.2.35}$$

Combining (9.2.33), (9.2.34) and (9.2.35) we get the required result. □

Property 19) Schumacher's theorem

Let $\boldsymbol{p}$ be a probability distribution on a finite set X. For $\epsilon > 0$ define

$$\nu(\boldsymbol{p}, \epsilon) = \min\{\#E \mid E \subset X,\ \Pr(E;\boldsymbol{p}) \geq 1 - \epsilon\}.$$

It is quite possible that $\#X$ is large in comparison with $\nu(\boldsymbol{p}, \epsilon)$. In other words, by omitting a set of probability at most ϵ we may have most of the statistical information packed in a set E of size much smaller than $\#X$. In the context of information theory it is natural to consider the ratio $\frac{\log_2 \nu(\boldsymbol{p},\epsilon)}{\log_2 \#X}$ as the information content of $\boldsymbol{p}$ upto a negligible set of probability at most ϵ. If now we replace the probability space $(X, \boldsymbol{p})$ by its n-fold cartesian product $(X^n, \boldsymbol{p}^{\otimes n})$ (n i.i.d. copies of $(X, \boldsymbol{p})$) and allow n to increase to infinity then an application of the law of large numbers leads to the following result:

$$\lim_{n\to\infty} \frac{\log \nu(\boldsymbol{p}^{\otimes n}, \epsilon)}{\log \#X^n} = \frac{H(\boldsymbol{p})}{\log X}.$$

Or equivalently,

$$\lim_{n\to\infty} \frac{\log \nu(\boldsymbol{p}^{\otimes n}, \epsilon)}{n} = H(\boldsymbol{p}) \quad \text{for all } \epsilon > 0 \tag{9.2.36}$$

where $H(\boldsymbol{p})$ is the Shannon entropy of $\boldsymbol{p}$. This is a special case of Macmillan's theorem in classical information theory. Our next result is a quantum analogue of (9.2.36), which also implies (9.2.36). Let $(\mathcal{H}, \rho)$ be a quantum probability space where $\mathcal{H}$ is a finite dimensional Hilbert space and ρ is a state. For any projection operator E on $\mathcal{H}$ denote by $\dim E$ the dimension of the range of E. For any $\epsilon > 0$ define

$$\nu(\rho, \epsilon) = \min\{\dim E \mid E \text{ is a projection in } \mathcal{H}, \operatorname{Tr} \rho E \geq 1 - \epsilon\}.$$

Theorem 9.2.37 (Schumacher) *For any* $\epsilon > 0$

$$\lim_{n\to\infty} \frac{\log \nu(\rho^{\otimes n},\ \epsilon)}{n} = S(\rho) \tag{9.2.38}$$

where $S(\rho)$ *is the von Neumann entropy of* ρ.

Proof By the spectral theorem ρ can be expressed as

$$\rho = \sum_x p(x)\ |x\rangle\langle x|$$

where x varies in a finite set X of labels, $\boldsymbol{p} = \{p(x), x \in X\}$ is a probability distribution with $p(x) > 0$ for every x and $\{|x\rangle,\ x \in X\}$ is an orthonormal set in $\mathcal{H}$. Then

$$\rho^{\otimes n} = \sum_{\boldsymbol{x}=(x_1,x_2,\ldots,x_n)} p(x_1)p(x_2)\ldots p(x_n)\ |\boldsymbol{x}\rangle\langle\boldsymbol{x}|$$

where x_i's vary in X and $|\boldsymbol{x}\rangle$ denotes the product vector $|x_1\rangle|x_2\rangle\ldots|x_n\rangle$. Write $p_n(\boldsymbol{x}) = p(x_1)\ldots p(x_n)$ and observe that $\boldsymbol{p}^{\otimes n} = \{p_n(\boldsymbol{x}),\ \boldsymbol{x} \in X^{\otimes n}\}$ is the probability distribution of n i.i.d. copies of $\boldsymbol{p}$. We have $S(\rho) = -\sum_x p(x) \log p(x) = H(\boldsymbol{p})$. From the strong law of large numbers for i.i.d. random variables it follows that

$$\lim_{n\to\infty} -\frac{1}{n} \log p(x_1)p(x_2)\ldots p(x_n) = \lim_{n\to\infty} -\frac{1}{n} \sum_{i=1}^{n} \log p(x_i) = S(\rho)$$

in the sense of almost sure convergence in the probability space $(X^\infty, \boldsymbol{p}^{\otimes\infty})$. This suggests that, in the search for a small set of high probability, we consider the set

$$T(n,\epsilon) = \left\{ \boldsymbol{x} : \left| -\frac{1}{n}\log p(x_1)p(x_2)\dots p(x_n) - S(\rho) \right| \le \epsilon \right\} \quad (9.2.39)$$

