

量子物理

阿笠博士

August 10, 2024

Contents

0	前言	4
0.1	可能需要的物理常数	5
1	波函数	6
1.1	定义	6
1.1.1	波函数是概率分布的体现	6
1.1.2	平方可积函数的性质	7
1.2	傅立叶变换*	7
1.2.1	动量表象	12
1.3	薛定谔方程	12
1.3.1	一维定态薛定谔方程的性质	15
1.3.2	一维定态薛定谔方程的解	16
1.4	晶体的热容*	27
1.5	热方程*	30
1.5.1	热方程的基本解*	31
1.5.2	传播子*	34
2	线性算子	36
2.1	定义	36
2.1.1	李代数*	38
2.2	伴随与厄米算子	39
2.2.1	测量	40
2.3	常用的线性算子	42
2.3.1	期望值与测不准原理	43
2.3.2	哈密顿算子*	46
2.4	向量场*	50
2.4.1	积分曲线*	52
2.4.2	流*	53
2.5	对偶空间	55

2.6	张量	58
2.6.1	张量场*	65
2.6.2	反对称张量*	66
2.7	微分形式*	73
2.7.1	\mathbb{R}^3 上的微分形式*	77
2.8	哈密顿系统*	79
2.8.1	守恒与对称*	82
2.9	狄拉克符号	86
2.9.1	线性映射的狄拉克符号表示	88
3	角动量与自旋	96
3.1	角动量算子	96
3.2	氢原子波函数	99
3.2.1	氢原子跃迁	100
3.3	自旋	103
3.3.1	四元数*	105
3.3.2	自旋波函数是 \mathbb{C}^2 上的向量	107
3.4	矩阵指数	110
3.4.1	矩阵李群*	120
3.4.2	自旋动力学	129
3.5	2个粒子的自旋态	130
3.6	氢原子的自旋-轨道耦合*	134
3.6.1	氢原子的自旋-轨道耦合	138
4	量子信息	140
4.1	量子比特	140
4.1.1	布洛赫球*	141
4.2	信息熵*	147
4.2.1	条件熵与联合熵*	154
4.2.2	密度矩阵*	159
4.3	量子门	163
4.4	双量子比特	164
4.4.1	多量子比特	171
4.5	量子力学受到的挑战	172
4.5.1	EPR佯谬	172
4.5.2	贝尔定理	173
4.6	不可克隆定理	176
4.7	纠缠态	177

5 例题	180
5.1 一维无限深方势阱	180
5.2 一维无限深方势阱	182
5.3 谐振子	184
5.4 谐振子	185
5.5 谐振子	185
5.6 动量表象、线性算子的期望值、测不准原理	186
5.7 狄拉克符号	189
5.8 氢原子能级	189
5.9 氢原子波函数	190
5.10 氢原子跃迁	191
5.11 泡利矩阵	192
5.12 自旋的测量	193
5.13 自旋体系的能量	194
5.14 自旋动力学	195
5.15 2个粒子的自旋态	197
5.16 自旋-自旋耦合	198
5.17 原子的磁矩	198
5.18 纠缠态	199
5.19 量子门	199

0 前言

众所周知，量子物理是中科大2020级(包含)以后所有信智学部(计算机、电子信息、大数据、网络安全、人工智能)的学生都绕不开的一道坎。毫无疑问，这门课程彻底激怒了所有信智学部的学生，但是无奈的是，培养方案在短期内根本没有改变的可能性。毕竟我也不是信智学部的学生，所谓的“正义感”显得我很虚伪，并且毫无作用：因为再怎么辱骂课程安排，你还是无法逃避。所以，我也只能在课程学习上给大家带来力所能及的帮助，至少不用完全依赖反人类的“官方资料”；此外，我也写了“官方资料”的习题答案，已经发表到评课社区了，[点击我就可以啦！ \$\epsilon = \epsilon = \(/ \geq \nabla \leq\) /\$](#) 这样能够节省大家的时间。

至于为什么量子物理这么令人反感，实际上，我在带这门课程的助教之前，我也没有什么概念，因为我本科阶段学的是杨焕雄老师授课的“量子力学A”，当时我并没有认为这门课程给我带来困难，而且信智学部的量子物理毫无疑问会简单很多。但是，带了这门课程的助教之后，看到了学校给出的量子物理所谓的“官方资料”，我就知道原因出在哪里了。至于我对“官方资料”的评论，在此不做展开，留给读者们自己练习，因为如果在此展开了，那么就不是几页纸能够放得下的了。

个人认为量子力学比较好的资料是Griffiths的Introduction to Quantum Mechanics，我当时购买了英文的原版书作为收藏。但是，其翻译的中文版，我自己虽然没有阅读过，但是我的朋友对它评价很差，因为它排版混乱、重点很不突出，所以我也不推荐大家购买。我阅读的是英文原版，虽然它逻辑非常清晰，但是，对于对这门课程根本不感兴趣的学生，花费大量时间阅读英文教材是根本不可能的。对于量子信息，比较好的资料是Thomas G. Wong的Introduction to Classical and Quantum Computing，当然了，也是英文的。至于我为什么列出这2份资料，主要原因有一些学生可能对这门课程有一定的兴趣，所以我也给他们拓展这方面兴趣的机会：如果你感兴趣，就阅读我列出的这2份资料，当然你要面对英文，但是幸运的是，英文难度很低。

但是，在写这份资料时，我遇到了瓶颈：毕竟我研究生也不会深究物理了，量子力学对我而言是完全没用的课程，在知识方面，我自然无法与学院的一些学生分庭抗礼；此外，我于2022年秋季学习量子力学A课程，时间较为久远，我已然力不从心。写“自旋-轨道耦合”这一部分花费了我相当长的时间，虽然这一部分看似简短，但是能够将这一抽象的知识点解释清楚是相当困难的。这一部分我没有找到非常好的参考资料，因为市面上的量子力学教材几乎都注重所谓的“物理解释”，而抛弃了数学的严谨。好在杨焕雄老师量子力学A的资料对其解释相当透彻，我仔细阅读了一下，因此后面还是将这一部分完成了。

不过，我并没有因为写这份资料而感到厌烦。我以前上学的时候，班级里有一些女生喜欢上课写一些小说，对我而言，写这份资料正如她们写小说一样，令我乐在其中：一方面，我能够让读者知道我对这门课程的理解程度，“展现自我”；另一方面，写资料就如同与读者对话，这也正好符合我内向的性格，毕竟我也不会在现实生活中与读者们相见。

最后，我非常感谢教授我理论力学A的潘海俊老师与量子力学A的杨焕雄老师。潘海俊老师是“艺术家”，他的课程节奏张弛有度，将理论力学的优雅之处演绎得淋漓尽致；杨焕雄老师则是“实干家”，每堂课程都会提供他亲自撰写的课程资料，他的资料给予我相当多的灵感，也是我写这份资料的原因之一——我并不要求获得名声与稿费，单纯只是致敬我的老师。

写到这里，就到此为止吧！

0.1 可能需要的物理常数

(A) 真空介电常数:

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ F} \cdot \text{m}^{-1}.$$

(B) 磁导率:

$$\mu_0 = 4\pi \times 10^{-7} \text{ N} \cdot \text{A}^{-1} = 1.26 \times 10^{-6} \text{ N} \cdot \text{A}^{-1}.$$

(C) 真空中的光速:

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3.00 \times 10^8 \text{ m/s}.$$

(D) 电子的电荷量:

$$e = 1.60 \times 10^{-19} \text{ C}.$$

(E) 电子伏特:

$$1 \text{ eV} = 1.60 \times 10^{-19} \text{ J}.$$

(F) 普朗克常数、约化普朗克常数:

$$\begin{aligned} h &= 6.63 \times 10^{-34} \text{ J} \cdot \text{s} = 4.14 \times 10^{-15} \text{ eV} \cdot \text{s}, \\ \hbar &= \frac{h}{2\pi} = 1.05 \times 10^{-34} \text{ J} \cdot \text{s} = 6.58 \times 10^{-16} \text{ eV} \cdot \text{s}. \end{aligned}$$

(G) 精细结构常数:

$$\alpha = \frac{e^2}{2\epsilon_0 \hbar c} = 7.30 \times 10^{-3} = \frac{1}{137}.$$

(H) 电子的静止质量:

$$m_e = 9.11 \times 10^{-31} \text{ kg} = 0.511 \text{ MeV}/c^2.$$

(I) 质子的静止质量:

$$m_p = 1.67 \times 10^{-27} \text{ kg} = 9.38 \times 10^2 \text{ MeV}/c^2.$$

(J) 里德伯常数:

$$R = \frac{\alpha^2 m_e c}{2h} = 1.10 \times 10^7 \text{ m}^{-1}.$$

(K) 玻尔半径:

$$a_0 = \frac{\hbar}{\alpha m_e c} = 5.29 \times 10^{-11} \text{ m}.$$

(L) 玻尔兹曼常数:

$$k = 1.38 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}.$$

1 波函数

注意：除非特殊声明，默认 $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ ，即 x 是 \mathbb{R}^n 上的向量；并且 $dV = dx^1 \cdots dx^n$ 代表 \mathbb{R}^n 上的体积形式。

波函数是贯穿量子力学的核心概念，它反映了微观世界中，基本粒子出现在空间各处的概率分布。它是坐标 $x = (x^1, \dots, x^n)$ 、时间 t 的函数，并满足薛定谔方程——与坐标、时间有关的偏微分方程。之后，我们将求解一些满足特定薛定谔方程对应的波函数。

1.1 定义

记 $\Psi : \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{C}$ 是连续的平方可积函数，定义为

$$\int_{\mathbb{R}^n} |\Psi(x, t)|^2 dV < \infty,$$

其中 $x = (x^1, \dots, x^n)$ 代表坐标， t 代表时间， $dV = dx^1 \cdots dx^n$ 代表 \mathbb{R}^n 上的体积形式。

对于平方可积函数，其归一化操作为

$$\tilde{\Psi} = \frac{\Psi}{\left(\int_{\mathbb{R}^n} |\Psi|^2 dV\right)^{1/2}}.$$

因此

$$\int_{\mathbb{R}^n} |\tilde{\Psi}|^2 dV = 1. \quad (1.1)$$

将满足(1.1)的平方可积函数 $\tilde{\Psi}$ 称为**波函数**。

1.1.1 波函数是概率分布的体现

波函数反映了粒子在 \mathbb{R}^n 的**概率分布**：当粒子的波函数为 $\Psi(x, t)$ ， $x \in \mathbb{R}^n$ 时，粒子出现在区域 $\Omega = \prod_{i=1}^n (x^i, x^i + dx^i) \subseteq \mathbb{R}^n$ 的概率为

$$\Pr(\Omega) = |\Psi(x, t)|^2 dx^1 \cdots dx^n.$$

给定2个线性相关的平方可积函数 Ψ, Φ 满足

$$\Phi = c\Psi,$$

其中 $c \neq 0$ 是常数。分别对其进行归一化可得

$$\begin{aligned} \tilde{\Psi} &= \frac{\Psi}{\left(\int_{\mathbb{R}^n} |\Psi|^2 dV\right)^{1/2}}, \\ \tilde{\Phi} &= \frac{\Phi}{\left(\int_{\mathbb{R}^n} |\Phi|^2 dV\right)^{1/2}} = \frac{c}{|c|} \frac{\Psi}{\left(\int_{\mathbb{R}^n} |\Psi|^2 dV\right)^{1/2}} = \lambda \tilde{\Psi}, \end{aligned}$$

其中 $\lambda = c/|c|, |\lambda| = 1$ 。当粒子A,B的波函数分别为 $\tilde{\Psi}, \tilde{\Phi}$ 时, 粒子A,B出现在区域 $\Omega = \prod_{i=1}^n (x^i, x^i + dx^i) \subseteq \mathbb{R}^n$ 的概率分别为

$$\begin{aligned} \Pr_A(\Omega) &= \left| \tilde{\Psi}(x, t) \right|^2 dx^1 \cdots dx^n, \\ \Pr_B(\Omega) &= \left| \tilde{\Phi}(x, t) \right|^2 dx^1 \cdots dx^n = |\lambda|^2 \left| \tilde{\Psi}(x, t) \right|^2 dx^1 \cdots dx^n = \left| \tilde{\Psi}(x, t) \right|^2 dx^1 \cdots dx^n. \end{aligned} \quad (1.2)$$

因此 $\Pr_A(\Omega) = \Pr_B(\Omega)$, $\tilde{\Psi}, \tilde{\Phi}$ 描述的概率分布是一致的。

可以得出, 只要给定一个平方可积函数, 我们就可以对其进行归一化, 从而确定粒子的概率分布。并且, 根据(1.2), 即使平方可积函数与任意一个非零常数相乘, 它对应的概率分布不会发生改变。不正式地说, 决定粒子概率分布的是平方可积函数的“方向”, 而不是平方可积函数的“大小”。

1.1.2 平方可积函数的性质

平方可积函数的集合(记为 $L^2(\mathbb{R}^n)$)是复数域 \mathbb{C} 上的向量空间, 因为对于任意平方可积函数 Ψ, Φ 与常数 $c \in \mathbb{C}$, 都具有

$$\begin{aligned} \int_{\mathbb{R}^n} |\Psi + \Phi|^2 dV &\leq 4 \left(\int_{\mathbb{R}^n} |\Psi|^2 dV + \int_{\mathbb{R}^n} |\Phi|^2 dV \right) < \infty, \\ \int_{\mathbb{R}^n} |c\Psi|^2 dV &= |c|^2 \int_{\mathbb{R}^n} |\Psi|^2 dV < \infty. \end{aligned}$$

根据这一性质, 我们可以将平方可积函数进行任意的线性叠加, 即对于平方可积函数 Ψ_1, Ψ_2, \dots 与常数 $c_1, c_2, \dots \in \mathbb{C}$, 线性叠加

$$\Psi = \sum_{i=1}^{\infty} c_i \Psi_i$$

也是平方可积函数。

1.2 傅立叶变换*

¹ 对于定义域为 \mathbb{R}^n , 并且平方可积的函数 $u(x), x \in \mathbb{R}^n$, 定义它的**傅立叶变换** $\mathcal{F}u = \hat{u}$ 为

$$\hat{u}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot y} u(x) dV_x, \quad y \in \mathbb{R}^n,$$

与**逆傅立叶变换** $\mathcal{F}^{-1}u = \check{u}$ 为

$$\check{u}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix \cdot y} u(x) dV_x, \quad y \in \mathbb{R}^n,$$

¹符号*表示这一章节可以跳过, 建议跳转到第12页的章节1.2.1。

其中 $x \cdot y = \sum_{i=1}^n x^i y^i$ 。很显然，傅立叶变换与逆傅立叶变换都是 \mathbb{C} -线性的，即对于任意 $\alpha, \beta \in \mathbb{C}$ ，都具有

$$\begin{aligned}(\alpha u + \beta v)^\wedge &= \alpha \hat{u} + \beta \hat{v}, \\(\alpha u + \beta v)^\vee &= \alpha \check{u} + \beta \check{v},\end{aligned}$$

其中 $(\alpha u + \beta v)^\wedge, (\alpha u + \beta v)^\vee$ 是 $\alpha u + \beta v$ 的傅立叶变换与逆傅立叶变换。

Theorem 1.1 (Plancherel's Theorem). \hat{u}, \check{u} 均为平方可积函数，并且满足

$$\int_{\mathbb{R}^n} |\hat{u}(y)|^2 dV_y = \int_{\mathbb{R}^n} |\check{u}(y)|^2 dV_y = \int_{\mathbb{R}^n} |u(x)|^2 dV_x.$$

证明这个定理之前，我们首先需要证明一个积分式：

Lemma 1.2.

$$\int_{\mathbb{R}^n} e^{-ix \cdot y - t|x|^2} dV_x = \left(\frac{\pi}{t}\right)^{n/2} e^{-|y|^2/4t}, \quad (1.3)$$

其中 $t > 0$ 是常数， $|x| = \sqrt{(x^1)^2 + \cdots + (x^n)^2}$ ， $|y| = \sqrt{(y^1)^2 + \cdots + (y^n)^2}$ 是向量 x, y 的范数。

Proof. 因为

$$\int_{\mathbb{R}^n} e^{-ix \cdot y - t|x|^2} dV_x = \prod_{j=1}^n \int_{-\infty}^{\infty} e^{-iy^j x^j - t(x^j)^2} dx^j,$$

所以只需证明

$$\int_{-\infty}^{\infty} e^{-iax - tx^2} dx = \left(\frac{\pi}{t}\right)^{1/2} e^{-a^2/4t}$$

即可，其中 $a, t \in \mathbb{R}, t > 0$ 。令复数 $z = t^{1/2}x + \frac{a}{2t^{1/2}}i$ ，因此

$$\int_{-\infty}^{\infty} e^{-iax - tx^2} dx = \frac{e^{-a^2/4t}}{t^{1/2}} \int_{\Gamma} e^{-z^2} dz,$$

其中 Γ 是复平面上的轨迹 $\left\{ \operatorname{Im}(z) = \frac{a}{2t^{1/2}} \right\}$ 。因为 e^{-z^2} 是全纯函数，并且 $\lim_{|z| \rightarrow \infty} e^{-z^2} = 0$ ，所以

$$\int_{\Gamma} e^{-z^2} dz = \int_{-\infty}^{\infty} e^{-x^2} dx = \pi^{1/2},$$

因此

$$\int_{-\infty}^{\infty} e^{-iax - tx^2} dx = \frac{e^{-a^2/4t}}{t^{1/2}} \int_{\Gamma} e^{-z^2} dz = \left(\frac{\pi}{t}\right)^{1/2} e^{-a^2/4t}.$$

□

接下来，我们可以证明Theorem 1.1:

Proof. 对于平方可积函数 v, w ，它们满足

$$\int_{\mathbb{R}^n} v(x) \hat{w}(x) dV_x = \int_{\mathbb{R}^n} \hat{v}(y) w(y) dV_y = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-ix \cdot y} v(x) w(y) dV_x dV_y. \quad (1.4)$$

对于 $\epsilon > 0$ ，令 $v_\epsilon(x) = e^{-\epsilon|x|^2}$ 。根据(1.3)可得 $\hat{v}_\epsilon(y) = \frac{1}{(2\epsilon)^{n/2}} e^{-|y|^2/4\epsilon}$ 。代入(1.4)可得

$$\int_{\mathbb{R}^n} \hat{w}(y) e^{-\epsilon|y|^2} dV_y = \frac{1}{(2\epsilon)^{n/2}} \int_{\mathbb{R}^n} w(x) e^{-|x|^2/4\epsilon} dV_x. \quad (1.5)$$

令 u 是平方可积函数，并定义 $v(x) = \bar{u}(-x)$ 。令

$$w(x) = u * v = \int_{\mathbb{R}^n} u(x-y) v(y) dV_y,$$

根据Theorem 1.3(C)可得

$$\hat{w} = (2\pi)^{n/2} \hat{u} \hat{v}.$$

然而

$$\hat{v}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot y} \bar{u}(-x) dV_x = \bar{\hat{u}}(y),$$

所以

$$\hat{w} = (2\pi)^{n/2} |\hat{u}|^2. \quad (1.6)$$

因为 w 是连续的，根据(1.3)可得

$$\lim_{\epsilon \rightarrow 0} \frac{1}{(2\epsilon)^{n/2}} \int_{\mathbb{R}^n} w(x) e^{-|x|^2/4\epsilon} dV_x = w(0) \lim_{\epsilon \rightarrow 0} \frac{1}{(2\epsilon)^{n/2}} \int_{\mathbb{R}^n} e^{-|x|^2/4\epsilon} dV_x = (2\pi)^{n/2} w(0).$$

对(1.5)取极限 $\epsilon \rightarrow 0$ 可得

$$\int_{\mathbb{R}^n} \hat{w}(y) dV_y = (2\pi)^{n/2} w(0).$$

根据(1.6)可得

$$\int_{\mathbb{R}^n} |\hat{u}(y)|^2 dV_y = w(0) = \int_{\mathbb{R}^n} u(x) v(-x) dV_x = \int_{\mathbb{R}^n} |u(x)|^2 dV_x.$$

对于 \hat{u} 的情况，证明类似，不再赘述。 □

我们可以引入傅立叶变换的一些性质:

Theorem 1.3. 对于平方可积函数 u, v ，都具有

$$(A) \int_{\mathbb{R}^n} u(x) \bar{v}(x) dV_x = \int_{\mathbb{R}^n} \hat{u}(y) \bar{\hat{v}}(y) dV_y.$$

(B) 如果 u 是光滑函数, 并且 u 的支撑

$$\text{supp } u = \overline{\{x : u(x) \neq 0\}}$$

(对于集合 A , 记 \bar{A} 为 A 的闭包, 是包含 A 的最小闭集)包含于有限半径的球(以原点为球心), 即存在 $R > 0$ 使得

$$\text{supp } u \subseteq \{x : |x| \leq R\},$$

那么 $(D^\alpha u)^\wedge = (iy)^\alpha \hat{u}$, 其中 α 是非负整数, $D^\alpha = \frac{\partial^\alpha}{\partial(x^1)^{\alpha_1} \dots \partial(x^n)^{\alpha_n}}$, $\{\alpha_1, \dots, \alpha_n\}$ 是任意满足 $\alpha_1 + \dots + \alpha_n = \alpha$ 的非负整数的数组; $(D^\alpha u)^\wedge$ 是 $D^\alpha u$ 的傅立叶变换。

(C) 定义

$$w(x) = u * v = \int_{\mathbb{R}^n} u(x-y)v(y)dV_y$$

是 u, v 的卷积, 那么

$$\hat{w}(y) = (2\pi)^{n/2} \hat{u}(y) \hat{v}(y).$$

(D) $u = (\hat{u})^\vee$, 其中 $(\hat{u})^\vee$ 是 \hat{u} 的逆傅立叶变换。

Proof. (A) 根据Theorem 1.1可得, 对于平方可积函数 u, v 与常数 $\alpha \in \mathbb{C}$, 都具有

$$\int_{\mathbb{R}^n} |u(x) + \alpha v(x)|^2 dV_x = \int_{\mathbb{R}^n} |\hat{u}(y) + \alpha \hat{v}(y)|^2 dV_y.$$

展开可得

$$\int_{\mathbb{R}^n} |u|^2 + |\alpha v|^2 + \bar{u}(\alpha v) + u(\overline{\alpha v}) dV_x = \int_{\mathbb{R}^n} |\hat{u}|^2 + |\alpha \hat{v}|^2 + \bar{\hat{u}}(\alpha \hat{v}) + \hat{u}(\overline{\alpha \hat{v}}) dV_y,$$

利用Theorem 1.1, 可以抵消前2项:

$$\int_{\mathbb{R}^n} \bar{u}(\alpha v) + u(\overline{\alpha v}) dV_x = \int_{\mathbb{R}^n} \bar{\hat{u}}(\alpha \hat{v}) + \hat{u}(\overline{\alpha \hat{v}}) dV_y.$$

令 $\alpha = 1, i$, 可得

$$\begin{aligned} \int_{\mathbb{R}^n} \bar{u}v + u\bar{v} dV_x &= \int_{\mathbb{R}^n} \bar{\hat{u}}\hat{v} + \hat{u}\bar{\hat{v}} dV_y, \\ \int_{\mathbb{R}^n} \bar{u}v - u\bar{v} dV_x &= \int_{\mathbb{R}^n} \bar{\hat{u}}\hat{v} - \hat{u}\bar{\hat{v}} dV_y, \end{aligned}$$

因此

$$\int_{\mathbb{R}^n} u\bar{v} dV_x = \int_{\mathbb{R}^n} \hat{u}\bar{\hat{v}} dV_y.$$

(B) 利用分部积分法:

$$\begin{aligned}
(D^\alpha u)^\wedge(y) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot y} D^\alpha u(x) dx \\
&= \frac{(-1)^\alpha}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{\partial e^{-ix \cdot y}}{\partial x^\alpha} u(x) dx \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot y} (iy)^\alpha u(x) dx \\
&= (iy)^\alpha \hat{u}(y).
\end{aligned}$$

(C)

$$\begin{aligned}
(u * v)^\wedge(y) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot y} \int_{\mathbb{R}^n} u(z) v(x - z) dV_z dV_x \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-iz \cdot y} u(z) \left(\int_{\mathbb{R}^n} e^{-i(x-z) \cdot y} v(x - z) dV_x \right) dV_z \\
&= \int_{\mathbb{R}^n} e^{-iz \cdot y} u(z) dz \hat{v}(y) \\
&= (2\pi)^{n/2} \hat{u}(y) \hat{v}(y).
\end{aligned}$$

(D) 注意到

$$\int_{\mathbb{R}^n} \check{u} v dV_x = \int_{\mathbb{R}^n} u \check{v} dV_x = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{ix \cdot y} u(y) v(x) dV_x dV_y.$$

此外

$$\begin{aligned}
\overline{(\check{v})}^\wedge &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \overline{e^{-ix \cdot y} \check{v}(x)} dV_x \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix \cdot y} v(x) dV_x \\
&= \check{v},
\end{aligned}$$

因此, 根据(A)可得

$$\int_{\mathbb{R}^n} (\hat{u})^\sim v dV_x = \int_{\mathbb{R}^n} \hat{u} \check{v} dV_x = \int_{\mathbb{R}^n} \hat{u} \overline{(\check{v})}^\wedge dV_x = \int_{\mathbb{R}^n} v \bar{\bar{v}} dV_x = \int_{\mathbb{R}^n} u v dV_x,$$

根据 v 的任意性, 完成证明。

□

1.2.1 动量表象

利用上一节的知识，将 $\Psi(x, t)$ 进行傅立叶变换可得

$$\Phi(p, t) = \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} e^{-ip \cdot x / \hbar} \Psi(x, t) dV_x. \quad (1.7)$$

根据Theorem 1.3(D)可得

$$\Psi(x, t) = \frac{1}{(2\pi\hbar)^{n/2}} \int_{\mathbb{R}^n} e^{ip \cdot x / \hbar} \Phi(p, t) dV_p.$$

根据Theorem 1.1可得

$$\int_{\mathbb{R}^n} |\Phi(p, t)|^2 dV_p = \int_{\mathbb{R}^n} |\Psi(x, t)|^2 dV_x,$$

因此，如果 $\int_{\mathbb{R}^n} |\Psi(x, t)|^2 dV_x = 1$ ，那么

$$\int_{\mathbb{R}^3} |\Phi(p, t)|^2 dV_p = 1,$$

表明 $\Phi(p, t)$ 同样也是归一化的。此时我们称 $\Phi(p, t)$ 为**动量表象波函数**，它反映了粒子动量在动量空间各处的概率分布，粒子动量位于区域 $\prod_{i=1}^n (p^i, p^i + dp^i) \subseteq \mathbb{R}^n$ 的概率为

$$|\Phi(p, t)|^2 dp^1 \cdots dp^n.$$

1.3 薛定谔方程

薛定谔方程是关于平方可积函数 $\Psi(x, t)$ 的偏微分方程：

$$-\frac{\hbar^2}{2m} \Delta \Psi(x, t) + V(x) \Psi(x, t) = i\hbar \frac{\partial \Psi}{\partial t}(x, t). \quad (1.8)$$

其中 $\Delta \Psi(x, t) = \sum_{i=1}^n \frac{\partial^2 \Psi}{\partial (x^i)^2}$ 。令 $\Psi(x, t) = \Psi(x, 0) f(t)$ ，可得

$$-\frac{\hbar^2}{2m} \frac{\Delta \Psi(x, 0)}{\Psi(x, 0)} + V(x) = i\hbar \frac{f'(t)}{f(t)}.$$

而左边是 x 的函数，右边是 t 的函数，因此左边与右边取值恒为常数，记其为 E 。因此

$$\begin{aligned} -\frac{\hbar^2}{2m} \Delta \Psi(x, 0) + V(x) \Psi(x, 0) &= E \Psi(x, 0), \\ f'(t) &= -\frac{iE}{\hbar} f(t). \end{aligned}$$

可得

$$f = e^{-iEt/\hbar},$$

因此

$$\Psi(x, t) = \Psi(x, 0)e^{-iEt/\hbar}.$$

我们只需求解

$$-\frac{\hbar^2}{2m}\Delta\Psi(x, 0) + V(x)\Psi(x, 0) = E\Psi(x, 0) \quad (1.9)$$

对应的 Ψ, E 即可。将该方程称为**定态薛定谔方程**， $\Psi(x, 0)$ 称为**定态波函数**，它是初始时刻 $t = 0$ 的波函数。为了表示方便，通常情况下，我们可以将定态波函数 $\Psi(x, 0)$ 简记为 $\Psi(x)$ 。

定义平方可积函数的**内积**

$$\langle\Phi(x, t), \Psi(x, t)\rangle = \int_{\mathbb{R}^n} \overline{\Phi}(x, t)\Psi(x, t)dV,$$

很显然，它们满足以下性质：

$$\begin{aligned} \langle\Phi, \Psi\rangle &= \overline{\langle\Psi, \Phi\rangle}, \\ \langle\Phi, \alpha\Psi + \beta\Theta\rangle &= \alpha\langle\Phi, \Psi\rangle + \beta\langle\Phi, \Theta\rangle, \\ \langle\alpha\Phi + \beta\Theta, \Psi\rangle &= \overline{\alpha}\langle\Phi, \Psi\rangle + \overline{\beta}\langle\Theta, \Psi\rangle, \\ \langle\Psi, \Psi\rangle &\geq 0. \end{aligned} \quad (1.10)$$

其中第4个表达式取等号当且仅当 $\Psi = 0$ 。根据第4个表达式，可以定义函数的**范数**

$$\|\Psi\| = \langle\Psi, \Psi\rangle^{1/2} = \int_{\mathbb{R}^n} |\Psi|^2 dV.$$

因此，归一化的波函数 Ψ 满足 $\|\Psi\| = 1$ 。范数满足以下性质：

$$\begin{aligned} \|\alpha\Psi\| &= |\alpha|\|\Psi\|, \\ |\langle\Phi, \Psi\rangle| &\leq \|\Phi\|\|\Psi\|, \\ \|\Phi + \Psi\| &\leq \|\Phi\| + \|\Psi\|. \end{aligned} \quad (1.11)$$

证明如下：

Proof. 第1个表达式是很显然的，根据(1.10)可得

$$\|\alpha\Psi\| = \sqrt{\langle\alpha\Psi, \alpha\Psi\rangle} = \sqrt{\alpha\overline{\alpha}\langle\Psi, \Psi\rangle} = |\alpha|\|\Psi\|.$$

对于第2个表达式，对于任意 $\lambda \in \mathbb{C}$ ，根据(1.10)可得

$$0 \leq \langle\Phi - \lambda\Psi, \Phi - \lambda\Psi\rangle = \langle\Phi, \Phi\rangle - \lambda\langle\Phi, \Psi\rangle - \overline{\lambda}\langle\Psi, \Phi\rangle + |\lambda|^2\langle\Psi, \Psi\rangle,$$

令 $\langle \Phi, \Psi \rangle = re^{i\theta}$, 选取 $\lambda = te^{-i\theta}$, 上面的不等式化为

$$0 \leq \langle \Phi, \Phi \rangle - 2rt + t^2 \langle \Psi, \Psi \rangle,$$

该二次不等式 $at^2 + bt + c \geq 0$ 的判别式满足 $\Delta = b^2 - 4ac \leq 0$, 即

$$4r^2 - 4 \langle \Phi, \Phi \rangle \langle \Psi, \Psi \rangle \leq 0,$$

因此

$$|\langle \Phi, \Psi \rangle| \leq \|\Phi\| \|\Psi\|.$$

对于第3个表达式, 根据(1.10)可得

$$\begin{aligned} \|\Phi + \Psi\|^2 &= \langle \Phi + \Psi, \Phi + \Psi \rangle = \langle \Phi, \Phi \rangle + \langle \Phi, \Psi \rangle + \langle \Psi, \Phi \rangle + \langle \Psi, \Psi \rangle \\ &= \|\Phi\|^2 + 2\operatorname{Re}\langle \Phi, \Psi \rangle + \|\Psi\|^2 \\ &\leq \|\Phi\|^2 + 2|\langle \Phi, \Psi \rangle| + \|\Psi\|^2 \\ &\leq \|\Phi\|^2 + 2\|\Phi\| \|\Psi\| + \|\Psi\|^2 \\ &= (\|\Phi\| + \|\Psi\|)^2, \end{aligned}$$

因此

$$\|\Phi + \Psi\| \leq \|\Phi\| + \|\Psi\|.$$

□

求解定态薛定谔方程(1.9), 我们可能会得到一组解 $\{\Psi_i(x), E_i\}, i = 1, 2, 3, \dots$ 满足

$$-\frac{\hbar^2}{2m} \Delta \Psi_i(x) + V(x) \Psi_i(x) = E_i \Psi_i(x), \quad (1.12)$$

我们将 $\{\Psi_i(x), E_i\}, i = 1, 2, 3, \dots$ 称为**能级**。它们是标准正交的:

$$\langle \Psi_i(x), \Psi_j(x) \rangle = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (1.13)$$

$\{\Psi_i(x)\}$ 是平方可积函数集合 $L^2(\mathbb{R}^n)$ 的一组标准正交基。因此, 对于初始时刻的任意波函数 $\Psi(x)$, 它都可以表示为

$$\Psi(x) = \sum_{i=1}^{\infty} c_i \Psi_i(x), \quad (1.14)$$

其中 $c_i \in \mathbb{C}$ 是常数。根据(1.13)可得

$$c_i = \langle \Psi_i(x), \Psi(x) \rangle = \int_{\mathbb{R}^n} \overline{\Psi_i(x)} \Psi(x) dV. \quad (1.15)$$

Theorem 1.4. 任意时刻 t 的波函数为

$$\Psi(x, t) = \sum_{i=1}^{\infty} c_i \Psi_i(x) e^{-iE_i t/\hbar}. \quad (1.16)$$

Proof. 对(1.16)微分可得

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t}(x, t) &= \sum_{i=1}^{\infty} c_i E_i \Psi_i(x) e^{-iE_i t/\hbar}, \\ -\frac{\hbar^2}{2m} \Delta \Psi(x, t) + V(x) \Psi(x, t) &= \sum_{i=1}^{\infty} c_i \left(-\frac{\hbar^2}{2m} \Delta \Psi_i(x) + V(x) \Psi_i(x) \right) e^{-iE_i t/\hbar}. \end{aligned}$$

根据(1.12)可得

$$-\frac{\hbar^2}{2m} \Delta \Psi(x, t) + V(x) \Psi(x, t) = i\hbar \frac{\partial \Psi}{\partial t}(x, t).$$

因此, (1.16)定义的 $\Psi(x, t)$ 满足薛定谔方程。 \square

1.3.1 一维定态薛定谔方程的性质

有时候为了简便, 我们经常研究一维定态薛定谔方程(粒子只在一维空间 \mathbb{R} 上运动)。此时, 定态薛定谔方程为

$$-\frac{\hbar^2}{2m} \Psi''(x) + V(x) \Psi(x) = E \Psi(x), \quad x \in \mathbb{R}.$$

我们将它变换为

$$\Psi''(x) = \frac{2m}{\hbar^2} (V(x) - E) \Psi(x).$$

因此, 对于波函数的连续性, 具有以下结论:

- (A) 当 $V(x)$ 处处连续, 那么 $\Psi(x), \Psi'(x), \Psi''(x)$ 均处处连续。
- (B) 当 $V(x)$ 在 $x = x_0$ 处不连续, 但是 $V(x)$ 在 $(x_0 - \epsilon, x_0 + \epsilon)$ 是有界的, 那么 $\Psi(x), \Psi'(x)$ 在 $x = x_0$ 处连续, 即

$$\begin{aligned} \lim_{x \nearrow x_0} \Psi(x) &= \lim_{x \searrow x_0} \Psi(x), \\ \lim_{x \nearrow x_0} \Psi'(x) &= \lim_{x \searrow x_0} \Psi'(x). \end{aligned} \quad (1.17)$$

$\Psi''(x)$ 则不一定。

(C) 当 $V(x)$ 在 $x = x_0$ 处不连续, 并且 $V(x)$ 在 $(x_0 - \epsilon, x_0 + \epsilon)$ 是无界的, 那么 $\Psi(x)$ 在 $x = x_0$ 处连续, 即

$$\lim_{x \nearrow x_0} \Psi(x) = \lim_{x \searrow x_0} \Psi(x). \quad (1.18)$$

$\Psi'(x), \Psi''(x)$ 则不一定。

为了保证 $\Psi(x)$ 是平方可积函数, 并且具有非平凡解 $\Psi(x) \neq 0$, 必须满足

$$E \geq \inf V.$$

因为此时 $\Psi(x)$ 可以被归一化, 对于归一化的 $\Psi(x)$, 具有

$$\begin{aligned} E &= E \int_{-\infty}^{\infty} |\Psi(x)|^2 dx \\ &= \int_{-\infty}^{\infty} \bar{\Psi}(x) (E\Psi(x)) dx \\ &= \int_{-\infty}^{\infty} \bar{\Psi}(x) \left(-\frac{\hbar^2}{2m} \Psi''(x) + V(x)\Psi(x) \right) dx \\ &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} \bar{\Psi}(x) \Psi''(x) dx + \int_{-\infty}^{\infty} V(x) |\Psi(x)|^2 dx \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} |\Psi'(x)|^2 dx + \int_{-\infty}^{\infty} V(x) |\Psi(x)|^2 dx \\ &\geq \int_{-\infty}^{\infty} V(x) |\Psi(x)|^2 dx \\ &\geq \inf V \int_{-\infty}^{\infty} |\Psi(x)|^2 dx \\ &= \inf V. \end{aligned}$$

其中4-5行利用了分部积分法:

$$\begin{aligned} \int_{-\infty}^{\infty} \bar{\Psi}(x) \Psi''(x) dx &= [\bar{\Psi}(x) \Psi'(x)]_{x \rightarrow -\infty}^{x \rightarrow \infty} - \int_{-\infty}^{\infty} \bar{\Psi}'(x) \Psi'(x) dx \\ &= - \int_{-\infty}^{\infty} |\Psi'(x)|^2 dx. \end{aligned} \quad (1.19)$$

1.3.2 一维定态薛定谔方程的解

(I) 一维无限深方势阱:

此时势能函数为

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ \infty, & \text{otherwise,} \end{cases}$$

其定态薛定谔方程为

$$-\frac{\hbar^2}{2m}\Psi''(x) + V(x)\Psi(x) = E\Psi(x),$$

所以

$$\begin{cases} \Psi(x) = 0, & x < 0, \\ -\frac{\hbar^2}{2m}\Psi''(x) = E\Psi(x), & 0 < x < a, \\ \Psi(x) = 0, & x > a. \end{cases}$$

记 $k = \frac{\sqrt{2mE}}{\hbar}$, 因此

$$\Psi(x) = \begin{cases} 0, & x < 0, \\ A \cos kx + B \sin kx, & 0 < x < a, \\ 0, & x > a. \end{cases}$$

根据(1.18)可得

$$\begin{aligned} \lim_{x \nearrow 0} \Psi(x) &= \lim_{x \searrow 0} \Psi(x), \\ \lim_{x \searrow a} \Psi(x) &= \lim_{x \nearrow a} \Psi(x), \end{aligned}$$

因此

$$\begin{aligned} A &= 0, \\ \sin ka &= 0. \end{aligned}$$

可得

$$k = \frac{n\pi}{a}, \quad n = 1, 2, \dots,$$

因此

$$\begin{aligned} \frac{\sqrt{2mE}}{\hbar} &= \frac{n\pi}{a}, \\ E &= \frac{n^2\pi^2\hbar^2}{2ma^2}. \end{aligned}$$

波函数为

$$\Psi(x) = \begin{cases} 0, & x < 0, \\ B \sin \frac{n\pi x}{a}, & 0 < x < a, \\ 0, & x > a. \end{cases}$$

而

$$\langle \Psi, \Psi \rangle = \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = B^2 \int_0^a \sin^2 \frac{n\pi x}{a} dx = \frac{aB^2}{n\pi} \int_0^{n\pi} \sin^2 u du = \frac{aB^2}{2} = 1,$$

因此

$$B = \sqrt{\frac{2}{a}}.$$

波函数为

$$\Psi(x) = \begin{cases} 0, & x < 0, \\ \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}, & 0 < x < a, \\ 0, & x > a. \end{cases} \quad (1.20)$$

(II) 一维有限深方势阱:

此时势能函数为

$$V(x) = \begin{cases} 0, & -a < x < a, \\ V_0, & \text{otherwise,} \end{cases}$$

其定态薛定谔方程为

$$-\frac{\hbar^2}{2m} \Psi''(x) + V(x)\Psi(x) = E\Psi(x),$$

所以

$$\begin{cases} -\frac{\hbar^2}{2m} \Psi''(x) + V_0\Psi(x) = E\Psi(x), & x < -a, \\ -\frac{\hbar^2}{2m} \Psi''(x) = E\Psi(x), & -a < x < a, \\ -\frac{\hbar^2}{2m} \Psi''(x) + V_0\Psi(x) = E\Psi(x), & x > a. \end{cases}$$

记 $k = \frac{\sqrt{2m(V_0-E)}}{\hbar}$, $l = \frac{\sqrt{2mE}}{\hbar}$, 因此

$$\begin{cases} \Psi''(x) - k^2\Psi(x) = 0, & x < -a, \\ \Psi''(x) + l^2\Psi(x) = 0, & -a < x < a, \\ \Psi''(x) - k^2\Psi(x) = 0, & x > a. \end{cases}$$

根据 $\lim_{|x| \rightarrow \infty} \Psi(x) = 0$ (否则无法满足平方可积的条件) 可得

$$\Psi(x) = \begin{cases} Ae^{kx}, & x < -a, \\ B \cos lx + C \sin lx, & -a < x < a, \\ De^{-kx}, & x > a. \end{cases}$$

根据(1.17)可得

$$\begin{aligned}\lim_{x \nearrow -a} \Psi(x) &= \lim_{x \searrow -a} \Psi(x), \\ \lim_{x \nearrow -a} \Psi'(x) &= \lim_{x \searrow -a} \Psi'(x), \\ \lim_{x \nearrow a} \Psi(x) &= \lim_{x \searrow a} \Psi(x), \\ \lim_{x \nearrow a} \Psi'(x) &= \lim_{x \searrow a} \Psi'(x),\end{aligned}$$

因此

$$\begin{aligned}Ae^{-ka} &= B \cos la - C \sin la, \\ De^{-ka} &= B \cos la + C \sin la, \\ Ake^{-ka} &= Bl \sin la + Cl \cos la, \\ -Dke^{-ka} &= -Bl \sin la + Cl \cos la.\end{aligned}\tag{1.21}$$

因为 $k^2 + l^2 = \frac{2mV_0}{\hbar^2} > 0$, 所以不可能出现 $k = l = 0$ 的情况, 只能是以下3种情况:

(A) $k = 0, l \neq 0$.

(B) $k \neq 0, l = 0$.

(C) $k, l \neq 0$.

我们分别对这3种情况进行分析:

(A) 此时

$$\begin{aligned}E &= V_0, \\ \Psi(x) &= \begin{cases} A, & x < -a, \\ B \cos lx + C \sin lx, & -a < x < a, \\ D, & x > a, \end{cases}\end{aligned}$$

根据 $\lim_{|x| \rightarrow \infty} \Psi(x) = 0$ 可得

$$A = D = 0.$$

根据(1.21)可得

$$\begin{aligned}0 &= B \cos la - C \sin la, \\ 0 &= B \cos la + C \sin la, \\ 0 &= Bl \sin la + Cl \cos la, \\ 0 &= -Bl \sin la + Cl \cos la,\end{aligned}$$

即

$$\begin{aligned}B \cos la &= 0, \\B \sin la &= 0, \\C \cos la &= 0, \\C \sin la &= 0.\end{aligned}$$

而 $\cos^2 la + \sin^2 la = 1$, 所以 $\cos la, \sin la$ 不能同时为0, 因此

$$B = C = 0,$$

这表明 $\Psi(x) = 0$, 不是我们关心的解。

(B) 此时

$$\begin{aligned}E &= 0, \\ \Psi(x) &= \begin{cases} Ae^{kx}, & x < -a, \\ B, & -a < x < a, \\ De^{-kx}, & x > a, \end{cases}\end{aligned}$$

根据(1.21)可得

$$\begin{aligned}Ae^{-ka} &= B, \\De^{-ka} &= B, \\Ake^{-ka} &= 0, \\-Dke^{-ka} &= 0.\end{aligned}$$

因为 $k \neq 0, e^{-ka} \neq 0$, 所以

$$A = D = 0,$$

可得

$$B = 0,$$

这表明 $\Psi(x) = 0$, 不是我们关心的解。

(C) 将(1.21)进行恒等变换可得

$$\begin{aligned}Ae^{-ka} &= B \cos la - C \sin la, \\De^{-ka} &= B \cos la + C \sin la, \\Ae^{-ka} &= \frac{l}{k} (B \sin la + C \cos la), \\De^{-ka} &= \frac{l}{k} (B \sin la - C \cos la),\end{aligned}$$

因此

$$\begin{aligned} B \cos la &= \frac{l}{k} B \sin la, \\ C \sin la &= -\frac{l}{k} C \cos la. \end{aligned} \tag{1.22}$$

如果 $B = C = 0$, 那么 $\Psi(x) = 0$, 不是我们关心的解。如果 $B, C \neq 0$, 那么

$$\begin{aligned} \cos la &= \frac{l}{k} \sin la, \\ \sin la &= -\frac{l}{k} \cos la. \end{aligned}$$

因此

$$1 = \cos^2 la + \sin^2 la = \left(\frac{l}{k} \sin la \right) \cos la + \left(-\frac{l}{k} \cos la \right) \sin la = 0,$$

这是不可能的。因此, 只有可能是以下2种情况:

(a) $B \neq 0, C = 0$.

(b) $B = 0, C \neq 0$.

我们分别对这2种情况进行分析:

(a) 此时

$$\Psi(x) = \begin{cases} Ae^{kx}, & x < -a, \\ B \cos lx, & -a < x < a, \\ De^{-kx}, & x > a, \end{cases}$$

根据(1.22)可得

$$\tan la = \frac{k}{l},$$

这对应

$$\tan \left(\frac{\sqrt{2mE}}{\hbar} a \right) = \sqrt{\frac{V_0}{E}} - 1.$$

根据(1.21)可得

$$A = D = Be^{ka} \cos la,$$

因此

$$\Psi(x) = \begin{cases} Be^{k(x+a)} \cos la, & x < -a, \\ B \cos lx, & -a < x < a, \\ Be^{-k(x-a)} \cos la, & x > a. \end{cases}$$

因此

$$\begin{aligned}
\langle \Psi, \Psi \rangle &= \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = B^2 \left(a + \frac{\cos^2 la}{k} + \frac{\sin 2la}{2l} \right) \\
&= B^2 \left(a + \frac{1}{k} \left(\cos^2 la + \frac{k}{2l} \sin 2la \right) \right) \\
&= B^2 \left(a + \frac{1}{k} \left(\cos^2 la + \frac{1}{2} \tan la \sin 2la \right) \right) \\
&= B^2 \left(a + \frac{1}{k} \right) \\
&= 1,
\end{aligned}$$

可得

$$B = \left(a + \frac{1}{k} \right)^{-1/2} = \left(a + \frac{\hbar}{\sqrt{2m(V_0 - E)}} \right)^{-1/2}.$$

(b) 此时

$$\Psi(x) = \begin{cases} Ae^{kx}, & x < -a, \\ C \sin lx, & -a < x < a, \\ De^{-kx}, & x > a, \end{cases}$$

根据(1.21)可得

$$\tan la = -\frac{l}{k},$$

这对应

$$\tan \left(\frac{\sqrt{2mE}}{\hbar} a \right) = -\sqrt{\frac{E}{V_0 - E}}.$$

根据(1.21)可得

$$A = -D = -Ce^{ka} \sin la,$$

因此

$$\Psi(x) = \begin{cases} -Ce^{k(x+a)} \sin la, & x < -a, \\ C \sin lx, & -a < x < a, \\ Ce^{-k(x-a)} \sin la, & x > a. \end{cases}$$

因此

$$\begin{aligned}
\langle \Psi, \Psi \rangle &= \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = C^2 \left(a + \frac{\sin^2 la}{k} - \frac{\sin 2la}{2l} \right) \\
&= C^2 \left(a + \frac{1}{k} \left(\sin^2 la - \frac{k}{2l} \sin 2la \right) \right) \\
&= C^2 \left(a + \frac{1}{k} \left(\sin^2 la + \frac{1}{2} \cot la \sin 2la \right) \right) \\
&= C^2 \left(a + \frac{1}{k} \right) \\
&= 1,
\end{aligned}$$

可得

$$C = \left(a + \frac{1}{k} \right)^{-1/2} = \left(a + \frac{\hbar}{\sqrt{2m(V_0 - E)}} \right)^{-1/2}.$$

(III) **概率流:** 定义概率流为

$$J = \frac{i\hbar}{2m} (\Psi \nabla \bar{\Psi} - \bar{\Psi} \nabla \Psi). \quad (1.23)$$

它的散度为

$$\operatorname{div} J = \frac{i\hbar}{2m} (\nabla \Psi \cdot \nabla \bar{\Psi} + \Psi \Delta \bar{\Psi} - \nabla \bar{\Psi} \cdot \nabla \Psi - \bar{\Psi} \Delta \Psi) = \frac{i\hbar}{2m} (\Psi \Delta \bar{\Psi} - \bar{\Psi} \Delta \Psi),$$

其中对于光滑函数 f, g , $\nabla f \cdot \nabla g = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^i}$ 。根据(1.8)可得

$$\begin{aligned}
\Psi \Delta \bar{\Psi} &= \frac{2mi}{\hbar} \Psi \frac{\partial \bar{\Psi}}{\partial t} + \frac{2mV}{\hbar^2} |\Psi|^2, \\
\bar{\Psi} \Delta \Psi &= -\frac{2mi}{\hbar} \bar{\Psi} \frac{\partial \Psi}{\partial t} + \frac{2mV}{\hbar^2} |\Psi|^2,
\end{aligned}$$

因此

$$\operatorname{div} J = -\Psi \frac{\partial \bar{\Psi}}{\partial t} - \bar{\Psi} \frac{\partial \Psi}{\partial t} = -\frac{\partial |\Psi|^2}{\partial t}.$$

根据Stokes定理可得, 对于任意有界闭集 $\Omega \subseteq \mathbb{R}^3$, 都具有

$$\int_{\Omega} -\frac{\partial |\Psi|^2}{\partial t} dV = \int_{\partial \Omega} J \cdot dA,$$

其中 $\partial \Omega$ 是区域 Ω 的边界。而左边正是粒子处于区域 Ω 的概率对时间的微分, 如Figure 1.1所示。

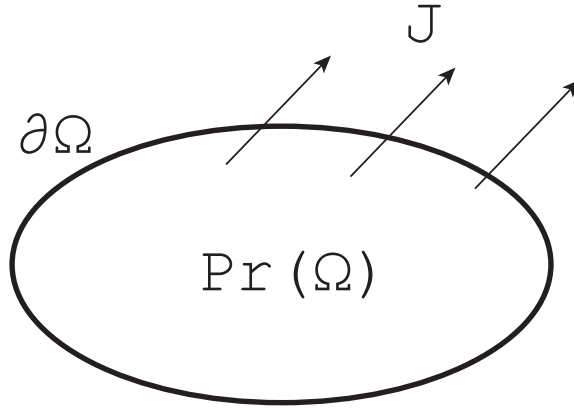


Figure 1.1: 概率流

(IV) 散射态:

对于一维势垒

$$V(x) = \begin{cases} 0, & x < 0, \\ V_0, & x \geq 0, \end{cases}$$

其中 $V_0 > 0$ 。当电子从左到右入射到该势垒上，我们需要计算它的反射率与透射率。具体步骤如下：

(A) $E > V_0$.

此时定态薛定谔方程为

$$\begin{cases} -\frac{\hbar^2}{2m}\Psi''(x) = E\Psi(x), & x < 0, \\ -\frac{\hbar^2}{2m}\Psi''(x) = (E - V_0)\Psi(x), & x \geq 0. \end{cases}$$

令 $k = \sqrt{\frac{2mE}{\hbar^2}}$, $l = \sqrt{\frac{2m(E-V_0)}{\hbar^2}}$, 因此

$$\Psi(x) = \begin{cases} e^{ikx} + re^{-ikx}, & x < 0, \\ te^{ilx}, & x \geq 0. \end{cases}$$

根据(1.17)可得

$$\begin{aligned} 1 + r &= t, \\ k(1 - r) &= lt. \end{aligned}$$

因此

$$\begin{aligned} r &= \frac{k - l}{k + l}, \\ t &= \frac{2k}{k + l}. \end{aligned}$$

我们将入射波、反射波、透射波分别记为

$$\begin{aligned}\Psi_i &= e^{ikx}, \\ \Psi_r &= re^{-ikx}, \\ \Psi_t &= te^{ilx}.\end{aligned}$$

根据(1.23)，它们的概率流分别为

$$\begin{aligned}J_i &= \frac{i\hbar}{2m} (\Psi_i \overline{\Psi'_i} - \overline{\Psi_i} \Psi'_i) = \frac{\hbar k}{m}, \\ J_r &= \frac{i\hbar}{2m} (\Psi_r \overline{\Psi'_r} - \overline{\Psi_r} \Psi'_r) = \frac{\hbar k}{m} r^2, \\ J_t &= \frac{i\hbar}{2m} (\Psi_t \overline{\Psi'_t} - \overline{\Psi_t} \Psi'_t) = \frac{\hbar l}{m} t^2.\end{aligned}$$

因此，反射率与透射率分别为

$$\begin{aligned}R &= \frac{J_r}{J_i} = r^2 = \frac{(\sqrt{E} - \sqrt{E - V_0})^2}{(\sqrt{E} + \sqrt{E - V_0})^2}, \\ T &= \frac{J_t}{J_i} = \frac{l}{k} t^2 = \frac{4\sqrt{E}\sqrt{E - V_0}}{(\sqrt{E} + \sqrt{E - V_0})^2}.\end{aligned}$$

因此

$$R + T = 1.$$

(B) $E < V_0$.

