量子信息简明教程

马雄峰 张行健 黄溢智 著

清華大學出版社
北京

内 容 简 介

本书从量子力学与信息论的基础出发，系统且完整地介绍了量子信息领域的基础知识，包括量子信息科学所需的线性代数与信息论数学基础、量子系统的基本描述与一般描述、量子系统中的现象与应用，以及量子信息论与量子纠缠的初步介绍。此外，本书还涵盖了一些近期量子信息领域研究方向的发展。

本书可作为量子信息专业的基础课程或物理专业的通识课程教材使用。

图书在版编目（CIP）数据

量子信息简明教程 / 马雄峰, 张行健, 黄溢智著.—北京：清华大学出版社，2023.5
ISBN 978-7-302-63114-9

Ⅰ. ①量… Ⅱ. ①马… ②张… ③黄… Ⅲ. ①量子力学–信息技术–教材 Ⅳ. ①O413.1

中国国家版本馆 CIP 数据核字（2023）第 047578 号

责任编辑：孙亚楠
封面设计：常雪影
责任校对：赵丽敏
责任印制：丛怀宇

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn, http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编：**100084
社 总 机：010-83470000　　**邮　　购：**010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市东方印刷有限公司
经　　销：全国新华书店
开　　本：170mm×240mm　　**印　张：**11.75　　**字　　数：**234 千字
版　　次：2023 年 7 月第 1 版　　**印　　次：**2023 年 7 月第 1 次印刷
定　　价：49.00 元

产品编号：100085-01

前言 Preface

19 世纪末，物理学界普遍认为基本的物理原则已被牢固地建立起来，物理学的未来“要在小数点的第 6 位找到”。但在解释黑体辐射的“紫外灾变”问题时，普朗克的能量量子化却颠覆了传统的统计力学理论，由此引出了物理学的新篇章——量子力学。量子理论无论在解释微观物理现象还是工程应用上都迅速发挥了巨大作用，激光的发明和大规模集成线路的应用更是让“量子”这一概念深入人心。尽管量子理论与我们所观测的物理现象高度吻合，但在如何理解其基本原理这一问题上，却始终存在着巨大的争论。其中最有名的，是爱因斯坦和玻尔对物理测量是否可以产生不可预测随机性的世纪之争。随着贝尔不等式的提出及其实验验证，今天大部分人接受了量子力学及其蕴含的内禀随机性。当然，经典力学，包括牛顿力学和麦克斯韦的电磁方程，在很多情况下，特别是对许多宏观物理现象的解释中，依然是非常好的一个近似。

20 世纪科技领域另外一件大事是计算机的发明和信息技术的发展。理论方面，图灵、冯·诺依曼建立起了通用计算任务的数学模型，明确了可计算与不可计算任务的分界线；香农将信息的概念从具体的文字、图像、声音中抽象为概率与信息熵；在对复杂动力系统的研究过程中，人们又逐渐建立了计算复杂性、通信复杂性的概念。与此同时，硬件制造上的突飞猛进也支持了信息技术的高速发展。随着半导体工艺、存储技术、光纤等的不断迭代更新，早期庞大、低效、单一的电子管计算机已经发展成了小型、高速、可以通过互联网连接的计算网络。无须多言，计算机与信息技术的力量已经深入生产生活的各个方面。

在量子力学和信息技术的发展过程中，人们逐渐意识到，信息是物理的，物理也是信息的，可以利用量子力学的手段来处理信息。在计算领域，这一可能性最早由苏联数学家马宁与美国物理学家费曼等人意识到。为了分析许多复杂的物理过程，比如黑洞的动力学演化，与其用电子计算机“离散”地求解方程，不妨将其自然地对应到相似的物理过程，比如凝聚态物理中的现象，直接用这一量子过程进行“模拟”计算。同一时期，在通信与密码领域，人们注意到利用量子力学的内禀随机性，可以实现量子密钥分发等具有信息论安全性的隐私通信——这在经典世界甚至是完全不可能的！随着研究的深入，人们逐渐发掘出基于量子原理进行安全信息传输的巨大潜力，发现了具有指数加速的量子算法，超越仅用经典物理过程进行精密测量的精度，等等。于是，量子信息科学这个交叉领域应运而生，并渐渐成为计算机科学与物理学的重要分支。你能听到计算机科学家们在讨论量子复杂性、量子货币，也能看到物理学家们在研究黑洞信息熵、量子人工

智能。2022 年的诺贝尔物理学奖授予了三位在贝尔不等式、量子纠缠等方向做出了开创性研究的物理学家，以表彰他们为奠定量子信息科学基础的重要贡献。量子信息无疑是近年来最火爆、最有活力的科学研究方向之一。

不止是理论研究本身，经过四十余年的发展，关于量子信息的一些初期理论构想、实验演示，现在已经成为快速发展的新兴产业。近期，谷歌、中国科学技术大学等团队展示了“量子优越性”，验证了量子计算相比于经典计算的优势；我国与欧盟多国之间已经建成了天地一体化的洲际量子通信网络。大量量子信息相关企业也陆续成立，国内的国盾量子、问天量子，国外的 MagiQ、ID Quantique、Xanadu、Rigetti Computing 等，都得到了投资人的青睐。而老牌的互联网和计算机硬件公司，谷歌、IBM、亚马逊、腾讯、阿里、百度等，也都不甘落于人后，纷纷成立了自己的量子实验室。一些量子应用，比如量子密钥分发、量子随机数产生、量子精密测量，不仅已经有了相对成熟的商业化产品，相关的标准化工作也已开始推进，为进一步的产业化铺垫道路。

在前沿研究之外，现在国内外许多一流高校、研究所已经开设了大量量子信息相关课程，作者在清华大学也已开展了近十年的量子信息教学。2021 年，在我国教育部新增的高校本科专业中，就包含了量子信息科学。在此背景下，系统且兼具前沿性的量子信息教学变得愈发重要。从 20 世纪 90 年代开始，国际上陆续有不少很好的量子信息方面的教材，国内的一些学者也引入了这些教材，并进行了很好的翻译。然而，在本书作者参与大学课程教学的时候发现，这些教材大多并不适合一学期的教学，特别是为初学者入门使用。经过近十年教学的积累，我们把相关的课件整理，形成了这一本书。这里，我们主要介绍量子信息科学的基本概念，而不深入讨论量子密码学、量子计算等专业方向知识。希望这样一本书可以作为理工科大学量子信息教材或者参考书，在大约一学期的课程时长内帮助同学们掌握量子信息理论的基础，扫清他们进一步了解甚至进入量子信息前沿领域研究的知识障碍。同时，也希望本书可以作为一本自学的书籍，特别为具有一定线性代数背景且对量子信息感兴趣的读者，能够帮助他们了解一些在文献阅读中可能会碰到的量子信息的专业术语与概念。

在本书的内容安排上，我们在第 0 章首先介绍必要的数学知识，以及量子信息领域内常用的符号表示，包括量子力学基本描述、线性代数、概率论、香农信息论基础等内容。对这些内容熟悉的读者，可以很快地浏览相关内容，然后进入对量子信息知识的学习中。在本书的主体部分，我们依次介绍量子系统的基本描述、量子系统的一般描述、一些有趣的量子现象和应用、量子信息论初步。具体而言，

（1）在第 1 章，通过单体量子系统，说明在量子力学中如何表示一个物理系统，给出可观测量的定义及封闭系统演化的数学描述，在此基础上，介绍实验中确定量子态的基本方法，以及说明量子力学所给出的信息守恒规律。

（2）在第 2 章，将考虑更为一般的多体量子系统，在对子系统的描述过程中，将引入必要的统计力学方法——密度矩阵，以及偏迹等数学运算，并由此给出最一般量子态的定义及性质；此外，也将给出最一般的测量过程、量子操作、量子系统演化的定义。

（3）在第 3 章，将综合运用第 1 章和第 2 章的内容，研究三种有趣的量子现象和应用：贝尔不等式、量子密集编码、量子隐形传态。通过这些内容，我们将更加深刻地认识到量子力学和经典力学之间的差异，并初步感受到量子力学在信息传输方面的价值。

（4）在第 4 章，将初步介绍量子信息论。我们首先将香农信息论中信息熵的概念推广到量子世界中。量子信息熵相比于经典信息熵有许多不同之处，导致这些差异的一个重要原因是量子纠缠，对此我们将简单介绍量子纠缠的一些核心概念。最后，将初步介绍量子通信与编码，并给出系统性研究此问题的框架。

若读者拥有良好的线性代数基础，大多数内容通过学习本书前序知识，便可以无障碍地逐渐深入。有些章节难度较高，或需要额外的前置知识帮助理解，对这些内容，我们用星号予以标注。在第一次学习或基础课程讲解时，可以跳过这些内容。此外，我们对每章内容都配以一些习题，鼓励读者们进行适当的练习，以巩固所学内容。

这里，我们要特别感谢过去这些年作者开展量子信息相关课程时的助教们：袁骁、张振、赵琦、彭天翼、周泓伊、曾培、陈森睿、吴蔚捷、张艺泓、唐一凡、余文峻、刘振寰、鄢语轩、陈俊杰、张辰逸、刘国定、刘鹏宇。他们很大程度上参与了课件底稿的准备。事实上，两位作者张行健和黄溢智做过多次相关课程的助教。同时我们也非常感谢参与这些课程的同学们的反馈。正是来自他们提出的很好的问题和关于教学的反馈，使得我们的教材能够更加全面、易读。

目录 Contents

第 0 章

线性代数与经典信息论基础

在本章中，将介绍量子信息的数学与信息论基础。量子信息的描述主要用到了线性代数中的向量和矩阵。我们从线性空间最基本的相关记号与定义出发，介绍线性代数的一些基本内容，包括线性空间和矩阵运算等。一般来讲，量子信息相关的讨论都是在复数域 $\mathbb{C}$ 上进行的。另外，还介绍了经典信息论中的基本概念，包括香农编码定理和不同的信息熵。这里，信息熵本质上就是热力学熵。量子信息科学一个重要分支就是用信息科学的语言来描述量子力学。

在本章中，0.1 节介绍了本书中常见的符号及其对应的含义；0.2 节阐述了希尔伯特空间的描述，特别是狄拉克符号的表示；0.3 节则回顾了量子信息中用到较多的矩阵运算；0.4 节主要介绍了信息论的一些基础知识。在这些小节中，0.2 节和 0.3.1 节是本书的基础，若不能很好掌握，建议从线性代数教材入手，系统性地补充相关知识。0.3 节中的其余小节将会在本书后续章节展开。0.3.6 节为选读，本书后续部分用到的较少。0.4 节主要应用在第 3 章和第 4 章。对于这些小节的内容，读者可以先选择性地进行阅读，在后续章节中遇到相应内容再返回第 0 章进行仔细的学习。

0.1 数学记号

本书采用的符号见表 0.1。

表 0.1　本书采用的数学记号

记号	描述
$[d]$	指标集 $\{0,1,2,\cdots,d-1\}$
$\mathbb{C}$	复数域
$\|i\rangle$	计算基矢的第 i 个向量, 第 i 个元素取值为 1, 其余为 0
$(\boldsymbol{A})_{ij}$	矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列的元素, 或简记为 $(\boldsymbol{A})_{ij}=a_{ij}$
$\boldsymbol{I}$	单位矩阵，对角元素全部为 1, 其余元素均为 0
$\boldsymbol{0}_{nm}$	复矩阵空间 $\mathbb{C}^{n\times m}$ 中的零矩阵，全部元素均为 0
厄米	$\boldsymbol{A}\in\mathbb{C}^{d\times d}$: $\boldsymbol{A}^\dagger=\boldsymbol{A}$
幺正	$\boldsymbol{U}\in\mathbb{C}^{d\times d}$: $\boldsymbol{U}\boldsymbol{U}^\dagger=\boldsymbol{U}^\dagger\boldsymbol{U}=\boldsymbol{I}$，有时也音译为酉
$\mathcal{H}_d$	d 维希尔伯特空间: 具有内积结构的复线性空间 $\mathbb{C}^d$
$\mathcal{H}_S$	系统 S 对应的希尔伯特空间
$\dim(\mathcal{H})$	希尔伯特空间 $\mathcal{H}$ 维度

续表

记号	描述
$\mathcal{L}(\mathcal{H})$	希尔伯特空间 $\mathcal{H}$ 到自身的线性映射算子集合
$\mathcal{D}(\mathcal{H})$	作用于希尔伯特空间 $\mathcal{H}$ 上的密度算子集合
$\left\|\Phi_d^+\right\rangle$	$\mathcal{H}_d$ 内最大纠缠态，$\sum_{i=0}^{d-1}\|ii\rangle/\sqrt{d}$
$\|\phi\rangle\in\mathcal{H}$	纯态: $\phi\equiv\|\phi\rangle\langle\phi\|\in\mathcal{D}(\mathcal{H})$

0.2 希尔伯特空间表示

在量子信息领域，通常采用狄拉克符号（Dirac's bra-ket notation）标记量子态、量子操作、量子测量等。这里，介绍如何利用狄拉克符号对希尔伯特空间进行表示。

0.2.1 狄拉克符号

对于由列向量组成的复线性空间 $\mathbb{C}^{d\times 1}$, 一组常用的正交归一基可以表示为

$$\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix},\ \cdots,\ \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix} \tag{0.1}$$

每个基向量均只有一个元素为 1，而其他元素为 0。如果将上述向量按顺序组合在一起，它们将构成 d 维单位矩阵 $\boldsymbol{I}\in\mathbb{C}^{d\times d}$。利用狄拉克符号，将这些列向量表示为右矢（ket）：

$$|0\rangle=\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\ |1\rangle=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\ |2\rangle=\begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix},\ \cdots,\ |d-1\rangle=\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix} \tag{0.2}$$

这里，为了方便量子信息中的讨论，从指标 0 开始排序。记指标集 $[d]\equiv\{0,1,2,\cdots,d-1\}$。这样可以用 $\{|i\rangle\}_{i=0}^{d-1}$ 或者 $\{|i\rangle\}_{i\in[d]}$ 来表示这个右矢集合，通常也将这组特殊的基矢称为计算基矢（computational basis）。线性空间 $\mathbb{C}^{d\times 1}$ 中的任一向量均可以表示为一个右矢 $|\psi\rangle$, 为上述选取的基向量的线性组合：

$$
\begin{aligned}
|\psi\rangle &= \sum_{i=0}^{d-1} x_i |i\rangle \\
&= x_0 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_1 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_{d-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{d-1} \end{pmatrix}
\end{aligned} \tag{0.3}
$$

其中，$x_i \in \mathbb{C}$。当然，除了计算基矢以外，希尔伯特空间中任意一组正交归一的向量都可以构成一组基矢。通常将计算基矢记为 $\{|i\rangle\}_{i=0}^{d-1}$，而其他基矢记为 $\{|\phi_i\rangle\}_{i=0}^{d-1}$。

与右矢相对应，将右矢的厄米共轭（Hermitian conjugate）向量称为**左矢**（bra），$\langle\psi| \equiv (|\psi\rangle)^\dagger$。在这里，厄米共轭操作 † 将一个行向量转换为一个列向量，并对向量的每一个元素取复数共轭（反之亦然）。类似地，可以写下由行向量组成的复线性空间 $\mathbb{C}^{1\times d}$ 的一组自然正交归一基：

$$
\begin{cases}
\langle 0| = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \\
\langle 1| = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \\
\qquad\qquad \vdots \\
\langle d-1| = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 \end{pmatrix}
\end{cases} \tag{0.4}
$$

线性空间 $\mathbb{C}^{1\times d}$ 中的任一行向量可以表示为

$$
\langle\psi| = \sum_{i=0}^{d-1} x_i^* \langle i| \tag{0.5}
$$

这里 x_i^* 表示 x_i 的复共轭。使用狄拉克符号的主要好处是，它可以帮助我们简单明了地区分列向量 $|\cdot\rangle$ 和行向量 $\langle\cdot|$，简化线性代数中运算的表示，例如，下面将要介绍的内积运算与投影运算。

0.2.2 希尔伯特空间

在量子力学中，希尔伯特空间是复数域 $\mathbb{C}$ 上具有内积结构的由列向量张成的线性空间，记为 $\mathcal{H}$。列向量由右矢表示，$|\psi\rangle \in \mathbb{C}^{d\times 1}$。两个右矢，例如 $|\psi\rangle = \sum\limits_i x_i |i\rangle$ 和 $|\phi\rangle = \sum\limits_i y_i |i\rangle$ 的内积，定义为

$$
\langle\psi|\phi\rangle \equiv \sum_{i=0}^{d-1} x_i^* y_i
$$

$$= \begin{pmatrix} x_0^* & x_1^* & \cdots & x_{d-1}^* \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{d-1} \end{pmatrix} \tag{0.6}$$

如果 $\langle\psi|\phi\rangle = 0$，那么这两个量子态 $|\psi\rangle, |\phi\rangle$ 被称为相互正交。内积运算具有下述性质：

（1）非负性（positivity）：$\langle\psi|\psi\rangle \geqslant 0, \forall |\psi\rangle \in \mathcal{H}$, 当且仅当 $|\psi\rangle = 0$ 时取等号。

（2）线性（linearity）：$\langle\phi| (a|\psi_1\rangle + b|\psi_2\rangle) = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle, \forall |\phi\rangle, |\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}, a, b \in \mathbb{C}$。

（3）共轭对称性（skew symmetry）：$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*, \forall |\phi\rangle, |\psi\rangle \in \mathcal{H}$。

希尔伯特空间维度 $\dim(\mathcal{H}) = d$，即为对应的线性空间维度。一个 d 维希尔伯特空间通常记为 $\mathcal{H}_d$。在量子信息中，当存在多个量子系统，也就是存在多个希尔伯特空间时，通常用下标来区分这些空间，例如 $\mathcal{H}_S$ 和 $\mathcal{H}_R$。

对于两个希尔伯特空间 $\mathcal{H}_S$ 和 $\mathcal{H}_R$，考虑它们各自的一组正交归一基，$\{|i\rangle_S\}_{i=0}^{m-1}$ 和 $\{|j\rangle_R\}_{j=0}^{n-1}$。两个空间上的线性运算，$\mathcal{H}_R \mapsto \mathcal{H}_S$, 对应的线性算子可以用矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 来表示。需要注意的是，两个空间中的零向量 $|0\rangle_S \in \mathcal{H}_S$ 和 $|0\rangle_R \in \mathcal{H}_R$ 表示不同的列向量。一般情况下，它们的维度可以是不同的。为了简便起见，在不致混淆的情况下将省略空间下标。这样，线性算子可以表示为

$$\boldsymbol{A} = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} a_{ij} |i\rangle\langle j| \tag{0.7}$$

这里矩阵元为 $a_{ij} = {}_S\langle i| A |j\rangle_R$。特别地，当线性算子将一个希尔伯特空间映射到其自身时，矩阵为方阵，$\boldsymbol{A} \in \mathbb{C}^{n\times n}$。我们将这种自映射（self-mapping）线性算子构成的集合记为 $\mathcal{L}(\mathcal{H}_n)$。这里沿用信息论的惯例，元素指标 i, j 从 0 开始，因此将会使用第 0 行和第 0 列的说法。

希尔伯特空间 $\mathcal{H}$ 和作用在其上的算子，包括向量加法、标量数乘及向量内积运算，共同构成了一个复代数（complex algebra）。事实上，对这一代数结构还有更严格的要求：这一代数中的算子应该在范数拓扑（norm topology）和伴随操作（adjoint operation）下封闭。这样，空间 $\mathcal{H}$ 和作用于其上的算子将构成一个 C^* 代数（C-star-algebra）, 感兴趣的读者可以阅读关于代数理论的教材来深入了解这部分内容。

0.2.3 射线

对于线性空间中只相差一个非零常数倍数的向量等价类，我们称这些向量构成一条以零向量为起点的射线（ray）。对于任意一条非零的射线，可以选取该等

价类的一个具有单位向量范数（norm）的代表元 $|\psi\rangle$。我们通常选取向量的欧几里得范数（Euclidean norm）：

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle} \tag{0.8}$$

简单地讲，向量范数表示了向量的长度。对于具有单位长度的代表元：

$$\langle\psi|\psi\rangle = 1 \tag{0.9}$$

0.3 矩阵运算

在量子信息中，常常需要进行矩阵运算。这里，只对本书涉及的关键矩阵运算进行简要回顾。

0.3.1 矩阵基本运算

首先，列出矩阵的若干基本运算。这里，我们经常用矩阵变量对应的小写字母或者大写字母加括号下角标来表示它们的构成元素。考虑矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times l}$ 和 $\boldsymbol{B} \in \mathbb{C}^{l\times m}$,

（1）矩阵乘法（multiplication）: $\boldsymbol{C} = \boldsymbol{AB}$, 矩阵 $\boldsymbol{C}$ 的元素为 $c_{ij} = \sum\limits_k a_{ik}b_{kj}$。

（2）转置（transpose）: $\boldsymbol{A}^{\mathrm{T}}$, 矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的元素为 $(\boldsymbol{A}^{\mathrm{T}})_{ij} = a_{ji}$。

（3）复共轭（complex conjugate）: $\boldsymbol{A}^*$, 矩阵 $\boldsymbol{A}^*$ 的元素为 $(\boldsymbol{A}^*)_{ij} = a^*_{ij}$。

（4）厄米共轭（Hermitian conjugate）: $\boldsymbol{A}^\dagger = (\boldsymbol{A}^{\mathrm{T}})^*$, 矩阵 $\boldsymbol{A}^\dagger$ 的元素为 $(\boldsymbol{A}^\dagger)_{ij} = a^*_{ji}$。$\boldsymbol{A}^\dagger$ 称为 $\boldsymbol{A}$ 的厄米共轭矩阵或者伴随矩阵（adjoint matrix）。如果 $\boldsymbol{A}^\dagger = \boldsymbol{A}$, $\boldsymbol{A}$ 被称为厄米矩阵（Hermitian matrix），或者被称作厄米的。

（5）幺正变换[①]（unitary transformation）: 当 $\boldsymbol{A}$ 为 n 维矩阵，即 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 时，矩阵 $\boldsymbol{A}$ 的幺正变换为 $\boldsymbol{UAU}^\dagger$, 其中 $\boldsymbol{U}$ 是幺正矩阵，满足 $\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{UU}^\dagger = \boldsymbol{I}$。

（6）等距变换（isometric transformation）：当 $\boldsymbol{A}$ 为 n 维方阵时，$\boldsymbol{VAV}^\dagger$, 其中 $\boldsymbol{V} \in \mathbb{C}^{k\times n}$ 被称为等距变换矩阵（isometry）或半幺正矩阵（semi-unitary matrix），满足 $\forall\,|\psi\rangle \in \mathbb{C}^{n\times 1}, \|\,|\psi\rangle\,\| = \|\boldsymbol{V}\,|\psi\rangle\,\|$。

矩阵的一个重要属性是秩（rank）。对于 $m\times n$ 维矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，它的秩定义为该矩阵中线性无关的行或列的数目，并表示为 $\mathrm{rank}(\boldsymbol{A})$。容易看出，在 m 维行或列向量集合中，最多存在 m 个线性无关的行或列向量，因此 $\boldsymbol{A} \leqslant \min\{m, n\}$。另外，任一矩阵的线性无关行数与线性无关列数是相同的。

对于两个方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 一般来说，$\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 不相等。如果 $\boldsymbol{AB} = \boldsymbol{BA}$, 称这两个矩阵对易（commute）。特别地，单位矩阵 $\boldsymbol{I}$ 与所有相应维度的矩阵对

① “幺正”对应实数空间下的正交变换和正交矩阵，有时也被音译为酉变换和酉矩阵。本书中，我们在叙述中有时候会混用变换和矩阵，并用同样的符号来表示一个变换与其对应的矩阵。

易，$\boldsymbol{IA}=\boldsymbol{AI}=\boldsymbol{A}$。事实上，如果一个矩阵和所有矩阵都对易，那么它一定是 $\boldsymbol{I}$ 的数乘，即常数矩阵，见习题 0.3。对于乘法和共轭操作，存在下面的等式关系：

$$\begin{cases}(\boldsymbol{AB})^{*}=\boldsymbol{A}^{*}\boldsymbol{B}^{*}\\(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\(\boldsymbol{AB})^{\dagger}=\boldsymbol{B}^{\dagger}\boldsymbol{A}^{\dagger}\end{cases}\tag{0.10}$$

例 0.1 证明幺正变换不会改变两个矩阵的对易性质。

解 假设矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 对易，即 $\boldsymbol{AB}=\boldsymbol{BA}$，经过幺正变换后，新的矩阵为 $\boldsymbol{UAU}^{\dagger}$ 和 $\boldsymbol{UBU}^{\dagger}$, 有

$$\begin{aligned}&\boldsymbol{UAU}^{\dagger}\boldsymbol{UBU}^{\dagger}\\=&\boldsymbol{UABU}^{\dagger}\\=&\boldsymbol{UBAU}^{\dagger}\\=&\boldsymbol{UBU}^{\dagger}\boldsymbol{UAU}^{\dagger},\end{aligned}\tag{0.11}$$

其中用到了幺正矩阵的性质 $\boldsymbol{U}^{\dagger}\boldsymbol{U}=\boldsymbol{UU}^{\dagger}=\boldsymbol{I}$。因此，有 $\boldsymbol{UAU}^{\dagger}\boldsymbol{UBU}^{\dagger}=\boldsymbol{UBU}^{\dagger}\boldsymbol{UAU}^{\dagger}$，经过幺正变换后的两个矩阵仍对易。同理可证，两个不对易的矩阵在经过幺正变换后仍不对易。

综上，幺正变换不会改变两个矩阵的对易性质。 □

幺正变换可以看作等距变换的一个特例。这可以从下面的一种等距变换等价定义方式看出。

例 0.2 考虑 n 维希尔伯特空间 $\mathcal{H}$ 以及作用于其中向量的线性变换矩阵 $\boldsymbol{V}\in\mathbb{C}^{k\times n},k\geqslant n$。请证明 $\boldsymbol{V}$ 是等距变换矩阵的充要条件是 $\boldsymbol{V}^{\dagger}\boldsymbol{V}=\boldsymbol{I}$，其中 $\boldsymbol{I}$ 是作用于 $\mathcal{H}$ 的 n 维单位矩阵。

证明 必要性（$\boldsymbol{V}$ 是等距变换矩阵 $\Rightarrow \boldsymbol{V}^{\dagger}\boldsymbol{V}=\boldsymbol{I}$）：根据定义，$\boldsymbol{V}$ 是等距变换矩阵，有 $\forall|\psi\rangle\in\mathbb{C}^{n\times 1}$,

$$\begin{aligned}\|\boldsymbol{V}|\psi\rangle\|&=\sqrt{\langle\psi|\boldsymbol{V}^{\dagger}\boldsymbol{V}|\psi\rangle}\\&=\||\psi\rangle\|\\&=\sqrt{\langle\psi|\psi\rangle}\end{aligned}\tag{0.12}$$

因此有 $\boldsymbol{V}^{\dagger}\boldsymbol{V}=\boldsymbol{I}$ 成立。

充分性（$\boldsymbol{V}^{\dagger}\boldsymbol{V}=\boldsymbol{I}\Rightarrow\boldsymbol{V}$ 是等距变换矩阵）：$\forall|\psi\rangle\in\mathbb{C}^{n\times 1}$, 有

$$\begin{aligned}\|\boldsymbol{V}|\psi\rangle\|&=\sqrt{\langle\psi|\boldsymbol{V}^{\dagger}\boldsymbol{V}|\psi\rangle}\\&=\sqrt{\langle\psi|\boldsymbol{I}|\psi\rangle}\\&=\sqrt{\langle\psi|\psi\rangle}\end{aligned}\tag{0.13}$$

按定义，$\boldsymbol{V}$ 是等距变换矩阵。 □

定义 0.1 (正规矩阵，normal matrix) 方阵 $\boldsymbol{A} \in \mathbb{C}^{d\times d}$ 被称作正规的，当且仅当它与自身的厄米共轭矩阵 $\boldsymbol{A}^\dagger$ 对易，$\boldsymbol{A}^\dagger\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\dagger$。

定义 0.2 (方阵在幺正运算下的对角化) 方阵 $\boldsymbol{A} \in \mathbb{C}^{d\times d}$ 在幺正运算下是可对角化的，当且仅当存在一个幺正矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger$ 是对角矩阵。

定理 0.1 (谱分解，spectral decomposition) 方阵 $\boldsymbol{A} \in \mathbb{C}^{d\times d}$ 在幺正运算下是可对角化的，当且仅当其为正规矩阵。

证明 必要性（可对角化 $\Rightarrow$ 正规）：这个方向可以直接予以验证。设存在一个幺正矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger$ 为对角矩阵，那么 $(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)^\dagger = \boldsymbol{U}\boldsymbol{A}^\dagger\boldsymbol{U}^\dagger$ 也是对角的。两个对角矩阵对易，

$$\begin{aligned}(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)^\dagger\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger &= \boldsymbol{U}\boldsymbol{A}^\dagger\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger = \boldsymbol{U}\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{U}^\dagger \\ = \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)^\dagger &= \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{A}^\dagger\boldsymbol{U}^\dagger = \boldsymbol{U}\boldsymbol{A}\boldsymbol{A}^\dagger U^\dagger \end{aligned} \tag{0.14}$$

因此，得到 $\boldsymbol{A}^\dagger\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\dagger$。

充分性（正规 $\Rightarrow$ 可对角化）：这一方向的证明相对复杂。首先，利用 Schur 分解，任一矩阵可以通过幺正变换变成上三角矩阵 $\boldsymbol{T}$，$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^\dagger$。设矩阵 $\boldsymbol{A}$ 正规，利用类似于式(0.14)的做法，我们得到 $\boldsymbol{T}\boldsymbol{T}^\dagger = \boldsymbol{T}^\dagger\boldsymbol{T}$。于是，两个矩阵的对角元是相同的，

$$\begin{aligned}(\boldsymbol{T}\boldsymbol{T}^\dagger)_{ii} &= \sum_j t_{ij}t_{ij}^* = \sum_{j\geqslant i}|t_{ij}|^2 \\ = (\boldsymbol{T}^\dagger\boldsymbol{T})_{ii} &= \sum_k t_{ki}^* t_{ki} = \sum_{k\leqslant i}|t_{ki}|^2 \end{aligned} \tag{0.15}$$

注意到对于第 0 行，有 $\sum\limits_{j\geqslant 0}|t_{0j}|^2 = |t_{00}|^2$, 因此对于 $j \geqslant 1$, $t_{0j} = 0$。通过数学归纳法，可以证明所有的非对角元素都是 0。因此，$\boldsymbol{T}$ 必须是对角矩阵。 □

在量子信息中，特别地，我们对两类正规矩阵感兴趣 —— 厄米矩阵和幺正矩阵。根据谱分解定理，它们都可以通过幺正变换，在合适的基向量下表示为对角矩阵。幺正矩阵的本征值为单位元在复共轭运算下的平方根。而由 $(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)^\dagger = \boldsymbol{U}\boldsymbol{A}^\dagger\boldsymbol{U}^\dagger$ 知, 厄米矩阵的本征值一定为实数。如果一个厄米矩阵所有的本征值均为非负的，我们称它为半正定的（positive semi-definite），记为 $\boldsymbol{A} \geqslant 0$。

例 0.3 一个正规矩阵可以表示成

$$\boldsymbol{A} = \sum_i a_i\,|\phi_i\rangle\langle\phi_i| \tag{0.16}$$

这里求和针对所有 $\boldsymbol{A}$ 的本征值 a_i 和本征向量 $|\phi_i\rangle$。

解 从谱分解定理我们知道，正规矩阵 $\boldsymbol{A}$ 在幺正运算下是可对角化的，即存在一个幺正矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{\Lambda} = \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger$ 为对角矩阵。由对角矩阵性质可知，$\boldsymbol{\Lambda} = \sum_i a_i |i\rangle\langle i|$，其中 $|i\rangle$ 是一组正交归一的基矢，所以有 $\boldsymbol{A} = \boldsymbol{U}^\dagger\boldsymbol{\Lambda}\boldsymbol{U} = \sum_i a_i \boldsymbol{U}^\dagger |i\rangle\langle i| \boldsymbol{U}$。令 $\boldsymbol{U}^\dagger |i\rangle = |\phi_i\rangle$，就可以得到式(0.16)。 □

另外，也可以得到 $\boldsymbol{A}|\phi_i\rangle = \boldsymbol{U}^\dagger\boldsymbol{\Lambda}\boldsymbol{U}\boldsymbol{U}^\dagger |i\rangle = \boldsymbol{U}^\dagger a_i |i\rangle = a_i |\phi_i\rangle$，即 $|\phi_i\rangle$ 是矩阵 $\boldsymbol{A}$ 对应本征值 a_i 的本征向量。

对于式(0.16)，还有一些补充说明：这里的分解可能包含零本征值和对应的本征向量。特别地，对于幺正矩阵，不存在零本征值。有时候为了简单起见，在不引起歧义的情况下，我们也会把本征向量 $|\phi_i\rangle$ 按对应的本征值记为 $|a_i\rangle$ 或者按指标记为 $|i\rangle$。对应不同的本征向量 $|\phi_i\rangle$ 和 $|\phi_j\rangle$ 的本征值 a_i 和 a_j 可能是相同的，我们称这种情况为本征值简并。存在本征值简并的情况下，正规矩阵的展开不唯一。

定义 0.3 (迹，trace) 对于 d 维方阵 $\boldsymbol{A} \in \mathbb{C}^{d\times d}$, 它的迹定义为矩阵所有对角元素之和：

$$\mathrm{tr}(\boldsymbol{A}) = \sum_i a_{ii} \tag{0.17}$$

对任意两个矩阵 $\boldsymbol{A} \in \mathbb{C}^{d\times k}$ 和 $\boldsymbol{B} \in \mathbb{C}^{k\times d}$, 根据式(0.17)和矩阵乘法，有下述等式：

$$\begin{aligned}\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) &= \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}) \\ &= \sum_{i,j} a_{ij} b_{ji}\end{aligned} \tag{0.18}$$

特别地，上述等式对于非对易的矩阵，即 $\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}$ 的情形，依然成立。对于任意幺正矩阵 $\boldsymbol{U}$，$\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{I}$，可以由式(0.18)推导出 $\mathrm{tr}\big(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\big) = \mathrm{tr}\big(\boldsymbol{A}\boldsymbol{U}^\dagger\boldsymbol{U}\big)$。因此，迹运算在幺正变换下保持不变：

$$\mathrm{tr}\big(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\big) = \mathrm{tr}(\boldsymbol{A}) \tag{0.19}$$

对于一组正交归一基 $\{|\psi_i\rangle\}$, 可以将其用狄拉克符号表示：

$$\mathrm{tr}(\boldsymbol{A}) = \sum_i \langle\psi_i| \boldsymbol{A} |\psi_i\rangle \tag{0.20}$$

从式(0.19)可以看出，迹运算与基矢选取无关。如果 $\boldsymbol{A}$ 是厄米的，可以选取其归一化本征向量构成式(0.20)的基。容易看出，$\boldsymbol{A}$ 的迹等于其所有本征值之和：

$$\mathrm{tr}(A) = \sum_i \alpha_i \tag{0.21}$$

其中，α_i 是 $\boldsymbol{A}$ 的本征值。

作为式 (0.18) 的一个特例，有

$$\mathrm{tr}(|\psi\rangle\langle\phi|) = \mathrm{tr}(\langle\phi|\psi\rangle) = \langle\phi|\psi\rangle \tag{0.22}$$

求迹操作的这种轮换不变的性质在后续会多次用到。

0.3.2 直和与张量积

在量子信息中，经常会遇到涉及多系统高维度的运算。在这一节，介绍两种常用的使系统维度扩大的系统之间的运算——直和与张量积。

定义 0.4 (直和，direct sum) 两个矩阵的直和，$\boldsymbol{A}\oplus\boldsymbol{B}$，其中 $\boldsymbol{A}\in\mathbb{C}^{k\times l},\boldsymbol{B}\in\mathbb{C}^{m\times n}$，定义为

$$\begin{aligned}&\begin{pmatrix} a_{0,0} & \cdots & a_{0,l-1} \\ \vdots & \ddots & \vdots \\ a_{k-1,0} & \cdots & a_{k-1,l-1} \end{pmatrix} \oplus \begin{pmatrix} b_{0,0} & \cdots & b_{0,n-1} \\ \vdots & \ddots & \vdots \\ b_{m-1,0} & \cdots & b_{m-1,n-1} \end{pmatrix} \\ =&\begin{pmatrix} a_{0,0} & \cdots & a_{0,l-1} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{k-1,0} & \cdots & a_{k-1,l-1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & b_{0,0} & \cdots & b_{0,n-1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{m-1,0} & \cdots & b_{m-1,n-1} \end{pmatrix}\end{aligned} \tag{0.23}$$

运算结果是分块对角矩阵。

两个矩阵的直和结构是分块对角的，简记为

$$\boldsymbol{A}\oplus\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}_{kn}\\ \boldsymbol{0}_{ml} & \boldsymbol{B}\end{pmatrix} \tag{0.24}$$

其中，$\boldsymbol{0}_{kn}$ 和 $\boldsymbol{0}_{ml}$ 分别是维度为 $k\times n$ 和 $m\times l$ 的零矩阵。矩阵 $\boldsymbol{A}\oplus\boldsymbol{B}$ 的维度为 $(k+m)\times(l+n)$。

思考题 0.1 请证明下述等式：对于任意合适维度的矩阵，以及常数 $x,y\in\mathbb{C}$，

$$(x\boldsymbol{A}_1+y\boldsymbol{A}_2)\oplus(x\boldsymbol{B}_1+y\boldsymbol{B}_2)=x(\boldsymbol{A}_1\oplus\boldsymbol{B}_1)+y(\boldsymbol{A}_2\oplus\boldsymbol{B}_2) \tag{0.25}$$

$$(\boldsymbol{A}_1\oplus\boldsymbol{B}_1)(\boldsymbol{A}_2\oplus\boldsymbol{B}_2)=(\boldsymbol{A}_1\boldsymbol{A}_2)\oplus(\boldsymbol{B}_1\boldsymbol{B}_2) \tag{0.26}$$

$$\mathrm{tr}(\boldsymbol{A}\oplus\boldsymbol{B})=\mathrm{tr}(\boldsymbol{A})+\mathrm{tr}(\boldsymbol{B}) \tag{0.27}$$

