量子计算和量子信息(二)

——量子信息部分

Quantum Computation and Quantum Information

Michael A. Nielsen
Isaac L. Chuang 著

郑大钟 赵千川 译

清华大学出版社

本书是剑桥大学出版社出版的Michael A. Nielsen和Isaac L. Chuang合著的Quantum Computation and Quantum Information的量子信息部分的中译本。

量子计算与量子信息是涉及物理学、计算机科学和数学等多学科的综合性交叉研究领域。本书首先介绍量子噪声和纠错，然后介绍熵，最后介绍量子信息论。

本书完整系统地介绍了量子计算与量子信息的最新成果和基本知识。本书内容深入浅出，层次分明，参考文献丰富，它既可作一般有兴趣的读者了解该领域的入门读物，也可用做大专院校的教材，或供大学高年级学生和研究生自学使用，对相关领域的研究人员也有很大的参考价值。

郑大钟　清华大学自动化系教授，博士生导师。长期从事控制理论教学。出版教材6部，《线性系统理论》获电子工业部优秀教材一等奖和国家级教学成果二等奖。2000年被亚洲控制教授协会(ACPA)授予“ACPA Myoung Sam Ko 控制教育奖”。主要研究线性系统理论、控制系统的鲁棒性、大系统分散控制、离散事件动态系统、混合动态系统、电力系统安全性等。曾主持国家级项目多项，发表论文100多篇。获“何潘清漪优秀论文奖”两次，获国家教委科技进步奖(理论)三等奖1次。

现任中国自动化学会荣誉理事和控制理论专业委员会副主任，《自动化学报》副主编，亚洲控制杂志(AJC)编委等。曾任第八届全国政协委员，亚洲控制教授协会副主席，教育部“高等教育面向21世纪教学内容和课程体系改革”顾问组成员等。

赵千川，清华大学自动化系智能与网络化系统研究中心教授，主要研究离散事件动态系统(DEDS)理论及其在制造、通信等领域的应用。1987年考入清华大学自动化系，1996年获得控制理论与应用专业工学博士学位，并留校任教。2000年3月至2001年4月得到国家留学基金资助，赴美国Carnegie Mellon大学从事访问研究。曾获2000年度“何潘清漪优秀论文奖”，该奖为DEDS领域对华人的最高奖。

ISBN 7-302-09756-9

9 787302 097563 >

定价：33.00元

量子计算和量子信息(二)

——量子信息部分

Quantum Computation and Quantum Information

Michael A. Nielsen
Isaac L. Chuang 著

郑大钟 赵千川 译

清华大学出版社

北 京

内 容 简 介

本书是剑桥大学出版社出版的 Michael A. Nielsen 和 Isaac L. Chuang 合著的 Quantum Computation and Quantum Information 的量子信息部分的中译本.

量子计算与量子信息是涉及物理学、计算机科学和数学等多学科的综合性交叉研究领域. 本书首先介绍量子噪声和纠错,然后介绍熵,最后介绍量子信息论.

本书完整系统地介绍了量子计算与量子信息的最新成果和基本知识. 本书内容深入浅出,层次分明,参考文献丰富. 它既可作一般有兴趣的读者了解该领域的入门读物,也可用做大专院校的教材,或供大学高年级学生和研究生自学使用,对相关领域的研究人员也有很大的参考价值.

Michael A. Nielsen, Isaac L. Chuang
Quantum Computation and Quantum Information
© Cambridge University Press 2000
First Published 2000
All Rights Reserved.
For sale in Main land China only.

本书翻译版由剑桥大学出版社授权清华大学出版社在中国境内独家出版、发行.
未经出版者书面许可,不得以任何方式复制或抄袭本书的任何部分.

北京市版权局著作权合同登记号 图字:01-2001-4170

版权所有,翻印必究。举报电话:010-62782989 13801310933

图书在版编目(CIP)数据

量子计算和量子信息(二)——量子信息部分/(美)尼尔森(Nielsen, M. A.),(美)庄(Chuang, I. L.)著;郑大钟,赵千川译. —北京:清华大学出版社,2005.1
书名原文:Quantum Computation and Quantum Information
ISBN 7-302-09756-9

Ⅰ.量… Ⅱ.①尼… ②庄… ③郑…④赵… Ⅲ.①量子力学—光通信 ②第五代计算机
Ⅳ.①TN929.1 ②TP387

中国版本图书馆 CIP 数据核字(2004)第 107116 号

出 版 者:清华大学出版社 **地 址**:北京清华大学学研大厦
http://www.tup.com.cn **邮 编**:100084
社 总 机:010-62770175 **客户服务**:010-62776969
组稿编辑:王一玲
文稿编辑:邹开颜
封面设计:常雪影
版式设计:刘祎淼
印 装 者:三河市春园印刷有限公司
发 行 者:新华书店总店北京发行所
开 本:175×245 **印张**:18.75 **字数**:360 千字
版 次:2005 年 2 月第 1 版 2005 年 2 月第1 次印刷
书 号:ISBN 7-302-09756-9/O·419
印 数:1～3000
定 价:33.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770175-3103 或(010)62795704

译者序

量子计算与量子信息的研究可以追溯到几十年前，但真正引起广泛关注的是20世纪90年代中期. 这期间发现了Shor量子因子分解算法和Grover量子搜索算法，这两类算法展示了量子计算从根本上超越经典计算机计算能力和在信息处理方面的巨大潜力. 与此同时，量子计算机和量子信息处理装置在物理实现的研究，成为继并行计算机、生物计算机等之后的非串行计算体系的又一热点.

量子计算与量子信息对人类社会最具影响也最为惊人的发现之一是，量子计算机能够迅速破解广泛采用的RSA密码系统. 掌握量子计算能力的制高点已成为关系信息安全的重要课题，不少国家已纷纷开始启动和资助相关的研究项目.

国外不少大学也已开设了有关课程. 译者2000年访问的美国Carnegie Mellon大学，他们在计算机系和物理系的研究生中开设了量子计算机课程. 所采用的教材正是剑桥大学出版社出版的Michael A. Nielsen和Isaac L. Chuang的英文版原著Quantum Computation and Quantum Information，本书是该著作量子计算部分的中译本.

原著共12章，分三个部分，分别介绍基础知识以及量子计算和量子信息. 由于篇幅宏大，为方便读者和照顾不同背景读者的需要，原著将内容安排为两个相对独立的主题：第一和第二部分一起构成学习量子计算的相对完整的材料；第一和第三部分构成量子信息相对完整的内容. 中译本继承原著者的思想，分为量子计算和量子信息两册单独出版，供读者根据需要选择. 本书限于量子信息内容，对应原著的第8～12章共5章内容. 其中8～10章由郑大钟译出，11～12章由赵千川译出. 赵千川负责统稿. 量子计算内容（原著第1～7章）已单独出版.

本书在写作上定位为教材，因此照顾到广大读者在背景知识上的差异，尽可能以浅显和自成体系的方式叙述主要研究思路，力图深入浅出. 在细节的处理上较好地保持了严谨性和启发性的折衷. 特别是在一般专业读者较为生疏的量子力学方面，大胆地采用了基于线性代数的公理化体系，大大简化了学习主题的途径. 这与我国物理专业量子力学教学改革中的类似尝试不谋

而合.

译者在翻译过程中可喜地看到,我国研究工作者也已经开始了相关领域的研究,并已取得了一些成果,如已有论文集和总结国内外研究成果的学术专著出版.应该说量子计算与量子信息的研究还远没有成熟,该领域的研究充满着令人兴奋的挑战性课题.译者衷心希望本书的出版能为量子计算与量子信息方面知识在我国的传播起到一定推动作用.

译者

2003 年 8 月于

清华大学

PREFACE

前　言

本书介绍量子计算和量子信息领域的主要思想和方法. 由于这个学科领域的迅速发展和其交叉学科性质，初学者对该领域最重要的方法和成果获得全面了解并不容易.

因而本书有两方面的目的. 首先介绍计算机科学、数学和物理方面必要的背景知识，读者需要具有三个学科中至少一个学科的相当于研究生入学水平；其中最重要的是要有一定的数学修养和希望了解量子计算与量子信息的愿望. 本书的第二个目的是详尽叙述量子计算与量子信息的核心成果. 通过深入学习，读者能够掌握这个令人激动的领域的基本工具和成果. 这可作为读者一般教育的一部分，或作为他独立从事量子计算与量子信息研究的准备.

本书结构

本书的基本框架如图 1 所示，共分为三个部分. 叙述的基本原则是从具体到抽象. 先讲量子计算后讲量子信息；先讲特殊的量子纠错码后讲量子信息论的一般结果；先讲例子后讲一般理论.

第一部分概述量子计算与量子信息领域的主要思想和成果，并介绍量子计算与量子信息所必需的计算机科学、数学和物理背景知识. 第 1 章是介绍性的，介绍该领域的发展历史和基本概念，着重介绍了历史上的若干重要的未解决问题(open problem). 这部分读者即使不具备计算机科学或物理学背景，也可以读懂. 第 2 章和第 3 章给出了更深入、详细的背景知识，分别详尽叙述量子力学和计算机科学的基本概念. 读者可根据个人的背景，重点阅读第一部分的某些章节，后面必要时可返回来阅读，来获得所需的量子力学和计算机科学知识.

第二部分详尽叙述量子计算. 第 4 章描述量子计算所需基本元素，给出更复杂应用中要用到的基本运算. 第 5 章、第 6 章描述两个已知的量子算法：量子 Fourier 变换和量子搜索算法. 第 5 章还解释量子 Fourier 变换如何用于解因子分解(factoring)和离散对数(discrete logarithm)问题，以及这些结果对密码系统的重要性. 第 7 章以已在实验室获得成功的几个实现为例，来阐述量

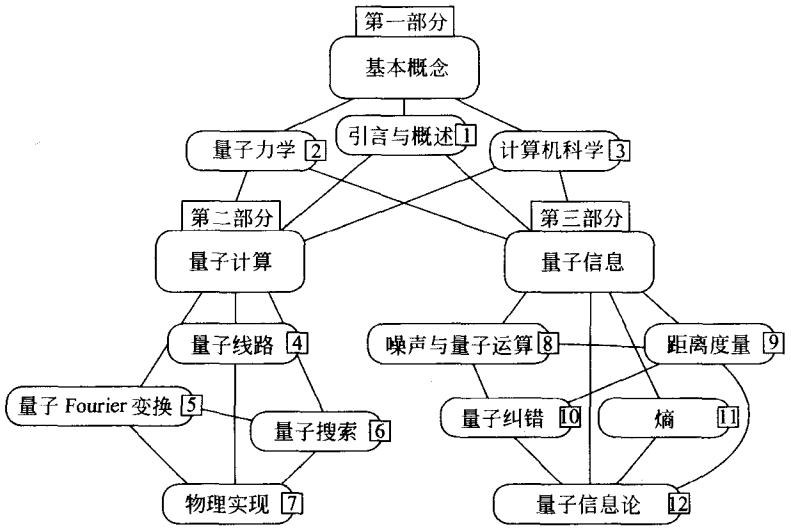


图 1　本书的结构

子计算机好的物理实现的一般原则.

第三部分有关量子信息:即如何用量子状态表示和传送信息,如何对付经典信息和量子信息的损失. 第 8 章描述了用来理解现实世界量子信息处理的量子噪声性质和对理解量子噪声非常有用的量子运算形式化. 第 9 章描述精确量化两个量子信息相似程度的距离度量. 第 10 章讲量子纠错码,量子纠错码可用来使量子计算避免受到噪声的影响. 这章的一个重要结果是阈值定理. 阈值定理表明,在真实的噪声模型中,噪声原则上不对量子计算构成严重妨碍. 第 11 章引入基础信息论的概念——熵,并给出经典和量子信息论中熵的许多性质. 最后,第 12 章讨论量子状态和量子信道的信息承载属性,详尽描述这类系统传送经典信息和量子信息以及机密信息时具有的许多特殊性质.

本书配有大量练习和问题. 练习贯穿在文中,为巩固对基本内容的理解而设,除个别情况,很容易在几分钟内完成这些练习. 问题则安排在每章后边,用于补充一些由于正文篇幅限制,而未给出的有趣的新材料. 问题常由几部分构成,目的是对特定的思路作一定深度的阐述. 有几个问题在本书付印时尚未解决,这在叙述时作了说明. 每章以整章主要结果的概要结束,并以"历史和进一步阅读的材料"为一节给出整章的主要思路、参考文献和推荐的阅读材料.

本书正文之前有目录、名词和记号.

本书正文之后包括五个附录和一个参考文献.

附录 A 复习初等概率论的一些基本概念、记号和结论. 我们假设读者熟悉这

部分内容，包括进来的目的只是便于参考. 同样为方便读者，附录 B 复习群论的基本概念. 附录 C 包含量子计算的一个重要结论 Solovay-Kitaev 定理的证明，该定理表明量子门的有限集合可以用来快速逼近任意的量子门. 附录 D 复习理解量子因子分解和离散对数算法以及 RSA 密码系统所必需的数论知识. RSA 密码系统在下册的附录中介绍. 附录 E 包括量子计算与量子信息中最重要的定理之一 Lieb 定理的证明，该定理是重要的熵不等式（如著名的次可加不等式）的雏形. 因为 Solovay-Kitaev 定理和 Lieb 定理的证明较长，所以需要独立于正文给出.

参考文献列出书中引用的全部文献，同时向由于疏忽而未被引用的作者表示歉意.

量子计算和量子信息领域发展非常迅速，这使得我们对所有的论题无法按照希望的深度展开. 但三个方面需特别提及. 第一个主题是纠缠（entanglement）测量，如书中所解释的，纠缠现象是量子隐形传态（teleportation）、快速量子算法和量子纠错等效应中的关键要素，简言之，是量子计算与量子信息的利器. 纠缠作为一种新的物理资源，寻求、驾驭它的规律和用途正成为一个兴起的研究方向. 我们认为尽管这方面的研究极富吸引力，但还没有达到像本书其他主题那样完整的程度，所以我们在第 12 章仅给出一个简述. 同样，考虑到极富吸引力的分布式量子计算（有时称量子通信复杂性）的研究非常活跃，为避免书未出版而内容过时，所以没有涉及. 量子信息处理机的实现也已成为一个有趣和成果丰富的方向，我们仅用一章的篇幅介绍，但物理实现有很多的内容，这涉及物理、化学和工程中更多的领域，因而不得不割爱.

如何使用本书

本书可用作多种用途，可以作为各类课程的基础教材，从用于讲授量子计算与量子信息的短期专题基础讲座到涉及整个领域的全年的正式课程. 只想对量子计算与量子信息稍作了解的读者可以自学；想进入研究前沿的读者也可采用本书. 本书的目的之一还在于作为该领域的一本参考书，特别希望它对初次接触这个领域的研究人员有价值.

致自学读者

本书考虑到自学读者的需要，文中准备了大量的练习，可用来理解正文内容，并进行自我测试. 目录和每章后边的提要可以帮助读者很快决定哪些章节需要透彻学习. 图 1 所示的关系图可帮助读者决定阅读的顺序.

致教师

本书覆盖了很宽范围的主题，可以作为多种课程的基础课本.

一个学期的量子计算课程可以根据学生的背景选择：第 1 到第 3 章部分内容、第 4 章量子线路、第 5 章、第 6 章量子算法的全部、第 7 章物理实现的一部分、第 8 到第 10 章特别是第 10 章关于量子纠错的全部内容.

一学期的量子信息课程也可以根据学生的背景选择：第 1 到第 3 章部分内容、第 8 到第 10 章关于量子纠错、第 11 章量子熵和第 12 章量子信息论.

一学年课程可以覆盖整本书的内容，还可以增加从部分章节的“历史和进一步阅读”中选择的额外阅读材料. 量子计算与量子信息领域本身还可为学生提供很好的研究题目.

除了用于量子计算与量子信息课程，我们还希望本书的内容作为物理系学生的量子力学专业的引论. 传统的量子力学课的引论非常强调偏微分方程的数学工具，我们认为这常常阻碍学生对基本思想的把握. 量子计算与量子信息为理解量子力学的基本概念和特殊现象提供了绝好的思想实验场所. 这样的课程可以集中于第 2 章量子力学引论、第 4 章量子线路的基本内容、第 5 章、第 6 章量子算法的一部分、第 7 章量子计算物理实现、然后根据兴趣选择本书第三部分的内容.

致学生

我们尽量使本书内容自成体系. 主要例外是我们偶尔会略去一些需要读者自己弄明白的论述，这些地方留作练习，当然我们假设读者愿意尝试解决书中所有的练习. 几乎所有练习都可在几分钟之内解决，如果读者对许多练习遇到很多困难，也许应该回到前面去掌握一些关键的概念.

进一步阅读的材料

如前所述，每章最后有一节“历史和进一步阅读的材料”，还有一些内容包括很广的参考文献. Preskill[Pre98b] 关于量子计算与量子信息的讲义与本书的角度有些不同. 关于专题的好的综述文章有（按在本书中出现的先后）：Aharonov 关于量子计算的综述[Aha99b]，Kitave 关于算法和纠错的综述[Kit97b]，Mosca 关于量子算法的学位论文[Mos99]，Fuch 关于量子信息中可区分性和距离测度的学位论文[Fuc96]，Gottesman 关于量子纠错的学位论文[Got97]，Preskill[Pre97] 关于量子纠错的综述，Nielsen 关于量子信息论的学位论文[Nie98] 和 Bennett 与 Shor[BS98]、Bennett 与 Divinceno 关于量子信息论的综述[BD00]. 其他有价值的参考资料包括 Gruska 的书[Gru99] 和 Lo、Spiller 和 Popescu 编辑的综述文集[LSP98].

更正

任何大篇幅的文本总有一些错误和疏忽，本书也不例外. 如果你发现任何错误或有任何关于本书的评论，请通过电子邮件发给：qci@squint. org. 本书英文版的更正信息可在以下网址找到：http://www. squint. org/qci/.（中译本电子邮件发给 zhaoqc@tsinghua. edu. cn）

NOMENCLATURE AND NOTATION

名词和记号

量子计算与量子信息中一些名词和记号有两个以上含义，这里收集了许多本书经常使用的条目及其在本书中的约定，以避免混淆.

线性代数与量子力学

除特别声明，所有向量空间均假定为有限维. 许多情况下该限制是不必要的，或者可以通过其他技巧去掉，不过作这样的全局性限制，可以使表达更易理解，也不会分散读者对运用结果的注意力.

半正定算子 A 满足对任意 $|\psi\rangle$，$\langle\psi|A|\psi\rangle\geqslant 0$. 正定算子 A 满足对任意 $|\psi\rangle\neq 0$，$\langle\psi|A|\psi\rangle>0$. 算子的支集定义为正交于其核的向量空间. 对 Hermite 算子即为由非零特征值的特征向量所张成的向量空间.

记号 U（或 V）一般用于表示酉算子或酉矩阵. H 通常用来表示 Hadamard 门，有时也表示量子系统的 Hamilton 函数.

向量有时写作列的形式，如

$$\begin{bmatrix}1\\2\end{bmatrix} \tag{1}$$

有时写成便于阅读的形式(1,2)，后一种形式应理解为列向量的简写. 对于用作量子比特（或称量子位，qubit）的双态量子系统，常将状态 $|0\rangle$ 等同于向量 (1,0)，状态 $|1\rangle$ 等同于(0,1). 这里采用 Pauli sigma 矩阵的传统定义（参见下面的“常用的量子门和线路的符号”）. 注意，习惯上 Pauli sigma z 矩阵为 $\sigma_z|0\rangle=|0\rangle$，$\sigma_z|1\rangle=-|1\rangle$，这与某些物理学家（通常不是计算机科学家或数学家）的直观期待相反. 这个不一致来源于 σ_z 的本征态 $+1$ 常被物理学家等同于激发态，因而很多人很自然地将其等同于 $|1\rangle$，而不是像本书中等同于 $|0\rangle$. 我们这样做是为了保持线性代数矩阵元素指标的一致性，也就是让 σ_z 的第一列自然地表示 σ_z 在 $|0\rangle$ 上的作用，第二列表示 σ_z 在 $|1\rangle$ 上的作用，这个做法在量子计算与量子信息领域内是通行的. 除了 σ_x，σ_y 和 σ_z 这些 Pauli sigma 矩阵的传统记号外，用 σ_1，σ_2，σ_3 来表示这三个矩阵也很方便. 我们用 σ_0 表示 2×2 单位阵. 不过最常用的记号还是用 I，X，Y 和 Z 分别表示 σ_0，σ_1，σ_2 和 σ_3.

信息论和概率

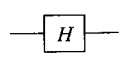
按照信息论的习惯，除非特别说明，对数总是取以 2 为底. 我们用 $\log(x)$ 表示以 2 为底的对数，在个别情况下用到 $\ln(x)$ 表示自然对数. **概率分布**指满足 $p_x \geqslant 0$ 和 $\sum_x p_x = 1$ 的有限实数集合 p_x. 半正定算子 A 相对于半正定算子 B 的**相对熵**定义为 $S(A \parallel B) \equiv \mathrm{tr}(A\log A) - \mathrm{tr}(A\log B)$.

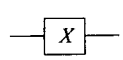
杂项

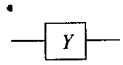
⊕记模 2 的加法. 全书中 z 念作“zed”.

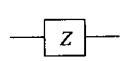
常用量子门和线路的符号

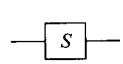
引入设计量子线路常用的表示酉变换的一些方框符号，为方便读者，收集如下：酉变换的行和列从左到右、从上到下标为 $00\cdots0, 00\cdots1$ 到 $11\cdots1$，最下方的线是重要性最小的线. 注意 $e^{i\pi/4}$ 是 i 的平方根，故 $\pi/8$ 门是相位门(phase gate)的平方根. 相位门本身是 Pauli-Z 门的平方根.

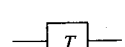
Hadamard 门  $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$

Pauli-X 门  $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$

Pauli-Y 门  $\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$

Pauli-Z 门  $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$

相位门  $\begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}$

$\frac{\pi}{8}$ 门  $\begin{bmatrix}1 & 0\\ 0 & e^{i\pi/4}\end{bmatrix}$

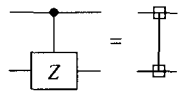
受控非门  $\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}$

交换门



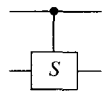
$$\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}$$

受控 Z 门



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

受控相位门



$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&\mathrm{i}\end{bmatrix}$$

Toffoli 门



$$\begin{bmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\end{bmatrix}$$

Fredkin(受控交换) 门

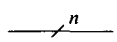


$$\begin{bmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\end{bmatrix}$$

测量运算  投影到$|0\rangle$和$|1\rangle$上

量子比特 承载单量子比特的线(时间为先左后右)

经典比特 承载单经典比特的线

n 量子比特  承载 n 量子比特的线

CONTENTS

目 录

译者序 …… 1
前言 …… 3
名词和记号 …… 9

第三部分 量子信息

第 8 章 量子噪声和量子运算 …… 3
8.1 经典噪声和 Markov 过程 …… 4
8.2 量子运算 …… 6
8.2.1 技术总述 …… 6
8.2.2 环境和量子运算 …… 7
8.2.3 算子和表示 …… 9
8.2.4 量子运算的公理化方法 …… 16
8.3 量子噪声和量子运算的例子 …… 22
8.3.1 迹与偏迹 …… 22
8.3.2 单量子比特量子运算的几何图像 …… 23
8.3.3 比特翻转和相位翻转信道 …… 24
8.3.4 去极化信道 …… 26
8.3.5 幅值阻尼 …… 28
8.3.6 相位阻尼 …… 32
8.4 量子运算的应用 …… 35
8.4.1 主方程 …… 35
8.4.2 量子过程层析 …… 37
8.5 量子运算体系的局限性 …… 43

第 9 章 量子信息的距离度量 …… 47
9.1 经典信息的距离度量 …… 47
9.2 两个量子状态有多接近 …… 51

9.2.1 迹距离 …… 51
9.2.2 保真度 …… 56
9.2.3 距离度量之间的关系 …… 62
9.3 量子信道对信息的保持 …… 63

第 10 章 量子纠错 …… 72
10.1 引言 …… 73
10.1.1 三量子比特的比特翻转码 …… 74
10.1.2 三量子比特相位翻转码 …… 77
10.2 Shor 码 …… 79
10.3 量子纠错的理论 …… 82
10.3.1 差错的离散化 …… 85
10.3.2 独立差错模型 …… 88
10.3.3 简并编码 …… 90
10.3.4 量子 Hamming 界 …… 91
10.4 量子码的构造 …… 92
10.4.1 经典线性码 …… 92
10.4.2 Calderbank-Shor-Steane 码 …… 96
10.5 稳定子码 …… 100
10.5.1 稳定子体系 …… 100
10.5.2 酉门和稳定子体系 …… 105
10.5.3 稳定子体系中的测量 …… 109
10.5.4 Gottesman-Knill 定理 …… 110
10.5.5 稳定子码构造 …… 111
10.5.6 例子 …… 113
10.5.7 稳定子码的标准形 …… 116
10.5.8 编码、解码和纠错的量子线路 …… 118
10.6 容错量子计算 …… 122
10.6.1 容错：大图像 …… 122
10.6.2 容错量子逻辑 …… 128
10.6.3 容错测量 …… 135
10.6.4 容错量子计算基础 …… 140

第 11 章 熵与信息 …… 146
11.1 Shannon 熵 …… 146

11.2 熵的基本属性 …… 149
11.2.1 二元熵 …… 149
11.2.2 相对熵 …… 150
11.2.3 条件熵和互信息 …… 151
11.2.4 数据处理不等式 …… 154
11.3 Von Neumann 熵 …… 155
11.3.1 量子相对熵 …… 157
11.3.2 熵的基本性质 …… 158
11.3.3 测量与熵 …… 159
11.3.4 次可加性 …… 160
11.3.5 熵的凹性 …… 161
11.3.6 混合量子状态的熵 …… 162
11.4 强次可加性 …… 163
11.4.1 强次可加性的证明 …… 164
11.4.2 强次可加性：基本应用 …… 166

第 12 章 量子信息论 …… 171
12.1 区分量子状态和可访问的信息 …… 172
12.1.1 Holevo 界 …… 175
12.1.2 应用 Holevo 界的例子 …… 176
12.2 数据压缩 …… 179
12.2.1 Shannon 无噪声信道编码定理 …… 180
12.2.2 Schumacher 量子无噪声信道编码定理 …… 184
12.3 带噪声量子信道上的经典信息 …… 189
12.3.1 带噪声经典信道上的通信 …… 189
12.3.2 带噪声量子信道上的通信 …… 195
12.4 带噪声量子信道上的量子信息 …… 201
12.4.1 熵交换与量子 Fano 不等式 …… 201
12.4.2 量子数据处理不等式 …… 203
12.4.3 量子单一界 …… 208
12.4.4 量子纠错、制冷和 Maxwell 妖 …… 209
12.5 作为物理资源的纠缠 …… 210
12.5.1 二部纯态纠缠的变换 …… 212
12.5.2 纠缠的蒸馏和稀释 …… 217
12.5.3 纠缠蒸馏与量子纠错 …… 219

12.6 量子密码术…… 221
12.6.1 私钥密码术…… 221
12.6.2 保密增强与信息调和…… 222
12.6.3 量子密钥分配…… 225
12.6.4 保密性与量子相干信息…… 230
12.6.5 量子密钥分配的安全性…… 231

附录 A Lieb 定理证明 …… 246

参考文献…… 251

第三部分

量子信息

CHAPTER 8

第 8 章 量子噪声和量子运算

迄今为止，我们还只是涉及到封闭量子系统的动力学过程. 封闭量子系统就是不容许与外部世界有任何交互作用的量子系统. 尽管在这类理想化系统中可以完成的那些信息处理任务，在原理上可以得出一些令人着迷的结论，但是，基于现实世界中不存在完全封闭系统(除非把宇宙看成整体)这一事实，这些结果大大减少了其价值. 实际的系统都会经受到与外部世界的不期望的交互作用. 可以把这些不需要的交互作用看成是量子信息处理系统中的噪声. 为了构筑有用的量子信息处理系统，我们需要理解和控制这些噪声过程. 这就是本书第三部分的一个中心课题. 这个课题以本章中量子运算形式化的描述作为开始. 对于描述量子噪声和开放量子系统行为，量子运算形式化是一组强有力的工具.

一个开放系统和一个封闭系统之间究竟有什么区别呢？某些机械式时钟中的悬摆可以当作为一个近似的封闭系统，摆与世界的其余部分——它的环境——主要通过摩擦来形成很轻微的交互作用. 但是，为了合理地描述摆的完整动力学过程和为什么悬摆最终停止其运动，必须要考虑空气摩擦的阻尼效应和摆悬置机构的不完善缺陷. 类似地，没有一个量子系统所有时候都是完全封闭的，特别是量子计算机不完全封闭，它必须用一个外部系统来巧妙地编程以执行某组期望的运算. 举例来说，如果将一个量子比特(qubit)的状态用一个电子的两个位置来表示，那么这个电子将会与其他带电粒子产生交互作用，这种作用成为影响这个量子比特状态的不能控制的噪声源. 一个开放系统就是与其他环境系统具有交互作用的系统，而这个环境系统的动力学过程我们希望予以忽略或均化.

量子运算的数学形式化对于我们描述开放量子系统的动力学过程是一个关键的工具. 这种工具非常强有力，利用它可同时来提供范围广泛的物理图像. 这种工具不仅可用来描述与其环境弱耦合的近似封闭系统，而且也可描述与其环境强耦合的系统，以及那些突然被打开并接受量测的封闭系统. 量子运算在应用于量子计算和量子信息时的一个另外的优点是，它们特别适合于描述离散状态变化，即在无需时间迁移显式参照下初始状态 ρ 和最末状态 ρ' 之间的变换. 这种离散时间分析

相当不同于物理学家为描述开放量子系统所传统采用的倾向于连续时间描述的工具，如主方程、Langevin 方程和随机微分方程等.

本章的构成如下. 8.1 节中我们以经典系统中如何描述噪声的讨论作为开始. 在学习如何考虑量子运算和量子噪声的过程中，通过理解经典噪声所获得的直觉是非常宝贵的. 8.2 节从三个不同的观点介绍量子运算的体系，以使我们能对量子运算的基本理论有更为透彻的了解. 8.3 节对采用量子运算的噪声提供几个重要的例子，包括像去极化、幅值阻尼和相位阻尼这样一些例子. 此外，在应用 Bloch（布洛兹）球同时，对理解单个量子比特上量子噪声的一种几何方法进行解释. 8.4 节则来解释量子运算的一些各种各样的应用：量子运算同物理学家传统地用于描述量子噪声的其他工具之间的联系，例如：主方程(master equation)；如何应用熟知的量子过程层析(quantum process tomography)方法来试验地确定一个量子系统所经受的动力学过程；解释这些量子运算如何可以被用于理解一个基本事实，这就是我们周围的世界看来是服从经典物理学规律的，而它们实际上是遵循量子力学定律的. 最后，8.5 节讨论以量子运算体系作为一般性方法来描述量子系统中噪声的局限性.

8.1 经典噪声和 Markov 过程

为了理解量子系统中的噪声，通过理解经典系统中的噪声来建立某些直觉将会是有帮助的. 我们应当如何来对经典系统中的噪声建模呢？让我们看一些简单的例子，以了解如何来做到这一点，以及有关量子系统中的噪声能教给我们一些什么.

想象一个比特正被存储到内置于普通经典计算机中的一个硬盘驱动器里. 这个比特出发于 0 或 1，但在一段长时间后很可能变为这样的情况，扰动磁场会对这个比特造成干扰，并可能引起其状态翻转. 通过引入使这个比特翻转的概率为 p 和使这个比特保持原状态的概率为 $1-p$，我们就可对这类问题建模. 这个过程说明于图 8.1 中.

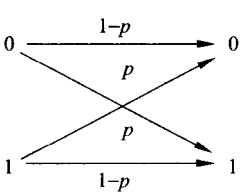


图 8.1 在一段长时间后硬盘驱动器上的一个比特可能以概率 p 翻转.

当然，真正发生什么取决于环境包含能造成这个比特翻转的磁场. 为计算这个比特翻转概率 p，我们需要了解两件事情. 第一，需要建立环境中磁场分布的模型. 假定硬盘驱动器使用者没有愚蠢到做像在靠近硬盘驱动器的地方使用强磁铁这样的事情，那么通过类似于驱动器所运行的环境中对磁场采样，我们就能

构筑一个接近实际的模型. 第二,我们需要建立环境中磁场如何与硬盘中比特产生交互作用的模型. 幸好,这种模型对于物理学家已是有所了解的,这就是 Maxwell(麦克斯韦)方程. 利用上述这两个要素,我们就能从原理上来计算在某个规定时间段内驱动器上一个比特会出现翻转的概率 p.

这个基本的过程(找出环境的模型以及系统-环境间交互作用的模型)是我们在研究经典噪声和量子噪声中将会反复遵循的过程. 与环境的交互作用是经典系统和量子系统两者中的基本噪声来源. 往往不容易找到针对环境或系统-环境间交互作用的精确模型;但是,通过在建模中引入保守性和限于研究所观测到的系统性质以判断其是否服从这个模型,那么还是有可能使现实物理系统中噪声的建模达到高的准确度.

硬盘上这个比特的行为可以用一个单一方程来简洁地表征. 设 p_0 和 p_1 为这个比特分别处于状态 0 和 1 的初始概率,令 q_0 和 q_1 为在噪声出现之后这个比特处于状态 0 和 1 的相应概率,令 X 为这个比特的初始状态,Y 为这个比特的最终状态. 那么,全概率公式(《量子计算和量子信息(一)——量子计算部分》附录 A)可陈述为

$$p(Y=y)=\sum_x p(Y=y\mid X=x)p(X=x) \tag{8.1}$$

条件概率 $p(Y=y|X=x)$,由于其表征系统中可能出现的变化而被称为转移概率. 显式地列写出硬盘上这个比特的方程,有

$$\begin{bmatrix} q_0 \\ q_1 \end{bmatrix}=\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}\begin{bmatrix} p_0 \\ p_1 \end{bmatrix} \tag{8.2}$$

让我们来看一个较为复杂一些的经典系统中噪声的例子. 假设,我们正试图来构筑一个经典线路以执行某个计算任务,不幸的是,我们被提供来组成这个线路的是有缺陷的元件. 这个有点人为的线路由作用于两个顺序的(有缺陷)非门的单输入比特 X、中间比特 Y 和最终比特 Z 所组成. 看来可以合理地假定,第二个非门是否正确工作独立于第一个非门是否正确工作. 这个假定(顺序噪声过程为独立起作用)对处理很多情形是一个物理上合理的假定. 这就导致一个熟知为 Markov(马尔可夫)过程的特殊类型随机过程 $X\to Y\to Z$. 物理上,这个 Markov 假定对应于如下假定:引起第一个非门中噪声的环境与引起第二个非门中噪声的环境是独立地作用的. 其合理性由两个门可认为在空间上有相当距离所提供.

概括起来,经典系统中的噪声可以采用随机过程理论来描述. 通常,在多级过程的分析中,采用 Markov 过程是一个好的假定. 对于单级过程,输出概率 $\vec{q}$ 通过如下方程相关于输入概率 $\vec{p}$:

$$\vec{q}=E\vec{p} \tag{8.3}$$

其中,E 为转移概率矩阵,我们将称其为演化矩阵. 因此,系统的最终状态跟起始状

态线性相关. 这种线性属性会重现于量子噪声的描述中，其中采用密度矩阵代替概率分布.

演化矩阵 E 必会具有什么样的性质呢？我们要求，如果输入概率 $\vec{p}$ 是一个合法的概率分布，那么 $E\vec{p}$ 必也是一个合法的概率分布. 使这个条件满足可化为等价于关于演化矩阵 E 的两个条件. 第一，E 的所有元必须为非负的，即所知为正性(positivity)要求的条件. 如果 E 的所有元不都是非负的，那么就有可能在 $E\vec{p}$ 中得到负概率. 第二，E 的所有列必须取和为 1，即所知为完备性(completeness)要求的条件. 设这一点不成立. 例如，E 的第一列不取和为 1. 令输入概率 $\vec{p}$ 的第一个元为 1 而其余元均为 0，则我们看到，在这种情况下 $E\vec{p}$ 不再是一个合法的概率分布.

概括起来，经典噪声的关键属性为如下几个方面：输入概率和输出概率之间具有线性关系；线性关系由转移矩阵所描述；转移矩阵具有非负元(正性)和所有列均取和为 1(完备性). 在规定噪声是由独立的环境所引起前提下，包含多级的经典噪声过程可以描述为 Markov 过程. 这些关键属性中的每一个都可在量子噪声理论中具有重要的相似属性. 当然，也会存在某些令人惊奇的新的量子噪声属性.

8.2 量子运算

8.2.1 技术总述

很像 Markov 过程描述经典状态的随机变化那样，对于描述广泛类型的量子系统的演化，包括量子状态的随机变化，量子运算体系是一个一般化的工具. 正如同经典状态采用概率向量来描述，我们将利用密度算子(密度矩阵)ρ来描述量子状态. 读者如果想复习它的一些有关属性，则可在继续学习本章之前，阅读 2.4 节(见《量子计算和量子信息(一)——量子计算部分》). 同时，类似于经典状态变换如何由式(8.3)所描述那样，量子状态变换可描述为

$$\rho' = \varepsilon(\rho) \tag{8.4}$$

这个方程中的映射 ε 是一个量子运算(quantum operation). 我们在第 2 章中已经先遇到过的量子运算的两个简单例子，分别为酉变换 $\varepsilon(\rho)=U\rho U^{\dagger}$ 和测量 $\varepsilon_m(\rho)=M_m\rho M_m^{\dagger}$(参看下面的练习 8.1 和练习 8.2). 量子运算会引起作为某个物理过程结果而出现的状态的动态变化；ρ为过程前的初始状态，$\varepsilon(\rho)$为过程出现后的最终状态，有可能要除去某个归一化因子.

贯穿如下各节，我们会结合酉演化、测量和甚至更一般的过程来展开量子运算的一般化理论. 我们会来介绍理解量子运算的三种相区别的方法，如图 8.2 中所说明的，所有这些方法是等价的. 第一种方法是基于这样的思想，把研究动力学过程归结为研究系统与环境间的交互作用，很像 8.1 节中描述经典噪声那样. 这种方法

是具体的，并易于与实际世界相联系. 遗憾的是，这种方法具有数学上不方便的弊端. 理解量子运算的第二种方法就是，在完全等价于第一种方法的同时，通过对量子运算提供所熟知的“算子和表示”(operator-sum representation)这一强有力的数学表示来克服这种数学上的不方便性. 这种方法相当抽象，但对计算和理论工作是很有用的. 理解量子运算的第三种方法是来源于物理上的一组公理，并等价于其他两种方法，它们就是量子力学中动力学映射所期望满足的公理. 这种方法的优点在于其非常的一般性——它显示出，在令人惊讶的广泛情况下，量子动力学过程都可由量子运算来描述. 但是，这种方法既不能提供像第二种方法的计算方便性，也不具有像第一种方法的具体属性. 归结起来，量子运算的这三种方法提供了一个强有力的工具，应用这些方法我们就能来理解量子噪声及其影响.

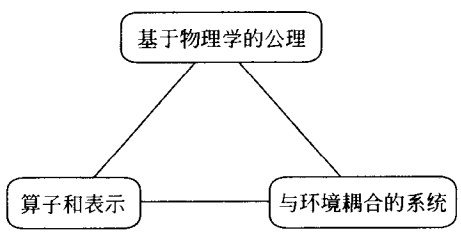


图 8.2　量子运算的等价的但取决于期望应用而具有不同优点的三种方法.

练习 8.1(作为量子运算的酉演化)　纯状态在酉变换下演化为 $|\psi\rangle \rightarrow U|\psi\rangle$，试证明，对 $\rho = |\psi\rangle\langle\psi|$，等价地可写为 $\rho \rightarrow \varepsilon(\rho) \equiv U\rho U^{\dagger}$.

练习 8.2(作为量子运算的测量)　回顾第 2.2.3 节，具有标号为 m 的结果的一个量子测量可用一组测量算子 M_m 来描述，使 $\sum_m M_m^{\dagger} M_m = I$ 成立. 令系统在测量前瞬间的状态为 ρ，试证明，对 $\varepsilon_m(\rho) \equiv M_m \rho M_m^{\dagger}$，系统在测量后瞬间的状态为

$$\frac{\varepsilon_m(\rho)}{\mathrm{tr}(\varepsilon_m(\rho))} \tag{8.5}$$

并证明，获得这个测量结果的概率为 $p(m) = \mathrm{tr}(\varepsilon_m(\rho))$.

8.2.2　环境和量子运算

封闭量子系统的动力学过程是用酉变换来描述的. 概念上，我们可以把酉变换想象为一个盒子，输入状态加入到盒子中，并从盒子中得到输出，如同图 8.3 的左边部分所指明的那样. 这里，盒子的内部运行机制并不为我们所关注，它可以是用一个量子线路、某个 Hamilton(哈密顿)系统或其他任何东西来实现的.

描述开放量子系统动力学过程的一个自然方法是，把其看成是从称之为主系统(principle system)的所感兴趣系统跟连同系统会形成为一个封闭量子系统的环境之



图 8.3 封闭量子系统(左)和开放量子系统(右)的模型,一个开放量子系统由主系统和环境两部分组成.

间的交互作用中所出现的,如图 8.3 中右边部分所指明的那样. 换句话说,设我们有一个处于状态 ρ 的系统,此状态被送到耦合于环境的一个盒子中. 一般来说,这个系统的最终状态 $\varepsilon(\rho)$ 有可能并不是通过酉变换而与初始状态 ρ 相联系的. 我们现在假定,系统-环境输入状态为一个积状态 $\rho\otimes\rho_{env}$. 在盒子的变换 U 后,系统不再与环境具有交互作用,从而我们就在环境上执行一个偏迹,以得到系统单独的约化状态

$$\varepsilon(\rho) = tr_{env}[U(\rho\otimes\rho_{env})U^{\dagger}] \tag{8.6}$$

当然,如果变换 U 不包括与环境的任何交互作用,那么就有 $\varepsilon(\rho)=\tilde{U}\rho\tilde{U}^{\dagger}$,其中 $\tilde{U}$ 为 U 中单独作用到系统的部分. 式(8.6)就是量子运算三个等价定义中的第一个.

这个定义中做了一个重要假定,系统和环境是从一个积状态出发的. 当然,一般地说这个假定是不成立的. 量子系统会持续地跟其环境产生交互作用,同时形成其相关性. 这种现象的一个表现途径是通过在系统跟其环境之间的热交换. 置于环境中的量子系统达到跟其相同的温度,这就造成了两者之间的相关性. 但是,在很多实际中有意义的情况中,可以合理地假定,系统及其环境是从积状态启动的. 当试验者来制备处于指定状态下的量子系统时,他们会取消那个系统和环境间的所有相关性. 理想地,相关性将会被完全破坏,留下处于纯态的系统. 即使是不属于这种情形,我们随后也将会看到,当系统和环境不是以积状态启动时,量子运算体系甚至还能用来描述量子动力学过程.

人们可能会提出的另一个问题是,如果环境具有几乎无穷维的自由度,变换 U 还如何可以来被指定呢? 非常有趣的是,可以断言,为了使这个模型能合理地来描述任何可能的变换 $\rho\to\varepsilon(\rho)$,如果主系统具有 d 维 Hilbert(希尔伯特)空间,那么就有能力在不大于 d^2 维的 Hilbert 空间中对环境进行建模. 还可断言,对于环境没有必要以混合态来启动;从纯态启动就够了. 我们将在 8.2.3 节中回到这些点上去.

作为式(8.6)应用的一个显式例子,考虑图 8.4 所示的双量子比特量子线路,线路中 U 为受控非门,主系统为控制量子比特,而环境作为目标量子比特初始时处于状态

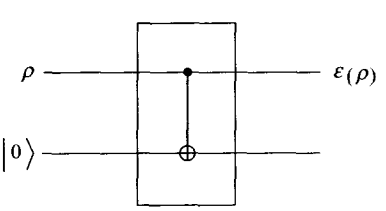


图 8.4 作为单量子比特量子运算基本例子的受控非门.

$\rho_{env}=|0\rangle\langle 0|$. 将此插入式(8.6)的同时，容易看出

$$\varepsilon(\rho)=P_0\rho P_0+P_1\rho P_1 \tag{8.7}$$

其中，$P_0=|0\rangle\langle 0|$ 和 $P_1=|1\rangle\langle 1|$ 为投影算子. 直观上，出现这个动力学过程是因为，仅当系统处于状态 $|0\rangle$ 时环境位于状态 $|0\rangle$，否则环境就被翻转到状态 $|1\rangle$. 8.2.3 节中，作为算子和表示的一个例子，我们会来给出这个方程的推导.

我们已经基于主系统和环境的交互作用来描述量子运算；还可方便地将这个定义稍作推广允许不同的输入和输出空间. 例如，想象一个标示为 A 的单量子比特被制备于某个未知状态 ρ，标示为 B 的三级量子系统(qutrit)被制备于某个标准状态 $|0\rangle$，然后通过酉交互作用 U 使其与系统 A 产生交互作用，以使联合系统演化到状态 $U(\rho\otimes|0\rangle\langle 0|)U^\dagger$. 随后，我们丢弃系统 A，留下处于某个最终状态 ρ' 的系统 B. 根据定义，描述这个过程的量子运算 ε 为

$$\varepsilon(\rho)=\rho'=\mathrm{tr}_A(U(\rho\otimes|0\rangle\langle 0|)U^\dagger) \tag{8.8}$$

注意，ε 将输入系统 A 的密度算子映射到输出系统 B 的密度算子. 我们下面对量子运算的大部分讨论只是关注某个系统 A 上的量子运算，也即它们将系统 A 的密度算子映射到系统 A 的密度算子. 但是，也允许更为一般的定义，这在应用中偶尔是会有用的. 通过将量子运算定义作为如下过程结果而出现的一类映射，就可来提供这样一个定义：某个初始系统被制备于未知量子状态 ρ，然后同其他制备于标准状态的系统连接，允许依据某个酉交互作用进行交互，随后将组合系统的某一部分丢弃，仅留下处于某个状态 ρ' 的最终系统. 定义这个过程的量子运算 ε 只是简单地映射 ρ 到 ρ'. 我们将会看到，允许不同输入和输出空间的这种推广，可以自然地跟通过算子和表示的量子运算的处理相融合，同时也跟我们的公理化研究相融合. 然而，对于大部分情况，如果我们假定量子运算的输入和输出空间为相同，并且采用“主系统”和“环境”之间这种方便的区分(这在一般情形中是看不到的)，就可使讨论得以简化；而当输入和输出空间为不同时，会同时给出少量的练习以指明必要的推广.

8.2.3　算子和表示

量子运算能够以优美形式来表示，即所知的算子和表示，它本质上就是单独根据主系统的 Hilbert 空间上的算子对式(8.6)的一种显式重述. 这个主要结果起源于如下的简单计算. 令 $|e_k\rangle$ 为环境的(有限维)状态空间的一个标准正交基，再令 $\rho_{env}=|e_0\rangle\langle e_0|$ 为环境的初始状态. 不失一般性，假定环境启动于纯态，因为如果其启动于混合态，我们完全可以来引入纯化环境的一个附加的外部系统(见 2.5 节). 尽管这个外部系统是虚构的，但对主系统所经历的动力学过程并没有造成任何区别，因而可被用来作为计算中的一个中间步骤. 式(8.6)可因之而被重新写为

$$\varepsilon(\rho)=\sum_k \langle e_k|U[\rho\otimes|e_0\rangle\langle e_0|]U^\dagger|e_k\rangle \tag{8.9}$$

$$=\sum_k E_k\rho E_k^\dagger \tag{8.10}$$

其中，$E_k\equiv\langle e_k|U|e_0\rangle$为主系统的状态空间上的一个算子. 式(8.10)称为 ε 的算子和表示，算子$\{E_k\}$称为量子运算 ε 的运算元. 算子和表示是重要的，它将反复地用于本书的剩余部分中.

运算元满足称为完备性关系(completeness relation)的重要约束，这个完备性类似于经典噪声描述中演化矩阵的完备性关系. 在经典情形中，完备性关系是由要求概率分布归化到 1 而提出的. 在量子情形中，完备性关系则是从类似地要求$\varepsilon(\rho)$的迹等于 1 而提出的：

$$1=\mathrm{tr}(\varepsilon(\rho)) \tag{8.11}$$

$$=\mathrm{tr}\left(\sum_k E_k\rho E_k^\dagger\right) \tag{8.12}$$

$$=\mathrm{tr}\left(\sum_k E_k^\dagger E_k\rho\right) \tag{8.13}$$

因为这个关系对所有 ρ都成立，所以我们必有

$$\sum_k E_k^\dagger E_k=I \tag{8.14}$$

这个方程对保迹(trace-preserving)的量子运算是满足的. 也存在非保迹的量子运算，对其成立 $\sum_k E_k^\dagger E_k\leqslant I$. 它们可来描述这样一些过程，其中出现于过程中的额外信息是通过测量得到的，如同我们稍后会更详细地解释的那样. 满足 $\sum_k E_k^\dagger E_k\leqslant I$ 的式(8.10)形式的映射 ε 可提供量子运算的第二个定义. 我们下面会来证明，这个定义本质上等价于第一个定义式(8.6)，且事实上还会稍微更一般化一些，因为它允许于非保迹运算. 今后我们将经常有机会在这两个定义之间切换；但从上下文应该可以清楚判断，我们在任何给定时刻正在采用哪个定义.

练习 8.3 我们对算子和表示的推导隐含地假定运算的输入和输出空间是相同的. 设初始处于未知量子状态 ρ复合的复合系统 AB 跟初始处于某个标准状态 $|0\rangle$的组复合系统 CD 相接触，两个系统根据酉作用 U 进行交互作用. 在交互作用后，我们丢弃系统 A 和 D，留下系统 BC 的状态 ρ'. 试证明，对从系统 AB 的状态空间到系统 CD 的状态空间的某组线性算子 E_k，成立 $\sum_k E_k^\dagger E_k=I$，则映射 $\varepsilon(\rho)=\rho'$ 满足

$$\varepsilon(\rho)=\sum_k E_k\rho E_k^\dagger \tag{8.15}$$

算子和表示之所以重要是因为它提供给我们表征主系统动力学过程的一个内在的手段. 算子和表示可以无需显式地考虑环境的性质而来描述主系统的动力学

过程；我们所需知道的一切都被塞进到仅仅作用于主系统的算子 E_k 中. 这既会简化计算，还通常会提供相当多的理论上的洞察力. 此外，许多不同的环境交互作用有可能在主系统上引起相同的动力学过程. 如果仅是对主系统动力学过程感兴趣，那么只选取一个不包含关于其他系统不重要信息的动力学过程的表示是正确的.

在本节的剩余部分中，我们来考察算子和表示的性质，特别是它的三个特性. 首先，我们根据运算元 E_k 来对它提供一个物理解释. 从这方面产生的一个自然的问题是，对任一开放量子系统（例如，给定系统-环境交互作用或其他的规范）如何来确定算子和表示. 这一点会在下面提出的第二个课题中给以回答. 反之，对任一算子和表示如何来构造开放量子系统的模型将在最后回答.

练习 8.4（测量）　设我们有一个单量子比特主系统，其通过变换

$$U = P_0 \otimes I + P_1 \otimes X \tag{8.16}$$

与单量子比特的环境产生交互作用. 其中，X 为（作用于环境上的）普通的 Pauli 矩阵，$P_0 = |0\rangle\langle 0|$ 和 $P_1 = |1\rangle\langle 1|$ 为（作用于系统上的）投影算子. 假定环境从状态 $|0\rangle$ 出发的同时，试给出该过程的量子运算的算子和表示.

练习 8.5（自旋翻转）　像前一个练习中那样，但现在令

$$U = \frac{X}{\sqrt{2}} \otimes I + \frac{Y}{\sqrt{2}} \otimes X \tag{8.17}$$

试给出该过程量子运算的算子和表示.

练习 8.6（量子运算的复合）　设 ε 和 $\mathcal{F}$ 为同一量子系统上的量子运算，试证明，复合运算 $\varepsilon \circ \mathcal{F}$ 是在具有算子和表示意义下的一个量子运算. 再陈述和证明上述这个结果的一个推广形式，这里，对 ε 和 $\mathcal{F}$ 不必要具有相同输入和输出空间。

1. 算子和表示的物理解释

对算子和表示可以给出一个漂亮的解释. 设想在酉变换 U 以后，环境的测量在基 $|e_k\rangle$ 上执行. 应用隐含测量原理，我们看到，这样一个测量仅只影响环境的状态，而不改变主系统的状态. 令 ρ_k 为结果 k 出现时的主系统状态，于是

$$\begin{aligned} \rho_k &\propto \mathrm{tr}_E(|e_k\rangle\langle e_k| U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger |e_k\rangle\langle e_k|) \\ &= \langle e_k| U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger |e_k\rangle \qquad (8.18) \\ &= E_k \rho E_k^\dagger \qquad (8.19) \end{aligned}$$

归一化 ρ_k，有

$$\rho_k = \frac{E_k \rho E_k^\dagger}{\mathrm{tr}(E_k \rho E_k^\dagger)} \tag{8.20}$$

我们发现，结果 k 的概率可给出为

$$\begin{aligned} p(k) &= \mathrm{tr}(|e_k\rangle\langle e_k| U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger |e_k\rangle\langle e_k|) \qquad (8.21) \\ &= \mathrm{tr}(E_k \rho E_k^\dagger) \qquad (8.22) \end{aligned}$$

因此

$$\varepsilon(\rho) = \sum_k p(k)\rho_k = \sum_k E_k \rho E_k^\dagger \tag{8.23}$$

这个方程对具有运算元$\{E_k\}$的量子运算的机理提供了一个漂亮的物理解释，量子运算的作用等价于取定状态ρ并随机地以概率$\mathrm{tr}(E_k\rho E_k^\dagger)$将它替换为$E_k\rho E_k^\dagger/\mathrm{tr}(E_k\rho E_k^\dagger)$. 在这个意义上，这与采用于经典信息论中的带噪声通信信道的概念是非常类似的. 按此思路，我们有时也将称描述量子噪声过程的某些量子运算为带噪声的量子信道.

算子和表示的这种解释，基于图 8.4，可由一个简单的例子来加以说明. 设我们选取状态$|e_k\rangle = |0_E\rangle$和$|1_E\rangle$，其中我们包含下标"E"是为了清楚地表明这个状态是环境的状态. 这一点可以解释为，如图 8.5 中所示那样，我们是在环境量子比特的计算基中来进行测量的. 进行这样一个测量当然不会改变主系统的状态. 在采用下标"P"来代表主系统的同时，受控非门就可被展开为

$$U = |0_P0_E\rangle\langle 0_P0_E| + |0_P1_E\rangle\langle 0_P1_E| + |1_P1_E\rangle\langle 1_P0_E| + |1_P0_E\rangle\langle 1_P1_E| \tag{8.24}$$

因此

$$E_0 = \langle 0_E|U|0_E\rangle = |0_P\rangle\langle 0_P| \tag{8.25}$$

$$E_1 = \langle 1_E|U|0_E\rangle = |1_P\rangle\langle 1_P| \tag{8.26}$$

从而有

$$\varepsilon(\rho) = E_0\rho E_0 + E_1\rho E_1 \tag{8.27}$$

这是跟式(8.7)相一致的.

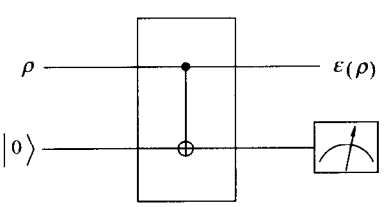


图 8.5 作为单量子比特测量的基本模型的受控非门.

2. 测量与算子和表示

给定开放量子系统的一个描述，我们如何来确定其动力学过程的算子和表示呢？已经找到的一个答案为：给定酉系统-环境变换运算U以及环境的一组基状态$|e_k\rangle$，那么运算元就为

$$E_k \equiv \langle e_k|U|e_0\rangle \tag{8.28}$$

在容许获得有关量子状态的信息的同时，利用允许在酉交互作用后在复合的系

统-环境上执行测量的可能性，有可能甚至进一步来推广这个结果. 由此导出，这种物理上的可能性与非保迹量子运算存在自然联系，也即映射 $\varepsilon(\rho)=\sum_k E_k\rho E_k^\dagger$ 使得 $\sum_k E_k^\dagger E_k \leqslant I$.

设主系统初始处于状态 ρ，为方便起见，我们用字母 Q 表示主系统，连同 Q 一起的是环境系统 E. 我们假设，Q 和 E 为初始独立的系统，且 E 启动于某个标准状态 σ，因此系统的联合状态初始为

$$\rho^{QE}=\rho\otimes\sigma \tag{8.29}$$

我们假设，系统是依据某个酉交互 U 来交互作用的. 在酉交互后，由 P_m 所描述的投影测量就在联合系统上被执行. 不进行测量的情形对应于一种特殊的情形，在此情形中只有对应于投影算子 $P_0\equiv I$ 的单个测量结果 $m=0$.

这种情况概括于图 8.6. 我们的目标是要确定以初始状态 ρ 为函数的 Q 的最终状态. 给定所出现的测量结果 m，QE 的最终状态为

$$\frac{P_mU(\rho\otimes\sigma)U^\dagger P_m}{\mathrm{tr}(P_mU(\rho\otimes\sigma)U^\dagger P_m)} \tag{8.30}$$

对环境 E 取迹，则单独 Q 的最终状态为

$$\frac{\mathrm{tr}_E(P_mU(\rho\otimes\sigma)U^\dagger P_m)}{\mathrm{tr}(P_mU(\rho\otimes\sigma)U^\dagger P_m)} \tag{8.31}$$

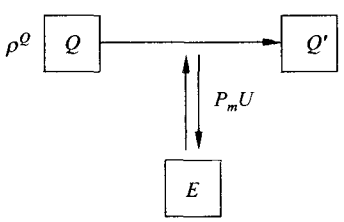


图 8.6　一个量子运算的环境模型.

这个最终状态的表示包含环境的初始状态 σ、交互作用 U 和测量算子 P_m. 定义一个映射

$$\varepsilon_m(\rho)\equiv \mathrm{tr}_E(P_mU(\rho\otimes\sigma)U^\dagger P_m) \tag{8.32}$$

于是单独 Q 的最终状态为 $\varepsilon_m(\rho)/\mathrm{tr}(\varepsilon_m(\rho))$. 注意到，$\mathrm{tr}[\varepsilon_m(\rho)]$ 是出现测量结果为 m 的概率. 令 $\sigma=\sum_j q_j|j\rangle\langle j|$ 为对 σ 的一个系综分解，对系统 E 引入一个标准正交基 $|e_k\rangle$，注意到

$$\varepsilon_m(\rho)=\sum_{jk} q_j\mathrm{tr}_E(|e_k\rangle\langle e_k|P_mU(\rho\otimes|j\rangle\langle j|)U^\dagger P_m|e_k\rangle\langle e_k|) \tag{8.33}$$

$$=\sum_{jk} E_{jk}\rho E_{jk}^\dagger \tag{8.34}$$

其中

$$E_{jk}\equiv\sqrt{q_j}\langle e_k|P_mU|j\rangle \tag{8.35}$$

该方程是式(8.10)的一个推广. 它在给定 E 的初始状态 σ 为已知和 Q 与 E 之间动力学过程为已知的条件下，为计算出现在 ε_m 的算子和表示中的那些算子提供了一

种显式的方法. 量子运算 ε_m 可被认为是定义了一类由推广第 2 章(见《量子计算和量子信息(一)》)中所给出的测量的描述所得到的测量过程.

练习 8.7 假设,代替在复合的主系统和环境上作投影测量,我们已经执行了由测量算子 $\{M_m\}$ 所描述的一般测量. 试对主系统上的相应量子运算 ε_m 来找出算子和表示,并证明相应的测量概率为 $\mathrm{tr}[\varepsilon(\rho)]$.

3. 任意算子和表示的系统-环境模型

我们已经证明,交互的量子系统可采用某种自然的方式得到量子运算的算子和表示. 相反的问题如何呢? 给定一组算子 $\{E_k\}$,是否存在某种合理的模型的环境系统和动力学过程,来得到具有这些运算元的量子运算? 所谓"合理"我们指的是,这个动力学过程必须为一个酉演化或为一个投影测量. 这里,我们要说明如何来构造这样一个模型. 我们将仅证明对映射输入空间到相同输出空间的量子运算如何来做到这一点,尽管将构造推广到更为一般情形主要是符号上的事情. 特别是,我们要证明,对任一具有运算元 $\{E_k\}$ 的保迹或非保迹的量子运算 ε,必存在启动于纯态 $|e_0\rangle$ 的一个模型环境 E,以及由酉算子 U 和 E 上的投影算子 P 所表征的模型动力学过程,使得有

$$\varepsilon(\rho) = \mathrm{tr}_E(PU(\rho\otimes|e_0\rangle\langle e_0|)U^\dagger P) \tag{8.36}$$

为了看清这一点,首先设 ε 为保迹量子运算,具有由满足完备性关系 $\sum_k E_k^\dagger E_k = I$ 的算子元 $\{E_k\}$ 所生成的算子和表示,于是我们只是试图来找到一个合适的酉算子 U 以对动力学过程建模. 令 $|e_k\rangle$ 为 E 的一个标准正交基组,且与算子 E_k 的下标指数 k 一一对应. 注意到,根据定义,E 具有这样的一个基;我们试图要来找到可提供得到由算子元 $\{E_k\}$ 所描述的动力学过程的模型环境. 定义一个算子 U,它对形如 $|\psi\rangle|e_0\rangle$ 的状态具有如下作用:

$$U|\psi\rangle|e_0\rangle \equiv \sum_k E_k|\psi\rangle|e_k\rangle \tag{8.37}$$

其中,$|e_0\rangle$ 只是这个模型环境的某个标准状态. 注意到,对于主系统的任意状态 $|\psi\rangle$ 和 $|\varphi\rangle$,根据完备性关系,有

$$\langle\psi|\langle e_0|U^\dagger U|\varphi\rangle|e_0\rangle = \sum_k \langle\psi|E_k^\dagger E_k|\varphi\rangle = \langle\psi|\varphi\rangle \tag{8.38}$$

因此,算子 U 可被扩展为作用于联合系统的整个状态空间的酉算子. 容易验证

$$\mathrm{tr}_E(U(\rho\otimes|e_0\rangle\langle e_0|)U^\dagger) = \sum_k E_k\rho E_k^\dagger \tag{8.39}$$

所以,这个模型提供了具有算子元 $\{E_k\}$ 的量子运算 ε 的一个实现. 这个结果说明示于盒子 8.1 中.

采用同样思路的构造(练习 8.8),非保迹量子运算也可容易地被建模. 这个构

造的更有意义的推广是对应于从测量得到的可能结果的一组量子运算$\{\varepsilon_m\}$的情况，由于对所有可能的输入ρ两两相异结果的概率总和为 1，即 $1=\sum_m p(m)=\mathrm{tr}\left[\left(\sum_m \varepsilon_m\right)(\rho)\right]$，所以量子运算 $\sum_m \varepsilon_m$ 是保迹的. 参看下面的练习 8.9.

练习 8.8（非保迹量子运算）　试解释如何对一个非保迹量子运算的系统-环境模型构造酉算子. 可以通过将外部算子 E_∞ 引入到选定的运算元组 E_k，且 E_k 满足包括 $k=\infty$ 在内的整个 k 范围内求和时得到 $\sum_k E_k^\dagger E_k = I$.

练习 8.9（测量模型）　如果我们被给定一组量子运算$\{\varepsilon_m\}$使得 $\sum_m \varepsilon_m$ 是保迹的，那么有可能来构造一个测量模型以得到这组量子运算. 对每个 m，令 E_{mk} 为 ε_m 的一组运算元. 现引入环境系统 E，其标准正交基$|m,k\rangle$一一对应于运算元的下标. 类似于先前的构造，定义一个算子 U，使得有

$$U|\psi\rangle|e_0\rangle=\sum_{mk}E_{mk}|\psi\rangle|m,k\rangle \tag{8.40}$$

下一步，定义环境系统 E 上的投影算子 $P_m\equiv\sum_k|m,k\rangle\langle m,k|$. 试证明，在$\rho\otimes|e_0\rangle\langle e_0|$上执行 U，则测量 P_m 将以概率 $\mathrm{tr}(\varepsilon_m(\rho))$给出 m，且主系统的相应的测量后状态必为 $\varepsilon_m(\rho)/\mathrm{tr}(\varepsilon_m(\rho))$.

盒子 8.1　模拟一个量子运算

给定以算子和表示 $\varepsilon(\rho)=\sum_k E_k\rho E_k^\dagger$ 表达的一个保迹量子运算，我们就可按照如下的方法来对其构造物理模型. 从式(8.10)，我们要求 U 满足

$$E_k=\langle e_k|U|e_0\rangle \tag{8.41}$$

其中，U 为某个酉算子，$|e_k\rangle$为环境系统的标准正交基向量. 这样一个 U 可以方便地在基$|e_k\rangle$上表示为分块矩阵：

$$U=\begin{bmatrix}[E_1] & \cdot & \cdot & \cdot & \cdots\\ [E_2] & \cdot & \cdot & \cdot & \cdots\\ [E_3] & \cdot & \cdot & \cdot & \cdots\\ [E_4] & \cdot & \cdot & \cdot & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots\end{bmatrix} \tag{8.42}$$

注意到，运算元 E_k 只能确定这个矩阵的第一个块列（不像其他地方，这里为了方便，状态的第一个标号为环境，第二个标号为主系统）. 这个矩阵的其余部分留给我们来确定；我们只需简单地选取元使 U 为是酉的. 注意，利用第 4 章的结果，U 可以用一个量子线路来实现.

8.2.4 量子运算的公理化方法

在此之前，我们研究量子运算的主要动机在于，量子运算为研究具有与环境交互的系统提供了一种优美的方法. 现在，我们打算转换到不同的观点，在此观点中我们试图写出期望量子运算应服从的来源于物理上的那些公理. 比之基于显式的系统-环境模型的先前方法，这种观点要更抽象，但也正是由于这种抽象性使这种方法非常强有力.

我们的打算如下. 首先，我们要忘掉已经学过的关于量子运算的所有一切，重新开始根据一组公理来定义量子运算，我们将会根据物理依据证明这些公理是合理的. 然后，我们将会证明映射 ε 满足这些公理当且仅当这个映射具有算子和表示，从而提供了抽象的公理化表示与我们先前的讨论之间所缺少的联系.

我们将量子运算 ε 定义为从输入空间 Q_1 的密度算子集合到输出空间 Q_2 的密度算子集合的一个映射，这个映射具有如下三个公理化性质(注意，为了证明中的符号简单起见，我们取 $Q_1=Q_2=Q$)：

(1) **A1**：第一，当 ρ 为初始状态时，$\mathrm{tr}[\varepsilon(\rho)]$ 为由 ε 表征的过程出现的概率，因此，对任意状态 ρ，$0\leqslant\mathrm{tr}[\varepsilon(\rho)]\leqslant 1$.

(2) **A2**：第二，ε 为密度矩阵集合上的一个凸线性映射，也即对概率 $\{p_i\}$，有

$$\varepsilon\left(\sum_i p_i\rho_i\right)=\sum_i p_i\varepsilon(\rho_i) \tag{8.43}$$

(3) **A3**：第三，ε 为完全正映射. 也即，如果 ε 将 Q_1 的密度算子映射到 Q_2 的密度算子，那么 $\varepsilon(A)$ 对任意半正定算子 A 必为是半正定的. 进而，如果我们引入一个任意维的附加系统 R，那么就必成立：$(\mathcal{I}\otimes\varepsilon)(A)$ 对组合系统 RQ_1 上的任意算子 A 必为是正的，其中 $\mathcal{I}$ 表示系统 R 上的单位映射.

第一个性质是为了数学上的方便性. 为了处理测量的情况，事实上如下约定非常方便，即 ε 不必要保持密度矩阵的迹性质 $\mathrm{tr}(\rho)=1$. 更恰当地说，我们要作出的约定是，ε 按这样的方式来定义，即 $\mathrm{tr}[\varepsilon(\rho)]$ 恰等于由 ε 描述的测量结果出现的概率. 举例来说，设我们在单量子比特的计算基上进行投影测量，那么，可以应用两个量子运算来描述由 $\varepsilon_0(\rho)\equiv|0\rangle\langle 0|\rho|0\rangle\langle 0|$ 和 $\varepsilon_1(\rho)\equiv|1\rangle\langle 1|\rho|1\rangle\langle 1|$ 所定义的过程. 注意到，两个各自结果的概率由 $\mathrm{tr}[\varepsilon_0(\rho)]$ 和 $\mathrm{tr}[\varepsilon_1(\rho)]$ 正确地给出. 利用这个约定，最后量子状态于是被正确地归一化为

$$\frac{\varepsilon(\rho)}{\mathrm{tr}[\varepsilon(\rho)]} \tag{8.44}$$

在过程为确定性的情况中，即在没有测量发生的情况中，这可简化为要求对所有的 ρ 成立 $\mathrm{tr}[\varepsilon(\rho)]=1=\mathrm{tr}(\rho)$. 如同先前讨论过的，在这种情况中，我们说量子运算是一个保迹量子运算，因为 ε 自身就提供了量子过程的一个完整描述. 另一方面，如果存在一个 ρ 使有 $\mathrm{tr}[\varepsilon(\rho)]<1$，那么量子运算是非保迹的，因为仅 ε 自身并不提供系统中可能出现的过程的完整描述(也即，其他测量结果也可能以某个概率出现). 一个物理上

的量子运算为这样的一种运算，它必满足概率从来不超过 1 即$\mathrm{tr}[\varepsilon(\rho)]\leqslant 1$的要求.

第二个性质起源于对量子运算的物理要求. 设加到量子运算的输入 ρ 是通过在量子状态的系综$\{p_i,\rho_i\}$中随机地选取一个状态而得到的，也即 $\rho=\sum_i p_i\rho_i$. 然后，我们期望，所得到的状态 $\varepsilon(\rho)/\mathrm{tr}[\varepsilon(\rho)]=\varepsilon(\rho)/p(\varepsilon)$ 对应于从系综$\{p(i\mid\varepsilon),\varepsilon(\rho_i)/\mathrm{tr}[\varepsilon(\rho_i)]\}$中的一个随机选取，其中 $p(i\mid\varepsilon)$是在 ε 表征的过程出现的前提下所制备的状态为 ρ_i 的概率. 因此，我们要求

$$\varepsilon(\rho)=p(\varepsilon)\sum_i p(i\mid\varepsilon)\frac{\varepsilon(\rho_i)}{\mathrm{tr}[\varepsilon(\rho_i)]} \tag{8.45}$$

其中，$p(\varepsilon)=\mathrm{tr}[\varepsilon(\rho)]$是由 ε 描述的过程出现在 ρ 的输入上的概率. 应用 Bayes 规则(《量子计算和量子信息(一)》附录 A)

$$p(i\mid\varepsilon)=p(\varepsilon\mid i)\frac{p_i}{p(\varepsilon)}=\frac{\mathrm{tr}[\varepsilon(\rho_i)]p_i}{p(\varepsilon)} \tag{8.46}$$

于是，方程(8.45)简化到式(8.43).

第三个性质也是起源于一个重要的物理要求. 它不仅要求，只要 ρ 是有效的，$\varepsilon(\rho)$就必须是有效的密度矩阵(除去归一化考虑)；而且进而，如果 $\rho=\rho_{RQ}$ 为 R 和 Q 的某个联合系统的密度矩阵，如果 ε 只作用于 Q，那么 $\varepsilon(\rho_{RQ})$必须仍然导出这个联合系统的一个有效的密度矩阵(除去归一化考虑). 盒子 8.2 中给出一个例子. 形式上，设我们引入一个另外一个(有限维的)系统 R. 令 $\mathcal{I}$ 表示系统 R 上的单位映射. 于是，映射 $\mathcal{I}\otimes\varepsilon$ 必须把半正定算子映为半正定算子.

盒子 8.2　完全正性与正性

单量子比特上的转置运算可提供一个例子，用以说明为什么完全正性对量子运算是一个重要的要求. 根据定义，这个映射在计算基中转置密度算子：

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\xrightarrow{T}\begin{bmatrix}a & c\\ b & d\end{bmatrix} \tag{8.47}$$

这个映射会保持单量子比特的半正定性. 但是，若设量子比特为一个初始处于纠缠状态

$$\frac{|00\rangle+\langle 11|}{\sqrt{2}} \tag{8.48}$$

的双量子比特系统的一部分，并将这双量子比特的第一个进行转置运算，而第二个量子比特服从于平凡的动力学过程. 那么，系统在动力学过程作用后的密度算子为

$$\frac{1}{2}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{8.49}$$

计算显示，这个算子具有特征值 1/2,1/2,1/2 和 −1/2，所以这不是一个合法的密度算子. 因此，转置运算就是正映射为不完全正的一个例子，也即，它可保持主系统上那些算子的半正定性，但当其作用于以主系统作为子系统的系统时就不再继续保持半正定性.

或许会令人惊讶，这三个公理对于定义量子运算已经是足够了. 下面的定理显示，这些公理等价于先前的系统-环境模型和根据算子和表示的定义.

定理 8.1 映射 ε 满足公理 **A1**，**A2** 和 **A3**，当且仅当对某个映射输入 Hilbert 空间到输出 Hilbert 空间算子集 $\{E_i\}$ 成立

$$\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger \tag{8.50}$$

且有 $\sum_i E_i^\dagger E_i \leqslant I$.

证 设 $\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger$. ε 显然是线性的，所以为检验 ε 是一个量子运算，我们只需证明它是完全正的. 令 A 为作用于增广系统 RQ 的状态空间上的任一半正定算子，令 $|\psi\rangle$ 为 RQ 的某个状态. 定义 $|\varphi_i\rangle \equiv (I_R \otimes E_i^\dagger)|\psi\rangle$，同时，应用算子 A 的正性，有

$$\langle\psi|(I_R \otimes E_i)A(I_R \otimes E_i^\dagger)|\psi\rangle = \langle|\varphi_i|A|\varphi_i\rangle \tag{8.51}$$

$$\langle\psi|(I_R \otimes E_i)A(I_R \otimes E_i^\dagger)|\psi\rangle \geqslant 0 \tag{8.52}$$

由此导出

$$\langle\psi|(\mathcal{I} \otimes \varepsilon)(A)|\psi\rangle = \sum_i \langle\varphi_i|A|\varphi_i\rangle \geqslant 0 \tag{8.53}$$

因此，对任一半正定算子 A，算子 $(\mathcal{I} \otimes \varepsilon)(A)$ 如所要求那样也是半正定的. 而要求 $\sum_i E_i^\dagger E_i \leqslant I$ 保证概率均小于或等于 1. 这就完成了第一部分的证明.

进而，设 ε 满足公理 **A1**，**A2** 和 **A3**. 我们的目标是对 ε 找到一个算子和表示. 设我们引入一个具有与原量子系统 Q 相同维数的系统 R. 令 $|i_R\rangle$ 和 $|i_Q\rangle$ 分别为 R 和 Q 的正交基. 为方便起见，对这两个基采用同一下标 i，这在 R 和 Q 具有相同的维数时肯定是可以做到的. 再定义系统 RQ 的联合状态 $|\alpha\rangle$ 为

$$|\alpha\rangle \equiv \sum_i |i_R\rangle|i_Q\rangle \tag{8.54}$$

这个状态 $|\alpha\rangle$ 除去归一化因子就为系统 R 和 Q 的一个最大纠缠状态. 把 $|\alpha\rangle$ 作为最大纠缠状态的这种解释对于理解下面的构造会是有帮助的. 进而，我们在 RQ 的状态空间上定义一个算子 σ 为

$$\sigma \equiv (\mathcal{I} \otimes \varepsilon)(|\alpha\rangle\langle\alpha|) \tag{8.55}$$

可以把它想象为是量子运算 ε 作用到系统 RQ 的最大纠缠状态的一半上的结果. 真正不同寻常的事实是，算子 σ 可用来完全表征量子运算 ε. 我们现在就来给出说明. 也即，为了弄清楚 ε 是如何作用于一个任意状态 Q 的，只要弄清楚 ε 是如何作用于与其他系统相连的 Q 的单个最大纠缠状态的.

允许我们从 σ 来恢复 ε 的技巧如下. 令 $|\psi\rangle = \sum_j \psi_j |j_Q\rangle$ 为系统 Q 的任一状态，再据等式

$$|\tilde{\psi}\rangle \equiv \sum_j \psi_j^* |j_R\rangle \tag{8.56}$$

来定义系统 R 的对应状态 $|\tilde{\psi}\rangle$. 注意到

$$\langle\tilde{\psi}|\sigma|\tilde{\psi}\rangle = \langle\tilde{\psi}|\left(\sum_{ij} |i_R\rangle\langle j_R| \otimes \varepsilon(|i_Q\rangle\langle j_Q|)\right)|\tilde{\psi}\rangle \tag{8.57}$$

$$= \sum_{ij} \psi_i \psi_j^* \varepsilon(|i_Q\rangle\langle j_Q|) \tag{8.58}$$

$$= \varepsilon(|\psi\rangle\langle\psi|) \tag{8.59}$$

令 $\sigma = \sum_i |s_i\rangle\langle s_i|$ 为 σ 的某个分解，其中向量 $|s_i\rangle$ 不未必为归一化的. 定义一个映射

$$E_i(|\psi\rangle) \equiv \langle\tilde{\psi}|s_i\rangle \tag{8.60}$$

稍加思考就可证明，这个映射是一个线性映射，所以 E_i 是 Q 的状态空间上的一个线性算子. 进而，我们有

$$\sum_i E_i |\psi\rangle\langle\psi| E_i^\dagger = \sum_i \langle\tilde{\psi}|s_i\rangle\langle s_i|\tilde{\psi}\rangle \tag{8.61}$$

$$= \langle\tilde{\psi}|\sigma|\tilde{\psi}\rangle \tag{8.62}$$

$$= \varepsilon(|\psi\rangle\langle\psi|) \tag{8.63}$$

因此，对 Q 的所有纯态 $|\psi\rangle$，

$$\varepsilon(|\psi\rangle\langle\psi|) = \sum_i E_i |\psi\rangle\langle\psi| E_i^\dagger \tag{8.64}$$

应用凸线性属性，一般可以导出

$$\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger \tag{8.65}$$

于是，由将 $\varepsilon(\rho)$ 的迹等同于概率的公理 **A1**，可以立即导出条件 $\sum_i E_i^\dagger E_i \leqslant I$. □

算子和表示的自由度

我们已经看到，算子和表示为开放量子系统的动力学过程提供了一个非常一般的描述. 那么，它是惟一的描述吗？

考虑作用于单量子比特上的量子运算 ε 和 $\mathcal{F}$，它具有算子和表示 $\varepsilon(\rho) = \sum_k E_k \rho E_k^\dagger$ 和 $\mathcal{F}(\rho) = \sum_k F_k \rho F_k^\dagger$，其中 ε 和 $\mathcal{F}$ 的运算元定义为

$$E_1 = \frac{I}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad E_2 = \frac{Z}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{8.66}$$

和

$$F_1 = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad F_2 = |1\rangle\langle 1| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{8.67}$$

这是一些看上去非常不同的量子运算. 有意思的是，ε 和 $\mathcal{F}$ 实际上是相同的量子运算. 为看清楚这一点，注意到 $F_1 = (E_1 + E_2)/\sqrt{2}$ 和 $F_2 = (E_1 - E_2)/\sqrt{2}$，因而，

$$\mathcal{F}(\rho)=\frac{(E_1+E_2)\rho(E_1^\dagger+E_2^\dagger)+(E_1-E_2)\rho(E_1^\dagger-E_2^\dagger)}{2} \tag{8.68}$$

$$=E_1\rho E_1^\dagger+E_2\rho E_2^\dagger \tag{8.69}$$

$$=\varepsilon(\rho) \tag{8.70}$$

这个例子显示，出现于一个量子运算的算子和表示中的运算元并不惟一.

算子和表示中的这种自由性是很有意思的. 设我们来掷一枚公平的硬币，并依据硬币掷得的结果，将酉算子 I 或 Z 作用于量子系统. 这个过程对应于 ε 的第一算子和表示. ε 的第二算子和表示(上面标记为 $\mathcal{F}$)对应于在$\{|0\rangle,|1\rangle\}$基上执行一个投影测量，其测量结果未知. 这两个明显很不相同的物理过程在主系统上引起完全相同的动力学过程.

两组运算元什么情况下给出同样的量子运算？理解这个问题是重要的，至少有两个理由. 第一，从物理的观点，理解算子和表示中的自由性让我们更加洞察到如何从不同的物理过程得到相同的系统动力学过程. 第二，理解算子和表示中的自由性，是更好理解量子纠错的关键.

直观上，算子和表示中必定存在很大的自由性，这一点是清楚的. 考虑描述如图 8.3所示那样系统的动力学过程的保迹量子运算 ε. 我们已经证明，对 ε 的运算元 $E_k=\langle e_k|U|e_k\rangle$可以相应于对环境的一个标准正交基$|e_k\rangle$. 设如图 8.7 所示那样，我们用仅在环境上的一个附加酉作用 U'补充交互作用 U. 很清楚，这样做不会改变主系统的状态. 对于这个新过程$(I\otimes U')U$ 对应的运算元是什么呢？我们得到

$$F_k=\langle e_k|(I\otimes U')U|e_k\rangle \tag{8.71}$$

$$=\sum_j[I\otimes\langle e_k|U'|e_j\rangle]\langle e_j|U|e_0\rangle \tag{8.72}$$

$$=\sum_j U'_{kj}E_j \tag{8.73}$$

其中，用到这样一个事实，$\sum_j|e_j\rangle\langle e_j|=I$ 和 U'_{kj} 为 U'的隶属于基$|e_k\rangle$的矩阵元. 由此推论，如同定理 8.2 中会证明的，由这种物理上激发的情景所得到的算子和表示中的自由性，抓住了算子和表示中可以利用的全部自由性的本质.

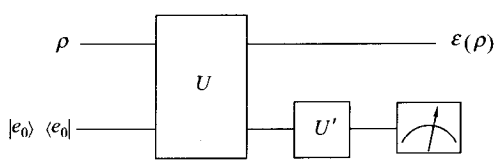


图 8.7 算子和表示中的酉自由性的原型.

定理 8.2(算子和表示中的酉自由性) 设$\{E_1,\cdots,E_m\}$和$\{F_1,\cdots,F_n\}$分别为量子运算 ε 和 $\mathcal{F}$ 的运算元，通过对较短那个运算元序列中添加零算子，我们可保证 $n=m$.

那么，$\varepsilon=\mathcal{F}$ 当且仅当存在复数 u_{ij} 使成立 $E_i=\sum_j u_{ij}F_j$，u_{ij} 为 m 乘 m 的酉矩阵.

证　证明关键是定理 2.6. 回顾先前，这个结果告诉我们，两个向量集 $|\psi_i\rangle$ 和 $|\varphi_i\rangle$ 生成同一算子，当且仅当

$$|\psi_i\rangle=\sum_j u_{ij}|\varphi_j\rangle \tag{8.74}$$

其中，u_{ij} 为复数的酉矩阵，且不管状态集 $|\psi_i\rangle$ 或 $|\varphi_i\rangle$ 更短，总可以通过“填补”附加的状态 0，以使两个集合具有相同数量的元. 这个结果允许我们可用来表征算子和表示中的自由性. 设 $\{E_i\}$ 和 $\{F_j\}$ 为对同一量子运算的两个运算元集合，$\sum_i E_i\rho E_i^\dagger=\sum_j F_j\rho F_j^\dagger$ 对所有 ρ 成立. 定义

$$|e_i\rangle\equiv\sum_k |k_R\rangle(E_i|k_Q\rangle) \tag{8.75}$$

$$|f_j\rangle\equiv\sum_k |k_R\rangle(F_j|k_Q\rangle) \tag{8.76}$$

回顾式(8.55)中 σ 的定义，由定义导出 $\sigma=\sum_i |e_i\rangle\langle e_i|=\sum_j |f_j\rangle\langle f_j|$，因此存在酉矩阵 u_{ij}，使下式成立：

$$|e_i\rangle=\sum_j u_{ij}|f_j\rangle \tag{8.77}$$

但是，对任意的 $|\psi\rangle$，有

$$E_i|\psi\rangle=\langle\tilde{\psi}|e_i\rangle \tag{8.78}$$

$$=\sum_j u_{ij}\langle\tilde{\psi}|f_j\rangle \tag{8.79}$$

$$=\sum_j u_{ij}F_j|\psi\rangle \tag{8.80}$$

从而

$$E_i=\sum_j u_{ij}F_j \tag{8.81}$$

反之，设 E_i 和 F_j 通过形为 $E_i=\sum_j u_{ij}F_j$ 的酉变换而相关联，则简单的代数运算显示，具有运算元 $\{E_i\}$ 的量子运算等同于具有运算元 $\{F_j\}$ 的量子运算. □

定理 8.2 可用于来回答另一个有意义的问题：对给定量子运算建模所需环境的最大范围是什么？

定理 8.3　一个维数 d 的 Hilbert 空间系统上的所有量子运算 ε 都可由包含最多 d^2 个元的一个算子和表示来生成：

$$\varepsilon(\rho)=\sum_{k=1}^{M} E_k\rho E_k^\dagger \tag{8.82}$$

其中，$1\leqslant M\leqslant d^2$.

这个定理的证明比较简单，留给读者作为一个练习.

练习 8.10 按如下方式，基于算子和表示中的自由性，给出定理 8.3 的证明. 令 $\{E_i\}$ 为 ϵ 的一组运算元，定义矩阵 $W_{jk} \equiv \mathrm{tr}(E_j^\dagger E_k)$，证明矩阵 W 是 Hermite(埃尔米特)矩阵且秩最大为 d^2，并因此存在酉矩阵 u 使 $uWu^\dagger$ 为最多有 d^2 个非零元的对角阵. 应用 u 来定义 ϵ 的一组最多有 d^2 个的非零运算元.

练习 8.11 设 ϵ 为将 d 维输入空间映射到 d' 维输出空间的一个量子运算，试证明，ϵ 可用一组最多 dd' 个运算元 $\{E_i\}$ 来描述.

算子和表示中的自由性有着令人惊讶的用途. 举例来说，我们将在第 10 章的量子纠错的研究中用到这一点. 在第 10 章中，我们将会看到，算子和表示中的某些算子集可以提供关于量子纠错过程的更为有用的信息，从而适合我们从那种观点来研究量子纠错. 如通常那样，以多种途径理解一个过程，会使我们对所从事的事情有更多的见识.

8.3 量子噪声和量子运算的例子

这节中，我们要来考察量子噪声和量子运算的一些具体例子. 这些模型说明，我们描述的量子运算体系是强有力的. 在理解噪声对量子系统的实际影响，以及理解如何采用如纠错的技术来控制噪声等，量子运算体系也是重要的.

在 8.3.1 节中，我们以考虑测量如何可被描述为一个量子运算作为开始，特别是要考虑迹运算和偏迹运算. 然后，我们将转到噪声过程，并以对理解单量子比特上量子运算的图方法的介绍作为 8.3.2 节的开端. 这种方法将被用于本节的剩余部分，以说明基本的比特翻转和相位翻转过程(8.3.3 节)、去极化信道(8.3.4 节)、幅值阻尼(8.3.5 节)以及相位阻尼(8.3.6 节)等. 幅值阻尼和相位阻尼是噪声的理想模型，这些模型能抓住出现于量子力学系统中噪声的许多重要特性，我们不仅要考虑它们的抽象的数学体系，而且要考虑这种过程在现实世界量子系统中是如何产生的.

8.3.1 迹与偏迹

量子运算体系的主要用途之一是描述测量的影响. 量子运算可被用来描述由量子系统上的测量得到特定结果的概率和由测量引起的系统中的状态的改变.

与测量相关的最简单的运算是迹映射 $\rho \rightarrow \mathrm{tr}(\rho)$. 我们可以按如下方式证明它事实上就是一个量子运算. 令 H_Q 为由标准正交基 $|1\rangle \cdots |d\rangle$ 张成的任一输入 Hilbert 空间，再令 H'_Q 为由状态 $|0\rangle$ 张成的一维输出空间，定义

$$\epsilon(\rho) \equiv \sum_{i=1}^{d} |0\rangle\langle i|\rho|i\rangle\langle 0| \tag{8.83}$$

使得，应用定理 8.1，ϵ 为一个量子运算. 注意到 $\epsilon(\rho) = \mathrm{tr}(\rho)|0\rangle\langle 0|$，于是，除去不

重要的倍数$|0\rangle\langle 0|$，这个量子运算等同于迹函数.

更为有用的一个结果是，认识到偏迹是一个量子运算. 设我们有一个联合系统QR，并希望对系统R取迹. 令$|j\rangle$为系统R的一个基，用

$$E_i\left(\sum_j \lambda_j |q_j\rangle |j\rangle\right) \equiv \lambda_i |q_i\rangle \tag{8.84}$$

来定义一个线性算子E_i：$H_{QR} \to H_Q$，其中，λ_j为复数，$|q_j\rangle$为系统Q的任意状态. 定义ε为具有运算元$\{E_i\}$的量子运算，也即

$$\varepsilon(\rho) \equiv \sum_i E_i \rho E_i^\dagger \tag{8.85}$$

应用定理 8.1，可得到这是一个从系统QR到系统Q的量子运算. 注意到

$$\varepsilon(\rho \otimes |j\rangle\langle j'|) = \rho \delta_{j,j'} = \mathrm{tr}_R(\rho \otimes |j\rangle\langle j'|) \tag{8.86}$$

其中，ρ为系统Q的状态空间上的任一 Hermite 算子；$|j\rangle$和$|j'\rangle$为系统R的标准正交基的成员. 利用ε和tr_R的线性属性，就可导出$\varepsilon = \mathrm{tr}_R$.

8.3.2　单量子比特量子运算的几何图像

有一种优美的几何方法可来图示单量子比特上的量子运算. 这种方法允许人们根据 Bloch 球面上量子运算的作用获得对量子运算行为的直觉感觉. 回顾练习 2.72，单量子比特的状态总是可被写成为 Bloch 表示：

$$\rho = \frac{I + \boldsymbol{r} \cdot \boldsymbol{\sigma}}{2} \tag{8.87}$$

其中，$\boldsymbol{r}$为三元实向量. 显式地，这就提供我们：

$$\rho = \frac{1}{2}\begin{bmatrix} 1 + r_z & r_x - \mathrm{i}r_y \\ r_x + \mathrm{i}r_y & 1 - r_z \end{bmatrix} \tag{8.88}$$

在这个表示式中，可以导出，一个任意的保迹量子运算等价于形如下式的一个映射：

$$\boldsymbol{r} \xrightarrow{\varepsilon} \boldsymbol{r}' = M\boldsymbol{r} + \boldsymbol{c} \tag{8.89}$$

其中，M为 3×3 实矩阵；$\boldsymbol{c}$为常向量. 这是一个将 Bloch 球面映射到其自身的仿射映射. 为看清这一点，设把生成ε的算子和表示的算子E_i写成为形式：

$$E_i = \alpha_i I + \sum_{k=1}^{3} a_{ik}\sigma_k \tag{8.90}$$

那么，不难检验

$$M_{jk} = \sum_l \left[a_{lj}a_{lk}^* + a_{lj}^* a_{lk} + \left(|\alpha_l|^2 - \sum_p a_{lp}a_{lp}^*\right)\delta_{jk} + i\sum_p \varepsilon_{jkp}\left(\alpha_l a_{lp}^* - \alpha_l^* a_{lp}\right)\right] \tag{8.91}$$

$$c_k = 2i\sum_l \sum_{jp} \varepsilon_{jpk} a_{lj} a_{lp}^* \tag{8.92}$$

其中，我们已经应用完备性关系$\sum_i E_i^\dagger E_i = I$以简化$\boldsymbol{c}$的表达式.

仿射映射(affine map)等式(8.89)的含义可通过考虑矩阵 M 的极分解 $M=U|M|$ 而使之更为清楚,其中 U 为酉矩阵. 因为 M 是实的,可以导出 $|M|$ 是实的和 Hermite 型的,也即 $|M|$ 是一个对称矩阵. 进而,由于 M 是实的,我们可以假定 U 是实的,且因此为一个正交矩阵,也即 $U^{\mathrm{T}}U=I$,其中 T 为转置运算. 于是,我们可以写为

$$M = OS \tag{8.93}$$

其中,O 是行列式为 1 的实正交矩阵,表示一个适当的旋转;S 是实对称矩阵. 按这种方式来看,式(8.89)恰好就是 Bloch 球面沿着由 S 确定的主轴的一种变形,然后加上适当的旋转 O,再加上位移 $\boldsymbol{c}$.

练习 8.12 为什么我们在分解式(8.89)中可以假定 O 具有行列式值 1?

练习 8.13 试证明,酉变换对应于 Bloch 球面的旋转.

练习 8.14 试证明,$\det(S)$ 未必为正.

8.3.3 比特翻转和相位翻转信道

上面所描述的几何图像可用来对单量子比特上的某些重要量子运算进行可视化,在后面它们将被用于纠错理论中. 比特翻转信道将量子比特的状态以概率 $1-p$ 从 $|0\rangle$ 翻转到 $|1\rangle$(或者相反),它具有运算元

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},$$

$$E_1 = \sqrt{1-p}X = \sqrt{1-p}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \tag{8.94}$$

比特翻转信道的作用如图 8.8 所示.

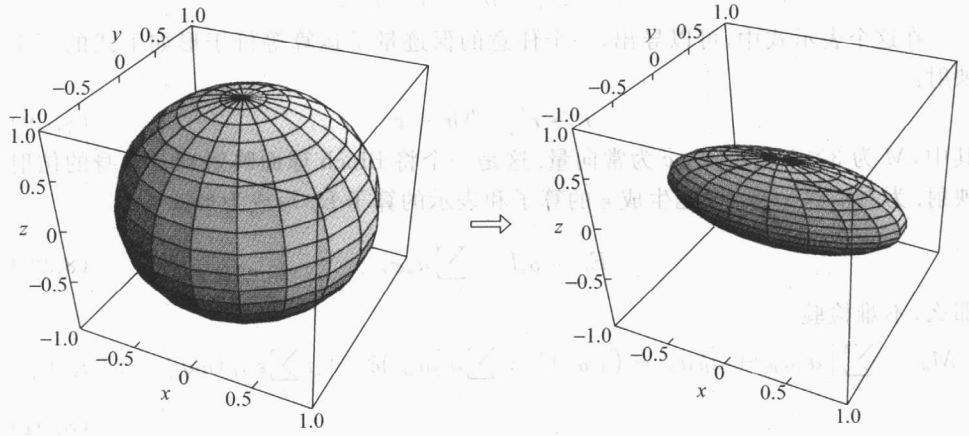


图 8.8 $p=0.3$ 时比特翻转信道对 Bloch 球面的作用. 左边的球面表示所有纯态的集合,右边的变形球面表示通过这个信道后的所有状态. 注意,$\hat{x}$ 轴上的状态被保留,而 $\hat{y}$-$\hat{z}$ 平面均匀地收缩一个 $1-2p$ 因子.

这种几何图像使之非常容易地来验证关于这个量子运算的某些事实. 举例来说，容易验证，单量子比特的量 $\mathrm{tr}(\rho^2)$ 等于 $(1+|r|^2)/2$，其中 $|r|$ 为 Bloch 向量的范数. 示于图 8.8 上的 Bloch 球面的收缩不会增加 Bloch 向量的范数，因此我们可以立刻得到结论，对比特翻转信道，$\mathrm{tr}(\rho^2)$ 只是减小. 这只是几何图像应用的一个例子；一旦对几何图像充分熟悉，它就会成为对有关单量子比特上量子运算性质认识的一个很大源泉.

相位翻转信道具有运算元

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},$$

$$E_1 = \sqrt{1-p}Z = \sqrt{1-p}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{8.95}$$

相位翻转信道的作用说明于图 8.9. 作为相位翻转信道的一种特殊情况，考虑当取 $p=1/2$ 时得到的量子运算. 应用算子和表示中的自由性，这个运算可被写成为

$$\rho \to \varepsilon(\rho) = P_0\rho P_0 + P_1\rho P_1 \tag{8.96}$$

其中，$P_0 = |0\rangle\langle 0|$ 和 $P_1 = |1\rangle\langle 1|$，对应于单量子比特在 $|0\rangle, |1\rangle$ 基上的一个测量，而测量的结果未知. 采用上面的规定，容易看到，Bloch 球面上的对应映射可为

$$(r_x, r_y, r_z) \to (0, 0, r_z) \tag{8.97}$$

几何上，Bloch 向量被沿着 z 轴投影，而 Bloch 向量在 x 和 y 方向上的分量丢失了.

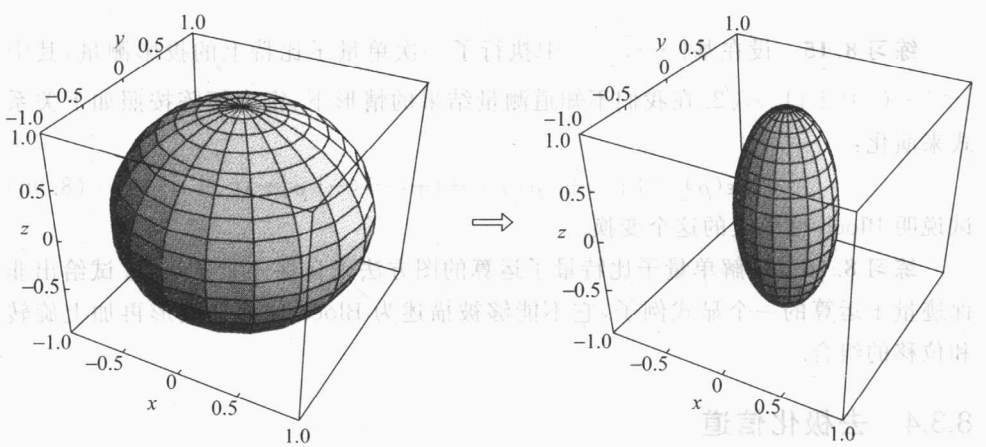


图 8.9　$p=0.3$ 时相位翻转信道对 Bloch 球面的作用. 注意，$\hat{z}$ 轴上的状态被保留，而 $\hat{x}$-$\hat{y}$ 均匀地收缩一个 $1-2p$ 因子.

比特-相位翻转信道具有运算元

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$E_1 = \sqrt{1-p}Y = \sqrt{1-p}\begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \tag{8.98}$$

如同其名称所指明的那样，这是一个相位翻转和一个比特翻转的组合，因为 $Y = \mathrm{i}XZ$. 比特-相位翻转信道的作用说明于图 8.10.

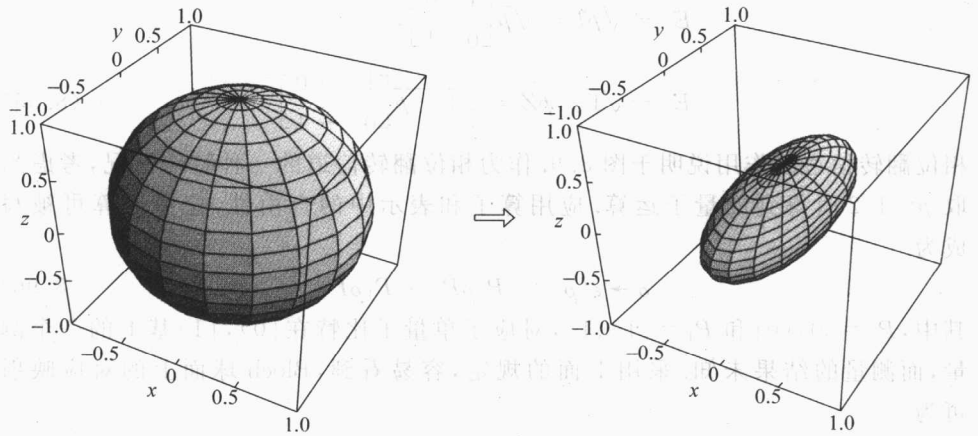


图 8.10 $p=0.3$ 时比特-相位翻转信道对 Bloch 球面的作用. 注意，$\hat{y}$ 轴上的状态被保留，而 $\hat{x}$-$\hat{z}$ 平面均匀地收缩一个 $1-2p$ 因子.

练习 8.15 设在基 $|+\rangle, |-\rangle$ 中执行了一次单量子比特上的投影测量，其中 $|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2}$. 在我们不知道测量结果的情形下，密度矩阵按照如下关系式来演化：

$$\rho \to \varepsilon(\rho) = |+\rangle\langle+|\,\rho\,|+\rangle\langle+| + |-\rangle\langle-|\,\rho\,|-\rangle\langle-| \tag{8.99}$$

试说明 Bloch 球面上的这个变换.

练习 8.16 理解单量子比特量子运算的图方法来自保迹量子运算. 试给出非保迹量子运算的一个显式例子，它不能够被描述为 Bloch 球面的变形再加上旋转和位移的组合.

8.3.4 去极化信道

去极化信道是一类重要的量子噪声. 想象我们取一个单量子比特，以概率 p 使量子比特去极化. 也即，单量子比特被完全混合态 $I/2$ 所替代. 单量子比特以概率 $1-p$ 保持不变. 量子系统在这个噪声后的状态为

$$\varepsilon(\rho)=\frac{pI}{2}+(1-p)\rho \tag{8.100}$$

去极化信道对 Bloch 球面的作用说明于图 8.11.

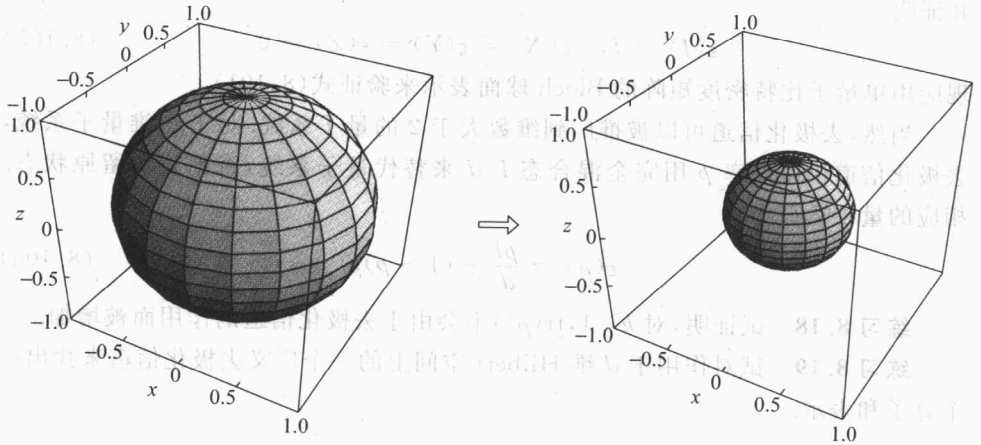


图 8.11　$p=0.5$ 时去极化信道对 Bloch 球面的作用. 注意，整个面以一个 p 的函数均匀地收缩.

一个模拟去极化信道的量子线路说明于图 8.12. 这个线路顶端的线是去极化信道的输入，而底部两根线是模拟信道的环境. 我们这里已经用到具有两个混合态输入的一个环境. 其思想是，第三量子比特作为控制，决定存储在第二量子比特中的完全混合状态 $I/2$ 是否会被交换到第一量子比特中去，第三量子比特初始是状态 $|0\rangle$ 和状态 $|1\rangle$ 的混合，状态 $|0\rangle$ 的概率为 $1-p$，状态 $|1\rangle$ 的概率为 p.

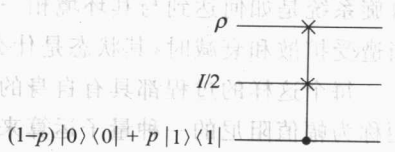


图 8.12　去极化信道的线路实现.

式(8.100)并不具有算子和表示的形式，但是，如果我们看到，对任意 ρ 有：

$$\frac{I}{2}=\frac{\rho+X\rho X+Y\rho Y+Z\rho Z}{4} \tag{8.101}$$

然后将 $I/2$ 关系式代入式(8.100)，就得到

$$\varepsilon(\rho)=\left(1-\frac{3p}{4}\right)+\frac{p}{4}(X\rho X+Y\rho Y+Z\rho Z) \tag{8.102}$$

这就显示，去极化信道具有运算元 $\{\sqrt{1-3p/4}\,I,\sqrt{p}X/2,\sqrt{p}Y/2,\sqrt{p}Z/2\}$. 注意，顺便提及，常可方便地以不同途径将去极化信道参数化，例如

$$\varepsilon(\rho)=(1-p)\rho+\frac{p}{3}(X\rho X+Y\rho Y+Z\rho Z) \tag{8.103}$$

这个关系式可解释为，状态 ρ 以概率 $1-p$ 被保留，算子 X，Y 和 Z 各以概率 $p/3$ 而作用其上.

练习 8.17 用如下方式来验证式(8.101). 定义

$$\varepsilon(A) \equiv \frac{A + XAX + YAY + ZAZ}{4} \tag{8.104}$$

并证明

$$\varepsilon(I) = I, \quad \varepsilon(X) = \varepsilon(Y) = \varepsilon(Z) = 0 \tag{8.105}$$

现应用单量子比特密度矩阵的 Bloch 球面表示来验证式(8.101).

当然,去极化信道可以被推广到维数大于 2 的量子系统. 对于 d 维量子系统,去极化信道仍以概率 p 用完全混合态 I/d 来替代量子系统,否则就保留原状态. 相应的量子运算为

$$\varepsilon(\rho) = \frac{pI}{d} + (1-p)\rho \tag{8.106}$$

练习 8.18 试证明,对 $k \geqslant 1$, $\mathrm{tr}(\rho^k)$ 不会由于去极化信道的作用而被增加.

练习 8.19 试对作用于 d 维 Hilbert 空间上的一个广义去极化信道来找出一个算子和表示.

8.3.5 幅值阻尼

量子运算的一个重要应用是描述能量耗散——一个由于能量从量子系统中失去的效应. 一个原子自发地发射一个光子时的动力学过程是什么? 高温下的一个自旋系统是如何达到与其环境相一致的平衡状态的? 干涉计或空腔中的一个光子当遭受扩散和衰减时,其状态是什么?

每个这样的过程都具有自身的惟一特性,但是所有这些过程的一般行为都可用称为幅值阻尼的一种量子运算来很好地表征,这一点我们可以通过考虑如下的方案来导出. 设有一个具有量子状态 $a|0\rangle + b|1\rangle$ 的单光模式,即零个或一个光子的叠加. 通过想象在光子的路径中插入一个部分镀银的镜子即分束器,就可以对这个模式的光子进行散射建模. 如同我们在 7.4.2 节中所见过的,这个分束器允许此光子与另一个单光模式(代表环境的)耦合,其方式是按照酉变换 $B = \exp[\theta(a^\dagger b - ab^\dagger)]$,其中 $a, a^\dagger$ 和 $b, b^\dagger$ 为光子在两个模式的湮没和产生算子. 分束器后的输出,在假定环境不引起任何光子的情形下,可应用式(7.34)简单地定出为 $B|0\rangle(a|0\rangle + b|1\rangle) = a|00\rangle + b(\cos\theta|01\rangle + \sin\theta|10\rangle)$. 对环境取迹给出量子运算:

$$\varepsilon_{\mathrm{AD}}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \tag{8.107}$$

其中,$E_k = \langle k|B|0\rangle$ 为

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}$$

$$E_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix} \tag{8.108}$$

它们就是幅值阻尼的运算元. 而 $\gamma=\sin^2\theta$ 可认为是丢失一个光子的概率.

注意到,不可能有 E_0 和 E_1 的线性组合来给出正比于单位阵的运算元(不妨与练习 8.23 相比较). E_1 运算把 $|1\rangle$ 状态变到 $|0\rangle$ 状态,对应于丢失一个能量量子到环境的物理过程. E_0 运算保持 $|0\rangle$ 状态不变,但会减少 $|1\rangle$ 状态的幅值;物理上,这种情况的发生是由于一个能量量子没有丢失到环境中,因而环境现时所感知到的更像是系统处于 $|0\rangle$ 状态而不是 $|1\rangle$ 状态.

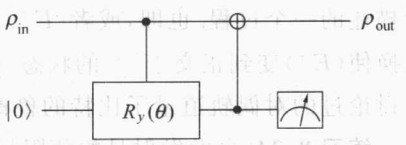


图 8.13　幅值阻尼的线路模型.