此时定态薛定谔方程为

$$\begin{cases} -\frac{\hbar^2}{2m} \Psi''(x) = E\Psi(x), & x < 0, \\ -\frac{\hbar^2}{2m} \Psi''(x) + (V_0 - E)\Psi(x) = 0, & x \geq 0. \end{cases}$$

令 $k = \sqrt{\frac{2mE}{\hbar^2}}$, $l = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}}$, 因此

$$\Psi(x) = \begin{cases} e^{ikx} + re^{-ikx}, & x < 0, \\ te^{-lx}, & x \geq 0. \end{cases}$$

根据(1.17)与(1.18)可得

$$\begin{aligned}1 + r &= t, \\ k(1 - r) &= illt.\end{aligned}$$

因此

$$r = \frac{k - il}{k + il},$$

$$t = \frac{2k}{k + il}.$$

我们将入射波、反射波、透射波分别记为

$$\Psi_i = e^{ikx},$$

$$\Psi_r = re^{-ikx},$$

$$\Psi_t = te^{-lx}.$$

根据(1.23)，它们的概率流分别为

$$J_i = \frac{i\hbar}{2m} (\Psi_i \overline{\Psi_i'} - \overline{\Psi_i} \Psi_i') = \frac{\hbar k}{m},$$

$$J_r = \frac{i\hbar}{2m} (\Psi_r \overline{\Psi_r'} - \overline{\Psi_r} \Psi_r') = \frac{\hbar k}{m} |r|^2,$$

$$J_t = \frac{i\hbar}{2m} (\Psi_t \overline{\Psi_t'} - \overline{\Psi_t} \Psi_t') = 0.$$

因此，反射率与透射率分别为

$$R = \frac{J_r}{J_i} = |r|^2 = 1,$$

$$T = \frac{J_t}{J_i} = 0.$$

因此

$$R + T = 1.$$

(V) 谐振子:

对于抛物线型势能函数

$$V(x) = \frac{1}{2}m\omega^2 x^2,$$

其中 ω 是常数，其物理含义是振动的圆频率(频率 f 与 2π 的乘积)。我们将不对该薛定谔方程进行具体的求解，因为过程特别复杂，因此，我们只给出结论：定态薛定谔方程

$$-\frac{\hbar^2}{2m}\Psi_n''(x) + V(x)\Psi_n(x) = E_n\Psi_n(x), \quad n = 0, 1, 2, \dots$$

的解为

$$\begin{aligned}\Psi_n(x) &= \left(\frac{\alpha}{\pi^{1/2} 2^n n!} \right)^{1/2} e^{-\alpha^2 x^2 / 2} H_n(\alpha x), \\ E_n &= \left(n + \frac{1}{2} \right) \hbar \omega,\end{aligned}\tag{1.24}$$

其中 $\alpha = \sqrt{\frac{m\omega}{\hbar}}$, $H_n(x)$ 是次数为 n 的多项式, 具体形式无需得知。

1.4 晶体的热容*

² 对于谐振子的能量, 我们虽然不关心它的推导过程, 但是它具有广泛的应用。玻尔兹曼分布表明, 在环境温度为 T 的情况下, 对于晶体中振动频率为 ω 的粒子, 其能量为 E_n 的数量正比于

$$e^{-E_n/kT} = e^{-\beta E_n},$$

其中 $E_n = (n + 1/2) \hbar \omega$ 是谐振子的能量, $k = 1.38 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$ 是玻尔兹曼常数, $\beta = 1/kT$ 是为了简便引入的辅助变量。因此, 对于所有振动频率为 ω 的粒子, 其平均能量为

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} E_n e^{-\beta E_n}}{\sum_{n=0}^{\infty} e^{-\beta E_n}}.$$

将分母视为 β 的函数:

$$\begin{aligned}Z(\beta) &= \sum_{n=0}^{\infty} e^{-\beta E_n} = \sum_{n=0}^{\infty} e^{-\beta \hbar \omega (n+1/2)} = e^{-\beta \hbar \omega / 2} \sum_{n=0}^{\infty} e^{-\beta \hbar \omega n} \\ &= \frac{e^{-\beta \hbar \omega / 2}}{1 - e^{-\beta \hbar \omega}}.\end{aligned}$$

因此

$$-\frac{\partial Z}{\partial \beta} = \sum_{n=0}^{\infty} E_n e^{-\beta E_n} = \frac{\hbar \omega e^{-\beta \hbar \omega / 2} (1 + e^{-\beta \hbar \omega})}{2 (1 - e^{-\beta \hbar \omega})^2}.$$

这表明

$$\begin{aligned}\langle E \rangle &= \frac{-\partial Z / \partial \beta}{Z} = \frac{\hbar \omega (1 + e^{-\beta \hbar \omega})}{2 (1 - e^{-\beta \hbar \omega})} = \frac{\hbar \omega}{2} + \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \\ &= \frac{\hbar \omega}{2} + \frac{\hbar \omega}{e^{\hbar \omega / kT} - 1}.\end{aligned}\tag{1.25}$$

爱因斯坦认为, 晶体中的所有电子都以相同并且固定的频率 ω_0 振动。这样一来, 假设晶体中的电子个数为 N , 那么晶体的内能, 即所有电子的总能量为

$$E = N \langle E \rangle = N \left(\frac{\hbar \omega_0}{2} + \frac{\hbar \omega_0}{e^{\hbar \omega_0 / kT} - 1} \right) = \frac{N \hbar \omega_0}{2} + \frac{N \hbar \omega_0}{e^{\hbar \omega_0 / kT} - 1}.$$

²符号*表示这一章节可以跳过, 建议跳转到第36页的章节2.1。

晶体的热容为

$$C_V = \frac{\partial E}{\partial T} = \frac{N\hbar^2\omega_0^2}{kT^2} \frac{e^{\hbar\omega_0/kT}}{(e^{\hbar\omega_0/kT} - 1)^2} = CF\left(\frac{T_0}{T}\right),$$

其中 C 是与温度无关的常数； $T_0 = \hbar\omega_0/k$ 是**爱因斯坦温度**，只与材料的性质有关(例如，对于金刚石， $T_0 \approx 1,320$ K)；函数 $F(x) = \frac{x^2 e^x}{(e^x - 1)^2}$ 。

在高温 $T \gg T_0$ 的情况下， $T_0/T \ll 1$ 。此时利用 $x \ll 1$ 的近似 $e^x \approx 1 + x$ 可得，

$$F\left(\frac{T_0}{T}\right) \approx 1,$$

因此，在高温条件下，热容与温度几乎无关。这与热力学的结论类似：例如，1 mol 的氢气在温度较高的时候，热容为 $C_V = \frac{3}{2}R$ ，与温度无关，其中 $R = 8.31 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ 是气体常数。

但是，在低温 $T \ll T_0$ 的情况下，情况就有所不同了。此时 $T_0/T \gg 1$ ，

$$F(x) = \frac{x^2 e^x}{(e^x - 1)^2} \approx x^2 e^{-x} = \left(\frac{T_0}{T}\right)^2 e^{-T_0/T},$$

热容随着温度的衰减速率是相当快的。可惜的是，这与实验结果并不相符——实验结果表明，当温度很低时，热容 C_V 与温度的3次方 T^3 成正比，而这一近似的计算结果与实验结果相比过于低了。这表明爱因斯坦的假设——所有电子的振动频率相同并且固定——是存在一定的问题的。

为了解释这一问题出现的原因，我们将分析晶体的结构。晶体是由**晶格**组成的，晶格是电子的“房间”——电子只能在晶格内部运动，而不会离开晶格。因此，我们可以将任意电子的定态薛定谔方程表示为

$$-\frac{\hbar^2}{2m}\Delta\Psi(x, y, z) + V(x, y, z)\Psi(x, y, z) = E\Psi(x, y, z),$$

其中 $V(x, y, z)$ 描述了晶格对电子的阻碍作用：

$$V(x, y, z) = \begin{cases} 0, & |x|, |y|, |z| \in (0, a), \\ \infty, & \text{otherwise.} \end{cases}$$

分离变量 $\Psi(x, y, z) = F(x)G(y)H(z)$ ，很容易得出

$$\Psi(x, y, z) = \begin{cases} \left(\frac{2}{a}\right)^{3/2} \sin \frac{n_1 \pi x}{a} \sin \frac{n_2 \pi y}{a} \sin \frac{n_3 \pi z}{a}, & |x|, |y|, |z| \in (0, a), \\ 0, & \text{otherwise.} \end{cases}$$

其中 $n_1, n_2, n_3 > 0$ 是正整数，我们将它记为 Ψ_{n_1, n_2, n_3} 。对应的能量为

$$E_{n_1, n_2, n_3} = \frac{(n_1^2 + n_2^2 + n_3^2) \pi^2 \hbar^2}{2ma^2}.$$

根据(1.16)可得，对于初始时刻为 Ψ_{n_1, n_2, n_3} 的波函数，它在任意时刻 t 将演化为

$$\Psi_{n_1, n_2, n_3} e^{-iE_{n_1, n_2, n_3} t / \hbar}.$$

它是一个频率为 $E_{n_1, n_2, n_3} / \hbar$ 的函数——这表明，波函数为 Ψ_{n_1, n_2, n_3} 的电子振动频率为

$$\frac{E_{n_1, n_2, n_3}}{\hbar} = \frac{(n_1^2 + n_2^2 + n_3^2) \pi^2 \hbar}{2ma^2}.$$

给定固定的频率 ω ，当电子振动频率小于 ω ，对应的 (n_1, n_2, n_3) 满足

$$n_1^2 + n_2^2 + n_3^2 < \frac{2ma^2\omega}{\pi^2\hbar}.$$

将 (n_1, n_2, n_3) 视为半径为 $2ma^2\omega/\pi^2\hbar$ 的球内部第一象限的格点。因为每个格点占据的体积为1，所以球内部第一象限的格点数量可以近似为它自身体积的1/8:

$$V(\omega) = \frac{1}{6}\pi \left(\frac{2ma^2\omega}{\pi^2\hbar} \right)^3 = \frac{4m^3a^6\omega^3}{3\pi^5\hbar^3}.$$

这表明，当电子振动频率在 $(\omega, \omega + d\omega)$ 之间，其对应的 (n_1, n_2, n_3) 所有可能的数量为

$$dV = V'(\omega)d\omega = \frac{4m^3a^6\omega^2}{\pi^5\hbar^3}d\omega.$$

因此，电子振动频率在 $(\omega, \omega + d\omega)$ 之间的数量正比于

$$C\omega^2 d\omega,$$

其中 C 是与频率、温度无关的常数。

德拜认为电子的振动频率是有上限的，并且电子的数量分布正如之前的推导所示。因此，其上限 ω_0 满足

$$\int_0^{\omega_0} C\omega^2 d\omega = N,$$

其中 N 是电子数量。容易得出 $\omega_0 = (3N/C)^{1/3}$ ，因此其取值只与材料的性质有关，将 ω_0 称为**德拜频率**。

在德拜的假设下，根据(1.25)，我们就可以得出所有电子的总能量为

$$\begin{aligned} E &= \int_0^{\omega_0} C\omega^2 \left(\frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1} \right) d\omega \\ &= \frac{1}{8}C\hbar\omega_0^4 + \frac{Ck^4T^4}{\hbar^3} \int_0^{\hbar\omega_0/kT} \frac{x^3}{e^x - 1} dx. \end{aligned}$$

在高温情况下, $\hbar\omega_0/kT \ll 1$ 。此时利用 $x \ll 1$ 的近似 $e^x \approx 1 + x$, 积分的原函数可以近似为 x^2 。因此

$$E \approx \frac{1}{8}C\hbar\omega_0^4 + \frac{Ck^4T^4}{\hbar^3} \int_0^{\hbar\omega_0/kT} x^2 dx = \frac{1}{8}C\hbar\omega_0^4 + \frac{1}{3}C\omega_0^3kT.$$

晶体的热容为

$$C_V = \frac{\partial E}{\partial T} = \frac{1}{3}C\omega_0^3k,$$

其取值与温度无关, 这与实验结果自然是相符的。

在低温情况下, $\hbar\omega_0/kT \rightarrow \infty$ 。而积分

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}$$

是收敛的, 因此

$$E \approx \frac{1}{8}C\hbar\omega_0^4 + \frac{1}{15} \frac{C\pi^4k^4T^4}{\hbar^3}.$$

晶体的热容为

$$C_V = \frac{\partial E}{\partial T} = \frac{4}{15} \frac{C\pi^4k^4T^3}{\hbar^3},$$

这与 T^3 成正比! 因此, 德拜的这一假设正好与实验结果相符。

1.5 热方程*

³ 热方程是关于函数 $u(x, t) : \mathbb{R}^n \times (0, \infty) \rightarrow \mathbb{R}$ 的偏微分方程:

$$u_t - \Delta u = 0.$$

其中 $u_t = \frac{\partial u}{\partial t}$, $\Delta u = \sum_{i=1}^n \frac{\partial^2 u}{\partial (x^i)^2}$ 。

热方程具有一定的物理含义。我们可以将 u 理解为空间中的温度分布, 因此, 对于有界闭集 $\Omega \subseteq \mathbb{R}^n$, 其内部的温度满足

$$\frac{d}{dt} \int_{\Omega} u dV = - \int_{\partial\Omega} F \cdot dA,$$

其中 F 是单位时间、单位面积向外交换的热量大小。根据Stokes定理可得

$$u_t = -\operatorname{div} F.$$

而 F 与温度梯度成正比, 因此

$$F = -\nabla u,$$

这表明

$$u_t = \operatorname{div}(\nabla u) = \Delta u.$$

³符号*表示这一章节可以跳过, 建议跳转到第36页的章节2.1。

1.5.1 热方程的基本解*

⁴ 因为热方程对空间坐标计算2阶导数，对时间计算1阶导数。所以如果 $u(x, t)$ 是热方程的解，那么 $u(\lambda x, \lambda^2 t)$ 也是热方程的解，其中 $\lambda \in \mathbb{R}$ 。通过这一现象，假设 u 可以被分离为空间坐标的函数与时间的函数之积：

$$u(x, t) = \frac{1}{t^\alpha} v\left(\frac{x}{t^{1/2}}\right),$$

其中 $\alpha \in \mathbb{R}$ 是常数， $v: \mathbb{R}^n \rightarrow \mathbb{R}$ 是空间坐标的函数。

因此

$$\begin{aligned} u_t &= -\frac{\alpha}{t^{\alpha+1}} v(y) - \frac{y}{2t^{\alpha+1}} \cdot \nabla v(y), \\ \Delta u &= \frac{1}{t^{\alpha+1}} \Delta v(y), \end{aligned}$$

其中 $y = t^{-1/2}x$ 。代入热方程 $u_t - \Delta u = 0$ 可得

$$\alpha v + \frac{1}{2} y \cdot \nabla v + \Delta v = 0.$$

进一步假设 v 是球对称的，即 $v(y) = w(|y|)$ 。因此

$$\alpha w + \frac{1}{2} r w' + w'' + \frac{n-1}{r} w' = 0,$$

其中 $r = |y|$, $w' = \frac{dw}{dr}$ 。令 $\alpha = n/2$ ，因此

$$(r^{n-1} w')' + \frac{1}{2} (r^n w)' = 0.$$

这表明

$$r^{n-1} w' + \frac{1}{2} r^n w = a,$$

其中 a 是常数。假设 $\lim_{r \rightarrow \infty} w = 0$, $\lim_{r \rightarrow \infty} w' = 0$ ，我们可以得出 $a = 0$ 。因此

$$w' = -\frac{1}{2} r w.$$

这表明

$$w = b e^{-r^2/4},$$

其中 b 是常数。因此 $\frac{b}{t^{n/2}} e^{-|x|^2/4t}$ 是热方程的一个解。

据此，定义热方程的**基本解**为

$$\Phi(x, t) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t}, \quad x \in \mathbb{R}^n, t > 0.$$

可以注意到 Φ 在 $(0, 0)$ 是奇异的，并且 Φ 在空间上是球对称的，只与 $|x|$ 有关。系数 $(4\pi)^{-n/2}$ 的选取是因为以下的引理：

⁴符号*表示这一章节可以跳过，建议跳转到第36页的章节2.1。

Lemma 1.5. 对于任意 $t > 0$,

$$\int_{\mathbb{R}^n} \Phi(x, t) dV = 1.$$

Proof.

$$\begin{aligned} \int_{\mathbb{R}^n} \Phi(x, t) dV &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x|^2/4t} dV \\ &= \frac{1}{(4\pi t)^{n/2}} \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-(x^i)^2/4t} dx^i \\ &= \frac{1}{\pi^{n/2}} \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-(y^i)^2} dy^i \\ &= 1, \end{aligned}$$

其中 $y^i = 2t^{1/2}x^i$ 。

□

基本解可以用于初值问题

$$\begin{aligned} u_t - \Delta u &= 0, & u &\in \mathbb{R}^n \times (0, \infty), \\ u &= g, & u &\in \mathbb{R}^n \times \{0\}. \end{aligned}$$

因为 $\Phi(x, t)$ 是热方程的一个解，所以对于任意固定的 $y \in \mathbb{R}^n$ ， $\Phi(x - y, t)$ 同样也是热方程的一个解。这表明卷积

$$\begin{aligned} u(x, t) &= \int_{\mathbb{R}^n} \Phi(x - y, t) g(y) dV_y \\ &= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/4t} g(y) dV_y, \quad x \in \mathbb{R}^n, t > 0 \end{aligned}$$

应该也是热方程的一个解。我们将通过严格的计算进行论证：

Theorem 1.6. 假设 g 是连续并有界 (存在 $M > 0$ 使得 $|g(x)| < M$ 对于任意 $x \in \mathbb{R}^n$ 都成立) 的函数， u 是前面定义的卷积。那么

(A) u 在 $t > 0$ 是光滑函数 (对 x, t 任意阶可导)，记为 $u \in C^\infty(\mathbb{R}^n \times (0, \infty))$ 。

(B) 对于任意 $x \in \mathbb{R}^n, t > 0$,

$$u_t(x, t) - \Delta u(x, t) = 0.$$

(C) $u(x, t)$ 在 $t \searrow 0$ 收敛于初始条件 $g(x)$ 。相当于，对于任意 $p \in \mathbb{R}^n$,

$$\lim_{(x, t) \rightarrow (p, 0)} u(x, t) = g(p).$$

Proof. 第1个命题是显然的。第2个命题则是很容易的计算:

$$\begin{aligned} u_t(x, t) - \Delta u(x, t) &= \int_{\mathbb{R}^n} (\Phi_t(x - y, t) - \Delta_x \Phi(x - y, t)) g(y) dy \\ &= 0, \end{aligned}$$

其中 Δ_x 相当于对变量 x 计算二阶导数。

对于第3个命题, 固定 $p \in \mathbb{R}^n, \epsilon > 0$ 。根据 g 的连续性, 我们可以找出 $\delta > 0$ 使得, 对于任意 $|y - p| < \delta$ 都具有

$$|g(y) - g(p)| < \epsilon.$$

因此, 如果 $|x - p| < \delta/2$, 根据Lemma 1.5可得

$$\begin{aligned} |u(x, t) - g(p)| &= \left| \int_{\mathbb{R}^n} \Phi(x - y, t) (g(y) - g(p)) dV_y \right| \\ &\leq \int_{\mathbb{R}^n} \Phi(x - y, t) |g(y) - g(p)| dV_y \\ &= \int_{B_\delta(p)} \Phi(x - y, t) |g(y) - g(p)| dV_y \\ &\quad + \int_{\mathbb{R}^n \setminus B_\delta(p)} \Phi(x - y, t) |g(y) - g(p)| dV_y \\ &= I + J. \end{aligned}$$

其中 $B_\delta(p)$ 是以 p 为球心, 半径为 δ 的球。而

$$I \leq \epsilon \int_{B_\delta(p)} \Phi(x - y, t) dV_y < \epsilon \int_{\mathbb{R}^n} \Phi(x - y, t) dV_y = \epsilon.$$

此外, 如果 $|x - p| \leq \delta/2, |y - p| \geq \delta$, 那么根据三角不等式可得

$$|y - p| \leq |y - x| + |x - p| \leq |y - x| + \frac{\delta}{2} \leq |y - x| + \frac{1}{2} |y - p|.$$

因此 $|y - x| \geq \frac{1}{2}|y - p|$ 。根据 g 的有界性，存在 $M > 0$ 使得 $|g(x)| < M$ 对于任意 $x \in \mathbb{R}^n$ 都成立。那么

$$\begin{aligned}
J &\leq \int_{\mathbb{R}^n \setminus B_\delta(p)} \Phi(x - y, t) (|g(y)| + |g(p)|) dV_y \\
&< 2M \int_{\mathbb{R}^n \setminus B_\delta(p)} \Phi(x - y, t) dV_y \\
&= \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B_\delta(p)} e^{-|x-y|^2/4t} dV_y \\
&\leq \frac{C}{t^{n/2}} \int_{\mathbb{R}^n \setminus B_\delta(p)} e^{-|y-p|^2/16t} dV_y \\
&= C \int_{\mathbb{R}^n \setminus B_{\delta/\sqrt{t}}(0)} e^{-|z|^2/16} dV_z,
\end{aligned}$$

当 $t \searrow 0$ 时， $B_{\delta/\sqrt{t}}(0)$ 倾向于“占据”整个空间 \mathbb{R}^n ，而 $e^{-|z|^2/16}$ 是指数衰减的，因此 $\lim_{t \searrow 0} J = 0$ 。这表明，如果 $|x - p| < \delta/2$ 并且 t 足够小， $|u(x, t) - g(p)| = I + J < 2\epsilon$ 。□

Theorem 1.7. 假设 u 是方程

$$\begin{aligned}
u_t - \Delta u &= 0, & u &\in \mathbb{R}^n \times (0, T), \\
u &= g, & u &\in \mathbb{R}^n \times \{0\}.
\end{aligned}$$

的一个解，并且 $u(x, t)$ 的增长速率不能太快，

$$u(x, t) \leq Ae^{a|x|^2}, \quad x \in \mathbb{R}^n, t \in [0, T],$$

其中 $A, a > 0$ 是常数。那么

$$\sup_{\mathbb{R}^n \times [0, T]} u = \sup_{\mathbb{R}^n} g.$$

这一定理的证明技巧性较强，我们不做展开。但是其直观性也是很强的：空间中的温度随时间变化是趋于平均的，因此在 $t > 0$ 的任意时刻，空间中的最高温度不可能高于 $t = 0$ 时刻空间中的最高温度。

1.5.2 传播子*

⁵ 通常情况下，薛定谔方程没有通用的求解方式：正如前面的章节所示，即使势能函数是非常简单的二次函数 $V = \frac{1}{2}m\omega^2 x^2$ ，方程解的形式还是相当复杂的。但是，如果令势能函数为0，那么薛定谔方程为

$$-\frac{\hbar}{2m} \Delta \Psi = i\Psi_t.$$

⁵符号*表示这一章节可以跳过，建议跳转到第36页的章节2.1。

令 $u(x, t) = \Psi\left(\sqrt{\frac{h}{2m}}x, t\right)$, 那么

$$iu_t + \Delta u = 0,$$

它与热方程是类似的, 只相差了一个常数 i 。那么我们可以进行假设, 将 t 替换为 it , 得出它的“基本解”(称为传播子)为

$$\Phi(x, t) = \frac{C}{t^{n/2}} e^{-|x|^2/4it}.$$

其中 $C \in \mathbb{C}$ 是常数。因此, 对于初值问题

$$\begin{aligned} iu_t + \Delta u &= 0, & u &\in \mathbb{R}^n \times (0, \infty), \\ u &= g, & u &\in \mathbb{R}^n \times \{0\}. \end{aligned}$$

卷积

$$\begin{aligned} u(x, t) &= \int_{\mathbb{R}^n} \Phi(x - y, t) g(y) dV_y \\ &= \frac{C}{t^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/4it} g(y) dV_y, & x \in \mathbb{R}^n, t > 0 \end{aligned}$$

是薛定谔方程的一个解。

可惜的是, $\Phi(x, t)$ 并不是随着 $|x| \rightarrow \infty$ 指数衰减的, 因为 $\left|e^{-|x|^2/4it}\right|$ 的指数部分是纯虚数, 所以 $\left|e^{-|x|^2/4it}\right| \equiv 1$ 。这表明积分 $\int_{\mathbb{R}^n} \Phi(x, t) dV$ 是发散的, 因此, 我并没有将常数 C 直接写为 $1/(4\pi)^{n/2}$ 。此外, 证明 $u(x, t)$ 在 $t \searrow 0$ 是否收敛于 $g(x)$ 也是一个难题: 它需要利用路径积分的方法, 其原理相当复杂, 并且我附近没有参考资料, 因此不做深入讨论。

2 线性算子

注意：除非特殊声明，默认 $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ ，即 x 是 \mathbb{R}^n 上的向量；并且 $dV = dx^1 \cdots dx^n$ 代表 \mathbb{R}^n 上的体积形式。

线性算子能对波函数进行操作，例如对波函数计算微分、将波函数与一个常数相乘。与线性代数的矩阵类似，线性算子同样具有本征态，以及对应的本征值，对应线性代数的特征向量与特征值，其含义也不难得知。但是，不同的是，线性算子具有本征值的个数可能是无穷的， n 阶矩阵具有本征值的个数为 n ，自然不可能是无穷大。因此，我们可以将线性算子视为“无穷阶的矩阵”。

量子力学之所以与经典力学存在差异，关键的原因如下：在量子力学体系中，很多物理量并不是连续的，而是离散的：例如氢原子核外电子的角动量只能为 $m\hbar$ 、谐振子的能量只能为 $(n + 1/2)\hbar\omega$ ，其中 m, n 是自然数。但是，如果引入了线性算子这一概念，那么这一现象就可以得到解释：这些物理量对应线性算子的本征值，它们是离散的。测量体系的物理量，对应于波函数往线性算子的本征态上投影，并且对于不同的本征态，投影到其上的概率也不尽相同，它和波函数与本征态“接近”的程度有关——当然，这是不严谨的表述。

2.1 定义

定义**线性算子** $A \in \text{End}_{\mathbb{C}}(C^\infty(\mathbb{R}^n))$ 为光滑函数集合的 \mathbb{C} -线性映射 $A : C^\infty(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^n)$ ，即对于任意光滑函数 $\Psi, \Phi \in C^\infty(\mathbb{R}^n)$ 与常数 $\alpha \in \mathbb{C}$ ，都具有

$$\begin{aligned} A(\Psi + \Phi) &= A\Psi + A\Phi, \\ A(\alpha\Psi) &= \alpha(A\Psi). \end{aligned} \tag{2.1}$$

定义**单位算子** I 为

$$I\Psi \equiv \Psi.$$

对于常数 $\alpha \in \mathbb{C}$ ， αA 也是线性算子，定义为

$$(\alpha A)\Psi = \alpha(A\Psi). \tag{2.2}$$

对于线性算子 A, B ，定义其加法运算与乘法运算为

$$\begin{aligned} (A + B)\Psi &= A\Psi + B\Psi, \\ (AB)\Psi &= A(B\Psi). \end{aligned} \tag{2.3}$$

很显然 $A + B, AB$ 是线性算子。我们将 n 个线性算子 A 的乘积简记为 A^n ，并且 $A^0 = I$ 。因此，线性算子的集合是复数域 \mathbb{C} 上的向量空间；但是，它具有比一般向量空间更强的性质，详见Theorem 2.2；并且，我们在下一小节将分析具有这种性质的向量空间。

Lemma 2.1. 对于任意线性算子 A, B, C 与常数 $\alpha, \beta \in \mathbb{C}$ ，都具有

$$\begin{aligned} (\alpha A + \beta B)C &= \alpha AC + \beta BC, \\ A(\alpha B + \beta C) &= \alpha AB + \beta AC. \end{aligned}$$

Proof. 对于任意光滑函数 $\Psi \in C^\infty(\mathbb{R}^n)$, 可得

$$\begin{aligned}
((\alpha A + \beta B)C)\Psi &= (\alpha A + \beta B)(C\Psi) && \text{According to (2.1)} \\
&= (\alpha A)(C\Psi) + (\beta B)(C\Psi) && \text{According to (2.3)} \\
&= \alpha(A(C\Psi)) + \beta(B(C\Psi)) && \text{According to (2.2)} \\
&= \alpha((AC)\Psi) + \beta((BC)\Psi) && \text{According to (2.3)} \\
&= (\alpha AC)\Psi + (\beta BC)\Psi && \text{According to (2.2)} \\
&= (\alpha AC + \beta BC)\Psi, && \text{According to (2.3)}
\end{aligned}$$

第2个表达式证明原理类似, 不做赘述。 \square

定义线性算子的对易子为

$$[A, B] = AB - BA,$$

很显然 $[A, B]$ 也是线性算子, 并且具有以下关系:

Theorem 2.2. 对于任意线性算子 A, B, C 与常数 $\alpha, \beta \in \mathbb{C}$, 都具有

$$\begin{aligned}
[A, A] &= 0, \\
[A, B] &= -[B, A], \\
[\alpha A + \beta B, C] &= \alpha[A, C] + \beta[B, C], \\
[A, \alpha B + \beta C] &= \alpha[A, B] + \beta[A, C], \\
[A, BC] &= [A, B]C + B[A, C], \\
[AB, C] &= [A, C]B + A[B, C], \\
[A, [B, C]] + [B, [C, A]] + [C, [A, B]] &= 0.
\end{aligned} \tag{2.4}$$

Proof. 1, 2表达式根据定义是显然的; 3, 4表达式根据Lemma 2.1是显然的; 根据第2个表达式, 如果能证明第5个表达式, 那么第6个表达式的证明就迎刃而解了。所以我们只需证明5, 7表达式即可。因为

$$\begin{aligned}
[A, BC] &= ABC - BCA, \\
[A, B]C + B[A, C] &= (AB - BA)C + B(AC - CA) = ABC - BCA.
\end{aligned}$$

所以第5个表达式成立。因为

$$\begin{aligned}
[A, [B, C]] &= A[B, C] - [B, C]A = ABC - ACB - BCA + CBA, \\
[B, [C, A]] &= B[C, A] - [C, A]B = BCA - BAC - CAB + ACB, \\
[C, [A, B]] &= C[A, B] - [A, B]C = CAB - CBA - ABC + BAC,
\end{aligned}$$

求和可以发现全部抵消, 所以第7个表达式成立。 \square

2.1.1 李代数*

⁶ 定义李代数为域 \mathbb{F} 上的向量空间 \mathfrak{g} ：它具有代数运算 $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ ，记为 $(X, Y) \mapsto [X, Y]$ ，我们将它称为李括号。并且，对于任意向量 $X, Y, Z \in \mathfrak{g}$ 与标量 $\alpha, \beta \in \mathbb{F}$ ，都满足

$$\begin{aligned} [X, Y] &= -[Y, X], \\ [\alpha X + \beta Y, Z] &= \alpha [X, Z] + \beta [Y, Z], \\ [X, \alpha Y + \beta Z] &= \alpha [X, Y] + \beta [X, Z], \\ [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] &= 0. \end{aligned} \tag{2.5}$$

如果 \mathfrak{g} 是李代数， \mathfrak{g} 的线性子空间 $\mathfrak{h} \subseteq \mathfrak{g}$ 被称为李代数 \mathfrak{g} 的子代数，当且仅当它在李括号下封闭：对于任意 $X, Y \in \mathfrak{h}$ ，都具有 $[X, Y] \in \mathfrak{h}$ 。如果 $\mathfrak{g}, \mathfrak{h}$ 是李代数， \mathbb{F} -线性映射 $A : \mathfrak{g} \rightarrow \mathfrak{h}$ 被称为李代数同态，当且仅当 A 保持李括号： $A[X, Y] = [AX, AY]$ 。当李代数同态 A 存在逆映射 $A^{-1} : \mathfrak{h} \rightarrow \mathfrak{g}$ ，并且 A^{-1} 也是李代数同态，那么称 A 为李代数同构。对于李代数 \mathfrak{g} 的子代数 \mathfrak{h} ，如果 $[X, Y] \in \mathfrak{h}$ 对于任意 $X \in \mathfrak{g}, Y \in \mathfrak{h}$ 都成立，那么称 \mathfrak{h} 为李代数 \mathfrak{g} 的理想。

我们将给出几个具体实例：

- (A) 线性算子的集合是 \mathbb{C} 上的向量空间，定义李括号为对易子 $[A, B] = AB - BA$ ，根据Theorem 2.2，它是李代数。
- (B) 令实数域 \mathbb{R} 上的向量空间 \mathbb{R}^3 的李括号为向量的外积 $[\mathbf{A}, \mathbf{B}] = \mathbf{A} \times \mathbf{B}$ 。很显然，它满足(2.5)的1-3表达式；对于第4个表达式，我们可以利用向量外积的性质

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C},$$

不难得出，第4个表达式是成立的。

- (C) 记 $M(n, \mathbb{C})$ 为所有 n 阶矩阵的集合，它是实数域 \mathbb{R} 上的 $2n^2$ 维向量空间。定义李括号为 $[A, B] = AB - BA$ 。类似于线性算子，它是李代数。
- (D) 记 $M(n, \mathbb{R})$ 为所有 n 阶实矩阵的集合，它是实数域 \mathbb{R} 上的 n^2 维向量空间，是 $M(n, \mathbb{C})$ 的线性子空间。采用相同的方式定义李括号，可以发现，如果 A, B 是实矩阵，那么 $[A, B] = AB - BA$ 也是实矩阵。因此， $M(n, \mathbb{R})$ 是李代数 $M(n, \mathbb{C})$ 的子代数。
- (E) \mathbb{R}^n 上的向量场定义为向量值函数 $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ，具体形式为

$$V(x^1, \dots, x^n) = (V^1(x^1, \dots, x^n), \dots, V^n(x^1, \dots, x^n)),$$

其中 $V^1, \dots, V^n : \mathbb{R}^n \rightarrow \mathbb{R}$ 。相当于将 \mathbb{R}^n 上的每一点都赋予一个向量：例如电场、磁场就是 \mathbb{R}^3 上的向量场，我们可以讨论 \mathbb{R}^3 上每一点的电场强度 $\mathbf{E} \in \mathbb{R}^3$ 、磁感应强度 $\mathbf{B} \in \mathbb{R}^3$ 。光滑向量场对应 $V^1, \dots, V^n : \mathbb{R}^n \rightarrow \mathbb{R}$ 均为光滑函数。

⁶符号*表示这一章节可以跳过，建议跳转到第39页的章节2.2。

记 $\mathcal{X}(\mathbb{R}^n)$ 为 \mathbb{R}^n 上的光滑向量场的集合。令向量场 $X, Y \in \mathcal{X}(\mathbb{R}^n)$ ，其分量分别记为 $X^1, \dots, X^n; Y^1, \dots, Y^n$ 。定义集合 $\mathcal{X}(\mathbb{R}^n)$ 的李括号 $[X, Y]$ 为另一个光滑向量场：

$$[X, Y] = \left(\sum_{i=1}^n X^i \frac{\partial Y^1}{\partial x^i} - Y^i \frac{\partial X^1}{\partial x^i}, \dots, \sum_{i=1}^n X^i \frac{\partial Y^n}{\partial x^i} - Y^i \frac{\partial X^n}{\partial x^i} \right), \quad (2.6)$$

因此，在该李括号的定义下， $\mathcal{X}(\mathbb{R}^n)$ 是李代数，其计算较为复杂，在此就不做验证了。

(F) 对于任意向量空间 V ，如果我们定义所有李括号都为0，即 $[A, B] = 0$ 对所有 $A, B \in V$ 均成立。那么很显然它是李代数，因为(2.5)的所有表达式取值均为0。我们将这个李代数称为**Abelian**。

注意：李代数没有结合律，即 $[[X, Y], Z]$ 与 $[X, [Y, Z]]$ 不一定相等，正如 \mathbb{R}^3 上向量的外积不满足 $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ 一样。

2.2 伴随与厄米算子

对于线性算子 A 与任意光滑函数 Ψ, Φ ，存在唯一的线性算子 A^\dagger 满足

$$\langle \Psi, A\Phi \rangle \equiv \langle A^\dagger \Psi, \Phi \rangle,$$

我们将 A^\dagger 称为线性算子 A 的**伴随算子**。

Theorem 2.3. 对于任意线性算子 A, B 与常数 $\lambda \in \mathbb{C}$ ，都具有

$$\begin{aligned} (A^\dagger)^\dagger &= A, \\ (A + B)^\dagger &= A^\dagger + B^\dagger, \\ (\lambda A)^\dagger &= \bar{\lambda} A^\dagger, \\ (AB)^\dagger &= B^\dagger A^\dagger. \end{aligned}$$

Proof. 根据(1.10)可得

$$\begin{aligned}
\langle \Psi, A\Phi \rangle &= \langle A^\dagger \Psi, \Phi \rangle = \overline{\langle \Phi, A^\dagger \Psi \rangle} = \overline{\langle (A^\dagger)^\dagger \Phi, \Psi \rangle} \\
&= \langle \Psi, (A^\dagger)^\dagger \Phi \rangle, \\
\langle (A+B)^\dagger \Psi, \Phi \rangle &= \langle \Psi, (A+B)\Phi \rangle = \langle \Psi, A\Phi + B\Phi \rangle = \langle \Psi, A\Phi \rangle + \langle \Psi, B\Phi \rangle \\
&= \langle A^\dagger \Psi, \Phi \rangle + \langle B^\dagger \Psi, \Phi \rangle = \langle A^\dagger \Psi + B^\dagger \Psi, \Phi \rangle \\
&= \langle (A^\dagger + B^\dagger) \Psi, \Phi \rangle, \\
\langle (\lambda A)^\dagger \Psi, \Phi \rangle &= \langle \Psi, \lambda A\Phi \rangle = \lambda \langle \Psi, A\Phi \rangle = \lambda \langle A^\dagger \Psi, \Phi \rangle \\
&= \langle \bar{\lambda} A^\dagger \Psi, \Phi \rangle, \\
\langle (AB)^\dagger \Psi, \Phi \rangle &= \langle \Psi, AB\Phi \rangle = \langle A^\dagger \Psi, B\Phi \rangle = \langle B^\dagger (A^\dagger \Psi), \Phi \rangle \\
&= \langle (B^\dagger A^\dagger) \Psi, \Phi \rangle.
\end{aligned}$$

□

当 $A^\dagger = A$ 时, A 被称为**厄米算子**, 即对于任意光滑函数 Ψ, Φ , 均满足

$$\langle \Psi, A\Phi \rangle = \langle A\Psi, \Phi \rangle. \quad (2.7)$$

本课程涉及的绝大多数线性算子均为厄米算子。

2.2.1 测量

对于厄米算子 A , 可能存在一组归一化的光滑函数 $\{\Psi_n\}$ 与一组常数 $\lambda_n (n = 1, 2, 3, \dots)$ 使得

$$\begin{aligned}
A\Psi_n &= \lambda_n \Psi_n, \\
\|\Psi_n\| &= 1.
\end{aligned} \quad (2.8)$$

我们将 Ψ_n 与 λ_n 分别称为 A 的**本征态**与**本征值**。它们是标准正交的:

$$\begin{aligned}
\lambda_n &\in \mathbb{R}, \\
\langle \Psi_i, \Psi_j \rangle &= \delta_{ij}.
\end{aligned} \quad (2.9)$$

我们将厄米算子 A 称为**可观测量**。假设粒子的波函数 $\Psi (\|\Psi\| = 1)$ 可以表示为本征态的展开

$$\Psi = \sum_{i=1}^{\infty} c_i \Psi_i,$$

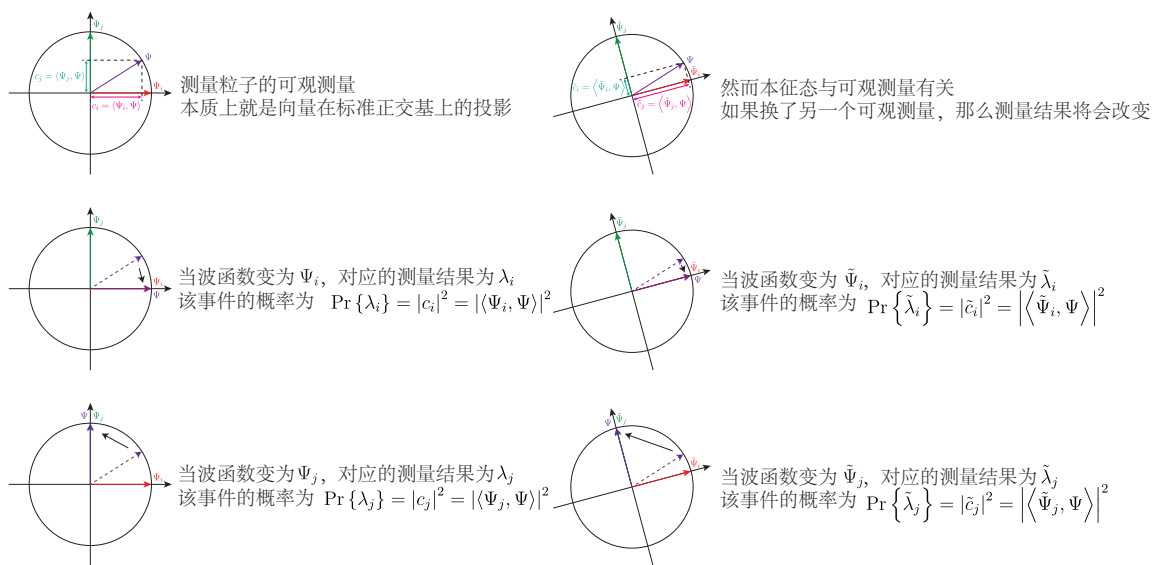


Figure 2.1: 测量

根据(1.10)与(2.9)可得

$$c_i = \langle \Psi_i, \Psi \rangle,$$

$$\sum_{i=1}^{\infty} |c_i|^2 = 1.$$

测量该可观测量，得到的结果只能为 A 的本征值 $\lambda_i, i = 1, 2, 3, \dots$ ，并且测量结果为 λ_i 的概率为

$$\Pr\{\lambda_i\} = |c_i|^2, \quad (2.10)$$

因此

$$\sum_{i=1}^{\infty} \Pr\{\lambda_i\} = \sum_{i=1}^{\infty} |c_i|^2 = 1.$$

测量的平均值为

$$\sum_{i=1}^{\infty} \lambda_i \Pr\{\lambda_i\} = \sum_{i=1}^{\infty} \lambda_i |c_i|^2. \quad (2.11)$$

如果测量结果为 λ_i ，那么在测量之后的瞬间，波函数将直接变为 Ψ_i 。

因此，“测量”的本质是向量在标准正交基上的投影，如Figure 2.1所示。

2.3 常用的线性算子

对于任意 $i, j \in \{1, \dots, n\}$, 定义位置算子 X_i 与动量算子 P_i 为

$$\begin{aligned} X_i \Psi(x, t) &= x^i \Psi(x, t), \\ P_i \Psi(x, t) &= -i\hbar \frac{\partial \Psi}{\partial x^i}(x, t), \end{aligned} \quad (2.12)$$

它们均为厄米算子。可以得出以下结论:

Theorem 2.4. 对于任意 $i, j \in \{1, \dots, n\}$, 都具有

$$\begin{aligned} [X_i, X_j] &= 0, \\ [P_i, P_j] &= 0, \\ [X_i, P_j] &= i\hbar \delta_{ij} I. \end{aligned} \quad (2.13)$$

Proof. 前2个表达式都是很显然的, 因为 $x^i x^j = x^j x^i$, $\frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i}$ 。
对于第3个表达式, 当 $i \neq j$ 时, 对于任意 $\Psi(x, t)$, 都具有

$$\begin{aligned} [X_i, P_j] \Psi &= X_i \left(-i\hbar \frac{\partial \Psi}{\partial x^j} \right) - P_j (x^i \Psi) \\ &= -i\hbar x^i \frac{\partial \Psi}{\partial x^j} + i\hbar \frac{\partial (x^i \Psi)}{\partial x^j} \\ &= i\hbar \left(-x^i \frac{\partial \Psi}{\partial x^j} + x^i \frac{\partial \Psi}{\partial x^j} \right) \\ &= 0, \end{aligned}$$

因此 $[X_i, P_j] = 0$ 。当 $i = j$ 时, 对于任意 $\Psi(x, t)$, 都具有

$$\begin{aligned} [X_i, P_i] \Psi &= X_i \left(-i\hbar \frac{\partial \Psi}{\partial x^i} \right) - P_i (x^i \Psi) \\ &= -i\hbar x^i \frac{\partial \Psi}{\partial x^i} + i\hbar \frac{\partial (x^i \Psi)}{\partial x^i} \\ &= i\hbar \left(-x^i \frac{\partial \Psi}{\partial x^i} + x^i \frac{\partial \Psi}{\partial x^i} + \Psi \right) \\ &= i\hbar \Psi, \end{aligned}$$

因此 $[X_i, P_i] = i\hbar I$ 。 □

因此, 根据(2.18)可得

$$\Delta X_i \Delta P_i \geq \frac{\hbar}{2}, \quad i = 1, \dots, n.$$

这是测不准原理的一个重要应用。

2.3.1 期望值与测不准原理

在波函数为 Ψ ($\|\Psi\| = 1$)的情况下, 厄米算子 A 的期望值为

$$\langle A \rangle = \langle \Psi, A\Psi \rangle = \int_{\mathbb{R}^n} \bar{\Psi}(A\Psi) dV. \quad (2.14)$$

因此, 单位算子 I 的期望值为

$$\langle I \rangle = \langle \Psi, I\Psi \rangle = \langle \Psi, \Psi \rangle = 1.$$

任意厄米算子 A 的期望值均为实数, 利用(1.10)与(2.7)不难得出:

$$\langle \Psi, A\Psi \rangle = \langle A\Psi, \Psi \rangle = \overline{\langle \Psi, A\Psi \rangle}.$$

因此, 位置算子 X_i, X_i^2 的期望值为

$$\begin{aligned} \langle X_i \rangle &= \int_{\mathbb{R}^n} \bar{\Psi}(X_i\Psi) dV = \int_{\mathbb{R}^n} x^i |\Psi|^2 dV, \\ \langle X_i^2 \rangle &= \int_{\mathbb{R}^n} \bar{\Psi}(X_i^2\Psi) dV = \int_{\mathbb{R}^n} (x^i)^2 |\Psi|^2 dV. \end{aligned} \quad (2.15)$$

动量算子 P_i, P_i^2 期望值为

$$\begin{aligned} \langle P_i \rangle &= \int_{\mathbb{R}^n} \bar{\Psi}(P_i\Psi) dV = (-i\hbar) \int_{\mathbb{R}^n} \bar{\Psi} \frac{\partial \Psi}{\partial x^i} dV, \\ \langle P_i^2 \rangle &= \int_{\mathbb{R}^n} \bar{\Psi}(P_i^2\Psi) dV = (-i\hbar)^2 \int_{\mathbb{R}^n} \bar{\Psi} \frac{\partial^2 \Psi}{\partial (x^i)^2} dV \\ &= -\hbar^2 \int_{\mathbb{R}^n} \bar{\Psi} \frac{\partial^2 \Psi}{\partial (x^i)^2} dV \\ &= \hbar^2 \int_{\mathbb{R}^n} \left| \frac{\partial \Psi}{\partial x^i} \right|^2 dV. \end{aligned}$$

其中最后一个等式利用了分部积分法, 如(1.19)所示。

根据Theorem 1.3, 那么这样一来, 我们就给出了 $\langle P_i \rangle, \langle P_i^2 \rangle$ 的另一种计算方式, 其表达式为

$$\begin{aligned}
\langle P_i \rangle &= (-i\hbar) \int_{\mathbb{R}^n} \bar{\Psi} \frac{\partial \Psi}{\partial x^i} dV \\
&= (-i\hbar) \int_{\mathbb{R}^n} \bar{\Phi} \left(\frac{ip^i}{\hbar} \right) \Phi dV_p \\
&= \int_{\mathbb{R}^n} p^i |\Phi|^2 dV_p, \\
\langle P_i^2 \rangle &= (-i\hbar)^2 \int_{\mathbb{R}^n} \bar{\Psi} \frac{\partial^2 \Psi}{\partial (x^i)^2} dV \\
&= (-i\hbar)^2 \int_{\mathbb{R}^n} \bar{\Phi} \left(\frac{ip^i}{\hbar} \right)^2 \Phi dV_p \\
&= \int_{\mathbb{R}^n} (p^i)^2 |\Phi|^2 dV_p.
\end{aligned} \tag{2.16}$$

其中 $dV_p = p^1 \cdots p^n$, Φ 的定义如(1.7)所示。

定义厄米算子 A 的不确定度为

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \tag{2.17}$$

测不准原理的描述如下:

Theorem 2.5.

$$(\Delta A)^2 (\Delta B)^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2. \tag{2.18}$$

Proof. 定义线性算子

$$\begin{aligned}
A' &= A - \langle \Psi, A\Psi \rangle I, \\
B' &= B - \langle \Psi, B\Psi \rangle I.
\end{aligned}$$

根据(1.10)可得

$$\begin{aligned}
\langle \Phi, A'\Theta \rangle &= \langle \Phi, A\Theta \rangle - \langle \Phi, \langle \Psi, A\Psi \rangle \Theta \rangle = \langle \Phi, A\Theta \rangle - \langle \Psi, A\Psi \rangle \langle \Phi, \Theta \rangle, \\
\langle A'\Phi, \Theta \rangle &= \langle A\Phi, \Theta \rangle - \langle \langle \Psi, A\Psi \rangle \Phi, \Theta \rangle = \langle A\Phi, \Theta \rangle - \overline{\langle \Psi, A\Psi \rangle} \langle \Phi, \Theta \rangle.
\end{aligned}$$

因为 A 是厄米算子, 所以它满足(2.7), 并利用(1.10)可得

$$\begin{aligned}
\langle \Phi, A\Theta \rangle &= \langle A\Phi, \Theta \rangle, \\
\langle \Psi, A\Psi \rangle &= \langle A\Psi, \Psi \rangle = \overline{\langle \Psi, A\Psi \rangle}.
\end{aligned}$$

因此

$$\langle \Phi, A' \Theta \rangle = \langle A' \Phi, \Theta \rangle,$$

这证明了 A' 是厄米算子。同理可得, B' 也是厄米算子。根据(1.11)可得

$$\begin{aligned} \langle A' \Psi, A' \Psi \rangle \langle B' \Psi, B' \Psi \rangle &\geq |\langle A' \Psi, B' \Psi \rangle|^2 \\ &\geq |\operatorname{Im} \langle A' \Psi, B' \Psi \rangle|^2 \\ &= \left| \frac{1}{2} \left(\langle A' \Psi, B' \Psi \rangle - \overline{\langle A' \Psi, B' \Psi \rangle} \right) \right|^2 \\ &= \frac{1}{4} |\langle A' \Psi, B' \Psi \rangle - \langle B' \Psi, A' \Psi \rangle|^2 \\ &= \frac{1}{4} |\langle \Psi, A' B' \Psi \rangle - \langle \Psi, B' A' \Psi \rangle|^2 \\ &= \frac{1}{4} |\langle \Psi, [A', B'] \Psi \rangle|^2 \\ &= \frac{1}{4} |\langle [A', B'] \rangle|^2. \end{aligned}$$

利用波函数 $\langle \Psi, \Psi \rangle = 1$, 与 A' 是厄米算子的性质, 可得

$$\begin{aligned} \langle A' \Psi, A' \Psi \rangle &= \langle A \Psi - \langle \Psi, A \Psi \rangle \Psi, A \Psi - \langle \Psi, A \Psi \rangle \Psi \rangle \\ &= \langle A \Psi, A \Psi \rangle - \langle A \Psi, \langle \Psi, A \Psi \rangle \Psi \rangle - \langle \langle \Psi, A \Psi \rangle \Psi, A \Psi \rangle + \langle \langle \Psi, A \Psi \rangle \Psi, \langle \Psi, A \Psi \rangle \Psi \rangle \\ &= \langle \Psi, A^2 \Psi \rangle - 2 \langle \Psi, A \Psi \rangle^2 + \langle \Psi, A \Psi \rangle^2 \\ &= \langle A^2 \rangle - \langle A \rangle^2 \\ &= (\Delta A)^2, \end{aligned}$$

因此

$$(\Delta A)^2 (\Delta B)^2 \geq \frac{1}{4} |\langle [A', B'] \rangle|^2.$$

我们只需证明 $[A', B'] = [A, B]$ 即可, 这是一个很容易的计算:

$$\begin{aligned} [A', B'] &= A' B' - B' A' \\ &= (A - \langle \Psi, A \Psi \rangle I)(B - \langle \Psi, B \Psi \rangle I) - (B - \langle \Psi, B \Psi \rangle I)(A - \langle \Psi, A \Psi \rangle I) \\ &= AB - \langle \Psi, B \Psi \rangle A - \langle \Psi, A \Psi \rangle B + \langle \Psi, A \Psi \rangle \langle \Psi, B \Psi \rangle I \\ &\quad - BA + \langle \Psi, A \Psi \rangle B + \langle \Psi, B \Psi \rangle A - \langle \Psi, B \Psi \rangle \langle \Psi, A \Psi \rangle I \\ &= AB - BA = [A, B]. \end{aligned}$$

□

2.3.2 哈密顿算子*

⁷ 给定势能函数 $V(x)$ ，定义哈密顿算子 H 为

$$H\Psi(x, t) = -\frac{\hbar^2}{2m}\Delta\Psi(x, t) + V(x)\Psi(x, t). \quad (2.19)$$

哈密顿算子是厄米算子。因此，薛定谔方程又可以写为

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t}. \quad (2.20)$$

哈密顿算子的本征值 E_n 即为粒子的能量。

Lemma 2.6. 对于 $i \in \{1, \dots, n\}$ ，位置算子 X_i 与动量算子 P_i 满足

$$\begin{aligned} [X_i, H] &= i\hbar\frac{P_i}{m}, \\ [P_i, H] &= -i\hbar\frac{\partial V}{\partial x^i}I. \end{aligned}$$

Proof. 对于任意 $\Psi(x, t)$ ，都具有

$$\begin{aligned} [X_i, H]\Psi &= X_i\left(-\frac{\hbar^2}{2m}\Delta\Psi + V\Psi\right) - H(x^i\Psi) \\ &= -\frac{\hbar^2}{2m}x^i\Delta\Psi + x^iV\Psi + \frac{\hbar^2}{2m}\Delta(x^i\Psi) - x^iV\Psi \\ &= -\frac{\hbar^2}{2m}(x^i\Delta\Psi - \Delta(x^i\Psi)), \end{aligned}$$

⁷符号*表示这一章节可以跳过，建议跳转到第55页的章节2.5。

而

$$\begin{aligned}
\Delta(x^i \Psi) &= \frac{\partial^2(x^i \Psi)}{\partial(x^i)^2} + x^i \sum_{j=1, j \neq i}^n \frac{\partial^2 \Psi}{\partial(x^j)^2} \\
&= \frac{\partial}{\partial x^i} \frac{\partial(x^i \Psi)}{\partial x^i} + x^i \sum_{j=1, j \neq i}^n \frac{\partial^2 \Psi}{\partial(x^j)^2} \\
&= \frac{\partial}{\partial x^i} \left(\Psi + x^i \frac{\partial \Psi}{\partial x^i} \right) + x^i \sum_{j=1, j \neq i}^n \frac{\partial^2 \Psi}{\partial(x^j)^2} \\
&= \frac{\partial \Psi}{\partial x^i} + \frac{\partial}{\partial x^i} \left(x^i \frac{\partial \Psi}{\partial x^i} \right) + x^i \sum_{j=1, j \neq i}^n \frac{\partial^2 \Psi}{\partial(x^j)^2} \\
&= 2 \frac{\partial \Psi}{\partial x^i} + x^i \frac{\partial^2 \Psi}{\partial(x^i)^2} + x^i \sum_{j=1, j \neq i}^n \frac{\partial^2 \Psi}{\partial(x^j)^2} \\
&= 2 \frac{\partial \Psi}{\partial x^i} + x^i \Delta \Psi,
\end{aligned}$$

因此

$$[X_i, H] \Psi = \frac{\hbar^2}{m} \frac{\partial \Psi}{\partial x^i} = i\hbar \frac{P_i \Psi}{m}.$$

此外

$$\begin{aligned}
[P_i, H] \Psi &= P_i \left(-\frac{\hbar^2}{2m} \Delta \Psi + V \Psi \right) - H \left(-i\hbar \frac{\partial \Psi}{\partial x^i} \right) \\
&= -i\hbar \frac{\partial}{\partial x^i} \left(-\frac{\hbar^2}{2m} \Delta \Psi + V \Psi \right) + i\hbar \left(-\frac{\hbar^2}{2m} \Delta \left(\frac{\partial \Psi}{\partial x^i} \right) + V \frac{\partial \Psi}{\partial x^i} \right) \\
&= \frac{i\hbar^3}{2m} \left(\frac{\partial}{\partial x^i} \Delta \Psi - \Delta \left(\frac{\partial \Psi}{\partial x^i} \right) \right) - i\hbar \left(\frac{\partial V}{\partial x^i} \Psi + V \frac{\partial \Psi}{\partial x^i} - V \frac{\partial \Psi}{\partial x^i} \right).
\end{aligned}$$

因为对于任意 $i, j \in \{1, \dots, n\}$, 都具有 $\frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i}$, 所以 $\frac{\partial}{\partial x^i} \Delta \Psi - \Delta \left(\frac{\partial \Psi}{\partial x^i} \right) = 0$, 即

$$[P_i, H] \Psi = -i\hbar \frac{\partial V}{\partial x^i} \Psi.$$

□

回忆函数内积的定义:

$$\langle \Phi(x, t), \Psi(x, t) \rangle = \int_{\mathbb{R}^n} \bar{\Phi}(x, t) \Psi(x, t) dV,$$

我们可以得出以下结论:

Lemma 2.7.

$$\frac{d}{dt} \langle \Phi, \Psi \rangle = \left\langle \frac{\partial \Phi}{\partial t}, \Psi \right\rangle + \left\langle \Phi, \frac{\partial \Psi}{\partial t} \right\rangle.$$

Proof.

$$\begin{aligned} \frac{d}{dt} \langle \Phi, \Psi \rangle &= \frac{d}{dt} \left(\int_{\mathbb{R}^n} \bar{\Phi} \Psi dV \right) \\ &= \int_{\mathbb{R}^n} \frac{\partial}{\partial t} (\bar{\Phi} \Psi) dV \\ &= \int_{\mathbb{R}^n} \frac{\partial \bar{\Phi}}{\partial t} \Psi + \bar{\Phi} \frac{\partial \Psi}{\partial t} dV \\ &= \left\langle \frac{\partial \Phi}{\partial t}, \Psi \right\rangle + \left\langle \Phi, \frac{\partial \Psi}{\partial t} \right\rangle. \end{aligned}$$

□

而波函数 $\Psi(x, t)$ 对应厄米算子 A 的期望值为 $\langle A \rangle = \langle \Psi(x, t), A\Psi(x, t) \rangle$ ，我们又可以得出以下结论：

Corollary 2.7.1. 如果 A 是位置算子 X_i 或动量算子 P_i ，其中 $i \in \{1, \dots, n\}$ ，那么

$$\frac{d}{dt} \langle A \rangle = \left\langle \frac{1}{i\hbar} [A, H] \right\rangle.$$

Proof. 根据Lemma 2.7可得

$$\frac{d}{dt} \langle A \rangle = \frac{d}{dt} \langle \Psi, A\Psi \rangle = \left\langle \frac{\partial \Psi}{\partial t}, A\Psi \right\rangle + \left\langle \Psi, \frac{\partial}{\partial t} (A\Psi) \right\rangle,$$

而

$$\begin{aligned} \frac{\partial}{\partial t} (x^i \Psi) &= x^i \frac{\partial \Psi}{\partial t}, \\ \frac{\partial^2 \Psi}{\partial x^i \partial t} &= \frac{\partial^2 \Psi}{\partial t \partial x^i}. \end{aligned}$$

因此，无论 A 是位置算子 X_i 还是动量算子 P_i ，都具有

$$\frac{\partial}{\partial t} (A\Psi) = A \frac{\partial \Psi}{\partial t},$$

即

$$\frac{d}{dt} \langle A \rangle = \left\langle \frac{\partial \Psi}{\partial t}, A\Psi \right\rangle + \left\langle \Psi, A \frac{\partial \Psi}{\partial t} \right\rangle.$$

根据(2.20)可得

$$\frac{\partial \Psi}{\partial t} = \frac{1}{i\hbar} H \Psi,$$

因此，利用 H 是厄米算子的性质可得

$$\begin{aligned} \frac{d}{dt} \langle A \rangle &= \left\langle \frac{1}{i\hbar} H \Psi, A \Psi \right\rangle + \left\langle \Psi, \frac{1}{i\hbar} A H \Psi \right\rangle \\ &= -\frac{1}{i\hbar} \langle H \Psi, A \Psi \rangle + \frac{1}{i\hbar} \langle \Psi, A H \Psi \rangle \\ &= -\frac{1}{i\hbar} \langle \Psi, H A \Psi \rangle + \frac{1}{i\hbar} \langle \Psi, A H \Psi \rangle \\ &= \frac{1}{i\hbar} \langle \Psi, [A, H] \Psi \rangle \\ &= \left\langle \frac{1}{i\hbar} [A, H] \right\rangle. \end{aligned}$$

□

因此，我们可以将Lemma 2.6改写为

$$\begin{aligned} \frac{d}{dt} \langle X_i \rangle &= \frac{\langle P_i \rangle}{m}, \\ \frac{d}{dt} \langle P_i \rangle &= -\frac{\partial V}{\partial x^i}. \end{aligned}$$

该表达式还是不够美观，我们需要进一步处理。可以发现

$$-\frac{\hbar^2}{2m} \Delta \Psi + V \Psi = \frac{1}{2m} \sum_{i=1}^n P_i^2 \Psi(x^1, \dots, x^n, t) + V(X_1, \dots, X_n) \Psi(x^1, \dots, x^n),$$

(势能函数 $V(x)$ 是 x^1, \dots, x^n 的表达式， $V(X_1, \dots, X_n)$ 将所有 x^i 替换为 X_i ，例如 $V(x^1, x^2) = \frac{1}{2}m\omega_1^2(x^1)^2 + \frac{1}{2}m\omega_2^2(x^2)^2$ 对应 $V(X_1, X_2) = \frac{1}{2}m\omega_1^2 X_1^2 + \frac{1}{2}m\omega_2^2 X_2^2$ 。)因此哈密顿算子又可以记为

$$H(X, P) = \frac{1}{2m} \sum_{i=1}^n P_i^2 + V(X_1, \dots, X_n),$$

其中 $X = (X_1, \dots, X_n)$, $P = (P_1, \dots, P_n)$ 。因此

$$\begin{aligned} \frac{\partial H}{\partial P_i}(X, P) &= \frac{P_i}{m}, \\ \frac{\partial H}{\partial X_i}(X, P) &= \frac{\partial V}{\partial X_i}, \end{aligned}$$

就可以引申出

$$\begin{aligned}\frac{d}{dt}\langle X_i \rangle &= \left\langle \frac{\partial H}{\partial P_i}(X, P) \right\rangle, \\ \frac{d}{dt}\langle P_i \rangle &= - \left\langle \frac{\partial H}{\partial X_i}(X, P) \right\rangle.\end{aligned}$$

可以发现 $\left\langle \frac{\partial H}{\partial P_i}(X, P) \right\rangle, \left\langle \frac{\partial H}{\partial X_i}(X, P) \right\rangle$ 均为 $\langle X_1 \rangle, \dots, \langle X_n \rangle, \langle P_1 \rangle, \dots, \langle P_n \rangle$ 的函数。令 $\alpha^i(t) = \langle X_i \rangle, \beta^i(t) = \langle P_i \rangle$, 那么可以进一步化为

$$\begin{aligned}\dot{\alpha}^i(t) &= \frac{\partial H}{\partial \beta^i}(\alpha(t), \beta(t)), \\ \dot{\beta}^i(t) &= - \frac{\partial H}{\partial \alpha^i}(\alpha(t), \beta(t)),\end{aligned}\tag{2.21}$$

其中 $\alpha(t) = (\alpha^1(t), \dots, \alpha^n(t)), \beta(t) = (\beta^1(t), \dots, \beta^n(t))$ 。我们将方程组(2.21)称为**哈密顿方程**。在深入讨论这一概念之前, 我们需要一些数学工具。

2.4 向量场*

⁸ 在本章节中, 我们将向量空间 \mathbb{R}^n 上的一组基记为 $\partial/\partial x^1, \dots, \partial/\partial x^n$, 将任意向量 $X = (X^1, \dots, X^n)$ 都表示为

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}.$$

虽然这一表示看似非常奇怪, 但是它能够为以后的分析带来方便, 我们后面会对其进行解释。

利用第141页对“连续”的定义, 我们定义连续映射 $\pi: M \rightarrow N$ 的**截面**为连续映射 $\sigma: N \rightarrow M$, 其中 $\pi \circ \sigma = \text{Id}_N$, 换句话说, 对于任意 $x \in N$, 都具有 $\pi(\sigma(x)) = x$ 。例如, 对于 $n > m$, 定义映射 $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 它将删除最后 $n - m$ 个坐标, 但是前面的坐标不发生改变。那么任意截面 $\sigma: \mathbb{R}^m \rightarrow \mathbb{R}^n$ 都可以被表示为

$$\sigma(x^1, \dots, x^m) = (x^1, \dots, x^m, c^{m+1}, \dots, c^n),$$

其中 $c^{m+1}, \dots, c^n \in \mathbb{R}$ 是任意常数, 很显然 $\pi(\sigma(x^1, \dots, x^m)) = (x^1, \dots, x^m)$ 。我们将映射 $\pi: M \rightarrow N$ 所有光滑截面(截面不仅是连续映射, 还是光滑映射) $\sigma: N \rightarrow M$ 的集合记为 $\Gamma(M)$ 。

对于 $U \subseteq \mathbb{R}^n$, 定义 U 的**切丛**为

$$TU = U \times \mathbb{R}^n = \{(p, v) : p \in U, v \in \mathbb{R}^n\},$$

⁸符号*表示这一章节可以跳过, 建议跳转到第55页的章节2.5。

记映射 $\pi : TU \rightarrow U$ 为

$$\pi(p, v) = p,$$

因此, 映射 π 的任意截面 $\sigma : U \rightarrow TU$ 都可以被表示为

$$\sigma(p) = (p, w),$$

其中 $w \in \mathbb{R}^n$ 是任意向量。这相当于将 U 上的任意一点 p 都赋予一个向量, 这与第39页定义的“向量场”是等价的。因此, U 上的向量场的另一定义为映射 $\pi : TU \rightarrow U$ 的截面。

任意向量场 X 在 p 点的取值可以被表示为

$$X(p) = \sum_{i=1}^n X^i(p) \frac{\partial}{\partial x^i},$$

其中 $p \in U$; X^i, ω_i 将 U 上的任意一点 p 映射到一个实数。我们称向量场是光滑的, 当且仅当对应的所有 $X^i : U \rightarrow \mathbb{R}$ 都是光滑函数, 将所有光滑向量场的集合记为 $\mathcal{X}(U)$ 。

将一个光滑函数 $f \in C^\infty(U)$ 与一个光滑向量场 $X \in \mathcal{X}(U)$ 相乘, 得到的结果还是光滑向量场, 其取值为

$$(fX)(p) = f(p)X(p) = \sum_{i=1}^n f(p)X^i(p) \frac{\partial}{\partial x^i}.$$

而 $C^\infty(U)$ 是一个交换环: 给定2个光滑函数 $f, g \in C^\infty(U)$, $-f, f+g, fg$ 还是光滑函数, 并且加法与乘法均有交换律、结合律。因此 $\mathcal{X}(U)$ 是交换环 $C^\infty(U)$ 上的模。

向量场 X 还可以作用于光滑函数 f , 其作用结果是连续函数(因为向量场不一定光滑):

$$Xf = \sum_{i=1}^n \left(X^i \frac{\partial}{\partial x^i} \right) f = \sum_{i=1}^n X^i \frac{\partial f}{\partial x^i}, \quad (2.22)$$

其在任意 p 点的取值, 反映了光滑函数 f 在 p 处的向量 $X(p)$ 的方向导数, 即

$$(Xf)(p) = \lim_{h \rightarrow 0} \frac{f(p + hX(p)) - f(p)}{h},$$

利用泰勒定理不难验证。因此, 在这个情况下, 将基表示为 $\partial/\partial x^i$ 更具有直观性。

Theorem 2.8. 向量场 X 是光滑的, 当且仅当, 对于任意光滑函数 f , Xf 也是光滑函数。

Proof. 充分条件: 假设 X 是光滑向量场, 那么任意 X^i 都是光滑函数。而 f 是光滑函数, 因此 $\frac{\partial f}{\partial x^i}$ 是光滑函数。再根据有限个光滑函数的加法与乘法依旧是光滑函数, 可得 $Xf = \sum_{i=1}^n X^i \frac{\partial f}{\partial x^i}$ 是光滑函数。

必要条件: 因为 $f(x^1, \dots, x^n) = x^i$ 是光滑函数, 在这个情况下, 计算 Xf :

$$Xf = \sum_{j=1}^n X^j \frac{\partial f}{\partial x^j} = \sum_{j=1}^n X^j \frac{\partial x^i}{\partial x^j} = X^i,$$