在更为严格的表述中，矩阵直和来源于其作用的线性空间的直和。考虑希尔伯特空间 $\mathcal{H}$ 和 $\mathcal{H}'$，如果 $\forall|\phi\rangle\in\mathcal{H}'$，则必然有 $|\phi\rangle\in\mathcal{H}$，我们称 $\mathcal{H}'$ 为 $\mathcal{H}$ 的子空间（subspace），记为 $\mathcal{H}'\subseteq\mathcal{H}$。设 $\mathcal{H}_n$ 和 $\mathcal{H}_m$ 是 $\mathcal{H}$ 的两个子空间，并满足 $\mathcal{H}_n\cap\mathcal{H}_m=\{0\}$，那么 $\mathcal{H}_n$ 和 $\mathcal{H}_m$ 的直和定义为

$$\mathcal{H}_n\oplus\mathcal{H}_m=\big\{|\varphi\rangle=|\psi\rangle+|\phi\rangle\,\big|\,|\psi\rangle\in\mathcal{H}_n,|\phi\rangle\in\mathcal{H}_m\big\} \tag{0.28}$$

直和空间的维度为

$$\dim(\mathcal{H}_n \oplus \mathcal{H}_m) = n + m \tag{0.29}$$

定义 0.5 (张量积，tensor product 或者 Kronecker product)　两个矩阵的张量积定义为 $\boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B}$：

$$\begin{aligned}\boldsymbol{C} &= \begin{pmatrix} a_{0,0} & \cdots & a_{0,l-1} \\ \vdots & \ddots & \vdots \\ a_{k-1,0} & \cdots & a_{k-1,l-1} \end{pmatrix} \otimes \begin{pmatrix} b_{0,0} & \cdots & b_{0,n-1} \\ \vdots & \ddots & \vdots \\ b_{m-1,0} & \cdots & b_{m-1,n-1} \end{pmatrix} \\ &= \begin{pmatrix} a_{0,0}\boldsymbol{B} & \cdots & a_{0,l-1}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{k-1,0}\boldsymbol{B} & \cdots & a_{k-1,l-1}\boldsymbol{B} \end{pmatrix}\end{aligned} \tag{0.30}$$

其中，$a_{ij}\boldsymbol{B}$ 是对矩阵 $\boldsymbol{B}$ 的数乘运算。

对于矩阵 $\boldsymbol{A}\otimes\boldsymbol{B}$，其维度为 $km \times ln$。可以将其视作有四个指标 i_A, i_B, j_A, j_B 的张量（这也是称该运算为张量积的原因）：

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{i_A i_B}_{j_A j_B} = (\boldsymbol{A})^{i_A}_{j_A} (\boldsymbol{B})^{i_B}_{j_B} \tag{0.31}$$

在这四个指标中，两个为行指标，标记为张量上标；两个为列指标，标记为下标。我们将在 0.3.4节中简单介绍张量的一种图像表示。

思考题 0.2　请证明下述结果的正确性：对于任意合适维度的矩阵，常数 $x, y \in \mathbb{C}$, 以及 d 维单位矩阵 $\boldsymbol{I}_d \in \mathbb{C}^{d\times d}$,

（1）两个矩阵张量积的迹是矩阵迹的乘积：

$$\mathrm{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A})\,\mathrm{tr}(\boldsymbol{B}) \tag{0.32}$$

（2）张量运算 $\otimes$ 与厄米共轭运算 $\dagger$ 对易：

$$(\boldsymbol{A} \otimes \boldsymbol{B})^\dagger = \boldsymbol{A}^\dagger \otimes \boldsymbol{B}^\dagger \tag{0.33}$$

（3）如果对同一矩阵进行若干次直和或张量积运算，可以简记为

$$\begin{cases} \boldsymbol{A}^{\oplus n} \equiv \underbrace{\boldsymbol{A} \oplus \boldsymbol{A} \oplus \cdots \oplus \boldsymbol{A}}_{n} \\ \boldsymbol{A}^{\otimes n} \equiv \underbrace{\boldsymbol{A} \otimes \boldsymbol{A} \otimes \cdots \otimes \boldsymbol{A}}_{n} \end{cases} \tag{0.34}$$

那么，我们可以将张量积运算 $\otimes$ 和直和运算 $\oplus$ 联系起来：

$$\boldsymbol{I}_n \otimes \boldsymbol{A} = \boldsymbol{A}^{\oplus n} \tag{0.35}$$

因此，在文献中鲜少出现 $\boldsymbol{A}^{\oplus n}$。

（4）张量积运算是双线性的（double linear）：

$$\begin{cases}(x\boldsymbol{A}_1+y\boldsymbol{A}_2)\otimes\boldsymbol{B}=x(\boldsymbol{A}_1\otimes\boldsymbol{B})+y(\boldsymbol{A}_2\otimes\boldsymbol{B})\\ \boldsymbol{A}\otimes(x\boldsymbol{B}_1+y\boldsymbol{B}_2)=x(\boldsymbol{A}\otimes\boldsymbol{B}_1)+y(\boldsymbol{A}\otimes\boldsymbol{B}_2)\end{cases}\tag{0.36}$$

与此相对的，直和运算不满足该性质，见式(0.25)。

作为张量积运算的特例，对于两个向量的张量积，通常简记为

$$|\phi\rangle\otimes|\psi\rangle\equiv|\phi\rangle|\psi\rangle\equiv|\phi\psi\rangle\tag{0.37}$$

其中，向量 $|\phi\rangle$ 和 $|\psi\rangle$ 一般维度不相同。设 $\{|i\rangle_S\}$ 和 $\{|j\rangle_R\}$ 分别是空间 $\mathcal{H}_S$ 和 $\mathcal{H}_R$ 的一组基，那么，这些基向量的张量积的内积运算为

$$(\langle i|_S\langle j|_R)(|i'\rangle_S|j'\rangle_R)=\delta_{ii'}\delta_{jj'}\tag{0.38}$$

其中，$\delta_{ii'}$ 和 $\delta_{jj'}$ 为 Kronecker 函数（Kronecker delta functions）。

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}\tag{0.39}$$

因此，量子态集合 $\{|i\rangle_S|j\rangle_R\}$ 构成了一个更大的希尔伯特空间的基向量，对应的空间为 $\mathcal{H}_S\otimes\mathcal{H}_R$, 其维度为

$$\dim(\mathcal{H}_S\otimes\mathcal{H}_R)=\dim(\mathcal{H}_S)\dim(\mathcal{H}_R)\tag{0.40}$$

在量子信息中，我们称 S 和 R 为联合系统 SR 的子系统（subsystem）。

0.3.3 偏迹

张量积运算是一种快速扩展系统维度的运算，迹运算则将矩阵缩并为一个标量。更进一步地，我们对两种运算的联系进行探讨。对于两个方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，有 $\mathrm{tr}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{tr}(\boldsymbol{A})\,\mathrm{tr}(\boldsymbol{B})$。对于联合系统的迹运算，可以将其视作先对第一个子系统求迹，之后再对第二个系统求迹，或反过来：

$$\mathrm{tr}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{tr}[\mathrm{tr}(\boldsymbol{A})\boldsymbol{B}]=\mathrm{tr}[\mathrm{tr}(\boldsymbol{B})\boldsymbol{A}]\tag{0.41}$$

矩阵 $\mathrm{tr}(\boldsymbol{A})\boldsymbol{B}$ 和 $\mathrm{tr}(\boldsymbol{B})\boldsymbol{A}$ 可以视为矩阵 $\boldsymbol{A}\otimes\boldsymbol{B}$ “部分的”迹。这种形式的矩阵在量子信息中有清晰的物理含义，并且具有广泛的应用。

定义 0.6 (偏迹，partial trace)　对于线性算子 $\boldsymbol{T}\in\mathcal{L}(\mathcal{H}_S\otimes\mathcal{H}_R)$, 定义关于系统 S 的偏迹，$\mathrm{tr}_S(\boldsymbol{T})$, 以及关于系统 R 的偏迹，$\mathrm{tr}_R(\boldsymbol{T})$, 为由下面元素构成的

矩阵：

$$\begin{cases} (\mathrm{tr}_S(\boldsymbol{T}))_{ij} = \sum\limits_{k=0}^{\dim(\mathcal{H}_S)-1} (T)_{kj}^{ki} \\ (\mathrm{tr}_R(\boldsymbol{T}))_{ij} = \sum\limits_{k=0}^{\dim(\mathcal{H}_R)-1} (T)_{jk}^{ik} \end{cases} \tag{0.42}$$

现在，我们在不引入过多张量表示的情况下，对偏迹运算的含义进行理解。首先，偏迹运算得到的矩阵

$$\mathrm{tr}_S(\boldsymbol{T}) \in \mathcal{L}(\mathcal{H})_R, \quad \mathrm{tr}_R(\boldsymbol{T}) \in \mathcal{L}(\mathcal{H})_S \tag{0.43}$$

按定义分别是对于子系统 S 和 R 求迹的结果。考虑一个简单的例子。设 $\boldsymbol{T} = \boldsymbol{A} \otimes \boldsymbol{B}$, 其中 $\boldsymbol{A} \in \mathcal{L}(\mathcal{H})_S$, $\boldsymbol{B} \in \mathcal{L}(\mathcal{H})_R$, 如式(0.41)所示。那么，

$$\mathrm{tr}_S(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A})\boldsymbol{B}, \quad \mathrm{tr}_R(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B})A \tag{0.44}$$

因此，偏迹运算可以看作张量积 $\otimes$ 的某种逆运算。尽管一般来说，对于张量希尔伯特空间上的方阵 $\boldsymbol{T} \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_R)$, 不一定能写成两个矩阵的张量形式：

$$\boldsymbol{T} \neq \boldsymbol{A} \otimes \boldsymbol{B} \tag{0.45}$$

但张量积的“逆运算”这一直观理解依然适用。

例 0.4 考虑一个四维希尔伯特空间，$\mathcal{H}_4 = \mathcal{H}_S \otimes \mathcal{H}_R$, 其中，$S$ 和 R 均为二维系统，

$$\boldsymbol{T} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \tag{0.46}$$

请计算矩阵 $\boldsymbol{T}$ 分别对 S 和 R 求偏迹的结果。

解 首先，将 $\boldsymbol{T}$ 表示为矩阵形式：

$$\boldsymbol{T} = \begin{pmatrix} (T)_{00}^{00} & (T)_{01}^{00} & (T)_{10}^{00} & (T)_{11}^{00} \\ (T)_{00}^{01} & (T)_{01}^{01} & (T)_{10}^{01} & (T)_{11}^{01} \\ (T)_{00}^{10} & (T)_{01}^{10} & (T)_{10}^{10} & (T)_{11}^{10} \\ (T)_{00}^{11} & (T)_{01}^{11} & (T)_{10}^{11} & (T)_{11}^{11} \end{pmatrix} \tag{0.47}$$

这样，可以按照式(0.42)对矩阵求偏迹：

$$\begin{cases} \mathrm{tr}_S(\boldsymbol{T}) = \begin{pmatrix} \sum\limits_k (T)^{k0}_{k0} & \sum\limits_k (T)^{k0}_{k1} \\ \sum\limits_k (T)^{k1}_{k0} & \sum\limits_k (T)^{k1}_{k1} \end{pmatrix} = \begin{pmatrix} a_{11}+a_{33} & a_{12}+a_{34} \\ a_{21}+a_{43} & a_{22}+a_{44} \end{pmatrix} \\ \mathrm{tr}_R(\boldsymbol{T}) = \begin{pmatrix} \sum\limits_k (T)^{0k}_{0k} & \sum\limits_k (T)^{0k}_{1k} \\ \sum\limits_k (T)^{1k}_{0k} & \sum\limits_k (T)^{1k}_{1k} \end{pmatrix} = \begin{pmatrix} a_{11}+a_{22} & a_{13}+a_{24} \\ a_{31}+a_{42} & a_{33}+a_{44} \end{pmatrix} \end{cases} \tag{0.48}$$

这里，容易看出 $\mathrm{tr}[\mathrm{tr}_S(\boldsymbol{T})] = \mathrm{tr}[\mathrm{tr}_R(\boldsymbol{T})] = \mathrm{tr}(\boldsymbol{T})$。 □

通过这个例子可以看到，多个系统的算子运算通常比较繁琐。幸运的是，可以利用狄拉克符号来简化书写。利用左/右矢符号，通过引入系统 R 的一组完备基 $\{|i\rangle\}$, 对于系统 R 求偏迹的结果可以表示为

$$\mathrm{tr}_R(\boldsymbol{T}_{SR}) = \sum_i (\boldsymbol{I}_S \otimes \langle i|_R)\boldsymbol{T}_{SR}(\boldsymbol{I}_S \otimes |i\rangle_R) \equiv \sum_i \langle i|\,\boldsymbol{T}_{SR}\,|i\rangle_R \tag{0.49}$$

严格地讲，按照矩阵乘法定义，记号 $\langle i|\,\boldsymbol{T}_{SR}\,|i\rangle$ 不是一个良定义的表示。这里请注意矩阵和向量的维度。但在书写中，我们经常用这种省略对另一系统进行恒等操作的表示方法来简化书写。关于这一表示的含义和证明，我们留作练习（习题 0.4）。

*0.3.4 张量网络

在这一节，介绍一种新的数学工具——张量网络（tensor network），这一工具对于张量计算很有帮助。这里，将采用张量指标的书写方法，上指标代表行，下指标代表列。将一个张量表示为一个有打开的“腿”的框图，这些腿表示张量的指标（index）。例如，一个三体的量子态 ρ_{ABC} 表示为一个有六条腿的框图，其中三条在左侧，表示行指标 i_A, i_B, i_C，即狄拉克记号中的 “$|\cdot\rangle$”，另外三条在右侧，表示列指标 j_A, j_B, j_C，即狄拉克记号中的 “$\langle\cdot|$”，如图 0.1 所示。

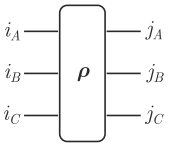

图 0.1 张量网络表示：$\rho_{ABC} = \sum\limits_{i_A,i_B,i_C,j_A,j_B,j_C} \rho^{i_A i_B i_C}_{j_A j_B j_C}\, |i_A i_B i_C\rangle\langle j_A j_B j_C|$

指标缩并是一个基本的张量运算，在图 0.1 中，我们将要缩并的两个指标对应的腿连接起来。利用这一记号，给出一些常用的矩阵运算的张量网络表示。矩阵乘积 $(\boldsymbol{\rho\sigma})^i_j = \sum\limits_k (\boldsymbol{\rho})^i_k (\boldsymbol{\sigma})^k_j$ 可以图像化地表示为 $\boldsymbol{\rho}$ 所对应的框图右侧的腿与 $\boldsymbol{\sigma}$ 左侧的腿相连接；矩阵的迹 $\mathrm{tr}(\boldsymbol{\rho}) = \sum\limits_i (\rho)^i_i$ 是 $\boldsymbol{\rho}$ 张量行和列指标的缩并，可

以用它的张量框图左侧与右侧的腿相连接来表示。偏迹操作就是把相应的子系统对应的指标缩并了，$(\mathrm{tr}_A(\rho_{AB}))^{i_B}_{j_B} = \sum_{i_A}(\rho_{AB})^{i_A,i_B}_{i_A,j_B}$，其他子系统的指标保持不变。此外，还有些运算并不涉及指标缩并，比如，矩阵的转置 $(\boldsymbol{\rho}^{\mathrm{T}})^i_j = (\boldsymbol{\rho})^j_i$ 可以表示为左右两边的指标互换；张量积，或者称 Kronecker 积（Kronecker product），$(\boldsymbol{\rho}\otimes\boldsymbol{\sigma})^{i_A,i_B}_{j_A,j_B} = (\boldsymbol{\rho})^{i_A}_{j_A}(\boldsymbol{\sigma})^{i_B}_{j_B}$ 的张量网络表示就是将 $\boldsymbol{\rho},\boldsymbol{\sigma}$ 的框图排列放在一起，如图 0.2 所示。

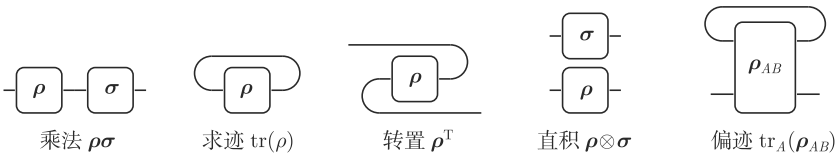

图 0.2　用张量网络表示一些矩阵运算

除了矩阵运算，也可以利用张量网络表示一些常用的量子态和算子。最简单的就是单位矩阵 $\boldsymbol{I}$，就是一条横线。再比如，（未归一化的）最大纠缠态（unnormalised maximally entangled state，UMES），对于 d 维希尔伯特空间中的 UMES 算子，

$$\sqrt{d}\left|\Phi_d^+\right\rangle = \sum_i |ii\rangle \tag{0.50}$$

它的张量网络表示为一个半圆，半圆圆弧的两端都朝左。我们也可以表示这个态的密度矩阵：

$$d\Phi_d^+ = \sum_{i,j} |ii\rangle\langle jj| = \left(\sum_i |ii\rangle\right)\left(\sum_j \langle jj|\right) \tag{0.51}$$

那就是由一个朝左和一个朝右的半圆组成。交换算子（swap operator，SWAP）是一个常用的算子，它会将作用的两个系统 A,B 的角标互换：

$$\begin{aligned}\mathrm{SWAP}(\rho_{AB}) &= \sum \rho^{i_A,i_B}_{j_A,j_B}\,|i_Bi_A\rangle\langle j_Bj_A| \\ &= \sum \rho^{i_B,i_A}_{j_B,j_A}\,|i_Ai_B\rangle\langle j_Aj_B|\end{aligned} \tag{0.52}$$

这里，第二个等式将指标重新定义排列了。由此，我们也可以直接写出 SWAP 的幺正变换表示：

$$S = \sum |i_Bi_A\rangle\langle i_Ai_B| \tag{0.53}$$

这些态和操作的图像化表示如图 0.3 所示。

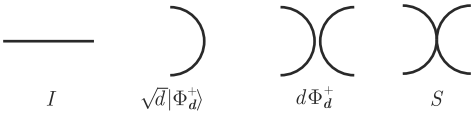

图 0.3　用张量网络表示一些量子态和操作，这里注意 $d\Phi_d^+$ 和 S 的区别，见式 (0.51) 和式 (0.53)

对比图 0.2 和图 0.3 中的图像，我们发现 UMES 这样的半圆也在求迹和转置中出现了，这是有意为之。事实上，考虑一个矩阵，$\boldsymbol{\rho}=\sum\limits_{i,j}\rho_j^i\,|i\rangle\langle j|$, 它的转置和迹可以分别表示为

$$\begin{cases} d(\langle\Psi^+|\otimes\boldsymbol{I})(\boldsymbol{I}\otimes\boldsymbol{\rho}\otimes\boldsymbol{I})(\boldsymbol{I}\otimes|\Psi^+\rangle)=\sum\limits_{i,j}\rho_j^i\sum\limits_{k,l}\langle k|i\rangle\langle j|l\rangle\,|l\rangle\langle k| \\ \quad=\sum\limits_{i,j}\rho_j^i\,|j\rangle\langle i|=\boldsymbol{\rho}^{\mathrm{T}} \\ d\langle\Psi^+|\boldsymbol{\rho}\otimes\boldsymbol{I}|\Psi^+\rangle=\sum\limits_{k}\langle kk|\sum\limits_{i,j}\rho_j^i\,|i\rangle\langle j|\otimes\boldsymbol{I}\sum\limits_{l}|ll\rangle \\ \quad=\sum\limits_{i,j}\rho_j^i\sum\limits_{k,l}\langle k|i\rangle\langle j|l\rangle\langle k|l\rangle=\sum\limits_{i}\rho_i^i=\mathrm{tr}(\boldsymbol{\rho}) \end{cases} \tag{0.54}$$

读者可以自行将两个图拼凑一下，可以很快得到上述等式。这是用作图来表示张量的优势。下面这个例子，可以用作图法很快证明。

例 0.5 矩阵乘法可以由 SWAP 操作完成：

$$\mathrm{tr}[S(\boldsymbol{\rho}\otimes\boldsymbol{\sigma})]=\mathrm{tr}(\boldsymbol{\rho}\boldsymbol{\sigma}) \tag{0.55}$$

解 直接作张量图来证明，如图 0.4 所示。

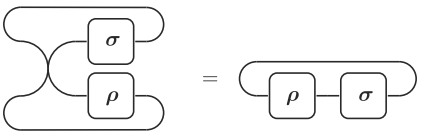

图 0.4 张量网络图证明：$\mathrm{tr}[S(\boldsymbol{\rho}\otimes\boldsymbol{\sigma})]=\mathrm{tr}(\boldsymbol{\rho}\boldsymbol{\sigma})$

注意，图 0.4 中的线可以拉直，并不影响最后结果。 □

在第 2 章中，会介绍两种量子信道的表示方法：Kraus 算子表示和蔡矩阵表示。利用张量网络，可以将两种表示方法图示化，如图 0.5 所示。

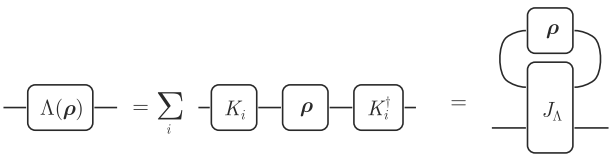

图 0.5 量子信道的两种表示方式, $\Lambda(\rho)=\sum\limits_i K_i\boldsymbol{\rho}K_i^\dagger=\mathrm{tr}_R[J_\Lambda(\boldsymbol{\rho}_R^{\mathrm{T}}\otimes\boldsymbol{I})]$

0.3.5 方阵的解析函数

矩阵乘法可以视作关于一个矩阵的简单函数。在这一节，我们定义基于乘法运算的更一般矩阵函数。首先，定义一个复矩阵的幂，$\boldsymbol{A}\in\mathbb{C}^{d\times d}$。幂次运算是同

一矩阵相乘多次：

$$\boldsymbol{A}^k = \underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}_{k\text{ 次}} \tag{0.56}$$

幂函数作为对矩阵的操作，与任一幺正变换对易：

$$\begin{aligned}\boldsymbol{U}\boldsymbol{A}^k\boldsymbol{U}^\dagger &= \boldsymbol{U}\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}\boldsymbol{U}^\dagger \\ &= \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\cdots\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger \\ &= (\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)^k\end{aligned} \tag{0.57}$$

这里注意到按照定义，$\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}$。这样，可以将标量函数的泰勒展开（Taylor expansion）推广至矩阵函数，并定义一个方阵的解析函数：

$$f(\boldsymbol{A}) \equiv \sum_k \frac{f^{(k)}(0)}{k!}\boldsymbol{A}^k \tag{0.58}$$

其中，$f^{(k)}(0)$ 是函数 $f(x)$ 在 $x=0$ 处的第 k 阶导数。如果 $f^{(k)}(0)$ 发散，例如对数函数 log 或者平方根函数，那么，可以通过对矩阵函数添加恒等矩阵的方式，在定义域上的其他点对函数进行展开。对于连续性很差的函数，这里就不展开讨论。

思考题 0.3 *方阵 $\boldsymbol{A}\in\mathbb{C}^{d\times d}$ 与自身的矩阵函数对易，即 $\boldsymbol{A}f(\boldsymbol{A}) = f(\boldsymbol{A})\boldsymbol{A}$。*

一般来说，式(0.58)的计算非常繁琐。幸运的是，对于正规矩阵 $\boldsymbol{A}$，有更简便的方法计算矩阵函数 $f(\boldsymbol{A})$。如式(0.57)所示，幂函数与任一幺正变换对易，因此在计算时，可以对矩阵乘法的顺序进行对换。这样，可以选取一个可以将矩阵 $\boldsymbol{A}$ 对角化的幺正矩阵 $\boldsymbol{U}$：

$$\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger = \begin{pmatrix}\alpha_0 & & & \\ & \alpha_1 & & \\ & & \ddots & \\ & & & \alpha_{d-1}\end{pmatrix} \tag{0.59}$$

其中，$\alpha_0,\alpha_1,\cdots,\alpha_{d-1}$ 是矩阵 $\boldsymbol{A}$ 的本征值，一般可以是复数。这样，可以将函数 $f(\cdot)$ 看作对 $\boldsymbol{A}$ 的本征值分别进行运算，随后再对这一变换后的对角矩阵通过 $\boldsymbol{U}^\dagger$ 旋转至待求解的矩阵：

$$\begin{aligned}f(\boldsymbol{A}) &= f(\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger\boldsymbol{U}) \\ &= \boldsymbol{U}^\dagger f(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\dagger)\boldsymbol{U}\end{aligned}$$

$$= \boldsymbol{U}^{\dagger} \begin{pmatrix} f(\alpha_0) & & & \\ & f(\alpha_1) & & \\ & & \ddots & \\ & & & f(\alpha_{d-1}) \end{pmatrix} \boldsymbol{U} \tag{0.60}$$

换句话说，对于可以在幺正变换下对角的矩阵，矩阵函数的求解问题本质上转化为对其本征值的求解。

经常地，我们用狄拉克符号将矩阵用其归一化本征向量 $|\psi_k\rangle$ 和相应的本征值 α_k 进行表示：

$$\boldsymbol{A} = \sum_{k=0}^{d-1} \alpha_k |\psi_k\rangle\langle\psi_k| \tag{0.61}$$

根据式(0.60), 有

$$f(\boldsymbol{A}) = \sum_{k=0}^{d-1} f(\alpha_k) |\psi_k\rangle\langle\psi_k| \tag{0.62}$$

这里，我们利用了正交归一投影算子（projector）的幂次依然是其自身的结论，即 $\forall k, l \in [d]$,

$$\begin{cases} (|\psi_k\rangle\langle\psi_k|)(|\psi_l\rangle\langle\psi_l|) = \delta_{kl} |\psi_k\rangle\langle\psi_k| \\ (|\psi_k\rangle\langle\psi_k|)^n = |\psi_k\rangle\langle\psi_k| \end{cases} \tag{0.63}$$

对任意正整数 n 成立，其中 δ_{kl} 是 Kronecker 函数。对于正规矩阵 $\boldsymbol{A}$，我们也将式(0.62)看作矩阵函数 $f(\boldsymbol{A})$ 的定义。

例 0.6 (矩阵直和的函数)　对于两个可对角化的矩阵，$\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 和 $\boldsymbol{B} \in \mathbb{C}^{m\times m}$, $f(\boldsymbol{A} \oplus \boldsymbol{B}) = f(\boldsymbol{A}) \oplus f(\boldsymbol{B})$。

解　将矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的本征值分别记为 α_i 和 β_j, 其中，$i \in [n]$, $j \in [m]$。可以用矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的归一化本征向量对矩阵直和进行表示：

$$\begin{aligned} \boldsymbol{A} \oplus \boldsymbol{B} &= \left(\sum_{i=0}^{n-1} \alpha_i |\psi_i\rangle\langle\psi_i| \right) \oplus \left(\sum_{j=0}^{m-1} \beta_j |\phi_j\rangle\langle\phi_j| \right) \\ &= \sum_i \alpha_i \left|\tilde{\psi}_i\right\rangle\left\langle\tilde{\psi}_i\right| + \sum_j \beta_j \left|\tilde{\phi}_j\right\rangle\left\langle\tilde{\phi}_j\right| \end{aligned} \tag{0.64}$$

其中，$\left|\tilde{\psi}_i\right\rangle \in \mathbb{C}^{n+m}$ 是由在 $|\psi_i\rangle$ 的表示后加上 m 个 0 得到的，而 $\left|\tilde{\phi}_j\right\rangle \in \mathbb{C}^{n+m}$ 是由在 $|\phi_j\rangle$ 的表示前面加上 n 个 0 得到的。这么做是为了能够在更大的空间里表示原本的 $|\psi_i\rangle$ 与 $|\phi_j\rangle$ 这两个小空间的态。这样，向量 $\left\{\left|\tilde{\psi}_i\right\rangle, \left|\tilde{\phi}_j\right\rangle\right\}$ 构成一组正交归一基，使得直和矩阵 $\boldsymbol{A} \oplus \boldsymbol{B}$ 在这组基上是对角化的。因此，对于矩阵直和

$\boldsymbol{A}\oplus\boldsymbol{B}$ 的函数，有下式给出：

$$\begin{aligned} f(\boldsymbol{A}\oplus\boldsymbol{B}) &= \sum_i f(\alpha_i)\left|\tilde{\psi}_i\right\rangle\!\left\langle\tilde{\psi}_i\right| + \sum_j f(\beta_j)\left|\tilde{\phi}_j\right\rangle\!\left\langle\tilde{\phi}_j\right| \\ &= \left(\sum_{i=0}^{n-1} f(\alpha_i)\left|\psi_i\right\rangle\!\left\langle\psi_i\right|\right) \oplus \left(\sum_{j=0}^{m-1} f(\beta_j)\left|\phi_j\right\rangle\!\left\langle\phi_j\right|\right) \\ &= f(\boldsymbol{A})\oplus f(\boldsymbol{B}) \end{aligned} \tag{0.65}$$

由此可以看出，在对正规矩阵直和结果进行函数运算时，等同于分别对两个矩阵进行该函数运算再直和。 □

例 0.7 对于两个可对角化矩阵 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{d\times d}$, $\mathrm{tr}(|\boldsymbol{A}||\boldsymbol{B}|)\leqslant \mathrm{tr}\,|\boldsymbol{A}|\,\mathrm{tr}\,|\boldsymbol{B}|$。

解

$$\begin{aligned} \mathrm{tr}(|\boldsymbol{A}||\boldsymbol{B}|) &= \mathrm{tr}\left(\sum_i |\alpha_i|\,|\psi_i\rangle\!\langle\psi_i| \sum_j |\beta_j|\,|\phi_j\rangle\!\langle\phi_j|\right) \\ &= \mathrm{tr}\left(\sum_{i,j} |\alpha_i\beta_j|\,|\psi_i\rangle\,\langle\psi_k|\phi_j\rangle\,\langle\phi_j|\right) \\ &= \sum_{i,j,k} |\alpha_i\beta_j|\,\langle\psi_k|\psi_i\rangle\,\langle\psi_i|\phi_j\rangle\,\langle\phi_j|\psi_k\rangle \\ &= \sum_{i,j} |\alpha_i\beta_j|\,|\langle\psi_i|\phi_j\rangle|^2 \\ &\leqslant \sum_{i,j} |\alpha_i\beta_j| \\ &= \mathrm{tr}\,|\boldsymbol{A}|\,\mathrm{tr}\,|\boldsymbol{B}| \end{aligned} \tag{0.66}$$

其中，在第三个等式的推导中，使用了求迹操作的轮换性质及求迹操作与求和操作可以交换顺序的性质。 □

*0.3.6 一般矩阵的解析函数

除方阵外，对于一般的矩阵，也可以定义它们的矩阵函数。基于 0.3.5 节中关于方阵的结果，可以首先将一般的矩阵转化为方阵，再进行函数定义。一种转化方式是利用矩阵的奇异值。

定义 0.7 (奇异值，singular value) 矩阵的奇异值，$\boldsymbol{A}\in\mathbb{C}^{n\times k}$, 定义为下述矩阵的本征值：

$$|\boldsymbol{A}| = \sqrt{\boldsymbol{A}^\dagger\boldsymbol{A}} \tag{0.67}$$

其中，矩阵 $\boldsymbol{A}^\dagger\boldsymbol{A}$ 是方阵且厄米。

可以看到，矩阵的奇异值是一个半正定矩阵的本征值，因此我们也称奇异值为矩阵的绝对值（absolute value）。对于在幺正变换下可以对角化的矩阵，矩阵绝对值就是其本征值的绝对值，可以由式(0.62)给出。作为直接的推论，对于任一幺正矩阵 $\boldsymbol{U}$,

$$|\boldsymbol{U}| = \sqrt{\boldsymbol{U}^\dagger\boldsymbol{U}} = \boldsymbol{I} \tag{0.68}$$

思考题 0.4 请证明下述结论：

（1）$\boldsymbol{A}\boldsymbol{A}^\dagger$ 和 $\boldsymbol{A}^\dagger\boldsymbol{A}$ 均为半正定矩阵，且 $\mathrm{tr}\big(\boldsymbol{A}^\dagger\boldsymbol{A}\big) = \mathrm{tr}\big(\boldsymbol{A}\boldsymbol{A}^\dagger\big)$。

（2）$|\mathrm{tr}(\boldsymbol{A})| \leqslant \mathrm{tr}(|\boldsymbol{A}|)$, 等号仅在 $\boldsymbol{A} \geqslant 0$ 或 $-\boldsymbol{A} \geqslant 0$ 时取到。

（3）一般地，$|\boldsymbol{A}| \neq |\boldsymbol{A}^\dagger|$, 但 $\mathrm{tr}(|\boldsymbol{A}|) = \mathrm{tr}\big(|\boldsymbol{A}^\dagger|\big)$。

类似于对向量的分析，我们可以定义矩阵的范数（norm），并利用泛函分析中的结果对矩阵进行研究。基于矩阵绝对值，我们可以定义一类矩阵范数。在量子信息中，几种常用的范数包括 Schatten 范数、l_p-范数、迹范数（trace norm），以及樊畿范数[1]。在本书中，主要用到 Schatten 范数。在量子信息研究中，人们经常简称其为 p-范数。需要注意的是，在不同的学科中，类似的名称可能指代不同的范数。由于定义上的相似性，我们在这里也简单介绍矩阵的 l_p-范数。

定义 0.8 (l_p-范数) 给定 $p \geqslant 1$, 对于复矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times k}$, 它的 l_p-范数为

$$\|\boldsymbol{A}\|_{l_p} = \left(\sum_{i,j} |A_{ij}|^p\right)^{1/p} \tag{0.69}$$

定义 0.9 (Schatten p-范数) 给定 $p \geqslant 1$, 对于复矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times k}$, 它的 Schatten p-范数为

$$\|\boldsymbol{A}\|_p = [\mathrm{tr}(|\boldsymbol{A}|^p)]^{1/p} \tag{0.70}$$

其中，$|\boldsymbol{A}|$ 由式(0.67)给出。

p-范数有许多良好的数学性质，包括次可乘性（sub-multiplicativity，向量范数自然拥有这一性质，但矩阵范数不一定有这一性质）和单调性（monotonicity），分别由下面两式表示：

$$\|\boldsymbol{A}\boldsymbol{B}\|_p \leqslant \|\boldsymbol{A}\|_p\|\boldsymbol{B}\|_p \tag{0.71}$$

$$\|\boldsymbol{A}\|_1 \geqslant \|\boldsymbol{A}\|_p \geqslant \|\boldsymbol{A}\|_q \geqslant \|\boldsymbol{A}\|_\infty \tag{0.72}$$

其中，$\boldsymbol{A} \in \mathbb{C}^{n\times k}$, $\boldsymbol{B} \in \mathbb{C}^{k\times m}$，$q \geqslant p \geqslant 1$。我们将证明留作习题。

思考题 0.5 (Hilbert-Schmidt 内积运算下的 Cauchy–Schwarz 不等式) 对于两个复矩阵 $\boldsymbol{A} \in \mathbb{C}^{k\times n}$ 和 $\boldsymbol{B} \in \mathbb{C}^{k\times m}$,

$$\mathrm{tr}\big(\boldsymbol{A}^\dagger\boldsymbol{B}\big) \leqslant \sqrt{\mathrm{tr}(\boldsymbol{A}^\dagger\boldsymbol{A})\,\mathrm{tr}(\boldsymbol{B}^\dagger\boldsymbol{B})} = \|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_2 \tag{0.73}$$

例 0.8 (* 矩阵在 p-范数下的 Hölder 不等式) 对于复矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times d}$, $\boldsymbol{B} \in \mathbb{C}^{m\times d}$, 以及满足关系 $p^{-1}+q^{-1}=1$ 的正实数 p,q,

$$\left\|\boldsymbol{A}\boldsymbol{B}^{\dagger}\right\|_1 \leqslant \|\boldsymbol{A}\|_p \|\boldsymbol{B}\|_q \tag{0.74}$$

解 记矩阵 $\boldsymbol{A}^{\dagger}\boldsymbol{A}$ 和 $\boldsymbol{B}^{\dagger}\boldsymbol{B}$ 的本征值和对应的归一化本征向量分别为 $\{\alpha_i^2, |\psi_i\rangle\}$ 和 $\{\beta_j^2, |\phi_j\rangle\}$。根据式(0.67), 有

$$\begin{cases} |\boldsymbol{A}| = \sum\limits_i |\alpha_i|\, |\psi_i\rangle\langle\psi_i| \\ |\boldsymbol{B}| = \sum\limits_j |\beta_j|\, |\phi_j\rangle\langle\phi_j| \end{cases} \tag{0.75}$$

利用奇异值分解（singular value decomposition），存在线性等距变换算子 $\boldsymbol{U} \in \mathbb{C}^{n\times d}$, $\boldsymbol{V} \in \mathbb{C}^{m\times d}$, 满足 $\boldsymbol{U}^{\dagger}\boldsymbol{U} = \boldsymbol{V}^{\dagger}\boldsymbol{V} = \boldsymbol{I}$, 分别使得 $\boldsymbol{A} = \boldsymbol{U}|\boldsymbol{A}|$, $\boldsymbol{B} = \boldsymbol{V}|\boldsymbol{B}|$。这一结果也被称为极分解（polar decomposition）。这样，式(0.74)左端可以表示为

$$\begin{aligned} \left\|\boldsymbol{A}\boldsymbol{B}^{\dagger}\right\|_1 &= \mathrm{tr}\big(|\boldsymbol{A}\boldsymbol{B}^{\dagger}|\big) \\ &= \mathrm{tr}\big(|(\boldsymbol{U}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger})|\big) \\ &= \mathrm{tr}\,\sqrt{(\boldsymbol{U}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger})^{\dagger}(\boldsymbol{U}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger})} \\ &= \mathrm{tr}\,\sqrt{\boldsymbol{V}|\boldsymbol{B}||\boldsymbol{A}||\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger}} \\ &= \mathrm{tr}\,\sqrt{\boldsymbol{V}|\boldsymbol{B}||\boldsymbol{A}|\boldsymbol{V}^{\dagger}\boldsymbol{V}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger}} \\ &= \mathrm{tr}\big(\boldsymbol{V}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{V}^{\dagger}\big) \\ &= \mathrm{tr}(|\boldsymbol{A}||\boldsymbol{B}|) \end{aligned} \tag{0.76}$$