练习 8.20(幅值阻尼的线路模型)　试证明,图 8.13 中的线路是对以 $\sin^2(\theta/2)=\gamma$ 表示的幅值阻尼量子运算的建模.

练习 8.21(谐波振荡器的幅值阻尼)　设我们的主系统,一个谐波振荡器,通过 Hamilton 量

$$H=\chi(a^\dagger b+b^\dagger a) \tag{8.109}$$

与建模为另一谐波振荡器的环境之间产生交互作用,其中 a 和 b 分别为如 7.3 节所定义的谐波振荡器的湮没算子.

(1) 应用 $U=\exp(-\mathrm{i}H\Delta t)$,并表示 $b^\dagger b$ 的本征态为 $|k_b\rangle$,同时选取真空状态 $|0_b\rangle$ 为环境的初态,试证明运算元 $E_k=\langle k_b|U|0_b\rangle$ 为

$$E_k=\sum_n\sqrt{\binom{n}{k}}\sqrt{(1-\gamma)^{n-k}\gamma^k}\,|n-k\rangle\langle n| \tag{8.110}$$

其中,$\gamma=1-\cos^2(\chi\Delta t)$ 为丢失单个量子能量的概率,而形如 $|n\rangle$ 的状态为 $a^\dagger a$ 的本征态.

(2) 试证明运算元 E_k 可定义一个保迹量子运算.

练习 8.22(单量子比特密度矩阵的幅值阻尼)　对一般的单量子比特状态

$$\rho=\begin{bmatrix}a & b\\ b^* & c\end{bmatrix} \tag{8.111}$$

试证明幅值阻尼可导致

$$\varepsilon_{AD}(\rho)=\begin{bmatrix}1-(1-\gamma)(1-a) & b\sqrt{1-\gamma}\\ b^*\sqrt{1-\gamma} & c(1-\gamma)\end{bmatrix} \tag{8.112}$$

练习 8.23(对偶轨道量子比特的幅值阻尼)　设一个单量子比特状态可以用两个量子比特来表示:

$$|\psi\rangle=a\,|01\rangle+b\,|10\rangle \tag{8.113}$$

试证明,作用于这个状态的 $\varepsilon_{AD}\otimes\varepsilon_{AD}$ 给出由运算元

$$E_0^{\mathrm{dr}} = \sqrt{1-\gamma}I \tag{8.114}$$

$$E_1^{\mathrm{dr}} = \sqrt{\gamma}[\,|00\rangle\langle 01| + |00\rangle\langle 10|\,] \tag{8.115}$$

所描述的一个过程. 也即,或者(E_0^{dr})对量子比特什么也没有发生,或者量子比特被变换使(E_1^{dr})变到正交于$|\psi\rangle$的状态$|00\rangle$. 这是一个简单的差错检测码,也是 7.4 节中讨论过的对偶轨道量子比特的鲁棒性的基础.

练习 8.24(自发发射是幅值阻尼) 如 7.6.1 节中描述过的那样,耦合到单模式电磁辐射的一个单原子经历了自发发射. 为了看清这个过程正好就是幅值阻尼,取由失调率为 $\delta=0$ 的 Jaynes-Cummings 交互作用导出的酉运算,即式(7.77),并给出由对场取迹所得到的量子运算.

量子运算的一个一般特征是在运算下状态的集合保持不变. 举例来说,我们已经看到过,相位翻转信道是如何保持 Bloch 球面的 $\hat{z}$ 轴不改变的;这对应于以任意概率 p 的形如 $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ 的那些状态. 在幅值阻尼的情况中,仅只有基态$|0\rangle$是保持不变的. 这是我们在环境建模时一个自然结论,环境起始于$|0\rangle$状态(就像处于零度).

什么量子运算可来描述有限温度下对环境的耗散作用呢? 这个称之为广义幅值阻尼的过程 ϵ_{GAD},可通过如下运算元,用单量子比特来定义:

$$E_0 = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \tag{8.116}$$

$$E_1 = \sqrt{p}\begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix} \tag{8.117}$$

$$E_2 = \sqrt{1-p}\begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{bmatrix} \tag{8.118}$$

$$E_3 = \sqrt{1-p}\begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix} \tag{8.119}$$

其中,定态

$$\rho_\infty = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix} \tag{8.120}$$

满足 $\epsilon_{\mathrm{GAD}}(\rho_\infty)=\rho_\infty$. 广义幅值阻尼可用来描述由自旋对其周围点阵的耦合引起的"T_1"弛豫过程,该点阵是通常处于温度远高于自旋温度的一个热平衡大系统. 这种情况与 NMR 量子计算相关,其中描述于盒子 8.3 中的 ϵ_{GAD} 的某些性质会变得很重要.

盒子 8.3　广义幅值阻尼和有效的纯状态

已经发现，7.7 节中所引入的“有效纯态”一词，在量子计算机的 NMR 实现中是很有用的. 这些状态的行为很像酉演化下的纯态和无痕迹观测量的测量. 它们的行为在量子运算下的表现如何呢？一般来说，非酉量子运算会破坏这些状态的有效性，但令人惊讶的是，它们在广义幅值阻尼下的表现是正常的.

考虑一个单量子比特有效纯态 $\rho=(1-p)I+(2p-1)|0\rangle\langle 0|$. 很清楚，作用于 $U\rho U^\dagger$ 的无痕迹测量的观测量所导致的结果，必正比于在纯态 $U|0\rangle\langle 0|U^\dagger$ 上导致的那些结果. 设 ρ 为 $\varepsilon_{\mathrm{GAD}}$ 的定态. 有意思的是，在这个情况中，

$$\varepsilon_{\mathrm{GAD}}U\rho U^\dagger=(1-p)I+(2p-1)\varepsilon_{\mathrm{AD}}(U\rho U^\dagger) \tag{8.121}$$

也即，在广义幅值阻尼下，一个有效的纯态可以保持这种特性，并且，ρ 的“纯”成分的行为表现类同于它在零温度下一个容器内的幅值阻尼过程.

练习 8.25　如果通过假定平衡状态下处于 $|0\rangle$ 和 $|1\rangle$ 状态的概率满足 Boltzmann(玻耳兹曼)分布，也即 $p_0=\mathrm{e}^{-E_0/k_{\mathrm{B}}T}/\mathcal{Z}$ 和 $p_1=\mathrm{e}^{-E_1/k_{\mathrm{B}}T}/\mathcal{Z}$，其中 E_0 为状态 $|0\rangle$ 的能量，E_1 为状态 $|1\rangle$ 的能量，而 $\mathcal{Z}=\mathrm{e}^{-E_0/k_{\mathrm{B}}T}+\mathrm{e}^{-E_1/k_{\mathrm{B}}T}$，我们定义一个量子比特的温度为 T，那么试问，状态 ρ_∞ 以什么温度来描述？

我们可以在 Bloch 表示中把幅值阻尼的作用看作 Bloch 向量变换：

$$(r_x,r_y,r_z)\to(r_x\sqrt{1-\gamma},r_y\sqrt{1-\gamma},\gamma+r_z(1-\gamma)) \tag{8.122}$$

如同经常的实际物理过程情况那样，当 γ 采用诸如 $1-\mathrm{e}^{-t/T_1}$（t 为时间，T_1 是表征过程的速度的某个常数）的时变函数来替代时，我们就可把幅值阻尼的作用想象为 Bloch 球上的一个流，它把单位球内的每个点移向一个位于北极的固定点 $|0\rangle$. 这个过程如图 8.14 所示.

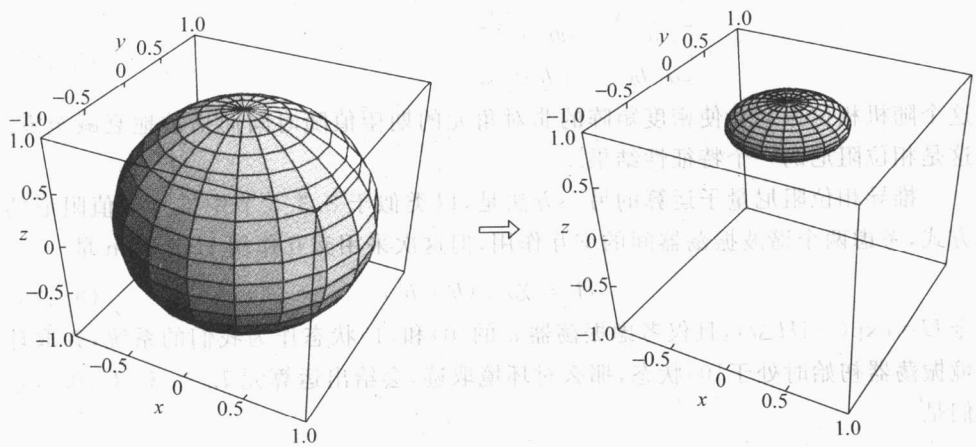


图 8.14　$p=0.8$ 时幅值阻尼信道对 Bloch 球面的作用. 注意，整个球面是如何朝向北极即状态 $|0\rangle$ 收缩的.

类似地，广义幅值阻尼执行变换：

$$(r_x, r_y, r_z) \to (r_x\sqrt{1-\gamma}, r_y\sqrt{1-\gamma}, \gamma(2p-1)+r_z(1-\gamma)) \quad (8.123)$$

比较式(8.122)和式(8.123)，可清楚地看出，幅值阻尼和广义幅值阻尼的差异仅在于流的固定点位置的不同；而最后状态是一个沿着 $\hat{z}$ 轴的、在点$(2p-1)$上的混合态.

8.3.6 相位阻尼

相位阻尼是纯粹量子力学性质的噪声过程，它描述没有能量损失下的量子信息的丢失. 相位阻尼在物理上可描述，举例来说，当一个光子通过波导传播发生的随机散射的情形，或者一个原子的电状态与远处电荷的交互作用发生摄动的情形. 一个量子系统的能量本征态不会像时间函数那样随时间改变，但确实会积累一个正比于特征值的相位. 当系统演化了一段时间后，有关这个量子相位的部分信息——能量本征态之间的相对相位——就会被丢失.

这类量子噪声的一个很简单的模型如下. 设我们有一个量子比特 $|\psi\rangle = a|0\rangle + b|1\rangle$，其上作用旋转运算 $R_z(\theta)$，其中旋转角 θ 是随机的. 举例来说，这个随机性可以是由与环境的确定性交互作用所引起的，而环境从不反过来与系统产生交互作用，因此是隐含地被测量的(参看 4.4 节). 我们将称这个随机的 R_z 运算为相位振动(phase kick). 假定，相位振动角 θ 可很好地表示为具有均值为 0 以及方差为 2λ 的高斯分布的一个随机变量.

这个过程的输出状态由在 θ 上取平均所得到的密度矩阵来给出：

$$\rho = \frac{1}{\sqrt{4\pi\lambda}}\int_{-\infty}^{\infty} R_z(\theta)\,|\psi\rangle\langle\psi|\,R_z^{\dagger}(\theta)\mathrm{e}^{-\theta^2/4\lambda}\mathrm{d}\theta \quad (8.124)$$

$$= \begin{bmatrix} |a|^2 & ab^*\mathrm{e}^{-\lambda} \\ a^*b\mathrm{e}^{-\lambda} & |b|^2 \end{bmatrix} \quad (8.125)$$

这个随机相位振动会使密度矩阵的非对角元的期望值随时间而指数地衰减至零. 这是相位阻尼的一个特征性结果.

推导相位阻尼量子运算的另一方法是，以类似于 8.3.5 节中推导幅值阻尼的方式，考虑两个谐波振荡器间的交互作用，但这次采用交互作用 Hamilton 量

$$H = \chi a^{\dagger}a(b+b^{\dagger}) \quad (8.126)$$

令 $U=\exp(-\mathrm{i}H\Delta t)$，且仅考虑振荡器 a 的$|0\rangle$和$|1\rangle$状态作为我们的系统，并取环境振荡器初始时处于$|0\rangle$状态，那么对环境取迹，会给出运算元 $E_k=\langle k_b|U|0_b\rangle$，它们是

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} \quad (8.127)$$

$$E_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{bmatrix} \tag{8.128}$$

其中，$\lambda = 1 - \cos^2(\chi \Delta t)$可以解释为来自系统的一个光子没有能量损失的散射的概率. 如像幅值阻尼的情况那样，E_0 会保持$|0\rangle$状态不改变，但会减小$|1\rangle$状态的幅值；然而，不像幅值阻尼，E_1 运算会破坏$|0\rangle$状态，并会减小$|1\rangle$状态的幅值但不会将其改变为$|0\rangle$状态.

应用定理 8.2 即量子运算的酉自由性，我们会发现，E_0 和 E_1 重新进行酉组合可对相位阻尼给出一组新的运算元：

$$\tilde{E}_0 = \sqrt{\alpha} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{8.129}$$

$$\tilde{E}_1 = \sqrt{1-\alpha} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{8.130}$$

其中，$\alpha = (1+\sqrt{1-\lambda})/2$. 因此，相位阻尼量子运算准确地等同于我们在 8.3.3 节中考虑过的相位翻转信道.

由于相位阻尼等同于相位翻转信道，故我们已经看到如何将其在图 8.9 的 Bloch 球上可视化. 这对应于 Bloch 向量变换：

$$(r_x, r_y, r_z) \rightarrow (r_x\sqrt{1-\lambda}, r_y\sqrt{1-\lambda}, r_z) \tag{8.131}$$

这个变换具有将球压缩为椭球的作用. 由于历史的原因，相位阻尼经常被称作为“T_2”(或自旋-自旋)弛豫过程，其中 $e^{-t/2T_2} = \sqrt{1-\lambda}$. 作为时间的函数，阻尼的值是随时间而递增的，对应于单位球内所有点朝向 $\hat{z}$ 轴的一个向内的流. 注意，沿着 $\hat{z}$ 轴的状态保持不变.

历史上，相位阻尼曾经几乎总是在物理上被视为由随机相位振动或散射过程所引起的一种过程. 直到它同相位翻转信道的联系被发现之前，量子纠错一直得不到发展，这是因为，总认为相位差错是连续的因而不能被描述为一个离散过程. 事实上，单量子比特相位差错总可以被视为是从一个特定的过程所导出的，其中或者量子比特以概率 α 什么都没有发生，或者量子比特以概率 $1-\alpha$ 通过 Pauli 运算 Z 被翻转. 尽管这可能并不是实际发生的微观物理学过程，但根据一个量子比特在一个离散时间区间上的变换要比内在的随机过程大的观点，这根本没有任何差别.

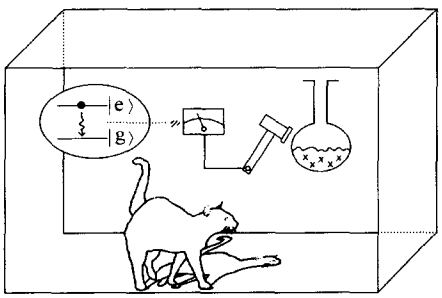
相位阻尼是量子计算和量子信息研究中最为微妙和重要的过程之一. 相位阻尼已成为大量研究和思考的一个课题，特别是关于为什么我们周围的世界具有根本不属于我们日常体验的叠加态，但看起来却如此经典. 也许，正是相位阻尼负责从我们每一天中去掉叠加态(练习 8.31)？先驱量子物理学家薛定谔(Schrödinger)或许是第一个提出这个问题的人，他以特别刻板的形式提出了这一点，如盒子 8.4 中所讨论的那样.

盒子 8.4 Schrödinger 猫

当我一听到 Schrödinger(薛定谔)猫时,我就会伸手取我的枪.

——*StephenHawkin*

Schrödinger 的声名狼藉的猫,由一个自动装置所左右,面临着生或死的可能事件.如同下图中的说明,如果一个被激发的原子状态被观测到衰减,自动装置会打破毒药瓶并把猫毒死.



Schrödinger 问,当原子位于叠加态时将会发生什么?这只猫是活还是死?为什么像这样的叠加态在每一天的世界中显然是不会出现的?这个难题只能通过认识到这在现实生活中是非常不大可能出现来加以解决,因为肉眼可见的叠加态对退相干极其敏感.设原子代表一个单量子比特.联合系统具有初态$|活\rangle|1\rangle$.设在原子的一个半衰期后,状态等于均匀叠加态$|活\rangle(|0\rangle+|1\rangle)/\sqrt{2}$(这代表了真实物理现象的简化,真实物理现象对这里处理的情况未免太复杂了些).如果原子处于$|0\rangle$状态,器械就会杀死猫;否则的话,猫就活着.这就给出状态$|\psi\rangle=[|死\rangle|0\rangle+|活\rangle|1\rangle]/\sqrt{2}$,在此状态下这只猫的状态已变为与原子的状态相纠缠.这看来是表明,猫同时可为活和死,但假设我们考虑这个状态的密度矩阵为

$$\rho=|\psi\rangle\langle\psi| \tag{8.132}$$

$$=\frac{1}{2}[|活,1\rangle\langle活,1|+|死,0\rangle\langle死,0|+|活,1\rangle\langle死,0|+|死,0\rangle\langle活,1|] \tag{8.133}$$

现在,要将猫和原子在它们的箱子中做到完全隔离,这在实际上是不可能的.因此,关于这个叠加态的信息将会泄漏到外部世界中.举例来说,来自猫身体的热量会透过墙体,并向外部世界提供其状态的某种指示.这种作用可以被建模为相位阻尼,它会指数地衰减到ρ中的最后两(非对角)项.作为直接的近似,我们可以将猫-原子系统建模为一个简谐振荡器.关于这样一个系统退相干的一个重要结果是,高能级差的状态之间的相干性衰减快于具有低能级差的状态之间的相干性的衰减(练习 8.31).因此,ρ会很快地被变换为一个接近于对角的状态,它代表猫-原子状态的系综(ensemble),相应于或为活或为死,但不是两个状态的叠加.

练习 8.26(相位阻尼的线路模型)　试证明,选取适当的 θ 后,图 8.15 中的线路可以被用来建模相位阻尼量子运算.

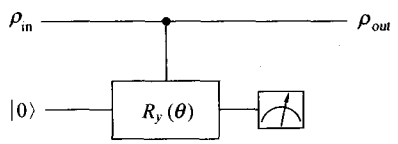


图 8.15　相位阻尼的线路模型.上部连线传送具有未知状态的输入量子比特,下部连线是一个辅助量子比特用于建模环境.

练习 8.27(相位阻尼=相位翻转信道)　给出将式(8.127)和式(8.128)的运算元同式(8.131)和式(8.131)的运算元联系起来的酉变换.也即,寻找一个 u,使有 $E_k = \sum_j u_{kj} E_j$.

练习 8.28(受控非门相位阻尼模型线路)　试证明,单受控非门可被用于作为相位阻尼的模型,如果令环境的初态为一个混合态,其中阻尼的值由状态混合的概率来确定.

练习 8.29(酉性)　一个量子过程 ε 是酉的,如果 $\varepsilon(I)=I$.试证明,去极化信道和相位阻尼信道是酉的,而幅值阻尼信道不是酉的.

练习 8.30($T_2 \leqslant T_1/2$)　相位相干弛豫率 T_2 正是量子比特密度矩阵中非对角元的指数衰减率,而 T_1 则为量子比特密度矩阵中对角元的指数衰减率(见式(7.144)).幅值阻尼同时具有非零的 T_1 和 T_2 率,证明对幅值阻尼有 $T_2 = T_1/2$.再证明:如果幅值阻尼和相位阻尼两者同时作用,那么必有 $T_2 \leqslant T_1/2$.

练习 8.31(相位阻尼的指数灵敏度)　应用式(8.126),证明谐波振荡器的密度矩阵中的元 $\rho_{nm} = \langle n|\rho|m\rangle$,在相位阻尼的作用下,对某个常数 λ 按指数 $e^{-\lambda(n-m)^2}$ 衰减.

8.4　量子运算的应用

作为一个强有力工具,量子运算体系有着众多的应用.本节中,我们要描述两种这类应用.8.4.1 节描述主方程(master equation)的理论,这是一种补充量子运算体系的量子噪声图像.主方程方法应用微分方程来描述连续时间中的量子噪声,它是物理学家们对量子噪声最为常用的方法.在 8.4.2 节中,我们要描述量子过程层析,这是一种试验确定量子系统动力学过程的方法.

8.4.1　主方程

开放量子系统出现在许多学科中,除了量子运算外的许多工具可以用于这些学科的研究中.这节中,我们要描述一个这样的工具,即主方程方法.

开放量子系统的动力学过程已在量子光学领域得到广泛研究. 这种背景中的主要目标是,应用恰当表征非酉行为的微分方程,以描述开放系统的时间演化. 这种描述是由主方程所提供的,应用 Lindblad 形式,它可以最一般地写成

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=-\frac{\mathrm{i}}{\hbar}[H,\rho]+\sum_j[2L_j\rho L_j^\dagger-\{L_j^\dagger L_j,\rho\}] \tag{8.134}$$

其中,$\{x,y\}=xy+yx$ 表示反对易式(anticommutator); H 为系统的 Hamilton 量,即表示动力学过程的相干部分的 Hamilton 算子; L_j 为表示系统耦合于其环境的 Lindblad 算子.

这个微分方程所以取为上面形式,是为了在类似于先前描述量子运算的意义下使过程为完全正的. 同样一般地假定,系统和环境是以一个积状态起始的. 进而,为了对一个过程推导主方程,人们通常以系统-环境模型 Hamilton 量作为起点,随后采用 Born 和 Markov 近似以确定 L_j. 注意,主方程方法中,在所有时刻均有 $\mathrm{tr}[\rho(t)]=1$.

作为 Lindblad 方程的一个例子,考虑经历自发发射的耦合到真空的一个双能级原子. 原子演化的相干部分由 Hamilton 量 $H=-\hbar\omega\sigma_z/2$ 来描述,$\hbar\omega$ 为原子能级的能级差. 自发发射,在过程中发射出一个光子同时,会引起使处于激发态($|1\rangle$)的一个原子降到基态($|0\rangle$). 这个发射由 Lindblad 算子 $\sqrt{\gamma}\sigma_-$ 来描述,其中 $\sigma_-\equiv|0\rangle\langle1|$ 为原子的下降算子(lowering operator),γ 为自发发射率. 描述这个过程的主方程为

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=-\frac{\mathrm{i}}{\hbar}[H,\rho]+\gamma[2\sigma_-\rho\sigma_+-\sigma_+\sigma_-\rho-\rho\sigma_+\sigma_-] \tag{8.135}$$

其中,$\sigma_+\equiv\sigma_-^\dagger$ 为原子的上升算子(raising operator).

为求解方程,有帮助的一个做法是转到交互作用的图像中,也即作变量的替换:

$$\tilde{\rho}\equiv\mathrm{e}^{\mathrm{i}Ht}\rho(t)\mathrm{e}^{-\mathrm{i}Ht} \tag{8.136}$$

容易得到 $\tilde{\rho}$ 的运动方程为

$$\frac{\mathrm{d}\tilde{\rho}}{\mathrm{d}t}=\gamma[2\tilde{\sigma}_-\tilde{\rho}\tilde{\sigma}_+-\tilde{\sigma}_+\tilde{\sigma}_-\tilde{\rho}-\tilde{\rho}\tilde{\sigma}_+\tilde{\sigma}_-] \tag{8.137}$$

其中

$$\tilde{\sigma}_-\equiv\mathrm{e}^{\mathrm{i}Ht}\sigma_-\mathrm{e}^{-\mathrm{i}Ht}=\mathrm{e}^{-\mathrm{i}\omega t}\sigma_- \tag{8.138}$$

$$\tilde{\sigma}_+\equiv\mathrm{e}^{\mathrm{i}Ht}\sigma_+\mathrm{e}^{-\mathrm{i}Ht}=\mathrm{e}^{\mathrm{i}\omega t}\sigma_+ \tag{8.139}$$

最后的运动方程为

$$\frac{\mathrm{d}\tilde{\rho}}{\mathrm{d}t}=\gamma[2\sigma_-\tilde{\rho}\sigma_+-\sigma_+\sigma_-\tilde{\rho}-\tilde{\rho}\sigma_+\sigma_-] \tag{8.140}$$

这个运动方程可应用 $\tilde{\rho}$ 的 Bloch 向量表示来容易地求解. 这个解就为

$$\lambda_x=\lambda_x(0)\mathrm{e}^{-\gamma t} \tag{8.141}$$

$$\lambda_y = \lambda_y(0)\mathrm{e}^{-\gamma t} \tag{8.142}$$

$$\lambda_z = \lambda_z(0)\mathrm{e}^{-2\gamma t} + 1 - \mathrm{e}^{-2\gamma t} \tag{8.143}$$

定义 $\gamma' = 1 - \exp(-2t\gamma)$，则我们可以容易地检验，这个演化等价于

$$\tilde{\rho}(t) = \varepsilon(\tilde{\rho}(0)) \equiv E_0\tilde{\rho}(0)E_0^\dagger + E_1\tilde{\rho}(0)E_1^\dagger \tag{8.144}$$

其中

$$E_0 \equiv \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma'} \end{bmatrix} \tag{8.145}$$

$$E_1 \equiv \begin{bmatrix} 0 & \sqrt{\gamma'} \\ 0 & 0 \end{bmatrix} \tag{8.146}$$

为定义量子运算 ε 的运算元. 注意，比较式(8.108)，ε 的作用为幅值阻尼. 我们考虑过的那个例子就是一个自旋-玻色子(s'pin-boson)模型的实例，在此例子中一个小的有穷维量子系统同容纳它的简谐振荡器交互作用. 物理上，这在描述原子与电磁辐射的交互作用中是重要的，如在腔 QED 或原子和离子阱中那样.

主方程方法不如量子运算体系方法更一般. 对主方程的求解可允许人们来确定密度矩阵的时间相关性. 了解这一点意味着，结果可以被表达为在算子和表示上的一个量子运算：

$$\rho(t) = \sum_k E_k(t)\rho(0)E_k^\dagger(t) \tag{8.147}$$

其中，$E_k(t)$ 为依赖于时间的运算元由主方程的解所确定. 但是，根据算子和表示描述的量子过程未必能被写为一个主方程. 举例来说，量子运算其所以可描述非 Markov 的动力学过程，只是因为它们只用来描述状态的改变而不是连续的时间演化. 然而，每一种方法都有它自己的位置. 事实上，甚至连量子运算都不能提供最一般的描述；我们在 8.5 节中将会考虑某些过程，它们是不能由量子运算来描述的.

8.4.2　量子过程层析

量子运算对开放量子系统提供了一个漂亮的数学模型，并且可以方便地被可视化(至少对量子比特如此)——但是，它们是如何与实验中可测量的量联系起来的呢？如果实验学家想要刻画一个量子系统的动力学过程，那么应当作什么测量呢？对于经典的系统，这个基本任务称为系统辨识. 这里，我们要说明，它的相似方法，即所谓的量子过程层析是如何对有穷维量子系统来执行的.

为了理解过程层析，我们首先需要了解称为量子状态层析的另一种方法. 状态层析是实验地确定未知量子状态的方法. 设给定了单量子比特的一个未知状态 ρ，如何通过实验确定这个状态 ρ？

如果只提供给我们 ρ 的一份单个备份，那么刻画 ρ 事实上是不可能的. 基本的问题是，没有一个量子测量可确定地来区别开像 $|0\rangle$ 和 $(|0\rangle + |1\rangle)/\sqrt{2}$ 那样的非正交量

子状态. 但是，如果我们拥有大量的 ρ 的备份，就有可能来估计 ρ. 例如，如果 ρ 是由某个实验所产生的量子状态，那么我们就可简单地将实验重复多次以产生多份状态 ρ.

设我们有单量子比特密度矩阵 ρ 的多份备份. 集合 $I/\sqrt{2}, X/\sqrt{2}, Y/\sqrt{2}, Z/\sqrt{2}$ 构成 Hilbert-Schmidt(希尔伯特-施密特)内积意义下的标准正交的矩阵集，所以 ρ 可表为

$$\rho = \frac{\mathrm{tr}(\rho) I + \mathrm{tr}(X\rho) X + \mathrm{tr}(Y\rho) Y + \mathrm{tr}(Z\rho) Z}{2} \tag{8.148}$$

回顾以前，像 $\mathrm{tr}(A\rho)$ 的表达式可解释为观测量的平均值. 例如，为估计 $\mathrm{tr}(Z\rho)$，我们来测量观测量 Z 很多次(如 m 次)，得到结果为 $z_1, z_2, \cdots, z_m$，它们都等于 $+1$ 或 -1. 这些量的实验平均值 $\sum_i z_i/m$ 就是对 $\mathrm{tr}(Z\rho)$ 的真值的一个估计. 我们可以应用中心极限定理来确定，对大的 m 这个估计的效果会有多好，这里估计近似地为具有等于 $\mathrm{tr}(Z\rho)$ 的均值和标准差为 $\triangle Z/\sqrt{m}$ 的高斯函数，其中 $\triangle Z$ 是 Z 的上界为 1 的单个测量值的标准差，所以我们的估计的标准差最大为 $1/\sqrt{m}$.

按类似的方式，我们能以高的置信度，通过对大量采样值取极限来估计 $\mathrm{tr}(X\rho)$ 和 $\mathrm{tr}(Y\rho)$，从而得到 ρ 的一个好的估计. 至少在原理上，不难将这种方法推广到多于一个量子比特的情况. 类似于单量子比特情况，n 个量子比特上的一个任意的密度矩阵可被表为

$$\rho = \sum_{\boldsymbol{v}} \frac{\mathrm{tr}(\sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n} \rho) \sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n}}{2^n} \tag{8.149}$$

其中，求和是在向量 $\boldsymbol{v} = (v_1, \cdots, v_n)$ 上作的元 v_i 取值于 0,1,2,3. 通过对 Pauli 矩阵乘积的观测量执行多次测量，我们就可估计这个和式中的每一项，从而得到 ρ 的一个估计.

我们已经描述了如何对由多量子比特组成的系统来作状态层析. 如果把非量子比特系统包括进去，那么又会什么样呢？毫不奇怪，容易将上面的做法推广到这样的系统. 我们在这里不会详细叙述这样的做法，读者可以参考本章结尾的“历史和进一步阅读材料”.

既然我们懂得如何做量子状态层析，我们又如何用它去做量子过程层析呢？这种实验方法可概述如下. 设系统的状态空间有 d 维；举例来说，对一个单量子比特为 $d=2$. 我们取定 d^2 个纯量子状态 $|\psi_1\rangle, \cdots, |\psi_{d^2}\rangle$，且选取得使相应的密度矩阵 $|\psi_1\rangle\langle\psi_1|, \cdots, |\psi_{d^2}\rangle\langle\psi_{d^2}|$ 构成矩阵空间的一个基组. 下面，将更详细地解释如何选取这样一个基组. 对每个状态 $|\psi_j\rangle$，将量子系统制备于那个状态，然后使其服从于我们希望表征的过程. 在过程已经运行完成以后，应用量子状态层析以确定来自过程输出的状态 $\varepsilon(|\psi_j\rangle\langle\psi_j|)$. 从纯粹理论的观点，现在算是做完了，因为原理上量子运算 ε 现在通过 ε 对所有状态的线性扩张而被确定.

实际上，我们当然更愿意有一种方法，能从实验可得到的数据中定出 ε 的一个有用的表示. 在对一个单量子比特的情况显式地设计出后，我们将会对这样做法来

说明一般的步骤. 目标是对 ε 确定一组运算元$\{E_i\}$,使有

$$\varepsilon(\rho) = \sum_i E_i \rho E_i^\dagger \tag{8.150}$$

然而,实验结果只包含数字而非算子,算子属于理论概念. 为了从可测量参数来确定 E_i,方便的做法是利用一组固定的算子 $\tilde{E}_i$ 来考虑 ε 的等价描述,$\tilde{E}_i$ 集合对状态空间上算子集合构成一个基,使对某个复数组 e_{im} 有

$$E_i = \sum_m e_{im} \tilde{E}_m \tag{8.151}$$

式(8.150)可因此重写成为

$$\varepsilon(\rho) = \sum_{mn} \tilde{E}_m \rho \tilde{E}_n^\dagger \chi_{mn} \tag{8.152}$$

其中,$\chi_{mn} \equiv \sum_i e_{im} e_{in}^*$ 为一个矩阵的元. 按定义此矩阵为半正定的 Hermite 矩阵. 这个称为 chi 矩阵表示的表达式表明,一旦这个算子 $\tilde{E}_i$ 的集合选定,ε 就可以用一个复数矩阵 χ 来完全描述.

一般而言,χ 将包含 $d^4 - d^2$ 个独立的实参数,因为从 $d \times d$ 复数矩阵到 $d \times d$ 矩阵的一般线性映射要用 d^4 个独立参数来描述,但由于 ρ 保持为具有迹 1 的 Hermite 阵,则共需 d^2 个附加约束; 也即,满足完备性关系:

$$\sum_i E_i^\dagger E_i = I \tag{8.153}$$

它给出了 d^2 个实约束. 我们将会说明如何实验地确定 χ,然后再说明一旦 χ 矩阵为已知,形如式(8.150)的算子和表示如何恢复.

令 ρ_j,$1 \leqslant j \leqslant d^2$ 为 $d \times d$ 矩阵空间的一个固定的线性独立基; 也即,任一 $d \times d$ 矩阵都可被写成为这组 ρ_j 的惟一的线性组合. 一个方便的选取为算子 $|n\rangle\langle m|$ 组. 实验上,通过制备输入状态 $|n\rangle$,$|m\rangle$,$|+\rangle = (|n\rangle + |m\rangle)/\sqrt{2}$ 和 $|-\rangle = (|n\rangle + \mathrm{i}|m\rangle)/\sqrt{2}$,并构成线性组合 $\varepsilon(|n\rangle\langle n|)$,$\varepsilon(|m\rangle\langle m|)$,$\varepsilon(|+\rangle\langle +|)$ 和 $\varepsilon(|-\rangle\langle -|)$,输出状态 $\varepsilon(|n\rangle\langle m|)$ 就可得到为

$$\varepsilon(|n\rangle\langle m|) = \varepsilon(|+\rangle\langle +|) + \mathrm{i}\varepsilon(|-\rangle\langle -|) - \frac{1+\mathrm{i}}{2}\varepsilon(|n\rangle\langle n|) - \frac{1+\mathrm{i}}{2}\varepsilon(|m\rangle\langle m|) \tag{8.154}$$

因此,对每个 ρ_j,都可以应用状态层析来确定 $\varepsilon(\rho_j)$.

进而,每个 $\varepsilon(\rho_j)$ 都可被表示为基状态的线性组合

$$\varepsilon(\rho_j) = \sum_k \lambda_{jk} \rho_k \tag{8.155}$$

并因为 $\varepsilon(\rho_j)$ 由状态层析获知,λ_{jk} 可应用标准的线性代数算法而确定. 为继续进行下去,可以写出

$$\tilde{E}_m \rho_j \tilde{E}_n^\dagger = \sum_k \beta_{jk}^{mn} \rho_k \tag{8.156}$$

其中，β_{jk}^{mn} 为复数，在给定 $\tilde{E}_m$ 算子和 ρ_j 算子后，它们可应用线性代数中标准算法而定出. 结合上述最后两个表达式和式(8.152)，有

$$\sum_k \sum_{mn} \chi_{mn} \beta_{jk}^{mn} \rho_k = \sum_k \lambda_{jk} \rho_k \tag{8.157}$$

由 ρ_k 的线性无关性，可导出对每个 k，有

$$\sum_{mn} \beta_{jk}^{mn} \chi_{mn} = \lambda_{jk} \tag{8.158}$$

这个关系式对于矩阵 χ 给出正确的量子运算 ϵ 是一个必要和充分的条件. 可以把 χ 和 λ 看作向量，β 看作以 mn 表列指数和以 jk 表行指数的 $d^4 \times d^4$ 矩阵. 为了说明 χ 是如何得到的，令 κ 为矩阵 β 的广义逆，并满足关系式

$$\beta_{jk}^{mn} = \sum_{st,xy} \beta_{jk}^{st} \kappa_{st}^{xy} \beta_{xy}^{mn} \tag{8.159}$$

大多数矩阵运算的计算机软件包都能提供这种广义逆计算. 我们现来证明，由

$$\chi_{mn} \equiv \sum_{jk} \kappa_{jk}^{mn} \lambda_{jk} \tag{8.160}$$

所定义的 χ 满足关系式(8.158).

验证由式(8.160)定义的 χ 满足式(8.158)的困难在于，一般说来，χ 不是由式(8.158)所惟一确定的. 为方便起见，我们以矩阵的形式重写这些方程为

$$\beta \vec{\chi} = \vec{\lambda} \tag{8.161}$$

$$\vec{\chi} \equiv \kappa \vec{\lambda} \tag{8.162}$$

我们从导致式(8.152)的构造中可知，对式(8.161)至少存在一个解，我们将其称为 $\vec{\chi}'$. 因此，$\vec{\lambda} = \beta \vec{\chi}'$. 广义逆满足 $\beta\kappa\beta = \beta$. 用 β 左乘 $\vec{\chi}$ 的定义式，给出

$$\beta \vec{\chi} = \beta \kappa \vec{\lambda} \tag{8.163}$$

$$= \beta \kappa \beta \vec{\chi}' \tag{8.164}$$

$$= \beta \vec{\chi}' \tag{8.165}$$

$$= \vec{\lambda} \tag{8.166}$$

因而，如同我们想要证明的，由式(8.162)所定义的 χ 满足式(8.161).

在定出 χ 的同时，人们可立刻以下面的方式来得到 ϵ 的算子和表示. 令酉矩阵 $U^{\dagger}$ 使 χ 对角化，

$$\chi_{mn} = \sum_{xy} U_{mx} d_x \delta_{xy} U_{ny}^* \tag{8.167}$$

基此，可以容易验证

$$E_i = \sqrt{d_i} \sum_j U_{ji} \tilde{E}_j \tag{8.168}$$

是 ϵ 的算子元. 我们的算法可以因此归纳如下：应用状态层析实验地定出 λ，通过方程 $\vec{\chi} = \kappa\lambda$ 转而由其定出 χ，它又可提供我们 ϵ 的一个完全的描述，包括一组运算元 E_i.

在单量子比特量子过程的情况中，只有 12 个参数必须加以确定（见盒子 8.5）. 但一个双量子比特量子黑箱 ϵ_2 的动力学过程将对我们的理解造成要大得多的挑战. 在这种情况中，多达 240 个参数需要被确定，以完全表征作用于量子系统上的量子运算. 确定这些参数显然是一件相当可观的任务. 但是，类似单量子比特的情况，可相当直观地来实现一种数字程序，这个程序能自动完成计算，实验所需的状态层析和状态制备操作可在实验室获得.

盒子 8.5　单量子比特的过程层析成像术

对单量子比特运算情况，过程层析的一般方法可被简化，以提供在实验条件下有用的显式表达. 通过选取具有对易性质的固定算子组 $\tilde{E}_i$，可使这种简化成为可能，对易性质能方便地允许应用直接的矩阵乘法来确定 χ 矩阵. 在单量子比特情况中，我们采用

$$\tilde{E}_0 = I \tag{8.169}$$

$$\tilde{E}_1 = X \tag{8.170}$$

$$\tilde{E}_2 = -\mathrm{i}Y \tag{8.171}$$

$$\tilde{E}_3 = Z \tag{8.172}$$

在对 χ 简化后，共有 12 个参数，它们可确定任意的单量子比特量子运算 ϵ.

这些参数可通过四组实验来测量. 作为一个特例，设将输入状态制备为 $|0\rangle, |1\rangle, |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ 和 $|-\rangle = (|0\rangle + \mathrm{i}|1\rangle)/\sqrt{2}$，并应用状态层析定出四个矩阵：

$$\rho_1' = \epsilon(|0\rangle\langle 0|) \tag{8.173}$$

$$\rho_4' = \epsilon(|1\rangle\langle 1|) \tag{8.174}$$

$$\rho_2' = \epsilon(|+\rangle\langle +|) - \mathrm{i}\epsilon(|-\rangle\langle -|) - (1-\mathrm{i})(\rho_1' + \rho_4')/2 \tag{8.175}$$

$$\rho_3' = \epsilon(|+\rangle\langle +|) + \mathrm{i}\epsilon(|-\rangle\langle -|) - (1+\mathrm{i})(\rho_1' + \rho_4')/2 \tag{8.176}$$

这些矩阵对应于 $\rho_j' = \epsilon(\rho_j)$，其中

$$\rho_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{8.177}$$

$\rho_2 = \rho_1 X$，$\rho_3 = X\rho_1$ 和 $\rho_4 = X\rho_1 X$. 从式(8.156)和式(8.169)～式(8.172)，我们可以确定 β，类似地 ρ_j' 可确定 λ. 但是，由于基的特定选择以及 $\tilde{E}_i$ 的 Pauli 矩阵表示，我们可以把 β 矩阵表为 Kronecker 积 $\beta = \Lambda \otimes \Lambda$，其中

$$\Lambda = \frac{1}{2}\begin{bmatrix} I & X \\ X & -I \end{bmatrix} \tag{8.178}$$

使得 χ 可方便地以分块矩阵形式表达为

$$\chi = \Lambda \begin{bmatrix} \rho_1' & \rho_2' \\ \rho_3' & \rho_4' \end{bmatrix} \Lambda \tag{8.179}$$

我们已经说明，对量子系统动力学过程的一个有用的描述，如何应用系统的步骤可通过实验来确定. 量子过程层析这种方法类似于经典控制理论中所形成的系统辨识步骤，并且它在理解和控制带噪声的量子系统中扮演着类似的角色.

练习 8.32 试解释，如何把量子过程层析推广到非保迹量子运算的情况，如出现于测量研究中的情况.

练习 8.33（表征量子过程） 设人们希望，通过描述 Bloch 球上的一个点集 $\{\vec{r}_k\}$ 在 ε 下的变换情况，以完全表征一个任意的单量子比特运算 ε. 试证明，这个集合至少包含有四个点.

练习 8.34（对双量子比特的过程层析） 试证明，描述两量子比特上的黑箱运算的 χ^2 可表为

$$\chi^2 = \Lambda_2 \vec{\rho}' \Lambda_2 \tag{8.180}$$

式中，$\Lambda_2 = \Lambda \otimes \Lambda$，$\Lambda$ 为如盒子 8.5 中所定义；ρ' 为 16 个测得的密度矩阵的分块矩阵

$$\vec{\rho}' = P^{\mathrm{T}} \begin{bmatrix} \rho'_{11} & \rho'_{12} & \rho'_{13} & \rho'_{14} \\ \rho'_{21} & \rho'_{22} & \rho'_{23} & \rho'_{24} \\ \rho'_{31} & \rho'_{32} & \rho'_{33} & \rho'_{34} \\ \rho'_{41} & \rho'_{42} & \rho'_{43} & \rho'_{44} \end{bmatrix} P \tag{8.181}$$

其中，$\rho'_{nm} = \varepsilon(\rho_{nm})$，$\rho_{nm} = T_n |00\rangle\langle 00| T_m$，$T_1 = I \otimes I$，$T_2 = I \otimes X$，$T_3 = X \otimes I$，$T_4 = X \otimes X$，而 $P = I \otimes [(\rho_{00} + \rho_{12} + \rho_{21} + \rho_{33}) \otimes I]$ 为置换阵.

练习 8.35（过程层析的例子） 考虑未知动力学过程 ε_1 的单量子比特黑箱. 设下述四个密度矩由根据式(8.173)～式(8.176)执行的实验测量所得到：

$$\rho'_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{8.182}$$

$$\rho'_2 = \begin{bmatrix} 0 & \sqrt{1-\gamma} \\ 0 & 0 \end{bmatrix} \tag{8.183}$$

$$\rho'_3 = \begin{bmatrix} 0 & 0 \\ \sqrt{1-\gamma} & 0 \end{bmatrix} \tag{8.184}$$

$$\rho'_4 = \begin{bmatrix} \gamma & 0 \\ 0 & 1-\gamma \end{bmatrix} \tag{8.185}$$

其中，γ 为一个数值参数. 从对每个这些输入-输出关系的独立研究中，人们可以作出一些重要的推测：基态 $|0\rangle$ 在 ε_1 下保持不变，激发态 $|1\rangle$ 部分地衰减至基态，而叠加态则被阻尼. 试确定这个过程的 χ 矩阵.

8.5　量子运算体系的局限性

是否存在我们感兴趣的量子系统，其动力学过程不是用量子运算所描述的？在本节中，我们要构造一个人为的例子，其系统的演化不是由量子运算所描述的，并试图了解这在何种情况下可能出现.

设一个单量比特位被制备于记作 ρ 的某个未知的量子状态. 这个量子比特的制备包括在制备这个量子比特的实验室中要完成的某些步骤. 设在这之中，实验室自由度为一个单量子比特，作为状态制备步骤的副产品，如果 ρ 为 Bloch 球的底半部上的状态，则其会保持在状态 $|0\rangle$，如果 ρ 为 Bloch 球的顶半部上的状态，则其保持在状态 $|1\rangle$. 也即，如果 ρ 是 Bloch 球的底半部上的状态，系统在制备后的状态就为

$$\rho\otimes|0\rangle\langle 0|\otimes \text{其他自由度} \tag{8.186}$$

如果 ρ 是 Bloch 球的顶半部上的状态，则为

$$\rho\otimes|1\rangle\langle 1|\otimes \text{其他自由度} \tag{8.187}$$

一旦状态制备完成，系统开始与环境产生交互作用，这就是在这种情况中的所有实验室自由度. 设这个交互作用是使一个受控非门被置于实验室系统中的主系统和外量子比特之间. 因此，如果系统的 Bloch 向量初始位于 Bloch 球的底半部，它就会在过程中保持不变；而如果系统的 Bloch 向量初始位于 Bloch 球的顶半部，则它会旋转到 Bloch 球的底半部.

显然，这个过程并不是作用于 Bloch 球上的一个仿射映射，因而根据 8.3.2 节中的结果，它不能作为一个量子运算. 从这个讨论中所学到的知识是，与用于制备那个系统的自由度产生交互作用的一个量子系统，在其准备完成以后，一般将会经历在量子运算体系内不能充分描述的一个动力学过程. 这得出了一个重要的结论，如其所指出的那样，必存在一些物理上合理的环境，在这些环境下量子运算体系不能充分地描述发生在一个量子系统中的过程. 这一点应当牢记在心，比如说，在 8.4 节讨论过的量子过程层析的应用中即是如此.

尽管如此，在本书的剩下部分中，我们仍将在量子运算体系内来讨论. 对于描述由量子系统经历的动力学过程，量子运算提供了一个强有力的和比较通用的工具. 更重要的，它提供了对量子信息处理相关问题获得具体进展的手段. 在超出量子运算体系范围以外，进一步来研究量子信息处理是一个饶有兴趣的问题.

问题 8.1（量子运算的 Lindblad 形）　在 8.4.1 节的注释中，显式地得到对 $\rho(t)$ 求解下述微分方程的步骤：

$$\dot{\rho} = -\frac{\lambda}{2}(\sigma_+\sigma_-\rho + \rho\sigma_+\sigma_- - 2\sigma_-\rho\sigma_+) \tag{8.188}$$

试表达映射 $\rho(0)\to\rho(t)$ 为 $\rho(t) = \sum_k E_k(t)\rho(t)E_k^\dagger(t)$.

问题 8.2(作为量子运算的隐形传态) 设 Alice 拥有记作系统 1 的单量子比特,她想要对它隐形传态(teleportation)给 Bob. 不幸的是,她和 Bob 仅分享一个不完美的纠缠量子比特对. 这个量子比特对中,Alice 的一半记为系统 2,Bob 的一半记为系统 3. 设 Alice 执行一次测量,它由一组量子运算 ε_m 所描述,且在系统 1 和系统 2 上具有结果 m. 试证明,这会导出将系统 1 的初态关联到系统 3 的终态的一个运算 $\tilde{\varepsilon}_m$. 再证明,若 Bob 能采用一个保迹量子运算 $\mathcal{R}_m$ 来逆转这个运算,得到

$$\mathcal{R}_m\left(\frac{\tilde{\varepsilon}_m(\rho)}{\mathrm{tr}[\tilde{\varepsilon}_m(\rho)]}\right) = \rho \tag{8.189}$$

则隐形传态就可以实现. 其中,ρ 为系统 1 的初始状态.

问题 8.3(随机酉信道) 一个诱人的想法是,使所有单位信道也即 $\varepsilon(I)=I$ 的那些信道,是由随机酉运算上取平均也即 $\varepsilon(\rho) = \sum_k p_k U_k \rho U_k^\dagger$ 来导出,其中 U_k 为酉算子,p_k 构成概率分布. 试证明,尽管这对单量子比特成立,但对更大的系统则不成立.

第 8 章小结 量子噪声和量子运算

- **算子和表示** 一个开放量子系统的行为可建模为

$$\varepsilon(\rho) = \sum_k E_k \rho E_k^\dagger \tag{8.190}$$

其中,如果量子运算是保迹的,那么 E_k 为满足 $\sum_k E_k^\dagger E_k = I$ 的运算元.

- **量子运算的环境模型** 一个保迹量子运算总是可看成为是从系统与一个初始不相关环境的酉交互作用中出现的,反之亦然. 非保迹量子运算可以类似地处理,除了在系统与环境的复合上作投影测量外,且对应不同的非保迹量子运算具有不同的结果.

- **量子过程的层析** 一个 d 维量子系统上的量子运算,通过实验地测量从 d^2 个纯态输入所产生的输出密度矩阵,而可以完全地被确定.

- **重要的单量子比特量子运算的运算元**

去极化信道

$$\sqrt{1-\frac{3p}{4}}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \quad \sqrt{\frac{p}{4}}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

$$\sqrt{\frac{p}{4}}\begin{bmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{bmatrix} \quad \sqrt{\frac{p}{4}}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

幅值阻尼	$\begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\gamma}\end{bmatrix}$	$\begin{bmatrix}0 & \sqrt{\gamma}\\ 0 & 0\end{bmatrix}$
相位阻尼	$\begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\gamma}\end{bmatrix}$	$\begin{bmatrix}0 & 0\\ 0 & \sqrt{\gamma}\end{bmatrix}$
相位翻转	$\sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	$\sqrt{1-p}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$
比特翻转	$\sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	$\sqrt{1-p}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$
比特 - 相位翻转	$\sqrt{p}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	$\sqrt{1-p}\begin{bmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{bmatrix}$

历史和进一步阅读的材料

量子噪声在好些领域中都是一个重要的课题，已有大量关于这个主题的参考文献. 我们必然只能限于引证关于这个课题可找到资源中的一小部分例子. 从相当的数学角度研究量子噪声的早期论文属于 Davies[Dav76]. Caldeira 和 Leggett[CL83] 采用基于 Feynman 路径积分的方法，最早的也是最为完整地研究了称为自旋-玻色子模型的一个重要模型. Gardiner[Gar91] 从量子光学的观点研究了量子噪声. 更为近期一点，量子光学界已经对量子噪声发展了称为量子轨道的方法. 这个主题的评述可见诸于 Zoller 和 Gardiner[ZG97] 以及 Plenio 和 Knight[PK98] 的文章.

关于量子运算主题已有大量的文献存在. 我们只会提到很少一部分关键性文献，首先是 Kraus 的书[Kra83]，此书包含了关于这个主题的非常早期工作的文献. 关于这个主题的有影响的早期论文，包括 Hellwig 和 Kraus[HK69, HK70] 以及 Choi[Cho75] 的那些论文. Lindblad[Lin76] 把量子运算形式化同连续时间量子演化的理论联系起来，同时引入现在称为 Lindblad 形的结果. Schumacher[Sch96b] 和 Caves[Gav99] 就量子运算体系，从量子纠错的观点，写出了非常出色的综述性材料.

量子状态层析是由 Vogel 和 Risken[VR89] 所建议的. Leonhardt[Leo97] 写了一篇近期评述，包括有对其他工作的多篇文献. 对量子过程层析的需要是在 Turchette, Hood, Lange, Mabuchi 和 Kimble[THL+95] 的论文中所指出的. 有关理论是由 Chuang 和 Nielson[CN97] 以及 Poyatos, Cirac 和 Zoller[PCZ97] 独立发展的. 而 Jones[Jon94] 较早地勾画出了量子过程层析的主要思想.

在退相干(decoherence)一词上存在着不幸的混乱局面. 历史上，这个词只是用于称呼相位阻尼过程，特别是 Zurek[Zur91]. Zurek 和其他研究者认识到，在从量子物理到经典物理的转换中，相位阻尼具有独特的作用；对某些环境耦合，相位阻

尼会出现在要远快于任何幅值阻尼过程的时间尺度上，相位阻尼因而在确定量子相干的丧失中会要重要得多. 这些研究的主要之点，由于环境的交互作用，已经成为这种经典性的体现. 但是，一般地，退相干在量子计算和量子信息的使用涉及到量子处理中任意噪声过程. 本书中，我们更喜欢更为一般的术语量子噪声，并且倾向于使用这个名词，尽管偶尔也会在文中的适当地方发现“退相干”一词.

量子运算体系的局限性的某些更为详细的讨论(以及特别是，系统和环境初始位于积状态的假定)是由 Royer[Roy96] 给出的.

问题 8.2 应归于 Nielson 和 Caves[NC97]. 问题 8.3 来源于 Landau 和 Streater[LS93] 深入研究双随机量子运算的凸集的极值点的一部分工作.

CHAPTER 9

第 9 章

量子信息的距离度量

两条信息相似是什么意思？信息通过某个过程而保持是什么意思？这些问题构成了量子信息处理理论的中心问题，本章的目的是要介绍可对这些问题提供定量答案的距离度量. 受到前面这两个问题的推动，我们将会关注两大类的距离度量，静态度量(static measure)和动态度量(dynamic measure). 静态度量可定量地表示两个量子状态有多近，而动态度量则可定量地表示在动态过程期间信息能多好地被保持. 我们所采取的策略是，以通过引入好的距离静态度量作为开端，然后再采用这些静态度量作为阐述距离的动态度量的基础.

不管是经典力学还是量子力学，在定义距离度量的方法中都存在有某种任意性，并且量子计算和量子信息界的研究者近几年中已经发现采用多样化的距离度量可带来方便. 这些度量中的两种，迹距离(trace distance)和保真度(fidelity)，现今使用得特别广泛. 我们在本章中会详细讨论这两种度量. 大多数情况中，两种距离度量的性质是很类似的；但是对于某些应用，一种度量可能会比另外一种度量更易于处理. 正是出于这一理由，并由于两种度量被广泛应用在量子计算和量子信息界，我们会对这两者都进行讨论.

9.1 经典信息的距离度量

区分概率分布的想法令人难以捉摸.

——Christopher Fuchs

让我们从可以容易地应用我们直觉知识的舞台——经典信息的距离度量——开始. 什么是经典信息理论中被比较的对象？我们来考虑比较如 00010 和 10011 这样的比特串. 一种定量表示这些比特串之间距离的方法是 Hamming 距离，它被定义为两个比特串间为不相等处的位所的个数. 举例来说，比特串 00010 和 10011 在第一个位所和最后一个位所处为不相同，所以它们之间的

Hamming 距离为 2. 不幸的是，虽然两个对象之间的 Hamming(汉明)距离只是简单地查一下标签的事，但是在量子力学的 Hilbert 空间舞台中并不先验地存在任何标签.

对量子信息的距离度量开展研究一个很好的起点是将其与经典的概率分布相比较. 事实上，在经典信息论中，信源通常被建模为随机变量，也即为某个源字母表上的概率分布. 举例来说，英文课本的一份未知源材料可以被建模为罗马字母表上的一个随机变量序列. 在阅读课本之前，我们可作的直接判断只能是如课本中各个字母出现的相对频率，以及它们之间的某种相关性，例如在英文课本中字母对“th”的出现要远普遍于字母对“zx”. 将信源表征为某个字母表上的概率分布这种表征，激励我们集中关注于距离度量搜寻中概率分布的比较.

说同一指标集 x 上的两个概率分布 $\{p_x\}$ 和 $\{q_x\}$ 为彼此类似是什么意思呢？很难对这样一个显然具有惟一正确答案的问题来给出解答，所以我们提出两种不同的答案，其中的每一个都已被广泛地为量子计算和量子信息界所应用. 第一种度量为迹距离，由如下方程来定义：

$$D(p_x,q_x)\equiv\frac{1}{2}\sum_x|p_x-q_x| \tag{9.1}$$

这个量有时被称为 L_1 距离或 Kolmogorov(柯尔莫哥洛夫)距离. 我们更喜欢迹距离这个名词，因为后面有这个量的量子力学相似版本. 迹距离事实上是关于概率分布的一种度量，(一个度量 $D(x,y)$ 必须是对称的，且满足三角不等式 $D(x,z)\leqslant D(x,y)+D(y,z)$)，所以采用名词距离是合适的.

练习 9.1 概率分布(1,0)和概率分布(1/2,1/2)之间的迹距离是什么？概率分布(1/2,1/3,1/6)和概率分布(3/4,1/8,1/8)之间的迹距离又是什么？

练习 9.2 试证明，概率分布 $(p,1-p)$ 和概率分布 $(q,1-q)$ 之间的迹距离为 $|p-q|$.

两个概率分布之间的第二种距离度量，即概率分布 $\{p_x\}$ 和 $\{q_x\}$ 的保真度，可用下式来定义：

$$F(p_x,q_x)\equiv\sum_x\sqrt{p_xq_x} \tag{9.2}$$

相比于迹距离，保真度是度量概率分布之间距离的一种很不同的方法. 首先，它并不是一个度量，尽管随后我们会讨论从保真度导出的一个度量. 为了看清保真度不是一个度量，注意到当分布 $\{p_x\}$ 和 $\{q_x\}$ 为相同时，有 $F(p_x,q_x)=\sum_x p_x=1$. 对保真度的更好的几何理解说明于图 9.1；保真度恰好就是位于单位球上的具有分量 $\sqrt{p_x}$ 和 $\sqrt{q_x}$ 的两个向量之间的内积.

练习 9.3 概率分布(1,0)和概率分布(1/2,1/2)之间的保真度是什么？概率

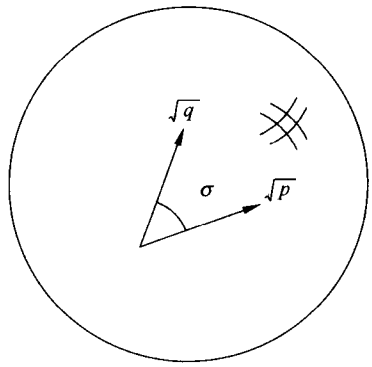


图 9.1　保真度在几何上解释为位于单位球上的向量 $\sqrt{p_x}$ 和 $\sqrt{q_x}$ 之间的内积(因为 $1=\sum_x(\sqrt{p_x})^2=\sum_x(\sqrt{q_x})^2$).

分布(1/2,1/3,1/6)和概率分布(3/4,1/8,1/8)之间的保真度又是什么?

迹距离和保真度都是数学上有用的手段,用来定义两个概率分布之间距离.这些度量是否有物理上的运算意义呢? 在迹距离的情况中,对这个问题的答案为"是".特别是,可简单地证明

$$D(p_x,q_x)=\max_S|\,p(S)-q(S)\,|=\max_S\Big|\sum_{x\in S}p_x-\sum_{x\in S}q_x\Big| \tag{9.3}$$

其中,极大取遍指标集$\{x\}$的所有子集 S.取极大值的量是事件 S 根据分布$\{p_x\}$出现的概率跟事件 S 根据分布$\{q_x\}$出现的概率之间的差别.事件 S 因而在某种意义上就是试图区分分布$\{p_x\}$和$\{q_x\}$时所要检验的最优事件,而迹距离则控制着进行这项区分的可能程度.

遗憾的是,对于保真度尚不知道有类似的清楚解释.但是,在 9.2 节中,我们会证明,在数学上,保真度是一个足够有用的量,即便其没有清楚的物理解释.此外,我们不能排除将来会发现保真度的清楚物理解释可能性.最后,可以导出,在保真度和迹距离之间存在着紧密的关系,因而一个量的性质常能被用来引出另一个量的一些性质.这是一个常常令人惊讶的有用事实.

练习 9.4　试证明式(9.3).

练习 9.5　试证明,绝对值符号可以从式(9.3)中去掉,也即

$$\begin{aligned}D(p_x,q_x)&=\max_S(p(S)-q(S))\\&=\max_S\Big(\sum_{x\in S}p_x-\sum_{x\in S}q_x\Big)\end{aligned} \tag{9.4}$$

迹距离和保真度是用来比较两个固定的概率分布的距离的静态度量.还可有距离的第三个概念,它是距离的一种动态度量用来度量信息可在多大程度上由某个物理过程所保持.设一个随机变量 X 经由带噪声信道发送出去,同时给出另一

个随机变量 Y 为输出，形成一个 Markov 过程 $X\to Y$. 为方便起见，我们假定 X 和 Y 两者具有相同的取值范围，表为 x. 那么，Y 不等于 X 的概率 $p(X\neq Y)$ 是一个用来表征信息能由过程所保持程度的明显和重要的度量.

令人惊讶的是，这个距离的动态度量可以被理解为静态迹距离的一种特殊情况. 想象，给定一个随机变量 X，且我们制作一份 X 的备份，并产生一个新的随机变量 $\widetilde{X}=X$. 如图 9.2 中指明的那样，随机变量 X 现在通过带噪声信道，留下作为输出的随机变量 Y. 初始完全相关的对 $(\widetilde{X},X)$ 会以多大程度接近最后的对 $(\widetilde{X},Y)$ 呢？采用迹距离作为我们的接近度的度量，再运用简单的代数运算，就得到

$$D((\widetilde{X},X),(X,Y))=\frac{1}{2}\sum_{xx'}\mid \delta_{xx'}p(X=x)-p(\widetilde{X}=x,Y=x')\mid \tag{9.5}$$

$$=\frac{1}{2}\sum_{x\neq x'}p(\widetilde{X}=x,Y=x')+\frac{1}{2}\sum_{x}\mid p(X=x)-p(\widetilde{X}=x,Y=x)\mid \tag{9.6}$$

$$=\frac{1}{2}\sum_{x\neq x'}p(\widetilde{X}=x,Y=x')+\frac{1}{2}\sum_{x}(p(X=x)-p(\widetilde{X}=x,Y=x)) \tag{9.7}$$

$$=\frac{p(\widetilde{X}\neq Y)+1-p(\widetilde{X}=Y)}{2} \tag{9.8}$$

$$=\frac{p(X\neq Y)+p(\widetilde{X}\neq Y)}{2} \tag{9.9}$$

$$=p(X\neq Y) \tag{9.10}$$

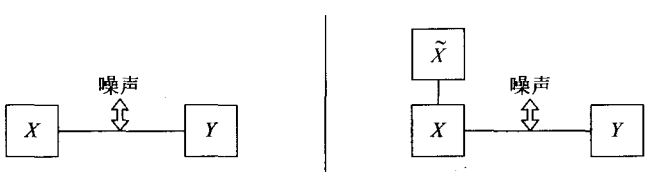


图 9.2 给定一个 Markov 过程 $X\to Y$，在 X 受到使其转化为 Y 的噪声作用之前，我们可首先制作一份 X 的备份 $\widetilde{X}$.

因此，如同图 9.3 中指明的那样，信道中出现一个差错的概率就等于 $(\widetilde{X},X)$ 的概率分布和 $(\widetilde{X},Y)$ 的概率分布之间的迹距离. 这是一个重要的解释，因为它将会成为类似量子解释的基础. 这一点之所以必要，是由于不存在概率 $p(X\neq Y)$ 的直接量子类似版本，因为量子力学中没有与存在于不同时间的变量 X 和 Y 的联合概率分布相类似的概念. 作为替代，我们采用与刚才给出的解释相类似的方法，并基于在量子信道的动力学过程期间重要的是保持量子纠缠而不是经典的相关思想，来定义量子距离的动态度量.

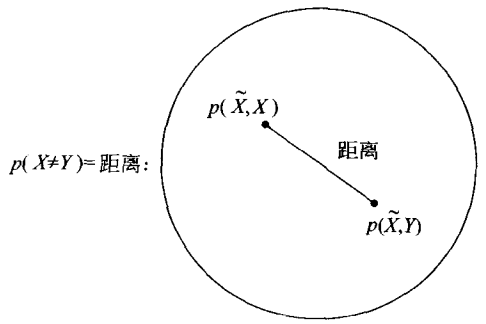


图 9.3　信道中出现一个差错的概率就等于$(\tilde{X},X)$的概率分布和$(\tilde{X},Y)$的概率分布之间的迹距离.

9.2　两个量子状态有多接近

两个量子状态有多接近？贯穿以下的几节中，我们要描述迹距离和保真度的经典概念的量子推广，并详细讨论这些量的性质.

9.2.1　迹距离

我们以定义量子状态 ρ 和 σ 之间的迹距离作为开端：

$$D(\rho,\sigma)\equiv\frac{1}{2}\mathrm{tr}\,|\rho-\sigma| \tag{9.11}$$

其中，按照通常惯例，我们定义 $|A|\equiv\sqrt{A^{\dagger}A}$ 为 $A^{\dagger}A$ 的正平方根. 注意，量子迹距离是在一定意义下推广了经典迹距离，即若 ρ 和 σ 对易，则 ρ 和 σ 之间的量子迹距离等于 ρ 和 σ 的特征值之间的经典迹距离. 更为明确地说，如果 ρ 和 σ 对易，那么它们在同一基上为对角阵，且对某个正交基 $|i\rangle$ 有

$$\rho=\sum_i r_i\,|i\rangle\langle i|,\quad \sigma=\sum_i s_i\,|i\rangle\langle i| \tag{9.12}$$

因此

$$D(\rho,\sigma)=\frac{1}{2}\mathrm{tr}\left|\sum_i(r_i-s_i)\,|i\rangle\langle i|\right| \tag{9.13}$$

$$=D(r_i,s_i) \tag{9.14}$$

练习 9.6　两个密度算子

$$\frac{3}{4}\,|0\rangle\langle 0|+\frac{1}{4}\,|1\rangle\langle 1|,\quad \frac{2}{3}\,|0\rangle\langle 0|+\frac{1}{3}\,|1\rangle\langle 1| \tag{9.15}$$

之间的迹距离是什么？两个密度算子

$$\frac{3}{4}|0\rangle\langle 0|+\frac{1}{4}|1\rangle\langle 1|, \quad \frac{2}{3}|+\rangle\langle +|+\frac{1}{3}|-\rangle\langle -| \tag{9.16}$$

之间的迹距离又是什么(回顾,$|\pm\rangle=(|0\rangle\pm|1\rangle)/\sqrt{2}$)?

理解迹距离的一个好的途径是,在 Bloch 球表示上,针对一个量子比特的特殊情况来对其理解. 设 ρ 和 σ 具有各自的 Bloch 向量 $\boldsymbol{r}$ 和 $\boldsymbol{s}$,

$$\rho=\frac{I+\boldsymbol{r}\cdot\boldsymbol{\sigma}}{2}, \quad \sigma=\frac{I+\boldsymbol{s}\cdot\boldsymbol{\sigma}}{2} \tag{9.17}$$

(回顾,$\boldsymbol{\sigma}$ 表示 Pauli 矩阵的向量;不要将其与状态 σ 相混淆). 于是,ρ 和 σ 之间的迹距离就可容易地计算:

$$D(\rho,\sigma)\equiv\frac{1}{2}\mathrm{tr}|\rho-\sigma| \tag{9.18}$$

$$=\frac{1}{4}\mathrm{tr}|(\boldsymbol{r}-\boldsymbol{s})\cdot\boldsymbol{\sigma}| \tag{9.19}$$

$(\boldsymbol{r}-\boldsymbol{s})\cdot\boldsymbol{\sigma}$ 具有特征值 $\pm|\boldsymbol{r}-\boldsymbol{s}|$,所以 $|(\boldsymbol{r}-\boldsymbol{s})\cdot\boldsymbol{\sigma}|$ 的迹为 $2|\boldsymbol{r}-\boldsymbol{s}|$,且我们看到

$$D(\rho,\sigma)=\frac{|\boldsymbol{r}-\boldsymbol{s}|}{2} \tag{9.20}$$

也即,两个单量子比特状态之间的距离等于 Bloch 球上它们之间普通 Euclid 距离的一半.

量子比特的迹距离的这种直觉几何图像,在试图来理解迹距离的一般性质时通常会是很有用的. 所猜想的性质,可以被建议,可以被驳倒,或可以是由观察 Bloch 球上的多个简单例子来得到的看似可能的结果. 举例来说,Bloch 球的旋转下 Euclid(欧几里得)距离保持不变. 这就提示,一般来说,迹距离在酉变换下可保持:

$$D(U\rho U^{\dagger},U\sigma U^{\dagger})=D(\rho,\sigma) \tag{9.21}$$

这是读者稍加思考就可容易验证的一个推测. 我们下面将回到距离度量研究中惯用的 Bloch 球的图像.

为了理解迹距离的性质,一个好的出发点是来证明推广经典迹距离的式(9.3)的迹距离公式:

$$D(\rho,\sigma)=\max_{P}\mathrm{tr}(P(\rho-\sigma)) \tag{9.22}$$

其中,极大化可以选取在所有投影算子 P 上或所有半正定算子 $P\leqslant I$ 上;这个公式在两种情况下都是有效的. 这个公式对迹距离引出一个吸引人的解释. 回顾 POVM(positive operator-valued measure)元为半正定算子 $P\leqslant I$,迹距离因而等于具有 POVM 元的测量输出在状态 ρ 和 σ 下可出现的概率上的极大差值,极大取遍所有的 POVM 元 P.

我们现就极大化取于投影算子上的情况证明式(9.22);对极大化取于半正定算子 $P\leqslant I$ 的情况可参照同样的推证. 这个证明是基于 $\rho-\sigma$ 可被表达为 $\rho-\sigma=$

$Q-S$的事实，其中 Q 和 S 为具有正交支集的半正定算子(参见练习 9.7). 这意味着 $|\rho-\sigma|=Q+S$，从而导出 $D(\rho,\sigma)=(\mathrm{tr}(Q)+\mathrm{tr}(S))/2$. 但是 $\mathrm{tr}(Q-S)=\mathrm{tr}(\rho-\sigma)=0$，所以 $\mathrm{tr}(Q)=\mathrm{tr}(S)$，并因而有 $D(\rho,\sigma)=\mathrm{tr}(Q)$. 令 P 为到 Q 的支集上的投影，则 $\mathrm{tr}(P(\rho-\sigma))=\mathrm{tr}(P(Q-S))=\mathrm{tr}(Q)=D(\rho,\sigma)$. 反之，令 P 为任一投影算子，则 $\mathrm{tr}(P(\rho-\sigma))=\mathrm{tr}(P(Q-S))\leqslant\mathrm{tr}(PQ)\leqslant\mathrm{tr}(Q)=D(\rho,\sigma)$. 这就完成了证明.

练习 9.7　试证明，对任意状态 ρ 和 σ，可以写出 $\rho-\sigma=Q-S$，其中 Q 和 S 为具有正交向量空间上支集的半正定算子(提示：应用谱分解 $\rho-\sigma=UDU^{\dagger}$，并将对角矩阵 D 分离为正和负部分. 这个事实延续到后面还会有用).

有一种将量子迹距离与经典迹距离更为密切联系起来的方法.

定理 9.1　令 $\{E_m\}$ 为一个 POVM，且以 $p_m=\mathrm{tr}(\rho E_m)$ 和 $q_m=\mathrm{tr}(\sigma E_m)$ 为由 m 标记的测量输出概率. 则有

$$D(\rho,\sigma)=\max_{\{E_m\}}D(p_m,q_m) \tag{9.23}$$

其中，极大取在所有 POVM $\{E_m\}$ 上.

证　注意到

$$D(p_m,q_m)=\frac{1}{2}\sum_m|\,\mathrm{tr}(E_m(\rho-\sigma))\,| \tag{9.24}$$

应用谱分解，我们可以写出 $\rho-\sigma=Q-S$，其中 Q 和 S 为具有正交支集的半正定算子. 因此，$|\rho-\sigma|=Q+S$，且有

$$|\,\mathrm{tr}(E_m(\rho-\sigma))\,|=|\,\mathrm{tr}(E_m(Q-S))\,| \tag{9.25}$$

$$\leqslant\mathrm{tr}(E_m(Q+S)) \tag{9.26}$$

$$\leqslant\mathrm{tr}(E_m|\,\rho-\sigma\,|) \tag{9.27}$$

从而

$$D(p_m,q_m)\leqslant\frac{1}{2}\sum_m\mathrm{tr}(E_m|\,\rho-\sigma\,|) \tag{9.28}$$

$$=\frac{1}{2}\mathrm{tr}(|\,\rho-\sigma\,|) \tag{9.29}$$

$$=D(\rho,\sigma) \tag{9.30}$$

其中，用到了 POVM 元的完备性关系 $\sum_m E_m=I$.

反之，利用选取一个测量使其 POVM 元包含到 Q 和 S 的支集上的投影，我们就看到存在测量使其发生概率分布为 $D(p_m,q_m)=D(\rho,\sigma)$. □

因此，如果两个密度算子是在迹距离意义下接近的，那么在量子状态上进行的任何测量都会引起迹距离经典意义下彼此接近的概率分布，同时还提供两个量子状态之间迹距离的第二个解释. 迹距离就是由两个量子状态上的测量所引起的概率分布之间的迹距离的一个可达的上界.

我们称迹距离为“距离”，所以我们应当来检验其是否具有作为密度算子空间

上的一个度量的属性. 根据我们的对一个单量子比特的几何图像，这对单量子比特显然是正确的；更一般地这是否正确呢？显然，$D(\rho,\sigma)=0$ 当且仅当 $\rho=\sigma$，$D(\cdot,\cdot)$为其输入的一个对称函数. 剩下需要检验的，只是如下的三角不等式成立：

$$D(\rho,\tau)\leqslant D(\rho,\sigma)+D(\sigma,\tau) \tag{9.31}$$

为了证明这一点，注意到从式(9.22)知，存在一个投影算子 P 使成立：

$$D(\rho,\tau)=\mathrm{tr}(P(\rho-\tau)) \tag{9.32}$$

$$=\mathrm{tr}(P(\rho-\sigma))+\mathrm{tr}(P(\sigma-\tau)) \tag{9.33}$$

$$\leqslant D(\rho,\sigma)+D(\sigma,\tau) \tag{9.34}$$

这就确立了迹距离是一个度量.

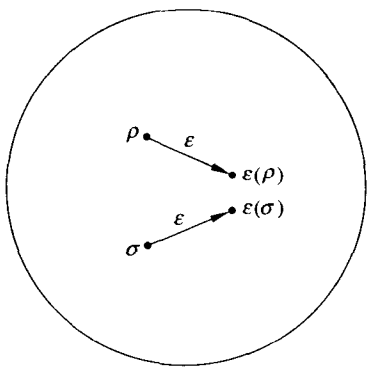


图 9.4 保迹量子运算引起密度算子空间上的压缩.

到现在，我们还没有了解关于迹距离全部一切. 但是，我们已能来证明某些确实引人入胜的结果，这些结果在很多情形中都是有用的. 最有兴趣的结果是，没有一个物理过程会增加两个量子状态之间的距离，这就是图 9.4 中所说明的结果. 我们现将其更为正式地陈述为一个定理.

定理 9.2(迹保持量子运算具有压缩性) 设 ε 为一个保迹量子运算，令 ρ 和 σ 为密度算子，则有

$$D(\varepsilon(\rho),\varepsilon(\sigma))\leqslant D(\rho,\sigma) \tag{9.35}$$

证 应用谱分解写出 $\rho-\sigma=Q-S$，其中 Q 和 S 为具有正交支集的半正定算子，并令 P 为使 $D(\varepsilon(\rho),\varepsilon(\sigma))=\mathrm{tr}[P(\varepsilon(\rho)-\varepsilon(\sigma))]$的一个投影算子. 注意到 $\mathrm{tr}(Q)-\mathrm{tr}(S)=\mathrm{tr}(\rho)-\mathrm{tr}(\sigma)=0$，所以 $\mathrm{tr}(Q)=\mathrm{tr}(S)$，从而有 $\mathrm{tr}(\varepsilon(Q))=\mathrm{tr}(\varepsilon(S))$. 应用这个事实，我们看到：

$$D(\rho,\sigma)=\frac{1}{2}\mathrm{tr}\,|\rho-\sigma| \tag{9.36}$$

$$=\frac{1}{2}\mathrm{tr}\,|Q-S| \tag{9.37}$$

$$=\frac{1}{2}\mathrm{tr}(Q)+\frac{1}{2}\mathrm{tr}(S) \tag{9.38}$$

$$=\frac{1}{2}\mathrm{tr}(\varepsilon(Q))+\frac{1}{2}\mathrm{tr}(\varepsilon(S)) \tag{9.39}$$

$$=\mathrm{tr}(\varepsilon(Q)) \tag{9.40}$$

$$\geqslant \mathrm{tr}(P\varepsilon(Q)) \tag{9.41}$$

$$\geqslant \mathrm{tr}(P(\varepsilon(Q)-\varepsilon(S))) \tag{9.42}$$

$$= \mathrm{tr}(P(\varepsilon(\rho)-\varepsilon(\sigma))) \tag{9.43}$$

$$= D(\varepsilon(\rho),\varepsilon(\sigma)) \tag{9.44}$$

这就完成了证明. □

这个结果有一个重要的特殊情况，它可通过下面的类比来理解. 想象某人向你展示画廊上的两幅不同的油画. 假设你眼力相当好，则你不应该会有任何困难来说出两幅画的区别. 另一方面，如同图 9.5 上所说明那样，如果把这两幅油画的大部分遮住，那么你很可能就有较大困难来说出两者的区别. 类似地，如果我们把两个量子状态的一部分“遮住”，那么我们能够证明这两个状态之间的距离是决不会增加的. 为证明这一点，回顾偏迹是一种保迹量子运算. 应用定理 9.2，如果我们取复合量子系统 AB 的量子状态 ρ^{AB} 和 σ^{AB}，那么 $\rho^A=\mathrm{tr}_B(\rho^{AB})$ 和 $\sigma^A=\mathrm{tr}_B(\sigma^{AB})$ 之间的距离决不会大于 ρ^{AB} 和 σ^{AB} 之间的距离，即

$$D(\rho^A,\sigma^A) \leqslant D(\rho^{AB},\sigma^{AB}) \tag{9.45}$$

在许多应用中，我们想要估计混合输入的迹距离. 利用下述定理，可对这样一些估计提供很大帮助.

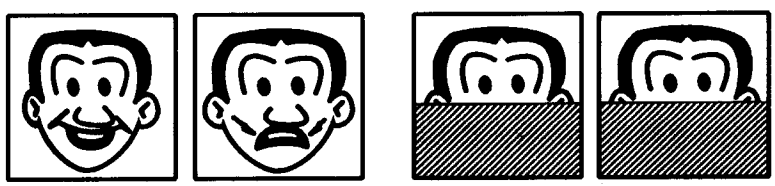


图 9.5　当只有部分信息可被利用时，物体就会变得较难区别.

定理 9.3（迹距离的强凸性）　令 $\{p_i\}$ 和 $\{q_i\}$ 为同一指标集上的概率分布，ρ_i 和 σ_i 为下标取自同一指标集上的密度算子，则

$$D\left(\sum_i p_i\rho_i,\sum_i q_i\sigma_i\right) \leqslant D(p_i,q_i)+\sum_i p_i D(\rho_i,\sigma_i) \tag{9.46}$$

其中，$D(p_i,q_i)$ 为概率分布 $\{p_i\}$ 和概率分布 $\{q_i\}$ 之间的经典迹距离.

这个结果可被用于对迹距离证明凸性结果，所以我们称这个性质为迹距离的强凸性性质.

证　应用式(9.22)，存在一个投影算子 P 使成立：

$$D\left(\sum_i p_i\rho_i,\sum_i q_i\sigma_i\right) = \sum_i p_i\mathrm{tr}(P\rho_i)-\sum_i q_i\mathrm{tr}(P\sigma_i) \tag{9.47}$$

$$= \sum_i p_i\mathrm{tr}(P(\rho_i-\sigma_i))+\sum_i (p_i-q_i)\mathrm{tr}(P\sigma_i) \tag{9.48}$$

$$\leqslant \sum_i p_i D(\rho_i,\sigma_i)+D(p_i,q_i) \tag{9.49}$$

其中，$D(p_i,q_i)$为概率分布$\{p_i\}$和概率分布$\{q_i\}$之间的迹距离，且在上述最后一行中我们用到式(9.22). □

作为这个结果的一种特殊情形，我们看到，迹距离为对其输入是联合凸的：

$$D(\sum_i p_i\rho_i,\sum_i p_i\sigma_i)\leqslant\sum_i p_iD(\rho_i,\sigma_i) \tag{9.50}$$

练习 9.8(迹距离的凸性) 试证明，迹距离为对其第一个输入是凸的：

$$D(\sum_i p_i\rho_i,\sigma)\leqslant\sum_i p_iD(\rho_i,\sigma) \tag{9.51}$$

由对称性，对第二个输入的凸性可由对第一个输入的凸性来导出.

练习 9.9(不动点的存在性) Schauder(绍德尔)不动点定理从数学角度是一个经典结果，它意味着 Hilbert 空间的一个凸的紧子集上的任一连续映射都具有一个不动点. 试用 Schauder 不动点定理来证明，任一保迹量子运算 ε 都具有一个不动点即 ρ，使成立 $\varepsilon(\rho)=\rho$.

练习 9.10 设 ε 为严格压缩(contractive)的保迹量子运算，即对任意 ρ和σ 有 $D(\varepsilon(\rho),\varepsilon(\sigma))<D(\rho,\sigma)$，试证明 ε 具有惟一不动点.

练习 9.11 设 ε 为保迹量子运算，且对其存在一个密度算子 ρ_0 和一个保迹量子运算 ε'，使对某个 p，$0<p\leqslant 1$，有下式成立：

$$\varepsilon(\rho)=p\rho_0+(1-p)\varepsilon'(\rho) \tag{9.52}$$

物理上，这个方程的意思为，输入状态以概率 p 被丢弃并为固定状态 ρ_0 所替代，同时运算 ε'以概率 $1-p$ 发生. 试应用联合凸性来证明，ε 是一个严格压缩的量子运算，且因而具有惟一的不动点.

练习 9.12 考虑 8.3.4 节中所引入的去极化信道 $\varepsilon(\rho)=pI/2+(1-p)\rho$，试对任意的 ρ和σ 应用 Bloch 表示来求 $D(\varepsilon(\rho),\varepsilon(\sigma))$，并显式地证明映射 ε 是严格压缩的，也即 $D(\varepsilon(\rho),\varepsilon(\sigma))<D(\rho,\sigma)$.

练习 9.13 试证明，比特翻转信道(8.3.3 节)是压缩的，但不是严格压缩的. 再来对比特翻转信道求出不动点的集合.

9.2.2 保真度

量子状态之间距离的第二个度量方法是保真度. 保真度并不是密度算子上的一个度量，但我们将会看到它确实会引出一个有用的度量. 这一节要来介绍保真度的定义和基本性质. 状态 ρ和状态σ 的保真度定义为

$$F(\rho,\sigma)\equiv \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \tag{9.53}$$

当然，这并不明显就是 ρ和σ 之间距离的一个有用的度量方法，甚至都看不出它是对称的. 然而，我们将会看到，保真度相对于其输入是对称的，并具有一个好的距离

度量方法所应具有的许多其他性质.

存在两种可对保真度提供更为显式的公式的重要特殊情况. 第一种是当 ρ 和 σ 为对易的情况，也即 ρ 和 σ 在同一基上是对角的，对某个标准正交基 $|i\rangle$ 有

$$\rho=\sum_i r_i\,|i\rangle\langle i|,\quad \sigma=\sum_i s_i\,|i\rangle\langle i| \tag{9.54}$$

在这种情况中，我们看到

$$F(\rho,\sigma)=\mathrm{tr}\sqrt{\sum_i r_i s_i\,|i\rangle\langle i|} \tag{9.55}$$

$$=\mathrm{tr}\Big(\sum_i \sqrt{r_i s_i}\,|i\rangle\langle i|\Big) \tag{9.56}$$

$$=\sum_i \sqrt{r_i s_i} \tag{9.57}$$

$$=F(r_i,s_i) \tag{9.58}$$

也即，当 ρ 和 σ 为对易时，量子保真度 $F(\rho,\sigma)$ 就还原为 ρ 和 σ 的特征值分布 r_i 和 s_i 之间的经典保真度 $F(r_i,s_i)$.

第二个例子是计算纯态 $|\psi\rangle$ 和任意状态 ρ 之间的保真度. 从式(9.53)，我们看到

$$F(|\psi\rangle,\rho)=\mathrm{tr}\sqrt{\langle\psi|\rho|\psi\rangle\,|\psi\rangle\langle\psi|} \tag{9.59}$$

$$=\sqrt{\langle\psi|\rho|\psi\rangle} \tag{9.60}$$

也即，保真度等于 $|\psi\rangle$ 和 ρ 之间重叠部分的平方根. 这是我们会经常用到的一个重要结果.

对于单量子比特的情况，我们能显式地来计算两个状态间的迹距离，并给出它的简单的几何解释，即为 Bloch 球上点之间的 Euclid 距离的一半. 遗憾的是，对于单量子比特的两状态之间的保真度，还不知道有类似那样清楚的几何解释.

不过，保真度确实满足许多与迹距离相同的性质. 举例来说，它在酉变换下是不变的：

$$F(U\rho U^\dagger,U\sigma U^\dagger)=F(\rho,\sigma) \tag{9.61}$$

练习 9.14(保真度在酉变换下的不变性)　试应用对任一半正定算子 A 有 $\sqrt{UAU^\dagger}=U\sqrt{A}\,U^\dagger$ 的事实来证明式(9.61).

保真度也存在一个与迹距离的特性式(9.22)相似的有用特性.

定理 9.4(Uhlmann 定理)　设 ρ 和 σ 为量子系统 Q 的状态，现引入为 Q 备份的另一量子系统 R，则

$$F(\rho,\sigma)=\max_{|\psi\rangle,|\varphi\rangle}|\langle\psi|\varphi\rangle| \tag{9.62}$$

其中，极大取遍 ρ 和 σ 在 RQ 中的所有纯化 $|\psi\rangle$ 和 $|\varphi\rangle$.

在给出 Uhlmann 定理的证明前，我们需要如下一个易于证明的引理.

引理 9.5 令 A 为任一算子，U 为酉算子，那么有

$$|\operatorname{tr}(AU)|\leqslant \operatorname{tr}|A| \tag{9.63}$$

其中，等号在选取 $U=V^{\dagger}$ 时成立，$A=|A|V$ 为 A 的极分解.

证 等式成立在所述条件下是清楚的. 注意到

$$\begin{aligned}|\operatorname{tr}(AU)| &= |\operatorname{tr}(|A|VU)| \\ &= |\operatorname{tr}(|A|^{1/2}|A|^{1/2}VU)|\end{aligned} \tag{9.64}$$

对 Hilbert-Schmidt 内积，Cauchy-Schwarz(柯西-施瓦茨)不等式给出

$$\begin{aligned}|\operatorname{tr}(AU)| &\leqslant \sqrt{\operatorname{tr}|A|\operatorname{tr}(U^{\dagger}V^{\dagger}|A|VU)} \\ &= \operatorname{tr}|A|\end{aligned} \tag{9.65}$$

这就完成了证明. □

证(Uhlmann 定理) 选定系统 R 和 Q 中正交基 $|i_R\rangle$ 和 $|i_Q\rangle$. 由于 R 和 Q 是同维的，指标 i 可假定为跑遍同一组值. 定义 $|m\rangle\equiv\sum_i |i_R\rangle|i_Q\rangle$，令 $|\psi\rangle$ 为 ρ 的任一纯化状态，那么，基于 Schmidt 分解易知，对系统 R 和 Q 上的某些酉算子 U_R 和 U_Q，有

$$|\psi\rangle=(U_R\otimes\sqrt{\rho}U_Q)|m\rangle \tag{9.66}$$

类似地，如果 $|\varphi\rangle$ 为 σ 的任一纯化状态，那么存在酉算子 V_R 和 V_Q 使成立：

$$|\varphi\rangle=(V_R\otimes\sqrt{\sigma}V_Q)|m\rangle \tag{9.67}$$

取内积，给出

$$|\langle\psi|\varphi\rangle|=|\langle m|(U_R^{\dagger}V_R\otimes U_Q^{\dagger}\sqrt{\rho}\sqrt{\sigma}V_Q)|m\rangle| \tag{9.68}$$

再应用下面的练习 9.16，我们看到

$$|\langle\psi|\varphi\rangle|=|\operatorname{tr}(V_R^{\dagger}U_RU_Q^{\dagger}\sqrt{\rho}\sqrt{\sigma}V_Q)| \tag{9.69}$$

记 $U\equiv V_QV_R^{\dagger}U_RU_Q^{\dagger}$，还可得到

$$|\langle\psi|\varphi\rangle|=|\operatorname{tr}(\sqrt{\rho}\sqrt{\sigma}U)| \tag{9.70}$$

应用引理 9.5，

$$|\langle\psi|\varphi\rangle|\leqslant\operatorname{tr}|\sqrt{\rho}\sqrt{\sigma}|=\operatorname{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \tag{9.71}$$

为明白等号可以成立，设 $\sqrt{\rho}\sqrt{\sigma}=|\sqrt{\rho}\sqrt{\sigma}|V$ 为 $\sqrt{\rho}\sqrt{\sigma}$ 的极分解. 选取 $U_R=U_Q=V_R=I$ 和 $V_Q=V^{\dagger}$，就可得到等式. □

练习 9.15 试证明：

$$F(\rho,\sigma)=\max_{|\varphi\rangle}|\langle\psi|\varphi\rangle| \tag{9.72}$$

其中，$|\psi\rangle$ 为 ρ 的任一固定的纯化状态，极大取遍 σ 的所有纯化状态.

练习 9.16(Hilbert-Schmidt 内积和纠缠)　设 R 和 Q 是具有相同 Hilbert 空间的两个量子系统，令 $|i_R\rangle$ 和 $|i_Q\rangle$ 为 R 和 Q 的标准正交基，令 A 为 R 上的一个算子，B 为 Q 上的一个算子，定义 $|m\rangle \equiv \sum_i |i_R\rangle |i_Q\rangle$. 试证明

$$\mathrm{tr}(A^\dagger B) = \langle m | (A \otimes B) | m \rangle \tag{9.73}$$

其中，等式左边的相乘是矩阵相乘，并理解为 A 的矩阵元是相对于基 $|i_R\rangle$ 取的，B 的矩阵元是相对于基 $|i_Q\rangle$ 取的.

Uhlmann 公式(9.62)并没有如式(9.53)那样提供确定保真度的一个计算工具. 但是，在许多例子中，应用 Uhlmann 公式比之式(9.53)，更易于来证明保真度的多条性质. 举例来说，Uhlmann 公式可使保真度的一些性质变得更为清楚，如保真度对其输入为对称即 $F(\rho,\sigma)=F(\sigma,\rho)$，保真度界于 0 和 1 之间即 $0 \leqslant F(\rho,\sigma) \leqslant 1$. 如果 $\rho=\sigma$，那么从 Uhlmann 公式可清楚地有 $F(\rho,\sigma)=1$. 如果 $\rho \neq \sigma$，那么对 ρ 和 σ 分别的任意纯化状态 $|\psi\rangle$ 和 $|\varphi\rangle$ 都有 $|\psi\rangle \neq |\varphi\rangle$，所以 $F(\rho,\sigma)<1$. 另一方面，作为理解保真度各种性质的一个手段，式(9.53)有时是有用的. 例如，我们看到 $F(\rho,\sigma)=0$ 当且仅当 ρ 和 σ 的支集位于正交的子空间上. 直观上，当 ρ 和 σ 支集位于正交的子空间上时，它们是完全可区分的，所以我们可以预期保真度在这些点上取为极小. 概括起来，保真度对其输入为对称；$0 \leqslant F(\rho,\sigma) \leqslant 1$；使前一个不等式中取等号当且仅当 ρ 和 σ 具有正交支集；而使后一个不等式中取等号当且仅当 $\rho=\sigma$.

我们已经看到，通过考虑由测量所导出的概率分布，量子迹距离可以与经典迹距离相联系. 按类似方法，可以证明

$$F(\rho,\sigma) = \min_{\{E_m\}} F(p_m, q_m) \tag{9.74}$$

其中，极小取遍所有 POVM$\{E_m\}$，$p_m \equiv \mathrm{tr}(\rho E_m)$ 和 $q_m \equiv \mathrm{tr}(\sigma E_m)$ 为 ρ 和 σ 的相应于 POVM$\{E_m\}$ 的概率分布. 为证明这一点，注意

$$F(\rho,\sigma) = \mathrm{tr}(\sqrt{\rho}\sqrt{\sigma}U) \tag{9.75}$$

$$= \sum_m \mathrm{tr}(\sqrt{\rho}\sqrt{E_m}\sqrt{E_m}\sqrt{\sigma}U) \tag{9.76}$$

其中 U 来自极分解 $\sqrt{\rho^{1/2}\sigma\rho^{1/2}} = \sqrt{\rho}\sqrt{\sigma}U$. 基于 Cauchy-Schwarz 不等式和简单的代数运算，给出

$$F(\rho,\sigma) \leqslant \sum_m \sqrt{\mathrm{tr}(\rho E_m)\mathrm{tr}(\sigma E_m)} \tag{9.77}$$

$$= F(p_m, q_m) \tag{9.78}$$

这就建立起关系式：

$$F(\rho,\sigma) \leqslant \min_{\{E_m\}} F(p_m, q_m) \tag{9.79}$$

为证明该不等式中等式可以成立，我们需要来找出一个 POVM$\{E_m\}$，使 Cauchy-

Schwarz 不等式对求和式中每一项都满足等式关系，也即对某组复数 α_m 成立 $\sqrt{E_m}\sqrt{\rho}=\alpha_m\sqrt{E_m}\sqrt{\sigma}U$. 但是 $\sqrt{\rho}\sqrt{\sigma}U=\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$，所以对可逆的 ρ，

$$\sqrt{\sigma}U=\rho^{-1/2}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \tag{9.80}$$

代入，我们就找到等式成立的条件为

$$\sqrt{E_m}(I-\alpha_m M)=0 \tag{9.81}$$

其中，$M\equiv\rho^{-1/2}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\rho^{-1/2}$. 如果 $M=\sum_m\beta_m|m\rangle\langle m|$ 为 M 的一个谱分解，那么我们就选取 $E_m=|m\rangle\langle m|$ 和 $\alpha_m=1/\beta_m$. ρ 不可逆的情况可从连续性来导出.

我们已证明迹距离的三个重要性质——度量属性、压缩性和严格凸性. 更为值得注意的是，对保真度相似的性质都成立. 而用于对保真度的证明方法相当不同于用于对迹距离的证明方法. 基于这个理由，值得稍为详细地来看一下这些结果.

保真度本身不是度量；但是，可以有一种简单的方法使保真度转化为一种度量. 注意到球面上两个点之间的角度是一种度量，其基本思想就可从图 9.6 看出. 对于量子的情况，Uhlmann 定理告诉我们，两个状态之间的保真度就等于这些状态的纯化状态之间的最大内积. 这就提示我们把状态 ρ 和 σ 之间的角度定义为

$$A(\rho,\sigma)\equiv\arccos F(\rho,\sigma) \tag{9.82}$$

显然，这个角度是非负的，相对于其输入为对称，且其等于零当且仅当 $\rho=\sigma$. 如果我们能证明这个角度服从三角不等式，那么我们就将确认这个角度是一种度量.

我们应用 Uhlmann 定理来证明这个三角不等式，并来证明关于三维向量的几个明显的事实. 令 $|\varphi\rangle$ 为 σ 的纯化状态，并选取 ρ 的纯化状态 $|\psi\rangle$ 和 τ 的纯化状态 $|\gamma\rangle$，使有

$$F(\rho,\sigma)=\langle\psi|\varphi\rangle \tag{9.83}$$

$$F(\sigma,\tau)=\langle\varphi|\gamma\rangle \tag{9.84}$$

且 $\langle\psi|\gamma\rangle$ 是实的和正的（注意，这一点总是可以做到的. 如果有必要，通过用适当的相位因子乘以 $|\psi\rangle$，$|\varphi\rangle$ 和 $|\gamma\rangle$）. 从图 9.6 可清楚看到

$$\arccos(\langle\psi|\gamma\rangle)\leqslant A(\rho,\sigma)+A(\sigma,\tau) \tag{9.85}$$

但据 Uhlmann 定理知 $F(\rho,\tau)\geqslant\langle\psi|\gamma\rangle$，所以 $A(\rho,\tau)\leqslant\arccos\langle\psi|\gamma\rangle$. 将其与前面的不等式相结合，就可给出如下三角形不等式：

$$A(\rho,\tau)\leqslant A(\rho,\sigma)+A(\sigma,\tau) \tag{9.86}$$

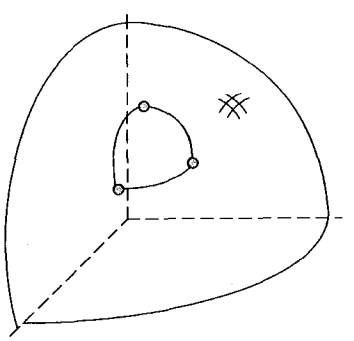


图 9.6 单位球面上点之间的角度是一种度量.

练习 9.17 试证明，$0\leqslant A(\rho,\tau)\leqslant\pi/2$，且

在第一个不等式中具有相等关系当且仅当 $\rho=\sigma$.

定性地，保真度表现为像是迹距离的一个“颠倒”版本，随两个状态的可区分性加强而减小，随两个状态的可区分性减弱而增大. 因此，我们不应期望，迹距离的压缩性和非增性对保真度成立. 而是类似的非减属性对保真度确实成立. 我们将称这一点为保真度在量子运算下的单调性.

定理 9.6(保真度的单调性)　设 ϵ 是一个保迹量子运算，令 ρ 和 σ 为密度算符，可以证明

$$F(\epsilon(\rho),\epsilon(\sigma)) \geqslant F(\rho,\sigma) \tag{9.87}$$

证　令 $|\psi\rangle$ 和 $|\varphi\rangle$ 为 ρ 和 σ 进入联合系统 RQ 的纯化状态，使成立 $F(\rho,\sigma)\geqslant|\langle\psi|\varphi\rangle|$. 对起始位于纯态 $|0\rangle$ 并通过酉交互 U 使与量子系统 Q 产生交互作用的量子运算 ϵ，引入一个模型环境 E. 注意到，$U|\psi\rangle|0\rangle$ 为 $\epsilon(\rho)$ 的一个纯化状态，$U|\varphi\rangle|0\rangle$ 为 $\epsilon(\sigma)$ 的一个纯化状态. 应用 Uhlmann 定理，可以导出

$$F(\epsilon(\rho),\epsilon(\sigma)) \geqslant |\langle\psi|\langle 0|U^{\dagger}U|\varphi\rangle|0\rangle| \tag{9.88}$$

$$= |\langle\psi|\varphi\rangle| \tag{9.89}$$

$$= F(\rho,\sigma) \tag{9.90}$$

这就建立起我们想要证明的性质. □

练习 9.18(角度的压缩性)　设 ϵ 是一个保迹量子运算，试证明

$$A(\epsilon(\rho),\epsilon(\sigma)) \leqslant A(\rho,\sigma) \tag{9.91}$$

应用 Uhlmann 定理，通过对保真度来证明相似于迹距离强凸性的一个结果，我们就算完成了对保真度的基本性质的研究.

定理 9.7(保真度的强凹性)　令 p_i 和 q_i 为同一指标集上的概率分布，ρ_i 和 σ_i 为也由相同指标集所标记的密度算子，则有

$$F\left(\sum_i p_i\rho_i,\sum_i q_i\sigma_i\right) \geqslant \sum_i \sqrt{p_iq_i}F(\rho_i,\sigma_i) \tag{9.92}$$

毫不令人惊讶，这个结果可以用来证明对保真度的凹性结果，因此我们称其为保真度的强凹性. 这个性质并不严格相似于迹距离的强凸性；但是，内涵上的相似性让我们采用类似的术语.

证　令 $|\psi_i\rangle$ 和 $|\varphi_i\rangle$ 为 ρ_i 和 σ_i 的纯化状态，并选取它使成立 $F(\rho_i,\sigma_i)=\langle\psi_i|\varphi_i\rangle$. 相应于概率分布的指标集 i，引入一个具有标准正交基状态 $|i\rangle$ 的辅助系统. 定义

$$|\psi\rangle \equiv \sum_i \sqrt{p_i}\,|\psi_i\rangle|i\rangle,$$

$$|\varphi\rangle \equiv \sum_i \sqrt{q_i}\,|\varphi_i\rangle|i\rangle \tag{9.93}$$

注意到，$|\psi\rangle$ 为 $\sum_i p_i\rho_i$ 的一个纯化状态，$|\varphi\rangle$ 为 $\sum_i q_i\sigma_i$ 的一个纯化状态. 所以，应用

Uhlmann 公式，得到

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geqslant |\langle\psi|\varphi\rangle| = \sum_i \sqrt{p_i q_i}\,|\langle\psi_i|\varphi_i\rangle| = \sum_i \sqrt{p_i q_i} F(\rho_i,\sigma_i) \tag{9.94}$$

这就获得了我们想要证明的结果. □

练习 9.19(保真度的联合凹性) 试证明保真度是联合凹的，有

$$F\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \geqslant \sum_i p_i F(\rho_i,\sigma_i) \tag{9.95}$$

练习 9.20(保真度的凹性) 试证明保真度相对于第一个元是凹的，有

$$F\left(\sum_i p_i\rho_i, \sigma\right) \geqslant \sum_i p_i F(\rho_i,\sigma) \tag{9.96}$$

根据对称性，保真度相对于第二个元也是凹的.

9.2.3 距离度量之间的关系

迹距离和保真度是密切相关的，尽管它们的形式很不相同. 定性地，对于许多的应用，它们可以认为是距离的等价度量方法. 这一节中，我们要更精确地定量确定迹距离和保真度之间的关系.

在纯态的情况中，迹距离和保真度是彼此完全等价的. 为看清这一点，考虑两个纯态 $|a\rangle$ 和 $|b\rangle$ 之间的迹距离. 应用 Gram-Schmidt 方法，我们可以找到正交状态 $|0\rangle$ 和 $|1\rangle$ 使 $|a\rangle = |0\rangle$ 和 $|b\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$. 注意到，$F(|a\rangle, |b\rangle) = |\cos\theta|$，进而，有

$$D(|a\rangle, |b\rangle) = \frac{1}{2}\mathrm{tr}\left|\begin{bmatrix} 1-\cos^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & -\sin^2\theta \end{bmatrix}\right| \tag{9.97}$$

$$= |\sin\theta| \tag{9.98}$$

$$= \sqrt{1 - F(|a\rangle, |b\rangle)^2} \tag{9.99}$$

因此，两个纯态之间的迹距离是这些状态保真度的一个函数，反之亦然. 纯态级别上的这个关系可以被用于推理出混合态级别上的关系. 令 ρ 和 σ 为任意两个量子状态，令 $|\psi\rangle$ 和 $|\varphi\rangle$ 为纯化状态，且使 $F(\rho,\sigma) = |\langle\psi|\varphi\rangle| F(|\psi\rangle, |\varphi\rangle)$. 回顾在偏迹下迹距离是非增的，我们看到

$$D(\rho,\sigma) \leqslant D(|\psi\rangle, |\varphi\rangle) \tag{9.100}$$

$$= \sqrt{1 - F(\rho,\sigma)^2} \tag{9.101}$$

因此，如果两个状态之间的保真度接近于 1，就可推论这些状态在迹距离下也是相

接近的. 反命题同样成立. 为看清这一点，令$\{E_m\}$为一个 POVM，使有

$$F(\rho,\sigma)=\sum_m\sqrt{p_m q_m} \tag{9.102}$$

其中，$p_m\equiv\mathrm{tr}(\rho E_m)$和$q_m\equiv\mathrm{tr}(\sigma E_m)$分别为对状态$\rho$和$\sigma$可得到结果$m$的概率. 首先，注意

$$\sum_m(\sqrt{p_m}-\sqrt{q_m})^2=\sum_m p_m+\sum_m q_m-2F(\rho,\sigma) \tag{9.103}$$

$$=2(1-F(\rho,\sigma)) \tag{9.104}$$

然而，同样成立$|\sqrt{p_m}-\sqrt{q_m}|\leqslant|\sqrt{p_m}+\sqrt{q_m}|$，所以

$$\sum_m(\sqrt{p_m}-\sqrt{q_m})^2\leqslant\sum_m|\sqrt{p_m}-\sqrt{q_m}||\sqrt{p_m}+\sqrt{q_m}| \tag{9.105}$$

$$=\sum_m|p_m-q_m| \tag{9.106}$$

$$=2D(p_m,q_m) \tag{9.107}$$

$$\leqslant 2D(\rho,\sigma) \tag{9.108}$$

比较式(9.104)和式(9.108)，我们看到

$$1-F(\rho,\sigma)\leqslant D(\rho,\sigma) \tag{9.109}$$

概括起来，有

$$1-F(\rho,\sigma)\leqslant D(\rho,\sigma)\leqslant\sqrt{1-F(\rho,\sigma)^2} \tag{9.110}$$

其含义是，在量子状态接近度的定性质上，迹距离和保真度是等价的度量方法. 事实上，在许多地方，并不在乎是采用迹距离还是保真度来定量确定距离，因为一种方法的结果可以推出另一种方法的等价结果.

练习 9.21　当比较纯态和混合态时，有可能来导出比式(9.110)更强的关于迹距离和保真度之间的关系表达式. 试证明

$$1-F(|\psi\rangle,\sigma)^2\leqslant D(|\psi\rangle,\sigma) \tag{9.111}$$

9.3　量子信道对信息的保持

朋友来又去，敌人却聚着不散.

——Jones 定律，属于 Thomas Jones

量子信道保持信息到什么程度？更为精确地，设量子系统处于状态$|\psi\rangle$，且出现某个物理过程，使量子系统改变到状态$\varepsilon(|\psi\rangle\langle\psi|)$. 量子信道$\varepsilon$多大程度上保持了量子系统的状态$|\psi\rangle$？前几节中讨论过的距离的静态度量，将被用于本节中，以研究对量子信道保持信息的程度的度量.

这种情形经常出现于量子计算和量子信息中. 举例来说，在量子计算机的存储

器中，$|\psi\rangle$为存储器的初始状态，ε代表存储器所经历的动力学过程，包括与环境交互而产生的噪声过程. 第二个例子出自于将状态$|\psi\rangle$从一个地方传输到另一个地方的量子通信信道. 没有一个信道在任何时候都是完美无缺的，所以信道的作用可用量子运算ε来描述.

定量确定状态$|\psi\rangle$可在多大程度上由信道所保持的一条明显途径是，应用9.2节中引入的静态距离度量. 举例来说，我们能够计算初始状态$|\psi\rangle$和终了状态$\varepsilon(|\psi\rangle\langle\psi|)$之间的保真度. 对于退极化信道的情况，得到

$$F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|))=\sqrt{\langle\psi|\left(p\frac{I}{2}+(1-p)|\psi\rangle\langle\psi|\right)|\psi\rangle} \tag{9.112}$$

$$=\sqrt{1-\frac{p}{2}} \tag{9.113}$$

这个结果与我们的直观符合得很好——退极化的概率p越高，终了状态与初始状态的保真度越低. 当规定p为很小时，保真度接近于1，状态$\varepsilon(\rho)$实际上与初始状态$|\psi\rangle$区别不开.

上面表达式中保真度的应用没有什么特别之处. 我们也可同样好地应用迹距离. 但是，对于本章的剩余部分，我们将只限于基于保真度及所导出的量来量度距离. 应用9.2节中所建立的迹距离的那些性质，对于大多数情形，都可毫无困难地给出基于迹距离的结果. 然而，事实上，保真度就计算而言是一种更为容易的工具，因此我们只限于基于保真度来考虑.