那么 X^i 是光滑函数。而这一命题对于任意 i 都成立, 这表明 X 是光滑向量场。 \square

2.4.1 积分曲线*

⁹ 空间中的电场与磁场均为 \mathbb{R}^3 上的光滑向量场。在电磁学中，我们经常使用数学模型——电场线(或磁感线)，以“可视化”电磁场：对于任意一根电场线(或磁感线)，其每一点上的切向量都反映了该点电场强度(或磁感应强度)的方向。

同样的，对于 \mathbb{R}^n 上的任意光滑向量场，我们可以利用“电场线”来对其进行可视化，严格的数学名词是**积分曲线**，正是我们马上就要定义的：对于向量场 $X : \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$X(p) = \sum_{i=1}^n X^i(p) \frac{\partial}{\partial x^i},$$

其积分曲线是一个光滑函数 $\gamma : J \rightarrow \mathbb{R}^n$ ，将 $\gamma(t)$ 写为分量形式 $\gamma(t) = (\gamma^1(t), \dots, \gamma^n(t))$ ，它满足

$$\dot{\gamma}^i(t) = X^i(\gamma^1(t), \dots, \gamma^n(t)), \quad i = 1, \dots, n, \quad (2.23)$$

其中 $J \subseteq \mathbb{R}$ 是一个连通的开区间。如果 $0 \in J$ ，那么将 $\gamma(0)$ 称为**积分曲线 γ 的起点**。我们将给出一些实例：

(A) 令 $V = \frac{\partial}{\partial x}$ 是 \mathbb{R}^2 上的向量场。记其积分曲线为 $\gamma(t) = (\gamma^1(t), \gamma^2(t))$ ，因此

$$\begin{aligned} \dot{\gamma}^1(t) &= 1, \\ \dot{\gamma}^2(t) &= 0, \end{aligned}$$

这表明

$$\gamma(t) = (a + t, b),$$

其中 a, b 是常数，并且 $\gamma(0) = (a, b)$ 。因此，对于 \mathbb{R}^2 上的任意一点，都有以它为起点的一条积分曲线，并且这样的积分曲线是唯一的。

(B) 令 $V = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}$ 是 \mathbb{R}^2 上的向量场。记其积分曲线为 $\gamma(t) = (\gamma^1(t), \gamma^2(t))$ ，因此

$$\begin{aligned} \dot{\gamma}^1(t) &= -\gamma^2(t), \\ \dot{\gamma}^2(t) &= \gamma^1(t), \end{aligned}$$

该方程组的解为

$$\begin{aligned} \gamma^1(t) &= a \cos t - b \sin t, \\ \gamma^2(t) &= a \sin t + b \cos t, \end{aligned}$$

其中 a, b 是常数，并且 $\gamma(0) = (a, b)$ 。因此，对于 \mathbb{R}^2 上的任意一点，都有以它为起点的一条积分曲线，并且这样的积分曲线是唯一的；而且，如果起点为 $a = b = 0$ ，那么对应的积分曲线为 $\gamma(t) \equiv (0, 0)$ ，相当于固定不动！

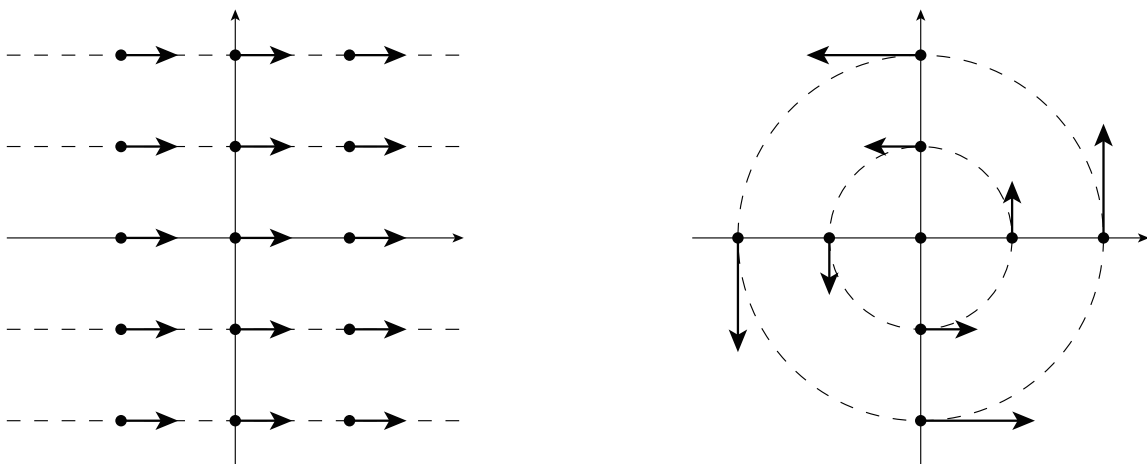


Figure 2.2: 左图为向量场 $V = \frac{\partial}{\partial x}$ 与其积分曲线，右图为向量场 $V = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$ 与其积分曲线。

Figure 2.2绘制了 \mathbb{R}^2 上的向量场与对应的积分曲线，其中向量场通过“箭头”表示，积分曲线通过“虚线”表示。

Theorem 2.9. $Xf \equiv 0$ 当且仅当光滑函数 $f \in C^\infty(\mathbb{R}^n)$ 在向量场 X 的任意积分曲线 $\gamma(t)$ 上都取到常数——换句话说， $f \circ \gamma$ 是常函数。

Proof. 根据(2.23)可得

$$\begin{aligned} \frac{d}{dt} f \circ \gamma(t) &= \sum_{i=1}^n \dot{\gamma}^i(t) \frac{\partial f}{\partial x^i} (\gamma^1(t), \dots, \gamma^n(t)) \\ &= \sum_{i=1}^n X^i (\gamma^1(t), \dots, \gamma^n(t)) \frac{\partial f}{\partial x^i} (\gamma^1(t), \dots, \gamma^n(t)) \\ &= (Xf) (\gamma^1(t), \dots, \gamma^n(t)) = 0. \end{aligned}$$

□

2.4.2 流*

¹⁰ 定义满足以下条件的光滑映射 $\theta : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$:

$$\begin{aligned} \theta(0, p) &= p, & p &\in \mathbb{R}^n, \\ \theta(t, \theta(s, p)) &= \theta(t+s, p), & s, t &\in \mathbb{R}, \end{aligned}$$

⁹ 符号*表示这一章节可以跳过，建议跳转到第55页的章节2.5。

¹⁰ 符号*表示这一章节可以跳过，建议跳转到第55页的章节2.5。

我们将这一连续映射称为**流**。它可以生成 \mathbb{R}^n 上的光滑向量场

$$X(p) = \left. \frac{\partial}{\partial t} \right|_{t=0} \theta(t, p),$$

(因为 $\theta(t, p) \in \mathbb{R}^n$ ，所以对时间求微分后得到的结果依然是 n 维向量。)我们将这一向量场称为 **θ 的无穷小生成元**。它具有以下性质：

Theorem 2.10. 对于任意 $p \in \mathbb{R}^n$ ，固定 p 不动，那么光滑函数 $\gamma(t) = \theta(t, p)$ 是 X 的积分曲线(当然，以 p 为起点)。其中 X 是 θ 的无穷小生成元。

Proof. 根据积分曲线的定义，我们需要证明

$$\dot{\gamma}^i(t) = X^i(\gamma^1(t), \dots, \gamma^n(t)), \quad i = 1, \dots, n.$$

将 $\theta(t, p)$ 写成分量形式 $(\theta^1(t, p), \dots, \theta^n(t, p))$ ， θ 的无穷小生成元 X 的分量满足

$$\left. \frac{\partial}{\partial s} \right|_{s=0} \theta^i(s, p) = X^i(p), \quad i = 1, \dots, n.$$

因此，这等价于证明

$$\dot{\gamma}^i(t) = \left. \frac{\partial}{\partial s} \right|_{s=0} \theta^i(s, \gamma^1(t), \dots, \gamma^n(t)), \quad i = 1, \dots, n.$$

根据 $\gamma(t)$ 的定义 $\gamma(t) = \theta(t, p)$ ，这等价于证明

$$\frac{\partial}{\partial t} \theta(t, p) = \left. \frac{\partial}{\partial s} \right|_{s=0} \theta(s, \theta(t, p)),$$

而 $\theta(t, \theta(s, p)) = \theta(t + s, p)$ ，因此这等价于证明

$$\frac{\partial}{\partial t} \theta(t, p) = \left. \frac{\partial}{\partial s} \right|_{s=0} \theta(s + t, p),$$

根据链式法则，这是显然的。 □

但是，对于任意光滑向量场 X ，是否存在流 θ ，使得 X 是 θ 的无穷小生成元呢？答案是肯定的，我们将对应的 θ 称为**向量场 X 的流**。不过，在这里就不做证明了，因为这是动力系统的一个结论：

Theorem 2.11. 令 X 为 \mathbb{R}^n 上的光滑向量场。那么，存在唯一的流 $\theta : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ 使得它的无穷小生成元为 X 。

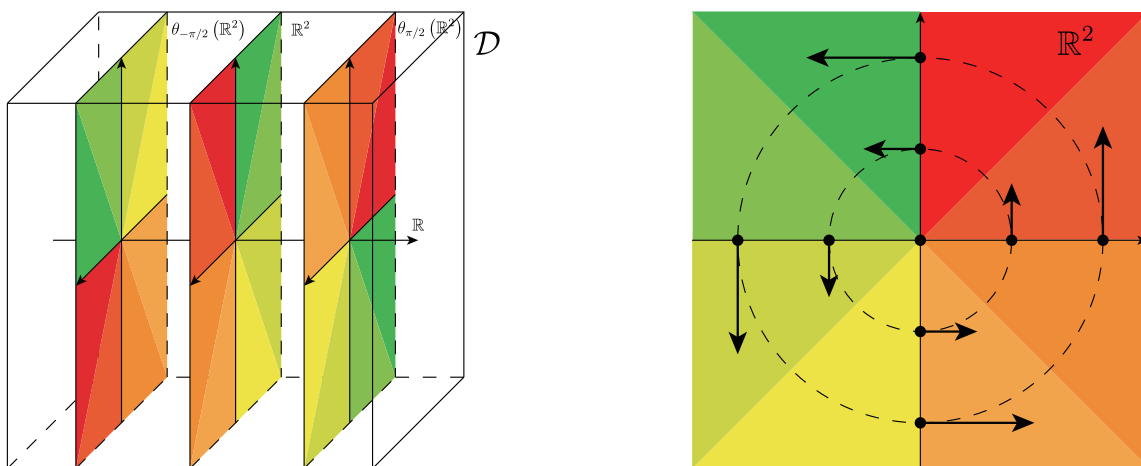


Figure 2.3: 流的几何意义

通常情况下，我们将 $\theta(t, p)$ 简记为 $\theta_t(p)$ 。根据Theorem 2.10可得，对于固定的 $t \in \mathbb{R}$ ， $\theta_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 满足

$$\theta_t(p) = \gamma(t) = (\gamma^1(t), \dots, \gamma^n(t)),$$

其中 $\gamma(t)$ 是以 p 为起点的积分曲线。

因此，对于固定的 $t \in \mathbb{R}$ ，我们可以将 $\theta_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 视为——任意 $p \in \mathbb{R}^n$ 沿着以 p 为起点的积分曲线“随波逐流”，经过时间 t 后将到达 $\theta_t(p)$ 。以第52页的(B)为例，对于任意 $p = (a, b) \in \mathbb{R}^2$ ， $\theta_t(a, b) = (a \cos t - b \sin t, a \sin t + b \cos t)$ 。因此，当 $t = \pi/2$ 时， $\theta_{\pi/2}(a, b) = (-b, a)$ ；当 $t = -\pi/2$ 时， $\theta_{-\pi/2}(a, b) = (b, -a)$ 。这表明 $\theta_{\pi/2}$ 将 \mathbb{R}^2 逆时针旋转90度， $\theta_{-\pi/2}$ 将 \mathbb{R}^2 顺时针旋转90度，如Figure 2.3所示。

2.5 对偶空间

对于 \mathbb{R} 上的有限维向量空间 V ，定义对偶空间 V^* ，它是所有线性映射 $\omega : V \rightarrow \mathbb{R}$ 的集合。因此，对于任意 $v, w \in V, \alpha, \beta \in \mathbb{R}$ ，

$$\omega(\alpha v + \beta w) = \alpha \omega(v) + \beta \omega(w). \quad (2.24)$$

并且，对于任意 $\omega, \nu \in V^*, \alpha, \beta \in \mathbb{R}$ ， $\alpha \omega + \beta \nu \in V^*$ ，定义

$$(\alpha \omega + \beta \nu)(v) = \alpha \omega(v) + \beta \nu(v). \quad (2.25)$$

Theorem 2.12. V^* 是 \mathbb{R} 上的向量空间。记有限维向量空间 V 的一组基为 E_1, \dots, E_n ，那么 $\epsilon^1, \dots, \epsilon^n \in V^*$ 是向量空间 V^* 上的一组基，其中

$$\epsilon^i(E_j) = \delta_j^i = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

因此， V^* 是有限维向量空间，并且 $\dim V^* = \dim V$ 。

Proof. 很显然 V^* 是向量空间。对于任意 $\omega \in V^*$, 记

$$\beta = \omega(E_1)\epsilon^1 + \cdots + \omega(E_n)\epsilon^n,$$

可以得出, $\beta(E_i) = \omega(E_i)$ 对于任意 i 都成立。因此, $\beta = \omega$, 这证明了 $\epsilon^1, \dots, \epsilon^n$ 生成向量空间 V^* 。

如果存在不完全为零的常数 c_1, \dots, c_n , 使得

$$\gamma = c_1\epsilon^1 + \cdots + c_n\epsilon^n = 0,$$

那么 $\gamma(E_i) = 0$ 对于任意 i 都成立, 而 $\gamma(E_i) = c_i$, 这表明 $c_i = 0$ 对于任意 i 都成立, 得到矛盾。因此, $\epsilon^1, \dots, \epsilon^n$ 是线性无关的。□

据此, 我们将向量空间 V 的任意元素称为**向量**, 其对偶空间 V^* 的任意元素称为**余向量**。

根据Theorem 2.12, \mathbb{R}^n 对偶空间的维数为 n , 记对应的基为 dx^1, \dots, dx^n , 因此

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_j^i. \quad (2.26)$$

务必注意, dx^1, \dots, dx^n 与 $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ 所处的向量空间是不同的, 正如我们无法将列向量与行向量线性组合一样。

对于 \mathbb{R}^n 的任意向量

$$v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}, \quad (2.27)$$

与余向量

$$\omega = \sum_{i=1}^n \omega_i dx^i, \quad (2.28)$$

其中 $v_i, \omega_i \in \mathbb{R}$ 。根据(2.24)与(2.25)可得

$$\begin{aligned} \omega(v) &= \omega \left(\sum_{i=1}^n v^i \frac{\partial}{\partial x^i} \right) = \sum_{i=1}^n v^i \omega \left(\frac{\partial}{\partial x^i} \right) \\ &= \sum_{i=1}^n v^i \left(\sum_{j=1}^n \omega_j dx^j \right) \left(\frac{\partial}{\partial x^i} \right) \\ &= \sum_{i,j=1}^n v^i \omega_j dx^j \left(\frac{\partial}{\partial x^i} \right) = \sum_{i,j=1}^n v^i \omega_j \delta_i^j \\ &= \sum_{i=1}^n v^i \omega_i. \end{aligned}$$

Theorem 2.12表明, 对于 \mathbb{R} 上的有限维向量空间 V , 其对偶空间 V^* 同样也是 \mathbb{R} 上的向量空间; 而任意向量空间都具有对偶空间, 那么理论上, 我们可以定义“对偶空间的对偶空间”, 我们暂时记它为 $V^{**} = (V^*)^*$ 。因此, V^{**} 的任意元素 $\chi \in V^{**}$ 将 V^* 的任意元素 $\omega \in V^*$ 映射到一个实数 $\chi(\omega) \in \mathbb{R}$ 。但是, 这一定义是没有必要的, 以下的定理会对其进行论证:

Theorem 2.13. 对于 \mathbb{R} 上的有限维向量空间 V , 定义映射 $\xi: V \rightarrow V^{**}$ 为

$$\xi(v)(\omega) = \omega(v),$$

其中 $\omega \in V^*$ 。那么 ξ 不仅是线性映射, 还是同构映射。

Proof. 对于 $\alpha, \beta \in \mathbb{R}; v, w \in V; \omega \in V^*$, 根据(2.24)可得

$$\xi(\alpha v + \beta w)(\omega) = \omega(\alpha v + \beta w) = \alpha\omega(v) + \beta\omega(w),$$

而根据(2.25)可得

$$(\alpha\xi(v) + \beta\xi(w))(\omega) = \alpha\xi(v)(\omega) + \beta\xi(w)(\omega) = \alpha\omega(v) + \beta\omega(w),$$

因此 $\xi(\alpha v + \beta w)(\omega) = (\alpha\xi(v) + \beta\xi(w))(\omega)$, 根据 ω 的任意性可得

$$\xi(\alpha v + \beta w) = \alpha\xi(v) + \beta\xi(w),$$

这证明了 ξ 是线性映射。

根据Theorem 2.12可得, $\dim V = \dim V^* = \dim V^{**}$, 因此我们只需证明 ξ 是单射即可: 换句话说, 对于任意 $v \neq 0 \in V$, 都具有 $\xi(v) \neq 0 \in V^{**}$ 。令 $E_1 = v, E_2, \dots, E_n$ 为 V 的一组基, $\epsilon^1, \dots, \epsilon^n$ 为对偶空间的基, 如Theorem 2.12所示。那么

$$\xi(v)(\epsilon^1) = \epsilon^1(v) = \epsilon^1(E_1) = 1 \neq 0.$$

□

该定理定义的线性映射 ξ 与基的选取是无关的, 我们可以直接省略 ξ , 将 $\xi(v)(\omega)$ 直接简记为

$$\xi(v)(\omega) = \langle v, \omega \rangle = \omega(v),$$

我们也可以将 $\omega(v)$ 记为 $\langle \omega, v \rangle$, 相当于 $\langle \cdot, \cdot \rangle$ 将向量与余向量进行“配对”, 但是可以交换顺序, 即 $\langle \omega, v \rangle = \langle v, \omega \rangle$ 。因此

$$\langle \epsilon^i, E_j \rangle = \langle E_j, \epsilon^i \rangle = \delta_j^i.$$

很显然, $\langle \cdot, \cdot \rangle$ 是双线性的, 即对于任意 $\alpha, \beta \in \mathbb{R}$ 都具有

$$\begin{aligned} \langle \alpha X + \beta Y, Z \rangle &= \alpha \langle X, Z \rangle + \beta \langle Y, Z \rangle, \\ \langle X, \alpha Y + \beta Z \rangle &= \alpha \langle X, Y \rangle + \beta \langle X, Z \rangle, \end{aligned}$$

它与“配对”的顺序无关。

2.6 张量

令 V_1, \dots, V_k 是实数域 \mathbb{R} 上的向量空间, 定义 $V_1 \otimes \dots \otimes V_k$ 为向量空间 V_1, \dots, V_k 的张量积, 它同样也是 \mathbb{R} 上的向量空间, 通过任意向量 $v_1 \in V_1, \dots, v_k \in V_k$ 的张量积 $v_1 \otimes \dots \otimes v_k$ 生成。向量的张量积具有以下性质: 对于任意 $a, b \in \mathbb{R}; i \in [1, k]$,

$$v_1 \otimes \dots \otimes (av_i + bv'_i) \otimes \dots \otimes v_k = a(v_1 \otimes \dots \otimes v_i \otimes \dots \otimes v_k) + b(v_1 \otimes \dots \otimes v'_i \otimes \dots \otimes v_k).$$

Theorem 2.14.

令 V_1, \dots, V_k 是实数域 \mathbb{R} 上的有限维向量空间。定义多线性映射 $T: V_1 \times \dots \times V_k \rightarrow W$ 为

$$T(v_1, \dots, av_i + bv'_i, \dots, v_k) = aT(v_1, \dots, v_i, \dots, v_k) + bT(v_1, \dots, v'_i, \dots, v_k),$$

那么存在唯一的线性映射 $V_1 \otimes \dots \otimes V_k \rightarrow W$, 记为 \tilde{T} , 其形式为

$$\tilde{T}(v_1 \otimes \dots \otimes v_k) = T(v_1, \dots, v_k). \quad (2.29)$$

如Figure 2.4所示, 其中 $\pi(v_1, \dots, v_k) = v_1 \otimes \dots \otimes v_k$ 。

这一定理我们就不做证明了, 但是它至关重要, 它表明了张量积与多线性映射的关系!

Corollary 2.14.1. 令 V_1, \dots, V_k 是实数域 \mathbb{R} 上的有限维向量空间, v_1, \dots, v_k 分别为 V_1, \dots, V_k 上的向量, 那么

$$v_1 \otimes \dots \otimes 0 \otimes \dots \otimes v_k = 0 \otimes \dots \otimes 0.$$

因此, 只要 v_1, \dots, v_k 有一个为0, 那么 $v_1 \otimes \dots \otimes v_k$ 就是 $V_1 \otimes \dots \otimes V_k$ 上的零元。但是, 如果 v_1, \dots, v_k 均是非零的, 那么

$$v_1 \otimes \dots \otimes v_k \neq 0 \otimes \dots \otimes 0.$$

相当于, $v_1 \otimes \dots \otimes v_k = 0$ 当且仅当 v_1, \dots, v_k 存在取值为0的元素。

Proof. 因为

$$\begin{aligned} v_1 \otimes \dots \otimes 0 \otimes \dots \otimes v_k &= v_1 \otimes \dots \otimes (0v_i) \otimes \dots \otimes v_k = 0(v_1 \otimes \dots \otimes v_k), \\ 0 \otimes \dots \otimes 0 &= (0v_1) \otimes \dots \otimes (0v_k) = (0 \dots 0)(v_1 \otimes \dots \otimes v_k) \\ &= 0(v_1 \otimes \dots \otimes v_k), \end{aligned}$$

所以 $v_1 \otimes \dots \otimes 0 \otimes \dots \otimes v_k = 0 \otimes \dots \otimes 0$ 。

对于后者表述, 我们可以利用反证法, 假设存在全部非零的 v_1^0, \dots, v_k^0 满足 $v_1^0 \otimes \dots \otimes v_k^0 = 0$ 。此时, 定义线性映射 $F_i: V_i \rightarrow \mathbb{R}$, 使得 $F_i(v_i^0) = 1$ 。再定义映射 $T: V_1 \times \dots \times V_k \rightarrow \mathbb{R}$ 为

$$T(v_1, \dots, v_k) = F_1(v_1) \dots F_k(v_k).$$

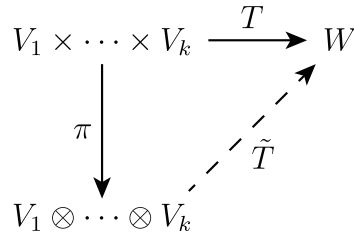


Figure 2.4: 乘积空间的多线性映射对应张量积空间的线性映射

显然它是多线性映射。因此，存在唯一的线性映射 $\tilde{T}: V_1 \otimes \cdots \otimes V_k \rightarrow W$ 为

$$\tilde{T}(v_1 \otimes \cdots \otimes v_k) = F_1(v_1) \cdots F_k(v_k).$$

因为 $v_1^0 \otimes \cdots \otimes v_k^0 = 0$ ，那么这样一来，根据 \tilde{T} 是线性的性质， $\tilde{T}(v_1^0 \otimes \cdots \otimes v_k^0) = 0$ 。而 $F_1(v_1^0) \cdots F_k(v_k^0) = 1$ ，我们就给出了矛盾。 \square

Corollary 2.14.2. 令 $V_1, \dots, V_k; W_1, \dots, W_k$ 是实数域 \mathbb{R} 上的有限维向量空间，所有 $T_i: V_i \rightarrow W_i$ 都是线性映射。那么存在唯一的线性映射 (记为 $T_1 \otimes \cdots \otimes T_k: V_1 \otimes \cdots \otimes V_k \rightarrow W_1 \otimes \cdots \otimes W_k$)，其形式为

$$T_1 \otimes \cdots \otimes T_k(v_1 \otimes \cdots \otimes v_k) = T_1(v_1) \otimes \cdots \otimes T_k(v_k).$$

Proof. 定义映射 $T: V_1 \times \cdots \times V_k \rightarrow W_1 \otimes \cdots \otimes W_k$,

$$T(v_1, \dots, v_k) = T_1(v_1) \otimes \cdots \otimes T_k(v_k).$$

它是多线性映射，因为

$$\begin{aligned} T(v_1, \dots, av_i + bv'_i, \dots, v_k) &= T_1(v_1) \otimes \cdots \otimes T_i(av_i + bv'_i) \otimes \cdots \otimes T_k(v_k) \\ &= T_1(v_1) \otimes \cdots \otimes (aT_i(v_i) + bT_i(v'_i)) \otimes \cdots \otimes T_k(v_k) \\ &= a(T_1(v_1) \otimes \cdots \otimes T_i(v_i) \otimes \cdots \otimes T_k(v_k)) \\ &\quad + b(T_1(v_1) \otimes \cdots \otimes T_i(v'_i) \otimes \cdots \otimes T_k(v_k)) \\ &= aT(v_1, \dots, v_i, \dots, v_k) + bT(v_1, \dots, v'_i, \dots, v_k). \end{aligned}$$

因此存在唯一线性映射 $\tilde{T}: V_1 \otimes \cdots \otimes V_k \rightarrow W_1 \otimes \cdots \otimes W_k$ 使得

$$\tilde{T}(v_1 \otimes \cdots \otimes v_k) = T(v_1, \dots, v_k) = T_1(v_1) \otimes \cdots \otimes T_k(v_k).$$

\square

通过以上的定理与引理，我们能够得出以下结论，该结论能够大致描绘出张量积空间的结构：

Theorem 2.15. 令 V_1, \dots, V_k 是实数域 \mathbb{R} 上的有限维向量空间, 其维数分别为 n_1, \dots, n_k 。记 $\{E_{1i_1}\}, \dots, \{E_{ki_k}\} : i_1 \in [1, n_1], \dots, i_k \in [1, n_k]$ 分别为 V_1, \dots, V_k 的一组基, 那么

$$\mathcal{A} = \{E_{1i_1} \otimes \dots \otimes E_{ki_k} : i_1 \in [1, n_1], \dots, i_k \in [1, n_k]\}$$

为 $V_1 \otimes \dots \otimes V_k$ 的一组基。因此, $\dim V_1 \otimes \dots \otimes V_k = n_1 \cdots n_k$,

$$V_1 \otimes \dots \otimes V_k = \left\{ \sum_{i_1, \dots, i_k} a^{i_1 \dots i_k} (E_{1i_1} \otimes \dots \otimes E_{ki_k}) : a^{i_1 \dots i_k} \in \mathbb{R} \right\}.$$

Proof. 因为任意向量都可以被表示为基的线性组合, 所以可以将任意 $v_1 \otimes \dots \otimes v_k$ 表示为 $E_{1i_1} \otimes \dots \otimes E_{ki_k}$ 的线性组合, 这表明 \mathcal{A} 生成 $V_1 \otimes \dots \otimes V_k$ 。但是还需要证明 \mathcal{A} 是线性无关的。假设一些形如

$$a^{i_1 \dots i_k} E_{1i_1} \otimes \dots \otimes E_{ki_k}$$

的线性组合为 0, 我们需要证明任意 $a^{i_1 \dots i_k} = 0$ 。

根据 Theorem 2.12, 令对偶空间为 V_1^*, \dots, V_k^* , 记 $\{\epsilon^{1i_1}\}, \dots, \{\epsilon^{ki_k}\} : i_1 \in [1, n_1], \dots, i_k \in [1, n_k]$ 分别为 V_1^*, \dots, V_k^* 的基。对于任意 $m_1 \in [1, n_1], \dots, m_k \in [1, n_k]$, 定义多线性映射 $\tau^{m_1 \dots m_k} : V_1 \times \dots \times V_k \rightarrow \mathbb{R}$ 为

$$\tau^{m_1 \dots m_k}(v_1, \dots, v_k) = \epsilon^{1m_1}(v_1) \cdots \epsilon^{km_k}(v_k).$$

因此, 存在唯一的线性映射 $\tilde{\tau}^{m_1 \dots m_k} : V_1 \otimes \dots \otimes V_k \rightarrow \mathbb{R}$, 使得

$$\tilde{\tau}^{m_1 \dots m_k}(v_1 \otimes \dots \otimes v_k) = \epsilon^{1m_1}(v_1) \cdots \epsilon^{km_k}(v_k).$$

将 $\tilde{\tau}^{m_1 \dots m_k}$ 作用于线性组合 $a^{i_1 \dots i_k} E_{1i_1} \otimes \dots \otimes E_{ki_k}$ 可得 $a^{m_1 \dots m_k} = 0$ 。根据 m_1, \dots, m_k 的任意性, 我们得出了 \mathcal{A} 是线性无关的结论。□

Theorem 2.16. 令 V, W, X 是实数域 \mathbb{R} 上的有限维向量空间。它们具有以下性质:

(A) 对于 $v \in V, w \in W$, $v \otimes w \mapsto w \otimes v$ 是同构映射。因此

$$V \otimes W \cong W \otimes V.$$

(B) 将 \mathbb{R} 视为 1 维向量空间。对于 $a \in \mathbb{R}, v \in V$, $a \otimes v \mapsto av$ 是同构映射。因此

$$\mathbb{R} \otimes V \cong V.$$

(C) 对于 $v \in V, w \in W, x \in X$, $(v \otimes w) \otimes x \mapsto v \otimes (w \otimes x), (v \otimes w) \otimes x \mapsto v \otimes w \otimes x$ 都是同构映射。因此

$$(V \otimes W) \otimes X \cong V \otimes (W \otimes X) \cong V \otimes W \otimes X.$$

Proof. 根据Theorem 2.15,

$$\begin{aligned}\dim V \otimes W &= \dim W \otimes V, \\ \dim \mathbb{R} \otimes V &= \dim V, \\ \dim(V \otimes W) \otimes X &= \dim V \otimes (W \otimes X) = \dim V \otimes W \otimes X.\end{aligned}$$

很显然这些映射全部都是线性映射, 并且是满射, 因此都是同构映射。 \square

Lemma 2.17. 对于维数分别为 n_1, \dots, n_k 的有限维向量空间 V_1, \dots, V_k , 记 $\{E_{1i_1}\}, \dots, \{E_{ki_k}\}$ 分别为 V_1, \dots, V_k 的基; 根据Theorem 2.12, 令对偶空间为 V_1^*, \dots, V_k^* , 记 $\{\epsilon^{1i_1}\}, \dots, \{\epsilon^{ki_k}\}$ 分别为 V_1^*, \dots, V_k^* 的基。其中 $i_1 \in [1, n_1], \dots, i_k \in [1, n_k]$ 。

令所有多线性映射 $V_1 \times \dots \times V_k \rightarrow \mathbb{R}$ 的集合为 $L(V_1, \dots, V_k; \mathbb{R})$, 那么

$$\mathcal{B} = \left\{ \epsilon^{1i_1} \otimes \dots \otimes \epsilon^{ki_k} : i_1 \in [1, n_1], \dots, i_k \in [1, n_k] \right\}$$

为 $L(V_1, \dots, V_k; \mathbb{R})$ 的一组基, 因此 $\dim L(V_1, \dots, V_k; \mathbb{R}) = n_1 \cdots n_k$ 。

Proof. 令 $F \in L(V_1, \dots, V_k; \mathbb{R})$ 是任意的多线性映射。对于任意 $i_1 \in [1, n_1], \dots, i_k \in [1, n_k]$, 定义

$$F_{i_1 \dots i_k} = F(E_{1i_1}, \dots, E_{ki_k}), \quad (2.30)$$

因此, 存在唯一的线性映射 $\tilde{F} : V_1 \otimes \dots \otimes V_k \rightarrow \mathbb{R}$ 使得

$$\tilde{F}(E_{1i_1} \otimes \dots \otimes E_{ki_k}) = F(E_{1i_1}, \dots, E_{ki_k}).$$

我们需要证明

$$\tilde{F} = \sum_{i_1 \dots i_k} F_{i_1 \dots i_k} \epsilon^{1i_1} \otimes \dots \otimes \epsilon^{ki_k}, \quad (2.31)$$

其中 $\sum_{i_1 \dots i_k}$ 代表 i_1, \dots, i_k 求和取遍 $i_1 \in [1, n_1], \dots, i_k \in [1, n_k]$ 的所有整数。对于任意向量 $v_1 \in V_1, \dots, v_k \in V_k$, 记 $v_1 = \sum_{i_1=1}^{n_1} v_1^{i_1} E_{1i_1}, \dots, v_k = \sum_{i_k=1}^{n_k} v_k^{i_k} E_{ki_k}$, 那么

$$\begin{aligned}\sum_{i_1 \dots i_k} F_{i_1 \dots i_k} \epsilon^{1i_1} \otimes \dots \otimes \epsilon^{ki_k} (v_1 \otimes \dots \otimes v_k) &= \sum_{i_1 \dots i_k} F_{i_1 \dots i_k} \epsilon^{1i_1}(v_1) \cdots \epsilon^{ki_k}(v_k) \\ &= \sum_{i_1 \dots i_k} F_{i_1 \dots i_k} v_1^{i_1} \cdots v_k^{i_k},\end{aligned}$$

根据 F 的多线性与(2.30)可得, 这一结果恰好与 $F(v_1, \dots, v_k)$ 相等, 这证明了(2.31)。因此, \mathcal{B} 生成向量空间 $L(V_1, \dots, V_k; \mathbb{R})$ 。

为了证明 \mathcal{B} 是线性无关的, 假设一些形如

$$F_{i_1 \dots i_k} \epsilon^{1i_1} \otimes \dots \otimes \epsilon^{ki_k}$$

的线性组合为0, 那么我们将这一线性组合作用于 $E_{1j_1} \otimes \dots \otimes E_{kj_k}$, 其中 $j_1 \in [1, n_1], \dots, j_k \in [1, n_k]$ 是任意整数。这样可以得到 $F_{j_1 \dots j_k} = 0$ 对于任意 j_1, \dots, j_k 均成立, 这证明了 \mathcal{B} 是线性无关的。 \square

Theorem 2.18. 对于有限维向量空间 V_1, \dots, V_k , 记其对偶空间为 V_1^*, \dots, V_k^* , 所有多线性映射 $V_1 \times \dots \times V_k \rightarrow \mathbb{R}$ 的集合为 $L(V_1, \dots, V_k; \mathbb{R})$, 那么

$$V_1^* \otimes \dots \otimes V_k^* \cong L(V_1, \dots, V_k; \mathbb{R}).$$

Proof. 定义映射 $\Phi: V_1^* \times \dots \times V_k^* \rightarrow L(V_1, \dots, V_k; \mathbb{R})$ 为

$$\Phi(\omega^1, \dots, \omega^k)(v_1, \dots, v_k) = \omega^1(v_1) \dots \omega^k(v_k),$$

其中 $\omega^1 \in V_1^*, \dots, \omega^k \in V_k^*; v_1 \in V_1, \dots, v_k \in V_k$ 。

很显然 Φ 是多线性映射, 因此, 存在线性映射 $\tilde{\Phi}: V_1^* \otimes \dots \otimes V_k^* \rightarrow L(V_1, \dots, V_k; \mathbb{R})$ 满足

$$\tilde{\Phi}(\omega^1 \otimes \dots \otimes \omega^k)(v_1, \dots, v_k) = \omega^1(v_1) \dots \omega^k(v_k),$$

我们需要证明 $\tilde{\Phi}$ 为同构映射。

记 $\dim V_1 = n_1, \dots, \dim V_k = n_k$, 根据Theorem 2.12可得, $\dim V_1^* = n_1, \dots, \dim V_k^* = n_k$, 因此 $\dim(V_1^* \otimes \dots \otimes V_k^*) = n_1 \dots n_k$ 。根据Lemma 2.17可得 $\dim L(V_1, \dots, V_k; \mathbb{R}) = n_1 \dots n_k$, 因此 $\dim L(V_1, \dots, V_k; \mathbb{R}) = n_1 \dots n_k$ 。

这表明 $\dim(V_1^* \otimes \dots \otimes V_k^*) = \dim L(V_1, \dots, V_k; \mathbb{R})$, 所以我们只需证明 $\tilde{\Phi}$ 是单射即可。

假设 $\tilde{\Phi}(\omega^1 \otimes \dots \otimes \omega^k)$ 是零映射, 换句话说 $\tilde{\Phi}(\omega^1 \otimes \dots \otimes \omega^k)(v_1, \dots, v_k) = \omega^1(v_1) \dots \omega^k(v_k) = 0$ 对任意 $v_1 \in V_1, \dots, v_k \in V_k$ 均成立。这表明 $\omega^1, \dots, \omega^k$ 至少有1个为0, 因此根据Corollary 2.14.1, $\omega^1 \otimes \dots \otimes \omega^k = 0 \otimes \dots \otimes 0$, 对应向量空间 $V_1^* \otimes \dots \otimes V_k^*$ 的零元。□

Corollary 2.18.1. 对于有限维向量空间 V_1, \dots, V_k , 记其对偶空间为 V_1^*, \dots, V_k^* , 所有多线性映射 $V_1^* \times \dots \times V_k^* \rightarrow \mathbb{R}$ 的集合为 $L(V_1^*, \dots, V_k^*; \mathbb{R})$, 那么

$$V_1 \otimes \dots \otimes V_k \cong L(V_1^*, \dots, V_k^*; \mathbb{R}).$$

Proof. 只需定义映射 $\Phi: V_1 \times \dots \times V_k \rightarrow L(V_1^*, \dots, V_k^*; \mathbb{R})$ 为

$$\Phi(v_1, \dots, v_k)(\omega^1, \dots, \omega^k) = \omega^1(v_1) \dots \omega^k(v_k)$$

即可, 后面原理完全相同。□

给定 \mathbb{R} 上的有限维向量空间 V , 因为对偶空间 V^* 也是向量空间, 所以我们可以对它们进行张量积。我们定义

$$\begin{aligned} T^k(V) &= \underbrace{V \otimes \dots \otimes V}_{k \text{ copies}}, \\ T^k(V^*) &= \underbrace{V^* \otimes \dots \otimes V^*}_{k \text{ copies}}, \\ T^{(k,l)}(V) &= \underbrace{V \otimes \dots \otimes V}_{k \text{ copies}} \otimes \underbrace{V^* \otimes \dots \otimes V^*}_{l \text{ copies}}, \end{aligned}$$

其中 $k, l \geq 0$ 。我们将 $T^k(V)$ 的任意元素称为向量空间 V 上的 k 阶逆变张量, $T^k(V^*)$ 的任意元素称为向量空间 V 上的 k 阶协变张量, $T^{(k,l)}(V)$ 的任意元素称为向量空间 V 上的 (k, l) -张量。对于特定的 $k, l \geq 0$, 我们可以将一些 (k, l) -张量简记为:

$$\begin{aligned} T^{(0,0)}(V) &= T^0(V^*) = T^0(V) = \mathbb{R}, \\ T^{(0,1)}(V) &= T^1(V^*) = V^*, \\ T^{(1,0)}(V) &= T^1(V) = V, \\ T^{(0,k)}(V) &= T^k(V^*), \\ T^{(k,0)}(V) &= T^k(V). \end{aligned}$$

因此, k 阶逆变张量是 $(k, 0)$ -张量, k 阶协变张量是 $(0, k)$ -张量, 当然我们没有必要采用后者的用语, 因为它较为繁琐。此外, 我们可以将任意向量视为 1 阶逆变张量, 将任意余向量视为 1 阶协变张量。

记 $\{E_i\}$ 为 V 的一组基, $\{\epsilon^i\}$ 为 V^* 的一组基, 因此

$$\begin{aligned} \mathcal{A} &= \{\epsilon^{i_1} \otimes \cdots \otimes \epsilon^{i_k}\}, \\ \mathcal{B} &= \{E_{i_1} \otimes \cdots \otimes E_{i_k}\}, \\ \mathcal{C} &= \{E_{i_1} \otimes \cdots \otimes E_{i_k} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_l}\}, \end{aligned}$$

分别为 $T^k(V^*), T^k(V), T^{(k,l)}(V)$ 的基, 其中 $i_1, \dots, i_k; j_1, \dots, j_l \in [1, n]$ 。因此 $\dim T^k(V^*) = \dim T^k(V) = n^k, \dim T^{(k,l)}(V) = n^{k+l}$ 。

根据 Theorem 2.18, Corollary 2.18.1, 并利用 Theorem 2.13 的结论 (V^* 的对偶空间同构于 V), 我们可以将任意一个 (k, l) -张量 $\alpha \in T^{(k,l)}(V)$ 都视为一个多线性映射

$$\alpha : \underbrace{V^* \times \cdots \times V^*}_{k \text{ copies}} \times \underbrace{V \times \cdots \times V}_{l \text{ copies}} \rightarrow \mathbb{R}.$$

将 α 表示为基的生成

$$\alpha = \sum_{i_1 \dots i_k j_1 \dots j_l} \alpha_{j_1 \dots j_l}^{i_1 \dots i_k} E_{i_1} \otimes \cdots \otimes E_{i_k} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_l},$$

其中 $i_1, \dots, i_k; j_1, \dots, j_l \in [1, n]$, 对于任意 $i_1, \dots, i_k; j_1, \dots, j_l$, $\alpha_{j_1 \dots j_l}^{i_1 \dots i_k} \in \mathbb{R}$ 。 $\sum_{i_1 \dots i_k j_1 \dots j_l}$ 代表 $i_1, \dots, i_k; j_1, \dots, j_l$ 求和取遍 $[1, n]$ 的所有整数。那么

$$\alpha_{j_1 \dots j_l}^{i_1 \dots i_k} = \alpha(\epsilon^{i_1}, \dots, \epsilon^{i_k}, E_{j_1}, \dots, E_{j_l}).$$

根据 Corollary 2.14.2, 我们可以定义 $F \in T^{(k,l)}(V)$ 与 $G \in T^{(k',l')}(V)$ 的张量积 $F \otimes G \in T^{(k+k',l+l')}(V)$ 为

$$\begin{aligned} F \otimes G(\omega^1, \dots, \omega^{k+k'}, v_1, \dots, v_{l+l'}) &= F(\omega^1, \dots, \omega^k, v_1, \dots, v_l) \\ &\quad \times G(\omega^{k+1}, \dots, \omega^{k+k'}, v_{l+1}, \dots, v_{l+l'}). \end{aligned}$$

Theorem 2.19. 对于有限维向量空间 V ，任意一个 $(1,1)$ -张量 $\mathcal{A} \in T^{(1,1)}(V)$ 都与 V 到自身的线性映射 $A : V \rightarrow V$ 一一对应。因此

$$T^{(1,1)}(V) \cong \text{End}(V),$$

其中 $\text{End}(V)$ 是 V 到自身的线性映射的集合。

Proof. 对于任意 $\omega \in V^*, v \in V$ ，定义线性映射 $\Phi : \text{End}(V) \rightarrow T^{(1,1)}(V)$ 为

$$\Phi(A)(\omega, v) = \omega(Av).$$

如果 $\Phi(A) = 0$ ，那么对于任意 $\omega \in V^*, v \in V$ ，都具有 $\omega(Av) = 0$ 。因此 $A = 0$ ，这表明 Φ 是单射。

但是还需要证明 Φ 是满射，即对于任意 $\mathcal{A} \in T^{(1,1)}(V)$ ，总是存在 $A \in \text{End}(V)$ 使得 $\Phi(A) = \mathcal{A}$ 。我们将 \mathcal{A} 表示为分量形式

$$\mathcal{A} = \sum_{i,j=1}^n \mathcal{A}_j^i E_i \otimes \epsilon^j.$$

其中 E_i, ϵ^j 分别为 V, V^* 的一组基。令线性映射 A 满足，对于任意 $j \in [1, n]$ ，都具有

$$AE_j = \sum_{i=1}^n \mathcal{A}_j^i E_i.$$

我们只需证明 A 是满足条件的线性映射即可，这样就完成了证明。

将任意 ω, v 表示为分量形式

$$\begin{aligned} \omega &= \sum_{k=1}^n \omega_k \epsilon^k, \\ v &= \sum_{l=1}^n v^l E_l. \end{aligned}$$

因此

$$\begin{aligned} \mathcal{A}(\omega, v) &= \sum_{i,j=1}^n \mathcal{A}_j^i E_i \left(\sum_{k=1}^n \omega_k \epsilon^k \right) \epsilon^j \left(\sum_{l=1}^n v^l E_l \right) = \sum_{i,j,k,l} \mathcal{A}_j^i \omega_k v^l \delta_i^k \delta_l^j \\ &= \sum_{i,j=1}^n \mathcal{A}_j^i \omega_i v^j. \end{aligned}$$

而

$$\begin{aligned}
\Phi(A)(\omega, v) &= \omega(Av) = \omega\left(A \sum_{j=1}^n v^j E_j\right) = \omega\left(\sum_{j=1}^n v^j A E_j\right) \\
&= \omega\left(\sum_{i,j=1}^n \mathcal{A}_j^i v^j E_i\right) = \sum_{i,j=1}^n \mathcal{A}_j^i v^j \omega(E_i) \\
&= \sum_{i,j=1}^n \mathcal{A}_j^i \omega_i v^j.
\end{aligned}$$

因此 $\Phi(A) = \mathcal{A}$ 。

□

2.6.1 张量场*

¹¹ 利用 \mathbb{R}^n 与其对偶空间基的记号, 可得

$$\begin{aligned}
\mathcal{A} &= \{dx^{i_1} \otimes \cdots \otimes dx^{i_k}\}, \\
\mathcal{B} &= \left\{ \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_k}} \right\}, \\
\mathcal{C} &= \left\{ \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_l} \right\},
\end{aligned}$$

分别为 $T^k(\mathbb{R}^{n*}), T^k(\mathbb{R}^n), T^{(k,l)}(\mathbb{R}^n)$ 的基。这样, \mathbb{R}^n 上的任意 (k, l) -张量 A 都可以表示为

$$A = \sum_{i_1 \dots i_k j_1 \dots j_l} A_{j_1 \dots j_l}^{i_1 \dots i_k} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_l},$$

其中 $i_1, \dots, i_k; j_1, \dots, j_l \in [1, n]$, 对于任意 $i_1, \dots, i_k; j_1, \dots, j_l$, $A_{j_1 \dots j_l}^{i_1 \dots i_k} \in \mathbb{R}$ 。 $\sum_{i_1 \dots i_k j_1 \dots j_l}$ 代表 $i_1, \dots, i_k; j_1, \dots, j_l$ 求和取遍 $[1, n]$ 的所有整数。那么

$$A_{j_1 \dots j_l}^{i_1 \dots i_k} = A\left(dx^{i_1}, \dots, dx^{i_k}, \frac{\partial}{\partial x^{j_1}}, \dots, \frac{\partial}{\partial x^{j_l}}\right).$$

与向量场的定义类似, 对于 $U \subseteq \mathbb{R}^n$, 我们定义 U 的 **(k, l)-张量丛** 为

$$T^{(k,l)}TU = U \times T^{(k,l)}(\mathbb{R}^n) = \left\{ (p, A) : p \in U, A \in T^{(k,l)}(\mathbb{R}^n) \right\},$$

记映射 $\pi : T^{(k,l)}TU \rightarrow U$ 为

$$\pi(p, A) = p,$$

¹¹ 符号*表示这一章节可以跳过, 建议跳转到第86页的章节2.9。

因此，映射 π 的任意截面 $\sigma : U \rightarrow T^{(k,l)}TU$ 都可以被表示为

$$\sigma(p) = (p, B),$$

其中 $B \in T^{(k,l)}(\mathbb{R}^n)$ 是任意 (k, l) -张量。我们将 (\mathbf{k}, \mathbf{l}) -张量场定义为映射 $\pi : T^{(k,l)}TU \rightarrow U$ 的截面。当然，你也可以认为它将 U 上的任意一点 p 都赋予一个 (k, l) -张量。任意 (k, l) -张量场 A 都可以被表示为

$$A = \sum_{i_1 \dots i_k j_1 \dots j_l} A_{j_1 \dots j_l}^{i_1 \dots i_k} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_l},$$

其中 $A_{j_1 \dots j_l}^{i_1 \dots i_k} : U \rightarrow \mathbb{R}$ 是 n^{k+l} 个定义域为 U 的函数。当这些函数全部都是光滑函数时，称张量场是光滑的。对于 \mathbf{k} 阶逆变张量场， \mathbf{k} 阶协变张量场，情况是相同的，因为 k 阶逆变张量是 $(k, 0)$ -张量， k 阶协变张量是 $(0, k)$ -张量。

因为余向量是1阶协变张量，即 $(0, 1)$ -张量，所以余向量场就是 $(0, 1)$ -张量场，我们将它记为 $\mathcal{T}(U)$ 。

正如我们可以将任意 (k, l) -张量视为 k 个余向量与 l 个向量到一个实数的多线性映射，同样的，我们可以将任意光滑 (k, l) -张量场视为将 k 个光滑余向量场与 l 个光滑向量场到一个 \mathbb{R}^n 上的光滑函数的多线性映射：

Lemma 2.20. 任意光滑 (k, l) -张量场 F 都与 $C^\infty(U)$ 上的多线性映射

$$F : \underbrace{\mathcal{T}(U) \times \dots \times \mathcal{T}(U)}_{k \text{ copies}} \times \underbrace{\mathcal{X}(U) \times \dots \times \mathcal{X}(U)}_{l \text{ copies}} \rightarrow C^\infty(U),$$

$$F(\omega^1, \dots, \omega^k, X_1, \dots, X_l)(p) = F(p)(\omega^1(p), \dots, \omega^k(p), X_1(p), \dots, X_l(p)), \quad p \in U$$

一一对应。 $C^\infty(U)$ 上的多线性映射定义为，对于任意光滑函数 $u, v \in C^\infty(U)$ ，都具有

$$\mathcal{F}(\dots, u\alpha + v\beta, \dots) = u\mathcal{F}(\dots, \alpha, \dots) + v\mathcal{F}(\dots, \beta, \dots).$$

但是，其证明超出本文范围，因此不做展开。

2.6.2 反对称张量*

¹² 向量空间 V 上的 k 阶协变张量是**反对称张量**，当且仅当对于任意向量 $v_1, \dots, v_k \in V$ 与任意整数 j, k 满足 $1 \leq i < j \leq k$ ，都具有

$$\alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k),$$

通俗来说，交换任意2个向量的位置，对称张量对应的映射值不变，反对称张量对应的映射值取到原来的相反数。我们将 k 阶协变反对称张量的集合记为 $\Lambda^k(V^*)$ 。当 $k = 1$ 时，规定 $\Lambda^1(V^*) = T^1(V^*) = V^*$ ，即所有余向量的集合；当 $k = 0$ 时，规定 $\Lambda^0(V^*) = \mathbb{R}$ 。

¹²符号*表示这一章节可以跳过，建议跳转到第86页的章节2.9。

$\Lambda^k(V^*)$ 是 \mathbb{R} 上的向量空间, 因为对于 $\omega, \eta \in \Lambda^k(V^*)$,

$$\begin{aligned}(a\omega + b\eta)(v_1, \dots, v_i, \dots, v_j, \dots, v_k) &= a\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) + b\eta(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \\ &= -a\omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k) - b\eta(v_1, \dots, v_j, \dots, v_i, \dots, v_k) \\ &= -(a\omega + b\eta)(v_1, \dots, v_j, \dots, v_i, \dots, v_k),\end{aligned}$$

这表明 $a\omega + b\eta \in \Lambda^k(V^*)$ 。

Lemma 2.21. 令 α 是有限维向量空间 V 上的 k 阶协变张量, 那么以下命题是等价的:

(A) α 是反对称张量, 即 $\alpha \in \Lambda^k(V^*)$ 。

(B) 如果向量 v_1, \dots, v_k 是线性相关的, 那么 $\alpha(v_1, \dots, v_k) = 0$ 。

(C) 当向量 v_1, \dots, v_k 出现2个相同的, 那么 $\alpha(v_1, \dots, v_k) = 0$ 。

Proof. (A) \Rightarrow (C)与(B) \Rightarrow (C)是显然的, 我们只需证明(C) \Rightarrow (A)与(C) \Rightarrow (B)即可。假设(C)是成立的, 那么对于向量 v_1, \dots, v_k ,

$$\begin{aligned}0 &= \alpha(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_k) \\ &= \alpha(v_1, \dots, v_i, \dots, v_i, \dots, v_k) + \alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \\ &\quad + \alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k) + \alpha(v_1, \dots, v_j, \dots, v_j, \dots, v_k) \\ &= \alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_k) + \alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k),\end{aligned}$$

这表明 α 是反对称张量, 因此(C) \Rightarrow (A)。

如果向量 v_1, \dots, v_k 是线性相关的, 为简便起见, 假设 $v_k = \sum_{j=1}^{k-1} a^j v_j$ 。那么

$$\alpha(v_1, \dots, v_k) = \sum_{j=1}^{k-1} a^j \alpha(v_1, \dots, v_{k-1}, v_j),$$

而在等式右边的任意一项, α 作用的 k 个向量当中都具有2个向量相同, 因此等式右边的任意一项取值均为0, 这表明 $\alpha(v_1, \dots, v_k) = 0$, 因此(C) \Rightarrow (B)。□

对于任意 k 阶协变张量 $\alpha \in T^k(V^*)$, 定义映射 $\text{Alt} : T^k(V^*) \rightarrow \Lambda^k(V^*)$, 它将任意 k 阶协变张量转换为反对称张量, 其具体形式为:

$$(\text{Alt } \alpha)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\text{sgn } \sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}),$$

其中 S_k 是 k 阶置换群, $\text{sgn } \sigma$ 反映了数组 $\sigma = (\sigma(1), \dots, \sigma(k))$ 变为 $(1, \dots, k)$ 所需两两对换的次数: 当次数是偶数, $\text{sgn } \sigma = 1$; 当次数是奇数, $\text{sgn } \sigma = -1$ 。例如, 令 $\sigma = (2, 5, 4, 3, 1) \in S_5$, 为了将它变为 $(1, 2, 3, 4, 5)$, 我们需要

(A) 将2与1对换, 得到(1, 5, 4, 3, 2)。

(B) 将5与2对换, 得到(1, 2, 4, 3, 5)。

(C) 将3与4对换, 得到(1, 2, 3, 4, 5)。

因此, 对换次数为3, 是奇数, 这表明 $\text{sgn } \sigma = -1$ 。但是, 对换的方式不止一种, 我们可以选取其他对换方式:

(A) 将5与1对换, 得到(2, 1, 4, 3, 5)。

(B) 将4与3对换, 得到(2, 1, 3, 4, 5)。

(C) 将2与1对换, 得到(1, 2, 3, 4, 5)。

但是 $\text{sgn } \sigma$ 依然为 -1 。群论已经证明, $\text{sgn } \sigma$ 的取值与对换的方式无关, 在此不做证明, 因此 $\text{sgn } \sigma$ 是良定义的。

因此, 如果 α 是2阶协变张量, 那么

$$(\text{Alt } \alpha)(v_1, v_2) = \frac{1}{2} (\alpha(v_1, v_2) - \alpha(v_2, v_1)),$$

如果 α 是3阶协变张量, 那么

$$\begin{aligned} (\text{Alt } \alpha)(v_1, v_2, v_3) = & \frac{1}{6} (\alpha(v_1, v_2, v_3) + \alpha(v_2, v_3, v_1) + \alpha(v_3, v_1, v_2) \\ & - \alpha(v_2, v_1, v_3) - \alpha(v_1, v_3, v_2) - \alpha(v_3, v_2, v_1)), \end{aligned}$$

对于更高阶的情况, 以此类推。

很容易验证 $\text{Alt } \alpha$ 是反对称张量, 并且 $\text{Alt } \alpha = \alpha$ 当且仅当 α 是反对称张量。

对于反对称张量 $\omega \in \Lambda^k(V^*)$ 与 $\eta \in \Lambda^l(V^*)$, 定义它们的楔积为 $k + l$ 阶反对称张量 $\omega \wedge \eta \in \Lambda^{k+l}(V^*)$:

$$\omega \wedge \eta = \frac{(k+l)!}{k!l!} \text{Alt}(\omega \otimes \eta).$$

因此, 对于余向量 $\omega, \eta \in \Lambda^1(V^*)$,

$$\omega \wedge \eta(v_1, v_2) = 2 \text{Alt}(\omega \otimes \eta)(v_1, v_2) = \omega(v_1)\eta(v_2) - \eta(v_1)\omega(v_2). \quad (2.32)$$

楔积具有以下性质:

Theorem 2.22. 令 $\omega, \omega', \eta, \eta'$ 与 ξ 都是反对称张量。那么它们满足

(A) 双线性:

$$\begin{aligned} (a\omega + b\omega') \wedge \eta &= a(\omega \wedge \eta) + b(\omega' \wedge \eta), \\ \eta \wedge (a\omega + b\omega') &= a(\eta \wedge \omega) + b(\eta \wedge \omega'). \end{aligned}$$

(B) 结合律:

$$\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \xi. \quad (2.33)$$

(C) 反对称性: 对于 $\omega \in \Lambda^k(V^*)$ 与 $\eta \in \Lambda^l(V^*)$,

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega. \quad (2.34)$$

特别的, 对于余向量 $\omega, \eta \in \Lambda^1(V^*)$,

$$\omega \wedge \eta = -\eta \wedge \omega,$$

因此, 对于任意余向量 $\omega \in \Lambda^1(V^*)$, 都具有

$$\omega \wedge \omega = 0.$$

(D) 对于任意余向量 $\omega^1, \dots, \omega^k$ 与向量 v_1, \dots, v_k ,

$$\omega^1 \wedge \dots \wedge \omega^k(v_1, \dots, v_k) = \det(\omega^j(v_i)),$$

其中 $\det(\omega^j(v_i))$ 代表第 i 行, 第 j 列元素为 $\omega^j(v_i)$ 的矩阵的行列式。

Proof. 注意: 该证明只能提供一个直观的思路, 关于严谨的证明请阅读微分几何方向的资料。

第1个性质是显然的, 因为张量积运算是双线性的, Alt运算是线性的。第2个性质计算较为繁琐, 而且对本文后面的内容帮助不大, 因此不做证明。第3个性质的直观理解如Figure 2.5所示。对于第4个性质, 我们利用数学归纳法: 当 $k = 1$ 时, 显然成立; 假设这一结论对于 $k - 1$ 是成立的, 记 $\omega^1 \wedge \dots \wedge \omega^{k-1} = \eta$, 那么根据假设

$$\eta(v_1, \dots, v_{k-1}) = \det \begin{pmatrix} \omega^1(v_1) & \dots & \omega^{k-1}(v_1) \\ \vdots & \ddots & \vdots \\ \omega^1(v_{k-1}) & \dots & \omega^{k-1}(v_{k-1}) \end{pmatrix},$$

因此

$$\begin{aligned} \omega^1 \wedge \dots \wedge \omega^k(v_1, \dots, v_k) &= \eta \wedge \omega^k(v_1, \dots, v_k) = k \text{Alt}(\eta \otimes \omega)(v_1, \dots, v_k) \\ &= \frac{1}{(k-1)!} \sum_{\sigma \in S_k} (\text{sgn } \sigma) \eta(v_{\sigma(1)}, \dots, v_{\sigma(k-1)}) \omega^k(v_{\sigma(k)}) \\ &= \frac{1}{(k-1)!} \sum_{\sigma \in S_k} (\text{sgn } \sigma) \det \begin{pmatrix} \omega^1(v_{\sigma(1)}) & \dots & \omega^{k-1}(v_{\sigma(1)}) \\ \vdots & \ddots & \vdots \\ \omega^1(v_{\sigma(k-1)}) & \dots & \omega^{k-1}(v_{\sigma(k-1)}) \end{pmatrix} \omega^k(v_{\sigma(k)}), \end{aligned}$$

令 $m \in [1, k]$ 为任意正整数。那么, 当 $\sigma = (1 \ \dots \ m-1 \ m+1 \ \dots \ k \ m)$ 时, $\text{sgn } \sigma = (-1)^{k-m}$ 。固定 $\sigma(k) = m$ 不动, 对换 $\sigma(1), \dots, \sigma(k-1)$ 其中的任意2个元素, 可以发现, 表

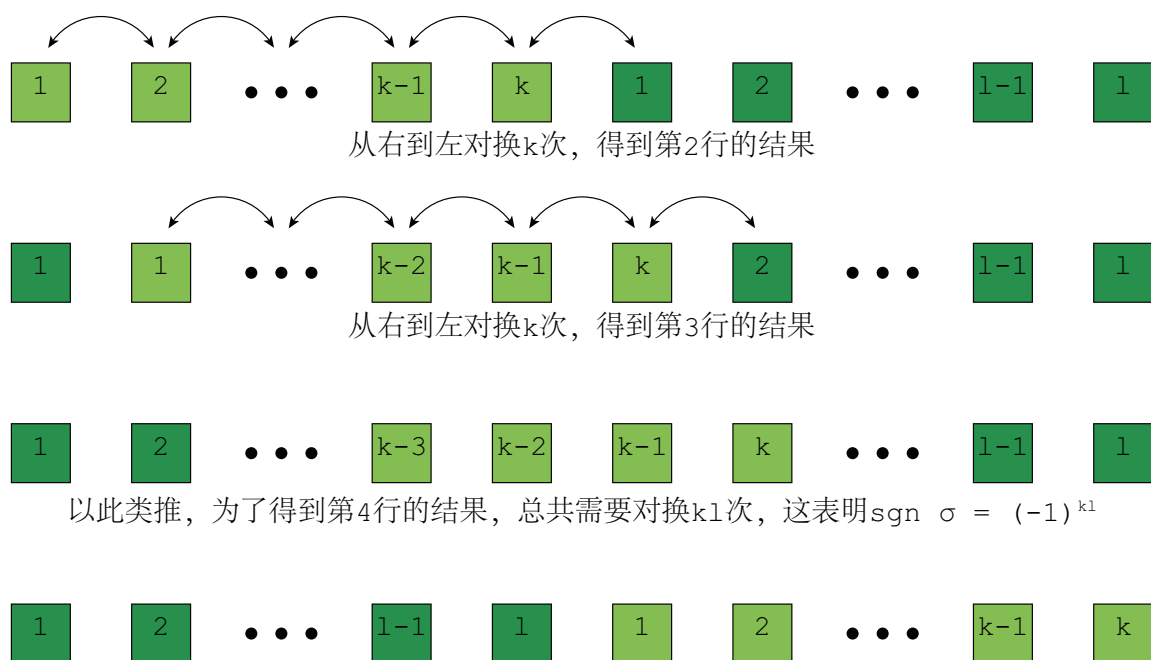


Figure 2.5: 交换 k 个元素的数组与 l 个元素的数组，需要的对换次数为 kl

达式 $\text{sgn } \sigma$ 的值取到相反数， \det 的值同样取到相反数，因此它们的乘积保持不变。这表明当 $\sigma(k) = m$ 时，求和项的取值恒为

$$(-1)^{k-m} \det \begin{pmatrix} \omega^1(v_1) & \cdots & \omega^{k-1}(v_1) \\ \vdots & & \vdots \\ \omega^1(v_{m-1}) & \cdots & \omega^{k-1}(v_{m-1}) \\ \omega^1(v_{m+1}) & \cdots & \omega^{k-1}(v_{m+1}) \\ \vdots & & \vdots \\ \omega^1(v_k) & \cdots & \omega^{k-1}(v_k) \end{pmatrix} \omega^k(v_m),$$