按照式(0.66), 可以对本征值构成的向量使用向量情形的 Hölder 不等式。对于满足 $p^{-1}+q^{-1}=1$ 的正实数 p,q，

$$\begin{aligned} \mathrm{tr}(|\boldsymbol{A}||\boldsymbol{B}|) &\leqslant \sum_{i,j} |\alpha_i\beta_j| \\ &\leqslant \left(\sum_i |\alpha_i|^p\right)^{1/p} \left(\sum_j |\beta_j|^q\right)^{1/q} \\ &= \|\boldsymbol{A}\|_p \|\boldsymbol{B}\|_q \end{aligned} \tag{0.77}$$

□

这一证明过程展示了在处理非方阵运算时，常用的几种数学技巧。关于一般的矩阵，还有许多有趣的函数。感兴趣的读者可以参考文献 [2]。

0.4 经典信息论简介

在介绍信息论之前，首先需要明确：数学上，信息到底代表什么。为了回答这一问题，先来看一些简单的例子，并通过数据压缩这一信息处理任务来量化信息。随后，我们总结一些经典信息论中的重要内容。

0.4.1 香农熵

想象这样一个情景：两个好友，甲和乙[①]，分别住在北京和上海。甲给乙发送关于北京天气的消息。方便起见，我们只考虑两种天气：晴天或雨天。在甲发送消息之前，乙想猜测一下北京的天气状况。直观上，乙通过接收甲的消息所能获取的有用信息取决于他的先验知识。比如说，如果乙了解到北京几天来一直晴空万里，而且短时间内周边地区也没有任何积雨云，那他几乎百分百确定北京的天气将是晴天。一般地，假设乙通过自己的了解对北京的降水概率进行如下的猜测：

（1）10% 降雨, 90% 晴天；

（2）50% 降雨, 50% 晴天；

（3）90% 降雨, 10% 晴天。

如果稍后甲给乙发送了“晴天”的消息，在哪种情况下乙可以获得最多的信息？显然，在第一种情形，乙获得的信息最少，因为他在早先就比较确定北京将是晴天；在第二种情形，乙事先完全不确定北京的天气；而在第三种情形，当乙收到甲的消息时，他大概会大吃一惊——也就是说，他得到了最多的信息。

通过这个例子，我们获得了这样的一个直观感受：信息应该与先验概率有着密切联系。你可能会好奇，在这个例子里，“晴天”和“雨天”这几个字包含了怎样的信息。这里，我们限定了乙只是想猜测北京的天气究竟是两种中的哪一个，事实上，你可以将两个词替换成任何其他的事物或者事件。比如说，你走到下一个路口时会不会遇到红灯，你从一个装了黑球和白球的袋子里拿出一个球的颜色是黑色还是白色。基于同样的分析，你对于是否遇到红灯或者是否拿出黑球这件事所能得到的信息同样只依赖于你对红灯或黑球出现概率的先验判断。因此，信息量应该与具体事件（晴天/雨天，红灯/绿灯，黑球/白球 ……）的表示细节没有关系。这样一种内容无关的抽象描述方法是香农信息论（Shannon information theory）的出发点。

在这一节，我们先严格地将讨论限定在经典信息论，即我们所分析的对象都是经典随机变量。我们将使用“比特”（bit）来表示信息量的基本单位。一个比特可以表示为一个数位，0 或者 1。一个两点分布（也称为伯努利分布，Bernoulli distribution）的二元随机变量（binary random variable），如果发生事件 0 和

① 在量子信息英文书籍中经常涉及一男一女两个角色，Alice 和 Bob，这里我们用甲和乙对应。所以，本书中甲是女性角色，相对应的角标为 A，乙是男性角色，相对应的角标为 B。

1 的概率相等，这一随机变量取值所包含的信息量为一个比特。对于一个一般的服从概率分布 $p(x)$ 的随机变量 X，其所包含的信息量可以用香农熵（Shannon entropy）量化，定义为

$$H(X) = -\sum_x p(x)\log[p(x)] \tag{0.78}$$

其中对 x 的求和遍历随机变量 X 的所有可能取值。本书中，如无特殊说明，对数函数的底数均为 2。另外，我们以极限的方式取 $0\log 0 = 0$。

思考题 0.6 (香农熵的数学合理性) 考虑一个随机实验中的事件，并将其对应的随机变量记为 E。让我们考虑这个随机变量的“信息函数”或“惊奇系数”，$I(E)$。我们要求这个函数满足以下数学性质：

（1）$I(E)$ 只是关于 E 的概率分布（概率密度函数）的函数，

$$I(E = e) = I[p_E(e)] \tag{0.79}$$

其中，$p_E(e) \in (0, 1]$，并且 $\sum_e p_E(e) = 1$。

（2）I 在区间 $(0, 1]$ 上平滑。

（3）两个独立事件所带来的总的惊奇系数等于各自惊奇系数的和：

$$I(pq) = I(p) + I(q) \tag{0.80}$$

其中，$p, q \in (0, 1]$ 是两个独立事件各自的发生概率。

在确定函数 I 的形式后，请计算随机变量 E 的“平均信息量”，即 $\mathbb{E}_{\Pr(E)}[I(E)]$，并将其与式(0.78)比较。

当考虑二元随机变量时，香农熵变成二元熵（binary entropy），通常用某一事件的概率表示。对于 $x \in [0, 1]$，可以定义二元熵函数（binary entropy function），

$$h(x) = -x\log x - (1-x)\log(1-x) \tag{0.81}$$

由前面 log 函数的极限规定 $0\log 0 = 0$，可以很快看到 $h(0) = h(1) = 0$。另外，对于 $p = 11\%$，二元熵的值为 $h(p) \approx 0.5$，换句话说，在 N 比特长的字符串中只有 $N/2$ 比特的信息。二元熵函数随概率的变化如图 0.6 所示。

0.4.2 数据压缩

前面我们介绍了香农熵的概念。除了数学上的“直觉”，在实际信息处理中，这样的度量又有怎样的操作含义？为了回答这一问题，我们考虑这样一个通信情景。甲希望通过一个比特信道向乙传送信息。假设甲的信息源会随机地从字母表 $\{a, b, c, d\}$ 中按照下面的概率分布选出一个字符，

$$\Pr(a) = \frac{1}{2}, \quad \Pr(b) = \frac{1}{8}, \quad \Pr(c) = \frac{1}{4}, \quad \Pr(d) = \frac{1}{8} \tag{0.82}$$

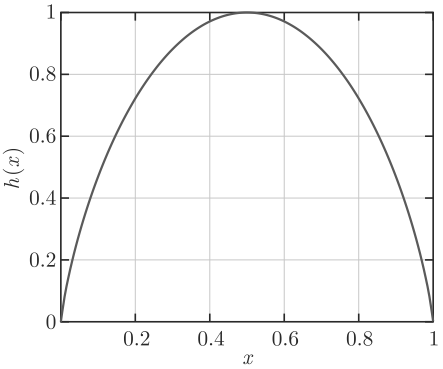

图 0.6　二元熵函数随概率的变化曲线

对于一个比特信道，它只能接收比特，即它不能将字符 a, b, c, d 直接作为输入。为此甲需要将她要传送的字符编码为比特串。甲可以用下面的编码方式：

$$a \to 00, \quad b \to 01, \quad c \to 10, \quad d \to 11 \tag{0.83}$$

这是一个定长编码，码字（code word）长度总是 2 个比特。除此之外，还有一种编码方案：

$$a \to 0, \quad b \to 110, \quad c \to 10, \quad d \to 111 \tag{0.84}$$

期望上，这个编码方案的平均码字长度为

$$\frac{1}{2} \times 1 + \frac{1}{8} \times 3 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 = \frac{7}{4} \tag{0.85}$$

假如这个通信任务重复许多次，并且每次信息源都是按照式(0.82)的概率分布独立且同分布（independently and identically distributed, i.i.d.）地发出信息，那么很明显，第二种方案大概率会节省通信所需发送的比特数。我们可以直观地看出，这一编码方案节省通信所需比特数的原因是用较短的码字编码频繁出现的字符，而用较长的码字编码较少出现的字符。而这一编码方案的平均信息量与香农熵有怎样的关系呢？按照式(0.78)的定义，

$$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{8}\log\frac{1}{8} = \frac{7}{4} \tag{0.86}$$

正好是式(0.85)的计算结果。

思考题 0.7　这一结果是巧合吗？还是有更为本质的原因？

香农对无噪声信道编码问题进行了严格的讨论，并得到了信源编码定理（Shannon's source coding theorem）[3]。这一定理指出，对于独立同分布的一串随机变量，在渐进极限下（即随机变量数量趋于无穷时），如果对这些随机变量所含的数

据进行压缩，压缩编码率（即平均每个字符所需的编码比特数）不可能少于信息源所产生的随机变量的香农熵，否则必然会导致信息丢失。另外，一定存在编码方案，使得在对足够多的随机变量进行编码时，编码率任意接近于香农熵。

香农熵的操作含义体现在一个具体的编码方案中。考虑二元随机变量 $X \in \{0,1\}$, 以及相应的比特串 $X^n \in \{0,1\}^n$，其中对任意 $i \in [n]$，第 i 个二元随机变量 X_i 是服从下面概率分布的独立同分布随机变量：

$$\begin{cases} \Pr(X_i = 1) = p \\ \Pr(X_i = 0) = 1 - p \end{cases} \tag{0.87}$$

其中，$p \in [0,1]$。我们希望将比特串 X^n 如实存储，即制备一个足够大的内存。如果不希望产生任何差错，显然我们需要将 X^n 的每一个比特都进行记录，因此内存大小至少是 n 比特。但在实际应用中，通常我们会允许一个足够小的可以忽略的失败概率。在这一前提下，我们注意到下面的事实。

观察点 所有的 n 比特字符串构成一个 2^n 维空间，但在 $p \neq 0.5$ 时，有些比特串相比于其他比特串的出现概率明显要低。

为了清晰地说明这一事实，我们考虑 $p < 0.5, n \gg 1$ 时，对于可能出现的字符串 0^n 和 1^n,

$$\frac{\Pr(X = 0^n)}{\Pr(X = 1^n)} = \frac{(1-p)^n}{p^n} \gg 1 \tag{0.88}$$

受到这一发现的启发，可以想象这样一个编码存储方案：我们只存储更可能发生的比特串，而忽略掉那些几乎不会出现的比特串。将这样的想法严格表述出来，我们可以定义 ε-最小概然集合。

定义 0.10 (ε-最小概然集合，ε-smallest probable set) 给定 $0 \leqslant \varepsilon < \frac{1}{2}$ 和服从概率分布 p_X 的有限值域随机变量 $X \in \mathcal{X}$，其中样本空间 $\mathcal{X}$ 有限大，$|\mathcal{X}| < \infty$, 随机变量 X 的 ε-最小概然集合定义为样本空间的最小子集 $\mathcal{T}_X^\varepsilon \subseteq \mathcal{X}$, 使得在一次随机试验中，$X$ 的取值以不超过 ε 的失败概率落在该集合中，

$$\begin{aligned} &\mathcal{T}_X^\varepsilon = \arg\min_{\mathcal{S}} |\mathcal{S}|, \\ &\text{s.t. } \Pr(X \in \mathcal{S}) \geqslant 1 - \varepsilon \end{aligned} \tag{0.89}$$

对于连续随机变量（continuous random variable），或者说无穷维随机变量，我们同样可以考虑类似于最小概然集合的概念，但此时需要对集合大小等概念进行合适的推广。本书不讨论这一内容，对此感兴趣的读者可以参考文献 [4] 等经典信息论教材。

给定允许的失败概率 ε, 前面讨论的数据存储问题转化为确定 $\mathcal{T}_{X^n}^\varepsilon$ 的问题，相应地，内存所需大小对应的信息量由 $\log|\mathcal{T}_{X^n}^\varepsilon|$ 确定。准确地找到 $\mathcal{T}_{X^n}^\varepsilon$ 这个集合

一般来说是困难的，但我们可以退而求其次，寻找一个较小的最概然集合，使得 X 的实现至少以 $(1-\varepsilon)$ 的概率落入其中。为了找到这样一个集合，我们可以寻找合适的函数，将寻找集合元素的问题转化为确定这些元素相应的函数取值问题。

对于我们现在考虑的独立同分布信息源，期望上，X^n 中含有字符 1 的数量为 np, 在整个空间中这样的字符串的个数由二项式系数给出，$\binom{n}{np}=\dfrac{n!}{(np)!(n-np)!}$。为了表述方便起见，我们定义 n 比特字符串 $x=(x_0,\cdots,x_{n-1})\in\{0,1\}^n$ 的权重为

$$wt(x)=\sum_{i=0}^{n-1}x_i \tag{0.90}$$

给定常数 $c>0$, 定义 $\mathcal{D}(c)$ 为所有权重在区间 $[np-c\sqrt{n},np+c\sqrt{n}]$ 的字符串集合。我们称这一集合为参数 c 的典型集（typical set）[①]。典型集、非典型集和整个样本空间的关系如图 0.7 所示。这样，我们的主要任务是计算最概然集合的大小及其失败概率 ε。下面例子给出如何估算一个典型集的大小。

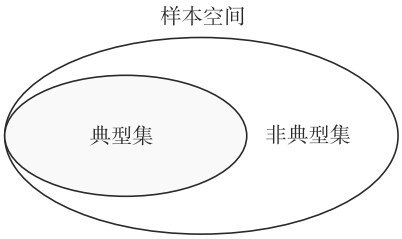

图 0.7　典型集的图像化示意图。在编码成功概率渐进收敛到 1 时，采样所得的比特串几乎必然落入典型集中。香农源编码定理告诉我们，典型集中包含的比特串数量和全空间中比特串数量的比值，即编码压缩率，在渐进极限下由数据源的香农熵给出

例 0.9　对于 $p\in(0,1)$, 集合 $\mathcal{D}(c)$ 的大小

$$|\mathcal{D}(c)|=\sum_{k=-c\sqrt{n}}^{c\sqrt{n}}\binom{n}{k+np} \tag{0.91}$$

可以给出上界估计：

$$|\mathcal{D}(c)|\leqslant 2^{nh(p)-c\sqrt{n}\log p(1-p)} \tag{0.92}$$

对于 $p=0$ 或者 $p=1$，根据 $\mathcal{D}(c)$ 的定义，式 (0.92) 自然成立。该例题中，主要关注 $0<p<1$ 的情况。另外，为了讨论方便，假设 $k+np$ 是整数，而且在区间 $[0,n]$ 内。如果不是，式子稍作调整还是成立。

---

① 在信息论文献中，“典型集”一般特指独立同分布随机变量多次重复时的情形，其定义也稍有不同，由渐进等分定理（asymptotic equipartition theorem）导出。在这里，我们不严格区分最概然集（probable set）和典型集。对于多次重复的独立同分布随机变量，这两个集合大小在渐进表现上是一致的[4]。

解 考虑 n 个独立同分布的两点分布随机变量 $X \in \{0,1\}$，其概率密度函数为

$$\Pr(X) = \begin{cases} p, & X = 1 \\ 1-p, & X = 0 \end{cases} \tag{0.93}$$

记 n 个随机变量 $X = (X_1, X_2, \cdots, X_n)$ 取得的字符串为 x。由于是独立随机分布，记 $wt(x) = k$，则 $X = x$ 的概率为

$$\begin{cases} \Pr(x) = p^k (1-p)^{n-k} \\ \log \Pr(x) = k \log(p) + (n-k) \log(1-p) \end{cases} \tag{0.94}$$

如果这一字符串落入典型集，$x \in \mathcal{D}(c)$，即 $k \in [np - c\sqrt{n}, np + c\sqrt{n}]$，于是有

$$\begin{cases} k \leqslant np + c\sqrt{n} \\ n - k \leqslant n(1-p) + c\sqrt{n} \end{cases} \tag{0.95}$$

对上述不等式两端分别乘以 $\log p$ 和 $\log(1-p)$，

$$\begin{cases} k \log p \geqslant (np + c\sqrt{n}) \log p \\ (n-k) \log(1-p) \geqslant [n(1-p) + c\sqrt{n}] \log(1-p) \end{cases} \tag{0.96}$$

代入式(0.94)得到：

$$\begin{aligned} \log [\Pr(x)] &\geqslant (np + c\sqrt{n}) \log p + [n(1-p) + c\sqrt{n}] \log(1-p) \\ &= -nh(p) + c\sqrt{n} \log p(1-p) \end{aligned} \tag{0.97}$$

这里，我们利用了二元熵定义简化这一表达式。典型集 $\mathcal{D}(c)$ 是集合 $\{0,1\}^n$ 的子集，由概率归一化我们知道，

$$\begin{aligned} 1 &= \sum_{x \in \{0,1\}^n} \Pr(x) \\ &\geqslant \sum_{x \in \mathcal{D}(c)} 2^{\log \Pr(x)} \\ &\geqslant 2^{-nh(p) + c\sqrt{n} \log p(1-p)} \cdot |\mathcal{D}(c)| \end{aligned} \tag{0.98}$$

在对不等式各项调整顺序后，便得到了 $\mathcal{D}(c)$ 集合大小的上界估计：

$$|\mathcal{D}(c)| \leqslant 2^{nh(p) - c\sqrt{n} \log p(1-p)} \tag{0.99}$$

□

对于一个典型集，我们可以采用二进制编码，这样信息源 X^n 编码储存所需内存大小为

$$\log|\mathcal{D}(c)| \leqslant nh(p) - c\sqrt{n}\log p(1-p) \tag{0.100}$$

给定这个典型集，下面来计算相应的失败概率 ε。记 $\delta = \dfrac{c}{p\sqrt{n}}, \mu = np$, 那么

$$\begin{aligned}
\varepsilon &= \Pr[X^n \notin \mathcal{D}(c)] \\
&= \Pr(|wt(X^n) - \mu| \geqslant \delta\mu) \\
&\leqslant 2\left[\frac{\mathrm{e}^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\mu} \\
&= \frac{2\mathrm{e}^{c\sqrt{n}}}{\left(1+\dfrac{c}{p\sqrt{n}}\right)^{(1+\frac{c}{p\sqrt{n}})np}} \\
&= \frac{2\mathrm{e}^{c\sqrt{n}}}{\left(1+\dfrac{c}{p\sqrt{n}}\right)^{\frac{p\sqrt{n}}{c}(c\sqrt{n}+\frac{c^2}{p})}}
\end{aligned} \tag{0.101}$$

在式 (0.101) 第三行，使用了双边切诺夫不等式 (Chernoff inequality)。对于 $n \to \infty$, 有

$$\frac{2\mathrm{e}^{c\sqrt{n}}}{\left(1+\dfrac{c}{\sqrt{np}}\right)^{\frac{\sqrt{n}p}{c}(c\sqrt{n}+\frac{c^2}{p})}} \overset{n\to\infty}{\longrightarrow} \frac{2\mathrm{e}^{c\sqrt{n}}}{\mathrm{e}^{c\sqrt{n}+\frac{c^2}{p}}} = 2\mathrm{e}^{-\frac{c^2}{p}} \tag{0.102}$$

这里使用了欧拉数的极限定义：

$$\mathrm{e} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \tag{0.103}$$

可以看到，当 c 取一个较大的值，比如 100 时，所对应的失败概率会指数地趋近于 0。与式(0.100)结合，我们可以看到，信息源编码储存单个比特所需内存大小在 $n \to \infty$ 时趋向于 $h(p)$，同时失败概率收敛到 0。

注意点 (非 iid 情形)　如果字符串中的各随机变量相互关联（即不是独立同分布变量），并且关联的形式是已知的，那么由于有更多的限制条件，典型集会比独立同分布情形小，因此编码所需比特数更少。

但需要注意的是，如果仅知道边际概率分布（marginal probability distribution），对于随机变量之间具体的关联形式未知，典型集大小可以是任意的，特别地，可以比独立同分布情形大。

0.4.3 其他的熵形式信息量

基于香农熵，我们可以将其他信息量，特别是多个随机变量之间关联，利用熵的形式表示出来。对于联合随机变量 (X,Y), 可以定义它们的联合熵（joint entropy）：

$$H(X,Y)=-\sum_{x,y}p_{X,Y}(x,y)\log[p_{X,Y}(x,y)] \tag{0.104}$$

一般来说，X,Y 不是相互独立的。假设甲持有随机变量 X, 乙持有随机变量 Y, 我们称乙拥有 Y 形式的关于 X 的侧信息（side information）。为了量化乙的侧信息，我们可以用条件熵（conditional entropy）：

$$H(X|Y)=-\sum_{x,y}p_{X,Y}(x,y)\log[p_{X|Y}(x|y)] \tag{0.105}$$

条件熵 $H(X|Y)$ 表示在乙已知随机变量 Y 的情况下，对于 X 的不确定度。类似于香农熵，条件熵的操作含义应该在独立同分布随机变量重复实验的渐进极限下理解。

思考题 0.8 验证 $H(X,Y)=H(X)+H(Y|X)=H(Y)+H(X|Y)$。

依然考虑上面的情形，即甲和乙分别拥有随机变量 X 和 Y。为了刻画两个随机变量之间的关联，可以考虑它们的互信息（mutual information），定义是一个边际随机变量的熵，如 $H(X)$，和相对应的条件熵 $H(X|Y)$ 之间的差：

$$I(X:Y)=H(X)-H(X|Y) \tag{0.106}$$

在下面的维恩图（Venn diagram）中（图 0.8），直观地画出边际随机变量的熵、联合熵、条件熵、互信息的关系。

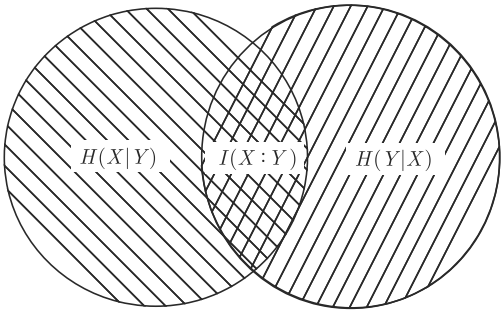

图 0.8 互信息、边际随机变量熵、条件熵的 Venn 关系图。图中两个圆分别代表 $H(X)$ 和 $H(Y)$, 它们的交叉的地方为 $I(X:Y)$, 它们的并集为 $H(X,Y)$, 两个半月牙形代表了 $H(X|Y)$ 和 $H(Y|X)$

思考题 0.9 证明互信息对于函数的两个自变量是对称的，即 $I(X:Y)=I(Y:X)$。

对任意随机变量 X 和 Y, 互信息 $I(X:Y)$ 总是非负的，$I(X:Y)\geqslant 0$。

此外，在信息处理任务中，我们经常希望量化两个随机变量之间的“距离”，或者说一个概率密度函数 $p_{X_1}(x)$ 相比于另一个概率密度函数 $p_{X_2}(x)$ 有多“远”。为此，可以使用相对熵（relative entropy）。在经典信息论等领域，这一信息量经常被称为 Kullback-Leibler 散度（Kullback-Leibler divergence）。定义相对熵 $D(p_{X_1}\|p_{X_2})$ 为

$$D(p_{X_1}\|p_{X_2})\equiv\sum_x p_{X_1}(x)\log\left[\frac{p_{X_1}(x)}{p_{X_2}(x)}\right] \tag{0.107}$$

对于相对熵的操作含义，一种理解方式是它刻画了错误编码的额外开销，即如果在对随机变量 X_1 进行编码时，错误地使用了概率密度函数 p_{X_2}。在这一情形下，编码平均所需比特数为 $[H(X_1)+D(p_{X_1}\|p_{X_2})]$。容易验证，如果编码所用概率密度函数是正确的，相对熵 $D(p_{X_1}\|p_{X_1})=0$, 则编码平均所需比特数变为 $H(X_1)$。

定理 0.2 (香农熵的次可加性)　两个随机变量 X 和 Y 的互信息等于联合随机变量概率分布 $p_{X,Y}(x,y)$ 与边际概率分布乘积 $p_X(x)p_Y(y)$ 的相对熵：

$$D[p_{X,Y}(x,y)||p_X(x)p_Y(y)]=I(X{:}Y)=H(X)+H(Y)-H(X,Y) \tag{0.108}$$

关于这个定理的证明会在 4.1 节具体给出。这一定理计算了两个随机变量 X 和 Y 定义的联合随机变量概率分布 $p_{X,Y}(x,y)$ 与边际概率分布乘积 $p_X(x)p_Y(y)$ 的距离。在这个意义上，互信息刻画了与相互独立的联合随机变量距离有多远。

需要注意的是，相对熵并不是真正的距离度量（distance measure）。此外，这一函数 $D(p_{X_1}\|p_{X_2})$ 在某些情况下存在不好的数学特征。如果第一个自变量概率分布 $p_{X_1}(x_1)$ 的支撑集（support）没有全部包含在第二个自变量概率分布 $p_{X_2}(x_2)$ 的支撑集中，即 $\exists X=x$ 使得 $p_{X_1}(x)\neq 0$ 但 $p_{X_2}(x)=0$, 相对熵函数会发散。

注意点　为了说明相对熵函数不是真正的距离度量，首先注意到相对熵对于两个自变量不是对称的，即一般情况下，$D(p_{X_1}||p_{X_2})\neq D(p_{X_2}||p_{X_1})$。比如，考虑 $p_{X_1=1}=p_{X_1=2}=\dfrac{1}{2}, p_{X_2=1}=1$, 那么 $D(p_{X_2}||p_{X_1})=1$ 但 $D(p_{X_1}||p_{X_2})$ 发散。尽管如此，我们经常会将其看作一种对概率分布距离的反映。特别地，相对熵函数与真正的距离度量有着紧密联系。比如，相对熵与迹距离可以由 Pinsker 不等式联系起来：

$$D(p_X\|p_Y)\geqslant\frac{1}{2\ln 2}\|p_X-p_Y\|_1^2 \tag{0.109}$$

其中，ln 是以欧拉数 e 为底的自然对数，$\|p_X-p_Y\|_1\equiv\sum\limits_a|p_X(a)-p_Y(a)|$。

0.4.4　信道编码

对于带有噪声的信道，香农也进行了研究，并得到了含噪声信道编码定理（noisy channel coding theorem）[3]。考虑利用带有噪声的信道进行通信的两个人，

甲和乙。假设信道是对称的，即无论信息比特是 0 或 1，均以概率 p 对其进行翻转，见表 0.2。

表 0.2　含噪的对称比特信道

发送信息比特	不同接收信息比特下的概率	
	0	1
0	$1-p$	p
1	p	$1-p$

通信中，甲依然希望如实地向乙传送信息。设甲需要传送的信息为 $X \in \{0,1\}^n$，在 X 经过信道后，乙收到的信息为 $Y \in \{0,1\}^n$。定义传输误码随机变量为 $E \in \{0,1\}^n$，如果 X 的第 i 个比特发生了翻转，则 $E_i = 1$，否则 $E_i = 0$。这样，有 $Y_i = X_i \oplus E_i$，这里 $\oplus$ 为异或操作，即比特求和取模 2。我们可以相应地定义比特串之间的异或操作，于是有 $Y = X \oplus E$。如果乙获得了关于 E 的额外信息，那么乙可以通过将发生错误的比特翻转回去以纠正错误。为此，我们设想甲和乙还共享了一个无噪声的理想信道，以此乙可以收取关于纠错的信息。不过这一信道的使用成本很高，因此，甲和乙希望找到一个信息编码方案，以尽可能少的代价传送关于纠错所需的信息。注意到 E 中每个比特服从独立且相同的概率分布：

$$E_i = \begin{cases} 0 & w.p. \quad 1-p \\ 1 & w.p. \quad p \end{cases} \tag{0.110}$$

将错误类型 E 看作重复实验中的信息源，通过应用前面阐述的信息源编码定理，在渐近极限 $n \to \infty$ 下，错误类型可以被压缩至大小约为 $nh(p)$ 比特的集合中。这样，只需要大约 $nh(p)$ 个比特，乙就可以以很高的概率完成错误纠正。

尽管信道编码原理上是可行的，但在实际中操作起来有很大的难度。主要的问题在于，信息的编码和解码复杂度极高。如果我们忽略这一问题，一个概念上容易理解的方案是使用通用随机哈希（universal random hashing）。甲和乙共享一个满秩的随机矩阵 $\boldsymbol{G}$，并且甲通过无噪声信道向乙传送 $\boldsymbol{G}X$，这里我们把字符串 X 看成一个列向量。乙可以计算矩阵 $\boldsymbol{G}$ 的广义逆并相应地进行错误纠正。稍后在介绍量子熵和通信任务中的量子信道模型时，我们会继续深入信道编码问题，并由此定义信道容量（channel capacity）。

注意点　错误纠正的核心是典型集的大小。在有限码长的实际情形中，需要考虑必要的冗余。

关于刚才的内容，你可能会想，实际中是否真的存在无噪声信道。事实上，可以利用信道编码方案来对抗噪声，这就是香农含噪声信道编码理论（Shannon's noisy channel coding theory）。一般地，给定一个带有噪声的信道 $\mathcal{N}$，如果希望传

输信息 M, 与其直接将其通过信道进行传输，可以首先对其进行编码，将信息编码至一个新的随机变量 X 中。这一随机变量经过带有噪声的信道变为 Y, 在收到这一随机变量后，接收端对其进行解码并获取信息 M'。这一过程如图 0.9 所示。

图 0.9　通用的信道编码方案示意图

利用合适的编码解码方案，解码收到的信息 M' 可以与原信息 M 有很高的保真度。尽管如此，信道编码的能力不是无限的。即使在最优编码方案下，由于带有噪声的信道本身的限制，可以保证的传输率存在理论上界。传输率是指通过合理的编码解码方案设计，以及可能趋向于无穷次地独立使用相同的含有噪声的信道，平均每次信道使用中准确传输的比特数的期望值。对于每一种编码解码方案，原则上都存在这样一个传输率。所有传输率的上确界被称为信道容量。香农的含噪声信道编码定理告诉我们，信道容量由下式给出：

$$C(\mathcal{N}) = \sup_{p_X} I(X:Y) \tag{0.111}$$

其中，X 和 Y 是信道输入和输出的随机变量，优化遍历 X 所有可能的概率分布 p_X。对于很一般的信道，通常来说，即使知道了信道噪声的严格描述（例如，一个概率转移矩阵），信道容量也很难准确计算。但对于一些特定的信道，信道容量可以有解析表达式，比如表 0.2 提到的对称比特翻转信道。

思考题 0.10　请尝试给出这一信道的信道容量表达式，并找到一种对应的编码解码方案，使得传输率可以达到信道容量。

0.4.5　Rényi 熵

香农熵可以用几条关于信息度量的公理导出。事实上，香农熵属于一类更大的可以由公理化体系导出的熵度量，即 Rényi 熵。在稍后介绍信息熵的量子推广时，我们会简单介绍这一公理化体系。在本书中，我们不会深入讨论这部分的数学。在这里，给出 Rényi 熵的定义及一些信息论中常用的 Rényi 熵。

定义 0.11　给定常数 α, 满足 $\alpha \geqslant 0$ 且 $\alpha \neq 1$, 定义 X 的概率分布为 $p_i{}_{i=1}^{n}, p_i > 0$（需要保证 max-entropy 定义），α 阶的 Rényi 熵定义为

$$H_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^\alpha\right) \tag{0.112}$$

下面讨论 Rényi 熵的一些特例。

- **Hartley 熵（Hartley entropy）或最大熵（max-entropy）**：

$$H_0 = \log n = \log |X| \tag{0.113}$$

- **香农熵**: 对于 H_α, 取极限 $\alpha \to 1$,

$$H(X) = -\sum_i p_i \log p_i \tag{0.114}$$

- **碰撞熵（collision entropy）**：这一熵度量有时被直接称作 “Rényi 熵”, 对应于 $\alpha = 2$,

$$H_2(X) = -\log \sum_i p_i^2 \tag{0.115}$$

- **最小熵（Min-entropy）**：在极限 $\alpha \to \infty$, Rényi 熵 H_α 收敛到最小熵 H_∞:

$$H_\infty(X) = -\max_i \log p_i \tag{0.116}$$

例 0.10 证明：

$$\lim_{\alpha\to 1} H_\alpha(X) = H(X) = -\sum_i p_i \log p_i \tag{0.117}$$

其中，$\{p_i\}_i$ 为随机变量 X 对应的概率分布。

解 使用洛必达法则（L'Hôspital's rule），有

$$\begin{aligned}
\lim_{\alpha\to 1} H_\alpha(X) &= \lim_{\alpha\to 1} \frac{1}{1-\alpha} \log\left(\sum_i p_i^\alpha\right) \\
&= \lim_{\alpha\to 1} \frac{\dfrac{\partial}{\partial\alpha} \log(\sum_i p_i^\alpha)}{\dfrac{\partial}{\partial\alpha}(1-\alpha)} \\
&= \lim_{\alpha\to 1} \frac{\dfrac{1}{\ln 2} \sum_i p_i^\alpha \ln p_i}{-\sum_i p_i^\alpha} \\
&= -\sum_i p_i \log p_i
\end{aligned} \tag{0.118}$$

□

习题

习题 0.1 (狄拉克符号)

(1) 请写出 $|0\rangle, |1\rangle, |\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}, |\pm\mathrm{i}\rangle = (|0\rangle \pm \mathrm{i}\,|1\rangle)/\sqrt{2}$ 的向量形式。

(2) 请写出 $|+\rangle\langle \mathrm{i}|-\rangle\langle 0| + |1\rangle\langle -\mathrm{i}|$ 对应的矩阵形式。

习题 0.2 (矩阵计算) 已知矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 和 $\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$，试计算 $|\boldsymbol{A}|$，$|\boldsymbol{B}|$ 和 $|\boldsymbol{B}^\dagger|$。

习题 0.3 (矩阵对易) 考虑一定维度的方矩阵，试证明，

(1) 如果一个矩阵和所有对角矩阵对易，那么它一定是对角矩阵。

(2) 如果一个矩阵和所有矩阵对易，那么它一定是 $\boldsymbol{I}$ 的数乘。

习题 0.4 (左右矢的偏迹计算) 给定 $T \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_R)$，$\mathcal{H}_R$ 的一组基矢 $\{|i\rangle_R\}$，以及在 $\mathcal{H}_S$ 下定义的单位矩阵 $\boldsymbol{I}_S$，尝试通过矩阵形式的具体计算说明下面表示的正确性：

$$\mathrm{tr}_R(T) = \sum_i (\boldsymbol{I}_S \otimes \langle i|_R) T (\boldsymbol{I}_S \otimes |i\rangle_R) \equiv \sum_i \langle i|\, T\, |i\rangle_R \tag{0.119}$$

习题 0.5 (密度矩阵)

(1) 矩阵 $\boldsymbol{\rho}$ 被称为密度矩阵，当且仅当

$$\begin{cases} \mathrm{tr}(\boldsymbol{\rho}) = 1 \\ \boldsymbol{\rho} = \boldsymbol{\rho}^\dagger \\ \langle \phi|\, \boldsymbol{\rho}\, |\phi\rangle \geqslant 0, \quad \forall\, |\phi\rangle \end{cases} \tag{0.120}$$

试证明存在 $\boldsymbol{\rho}$ 的谱分解 $\boldsymbol{\rho} = \sum_i \lambda_i |i\rangle\langle i|$ 满足 $\sum_i \lambda_i = 1, \lambda_i = \lambda_i^*, \lambda_i \geqslant 0$。

(2) 请讨论 $\mathrm{tr}(\boldsymbol{\rho}^2) = 1$ 的充要条件。

习题 0.6 (p-范数的性质) 试证明 p-范数的次可乘性和单调性，即

$$\|\boldsymbol{AB}\|_p \leqslant \|\boldsymbol{A}\|_p \|\boldsymbol{B}\|_p \tag{0.121}$$

$$\|\boldsymbol{A}\|_1 \geqslant \|\boldsymbol{A}\|_p \geqslant \|\boldsymbol{A}\|_q \geqslant \|\boldsymbol{A}\|_\infty \tag{0.122}$$

其中，$\boldsymbol{A} \in \mathbb{C}^{n\times k}$, $\boldsymbol{B} \in \mathbb{C}^{k\times m}$，以及 $q \geqslant p \geqslant 1$。

习题 0.7 (香农熵的性质)

(1) 考虑抛掷一枚无偏的硬币，并将朝上结果表示为一个二值的随机变量，这一变量的香农熵取值是多少？类似地，考虑抛掷一个重量均匀的骰子，将其朝上一面点数表示为一个随机变量，这一变量的香农熵取值是多少？

（2）考虑抛掷一枚重量不均匀的硬币，其中正面朝上的概率为 p，请计算此时抛掷硬币朝上结果对应的随机变量的香农熵取值，并画出这一取值随 p 变化的曲线。

（3）请证明，对任一随机变量，香农熵的取值总是非负的。

习题 0.8 (Rényi 熵关于 α 的单调性)

（1）考虑一个一般的随机变量 X，请证明，对于 Rényi 熵 $H_\alpha(X)$，$\beta \geqslant \alpha \geqslant 0$，有

$$H_\alpha(X) \geqslant H_\beta(X) \tag{0.123}$$

其中，$H_1(X)$ 在极限意义下，认为是香农熵 $H(X)$。

（2）考虑一个一般的随机变量 X，请证明，

$$H_2(X) \leqslant 2H_\infty(X) \tag{0.124}$$

第 1 章

量子系统的基本描述

在本章中，将讨论量子系统的基础——纯态、复合系统、投影测量、幺正演化。重点关注量子信息中最简单的系统——量子比特。一个量子比特可以用一个量子态来表示，它可以被算符作用，并可以通过测量读出。这些结果可以拓展到任意维度的量子系统上。还将介绍一些相关的概念：玻恩定律、密度矩阵、布洛赫球、泡利矩阵，并利用这些概念，讨论一些有趣的应用和量子世界的基本量子定理，包括量子态层析、不可克隆定理、不可删除定理、普适非门不存在定理等。

1.1 节 ∼1.3 节主要是对一个量子系统的基本描述，包括 1.1 节介绍的对量子状态的刻画、1.2 节介绍的测量及 1.3 节介绍的封闭系统的演化。作为 1.1 节 ∼1.3 节知识的直接应用，1.4 节和 1.5 节将分别介绍量子态层析术和量子系统中的信息守恒现象。这一章的内容是整个量子信息的基础。对于一个有志于从事量子信息科学研究的读者，这一章应该完全掌握。这里我们经常以量子比特系统作为例子，但是里面的结论对于一般情况通用。