信息保持的原型度量，即保真度$F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|))$，具有一些需要加以补正的缺点. 在实际的量子存储器或量子通信信道中，我们事先并不知道系统的初始状态$|\psi\rangle$是什么. 但是，通过在所有可能的初始状态上取极小：

$$F_{\min}(\varepsilon)\equiv\min_{|\psi\rangle}F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|)) \tag{9.114}$$

我们可以定量确定系统的最坏情况的行为. 举例来说，对p退极化信道$F_{\min}=\sqrt{1-p/2}$，因为信道的保真度对所有输入状态$|\psi\rangle$都是相同的. 一个更有兴趣的例子是相位阻尼信道，

$$\varepsilon(\rho)=p\rho+(1-p)Z\rho Z \tag{9.115}$$

对于相位阻尼信道，保真度为

$$F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|))=\sqrt{\langle\psi|(p|\psi\rangle\langle\psi|+(1-p)Z|\psi\rangle\langle\psi|Z)|\psi\rangle} \tag{9.116}$$

$$=\sqrt{p+(1-p)\langle\psi|Z|\psi\rangle^2} \tag{9.117}$$

平方根符号中的第二项是非负的，且当$|\psi\rangle=(|0\rangle+|1\rangle)/2$时等于0. 因此，对于相位阻尼信道，最小保真度为

$$F_{\min}(\varepsilon)=\sqrt{p} \tag{9.118}$$

读者可能会感到奇怪，为什么 $F_{\min}$ 定义中我们是在纯态上取极小化. 难道，感兴趣的量子系统不会从混合态 ρ 出发吗？举例来说，量子存储器可能会与量子计算机的剩余部分相纠缠，从而它会从混合态出发. 幸运的是，可借助保真度的联合凹性来证明，允许混合态这一点并不改变 $F_{\min}$. 为看清这点，设 $\rho=\sum_i \lambda_i |i\rangle\langle i|$ 为量子系统的初始状态. 那么，我们有

$$F(\rho,\varepsilon(\rho))=F\left(\sum_i \lambda_i |i\rangle\langle i|,\sum_i \lambda_i \varepsilon(|i\rangle\langle i|)\right) \tag{9.119}$$

$$\geqslant \sum_i \lambda_i F(|i\rangle,\varepsilon(|i\rangle\langle i|)) \tag{9.120}$$

这就导出，对状态 $|i\rangle$ 的至少一个，有

$$F(\rho,\varepsilon(\rho)) \geqslant F(|i\rangle,\varepsilon(|i\rangle\langle i|)) \tag{9.121}$$

因此 $F(\rho,\varepsilon(\rho))\geqslant F_{\min}$.

当然，我们感兴趣的，不仅在于当通过量子通信信道传输时，而且在于当动态地经历计算时，要保持量子状态. 设想，举例来说，作为量子计算的组成部分，我们试图实现由酉算子 U 所描述的一个量子门. 如同第 8 章中描述过的，任何这样的尝试都会不可避免地遇到某些(希望不是太严重的)噪声，所以正确描述这个门要使用保迹量子运算 ε. 衡量我们的门成功与否的一个自然度量是门保真度：

$$F(U,\varepsilon)\equiv \min_{|\psi\rangle} F(U|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|)) \tag{9.122}$$

举例来说，设我们试图在单量子比特上来实现一个非门，但实现中代之用对某个小噪声参数 p 的受噪声污染的运算 $\varepsilon(\rho)=(1-p)X\rho X+pZ\rho Z$. 那么，这个运算的门保真度可为

$$F(X,\varepsilon)=\min_{|\psi\rangle}\sqrt{\langle\psi|X((1-p)X|\psi\rangle\langle\psi|X+pZ|\psi\rangle\langle\psi|Z)X|\psi\rangle} \tag{9.123}$$

$$=\min_{|\psi\rangle}\sqrt{(1-p)+p\langle\psi|Y|\psi\rangle^2} \tag{9.124}$$

$$=\sqrt{(1-p)} \tag{9.125}$$

在练习 9.22 中，将要求读者证明，执行每个都具有高保真度的一系列门运算，对保证整个运算具有高保真度是充分的，因而为了量子计算的目的，在计算中每个门都具有高保真度是充分的(请与第 4 章中关于近似量子线路的类似但较少一般性的论点相比较).

练习 9.22(保真度度量的链性质)　设 U 和 V 为酉算子，ε 和 $\mathcal{F}$ 为意欲近似 U 和 V 的保迹量子运算，令 $d(\cdot,\cdot)$ 为密度矩阵空间上的任一度量(如角度 $\arccos(F(\rho,\sigma))$). 假设该距离是酉不变的，即 $d(U\rho U^\dagger,U\sigma U^\dagger)=d(\rho,\sigma)$ 对任意密度矩阵 ρ,σ 和酉矩阵 U 成立. 定义相应的误差 $E(U,\varepsilon)$ 为

$$E(U,\varepsilon)\equiv \max_{\rho} d(U\rho U^\dagger,\varepsilon(\rho)) \tag{9.126}$$

试证明 $E(VU,\mathcal{F}\circ\varepsilon)\leqslant E(U,\varepsilon)+E(V,\mathcal{F})$. 因此，为执行一个具有高保真度的量子计算，只要以高保真度完成这个计算的每一步就足够了.

量子信源和纠缠保真度

在没有准确定义量子信源指什么的情况下，我们已经在谈论信息保持的动态度量了. 我们现在要来解释这个概念的两种可能定义，并用这些定义来引出信息保持的某些动态度量. 先验地，还不完全清楚的是，如何着手把定义量子信源这个概念做得最好. 经典做法上，这种定义问题的最好解答完全不是明显的，并有可能会提出多种不等价的定义，每个都会得到丰富和有用的信息理论. 由于量子信息包含经典信息作为其子领域，如果量子力学上会有多种方法来定义信源这个概念，这应当不会是令人惊讶的. 作为本章的结束，我们对信源这个概念引入两种可能的量子定义，解释它们是如何启发得到相应的信息保持的距离度量的，并会证明这些距离度量的一些基本性质. 对量子信源的进一步的讨论会被推迟到第 12 章.

量子信源的一个有吸引力的定义是，想象由某个物理过程产生的一串同样的量子系统(譬如说，一串量子比特)，其中，各个组成系统的状态由 ρ_{X_1}，ρ_{X_2}，…给定；X_j 为独立同分布的随机变量；ρ_j 为某个密度算子的固定集. 举例来说，读者可以想象一串量子比特，其每个量子比特以 1/2 的概率制备于状态 $|0\rangle$，或以 1/2 的概率制备于状态 $(|0\rangle+|1\rangle)/\sqrt{2}$.

量子信源的这种系综概念会自然地导出系综平均保真度(ensemble average fidelity)的概念. 此概念抓住了这样一个思想，就是量子信源在由保迹量子运算 ε 描述的噪声信道的作用下会很好地被保持，也即

$$F=\sum_j p_j F(\rho_j,\varepsilon(\rho_j))^2 \tag{9.127}$$

其中，p_j 为对系统制备为的不同状态 ρ_j 的概率. 显然，$0\leqslant\bar{F}\leqslant 1$，且在已知 $\bar{F}\approx 1$ 条件下我们可以确信，在平均意义上，信道 ε 会高准确度地来保持由量子信源所发出的状态. 读者可能会感到奇怪，为什么出现在定义式右边的保真度是取平方的. 对这个问题有两种答案，一种简单而一种复杂. 简单的答案是，如同下面将定义的，包含这个平方项会使整体保真度更为自然地与纠缠保真度联系起来. 较为复杂的答案是，量子信息目前尚处于起步阶段，并且还不完全清楚不少概念如信息保持等的“正确”定义是什么. 然而，如同我们将在第 12 章中会看到的，系综平均保真度和纠缠保真度会引出量子信息的丰富理论，这导致我们确信这些度量已经走上正确的轨道，尽管量子信息的完整理论至今还没有发展完成.

练习 9.23 试证明，当且仅当 $\varepsilon(\rho_j)=\rho_j$ 对所有使成立 $p_j>0$ 的 j，$\bar{F}=1$.

量子信源还存在我们可以考虑的第二个概念，它是由这样一种思想所引发的，一个能很好保持信息的信道一定也是一个能很好保持纠缠的信道. 这种基本思想

来自于 9.1 节中对差错的经典概率的讨论. 正像在那里所说的，因为定义在两个时间上的概率分布的直接量子类似版本不存在，量子过程不能定义差错概率 $p(X\neq Y)$ 的直接类似版本. 作为替代，我们将利用说明于图 9.7 中的思想的量子类似版本，它是利用首先复制随机变量 X 到 $\widetilde{X}$ 以定义距离的动态度量，然后使 X 受噪声的影响以得到 Y，并采用 $(\widetilde{X},X)$ 和 $(\widetilde{X},Y)$ 的各自联合分布之间的某个度量量 $D[(\widetilde{X},X),(\widetilde{X}',Y)]$ 作为我们的距离度量.

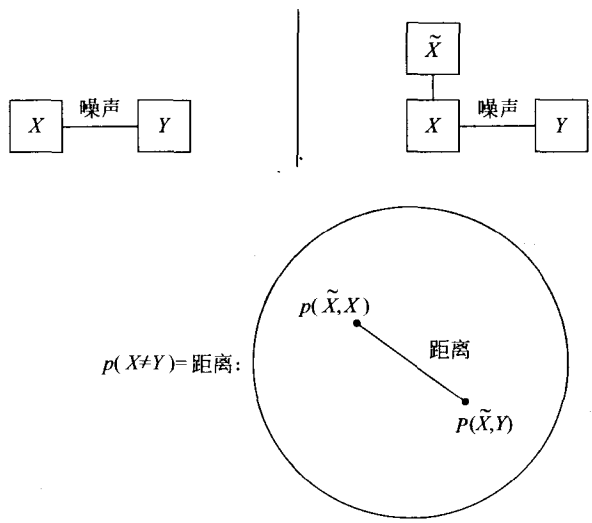


图 9.7　信道中出现一个差错的概率等于 $(\widetilde{X},X)$ 和 $(\widetilde{X},Y)$ 的概率分布之间的迹距离.

这个模型的量子类似版本如下. 量子系统 Q 制备于状态 ρ，Q 的这个状态假定为以某种方式与外部世界相纠缠. 这个纠缠取代经典模型中 X 和 $\widetilde{X}$ 之间的相关性. 我们通过引入一个假想的系统 R 来表示纠缠，使得 RQ 的联合状态为一个纯态. 事实上，我们所证明的所有结果不以任何方式依赖于这个纯化状态是如何被形成的，所以我们还可以合理地设想这是与外部世界的一种任意的纠缠. 系统 Q 于是属于由量子运算 ϵ 所描述的一个动力学过程，基本情况说明于图 9.8.

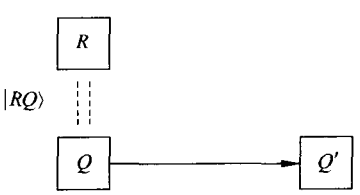


图 9.8　量子信道的 RQ 图像，RQ 的初始状态为纯态.

量子运算 ϵ 保持 Q 和 R 之间的纠缠到什么程度？我们应用纠缠保真度 $F(\rho,\epsilon)$ 来定量地确定这一点. 纠缠保真度是对保迹量子运算 ϵ 定义的 ρ 和 ϵ 的一个函数，为

$$F(\rho,\varepsilon)\equiv F(RQ,R'Q')^2 \tag{9.128}$$

$$=\langle RQ|[(\mathcal{I}_R\otimes\varepsilon)(|RQ\rangle\langle RQ|)]|RQ\rangle \tag{9.129}$$

其中，带“′”的符号表示量子运算作用后的系统状态，而不带“′”的符号表示量子运算作用前的系统状态. 这个定义式右边的量是 RQ 的初始状态和最终状态之间的静态保真度的平方. 因为它会简化纠缠保真度的某些性质，所以应用静态保真度的平方纯粹是为了方便一些. 注意到，纠缠保真度只依赖于 ρ 和 ε，而不(因为有可能出现)依赖于纯化状态 $|RQ\rangle$ 的细节. 为看清这一点，我们利用练习 2.81 所证明的事实，ρ 的任意两个纯化状态 $|R_1Q_1\rangle$ 和 $|R_2Q_2\rangle$ 是由单独作用于 R 的一个酉运算 U 相关联的，即 $|R_2Q_2\rangle=U|R_1Q_1\rangle$，因而

$$F(|R_2Q_2\rangle,\rho^{R_2'Q_2'})=F(|R_1Q_1\rangle,\rho^{R_1'Q_1'}) \tag{9.130}$$

这就得到了结果. 纠缠保真度提供了一个度量，以反映 R 和 Q 之间的纠缠多大程度上由过程 ε 所保持，值接近于 1 指示纠缠已很好地保持；值接近于 0 指示绝大部分纠缠已破坏. 选取静态保真度的平方还是静态保真度本身基本上是随意的；现在这个定义会导出稍微更具吸引力的一些数学性质.

纠缠保真度有吸引力的性质之一是具有能进行准确计算的一个非常简单的公式. 设 E_i 为量子运算 ε 的一组运算元，则

$$F(\rho,\varepsilon)=\langle RQ|\rho^{R'Q'}|RQ\rangle$$

$$=\sum_i|\langle RQ|E_i|RQ\rangle|^2 \tag{9.131}$$

设我们可写出 $|RQ\rangle=\sum_j\sqrt{p_j}|j\rangle|j\rangle$，其中 $\rho=\sum_j p_j|j\rangle\langle j|$，那么

$$\langle RQ|E_i|RQ\rangle=\sum_{jk}\sqrt{p_jp_k}\langle j|k\rangle\langle j|E_i|k\rangle \tag{9.132}$$

$$=\sum_j p_j\langle j|E_i|j\rangle \tag{9.133}$$

$$=\mathrm{tr}(E_i\rho) \tag{9.134}$$

将这个表达式代入式(9.131)，我们就得到如下有用的计算公式：

$$F(\rho,\varepsilon)=\sum_i|\mathrm{tr}(\rho E_i)|^2 \tag{9.135}$$

因此，举例来说，对相位阻尼信道 $\varepsilon(\rho)=p\rho+(1-p)Z\rho Z$，纠缠保真度为

$$F(\rho,\varepsilon)=p|\mathrm{tr}(\rho)|^2+(1-p)|\mathrm{tr}(\rho Z)|^2$$

$$=p+(1-p)\mathrm{tr}(\rho Z)^2 \tag{9.136}$$

所以我们看到，正如我们直觉上期望的，纠缠保真度会随着 p 的减少而随之减少.

我们已经定义了量子信源和相应距离度量的两个概念，一个基于的思想是我们以高平均保真度来保持量子状态的系综，另一个基于的思想是它就为我们想要保持的信源与某个参考系统之间的纠缠. 或许会令人惊讶，这两个定义实际上有着密切的关系. 这一点的理由在于纠缠保真度的两条有用的性质. 第一，纠缠保真度

为关于过程的输入和输出之间的静态保真度的平方的下界：

$$F(\rho,\varepsilon)\leqslant[F(\rho,\varepsilon(\rho))]^2 \tag{9.137}$$

直观上，这个结果是说，要保持状态再加上与外部世界的纠缠会难于仅只要单独保存状态. 它的证明只是静态保真度在偏迹下的单调性的一个基本应用，即 $F(\rho,\varepsilon)=F(|RQ\rangle,\rho^{R'Q'})^2\leqslant F(\rho^Q,\rho^{Q'})^2$.

纠缠保真度的第二个我们需要的将其与系综平均定义联系起来的性质为，纠缠保真度是 ρ 的一个凸函数. 为看清这一点，定义 $f(x)\equiv F(x\rho_1+(1-x)\rho_2,\varepsilon)$，并应用式(9.135)经过简单计算，我们得到

$$f''(x)=\sum_i|\operatorname{tr}((\rho_1-\rho_2)E_i)|^2 \tag{9.138}$$

所以 $f''(x)\geqslant 0$，这意味着纠缠保真度的凸性，如所要求那样.

结合这两个结果，我们看到

$$F\Big(\sum_j p_j\rho_j,\varepsilon\Big)\leqslant\sum_j p_jF(\rho_j,\varepsilon) \tag{9.139}$$

$$\leqslant\sum_j p_jF(\rho_j,\varepsilon(\rho_j))^2 \tag{9.140}$$

因此

$$F\Big(\sum_j p_j\rho_j,\varepsilon\Big)\leqslant\bar{F} \tag{9.141}$$

于是，任一量子信道 ε，可很好地保持由密度算子 ρ 描述的信源和一个参考系统之间的纠缠，就将会自动地很好保持由概率 p_j 和状态 ρ_j 所描述且使 $\rho=\sum_j p_j\rho_j$ 的系综信源. 在这个意义下，基于纠缠保真度的量子信源概念是比系综定义更为苛刻的一个概念. 因此，在第 12 章中，在量子信息理论研究中我们更喜欢基于纠缠保真度的定义.

我们现以简短归纳纠缠保真度的一些容易证明的性质来结束本章，这些性质在随后各章中将会是有用的.

(1) $0\leqslant F(\rho,\varepsilon)\leqslant 1$，这可立即由静态保真度的性质得到.

(2) 纠缠保真度相对于量子运算输入是线性的，这可立即由纠缠保真度的定义来导出.

(3) 对纯态输入，纠缠保真度等于输入和输出之间静态保真度的平方：

$$F(|\psi\rangle,\varepsilon)=F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|))^2 \tag{9.142}$$

这可立即由状态 $|\psi\rangle$ 为其自身的纯化状态的事实，以及纠缠保真度的定义来得到.

(4) $F(\rho,\varepsilon)=1$，当且仅当对 ρ 的支集上的所有纯态 $|\psi\rangle$，有

$$\varepsilon(|\psi\rangle\langle\psi|)=|\psi\rangle\langle\psi| \tag{9.143}$$

为证明这点，设 $F(\rho,\varepsilon)=1$，$|\psi\rangle$ 为 ρ 的支集上的一个纯态. 定义 $p\equiv 1/\langle\psi|\rho^{-1}|\psi\rangle>0$ (对比练习 2.73)并定义 σ 为密度算子使有 $(1-p)\sigma=\rho-p|\psi\rangle\langle\psi|$. 那么，利用凸

性，有

$$1 = F(\rho,\varepsilon) \leqslant pF(|\psi\rangle,\varepsilon) + (1-p) \tag{9.144}$$

因此 $F(|\psi\rangle,\varepsilon)=1$，这就获得了正方向的结果. 反方向的证明只是纠缠保真度定义的直截了当的应用.

(5) 设对某个 η，对 ρ 的支集上的所有 $|\psi\rangle$ 有 $\langle\psi|\varepsilon(|\psi\rangle\langle\psi|)|\psi\rangle \geqslant 1-\eta$，那么，必成立 $F(\rho,\varepsilon)\geqslant 1-(3\eta/2)$(参见问题 9.3).

问题 9.1(保真度的不同刻画) 试证明

$$F(\rho,\sigma) = \inf_P \sqrt{\mathrm{tr}(\rho P)\mathrm{tr}(\sigma P^{-1})} \tag{9.145}$$

其中，下确界取遍所有可逆的半正定矩阵 P.

问题 9.2 令 ε 为一个保迹量子运算，试证明，对每个 ρ，对 ε 存在一组运算元 $\{E_i\}$，使有

$$F(\rho,\varepsilon) = |\mathrm{tr}(\rho E_1)|^2 \tag{9.146}$$

问题 9.3 试证明事实(5).

第 9 章小结 量子信息的距离度量

- **迹距离** $D(\rho,\sigma)=\frac{1}{2}\mathrm{tr}|\rho-\sigma|$. 量子运算下压缩的密度算子上的双重凸度量.
- **保真度**

$$F(\rho,\sigma) \equiv \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}} = \max_{|\psi\rangle,|\varphi\rangle} |\langle\psi|\varphi\rangle|$$

强凹性 $F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geqslant \sum_i \sqrt{p_i q_i}F(\rho_i,\sigma_i)$.

- **纠缠保真度** $F(\rho,\varepsilon)$. 量子力学过程中纠缠保持程度的度量，其中系统 Q 初始状态为 ρ，并假定系统 Q 与另一个量子系统 R 相纠缠，且对系统 Q 作用以量子运算 ε.

历史和进一步阅读的材料

对想学习更多量子信息的距离度量的读者，建议可从 Fuchs 1996 年的博士学位论文[Fuc96]入手. 该文包含有大量的有关量子信息的距离度量的材料，包括一个 528 篇的有关组织于各主题领域的相关课题的参考文献目录. 特别是，式(9.74)的证明以及许多其他有价值的证明，都可以在那里找到. 迹距离的压缩性是由 Ruskai[Rus94] 证明的. 保真度的单调性由 Barnum, Caves, Fuchs, Jozsa,

Schumacher[BCF+96]证明. 在文献中，我们称为保真度的这个量及其平方两者都叫作保真度. Uhlmann 的证明以其名字命名定理的论文[Uhl76]也包含对保真度基本性质的广泛讨论. 本章中给出的 Uhlmann 定理的证明应归于 Jozsa[Joz94]. 对保真度度量的链性质及其与带噪声的量子计算间关系的更为详细的讨论由 Aharonov，Kitaev 和 Nisan[AKN98]给出. Schumacher[Sch96b]引入纠缠保真度并证明许多基本性质. Knill 和 Laflamme[KL97]建立了子空间保真度和纠缠保真度之间的联系，即问题 9.3. 这一事实的更为详细的证明见于 Barnum，Knill 和 Nielson 的论文[BKN98]. 问题 9.1 来源于 Alberti[Alb83].

CHAPTER 10
第 10 章
量子纠错

我们都会听说过，用纠缠来对抗纠缠是可能的.

——John Preskill

本章要来解释，在存在噪声的情况下，如何可靠地进行量子信息处理. 本章涉及三个主要的问题：量子纠错码的基本理论、容错量子计算和阈值定理. 我们将从介绍量子纠错码的基本理论开始，它能抵抗噪声以保护量子信息. 这些代码的工作机理为，先以特殊方式编码量子状态使其能抵消噪声的影响，然后在想要恢复原来的状态时进行解码. 10.1 节要来说明经典纠错的基本思想，以及为使量子纠错成为可能所必须克服的某些概念上的挑战. 10.2 节说明量子纠错码的一个简单例子，随后我们还会将其推广到 10.3 节的量子纠错码的理论中去. 10.4 节要来说明来自线性码经典理论的一些思想，以及它们如何来导出称为 Calderbank-Shor-Steane(CSS)码的一类有趣的量子码. 10.5 节讨论与经典纠错码有紧密联系的稳定子码(stabilizer code)这样一类结构丰富的码，来结束对量子纠错码的引论性概述.

我们对量子纠错的讨论假定，量子状态的编码和解码总是可以没有差错地完美地来做到的. 举例来说，如果我们想要把量子状态送到带噪声量子信道上，并能够使用几乎无噪声的量子计算机来对每个信道终端的状态执行完好的编码和解码，这一点是很有用的. 但是，如果用于编码和解码的量子门自身就受噪声污染，就不能作出这样的假设. 幸运的是，10.6 节中所阐述的容错量子计算的理论，会允许我们去掉完美的编码和解码这个假设. 甚至更为令人印象深刻的是，容错量子计算会允许我们以对相关的门运算容错的方式，对编码后的量子状态执行逻辑运算. 本章的重点是量子计算的阈值定理，出现在 10.6.4 节中：在单个量子门中的噪声低于一个确定阈值前提下，有可能来有效地执行任意规模的量子计算. 当然，也存在对这个结果的警告，对此我们会花一些篇幅讨论. 尽管如此，阈值定理仍是一个值得注意的结果，它表明噪声对大规模量子计算的性能看来并不会造成根本障碍.

10.1 引　　言

噪声是信息处理系统的一大祸害. 只要有可能,我们总要使我们构造的系统完全避免噪声,而对不可能的情形,我们会试图抵消噪声的影响. 举例来说,现代计算机中的元部件都是非常可靠的,典型故障率低于每 10^{17} 次运算出 1 次差错. 对于绝大多数的实际用途,我们可以认为计算机元部件完全没有噪声. 另一方面,许多广泛应用中的系统确实面临真正的噪声问题. 调制解调器和光盘播放装置两者都采用纠错码保护信息免受噪声影响. 实际中用于针对噪声的保护技术的细节有时还是相当复杂的,但其基本原理是容易理解的. 关键的思想是,如果我们想要针对噪声的影响保护一个消息,那么我们应当通过对这个消息加入一些冗余信息来编码消息. 采用这种方法,即使编码消息中的某些信息为噪声所污染,在编码消息中仍将有足够的冗余度使有可能恢复或解码消息,使得原来消息中的所有信息得以恢复.

举例来说,设我们想要通过带噪声经典信道从一个位置发送一个比特到另一个位置. 信道中噪声的影响是以概率 $p>0$ 将正在被传输的比特翻转,而以概率 $1-p$ 使比特无差错地传输. 这种信道称为二元对称信道(binary symmetric channel),并说明于图 10.1. 针对二元对称信道中噪声的影响来保护比特的一个简单手段是,把我们所想保护的比特替换为其自身的三份备份:

$$0 \rightarrow 000 \tag{10.1}$$

$$1 \rightarrow 111 \tag{10.2}$$

这个比特串 000 和 111 有时被叫做逻辑 0 和逻辑 1,因为它们分别扮演了 0 和 1 的角色. 我们现在通过信道发送所有这三个比特. 在信道的接收方三个比特均为输出,且接收方必须确定原来的比特的值是什么. 设从信道的输出为 001. 规定一个比特翻转的概率 p 为不太高,非常可能是第三比特被信道翻转,而 0 为所发送的比特.

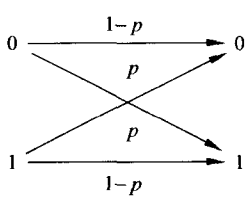


图 10.1　二元对称信道.

这种类型的解码被称为多数判决(majority voting),因为信道的解码输出不论是 1 还是 0,在实际信道输出都占多数. 如果通过信道发送的比特中两个或多个被翻转那么多数判决失败,否则多数判决成功. 所有比特中两个或多个被翻转的概率为 $3p^2(1-p)+p^3$,所以差错的概率为 $p_e=3p^2-2p^3$. 要是没有编码,出现一个差错的概率为 p,所以只要 $p_e<p$ 这种编码会使传输更为可靠,而 $p<1/2$ 时就是这种情况.

因为我们是通过将所发送的消息重复多次而来对其编码的，所以刚才描述的这类码称为重复码(repetition code)．类似的技术作为日常谈话的一部分已被用了上千年：如果我们听某个人所说的话语有困难，或许是因为他们有外国口音，我们会请他们重复一下他们刚才说的是什么．我们可能在任一时间里没有听清楚所有的词，但是我们可把重复的话合起来以理解一个连贯的消息．在经典纠错码的理论中，发展了许多聪明的技术；但是，关键的思想始终是利用加入足够的冗余来编码消息，使得有噪声的编码消息仍可以恢复原来的消息，至于需要添加的冗余量则依赖于信道中噪声的严重程度．

10.1.1 三量子比特的比特翻转码

为了抵消噪声的影响保护量子状态，我们将会基于类似的原理来引入量子纠错码．经典信息和量子信息之间存在着一些重要的区别，这就需要引入一种新的思想以使这样的量子纠错码成为可能．特别是，粗粗看一下，我们会涉及三个相当严重的困难：

(1) 不可克隆：有人可能试图通过将量子状态复制三次或多次，以量子力学方式来实现重复码．但据盒子 12.1 中的不可克隆(no-cloning)定理，这种作法是不允许的．即使复制是可能的，也不可能来度量和比较来自信道的三个量子状态输出．

(2) 差错是连续的：连续的不同差错可能出现在单量子比特上．为确定哪个差错出现以便来纠正它，看来需要无穷好的精度，因此要求无穷多的资源．

(3) 测量会破坏量子信息：在经典纠错中，我们会观测来自信道的输出，并决定采用什么样的解码步骤．量子力学中的观测一般会破坏所观测的量子状态并使恢复成为不可能．

幸运的是，如同我们将会论证的，这三个问题中没有一个是致命的．设我们通过一个信道发送量子比特，信道以概率 $1-p$ 保持量子比特不改变，以概率 p 使量子比特翻转．也即，以概率 p 状态 $|\psi\rangle$ 被取为状态 $X|\psi\rangle$，其中 X 为通常的 Pauli sigma x 算子或比特翻转算子．这种信道被称为比特翻转信道(bit flip channel)，我们现在来解释比特翻转码(bit flip code)，这种码可被用来针对来自这种信道的噪声的影响而保护量子比特．

设我们将单量子比特状态 $a|0\rangle+b|1\rangle$ 用三个量子比特编码为 $a|000\rangle+b|111\rangle$．一个方便的方法是把这个编码写为

$$|0\rangle \rightarrow |0_L\rangle \equiv |000\rangle \tag{10.3}$$

$$|1\rangle \rightarrow |1_L\rangle \equiv |111\rangle \tag{10.4}$$

其中，可以理解，基状态的叠加被取为相应编码状态的叠加．符号 $|0_L\rangle$ 和 $|1_L\rangle$ 表示

这些是逻辑$|0\rangle$和逻辑$|1\rangle$状态，而不是物理的 0 和 1 状态. 执行这种编码的线路说明于图 10.2.

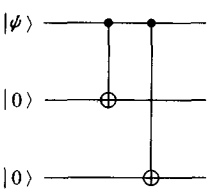


图 10.2　三量子比特比特翻转码的编码线路. 被编码的数据在顶部连线上加入线路.

练习 10.1　试验证图 10.2 中的编码线路可按要求工作.

设初始状态$a|0\rangle+b|1\rangle$已被完美地编码为$a|000\rangle+b|111\rangle$. 这三个量子比特中的每一个都通过一个比特翻转信道的独立备份. 设一个或更少的量子比特出现了一个比特翻转. 有一种简单的两阶段纠错方法，可用于恢复和纠正这种情况中的量子状态.

(1) 差错检测(error-detection)或症状诊断(syndrome diagnosis)：我们执行一次测量，它会告诉我们，什么差错(如果存在的话)出现在量子状态上. 这个测量结果被称为差错症状(error syndrome). 对于比特翻转信道，对应于四个投影算子，可有四种差错症状：

$$P_0 \equiv |000\rangle\langle 000| + |111\rangle\langle 111| \quad \text{没有差错} \tag{10.5}$$

$$P_1 \equiv |100\rangle\langle 100| + |011\rangle\langle 011| \quad \text{第一量子比特上比特翻转} \tag{10.6}$$

$$P_2 \equiv |010\rangle\langle 010| + |101\rangle\langle 101| \quad \text{第二量子比特上比特翻转} \tag{10.7}$$

$$P_3 \equiv |001\rangle\langle 001| + |110\rangle\langle 110| \quad \text{第三量子比特上比特翻转} \tag{10.8}$$

设举例来说比特翻转出现在第一个量子比特上，所以破坏后的状态为$a|100\rangle+b|011\rangle$. 注意到在这种情况中$\langle\psi|P_1|\psi\rangle=1$，所以测量结果(差错症状)的输出肯定为 1. 进而，差错症状测量不会引起状态的任何改变，在差错症状测量之前和之后的状态都为$a|100\rangle+b|011\rangle$. 注意，差错症状所包含的只是有关出现什么差错的信息，而不允许我们来推断有关 a 或 b 的值的任何事情，也即它不包含所被保护状态的信息. 这是差错症状测量的一个普遍特征，因为为了得到有关量子状态身份的信息一般地说有必要对那个状态进行扰动.

(2) 恢复(recovery)：我们采用差错症状的值来了解采用什么方法来恢复初始状态. 举例来说，如果差错症状为 1，指示第一个量子比特上比特翻转，则我们只要再一次翻转那个量子比特，就以完全准确地恢复到原状态$a|000\rangle+b|111\rangle$. 这四种可能的差错症状和每种情况中的恢复方法为：0(没有差错)——什么也不用做；1(第一量子比特上比特翻转)——再一次翻转第一量子比特；2(第二量子比特上比特翻转)——再一次翻转第二量子比特；3(第三量子比特上比特翻转)——再一次翻转第三量子比特. 对于差错症状的每个值，在给定相应所出现的差错后，容易看出原状态可得以完全准确的恢复.

这种纠错方法，只要在三个量子比特出现不超过一个的比特翻转，就可完美地来工作. 这种情况以概率$(1-p)^3+3p(1-p)^2=1-3p^2+2p^3$出现. 剩下一个差错

没有纠正的概率因而为 $3p^2-2p^3$，正好如我们以前研究过的经典重复码. 同样，对 $p<1/2$，编码和解码会改善量子状态的存储可靠性.

改善差错分析

上述这种差错分析并不是完全足够的. 因为并非所有的差错和量子力学中的状态都是同等地作用的：量子状态位于连续空间中，所以有可能某些差错只会使状态造成轻微破坏，而其他差错则使状态完全弄乱. 比特翻转差错 X 是一个极端的例子，这个差错根本不影响状态 $(|0\rangle+|1\rangle)/\sqrt{2}$，但会翻转 $|0\rangle$ 状态而使其变为 $|1\rangle$. 在前一种情况中我们无需担心比特翻转差错的出现，而在后一种情况中我们显然会非常担心.

为讨论这个问题，我们要用到引入于第 9 章中的保真度. 回顾，一个纯态和一个混合态之间的保真度为 $F(|\psi\rangle,\rho)=\sqrt{\langle\psi|\rho|\psi\rangle}$. 量子纠错的目标是来提高保真度，使以相当接近于最大可能保真度 1 来存储(或传送)量子信息. 让我们对利用三量子比特比特翻转码所能达到的最小保真度和当不执行纠错时的保真度来作一比较. 设所感兴趣的量子状态为 $|\psi\rangle$. 没有采用纠错码时，量子比特在经由信道发送后的状态为

$$\rho=(1-p)\,|\psi\rangle\langle\psi|+pX\,|\psi\rangle\langle\psi|\,X \tag{10.9}$$

保真度为

$$F=\sqrt{\langle\psi|\rho|\psi\rangle}=\sqrt{(1-p)+p\langle\psi|X|\psi\rangle\langle\psi|X|\psi\rangle} \tag{10.10}$$

平方根号中的第二项为非负且当 $|\psi\rangle=|0\rangle$ 时等于 0，所以我们看到最小保真度为 $F=\sqrt{(1-p)}$. 设三量子比特纠错码被用来保护 $|\psi\rangle=a|0_{\mathrm{L}}\rangle+b|1_{\mathrm{L}}\rangle$. 噪声和纠错两者作用后的量子状态为

$$\rho=[(1-p)^3+3p(1-p)^2]\,|\psi\rangle\langle\psi|+\cdots \tag{10.11}$$

省略号表示的项代表两个或三个量子比特上比特翻转的贡献. 所有这些省略项都是半正定算子，所以我们所计算的保真度将是真正保真度的一个下界. 我们看到，$F=\sqrt{\langle\psi|\rho|\psi\rangle}\geqslant\sqrt{(1-p)^3+3p(1-p)^2}$. 也即，保真度至少为 $\sqrt{1-3p^2+2p^3}$，所以存储量子状态的保真度在规定 $p<1/2$ 下会被改善，这与我们早先基于粗糙得多的分析得到的结论是相同的.

练习 10.2 比特翻转信道的作用可用量子运算 $\varepsilon(\rho)=(1-p)\rho+pX\rho X$ 来描述. 试证明，这可以有一个替代的算子和表示为 $\varepsilon(\rho)=(1-2p)\rho+2pP_+\rho P_++2pP_-\rho P_-$，其中 P_+ 和 P_- 分别为到 X 的 $+1$ 和 -1 本征态 $(|0\rangle+|1\rangle)/\sqrt{2}$ 和 $(|0\rangle-|1\rangle)/\sqrt{2}$ 上的投影算子. 这后一个表达式可理解为是量子比特以概率 $1-2p$ 保留，而以概率 $2p$ 被 $|+\rangle$ 和 $|-\rangle$ 基中的环境来测量的一个模型.

存在另一种理解差错症状测量的不同方法，它在推广三量子比特码中是会有

用处的. 假设,代替测量四个投影算子 P_0, P_1, P_2, P_3,我们来执行两个测量,第一个是对可观测量 Z_1Z_2(也即 $Z\otimes Z\otimes I$)的,第二个是对可观测量 Z_2Z_3 的. 这些观测量中的每一个都具有特征值 ± 1,所以每个测量都提供一个单量子比特信息,总共两个量子比特信息,即四个可能的差错症状,恰好等同于早先描述中的结果. 第一个测量即对 Z_1Z_2 的测量,可被想象为来比较第一个量子比特和第二个量子比特,看它们是否相同. 为看清为什么就是这回事,注意到 Z_1Z_2 具有谱分解

$$Z_1Z_2 = (|00\rangle\langle 00| + |11\rangle\langle 11|)\otimes I - (|01\rangle\langle 01| + |10\rangle\langle 10|)\otimes I \tag{10.12}$$

它对应于具有投影算子 $(|00\rangle\langle 00| + |11\rangle\langle 11|)\otimes I$ 和 $(|01\rangle\langle 01| + |10\rangle\langle 10|)\otimes I$ 的一个投影测量. 因此,测量 Z_1Z_2 可被想象为来比较第一个量子比特和第二个量子比特的值,若它们相同给出 $+1$,若它们不同给出 -1. 类似地,测量 Z_2Z_3 被想象为比较第二个量子比特和第三个量子比特的值,若它们相同给出 $+1$,若它们不同给出 -1. 组合这两个测量结果,我们就能确定是否有比特翻转出现在量子比特中的一个比特上,如果有那么还可来确定是哪一种:若两者测量结果都给出 $+1$,则以高概率没有比特翻转出现;若测量 Z_1Z_2 给出 $+1$ 而测量 Z_2Z_3 给出 -1,则以高概率只有第三个量子比特翻转;若测量 Z_1Z_2 给出 -1 而测量 Z_2Z_3 给出 $+1$,则以高概率只有第一个量子比特翻转;最后,若两者测量结果都给出 -1,则以高概率只有第二个量子比特翻转. 这些测量成功的关键是,两个测量都不会给出有关编码后量子状态的幅值 a 和 b 的任何信息,因而两个测量都不会破坏我们想用这个码来保护的量子状态的叠加.

练习 10.3　试用显式计算证明,在两种方法导致相同的测量统计结果和测量后状态的意义下,在不计测量结果标号情况下,测量由式(10.5)～式(10.8)定义的四个投影算子,与测量 Z_2Z_3 后测量 Z_1Z_2 是等价的.

10.1.2　三量子比特相位翻转码

比特翻转码是令人感兴趣的,但它看来并不是超越经典纠错码的一种重要的革新,并且留下许多未解决的问题(例如,区别于比特翻转的许多类型的差错可能对量子比特发生). 一个更有兴趣的带噪声量子信道是单量子比特的相位翻转差错模型. 在这种差错模型中,量子比特以概率 $1-p$ 得以保留,而状态 $|0\rangle$ 和 $|1\rangle$ 的相对相位以概率 p 被翻转. 更确切地说,相位翻转算子 Z 以概率 $p>0$ 作用于量子比特,所以在相位翻转下状态 $a|0\rangle+b|1\rangle$ 变为状态 $a|0\rangle-b|1\rangle$. 相位翻转信道没有经典的等价物,因为经典信道不具有任何相位的等价性质. 不过,有一种简单的方法把相位翻转信道转化为比特翻转信道. 设我们考虑的是量子比特基 $|+\rangle \equiv (|0\rangle+|1\rangle)/\sqrt{2}$ 和 $|-\rangle \equiv (|0\rangle-|1\rangle)/\sqrt{2}$. 关于这个基,算子 Z 变 $|+\rangle$ 为 $|-\rangle$ 并且

反之亦然，也即标志“+”和“−”只是起到像比特翻转的作用. 这就提示我们，采用状态$|0_L\rangle\equiv|+++\rangle$和$|1_L\rangle\equiv|---\rangle$作为逻辑 0 状态和逻辑 1 状态来保护针对相位翻转的差错. 纠错需要的所有运算——编码、差错检测和恢复——都是像比特翻转信道那样来执行，只是要以$|+\rangle$，$|-\rangle$基代替$|0\rangle$，$|1\rangle$基. 为了实现这种基的转换，我们只要在纠错过程的适当位置上应用 Hadamard 门及其逆(也为 Hadamard 门)，因为 Hadamard 门可来实现$|0\rangle$，$|1\rangle$基和$|+\rangle$，$|-\rangle$基之间来回的转换.

更明确地，对相位翻转信道的编码可按两步来执行：第一，完全准确地对三个量子比特按比特翻转信道那样编码；第二，对每个量子比特作用 Hadamard 门(图 10.3). 差错检测可通过如前相同的投影测量而来达到，但要由 Hadamard 门取共轭：$P_j \to P'_j \equiv H^{\otimes 3} P_j H^{\otimes 3}$. 等价地，差错症状测量可以通过测量观测量$H^{\otimes 3} Z_1 Z_2 H^{\otimes 3} = X_1 X_2$和$H^{\otimes 3} Z_2 Z_3 H^{\otimes 3} = X_2 X_3$来执行. 令人感兴趣的是，按照比特翻转码的$Z_1 Z_2$和$Z_2 Z_3$的测量相类似思路，可对这些测量作出解释. 对观测量$X_1 X_2$和$X_2 X_3$的测量，分别对应于比较第一和第二量子比特，以及比较第二和第三量子比特的正负号. 其含义为，在$X_1 X_2$的测量例如对形如$|+\rangle|+\rangle\otimes(\cdot)$或$|-\rangle|-\rangle\otimes(\cdot)$的状态给出+1，而对形如$|+\rangle|-\rangle\otimes(\cdot)$或$|-\rangle|+\rangle\otimes(\cdot)$的状态给出−1. 最后，纠错可用恢复运算来完成，这种运算就是从比特翻转码导出的 Hadamard 共轭(Hadamard-conjugated)恢复运算. 举例来说，设我们在第一量子比特的符号中检测到一个从$|+\rangle$到$|-\rangle$的翻转，那么，我们就可通过对第一量子比特作用$HX_1H=Z_1$而来恢复. 对其他的差错症状可应用类似的方法.

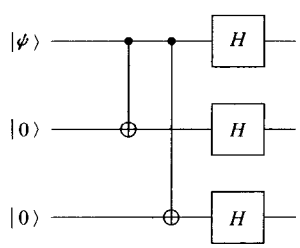


图 10.3 相位翻转码的编码线路.

显然，这种相位翻转信道的码具有像比特翻转信道的码相同的特性. 特别是，相位翻转码的最小保真度等同于比特翻转码的最小保真度，并且对没有纠错的情况能提供改善的码我们有着同样的判据. 我们说，这两个信道是酉等价的，因为存在一个酉算子U(这种情况中的 Hadamard 门)，在第一个信道先用U而跟随的通道用$U^\dagger$情况下，使得一个信道的作用同于另一个信道的作用. 这些运算可以平凡地合并到编码和纠错运算中. 对一般的酉算子，这些思想在问题 10.1 讨论.

练习 10.4 考虑三量子比特比特翻转码. 设通过测量对应于到 8 个计算基状态上投影的 8 个正交投影算子，我们执行了对差错症状的测量.

(1) 试写出对应于这个测量的投影算子，并解释测量结果如何可被用于诊断差错症状：要么所有比特没有翻转，要么比特数为j的比特翻转，其中j取值于 1 到 3 之内.

(2) 试证明恢复方法仅适用于计算基状态.

(3) 纠错方法的最小保真度是什么?

10.2　Shor 码

有一种简单的量子码能针对单量子比特上的任意差错的影响进行保护. 这种码就是以其发明者命名的 Shor 码，它是三量子比特相位翻转码和三量子比特比特翻转码的组合. 我们首先用相位翻转码来编码量子比特：$|0\rangle \to |+++\rangle$，$|1\rangle \to |---\rangle$. 其次，我们用三个量子比特比特翻转码来编码这些量子比特中的每个：$|+\rangle$编码为$(|000\rangle + |111\rangle)/\sqrt{2}$，$|-\rangle$编码为$(|000\rangle - |111\rangle)/\sqrt{2}$. 这个结果为 9 量子比特码，其码字为

$$|0\rangle \to |0_L\rangle \equiv \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$

$$|1\rangle \to |1_L\rangle \equiv \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}} \tag{10.13}$$

编码 Shor 码的量子线路如图 10.4 所示. 如上面所描述的，线路的第一部分用三个量子比特相位翻转码来编码量子比特；与图 10.3 的比较显示，两个线路是相同的. 线路的第二部分用比特翻转码来编码这三个量子比特的每一个，也即同时用三份图 10.2 的比特翻转码编码线路. 这种采用多层级的编码方法称为串联. 它是从老的码得到新的码的一个重要技巧，随后我们还会再一次用它来证明关于量子纠错的一些重要结果.

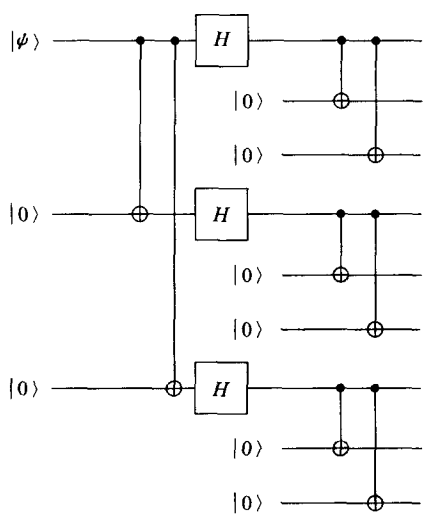


图 10.4　9 量子比特 Shor 码的编码线路. 缩进的那些 $|0\rangle$ 状态只是为强调编码的串联本质.

Shor 码能对任一量子比特上的相位翻转差错和比特翻转差错进行保护. 为看清这点,设比特翻转出现在第一个量子比特上. 就比特翻转码来说,我们执行对 Z_1Z_2 的一次测量,并比较前两个量子比特,发现它们为不同. 基此我们得出结论,比特翻转差错出现在第一或第二量子比特上. 下一步,我们通过执行对 Z_2Z_3 的一次测量来比较第二和第三量子比特. 我们发现它们为相同,所以第二量子比特不可能出现翻转. 据此得到结论,第一量子比特必出现了翻转,只要再一次翻转第一量子比特就会从差错中恢复,回复到原来的状态. 按类似的方法,我们就能从这个码中检测和恢复出 9 个量子比特的任意一个受比特翻转差错影响的比特.

我们现以类似的方式来处理量子比特上的相位翻转. 设相位翻转出现在第一量子比特上. 这个相位翻转使第一量子比特块中的符号翻转,变$(|000\rangle+|111\rangle)$为$(|000\rangle-|111\rangle)$,反之亦然. 事实上,前三个量子比特中任意一个上的相位翻转都具有这种影响,我们所描述的纠错方法对这三种可能差错中的任意一种都能奏效. 差错症状测量开始于比较第一个和第二个三量子比特块的符号,就像对相位翻转码的差错症状测量开始于比较第一和第二量子比特一样. 举例来说,$(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)$在两个量子比特块中具有相同符号(−),而$(|000\rangle-|111\rangle)(|000\rangle+|111\rangle)$具有不同符号. 当相位翻转出现在前三个量子比特中任意一个上时,我们发现第一和第二量子比特块的符号为不同. 差错症状测量的第二阶段和最后阶段是来比较第二和第三量子比特块的符号. 我们发现这些符号为相同,并得到结论在第一个三量子比特块中必有翻转. 通过翻转第一个三量子比特块中的符号,我们就会将其恢复到它的原来的值. 以类似的方式,我们还可恢复 9 个量子比特中任意一个上的相位翻转.

练习 10.5 试证明,检测 Shor 码中相位翻转差错的差错症状测量对应于测量观测量 $X_1X_2X_3X_4X_5X_6$ 和 $X_4X_5X_6X_7X_8X_9$.

练习 10.6 试证明,前三个量子比特中任意一个上的相位翻转恢复可以通过应用算子 $Z_1Z_2Z_3$ 而来实现.

设比特翻转和相位翻转差错两者同时出现在第一量子比特上,也即算子 Z_1X_1 作用于那个量子比特上. 那么,容易看出,检测比特翻转差错的方法将可来检测第一量子比特上的比特翻转并对其纠正,检测相位翻转差错的方法将可来检测第一个三量子比特块上的相位翻转并对其纠正. 因此,Shor 码也能来纠正单量子比特上比特翻转和相位翻转的组合差错.

事实上,Shor 码对单量子比特上的保护要比比特翻转和相位翻转差错多得多——我们现来证明,Shor 码可对完全任意的差错进行保护,只要规定它们仅只影响一个单量子比特. 差错可以是微小的,比如说,沿 Bloch 球的 z 轴旋转 $\pi/263$ 弧度;或者可以为明显严重的差错,像完全移去量子比特并将其替换为垃圾. 令人感兴趣的事情是,为了对任意的差错进行保护,无需做任何附加的工作——已经描

述过的这个方法就可以了. 这是一个特别惊人的例子，可能出现在单量子比特上的明显连续差错都可通过只纠正这些差错的一个离散子集而被纠正；所有其他可能的差错都可用这种方法自动纠正. 差错的这种离散化对于量子纠错为什么能工作是很重要的，并且应将其看作为是与模拟系统的经典纠错的一个对照，模拟系统中这样的差错离散化是不可能的.

为简化分析，设一个任意类型的噪声只出现在第一量子比特上；我们以后会回来讨论当噪声也影响到其他量子比特时的情形. 遵循第 8 章，我们用保迹量子运算 ε 来描述噪声. 分析纠错的最为方便的做法是将 ε 展开为具有运算元 $\{E_i\}$ 的算子和表示. 设噪声作用前，编码后的量子比特状态为 $|\psi\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$，则噪声作用以后这个状态为 $\varepsilon(|\psi\rangle\langle\psi|) = \sum_i E_i|\psi\rangle\langle\psi|E_i^\dagger$. 为分析纠错的作用，最容易的做法是把纠错作用集中到这个和式的一个单项上，比如说 $E_i|\psi\rangle\langle\psi|E_i^\dagger$. 作为是第一量子比特上的一个单独的算子 E_i，可以被展开为单位阵 I、比特翻转 X_1、相位翻转 Z_1 以及组合位与相位翻转 X_1Z_1 的一个线性组合：

$$E_i = e_{i0}I + e_{i1}X_1 + e_{i2}Z_1 + e_{i3}X_1Z_1 \tag{10.14}$$

(非归一化)量子状态 $E_i|\psi\rangle$ 因而可写成 $|\psi\rangle, X_1|\psi\rangle, Z_1|\psi\rangle, X_1Z_1|\psi\rangle$ 四项的叠加. 测量差错症状会将这个叠加结果坍缩为四个状态 $|\psi\rangle, X_1|\psi\rangle, Z_1|\psi\rangle, X_1Z_1|\psi\rangle$ 之一，恢复过程由随后通过作用适当的逆运算而执行，并获得最终状态 $|\psi\rangle$. 这些对于所有其他运算元 E_i 也是正确的. 因此，纠错会使原来状态 $|\psi\rangle$ 恢复，尽管事实上第一量子比特上的差错是任意的. 这是关于量子纠错的一个基本和深刻的事实，通过只纠正差错的一个离散集——在这个例子中，为比特翻转、相位翻转以及组合比特与相位翻转——量子纠错码就会自动地纠正要比这个离散集明显大得多的(连续)差错类.

当噪声不只影响第一个量子比特时会发生什么呢？处理这个问题有两个基本思想. 首先，在很多情况中，一个很好的近似是假定噪声为独立地作用于量子比特. 在规定噪声在一个量子比特上的影响为相当小后，我们就可把噪声的总影响展开为包括零量子比特上的差错、单量子比特上的差错、双量子比特上的差错等各项之和，并以零量子比特上的差错项和单量子比特上的差错项控制高阶项. 执行纠错会带来对零阶项和一阶项的合理纠正，剩下的只是小得多的二阶和高阶差错，这就实现了对差错的基本抑制. 基于这种思想的更为详细的分析将会在以后给出. 当然，有些时候，假定噪声为独立地作用于量子比特并不合理. 当出现这种情形时，我们会采用不同的思想——能够纠正多于一个单量子比特上差错的纠错码. 这种码可以沿着类似于 Shor 码的思路来构造，并且我们会在本章随后部分中去解释如何能做到这点的那些基本思想.

10.3 量子纠错的理论

我们能构造量子纠错码的一般理论吗？这一节将来介绍研究量子纠错的一个总的框架，其中包括量子纠错条件，这是一组使量子纠错成为可能所必须满足的方程. 当然，提出这一个框架并不能保证好的量子纠错码一定存在，这个课题会在10.4节中加以处理. 但是，这个框架确实可以提供背景知识，能使我们来找到好的量子纠错码.

量子纠错理论的基本思想自然地推广了由Shor码引入的思想. 量子状态通过酉运算被编码为量子纠错码，其形式定义为某个较大Hilbert空间中的一个子空间C. 为方便起见，我们采用符号P表示到码空间C上的投影算子；且对三量子比特比特翻转码，$P=|000\rangle\langle 000|+|111\rangle\langle 111|$. 在编码以后，这个码会受到噪声的影响，紧接着执行差错症状测量以检测所出现的差错类型. 一旦差错症状确定，恢复运算就会执行，以使量子系统回复到这个码的原来状态. 其基本图像说明于图10.5，不同的差错症状对应于整个Hilbert空间中保形的和正交的子空间. 这些

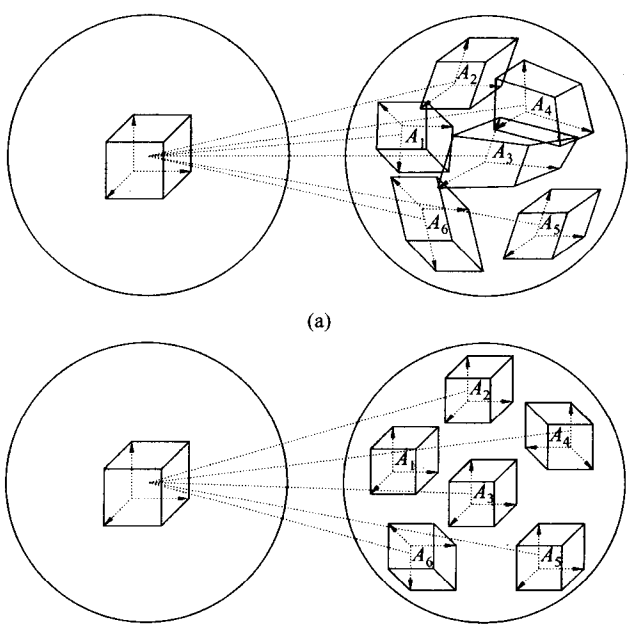


图10.5 量子编码中Hilbert空间的封装：(a)坏码，具有非正交的、变形的空间；(b)好码，具有正交的(可区别的)、保形的空间.

子空间必是正交的，否则它们就不能被差错症状测量可靠地区分. 进而，由于到不同子空间的差错映射必将(正交)码字映射到正交状态，因此这些不同的子空间必为是原来码空间的保形版本，这样就能使其从差错中恢复. 这个直观的图像基本上就是下面所要讨论的量子纠错条件的要旨.

为发展量子纠错的一般理论，就关于噪声的本质和关于用于处理纠错的程序，作一些尽可能少的假设，这对我们来说是适宜的. 也即，我们没有必要假定纠错是通过检测-恢复两阶段方法来做到的，我们也无需对量子系统的噪声作任何假定或减弱假定. 代之为，我们仅只作两个很宽泛的假定：噪声由量子运算 ε 所描述，整个纠错方法由我们称之为纠错运算的一个保迹量子运算 $\mathcal{R}$ 承担. 这个纠错运算把我们上面称为差错检测和恢复的两个步骤合到一起了. 为确保纠错是成功的，我们要求对任何状态 ρ，其支集位于码空间 C 中，有

$$(\mathcal{R}\circ\varepsilon)(\rho)\propto\rho \tag{10.15}$$

读者可能想知道，为什么我们在上述方程中写为“$\propto$”而不是“$=$”. 如果 ε 为保迹量子运算，那么通过对方程两边取迹我们看到“$\propto$”会变为“$=$”. 但是，有时我们可能感兴趣于纠错的非保迹量子运算如测量，对这种情形取“$\propto$”是合适的. 当然，纠错步骤 $\mathcal{R}$ 必须以概率 1 成功，这就是为什么我们要求 $\mathcal{R}$ 为保迹的原因.

量子纠错条件是一个简单方程组，它们可被检验以确定量子纠错码是否能对抗特殊类型的噪声 ε. 我们将应用这些条件来构造大量的量子码，并将研究量子纠错码的一些普遍性质.

定理 10.1(量子纠错条件)　令 C 为一个量子码，令 P 为到 C 的投影算子. 设 ε 为具有运算元 $\{E_i\}$ 的量子运算. 则纠正 C 上 ε 的纠错运算 $\mathcal{R}$ 存在的充分必要条件为，对某个复数 Hermite 矩阵 α 成立

$$PE_i^{\dagger}E_jP=\alpha_{ij}P \tag{10.16}$$

我们称运算元 $\{E_i\}$ 为噪声 ε 的差错，且如果这样一个 $\mathcal{R}$ 存在，我们就说 $\{E_i\}$ 组成一个可纠正的差错集合.

证　在式(10.16)成立的前提下，我们首先通过构造一个显式纠错运算来证明充分性. 我们采用的构造方法就是用于 Shor 码的两步骤形式——差错检测和后续的恢复——所以这个证明也证明了，纠错总是可以应用这种两步骤方法来实现. 设 $\{E_i\}$ 为满足量子纠错条件式 (10.16) 的一组运算元. 根据假定，α 为 Hermite 矩阵，因而可被对角化为 $d=u^{\dagger}\alpha u$，其中 u 为酉矩阵和 d 为对角矩阵. 定义算子 $F_k\equiv\sum_i u_{ik}E_i$. 回顾定理 8.2，我们看到 $\{F_i\}$ 也是 ε 的一组运算元. 通过直接的代入，得到

$$PF_k^{\dagger}F_lP=\sum_{ij}u_{ki}^{\dagger}u_{jl}PE_i^{\dagger}E_jP \tag{10.17}$$

代入式(10.16),可将其简化为 $PF_k^\dagger F_l P = \sum_{ij} u_{ki}^\dagger \alpha_{ij} u_{jl} P$,且由于 $d = u^\dagger \alpha u$,我们得到

$$PF_k^\dagger F_l P = d_{kl} P \tag{10.18}$$

因为 d_{kl} 是对角化的,所以这可被看作是量子纠错条件式(10.16)的简化形式.

我们应用简化条件式(10.18)和极分解(2.1.10 节)来定义差错症状测量.从极分解,我们看到,对某个酉 U_k 成立 $F_k P = U_k \sqrt{PF_k^\dagger F_k P} = \sqrt{d_{kk}} U_k P$. F_k 的作用因而是把编码子空间旋转到由投影算子 $P_k \equiv U_k P U_k^\dagger = F_k P U_k^\dagger / \sqrt{d_{kk}}$ 所定义的这个子空间中.式(10.18)意味着这些子空间是正交的,因为当 $k \neq l$ 时有

$$P_l P_k = P_l^\dagger P_k = \frac{U_l P F_l^\dagger F_k P U_k^\dagger}{\sqrt{d_{ll} d_{kk}}} = 0 \tag{10.19}$$

差错症状测量是由投影算子 P_k 所定义的一个投影测量,算子 P_k 若必要可通过附加投影算子而被增广以满足完备性关系 $\sum_k P_k = I$. 恢复可简单地通过作用 $U_k^\dagger$ 来实现.为看清这个纠错方法可用,注意到检测 - 恢复组合步骤对应于量子运算 $\mathcal{R}(\sigma) = \sum_k U_k^\dagger P_k \sigma P_k U_k$. 对于这个码中的状态 ρ,简单的代数运算和定义的应用显示:

$$U_k^\dagger P_k F_l \sqrt{\rho} = U_k^\dagger P_k^\dagger F_l P \sqrt{\rho} \tag{10.20}$$

$$= \frac{U_k^\dagger U_k P F_k^\dagger F_l P \sqrt{\rho}}{\sqrt{d_{kk}}} \tag{10.21}$$

$$= \delta_{kl} \sqrt{d_{kk}} P \sqrt{\rho} \tag{10.22}$$

$$= \delta_{kl} \sqrt{d_{kk}} \sqrt{\rho} \tag{10.23}$$

因此,如所要求那样,有

$$\mathcal{R}(\varepsilon(\rho)) = \sum_{kl} U_k^\dagger P_k F_l \rho F_l^\dagger P_k U_k \tag{10.24}$$

$$= \sum_{kl} \delta_{kl} d_{kk} \rho \tag{10.25}$$

$$\propto \rho \tag{10.26}$$

为证明量子纠错条件式(10.16)的必要性,设 $\{E_i\}$ 是通过具有算子元 $\{R_j\}$ 的纠错运算 $\mathcal{R}$ 可完美地纠正的一组差错.定义量子运算 ε_C 为 $\varepsilon_C(\rho) \equiv \varepsilon(P\rho P)$. 因为对所有状态 ρ,$P\rho P$ 都属于码空间,于是对所有状态 ρ,

$$\mathcal{R}(\varepsilon_C(\rho)) \propto P\rho P \tag{10.27}$$

进而,若右边和左边两者都是线性的,比例因子必为不依赖于 ρ 的常数 c. 根据运算元显式地重写这最后一个方程,给出

$$\sum_{ij} R_j E_i P \rho P E_i^\dagger R_j^\dagger = c P \rho P \tag{10.28}$$

这个方程对所有 ρ 都成立. 且可导出，具有运算元 $\{R_j E_i\}$ 的量子运算等同于具有单个运算元 $\sqrt{c}P$ 的量子运算. 而定理 8.2 意味着，存在复数 c_{ki} 使有

$$R_k E_i P = c_{ki} P \tag{10.29}$$

取这个方程的伴随形式得 $PE_i^\dagger R_k^\dagger = c_{ki}^* P$，因此 $PE_i^\dagger R_k^\dagger R_k E_j P = c_{ki}^* c_{kj} P$. 而 $\mathcal{R}$ 为保迹运算，故 $\sum_k R_k^\dagger R_k = I$. 将方程 $PE_i^\dagger R_k^\dagger R_k E_j P = c_{ki}^* c_{kj} P$ 对 k 求和同时，我们推出

$$PE_i^\dagger E_j P = \alpha_{ij} P \tag{10.30}$$

其中，$\alpha_{ij} \equiv \sum_k c_{ki}^* c_{kj}$ 恰好就为复数的 Hermite 矩阵. 这些就是量子纠错条件. □

量子纠错条件的直接验证是一件容易但费时的事. 在 10.4 节和 10.5 节中，我们会来描述一个形式化理论，它采用量子纠错条件作为构造许多有趣码类的切入点，且可避免直接验证量子纠错条件的许多相关困难. 现在，读者应当自己做一下下面的例子，它说明了实际的量子纠错条件.

练习 10.7　考虑 10.1.1 节的三量子比特比特翻转码，其具有相应的投影算子为 $P = |000\rangle\langle 000| + |111\rangle\langle 111|$. 这个码针对的噪声过程具有运算元

$$\{\sqrt{(1-p)^3}I, \sqrt{p(1-p)^2}X_1, \sqrt{p(1-p)^2}X_2, \sqrt{p(1-p)^2}X_3\}$$

其中 p 为一个比特翻转的概率. 注意，这个量子运算并不是保迹的，因为我们已经去掉了相应于两个和三个量子比特上的比特翻转的运算元. 试验证，对这个码和噪声过程的量子纠错条件.

10.3.1　差错的离散化

我们已经讨论了针对一种特定噪声过程的量子信息的保护. 但是，一般来说，我们并不准确知道量子系统遭受的是什么噪声. 如果特定码 C 和纠错运算 $\mathcal{R}$ 能被用于针对全部类型噪声过程的保护，那么这将会非常有用. 幸运的是，量子纠错条件很容易用于严格地提供这类保护.

定理 10.2　设 C 为量子码，$\mathcal{R}$ 为定理 10.1 证明中所构造的纠错运算，用以从具有算子元 $\{E_i\}$ 的噪声过程中恢复. 设 $\mathcal{F}$ 为具有运算元 $\{F_j\}$ 的量子运算，运算元 $\{F_j\}$ 为 E_i 的线性组合，即对某个复数矩阵 m_{ji} 有 $F_j = \sum_i m_{ji} E_i$. 那么，纠错运算 $\mathcal{R}$ 也可对码 C 上的噪声过程 $\mathcal{F}$ 的作用来进行纠正.

盒子 10.1 无测量的量子纠错

在正文中，我们把量子纠错描述为一个两阶段过程：首先采用量子测量来实现的一个差错检测步骤，然后采用条件酉运算来实现的一个恢复步骤. 有可能仅采用酉运算和制备于标准状态的辅助系统来执行量子纠错. 知道如何来做这一点的好处在于，对某些现实世界量子系统，执行对量子纠错所需要的量子测量是很困难的，所以需要有一个替代的方法. 用于做到这点的技术基本上同于第 8 章里对任意量子运算建立模型中所描述的那些；我们现在在量子纠错背景中来重新叙述其基本思想.

设主系统上的差错症状——会被纠正的差错症状——的测量是由测量算子 M_i 所描述的，且相应的条件酉运算为 U_i. 对应于可能的差错症状，引入具有基状态 $|i\rangle$ 的一个辅助系统. 在纠错之前，辅助系统开始于标准纯态 $|0\rangle$. 定义主系统加附属系统上的酉算子为

$$U|\psi\rangle|0\rangle \equiv \sum_i (U_i M_i|\psi\rangle)|i\rangle \tag{10.31}$$

这个定义式可以被推广到作用于整个空间上的酉算子，因为

$$\langle\varphi|\langle 0|U^\dagger U|\psi\rangle|0\rangle = \sum_{ij}\langle\varphi|M_i^\dagger M_j|\psi\rangle\delta_{ij} \tag{10.32}$$

$$= \sum_i \langle\varphi|M_i^\dagger M_i|\psi\rangle \tag{10.33}$$

$$= \langle\varphi|\psi\rangle \tag{10.34}$$

也即，U 保持内积，因而能被扩张为整个空间上的酉算子. U 的作用在于影响正在被纠错的系统上的变换 $\mathcal{R}(\sigma) = \sum_i U_i M_i \sigma M_i^\dagger U_i^\dagger$，这个变换正是正文中，描述的执行量子纠错的量子运算. 注意，为使这个量子纠错方法能工作，有必要在执行每次纠错时采用新的辅助系统.

证 据定理 10.1，运算元 $\{E_i\}$ 必满足量子纠错条件 $PE_iE_j^\dagger P = \alpha_{ij}P$. 如在定理 10.1的证明中所证得的，不失一般性，我们可以假定所选的 ϵ 运算元，使 $\alpha_{ij} = d_{ij}$ 为具有实数元的对角矩阵. 纠错运算 $\mathcal{R}$ 具有运算元 $U_k^\dagger P_k$，其中据式(10.23)，选择 U_k 和 P_k 使对码空间中的任一 ρ 有

$$U_k^\dagger P_k E_i \sqrt{\rho} = \delta_{ki}\sqrt{d_{kk}}\sqrt{\rho} \tag{10.35}$$

代入 $F_j = \sum_i m_{ji}E_i$，得出

$$U_k^\dagger P_k F_j \sqrt{\rho} = \sum_i m_{ji}\delta_{ki}\sqrt{d_{kk}}\sqrt{\rho} \tag{10.36}$$

$$= m_{jk}\sqrt{d_{kk}}\sqrt{\rho} \tag{10.37}$$

因而，如所要求的那样：

$$\mathcal{R}(\mathcal{F}(\rho)) = \sum_{kj} U_k^\dagger P_k F_j \rho F_j^\dagger P_k U_k \tag{10.38}$$

$$= \sum_{kj} |m_{jk}|^2 d_{kk}\rho \tag{10.39}$$

$$\propto \rho \tag{10.40}$$

□

这个结果使得能够引入更为有力的语言以描述量子纠错码. 我们不再讨论由码 C 纠正的差错过程类 ε 和纠错算子 $\mathcal{R}$，而是讨论可纠正的差错算子（或简称为差错）的集合 $\{E_i\}$. 通过这一点，意味着，量子纠错条件对这些算子成立：

$$PE_iE_j^\dagger P = \alpha_{ij} P \tag{10.41}$$

定理 10.1 和定理 10.2 意味着，任一噪声过程 ε，其运算元由这些差错算子 $\{E_i\}$ 的线性组合而成，都将通过恢复运算 $\mathcal{R}$ 可被纠正.

让我们来看应用这个强有力新观点的一个例子. 设 ε 为作用于单量子比特上的量子运算. 那么，其每个运算元 $\{E_i\}$ 都可以被写成为 Pauli 矩阵 $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ 的线性组合. 因此，为检验 Shor 码能对第一量子比特上的单量子比特差错进行纠错，只需验证满足方程

$$P\sigma_i^1\sigma_j^1 P = \alpha_{ij} P \tag{10.42}$$

就足够了，其中 σ_i^1 为作用于第一量子比特上的 Pauli 矩阵（I, X, Y 和 Z）. 一旦满足这点，就会保证第一量子比特上的任何差错过程都可以被纠正（实际的计算是十分简单的，因而将其作为练习 10.10 的一部分）. 事实上，这个例子解释了这样一点，即第一次接触量子纠错文献时的一种神秘感：许多作者都对退极化信道 $\varepsilon(\rho) = (1-p)\rho + p(X\rho X + Y\rho Y + Z\rho Z)/3$ 有一种看起来可疑的喜爱. 很容易会设想，这会大大地限制它们的纠错模型的有效性，但情况并非如此，对像现在刚作的讨论意味着，退极化信道的纠错能力自动地暗示了对任意单量子比特的量子运算的纠错能力.

概括起来，我们已经认识到，有可能使量子差错离散化；为对抗单量子比特上可能的连续差错，只要赢得对有限差错集即四个 Pauli 矩阵的“战争”就足够了. 类似结果对于高维量子系统同样成立. 这反映了与经典模拟系统的纠错理论的显著差异. 这类系统中的纠错是非常不同的，因为原理上存在有无穷数目的不同差错症状. 经典信息处理的数字纠错要成功得多，因为它只包含有限数目的差错症状. 我们已经学过的令人惊讶的事情是，相比于与经典模拟纠错，量子纠错看起来与经典数字纠错类似得多.

练习 10.8　试验证，三量子比特相位翻转码 $|0_L\rangle = |+++\rangle$，$|1_L\rangle = |---\rangle$ 满足对差错算子集合 $\{I, Z_1, Z_2, Z_3\}$ 的量子纠错条件.

练习 10.9 再一次考虑三量子比特相位翻转码，令 P_i 和 Q_i 为分别到第 i 量子比特的 $|0\rangle$ 和 $|1\rangle$ 状态上的投影算子，试证明，三量子比特相位翻转码可对抗差错集 $\{I, P_1, Q_1, P_2, Q_2, P_3, Q_3\}$.

练习 10.10 试对包含 I 和差错算子 X_j, Y_j, Z_j 且 $j=1$ 到 9 的差错集，显式地验证 Shor 码的量子纠错条件.

练习 10.11 试构造单量子比特量子运算 ε 的运算元，使得根据任一状态 ρ 的输入以完全随机的状态 $I/2$ 来替换它. 令人吃惊的是，甚至像这样的噪声模型都可以通过如 Shor 码这样的码而被纠正.

10.3.2 独立差错模型

我们如何能在量子纠错和第 9 章引入的实现可靠量子信息处理的判据之间建立起联系呢？这一节中，利用不同量子比特上差错为独立的假设，我们会解释如何来做到这点的基本思想. 直观上，如果噪声过程独立地作用于码中的不同量子比特上，那么在规定噪声为足够弱的情况下，与非编码状态相比，纠错应当更能改善编码状态的存储保真度. 为说明这一点，我们以退极化信道的例子作为开始，它会对基本思想提供一个特别简单的实证，随后再扩展这种思想来包括其他的重要信道.

回顾，退极化信道可以由一个单参数即概率 p 来描述. 单量子比特上退极化信道的作用可由方程 $\varepsilon(\rho)=(1-p)\rho+(p/3)[X\rho X+Y\rho Y+Z\rho Z]$ 来定义，并可以被解释为一句话，对量子比特什么都没有发生的概率为 $1-p$，算子 X, Y 和 Z 中的每一个作用于量子比特的概率为 $p/3$. 退极化信道特别易于在量子纠错的背景下进行分析，因为它具有根据四个基本差错 I, X, Y 和 Z 来表述的一个方便的解释，而这些基本差错已被最为广泛地应用于量子码的分析中. 我们将会来解释这种分析是如何进行的，然后再回到如下问题，当我们考虑一个没有由 I, X, Y 和 Z 运算简单解释的过程时，情况会怎样. 简单的计算显示，对通过退极化信道发送的状态，最小保真度给出为 $F=\sqrt{1-2p/3}=1-p/3+O(p^2)$.

练习 10.12 试证明，状态 $|0\rangle$ 和 $\varepsilon(|0\rangle\langle 0|)$ 之间的保真度为 $\sqrt{1-2p/3}$. 并用它来论证，退极化信道的保真度为 $\sqrt{1-2p/3}$.

设我们用能纠正任何单量子比特上差错的一个 n 量子比特量子码来对一个单量子比特信息编码. 设具有参数 p 的退极化信道独立地作用于这个量子比特中的每一个，则在所有 n 个量子比特上所引起的联合作用为

$$\varepsilon^{\otimes n}(\rho)=(1-p)^n\rho+\sum_{j=1}^{n}\sum_{k=1}^{3}(1-p)^{n-1}\frac{p}{3}\sigma_k^j\rho\sigma_k^j+\cdots \tag{10.43}$$

其中，“…”表示均为正且在分析时将丢弃的高阶项. 纠错执行以后，只要 ρ 原来处

于码中，出现在这个和式中的所有项将会被回复到状态 ρ：

$$(\mathcal{R}\otimes\varepsilon^{\otimes n})(\rho)=[(1-p)^n+n(1-p)^{n-1}p]\rho+\cdots \tag{10.44}$$

所以保真度满足

$$F\geqslant\sqrt{(1-p)^{n-1}(1-p+np)}=1-\frac{\binom{n}{2}}{2}p^2+O(p^3) \tag{10.45}$$

因此，在规定差错的概率 p 为充分小后，采用量子纠错码会改善被这个码所保护的量子状态保真度.

并非所有噪声污染的信道都可如此容易地解释为由无差错、比特翻转、相位翻转及其两两组合的一个随机组合. 许多自然出现的量子信道都不具备这样的解释. 考虑振幅阻尼的例子(8.3.5 节)，其具有运算元 E_0 和 E_1：

$$E_0=\begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\gamma}\end{bmatrix},\quad E_1=\begin{bmatrix}0 & \sqrt{\gamma}\\ 0 & 0\end{bmatrix} \tag{10.46}$$

参数 γ 为表征幅值阻尼过程强度的一个小的正参数——强度随着 γ 趋于 0 而减少，直到最终我们得到基本无噪声量子信道. 我们可能会合理地猜测，根据一组包括正比于单位阵的项在内的运算元 $\{f(\gamma)I,E_1',E_2',\cdots\}$，幅值阻尼信道具有一个等价的描述，其中随 $\gamma\to0$ 而 $f(\gamma)\to1$. 如果果真如此，那么对独立作用于多个量子比特上的幅值阻尼信道，其纠错的分析可以类似于对退极化信道的纠错分析那样来做. 令人惊讶的是，可以推证，这样的描述根本不可能存在. 这一点可简单地由定理 8.2 来导出，因为对 $\gamma>0$ 没有一个 E_0 和 E_1 的线性组合能在任何时候都正比于单位阵，因而对幅值阻尼信道没有一组运算元可在任何时候都包含正比于单位阵的项.

类似地，量子力学中的许多其他噪声过程在物理意义下总是接近于单位阵的. 然而，没有一个这种过程的算子和描述包含大的单位阵组元. 直观上，似乎是正确的，只要噪声为足够弱，这种形式纠错应当会在量子信息的存储保真度中导致一个净增益. 为具体起见，我们现在可通过，对幅值阻尼信道的特例而获得证明. 简单的计算显示，作用于一个单量子比特的幅值阻尼信道的最小保真度为$\sqrt{1-\gamma}$. 设这个量子比特采用能纠正单量子比特上任意差错的一个 n 量子比特量子码来编码，设参数为 γ 的幅值阻尼信道独立地作用于每个量子比特上. 我们来概略说明这个基本思想，即量子纠错的作用在于改变存储的保真度到 $1-O(\gamma^2)$，所以对小的 γ 以量子码来编码这个量子比特会导致一个对差错的净抑制.

练习 10.13　试证明，当 ε 是具有参数 γ 的幅值阻尼信道时，最小保真度 $F(|\psi\rangle,\varepsilon(|\psi\rangle\langle\psi|))$为$\sqrt{1-\gamma}$.

用 $E_{j,k}$ 表示 E_j 在第 j 个量子比特上的作用，则噪声在编码后量子比特上的影响可写为

$$\varepsilon^{\otimes n}(\rho)=(E_{0,1}\otimes E_{0,2}\otimes\cdots\otimes E_{0,n})\rho(E_{0,1}^{\dagger}\otimes E_{0,2}^{\dagger}\otimes\cdots\otimes E_{0,n}^{\dagger})+$$
$$\sum_{j=1}^{n}\left[E_{1,j}\otimes\left(\bigotimes_{k\neq j}E_{0,k}\right)\right]\rho\left[E_{1,j}^{\dagger}\otimes\left(\bigotimes_{k\neq j}E_{0,k}^{\dagger}\right)\right]+O(\gamma^2) \quad (10.47)$$

设 $E_0=(1-\gamma/4)I+\gamma Z/4+O(\gamma^2)$ 和 $E_1=\sqrt{\gamma}(X+\mathrm{i}Y)/2$，将这些表达式代入到式(10.47)，就给出

$$\varepsilon^{\otimes n}(\rho)=\left(1-\frac{\gamma}{4}\right)^{2n}\rho+\frac{\gamma}{4}\left(1-\frac{\gamma}{4}\right)^{2n-1}\sum_{j=1}^{n}(Z_j\rho+\rho Z_j)+$$
$$\frac{\gamma}{4}\left(1-\frac{\gamma}{4}\right)^{2n-2}\sum_{j=1}^{n}(X_j+\mathrm{i}Y_j)\rho(X_j-\mathrm{i}Y_j)+O(\gamma^2) \quad (10.48)$$

设 ρ 为这个码的状态. 显然，ρ 上的纠错作用是保持其不变. 容易通过考虑在 $Z_j|\psi\rangle\langle\psi|$ 上的作用而来理解像在 $Z_j\rho$ 和 ρZ_j 项上的纠错作用，其中 $|\psi\rangle$ 为这个码的状态. 我们设这个码使差错 Z_j 把 $|\psi\rangle$ 变到正交于码的一个子空间，以便当执行差错症状测量时像 $Z_j|\psi\rangle\langle\psi|$ 那样项就不出现(注意，即使不引入这个正交性假定，通过研究将这个码变到正交子空间上的差错算子，类似的分析仍然可以进行). 因此，像 $Z_j\rho$ 那样的项在纠错以后会消失，像 $Z_j\rho$，$X_j\rho Y_j$ 和 $Y_j\rho X_j$ 那样的项也是如此. 进而，纠错会将 $X_j\rho Y_j$ 和 $Y_j\rho X_j$ 变回到 ρ，因为这个码能纠正一个量子比特上的差错. 因而，在纠错以后系统的状态为

$$\left(1-\frac{\gamma}{4}\right)^{2n}\rho+2n\frac{\gamma}{4}\left(1-\frac{\gamma}{4}\right)^{2n-2}\rho+O(\gamma^2)=\rho+O(\gamma^2) \quad (10.49)$$

因此，准确到阶 γ^2，纠错会使量子系统返回到它的原始状态 ρ，而对弱噪声(小 γ)，如同退极化信道那样，纠错会对差错产生净抑制. 我们这里的分析是针对幅值阻尼噪声模型的，但不难推广这个讨论到其他噪声模型并得到类似的结论. 因此，一般情况下，本章的剩下部分，主要是针对那些特定噪声模型进行讨论. 这些噪声模型可被理解为对应于 Pauli 矩阵差错的随机应用，这些应用类似于退极化信道，这样允许我们采用经典概率论中所熟悉的概念来进行分析. 切记，应用类似于我们刚刚概述过的那些原理，我们所描述的这些思想可以被推广到比这种简单差错模型的应用范围要宽得多的一类差错模型.

10.3.3 简并编码

量子纠错码在许多方面十分类似于经典码——首先通过测量差错症状来识别差错，然后再进行适当地纠正. 但是，有一类称为简并编码(degenerate code)的有趣量子码，其具有经典码中不曾有的一个引人注目的性质. 其思想最容易通过 Shor 码例子来说明. 考虑差错 Z_1 和 Z_2 在 Shor 码的码字上的影响. 如同我们已经表明的，这些差错的影响在两个码字上是相同的. 对经典纠错码，不同比特上的差

错会导致不同的掺杂讹误的码字. 简并量子码的现象是好坏参半的量子码类. 坏的方面是指经典上用于证明纠错界的某些证明技术会失效，因为它们不能被应用于简并编码. 10.3.4 节中，我们将会看到这样一个例子和量子 Hamming 界. 好的方面则指简并量子码看来位于最令人感兴趣的量子码之列. 在某种意义下，简并量子码比之经典码能够“把更多信息封装起来”，因为不同的差错并不必然把码空间映到正交空间，并有可能(尽管还没有被证明)，这种附加的能力会导致简并编码比之任何非简并编码能更有效地来存储量子信息.

10.3.4　量子 Hamming 界

在应用中，我们会愿意采用可能的“最好”量子码. 在给定的情形中，“最好”的意思是依赖于应用的. 由于这个原因，我们希望有判据来判断一种具有特定特性的码是否存在. 在本节中，我们来介绍量子 Hamming 界. 这个简单界提供了关于量子码一般特性的一些灵感. 不幸的是，Hamming 界适用于非简并编码，但它告诉我们更一般的界可能的样子. 设想，以能纠正 t 或更少数目的量子比特上差错的方式，一个非简并编码将 k 量子比特编码为 n 量子比特. 设出现 j 个差错，其中 $j\leqslant t$. 总共有 $\binom{n}{j}$ 组差错的可能出现位置. 对应每组这样的位置，会有三个可能的差错——三个 Pauli 矩阵 X,Y,Z——它们可能出现在每个量子比特中，总共有 3^j 个可能的差错. 在 t 或更少量子比特上出现差错的总个数因而为

$$\sum_{j=0}^{t}\binom{n}{j}3^j \tag{10.50}$$

(注意，$j=0$ 对应于“差错”I，即任何量子比特上没有差错的情况). 为了以非简并编码方式来编码 k 个量子比特，这些差错中的每个都必须对应于一个正交的 2^k 维子空间. 所有这些子空间必须置于对 n 个量子比特可利用的整个 2^n 维空间中，从而导致不等式：

$$\sum_{j=0}^{t}\binom{n}{j}3^j 2^k \leqslant 2^n \tag{10.51}$$

这就是量子 Hamming 界. 举例来说，考虑这样一种情形，其中我们希望以一个量子比特上差错可被容忍的方式，用 n 个量子比特来编码一个量子比特. 在这种情况中，量子 Hamming 界为

$$2(1+3n)\leqslant 2^n \tag{10.52}$$

变换后表明，这个不等式对 $n\leqslant 4$ 不满足，而对 $n\geqslant 5$ 的值满足. 因此，不存在用少于 5 个量子比特对一个量子比特编码的一种非简并编码，这种编码方式能够对抗单量子比特上所有可能差错.

当然，并非所有量子码都是非简并的，所以，量子 Hamming 界主要作为一种

经验来使用，而不是作为量子码存在性的确实的和快速的界（特别是，在本书写作的时候，还没有发现违反量子 Hamming 界的码，即使甚至允许简并编码）. 后面，我们将会有机会来看某些适用于所有量子码的界，而不仅是非简并编码. 举例来说，在 12.4.3 节中，我们会证明量子单一界，它意味着，用 n 个量子比特来编码 k 个量子比特并能纠正任何 t 个量子比特上差错的任何量子码都必须满足 $n \geqslant 4t+k$. 基此导出，编码一个量子比特并能纠正一个量子比特上任意差错的最小码必须满足 $n \geqslant 4+1=5$，而事实上我们不久的将会来展示这样一个五量子比特码.

10.4 量子码的构造

我们现在对研究量子纠错码有了一个理论框架，但是我们至今还没有这些码的多个例子. 10.4.1 节中，通过对经典线性码理论所作的简短回顾，我们开始来弥补这个缺陷. 随后，10.4.2 节解释，经典线性码的思想如何用来构造一大类称为 Calderbank-Shor-Steane(CSS)码的量子码. 10.5 节引入稳定子码(stabilizer code)来结束本章内容. 这是一类甚至比之 CSS 码更为一般的码，它为构造各种不同类型的量子码提供强有力的手段.

10.4.1 经典线性码

经典纠错码有着许多各种各样的重要技术应用，所以毫不奇怪对这种码已发展出一套强有力的理论. 我们在经典纠错码的技术方面的兴趣在于，这些技术中的许多对量子纠错都具有重要的意义，特别是经典线性码理论，可以用于发展很多种好的量子纠错码. 在本节中，我们来对经典线性码作些复习，特别强调对量子纠错来说非常重要的那些思想.

编码 k 个量子比特信息到 n 个比特码空间的一个线性码 C，可由一个 $n \times k$ 生成矩阵 G 来指定，G 的所有元属于 $\boldsymbol{Z}_2$ 即 0 和 1. 矩阵 G 将消息映射到其编好的码. 因此，k 个比特消息 x 被编码为 Gx，其中消息 x 被显式地当成为一个列向量. 幸运的是，乘法运算以及本节中的所有其他算术运算都是按模 2 来做的. 一个简单的例子是，将单量子比特映射到三个重复比特的重复码可由如下生成矩阵来给定：

$$G = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{10.53}$$

因为 G 映射可能的消息 0 和 1 到其编码后的形式 $G[0]=(0,0,0)$ 和 $G[1]=(1,1,1)$.（回忆先前，$(a,b,\cdots,z)$ 为我们对列向量的简写符号）. 我们说，用 n 个比特编码 k 个比特信息的码为一个 $[n,k]$ 码；因而这个例子就是一个 $[3,1]$ 码. 稍微复杂一

些的例子是，用每个比特的三次重复来编码两个比特——一个[6,2]码. 这个例子具有生成矩阵：

$$G=\begin{bmatrix}1 & 0\\ 1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 1\\ 0 & 1\end{bmatrix} \tag{10.54}$$

由此，我们看到

$$G(0,0)=(0,0,0,0,0,0),\quad G(0,1)=(0,0,0,1,1,1) \tag{10.55}$$

$$G(1,0)=(1,1,1,0,0,0),\quad G(1,1)=(1,1,1,1,1,1) \tag{10.56}$$

这恰好就是我们所期望的. 这个码的可能码字的集合对应于由 G 的列所张成的向量空间，所以为使所有消息都可被惟一地编码，我们要求 G 的列为线性无关，除此之外其他方面对 G 没有限制.

练习 10.14　试写出，采用重复每个比特 r 次来编码 k 个比特的生成矩阵的表达式. 这是一个$[rk,k]$线性码，并应具有 $rk\times k$ 的生成矩阵.

练习 10.15　试证明，在一个生成矩阵中把 G 的一个列加到另一个列上去会导致生成相同码的生成矩阵.

线性码比之一般纠错码的一大优点是它们的紧凑表示. 一个用 n 个比特编码 k 个比特的一般码要求 2^k 个长度均为 n 的码字来指定编码，总共要 $n2^k$ 个比特来指定这个码的描述. 对于线性码，我们只需要指定生成矩阵的 kn 个比特，这在所要求的存储量上是一个指数的节省. 这种紧凑的描述反映在进行有效的编码和解码的能力上，这是经典线性码与其量子对应物稳定子码所共有的重要特性. 我们已可看到，如何来执行一个经典线性码的有效编码：简单地用 $n\times k$ 生成矩阵 G 乘以 k 比特消息来得到 n 比特编码后的消息，这是一种用 $O(nk)$ 步运算能够完成的方法.

线性码的生成矩阵定义的有吸引力的特性之一是，在我们想要编码的消息和它们如何被编码之间有一种明显的关系，但执行纠错不是很清楚. 按线性码的另一种替代(但等价)的奇偶检验矩阵形式，线性码的纠错最容易理解. 在这个定义中，一个$[n,k]$码定义为由 $\boldsymbol{Z}_2$ 上使

$$Hx=0 \tag{10.57}$$

成立的所有 n 元向量 x 来组成. 其中 H 称为奇偶检验矩阵(parity check matrix)的一个$(n-k)\times n$ 矩阵，其元均为 0 和 1. 等价地，但更为简洁地，这个码可被定义为 H 的核. 一个编码 k 比特的码具有 2^k 个可能的码字，所以 H 的核必是 k 维的，因此我们要求 H 为行线性无关.

练习 10.16　试证明，把奇偶检验矩阵的一行加到另一行上去不会改变这个

码. 应用高斯消去法和比特的交换，就能够来假定，奇偶检验矩阵具有标准形 $[A|I_{n-k}]$，其中 A 为 $(n-k)\times k$ 矩阵.

为了将线性码的奇偶检验描述与生成矩阵描述联系起来，我们需要来研究一种方法，以使我们能在奇偶检验矩阵 H 和生成矩阵 G 之间进行双向转换. 为从奇偶检验矩阵转换到生成矩阵，挑选 k 个线性无关的向量 $y_1,\cdots,y_k$ 来张成 H 的核，并使 G 具有列 y_1 直到 y_k. 为从生成矩阵去到奇偶检验矩阵，挑选 $n-k$ 个正交于 G 各列的线性无关的向量 $y_1,\cdots,y_{n-k}$，并使 H 的行为 $y_1^{\mathrm{T}},\cdots,y_{n-k}^{\mathrm{T}}$（这里正交指模 2 内积必须为 0）. 作为一个例子，考虑由生成矩阵式(10.53)定义的[3,1]重复码. 为构造 H，我们挑选 3−1=2 个正交于 G 各列的线性无关向量，比如说(1,1,0)和(0,1,1)，并定义奇偶检验矩阵为

$$H\equiv\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 1\end{bmatrix} \tag{10.58}$$

容易检验，$Hx=0$ 仅对码字 $x=(0,0,0)$ 和 $x=(1,1,1)$ 成立.

练习 10.17 试对由式(10.54)中生成矩阵所定义的[6,2]重复码求一个奇偶检验矩阵.

练习 10.18 试证明，同一个线性码的奇偶检验矩阵 H 和生成矩阵 G 满足 $HG=0$.

练习 10.19 设对某个 $(n-k)\times k$ 矩阵 A，一个[n,k]线性码 C 具有形为 $H=[A|I_{n-k}]$ 的奇偶检验矩阵. 试证明，相应的生成矩阵为

$$G=\begin{bmatrix}I_k\\ \hline -A\end{bmatrix} \tag{10.59}$$

（注意 $-A=A$，因为我们考虑的是模 2 情形；但是，这个方程对比 $\boldsymbol{Z}_2$ 更为一般的域上的线性码也同样成立）.

奇偶检验矩阵使差错检测和恢复变得十分明显. 设我们编码消息 x 为 $y=Gx$，但由噪声引起的一个差错 e 使 y 受到污染，并给出污染后的码字 $y'=y+e$（注意，“+”号在这里表示比特方式的模 2 加）. 因为对所有码字 $Hy=0$，由此导出 $Hy'=He$. 我们称 Hy' 为差错症状，它扮演类似于量子纠错中差错症状所扮演的角色；它是污染后状态 y' 的函数 Hy'，如同量子差错症状是通过测量污染后量子状态所决定的. 由于关系 $Hy'=He$，差错症状包含有关出现的差错的信息，将能像所希望那样来恢复到原来的码字 y. 为看清这如何会是可能的，设没有差错或仅有一个差错出现. 那么，差错症状 Hy' 在没有差错的情况中等于 0，而当一个差错出现在第 j 个比特上时等于 He_j，其中 e_j 是第 j 个元为 1 的单位向量. 如果我们假定差错最多出现在一个比特上，那么有可能通过计算差错症状 Hy' 并将其与不同的 He_j 值比较，以确定哪个比特（如果有的话）需要被纠正，而来执行纠错.

更为一般地，要深入了解纠错是如何利用线性码可被执行的，这可以通过距离

的概念而获知. 设 x 和 y 各为 n 比特的字. x 和 y 之间的(Hamming)距离 $d(x,y)$ 定义为 x 和 y 有差异的位置的数目. 举例来说,$d((1,1,0,0),(0,1,0,1))=2$. 字 x 的(Hamming)权重定义为全零字串到其的距离 $\mathrm{wt}(x)\equiv d(x,0)$,也即 x 中非零位置的数目. 注意,$d(x,y)=\mathrm{wt}(x+y)$. 为理解与纠错的联系,设我们采用一种线性纠错码,对 x 编码为 $y=Gx$. 经噪声污染后输出 $y'=y+e$. 设一个量子比特翻转的概率小于 1/2,已编码的最可能码字为 y,这个码字 y 得到 y' 所需要的比特翻转数目最少,即使 $\mathrm{wt}(e)=d(y,y')$ 取极小. 原则上,应用线性码的纠错,可通过简单地以这样一个 y 替换 y' 来实现. 实际上,这种做法是相当没有效率的,因为确定最小距离 $d(y,y')$ 一般来说需要搜索所有 2^k 个可能的码字 y. 经典码理论中,在构造具有特殊结构的码上已经作出了大量的努力,这些码能使纠错更有效地来执行. 这些构造已超出了本书的范围.

这个码的全局性质也可以用 Hamming 距离来理解. 我们定义,一个码的距离为其任意两个码字之间的最小距离:

$$d(C)\equiv \min_{x,y\in C,x\neq y} d(x,y) \tag{10.60}$$

但是,$d(x,y)=\mathrm{wt}(x+y)$. 由于码是线性的,若 x 和 y 为码字则 $x+y$ 也是码字,所以我们看到

$$d(C)=\min_{x\in C,x\neq 0}\mathrm{wt}(x) \tag{10.61}$$

令 $d\equiv d(C)$,则 C 为一个 $[n,k,d]$ 码. 距离的重要性在于,对某个整数 t,简单地通过解码受污染的编码消息 y' 为满足 $d(y,y')\leqslant t$ 的惟一码字 y,具有距离至少为 $2t+1$ 的一个码就能来纠正最多 t 个比特上的差错.

练习 10.20　令 H 是任意 $d-1$ 个列为线性无关的一个奇偶检验矩阵,但存在一组 d 个线性相关列,试证明,由 H 所定义的码具有距离 d.

练习 10.21(单一界)　试证明,一个 $[n,k,d]$ 码必满足 $n-k\geqslant d-1$.

Hamming 码是一类便于说明的线性纠错码. 设 $r\geqslant 2$,为一个整数. 令 H 为一个矩阵,其列为全部长度为 r 的 2^r-1 个不全为 0 的比特串. 这个奇偶检验矩阵定义了一个称为 Hamming 码的 $[2^r-1,2^r-r-1]$ 线性码. 对于量子纠错,一个特别重要的例子是 $r=3$ 的情况,它是一个 $[7,4]$ 码,具有如下奇偶检验矩阵:

$$H=\begin{bmatrix}0&0&0&1&1&1&1\\0&1&1&0&0&1&1\\1&0&1&0&1&0&1\end{bmatrix} \tag{10.62}$$

H 的任意两列为不相同,因而为线性无关;前三列为线性相关,所以据练习 10.20 知这个码的距离为 3. 因此,这个码能来纠正任意单比特上的差错. 事实上,这个纠错方法非常简单. 设一个差错出现在第 j 个比特上,对式(10.62)的检查显示,差错症状 He_j 正好就是 j 的二进位表示,它告诉了哪个比特翻转以纠正差错.

练习 10.22 试证明，所有 Hamming 码都具有距离 3，因而能来纠正单比特上的差错，Hamming 码因此为$[2^r-1,2^r-r-1,3]$码.

线性码更为一般的性质还有什么呢？特别是，我们希望得到能够告诉我们具有特殊码参数的码是否存在的条件. 毫不惊讶，有许多证明这些条件的方法. 所知的 Gilbert-Vershamov 界就是这样的一组条件. 它规定，对大的 n，必存在对某个 k 防止 t 比特上差错的$[n,k]$纠错码，使成立：

$$\frac{k}{n} \geqslant 1 - H\left(\frac{2t}{n}\right) \tag{10.63}$$

其中 $H(x) \equiv -x\log(x)-(1-x)\log(1-x)$为二元 Shannon 熵，第 11 章会对其作详细研究. Gilbert-Vershamov 界的重要性在于，在假定人们并不试图编码太多比特(k)到太少数目的比特(n)情况下，它会保证好码的存在性. Gilbert-Vershamov 界的证明十分简单，留作为一个练习.

练习 10.23 试证明 Gilbert-Vershamov 界.

我们现在通过对一些码的重要构造即对偶构造的解释，来结束对经典纠错码的概述. 设 C 为一个$[n,k]$码，其具有生成矩阵 G 和奇偶检验矩阵 H，那么，可定义另一个码，C 的对偶码并记为 $C^{\perp}$，其具有生成矩阵 G^{T} 和奇偶检验矩阵 H^{T}. 等价地，C 的对偶码由正交于 C 中所有码字的全部码字 y 组成. 称一个码为弱自对偶(weakly self-dual)，如果 $C \subseteq C^{\perp}$；称一个码为严格自对偶(strictly self-dual)，如果 $C=C^{\perp}$. 相当引人注目的是，对经典线性码的对偶构造会自然地出现于量子纠错的研究中，而且是构造 CSS 码的一类重要量子码的关键.

练习 10.24 试证明，具有生成矩阵 G 的一个码为弱自对偶，当且仅当 $G^{\mathrm{T}}G=0$.

练习 10.25 令 C 为一个线性码. 试证明，如果 $x \in C^{\perp}$，那么 $\sum_{y \in C}(-1)^{xy} = |C|$；如果 $x \notin C^{\perp}$，那么 $\sum_{y \in C}(-1)^{xy} = 0$.

10.4.2 Calderbank-Shor-Steane 码

量子纠错码大类中的第一个例子是 Calderbank-Shor-Steane 码，通常更多地被称为 CSS 码，这是以码发明者姓名的首字母所命名的. CSS 码是更为一般类稳定子码的一个重要子类.

设 C_1 和 C_2 为$[n,k_1]$和$[n,k_2]$经典线性码，使有 $C_2 \subset C_1$，且 C_1 和 $C_2^{\perp}$，两者可纠正 t 个差错. 通过下面的构造，我们将要定义能纠正 t 个量子比特上差错的一个$[n,k_1-k_2]$量子码 $\mathrm{CSS}(C_1,C_2)$，即 C_2 上 C_1 的 CSS 码. 设 $x \in C_1$ 为 C_1 中的任一码字，那么，我们就定义量子状态$|x+C_2\rangle$为

$$|x+C_2\rangle \equiv \frac{1}{\sqrt{|C_2|}}\sum_{y \in C_2}|x+y\rangle \tag{10.64}$$

其中，“+”为按比特的模 2 方式加. 设 x' 为 C_1 的一个元，使有 $x-x'\in C_2$. 那么，容易看到 $|x+C_2\rangle=|x'+C_2\rangle$，因此状态 $|x+C_2\rangle$ 只依赖于 x 所在的陪集 C_1/C_2，这同时解释了我们已用于 $|x+C_2\rangle$ 的陪集符号. 进而，如果 x 和 x' 属于 C_2 的不同陪集，那么不存在 $y,y'\in C_2$，使得 $x+y=x'+y'$，因而 $|x+C_2\rangle$ 和 $|x'+C_2\rangle$ 为正交状态. 量子码 $\mathrm{CSS}(C_1,C_2)$ 就定义为由所有 $x\in C_1$ 的状态 $|x+C_2\rangle$ 所张成的向量空间. C_1 中 C_2 的陪集的数目为 $|C_1|/|C_2|$，所以 $\mathrm{CSS}(C_1,C_2)$ 的维数为 $|C_1|/|C_2|=2^{k_1-k_2}$，因此 $\mathrm{CSS}(C_1,C_2)$ 是一个 $[n,k_1-k_2]$ 量子码.

可以利用 C_1 和 $C_2^{\perp}$ 的经典纠错性质来检测和纠正量子差错. 事实上，通过分别利用 C_1 和 $C_2^{\perp}$ 的纠错性质，有可能对 $\mathrm{CSS}(C_1,C_2)$ 上最多 t 个比特翻转差错和相位翻转差错进行纠错. 设比特翻转差错由 n 比特向量 e_1 来描述，且在比特翻转出现的比特上为 1，在其他比特上为 0. 相位翻转差错由 n 比特向量 e_2 来描述，且在相位翻转出现的比特上为 1，在其他位上为 0. 如果 $|x+C_2\rangle$ 为原始状态，那么受污染后的状态为

$$\frac{1}{\sqrt{|C_2|}}\sum_{y\in C_2}(-1)^{(x+y)e_2}\,|x+y+e_1\rangle \tag{10.65}$$

为检测比特翻转出现在什么地方，方便的做法是引入一个辅助码，包含足够多量子比特来存储码 C_1 的差错症状，且初始处于全 0 状态 $|0\rangle$. 我们采用可逆计算，即对码 C_1 应用奇偶矩阵 H_1，把 $|x+y+e_1\rangle|0\rangle$ 变到 $|x+y+e_1\rangle|H_1(x+y+e_1)\rangle=|x+y+e_1\rangle|H_1e_1\rangle$，因为通过奇偶检验矩阵 $x+y\in C_1$ 可消去. 这个运算的作用是来产生状态

$$\frac{1}{\sqrt{|C_2|}}\sum_{y\in C_2}(-1)^{(x+y)e_2}\,|x+y+e_1\rangle\,|H_1e_1\rangle \tag{10.66}$$

练习 10.26　设 H 为一个奇偶检验矩阵，试解释，用完全由受控非门组成的线路如何来计算变换 $|x\rangle|0\rangle\to|x\rangle|Hx\rangle$.

对比特翻转差错的差错检测是通过测量辅助码得到结果 H_1e_1 并消去辅助码而完成的，并给出状态为

$$\frac{1}{\sqrt{|C_2|}}\sum_{y\in C_2}(-1)^{(x+y)e_2}\,|x+y+e_1\rangle \tag{10.67}$$

得到差错症状 H_1e_1 后，就可以推断差错 e_1，因为 C_1 能纠正最多 t 个差错，这就完成了差错检测. 恢复可简单地通过对差错 e_1 中出现比特翻转的位置上应用非门而执行，消去所有比特翻转差错后给出状态

$$\frac{1}{\sqrt{|C_2|}}\sum_{y\in C_2}(-1)^{(x+y)e_2}\,|x+y\rangle \tag{10.68}$$

为检测相位翻转差错，我们对每个量子比特应用 Hadamard 门，把状态变为

$$\frac{1}{\sqrt{|C_2|\,2^n}}\sum_{z}\sum_{y\in C_2}(-1)^{(x+y)(e_2+z)}\,|z\rangle \tag{10.69}$$

其中，求和取遍 n 个量子比特 z 的所有可能值. 令 $z' \equiv z + e_2$，则这个状态可重写为

$$\frac{1}{\sqrt{|C_2|2^n}} \sum_{z'} \sum_{y \in C_2} (-1)^{(x+y)z'} |z' + e_2\rangle \tag{10.70}$$

（下一步出现在练习 10.25 中）设 $z' \in C_2^{\perp}$，则容易看到 $\sum_{y \in C_2} (-1)^{y \cdot z'} = |C_2|$，而若 $z' \notin C_2^{\perp}$ 则 $\sum_{y \in C_2} (-1)^{yz'} = 0$. 因此，状态可重写为

$$\frac{1}{\sqrt{2^n/|C_2|}} \sum_{z' \in C_2^{\perp}} (-1)^{xz'} |z' + e_2\rangle \tag{10.71}$$

这个关系式看上去正好就是由向量 e_2 描述的比特翻转差错. 至于比特翻转的差错检测，我们引入一个辅助码并对 $C_2^{\perp}$ 逆向地应用奇偶检验矩阵 H_2 以得到 $H_2 e_2$，并纠正比特翻转差错 e_2，得到状态为

$$\frac{1}{\sqrt{2^n/|C_2|}} \sum_{z' \in C_2^{\perp}} (-1)^{xz'} |z'\rangle \tag{10.72}$$

纠错可通过再一次对每个量子比特应用 Hadamard 门而来完成；我们可以或者直接计算这些门的结果，或者注意到 $e_2 = 0$ 时式(10.71)中的状态应用 Hadamard 门的作用；由于 Hadamard 门是自可逆的，这就回到了 $e_2 = 0$ 时式(10.68)中的状态：

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} |x + y\rangle \tag{10.73}$$

这就是初始的编码后的状态.

CSS 码的一个重要应用是证明 Gilbert-Varshamov 界的量子版本，这个 Gilbert-Varshamov 界可保证好量子码的存在. 这可陈述为，当 n 变大达到极限时，对某个 k 存在防止最多 t 个量子比特上差错的$[n,k]$量子码，使有

$$\frac{k}{n} \geqslant 1 - 2H\left(\frac{2t}{n}\right) \tag{10.74}$$

因此，假设人们并不试图填塞太多的量子比特 k 到 n 个量子比特码中，那么好的纠错量子码是存在的. CSS 码的 Gilbert-Varshamov 界的证明要比经典 Gilbert-Varshamov 界的证明更为复杂，原因是经典码 C_1 和 C_2 中引入的一些限制，其证明留作为一个章末问题.

概括起来，设 C_1 和 C_2 分别为$[n,k_1]$和$[n,k_2]$经典线性码并使 $C_2 \subset C_1$，C_1 和 $C_2^{\perp}$ 两者可纠正最多 t 个比特上的差错. 则 CSS(C_1, C_2)是一个$[n, k_1 - k_2]$量子纠错码，它能来纠正最多 t 个比特上的任意差错. 进而，差错检测和纠正步骤仅要求应用 Hadamard 门和受控非门，在每种情况中门的数目线性于码的大小. 编码和解码也可应用一些线性于码大小的门，但这里不进行讨论；对它将在 10.5.8 节中进行一般性的讨论.

练习 10.27　试证明，由下式定义的码：

$$|x+C_2\rangle \equiv \frac{1}{\sqrt{|C_2|}}\sum_{y\in C_2}(-1)^{uy}|x+y+v\rangle \tag{10.75}$$

及其通过 u 和 v 的参数化，在具有同样纠错性质的意义下，等价于 $\mathrm{CSS}(C_1,C_2)$. 这些码称为 $\mathrm{CSS}_{u,v}(C_1,C_2)$，它们在后续 12.6.5 节里量子密钥分配研究中将会有用.

Steane 码

CSS 码的一个重要例子可以采用[7,4,3]Hamming 码来构造，其奇偶检验矩阵重新给出为

$$H=\begin{bmatrix}0&0&0&1&1&1&1\\0&1&1&0&0&1&1\\1&0&1&0&1&0&1\end{bmatrix} \tag{10.76}$$

设我们标记这个码为 C，定义 $C_1\equiv C$ 和 $C_2\equiv C^{\perp}$. 为使用这些码来定义一个 CSS 码，首先需要检查 $C_2\subset C_1$. 根据定义，$C_2=C^{\perp}$ 的奇偶检验矩阵等于 $C_1=C$ 的生成矩阵的转置即

$$H[C_2]=G[C_1]^{\mathrm{T}}=\begin{bmatrix}1&0&0&0&0&1&1\\0&1&0&0&1&0&1\\0&0&1&0&1&1&0\\0&0&0&1&1&1&1\end{bmatrix} \tag{10.77}$$

练习 10.28　试验证，式(10.77)中的矩阵的转置是[7,4,3]Hamming 码的生成矩阵.

与式(10.76)比较，我们看到 $H[C_2]$ 的行张成的空间严格包含 $H[C_1]$ 的行张成的空间，且由于相应的码是 $H[C_2]$ 和 $H[C_1]$ 的核，因此我们得出结论 $C_2\subset C_1$. 进而 $C_2^{\perp}=(C^{\perp})^{\perp}=C$，所以 C_1 和 C_2 两者均为距离 3 码，它可来纠正 1 比特上的差错. 因为 C_1 是[7,4]码，而 C_2 是[7,3]码，基此导出 $\mathrm{CSS}(C_1,C_2)$ 是能纠正单量子比特上差错的一个[7,1]量子码.