Any element of $T(n,\epsilon)$ is called an *ϵ-typical sequence* of length n. It is a consequence of the large deviation principle that there exist constants $A > 0$, $0 < c < 1$ such that

$$\Pr(T(n,\epsilon)) \ge 1 - Ac^n, \quad (9.2.40)$$

Pr denoting probability but according to the distribution $\boldsymbol{p}^{\otimes n}$. This says but for a set of sequences of total probability $< Ac^n$ every sequence is ϵ-typical. It follows from (9.2.39) that for any ϵ-typical sequence $\boldsymbol{x}$

$$2^{-n(S(\rho)+\epsilon)} \le p_n(\boldsymbol{x}) \le 2^{-n(S(\rho)-\epsilon)}. \quad (9.2.41)$$

Define the projection

$$E(n,\epsilon) = \sum_{\boldsymbol{x}\in T(n,\epsilon)} |\,\boldsymbol{x}\rangle\langle\boldsymbol{x}\,| \quad (9.2.42)$$

and note that $\dim E(n,\epsilon) = \#T(n,\epsilon)$. Summing over $\boldsymbol{x} \in T(n,\epsilon)$ in (9.2.41) we conclude that

$$2^{-n(S(\rho)+\epsilon))} \dim E(n,\epsilon) \le \Pr(T(n,\epsilon)) \le 2^{-n(S(\rho)-\epsilon))} \dim E(n,\epsilon)$$

and therefore by (9.2.40) and the fact that probabilities never exceed 1, we get

$$2^{n(S(\rho)-\epsilon))}(1 - Ac^n) \le \dim E(n,\epsilon) \le 2^{n(S(\rho)+\epsilon))}$$

for all $\epsilon > 0, n = 1, 2, \dots$. In particular

$$\frac{\log \dim E(n,\epsilon)}{n} \le S(\rho) + \epsilon.$$

Fix ϵ and let $\delta > 0$ be arbitrary. Choose n_0 so that $Ac^{n_0} < \delta$. Note that $\operatorname{Tr}\rho^{\otimes n}E(n,\epsilon) = \Pr(T(n,\epsilon)) \ge 1 - \delta$ for $n \ge n_0$. By the definition of $\nu(\rho^{\otimes n}, \delta)$ we have

$$\frac{\log \nu(\rho^{\otimes n}, \delta)}{n} \le \frac{\log \dim E(n,\epsilon)}{n} \le S(\rho) + \epsilon, \text{ for } n \ge n_0.$$

Letting $n \to \infty$ we get

$$\overline{\lim}_{n\to\infty} \frac{\log \nu(\rho^{\otimes n}, \delta)}{n} \leq S(\rho) + \epsilon.$$

Since ϵ is arbitrary we get

$$\overline{\lim}_{n\to\infty} \frac{\log \nu(\rho^{\otimes n}, \delta)}{n} \leq S(\rho).$$

Now we shall arrive at a contradiction by assuming that

$$\underline{\lim}_{n\to\infty} \frac{\log \nu(\rho^{\otimes n}, \delta)}{n} < S(\rho).$$

Under such a hypothesis there would exist an $\eta > 0$ such that

$$\frac{\log \nu(\rho^{\otimes n}, \delta)}{n} \leq S(\rho) - \eta$$

for infinitely many n, say $n = n_1, n_2, \ldots$ where $n_1 < n_2 < \cdots$. In such a case there exists a projection F_{n_j} in $\mathcal{H}^{\otimes n_j}$ such that

$$\begin{aligned} \dim F_{n_j} &\leq 2^{n_j(S(\rho)-\eta))} \\ \operatorname{Tr} \rho^{\otimes n_j} F_{n_j} &\geq 1 - \delta \end{aligned} \tag{9.2.43}$$

for $j = 1, 2, \ldots$. Choosing $\epsilon < \eta$ and fixing it we have

$$\begin{aligned} 1 - \delta &\leq \operatorname{Tr} \rho^{\otimes n_j} F_{n_j} \\ &= \operatorname{Tr} \rho^{\otimes n_j} E(n_j, \epsilon) F_{n_j} + \operatorname{Tr} \rho^{\otimes n_j} (I - E(n_j, \epsilon)) F_{n_j}. \end{aligned} \tag{9.2.44}$$