它们与 $\sigma(1), \dots, \sigma(k-1)$ 的取值无关。而 $\sigma(1), \dots, \sigma(k-1)$ 共有 $(k-1)!$ 个不同的排列，因此

$$\begin{aligned}
& \omega^1 \wedge \dots \wedge \omega^k(v_1, \dots, v_k) \\
&= \frac{1}{(k-1)!} \sum_{m=1}^k \sum_{\sigma \in S_k, \sigma(k)=m} (-1)^{k-m} \det \begin{pmatrix} \omega^1(v_1) & \dots & \omega^{k-1}(v_1) \\ \vdots & & \vdots \\ \omega^1(v_{m-1}) & \dots & \omega^{k-1}(v_{m-1}) \\ \omega^1(v_{m+1}) & \dots & \omega^{k-1}(v_{m+1}) \\ \vdots & & \vdots \\ \omega^1(v_k) & \dots & \omega^{k-1}(v_k) \end{pmatrix} \omega^k(v_m) \\
&= \sum_{m=1}^k (-1)^{k-m} \det \begin{pmatrix} \omega^1(v_1) & \dots & \omega^{k-1}(v_1) \\ \vdots & & \vdots \\ \omega^1(v_{m-1}) & \dots & \omega^{k-1}(v_{m-1}) \\ \omega^1(v_{m+1}) & \dots & \omega^{k-1}(v_{m+1}) \\ \vdots & & \vdots \\ \omega^1(v_k) & \dots & \omega^{k-1}(v_k) \end{pmatrix} \omega^k(v_m) \\
&= \det \begin{pmatrix} \omega^1(v_1) & \dots & \omega^k(v_1) \\ \vdots & \ddots & \vdots \\ \omega^1(v_k) & \dots & \omega^k(v_k) \end{pmatrix},
\end{aligned}$$

因此，这一结论对于 k 是成立的。 \square

利用第4个性质，我们很容易得出以下结论：

Corollary 2.22.1. 令 $\epsilon^1, \dots, \epsilon^n$ 是 V^* 上的一组基， E_1, \dots, E_n 是 V 上的一组基，并且 $\epsilon^i(E_j) = \delta_j^i$ ，那么对于 $i_1, \dots, i_k; j_1, \dots, j_k \in [1, n]$ ，

$$\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k}(E_{j_1}, \dots, E_{j_k}) = \det \left(\delta_{j_p}^{i_q} \right),$$

其中 $\det \left(\delta_{j_p}^{i_q} \right)$ 代表第 p 行，第 q 列元素为 $\delta_{j_p}^{i_q}$ 的矩阵的行列式。这表明，如果 i_1, \dots, i_k 出现2个相同的，那么矩阵 $\left(\delta_{j_p}^{i_q} \right)$ 会有2列相同，因此 $\det \left(\delta_{j_p}^{i_q} \right) = 0$ 对于任意 j_1, \dots, j_k 的选取都成立，意味着 $\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} = 0$ 。

Theorem 2.23. 令 V 是 n 维向量空间， $\epsilon^1, \dots, \epsilon^n$ 是对偶空间 V^* 上的一组基。那么

(A) 当 $k \leq n$ 时，

$$\mathcal{E} = \{ \epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} : 1 \leq i_1 < \dots < i_k \leq n \}$$

是向量空间 $\Lambda^k(V^*)$ 的一组基，因此

$$\dim \Lambda^k(V^*) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

(B) 当 $k > n$ 时, $\Lambda^k(V^*)$ 只有元素 $\{0\}$, 因此 $\dim \Lambda^k(V^*) = 0$ 。

Proof. 当 $k > n$ 时, 对于 $i_1, \dots, i_k \in [1, n]$, i_1, \dots, i_k 必定出现2个相同的, 根据Corollary 2.22.1可得 $\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} = 0$ 。

当 $k \leq n$ 时, 令 $\alpha \in \Lambda^k(V^*)$ 是任意 k 阶反对称张量。记 $\alpha_{i_1 \dots i_k} = \alpha(E_{i_1}, \dots, E_{i_k})$, 其中 E_1, \dots, E_n 是 V 上的一组基, 并且 $\epsilon^i(E_j) = \delta_j^i$; 此外 $1 \leq i_1 < \dots < i_k \leq n$ 。记 $j_1, \dots, j_k \in [1, n]$, 根据Corollary 2.22.1可得

$$\begin{aligned} & \sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1 \dots i_k} \epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} (E_{j_1}, \dots, E_{j_k}) \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1 \dots i_k} \det \begin{pmatrix} \delta_{j_1}^{i_1} & \dots & \delta_{j_1}^{i_k} \\ \vdots & \ddots & \vdots \\ \delta_{j_k}^{i_1} & \dots & \delta_{j_k}^{i_k} \end{pmatrix} \\ &= \alpha_{m_1 \dots m_k} \det \begin{pmatrix} \delta_{j_1}^{m_1} & \dots & \delta_{j_1}^{m_k} \\ \vdots & \ddots & \vdots \\ \delta_{j_k}^{m_1} & \dots & \delta_{j_k}^{m_k} \end{pmatrix} \\ &= \alpha(E_{m_1}, \dots, E_{m_k}) \det \begin{pmatrix} \delta_{j_1}^{m_1} & \dots & \delta_{j_1}^{m_k} \\ \vdots & \ddots & \vdots \\ \delta_{j_k}^{m_1} & \dots & \delta_{j_k}^{m_k} \end{pmatrix}, \end{aligned}$$

其中 $1 \leq m_1 < \dots < m_k \leq n$ 是将 j_1, \dots, j_k 从小到大排列的结果。令 j_1, \dots, j_k 变换到 m_1, \dots, m_k 所需的对换次数为 τ , 那么

$$\det \begin{pmatrix} \delta_{j_1}^{m_1} & \dots & \delta_{j_1}^{m_k} \\ \vdots & \ddots & \vdots \\ \delta_{j_k}^{m_1} & \dots & \delta_{j_k}^{m_k} \end{pmatrix} = (-1)^\tau,$$

因为矩阵每2列发生对换, 行列式取值变为原来的相反数。同样的,

$$\alpha(E_{m_1}, \dots, E_{m_k}) = (-1)^\tau \alpha(E_{j_1}, \dots, E_{j_k}),$$

因为 α 是反对称张量。这表明

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1 \dots i_k} \epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} (E_{j_1}, \dots, E_{j_k}) = (-1)^{2\tau} \alpha(E_{j_1}, \dots, E_{j_k}) = \alpha(E_{j_1}, \dots, E_{j_k}),$$

因此 $\sum_{1 \leq i_1 < \dots < i_k \leq n} \alpha_{i_1 \dots i_k} \epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k} = \alpha$, 这表明 $\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k}$ 生成向量空间 $\Lambda^k(V^*)$ 。
假设一些形如

$$\alpha_{i_1 \dots i_k} \epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k}, \quad 1 \leq i_1 < \dots < i_k \leq n$$

的线性组合为0, 那么我们将这一线性组合作用于 E_{j_1}, \dots, E_{j_k} , 其中 $j_1, \dots, j_k \in [1, n]$ 是任意正整数。将 j_1, \dots, j_k 从小到大排列的结果记为 m_1, \dots, m_k , 这样可以得出 $\alpha_{m_1 \dots m_k} = 0$ 对于任意 $1 \leq m_1 < \dots < m_k \leq n$ 恒成立, 因此 $\epsilon^{i_1} \wedge \dots \wedge \epsilon^{i_k}$ 是线性无关的。□

定义向量空间

$$\Lambda(V^*) = \bigoplus_{k=0}^n \Lambda^k(V^*).$$

根据Theorem 2.23可得, 向量空间 $\Lambda(V^*)$ 的维数为

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

可以得出, 楔积 \wedge 是 $\Lambda(V^*)$ 上的代数运算 $\wedge : \Lambda(V^*) \times \Lambda(V^*) \rightarrow \Lambda(V^*)$: 它将 $\omega \in \Lambda^k(V^*) \subseteq \Lambda(V^*)$ 与 $\eta \in \Lambda^l(V^*) \subseteq \Lambda(V^*)$ 映射到 $\omega \wedge \eta \in \Lambda^{k+l}(V^*) \subseteq \Lambda(V^*)$ 。对于 $\omega \in \Lambda^k(V^*)$, 记 ω 的**次数**为 $|\omega| = k$, 因此, \wedge 将次数为 k 的元素与次数为 l 的元素映射到次数为 $k + l$ 的元素。此外, 根据(2.33), \wedge 满足结合律; 根据(2.34), $\omega \wedge \eta = (-1)^{|\omega||\eta|} \eta \wedge \omega$ 。我们将 \wedge 运算下的向量空间 $\Lambda(V^*)$ 称为**外代数**。

2.7 微分形式*

¹³ 对于 $U \subseteq \mathbb{R}^n$, **k-微分形式**定义为 U 上的 k 阶反对称张量场——将 U 上的任意一点都赋予一个 k 阶反对称张量。通常, 我们将 k -微分形式简称为 k -形式。根据Theorem 2.23, 我们可以将 k 形式表示为

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

其中 $\omega_{i_1 \dots i_k} \in C^\infty(U)$ 。我们将 k -形式的集合记为 $\Omega^k(U)$ 。而1阶反对称张量就是余向量, 因此1-形式是余向量场, 即 $\Omega^1(U) = \mathcal{T}(U)$ 。规定0-形式是定义域为 U 的光滑函数, 即 $\Omega^0(U) = C^\infty(U)$ 。

规定光滑函数 $f \in C^\infty(U)$ 与 k -形式的乘积为

$$f\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} f\omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

因为 $\omega_{i_1 \dots i_k}$ 是光滑函数, 所以 $f\omega_{i_1 \dots i_k}$ 还是光滑函数, 这表明乘积得到的结果还是 k -形式, 因此 $\Omega^k(U)$ 是交换环 $C^\infty(U)$ 上的模。

k -形式 ω 与 l -形式 η 楔积之后的结果是 $(k + l)$ -形式 $\omega \wedge \eta$, 其在任意 $p \in U$ 上的取值为

$$(\omega \wedge \eta)(p) = \omega(p) \wedge \eta(p),$$

相当于, $\omega \wedge \eta$ 在每一点的取值(它是 $(k + l)$ -阶反对称张量)都对应于 ω, η 在该点对应取值(它们分别是 k, l 阶反对称张量)的楔积。微分形式的楔积在交换环 $C^\infty(U)$ 上是双线性的: 因为光滑函数与微分形式的乘积, 相当于在对应的每一点乘上一个常数, 而根据Theorem

¹³符号*表示这一章节可以跳过, 建议跳转到第86页的章节2.9。

2.22可得，反对称张量的楔积是 \mathbb{R} -双线性的。因此，对于 $f, g \in C^\infty(U)$ 与微分形式 ω, η ，都具有

$$(f\omega) \wedge (g\eta) = fg\omega \wedge \eta. \quad (2.35)$$

定义 k -形式的外微分 $d: \Omega^k(U) \rightarrow \Omega^{k+1}(U)$ 为

(A) 对于0-形式，即 U 上的光滑函数 $f: U \rightarrow \mathbb{R}$ ，其外微分为

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i.$$

因此，对于向量场 $X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i} \in \mathcal{X}(U)$ ，

$$df(X) = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i(X) = \sum_{i=1}^n \frac{\partial f}{\partial x^i} X^i = Xf. \quad (2.36)$$

(B) 对于 k -形式(对于 $k=0$ 同样适用)，其外微分为

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

其外微分为

$$\begin{aligned} d\omega &= \sum_{1 \leq i_1 < \dots < i_k \leq n} d\omega_{i_1 \dots i_k} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{j=1}^n \frac{\partial \omega_{i_1 \dots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}. \end{aligned} \quad (2.37)$$

观察0-形式 f 的外微分，你是否对此感到熟悉呢？因此，这也解释了我们为什么将 \mathbb{R}^{n*} 上的基记为 dx^1, \dots, dx^n ，因为它确实能带来一定的直观性！

关于外微分，它自然有一些性质：

Theorem 2.24. (A) d 在 \mathbb{R} 上是线性的，即对于任意 $a, b \in \mathbb{R}$ 与 $\omega, \eta \in \Omega^k(U)$ ，都满足

$$d(a\omega + b\eta) = ad\omega + bd\eta.$$

(B) 将任意 k -形式计算2次及以上的外微分后，得到的结果恒为0：

$$d \circ d \equiv 0. \quad (2.38)$$

(C) 对于 k -形式 $\omega \in \Omega^k(U)$ 与 l -形式 $\eta \in \Omega^l(U)$ ，

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta.$$

Proof. 第1个性质是显然的。对于第2个性质,

$$\begin{aligned}
& d \circ d \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \right) \\
&= d \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{i=1}^n \frac{\partial \omega_{i_1 \dots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \right) \\
&= \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{i,j=1}^n \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} dx^i \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&= \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{1 \leq i < j \leq n} \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} dx^i \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&\quad + \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{1 \leq j < i \leq n} \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} dx^i \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&= \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{1 \leq i < j \leq n} \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} dx^i \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&\quad + \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{1 \leq j < i \leq n} \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} (-dx^j \wedge dx^i) \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&= \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{1 \leq i < j \leq n} \left(\frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^i \partial x^j} - \frac{\partial^2 \omega_{i_1 \dots i_k}}{\partial x^j \partial x^i} \right) dx^i \wedge dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\
&\equiv 0.
\end{aligned}$$

对于3-4行, 根据(2.34)可得 $dx^i \wedge dx^i = 0$, 因此 $i = j$ 对应的项取值均为0, 对于 $i \neq j$, 我们可以将情况分为 $i < j$ 或 $j < i$ 。对于5-6行, 只需将第5行加法的第2项 i, j 进行交换就可以得到第6行。

对于第3个性质, 因为我们可以将 k -形式表示为

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} \omega_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

l -形式同样如此。直接计算是较为复杂的, 但是根据第1个性质(d 是 \mathbb{R} -线性的)以及(2.35)(微分形式的楔积是 $C^\infty(U)$ -双线性的, 自然是 \mathbb{R} -双线性的, 在这里将实数 \mathbb{R} 视为常函数 $f \equiv C \in \mathbb{R}$), 我们只需要证明, 该性质在

$$\begin{aligned}
\omega &= f dx^{i_1} \wedge \dots \wedge dx^{i_k}, \\
\eta &= g dx^{j_1} \wedge \dots \wedge dx^{j_l},
\end{aligned}$$

的情况下是成立的即可。其中 $f, g \in C^\infty(U)$, 并且 $1 \leq i_1 < \cdots < i_k \leq n, 1 \leq j_1 < \cdots < j_l \leq n$ 。根据(2.35)可得

$$\omega \wedge \eta = fg dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l},$$

那么

$$\begin{aligned} d(\omega \wedge \eta) &= \sum_{m=1}^n \frac{\partial(fg)}{\partial x^m} dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \\ &= \sum_{m=1}^n \left(\frac{\partial f}{\partial x^m} g + f \frac{\partial g}{\partial x^m} \right) dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l}, \end{aligned}$$

而

$$\begin{aligned} &\sum_{m=1}^n \frac{\partial f}{\partial x^m} g dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \\ &= \sum_{m=1}^n \left(\frac{\partial f}{\partial x^m} dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \right) \wedge (g dx^{j_1} \wedge \cdots \wedge dx^{j_l}) \\ &= d\omega \wedge \eta, \end{aligned}$$

根据(2.34), 对于1阶反对称张量 dx, dy , $dx \wedge dy = -dy \wedge dx$, 因此

$$\begin{aligned} &\sum_{m=1}^n f \frac{\partial g}{\partial x^m} dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \\ &= (-1)^k \sum_{m=1}^n (f dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \wedge \left(\frac{\partial g}{\partial x^m} dx^m \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \right) \\ &= (-1)^k \omega \wedge d\eta. \end{aligned}$$

□

对于 $\omega \in \Omega^k(U)$, 如果 $d\omega = 0$, 那么称 ω 为**闭形式**; 如果存在 $\eta \in \Omega^{k-1}(U)$ 使得 $\omega = d\eta$, 那么称 ω 为**恰当形式**。当 $\omega \in \Omega^n(U)$ 时, 根据Theorem 2.23可得 $\Omega^{n+1}(U) = 0$, 因此 $d\omega \equiv 0$, 这表明所有 n -形式都是闭形式。

Theorem 2.25. 任意恰当形式都是闭形式。

Proof. 如果 $\omega \in \Omega^k(U)$ 是恰当形式, 那么存在 $\eta \in \Omega^{k-1}(U)$ 使得 $\omega = d\eta$ 。而根据(2.38)可得 $d\omega = d(d\eta) = 0$, 因此 ω 是闭形式。 □

但是，这不代表任何闭合形式都是恰当形式，这与 U 的性质有关：如果 U 的 p 阶de Rham上同调群(de Rham cohomology group) $H_{\text{dR}}^p(U)$ 为0，那么任意闭合的 p -形式 $\omega \in \Omega^p(U)$ 都是恰当形式。所以计算 U 的de Rham上同调群是关键的一步——de Rham定理表明，对于任意整数 $p \geq 0$ ， U 的 p 阶de Rham上同调群同构于它的 p 阶奇异上同调群(Singular cohomology group)！因此，这个问题就可以转化为代数拓扑的问题，但是篇幅有限，这些话题我们都无法展开。以下的定理自然就无法证明了，因为它需要计算 \mathbb{R}^n 的de Rham上同调群：

Theorem 2.26. 对于任意正整数 $p \geq 1$ ， \mathbb{R}^n 上的任意闭合的 p -形式都是恰当形式。

2.7.1 \mathbb{R}^3 上的微分形式*

¹⁴ 我们将在 \mathbb{R}^3 上分析其1-形式与2-形式。任意1-形式都可以表示为

$$\omega = Pdx + Qdy + Rdz,$$

其中 $P, Q, R \in C^\infty(\mathbb{R}^3)$ 。根据(2.37)计算它的外微分，并且根据(2.34)得到 $dx \wedge dx = dy \wedge dy = dz \wedge dz = 0$ ， $dx \wedge dy = -dy \wedge dx$ ：

$$\begin{aligned} d\omega &= dP \wedge dx + dQ \wedge dy + dR \wedge dz \\ &= \left(\frac{\partial P}{\partial x} dx + \frac{\partial P}{\partial y} dy + \frac{\partial P}{\partial z} dz \right) \wedge dx + \left(\frac{\partial Q}{\partial x} dx + \frac{\partial Q}{\partial y} dy + \frac{\partial Q}{\partial z} dz \right) \wedge dy \\ &\quad + \left(\frac{\partial R}{\partial x} dx + \frac{\partial R}{\partial y} dy + \frac{\partial R}{\partial z} dz \right) \wedge dz \\ &= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy \wedge dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz \wedge dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy. \end{aligned}$$

任意2-形式都可以表示为

$$\eta = udy \wedge dz + vdz \wedge dx + wdx \wedge dy,$$

其中 $u, v, w \in C^\infty(\mathbb{R}^3)$ 。同理可得，它的外微分为

$$d\eta = \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) dx \wedge dy \wedge dz.$$

1-形式的集合 $\Omega^1(\mathbb{R}^3)$ 与2-形式的集合 $\Omega^2(\mathbb{R}^3)$ 均与 \mathbb{R}^3 上的向量场的集合 $\mathcal{X}(\mathbb{R}^3)$ 同构，对应的同构映射为

$$\begin{aligned} F : \Omega^1(\mathbb{R}^3) &\rightarrow \mathcal{X}(\mathbb{R}^3), & Pdx + Qdy + Rdz &\mapsto P \frac{\partial}{\partial x} + Q \frac{\partial}{\partial y} + R \frac{\partial}{\partial z}, \\ G : \Omega^2(\mathbb{R}^3) &\rightarrow \mathcal{X}(\mathbb{R}^3), & udy \wedge dz + vdz \wedge dx + wdx \wedge dy &\mapsto u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}. \end{aligned}$$

¹⁴ 符号*表示这一章节可以跳过，建议跳转到第86页的章节2.9。

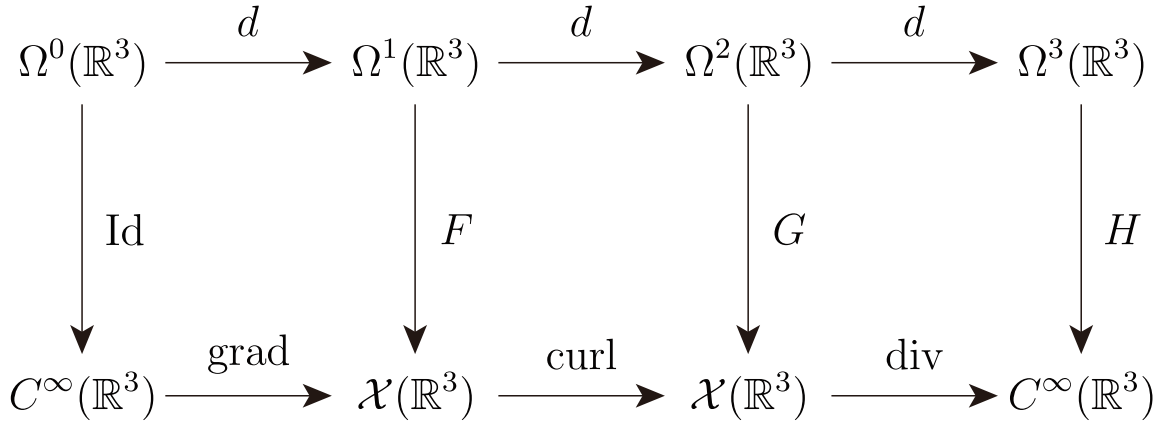


Figure 2.6: \mathbb{R}^3 上的向量与微分形式的关系

0-形式的集合 $\Omega^0(\mathbb{R}^3)$ 本身就是光滑函数的集合 $C^\infty(\mathbb{R}^3)$ 。3-形式可以表示为

$$f dx \wedge dy \wedge dz,$$

其中 $f \in C^\infty(\mathbb{R}^3)$ 。因此 $\Omega^3(\mathbb{R}^3)$ 与 $C^\infty(\mathbb{R}^3)$ 同构，对应的同构映射为

$$H : \Omega^3(\mathbb{R}^3) \rightarrow C^\infty(\mathbb{R}^3), \quad f dx \wedge dy \wedge dz \mapsto f.$$

根据多变量微积分的知识，向量场 $X = P \frac{\partial}{\partial x} + Q \frac{\partial}{\partial y} + R \frac{\partial}{\partial z}$ 的旋度为

$$\text{curl } X = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \frac{\partial}{\partial x} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \frac{\partial}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \frac{\partial}{\partial z},$$

向量场 $Y = u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}$ 的散度为

$$\text{div } Y = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}.$$

因此，我们可以得到交换图Figure 2.6，其中grad代表梯度运算，相当于 ∇f ；Id代表恒映射。根据该交换图，因为 $d \circ d = 0$ ，所以我们可以很容易得出 $\text{div} \circ \text{curl} = 0, \text{curl} \circ \text{grad} = 0$ 。

对于空间中不随时间发生变化的电磁场(视为 \mathbb{R}^3 上的向量场 $E = E^1 \frac{\partial}{\partial x} + E^2 \frac{\partial}{\partial y} + E^3 \frac{\partial}{\partial z}$, $B = B^1 \frac{\partial}{\partial x} + B^2 \frac{\partial}{\partial y} + B^3 \frac{\partial}{\partial z}$)，根据麦克斯韦方程组可得

$$\begin{aligned}
\text{curl } E &= 0, \\
\text{div } B &= 0.
\end{aligned}$$

利用同构关系，将电场视为1-形式

$$E = E^1 dx + E^2 dy + E^3 dz \in \Omega^1(\mathbb{R}^3),$$

将磁场视为2-形式

$$B = B^1 dy \wedge dz + B^2 dz \wedge dx + B^3 dx \wedge dy \in \Omega^2(\mathbb{R}^3),$$

那么 $dE = 0, dB = 0$, 因此 $E \in \Omega^1(\mathbb{R}^3), B \in \Omega^2(\mathbb{R}^3)$ 都是闭形式。根据 Theorem 2.26 可得, \mathbb{R}^3 上任意闭合的1-形式与2-形式都是恰当形式, 因此, 存在光滑函数 $V \in C^\infty(\mathbb{R}^3)$ 与1-形式 $A \in \Omega^1(\mathbb{R}^3)$, 使得

$$\begin{aligned} E &= dV, \\ B &= dA, \end{aligned}$$

我们将 V, A 分别称为**电势**与**磁矢势**。

2.8 哈密顿系统*

¹⁵ 在本章节中, 我们将研究偶数维的向量空间 \mathbb{R}^{2n} , 并将坐标记为 $(x^1, \dots, x^n, y^1, \dots, y^n)$ 。我们定义**辛形式**为 \mathbb{R}^{2n} 上的2-形式

$$\omega = \sum_{i=1}^n dx^i \wedge dy^i.$$

令 $H \in C^\infty(\mathbb{R}^{2n})$ 是光滑函数, 定义**哈密顿向量场**为

$$X_H = \sum_{i=1}^n \frac{\partial H}{\partial y^i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial y^i}.$$

对于光滑向量场 $V \in \mathcal{X}(\mathbb{R}^{2n})$, 如果存在光滑函数 f 使得 $V = X_f$, 那么该向量场是**哈密顿**的。

Theorem 2.27. 对于 \mathbb{R}^{2n} 上的任意光滑向量场 $Y \in \mathcal{X}(\mathbb{R}^{2n})$, 都具有

$$\omega(X_H, Y) = dH(Y) = YH.$$

Proof. 将 Y 表示为分量形式:

$$Y = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i} + b^i \frac{\partial}{\partial y^i},$$

¹⁵ 符号*表示这一章节可以跳过, 建议跳转到第86页的章节2.9。

其中 $a^i, b^i \in C^\infty(\mathbb{R}^{2n})$ 。根据(2.32)可得

$$\begin{aligned}\omega(X_H, Y) &= \sum_{i=1}^n dx^i(X_H) dy^i(Y) - dx^i(Y) dy^i(X_H) \\ &= \sum_{i=1}^n (X_H x^i) (Y y^i) - (Y x^i) (X_H y^i) \\ &= \sum_{i=1}^n a^i \frac{\partial f}{\partial x^i} + b^i \frac{\partial f}{\partial y^i} = YH.\end{aligned}$$

□

我们将 \mathbb{R}^{2n} , 辛形式 ω 与光滑函数 H 三者 $(\mathbb{R}^{2n}, \omega, H)$ 称为**哈密顿系统**。将 H 称为哈密顿系统的**哈密顿量**；向量场 X_H 的流被称为**哈密顿流**， X_H 的积分曲线被称为哈密顿系统的**轨迹**。记轨迹为 $\gamma(t) = (x(t), y(t))$ ，根据积分曲线的定义可得

$$\begin{aligned}\dot{x}^i(t) &= \frac{\partial H}{\partial y^i}(x(t), y(t)), \\ \dot{y}^i(t) &= -\frac{\partial H}{\partial x^i}(x(t), y(t)).\end{aligned}$$

该方程组被称为**哈密顿方程**。

我们将通过经典力学的实例来体现哈密顿系统的具体应用。考虑 n 个物体在三维空间 \mathbb{R}^3 运动，其质量分别为 m_1, \dots, m_n 。为简便起见，我们将每个物体都视为 \mathbb{R}^3 上的一个点，记其任意时刻 t 的坐标为 $\mathbf{q}_1(t), \dots, \mathbf{q}_n(t)$ ，写成分量形式为 $\mathbf{q}_k(t) = (q_k^1(t), q_k^2(t), q_k^3(t))$ 。我们将整个体系的演化视为 \mathbb{R}^{3n} 的轨迹，其含义为 n 个物体在任意时刻 t 的坐标：

$$q(t) = (q_1^1(t), q_1^2(t), q_1^3(t), \dots, q_n^1(t), q_n^2(t), q_n^3(t)).$$

假设每个物体受到的作用力都只与所有物体的坐标有关(例如万有引力)，那我们可以将第 k 个粒子受到的作用力表示为 $\mathbf{F}_k(q) = (F_k^1(q), F_k^2(q), F_k^3(q))$ 。因此，根据牛顿第二定律可得，对于任意 $k \in \{1, \dots, n\}$ ，都具有

$$\begin{aligned}m_k \ddot{q}_k^1(t) &= F_k^1(q(t)), \\ m_k \ddot{q}_k^2(t) &= F_k^2(q(t)), \\ m_k \ddot{q}_k^3(t) &= F_k^3(q(t)).\end{aligned}$$

可以将方程组通过矩阵形式进一步化简。假设 $M = (M_{ij})$ 是 $3n \times 3n$ 的对角矩阵，其对角元素为 $(m_1, m_1, m_1, m_2, m_2, m_2, \dots, m_n, m_n, m_n)$ ；将 n 个物体的坐标表示为 $3n$ 维的向量 $q(t) = (q^1(t), \dots, q^{3n}(t))$ ，受到的力表示为 $3n$ 维的向量 $F(q) = (F_1(q), \dots, F_{3n}(q))$ ，其中 $q^{3k-2} = q_k^1, q^{3k-1} = q_k^2, q^{3k} = q_k^3$ ， $F_{3k-2} = F_k^1, F_{3k-1} = F_k^2, F_{3k} = F_k^3$ 。那么该方程组可以表示为

$$\sum_{j=1}^{3n} M_{ij} \ddot{q}^j(t) = F_i(q(t)).$$

而第 k 个粒子的动量与其坐标关系为 $\mathbf{p}_k(t) = (p_k^1(t), p_k^2(t), p_k^3(t)) = m_k \dot{\mathbf{q}}_k(t)$, 将 n 个物体的动量表示为 $3n$ 维的向量 $p(t) = (p_1(t), \dots, p_{3n}(t))$, 其中 $p^{3k-2} = p_k^1, p^{3k-1} = p_k^2, p^{3k} = p_k^3$ 。因此

$$p_i(t) = \sum_{j=1}^{3n} M_{ij} \dot{q}^j(t),$$

对时间微分可得

$$\dot{p}_i(t) = \sum_{j=1}^{3n} M_{ij} \ddot{q}^j(t) = F_i(q(t)).$$

我们将进一步假设作用力是保守力, 即存在势能函数 $V \in C^\infty(\mathbb{R}^{3n})$ 使得

$$F_i(q) = -\frac{\partial V}{\partial q^i}(q),$$

因此

$$\dot{p}_i(t) = -\frac{\partial V}{\partial q^i}(q(t)).$$

令 (M^{ij}) 是 (M_{ij}) 的逆矩阵, 即 $(M^{ij}) = \text{diag}(m_1^{-1}, m_1^{-1}, m_1^{-1}, \dots, m_n^{-1}, m_n^{-1}, m_n^{-1})$ 。那么我们将 $\sum_{j=1}^{3n} M_{ij} \ddot{q}^j(t) = F_i(q(t))$ 改写为

$$\ddot{q}^i(t) = \sum_{j=1}^{3n} M^{ij} p_j(t).$$

因此

$$\begin{aligned} \ddot{q}^i(t) &= \sum_{j=1}^{3n} M^{ij} p_j(t), \\ \dot{p}_i(t) &= -\frac{\partial V}{\partial q^i}(q(t)). \end{aligned}$$

定义体系的动能为

$$K(p) = \sum_{i,j=1}^{3n} \frac{1}{2} M^{ij} p_i p_j,$$

定义体系的总能量为

$$E(q, p) = V(q) + K(p),$$

因此

$$\begin{aligned} \dot{q}^i(t) &= \frac{\partial E}{\partial p_i}(q(t), p(t)), \\ \dot{p}_i(t) &= -\frac{\partial E}{\partial q^i}(q(t), p(t)). \end{aligned}$$

这与哈密顿方程的形式相同！其对应的哈密顿系统为 $(\mathbb{R}^{6n}, \omega, E)$ ，此时辛形式为

$$\omega = \sum_{i=1}^{3n} dq^i \wedge dp_i.$$

该哈密顿系统的轨迹 $(q(t), p(t))$ 反映了 n 个粒子位置、动量随时间的演化。

2.8.1 守恒与对称*

¹⁶ 对于光滑函数 $f, g \in C^\infty(\mathbb{R}^{2n})$ ，定义泊松括号为

$$\begin{aligned} \{f, g\} &= \omega(X_f, X_g) = df(X_g) = X_g f \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial y^i} - \frac{\partial f}{\partial y^i} \frac{\partial g}{\partial x^i}. \end{aligned} \quad (2.39)$$

泊松括号具有以下性质：

Theorem 2.28. 对于任意光滑函数 $f, g, h \in C^\infty(\mathbb{R}^{2n})$ ，都具有

(A) 对于任意 $\alpha, \beta \in \mathbb{R}$ ，都具有

$$\begin{aligned} \{\alpha f + \beta g, h\} &= \alpha \{f, h\} + \beta \{g, h\}, \\ \{f, \alpha g + \beta h\} &= \alpha \{f, g\} + \beta \{f, h\}. \end{aligned}$$

(B) $\{f, g\} = -\{g, f\}$.

(C) $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.

因此，向量空间 $C^\infty(\mathbb{R}^{2n})$ 在泊松括号下是一个李代数。

Proof. 前2个结论根据定义是显然的。第3个结论，直接计算即可，但是较为繁琐，因此不做展开。 \square

令 $(\mathbb{R}^{2n}, \omega, H)$ 为哈密顿系统，其哈密顿向量场为 X_H 。如果光滑函数 $f \in C^\infty(\mathbb{R}^{2n})$ 在哈密顿向量场 X_H 的任意积分曲线上取到常数，那么它被称为**守恒量(Conserved quantity)**；对于光滑向量场 $V \in \mathcal{X}(\mathbb{R}^{2n})$ ，如果它是哈密顿的，并且 H 在 V 的任意积分曲线上取到常数，那么该光滑向量场被称为**对称变换(Infinitesimal symmetry)**。

Lemma 2.29. f 是守恒量当且仅当 $\{f, H\} = 0$ ； V 是对称变换当且仅当 V 是哈密顿的，并且 $VH = 0$ 。

Proof. 根据Theorem 2.9，这是显然的。 \square

¹⁶ 符号*表示这一章节可以跳过，建议跳转到第86页的章节2.9。

诺特在1918年证明了以下定理(实际上是一个变式), 它揭示了守恒量与对称变换之间的关系, 对理论物理学产生了极其深远的影响:

Theorem 2.30 (Noether's Theorem). 令 $(\mathbb{R}^{2n}, \omega, H)$ 为哈密顿系统。如果光滑函数 $f \in C^\infty(\mathbb{R}^{2n})$ 是守恒量, 那么, 其哈密顿向量场 X_f 是对称变换。此外, 任意对称变换 V 都对应一个守恒量的哈密顿向量场, 而且具有唯一性: 如果 f, \tilde{f} 是守恒量并且满足 $X_f = X_{\tilde{f}} = V$, 那么 $f - \tilde{f} \equiv C$, 即 $f - \tilde{f}$ 取到常数。

Proof. 我们需要利用Lemma 2.29的结论。如果 $\{f, H\} = 0$, 那么根据(2.39), $X_f H = \{H, f\} = -\{f, H\} = 0$ 。而根据定义, X_f 是哈密顿的, 因此 X_f 是对称变换。

对于对称变换 V , 因为它是哈密顿的, 假设光滑函数 f 满足 $X_f = V$, 那么 $VH = X_f H = 0$ 。而 $X_f H = \{H, f\} = -\{f, H\} = 0$, 因此 $\{f, H\} = 0$, 这表明 f 是守恒量。如果守恒量 f, \tilde{f} 满足 $X_f = X_{\tilde{f}} = V$, 那么对于任意光滑函数 g , 都具有

$$X_f g = X_{\tilde{f}} g.$$

因此

$$\{f, g\} - \{\tilde{f}, g\} = \{f - \tilde{f}, g\} = 0.$$

这表明

$$\sum_{i=1}^n \frac{\partial(f - \tilde{f})}{\partial x^i} \frac{\partial g}{\partial y^i} - \frac{\partial(f - \tilde{f})}{\partial y^i} \frac{\partial g}{\partial x^i} = 0.$$

取 $g(x^1, \dots, x^n, y^1, \dots, y^n) = x^i, g(x^1, \dots, x^n, y^1, \dots, y^n) = y^i$ 可得, 对于任意 $i \in [1, n]$, 都具有 $\frac{\partial(f - \tilde{f})}{\partial x^i} = 0, \frac{\partial(f - \tilde{f})}{\partial y^i} = 0$ 。因此, $f - \tilde{f} \equiv C$ 。□

现在, 我们将只分析2个物体的情况, 并且假设势能函数只与2个物体之间的距离有关, 即势能函数可以被表示为 $V(q) = v(r(q))$, 其中 $v: (0, \infty) \rightarrow \mathbb{R}$ 是光滑函数,

$$\begin{aligned} r(q) &= |\mathbf{q}_1 - \mathbf{q}_2| = \sqrt{(q_1^1 - q_2^1)^2 + (q_1^2 - q_2^2)^2 + (q_1^3 - q_2^3)^2} \\ &= \sqrt{(q^1 - q^4)^2 + (q^2 - q^5)^2 + (q^3 - q^6)^2} \end{aligned}$$

是2个物体之间的距离。因此, 哈密顿量为

$$\begin{aligned} H(q, p) &= V(q) + K(p) = v(r(q)) + \sum_{i,j=1}^6 \frac{1}{2} M^{ij} p_i p_j, \\ &= v(r(q)) + \frac{(p_1)^2 + (p_2)^2 + (p_3)^2}{2m_1} + \frac{(p_4)^2 + (p_5)^2 + (p_6)^2}{2m_2}. \end{aligned}$$

(A) 定义总能量为哈密顿量本身

$$H(q, p).$$

而 $\{H, H\} = 0$ ，因此 H 是守恒量，这就是**能量守恒定律**。而 $X_H H = \{H, H\} = 0$ ，根据定义， X_H 显然是哈密顿的，因此 X_H 是对称变换，这符合诺特定理。而哈密顿流反映了体系随时间的演化，因此，通俗来说，能量守恒对应的对称性是时间的演化，如Figure 2.7所示。

(B) 令 $\mathbf{u} = (u^1, u^2, u^3)$ 是 \mathbb{R}^3 上的单位向量，即 $|\mathbf{u}| = 1$ 。定义 \mathbf{u} 方向的总动量为

$$\begin{aligned} P(q, p) &= \mathbf{u} \cdot \mathbf{p}_1 + \mathbf{u} \cdot \mathbf{p}_2 \\ &= u^1 p_1^1 + u^2 p_1^2 + u^3 p_1^3 + u^1 p_2^1 + u^2 p_2^2 + u^3 p_2^3 \\ &= u^1 p_1 + u^2 p_2 + u^3 p_3 + u^1 p_4 + u^2 p_5 + u^3 p_6, \end{aligned}$$

因为 P 与 q 无关，所以

$$\begin{aligned} \{P, H\} &= \sum_{i=1}^6 \frac{\partial P}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial P}{\partial p_i} \frac{\partial H}{\partial q^i} = - \sum_{i=1}^6 \frac{\partial P}{\partial p_i} \frac{\partial H}{\partial q^i} \\ &= -u^1 \frac{\partial V}{\partial q^1} - u^2 \frac{\partial V}{\partial q^2} - u^3 \frac{\partial V}{\partial q^3} - u^1 \frac{\partial V}{\partial q^4} - u^2 \frac{\partial V}{\partial q^5} - u^3 \frac{\partial V}{\partial q^6} \\ &= -u^1 \left(\frac{\partial V}{\partial q^1} + \frac{\partial V}{\partial q^4} \right) - u^2 \left(\frac{\partial V}{\partial q^2} + \frac{\partial V}{\partial q^5} \right) - u^3 \left(\frac{\partial V}{\partial q^3} + \frac{\partial V}{\partial q^6} \right). \end{aligned}$$

其中

$$\begin{aligned} \frac{\partial V}{\partial q^1} + \frac{\partial V}{\partial q^4} &= v'(r(q)) \left(\frac{\partial r}{\partial q^1} + \frac{\partial r}{\partial q^4} \right) \\ &= v'(r(q)) \left(\frac{q^1 - q^4}{r(q)} + \frac{-(q^1 - q^4)}{r(q)} \right) \\ &= 0, \end{aligned}$$

后2项同样为0，因此 $\{P, H\} = 0$ ，表明任意方向 \mathbf{u} 的总动量 P 是守恒量，这就是**动量守恒定律**。

根据诺特定理， X_P 是对称变换，对其进行计算可得：

$$\begin{aligned} X_P &= \sum_{i=1}^6 \frac{\partial P}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial P}{\partial q^i} \frac{\partial}{\partial p_i} = \sum_{i=1}^6 \frac{\partial P}{\partial p_i} \frac{\partial}{\partial q^i} \\ &= u^1 \frac{\partial}{\partial q^1} + u^2 \frac{\partial}{\partial q^2} + u^3 \frac{\partial}{\partial q^3} + u^1 \frac{\partial}{\partial q^4} + u^2 \frac{\partial}{\partial q^5} + u^3 \frac{\partial}{\partial q^6}, \end{aligned}$$

其积分曲线 $\gamma(t) = (q^1(t), \dots, q^6(t), p_1(t), \dots, p_6(t))$ 满足

$$\begin{aligned} \dot{q}^1(t) &= \dot{q}^4(t) = u^1, \\ \dot{q}^2(t) &= \dot{q}^5(t) = u^2, \\ \dot{q}^3(t) &= \dot{q}^6(t) = u^3, \\ \dot{p}_1(t) &= \dots = \dot{p}_6(t) = 0. \end{aligned}$$

因此，向量场 X_P 的流为

$$\begin{aligned} \theta_t(q^1, \dots, q^6, p_1, \dots, p_6) &= (q^1 + tu^1, q^2 + tu^2, q^3 + tu^3, q^4 + tu^1, q^5 + tu^2, q^6 + tu^3, \\ &\quad p_1, p_2, p_3, p_4, p_5, p_6). \end{aligned}$$

这表明动量守恒对应的对称性是坐标空间发生“平移”，动量空间“保持不变”，如Figure 2.7所示。

(C) 定义 z 方向的角动量为

$$\begin{aligned} L(q, p) &= q_1^1 p_1^2 - q_1^2 p_1^1 + q_2^1 p_2^2 - q_2^2 p_2^1 \\ &= q^1 p_2 - q^2 p_1 + q^4 p_5 - q^5 p_4, \end{aligned}$$

因此

$$\begin{aligned} \{L, H\} &= \sum_{i=1}^6 \frac{\partial L}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial L}{\partial p_i} \frac{\partial H}{\partial q^i} \\ &= \left(p_2 \frac{p_1}{m_1} - p_1 \frac{p_2}{m_1} + p_5 \frac{p_4}{m_2} - p_4 \frac{p_5}{m_2} \right) - \sum_{i=1}^6 \frac{\partial L}{\partial p_i} \frac{\partial H}{\partial q^i} \\ &= - \sum_{i=1}^6 \frac{\partial L}{\partial p_i} \frac{\partial H}{\partial q^i} = q^2 \frac{\partial V}{\partial q^1} - q^1 \frac{\partial V}{\partial q^2} + q^5 \frac{\partial V}{\partial q^4} - q^4 \frac{\partial V}{\partial q^5} \\ &= q^2 v'(r(q)) \frac{\partial r}{\partial q^1} - q^1 v'(r(q)) \frac{\partial r}{\partial q^2} + q^5 v'(r(q)) \frac{\partial r}{\partial q^4} - q^4 v'(r(q)) \frac{\partial r}{\partial q^5} \\ &= v'(r(q)) \left(q^2 \frac{q^1 - q^4}{r(q)} - q^1 \frac{q^2 - q^5}{r(q)} + q^5 \frac{-(q^1 - q^4)}{r(q)} - q^4 \frac{-(q^2 - q^5)}{r(q)} \right) \\ &= \frac{v'(r(q))}{r(q)} (q^2(q^1 - q^4) - q^1(q^2 - q^5) - q^5(q^1 - q^4) + q^4(q^2 - q^5)) \\ &= 0, \end{aligned}$$

这证明了 L 是守恒量，这就是**角动量守恒定律**。

根据诺特定理, X_L 是对称变换, 对其进行计算可得:

$$\begin{aligned} X_L &= \sum_{i=1}^6 \frac{\partial L}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial L}{\partial q^i} \frac{\partial}{\partial p_i} \\ &= \left(-q^2 \frac{\partial}{\partial q^1} + q^1 \frac{\partial}{\partial q^2} - q^5 \frac{\partial}{\partial q^4} + q^4 \frac{\partial}{\partial q^5} \right) - \left(p_2 \frac{\partial}{\partial p_1} - p_1 \frac{\partial}{\partial p_2} + p_5 \frac{\partial}{\partial p_4} - p_4 \frac{\partial}{\partial p_5} \right), \end{aligned}$$

其积分曲线 $\gamma(t) = (q^1(t), \dots, q^6(t), p_1(t), \dots, p_6(t))$ 满足

$$\begin{aligned} \dot{q}^1(t) &= -q^2(t), & \dot{q}^2(t) &= q^1(t), \\ \dot{q}^4(t) &= -q^5(t), & \dot{q}^5(t) &= q^4(t), \\ \dot{q}^3(t) &= \dot{q}^6(t) = 0, \\ \dot{p}_1(t) &= -p_2(t), & \dot{p}_2(t) &= p_1(t), \\ \dot{p}_4(t) &= -p_5(t), & \dot{p}_5(t) &= p_4(t), \\ \dot{p}_3(t) &= \dot{p}_6(t) = 0. \end{aligned}$$

因此, 向量场 X_L 的流为

$$\begin{aligned} \theta_t(q^1, \dots, q^6, p_1, \dots, p_6) &= (q^1 \cos t - q^2 \sin t, q^1 \sin t + q^2 \cos t, q^3, \\ &\quad q^4 \cos t - q^5 \sin t, q^4 \sin t + q^5 \cos t, q^6, \\ &\quad p_1 \cos t - p_2 \sin t, p_1 \sin t + p_2 \cos t, p_3, \\ &\quad p_4 \cos t - p_5 \sin t, p_4 \sin t + p_5 \cos t, p_6). \end{aligned}$$

这表明角动量守恒对应的对称性是坐标空间与动量空间均发生“旋转”, 如Figure 2.7所示。

Figure 2.7大致反映了守恒量 H, P, L 对应的对称性。

2.9 狄拉克符号

狄拉克符号, 可以将复数域 \mathbb{C} 上的内积空间(定义内积运算的向量空间, 满足性质(1.10)) V 的任意向量 Ψ 都通过右矢 $|\Psi\rangle$ 表示。记它的左矢为 $\langle\Psi|$, 它是对偶空间 V^* 上的一个向量, 其运算规则满足

$$\langle\Psi|(|\Phi\rangle) = \langle\Psi|\Phi\rangle,$$

其中, $\langle\Psi|\Phi\rangle$ 是向量 $|\Psi\rangle, |\Phi\rangle$ 之间的内积。

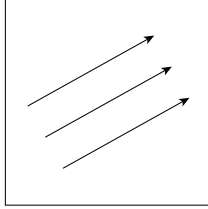
因此, 如果选取 V 上的一组标准正交基 $|1\rangle, \dots, |n\rangle$ (在之后的篇幅中, 我们默认形式为 $|i\rangle, i \in \{1, \dots, n\}$ 的向量是 V 上的一组标准正交基), 那么

$$\langle i|j\rangle = \delta_j^i. \quad (2.40)$$

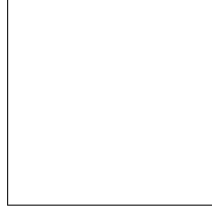
根据Theorem 2.12可得, $\langle 1|, \dots, \langle n|$ 也是对偶空间 V^* 上的一组基。

能量守恒对应的对称性是时间的演化 过去  将来

动量守恒对应的对称性是
坐标空间平移，动量空间不变

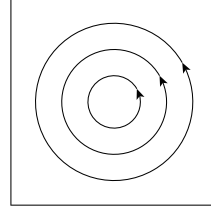


坐标演化

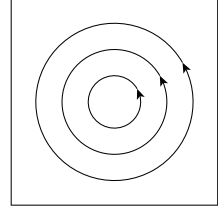


动量演化

角动量守恒对应的对称性是
坐标空间旋转，动量空间旋转



坐标演化



动量演化

Figure 2.7: 守恒与对称

根据线性代数的知识，我们可以将 V 上任意向量 $|\Psi\rangle$ 展开为

$$|\Psi\rangle = \sum_{i=1}^n \Psi^i |i\rangle, \quad (2.41)$$

我们可以将它视为列向量

$$\begin{pmatrix} \Psi^1 \\ \vdots \\ \Psi^n \end{pmatrix},$$

其中 $\Psi^i \in \mathbb{C}$ 。因为 $\langle\Psi|$ 是对偶空间 V^* 上的向量， $\langle 1|, \dots, \langle n|$ 也是对偶空间 V^* 上的一组基，所以它也可以被展开为

$$\langle\Psi| = \sum_{i=1}^n \Psi_i \langle i|. \quad (2.42)$$

我们可以将它视为行向量

$$(\Psi_1 \quad \dots \quad \Psi_n),$$

其中 $\Psi_i \in \mathbb{C}$ 。根据(2.24), (2.25)与(2.40)可得

$$\langle\Psi|\Phi\rangle = \sum_{i,j} \Psi_i \Phi_j \langle i|j\rangle = \sum_{i,j} \Psi_i \Phi_j \delta_j^i = \sum_{i=1}^n \Psi_i \Phi^i, \quad (2.43)$$

而 $\langle\Psi|\Phi\rangle$ 又是 $|\Psi\rangle, |\Phi\rangle$ 的内积，因此

$$\langle\Psi|\Phi\rangle = \sum_{i=1}^n \overline{\Psi^i} \Phi^i. \quad (2.44)$$

(注意, 复空间上的内积是 $\Psi^\dagger \Phi$, 而不是 $\Psi^T \Phi$ 。其中 \dagger 代表**共轭转置**, 相当于将矩阵或向量进行转置之后, 再取所有元素的共轭复数。)根据 Φ^i 的任意性可得

$$\Psi_i = \overline{\Psi^i}. \quad (2.45)$$

例如 $|\Psi\rangle = i|1\rangle + (1+i)|2\rangle + |3\rangle$, 那么根据(2.42)与(2.45)可得

$$\langle\Psi| = \sum_{i=1}^3 \Psi_i \langle i| = \bar{i} \langle 1| + \overline{1+i} \langle 2| + \bar{1} \langle 3| = -i \langle 1| + (1-i) \langle 2| + \langle 3|.$$

根据(2.44), 我们就可以得出以下结论: 对于任意向量

$$\begin{aligned} |\Psi\rangle &= \sum_{i=1}^n \Psi^i |i\rangle, \\ |\Phi\rangle &= \sum_{i=1}^n \Phi^i |i\rangle, \end{aligned}$$

都具有

$$\langle\Phi|\Psi\rangle = \overline{\langle\Psi|\Phi\rangle}. \quad (2.46)$$

因此, 我们可以根据(2.24), (2.25), (2.46)与内积的性质(1.10), 得到以下结论:

$$\begin{aligned} \langle\Phi|(\alpha|\Psi\rangle + \beta|\Theta\rangle) &= \alpha\langle\Phi|\Psi\rangle + \beta\langle\Phi|\Theta\rangle, \\ (\alpha\langle\Phi| + \beta\langle\Theta|)|\Psi\rangle &= \alpha\langle\Phi|\Psi\rangle + \beta\langle\Theta|\Psi\rangle, \\ \langle\Phi|\Psi\rangle &= \overline{\langle\Psi|\Phi\rangle}, \\ \langle\Psi|\Psi\rangle &\geq 0. \end{aligned} \quad (2.47)$$

其中 $\alpha, \beta \in \mathbb{C}$, 第4个表达式取等号当且仅当 $\Psi = 0$ 。

2.9.1 线性映射的狄拉克符号表示

如果将向量 $|\Psi\rangle \in V$ 与对偶向量 $\langle\Phi| \in V^*$ 并列, 那会得到什么结果呢? 我们将它记为 $|\Psi\rangle\langle\Phi|$ 。它的作用是, 将一个向量变成另一个向量:

$$|\Psi\rangle\langle\Phi|(|\Theta\rangle) = |\Psi\rangle\langle\Phi|\Theta\rangle = \langle\Phi|\Theta\rangle|\Psi\rangle.$$

(因为 $\langle\Phi|\Theta\rangle$ 是一个标量, 所以可以被提取出来)很显然, $|\Psi\rangle\langle\Phi|$ 是线性映射。这与Theorem 2.19的结论是不谋而合的!