这个[7,1]量子码具有很好的性质，这使它很容易驾驭，并将被用于本章其余部分的许多例子中. 这个码就是以发明者命名的 Steane 码. C_2 的码字可通过式(10.77)和练习 10.28 来容易地确定. 我们现在非显式地来写出它们，如对 Steane 码的逻辑 $|0_{\mathrm{L}}\rangle$ 元 $|0+C_2\rangle$，有

$$\begin{aligned}|0_{\mathrm{L}}\rangle=\frac{1}{\sqrt{8}}[&|0000000\rangle+|1010101\rangle+|0110011\rangle+|1100110\rangle+\\&|0001111\rangle+|1011010\rangle+|0111100\rangle+|1101001\rangle]\end{aligned} \tag{10.78}$$

为确定其他的逻辑码字，我们需要来找出 C_1 的不在 C_2 中的一个元. 这样元的一个例子为(1,1,1,1,1,1,1)，给出

$$|1_L\rangle = \frac{1}{\sqrt{8}}[|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle + |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle] \quad (10.79)$$

10.5 稳定子码

我们不能克隆，我们来分裂
相干性，保护它避免错误
那会破坏我们有价值的量子比特
还会使我们的计算花时太长.

纠正翻转和相位——那已足够
如果另外的差错出现，在我们的码中
那就简单地测量它，再让上帝去掷骰子
让它坍缩到 X 或 Y 或 Z.

我们起始于噪声污染的七，九或五
结束于完美的一. 为更好消除
那些我们必须避免的错误，我们首先必须努力
以找出哪些可以对易，哪些不能.

利用群和本征态，我们已经学会确定
你的量子差错，用我们的量子窍门.

——Daniel Gottesman 的《量子纠错十四行诗》

稳定子码，有时也被称为加性量子码，是一类重要的量子码，其构造类似于经典线性码. 为了理解稳定子码，首先有必要介绍稳定子体系(stabilizer formalism). 稳定子体系，是一种强有力的方法，用以理解量子力学中运算类. 稳定子体系的应用远超量子纠错范围；但是，在本书中我们的主要关注点是在这种特殊应用上. 在定义稳定子体系以后，我们会来解释，如何可以用它来描述酉门和测量，以及可定量确定稳定子运算局限性的一个重要定理. 随后，我们会来介绍稳定子码的稳定子构造，以及一些显式例子、一种有用的标准形和用以编码、解码和纠错的线路.

10.5.1 稳定子体系

稳定子体系的中心思路可通过一个例子容易地来说明. 考虑双量子比特的 EPR 态：

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{10.80}$$

容易验证，这个状态满足恒等式 $X_1X_2|\psi\rangle = |\psi\rangle$ 和 $Z_1Z_2|\psi\rangle = |\psi\rangle$；我们就说，这个状态被算子 X_1X_2 和 Z_1Z_2 稳定. 有点不太显然的是，状态 $|\psi\rangle$ 是用这些算子 X_1X_2 和 Z_1Z_2 稳定的惟一量子状态(除了一个全局相位). 稳定子体系的基本思想是，比之显式地研究状态自身，许多量子状态可通过稳定它们的算子更容易地来描述. 这个论断初次看来或许会令人惊讶，然而它确实是正确的. 这就导出，比之用状态向量描述，许多量子码(包括 CSS 码和 Shor 码)可用稳定子得到更为简洁的描述. 甚至更为重要的是，量子比特上的差错，各种运算如 Hadamard 门、相位门和甚至受控非门，以及计算基中的测量，所有这一切都可以用稳定子形式来容易地描述.

使稳定子体系强有力的关键在于群论的灵活运用，群论基本原理的复习请见于《量子计算与量子信息(一)》附录 B. 最主要的群是 n 个量子比特上的 Pauli 群 G_n. 对于单量子比特，Pauli 群定义为由所有 Pauli 矩阵与 $\pm 1, \pm i$ 相乘所组成：

$$G_1 \equiv \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\} \tag{10.81}$$

这个矩阵集合构成矩阵乘运算下的一个群. 读者可能会感到奇怪，为什么我们不去掉这些乘子 $\pm 1, \pm i$；之所以包含这些因子的理由是，来保证 G_1 在乘法下为封闭，并因此构成一个合法的群. n 个量子比特上的一般 Pauli 群定义为由 Pauli 矩阵的所有 n 重张量积所组成，且我们再次允许有乘子 $\pm 1, \pm i$.

现在我们就能更为精确地来定义稳定子. 设 S 为 G_n 的一个子群，定义 V_S 为由 S 的每个元所固定的 n 量子比特状态的集合. V_S 为由 S 所稳定的向量空间，S 被称为空间 V_S 的稳定子，因为 V_S 的每个元为在 S 中元的作用下是稳定的. 读者应当确信如下简单练习所叙述的事实.

练习 10.29　试证明，V_S 的任意两个元的任意线性组合也必位于 V_S 中，因此，V_S 是 n 量子比特状态空间的一个子空间. 再证明，V_S 为由 S 中每个算子所固定的子空间的交(也即是 S 的特征值 1 的特征空间的元素).

让我们来看应用稳定子体系的一个简单例子，这是一种 $n=3$ 量子比特和 $S=\{I, Z_1Z_2, Z_2Z_3, Z_1Z_3\}$ 的情况. Z_1Z_2 所固定的子空间由 $|000\rangle, |001\rangle, |110\rangle$ 和 $|111\rangle$ 来张成，Z_2Z_3 所固定的子空间由 $|000\rangle, |100\rangle, |011\rangle$ 和 $|111\rangle$ 来张成. 注意，对于这两个列表，元 $|000\rangle$ 和 $|111\rangle$ 是共同的. 利用这些事实稍加思考容易知道，V_S 必是由状态 $|000\rangle$ 和 $|111\rangle$ 所张成的子空间.

这个例子中，我们通过考虑由 S 中两个算子所稳定的子空间而来简单地确定 V_S. 这显示了一个重要的一般现象——通过其生成元来描述群. 如同《量子计算与量子信息(一)》附录 B 中解释的那样，如果 G 的每个元都可被写为序列 $g_1, \cdots, g_l$ 的元的一个乘积，群 G 中的一组元 $g_1, \cdots, g_l$ 被说成为来生成这个群 G，写为 $G=\langle g_1, \cdots, g_l\rangle$. 在例子中，由于 $Z_1Z_3 = (Z_1Z_2)(Z_2Z_3)$ 和 $I = (Z_1Z_2)^2$，所以 $S=\langle Z_1Z_2,$

$Z_2Z_3\rangle$. 采用生成元描述群的一大优点在于，它们为描述群提供了一个简洁的手段. 事实上，我们在《量子计算与量子信息(一)》附录 B 中证得，大小为 $|G|$ 的一个群 G 具有最多为 $\log(|G|)$ 个的一组生成元. 进而，为看清一个特定的向量可用群 S 来镇定，我们只需要检验向量可用生成元稳定，因为它自动就会由生成元的乘积稳定，这就对其提供了一个最为方便的表示(我们对群生成元所采用的符号$\langle\cdots\rangle$，可能会与 2.2.5 节开始处所引入的可观测平均量的符号相混乱；但是，实际上，从符号所应用的上下文关系总还是清楚的).

并非 Pauli 群的任一子群 S 都可被用作非平凡向量空间的稳定子. 举例来说，考虑由$(\pm I,\pm X)$组成的 G_1 的子群，显然$(-I)|\psi\rangle=|\psi\rangle$只有解$|\psi\rangle=0$，因而，$(\pm I,\pm X)$是平凡向量空间的稳定子. 为使 S 稳定一个非平凡向量空间 V_S，S 必须满足什么条件呢？容易看到，两个必要的条件是：(1)S 的元为可对易；(2)$-I$ 不是 S 的一个元. 虽然我们尚不具备证明的所有工具，但是随后将会证明这两个条件对于 V_S 为非平凡也是充分条件.

练习 10.30 试证明，$-I\notin S$ 意味着 $\pm iI\notin S$.

为证明这两个条件是必要的，设 V_S 为非平凡，所以它包含一个非零向量$|\psi\rangle$. 令 M 和 N 为 S 的元，则 M 和 N 为 Pauli 矩阵的张量积，且可能具有一个总乘子. 由于 Pauli 矩阵全部均与另一个为对易或反对易，基此导出，M 和 N 必或为对易或反对易. 为建立条件(1)即它们为对易，我们反设 M 和 N 为反对易，并来证明这会导致矛盾. 据假定 $-NM=MN$，所以我们有 $-|\psi\rangle=-NM|\psi\rangle=MN|\psi\rangle=|\psi\rangle$，其中第一个等式和最后一个等式由 M 和 N 可稳定$|\psi\rangle$的事实所导出. 于是，我们有 $-|\psi\rangle=|\psi\rangle$，这意味着$|\psi\rangle$为零向量，这就如预期的那样，导出了矛盾. 为建立条件(2)即 $-I\notin S$，只需注意，如果 $-I$ 为 S 的一个元，那么我们有 $-|\psi\rangle=|\psi\rangle$，这又一次会导致矛盾.

练习 10.31 设 S 为由元 $g_1,\cdots,g_l$ 生成的 G_n 的一个子群. 试证明，当且仅当 g_i 和 g_j 对每个对 i,j 为对易 S 的所有元对易.

稳定子体系的一个精彩例子是 7 量子比特 Steane 码. 事实上列于图 10.6 中的 6 个生成元 g_1 到 g_6 对 Steane 码的码空间生成一个稳定子. 当与由状态向量描述的、有点凌乱的表达式(10.78)和式(10.79)相比时，可以观察到这个描述是如何的清晰和简洁；当我们从稳定子体系来考察量子纠错时，甚至还会发现不少更多的优点. 还可注意到，图 10.6 中的生成元和用于构造 Steane 码中的线性经典码 C_1 和 $C_2^{\perp}$ 的奇偶检验矩阵之间的相似性(回顾，对于 Steane 码，$C_1=C_2^{\perp}$ 就为 Hamming[7,4,3]码，其奇偶检验矩阵由式(10.76)给出). 这个稳定子的前三个生成元中 X 的位置，与 C_1 的奇偶检验矩阵中 1 的位置相对应，后三个生成元 g_4 到 g_6 中 Z 的位置，与 $C_2^{\perp}$ 的奇偶检验矩阵中 1 的位置相对应. 有了这些结论，下面这个练习的解几乎就是不证自明的了.

练习 10.32　试验证，如同 10.4.2 节中所描述的，图 10.6 中的生成元可稳定 Steane 码的码字.

名称	算　子						
g_1	I	I	I	X	X	X	X
g_2	I	X	X	I	I	X	X
g_3	X	I	X	I	X	I	X
g_4	I	I	I	Z	Z	Z	Z
g_5	I	Z	Z	I	I	Z	Z
g_6	Z	I	Z	I	Z	I	Z

图 10.6　Steane 7 量子比特码的稳定子生成元. 每一项代表各个量子比特上的张量积；举例来说，$ZIZIZIZ=Z\otimes I\otimes Z\otimes I\otimes Z\otimes I\otimes Z=Z_1Z_3Z_5Z_7$.

稳定子体系对描述量子码的这种应用预示着我们随后的稳定子对描述很宽一类量子码的应用. 但就现时，重要的是去体会关于作为一种量子码的 Steane 码没有什么东西是特殊的——它只不过是向量空间的一个子空间，碰巧具有稳定子的一种描述.

实际上，我们想要生成元 $g_1,\cdots,g_l$ 在移去任一生成元 g_i 都会使所生成的群变小的意义下为独立：

$$\langle g_1,\cdots,g_{i-1},g_{i+1},\cdots,g_l\rangle \neq \langle g_1,\cdots,g_l\rangle \tag{10.82}$$

根据我们现在的知识，确定生成元的一个特殊集合是否独立是相当费时的事情；幸运的是，存在一种简单的方法，这种方法称为检验矩阵，取这样名字的原因是它在稳定子码理论中扮演的角色类似于经典线性码中的奇偶检验矩阵.

设 $S=\langle g_1,\cdots,g_l\rangle$. 用检验矩阵来表示生成元 $g_1,\cdots,g_l$ 是一种非常有用的方法. 检验矩阵是一个 $l\times 2n$ 的矩阵，其行对应于生成元 g_1 到 g_l；矩阵左边的 1 表明哪些生成元包含 X，矩阵右边的 1 表明哪些生成元包含 Z；矩阵两边都出现 1 则表明生成元中有一个 Y. 更确切地，第 i 行可如下构造. 如果 g_i 在第 j 个量子比特上包含一个 I，那么第 j 列元和第 $n+j$ 列元均为 0；如果 g_i 在第 j 个量子比特上包含一个 X，那么第 j 列元为 1 而第 $n+j$ 列元为 0；如果 g_i 在第 j 个量子比特上包含一个 Z，那么第 j 列元为 0 而第 $n+j$ 列元为 1；如果 g_i 在第 j 个量子比特上包含一个 Y，那么第 j 列元和第 $n+j$ 列两者均为 1. 在 Steane 7 量子比特码情况中，我们可以从图 10.6 得出检验矩阵为

$$\left[\begin{array}{ccccccc|ccccccc}
0&0&0&1&1&1&1&0&0&0&0&0&0&0\\
0&1&1&0&0&1&1&0&0&0&0&0&0&0\\
1&0&1&0&1&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&1&1&1\\
0&0&0&0&0&0&0&0&1&1&0&0&1&1\\
0&0&0&0&0&0&0&1&0&1&0&1&0&1
\end{array}\right] \tag{10.83}$$

检验矩阵并不包含有关生成元前面的乘子的任何信息，但它确实包含很多其他有用的信息，以至多到我们要用 $r(g)$ 来指示 Pauli 群一个元 g 的 $2n$ 维行向量表示. 设我们定义一个 $2n\times 2n$ 矩阵 Λ 为

$$\Lambda=\begin{bmatrix}0 & I\\ I & 0\end{bmatrix} \tag{10.84}$$

其中，位于非对角线上的 I 矩阵是 $n\times n$ 的. 容易看出，当且仅当 $r(g)\Lambda r(g')^{\mathrm{T}}=0$，Pauli 群的元 g 和 g' 对易；公式 $x\Lambda y^{\mathrm{T}}$ 定义行矩阵 x 和 y 之间的一类“扭曲的”内积，表示对应于 x 和 y 的 Pauli 群的元是否为对易或不对易.

练习 10.33 试证明，当且仅当 $r(g)\Lambda r(g')^{\mathrm{T}}=0$，$g$ 和 g' 对易（在检验矩阵表示中，算术计算按模 2 做）.

练习 10.34 令 $S=\langle g_1,\cdots,g_l\rangle$. 试证明，当且仅当对所有 j 有 $g_j^2=I$ 和对所有 j 有 $g_j\neq -I$，$-I$ 不是 S 的一个元.

练习 10.35 令 S 为 G_n 的一个子群使得 $-I$ 不是 S 的一个元. 试证明，对所有 $g\in S$ 有 $g^2=I$，并有 $g^{\dagger}=g$.

生成元的独立性和检验矩阵之间的一个有用联系可通过下面的命题来建立.

命题 10.3 令 $S=\langle g_1,\cdots,g_l\rangle$，且使 $-I$ 不是 S 的一个元，则当且仅当相应检验矩阵的行线性独立，生成元 g_1 到 g_l 为独立.

证 我们证明逆否命题. 首先，根据练习 10.35，g_i^2 对所有 i 必等于 I. 注意到 $r(g)+r(g')=r(gg')$，所以行向量表示中的加对应于群元的乘. 因此，检验矩阵的行是线性独立的，即 $\sum_i a_i r(g_i)=0$，当且仅当 $\prod_i g_i^{a_i}$ 除去一个总乘子等于单位矩阵，对某个 j 有 $a_j\neq 0$. 但是 $-I\notin S$，所以乘子必为 1，而最后一个条件对应于条件 $g_j=g_j^{-1}=\prod_{i\neq j} g_i^{a_i}$，因而 $g_1,\cdots,g_l$ 不是独立的生成元. □

下面这个看上去无伤大雅的命题有着惊人的用处，在对当 S 由 $l=n-k$ 个独立的可对易生成元来生成且 $-I\notin S$ 时，V_S 是 2^k 维的证明中，这个命题能够立刻起到关键的作用. 贯穿本章的其余部分，我们将会反复应用这个命题. 并且，再一次选取检验矩阵表示来作为证明中的工具.

命题 10.4 令 $S=\langle g_1,\cdots,g_l\rangle$，由 l 个独立的生成元所生成，且满足 $-I\notin S$. i 在范围 $1,\cdots,l$ 内，那么，存在 $g\in G_n$ 使有 $gg_ig^{\dagger}=-g_i$，并对所有 $j\neq i$ 有 $gg_jg^{\dagger}=g_j$.

证 令 G 为相应于 $g_1,\cdots,g_l$ 的检验矩阵. 根据命题 10.3 知 G 的行是线性独立的，所以存在一个 $2n$ 维向量 x 使成立 $G\Lambda x=e_i$，其中 e_i 是第 i 位置为 1 其余均为 0 的 l 维向量. 令 g 为使成立 $r(g)=x^{\mathrm{T}}$. 那么，根据 x 的定义，对所有 $j\neq i$

有 $r(g_j)\Lambda r(g)^{\mathrm{T}}=0$ 而 $r(g_i)\Lambda r(g)^{\mathrm{T}}=1$，因此 $gg_ig^{\dagger}=-g_i$ 并对所有 $j\neq i$ 有 $gg_jg^{\dagger}=g_j$. □

通过履行我们先前的承诺，我们结束对稳定子体系的基本内容的考察，可得到结论，即在 S 由独立的可对易生成元生成且 $-I\notin S$ 的前提下，V_S 是非平凡的. 事实上，如果有 $l=n-k$ 个生成元，那么至少似乎有可能(我们将会证明)V_S 是 2^k 维的. 这是基于直觉的论据，即稳定子的每个附加生成元会消去 1/2 的 V_S 维数. 这如同我们会单纯地期望一样，因为 Pauli 矩阵张量积的 $+1$ 和 -1 特征空间把整个 Hilbert 空间分成为维数相等的两个子空间.

命题 10.5　令 $S=\langle g_1,\cdots,g_{n-k}\rangle$，由 G_n 的 $n-k$ 个独立和可对易的元所生成，且 $-I\notin S$，那么，V_S 是一个 2^k 维的向量空间.

在后面对稳定子体系的所有讨论中，我们约定，稳定子总是利用独立可对易生成元来描述并有 $-I\notin S$.

证　令 $x=(x_1,\cdots,x_{n-k})$ 为 $\boldsymbol{Z}_2$ 的 $n-k$ 个元的向量，定义

$$P_S^x\equiv\frac{\prod_{j=1}^{n-k}(I+(-1)^{x_j}g_j)}{2^{n-k}} \tag{10.85}$$

由于 $(I+g_j)/2$ 为到 g_j 的 $+1$ 特征空间上的投影算子，容易看出 $P_S^{(0,\cdots,0)}$ 必为到 V_S 上的投影算子. 根据命题 10.4，对每个 x，存在属于 G_n 的 g_x，有 $g_xP_S^{(0,\cdots,0)}(g_x)^{\dagger}=P_S^x$，则 P_S^x 的维数等同于 V_S 的维数. 进而，对两两相异的 x，容易看出 P_S^x 是正交的. 结合代数关系

$$I=\sum_x P_S^x \tag{10.86}$$

证明完成. 等式左边为到 2^n 维空间上的投影算子，而右边为维数同于 V_S 的 2^{n-k} 个正交投影算子上的求和，因此 V_S 的维数必为 2^k. □

10.5.2　酉门和稳定子体系

我们刚刚讨论过稳定子体系在描述向量空间中的应用. 稳定子体系也可用于描述在各种各样感兴趣的量子运算下，更大状态空间中的那些向量空间的动力学过程. 撇开了解量子动态运算的固有兴趣，这个目标也是非常重要的，因为我们将会采用稳定子体系来描述量子纠错，并且想要有一个优雅的方法来理解噪声和其他动态过程对那些码的影响. 设对由群 S 稳定的一个向量空间 V_S 作用一个酉运算 U，令 $|\psi\rangle$ 为 V_S 的任一元，那么，对 S 的任意一个元 g 有

$$U|\psi\rangle=Ug|\psi\rangle=UgU^{\dagger}U|\psi\rangle \tag{10.87}$$

因而状态 $U|\psi\rangle$ 可由 $UgU^{\dagger}$ 所稳定，由此我们推断，向量空间 UV_S 可由群 $USU^{\dagger}\equiv$

$\{UgU^\dagger \mid g \in S\}$所稳定. 进而，若 $g_1, \cdots, g_l$ 生成 S，则 $Ug_1U^\dagger, \cdots, Ug_lU^\dagger$ 生成 $USU^\dagger$，所以要计算稳定子中的改变，只需计算它是如何来影响稳定子的生成元的.

用这种方法处理动力学过程的一大优点在于，对某个特殊的酉运算 U，生成元的这种变换具有特别吸引人的形式. 举例来说，设我们将一个 Hadamard 门作用于单量子比特. 注意到

$$HXH^\dagger = Z, \quad HYH^\dagger = -Y, \quad HZH^\dagger = X \tag{10.88}$$

作为推论，我们可正确地推断，在 Hadamard 门作用于由 $Z(|0\rangle)$镇定的量子状态后，所得到的状态将可由 $X(|+\rangle)$所稳定.

运算	输入	输出
受控非	X_1	X_1X_2
	X_2	X_2
	Z_1	Z_1
	Z_2	Z_1Z_2
H	X	Z
	Z	X
S	X	Y
	Z	Z
X	X	X
	Z	$-Z$
Y	X	$-X$
	Z	$-Z$
Z	X	$-X$
	Z	Z

图 10.7 在各种通常运算的共轭作用下 Pauli 群的元的变换性质. 受控非门以第 1 个量子比特为控制和以第 2 个量子比特为目标.

你可能会认为，这并不令人印象深刻. 设想，我们有 n 个量子比特在一个状态中，其稳定子为$\langle Z_1, Z_2, \cdots, Z_n\rangle$. 容易看出，这就是状态$|0\rangle^{\otimes n}$. 将 Hadamard 门作用于 n 个量子比特的每一个，随后的这个状态就具有稳定子$\langle X_1, X_2, \cdots, X_n\rangle$；又容易看出，这只不过是个熟悉的状态，即所有计算基态的一个均匀叠加态. 关于这个例子值得注意之处是，最后状态的通常(状态向量)描述需要确定 2^n 个振幅，而由生成元$\langle X_1, X_2, \cdots, X_n\rangle$所进行的描述与 n 成线性. 读者或许仍然会说，在将 Hadamard 门作用于 n 个量子比特的每个后，量子计算机中并没有纠缠，所以毫不奇怪可以得到简洁的描述. 但是，在稳定子体系之内，还有很多的可能性，包括受控非门的有效描述，它连同 Hadamard 门一起可能生成纠缠. 为了解这是怎么一回事，考虑算子 X_1, X_2, Z_1 和 Z_2 在受控非门共轭作用下的行为. 用 U 表示以第 1 个量子比特作为控制和第 2 个量子比特作为目标的受控非门，有

$$UX_1U^\dagger = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$$

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

$$= X_1X_2 \tag{10.89}$$

类似的计算显示，$UX_2U^\dagger = X_2$，$UZ_1U^\dagger = Z_1$ 和 $UZ_2U^\dagger = Z_1Z_2$. 为看清楚双量子比特 Pauli 群中 U 如何共轭作用到其他算子的，我们只需对已知结果取乘积. 举例来说，为计算 $UX_1X_2U^\dagger$，我们观察到 $UX_1X_2U^\dagger = UX_1U^\dagger UX_2U^\dagger = (X_1X_2)X_2 = X_1$. Y 型 Pauli矩阵可类似地进行处理，比如 $UY_2U^\dagger = \mathrm{i}UX_2Z_2U^\dagger = \mathrm{i}UX_2U^\dagger UZ_2U^\dagger = \mathrm{i}X_1(Z_1Z_2) = Z_1Y_2$.

练习 10.36　试显式地验证，$UX_1U^\dagger = X_1X_2$，$UX_2U^\dagger = X_2$，$UZ_1U^\dagger = Z_1$，和 $UZ_2U^\dagger = Z_1Z_2$. 图 10.7 总结了 Hadamard 门、相位门和 Pauli 门的这些和其他一些有用的共轭关系.

练习 10.37　求 $UY_1U^\dagger$.

应用稳定子体系来理解酉动力学过程的一个例子是，考虑第 1.3.4 节中所引入的交换线路；这里为方便起见，将该线路说明于图 10.8 中. 考虑在线路中门的共轭作用下 Z_1 和 Z_2 变换的方式. 算子 Z_1 按顺序 $Z_1 \to Z_1 \to Z_1Z_2 \to Z_2$ 进行变换，算子 Z_2 按顺序 $Z_2 \to Z_1Z_2 \to Z_1 \to Z_1$ 进行变换；类似地，在这个线路下，有 $X_1 \to X_2$ 和 $X_2 \to X_1$. 当然，如果我们取 U 为交换算子，那么显然有 $UZ_1U^\dagger = Z_2$ 和 $UZ_2U^\dagger = Z_1$，且类似地对 X_1 和 X_2 也有相应关系，如对图 10.8 中的线路那样. 证明这意味着这个线路会实现 U，这留作为一个练习.

练习 10.38　设 U 和 V 为双量子比特上的酉算子，它们以同样的方式来对 Z_1，Z_2，X_1 和 X_2 进行共轭变换. 试证明，这意味着 $U = V$.

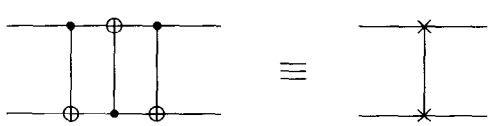


图 10.8　交换两个量子比特的线路.

这个交换线路的例子是令人感兴趣的，但该线路对稳定子体系的真正有用特性——描述某种形式量子纠缠的能力，并非十分合适. 我们已看到过，稳定子体系

可用来描述 Hadamard 门和受控非门，当然这些门的连接也用于产生纠缠状态（与《量子计算和量子信息(一)》1.3.6 节相比较）. 我们将会看到，稳定子体系事实上可用来描述很宽一类纠缠状态，包括许多量子纠错码.

在 Hadamard 门和受控非门以外，有什么样的门可以在稳定子体系内来描述呢？对这个集合最重要的补充是相位门即单量子比特门，我们现来回顾其定义为

$$S=\begin{bmatrix}1 & 0\\ 0 & \mathrm{i}\end{bmatrix} \tag{10.90}$$

相位门在 Pauli 矩阵上共轭作用可容易地计算为

$$SXS^{\dagger}=Y,\quad SZS^{\dagger}=Z \tag{10.91}$$

练习 10.39 试验证式(10.91).

事实上，共轭作用下的任何把 G_n 的元变为 G_n 的元的酉运算都可由 Hadamard 门、相位门和受控非门来组成. 根据定义，使 $UG_nU^{\dagger}=G_n$ 成立的 U 集合为 G_n 的正规化子(normalizer)并表为 $N(G_n)$，因此，我们断言 G_n 的正规化子由 Hadamard 门、相位门和受控非门来生成. 有鉴于此，Hadamard 门、相位门和受控非门有时简称为正规化子门. 这个结果的证明简单但有指导性，读者可在练习 10.40 中详细展开对它的讨论.

定理 10.6 设 U 为 n 量子比特上的任一酉算子，并具有性质：如果 $g\in G_n$，则 $UgU^{\dagger}\in G_n$；那么除去一个全局相位，U 可由 $O(n^2)$ 个 Hadamard 门、相位门和受控非门来组成.

练习 10.40 试对定理 10.6，提供一个如下归纳证明法.

(1) 证明，Hadamard 门和相位门可用于执行单量子比特上的任意一个正规化子运算.

(2) 设 U 为 $N(G_{n+1})$ 中的一个 $n+1$ 量子比特门，使得对某个 $g,g'\in G_n$ 有 $UZ_1U^{\dagger}=X_1\otimes g$ 和 $UX_1U^{\dagger}=Z_1\otimes g'$，通过 $U'|\psi\rangle\equiv\sqrt{2}\langle 0|U(|0\rangle\otimes|\psi\rangle)$ 来定义 n 量子比特上的 U'. 采用归纳假定来证明，图 10.9 中对 U 的构造可用 $O(n^2)$ 个 Hadamard 门、相位门和受控非门而实现.

(3) 证明，任一门 $U\in N(G_{n+1})$ 都可用 $O(n^2)$ 个 Hadamard 门、相位门和受控非门实现.

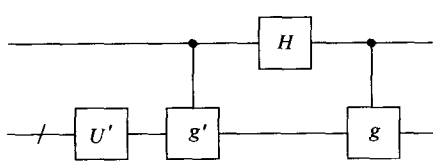


图 10.9 用以证明 Hadamard 门、相位门和受控非门来生成正规化子 $N(G_n)$ 的构造.

我们已经看到，许多令人感兴趣的量子门都在正规化子 $N(G_n)$ 中；是否有不属于正规化子的任何门呢？结论是大多数量子门都不属于. 两种有特殊意义而不属于正规化子的门是 $\pi/8$ 门和 Toffoli 门. 令 U 表示量子比特 1 和 2 为控制而量子比特 3 为目标的 Toffoli 门，并用 T 表示 $\pi/8$ 门，则我们可容易地计算 $\pi/8$ 门和 Toffoli 门在 Pauli 矩阵上的共轭作用为

$$TZT^\dagger = Z, \quad TXT^\dagger = \frac{X+Y}{\sqrt{2}} \tag{10.92}$$

和

$$UZ_1U^\dagger = Z_1, \quad UX_1U^\dagger = X_1 \otimes \frac{I+Z_2+X_3-Z_2X_3}{2} \tag{10.93}$$

$$UZ_2U^\dagger = Z_2, \quad UX_2U^\dagger = X_2 \otimes \frac{I+Z_1+X_3-Z_1X_3}{2} \tag{10.94}$$

$$UX_3U^\dagger = X_3, \quad UZ_3U^\dagger = Z_3 \otimes \frac{I+Z_1+Z_2-Z_1Z_2}{2} \tag{10.95}$$

遗憾的是，这使得比之只包含 Hadamard 门、相位门和受控非门的线路，通过稳定子体系来分析包含 $\pi/8$ 门和 Toffoli 门的量子线路要不方便得多. 幸运的是，对稳定子量子码的编码、解码、纠错和恢复都可只采用这种正规化子门来完成，所以稳定子体系对于分析这样一些码是很方便的.

练习 10.41　试通过式(10.95)来验证式(10.92).

10.5.3　稳定子体系中的测量

我们已经解释过有限一类酉门如何在稳定子体系内可被方便地描述，其实其他的也可以. 计算基中的测量在稳定子体系内也可容易地描述. 为了解具体做法，设想我们做一次 $g \in G_n$ 的测量(回忆，g 为一个 Hermite 算子，因而可以被看成是《量子计算和量子信息(一)》2.2.5 节意义下一个可观测量). 为方便起见，不失一般性假定，g 是前面没有乘子 -1 或 $\pm i$ 的 Pauli 矩阵的一个乘积，系统假定为处于具有稳定子 $\langle g_1, \cdots, g_l \rangle$ 的状态 $|\psi\rangle$. 状态的稳定子在这个测量下如何来变换？这里有两种可能性：

- g 与稳定子的所有生成元对易.
- g 与稳定子的一个或多个生成元反对易. 设稳定子具有生成元 $g_1, \cdots, g_l$，且 g 与 g_1 反对易. 不失一般性我们可以假定 g 与 $g_2, \cdots, g_l$ 对易，因为如果它与这些元中的一个(例如 g_2)不对易，那么容易验证 g 确实与 g_1g_2 对易，因而我们只需简单地在稳定子的生成元序列中用 g_1g_2 替换 g_2.

第一种情形，据下面的论证，可以导出 g 或者 $-g$ 为稳定子的一个元. 因为对每个稳定子生成元 g_j，$gg_j|\psi\rangle = g_jg|\psi\rangle = g|\psi\rangle$，$g|\psi\rangle$ 位于 V_S 中并因此是 $|\psi\rangle$ 的倍数.

由于 $g^2=I$，基此导出 $g|\psi\rangle=\pm|\psi\rangle$，由此 g 或者 $-g$ 必位于稳定子中. 我们假定 g 位于稳定子中，对 $-g$ 的讨论可类似进行. 这种情况下，$g|\psi\rangle=|\psi\rangle$ 并 g 的测量以概率 1 得到 $+1$，且这个测量不会扰动系统的状态并因此保持稳定子不变.

当 g 或 $-gg_1$ 反对易而与稳定子的所有其他生成元对易时的第二种情形，结果会如何呢？注意，g 具有特征值 ±1，所以测量结果 ±1 的投影算子分别由 $(I\pm g)/2$ 所给出，因此测量概率为

$$p(+1)=\mathrm{tr}\left(\frac{I+g}{2}|\psi\rangle\langle\psi|\right) \tag{10.96}$$

$$p(-1)=\mathrm{tr}\left(\frac{I-g}{2}|\psi\rangle\langle\psi|\right) \tag{10.97}$$

利用事实 $g_1|\psi\rangle=|\psi\rangle$ 和 $gg_1=-g_1g$，给出

$$p(+1)=\mathrm{tr}\left(\frac{I+g}{2}g_1|\psi\rangle\langle\psi|\right) \tag{10.98}$$

$$=\mathrm{tr}\left(g_1\frac{I-g}{2}|\psi\rangle\langle\psi|\right) \tag{10.99}$$

应用迹的循环性质，将 g_1 放到迹的右边末端，并利用 $g_1=g_1^\dagger$ 把它吸收到 $\langle\psi|$ 中（练习 10.35），从而给出

$$p(+1)=\mathrm{tr}\left(\frac{I-g}{2}|\psi\rangle\langle\psi|\right)=p(-1) \tag{10.100}$$

因为 $p(+1)+p(-1)=1$，我们推出 $p(+1)=p(-1)=1/2$. 设出现结果 $+1$，在这种情况下，系统的新状态为 $|\psi^+\rangle\equiv(I+g)|\psi\rangle/\sqrt{2}$，它具有稳定子 $\langle g,g_2,\cdots,g_n\rangle$. 类似地，如果出现结果 -1，以后的状态可由 $\langle -g,g_2,\cdots,g_n\rangle$ 来镇定.

10.5.4 Gottesman-Knill 定理

关于应用稳定子来描述酉动力学过程和测量的结果可用著名的 Gottesman-knill 定理来概括.

定理 10.7（Gottesman-knill 定理） 设执行一个量子计算，其只包含如下的一些元：计算基中的状态制备、Hadamard 门、相位门、受控非门、Pauli 门、Pauli 群中观测量的测量（这包括作为特殊情况的在计算基中的测量），连同可能以这样测量的结果为条件的经典控制，则这样一个计算可以在经典计算机上有效地模拟.

我们已经非显式地证明过 Gottesman-knill 定理. 经典计算机执行仿真的方法是，当各种各样的运算在计算中执行时，它只简单地保存稳定子的生成元. 举例来说，为了模拟 Hadamard 门，我们只要更新描述量子状态的 n 个生成元中每一个. 类似地，状态制备、相位门、受控非门、Pauli 门以及 Pauli 群中观测量的测量，所有这一切都可在经典计算机上用 $O(n^2)$ 步来做到，所以包含来自这个集合中 m 个运算的量子计算可以用经典计算机上的 $O(mn^2)$ 个运算来模拟.

Gottesman-knill 定理突出表明量子计算的能力是如何的难以捉摸. 它显示出,包含高度纠缠态的某些量子计算可以在经典计算机上被有效地模拟. 当然,并非所有量子计算(也即并非所有形式的纠缠)都可以在稳定子体系内被有效地描述,但是一类给人印象深刻的量子计算是能被有效地描述的. 考虑到令人感兴趣的量子信息处理任务如量子隐形传态(《量子计算和量子信息(一)》1.3.7 节)和超密度编码(《量子计算和量子信息(一)》2.3 节),都能只应用 Hadamard 门、受控非门和计算基中的测量来执行; 因而,根据 Gottesman-knill 定理,它们也可在经典计算机上有效地模拟. 进而,我们不久将会看到,广泛的各种各样的量子纠错码可以在稳定子体系内描述. 比之仅仅是量子纠缠所带来的能力,量子计算所拥有的要多得多.

练习 10.42　采用稳定子体系来验证,图 1.13 的线路会如要求那样隐形传态量子比特. 注意稳定子体系会限制正在被传送的状态类型,所以在某种意义下这并不是对隐形传态的一个完全描述,然而它确实提供了对隐形传态动力学过程的一个了解.

10.5.5　稳定子码构造

稳定子体系极其适合描述量子码. 这一节中,我们来解释这一点是如何来做到的,并用它来说明几种重要的码,包括 Shor 的 9 量子比特码、CSS 码和 5 量子比特码,后者是一个能被用于对防止单量子比特上任意差错的影响的最小码. 其基本思想非常简单: 一个$[n,k]$稳定子码被定义为由 G_n 的子群 S 可稳定的向量空间 V_S,使得 $-I\notin S$ 且 S 具有 $n-k$ 个独立的和对易的生成元,$S=\langle g_1,\cdots,g_{n-k}\rangle$,我们记该码为 $C(S)$.

码 $C(S)$的逻辑基状态是什么呢? 原理上,给定稳定子 S 的 $n-k$ 个生成元,我们可在码 $C(S)$中选取任意 2^k 个正交归一向量以作为我们的逻辑计算基态. 实际上,更有意义的是,用更为系统的方法来选取状态. 一种方法如下. 首先,我们选取算子 $\overline{Z}_1,\cdots,\overline{Z}_k\in G_n$,使得 $g_1,\cdots,g_{n-k},\overline{Z}_1,\cdots,\overline{Z}_k$ 形成一个独立的和对易的集合(稍后我们会详细解释这是如何做到的). 算子 $\overline{Z}_j$ 扮演逻辑量子比特数 j 上的逻辑 Pauli sigma z 算子的角色,所以逻辑计算基状态 $|x_1,\cdots,x_k\rangle_L$ 因此定义为具有如下稳定子的状态:

$$\langle g_1,\cdots,g_{n-k},(-1)^{x_1}\overline{Z}_1,\cdots,(-1)^{x_k}\overline{Z}_k\rangle \tag{10.101}$$

类似地,定义 $\overline{X}_j$ 为 Pauli 矩阵乘积,它在共轭作用下将 $\overline{Z}_j$ 变到 $-\overline{Z}_j$,而其他的 $\overline{Z}_i$ 和 g_i 保持不变. 很清楚,$\overline{X}_j$ 对编码后的第 j 个量子比特起到了量子非门的作用. 算子 $\overline{X}_j$ 满足 $\overline{X}_j g_k \overline{X}_j^\dagger = g_k$,因而可与稳定子的所有生成元对易. 也容易检验,$\overline{X}_j$ 与除 $\overline{Z}_j$ 以外的所有 $\overline{Z}_i$ 对易,与 $\overline{Z}_j$ 则为反对易.

稳定子码的纠错性质与其稳定子的生成元有何联系呢？设我们用具有稳定子 $S=\langle g_1,\cdots,g_{n-k}\rangle$ 的 $[n,k]$ 稳定子码 $C(S)$ 来编码一个状态，且一个差错 E 出现而污染了数据. 在三阶段分析中，我们将会来确定，什么类型的差错能够用 $C(S)$ 来检测，什么时候能够进行恢复. 首先，我们来关注在码空间上不同类型差错所具有的影响，以简单地获取有关什么类型的差错可以检测和纠正的某些直观认识；但将不会有任何证明，因为这个阶段只是简单地来建立直觉. 第二阶段是对一个一般性定理的叙述和证明，这个定理会告诉我们基于量子纠错条件，什么样类型的差错能用稳定子码来检测和纠正. 第三阶段是，应用诸如差错症状等概念，对执行差错检测和恢复来提供一个实际的“处方”.

设 $C(S)$ 为受差错 $E\in G_n$ 污染的一个稳定子码，当 E 与稳定子的一个元为反对易时，在这个码空间会发生什么呢？在这种情况下，E 把 $C(S)$ 变到一个正交子空间，且通过执行一个适当的投影测量，该差错可以在原理上检测（以及也许在检测后纠正）. 如果 $E\in S$，我们就不需要担心，因为差错 E 根本就不会污染这个空间. 真正的危险来自于当 E 可与 S 的所有元对易却不是真的位于 S 中时，也即对所有 $g\in S$ 成立 $Eg=gE$ 时. 使对所有 $g\in S$ 成立 $Eg=gE$ 的集合 $E\in G_n$ 称为 G_n 中 S 的中心化子（centralizer）并将其表为 $Z(S)$. 事实上，对于我们所关心的稳定子群 S，中心化子等同于更为熟悉的群即 S 的记作 $N(S)$ 的正规化子，正规化子定义为由使对所有 $g\in S$ 成立 $EgE^{\dagger}\in S$ 的 G_n 的所有元 S 所组成.

练习 10.43 试证明，对 G_n 的任一子群 S，有 $S\in N(S)$.

练习 10.44 试证明，对不包含 $-I$ 的 G_n 的任一子群 S，有 $N(S)=Z(S)$.

关于各种类型差错算子 E 的影响的这些观察结果，启发出了下面定理的陈述和证明. 这个定理本质上只是量子差错条件（定理 10.1）的稳定子码术语翻译.

定理 10.8（稳定子码的纠错条件） 令 S 为一个稳定子码 $C(S)$ 的稳定子，设 $\{E_j\}$ 为 G_n 中使对所有 j 和 k 成立 $E_j^{\dagger}E_k\notin N(S)-S$ 的算子的一个集合，那么，$\{E_j\}$ 为对码 $C(S)$ 的一个可纠正差错的集合.

不失一般性，我们只考虑 G_n 中 $E_j^{\dagger}=E_j$ 成立的差错 E_j，这就把对稳定子码的纠错条件简化为对所有 j 和 k 具有 $E_j^{\dagger}E_k\notin N(S)-S$.

证 令 P 为到码空间 $C(S)$ 的投影算子. 当给定 j 和 k 时有两种可能性：或者 $E_j^{\dagger}E_k$ 位于 S 中，或者 $E_j^{\dagger}E_k$ 位于 $G_n-N(S)$ 中. 考虑第一种情况. 由于 P 与 S 的元相乘 P 不变，所以有 $PE_j^{\dagger}E_kP=P$. 设 $E_j^{\dagger}E_k\in G_n-N(S)$，使得 $E_j^{\dagger}E_k$ 与 S 的某个元 g_1 必为反对易. 令 $g_1,\cdots,g_{n-k}$ 为 S 的一组生成元，使得

$$P=\frac{\prod_{l=1}^{n-k}(I+g_l)}{2^{n-k}} \tag{10.102}$$

应用反对易性，给出

$$E_j^\dagger E_k P = (I - g_1) E_j^\dagger E_k \frac{\prod_{l=2}^{n-k} (I + g_l)}{2^{n-k}} \tag{10.103}$$

但是，由于$(I+g_1)(I-g_1)=0$而有 $P(I-g_1)=0$，因此每当 $E_j^\dagger E_k \in G_n - N(S)$就有 $PE_j^\dagger E_k P=0$，这就导出，差错的集合$\{E_j\}$满足量子纠错条件，因而构成一个可纠正的差错集合. □

定理 10.8 的内容和证明是很棒的理论结果，但是，它们并没有显式地告诉我们，当实际可能时，如何来执行纠错运算. 为理解这是如何做到的，设 $g_1, \cdots, g_{n-k}$为一个$[n,k]$稳定子码的稳定子的一组生成元，$\{E_j\}$为对这个码的一个可纠正差错的集合. 差错检测可以通过依次测量稳定子的生成元 g_1 到 g_{n-k}来进行，得到的差错症状由测量的结果 β_1 到 β_{n-k}所组成. 如果差错 E_j 出现，那么差错症状由 β_l 来给出且使得 $E_j g_l E_j^\dagger = \beta_l g_l$. 如果 E_j 为具有这个差错症状的惟一差错算子，恢复可简单地应用 $E_j^\dagger$ 来实现. 如果有两个不相同的差错 E_j 和 $E_{j'}$ 导致同一个差错症状，就会有 $E_j P E_j^\dagger = E_{j'} P E_{j'}^\dagger$，其中 P 为到码空间的投影算子. 所以，只要 $E_j^\dagger E_{j'} \in S$ 成立，就有 $E_j^\dagger E_{j'} P E_{j'}^\dagger E_j = P$，因此在差错 $E_{j'}$ 发生后应用 $E_j^\dagger$ 会使得恢复成功. 因而，对每个可能的差错症状，我们只是选择一个与该差错症状相一致的单差错 E_j，在当该差错症状被观测到时，应用 $E_j^\dagger$ 以实现恢复.

定理 10.8 引发出类似于经典码距离的量子码距离的概念定义. 我们定义一个差错 $E \in G_n$ 的权重为在张量积中不等于单位阵的项数目. 举例来说，$X_1 Z_4 Y_8$ 的权重为 3. 一个稳定子码 $C(S)$的距离定义为 $N(S) - S$ 的元的最小权重. 如果 $C(S)$是具有距离 d 的一个$[n,k]$码，那么我们说 $C(S)$是一个$[n,k,d]$稳定子码. 根据定理 10.8，如同在经典码中那样，一个具有距离至少为 $2t+1$ 的码必能来纠正任意 t 量子比特上的任意差错.

练习 10.45(纠正定位差错)　设 $C(S)$为一个$[n,k,d]$稳定子码，设应用这种码将 k 个量子比特以 n 个量子比特进行编码，其随后受到噪声的污染. 但是，幸运的是，我们被告知量子比特中只有 $d-1$ 个可被噪声所影响，进而，我们还被告知，确切哪 $d-1$ 个量子比特受到了影响. 试证明，有可能来纠正这种定位差错的影响.

10.5.6　例子

我们现在给出稳定子码的一些简单例子，包括已经熟悉的码如 Shor 的 9 量子比特码和 CSS 码，但这里是从稳定子体系的新观点出发对它们进行介绍. 在每种情况中，码的性质可通过对稳定子的生成元应用定理 10.8 而容易导出. 掌握这些例子后，我们将把注意力集中在寻找执行编码和解码的量子线路上.

1.3 量子比特比特翻转码

考虑熟悉的3量子比特比特翻转码. 它由状态$|000\rangle$和$|111\rangle$所张成,并具有由Z_1Z_2和Z_2Z_3生成的稳定子. 容易验证,来自差错集$\{I, X_1, X_2, X_3\}$的两个元的每种可能的积——$I, X_1, X_2, X_3, X_1X_2, X_1X_3, X_2X_3$——它们与稳定子的生成元中的至少一个反对易(除位于S中的I以外),因而根据定理10.8,集合$\{I, X_1, X_2, X_3\}$对具有稳定子$\{Z_1Z_2, Z_2Z_3\}$的3量子比特比特翻转码构成差错的一个可纠正集.

比特翻转码的差错检测可通过测量稳定子生成元Z_1Z_2和Z_2Z_3来实现. 举例来说,如果差错X_1出现,那么稳定子变换到$\langle -Z_1Z_2, Z_2Z_3\rangle$,因此差错症状测量给出结果$-1$和$+1$. 类似地,差错$X_2$给出差错症状$-1$和$-1$,差错$X_3$给出差错症状$+1$和$-1$,而平凡差错$I$给出差错症状$+1$和$+1$. 在每种情况下,通过对由差错症状所指示的差错应用逆运算,就可以显然的方式简单地实现恢复. 比特翻转码的纠错运算总结于图10.10.

当然,我们所概述的过程与早先对3量子比特比特翻转码描述的方法完全一致. 如果这就是我们获得的所有见识,那么所有这种群论的分析很难说是值得的,只有当我们转到更为复杂的例子时,稳定子体系的真正效用才开始显现.

练习10.46 试证明,3量子比特相位翻转码的稳定子由X_1X_2和X_2X_3所生成.

Z_1Z_2	Z_2Z_3	差错类型	动作
$+1$	$+1$	无差错	无动作
$+1$	-1	比特3被翻转	翻转比特3
-1	$+1$	比特1被翻转	翻转比特1
-1	-1	比特2被翻转	翻转比特2

图10.10 以稳定子码语言描述的3量子比特相位翻转码纠错.

名称	算子								
g_1	Z	Z	I	I	I	I	I	I	I
g_2	I	Z	Z	I	I	I	I	I	I
g_3	I	I	I	Z	Z	I	I	I	I
g_4	I	I	I	I	Z	Z	I	I	I
g_5	I	I	I	I	I	I	Z	Z	I
g_6	I	I	I	I	I	I	I	Z	Z
g_7	X	X	X	X	X	X	I	I	I
g_8	X	I	I	I	X	X	X	X	X
$\bar{Z}$	X	X	X	X	X	X	X	X	X
$\bar{X}$	Z	Z	Z	Z	Z	Z	Z	Z	Z

图10.11 Shor的9量子比特码的8个生成元,以及逻辑Z和逻辑X运算(是的,它们实际上正好和人们可能天真的期望相反).

2. 9 量子比特 Shor 码

Shor 码的稳定子具有 8 个生成元，如图 10.11 所示. 对包含 I 和所有单量子比特差错的差错集，容易检验定理 10.8 的条件. 举例来说，考虑像 X_1 和 Y_4 的单量子比特差错，它们的积 X_1Y_4 与 Z_1Z_2 为反对易，因而不位于 $N(S)$内. 类似地，所有来自这个差错集的其他两个差错的积或是位于 S 内，或是与 S 的至少一个元为反对易并因而不位于 $N(S)$内，这同时意味着 Shor 码可被用于来纠正任意的单量子比特差错.

练习 10.47　试验证，图 10.11 的那些生成元可生成式(10.13)的两个码字.

练习 10.48　试证明，运算 $\bar{Z}=X_1X_2X_3X_4X_5X_6X_7X_8X_9$ 和 $\bar{X}=Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9$ 在 Shor 码编码后的量子比特上起到逻辑 Z 和逻辑 X 运算的作用. 也即，要来证明，这个 $\bar{Z}$ 独立于 Shor 码的生成元并与其对易；$\bar{X}$ 独立于 Shor 码的生成元并与其对易，而与 $\bar{Z}$ 为反对易.

3. 5 量子比特码

对单量子比特编码的一个量子码最小尺寸是多少？它使编码后状态中单量子比特上的任意差错都能被检测和纠正. 事实上，问题的答案为 5 个量子比特(参见练习 12.4.3). 5 量子比特码的稳定子具有图 10.12 给出的那些生成元. 因为 5 量子比特码是有能力防止单个差错的最小码，或许会认为它是最为有用的码；但是，对于许多应用，采用 Steane 的 7 量子比特码要更为直接.

名称	算　子				
g_1	X	Z	Z	X	I
g_2	I	X	Z	Z	X
g_3	X	I	X	Z	Z
g_4	Z	X	I	X	Z
$\bar{Z}$	Z	Z	Z	Z	Z
$\bar{X}$	X	X	X	X	X

图 10.12　5 量子比特码的四个生成元，以及逻辑 Z 运算和逻辑 X 运算. 注意，后三个生成元可以通过第一个逐次右移得到.

练习 10.49　应用定理 10.8 来验证，5 量子比特码能防止任意单量子比特差错. 5 量子比特码的逻辑码字为

$$\begin{aligned}|0_L\rangle=\frac{1}{4}[&|00000\rangle+|10010\rangle+|01001\rangle+|10100\rangle+\\&|01010\rangle-|11011\rangle-|00110\rangle-|11000\rangle-\\&|11101\rangle-|00011\rangle-|11110\rangle-|01111\rangle-\\&|10001\rangle-|01100\rangle-|10111\rangle+|00101\rangle]\end{aligned}\tag{10.104}$$

$$
\begin{aligned}
|1_L\rangle = \frac{1}{4}[&|11111\rangle + |01101\rangle + |10110\rangle + |01011\rangle + \\
&|10101\rangle - |00100\rangle - |11001\rangle - |00111\rangle - \\
&|00010\rangle - |11100\rangle - |00001\rangle - |10000\rangle - \\
&|01110\rangle - |10011\rangle - |01000\rangle + |11010\rangle] \qquad (10.105)
\end{aligned}
$$

练习 10.50 试证明，5 量子比特码会使量子 Hamming 界饱和，也即它使不等式(10.51)取等号.

4. CSS 码和 7 量子比特码

CSS 码是一类稳定子码中的突出例子，它能清晰地说明用稳定子体系来理解量子码的结构是多么的容易. 设 C_1 和 C_2 分别为$[n,k_1]$和$[n,k_2]$经典线性码，使得 $C_2 \subset C_1$ 且 C_1 和 $C_2^{\perp}$ 两者都可纠正 t 个差错. 定义具有如下形式的一个检验矩阵：

$$
\left[\begin{array}{c|c} H(C_2^{\perp}) & 0 \\ 0 & H(C_1) \end{array}\right] \qquad (10.106)
$$

为看出这定义了稳定子码，需要这个检验矩阵满足对易条件 $H(C_2^{\perp})H(C_1)^{\mathrm{T}}=0$，但是，由于假定 $C_2 \subset C_1$，我们有 $H(C_2^{\perp})H(C_1)^{\mathrm{T}}=[H(C_1)G(C_2)]^{\mathrm{T}}=0$，确实，这个码正是 CCS$(C_1,C_2)$且具有纠正 t 量子比特上任意差错的能力，这只是一个容易的练习.

7 量子比特 Steane 码是 CSS 码的一个例子，其检验矩阵我们已经在式(10.83)中见到过. 对 Steane 码，编码后的 Z 和 X 算子可分别定义为

$$
\begin{aligned}
\overline{Z} &\equiv Z_1 Z_2 Z_3 Z_4 Z_5 Z_6 Z_7, \\
\overline{X} &\equiv X_1 X_2 X_3 X_4 X_5 X_6 X_7 \qquad (10.107)
\end{aligned}
$$

练习 10.51 试验证，式(10.106)中定义的这个检验矩阵对应于 CCS 码 CCS(C_1,C_2)的稳定子；再用定理 10.8 来证明，最多 t 个量子比特上的任意差错可以用这个码来纠正.

练习 10.52 试通过在码字上直接运算来验证，式(10.107)的运算起到如逻辑 Z 和 X 的作用.

10.5.7 稳定子码的标准形

对稳定子码，如果我们将码化为标准形，逻辑 Z 和 X 算子就会变得容易理解得多. 为了解标准形是什么，考虑$[n,k]$稳定子码 C 的检验矩阵：

$$
G = [G_1 \mid G_2] \qquad (10.108)
$$

这个矩阵具有 $n-k$ 个行. 这个矩阵行的对换对应于重新标记生成元，这个矩阵两边相应列的对换对应于重新标记量子比特，将两行相加对应于乘以生成元；容易

看出，当 $i\neq j$ 时，我们总是可以用 g_ig_j 来替换 g_i. 因此，存在具有不同生成元集合的一个等价码，其相应的检验矩阵对应于矩阵 G，其中已对 G_1 应用 Gauss 消去法，且有必要时对换量子比特：

$$\begin{array}{c} \\ r\{ \\ n-k-r\{ \end{array} \begin{array}{c} \overbrace{}^{r}\ \overbrace{}^{n-r} \\ \left[\begin{array}{cc|cc} I & A & B & C \\ 0 & 0 & D & E \end{array}\right] \end{array} \tag{10.109}$$

其中 r 是 G_1 的秩. 下一步，当有必要时对换量子比特，我们对 E 执行 Gauss 消去法以得到

$$\begin{array}{c} \\ r\{ \\ n-k-r-s\{ \\ s\{ \end{array} \begin{array}{c} \overbrace{}^{r}\ \overbrace{}^{n-k-r-s}\ \overbrace{}^{k+s} \\ \left[\begin{array}{ccc|ccc} I & A_1 & A_2 & B & C_1 & C_2 \\ 0 & 0 & 0 & D_1 & I & E_2 \\ 0 & 0 & 0 & D_2 & 0 & 0 \end{array}\right] \end{array} \tag{10.110}$$

最后 s 个生成元不能与最前 r 个生成元相对易，除非 $D_2=0$，因此，我们可以假定 $s=0$. 进而，通过对行取适当的线性组合，我们也可使 $C_1=0$. 所以，我们的检验矩阵具有形式：

$$\begin{array}{c} \\ r\{ \\ n-k-r\{ \end{array} \begin{array}{c} \overbrace{}^{r}\ \overbrace{}^{n-k-r}\ \overbrace{}^{k} \\ \left[\begin{array}{ccc|ccc} I & A_1 & A_2 & B & 0 & C \\ 0 & 0 & 0 & D & I & E \end{array}\right] \end{array} \tag{10.111}$$

其中，我们已经重新标记 E_2 为 E，D_1 为 D. 不难看出，这种方法不是惟一的；但是，我们说，具有形为式(10.111)的检验矩阵为处于标准形.

给定量子码的标准形后，可容易为这个码定义编码后的 Z 算子. 也即，我们必须选择 k 个算子，它们彼此相独立并独立于稳定子的生成元，它们彼此对易并与稳定子的生成元对易. 设对这些 k 个编码后的 Z 算子，我们写出检验矩阵为 $G_z=[F_1F_2F_3|F_4F_5F_6]$，其中所有矩阵均具有 k 个行，而各自的列维数分别为 r，$n-k-r$，k，r，$n-k-r$ 和 k，我们选取这些矩阵使得 $G_z=[000|A_2^{\mathrm{T}}0I]$. 这些编码后的 Z 算子与稳定子元的对易性是由方程 $I\times(A_2^{\mathrm{T}})^{\mathrm{T}}+A_2=0$ 来导出的. 很清楚，编码后 Z 算子相互对易，因为它们只包含 Z 算子的积. 编码后 Z 的算子与稳定子的前 r 个生成元的独立性出自这样一个事实，即没有任何 X 项出现在编码后 Z 算子的定义中. 编码后 Z 算子与 $n-k-r$ 生成元集合的独立性则出于这样的事实，即出现在那些生成元的检验矩阵中的为 $(n-k-r)\times(n-k-r)$ 单位矩阵，而编码后的 Z 算子检验矩阵中没有相应的项. 采用类似的方法，我们可来选择具有 $k\times 2n$ 检验

矩阵$[0E^{\mathrm{T}}I|C^{\mathrm{T}}00]$的编码后的 X 算子.

练习 10.53 试证明,编码后的 Z 算子相互独立.

练习 10.54 试证明,若具有如上定义的编码后的 X 算子的检验矩阵,则编码后的 X 算子具有如下性质:相互独立且与所有生成元相独立,与稳定子的所有生成元对易以及相互对易;并且,$\overline{X}_j$ 为与除 $\overline{Z}_j$ 而外的所有 $\overline{Z}_k$ 对易,而与 $\overline{Z}_j$ 反对易.

作为一个例子,我们把 Steane 码的检验矩阵(式(10.83))变换为标准形.对这个码,我们有 $n=7$ 和 $k=1$;而对这个检验矩阵的分析显示,σ_z 部分的秩为 $r=3$.由此,通过对换量子比特 1 和 4,3 和 4,以及 6 和 7,然后通过将行 6 加到行 4,再将行 6 加到行 5,最后将行 5 和行 4 加到行 6,矩阵就可被变换到标准形.得到的标准形为

$$\left[\begin{array}{ccccccc|ccccccc}
1&0&0&0&1&1&1&0&0&0&0&0&0&0\\
0&1&0&1&0&1&1&0&0&0&0&0&0&0\\
0&0&1&1&1&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&1&1&0&0&1\\
0&0&0&0&0&0&0&0&1&1&0&1&0&1\\
0&0&0&0&0&0&0&1&1&1&0&0&1&0
\end{array}\right] \tag{10.112}$$

我们可读出 $A_2=(1,1,0)$,因此编码后 Z 具有检验矩阵$[0000000|11\,00001]$,它对应于 $\overline{Z}=Z_1Z_2Z_7$.回顾,量子比特 1 和 4,3 和 4,以及 6 和 7 已对换,这对应于原来码中的 $\overline{Z}=Z_2Z_4Z_6$ 的编码后 Z.根据式(10.107),这看起来会有点使人感到迷惑,方程指出编码后 Z 为 $\overline{Z}=Z_1Z_2Z_3Z_4Z_5Z_6Z_7$;但是,这种迷惑只要注意到如下事实就可被解决:两个"不同的"编码后 Z 算子的区别仅在于因子 $Z_1Z_3Z_5Z_7$ 即 Steane 码的稳定子的一个元,因而两者对 Steane 码具有相同的作用.

练习 10.55 试求 Steane 码标准形的 $\overline{X}$ 算子.

练习 10.56 试证明,用 g 乘以算子来替换一个编码后 X 或 Z 算子,其中 g 为稳定子的一个元,它不会改变这个算子对这个码的作用.

练习 10.57 试对处于标准形的 5 量子比特码和 9 量子比特码来给出它们的检验矩阵.

10.5.8 编码、解码和纠错的量子线路

稳定子码的特性之一在于,它们的结构使能系统地来构造编码、解码和纠错的步骤.我们首先描述一般的方法,然后作为例子介绍一些显式的线路构成.让我们

从一般情况，即具有生成元 $g_1,\cdots,g_{n-k}$ 和逻辑 Z 算子 $\bar{Z}_1,\cdots,\bar{Z}_k$ 的 $[n,k]$ 稳定子码开始.

制备编码后的 $|0\rangle^{\otimes k}$ 状态是非常简单的，这个状态对于启动量子计算是标准状态. 为此，我们可从任何易于制备的状态——例如 $|0\rangle^{\otimes n}$——开始，并依次测量观测量 $g_1,\cdots,g_{n-k},\bar{Z}_1,\cdots,\bar{Z}_k$ 中的每一个. 取决于测量的结果，所得到的量子状态将具有稳定子 $\langle \pm g_1,\cdots,\pm g_{n-k},\pm\bar{Z}_1,\cdots,\pm\bar{Z}_k\rangle$，其中不同的符号（+或−）由各自的测量结果所确定. 所有稳定子生成元和 $\bar{Z}_j$ 的符号可在随后应用 Pauli 算子积来确定，如同命题 10.4 的证明中所描述的，这就产生具有稳定子 $\langle g_1,\cdots,g_{n-k},\bar{Z}_1,\cdots,\bar{Z}_k\rangle$ 的状态，也即到达编码后的 $|0\rangle^{\otimes k}$. 一旦这个状态制备好，就有可能应用取自集合 $X_1,\cdots,X_k$ 的一些适当的算子，来将状态改变到任意编码后的计算基态 $|x_1,\cdots,x_k\rangle$. 当然，这种编码的方法也有缺点，即它不是酉的. 为得到完全酉的编码，可来采用基于检验矩阵标准形的一种另外的方法，这种方法将概述于问题 10.3. 同样，如果你想要编码一个未知状态，从编码后的 $|0\rangle^{\otimes k}$ 状态开始，这也可以系统地来做到，如同问题 10.4 中所解释的. 至于对于我们的目的，制备编码后的 $|0\rangle^{\otimes k}$ 状态已是足够了.

解码量子码也是十分简单的，但是值得来解释，为什么对于许多目的完全的解码没有必要. 事实上，容错量子计算的方法可用来直接对编码后的数据执行逻辑运算，而无需对数据进行解码，而且，仅仅通过测量逻辑 Z 算子，而无需解码和在计算基中测量，以这种方法执行的一个计算输出就可直接确定. 因此，对于我们的目的而言，进行保持被编码的量子信息的完全酉解码并不是太重要. 如果出于某种原因期望这样一个解码的过程——或许有人正在带噪声的信道上用量子纠错码来传送信息——那么，通过反向运行问题 10.3 中的酉编码线路就可来实现这一点.

稳定子码的纠错方法已在 10.5.5 节中描述过. 这种方法很像编码过程：只是依次测量生成元 $g_1,\cdots,g_{n-k}$ 中的每一个，得到差错症状 $\beta_1,\cdots,\beta_{n-k}$，然后，采用经典计算来由 β_j 确定所要求恢复的运算 $E_j^{\dagger}$.

在上面的每个描述中，构造编码、解码和纠错线路的关键在于，了解如何来测量算子. 回顾，这就是我们所已广泛应用的标准投影测量的推广，它的目的是把状态投影到算子的本征态，并得到投影后状态和本征值的一个指示. 如果这使你回想起第 5 章的相位估计算法，并非巧合. 回顾第 5 章及练习 4.34，在给定执行受控 M 运算的一个门后，示于图 10.13 的线路可被用来测量单量子比特算子 M（具有特征值 ± 1）. 这种线路的两个有用的版本，可用于测量 X 和 Z，如图 10.14 和图 10.15所示.

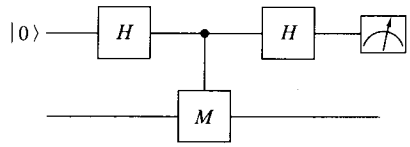


图 10.13 测量具有特征值±1 的单量子比特算子 M 的量子线路，顶部量子比特为用于辅助测量，底部量子比特则被测量.

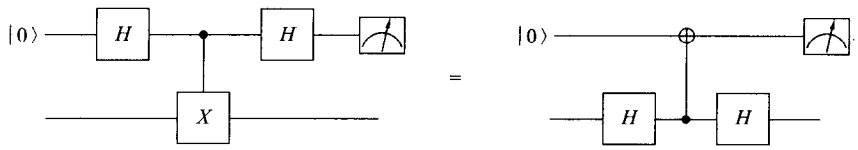


图 10.14 测量 X 算子的量子线路. 这里给出了两种等价线路：左边线路为通常的结构(如图 10.13 中那样)，右边线路为有用的等价线路.

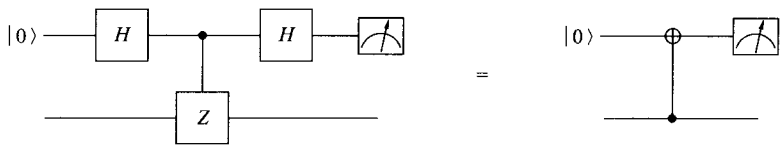


图 10.15 测量 Z 算子的量子线路. 这里给出了两种等价线路：左边线路为通常的结构(如图 10.13 中那样)，右边线路为有用的简化.

当然，M 为单量子比特算子的事实，没有什么特殊之处：如果我们用一束量子比特来替换第二量子比特，而 M 为具有特征值±1 的任意 Hermite 算子，那么图 10.13中的线路会照样工作. 这样一些算子包括例如在稳定子码的编码、解码和纠错过程期间我们需要测量的 Pauli 算子的那些积.

作为一个具体例子，考虑差错症状测量和 7 量子比特 Steane 码的编码方法. 从这个码的检验矩阵的标准形即式(10.112)开始讨论比较方便，因为从这个矩阵我们能够立刻得到需要测量的生成元. 特别地，回顾，左边方块对应于 X 生成元，右边方块对应于 Z 生成元，所以就立即导出示于图 10.16 的量子线路. 注意，矩阵中 0 和 1 的位置如何对应于左半部分(测量 X)中这些门的目的位置和右半部分(测量 Z)中这些门的目的位置. 通过根据 Pauli 算子积在码量子比特上测量结果来纠正差错，这个线路就能被用于执行纠错. 或者，如同先前描述过的，通过增加 $\overline{Z}$ 的一次附加测量和固定稳定子的生成元中的正负号，这个线路就可被用于制备编码后逻辑状态 $|0_L\rangle$.

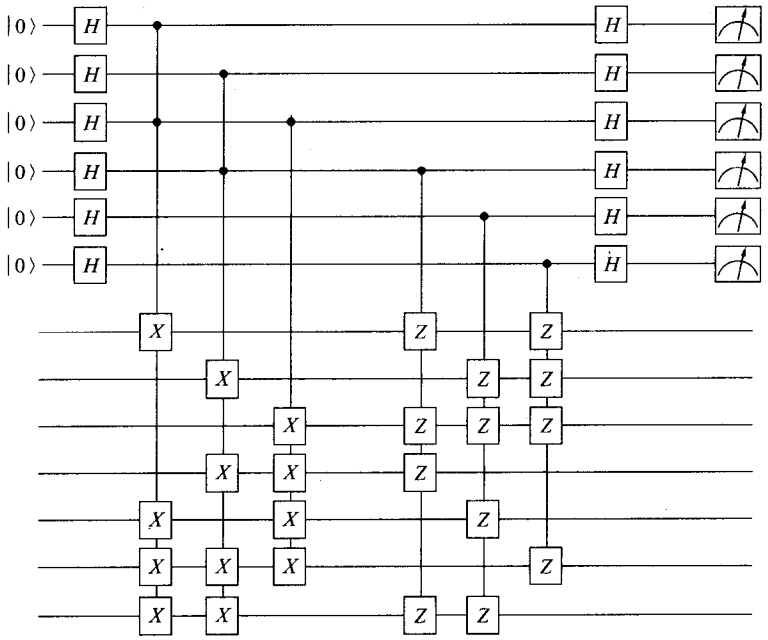


图 10.16　用于给出差错症状的测量 Steane 码的生成元的量子线路，顶部 6 个量子比特为用于辅助测量，底部 7 个量子比特为码量子比特.

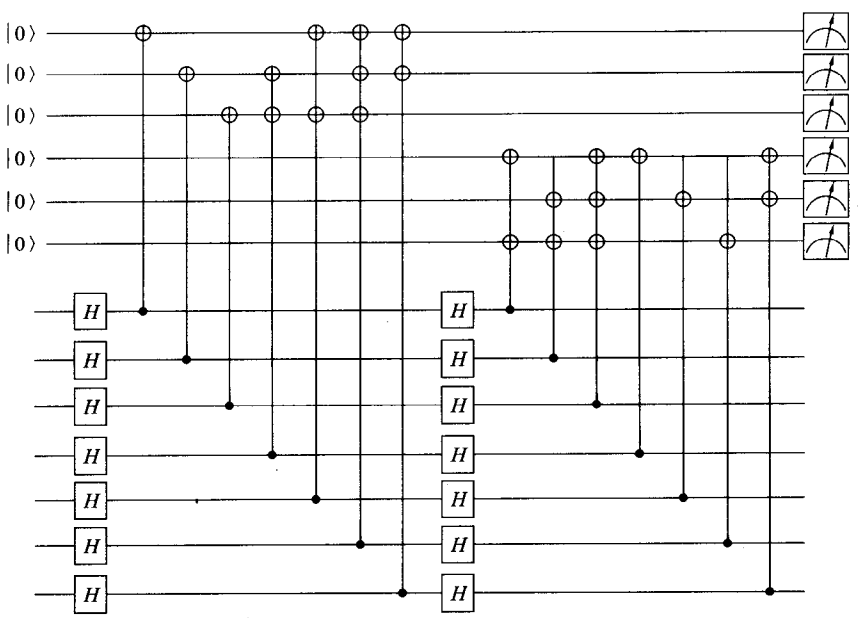


图 10.17　等价于图 10.16 中线路的量子线路 .

练习 10.58 试验证图 10.13～图 10.15 中的那些线路会像所描述的那样工作，并来检验所宣称的线路等价性.

练习 10.59 试证明，利用图 10.14 和 10.15 的等同性，图 10.16 的差错症状线路可用图 10.17 的线路来替换.

练习 10.60 试构造类似于图 10.16 中线路，但针对 9 量子比特码和 5 量子比特码的差错症状测量线路.

练习 10.61 试显式地描述对应于不同可能的差错症状的恢复算子 $E_j^{\dagger}$，而差错症状可采用图 10.16 中的线路而被测量.

10.6 容错量子计算

量子纠错的最强有力的应用之一并非仅在于检测存储的或传输的量子信息，而且还在于保护动态地进行计算时的量子信息. 引人注目的是，事实上，甚至使用会出错的逻辑门，仍能达成任意精度的量子计算，只要每个门的差错概率低于某个确定常阈值. 下面一些节中，我们会解释用于达到这种引人注目结果的容错量子计算的方法. 在 10.6.1 节中，我们首先来浏览一下大的图像，对容错量子计算各个要素的详细考察则见 10.6.2 节和 10.6.3 节，10.6.4 节为结束，讨论了对容错结构的某些局限性和可能推广. 注意，对容错量子计算的许多细节的严格讨论已稍微超出我们的范围，有兴趣的读者可参看章末的“历史和进一步阅读的材料”.

10.6.1 容错：大图像

容错量子计算理论是在获得阈值条件过程中综合各种不同想法而成的. 我们现在来依次描述这些思想. 我们以编码后数据上的计算概念作为开始，并要解释差错传播和差错积累这些基本问题如何要求编码后数据上计算的线路满足某些容错准则. 我们随后引入量子线路的基本噪声模型，这个模型允许我们对容错运算概念提供更为精确的定义. 我们将通过容错运算的一个特别的实例——容错受控非门——来开展讨论，并解释它如何用于防止差错的传播和积累. 最后通过解释容错运算如何与称为串联(concatenation)的方法相结合以得到量子计算的阈值定理，并对阈值给出一个简单估计.

1. 基本问题

容错量子计算的基本思想是，对编码后的量子状态以不要求解码的方式来进行直接计算. 假定我们有一个如图 10.18 所示的简单量子线路. 遗憾的是，噪声会影响用于构成这个线路的每个部分——状态制备过程、量子逻辑门、输出的测量以及甚至量子信息沿着量子连线的简单传输等. 为扼制噪声的影响，我们应用纠错码

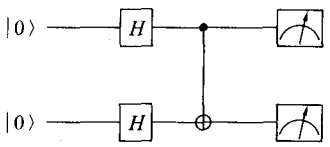


图 10.18　一个简单的量子线路. 如果线路中的每个元件以概率 p 失效，那么输出出现一个差错的概率为 $O(p)$.

如 7 量子比特 Steane 码，以一个编码后量子比特块来代替原线路中的每个量子比特，并以作用于编码后状态上的编码后门(encoded gate)的一个过程来代替原线路中的每个门，如图 10.19 所示. 通过周期性地在编码后状态上执行纠错，我们就可防止这个状态中的差错积累. 当然，仅只周期性地执行纠错对于防止差错的积累还是不够的，甚至纠错作用于每个编码后门以后. 其原因是双重的. 首先，也是最为重要的，编码后门可以引起差错的传播. 举例来说，如图 10.20所示的编码后受控非门，会把编码后控制量子比特上的差错传播到编码后目标量子比特上，因此，形成编码后控制量子比特的量子比特中的差错，会传播成为编码后目标量子比特中的差错. 为使纠错能有效消去差错，编码后门应该非常仔细地设计，从而在执行编码后门的过程期间，任何地方的一个差错都只能传播到每个编码后数据块中很少数量的量子比特上. 执行编码后门的这种过程称为容错过程. 我们将会证明，应用容错过程，可以来执行一组通用逻辑运算——Hadamaed 门、相位门、受控非门以及 $\pi/8$ 门等. 第二个必须解决的问题是，纠错本身也会在编码后量子比特上引入差错，所以必须仔细设计纠错过程，使之不会引入太多差错到编码后的数据中. 这一点可采用以编码后门来防止差错传播的类似技术而来实现，来确保纠错过程期间的差错不会传播到编码后数据中引起太多差错.

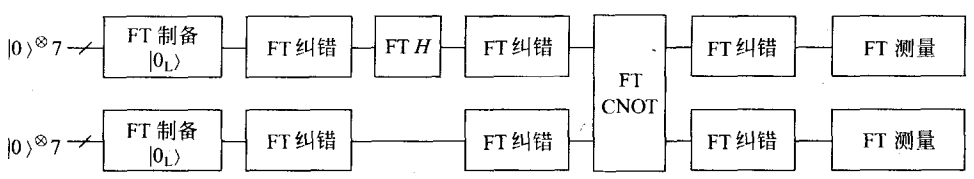


图 10.19　采用编码后量子比特和编码后运算，对图 10.18 中的线路进行的仿真.



图 10.20　受控非门会引起一个差错传播，从而不仅影响一个量子比特，而且会影响两个量子比特. 这在应用编码后量子比特，且一个编码后受控非门被实现时也同样成立，如同文中所讨论的.

如果应用容错过程来执行所有这些运算，那么输出上差错的概率为 $O(p^2)$，其中 p 为线路中任一单个元件失效的概率. 一个有意思的特性是，第二个纠错步骤

执行于第二个量子比特上. 通过这个步骤被纠错的“运算”看来是平凡的：量子比特什么也没有发生. 然而，仅仅把量子比特存储一个周期时间，就会引入差错到量子比特中；因此，应当周期地进行纠错，以防止差错的积累.

2. 容错运算：定义

让我们更为准确地阐述实现一个容错编码后量子门的一个特定过程是什么意思. 我们定义一个过程的容错性具有这样一个性质，如果过程中只有一个元件失效，那么失效在过程的每个编码后量子比特输出块中最多只引起一个差错. 举例来说，量子纠错的容错恢复过程中单个元件的失效，导致恢复过程会被正确地执行，除了在输出的一个单量子比特上有一个差错出现. 所谓“元件”，我们意指是用于编码后门中的任何基本运算，可以包括带噪声的门、带噪声的测量、带噪声的量子连线以及带噪声的状态制备等. 量子门的容错过程的这个定义，在文献中有时还推广来处理出现在容错计算理论中的某些更为微妙的问题，但对我们的详细程度而言这就足够了.

当然，在量子计算中，我们并不希望只执行编码后量子门，定义容错测量过程和容错状态制备的概念同样是有用的. 如果过程中任一单个元件上的失效只在过程输出的每个编码后量子比特块中的最多一个量子比特上导致一个差错，则测量一组编码后量子比特上的观测量的过程认为是容错的. 进而，我们要求，仅当一个元件失效时，则得到的测量结果必有 $O(p^2)$ 的差错概率，其中 p 为用于实现这个测量过程的任一元件中的(最大)差错症状概率. 如果给定单个元件在过程期间失效，来自过程的每个编码后量子比特输出块中最多只有一个量子比特处于差错下，则一个制备固定的编码后状态的过程认为是容错的.

为使这些容错的概念更为准确，我们需要更明确差错模型. 主要简化之一是将量子比特上的差错描述为如下四种类型之一：以一定概率随机出现的 I, X, Y 或 Z. 当执行如受控非门等的门时，我们就允许相关差错以某个概率出现在两个量子比特上，但再次假定它们具有 Pauli 矩阵的张量积形式. 这种概率性的分析，能使我们应用熟悉的经典概率论的概念，来确定线路的输出为正确或不正确的总概率. 在容错更为复杂的表述中(参见“历史和进一步阅读材料”)，可以考虑更为一般得多的差错模型，例如允许任意形式的相关差错出现在若干量子比特中. 不过，结合本章早些时候我们所获得的知识，即纠正一组离散的差错对于执行可能的连续差错的纠错是足够的，而用于那些比较复杂分析的技术基本上就是我们所描述的分析的推广.

利用准备就绪的噪声模型，就可更为精确地来说明，当我们说一个差错通过线路“传播”指的是什么意思. 举例来说，考虑图 10.20 中的受控非门. 设想，恰在受控非门作用前瞬时，一个 X 差错出现在第一量子比特上. 如果受控非门的酉算子记

为 U，那么这个线路的有效作用为 $UX_1=UX_1U^{\dagger}U=X_1X_2U$，也即好像受控非门正确地作用了，但一个 X 差错出现在受控非门后的两个量子比特上. 贯穿本章的其余部分，我们会重复应用通过门对差错取共轭这个技巧，以研究差错是如何通过线路来传播的. 差错传播的一个稍微具挑战性的例子是假定失效的为受控非门自身，那么，将会发生什么呢？设带噪声的受控非门来实现量子运算 ε，则这可以重写为 $\varepsilon=\varepsilon\circ\mathcal{U}^{-1}\circ\mathcal{U}$，其中 $\mathcal{U}$ 是实现完美的受控非门的量子运算. 因此，带噪声的受控非门等价于，紧随着一个完美的受控非门来应用运算 $\varepsilon\circ\mathcal{U}^{-1}$. 如果这个带噪声的受控非门相当的好，运算 $\varepsilon\circ\mathcal{U}^{-1}$ 近似地为单位阵，并可理解为通常的张量积形式的差错模型（如以小概率 p 出现在两个量子比特上的 $X\otimes Z$）.

贯穿在下面几节中，我们会来详细解释执行我们所描述过的每类容错运算——具有通用门集合的容错量子逻辑、容错测量以及容错状态制备等——的过程. 我们所描述的实际结构是针对 Steane 码的，但是它们可以相当容易地推广到更为一般的稳定子码的情况. 就现时而言，我们不妨想象在我们的处理上，我们有所有这些过程，我们能否将它们组合在一起来执行量子计算呢？

3. 例子：容错受控非门

首先考察如图 10.21 所示的，紧随着容错纠错步骤的一个实现容错受控非门的过程. 对这个线路的分析可用 4 个步骤来进行. 步骤 1 是线路的进入点，步骤 2 是在编码后受控非门执行之后，步骤 3 是在差错症状测量之后，而步骤 4 则是在恢复运算作用之后. 我们的目标是来证明，这个线路引入两个或多个差错到第一个编码后块中的概率为 $O(p^2)$，其中 p 为线路中单个元件失效的概率. 因为仅当第一量子比特块中存在两个或多个差错时，这个块的（假设的）完美解码才会失效，这就导出完美解码后状态在线路作用以后，包含差错的概率最多比线路作用之前大 $O(p^2)$.

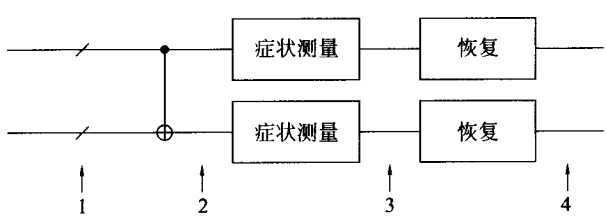


图 10.21　包含纠错的容错过程结构的框图

为了证明这个过程，以概率 $O(p^2)$ 引入两个差错到第一个编码后块中，然后让我们来找出这个线路能引入两个或更多差错到输出量子比特第一个编码块中的所有可能途径.

（1）在量子比特的每个编码后块中，存在一个事先存在的单差错进入到步骤 1

上的线路.这会在第一块的输出中引起两个差错,这是因为,举例来说,第二块上的差错可能通过编码后受控非门线路在量子比特的第一块上引起一个差错.规定到此阶段所有运算都是容错地进行的,我们就能认为这样一个差错进入到第一个块的概率最多为 $c_0 p$,其中 c_0 为某个常数.因为,在这个线路先前阶段中的差错症状测量或恢复阶段期间,这样一个差错必就已经出现.c_0 就是在这个线路先前阶段中的差错症状测量或恢复阶段期间可能出现失效的位置总数.如果简单起见,我们假定,单个事先存在的差错进入到第二个块上步骤 1 的概率也为 $c_0 p$,且这两个差错为独立地出现,那么这个事件的概率最多为 $c_0^2 p^2$.对于下面所描述的 Steane 码结构,共有来自于 6 个分立的差错症状对 c_0 的贡献,每个差错症状近似地各有 10^1 个可出现失效的位置,以及一个恢复运算包含 7 个元件,则总体近似地为 $c_0 \approx 70$.

(2) 一个事先存在的单差错在步骤 1 中进入到量子比特的第一块或第二块上,且单个失效出现在容错受控非门期间,这种情形的概率为 $c_1 p^2$,其中 c_1 是常数并定义为可能出现失效的点对的数目.对于下面所描述的 Steane 码结构,我们先前宣称过,粗略地共有 70 个位置乘以 2 个块即总起来 140 个位置,那里已经出现的失效会引起一个差错进入线路.线路中还有另外 7 个位置可能出现失效,则总共 $c_1 \approx 7 \times 140 \approx 10^3$ 个可出现两个失效的位置.

(3) 两个失效在容错受控非门期间出现.这种情形发生的概率最多 $c_2 p^2$,其中 c_2 为可能出现失效点对的数目.对 Steane 码,$c_2 \approx 10^2$.

(4) 一个差错症状在受控非门期间和差错症状测量期间出现.两个或多个差错可以出现在输出上的惟一途径是,如果差错症状测量以概率 $c_3 p^2$ 提供不正确结果,c_3 为某个常数(对 Steane 码,$c_3 \approx 10^2$).另一种看起来有兴趣但并不紧要的情形是,当差错症状测量提供正确的结果,此情形中通过受控非门引入的差错正确地诊断并应用恢复而纠正,剩下的只有差错症状测量期间引入的一个输出上的单差错.

(5) 两个或多个差错症状在差错症状测量期间出现.这种情形发生概率最多为 $c_4 p^2$,其中 c_4 为可能出现失效点对的数目.对 Steane 码,$c_4 \approx 70^2 \approx 5 \times 10^2$.

(6) 一个差错症状在差错症状测量期间出现,一个差错症状在恢复期间出现.这种情形发生概率最多为 $c_5 p^2$,其中 c_5 为可能出现失效点对的数目.对 Steane 码,$c_5 \approx 70 \times 7 \approx 500$.

(7) 两个或多个差错症状在恢复期间出现.这种情形发生概率最多为 $c_6 p^2$,其中 c_6 为可能出现失效点对的数目.对 Steane 码,$c_6 \approx 7^2 \approx 50$.

因此,这个线路引入两个或多个差错到量子比特的编码后第一个块的概率最多为 cp^2,且常数 $c = c_0^2 + c_1 + c_2 + c_3 + c_4 + c_5 + c_6$,对 Steane 码这个常数近似等于 10^4.如果一个完美的解码在线路的末端执行,那么一个差错的概率最多为 cp^2.这是一个真正引人注目的结果:我们已经设法对一类受控非门找到了一个实现,这类受控非门具有性质即单个元件以概率 p 可失效而编码后过程以概率 $1-cp^2$ 成

功，从而如果 p 足够小，例如，$p<10^{-4}$，那么通过编码过程就得到一个纯增益. 对量子计算中所有其他运算也可导出类似的结论，从而通过容错地进行我们的运算中的任何一种，我们都把失效的概率从 p 减少到 cp^2，其中 c 为某个常数. 尽管只对受控非门的 c 作过估计，但是对其他容错运算的估计并没有太大的区别，因而我们在数值估计中将继续采用 $c\approx 10^4$.

4. 串联码和阈值定理

存在一个基于串联码(concatenated code)的巧妙结构，串联码可以用来进一步减少实际的计算差错率. 其思想是迭代地应用上述以编码后线路来模拟一个线路的方案，编码后线路构成分层的量子线路 C_0(我们所希望模拟的原线路)，C_1，C_2，…. 在这个构造的第一阶段，原线路中的每个量子比特用一个量子码来编码，此量子码的量子比特自身又用量子码来编码，此量子码自己的量子比特自身又是被编码的，如此下去，如图 10.22 中所示. 在这个构造的第二阶段，原线路 C_0 中的任一给定的门，如 Hadamard 门，在线路 C_1 中代之以一个实现编码后 Hadamard 门和纠错的容错过程. 而线路 C_1 中使用的每个组成元件，在线路 C_2 中则代之以一个实现元件的编码后版本和纠错的容错过程. 如此无限地做下去. 设如说明那样，我们来做两层的串联. 如果这个码最下层的元件——实际的物理量子比特——的失效概率为 p，那么中间层(一个编码层)的失效概率最多为 cp^2，而最高层(两个编码层)——这一层中，若计算得到正确的输出，则线路必正确运行——的失效概率为 $c(cp^2)^2$. 因此，如果我们串联 k 层，那么最高层的过程失效概率为 $(cp)^{2^k}/c$，而模拟线路的尺寸则比原线路的尺寸高 d^k 倍，其中 d 为一个常数，表示为编码一个门并进行纠错的容错过程所用运算数目的最大值.

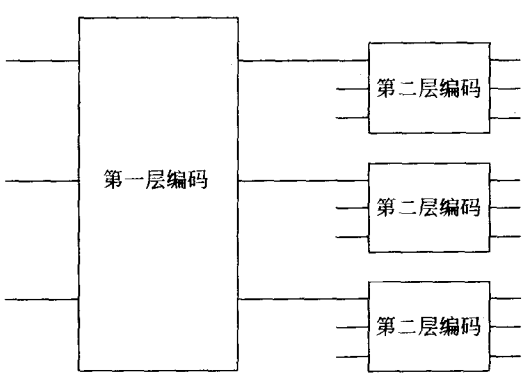


图 10.22　用 9 量子比特编码单量子比特的两层串联码. 为使图简单一些，我们这里采用 3 量子比特码；实际上，只要能纠正一个或多个量子比特上任意差错的一个码如 Steane 码均都可用.

然后，设我们想要来模拟一个含 $p(n)$ 个门的线路，其中 n 表征某个问题的尺寸，而 $p(n)$ 为 n 的一个多项式函数. 举例来说，这可以是量子分解因子算法的线路. 设在这个算法的仿真中，我们想要得到 ϵ 的最后精度. 为做到这一点，算法中每个门的仿真都必须精确到 $\epsilon/p(n)$，串联的级数 k 满足

$$\frac{(cp)^{2^k}}{c} \leqslant \frac{\epsilon}{p(n)} \tag{10.113}$$

如果 $p<p_{\text{th}}\equiv 1/c$，那么这样的 k 可以找到. 这个条件——$p<p_{\text{th}}$——称为量子计算的阈值条件，因为假定它被满足后，就能在量子计算中实现任意的精度. 需要多大规模的模拟线路才能得到这种精度呢？选择满足式(10.113)的最小的 k，使得(10.113)的不等式接近于饱和，并整理式(10.113)得到：

$$d^k \approx \left(\frac{\log(p(n)/c\epsilon)}{\log(1/pc)}\right)^{\log d} = O(\text{poly}(\log p(n)/\epsilon)) \tag{10.114}$$

其中 poly 表示固定次数的多项式. 因此，模拟线路包含门的个数为

$$O(\text{poly}(\log p(n)/\epsilon)p(n)) \tag{10.115}$$

它只是多项式地大于原线路的尺寸. 概括起来，我们有如下的量子计算的阈值定理.

量子计算的阈值定理 一个包含 $p(n)$ 门的量子线路，可以在硬件上应用

$$O(\text{poly}(\log p(n)/\epsilon)p(n)) \tag{10.116}$$

个门来模拟，其发生差错的概率最多为 ϵ. 其中，硬件元件的失效概率最多为 p，这里假设 p 低于某个定常阈值即 $p<p_{\text{th}}$，并且已对硬件中噪声做出了合理的假定.

p_{th} 的值是什么？对 Steane 码，根据我们的计算 $c\approx 10^4$，所以一个非常粗糙的估计为 $p_{\text{th}}\approx 10^{-4}$. 需要强调的是，我们的估计相比于严格的计算结果相差很远，但是对阈值复杂得多的计算已得到其典型值在 $10^{-5}\sim 10^{-6}$ 范围内. 注意，阈值的精确值在很大程度上依赖于对有关计算能力所作的假定. 举例来说，如果不可能采用并行运算，那么阈值条件是不可能达到的，因为对于纠错处理而言，差错在线路中积累得太快了. 在量子运算以外，对差错症状测量结果的处理，以及确定应用什么样的量子门来纠正差错等过程还需要经典计算. 对阈值估计的局限性的一些讨论在 10.6.4 节给出.

练习 10.62 试通过显式构造稳定子的生成元来证明，$[n_2,1]$ 稳定子码与 $[n_1,1]$ 稳定子码的串联会给出一个 $[n_1 n_2,1]$ 稳定子码.

10.6.2 容错量子逻辑

容错量子线路构造中的一个关键技术是，构造在编码后状态上作逻辑运算的容错运算方法. 在第 4 章的 4.5.3 节中，我们曾经学过，Hadamard 门、相位门、受

控非门以及 $\pi/8$ 门等组成一个通用的集合，任一量子计算都由这个集合描述. 我们现来解释这些门中的每个门可以如何容错地来实现.

1. 正规化子运算

我们从 Steane 码的具体例子——Hadamard 门、相位门和受控非门——的正规化子运算的容错构造开始. 通过了解这个具体例子中容错构造的基本原理，容易将其推广到任何一个稳定子码. 回顾式(10.107)，对 Steane 码，编码后状态上的 Pauli $\bar{Z}$ 和 $\bar{X}$ 算子可以根据未编码量子比特上的算子被写为

$$\bar{Z} = Z_1 Z_2 Z_3 Z_4 Z_5 Z_6 Z_7, \quad \bar{X} = X_1 X_2 X_3 X_4 X_5 X_6 X_7 \tag{10.117}$$

正如 Hadamard 门在共轭作用下对换 Z 和 X 那样，一个编码后 Hadamard 门 $\bar{H}$ 在共轭作用下应当对换 $\bar{Z}$ 和 $\bar{X}$. $\bar{H} = H_1 H_2 H_3 H_4 H_5 H_6 H_7$ 会实现这个任务，从而编码后量子比特上的 Hadamard 门可如图 10.23 所示那样而来实现.

练习 10.63　设 U 为映射 Steane 码到其自身的任一酉运算，并使成立 $U\bar{Z}U^\dagger = \bar{X}$ 且 $U\bar{X}U^\dagger = \bar{Z}$. 试证明，除了一个全局相位，$U$ 在编码后状态 $|0_L\rangle$ 和 $|1_L\rangle$ 上的作用为 $|0_L\rangle \to (|0_L\rangle + |1_L\rangle)/\sqrt{2}$ 和 $|1_L\rangle \to (|0_L\rangle - |1_L\rangle)/\sqrt{2}$.

这是一个好的开端，但只在编码后状态上作逻辑运算还不足以使这个运算容错，我们还需要了解差错是如何传播的. 因为实现 $\bar{H} = H^{\otimes 7}$ 的线路不包括多于一个的交互的编码后块中的量子比特，在物理上似乎可合理地假定，这个线路中，单个元件的失效在过程输出的量子比特块中可以最多引起一个差错. 为明白这是正确的，想象一个差错在编码后 H 门作用前瞬时出现在第一量子比特上. 为了明确起见，设这个差错为 Z 差错，则这个量子比特上的合运算为 HZ. 如在对受控非门的差错传播的先前分析中那样，插入单位阵 $H^\dagger H = I$，得出 $HZ = HZH^\dagger H = XH$，从而这样一个差错等价于首先作用 H 随后出现差错 X. 类似地，门运算本身期间的差错症状等价于一个完美的门跟随一个作用于量子比特上的小噪声，对通常的 X，Y 和 Z 模型我们认为所有的噪声都以某个小概率出现. 图 10.23 中的这个线路因此实际上就是一个容错运算，因为出现在过程中任何地方的单个失效不会传播去影响其他的量子比特，因而在过程输出的量子比特块中最多引起一个差错.

是否可从图 10.23 的线路中提取出一般原理呢？一个有用的事实是，如果编码后的那些门能用逐个比特方式来实现，这些门就必是自动容错的. 因为这个特性保证编码后门中任何地方的单个失效对这个码的每个块最多引入一个差错，因而差错概率不会超出纠错码的控制. 这个特性，即一个编码后门能用逐个比特方式来实现，称为编码后量子门的横截(transversality)特性. 横截性是令人感兴趣的，因为它对寻找容错量子线路会提供一般的设计原理. 我们下面会看到，Hadamard 门以外的许多门都能给出横截实现. 值得注意的是，还有可能找到非横截性的容错构造，例如我们下面将会看到的容错 $\pi/8$ 门的例子.

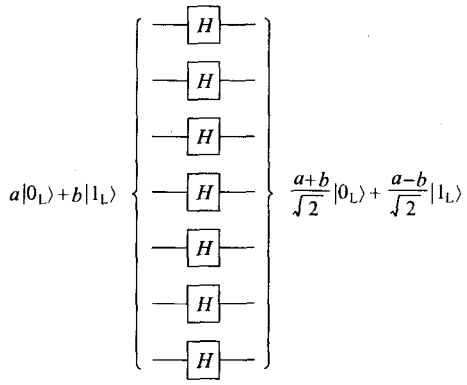


图 10.23 用 Steane 码编码的量子比特上的横截 Hadamaed 门.

应用 Steane 码，Hadamard 门以外的许多门都可容易地给出横截(因而是容错的)实现. 除了 Hadamard 门，三种最为令人感兴趣的门是相位门、Pauli X 门和 Pauli Z 门. 设我们对 Steane 码的 7 个量子比特的每一个逐个比特地作用 X 门. 这会在共轭作用下变换每个 Z 算子到 $-Z$ 算子，所以通过逐个比特共轭地作用 X 得到 $\overline{Z}\rightarrow(-1)^7\overline{Z}=-\overline{Z}$ 和 $\overline{X}\rightarrow\overline{X}$，因而，这个线路实现了 Steane 码状态上的编码后 X 运算. 这个线路是横截的，因而自动地是容错的. 按类似的方式，对 Steane 码的状态逐个比特地作用 Z 运算，就会给出编码后 Z 的一个容错实现. 相位门的横截实现稍有一点挑战性. 在共轭作用下，$\overline{S}$ 必须把 $\overline{Z}$ 变到 $\overline{Z}$ 而把 $\overline{X}$ 变到 $\overline{Y}=\mathrm{i}\overline{X}\overline{Z}$. 但是，应用显然的猜测，$\overline{S}=S_1S_2S_3S_4S_5S_6S_7$ 在共轭作用下把 $\overline{Z}$ 变到 $\overline{Z}$，而把 $\overline{X}$ 变到 $-\overline{Y}$. $-\overline{Y}$ 前面的这个负号可以通过作用 $\overline{Z}$ 来修正. 因此，对码中每个量子比特作用运算 ZS 就实现了横截的，也即是容错的编码后相位门.

与 Hadamard 门、Pauli 门和相位门形成对照的是，容错地实现受控非门乍看来似乎有挑战性，因为它包含两个分开的 7 量子比特码块. 我们如何来实现对这个码的每块不引入多于一个差错的受控非门呢？ 幸运的是，当采用 Steane 码时，做到这点是很简单的，如图 10.24 中所说明的：容易看出，通过 7 个受控非门成对地作用于两个块中的 7 量子比特之间而来实现. 读者可能会担心这种横截构造违反了我们自己的规则；难道，我们正在实现的这个受控非门不会在单量子比特之后引起差错传播吗？ 这是正确的，不过没有问题，因为差错传播仅影响另一个块中的最多一个量子比特；它不会对同一个块内的量子比特带来不利影响. 记住，影响其他块中的量子比特还不是很糟糕，因为每个块都能处理单量子比特上的差错.

更为精确地，设刚好在每块的第一量子比特之间的受控非门前，X 差错出现在第一量子比特上，将这些量子比特标记为 1 到 8. 如果表这个受控非门为 U，那么实际的作用为 $UX_1 = UX_1U^\dagger U = X_1X_8U$，也即受控非门仿佛正确地被作用，但一个 X 差错出现在编码后量子比特两块的第一量子比特上. 更具挑战性的是，设受控非门中的一个失效. 那么将会发生什么呢？设带噪声受控非门实现量子运算 ε，则其可重写为 $\varepsilon = \varepsilon \circ \mathcal{U}^{-1} \circ \mathcal{U}$，其中 $\mathcal{U}$ 为实现完美受控非门的量子运算. 因此，噪声污染受控非门等价于一个完美受控非门紧随运算 $\varepsilon \circ \mathcal{U}^{-1}$. 如果带噪声受控非门还不错，那么运算 $\varepsilon \circ \mathcal{U}^{-1}$ 近似地为单位阵，且可用通常的张量积的差错模型来理解它，例如两个量子比特上以某个小概率出现 $X \otimes Z$. 幸运的是，尽管这些差错包含两个量子比特，它们在编码后量子比特的每个块中却仅包含一个单量子比特. 关于其他位置上的差错传播可得到类似的结论. 这就导出，我们已描述过的过程内任何地方的单个元件的失效传播会在编码后量子比特的每块中造成不多于一个的差错，因而编码后受控非门的这种实现是容错的.

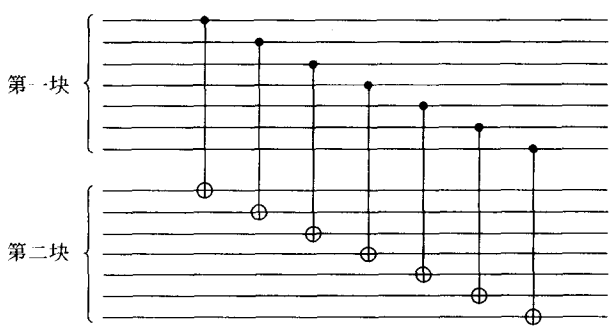


图 10.24　在分立块中用 Steane 码编码的两量子比特之间的横截受控非门.

在找到 Hadamard 门、相位门和受控非门的容错实现后，由定理 10.6 可以导出，正规化子中的任何运算都能容错实现. 当然正规化子运算没有穷尽为进行量子计算所要求的酉门的完全集合，但这是一个有希望的开端.

练习 10.64（差错的反向传播）　很清楚，受控非门的控制量子比特上的一个 X 差错会传播到目标量子比特，除此而外，事实上，目标量子比特上的一个 Z 差错也会反向传播到控制量子比特. 试用稳定子体系来证明这一点，再直接应用量子线路同一性证明这一点. 读者可能会发现，练习 10.20 也许是有用的.

2. 容错 $\pi/8$ 门

对于通用量子计算，我们所要求来完成的门的标准集合中，剩下的一个门就是 $\pi/8$ 门. 作为一种替代，如同 4.5.3 节中所注意到的，将容错 Toffoli 门加入到我们

现有容错的 Hadamard 门、相位门和受控非门的集合中也可提供我们一个通用的门集合，允许我们以容错的方式来执行量子计算机所要求的所有门. 事实上容错 $\pi/8$ 门的实现是很简单的，应用类似的但要较为精细的构造就能实现一个 Toffoli 门.

构造容错 $\pi/8$ 门的基本策略是将构造分为三个部分. 第一部分是一个简单的线路，这个线路采用我们已经知道如何容错的元件如受控非门、相位门和 X 门来模拟 $\pi/8$ 门. 但是，这个线路中还有两个部分我们尚不知道如何来做到容错. 第一是对加到线路中的输入来制备一个辅助状态. 为使辅助比特是合适的，要求辅助比特制备期间任一元件的失效只能在组成辅助部分的量子比特块中最多导致一个单差错. 本节的后面部分中，我们会来解释，如何制备这样的容错辅助比特. 我们所需要的第二种运算是测量. 为了使测量容错，我们要求测量过程期间的单个元件的失效应当不影响测量结果. 如果影响测量结果，那么差错就会传播而在第一块中的多个量子比特上引起差错，因为编码后 SX 运算为执行或不执行这一点是由测量结果来确定的. 如何来做这样一个容错测量会在下一节中进行描述(严格地说，对于我们描述的容错测量过程，实际上不正确的测量结果概率可能为 $O(p^2)$，其中 p 为单个元件失效的概率. 对于现在的讨论，我们将忽略这一点；沿类似的路线，应用稍微复杂的分析，不加忽略也是容易处理的).

图 10.25 显示了一个实现 $\pi/8$ 门的线路. 或许除虚线方框表示的元件和测量以外，线路中的所有元件都可容错地执行. 这个线路以两个编码后量子比特开始，其中一个是我们希望在其上运算的量子比特 $|\psi\rangle = a|0\rangle + b|1\rangle$(这里，令 $|0\rangle$ 和 $|1\rangle$ 表示逻辑状态). 另一个量子比特被制备于状态：

$$|\Theta\rangle = \frac{|0\rangle + \exp(i\pi/4)|1\rangle}{\sqrt{2}} \tag{10.118}$$

这是由图上虚线方框中的线路所生成的状态. 我们很快会来解释，这个辅助量子比特制备步骤是如何做到容错的. 下一步，执行一个容错受控非运算，给出

$$\begin{aligned}&\frac{1}{\sqrt{2}}[|0\rangle(a|0\rangle + b|1\rangle) + \exp(i\pi/4)|1\rangle(a|1\rangle + b|0\rangle)]\\ =&\frac{1}{\sqrt{2}}[(a|0\rangle + b\exp(i\pi/4)|1\rangle)|0\rangle + (b|0\rangle + a\exp(i\pi/4)|1\rangle)|1\rangle].\end{aligned} \tag{10.119}$$

最后，测量第二量子比特，若其为 0 则我们就完成；否则，对剩余的量子比特，执行运算

$$SX = \begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \tag{10.120}$$

在两种情况下，我们都得到状态 $a|0\rangle+b\exp(i\pi/4)|1\rangle$，除去一个无关的全局相位，这就如对 $\pi/8$ 门所要求的. 这个漂亮的结果可能看来不知从何而来，但事实上它是下面练习解释的一个系统构造的结果. 如同练习 10.68 中所显示的，同样的构造也用于实现容错 Toffoli 门.

容错 $\pi/8$ 门的构造要求一种生成辅助状态 $|\Theta\rangle$ 的容错方法. 这个制备可以应用容错测量技术来实现，下节会详细解释. 至于现在，我们要来解释它与容错测量的联系. 如图 10.25 所示，$|\Theta\rangle$ 可以由应用 Hadamard 门和然后再对状态 $|0\rangle$ 应用 $\pi/8$ 门而来生成. 状态 $|0\rangle$ 是 Z 的一个 $+1$ 本征态，所以导出 $|\Theta\rangle$ 是 $THZHT^\dagger = TXT^\dagger = e^{-i\pi/4}SX$ 的一个 $+1$ 本征态. $|\Theta\rangle$ 因此可通过首先制备编码后 $|0\rangle$，然后再容错地测量 $e^{-i\pi/4}SX$ 而制备. 如果得到结果 $+1$，那么我们断言状态已被正确地制备. 如果得到结果 -1，那么我们有两种选择. 或者重来，重复这个过程，直到 $e^{-i\pi/4}SX$ 的容错测量给出结果 $+1$；或者可以利用更为精致和有效的事实，由于 $ZSXZ=-SX$，则应用容错 Z 运算会改变状态，从 $e^{-i\pi/4}SX$ 的 -1 本征态到 $e^{-i\pi/4}SX$ 的 $+1$ 特征状态 $|\Theta\rangle$. 无论采用哪个过程，过程中任何地方的单个失效只在辅助状态 $|\Theta\rangle$ 的最多一个量子比特上生成一个差错.

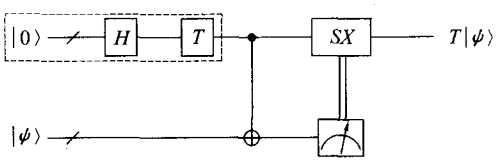
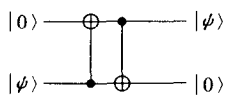


图 10.25　容错地实现 $\pi/8$ 门的量子线路. 虚线方框表示辅助状态 $(|0\rangle+\exp(i\pi/4)|1\rangle)/\sqrt{2}$ 的(非容错)制备过程，如何容错地做这个准备会在正文中加以解释. 连线上的斜线表示一个 7 量子比特束，双线连线表示从测量得到的经典比特. 注意，最后的 SX 运算是由测量结果所控制的.

不难看出，我们已经描述的过程作为整体是容错的；为看清这一点，不妨来看一个显式的例子，这可能会有用. 设一个单元件失效出现在辅助量子比特构造期间，并导致辅助量子比特中一个单量子比特上的一个差错. 这个差错通过编码后受控非门传播，在量子比特的第一个块和第二个块的每个量子比特中引起一个差错. 幸运的是，第二个编码后量子比特中单量子比特上的差错不会影响容错测量过程的结果，所以 SX 会适当地应用或不应用，因而量子比特的第一个块上的差错会通过传播在编码后门的输出中引起一个单差错. 类似地，不难相信，在编码后 $\pi/8$ 门的这个过程中其他任何地方的单个失效，只会在编码后量子比特输出块中的单量子比特上导致一个差错.

练习 10.65　状态为 $|\psi\rangle$ 的一个未知量子比特，只应用两个受控非门，就可与

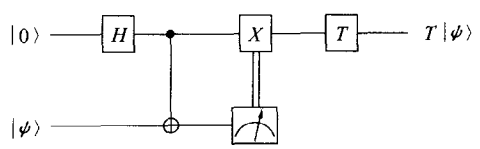
状态为$|0\rangle$的第二量子比特相对换，如下所示：



试证明：下面的只采用单个受控非门的两个线路，并有测量和经典的受控单量子比特运算，也会完成同样的任务.