From (9.2.40) and the fact that $\rho^{\otimes n}$ and $E(n, \epsilon)$ commute with each other we have

$$\begin{aligned} \operatorname{Tr} \rho^{\otimes n_j} (I - E(n_j, \epsilon)) F_{n_j} &\leq \operatorname{Tr} \rho^{\otimes n_j} (I - E(n_j, \epsilon)) \\ &= 1 - \Pr(T(n_j, \epsilon)) \\ &< A c^{n_j}. \end{aligned} \tag{9.2.45}$$

Furthermore from (9.2.41) we have

$$\rho^{\otimes n_j} E(n_j, \epsilon) = \sum_{\boldsymbol{x} \in T(n_j, \epsilon)} p_{n_j}(\boldsymbol{x}) \, |\boldsymbol{x}\rangle\langle \boldsymbol{x}| \leq 2^{-n_j(S(\rho)-\epsilon))} I.$$

Thus by (9.2.43) we get

$$\begin{aligned} \operatorname{Tr} \rho^{\otimes n_j} E(n_j, \epsilon) F_{n_j} &\leq 2^{-n_j(S(\rho)-\epsilon))} \dim F_{n_j} \\ &\leq 2^{-n_j(S(\rho)-\epsilon))+n_j(S(\rho)-\eta))} \\ &= 2^{-n_j(\eta-\epsilon)}. \end{aligned} \tag{9.2.46}$$

Now combining (9.2.44), (9.2.45) and (9.2.46) we get

$$1 - \delta \leq 2^{-n_j(\eta-\epsilon)} + Ac^{n_j},$$

where the right side tends to 0 as $j \to \infty$, a contradiction.

□

Property 20) Feinstein's fundamental lemma

Consider a classical information channel $\mathcal{C}$ equipped with an input alphabet A, an output alphabet B and a transition probability $\{p_x(V), x \in A, V \subset B\}$. We assume that both A and B are finite sets. If a letter $x \in A$ is transmitted through the channel $\mathcal{C}$ then any output $y \in B$ is possible and $p_x(V)$ denotes the probability that the output letter belongs to V under the condition that x is transmitted. For such a channel we define a *code* of *size* N and *error probability* $\leq \epsilon$ to be a set $C = \{c_1, c_2, \ldots, c_N\} \subset A$ together with a family $\{V_1, V_2, \ldots, V_N\}$ of disjoint subsets of B satisfying the condition $p_{c_i}(V_i) \geq 1 - \epsilon$ for all $i = 1, 2, \ldots, N$. Let

$$\nu(\mathcal{C}, \epsilon) = \max \left\{ N \;\middle|\; \begin{array}{l} \text{there exists a code of size } N \text{ and error proba-} \\ \text{bility} \leq \epsilon \end{array} \right\}.$$

Our aim is to estimate $\nu(\mathcal{C}, \epsilon)$ in terms of information theoretic parameters concerning the conditional distributions $p_x(\cdot)$, $x \in A$, denoted by $\boldsymbol{p}_x$. To this end consider an input probability distribution $p(x)$, $x \in A$, denoted by $\boldsymbol{p}$ and define the joint input-output distribution $\boldsymbol{P}$ such that $\Pr(x, y; \boldsymbol{P}) = p(x)p_x(\{y\})$. From now onwards we write $\Pr(x, y)$ instead of $\Pr(x, y, \boldsymbol{P})$. Denote by $H_{\boldsymbol{p}}(A : B)$ the mutual information between the input and the output according to the joint distribution $\boldsymbol{P}$. Put

$$C = \sup_{\boldsymbol{p}} H_{\boldsymbol{p}}(A : B) \tag{9.2.47}$$

where the supremum is taken over all input distributions $\boldsymbol{p}$. For a fixed input distribution $\boldsymbol{p}$, put

$$\sigma^2_{\boldsymbol{p}} = \sum_{x \in A, y \in B} \Pr(x, y) \left\{ \log \frac{\Pr(x, y)}{p(x)q(y)} - H_{\boldsymbol{p}}(A : B) \right\}^2 \tag{9.2.48}$$

where $\boldsymbol{q}$ is the B-marginal distribution determined by $\boldsymbol{P}$. Thus $q(y) = \sum_x \Pr(x, y)$. With these notations we have the following lemma.