1.1 量子纯态

1.1.1 量子态空间与态叠加原理

我们从希尔伯特空间开始讨论量子力学中的物理系统和操作的描述。在量子力学中，一个物理系统的状态由量子态（quantum state）描述。关于物理系统，我们给出量子力学的态公理。

公理 1.1 (量子态公理)　一个孤立物理系统的所有可能状态由一个希尔伯特空间所描述，称为系统的态空间（state space）。系统物理状态由一个单位长度的向量所完全刻画，该向量被称为态向量（state vector）。

在接下来的叙述中，直接称态向量为一个量子态，用一个狄拉克符号中的右矢 $|\psi\rangle$ 来表示。在量子力学中，对于一个由 $|\psi\rangle$ 描述的物理系统，当态向量整体乘以一个非零常数 c 时，得到的向量 $c|\psi\rangle$ 同样描述该系统。换句话说，量子态的全局相位不具有物理含义。所以，一种可能的物理状态对应希尔伯特空间中的一条射线（ray）。

给定两个量子态 $|\phi\rangle$，$|\psi\rangle$，可以通过线性叠加给出另一个可能的量子态，$a|\phi\rangle + b|\psi\rangle$。这种性质被称作态叠加（state superposition）原理。在这一叠加态中，由 a, b 取值所决定的 $|\phi\rangle, |\psi\rangle$ 间的相对相位具有物理含义。例如，我们认为 $a|\phi\rangle + b|\psi\rangle$ 和 $\mathrm{e}^{\mathrm{i}\alpha}(a|\phi\rangle + b|\psi\rangle)$ 的物理状态相同，但与 $a|\phi\rangle + \mathrm{e}^{\mathrm{i}\alpha}b|\psi\rangle$ 的状态不同。

在一个线性空间里面，线性叠加是一个很基本的数学操作。这里需要指出的是，线性叠加操作具有非常深刻的物理含义，区分经典力学的所有量子特性都与量子态叠加相关。事实上，对于叠加态原理的理解是量子力学里面的一个难点。后面会慢慢展开讲述。

量子系统和量子态这两个概念我们在这里清晰化一下。一般将量子系统理解为一个“实物”，比如一个电子、一个光子甚至一个人，而量子态是实物的状态。不同的量子系统可以处在相同的状态上，一个量子系统也可以处在不同的量子态上。比如我们说一个人甲，她可以处在“高兴”“愤怒”等状态。同样，另一个人乙也可以处在这些状态上。从这个角度来看，量子态更像是一种“虚的信息”——量子信息[①]。量子信息科学很多时候是脱离具体量子系统抽象地研究这些量子态。我们会在 1.1.5节简单提及实际物理系统上的量子态。

1.1.2 量子比特

在经典信息论中，信息载体——比特（bit）——表示一个取值为 0 或 1 的随机变量。例如，一个电容器的状态可以离散表示为一个比特：当电容器处于高电平时，将其状态记为 1；处于低电平时，将其状态记为 0。在基于经典物理理论的信息论中，认为 0 和 1 两个状态是可以被准确无误地区分开来的。

在量子信息论中，量子比特（qubit）是比特概念的一种量子对应。类似地，一个量子比特描述了一个由 $|0\rangle, |1\rangle$ 表示的二能级量子系统的状态。与经典比特的区别在于，一个量子比特可以处在两个能级的叠加态上，由 $|0\rangle, |1\rangle$ 的线性组合所描述：

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{1.1}$$

其中，$a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1$。对于所有具有这种叠加形式的量子态（包括 $|0\rangle, |1\rangle$），我们称其为纯态（pure state）。所有的量子比特纯态张成了二维希尔伯特空间 $\mathcal{H}_2$。在接下来的讨论中，将默认 $|0\rangle$ 和 $|1\rangle$ 是正交归一的，即 $\langle 0|0\rangle = \langle 1|1\rangle = 1$，$\langle 0|1\rangle = \langle 1|0\rangle = 0$。称它们构成了 $\mathcal{H}_2$ 的一组基。经典信息处理可以看作量子信息处理的一个特例，其所涉及的状态仅包括正交的两个基向量。由于态叠加原理，

① 这里我们似乎默认量子信息和量子系统是不同的存在。一个很深刻的问题是，这个世界除了量子信息，还有“实物”吗？想象一下，我们平时所接触的不同的实物本质上是能量的不同状态而已。等有一天物理学把所有基本粒子大统一了，是不是只剩下量子信息了？

不同的量子态之间一般不再相互正交，我们将逐渐看到，量子信息处理相比于经典信息处理的优势来源最终都可以归结为这种非正交的特性。

例 1.1 写出下面三个矩阵的本征态，并证明每个矩阵的本征态构成了二维希尔伯特空间的一组基矢。

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \tag{1.2}$$

解 对于 σ_z，本征态为 $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 和 $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$，有 $\langle 0|0\rangle = \langle 1|1\rangle = 1$，$\langle 0|1\rangle = \langle 1|0\rangle = 0$ 因而构成一组基矢。

对于 σ_x，本征态为 $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \equiv |+\rangle$ 和 $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv |-\rangle$，有 $\langle +|+\rangle = \langle -|-\rangle = 1$，$\langle +|-\rangle = \langle -|+\rangle = 0$ 因而构成一组基矢。

对于 σ_y，本征态为 $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + \mathrm{i}\,|1\rangle) \equiv |+\mathrm{i}\rangle$ 和 $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -\mathrm{i} \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - \mathrm{i}\,|1\rangle) \equiv |-\mathrm{i}\rangle$，有 $\langle +\mathrm{i}|+\mathrm{i}\rangle = \langle -\mathrm{i}|-\mathrm{i}\rangle = 1$，$\langle +\mathrm{i}|-\mathrm{i}\rangle = \langle -\mathrm{i}|+\mathrm{i}\rangle = 0$ 因而构成一组基矢。 □

后面我们会看到，这三个矩阵被称为泡利矩阵，它们的本征态构成的三组基矢是二维希尔伯特空间中最常用的三组基矢，分别将其记为 Z 基矢、X 基矢和 Y 基矢。

在量子态公理中，已经假设了量子态的模长为 1，这样做的合理性是怎样的？另外，对于处于态叠加状态的系统，相对相位的物理含义是什么？这两个问题的答案与量子力学中的测量公理或玻恩概率波解释规则（Born's rule）有关，将在 1.2节介绍这一内容。在这里，首先介绍纯态量子比特系统测量基向量的情形。类似于在经典信息处理中，可以通过观测系统处于高低电平以确定比特取值，在量子信息处理中，也可以考虑这样的问题：对于式(1.1)所描述的量子比特，当对其处于 $|0\rangle, |1\rangle$ 中的哪一个状态进行观测时，观测结果将是怎样的？不同的是，在量子信息处理中，观测结果将是概率性的，概率由叠加系数模二次方决定：

$$\mathrm{Pr} = \begin{cases} |a|^2, & \text{状态 “0”} \\ |b|^2, & \text{状态 “1”} \end{cases} \tag{1.3}$$

前面对于量子态系数模长为 1 的要求，$|a|^2 + |b|^2 = 1$，便是对于概率的归一化要求。

基于玻恩规则的量子力学随机性解释是以玻尔为代表的哥本哈根学派的核心观点。从玻恩规则，可以看出量子态公理的合理性。但在上面的描述中，隐含了

这样的假设：在量子力学中，虽然一般的物理状态由可以处于叠加态的量子态描述，但作为观测者，所能测量的物理量却是“经典的”。而且与经典物理理论不同，对于一个完全确定的量子系统，当我们对其进行观测时，观测后其所处物理状态受到测量影响，且测量结果可以是完全随机的。历史上，包括爱因斯坦在内的许多人对于这一结果并不接受。爱因斯坦曾经留下这样一句名言：“上帝不掷骰子（The theory delivers much but it hardly brings us closer to the Old One's secret. In any event, I am convinced that He is not playing dice）。”而后面我们将会看到，量子态公理可以通过实验验证。在第 3 章会介绍可以裁决玻尔与爱因斯坦争论的方法——贝尔不等式。

正如在经典信息论中可以使用高维信息载体，在量子信息论中，也可以考虑高维量子比特。在 d 维的希尔伯特空间 $\mathcal{H}_d$ 中，一个纯的高维量子比特 $|\psi\rangle$ 可以被表示为

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |i\rangle \tag{1.4}$$

其中，$\{|i\rangle\}_{i\in[d]}$ 构成了一组互相正交的基矢 $\mathcal{H}_d$，$c_i \in \mathbb{C}$。与量子比特类似，当在基矢 $\{|i\rangle\}_{i\in[d]}$ 下测量时，得到结果 i 的概率为 $|c_i|^2/(\sum\limits_i |c_i|^2)$。对于一个归一化的高维量子比特，有 $\sum\limits_i |c_i|^2 = 1$。这一章中接下来要介绍的内容并不限于特定维度的系统。为方便起见，我们将重点围绕量子比特展开讨论，在涉及与维度相关的内容时会予以说明。

例 1.2 $\mathcal{H}_d$ 中的一个纯态有多少个自由实参数？

解 从式 (1.4)中可以看到，复系数 c_i 共有 d 项。由于全局相位在量子力学中并不重要，我们总可以取一个非零的 c_i，并将 $|\psi\rangle$ 除以它，而不改变所表示的量子状态。也就是说，我们总是可以假设其中一个系数 c_i 是 1。因此，剩余的自由实参数的数量为 $2d-2$。 □

1.1.3 复合系统

我们已经初步讨论了一个孤立系统的描述，但在实际中，经常会遇到多个物理系统。比如有两个通过光纤和设备管道相连的物理实验室，几个实验者各自处在其中一个实验室中，合作进行物理实验。对于这个合作完成的实验而言，可以将两个实验室整体看作一个复合在一起的更大的实验室。此外，对于一个复杂的物理系统，比如我们所处的太阳系，为了天文探测和研究的方便，我们也会经常将其人为地划分成若干小系统——八大行星。我们称这样由多个部分构成的整体系统为复合系统。通常，将这些整体系统中的组成部分称为子系统（subsystem）。

现在，假设有一个系统由两部分组成，分别记为子系统 A 和子系统 B。对于

这一复合系统，应该如何描述呢？这由量子力学的复合系统公理所给出。

公理 1.2 (量子复合系统空间) 设两个子系统所对应的希尔伯特空间分别为 $\mathcal{H}_A$ 和 $\mathcal{H}_B$，整体系统的希尔伯特空间为 $\mathcal{H}_A \otimes \mathcal{H}_B$。同时，假设两个子系统分别处在量子态 $|\psi\rangle_A$ 和 $|\phi\rangle_B$。那么，整体系统处在量子态 $|\psi\rangle_A \otimes |\phi\rangle_B$。

这里的 $\otimes$ 是张量积（tensor product）运算，不熟悉的读者可以阅读 0.3 节中的相关定义。需要注意的是，子系统 $\mathcal{H}_A$ 和 $\mathcal{H}_B$ 的维度可以不相同，同时它们本身也可以视作一个系统，拥有各自的基矢。复合系统公理表明，对于两个相互“独立” 的态 $|\psi\rangle_A$ 和 $|\phi\rangle_B$，整体的态可以简单地用二者的直积 $|\psi\rangle \otimes |\phi\rangle$ 来表示，大多数时候，会把这个态简写为 $|\psi\rangle\,|\phi\rangle$ 或者 $|\psi\phi\rangle$。

现在考虑一个简单的情形：两个子系统都是二维的。那么，复合系统可以看作一个四维量子系统。如果对子系统分别独立地选取一组基并记作 $\{|0\rangle, |1\rangle\}$，那么对于复合系统，可以用 $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ 作为一组正交归一基。这里，省略了子系统的下标。利用线性叠加原理，可以构造出这个四维空间中的纯量子态。比如，可以构造下面的一组量子态：

$$\begin{cases} |\Phi^+\rangle = \dfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ |\Phi^-\rangle = \dfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ |\Psi^+\rangle = \dfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |\Psi^-\rangle = \dfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \end{cases} \tag{1.5}$$

由于历史的原因，我们称这四个量子态为贝尔态（Bell state），并将在接下来的章节进一步讨论它们的性质。

思考题 1.1 贝尔态能否写成 $|\psi\rangle_A \otimes |\phi\rangle_B$ 的形式？

1.1.4 量子态的密度矩阵表示

量子态的另一种表示方式是密度矩阵（density matrix）。对于 $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$，其对应的密度矩阵算符表示为

$$\begin{aligned} \boldsymbol{\rho} &= |\psi\rangle\langle\psi| \\ &= (a\,|0\rangle + b\,|1\rangle)(a^*\,\langle 0| + b^*\,\langle 1|) \\ &= |a|^2\,|0\rangle\langle 0| + |b|^2\,|1\rangle\langle 1| + ab^*\,|0\rangle\langle 1| + a^*b\,|1\rangle\langle 0| \\ &= \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \end{aligned} \tag{1.6}$$

对于一个纯态，其对应的密度矩阵具有下面的性质。

方框 1：纯态密度矩阵的性质

1. 自伴性，即厄米矩阵（Hermitian）：$\boldsymbol{\rho}^\dagger = \boldsymbol{\rho}$；
2. 归一化：$\mathrm{tr}(\boldsymbol{\rho}) = 1$；
3. 半正定性：$\boldsymbol{\rho} \geqslant 0$，即 $\forall |\phi\rangle \in \mathcal{H}_d, \langle\phi| \boldsymbol{\rho} |\phi\rangle \geqslant 0$；
4. 纯性：$\mathrm{tr}(\boldsymbol{\rho}^2) = 1$。

例 1.3 满足方框 1中性质的矩阵一定可以写成如下形式：

$$\boldsymbol{\rho} = |\psi\rangle\langle\psi| \tag{1.7}$$

解 由于 $\boldsymbol{\rho}$ 是一个厄米矩阵，可以找到一个酉矩阵 $\boldsymbol{U}$ 将其对角化：

$$\boldsymbol{U}\boldsymbol{\rho}\boldsymbol{U}^\dagger = \begin{pmatrix} \lambda_0 & & & \\ & \lambda_1 & & \\ & & \lambda_2 & \\ & & & \ddots \end{pmatrix} \tag{1.8}$$

由性质 2∼ 性质 4 知道，$\sum_i \lambda_i = 1$，$\lambda_i \geqslant 0$，$\sum_i \lambda_i^2 = 1$。于是我们知道只有一个本征值 $\lambda_k = 1$，其余均为 0。记 $|\psi\rangle = \boldsymbol{U}|\lambda_k\rangle$，有 $\boldsymbol{\rho} = |\psi\rangle\langle\psi|$。 □

思考题 1.2 当考虑式 (1.6)对应的量子态的测量问题时，由玻恩规则知，测量结果的概率只由密度矩阵的对角项决定。这是否说明密度矩阵的非对角项是可有可无的？为什么？

1.1.5 实际物理系统

有很多方法可以将量子比特“编码”到实际系统中：电子自旋方向、一个原子或者离子的双能级、光子的路径或者偏振等。物理上，一个自旋 $\frac{1}{2}$ 的系统[①]可以被看作一个简单的量子比特系统。量子比特的许多性质都是由自旋 $\frac{1}{2}$ 系统派生出来的。从历史上看，量子比特的大部分性质是在自旋 $\frac{1}{2}$ 系统中被证明的。

在量子光学中，单个光子的偏振可以用一个量子比特来很好地描述。注意，这并不是一个平凡的结论。光子自旋为 1，但是由于其没有静止质量，完全对称的维度 $s_z = 0$ 被禁止了，换句话说，只允许有两个 Z 方向取值 $s_z = \pm 1$，对应光

① 自旋是基本粒子的内禀性质之一，对于电子来说，自旋等于 $\frac{1}{2}$。

子偏振方向。这样，可以将单个光子视为一个二维量子系统，等价于一个自旋 $\frac{1}{2}$ 的电子系统。

沿不同轴的偏振刚好构成了不同的基矢。光子的偏振：水平和竖直偏振分别用 $|H\rangle$ 和 $|V\rangle$ 表示；对角偏振则由 $|+45°\rangle$ 和 $|-45°\rangle$ 表示或者表示为 $|+\rangle$ 和 $|-\rangle$；圆偏振由 $|R\rangle$ 和 $|L\rangle$ 表示。如果令 $|H\rangle = |0\rangle$，$|V\rangle = |1\rangle$，那么

$$\begin{cases} |R\rangle = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} \\ |L\rangle = \dfrac{1}{\sqrt{2}} \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix} \end{cases} \tag{1.9}$$

这三种偏振态正好对应三组相互正交的基矢。

对于量子信息来说，这些具体的物理系统都是信息的载体。本书的重点在于量子信息本身，并不会深入探讨这些物理系统的性质。

1.2 测量

如果我们观测或者说测量实验室中制备的一个量子态，会获得怎样的测量结果？同时，它们与测量前的系统有什么关系？对于这一问题，量子力学的解释由测量公理给出。到目前为止，所有的实验结果都与该公理相一致。后面小节中，将详细说明可观测量和测量的具体含义和数学描述。

公理 1.3 (测量公理) 测量是观测者对一个物理系统所处状态获取信息的过程。在量子力学中，对系统的可观测量 A 进行测量的结果是将物理系统制备到 A 的一个本征态上，同时，观测者将获取相应的本征值信息。取得某本征值的概率由其对应的本征态与待测系统量子态内积给出。

1.2.1 可观测量

在量子力学中，可观测量，例如系统的位置、动量、能量，由希尔伯特空间中的厄米算符表示。这是我们在原则上对于物理系统所能观测的最一般的性质。同时，测量公理也隐含了如下的含义：所有的物理量都可以由厄米算符表示。数学上，厄米算符是具有如下性质的算符。

（1）线性算符：对于希尔伯特空间 $\mathcal{H}$ 中的向量 $|\psi\rangle$，线性算符 A 将其线性映射到 $\mathcal{H}$ 中的另一个向量，满足 $\forall a, b, \in \mathbb{C}$，

$$\begin{cases} A : |\psi\rangle \mapsto A|\psi\rangle \\ A(a|\phi\rangle + b|\psi\rangle) = aA|\phi\rangle + bA|\psi\rangle \end{cases} \tag{1.10}$$

（2）自反性，即厄米性：对于线性算符，其伴随算符 $A^\dagger$ 由下式定义，$\forall |\psi\rangle, |\phi\rangle \in \mathcal{H}$，

$$\langle\psi|A\phi\rangle = \langle A^\dagger\psi|\phi\rangle \tag{1.11}$$

当且仅当 $A = A^\dagger$，即 $\langle\psi|\,A\,|\phi\rangle = \langle\psi|\,A^\dagger\,|\phi\rangle = \langle\phi|\,A\,|\psi\rangle^*$ 对于任意向量 $|\psi\rangle, |\phi\rangle$ 都成立时，线性算符 A 被称为厄米算符。

厄米算符可以用相应希尔伯特空间上的厄米矩阵来表示。这里给出一些伴随矩阵和厄米矩阵的数学性质，在接下来的章节中将会有帮助。

- 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是厄米矩阵，那么 $\boldsymbol{A}+\boldsymbol{B}$ 也是厄米矩阵，因为 $(\boldsymbol{A}+\boldsymbol{B})^\dagger = \boldsymbol{A}^\dagger + \boldsymbol{B}^\dagger$。此外，$\boldsymbol{AB}+\boldsymbol{BA}$ 和 $i(\boldsymbol{AB}-\boldsymbol{BA})$ 也是厄米矩阵。
- $(\boldsymbol{AB})^\dagger = \boldsymbol{B}^\dagger\boldsymbol{A}^\dagger$，因此如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是厄米矩阵且相互对易（commuting），即 $\boldsymbol{AB} = \boldsymbol{BA}$，那么 $\boldsymbol{AB}$ 是厄米矩阵。
- 厄米矩阵是正规矩阵（normal matrix），因此可以做谱分解（spectral decomposition）。

对于二维量子系统，一类特殊的可观测量可以用泡利矩阵（Pauli matrices）来表示，它们是一组 2×2 维复矩阵：

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.12}$$

文献中，三个泡利矩阵有时也被记作 $\{\sigma_1, \sigma_2, \sigma_3\}$ 或者 $\{X, Y, Z\}$。容易看出，每个泡利矩阵都等于自身的伴随矩阵，因此它们是厄米矩阵。此外，它们的迹均为 0。泡利矩阵之间都是反对易的，$\forall i \neq j \in \{x, y, z\}$，

$$\sigma_i\sigma_j = -\sigma_j\sigma_i \tag{1.13}$$

另外，它们有如下关系：

$$\sigma_y = \mathrm{i}\sigma_x\sigma_z \tag{1.14}$$

通过计算很容易验证，泡利矩阵也都是幺正的。

历史上，泡利矩阵的形式选取具有明确的物理含义。泡利矩阵与李群（Lie group, 一种具有良好微分形式的拓扑群）和李代数（Lie algebra）有着密切的联系，特别是一些常见的群，比如 $SU(2)$ 和 $SO(3)$。对于这一部分感兴趣的读者可以参考文献 [5] 的第 4 章。

对于复合系统，如果只对一个子系统进行观测，那么对于整个系统而言，应该如何描述这一可观测量呢？比如由两个子系统 A 和 B 构成的复合系统，仅对 A 系统观测物理量 σ_x。那么，对于复合系统而言，可以用 $\sigma_x \otimes I$ 来描述这一观测量。

思考题 1.3 对于这一两体复合系统，验证 $\sigma_x \otimes \sigma_x$ 也是一个可观测量。如果对这个观测量进行测量，它的物理含义是什么？

1.2.2 投影测量

从谱分解定理，我们知道厄米矩阵的本征向量构成相应希尔伯特空间的完备基。考虑一个可观测量对应的厄米矩阵 $\boldsymbol{A}$，对其进行谱分解：

$$\boldsymbol{A} = \sum_i \lambda_i |i\rangle\langle i| \tag{1.15}$$

其中，$\{|i\rangle\}$ 是 $\boldsymbol{A}$ 的本征向量谱，满足特征方程 $\boldsymbol{A}|i\rangle = \lambda_i |i\rangle$。注意，这里可能包含零本征值对应的本征向量。先来考虑一个自然而简单的情形是分解不含简并，即矩阵不同的本征态对应于不同的本征值。对这样的厄米矩阵进行的测量被称作冯 • 诺依曼测量（von Neumann measurement）。由于 $\boldsymbol{A}$ 的本征态同时给出了其所作用的希尔伯特空间的一组正交归一基矢，因此，对于待测量的物理系统，可以用这组正交归一基矢进行展开：

$$|\psi\rangle = \sum_i c_i |i\rangle \tag{1.16}$$

对这一量子态测量 $\boldsymbol{A}$，得到结果 λ_i 的概率为

$$p(\lambda_i) = |\langle\psi|i\rangle|^2 = |c_i|^2 \tag{1.17}$$

可以看出，冯 • 诺依曼测量得到结果 λ_i 的含义是提取出 $|\psi\rangle$ 在相应基矢 $|i\rangle$ 上的叠加系数，也可以说是将 $|\psi\rangle$ 向 $|i\rangle$ 这一向量方向上进行投影。

一般来讲，一个厄米矩阵的本征态可能出现简并情况，即在式(1.15)里，对于不同的 i, j，$\lambda_i = \lambda_j$。对于简并于相同本征值的本征态，它们的线性组合依然满足相同本征值的特征方程，因此，谱分解不唯一。对这样的厄米矩阵进行测量，结果是怎样的？我们可以这样理解：考虑 $\boldsymbol{A}$ 的一个特定分解，以及在这一分解下与 $\boldsymbol{A}$ 具有相同本征态的另一个厄米矩阵 $\boldsymbol{A}'$，但 $\boldsymbol{A}'$ 的本征值 $\{\lambda_i'\}$ 与 $\boldsymbol{A}$ 有些区别，使得 $\boldsymbol{A}'$ 的谱分解不含简并。我们首先对量子态进行了 $\boldsymbol{A}'$ 所对应的冯 • 诺依曼测量，随后再对测量结果 λ_i' 重新标记为 $\boldsymbol{A}$ 在相应本征态 $|i\rangle$ 上对应的本征值 λ_i。换句话说，对含有本征值简并的 $\boldsymbol{A}$ 进行测量的结果，可以看作对一冯 • 诺依曼测量结果进行了“粗粒化”（coarse grain），将一些测量结果进行了合并和重新标记。这里，需要注意的是，尽管测量结果是相同的，但这两种实现并不总是等价，A' 测量及后续的粗粒化处理会破坏简并子空间除了本征值外的一些信息，而 A 测量不会如此。

上面的这种测量被称作投影测量（projection-valued measure, PVM），是冯 • 诺依曼测量的一种推广。后面将会看到，这种测量过程涵盖了大多数我们要考虑的问题。为了从数学上理解投影测量，我们从另一个角度来看可观测量 $\boldsymbol{A}$ 的谱

分解。考虑 $\boldsymbol{A}$ 所有不同的本征值 λ_i，对于简并于 λ_i 这一本征值的所有本征向量 $|i_n\rangle$，记 $E_i = \sum\limits_n |i_n\rangle\langle i_n|$，则

$$\boldsymbol{A} = \sum_i \lambda_i E_i \tag{1.18}$$

容易看出，$\forall i, j$，

$$\begin{cases} E_i E_j = \delta_{ij} E_i \\ \sum\limits_i E_i = \boldsymbol{I} \end{cases} \tag{1.19}$$

我们称满足这些性质的算符 E_i 为投影算符（projective operator），其本征值为 0 或者 1。投影算符所对应的测量也就是投影测量。由于上述谱分解中不同本征值对应的投影算符之间相互正交，因此投影测量也被称作一种正交测量。冯 • 诺依曼测量是 E_i 秩为 1 的特殊情形。

由测量公理可以推出，当对一个系统进行投影算符 E_i 所对应的投影测量时，可以得到下面这些结果：

（1）如果测量前的量子态处于 $|\psi\rangle$，那么测量得到结果 λ_i 的概率为

$$p(\lambda_i) = \| E_i |\psi\rangle \|^2 = \langle\psi| E_i |\psi\rangle \tag{1.20}$$

物理上，这相当于将量子态 $|\psi\rangle$ 投影到 E_i 这一子空间上，这也是投影测量名称的来源。

（2）如果测量得到结果 λ_i，测量后的系统将演化为下面的归一化量子态：

$$\frac{E_i |\psi\rangle}{\| E_i |\psi\rangle \|} \tag{1.21}$$

（3）如果一次测量后同样的测量又立刻重复了一次，那么根据测量公理（在这里即玻恩规则），第二次测量结果和第一次相同，即如果第一次测量得到 λ_i，第二次将以 1 的概率再次得到该结果。

（4）投影测量有很多良好的性质。特别地，测量结果的期望值是容易计算的。如果对很多份完全相同且独立的系统 $|\psi\rangle$ 分别进行对可观测量 A 的测量，那么测量结果的平均值将趋近于

$$\langle A\rangle \equiv \sum_i \lambda_i p(\lambda_i) = \sum_i \lambda_i \langle\psi| E_i |\psi\rangle = \langle\psi| A |\psi\rangle \tag{1.22}$$

如果用密度矩阵的表示方法，

$$\langle \boldsymbol{A}\rangle = \mathrm{tr}(\boldsymbol{\rho A}) \tag{1.23}$$

例 1.4 写出泡利矩阵测量对应的投影算符。

解 首先，注意到 σ_z 的本征态为 $|0\rangle$ 和 $|1\rangle$，所以对应的投影算符为 $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$。记 σ_x 的本征态为

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle) \tag{1.24}$$

则对应的投影算符为 $\{|+\rangle\langle +|, |-\rangle\langle -|\}$。记 σ_y 的本征态为

$$|\pm \mathrm{i}\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm \mathrm{i}\,|1\rangle) \tag{1.25}$$

则对应的投影算符为 $\{|+\mathrm{i}\rangle\langle +\mathrm{i}|, |-\mathrm{i}\rangle\langle -\mathrm{i}|\}$。 □

1.3 幺正变换

当我们对物理系统进行观测时，实际上是我们与物理系统之间发生了相互作用，或者某种关联。而如果一个物理系统没有与外部进行相互作用，其自身状态也可以发生改变。在量子力学中，封闭系统的状态变化由下面的量子演化公理描述。

公理 1.4 (量子演化) *在量子力学中，量子态随时间的演化是由薛定谔方程决定的：*

$$\mathrm{i}\hbar\frac{\partial \Psi}{\partial t} = \hat{H}\Psi \tag{1.26}$$

其中，Ψ 是波函数，$\hat{H}$ 是哈密顿量（Hamiltonian）。

1.3.1 量子演化的数学表示

薛定谔方程描述了量子演化的微分性质。在凝聚态物理、高能物理、量子原子学等领域，这一方程在研究中至关重要。不过在量子信息中，经常考虑的是离散的问题，即我们关心一个系统在经过一段时间后变成了怎样的量子态。考虑量子态从 $t = 0$ 演化到 t，当哈密顿量不随时间变化时，有

$$\Psi(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}\Psi(t=0) \tag{1.27}$$

这里，我们不会讨论薛定谔方程的细节问题。唯一需要明确的是，这里我们假设 $\hat{H}$ 是一个厄米算符，对应我们研究的系统是一个封闭系统，如果 $\hat{H}$ 不是厄米算符，对应了开放系统，需要用后面在 2.4.2 节中介绍的方法来研究。在本书中，演化 = 操作 = 变换。如果想设计在特定系统中，例如在超导量子系统或离子阱系统中，量子操作的实现，那么薛定谔方程将是我们的重要参考。

现在将 $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ 记作一个新的算符 U。按照定义，我们发现

$$\begin{cases} UU^\dagger = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\hat{H}t} = \boldsymbol{I} \\ U^\dagger U = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\hat{H}t}\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} = \boldsymbol{I} \end{cases} \tag{1.28}$$

数学上，称具有这种性质的算符为幺正算符（unitary operator），其相应的矩阵表示为幺正矩阵。在量子计算中，这一算符的作用也被称作门（gate）。封闭量子系统的演化也被称作是幺正的。对于一个纯态 $|\psi\rangle$，幺正演化可以表示为

$$U : |\psi\rangle \mapsto U|\psi\rangle \tag{1.29}$$

如果用密度矩阵来表示量子态，则幺正演化可以表示为 $\rho \mapsto U\rho U^\dagger$。

对于复合系统而言，如果有多个幺正矩阵独立地作用在每个子系统上，那么对于整个系统，作用效果是这些矩阵的张量积。例如，考虑一个两体量子系统，其由两个子系统 A 和 B 构成。幺正矩阵 $\boldsymbol{U}$ 作用在子系统 A 上，$\boldsymbol{V}$ 作用在子系统 B 上，那么对于整个系统，量子演化可以由 $\boldsymbol{U}\otimes\boldsymbol{V}$ 描述。

1.3.2 典型幺正矩阵

现在介绍一些常用的幺正矩阵。这些矩阵将在后面的章节频繁出现，因此，熟悉它们的表示和性质将对后续的学习非常有帮助。

首先，从简单的量子比特系统出发。前面所引入的泡利矩阵不仅是厄米的，同时还是幺正的。不难验证，

$$\begin{cases} \sigma_x\sigma_x^\dagger = \sigma_x^\dagger\sigma_x = \sigma_x^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I} \\ \sigma_y\sigma_y^\dagger = \sigma_y^\dagger\sigma_y = \sigma_y^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I} \\ \sigma_z\sigma_z^\dagger = \sigma_z^\dagger\sigma_z = \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I} \end{cases} \tag{1.30}$$

所以，泡利矩阵不仅可以作为可观测量出现，也可以作为幺正变换出现。当然，变换和测量的物理含义并不相同。泡利操作 σ_x 有时候也被称为比特翻转操作，因为

$$\begin{cases} \sigma_x|0\rangle = |1\rangle \\ \sigma_x|1\rangle = |0\rangle \end{cases} \tag{1.31}$$

泡利操作 σ_z 有时候也被称为相位翻转操作，因为

$$\begin{cases} \sigma_z |0\rangle = |0\rangle \\ \sigma_z |1\rangle = -|1\rangle \end{cases} \tag{1.32}$$

Hadamard 变换是一种被广泛使用的单量子比特幺正运算[①]，其矩阵形式如下：

$$\boldsymbol{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (\sigma_x + \sigma_z) \tag{1.33}$$

由于

$$\boldsymbol{H} \sigma_x \boldsymbol{H}^\dagger = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z \tag{1.34}$$

所以这个操作实现了量子态在 X 和 Z 基之间的旋转。

不难看出，Hadamard 门既是厄米的又是幺正的。当然也存在一些非厄米的幺正门，例如 $\pi/2$ 相位门 (S) 和 $\pi/4$ 相位门 (T)，

$$S = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\pi \mathrm{i}/4} \end{pmatrix} \tag{1.35}$$

在量子线路中，一个单比特量子门可以像图 1.1那样表示。

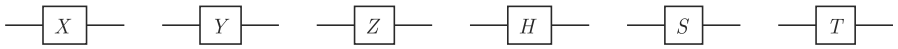

图 1.1　一些典型的量子比特门，分别是 σ_x，σ_y，σ_z，H，S 和 T

还有一个在两体量子比特系统中十分常见的操作——受控非门（controlled-NOT gate，CNOT），或者说受控 X 门（controlled-X gate）。由于 CNOT 作用在两个量子比特上，所以它是一个四维的幺正算子，在计算基矢上，CNOT 可以表示为

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.36}$$

对于 CNOT，一个更有趣也更具物理操作含义的理解方式是将其看作对两个二维量子系统的联合作用。如在 1.1.3节讨论的那样，当这个四维量子系统是由两个子

① 尽管我们都使用字母 H 来表示，但这里与物理系统的哈密顿函数 $\hat{H}$ 或希尔伯特空间的 $\mathcal{H}$ 无关，请注意区分。

系统构成时，如果我们选取计算基矢是由两个子系统各自的计算基矢构成，那么，CNOT 可以表示为

$$\mathrm{CNOT} = |0\rangle\langle 0| \otimes \boldsymbol{I} + |1\rangle\langle 1| \otimes \sigma_x \tag{1.37}$$

这可以这样来理解：当第一个量子比特处于量子态 $|0\rangle$ 时，第二个量子比特保持不变；当第一个量子比特处于量子态 $|1\rangle$ 时，对第二个量子比特作用 σ_x 操作。也就是说，第一个量子比特起到了“控制”第二个量子比特的作用。这也是受控非门这一名称的来源。

类似地，也可以定义受控相位门（controlled-phase gate）或者说受控 Z 门（controlled-Z gate，简记为 CZ 门）：

$$\mathrm{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = |0\rangle\langle 0| \otimes \boldsymbol{I} + |1\rangle\langle 1| \otimes \sigma_z \tag{1.38}$$

如果控制比特是 $|0\rangle$，第二个量子比特保持不变。否则，如果控制比特是 $|1\rangle$，这个算符会在第二个比特上作用相位翻转操作，即 σ_z。

在量子线路中，一个两体量子比特门可以用图 1.2中的形式来表示。

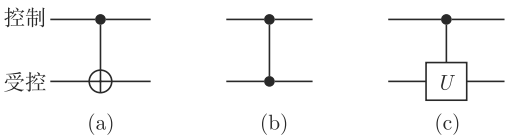

图 1.2　一些典型的两体量子比特门：（a）CNOT; (b) CZ; （c）一般的受控幺正门（control-U gate）

思考题 1.4　看上去似乎上述两种受控幺正门，CNOT 和 CZ，对控制比特没有进行任何操作。这是否意味着这些算符对于控制比特没有任何影响？

1.3.3　高维幺正变换

在 $\mathcal{H}_d$ 下的幺正算符被定义为满足 $UU^\dagger = U^\dagger U = I_d$ 的线性算符 $U \in \mathcal{L}(\mathcal{H}_d)$，其中 I_d 是 $\mathcal{H}_d$ 上的单位算符。在这一含义下，可以将量子比特系统中的 σ_x, σ_z, H 推广到高维量子比特系统。对于正交基矢 $\{|i\rangle\}_{i\in[d]}$，推广后的 $\sigma_x(x)$ 可以被定义为

$$\sigma_x(x)\,|j\rangle = |x \oplus j\rangle \tag{1.39}$$

其中，$x \oplus j \equiv x + j \mod d$。推广后的 $\sigma_z(z)$ 可以被定义为

$$\sigma_z(z)\,|j\rangle = \mathrm{e}^{2\pi \mathrm{i} z j/d}\,|j\rangle \tag{1.40}$$

为了防止混淆，将虚数单位表示为罗马体 i。在这一小节中，采用如下标记：

$$\begin{cases} X = \sigma_x(1) \\ Z = \sigma_z(1) \end{cases} \tag{1.41}$$

那么，不难看出

$$\begin{cases} \sigma_x(x) = X^x \\ \sigma_z(z) = Z^z \end{cases} \tag{1.42}$$

其中，$x, z \in [d]$。算符 $\{X^x Z^z\}_{x,z\in[d]}$ 被称为 Heisenberg-Weyl 算符，有时也被直接称作 Weyl 算符。

现在，让我们看一下 Weyl 算子的本征态和本征值。为简单起见，将使用狄拉克符号来表示它们的表示矩阵。首先，可以写出 X 和 Z 对应的算符：

$$\begin{cases} X = \sum_{j=0}^{d-1} |j \oplus 1\rangle\langle j| \\ Z = \sum_{j=0}^{d-1} \mathrm{e}^{2\pi \mathrm{i} j/d} |j\rangle\langle j| \end{cases} \tag{1.43}$$

类似地，从式(1.39)和式(1.40)不难得出，X^x, Z^z 的矩阵形式如下所示：

$$\begin{cases} X^x = \sum_{j=0}^{d-1} |j \oplus x\rangle\langle j| \\ Z^z = \sum_{j=0}^{d-1} \mathrm{e}^{2\pi \mathrm{i} z j/d} |j\rangle\langle j| \end{cases} \tag{1.44}$$

显然，这些算子并不是厄米的，但它们是幺正的。

Z 和 Z^z 的本征态和本征值非常直接就可以写出来：

$$\begin{cases} Z|j\rangle = \mathrm{e}^{2\pi \mathrm{i} j/d} |j\rangle \\ Z^z|j\rangle = \mathrm{e}^{2\pi \mathrm{i} j z/d} |j\rangle \end{cases} \tag{1.45}$$