Theorem 2.31. 令 $|\Psi\rangle, |\Phi\rangle$ 展开的形式为

$$\begin{aligned} |\Psi\rangle &= \sum_{i=1}^n \Psi^i |i\rangle, \\ |\Phi\rangle &= \sum_{j=1}^n \Phi^j |j\rangle, \end{aligned}$$

其中 $\Psi^i, \Phi^j \in \mathbb{C}$ 。那么

$$|\Psi\rangle \langle \Phi| = \sum_{i,j} \Psi^i \Phi_j^* |i\rangle \langle j|. \quad (2.48)$$

Proof. 对于任意 $|\Theta\rangle = \sum_{k=1}^n \Theta^k |k\rangle$, 根据(2.43)可得

$$\begin{aligned} |\Psi\rangle \langle \Phi| \Theta\rangle &= \langle \Phi| \Theta\rangle |\Psi\rangle \\ &= \left(\sum_{i=1}^n \Phi_i \Theta^i \right) \sum_{j=1}^n \Psi^j |j\rangle \\ &= \sum_{i,j} \Psi^j \Phi_i \Theta^i |j\rangle. \end{aligned}$$

而

$$\begin{aligned} \left(\sum_{i,j} \Psi^i \Phi_j |i\rangle \langle j| \right) |\Theta\rangle &= \sum_{i,j,k} \Psi^i \Phi_j \Theta^k |i\rangle \langle j|k\rangle \\ &= \sum_{i,j,k} \Psi^i \Phi_j \Theta^k \langle j|k\rangle |i\rangle \\ &= \sum_{i,j,k} \Psi^i \Phi_j \Theta^k \delta_k^j |i\rangle \\ &= \sum_{i,j} \Psi^i \Phi_j \Theta^j |i\rangle. \end{aligned}$$

交换 i, j 即可完成证明。 □

Theorem 2.32. 任意线性映射 A 都可以表示为以下形式:

$$A = \sum_{i,j} A_j^i |i\rangle \langle j|,$$

其中

$$A_j^i = \langle i|A|j\rangle.$$

因此，任意线性映射都可以视为一个 n 阶矩阵，其第 i 行，第 j 列的元素为 A_j^i ，但是它的取值并不是固定的，取决于标准正交基的选取。我们将这个行为称为**矩阵表示**。可以将它视为矩阵

$$\begin{pmatrix} A_1^1 & \cdots & A_n^1 \\ \vdots & \ddots & \vdots \\ A_1^n & \cdots & A_n^n \end{pmatrix}.$$

恒映射记为 I ，它将任意向量映射到它本身，具体来说，对于任意向量 $|\Psi\rangle$ ，都具有 $I|\Psi\rangle \equiv \Psi$ 。而 $\langle i|I|j\rangle = \langle i|j\rangle = \delta_j^i$ ，因此

$$I = \sum_{i=1}^n |i\rangle \langle i|.$$

它的矩阵表示是单位矩阵，当然，这并不难猜测。

因此，根据(2.48)， $|\Psi\rangle \langle \Phi|$ 的矩阵表示为

$$\begin{pmatrix} \Psi^1 \\ \vdots \\ \Psi^n \end{pmatrix} (\Phi_1 \quad \cdots \quad \Phi_n) = \begin{pmatrix} \Psi^1 \Phi_1 & \cdots & \Psi^1 \Phi_n \\ \vdots & \ddots & \vdots \\ \Psi^n \Phi_1 & \cdots & \Psi^n \Phi_n \end{pmatrix}.$$

Proof. 因为 $A|j\rangle$ 是一个向量，所以可以展开为以下形式：

$$A|j\rangle = \sum_{k=1}^n A_j^k |k\rangle, \quad (2.49)$$

将 $\langle i|$ 作用于等式的两边，可得

$$\langle i|A|j\rangle = \sum_{k=1}^n A_j^k \langle i|k\rangle = \sum_{k=1}^n A_j^k \delta_k^i = A_j^i.$$

根据(2.41)，记 $|\Psi\rangle = \sum_j \Psi^j |j\rangle$ ，因此，根据(2.49)可得

$$A|\Psi\rangle = \sum_{j=1}^n \Psi^j A|j\rangle = \sum_{j,k} A_j^k \Psi^j |k\rangle. \quad (2.50)$$

而

$$\left(\sum_{i,j} A_j^i |i\rangle \langle j| \right) |\Psi\rangle = \sum_{i,j,k} A_j^i \Psi^k |i\rangle \langle j|k\rangle = \sum_{i,j,k} A_j^i \Psi^k \delta_k^j |i\rangle = \sum_{i,k} A_k^i \Psi^k |i\rangle,$$

所以

$$A|\Psi\rangle = \left(\sum_{i,j} A_j^i |i\rangle \langle j| \right) |\Psi\rangle.$$

根据 $|\Psi\rangle$ 的任意性可得

$$A = \sum_{i,j} A_j^i |i\rangle \langle j|.$$

□

Corollary 2.32.1. 形式为 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$ 的线性映射作用于形式为 $|\Psi\rangle = \sum_k \Psi^k |k\rangle$ 的向量上，得到的结果为

$$A|\Psi\rangle = \sum_{i,j} A_j^i \Psi^j |i\rangle. \quad (2.51)$$

因此，线性算子作用于向量可以视为对应的矩阵与对应的向量相乘！ $A|\Psi\rangle$ 的矩阵表示为

$$\begin{pmatrix} A_1^1 & \cdots & A_n^1 \\ \vdots & \ddots & \vdots \\ A_1^n & \cdots & A_n^n \end{pmatrix} \begin{pmatrix} \Psi^1 \\ \vdots \\ \Psi^n \end{pmatrix}.$$

Proof. 请参考(2.50)。

□

Theorem 2.33. 对于形式为 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$, $B = \sum_{k,l} B_l^k |k\rangle \langle l|$ 的线性映射， AB 则可以表示为

$$AB = \sum_{i,j,l} A_j^i B_l^j |i\rangle \langle l|. \quad (2.52)$$

因此， AB 对应的矩阵可以视为 A 对应的矩阵与 B 对应的矩阵相乘！ AB 的矩阵表示为

$$\begin{pmatrix} A_1^1 & \cdots & A_n^1 \\ \vdots & \ddots & \vdots \\ A_1^n & \cdots & A_n^n \end{pmatrix} \begin{pmatrix} B_1^1 & \cdots & B_n^1 \\ \vdots & \ddots & \vdots \\ B_1^n & \cdots & B_n^n \end{pmatrix}.$$

Proof. 记 $|\Psi\rangle = \sum_i \Psi^i |i\rangle$ ，因为 $AB|\Psi\rangle = A(B|\Psi\rangle)$ ，利用(2.49)可得

$$\begin{aligned} AB|\Psi\rangle &= \sum_{k,l} A \left(B_l^k \Psi^l |k\rangle \right) \\ &= \sum_{k,l} B_l^k \Psi^l A|k\rangle \\ &= \sum_{i,k,l} A_k^i B_l^k \Psi^l |i\rangle, \end{aligned}$$

而

$$\left(\sum_{i,j,l} A_j^i B_l^j |i\rangle \langle l| \right) |\Psi\rangle = \sum_{i,j,l} A_j^i B_l^j \Psi^l |i\rangle = AB|\Psi\rangle,$$

根据 $|\Psi\rangle$ 的任意性，完成证明。

□

类比线性代数，我们就可以很容易地得出：

$$(AB)C = A(BC).$$

虽然该定理的证明在数学上是没有问题的，但是它略显复杂，因为我们借助了向量 $|\Psi\rangle$ ，并且这个证明过程仿佛已经知道了 $AB = \sum_{i,j,l} A_j^i B_l^j |i\rangle \langle l|$ 这个事实，如同“自圆其说”一般。那么，有什么办法，可以更加方便地去计算 AB 呢？

可以看到，如果我们尝试将 AB “拼在一起”，那么就会得到：

$$AB = \sum_{i,j,k,l} A_j^i B_l^k (|i\rangle \langle j|) (|k\rangle \langle l|),$$

是不是很熟悉呢？可能你已经想到了，将中间的 $\langle j|$ 与 $|k\rangle$ “拼在一起”，得到 $\langle j|k\rangle = \delta_k^j$ ，那我们就试试看这样做是否可行：

$$AB = \sum_{i,j,k,l} A_j^i B_l^k \delta_k^j |i\rangle \langle l| = \sum_{i,j,l} A_j^i B_l^j |i\rangle \langle l|,$$

这与(2.52)的结果刚好相等！这样，你就完全没有必要特意去区分向量与线性映射的计算规则，直接将对偶向量与向量“拼在一起”即可，如果产生了 $\langle \Psi | \Phi \rangle$ 这样的标量，再把它提取出来，这样就能缩短表达式的长度了！但是要当心，你不能随意交换左右矢的排列次序，正如矩阵相乘，你不能随意交换矩阵的排列次序一样。例如，

$$\begin{aligned} (|\alpha\rangle \langle \beta|) (|\gamma\rangle \langle \delta|) (|\theta\rangle \langle \psi|) &= |\alpha\rangle \langle \beta|\gamma\rangle \langle \delta|\theta\rangle \langle \psi| \\ &= \langle \beta|\gamma\rangle \langle \delta|\theta\rangle |\alpha\rangle \langle \psi|. \end{aligned}$$

第一个等号就是“拼在一起”，第二个等号就是把标量 $\langle \beta|\gamma\rangle, \langle \delta|\theta\rangle$ 提取出来。所以，狄拉克符号强大之处就体现在这里，它非常的直观！

在线性代数理论中，线性映射都是“向右作用”的，即 $(AB)v = A(Bv)$ ，我们首先需要计算 Bv ，再计算 $A(Bv)$ 。不过有些时候，我们希望线性映射“向左作用”，因为在某些情况下向左作用的比向右作用更加容易计算。此时需要引入**伴随**的概念。 V 上的线性映射 A 的伴随同样也是 V 上的线性映射，记为 A^\dagger 。它满足以下关系：对于任意向量 $|\Psi\rangle, |\Phi\rangle$ ，都具有

$$\langle \Psi | A | \Phi \rangle = \overline{\langle \Phi | A^\dagger | \Psi \rangle}. \quad (2.53)$$

因此

$$\langle \Psi | A | \Phi \rangle = \overline{\langle \Phi | A^\dagger | \Psi \rangle} = \overline{\overline{\langle \Psi | (A^\dagger)^\dagger | \Phi \rangle}} = \langle \Psi | (A^\dagger)^\dagger | \Phi \rangle,$$

根据 $|\Psi\rangle, |\Phi\rangle$ 的任意性可得

$$A = (A^\dagger)^\dagger.$$

当 $A = A^\dagger$ 时， A 被称为**自伴随**的。从上面可以看出，如果能求出线性映射的伴随，那么就可以实现“向左作用”，如果该线性映射是自伴随的，那情况就更为简单了！

以下定理虽然没有直接求出线性映射的伴随，但是它可以提供一种较为简单的计算思路，并且具有一定的直观性。

Theorem 2.34. 形式为 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$ 的线性映射，当处于对偶向量的右方时，可以利用其 $|i\rangle$ 的部分直接向左作用于对偶向量。

换言之，如果向右作用，那么

$$\langle \Psi | \left(\sum_{i,j} A_j^i |i\rangle \langle j| \right) | \Phi \rangle = \langle \Psi | \left(\sum_{i,j} A_j^i |i\rangle \langle j| \Phi \rangle \right) = \sum_{i,j} A_j^i \langle \Psi | i \rangle \langle j | \Phi \rangle,$$

如果向左作用，那么

$$\langle \Psi | \left(\sum_{i,j} A_j^i |i\rangle \langle j| \right) | \Phi \rangle = \left(\sum_{i,j} A_j^i \langle \Psi | i \rangle \langle j| \right) | \Phi \rangle = \sum_{i,j} A_j^i \langle \Psi | i \rangle \langle j | \Phi \rangle,$$

Proof. 根据(2.41)，记 $|\Psi\rangle = \sum_i \Psi_i |i\rangle$, $|\Phi\rangle = \sum_j \Phi_j |j\rangle$ ，因此

$$\begin{aligned} \langle \Psi | \left(\sum_{i,j} A_j^i |i\rangle \langle j| \Phi \rangle \right) &= \langle \Psi | \left(\sum_{i,j} A_j^i \Phi_j |i\rangle \right) = \sum_{i,j} A_j^i \Psi_i \Phi_j, \\ \left(\sum_{i,j} A_j^i \langle \Psi | i \rangle \langle j| \right) | \Phi \rangle &= \left(\sum_{i,j} A_j^i \Psi_i \langle j| \right) | \Phi \rangle = \sum_{i,j} A_j^i \Psi_i \Phi_j, \end{aligned}$$

所以这两者相等，我们可以直接记为

$$\langle \Psi | A | \Phi \rangle = \sum_{i,j} A_j^i \Psi_i \Phi_j. \quad (2.54)$$

其矩阵表示为

$$(\Psi_1 \quad \cdots \quad \Psi_n) \begin{pmatrix} A_1^1 & \cdots & A_n^1 \\ \vdots & \ddots & \vdots \\ A_1^n & \cdots & A_n^n \end{pmatrix} \begin{pmatrix} \Phi^1 \\ \vdots \\ \Phi^n \end{pmatrix}.$$

□

根据这一定理，我们可以很方便地写出形式为 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$ 的线性映射的伴随。

Corollary 2.34.1.

$$A^\dagger = \sum_{i,j} \overline{A_j^i} |j\rangle \langle i|. \quad (2.55)$$

因此，线性映射的伴随可以视为其对应矩阵的共轭转置(先对任意元素取共轭复数，再对矩阵进行转置)！ A^\dagger 的矩阵表示为

$$\begin{pmatrix} \overline{A_1^1} & \cdots & \overline{A_1^n} \\ \vdots & \ddots & \vdots \\ \overline{A_n^1} & \cdots & \overline{A_n^n} \end{pmatrix}.$$

Proof. 记 $|\Psi\rangle = \sum_i \Psi^i |i\rangle$, $|\Phi\rangle = \sum_j \Phi^j |j\rangle$, 根据(2.53)与(2.54)可得

$$\langle \Psi | A | \Phi \rangle = \sum_{i,j} A_j^i \Psi_i \Phi^j,$$

因此

$$\langle \Phi | A^\dagger | \Psi \rangle = \overline{\langle \Psi | A | \Phi \rangle} = \sum_{i,j} \overline{A_j^i \Psi_i \Phi^j}.$$

令 $B = \sum_{i,j} \overline{A_i^j} |i\rangle \langle j|$, 因此

$$\langle \Phi | B | \Psi \rangle = \sum_{i,j} \overline{A_i^j} \Phi_i \Psi^j = \sum_{i,j} \overline{A_j^i} \Psi^j \Phi_i,$$

根据(2.45)可得

$$\langle \Phi | B | \Psi \rangle = \sum_{i,j} \overline{A_j^i \Psi_j \Phi^i},$$

根据 $|\Psi\rangle, |\Phi\rangle$ 的任意性可得

$$B = A^\dagger.$$

□

根据这一定理, 我们可以很容易得出以下结论:

Corollary 2.34.2. 对于任意 $\lambda \in \mathbb{C}$,

$$\begin{aligned} (A + B)^\dagger &= A^\dagger + B^\dagger, \\ (\lambda A)^\dagger &= \bar{\lambda} A^\dagger. \end{aligned}$$

因此, 伴随运算只在实数域 \mathbb{R} 上是线性的, 在复数域 \mathbb{C} 上不是线性的。

对于向量 $|\Psi\rangle$ 的对偶向量 $\langle \Psi |$ 与线性算子 A , $\langle \Psi | A$ 也是对偶空间上的一个向量, 其运算满足

$$(\langle \Psi | A) | \Phi \rangle = \langle \Psi | A | \Phi \rangle. \quad (2.56)$$

Theorem 2.35. 向量 $A | \Psi \rangle$ 的对偶向量为 $\langle \Psi | A^\dagger$ 。

Proof. 记 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$, $|\Psi\rangle = \sum_k \Psi^k |k\rangle$, 根据(2.51)可得

$$A | \Psi \rangle = \sum_{i,j} A_j^i \Psi^j |i\rangle,$$

记 $A | \Psi \rangle$ 的对偶向量为 $\langle \Theta |$, $\overline{A_j^i}$ 记为 $(A^\dagger)_i^j$, 根据(2.42)与(2.45)可得

$$\langle \Theta | = \sum_{i,j} (A^\dagger)_i^j \Psi_j \langle i|,$$

记任意向量 $|\Phi\rangle = \sum_i \Phi^i |i\rangle$ ，因此

$$\langle\Theta|\Phi\rangle = \sum_{i,j} (A^\dagger)_i^j \Psi_j \Phi^i,$$

而根据(2.55)与(2.56)可得

$$\left(\langle\Psi|A^\dagger\right)|\Phi\rangle = \langle\Psi|A^\dagger|\Phi\rangle = \sum_{i,j} (A^\dagger)_j^i \Psi_i \Phi^j,$$

交换 i, j 可得

$$\langle\Theta|\Phi\rangle = \left(\langle\Psi|A^\dagger\right)|\Phi\rangle.$$

根据 $|\Phi\rangle$ 的任意性可得

$$\langle\Theta| = \langle\Psi|A^\dagger.$$

□

Theorem 2.36.

$$(AB)^\dagger = B^\dagger A^\dagger.$$

这与线性代数公式 $(AB)^\dagger = B^\dagger A^\dagger$ 相对应，当然你也可以类比为转置矩阵之间的关系 $(AB)^T = B^T A^T$ 。

Proof. 将 A, B 表示为 $A = \sum_{i,j} A_j^i |i\rangle \langle j|$, $B = \sum_{k,l} B_l^k |k\rangle \langle l|$ ，可得

$$A^\dagger = \sum_{i,j} (A^\dagger)_j^i |i\rangle \langle j|,$$

$$B^\dagger = \sum_{k,l} (B^\dagger)_l^k |k\rangle \langle l|,$$

因此

$$B^\dagger A^\dagger = \sum_{j,k,l} (B^\dagger)_l^k (A^\dagger)_j^l |k\rangle \langle j| = \sum_{j,k,l} \overline{A_l^j B_k^l} |k\rangle \langle j|,$$

而

$$AB = \sum_{i,j,l} A_j^i B_l^j |i\rangle \langle l|,$$

因此

$$(AB)^\dagger = \sum_{i,j,l} \overline{A_j^l B_i^j} |i\rangle \langle l|,$$

重组上下标即可完成证明。

□

根据以上2个定理，我们可以绘制交换图Figure 2.8。

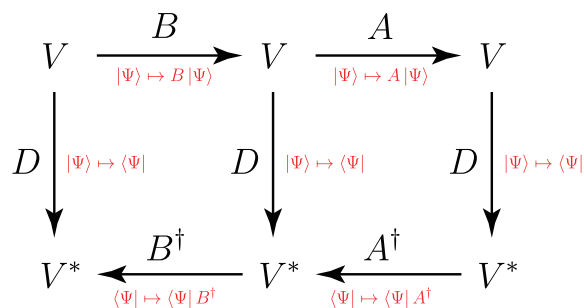


Figure 2.8: 交换图

3 角动量与自旋

注意：除非特殊声明，默认 $x = (x^1, x^2, x^3) \in \mathbb{R}^3$ ，即 x 是 \mathbb{R}^3 上的向量。

人们对于原子的逐步认识，正是物理学一直发展的体现。量子力学理论建立的原因之一，就是科学家发现氢原子的光谱是离散的，而不是连续的。我们在这部分分析这一现象的成因。但是，更加神奇的是，满足氢原子势能分布的波函数同时为3个线性算子的本征态！但是，对应的本征值是不同的，这些本征值对应不同的物理量——氢原子的能量、氢原子的角动量大小、氢原子角动量的 z 分量。

自旋是连接量子力学与量子信息的桥梁，它是一个全新的体系，在该体系下，我们可以将线性算子视为2阶矩阵，将波函数视为2维向量，这样能够极大简化计算！

3.1 角动量算子

利用经典力学角动量公式

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} = (x^2 p^3 - x^3 p^2, x^3 p^1 - x^1 p^3, x^1 p^2 - x^2 p^1),$$

定义角动量算子 L_1, L_2, L_3 与角动量平方算子为

$$\begin{aligned} L_1 &= X_2 P_3 - X_3 P_2, \\ L_2 &= X_3 P_1 - X_1 P_3, \\ L_3 &= X_1 P_2 - X_2 P_1, \\ L^2 &= L_1^2 + L_2^2 + L_3^2. \end{aligned} \tag{3.1}$$

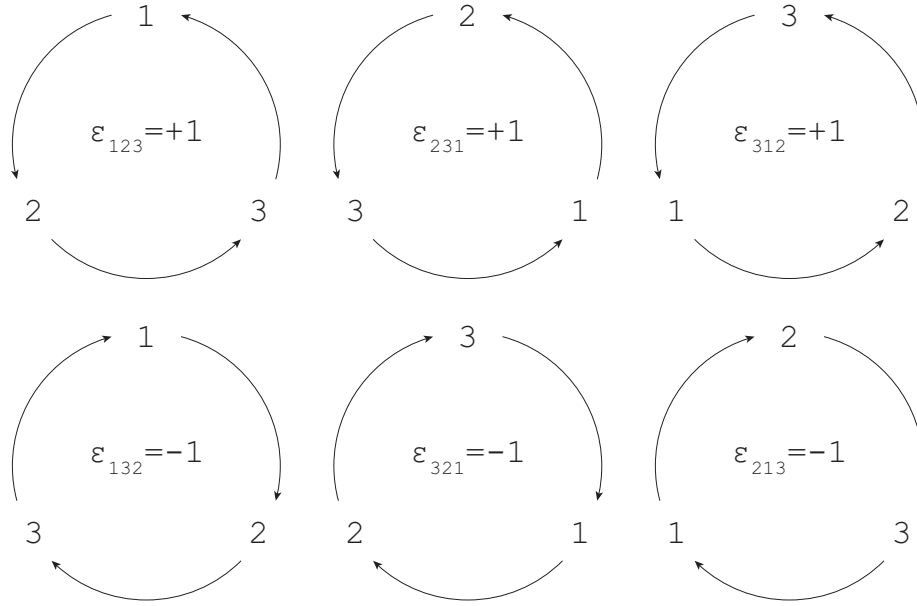


Figure 3.1: ϵ_{ijk} 的直观理解

它们均为厄米算子，并满足

$$\begin{aligned}
 [L_i, L_j] &= i \sum_{k=1}^3 \epsilon_{ijk} L_k, \\
 [L^2, L_i] &= 0, \\
 [L_i, X_j] &= i \sum_{k=1}^3 \epsilon_{ijk} X_k, \\
 [L_i, P_j] &= i \sum_{k=1}^3 \epsilon_{ijk} P_k,
 \end{aligned} \tag{3.2}$$

其中 $\epsilon_{ijk}(i, j, k = 1, 2, 3)$ 的取值可以通过Figure 3.1记忆：当 i, j, k 有任意2个是相等的，那么 $\epsilon_{ijk} = 0$ ；否则，将 i, j, k 逆时针排成一圈，然后绘制箭头，从1指向2，从2指向3，从3指向1。如果箭头方向呈现逆时针，那么 $\epsilon_{ijk} = 1$ ，如果是顺时针，那么 $\epsilon_{ijk} = -1$ 。很显然

$$\begin{aligned}
 \epsilon_{ijk} &= -\epsilon_{ikj}, \\
 &= -\epsilon_{kji}, \\
 &= -\epsilon_{jik}.
 \end{aligned}$$

即交换任意2个下标，取值将变为原来的相反数。以第1个表达式为例，它等价于

$$\begin{aligned}[L_1, L_2] &= i\hbar L_3, \\ [L_2, L_3] &= i\hbar L_1, \\ [L_3, L_1] &= i\hbar L_2.\end{aligned}$$

我们只证明(3.2)的第1个表达式，其余表达式原理类似。根据(2.4)与(2.13)可得

$$\begin{aligned}[L_1, L_2] &= [X_2 P_3 - X_3 P_2, X_3 P_1 - X_1 P_3] \\ &= [X_2 P_3, X_3 P_1] - [X_2 P_3, X_1 P_3] - [X_3 P_2, X_3 P_1] + [X_3 P_2, X_1 P_3] \\ &= [X_2 P_3, X_3] P_1 + X_3 [X_2 P_3, P_1] \\ &\quad - [X_2 P_3, X_1] P_3 - X_1 [X_2 P_3, P_3] \\ &\quad - [X_3 P_2, X_3] P_1 - X_3 [X_3 P_2, P_1] \\ &\quad + [X_3 P_2, X_1] P_3 + X_1 [X_3 P_2, P_3] \\ &= \text{I} - \text{II} - \text{III} + \text{IV},\end{aligned}$$

而

$$\begin{aligned}\text{I} &= -[X_3, X_2 P_3] P_1 - X_3 [P_1, X_2 P_3] \\ &= -([X_3, X_2] P_3 + X_2 [X_3, P_3]) P_1 - X_3 ([P_1, X_2] P_3 + X_2 [P_1, P_3]) \\ &= -i\hbar X_2 P_1, \\ \text{II} &= -[X_1, X_2 P_3] P_3 - X_1 [P_3, X_2 P_3] \\ &= -([X_1, X_2] P_3 + X_2 [X_1, P_3]) P_3 - X_1 ([P_3, X_2] P_3 + X_2 [P_3, P_3]) \\ &= 0, \\ \text{III} &= -[X_3, X_3 P_2] P_1 - X_3 [P_1, X_3 P_2] \\ &= -([X_3, X_3] P_2 + X_3 [X_3, P_2]) P_1 - X_3 ([P_1, X_3] P_2 + X_3 [P_1, P_2]) \\ &= 0, \\ \text{IV} &= -[X_1, X_3 P_2] P_3 - X_1 [P_3, X_3 P_2] \\ &= -([X_1, X_3] P_2 + X_3 [X_1, P_2]) P_3 - X_1 ([P_3, X_3] P_2 + X_3 [P_3, P_2]) \\ &= i\hbar X_1 P_2,\end{aligned}$$

因此

$$[L_1, L_2] = i\hbar (X_1 P_2 - X_2 P_1) = i\hbar L_3.$$

情况 $[L_2, L_3] = i\hbar L_1, [L_3, L_1] = i\hbar L_2$ 的原理类似。

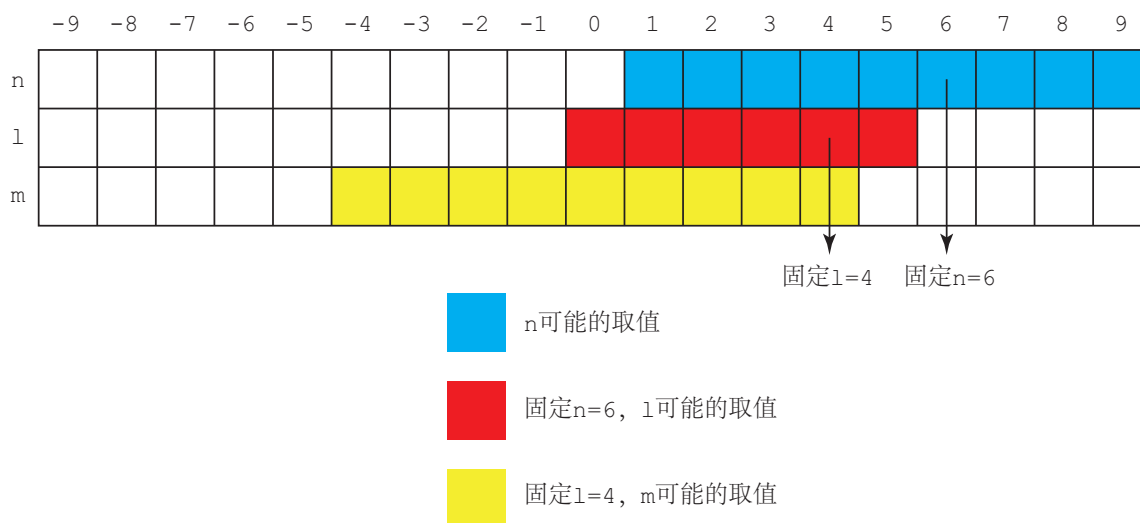


Figure 3.2: n, l, m 可能的取值

3.2 氢原子波函数

氢原子的势能函数为

$$V(\mathbf{x}) = -\frac{e^2}{4\pi\epsilon_0 |\mathbf{x}|},$$

记氢原子核外电子的质量为 m_e ，核外电子满足的定态薛定谔方程为

$$-\frac{\hbar^2}{2m_e} \Delta \Psi(\mathbf{x}) - \frac{e^2}{4\pi\epsilon_0 |\mathbf{x}|} \Psi(\mathbf{x}) = E \Psi(\mathbf{x}),$$

该方程的解为

$$\begin{aligned} \Psi_{n,l,m}(\mathbf{x}) &= \Psi_{n,l,m}, \\ E_{n,l,m} &= -\frac{E_0}{n^2}, \end{aligned}$$

$\{\Psi_{n,l,m}, E_{n,l,m}\}$ 是这一定态薛定谔方程的能级。 n, l, m 均为整数，其取值满足

$$\begin{aligned} n &= 1, 2, 3, \dots, \\ l &= 0, 1, 2, \dots, n-1, \\ m &= -l, -l+1, \dots, l-1, l. \end{aligned} \tag{3.3}$$

如Figure 3.2所示。 $E_0 = \frac{m_e}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 = 13.6 \text{ eV}$ 为一个常数。可以得出， $E_{n,l,m}$ 的取值仅

与 n 有关，将它简记为 $E_n = -E_0/n^2$ 。我们将 $\Psi_{n,l,m}$ 称为**氢原子波函数**，其具体形式为

$$\begin{aligned}\Psi_{n,l,m} &= R_{n,l}(\rho)Y_l^m(\phi, \theta) \\ &= \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} e^{-\rho/na_0} \left(\frac{2\rho}{na_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2\rho}{na_0}\right) Y_l^m(\phi, \theta).\end{aligned}$$

其中 $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 5.29 \times 10^{-11} \text{ m}$ 是**玻尔半径**， $L_{n-l-1}^{2l+1}(x)$ 是次数为 $n-l-1$ 的多项式， $Y_l^m(\phi, \theta)$ 是球谐函数，具体形式均无需得知。 (ρ, ϕ, θ) 是球坐标，它与 \mathbb{R}^3 的坐标 (x, y, z) 的转换关系为

$$(x, y, z) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi).$$

利用例题5.5定义的“简并度”的概念，可以得出，对于正整数 n ，能量 E_n 的简并度为

$$\sum_{l=0}^{n-1} (2l+1) = n^2.$$

我们将列举几个较为简单的氢原子波函数，但是均无需记忆：

$$\begin{aligned}\Psi_{1,0,0} &= \frac{1}{\sqrt{\pi}a_0^{3/2}} e^{-\rho/a_0}, \\ \Psi_{2,0,0} &= \frac{1}{4\sqrt{2\pi}a_0^{3/2}} \left(2 - \frac{\rho}{a_0}\right) e^{-\rho/2a_0}, \\ \Psi_{2,1,0} &= \frac{1}{4\sqrt{2\pi}a_0^{3/2}} \frac{\rho}{a_0} e^{-\rho/2a_0} \cos \phi, \\ \Psi_{2,1,\pm 1} &= \mp \frac{1}{8\sqrt{\pi}a_0^{3/2}} \frac{\rho}{a_0} e^{-\rho/2a_0} \sin \phi e^{\pm i\theta}.\end{aligned}$$

Figure 3.3形象描述了不同 n, l, m 的氢原子波函数模平方 $|\Psi_{n,l,m}|^2$ ，即概率分布的图像。

3.2.1 氢原子跃迁

在某些情况下，氢原子波函数可能从 $\Psi_{n,l,m}$ 变化为 $\Psi_{\tilde{n},\tilde{l},\tilde{m}}$ ，其中 $(n, l, m) \neq (\tilde{n}, \tilde{l}, \tilde{m})$ ，并且 $n > \tilde{n}$ 。这个现象称为**氢原子跃迁**。但是氢原子跃迁需要遵守**选择定则**，其表达式为

$$l - \tilde{l} = \pm 1.$$

在这个情况下，氢原子将向外释放光子，光子的能量为

$$E = E_n - E_{\tilde{n}} = E_0 \left(\frac{1}{\tilde{n}^2} - \frac{1}{n^2} \right).$$

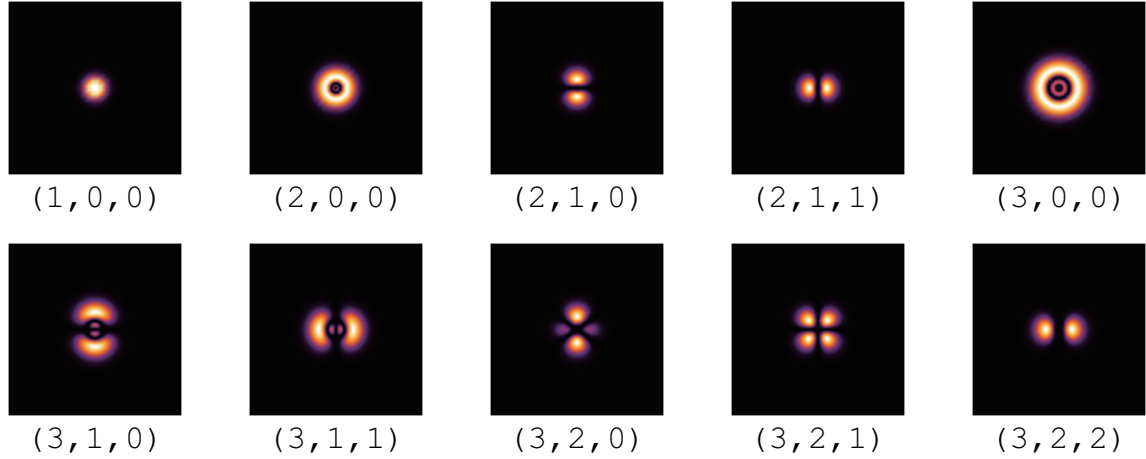


Figure 3.3: 氢原子波函数的图像

而光子的能量 E 与频率 f 的关系为 $E = hf$ ，频率 f 与波长 λ 的关系为 $f = c/\lambda$ ，其中 h, c 分别为普朗克常数与真空中的光速。因此光子的波长满足

$$\frac{1}{\lambda} = \frac{E_0}{hc} \left(\frac{1}{\tilde{n}^2} - \frac{1}{n^2} \right) = R \left(\frac{1}{\tilde{n}^2} - \frac{1}{n^2} \right).$$

其中 $R = E_0/hc = 1.10 \times 10^7 \text{ m}^{-1}$ 是里德伯常数。

可以得出，光子的波长仅仅与 n, \tilde{n} 有关，我们又可以将上述过程简称为“氢原子从能级 n 跃迁到能级 \tilde{n} ”。当氢原子从任意能级 $n > 1$ 跃迁到能级1，我们将其发射的光谱称为**莱曼系**，它们都是紫外线；当氢原子从任意能级 $n > 2$ 跃迁到能级2，我们将其发射的光谱称为**莱曼系**，它们由可见光与紫外线组成；当氢原子从任意能级 $n > 3$ 跃迁到能级3，我们将其发射的光谱称为**巴尔末系**，它们都是红外线。详见Figure 3.4。能级1被称为**氢原子基态**，任意能级 $n > 1$ 都被称为**氢原子激发态**。

通过复杂的计算得出，氢原子波函数 $\Psi_{n,l,m}$ 同时为哈密顿算子、角动量平方算子、角动量算子(见定义(3.1)(2.19)) $\{H, L^2, L_3\}$ 的本征态，其本征值分别为

$$\begin{aligned} H\Psi_{n,l,m} &= -\frac{E_0}{n^2}\Psi_{n,l,m}, \\ L^2\Psi_{n,l,m} &= l(l+1)\hbar^2\Psi_{n,l,m}, \\ L_3\Psi_{n,l,m} &= m\hbar\Psi_{n,l,m}. \end{aligned} \tag{3.4}$$

因此， $\Psi_{n,l,m}$ 对应的能量为 $-E_0/n^2$ ，角动量平方为 $l(l+1)\hbar^2$ ，角动量 z 轴分量为 $m\hbar$ 。

氢原子波函数是归一化的，并满足

$$\begin{aligned} \int_{\mathbb{R}^3} |\Psi_{n,l,m}|^2 dV &= 1, \\ \langle \Psi_{n,l,m}, \Psi_{\tilde{n},\tilde{l},\tilde{m}} \rangle &= \int_{\mathbb{R}^3} \overline{\Psi_{n,l,m}} \Psi_{\tilde{n},\tilde{l},\tilde{m}} dV = \delta_{n\tilde{n}} \delta_{l\tilde{l}} \delta_{m\tilde{m}}. \end{aligned}$$

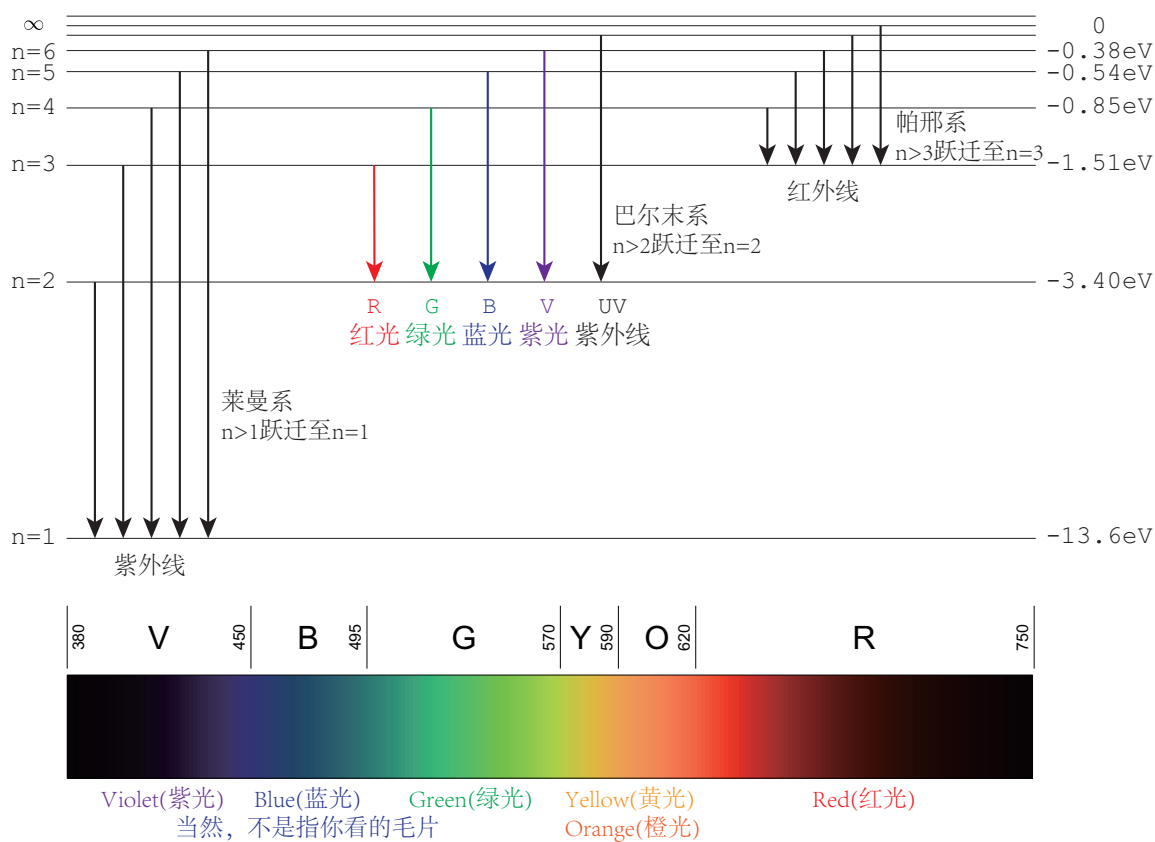


Figure 3.4: 氢原子光谱

对于电荷数 $Z > 1$ (元素周期表的原子序数, 例如Ag是47) 的类氢离子 (例如 He^+ , Li^{2+} , 它们都只有1个核外电子, 分别对应 $Z = 2, 3$), 势能函数变为

$$V(\mathbf{x}) = -\frac{Ze^2}{4\pi\epsilon_0 |\mathbf{x}|},$$

此时哈密顿算子的本征值(能量)将变为

$$E_n = -\frac{Z^2 E_0}{n^2}.$$

类氢离子跃迁的原理与氢原子是类似的, 无需赘述。因此, 跃迁产生的光子的波长满足

$$\frac{1}{\lambda} = Z^2 R \left(\frac{1}{\tilde{n}^2} - \frac{1}{n^2} \right). \quad (3.5)$$

其中 $R = E_0/hc = 1.10 \times 10^7 \text{ m}^{-1}$ 是里德伯常数。

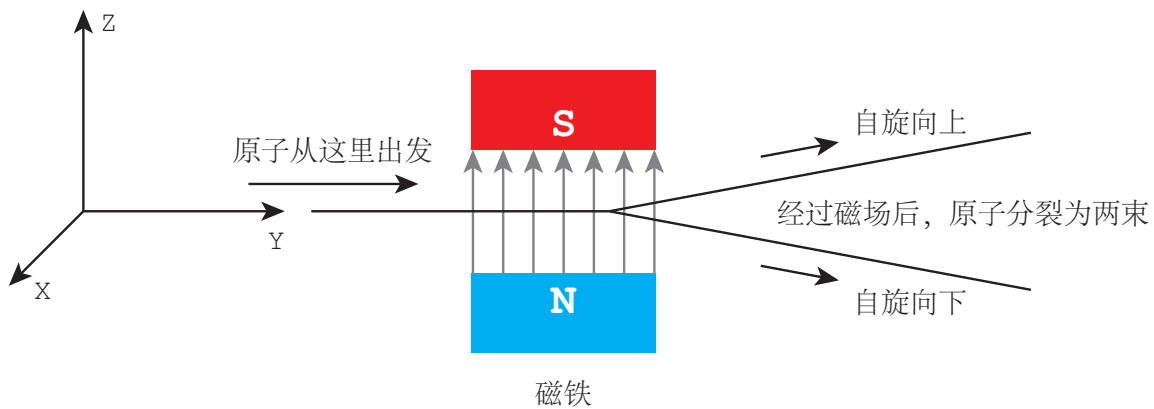


Figure 3.5: Stern-Gerlach实验

3.3 自旋

如Figure 3.5所示，Stern-Gerlach让银原子经过强磁场，发现银原子刚好分成两束，他们根据这个实验提出了**自旋**这一概念。定义**自旋算子** S_1, S_2, S_3 ，它们之间的对易子为

$$\begin{aligned} [S_1, S_2] &= i\hbar S_3, \\ [S_2, S_3] &= i\hbar S_1, \\ [S_3, S_1] &= i\hbar S_2. \end{aligned} \quad (3.6)$$

很显然，也可以简记为

$$[S_i, S_j] = i\hbar \sum_{k=1}^3 \epsilon_{ijk} S_k.$$

其中 S_3 的本征态只有2个，分别记为 $\chi_0(3), \chi_1(3)$ (括号3代表它们是 S_3 的本征态)，它们是标准正交的，满足

$$\langle \chi_i(3), \chi_j(3) \rangle = \delta_{ij}.$$

其本征值为

$$\begin{aligned} S_3 \chi_0(3) &= \frac{\hbar}{2} \chi_0(3), \\ S_3 \chi_1(3) &= -\frac{\hbar}{2} \chi_1(3). \end{aligned} \quad (3.7)$$

我们可以将 S_1, S_2, S_3 视为矩阵:

$$\begin{aligned} S_1 &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ S_2 &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ S_3 &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{3.8}$$

将 $\chi_0(3), \chi_1(3)$ 视为向量:

$$\begin{aligned} \chi_0(3) &\longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \chi_1(3) &\longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

容易验证 S_1, S_2, S_3 满足表达式(3.6), 并且(3.7)也是成立的。
因此

$$\begin{aligned} S_1 \chi_0(3) &= \frac{\hbar}{2} \chi_1(3), \\ S_1 \chi_1(3) &= \frac{\hbar}{2} \chi_0(3), \\ S_2 \chi_0(3) &= \frac{i\hbar}{2} \chi_1(3), \\ S_2 \chi_1(3) &= -\frac{i\hbar}{2} \chi_0(3). \end{aligned} \tag{3.9}$$

记泡利矩阵为

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{3.10}$$

因此

$$\begin{aligned} S_1 &= \frac{\hbar}{2} \sigma_1, \\ S_2 &= \frac{\hbar}{2} \sigma_2, \\ S_3 &= \frac{\hbar}{2} \sigma_3. \end{aligned} \tag{3.11}$$

泡利矩阵在群表示论中具有广泛的应用。泡利矩阵具有良好的性质，根据(3.6)与(3.11)可得，

$$[\sigma_i, \sigma_j] = \sigma_i \sigma_j - \sigma_j \sigma_i = 2i \sum_{k=1}^3 \epsilon_{ijk} \sigma_k. \quad (3.12)$$

我们还可以得到(自行验证):

$$\begin{aligned} \{\sigma_i, \sigma_j\} &= \sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} I, \\ \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 = I. \end{aligned} \quad (3.13)$$

3.3.1 四元数*

¹⁷ 定义实数域 \mathbb{R} 上的4维向量空间 $\mathbb{H} = \mathbb{C} \times \mathbb{C}$ ，并定义代数运算 $\mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}$ 为

$$(a, b)(c, d) = (ac - d\bar{b}, \bar{a}d + cb), \quad a, b, c, d \in \mathbb{C}.$$

可以得出，该代数运算并不是交换的，因此 \mathbb{H} 不是一个域。我们将向量空间 \mathbb{H} 中的元素称为**四元数**。对于任意四元数 $p = (a, b) \in \mathbb{H}$ ，定义 $p^* = (\bar{a}, -b)$ 。可以得出，对于任意四元数 p, q, r ，都具有

$$\begin{aligned} (pq)^* &= q^* p^*, \\ p(qr) &= (pq)r. \end{aligned}$$

因此，四元数的代数运算是结合的， \mathbb{H} 是一个环。

定义

$$\begin{aligned} \mathbf{1} &= (1, 0), \\ \mathbf{i} &= (i, 0), \\ \mathbf{j} &= (0, 1), \\ \mathbf{k} &= (0, -i). \end{aligned}$$

显然，它们是 \mathbb{H} 上的一组基，换句话说，任意四元数都可以被表示为 $a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ ，其中 $a, b, c, d \in \mathbb{R}$ 。

¹⁷ 符号*表示这一章节可以跳过，建议跳转到第107页的章节3.3.2。

根据计算可得

$$\begin{aligned}
\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 &= -\mathbf{1}, \\
\mathbf{1}q &= q\mathbf{1}, & \text{for all } q \in \mathbb{H}, \\
\mathbf{ij} &= -\mathbf{ji} = \mathbf{k}, \\
\mathbf{jk} &= -\mathbf{kj} = \mathbf{i}, \\
\mathbf{ki} &= -\mathbf{ik} = \mathbf{j}, \\
\mathbf{1}^* &= \mathbf{1}, \\
\mathbf{i}^* &= -\mathbf{i}, \\
\mathbf{j}^* &= -\mathbf{j}, \\
\mathbf{k}^* &= -\mathbf{k}.
\end{aligned}$$

其中 $\mathbf{i}^2 = (i, 0)(i, 0) = -(1, 0)$, $\mathbf{j}^2, \mathbf{k}^2$ 的计算原理类似。如果我们将形如 $x\mathbf{1}, x \in \mathbb{R}$ 的四元数视为实数 x , 那么对于任意四元数 p , 我们可以得出

$$p^*p \geq 0,$$

取等号当且仅当 $p = (0, 0)$ 。因此, 我们可以定义四元数 p 的范数为

$$|p| = \sqrt{p^*p}.$$

任意非零的四元数 p 都具有乘法逆元 p^{-1} , 其取值为

$$p^{-1} = |p|^{-2}p^*,$$

可以得出

$$p^{-1}p = pp^{-1} = (1, 0).$$

因此, 四元数虽然不是一个域, 但是可以定义除法运算: 对于任意四元数 $a \neq (0, 0)$ 与 b , 有关四元数 x 的方程 $ax = b, xa = b$ 均有解。

对于泡利矩阵, 我们记

$$\begin{aligned}
I &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\
X &= i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\
Y &= i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \\
Z &= i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.
\end{aligned}$$

而所有2阶矩阵的集合 $M(2, \mathbb{C})$ 是实数域 \mathbb{R} 上的8维向量空间，记

$$S = \text{span}(I, X, Y, Z) = \{aI + bX + cY + dZ : a, b, c, d \in \mathbb{R}\} \subseteq M(2, \mathbb{C})$$

为向量空间 $M(2, \mathbb{C})$ 的一个子空间。定义双线性运算 $S \times S \rightarrow S$ 为矩阵的乘积，根据计算可得

$$\begin{aligned} X^2 &= Y^2 = Z^2 = -I, \\ IA &= AI, & \text{for all } A \in S, \\ XY &= -YX = Z, \\ YZ &= -ZY = X, \\ ZX &= -XZ = Y, \\ I^\dagger &= I, \\ X^\dagger &= -X, \\ Y^\dagger &= -Y, \\ Z^\dagger &= -Z. \end{aligned}$$

很显然，泡利矩阵与四元数有着千丝万缕的联系。

我们还可以进一步推广到**八元数**，它的集合是实数域 \mathbb{R} 上的8维向量空间 $\mathbb{O} = \mathbb{H} \times \mathbb{H}$ ，并定义代数运算 $\mathbb{O} \times \mathbb{O} \rightarrow \mathbb{O}$ 为

$$(p, q)(r, s) = (pr - sq^*, p^*s + rq), \quad p, q, r, s \in \mathbb{H}.$$

可以得出，该代数运算既不是交换的，又不是结合的，这表明 \mathbb{O} 甚至不是一个环。但是，我们可以定义更弱的结构——**可除代数**：它是一个含有乘法单位元，并且没有零因子(换句话说， $ab = 0$ 当且仅当 $a = 0$ 或 $b = 0$)的代数。

很显然，实数 \mathbb{R} 、复数 \mathbb{C} 、四元数 \mathbb{H} 均为可除代数，它们的乘法单位元分别为 $1, 1, \mathbf{1}$ ，并且任意非零实数、复数、四元数都具有乘法逆元。Frobenius在1877年证明了有限维结合的可除代数只能是 $\mathbb{R}, \mathbb{C}, \mathbb{H}$ ；之后，Hurwitz在1898年证明了有限维可除代数只能是 $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ 。因此，我们将不再往下讨论“十六元数”“三十二元数”等，因为它们几乎失去了所有性质，研究价值很低。

3.3.2 自旋波函数是 \mathbb{C}^2 上的向量

任意**自旋波函数**都可以被表示为 $\chi_0(3), \chi_1(3)$ 的非零线性组合 $\Psi = c_0\chi_0(3) + c_1\chi_1(3)$ 。我们可以将它表示为 \mathbb{C}^2 上的向量

$$c_0\chi_0(3) + c_1\chi_1(3) \longleftrightarrow \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}. \quad (3.14)$$

而波函数是归一化的，因此

$$|c_0|^2 + |c_1|^2 = 1. \quad (3.15)$$

在自旋体系下，我们可以将任意线性算子 A 视为2阶矩阵

$$\begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix}, \quad (3.16)$$

其中 $A_{ij} = \langle \chi_i(3), A\chi_j(3) \rangle$ 。记它的本征态为 $\Psi = c_0\chi_0(3) + c_1\chi_1(3)$ ，对应的本征值为 λ ，那么

$$\begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}. \quad (3.17)$$

因此，在自旋体系下，我们可以很容易地求出任意线性算子 A 的本征态与对应的本征值！

我们以 S_1 为例，那么

$$\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \end{pmatrix},$$

该方程具有2个解(线性无关的意义下)，分别为

$$\begin{aligned} c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= \frac{1}{\sqrt{2}}, & \lambda &= \frac{\hbar}{2}, \\ c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= -\frac{1}{\sqrt{2}}, & \lambda &= -\frac{\hbar}{2}. \end{aligned}$$

因此， S_1 具有2个本征态，分别记为

$$\begin{aligned} \chi_0(1) &= \frac{1}{\sqrt{2}}(\chi_0(3) + \chi_1(3)), \\ \chi_1(1) &= \frac{1}{\sqrt{2}}(\chi_0(3) - \chi_1(3)), \end{aligned}$$

对应的本征值分别为

$$\begin{aligned} S_1\chi_0(1) &= \frac{\hbar}{2}\chi_0(1), \\ S_1\chi_1(1) &= -\frac{\hbar}{2}\chi_1(1). \end{aligned}$$

以 S_2 为例，那么

$$\frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \end{pmatrix},$$

该方程同样具有2个解(线性无关的意义下)，分别为

$$\begin{aligned} c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= \frac{i}{\sqrt{2}}, & \lambda &= \frac{\hbar}{2}, \\ c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= -\frac{i}{\sqrt{2}}, & \lambda &= -\frac{\hbar}{2}. \end{aligned}$$

因此, S_2 同样具有2个本征态, 分别记为

$$\begin{aligned}\chi_0(2) &= \frac{1}{\sqrt{2}}(\chi_0(3) + i\chi_1(3)), \\ \chi_1(2) &= \frac{1}{\sqrt{2}}(\chi_0(3) - i\chi_1(3)),\end{aligned}$$

对应的本征值分别为

$$\begin{aligned}S_2\chi_0(2) &= \frac{\hbar}{2}\chi_0(2), \\ S_2\chi_1(2) &= -\frac{\hbar}{2}\chi_1(2).\end{aligned}$$

根据(3.14), 通过向量表示, 可得

$$\begin{aligned}\chi_0(1) &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \chi_1(1) &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \\ \chi_0(2) &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \\ \chi_1(2) &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \\ \chi_0(3) &\longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \chi_1(3) &\longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}\tag{3.18}$$

容易得出, 对于任意 $u \in \{1, 2, 3\}$ 与 $i, j \in \{0, 1\}$, 都具有

$$\chi_i(u) \cdot \chi_j(u) = \chi_i(u)^\dagger \chi_j(u) = \delta_{ij}.$$

(注意, 复空间上的内积是 $v_i^\dagger v_j$, 而不是 $v_i^T v_j$)

定义自旋平方算子 S^2 为

$$S^2 = S_1^2 + S_2^2 + S_3^2,\tag{3.19}$$

根据(3.8)可得

$$S^2 \longleftrightarrow \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\tag{3.20}$$

因此

$$S^2 = \frac{3\hbar^2}{4} I,$$

对于任意 $u \in \{1, 2, 3\}$ 与 $i \in \{0, 1\}$, 都具有

$$S^2 \chi_i(u) = \frac{3\hbar^2}{4} \chi_i(u). \quad (3.21)$$

3.4 矩阵指数

记 $\mathfrak{gl}(n, \mathbb{C})$ 是所有 n 阶矩阵的集合, 它是实数域 \mathbb{R} 上的 $2n^2$ 维向量空间; $\mathrm{GL}(n, \mathbb{C})$ 是所有 n 阶可逆矩阵的集合, 它是一个群。定义映射

$$\exp : \mathfrak{gl}(n, \mathbb{C}) \rightarrow \mathrm{GL}(n, \mathbb{C}), A \mapsto \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

根据 Weierstrass M 判别法可得, 该级数在任意有界闭集 $\Omega \subseteq \mathfrak{gl}(n, \mathbb{C})$ 上都是一致收敛的, 类比于微积分的级数公式 $e^x = \sum_{k=0}^{\infty} x^k/k!$ 。显然

$$\begin{aligned} \exp 0 &= I, \\ \exp A^T &= (\exp A)^T, \\ \exp A^\dagger &= (\exp A)^\dagger. \end{aligned} \quad (3.22)$$

Theorem 3.1. 对于任意矩阵 $A \in \mathfrak{gl}(n, \mathbb{C})$ 与任意整数 $n \in \mathbb{Z}$, 都具有

$$\begin{aligned} (\exp A)^{-1} &= \exp(-A), \\ (\exp A)^n &= \exp(nA). \end{aligned}$$

甚至, 我们可以推广结论: 对于任意实数 $s, t \in \mathbb{R}$, 都具有

$$(\exp sA)(\exp tA) = \exp(s+t)A. \quad (3.23)$$

Proof. 第1个表达式等价于

$$(\exp A)(\exp(-A)) = I,$$

我们将证明这个表达式。

$$\begin{aligned} (\exp A)(\exp(-A)) &= \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) \left(\sum_{l=0}^{\infty} (-1)^l \frac{A^l}{l!} \right) \\ &= \sum_{k,l=0}^{\infty} (-1)^l \frac{A^{k+l}}{k!l!} \\ &= \sum_{k,l=0}^{\infty} (-1)^l \binom{k+l}{k} \frac{A^{k+l}}{(k+l)!}, \end{aligned}$$

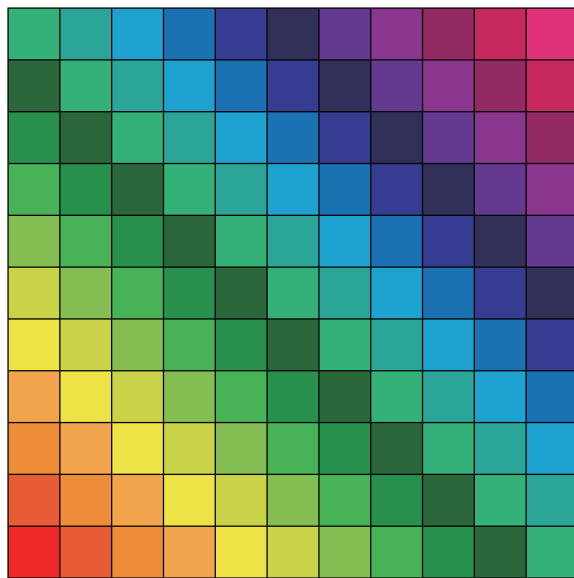


Figure 3.6: 求和方式

如Figure 3.6所示，我们可以“斜向”求和，令 $m = k + l$ ，并利用组合数的性质 $\binom{m}{k} = \binom{m}{m-k} = \binom{m}{l}$ ，可得

$$(\exp A)(\exp(-A)) = \sum_{m=0}^{\infty} \frac{A^m}{m!} \sum_{l=0}^m (-1)^l \binom{m}{l},$$

对于 $\sum_{l=0}^m (-1)^l \binom{m}{l}$ ，当 $m = 0$ 时， $\sum_{l=0}^m (-1)^l \binom{m}{l} = (-1)^0 \binom{0}{0} = 1$ ；当 $m > 0$ 时，利用二项式定理 $(a + b)^m = \sum_{k=0}^m a^k b^{m-k} \binom{m}{k}$ 可得

$$\sum_{l=0}^m (-1)^l \binom{m}{l} = \sum_{l=0}^m (-1)^l 1^{m-l} \binom{m}{l} = (-1 + 1)^m = 0.$$

因此

$$(\exp A)(\exp(-A)) = \frac{A^0}{0!} = I.$$

对于第2个表达式，我们只需证明 $n \geq 0$ 的情况即可，因为 $n < 0$ 的情况可以通过第1个表达式直接证明。当 $n = 0$ ，显然成立，我们假设该表达式对于 $n \geq 0$ 均成立，那么

$$\begin{aligned}
 (\exp A)^n (\exp A) &= (\exp(nA))(\exp A) \\
 &= \left(\sum_{k=0}^{\infty} \frac{(nA)^k}{k!} \right) \left(\sum_{l=0}^{\infty} \frac{A^l}{l!} \right) \\
 &= \sum_{k,l=0}^{\infty} n^k \binom{k+l}{k} \frac{A^{k+l}}{(k+l)!} \\
 &= \sum_{m=0}^{\infty} \frac{A^m}{m!} \sum_{l=0}^m n^{m-l} \binom{m}{l} \\
 &= \sum_{m=0}^{\infty} \frac{A^m}{m!} (n+1)^m \\
 &= \exp((n+1)A),
 \end{aligned}$$

即该表达式对于 $n+1$ 也成立。根据数学归纳法，完成证明。

第3个表达式的证明原理类似，不做赘述。

□

接下来的引理对矩阵指数的计算至关重要：

Lemma 3.2. 如果 S 是可逆矩阵，那么

$$\exp(SAS^{-1}) = S(\exp A)S^{-1}, \quad (3.24)$$

Proof.

$$\begin{aligned}
 \exp(SAS^{-1}) &= \sum_{k=0}^{\infty} \frac{(SAS^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{(SAS^{-1})(SAS^{-1}) \cdots (SAS^{-1})}{k!} \\
 &= \sum_{k=0}^{\infty} \frac{SA^k S^{-1}}{k!} \\
 &= S \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) S^{-1} \\
 &= S(\exp A)S^{-1}.
 \end{aligned}$$

□

据此，我们可以很容易地计算矩阵的指数：一般情况下， A 的本征态形成向量空间 \mathbb{C}^n 的一组基，它们是标准正交的。记 A 的特征值为 $\lambda_1, \dots, \lambda_n$ ，特征向量为 $v_1 = (v_1^1, \dots, v_1^n)^T$ ，

$\dots, v_n = (v_n^1, \dots, v_n^n)^T$, 并且满足

$$v_i \cdot v_j = \sum_{k=1}^n \overline{v_i^k} v_j^k = \delta_{ij}. \quad (3.25)$$

(注意, 复空间上的内积是 $v_i^\dagger v_j$, 而不是 $v_i^T v_j$) 因此 $V^\dagger V = I$, V 是酉矩阵。并且对于任意 $i \in \{1, \dots, n\}$, 都具有

$$A \begin{pmatrix} v_i^1 \\ v_i^2 \\ \vdots \\ v_i^n \end{pmatrix} = \lambda_i \begin{pmatrix} v_i^1 \\ v_i^2 \\ \vdots \\ v_i^n \end{pmatrix}.$$

这等价于

$$A \begin{pmatrix} v_1^1 & v_2^1 & \cdots & v_n^1 \\ v_1^2 & v_2^2 & \cdots & v_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ v_1^n & v_2^n & \cdots & v_n^n \end{pmatrix} = \begin{pmatrix} v_1^1 & v_2^1 & \cdots & v_n^1 \\ v_1^2 & v_2^2 & \cdots & v_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ v_1^n & v_2^n & \cdots & v_n^n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

因此

$$A = V \Lambda V^{-1}, \quad (3.26)$$

其中

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \quad V = \begin{pmatrix} v_1^1 & v_2^1 & \cdots & v_n^1 \\ v_1^2 & v_2^2 & \cdots & v_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ v_1^n & v_2^n & \cdots & v_n^n \end{pmatrix}.$$

因为 V 是酉矩阵, 所以 $V^{-1} = V^\dagger$ 。那么根据(3.24)可得

$$\begin{aligned}
\exp A &= V(\exp \Lambda)V^{-1} = V \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!} & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{\lambda_2^k}{k!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{\infty} \frac{\lambda_n^k}{k!} \end{pmatrix} V^{-1} \\
&= V \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} V^{-1} = V \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} V^\dagger \\
&= \begin{pmatrix} v_1^1 & v_2^1 & \cdots & v_n^1 \\ v_1^2 & v_2^2 & \cdots & v_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ v_1^n & v_2^n & \cdots & v_n^n \end{pmatrix} \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} \begin{pmatrix} \overline{v_1^1} & \overline{v_1^2} & \cdots & \overline{v_1^n} \\ \overline{v_2^1} & \overline{v_2^2} & \cdots & \overline{v_2^n} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{v_n^1} & \overline{v_n^2} & \cdots & \overline{v_n^n} \end{pmatrix}.
\end{aligned} \tag{3.27}$$

据此, 我们可以得出以下结论:

Corollary 3.2.1. 对于任意矩阵 $A \in \mathfrak{gl}(n, \mathbb{C})$, 矩阵 $\exp A$ 的特征值为 $e^{\lambda_1}, \dots, e^{\lambda_n}$ 。
因此, $\exp A$ 的行列式为

$$\det(\exp A) = e^{\text{tr}(A)}, \tag{3.28}$$

其中 $\text{tr}(A)$ 为矩阵的迹, 即矩阵对角元素之和。

Proof. 根据(3.27)可得

$$\exp A = V \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} V^{-1},$$

因此

$$\begin{aligned}
\exp A - e^{\lambda_1} I &= V \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} V^{-1} - V V^{-1} \\
&= V \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix} V^{-1},
\end{aligned}$$

这表明 $\det(\exp A - e^{\lambda_1} I) = 0$ ，因此 e^{λ_1} 是 $\exp A$ 的特征值；对于 $\lambda_2, \dots, \lambda_n$ ，情况相同。而 $\exp A$ 是 n 阶矩阵，特征值不超过 n 个，所以 $e^{\lambda_1}, \dots, e^{\lambda_n}$ 是 $\exp A$ 的全部特征值。

而矩阵的行列式为全部特征值之积，迹为全部特征值之和，因此 $\det(\exp A) = e^{\lambda_1} \dots e^{\lambda_n} = e^{\lambda_1 + \dots + \lambda_n} = e^{\text{tr}(A)}$ 。□

同时，我们可以定义映射

$$\log : \text{GL}(n, \mathbb{C}) \rightarrow \mathfrak{gl}(n, \mathbb{C}), A \mapsto \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(A - I)^k}{k}.$$

该级数只在集合

$$\{A : \|A - I\| < 1\} \subseteq \text{GL}(n, \mathbb{C})$$

上收敛(在此不做证明)，其中

$$\|A\| = \sqrt{\text{tr}(A^\dagger A)}, \quad (3.29)$$

是矩阵的**Hilbert 范数**。可以得出(在此不做证明)

$$\begin{aligned} \|A\| &\geq 0, \\ \|\lambda A\| &= |\lambda| \|A\|, \\ \|A + B\| &\leq \|A\| + \|B\|, \\ \|AB\| &\leq \|A\| \|B\|. \end{aligned}$$

类比于微积分的级数公式 $\log x = \sum_{k=1}^{\infty} (-1)^{k-1} (x-1)^k / k$ ，它的收敛半径为1。

Theorem 3.3. 对于任意矩阵 A 满足 $\|A\| < \log 2$ ，都满足

$$\log(\exp A) = A.$$

同样的，对于任意矩阵 A 满足 $\|A - I\| < 1$ ，都满足

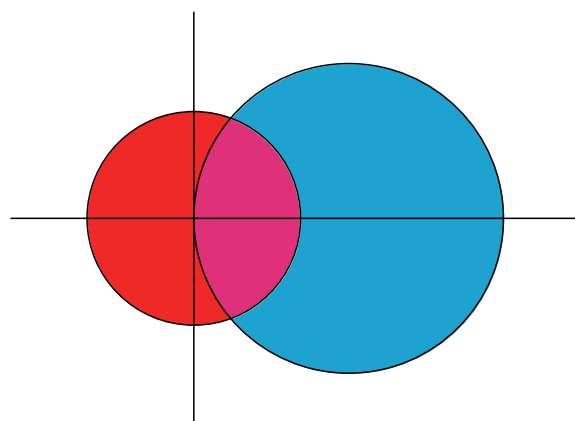
$$\exp(\log A) = A.$$

因此，在集合 $\{A : \|A\| < \log 2, \|A - I\| < 1\}$ ， \log 与 \exp 互为逆映射，如Figure 3.7所示。

1, 2表达式原理类似，我们将只证明第1个表达式。在证明它之前，我们首先证明 $\|A\| < \log 2$ 能推导出 $\|\exp A - I\| < 1$ 。假设 $\|A\| < \epsilon$ ，那么

$$\|\exp A - I\| = \left\| \sum_{k=1}^{\infty} \frac{A^k}{k!} \right\| \leq \sum_{k=1}^{\infty} \left\| \frac{A^k}{k!} \right\| = \sum_{k=1}^{\infty} \frac{\|A^k\|}{k!} \leq \sum_{k=1}^{\infty} \frac{\|A\|^k}{k!} < \sum_{k=1}^{\infty} \frac{\epsilon^k}{k!} = e^\epsilon - 1.$$

因此，如果 $\epsilon \leq \log 2$ ，那么 $\|\exp A - I\| < 1$ 。



$$\begin{aligned} &||A|| < \log 2 = 0.7 \\ &||A - I|| < 1 \\ &||A|| < \log 2 = 0.7, ||A - I|| < 1 \end{aligned}$$

Figure 3.7: A 需要满足的条件

Proof. 我们首先计算任意 $||A - I|| < 1$ 对应 $\log A$ 的取值。记 A 的特征值为 $\lambda_1, \dots, \lambda_n$ ，根据(3.26)，我们将 A 对角化为

$$A = V\Lambda V^{-1} = V \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} V^{-1}$$

因此

$$A - I = V\Lambda V^{-1} - VV^{-1} = V(\Lambda - I)V^{-1} = V \begin{pmatrix} \lambda_1 - 1 & 0 & \cdots & 0 \\ 0 & \lambda_2 - 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - 1 \end{pmatrix} V^{-1}$$

因为 $\|A - I\| < 1$, 所以任意特征值均满足 $|\lambda_i - 1| < 1$ (在此不做证明, 较为复杂)。因此

$$\begin{aligned}\log A &= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(A - I)^k}{k} \\ &= V \begin{pmatrix} \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\lambda_1 - 1)^k}{k} & 0 & \cdots & 0 \\ 0 & \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\lambda_2 - 1)^k}{k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\lambda_n - 1)^k}{k!} \end{pmatrix} V^{-1} \\ &= V \begin{pmatrix} \log \lambda_1 & 0 & \cdots & 0 \\ 0 & \log \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \log \lambda_n \end{pmatrix} V^{-1}.\end{aligned}$$