Lemma 9.2.49 *Let $\eta > 0$, $\delta > 0$ be positive constants and let $\boldsymbol{p}$ be any input distribution on A. Then there exists a code of size N and error probability $\leq \eta$ where*

$$N \geq \left(\eta - \frac{\sigma_p^2}{\delta^2}\right) 2^{H_p(A:B)-\delta}.$$

Proof Put $R = H_p(A : B)$. Define the random variable ξ on the probability space $(A \times B, \boldsymbol{P})$ by

$$\xi(x, y) = \log \frac{\Pr(x, y)}{p(x)q(y)}.$$

Then ξ has expectation R and variance $\sigma_{\boldsymbol{p}}^2$ defined by (9.2.48). Let

$$V = \left\{(x, y) : \left|\log \frac{\Pr(x, y)}{p(x)q(y)} - R\right| \leq \delta\right\}. \tag{9.2.50}$$

Then by Chebyshev's inequality for the random variable ξ we have

$$\Pr(V; \boldsymbol{P}) \geq 1 - \frac{\sigma_p^2}{\delta^2}. \tag{9.2.51}$$

Define $V_x = \{y \mid (x, y) \in V\}$. Then (9.2.51) can be expressed as

$$\sum_{x \in A} p(x)p_x(V_x) \geq 1 - \frac{\sigma_p^2}{\delta^2}. \tag{9.2.52}$$

This shows that for a $\boldsymbol{p}$-large set of x's the conditional probabilities $p_x(V_x)$ must be large. When $(x, y) \in V$ we have from (9.2.50)

$$R - \delta \leq \log \frac{\Pr(x, y)}{p(x)q(y)} \leq R + \delta$$

or equivalently

$$q(y)2^{R-\delta} \leq p_x(y) \leq q(y)2^{R+\delta}.$$

Summing over $y \in V_x$ we get

$$q(V_x)2^{R-\delta} \leq p_x(V_x) \leq q(V_x)2^{R+\delta}.$$

In particular,

$$q(V_x) \leq p_x(V_x)2^{-(R-\delta)} \leq 2^{-(R-\delta)}. \tag{9.2.53}$$

In other words V_x's are q-small. Now choose x_1 in A such that $p_{x_1}(V_{x_1}) \geq 1-\eta$ and set $V_1 = V_{x_1}$. Then choose x_2 such that $p_{x_2}(V_{x_2} \cap V_1') > 1-\eta$, where the prime $'$ denotes complement in B. Put $V_2 = V_{x_2} \cap V_1'$. Continue this procedure till we have an x_N such that

$$p_{x_N}(V_{x_N} \cap V_1' \cap \cdots \cap V_{N-1}') > 1-\eta,$$

and for any $x \notin \{x_1, x_2, \ldots, x_N\}$,

$$p_x(V_x \cap (\cup_{j=1}^N V_j)') \leq 1-\eta$$

where $V_N = V_{x_N} \cap V_1' \cap \cdots \cap V_{N-1}'$. By choice the sets $V_1, V_2, \ldots, V_N$ are disjoint, $\cup_{i=1}^N V_i = \cup_{i=1}^N V_{x_i}$ and therefore

$$p_x(V_x \cap (\cup_{j=1}^N V_j)') \leq 1-\eta \quad \text{for all } x \in A. \tag{9.2.54}$$

From (9.2.52), (9.2.53) and (9.2.54) we have

$$\begin{aligned}
1 - \frac{\sigma_p^2}{\delta^2} &\leq \sum_x p(x)p_x(V_x) \\
&= \sum_x p(x)p_x(V_x \cap (\cup_{i=1}^N V_i)') + \sum_x p(x)p_x(V_x \cap (\cup_{i=1}^N V_i)) \\
&\leq 1 - \eta + \sum_x p(x)p_x(V_x \cap (\cup_{i=1}^N V_i)) \\
&= 1 - \eta + q(\cup_{i=1}^N V_i) \\
&\leq 1 - \eta + \sum_{i=1}^N q(V_i) \\
&\leq 1 - \eta + \sum_{i=1}^N q(V_{x_i}) \\
&\leq 1 - \eta + N2^{-(R-\delta)}.
\end{aligned}$$

Thus $N \geq \left(\eta - \frac{\sigma_p^2}{\delta^2}\right) 2^{(R-\delta)}$.