也就是说，Z 的本征态组成了计算基矢。当 z 和 d 有一个素数公因子时，Z^z 的本征态是简并的。

X 的本征态有点复杂，这里用 $\{|\tilde{j}\rangle\}_{j\in[d]}$ 表示。获取 $|\tilde{j}\rangle$ 具体表达式的标准方法是在基 $\{|j\rangle\}_{j\in[d]}$ 中写下 X 的矩阵元素，并对角化为 $X = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\dagger$，其中 $\boldsymbol{U}$

是幺正的，$\boldsymbol{\Lambda}$ 是对角的。那么，$X\boldsymbol{U}=\boldsymbol{U}\boldsymbol{\Lambda}$ 意味着 $\boldsymbol{U}$ 的所有列都是 X 的本征态。我们把这个标准计算过程留作习题 1.9。

在这里，我们就像以前的物理学家一样，根据量子比特的结果大胆猜测 $\{|\tilde{j}\rangle\}_{j\in[d]}$ 的表达式，然后对其进行验证。我们猜测 X 的本征态，或者说 $\{|\tilde{j}\rangle\}_{j\in[d]}$ 为

$$|\tilde{j}\rangle=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}\,|k\rangle \tag{1.46}$$

因为对于任意 $j\in[d]$，

$$\begin{aligned}X|\tilde{j}\rangle&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}X\,|k\rangle\\&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}\,|k\oplus 1\rangle\\&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}j(k-1)/d}\,|k\rangle\\&=\mathrm{e}^{-2\pi\mathrm{i}j/d}\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}\,|k\rangle\\&=\mathrm{e}^{-2\pi\mathrm{i}j/d}\,|\tilde{j}\rangle\end{aligned} \tag{1.47}$$

在第三个等于号我们用了 $\mathrm{e}^{2\pi\mathrm{i}j(d-1)/d}=\mathrm{e}^{-2\pi\mathrm{i}j/d}$ 这一结论。类似地，可以证明 $|\tilde{j}\rangle$ 是 X^x 对应本征值 $\mathrm{e}^{-2\pi\mathrm{i}jx/d}$ 的本征态。综上，

$$\begin{cases}X|\tilde{j}\rangle=\mathrm{e}^{-2\pi\mathrm{i}j/d}\,|\tilde{j}\rangle\\X^x|\tilde{j}\rangle=\mathrm{e}^{-2\pi\mathrm{i}jx/d}\,|\tilde{j}\rangle\end{cases} \tag{1.48}$$

有趣的是，我们可以计算

$$\begin{aligned}Z|\tilde{j}\rangle&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}Z\,|k\rangle\\&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}jk/d}\mathrm{e}^{2\pi\mathrm{i}k/d}\,|k\rangle\\&=\frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\mathrm{e}^{2\pi\mathrm{i}(j+1)k/d}\,|k\rangle\\&=\left|\widetilde{j+1}\right\rangle\end{aligned} \tag{1.49}$$

从而得出：

$$Z=\sum_{j=0}^{d-1}\left|\widetilde{j\oplus 1}\right\rangle\!\left\langle\tilde{j}\right| \tag{1.50}$$

这与式 (1.43)的形式十分相似。

思考题 1.5 证明下式

$$X^xZ^z=\mathrm{e}^{2\pi\mathrm{i}xz/d}Z^zX^x \tag{1.51}$$

同样地，可以推广 Hadamard 算子得到量子傅里叶变换。也就是说，量子比特系统上的 Hadamard 算子 H 是作用于希尔伯特空间上的量子傅里叶变换算子的一个特例。

定义 1.1 (量子傅里叶变换，quantum Fourier transformation) 将 Z,X 的本征态分别记为 $\{|j\rangle\}_{j\in[d]},\{|\tilde{j}\rangle\}_{j\in[d]}$。傅里叶变换算子的定义如下：

$$F\equiv\sum_{j=0}^{d-1}\left|\tilde{j}\right\rangle\!\left\langle j\right| \tag{1.52}$$

这就是量子傅里叶变换，就像量子比特情况中 Hadamard 门那样，交换了 X 和 Z 的基矢态：

$$F\left|j\right\rangle=\left|\tilde{j}\right\rangle \tag{1.53}$$

例 1.5 证明 F 可以旋转算子 X 和 Z，

$$\begin{cases}F\sigma_x(x)F^\dagger=\sigma_z(x)\\ F\sigma_z(z)F^\dagger=\sigma_x(-z)\end{cases} \tag{1.54}$$

解 首先，我们注意到

$$\begin{aligned}XFZ&=X\left(\sum_{j=0}^{d-1}\left|\tilde{j}\right\rangle\!\left\langle j\right|\right)Z\\&=\sum_{j=0}^{d-1}\mathrm{e}^{-2\pi\mathrm{i}j/d}\left|\tilde{j}\right\rangle\!\left\langle j\right|\mathrm{e}^{2\pi\mathrm{i}j/d}\\&=F\end{aligned} \tag{1.55}$$

第二个等号我们使用了式(1.45)和式(1.48)，$(\langle j|\,Z)^\dagger=Z^\dagger\left|j\right\rangle=\mathrm{e}^{-2\pi\mathrm{i}j/d}\left|j\right\rangle$。那么可以得到：

$$FZF^\dagger=X^\dagger \tag{1.56}$$

然后，就可以很快得到 Weyl 算子之间的变换：

$$F\sigma_z(z)F^\dagger = (FZF^\dagger)^z = (X^\dagger)^z = X^{-z} = \sigma_x(-z) \tag{1.57}$$

类似地，由式(1.50)，有

$$\begin{aligned} Z^\dagger FX &= Z^\dagger \left(\sum_{j=0}^{d-1} \left|\tilde{j}\right\rangle\!\left\langle j\right| \right) X \\ &= \sum_{j=0}^{d-1} \left|\widetilde{j-1}\right\rangle\!\left\langle j-1\right| \\ &= F \end{aligned} \tag{1.58}$$

那么，就可以得到另一边的变换：

$$F\sigma_x(x)F^\dagger = (FXF^\dagger)^x = (Z)^x = X^z = \sigma_z(x) \tag{1.59}$$

□

除了空间维数从 2 变为了 d 外，高维量子系统的测量定义与量子比特相同。然而，在一般的 d 系统中 $\sigma_x(1), \sigma_z(1)$ 并不是厄米的（请证明这一点!），这导致了我们不能像量子比特系统中测量 σ_x, σ_z 一样直接测它们。但是对于 $\sigma_x(1)$ 的本征态 $\{\left|\tilde{j}\right\rangle\}_j$ 和 $\sigma_z(1)$ 的本征态 $\{|j\rangle\}_j$，可以定义如下观测量：

$$\begin{cases} M_{\sigma_x(1)} \equiv \sum_{j=0}^{d-1} j \left|\tilde{j}\right\rangle\!\left\langle\tilde{j}\right| \\ M_{\sigma_z(1)} \equiv \sum_{j=0}^{d-1} j \left|j\right\rangle\!\left\langle j\right| \end{cases} \tag{1.60}$$

例 1.6 计算 Z, X 的本征态 $\{|j\rangle\}_{j\in[d]}, \{\tilde{j}\}_{j\in[d]}$ 对这两个可观察量的期望值。

解 首先，证明 $\left|\left\langle i|\tilde{j}\right\rangle\right|$ 与 i 和 j 无关：

$$\left|\left\langle i|\tilde{j}\right\rangle\right| = \left|\langle i| X \left|\tilde{j}\right\rangle\right| = \left|(\langle i| X) \left|\tilde{j}\right\rangle\right| = \left|\left\langle i \oplus -1|\tilde{j}\right\rangle\right| \tag{1.61}$$

同样，对于不同的 j，可以证明这个关系，从而可以得到 $\left|\left\langle i|\tilde{j}\right\rangle\right|^2 = \frac{1}{d}, \forall i, j$。这样就很容易验证

$$\begin{cases} \left\langle\tilde{j}\right| M_{\sigma_x(1)} \left|\tilde{j}\right\rangle = j \\ \left\langle\tilde{j}\right| M_{\sigma_z(1)} \left|\tilde{j}\right\rangle = \frac{d-1}{2} \\ \langle j| M_{\sigma_x(1)} |j\rangle = \frac{d-1}{2} \\ \langle j| M_{\sigma_z(1)} |j\rangle = j \end{cases} \tag{1.62}$$

□

1.4 确定量子态

1.4.1 布洛赫球面

从前面小节我们知道，一个量子比特纯态可以表示为 $a|0\rangle + b|1\rangle$。当 a, b 为实数时，这个态可以由二维平面上的射线表示。那么如果 a, b 是一般的复数，是否有相应的几何表示方法呢？为了讨论这一问题，首先需要确定一个量子比特纯态有几个独立的自由度。由于 $|a|^2 + |b|^2$ 的取值不影响系统的量子态，可以采用归一化表示方法将其固定为 1。同时，系统的全局相位不影响量子态，$\mathrm{e}^{\mathrm{i}\theta}|\phi\rangle$ 与 $|\phi\rangle$ 表示同一个系统。因此，一个量子比特纯态有 2 个独立的自由度。归一化的纯态量子比特可以表示为下面的形式：

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \mathrm{e}^{\mathrm{i}\varphi}\sin\frac{\theta}{2}|1\rangle \tag{1.63}$$

可以将纯态表示为在单位球面 $(r = 1, \theta, \varphi)$ 上的一个点，其中 θ 和 φ 分别是量子态与 z 轴的夹角，以及其在 xy 平面上投影与 x 轴的夹角。称该单位球面为布洛赫球面，如图 1.3所示。

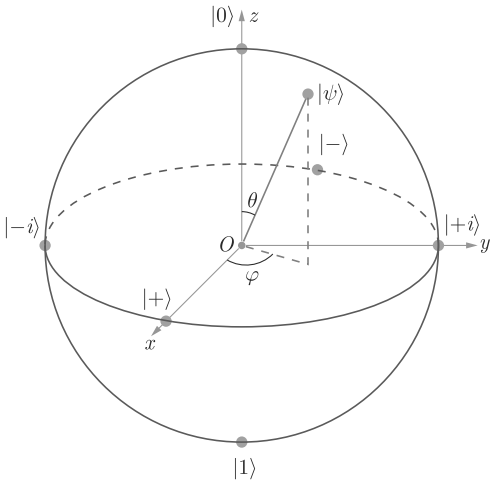

图 1.3 布洛赫球面

由式(1.63)可以看出，$\{|0\rangle, |1\rangle\}$ 是 z 轴上的两个端点量子态，$\{|+\rangle, |-\rangle\}$ 是 x 轴上的两个端点量子态，$\{|+\mathrm{i}\rangle, |-\mathrm{i}\rangle\}$ 是 y 轴上的两个端点量子态：

$$\begin{cases} |\pm\rangle = \dfrac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle) \\ |\pm\mathrm{i}\rangle = \dfrac{1}{\sqrt{2}}(|0\rangle \pm \mathrm{i}\,|1\rangle) \end{cases} \tag{1.64}$$

对比式 (1.63) 和式 (1.64) 可以看出，这些端点的量子态正好对应 3 个泡利矩阵的本征态。

例 1.7 (布洛赫球面上的纯态)　在文献中，有时会将量子比特用布洛赫球上的角度或者坐标轴表示。

（1）写下 x，y，z 三根轴的端点的量子态：$|x+\rangle$，$|x-\rangle$，$|y+\rangle$，$|y-\rangle$，$|z+\rangle$，$|z-\rangle$；

（2）按照式(1.63)的形式，考虑 $\theta=90°$，用 φ 表示量子态为 $|\varphi\rangle$，写下量子态 $|+45°\rangle$，$|-45°\rangle$，这两个态是否相互正交？

（3）在布洛赫球面表示方法下，一对相互正交的量子态应该如何表示？

解　（1）将相应的立体角代入式(1.63)，可以得到：$|x+\rangle=|+\rangle$，$|x-\rangle=|-\rangle$，$|y+\rangle=|+\mathrm{i}\rangle$，$|y-\rangle=|-\mathrm{i}\rangle$，$|z+\rangle=|0\rangle$，$|z-\rangle=|1\rangle$，其中 $|\pm\rangle$ 和 $|\pm\mathrm{i}\rangle$ 由式(1.64)给出。

（2）

$$\begin{cases}|+45°\rangle=\cos\dfrac{\pi}{4}|0\rangle+\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}\sin\dfrac{\pi}{4}|1\rangle=\dfrac{\sqrt{2}}{2}\left(|0\rangle+\dfrac{1+\mathrm{i}}{\sqrt{2}}|1\rangle\right)\\ |-45°\rangle=\cos\dfrac{\pi}{4}|0\rangle+\mathrm{e}^{-\mathrm{i}\frac{\pi}{4}}\sin\dfrac{\pi}{4}|1\rangle=\dfrac{\sqrt{2}}{2}\left(|0\rangle+\dfrac{1-\mathrm{i}}{\sqrt{2}}|1\rangle\right)\end{cases}\tag{1.65}$$

两个量子态不正交，因为

$$\langle+45°|-45°\rangle=\frac{1}{2}\left(1+\frac{1-\mathrm{i}}{\sqrt{2}}\frac{1-\mathrm{i}}{\sqrt{2}}\right)\neq 0\tag{1.66}$$

注意，在用偏振构建量子比特的光学系统中，$|\pm45°\rangle$ 表示的含义会与本题不同，这种情况下，角度通常指偏振的角度，读者需要注意区分。

（3）两个正交的量子态由同一条直线上指向相反方向的两条射线表示。这也从侧面说明了布洛赫球和我们常用的希尔伯特空间有很大的区别。　□

通过下面这个简单的问题，可以看到幺正矩阵作用在量子态上时，具有旋转的含义。我们以泡利操作作为例子。

例 1.8　考虑一个量子比特系统，定义基矢为 $\{|0\rangle,|1\rangle\}$。考虑一个一般的量子态 $|\psi\rangle=a|0\rangle+b|1\rangle$，其中 $|a|^2+|b|^2=1$。计算对这一量子态作用 σ_z 的演化结果 $\sigma_z|\psi\rangle$，并用布洛赫球的几何方式予以表示。特别地，当 $|\psi\rangle$ 是泡利矩阵 $\sigma_x,\sigma_y,\sigma_z$ 的本征态时，结果是怎样的？

解

$$\begin{aligned}\sigma_z|\psi\rangle&=a\sigma_z|0\rangle+b\sigma_z|1\rangle\\&=a|0\rangle-b|1\rangle\end{aligned}\tag{1.67}$$

对比式(1.63)，可以看出 σ_z 将 φ 变成了 $\varphi+\pi$，即在布洛赫球上绕 Z 轴旋转了一个角度 π。具体对于各泡利矩阵的本征态的操作如何，请读者在布洛赫球面上自行给出。 □

思考题 1.6 对于高维量子态，我们怎样用类似于布洛赫球面的几何图像予以描述？

1.4.2 量子比特基

从布洛赫球面的讨论出发，我们来考虑更为一般性的数学描述。前面提到的泡利矩阵 $\{\sigma_x,\sigma_y,\sigma_z\}$ 构成二维零迹厄米矩阵空间的一组正交基，即对于任意的一个该空间中的矩阵 $\boldsymbol{M}$，可以将其线性展开为泡利矩阵的叠加。为了表示方便，通常将三个泡利矩阵表示成向量形式 $\sigma=(\sigma_x,\sigma_y,\sigma_z)$，注意这个“向量”的元素是矩阵。再加上二维单位矩阵 $\boldsymbol{I}$，为方便起见，这里记为

$$\sigma_0=\begin{pmatrix}1&0\\0&1\end{pmatrix} \tag{1.68}$$

可以将二维矩阵空间中的任一元素表示为 $\{\sigma_0,\sigma_x,\sigma_y,\sigma_z\}$ 的线性展开：

$$\begin{aligned}M&=\sum_{i\in\{0,x,y,z\}}r_i\sigma_i\\&=r_0\sigma_0+r\cdot\sigma\end{aligned} \tag{1.69}$$

这里，$r\cdot\sigma=r_x\sigma_x+r_y\sigma_y+r_z\sigma_z$ 类似于普通的向量点乘运算。如果 M 是厄米的，则 r 为三维实向量。在我们的讨论中，主要关注量子态，即迹为 1 的半正定矩阵，这样 $r_0=1/2$。

我们称泡利矩阵 σ_x，σ_y，σ_z 各自的本征向量对应于 X,Y,Z 基矢，$\{|+\rangle,|-\rangle\}$，$\{|+\mathrm{i}\rangle,|-\mathrm{i}\rangle\}$，$\{|0\rangle,|1\rangle\}$。这些态的定义和前面小节中布洛赫球面涉及的量子态一一对应，见式(1.64)。

思考题 1.7 证明一个量子比特密度矩阵用式 (1.69) 展开得到的 $\boldsymbol{r}$ 正好是该量子态在前面小节提到的布洛赫球表示的坐标。

如果从 X,Y,Z 这三组基矢，$\{|+\rangle,|-\rangle\}$，$\{|+\mathrm{i}\rangle,|-\mathrm{i}\rangle\}$，$\{|0\rangle,|1\rangle\}$，取出不同基矢的态求内积，其模平方均为 $1/2$，

$$|\langle\pm|0\rangle|^2=|\langle\pm\mathrm{i}|0\rangle|^2=|\langle\pm\mathrm{i}|\pm\rangle|^2=\frac{1}{2} \tag{1.70}$$

我们称这些基矢相互是无偏的。更一般地，无偏基矢（mutually unbiased bases, MUB）的定义如下。

定义 1.2 (无偏基矢) d 维希尔伯特空间 $\mathcal{H}_d$ 中的两组相互无偏基矢是正交归一的两组完备向量 $\{\phi_0,\phi_1,\cdots,\phi_{d-1}\}$ 和 $\{\psi_0,\psi_1,\cdots,\psi_{d-1}\}$，使得从两组基矢中任取一个向量，它们内积模长的平方等于系统维度的倒数，$\forall i,j$，

$$|\langle\psi_i|\phi_j\rangle|^2=\frac{1}{d} \tag{1.71}$$

思考题 1.8 对于 $d\geqslant 2$，给定希尔伯特空间 $\mathcal{H}_d$，

（1）应该如何定义类似于泡利矩阵的高维量子空间的基？提示：请思考泡利矩阵具有哪些良好的数学性质，参考 1.3.3节。

（2）对于这一高维希尔伯特空间，是否一定有相互无偏基矢？

（3）如果存在相互无偏基矢，一共有多少个？

对于最后一个问题，最一般的情形仍然是待解决的数学难题，但对于某些特定的 d，比如当其为质数幂次时，问题已经得到了完整的解答。

1.4.3 量子层析术

如果想确切地知道 $|\psi\rangle$ 究竟是怎样一个量子态，要怎么做呢？例如，在实验室里，实验员试图确定自己究竟制备了怎样的量子系统。我们不妨从简单的情形入手，假设实验员制备了很多份相同的纯量子态 $|\psi\rangle$，把纯态 $|\psi\rangle$ 对应的密度矩阵记为 ψ，并且 $\psi=|\psi\rangle\langle\psi|$。

回想一下，在式(1.63)中，纯态可以由两个角度 ϕ 和 θ 来表示。这一灵感来自对应的布洛赫球上的点，如图 1.3所表示的量子比特，所以我们的主要任务就是找出量子态在布洛赫球上对应的位置坐标 (x,y,z) 或 (θ,ϕ)，x,y,z 是在笛卡儿坐标系下对应的坐标，而其对应的标准正交基的 X,Y,Z 是由泡利矩阵 $\sigma_x,\sigma_y,\sigma_z$ 决定的。通过简单的 PVM 计算（请核对它们的正确性!），有

$$\begin{cases} x=\langle\sigma_x\rangle=\langle\psi|\,\sigma_x\,|\psi\rangle=\mathrm{tr}(\sigma_x\psi) \\ y=\langle\sigma_y\rangle=\langle\psi|\,\sigma_y\,|\psi\rangle=\mathrm{tr}(\sigma_y\psi) \\ z=\langle\sigma_z\rangle=\langle\psi|\,\sigma_z\,|\psi\rangle=\mathrm{tr}(\sigma_z\psi) \end{cases} \tag{1.72}$$

因此，想要知道 (x,y,z)，只需要测量出 $\langle\sigma_x\rangle,\langle\sigma_y\rangle,\langle\sigma_z\rangle$ 这些期望值。于是通过这三个测量，便可以对 $|\psi\rangle$ 进行“层析”。

对于纯量子态，除了向量表示方法，正如在 1.4.2 节讨论的，其密度矩阵 $\boldsymbol{\rho}$ 应该可以用泡利矩阵 $\{\sigma_0,\sigma_x,\sigma_y,\sigma_z\}$ 来表示。由于 $\mathrm{tr}(\boldsymbol{\rho})=1$，可以将 $\boldsymbol{\rho}$ 扩展表示成

$$\boldsymbol{\rho}=\frac{1}{2}(\sigma_0+\boldsymbol{P}\cdot\boldsymbol{\sigma})$$

$$= \frac{1}{2}\begin{pmatrix} 1+P_z & P_x - \mathrm{i}P_y \\ P_x + \mathrm{i}P_y & 1-P_z \end{pmatrix} \tag{1.73}$$

其中，$\boldsymbol{P} = (P_x, P_y, P_z)$ 是一个向量并且 $|\boldsymbol{P}| \leqslant 1$。当 $|\boldsymbol{P}| = 1$ 时，式 (1.73)可以表示在二维希尔伯特空间中的一个纯态。正如式 (1.73)给出的那样，Z 基矢的测量提供了关于 P_z 的信息。类似地，X, Y 基矢的测量提供了关于 P_x, P_y 的信息。因此 X，Y 和 Z 基矢的测量在对于一个量子比特的层析上是完整的。或者也可以这么说，

$$\boldsymbol{\rho} = \frac{1}{2}[\mathrm{tr}(\boldsymbol{\rho})\sigma_0 + \mathrm{tr}(\sigma_x\boldsymbol{\rho})\sigma_x + \mathrm{tr}(\sigma_y\boldsymbol{\rho})\sigma_y + \mathrm{tr}(\sigma_z\boldsymbol{\rho})\sigma_z] \tag{1.74}$$

对于高维量子比特，也可以采用类似的方式进行量子态层析，进而确定量子态的具体形式。考虑一个 2^n 维的纯量子态 $|\psi\rangle$，可以将其看作由 n 个量子比特组成的复合系统。对于这一空间上的所有厄米矩阵，可以将其用子系统上的泡利矩阵进行展开。对于用密度矩阵表示的量子态 ρ，

$$\rho = \sum_{i_1, i_2, \cdots, i_n} P_{i_1 i_2 \cdots i_n} \bigotimes_{t=1}^{n} \sigma_{i_t} \tag{1.75}$$

其中，$\sigma_{i_t} \in \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$，下标 t 表示这一泡利矩阵是作用在第 t 个子系统上的。由于量子态要满足迹为 1 的条件，一般地，为了确定一个 2^n 维量子态的密度矩阵，其中有 $4^n - 1$ 个参数。

思考题 1.9 读到这里，你可能会好奇，为什么在向量表示之外，要用密度矩阵来重新讨论量子态层析的问题。对此，请考虑下面的问题。

（1）在二维量子比特情形，对于式(1.73)，如果测量结果 (x, y, z) 落在了布洛赫球（体）的内部，而不是表面上呢？这个点代表了怎样的物理意义呢？

（2）对于高维量子比特，假设所研究的 2^n 维量子态可以写成 $\rho = |\psi\rangle\langle\psi|$ 的形式，即量子态是纯态。对于这一情况，式(1.75)这一展开式中一般有多少个自由参数？如果这一情况的自由参数数量小于 $4^n - 1$，对于缺少的那些参数，你有什么想法？

第 2 章内容剧透：对 “混合” 的量子态，不能简单地用向量 $|\psi\rangle$ 来表示，但是仍然可以用密度矩阵 $\boldsymbol{\rho}$ 表示。这个矩阵是一个正半定的对易化厄米算符。我们将在第 2 章讨论 “混合” 量子态。

量子态层析术（state tomography）是通过适当的测量来重建量子系统的量子态的过程。这里的量子态可以是纯的（可以用矢量表示），也可以是混的（只能用密度矩阵表示)。如果一组测量能够唯一地识别量子态，就称为层析完整（tomographically complete）。也就是说，测量的结果能够提供关于量子态的所有信息。在物理学中，这组测量对应于系统标定。量子态层析术背后蕴含的一般原理是，通过对相同密度矩阵描述的量子系统重复执行许多不同的测量，可以用频率来推断概率，这些概率则与玻恩定则相结合，以确定最符合测量值的密度矩阵。

另一个与之相关的概念是量子过程层析术（quantum process tomography），一些已知的量子态被用来探测一个量子过程，以找出如何具体地描述这个过程。类似地，量子测量断层术的目的是找出进行的测量是什么。

1.5 量子信息守恒

1982 年，Nick Herbert 发表了一篇文章，提出使用量子纠缠实现超光速通信的设想。这一明显有悖于相对论的文章很快引起了物理学界的广泛关注。同年，William Wootters 与 Wojciech Zurek 及 Dennis Dieks 分别发表了题为 *A Single Quantum Cannot be Cloned* 和 *Communication by EPR devices* 的文章，指出量子态具有不可克隆的特性，而所谓的超光速通信违背了这一性质，因此是错误的[①]。这一系列讨论促使人们更加认真地思考量子力学与信息论之间的关系[②]，也在很大程度上推动了包括利用范畴论等数学工具理解量子力学理论本身的研究。在这一节，我们讨论量子力学中的信息守恒相关结论。除了不可克隆定理外，我们还会介绍它的“时间反演”版本——量子不可删除定理，以及关于通用量子非门存在性等结论。

1.5.1 量子不可克隆定理

量子不可克隆定理有许多种等价的表述方式。在这里，我们用类似于文献 [6] 的语言予以描述。

定理 1.1 (量子不可克隆定理，no-cloning theorem)　不可能构造一个能够确定性地复制任意量子态且不改变原始量子态的克隆机器。

假设存在一个“完美”的克隆机器，即存在幺正演化 U，满足 $\forall|\psi\rangle$，

$$|\psi\rangle|0\rangle \to U(|\psi\rangle|0\rangle) = |\psi\rangle|\psi\rangle \tag{1.76}$$

那么对于任意两个量子态 $|\psi_1\rangle$ 和 $|\psi_2\rangle$，有

$$\begin{cases} U(|\psi_1\rangle|0\rangle) = |\psi_1\rangle|\psi_1\rangle \\ U(|\psi_2\rangle|0\rangle) = |\psi_2\rangle|\psi_2\rangle \end{cases} \tag{1.77}$$

将这两个式子做内积，有

---

① 历史上，除了这两篇发表的工作，量子不可克隆定理在此之前已经被注意到。在对 Herbert 文章的审稿过程中，审稿人 Giancarlo Ghirardi 已经发现了这一问题。而在更早的 1970 年，James Park 在解释量子测量对物理系统状态的扰动时，从不同的角度说明了量子不可克隆的性质。

② 需要注意的是，直到今天，Herbert 错误的“超光速量子通信”依然被很多科普文章所采用！在本书后续章节，我们将说明一个正确的通信协议中，究竟何为信息，通信者又应该如何利用量子态和量子操作完成信息的传递。

$$\begin{cases}(\langle 0|\langle\psi_1|U^*)(U|\psi_2\rangle|0\rangle)=(\langle 0|\langle\psi_1|)(|\psi_2\rangle|0\rangle)=\langle\psi_1|\psi_2\rangle\\(\langle 0|\langle\psi_1|U^*)(U|\psi_2\rangle|0\rangle)=(\langle\psi_1|\langle\psi_1|)(|\psi_2\rangle|\psi_2\rangle)=\langle\psi_1|\psi_2\rangle^2\end{cases}\tag{1.78}$$

那么可以得到 $\langle\psi_1|\psi_2\rangle=\langle\psi_1|\psi_2\rangle^2$，所以 $\langle\psi_1|\psi_2\rangle$ 只能等于 0 或 1。因此 $|\psi_1\rangle$ 和 $|\psi_2\rangle$ 要么相等，要么互相正交，这与式(1.76)中两个任意态相矛盾。由此导出了不可克隆定理。

思考题 1.10 如果允许克隆后的系统多一个全局相位，即 $U(|\psi\rangle|0\rangle)=\mathrm{e}^{i\alpha(\psi)}|\psi\rangle|\psi\rangle$，那么不可克隆原理是否还成立？

思考题 1.11 不可克隆定理是否与海森堡不确定性关系相符？

思考题 1.12 不可克隆定理排除了适用于所有量子态的“通用”克隆机的可能性；或者说，在我们的证明过程中，我们事先不知道想要复制的量子态的具体形式。但是，考虑一下，是否有可能使用一个幺正演化来克隆某些特定的状态？这里，让我们考虑两个简单的情形。

（1）假设已经确定了一个量子态 $|\psi\rangle$。尝试构造一个幺正算符 U，使其满足 $U(|\psi\rangle|0\rangle)=|\psi\rangle|\psi\rangle$。

（2）在 $|\psi\rangle$ 外，同时考虑与其正交的一个量子态 $|\psi^\perp\rangle$，即 $\langle\psi|\psi^\perp\rangle=0$。尝试构造一个幺正算符 U，使其满足 $U(|\psi\rangle|0\rangle)=|\psi\rangle|\psi\rangle$，$U\left(|\psi^\perp\rangle|0\rangle\right)=|\psi^\perp\rangle|\psi^\perp\rangle$。

思考题 1.13 如果放宽不可克隆定理的叙述，会得到一些不同的结果吗？请尝试考虑下面的一些可能性。

（1）对于任意的量子态，能否按照一个给定的成功概率将其克隆？

（2）对于任意的量子态，如果允许克隆机器对其稍微有些破坏，能否克隆出一份量子态？

（3）对于一个未知的量子态 $|\psi\rangle$，假设有很多份相同的拷贝，即 $|\psi\rangle^{\otimes n}$。在这种情况下，能否找到一个克隆机器，在不破坏这些已有拷贝的情况下，多复制出一份完美的拷贝？如果允许这样的复制有可能失败，或者对原有拷贝有一些破坏呢？

为了回答这些问题，你可能需要首先考虑“成功概率”、对量子态“稍微有些破坏”的含义。前面两个问题的答案又被称为“不完美量子克隆”——当我们不再要求完美的量子态复制时，量子克隆又重新变为可能，不过这样做的成功概率存在着一个上限。第三个问题又被称为密码学意义的量子克隆问题，与量子复杂性理论密切相关。这几个问题在量子密码学中都得到了充分的研究，并被巧妙地用于量子密码分析之中，感兴趣的读者可以参考文献 [7]。

在实际物理系统中，由于环境噪声的影响，量子态通常十分脆弱。在经典信息处理中，会通过复制信息的方式构造纠错码，通过探测和纠正错误，从而保护信息；但现在，由于量子不可克隆定理的限制，不能通过复制量子态来抵抗噪声[①]。

① 虽然有这样的根本限制，我们是否就无法构造量子纠错码了呢？答案并非是悲观的，虽然不可克隆定理限制了很多纠错的方法，但事实上，纠错不一定需要复制！

这让我们产生这样的一种印象：量子态很难被保存下来。在实验上，高品质的量子存储确实是一件非常具有挑战性的事情。

1.5.2 量子不可删除定理

现在来考虑量子克隆的一个“反问题”：对于任意一个量子态的两份拷贝，能否通过量子操作删除其中一份呢？或许有些令人惊讶，这一问题的答案也是否定的：量子态还具有某种“鲁棒性”，让我们不能没有代价地将其消除。这便是量子不可删除定理（the no-deletion theorem），其具体描述如下。

定理 1.2 给定任意量子态的两个副本，不可能构造一个删除机器，能够删除其中一份量子态且其随后的状态与被删除的量子态无关。

一个更数学的表述是，对于任意的未知量子态 $|\psi\rangle = a|0\rangle + b|1\rangle$，不存在一个通用的线性等距变换 U，使得

$$U|\psi\rangle|\psi\rangle|A\rangle = |\psi\rangle|0\rangle|A'\rangle \tag{1.79}$$

其中描述量子删除机器的辅助量子比特最终状态 $|A'\rangle$ 独立于被删除的量子态 $|\psi\rangle$。

假设

$$\begin{cases} U|0\rangle|0\rangle|A\rangle = |0\rangle|0\rangle|A_0\rangle \\ U|1\rangle|1\rangle|A\rangle = |1\rangle|0\rangle|A_1\rangle \\ U(|1\rangle|0\rangle|A\rangle + |0\rangle|1\rangle|A\rangle) = |\Psi\rangle \end{cases} \tag{1.80}$$

对于一个未知的量子态 $|\psi\rangle$，设 $|\psi\rangle = a|0\rangle + b|1\rangle$，

$$\begin{aligned} U|\psi\rangle|\psi\rangle|A\rangle &= U(a^2|0\rangle|0\rangle|A\rangle + b^2|1\rangle|1\rangle|A\rangle + ab|1\rangle|0\rangle|A\rangle + ab|0\rangle|1\rangle|A\rangle) \\ &= a^2|0\rangle|0\rangle|A_0\rangle + b^2|1\rangle|0\rangle|A_1\rangle + ab|\Psi\rangle \end{aligned} \tag{1.81}$$

同时由式(1.79)，有

$$U|\psi\rangle|\psi\rangle|A\rangle = |\psi\rangle|0\rangle|A'\rangle = a|0\rangle|0\rangle|A'\rangle + b|1\rangle|0\rangle|A'\rangle \tag{1.82}$$

对于任意的 a, b，式(1.81)与式(1.82)相等，因此有

$$\begin{cases} |\Psi\rangle = |0\rangle|0\rangle|A_1\rangle + |1\rangle|0\rangle|A_0\rangle \\ |A'\rangle = a|A_0\rangle + b|A_1\rangle \end{cases} \tag{1.83}$$

另外，因为 $|a|^2 + |b|^2 = 1$，由对量子态 $|A'\rangle$ 的归一化要求可知，$|A_0\rangle$ 和 $|A_1\rangle$ 相互正交。这也就意味着尽管 $|\psi\rangle$ 从我们所关心的系统中被删除了，但它的信息却被转移到删除机器的末态 $|A'\rangle$ 中。

思考题 1.14 不可克隆加上不可删除是否构成信息的守恒定律？

如果将量子态看作量子信息的一种载体，量子不可克隆定理和量子不可删除定理说明，我们不能没有任何代价地凭空增加或彻底消除这种资源。现在我们所说的“信息”，指的是由态向量描述的量子态，也就是纯态；到目前为止，也许你会对公理 1.1 感到十分自然——虽然介绍了密度矩阵的概念，但本章的所有推导仅用态向量就可以完成。不过在第 2 章将进一步加深这一概念：纯态量子态代表着对系统的完全刻画。

不可克隆定理和不可删除定理与量子引力和黑洞的前沿理论有着有趣的关系。关于这个话题的最著名的争论之一就是黑洞信息悖论。这一悖论最早是由 Stephen Hawking 和 Kip Thorne 提出的，他们坚信被黑洞吞噬的信息永远不会被外界所知。即使黑洞蒸发并完全消失，被吞噬的信息也永远不会被揭示出来。但是 John Preskill 坚信，在正确的量子引力理论中，一定也会找到一种使信息由蒸发的黑洞释放出来的机制。因此，Preskill 和 Hawking 及 Thorne 打了一个赌：当一个初始的纯量子态经过引力坍塌形成黑洞时，黑洞蒸发结束时的最终状态将会是一个纯量子态。

有关这一悖论的辩论也在不断更新。Hawking 在 2004 年承认赌输了（而 Thorne 没有）；近二十余年，为了构造出 Preskill 所坚信的“正确理论”，大量的量子引力理论被提出。但在今天，黑洞信息悖论仍然是理论物理学中一个悬而未决的问题。

思考题 1.15 你对黑洞信息悖论有什么看法？

1.5.3 不存在通用的非门

在量子不可克隆定理和不可删除定理之后，人们还发现了一系列与信息的守恒性质相关的结论。在这里，简单介绍一个简单且广为人知的结论：量子力学中不存在通用的非门（not-gate）。一个通用的非门 U 要求对于任意的输入量子态 $|\phi\rangle$，有 $\langle\phi|U|\phi\rangle = 0$。通过前面介绍的投影测量，可以准确地区分 $|\psi\rangle$ 和 $U|\psi\rangle$ 两个量子态，因此在区分物理状态的意义上，U 翻转了量子态。

下面证明这样的 U 并不存在。在 d 维希尔伯特空间中，因为 $\boldsymbol{U}$ 是一个幺正矩阵，对其进行谱分解，设

$$\boldsymbol{U} = \mathrm{e}^{\mathrm{i}\theta_0}|\psi\rangle\langle\psi| + \sum_{n=1}^{d-1}\mathrm{e}^{\mathrm{i}\theta_n}\left|\psi_n^{\perp}\right\rangle\left\langle\psi_n^{\perp}\right| \tag{1.84}$$

其中，$\forall n, \left\langle\psi\middle|\psi_n^{\perp}\right\rangle = 0$。那么

$$\langle\psi|\boldsymbol{U}|\psi\rangle = \mathrm{e}^{\mathrm{i}\theta_1} \neq 0 \tag{1.85}$$

这与 $\boldsymbol{U}$ 是通用的非门这一假设矛盾。

需要特别指出的是，针对计算基矢 $|0\rangle, |1\rangle$ 的非门是存在的。比如 σ_x 门就是一个例子，

$$\begin{cases} \sigma_x |0\rangle = |1\rangle \\ \sigma_x |1\rangle = |0\rangle \end{cases} \tag{1.86}$$

在仅考虑 $|0\rangle, |1\rangle$ 两个量子态的情形下，我们所关心的物理系统实际上是一个经典比特，σ_x 也变成了经典信息处理中的非门。

习题

习题 1.1 (泡利矩阵)　两个矩阵的对易子定义为

$$[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{AB}-\boldsymbol{BA} \tag{1.87}$$

试证明泡利矩阵之间的对易关系为

$$[\sigma_i,\sigma_j]=2\mathrm{i}\sum_{k=1}^{3}\epsilon_{ijk}\sigma_k \tag{1.88}$$