根据Theorem 3.2.1可得

$$\log(\exp A) = V \begin{pmatrix} \log e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & \log e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \log e^{\lambda_n} \end{pmatrix} V^{-1} = V \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} V^{-1} = A.$$

□

对于 n^2 个光滑函数 $A_j^i(t) \in \mathbb{C}, i, j \in \{1, \dots, n\}$, 定义含时间演化的矩阵

$$A(t) = \begin{pmatrix} A_1^1(t) & \cdots & A_n^1(t) \\ \vdots & \ddots & \vdots \\ A_1^n(t) & \cdots & A_n^n(t) \end{pmatrix},$$

定义时间导数为

$$\frac{d}{dt} A(t) = \begin{pmatrix} \dot{A}_1^1(t) & \cdots & \dot{A}_n^1(t) \\ \vdots & \ddots & \vdots \\ \dot{A}_1^n(t) & \cdots & \dot{A}_n^n(t) \end{pmatrix}.$$

据此, 我们可以得出以下定理:

Theorem 3.4. 对于任意含时间演化的矩阵 $A(t), B(t)$ 与常数 $\lambda \in \mathbb{C}$, 都具有

$$\begin{aligned}\frac{d}{dt}(A(t) + B(t)) &= \frac{d}{dt}A(t) + \frac{d}{dt}B(t), \\ \frac{d}{dt}(\lambda A(t)) &= \lambda \frac{d}{dt}A(t), \\ \frac{d}{dt}\bigg|_{t=t_0}(A(t)B(t)) &= \left(\frac{d}{dt}\bigg|_{t=t_0}A(t)\right)B(t_0) + A(t_0)\left(\frac{d}{dt}\bigg|_{t=t_0}B(t)\right).\end{aligned}$$

Proof. 1, 2表达式是显然的, 只需证明第3个表达式。因为等式两边都是矩阵, 所以只需证明, 对于任意 $i, j \in \{1, \dots, n\}$, 其第 i 行, 第 j 列元素相等即可。等式左边为

$$\frac{d}{dt}\bigg|_{t=t_0}\left(\sum_{k=1}^n A_k^i(t)B_j^k(t)\right) = \sum_{k=1}^n \left(\frac{d}{dt}\bigg|_{t=t_0}A_k^i(t)\right)B_j^k(t_0) + A_k^i(t_0)\left(\frac{d}{dt}\bigg|_{t=t_0}B_j^k(t)\right),$$

这恰好与等式右边第 i 行, 第 j 列元素相等。 \square

因此, 不难得出

$$\frac{d}{dt}\bigg|_{t=0}\exp tA = \sum_{k=0}^{\infty} \frac{d}{dt}\bigg|_{t=0} \frac{(tA)^k}{k!} = \sum_{k=0}^{\infty} \left(\frac{d}{dt}\bigg|_{t=0} t^k\right) \frac{A^k}{k!} = A. \quad (3.30)$$

Lemma 3.5. 如果含时间演化的矩阵 $A(t)$ 满足

$$A(0) = I,$$

那么

$$\frac{d}{dt}\bigg|_{t=0} \log A(t) = \frac{d}{dt}\bigg|_{t=0} A(t).$$

Proof. 记 $B(t) = A(t) - I$, 那么 $B(0) = 0$ 。根据定义可得

$$\log A(t) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{B(t)^k}{k}.$$

因此

$$\frac{d}{dt}\bigg|_{t=0} \log A(t) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \frac{d}{dt}\bigg|_{t=0} B(t)^k,$$

对于任意 $m > 1$, 根据Theorem 3.4不难发现, $\frac{d}{dt}\big|_{t=0} B(t)^m$ 的每项展开式必然包含 $B(0)$ 。因此, $\frac{d}{dt}\big|_{t=0} B(t)^m = 0$ 对于任意 $m > 1$ 均成立, 这表明

$$\frac{d}{dt}\bigg|_{t=0} \log A(t) = \frac{d}{dt}\bigg|_{t=0} B(t) = \frac{d}{dt}\bigg|_{t=0} (A(t) - I) = \frac{d}{dt}\bigg|_{t=0} A(t).$$

\square

注意：我们始终没有说明 $(\exp X)(\exp Y) = \exp(X + Y)$ ，因为它不正确。我们给出一个反例：将泡利矩阵 $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ 对角化可得

$$\begin{aligned}\sigma_1 &= V_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} V_1^\dagger, \\ \sigma_2 &= V_2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} V_2^\dagger,\end{aligned}$$

其中 $V_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, V_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ 。根据(3.27)可得

$$\begin{aligned}\exp \sigma_1 &= V_1 \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} V_1^\dagger = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix}, \\ \exp \sigma_2 &= V_2 \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} V_2^\dagger = \begin{pmatrix} \cosh 1 & -i \sinh 1 \\ i \sinh 1 & \cosh 1 \end{pmatrix}.\end{aligned}$$

将 $\sigma_1 + \sigma_2 = \begin{pmatrix} 0 & 1-i \\ 1+i & 0 \end{pmatrix}$ 对角化可得

$$\sigma_1 + \sigma_2 = V_3 \begin{pmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{pmatrix} V_3^\dagger,$$

其中 $V_3 = \begin{pmatrix} \frac{1-i}{2} & -\frac{1-i}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ 。根据(3.27)可得

$$\exp(\sigma_1 + \sigma_2) = V_3 \begin{pmatrix} e^{\sqrt{2}} & 0 \\ 0 & e^{-\sqrt{2}} \end{pmatrix} V_3^\dagger = \begin{pmatrix} \cosh \sqrt{2} & \frac{1-i}{\sqrt{2}} \sinh \sqrt{2} \\ \frac{1+i}{\sqrt{2}} \sinh \sqrt{2} & \cosh \sqrt{2} \end{pmatrix}.$$

显然 $(\exp \sigma_1)(\exp \sigma_2) \neq \exp(\sigma_1 + \sigma_2)$ 。但是，在 X, Y 非常小(严格来说是 $\|X\|, \|Y\|$ 非常小)的时候， $(\exp X)(\exp Y)$ 的取值近似为 $\exp(X + Y)$ ，我们将严格叙述这一现象：

Theorem 3.6. 对于任意矩阵 X, Y ，存在含时间演化的矩阵 $Z(t)$ ，使得对于任意 $t \in (-\epsilon, \epsilon)$ ，都具有

$$(\exp tX)(\exp tY) = \exp(t(X + Y) + t^2 Z(t)), \quad (3.31)$$

其中 $\epsilon > 0$ 是与 X, Y 有关的常数：对于任意 $t \in (-\epsilon, \epsilon)$ ，都具有 $\|(\exp tX)(\exp tY) - I\| < 1$ ，即 $\log((\exp tX)(\exp tY))$ 是可以被定义的。

Proof. 定义含时间演化的矩阵 $\phi(t)$ 为

$$\phi(t) = \log((\exp tX)(\exp tY)).$$

根据Theorem 3.4, Lemma 3.5, (3.30)可得

$$\left. \frac{d}{dt} \right|_{t=0} \phi(t) = \left. \frac{d}{dt} \right|_{t=0} (\exp tX)(\exp tY) = \left. \frac{d}{dt} \right|_{t=0} \exp tX + \left. \frac{d}{dt} \right|_{t=0} \exp tY = X + Y,$$

因此, 根据泰勒定理, 存在含时间演化的矩阵 $Z(t)$, 满足

$$\phi(t) = t(X + Y) + t^2 Z(t).$$

□

Corollary 3.6.1.

$$\lim_{n \rightarrow \infty} \left(\left(\exp \frac{t}{n} X \right) \left(\exp \frac{t}{n} Y \right) \right)^n = \exp t(X + Y). \quad (3.32)$$

Proof. 根据(3.31)可得

$$\left(\exp \frac{t}{n} X \right) \left(\exp \frac{t}{n} Y \right) = \exp \left(\frac{t}{n} (X + Y) + \frac{t^2}{n^2} Z \left(\frac{t}{n} \right) \right),$$

根据Theorem 3.1第2个表达式可得

$$\left(\left(\exp \frac{t}{n} X \right) \left(\exp \frac{t}{n} Y \right) \right)^n = \exp \left(t(X + Y) + \frac{t^2}{n} Z \left(\frac{t}{n} \right) \right),$$

固定 t 的取值, 取极限 $n \rightarrow \infty$ 即可完成证明。

□

3.4.1 矩阵李群*

¹⁸ 很多矩阵集合在矩阵乘法运算下是一个群, 我们将这样的矩阵集合称为**矩阵李群**。我们将给出矩阵李群的一些实例:

- (A) 定义**一般线性群** $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{GL}(n, \mathbb{C})$ 分别为所有可逆实矩阵、复矩阵的集合。很显然, 它们在乘法运算下是一个群, 因为矩阵 A 可逆当且仅当 $\det A \neq 0$, 并且根据线性代数的知识, $(\det A)(\det B) = \det AB$, $\det(A^{-1}) = 1/\det A$, 我们可以得到2个可逆矩阵的乘积还是可逆矩阵, 可逆矩阵的逆元还是可逆矩阵。很显然, $\mathrm{GL}(n, \mathbb{R})$ 是 $\mathrm{GL}(n, \mathbb{C})$ 的子群。
- (B) 定义**特殊线性群** $\mathrm{SL}(n, \mathbb{R})$, $\mathrm{SL}(n, \mathbb{C})$ 分别为所有行列式为1的实矩阵、复矩阵的集合。利用 $(\det A)(\det B) = \det AB$, $\det(A^{-1}) = 1/\det A$, 它们同样在乘法运算下是一个群。很显然, $\mathrm{SL}(n, \mathbb{R})$ 是 $\mathrm{GL}(n, \mathbb{R})$ 的子群, $\mathrm{SL}(n, \mathbb{C})$ 是 $\mathrm{GL}(n, \mathbb{C})$ 的子群, 因为行列式为1的矩阵均为可逆矩阵。

¹⁸符号*表示这一章节可以跳过, 建议跳转到第129页的章节3.4.2。

(C) 定义正交群与么正群 $O(n)$, $U(n)$ 分别为所有正交矩阵(注意是实矩阵)、酉矩阵的集合。即

$$O(n) = \{A : A \in GL(n, \mathbb{R}), A^T A = AA^T = I\},$$

$$U(n) = \{A : A \in GL(n, \mathbb{C}), A^\dagger A = AA^\dagger = I\}.$$

我们将验证 $O(n)$ 是群, $U(n)$ 的情况类似: 令 $A, B \in O(n)$, 那么

$$(AB)^T(AB) = B^T A^T AB = B^T B = I,$$

$$(A^{-1})^T A^{-1} = (A^T)^{-1} A^{-1} = (AA^T)^{-1} = I,$$

所以 $AB, A^{-1} \in O(n)$ 。很显然, $O(n)$ 是 $U(n)$ 的子群, 因为当 A 为实矩阵时, $A^\dagger = A^T$; 并且毫无疑问, $O(n)$ 是 $GL(n, \mathbb{R})$ 的子群, $U(n)$ 是 $GL(n, \mathbb{C})$ 的子群。

(D) 定义特殊正交群与特殊么正群分别为 $SO(n) = O(n) \cap SL(n, \mathbb{R})$, $SU(n) = U(n) \cap SL(n, \mathbb{C})$ 。即

$$SO(n) = \{A : A \in GL(n, \mathbb{R}), A^T A = AA^T = I, \det A = 1\},$$

$$SU(n) = \{A : A \in GL(n, \mathbb{C}), A^\dagger A = AA^\dagger = I, \det A = 1\}.$$

同样的, 它们是群, 并且 $SO(n)$ 是 $SU(n)$ 的子群; 此外, $SO(n)$ 是 $O(n)$ 的子群, $SU(n)$ 是 $U(n)$ 的子群。

因为群的任意元素都有逆元, 所以矩阵李群的任意元素都是可逆矩阵, 换言之, 任意矩阵李群均为 $GL(n, \mathbb{C})$ 的子群。

对于一个矩阵李群 $G \subseteq GL(n, \mathbb{C})$, 定义集合

$$\text{Lie}(G) = \{X : X \in \mathfrak{gl}(n, \mathbb{C}), \exp tX \in G \text{ for all } t \in \mathbb{R}\},$$

其中 $\mathfrak{gl}(n, \mathbb{C})$ 是所有 n 阶矩阵的集合。很显然 $0 \in \text{Lie}(G)$, 因为 $\exp t \cdot 0 \equiv \exp 0 = I \in G$ 。该集合具有很好的性质:

Theorem 3.7. 对于任意 $X, Y \in \text{Lie}(G)$ 与任意实数 $s \in \mathbb{R}$, 都具有

$$X + Y \in \text{Lie}(G),$$

$$sX \in \text{Lie}(G).$$

Proof. 对于第1个表达式, 根据(3.32)可得

$$\exp t(X + Y) = \lim_{n \rightarrow \infty} \left(\left(\exp \frac{t}{n} X \right) \left(\exp \frac{t}{n} Y \right) \right)^n,$$

对于任意 $t \in \mathbb{R}$, 因为 $X, Y \in \text{Lie}(G)$, 所以根据定义, $\exp \frac{t}{n} X, \exp \frac{t}{n} Y \in G$ 。因为矩阵李群在矩阵乘法运算下封闭, 所以 $\lim_{n \rightarrow \infty} \left(\left(\exp \frac{t}{n} X \right) \left(\exp \frac{t}{n} Y \right) \right)^n \in G$, 即 $\exp t(X + Y) \in G$ 对任意 $t \in \mathbb{R}$ 均成立, 这表明 $X + Y \in \text{Lie}(G)$ 。

对于第2个表达式, 给定 $s \in \mathbb{R}$, 因为 $\exp tX \in G$ 对任意 $t \in \mathbb{R}$ 均成立, 所以 $\exp tsX \in G$ 对任意 $t \in \mathbb{R}$ 均成立, 这表明 $sX \in \text{Lie}(G)$ 。□

但是, $\text{Lie}(G)$ 还有更强的结构:

Theorem 3.8. 对于任意 $X, Y \in \text{Lie}(G)$, 定义李括号为 $[X, Y] = XY - YX$ 。因此

$$[X, Y] \in \text{Lie}(G).$$

类比于 Theorem 2.2 的证明, 容易得出

$$\begin{aligned} [X, Y] &= -[Y, X], \\ [\alpha X + \beta Y, Z] &= \alpha[X, Z] + \beta[Y, Z], \\ [X, \alpha Y + \beta Z] &= \alpha[X, Y] + \beta[X, Z], \\ [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] &= 0, \end{aligned}$$

其中 $\alpha, \beta \in \mathbb{R}$ 。因此, 根据 (2.5), $\text{Lie}(G)$ 是李代数。

Proof. 根据 (3.24) 可得, 对于任意 $t \in \mathbb{R}, A \in G, X \in \text{Lie}(G)$, 因为 $\exp tX \in G$, 所以 $\exp(tAXA^{-1}) = \exp(A(tX)A^{-1}) = A(\exp tX)A^{-1} \in G$, 因此 $AXA^{-1} \in \text{Lie}(G)$ 。这表明, 对于任意 $t \in \mathbb{R}, X, Y \in \text{Lie}(G)$, 因为 $\exp tX \in G$, 所以 $(\exp tX)Y(\exp(-tX)) \in \text{Lie}(G)$ 。定义含时间演化的矩阵

$$\Gamma(t) = (\exp tX)Y(\exp(-tX)),$$

因此, $\Gamma(t) \in \text{Lie}(G)$ 。根据 Theorem 3.4 的第 3 个表达式与 (3.30) 可得

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \Gamma(t) &= \left(\left. \frac{d}{dt} \right|_{t=0} \exp tX \right) Y + Y \left(\left. \frac{d}{dt} \right|_{t=0} \exp(-tX) \right) \\ &= XY + Y(-X) = XY - YX \\ &= [X, Y]. \end{aligned}$$

而根据微分的定义

$$\left. \frac{d}{dt} \right|_{t=0} \Gamma(t) = \lim_{h \rightarrow 0} \frac{\Gamma(h) - \Gamma(0)}{h} = \lim_{h \rightarrow 0} \frac{\Gamma(h) - Y}{h},$$

对于很小的正数 h , $\Gamma(h) \in \text{Lie}(G)$, 但是 $Y \in \text{Lie}(G)$ 。根据 Theorem 3.7 可得, $\text{Lie}(G)$ 是 \mathbb{R} 上的向量空间, 因此 $\frac{\Gamma(h) - Y}{h} \in \text{Lie}(G)$, 这表明 $[X, Y] = \left. \frac{d}{dt} \right|_{t=0} \Gamma(t) \in \text{Lie}(G)$ 。□

我们将 $\text{Lie}(G)$ 称为 **矩阵李群 G 的李代数**。对于任意 $X \in \text{Lie}(G)$, 根据 (3.30), 轨迹 $\gamma(t) = \exp tX$ 在 $t = 0$ 的切向量为 $\gamma'(0) = \left. \frac{d}{dt} \right|_{t=0} \exp tX = X$, 因此, 李代数可以理解为矩阵李群在单位元 I 的处切平面, 如 Figure 3.8 所示。

我们将计算一些矩阵李群的李代数:

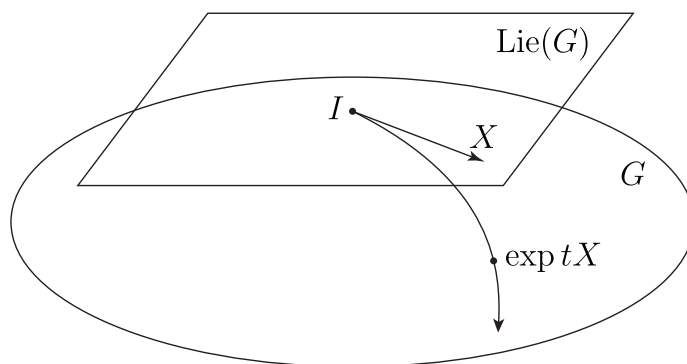


Figure 3.8: 矩阵李群 G 与它的李代数 $\text{Lie}(G)$

Theorem 3.9.

$$\begin{aligned}
 \text{Lie}(\text{GL}(n, \mathbb{R})) &= \mathfrak{gl}(n, \mathbb{R}), \\
 \text{Lie}(\text{GL}(n, \mathbb{C})) &= \mathfrak{gl}(n, \mathbb{C}), \\
 \text{Lie}(\text{SL}(n, \mathbb{R})) &= \mathfrak{sl}(n, \mathbb{R}), \\
 \text{Lie}(\text{SL}(n, \mathbb{C})) &= \mathfrak{sl}(n, \mathbb{C}), \\
 \text{Lie}(\text{O}(n)) &= \text{Lie}(\text{SO}(n)) = \mathfrak{o}(n), \\
 \text{Lie}(\text{U}(n)) &= \mathfrak{u}(n), \\
 \text{Lie}(\text{SU}(n)) &= \mathfrak{su}(n),
 \end{aligned}$$

其中 $\mathfrak{gl}(n, \mathbb{R})$, $\mathfrak{gl}(n, \mathbb{C})$ 是所有 n 阶实矩阵、复矩阵的集合； $\mathfrak{sl}(n, \mathbb{R})$, $\mathfrak{sl}(n, \mathbb{C})$ 是所有 n 阶迹(矩阵对角元素之和)为0的实矩阵、复矩阵的集合； $\mathfrak{o}(n)$, $\mathfrak{u}(n)$ 是所有反对称实矩阵($A^T + A = 0$)、反厄米矩阵($A^\dagger + A = 0$)的集合； $\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$ 。

Proof. 假设 $t \in \mathbb{R}$ 是任意实数。

对于任意实矩阵 $X \in \mathfrak{gl}(n, \mathbb{R})$, $\exp tX$ 是可逆的实矩阵, 其逆元为 $\exp(-tX)$, 因此 $X \in \text{Lie}(\text{GL}(n, \mathbb{R}))$, 这表明 $\mathfrak{gl}(n, \mathbb{R}) \subseteq \text{Lie}(\text{GL}(n, \mathbb{R}))$ 。对于任意矩阵 $X \in \text{Lie}(\text{GL}(n, \mathbb{R}))$, 根据李代数的定义可得, $\exp tX \in \text{GL}(n, \mathbb{R})$ 是可逆实矩阵, 自然是实矩阵; 根据(3.30)可得,

$$X = \left. \frac{d}{dt} \right|_{t=0} \exp tX = \lim_{h \rightarrow 0} \frac{\exp hX - I}{h}$$

是实矩阵的线性组合, 因此 X 是实矩阵, 这表明 $\text{Lie}(\text{GL}(n, \mathbb{R})) \subseteq \mathfrak{gl}(n, \mathbb{R})$ 。这样就证明了 $\text{Lie}(\text{GL}(n, \mathbb{R})) = \mathfrak{gl}(n, \mathbb{R})$ 。Lie($\text{GL}(n, \mathbb{C})$)的情况相同, 不做赘述。

对于任意迹为0的实矩阵 $X \in \mathfrak{sl}(n, \mathbb{R})$, 根据(3.28)可得

$$\det(\exp tX) = e^{\text{tr}(tX)} = e^{t \cdot \text{tr}(X)} = e^0 = 1,$$

并且 $\exp tX$ 是实矩阵, 因此 $X \in \text{Lie}(\text{SL}(n, \mathbb{R}))$, 这表明 $\mathfrak{sl}(n, \mathbb{R}) \subseteq \text{Lie}(\text{SL}(n, \mathbb{R}))$ 。对于任意矩阵 $X \in \text{Lie}(\text{SL}(n, \mathbb{R}))$, $\exp tX \in \text{SL}(n, \mathbb{R})$ 是行列式为1的实矩阵, 而

$$\text{tr}(X) = \frac{d}{dt} \Big|_{t=0} e^{t \cdot \text{tr}(X)} = \frac{d}{dt} \Big|_{t=0} \det(\exp tX) = \frac{d}{dt} \Big|_{t=0} 1 = 0,$$

因此 $X \in \mathfrak{sl}(n, \mathbb{R})$, 这表明 $\text{Lie}(\text{SL}(n, \mathbb{R})) \subseteq \mathfrak{sl}(n, \mathbb{R})$ 。Lie $(\text{SL}(n, \mathbb{C}))$ 的情况相同, 不做赘述。

对于任意反对称实矩阵 $X \in \mathfrak{o}(n)$, 根据(3.22)第2个表达式可得

$$(\exp tX)^T (\exp tX) = (\exp tX^T) (\exp tX) = (\exp(-tX)) (\exp tX) = I,$$

因此 $\exp tX$ 是正交矩阵, 它又是实矩阵, 因此 $X \in \text{Lie}(\text{O}(n))$, 这表明 $\mathfrak{o}(n) \subseteq \text{Lie}(\text{O}(n))$ 。对于任意矩阵 $X \in \text{Lie}(\text{O}(n))$, $\exp tX \in \text{O}(n)$ 是正交矩阵, 因此 $\exp tX^T = (\exp tX)^T = (\exp tX)^{-1} = \exp(-tX)$, 而根据(3.28),

$$X^T = \frac{d}{dt} \Big|_{t=0} \exp tX^T = \frac{d}{dt} \Big|_{t=0} \exp(-tX) = -X,$$

因此 $X \in \mathfrak{o}(n)$, 这表明 $\text{Lie}(\text{O}(n)) \subseteq \mathfrak{o}(n)$ 。Lie $(\text{U}(n))$ 的情况相同。对于Lie $(\text{SO}(n))$, 因为任意反对称实矩阵的对角元素都为0, 所以迹为0, 根据(3.28)可得 $\det(\exp tX) = 1$, 因此上述证明对Lie $(\text{SO}(n))$ 是同样成立的, 不做赘述。Lie $(\text{SU}(n))$ 的原理几乎完全一样, 留给读者自己练习(我读数学书的时候最讨厌这句话了)。□

Theorem 3.9的形式虽然整体而言较为美观, 但仍具瑕疵: 我们可能会想, Lie $(\text{SO}(n))$ 应该是 $\mathfrak{o}(n) \cap \mathfrak{sl}(n, \mathbb{R})$, 但是这一定义完全是多余的: 因为反对称实矩阵的所有对角元素都为0, 因此迹为0, 这表明 $\mathfrak{o}(n) \subseteq \mathfrak{sl}(n, \mathbb{R})$, 因此 $\mathfrak{o}(n) \cap \mathfrak{sl}(n, \mathbb{R}) = \mathfrak{o}(n)$ 。但是, 反厄米矩阵的对角元素可以不为0, 迹同样也不一定为0。比如 $A = \begin{pmatrix} i & i \\ i & i \end{pmatrix}$ 满足 $A^\dagger + A = 0$, 但是 $\text{tr}(A) = 2i \neq 0$ 。

回到第38页的定义。因为 $\mathfrak{su}(n)$ 是 $\mathfrak{u}(n)$ 的线性子空间, $\mathfrak{u}(n)$ 是 $\mathfrak{gl}(n, \mathbb{C})$ 的线性子空间(视为实数域 \mathbb{R} 上的向量空间), 并且这些李代数的李括号定义是相同的, 所以它们具有子代数的关系。但是 $\text{SU}(n)$ 是 $\text{U}(n)$ 的子群, $\text{U}(n)$ 是 $\text{GL}(n, \mathbb{C})$ 的子群, 所以可以引出以下定理:

Theorem 3.10. 对于矩阵李群 G, H , 如果 H 是 G 的子群, 那么Lie (H) 是Lie (G) 的子代数。

但是可惜的是, 其证明超出本文范围, 在此不做展开。矩阵李群与李代数的关系远不止这些, 我们将叙述一些更加美妙的定理, 但是均不做证明:

Theorem 3.11. 对于矩阵李群 G, H 与群同态映射 $\Phi: G \rightarrow H$ (对于任意 $X, Y \in G$, 都具有 $\Phi(XY) = \Phi(X)\Phi(Y)$), 那么存在李代数同态 $\Phi_*: \text{Lie}(G) \rightarrow \text{Lie}(H)$, 使得Figure 3.9交换。换句话说, 对于任意 $X \in \text{Lie}(G)$, 都具有 $\exp(\Phi_*(X)) = \Phi(\exp(X)) \in H$ 。

$$\begin{array}{ccc}
\mathrm{Lie}(G) & \xrightarrow{\Phi_*} & \mathrm{Lie}(H) \\
\exp \downarrow & & \downarrow \exp \\
G & \xrightarrow{\Phi} & H
\end{array}$$

Figure 3.9: 交换图

我们定义**正规子群**为条件更强的子群：对于群 G 与 G 的子群 H ，如果对于任意 $g \in G, h \in H$ 都具有 $ghg^{-1} \in H$ ，那么 H 是 G 的正规子群。而第38页定义的“理想”，很显然条件强于“子代数”。据此，我们有比Theorem 3.10条件更强的定理：

Theorem 3.12. 对于矩阵李群 G, H ，如果 G, H 是连通的，那么 H 是 G 的正规子群当且仅当 $\mathrm{Lie}(H)$ 是 $\mathrm{Lie}(G)$ 的理想。

但是我们并没有定义**连通**的具体含义，其定义为：如果矩阵李群 G 是连通的，那么对于任意元素 $X, Y \in G$ ，存在含时间演化的矩阵 $A(t), t \in [0, 1]$ 使得 $A(t)$ 是连续函数(即 $A(t)$ 的每个元素 $A_j^i(t)$ 均为连续函数)，并满足 $A(0) = X, A(1) = Y$ 。通俗来说，任意两点都能通过一条“道路”连接。

但是我们可以给出等价的定义，使得有关的证明更为简便：对于任意元素 $X, Y, Z \in G$ ，如果 $A(t)$ 将 X, Y 相连， $B(t)$ 将 Y, Z 相连，那么 $A(0) = X, A(1) = B(0) = Y, B(1) = Z$ 。可以定义

$$C(t) = \begin{cases} A(2t), & t \in [0, 1/2], \\ B(2t - 1), & t \in [1/2, 1], \end{cases}$$

很显然 $C(t)$ 将 X, Z 相连。因此， G 是连通的，当且仅当对于任意元素 $X \in G$ ，都存在连续函数 $A(t)$ ，使得 $A(0) = X, A(1) = I$ ，其中 I 是单位矩阵。

接下来，我们将分析一些矩阵李群的连通性：

Theorem 3.13. $\mathrm{GL}(n, \mathbb{C}), \mathrm{SL}(n, \mathbb{C}), \mathrm{U}(n), \mathrm{SU}(n), \mathrm{SO}(n)$ 均是连通的，但是 $\mathrm{O}(n)$ 不是连通的。

Proof. 线性代数可以证明(在此不做展开)，任意矩阵都相似于上三角矩阵，即对于任意矩阵 A ，都存在上三角矩阵 B 与可逆矩阵 C 使得 $A = CBC^{-1}$ ，将 B 记为

$$B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

并且，如果 $A \in \mathrm{GL}(n, \mathbb{C})$ 是可逆矩阵，那么任意 λ_i 都是非零的(否则 A 的行列式为0，就不是可逆矩阵)。令 $B(t)$ 为含时间演化的矩阵，其中 $B(0) = B, B(1/2) = \mathrm{diag}(\lambda_1, \dots, \lambda_n), B(1)$

$= I$, 对于 $t \in [0, 1/2]$, 我们只需将 B 的右上角(不包含对角线)的元素与 $1 - 2t$ 相乘即可; 对于 $t \in [1/2, 1]$,

$$B(t) = \begin{pmatrix} 2(1 - \lambda_1)t + 2\lambda_1 - 1 & & * \\ & \ddots & \\ 0 & & 2(1 - \lambda_n)t + 2\lambda_n - 1 \end{pmatrix}.$$

定义 $A(t) = CB(t)C^{-1}$, 因此 $A(0) = CB(0)C^{-1} = CBC^{-1} = A$, $A(1) = CC^{-1} = I$ 。这证明了我们可以定义连续函数 $A(t)$ 将任意可逆矩阵 A 与单位矩阵 I 相连, 因此 $\text{GL}(n, \mathbb{C})$ 是连通的。

对于 $\text{SL}(n, \mathbb{C})$, 情况类似, 同样可以定义 $B(t)$ 为含时间演化的矩阵, 其中 $B(0) = B$, $B(1/2) = \text{diag}(\lambda_1, \dots, \lambda_n)$, $B(1) = I$, 只是此时需要保证 $A(t) = CB(t)C^{-1}$ 的行列式恒等于 1, 这表明 $\lambda_1 \cdots \lambda_n = 1$ 。在 $t \in [0, 1/2]$ 时, 因为对角元素没有发生变化, 所以 $\det A(t) = (\det C)(\det B(t))(\det C^{-1}) = \det B(t) \equiv \lambda_1 \cdots \lambda_n = 1$ 对于 $t \in [0, 1/2]$ 恒成立。在 $t \in [1/2, 1]$ 时, 情况稍微严苛一些, 需要令

$$B(t) = \begin{pmatrix} \lambda_1(t) & & * \\ & \ddots & \\ 0 & & \lambda_n(t) \end{pmatrix},$$

对应的 $\lambda_1(t) \cdots \lambda_n(t) \equiv 1$ 。这个同样不难, 我们可以令 $\lambda_1(t), \dots, \lambda_{n-1}(t) \in \mathbb{C}^*$ 使得 $\lambda_1(0), \dots, \lambda_{n-1}(0) = \lambda_1, \dots, \lambda_{n-1}$; $\lambda_1(1), \dots, \lambda_{n-1}(1) = 1, \dots, 1$, 其中 $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, 相当于在挖去原点的复平面上找到任意复数与 1 之间的轨迹。对于 $\lambda_n(t)$, 只需令 $\lambda_n(t) = 1/(\lambda_1(t) \cdots \lambda_{n-1}(t))$ 即可, 因为 $\lambda_1(t), \dots, \lambda_{n-1}(t)$ 始终非零, 所以 $\lambda_n(t)$ 始终非零, 并且是连续函数, 这样就能满足条件。

而任意酉矩阵都可以被分解为 $U = V\Lambda V^{-1}$, 其中 V 也是酉矩阵(在此不做证明), Λ 是对角矩阵, 其对角元素可以表示为 $e^{i\lambda_1}, \dots, e^{i\lambda_n}$ 的形式, 其中 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ 。对于任意 $t \in [0, 1]$, 定义

$$U(t) = V\Lambda(t)V^{-1} = V \begin{pmatrix} e^{(1-t)i\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{(1-t)i\lambda_n} \end{pmatrix} V^{-1},$$

很显然 $\Lambda(t)$ 在任意时刻都是酉矩阵。因此 $U(t)$ 在任意时刻都是酉矩阵, 其原因为

$$U(t)^\dagger = (V^{-1})^\dagger \Lambda(t)^\dagger V^\dagger = (V^\dagger)^\dagger \Lambda(t)^{-1} V^{-1} = V\Lambda(t)^{-1} V^{-1} = (V\Lambda(t)V^{-1})^{-1} = U(t)^{-1}.$$

因此 $U(t)$ 将任意酉矩阵 U 与单位矩阵 I 相连。

对于 $\text{SU}(n)$, 情况还是类似, 对于任意 $t \in [0, 1]$, 令

$$U(t) = V\Lambda(t)V^{-1} = V \begin{pmatrix} e^{i\lambda_1(t)} & & 0 \\ & \ddots & \\ 0 & & e^{i\lambda_n(t)} \end{pmatrix} V^{-1},$$

其中 $\lambda_1(t), \dots, \lambda_n(t) \in \mathbb{R}, \lambda_1(0), \dots, \lambda_n(0) = \lambda_1, \dots, \lambda_n; \lambda_1(1), \dots, \lambda_n(1) = 0, \dots, 0$ 。很显然 $\Lambda(t)$ 在任意时刻都是酉矩阵。因为 $U(t)$ 的行列式为 $e^{i(\lambda_1(t) + \dots + \lambda_n(t))}$ ，为了确保行列式始终为1，我们需要让 $\lambda_1(t) + \dots + \lambda_n(t)$ 始终为0。这同样是很容易的，只需连续调节 $\lambda_1(t), \dots, \lambda_{n-1}(t)$ 从 $\lambda_1, \dots, \lambda_{n-1}$ 到 $0, \dots, 0$ ，然后让 $\lambda_n(t) = -\lambda_1(t) - \dots - \lambda_{n-1}(t)$ 即可。

与前面不同的是， $SO(n)$ 连通性的证明具有很强的几何意义。因为 $SO(n)$ 的任意元素都是行列式为1的正交矩阵，所以我们可以将它视为“旋转” \mathbb{R}^n 的标准正交基。同样道理，我们还是需要证明，存在 $A(t)$ 将任意矩阵 $A \in SO(n)$ 与单位矩阵 I 相连。记任意矩阵 $A \in SO(n)$ 为

$$A = (v_1 \quad \cdots \quad v_n),$$

其中 $v_1, \dots, v_n \in \mathbb{R}^n$ 为列向量，满足 $v_i \cdot v_j = \delta_{ij}$ 。同时，记

$$e_1 = (1, 0, \dots, 0)^T, e_2 = (0, 1, 0, \dots, 0)^T, \dots, e_n = (0, \dots, 0, 1)^T$$

为单位矩阵 I 的 n 个列向量。在 $t \in [0, 1/(n-1)]$ 时，我们可以将 v_1 转到 e_1 的位置；在 $t \in [1/(n-1), 2/(n-1)]$ 时，固定 v_1 不动，将 v_2 转到 e_2 的位置；在 $t \in [2/(n-1), 3/(n-1)]$ 时，固定 v_1, v_2 不动，将 v_3 转到 e_3 的位置。……以此类推，最后，在 $t \in [(n-2)/(n-1), 1]$ 时，固定 v_1, v_2, \dots, v_{n-2} 不动，将 v_{n-1} 转到 e_{n-1} 的位置，此时根据 A 的标准正交性， v_n 自动与 e_n 重合，因此我们完成了操作，如Figure 3.10所示，我们给出了 $n=3$ 的图解。

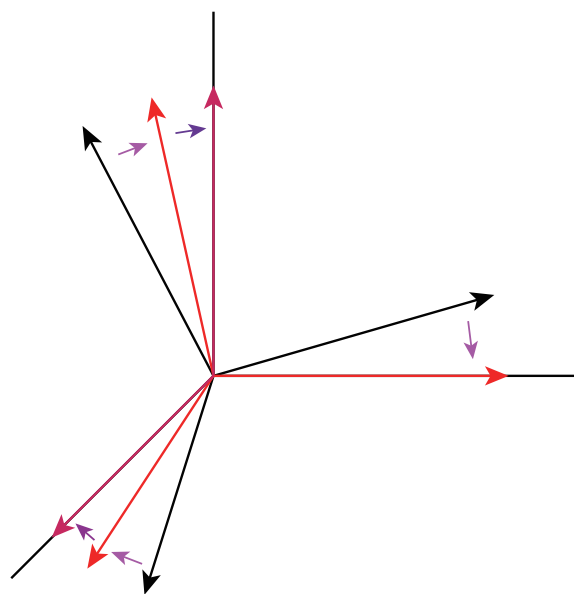
但是对于 $O(n)$ ，情况就完全不同了。根据 $O(n)$ 的定义，可得任意 $A \in O(n)$ 都满足 $A^T A = I$ ，因此 $\det(A^T A) = (\det A)^2 = 1$ ，这表明 $\det A = \pm 1$ 。选取 A 使得 $\det A = -1$ ，假设存在连续函数 $A(t)$ 使得 $A(0) = A, A(1) = I$ 。因为 $\det A(t)$ 是次数为 $n!$ 的多项式函数，换句话说，它由有限个 $A_j^i(t)$ 组成，所以 $\det A(t)$ 是连续函数。而 $\det A(0) = \det A = -1, \det A(1) = \det I = 1$ ， $O(n)$ 任意元素的行列式只能为 ± 1 ，而不能是中间的任意值，所以这与 $\det A(t)$ 是连续函数的事实矛盾。□

根据 $\det A = \pm 1$ 的性质，我们可以将 $O(n)$ 分为2个集合的并集：

$$\begin{aligned} O(n) &= \{A : A \in O(n), \det A = 1\} \cup \{A : A \in O(n), \det A = -1\} \\ &= SO(n) \cup \{A : A \in O(n), \det A = -1\}, \end{aligned}$$

我们已经证明 $SO(n)$ 是连通的；对于后面的集合，它同样也是连通的：我们只需将1个向量的取值取反即可，相当于“镜像”。因此， $O(n)$ 由2个连通的集合组成：利用严格的数学术语，我们称 $O(n)$ 的**连通分量**有2个；并且， $SO(n)$ 对应其中的1个连通分量，这相当于 $SO(n)$ 不仅包含于 $O(n)$ ，而且与 $O(n)$ 的某一部分完全“重合”！利用(3.29)定义的Hilbert范数的概念，如果我们分别在 $O(n), SO(n)$ 单位元的周围绘制“半径”为 R 的“圆”，对应集合

$$\begin{aligned} U &= \{A : A \in O(n), \|A - I\| \leq R\}, \\ V &= \{A : A \in SO(n), \|A - I\| \leq R\}, \end{aligned}$$



在 $t = [0, 1/2]$ 时, 将 v_1 转到 e_1
 在 $t = [1/2, 1]$ 时, 固定 v_1 不动, 将 v_2 转到 e_2 ,
 此时 v_3 自动与 e_3 重合

Figure 3.10: $SO(n)$ 的连通性

那么 U, V 是“相同”的！严格来说，利用第141页的定义，我们称 U, V 是同胚的。但是，对于 $U(n)$ 与 $SU(n)$ ，情况就没有这么理想了，因为它们都是连通的，而且很显然 $SU(n) \subsetneq U(n)$ 。光是文字叙述可能无法清楚表达我的意思，所以我就给出图解，如Figure 3.11所示。

利用严格的数学术语，我们称 $O(n), SO(n)$ 为嵌入 $GL(n, \mathbb{R})$ 的 $\frac{n(n-1)}{2}$ 维子流形； $U(n)$ 为嵌入 $GL(n, \mathbb{C})$ 的 n^2 维子流形， $SU(n)$ 为嵌入 $GL(n, \mathbb{C})$ 的 $n(n-1)$ 维子流形。篇幅受限，我们虽然不会解释这些数学术语的具体含义，但是通过它们，这一现象可以得到比较直观的反映：因为 $O(n), SO(n)$ 的维数相同，并且 $SO(n) \subseteq O(n)$ ，所以在单位元 I 附近，它们是完全重合的；因为 $U(n)$ 的维数大于 $SU(n)$ ，所以在单位元 I 附近，看起来像 $SU(n)$ 嵌入 $U(n)$ 。

第122页的叙述表明，我们可以将矩阵李群的李代数视为矩阵李群在单位元 I 处的切平面。根据几何上的理解，切平面只与单位元 I 附近的区域有关：正如计算函数 $y = f(x)$ 在给定点 $x = x_0$ 处的切线，我们只需知道 $f(x)$ 在 x_0 处的取值与一阶导数即可，而对于一阶导数，我们只需知道 $f(x)$ 在 x_0 附近的取值即可。而在单位元 I 附近， $O(n), SO(n)$ 是完全重合的，因此，它们在单位元 I 的切平面是完全相同的。这表明， $O(n), SO(n)$ 的李代数是完全相同的！当然，Theorem 3.9已经计算过了。

行文至此，这一部分也应该画上句号，否则根本没完没了。但是，有关这一部分的内容远远不止这些：如果你学过微分流形理论，那么你可以把“矩阵李群”前面的“矩

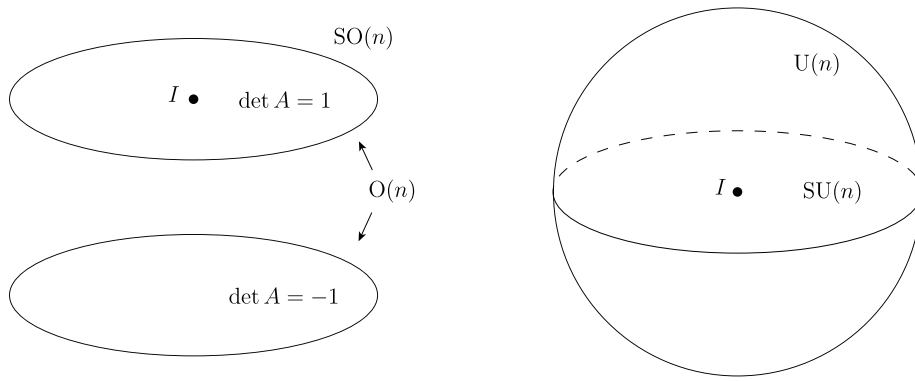


Figure 3.11: $O(n)$, $SO(n)$, $U(n)$, $SU(n)$ 的直观理解

阵”这两个字去除，将这些定理推广到广义的**李群**上。矩阵李群是李群的一种，但是李群的种类远远不止矩阵李群；并且结合微分流形理论，李群的几何意义远远强于矩阵李群，但是在此我们就不做介绍了。幸运的是，对于李代数的定义，我们在本文已经完善了：同样的，每个李群都具有对应的李代数，我们还是可以将其理解为李群的切平面。

3.4.2 自旋动力学

在矩阵表示下，薛定谔方程 $H\Psi = i\hbar \frac{\partial \Psi}{\partial t}$ 将表示为

$$H \begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix} = i\hbar \begin{pmatrix} \dot{c}_0(t) \\ \dot{c}_1(t) \end{pmatrix}.$$

其中 H 是哈密顿算子。将上述表达式化为

$$\begin{pmatrix} \dot{c}_0(t) \\ \dot{c}_1(t) \end{pmatrix} = -\frac{iH}{\hbar} \begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix},$$

并且记初始时刻的波函数向量表示为

$$\begin{pmatrix} c_0(0) \\ c_1(0) \end{pmatrix},$$

可得

$$\begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix} = \exp\left(-\frac{iH}{\hbar}t\right) \begin{pmatrix} c_0(0) \\ c_1(0) \end{pmatrix}. \quad (3.33)$$

而哈密顿算子的本征值代表粒子的能量，记其本征值为 E_0, E_1 ，本征态为 $\Psi_0 = (\Psi_0^0, \Psi_0^1)^T$, $\Psi_1 = (\Psi_1^0, \Psi_1^1)^T$ ，并且满足 $\Psi_i \cdot \Psi_j = \overline{\Psi_i^0} \Psi_j^0 + \overline{\Psi_i^1} \Psi_j^1 = \delta_{ij}$ ，因此

$$-\frac{iH}{\hbar}t = \begin{pmatrix} \Psi_0^0 & \Psi_1^0 \\ \Psi_0^1 & \Psi_1^1 \end{pmatrix} \begin{pmatrix} -\frac{iE_0}{\hbar}t & 0 \\ 0 & -\frac{iE_1}{\hbar}t \end{pmatrix} \begin{pmatrix} \overline{\Psi_0^0} & \overline{\Psi_0^1} \\ \overline{\Psi_1^0} & \overline{\Psi_1^1} \end{pmatrix}.$$

根据(3.27)可得

$$\exp\left(-\frac{iH}{\hbar}t\right) = \begin{pmatrix} \Psi_0^0 & \Psi_1^0 \\ \Psi_0^1 & \Psi_1^1 \end{pmatrix} \begin{pmatrix} e^{-iE_0t/\hbar} & 0 \\ 0 & e^{-iE_1t/\hbar} \end{pmatrix} \begin{pmatrix} \overline{\Psi_0^0} & \overline{\Psi_1^0} \\ \overline{\Psi_0^1} & \overline{\Psi_1^1} \end{pmatrix}.$$

因此

$$\begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix} = \begin{pmatrix} \Psi_0^0 & \Psi_1^0 \\ \Psi_0^1 & \Psi_1^1 \end{pmatrix} \begin{pmatrix} e^{-iE_0t/\hbar} & 0 \\ 0 & e^{-iE_1t/\hbar} \end{pmatrix} \begin{pmatrix} \overline{\Psi_0^0} & \overline{\Psi_1^0} \\ \overline{\Psi_0^1} & \overline{\Psi_1^1} \end{pmatrix} \begin{pmatrix} c_0(0) \\ c_1(0) \end{pmatrix}. \quad (3.34)$$

3.5 2个粒子的自旋态

对于2个粒子A,B, 记它们的自旋算子分别为 $S_1^A, S_2^A, S_3^A; S_1^B, S_2^B, S_3^B$, 其矩阵表示分别为

$$\begin{aligned} S_1^A &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & S_2^A &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & S_3^A &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ S_1^B &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & S_2^B &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & S_3^B &\longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

以及 S_3^A, S_3^B 的本征向量表示分别为

$$\begin{aligned} |\uparrow\rangle^A &\longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & |\downarrow\rangle^A &\longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ |\uparrow\rangle^B &\longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & |\downarrow\rangle^B &\longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

对应的本征值为

$$\begin{aligned} S_3^A |\uparrow\rangle^A &= \frac{\hbar}{2} |\uparrow\rangle^A, \\ S_3^A |\downarrow\rangle^A &= -\frac{\hbar}{2} |\downarrow\rangle^A, \\ S_3^B |\uparrow\rangle^B &= \frac{\hbar}{2} |\uparrow\rangle^B, \\ S_3^B |\downarrow\rangle^B &= -\frac{\hbar}{2} |\downarrow\rangle^B. \end{aligned}$$

定义**Kronecker Product**运算 $\otimes : M(m \times n, \mathbb{C}) \times M(p \times q, \mathbb{C}) \rightarrow M(mp \times nq, \mathbb{C})$ 为

$$\begin{aligned}
 & \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \otimes \begin{pmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & \ddots & \vdots \\ B_{p1} & \cdots & B_{pq} \end{pmatrix} \\
 &= \begin{pmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{m1}B & \cdots & A_{mn}B \end{pmatrix} \\
 &= \begin{pmatrix} A_{11}B_{11} & \cdots & A_{11}B_{1q} & \cdots & A_{1n}B_{11} & \cdots & A_{1n}B_{1q} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ A_{11}B_{p1} & \cdots & A_{11}B_{pq} & \cdots & A_{1n}B_{p1} & \cdots & A_{1n}B_{pq} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_{m1}B_{11} & \cdots & A_{m1}B_{1q} & \cdots & A_{mn}B_{11} & \cdots & A_{mn}B_{1q} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ A_{m1}B_{p1} & \cdots & A_{m1}B_{pq} & \cdots & A_{mn}B_{p1} & \cdots & A_{mn}B_{pq} \end{pmatrix}.
 \end{aligned}$$

2个粒子组成的自旋波函数可以表示为任意非零线性组合

$$\Psi = c_0 |\uparrow\uparrow\rangle + c_1 |\uparrow\downarrow\rangle + c_2 |\downarrow\uparrow\rangle + c_3 |\downarrow\downarrow\rangle,$$

其中

$$\begin{aligned}
 |\uparrow\uparrow\rangle &= |\uparrow\rangle^A \otimes |\uparrow\rangle^B \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\
 |\uparrow\downarrow\rangle &= |\uparrow\rangle^A \otimes |\downarrow\rangle^B \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\
 |\downarrow\uparrow\rangle &= |\downarrow\rangle^A \otimes |\uparrow\rangle^B \longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
 |\downarrow\downarrow\rangle &= |\downarrow\rangle^A \otimes |\downarrow\rangle^B \longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
 \end{aligned}$$

因此, Ψ 可以被表示为向量

$$c_0 |\uparrow\uparrow\rangle + c_1 |\uparrow\downarrow\rangle + c_2 |\downarrow\uparrow\rangle + c_3 |\downarrow\downarrow\rangle \longleftrightarrow \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

其中 $|c_0|^2 + |c_1|^2 + |c_2|^2 + |c_3|^2 = 1$ 。

定义总自旋算子 $S_3(A, B)$ 与总自旋平方算子 $S^2(A, B)$ 为

$$S_3(A, B) = S_3^A \otimes I + I \otimes S_3^B,$$

$$S^2(A, B) = (S_1^A \otimes I + I \otimes S_1^B)^2 + (S_2^A \otimes I + I \otimes S_2^B)^2 + (S_3^A \otimes I + I \otimes S_3^B)^2.$$

其中

$$S_3^A \otimes I \longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

$$I \otimes S_3^B \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

因此

$$S_3(A, B) \longleftrightarrow \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

同理可得

$$S_1^A \otimes I + I \otimes S_1^B \longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

$$S_2^A \otimes I + I \otimes S_2^B \longleftrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i & -i & 0 \\ i & 0 & 0 & -i \\ i & 0 & 0 & -i \\ 0 & i & i & 0 \end{pmatrix}.$$

因此

$$\begin{aligned}
 S^2(A, B) &\longleftrightarrow \frac{\hbar^2}{4} \left(\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}^2 + \begin{pmatrix} 0 & -i & -i & 0 \\ i & 0 & 0 & -i \\ i & 0 & 0 & -i \\ 0 & i & i & 0 \end{pmatrix}^2 + \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}^2 \right) \\
 &= \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.
 \end{aligned}$$

$S_3(A, B)$ 与 $S^2(A, B)$ 的本征向量表示满足

$$\begin{aligned}
 \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} &= \lambda \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix}, \\
 \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix} &= \mu \begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix}.
 \end{aligned}$$

定义

$$\begin{aligned}
 \chi_{1,1} &= |\uparrow\uparrow\rangle \longleftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\
 \chi_{1,0} &= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \\
 \chi_{1,-1} &= |\downarrow\downarrow\rangle \longleftrightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \\
 \chi_{0,0} &= \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.
 \end{aligned} \tag{3.35}$$

可以发现，它们同时为 $S_3(A, B)$ 与 $S^2(A, B)$ 的本征态，并且是标准正交的。其本征值分别为

$$\begin{aligned}
S_3(A, B)\chi_{1,1} &= \hbar\chi_{1,1}, \\
S_3(A, B)\chi_{1,0} &= 0, \\
S_3(A, B)\chi_{1,-1} &= -\hbar\chi_{1,-1}, \\
S_3(A, B)\chi_{0,0} &= 0, \\
S^2(A, B)\chi_{1,1} &= 2\hbar^2\chi_{1,1}, \\
S^2(A, B)\chi_{1,0} &= 2\hbar^2\chi_{1,0}, \\
S^2(A, B)\chi_{1,-1} &= 2\hbar^2\chi_{1,-1}, \\
S^2(A, B)\chi_{0,0} &= 0.
\end{aligned} \tag{3.36}$$

可以得出， $\chi_{s,m}$ 的第一个下标 s ，对应 $S^2(A, B)$ 的本征值为 $s(s+1)\hbar^2$ ；第二个下标 m ，对应 $S_3(A, B)$ 的本征值为 $m\hbar$ 。对于 $S^2(A, B)$ ，由于前三者的本征值均为 $2\hbar^2$ ，我们将前三者称为**三重态**。后者的本征值为0，与前三者不同，我们将后者称为**单态**。

$S^2(A, B), S_3(A, B)$ 本征值的物理意义是A,B粒子组成系统的总角动量平方与总角动量 z 分量。

3.6 氢原子的自旋-轨道耦合*

¹⁹ 如果不考虑氢原子核外电子的自旋，那么氢原子波函数正是满足表达式(3.4)的 $\Psi_{n,l,m}$ 。但是，如果考虑了氢原子核外电子的自旋，那么波函数将会发生些微的改变，把这一现象称为**自旋-轨道耦合**。我们将在这一小节将对其进行简要的分析。

根据玻尔模型，氢原子的核外电子绕着氢原子核公转。我们可以将核外电子类比为地球，氢原子核类比为太阳，(3.1)定义的角动量算子 L_1, L_2, L_3 描述了氢原子核外电子公转的角动量各方向的分量，(3.8)定义的自旋算子 S_1, S_2, S_3 描述了氢原子核外电子自旋角动量各方向的分量。

定义总角动量算子 J_3 与总角动量平方算子 J^2 为

$$\begin{aligned}
J_3 &= L_3 \otimes I + I \otimes S_3, \\
J^2 &= (L_1 \otimes I + I \otimes S_1)^2 + (L_2 \otimes I + I \otimes S_2)^2 + (L_3 \otimes I + I \otimes S_3)^2.
\end{aligned}$$

¹⁹符号*表示这一章节可以跳过，建议跳转到第138页的章节3.6.1。

利用 $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, 我们可以将 J^2 进行化简:

$$\begin{aligned}
J^2 &= \sum_{i=1}^3 (L_i \otimes I + I \otimes S_i)^2 \\
&= \sum_{i=1}^3 (L_i \otimes I)^2 + (L_i \otimes I)(I \otimes S_i) + (I \otimes S_i)(L_i \otimes I) + (I \otimes S_i)^2 \\
&= \sum_{i=1}^3 L_i^2 \otimes I + 2L_i \otimes S_i + I \otimes S_i^2 \\
&= L^2 \otimes I + I \otimes S^2 + 2 \sum_{i=1}^3 L_i \otimes S_i,
\end{aligned} \tag{3.37}$$

其中 $L^2 = L_1^2 + L_2^2 + L_3^2, S^2 = S_1^2 + S_2^2 + S_3^2$ 。

定义

$$\begin{aligned}
|l, m, \frac{1}{2}\rangle_0 &= \frac{1}{\sqrt{2l+1}} \left(\sqrt{l+m+1} \Psi_{n,l,m} \otimes \chi_0(3) + \sqrt{l-m} \Psi_{n,l,m+1} \otimes \chi_1(3) \right), \\
|l, m, \frac{1}{2}\rangle_1 &= \frac{1}{\sqrt{2l+1}} \left(-\sqrt{l-m} \Psi_{n,l,m} \otimes \chi_0(3) + \sqrt{l+m+1} \Psi_{n,l,m+1} \otimes \chi_1(3) \right),
\end{aligned}$$

根据(3.4),(3.7)可得

$$\begin{aligned}
&(L_3 \otimes I) |l, m, \frac{1}{2}\rangle_0 \\
&= \frac{1}{\sqrt{2l+1}} \left(\sqrt{l+m+1} (L_3 \Psi_{n,l,m}) \otimes (I \chi_0(3)) + \sqrt{l-m} (L_3 \Psi_{n,l,m+1}) \otimes (I \chi_1(3)) \right) \\
&= \frac{m\hbar}{\sqrt{2l+1}} \sqrt{l+m+1} \Psi_{n,l,m} \otimes \chi_0(3) + \frac{(m+1)\hbar}{\sqrt{2l+1}} \sqrt{l-m} \Psi_{n,l,m+1} \otimes \chi_1(3), \\
&(I \otimes S_3) |l, m, \frac{1}{2}\rangle_0 \\
&= \frac{1}{\sqrt{2l+1}} \left(\sqrt{l+m+1} (I \Psi_{n,l,m}) \otimes (S_3 \chi_0(3)) + \sqrt{l-m} (I \Psi_{n,l,m+1}) \otimes (S_3 \chi_1(3)) \right) \\
&= \frac{\hbar}{2\sqrt{2l+1}} \left(\sqrt{l+m+1} \Psi_{n,l,m} \otimes \chi_0(3) - \sqrt{l-m} \Psi_{n,l,m+1} \otimes \chi_1(3) \right),
\end{aligned}$$

因此

$$J_3 |l, m, \frac{1}{2}\rangle_0 = \left(m + \frac{1}{2} \right) \hbar |l, m, \frac{1}{2}\rangle_0.$$

同理可得

$$J_3 |l, m, \frac{1}{2}\rangle_1 = \left(m - \frac{1}{2} \right) \hbar |l, m, \frac{1}{2}\rangle_1.$$

对于 J^2 ，直接计算非常麻烦，我们先引入2个线性算子

$$\begin{aligned} L_+ &= L_1 + iL_2, \\ L_- &= L_1 - iL_2, \end{aligned}$$

可以得出(在这里不做展开)，并结合(3.3)辅助理解，当 $-l+1 \leq m \leq l-1$ 时，

$$\begin{aligned} L_+ \Psi_{n,l,m} &= \hbar \sqrt{(l+m+1)(l-m)} \Psi_{n,l,m+1}, \\ L_- \Psi_{n,l,m} &= \hbar \sqrt{(l-m+1)(l+m)} \Psi_{n,l,m-1}, \end{aligned}$$

当 $m = l$ 时，

$$L_+ \Psi_{n,l,m} = 0,$$

当 $m = -l$ 时，

$$L_- \Psi_{n,l,m} = 0.$$

因此，当 $-l+1 \leq m \leq l-1$ 时，

$$\begin{aligned} L_1 \Psi_{n,l,m} &= \frac{1}{2}(L_+ + L_-) \Psi_{n,l,m} \\ &= \frac{\hbar}{2}(\sqrt{(l+m+1)(l-m)} \Psi_{n,l,m+1} + \sqrt{(l-m+1)(l+m)} \Psi_{n,l,m-1}), \\ L_2 \Psi_{n,l,m} &= -\frac{i}{2}(L_+ - L_-) \Psi_{n,l,m} \\ &= -\frac{i\hbar}{2}(\sqrt{(l+m+1)(l-m)} \Psi_{n,l,m+1} - \sqrt{(l-m+1)(l+m)} \Psi_{n,l,m-1}), \end{aligned} \tag{3.38}$$

当 $m = \pm l$ 时，仅需令对应的项为0即可，但是与本文无关，在此不做展开。

根据(3.4),(3.7),(3.9),(3.21)与(3.38)可得(仅分析 $-l+1 \leq m \leq l-2$ 的情况, $m = \pm l$ 或 $m = l-1$ 的情况类似, 不做展开),

$$\begin{aligned}
& (L^2 \otimes I) |l, m, \frac{1}{2}\rangle_0 \\
&= \frac{1}{\sqrt{2l+1}} \left(\sqrt{l+m+1} (L^2 \Psi_{n,l,m}) \otimes (I \chi_0(3)) + \sqrt{l-m} (L^2 \Psi_{n,l,m+1}) \otimes (I \chi_1(3)) \right) \\
&= \frac{l(l+1)\hbar^2}{\sqrt{2l+1}} \left(\sqrt{l+m+1} \Psi_{n,l,m} \otimes \chi_0(3) + \sqrt{l-m} \Psi_{n,l,m+1} \otimes \chi_1(3) \right) \\
&= l(l+1)\hbar^2 |l, m, \frac{1}{2}\rangle_0, \\
& (L^2 \otimes I) |l, m, \frac{1}{2}\rangle_1 \\
&= \frac{1}{\sqrt{2l+1}} \left(-\sqrt{l-m} (L^2 \Psi_{n,l,m}) \otimes (I \chi_0(3)) + \sqrt{l+m+1} (L^2 \Psi_{n,l,m+1}) \otimes (I \chi_1(3)) \right) \\
&= \frac{l(l+1)\hbar^2}{\sqrt{2l+1}} \left(-\sqrt{l-m} \Psi_{n,l,m} \otimes \chi_0(3) + \sqrt{l+m+1} \Psi_{n,l,m+1} \otimes \chi_1(3) \right) \\
&= l(l+1)\hbar^2 |l, m, \frac{1}{2}\rangle_1,
\end{aligned}$$

类似的, 经过复杂的计算, 我们可以得到

$$\begin{aligned}
(I \otimes S^2) |l, m, \frac{1}{2}\rangle_0 &= \frac{3\hbar^2}{4} |l, m, \frac{1}{2}\rangle_0, \\
(I \otimes S^2) |l, m, \frac{1}{2}\rangle_1 &= \frac{3\hbar^2}{4} |l, m, \frac{1}{2}\rangle_1, \\
(L_1 \otimes S_1 + L_2 \otimes S_2 + L_3 \otimes S_3) |l, m, \frac{1}{2}\rangle_0 &= \frac{l\hbar^2}{2} |l, m, \frac{1}{2}\rangle_0, \\
(L_1 \otimes S_1 + L_2 \otimes S_2 + L_3 \otimes S_3) |l, m, \frac{1}{2}\rangle_1 &= -\frac{(l+1)\hbar^2}{2} |l, m, \frac{1}{2}\rangle_1,
\end{aligned}$$