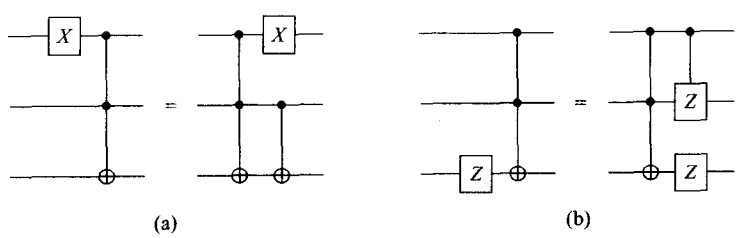


练习 10.66(容错 $\pi/8$ 门构造) 实现 $\pi/8$ 门的一种方法是，首先对想要变换的量子比特状态$|\psi\rangle$用某个已知状态$|0\rangle$对换，然后应用 $\pi/8$ 门在所得到的量子比特之上. 下面就是做到这个的一个量子线路：



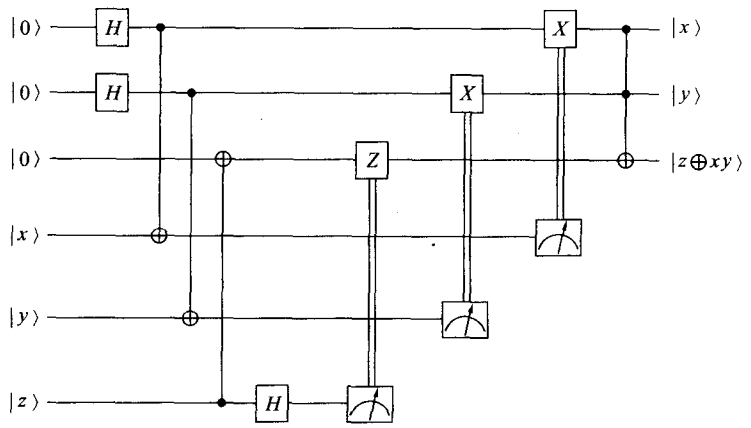
虽然只这样做看来并不是特别有用，其实会导致某些特别有用的东西. 试证明，应用关系 $TXT^{\dagger}=\exp(-i\pi/4)SX$ 和 $TU=UT$(U 为受控非门，T 作用于控制量子比特上)，我们可得到图 10.25 的线路.

练习 10.67 试证明，下述的线路等同：

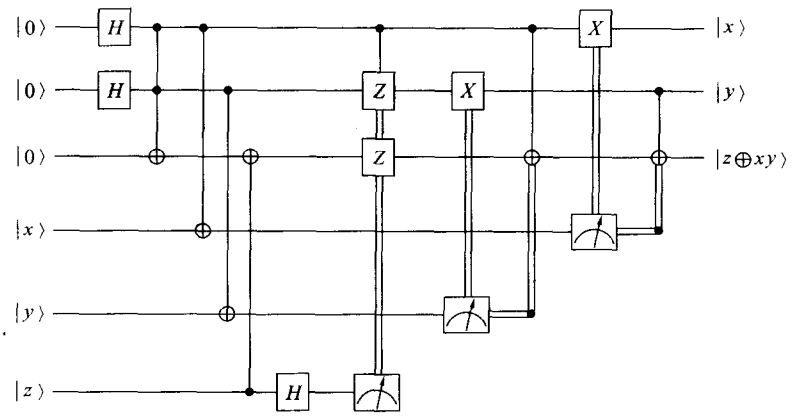


练习 10.68(容错 Toffoli 门构造) 类似于上面对 $\pi/8$ 门的一系列练习中的过程，会给出一个容错 Toffoli 门.

(1) 首先，用某个已知状态$|000\rangle$对换想要变换的量子比特状态$|xyz\rangle$，然后应用 Toffoli 门到得到的量子比特上. 试证明，下述线路可完成这个任务：



(2) 应用练习 10.67 中的对易规则，试证明移动最后的 Toffoli 门到最左边可给出如下的线路：



(3) 假定示于最左边虚线方框中的辅助量子码制备可做到容错，试证明应用 Steane 码，这个线路可以给出 Toffoli 门的一个容错实现.

10.6.3　容错测量

容错线路的构造中一个非常有用和重要的工具是测量 M 算子的能力. 测量用于编码，读出量子计算的结果，检测纠错中的差错症状以及在容错的 $\pi/8$ 门与 Toffoli 门的构造中进行辅助状态制备，因此测量对于容错量子计算绝对是非常关键的. 为使执行编码后测量的过程可认为是容错的，因此为防止差错传播，我们要求有两件事情应成立. 第一，过程中任何地方的单失效只能在过程末量子比特的任

一块中最多导致一个差错. 第二,即使单失效出现在过程中,我们要求测量结果正确的概率为 $1-O(p^2)$. 这后一个要求是非常重要的,因为测量结果可能用于控制量子计算机中的其他运算. 如果测量结果不正确,那么它就可能通过传播影响编码后量子比特的其他块中的许多量子比特.

回顾,单量子比特观测量 M 的测量可应用如图 10.26 所示的线路来执行. 设当逐个比特地对码的每个量子比特应用门 M',可给出 M 在量子码上的一个横截的编码后实现. 举例来说,对于 Steane 码,当逐个比特地应用 $M'=H$,可给出 $M=H$ 一个横截的实现,而 $M=S$ 的一个横截的实现采用逐个比特地应用 $M'=ZS$. 这就提出测量编码后数据上编码后 M 的一个可能线路,如图 10.27 所示. 注意,一个现实的量子码如 Steane 码会要求更多的量子比特. 不幸的是,图 10.27 中的线路并不是容错的. 为看清这一点,想象一个单失效出现在线路的最开始端即辅助量子比特上,它会向前传播并影响所有编码后量子比特,所以这个线路是非容错的.

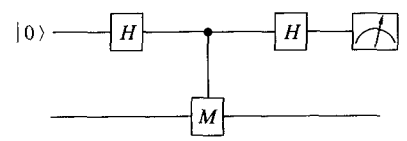


图 10.26 测量具有特征值±1 的单量子比特算子 M 的量子线路. 顶部量子比特为用于测量的辅助部分,底部量子比特正在被测量.

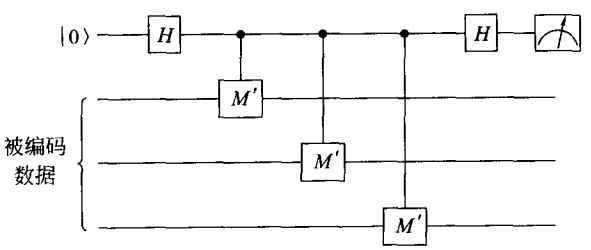


图 10.27 当逐个比特地应用 M',执行具有横截实现的编码后观测量 M 的测量的图解过程. 这个线路是非容错的. 注意,一个实际的码要求量子比特多于 3 个.

一个使测量线路为容错的好方法如图 10.28 所示. 为简单起见,这个图中仅用三个量子比特编码将被测量的数据;实际中将采用更多的量子比特,如 7 量子比特 Steane 码. 为具体起见,我们想象,这里用的正是 Steane 码. 除编码后数据外,这个线路对每个数据量子比特引入一个初始时各均处于状态 $|0\rangle$ 的辅助量子比特. 第一步是制备这个辅助量子比特于"猫"态 $|00\cdots0\rangle+|11\cdots1\rangle$. (注意,这个猫态没有用任何码来编码). 用来进行这个制备的线路自身不是容错的,因为线路期间的单失

效会在猫态中的多重量子比特上引起差错，然而，这不会影响整个过程的容错，因为紧随着辅助量子比特的制备我们应用了若干验证步骤(图中只显示一个验证步骤).

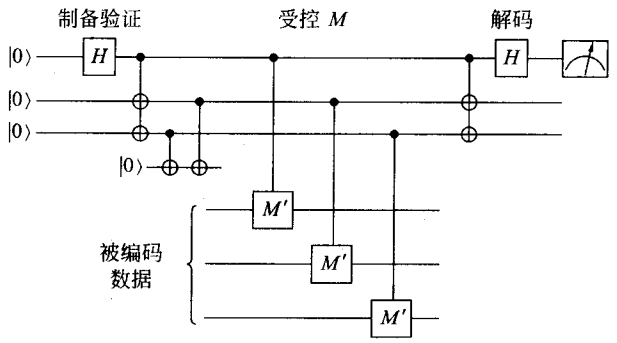


图 10.28　执行于编码后数据上的观测量 M 的容错测量的图解过程. 这个过程重复三次并对测量结果采用多数判决，结果中测量结果以概率 $O(p^2)$ 为错，其中 p 为任一单元件的失效概率. 线路中任何地方的单差错会在数据中最多生成一个差错.

验证按如下方式工作. 基本思想是，为验证状态为猫态，只要对猫态中所有量子比特对 i 和 j 的 Z_iZ_j 的测量给出 1 就足够了；也即，猫态中任一量子比特对的奇偶性为偶数. 为对一个特定的量子比特对 Z_iZ_j(例子中为 Z_2Z_3)来验证这点，我们引入初始状态为 $|0\rangle$ 的一个额外量子比特. 在测量这个额外量子比特之前，通过实现以辅助量子比特为控制和额外量子比特为目标的两个受控非门，我们就可计算辅助量子比特的每两个的奇偶性. 如果测量到的奇偶性为 1，那么我们就知道辅助部分没有处于猫态，将其丢弃并重新开始. 设一个单元件的失效出现在这一系列奇偶检验期间的某个地方. 这个过程不是容错的，因为容易证明存在单元件的失效导致在辅助状态中产生多于一个的相位翻转. 举例来说，如果受控非门之间的额外量子比特上存在一个 Z 差错，那么这个差错就会向前传播而在两个辅助量子比特引起 Z 差错. 幸运的是，容易证明，辅助量子比特中的多重 Z 差错不会传播到编码后数据，尽管它们可能造成最后的测量结果为不正确. 为解决这个问题，如同下面会较详细描述的，我们重复这个测量过程三次并取多数判决，所以这种方法中测量错两次或两次以上的概率最多为 $O(p^2)$，其中 p 为单元件失效的概率. 关于 X 差错和 Y 差错又会什么样呢？这些差错可以传播引起编码后数据中的差错，但一个幸运的事实是，猫态制备和验证期间的单失效只能在验证后的辅助量子比特中引起最多一个 X 差错或 Y 差错，因而在编码后数据中最多引起一个差错，这就保证了容错性.

练习 10.69　试证明，辅助量子比特制备和验证中任何地方的单失效可在辅助量子比特输出中最多导致一个 X 差错或 Y 差错.

练习 10.70 试证明，辅助量子比特中的 Z 差错不会传播以致影响编码后数据，但会导致观测到一个不正确的测量结果.

在猫态验证过后，受控 M' 门执行于辅助量子比特和数据量子比特对之间，其中没有辅助量子比特应用多于一次. 因此，如果辅助量子比特处于状态 $|00\cdots0\rangle$，这就导致对编码后数据什么也没有做；而如果辅助量子比特处于状态 $|11\cdots1\rangle$，编码后 M 运算就会作用于数据. 猫态的价值在于，它可保证差错不会从一个受控 M' 门传播到另一个受控 M' 门，所以验证阶段中或受控 M' 门序列中的单差错在编码后数据中最多导致一个单差错. 最后，测量结果可以通过应用一系列受控非门和一个 Hadamard 门对猫态解码而得到；所得到的量子比特为 0 或 1 取决于这个数据的状态的特征值. 这些最后的门不包含数据，因此这些门中的差错根本不会传播来影响数据. 但如果这个最后门序列中的差错导致不正确的测量结果，则又会什么样呢？通过重复测量过程三次并取结果的多数判决，我们就能保证，测量结果中有一个差错的概率为 $O(p^2)$，其中 p 为一个单元件失效的概率.

我们已经描述了执行容错测量的一种方法，使得测量以概率 $O(p^2)$ 给出不正确的结果，其中 p 为所有单元件的失效概率. 过程中任何地方的单失效最多只在编码后数据的一个量子比特上导致一个差错. 这个构造可应用于以横截方式来实现任何单量子比特观测量 M. 对于 Steane 码，这个横截方式包括 Hadamard 门、相位门和 Pauli 门，以及稍加修改的观测量 $M=e^{-i\pi/4}SX$. 为对 M 的这种选取来执行 Steane 码上的受控 M 运算，我们对辅助量子比特和码中的每个量子比特对横截地应用受控 ZSX 门，紧接着横截地作用于辅助量子比特的 7 个 T 门. 如同 10.6.2 节中所描述的，这个观测量的容错测量可被用来制作用于 $\pi/8$ 门的容错线路中的辅助量子比特.

练习 10.71 试验证，当 $M=e^{-i\pi/4}SX$ 时，我们已经描述的过程可给出测量 M 的一个容错方法.

练习 10.72（容错 Toffoli 辅助状态的构造） 说明如何容错地来制备由练习 10.68 的虚线方框中的线路所制作的状态，也即

$$\frac{|000\rangle+|010\rangle+|100\rangle+|111\rangle}{2} \tag{10.121}$$

你可能会发现，它有助于来首先给出这个状态的那些稳定子生成元. 如下介绍的是稳定子生成元的测量.

我们已描述了一个单量子比特的编码后观测量 M 的容错测量过程，这种技术可以立刻推广到其他的例子. 对于我们的目的，要能测量具有 Pauli 矩阵张量积形式的稳定子生成元就够了. 这样的测量允许我们执行容错纠错，对量子计算机的初始编码，以及测量计算的最后读出阶段的编码后 Z 算子.

作为一个简单例子，设我们想要测量一个例如 $X_1Z_2X_3$ 算子，它是用 Steane 码编码的 7 量子比特的一个块的前三个量子比特. 可以用图 10.28 的一个直接推

广来执行这个测量，如图 10.29 所示. 为实现对算子 $X_1Z_2X_3$ 的容错测量过程，在编码后数据上应用横截的受控运算之前，我们再一次来执行已验证的猫态制备. 具有了容错测量这样一些观测量的能力，我们就自动地获得执行编码、差错症状测量以及执行量子计算所要求的在(逻辑)计算基中测量等步骤的能力. 为了编码的目的，对量子计算而言，制备一个编码后 $|0\rangle$ 状态就足够了. 对于稳定子码如 Steane 码，根据 10.5.1 节的命题 10.4 证明中的步骤，通过容错测量所有稳定子生成元和编码后 $\bar{Z}$ 算子，然后应用适当的容错运算来调整稳定子生成元和编码后 Z 的正负号，就可完成这样一个制备. 一个说明 Steane 码编码后 $|0\rangle$ 状态如何可容错制备的例子在练习 10.73 中给出. 对纠错的差错症状测量，以及在量子计算机的编码后计算基中最后读出，可以用类似的路线来容错地实现.

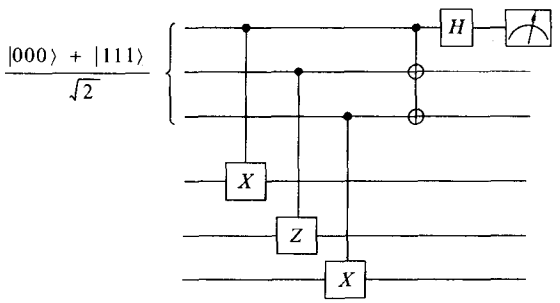


图 10.29　执行三个量子比特上算子 XZX 的容错测量的图解过程.

练习 10.73(容错编码后状态的构造)　试证明，Steane 码编码后 $|0\rangle$ 状态可按如下方式来容错地构造：

(1) 从图 10.16 的线路出发，并如图 10.30 所示那样，用猫态 $|00\cdots0\rangle+|11\cdots1\rangle$ 的每个辅助量子比特替换每个生成元的测量结果，且重新安排所有运算使它们的控制在不同量子比特上，从而使差错不会在码块内传播.

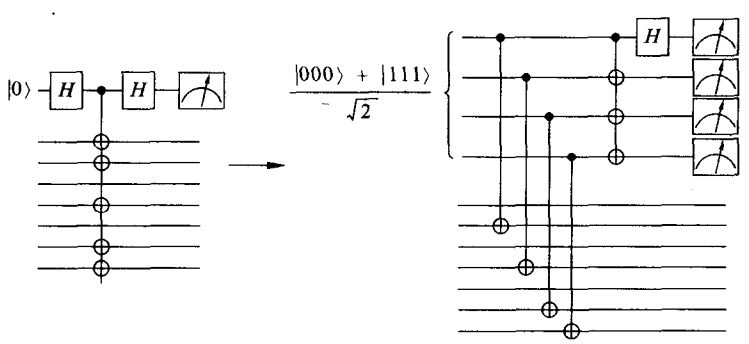


图 10.30　容错地生成 Steane 码编码后 $|0\rangle$ 状态中的一个步骤.

(2) 加入一个容错测量 $\bar{Z}$ 的阶段.

(3) 计算这个线路的差错概率,并当生成元测量重复三次并取多数判决时,计算这个线路的差错概率.

(4) 列举以测量结果为条件的运算,并证明这些运算可以做到容错.

练习 10.74 试构造一个量子线路,它对 5 量子比特码(10.5.6 节)容错地生成编码后 $|0\rangle$ 状态.

10.6.4 容错量子计算基础

量子纠错码最引人入胜的成功——量子计算的阈值——是当单独的量子门中的噪声低于某个定常阈值后,就有可能有效地执行任意的大量子计算. 换言之,对于量子计算,噪声在原则上并不是一个严重的问题. 阈值证明中的基本思想,如同 10.6.1 节中所概括的,是直接在编码后状态上执行容错运算,其中夹杂着纠错步骤,从而使差错概率从 p 净减为 $O(p^2)$. 通过把我们的码多次串联并制成分层的多个容错过程,只要原来的差错概率 p 小于某个阈值 p_{th},差错概率可进一步减少到 $O(p^4)$,再到 $O(p^8)$,如此下去,最后减少到所期望的低水平. 应用我们已经描述过的过程,可估计出阈值近似地为 $p_{\text{th}} \approx 10^{-5} \sim 10^{-6}$.

关于阈值定理的大胆断言显然需要限定词,即有可能完全防止任意噪声的影响不属于这种情况. 阈值定理依赖于物理上一些合理的假定,这些假定包括关于出现在量子计算机中噪声的类型,以及为达到其强有力结论而可利用的量子计算机结构. 已介绍的差错模型是相当简化的,实际量子计算机将会经历比这里所述的要多得多的各种类型噪声. 然而,可喜的是,我们这里所引入的这些技术,当与比较复杂的量子纠错码和与比较复杂的分析工具相结合时,可以导出一个比我们所述的更广泛环境中的量子计算的阈值.

我们这里没有篇幅去涉及更为复杂的分析,但是给出若干观察的事实还是适宜的. 第一,有趣的是阈值结果要求线路有很高度的并行性. 即使我们想做的一切只是存储量子信息到一个量子存储器中,这种运算也将要求需要高度并行的周期性纠错. 因此,对于愿意成为量子计算机设计者的人而言,一个能期望的目标是开发可并行化的结构,为的是使容错量子计算的技术可应用. 第二,我们注意到,我们对阈值的陈述完全忽略了在状态制备、差错症状测量和恢复期间所做的经典计算和通信的花费,这些花费可能是相当高的;举例来说,在串联码的最高层面上执行恢复,要求量子系统的所有部件之间进行通信. 如果这种通信的完成不能比差错在系统中出现更快,那么差错就将会蔓延开来而抵消纠错的效果. 更为复杂的分析就会涉及这个问题;但是,同其他复杂性相一致,对量子计算采用更严格的阈值形式会产生伴随花费. 第三,对测量和 $\pi/8$ 门的容错结构利用了状态 $|0\rangle$(或许还有较为

轻微附加噪声)的辅助量子比特.事实上,可以证明持续地提供这样的新辅助量子比特对于阈值定理的应用是必要的,因而量子计算机设计者必须提供的结构,不仅要可并行化,而且要允许新的辅助量子比特可不断地加入.

我们已做过的介绍只集中在基本原理上,而并非在优化所使用的方法上,很有可能实际中会采用更有效率的结构版本.一个简单但重要的指导性原理是选择合适的码.我们所以集中在 Steane 码,是因为它容易运行和证明所有基本原理;但是,实际上,其他码可能会工作得更好.举例来说,针对用于实现的实际物理系统,使用对已知类型噪声进行优化的码,就有可能带来在第一层串联上的很大好处.

尽管阈值定理中的理论思想,对于量子计算的特定实现,有着各种各样不同的使用方法,但怀疑论者可能仍然宣称,对所有这些阈值可被证明的噪声模型,都是过于保守的,而且将不能在任何实际物理系统范围内实现.这种怀疑最后只能用演示大规模容错量子计算的实验来回答.现有结果的惊人之处在于,就我们现有知识而言,没有一个物理原理会限制量子计算机在将来某一天被实现.

概括起来,本章中我们已经概要指出,量子信息处理的基本原理是以容错方式执行,并集中在量子计算的具体例子中.同样的基本技术也可用于其他能执行量子信息处理的系统,如执行像量子密码任务的量子信道.所有已知系统中的量子信息的极度脆弱性,使得看起来任何实际的量子信息处理系统都需要使用某种形式的量子纠错.但令人惊讶的是,这些技术工作得如此之好,以至于任意可靠的量子计算都可应用带噪声的元件来执行,只要这些元件中的差错概率小于某个定常阈值.

问题 10.1　若存在酉信道 $\mathcal{U}$ 和 $\mathcal{V}$,使成立 $\varepsilon_2=\mathcal{U}\circ\varepsilon_1\circ\mathcal{V}$,则信道 ε_1 和 ε_2 被称为等价.

(1) 试证明,信道等价的关系是一种等价关系.

(2) 试说明如何把 ε_1 的纠错码变为 ε_2 的纠错码.假定,ε_1 的纠错过程是通过一个投影测量及一个条件酉运算来进行的.试解释,ε_2 的纠错过程如何可用同样方式来执行.

问题 10.2(Gilbert-Vershamov 界)　试证明 CSS 码的 Gilbert-Vershamov 界,也即对某个 k,存在可纠正 t 个差错的一个$[n,k]$CSS 码,使成立

$$\frac{k}{n}\geqslant 1-2H\left(\frac{2t}{n}\right) \tag{10.122}$$

作为一个挑战,读者可能会愿意尝试证明对一般的稳定子码的 Gilbert-Vershamov 界,也即存在可纠正 t 个量子比特上差错的一个$[n,k]$稳定子码,使得

$$\frac{k}{n}\geqslant 1-\frac{2\log(3)t}{n}-2H\left(\frac{2t}{n}\right) \tag{10.123}$$

问题 10.3(编码稳定子码) 假定码的生成元处于标准形，并且编码后 Z 算子和 X 算子已被构造处于标准形. 试求一个线路，把对应于这个码的所有生成元的列表并连同编码后 Z 运算的 $n\times 2n$ 检验矩阵从

$$G=\left[\begin{array}{ccc|ccc}0&0&0&I&0&0\\0&0&0&0&I&0\\0&0&0&0&0&I\end{array}\right] \tag{10.124}$$

变到标准形

$$\left[\begin{array}{ccc|ccc}I&A_1&A_2&B&0&C_2\\0&0&0&D&I&E\\0&0&0&A_2^T&0&I\end{array}\right] \tag{10.125}$$

问题 10.4(通过隐形传态进行编码) 给定一个要用稳定子码编码的量子比特 $|\psi\rangle$，但我们对 $|\psi\rangle$ 是如何构造的一无所知：它是一个未知的状态. 试构造一个线路按下述方式来执行编码：

(1) 试说明，通过将其写成为稳定子状态，如何容错地构造这个部分编码后的状态：

$$\frac{|0\rangle|0_{\mathrm{L}}\rangle+|1\rangle|1_{\mathrm{L}}\rangle}{\sqrt{2}} \tag{10.126}$$

以使它可通过测量稳定子生成元而被制备.

(2) 试说明如何来对这个状态容错地执行一个 Bell 基测量，这个基具有 $|\psi\rangle$ 和这个状态的未编码量子比特.

(3) 试给出在这个测量之后来修正剩余量子比特的 Pauli 运算，从而它变为 $|\psi\rangle$，如同在通常的量子隐形传态方案中那样.

试计算，这个线路的差错概率，并说明如何修改这个线路以执行容错解码.

问题 10.5 设 $C(S)$ 为一个具有纠正单量子比特上差错能力的 $[n,1]$ 稳定子码. 试解释，仅采用容错稳定子状态制备、稳定子的元的容错测量以及横截作用的正规化子门，受控非门的一个容错实现，是如何在两个用这个码所编码的逻辑量子比特之间被实现的.

第 10 章小结 量子纠错

- **量子纠错码**：一个 $[n,k,d]$ 量子纠错码采用 n 个量子比特来编码 k 个量子比特，且具有距离 d.
- **量子纠错条件**：令 C 为一个量子纠错码，P 为到 C 上的投影算子，那么当且仅当对某个 Hermite 复数矩阵 α 有

$$PE_i^{\dagger}E_jP=\alpha_{ij}P \tag{10.127}$$

这个码能够纠正差错集合 $\{E_i\}$.

• **稳定子码**：令 S 为稳定子码 $C(S)$ 的稳定子，设 $\{E_j\}$ 为 Pauli 群中差错的集合，使对所有 j 和 k 有 $E_j^{\dagger}E_k \notin N(S)-S$，那么，$\{E_j\}$ 为码 $C(S)$ 的一个可纠正差错的集合.

• **容错量子计算**：编码后量子状态上的逻辑运算的一个通用集合可以用这样一种方式来执行，即编码后状态中的实际失效概率的大小如 $O(p^2)$，其中 p 为底层门的失效概率.

• **阈值定理**：规定单量子门中的噪声低于某个定常阈值，且噪声满足物理上的一定的合理假定，则在线路规模上只用很少的为保证可靠性所必要的一些附加开销，就有可能可靠地执行任意长的量子计算.

历史和进一步阅读的材料

在经典信息论中，已有许多关于纠错码的优秀教材. 我们要特别推荐 MacWilliams 和 Sloane 的极好的教材[MS77]. 这本教材的开始非常基础，但很快就平滑地进入到比较高级的课题，包括了很多的材料. 比较新一点的引论性著作，同时也是一本好著作，是由 Welsh 编著的教材[Wel88].

量子纠错是由 Shor[Sho95]（他发现了本章 10.2 节介绍的 9 量子比特码）和 Steane[MS77]（他采用不同方法并以此方法研究了多重粒子纠缠状态的干涉性质）所独立地发现的. 量子纠错条件是由 Bennett，Divincenro，Smolin 和 Wootters[BDSW96] 与 Knill 和 Laflamme[KL97] 分别独立证明的，它建立在 Ekert 和 Macchiavello 的早期工作[EM96] 上. 5 量子比特码由 Bennett，Divincenro，Smolin 和 Wootters[BDSW96] 所发现，并独立地由 Laflamme，Miquel，Paz 和 Zurek[LMPZ96] 给出.

Calderbank 和 Shor[CS96] 以及 Steane[Ste96b]，采用经典纠错思想来发展 CSS（Calderbank-Shor-Steane）码. Calderbank 和 Shor 也陈述和证明 CSS 码的 Gilbert-Varshamov 界. Gottesman[Got96] 发明了稳定子体系，并用其来定义稳定子码，还研究了一些特殊码的某些性质. Galderbank，Rains，Shor 和 Slone[CRSS97] 基于经典编码理论的思想，独立地发明了一种实质上等价于量子纠错的方法. 他们应用 $GF(4)$ 正交几何方法[CRSS98] 来分类几乎所有已知的量子码，并对一般稳定子码的 Gilbert-Varshamov 界提供了第一个证明，这个界较早时候由 Ekert 和 Macchiavello 所提出[EM96]. Gottesman-Knill 定理似乎由 Gottesman 在文献 [Got97]中所首先陈述，连同基于 Gottesman 所引入的稳定子体系的证明，文中他把结果归功于 Knill. Gottesman 还相当成功地把稳定子体系应用于广泛的各种各样的问题，例如参看文献[Got97]及其中的进一步参考文献. 我们对稳定子体系的阐述主要基于文献[Got97]，其中可找到我们所描述过的大部分结果，包括

Hadamard 门、相位门以及受控非门生成正规化子 $N(G_n)$ 等结果.

对于特殊类型的量子码的许多构造已为所知. 我们这里指出的只是很少的一些. Rains, Hardin, Shor 和 Sloane[RHSS97], 对我们所考虑过的稳定子码以外的那些量子码, 构造了一些感兴趣的例子. 许多人基于不同量子比特系统考虑了量子码; 我们要特别提到 Gottesman[Got98a] 和 Rains[Rai99b] 的工作, 这些工作构造了非二进制码, 并考虑了应用这类码的容错计算. Aharonov 和 Ben-Or[ABO99] 应用基于有限域上的多项式的有趣的技术构造了非二进制码, 并研究了应用这类码的容错计算. 近似量子纠错是我们没有涉及到的另一个课题; 近似量子纠错可以导致改善的码这一点是由 Leung, Nielson, Chuang 和 Yamamoto 所证明的[LNCY97].

一大类有趣的量子纠错码(但它已超出本章的范围), 有无噪声量子码(noiseless quantum code)和无退相干子空间(decoherence free subspace)等不同的名字. 这些课题(以及它们之间的联系)有着大量的工作要做. 入门性文献可通过 Zanardi 和 Rasetti[ZR98, Zan99], Lidar, Chuang 和 Whaley[LCW98], Bacon, Kempe, Lidar 和 Whaley[BKLW99, LBW99], 以及 Knill, Laflamme 和 Viola[KLV99] 的工作来找到.

关于量子纠错码的多个界已为所知, 通常是从类似的经典界修改而来的. Ekert 和 Macchiavello[EM96] 指出证明相似于 Hamming 界的量子结果的可能性; 这种构造和"退化"量子码的作用由 Gottesman[Got96] 随后所阐明. Shor 和 Laflamme[SL97] 证明了相应于经典编码理论中的一个结果, 即 MacWilliams 恒等式的量子结果, 这个恒等式激起过研究与量子码相关的某些多项式性质的大量工作(权重枚举)以及与量子码界问题有关的更为一般的工作, 包括 Ashikhmin[Ash97], Ashikhmin 和 Lytsin[AL99], 以及 Rains[Rai98, Rai99c, Rai99a] 关于这个课题的若干论文.

经典计算机的容错计算理论是由 von Neumann[Von56] 所作出的, 并在 Winograd 和 Cowan[WC67] 的专著中加以讨论. Shor[Sho96] 引入容错计算的思想到量子计算中, 并证明如何来执行所有的基本容错步骤(状态制备、量子逻辑、纠错以及测量等). Kitaev[Kit97a, Kit97b] 独立地发展了许多类似的思想, 包括对许多基本量子逻辑门的容错构造. Cirac, Pellizzari 和 Zoller[CPZ96], 以及 Zurek 和 Laflamme[ZL96], 也朝着容错量子计算采取了较早的步骤. DiVincenzo 和 Shor 推广了 Shor 的原来的构造, 以说明对任一稳定子码如何来执行差错症状的容错测量[DS96], 而 Gottesman[Got98b] 推广了所有容错构造并证明如何应用任一稳定子码来执行容错计算. 对这个工作的一般性评述以及许多其他综述材料可以在[Got97]中找到; 这包括为求解问题 10.5 的一个构造. 容错 $\pi/8$ 门和容错 Toffoli 门的构造是建立在由 Gottesman 和 Chuang[GC99] 与 Zhou 和 Chuang[ZC00] 所发展的思想基础上的; 描述于练习 10.68 中的容错 Toffoli 门的线路实际上就是 Shor 的原来的构造[Sho96]. Steane[Ste99] 对容错过程发展了许多有独创性的构造.

Kitaev[Kit97a, Kit97b] 采用拓扑方法辅助执行量子纠错, 对实现容错引入了一组漂

亮的思想. 其基本思想是，如果信息被存储于系统的拓扑中，那么信息对噪声的影响将自然地是鲁棒的. 这些以及许多其他极好的思想，已在 Bravyi 和 Kitaev[BK98b] 与 Freedman 和 Meyer[FM98] 的进一步的论文中得到发展. Preskill 的论文[Pre97] 是量子纠错领域中一篇优秀的综述，并有对拓扑量子纠错的一个特别漂亮的描述，以及争论性的讨论是否拓扑量子纠错可用来认识有关黑洞和量子重力等基础问题.

许多不同的小组证明过量子计算的阈值结果. 这些结果对广泛的各种各样的假定成立，并给出本质上不同的阈值定理. Aharonov 和 Ben-Or[ABO97, ABO99] 以及 Kitaev[Kit97c, Kit97b] 的阈值证明并不要求快速的和可靠的经典计算. Aharonov 和 Ben-Or 也证明，为使阈值结果成立，在每个时间步上量子计算机中必存在持续的并行性[ABO97]. Gottesman[Got97] 和 Preskill[Pre98c, GP00] 在他们的证明中，对阈值的值提供了一个特别详细的优化. Knill, Laflamme 和 Zurek[KLZ98a, KLZ98b] 的结果集中在对一类广泛的差错模型来证明阈值定理. Aharonov, Ben-Or, Impagliazzo 和 Nisan 也已证明[ABOIN96]，提供新鲜的量子比特对于阈值是必要的. 进一步的参考文献和历史材料可以在所引用的工作中找到. 实际上，每个组的研究都是建立在 Shor 的关于容错量子计算的先驱性工作[Sho96] 基础上的.

人们从各种各样的不同观点，已对容错量子计算写出了众多的优秀评述，发展了比我们所述的要详细得多的基本思想. Aharonov 的博士学位论文[Aha99a] 发展了阈值定理和自成体系的许多相当有兴趣的材料. Gottesman 的博士学位论文[Got97]，从更为强调量子码特性角度也对容错量子计算提供了一个评述，并对广泛的各种各样不同的码发展了容错结构. Knill, Laflamme 和 Zurek 就阈值结果写出了不完全流行的一个总结性看法[KLZ98a]. 最后，Preskill 写了两篇极好的文章[Pre98c, Pre98a] 来解释量子纠错和容错量子计算.

CHAPTER 11
第 11 章
熵与信息

熵是量子信息理论的关键概念之一，它用来度量物理系统的状态所包含的不确定性. 本章中我们将综述经典和量子信息理论中熵的定义和基本性质. 本章部分内容包含相当详细和篇幅较大的数学论述. 首次阅读这几节可以浏览，以后需要参考时再回来仔细阅读.

11.1 Shannon 熵

Shannon 熵是经典信息论的关键概念. 假定我们要获取随机变量 X 的值，X 的 Shannon 熵是对平均而言在获取 X 的值的过程中我们得到信息多少的度量. 换个说法，X 的 Shannon 熵是在我们得到 X 的值之前关于 X 的不确定性的测度. 这两种观点是互补的. 我们既可以把熵视为在得到 X 的值之前不确定性的一种测度，又可以把它视为得到 X 值之后我们得到信息多少的一种测度.

直观上，随机变量包含的信息应不依赖于随机变量可取值的标号. 例如，我们认为，一个取值“正面”和“反面”的概率分别为 1/4 和 3/4 的随机变量与一个取值 0 和 1 的概率分别为 1/4 和 3/4 的随机变量包含等量的信息. 因此，随机变量的熵定义为该随机变量取不同值的概率的一个函数，而不受这些值的标号影响. 我们常将熵写作概率分布 $p_1,\cdots,p_n$ 的函数，与该概率分布相联系的 Shannon 熵定义为

$$H(X)\equiv H(p_1,\cdots,p_n)\equiv-\sum_x p_x\log p_x \tag{11.1}$$

下面会很快对该定义给出解释. 注意定义中——并且贯穿全书——对数符号“log”是以 2 为底的，而对数符号“ln”指自然对数. 按照这种约定，习惯上我们说熵是以“比特”测度的. 人们也许会问 $p_x=0$ 的时候怎么办，因为 log0 没有定义. 直观上讲，从不发生的事件应该对熵没有贡献，故约定 $0\log 0\equiv 0$. 更正规地，注意到 $\lim_{x\to 0} x\log x=0$，这进一步支持了我们的约定.

为什么这样来定义熵？本节后面的练习 11.2 基于对信息度量所期望的一

组合理的公理，为此定义提供了一个直观解释. 这个直观说明起到加强的作用，但还不完整. 采用熵的该定义的最好理由是它能够用于量化存储信息所需要的资源. 更具体地，假定有某个源（或许是无线电天线）在产生某种信息，比如说是比特串的形式. 考虑一个源的非常简单的模型：源的模型是产生独立同分布随机变量的序列 $X_1, X_2, \cdots$，大多数实际的信源并不完全是这样，但这确实常常是一个好的近似. Shannon 的研究是问，为将来能重构出该信源产生的信息最少需要多少物理资源来存储这些信息？这个问题的答案恰好就是熵，即每个信源符号需要 $H(X)$ 比特，其中 $H(X) \equiv H(X_1) = H(X_2) = \cdots$ 是描述信源的每个随机变量的熵. 这个结论称为 Shannon *无噪声信道编码定理*，我们会在第 12 章证明它的经典和量子版本.

作为 Shannon 无噪声信道编码定理的具体例子，假设一个信源产生四个符号 1,2,3 和 4. 不经压缩时每次使用该信源，为存储四个可能的输出要消耗两比特的存储空间. 然而假设信源产生符号 1 的概率是 1/2，产生符号 2 的概率是 1/4，而产生符号 3 和 4 的概率各为 1/8. 我们就有可能利用信源在输出上的不对称性，以较少的比特来存储常用的符号如 1，而用较多的比特来存储较少使用的符号如 3 和 4. 一种方案是将 1 编码为比特串 0，将 2 编为 10，3 编为比特串 110，而 4 编为比特串 111. 注意压缩后每次使用信源串的平均长度为 $\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3 = 7/4$ 比特的信息. 这比存储该信源的直接方法需要的比特数要少. 惊人的是，这与信源的熵完全一致，$H(X) = -1/2\log(1/2) - 1/4\log(1/4) - 1/8\log(1/8) - 1/8\log(1/8) = 7/4$！而且事实上任何进一步的压缩都会带来不可恢复的损失；熵定量描述了可以达到的最优压缩.

这里从数据压缩角度对熵的定义作的操作性分析表达了量子和经典信息论共同的一个关键思想：*信息的基本测度源自对解决某些信息处理问题所需要物理资源的基本问题的回答*.

练习 11.1（熵的简单计算）　求与投掷一枚均匀硬币相关联的熵，求与掷均匀骰子相关联的熵. 如果硬币或骰子是不均匀的，熵有何不同？

练习 11.2（熵的定义的直观解释）　我们要定量描述在一个概率实验中发生的事件 E 提供多少信息，利用取值决定于该事件 E 的信息函数 $I(E)$ 来做. 我们对这个函数作如下假设：

(1) $I(E)$ 仅是事件 E 的概率的函数，故可以写作 $I = I(p)$，其中 p 是概率，其值在 0 到 1 范围之内.

(2) I 是概率的一个光滑函数.

(3) 当 $p, q > 0$，$I(pq) = I(p) + I(q)$（解释：从两个分别具有概率 p 和 q 的独立事件得到的信息是从单独事件得到信息之和）.

证明 $I(p)=k\log p$ 对某个常数 k 成立. 从而从互斥的一组发生概率为 $p_1,\cdots,p_n$ 的事件得到的平均信息为 $k\sum_i p_i\log p_i$，除一个常数因子，这正是 Shannon 熵.

盒子 11.1 熵量子测不准原理

可以用熵的形式给出量子力学测不准原理的一个优雅的表述. 回忆盒子 2.4 中的 Heisenberg 测不准原理. 它断言对处于状态 $|\psi\rangle$ 的量子系统，可观测量 C 和 D 的标准差 $\Delta(C)$ 和 $\Delta(D)$ 必满足关系：

$$\Delta(C)\Delta(D)\geqslant\frac{|\langle\psi|[C,D]|\psi\rangle|}{2} \tag{11.2}$$

令 $C=\sum_c c|c\rangle\langle c|$ 和 $D=\sum_d d|d\rangle\langle d|$ 分别为 C 和 D 的谱分解. 定义 $f(C,D)\equiv\max_{c,d}|\langle c|d\rangle|$ 为任意两个特征向量 $|c\rangle$ 和 $|d\rangle$ 之间的最大忠实度，例如对 Pauli 矩阵 X 和 Z 有 $f(X,Z)=1/\sqrt{2}$.

设量子系统状态被制备为 $|\psi\rangle$，且令 $p(c)$ 为与测量 C 所关联的概率分布，相应的熵为 $H(C)$，$q(d)$ 为与测量 D 所关联的概率分布，相应的熵为 $H(D)$. 熵测不准原理表述为

$$H(C)+H(D)\geqslant 2\log\left(\frac{1}{f(C,D)}\right) \tag{11.3}$$

该结果的完整证明偏离我们的主题太远(请参考“历史和进一步阅读的材料”)；不过，我们可以给出一个较弱结论的证明：

$$H(C)+H(D)\geqslant -2\log\frac{1+f(C,D)}{2} \tag{11.4}$$

为证明该式，注意到

$$H(C)+H(D)=-\sum_{cd}p(c)q(d)\log(p(c)q(d)) \tag{11.5}$$

我们要估计 $p(c)q(d)=|\langle c|\psi\rangle\langle\psi|d\rangle|^2$ 的上界. 为此，令 $|\tilde{\psi}\rangle$ 为 $|\psi\rangle$ 在 $|c\rangle$ 和 $|d\rangle$ 张成的平面上的投影，于是 $|\tilde{\psi}\rangle$ 的模 λ 不超过 1. 若 θ 为 $|d\rangle$ 与 $|c\rangle$ 在该平面成的夹角，φ 为 $|\tilde{\psi}\rangle$ 与 $|d\rangle$ 的夹角，则可知 $p(c)q(d)=|\langle c|\tilde{\psi}\rangle\langle\tilde{\psi}|d\rangle|^2=\lambda^2\cos^2(\theta-\varphi)\cdot\cos^2(\varphi)$. 计算表明最大值在 $\lambda=1$ 且 $\varphi=\theta/2$ 时达到，此时 $p(c)q(d)=\cos^4(\theta/2)$，并可写作

$$p(c)q(d)=\left(\frac{1+|\langle c|d\rangle|}{2}\right)^2 \tag{11.6}$$

与式(11.5)合并得到

$$H(C)+H(D)\geqslant -2\log\frac{1+f(C,D)}{2} \tag{11.7}$$

如所欲证.

11.2　熵的基本属性

11.2.1　二元熵

二输出随机变量的熵非常有用，因此我们给它起一个特殊的名称，叫做二元熵，定义为

$$H_{\text{bin}}(p) \equiv -p\log p-(1-p)\log(1-p) \tag{11.8}$$

其中 p 和 $1-p$ 是两个输出的概率. 当上下文清楚的时候，用 $H(p)$ 代替 $H_{\text{bin}}(p)$. 二元熵函数如图 11.1 所示. 注意 $H(p)=H(1-p)$ 且 $H(p)$ 在 $p=1/2$ 处取到最大值 1.

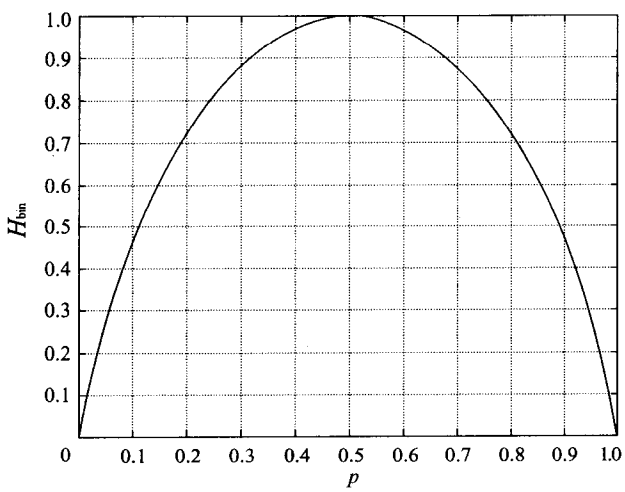


图 11.1　二元熵函数 $H(p)$.

二元熵为理解熵的一般性质提供了极好的例子. 熵在两个或更多概率分布混合时的行为如何是一个非常重要的性质. 设想 Alice 有两枚硬币，一枚是 25 美分，另一枚是一澳元. 两枚硬币都不均匀，其中美分出现正面的概率是 p_{U}，而澳币出现正面的概率是 p_{A}，假设 Alice 投美分的概率是 q，投澳币的概率是 $1-q$，Alice 告诉 Bob 结果是正面还是反面，平均而言 Bob 获得了多少信息？直观上，Bob 至少会得到从投美分或澳币所获得的信息. 用不等式表达即为

$$H(qp_{\text{U}}+(1-q)p_{\text{A}}) \geqslant qH(p_{\text{U}})+(1-q)H(p_{\text{A}}) \tag{11.9}$$

有时不等式是严格的，因为 Bob 不仅获得了硬币值（正或反）的信息，还获得了硬币类型的某些附加信息. 例如，如果 $p_{\text{U}}=1/3$，$p_{\text{A}}=5/6$，而且出现了正面，那么这就

指示 Bob 该币很可能是澳币.

容易证明式(11.9)是成立的,它是更一般的凹性概念的一个例子. 我们在第 9 章中讨论距离度量时曾遇到过凹性. 回忆一个实函数 f 称为凹,若对 0 到 1 中任意 p,有

$$f(px+(1-p)y) \geqslant pf(x)+(1-p)f(y) \tag{11.10}$$

从图 11.1 上易见二元熵是凹的,因为二元熵的图总位于图截线之上. 我们将非常重视经典和量子熵的凹性. 不要被上述直观说明所误导:量子信息的许多深刻结果都植根于经典和量子熵凹性的巧妙应用. 而且,对量子熵有时很难从直观上说清熵应该具备的凹性.

练习 11.3 证明二元熵 $H_{\text{bin}}(p)$ 在 $p=1/2$ 处达到其最大值 1.

练习 11.4(二元熵的凹性) 从图 11.1 可见二元熵是凹函数. 证明确实如此,即

$$H_{\text{bin}}(px_1+(1-p)x_2) \geqslant pH_{\text{bin}}(x_1)+(1-p)H_{\text{bin}}(x_2) \tag{11.11}$$

其中 $0 \leqslant p, x_1, x_2 \leqslant 1$. 证明此外二元熵还是*严格凹的*,即上述不等式仅在 $x_1=x_2$ 或 $p=0$,或 $p=1$ 的平凡情形为等式.

11.2.2 相对熵

相对熵是一个非常有用的熵型度量,它度量了两个定义在同一指标集 x 上的概率分布 $p(x)$ 和 $q(x)$ 之间接近程度. $p(x)$ 到 $q(x)$ 的相对熵定义为

$$\begin{aligned} H(p(x) \| q(x)) &\equiv \sum_x p(x)\log\frac{p(x)}{q(x)} \\ &\equiv -H(X)-\sum_x p(x)\log q(x) \end{aligned} \tag{11.12}$$

我们定义 $-0\log 0 \equiv 0$,且若 $p(x)>0$ 则 $-p(x)\log 0 \equiv +\infty$.

相对熵的用途,甚至为什么说它是两个分布间距离的好的度量,这并不是非常显然. 下面的定理提供了相对熵为什么作为某种距离测度的一些背景.

定理 11.1(相对熵的非负性) 相对熵是非负的,$H(p(x) \| q(x)) \geqslant 0$,且当且仅当 $p(x)=q(x)$ 对所有 x 成立时取等号.

证 信息论中有一个非常有用的不等式,$\log x \ln 2 = \ln x \leqslant x-1$,对所有正数 x 成立,且当且仅当 $x=1$ 取等号. 这里我们需要对该结果稍作整理,变成 $-\log x \geqslant (1-x)/\ln 2$,并注意到

$$H(p(x) \| q(x)) = -\sum_x p(x)\log\frac{q(x)}{p(x)} \tag{11.13}$$

$$\geqslant \frac{1}{\ln 2}\sum_x p(x)\left(1-\frac{q(x)}{p(x)}\right) \tag{11.14}$$

$$= \frac{1}{\ln 2}\sum_x (p(x) - q(x)) \tag{11.15}$$

$$= \frac{1}{\ln 2}(1-1) = 0 \tag{11.16}$$

为所欲证. 注意到第二行等式成立当且仅当 $q(x)/p(x)=1$ 对所有 x 成立，即两个分布完全相同. □

相对熵的用途常不在于其本身，而在于其他熵型量可以被视为相对熵的特例. 相对熵的结果就可导出其他熵型量的特殊结果. 例如，我们可以用相对熵的非负性证明如下关于熵的基本事实. 设 $p(x)$ 是 X 的一个具有 d 个结果的概率分布，令 $q(x)\equiv 1/d$ 为这些结果上的均匀分布，则

$$\begin{aligned} H(p(x) \parallel q(x)) &= H(p(x) \parallel 1/d) \\ &= -H(X) - \sum_x p(x)\log(1/d) \\ &= \log d - H(X) \end{aligned} \tag{11.17}$$

由相对熵的非负性，即定理 11.1，可见 $\log d - H(X) \geqslant 0$，并且当且仅当 X 为均匀分布时取等号. 这是一个基本事实. 由于其重要性，我们将其重新叙述为一个定理.

定理 11.2　设 X 是一个具有 d 个结果的随机变量，则 $H(X) \leqslant \log d$，并且当且仅当 X 为这 d 个结果上的均匀分布时取等号.

练习 11.5（Shannon 熵的次可加性）　证明 $H(p(x,y) \parallel p(x)p(y)) = H(p(x)) + H(p(y)) - H(p(x,y))$；由此证明 $H(X,Y) \leqslant H(X) + H(Y)$，并且当且仅当 X 和 Y 是独立随机变量时取等号.

11.2.3　条件熵和互信息

设 X 和 Y 是两个随机变量，X 包含的信息与 Y 包含的信息有什么关系呢？本节引入两个概念——*条件熵*和*互信息*——来回答这个问题. 我们给出的定义相当形式化，或许读者会困惑，一个特别的量如条件熵为什么可以按我们指出的方式来解释. 请牢记这些定义的最终合理性是它们回答了资源问题. 我们将在第 12 章作出更详细的讨论，对这些量的解释也和所要回答的资源问题性质有关.

上节我们已经间接地遇到了一对随机变量的联合熵，为清楚起见，我们显式地给出定义. X 和 Y 的联合熵直接定义为

$$H(X,Y) \equiv -\sum_{x,y} p(x,y)\log p(x,y) \tag{11.18}$$

且可以直接扩展到任意由随机变量组成的向量. 联合熵测量了我们关于 (X,Y) 对的整体不确定性. 假设我们知道 Y 的值，那么我们已得到了关于该对 (X,Y) 的 $H(Y)$ 比特的信息. 关于 (X,Y) 对的剩余不确定性与我们所缺乏的关于 X 的知识相联系，即使我们知道 Y. 于是*已知 Y 条件下 X 的熵*定义为

$$H(X\mid Y)\equiv H(X,Y)-H(Y) \tag{11.19}$$

条件熵是已知 Y 值条件下，平均而言我们关于 X 值的不确定程度的度量.

第二个量，X 和 Y 的互信息量，测量 X 和 Y 包含多少共同信息. 假设我们把 X 包含的信息 $H(X)$ 加到 Y 包含的信息中，X 和 Y 中的共同信息将被计算两次，而不同的信息恰好被计算一次. 减去 (X,Y) 的联合信息 $H(X,Y)$，就得到 X 和 Y 的共同信息或称互信息

$$H(X:Y)\equiv H(X)+H(Y)-H(X,Y) \tag{11.20}$$

注意这个将条件熵和互信息联系起来的有用等式 $H(X:Y)=H(X)-H(X|Y)$.

为得到对 Shannon 熵行为的一些认识，我们给出不同熵之间的若干简单关系.

定理 11.3(Shannon 熵的基本性质)

(1) $H(X,Y)=H(Y,X)$，$H(X:Y)=H(Y:X)$.

(2) $H(Y|X)\geqslant 0$ 从而 $H(X:Y)\leqslant H(Y)$，且当且仅当 Y 是 X 的函数 $Y=f(X)$ 时取等号.

(3) $H(X)\leqslant H(X,Y)$，且当且仅当 Y 是 X 的函数时取等号.

(4) 次可加性 $H(X,Y)\leqslant H(X)+H(Y)$，且当且仅当 X 和 Y 是独立随机变量时取等号.

(5) $H(Y|X)\leqslant H(Y)$ 从而 $H(X:Y)\geqslant 0$，且当且仅当 X 和 Y 是独立随机变量时两式取等号.

(6) 强次可加性 $H(X,Y,Z)+H(Y)\leqslant H(X,Y)+H(Y,Z)$，且当且仅当 $Z\to Y\to X$ 构成 Markov 链时取等号.

(7) 条件减小熵 $H(X|Y,Z)\leqslant H(X|Y)$.

大部分的证明要么显然，要么属于简单练习，下面给出一些提示.

证 (1) 根据有关定义，结果很显然.

(2) 由 $p(x,y)=p(x)p(y|x)$ 知

$$H(X,Y)=-\sum_{xy}p(x,y)\log p(x)p(y\mid x) \tag{11.21}$$

$$=-\sum_{x}p(x)\log p(x)-\sum_{xy}p(x,y)\log p(y\mid x) \tag{11.22}$$

$$=H(X)-\sum_{xy}p(x,y)\log p(y\mid x) \tag{11.23}$$

因此 $H(Y\mid X)=-\sum_{xy}p(x,y)\log p(y\mid x)$. 但 $-\log p(y|x)\geqslant 0$，故 $H(Y|X)\geqslant 0$，且当且仅当 Y 是 X 的一个确定性函数时取等号.

(3) 由前面结果直接得到.

(4) 为证次可加性和随后的强次可加性，我们再次使用对所有正的 x，$\log x\leqslant$

$(x-1)/\ln 2$，且当且仅当 $x=1$ 时取等号的事实，我们有

$$\sum_{x,y} p(x,y)\log\frac{p(x)p(y)}{p(x,y)} \leqslant \frac{1}{\ln 2}\sum_{x,y} p(x,y)\left(\frac{p(x)p(y)}{p(x,y)}-1\right) \tag{11.24}$$

$$= \frac{1}{\ln 2}\sum_{x,y} p(x)p(y)-p(x,y)$$

$$= \frac{1-1}{\ln 2} = 0 \tag{11.25}$$

于是得到次可加性. 注意当且仅当 $p(x,y)=p(x)p(y)$ 对所有 x,y 成立，该式取等号，即当且仅当 X 和 Y 独立次可加不等式达到饱和.

(5) 从次可加性和相关定义可得.

(6) Shannon 熵的强次可加性证明技术与证明次可加性相同；证明难度稍高. 练习 11.6 请读者提供这个证明.

(7) 直观上，我们期望在已知 Y,Z 的值条件下，关于 X 的不确定性会低于仅仅知道 Y 时 X 的不确定性. 更形式化来看，代入有关定义，条件减小熵的结论等价于

$$H(X,Y,Z)-H(Y,Z) \leqslant H(X,Y)-H(Y) \tag{11.26}$$

它是从强次可加不等式整理得到的. □

练习 11.6(经典强次可加性的证明)　证明 $H(X,Y,Z)+H(Y)\leqslant H(X,Y)+H(Y,Z)$，且当且仅当 $Z\to Y\to X$ 构成 Markov 链时取等号.

练习 11.7　练习 11.5 中，我们间接证明了互信息 $H(X:Y)$ 可以表为两个概率分布相对熵，即 $H(X:Y)=H(p(x,y)\,\|\,p(x)p(y))$. 求条件熵 $H(Y|X)$ 为两个概率分布间相对熵的表达式，以此式推断 $H(Y|X)\geqslant 0$，并求相等的条件.

熵的各种关系大部分可以从图 11.2 所示熵 Venn 图导出. 虽然这种图在引导熵的性质时并不完全可靠，但是它们可以帮助记忆熵的各种定义和性质.

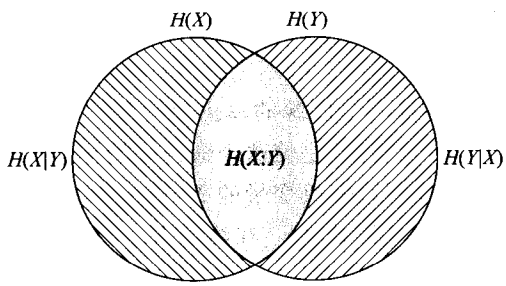


图 11.2　各种熵之间的关系.

我们以如下关于条件熵的简单而有用的链式法则来结束本节条件熵和互信息基本性质的讨论.

定理 11.4(条件熵的链式法则) 令 $X_1,\cdots,X_n$ 和 Y 为任意一组随机变量,则

$$H(X_1,\cdots,X_n \mid Y)=\sum_{i=1}^{n}H(X_i \mid Y,X_1,\cdots,X_{i-1}) \tag{11.27}$$

证 我们先对 $n=2$ 证明,再对 n 用归纳法. 直接从定义出发,经过简单计算得

$$H(X_1,X_2 \mid Y)=H(X_1,X_2,Y)-H(Y) \tag{11.28}$$

$$=H(X_1,X_2,Y)-H(X_1,Y)+H(X_1,Y)-H(Y) \tag{11.29}$$

$$=H(X_2 \mid Y,X_1)+H(X_1 \mid Y) \tag{11.30}$$

即 $n=2$ 时成立. 假设结果对一般的 n 成立,然后证结果对 $n+1$ 成立. 根据已证明的 $n=2$ 的情形,得

$$H(X_1,\cdots,X_{n+1} \mid Y)=H(X_2,\cdots,X_{n+1} \mid Y,X_1)+H(X_1 \mid Y) \tag{11.31}$$

对右边第一项应用归纳假设导出

$$H(X_1,\cdots,X_{n+1} \mid Y)=\sum_{i=2}^{n+1}H(X_i \mid Y,X_1,\cdots,X_{i-1})+H(X_1 \mid Y) \tag{11.32}$$

$$=\sum_{i=1}^{n+1}H(X_i \mid Y,X_1,\cdots,X_{i-1}) \tag{11.33}$$

归纳完成. □

练习 11.8(互信息未必总是次可加的) 令 X 和 Y 为独立同分布随机变量,值为 0 和 1 概率分别为 1/2,令 $Z\equiv X\oplus Y$,其中$\oplus$为模 2 加,证明在这种情况下,互信息不是次可加的,

$$H(X,Y:Z)\not\leqslant H(X:Z)+H(Y:Z) \tag{11.34}$$

练习 11.9(互信息未必总是超可加的) 令 X_1 为以 1/2 概率取值 0 和 1 的随机变量且 $X_2\equiv Y_1\equiv Y_2\equiv X_1$,证明这种情况下互信息是非超可加的,

$$H(X_1:Y_1)+H(X_2:Y_2)\not\leqslant H(X_1,X_2:Y_1,Y_2) \tag{11.35}$$

11.2.4 数据处理不等式

在许多应用中,我们在包含噪声的不完整信息上进行计算. 信息论的一个基本不等式*数据处理不等式*,断言一个信源的输出包含的信息只能随时间递减:信息一旦丢失,将永远消失. 本节的目的是使这个命题确切化.

*信息处理*的直观概念体现在随机变量的 Markov 链模型中. Markov 链是随机变量的一个序列 $X_1\to X_2\to\cdots$,满足 X_{n+1} 在给定 X_n 条件下与 $X_1,\cdots,X_{n-1}$ 独立,其形式为

$$p(X_{n+1}=x_{n+1} \mid X_n=x_n,\cdots,X_1=x_1)$$

$$=p(X_{n+1}=x_{n+1} \mid X_n=x_n) \tag{11.36}$$

随时间推进,在什么条件下一个 Markov 链会丢失它以前值的信息?下面的*信息处理不等式*从信息论角度回答了这个问题.

定理 11.5(数据处理不等式)　设 $X\to Y\to Z$ 是一个 Markov 链,则

$$H(X)\geqslant H(X:Y)\geqslant H(X:Z) \tag{11.37}$$

并且,第一个不等式达到饱和当且仅当在给定 Y 条件下可以重构 X.

直观上这个结果是合理的:它告诉我们,如果一个随机变量 X 受噪声影响产生 Y,那么我们采取的进一步行动(数据处理)不能用来增加过程的输出与原始信息 X 之间的互信息.

证　第一个不等式在定理 11.3 中已证.由定义可知,$H(X:Z)\leqslant H(X:Y)$ 等价于 $H(X|Y)\leqslant H(X|Z)$.根据 $X\to Y\to Z$ 是 Markov 链的事实易证(练习11.10) $Z\to Y\to X$ 亦为 Markov 链,于是 $H(X|Y)=H(X|Y,Z)$.于是问题归结为证明 $H(X,Y,Z)-H(Y,Z)=H(X|Y,Z)\leqslant H(X|Z)=H(X,Z)-H(Z)$,而这正是已证明的强次开加不等式.

设 $H(X:Y)<H(X)$,则不可能从 Y 重构出 X.因为若 Z 为仅仅根据 Y 的知识所进行的重构,则 $X\to Y\to Z$ 必为一个 Markov 链,于是根据数据处理不等式必有 $H(X)>H(X:Z)$,因此 $Z\neq X$.另一方面,若 $H(X:Y)=H(X)$,则我们有 $H(X|Y)=0$,并且只要 $p(X=x,Y=y)>0$ 就有 $p(X=x|Y=y)=1$.即,若 $Y=y$ 则必以概率 1 导出 X 等于 x,使我们能够重构出 X.　□

如上面提到的,如果 $X\to Y\to Z$ 是 Markov 链,那么 $Z\to Y\to X$ 也是 Markov 链,于是作为数据处理不等式的推论,我们得到如果 $X\to Y\to Z$ 是 Markov 链,那么

$$H(Z:Y)\geqslant H(Z:X) \tag{11.38}$$

我们把这个结果称为*数据管道不等式*.直观上,它断言 Z 与 X 共享的任何信息必然也为 Z 和 Y 所共享;这部分信息从 X 通过管道 Y 到达 Z.

练习 11.10　证明如果 $X\to Y\to Z$ 是 Markov 链,那么 $Z\to Y\to X$ 也是 Markov 链.

11.3　von Neumann 熵

Shannon 熵量测的不确定性与经典概率分布相关联.描述量子状态的方式是类似的,只是用密度算子代替了概率分布.本节我们把 Shannon 熵的定义推广到量子状态.

von Neumann 用下式定义量子状态 ρ 的熵:

$$S(\rho)\equiv-\operatorname{tr}(\rho\log\rho) \tag{11.39}$$

此式中对数是以 2 为底的.若 λ_x 是 ρ 的特征值,则 von Neumann 的定义可重写为

$$S(\rho)=-\sum_x\lambda_x\log\lambda_x \tag{11.40}$$

其中如对 Shannon 熵那样,定义 $0\log0\equiv0$.从计算角度看,后式更有用.例如在 d 维空间中的完全混合密度算子 I/d 具有熵 $\log d$.

从现在起，当我们提到熵，从上下文会明白我们指的是 Shannon 熵还是 von Neumann 熵.

练习 11.11(计算熵的例) 计算 $S(\rho)$，其中

$$\rho = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{11.41}$$

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{11.42}$$

$$\rho = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \tag{11.43}$$

练习 11.12(量子和经典熵的比较) 设 $\rho = p|0\rangle\langle 0| + (1-p)\frac{(|0\rangle+|1\rangle)(\langle 0|+\langle 1|)}{2}$，计算 $S(\rho)$，比较 $S(\rho)$ 和 $H(p,1-p)$ 的值.

盒子 11.2 熵的连续性

设想 ρ 改变一个小量，$S(\rho)$ 会变化多少？Fannes 不等式告诉我们变化不大，它甚至给出该变化的一个界.

定理 11.6(Fannes 不等式) 设 ρ 和 σ 是迹距离满足 $T(\rho,\sigma) \leqslant 1/\mathrm{e}$ 的密度矩阵，则

$$|S(\rho) - S(\sigma)| \leqslant T(\rho,\sigma)\log d + \eta(T(\rho,\sigma)) \tag{11.44}$$

其中 d 是 Hilbert 空间的维数，且 $\eta(x) \equiv -x\log x$. 去掉 $T(\rho,\sigma) \leqslant 1/\mathrm{e}$ 的限制，有较弱的不等式

$$|S(\rho) - S(\sigma)| \leqslant T(\rho,\sigma)\log d + \frac{1}{\mathrm{e}} \tag{11.45}$$

证 为证 Fannes 不等式，我们需要联系两算子迹距离与它们特征值的一个简单结果. 令 $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_d$ 为 ρ 的按递减序排列的特征值，而 $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_d$ 为 σ 的按递减序排列的特征值. 根据谱分解，我们有分解 $\rho - \sigma = Q - R$，其中 Q 和 R 是具有正交支集的半正定算子，于是 $T(\rho,\sigma) = \mathrm{tr}(R) + \mathrm{tr}(Q)$. 定义 $V \equiv R + \rho = Q + \sigma$，可知 $T(\rho,\sigma) = \mathrm{tr}(R) + \mathrm{tr}(Q) = \mathrm{tr}(2V) - \mathrm{tr}(R) - \mathrm{tr}(Q)$. 令 $t_1 \geqslant t_2 \geqslant \cdots \geqslant t_d$ 为 T 的特征值. 注意 $t_i \geqslant \max(r_i, s_i)$，于是 $2t_i \geqslant r_i + s_i + |r_i - s_i|$，从而

$$T(\rho,\sigma) \geqslant \sum_i |r_i - s_i| \tag{11.46}$$

根据计算只要 $|r-s| \leqslant 1/2$ 就有 $|\eta(r) - \eta(s)| \leqslant \eta(|r-s|)$. 易见 $|r_i - s_i| \leqslant 1/2$ 对所有 i 成立，故

$$\begin{aligned} |S(\rho) - S(\sigma)| &= \left|\sum_i (\eta(r_i) - \eta(s_i))\right| \\ &\leqslant \sum_i \eta(|r_i - s_i|) \end{aligned} \tag{11.47}$$

置 $\Delta \equiv \sum_i |r_i - s_i|$，并注意到 $\eta(|r_i - s_i|) = \Delta\eta(|r_i - s_i|/\Delta) - |r_i - s_i| \cdot \log(\Delta)$，可导出

$$|S(\rho) - S(\sigma)| \leqslant \Delta \sum \eta(|r_i - s_i|/\Delta) + \eta(\Delta) \leqslant \Delta \log d + \eta(\Delta) \tag{11.48}$$

其中第二个不等式用到定理 11.2. 但由式(11.46)，$\Delta \leqslant T(\rho,\sigma)$，故由 $\eta(\cdot)$ 在区间[0,1/e]上的单调性，

$$|S(\rho) - S(\sigma)| \leqslant T(\rho,\sigma)\log d + \eta(T(\rho,\sigma)) \tag{11.49}$$

其中 $T(\rho,\sigma) \leqslant 1/e$，即 Fannes 不等式. 对上述证明稍作修改可得 $T(\rho,\sigma)$ 一般情形下 Fannes 不等式的较弱形式. □

11.3.1　量子相对熵

就像经典的 Shannon 熵，定义量子版本的相对熵极其有用. 设 ρ 和 σ 是密度算子，ρ 到 σ 的相对熵定义为

$$S(\rho \| \sigma) \equiv \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) \tag{11.50}$$

正如经典的相对熵，量子相对熵有时会是无穷的. 特别地，如果 σ 的核(由 σ 属于 0 特征值的特征向量张成的空间)与 ρ 的支集(属于 ρ 非零特征值的特征向量张成的空间)有非平凡的交，则相对熵定义为 $+\infty$，否则定义为有限. 量子相对熵是非负的，该结果称为 Klein 不等式.

定理 11.7(Klein 不等式)　量子相对熵是非负的，即

$$S(\rho \| \sigma) \geqslant 0 \tag{11.51}$$

且当且仅当 $\rho = \sigma$ 时取等号.

证　令 $\rho = \sum_i p_i |i\rangle\langle i|$ 和 $\sigma = \sum_j q_j |j\rangle\langle j|$ 分别为 ρ 和 σ 的标准正交分解，根据相对熵的定义得

$$S(\rho \| \sigma) = \sum_i p_i \log p_i - \sum_i \langle i | \rho\log\sigma | i\rangle \tag{11.52}$$

此式代入等式 $\langle i|\rho = p_i\langle i|$ 和

$$\langle i | \log\sigma | i\rangle = \langle i | \left(\sum_j \log(q_j) |j\rangle\langle j|\right) | i\rangle = \sum_j \log(q_j) P_{ij} \tag{11.53}$$

其中 $P_{ij} \equiv \langle i|j\rangle\langle j|i\rangle \geqslant 0$，导出

$$S(\rho \| \sigma) = \sum_i p_i \left(\log p_i - \sum_j P_{ij} \log(q_j)\right) \tag{11.54}$$

注意 P_{ij} 满足 $P_{ij}\geqslant 0$，$\sum_i P_{ij}=1$ 且 $\sum_j P_{ij}=1$（按矩阵论的语言，P_{ij} 的这条性质称为双随机性）. 由于 $\log(\cdot)$ 是严格凹函数，故 $\sum_j P_{ij}\log q_j \leqslant \log r_i$，其中 $r_i \equiv \sum_j P_{ij}q_j$，且当且仅当存在某个 j 使 $P_{ij}=1$ 时取等号. 于是

$$S(\rho\|\sigma)\geqslant \sum_i p_i \log\frac{p_i}{r_i} \tag{11.55}$$

且当且仅当对每个 i 都存在一个 j 使 $P_{ij}=1$ 时取等号，即当且仅当 P_{ij} 是置换阵. 这里具有经典相对熵的形式. 由经典相对熵的非负性，定理 11.1，可推断

$$S(\rho\|\sigma)\geqslant 0 \tag{11.56}$$

且当且仅当对所有 i，$p_i=r_i$，并且 P_{ij} 是置换阵. 为进一步简化等式条件，注意通过必要时对 σ 的本征态重新编号，我们总可以假定 P_{ij} 是单位阵，因此 ρ 和 σ 在同一基底下是对角形的. 条件 $p_i=r_i$ 意味着 ρ 和 σ 相应的特征值相同，于是 $\rho=\sigma$ 就是取等号的条件. □

11.3.2 熵的基本性质

von Neumann 熵有许多有趣和有用的性质：

定理 11.8（von Neumann 熵的基本性质）

(1) 熵是非负的，当且仅当状态为纯态，熵为零.

(2) 在 d 维 Hilbert 空间中熵最多为 $\log d$，当且仅当系统处于完全混合态 I/d 熵等于 $\log d$.

(3) 设复合系统 AB 处于纯态，则 $S(A)=S(B)$.

(4) 设 p_i 是概率，而状态 ρ_i 在正交子空间上支集，则

$$S\left(\sum_i p_i\rho_i\right)=H(p_i)+\sum_i p_i S(\rho_i) \tag{11.57}$$

(5) **联合熵定理** 设 p_i 是概率，$|i\rangle$ 是子系统 A 的正交状态，ρ_i 是另一系统 B 的任意一组密度算子，则

$$S\left(\sum_i p_i\,|i\rangle\langle i|\otimes\rho_i\right)=H(p_i)+\sum_i p_i S(\rho_i) \tag{11.58}$$

证 (1) 可从定义直接得到.

(2) 结论由下面的相对熵的非负性得到，

$$0\leqslant S(\rho\|I/d)=-S(\rho)+\log d$$

(3) 由 Schmidt 分解可知，系统 A 和 B 的密度算子的特征值相同（回忆定理 2.7 后面的讨论），而熵是完全由特征值决定的，故 $S(A)=S(B)$.

(4) 令 λ_i^j 和 $|e_i^j\rangle$ 是相应于 ρ_i 的特征值和特征向量. 注意到 $p_i\lambda_i^j$ 和 $|e_i^j\rangle$ 是 $\sum_i p_i\rho_i$ 的特征值和特征向量，因此

$$S\left(\sum_i p_i\rho_i\right)=-\sum_{ij}p_i\lambda_i^j\log p_i\lambda_i^j \tag{11.59}$$

$$=-\sum_i p_i\log p_i-\sum_i p_i\sum_j\lambda_i^j\log\lambda_i^j \tag{11.60}$$

$$=H(p_i)+\sum_i p_iS(\rho_i) \tag{11.61}$$

(5) 从前面结果直接可以得到. □

练习 11.13(张量积的熵)　利用联合熵定理证明 $S(\rho\otimes\sigma)=S(\rho)+S(\sigma)$,从熵的定义出发直接证明该结果.

类似于 Shannon 熵,可以对复合量子系统定义量子联合熵和量子条件熵,以及量子互信息. 对由两部分 A 和 B 组成的复合系统联合熵直接定义为 $S(A,B)\equiv-\mathrm{tr}(\rho^{AB}\log\rho^{AB})$,其中 ρ^{AB} 是系统 AB 的密度矩阵. 条件熵和互信息定义为

$$S(A\mid B)\equiv S(A,B)-S(B) \tag{11.62}$$

$$S(A:B)\equiv S(A)+S(B)-S(A,B) \tag{11.63}$$

$$=S(A)-S(A\mid B)$$

$$=S(B)-S(B\mid A) \tag{11.64}$$

Shannon 熵的某些性质对 von Neumann 熵不成立,这对量子信息有许多重要影响. 例如,对随机变量 X 和 Y,不等式 $H(X)\leqslant H(X,Y)$ 成立. 直观上很明显:我们关于 X 和 Y 联合状态的不确定性肯定不超过对 X 状态的不确定性. 而这个观念对量子状态是不对的. 考虑处于纠缠态 $(|00\rangle+|11\rangle)/\sqrt{2}$ 的双量子比特系统 AB. 这是一个纯态,因此 $S(A,B)=0$. 另一方面,系统 A 具有密度矩阵 $I/2$,因此熵为 1. 这个结果的另一种表述方式为,对该系统,量 $S(B|A)=S(A,B)-S(A)$ 是负的.

练习 11.14(纠缠与负的条件熵)　设 $|AB\rangle$ 为属于 Alice 和 Bob 的复合系统的一个纯态,证明当且仅当 $S(B|A)<0$,$|AB\rangle$ 是纠缠的.

11.3.3　测量与熵

当我们进行测量时,量子系统的熵会如何变化? 不出意料,这个问题的答案依赖于测量的类型. 不过,关于熵的行为,我们还是有一些惊人的普遍结论.

例如假设在量子系统上进行由投影 P_i 描述的投影测量,但我们完全不知道测量的结果. 若测量前系统状态为 ρ,则测量后状态由

$$\rho'=\sum_i P_i\rho P_i \tag{11.65}$$

给出. 下面的结果表明在这个过程中熵不会减小,而且只有在状态没有被测量改变情况下才保持不变.

定理 11.9(投影测量增加熵) 设 P_i 是一组完备正交投影算子而 ρ 是一个密度算子,则系统测量后的状态 $\rho' \equiv \sum_i P_i \rho P_i$ 所具有的熵至少与原来的熵一样大,

$$S(\rho') \geqslant S(\rho) \tag{11.66}$$

且当且仅当 $\rho=\rho'$ 时取等号.

证 证明过程是将 Klein 不等式应用到 ρ 和 ρ' 上,

$$0 \leqslant S(\rho' \parallel \rho) = -S(\rho) - \mathrm{tr}(\rho \log \rho') \tag{11.67}$$

如果我们能证明 $-\mathrm{tr}(\rho\log\rho')=S(\rho')$,则结果就得证. 为此,我们应用完备性关系 $\sum_i P_i = I$,关系 $P_i^2=P_i$ 和迹的循环性质,得到

$$-\mathrm{tr}(\rho\log\rho') = -\mathrm{tr}\left(\sum_i P_i \rho \log \rho'\right) \tag{11.68}$$

$$= -\mathrm{tr}\left(\sum_i P_i \rho \log \rho' P_i\right) \tag{11.69}$$

注意 $\rho' P_i = P_i \rho P_i = P_i \rho'$,即 P_i 与 ρ' 可对易从而与 $\log\rho'$ 可对易,故

$$-\mathrm{tr}(\rho\log\rho') = -\mathrm{tr}\left(\sum_i P_i \rho P_i \log \rho'\right) \tag{11.70}$$

$$= -\mathrm{tr}(\rho' \log \rho') = S(\rho') \tag{11.71}$$

证明完成. □

练习 11.15(广义测量可以减小熵) 用测量算子 $M_1=|0\rangle\langle 0|$ 和 $M_2=|0\rangle\langle 1|$ 对处在状态 ρ 的量子比特进行测量. 如果我们不知道测量结果,则随后系统的状态为 $M_1\rho M_1^\dagger + M_2 \rho M_2^\dagger$. 证明该过程能够减少该量子比特的熵.

11.3.4 次可加性

设不同的量子系统 A 和 B 具有联合状态 ρ^{AB},则两个系统的联合熵满足不等式

$$S(A,B) \leqslant S(A) + S(B) \tag{11.72}$$

$$S(A,B) \geqslant |\, S(A) - S(B)\,| \tag{11.73}$$

第一个不等式称为 von Neumann 熵的**次可加性**不等式. 等式成立的充要条件是 A 和 B 是非相关的,即 $\rho^{AB}=\rho^A \otimes \rho^B$. 第二个不等式称为三角不等式,有时称为 Araki-Lieb 不等式; 它是 Shannon 熵 $H(X,Y) \geqslant H(X)$ 的量子类似.

次可加性的证明是 Klein 不等式 $S(\rho) \leqslant -\mathrm{tr}(\rho\log\sigma)$ 的简单应用. 令 $\rho \equiv \rho^{AB}$ 和 $\sigma \equiv \rho^A \otimes \rho^B$,注意

$$-\mathrm{tr}(\rho\log\sigma) = -\mathrm{tr}(\rho^{AB}(\log\rho^A + \log\rho^B)) \tag{11.74}$$

$$= -\mathrm{tr}(\rho^A \log \rho^A) - \mathrm{tr}(\rho^B \log \rho^B) \tag{11.75}$$

$$= S(A) + S(B) \tag{11.76}$$

于是 Klein 不等式蕴含 $S(A,B) \leqslant S(A)+S(B)$,如所欲证. Klein 不等式取等号的

条件 $\rho=\sigma$ 导出次可加性的等式条件 $\rho^{AB}=\rho^{A}\otimes\rho^{B}$.

为证三角不等式,引入一个如 2.5 节(见《量子计算和量子信息(一)》)的纯化系统 A 和 B 的系统 R,应用次可加性有

$$S(R)+S(A)\geqslant S(A,R) \tag{11.77}$$

因为 ABR 处于纯态,$S(A,R)=S(B)$,且 $S(R)=S(A,B)$. 上面的不等式可以整理得

$$S(A,B)\geqslant S(B)-S(A) \tag{11.78}$$

该不等式取等号的条件并不像次可加性的等式条件那么容易导出. 形式上,取等号的条件是 $\rho^{AR}=\rho^{A}\otimes\rho^{R}$. 直观上,这意味着在给定现存的与系统 B 的关联条件下,A 已经尽可能与外界发生纠缠. 练习 11.16 中给出该等式条件更详细的数学描述.

由系统 A 和 B 的对称性,我们还有 $S(A,B)\geqslant S(A)-S(B)$. 与 $S(A,B)\geqslant S(B)-S(A)$ 结合就导出三角不等式.

练习 11.16($S(A,B)\geqslant S(B)-S(A)$ 取等号的条件)　令 $\rho^{AB}=\sum_i\lambda_i\,|\,i\rangle\langle i\,|$ 为 ρ^{AB} 的谱分解,证明当且仅当算子 $\rho_i^{A}\equiv\mathrm{tr}_B(|i\rangle\langle i|)$ 具有共同的特征基底,而 $\rho_i^{B}\equiv\mathrm{tr}_A(|i\rangle\langle i|)$ 具有正交支集时,$S(A,B)=S(B)-S(A)$.

练习 11.17　对 AB 求一个非平凡混合态 ρ,使得 $S(A,B)=S(B)-S(A)$.

11.3.5　熵的凹性

熵是输入的凹函数. 即给定一组概率 p_i——满足 $\sum_i p_i=1$ 的非负实数——和相应的密度算子 ρ_i,熵满足不等式

$$S\Big(\sum_i p_i\rho_i\Big)\geqslant\sum_i p_iS(\rho_i) \tag{11.79}$$

直观上 $\sum_i p_i\rho_i$ 表示量子系统的状态以概率 p_i 处于未知状态 ρ_i,我们对状态的这种方式混合的不确定性应该高于状态 ρ_i 的平均不确定性,因为状态 $\sum_i p_i\rho_i$ 表达的不仅是对状态 ρ_i 的未知,而且还包含对指标 i 的未知.

设 ρ_i 是系统 A 的一组状态. 引入辅助系统 B,其状态空间具有对应密度算子 ρ_i 的标准正交基 $|i\rangle$,定义一个联合状态 ρ^{AB} 如下:

$$\rho^{AB}\equiv\sum_i p_i\rho_i\otimes|\,i\rangle\langle i\,| \tag{11.80}$$

为证凹性,我们利用熵的次可加性. 注意对密度矩阵 ρ^{AB} 有

$$S(A)=S\Big(\sum_i p_i\rho_i\Big) \tag{11.81}$$

$$S(B)=S\Big(\sum_i p_i\,|\,i\rangle\langle i\,|\Big)=H(p_i) \tag{11.82}$$

$$S(A,B) = H(p_i) + \sum_i p_i S(\rho_i) \tag{11.83}$$

应用次可加性 $S(A,B) \leqslant S(A)+S(B)$ 得到

$$\sum_i p_i S(\rho_i) \leqslant S\left(\sum_i p_i \rho_i\right) \tag{11.84}$$

即为凹性. 注意等号成立当且仅当所有 $p_i>0$ 的状态 ρ_i 都相同；即熵是输入的严格凹函数.

值得总结一下我们用来证明凹性和类似的用于证明三角不等式的技巧. 为证明系统 A 的结果，我们引入了一个**辅助系统** B. 引入辅助系统是量子信息论中反复应用的技巧. 这里引入 B 的原因是我们希望找到一个系统，其部分状态为 $\sum_i p_i \rho_i$ 而 i 的值为未知. 系统 B 实际上存储着 i 的真值；若 A“真地”处于状态 ρ_i，则系统 B 将处于状态 $|i\rangle\langle i|$，并且在 $|i\rangle$ 基底下观测系统 B 将揭示这一事实. 利用辅助系统将我们的直观想法严格化是一种艺术，但它确实构成量子信息论许多证明的实质部分.

练习 11.18 证明当且仅当所有 ρ_i 相同凹性不等式(11.79)取等号.

练习 11.19 证明存在一组酉矩阵 U_j 和概率 p_j，使得对任意矩阵 A 有

$$\sum_i p_i U_i A U_i^\dagger = \mathrm{tr}(A)\frac{I}{d} \tag{11.85}$$

其中 d 是 A 所在的 Hilbert 空间的维数. 据此和熵的严格凹性给出完全混合态 I/d，是 d 维空间中惟一具有最大熵状态的新证明.

练习 11.20 令 P 为一个投影且 $Q=I-P$ 为其补投影，证明存在酉算子 U_1 和 U_2 以及概率 p，使得对所有 ρ，有 $P\rho P+Q\rho Q=pU_1\rho U_1^\dagger+(1-\rho)U_2\rho U_2^\dagger$. 据此给出定理 11.9 基于凹性的新证明.

练习 11.21(Shannon 熵的凹性) 利用 von Neumann 熵的凹性推导 Shannon 熵对概率分布为凹的.

练习 11.22(凹性的新证明) 定义 $f(p)\equiv S(p\rho+(1-p)\sigma)$，论证为证凹性只需证 $f''(p)\leqslant 0$. 先对 ρ 和 σ 可逆的情形证明 $f''(p)\leqslant 0$，再对其他情况证明.

11.3.6 混合量子状态的熵

与凹性相对的一面是如下提供混合量子状态熵的上界估计的有用定理. 两个结果一起构成了量子状态 ρ_i 的混合 $\sum_i p_i\rho_i$ 的如下不等式：

$$\sum_i p_i S(\rho_i) \leqslant S\left(\sum_i p_i\rho_i\right) \leqslant \sum_i p_i S(\rho_i) + H(p_i) \tag{11.86}$$

其中右边上界的含义是，状态 $\sum_i p_i\rho_i$ 的不确定性总不超过状态 ρ_i 的平均不确定性加上 $H(p_i)$ 的附加贡献. $H(p_i)$ 代表指标 i 的不确定性对总的不确定性的最大

贡献. 下面就来证明该上界.

定理 11.10　设 $\rho=\sum_i p_i\rho_i$，其中 p_i 是一组概率，ρ_i 是密度算子，则

$$S(\rho)\leqslant\sum_i p_iS(\rho_i)+H(p_i) \tag{11.87}$$

且当且仅当状态 ρ_i 具有正交子空间上的支集时取等号.

证　我们从纯态 $\rho_i=|\psi_i\rangle\langle\psi_i|$ 的情况出发. 设 ρ_i 是系统 A 的状态，引入具有对应概率 p_i 指标 i 的标准正交基底 $|i\rangle$ 的辅助系统 B，定义

$$|AB\rangle\equiv\sum_i\sqrt{p_i}\,|\psi_i\rangle|i\rangle \tag{11.88}$$

因为 $|AB\rangle$ 是纯态，故

$$S(B)=S(A)=S\Big(\sum_i p_i|\psi_i\rangle\langle\psi_i|\Big)=S(\rho) \tag{11.89}$$

在系统 B 的 $|i\rangle$ 基底上进行投影测量，测量后系统 B 的状态为

$$\rho^{B'}=\sum_i p_i|i\rangle\langle i| \tag{11.90}$$

但由定理 11.9 投影测量不减少熵，于是 $S(\rho)=S(B)\leqslant S(B')=H(p_i)$. 注意对纯态情形 $S(\rho_i)=0$，我们就证明了

$$S(\rho)\leqslant H(p_i)+\sum_i p_iS(\rho_i) \tag{11.91}$$

其中状态 ρ_i 为纯态. 进而当且仅当 $B=B'$ 时等号成立，而易见这等价于状态 $|\psi_i\rangle$ 正交.

现在不难证明混合态的情形. 令 $\rho_i=\sum_j p_j^i|e_j^i\rangle\langle e_j^i|$ 为状态 ρ_i 的标准正交分解，于是 $\rho=\sum_{ij}p_ip_j^i|e_j^i\rangle\langle e_j^i|$. 运用纯态的结果和对每个 i 成立的 $\sum_j p_j^i=1$，得到

$$S(\rho)\leqslant-\sum_{ij}p_ip_j^i\log(p_ip_j^i) \tag{11.92}$$

$$=-\sum_i p_i\log p_i-\sum_i p_i\sum_j p_j^i\log p_j^i \tag{11.93}$$

$$=H(p_i)+\sum_i p_iS(\rho_i) \tag{11.94}$$

即为所欲证. 对混合态取等号的条件可从纯态取等号的条件直接得到. □

11.4　强次可加性

对两个量子系统的次可加性和三角不等式可以推广到三量子系统. 基本结果称为**强次可加性**不等式，这是量子信息论最重要和有用的结论之一. 不等式针对三量子系统 A,B,C，有

$$S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C) \tag{11.95}$$

不幸的是，不像经典情形，量子强次可加性不等式的所有已知证明都很难. 不过由于在量子信息论中非常有用，我们还是要给出该结果的完整证明. 11.4.1 节描述证明的基本结构，部分细节留到附录.

11.4.1 强次可加性的证明

我们要给出的强次可加性不等式的证明基于数学上一个深刻的结果，称为 Lieb 定理. 为应用 Lieb 定理，我们需要引入一个定义.

设 $f(A,B)$是两个矩阵 A,B 的一个实值函数，如果对所有 $0\leqslant\lambda\leqslant1$，有

$$f(\lambda A_1+(1-\lambda)A_2,\lambda B_1+(1-\lambda)B_2)\geqslant\lambda f(A_1,B_1)+(1-\lambda)f(A_2,B_2) \tag{11.96}$$

则 f 称为关于 A 和 B 联合凹.

练习 11.23(联合凹蕴含对每个输入凹) 令 $f(A,B)$为一个联合凹函数，证明对给定的 B，$f(A,B)$关于 A 是凹的. 给出一个两变量函数，对每个输入为凹但不是联合凹的.

定理 11.11(Lieb 定理) 令 X 为一矩阵，且 $0\leqslant t\leqslant1$，则函数

$$f(A,B)\equiv\mathrm{tr}(X^\dagger A^t X B^{1-t}) \tag{11.97}$$

关于半正定矩阵 A 和 B 是联合凹的.

证 见附录 Lieb 定理的证明. □

Lieb 定理可以推出一系列有独立价值的结果，最终可导出强次可加性的证明. 我们从相对熵的凸性开始.

定理 11.12(相对熵的凸性) 相对熵 $S(\rho\|\sigma)$对其自变量是联合凸的.

证 对任意作用在同一空间上的矩阵 A 和 X 定义

$$I_t(A,X)\equiv\mathrm{tr}(X^\dagger A^t X A^{1-t})-\mathrm{tr}(X^\dagger X A) \tag{11.98}$$

据 Lieb 定理，表达式中第一项对 A 是凹的，而第二项对 A 是线性的，于是 $I_t(A,X)$对 A 是凹的. 定义

$$I(A,X)\equiv\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}I_t(A,X)=\mathrm{tr}(X^\dagger(\log A)XA)-\mathrm{tr}(X^\dagger X(\log A)A) \tag{11.99}$$

注意到 $I_0(A,X)=0$，利用 $I_t(A,X)$对 A 的凸性可得

$$I(\lambda A_1+(1-\lambda)A_2,X)=\lim_{\Delta\to0}\frac{I_\Delta(\lambda A_1+(1-\lambda)A_2,X)}{\Delta} \tag{11.100}$$

$$\geqslant\lambda\lim_{\Delta\to0}\frac{I_\Delta(A_1,X)}{\Delta}+(1-\lambda)\lim_{\Delta\to0}\frac{I_\Delta(A_2,X)}{\Delta} \tag{11.101}$$

$$=\lambda I(A_1,X)+(1-\lambda)I(A_2,X) \tag{11.102}$$

即 $I(A,X)$是 A 的凹函数. 定义分块矩阵

$$A \equiv \begin{bmatrix} \rho & 0 \\ 0 & \sigma \end{bmatrix}, \quad X \equiv \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix} \tag{11.103}$$

易验证 $I(A,X)=-S(\rho\|\sigma)$. $S(\rho\|\sigma)$的联合凸性可由 $I(A,X)$关于 A 的凹性得到. □

推论 11.13(量子条件熵的凹性)　令 AB 为由 A 和 B 组成的复合量子系统，则对 AB 的状态 ρ^{AB} 条件熵 $S(A|B)$是凹的.

证　令 d 为系统 A 的维数，注意到

$$S\left(\rho^{AB}\left\|\frac{I}{d}\otimes\rho^{B}\right.\right)=-S(A,B)-\mathrm{tr}\left(\rho^{AB}\log\left(\frac{I}{d}\otimes\rho^{B}\right)\right) \tag{11.104}$$

$$=-S(A,B)-\mathrm{tr}(\rho^{B}\log\rho^{B})+\log d \tag{11.105}$$

$$=-S(A\mid B)+\log d \tag{11.106}$$

因此 $S(A|B)=\log d-S(\rho^{AB}\|I/d\otimes\rho^{B})$. $S(A|B)$的凹性可从相对熵的联合凸性导出. □

定理 11.14(强次可加性)　对任意的三量子系统 A,B,C，成立不等式

$$S(A)+S(B)\leqslant S(A,C)+S(B,C) \tag{11.107}$$

$$S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C) \tag{11.108}$$

证　这两个不等式实际上是等价的. 我们将用条件熵的凹性来证明第一个不等式，再证明第二个也成立. 定义 $T(\rho^{ABC})$为系统 ABC 的密度算子的函数，

$$T(\rho^{ABC})\equiv S(A)+S(B)-S(A,C)-S(B,C)$$

$$=-S(C\mid A)-S(C\mid B) \tag{11.109}$$

由条件熵的凹性可知 $T(\rho^{ABC})$是 ρ^{ABC} 的凸函数. 令 $\rho^{ABC}=\sum_i p_i\mid i\rangle\langle i\mid$ 为 ρ^{ABC} 的谱分解，由 T 的凹性，$T(\rho^{ABC})\leqslant\sum_i p_i T(\mid i\rangle\langle i\mid)$. 但 $T(|i\rangle\langle i|)=0$ 因为对一个纯态 $S(A,C)=S(B)$且 $S(B,C)=S(A)$，于是 $T(\rho^{ABC})\leqslant 0$，故

$$S(A)+S(B)-S(A,C)-S(B,C)\leqslant 0 \tag{11.110}$$

这正是我们要证明的第一个不等式.

为证明第二个不等式，引入一个纯化系统 ABC 的辅助系统 R，利用刚刚证明的不等式得

$$S(R)+S(B)\leqslant S(R,C)+S(B,C) \tag{11.111}$$

因为 $ABCR$ 是纯态，$S(R)=S(A,B,C)$且 $S(R,C)=S(A,B)$，故式(11.111)变为

$$S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C) \tag{11.112}$$

如所欲证. □

练习 11.24　通过不等式 $S(A)+S(B)\leqslant S(A,C)+S(B,C)$，我们证明了强次可加性，该不等式的证明也可由强次可加性推出.

练习 11.25　我们从条件熵 $S(A|B)$的凹性导出了强次可加性，条件熵的凹性的证明可由强次可加性推出(提示：可以引入辅助系统).

11.4.2 强次可加性：基本应用

强次可加性和相关结果在量子信息论中有许多有价值的应用，下面来看几个基本结论.

首先，值得强调不等式 $S(A)+S(B)\leqslant S(A,C)+S(B,C)$ 的成立是多么重要. 相应的不等式对 Shannon 熵也成立，但原因不同. 对 Shannon 熵，$H(A)\leqslant H(A,C)$ 和 $H(B)\leqslant H(B,C)$ 成立，故两不等式的和一定也成立. 在量子情形，可能出现 $S(A)>S(A,C)$ 或 $S(B)>S(B,C)$ 的情况，然而自然的设计却为保证条件 $S(A)+S(B)\leqslant S(A,C)+S(B,C)$ 始终成立而使两个不等式不可能同时为真. 也可以用条件熵和互信息的语言来描述这个事实，

$$0\leqslant S(C\mid A)+S(C\mid B) \tag{11.113}$$

$$S(A:B)+S(A:C)\leqslant 2S(A) \tag{11.114}$$

出于类似的原因，它们也是两个重要不等式. 不过注意，不等式 $0\leqslant S(A\mid C)+S(B\mid C)$ 的真实性并不像人们或许想像的可基于式(11.114)得到，这可由选 ABC 为 A 的纯态和 BC 的 EPR 态的积看出.

练习 11.26 证明 $S(A:B)+S(A:C)\leqslant 2S(A)$. 注意相应的不等式对 Shannon 熵成立因为 $H(A:B)\leqslant H(A)$. 给出一个 $S(A:B)>S(A)$ 的例子.

在实践中，用条件或互信息语言叙述的强次可加性常常更容易应用. 下面的定理列出强次可加性的三个非常简单的形式，为量子熵属性提供了直观的向导.

定理 11.15 (1) **条件减少熵** 设 ABC 是复合量子系统，则 $S(A\mid B,C)\leqslant S(A\mid B)$.

(2) **丢弃量子系统从不增加互信息** 设 ABC 是复合量子系统，则

$$S(A:B)\leqslant S(A:B,C)$$

(3) **量子运算从不增加互信息** 设 AB 是复合量子系统而 ε 是系统 B 上保持迹的量子运算，令 $S(A:B)$ 表示在 ε 作用到 B 之前系统 A 和 B 之间的互信息，而 $S(A':B')$ 为 ε 作用到 B 之后的互信息，则 $S(A':B')\leqslant S(A:B)$.

证 (1) 证明与经典情形的证明相同(定理 11.3 的部分)，为方便起见，重述如下：$S(A\mid B,C)\leqslant S(A\mid B)$ 等价于 $S(A,B,C)-S(B,C)\leqslant S(A,B)-S(B)$，而这又等价于 $S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C)$，即强次可加性.

(2) $S(A:B)\leqslant S(A:B,C)$ 等价于 $S(A)+S(B)-S(A,B)\leqslant S(A)+S(B,C)-S(A,B,C)$，也等价于 $S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C)$，即强次可加性.

(3) 由第 8 章的构造，ε 在 B 上的作用可由引入的第三个系统 C，设其初态为 $|0\rangle$，以及引入 B 和 C 之间的一个酉相互作用来模拟. ε 在 B 上的作用等价于 U 的作用之后丢弃系统 C. 用“′”符号表示 U 作用之后系统的状态，得到 $S(A:B)=S(A:B,C)$. 因为 C 开始时 C 在与 AB 的积状态中，而且显然 $S(A:B,C)=$

$S(A'\colon B',C')$，丢弃系统不会增加互信息，故 $S(A'\colon B')\leqslant S(A'\colon B',C')$. 总结起来得到所需的 $S(A'\colon B')\leqslant S(A\colon B)$. □

关于量子条件熵的强次可加性有一系列有趣的问题. 前面我们看到 Shannon 互信息是非次可加的，因此量子互信息也是非次可加的. 条件熵的次可加性如何呢？即

$$S(A_1,A_2\mid B_1,B_2)\leqslant S(A_1\mid B_1)+S(A_2\mid B_2) \tag{11.115}$$

是否对四个量子系统 A_1,A_2,B_1 和 B_2 成立？事实上，该不等式是成立的，而且，条件熵对第一和第二项还是次可加的. 这些事实的证明是应用强次可加性的很好练习.

定理 11.16（条件熵的次可加性）　令 $ABCD$ 是四量子系统的复合，则条件熵对第一和第二项是联合次可加的：

$$S(A,B\mid C,D)\leqslant S(A\mid C)+S(B\mid D) \tag{11.116}$$

令 ABC 是三量子系统的复合，则条件熵对第一和第二项都是次可加的：

$$S(A,B\mid C)\leqslant S(A\mid C)+S(B\mid C) \tag{11.117}$$

$$S(A\mid B,C)\leqslant S(A\mid B)+S(A\mid C) \tag{11.118}$$

证　为证对两项的联合次可加性，注意由强次可加性，得

$$S(A,B,C,D)+S(C)\leqslant S(A,C)+S(B,C,D) \tag{11.119}$$

不等式两边同时加上 $S(D)$，得

$$S(A,B,C,D)+S(C)+S(D)\leqslant S(A,C)+S(B,C,D)+S(D) \tag{11.120}$$

对右边最后两项应用强次可加性得到

$$S(A,B,C,D)+S(C)+S(D)\leqslant S(A,C)+S(B,D)+S(C,D) \tag{11.121}$$

整理此式得到

$$S(A,B\mid C,D)\leqslant S(A\mid C)+S(B\mid D) \tag{11.122}$$

这即为条件熵的联合次可加性.

条件熵对第一项的次可加性，$S(A,B|C)\leqslant S(A|C)+S(B|C)$ 显然等价于强次可加性. 对第二项的次可加性的证明更难些，要证的是 $S(A|B,C)\leqslant S(A|B)+S(A|C)$. 注意这等价于不等式

$$S(A,B,C)+S(B)+S(C)\leqslant S(A,B)+S(B,C)+S(A,C) \tag{11.123}$$

为此，注意不等式 $S(C)\leqslant S(A,C)$ 和 $S(B)\leqslant S(A,B)$ 中至少有一个为真，因为由定理 11.14 知，$S(A|B)+S(A|C)\geqslant 0$. 设 $S(C)\leqslant S(A,C)$，将此不等式加到强次可加不等式 $S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C)$ 上，就得到所需结果. $S(B)\leqslant S(A,B)$ 的证明类似. □

当引入相对熵时，我们看到它非常像概率分布或密度算子之间的一个距离测度. 设量子系统由标记为 A 和 B 的两部分组成且有两个密度算子 ρ^{AB} 和 σ^{AB}. 为满足期望的类似距离的性质，关于 $S(\cdot\|\cdot)$ 有一条非常期望的性质，即当忽略系统的一部分时，它应该减小，即

$$S(\rho^A \| \sigma^A) \leqslant S(\rho^{AB} \| \sigma^{AB}) \tag{11.124}$$

该结果称为相对熵的单调性. 直观上这是一个距离测度具有的一条非常合理的性质；我们预期忽略物理系统的一部分会使系统的两个状态更难区分(对照 9.2.1 节)，因此使它们之间任何合理的距离测度都减小.

定理 11.17(相对熵的单调性) 令 ρ^{AB} 和 σ^{AB} 为复合系统 AB 的任意两个密度矩阵，则

$$S(\rho^A \| \sigma^A) \leqslant S(\rho^{AB} \| \sigma^{AB}) \tag{11.125}$$

证 练习 11.19 蕴含着存在 B 空间上的酉变换 U_j 和概率 p_j，使得对所有 ρ^{AB} 有

$$\rho^A \otimes \frac{I}{d} = \sum_j p_j U_j \rho^{AB} U_j^\dagger \tag{11.126}$$

由相对熵的凸性可得

$$S\left(\rho^A \otimes \frac{I}{d} \middle\| \sigma^A \otimes \frac{I}{d}\right) \leqslant \sum_j p_j S(U_j \rho^{AB} U_j^\dagger \| U_j \sigma^{AB} U_j^\dagger) \tag{11.127}$$

但相对熵在酉变换下是不变的，于是可得

$$S\left(\rho^A \otimes \frac{I}{d} \middle\| \sigma^A \otimes \frac{I}{d}\right) \leqslant \sum_j p_j S(\rho^{AB} \| \sigma^{AB}) = S(\rho^{AB} \| \sigma^{AB}) \tag{11.128}$$

此式结合下面容易验证的事实

$$S\left(\rho^A \otimes \frac{I}{d} \middle\| \sigma^A \otimes \frac{I}{d}\right) = S(\rho^A \| \sigma^A) \tag{11.129}$$

就导出相对熵的单调性. □

问题 11.1(广义 Klein 不等式) 设 $f(\cdot)$ 为从实数到实数的凸函数，则如 2.1.8节(见《量子计算和量子信息(一)》)那样，f 诱导出一个自然的定义在 Hermite 算子上的函数 $f(\cdot)$，证明

$$\mathrm{tr}(f(A) - f(B)) \geqslant \mathrm{tr}((A-B)f'(B)) \tag{11.130}$$

利用该结果证明相对熵为非负.

问题 11.2(广义相对熵) 相对熵的定义可以推广到任意两个半正定算子 r 和 s 上，

$$S(r \| s) \equiv \mathrm{tr}(r \log r) - \mathrm{tr}(r \log s) \tag{11.131}$$

前面相对熵联合凸性的证明可直接用到如下推广的定义：

(1) 对 $\alpha, \beta > 0$，证明

$$S(\alpha r \| \beta s) = \alpha S(r \| s) + \alpha \mathrm{tr}(r) \log(\alpha/\beta) \tag{11.132}$$

(2) 证明相对熵的联合凸性蕴含相对熵的次可加性

$$S(r_1+r_2\|s_1+s_2)\leqslant S(r_1\|s_1)+S(r_2\|s_2) \tag{11.133}$$

(3) 证明相对熵的次可加性蕴含相对熵的联合凸性.

(4) 令 p_i 和 q_i 为同一指标集上的概率分布,证明

$$S\Big(\sum_i p_i r_i\Big\|\sum_i q_i s_i\Big)\leqslant\sum_i p_i S(r_i\|s_i)+\sum_i p_i\mathrm{tr}(r_i)\log(p_i/q_i) \tag{11.134}$$

当由于 r_i 是密度算子因此 $\mathrm{tr}(r_i)=1$ 时,该式导出漂亮的公式

$$S\Big(\sum_i p_i r_i\Big\|\sum_i q_i s_i\Big)\leqslant\sum_i p_i S(r_i\|s_i)+H(p_i\|q_i) \tag{11.135}$$

其中 $H(\cdot\|\cdot)$ 是 Shannon 相对熵.

问题 11.3(条件熵对应的三角不等式)　(1) 证明 $H(X,Y|Z)\geqslant H(X|Z)$.

(2) 证明 $S(A,B|C)\geqslant S(A|C)$ 并非总是成立.

(3) 证明条件版本的三角不等式

$$S(A,B\mid C)\geqslant S(A\mid C)-S(B\mid C) \tag{11.136}$$

问题 11.4(条件形式的强次可加性)　(1) 证明 $S(A,B,C|D)+S(B|D)\leqslant S(A,B|D)+S(B,C|D)$.

(2) 举例说明下式不总是成立:

$$H(D|A,B,C)+H(D|B)\leqslant H(D|A,B)+H(D|B,C)$$

问题 11.5(研究强次可加性)　给出量子熵强次可加性不等式的一个简单证明.

第 11 章的总结　熵与信息

- 信息的基本度量来源于对用来解决某些信息处理问题所需要的物理资源问题的解决.
- 基本定义:

(熵)　$S(A)=-\mathrm{tr}(\rho^A\log\rho^A)$

(相对熵)　$S(\rho\|\sigma)=-S(\rho)-\mathrm{tr}(\rho\log\sigma)$

(条件熵)　$S(A|B)=S(A,B)-S(B)$

(互信息)　$S(A:B)=S(A)+S(B)-S(A,B)$

- 强次可加性:$S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C)$. 另一个熵不等式是该不等式的推论,即相对熵的联合凸性.
- 相对熵对其输入是联合凸的.
- 相对熵的单调性:$S(\rho^A\|\sigma^A)\leqslant S(\rho^{AB}\|\sigma^{AB})$.

历史和进一步阅读的材料

历史上，熵的概念首先出现在热力学和统计力学的研究中. 但研究熵的现代信息论基础产生于 Shannon 关于信息论的著名论文中[Sha48]. Cover 和 Thomas 的著作[CT91]的第 2 章和第 16 章为 Shannon 熵性质提供了一个很好的一般性参考. von Neumann 熵的一般文献是 Wehrl 的综述文章[Weh78]和 Ohya，Petz 的书[OP93].

我们证明的熵测不准原理来自 Deutsch[Deu83]. 许多学者曾研究过熵不确定性关系，这里只提及另外两篇论文. Kraus[Kra87]对一类特殊的测量猜想了比 Deutsch 的更强的熵不确定性关系，而 Maassen 和 Uffink[MU88]证明了 Kraus 的猜想. 相对熵是 Kullback 和 Leibler[KL51]提出的，它的量子推广来源于 Umegaki[Ume62]. Fanne 不等式出现在文献[Fan73]中. Klein 不等式是在文献[Kle31]中证明的. 三角不等式来自 Araki 和 Lieb[AL70]. 强次可加性的历史很有趣. Robinson 和 Ruelle[RR67]首先注意到统计物理中经典强次可加性的重要性，随后 Robinson 和 Ruelle[RR68]在 1968 年提出量子版本的猜想. 不过该结果的证明相当困难. 最终在 1973 年该定理在两篇论文中被证明：在文献[Lie73]中以 Lieb 命名的定理，以及在 Lieb 和 Ruskai[LR73b]中建立的与强次可加性之间惊人的关系；也可参考文献[LR73a]. Lieb 定理是 1963 年由 Wigner 和 Yanase[WY63]及随后 Dynson(未发表)的推广提出的 Winger-Yanase-Dyson 猜想的扩展，而在 1973 年之前人们不知道 Winger-Yanase-Dyson 猜想和强次可加性有关系. 关于 Winger-Yanase-Dyson 猜想可以参考 Wehrl[Weh78]. 我们给出的 Lieb 定理的证明来自 Simon[Sim79]，是 Uhlmann[Uhl77]给出证明的变形. Lieb 定理还有一些其他的证明，例如见 Epstein[Eps73]，Ando[And79]和 Petz[Pet86]的相关工作. 相对熵对第一项和第二项的强次可加性是 Lieb[Lie75]证明的. 量子条件熵的联合次可加性是 Nielsen[Nie98]证明的. 相当熵的单调性是 Lindblad[Lin75]首先注意到的. 问题 11.2 来源于 Ruskai[Rus94].

CHAPTER 12

第12章 量子信息论

经典信息论主要关心通过服从经典物理学原理的信道传送经典信息，例如字母表中的字母、语音和比特串的问题. 如果我们建立量子力学信道，问题会发生什么变化？我们能够用量子力学来传送秘密信息而不被窃听吗？这只是允许使用量子力学信道时可能问的两个问题. 这里对信道的重新定义促使我们重新审视激发经典信息论的那些基本问题，以便探索新的答案. 本章综述量子信息论的知识，包括量子信道带来的某些令人吃惊和迷人的机会.

量子信息论产生于对信道的研究，但它的应用领域却宽得多，并且用简单的语言去概括这个领域的目标并非易事. 如 1.6 节（见《量子计算和量子信息（一）》）所描述的，我们可以对量子信息论的研究概括出三个基本目标：识别出量子力学中静态资源的基本类型（我们归类为不同类型的信息），识别出量子力学中动态过程的基本类型（我们归类为不同类型的信息处理），以及执行基本动态过程涉及的各类资源间转换的量化. 量子信息论从实质上比经典信息论丰富，因为量子力学包含了如此多类的基本静动态资源——不仅仅是它支持所有熟悉的经典类型，而且具有全新的新类型，如纠缠的静态资源，使该领域比经典世界有趣得多.

本章的标题是量子信息论. 读者也许会怀疑在一章中如何覆盖量子信息论的各个方面. 事实上，量子信息论包含着许多这里未描述的方面，如量子运算的研究、忠实度的测度定义和研究、量子纠错码以及熵的各种概念，所有这些内容我们都已在前面的章节中详细地介绍了. 本章的目的是用最纯粹的形式来描述量子信息论；其他章节则集中在研究具体的工具，而这里我们关心事情的全貌，对量子信息的性质做最一般的论述.