这里将泡利矩阵的下标重新标记为 $i,j,k\in\{1,2,3\}\equiv\{x,y,z\}$，Kronecker 符号值为 $\epsilon_{123}=\epsilon_{231}=\epsilon_{312}=1$，$\epsilon_{321}=\epsilon_{213}=\epsilon_{132}=-1$，其他情况下 $\epsilon_{ijk}=0$。

习题 1.2 (量子态及可观测量的计算)

(1) 写出 $|+\rangle\langle 0|-\rangle\langle +|\ +\ |-i\rangle\langle -i|$ 的矩阵形式。这是一个合法的量子态吗？如果不是的话，请尝试将其归一化。

(2) 用可观测量 H (Hadamard 门) 测量 (1) 中的 (归一化) 的量子态的期望是什么？

(3) 用可观测量 H (Hadamard 门) 测量 (1) 中的 (归一化) 的量子态的结果及相应的输出量子态是什么？

习题 1.3 (布洛赫球表示)

(1) 找出下列四个态在布洛赫球面上对应的角度 θ 和 ϕ: $|0\rangle$，$|+\rangle=(|0\rangle+|1\rangle)/\sqrt{2}$，$|-\mathrm{i}\rangle=(|0\rangle-\mathrm{i}\,|1\rangle)/\sqrt{2}$，$(|1\rangle+|-\rangle+|+\mathrm{i}\rangle)/\sqrt{5}$。

(2) 对一个任意的量子比特态 (可能是纯态或是混态) $\boldsymbol{\rho}$, 证明它可以被表示为

$$\boldsymbol{\rho}=\frac{\boldsymbol{I}+\boldsymbol{n}\cdot\boldsymbol{\sigma}}{2} \tag{1.89}$$

其中，$\boldsymbol{\sigma}=\left\{\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\sigma_y=\begin{pmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{pmatrix},\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right\}$ 是泡利矩阵组成的向量。此外，写出 $\boldsymbol{n}$ 在布洛赫球表示中的含义。

(3) 对于任意两个互相垂直的纯态，证明这两个态在布洛赫球上对应点的连线过球心。

习题 1.4 (量子态的计算)　考虑下面这两个量子态

$$\begin{cases}|\psi\rangle\langle\psi|=\dfrac{1}{2^2}\displaystyle\sum_{i,j=0}^{3}r_{ij}\sigma_i\otimes\sigma_j\\[2ex]|\phi\rangle\langle\phi|=\dfrac{1}{2^2}\displaystyle\sum_{i,j=0}^{3}r'_{ij}\sigma_i\otimes\sigma_j\end{cases} \tag{1.90}$$

其中，$r_{00}=r'_{00}=1$。定义两个向量 $\boldsymbol{r}_1=(r_{01},r_{02},\cdots,r_{33})$ 及 $\boldsymbol{r}_2=(r'_{01},r'_{02},\cdots,r'_{33})$。尝试找出 $\cos\theta=\dfrac{\boldsymbol{r}_1\cdot\boldsymbol{r}_2}{|\boldsymbol{r}_1||\boldsymbol{r}_2|}$ 的最小值。

习题 1.5 (保真度)

(1) 假设有一个单量子比特（自旋 1/2）处于未知的纯态 $|\psi\rangle$，这个态是从均匀分布在布洛赫球面上的量子集合中随机选取的。随机猜测状态是 $|\phi\rangle$。试计算这一猜测平均保真度 F。保真度（fidelity）定义为

$$F=|\langle\phi|\psi\rangle|^2 \tag{1.91}$$

(2) 在 (1) 小题中随机选择一个单量子比特纯态之后，对自旋沿 $\hat{z}$-轴进行测量。这种测量制备了一个可以由下述密度矩阵描述的量子态：

$$\boldsymbol{\rho}=P_\uparrow\langle\psi|P_\uparrow|\psi\rangle+P_\downarrow\langle\psi|P_\downarrow|\psi\rangle \tag{1.92}$$

其中，$P_{\uparrow,\downarrow}$ 表示对 $\hat{z}$ 轴上自旋向上和自旋向下的态的投影。从平均的保真度（On the average, with what fidelity）

$$F\equiv\langle\psi|\rho|\psi\rangle \tag{1.93}$$

角度来说，这个矩阵能够多好地表示初始的态 $|\psi\rangle$? 与 (1) 的答案相比，保真度 F 的提升可以粗略衡量我们通过测量操作了解到多少信息。

习题 1.6 (量子不可克隆定理)

(1) 假设有两个互相正交的量子态 $|\psi\rangle$ 和 $|\psi^\perp\rangle$，且满足 $\langle\psi|\psi^\perp\rangle=0$。构造一个可以克隆这两个态的 2 量子比特的幺正变换，也就是说，找到满足下式的幺正算符 U：

$$U|\psi\rangle|0\rangle=|\psi\rangle|\psi\rangle \tag{1.94}$$

$$U|\psi^\perp\rangle|0\rangle=|\psi^\perp\rangle|\psi^\perp\rangle \tag{1.95}$$

(2) 证明无法以 100% 的正确率分辨两个不互相正交的态。

习题 1.7 (由克隆导出超光速通信)

(1) 证明

$$|\Phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)=\frac{1}{\sqrt{2}}(|{+}{+}\rangle+|{-}{-}\rangle) \tag{1.96}$$

(2) 如果存在通用的量子克隆器，那么甲就有可能通过利用和乙之间共享的 EPR 对 $|\Phi^+\rangle$，来实现以比光速更快的速度向乙发出信号。也就是说，证明存在一种协议，可以利用量子克隆器实现这种超光速通信。

习题 1.8 (相位随机化的相干态[8]) 在量子光学中，光子数基矢（Fock basis）$\{|n\rangle\}_{n=0}^{\infty}$ 被广泛应用，其中 k 表示光子数。相干态的定义如下：

$$|\alpha\rangle \equiv \mathrm{e}^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{1.97}$$

其中，$\alpha \in \mathbb{C}$。

（1）如果在光子数基矢下测量式 (1.97)中的相干态，那么测量结果（光子数）的概率分布是什么？

（2）有时候我们会将一个复数 α 写成 $|\alpha|\mathrm{e}^{\mathrm{i}\theta}$ 的形式，其中 $\theta \in [0, 2\pi)$。证明如果相干态的相位 θ 是随机化的，那么总的态可以写成如下形式：

$$\int_0^{2\pi} |\alpha\rangle\langle\alpha| \, d\theta = \sum_{n=0}^{\infty} P(n) \, |n\rangle\langle n| \tag{1.98}$$

其中，$P(n)$ 服从泊松分布，并且平均光子数为 $\mu = |\alpha|^2$。

习题 1.9 求 Weyl 算符 X 的本征态。

习题 1.10 (相互无偏基) 对于两组在希尔伯特空间 $\mathcal{H}_d$ 上的相互无偏基 $\{|e_1\rangle, |e_2\rangle, \cdots, |e_d\rangle\}$ 和 $\{|f_1\rangle, |f_2\rangle, \cdots, |f_d\rangle\}$，它们满足各自的任意基量子态 $|e_j\rangle$ 和 $|f_k\rangle$ 的内积大小的平方等于维数 d 的倒数，即 $\forall j, k \in \{1, 2, \cdots, d\}$，

$$|\langle e_j|f_k\rangle|^2 = \frac{1}{d} \tag{1.99}$$

一组基中的每两个都是相互无偏的，那么它们被称为相互无偏基（mutually unbiased bases，MUB）。

（1）* 找出量子比特 $\mathcal{H}_2$ 情况下最多可以有的相互无偏基的数量。

（2）** 找出维数 d 为素数幂的情况下最多可以有的相互无偏基的数量[9]。

（3）*** 找出任意维数 d 下最多可以有的相互无偏基的数量。事实上 $d = 6$ 时这个问题已经很难解决了。

第 2 章

量子系统的一般描述

在本章，将介绍两体量子系统，并从两体系统出发来介绍一般量子态，以及对它的测量和演化的数学描述。本章中，会经常以两个量子比特构成的系统为例子，当然这些结果也可以扩展到高维多体系统中。当我们考虑两体量子系统时，会出现很多有趣的问题。特别是，会发现第 1 章中介绍的希尔伯特空间中的射线不足以来表示一个量子态，一般的测量也不是投影测量，一般量子演化的描述也会比幺正矩阵复杂不少。

本章内容是对第 1 章单体系统的拓展，也是量子信息科学的基础。2.1 节可以视作第 1 章介绍的高维量子系统的一个例子，主要介绍了两体量子纯态；2.2 节引入了更为一般的量子态，即混态；2.3 节则介绍了一般量子测量的描述和性质；2.4 节相应地介绍了一般的量子操作的描述，即量子信道；2.5 节给出一般量子演化的一些例子。如果读者发现 2.1 节理解起来有困难，应该重新巩固第 1 章的知识。一般来讲，2.2 节作为量子信息基础应该完全掌握。2.3 节除非作相关领域研究，一般基本掌握即可。2.4 节与 2.5 节相对来说难度较高，只需掌握相关的概念，知道幺正矩阵演化的拓展即可。对于一般的量子演化，掌握 2.5 节中的具体例子即可。另外，本章在数学上大量用到了张量积和偏迹，不熟悉的读者可以参考第 0 章数学基础部分。

2.1 两体量子纯态

2.1.1 两体量子系统与量子态

在介绍两体量子系统中的量子态之前，先来看一下这个系统本身及其表示。两体量子系统，顾名思义，可以看作是由两个部分组成的。当然，相应的也有多体量子系统，由多个部分组成。通常，将这些整体系统中的组成部分称为子系统（subsystem）。

现在，假设有个系统由两部分组成，分别记为子系统 A 和子系统 B。两个子系统所对应的希尔伯特空间分别为 $\mathcal{H}_A$ 和 $\mathcal{H}_B$。这样整体系统的希尔伯特空间为 $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$。这里的 $\otimes$ 是张量积（tensor product）运算，不熟悉的读者可以阅读 0.3.2 节中的相关定义。需要注意的是，子系统 $\mathcal{H}_A$ 和 $\mathcal{H}_B$ 的维度可以不

相同，同时它们本身也可以视作一个系统，拥有各自的基矢。同样地，对于这两个子系统的状态，也可以用各自的量子态描述，例如子系统 A 的状态可以由量子态 $|\psi\rangle_A$ 描述，子系统 B 的状态可以由 $|\phi\rangle_B$ 描述，我们通常会用下标来标记这个量子态属于哪一个子系统。对于两个相互 “独立” 的态 $|\psi\rangle_A, |\phi\rangle_B$，整体的态可以简单地用二者的直积 $|\psi\rangle \otimes |\phi\rangle$ 来表示，大多数时候，会把这个态简写为 $|\psi\rangle|\phi\rangle$ 或是 $|\psi\phi\rangle$。这一点来源于量子力学中关于**复合系统**的公理，是个很自然的想法。很可惜的是，从后面的一些例子可以看出：这个公理不能够描述所有两体量子系统。从这一章开始，将会看到一些真正的 “量子” 性质。

搞清楚两体量子系统后，考虑一个最简单的两体量子态——双量子比特。这样该量子态对应的希尔伯特空间的维度是 4，因此，这个空间对应有 4 个基矢态。最简单的一组基矢就是 $|00\rangle, |01\rangle, |10\rangle, |11\rangle$。我们也把由 $|0\rangle$ 和 $|1\rangle$ 直积得到的态组成的这组基矢称为计算基矢（computational basis）。对于更高维的系统，也可以类似地定义计算基矢。

除了计算基矢外，贝尔态基矢是其中被广泛使用的一组基矢，后面会经常用到。在 Z 基矢下，这 4 个贝尔态可以写成如下形式：

$$\begin{cases} \left|\Phi^+\right\rangle = \dfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ \left|\Phi^-\right\rangle = \dfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \\ \left|\Psi^+\right\rangle = \dfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ \left|\Psi^-\right\rangle = \dfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \end{cases} \tag{2.1}$$

在 X 基矢下则为

$$\begin{cases} \left|\Phi^+\right\rangle = \dfrac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle) \\ \left|\Phi^-\right\rangle = \dfrac{1}{\sqrt{2}}(|{-}{+}\rangle + |{+}{-}\rangle) \\ \left|\Psi^+\right\rangle = \dfrac{1}{\sqrt{2}}(|{+}{+}\rangle - |{-}{-}\rangle) \\ \left|\Psi^-\right\rangle = \dfrac{1}{\sqrt{2}}(|{-}{+}\rangle - |{+}{-}\rangle) \end{cases} \tag{2.2}$$

在 Y 基矢下则为

$$\begin{cases} \left|\Phi^{+}\right\rangle = \dfrac{1}{\sqrt{2}}(|{+}{\rm i}-{\rm i}\rangle + |{-}{\rm i}+{\rm i}\rangle) \\ \left|\Phi^{-}\right\rangle = \dfrac{1}{\sqrt{2}}(|{+}{\rm i}+{\rm i}\rangle + |{-}{\rm i}-{\rm i}\rangle) \\ \left|\Psi^{+}\right\rangle = \dfrac{-{\rm i}}{\sqrt{2}}(|{+}{\rm i}+{\rm i}\rangle - |{-}{\rm i}-{\rm i}\rangle) \\ \left|\Psi^{-}\right\rangle = \dfrac{\rm i}{\sqrt{2}}(|{+}{\rm i}-{\rm i}\rangle - |{-}{\rm i}+{\rm i}\rangle) \end{cases} \tag{2.3}$$

这里 $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$，$|\pm{\rm i}\rangle = (|0\rangle \pm {\rm i}\,|1\rangle)/\sqrt{2}$。我们在 $|\Psi^{\pm}\rangle$ 加了全局相位 i，这么做仅仅是为了数学形式上的统一，并无特殊物理含义。

很多简单有趣的量子信息现象都与贝尔态有关，例如贝尔不等式（Bell's inequality）、隐形传态（teleportation）、超密编码（superdense coding）等，会在后面章节一一展开介绍。

2.1.2 施密特分解

从 2.1.1 节的式(2.1)∼ 式(2.3)，我们发现 4 个贝尔态都可以写成如下形式：

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}\,|i\rangle_A\,|i\rangle_B \tag{2.4}$$

$\{|i\rangle_A\}$ 和 $\{|i\rangle_B\}$ 分别是两个子系统 A, B 的两组正交基矢。这种分解的方式可以推广到任意两体系统中的纯态，得到施密特分解（Schmidt decomposition）定理。

定理 2.1 (施密特分解) 对任意一个两体系统的纯态 $|\Psi\rangle_{AB}$，存在两组正交基矢 $\{|\phi_i\rangle_A\}$ 和 $\{|\psi_i\rangle_B\}$ 使得

$$|\Psi\rangle_{AB} = \sum_i \sqrt{p_i}\,|\phi_i\rangle_A\,|\psi_i\rangle_B \tag{2.5}$$

子系统 A 和 B 具有相同的本征值 $p_i \geqslant 0$ 且满足 $\sum\limits_i p_i = 1$，其中非零本征值的数量被称为 $|\Psi\rangle_{AB}$ 的施密特数（Schmidt number）。

证明 这里给出在子系统 A 和 B 的维度相同时施密特分解正确性的证明，更一般的维度不同的情况下的证明留给读者作为练习。取子系统 A 和 B 各自的一组任意的正交基矢 $\{|j\rangle_A\}$ 和 $\{|k\rangle_B\}$，那么 $|\Psi\rangle_{AB}$ 就可以写为如下形式：

$$|\Psi\rangle_{AB} = \sum_{jk} a_{jk}\,|j\rangle_A\,|k\rangle_B \tag{2.6}$$

其中，a_{jk} 可以看作构成了一个矩阵 $\boldsymbol{A}$。可以对这个矩阵进行奇异值分解，$\boldsymbol{A} = \boldsymbol{UDV}$，其中 $\boldsymbol{D}$ 是一个由非负元素组成的归一化的对角矩阵，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是两个幺

正矩阵。因此有

$$|\Psi\rangle_{AB} = \sum_{ijk} u_{ji} d_{ii} v_{ik} |j\rangle_A |k\rangle_B \tag{2.7}$$

接下来令 $|\phi_i\rangle_A = \sum_j u_{ji} |j\rangle_A$，$|\psi_i\rangle_B = \sum_k v_{ik} |k\rangle_B$，$\sqrt{p_i} = d_{ii}$，就能得到：

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i} |\phi_i\rangle_A |\psi_i\rangle_B \tag{2.8}$$

由 $\boldsymbol{U}$ 的幺正性和 $\{|j\rangle_A\}$ 的正交性不难得到 $\{|\phi_i\rangle_A\}$ 构成了一组正交基矢，同样地，$\{|\psi_i\rangle_B\}$ 也构成了一组正交基矢。 □

后面会看到，当施密特数大于 1 时，这个两体系统的态是纠缠的。不难看出，上文提到的四个贝尔态都是纠缠态。

2.1.3 有趣的“悖论”

2.1.2 节的施密特分解告诉我们，两体系统的纯态可以写成一系列量子态直积的求和。那么，问一个很简单的问题，是不是可以通过选取合适的基矢，让任意的两体系统的态 $|\Psi\rangle_{AB}$ 写成两个子系统中态的直积形式，$|\Psi\rangle_{AB} = |\psi\rangle_A |\phi\rangle_B$？如果你使用上面介绍的贝尔态尝试一下就可以看出，答案是不能。事实上，由线性代数可知，所有施密特数大于 1 的两体量子态都不能写成直积形式。让我们看一下其中一个贝尔态，

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B) \tag{2.9}$$

当我们在 A 系统进行 Z 基矢测量时，有 1/2 的概率测量的结果是 $|0\rangle$ 并且测量之后的态变为了 $|0\rangle_A |0\rangle_B$；另外 1/2 的概率，测量结果为 $|1\rangle$ 且测量之后的态变为 $|1\rangle_A |1\rangle_B$。一个朴素的想法是，根据第 1 章量子态的表示方法，A 系统的态可以表示为

$$|\psi\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle + \mathrm{e}^{\mathrm{i}\alpha} |1\rangle) \tag{2.10}$$

这里有个相对相位的自由度 $\alpha \in [0, 2\pi)$。考虑这个贝尔态在 X 和 Y 基矢下的表示，

$$\begin{cases} \left|\Phi^+\right\rangle_{AB} = \dfrac{1}{\sqrt{2}}(|+\rangle_A |+\rangle_B + |-\rangle_A |-\rangle_B) \\ \left|\Phi^+\right\rangle_{AB} = \dfrac{1}{\sqrt{2}}(|+\mathrm{i}\rangle_A |-\mathrm{i}\rangle_B + |-\mathrm{i}\rangle_A |+\mathrm{i}\rangle_B) \end{cases} \tag{2.11}$$

同样通过上述的简单想法来计算，有 $\beta,\gamma\in[0,2\pi)$,

$$|\psi\rangle_A=\frac{1}{\sqrt{2}}(|+\rangle+\mathrm{e}^{\mathrm{i}\beta}|-\rangle) \tag{2.12}$$

$$|\psi\rangle_A=\frac{1}{\sqrt{2}}(|+\mathrm{i}\rangle+\mathrm{e}^{\mathrm{i}\gamma}|-\mathrm{i}\rangle) \tag{2.13}$$

很遗憾，式(2.10)、式(2.12)、式(2.13)联立并没有对于相对相位 α,β,γ 的解。事实上，我们可以很快看到，式(2.10)和式(2.12)联立就可以得出 $|\psi\rangle_A=|\pm\mathrm{i}\rangle$，而这个结论与式(2.13)矛盾。哪里出问题了？

思考题 2.1 这个贝尔态不管用 X,Y,Z 哪组基矢下做测量，得到两个不同结果的概率都是完全随机的 1:1，这样一个态如果在布洛赫球里面表示，应该在哪个位置？该位置还是一个希尔伯特空间中的射线吗？

事实上，这个 “悖论” 来源于我们不能将其中的子系统 A 中的量子态当作一个射线，或者 “纯态”。在第 1 章中讲的态线性叠加是一个非平凡的操作。这里，简单地把两个会出现的态线性叠加起来，出现了矛盾。为了解决这个矛盾，会在下面用偏迹来引入密度矩阵。

2.2 一般态

前面提到，对于两个相互独立子系统态 $|\psi\rangle,|\phi\rangle$，整体的态可以简单地用二者的直积 $|\psi\rangle|\phi\rangle$ 来表示。这一点来源于量子力学中关于复合系统的公理，是个很自然的想法。但这个公理能够描述所有两体量子系统吗？即是不是所有的两体系统的态 $|\Psi\rangle_{AB}$ 都可以像上面那样写成两个态的直积形式？从 2.1.3 节例子中可以看出，有一些态是不能写成直积形式的。而无论是在理论还是在实验上，物理学家们都已经发现了存在不能被分解成直积形式的量子态 $|\Psi\rangle_{AB}$。那么在给定一个这样不可分的态 $|\Psi\rangle_{AB}$ 时，应该如何表示子系统的状态呢？这里，将引入密度矩阵的偏迹来解决这个问题。

2.2.1 从偏迹到子系统的密度矩阵

先来看一下复合系统的密度矩阵，对于 $|\Psi\rangle_{AB}=|\psi\rangle_A|\phi\rangle_B$，其密度矩阵写为

$$\begin{aligned}\boldsymbol{\rho}_{AB}&=|\Psi\rangle\langle\Psi|_{AB}\\&=(|\psi\rangle_A|\phi\rangle_B)(\langle\psi|_A\langle\phi|_B)\\&=|\psi\rangle\langle\psi|_A\otimes|\phi\rangle\langle\phi|_B\\&=\boldsymbol{\rho}_A\otimes\boldsymbol{\rho}_B\end{aligned} \tag{2.14}$$

由偏迹和直积的定义可以看出，$\boldsymbol{\rho}_A = \mathrm{tr}_B(\boldsymbol{\rho}_{AB})$，$\boldsymbol{\rho}_B = \mathrm{tr}_A(\boldsymbol{\rho}_{AB})$。

对于任意一个复合系统的量子态 $|\Psi\rangle_{AB}$，都可以把它表示成一个密度矩阵。那么，一般地，是不是其子系统的密度矩阵都可以从复合系统密度矩阵的偏迹给出？答案是肯定的。事实上，复合系统的密度矩阵和子系统的密度矩阵就是用偏迹联系起来的，这是量子力学的公理之一。

公理 2.1 给定一个孤立的复合系统量子态 $|\Psi\rangle_{AB}$，其子系统的态由偏迹给出：

$$\begin{cases} \rho_A = \mathrm{tr}_B(|\Psi\rangle\langle\Psi|_{AB}) \\ \rho_B = \mathrm{tr}_A(|\Psi\rangle\langle\Psi|_{AB}) \end{cases} \tag{2.15}$$

一个量子态最一般的形式可以用密度矩阵来表示。这里用 $\mathcal{D}(\mathcal{H}_A)$ 来表示作用在希尔伯特空间 $\mathcal{H}_A$ 上所有密度矩阵的集合。在量子信息中，当我们说量子态的时候，往往指的是系统的密度矩阵而不是右矢形式。可以尝试计算一下 2.1 节中提到的贝尔态在子系统 A 的密度矩阵：

$$\begin{aligned} \boldsymbol{\rho}_A &= \mathrm{tr}_B(|\Phi^+\rangle\langle\Phi^+|_{AB}) \\ &= \frac{1}{2}\,\mathrm{tr}_B[(|00\rangle + |11\rangle)(\langle 00| + \langle 11|)] \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \end{aligned} \tag{2.16}$$

从结果不难看出，对于一个一般的密度矩阵 $\boldsymbol{\rho}$，不能写成左右矢的形式，$\boldsymbol{\rho} \neq |\phi\rangle\langle\phi|$。

你可能很好奇为什么用偏迹操作来得到子系统的态。这是因为偏迹操作是唯一能给出与观测量相符的物理量的线性算符。如果你想对这一点有更深的理解，可以考虑一下下面的问题。这一点会在 2.3.1节定义好一般测量之后进一步讨论。

思考题 2.2 验证如果乙执行幺正算符或投影测量时没有通知甲测量结果，甲的局部密度矩阵不会改变。

另外，如果复合系统不是一个孤立系统，复合系统和子系统的量子态之间的关系由下面的推论给出。

推论 2.1 给定一个复合系统量子态 ρ_{AB}，其子系统的态由偏迹给出：

$$\begin{cases} \rho_A = \mathrm{tr}_B(\rho_{AB}) \\ \rho_B = \mathrm{tr}_A(\rho_{AB}) \end{cases} \tag{2.17}$$

证明 把所有和联合系统 AB 有相互作用的系统记为 C，这里如果有必要把整个世界的系统包含进来，这样 ABC 系统形成了一个孤立系统，可以由一个希尔伯特空间中的态 $|\Psi\rangle_{ABC}$ 给出。根据公理 2.1，可以将 BC 看作一个整体求偏迹来求子系统 A 中的态：

$$\rho_A = \mathrm{tr}_{BC}(|\Psi\rangle\langle\Psi|_{ABC}) \tag{2.18}$$

由偏迹操作的定义有

$$\begin{aligned}\rho_A &= \mathrm{tr}_{BC}(|\Psi\rangle\langle\Psi|_{ABC}) \\ &= \mathrm{tr}_B[\mathrm{tr}_C(|\Psi\rangle\langle\Psi|_{ABC})]\end{aligned} \tag{2.19}$$

继续采用公理 2.1，如果将系统 AB 看作一个整体，有 $\rho_{AB} = \mathrm{tr}_C(|\Psi\rangle\langle\Psi|_{ABC})$，所以

$$\begin{aligned}\rho_A &= \mathrm{tr}_B[\mathrm{tr}_C(|\Psi\rangle\langle\Psi|_{ABC})] \\ &= \mathrm{tr}_B(\rho_{AB})\end{aligned} \tag{2.20}$$

同样地，可以得到 $\rho_B = \mathrm{tr}_A(\rho_{AB})$。 □

在方框 2中，列举了由公理 2.1给出的密度矩阵 $\boldsymbol{\rho}$ 的性质，相关的证明留做习题。事实上，满足这三条性质的矩阵一定是一个量子态，即可以由公理 2.1给出，其证明将在 2.2.3节中给出。

方框 2：密度矩阵的性质

1. 厄米性：$\boldsymbol{\rho}^\dagger = \boldsymbol{\rho}$。
2. 半正定：$\boldsymbol{\rho} \geqslant 0$。
3. 归一化：$\mathrm{tr}(\boldsymbol{\rho}) = 1$。

我们注意到，除了不满足 $\rho^2 = \rho$，这里的密度矩阵 ρ_A 和之前介绍的纯态的密度矩阵 $|\psi\rangle\langle\psi|$ 有一样的性质。你可以很快计算一下这条性质发生了怎样的改变，然后证明一下：$\rho^2 = \rho \Leftrightarrow \mathrm{tr}(\rho^2) = 1$。根据这个性质，可以定义一个量来区分纯态与非纯态，即纯度（purity）。

定义 2.1 (纯度) 一个态 ρ 的纯度定义为 $\mathrm{tr}(\rho^2)$。

思考题 2.3 证明满足密度矩阵性质方框 2的 ρ 都有 $\mathrm{tr}(\rho^2) \leqslant 1$。

从矩阵的秩出发，不难证明，一个态可以写成左右矢的形式，$\rho = |\psi\rangle\langle\psi|$，当且仅当它的纯度为 1，即 $\mathrm{tr}(\rho^2) = 1$。这个时候，量子态可以用希尔伯特空间的一条射线来表示，我们说这个态是纯态。因此，纯度可以用来区分纯态与非纯态。

例 2.1 一个复合系统处在量子态 $|\Psi\rangle_{AB} = a\,|0\rangle_A\,|0\rangle_B + b\,|1\rangle_A\,|1\rangle_B$ 上，求子系统的量子态及其纯度。

解 首先写出复合系统量子态的密度矩阵形式，

$$\rho_{AB} = aa^*\,|00\rangle\langle 00| + bb^*\,|11\rangle\langle 11| + ab^*\,|00\rangle\langle 11| + a^*b\,|11\rangle\langle 00| \tag{2.21}$$

子系统量子态由偏迹给出，

$$\begin{cases}\rho_A = \mathrm{tr}_B(\rho_{AB}) = aa^*\,|0\rangle\langle 0| + bb^*\,|1\rangle\langle 1| \\ \rho_B = \mathrm{tr}_A(\rho_{AB}) = aa^*\,|0\rangle\langle 0| + bb^*\,|1\rangle\langle 1|\end{cases} \tag{2.22}$$

子系统的纯度为

$$\begin{cases} \mathrm{tr}\left(\rho_A^2\right) = \mathrm{tr}\left[(aa^* |0\rangle\langle 0| + bb^* |1\rangle\langle 1|)(aa^* |0\rangle\langle 0| + bb^* |1\rangle\langle 1|)\right] = |a|^4 + |b|^4 \\ \mathrm{tr}\left(\rho_B^2\right) = \mathrm{tr}\left[(aa^* |0\rangle\langle 0| + bb^* |1\rangle\langle 1|)(aa^* |0\rangle\langle 0| + bb^* |1\rangle\langle 1|)\right] = |a|^4 + |b|^4 \end{cases} \tag{2.23}$$

□

这里，我们看到 $\rho_A = \rho_B$，实际上，从 2.1.2 节施密特分解定理 2.1不难看出，对纯态 $|\psi\rangle_{AB}$，总是能找到 A 和 B 系统的一组合适基矢，$\{|\phi_i\rangle_A\}$，$\{|\psi_i\rangle_B\}$，使得

$$\begin{cases} \rho_A = \sum\limits_i p_i |\phi_i\rangle\langle\phi_i|_A \\ \rho_B = \sum\limits_i p_i |\psi_i\rangle\langle\psi_i|_B \end{cases} \tag{2.24}$$

于是，我们看到，ρ_A 和 ρ_B 有相同的非零本征值。需要强调的是，$|\phi_i\rangle_A$ 和 $|\psi_i\rangle_B$ 代表了不同系统的量子态，它们可以代表不同的量子态。

当然，一般情况下，我们没有 $\rho_A = \rho_B$，比如 $\rho_{AB} = \rho_A \otimes \rho_B$，这样两个系统的量子态可以没有任何关系。那么，为什么当 ρ_{AB} 是一个纯态时，子系统 A 和 B 的量子态有如此紧密的联系，这之中是不是还有什么更深层次的原因呢？后面章节中，会从信息的角度来进一步讨论。

2.2.2 纯态的混合

由量子态性质——方框 2，我们知道，密度矩阵是厄米的，所以总是可以被对角化，即有如下谱分解（spectrum decomposition）定理，其证明见第 0 章矩阵运算部分，正规矩阵的定义。

定理 2.2 (谱分解定理) 对于任意归一化的正定厄米算符 ρ, 总存在一组正交归一的纯态及其对应的非负实数 $\{p_i, |i\rangle\}$, 满足 $p_i \geqslant 0$, $\sum\limits_i p_i = 1$, 并且 $\langle i|j\rangle = \delta_{ij}$, 使得

$$\rho = \sum_i p_i |i\rangle\langle i| \tag{2.25}$$

对于前面提到的量子态纯度的定义 $\mathrm{tr}(\rho^2)$，现在从谱分解定理，有

$$\mathrm{tr}\left(\rho^2\right) = \sum_i p_i^2 \leqslant \sum_i p_i = 1 \tag{2.26}$$

显然，如果 $\rho = \rho^2$，那么只有一个 $\{p_i\}$ 等于 1，其他都应该是 0，因此 ρ 是一个纯态。这样就证明了思考题 2.3。

我们来看一下谱分解是否唯一。首先，一个正规矩阵的本征值是唯一的，也就是说，上述分解的 $\{p_i\}$ 是唯一的。如果这些本征值不简并，即均不相同，那么相应的本征态也是唯一确定的，于是谱分解是唯一的。但是如果本征值是简并的，即某个本征值对应了两个以上的本征态，也就是对应了一个希尔伯特空间的子空间，而我们可以采用这个子空间中的任意一组基矢来用作谱分解的 $|i\rangle$，这样，谱分解并不唯一。显然，简并情况下不同的分解之间可以用幺正变换来相互转换。

例 2.2 对于 2.1.3 节中的贝尔态，

$$\begin{aligned}\left|\Phi^+\right\rangle_{AB} &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ &= \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle) \\ &= \frac{1}{\sqrt{2}}(|{+}\mathrm{i}\,{-}\,\mathrm{i}\rangle + |{-}\mathrm{i}\,{+}\,\mathrm{i}\rangle)\end{aligned} \tag{2.27}$$

写出子系统 A 的量子态，讨论其谱分解。

解 子系统 A 的量子态由偏迹给出，

$$\begin{aligned}\rho_A &= \mathrm{tr}_B\left(\left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|_{AB}\right) \\ &= \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) \\ &= \frac{1}{2}\left(|+\rangle\langle +| + |-\rangle\langle -|\right) \\ &= \frac{1}{2}\left(|+\mathrm{i}\rangle\langle +\mathrm{i}| + |-\mathrm{i}\rangle\langle -\mathrm{i}|\right) \\ &= \frac{\boldsymbol{I}}{2}\end{aligned} \tag{2.28}$$

该式已经给出了 ρ_A 的三种不同的谱分解，对应贝尔态在三种不同基矢下的表示。从这里可以看出，由于本征值简并，谱分解并不唯一。事实上，对于任意两个正交的量子比特 $|\phi\rangle, \left|\phi^\perp\right\rangle \in \mathcal{H}_2$，$\left\langle\phi\middle|\phi^\perp\right\rangle = 0$，可以作为其谱分解，

$$\rho_A = \frac{\boldsymbol{I}}{2} = \frac{1}{2}\left(|\phi\rangle\langle\phi| + \left|\phi^\perp\right\rangle\!\left\langle\phi^\perp\right|\right) \tag{2.29}$$

□

对于上述例子中的 ρ_A，我们注意到，$\mathrm{tr}(\rho^2) = 1/2 < 1$，其纯度小于 1，所以并非是一个纯态。我们把所有纯度小于 1 的非纯态，$\mathrm{tr}(\rho^2) < 1$，称为混态（mixed state）。从谱分解定理可以看出，如果分解的非零本征值数量为 1，那么这个态是纯态，如果有超过一个非零本征值，那么这个态是混态。

实际上，混态并不总是需要像谱分解定理里那样写成一组相互正交的纯态。例如，$\rho=\frac{1}{2}(|0\rangle\langle 0|+|+\rangle\langle +|)$ 也是一个混态。所以，更一般地，一个混态 ρ 的密度矩阵总可以写成如下形式[①]：

$$\rho=\sum_i p_i\,|\phi_i\rangle\langle\phi_i| \tag{2.30}$$

其中，$p_i\geqslant 0,\sum_i p_i=1$，不同的纯态 $|\phi_i\rangle\langle\phi_i|$ 之间没有必然的关系。这里，可以把 $\{p_i\}$ 看成一个概率分布。从数学形式上可以看出，混态可以看作由多个纯态按一定的概率混合得到，这也是为什么我们将之称为混态。

在实验上，考虑如何制备一个量子态。对于式(2.30)这样的量子态 ρ，可以用下面三个不同的角度来看该态的制备。

（1）可以直接从公理 2.1出发来制备 ρ。考虑有一个仪器，先制备了下面这个 AB 联合系统的纯态：

$$|\Psi\rangle_{AB}=\sum_i \sqrt{p_i}\,|\phi_i\rangle_A\,|i\rangle_B \tag{2.31}$$

这里，$\{|i\rangle_B\}$ 是一组正交归一基矢。之后单独将 A 系统输出，就会得到 $\rho_A=\mathrm{tr}_B\left(|\Psi\rangle\langle\Psi|_{AB}\right)=\sum_i p_i\,|\phi_i\rangle\langle\phi_i|$，就是式(2.30)量子态 ρ。

（2）这个仪器在制备 $|\Psi\rangle_{AB}$ 之后，可以对系统 B 在基矢 $\{|i\rangle_B\}$ 上进行测量。不难看出，测量将会以 p_i 的概率得到 $|i\rangle_B$ 的结果，同时 A 系统上的态对应的密度矩阵会相应地变为 $|\phi_i\rangle\langle\phi_i|$。如果这个仪器不会向外界透露测量结果，那么该测量是不会影响输出系统 A 的量子态[②]的。因此外界用户得到的系统 A 的量子态还是由式(2.30)给出。

（3）更进一步地，既然外界用户无法得知仪器有没有进行测量，也无法区分有测量操作与无测量操作时输出的量子态，那么，这个仪器可以直接省略制备 $|\Psi\rangle_{AB}$ 和测量系统 B 的步骤，以 p_i 的概率输出一个量子态 $|\phi_i\rangle$，这样也能制备出式(2.30)中的量子态 ρ。

对比上述三种情况下制备的量子态，从一个外界用户看来在物理上无法区别。于是，对于用户来讲，系统 A 的量子态均由式(2.30)给出。反过来，也找到了式(2.30)的物理含义：各种可能的态 $|\phi_i\rangle$ 都有一定的概率 p_i 出现，这也是我们上面所提到的“概率混合”的物理含义，也是混态名称的由来。更一般地，如果设备以 p_i 的概率输出量子态 ρ_i，那么这个输出的量子态应该写为

---

① 严格证明这个将混态分解为纯态的关键点在于：密度矩阵空间 $\mathcal{D}(\mathcal{H}_A)$ 是一个凸空间，而且纯态是这个空间的极点（extreme points）。定理 2.2可以看作这个分解的特例。

② 可以想象一下，仪器在做测量的时候，A,B 两个子系统已经类空间隔了。外界用户甚至不知道仪器有没有对系统 B 进行测量。

$$\rho = \sum_i p_i \rho_i \tag{2.32}$$

需要强调的是，上述后两种制备方案中，仪器均不能向外界透露测量结果或者选择结果 i。如果仪器稍后将 i 的信息发给外界用户，用户手里的态就不再是混态 ρ 了，而变成了 $|\phi_i\rangle$。这个从信息的角度来看并不意外。量子态本来就是描述一个系统的状态，即系统的量子信息，也就是一个观察者对一个系统的认识。于是，量子态也就可以随着观察者获得更多信息而改变。这也是为什么我们说“物理是信息的”。关于这一点，可以看下面这两个问题。