因此, 根据(3.37)可得

$$\begin{aligned}
J^2 |l, m, \frac{1}{2}\rangle_0 &= \left(l + \frac{1}{2}\right) \left(l + \frac{3}{2}\right) \hbar^2 |l, m, \frac{1}{2}\rangle_0, \\
J^2 |l, m, \frac{1}{2}\rangle_1 &= \left(l - \frac{1}{2}\right) \left(l + \frac{1}{2}\right) \hbar^2 |l, m, \frac{1}{2}\rangle_1.
\end{aligned}$$

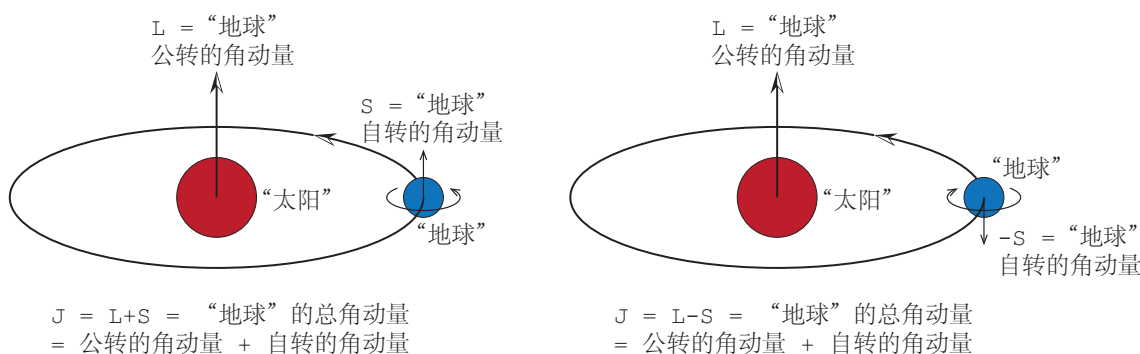


Figure 3.12: 总角动量

综上所述

$$\begin{aligned}
 J_3 |l, m, \frac{1}{2}\rangle_0 &= \left(m + \frac{1}{2}\right) \hbar |l, m, \frac{1}{2}\rangle_0, \\
 J_3 |l, m, \frac{1}{2}\rangle_1 &= \left(m - \frac{1}{2}\right) \hbar |l, m, \frac{1}{2}\rangle_1, \\
 J^2 |l, m, \frac{1}{2}\rangle_0 &= \left(l + \frac{1}{2}\right) \left(l + \frac{3}{2}\right) \hbar^2 |l, m, \frac{1}{2}\rangle_0, \\
 J^2 |l, m, \frac{1}{2}\rangle_1 &= \left(l - \frac{1}{2}\right) \left(l + \frac{1}{2}\right) \hbar^2 |l, m, \frac{1}{2}\rangle_1.
 \end{aligned}$$

因此， $|l, m, \frac{1}{2}\rangle_0, |l, m, \frac{1}{2}\rangle_1$ 同时为 J_3, J^2 的本征态，我们可以将 $|l, m, \frac{1}{2}\rangle_0$ 理解为“地球”的自转方向与公转方向相同，将 $|l, m, \frac{1}{2}\rangle_1$ 理解为“地球”的自转方向与公转方向相反。如Figure 3.12所示，在这个情况下，“太阳”的角动量为 $m\hbar$ ，“地球”的角动量为 $\frac{1}{2}\hbar$ 。

根据该表达式，可以得出， $|l, m, \frac{1}{2}\rangle_0$ 对应总角动量 z 方向分量为 $(m + \frac{1}{2})\hbar$ ，总角动量平方为 $(l + \frac{1}{2})(l + \frac{3}{2})\hbar^2$ ； $|l, m, \frac{1}{2}\rangle_1$ 对应总角动量 z 方向分量为 $(m - \frac{1}{2})\hbar$ ，总角动量平方为 $(l - \frac{1}{2})(l + \frac{1}{2})\hbar^2$ 。

3.6.1 氮原子的自旋-轨道耦合

氮原子具有2个核外电子，分析其自旋-轨道耦合，原理上与氢原子是类似的，但是计算极其复杂，在此不做展开，直接记忆结论即可。

首先引入一个概念，在结构化学中，我们经常遇到s, p, d, f这些符号，这些符号表明的是电子的角动量，正如(3.3)所示：

$$\begin{aligned}
 s &\Longleftrightarrow l = 0, \\
 p &\Longleftrightarrow l = 1, \\
 d &\Longleftrightarrow l = 2, \\
 f &\Longleftrightarrow l = 3.
 \end{aligned} \tag{3.39}$$

当2个电子的角动量分别为 l_1, l_2 ，原子的角动量可能的取值为

$$L = l_1 + l_2, l_1 + l_2 - 1, \dots, |l_1 - l_2|, \quad (3.40)$$

原子的自旋可能的取值为

$$S = 0, 1. \quad (3.41)$$

选取固定的 L, S ，原子的总角动量可能的取值为

$$J = L + S, L + S - 1, \dots, |L - S|. \quad (3.42)$$

选取固定的 J ，内积 $\mathbf{L} \cdot \mathbf{S}$ 的取值为

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} (J(J+1) - L(L+1) - S(S+1)) \hbar^2. \quad (3.43)$$

在原子光谱分析中，将参数分别为 L, S, J 的原子置入 z 方向的磁场后，原子会往外发射光谱，总共有 $2J + 1$ 条谱线。定义

$$m_J = -J, -J + 1, \dots, J - 1, J,$$

它具有 $2J + 1$ 个离散的取值，对应光谱的谱线数量。**朗德因子**定义为

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}, \quad (3.44)$$

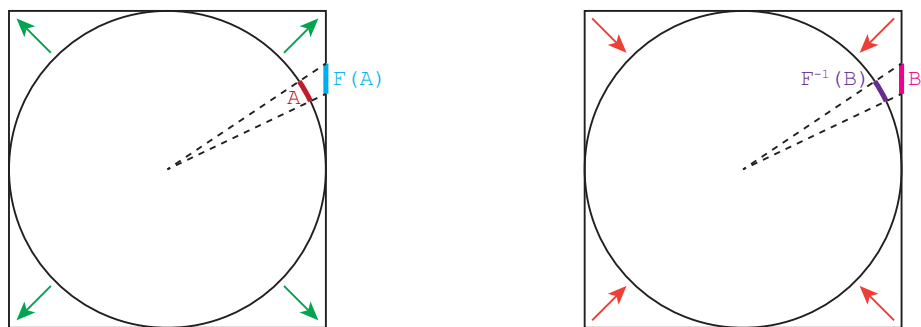
实验测量得到的原子磁矩值为

$$\mu_J = -m_J g \mu_B = J g \mu_B, (J-1) g \mu_B, \dots, (-J+1) g \mu_B, -J g \mu_B, \quad (3.45)$$

其中 $\mu_B = \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \text{ A} \cdot \text{m}$ 是**玻尔磁子**。通常，我们会用下列符号表示原子的参数：

$$^{2S+1}L_J, \quad (3.46)$$

其中，左上标与右下标用数字表示，中间部分用字母表示： $L = 0, 1, 2, 3$ 分别对应字母S, P, D, F。例如， $L = 2, S = 1, J = 3, 2, 1$ 分别对应符号为 $^3D_3, ^3D_2, ^3D_1$ 。



映射 F : 圆形 \rightarrow 正方形

逆映射 F^{-1} : 正方形 \rightarrow 圆形

Figure 4.1: 圆与正方形同胚

4 量子信息

量子信息！这对应“官方资料”完全不说人话的6-7章，我也能体会到学生阅读这部分资料，那种极度痛苦并且扭曲的心情。实际上，量子信息与自旋关系类似，都可以将线性算子与波函数视为2阶矩阵与2维向量。但是此时，我们将线性算子称为“量子门”，波函数称为“量子态”。线性算子作用于波函数，可以类比为电路的逻辑门作用于电路信号，例如，非门将信号0变为信号1，将信号1变为信号0。

4.1 量子比特

量子比特是量子信息的计量单位。与自旋波函数类似，任意量子比特都可以被表示为 $|0\rangle, |1\rangle$ (利用狄拉克符号)的非零线性组合 $|\Psi\rangle = c_0 |0\rangle + c_1 |1\rangle$ ，其中 $|0\rangle, |1\rangle$ 是自旋算子 S_3 的本征态，分别对应 $\chi_0(3), \chi_1(3)$ 。我们可以将 $|\Psi\rangle = c_0 |0\rangle + c_1 |1\rangle$ 表示为向量

$$c_0 |0\rangle + c_1 |1\rangle \longleftrightarrow \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}.$$

根据归一化可得

$$|c_0|^2 + |c_1|^2 = 1.$$

根据(1.2)可得，一个量子比特与一个非零常数 $\lambda \in \mathbb{C}$ 相乘，所表述的量子信息是不变的。

根据这个思路，我们可以引入**复射影空间**的概念：对于去除原点的 $n+1$ 维复空间 $\mathbb{C}^{n+1} \setminus \{0\}$ ，定义等价关系 \sim ：对于任意非零常数 $\lambda \in \mathbb{C}$ ，都具有 $v \sim \lambda v$ ，其中 $v \in \mathbb{C}^{n+1} \setminus \{0\}$ 。

复射影空间 \mathbb{CP}^n 定义为商集 $(\mathbb{C}^{n+1} \setminus \{0\}) / \sim$ ：我们将 $\mathbb{C}^{n+1} \setminus \{0\}$ 上的任意元素都用 $n+1$ 个不全为0的复数 (z_1, \dots, z_{n+1}) 表示，它对应复射影空间 \mathbb{CP}^n 上的等价类，记为 $[z_1, \dots, z_{n+1}]$ 。根据等价关系的定义，对于任意非零常数 λ ，都具有

$$[z_1, \dots, z_{n+1}] = [\lambda z_1, \dots, \lambda z_{n+1}].$$

根据以上信息，量子比特能够表示信息的集合是 \mathbb{CP}^1 。

复射影空间 \mathbb{CP}^n 在代数几何中具有广泛的应用。对于次数为 $d > 0$ 的齐次多项式(每项的次数均为 d ，例如 $x^2y - 2xyz + 3z^3$ ，此时次数 $d = 3$ ； $x^3 - y^2 + z$ 就不是齐次多项式，因为第1项次数为3，第2项次数为2，第3项次数为1) $F(z_1, \dots, z_{n+1})$ ，给定任意非零常数 λ ，可以得出

$$F(\lambda z_1, \dots, \lambda z_{n+1}) = \lambda^d F(z_1, \dots, z_{n+1}),$$

因此， $F(\lambda z_1, \dots, \lambda z_{n+1}) = 0$ 当且仅当 $F(z_1, \dots, z_{n+1}) = 0$ 。而 $[z_1, \dots, z_{n+1}]$ 与 $[\lambda z_1, \dots, \lambda z_{n+1}]$ 是复射影空间 \mathbb{CP}^n 上相同的两点，因此，我们可以在 \mathbb{CP}^n 上分析齐次多项式的零点。当 $n = 2$ 时，零点的集合称为**代数曲线**。

4.1.1 布洛赫球*

²⁰ 我们需要引入拓扑学中至关重要的概念，由于篇幅有限，本文对其解释并不严格。给定非空集合 U, V (严格来说是“拓扑空间”)与映射 $F : U \rightarrow V$ ，对于 U 上任意一点 x ，如果我们总是能找到包含 x 的“很小”的连通子集 $A \subseteq U$ ，使得 $F(A) \subseteq V$ 同样也是“很小”的连通子集，那么我们称映射 F 是**连续**的：正如微积分课程中，对于连续函数 $f : \mathbb{R} \rightarrow \mathbb{R}$ ，我们总能找到“很小”的区间 $(x_0 - \delta, x_0 + \delta)$ ，使得 $f(x_0 - \delta, x_0 + \delta)$ 包含于“很小”的区间 $(f(x_0) - \epsilon, f(x_0) + \epsilon)$ 。如果映射 F 不仅是连续的，而且具有连续的逆映射 $F^{-1} : V \rightarrow U$ ，那我们称映射 F 是**同胚映射**， U 与 V 是**同胚**的。(“很小”这一词汇显然是不规范的说法，但是从“拓扑空间”开始解释过于复杂，由于篇幅有限，因此不做展开。)

因为2个连续映射的复合映射还是连续映射，所以同胚具有传递性：如果 U 与 V 是同胚的，对应同胚映射为 $F : U \rightarrow V$ ； V 与 W 也是同胚的，对应同胚映射为 $G : V \rightarrow W$ ，那么 U 与 W 是同胚的，对应同胚映射为 $G \circ F : U \rightarrow W$ ，此时逆映射为 $F^{-1} \circ G^{-1} : W \rightarrow U$ 。

我们给出一个具体实例。如Figure 4.1所示，我们将一个圆往外拉伸成正方形，这个拉伸的过程就是一个映射，记其为 F 。对于圆上任意一点，我们都可以选取“很小”的子集 A ，使得 $F(A)$ 也是“很小”的子集。 F 存在连续的逆映射 F^{-1} ，因为我们同样可以将正方形压缩成圆形，并且，对于正方形上任意一点，我们同样可以选取“很小”的子集 B ，使得 $F^{-1}(B)$ 也是“很小”的子集。因此，圆与正方形是同胚的。

我们继续给出一些实例，以加深读者的理解：

(A) 我们将圆盘视为复平面上的子集

$$D = \{z : z \in \mathbb{C}, |z| < 1\},$$

定义映射 $F : D \rightarrow \mathbb{C}$ 为

$$F(z) = \frac{z}{1 - |z|},$$

它是连续映射，因为对于任意 $z \in \mathbb{C}, |z| < 1$ ，当 z 发生微小的改变， $F(z)$ 同样也发生微小的改变。定义映射 $G : \mathbb{C} \rightarrow D$ 为

$$G(z) = \frac{z}{1 + |z|},$$

²⁰ 符号*表示这一章节可以跳过，建议跳转到第163页的章节4.3。

同样它也是连续映射，因为对于任意 $z \in \mathbb{C}$ ，当 z 发生微小的改变， $F(z)$ 同样也发生微小的改变。而

$$F \circ G(z) = \frac{\frac{z}{1+|z|}}{1 - \left| \frac{z}{1+|z|} \right|} = \frac{\frac{z}{1+|z|}}{1 - \frac{|z|}{1+|z|}} = z, \quad z \in \mathbb{C},$$

$$G \circ F(z) = \frac{\frac{z}{1-|z|}}{1 + \left| \frac{z}{1-|z|} \right|} = \frac{\frac{z}{1-|z|}}{1 + \frac{|z|}{1-|z|}} = z, \quad z \in \mathbb{C}, |z| < 1,$$

因此 G 是 F 的逆映射，这表明 F 是同胚映射，圆盘 D 与复平面 \mathbb{C} 是同胚的。

(B) 如果 M 是 n 维流形，那么对于任意 $p \in M$ ，都存在包含 p 点的子集 $U \subseteq M$ ，它同胚于 \mathbb{R}^n 上的一个子集 $\hat{U} \subseteq \mathbb{R}^n$ 。单位球面

$$\mathbb{S}^n = \{(x^1, \dots, x^{n+1}) : (x^1)^2 + \dots + (x^{n+1})^2 = 1\}$$

是 n 维流形，我们将对其进行分析。假设 \mathbb{S}^n 上的任意一点 p 的坐标为 (y^1, \dots, y^{n+1}) ，因为 $(y^1)^2 + \dots + (y^{n+1})^2 = 1$ ，所以 y^1, \dots, y^{n+1} 不能全部为0。记非零的坐标分量为 y^k ，并假设 $y^k > 0$ (小于0的情况类似)。那么，我们可以找出包含 p 点的子集

$$U = \{(x^1, \dots, x^{n+1}) : x^k > 0\},$$

令 \mathbb{R}^n 上的子集为

$$\hat{U} = \{(a^1, \dots, a^n) : (a^1)^2 + \dots + (a^n)^2 < 1\},$$

定义映射 $\phi : U \rightarrow \hat{U}$ 为

$$\phi(x^1, \dots, x^{n+1}) = (x^1, \dots, x^{k-1}, x^{k+1}, \dots, x^{n+1}),$$

定义映射 $\psi : \hat{U} \rightarrow U$ 为

$$\psi(a^1, \dots, a^n) = (a^1, \dots, a^{k-1}, \sqrt{1 - (a^1)^2 - \dots - (a^n)^2}, a^k, \dots, a^n),$$

显然 $\phi \circ \psi$ 是恒映射，对于 $\psi \circ \phi$:

$$\begin{aligned} \psi \circ \phi(x^1, \dots, x^{n+1}) &= \psi(x^1, \dots, x^{k-1}, x^{k+1}, \dots, x^{n+1}) \\ &= (x^1, \dots, x^{k-1}, \\ &\quad \sqrt{1 - (x^1)^2 - \dots - (x^{k-1})^2 - (x^{k+1})^2 - \dots - (x^{n+1})^2}, \\ &\quad x^{k+1}, \dots, x^{n+1}) \end{aligned}$$

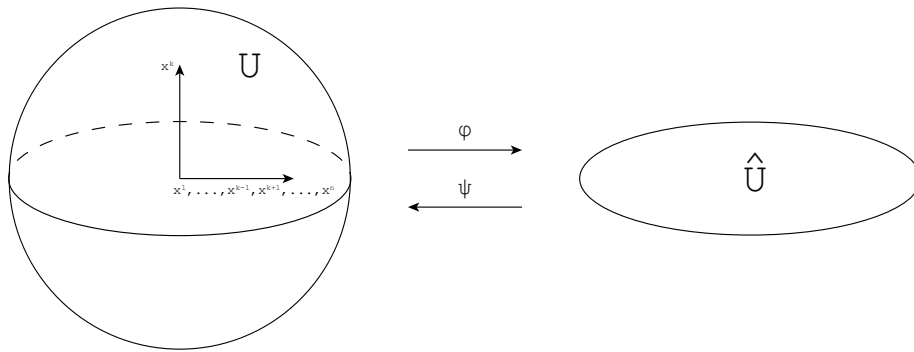


Figure 4.2: \mathbb{S}^n 是 n 维流形

而 $(x^1)^2 + \dots + (x^{n+1})^2 = 1, x^k > 0$, 因此 $\psi \circ \phi(x^1, \dots, x^{n+1}) = (x^1, \dots, x^{n+1})$, 这表明 $\psi \circ \phi$ 也是恒映射。而 ϕ, ψ 很显然是连续的, 因此 U 与 \hat{U} 同胚。如 Figure 4.2 所示, 我们可以认为映射 ϕ 将半球面 U “压平”, 映射 ψ 将圆盘 \hat{U} “拱起”。

那么, 我们可能会想, \mathbb{S}^n 本身是否与 \mathbb{R}^n 上的一个子集同胚? 用严格的数学术语, \mathbb{S}^n 是否嵌入 \mathbb{R}^n ? 我们可以大致构想 $n = 2$ 的情况: \mathbb{S}^2 相当于三维空间的空心球壳, 比如气球。如果它嵌入 \mathbb{R}^2 , 那么我们需要将“空心球壳”展开为一个平面, 并且不能破坏它的完整性(否则连续的逆映射将不存在)。但这是不可能的, 就好比, 你必须将气球戳破, 它才能展开为平面。事实上, 对于任意 $n \geq 1$, \mathbb{S}^n 都无法嵌入 \mathbb{R}^n , 但是其证明需要代数拓扑的知识, 我们无法在此展开。

根据同胚的定义, 我们可以引出以下定理:

Theorem 4.1. \mathbb{CP}^1 与单位球面 $\mathbb{S}^2 = \{(x, y, z) : x, y, z \in \mathbb{R}, x^2 + y^2 + z^2 = 1\}$ 同胚。

Proof. 记 $[z_1, z_2]$ 为 \mathbb{CP}^1 上的任意元素, 其中 $(z_1, z_2) \neq (0, 0)$ 。

(A) 当 $z_1 \neq 0$ 时, 记集合 $U = \{[z_1, z_2] : z_1 \neq 0\} \subseteq \mathbb{CP}^1$ 。令 $w = z_2/z_1$, 因此 $[z_1, z_2] = [1, w]$, 此时 w 可以是任意复数。因此

$$F : U \rightarrow \mathbb{R}^2, [z_1, z_2] = [1, w] \mapsto (\operatorname{Re}(w), \operatorname{Im}(w))$$

是同胚映射, 相当于, 任意复数 $x + iy$ 都与 2 个实数 (x, y) 一一对应。其逆映射为

$$F^{-1} : \mathbb{R}^2 \rightarrow U, (x, y) \mapsto [1, x + iy].$$

如 Figure 4.3 所示, 对于 \mathbb{R}^2 上的任意一点 A , 构建经过 A 与“北极” N 的直线, 那么该直线与球面会有一个交点 $P(A)$ 。记对应的映射为

$$G : \mathbb{R}^2 \rightarrow \mathbb{S}^2 \setminus N, (u, v) \mapsto \frac{(2u, 2v, u^2 + v^2 - 1)}{u^2 + v^2 + 1},$$

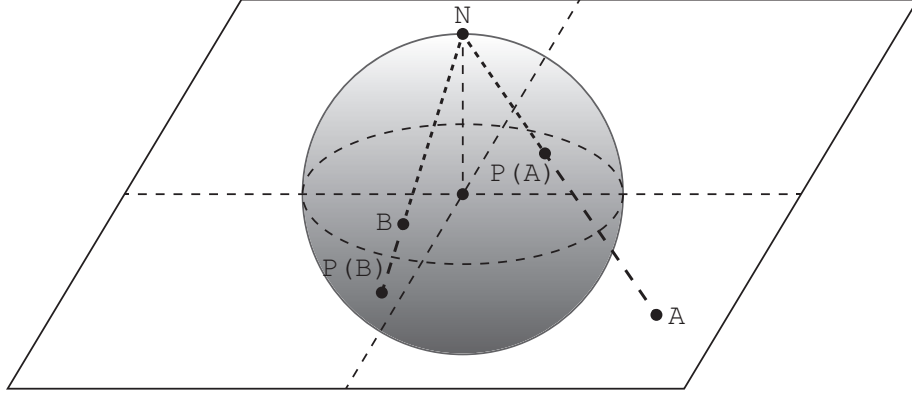


Figure 4.3: G 是同胚映射

它将 A 映射到对应的 $P(A)$ 。因为它具有连续的逆映射

$$G^{-1} : \mathbb{S}^2 \setminus N \rightarrow \mathbb{R}^2, (x, y, z) \mapsto \left(\frac{x}{1-z}, \frac{y}{1-z} \right),$$

所以 G 是同胚映射。根据同胚的传递性可得,

$$H = G \circ F : U \rightarrow \mathbb{S}^2 \setminus N, [z_1, z_2] \mapsto \frac{(2\operatorname{Re}(z_2/z_1), 2\operatorname{Im}(z_2/z_1), |z_2/z_1|^2 - 1)}{|z_2/z_1|^2 + 1} \quad (4.1)$$

是同胚映射, 其逆映射为

$$H^{-1} = F^{-1} \circ G^{-1} : \mathbb{S}^2 \setminus N \rightarrow U, (x, y, z) \mapsto \left[1, \frac{x + iy}{1 - z} \right]. \quad (4.2)$$

(B) 当 $z_1 = 0, z_2 \neq 0$ 时, 记集合 $V = \{[z_1, z_2] : z_1 = 0, z_2 \neq 0\} \subseteq \mathbb{CP}^1$ 。此时 $[z_1, z_2] = [0, 1]$, 因此, 集合 V 只有一个元素 $[0, 1]$ 。

容易得出 $U \cap V = \emptyset, U \cup V = \mathbb{CP}^1$, 定义映射 $D : \mathbb{CP}^1 \rightarrow \mathbb{S}^2$,

$$D(x) = \begin{cases} H(x), & x \in U, \\ N, & x \in V. \end{cases} \quad (4.3)$$

D 是一一映射, 因此具有逆映射。我们需要证明 D 是同胚映射, 即映射 D 与映射 D^{-1} 均为连续映射。

(1) 首先证明 $D : \mathbb{CP}^1 \rightarrow \mathbb{S}^2$ 是连续映射。根据(4.1)与同胚的定义可得, $D|_U$ 是连续映射, 所以我们只需选取包含 $[0, 1] \in V$ 的“很小”的子集

$$\Omega = \{[z, 1] : |z| < \epsilon\} \subseteq \mathbb{CP}^1,$$

其中 ϵ 是任意小的正数。根据(4.1)，并进行化简可得，

$$D(\Omega \setminus [0, 1]) = \left\{ \frac{(2\operatorname{Re}(z), -2\operatorname{Im}(z), 1 - |z|^2)}{1 + |z|^2} : 0 < |z| < \epsilon \right\},$$

$$D([0, 1]) = N = (0, 0, 1),$$

因此

$$D(\Omega) = D(\Omega \setminus [0, 1]) \cup D([0, 1]) = \left\{ \frac{(2\operatorname{Re}(z), -2\operatorname{Im}(z), 1 - |z|^2)}{1 + |z|^2} : |z| < \epsilon \right\},$$

该集合是包含 \mathbb{S}^2 的北极 N 的“很小”的子集。

- (2) 其次证明 $D^{-1} : \mathbb{S}^2 \rightarrow \mathbb{CP}^1$ 是连续映射。根据(4.1)与同胚的定义可得， $D^{-1}|_{\mathbb{S}^2 \setminus N}$ 是连续映射，所以我们只需选取包含 $N \in \mathbb{S}^2$ 的“很小”的子集

$$\Gamma = \left\{ (x, y, \sqrt{1 - x^2 - y^2}) : x^2 + y^2 < \min(\epsilon, 1) \right\} \subseteq \mathbb{S}^2,$$

其中 ϵ 是任意小的正数。根据(4.2)，并进行化简可得，

$$D^{-1}(\Gamma \setminus N) = \left\{ \left[\frac{x - iy}{1 + \sqrt{1 - x^2 - y^2}}, 1 \right] : 0 < x^2 + y^2 < \min(\epsilon, 1) \right\},$$

$$D^{-1}(N) = [0, 1],$$

因此

$$D^{-1}(\Gamma) = D^{-1}(\Gamma \setminus N) \cup D^{-1}(N) = \left\{ \left[\frac{x - iy}{1 + \sqrt{1 - x^2 - y^2}}, 1 \right] : x^2 + y^2 < \min(\epsilon, 1) \right\},$$

该集合是包含 $[0, 1]$ 的“很小”的子集。

□

因此，任意一个量子比特都与单位球面上的点一一对应，我们将这个单位球面称为**黎曼球面**(又称为**布洛赫球**)。而单位球面上的任意一点都可以表示为 $\cos \frac{\theta}{2} N + e^{i\phi} \sin \frac{\theta}{2} S$ ，其中 $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, N, S 是“北极”与“南极”的坐标。因此，任意一个量子比特都可以表示为以下形式：

$$|\Psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle,$$

其中 $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$ 。如Figure 4.4所示，当 $\theta = 0$ 时， $|\Psi\rangle = |0\rangle$ ；当 $\theta = \pi$ 时， $|\Psi\rangle = |1\rangle$ ，因此，该图提供了一个较为直观的表述。此时量子比特可以表示为向量

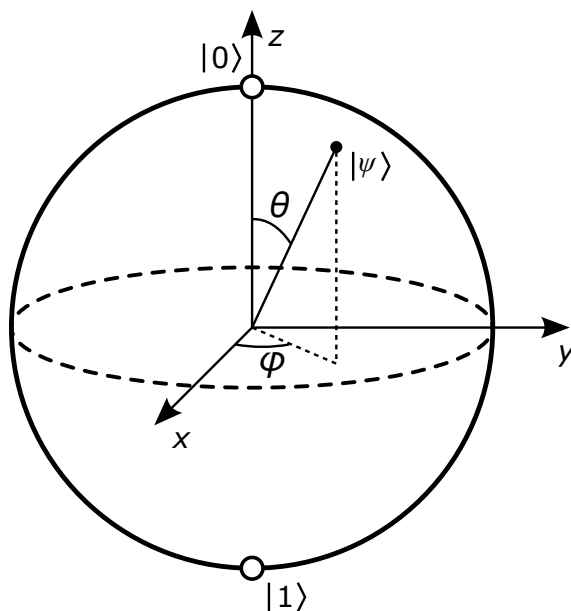


Figure 4.4: 布洛赫球

$$|\Psi\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix}.$$

通常来说，证明“不同胚”往往比证明“同胚”要困难：因为对于前者，你需要证明2个拓扑空间之间不存在连续的一一映射。但是，幸运的是，对于2个同胚的拓扑空间，它们之间具有一些相同的参数，称为**拓扑不变量**，例如**基本群(Fundamental group)**、**同调群(Homology group)**、**上同调群(Cohomology group)**等。因此，如果拓扑空间 A, B 的某一拓扑不变量是不相同的——例如，当 A, B 的上同调群 $H^*(A) \neq H^*(B)$ ，那么 A, B 就不是同胚的。

但是，这不代表 A, B 之间如果存在某一个相同的拓扑不变量，就能推断出 A, B 同胚。因为很有可能出现， A, B 之间存在一个拓扑不变量相同，而另一个拓扑不变量不同的情况——例如， S^2, S^3 的基本群是相同的： $\pi_1(S^2) = \pi_1(S^3) = 0$ ；但是它们的同调群是不相同的： $H_2(S^2) = \mathbb{Z}, H_2(S^3) = 0$ 。所以 S^2, S^3 不是同胚的。

找出，并证明拓扑空间的某一参数是拓扑不变量，亦或是计算给定拓扑空间的一些拓扑不变量，是拓扑学的核心内容。即使到了现在，一些拓扑不变量依然难以计算：例如 S^n 的**同伦群(Homotopy group)** $\pi_k(S^n)$ ，在 k 很大的时候，目前没有任何的计算方法。

我们将给出一个对读者较为友好的拓扑不变量：对于任意没有边界的曲面，我们将它剖分成“碎片”，每一块“碎片”都是一个曲面三角形(严格来说，同胚于2-单形，2-单形就是三角形)，并且任意2个曲面三角形要么不相交，要么有1条公共边。我们将这一过程称为**三角剖分**，如Figure 4.5所示。很显然，三角剖分并不是唯一的——可以剖分“精细”一点，也可以剖分“粗糙”一点。

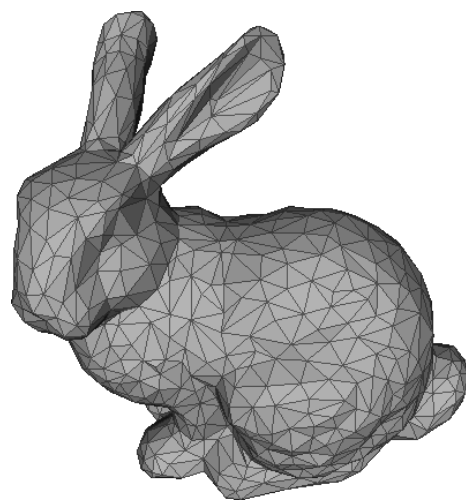


Figure 4.5: 三角剖分

记三角剖分的曲面 M 共有 V 个“顶点”， E 个“边”， F 个“面”，那么曲面 M 的**欧拉示性数(Euler characteristic)**为

$$\chi(M) = V - E + F.$$

欧拉示性数是一个拓扑不变量，它与三角剖分的方式无关；根据拓扑不变量的定义，如果曲面 M 与 N 同胚，那么 $\chi(M) = \chi(N)$ 。(在这里不做证明)

对于球面与环面

$$\begin{aligned} \mathbb{S}^2 &= \{(x, y, z) : x^2 + y^2 + z^2 = 1\}, \\ \mathbb{T}^2 &= \left\{ (x, y, z) : \left(\sqrt{x^2 + y^2} - 2 \right)^2 + z^2 = 1 \right\}. \end{aligned}$$

我们将计算它们的欧拉示性数。如Figure 4.6所示，我们将球面剖分为8个曲面三角形，将环面剖分为32个曲面三角形(没有完全绘制出来，因为太难绘制了。只需在每个曲面四边形上绘制对角线即可)。因此，对于球面， $V = 6, E = 12, F = 8$ ；对于环面， $V = 16, E = 48, F = 32$ 。它们的欧拉示性数分别为

$$\begin{aligned} \chi(\mathbb{S}^2) &= V - E + F = 6 - 12 + 8 = 2, \\ \chi(\mathbb{T}^2) &= V - E + F = 16 - 48 + 32 = 0. \end{aligned}$$

它们是不相同的，这表明 \mathbb{S}^2 与 \mathbb{T}^2 不是同胚的。

4.2 信息熵*

²¹ 以下是3个关于我(作者)的生日的命题:

²¹符号*表示这一章节可以跳过，建议跳转到第163页的章节4.3。

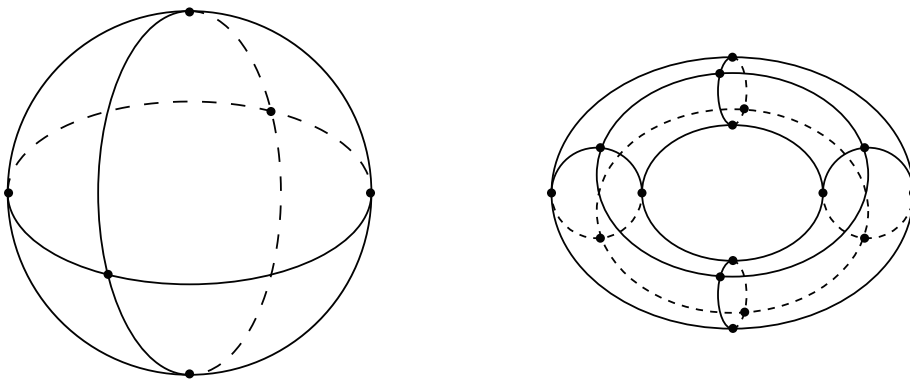


Figure 4.6: 球面与环面的三角剖分

(A) 我的生日在一年当中的某天。

(B) 我的生日在1月。

(C) 我的生日在1月5日。

虽然它们都是正确的，但是它们所表述的信息详细程度是不一样的：例如，命题(A)的信息量为0——因为任何人的情况都是如此；另外，命题(B)的信息量小于命题(C)的信息量——因为通过命题(B)只能确定我生日的大致范围，通过命题(C)能够完全确定我的生日！但是，这是秘密，就不要告诉别人了。

这一现象可以利用概率论的知识解释。假设事件 A_1, A_2, A_3 分别为“命题(A), (B), (C)是正确的”，那么，事件 A_1 发生的概率为 $\Pr\{A_1\} = 1$ ——因为满足这一条件人群的比例是1；同理，事件 A_2, A_3 发生的概率分别为 $\Pr\{A_2\} = \frac{1}{12}, \Pr\{A_3\} = \frac{1}{365}$ 。我们可以看出， $\Pr\{A_1\} > \Pr\{A_2\} > \Pr\{A_3\}$ ，但是对应的信息量大小顺序则完全相反——概率越低的事件对应的信息量越大。

据此，我们可以引出信息论的一些概念。假设离散型随机变量 X 所有可能取值的集合是可列集合

$$\Omega = \{x_i : i \in \mathbb{Z}^+\}.$$

对于任意正整数 $i \in \mathbb{Z}^+$ ，记 $\Pr\{x_i\}$ 为 X 取到 x_i 的概率，根据概率论的公理可得

$$\begin{aligned} \Pr\{x_i\} &\in [0, 1], \quad i \in \mathbb{Z}^+, \\ \sum_{j=1}^{\infty} \Pr\{x_j\} &= 1. \end{aligned} \tag{4.4}$$

我们定义变量 x_i 的信息量为

$$I\{x_i\} = -\log_2 \Pr\{x_i\},$$

很显然，当 $\Pr\{x_i\}$ 越小时，对应的信息量越大；并且，当 $\Pr\{x_i\} = 1$ 时，对应的信息量为0。这个定义正好符合了我们之前的讨论：命题(A)到(C)的概率越来越小，对应的信息

量越来越大；而命题(A)的概率为1，信息量为0。离散型随机变量 X 的**信息熵**是信息量的加权平均：

$$H(X) = \sum_{i=1}^{\infty} \Pr\{x_i\} I\{x_i\} = - \sum_{i=1}^{\infty} \Pr\{x_i\} \log_2 \Pr\{x_i\}. \quad (4.5)$$

因为 $\Pr\{x_i\} \in [0, 1]$ ，所以信息熵的取值是非负的。

注意：很多时候，离散型随机变量 X 所有可能取值的数量是有限的，即 $|\Omega| < \infty$ 。此时，按理来说，(4.4)与(4.5)应该修改为

$$\begin{aligned} \Pr\{x_i\} &\in [0, 1], \quad i \in [1, |\Omega|], \\ \sum_{j=1}^{|\Omega|} \Pr\{x_j\} &= 1, \\ H(X) &= - \sum_{i=1}^{|\Omega|} \Pr\{x_i\} \log_2 \Pr\{x_i\}, \end{aligned}$$

这当然是正确的。但是，我们可以将 Ω 扩充到可列无限集合，只需令

$$\Pr\{x_i\} = 0, \quad i > |\Omega|$$

即可，换句话说，将 X 不可能的取值加入集合 Ω 。此时前2个表达式与(4.4)是等价的，关于第3个表达式是否与(4.5)等价，我们可以进行一个简单的计算：虽然函数 $f(t) = -t \log_2 t$ 无法在 $t = 0$ 时定义，但是根据洛必达法则，

$$\lim_{t \searrow 0} f(t) = \lim_{\underbrace{u \rightarrow \infty}_{u = 1/t}} \frac{\log_2 u}{u} = \lim_{u \rightarrow \infty} \frac{1}{u \log 2} = 0,$$

因此，对于任意 $i > |\Omega|$ ， $-\Pr\{x_i\} \log_2 \Pr\{x_i\}$ 对信息熵的大小是没有影响的，采用定义(4.5)是没有问题的。

我们将给出一些具体实例。

(A) 投掷一个骰子，将上面的数字 X 所有可能取值的集合记为 Ω ，那么

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

添加 X 不可能的取值，将 Ω 扩充到可列无限集合，

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, \dots\}.$$

那么

$$\Pr\{x_i\} = \begin{cases} \frac{1}{6}, & i \in [1, 6], \\ 0, & i \geq 7. \end{cases}$$

因此，信息熵为

$$H(X) = - \sum_{i=1}^{\infty} \Pr \{x_i\} \log_2 \Pr \{x_i\} = \log_2 6. \quad (4.6)$$

- (B) 我在2023秋季学期带电磁学C助教的时候，将唱歌 X 所有可能取值的集合记为 Ω ，那么

$$\Omega = \{\text{Zood, I got a smoke, See you again}\}.$$

添加 X 不可能的取值，将 Ω 扩充到可列无限集合，

$$\Omega = \{\text{Zood, I got a smoke, See you again, 猫中毒, 敢测我的马, ...}\},$$

之所以不可能取到 x_4 ，是因为我厌恶嘉然(不可以吗)；对于 x_5 ，则是因为我带电磁学C助教的时候，这首歌曲还没有发布。那么

$$\Pr \{x_i\} = \begin{cases} \frac{99}{100}, & i = 1, \\ \frac{99}{10,000}, & i = 2, \\ \frac{1}{10,000}, & i = 3, \\ 0, & i \geq 4. \end{cases}$$

因此，信息熵为

$$H(X) = - \sum_{i=1}^{\infty} \Pr \{x_i\} \log_2 \Pr \{x_i\} = 0.0816. \quad (4.7)$$

- (C) 我在2024春季学期带量子物理助教的时候，将黑板画 X 所有可能取值的集合记为 Ω ，那么

$$\Omega = \{\text{柯南, 新一与小兰接吻, 安室透与赤井秀一结婚}\}.$$

添加 X 不可能的取值，将 Ω 扩充到可列无限集合，

$$\Omega = \{\text{柯南, 新一与小兰接吻, 安室透与赤井秀一结婚, 柯哀CP, 平次向和叶告白, ...}\},$$

之所以不可能取到 x_4 ，是因为柯哀本身就不是官配(^ ^)；对于 x_5 ，则是因为这件事情本身就是不可能的。那么

$$\Pr \{x_i\} = \begin{cases} \frac{1}{3}, & i = 1, \\ \frac{1}{2}, & i = 2, \\ \frac{1}{6}, & i = 3, \\ 0, & i \geq 4. \end{cases}$$

(这都能三选一)因此，信息熵为

$$H(X) = - \sum_{i=1}^{\infty} \Pr \{x_i\} \log_2 \Pr \{x_i\} = 1.46.$$

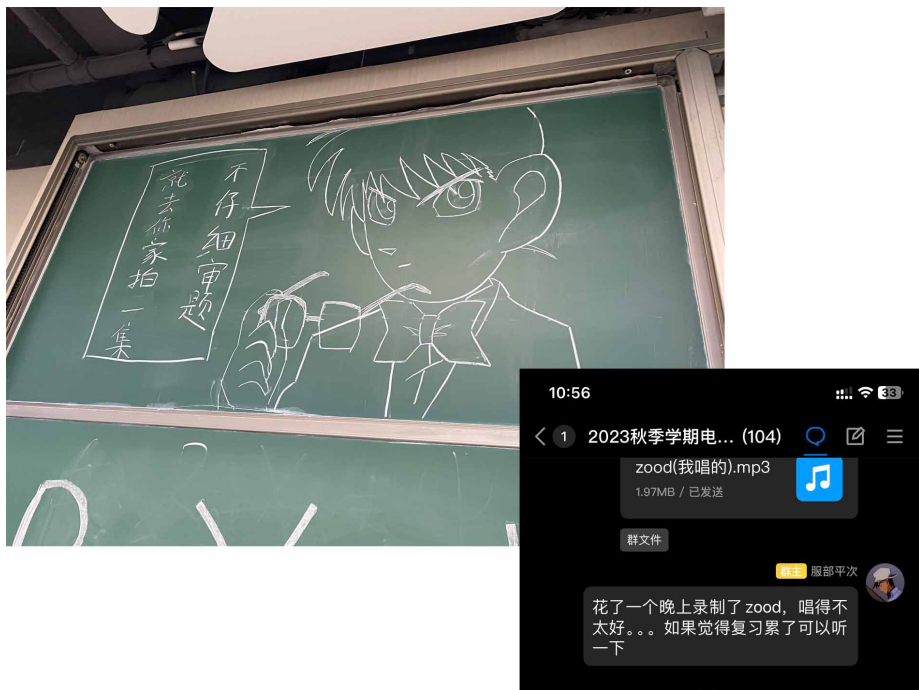


Figure 4.7: 我的作品

情况(B)与(C)如Figure 4.7所示。

对于开区间 $J \subseteq \mathbb{R}$, 我们定义函数 $f : J \rightarrow \mathbb{R}$ 是凸函数, 如果对于所有 $x, y \in J$ 与任意 $0 \leq \tau \leq 1$, 都具有

$$f(\tau x + (1 - \tau)y) \leq \tau f(x) + (1 - \tau)f(y). \quad (4.8)$$

Theorem 4.2. 对于二阶可导函数 $f \in C^2(J)$, f 是凸函数当且仅当 $f''(x) \geq 0$ 对于任意 $x \in J$ 都成立。

Proof. 根据 J 是开区间这一性质, 对于任意 $x \in J$, 我们总是能够找到正数 $h > 0$ 使得 $(x - 2h, x + 2h) \subseteq J$ 。根据泰勒定理, 存在 $c \in (x, x + h)$, $d \in (x - h, x)$ 使得

$$\begin{aligned} f(x + h) &= f(x) + f'(x)h + \frac{1}{2}f''(c)h^2, \\ f(x - h) &= f(x) - f'(x)h + \frac{1}{2}f''(d)h^2, \end{aligned}$$

进行相加可得

$$f(x + h) + f(x - h) = 2f(x) + \frac{1}{2}h^2 (f''(c) + f''(d)),$$

因此

$$f''(c) + f''(d) = 2 \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}.$$

取极限 $h \searrow 0$, 因为 $c \in (x, x+h)$, $d \in (x-h, x)$, 所以此时 $c, d \rightarrow x$. 这表明 $f''(c) + f''(d) \rightarrow 2f''(x)$, 因此

$$f''(x) = \lim_{h \searrow 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}.$$

将表达式(4.8)取 $\tau = 1/2$, 可得

$$f(x+h) + f(x-h) \geq 2f\left(\frac{1}{2}(x+h) + \frac{1}{2}(x-h)\right) = 2f(x).$$

因此

$$f''(x) = \lim_{h \searrow 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} \geq \lim_{h \searrow 0} \frac{2f(x) - 2f(x)}{h^2} = 0,$$

这证明了充分条件。

现在需要证明必要条件。假设 $f''(x) \geq 0$ 对于任意 $x \in J$ 都成立, 那么对于任意 $a, b \in J$, 根据泰勒定理, 存在 $c \in (\min(a, b), \max(a, b))$ 使得

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2}f''(c)(b-a)^2 \geq f(a) + f'(a)(b-a).$$

选取 $a = \tau x + (1-\tau)y$, $b = x$, 我们可以得到

$$f(x) \geq f(\tau x + (1-\tau)y) + (1-\tau)(x-y)f'(\tau x + (1-\tau)y). \quad (4.9)$$

选取 $a = \tau x + (1-\tau)y$, $b = y$, 我们可以得到

$$f(y) \geq f(\tau x + (1-\tau)y) - \tau(x-y)f'(\tau x + (1-\tau)y). \quad (4.10)$$

将(4.9)两边与 τ 相乘, (4.10)两边与 $1-\tau$ 相乘, 之后再将它们相加, 我们就得到了(4.8)。

□

Corollary 4.2.1. 如果 $f: J \rightarrow \mathbb{R}$ 是凸函数, 那么对于所有 $x_1, \dots, x_n \in J$ 与任意

$$a_1 + \dots + a_n = 1, \quad a_1, \dots, a_n \in [0, 1],$$

都具有

$$f(a_1x_1 + \dots + a_nx_n) \leq a_1f(x_1) + \dots + a_nf(x_n). \quad (4.11)$$

特别的, 如果 $x_1 + \dots + x_n = 1$, 并且 $a_1 = \dots = a_n = 1/n$, 那么

$$nf\left(\frac{1}{n}\right) \leq f(x_1) + \dots + f(x_n). \quad (4.12)$$

Proof. 我们利用数学归纳法。根据(4.8), 表达式(4.11)对 $n = 2$ 是成立的。假设(4.11)对 $n = k - 1$ 是成立的, 我们需要分析它对 $n = k$ 是否成立。

(A) 当 $a_k = 1$ 时, $a_1 = \cdots = a_{k-1} = 0$ 。那么

$$\begin{aligned} f(a_1x_1 + \cdots + a_nx_n) &= f(x_n), \\ a_1f(x_1) + \cdots + a_nf(x_n) &= f(x_n), \end{aligned}$$

(4.11)显然成立。

(B) 当 $a_k \neq 1$ 时, 可得

$$\frac{a_1}{1-a_k} + \cdots + \frac{a_{k-1}}{1-a_k} = 1.$$

根据(4.8), 令 $x = \frac{a_1}{1-a_k}x_1 + \cdots + \frac{a_{k-1}}{1-a_k}x_{k-1}, y = x_k, \tau = 1 - a_k$:

$$\begin{aligned} f(a_1x_1 + \cdots + a_kx_k) &= f(\tau x + (1-\tau)y) \leq \tau f(x) + (1-\tau)f(y) \\ &= (1-a_k)f\left(\frac{a_1}{1-a_k}x_1 + \cdots + \frac{a_{k-1}}{1-a_k}x_{k-1}\right) + a_kf(x_k) \\ &\leq (1-a_k)\left(\frac{a_1}{1-a_k}f(x_1) + \cdots + \frac{a_{k-1}}{1-a_k}f(x_{k-1})\right) + a_kf(x_k) \\ &= a_1f(x_1) + \cdots + a_kf(x_k). \end{aligned}$$

因此, (4.11)对 $n = k$ 依旧是成立的, 我们完成了证明。

□

根据Theorem 4.2, 我们可以通过计算函数 $F(x) = x \log_2 x (x > 0)$ 的二阶导数来得知 $F(x)$ 是凸函数。假设随机变量 X 所有可能取值的数量是有限的, 记它们为 x_1, \dots, x_n , 那么根据(4.5), 信息熵为

$$H(X) = - \sum_{i=1}^n F(\Pr\{x_i\}).$$

而根据概率论的公理, 又具有

$$\sum_{i=1}^n \Pr\{x_i\} = 1.$$

根据(4.12)可得,

$$H(X) \leq -nF\left(\frac{1}{n}\right) = \log_2 n,$$

这恰好对应 $\Pr\{x_1\} = \cdots = \Pr\{x_n\} = 1/n$ 的信息熵! 这表明, 当 X 得到所有取值的概率都相等的时候, 对应的信息熵是最大的: 这从定量的角度解释了你们以前在中学物理遇到的未解之谜, 为什么“熵”能够与“无序”联系在一起? 因为“无序”, 意味着物体可能处于空间

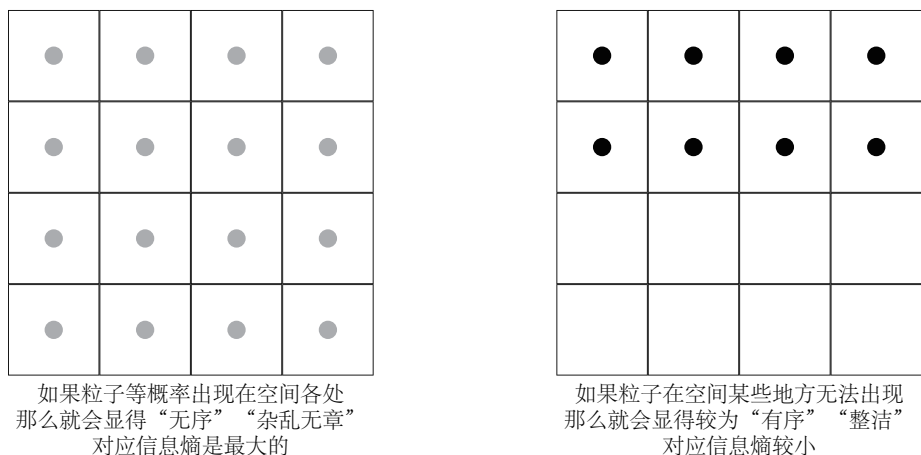


Figure 4.8: “无序”对应熵值较大，“有序”对应熵值较小

中的任何位置；如果给物体运动加上了一定的限制，使之无法出现于特定的区域(代表物理在该区域出现的概率为0)，那么整个空间将会变得“有序”。Figure 4.8较为直观地解释了这一现象。

这也正好与我们的计算结果(4.6)与(4.7)相互照应：因为对于前者，我唱第一首歌的意愿是很强烈的，意味着我对“唱歌”这一信息掌握程度很高，这代表了“有序”；对于后者，我画三幅画的意愿有点“模棱两可”，意味着我对“绘画”这一信息掌握程度很低，这代表了“无序”。

4.2.1 条件熵与联合熵*

²² 有些时候，即使给定一个看似信息量很大的描述，我们还是无法得到想要的结果：例如，假设某人的体重是70 kg，那我们将无法得知此人体重是否正常，这与他的身高有关——如果他的身高是165 cm，那么显然偏重；如果是175 cm，那么较为正常；如果是185 cm，那么显然偏轻。因此，这一问题涉及2个变量，分别为“身高”与“体重”：我们必须得知道他的“身高”是多少，这样，“体重”这一变量才具有信息量可言。

据此，我们需要引入新的概念。假设2个离散型随机变量 X, Y 所有可能取值的集合都是可列集合

$$\Omega_1 = \{x_i : i \in \mathbb{Z}^+\},$$

$$\Omega_2 = \{y_j : j \in \mathbb{Z}^+\}.$$

记**条件概率** $\Pr\{y_j|x_i\}$ 为，在 X 取到 x_i 的情况下， Y 取到 y_j 的概率；**联合概率** $\Pr\{x_i, y_j\}$ 为， X 取到 x_i 与 Y 取到 y_j 这2个条件同时满足的概率； $\Pr\{x_i\}$ 为 X 取到 x_i 的概率，无论 Y 的取值

²²符号*表示这一章节可以跳过，建议跳转到第163页的章节4.3。

如何； $\Pr\{y_j\}$ 为 Y 取到 y_j 的概率，无论 X 的取值如何。因此

$$\begin{aligned}\Pr\{x_i, y_j\} &= \Pr\{x_i\} \Pr\{y_j|x_i\}, \\ \sum_{i=1}^{\infty} \Pr\{x_i, y_j\} &= \Pr\{y_j\}, \\ \sum_{j=1}^{\infty} \Pr\{x_i, y_j\} &= \Pr\{x_i\},\end{aligned}\tag{4.13}$$

并规定，当 $\Pr\{x_i\} = 0$ 时， $\Pr\{y_j|x_i\} = 0$ 。很显然，如果 $\Pr\{y_j\} = 0$ ，那么即使加上 X 取到 x_i 这一条件， Y 取到 y_j 的概率还是为0，因此 $\Pr\{y_j|x_i\} = 0$ 。总而言之，当 $\Pr\{x_i\} = 0$ 或 $\Pr\{y_j\} = 0$ 任意一个条件成立时， $\Pr\{y_j|x_i\} = 0$ 。

很显然 $\Pr\{y_j|x_i\}$ 一定是非负的，因为等式右边要么是一个非负数与一个正数相除，要么是0。另外， $\Pr\{y_j|x_i\}$ 不会大于1，因为“ X 取到 x_i 与 Y 取到 y_j 这2个条件同时满足”这一要求很显然比“ X 取到 x_i ，无论 Y 的取值如何”要来得苛刻，因此

$$\Pr\{x_i, y_j\} \leq \Pr\{x_i\}.$$

因为 X 取到 x_i 时， Y 所有可能的取值依然为

$$\{y_j : j \in \mathbb{Z}^+\},$$

所以，对于任意正整数 $i \in \mathbb{Z}^+$ ，都具有

$$\begin{aligned}\Pr\{y_j|x_i\} &\in [0, 1], \quad i \in \mathbb{Z}^+, \\ \sum_{j=1}^{\infty} \Pr\{y_j|x_i\} &= 1.\end{aligned}\tag{4.14}$$

根据(4.5)，在 $X = x_i$ 的情况下，条件概率对应的信息熵为

$$H(Y|x_i) = - \sum_{j=1}^{\infty} \Pr\{y_j|x_i\} \log_2 \Pr\{y_j|x_i\}.$$

我们定义**条件熵**为条件概率对应信息熵的加权平均，并利用(4.13)：

$$\begin{aligned}H(Y|X) &= \sum_{i=1}^{\infty} \Pr\{x_i\} H(Y|x_i) \\ &= - \sum_{i,j=1}^{\infty} \Pr\{x_i\} \Pr\{y_j|x_i\} \log_2 \Pr\{y_j|x_i\} \\ &= - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}.\end{aligned}$$

同样的, Ω_1, Ω_2 有可能是有限集合, 即 $|\Omega_1| < \infty$ 或 $|\Omega_2| < \infty$ 。我们需要得知, 对于不可能取到的 $x_i (i > |\Omega_1|, \Pr\{x_i\} = 0)$ 或 $y_j (j > |\Omega_2|, \Pr\{y_j\} = 0)$, $-\Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}$ 这一项是否会对条件熵产生影响: 此时 $\Pr\{y_i|x_j\} = 0$, 利用 $\Pr\{x_i\} \leq 1$ 与(4.13)可得

$$\begin{aligned} -\Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\} &= -\Pr\{x_i\} \Pr\{y_j|x_i\} \log_2 \Pr\{y_j|x_i\} \\ &\leq -\Pr\{y_j|x_i\} \log_2 \Pr\{y_j|x_i\} \\ &= \lim_{t \searrow 0} (-t \log_2 t) \\ &= 0, \end{aligned}$$

而 $-\Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}$ 是非负的, 因此 $-\Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\} = 0$, 这表明它并不会对条件熵产生影响。

正如(4.13), 我们将条件概率 $\Pr\{y_j|x_i\}$ 与独立变量的概率 $\Pr\{x_i\}$ 通过联合概率 $\Pr\{x_i, y_j\}$ 联系起来。那么, 条件熵 $H(Y|X)$ 与信息熵 $H(X)$ 之间, 是否可以通过“联合熵(记为 $H(X, Y)$)”联系起来呢? 我们需要定义随机变量 $X \times Y$, 其所有可能取值的集合同样也是可列集合

$$\Omega_1 \times \Omega_2 = \{(x_i, y_j) : i, j \in \mathbb{Z}^+\}.$$

对于任意正整数 $i, j \in \mathbb{Z}^+$, 根据联合概率的定义, $\Pr\{x_i, y_j\}$ 为 $X \times Y$ 取到 (x_i, y_j) 的概率。根据概率论的公理可得

$$\begin{aligned} \Pr\{x_i, y_j\} &\in [0, 1], \quad i, j \in \mathbb{Z}^+, \\ \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} &= 1. \end{aligned}$$

类比(4.5), 定义**联合熵**为

$$H(X, Y) = - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 \Pr\{x_i, y_j\}. \quad (4.15)$$

以下定理证明我们的猜想是正确的:

Theorem 4.3.

$$H(X, Y) = H(X) + H(Y|X). \quad (4.16)$$

Proof. 因为

$$\begin{aligned} H(X) &= - \sum_{i=1}^{\infty} \Pr\{x_i\} \log_2 \Pr\{x_i\}, \\ H(Y|X) &= - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}. \end{aligned}$$

根据(4.13)与(4.14)可得

$$\begin{aligned}
H(X) + H(Y|X) &= - \sum_{i=1}^{\infty} \Pr\{x_i\} \log_2 \Pr\{x_i\} - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\} \\
&= - \sum_{i,j=1}^{\infty} (\Pr\{x_i\} \Pr\{y_j|x_i\} \log_2 \Pr\{x_i\} + \Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}) \\
&= - \sum_{i,j=1}^{\infty} (\Pr\{x_i, y_j\} \log_2 \Pr\{x_i\} + \Pr\{x_i, y_j\} \log_2 \Pr\{y_j|x_i\}) \\
&= - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 (\Pr\{x_i\} \Pr\{y_j|x_i\}) \\
&= - \sum_{i,j=1}^{\infty} \Pr\{x_i, y_j\} \log_2 \Pr\{x_i, y_j\} \\
&= H(X, Y).
\end{aligned}$$

□

回到我们所举的实例，记“身高”为离散型随机变量 X ，“体重”为离散型随机变量 Y 。因为“身高”给予“体重”一定的信息量——我们必须知道一个人的身高，才能判断他的体重是否正常，这表明，“身高未知”与“身高已知”带来“对信息的掌握程度”差距是巨大的。而信息熵又可以理解为“无序”，即“对信息的掌握程度”(对信息掌握程度越低，无序程度越高，信息熵越大；反之，对信息掌握程度越高，无序程度越低，信息熵越小)。因此，当我们知道身高以后，信息熵的改变

$$H(Y|X) - H(Y)$$

是一定小于0的。

据此，对于任意离散型随机变量 X, Y ，我们可以引入**互信息**的概念：

$$I(X; Y) = H(Y) - H(Y|X). \quad (4.17)$$

直觉上可以得知，它一定是非负的：无论 X 与 Y 是否具有关联，在知道 Y 的取值后，我们对 X 的掌握程度也不可能减小；并且，如果 X 与 Y 具有关联(比如“身高”与“体重”，较高的人往往较重)，那么在知道 Y 的取值后，我们对 X 的掌握程度会有变化，互信息 $I(X; Y)$ 也大于0。

但是，这一表述缺乏数学上的严谨。因此，我们将从严格的角度出发，来印证我们的论证：

Theorem 4.4. 对于任意离散型随机变量 X, Y ，互信息 $I(X; Y)$ 都是非负的。

Proof. 因为

$$\begin{aligned} H(Y) &= - \sum_{j=1}^{\infty} \Pr \{y_j\} \log_2 \Pr \{y_j\}, \\ H(Y|X) &= - \sum_{i,j=1}^{\infty} \Pr \{x_i, y_j\} \log_2 \Pr \{y_j|x_i\}. \end{aligned}$$

所以我们只需证明

$$\sum_{i,j=1}^{\infty} \Pr \{x_i, y_j\} \log_2 \Pr \{y_j|x_i\} \geq \sum_{j=1}^{\infty} \Pr \{y_j\} \log_2 \Pr \{y_j\}$$

即可。

令凸函数 $F(x) = x \log_2 x (x > 0)$, 那么根据(4.13)可得

$$\begin{aligned} \sum_{i,j=1}^{\infty} \Pr \{x_i, y_j\} \log_2 \Pr \{y_j|x_i\} &= \sum_{i,j=1}^{\infty} \Pr \{x_i\} \Pr \{y_j|x_i\} \log_2 \Pr \{y_j|x_i\} \\ &= \sum_{i,j=1}^{\infty} \Pr \{x_i\} F(\Pr \{y_j|x_i\}), \\ \sum_{j=1}^{\infty} \Pr \{y_j\} \log_2 \Pr \{y_j\} &= \sum_{j=1}^{\infty} F(\Pr \{y_j\}). \end{aligned}$$

这等价于证明, 对于任意 $j \in \mathbb{Z}^+$, 都具有

$$\sum_{i=1}^{\infty} \Pr \{x_i\} F(\Pr \{y_j|x_i\}) \geq F(\Pr \{y_j\}). \quad (4.18)$$

因为 $\sum_{i=1}^{\infty} \Pr \{x_i\} = 1$, 所以根据(4.11)与(4.13)可得

$$\begin{aligned} \sum_{i=1}^{\infty} \Pr \{x_i\} F(\Pr \{y_j|x_i\}) &\geq F\left(\sum_{i=1}^{\infty} \Pr \{x_i\} \Pr \{y_j|x_i\}\right) \\ &= F\left(\sum_{i=1}^{\infty} \Pr \{x_i, y_j\}\right) \\ &= F(\Pr \{y_j\}). \end{aligned}$$