我们从 12.1 节用信息论语言关于量子状态与经典状态相比的某些独特性质的讨论开始. 量子状态不仅一般而言是无法复制的，而且它们还是无法彻底区分的. 这由 Holevo 界来定量刻画. 12.2 节考虑量子信息论的一项基本任务——*数据压缩*，并描述量子状态像经典状态那样如何被大大压缩. 这是通过将典型序列定理与典型子空间定理并列，以证明 Schumacher 量子无噪声信道编码定理来做到的，

这定理与经典的 Shannon 无噪声信道编码定理相对应. 此问题的一个自然推广是经典信息的带噪声信道容量问题,在 12.3 节,我们定义并证明对应于 Shannon 带噪声信道编码定理的称为 Holevo-Schumacher-Westmoreland 定理的结果. 对量子信息而言,最困难的挑战是带噪声量子信道的容量,这是 12.4 节的主题. 该节将定义熵交换、量子 Fano 不等式和量子数据处理不等式,但作为公开的容量问题未被解决. 该节给出带噪声信道关系的两个应用:单量子界和 Maxwell 妖驱除,并总结前半章的内容. 在量子信息的论述过程中反复提到的两个主题是纠缠和非正交性,它们是后边两小节的主要论题. 12.5 节描述如何将纠缠视为物理资源,并解释如何对其进行变换、蒸馏和稀释. 12.6 节描述量子密码术,它是从本章量子信息的许多性质产生的可证明为安全的通信方式.

12.1 区分量子状态和可访问的信息

为说明量子和经典信息之间巨大的差异,我们可以做一个小游戏. 我们用两个虚构的人物 Alice 和 Bob 来描述这个游戏,当然可以用更抽象的语言来叙述结果,但拟人化的语言使结果更容易想象(和书写).

设 Alice 有一个经典信源,能够按概率 $p_0,\cdots,p_n$ 产生符号 $X=0,\cdots,n$,游戏的目的是让 Bob 尽他所能地确定 X 值. 为此,Alice 制备一个从某个固定集合 $\rho_0,\cdots,\rho_n$ 中选出的量子态 ρ_X,并将该状态交给 Bob. Bob 要对得到的状态进行量子测量,然后根据他的测量结果 Y 给出 X 值的最好猜测.

Bob 从测量得到的关于 X 的信息的一个好的度量是,如第 11 章定义的 X 和测量输出 Y 之间的互信息 $H(X:Y)$. 据信息处理不等式可知,当且仅当 $H(X:Y)=H(X)$时,Bob 可以从 Y 推断出 X,而一般有 $H(X:Y)\leqslant H(X)$. 后面我们将会看到,$H(X:Y)$和 $H(X)$的接近程度实际上提供了 Bob 可以确定 X 的程度的一个量化测度. Bob 的目标是选择一个测量使 $H(X:Y)$最大化,使它尽可能接近$H(X)$. 为此,我们把 Bob 可访问的信息定义为取遍所有测量方案情况下互信息的最大值 $H(X:Y)$,可访问的信息是 Bob 能够在多大程度上推断出 Alice 制备状态的一种度量.

经典信息论中,可访问的信息没有太多价值. 尽管实践中难以区分两个经典状态——比如我们阅读写得不好的字迹时遇到的困难——但原则上,区分两个经典状态总是可能的. 与此对照,在量子力学中并不总是可以区分不同状态,即使是原则上也不行. 例如,我们在盒子 2.3 中指出不存在能够可靠区分两个非正交量子状态的量子力学过程. 用可访问信息的语言叙述,即如果 Alice 以概率 p 制备状态 $|\psi\rangle$,以概率 $1-p$ 制备非正交的状态 $|\varphi\rangle$,则该制备的可访问信息将严格小于 $H(p)$,因为 Bob 不可能完全有把握地确定状态的值. 在经典世界,若干 Alice 制备

两个经典状态之一——如以概率 p 取一个 0 状态的比特，或以概率 $1-p$ 取 1 状态——则区分这两个状态没有什么原则性困难，因而可访问信息与制备的熵 $H(p)$ 相同.

上述讨论有一点需要澄清，即在一种经典场合下可访问信息有其重要性，该场合是有关区分概率分布的. 设想 Alice 按两种概率分布中的一种（$p, 1-p$ 或 q, $1-q$）来制备状态 0 和 1. 给定状态后，Bob 的目标是识别 Alice 用哪种概率分布制备了该状态. 显然，Bob 不能总是可靠地进行这种识别. 然而，这个例子（类似一组混合态中的一个量子系统时的可访问信息）不是太重要. 最重要的是量子力学的基本对象——量子纯态——具有与经典信息论中类似 0 和 1 的基本对象显著不同的而远为丰富的可区分性.

*不可克隆定理*提供了与经典信息相比，量子信息缺乏可访问性的另一个侧面. 经典信息当然是可以复制的. 这可以用数字信息精确实现，像生成此书的多重备份的 $\mathrm{L^AT_EX}$ 文件，或近似地，像在销售之前印刷厂复制出的本书的每一页上的模拟图像. 令人吃惊的是，不可克隆定理断言量子力学不允许未知量子状态被精确复制，并给我们的近似复制定出了严格的限制. 不可克隆定理将在盒子 12.1 中证明.

不可克隆定理初看上去似乎相当离奇，难道经典物理学不是量子力学的特例吗？如果我们不能复制量子状态，那怎能复制经典信息呢？答案是不可克隆定理并未禁止所有状态的复制，它只是断言非正交量子状态不能复制. 更确切地，设 $|\psi\rangle$ 和 $|\varphi\rangle$ 为两个非正交的量子状态，那么不可克隆定理蕴含不可能建造一个量子设备，在输入 $|\psi\rangle$ 或 $|\varphi\rangle$ 时，会输出输入状态的两个备份 $|\psi\rangle|\psi\rangle$ 或 $|\varphi\rangle|\varphi\rangle$. 另一方面，如果 $|\psi\rangle$ 和 $|\varphi\rangle$ 是正交的，那么不可克隆定理并不禁止它们的克隆. 事实上，很容易设计复制这些状态的量子电路. 这就解释了不可克隆定理与经典信息复制的可能性之间似乎存在的矛盾，因为经典信息的不同状态可以被视为仅仅是正交的量子状态.

练习 12.1　设 $|\psi\rangle$ 和 $|\varphi\rangle$ 为一个单量子比特的两个正交量子状态，设计一个具有双量子比特输入（数据和目标量子比特）的量子线路. 数据量子比特要么处在 $|\psi\rangle$ 状态要么处在 $|\varphi\rangle$ 状态，目标量子比特制备为标准状态 $|0\rangle$. 根据输入数据是 $|\psi\rangle$ 还是 $|\varphi\rangle$，该线路输出 $|\psi\rangle|\psi\rangle$ 或 $|\varphi\rangle|\varphi\rangle$.

克隆和可访问信息之间有什么联系？设 Alice 以概率 p 和 $1-p$ 制备了两个非正交量子状态 $|\psi\rangle$ 和 $|\varphi\rangle$ 中的一个. 假设 Bob 关于这些状态的可访问信息是 $H(p)$，即量子力学原理允许 Bob 通过测量获得识别 Alice 制备了 $|\psi\rangle$ 和 $|\varphi\rangle$ 的哪个状态的足够信息，那么 Bob 将能够可以以非常简单的方式克隆这些状态：他可以进行一个测量以确定 Alice 制备的是 $|\psi\rangle$ 和 $|\varphi\rangle$ 中的哪个状态，一旦完成了识别，他可以任意制备 Alice 给他的不论是 $|\psi\rangle$ 还是 $|\varphi\rangle$ 的多重备份. 因此不可克隆定理可以看成是这些状态的可访问信息严格小于 $H(p)$ 的一个结论：反过来看也是可以

的，我们可以把可访问信息小于 $H(p)$ 看作是不可克隆定理的一个结论. 证明过程如下. 设想可以克隆非正交状态. 从 Alice 那里收到状态 $|\psi\rangle$ 或 $|\varphi\rangle$ 之后，Bob 可以反复使用该克隆装置，得到状态 $|\psi\rangle^{\otimes n}$ 或是 $|\varphi\rangle^{\otimes n}$. 在 n 非常大的极限情况，这两个状态将非常接近正交，从而可以通过投影测量以任意高的可靠性区分它们. 即如果克隆是可能的，那么 Bob 将能以任意高的成功概率识别制备的是 $|\psi\rangle$ 或是 $|\varphi\rangle$，因而可访问信息将是 $H(p)$. 可以把不可克隆定理看成，等价于在量子力学中非正交状态的可访问量子信息一般而言小于制备的熵的命题.

盒子 12.1 不可克隆定理

是否可能制作未知量子状态的备份？令人吃惊的是，该问题的答案事实上竟是否定的. 在本盒子中，我们给出一个抓住了本质原因的这个事实的一个初等证明. 设有一台具有标为 A 和 B 的两个插槽的量子机器. 槽 A，即数据槽，开始处在未知但为纯态的量子状态 $|\psi\rangle$，这是要被复制到目标槽槽 B 中的状态. 我们假定目标槽开始处于某个标准纯态 $|s\rangle$，因此复制机的初始状态为

$$|\psi\rangle\otimes|s\rangle \tag{12.1}$$

现在某个酉演化实现复制过程，理想情况下，

$$|\psi\rangle\otimes|s\rangle \xrightarrow{U} U(|\psi\rangle\otimes|s\rangle)=|\psi\rangle\otimes|\psi\rangle \tag{12.2}$$

设该复制过程对两个特别的纯态 $|\psi\rangle$ 和 $|\varphi\rangle$ 有效，则可得

$$U(|\psi\rangle\otimes|s\rangle)=|\psi\rangle\otimes|\psi\rangle \tag{12.3}$$

$$U(|\varphi\rangle\otimes|s\rangle)=|\varphi\rangle\otimes|\varphi\rangle \tag{12.4}$$

两个等式作内积得到

$$\langle\psi|\varphi\rangle=(\langle\psi|\varphi\rangle)^2 \tag{12.5}$$

但 $x=x^2$ 只有两个解，$x=0$ 和 $x=1$，故要么 $|\psi\rangle=|\varphi\rangle$，要么 $|\psi\rangle$ 与 $|\varphi\rangle$ 正交. 因此克隆装置只能克隆互相正交的状态，进而通用量子克隆装置是不存在的. 潜在的克隆装置不能克隆如 $|0\rangle$ 和 $|\psi\rangle=(|0\rangle+|1\rangle)/\sqrt{2}$ 的量子状态，因为它们是非正交的.

我们证明的是用酉演化不可能完全克隆未知量子状态. 自然引出几个问题：如果复制混合态会怎样？如果允许克隆装置是非酉的会怎样？如果我们允许不完美复制，只要按某种有价值的忠实性度量是“好”的，这会是什么呢？这些都是非常好的问题，如我们在章末的“历史和进一步阅读的材料”所看到，已有许多相关的研究. 简单来说，即使允许非酉克隆装置，非正交纯态的克隆仍是不可能的，除非我们允许复制状态在忠实度的有限损失. 类似结论对混合态也成立，尽管即使定义克隆混合态的概念都要用更复杂的方法.

我们贯穿本书强调的是，量子信息的隐含性质处在量子计算与量子信息能力的核心，可访问信息用量化的方式描述了量子信息的隐含性质. 遗憾的是，还不知道计算可访问信息的通用方法；不过可以证明一些重要的界，其中最重要的是 Holevo 界.

12.1.1　Holevo 界

Holevo 界是可访问信息的一个极常用的上界，并且在量子信息论的许多应用中扮演重要角色.

定理 12.1（Holevo 界）　设 Alice 以概率 $p_0,\cdots,p_n$ 制备状态 ρ_X，其中 $X=0,\cdots,n$. Bob 进行 POVM 元 $\{E_y\}=\{E_0,\cdots,E_m\}$ 描述的测量，测量结果是 Y. Holevo 界断言对 Bob 可以进行的任何此类测量有

$$H(X:Y)\leqslant S(\rho)-\sum_x p_x S(\rho_x) \tag{12.6}$$

其中 $\rho=\sum_x p_x\rho_x$.

因此 Holevo 界是可访问信息的一个上界. Holevo 界右边的量在量子信息论中非常有用，有一个专门的名称，为 Holevo χ 量，有时记为 χ.

证　Holevo 界的证明是通过构造漂亮的三量子系统而获得，这三量子系统分别记为 P，Q 和 M. 系统 Q 是 Alice 给 Bob 的量子系统；正如第 11 章中许多熵不等式的证明那样，P 和 M 是为证明引入的辅助系统. 直观上，可认为 P 是制备系统. 根据定义，它具有标准正交基底 $|x\rangle$，它的元素对应量子系统 Q 的可能制备的标号 $0,\cdots,n$. 可以从直观上把 M 看作 Bob 的测量装置，它有一个基底 $|y\rangle$，其元素对应 Bob 测量的可能输出 $1,\cdots,n$. 假设全系统的初态为

$$\rho^{PQM}=\sum_x p_x|x\rangle\langle x|\otimes\rho_x\otimes|0\rangle\langle 0| \tag{12.7}$$

其中按 PQM 顺序书写张量积分解. 直观上，该状态表示 Alice 以概率 p_x 选择 x 的一个值，制备一个相应的状态 ρ_x 并交给 Bob 的情形. Bob 将使用他的测量设备进行测量. 测量设备的初态为标准状态 $|0\rangle$. 为描述测量，我们引入仅作用在 Q 和 M 上（不在 P）的量子运算 ε，其作用是在系统 Q 上进行具有元 $\{E_y\}$ 的 POVM 测量，测量结果保存在系统 M 中：

$$\varepsilon(\sigma\otimes|0\rangle\langle 0|)\equiv\sum_y\sqrt{E_y}\,\sigma\sqrt{E_y}\otimes|y\rangle\langle y| \tag{12.8}$$

其中 σ 是系统 Q 的任意状态，$|0\rangle$ 是测量设备的初态. 在下面的练习中，读者将证明 ε 是一个保持迹的量子运算.

练习 12.2　定义 U_y 为作用在系统 M 上的酉算子，它在基底上的作用为 $U_y|y'\rangle\equiv|y'+y\rangle$，其中加法是模 $n+1$ 的，证明 $\{\sqrt{E_y}\otimes U_y\}$ 定义了保迹量子运算 ε

的一组运算元，其中 ε 在形如 $\sigma\otimes|0\rangle\langle 0|$ 状态上的作用与式(12.8)一致.

Holevo 界的证明过程如下. 用“′”表示应用 ε 后 PQM 的状态，不带撇的状态表示应用 ε 前. 我们有 $S(P:Q)=S(P:Q,M)$，因为 M 开始与 P 和 Q 不相关；而 $S(P:Q,M)\geqslant S(P':Q',M')$，因为应用 ε 到 QM 不能增加 P 和 QM 之间的互信息(定理 11.15)；最后 $S(P':Q',M')\geqslant S(P':M')$，因为丢弃系统不会增加互信息(仍参考定理 11.15). 这些结果合起来给出

$$S(P':M')\leqslant S(P:Q) \tag{12.9}$$

经过简单代数计算，容易看出这就是 Holevo 界. 首先计算右边的量. 注意到

$$\rho^{PQ}=\sum_x p_x|x\rangle\langle x|\otimes\rho_x \tag{12.10}$$

从而可知 $S(P)=H(p_x)$，$S(Q)=S(\rho)$ 和 $S(P,Q)=H(p_x)+\sum_x p_x S(\rho_x)$(根据定理 11.10)，于是

$$S(P:Q)=S(P)+S(Q)-S(P,Q)=S(\rho)-\sum_x p_x S(\rho_x) \tag{12.11}$$

恰好是我们希望的 Holevo 界的右边. 为计算式(12.9)左边的量，注意

$$\rho^{P'Q'M'}=\sum_{xy}p_x|x\rangle\langle x|\otimes\sqrt{E_y}\,\rho_x\sqrt{E_y}\otimes|y\rangle\langle y| \tag{12.12}$$

对 Q' 求迹，并注意到 (X,Y) 对的联合分布 $p(x,y)$ 满足 $p(x,y)=p_x p(y|x)=p_x\mathrm{tr}(\rho_x E_y)=p_x\mathrm{tr}(\sqrt{E_y}\,\rho_x\sqrt{E_y})$，可得

$$\rho^{P'M'}=\sum_{xy}p(x,y)|x\rangle\langle x|\otimes|y\rangle\langle y| \tag{12.13}$$

从而 $S(P':M')=H(X:Y)$，正是我们希望的 Holevo 界的左边. Holevo 界的证明至此完成. □

12.1.2 应用 Holevo 界的例子

Holevo 界是证明量子信息论许多结果的基石. 这里，我们只是给出应用该重要结果的皮毛. 回忆定理 11.10，它意味着

$$S(\rho)-\sum_x p_x S(\rho_x)\leqslant H(X) \tag{12.14}$$

而且当且仅当状态 ρ_x 具有正交支集时取等号. 设状态 ρ_x 不具有正交支集，则不等式(12.14)是严格的. 于是 Holevo 界蕴含 $H(X:Y)$ 严格小于 $H(X)$，从而基于测量结果 Y，Bob 不可能以完全的可靠性确定 X. 这推广了我们已经知道的知识，若 Alice 制备的状态不是正交的，则 Bob 不可能完全确定 Alice 制备的是哪个状态.

一个具体例子是，Alice 按照投掷一枚均匀硬币的结果制备处于两个量子态之一的单量子比特. 若投出硬币的正面，则 Alice 制备状态 $|0\rangle$，如果投出反面，则 Alice 制备状态 $\cos\theta|0\rangle+\sin\theta|1\rangle$，其中 θ 是某个实参数. 在 $|0\rangle$，$|1\rangle$ 基底中，ρ 可以

写作

$$\rho=\frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}+\frac{1}{2}\begin{bmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{bmatrix} \tag{12.15}$$

简单计算表明，ρ 的特征值为 $(1\pm\cos\theta)/2$，因此 Holevo 界可由二元熵 $H((1+\cos\theta)/2)$ 给出，如图 12.1 所示. 注意，当 $\theta=\pi/2$ 时，Holevo 界达到最大值 1 比特，对应于 Alice 制备的状态是从正交集中所选择的，此时 Bob 可以完全断定 Alice 制备的是哪个状态.

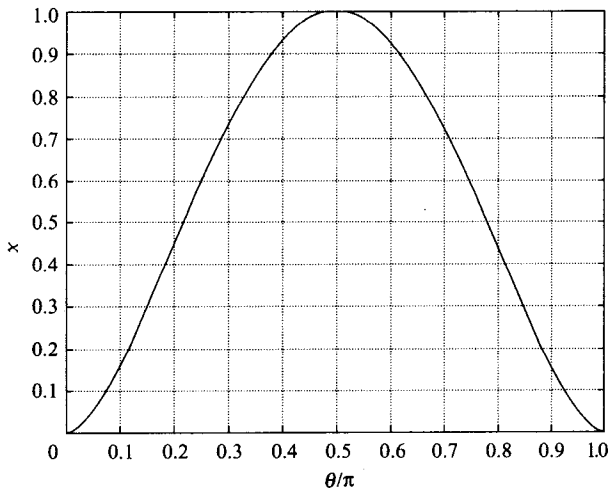


图 12.1　当以等概率制备状态 $|0\rangle$ 和 $\cos\theta|\rangle+\sin\theta|1\rangle$ 时，Holevo 界 χ 作为 θ 函数的图. 注意当 $\theta=\pi/2$ 时，相应于正交状态情形，Holevo 界达到最大值. 只有在这种情况下，Bob 能够完全确定 Alice 制备了哪个状态.

利用 Fano 不等式(参见盒子 12.2 的推导)，可以给予 Holevo 界更多的操作

盒子 12.2　Fano 不等式

假设我们希望从另一随机变量 Y 的知识推断一个随机变量 X 的取值. 直观上可以预期条件熵 $H(X|Y)$ 限制着我们作这种推断的能力. Fano 不等式是这个直观想法严格的表述，并且为给定 Y 条件下在多大程度上可以推断 X 给出了一个有用的界.

设 $\widetilde{X}\equiv f(Y)$ 是 Y 的某个函数，是我们关于 X 的最好猜测，令 $p_e\equiv p(X\neq\widetilde{X})$ 为这个猜测不正确的概率，则 Fano 不等式断言

$$H(p_e)+p_e\log(|X|-1)\geqslant H(X|Y) \tag{12.16}$$

其中 $H(\cdot)$ 是二元熵，而 $|X|$ 是假设的 X 可能取值的个数. 该不等式的定性含

义是，如果 $H(X|Y)$ 大（即与 $\log(|X|-1)$ 的大小可比），则推断中出错的概率 p_e 也将是大的.

为证明 Fano 不等式，定义一个误差随机变量，$E\equiv 1$ 若 $X\neq\widetilde{X}$，$E\equiv 0$ 若 $X=\widetilde{X}$. 注意 $H(E)=H(p_e)$. 利用条件熵的链式法则，可得 $H(E,X|Y)=H(E|X,Y)+H(X|Y)$. 但一旦 X 和 Y 已知，E 将完全确定，故 $H(E|X,Y)=0$ 从而 $H(E,X|Y)=H(X|Y)$. 再次应用链式法则到不同的变量上，可得 $H(E,X|Y)=H(X|E,Y)+H(E|Y)$. 条件减小熵，故 $H(E|Y)\leqslant H(E)=H(p_e)$，于是 $H(X|Y)=H(E,X|Y)\leqslant H(X|E,Y)+H(p_e)$. 按如下方式对 $H(X|E,Y)$ 估界，就可完成 Fano 不等式的证明（我们省略了一些简单的代数过程，读者很容易补上）：

$$\begin{aligned} H(X\mid E,Y) &= p(E=0)H(X\mid E=0,Y)+ \\ &\quad p(E=1)H(X\mid E=1,Y) \qquad (12.17)\\ &\leqslant p(E=0)\times 0+p_e\log(|X|-1) \\ &= p_e\log(|X|-1) \qquad (12.18)\end{aligned}$$

其中 $H(X|E=1,Y)\leqslant\log(|X|-1)$ 的根据是当 $E=1$ 时 $X\neq Y$，并且 X 最多取 $|X|-1$ 个值，这限制了它的熵，进而给出了条件熵的上界 $\log(|X|-1)$. 把 $H(X|E,Y)\leqslant p_e\log(|X|-1)$ 代入 $H(X|Y)\leqslant H(X|E,Y)+H(p_e)$，就给出 Fano 不等式 $H(X|Y)\leqslant H(p_e)+p_e\log(|X|-1)$.

性含义. 设基于测量 Y 和某种猜测规则，使用函数 $f(\cdot)$，Bob 对 Alice 制备的状态做出形如 $\widetilde{X}=f(Y)$ 的猜测，那么按照 Fano 不等式和 Holevo 界，

$$\begin{aligned} H(p(\widetilde{X}\neq X))+p(\widetilde{X}\neq X)\log(|X|-1) &\geqslant H(X\mid Y) \\ &= H(X)-H(X:Y) \\ &\geqslant H(X)-\chi \qquad (12.19)\end{aligned}$$

这就使我们可以对 Bob 对 X 值的推断进行估界. 粗略地说，χ 越小，Bob 越难以确定 Alice 制备的是哪个状态. 图 12.2 对 Alice 以一半的概率制备 $|0\rangle$，一半概率制备 $\cos\theta|\rangle+\sin\theta|1\rangle$ 说明了这一点. 如我们已看到的，此时界减小为 $H(p(\widetilde{X}\neq X))\geqslant 1-\chi$ 且 $\chi=H((1+\cos(\theta))/2)$. 注意当 $\theta\neq\pi/2$ 时，Bob 的猜测以某个有限的概率出错. 该错误概率随 θ 接近零而增大，最后，当 $\theta=0$ 时，两个状态将无法区分. 下界告诉我们 Bob 出错的概率至少为一半——如我们可预料的，他对 Alice 制备状态的猜测不能比碰运气做得更好.

练习 12.3 用 Holevo 界论证，n 量子比特不能用于传送多于 n 比特的经典信息.

练习 12.4 设 Alice 发送给 Bob 如下四个纯态的均匀混合：

$$|X_1\rangle = |0\rangle \tag{12.20}$$

$$|X_2\rangle = \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}|1\rangle\right] \tag{12.21}$$

$$|X_3\rangle = \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}e^{2\pi i/3}|1\rangle\right] \tag{12.22}$$

$$|X_4\rangle = \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}e^{4\pi i/3}|1\rangle\right] \tag{12.23}$$

证明 Bob 的测量和 Alice 传送的状态之间的最大互信息小于 1 比特. 已知有一个约为0.415 比特的 POVM. 读者能够构造该测量，甚至更好的测量，以达到 Holevo 界吗?

12.2　数据压缩

让我们转而考察一个经典和量子信息论都有的基本动态过程——数据压缩. 数据压缩问题的最一般形式是确定存储一个信源在物理上的最低要求是什么. 这是信息论最基本的问题之一，其影响远远超过直接的应用范围. 无论在经典还是在量子信息量论中，解决这个问题所用到的技术事实上具有比仅仅是数据压缩广得多的应用范围，不过在数据压缩上它们的表现形式或许最简单和最优雅. 本节就来详细考察量子的和经典的数据压缩.

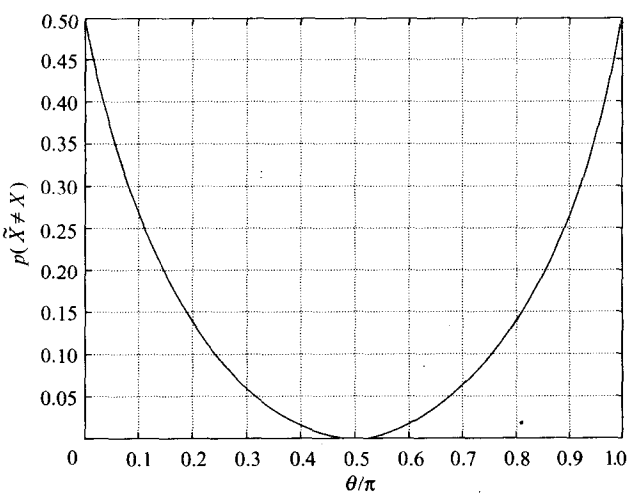


图 12.2　Bob 推断 Alice 制备状态 $|0\rangle$ 还是 $\cos\theta|\rangle + \sin\theta|1\rangle$ 出错概率的一个下界. 该下界是将 Fano 不等式与 Holevo 界结合得到的. 注意该界随 θ 接近 $\pi/2$ 减小到 0，此时状态可以可靠区分.

12.2.1 Shannon 无噪声信道编码定理

Shannon 无噪声信道编码定理量化了由经典信源产生的信息的压缩程度. 什么是经典信源? 信源有许多模型. 一个简单而有用的模型是随机变量序列 X_1, X_2,…构成的源,其随机变量的值表示该源的输出. 为方便起见,假定随机变量取值于有限字母表的符号,尽管扩展到无穷字母表也是对的. 进而假设对该源的不同使用是独立和同分布的; 即该源是所谓 i. i. d 信源. 易见,例如,读者阅读的英文文本中的字母并非独立; 不同字母之间存在很强的相关性. 举个简单的例子来说,字母"t"后边跟一个字母"h"的频繁程度要高于在通常英文中字母"h"出现的总频率; 我们说"t"和"h"的出现不独立,而是相关的. 对大量不同类型的信源(包括英文文本),i. i. d 信源的假定在实践中应用得很好,并且为处理 i. i. d 信源这类特殊情况引入的想法可以推广到更复杂的信源.

在讨论 Shannon 定理的技术细节之前,让我们通过例子了解该结果的直观含义. 设 i. i. d 信源产生比特 $X_1, X_2, X_3, \cdots$. 每个比特以概率 p 等于 0,以概率 $1-p$ 等于 1. Shannon 定理的关键思想是把对随机变量 $X_1, \cdots, X_n$ 的值 $x_1, \cdots, x_n$ 的可能序列划分为两类——非常可能出现的序列,称为典型序列和很少出现的序列,称为非典型序列. 这是怎样做的呢? 当 n 增大时,我们可以以很高的概率预期从该信源输出中比例为 p 的符号将等于 0,而比例为 $1-p$ 的符号将为 1. 使这项假定为真的序列 $x_1, \cdots, x_n$ 称为典型序列. 把这个定义和对信源的独立性假设结合起来,得到

$$p(x_1, \cdots, x_n) = p(x_1)p(x_2)\cdots p(x_n) \approx p^{np}(1-p)^{(1-p)n} \tag{12.24}$$

这对典型序列成立. 两边取对数得

$$\begin{aligned} -\log p(x_1, \cdots, x_n) &\approx -np\log p - n(1-p)\log(1-p) \\ &= nH(X) \end{aligned} \tag{12.25}$$

其中 X 是服从信源分布的随机变量且 $H(X) = -p\log(p) - (1-p)\log(1-p)$ 是源分布的熵,也称为源的熵率,因此 $p(x_1, \cdots, x_n) \approx 2^{-nH(X)}$. 由此可见至多有 $2^{nH(X)}$ 个典型序列,因为所有典型序列的总概率不会超过 1.

现在我们已有了理解数据压缩的一种简单方案的工具. 设信源的输出为 $x_1, \cdots, x_n$,为压缩该输出,我们检查一下 $x_1, \cdots, x_n$ 是否为典型序列. 如果不是,我们就放弃,即报告错误. 幸运的是,当 n 变大时,这种情况会非常少见,因为几乎所有的序列在 n 为极大情况下都是典型的. 若输出是典型的,我们记录这一事实. 由于最多有 $2^{nH(X)}$ 个典型序列,仅需要 $nH(X)$ 比特来惟一识别一个特定的典型序列. 我们选取某种这样的方案辨别信源的输出,并将其压缩到相应的典型序列的 $nH(X)$ 比特串描述. 该串以后可以解压缩. 随 n 的增大这个方案成功的概率

将接近 1.

这种方案有几点可挑剔之处：(1)它有一个小的且有限的失效概率. 采用类似思想的稍微复杂些的方案可以避免出错的可能. (2)为压缩，必须等待信源发出符号的量要大到 n 个. 同样可以修改方案，使得处理在信源发出符号的同时进行. (3)该方案未给出从信源的输出到压缩序列的显式映射. 我们再次指出，可以给出稍微复杂些的方案来解决这个问题. (4)用于数据压缩的具体过程依赖于源的分布. 如果不知道会如何？巧妙的通用压缩算法可以用来处理这种情况. 对这些问题感兴趣的读者可以参考本章末“历史和进一步阅读的材料”中列出的 Cover 和 Thomas 的书.

让我们把典型序列的概念推广到二进制以外的情形. 设 $X_1, X_2, \cdots$ 是 i. i. d 信源. 一般地，任意给定字母 x 出现在信源输出序列中的频率，都接近于信源任何一次使用中该字母出现的概率 $p(x)$. 基于这种直观解释，下面给出典型序列的严格定义. 给定 $\varepsilon>0$，信源的一串符号 $x_1 x_2 \cdots x_n$ 称为 ε 典型，若

$$2^{-n(H(X)+\varepsilon)} \leqslant p(x_1, \cdots, x_n) \leqslant 2^{-nH((X)-\varepsilon)} \tag{12.26}$$

我们用 $T(n,\varepsilon)$ 来记所用这种 ε 典型序列的集合. 该定义有一个有用的等价表述

$$\left| \frac{1}{n} \log \frac{1}{p(x_1, \cdots, x_n)} - H(X) \right| \leqslant \varepsilon \tag{12.27}$$

利用大数律(在盒子 12.3 中叙述和证明)，我们可以证明典型序列定理. 它严格化了如下思想，在 n 充分大时，信源输出的大多数序列是典型序列.

定理 12.2(典型序列定理)　(1) 固定 $\varepsilon>0$，则对任意的 $\delta>0$ 和充分大的 n，一个序列为 ε 典型的概率至少是 $1-\delta$.

(2) 对任意固定的 $\varepsilon>0$ 和 $\delta>0$，对充分大的 n，ε 典型序列的数目 $|T(n,\varepsilon)|$ 满足

$$(1-\delta)2^{n(H(X)-\varepsilon)} \leqslant |T(n,\varepsilon)| \leqslant 2^{n(H(X)+\varepsilon)} \tag{12.28}$$

(3) 令 $S(n)$ 为信源发出的长为 n 的序列的大小至多为 2^{nR} 的集合，其中 $R<H(X)$ 为固定，则对任意 $\delta>0$ 和充分大的 n，有

$$\sum_{x \in S(n)} p(x) \leqslant \delta \tag{12.29}$$

证　第 1 部分：大数律的直接应用. 注意 $-\log p(X_i)$ 为独立同分布随机变量，由大数律，对任意 $\varepsilon>0$ 和 $\delta>0$ 以及充分大的 n，有

$$p\left(\left| \sum_{i=1}^{n} \frac{-\log p(X_i)}{n} - E(-\log p(X)) \right| \leqslant \varepsilon \right) \geqslant 1-\delta \tag{12.30}$$

但 $E(\log p(X)) = -H(X)$ 且 $\sum_{i=1}^{n} \log p(X_i) = \log(p(X_1, \cdots, X_n))$，因此

$$p(|-\log(p(X_1, \cdots, X_n))/n - H(X)| \leqslant \varepsilon) \geqslant 1-\delta \tag{12.31}$$

即，一个序列是 ε 典型的概率至少为 $1-\delta$.

第 2 部分：由典型性定义，注意典型序列概率之和必然在 $1-\delta$（由第 1 部分）和 1 之间（因为概率加起来不会超过 1），因此

$$\begin{aligned} 1 &\geqslant \sum_{x\in T(n,\varepsilon)} p(x) \\ &\geqslant \sum_{x\in T(n,\varepsilon)} 2^{-n(H(X)+\varepsilon)} \\ &= |\ T(n,\varepsilon)\ | \ 2^{-n(H(X)+\varepsilon)} \end{aligned} \tag{12.32}$$

从而可知，$|T(n,\varepsilon)|\leqslant 2^{n(H(X)+\varepsilon)}$，以及

$$\begin{aligned} 1-\delta &\leqslant \sum_{x\in T(n,\varepsilon)} p(x) \\ &\leqslant \sum_{x\in T(n,\varepsilon)} 2^{-n(H(X)-\varepsilon)} \\ &= |\ T(n,\varepsilon)\ | \ 2^{-n(H(X)-\varepsilon)} \end{aligned} \tag{12.33}$$

进而得到 $|T(n,\varepsilon)|\geqslant(1-\delta)2^{n(H(X)-\varepsilon)}$.

盒子 12.3 大数律

大量重复一项实验，每次都测量某个参数的值 X. 我们把实验的结果标为 $X_1, X_2, \cdots$. 假设实验结果是独立的，直觉上我们预期均值 $E(X)$ 的估计 $S_n \equiv \sum_{i=1}^{n} X_i/n$，当 $n\to\infty$ 时应该趋向 $E(X)$. 大数律是对这一直观现象的严格表述.

定理 12.3（大数律） 设 $X_1, X_2, \cdots$ 是独立随机变量且都具有与随机变量 X 相同的分布，X 具有有限的一阶和二阶矩，$|E(X)|<\infty$ 和 $E(X^2)<\infty$，则对任意 $\varepsilon>0$，当 $n\to\infty$ 时 $p(|S_n-E(X)|>\varepsilon)\to 0$.

证 先假定 $E(X)=0$，在证明完成后讨论 $E(X)\neq 0$ 的情形. 因为随机变量均值为零，且独立，故对 $i\neq j$，$E(X_iX_j)=E(X_i)E(X_j)=0$，从而

$$E(S_n^2)=\frac{\sum_{i,j=1}^{n}E(X_iX_j)}{n^2}=\frac{\sum_{i=1}^{n}E(X_i^2)}{n^2}=\frac{E(X^2)}{n} \tag{12.34}$$

其中最后一个等式来源于 $X_1, \cdots, X_n$ 与 X 同分布的事实. 根据同样道理，从期望的定义可得

$$E(S_n^2)=\int \mathrm{d}P S_n^2 \tag{12.35}$$

其中 $\mathrm{d}P$ 是相应的概率测度. 显然，要么 $|S_n|\leqslant\varepsilon$，要么 $|S_n|>\varepsilon$，于是可将积分分成两部分，并进而舍弃一部分，因为其为非负，可得

$$E(S_n^2)=\int_{|S_n|\leqslant\varepsilon}\mathrm{d}PS_n^2+\int_{|S_n|>\varepsilon}\mathrm{d}PS_n^2\geqslant\int_{|S_n|>\varepsilon}\mathrm{d}PS_n^2 \tag{12.36}$$

在积分区域内 $S_n^2>\varepsilon^2$，故

$$E(S_n^2)\geqslant\varepsilon^2\int_{|S_n|>\varepsilon}\mathrm{d}P=\varepsilon^2 p(|S_n|>\varepsilon) \tag{12.37}$$

将此不等式与式(12.34)相比较，可知

$$p(|S_n|>\varepsilon)\leqslant\frac{E(X^2)}{n\varepsilon^2} \tag{12.38}$$

令 $n\to\infty$就完成证明. 当 $E(X)\neq0$ 的情形，通过定义

$$Y_i\equiv X_i-E(X),\quad Y\equiv X-E(X) \tag{12.39}$$

容易得到相应的结果. Y 和 $Y_1,Y_2,\cdots$是一列独立同分布随机变量，满足 $E(Y)=0$ 和 $E(Y^2)<\infty$. 期望的结果可以从前面的证明得到. □

第 3 部分：基本思想是把 $S(n)$中的序列划分成典型和非典型序列. 当 n 充分大时，非典型序列的概率变小. 显然 $S(n)$中典型序列的数目不会超过 $S(n)$中序列的总数，即最多 2^{nR}. 而每个典型序列出现的概率约为 $2^{-nH(X)}$，因此 $S(n)$中典型序列的总概率的大小接近 $2^{n(R-H(X))}$，它在 $R<H(X)$时趋向于零.

更严格地，选择 ε 使得 $R<H(X)-\delta$ 且 $0<\varepsilon<\delta/2$. 把 $S(n)$中序列分为 ε 典型和 ε 非典型序列. 由第 1 部分，对充分大的 n，非典型序列的总概率可以做到小于 $\delta/2$. 在 $S(n)$中最多有 2^{nR} 个典型序列，每个出现的概率最多是 $2^{-n(H(X)-\varepsilon)}$，因此典型序列的概率最多是 $2^{-n(H(X)-\varepsilon-R)}$. 随 n 趋向无穷，它将趋向于零，因此对充分大的 n，$S(n)$中序列的总概率小于 δ. □

Shannon 无噪声信道编码定理是典型序列定理的简单应用. 这里给出无噪声信道编码定理的一个非常简单的版本；更复杂的版本留到练习和本章末“历史和进一步阅读的材料”中. 基本假设是设 $X_1,X_2,\cdots$，为包含 d 个符号的字母表上的一个 i. i. d 经典信源. 一个比率为 R 的压缩方案把可能的序列 $x=(x_1,\cdots,x_n)$映射为长度为 nR 的比特串，记为 $C^n(x)=C^n(x_1,\cdots,x_n)$(注意 nR 可能不是整数；为简化记号，约定此时 $nR=\lfloor nR\rfloor$). 相匹配的解压缩方案将压缩的 nR 比特映射回到字母表的 n 字母串，$D^n(C^n(x))$. 如果 n 趋于∞时，$D^n(C^n(x))=x$ 的概率趋于 1 则称压缩-解压缩方案(C^n,D^n)为可靠的. Shannon 无噪声信道编码定理说明了对某比率 R 值可靠压缩方案是存在的，为熵率 $H(X)$提供了著名的操作性解释：即它正是可靠存储一个信源输出充分和必要的最少物理资源.

定理 12.4(Shannon 无噪声信道编码定理)　设$\{X_i\}$是一个具有熵率 $H(X)$ 的 i. i. d 信源. 假设 $R>H(X)$，则对该信源存在比率为 R 的可靠压缩方案；反之，若 $R<H(X)$，则任何压缩方案都不是可靠的.

证　设 $R>H(R)$. 选择 $\varepsilon>0$，使得 $H(X)+\varepsilon<R$. 考虑 ε 典型序列的集合 $T(n,\varepsilon)$. 对任意 $\delta>0$ 和充分大的 n，最多有 $2^{n(H(x)+\varepsilon)}<2^{nR}$ 个这样的序列，且信源产

.生这样序列的概率至少为 $1-\delta$，从而压缩方法可以简单地检查信源的输出是否为 ε 典型的. 如果不是，则压缩到某个固定的指示失败的 nR 比特串；解压缩运算只是输出一个随机序列 $x_1,\cdots,x_n$ 作为信源产生信息的一个猜测；实际上我们在这种情况下放弃压缩. 如果信源的输出是典型的，那么我们通过用 nR 比特以显然的方式存储特定序列的索引来直接压缩输出，以便后来恢复.

设 $R<H(X)$. 压缩-解压缩运算的合成最多有 2^{nR} 个输出，故信源的输出序列中最多有 2^{nR} 个可以无差错地进行压缩-解压缩. 由典型序列定理，对充分大的 n，对 $R<H(X)$，从信源输出的序列落在一个 2^{nR} 序列的子集中的概率趋于 0，因此任何这样的压缩方案不可能是可靠的. □

练习 12.5（可变长度零差错数据压缩） 考虑如下可变长度数据压缩方案的粗略解释. 令 $x_1,\cdots,x_n$ 是从一个熵率为 $H(X)$ 的 i. i. d 信源的 n 次使用得到的输出. 若 $x_1,\cdots,x_n$ 是典型的，则发送指示它的 $H(X)$ 比特索引. 如果 $x_1,\cdots,x_n$ 是非典型的，则为此序列发送一个未压缩的 $\log d^n$ 比特索引（回忆 d 是字母表的大小）. 试把这个粗略方案严格化为，对任意 $R>H(X)$，该信源可以以零差错概率被压缩到平均每个源符号使用 R 比特.

12.2.2 Schumacher 量子无噪声信道编码定理

量子信息论概念上的一个伟大突破是，认识到我们可以把量子状态视为信息，并关于量子状态提出信息论问题. 本节中我们要定义量子信源，并研究如下问题：信源产生的信息——量子状态——在多大程度上可以被压缩.

如何定义量子信源的概念？就像经典信源的定义，没有做出该定义的最佳方式，并且可以有几种不同的定义方式，且未必都是等价的. 我们采用的定义的基本想法是，把纠缠作为压缩和解压缩的对象. 更形式化地，一个（i. i. d）量子信源由一个 Hilbert 空间 H 和该 Hilbert 空间上的一个密度矩阵 ρ 所描述. 我们设想系统的状态 ρ 仅仅是一个处在纯态的更大系统的一部分，而 ρ 的混合性质是 H 和系统剩余部分的纠缠造成的. 对该信源的比率为 R 的压缩方案由两族运算 $\mathcal{C}^n$ 和 $\mathcal{D}^n$ 组成，这类似于经典情形的压缩解压缩方案. $\mathcal{C}^n$ 是压缩运算，把 $H^{\otimes n}$ 中的状态映射到 2^{nR} 维状态空间，压缩后空间中的状态. 我们可以把压缩后空间看成代表 nR 量子比特. 运算 $\mathcal{D}^n$ 是一个解压缩运算，把压缩后空间的状态映射为原来的状态空间. 因此压缩-解压缩运算的合成是 $\mathcal{D}^n\circ\mathcal{C}^n$. 我们关于可靠性的准则是对充分大的 n，纠缠忠实度 $F(\rho^{\otimes n},\mathcal{D}^n\circ\mathcal{C}^n)$ 应当趋于 1. 量子数据压缩的基本思路如图 12.3 所示.

使量子无噪声信道编码定理成为可能的关键技术思想是，量子版本的典型序列概念. 设与一个量子信源相关联的密度算子 ρ 具有标准正交分解：

$$\rho=\sum_x p(x)\,|x\rangle\langle x| \tag{12.40}$$

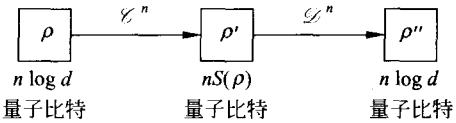


图 12.3　量子数据压缩. 压缩运算 $\mathscr{C}^n$ 把一个存储在 $n\log d$ 量子比特的量子源 ρ 压缩为 $nS(\rho)$ 量子比特，该源通过解压缩运算 $\mathscr{D}^n$ 被准确恢复.

其中 $|x\rangle$ 是标准正交集，$p(x)$ 是 ρ 的特征值. ρ 的特征值服从与概率分布相同的规则：它们是非负的，且和为 1，而且，$H(p(x))=S(\rho)$. 于是，可以和经典定义一样，ε 典型序列 $x_1,\cdots,x_n$ 满足

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)p(x_2)\cdots p(x_n)}\right)-S(\rho)\right|\leqslant\varepsilon \tag{12.41}$$

是有意义的. ε 典型序列是一个使得 $x_1,x_2,\cdots,x_n$ 成为 ε 典型的状态 $|x_1\rangle|x_2\rangle\cdots|x_n\rangle$. 定义 ε 典型子空间为由所有 ε 典型状态 $|x_1\rangle|x_2\rangle\cdots|x_n\rangle$ 张成的子空间. 我们把 ε 典型子空间记作 $T(n,\varepsilon)$，并把到 ε 典型子空间上的投影记作 $P(n,\varepsilon)$. 注意到

$$P(n,\varepsilon)=\sum_{x_i\in\text{典型}}|x_1\rangle\langle x_1|\otimes|x_2\rangle\langle x_2|\otimes\cdots|x_n\rangle\langle x_n| \tag{12.42}$$

现在可把典型序列定理翻译到如下的等价量子形式，典型子空间定理.

定理 12.5（典型子空间定理）　(1) 固定 $\varepsilon>0$. 则对任意 $\delta>0$ 和充分大的 n，

$$\mathrm{tr}(P(n,\varepsilon)\rho^{\otimes n})\geqslant 1-\delta \tag{12.43}$$

(2) 对任意固定的 $\varepsilon>0$ 和 $\delta>0$，以及充分大的 n，$T(n,\varepsilon)$ 的维数 $|T(n,\varepsilon)|=\mathrm{tr}(P(n,\varepsilon))$，满足

$$(1-\delta)2^{n(S(\rho)-\varepsilon)}\leqslant|T(n,\varepsilon)|\leqslant 2^{n(S(\rho)+\varepsilon)} \tag{12.44}$$

(3) 令 $S(n)$ 为到 $H^{\otimes n}$ 的任意至多 2^{nR} 维的子空间的一个投影，其中 $R<S(\rho)$ 为固定，则对任意 $\delta>0$ 和充分大的 n，有

$$\mathrm{tr}(S(n)\rho^{\otimes n})\leqslant\delta \tag{12.45}$$

对每种情况，结果都可以通过直接运用大数律得到，但我们宁愿利用典型序列定理以强调与证明 Shannon 无噪声信道编码定理的技术的紧密联系.

证　第 1 部分：注意

$$\mathrm{tr}(P(n,\varepsilon)\rho^{\otimes n})=\sum_{x_i\in\text{典型}}p(x_1)p(x_2)\cdots p(x_n) \tag{12.46}$$

结果可以立刻从典型序列定理的第 1 部分得到.

第 2 部分：可直接从典型序列定理的第 2 部分得到.

第 3 部分：我们把迹分为在典型子空间上的迹和非典型子空间上的迹，

$$\mathrm{tr}(S(n)\rho^{\otimes n})=\mathrm{tr}(S(n)\rho^{\otimes n}P(n,\varepsilon))+\mathrm{tr}(S(n)\rho^{\otimes n}(I-P(n,\varepsilon))) \tag{12.47}$$

且对每项分别估界. 对第一项可知,

$$\rho^{\otimes n}P(n,\varepsilon) = P(n,\varepsilon)\rho^{\otimes n}P(n,\varepsilon) \tag{12.48}$$

因为 $P(n,\varepsilon)$是与 $\rho^{\otimes n}$可对易的投影,但

$$\mathrm{tr}(S(n)P(n,\varepsilon)\rho^{\otimes n}P(n,\varepsilon)) \leqslant 2^{nR}2^{-n(S(\rho)-\varepsilon)} \tag{12.49}$$

因为 $P(n,\varepsilon)\rho^{\otimes n}P(n,\varepsilon)$的特征值有上界 $2^{-n(S(\rho)-\varepsilon)}$. 令 $n\to\infty$,可以看到第 1 项趋于 0. 对第 2 项,注意 $S(n)\leqslant I$. 由于 $S(n)$和 $\rho\otimes(I-P(n,\varepsilon))$都是半正定算子,故 $0\leqslant \mathrm{tr}(S(n)\rho^{\otimes n}(I-P(n,\varepsilon)))\leqslant \mathrm{tr}(\rho^{\otimes n}(I-P(n,\varepsilon)))\to 0$,当 $n\to\infty$. 于是第 2 项当 n 增大时也趋于 0,至此导出结论. □

有了典型子空间定理,我们不难证明 Shannon 无噪声信道编码定理的量子形式. 证明的主要思想是类似的,但由于证明中非对易算子的出现使技术上的分析更困难些. 这种情况没有经典的对应物.

定理 12.6(Schumacher 无噪声信道编码定理) 令$\{H,\rho\}$是 i. i. d 量子信源. 若 $R>S(\rho)$,则对该源$\{H,\rho\}$存在比率为 R 的可靠压缩方案. 若 $R<S(\rho)$,则比率 R 的任何压缩方案都不是可靠的.

证 设 $R>S(\rho)$且取 $\varepsilon>0$,使满足 $S(\rho)+\varepsilon\leqslant R$. 根据典型子空间定理,对任意的 $\delta>0$ 和充分大的 n,$\mathrm{tr}(\rho^{\otimes n}P(n,\varepsilon))\geqslant 1-\delta$,且 $\dim(T(n,\varepsilon))\leqslant 2^{nR}$. 令 H_c^n 为包含 $T(n,\varepsilon)$的任意 2^{nR}维 Hilbert 子空间,编码按如下方式进行. 首先进行由正交投影的完备集 $P(n,\varepsilon)$,$I-P(n,\varepsilon)$描述的测量. 相应的输出结果记为 0 和 1. 如果出现结果 0,什么也不做,状态留在典型子空间中. 如果出现结果 1,则将状态替换为从典型子空间中选出的某个标准状态$|0\rangle$;使用哪个状态并不重要. 事实上编码是一个映射 $\mathscr{C}^n$: $H^{\otimes n}\to H_c^n$,映射到 2^{nR}维子空间 H_c^n. 它具有算子和表示

$$\mathscr{C}^n(\sigma) \equiv P(n,\varepsilon)\sigma P(n,\varepsilon) + \sum_i A_i\sigma A_i^\dagger \tag{12.50}$$

其中 $A_i\equiv|0\rangle\langle i|$而$|i\rangle$是典型子空间正交补的标准正交基底.

解码运算 $\mathscr{D}^n$: $H_c^n\to H^{\otimes n}$定义为在 H_c^n 上恒等,$\mathscr{D}^n(\sigma)=\sigma$. 由这些编解码的定义,我们有

$$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{C}^n) = |\ \mathrm{tr}(\rho^{\otimes n}P(n,\varepsilon))\ |^2 + \sum_i |\ \mathrm{tr}(\rho^{\otimes n}A_i)\ |^2 \tag{12.51}$$

$$\geqslant |\ \mathrm{tr}(\rho^{\otimes n}P(n,\varepsilon))\ |^2 \tag{12.52}$$

$$\geqslant |\ 1-\delta\ |^2 \geqslant 1-2\delta \tag{12.53}$$

其中最后一行根据典型子空间定理得出. 但 δ 对充分大的 n 可变得任意小,故可知只要 $S(\rho)<R$ 总存在一个比率为 R 的可靠压缩方案$\{\mathscr{C}^n,\mathscr{D}^n\}$.

为证反过来的结论,设 $R<S(\rho)$. 不失一般性,设压缩运算把 $H^{\otimes n}$通过相应的投影 $S(n)$映射到一个 2^{nR}维子空间. 令 C_j 为压缩运算 $\mathscr{C}^n$ 的运算元,而 D_k 为解压缩运算的运算元,则我们有

$$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{E}^n)=\sum_{jk}|\operatorname{tr}(D_kC_j\rho^{\otimes n})|^2 \tag{12.54}$$

每个 C_j 算子都用投影 $S(n)$ 映射到子空间中，故 $C_j=S(n)C_j$. 令 $S^k(n)$ 为到 $S(n)$ 被 D_k 所映射到的子空间上的投影，则有 $S^k(n)D_kS(n)=D_kS(n)$ 且 $D_kC_j=D_kS(n)C_j=S^k(n)D_kS(n)C_j=S^k(n)D_kC_j$，其中

$$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{E}^n)=\sum_{jk}|\operatorname{tr}(D_kC_j\rho^{\otimes n}S^k(n))|^2 \tag{12.55}$$

应用 Cauchy-Schwarz 不等式得到

$$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{E}^n)\leqslant\sum_{jk}\operatorname{tr}(D_kC_j\rho^{\otimes n}C_j^\dagger D_k^\dagger)\operatorname{tr}(S^k(n)\rho^{\otimes n}) \tag{12.56}$$

根据典型子空间定理的第 3 部分可知，对任意的 $\delta>0$ 和充分大的 n，$\operatorname{tr}(S^k(n)\rho^{\otimes n})\leqslant\delta$. 进而典型子空间定理的证明蕴含，为使这一点成立的 n 不依赖于 k. 因此

$$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{E}^n)\leqslant\delta\sum_{jk}\operatorname{tr}(D_kC_j\rho^{\otimes n}C_j^\dagger D_k^\dagger) \tag{12.57}$$

$$=\delta \tag{12.58}$$

因为 $\mathscr{E}^n$ 和 $\mathscr{D}^n$ 是保迹的. 由于 δ 是任意的，故当 $n\to\infty$ 时，$F(\rho^{\otimes n},\mathscr{D}^n\circ\mathscr{E}^n)\to0$，从而该压缩方案是不可靠的. □

Schumacher 定理不仅讨论了可靠压缩方案的存在性，而且给出如何实际构造压缩方案的线索. 关键是要能够有效进行映入 2^{nR} 维典型子空间 H_c^n 的映射 $\mathscr{E}^n$：$H^{\otimes n}\to H_c^n$. 像枚举编码、Huffman 编码和算术编码那样的经典压缩技术虽然可以应用，但有一个很强的限制：编码线路必须是完全可逆的，并且在产生压缩后编码的过程中要完全擦除原来的状态！因为根据不可克隆定理，原状态无法复制，故不可能像通常的经典压缩方案那样在压缩后保持状态. 盒子 12.4 给出量子压缩工作过程的一个示例.

盒子 12.4　Schumacher 压缩

考虑由单量子比特密度矩阵

$$\rho=\frac{1}{4}\begin{bmatrix}3&1\\1&1\end{bmatrix} \tag{12.59}$$

刻画的 i. i. d 量子信源. 它可能来自比如说更大纠缠系统的一小部分. 从另一角度看这个信源（对比 9.3 节），它以各为一半的等概率产生状态 $|\psi_0\rangle=|0\rangle$ 或 $|\psi_1\rangle=(|0\rangle+|1\rangle)/\sqrt{2}$（见练习 12.8）. ρ 具有标准正交分解 $p|\bar0\rangle\langle\bar0|+(1-p)|\bar1\rangle\langle\bar1|$，其中 $|\bar0\rangle=\cos\frac{\pi}{8}|0\rangle+\sin\frac{\pi}{8}|1\rangle$，$|\bar1\rangle=-\sin\frac{\pi}{8}|0\rangle+\cos\frac{\pi}{8}|1\rangle$，而 $p=[3+\tan(\pi/8)]/4$. 在这个基底下，一组 n 个量子比特可写作状态

$$\sum_{X=\{\bar0\bar0\cdots\bar0,\bar0\cdots\bar0\bar1,\bar1\bar1\cdots\bar1\}}C_X|X\rangle \tag{12.60}$$

由定理 12.6，为以高忠实度可靠重构原始状态，只需要传送 Hamming 权重近似等于 np（即典型子空间的基底）的 $|X\rangle$. 这很容易理解，因为 $|\langle\bar{0}|\psi_k\rangle|=\cos(\pi/8)$（对 $k=\{0,1\}$）远大于 $|\langle\bar{1}|\psi_k\rangle|=\sin(\pi/8)$，且对具有大 Hamming 权重的 X，系数 C_X 非常小.

如何实现这一压缩方案？下面是一种近似方法. 设我们有量子线路 U_n. 它对基底状态 $|X\rangle$ 进行置换，使得状态按照 Hamming 权重进行字典排序. 例如，对 $n=4$，就是

$$\begin{array}{cccc}0000\to0000 & 1000\to0100 & 1001\to1000 & 1011\to1100\\ 0001\to0001 & 0011\to0101 & 1010\to1001 & 1101\to1101\\ 0010\to0010 & 0101\to0110 & 1100\to1010 & 1110\to1110\\ 0100\to0011 & 0110\to0111 & 0111\to1011 & 1111\to1111\end{array}$$

这样的变换，可以仅利用受控非门和 Toffoli 门实现，以可逆方式把典型子空间封装到前约为 $nH(p)$ 量子比特（从左到右）. 为完成该方案，还需要一个把单量子比特旋转到 $|\bar{0}\rangle,|\bar{1}\rangle$ 基底的量子门 V. 于是期望的压缩方案为 $\mathscr{C}^n=(V^\dagger)^{\otimes n}\cdot U_nV^{\otimes n}$，并且只需要发送 $\mathscr{C}^n$ 输出的前 $nH(p)$ 量子比特. 利用该线路的逆作为解码器，就能够以高忠实度重构出来自信源的状态序列. 一种更有效的编码方案是仅把 Hamming 权重约为 np 的状态封装到前 $nH(p)$ 量子比特空间；这可以用例如算术编码的量子版本来实现.

练习 12.6 在盒子 12.4 中，给出了一个 C_X 的关于 X 的具体表达式，同时描述了如何对于任意 n 构造执行 U_n 的量子线路. 作为 n 的函数，你需要多少个基本运算？

练习 12.7（数据压缩线路） 对任意 $R>S(\rho)=H(p)$，给出把一个满足 $\rho=p|0\rangle\langle0|+(1-p)|1\rangle\langle1|$ 的量子比特信源可靠压缩到 nR 量子比特的量子线路结构的框架.

练习 12.8（量子状态系综的压缩） 若采用下述系综定义，而不采用基于单一密度矩阵 ρ 和纠缠忠实度定义量子信源，则一个 i. i. d 量子信源将由量子状态的系综 $\{p_j,|\psi_j\rangle\}$ 给定，并且该源的连续使用是独立的并以概率 p_j 产生状态 $|\psi_j\rangle$. 一个压缩-解压缩方案 $(\mathscr{C}^n,\mathscr{D}^n)$ 称为是可靠的，如果系综平均忠实度当 $n\to\infty$ 时趋于 1：

$$\bar{F}\equiv\sum_J p_{j_1}\cdots p_{j_n}F(\rho_J,(\mathscr{D}^n\circ\mathscr{C}^n)(\rho_J))^2 \tag{12.61}$$

其中 $J=(j_1,\cdots,j_n)$ 且 $\rho_J\equiv|\psi_{j_1}\rangle\langle\psi_{j_1}|\otimes\cdots\otimes|\psi_{j_n}\rangle\langle\psi_{j_n}|$. 定义 $\rho\equiv\sum_j p_j|\psi_j\rangle\langle\psi_j|$，并证明在 $R>S(\rho)$ 条件下，存在对于如此定义的忠实度可靠的比率 R 压缩方案.

12.3　带噪声量子信道上的经典信息

可能出错的事情都会出错.

——引自 Edward A. Murphy, Jr.

我们打电话时偶尔会遇到麻烦. 当电话另一方的话极其难懂时，我们说遇到了坏线. 这是所有信息处理系统中某种程度上普遍存在的噪声现象的一个例子. 如第 10 章描述的，纠错码可以用来对抗噪声的影响，使得即使在噪声相当严重时仍能进行可靠通信和计算. 给定一条特定的带噪声信道 $\mathcal{N}$，一个有趣的问题是通过该信道可以可靠传输多少信息？ 例如，利用合适的纠错码，信道的 1000 次使用也许可以用来传送 500 比特的信息，这保证从任何由信道带来的差错中恢复的概率很高. 我们说这样的码具有比率 $500/1000=1/2$. 信息论的一个基本问题是要确定通过信道 $\mathcal{N}$ 可靠通信的最大传送率，即信道的容量.

对带噪声的经典信道，信道容量可以用称为 Shannon 带噪声信道编码定理的漂亮结果来计算. 下面就来考察在有噪声情况下经典信息的通信. 12.3.1 节讨论 Shannon 带噪声信道编码定理背后的一些主要思想. 不过我们不会涉及太多细节，因为 12.3.2 节将详细探讨两方试图使用带噪声量子信道传送经典信息这一更一般的问题.

12.3.1　带噪声经典信道上的通信

无论是量子还是经典的带噪声信道编码的许多主要思想，都可以通过研究二元对称信道来了解. 回忆 10.1 节的内容，二元对称信道是针对一个单比特信息的带噪声信道. 如图 12.4 所示，它的作用是以概率 $p>0$ 把传送的比特翻转，而以 $1-p$ 的概率无差错地传送该比特.

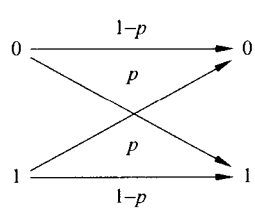


图 12.4　二元对称信道

每次使用二元对称信道可以可靠传送多少信息？ 尽管用纠错码可以通过该信道传送信息，但为完成通信需要付出额外的比特代价. 我们断言信息可以可靠传输的最大比率是 $1-H(p)$，这里 $H(\cdot)$ 是 Shannon 熵.

传送可靠完成是什么含义？ 这个问题很好，因为不同的答案会导致不同的比率. 我们采用如下的可靠性定义：我们假定信道的输入可以按大块一次性编码，并要求随块尺寸的增大，采用纠错码传送的差错会趋于 0. 可靠性的另一种定义是同样假定编码按块进行，但假定随块的增大，差错概率变得严格为 0. 遗憾的是，这个定义在纠错可以到达的性能上过于乐观，并且对二元对称信道导出零容量. 类似

地，如果不允许编码以大的块方式进行，容量也为 0. 的确相当令人惊奇（并且根本不显然）的是，甚至采用较保守的可靠性定义，都可以取得非零的信息传送比率. 为说明这一点需要若干巧妙的思想.

1. 二元对称信道的随机编码

设我们希望通过二元对称信道的 n 次使用来传送 nR 比特的信息；即希望以比率 R 经过信道传送信息. 我们将给出在 $R<1-H(p)$ 的条件下，概述对充分大的 n，存在有低出错概率纠错码的证明. 我们用到的第一个思想是构造纠错码的随机编码方法. 设 $(q,1-q)$ 是信道输入（0 和 1）上的任意固定的概率分布（该分布常称为码的先验分布——引入这个分布只是为随机编码方法能够工作，请不要将该分布中的随机性与信道的随机性混淆）. 接着选取一个码字 $x=(x_1,\cdots,x_n)$，为简单起见，独立地对 $j=1,\cdots,n$ 以概率 q 选择 $x_j=0$，以概率 $1-q$ 选择 $x_j=1$. 将这个过程重复 2^{nR} 次，产生具有 2^{nR} 项的码簿 C；我们把码簿的一项记作 x^j.

用这个过程显然可能构造出非常糟糕的纠错码. 我们可能非常不走运，所有码字由 n 个零组成，这显然对传送信息没有多少用处. 不过，事实上平均来看，这个随机编码过程给出了相当不错的纠错码. 为了解其中原因，让我们检查信道对编码的单个码字的影响. 由于所有码字是采用同样方式构造的，不妨检查第一个码字 x^1.

二元对称信道对 x^1 有什么影响？在一个长度为 n 的码字上，我们预期大概有 np 个比特被翻转，如图 12.5 所示，信道的输出有很高的概率与码字 x^1 具有大约 np 的 Hamming 距离. 我们问这样的输出在围绕 x^1 的半径为 np 的 Hamming 球

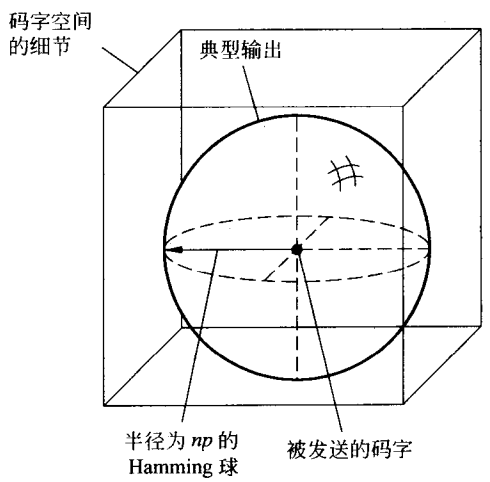


图 12.5 设码字 x^1 被二元对称信道的 n 次使用所传送. 那么信道的一个典型输出就是围绕被发送序列的半径为 np 的 Hamming 球（本图是图 12.6 的局部放大）.

上. 该 Hamming 球有多少元素？答案是大约 $2^{nH(p)}$ 个，因为 Hamming 球是由全部的具有 $y=x^1\oplus e$ 形式的典型输出序列构成. 这里 e 是信道发生的差错，$\oplus$ 表示模 2 加，而根据典型序列定理这样的典型差错 e 的数目约为 $2^{nH(p)}$.

刚才集中讨论了单个码字的情况，当然同类的差错会发生在所有的码字上. 图 12.6 描绘了所有码字和围绕它们的 Hamming 球的空间. 如果如图示那样，Hamming 球互不相交，那么 Bob 就有很简单的办法从信道的输出进行解码. 他只要检查输出是否落在其中一个 Hamming 球内，如果是这样就输出相应的码字，如果不是，就输出错误. 因为我们已经假定球互不相交，输入任何码字都有很高概率成功解码. 其实，即使球有些轻微的重叠，只要重叠不大，Bob 仍能够以较大概率进行成功解码——信道输出以很高的概率属于一个（不是零个或两个或更多）Hamming 球，从而保证成功解码.

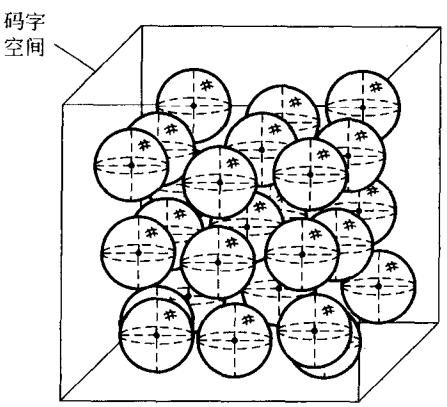


图 12.6　典型输出的 Hamming 球包围的为二元对称信道随机选取的码字. 图 12.5 给出一个单个码字的局部.

小的重叠出现的条件是什么？为了解这一点，我们需要更好地理解信道可能的输出结构. 从对一组独立同分布随机变量 $(X_1,\cdots,X_n)$ 的 2^{nR} 次采样中我们得到了码字. 随机变量以概率 q 取 $X_j=0$，以概率 $1-q$ 取 $X_j=1$. 以 Y_j 表示传送 X_j 通过二元对称信道得到的结果. 典型序列定理意味着对 $(Y_1,\cdots,Y_n)$ 的典型值大约有 $2^{nH(Y)}$ 个，其中 Y 与每个 Y_j 的分布相同. 而且，每个这样的典型输出值具有大致相同的概率.

现在，若从一百万大小的空间中均匀采样一百次，则不大可能会出现任何重复. 事实上，即使采样十万次出现重复的次数都很小，直到一百万次采样之后，重复的次数才开始相对于样本大小变得较大. 类似地，半径 np 的 Hamming 球之间的重叠直到所有球的元素加起来接近实际采样空间的大小 $2^{nH(Y)}$ 才开始变大. 因为

每个球大致包含 $2^{nH(p)}$ 个元素，这意味着在

$$2^{nR} \times 2^{nH(p)} < 2^{nH(Y)} \tag{12.62}$$

条件下，很容易获得一个好的纠错码. 这对应

$$R < H(Y) - H(p) \tag{12.63}$$

的条件. 这里熵 $H(Y)$ 依赖于为 X_j 选择的先验分布 $(q, 1-q)$. 为使比率尽可能大，我们使 $H(Y)$ 最大化. 简单计算表明，采用相应于 $q=1/2$ 的均匀先验分布可以达到最大，此时 $H(Y)=1$，从而小于 $1-H(p)$ 的任意比率 R 都是可以达到的.

我们已经概述了，可以通过二元对称信道以任意直到 $1-H(p)$ 的比率可靠传送信息的证明. 证明相当简要，但包含了甚至是量子情形下严格证明所需要的许多关键思想. 实际上我们所证明的比率也是通过二元对称信道所可能的最快比率；任何超过 $1-H(p)$ 比率的，无论怎样选择这些码字，Hamming 球将重叠太多，以致无法确定发送的是哪个码字！ 因此，$1-H(p)$ 是二元对称信道的容量.

随机编码对二元对称信道达到高比率编码的实用性如何？ 虽然实用随机码可以以高概率运行在接近容量的比率. 遗憾的是，该过程存在一个主要的困难. 为进行编码和解码，发送方和接收方(Alice 和 Bob)必须在进行这些任务时首先采用一致的策略. 对随机码的情况，这意味着 Alice 必须发送给 Bob 她的所有随机码字的清单. 做这件事，Alice 和 Bob 花费的通信要远高于他们从带噪声信道所得到的. 显然，对很多的应用，这是不希望的. 随机编码方法只是证明高比率码存在性的一种方法，而并不是实用的构造方法. 在大量的实际应用中，我们希望的是达到接近信道容量的比率而不引入 Alice 和 Bob 之间不可承受的附加通信开销. 即使对经典带噪声信道，经过数十年的巨大努力，最近才发现构造这样的码方法. 对带噪声量子信道构造类似的编码是一个有趣的公开问题.

2. Shannon 带噪声信道编码定理

Shannon 带噪声信道编码定理把对二元对称信道的容量结果推广到离散无记忆信道的情形. 这类信道具有有限的输入字母表 $\mathcal{I}$ 和有限的输出字母表 $\mathcal{O}$. 对二元对称信道，$\mathcal{I}=\mathcal{O}=\{0,1\}$. 信道的作用由一组条件概率 $p(y|x)$ 所描述，其中 $x\in\mathcal{I}$ 而 $y\in\mathcal{O}$. 它们代表给定输入是 x 的条件下，从信道输出不同 y 的概率，并满足规则

$$p(y \mid x) \geqslant 0 \tag{12.64}$$

$$\sum_y p(y \mid x) = 1, \text{对所有的 } x \tag{12.65}$$

信道无记忆是指每次使用信道时它的作用都相同，并且不同的使用之间是独立的. 我们用符号 $\mathcal{N}$ 来表示经典带噪声信道.

当然，许多有意义的通信信道并不是离散无记忆信道，如我们前面给出的电话线的例子，就具有连续的输入和输出. 更一般的信道可能比离散无记忆信道理解起

来更困难，但隐含的思想是一致的，所以我们推荐读者参考本章末的“历史和进一步阅读的材料”所提供的该主题方面的书籍.

让我们给出 Shannon 带噪声信道编码定理的叙述本身. 这里不会给出证明的细节，因为下节要证明量子信道更一般的结果. 但经典结果的叙述是必要的. 首先，我们需要使可靠信息传输的概念更确切. 基本思想如图 12.7 所示. 第一阶段，Alice 从 2^{nR} 个可能的消息中产生一个消息 M，并用映射 C^n：$\{1,\cdots,2^{nR}\}\to\mathscr{I}^n$ 进行编码. 该映射为 Alice 的每条消息分配一个输入串，该串通过信道的 n 次使用被传给 Bob. Bob 对信道的输出用映射 D^n：$\mathscr{O}^n\to\{1,\cdots,2^{nR}\}$ 进行解码，输出映射为信道的每个可能输出分配一个消息. 对于给定的编码-解码对，*差错概率*定义为所有消息 M 上信道解码输出 $D(Y)$ 不等于消息 M 的最大概率：

$$p(C^n,D^n)\equiv\max_M p(D^n(Y)\neq M\mid X=C^n(M)) \tag{12.66}$$

我们称一个比率 R 是可达到的，如果这样一系列的编码-解码对 (C^n,D^n) 是存在的，并且满足附加的要求，当 $n\to\infty$，$p(C^n,D^n)\to 0$. 一个给定的带噪声信道 $\mathcal{N}$ 的容量 $C(\mathcal{N})$ 定义为信道可达到的比率的上确界.

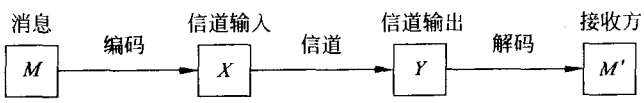


图 12.7　经典消息的带噪声编码问题. 我们要求 2^{nR} 个可能消息中的每一个消息都能以高概率无差错地通过信道.

如何直接计算信道容量完全不显然——直接根据定义的计算将面临在一大类（无穷多）可能的编码和解码方法上取上确界，因而看来是并非特别有用的办法. Shannon 带噪声信道编码定理大大简化了容量的计算，在一个简单和良好定义的很多情况下，可以将这个任务归结为精确求解的优化问题. 该优化问题即使没有精确解，在数值计算上也很容易.

定理 12.7（Shannon 带噪声信道编码定理）　对一个带噪声信道 $\mathcal{N}$，其容量由

$$C(\mathcal{N})=\max_{p(x)} H(X;Y) \tag{12.67}$$

给出，其中最大是在 X 的一次使用的所有分布 $p(x)$ 上取的，Y 是信道输出端得到的相应随机变量.

作为带噪声信道编码定理的例子，考虑以概率 p 翻转比特，且具有输入概率 $p(0)=q$，$p(1)=1-q$ 的二元对称信道，我们有

$$H(X;Y)=H(Y)-H(Y\mid X) \tag{12.68}$$

$$=H(Y)-\sum_x p(x)H(Y\mid X=x) \tag{12.69}$$

而对每个 x，$H(Y|X=x)=H(p)$，故 $H(X;Y)=H(Y)-H(p)$. 该式在选择 $q=$

1/2 时达到最大，故 $H(Y)=1$ 并且进而据 Shannon 带噪声信道编码定理，$C(\mathcal{N})=1-H(p)$. 这正如我们前面对二元对称信道容量直观计算所推导的.

练习 12.9 擦除性信道具有两个输入 0 和 1，三个输出 0，1 和 e. 输入不变的概率为 $1-p$. 输入擦除的概率为 p，并被替换为 e.

(1) 证明擦除性信道的容量为 $1-p$.

(2) 证明擦除性信道的容量比二元对称信道的容量大，直观上该结果为什么是合理的?

练习 12.10 设 $\mathcal{N}_1$ 和 $\mathcal{N}_2$ 是两个离散无记忆信道，满足 $\mathcal{N}_2$ 的输入字母表与 $\mathcal{N}_1$ 的输出字母表相同，证明

$$C(\mathcal{N}_2\circ\mathcal{N}_1)\leqslant\min(C(\mathcal{N}_1),C(\mathcal{N}_2)) \tag{12.70}$$

给出不等式是严格的一个例子.

上述带噪声信道编码定理的一个略显奇怪之处在于，经典信源的概念根本没有出现. 回忆较早时我们把经典信源定义为独立同分布随机变量的序列，就可以用一个有趣的方式把这个信源的概念同带噪声信道编码定理结合起来，而得到所谓信源-信道编码定理. 基本思想如图 12.8 所示. 一个具有熵率 $H(X)$ 产生信息. 由 Shannon 无噪声信道编码定理，可以对从该信源产生的信息进行压缩使得仅用 $nH(X)$ 比特就可以描述这些信息；这个步骤有时称为信源编码. 信源被压缩后的输出接着被用作带噪声信道的输入消息. 为以小于容量的比率传送，需要使用信道 $nH(X)/R$ 次，这样压缩数据才能可靠地传输到接收方. 接收方接下来对其进行解压缩，以恢复从信源而来的原始输出.

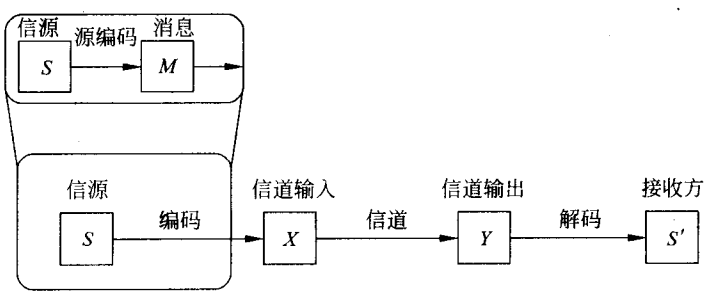


图 12.8 经典信源的带噪声编码问题，有时称为信源-编码模型.

读者也许会问，是否存在从信源经过带噪声信道的更好传送方案，能够比这种压缩-编码和解码-解压缩的两阶段方法做得更有效? 事实上，情况并非如此，这里描述的信源-信道编码实际上是最优的，不过这个事实的证明超出本书的范围；请参考本章末“历史和进一步阅读的材料”中更多细节.

12.3.2　带噪声量子信道上的通信

假设 Alice 和 Bob 使用带噪声量子通信信道进行通信而不是使用带噪声经典通信信道，更确切地，Alice 有某个消息 M 希望发送给 Bob. 正如对经典情况所做的，她要对消息进行编码，但现在是用量子状态进行编码，还要经过带噪声量子信道传送. 采用正确的编码方式，我们希望 Bob 能够以低的错误概率确定 Alice 的消息是什么. 进而，我们要求 Alice 向 Bob 发送信息的比率越高越好. 换言之，我们的目的是得到一个计算带噪声量子信道的经典信息容量的方法. 这个问题尚未得到完全解决，但已取得相当大的进展，本节就来考察这方面的进展.

已经知道的是如何计算一个信道 ε 的容量，假定 Alice 把她的消息用积状态 $\rho_1\otimes\rho_2\otimes\cdots$ 编码，其中每个 $\rho_1,\rho_2,\cdots$ 都是信道 ε 的输入. 我们把带有这个限制的容量称为积状态容量，并将其记作 $C^{(1)}(\varepsilon)$，以表示输入状态不能是信道的两次或更多次使用的纠缠. 注意 Alice 和 Bob 之间的这种受限通信模型其实允许 Bob 使用与信道的多次使用纠缠的测量来进行解码；事实上，这是实质性的. 惟一的(而且是遗憾的)限制是 Alice 只能制备积状态输入. 许多学者相信，但尚未证明，允许纠缠并不增加容量. 允许我们计算积状态容量的结果是 Holevo-Schumacher-Westmoreland(HSW)定理，其名称来源于发现者. 正如 Shannon 带噪声信道定理对于经典的带噪声信道那样，HSW 定理为特定的带噪声信道 ε 的积状态容量提供了有效的计算方法，在某些情况下甚至允许导出精确的表达式.

定理 12.8(Holevo-Schumacher-Westmoreland(HSW)定理)　设 ε 是一个保迹量子运算，定义

$$\chi(\varepsilon)\equiv\max_{\{p_j,\rho_j\}}\left[S\left(\varepsilon\left(\sum_j p_j\rho_j\right)\right)-\sum_j p_j S(\varepsilon(\rho_j))\right] \tag{12.71}$$

其中最大值是在信道所有可能的输入状态 ρ_j 的系综 $\{p_j,\rho_j\}$ 上取的，则 $\chi(\varepsilon)$ 就是信道 ε 的积状态容量，即 $\chi(\varepsilon)=C^{(1)}(\varepsilon)$.

潜在地，在式(12.71)中的最大值可能要在一个无界集合上取. 实际上，我们利用下面练习的结果，将最大值限制在至多包含 d^2 个元的纯态系综上，其中 d 是信道输入的维数.

练习 12.11　证明表达式(12.71)中的最大值可以在纯态的系综上达到，进而证明最多只需要考虑 d^2 个纯态的系综，其中 d 是信道输入的维数.

HSW 定理的证明包含几个不同的思想. 最容易理解该证明的方式是，将证明分成小的部分再放在一起得到 HSW 定理.

1. 随机编码

设 ρ_j 是信道 ε 的一组可能的输入且令 $\sigma_j\equiv\varepsilon(\rho_j)$ 为相应的输出. 我们将发展类

似于对二元对称信道所述的随机编码技术，以允许 Alice 和 Bob 使用状态 ρ_j 的积构成的码字进行通信. 令 p_j 是在指标 j 上的一个概率分布，即先验分布. Alice 想把从集合 $\{1,\cdots,2^{nR}\}$ 中选出的一个消息 M 发给 Bob. 她为每个消息 M 关联一个码字 $\rho_{M_1}\otimes\rho_{M_2}\otimes\cdots\otimes\rho_{M_n}$，其中 $M_1,\cdots,M_n$ 为从指标集 $\{j\}$ 中选取（$M_1,\cdots,M_n$ 并不意味着是 M 的十进制表示或其他类似的表示）. 对每个消息 M，Alice 通过对分布 $\{p_j\}$ 的采样来选择 M_1，类似地选择 $M_2,\cdots,M_n$. 这就完成了码字的确定. 稍作符号上的混用，我们书写 $\rho_M\equiv\rho_{M_1}\otimes\cdots\otimes\rho_{M_n}$. 相应的输出状态直接用 σ 代替 ρ 来记，故例如我们有 $\sigma_{M_1}=\varepsilon(\rho_{M_1})$ 且 $\sigma_M=\varepsilon^{\otimes n}(\rho_M)$.

当 Bob 收到特定状态 σ_M（对应 Alice 希望传送的消息 M）时，他进行测量以便确定消息是什么. 因为我们只对测量的统计结果感兴趣，而不对 Bob 系统测量后的状态感兴趣，所以只要用 POVM 形式描述该测量就足够了. 设对每个可能的消息 M，Bob 都有一个相应的 POVM 元 E_M. Bob 还可能有不对应 Alice 发送的任何消息的一个（或更多）POVM 元；显然所有这些元可以合成一个满足 $E_0=I-\sum_{M\neq 0}E_M$ 的单个的 POVM 元 E_0. Bob 成功识别 M 的概率是 $\mathrm{tr}(\sigma_M E_M)$，从而对消息 M 出错的概率是 $p_M^e\equiv 1-\mathrm{tr}(\sigma_M E_M)$.

我们希望证明的是，存在高比率编码使得出错概率 p_M^e 对所有消息 M 都是小的. 为此，采用一个 Shannon 对经典问题引入的与直观不符而且相当巧妙的技巧. 想像 Alice 通过从集合 $\{1,\cdots,2^{nR}\}$ 中均匀抽取产生消息 M，分析平均出错概率

$$p_{\mathrm{av}}\equiv\frac{\sum_M p_M^e}{2^{nR}}=\frac{\sum_M(1-\mathrm{tr}(\sigma_M E_M))}{2^{nR}} \tag{12.72}$$

证明的第一步是，证明存在高比率编码，满足随 n 变大 p_{av} 趋于 0. 完成这步之后，我们将用 Shannon 的办法来证明，这意味着存在具有本质上相同比率的编码使得对所有的 M，p_M^e 接近于 0. 我们从构造一个很好的（尽管不一定是最优）表示 Bob 解码信道输出 σ_M 方法的 POVM $\{E_M\}$ 出发. 构造的关键思想，正如对经典的二元对称信道那样，是典型性的思想.

令 $\varepsilon>0$. 定义 $\bar{\sigma}\equiv\sum_j p_j\sigma_j$，并令 P 为到 ε 典型子空间 $\bar{\sigma}^{\otimes n}$ 上的投影. 根据典型序列定理可知，对任意 $\delta>0$ 和充分大的 n，有

$$\mathrm{tr}(\bar{\sigma}^{\otimes n}(I-P))\leqslant\delta \tag{12.73}$$

对给定的消息 M，根据典型的 σ_M 约是 np_1 个 ρ_1 的备份，np_2 个 ρ_2 的备份等的张量积. 我们还要定义对 σ_M 的 ε 典型子空间的概念，定义 $\bar{S}\equiv\sum_j p_j S(\sigma_j)$. 设 σ_j 具有谱分解 $\sum_k\lambda_k^j|e_k^j\rangle\langle e_k^j|$，于是

$$\sigma_M=\sum_K\lambda_K^M|E_K^M\rangle\langle E_K^M| \tag{12.74}$$

其中 $K=(K_1,\cdots,K_n)$，并且为方便起见，定义 $\lambda_K^M\equiv\lambda_{K_1}^{M_1}\lambda_{K_2}^{M_2}\cdots\lambda_{K_n}^{M_n}$ 和 $|E_K^M\rangle\equiv|e_{K_1}^{M_1}\rangle|e_{K_2}^{M_2}\rangle\cdots|e_{K_n}^{M_n}\rangle$，定义 P_M 为到由所有满足

$$\left|\frac{1}{n}\log\frac{1}{\lambda_K^M}-\bar{S}\right|\leqslant\varepsilon \tag{12.75}$$

的 $|E_K^M\rangle$ 张成的空间上的投影(用 T_M 记所有使该条件满足的 K). 类似于典型序列定理的证明方法，大数律蕴含对任意 $\delta>0$ 和充分大的 n，有 $E(\mathrm{tr}(\sigma_M P_M))\geqslant 1-\delta$，其中期望是对从随机编码产生的码字 ρ_M(对固定的消息 M)上的分布取的，故对每个 M，

$$E[\mathrm{tr}(\sigma_M(I-P_M))]\leqslant\delta \tag{12.76}$$

同时注意由定义式(12.75) P_M 所投影到的子空间的维数至多为 $2^{n(\bar{S}+\varepsilon)}$，故

$$E(\mathrm{tr}(P_M))\leqslant 2^{n(\bar{S}+\varepsilon)} \tag{12.77}$$

现在用典型性概念来定义 Bob 的解码 POVM，定义

$$E_M\equiv\Big(\sum_{M'}PP_{M'}P\Big)^{-1/2}PP_MP\Big(\sum_{M'}PP_{M'}P\Big)^{-1/2} \tag{12.78}$$

其中 $A^{-1/2}$ 表示 $A^{1/2}$ 的广义逆，即在 A 的支集为 $A^{1/2}$ 的逆而其他为 0. 易知 $\sum_M E_M\leqslant I$. 并且我们可以定义另一个半正定算子 $E_0\equiv I-\sum_M E_M$，以将 POVM 补充完备. 该构造在直观上类似于已描述的二元对称信道的解码方法. 特别地，除去小的修正，E_M 等同于 P_M，而且 Bob 的 $\{E_M\}$ 测量本质上对应于检查信道的输出是否落在 P_M 所投影的空间上；该投影所到达的空间可以视为类似于半径为 np 的围绕对二元对称信道所用的码字的 Hamming 球.

随机编码有效性证明的主要技术部分在于获得平均出错概率 p_{av} 的一个上界，具体过程在盒子 12.5 中给出. 其结果是

$$p_{\mathrm{av}}\leqslant\frac{1}{2^{nR}}\sum_M\Big[3\mathrm{tr}(\sigma_M(I-P))+\sum_{M'\neq M}\mathrm{tr}(P\sigma_M PP_{M'})+\mathrm{tr}(\sigma_M(I-P_M))\Big] \tag{12.79}$$

量 p_{av} 是针对码字的特别选取定义的. 我们要计算该量在所有随机码上的期望. 根据构造 $E(\sigma_M)=\bar{\sigma}^{\otimes n}$，且当 $M'\neq M$ 时 σ_M 与 $P_{M'}$ 是独立的，故得到

$$\begin{aligned}E(p_{\mathrm{av}})\leqslant&3\mathrm{tr}(\bar{\sigma}^{\otimes n}(I-P))+(2^{nR}-1)\mathrm{tr}(P\bar{\sigma}^{\otimes n}PE(P_1))+\\&E(\mathrm{tr}(\sigma_1(I-P_1)))\end{aligned} \tag{12.80}$$

代入式(12.73)和式(12.76)，得

$$E(p_{\mathrm{av}})\leqslant 4\delta+(2^{nR}-1)\mathrm{tr}(P\bar{\sigma}^{\otimes n}PE(P_1)) \tag{12.81}$$

而 $P\bar{\sigma}^{\otimes n}P\leqslant 2^{-n(S(\bar{\sigma})-\varepsilon)}I$，且由式(12.77)有 $E(\mathrm{tr}(P_1))\leqslant 2^{n(\bar{S}+\varepsilon)}$，从而得

$$E(p_{\mathrm{av}})\leqslant 4\delta+(2^{nR}-1)2^{-n(S(\bar{\sigma})-\bar{S}-2\varepsilon)} \tag{12.82}$$

在条件 $R<S(\bar{\sigma})-\bar{S}$ 条件下，可知 $E(p_{\mathrm{av}})\to 0$ 当 $n\to\infty$. 事实上，通过选择系综 $\{p_j,\rho_j\}$ 以到达式(12.71)中的最大值，我们看到只要 $R<\chi(\varepsilon)$ 这必然是真的. 因此

必存在一系列的比率为 R 的编码，使得随编码的块尺寸 n 的增大 $p_{\text{av}}\to 0$. 于是对任意固定的 $\epsilon>0$(注意这里用 ϵ 的新含义取代原来的，原来的已不再需要)和充分大的 n，有

$$p_{\text{av}}=\frac{\sum_M p_M^{\epsilon}}{2^{nR}}<\epsilon \tag{12.83}$$

显然为保证这是真的，至少一半的消息 M 必须满足 $p_M^{\epsilon}<2\epsilon$. 我们通过从满足比率 R 和 $p_{\text{av}}<\epsilon$ 的编码中删除一半的码字(具有高 p_M^{ϵ} 的码字)来构造一个新的编码，得到有 $2^{nR}/2=2^{n(R-1/n)}$ 个码字，且对所有消息 M 满足 $p_M^{\epsilon}<2\epsilon$ 的新编码. 显然这个编码也具有渐近比率 R，且当 n 变大时，差错概率对所有的码字可以做到任意小，而不仅仅为平均的小.

盒子 12.5 HSW 定理：差错估计

HSW 定理证明中最复杂的部分是得到 p_{av} 的一个估计. 这里给出这个过程细节的概要；省略的步骤作为练习留待读者补充. 定义 $|\tilde{E}_K^M\rangle\equiv P|E_K^M\rangle$，则

$$E_M=\Big(\sum_{M'}\sum_{K\in T_{M'}}|\tilde{E}_K^{M'}\rangle\langle\tilde{E}_K^{M'}|\Big)^{-1/2}\sum_{K\in T_M}|\tilde{E}_K^M\rangle\langle\tilde{E}_K^M|\cdot \Big(\sum_{M'}\sum_{K\in T_{M'}}|\tilde{E}_K^{M'}\rangle\langle\tilde{E}_K^{M'}|\Big)^{-1/2} \tag{12.84}$$

若定义

$$\alpha_{(M,K),(M',K')}\equiv\langle\tilde{E}_K^M|\Big(\sum_{M'}\sum_{K''\in T_{M'}}|\tilde{E}_{K''}^{M'}\rangle\langle\tilde{E}_{K''}^{M'}|\Big)^{-1/2}|\tilde{E}_{K'}^{M'}\rangle \tag{12.85}$$

就可将平均差错概率写作

$$p_{\text{av}}=\frac{1}{2^{nR}}\sum_M\Big[1-\sum_K\sum_{K'\in T_M}\lambda_K^M|\alpha_{(M,K),(M',K')}|^2\Big] \tag{12.86}$$

利用 $\sum_K\lambda_K^M=1$ 并略去非正项，可得

$$p_{\text{av}}\leqslant\frac{1}{2^{nR}}\sum_M\Big[\sum_{K\in T_M}\lambda_K^M(1-\alpha^2_{(M,K),(M',K')})+\sum_{K\notin T_M}\lambda_K^M\Big] \tag{12.87}$$

定义矩阵 Γ，其元为 $\gamma_{(M,K),(M',K')}\equiv\langle\tilde{E}_K^M|\tilde{E}_{K'}^{M'}\rangle$，其中指标满足 $K\in T_M$ 和 $K'\in T_{M'}$. 在根据这些指标的约定所定义的矩阵空间中展开讨论是方便的. 令 E 为关于这些指标的单位阵，并用 sp(spur 的含义)表示关于这些指标的迹运算. 计算表明 $\Gamma^{1/2}=[\alpha_{(M,K),(M',K')}]$，从而可知 $\alpha^2_{(M,K),(M,K)}\leqslant\gamma_{(M,K),(M,K)}\leqslant 1$. 根据当 $0\leqslant x\leqslant 1$ 时 $1-x^2=(1+x)(1-x)\leqslant 2(1-x)$ 的事实，并结合式(12.87)，可得

$$p_{\text{av}}\leqslant\frac{1}{2^{nR}}\sum_M\Big[2\sum_{K\in T_M}\lambda_K^M(1-\alpha_{(M,K),(M,K)})+\sum_{K\notin T_M}\lambda_K^M\Big] \tag{12.88}$$

定义对角阵 $\Lambda\equiv\mathrm{diag}(\lambda_K^M)$，并注意

$$2(E-\Gamma^{1/2})=(E-\Gamma^{1/2})^2+(E-\Gamma) \tag{12.89}$$

$$=(E-\Gamma)^2(E+\Gamma^{1/2})^{-2}+(E-\Gamma) \tag{12.90}$$

$$\leqslant(E-\Gamma)^2+(E-\Gamma) \tag{12.91}$$

因此

$$2\sum_M\sum_{K\in T_M}\lambda_K^M(1-\alpha_{(M,K),(M,K)})=2\mathrm{sp}(\Lambda(E-\Gamma^{1/2})) \tag{12.92}$$

$$\leqslant\mathrm{sp}(\Lambda(E-\Gamma)^2)+\mathrm{sp}(\Lambda(E-\Gamma)) \tag{12.93}$$

计算右边项的迹，代入式(12.88)，经过简单计算得到

$$p_{\mathrm{av}}\leqslant\frac{1}{2^{nR}}\sum_M\Big[\sum_K\lambda_K^M\Big(2-2\gamma_{(M,K),(M,K)}+\sum_{K'\neq K}|\gamma_{(M,K),(M,K')}|^2+$$

$$\sum_{M'\neq M,K'\in T_{M'}}|\gamma_{(M,K),(M',K')}|^2\Big)+\sum_{K\notin T_M}\lambda_K^M\Big] \tag{12.94}$$

代入定义并经简单计算得出

$$p_{\mathrm{av}}\leqslant\frac{1}{2^{nR}}\sum_M\big[2\mathrm{tr}(\sigma_M(I-P))+\mathrm{tr}(\sigma_M(I-P)P_M(I-P))+$$

$$\sum_{M'\neq M}\mathrm{tr}(P\sigma_MPP_{M'})+\mathrm{tr}(\sigma_M(I-P_M))\big] \tag{12.95}$$

第二项小于 $\mathrm{tr}(\sigma_M(I-P))$，这给定所期望的差错估计，即式(12.79).

总结起来，我们证明对任意小于式(12.71)所定义的 $\chi(\varepsilon)$ 的比率 R，总存在使用积状态输入的编码，使得经过信道 ε 的传输可以以比率 R 进行. 我们的证明和 Shannon 经典带噪声信道编码定理的随机编码证明具有同样的不足，即它没有提供进行编码的构造过程，不过至少它说明了编码的存在，这个码的比率最大为容量.

2. 上界的证明

设 R 大于式(12.71)所定义的 $\chi(\varepsilon)$，证明 Alice 不可能通过信道 ε 以此比率向 Bob 发送信息. 我们的总体策略是，想象 Alice 均匀随机地从集合 $\{1,\cdots,2^{nR}\}$ 中产生消息 M，并证明她的平均出错概率必定有大于零的下界，故而最大出错概率必定也具有大于零的下界.

设 Alice 把消息 M 编码为 $\rho_M=\rho_1^M\otimes\cdots\rho_n^M$，而相应的输出用 σ 代替 ρ. Bob 用 POVM $\{E_M\}$ 进行解码. 不失一般性，假设它对每个消息 M 包含一个元 E_M，还可能包含一个额外的元 E_0 以保证完备性关系 $\sum_M E_M=I$ 得到满足，这就导出平均差错概率为

$$p_{\mathrm{av}}=\frac{\sum_M(1-\mathrm{tr}(\sigma_ME_M))}{2^{nR}} \tag{12.96}$$

由练习 12.3 可知 $R<\log d$，其中 d 为信道输入的维数，因此 POVM $\{E_M\}$ 最多包含 d^n+1 个元. 根据 Fano 不等式，可知

$$H(p_{\text{av}})+p_{\text{av}}\log d^n \geqslant H(M\mid Y) \tag{12.97}$$

其中 Y 是 Bob 解码的测量结果，于是

$$\begin{aligned} np_{\text{av}}\log d &\geqslant H(M)-H(M:Y)-H(p_{\text{av}}) \\ &= nR-H(M:Y)-H(p_{\text{av}}) \end{aligned} \tag{12.98}$$

先应用 Holevo 界，再利用熵的次可加性，可得

$$H(M:Y)\leqslant S(\bar{\sigma})-\sum_M \frac{S(\sigma_1^M\otimes\cdots\otimes\sigma_n^M)}{2^{nR}} \tag{12.99}$$

$$\leqslant \sum_{j=1}^{n}\left(S(\bar{\sigma}^j)-\sum_M \frac{S(\sigma_j^M)}{2^{nR}}\right) \tag{12.100}$$

其中 $\bar{\sigma}^j\equiv\sum_M \sigma_j^M/2^{nR}$. 不等式右边和中间 n 项的每一项都不大于式(12.71)所定义的 $\chi(\varepsilon)$，于是

$$H(M:Y)\leqslant n\chi(\varepsilon) \tag{12.101}$$

代入式(12.98)，导出 $np_{\text{av}}\log d\geqslant n(R-\chi(\varepsilon))-H(p_{\text{av}})$，从而当 n 增大到极限情况时，得到

$$p_{\text{av}}\geqslant\frac{(R-\chi(\varepsilon))}{\log d} \tag{12.102}$$

在 $R>\chi(\varepsilon)$ 时，该式大于 0，从而完成 $\chi(\varepsilon)$ 是积状态容量的一个上界的证明.

3. 例子

HSW 定理的一个有趣推论是，任意量子信道 ε 都可以用来传输经典信息，只要该信道不只是个常数. 因为若信道不是常数，则存在纯态 $|\psi\rangle$ 和 $|\varphi\rangle$，使得 $\varepsilon(|\psi\rangle\langle\psi|)\neq\varepsilon(|\varphi\rangle\langle\varphi|)$. 把由这两个状态以等概率 1/2 组成的系综代入针对积状态容量的表达式(12.71)，可知

$$\begin{aligned} C^{(1)}(\varepsilon)\geqslant & S\left(\frac{\varepsilon(|\psi\rangle\langle\psi|)+\varepsilon(|\varphi\rangle\langle\varphi|)}{2}\right)- \\ & \frac{1}{2}\varepsilon(|\psi\rangle\langle\psi|)-\frac{1}{2}\varepsilon(|\varphi\rangle\langle\varphi|)>0 \end{aligned} \tag{12.103}$$

其中第二个表达式来源于 11.3.5 节建立的熵的严格凹性.

让我们看一个积状态容量可以精确计算的简单例子，即具有参数 p 的去极化信道的情形. 令 $\{p_j,|\psi_j\rangle\}$ 为量子状态的系综，则我们有一个量子状态

$$\varepsilon(|\psi_j\rangle\langle\psi_j|)=p|\psi_j\rangle\langle\psi_j|+(1-p)\frac{I}{2} \tag{12.104}$$

它具有特征值 $(1+p)/2$ 和 $(1-p)/2$. 由此可导出

$$S(\varepsilon(|\psi_j\rangle\langle\psi_j|)) = H\left(\frac{1+p}{2}\right) \tag{12.105}$$

这完全不依赖于$|\psi_j\rangle$. 因此只要直接选择$|\psi_j\rangle$构成单个量子比特状态空间的标准正交基底(如$|0\rangle$和$|1\rangle$),给出 1 比特的熵值,以及带有参数 p 的去极化信道的积状态容量

$$C(\varepsilon) = 1 - H\left(\frac{1+p}{2}\right) \tag{12.106}$$

式(12.71)中的最大值可以通过最大化熵 $S(\sum_j \varepsilon(|\psi_j\rangle\langle\psi_j|))$ 达到.

练习 12.12　利用 HSW 定理的证明寻找 Shannon 噪声信道编码定理的一种证明,在可能之处简化证明过程。

12.4　带噪声量子信道上的量子信息

带噪声量子信道能够可靠传输多少量子信息? 我们对这个量子信道容量问题的认识不及对通过带噪声量子信道发送经典信息的容量问题的认识. 下面介绍一些已被用于研究量子信道传送量子信息的容量的信息论工具,最著名的包括 Fano 不等式(盒子 12.2)、数据处理不等式(11.2.4)和单一界(练习 10.21)在量子信息论中的对应.

关于量子数据压缩,我们的观点是把量子信源视为处于混合态 ρ 的与别的量子系统纠缠的量子系统,而由量子信息通过量子运算 ε 传输的可靠性测度是纠缠忠实度 $F(\rho,\varepsilon)$. 如第 9 章那样,有必要引入标号 Q 表示 ρ 所在的系统,标号 R 表示初始纯化 Q 的参考系统. 在这个框架下,纠缠忠实度是 Q 和 R 之间的纠缠在系统 Q 上的 ε 作用下保持程度的一种度量.

12.4.1　熵交换与量子 Fano 不等式

量子运算应用到量子系统 Q 的状态 ρ 上会引起多少噪声? 一种测量方法是扩展到系统 RQ. 它的开始状态是纯态,在量子运算作用下变成混合态. 定义运算 ε 在输入 ρ 上的熵交换为

$$S(\rho,\varepsilon) \equiv S(R',Q') \tag{12.107}$$

假设量子运算 ε 的作用如第 8 章那样,可通过引入起始处于纯态的环境 E,然后引入 Q 和 E 之间的酉相互作用来描述. 相互作用之后,RQE 的状态是一个纯态,从而 $S(R',Q')=S(E')$,故熵交换也等同于运算 ε 引入到初始为纯态的环境 E 中的熵的量.

注意,熵交换不依赖于 Q 的初态 ρ 纯化到 RQ 的方式. 原因是如练习 2.81

已证明的，Q 到 RQ 的任意两个纯化都由系统 R 上的一个酉运算联系着. R 上的这个酉运算显然与 Q 上的量子运算对易，因此两个不同的纯化所导出的终了状态 $R'Q'$ 由 R 上的酉变换所关联，进而给出熵交换的相同取值. 更进一步，由这些结果可知在模型从处于纯态的 E 出发的条件下，$S(E')$ 不依赖于 ε 的特定环境模型.

基于量子运算的算子和表示可给出熵交换的一个有用的显式公式. 设保迹量子运算 ε 具有运算元 $\{E_j\}$，则如 8.2.3 已说明的，该量子运算的一个酉模型可由 QE 上定义的酉算子 U 给出，U 满足

$$U|\psi\rangle|0\rangle = \sum_i E_i|\psi\rangle|i\rangle \tag{12.108}$$

这里 $|0\rangle$ 是环境的初始状态，$|i\rangle$ 是环境的标准正交基底. 注意应用 ε 之后，E' 的状态为

$$\rho^{E'} = \sum_{i,j} \mathrm{tr}(E_i \rho E_j^\dagger)|i\rangle\langle j| \tag{12.109}$$

即 $\mathrm{tr}(E_i\rho E_j^\dagger)$ 是 E' 在 $|i\rangle$ 基底下的矩阵元素. 于是给定一个带有运算元 $\{E_j\}$ 的量子运算后，可自然地定义具有矩阵元素 $W_{ij} \equiv \mathrm{tr}(E_i\rho E_j^\dagger)$ 的矩阵 W，即 W 是 E' 在适当基底下的矩阵. $\rho^{E'}$ 的这个表示导出了使计算显式化的熵交换公式，

$$S(\rho,\varepsilon) = S(W) \equiv -\mathrm{tr}(W\log W) \tag{12.110}$$

给定量子运算 ε 和状态 ρ，总可以为 ε 选取运算元 $\{F_j\}$，使得 W 是对角的；我们称 W 具有规范形. 为明确这样一组运算元总是存在的，回忆第 8 章的内容：一个量子运算可能有许多不同组的运算元. 特别地，量子运算元 $\{E_i\}$ 和 $\{F_j\}$ 是同一量子运算的运算元，当且仅当 $F_j = \sum_i u_{ji}E_i$，其中 u 是复数上的酉矩阵，并且为使 u 成为方阵，可能需要在 $\{E_i\}$ 或 $\{F_j\}$ 中添加 0 算子. 令 W 为 ε 特别选定的运算元 $\{E_i\}$ 相关联的 w 矩阵. W 是环境密度算子的矩阵表示，故是半正定的，可通过某个酉矩阵 v 对角化，$D = vWv^\dagger$，其中 D 是具有非负项的对角阵. 用方程 $F_j \equiv \sum_i v_{ji}E_i$ 定义算子 F_j，故 F_j 也是 ε 的一组运算元，且给出具有矩阵元

$$\widetilde{W}_{kl} = \mathrm{tr}(F_k\rho F_l^\dagger) = \sum_{mn} v_{km}v_{ln}^* W_{mn} = D_{kl} \tag{12.111}$$

的新 w 矩阵 $\widetilde{W}$. 于是，若关于运算元 $\{F_j\}$ 计算，该 w 矩阵是对角形的. ε 的任何这样的一组运算元都称为 ε 关于输入 ρ 的一个规范表示，后面将看到规范表示对量子纠错具有特殊的重要性.

熵交换的许多重要性质源自第 11 章中熵的性质. 例如，对 d 维空间上的保迹量子运算 ε，在规范表示下研究，立刻会发现 $S(I/d,\varepsilon)=0$ 当且仅当 ε 是一个酉量子运算. 因此，$S(I/d,\varepsilon)$ 可被视为系统整体上出现量子噪声的程度的一种度量. 另一个例子是矩阵 W 对 ρ 是线性的，并由熵的凹性可知 $S(\rho,\varepsilon)$ 对 ρ 是凹的.

由于系统 RQ 总可以被选得至多为 d^2 维，其中 d 是 Q 的维数，易见熵交换有上界 $2\log d$.

练习 12.13　证明熵交换对量子运算 ε 是凹的.

直观上，如果量子信源 Q 包含噪声带来纠缠 RQ 变为混合态，那么从初态 RQ 出发得到的终了态 $R'Q'$ 的忠实度不可能是完全的，而且噪声越大忠实度越差. 12.1.1 节在研究经典信道时就出现了类似的情形. 那里给定输出 Y 条件下，关于信道的输入 X 的不确定性 $H(X|Y)$ 通过 Fano 不等式与从 Y 能够恢复 X 的状态的概率联系起来. 该结果有一个非常有用的量子对应，它将熵交换 $S(\rho,\varepsilon)$ 与纠缠忠实度 $F(\rho,\varepsilon)$ 联系起来.

定理 12.9（量子 Fano 不等式）　令 ρ 为一量子状态而 ε 为一保迹量子运算，那么

$$S(\rho,\varepsilon) \leqslant H(F(\rho,\varepsilon)) + (1-F(\rho,\varepsilon))\log(d^2-1) \tag{12.112}$$

其中 $H(\cdot)$ 是二元 Shannon 熵.

量子 Fano 不等式揭示了一个有趣的直观含义：如果一个过程的熵交换大，那么该过程的纠缠忠实度必然小，显示 R 和 Q 之间的纠缠没有得到很好的保持. 进而，我们注意到在量子 Fano 不等式中熵交换 $S(\rho,\varepsilon)$ 类似于经典信息论中条件熵 $H(X|Y)$ 的角色.

证　为证量子 Fano 不等式，令 $|i\rangle$ 为系统 RQ 的标准正交基底，并使该集合中首个状态满足 $|1\rangle = |RQ\rangle$. 如果我们引入量 $p_i \equiv \langle i|\rho^{R'Q'}|i\rangle$，那么由 11.3.3 节的结果可得

$$S(R',Q') \leqslant H(p_1,\cdots,p_{d^2}) \tag{12.113}$$

其中 $H(p_i)$ 是集合 $\{p_i\}$ 的 Shannon 信息. 简单的代数运算显示

$$H(p_1,\cdots,p_{d^2}) = H(p_1) + (1-p_1)H\left(\frac{p_2}{1-p_1},\cdots,\frac{p_{d^2}}{1-p_1}\right) \tag{12.114}$$

此式并结合 $H\left(\frac{p_2}{1-p_1},\cdots,\frac{p_{d^2}}{1-p_1}\right) \leqslant \log(d^2-1)$ 的事实，以及根据定义 $p_1 = F(\rho,\varepsilon)$，可导出

$$S(\rho,\varepsilon) \leqslant H(F(\rho,\varepsilon)) + (1-F(\rho,\varepsilon))\log(d^2-1) \tag{12.115}$$

即量子 Fano 不等式.