例 2.3 （1）如果一个系统处在一个纯态上，是否还会因为观察者得到信息而改变？

（2）既然得到信息能够改变量子态，是否存在“失去”信息而改变量子态的情况？

解 （1）物理上，一个系统如果处在纯态，它是一个孤立系统，即数学上来讲，该系统和任何其他系统组成的联合系统的量子态一定是一个简单的直积形式，$|\phi\rangle\langle\phi| \otimes \rho_E$。从信息的角度来看，一个纯态已经包含了该系统的所有信息，观察者不可能从外界得到另外“有用”的信息。所以，如果一个系统处在一个纯态上，不会因为观察者得到信息而改变。

（2）第 1 章量子演化中我们知道，所有的量子操作都是幺正的，也就是可逆的。获得信息也可以看成一种量子操作。那么，既然有“获得信息”而改变量子态的操作，那么是否存在与之相反的“失去信息”而改变量子态的操作。这个就是信息擦除实验的基础。对此更深刻的理解，当前学术界的主流是多宇宙解释，有兴趣的读者可以找相关文献阅读。 □

思考题 2.4 在第 1 章我们了解了量子不可克隆定理，但是当时的证明仅仅是对纯态而言的。对于混态来说还有不可克隆定理成立吗？

2.2.3 混态的纯化

从 2.2.2 节我们知道，对于任意一个密度矩阵 $\boldsymbol{\rho}_A$，总可以找到它的一个纯态分解：

$$\boldsymbol{\rho}_A = \sum_i p_i |\phi_i\rangle\langle\phi_i| \tag{2.33}$$

这里，如果 $\{|\phi_i\rangle\}$ 组成一个正交归一基矢，那么该分解也称为谱分解，由定理 2.2给出。应该如何理解这种对于 $\boldsymbol{\rho}_A$ 的分解？一方面，可以从信息缺失角度来看。这里的信息缺失，以一种概率混合的“不确定”的形式记录下来，即如果观测者甲不确定系统 A 确切的量子态是什么，只知道这个态是某个集合中一个特定的量子态 $|\phi_i\rangle$ 的概率为 p_i，那么在她看来，系统量子态由式 (2.33) 给出。

在大多数情况下，我们并没有雄心勃勃地试图了解整个宇宙的物理描述。我们只是满足于观察自己的小角落。因此，在实践中，我们所做的观测总是局限于一个大得多的量子系统的一小部分。所以，式(2.33)可以看成一个大系统中的子系统量子态描述。因此，找到一个这部分系统对应的、虚拟的总系统有助于更直观地了解整个系统的演化。于是有如下量子态纯化（state purification）定理。

定理 2.3 (量子态纯化定理) 对于系统 A 的任意一个量子态 ρ_A，可以构造出 ρ_A 的一个“纯化”$|\Psi\rangle_{AB}$，使得

$$\rho_A = \mathrm{tr}_B(|\Psi\rangle\langle\Psi|_{AB}) \tag{2.34}$$

证明 首先，$\boldsymbol{\rho}_A$ 是迹为 1 的半正定厄米矩阵。可以找到 $\boldsymbol{\rho}_A$ 的一个纯态分解，见式(2.33)，比如谱分解。然后取系统 B 的一组正交归一量子态 $\{|i\rangle_B\}$，最后构造出一个纯化量子态：

$$|\Psi\rangle_{AB} = \sum_i \sqrt{p_i}\,|\phi_i\rangle_A\,|i\rangle_B \tag{2.35}$$

这里 $p_i > 0$，$\sum\limits_i p_i = 1$。 □

纯化之后得到的态 $|\Psi\rangle_{AB}$ 是个纯态，所以我们知晓复合系统 AB 的全部信息。这个点物理上也非常有趣。一定程度我们假设了整个世界是一个孤立系统，处在一个纯态上，那么，系统 B 是客观上存在的。

当然对于一个混态来说，纯化的构造并不唯一，两种不同的纯化 $|\Psi_1\rangle_{AB}$ 和 $|\Psi_2\rangle_{AB}$ 之间的关系可以由下式给出：

$$|\Psi_1\rangle_{AB} = (I_A \otimes U_B)\,|\Psi_2\rangle_{AB} \tag{2.36}$$

这两个态的差异可以由一个单独作用于 $\mathcal{H}_B$ 的幺正变换给出。证明留作习题 2.5。这里的物理含义也很清楚，既然得不到系统 B，那么并不能确定 B 上是否做了量子操作。反过来可以看出来，无论对系统 B 做何种量子操作，系统 A 的量子态不变。这个和非讯令性（no-signaling）相吻合。

例 2.4 纯化下面这个密度矩阵：

$$\boldsymbol{\rho} = \begin{pmatrix} \dfrac{4}{9} & \dfrac{1}{100} \\ \dfrac{1}{100} & \dfrac{5}{9} \end{pmatrix} \tag{2.37}$$

解 第一步，找到 $\boldsymbol{\rho}$ 的谱分解：

$$\boldsymbol{\rho} = \lambda_0\,|\phi_0\rangle\langle\phi_0| + \lambda_1\,|\phi_1\rangle\langle\phi_1| \tag{2.38}$$

其中，λ_0, λ_1 是 $\boldsymbol{\rho}$ 的本征值，$|\phi_0\rangle, |\phi_1\rangle$ 则是对应的本征态。

第二步，$\boldsymbol{\rho}$ 的纯化可以写成

$$|\psi_{AB}\rangle = \sqrt{\lambda_0}\,|\phi_0\rangle_A\,|0\rangle_B + \sqrt{\lambda_1}\,|\phi_1\rangle_A\,|1\rangle_B \tag{2.39}$$

这里，简单地取拓展的纯化 B 系统的计算基矢 $|0\rangle_B, |1\rangle_B$ 作为纯化态。具体数值请自行代入计算。 □

例 2.5 考虑量子态 $|\Phi^+\rangle_{AB} = \dfrac{1}{\sqrt{2}}(|0\rangle_A\,|0\rangle_B + |1\rangle_A\,|1\rangle_B)$，系统 A, B 分别在甲和乙手中。现在乙对手中的系统 B 进行了 Z 基矢测量。

（1）如果乙测量完之后不告诉甲测量的结果，甲手中的态是什么？

（2）如果乙告诉了甲测量的结果，$|0\rangle$ 或是 $|1\rangle$，甲手中的态是什么？

解 （1）由于甲不知道乙的测量结果，而乙的测量结果有一半概率是 $|0\rangle_B$，一半概率是 $|1\rangle_B$，所以甲手中的态也有一半的概率为 $|0\rangle_A$，一半概率为 $|1\rangle_A$，概率混合后为 $\dfrac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$。

（2）如果乙告诉甲测量结果为 $|0\rangle_B$，则原来的态坍缩到 $|00\rangle_{AB}$，甲手中的态为 $|0\rangle_B$；同理，如果乙告诉甲测量结果为 $|1\rangle_B$，甲手中的态为 $|1\rangle_A$。 □

从例 2.5中可以看出，乙在得到测量结果后就能够知道甲手中的纯态究竟是哪一个，但在甲看来，如果乙没有告诉她结果，系统 A 的一些信息缺失了，她手中的态就是一个混态。当乙告诉甲关于测量结果的信息时，这个混态就变成了纯态：信息是物理的，并且物理也是信息的！

在量子信息中，我们认为所有的混态 ρ_A 是由对一个更大的、不可分的系统求偏迹得到的。这里我们可以看到整个系统的不可分性和子系统中信息的缺失之间的紧密联系。我们将在后面章节从信息论的角度更清晰地阐述这一点。

有了纯化的概念，就可以来比较一下量子态的相干叠加与经典混合之间的区别。考虑这样一个相干叠加得到的态 $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$，它对应的密度矩阵是

$$|\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{2.40}$$

现在将它与另一个由经典混合得到的密度矩阵

$$\boldsymbol{\rho} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.41}$$

进行对比。如果在 Z 基矢上测量它们，这两个态的测量结果都会是以相等的概率得到 $z = 0$ 和 $z = 1$。看上去似乎这两个态有相同的产生随机性的能力。但是，如果在 X 基矢上对这两个态进行测量，那么它们是可以被分辨的。事实上，对于两个不同的量子态，总可以找到一个合适的测量来区分它们。

例 2.6 如果现在要在 $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ 和 $\rho=\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|$ 中选择一个态来产生随机数，那么应该选择哪一个态？

解 由于 $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ 是一个纯态，任意的额外信息都不会改变这个态，然而窃听者可以通过测量混态 $\rho=\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|$ 的纯化系统来对它进行攻击。因此，我们更倾向于使用前一个态来产生随机数。对这个问题感兴趣的同学，可以阅读文献 [10]。 □

这里，开始提到了“信息”的概念，并且发现它与量子态的纯度有着密切的关系。关于“信息”，特别是“量子信息”，将在第 4 章进一步解释。

2.2.4 混态的布洛赫球表示和量子层析

现在我们已经熟悉了混态 ρ，让我们重新考虑布洛赫球面上的表示。混态仍然可以表示在布洛赫球上吗？回忆一下第 1 章的表示，一个量子比特 ρ 可以写成

$$\begin{aligned}\rho&=\frac{1}{2}(\sigma_0+r\cdot\sigma)\\&=\frac{1}{2}[\mathrm{tr}(\rho)\sigma_0+\mathrm{tr}(\sigma_x\rho)\sigma_x+\mathrm{tr}(\sigma_y\rho)\sigma_y+\mathrm{tr}(\sigma_z\rho)\sigma_z]\end{aligned}\tag{2.42}$$

因此 $r=(x,y,z)$ 可以被看作 ρ 的坐标。这个表示并没有限定 ρ 必须是一个纯态，对于纯态，$|r|^2=2\,\mathrm{tr}(\rho^2)-1=1$，所以对应的点在布洛赫球面上，而对于一个混态 ρ_A^2，

$$|r|^2=2\,\mathrm{tr}\left(\rho_A^2\right)-1=x^2+y^2+z^2<1\tag{2.43}$$

这个可以由混态的纯度小于 1 来证明。因此，(x,y,z) 所代表的点是在布洛赫“球”内部而不是在球面上。

通过布洛赫球，也能很直观地表示混态分解，见式 (2.32)。将这个分解代入式(2.42)中，可以得到：

$$\begin{aligned}\rho&=\frac{1}{2}[\mathrm{tr}(\rho)\sigma_0+\mathrm{tr}(\sigma_x\rho)\sigma_x+\mathrm{tr}(\sigma_y\rho)\sigma_y+\mathrm{tr}(\sigma_z\rho)\sigma_z]\\&=\frac{1}{2}\left[\mathrm{tr}\left(\sum_i p_i\rho_i\right)\sigma_0+\mathrm{tr}\left(\sigma_x\sum_i p_i\rho_i\right)\sigma_x+\right.\\&\quad\left.\mathrm{tr}\left(\sigma_y\sum_i p_i\rho_i\right)\sigma_y+\mathrm{tr}\left(\sigma_z\sum_i p_i\rho_i\right)\sigma_z\right]\\&=\frac{1}{2}\left[\sigma_0+\sum_i p_i\,\mathrm{tr}(\sigma_x\rho_i)\sigma_x+\sum_i p_i\,\mathrm{tr}(\sigma_y\rho_i)\sigma_y+\sum_i p_i\,\mathrm{tr}(\sigma_z\rho_i)\sigma_z\right]\end{aligned}\tag{2.44}$$

所以 ρ 对应的向量 r 满足：

$$r = \left(\sum_i p_i x_i, \sum_i p_i y_i, \sum_i p_i z_i \right) = \sum_i p_i r_i \tag{2.45}$$

其中，r_i 是 ρ_i 在布洛赫球上的向量表示，x_i, y_i, z_i 则是对应的坐标，p_i 满足 $p_i > 0, \sum_i p_i = 1$。可以看出，这种分解就对应于布洛赫球表示里的分解，如图 2.1所示。

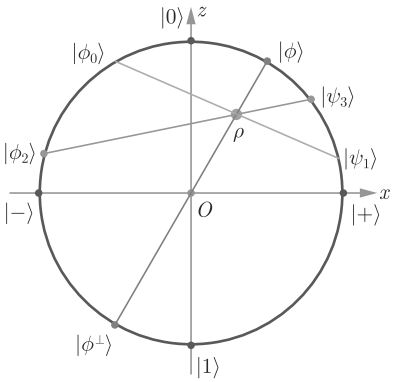

图 2.1　混态分解的布洛赫球表示。为了更加清晰地展示混态与纯态在布洛赫球的位置，图中取了布洛赫球在 x 轴与 z 轴所在平面上的截面作为一个例子。更一般的分解所涉及的混态与纯态，并不一定局限在这一平面上，甚至都不一定在同一平面上

对于将混态分解为两个纯态的概率混合，即 $\rho = p_0 |\psi_0\rangle\langle\psi_0| + p_1 |\psi_1\rangle\langle\psi_1|$ 的情况，由于 p_i 的性质，这三个态对应的点一定在同一条直线上。当这条直线经过布洛赫球的球心时，$|\phi\rangle, |\phi^\perp\rangle$ 互相正交，对应将 ρ 进行谱分解的情况。当这条直线没有经过球心时，就对应更一般的纯态分解的情况。显然，这种分解不唯一，图中显示了另外一种分解方法，$\rho = p_2 |\psi_2\rangle\langle\psi_2| + p_3 |\psi_3\rangle\langle\psi_3|$。事实上，一个混态可以写成多个纯态的分解，比如，$\rho$ 可以写成 $|\psi_0\rangle, |\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle$ 的混合。

类似地，之前学到的纯态量子层析方法也适用于混态。只需要测量三个可观测量 $\sigma_x, \sigma_y, \sigma_z$ 的平均值 $\mathrm{tr}(\sigma_x\rho), \mathrm{tr}(\sigma_y\rho), \mathrm{tr}(\sigma_z\rho)$，这样就可以代入式(2.42)得到相应的密度矩阵。

这里，还可以把结果拓展到一个 n 比特的系统。首先，引入 n 比特泡利算符，对一个矢量 $v \in \{I, x, y, z\}^n$，有

$$P_{\boldsymbol{v}} = \bigotimes_{i=1}^{n} \sigma_{v_i} \tag{2.46}$$

其中，v_i 是矢量 v 的第 i 个元素。那么，对 n 比特系统的层析可以由下式给出：

$$\rho = 2^{-n} \sum_{\boldsymbol{v} \in \{I,x,y,z\}^n} \mathrm{tr}(\rho P_{\boldsymbol{v}}) P_{\boldsymbol{v}} \tag{2.47}$$

求和式中有 4^n 项，所以应该有 4^n-1 的测量值。这种对密度矩阵的表示方法被称作 Pauli-Liouville 表示，感兴趣的同学可以阅读文献 [11] 和文献 [12]。注意，对 I^n 的测量实际上并不需要执行。这种层析的过程可能不是最佳的，因为任何局部 I 测量都可以省略。总的来说，只需要执行 3^n 次测量，但相对于 n 还是指数多的。

2.2.5 量子态距离的度量

从上述布洛赫球中我们看出，有些量子态比较接近，有些比较远。在量子信息中，时不时会讨论两个量子态有多接近。两个量子态相似度有许多不同的度量。下面讨论几种常用的。

为了度量两个密度矩阵之间的距离，先来看看概率论中的距离度量。给定两个概率分布 $\{p_x\}$ 和 $\{q_x\}$，一种度量距离的方式是迹距离（trace distance）：

$$D(p_x,q_x)\equiv\frac{1}{2}\sum_x|p_x-q_x| \tag{2.48}$$

另一种度量的方式是这两个概率分布间的保真度（fidelity）：

$$F(p_x,q_x)\equiv\sum_x\sqrt{p_xq_x} \tag{2.49}$$

从谱分解定理不难看出，密度矩阵和概率分布有很强的相通性。于是可以类似地定义量子态之间的迹距离和保真度。

定义 2.2 (迹距离)　两个量子态 ρ 和 σ 之间的迹距离定义如下：

$$D(\rho,\sigma)\equiv\frac{1}{2}\operatorname{tr}|\rho-\sigma| \tag{2.50}$$

定义 2.3 (保真度)　两个量子态 ρ 和 σ 之间的保真度定义如下：

$$F(\rho,\sigma)\equiv\operatorname{tr}\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}} \tag{2.51}$$

当 $\rho=|\phi\rangle\langle\phi|$ 是纯态时，有

$$F(\rho,\sigma)=\sqrt{\langle\phi|\,\sigma\,|\phi\rangle} \tag{2.52}$$

定理 2.4 (Uhlmann 定理)　给定两个量子态 ρ 和 σ，

$$F(\rho,\sigma)\equiv\max_{|\psi\rangle,|\phi\rangle}|\,\langle\psi|\phi\rangle\,| \tag{2.53}$$

其中最大值在所有 ρ 的纯化 $|\psi\rangle$ 和 σ 的纯化 $|\phi\rangle$ 中取。

迹距离和保真度之间的关系如下：

$$1 - F(\rho,\sigma) \leqslant D(\rho,\sigma) \leqslant \sqrt{1 - F(\rho,\sigma)^2} \tag{2.54}$$

不等式的证明留作习题 2.11。

需要注意的是，保真度并不是一个真正的距离度量（metric），因为它不满足三角不等式。但是 ρ 和 σ 之间的夹角是一个距离度量。

$$A(\rho,\sigma) = \arccos F(\rho,\sigma) \tag{2.55}$$

除了迹距离和保真度之外，还经常使用量子相对熵（quantum relative entropy）来度量两个量子态有多接近。

定义 2.4 (量子相对熵)　定义两个密度矩阵 $\boldsymbol{\rho}$ 和 $\boldsymbol{\sigma}$ 之间的相对熵为

$$S(\boldsymbol{\rho}||\boldsymbol{\sigma}) \equiv \mathrm{tr}(\boldsymbol{\rho}\log\boldsymbol{\rho} - \boldsymbol{\rho}\log\boldsymbol{\sigma}) \tag{2.56}$$

量子相对熵是相对熵在量子力学中的扩展产物，在数学意义上，它不是严格意义上的距离度量，因为它不对称，也不服从三角不等式。我们还没有真正涉及熵的概念，将在后面章节中再回到这个定义。

2.3 一般测量

在第 1 章中，已经学习了投影测量的内容，特别是秩为 1 的投影测量——冯·诺依曼测量。下面介绍量子力学中对于测量过程的最一般描述，从中会发现，冯·诺依曼测量是一种特殊但常见的情形。

2.3.1 正定算子测量

在前几节中，引入了一个非常重要的概念——混态，并且通过概率混合和纯化将其与第 1 章的重点——纯态联系了起来。那么对于第 1 章介绍的投影测量来说，是否也可以类似地进行混合呢？答案是肯定的，只要以不同的概率进行不同的测量就可以了，那这种概率混合后的测量是不是还是投影测量呢？不妨让我们考虑下面这个场景：

例 2.7　观测者甲将一个未知的量子比特态 ρ 输入一个用来进行量子测量的设备，这个设备会以 50% 的概率对输入的量子比特进行 Z 基矢测量，另外 50% 的概率则会进行 X 基矢测量，并输出测量结果 $\{1,-1\}$，这个设备进行的测量是投影测量吗？

解　这个设备进行的测量不是投影测量，当输出的测量结果是 1 时，甲有可能进行了 Z 测量，对应测量结果的态为 $|0\rangle$，也有可能进行了 X 测量，对应测量

结果的态为 $|+\rangle$，因此测量结果是 1 时对应的测量后的态为 $|0\rangle$ 与 $|+\rangle$ 按一定概率混合得到的混态，具体概率取决于被测量的态 ρ。同理，测量结果是 -1 时对应的测量后的态为 $|1\rangle$ 与 $|-\rangle$ 按一定概率混合得到的混态。一般情况下，对应不同测量结果的态并不正交，所以这不是一个投影测量。 □

从这个例子就可以看出，单纯的投影测量是无法描述所有的测量操作的，因此需要更严格地表述一下第 1 章中的测量公理。

公理 2.2 量子测量由一组测量算子 $\{M_m\}_m$ 描述，作用于待测量系统所处的希尔伯特空间上，指标 m 对应于实验中可能的测量结果。如果测量过程刚开始时刻量子系统处于状态 ρ，那么结果 m 发生的概率为

$$\Pr(m) = \mathrm{tr}\big(M_m \rho M_m^\dagger\big) \tag{2.57}$$

测量后相应的系统末态为

$$\frac{M_m \rho M_m^\dagger}{\mathrm{tr}\Big(M_m \rho M_m^\dagger\Big)} \tag{2.58}$$

根据概率公式(2.57)的求和归一性，我们可以得出，测量算子满足完备性条件：

$$\sum_m M_m^\dagger M_m = \boldsymbol{I} \tag{2.59}$$

对于这个公理，我们还有一些话想说。首先，这个公理告诉我们，测量是一个随机实验。而在使用密度矩阵的描述中，在物理上，实际上涉及两种类型的随机性，我们有时表示为内禀随机性（intrinsic randomness）和外在随机性（extrinsic randomness），这两者合在一起称为名义随机性（nominal randomness）。内禀随机性源于量子力学的不可预测性，而外在随机性也包含了信息不完全导致的不可预测性。内禀随机性的概念与量子相干性和叠加性有着深刻联系，感兴趣的同学可以阅读文献 [13]。

其次，我们在表示系统时有点草率。最终系统可以不同于原始系统，特别是态空间（希尔伯特空间）的维数可能会变化。如果明确地用 $\rho \in \mathcal{D}(\mathcal{H}_A)$, $M_m : \mathcal{H}_A \mapsto \mathcal{H}_{A'}$ 来表示的话，有

$$\begin{cases} \dfrac{M_m \rho M_m^\dagger}{\mathrm{tr}\Big(M_m \rho M_m^\dagger\Big)} \in \mathcal{D}(\mathcal{H}'_A) \\ \displaystyle\sum_m M_m^\dagger M_m = I_A \end{cases} \tag{2.60}$$

这里也反映出来，一般情况下，矩阵 M_m 并不一定是方阵。

思考题 2.5 证明式(2.57)对于任意密度矩阵 $\boldsymbol{\rho}$ 和测量算符都是非负的。

在很多情况下，系统的最终状态并不重要，只需要关注得到每种结果的概率。这种情况下可以用正定算子测量（positive operator-valued measure, POVM）来很好地描述测量。在测量假设中，得到测量结果 m 的概率为 $\mathrm{tr}\left(M_m\rho M_m^\dagger\right)$。利用求迹算子的循环性质，其概率可以写成 $\mathrm{tr}\left(M_m^\dagger M_m\rho\right)$。由此，可以定义一个新的算符：

$$E_m = M_m^\dagger M_m \tag{2.61}$$

很容易看出 $E_m \geqslant 0$ 是一个半正定算子。将集合 $\{E_m\}_m$ 称为一个 POVM，并将 E_m 称为 POVM 元素（element，注意不要与测量算符混淆）。在上述的定义下，不难验证，这些 POVM 元素满足性质见方框 3。其中前面两条由式 (2.61) 可以很快验证，最后一条可由概率归一化得出。

方框 3: 一般测量 POVM 的性质

1. 厄米性：$\boldsymbol{E}^\dagger = \boldsymbol{E}$。
2. 半正定：$\forall m,\ E_m \geqslant 0$。
3. 归一化：$\sum\limits_m E_m = \boldsymbol{I}$。

例 2.8 写出例 2.7中测量对应的 POVM 元素，并验证其满足 POVM 元素的性质。

解 由例 2.7中可能的测量结果，不难得出：

$$\begin{cases} \boldsymbol{E}_1 = \dfrac{1}{2}(|0\rangle\langle 0| + |+\rangle\langle +|) = \begin{pmatrix} \dfrac{3}{4} & \dfrac{1}{4} \\ \dfrac{1}{4} & \dfrac{1}{4} \end{pmatrix} \\ \boldsymbol{E}_{-1} = \dfrac{1}{2}(|1\rangle\langle 1| + |-\rangle\langle -|) = \begin{pmatrix} \dfrac{1}{4} & -\dfrac{1}{4} \\ -\dfrac{1}{4} & \dfrac{3}{4} \end{pmatrix} \end{cases} \tag{2.62}$$

容易验证，$\boldsymbol{E}_1, \boldsymbol{E}_{-1}$ 均为正定矩阵，且 $\boldsymbol{E}_1 + \boldsymbol{E}_{-1} = \boldsymbol{I}$。 □

思考题 2.2中，只考虑了投影测量，而现在已经有了一般测量算符的定义。考虑下面的问题，我们可以看到，即使将测量扩展到一般的测量，偏迹操作也能给出与观测量相符的物理量。

思考题 2.6 假设甲和乙共享了一个可以由密度矩阵 $\boldsymbol{\rho}_{AB}$ 来描述的量子系统。考虑一个甲可能会在她的系统上进行的局部测量，其测量算符为 $\{M_m\}_m$。那么总体的测量算符可以写成 $\{M_m^A \otimes I^B\}_m$。证明由全体的密度矩阵预测的测量结果

概率分布和局部的密度矩阵 $\boldsymbol{\rho}_A$ 给出的预测结果是一样的，其中，$\boldsymbol{\rho}_A = \mathrm{tr}_B(\boldsymbol{\rho}_{AB})$，

$$\mathrm{tr}\left[(M_m^A \otimes I^B)\boldsymbol{\rho}_{AB}\right] = \mathrm{tr}\left(M_m^A \boldsymbol{\rho}_A\right) \tag{2.63}$$

因此，全局量子理论的预测与局部量子理论的预测是一致的。

2.3.2 Naimark 定理

我们先重新表述一下投影测量（projection-valued measure, PVM），通常记为 $\{M_m\}_m$，满足：

（1）$M_m^\dagger = M_m$；

（2）$M_m M_{m'} = \delta_{m,m'} M_m$；

（3）$\sum\limits_m M_m^\dagger M_m = \boldsymbol{I}$。

相对于一般的测量，投影测量有两个额外的要求，即测量算子是厄米和相互正交的投影算符。例如，$\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ 是单比特系统两输出的投影值测量，而 $\{|0\rangle\langle 0| + |1\rangle\langle 1|, |2\rangle\langle 2|\}$ 是三维系统两输出的投影测量。

不难验证，投影测量是 POVM 的特殊情况，即测量算子与 POVM 元素相同，$E_m = M_m^\dagger M_m = M_m$。我们可以从密度矩阵和它纯化后的纯态之间的关系类似角度理解 POVM 测量和投影测量之间的区别。一个更有趣的事实是，任何 POVM 都可以看作一个 PVM 在一个更大的系统上的实现，这是由 Naimark[①]定理给出的[14]。

定理 2.5 (Naimark 定理) 对任意一个包含 n 个半正定算符 E_a，且满足 $\sum\limits_{a=1}^{n} E_a = \boldsymbol{I}$ 的 POVM，可以通过将希尔伯特空间扩展到一个更大的空间，并在更大的空间中进行投影测量来实现。

证明 通常来说，POVM 可以有不是秩为 1 的元素。但在这种情况下可以先对这些 POVM 元素执行谱分解，并将 POVM 视为一个只具有秩为 1 的元素的 POVM，然后将结果组合。因此，不失一般性，认为 POVM 元素 E_a 都是秩为 1 的。

让我们考虑一个希尔伯特空间 $\mathcal{H}, \dim \mathcal{H} = N$。POVM 可以由秩为 1 的正定算符，$\{E_a\}_a, a = 1, 2, \cdots, n, n \geqslant N$，来描述。$E_a$ 可以写成

$$E_a = |\psi_a\rangle\langle\psi_a| \tag{2.64}$$

其中，$|\psi_a\rangle\langle\psi_a|$ 一般对于 $\mathcal{H}$ 的一组正交基矢 $\{|i\rangle\}$ 是次归一化的（迹小于 1）。也就是说，$|\psi_a\rangle = \sum\limits_{i=1}^{N} \psi_{ai} |i\rangle$。那么基于 POVM 元素的归一性，有

① 由于苏联期刊翻译中使用的俄语罗马化，有时它也被称为 Neumark 定理。

$$\sum_{a=1}^{n}(E_a)_{ij}=\sum_{a=1}^{n}\psi_{ai}^{*}\psi_{aj}=\delta_{ij} \tag{2.65}$$

现在，把式 (2.65) 中的 $\sum\limits_{a=1}^{n}\psi_{ai}^{*}\psi_{aj}$ 分解为两个 n 维列向量 $\boldsymbol{\psi}_i$ 和 $\boldsymbol{\psi}_j$ 的内积，其中 $\boldsymbol{\psi}_i=(\psi_{1i},\psi_{2i},\cdots,\psi_{ni})^{\mathrm{T}}$, $\boldsymbol{\psi}_j=(\psi_{1j},\psi_{2j},\cdots,\psi_{nj})^{\mathrm{T}}$。由式 (2.65) 可知，这 N 个向量构成一组标准正交基。可以将这些向量扩展到 n 维空间的标准正交基，也就是说，可以找到 n 个 n 维列向量 $\boldsymbol{u}_i$，这样 $\forall a\in[n]$，有

$$\sum_{a=1}^{n}u_{ai}^{*}u_{aj}=\delta_{ij} \tag{2.66}$$

其中，$u_{ai}=\psi_{ai}$, $\forall a\in[n]$, $i\in[N]$。

很容易验证 (u_{ai}) 组成了一个幺正矩阵。因此，它的列构成一组标准正交基。注意，对于每一列，由前 N 项构成的向量与次归一化向量 $|\psi_a\rangle$ 一致。这表明，通过扩大希尔伯特空间，$\mathcal{H}'=\mathcal{H}\oplus\mathcal{H}^{\perp}$，可以构造 $\mathcal{H}'$ 的一组标准正交基 $|u\rangle$：

$$|u\rangle=\left|\tilde{\psi}\right\rangle+\left|\tilde{\psi}^{\perp}\right\rangle \tag{2.67}$$

其中，$\left|\tilde{\psi}\right\rangle,\left|\tilde{\psi}^{\perp}\right\rangle\in\mathcal{H}'$ 是由 $|\psi\rangle$ 和 $\left|\psi^{\perp}\right\rangle$ 分别在对方空间对应的位置添加 0 得到的扩大后的希尔伯特空间上的态。因此，POVM 被扩展为 PVM，这个 PVM 是由秩为 1 的算符 $\{|u_a\rangle\langle u_a|\}_a$ 给出的。 □

这个 Naimark 定理可以看作量子态纯化在测量中的类比，即总可以把一个一般的测量“纯化”成一个投影测量。这样，可以重新审视一下例 2.7中仪器所作的操作，实际上，由于甲完全不知道仪器做了怎样的操作，我们总可以假设仪器内部还有一个辅助空间 $\mathcal{H}_D$，处在量子态 $|+\rangle_D$ 上，这个仪器实际上是先对 $|+\rangle_D$ 进行了 Z 基矢的测量，根据测量的结果，1 的话就对输入的态进行 Z 基矢测量，-1 的话就进行 X 基矢测量。那么对于总体的态 $|+\rangle_D\otimes\rho$，所做的测量就是一组投影测量 $\{|00\rangle\langle 00|,|01\rangle\langle 01|,|1+\rangle\langle 1+|,|1-\rangle\langle 1-|\}$。

同样地，一般的测量也可以看作缺失了一部分信息的投影测量。如果甲知道了仪器对 $|+\rangle_D$ 的测量结果，那么她就确定地知道仪器对 ρ 进行了 Z 基矢测量还是 X 基矢测量，这种情况下，仪器做的测量在甲看来就是投影测量了。

在量子态中，还学习到与纯态的概率混合相对应的混态的纯态分解，例 2.7也表明，投影测量的概率混合可以是一个非投影测量。那么，是不是也可以类比一下式(2.32)，得出任何一个非投影测量都可以写成几个投影测量的混合的结论呢？有趣的是，答案是否定的。有不少非投影测量也不能写成其他测量的混合。有兴趣的读者可以参考 extremal POVM 相关文献[15-16]。

2.3.3 量子仪器

在很多情形中，不对测量算子的形式进行限制。根据测量公理，它们仅需要满足线性和完备性的要求。为了同时刻画观测者能够观测的经典结果和量子态演化后的系统状态，有时会用量子仪器的图像来描述量子测量过程。

在 2.3.2 节中，我们已经看到，如果只关注测量的经典结果，一个量子仪器可以对应于一个 POVM，而这一对应展示了 POVM 与投影测量之间的深刻联系。现在我们说明，对于一些问题而言，投影测量足以描述最一般的量子测量结果。考虑一组 POVM$\{M_m\}_m$ 作用于系统 $\mathcal{H}_A$ 上的量子态 ρ, 这个过程可以由图 2.2 描述的操作实现：一开始，一个辅助系统 $\mathcal{H}_E$ 被制备到量子态 $|0\rangle\langle 0|$；随后一个线性算子 U 作用在两体系统 $\mathcal{H}_A \otimes \mathcal{H}_E$ 上，

$$U(\rho \otimes |0\rangle\langle 0|)U^\dagger = \sum_m M_m \rho M_m^\dagger \otimes |m\rangle\langle m| \tag{2.68}$$

其中，$\{|m\rangle\}$ 构成了系统 $\mathcal{H}_E$ 的一组正交完备基，U 是一个等距变换（isometry）算子。可以证明，U 可以扩展为更大空间中的幺正算子（证明留作习题 2.12）。随后，如果对系统进行投影测量 $\{P_m\}$，其中投影算子为 $P_m = I_A \otimes |m\rangle\langle m|$，那么测量结果为 m 的概率为

$$\Pr(m) = \mathrm{tr}\big[P_m U(\rho \otimes |0\rangle\langle 0|)U^\dagger\big] = \mathrm{tr}\big(M_m \rho M_m^\dagger\big) \tag{2.69}$$

相应的系统末态演化为

$$\frac{P_m U(\rho \otimes |0\rangle\langle 0|)U^\dagger P_m^\dagger}{\mathrm{tr}[P_m U(\rho \otimes |0\rangle\langle 0|)U^\dagger]} = \frac{M_m \rho M_m^\dagger \otimes |m\rangle\langle m|}{\mathrm{tr}\Big(M_m \rho M_m^\dagger\Big)} \tag{2.70}$$

在对辅助系统 E 求偏迹后，我们可以看到，A 上的量子态与测量公理的描述是一致的。因此，这样一种利用投影测量的描述方法也可以看作对测量过程的一种等价描述。

在后面讨论量子信道的内容时，我们会说明，量子仪器可以理解为一个量子信道，而 Naimark 定理可以看作 Stinespring 延拓的一种特殊情况。

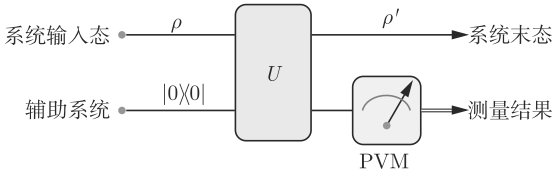

图 2.2　通过幺正算子和投影测量实现一个量子仪器。这里投影测量可以是秩为 1 的，$\{|m\rangle\langle m|\}$。一般情况下，ρ 和 ρ' 的维度是不一样的

2.3.4 联合测量和贝尔态测量

在实际量子信息处理任务中，经常会涉及同时对两个或者多个子系统同时进行测量的任务。这种同时作用于多个系统的测量被称作联合测量（joint measurement）。其中一种重要的类型是贝尔态测量（Bell state measurement）。完备的双比特贝尔态测量是一种投影测量，其效果是将量子态投影到贝尔基中的某一个。可以用 Hadamard 操作和 CNOT 操作来实现贝尔态测量。当 CNOT 门作用于计算基矢态上时，可以很快得到：

$$\begin{cases} \mathrm{CNOT}\,|0\rangle\,|0\rangle = |0\rangle\,|0\rangle \\ \mathrm{CNOT}\,|0\rangle\,|1\rangle = |0\rangle\,|1\rangle \\ \mathrm{CNOT}\,|1\rangle\,|0\rangle = |1\rangle\,|1\rangle \\ \mathrm{CNOT}\,|1\rangle\,|1\rangle = |1\rangle\,|0\rangle \end{cases} \tag{2.71}$$

为了对贝尔态进行投影测量，考虑幺正矩阵

$$U = (H \otimes I)\mathrm{CNOT} \tag{2.72}$$

容易验证，当该矩阵作用于贝尔态时，

$$\begin{cases} U(|00\rangle + |11\rangle) = (H \otimes I)\,|+0\rangle = |00\rangle \\ U(|00\rangle - |11\rangle) = (H \otimes I)\,|-0\rangle = |10\rangle \\ U(|01\rangle + |10\rangle) = (H \otimes I)\,|+1\rangle = |01\rangle \\ U(|01\rangle - |10\rangle) = (H \otimes I)\,|-1\rangle = |11\rangle \end{cases} \tag{2.73}$$

这四种量子态可以用局域 Z 测量进行分辨，也就是计算基矢态。在实验中，如果要分辨一个贝尔态是贝尔基中的哪一个，可以使用上面这一过程来完成态区分的任务。贝尔态测量可以用图 2.3所示的量子线路实现。

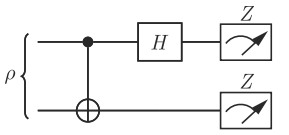

图 2.3 贝尔态测量的实现

2.4 一般的量子操作

在量子信息的语言中，量子操作、量子信道、量子态演化本质上都是同一件事——将一个量子态映射到另一个量子态：

$$\Lambda(\rho) = \rho' \tag{2.74}$$

一般来说，ρ 和 ρ' 的维度不一定相同。对于映射 Λ，它应该满足什么样的数学性质？直观上，它至少需要满足下面 3 个条件：

（1）线性: $\Lambda(\rho_1 + \rho_2) = \Lambda(\rho_1) + \Lambda(\rho_2)$；

（2）保正定性: $\forall \rho \geqslant 0, \Lambda(\rho) \geqslant 0$；

（3）保迹: $\mathrm{tr}(\Lambda(\rho)) = \mathrm{tr}(\rho)$。

线性的要求来自量子力学关于量子态演化的基本公理。稍后通过 Stinespring 扩展会更加深入地理解这一性质。由于演化是真实且物理的，只要映射的原象是量子态，映射的象也应该是量子态——这反映在另外的两个性质上。但有趣的是，保正定性并不是一个操作是量子演化的充分条件。接下来我们要说明，如果一个映射对应于一个实际的物理操作，它必须是完全正定（completely positive）的。

2.4.1 量子信道

量子信道是一个可以传输量子和经典信息的通信通道。在日常中，经典信息可以通过文档的形式在互联网上传输，这便可以看作信道的一个例子。在量子信息中，量子信道的严格定义由两个算子空间之间的完全正定保迹映射所描述。