□

Now we consider the n-fold product $\mathcal{C}^{(n)}$ of the channel $\mathcal{C}$ with input alphabet A^n, output alphabet B^n and transition probability $\{p_{\boldsymbol{x}}^{(n)}(V), \boldsymbol{x} \in A^n, V \subset B^n\}$ where for $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$,

$$p_{\boldsymbol{x}}^{(n)}(\{\boldsymbol{y}\}) = \prod_{i=1}^{n} p_{x_i}(\{y_i\}).$$

We now choose and fix an input distribution p on A and define the product probability distribution $\boldsymbol{P}^{(n)}$ on $A^n \times B^n$ by

$$P^{(n)}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} p(x_i) p_{x_i}(\{y_i\}).$$

Then the A^n marginal of $\boldsymbol{P}^{(n)}$ is given by

$$p^{(n)}(\boldsymbol{x}) = \prod_{i=1}^{n} p(x_i)$$

and $H_{\boldsymbol{p}^{(n)}}(A^n : B^n) = nH_{\boldsymbol{p}}(A : B)$, $\sigma^2_{\boldsymbol{p}^{(n)}} = n\sigma^2_{\boldsymbol{p}}$ where $\sigma^2_{\boldsymbol{p}}$ is given by (9.2.48). Choose $\eta > 0$, $\delta = n\epsilon$ and apply the Lemma 9.2.49 to the product channel. Then it follows that there exists a code of size N and error probability $\leq \eta$ with

$$N \geq \left(\eta - \frac{n\sigma^2_{\boldsymbol{p}}}{n^2\epsilon^2}\right) 2^{n(H_{\boldsymbol{p}}(A:B)-\epsilon)} = \left(\eta - \frac{\sigma^2_{\boldsymbol{p}}}{n\epsilon^2}\right) 2^{n(H_{\boldsymbol{p}}(A:B)-\epsilon)}.$$

Thus

$$\frac{1}{n} \log \nu(\mathcal{C}^{(n)}, \eta) \geq \frac{1}{n} \log \left(\eta - \frac{\sigma^2_{\boldsymbol{p}}}{n\epsilon^2}\right) + H_{\boldsymbol{p}}(A : B) - \epsilon.$$

In other words

$$\underline{\lim}_{n\to\infty} \frac{1}{n} \log \nu(\mathcal{C}^{(n)}, \eta) \geq H_{\boldsymbol{p}}(A : B) - \epsilon.$$

Here the positive constant ϵ and the initial distribution p on the input alphabet A are arbitrary. Hence we conclude that

$$\underline{\lim}_{n\to\infty} \frac{1}{n} \log \nu(\mathcal{C}^{(n)}, \eta) \geq C.$$

It has been shown by J. Wolfowitz ([16])that

$$\overline{\lim}_{n\to\infty}\frac{1}{n}\log\nu(\mathcal{C}^{(n)},\eta)\leq C.$$

The proof of this assertion is long and delicate and we refer the reader to [16]. We summarize our discussions in the form of a theorem.

Theorem 9.2.55 (Shannon-Wolfowitz) *Let $\mathcal{C}$ be a channel with finite input and output alphabets A and B respectively and transition probability $\{p_x(V), x\in A, V\subset B\}$. Define the constant C by* (9.2.47)*. Then*

$$\lim_{n\to\infty}\frac{1}{n}\log\nu(\mathcal{C}^{(n)},\eta)=C \quad \textit{for all } 0<\eta<1.$$

Remark 9.2.56 The constant C deserves to be and is called the *capacity* of the discrete memory less channel determined by the product of copies of $\mathcal{C}$.

A quantum information channel is characterized by an input Hilbert space $\mathcal{H}_A$, an output Hilbert space $\mathcal{H}_B$ and a quantum operation $\mathcal{E}$ which maps states on $\mathcal{H}_A$ to states on $\mathcal{H}_B$. We assume that $\mathcal{H}_A$ and $\mathcal{H}_B$ are finite dimensional. The operation $\mathcal{E}$ has the form