12.4.2　量子数据处理不等式

11.2.4 节中讨论了经典的数据处理不等式. 回忆数据处理不等式断言对于一个 Markov 过程 $X \to Y \to Z$，有

$$H(X) \geqslant H(X:Y) \geqslant H(X:Z) \tag{12.116}$$

当且仅当可以从 Y 恢复随机变量 X 的概率为 1 时取等号. 因此数据处理不等式为纠错的可能性提供了信息论方面的充要条件.

对应该数据处理不等式,有一个可以应用到由量子运算 ε_1 和 ε_2 描述的两阶段量子过程:

$$\rho \xrightarrow{\varepsilon_1} \rho' \xrightarrow{\varepsilon_2} \rho'' \tag{12.117}$$

定义量子相干信息为

$$I(\rho,\varepsilon) \equiv S(\varepsilon(\rho)) - S(\rho,\varepsilon) \tag{12.118}$$

人们猜测(没有证明)相干信息这个量在量子信息论中扮演类似经典信息论中的互信息 $H(X:Y)$ 角色. 这种看法的一个理由是,相干信息满足类似于经典数据处理不等式的量子数据处理不等式.

定理 12.10(量子数据处理不等式) 令 ρ 为一量子状态而 ε_1 和 ε_2 是保迹的量子运算,那么

$$S(\rho) \geqslant I(\rho,\varepsilon_1) \geqslant I(\rho,\varepsilon_2 \circ \varepsilon_1) \tag{12.119}$$

当且仅当能够完全*逆转*运算 ε_1 时,第一个不等式取等号. 完全逆转指存在保迹逆运算 $\mathcal{R}$,使得 $F(\rho,\mathcal{R}\circ\varepsilon)=1$.

与经典数据处理不等式对比显示,相干信息在量子数据处理不等式中的角色等同于互信息在经典数据处理不等式中的角色. 当然,这种粗略的论证不能看成相干信息是经典互信息的正确量子对应物观点的严格论述. 为建立这样的论述,相干信息应该以类似于经典互信息和经典信道容量关系的方式与量子信道容量相联系. 而这种联系尚未建立(参考本章末"历史和进一步阅读的材料"中介绍的部分发展).

量子数据处理不等式中定义的完全可逆性和较熟悉的概念如量子纠错中的概念有什么联系呢? 根据定义,我们说保迹量子运算 ε 在输入 ρ 上是*完全可逆*的,即如果存在保迹量子运算 $\mathcal{R}$,使得

$$F(\rho,\mathcal{R}\circ\varepsilon) = 1 \tag{12.120}$$

但从第 9 章式(9.143)所在的第(4)项可知,量子运算完全可逆当且仅当对 ρ 的支集中的每个状态 $|\psi\rangle$,有

$$(\mathcal{R}\circ\varepsilon)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi| \tag{12.121}$$

这个事实把完全可逆性概念和量子纠错码联系起来. 回忆量子纠错码是有逻辑码字张成的某个更大 Hilbert 空间的一个子空间,为对抗由量子运算 ε 导致的噪声,该量子运算 ε 必须可由保迹逆运算 $\mathcal{R}$ 在对编码中所有状态 $|\psi\rangle$ 成立 $(\mathcal{R}\circ\varepsilon)(|\psi\rangle\langle\psi|)=|\psi\rangle\langle\psi|$ 的意义下使其可逆. 该条件等价于数据处理不等式叙述中的完全可逆性准则,即对支集为编码空间的某个 ρ,$F(\rho,\mathcal{R}\circ\varepsilon)=1$.

证　量子数据处理不等式是采用一个四系统构造证明的：包括我们熟悉的 R 和 Q，起始处于纯态的系统 E_1 和 E_2. 选择它们，使 Q 和 E_1 间的一个酉相互作用产生 ε_1 的动态，Q 和 E_2 间的一个酉相互作用产生 ε_2 的动态. 量子数据处理不等式证明的第一阶段是，应用次可加性不等式 $S(R',E_1')\leqslant S(R')+S(E_1')$ 得到

$$I(\rho,\varepsilon_1)=S(\varepsilon_1(\rho))-S(\rho,\varepsilon_1) \tag{12.122}$$

$$=S(Q')-S(E_1') \tag{12.123}$$

$$=S(R',E_1')-S(E_1') \tag{12.124}$$

$$\leqslant S(R')+S(E_1')-S(E_1')=S(R') \tag{12.125}$$

$$=S(R)=S(Q)=S(\rho) \tag{12.126}$$

数据处理不等式证明的第二部分是应用强次可加性不等式，

$$S(R'',E_1'',E_2'')+S(E_1'')\leqslant S(R'',E_1'')+S(E_1'',E_2'') \tag{12.127}$$

从完整状态 $R''Q''E_1''E_2''$ 的纯性可知

$$S(R'',E_1'',E_2'')=S(Q'') \tag{12.128}$$

系统 R 和 E_1 都不涉及 Q 和 E_2 发生酉相互作用的第二阶段的动态，因此它们的状态在这个阶段中不变：$\rho^{R''E_1''}=\rho^{R'E_1'}$. 但从第一阶段动态之后 RQE_1 的纯性可知

$$S(R'',E_1'')=S(R',E_1')=S(Q') \tag{12.129}$$

现在式(12.127)中剩余的两项为熵交换

$$S(E_1'')=S(E_1')=S(\rho,\varepsilon_1);\quad S(E_1'',E_2'')=S(\rho,\varepsilon_2\circ\varepsilon_1) \tag{12.130}$$

把这些代入式(12.127)，得到

$$S(Q'')+S(\rho,\varepsilon_1)\leqslant S(Q')+S(\rho,\varepsilon_2\circ\varepsilon_1) \tag{12.131}$$

它可以整理为数据处理不等式的第二段，即 $I(\rho,\varepsilon_1)\geqslant I(\rho,\varepsilon_2\circ\varepsilon_1)$.

为完成证明，我们需要证明，当且仅当量子数据处理不等式中第一个不等式取等号，

$$S(\rho)=I(\rho,\varepsilon)=S(\rho')-S(\rho,\varepsilon) \tag{12.132}$$

ε 在输入 ρ 上完全可逆. 为证该条件对可逆性的必要性，设 ε 在输入 ρ 上完全可逆，具有逆运算 $\mathcal{R}$. 从量子数据处理不等式的第二段可知

$$S(\rho')-S(\rho,\varepsilon)\geqslant S(\rho'')-S(\rho,\mathcal{R}\circ\varepsilon) \tag{12.133}$$

从可逆性要求知 $\rho''=\rho$. 而且，从量子 Fano 不等式(12.112)和完全可逆性要求 $F(\rho,\mathcal{R}\circ\varepsilon)=1$ 可知，$S(\rho,\mathcal{R}\circ\varepsilon)=0$. 因此当应用到 $\rho\to\varepsilon(\rho)\to(\mathcal{R}\circ\varepsilon)(\rho)$ 时，量子数据处理不等式可重写为

$$S(\rho')-S(\rho,\varepsilon)\geqslant S(\rho) \tag{12.134}$$

将此与量子数据处理不等式的第一段 $S(\rho)\geqslant S(\rho')-S(\rho,\varepsilon)$ 结合，可推断

$$S(\rho) = S(\rho') - S(\rho,\varepsilon) \tag{12.135}$$

对任意在输入 ρ 上完全可逆的 ε 成立.

下面我们给出一个构造性证明,满足条件

$$S(\rho) = S(\rho') - S(\rho,\varepsilon) \tag{12.136}$$

蕴含量子运算在输入 ρ 上可逆. 注意 $S(\rho)=S(Q)=S(R)=S(R')$, $S(\rho')=S(Q')=S(R',E')$ 且 $S(\rho,\varepsilon)=S(E')$,可以看出 $S(R')+S(E')=S(R',E')$. 根据 11.3.4 节,这等价于 $\rho^{R'E'}=\rho^{R'}\otimes\rho^{E'}$. 设 Q 的初态为 $\rho=\sum_i p_i|i\rangle\langle i|$,我们将该状态纯化到 RQ 中,为 $|RQ\rangle=\sum_i\sqrt{p_i}|i\rangle|i\rangle$,其中第一个系统是 R,第二个系统是 Q. 注意 $\rho^{R'}=\rho^{R}=\sum_i p_i|i\rangle\langle i|$. 进而设对某个标准正交基底集合 $|j\rangle$, $\rho^{E'}=\sum_j q_j|j\rangle\langle j|$,使得

$$\rho^{R'E'}=\sum_{ij}p_iq_j|i\rangle\langle i|\otimes|j\rangle\langle j| \tag{12.137}$$

它具有特征值 $|i\rangle|j\rangle$,故根据 Schmidt 分解,可以把应用了量子运算 ε 之后的整体状态 $R'Q'E'$ 写作

$$|R'Q'E'\rangle=\sum_{ij}\sqrt{p_iq_j}|i\rangle|i,j\rangle|j\rangle \tag{12.138}$$

其中 $|i,j\rangle$ 是系统 Q 状态的某个标准正交基底集合. 按 $P_j\equiv\sum_i|i,j\rangle\langle i,j|$ 定义投影 P_j. 恢复运算的思想是先进行投影 P_j 所描述的测量,以揭示环境的状态 $|j\rangle$,然后作以 j 为条件的酉旋转 U_j,把状态 $|i,j\rangle$ 恢复到 $|i\rangle$: $U_j|i,j\rangle\equiv|i\rangle$. 即 j 是测量结果,而 U_j 是相应的恢复运算. 完整的恢复运算可以写作

$$\mathcal{R}(\sigma)\equiv\sum_j U_jP_j\sigma P_jU_j^\dagger \tag{12.139}$$

根据状态 $|i,j\rangle$ 的正交性,投影 P_j 是正交的,但未必是完备的. 若出现这种情况,为保证量子运算 $\mathcal{R}$ 是保迹的,需要增加一个附加的投影 $\tilde{P}\equiv I-\sum_j P_j$ 到投影集合中以使运算保迹.

逆运算之后,系统 RQE 的终了状态由

$$\sum_j U_jP_j|R'Q'E'\rangle\langle R'Q'E'|P_jU_j^\dagger$$

$$=\sum_j\sum_{i_1i_2}\sqrt{p_{i_1}p_{i_2}}q_j|i_1\rangle\langle i_2|\otimes(U_j|i_1,j\rangle\langle i_2,j|U_j^\dagger)\otimes|j\rangle\langle j| \tag{12.140}$$

$$=\sum_{i_1,i_2}\sqrt{p_{i_1}p_{i_2}}|i_1\rangle\langle i_2|\otimes|i_1\rangle\langle i_2|\otimes\rho^{E'} \tag{12.141}$$

给出. 由此可知 $\rho^{R'Q''}=\rho^{RQ}$，于是 $F(\rho,\mathcal{R}\circ\varepsilon)=1$，即运算 ε 在输入 ρ 上完全可逆，如所欲证. □

这就完成了信息论中针对保迹量子运算的可逆性条件的证明. 该结果的直观含义可以通过如下设想来得到：Q 为量子计算机的存储单元，R 为量子计算机的其他部分，E 为与 Q 相互作用带来噪声的环境. 信息论可逆性条件的最优解释是，噪声发生后的环境 E' 的状态不应该和量子计算其他部分 R' 的状态相关. 用更拟人化的方式表达，当环境通过和 Q 的相互作用学不到关于量子计算机其他部分的任何东西时，纠错是可能精确的.

更具体地，设 Q 是 n 量子比特系统，C 是系统 Q 中的 $[n,k]$ 量子纠错码，具有标准正交码字 $|x\rangle$ 和到码空间上的投影 P. 考虑密度矩阵 $P/2^k$，它可以纯化为 RQ 的一个纯态

$$\frac{1}{\sqrt{2^k}}\sum_x |x\rangle|x\rangle \tag{12.142}$$

假想这个编码能够纠正量子比特的某个子集 Q_1 上的任何错误，那么，特别地，它必能纠正直接把量子比特交换到环境中而代之以某个标准状态的错误. 这种情况下，信息论可逆性条件 $\rho^{R'E'}=\rho^{R'}\otimes\rho^{E'}$ 可写成条件 $\rho^{RQ_1}=\rho^{R}\otimes\rho^{Q_1}$. 因此为使纠错成为可能，参考系统 R 和可以纠错的子系统 Q_1 起始时必须是不相关的！

练习 12.14　证明条件 $\rho^{RQ_1}=\rho^{R}\otimes\rho^{Q_1}$ 对在子系统 Q_1 上可以纠错也是充分的.

量子数据处理不等式证明中用到的推理可用来证明更宽范围的其他类型的不等式. 例如，设有接受量子运算 ε 的一个处于状态 ρ 的量子系统 Q. 数据处理不等式的第一段是从应用熵的次可加不等式到系统 $R'E'$ 得来的，代之，把次可加不等式应用到系统 $Q'E'$，所得到将是

$$\begin{aligned} S(\rho)=S(R)&=S(R')=S(Q',E')\leqslant S(Q')+S(E')\\ &=S(\varepsilon(\rho))+S(\rho,\varepsilon) \end{aligned} \tag{12.143}$$

即

$$\Delta S+S(\rho,\varepsilon)\geqslant 0 \tag{12.144}$$

其中 $\Delta S\equiv S(\varepsilon(\rho))-S(\rho)$ 是由过程 ε 引起的熵交换. 粗略说，该不等式说的是系统的熵交换加上环境的熵变化必为非负，这显然符合热力学第二定律. 它将帮助我们在 12.4.4 节进行量子纠错的热力学分析.

练习 12.15　对两阶段量子过程 $\rho\to\rho'=\varepsilon_1(\rho)\to\rho''=(\varepsilon_2\circ\varepsilon_1)(\rho)$ 应用次可加性和强次可加性不等式的所有可能组合，来推出其他的不等式，在可能的情况下用熵交换和熵 $S(\rho)$，$S(\rho')$，$S(\rho'')$ 表示结果. 当不能用这些量表示结果时，给出仅用 ρ 和 ε_1 的运算元 $\{E_j\}$，ε_2 的运算元 $\{F_k\}$ 的计算公式.

12.4.3 量子单一界

量子纠错的信息论方法可用于证明量子纠错码纠错能力的一个漂亮的估界，即量子单一界. 回忆一个$[n,k,d]$编码使用n量子比特对k量子比特进行编码，并能纠正被定位的多到$d-1$量子比特上的错误(练习 10.45). 量子单一界断言必成立$n-k\geqslant 2(d-1)$. 对比经典单一界，即对$[n,k,d]$经典编码成立$n-k\geqslant d-1$(练习 10.21). 因为量子编码对多到t量子比特的差错进行纠错必须具有至少$2t+1$量子比特，所以$n-k\geqslant 4t$. 因此，对于$k=1$量子比特进行编码且能够纠正量子比特上$t=1$的差错的编码必须满足$n-1\geqslant 4$，即n至少为 5，故第 10 章中描述的 5 量子比特编码是为完成此任务的可能的最小编码.

量子单一界的证明是我们已用于分析量子纠错的信息论技术的扩展. 设编码是与系统Q相关联的2^k维子空间，具有记作$|x\rangle$的标准正交基底. 引入同样具有记作$|x\rangle$的标准正交基底的2^k维参考系统R，并考虑RQ的纠缠状态

$$|RQ\rangle=\frac{1}{\sqrt{2^k}}\sum_x |x\rangle|x\rangle \tag{12.145}$$

我们把Q的n量子比特划分为不相交的三块，分别由$d-1$量子比特组成的Q_1和Q_2，以及由剩余的$n-2(d-1)$量子比特组成的第三块Q_3. 由于编码具有距离d，任意一组被定位的$d-1$量子比特差错可以纠正，从而可能纠正Q_1或Q_2上的差错. 易知R和Q_1必为不相关的，R和Q_1也是不相关的. 基于这个事实，状态$RQ_1Q_2Q_3$的纯性，以及熵的次可加性，有

$$S(R)+S(Q_1)=S(R,Q_1)=S(Q_2,Q_3)\leqslant S(Q_2)+S(Q_3) \tag{12.146}$$

$$S(R)+S(Q_2)=S(R,Q_2)=S(Q_1,Q_3)\leqslant S(Q_1)+S(Q_3) \tag{12.147}$$

将两不等式相加得

$$2S(R)+S(Q_1)+S(Q_2)\leqslant S(Q_1)+S(Q_2)+2S(Q_3) \tag{12.148}$$

经过消去和代入$S(R)=k$，得到$k\leqslant S(Q_3)$. 但Q_3的大小是$n-2(d-1)$量子比特，于是$S(Q_3)\leqslant n-2(d-1)$，从而$k\leqslant n-2(d-1)$，故$2(d-1)\leqslant n-k$，即量子单一界.

作为量子单一界的一个应用例子，考虑去极化信道$\varepsilon(\rho)=p\rho+(1-p)/3(X\rho X+Y\rho Y+Z\rho Z)$，设去极化信道在大数量$n$量子比特上独立作用. 如果$p<3/4$，那么多于 1/4 的量子比特将发生差错，因此任何可以从差错中恢复的编码必有$t>n/4$. 但量子单一节蕴含$n-k\geqslant 4t>n$，且k必为非负，即此时不可能对任何量子比特进行编码. 因此，当$p<3/4$时，量子单一界蕴含去极化信道的量子信息容量为零.

12.4.4　量子纠错，制冷和 Maxwell 妖

量子纠错可视为一类制冷过程，尽管噪声过程的影响倾向于改变系统的熵，可以把量子系统保持在定常熵. 事实上，从这个角度看量子纠错可能甚至相当令人困惑，因为它表面上允许量子系统熵的减小，明显违背了热力学第二定律！为理解为什么没有违背热力学第二定律，我们对量子纠错做出类似盒子 3.5（见《量子计算和量子信息（一）》）中 Maxwell 妖的分析. 量子纠错实质上是 Maxwell 妖的一个特殊类型——我们可以想像一个妖，在量子系统上执行特征测量，然后按照特征测量的结果进行纠错. 正如对经典 Maxwell 妖的分析，根据 Landauer 原理，妖的记忆中对特征的存储会带来热力学代价. 特别地，由于任何记忆都是有限的，为了为新的测量结果预备空间，妖最终需要从记忆中擦除信息. Landauer 原理断言从内存中擦除一比特信息带来整个系统——量子系统，妖和环境——熵的至少一比特的增加.

更确切地，我们可以考虑如图 12.9 所示的四阶段纠错循环：

（1）系统初始状态为 ρ，在带噪声演化的作用下到达状态 ρ'. 在典型的纠错过程中，我们对系统熵增加的情况，即 $S(\rho')>S(\rho)$ 感兴趣，尽管这不是必然的.

（2）妖在状态 ρ' 上执行由算子 $\{M_m\}$ 描述的（特征）测量，以概率 $p_m=\mathrm{tr}(M_m\rho' M_m^\dagger)$ 得到结果 m，这导致测量后状态 $\rho'_m=M_m\rho' M_m^\dagger/p_m$.

（3）妖应用一个酉运算 V_m（恢复运算），产生终了系统状态

$$\rho''_m=V_m\rho'_m V_m^\dagger=\frac{V_m M_m\rho' M_m^\dagger V_m^\dagger}{p_m} \tag{12.149}$$

（4）循环重新开始. 为了保证确实是循环，并且纠错是成功的，我们对每个测量结果 m 都必须有 $\rho''_m=\rho$.

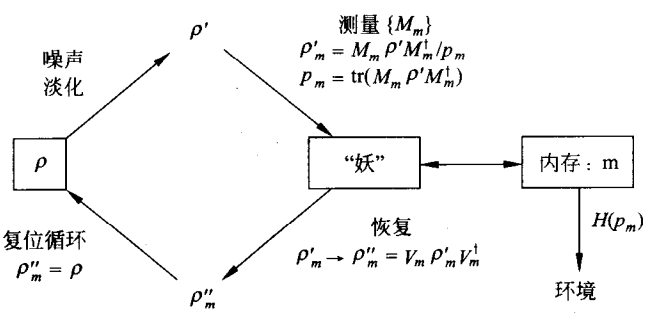


图 12.9　量子纠错循环.

我们现在来说明，在第二和第三阶段——纠错阶段——的任何熵的减少都是以环境中至少与量子纠错过程中熵的减少等量的熵的产生为代价的. 第三阶段后，妖的记忆中仅保存的记录是测量结果 m 的记录. 为了下一循环内存复位，妖必须擦除测量结果的记录，根据 Landauer 原理，这引起环境熵的净增加. 必须擦除的比特数决定于妖用于存储测量结果 m 的表示；根据 Shannon 无噪声信道编码定理，平均而言，至少需要 $H(p_m)$ 比特来存储测量结果，因此当测量记录被擦除时，一个纠错循环平均引起 $H(p_m)$ 比特的熵耗散到环境中.

在量子纠错前量子系统状态为 ρ'，量子纠错后量子系统状态为 ρ，于是纠错引起的系统熵的净变换为 $\Delta S \equiv S(\rho)-S(\rho')$. 与擦除测量记录相关联的附加熵消耗(平均)为 $H(p_m)$，而总的消耗为 $\Delta(S)+H(p_m)$. 我们的目标是估计这个热力学消耗的界，并以此说明热力学第二定律从来都没有被违背. 为此，引入两个记号：令 ε 表示纠错循环第一阶段发生的噪声过程，$\rho \to \rho' = \varepsilon(\rho)$，并令 $\mathcal{R}$ 为表示纠错运算的量子运算，

$$\mathcal{R}(\sigma) \equiv \sum_m V_m M_m \sigma M_m^\dagger V_m^\dagger \tag{12.150}$$

对输入 ρ'，该过程的 w 矩阵具有元 $W_{mn} = \mathrm{tr}(V_m M_m \rho' M_n^\dagger V_n^\dagger)$，因此具有对角元 $W_{mm} = \mathrm{tr}(V_m M_m \rho' M_m^\dagger V_m^\dagger) = \mathrm{tr}(M_m \rho' M_m^\dagger)$，这正是妖测量差错特征时得到的测量结果 m 的概率 p_m. 由定理 11.9 知，W 对角元的熵至少和 W 的熵一样大，故

$$H(p_m) \geqslant S(W) = S(\rho', \mathcal{R}) \tag{12.151}$$

当且仅当算子 $V_m M_m$ 是 $\mathcal{R}$ 关于 ρ' 的规范分解，且使得 W 中非对角元素均为 0 时，取等号. 由式(12.144)可知

$$\Delta S + S(\rho', \mathcal{R}) = S(\rho) - S(\rho') + S(\rho', \mathcal{R}) \geqslant 0 \tag{12.152}$$

把这个结果与式(12.151)结合，可推知 $\Delta S + H(p_m) \geqslant 0$. 但 $\Delta S + H(p_m) \geqslant 0$ 是由纠错过程引起的总熵变，可见纠错只能引起总熵的净增加，由纠错带来系统熵减少的代价是纠错中产生的差错特征被擦除时熵的产生.

练习 12.16 证明当 $\mathcal{R}$ 对输入 ρ 完全纠正 ε 时，不等式

$$S(\rho) - S(\rho') + S(\rho', \mathcal{R}) \geqslant 0 \tag{12.153}$$

实际上必取等号.

12.5 作为物理资源的纠缠

到目前为止，我们对量子信息的研究集中在与经典信息论中考虑的差别不大的资源上，为方便起见，图 12.10 总结了量子和经典形式的许多成果. 量子计算与量子信息的一个喜人之处在于，量子力学还包含了本质上为新类型的资源，这些与经典信息论中传统上视为信息的资源截然不同. 也许其中理解得最好的是量子纠缠，这正是我们要考察的资源.

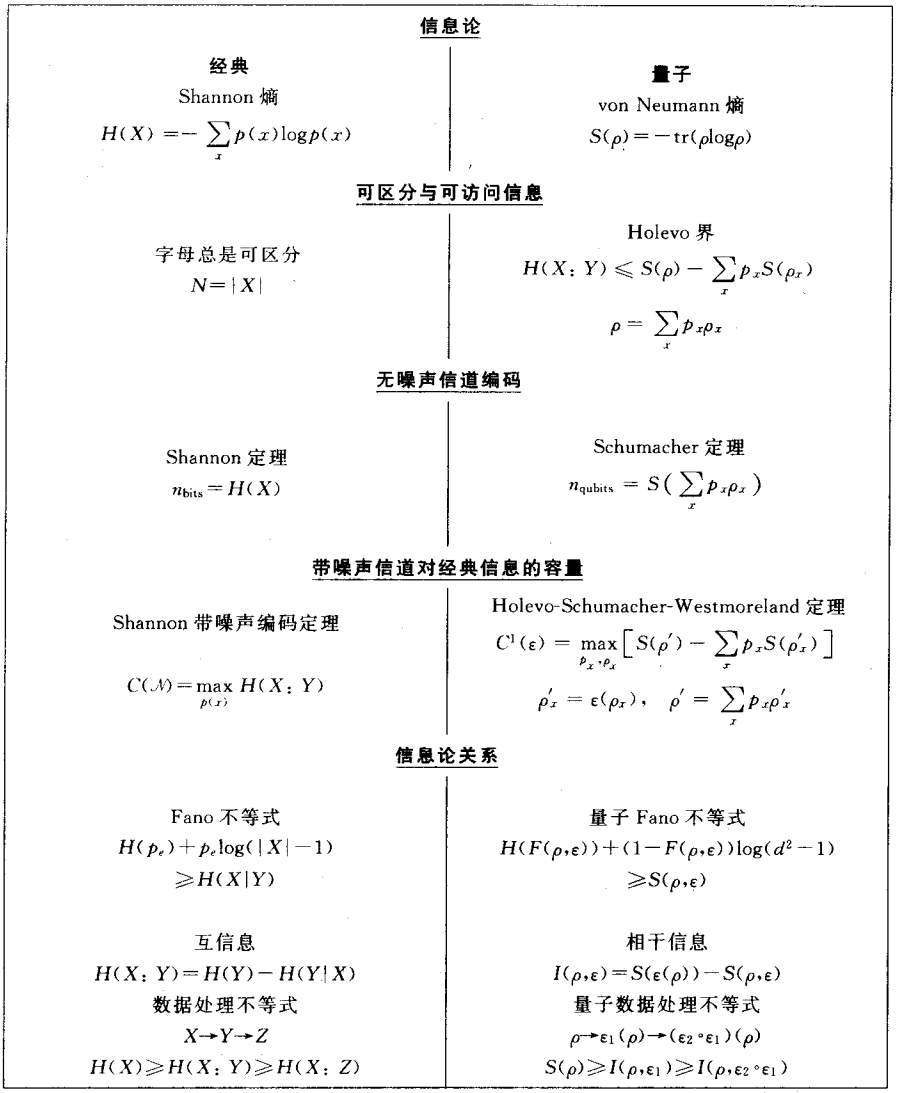


图 12.10　一些重要的经典信息关系的总结，以及这些关系的量子类比.

我们说“理解得最好”并不意味着理解了很多，建立一般的量子纠缠理论还有很长的路要走. 不过在这个建立一般理论的方向上已取得一些令人鼓舞的进展，揭示出了纠缠状态的内在结构，以及带噪声量子信道的性质和各类纠缠变换之间相当引人注目的联系. 现在只是简要概述已知的结果，主要集中在 Alice 和 Bob 双系统之间纠缠(二部纠缠)的变换性质. 当然，还有许多致力于发展多部系统纠缠的一般理论，但对如何去做尚未有很好的理解.

12.5.1 二部纯态纠缠的变换

我们从下列简单的问题开始：假定 Alice 和 Bob 共享一个纠缠纯态$|\psi\rangle$，设他们可以在本地系统上执行包括测量在内的任何运算，但只能使用经典通信，他们能够将$|\psi\rangle$变换为何种其他类型的纠缠$|\varphi\rangle$吗？Alice 和 Bob 之间不允许量子通信，就限制了他们可以完成变换的类型.

例如，Alice 和 Bob 共享处于 Bell 态$(|00\rangle+|11\rangle)/\sqrt{2}$的纠缠量子比特对. Alice 执行由测量算子 M_1 和 M_2 描述的双输出测量，

$$M_1=\begin{bmatrix}\cos\theta & 0\\ 0 & \sin\theta\end{bmatrix},\quad M_1=\begin{bmatrix}\sin\theta & 0\\ 0 & \cos\theta\end{bmatrix} \tag{12.154}$$

测量后，测量结果成为 1 或为 2，状态成为 $\cos\theta|00\rangle+\sin\theta|11\rangle$成为 $\cos\theta|11\rangle+\sin\theta|00\rangle$. 后一种情况下，Alice 在测量后应用一个非门，导致状态 $\cos\theta|01\rangle+\sin\theta|10\rangle$，然后她把测量结果(1 或 2)发送给 Bob. 如果测量结果是 1，Bob 就对状态不做什么；如果测量结果是 2，他就应用一个非门. 于是无论 Alice 的测量结果如何，联合系统的状态的终了状态都是 $\cos\theta|00\rangle+\sin\theta|11\rangle$. 即 Alice 和 Bob 只利用他们各自系统上的本地运算和经典通信，就把他们的初始纠缠资源$(|00\rangle+|11\rangle)/\sqrt{2}$变换为 $\cos\theta|00\rangle+\sin\theta|11\rangle$.

纠缠变换的重要性或许并非直截了当. 我们所允许类型的变换——本地运算和经典通信(LOCC)——具有特定的自身价值，不过，还完全看不出这是一个真正重要的问题. 然而事实上，该纠缠变换问题的推广展示出与量子纠错的深刻的和出人意料的联系，而且，为解决该问题引入的技术有相当的价值，并为纠缠的性质提供了意想不到的洞察力. 特别地，我们会发现纠缠与控制不等式(majorization)理论之间的密切的联系. 控制不等式理论实际上是在量子力学之前出现的一个数学领域.

在介绍纠缠变换的研究之前，让我们首先熟悉一下与控制不等式有关的一些事实. 控制是 d 维实向量上的一个序，其目的是刻画一个向量比另一个向量更无序(有序)的概念. 更确切地，设 $x=(x_1,\cdots,x_d)$和 $y=(y_1,\cdots,y_d)$为两个 d 维向量，我们用 $x^{\downarrow}$ 表示 x 的元按递减序的重排，故举例说 $x_1^{\downarrow}$ 就是 x 的最大元. 如果 $\sum_{j=1}^{k}x_j^{\downarrow}\leqslant\sum_{j=1}^{k}y_j^{\downarrow}$ 对 $k=1,\cdots,d$ 成立，且当 $k=d$ 时取等号，则说 x 被 y 控制(majorized)，记作 $x\prec y$. 这个概念与无序性之间的关系很快会清楚！

控制不等式与纠缠变换的联系很容易叙述但相当令人吃惊. 设$|\psi\rangle$和$|\varphi\rangle$是 Alice-Bob 联合系统的状态. 定义 $\rho_\psi\equiv\mathrm{tr}_B(|\psi\rangle\langle\psi|)$，$\rho_\varphi\equiv\mathrm{tr}_B(|\varphi\rangle\langle\varphi|)$为 Alice 系统的相应约化密度矩阵，且令 λ_ψ 和 λ_φ 为分量为 ρ_ψ 和 ρ_φ 特征值的向量. 当且仅当

$\lambda_\psi \prec \lambda_\varphi$ 时，我们将看到 $|\psi\rangle$ 可通过 LOCC 变换为 $|\varphi\rangle$. 为此，我们首先需要关于控制不等式的一些简单事实.

练习 12.17　证明当且仅当对所有实数 t，

$$\sum_{j=1}^{d}\max(x_j-t,0)\leqslant\sum_{j=1}^{d}\max(y_j-t,0),\ \text{且}\ \sum_{j=1}^{d}x_j=\sum_{j=1}^{d}y_j\ \text{时},x\prec y$$

练习 12.18　利用前一练习证明使得 $x\prec y$ 的 x 集合是凸的.

下面的命题给出控制概念的更直观含义，表明当且仅当 x 可被写作 y 的置换的一个凸组合时，$x\prec y$. 因此直观上，在 x 可由对 y 的元置换并将得到的向量混合得到的意义下，如果 x 比 y 更无序则 $x\prec y$. 该表示定理是控制不等式的研究中最有用的结果之一.

命题 12.11　当且仅当 $x=\sum_j p_j P_j y$ 对某个概率分布 p_j 和置换矩阵组 P_j 成立，$x\prec y$.

证　设 $x\prec y$. 不失一般性，设 $x=x^\downarrow$ 和 $y=y^\downarrow$. 我们将用维数 d 上的归纳法证明 $x=\sum_j p_j P_j y$. 对 $d=1$，结果显然. 假设 x 和 y 是 $d+1$ 维向量使得 $x\prec y$，则 $x_1\leqslant y_1$. 选择 j 使得 $y_j\leqslant x_1\leqslant y_{j-1}$，且定义 t 为 $[0,1]$ 范围内，使得 $x_1=ty_1+(1-t)y_j$. 定义置换的凸组合 $D\equiv tI+(1-t)T$，其中 T 为交换第 1 和第 j 个元的转置矩阵. 于是

$$Dy=(x_1,y_2,\cdots,y_{j-1},(1-t)y_1+ty_j,y_{j+1},\cdots,y_{d+1})\qquad(12.155)$$

定义 $x'\equiv(x_2,\cdots,x_{d+1})$ 和 $y'\equiv(y_2,\cdots,y_{j-1},(1-t)y_1+ty_j,y_{j+1},\cdots,y_{d+1})$. 练习 12.19 中证明 $x'\prec y'$，故由归纳假设 $x'=\sum_j p'_j P'_j y'$ 对概率 p'_j 和置换矩阵组 P'_j 成立，因此 $x=(\sum_j p'_j P'_j)Dy$，其中通过约定 P'_j 在第一项上平凡的作用扩展到 $d+1$ 维上. 因为 $D=(tI+(1-t)T)$，并且置换矩阵的乘积仍为置换矩阵，所欲证的结果成立. □

练习 12.19　验证 $x'\prec y'$.

反过来，设 $x=\sum_j p_j P_j y$，显然 $P_j y\prec y$，且由练习 12.18 可知 $x=\sum_j p_j P_j y\prec y$. □

作为置换矩阵的凸组合的矩阵有很多有趣的性质，例如，此类矩阵的元必为非负，且每行和每列的和必为 1. 具有这些性质的矩阵称为**双随机矩阵**，并且有一个称为 Birkhoff 定理的结果蕴含，双随机矩阵集合恰好与能写成置换矩阵的凸组合的矩阵集合相对应. 这里我们不证明 Birkhoff 定理（参看本章末“历史和进一步阅读的材料”），只给出定理的叙述.

定理 12.12（Birkhoff 定理）　d 乘 d 的矩阵 D 为双随机（即具有非负元且每

行每列和为 1)，当且仅当 D 可写作置换矩阵的凸组合，$D=\sum_j p_j P_j$.

从 Birkhoff 定理和命题 12.11 可知，当且仅当对某个双随机矩阵 D 成立 $x=Dy$，$x\prec y$. 该结果允许我们证明命题 12.11 的一个惊人和有用的算子推广. 假设 H 和 K 是两个 Hermite 算子，我们可以说如果 $\lambda(H)\prec\lambda(K)$ 则 $H\prec K$，其中用 $\lambda(H)$ 来记 Hermite 算子 H 特征值的向量. 于是我们有：

定理 12.13 令 H 和 K 为 Hermite 算子，则当且仅当存在概率分布 p_j 和酉矩阵 U_j，使得

$$H=\sum_j p_j U_j K U_j^\dagger \tag{12.156}$$

$H\prec K$

证 设 $H\prec K$，于是根据命题 12.11，$\lambda(H)=\sum_j p_j P_j \lambda(K)$. 令 $\Lambda(H)$ 表示元为 H 特征值的对角矩阵，于是向量方程 $\lambda(H)=\sum_j p_j P_j \lambda(K)$ 可以重新表示为

$$\Lambda(H)=\sum_j p_j P_j \Lambda(K) P_j^\dagger \tag{12.157}$$

但 $H=V\Lambda(H)V^\dagger$ 和 $\Lambda(K)=WKW^\dagger$ 对某酉矩阵 V 和 W 成立，给出 $H=\sum_j p_j U_j K U_j^\dagger$，其中 $U_j\equiv VP_jW$ 为酉矩阵. 这就完成证明向前的方向.

反过来，设 $H=\sum_j p_j U_j K U_j^\dagger$. 类似于前面，这等价于 $\Lambda(H)=\sum_j p_j V_j \Lambda(K) V_j^\dagger$ 对若干酉矩阵 V_j 成立. 把矩阵 V_j 的元写成 $V_{j,kl}$，有

$$\lambda(H)_k=\sum_{jl} p_j V_{j,kl}\lambda(K)_l V_{j,lk}^\dagger=\sum_{jl} p_j |V_{j,kl}|^2 \lambda(K)_l \tag{12.158}$$

定义具有元 $D_{kl}\equiv\sum_j p_j |V_{j,kl}|^2$ 的矩阵 D，于是有 $\lambda(H)=D\lambda(K)$. 根据定义，D 的元为非负，且因为酉矩阵 V_j 的行和列都是单位向量，所以 D 的行和列都是和为 1，于是 D 是双随机的，从而 $\lambda(H)\prec\lambda(K)$. □

我们已具备研究二部纯态纠缠的 LOCC 变换所需的关于控制不等式的所有事实. 论证的第一步是把研究双向经典通信的一般协议简化为只有单向经典通信的协议.

命题 12.14 设 $|\psi\rangle$ 可被 LOCC 变换为 $|\varphi\rangle$，于是该变换可由仅包含如下步骤的协议完成. Alice 进行一次由测量算子 M_j 描述的测量，把结果 j 发送给 Bob，Bob 在他的系统上进行酉运算 U_j.

证 不失一般性，设协议的组成为：Alice 进行一次测量，把结果发送给 Bob；Bob 进行一次测量(测量的性质可能依赖于从 Alice 收到的信息)，并把测量结果送回给 Alice；Alice 进行一次测量……；如此进行下去. 证明的思路仅为说明，Bob 可进行的任何测量的效果都能够被 Alice 所模拟(需要一点注意)，以致

Bob 的所有行动实际上都可以被 Alice 的行动代替. 为说明这一点，设想 Bob 在纯态$|\psi\rangle$进行具有测量算子 M_j 的测量，设该纯态具有 Schmidt 分解 $|\psi\rangle = \sum_l \sqrt{\lambda_l}|l_A\rangle|l_B\rangle$，并定义 Alice 系统上的算子 N_j 在 Alice 的 Schmidt 基底上的矩阵表示，与 Bob 算子 M_j 在 Bob 的 Schmidt 基底上的矩阵表示相同. 即若 $M_j = \sum_{kl} M_{j,kl}|k_B\rangle\langle l_B|$，则定义

$$N_j \equiv \sum_{kl} M_{j,kl}|k_A\rangle\langle l_A| \tag{12.159}$$

设 Bob 进行由测量算子 M_j 定义的测量，那么测量后的状态以概率 $\sum_{kl}\lambda_l|M_{j,kl}|^2$ 为 $|\psi_j\rangle \propto M_j|\psi\rangle = \sum_{kl} M_{j,kl}\sqrt{\lambda_l}|l_A\rangle|k_B\rangle$. 另一方面，如果 Alice 曾进行过 N_j 测量，那么测量后状态以概率 $\sum_{kl}\lambda_l|M_{j,kl}|^2$ 为 $|\varphi_j\rangle \propto N_j|\psi\rangle = \sum_{kl} M_{j,kl}\sqrt{\lambda_l}\cdot|k_A\rangle|l_B\rangle$. 进一步注意，除去通过映射$|k_A\rangle \leftrightarrow |k_B\rangle$互换 Alice 的系统和 Bob 的系统以外，$|\psi_j\rangle$和$|\varphi_j\rangle$为相同的状态，从而必有相同的 Schmidt 分量. 根据练习 2.80 可知，存在 Alice 系统上酉的 U_j 和 Bob 系统上的 V_j，使得$|\psi_j\rangle = (U_j \otimes V_j)|\varphi_j\rangle$. 于是 Bob 进行测量算子 M_j 描述的测量，等价于 Alice 进行测量算子 $U_j N_j$ 描述的测量加上随后 Bob 进行的酉变换 V_j. 总之，Bob 在已知纯态上的测量，除去 Bob 上的一个酉变换，能够被 Alice 的测量所模拟.

接着设想 Alice 和 Bob 参与把$|\psi\rangle$变换为$|\varphi\rangle$的多轮协议. 不失一般性，我们可设协议首轮包括 Alice 进行一次测量并把结果发送给 Bob. 第二轮由 Bob 进行一次测量(测量的类型或许决定于第一轮的结果)并把结果发送给 Alice. 不过，除去 Bob 的一个酉变换，我们可设该测量由 Alice 的一个测量所模拟. 事实上，我们可以用 Alice 的测量加上依赖于 Alice 测量结果的 Bob 的一个酉变换来代替 Bob 的所有测量和从 Bob 到 Alice 的通信. 最后，Alice 进行的所有测量可以合成为一个单次测量(练习 2.57)，其结果决定了 Bob 所执行的酉变换；该协议的净效果与原始的双向通信协议完全相同. □

定理 12.15　当且仅当 $\lambda_\psi \prec \lambda_\varphi$ 时，一个二部纯态$|\psi\rangle$可以被 LOCC 变换为另一纯态$|\varphi\rangle$.

证　设$|\psi\rangle$可被 LOCC 变换为$|\varphi\rangle$. 由命题 12.14，可假设变换的实现过程为 Alice 进行测量算子 M_j 的测量，然后把结果发送给 Bob，Bob 进行酉变换 U_j. 从 Alice 的观点看，她从状态 ρ_ψ 开始并结束在状态 ρ_φ，不管测量的结果，于是我们必有

$$M_j \rho_\psi M_j^\dagger = p_j \rho_\varphi \tag{12.160}$$

其中 p_j 是输出 j 的概率. 极分解 $M_j\sqrt{\rho_\psi}$蕴含存在一个酉的 V_j，使得

$$M_j\sqrt{\rho_\psi}=\sqrt{M_j\rho_\psi M_j^\dagger}\,V_j=\sqrt{p_j\rho_\varphi}\,V_j \tag{12.161}$$

等式前面乘上其伴随算子，得到

$$\sqrt{\rho_\psi}\,M_j^\dagger M_j\sqrt{\rho_\psi}=p_jV_j^\dagger\rho_\varphi V_j \tag{12.162}$$

对 j 求和并利用完备性关系 $\sum_j M_j^\dagger M_j=I$，得到

$$\rho_\psi=\sum_j p_jV_j^\dagger\rho_\varphi V_j \tag{12.163}$$

从而根据定理 12.13，得 $\lambda_\psi\prec\lambda_\varphi$.

反方向的证明本质上是将向前方向逆过来. 设 $\lambda_\psi\prec\lambda_\varphi$，故 $\rho_\psi\prec\rho_\varphi$，且由定理 12.13 必存在概率 p_j 和酉算子 U_j，使得 $\rho_\psi=\sum_j p_jU_j\rho_\varphi U_j^\dagger$. 暂且设 ρ_ψ 为可逆(该假设很容易取消；见练习 12.20)，并用

$$M_j\sqrt{\rho_\psi}\equiv\sqrt{p_j\rho_\varphi}\,U_j^\dagger \tag{12.164}$$

对 Alice 系统定义算子 M_j. 为明了这些算子定义了一个测量，我们需要检验完备性关系. 有 $M_j=\sqrt{p_j\rho_\varphi}U_j^\dagger\rho_\psi^{-1/2}$，因此得

$$\sum_j M_j^\dagger M_j=\rho_\psi^{-1/2}\Big(\sum_j p_jU_j\rho_\varphi U_j^\dagger\Big)\rho_\psi^{-1/2}=\rho_\psi^{-1/2}\rho_\psi\rho_\psi^{-1/2}=I \tag{12.165}$$

即为完备性关系. 设 Alice 进行由算子 M_j 描述的测量，得到结果 j 和相应的状态 $|\psi_j\rangle\propto M_j|\psi\rangle$. 令 ρ_j 表示 Alice 的对应于状态 $|\psi_j\rangle$ 的约化密度矩阵，于是代入式(12.164)

$$\rho_j\propto M_j\rho_\psi M_j^\dagger=p_j\rho_\varphi \tag{12.166}$$

从而 $\rho_j=\rho_\varphi$. 由练习 2.81 可知，Bob 通过应用适当的酉变换 V_j 可把 $|\psi_j\rangle$ 转换为 $|\varphi\rangle$. □

练习 12.20 证明，定理 12.15 反向部分的证明中可以取消 ρ_ψ 为可逆的假设.

练习 12.21(纠缠催化) 假设 Alice 和 Bob 共享一对处于状态 $|\psi\rangle=\sqrt{0.4}\cdot|00\rangle+\sqrt{0.4}|11\rangle+\sqrt{0.1}|22\rangle+\sqrt{0.1}|33\rangle$ 的四能级系统，证明该状态不可能被 LOCC 转换为状态 $|\varphi\rangle=\sqrt{0.5}|00\rangle+\sqrt{0.25}|11\rangle+\sqrt{0.25}|22\rangle$. 不过，设想如果有一家友善的银行愿意提供给他们“催化剂”的借贷，即处于状态 $|c\rangle=\sqrt{0.6}|00\rangle+\sqrt{0.4}|11\rangle$ 的量子比特的纠缠对. 证明 Alice 和 Bob 可以通过本地运算和经典通信把状态 $|\psi\rangle|c\rangle$ 转换为 $|\varphi\rangle|c\rangle$，在变换完成后将催化剂 $|c\rangle$ 还给银行.

练习 12.22(不带通信的纠缠转换) 设 Alice 和 Bob 试图仅利用本地运算——不用经典通信——把纯态 $|\psi\rangle$ 转换为纯态 $|\varphi\rangle$，证明，当且仅当 $\lambda_\psi\cong\lambda_\varphi\otimes x$，这是可能的，其中 x 是和为 1 的具有非负项的某个实向量，“$\cong$”表示左右两边具有相同的非零项.

12.5.2　纠缠的蒸馏和稀释

假设代替具有状态$|\psi\rangle$的单个备份，Alice 和 Bob 可用大量的备份. 用所有这些备份能够完成何种类型的纠缠变换？我们将集中在两类特定的纠缠变换，称为纠缠蒸馏和纠缠稀释. 纠缠蒸馏是指 Alice 和 Bob 用本地运算和经典通信把已知纯态$|\psi\rangle$的大量备份转换为 Bell 态$(|00\rangle+|11\rangle)/\sqrt{2}$的尽可能多的备份，且并非要求精确成功，而只要求具有高忠实度. 纠缠稀释指利用 LOCC 把 Bell 态$(|00\rangle+|11\rangle)/\sqrt{2}$的大量备份转换为$|\psi\rangle$的备份的逆过程，在起始时 Bell 态的备份数量充分大时同样具有高忠实度.

纠缠蒸馏和稀释研究的动机是什么？假设我们真正把纠缠作为一种物理资源，并认为作为物理资源能够对纠缠进行量化到与其他物理资源如能量或熵相当的量化程度. 假设我们决定选择 Bell 态$(|00\rangle+|11\rangle)/\sqrt{2}$作为纠缠的标准单位——基本度量，非常类似标准千克或标准米，我们能够以类似于将质量与物体关联的方式将纠缠的测度与量子状态$|\psi\rangle$关联. 例如，假设标准千克相当于 15 块某个特定牌子的巧克力饼干；我们说饼干具有 1/15kg 的质量. 严格地说，设想巧克力饼干具有的 1/14.8kg 质量，我们会遇到一点麻烦. 因为没有整块数的饼干可以与标准千克平衡，并且如何定义巧克力饼干的非整块数也并非显然. 幸运的是，我们所做的是注意到 148 块巧克力饼干严格等于 10kg，于是巧克力饼干的质量是 10/148kg. 但如果真实质量不是 1/14.8kg，而是更深奥的量比如 1/14.7982…kg 又怎样？那么，我们只要用巧克力饼干的一个充分大数目的 m 去平衡另外一个标准千克的大数 n，并声明当 m 和 n 都趋向极大时，一块巧克力饼干的质量为 n/m.

类似地，定义一个纯态$|\psi\rangle$中的纠缠量一种可能办法是设想被给定了一个大数目 n 的 Bell 态$(|00\rangle+|11\rangle)/\sqrt{2}$，并被要求用本地运算和经典通信产生$|\psi\rangle$的尽可能多(高忠实度)的备份. 如果可以产生的$|\psi\rangle$备份的数目为 m，那么我们定义极限情况下的比 n/m 为状态$|\psi\rangle$形成的纠缠. 换一种方式，我们可以设想进行逆向过程，从$|\psi\rangle$的 m 个备份用 LOCC 转换为$(|00\rangle+|11\rangle)/\sqrt{2}$的 n 个备份，并定义极限比 n/m 为状态$|\psi\rangle$的可蒸馏纠缠. 这两个定义给出相同数值的纠缠量并非显然；我们将看到对纯态$|\psi\rangle$形成纠缠和可蒸馏纠缠事实上完全相同！

让我们来看一个纠缠稀释的简单协议，以及另一个纠缠蒸馏的简单协议. 设纠缠状态$|\psi\rangle$具有 Schmidt 分解：

$$|\psi\rangle=\sum_x \sqrt{p(x)}\,|x_A\rangle\,|x_B\rangle \tag{12.167}$$

我们把平方后的 Schmidt 系数写作 $p(x)$，具有我们常保留给概率分布的形式，一方面是因为这些系数满足通常概率分布的规则(非负且和为 1)，也因为实际上概

率论的思想对理解纠缠蒸馏和稀释有用处. m 重张量积 $|\psi\rangle^{\otimes m}$ 可以写成

$$|\psi\rangle^{\otimes m} \equiv \sum_{x_1,x_2,\cdots,x_m} \sqrt{p(x_1)p(x_2)\cdots p(x_m)}\,|x_{1A}x_{2A}\cdots x_{mA}\rangle\,|x_{1B}x_{2B}\cdots x_{mB}\rangle \tag{12.168}$$

形式. 忽略所有那些 12.2.1 节意义下非 ε 典型的项 $x_1,\cdots,x_m$，定义一个新状态 $|\varphi_m\rangle$：

$$|\psi\rangle \equiv \sum_{x\in\text{典型}} \sqrt{p(x_1)p(x_2)\cdots p(x_m)}\,|x_{1A}x_{2A}\cdots x_{mA}\rangle\,|x_{1B}x_{2B}\cdots x_{mB}\rangle \tag{12.169}$$

该状态 $|\varphi_m\rangle$ 并不是适当地归一化的量子状态；为将其归一化，定义 $|\varphi'_m\rangle \equiv |\varphi_m\rangle / \sqrt{\langle\varphi_m|\varphi_m\rangle}$. 由典型序列定理的第 1 部分，忠实度 $F(|\psi\rangle^{\otimes m}, |\varphi'_m\rangle) \to 1$ 当 $m\to\infty$. 进而，由典型序列定理的第 2 部分，和式(12.169)中的项数至多为 $2^{m(H(p(x))+\varepsilon)} = 2^{m(S(\rho_\psi)+\varepsilon)}$，其中 ρ_ψ 是对 $|\psi\rangle$ 的 Bob 部分求迹的结果.

假设 Alice 和 Bob 共同拥有 $n = m(S(\rho_\psi)+\varepsilon)$ 个 Bell 态. Alice 在本地制备 $|\varphi'_m\rangle$ 的全部两个部分，然后用与 Bob 共享的 Bell 态，通过隐形传态把 $|\varphi'_m\rangle$ 中 Bob 应有的一半传给 Bob. 这样 Alice 和 Bob 可以稀释他们的 nBell 态得到 $|\varphi'_m\rangle$，这是 $|\psi\rangle^{\otimes m}$ 的一个很好的近似. 这个纠缠稀释过程的 $n = m(S(\rho_\psi)+\varepsilon)$，故比 n/m 趋向于 $S(\rho_\psi)+\varepsilon$. 我们可以选择任意小的 ε，于是可知状态 $|\psi\rangle$ 的形成纠缠不超过 $S(\rho_\psi)$，因为我们刚刚(近似地)说明 $S(\rho_\psi)$ 个 Bell 态可以被转换成 $|\psi\rangle$ 的一个单个备份.

一个把 $|\psi\rangle$ 的备份转换成 Bell 态纠缠蒸馏协议可沿类似思路给出. 设 Alice 和 Bob 共同拥有 $|\psi\rangle$ 的 m 个备份. Alice 通过对 ρ_ψ 的 ε 典型子空间上的测量，可以以高忠实度把状态 $|\psi\rangle^{\otimes m}$ 转换为状态 $|\varphi'_m\rangle$. 由典型序列的定义，在 $|\varphi_m\rangle$ 中最大的 Schmidt 系数至多为 $2^{-m(S(\rho_\psi)-\varepsilon)}$. 归一化后的状态 $|\varphi'_m\rangle$ 的 Schmidt 系数至多大一个因子 $1/\sqrt{(1-\delta)}$，因为典型序列定理告诉我们 $(1-\delta)$ 是序列为 ε 典型的概率的一个下界，并且对充分大的 m 可以任意接近 1. 因此，状态 $\rho_{\varphi'_m}$ 的最大特征值至多为 $2^{-m(S(\rho_\psi)-\varepsilon)}/(1-\delta)$. 选择任意 n，使得

$$\frac{2^{-m(S(\rho_\psi)-\varepsilon)}}{1-\delta} \leqslant 2^{-n} \tag{12.170}$$

则 $\rho_{\varphi'_m}$ 的特征值的向量被向量 $(2^{-n}, 2^{-n}, \cdots, 2^{-n})$ 控制，从而据定理 12.15，状态 $\rho_{\varphi'_m}$ 通过本地运算和经典通信可以转换为 n 个 Bell 状态. 检查式(12.170)，可以看到在 $n \approx mS(\rho_\psi)$ 条件下这是可能的，于是纠缠的蒸馏至少为 $S(\rho_\psi)$.

我们展示了蒸馏 $|\psi\rangle$ 到 $S(\rho_\psi)$ 个 Bell 态和把 $S(\rho_\psi)$ Bell 态稀释为 $|\psi\rangle$ 的一个备份的方法. 事实上，不难看出我们已描述的方法实际上是纠缠稀释和蒸馏的最优方法！比如，假设还有更有效的纠缠稀释协议，能够把 $|\psi\rangle$ 稀释到 $S > S(\rho_\psi)$ 个 Bell

态. 则从 $S(\rho_\psi)$ 个 Bell 态开始，Alice 和 Bob 能用已描述的协议产生 $|\psi\rangle$ 的一个备份，然后用假设的协议产生 S 个 Bell 态. 因此，通过本地运算和经典通信，Alice 和 Bob 已经把 $S(\rho_\psi)$ 个 Bell 态变成了 $S>S(\rho_\psi)$ 个 Bell 态. 容易相信(并参看练习 12.24)，用本地运算和经典通信增加 Bell 态的数目是不可能的，因此假设的稀释协议不可能存在. 类似地，我们可以看到所给的纠缠蒸馏过程也是最优的. 因此形成的纠缠和蒸馏的纠缠对状态 $|\psi\rangle$ 是相同的，并都等于 $S(\rho_\psi)$.

练习 12.23　证明已描述的纠缠蒸馏过程是最优的.

练习 12.24　回忆二部纯态的 Schmidt 数为非零 Schmidt 元的数目，证明一个纯的量子状态的 Schmidt 数不可能通过本地运算和经典通信被增加. 利用这个结果论证，Alice 和 Bob 共享的 Bell 态数目不能通过本地运算和经典通信增加.

我们已经学了如何最优地将二部量子系统的 Bell 态变到另外的纠缠状态 $|\psi\rangle$ 的备份，并再变回来. 这启发我们定义一个量子状态的纠缠量，该量为可与 $|\psi\rangle$ 备份相互转换的 Bell 态数目，即 $S(\rho_\psi)$. 从该定义我们学到什么？下面我们将看到，通过进一步推广可蒸馏纠缠的概念可以获得关于量子纠错的一些有价值的见解. 不过，目前，应当说关于纠缠的研究仍处于萌芽阶段，现在还不完全清楚定量测量纠缠的研究结果对量子计算与量子信息的理解能够预期什么进展. 我们对二部量子系统的纯态性质有相当的了解，但对包含三个或更多部分的系统或者甚至是二部系统的混合态的了解却非常有限. 加深对纠缠的理解，以及把它同包括量子算法、量子纠错和量子通信等问题联系起来，是量子计算与量子信息未来一项主要的研究任务.

12.5.3　纠缠蒸馏与量子纠错

我们对纯态定义了量子蒸馏，但没有理由使该定义不能扩展到混合态. 更确切地，设 ρ 为属于 Alice 和 Bob 的二部量子系统的一个一般状态. 他们被提供了大数量 m 个状态备份，并且使用本地运算和经典通信，试图以高忠实度把这些状态转换为最大 n 个 Bell 态. ρ 的可蒸馏纠缠 $D(\rho)$ 是可能的最好蒸馏协议的比 n/m 的极限值. 对纯态 $|\psi\rangle$，我们已证明 $D(|\psi\rangle)=S(\rho_\psi)$，但我们尚不知对混合态如何求 $D(\rho)$.

已有许多方法进行纠缠蒸馏，对特定类型的状态 ρ，它们给出 $D(\rho)$ 值的下界. 这里不讨论这些方法(参考章末的“历史和进一步阅读的材料”)，我们要描述的是纠缠蒸馏与量子纠错之间迷人的联系.

设想 Alice 要通过带噪声量子信道 ε 给 Bob 发送量子信息. 我们假设信道是量子比特信道，例如去极化信道，尽管基本思想很容易用到非量子比特信道. 通过信道发送量子信息的一种方法如下所述. Alice 制备大数目 m 个 Bell 态，并通过信道把每个 Bell 对的一半发送给 Bob. 设把 ε 应用到 Bell 对的一半的结果是产生状

态 ρ，于是最终 Alice 和 Bob 共享 ρ 的 m 个备份. Alice 和 Bob 进行纠缠蒸馏，产生 $mD(\rho)$个 Bell 对. Alice 现在可以制备一个 $mD(\rho)$量子比特的状态，并用他们所共享的 $mD(\rho)$个 Bell 对把它隐形传态给 Bob.

因此量子蒸馏协议可用于包含 Alice 和 Bob 的两方量子通信信道的一类纠错方法，使得 Alice 可以可靠地发送 $mD(\rho)$量子比特信息给 Bob，其中 $D(\rho)$是 ρ 的可蒸馏纠缠，ρ 是一个 Bell 对的一半通过 Alice 和 Bob 之间带噪声信道 ε 发送所导致的状态.

真正值得注意的是，这种采用纠缠蒸馏的通信方法甚至在传统量子纠错技术失败的某些情况下仍是可用的. 例如，我们在 12.4.3 节看到对 $p=3/4$ 的去极化信道，没有量子信息可以通过该信道传送. 然而，即使对这个信道，业已知道纠缠蒸馏协议仍能产生非零的传输比率 $D(\rho)$！这成为可能的原因是，纠缠蒸馏协议允许 Alice 和 Bob 之间往返的经典通信，而传统的量子纠错却不允许任何这样的经典通信.

此例允许我们解释在第 1 章所作的断言，如图 12.11 所说明的，就是存在对量子信息容量为零的信道，当一条这样的信道连接 Alice 到 Bob，且另一条连接 Bob 到 Alice 时，可以用于得到量子信息的净流动. 做到这一点的方法很简单，是基于纠缠蒸馏的. 为使纠缠成为可能，我们需要 Alice 和 Bob 能够进行经典通信，因此我们取出信道的前向使用的一半和反向通道的全部，用作蒸馏协议所使用的经典信息的传输. 由 HSW 定理，这些信道对经典信息传输具有非零比率. 信道前向使用剩下的一半用于从 Alice 到 Bob 传输 Bell 对的一半，而纠缠蒸馏用于从结果状态中抽取好的 Bell 对，然后把好的 Bell 对进行隐形传态以达成量子信息的净传输. 这是量子信息惊人性质的又一生动展示！

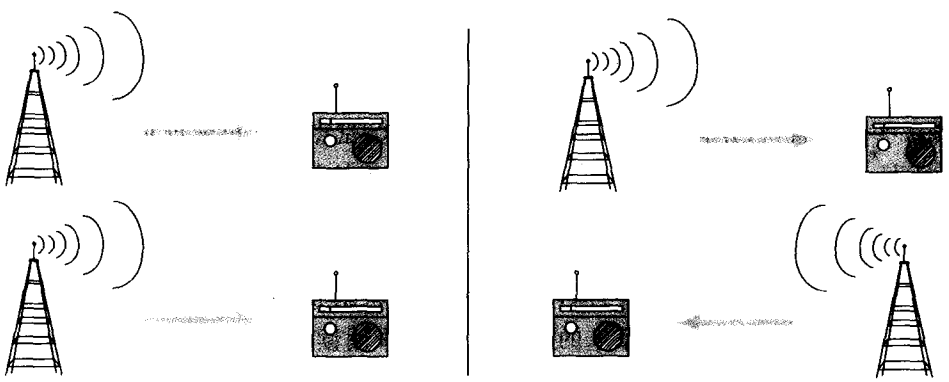


图 12.11 经典情形下，如果我们有两个并排的具有零容量的噪声很大的信道，那么两个信道合起来发送信息的容量为零. 不出意料，如果我们翻转一个信道的方向，我们发送信息的容量仍为零. 量子力学情形，翻转一个零容量信道实际上可以允许我们发送信息.

12.6　量子密码术

以量子信息最值得关注的一项应用作为本章的结束是非常合适的. 如第 5 章(见《量子计算和量子信息(一)》)所看到的,量子计算机可用于破解某些最优秀的公钥密码系统. 幸运的是,量子力学剥夺了一方面,它在另一方面给出了补偿:称作量子密码术或量子密钥分配的过程利用量子力学原理来保证秘密信息的可证明的安全分配. 本节来描述该过程,并讨论其安全性. 我们首先在 12.6.1 节中解释经典技术私钥密码术的基本思想. 私钥密码术是远早于公钥密码术(第 5 章提到过)的一种形式,并且私钥密码术的原理被用到了量子密码系统中. 量子系统中使用的另外两项经典技术,保密增强和信息调和,在 12.6.2 节中描述. 12.6.3 节给出量子密钥分配的三种不同协议. 这些协议的安全性如何? 结果如 12.6.4 节将看到的,12.4.1节首次遇到的量子信息的一个度量,相干信息,给出了原则上通过量子通信信道发送秘密信息能力的一个信息论下界. 这暗示量子信息的思想可能对证明特定量子密钥分配协议的安全性有用处,而事实上的确如此:12.6.5 节以概述量子纠错理论如何提供量子密码术安全性的证明来结束本章.

12.6.1　私钥密码术

直到 20 世纪 70 年代公钥密码术的发明,所有的密码术都基于另一不同的原理,现在其称为私钥密码术. 在私钥系统中,如果 Alice 希望向 Bob 发送消息,那么必须有一个加密密钥,使她能对她的消息进行加密,而 Bob 必须有一个相匹配的解密密钥,使 Bob 可以对加密的消息进行解密. Vernam 密码是一个简单但高度有效的私钥密码系统,有时也称为一次性便笺. Alice 和 Bob 用完全相同的 n 比特密钥串开始. Alice 通过将消息和密钥加起来对 n 比特消息进行加密,而 Bob 通过减法逆转加密过程来解密,如图 12.12 所示.

该系统的重要特征是只要密钥串是真正保密的,它的安全性是可证明的. 即,当该协议被 Alice 和 Bob 成功使用时,它成功的概率可以任意高(窃听者 Eve 总可以阻塞通信信道,但 Alice 和 Bob 可以检测到该阻塞并宣布失败). 并且对 Eve 采用的任何窃听方法,Alice 和 Bob 可以保证他们的未加密消息和 Eve 的互信息可以做到任意小. 与之对照,公钥密码术(《量子计算和量子信息(一)》附录 E)的安全性依赖于关于求解某些问题如求因子(在经典计算机上)的困难性,尽管它被广泛采用并且更方便.

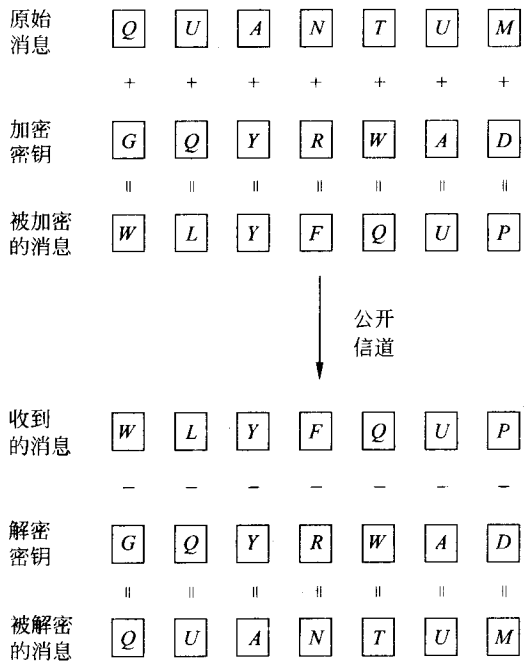


图 12.12 Verman 密码. Alice 用随机密钥比特(或在此例中,字母表中的字母)加到原始消息上进行加密. Bob 通过减去密钥比特来恢复消息.

私钥密码系统的主要困难是密钥比特的安全分配. 特别地,只有当密钥比特的数目至少与被加密消息的尺寸一样大,而且密钥比特不被重用,Vernam 密码才是安全的,因此,密钥比特数目的庞大使得这种方案一般来说不实用. 而且,密钥比特必须事先分配,直到使用前都要努力看管,使用后销毁;否则,原则上,这类经典信息可以在不干扰原始信息情况下被复制,因此使整个协议的安全性打了折扣. 尽管有这样的不足,像 Vernam 密码这样的私钥密码系统因其可证明的安全性而继续被采用,密钥材料的发放则是通过秘密会议、可信的信使,或专用保密通信链路.

练习 12.25 考虑有 n 个用户的系统,任意一对用户都希望能进行保密通信. 采用公钥密码术需要多少个密钥? 采用私钥密码术需要多少个密钥?

12.6.2 保密增强与信息调和

私钥密码术的第一步是密钥串的分配. 如果 Alice 和 Bob 开始时的密钥不完整会怎样? 具体地,假设 Alice 和 Bob 共享相关的随机比特串 X 和 Y,并且他们也

有 Eve 和 X 与 Y 互信息的一个上界. 从这些不完整密钥，他们如何获得一个可以执行安全密码协议的足够好的密钥？我们现在说明通过信息调和及保密增强两个步骤，他们可以系统地增加他们密钥串之间的相关性，同时减少窃听者 Eve 关于结果的互信息，达到任意的安全水平.

这些经典步骤将应用到下节量子密钥分配协议中.

信息调和只不过是在公共信道上执行的纠错，它调和 X 和 Y 之间的差错以获得共享的比特串 W，同时尽可能少地泄露给 Eve. 这个过程之后，假设 Eve 得到一个与 W 部分相关的随机变量 Z，接着 Alice 和 Bob 使用保密增强从 W 蒸馏出与 Z 的相关性低于期望阈值的一组较小的比特 S. 因为这个最后的步骤概念上是新的，让我们先来考虑它.

为什么保密增强能够成功的详细证明超出了本书的范围，但我们将描述基本方法和主要定理. 完成保密增强的一种方式是采用*通用散列函数* $\mathcal{G}$ 的类型，该类把 n 比特串的集合 $\mathcal{A}$ 映射到 m 比特串的集合 $\mathcal{B}$，使得对任意不同的 $a_1, a_2 \in \mathcal{A}$，当 g 从 $\mathcal{G}$ 随机均匀选取时，$g(a_1)=g(a_2)$ 的概率至多为 $1/|\mathcal{B}|$.

具有概率分布 $p(x)$ 的随机变量 X 的*碰撞熵*定义为

$$H_c(X) = -\log\left[\sum_x p(x)^2\right] \tag{12.171}$$

(有时称这为二阶 Rènyi 熵). 利用 log 函数的凹性，不难证明 Shannon 熵为该量提供了一个上界：$H(X) \geqslant H_c(X)$. H_c 在下面关于通用散列函数的定理中很重要.

定理 12.16　令 X 为字母表 $\mathcal{X}$ 上的随机变量，具有概率分布 $p(x)$ 和碰撞熵 $H_c(X)$，且令 G 为对应于从自 $\mathcal{X}$ 到 $\{0,1\}^m$ 的通用散列函数类中随机(均匀)选择一个成员的随机变量，则

$$H(G(X) \mid G) \geqslant H_c(G(X) \mid G) \geqslant m - 2^{m-H_c(X)} \tag{12.172}$$

定理 12.16 可按如下方式用于保密增强. Alice 和 Bob 公开选择 $g \in \mathcal{G}$，并都把它应用于 W，得到一个新的比特串 S，他们就选它作秘钥. 如果 Eve 关于 W 的不确定性在给定她的知识 $Z=z$(关于协议的一个特定实例)条件下已知是以碰撞熵以某个数作下界形式给出，即 $H_c(W|Z=z)>d$，那么从定理 12.16 可知

$$H_c(S \mid G, Z=z) \geqslant m - 2^{m-d} \tag{12.173}$$

换句话说，m 可以选得足够小使得 $H_c(S|G,Z=z)$ 约等于 m. 这就最大化了 Eve 关于密钥 S 的不确定性，使其成为安全的秘密.

信息调和进一步减少 Alice 和 Bob 能够得到的比特数，但可按如下方式估界. 通过计算她的比特串 X 的子集上一系列的奇偶校验结果，Alice 可以写出一条由子集规范和奇偶性组成的(经典)消息 u，并发送给 Bob，以允许 Bob 纠正他的串 Y

的差错，之后两人具有相同的串 W. 显然这需要在 u 中发送 $k>H(W|Y)$ 比特的信息. 然而，该过程给了 Eve 额外的知识 $U=u$，因此将她的碰撞熵增加到 $H_c(W|Z=z, U=u)$（在所有可能的调和消息 u 上）. 平均而言，这个增量有下界 $H_c(W|Z=z, U=u)\geqslant H_c(W|Z=z)-H(U)$，其中 $H(U)$ 是 U 的普通 Shannon 熵. 但这个界太弱，因为它意味着泄露的信息 $U=u$ 超过 $mH(U)$ 减少 H_c 的概率只有 $1/m$. 下述定理提供了一个更强的界.

定理 12.17 令 X 和 U 分别是具有字母表 $\mathcal{X}$ 和 $\mathcal{U}$ 的随机变量，其中 X 具有概率分布 $p(x)$，而 U 和 X 的联合分布为 $p(x,u)$，同时令 $s>0$ 为任意参数，那么，至少以 U 取某个值 u 使得

$$H_c(X\mid U=u)\geqslant H_c(X)-2\log|\mathcal{U}|-2s \tag{12.174}$$

的概率至少为 $1-2^{-s}$. 这里，s 称为安全性参数. 把这个定理应用到调和协议得出结论，Alice 和 Bob 可以选择 s，使得 Eve 的碰撞熵以好于 $1-2^{-s}$ 的概率具有下界 $H_c(W|Z=z,U=u)\geqslant d-2(k+s)$. 接着这个步骤，利用保密增强允许他们蒸馏出 m 个密钥比特 S，对于它，Eve 的全部信息少于 $2^{m-d+2(k+s)}$ 比特.

下面，将具体地讨论 CSS 码保密增强与信息调和的关系. 如我们上面注意的，信息调和不过是纠错；实际上保密增强也与纠错有密切关系，且这两项任务都可以用经典码实现. 在 12.6.5 节中，这个观点提供了一个简单的概念性图画，它在量子密钥分配安全性证明中有用，因为我们有成熟的量子纠错码理论. 有了这些，对下列观察是有用的.

从随机选定的 CSS 码（见 10.4.2 节）解码可以看作进行信息调和和保密增强. 虽然 CSS 码通常用于编码量子信息，但对于当前的目的，我们可以只考虑它们的经典性质. 考虑两个经典的线性码 C_1 和 C_2，它们对一个 t 纠错 $[n,m]$CSS 码满足条件：$C_2\subset C_1$ 且 C_1 和 $C_2^{\perp}$ 均可纠正 t 个差错. Alice 选择一个随机的 n 比特串 X 并把它传输给 Bob，他收到 Y.

让我们假设事先已知沿 Alice 和 Bob 之间的经典信道，每个码块由包括窃听在内的所有噪声源引起的平均差错数小于 t；实践中，这可以通过随机测试信道来确认. 进而，假设 Eve 对码 C_1 和 C_2 一无所知；这可通过 Alice 对码的随机选择来保证. 最后，假设 Alice 和 Bob 有一个关于 Eve 数据 Z 和他们自己数据 X 和 Y 之间互信息的上界.

Bob 收到 $Y=X+\varepsilon$，其中 ε 是差错. 因为已知发生的差错少于 t 个，如果 Alice 和 Bob 都把他们的状态纠正到 C_1 中的最近码字，他们的结果 $X',Y'\in C_1$ 将相同，$W=X'=Y'$. 这个步骤只不过是信息调和. 当然，Eve 关于 W 的互信息可能仍然很大，不可让人接受. 为减少它，Alice 和 Bob 辨别出他们的状态 W 属于 C_2 在 C_1 中

的 2^m 个陪集的哪一个；即他们计算 $W+C_2$ 在 C_1 中的陪集. 结果是他们的 m 比特密钥串 S. 利用 Eve 对 C_2 知识的缺乏和 C_2 的纠错性质，该过程可以把 Eve 关于 S 的互信息减少到可接受的程度，实现了保密增强.

12.6.3　量子密钥分配

量子密钥分配(QKD)是一个可证明为安全的协议，通过它，两方可以在公开信道上创建若干私钥比特. 这些密钥比特则可用于实现经典的私钥密码系统，使他们可以安全地通信. 对 QKD 协议的惟一要求是，量子比特在公开信道上可以以低于某个阈值的差错率通信. 结果密钥的安全性由量子信息的性质保证，这仅以物理学基本定律的正确性为前提！

QKD 背后的基本思想是遵循如下基本的观察：Eve 在不干扰 Alice 和 Bob 状态的情况下，不能得到任何关于从 Alice 传输到 Bob 的量子比特的信息. 首先，由不可克隆原理(盒子 12.1)，Eve 不能克隆 Alice 的量子比特，其次，我们有下述命题.

命题 12.18(获得信息蕴含干扰)　在区分两个非正交量子状态的任何尝试中，只有以信号引入干扰为代价，信息的获得才可能.

证　令 $|\psi\rangle$ 和 $|\varphi\rangle$ 为 Eve 试图获得相关信息的非正交量子状态. 根据 8.2 节的结果，不失一般性假设，她用于获得信息的过程是在状态($|\psi\rangle$ 或 $|\varphi\rangle$)上的酉相互作用以及制备为标准状态 $|u\rangle$ 的一个辅助单元. 假设该过程不干扰状态，在两种情况下得到

$$|\psi\rangle|u\rangle \to |\psi\rangle|v\rangle \tag{12.175}$$

$$|\varphi\rangle|u\rangle \to |\varphi\rangle|v'\rangle \tag{12.176}$$

Eve 希望 $|v\rangle$ 与 $|v'\rangle$ 不同，使得她可以得到用于识别状态的信息，然而，因为酉变换保持内积，必有

$$\langle v|v'\rangle\langle\psi|\varphi\rangle = \langle u|u\rangle\langle\psi|\varphi\rangle \tag{12.177}$$

$$\langle v|v'\rangle = \langle u|u\rangle = 1 \tag{12.178}$$

这意味着 $|v\rangle$ 与 $|v'\rangle$ 必然相同. 因此为区分 $|\psi\rangle$ 和 $|\varphi\rangle$，必然不可避免地要干扰至少其中一个状态. □

我们通过在 Alice 和 Bob 之间传输非正交量子比特状态来利用这一思想. 通过检查他们传输状态中的干扰，他们对通信信道中出现的任何噪声或窃听建立一个上界. 这些校验量子比特随机地插入到数据量子比特(后来会从中抽取密钥比特)中，使得上界对数据量子比特也适用. Alice 和 Bob 于是进行信息调和和保密增强，以蒸馏出共享的密码密钥串. 因此最大容许差错率取决于最佳信

息调和和保密增强协议的能力. 下面给出按这种方式工作的三个不同的 QKD 协议.

1. BB84 协议

Alice 从 a 和 b 开始,它们两个各有 $(4+\delta)n$ 位的随机经典比特串. 她把这些串编码为具有 $(4+\delta)n$ 量子比特的一个块,

$$|\psi\rangle = \bigotimes_{k=1}^{(4+\delta)n} |\psi_{a_k b_k}\rangle \tag{12.179}$$

其中 a_k 是 a 的第 k 比特(对 b 类似),且每个量子比特为四个状态之一:

$$|\psi_{00}\rangle = |0\rangle \tag{12.180}$$

$$|\psi_{10}\rangle = |1\rangle \tag{12.181}$$

$$|\psi_{01}\rangle = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \tag{12.182}$$

$$|\psi_{11}\rangle = |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2} \tag{12.183}$$

该过程的效果是把 a 按 b 确定,在基底 X 或 Z 中编码. 注意四个状态不是都相互正交的,因此没有测量可以完全确定地区分它们(的全部). 接着,Alice 在公共量子通信信道上发送 $|\psi\rangle$ 给 Bob.

Bob 收到 $\varepsilon(|\psi\rangle\langle\psi|)$,其中 ε 描述信道和 Eve 行动组合作用导致的量子运算,他随后公开宣布这个事实. 此时,Alice,Bob 和 Eve 均各自持有由分别的密度矩阵描述的各自状态. 同时注意,此时由于 Alice 还没有透露 b,Eve 并不知道她应该在哪个基底下测量以窃听通信;她最多只能猜测,并且如果猜测错误,那么她将干扰 Bob 收到的状态. 而且,现实中噪声 ε 不仅来源于 Eve 的窃听还或许部分地源自环境(糟糕的信道),因此它并不能帮助 Eve 完全控制信道,使得她完全对 ε 负责.

当然,此时 Bob 也发现 $\varepsilon(|\psi\rangle\langle\psi|)$ 不给出信息,因为他不具有关于 b 的知识. 然而,他继续进行,按他自己创建的随机 $(4+\delta)n$ 比特串 b' 所确定的,在基底 X 或 Z 中测量每个量子比特. 令 Bob 的测量结果是 a'. 这之后,Alice 公开宣布 b,并且通过在一个公开信道上的讨论,Alice 和 Bob 丢弃 $\{a', a\}$ 中所有除对应 b' 和 b 相等比特以外的比特. 他们剩余的那些比特满足 $a' = a$,因为对这些比特,Bob 测量和 Alice 制备是在同样的基底中. 注意 b 没有揭示关于 a 或从 Bob 测量得到的比特 a' 的任何事情,但 Alice 直到 Bob 宣布收到 Alice 的量子比特之后才公布 b 这点很重要. 为在下面的解释中简单起见,令 Alice 和 Bob 只保留他们结果的 $2n$ 比特;δ 可以选得充分大,使得这样做具有指数高的概率.

现在 Alice 和 Bob 进行一些测试,以确定在他们通信过程中出现了多少噪声或窃听. Alice 随机选择(他们的 $2n$ 比特的)n 比特,并公开宣布该选择. Bob

和 Alice 接着公开并比较这些校验比特的值. 如果有多于 t 比特的不同，那么他们中止并从开始重试协议. t 选得使如果测试通过，那么他们可以从剩余的 n 比特中应用信息调和与保密增强算法，来获得 m 个可接受的安全的共享密钥比特.

该协议称为 BB84，据发明者命名(见章末“历史和进一步阅读的材料”)，总结在图 12.13 中，并且在盒子 12.7 中描述一个实验上的实现. 该协议的相关版本，例如使用较少校验比特的，也具有相同的名称.

BB84 QKD 协议

1：Alice 随机选择 $(4+\delta)n$ 个随机数据比特.

2：Alice 随机选择一个 $(4+\delta)n$ 比特的串 b. 如果 b 中相应的比特为 0，她把数据比特的每个比特编码为 $\{|0\rangle,|1\rangle\}$；或者如果 b 的相应比特为 1 则编码为 $\{|+\rangle,|-\rangle\}$.

3：Alice 把这样得到的状态发送给 Bob.

4：Bob 收到 $(4+\delta)n$ 个量子比特，宣布这一事实，并随机地用 X 或 Z 测量每个量子比特.

5：Alice 宣布 b.

6：Alice 和 Bob 丢弃所有 Bob 用不同于 Alice 制备所用的基测量的比特. 以很高的概率，至少有 $2n$ 个比特会剩下(否则终止协议). 他们保持 $2n$ 个比特.

7：Alice 选择一个 n 比特子集合，以用作检查 Eve 的干扰，并告诉 Bob 她选择了哪些比特.

8：Alice 和 Bob 宣布并比较 n 个校验比特的值. 如果多于可接受的数目不同，他们终止协议.

9：Alice 和 Bob 在剩下的 n 比特上进行信息调和和保密增强以获得 m 比特的共享密钥.

图 12.13　称为 BB84 的四状态量子密钥分配协议.

练习 12.26　令 a'_k 为 Bob 对量子比特 $|\psi_{a_k b_k}\rangle$ 的测量结果，这里假设无噪声信道且未被窃听. 证明当 $b'_k \neq b_k$ 时，a'_k 为随机且与 a_k 完全不相关；但当 $b'_k = b_k$ 时，$a'_k = a_k$.

练习 12.27(随机采样测试)　随机测试 $2n$ 个测试比特的 n 比特，允许 Alice 和 Bob 以高概率为他们未测试的比特中差错的数目定出一个上界. 具体地，对任意的 $\delta > 0$，对大的 n，在测试比特上得到少于 δn 个差错，而在剩下 n 比特上存在 $(\delta+\varepsilon)n$ 个差错的概率渐近地小于 $\exp[-O(\varepsilon^2 n)]$. 我们在这里证明该断言.

(1) 不失一般性，可以假定在 $2n$ 比特中存在 μn 个差错，其中 $0 \leqslant \mu \leqslant 2$，则如果校验比特上有 δn 个差错，而在其余比特上有 $(\delta+\varepsilon)n$ 个差错，那么 $\delta = (\mu-\varepsilon)/2$. 断言中的两个条件于是蕴含下面的结论：

$$< \delta n \text{ 个差错在校验比特} \Rightarrow < \delta n \text{ 个差错在校验比特} \tag{12.184}$$

$$> (\delta+\varepsilon)n \text{ 个差错在其他比特} \Rightarrow > (\mu-\delta)n \text{ 个差错在其他比特} \tag{12.185}$$

并且事实上，上面右边的断言蕴含下面右边的断言. 基此，证明我们希望估界的概率 p 满足

$$p<\binom{2n}{2n}^{-1}\binom{\mu n}{\delta n}\binom{(2-\mu)n}{(1-\delta)n}\delta n \tag{12.186}$$

(2) 证明对大的 n，有估界

$$\frac{1}{an+1}2^{anH(b/a)}\leqslant\binom{an}{bn}\leqslant 2^{anH(b/a)} \tag{12.187}$$

其中 $H(\cdot)$是二元熵函数，即式(11.8). 把此式子用到上述 p 的界中.

(3) 用界 $H(x)<1-2(x-1/2)^2$ 来获得最终结果 $p<\exp[-O(\varepsilon^2 n)]$. 可以用一个表示可能的最坏情况的常数代替 μ.

(4) 将结果与盒子 3.4 的 Chernoff 界比较. 你能够想出 p 的上界的不同推导方式吗?

2. B92 协议

BB84 协议可以推广，使用其他状态和基底，类似的结论成立. 事实上，存在仅使用两个状态的特别的简单协议. 为简单起见，每次只考虑单比特的情形就够了，如 BB84 所做的，这个描述很容易推广到块测试.

设 Alice 制备了一个随机经典比特 a，并且，依赖于结果，发送给 Bob

$$|\psi\rangle=\begin{cases}|0\rangle, & a=0\\ \dfrac{|0\rangle+|1\rangle}{\sqrt{2}} & a=1\end{cases} \tag{12.188}$$

依赖于他所产生的随机经典比特 a'，Bob 接着测量他从 Alice 收到的量子比特，要么在 Z 基底$|0\rangle,|1\rangle$(若 $a'=0$)，要么在 X 基底$|\pm\rangle=(|0\rangle\pm|1\rangle)/\sqrt{2}$(若 $a'=1$). 从他的测量，对应于 X 和 Z 的本征态 -1 和 $+1$，他得到的结果 b 分别是 0 或 1. 然后 Bob 公开宣布 b(但对 a'保密)，而 Alice 和 Bob 进行公开讨论来保留那些 $b=1$ 的$\{a,a'\}$对. 注意当 $a=a'$时，总有 $b=0$. 只有当 $a'=1-a$ 时，Bob 得到 $b=1$，而且概率是 1/2. 最终 a 是 Alice 的密钥，而 $1-a'$是 Bob 的密钥.

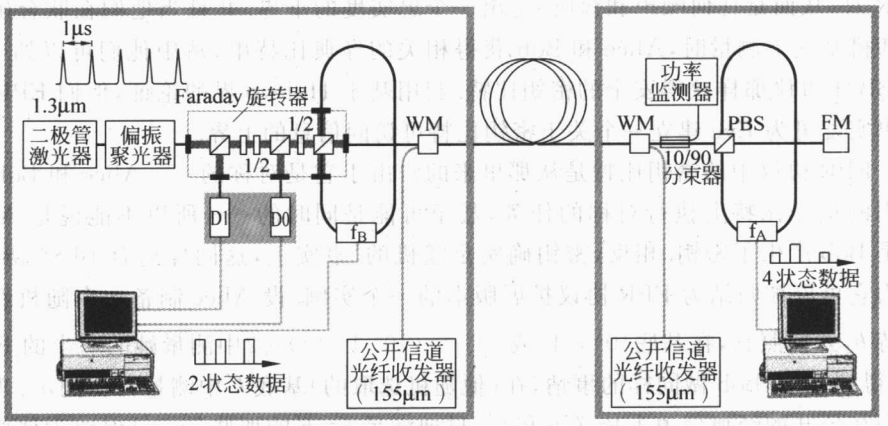
该协议，称为 B92(见章末“历史和进一步阅读的材料”)，凸现了完全区分非正交状态的不可能性如何处于量子密码术的中心地位. 正如在 BB84 中，因为任何窃听者都不可能在不干扰 Alice 和 Bob 最终保留的比特的情况下区分 Alice 的不同状态，该协议允许 Alice 和 Bob 创建共享的密钥比特同时，也定出他们通信过程中噪声和窃听的一个上界. 他们于是可以应用信息调和与保密增强来从他们得到的相关随机比特串中抽取安全比特.

练习 12.28　证明当 $b=1$ 时，a 和 a' 就彼此完全相关.

练习 12.29　给出一个使用 6 状态，X，Y 和 Z 的本征态的协议，并论证为什么它也是安全的. 对比 BB84 和 B92，讨论该协议对噪声和窃听的灵敏性.

盒子 12.7　实验量子密码术

量子密钥分配特别有趣也特别惊人，因为它很容易通过实验实现. 如下是 IBM 建造的，采样商用光纤元件在 10km 距离上传送密钥比特的一个系统的原理图.



Bob 起初用发射波长为 1.3μm 的二极管激光器产生若干强相干状态 $|\alpha\rangle$，并把它们传送给 Alice，Alice 衰减它们来产生（约）一个单光子. 她用水平和垂直偏振作为 $|0\rangle$ 和 $|1\rangle$ 状态，极化该光子到 BB84 协议中的四个状态之一. 她随后把该光子返回给 Bob，Bob 用偏振分析器在随机的基底中测量它. 通过采用这种光子穿过同一路径两次的特殊配置，设备可以做到光纤链路的缺陷（如路径长度的缓慢起伏和极化的漂移）的自补偿. Alice 和 Bob 随后选择他们使用了同样基底的那部分结果，调和他们的信息，并执行保密增强，在光子（在同一光纤上）的 1.55μm 波长的公开信道上通信. 密钥比特可以以每秒几百的速率交换. 最终，光源和检测器的提高应该会允许该速率有几个数量级的提高. 超过 40km 距离，也是搭建在远程通信光纤（在日内瓦湖底）中的量子密钥分配也已被演示出来.

3. EPR 协议

在 BB84 和 B92 协议中的密钥比特产生看起来是由 Alice 提出的，然而，事实上该密钥可被视为来源于包含纠缠性质的一个基本随机过程. 这由下述协议说明.

设 Alice 和 Bob 共享一组 n 个处于状态，

$$\frac{|00\rangle+|11\rangle}{\sqrt{2}} \tag{12.189}$$

的量子比特纠缠对，这类状态被称为 EPR 对. 可以从很多不同途径得到这些状态. 例如，Alice 可以制备这些对并随后把每对的一半发送给 Bob，或反之；或者，第三方可以制备这些对并把它们的一半发送给 Alice 和 Bob；或者，他们可以在很久以前相会并共享它们，并存储到如今. Alice 和 Bob 于是可以选择 EPR 对的一个随机子集，并测试看它们是否违背了 Bell 不等式(式(2.225)，2.6 节)，或进行其他忠实度的适当测试. 通过测试则确认他们继续持有充分纯的纠缠量子状态，为剩余 EPR 对(从而是任何噪声和窃听)定出一个忠实度的下界. 并且当他们在联合确定的随机基底下测量时，Alice 和 Bob 获得相关的经典比特串，从中他们可以如 B92 和 BB84 协议那样获得安全的密钥比特. 利用基于 Holevo 界的论证，他们 EPR 对的忠实度可为 Eve 建立一个关于密钥比特可访问信息的上界.

EPR 协议中的密钥比特是从那里来的？由于它是对称的——Alice 和 Bob 在他们的量子比特上执行对称的任务，甚至可能是同时的——所以不能说是 Alice 还是 Bob 产生了密钥. 相反，密钥确实是随机的. 事实上，这同样适合 BB84 协议，因为它可以被归结为 EPR 协议扩展版本的一个实例. 设 Alice 制备一个随机经典比特 b，并按照它，在基底 $|0\rangle$，$|1\rangle$ 或 $|\pm\rangle=(|0\rangle\pm|1\rangle)/\sqrt{2}$ 中测量她的一半的 EPR 对，得到 a. 让 Bob 做同样的事情，在(他随机选取的)基底 b' 中测量并得到 a'. 现在他们在公开的经典信道上传送 b 和 b'，只把满足 $b=b'$ 的那些 $\{a,a'\}$ 保留为他们的密钥. 直到 Alice 和 Bob 对他们的 EPR 部分进行测量之前，注意该密钥是*未确定*的. 对 B92 协议可作类似的观察. 出于这个原因，量子密码术有时不被认为是密钥交换或传送，而是被看作密钥生成，因为根本上来说 Alice 和 Bob 在协议最终完成之前都不能事先确定密钥.

12.6.4 保密性与量子相干信息

止此，我们已描述了 QKD 的基本协议并声称它是安全的，但我们没有提供定量的界，它安全到什么程度？事实上，本章已讨论的量子信息的基本量化度量与下面所讨论的量子密码术原则上的可得到安全性之间存在一个有趣和基本的联系.

量子相干信息 $I(\rho,\varepsilon)$ 给出量子信道传送保密信息能力的一个下界. 在最一般的情况下，Alice 制备状态 ρ_k^A，其中 $k=0,1,\cdots$，表示她可能发送的状态，每个的概率是 p_k. Bob 收到状态 $\rho_k^B=\varepsilon(\rho_k^A)$，它可能因为有信道噪声或窃听者 Eve 而与 ρ_k^A 不同. Bob 进行测量的结果与 Alice 的值 k 之间的互信息 $H_{\text{bob; alice}}$，由 Holevo 界式(12.6)给出上界，

$$H_{\text{bob; alice}} \leqslant \chi^B = S(\rho^B)-\sum_k p_k S(\rho_k^B) \tag{12.190}$$

其中 $\rho^B=\sum_k p_k\rho_k^B$. 类似地，Eve 的互信息有上界，

$$H_{\text{eve; alice}}\leqslant\chi^E=S(\rho^E)-\sum_k p_k S(\rho_k^E) \tag{12.191}$$

因为 Bob 相对 Eve 的过量信息（至少在某一个阈值之上）原则上可以被 Bob 和 Alice 利用，通过例如保密增强的技术来蒸馏一个共享的密钥，定义量

$$\mathscr{P}=\sup[H_{\text{bob; alice}}-H_{\text{eve; alice}}] \tag{12.192}$$

为信道可靠的保密性是有意义的，其中上确界取遍 Alice 和 Bob 对信道可用的全体策略. 这是 Bob 相对于 Eve 关于 Alice 的量子信号可获得的最大额外经典信息. 根据 HSW 定理，Alice 和 Bob 可采用一个策略，使得 $H_{\text{bob; alice}}=\chi^B$，而对 Eve 可能采取的任何策略，$H_{\text{eve; alice}}\leqslant\chi^E$. 因此，对适当选取的策略成立，$\mathscr{P}\geqslant\chi^B-\chi^E$.

从练习 12.11 可知，通过假定 Alice 的信号状态 $\rho_k^A=|\psi_k^A\rangle\langle\psi_k^A|$ 均为纯态且开始与 Eve 没有纠缠，而 Eve 起初处于某个状态 $|0^E\rangle$（不失一般性也可设为纯态），可以得到保密性 $\mathscr{P}$ 的一个下界. 一般地，从 Alice 到 Bob 的信道将包含与除 Eve 之外的某种环境的相互作用，但 Eve 有最大限度的优势，因此全部这类相互作用都可以归因于她，使得经过传输，Eve 和 Bob 收到的最终联合状态为

$$|\psi^{EB}\rangle=U|\psi_k^A\rangle|0^E\rangle \tag{12.193}$$

因为这是一个纯态，约化密度矩阵 ρ_k^E 和 ρ_k^B 将具有相同的非零特征值和相同的熵，即 $S(\rho_k^E)=S(\rho_k^B)$，于是，

$$\mathscr{P}\geqslant\chi^B-\chi^E \tag{12.194}$$

$$=S(\rho^B)-\sum_k p_k S(\rho_k^B)-S(\rho^E)+\sum_k p_k S(\rho_k^E) \tag{12.195}$$

$$=S(\rho^B)-S(\rho^E) \tag{12.196}$$

$$=I(\rho,\varepsilon) \tag{12.197}$$

即信道 ε 一个可靠的保密性下界由式(12.118)所定义的量子相干信息 $I(\rho,\varepsilon)$ 给出. 注意这个结果并非针对某个特定协议（它本身可能有自己的安全漏洞）. 同时，协议必须进行测试以实际确定信道 ε 的属性，之后这个界才能应用，而在此处的计算中并未考虑这些. 故尽管我们这里给出的信息论估界非常优雅，在量化 QKD 的安全性之前我们仍有工作要做！

12.6.5　量子密钥分配的安全性

量子密钥分配安全到什么程度？因为通信状态干扰的不可避免性，在窃听者得到的信息上，我们有理由相信 QKD 的安全性. 然而，为断定协议是真正安全的，我们所需要的是安全性的一个可量化的定义，该定义显式地给出在给定

Alice 和 Bob 努力的某种度量条件下，Eve 关于最终密钥知识的估界. 下述准则是可接受的：QKD 协议定义为安全，如果对由 Alice 和 Bob 选择的任意安全性参数 $s>0$ 和 $l>0$，以及对任意的窃听策略，该协议要么中止，要么至少以 $1-O(2^{-s})$ 概率成功，且保证 Eve 关于最终密钥的互信息小于 2^{-l}. 密钥串也必须是本质随机的.

在这最后一节里，我们给出 BB84 为安全的证明要点. 该证明作为本章的结束是合适的，因为它优雅地使用了量子信息的许多概念来提供一个无法比拟的简洁而清晰的论证. 该证明的起源本质上来自这样的观察：经过信息调和与保密增强，最终可获得的密钥比率事实上与 CSS 码(10.4.2 节)在带噪声通信信道上可达到的量子比特传输率一致.

主要思想梗概如下. 如果 Eve 每次只攻击一个量子比特的传输，那么可相对直接地逐一建立 BB84，B92 和 EPR 协议是安全的结论. 困难在于对付组合攻击的可能性，其中 Eve 操纵并可能存储被传输量子比特的大块. 为处理这个问题，我们需要更一般和更有力的论证. 假设我们通过某种方式知道 Eve 在每块从不引入超过 t 量子比特的差错，那么 Alice 可以把她的量子比特编码为 t 纠错量子码，使得 Eve 所有的干扰可以在 Bob 解码时被消除. 为使这样可行，必须建立两点：首先，如何定出 t 的上界？实际上可以通过对信道以合适方式采样完成，给我们一个安全的协议，甚至能对抗组合攻击. 遗憾的是，该协议一般要求容错量子计算机来对量子比特进行鲁棒编码和解码. 第二个挑战是，要选择量子码使得所有的序列编码、解码和测量都能用非量子计算或存储——仅仅是单量子比特制备和测量——进行. 用 CSS 码实现这个技巧(经过某种简化)，并且事实上正好给出 BB84 协议. 下面，我们从基于 QKD 协议的明显安全的 EPR 对出发，然后应用对这两个挑战的解答来系统地把初始协议简化为 BB84.

1. 安全 QKD 协议的要求

假设 Alice 有 n 对纠缠的量子比特，每个都处于状态

$$|\beta_{00}\rangle=\frac{|00\rangle+|11\rangle}{\sqrt{2}} \tag{12.198}$$

把这个状态记作 $|\beta_{00}\rangle^{\otimes n}$. Alice 把每对的一半传输给 Bob；由于信道上的噪声和窃听，结果状态可能是不纯的，并描述为密度矩阵 ρ. Alice 和 Bob 接着进行本地测量以获得一个密钥，如前所述. 下面的引理可用来说明，ρ 相对 $|\beta_{00}\rangle^{\otimes n}$ 的忠实度为 Eve 关于密钥的互信息设置了一个上界.

引理 12.19(高忠实度蕴含低熵) 如果 $F(\rho,|\beta_{00}\rangle^{\otimes n})^2>1-2^{-s}$，那么 $S(\rho)<(2n+s+1/\ln 2)2^{-s}+O(2^{-2s})$.

证 如果 $F(\rho,|\beta_{00}\rangle^{\otimes n})^2={}^{\otimes n}\langle\beta_{00}|\rho|\beta_{00}\rangle^{\otimes n}>1-2^{-s}$，那么 ρ 的最大特征值必然大于 $1-2^{s}$. 于是，ρ 的熵以具有对角元 $1-2^{-s}$，$2^{-s}/(2^{2n}-1)$，$2^{-s}/(2^{2n}-1)$，…，

$2^{-s}/(2^{2n}-1)$的对角型密度矩阵 $\rho_{\max}$的熵为上界. 即 $\rho_{\max}$具有最大项 $1-2^{s}$，且其他概率均匀地分配在剩余的 $2^{2n}-1$ 项上. 因为

$$S(\rho_{\max}) = -(1-2^{-s})\log(1-2^{-s}) - 2^{-s}\log\frac{2^{-s}}{2^{2n}-1} \tag{12.199}$$

所期望的结果成立. □

由 Holevo 界式(12.6)，$S(\rho)$是 Eve 可访问的信息的上界，这个信息来源于 Alice 和 Bob 对 ρ 的测量. 这意味着，如果一个 QKD 协议能够为 Alice 和 Bob 提供忠实度至少为 $1-2^{-s}$(以高的概率)的 EPR 对，那么它是安全的.

练习 12.30　简化式(12.199)以获得如引理所叙述的 $S(\rho)$的公式.

练习 12.31　读者或许不清楚为什么 $S(\rho)$是 Eve 关于 Alice 和 Bob 测量结果互信息的界，证明这可在假定 Eve 的最坏情形，即给予她对信道的完全控制时得到.

2. 随机采样可建立窃听的上界

一个协议如何能为 Alice 和 Bob 的 EPR 对的忠实度设置下界？这里的关键思想是一个经典论证，它基于随机采样，我们已在 BB84 协议的描述中遇到(练习 12.27). 然而当考虑量子测量结果时，基于经典概率的论证却不一定成立，Bell 不等式(2.6 节)就是一个生动的例子. 另一方面，如果测量的可观测量仅参照一个基底，量子实验的确允许经典解释. 而且幸运的是，为 Alice 和 Bob 对他们 EPR 对的忠实度估界，恰好要求仅在一个基底中进行测量.

按照式(10.14)，通过带噪声量子信道传输一个量子比特可描述为如下四种情况之一：什么也不发生(I)、比特翻转(X)、相位翻转(Z)或比特和相位翻转的组合(Y). 回忆 Bell 基底由如下四个状态定义：

$$|\beta_{00}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}, \quad |\beta_{10}\rangle = \frac{|00\rangle-|11\rangle}{\sqrt{2}},$$
$$|\beta_{01}\rangle = \frac{|01\rangle+|10\rangle}{\sqrt{2}}, \quad |\beta_{11}\rangle = \frac{|01\rangle-|10\rangle}{\sqrt{2}} \tag{12.200}$$

令每对中第二量子比特是 Alice 发送给 Bob 的. 如果该量子比特发生了比特翻转差错，那么$|\beta_{00}\rangle$就被变成$|\beta_{01}\rangle$. 类似地，相位翻转给出$|\beta_{10}\rangle$，两种差错的组合给出$|\beta_{11}\rangle$(除去一个无关的全局相位). 检测是否发生比特翻转的一个自然测量由 $\Pi_{\text{bf}} = |\beta_{01}\rangle\langle\beta_{01}| + |\beta_{11}\rangle\langle\beta_{11}|$和 $I-\Pi_{\text{bf}}$描述；类似地，由 $\Pi_{\text{pf}} = |\beta_{10}\rangle\langle\beta_{10}| + |\beta_{11}\rangle\langle\beta_{11}|$和 $I-\Pi_{\text{pf}}$描述的投影检测相位翻转. 因为这些测量与 Bell 基底对易，它们的结果服从经典的概率推理. 事实上，任何与 Bell 基底对易的测量也都满足同样的经典推理.

更确切地，Alice 和 Bob 通过随机采样其一个子集来对他们的 EPR 对忠实度估界. 假设 Alice 发送 $2n$ 个 EPR 对的一半给 Bob，他们随后随机选择其中 n 个，并

通过 Π_{bf} 或 Π_{pf}（同样随机选择）联合测量来检查这些量子比特. 采用与 BB84 中随机采样测试同样的经典推理（练习 12.27），如果检测到 δn 个比特或相位翻转差错，那么剩余的 n 个 EPR 对将指数地确定具有同样数目的差错，如果它们也在 Bell 基底中被测量.

Bell 态是非局部的，而在 Bell 基底中的一般测量要求非本地运算，因而可能是困难. 不过幸好，现有的方案中没有这个要求，因为 $\Pi_{\mathrm{bf}}=(I\otimes I-Z\otimes Z)/2$ 且 $\Pi_{\mathrm{pf}}=(I\otimes I-X\otimes X)/2$. 因此，Alice 和 Bob 可以用 Pauli 算子的本地测量来进行期望的检查，要么都用 Z 测量，要么都用 X 测量.

练习 12.32 注意 Alice 和 Bob 进行本地测量，如 $I\otimes X$ 和 $X\otimes I$，与 Bell 基底不对易. 证明尽管如此，Alice 和 Bob 从他们测量得到的统计与他们实际测量了 Π_{bf} 和 Π_{pf} 得到的统计相同.

3. 修改的 Lo-Chau 协议

因此在 Bell 基底中的随机采样为 Alice 和 Bob 提供对理想状态 $|\beta_{00}\rangle^{\otimes n}$ 具有已知忠实度的 EPR 对 ρ，并如前所述，这限制了 Eve 在 ρ 上可能进行的任何测量的互信息. 不过要 ρ 对密钥生成有用，Alice 和 Bob 必须减少 Eve 关于他们状态的互信息，小到指数量级. 这个任务可由应用经典的保密增强到他们的测量结果来完成. 等价地，Alice 和 Bob 可以首先进行纠缠蒸馏，如 12.5.2 节介绍的，来获得对某个 $m<n$ 非常接近 $|\beta_{00}\rangle^{\otimes m}$ 的 ρ'，然后测量最终状态. 这类量子保密增强将对我们有用.

大概的论述如下. 纠缠蒸馏可通过执行量子纠错来完成. 因为 ρ 几乎确定地含有 δn 个差错，将这些量子比特用 δn 量子纠错码编码，它允许最多为 δn 个差错被完全纠正. 如我们在 10.5.5 节和 10.5.8 节看到的，如果使用 $[n,m]$ 移定子码，那么编码、差错症状测量和差错恢复可以通过 Pauli 算子的测量来决定，这些算子决定于编码的校验矩阵的行. Alice 和 Bob 只是在他们相应的 n 量子比特即 ρ 的一半上进行相同测量和恢复运算，产生一个被纠正的状态，相对于 $|\beta_{00}\rangle^{\otimes m}$ 忠实度在 1 减去出现多于 δn 个差错的概率数量级. 通过构造，因为 Alice 和 Bob 执行相同的任务，差错症状测量实际上与 Bell 基底对易.

把随机采样和纠缠蒸馏结合起来就给出修改的 Lo-Chau 协议，如图 12.14 所示. 下面是一些对这个协议的注释. 第 3 步和第 7 步的随机 Hadamard 变换对 Eve 在 X 和 Z 基底中（因此引起 X 和 Z 差错）编码信息的检测产生了一个对称. 它们也允许在校验量子比特上进行 Π_{bf} 或 Π_{pf} 测量. 第 9 步的特殊过程可以从任何稳定子码的情形得到证实，如练习 12.34. 对 CSS 码的 Gilbert-Varshamov 界，即式(10.74)，表明大的块长存在好的量子编码，因此对一个 δn 纠错的 $[n,m]$ 量子码，可选充分大的 n，使安全性准则满足.

QKD 协议：修改的 Lo-Chau 协议

1：Alice 创建 $2n$ 个处于 $|\beta_{00}\rangle^{\times 2n}$ 状态的 EPR 对.

2：Alice 随机从 $2n$ 个 EPR 对中选择 n 个用于检查 Eve 干扰的校验. 她还未用它们做什么.

3：Alice 随机选择一个 $2n$ 比特串 b，并对 b 为 1 的每个对的第二量子比特执行 Hadamard 变换.

4：Alice 把每对的第二量子比特发送给 Bob.

5：Bob 接收到这些量子比特并公开宣布这一事实.

6：Alice 宣布 b 以及哪 n 个量子比特将作为校验比特.

7：Bob 在 b 为 1 的量子比特上执行 Hadamard 变换.

8：Alice 和 Bob 各自在 $|0\rangle,|1\rangle$ 基底中测量他们的 n 校验量子比特，并公开分享结果. 如果结果有多于 t 个不同，他们终止协议.

9：Alice 和 Bob 按照预先确定的最多纠正 t 个差错的 $[n,m]$ 量子码的校验阵测量他们的剩余 n 量子比特. 他们分享结果，计算差错症状，并纠正状态，得到 m 个接近完全的 EPR 对.

10：Alice 和 Bob 在 $|0\rangle,|1\rangle$ 基底中测量 m 个 EPR 对，以获得共享的密钥.

图 12.14　一个安全的 QKD 协议，因为使用了完美的量子计算机、纠错和 EPR 对的随机测试.

练习 12.33　令 $\{M_1,M_2,\cdots,M_n\}$ 为一组测量观测量，当输入状态 ρ 被测量时，产生相应结果 X_i. 证明如果 $[M_i,M_j]=0$，即它们彼此对易，则随机变量 X_i 服从经典的概率推理.

练习 12.34（使用纠错的纠缠蒸馏）　10.5.8 节中，我们看到，可通过在任意 n 量子比特量子状态上的测量它的生成元 $g_1,\cdots,g_{n-m}$，再应用 Pauli 运算把结果变到生成元的共同本征态 $+1$，来构造 $[n,m]$ 量子比特稳定子码的码字. 基此想法，如果我们从处于 $|\beta_{00}\rangle^{\otimes n}$ 状态的 n 个 EPR 对出发，在两个 n 量子比特对的一半进行相同的生成元测量，接着用 Pauli 运算来纠正这些对在测量结果上的不同，然后我们得到一个经过编码的 $|\beta_{00}\rangle^{\otimes m}$ 状态. 再证明如果稳定子码最多纠正 δn 个差错，那么即使一个 n 量子比特部分发生了 δn 个差错，我们仍能得到 $|\beta_{00}\rangle^{\otimes m}$.