为了数学内容的完整性，首先介绍 C^* 代数（C-star algebra），这是量子信道理论的代数基础。

定义 2.5 (C^* 代数)　一个 C^* 代数 A 是复数域上定义了映射 $*$ 的 Banach 代数。代数映射原象集中的一个元素 x 的象记为 x^*，映射 $*$ 具有下述性质：

（1）$*$ 是对合映射（involution），即对于代数 A 中的任一元素 x，都有 $x^{**} = (x^*)^* = x$；

（2）对于代数 A 中的任意元素 x, y: $(x+y)^* = x^* + y^*, (xy)^* = y^* x^*$；

（3）对于任一复数 $\lambda \in \mathcal{C}$ 和代数 A 中的任一元素 x: $(\lambda x)^* = \overline{\lambda} x^*$, 其中 $\overline{\lambda}$ 表示 λ 的复共轭；

（4）对于代数 A 中的任一元素 x: $\|x^* x\| = \|x\| \|x^*\|$, 其中 $\|\cdot\|$ 是代数 A 上定义的范数。

在接下来的讨论中，不会进一步讨论抽象的 C^* 代数。事实上，量子信息理论中的许多问题可以限定在下面这个具体的 C^* 代数情境中：在定义欧几里得距离的 n 维复数线性空间 $\mathbb{C}^n$ 上，如果考虑作用于上述所有 $n \times n$ 维矩阵所定义的矩阵空间，并选取矩阵空间上的矩阵范数 $\|\cdot\|$, 那么全部 $n \times n$ 维复矩阵将张成一个 C^* 代数。在本书中，将简单地把 C^* 代数看成上述复数域上的矩阵空间。在这个代数上，首先定义关于映射正定性质的一些基本概念，包括正定映射（positive map）、k-正定映射（k-positive map）和完全正定映射（completely positive map）。

定义 2.6 (正定映射、k-正定映射、完全正定映射)　考虑两个 C^* 代数 A, B

及其之间一个线性映射 $\Lambda : A \to B$，可以自然地导出另一个映射，$id_k \otimes \Lambda : \mathbb{C}^{k\times k} \otimes A \to \mathbb{C}^{k\times k} \otimes B$，其中 id_k 是空间 $\mathbb{C}^{k\times k}$ 上的恒等变换。

- Λ 被称作是正定的，如果 Λ 将 A 中的任一半正定算子映射到 B 中的半正定算子，$\forall a \geqslant 0 \Rightarrow \Lambda(a) \geqslant 0$。
- Λ 被称作是 k-正定的，如果 $id_k \otimes \Lambda$ 是一个正定映射。
- Λ 被称作是完全正定的，如果对于任意正整数 k，$id_k \otimes \Lambda$ 都是正定映射。

有了上面关于映射正定性的概念，我们给出量子信道的数学定义：

定义 2.7 (量子信道)　量子信道是两个线性算子空间上的线性完全正定保迹（completely positive and trace preserving, CPTP）映射。

量子信道又被称为“超算子”（super operator）。“超”代表该映射的作用对象是算子，而不是在希尔伯特空间上的向量。为什么要考虑定义 2.6中导出的映射？实际物理实验中，所实际关心的系统可能是一个更大系统的一部分。不失一般性，总可以考虑所观测系统的纯化系统。按照量子力学公理，在整个系统上的演化应该满足幺正变换。一般来说，对于作用在 $\mathcal{H}_A \otimes \mathcal{H}_B$ 上的幺正算子，如果仅关注它作用在 $\mathcal{H}_A$ 的效果，它对 $\mathcal{H}_A$ 中元素的作用效果不一定是幺正变换。事实上，A 系统上的演化可以描述为一个量子信道。

对于量子信道 $\Lambda(\cdot)$，可以将其表示为算子求和的形式：

$$\begin{cases} \Lambda(\rho) = \sum\limits_i F_i \rho F_i^\dagger \\ \sum\limits_i F_i^\dagger F_i = I \end{cases} \tag{2.75}$$

其中，F_i 被称作信道 Λ 的 Kraus 算子（Kraus operator）。

此外，还有一种信道的表示方式——蔡矩阵（Choi matrix）。信道 $\Lambda : \mathbb{C}^{n\times n} \to \mathbb{C}^{m\times m}$ 所对应的蔡矩阵由下式给出：

$$\boldsymbol{J}_\Lambda = (id_{\mathbb{C}^{n\times n}} \otimes \Lambda)\left(\sum_{i,j} |i\rangle\langle j| \otimes |i\rangle\langle j|\right) = \sum_{i,j} |i\rangle\langle j| \otimes \Lambda(|i\rangle\langle j|) \tag{2.76}$$

这里的蔡矩阵 $\boldsymbol{J}_\Lambda \in \mathbb{C}^{n\times n} \otimes \mathbb{C}^{m\times m}$ 记录了如何将一个 $\mathbb{C}^{n\times n}$ 的矩阵元素映射到 $\mathbb{C}^{m\times m}$。$\boldsymbol{J}_\Lambda$ 的秩被称作信道 Λ 的蔡秩（Choi rank）。我们也会将一个归一化蔡矩阵称作蔡态（Choi state）：

$$\chi_\Lambda = \frac{1}{n} J_\Lambda \tag{2.77}$$

可以看到这个矩阵满足量子态的所有条件。事实上，对于给定的一个信道 Λ，可以通过下面的线路图（图 2.4）产生蔡态。具体的证明留作习题 2.13。

下面的定理可以帮助我们判断一个映射是否是完全正定的。

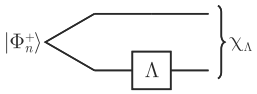

图 2.4　由信道 Λ 产生蔡态，其中 $|\Phi_n^+\rangle$ 为 n 维希尔伯特空间上的最大纠缠态

定理 2.6 (蔡文端定理)[17]　考虑一个正定映射 $\Lambda: C^{n\times n} \mapsto C^{m\times m}$，下面的几个叙述是相互等价的：

（1）Λ 是 n 正定的；

（2）Λ 是完全正定映射；

（3）由 Λ 通过式(2.76)定义的矩阵 $\boldsymbol{J}_\Lambda$ 是半正定的。

有了蔡矩阵我们可以计算信道作用后的结果 $\Lambda(\boldsymbol{\rho})$。不过，这种作用不能直接计算，而需要对蔡矩阵和量子态的密度矩阵进行一些预处理（张量指标重排，tensor index realignment）：

$$\begin{cases} \boldsymbol{J}'_\Lambda = \sum\limits_{i,j} \langle ij| \otimes \Lambda(|i\rangle\langle j|) \\ \boldsymbol{\rho}' = \sum\limits_{i,j} \rho_{ij} |ij\rangle \end{cases} \tag{2.78}$$

那么我们可以用向量乘法得到作用后的量子态 $\Lambda(\boldsymbol{\rho})$：

$$\boldsymbol{J}'_\Lambda \boldsymbol{\rho}' = \sum_{i,j} \rho_{ij} \Lambda(|i\rangle\langle j|) = \Lambda(\boldsymbol{\rho}) \tag{2.79}$$

如果要更直观地理解这个过程，我们用第 0 章中所提到的张量网络的方法来表示，如图 2.5所示。这里为了清晰地表示作用的过程，把原像空间记为 A，像空间记为 B。蔡矩阵可以看作四指标张量，其元素为 $(\boldsymbol{J}_\Lambda)_{j_A j_B}^{i_A i_B}$，张量上标 i_A, i_B 为行指标，下标 j_A, j_B 为列指标，满足 $i_A, j_A \in [n]$ 和 $i_B, j_B \in [m]$。密度矩阵 $\boldsymbol{\rho} \in \mathbb{C}^{n\times n}$ 是二指标张量，其元素为 $(\boldsymbol{\rho})_{j_A}^{i_A}$。首先将 $\boldsymbol{\rho}$ 向量化，变为一个 $n^2 \times 1$ 的列向量，其元素为 $(\boldsymbol{\rho})_{i_A j_A}$。这种向量化的方法对应于按顺序将密度矩阵的每一行进行转置后从上到下排列为一个列向量。同时，需要将蔡矩阵中的元素重排，变为 $(\boldsymbol{J}_\Lambda)_{i_A j_A}^{i_B j_B}$，这样蔡矩阵就由原来的 $nm \times nm$ 矩阵变为了 $m^2 \times n^2$ 矩阵。将这个矩阵左乘向量化后的密度矩阵，可以得到 $m^2 \times 1$ 的新列向量，即

$$(\Lambda(\boldsymbol{\rho}))_{i_A j_A} = \sum_{i_A j_A} (\boldsymbol{J}_\Lambda)_{i_A j_A}^{i_B j_B} (\boldsymbol{\rho})_{i_A j_A} \tag{2.80}$$

$(\Lambda(\boldsymbol{\rho}))_{i_A j_A}$ 就是 $\Lambda(\boldsymbol{\rho})$ 按同样的方法进行向量化后的结果。

思考题 2.7　其实还有另外一种对密度矩阵进行向量化的方式，即直接将每一列按顺序从上到下排列得到列向量，$(\rho)_{j_A}^{i_A} \to (\rho)_{j_A i_A}$。在这种向量化的方法下试着写出蔡矩阵重排后的元素及它们的张量网络表示。

量子信道还有多种其他表示方法，比如 Pauli-Liouville 表示、chi 矩阵表示等。不同的表示方法在不同的量子问题里面具有优势，这里就不一一展开讨论。一般来讲，上述 Kraus 算子和蔡矩阵是最常用最基本的两种表示方法。在大多数情况下，不区分量子态的演化和量子态经过量子信道作用这两种表述方法。

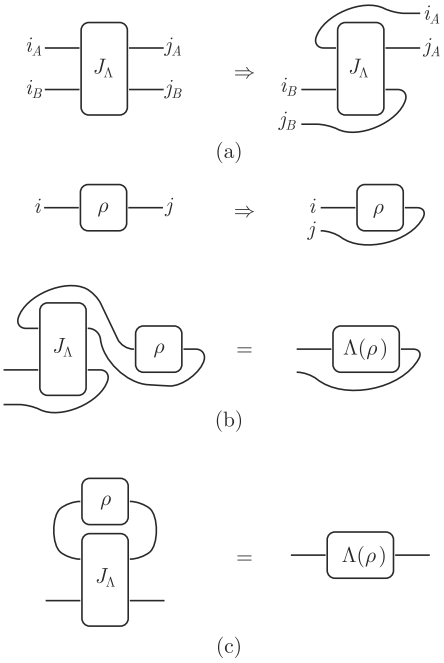

图 2.5　蔡矩阵作用的张量网络表示：（a）表示蔡矩阵 $\boldsymbol{J}_\Lambda$ 和密度矩阵 $\boldsymbol{\rho}$ 指标重排的过程；（b）$\boldsymbol{J}_\Lambda$ 和 $\boldsymbol{\rho}$ 指标重排后的乘积；（c）对应由（b）给出的矩阵指标重排反映了蔡矩阵和信道作用的关系

2.4.2　主方程

量子信道本质上是一种将开放量子系统动力学演化统一到封闭系统幺正演化的描述方式。在量子信息之外，比如量子光学和凝聚态领域，开放系统动力学演化问题得到了长时间的关注和研究。由于任何真正的量子系统都不是绝对封闭的，因而会与环境产生相互作用，从而导致衰减（decay）和退相干（decoherence）现象。因此，对于开放量子系统，不仅要关注量子系统本身，还要考虑环境对系统的影响，因此上述的描述方式就不合适了。为了解决开放量子系统的问题，需要寻找对于非幺正变换开放系统动力学过程的合适的微分方程描述，现在已经发展出了很多数学上等价，但在物理操作含义上具有不同特点的描述方式，其中一种便是使用密度矩阵及其对应的*主方程*（master equation）。其中，Lindblad 方程是最常用的主方程之一，用来描述密度矩阵的含时演化，其具体形式如下：

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{i}{\hbar}[H,\rho] + \sum_j [2L_j\rho L_j^\dagger - \{L_j^\dagger L_j, \rho\}] \tag{2.81}$$

其中，$\{x,y\} = xy + yx$ 表示反对易子；H 是系统哈密顿量，刻画了封闭系统的动力学演化，代表了系统相干（coherent）演化的部分, 从数学上来说，H 是一个厄米算子；L_j 是 Lindblad 算子，反映了系统与环境的耦合相互作用。关于这个问题的更多描述参见文献 [18] 的 8.4.1 节。

2.4.3 Stinespring 延拓

类似于量子态的纯化，可以将量子信道进行纯化。这一过程被称作量子信道的等距变换延拓（isometric extension），或称为 Stinespring 延拓（Stinespring dilation）。需要说明的是，Stinespring 延拓不是唯一的信道延拓方式。在不同的场景中，其他延拓方式可能会更加方便。

定理 2.7 (Stinespring 延拓)　对任意一个量子信道，$\Lambda : \mathcal{D}(\mathcal{H}_A) \mapsto \mathcal{D}(\mathcal{H}_B)$，和一个维度为 Λ 的蔡秩的希尔伯特空间 $\mathcal{H}_E$, 存在一个等距变换，$U : \mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_E$，满足 $\forall \rho_A \in \mathcal{D}(\mathcal{H}_A)$，

$$\mathrm{tr}_E(U\rho_A U^\dagger) = \Lambda(\rho_A) \tag{2.82}$$

等距变换 U 定义为 $U^\dagger U = I_A$。

证明　考虑量子信道的 Kraus 表示，$\Lambda(\rho) = \sum_{i=1}^{r} K_i \rho K_i^\dagger$, 其中 r 是 Λ 的蔡秩。考虑下面的算子：

$$\tilde{U} = \sum_{i=1}^{r} K_i \otimes |0\rangle\langle 0|^R + I^A \otimes \sum_{j=1}^{r-1} |j\rangle\langle j|^R \tag{2.83}$$

其中，K_i 作用在系统 A 上，这里引入了一个 r 维辅助系统 R。容易验证 $\tilde{U}$ 是等距变换：

$$\begin{aligned}
\tilde{U}^\dagger\tilde{U} = &\left(\sum_{i=1}^{r} K_i^\dagger \otimes |0\rangle\langle 0|^R + I^A \otimes \sum_{j=1}^{r-1} |j\rangle\langle j|^R\right)\cdot \\
&\left(\sum_{i=1}^{r} K_i \otimes |0\rangle\langle 0|^R + I^A \otimes \sum_{j=1}^{r-1} |j\rangle\langle j|^R\right) \\
= &\sum_{i=1}^{r} (K_i^\dagger K_i) \otimes |0\rangle\langle 0|^R + I^A \otimes \sum_{j=1}^{r-1} |j\rangle\langle j|^R \\
= &\, I^A \otimes |0\rangle\langle 0|^R + I^A \otimes \sum_{j=1}^{r-1} |j\rangle\langle j|^R
\end{aligned}$$

$$= I^A \otimes I^R \tag{2.84}$$

由此定义出来的映射，$U'(\rho) := \tilde{U}(\rho \otimes |0\rangle\langle 0|)\tilde{U}^\dagger$，我们有 $\Lambda(\rho) = \mathrm{tr}_R(U'(\rho))$。这样，便完成了证明。 □

在这个定理中，等距变换 U 等价于引入辅助系统后，联合作用于 $\mathcal{H}_A$ 和辅助系统的一个幺正算子。不难看出，对量子信道的延拓就像对量子态的纯化，总可以通过引入辅助系统来寻找更大的空间，使得研究对象在这个空间中是 “纯” 的，可以参考 2.3.2 节的 Naimark 定理。

2.5 带有噪声的量子演化

实际实验中，量子态的演化过程与理想过程总会有所偏差。对于实验观测者而言，他所能观测的量子系统演化不会是封闭系统的幺正演化。实际上，观测者的观测也会对系统产生影响。在这一节，我们从量子系统信息丢失的角度来解释演化中的噪声影响。

2.5.1 随机幺正演化导致的演化过程

我们从一个具体的例子开始讨论——量子比特翻转信道（quantum bit-flip channel）。考虑实验中制备一个量子态 $|\psi\rangle$ 的过程。由于噪声的影响，最终制备出来的量子态以一定概率发生了量子比特翻转，即制备结果有 $1-p$ 的概率得到预期目标态 $|\psi\rangle$，有 p 的概率得到 $\boldsymbol{X}|\psi\rangle$，这里 $\boldsymbol{X}$ 是泡利矩阵。这样的量子态可以由一个混态来描述：

$$(1-p)|\psi\rangle\langle\psi| + p\boldsymbol{X}|\psi\rangle\langle\psi|\boldsymbol{X}^\dagger \tag{2.85}$$

一般地，一个量子比特翻转信道可以表示为

$$\rho \mapsto (1-p)\rho + p\boldsymbol{X}\rho\boldsymbol{X}^\dagger \tag{2.86}$$

这一具体噪声模型可以推广到更一般的随机幺正演化，即按照一定的概率，一个量子态 ρ 经历了一个集合中某一个幺正算子的作用，$\{p_k, U_k\}$。容易验证，作用后的量子系统对应的密度矩阵为

$$\boldsymbol{\rho} \mapsto \sum_k p_k U_k \rho U_k^\dagger \tag{2.87}$$

这是最常见的量子信道形式。下面我们介绍一些常用的以幺正演化混合出来的量子信道。

与刚刚提到的量子比特翻转信道类似，另一个重要的信道模型是**量子相位翻转信道**（quantum phase-flip channel）。相位翻转，顾名思义就是将量子态 $|+\rangle$ 变成 $|-\rangle$，而 $|-\rangle$ 变成 $|+\rangle$，也就是泡利 Z 操作。对于该信道的表示，只要将式(2.86)中的 X 替换成 Z：

$$\rho \mapsto (1-p)\rho + pZ\rho Z \tag{2.88}$$

计算基矢态上的相位翻转对应于其共轭基矢态上的比特翻转。

上面介绍的比特翻转和相位翻转属于一类更一般的信道模型——**泡利信道**（Pauli channel）。这类信道的作用效果是对量子态按照一定概率分布随机进行泡利算子旋转。对于 n-qubit 系统，泡利信道可以表示为

$$\rho \mapsto \sum_{i\in\{0,1\}^{2n}} p_i P_i \rho P_i^\dagger \tag{2.89}$$

更一般地，这里 P_i 是泡利群元素。泡利群是由泡利算符作为生成元生成的群。作用在 n 个量子比特上的泡利群 $\mathcal{G}_n$ 由泡利算子 $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}$ 的 n 次张量积加上相位因子 $\pm 1, \pm \mathrm{i}$ 构成：

$$\mathcal{G}_n = \{\pm 1, \pm \mathrm{i}\} \otimes \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n} \tag{2.90}$$

从式(2.89)不难看出，泡利群元素的相位因子在这里是不起任何作用的，所以实际会用到的元素一共就只有 4^n 个，这也对应于角标 i 的取值范围 $i \in \{0,1\}^{2n}$。更进一步地，如果将这个式子与 2.2.4 节中提到的量子比特层析术的式子进行比较，不难看出，二者十分相似。实际上，对于泡利信道，完全可以像量子比特那样，用层析术来得到对这个信道的具体描述。更一般地，对于一个 d 维的希尔伯特空间，也可以用推广后的泡利矩阵来定义泡利信道。

思考题 2.8 一般的泡利信道，如果对于所有 i，$p_i = d^{-2}$，即所有泡利操作等概率出现，那么不管输入什么量子态，输出都是最大混态，$\boldsymbol{I}/d$，这里 d 是希尔伯特空间的维度，$\boldsymbol{I}$ 是单位矩阵。

对于二维情形，泡利信道可以表示为

$$\rho \mapsto \sum_{i,j=0}^{1} p_{ij} \sigma_z^i \sigma_x^j \rho \sigma_x^j \sigma_z^i \tag{2.91}$$

对此，我们来考虑一个特殊情况，$p_{00} = 1-p$ 且所有其他 $p_{ij} = p/3$，这时称为**去极化信道**（depolarising channel），表示为

$$\rho \mapsto (1-p)\rho + p\frac{\boldsymbol{I}}{2} \tag{2.92}$$

在这一信道模型中，我们以一定概率完全丢失初始量子态的全部信息，即按照一定的概率将初始量子态替换为最大混态。这个信道模型可以很自然地推广到 d 维的情形。

2.5.2 信息丢失导致的演化过程

从实验观测的角度，噪声也可以看作测量信息丢失导致的结果。考虑对初始量子态 ρ，进行了 POVM 测量 $\{E_k\}$，其中测量算子为半正定算子并满足完备性条件 $\sum_k E_k = I$, 对应于 Kraus 算子为 $E_k = M_k^\dagger M_k$。根据量子测量公理，测量结果为 k 的概率为

$$p_k = \mathrm{tr}\Big(M_k \rho M_k^\dagger\Big) \tag{2.93}$$

相应的测量末态为

$$\frac{M_k \rho M_k^\dagger}{\mathrm{tr}\Big(M_k \rho M_k^\dagger\Big)} \tag{2.94}$$

但如果丢失了测量结果 k 这一信息，那么测量后的系统将变成服从一定概率分布的系综，也就是混态。相应的密度矩阵将变成

$$\sum_k p_K(k) \frac{M_k \rho M_k^\dagger}{p_K(k)} = \sum_k M_k \rho M_k^\dagger \tag{2.95}$$

可以将这样一个演化结果表示为一个带噪声的信道 $\mathcal{N}(\rho)$：

$$\mathcal{N}(\rho) = \sum_k M_k \rho M_k^\dagger \tag{2.96}$$

需要说明的是，虽然将式(2.96)解释为量子测量结果丢失，这一表示方法实际上代表了一个密度矩阵的一般演化过程，其中 M_k 就是前面所描述的 Kraus 算子。事实上，任意的带有噪声的演化过程都可以表示为式(2.96)的形式。另外，密度矩阵 $\boldsymbol{\rho}$ 的演化应该是保迹的，因为演化结果也是密度矩阵，迹为 1：

$$\begin{aligned} &\mathrm{tr}[\mathcal{N}(\boldsymbol{\rho})] \\ &= \mathrm{tr}\left(\sum_k M_k \boldsymbol{\rho} M_k^\dagger\right) \\ &= 1 \end{aligned} \tag{2.97}$$

对于一个冯•诺依曼测量，如果测量结果丢失了，称为退相干信道（dephasing channel）。考虑 d 维系统，选取计算基矢为 $|0\rangle, \cdots, |n-1\rangle$, 退相干信道可以表示为

$$\rho \mapsto \sum_{i=0}^{d-1} |i\rangle\langle i| \, \rho \, |i\rangle\langle i| \tag{2.98}$$

擦除信道（erasure channel）是另一种重要信息丢失模型。它有简单的描述方式，并将在稍后研究量子信道容量问题时发挥重要作用。在光学实验中，擦除信道也是一种刻画光子传输损耗的简化模型。一个经典擦除信道以 $0 \leqslant 1-\varepsilon \leqslant 1$ 的概率如实传输一个比特，以 ε 的概率将其替换为一个擦除标记 e。信道输出结果比输入的字母表要多一个字符，即擦除字符 e。将擦除信道进行量子推广是非常直接的，它可以表示为

$$\rho \mapsto (1-\varepsilon)\rho + \varepsilon \left|e\right\rangle\!\left\langle e\right| \tag{2.99}$$

其中，擦除标记 $|e\rangle$ 不在原先的态空间内，$|e\rangle\langle e| \perp \mathrm{supp}(\rho)$。

在量子通信理论分析中，一种具有重要作用的信道理论模型是经典至量子信道（classical-to-quantum channel），也被称为**经典-量子信道**（classical-quantum channel）。经典-量子信道的作用效果是，首先将输入的量子态在某一指定的正交归一基上进行投影测量，再根据测量结果，制备并输出一个量子密度矩阵。假设信道的输入是一个密度矩阵 $\boldsymbol{\rho}$, 它所作用的希尔伯特空间有一正交归一基 $\{|k\rangle\}$, 经典-量子信道首先将输入量子态在这一基上进行测量。对于测量结果 k, 相应的测量末态为

$$\frac{\left|k\right\rangle\!\left\langle k\right| \boldsymbol{\rho} \left|k\right\rangle\!\left\langle k\right|}{\left\langle k\right| \boldsymbol{\rho} \left|k\right\rangle} \tag{2.100}$$

经典-量子信道将这一测量末态与一个密度矩阵 $\boldsymbol{\sigma}_k$ 关联起来：

$$\frac{\left|k\right\rangle\!\left\langle k\right| \boldsymbol{\rho} \left|k\right\rangle\!\left\langle k\right|}{\left\langle k\right| \boldsymbol{\rho} \left|k\right\rangle} \otimes \boldsymbol{\sigma}_k \tag{2.101}$$

随后信道对第一个系统取偏迹运算，只将第二个系统输出。这样，信道的作用效果可以描述为

$$\mathcal{N}(\boldsymbol{\rho}) = \sum_k \left\langle k\right| \boldsymbol{\rho} \left|k\right\rangle \boldsymbol{\sigma}_k \tag{2.102}$$

2.5.3 去极化信道

作为这一章内容的例子，我们从不同角度解释去极化信道。考虑一个二维去极化信道 Λ, 它以概率 $1-p$ 如实传输一个量子比特，以 p 的概率发生错误。指定计算基矢态为 $\{|0\rangle, |1\rangle\}$，传输过程中可能发生三种错误：

（1）比特翻转错误: $|\psi\rangle \mapsto \sigma_x |\psi\rangle$。

（2）相位翻转错误: $|\psi\rangle \mapsto \sigma_z |\psi\rangle$。

（3）同时发生比特翻转和相位翻转错误: $|\psi\rangle \mapsto \sigma_y |\psi\rangle$。

对于去极化信道，可以理解为这三种错误以相同的概率发生。

Stinespring 延拓: 考虑信道作用在量子比特空间 A 上。它可以通过在一个更大的空间 $\mathcal{H}_A \otimes \mathcal{H}_E$ 上进行幺正变换来实现，其中 E 是环境空间，从而满足

$$U(|\psi\rangle_A \otimes |0\rangle_E) = \sqrt{1-p}\,|\psi\rangle_A \otimes |0\rangle_E + \sqrt{p/3}(\sigma_x |\psi\rangle_A \otimes |1\rangle_E + \sigma_y |\psi\rangle_A \otimes |2\rangle_E + \sigma_z |\psi\rangle_A \otimes |3\rangle_E) \tag{2.103}$$

环境的演化同时记录了系统 A 中的变化。

Kraus 算子: 通过下面的 Kraus 算子：

$$\begin{cases} M_0 = \sqrt{1-p}I \\ M_1 = \sqrt{\dfrac{p}{3}}\sigma_x \\ M_2 = \sqrt{\dfrac{p}{3}}\sigma_y \\ M_3 = \sqrt{\dfrac{p}{3}}\sigma_z \end{cases} \tag{2.104}$$

（作为习题 2.9，请检查这些算子满足完备性条件）信道 Λ 的作用效果为

$$\Lambda(\rho) = \sum_{i=0}^{3} M_i \rho M_i^\dagger \tag{2.105}$$

Choi 算子: 考虑一个在系统 A 和辅助系统 R 上的最大纠缠态 $|\Phi^+\rangle$。其中，辅助系统 R 保持不变，系统 A 经历了信道 Λ 的作用。对于整个联合系统，它的演化结果为

$$\mathcal{I} \otimes \Lambda(\left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|) = (1-p)\left|\Phi^+\right\rangle\!\left\langle\Phi^+\right| + \frac{p}{3}\left(\left|\Phi^-\right\rangle\!\left\langle\Phi^-\right| + \left|\Psi^+\right\rangle\!\left\langle\Psi^+\right| + \left|\Psi^-\right\rangle\!\left\langle\Psi^-\right|\right) \tag{2.106}$$

当 $p = 3/4$ 时，末态演化为最大混态 $I/4$。

布洛赫球: 考虑经历信道作用前的量子比特为

$$\rho = \frac{I + r \cdot \sigma}{2} \tag{2.107}$$

在演化后，量子态变为

$$\begin{cases} \rho = \dfrac{I + r' \cdot \sigma}{2} \\ r' = \left(1 - \dfrac{4p}{3}\right) r \end{cases} \tag{2.108}$$

在布洛赫球的几何图像中，这个量子态的布洛赫向量“缩短”了。

习题

习题 2.1 (偏迹)

(1) 假设 Alice 和 Bob 共享一个量子系统 ρ^{AB}。考虑 Alice 可能在她的系统上进行的一个局域测量，其测量算符为 $\{M_m\}_m$。因此全局测量算符是 $\{M_m^A \otimes I^B\}_m$。证明全局密度矩阵所预测的概率与局域密度矩阵 ρ^A 所预测的概率是相同的。

$$\mathrm{tr}\big[(M_m^A \otimes I^B)\rho^{AB}\big] = \mathrm{tr}\big(M_m^A \rho^A\big) \tag{2.109}$$

因而，全局量子理论的预测与局域量子理论的预测是一致的。

(2) 证明如果 Bob 在没有通知 Alice 测量结果的情况下对他的系统执行幺正操作或测量，Alice 的局域密度矩阵不会改变。

习题 2.2 (偏迹与态分解相互对易) 对于两体量子态 ρ^{AB}，子系统 A 的状态由求偏迹后的态 $\rho^A = \mathrm{tr}_B(\rho^{AB})$ 给出。对于联合态的任意分解，$\rho_i^{AB} = \sum_i p_i \rho_i^{AB}$，其中 $p_i > 0$ 及 $\sum p_i = 1$，尝试将 ρ_A 分解为具有相同混合概率 $\{p_i\}$ 的态的叠加。反过来呢？已知 ρ^A 的态分解，能把 ρ^{AB} 分解成以一定概率混合的态吗？

习题 2.3 (密度矩阵的性质) 试证明由公理 2.1给出的密度矩阵 $\boldsymbol{\rho}$ 有如下性质：

(1) 厄米性：$\boldsymbol{\rho}^\dagger = \boldsymbol{\rho}$。

(2) 半正定：$\boldsymbol{\rho} \geqslant 0$。

(3) 归一化：$\mathrm{tr}(\boldsymbol{\rho}) = 1$。

习题 2.4 (量子态的纯化)

(1) 找出拥有如下形式的谱分解的态 ρ 的一个纯化：

$$\rho = \sum_i \lambda_i \left|i\right\rangle\!\left\langle i\right| \tag{2.110}$$

并证明所有的纯化都可以通过在参考系上的幺正操作互相联系起来。

(2) 找到以下经典-量子态的纯化：

$$\rho = \sum_x p(x) \left|x\right\rangle\!\left\langle x\right| \otimes \rho_x \tag{2.111}$$

习题 2.5 (不同纯化间的关系) 对于系统 A 的一个量子态 ρ_A，可以构造出 ρ_A 的两个不同的纯化 $\left|\Psi_1\right\rangle_{AB}$ 和 $\left|\Psi_2\right\rangle_{AB}$，试证明它们之间的关系可以由下式给出：

$$\left|\Psi_1\right\rangle_{AB} = (I_A \otimes U_B)\left|\Psi_2\right\rangle_{AB} \tag{2.112}$$

即这两个态的差异可以由一个单独作用于 $\mathcal{H}_B$ 的幺正变换给出。

习题 2.6 (施密特分解) 令 $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ 为复合系统 AB 上的一个纯态。在 A 和 B 的正交基矢上表示 $|\psi\rangle$ 为

$$|\psi\rangle = \sum_{i,j=1}^{d_A,d_B} c_{ij} |a_i\rangle |b_j\rangle \tag{2.113}$$

可以看出 $|\psi\rangle$ 的施密特分解为

$$|\psi\rangle = \sum_{k=1}^{d} \lambda_k |\alpha_k\rangle |\beta_k\rangle \tag{2.114}$$

其中，$d \leqslant \min\{d_A, d_B\}$ 是施密特数。

(1) 从矩阵 $\boldsymbol{C}\boldsymbol{C}^\dagger$ 中计算 λ_k，其中 $\boldsymbol{C} = (c_{ij})$ 是由式(2.113)中的系数组成的矩阵。

(2) 证明如果 λ_k 互不相同（非简并），两个系统的基矢 $|\alpha_k\rangle$ 和 $|\beta_k\rangle$ 对任意的 k 都可以用同一个等距变换联系起来。

习题 2.7 (施密特数的性质) 假设 $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ 是由系统 A 和 B 构成的复合系统上的一个纯态，

(1) 证明 $|\psi\rangle$ 的施密特数 $\mathrm{Sch}(\psi)$ 等于约化密度矩阵 $\boldsymbol{\rho}_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|)$ 的秩（注意厄米算子的秩等于其支撑集的维数）。

(2) 假设 $|\psi\rangle = \sum\limits_j |\alpha_j\rangle |\beta_j\rangle$ 是 $|\psi\rangle$ 的一个表示，其中 $|\alpha_j\rangle$ 和 $|\beta_j\rangle$ 分别是系统 A 和 B 上（未归一化）的态。证明这样一个分解中的项数大于或等于施密特数 $\mathrm{Sch}(\psi)$。

(3) 假设 $|\psi\rangle = \alpha |\phi\rangle + \beta |\gamma\rangle$，试证明

$$|\mathrm{Sch}(\phi) + \mathrm{Sch}(\gamma)| \geqslant \mathrm{Sch}(\psi) \geqslant |\mathrm{Sch}(\phi) - \mathrm{Sch}(\gamma)| \tag{2.115}$$

习题 2.8 (偏迹) 对于两个量子态 $\rho, \tau \in \mathcal{D}(\mathcal{H})$，它们之间的归一化迹距离定义如下：

$$d(\rho, \tau) = \frac{1}{2} \parallel \rho - \tau \parallel_1 \tag{2.116}$$

其中，$\| \cdot \|_1$ 表示 1 范数 —— 矩阵特征值的绝对值之和。

(1) 对于两个量子比特态 ρ, τ, 它们可以在布洛赫球上表示为

$$\begin{cases} \rho = \dfrac{1}{2}(I + r \cdot \sigma) \\ \tau = \dfrac{1}{2}(I + s \cdot \sigma) \end{cases} \tag{2.117}$$

证明它们的归一化迹距离可以由它们对应的布洛赫向量之间的欧几里得距离得到，即

$$2d(\rho, \tau) = \parallel \rho - \tau \parallel_1 = \parallel r - s \parallel_2 \tag{2.118}$$

（2）在一些量子信息处理任务，如量子态区分中，考虑迹距离的其他等价定义会让任务处理更加方便。其中一个常用的定义是最大概率差。对于两个量子态 $\rho,\tau\in\mathcal{D}(\mathcal{H})$，证明下述等式：

$$d(\rho,\tau)=\max_{0\leqslant\Lambda\leqslant I}\mathrm{tr}[\Lambda(\rho-\tau)] \tag{2.119}$$

Λ 是希尔伯特空间 $\mathcal{H}$ 上的半正定算子。在这里，$\mathcal{H}$ 的维度是固定的，但不一定是二维的。

习题 2.9 (去极化信道的 Kraus 算子)　试证明去极化信道的 Kraus 算子

$$\begin{cases} M_0=\sqrt{1-p}I \\ M_1=\sqrt{\dfrac{p}{3}}\sigma_x \\ M_2=\sqrt{\dfrac{p}{3}}\sigma_y \\ M_3=\sqrt{\dfrac{p}{3}}\sigma_z \end{cases} \tag{2.120}$$

满足归一化条件。

习题 2.10 (* 简化的量子边际问题[19])　量子边际问题描述如下: 给定一组局域密度矩阵 $\{\rho_j\}$，确定是否存在 n 体态 $\rho^{(n)}$，使 $\{\rho_j\}$ 为 $\rho^{(n)}$ 的约化密度矩阵。

（1）证明对任意给定的 d 维密度矩阵 $\boldsymbol{\rho}_A$，总存在一个两体量子态 ρ_{AB}，使得 $\mathrm{tr}_B(\rho_{AB})=\boldsymbol{\rho}_A$。

（2）考虑一个二量子比特态 ρ_{AB}，称 ρ_{ABC} 是 ρ_{AB} 的一种对称扩展，当且仅当其满足

$$\rho_{AB}=\mathrm{tr}_C(\rho_{ABC})=\mathrm{tr}_B(\rho_{ABC})=\rho_{AC} \tag{2.121}$$

证明当且仅当

$$\mathrm{tr}\left(\rho_B^2\right)\geqslant\mathrm{tr}\left(\rho_{AB}^2\right)-4\sqrt{\det(\rho_{AB})} \tag{2.122}$$

其中，$\rho_B=\mathrm{tr}_A(\rho_{AB})$ 时，ρ_{AB} 有对称扩展 ρ_{ABC}。

提示: 考虑这样的情况，有一个纯态的扩展 $|\psi_{ABC}\rangle$，使用量子比特 A,B 和量子比特 C 之间的施密特分解; 再证明 $\det(\rho_{AB})=0$。

习题 2.11　证明 Uhlmann 定理，即给定两个量子态 ρ 和 σ，证明

$$F(\rho,\sigma)\equiv\max_{|\psi\rangle,|\phi\rangle}|\langle\psi|\phi\rangle| \tag{2.123}$$

其中最大值在所有 ρ 的纯化 $|\psi\rangle$ 和 σ 的纯化 $|\phi\rangle$ 中取。

习题 2.12 考虑一组 POVM 测量 $\{M_m\}_m$ 作用于系统 $\mathcal{H}_A$ 上的量子态 ρ, 这个过程可以由下面的操作实现：一开始，一个辅助系统 $\mathcal{H}_E$ 被制备到量子态 $|0\rangle\langle 0|$。随后一个线性算子 U 作用在两体系统 $\mathcal{H}_A \otimes \mathcal{H}_E$ 上，

$$U(\rho \otimes |0\rangle\langle 0|)U^\dagger = \sum_m M_m \rho M_m^\dagger \otimes |m\rangle\langle m| \tag{2.124}$$

其中，$\{|m\rangle\}$ 构成了系统 $\mathcal{H}_E$ 的一组正交完备基，U 是一个等距变换 (isometry) 算子，具有保持向量内积不变的效果。证明，U 可以扩展为更大空间中的幺正算子（提示：尝试按照 Naimark 定理的分析方法证明该结论）。

习题 2.13 证明给定一个信道 Λ，可以用图 2.4 所示的量子线路产生对应的蔡态，其中 $|\Phi_n^+\rangle$ 为 n 维希尔伯特空间上的最大纠缠态。