$$\mathcal{E}(\rho)=\sum_{i=1}^{k}L_i\rho L_i^\dagger \tag{9.2.57}$$

where $L_1,\ldots,L_k$ are operators from $\mathcal{H}_A$ to $\mathcal{H}_B$ obeying the condition $\sum_i L_i^\dagger L_i=I_A$. A message encoded as the state ρ on $\mathcal{H}_A$ is transmitted through the channel and received as a state $\mathcal{E}(\rho)$ in $\mathcal{H}_B$ and the aim is to recover ρ as accurately as possible from $\mathcal{E}(\rho)$. Thus $\mathcal{E}$ plays the role of transition probability in the classical channel. The recovery is implemented by a recovery operation which maps states on $\mathcal{H}_B$ to states on $\mathcal{H}_A$. A *quantum code* $\mathcal{C}$ of *error not exceeding* ϵ can be defined as a subspace $\mathcal{C}\subset\mathcal{H}_A$ with the property that there exists a recovery operation $\mathcal{R}$ of the form

$$\mathcal{R}(\rho')=\sum_{j=1}^{\ell}M_j\rho' M_j^\dagger \quad \text{for any state } \rho' \text{ on } \mathcal{H}_B$$

where the following conditions hold:

1. $M_1, \ldots, M_\ell$ are operators from $\mathcal{H}_A$ to $\mathcal{H}_B$ satisfying $\sum_{j=1}^{\ell} M_j^\dagger M_j = I_B$;
2. for any $\psi \in \mathcal{C}$, $\langle\psi|\mathcal{R} \circ \mathcal{E}(|\,\psi\rangle\langle\psi\,|)|\psi\rangle \geq 1 - \epsilon$.

Now define

$$\nu(\mathcal{E}, \epsilon) = \max\{\dim \mathcal{C} \mid \mathcal{C} \text{ is a quantum code of error not exceeding } \epsilon\}.$$

We may call $\nu(\mathcal{E}, \epsilon)$ the maximal size possible for a quantum code of error not exceeding ϵ. As in the case of classical channels one would like to estimate $\nu(\mathcal{E}, \epsilon)$.

If $n > 1$ is any integer define the n-fold product $\mathcal{E}^{\otimes n}$ of the operation $\mathcal{E}$ by

$$\mathcal{E}^{\otimes n} = \sum_{i_1, i_2, \ldots, i_n} L_{i_1} \otimes L_{i_2} \otimes \cdots \otimes L_{i_n} \rho L_{i_1}^\dagger \otimes L_{i_2}^\dagger \otimes \cdots \otimes L_{i_n}^\dagger$$

for any state ρ on $\mathcal{H}_A^{\otimes n}$, where the L_i's are as in (9.2.57). It is an interesting problem to analyze the asymptotic behavior of the sequence $\left\{\frac{1}{n} \log \nu(\mathcal{E}^{\otimes n}, \epsilon)\right\}$ as $n \to \infty$.

Bibliography

[1] J. Aczel and Z. Daroczy, *On Measures of Information and Their Characterizations*, Academic Pub., New York, 1975.

[2] A. Aho, J. Hopcroft and J. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, Massachusetts, 1974.

[3] M. Artin, *Algebra*, Prentice Hall of India Pvt. Ltd., 1996

[4] I. L. Chuang and M. A. Nielsen, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[5] T. H. Cormen, C. E. Leiserson and R. L. Rivest, *Introduction to Algorithms*, McGraw-Hill Higher Education, March 1, 1990.

[6] G. H. Hardy and E. M. Wright, *Introduction to the Theory of Numbers*, ELBS and Oxford University Press, 4th edition, 1959.

[7] W. C. Huffman and Vera Pless, *Fundamentals of Error-correcting Codes*, Cambridge University Press, Cambridge, 2003.

[8] N. Jacobson, *Basic Algebra I, II*, Freeman, San Francisco, 1974, 1980.

[9] A. I. Khinchin, *Mathematical Foundations of Information Theory*, New York, 1957.

[10] D. E. Knuth, *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*, 3rd edition, Addison-Wesley, 1997.

[11] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933 (*Foundations of the Theory of Probability*, Chelsea, New york, 1950).

[12] D. C. Kozen, *The Design and Analysis of Algorithms*, Springer-Verlag, 1992.

[13] F. J. Macwilliams and N. J. A. Sloane, *Theory of Error-correcting Codes*, North Holland, Amsterdam, 1978.

[14] J. von Neumann, *Mathematical Foundations of Quantum Mechanics (translated from German)*, Princeton University Press, 1955. Original in, Collected Works, Vol 1, pp. 151–235, edited by A.H. Taub, Pergamon Press 1961.

[15] K. R. Parthasarathy, *Lectures on Error-correcting Codes*, Indian Statistical Institute, New Delhi.

[16] J. Wolfowitz, *Coding Theorems of Information Theory*, Springer Verlag, 3rd edition, 1978.