4. 量子纠错协议

修改的 Lo-Chau 协议用量子纠错来进行纠缠蒸馏，而且本质上建立在 EPR 协议上. 纠缠是一种脆弱的资源，而量子纠错一般要求鲁棒量子计算机，这带来实现上的挑战. 不过幸而，这个协议可以通过一系列步骤被系统地简化，可证明每一步都不损失方案的安全性. 让我们从取消分配 EPR 对的需要开始.

注意 Alice 在修改的 Lo-Chau 协议最后进行的测量可以在最开始时进行，而不带来其他地方的状态的任何改变. 在第 8 步，Alice 对她部分的校验 EPR 对测量把这些对变成 n 单量子比特，因此 Alcie 可直接发送单量子比特，来代替发送纠缠状态. 这给我们提供了如下修改的步骤：

1′：Alice 创建 n 个随机校验比特和处于状态 $|\beta_{00}\rangle^{\otimes n}$ 的 n 个 EPR 对，她还按校验比特把 n 量子比特编码为 $|0\rangle$ 或 $|1\rangle$.

2′：Alice 随机选择 n 个位置(从 $2n$ 中)，把校验量子比特放置到这些位置，把每个 EPR 对的一半放置到其余位置.

8′：Bob 在$|0\rangle$，$|1\rangle$基底中测量 n 个校验量子比特，并公开与 Alice 共享结果. 如果不同多于 t 个，他们就中止协议.

类似地，Alice 在第 9 步和第 10 步的测量使 EPR 对变成编码为随机量子码的随机量子比特. 这可以从下面的方式看出. C_1 在 C_2 上的$[n,m]$ CSS 码 $\mathrm{CSS}(C_1, C_2)$，是一个特别方便的编码选择，也是本节其余部分将使用的，它把 m 量子比特编码成 n 量子比特并最多纠正 t 个差错. 回忆 10.4.2 节，对此编码，H_1 和 $H_2^{\perp}$ 是对应于经典编码 C_1 和 $C_2^{\perp}$ 的奇偶校验矩阵，其中每个码字状态为

$$\frac{1}{\sqrt{|C_2|}}\sum_{w\in C_2}|v_k+w\rangle \tag{12.201}$$

v_k 为 C_2 在 C_1 中的 2^m 个陪集中的一个表示(v_k 的符号表示由密钥串 k 索引的向量 v). 再回忆存在一族等价于此码的编码 $\mathrm{CSS}_{z,x}(C_1, C_2)$，具有码字状态

$$|\xi_{v_k,z,x}\rangle=\frac{1}{\sqrt{|C_2|}}\sum_{w\in C_2}(-1)^{zw}|v_k+w+x\rangle \tag{12.202}$$

这些状态形成 2^n 维 Hilbert 空间的一个标准正交基底(见练习 12.35)，于是可将 Alice 的 n 个 EPR 对状态写作

$$|\beta_{00}\rangle^{\otimes n}=\sum_{j=0}^{2^n}|j\rangle|j\rangle=\sum_{v_k,z,x}|\xi_{v_k,z,x}\rangle|\xi_{v_k,z,x}\rangle \tag{12.203}$$

注意在这个表达式中，我们把标号分成了两个右矢，其中第一个表示 Alice 持有的量子比特，第二个表示发送给 Bob 的量子比特. 当 Alice 测量对应于第 9 步中她量子比特上的 H_1 和 $H_2^{\perp}$ 的稳定子生成元时，她得到 x 和 z 的随机值，并且类似地，在第 10 步的最终测量值给出 v_k 的一个随机选择. 于是剩余的 n 量子比特处于状态$|\xi_{v_k,z,x}\rangle$，为v_k 在 $\mathrm{CSS}_{z,x}(C_1, C_2)$中的码字，这正是一个 2^m 量子比特状态$|k\rangle$的编码结果. 因此，如上所断言，Alice 的测量产生编码到随机码中的随机量子比特.

因此，代替发送 EPR 对的一半，Alice 可以等价地随机选择 x，z 和 k，然后把$|k\rangle$编码到编码 $\mathrm{CSS}_{z,x}(C_1, C_2)$，并把编码后的 n 量子比特发送给 Bob. 下面给出修改的步骤：

1″：Alice 创建 n 个随机校验比特、一个随机 m 比特密钥 k、两个随机 n 比特串 x 和 z，她把$|k\rangle$编码到码 $\mathrm{CSS}_{z,x}(C_1, C_2)$，又按校验比特把 n 量子比特编码为$|0\rangle$或$|1\rangle$.

2″：Alice 随机选择 n 个位置(从 $2n$ 中)，在这些位置放置校验量子比特，对剩余位置中的量子比特进行编码.

6′：Alice 宣布 b，x，z 以及提供校验比特的相应 n 个量子比特.

9′：Bob 从 $\mathrm{CSS}_{z,x}(C_1, C_2)$解码剩余的 n 量子比特.

10′：Bob 测量他的量子比特来得到共享的密钥 k.

得到的方案，称为 CSS 码协议，如图 12.15 所示.

QKD 协议：CSS 码

1″：Alice 创建 n 给随机校验比特，一个随机 m 比特密钥 k，和两个随机 n 比特串 x 和 z. 她把 $|k\rangle$ 编码到码 $\mathrm{CSS}_{z,x}(C_1,C_2)$. 她又按校验比特把 n 量子比特编码为 $|0\rangle$ 或 $|1\rangle$.

2″：Alice 随机选择 n 个位置（从 $2n$ 中），并在这些位置放置校验量子比特，对剩余位置中的量子比特进行编码.

3：Alice 随机选择一个 $2n$ 比特串 b，并对 b 为 1 的每个对的第二量子比特执行 Hadamard 变换.

4：Alice 把每对的第二量子比特发送给 Bob.

5：Bob 接收到这些量子比特并公开宣布这一事实.

6′：Alice 宣布 b,x,z 以及哪 n 个量子比特提供校验比特.

7：Bob 在 b 为 1 的量子比特上执行 Hadamard 变换.

8′：Bob 在 $|0\rangle,|1\rangle$ 基底中测量 n 个校验量子比特，并公开把结果与 Alice 共享. 如果它们有多于 t 个不同，他们就中止协议.

9′：Bob 从 $\mathrm{CSS}_{z,x}(C_1,C_2)$ 解码剩余的 n 量子比特.

10′：Bob 测量他的量子比特来得到共享的密钥 k.

图 12.15　一个安全的 QKD 协议，因为是通过 CSS 码简化的修改的 Lo-Chau 协议.

练习 12.35　证明在式(12.202)中定义的状态 $|\xi_{v_k,z,x}\rangle$ 形成一个 2^n 维 Hilbert 空间的标准正交基底，即

$$\sum_{v_k,z,x}|\xi_{v_k,z,x}\rangle\langle\xi_{v_k,z,x}|=I \tag{12.204}$$

提示　对 $[n,k_1]$ 码 C_1，$[n,k_2]$ 码 C_2 及 $m=k_1-k_2$，注意有 2^m 个不同的 v_k 值，2^{n-k_1} 不同的 x，2^{k_2} 不同的 z.

练习 12.36　验证式(12.203).

练习 12.37　这里有另外一个方法，来理解为什么在第 9 步和第 10 步，Alice 测量能把 EPR 对变成编码到一个随机量子编码的随机量子比特. 设 Alice 有一个 EPR 对 $(|00\rangle+|11\rangle)/\sqrt{2}$，证明如果她在 X 基底中测量第一量子比特，那么第二量子比特变成由测量结果决定的 X 的一个本征态. 类似地，证明如果她在 Z 基底中测量，那么第二量子比特被变成由测量结果标记的 Z 的本征态. 利用这个观察和 10.5.8 节的结果，得出结论 Alice 在 EPR 对她的部分测量 H_1，$H_2^{\perp}$ 和 $\overline{Z}$ 导致由她的测量结果决定的一个 $\mathrm{CSS}_{z,x}(C_1,C_2)$ 的随机码字.

5. 简化为 BB84

由于直接从修改的 Lo-Chau 协议简化而来，CSS 码 QKD 协议是安全的，并且因为它根本没有直接使用 EPR 对，所以它很简单. 遗憾的是，它仍不完全令人满意，因为它要求用完美的量子计算来执行编码和解码（而不仅仅是单量子比特制备和测

量),且 Bob 等待与 Alice 通信时,需要在量子内存中临时存储量子比特.不过 CSS 码的应用消除了这两项要求,实质上是因为它们解耦了相位翻转纠错和比特翻转纠错.

首先注意,Bob 紧接着解码过程,在 Z 基底中测量他的量子比特;因此 Alice 发送的相位纠错信息 z 是不必要的.于是因为 C_1 和 C_2 是经典编码,不需要解码后再测量,他可以立刻测量以得到 $v_k+w+x+\epsilon$(其中 ϵ 代表由信道和 Eve 带来的可能差错),然后进行经典性解码:他减去 Alice 宣布的 x 值,然后纠正结果到 C_1 中的一个码字,如果没有超出码的距离,它肯定是 v_k+w.最终密钥 k 是 v_k+w+C_2 在 C_1 中的陪集(见《量子计算和量子信息(一)》附录 B 对陪集的解释和这里的概念).这给出:

9″:Bob 测量剩余的量子比特以得到 $v_k+w+x+\epsilon$,并从结果中减去 x,用码 C_1 来纠正它以得到 v_k+w.

10″:Bob 计算 v_k+w+C_2 在 C_1 中的陪集以得到密钥 k.

其次,因为 Alice 不需要泄露 z,她实际上发送的状态是一个混合态,为 z 的随机值上的平均,

$$\rho_{v_k,x}=\frac{1}{2^n}\sum_z |\xi_{v_k,z,x}\rangle\langle\xi_{v_k,z,x}| \tag{12.205}$$

$$=\frac{1}{2^n|C_2|}\sum_{w_1,w_2\in C_2}(-1)^{z(w_1+w_2)}|v_k+w_1+x\rangle\langle v_k+w_2+x| \tag{12.206}$$

$$=\frac{1}{|C_2|}\sum_{w\in C_2}|v_k+w+x\rangle\langle v_k+w+x| \tag{12.207}$$

该状态很容易创建:Alice 只需要经典地随机选择 $w\in C_2$,并用她随机确定的 x 和 k 构造 $|v_k+w+x\rangle$.于是我们有

1‴:Alice 创建 n 个随机校验比特、一个随机 n 比特串 x、一个随机的 $v_k\in C_1/C_2$ 和一个随机的 $w\in C_2$.她按照 v_k+w+x 把 n 量子比特编码为 $|0\rangle$ 或 $|1\rangle$,并且类似地按照校验比特安排 n 量子比特.

通过对步骤 6′稍作变化,可进一步简化步骤 1‴和 9″.目前 Alice 发送 $|v_k+w+x\rangle$,Bob 接收并测量得到 $v_k+w+x+\epsilon$,然后 Alice 发送 x,Bob 减去它得到 $v_k+w+\epsilon$.但如果 Alice 选择 $v_k\in C_1$(而非 C_1/C_2),那么 w 就是不必要的.进而,v_k+x就是完全随机的 n 比特串,而且不使用上述协议也是等价的,即如果 Alice 随机选择 x,发送 $|x\rangle$,Bob 接收并测量以得到 $x+\epsilon$,然后 Alice 发送 $x-v_k$,Bob 减去它,得到$v_k+\epsilon$.现在,随机校验比特和编码比特之间不存在任何差别!这给出:

1⁗:Alice 选择随机的 $v_k\in C_1$,并按照 $2n$ 随机比特创建处于状态 $|0\rangle$ 或 $|1\rangle$ 的 $2n$ 量子比特.

2‴:Alice 随机选择 n 个位置(从 $2n$ 中),指定它们为校验量子比特,其余的为 $|x\rangle$.

6″：Alice 宣布 $b, x-v_k$，以及提供校验比特的相应 n 个量子比特.

9‴：Bob 测量剩余的量子比特以得到 $x+\varepsilon$，并从结果中减去 $x-v_k$，用 C_1 对其纠正以获得 v_k.

10″：Alice 和 Bob 计算 v_k+C_2 在 C_1 中的陪集以获得密钥 k.

接下来，注意 Alice 不需要执行 Hadamard 运算（尽管实际中，用光子不难完成单量子比特运算）. 根据 b 的比特，她可以或者在 $|0\rangle, |1\rangle$ (Z) 基底，或者在 $|+\rangle, |-\rangle$ (X) 基底中直接对她的量子比特进行编码：

1‴″：Alice 创建 $(4+\delta)n$ 个随机比特. 对每个比特，她根据一个随机比特串 b，要么在 $|0\rangle, |1\rangle$ 基底中，要么在 $|+\rangle, |-\rangle$ 基底中创建一个量子比特.

我们几乎就要完成任务了：编码和解码现在都是经典地执行了. 剩下的问题是消除量子内存的需要. 为解决这个问题，设 Bob 从 Alice 收到量子比特后立即进行测量，随机选择在 X 或 Z 基底中进行测量. 当 Alice 随后宣布 b 时，他们可以只保留那些其基底实际上相同的比特. 这使 Bob 可以完全放弃他的量子存储装置. 注意他们以高的概率丢弃其一半的比特，故为获得与前面相同数目的密钥比特，他们可以从略微（比如 δ）多出原始随机比特数目两倍开始. 当然，Alice 现在必须推迟选择哪些比特是校验比特，直到丢弃步骤之后. 这给出了我们的最终协议，如图 12.16所示. 该协议除了一点表面的差别，恰好正是 BB84. 注意，经典编码 C_1 的使用如何执行信息调和，以及计算 v_k+C_2 在 C_1 中的陪集如何执行保密增强（见 12.6.2 节）.

QKD 协议：CSS 码

1：Alice 创建 $(4+\delta)n$ 个随机比特.

2：对每个比特，她按照一个随机比特串 b，要么在 Z 基，要么在 X 基中创建一个量子比特.

3：Alice 将得到的量子比特发送给 Bob.

4：Alice 随机选择一个 $v_k \in C_1$.

5：Bob 接收到这些量子比特并公开宣布这一事实，并随机地在 Z 或 X 基中测量每个量子比特.

6：Alice 宣布 b.

7：Alice 和 Bob 丢弃那些 Bob 在不同于 b 的基中测量的比特. 以高的概率，有至少 $2n$ 个比特剩下；否则，终止协议. Alice 随机地决定 $2n$ 个继续使用的比特，并随机地选择其中 n 个作为校验比特，并宣布这一选择.

8：Alice 和 Bob 公开比较他们的校验比特. 如果多于 t 个不同，他们就终止协议. Alice 剩下的是 n 比特串 x，而 Bob 是 $x+\varepsilon$.

9：Alice 宣布 $x-v_k$，Bob 从他的结果中将其减去，用码 C_1 来纠正以得到 v_k.

10：Alice 和 Bob 计算 v_k+C_2 在 C_1 中的培集以得到密钥 k.

图 12.16　通过 CSS 码协议的系统性简化得到的最终 QKD 协议，它与 BB84 完全相同（只有一点表面差别）. 为清楚起见，我们略去了′号.

总结起来，我们已系统性地证明 BB84 量子密钥分配协议的安全性. 我们从一个明显安全的要求完美量子计算和量子内存的方案出发，并系统地把它简化到 BB84. 由于仅作出明显不改变 Eve 量子状态（以所有公布的经典信息为条件）的修

改，我们推出 BB84 是安全的. 很自然，有些需要说明. 该证明仅适合一种理想情形，被发送的状态如所描述的情形. 在实践中，量子比特源是不完美的；例如这类源常常是调整到近似产生代表量子比特(如 7.4.1 节所描述)的单光子激光. 而且，该证明没有给出 Alice 和 Bob 解码代价的界；实际的密钥分配中，C_1 必须是有效可解码的. 该证明也没有提供可容忍的窃听的上界；它使用的 CSS 编码，并不是最优的. 据估计，用类似于 BB84 的协议 11%的比特和相位差错率是可接受的，但借助量子计算机的编码和解码，或许可容忍更高的差错率. 量子密码术的终极能力是一个有趣的公开问题，我们期望关于计算和通信的物理极限这样基本的问题继续吸引和挑战未来的研究者.

练习 12.38 证明如果我们有能力区分非正交状态，那么就可能破坏 BB84，实际上是我们所描述的所有 QKD 协议的安全性.

问题 12.1 本问题中，我们来看 Holevo 界的另一个证明. 定义Holevo chi量：

$$\chi \equiv S(\rho) - \sum_x p_x S(\rho_x) \tag{12.208}$$

(1) 设量子系统由两部分 A 和 B 组成，证明

$$\chi_A \leqslant \chi_{AB} \tag{12.209}$$

(提示：引入与 AB 相关的一个附加系统，应用强次可加性)

(2) 令 ε 为一量子运算，用前面的结果证明

$$\chi' \equiv S(\varepsilon(\rho)) - \sum_x p_x S(\varepsilon(\rho_x)) \leqslant \chi \equiv S(\rho) - \sum_x p_x S(\rho_x) \tag{12.210}$$

即，Holevo chi 量在量子运算下递减. 这是一个有独立重要性和用途的事实.

(3) 令 E_y 为一组 POVM 元，在所考虑的量子系统增加一个具有标准正交基底 $|y\rangle$ 的设备系统 M，定义一量子运算为

$$\varepsilon(\rho \otimes |0\rangle\langle 0|) \equiv \sum_y \sqrt{E_y}\,\rho\sqrt{E_y} \otimes |y\rangle\langle y| \tag{12.211}$$

其中 $|0\rangle$ 是 M 的某个标准纯态. 证明 ε 作用后，$\chi_M = H(X:Y)$. 基此及前面的两个结果证明

$$H(X:Y) \leqslant S(\rho) - \sum_x p_x S(\rho_x) \tag{12.212}$$

即 Holevo 界.

问题 12.2 该结果是前面问题的推广. 通过证明对非正交纯态的克隆过程必然增加 χ，提供了一个不可克隆定理的证明.

问题 12.3 对固定的量子信源和比率 $R > S(\rho)$，设计一个量子线路实现比率 R 的压缩方案.

问题 12.4(线性禁止克隆) 假设我们有一个具有两个槽 A 和 B 的量子机器. 槽 A 为数据槽，始于未知量子状态 ρ，这是待复制的状态. 槽 B 为目标槽，始于

某个标准量子状态 σ. 我们假定候选的任何复制过程对初始状态是线性的，

$$\rho \otimes \sigma \to \varepsilon(\rho \otimes \sigma) = \rho \otimes \rho \tag{12.213}$$

其中 ε 是某个线性函数. 证明如果 $\rho_1 \neq \rho_2$ 是满足

$$\varepsilon(\rho_1 \otimes \sigma) = \rho_1 \otimes \rho_1 \tag{12.214}$$

$$\varepsilon(\rho_2 \otimes \sigma) = \rho_2 \otimes \rho_2 \tag{12.215}$$

的密度算子，那么 ρ_1 和 ρ_2 的任何混合都不能被该过程正确复制.

问题 12.5(量子信道的经典容量*)　积状态容量式(12.71)是带噪声量子信道对经典信息的容量，即当对信道允许有纠缠的输入时的容量吗?

问题 12.6(达到容量的方法*)　找出一种达到接近带噪声量子信道对经典信息的积状态容量式(12.71)的编码的有效构造.

问题 12.7(量子信道容量*)　找出一种方法，来评价对给定量子信道 ε 传输量子信息容量.

第 12 章的总结　量子信息论

- 不可克隆：不可能构造这样的量子设备，它对任意 $|\psi\rangle$，在给定 $|\psi\rangle$ 条件下，输出 $|\psi\rangle|\psi\rangle$.
- Holevo 界：当试图区分以概率 p_x 发送的量子状态 ρ_x 时的最大可访问经典信息是

$$H(X:Y) \leqslant \chi \equiv S\left(\sum_x p_x \rho_x\right) - \sum_x p_x S(\rho_x)$$

- Schumacher 量子带噪声信道编码定理：$S(\rho)$ 可被解释为，忠实表示一个由 ρ 描述的量子信源所需要的量子比特数.
- Holevo-Schumacher-Westmoreland 定理：带噪声量子信道 ε 对经典信息的容量由

$$C(\varepsilon) = \max_{\{p_x, |\psi_x\rangle\}} S\left(\sum_x p_x \varepsilon(|\psi_x\rangle\langle\psi_x|)\right) - \sum_x p_x S(\varepsilon(|\psi_x\rangle\langle\psi_x|)) \tag{12.216}$$

给出.

- 纠缠变换的控制不等式条件：当且仅当 $\lambda_\psi \prec \lambda_\varphi$，Alice 和 Bob 可以通过本地运算和经典通信把 $|\psi\rangle$ 变到 $|\varphi\rangle$，其中 λ_ψ 是 $|\psi\rangle$ 的约化密度矩阵的特征值的向量(λ_φ 类似).
- 纯态纠缠蒸馏和稀释：随 $n \to \infty$，Alice 和 Bob 能仅通过本地运算和经典通信，在联合状态 $|\psi\rangle$ 的 n 和 $nS(\rho)$ Bell 对之间转换，其中 ρ 是约化密度矩阵.
- 量子密码术：用类似 BB84 的协议，采用非正交量子状态的通信可使证明为安全的密钥分配成为可能. 信道上的窃听将引起可检测的差错率上升，因为获取信息意味着干扰.

历史和进一步阅读的材料

Cover 和 Thomas 的书[CT91]是经典信息论的极好引论. 读者在信息论上希望了解更深入但要可读的信息论，应该请教 Csiszár 和 Körner[CK81]. 在 20 世纪科学上最有影响论文中，Shannon 的原始论文也值得阅读，这些论文被 Shannon 和 Weaver 重印为单卷本[SW49]. Bennet 和 Shor[BS98]，Bennet 和 DiVincenzo[BD00]写了量子信息论方面很好的综述文章.

不可克隆定理出自 Dieks[Die82]，Wootters 和 Zurek[WZ82]，有非常多的研究致力于推广这些结果. 论文的大部分主要考虑克隆装置的不同方案，这些方案在某些特定方面是有趣的——它们优化克隆忠实度的某种度量，或希望克隆装置具有的某种属性. 我们此处不对这项研究给出全面的综述，但指出许多论文可以在互联网上 http://arXiv.org/位置的 quant-ph archive 找到. 特别有趣的一些论文，包括 Barnum，Cave，Fuchs，Jozsa 和 Schumacher[BCF+96]扩展不可克隆定理到混合态和非酉克隆的工作；Mor[Mor98]在复合系统状态克隆上的工作；Westmoreland 和 Schumacher[WS98]在克隆与超光速通信之间可能的等价性上的工作和 van Enk[van98b]的反驳.

Holevo 界是 Gordon 在 1964[Gor64]猜想的，且由 Holeov 于 1973[Hol73]所证明. 基于难证明的强次可加不等式，我们给出了概念上简单的证明，但是 Holevo 采用了更直接的方法，且已由 Fuchs 和 Caves[FC94]简化. 用强次可加不等式的证明来源于 Yuen 和 Ozawa[YO93]，还可以从 Schumacher，Westmoreland 和 Wootters[SWW96]处看到.

经典无噪声信道编码定理来自 Shannon[Sha48，SW49]. 量子无噪声信道编码定理来自 Schumacher[Sch95]，并在一篇开创性论文中得到描述. 这篇论文以系统的方式引入许多量子信息论的基本概念，包括信源、忠实度度量和量子状态作为信息论可处理的资源的概念. 这个最后的观察，虽然简单但非常深刻，归功于 Schumacher 和 Wootters 之间的一次交谈，冠以现在处处可见的术语量子比特，被 Schumacher 在该论文的引言中固定下来. Jozsa 和 Schumacher 的一篇论文[JS94]简化了 Schumacher 原来的方法；这篇文章早于文献[Sch95]发表，但是后来写的. 这些论文基于练习 12.8 中讨论的系综平均忠实度度量，而不是基于我们这里给出证明的纠缠忠实度，纠缠忠实度基于 Nielsen[Nie98]的方法. Schumacher，Schumacher 和 Jozsa 的原始论文中的一个小的漏洞由 Barnum，Fuchs，Jozsa 和 Schumacher[BFJS96]的工作填补. M. Horodecki[Hor97]随后为同一结果提供了更有力的证明，还指出了

发展混合态系综量子数据压缩理论的途径. 盒子 12.4 描述的压缩方案，是 Cover 的枚举编码方法[CT91]的量子对应物，来源于 Schumacher[Sch95]，它的量子线路由 Cleve 和 DiVincenzo[CD96]显式地给出. Braunstein, Fuchs, Gottesman 和 Lo 把它推广到提供 Huffman 编码的量子类比[BFGL98]，而 Chuang 和 Modha 把它推广到算术编码[CM00].

Holevo-Schumacher-Westmoreland(HSW)定理的历史很有趣. 它针对的问题最先是 Holevo[Hol79]在 1979 年讨论的，并取得了部分进展. 在不知道该工作的情况下，Hausladen, Jozsa, Schumacher, Westermoreland 和 Wootters[HJS+96]在 1996 年解决了该问题的一种特殊情况. 稍后，Holevo[Hol98]及 Schumacher 和 Westermoreland[SW97]独立地证明 HSW 定理给出带噪声量子信道对经典信息的积状态容量. Fuchs[Fuc97]描述了积状态容量的一些有趣的例子，其中最大化的容量表达式(12.71)的状态系综包含非正交成员. King 和 Ruskai[KR99]在证明积状态容量等于不限于积状态容量的问题上取得一些可喜的进展，但一般的问题仍未解决.

熵交换是 Lindblad[Lin91]定义的，而由 Schumacher[Sch96b]重新发现，他证明了量子 Fano 不等式. 相干信息由 Lloyd[Llo97]及 Schumacher 和 Nielsen[SN96]在带噪声量子信道容量讨论中引入；文献[SN96]证明了量子信息处理不等式. 练习 12.15 中提到的包含这些不等式的表可以在 Nielsen 的博士论文[Nie98]中找到. 尚未解决的确定量子信道容量的问题(问题 12.7)有一段有趣的历史. 该问题的原始工作来自若干不同方面，可以从如下的论文看出，包括 Barnum, Nielsen 和 Schumacher 的论文[BNS98]，Bennett, DiVincenzo, Smolin 和 Wootters 的论文[BDSW96]，Lloyd 的论文[Lloyd97]，Schumacher 的论文[Sch96b]以及 Schumacher 和 Nielsen 的论文[SN96]. 这些若干观点的等价性已通过 Barnum, Knill 和 Nielsen[BKN98]以及 Barnum, Smolin 和 Terhal[BST98]的工作得到理解. Bennett, DiVincenzo 和 Smolin[BDS97]已得到了某些特殊信道(最值得关注的擦除信道)的容量. 而令人着迷地使用退化量子编码的去极化信道的容量的一个下界已由 Shor 和 Smolin[SS96]得到，由 DiVincenzo, Shor 和 Smolin[DSS98]改进. Zurek[Zur89]，Milburn[Mil96]，及 Lloyd[Llo96]分析了不在纠错问题中的量子 Maxwell 妖. 这里的分析基于 Nielsen, Cave, Schumacher 和 Barnum 的工作[NCSB98]. 这个观点还被 Vedral[Ved99]用到获得量子蒸馏过程的极限上. 量子单一界来自 Knill 和 Laflamme[KL97]，我们给的证明来自 Preskill[Pre98b].

纠缠的研究已发展成一个主要的研究方向，这方面研究论文非常多，我们无法在这里一一说明，再次请读者参考 http://arXiv.org/的 quant-pharchive. 基于控制不等式的纠缠变换条件(定理 12.15)属于 Nielsen[Nie99a]. 定理 12.13 来自

Uhlmann[Uhl71,Uhl72,Uhl73]. 命题 12.14 来自 Lo 和 Popescu[LP97]. 纠缠催化是 Jonathan 和 Plenio[JP99]发现的. Marshall 和 Olkin[MO79]是控制不等式的一个综合性的引论，包含了 Birkhoff 定理的证明. 纠缠稀释和蒸馏的极限来自 Bennett, Bernstein, Popescu 和 Schumacher[BBPS96]. 混合态的纠缠蒸馏协议由 Bennett, Brassard, Popescu, Schumacher, Smolin 和 Wootters[BBP+96]引入，与纠错的联系是 Bennett, DiVincenzo, Smolin 和 Wootters 的开拓性论文[BDSW96]中发展的，这篇论文引发许多后续的研究. 图 12.11 中的例子是 Gottesman 和 Nielsen(未发表)注意到的. 我们再提及几篇关于纠缠的非常有趣的论文，它们或许可以作为文献的入门. 遗憾的是许多有价值的论文被省略了. Horodecki 家族成员(Michal, Pawel 和 Ryszard)的系列论文深入地研究了纠缠的性质；特别值得注意的是文献[HHH96, HHH98, HHH99a, HHH99b, HHH99c]. Vedral 和 Plenio 的论文[VP98]及 Vidal 的论文[Vid98]也具有重要意义.

量子密码术的(早期)出色的基础综述参见 Bennett, Brassard 和 Ekert 在 Scienfific American 上的文章[BBE92]. 量子密码术的思想是 Wiesner 在 20 世纪 60 年代首先提出的. 遗憾的是，Wiesner 的思想未被接受发表，直到 20 世纪 80 年代早期这些思想才被所知. Wiesner 提出的(纠缠)量子状态，如果能长期存储，可被用作一类防伪货币[Wie,Wie83]. Bennett 开发了若干进一步的协议，其中一个导致 Bennett, Bessette, Brassard, Salvail 和 Smolin[BBB+82]的第一个实验实现. 它具有历史意义(在原理上)，因为它传输信息不超过一米，而且，窃听采用的是每当 1 被发送时电源释放出响亮的嗡嗡声. 保密增强的概念是 Bennett, Brassard 和 Robert[BBR88]首先引入的. 关于信息调和参见文献[BBB+92]和文献[BS94]. 定理 12.16 的叙述和证明在 Bennett, Brassard, Crèpeau 和 Maurer[BBCM95]对保密增强的可读性很高的一般处理中. 注意在调和过程中泄露的信息在保密增强的阈值上有重要影响，如定理 12.17 中的估界，被 Cachin 和 Maurer[CM97]证明. 通过使用遥远的相关随机信源，如卫星感知的星光，保密增强已用到经典的密钥生成[Mau93]. 称为 BB84 的四状态协议根据作者的名字命名，即 Bennett 和 Brassard[BB84]，类似地，两状态 B92 协议是根据 Bennett[Ben92]命名的. EPR 协议由 Ekert[Eke91]设计. 练习 12.27 中的随机采样界来自 Ambainis. 量子密码术的局限和安全性在许多文献中有深入讨论. 例如，参见 Barnnet 和 Phoenix[BP93]；Brassard[Bra93]；Ekert, Huttner, Palma 和 Peres[EHPP94]的工作；还有文献[Phy92]. 相干信息和保密性之间的联系由 Schumacher 和 Westmoreland[SW98]认识到的. 量子密码系统的实验实现方面发表了无数论文. 好的引论可参见 Hughes, Alde, Dyer, Luther, Morgan 和 Schauer[HAD+95]；日内瓦湖底的量子密码术演示是 Muller, Zbinden 和 Gisin[MZG96]

做的. 盒子 12.7 中描述的实验是 Bethune 和 Risk[BR98,BR00]在 IBM 做的，我们要感谢他们提供的设备原理图. 各种量子密钥分配协议的安全性的大量证明，在不同场合已给出. 特别值得注意的是采用 BB84 的 QKD 安全性的完整(详细但有些复杂的)证明由 Mayers[May98]给出. Biham, Boyer, Brassard, van de Graaf 和 Mor 还考虑了对 BB84 的攻击[BBB+98]. 采用 EPR 状态和要求完美量子计算的一个较简单的证明由 Lo 和 Chau[LC99]给出；这是 12.6.5 节开始时的协议. Lo 把它简化为在传输密钥材料之前，从一个确定差错率的测试开始[Lo99]. 12.6.5 节中更简单(而且漂亮)的证明来自 Shor 和 Preskill[SP00]，他们还给出了 12.6.5 节提到的 11%估计. 我们给出的那个证明还非常得益于与 Gottesman 的交谈.

APPENDIX A

附 录 A

Lieb 定理证明

量子信息论中最重要和有用的定理之一是，对 von Neumann 熵的强次可加性不等式. 它断言对一个三元量子系统 A,B,C，有

$$S(A,B,C)+S(B)\leqslant S(A,B)+S(B,C) \tag{A.1}$$

遗憾的是，尚不知道强次可加性的简单证明. 第 11 章给出了一个相对简单的证明，它基于一个称为 Lieb 定理的深刻数学结果. 本附录我们就来证明 Lieb 定理，首先从若干简单的概念和定义开始.

设 $f(A,B)$是两个矩阵 A 和 B 的实值函数，如果对所有的 $0\leqslant\lambda\leqslant1$，

$$f(\lambda A_1+(1-\lambda)A_2,\lambda B_1+(1-\lambda)B_2)\geqslant\lambda f(A_1,B_1)+(1-\lambda)f(A_2,B_2) \tag{A.2}$$

那么 f 称对 A 和 B 为联合凹. 对矩阵 A 和 B，如果 $B-A$ 是半正定矩阵，我们说 $A\leqslant B$. 如果 $B\leqslant A$ 我们说 $A\geqslant B$. 令 A 为任一矩阵，我们定义矩阵 A 的范数为

$$\|A\|\equiv\max_{\langle u|u\rangle=1}|\langle u|A|u\rangle| \tag{A.3}$$

在我们的 Lieb 定理证明中我们偶尔会用到如下容易验证的事实：

练习 A.1(共轭保持≤)　如果 $A\leqslant B$，证明对所有矩阵 $XAX^\dagger\leqslant XBX^\dagger$.

练习 A.2　证明当且仅当 A 为半正定，$A\geqslant0$.

练习 A.3(≤为一偏序)　证明关系“≤”是算子上的偏序——即它是传递的($A\leqslant B$ 和 $B\leqslant C$ 蕴含 $A\leqslant C$)、反对称的($A\leqslant B$ 和 $B\leqslant A$ 蕴含 $A=B$)和自反的($A\leqslant A$).

练习 A.4　设 A 具有特征值 λ_i，定义 λ 为集合 $|\lambda_i|$ 的最大值，证明

(1) $\|A\|\geqslant\lambda$；

(2) 当 A 为 Hermite 的，$\|A\|=\lambda$；

(3) 当

$$A=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix} \tag{A.4}$$

时，$\|A\|=3/2>1=\lambda$.

练习 A.5(AB 和 BA 具有相同的特征值) 证明 AB 和 BA 具有相同的特征值.(提示:对可逆的 A,证明 $\det(xI-AB)=\det(xI-BA)$,因此 AB 和 BA 的特征值相同.由连续性,这对 A 不可逆时也成立)

练习 A.6 假设 A 和 B 使得 AB 为 Hermite 的,用前两个事实证明 $\|AB\| \leqslant \|BA\|$.

练习 A.7 假设 A 为半正定,证明当且仅当 $A \leqslant I$, $\|A\| \leqslant 1$.

练习 A.8 令 A 为半正定矩阵,以方程 $\mathscr{A}(X)\equiv AX$ 定义一个超算子(矩阵上的线性算子),证明 $\mathscr{A}$ 关于 Hilbert-Schmidt 内积为半正定.即对所有 X, $\mathrm{tr}(X^\dagger \mathscr{A}(X)) \geqslant 0$.类似地,证明由 $\mathscr{A}(X)\equiv XA$ 定义的超算子关于矩阵上的 Hilbert-Schmidt 内积为半正定.

有了这些结果,我们现在来叙述并证明 Lieb 定理.

定理 A.1(Lieb 定理) 令 X 为一矩阵,且 $0 \leqslant t \leqslant 1$,则函数

$$f(A,B) \equiv \mathrm{tr}(X^\dagger A^t X B^{1-t}) \tag{A.5}$$

对半正定矩阵 A 和 B 是联合凹的.

Lieb 定理是如下引理的简单推论.

引理 A.2 令 $R_1, R_2, S_1, S_2, T_1, T_2$ 为半正定算子,使得 $0=[R_1,R_2]=[S_1,S_2]=[T_1,T_2]$,且

$$R_1 \geqslant S_1 + T_1 \tag{A.6}$$

$$R_2 \geqslant S_2 + T_2 \tag{A.7}$$

那么对所有的 $0 \leqslant t \leqslant 1$,

$$R_1^t R_2^{1-t} \geqslant S_1^t S_2^{1-t} + T_1^t T_2^{1-t} \tag{A.8}$$

矩阵不等式成立.

证明 首先对 $t=1/2$ 证明,然后以此对一般的 t 建立结果.为方便,假设 R_1 和 R_2 可逆,当不是这种情况时,证明需要在技术上有小的修改,这留作练习.

令 $|x\rangle$ 和 $|y\rangle$ 为两个任意向量.应用 Cauchy-Schwarz 不等式两次并经过简单计算得到

$$\begin{aligned}
&|\langle x|(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})|y\rangle| \\
&\leqslant |\langle x|S_1^{1/2}S_2^{1/2}|y\rangle| + |\langle x|T_1^{1/2}T_2^{1/2}|y\rangle| && \text{(A.9)} \\
&\leqslant \||S_1^{1/2}|x\rangle\|\,\||S_2^{1/2}|y\rangle\| + \||T_1^{1/2}|x\rangle\|\,\||T_2^{1/2}|y\rangle\| && \text{(A.10)} \\
&\leqslant \sqrt{(\||S_1^{1/2}|x\rangle\|^2 + \||T_1^{1/2}|x\rangle\|^2)(\||S_2^{1/2}|y\rangle\|^2 + \||T_2^{1/2}|y\rangle\|^2)} && \text{(A.11)} \\
&= \sqrt{\langle x|(S_1+T_1)|x\rangle\langle y|(S_2+T_2)|y\rangle} && \text{(A.12)}
\end{aligned}$$

由假设,$S_1+T_1 \leqslant R_1$ 和 $S_2+T_2 \leqslant R_2$,故

$$|\langle x|(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})|y\rangle| \leqslant \sqrt{\langle x|R_1|x\rangle\langle y|R_2|y\rangle} \tag{A.13}$$

令$|u\rangle$为任一单位向量，那么应用式(A.13)及$|x\rangle\equiv R_1^{-1/2}|u\rangle$和$|y\rangle\equiv R_2^{-1/2}|u\rangle$，导出

$$\langle u|R_1^{-1/2}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}|u\rangle$$

$$\leqslant\sqrt{\langle u|R_1^{-1/2}R_1R_1^{-1/2}|u\rangle\langle u|R_2^{-1/2}R_2R_2^{-1/2}|u\rangle} \tag{A.14}$$

$$=\sqrt{\langle u|u\rangle\langle u|u\rangle}=1 \tag{A.15}$$

因此

$$\|R_1^{-1/2}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}\|\leqslant 1 \tag{A.16}$$

定义

$$A\equiv R_1^{-1/4}R_2^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2} \tag{A.17}$$

$$B\equiv R_2^{1/4}R_1^{-1/4} \tag{A.18}$$

注意AB是Hermite的，故由练习A.6，

$$\|R_1^{-1/4}R_2^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4}\|$$

$$=\|AB\|\leqslant\|BA\| \tag{A.19}$$

$$=\|R_1^{-1/2}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}\| \tag{A.20}$$

$$\leqslant 1 \tag{A.21}$$

其中最后一个不等式正是式(A.16)．AB是半正定算子，故由练习A.7和前面的不等式，得

$$R_1^{-1/4}R_2^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4}\leqslant I \tag{A.22}$$

最后，由练习A6.1以及R_1和R_2的可对易性，有

$$S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2}\leqslant R_1^{1/2}R_2^{1/2} \tag{A.23}$$

这就证明式(A.8)对$t=1/2$成立．

令I为使式(A.8)成立的所有t的集合．据观察，我们知道0和1都是I的元，且刚刚证明1/2是I的元．现在我们用$t=1/2$的情形来对任意$0\leqslant t\leqslant 1$的t证明结果．假设μ和η为I的两个元，使得

$$R_1^{\mu}R_2^{1-\mu}\geqslant S_1^{\mu}S_2^{1-\mu}+T_1^{\mu}T_2^{1-\mu} \tag{A.24}$$

$$R_1^{\eta}R_2^{1-\eta}\geqslant S_1^{\eta}S_2^{1-\eta}+T_1^{\eta}T_2^{1-\eta} \tag{A.25}$$

这些不等式具有已证明的$t=1/2$情形的形式(A.6)和(A.7)．利用$t=1/2$的结果，看到

$$(R_1^{\mu}R_2^{1-\mu})^{1/2}(R_1^{\eta}R_2^{1-\eta})^{1/2}\geqslant(S_1^{\mu}S_2^{1-\mu})^{1/2}\cdot(S_1^{\eta}S_2^{1-\eta})^{1/2}$$

$$+(T_1^{\mu}T_2^{1-\mu})^{1/2}\cdot(T_1^{\eta}T_2^{1-\eta})^{1/2} \tag{A.26}$$

利用对易性，假定$0=[R_1,R_2]=[S_1,S_2]=[T_1,T_2]$，看到对$\nu\equiv(\mu+\eta)/2$，得到

$$R_1^{\nu}R_2^{1-\nu}\geqslant S_1^{\nu}S_2^{1-\nu}+T_1^{\nu}T_2^{1-\nu} \tag{A.27}$$

因此只要μ和η属于I，$(\mu+\eta)/2$也属于I．因为0和1属于I，易见0和1之间的任何具有有限二进制展开的数t都在I中．因此I在$[0,1]$中稠密．现在命题可以

从结论(A.8)对 t 的连续性得出. □

Lieb 定理的证明是引理 A.2 的简单应用. 使这点成为可能的巧妙想法是,把引理 A.2 中的算子选成*超算子*——算子上的线性映射. 这些算子的选择方式是,使其对 Hilbert-Schmidt 内积 $(A,B)\equiv \mathrm{tr}(A^\dagger B)$ 为半正定.

证明(Lieb 定理) 令 $0\leqslant\lambda\leqslant 1$,并定义超算子 $\mathcal{S}_1,\mathcal{S}_2,\mathcal{T}_1,\mathcal{T}_2,\mathcal{R}_1,\mathcal{R}_2$ 如下:

$$\mathcal{S}_1(X)\equiv \lambda A_1 X \tag{A.28}$$

$$\mathcal{S}_2(X)\equiv \lambda X B_1 \tag{A.29}$$

$$\mathcal{T}_1(X)\equiv (1-\lambda)A_2 X \tag{A.30}$$

$$\mathcal{T}_2(X)\equiv (1-\lambda)X B_2 \tag{A.31}$$

$$\mathcal{R}_1\equiv \mathcal{S}_1+\mathcal{T}_1 \tag{A.32}$$

$$\mathcal{R}_2\equiv \mathcal{S}_2+\mathcal{T}_2 \tag{A.33}$$

注意 $\mathcal{S}_1$ 和 $\mathcal{S}_2$ 对易,并且 $\mathcal{T}_1$ 和 $\mathcal{T}_2$ 对易,$\mathcal{R}_1$ 和 $\mathcal{R}_2$ 对易. 回忆练习 A.8,这些算子关于 Hilbert-Schmidt 内积是半正定的,则由引理 A6.2,有

$$\mathcal{R}_1^t\mathcal{R}_2^{1-t}\geqslant \mathcal{S}_1^t\mathcal{S}_2^{1-t}+\mathcal{T}_1^t\mathcal{T}_2^{1-t} \tag{A.34}$$

应用 Hilbert-Schmidt 内积,取上述不等式的矩阵 $X\cdot X$ 元得到

$$\begin{aligned}&\mathrm{tr}[X^\dagger(\lambda A_1+(1-\lambda)A_2)^t X(\lambda B_1+(1-\lambda)B_2)^{1-t}]\\ &\geqslant \mathrm{tr}[X^\dagger(\lambda A_1)^t X(\lambda B_1)^{1-t}]+\mathrm{tr}[X^\dagger((1-\lambda)A_2)^t X((1-\lambda)B_2)^{1-t}] \qquad (A.35)\\ &=\lambda\mathrm{tr}(X^\dagger A_1^t X B_1^{1-t})+(1-\lambda)\mathrm{tr}(X^\dagger A_2^t X B_2^{1-t}) \qquad (A.36)\end{aligned}$$

即为期望的联合凹性命题. □

历史和进一步阅读的材料

Lieb 定理的历史与量子熵的强次可加性不等式的证明紧密联系在一起,可以和该不等式证明的历史一起在第 11 章的"历史和进一步阅读的材料"找到.

参考文献

[ABO97] D. Aharonov and M. Ben-Or. Fault tolerant computation with constant error. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing*, pages 176—188, 1997.

[ABO99] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM J. Comp.*, page to appear, 1999. *arXive e-print quant-ph/9906129*[1]

[ABOIN96] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan. Limitations of noisy reversible computation, *arXive e-print-quant-ph/9611028*, 1996.

[ADH97] L. Adleman, J. Demarrais, and M. A. Huang. Quantum computability. *SIAM J. Comp.*, 26(5):1524—1540, 1997.

[Adl94] L. M. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266:1021, 1994.

[Adl98] L. M. Adleman. Computing with DNA. *Sci. Am.*, 279:54—61, Aug. 1998.

[AE75] L. Allen and J. H. Eberly. *Optical Resonance and Two-level Atoms*. Dover, New York, 1975.

[Aha99a] D. Aharonov. *Noisy Quantum Computation*. Ph. D. thesis, The Hebrew University, Jerusalem, 1999.

[Aha99b] D. Aharonov. Quantum computation. In D. Stauffer, editor, *Annual Reviews of Computational Physics VI*. World Scientific, Singapore, 1999.

[AKN98] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. *STOC 1997*, 1998. *arXive e-print quant-ph/9806029*.

[AL70] H. Araki and E. H. Lieb. Entropy inequalities. *Comm. Math. Phys.*, 18:160—170, 1970.

[AL97] D. S. Abrams and S. Lloyd. Simulation of many-body Fermi systems on a quantum computer. *Phys. Rev. Lett.*, 79(13): 2586 — 2589, 1997. *arXive e-print quant-ph/9703054*.

[1] 带有"arXive e-print quant-ph/xxxxxxx"的文献可以在互联网址 http://www.arXiv.org 找到.

[AL99] A. Ashikhmin and S. Lytsin. Upper bounds on the size of quantum codes. *IEEE Trans. Inf. Theory*, 45(4):1206－1215, 1999.

[Alb83] P. M. Alberti. A note on the transition-probability over c-* algebras. *Lett. In Math. Phys.*, 7(1):25－32, 1983.

[Amb00] A. Ambainis. Quantum lower bounds by quantum arguments, *arXive e-print quant-ph/O002066*, 2000.

[And79] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.*, 26: 203－241, 1979.

[Ash97] A. Ashikhmin. Remarks on bounds for quantum codes, *arXive e-print quant-ph/9705037*, 1997.

[Bar78] E. Barton. A reversible computer using conservative logic. Unpublished MIT 6. 895 term paper, 1978.

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175－179, IEEE, New York, 1984. Bangalore, India, December 1984.

[BBB+92] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin. Experimental quantum cryptography. *J. Cryptology*, 5:3－28, 1992.

[BBB+98] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor. Security of quantum key distribution against all collective attacks, *arXive e-print quantph/9801022*, 1998.

[BBBV97] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510－1523, 1997. *arXive e-print quant-ph/9701001*.

[BBBW82] C. H. Bennett, G. Brassard, S. Breidbart, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In D. Chaum, R. L. Rivest, and A. T. Sherman, editors, *Advances in Cryptology: Proceedings of Crypto* 82, pages 267 － 275, Plenum Press, New York, 1982.

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and EPR channels. *Phys. Rev. Lett.*, 70:1895－1899, 1993.

[BBC+95] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H.

Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52: 3457 — 3467, 1995. *arXive e-print quant-ph*/9503016.

[BBC+98] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science (FOCS'98)*, pages 352 — 361, IEEE, Los Alamitos, California, November 1998. *arXive e-print quant-ph*/9802049.

[BBCM95] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Trans. Inf. Theory*, 41: 1915 — 1923, 1995.

[BBE92] C. H. Bennett, G. Brassard, and A. K. Ekert. Quantum cryptography. *Sci. Am.*, 267 (4):50, Oct. 1992.

[BBHT98] M. Boyer, G. Brassard, P. Hoyer, and A. Tapp. Tight bounds on quantum searching. *Fortsch. Phys.-Prog. Phys.*, 46(4—5):493—505, 1998.

[BBM+98] D. Boschi, S. Branca, F. D. Martini, L. Hardy, and S. Popescu. Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolski-Rosen channels. *Phys. Rev. Lett.*, 80: 1121 — 1125, 1998. *arXive e-print quant-ph*/9710013.

[BBP+96] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76: 722, 1996. *arXive e-print quant-ph*/9511027.

[BBPS96] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53 (4):2046—2052, 1996. *arXive e-print quant-ph*/9511030.

[BBR88] C. H. Bennett, G. Brassard, and J. M. Robert. Privacy amplification by public discussion. *SIAM J. Comp.*, 17: 210—229, 1988.

[BCDP96] D. Beckman, A. N. Chari, S. Devabhaktuni, and J. Preskill. Efficient networks for quantum factoring. *Phys. Rev. A*, 54(2):1034, 1996. *arXive e-print quant-ph*/9602016.

[BCF+96] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.*, 76 (15): 2828 — 2821, 1996. *arXive e-print quant-ph*/9511010.

[BCJ+99] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing. *Phys. Rev. Lett.*, 83(5): 1054—1057, 1999.

[BCJD99] G. K. Brennen, C. M. Caves, P. S. Jessen, and I. H. Deutsch. Quantum logic gates in optical lattices. *Physical Review Letters*, 82: 1060—1063, 1999.

[BD00] C. H. Bennett and D. P. DiVincenzo. Quantum information and computation. *Nature*, 404:247—255, 2000.

[BDG88a] J. L. Balcázar, J. Diaz, and J. Gabarró. *Structural Complexity*, Volume I. Springer-Verlag, Berlin, 1988.

[BDG88b] J. L. Balcázar, J. Diaz, and J. Gabarró. *Structural Complexity*, Volume II Springer-Verlag, Berlin, 1988.

[BDK92] R. G. Brewer, R. G. DeVoe, and R. Kallenbach. Planar ion microtraps. *Phys. Rev. A*, 46 (11):R6781—6784, 1992.

[BDS97] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin. Capacities of quantum erasure channels. *Phys. Rev. Lett.*, 78(16): 3217—3220, 1997. *arXive e-print quant-ph/*9701015.

[BDSW96] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed state entanglement and quantum error correction. *Phys. Rev. A*, 54: 3824, 1996. *arXive e-print quant-ph*/9604024.

[Bel64] J. S. Bell. On the Einstein-Podolsy-Rosen paradox. *Physics*, 1: 195—200, 1964. Reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, 1987.

[Ben73] C. H. Bennett. Logical reversibility of computation. *IBM J. Res. Dev.*, 17(6):525—532, 1973.

[Ben80] P. Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *J. Stat. Phys.*, 22 (5):563—591, 1980.

[Ben82a] C. H. Bennett. The thermodynamics of computation—A review. *Int. J. Theor. Phys.* 21, 905—940, 1982.

[Ben87] C. H. Bennett. Demons, engines and the second law. *Sci. Am.*, 295(5):108, 1987.

[Ben89] C. H. Bennett. Time-space tradeoffs for reversible computation. *SIAM J. Comput.*, 18: 766—776, 1989.

[Ben92] C. H. Bennett. Quantum cryptography using any two non-

orthogonal states. *Phys. Ret. Lett.*, 68(21): 3121 — 3124, 1992.

[Bet84] T. Beth. *Methoden der Schnelle Fouriertransformation*. Teubner Leipzig, 1984.

[BFGL98] S. L. Braunstein, C. A. Fuchs, D. Gottesman, and H. Lo. A quantum analog of Huffman coding, *arXive e-print quant-ph*/9805080, 1998.

[BFJS96] H. Barnum, C. A. Fuchs, R. Jozsa, and B. Schumacher. General fidelity limit for quantum channels. *Phys. Rev. A*, 54:4707, 1996. *arXive e-print quant-ph*/9603014.

[Bha97] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.

[BHT98] G. Brassard, P. Hoyer, and A. Tapp. Quantum counting, *arXive e-print quant-ph*/9805082, 1998.

[BK92] V. B. Braginsky and F. Y. Khahili. *Quantum Measurement*. Cambridge University Press, Cambridge, 1992.

[BK98a] S. L. Braunstein and H. J. Kimble. Teleportation of continuous quantum variables. *Phys. Rev. Lett.*, 80: 869 — 872, 1998.

[BK98b] S. B. Bravyi and A. Y. Kitaev. Quantum codes on a lattice with boundary, *arXive e-print quant-ph*/9811052, 1998.

[BK99] S. L. Braunstein and H. J. Kimble. Dense coding for continuous variables, *arXive e-print quant-ph*/9910010, 1999.

[BKLW99] D. Bacon, J. Kempe, D. A. Lidar, and K. B. Whaley. Universal fault tolerant computation on decoherence-free subspaces, *arXive e-print quant-ph*/9909058, 1999.

[BKN98] H. Barnum, E. Knill, and M. A. Nielsen. On quantum fidelities and channel capacities. *arXive e-print quant-ph*/9809010, 1998.

[BL95] D. Boneh and R. J. Lipton. Quantum cryptoanalysis of hidden linear functions (extended abstract). In Don Coppersmith, editor, *Lecture notes in computer science -- Advances in Cryptology -- CRYPTO'95*, pages 424 — 437, Springer-Verlag, Berlin, 1995.

[BMP+99] P. O. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan. On universal and fault-tolerant quantum computing. *arXive e-print quant-ph*/9906054, 1999.

[BNS98] H. Barnum, M. A. Nielsen, and B. W. Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57:4153, 1998.

[Boh51] D. Bohm. *Quantum Theory*. Prentice-Hall, Englewood Cliffs, New Jersey, 1951.

[BP93] S. M. Barnett and S. J. D. Phoenix. Information-theoretic

limits to quantum cryptography. *Phys. Rev. A*, 48(1): R5—R8, 1993.

[BPM+97] D. Bouwmeester, J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. Experimental quantum teleportation. *Nature*, 390 (6660):575—579, 1997.

[BR98] D. S. Bethune and W. P. Risk. An autocompensating quantum key distribution system using polarization splitting of light. In *IQEC '98 Digest of Post deadline Papers*, pages QPD12—2, Optical Society of America, Washington, DC, 1998.

[BR00] D. S. Bethune and W. P. Risk. An autocompensating fiber-optic quantum cryptography system based on polarization splitting of light. *J. Quantum Electronics*, 36(3):100, 2000.

[Bra93] G. Brassard. A bibliography of quantum cryptography. *Université de Montréal preprint*, pages 1—10, 3 December 1993. A preliminary version of this appeared in *Sigact News*, vol. 24 (3), 1993, pages 16—20.

[Bra98] S. L. Braunstein. Error correction for continuous quantum variables. *Phys. Rev. Lett.*, 80:4084—4087, 1998. *arXive e-print quant-ph*/9711049.

[BS94] G. Brassard and L. Salvail. Secret-key reconciliation by public discussion. In T. Helleseth, editor, *Lecture Notes in Computer Science: Advances in Cryptology - EUROCRYPT'93*, Volume 765, pages 410423, Springer-Verlag, New York, 1994.

[BS98] C. H. Bennett and P. W. Shor. Quantum information theory, *itit*, 44 (6): 2724 — 2742, 1998.

[BST98] H. Barnum, J. A. Smolin, and B. Terhal. Quantum capacity is properly defined without encodings. *Phys. Rev. A*, 58 (5):3496—3501, 1998.

[BT97] B. M. Boghosian and W. Taylor. Simulating quantum mechanics on a quantum computer, *arXive e-print quant-ph*/9701019, 1997.

[BV97] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM J. Comput.*, 26(5): 1411—1473, 1997. *arXive e-print quant-ph*/9701001.

[BW92] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69(20):2881—2884, 1992.

[CAK98] N. J. Cerf, C. Adami, and P. Kwiat. Optical simulation of quantum logic. *Phys. Rev. A*, 57:R1477, 1998.

[Cay99] C. M. Caves. Quantum error correction and reversible operations. *Journal of Superconductivity*, 12(6): 707 — 718, 1999.

[CD96] R. Cleve and D. P. DiVincenzo. Schumacher's quantum data compression as a quantum computation. *Phys. Rev. A*, 54: 2636, 1996. *arXive e-print quant-ph*/9603009.

[CEMM98] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proc. R. Soc. London A*, 454 (1969):339—354, 1998.

[CFH97] D. G. Cory, A. F. Fahmy, and T. F. Havel. Ensemble quantum computing by NMR spectroscopy. *Proc. Nat. Acad. Sci. USA*, 94: 1634 — 1639, 1997.

[CGK98] I. L. Chuang, N. Gershenfeld, and M. Kubinec. Experimental implementation of fast quantum searching. *Phys. Rev. Lett.*, 18(15): 3408 — 3411, 1998.

[CGKL98] I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung. Bulk quantum computation with nuclear-magnetic-resonance: theory and experiment. *Proc. R. Soc. London A*, 454 (1969): 447 — 467, 1998.

[Che68] P. R. Chernoff. Note on product formulas for operator semigroups. *J. Functional Analysis*, 2:238—242, 1968.

[Cho75] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and Its Applications*, 10: 285 — 290, 1975.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 49:1804—1807, 1969.

[Chu36] A. Church. An unsolvable problem of elementary number theory. *Am. J. Math.* (*reprinted in* [*Dav*65]), 58:345, 1936.

[CK81] I. Csiszár and J. Körner. Information Theory: Coding Theorems for Discrete Memoryless Systems. *Academic Press*, New York, 1981.

[CL83] A. O. Caldeira and A. J. Leggett. Quantum tunnelling in a dissipative system. *Ann. Phys.*, 149 (2):374—456, 1983.

[Cla89] M. Clausen. Fast generalized Fourier transforms. *Theor. Comput. Sci.*, 67:55—63, 1989.

[Cle99] R. Cleve. The query complexity of order-finding, *arXive e-print quant-ph*/9911124, 1999.

[CLR90] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, Cambridge, Mass., 1990.

[CM97] C. Cachin and U. M. Maurer. Linking information reconciliation and privacy amplification. *J. Cryptology*, 10: 97 — 110, 1997.

[CM00] I. L. Chuang and D. Modha. Reversible arithmetic coding for quantum data compression. *IEEE Trans. Inf. Theory*, 46(3): 1104, May 2000.

[CMP+98] D. G. Cory, W. Mass, M. Price, E. Knill, R. Lafiamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo. Experimental quantum error correction, *arXive e-print quant-ph*/9802018, 1998.

[CN97] I. L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.*, 44(11—12):2455—2467, 1997. *arXive e-print quant-ph*/9610001.

[Con72] J. H. Conway. Unpredictable iterations. In *Proceedings of the Number Theory Conference*, pages 49—52, Boulder, Colorado, 1972.

[Con86] J. H. Conway. Fractran: a simple universal programming language. In T. M. Cover and B. Gopinath, editors, *Open Problems in Communication and Computation*, pages 4 — 26, Springer-Verlag, New York, 1986.

[Coo71] S. A. Cook. The complexity of theorem-proving procedures. In *Proc. 3rd Ann. ACM Symp. on Theory of Computing*, pages 151—158, Association for Computing Machinery, New York, 1971.

[Cop94] D. Coppersmith. An approximate Fourier transform useful in quantum factoring. *IBM Research Report RC* 19642, 1994.

[CPZ96] J. I. Cirac, T. Pellizzari, and P. Zoller. Enforcing coherent evolution in dissipative quantum dynamics. *Science*, 273: 1207,1996.

[CRSS97] A. R. Calderbank, E. M. Rains P. W. Shor, and N. J. A. Sloane. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405—408,1997.

[CRSS98] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction via codes over GF (4). *IEEE Trans. Inf. Theory*, 44(4): 1369—1387, 1998.

[CS96] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098, 1996. *arXive e-print quant-ph*/9512032.

[CST89] R. A. Campos, B. E. A. Saleh, and M. C. Tiech. Quantum-mechanical lossless beamsplitters: SU(2) symmetry and photon statistics. *Phys. Rev. A*, 40:1371, 1989.

[CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, New York, 1991.

[CTDL77a] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics, Vol. I.* John Wiley and Sons, New York, 1977.

[CTDL77b] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics, Vol. II.* John Wiley and Sons, New York, 1977.

[CVZ+98] I. L. Chuang, L. M. K. Vandersypen, X. L. Zhou, D. W. Leung, and S. Lloyd. Experimental realization of a quantum algorithm. *Nature*, 393(6681): 143 — 146, 1998.

[CW95] H. F. Chau and F. Wilczek. Simple realization of the Fredkin gate using a series of two-body operators. *Phys. Rev. Lett.*, 75(4): 748 — 750, 1995. *arXive e-print quant-ph*/9503005.

[CY95] I. L. Chuang and Y. Yamamoto. Simple quantum computer. *Phys. Rev. A*, 52: 3489 — 3496, 1995. *arXive e-print quant-ph*/9505011.

[CZ95] J. I. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74:4091, 1995.

[Dav65] M. D. Davis. *The Undecidable.* Raven Press, Hewlett, New York, 1965.

[Day76] E. B. Davies. *Quantum Theory of Open Systems.* Academic Press, London, 1976.

[DBE95] D. Deutsch, A. Barenco, and A. Ekert. Universality in quantum computation. *Proc. R. Soc. London A*, 449(1937): 669 — 677, 1995.

[Deu83] D. Deutsch. Uncertainty in quantum measurements. *Phys. Rev. Lett.*, 50(9): 631 — 633, 1983.

[Deu85] D. Deutsch. Quantum theory, the Church-Turing Principle and the universal quantum computer. *Proc. R. Soc. Lond. A*, 400:97, 1985.

[Deu89] D. Deutsch. Quantum computational networks. *Proc. R. Soc. London A*, 425:73, 1989.

[DG98] L.-M. Duan and G.-C. Guo. Probabilistic cloning and identification of linearly independent quantum states. *Phys. Rev. Lett.*, 80: 4999 — 5002, 1998. *arXive e-print quant-ph*/9804064.

[DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. Inf. Theory*, IT-22(6): 644 — 654, 1976.

[DH96] C. Dtirr and P. Hoyer. A quantum algorithm for finding the minimum, *arXive e-print quant-ph*/9607014, 1996.

[Die82] D. Dieks. Communication by EPR devices. *Phys. Lett. A*, 92(6):271—272, 1982.

[DiV95a] D. P. DiVincenzo. Quantum computation. *Science*, 270:255, 1995. *arXive e-print quant-ph*/9503016

[DiV95b] D. P. DiVincenzo. Two-bit gates

are universal for quantum computation. *Phys. Rev. A*, 51(2): 1015—1022, 1995.

[DiV98] D. P. DiVincenzo. Quantum gates and circuits. *Proc. R. Soc. London A*, 454:261—276, 1998.

[DJ92] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proc. R. Soc. London A*, 439:553, 1992.

[DL98] W. Diffie and S. Landau. *Privacy on the Line: the Politics of Wiretapping and Encryption*. MIT Press, Cambridge Massachusetts, 1998.

[DMB+93] L. Davidovich, A. Maali, M. Brune, J. M. Raimond, and S. Haroche. *Phys. Rev. Lett.*, 71:2360, 1993.

[DR90] P. Diaconis and D. Rockmore. Efficient computation of the Fourier transform on finite groups. *J. Amet. Math. Soc.*, 3(2):297—332, 1990.

[DRBH87] L. Davidovich, J. M. Raimond, M. Brune, and S. Haroche. *Phys. Rev. A*, 36: 3771, 1987.

[DRBH95] P. Domokos, J. M. Raimond, M. Brune, and S. Haroche. Simple cavity-QED, two-bit universal quantum logic gate: The principle and expected performances. *Phys. Rev. Lett.*, 52: 3554, 1995.

[DS96] D. P. DiVincenzo and P. W. Shor. Fault-tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.*, 77: 3260, 1996.

[DSS98] D. P. DiVincenzo, P. W. Shor, and J. Smolin. Quantum-channel capacities of very noisy channels. *Phys. Rev. A*, 57(2):830—839, 1998.

[Ear42] S. Earnshaw. On the nature of the molecular forces which regulate the constitution of the luminiferous ether. *Trans. Camb. Phil. Soc.*, 7:97—112, 1842.

[EBW87] R. R. Ernst, G. Bodenhausen, and A. Wokaun. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford University Press, Oxford, 1987.

[EH99] M. Ettinger and P. Hoyer. On quantum algorithms for noncommutative hidden subgroups. In *Symposium on The oretical Aspects in Computer Science*. University of Trier, 1999. *arXive e-print quant-ph*/9807029.

[EHK99] M. Ettinger, P. Hoyer, and E. Knill. Hidden subgroup states are almost orthogonal. *arXive e-print quant* \ | *ph*/9901034, 1999.

[EHPP94] A. K. Ekert, B. Hutmer, G. M. Palma, and A. Peres. Eavesdropping on quantum-cryptographical systems. *Phys. Rev. A*, 50(2): 1047—1056, 1994.

[EJ96] A. Ekert and R. Jozsa. Quantum computation and Shor's

factoring algorithm. *Rev. Mod. Phys.*, 68:1, 1996.

[EJ98] A. Ekert and R. Jozsa. Quantum algorithms: Entanglement enhanced information processing. *Proc. R. Soc. London A*, 356(1743): 1769 — 1782, Aug. 1998. *arXive e-print quant-ph*/9803072.

[Eke91] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661—663, 1991.

[EM96] A. Ekert and C. Macchiavllo. Error correction in quantum communication. *Phys. Rev. Lett.*, 77:2585, 1996. *arXive e-print quant-ph*/9602022.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47: 777 — 780, 1935.

[Eps73] H. Epstein. *Commun. Math. Phys.*, 31:317—325, 1973.

[Fan73] M. Fannes. A continuity property, of the entropy density for spin lattice systems. *Commun. Math. Phys.*, 31:291—294, 1973.

[FC94] C. A. Fuchs and C. M. Caves. Ensemble-dependent bounds for accessible information in quantum mechanics. *Phys. Rev. Lett.*, 73(23): 3047 — 3050, 1994.

[Fel68a] W. Feller. *An Introduction to Probability Theory and its Applications*, Volume 1. Wiley, New York, 1968.

[Fel68b] W. Feller. *An Introduction to Probability Theory and its Applications*, Volume 2. Wiley, New York, 1968.

[Fey82] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, 21:467,1982.

[FG98] E. Farhi and S. Gutmann. An analog analogue of a digital quantum computation. *Phys. Rev. A*, 57(4):2403 — 2406, 1998. *arXive e-print quant-ph*/9612026.

[FLS65a] R. P. Feynman, R. B. Leighton, and M. Sands. Volume Ⅲ of *The Feynman Lectures on Physics*. Addison-Wesley, Reading, Mass., 1965.

[FLS65b] R. P. Feynman, R. B. Leighton, and M. Sands. Volume Ⅰ of *The Feynman Lectures on Physics*. Addison-Wesley, Reading, Mass., 1965.

[FM98] M. H. Freedman and D. A. Meyer. Projective plane and planar quantum codes. *arXive e-print quant-ph*/9810055,1998.

[FS92] A. Fässler and E. Stiefel. *Group Theoretical Methods and Their Applications*. Birkhaüser, Boston, 1992.

[FSB+98] A. Furusawa, J. L. Sorensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik. Unconditional quantum telepor-

tation. *Science*, 282:706－709, 1998.

[FT82] E. Fredkin and T. Toffoli. Conservative logic. *Int. J. Theor. Phys.*, 21 (3/4): 219－253, 1982.

[Fuc96] C. A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. Ph. D. thesis, The University of New Mexico, Albuquerque, NM, 1996. *arXive e-print quant-ph/9601020*.

[Fuc97] C. A. Fuchs. Nonorthogonal quantum states maximize classical information capacity. *Phys. Rev. Lett.*, 79 (6): 1162－1165, 1997.

[Gar91] C. W. Gardiner. *Quantum Noise*. Springer-Verlag, Berlin, 1991.

[GC97] N. Gershenfeld and I. L. Chuang. Bulk spin resonance quantum computation. *Science*, 275: 350, 1997.

[GC99] D. Gottesman and I. L. Chuang. Quantum teleportation is a universal computational primitive. *Nature*, 402: 390－392, 1999. *arXive e-print quant-ph/9908010*.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractibility*. W. H. Freeman and Company, New York, 1979.

[GN96] R. B. Griffiths and C.-S. Niu. Semiclassical Fourier transform for quantum computation. *Phys. Rev. Lett.*, 76 (17): 3228－3231, 1996. *arXive e-print quant-ph/9511007*.

[Gor64] J. P. Gordon. Noise at optical frequencies; information theory. In P. A. Miles, editor, *Quantum Electronics and Coherent Light*, Proceedings of the International School of Physics 'Enrico Fermi' XXXI, Academic Press, New York, 1964.

[Got96] D. Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev. A*, 54: 1862, 1996.

[Got97] D. Gottesman. Stabilizer Codes and *Quantum Error Correction*. Ph. D. thesis, California Institute of Technology, Pasadena, CA, 1997.

[Got98a] D. Gottesman. Fault-tolerant quantum computation with higher-dimensional systems, *arXive e-print quant-ph/9802007*, 1998.

[Got98b] D. Gottesman. Theory of fault-tolerant quantum computation. *Phys. Rev. A*, 57(1):127－137, 1998. *arXive e-print quant-ph/9702029*.

[GP10] D. Gottesman and J. Preskill. The Hitchiker's guide to the threshold theorem. *Eternally in preparation*, 1: 1－9120, 2010.

[Gro96] L. Grover. In *Proc.* 28th *Annual ACM Symposium on the Theory of Computation*, pa-

ges 212 — 219, ACM Press, New York, 1996.

[Gro97] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79 (2): 325, 1997. *arXive e-print quant-ph*/9706033.

[Gru99] J. Gruska. *Quantum Computing*. McGraw-Hill, London, 1999.

[GS92] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes*. Clarendon Press, Oxford, 1992.

[HAD+95] R. J. Hughes, D. M. Aide, P. Dyer, G. G. Luther, G. L. Morgan, and M. Schauer. Quantum cryptography. *Contemp. Phys.*, 36(3):149—163, 1995. *arXive e-print quant-ph*/9504002.

[Hal58] P. R. Halmos. *Finite-dimensional Vector Spaces*. Van Nostrand, Princeton, N. J., 1958.

[Ham89] M. Hammermesh. *Group Theory and its Application to Physical Problems*. Dover, New York, 1989.

[HGP96] J. L. Hennessey, D. Goldberg, and D. A. Patterson. Computer Architecture: *A Quantitative Approach*. Academic Press, New York, 1996.

[HHH96] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223(1 — 2): 1 — 8, 1996.

[HHH98] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-state entanglement and distillation: is there a 'bound' entanglement in nature? *Phys. Rev. Lett.*, 80(24): 5239 — 5242, 1998.

[HHH99a] M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, 60(3):1888—1898, 1999.

[HHH99b] M. Horodecki, P. Horodecki, and R. Horodecki. Limits for. Entanglement measures, *arXive e-print quant-ph*/9908065, 1999.

[HHH99c] P. Horodecki, M. Horodecki, and R. Horodecki. Bound entanglement can be activated. *Phys. Rev. Lett.*, 82 (5): 1056 — 1059, 1999.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[HJ91] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[HJS+96] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters. Classical information capacity of a quantum channel. *Phys. Rev. A*, 54:1869, 1996.

[HJW93] L. P. Hughston, R. Jozsa, and

W. K. Wootters. A complete classification of quantum ensembles having a given density matrix. *Phys. Lett. A*, 183: 14—18, 1993.

[HK69] K.-E. Hellwig and K. Kraus. Pure operations and measurements. *Commun. Math. Phys.*, 11:214—220, 1969.

[HK70] K.-E. Hellwig and K. Kraus. Operations and measurements. Ⅱ. *Commun. Math. Phys.*, 16:142—147, 1970.

[Hof79] D. R. Hofstadter. *Gödel, Escher, Bach: an Eternal Golden Braid*. Basic Books, New York, 1979.

[Hol73] A. S. Holevo. Statistical problems in quantum physics. In Gisiro Maruyama and Jurii V. Prokhorov, editors, *Proceedings of the Second Japan-USSR Symposium on Probability Theory*, pages 104—119, Springer Verlag, Berlin, 1973. Lecture Notes in Mathematics, vol. 330.

[Hol79] A. S. Holevo. Capacity of a quantum communications channel. *Problems of Inf. Transm.*, 5(4):247—253, 1979.

[Hol98] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44(1):269—273, 1998.

[Hor97] M. Horodecki. Limits for compression of quantum information carried by ensembles of mixed states. *Phys. Rev. A*, 57:3364—3369, 1997.

[HSM+98] A. G. Huibers, M. Switkes, C. M. Marcus, K. Campman, and A. C. Gossard. Dephasing in open quantum dots. *Phys. Rev. Lett.*, 82:200, 1998.

[HW60] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers, Fourth Edition*. Oxford University Press, London, 1960.

[IAB+99] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small. Quantum information processing using quantum dot spins and cavity qed. *Phys. Rev. Lett.*, 83(20): 4204—4207, 1999.

[IY94] A. Imamoglu and Y. Yamamoto. Turnstile device for heralded single photons: Coulomb blockade of electron and hole tunneling in quantum confined p-i-n heterojunctions. *Phys. Rev. Lett.*, 72(2):210—213, 1994.

[Jam98] D. James. The theory of heating of the quantum ground state of trappedions. *arXive e-print quant-ph/9804048*, 1998.

[Jay57] E. T. Jaynes. Information theory and statistical mechanics, ii. *Phys. Rev.*, 108(2): 171—190, 1957.

[JM98] J. A. Jones and M. Mosca. Im-

plementation of a quantum algorithm to solve Deutsch's problem on a nuclear magnetic resonance quantum computer, *arXive e-print quant-ph*/9801027, 1998.

[JMH98] J. A. Jones, M. Mosca, and R. H. Hansen. Implementation of a quantum search algorithm on a nuclear magnetic resonance quantum computer. *Nature*, 393(6683):344, 1998. *arXive e-print quant-ph*/9805069.

[Jon94] K. R. W. Jones. Fundamental limits upon the measurement of state vectors. *Phys. Rev. A*, 50:3682—3699, 1994.

[Joz94] R. Jozsa. Fidelity for mixed quantum states. *J. Mod. Opt.*, 41:2315—2323, 1994.

[Joz97] R. Jozsa. Quantum algorithms and the Fourier transform, *arXive e-print quant-ph*/9707033, 1997.

[JP99] D. Jonathan and M. B. Plenio. Entanglement-assisted local manipulation of pure states. *Phys. Rev. Lett.*, 83:3566—3569, 1999.

[JS94] R. Jozsa and B. Schumacher. A new proof of the quantum noiseless coding theorem. *J. Mod. Opt.*, 41:2343—2349, 1994.

[Kah96] D. Kahn. *Codebreakers: the Story of Secret Writing*. Scribner, New York, 1996.

[Kan98] B. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393: 133 — 137, 1998.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85 — 103, Plenum Press, New York, 1972.

[KCL98] E. Knitl, I. Chuang, and R. Laflamme. Effective pure states for bulk quantum computation. *Phys. Rev. A*, 57(5):3348—3363, 1998. *arXive e-print quant-ph*/9706053.

[Kit95] A. Y. Kitaev. Quantum measurements and the Abelian stabilizer problem, *arXive e-print quant-ph*/9511026, 1995.

[Kit97a] A. Y. Kitaev. Fault-tolerant quantum computation by anyons, *arXive e-print quant-ph*/9707021, 1997.

[Kit97b] A. Y. Kitaev. Quantum computations: algorithms and error correction. *Russ. Math. Surv.*, 52(6): 1191 — 1249, 1997.

[Kit97c] A. Y. Kitaev. Quantum error correction with imperfect gates. In A. S. Holevo O. Hirota and C. M. Caves, editors, *Quantum Communication, Computing, and Measurement*, pages 181 — 188, Plenum Press, New York, 1997.

[KL51] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22: 79—86, 1951.

[KL97] E. Knill and R. Laflamme. A

theory of quantum error-correcting codes. *Phys. Rev. A*, 55:900, 1997. *arXive e-print quant-ph*/9604034.

[KL99] E. Knill and R. Laflamme. Quantum computation and quadratically signed weight enumerators, *arXive e-print quant-ph*/9909094, 1999.

[Kle31] O. Klein. *Z. Phys.*, 72: 767—775, 1931.

[KLV99] E. Knill, R. Laflamme, and L. Viola. Theory of quantum error correction for general noise, *arXive e-print quant-ph*/9908066, 1999.

[KLZ98a] E. Knill, R. Laflamme, and W. H. Zurek. Resilient quantum computation. *Science*, 279 (5349):342—345, 1998. *arXive e-print quant-ph*/9702058.

[KLZ98b] E. Knill, R. Laflamme, and W. H. Zurek. Resilient quantum computation: error models and thresholds. *Proc. R. Soc.* London A, 454(1969): 365 — 384, 1998. *arXive e-print quant-ph*/9702058.

[KMSW99] P. G. Kwiat, J. R. Mitchell, P. D. D. Schwindt, and A. G. White. Grover's search algorithm: An optical approach, *arXive e-print quant-ph*/9905086, 1999.

[Kni95] E. Knill. Approximating quantum circuits, *arXive e-print quant-ph*/9508006, 1995.

[Knu97] D. E. Knuth. *Fundamental Algorithms 3rd Edition*, Volume 1 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, 1997.

[Knu98a] D. E. Knuth. *Seminumerical Algorithms 3rd Edition*, Volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, 1998.

[Knu98b] D. E. Knuth. *Sorting and Searching 2nd Edition*, Volume 3 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, 1998.

[Kob94] N. Koblitz. *A Course in Number Theory and Cryptography*. Springer-Verlag, New York, 1994.

[KR99] C. King and M. B. Ruskai. Minimal entropy of states emerging from noisy quantum channels, *arXive e-print quant-ph*/9911079, 1999.

[Kra83] K. Kraus. *States, Effects, and Operations: Fundamental Notions of Quantum Theory. Lecture Notes in Physics*, Vol. 190. Springer-Verlag, Berlin, 1983.

[Kra87] K. Kraus. Complementary observables and uncertainty relations. *Phys. Rev. D*, 35(10): 3070—3075, 1987.

[KSC+94] P. G. Kwiat, A. M. Steinberg, R. Y. Chiao, P. H. Eberhard, and M. D. Petroff. Absolute efficiency and time-response

measurement of single-photon detectors. *Appl. Opt.*, 33 (10):1844—1853, 1994.

[KU91] M. Kitagawa and M. Ueda. Nonlinear-interferometric generation of number-phase correlated Fermion states. *Phys. Rev. Lett*, 67 (14): 1852, 1991.

[Lan27] L. Landau. Das dämpfungsproblem in der wellenmechanik. *Z. Phys.*, 45:430—441, 1927.

[Lan61] R. Landauer. Irreversibility and heat generation of the computing process. *IBM J. Res. Dev.*, 5: 183, 1961.

[LB99] S. Lloyd and S. Braunstein. Quantum computation over continuous variables. *Phys. Rev. Lett.*, 82: 1784 — 1787, 1999. *arXive e-print quant-ph/* 9810082.

[LBW99] D. A. Lidar, D. A. Bacon, and K. B. Whaley. Concatenating decoherence free subspaces with quantum error correcting codes. *Phys. Rev. Lett.*, 82 (22):4556—4559, 1999.

[LC99] H. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283:2050—2056, 1999. *arXive e-print quant-ph*/9803006.

[LCW98] D. A. Lidar, I. L. Chuang, and K. B. Whaley. Decoherence-free subspaces for quantum computation. *Phys. Rev. Lett.*, 81 (12):2594—2597, 1998.

[LD98] D. Loss and D. P. DiVincenzo. Quantum computation with quantum dots. *Phys. Rev. A*, 57:120—126, 1998.

[Lec63] Y. Lecerf. Machines de Turing réversibles. *Comptes Rendus*, 257:2597—2600, 1963.

[Leo97] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, New York, 1997.

[Lev73] L. Levin. Universal sorting problems. *Probl. Peredaci Inf.*, 9: 115—116, 1973. Original in Russian. English translation in *Probl. Inf. Transm. USSR* 9:265—266,(1973).

[Lie73] E. H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Ad. Math.*, 11:267—288, 1973.

[Lie75] E. H. Lieb. *Bull. AMS*, 81: 1—13, 1975.

[Lin75] G. Lindblad. Completely positive maps and entropy inequalities. *Commun. Math. Phys.*, 40:147—151, 1975.

[Lin76] G. Lindblad. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.*, 48:199, 1976.

[Lin91] G. Lindblad. Quantum entropy and quantum measurements. In C. Bendjaballah, O. Hirota, and S. Reynaud, editors, *Quantum Aspects of Optical*

Communications, Lecture Notes in Physics, vol. 378, pages 71—80, Springer-Verlag, Berlin, 1991.

[Lip95] R. Lipton. DNA solution of hard computational problems. *Science*, 268: 542 — 525, 1995.

[LKF99] N. Linden, E. Kupce, and R. Freeman. NMR quantum logic gates for homonu-clear spin systems, *arXive e-print quant-ph*/9907003, 1999.

[LL93] A. K. Lenstra and H. W. Lenstra Jr., editors. *The Development of the Number Field Sieve*. Springer-Verlag, New York, 1993.

[Llo93] S. Lloyd. A potentially realizable quantum computer. *Science*, 261:1569, 1993.

[Llo94] S. Lloyd. Necessary and sufficient conditions for quantum computation. *J. Mod. Opt.*, 41(12):2503, 1994.

[Llo95] S. Lloyd. Almost any quantum logic gate is universal. *Phys. Rev. Lett.*, 75(2):346, 1995.

[Llo96] S. Lloyd. Universal quantum simulators. *Science*, 273: 1073, 1996.

[Llo97] S. Lloyd. The capacity of the noisy quantum channel. *Phys. Rev.*, 4, 56:1613,1997.

[LLS75] R. E. Ladner, N. A. Lynch, and A. L. Selman. A comparison of polynomial-time reducibilities. *Theor. Comp. Sci.*, 1:103—124, 1975.

[LMPZ96] R. Lafiamme, C. Miquel, J.-P. Paz, and W. H. Zurek. Perfect quantum error correction code. *Phys. Rev. Lett.*, 77: 198, 1996. *arXive e-print quant-ph*/9602019.

[LNCY97] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto. Approximate quantum error correction can lead to better codes. *Phys. Rev. A*, 56: 2567—2573, 1997. *arXive e-print quant-ph*/9704002.

[Lo99] H. Lo. A simple proof of the unconditional security of quan tum key distribution. *arXive e-print quant-ph*/9904091,1999.

[Lom87] J. S. Lomont. *Applications of Finite Groups*. Dover, New York, 1987.

[Lou73] W. H. Louisell. *Quantum Statistical Properties of Radiation*. Wiley, New York, 1973.

[LP97] H.-K. Lo and S. Popescu. Concentrating local entanglement by local actions-beyond mean values, *arXive e-print quant-ph*/9707038, 1997.

[LP99] N. Linden and S. Popescu. Good dynamics versus bad kinematics. Is entanglement needed for quantum computation? *arXive e-print quant-ph*/9906008, 1999.

[LR68] O. E. Lanford and D. Robinson. Mean entropy of states in quantum-statistical mechan-

ics. *J. Math. Phys.*, 9(7): 1120—1125, 1968.

[LR73a] E. H. Lieb and M. B. Ruskai. A fundamental property of quantum-mechanical entropy. *Phys. Rev. Lett.*, 30(10): 434—436, 1973.

[LR73b] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum mechanical entropy. *J. Math. Phys.*, 14:1938—1941, 1973.

[LR90] H. Left and R. Rex. *Maxwell's Demon: Entropy, Information, Computing*. Princeton University Press, Princeton, NJ, 1990.

[LS93] L. J. Landau and R. F. Streater. On Birkhoff theorem for doubly stochastic completely positive maps of matrix algebras. *Linear Algebra Appl.*, 193:107—127, 1993.

[LS98] S. Lloyd and J. E. Slotine. Analog quantum error correction. *Phys. Rev. Lett.*, 80: 4088—4091, 1998.

[LSP98] H.-K. Lo, T. Spiller, and S. Popescu. Quantum information and computation. *World Scientific*, Singapore, 1998.

[LTV98] M. Li, J. Tromp, and P. Vitanyi. Reversible simulation of irreversible computation by pebble games. *Physica D*, 120:168—176, 1998.

[LV96] M. Li and P. Vitanyi. Reversibility and adiabatic computation: trading time and space for energy. *Proc. R. Soc. London A*, 452:76.

[LVZ+99] D. W. Leung, L. M. K. Vandersypen, X. Zhou, M. Sherwood, C. Yannoni, M. Kubinec, and I. L. Chuang. Experimental realization of a two-bit phase damping quantum code. *Phys. Rev. A*, 60: 1924, 1999.

[Man80] Y. Manin. *Computable and Uncomputable* (*in Russian*). Sovetskoye Radio, Moscow, 1980.

[Man99] Y. I. Manin. Classical computing, quantum computing, and Shor's factoring algorithm, *arXive e-print quant-ph/9903008*, 1999.

[Mau93] U. M. Maurer. Secret key agreement by public discussion from common information. *IEEE Trans. Inf. Theory*, 39:733—742, 1993.

[Max71] J. C. Maxwell. *Theory of Heat*. Longmans, Green, and Co., London, 1871.

[May98] D. Mayers. Unconditional security in quantum cryptography, *arXive e-print quant-ph/9802025*, 1998.

[ME99] M. Mosca and A. Ekert. The hidden subgroup problem and eigenvalue estimation on a quantum computer, *arXive e-print quant-ph/9903071*, 1999.

[Mer78] R. Merkle. Secure communications over insecure channels.

Comm. of the ACM, 21: 294－299, 1978.

[Mil76] G. L. Miller. Riemann's hypothesis and tests for primality. *J. Comput. Syst. Sci.*, 3 (3): 300－317, 1976.

[Mil89a] G. J. Milburn. Quantum optical Fredkin gate. *Phys. Rev. Lett.*, 62(18): 2124, 1989.

[Mil89b] D. A. B. Miller. Optics for low energy communications inside digital processors: quantum detectors, sources, and modulators as efficient impedance converters. *Opt. Lett.*, 14: 146, 1989.

[Mil96] G. J. Milburn. A quantum mechanical Maxwell's demon. Unpublished, 1996.

[Mil97] G. J. Milburn. Scr6dinger's Machines: *the Quantum Technology Reshaping Everyday Life*. W. H. Freeman, New York, 1997.

[Mil98] G. J. Milburn. *The Feynman Processor: Quantum Entanglement and the Computing Revolution*. Perseus Books, Reading, Mass., 1998.

[Min67] M. L. Minsky. *Computation: finite and infinite machines*. Prentice-Hall, Englewood Cliffs, N. J., 1967.

[MM92] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix InequaliTies*. Dover, New York, 1992.

[MMK+95] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland. Demonstration of a fundamental quantum logic gate. *Phys. Rev. Lett.*, 75: 4714, 1995.

[MO79] A. W. Marshall and I. Olkin. Inequalities: *Theory of Majorization and its Applications*. Academic Press, New York, 1979.

[MOL+99] J. E. Mooij, T. P. Orlando, L. Levitov, L. Tian, C. H. van der Waal, and S. Lloyd. Josephson persistent-current qubit. *Science*, 285: 1036－1039, 1999.

[Mor98] T. Mor. No-cloning of orthogonal states in composite systems. *Phys. Rev. Lett.*, 80: 3137－3140, 1998.

[Mos98] M. Mosca. Quantum searching, counting and amplitude amplification by eigenvector analysis. In R. Freivalds, editor, *Proceedings of International Workshop on Randomized Algorithms*, pages 90－100, 1998.

[Mos99] M. Mosca. *Quantum Computer Algorithms*. Ph. D. thesis, University of Oxford, 1999.

[MR95] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, Cambridge, 1995.

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-correcting Codes*. North-Holland, Amsterdam, 1977.

[MU88] H. Maassen and J. H. B. Uffink. Generalized entropic uncertainty relations. *Phys. Rev. Lett.*, 60(12): 1103 — 1106, 1988.

[MvOV96] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.

[MWKZ96] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger. Dense coding in experimental quantum communication. *Phys. Rev. Lett.*, 76(25): 4656 — 4659, 1996.

[MZG96] A. Muller, H. Zbinden, and N. Gisin. Quantum cryptography over 23 km in installed under-lake telecom fibre. *Europhys. Lett.*, 33: 334 — 339, 1996.

[NC97] M. A. Nielsen and C. M. Caves. Reversible quantum operations and their application to teleportation. *Phys. Rev.*, 4, 55(4): 2547—2556, 1997.

[NCSB98] M. A. Nielsen, C. M. Caves, B. Schumacher, and H. Barnum. Information-theoretic approach to quantum error correction and reversible measurement. *Proc. R. Soc. London A*, 454(1969): 277—304, 1998.

[Nie98] M. A. Nielsen. *Quantum Information Theory. Ph. D. thesis*, University of New Mexico, 1998.

[Nie99a] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83(2): 436—439, 1999.

[Nie99b] M. A. Nielsen. Probability distributions consistent with a mixed state, *arXive e-print quant-ph/9909020*, 1999.

[NKL98] M. A. Nielsen, E. Knill, and R. Laflamme. Complete quantum teleportation using nuclear magnetic resonance. *Nature*, 396(6706): 52—55, 1998.

[NPT99] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai. Coherent control of macroscopic quantum states in a single-cooper-pair box. *Natur*, 398: 786 — 788, 1999.

[OP93] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer-Verlag, Berlin, 1993.

[Pai82] A. Pals. Subtle is the Lord: *The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982.

[Pai86] A. Pals. *Inward Bound: Of Matter and Forces in the Physical World*. Oxford University Press, Oxford, 1986.

[Pai91] A. Pals. *Niels Bohr's Times: In Physics, Philosophy, and Polity*. Oxford University Press, Oxford, 1991.

[Pap94] C. M. Papadimitriou. *Computational Complexity*. Addison-Wesley, Reading, Massachusetts, 1994.

[Pat92] R. Paturi. On the degree of

polynomials that approximate symmetric Boolean functions (preliminary version). *Proc. 24th Ann. ACM Syrup. on Theory of Computing (STOC'92)*, pages 468—474, 1992.

[PCZ97] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: the two-bit quantum gate. *Phys. Rev. Lett.*, 78(2):390—393, 1997.

[PD99] P. M. Platzman and M. I. Dykman. Quantum computing with electrons floating on liquid helium. *Science*, 284:1967, 1999.

[Pen89] R. Penrose. *The Emperor's New Mind*. Oxford University Press, Oxford, 1989.

[Per52] S. Perils. *Theory of Matrices. Addison-Wesley, Reading, Mass.*, 1952.

[Per88] A. Peres. How to differentiate between non-orthogonal states. *Phys. Lett. A*, 128:19, 1988.

[Per93] A. Peres. *Quantum Theory: Concepts and Methods. Kluwer Academic*, Dordrecht, 1993.

[Per95] A. Peres. Higher order schmidt decompositions. *Phys. Lett. A*, 202:16—17, 1995.

[Pet86] D. Petz. Quasientropies for finite quantum systems. *Rep. Math. Phys.*, 23(1):57—65, 1986.

[Phy92] Physics Today Editor. Quantum cryptography defies eavesdropping. *Physics Today*, page 21, November 1992.

[PK96] M. B. Plenio and P. L. Knight. Realistic lower bounds for the factorization time of large numbers on a quantum computer. *Phys. Rev. A*, 53:2986—2990, 1996.

[PK98] M. B. Plenio and P. L. Knight. The quantum-jump approach to dissipative dynamics in quantum optics. *Rev. 34od. Phys.*, 70(1): 101 — 144, 1998.

[Pop75] R. P. Poplavskii. Thermodvnamical models of information processing (in Russian). *Usp. Fiz. Nauk*, 115(3): 465 — 501, 1975.

[PRB98] M. Pueschel, M. Roetteler, and T. Beth. Fast quantum Fourier transforms for a class of non-abelian groups, *arXive e-print quant-ph/9807064*, 1998.

[Pre97] J. Preskill. Fault-tolerant quantum computation, *arXive e-print quant-ph/9712048*, 1997.

[Pre98a] J. Preskill. Fault-tolerant quantum computation. In H.-K. Lo, T. Spiller, and S. Popescu, editors, *Quantum information and computation*. World Scientific, Singapore, 1998.

[Pre98b] J. Preskill. *Physics 229: Advanced Mathematical Methods of Physics--Quantum Computation and Information*. California Institute of Technology,

1998. *URL: htlp://www.theory. caltech. edu/people/preskill/ph229/*

[Pre98c] J. Preskill. Reliable quantum computers. *Proc. R. Soc. London A*, 454(1969):385—410, 1998.

[Rab80] M. O. Rabin. Probabilistic algorithm for testing primality. *J. Number Theory*, 12:128—138, 1980.

[Rah99] H. Z. Rahim. Richard Feynman and Bill Gates: an imaginary encounter. 1999. *URL: http://www. trnsoft. com/features/lrfbg. Htm*

[Rai98] E. M. Rains. Quantum weight enumerators. *IEEE Trans. Inf. Theory*, 44(4):1388—1394, 1998.

[Rai99a] E. M. Rains. Monotonicity of the quantum linear programming bound. *IEEE Trans. Inf. Theory*, 45(7):2489—2492, 1999.

[Rai99b] E. M. Rains. Nonbinary quantum codes. *IEEE Trans. Inf. Theory*, 45(6):1827—1832, 1999.

[Rai99c] E. M. Rains. Quantum shadow enumerators. *IEEE Trans. Inf. Theory*, 45(7):2361—2366, 1999.

[RB98] M. Roetteler and T. Beth. Polynomialtime solution to the hidden subgroup problem for a class of nonabelian groups. *arXive e-print quant-ph*/9812070, 1998.

[Res81] A. Ressler. *The Design of a Conservative Logic Computer and A Graphical Editor Simulator*. Master's thesis, Massachusetts Institute of Technology, 1981.

[RHSS97] E. M. Rains, R. H. Hardin, P. W. Shor, and N. J. A. Sloane. Nonadditive quantum code. *Phys. Rev. Lett*, 79(5):953—954, 1997.

[Roy96] A. Royer. Reduced dynamics with initial correlations, and time-dependent environment and Hamiltonians. *Phys. Rev. Lett.*, 77(16):3272—3275, 1996.

[RR67] D. W. Robinson and D. Ruelle. *Commun. Math. Phys.*, 5:288, 1967.

[RSA78] R. L. Rivest, A. Shamir, and L. M. Adleman. A method of obtaining digital signatures and public-key cryptosystems. *Comm. ACM*, 21(2):120—126, 1978.

[Rus94] M. B. Ruskai. Beyond strong subadditivity: improved bounds on the contraction of generalized relative entropy. *Rev. Math. Phys.*, 6(5A):1147—1161, 1994.

[RWvD84] S. Ramo, J. R. Whinnery, and T. van Duzer. *Fields and waves in communication electronics*. Wiley, New York, 1984.

[RZBB94] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental realization of any dis-

crete unitary operator. *Phys. Rev. Lett.*, 73 (1): 58431, 1994.

[Sak95] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, Reading, Mass., 1995.

[SC99] R. Schack and C. M. Caves. Classical model for bulk-ensemble NMR quantum computation. *Phys. Rev. A*, 60(6): 4354—4362, 1999.

[Sch06] E. Schmidt. Zur theorie der linearen und nichtlinearen integralgleighungen. *Math. Annalen.*, 63:433—476, 1906.

[Sch36] E. Schr6dinger. Probability relations between separated systems. *Proc. Cambridge Philos. Soc.*, 32:446—452, 1936.

[Sch95] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738—2747, 1995.

[Sch96a] B. Schneier. *Applied Cryptography*. John Wiley and Sons, New York, 1996.

[Sch96b] B. W. Schumacher. Sending entanglement through noisy quantum channels. *Phys. Rev. A*, 54:2614, 1996.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27: 379—423, 6234356, 1948.

[Sho94] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, IEEE Press, Los Alamitos, CA, 1994.

[Sho95] P. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:2493, 1995.

[Sho96] P. W. Shor. Fault-tolerant quantum computation. In *Proceedings, 37th Annual Symposium on Fundamentals of Computer Science*, pages 56—65, IEEE Press, Los Alamitos, CA, 1996.

[Sho97] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM. J. Comp.*, 26 (5): 1484—1509, 1997.

[Sim79] B. Simon. *Trace Ideals and Their Applications*. Cambridge University Press Cambridge, 1979.

[Sim94] D. Simon. On the power of quantum computation. In *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, pages 116—123, IEEE Press, Los Alamitos, CA, 1994.

[Sim97] D. R. Simon. On the power of quantum computation. *SIAM J. Comput.*, 26(5): 1474—1483, 1997.

[SL97] P. W. Shor and R. Laflamme. Quantum analog of the MacWilliams identities for classical coding theory. *Phys. Rev.*

Lett., 78(8): 1600 — 1602, 1997.

[SL98] D. Shasha and C. Lazere. *Out of Their Minds: The Lives and Discoveries 15 Great Computer Scientists*. Springe: Verlag, New York, 1998.

[Sle74] D. Slepian, editor. *Keys Papers in the Development of Information Theory IEEE Press*, New York, 1974.

[Sli96] C. P. Slichter. *Principles of Magnet Resonance. Springer*, Berlin, 1996.

[SN96] B. W. Schumacher and M. A. Nielse. Quantum data processing and error correction. *Phys. Rev. A*, 54(4): 2629, 1996. *arXive e-print quant-ph*/9604022.

[SP00] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol, *arXive e-print quant-ph*/003004, 2000.

[SS76] R. Solovay and V. Strassen. A fast MonteCarlo test for primality. *SIAM J. Comput.*, 6: 84—85, 1976.

[SS96] P. W. Shor and J. A. Smolin. Quantum error-correcting codes need not completely reveal the error syndrome, *arXive e-print quant-ph*/9604006, 1996.

[SS99] A. T. Sornborger and E. D. Stewart. Higher order methods for simulations on quantum computers. *Phys. Rev. A*, 60(3): 1956 — 1965, 1999. *arXive e-print quant-ph*/9903055.

[ST91] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. Wiley, NY, 1991.

[Ste96a] A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77: 793, 1996.

[Ste96b] A. M. Steane. Multiple particle interference and quantum error correction. *Proc. R. Soc. London A*, 452: 2551 — 2576, 1996.

[Ste97] A. Steane. The ion-trap quantum information processor. *Appl. Phys. B-Lasers and Optics*, 64(6): 623 — 642, 1997.

[Ste99] A. M. Steane. Efficient fault-tolerant quantum computing. *Nature*, 399: 124 — 126, May 1999.

[STH+99] S. Somaroo, C. H. Tseng, T. F. Havel, R. Lafiamme, and D. G. Cory. Quantum simulations on a quantum computer. *Phys. Rev. Lett.*, 82: 5381 — 5384, 1999.

[Str76] G. Strang. *Linear Algebra and Its Applications*. Academic Press, New York, 1976.

[SV99] L. J. Schulman and U. Vazirani. Molecular scale heat engines and scalable quantum computation. *Proc. 31st Ann. ACM Symp. on Theory of Computing* (*STOC*'99), pages

322－329, 1999.

[SW49] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[SW93] N. J. A. Sloane and A. D. Wyner, editors. *Claude Elwood Shannon: Collected Papers*. IEEE Press, New York, 1993.

[SW97] B. Schumacher and M. D. Westmoreland. Sending classical information vianoisy quantum channels. *Phys. Rev. A*, 56(1):131－138, 1997.

[SW98] B. Schumacher and M. D. Westmoreland. Quantum privacy and quantum coherence. *Phys. Rev. Lett.*, 80 (25): 5695－5697, 1998.

[SWW96] B. W. Schumacher, M. Westmoreland, and W. K. Wootters. Limitation on the amount of accessible information in a quantum channel. *Phys. Rev. Lett.*, 76:3453, 1996.

[Szi29] L. Szilard. Uber die entropieverminderung in einen thermodynamischcn system bei eingriffen intelligenter wesen. *Z. Phys.*, 53:840－856, 1929.

[TD98] B. M. Terhal and D. P. DiVincenzo. The problem of equilibration and the computation of correlation functions on a quantum computer, *arXive e-print quant-ph*/9810063, 1998.

[THL+95] Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble. Measurement of conditional phase shifts for quantum logic. *Phys. Rev. Lett.*, 75:4710, 1995.

[Tro59] H. F. Trotter. On the product of semigroups of operators. *Proc. Am. Math. Soc.*, 10: 545－551, 1959.

[Tsia0] B. S. Tsirelson. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.*, 4:93, 1980.

[Tur36] A. M. Turing. On computable numbers, with an application to the Entschei-dungsproblem. *Proc. Lond. Math. Soc.* 2 (*reprinted in* [*Dav65*]), 42:230, 1936.

[Tur97] Q. A. Turchette. *Quantum optics with single atoms and single photons*. Ph. D. thesis, California Institute of Technology, Pasadena, California, 1997.

[Uhl70] A. Uhlmann. On the Shannon entropy and related functionals on convex sets. *Rep. Math. Phys.*, 1(2):147－159, 1970.

[Uhl71] A. Uhlmann. Sätze über dichtematrizen. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 20: 633－637, 1971.

[Uhl72] A. Uhlmann. Endlich-dimensionale dichtematrizen I. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 21:421－452, 1972.

[Uhl73] A. Uhlmann. Endlich-dimen-

sionale dichtematrizen Ⅱ. *Wiss. Z. Karl-Marx-Univ. Leipzig*, 22:139—177, 1973.

[Uhl76] A. Uhlmann. The 'transition probability' in the state space of a *-algebra. *Rep. Math. Phys.*, 9:273—279, 1976.

[Uhl77] A. Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Commun. Math. Phys.*, 54:21—32, 1977.

[Ume62] H. Umegaki. *Kōdai Math. Sem. Rep.*, 14:59—85, 1962.

[Vai94] L. Vaidman. Teteportation of quantum states. *Phys. Rev. A*, 49(2):1473—1476, 1994.

[van98a] W. van Dam. Quantum oracle interrogation: getting all information for half the price. In *Proceedings of the* 39th *FOCS*, 1998. *arXive e-print quant-ph*/9805006.

[van98b] S. J. van Enk. No-cloning and superluminal signaling, *arXive e-print quant-ph*/9803030, 1998.

[Ved99] V. Vedral. Landauer's erasure, error correction and entanglement, *arXive e-print quant-ph*/9903049, 1999.

[Vid98] G. Vidal. Entanglement monotones. *arXive e-print quant-ph*/9807077, 1998.

[Vid99] G. Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83(5):1046—1049, 1999.

[von27] J. von Neumann. *Göttinger Nachrichten*, page 245, 1927.

[von56] J. von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In *Automata Studies*, pages 329—378, Princeton University Press, Princeton, NJ, 1956.

[von66] J. yon Neumann. Fourth University of Illinois lecture. In A. W. Burks, editor, *Theory of Self-Reproducing Automata*, page 66, University of Illinois Press, Urbana, 1966.

[VP98] V. Vedral and M. B. Plenio. Entanglement measures. *Phys. Rev. A*, 57(3):1619—1633, 1998.

[VR89] K. Vogel and H. Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 40(12):7113—7120, 1989.

[VYSC99] L. M. K. Vandersypen, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang. Realization of effective pure states for bulk quantum computation. *Phys. Rev. Lett.*, 83:3085—3088, 1999.

[VYW+99] R. Vrijen, E. Yablonovitch, K. Wang, H. W. Jiang, A. Balandin, V. Roychowdhury, T. Mor, and D. DiVincenzo. Electron spin resonance tran-

sistors for quantum computing in silicon-germanium heterostructures, *arXive e-print quant-ph*/9905096, 1999.

[CWar97] W. Warren. The usefulness of NMR quantum computing. *Science*, 277(5332):1688, 1997.

[Wat99] J. Watrous. PSPACE has around quantum interactive proof systems, *arXive e-print cs*/9901015, 1999.

[WC67] S. Winograd and J. D. Cowan. *Reliable Computation in the Presence of Noise*. MIT Press, Cambridge, MA, 1967.

[Weh78] A. Wehrt. General properties of entropy. *Rev. Mod. Phys.*, 50:221, 1978.

[Wel188] D. J. A. Welsh. *Codes and Cryptography*. Oxford University Press, New York, 1988.

[Wie] S. Wiesner. Unpublished manuscript circa 1969, appeared as [Wie83].

[Wie83] S. Wiesner. Conjugate coding. *SIGACT News*, 15:77, 1983.

[Wie96] S. Wiesner. Simulations of manybody quantum systems by a quantum computer, *arXive e-print quant-ph*/9603028, 1996.

[Wil91] D. Williams. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

[Win98] E. Winfree. *Algorithmic Self-Assembly of DNA*. Ph. D. thesis, California Institute of Technology, Pasadena, California, 1998.

[WMI+98] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof. Experimental issues incoherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl. Inst. Stand. Tech.*, 103: 259, 1998.

[WS98] M. D. Westmoreland and B. Schumacher. Quantum entanglement and the non-existence of superluminal signals, *arXive e-print quant-ph*/9801014, 1998.

[WY63] E. P. Wigner and M. M. Yanase. *Proc. Natl. Acad. Sci.* (*U. S. A.*), 49:910—918, 1963.

[WY90] K. Watanabe and Y. Yamamoto. Limits on tradeoffs between third-order optical nonlinearity, absorption loss, and pulse duration in self-induced transparency and real excitation. *Phys. Rev. A*, 42(3):1699—1702, 1990.

[WZ82] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299: 802—803, 1982.

[Yao93] A. C. Yao. Quantum circuit complexity. *Proc. of the 34th Ann. IEEE Symp. on Foundations of Computer Science*, pages 352—361, 1993.

[YK95] S. Younis and T. Knight. Non dissipative rail drivers for adiabatic circuits. In *Proceedings, Sixteenth Conference on*

Advanced Research in VLSI 1995, pages 404－414, IEEE Computer Society Press, Los Alamitos, CA, 1995.

[YKI88] Y. Yamamoto, M. Kitagawa, and K. Igeta. In *Proc. 3rd Asia-Pacific Phys. Conf.*, World Scientific, Singapore, 1988.

[YO93] H. P. Yuen and M. Ozawa. Ultimate information carrying limit of quantum systems. *Physical Review Letters*, 70: 363－366, 1993.

[YY99] F. Yamaguchi and Y. Yamamoto. Crystal lattice quantum computer. *Appl. Phys. A*, pages 1－8, 1999.

[Zal98] C. Zalka. Simulating quantum systems on a quantum computer. *Proc. R. Soc. London A*, 454(1969):313－322, 1998.

[Zal98] C. Zalka. Grover's quantum searching algorithm is optimal. *Physical Review A*, 60: 2746－2751, 1999.

[Zan99] P. Zanardi. Stabilizing quantum information. *arXive e-print quant-ph/9910016*, 1999.

[ZG97] P. Zoller and C. W. Gardiner. Quantum noise in quantum optics: the stochastic Schrödinger equation. In S. Reynaud, E. Giacobino, and J. Zinn-Justin, editors, *Quantum Fluctuations: Les Houches Summer School LXIII*, Elsevier, Amsterdam, 1997.

[ZHSL99] K. Zyczkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58(2):883－892, 1999.

[ZL96] W. H. Zurek and R. Lafiamme. Quantum logical operations on encoded qubits. *Phys. Rev. Lett.*, 77(22): 4683－4686, 1996.

[ZLC00] X. Zhou, D. W. Leung, and I. L. Chuang. Quantum logic gate constructions with one-bit "teleportation". *arXive e-print quant-ph/O* 002039, 2000.

[ZR98] P. Zanardi and M. Rasetti. Noiseless quantum codes. *Phys. Rev. Lett.*, 79(17): 3306－3309, 1998.

[Zur89] W. H. Zurek. Thermodynamic cost of computation, algorithmic complexity and the information metric. *Nature*, 341: 119, 1989.

[Zur91] W. H. Zurek. Decoherence and the transition from quantum to classical. *Phys. Today*, October 1991.