□

此外,“关联性”这一用语,也并没有得到严格的数学解释。这一解释源自于概率论的词汇——**相互独立**,指的是给定离散型随机变量 X, Y , 如果对于任意 $i, j \in \mathbb{Z}^+$, 都具有

$$\Pr \{y_j|x_i\} = \Pr \{y_j\}.$$

换句话说,即使给定条件 x_i , Y 取到任意 y_j 的概率还是不变。根据(4.13), 我们得到等价表述:

$$\Pr \{x_i, y_j\} = \Pr \{x_i\} \Pr \{y_j\}.$$

正如第一个孩子的性别(视为离散型随机变量 $X = \{\text{男}, \text{女}\}$)与第二个孩子的性别(视为离散型随机变量 $Y = \{\text{男}, \text{女}\}$)是相互独立的: 因为即使第一个孩子是男的, 在怀第二胎的时候, 胎儿是男性、是女性的概率依然都为 $1/2$ 。有了这一数学工具, 我们就可以将之前的表达进行严格的解释:

Theorem 4.5. 对于离散型随机变量 X, Y , 当 X, Y 相互独立时, 互信息 $I(X, Y) = 0$ 。

Proof. 我们只需证明, 当 X, Y 相互独立时, (4.18)取到等号。此时

$$\begin{aligned} \sum_{i=1}^{\infty} \Pr \{x_i\} F(\Pr \{y_j|x_i\}) &= \sum_{i=1}^{\infty} \Pr \{x_i\} F(\Pr \{y_j\}) \\ &= \left(\sum_{i=1}^{\infty} \Pr \{x_i\} \right) F(\Pr \{y_j\}) \\ &= F(\Pr \{y_j\}). \end{aligned}$$

□

根据(), 容易得知 $H(X, Y) = H(Y, X)$; 根据(4.16), 容易得知 $H(X) + H(Y|X) = H(Y) + H(X|Y)$ 。因此, 根据(4.16)与(4.17)可得

$$\begin{aligned} I(X; Y) &= H(Y) - H(Y|X) \\ &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X, Y). \end{aligned}$$

我们可以利用文氏图Figure 4.9表示以上关系。

4.2.2 密度矩阵*

²³ 如果一个体系的量子态为线性组合

$$|\Psi\rangle = \sum_{i=1}^n c^i |\Psi_i\rangle,$$

²³符号*表示这一章节可以跳过, 建议跳转到第163页的章节4.3。

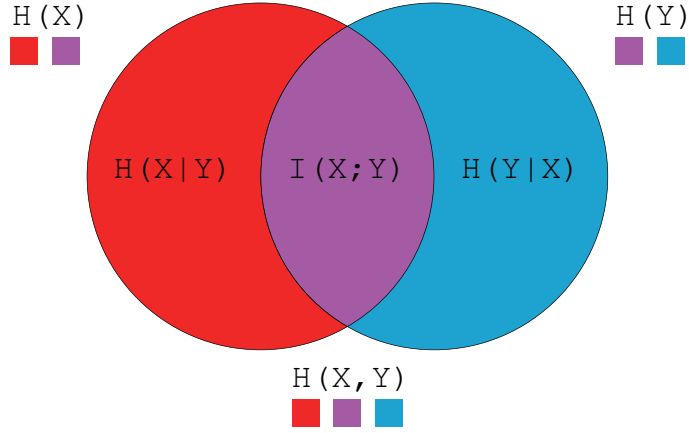


Figure 4.9: 信息熵之间的关系

其中 $c_i \in \mathbb{C}$, $|\Psi_i\rangle$ 为某一可观测量(厄米算子) A 的本征态(在这里, 我们假设 A 的本征态数量是有限的), $i \in [1, n]$ 。这些本征态是标准正交的, 相当于

$$\langle \Psi_i | \Psi_j \rangle = \delta_j^i. \quad (4.19)$$

并且具有归一化条件

$$|c^1|^2 + \dots + |c^n|^2 = 1. \quad (4.20)$$

将任意本征态 $|\Psi_k\rangle$ 表示为 $|1\rangle, \dots, |n\rangle$ 的线性组合:

$$|\Psi_k\rangle = \sum_{j=1}^n U_k^j |j\rangle.$$

因此

$$\begin{aligned} \langle \Psi_i | \Psi_j \rangle &= \left(\sum_{k=1}^n \overline{U_i^k} \langle k| \right) \left(\sum_{l=1}^n U_j^l |l\rangle \right) \\ &= \sum_{k,l=1}^n \overline{U_i^k} U_j^l \langle k|l\rangle = \sum_{k,l=1}^n \overline{U_i^k} U_j^l \delta_l^k \\ &= \sum_{k=1}^n \overline{U_i^k} U_j^k = \delta_j^i. \end{aligned}$$

这表明矩阵 (U_j^i) 是酉矩阵。

根据(2.10), 在测量该可观测量之后, 体系量子态变为 $|\Psi_i\rangle$ 的概率为

$$\Pr \{ \Psi_i \} = |c^i|^2.$$

因此, 根据(4.20)可得

$$\sum_{i=1}^n \Pr \{ \Psi_i \} = 1.$$

据此, 定义**密度矩阵**(实际上应该称为“密度算子”, 但是“密度矩阵”这一用语较为常见)为

$$\rho = \sum_{k=1}^n \Pr \{ \Psi_k \} |\Psi_k\rangle \langle \Psi_k| = \sum_{k=1}^n |c^k|^2 |\Psi_k\rangle \langle \Psi_k|. \quad (4.21)$$

von Neumann熵与信息熵的定义是类似的:

$$S(\rho) = - \sum_{k=1}^n \Pr \{ \Psi_k \} \log \Pr \{ \Psi_k \}.$$

注意: von Neumann熵表达式出现的对数是以 e 为底数的, 信息熵表达式出现的对数是以2为底数的。这与应用范围有关, 因为信息论经常接触二进制编码, 以2为底数会对信息熵的计算带来很大的方便; 而量子力学更着重数学表达式的推导, 在这个情况下以 e 为底数显然是更佳的选择。

因为 $\lim_{t \searrow 0} -t \log t = 0$, 所以就算是存在某些常数 $c^i = 0$, 它对von Neumann熵还是没有影响的。

将密度矩阵表示为 $|i\rangle \langle j|$ 的线性组合:

$$\begin{aligned} \rho &= \sum_{k=1}^n \Pr \{ \Psi_k \} \left(\sum_{i=1}^n U_k^i |i\rangle \right) \left(\sum_{j=1}^n \overline{U_k^j} \langle j| \right) \\ &= \sum_{i,j,k=1}^n \Pr \{ \Psi_k \} U_k^i \overline{U_k^j} |i\rangle \langle j|. \end{aligned}$$

因此, 密度矩阵的矩阵表示为

$$\begin{pmatrix} \rho_1^1 & \cdots & \rho_n^1 \\ \vdots & \ddots & \vdots \\ \rho_1^n & \cdots & \rho_n^n \end{pmatrix},$$

其中

$$\rho_j^i = \sum_{k=1}^n \Pr \{ \Psi_k \} U_k^i \overline{U_k^j}. \quad (4.22)$$

Theorem 4.6. (A) 密度矩阵是厄米算子:

$$\rho = \rho^\dagger.$$

(B) 密度矩阵具有正定性，即对于任意 $|\Phi\rangle$ ，都具有

$$\langle\Phi|\rho|\Phi\rangle \geq 0.$$

(C) 密度矩阵的迹为1，即

$$\text{tr}(\rho) = \sum_{i=1}^n \rho_i^i = 1.$$

其中 ρ_j^i 的表达式如(4.22)所示。

(D) 如果 $\text{Pr}\{\Psi_1\}, \dots, \text{Pr}\{\Psi_n\}$ 当中存在一个取值为1(假设它是 $\text{Pr}\{\Psi_k\}$)的元素，其他元素取值全部为0，那么密度矩阵即为

$$\rho = |\Psi_k\rangle \langle\Psi_k|,$$

我们将它称为**纯态**。它满足

$$\rho^2 = \rho.$$

反之，如果 $\text{Pr}\{\Psi_1\}, \dots, \text{Pr}\{\Psi_n\}$ 存在多个取值不为0的元素(假设它们是 $\text{Pr}\{\Psi_{i_1}\}, \dots, \text{Pr}\{\Psi_{i_p}\}$)，那么密度矩阵即为

$$\rho = \sum_{p=1}^k \text{Pr}\{\Psi_{i_p}\} |\Psi_{i_p}\rangle \langle\Psi_{i_p}|,$$

我们将它称为**混合态**。但是，此时

$$\rho^2 \neq \rho.$$

Proof. 对于第1个性质，根据(2.55)，我们只需证明 $\rho_j^i = \overline{\rho_i^j}$ 即可，这根据(4.22)是显然的。

对于第2个性质，如果利用它的定义(4.21)，就不难得知了：

$$\langle\Phi|\rho|\Phi\rangle = \sum_{k=1}^n \text{Pr}\{\Psi_k\} \langle\Phi|\Psi_k\rangle \langle\Psi_k|\Phi\rangle = \sum_{k=1}^n \text{Pr}\{\Psi_k\} |\langle\Psi_k|\Phi\rangle|^2 \geq 0.$$

对于第3个性质，根据(4.22)，

$$\sum_{i=1}^n \rho_i^i = \sum_{i,k=1}^n \text{Pr}\{\Psi_k\} U_k^i \overline{U_k^i} = \sum_{k=1}^n \text{Pr}\{\Psi_k\} \left(\sum_{i=1}^n |U_k^i|^2 \right).$$

而 (U_j^i) 是酉矩阵，因此对于任意 $k \in [1, n]$ ，都具有 $\sum_{i=1}^n |U_k^i|^2 = 1$ 。这表明

$$\sum_{i=1}^n \rho_i^i = \sum_{k=1}^n \text{Pr}\{\Psi_k\} = 1.$$

对于第4个性质，当 ρ 为纯态时，根据(4.19)：

$$\rho^2 = |\Psi_k\rangle \langle \Psi_k | \Psi_k\rangle \langle \Psi_k| = \langle \Psi_k | \Psi_k\rangle |\Psi_k\rangle \langle \Psi_k| = |\Psi_k\rangle \langle \Psi_k| = \rho.$$

当 ρ 为混合态时，根据(4.19)：

$$\begin{aligned} \rho^2 &= \left(\sum_{p=1}^k \Pr \{ \Psi_{i_p} \} |\Psi_{i_p}\rangle \langle \Psi_{i_p}| \right) \left(\sum_{q=1}^k \Pr \{ \Psi_{i_q} \} |\Psi_{i_q}\rangle \langle \Psi_{i_q}| \right) \\ &= \sum_{p,q=1}^k \Pr \{ \Psi_{i_p} \} \Pr \{ \Psi_{i_q} \} |\Psi_{i_p}\rangle \langle \Psi_{i_p} | \Psi_{i_q}\rangle \langle \Psi_{i_q}| \\ &= \sum_{p,q=1}^k \Pr \{ \Psi_{i_p} \} \Pr \{ \Psi_{i_q} \} \langle \Psi_{i_p} | \Psi_{i_q}\rangle |\Psi_{i_p}\rangle \langle \Psi_{i_q}| \\ &= \sum_{p,q=1}^k \Pr \{ \Psi_{i_p} \} \Pr \{ \Psi_{i_q} \} \delta_{i_q}^{i_p} |\Psi_{i_p}\rangle \langle \Psi_{i_q}| \\ &= \sum_{p=1}^k \Pr \{ \Psi_{i_p} \}^2 |\Psi_{i_p}\rangle \langle \Psi_{i_p}|, \end{aligned}$$

而对于任意 $p \in [1, k]$ ， $\Pr \{ \Psi_{i_p} \}$ 都严格小于1：否则根据 $\sum_{i=1}^n \Pr \{ \Psi_i \} = 1$ ，其余 $\Pr \{ \Psi_i \}$ 都为0， ρ 将不是混合态。这表明对于任意 $p \in [1, k]$ ， $\Pr \{ \Psi_{i_p} \}^2 < \Pr \{ \Psi_{i_p} \}$ ，因此 $\rho^2 \neq \rho$ 。□

与信息熵类似，纯态对应von Neumann熵是0，取到最小值；混合态对应von Neumann熵大于0，并且当 $\Pr \{ \Psi_1 \} = \dots = \Pr \{ \Psi_n \} = 1/n$ 时，von Neumann熵取到最大值 $\log n$ 。

4.3 量子门

如同经典计算机对低电平0与高电平1执行操作一般，量子计算机也会对量子比特 $|\psi\rangle$ 执行操作。我们将执行操作的单元称为**量子门**，它只能对量子比特进行么正变换，即

$$|\Psi\rangle \mapsto U |\Psi\rangle,$$

其中 U 是酉矩阵，定义为 $U^\dagger U = U U^\dagger = I$ 。

常见的量子门如下：

(A) Identity gate:

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

(B) Hadamard gate:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (4.23)$$

(C) T-gate:

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}.$$

(D) Phase gate:

$$S = T^2 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}.$$

可以自行验证一下它们是否为酉矩阵。

与经典的数字电路图一样，我们同样可以制作**量子线路图**，例如：

$$|0\rangle \text{ --- } \boxed{H} \text{ --- } \boxed{T} \text{ --- } \boxed{S} \text{ --- } \boxed{\text{测量}},$$

与数字电路图类似，量子线路图也是从左往右读的，该图的含义为：量子比特 $|0\rangle$ 依次经过了Hadamard gate, T-gate, Phase gate, 最后被测量。经过计算可得

$$\begin{aligned} STH|0\rangle &= \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ ie^{i\pi/4} \end{pmatrix}. \end{aligned}$$

因此，根据(2.10)，量子比特变为 $|0\rangle, |1\rangle$ 的概率分别为

$$\begin{aligned} \Pr\{0\} &= \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}, \\ \Pr\{1\} &= \left| \frac{1}{\sqrt{2}} ie^{i\pi/4} \right|^2 = \frac{1}{2}. \end{aligned}$$

4.4 双量子比特

经典计算机的逻辑运算不仅含有单目运算(自变量只有一个，例如NOT)，也含有双目运算(自变量有2个，例如XOR)。类比到量子计算机上，我们同样需要定义“双目运算”，此时需要2个量子比特，因此就引出了**双量子比特状态**的概念。

首先，对于2个状态分别为 $|\Psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, |\Phi\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$ 的量子比特，它们的Kronecker Product为

$$|\Psi\rangle \otimes |\Phi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \\ \beta \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix},$$

它是一个4维向量。有些时候，为了表示方便，我们可以把它记为 $|\Psi\rangle|\Phi\rangle$ 。

我们可以利用任意一个归一化的4维向量表示双量子比特状态，记它为

$$|\alpha\rangle = \begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix},$$

其中 $|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$ 。

我们记

$$\begin{aligned} |00\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\ |01\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\ |10\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\ |11\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

很显然，它是向量空间 \mathbb{C}^4 上的一组标准正交基，因此，可以将 α 表示为

$$|\alpha\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle,$$

与前面的论述类似，测量该双量子比特，它变为 $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ 的概率分别为

$$\begin{aligned} \Pr\{00\} &= |\alpha_{00}|^2, \\ \Pr\{01\} &= |\alpha_{01}|^2, \\ \Pr\{10\} &= |\alpha_{10}|^2, \\ \Pr\{11\} &= |\alpha_{11}|^2. \end{aligned}$$

虽然我们可以计算2个量子比特的Kronecker Product来得到一个4维向量，但是这并不

代表任意一个4维向量都可以被分解为2个量子比特的Kronecker Product，例如

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

它无法被表示为 $|\Psi\rangle \otimes |\Phi\rangle$ 的形式。利用反证法，假设它可以，那么我们记

$$|\Phi^+\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \\ \beta \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix},$$

因此

$$\begin{aligned} \alpha\gamma &= \frac{1}{\sqrt{2}}, \\ \alpha\delta &= 0, \\ \beta\gamma &= 0, \\ \beta\delta &= \frac{1}{\sqrt{2}}. \end{aligned}$$

根据第2个表达式可得 $\alpha = 0$ 或 $\delta = 0$ ，但是，此时第1个表达式或第4个表达式必有一个结果为0。因此，这是不可能的。我们称这样的双量子比特状态为**纠缠态**，在后面的章节里会涉及。

同样道理，量子计算机对双量子比特状态的操作也是对其进行么正变换，不过此时对应的酉矩阵是4阶的。例如，以下的4阶矩阵

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (4.24)$$

是酉矩阵，称为CNOT gate。可以计算得出

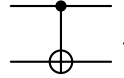
$$\begin{aligned} \text{CNOT} |00\rangle &= |00\rangle, \\ \text{CNOT} |01\rangle &= |01\rangle, \\ \text{CNOT} |10\rangle &= |11\rangle, \\ \text{CNOT} |11\rangle &= |10\rangle. \end{aligned}$$

CNOT gate之所以很重要，是因为它可以制备出纠缠态。定义 $|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)$ ，因此

$$\text{CNOT} |+\rangle |0\rangle = \text{CNOT} \left(\frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |10\rangle \right) = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) = |\Phi^+\rangle.$$

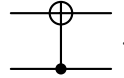
注意：之后的语句仅对“输入值为 $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ 这四者中的其一”而言是成立的，即第一个量子比特与第二个量子比特都只能为 $|0\rangle$ 或 $|1\rangle$ 。对于一般的情况 $|\alpha\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$ ，我们可以通过量子线路图计算出 $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ 对应的输出值，再进行线性组合。

因此，我们可以得出，CNOT gate不会改变第一个量子比特的状态。而对于第二个量子比特的状态是否会发生改变，取决于第一个量子比特的状态：当第一个量子比特的状态为 $|0\rangle$ 时，第二个量子比特的状态不会改变；当第一个量子比特的状态为 $|1\rangle$ 时，第二个量子比特的状态会取反，即 $|0\rangle \mapsto |1\rangle, |1\rangle \mapsto |0\rangle$ 。我们将CNOT gate对应的量子线路图记为



可以看到，该图的第一行是一个实心黑点，在电路图里等效于电线连接点，因此第一行的输出值与输入值始终相等；该图的第二行是一个符号 \oplus ，并且上面接了一根与第一行相连的竖线，因此第二行的输出值是第一行的输出值与第二行的输出值的“异或(XOR)”。这也很容易验证，因为 $0 \oplus 1 = 1, 0 \oplus 0 = 0, 1 \oplus 0 = 1, 1 \oplus 1 = 0$ 。

如果我们将第一个量子比特与第二个量子比特的“身份”交换一下，并且，记对应的4阶矩阵为CNOT₁₀ gate，那么上述文字内容可以被改为：CNOT₁₀ gate不会改变第二个量子比特的状态。而对于第一个量子比特的状态是否会发生改变，取决于第二个量子比特的状态：当第二个量子比特的状态为 $|0\rangle$ 时，第一个量子比特的状态不会改变；当第二个量子比特的状态为 $|1\rangle$ 时，第一个量子比特的状态会取反，即 $|0\rangle \mapsto |1\rangle, |1\rangle \mapsto |0\rangle$ 。我们将CNOT₁₀ gate对应的量子线路图记为



可以得出

$$\begin{aligned} \text{CNOT}_{10} |00\rangle &= |00\rangle, \\ \text{CNOT}_{10} |01\rangle &= |11\rangle, \\ \text{CNOT}_{10} |10\rangle &= |10\rangle, \\ \text{CNOT}_{10} |11\rangle &= |01\rangle, \end{aligned}$$

$$\text{CNOT}_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \quad (4.25)$$

我们也可以通过计算2个量子门的Kronecker Product来得到一个4阶酉矩阵。对于2阶酉矩阵

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad N = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

它们的Kronecker Product为

$$M \otimes N = \begin{pmatrix} a \begin{pmatrix} e & f \\ g & h \end{pmatrix} & b \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ c \begin{pmatrix} e & f \\ g & h \end{pmatrix} & d \begin{pmatrix} e & f \\ g & h \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}.$$

根据这一概念，我们可以得到一些4阶酉矩阵，例如

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

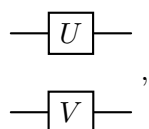
如果 U, V 是量子门， $|\Psi\rangle, |\Phi\rangle$ 是量子比特，可以得到

$$(U \otimes V) (|\Psi\rangle \otimes |\Phi\rangle) = (U |\Psi\rangle) \otimes (V |\Phi\rangle).$$

例如，因为 $H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ，所以

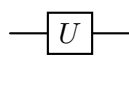
$$\begin{aligned} (H \otimes H) |00\rangle &= (H \otimes H) (|0\rangle \otimes |0\rangle) \\ &= (H |0\rangle) \otimes (H |0\rangle) \\ &= \frac{1}{2} (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \\ &= \frac{1}{2} (|00\rangle + |01\rangle + |10\rangle + |11\rangle). \end{aligned}$$

因此，仅通过 $U \otimes V$ 形式的4阶矩阵是无法制备出纠缠态的。但是CNOT可以制备出纠缠态，所以它无法被表示为 $U \otimes V$ 的形式。我们将 $U \otimes V$ 的量子线路图记为

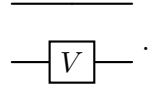


可以看到，该图的第一行与第二行互不相连，可以认为是“各走各的”，所以你完全可以将它视为2个不相关的线路：上面的量子比特经过量子门 U ，下面的量子比特经过量子门 V ，两者互不干涉。因此，第一行的输出值为 $|\Psi\rangle \mapsto U |\Psi\rangle$ ，第二行的输出值为 $|\Phi\rangle \mapsto V |\Phi\rangle$ ，其中 $|\Psi\rangle, |\Phi\rangle \in \{|0\rangle, |1\rangle\}$ 。

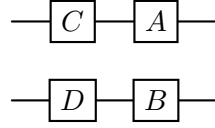
对于 $U \otimes I$ 与 $I \otimes V$ 的情况，通常将 I 略去不写，分别记为



与



而 $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ (在这里不做证明), 所以量子线路图



既可以视为 $(A \otimes B)(C \otimes D)$, 又可以视为 $(AC) \otimes (BD)$ 。因此, 第一行的输出值为 $|\Psi\rangle \mapsto AC|\Psi\rangle$, 第二行的输出值为 $|\Phi\rangle \mapsto BD|\Phi\rangle$, 其中 $|\Psi\rangle, |\Phi\rangle \in \{|0\rangle, |1\rangle\}$ 。

根据CNOT gate的思路, 我们可以定义controlled-U gate, 其中 U 是量子门:

$$\begin{aligned} CU|00\rangle &= |00\rangle, \\ CU|01\rangle &= |01\rangle, \\ CU|10\rangle &= |1\rangle \otimes (U|0\rangle), \\ CU|11\rangle &= |1\rangle \otimes (U|1\rangle). \end{aligned}$$

因此, 我们可以得出, controlled-U gate不会改变第一个量子比特的状态。而对于第二个量子比特的状态是否会发生改变, 取决于第一个量子比特的状态: 当第一个量子比特的状态为 $|0\rangle$ 时, 第二个量子比特的状态不会改变; 当第一个量子比特的状态为 $|1\rangle$ 时, 第二个量子比特的状态会变为 $|\Psi\rangle \mapsto U|\Psi\rangle$, 其中 $|\Psi\rangle \in \{|0\rangle, |1\rangle\}$ 。因此, 当

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

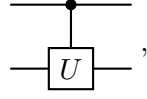
时, 那么

$$\begin{aligned} CU|00\rangle &= |00\rangle, \\ CU|01\rangle &= |01\rangle, \\ CU|10\rangle &= |1\rangle \otimes (a|0\rangle + b|1\rangle) = a|10\rangle + b|11\rangle, \\ CU|11\rangle &= |1\rangle \otimes (c|0\rangle + d|1\rangle) = c|10\rangle + d|11\rangle, \end{aligned}$$

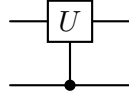
因此

$$CU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix}.$$

我们将controlled-U gate的量子线路图记为



可以看到，第一行的输出值与输入值始终相等；在该图的第二行，U gate的侧面接了一根与第一行相连的竖线。可以将它理解为“开关”：当第一行的量子比特为 $|0\rangle$ 时，“开关”是关闭的，因此U gate不起作用，controlled-U gate也不起作用；当第一行的量子比特为 $|1\rangle$ 时，“开关”是开启的，因此U gate起到作用，第二行的输出值为 $|\Psi\rangle \mapsto U|\Psi\rangle$ ，其中 $|\Psi\rangle \in \{|0\rangle, |1\rangle\}$ 。因此，量子线路图



的含义也就可想而知了，记其为 CU_{10} ，因此

$$CU_{10} |00\rangle = |00\rangle,$$

$$CU_{10} |01\rangle = (U|0\rangle) \otimes |1\rangle,$$

$$CU_{10} |10\rangle = |10\rangle,$$

$$CU_{10} |11\rangle = (U|1\rangle) \otimes |1\rangle,$$

$$CU_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & c \\ 0 & 0 & 1 & 0 \\ 0 & b & 0 & d \end{pmatrix}.$$

我们还可以定义SWAP gate:

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

顾名思义，它很显然是交换2个量子比特的位置。因此

$$\text{SWAP} |00\rangle = |00\rangle,$$

$$\text{SWAP} |01\rangle = |10\rangle,$$

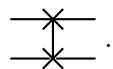
$$\text{SWAP} |10\rangle = |01\rangle,$$

$$\text{SWAP} |11\rangle = |11\rangle.$$

我们将SWAP gate的量子线路图记为



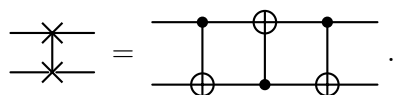
(非常直观吧) 但是, 有些时候, 为了避免造成不必要的麻烦, 我们又可以将它记为



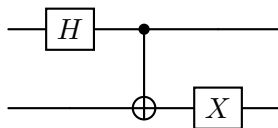
根据计算得出

$$\begin{aligned} \text{SWAP} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= (\text{CNOT})(\text{CNOT}_{10})(\text{CNOT}), \end{aligned}$$

因此, SWAP gate可以被三个CNOT gate替代:

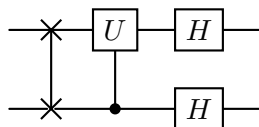


结合以上信息, 我们现在分析一些比较复杂的量子线路图, 例如



等效为

$$(I \otimes X)(\text{CNOT})(H \otimes I), \quad (4.26)$$



等效为

$$(H \otimes H)(CU_{10})(\text{SWAP}).$$

4.4.1 多量子比特

与双量子比特状态类似, 多量子比特状态可以被表示为

$$\bigotimes_{i=1}^n |\Psi_i\rangle,$$

其中 n 为量子比特的个数, $|\Psi_i\rangle = \alpha_i |0\rangle + \beta_i |1\rangle$, $|\alpha_i|^2 + |\beta_i|^2 = 1$ 是第 i 个量子比特的状态。因此, 它可以表示为一个 2^n 维的向量, 并且可以表示为向量空间 \mathbb{C}^{2^n} 的标准正交基

$$\bigotimes_{i=1}^n |\chi_i\rangle$$

的线性组合, 其中 $|\chi_i\rangle \in \{|0\rangle, |1\rangle\}$ 。例如, 若 $n = 3$, 那么多量子比特状态可以被表示为

$$|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle, |101\rangle, |110\rangle, |111\rangle$$

的线性组合。

同样道理, 量子计算机对多量子比特状态的操作也是对其进行么正变换, 不过此时对应的酉矩阵是 2^n 阶的。由于篇幅有限, 在此不做展开。

4.5 量子力学受到的挑战

在这一部分, 我们只讨论双量子比特状态。这也是我将多量子比特状态这一小节一笔带过的其中一部分原因。

4.5.1 EPR佯谬

考虑 π^0 粒子的弱衰变:

$$\pi^0 \rightarrow e + e^+,$$

其中 e, e^+ 分别为电子与正电子, 它们的静止质量是相同的。如果一开始 π^0 粒子是静止的, 那么根据动量守恒定律, 电子与正电子的运动速度大小相同, 方向相反; 因为 π^0 的自旋为0, 根据角动量守恒定律, 电子与正电子组成的体系的波函数对应总角动量平方与总角动量 z 分量均为0。因此, 根据(3.36), 体系的波函数只能为 $\chi_{0,0}$:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle).$$

假设在某个时间点, 电子与正电子相距 c 米远。此时去测量电子的自旋, 很显然得到自旋向上 $|0\rangle$ 的概率为 $1/2$, 自旋向下 $|1\rangle$ 的概率也为 $1/2$ 。但是, 如果得到自旋向上, 那么波函数将变为

$$|\tilde{\Psi}\rangle = |01\rangle,$$

正电子的自旋瞬间被确定下来, 它是向下的!

但是, 根据爱因斯坦的狭义相对论, 任何信息的传播速度都不能超过光速, 所以在测量之后的瞬间, 至少得再过1秒, 正电子的自旋状态才能发生改变。爱因斯坦等人根据这一矛盾(称为**EPR佯谬**, E对应爱因斯坦的名称, P, R对应爱因斯坦两位助手的名称), 在1935年(比广义相对论的提出还要晚!)发布了一篇文献“Can quantum mechanical description of physical reality be considered complete?”, 他们认为, 量子力学并不是一个完备的理论。

关于该矛盾，有人设想，波函数的变化不是瞬时的，而是以不超过光速的速度传播的：比如，在上面的例子中，可以设想，波函数在测量之后的瞬间，至少得再过1秒，才能变化为

$$|\tilde{\Psi}\rangle = |01\rangle,$$

但是该设想马上被推翻了：假设在这1秒之内(记时间点为 t_0 ，其中 $0 < t_0 < 1$)去测量正电子的自旋，因为此时波函数并未发生改变，所以，得到自旋向上 $|0\rangle$ 的概率为 $1/2$ ，自旋向下 $|1\rangle$ 的概率也为 $1/2$ 。如果得到自旋向上，那么在时间段 $(t_0, 1)$ ，电子与正电子的自旋都是向上的，这与角动量守恒定律相悖(动量守恒定律、能量守恒定律、角动量守恒定律是实验得知的结果，永远不可能被推翻)。因此，无论如何，波函数的变化都是瞬时发生的。

对于EPR佯谬，当时有2个不同的主流观点：

OPINION 1: 在 π^0 粒子衰变之后的瞬间，电子与正电子的自旋方向就已经是相反的了，只是量子力学理论无法告知而已。因此，该观点是对量子力学的挑战。

OPINION 2: 在 π^0 粒子衰变之后的瞬间，电子与正电子的自旋方向是不确定的，换言之，我们并不知道它们相反，只能通过测量才能得知。在测量电子自旋状态之后的瞬间，波函数直接变化。因此，该观点是对相对论的挑战。

最后通过生活中的实例进行类比。假设Alice与Bob两人参加一场考试，如果Alice与Bob的作答结果是完全相同的，那么老师肯定会猜测，这两人考试之前是否有过串通？OPINION 1对应，老师的猜测是正确的；OPINION 2对应，老师的猜测是错误的。

4.5.2 贝尔定理

Theorem 4.7. 对于 \mathbb{R}^3 上长度为1的向量(单位向量) $\mathbf{N} = (N^1, N^2, N^3)$ ，其中 $N^1, N^2, N^3 \in \mathbb{R}, (N^1)^2 + (N^2)^2 + (N^3)^2 = 1$ 。那么矩阵

$$\sigma(\mathbf{N}) = N^1\sigma_1 + N^2\sigma_2 + N^3\sigma_3$$

的本征值只可能为 ± 1 。其中 $\sigma_1, \sigma_2, \sigma_3$ 为泡利矩阵。

Proof. 根据(3.10)，可以计算得出

$$\begin{aligned}\sigma_1|0\rangle &= |1\rangle, \\ \sigma_1|1\rangle &= |0\rangle, \\ \sigma_2|0\rangle &= i|1\rangle, \\ \sigma_2|1\rangle &= -i|0\rangle, \\ \sigma_3|0\rangle &= |0\rangle, \\ \sigma_3|1\rangle &= -|1\rangle.\end{aligned}$$

因为 $\sigma(\mathbf{N})$ 是2阶矩阵，所以线性无关的本征态最多有2个，分别记为

$$\begin{aligned} |\Psi\rangle &= \begin{pmatrix} a \\ b \end{pmatrix} = a|0\rangle + b|1\rangle, \\ |\Phi\rangle &= \begin{pmatrix} c \\ d \end{pmatrix} = c|0\rangle + d|1\rangle, \end{aligned}$$

因为本征态与一个非零的常数相乘还是本征态，所以我们假设 $|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1$ 。记 $|\Psi\rangle, |\Phi\rangle$ 对应的本征值为 λ, μ ，因此

$$\begin{aligned} \sigma(\mathbf{N})|\Psi\rangle &= \lambda|\Psi\rangle, \\ \sigma(\mathbf{N})|\Phi\rangle &= \mu|\Phi\rangle, \end{aligned}$$

因此

$$\begin{aligned} (N^1\sigma_1 + N^2\sigma_2 + N^3\sigma_3)(a|0\rangle + b|1\rangle) &= \lambda(a|0\rangle + b|1\rangle), \\ (N^1\sigma_1 + N^2\sigma_2 + N^3\sigma_3)(c|0\rangle + d|1\rangle) &= \mu(c|0\rangle + d|1\rangle), \end{aligned}$$

展开可得

$$\begin{aligned} N^1a|1\rangle + N^1b|0\rangle + iN^2a|1\rangle - iN^2b|0\rangle + N^3a|0\rangle - N^3b|1\rangle &= \lambda a|0\rangle + \lambda b|1\rangle, \\ N^1c|1\rangle + N^1d|0\rangle + iN^2c|1\rangle - iN^2d|0\rangle + N^3c|0\rangle - N^3d|1\rangle &= \mu c|0\rangle + \mu d|1\rangle, \end{aligned}$$

因此

$$\begin{aligned} N^1b - iN^2b + N^3a &= \lambda a, \\ N^1a + iN^2a - N^3b &= \lambda b, \\ N^1d - iN^2d + N^3c &= \mu c, \\ N^1c + iN^2c - N^3d &= \mu d, \end{aligned}$$

因此

$$\begin{aligned} (N^3 - \lambda)a + (N^1 - iN^2)b &= 0, \\ (N^1 + iN^2)a - (N^3 + \lambda)b &= 0, \\ (N^3 - \mu)c + (N^1 - iN^2)d &= 0, \\ (N^1 + iN^2)c - (N^3 + \mu)d &= 0, \end{aligned} \tag{4.27}$$

因为 $(a, b) \neq (0, 0), (c, d) \neq (0, 0)$ ，而(4.27)存在非零解当且仅当

$$\begin{aligned} \det \begin{pmatrix} N^3 - \lambda & N^1 - iN^2 \\ N^1 + iN^2 & -N^3 - \lambda \end{pmatrix} &= 0 \\ \det \begin{pmatrix} N^3 - \mu & N^1 - iN^2 \\ N^1 + iN^2 & -N^3 - \mu \end{pmatrix} &= 0, \end{aligned}$$

根据 $(N^1)^2 + (N^2)^2 + (N^3)^2 = 1$ 可得 $\lambda^2 = \mu^2 = 1$, 因此 $\lambda, \mu = \pm 1$ 。

如果 $\lambda = \mu$, 记它们的取值均为 θ , 那么 $\sigma(\mathbf{N})|\Psi\rangle = \theta|\Psi\rangle, \sigma(\mathbf{N})|\Phi\rangle = \theta|\Phi\rangle$, 因此 $|\Psi\rangle$ 与 $|\Phi\rangle$ 的任意线性组合也是 $\sigma(\mathbf{N})$ 的本征向量。但是 $|\Psi\rangle$ 与 $|\Phi\rangle$ 是线性无关的, 因此向量空间 \mathbb{C}^2 上的所有向量都是 $\sigma(\mathbf{N})$ 的本征向量, 这显然是不可能的。

因此, $\lambda = 1, \mu = -1$ ($\lambda = -1, \mu = 1$ 是完全一样的, 没必要重复分析), 对应的特征向量为

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2(1-N^3)}} \begin{pmatrix} N^1 - iN^2 \\ 1 - N^3 \end{pmatrix}, \\ |\Phi\rangle &= \frac{1}{\sqrt{2(1+N^3)}} \begin{pmatrix} N^1 - iN^2 \\ -1 - N^3 \end{pmatrix}. \end{aligned} \quad (4.28)$$

□

根据这个定理, 对于任意单位向量 $\mathbf{N} = (N^1, N^2, N^3)$, 我们可以定义可观测量 $\sigma(\mathbf{N}) = N^1\sigma_1 + N^2\sigma_2 + N^3\sigma_3$ (表明 $\sigma(\mathbf{N})$ 是 \mathbf{N} 的函数), 测量它得到的结果只能为 ± 1 。

回到上一小节对于电子与正电子的测量, 不过此时测量的可观测量不是 σ_3 , 而是 $\sigma(\mathbf{N})$ 。我们将列出以下步骤:

STEP 1: 给定单位向量 \mathbf{A}, \mathbf{B} , 首先测量电子的可观测量 $\sigma(\mathbf{A})$, 测量结果记为 a , a 只能为 ± 1 。

STEP 2: 在测量结束后, 立刻测量正电子的可观测量 $\sigma(\mathbf{B})$, 测量结果记为 b , b 只能为 ± 1 。

STEP 3: 记 $c = ab$ 为两次测量结果的乘积, 因此, c 也只能为 ± 1 。

STEP 4: 反复执行以上操作, 每次操作结束后, 都将对应的 c 值记下来。记 $\Pr\{\mathbf{A}, \mathbf{B}\}$ 为这些 c 值的平均数。

当 $\mathbf{A} = \mathbf{B} = (0, 0, 1)$ 时, $\sigma(\mathbf{A}) = \sigma(\mathbf{B}) = \sigma_3$, 因此测量电子与正电子对应的可观测量均为 σ_3 , 这与上一小节的情况是相同的。

如果认为上一小节的OPINION 2是正确的(认为量子力学的假设是合理的), 那么, 当 $a = 1$ 时, $b = -1$; 当 $a = -1$ 时, $b = 1$ 。所以无论测量多少次, 对应的 c 值都为 -1 , 因此 $\Pr\{\mathbf{A}, \mathbf{B}\} = -1$ 。

Theorem 4.8. 对于任意单位向量 $\mathbf{A} = (A^1, A^2, A^3), \mathbf{B} = (B^1, B^2, B^3)$, 都具有

$$\Pr\{\mathbf{A}, \mathbf{B}\} = -\mathbf{A} \cdot \mathbf{B} = -A^1B^1 - A^2B^2 - A^3B^3,$$

因为计算难度较大, 所以我们不做证明了。因此, 当 $\mathbf{A} = \mathbf{B}$ 时, $-\mathbf{A} \cdot \mathbf{B} = -1$, 这表明电子的测量结果 a 与正电子的测量结果 b 一定满足 $a + b = 0$! 这与 \mathbf{A}, \mathbf{B} 是否等于 $(0, 0, 1)$ 是完全无关的。

但是, 贝尔认为上一小节的OPINION 1是正确的(认为相对论的假设是合理的)。据此, 他推导出了贝尔定理:

Theorem 4.9. 对于任意单位向量 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ ，都具有

$$|\Pr\{\mathbf{A}, \mathbf{B}\} - \Pr\{\mathbf{A}, \mathbf{C}\}| \leq 1 + \Pr\{\mathbf{B}, \mathbf{C}\}.$$

该定理就不做证明了(因为我根本看不懂，太抽象了)。可以看到，如果Theorem 4.8是成立的，那么选取 $\mathbf{A} = (1, 0, 0)$, $\mathbf{B} = (0, 1, 0)$, $\mathbf{C} = \frac{1}{\sqrt{2}}(\mathbf{A} + \mathbf{B})$ ，可以计算得出

$$\begin{aligned}\Pr\{\mathbf{A}, \mathbf{B}\} &= -\mathbf{A} \cdot \mathbf{B} = 0, \\ \Pr\{\mathbf{A}, \mathbf{C}\} &= -\mathbf{A} \cdot \mathbf{C} = -\frac{1}{\sqrt{2}}, \\ \Pr\{\mathbf{B}, \mathbf{C}\} &= -\mathbf{B} \cdot \mathbf{C} = -\frac{1}{\sqrt{2}},\end{aligned}$$

因此

$$\begin{aligned}|\Pr\{\mathbf{A}, \mathbf{B}\} - \Pr\{\mathbf{A}, \mathbf{C}\}| &= \frac{1}{\sqrt{2}} \approx 0.7, \\ 1 + \Pr\{\mathbf{B}, \mathbf{C}\} &= 1 - \frac{1}{\sqrt{2}} \approx 0.3,\end{aligned}$$

这很显然与贝尔定理矛盾。

在1970年代，物理学家做了很多实验以验证贝尔定理是否成立，但是可惜的是，有些实验的结果推翻了贝尔定理：相当于，贝尔的猜想是错误的，量子力学的假设是合理的。这些实验在物理学界产生了极大的震动。据此，有些人认为，量子力学允许所谓“超距作用”的存在，从而量子力学与相对论在本质上是矛盾的；但是，也有些人认为，这两次测量它并没有所谓的因果关系，因此并不能算得上是传递信息，那超过光速也是情有可原的，从而量子力学与相对论在本质上并不矛盾。

目前一些物理学家正在尝试如何将量子力学与相对论进行统一，其中弦理论、量子场论等理论就在尝试进行这个工作，这些理论是目前物理学界非常热门的话题。

4.6 不可克隆定理

对于经典计算机，我们可以很轻松地将信息“克隆”出来—只需读取该信息的比特，然后复制一份即可：如果在键盘上按下复制粘贴快捷键Command + C与Command + V，那么就实现了信息的“克隆”。

对于量子计算机，所谓的“克隆”就几乎不可实现了。试想，如果给定一个未知的量子比特 $|\Psi\rangle$ ，你该如何克隆它呢？换句话说，假设给定任意一个双量子比特状态 $|\Psi\rangle \otimes |0\rangle$ ，能否找到一个量子门 U ，使得 $U|\Psi\rangle \otimes |0\rangle = |\Psi\rangle \otimes |\Psi\rangle$ 呢？

为了得知这个等式是否能够成立，假设 $U = (U_{ij})$ ，其中第 i 行，第 j 列的元素记为 U_{ij} ，并假设 $|\Psi\rangle = \alpha|1\rangle + \beta|0\rangle$ 。因此

$$\begin{pmatrix} U_{11} & U_{12} & U_{13} & U_{14} \\ U_{21} & U_{22} & U_{23} & U_{24} \\ U_{31} & U_{32} & U_{33} & U_{34} \\ U_{41} & U_{42} & U_{43} & U_{44} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

$$\begin{pmatrix} U_{11} & U_{12} & U_{13} & U_{14} \\ U_{21} & U_{22} & U_{23} & U_{24} \\ U_{31} & U_{32} & U_{33} & U_{34} \\ U_{41} & U_{42} & U_{43} & U_{44} \end{pmatrix} \begin{pmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha^2 \\ \alpha\beta \\ \alpha\beta \\ \beta^2 \end{pmatrix},$$

$$U_{11}\alpha + U_{13}\beta = \alpha^2,$$

$$U_{21}\alpha + U_{23}\beta = \alpha\beta,$$

$$U_{31}\alpha + U_{33}\beta = \alpha\beta,$$

$$U_{41}\alpha + U_{43}\beta = \beta^2.$$

可以得出，该方程组不存在常数解，即 U_{ij} 的取值与 $|\Psi\rangle$ 有关，但是 $|\Psi\rangle$ 是未知的，所以我们无法制造出量子门 U 。因此，未知的量子比特是无法克隆的，该命题称为**不可克隆定理**。

4.7 纠缠态

对于无法表示成形式 $|\Psi\rangle \otimes |\Phi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$ 的双量子比特，我们将它称为**纠缠态**。常见的纠缠态如下：

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle), \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle), \\ |\Psi^-\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle), \end{aligned} \tag{4.29}$$

这4个双量子比特状态被称为**贝尔态**。

量子隐形传态是一种利用量子纠缠来传送量子态至任意距离的技术。它并不会传递任何的物质或能量，因此，在这个情况下，谈论超光速是没有意义的。我们将分析它的步骤：

- (A) 定义 V_1, V_2, V_3 为三个量子线路上量子比特所有状态的集合，它们均为向量空间。发送方(记为Alice)将任意量子态 $|\Psi\rangle = a|0\rangle + b|1\rangle \in V_1$ 与贝尔态 $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in$

$V_2 \otimes V_3$ 进行张量积, 得到 $V_1 \otimes (V_2 \otimes V_3)$ 上的一个向量:

$$\begin{aligned} |\Psi\rangle \otimes |\Phi^+\rangle &= (a|0\rangle + b|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ &= \frac{1}{\sqrt{2}}(a|0\rangle \otimes |00\rangle + a|0\rangle \otimes |11\rangle + b|1\rangle \otimes |00\rangle + b|1\rangle \otimes |11\rangle). \end{aligned}$$

(B) 利用同构 $V_1 \otimes (V_2 \otimes V_3) \cong (V_1 \otimes V_2) \otimes V_3$, 我们可以将 $|\Psi\rangle \otimes |\Phi^+\rangle$ 视为 $(V_1 \otimes V_2) \otimes V_3$ 上的向量:

$$|\Psi\rangle \otimes |\Phi^+\rangle \longleftrightarrow \frac{1}{\sqrt{2}}(a|00\rangle \otimes |0\rangle + a|01\rangle \otimes |1\rangle + b|10\rangle \otimes |0\rangle + b|11\rangle \otimes |1\rangle).$$

(C) 令 $\omega^1, \omega^2, \omega^3, \omega^4$ 是向量空间 $V_1 \otimes V_2$ 到复数域 \mathbb{C} 上的线性映射, 并且满足

$$\begin{aligned} \omega^1(|\Phi^+\rangle) &= 1, & \omega^1(|\Phi^-\rangle) &= 0, & \omega^1(|\Psi^+\rangle) &= 0, & \omega^1(|\Psi^-\rangle) &= 0, \\ \omega^2(|\Phi^+\rangle) &= 0, & \omega^2(|\Phi^-\rangle) &= 1, & \omega^2(|\Psi^+\rangle) &= 0, & \omega^2(|\Psi^-\rangle) &= 0, \\ \omega^3(|\Phi^+\rangle) &= 0, & \omega^3(|\Phi^-\rangle) &= 0, & \omega^3(|\Psi^+\rangle) &= 1, & \omega^3(|\Psi^-\rangle) &= 0, \\ \omega^4(|\Phi^+\rangle) &= 0, & \omega^4(|\Phi^-\rangle) &= 0, & \omega^4(|\Psi^+\rangle) &= 0, & \omega^4(|\Psi^-\rangle) &= 1. \end{aligned}$$

而

$$\begin{aligned} |\Psi\rangle \otimes |\Phi^+\rangle &= \frac{1}{\sqrt{2}}(a|00\rangle \otimes |0\rangle + a|01\rangle \otimes |1\rangle + b|10\rangle \otimes |0\rangle + b|11\rangle \otimes |1\rangle) \\ &= \frac{1}{2}(a(|\Phi^+\rangle + |\Phi^-\rangle) \otimes |0\rangle + a(|\Psi^+\rangle + |\Psi^-\rangle) \otimes |1\rangle \\ &\quad + b(|\Psi^+\rangle - |\Psi^-\rangle) \otimes |0\rangle + b(|\Phi^+\rangle - |\Phi^-\rangle) \otimes |1\rangle). \end{aligned}$$

利用同构关系 $\mathbb{C} \otimes V_3 \cong V_3, a \otimes v \mapsto av$, 有4种情况:

(a) 如果接收方(记为Bob)选取了线性映射 $\omega^1 \otimes I$, 那么

$$(\omega^1 \otimes I)(|\Psi\rangle \otimes |\Phi^+\rangle) = \frac{1}{2}(a|0\rangle + b|1\rangle).$$

它是 V_3 上的向量, 对其进行归一化, 并且转换为矩阵表示:

$$\begin{pmatrix} a \\ b \end{pmatrix}.$$

这就是Alice发送的量子态。

(b) 如果接收方(记为Bob)选取了线性映射 $\omega^2 \otimes I$, 那么

$$(\omega^2 \otimes I)(|\Psi\rangle \otimes |\Phi^+\rangle) = \frac{1}{2}(a|0\rangle - b|1\rangle).$$

它是 V_3 上的向量, 对其进行归一化, 并且转换为矩阵表示:

$$\begin{pmatrix} a \\ -b \end{pmatrix}.$$

令其经过量子门 $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, 就能返回到Alice发送的量子态:

$$\begin{pmatrix} a \\ -b \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ -b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

(c) 如果接收方(记为Bob)选取了线性映射 $\omega^3 \otimes I$, 那么

$$(\omega^3 \otimes I)(|\Psi\rangle \otimes |\Phi^+\rangle) = \frac{1}{2}(b|0\rangle + a|1\rangle).$$

它是 V_3 上的向量, 对其进行归一化, 并且转换为矩阵表示:

$$\begin{pmatrix} b \\ a \end{pmatrix}.$$

令其经过量子门 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, 就能返回到Alice发送的量子态:

$$\begin{pmatrix} b \\ a \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

(d) 如果接收方(记为Bob)选取了线性映射 $\omega^4 \otimes I$, 那么

$$(\omega^4 \otimes I)(|\Psi\rangle \otimes |\Phi^+\rangle) = \frac{1}{2}(-b|0\rangle + a|1\rangle).$$

它是 V_3 上的向量, 对其进行归一化, 并且转换为矩阵表示:

$$\begin{pmatrix} -b \\ a \end{pmatrix}.$$

令其经过量子门 $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, 就能返回到Alice发送的量子态:

$$\begin{pmatrix} -b \\ a \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

我们可以通过图像来辅助理解, 如Figure 4.10所示。实际上, 它并不会与量子不可克隆定理产生矛盾: 因为Alice已经将发送的量子态与贝尔态进行张量积, Bob再通过张量积的结果还原Alice发送的量子态, 此时Alice这里已经不存在这一量子态。因此, 量子隐形传态是将量子态进行传输, 而不是对量子态进行克隆: 就像一辆车从起点驶向终点, 当车到达终点时, 起点自然是不会存在这辆车的。

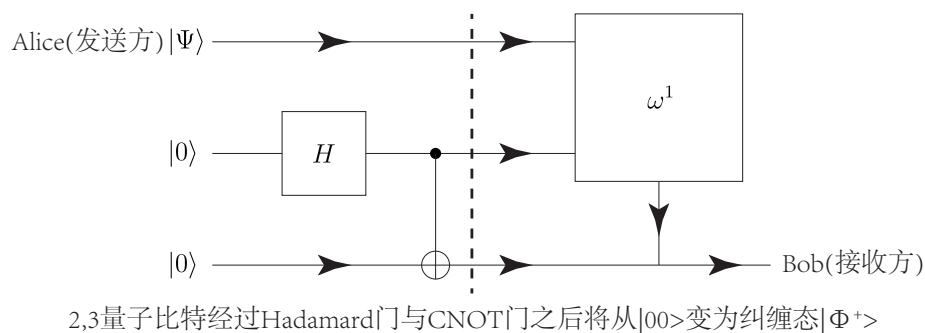


Figure 4.10: 量子隐形传态

5 例题

大量的习题，甚至还包含了一些考试的题目。主要是它们可以与之前的章节学习相结合，达到查漏补缺的作用。“官方资料”的习题答案我也写了一份，链接为[点击我就可以啦！](#)

这份资料也就到此为止了。可惜的是，至于光学部分，本人对此完全没有兴趣，也完全不知道怎么写，因此本文没有光学部分，在此感到抱歉。无论如何，抱着轻松的心情阅读这份资料吧！（不过考试周“预习”这门课程的时候就没有那么轻松了）

量子计算部分我尝试写过一些，但是后来发现自己无法将其解释清楚，因此全部删除了，毕竟我自己也无法理解其中的概念。

5.1 一维无限深方势阱

对于一维无限深方势阱

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ \infty, & \text{otherwise,} \end{cases}$$

已知粒子初始时刻的波函数为

$$\Psi(x) = \sqrt{\frac{8}{5a}} \left(1 + \cos \frac{\pi x}{a} \right) \sin \frac{\pi x}{a}, \quad 0 < x < a.$$

- (1) 计算在任意时刻 t 粒子的波函数。
- (2) 在任意时刻 t 测量粒子的能量，列出所有可能得到的结果，以及对应的概率。
- (3) 计算在任意时刻 t 粒子出现在 $0 < x < a/2$ 的概率。

求解：根据(1.20)可得，该势阱的能级为

$$\Psi_n(x) = \begin{cases} \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}, & 0 < x < a, \\ 0, & \text{otherwise.} \end{cases}$$

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}.$$

而

$$\Psi(x) = \sqrt{\frac{8}{5a}} \sin \frac{\pi x}{a} + \sqrt{\frac{2}{5a}} \sin \frac{2\pi x}{a},$$

根据(1.14)可得

$$c_1 = \frac{2}{\sqrt{5}},$$

$$c_2 = \frac{1}{\sqrt{5}}.$$

根据(1.16)可得，任意时刻 t 的波函数 $\Psi(x, t)$ 为

$$\Psi(x, t) = c_1 \Psi_1 e^{-iE_1 t/\hbar} + c_2 \Psi_2 e^{-iE_2 t/\hbar} = \sqrt{\frac{8}{5a}} \sin \frac{\pi x}{a} e^{-\pi^2 \hbar i/2ma^2} + \sqrt{\frac{2}{5a}} \sin \frac{2\pi x}{a} e^{-2\pi^2 \hbar i/ma^2}.$$

因此，测量粒子的能量，所有可能的结果为 E_1, E_2 ，根据(2.10)，可得对应的概率为

$$\Pr \{E_1\} = |c_1|^2 = \frac{4}{5},$$

$$\Pr \{E_2\} = |c_2|^2 = \frac{1}{5}.$$

粒子出现在 $0 < x < a/2$ 的概率为

$$\begin{aligned} \int_0^{a/2} |\Psi(x, t)|^2 dx &= \int_0^{a/2} \bar{\Psi} \Psi dx \\ &= \int_0^{a/2} \left(\bar{c}_1 \bar{\Psi}_1 e^{iE_1 t/\hbar} + \bar{c}_2 \bar{\Psi}_2 e^{iE_2 t/\hbar} \right) \left(c_1 \Psi_1 e^{-iE_1 t/\hbar} + c_2 \Psi_2 e^{-iE_2 t/\hbar} \right) dx \\ &= \int_0^{a/2} |c_1|^2 |\Psi_1|^2 + |c_2|^2 |\Psi_2|^2 + \bar{c}_1 c_2 \bar{\Psi}_1 \Psi_2 e^{i(E_1 - E_2)t/\hbar} \\ &\quad + c_1 \bar{c}_2 \Psi_1 \bar{\Psi}_2 e^{-i(E_1 - E_2)t/\hbar} dx \\ &= \frac{1}{2} + \frac{16}{15\pi} \cos \frac{(E_1 - E_2)t}{\hbar}. \end{aligned}$$

5.2 一维无限深方势阱

对于一维无限深方势阱

$$V(x) = \begin{cases} 0, & 0 < x < a, \\ \infty, & \text{otherwise,} \end{cases}$$

已知粒子初始时刻的波函数处于基态, 即

$$\Psi(x) = \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a}, \quad 0 < x < a. \quad (5.1)$$

假设势阱宽度突然变为原来的2倍, 换句话说, 势能函数变为

$$\tilde{V}(x) = \begin{cases} 0, & 0 < x < 2a, \\ \infty, & \text{otherwise,} \end{cases}$$

在势阱宽度变化后的瞬间, 假设粒子的波函数依然为(5.1)。此时, 测量粒子的能量, 列出所有可能得到的结果, 以及对应的概率。最后, 平均值是多少?

求解: 根据(1.20)可得, 在势阱宽度变化后的瞬间, 能级为

$$\begin{aligned} \tilde{\Psi}_n(x) &= \sqrt{\frac{1}{a}} \sin \frac{n\pi x}{2a}, & 0 < x < 2a, \\ \tilde{E}_n &= \frac{n^2 \pi^2 \hbar^2}{8ma^2}. \end{aligned}$$

根据(1.14), 我们需要将波函数(5.1)表示成以下形式:

$$\Psi(x) = \sum_{n=1}^{\infty} \tilde{c}_n \tilde{\Psi}_n(x),$$

根据(1.15)可得

$$\begin{aligned} c_n &= \langle \tilde{\Psi}_n(x), \Psi(x) \rangle \\ &= \int_{-\infty}^{\infty} \overline{\tilde{\Psi}_n(x)} \Psi(x) dx \\ &= \frac{\sqrt{2}}{a} \int_0^a \sin \frac{\pi x}{a} \sin \frac{n\pi x}{2a} dx \\ &= \frac{\sqrt{2}}{\pi} \int_0^{\pi} \sin t \sin \frac{nt}{2} dt \\ &= \begin{cases} (-1)^{\frac{n+1}{2}} \frac{4\sqrt{2}}{(n^2-4)\pi}, & n = 1, 3, 5, \dots, \\ \frac{\sqrt{2}}{2}, & n = 2, \\ 0, & n = 4, 6, 8, \dots, \end{cases} \end{aligned}$$

根据(2.10)可得, 测量得到能量 $\tilde{E}_n = \frac{n^2\pi^2\hbar^2}{8ma^2}$ 的概率为

$$\Pr\{\tilde{E}_n\} = \begin{cases} \frac{32}{(n^2-4)^2\pi^2}, & n = 1, 3, 5, \dots, \\ \frac{1}{2}, & n = 2, \\ 0, & n = 4, 6, 8, \dots, \end{cases}$$

根据(2.11), 平均值为

$$\begin{aligned} \sum_{n=1}^{\infty} E_n |c_n|^2 &= \frac{1}{2} \frac{\pi^2 \hbar^2}{2ma^2} + \sum_{k=1}^{\infty} \frac{32}{((2k-1)^2-4)^2\pi^2} \frac{(2k-1)^2\pi^2\hbar^2}{8ma^2} \\ &= \frac{1}{2} \frac{\pi^2 \hbar^2}{2ma^2} + \frac{4\hbar^2}{ma^2} \sum_{k=1}^{\infty} \frac{(2k-1)^2}{((2k-1)^2-4)^2} \\ &= \frac{1}{2} \frac{\pi^2 \hbar^2}{2ma^2} + \frac{\hbar^2}{ma^2} \sum_{k=1}^{\infty} \frac{1}{(2k-3)^2} + \frac{1}{(2k+1)^2} + \frac{2}{(2k-3)(2k+1)} = \frac{\pi^2 \hbar^2}{2ma^2}. \end{aligned}$$

具体计算过程下。因为

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots = \frac{\pi^2}{6},$$

因此

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{(2k-3)^2} &= 1 + 1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots \\ &= 1 + \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} \\ &= 1 + \frac{\pi^2}{6} - \frac{\pi^2}{24} = 1 + \frac{\pi^2}{8}, \\ \sum_{k=1}^{\infty} \frac{1}{(2k+1)^2} &= \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots \\ &= \sum_{k=1}^{\infty} \frac{1}{(2k-3)^2} - 2 = \frac{\pi^2}{8} - 1, \\ \sum_{k=1}^{\infty} \frac{2}{(2k-3)(2k+1)} &= \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{2k-3} - \frac{1}{2k+1} \\ &= \frac{1}{2} \left(\frac{1}{-1} - \frac{1}{3} + \frac{1}{1} - \frac{1}{5} + \frac{1}{3} - \frac{1}{7} + \dots \right) \\ &= \frac{1}{2} \left(\frac{1}{-1} + \frac{1}{1} \right) = 0. \end{aligned}$$

5.3 谐振子

一个粒子在势能为 $V(x) = \frac{1}{2}m\omega^2 x^2$ 的势场中运动，其初始时刻的波函数为

$$\Psi(x) = \cos \frac{\theta}{2} \Psi_0(x) + \sin \frac{\theta}{2} \Psi_1(x),$$

其中 Ψ_0, Ψ_1 分别为谐振子的基态与第一激发态。

- (1) 计算任意时刻 t 的波函数 $\Psi(x, t)$ 。
- (2) 证明 $\Psi(x, t + \frac{2\pi}{\omega}) = -\Psi(x, t)$ 。
- (3) 在任意时刻 t 测量粒子的能量，会得到什么结果，并且相应的概率是多少？最后，平均值是多少？

求解： 根据(1.24)可得， Ψ_0, Ψ_1 对应的能量分别为

$$\begin{aligned} E_0 &= \left(0 + \frac{1}{2}\right) \hbar\omega = \frac{1}{2}\hbar\omega, \\ E_1 &= \left(1 + \frac{1}{2}\right) \hbar\omega = \frac{3}{2}\hbar\omega. \end{aligned}$$

根据(1.16)可得，任意时刻 t 的波函数 $\Psi(x, t)$ 为

$$\Psi(x, t) = \cos \frac{\theta}{2} \Psi_0(x) e^{-iE_0 t/\hbar} + \sin \frac{\theta}{2} \Psi_1(x) e^{-iE_1 t/\hbar} = \cos \frac{\theta}{2} \Psi_0(x) e^{-i\omega t/2} + \sin \frac{\theta}{2} \Psi_1(x) e^{-3i\omega t/2}.$$

因此

$$\Psi\left(x, t + \frac{2\pi}{\omega}\right) = \cos \frac{\theta}{2} \Psi_0(x) e^{-i\omega t/2} e^{-\pi i} + \sin \frac{\theta}{2} \Psi_1(x) e^{-3i\omega t/2} e^{-3\pi i} = -\Psi(x, t).$$

因此，根据(2.10)，得到结果为 $\frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega$ 的概率分别为

$$\begin{aligned} \Pr\left\{\frac{1}{2}\hbar\right\} &= \left|\cos \frac{\theta}{2} e^{-iE_0 t/\hbar}\right|^2 = \cos^2 \frac{\theta}{2}, \\ \Pr\left\{\frac{3}{2}\hbar\right\} &= \left|\sin \frac{\theta}{2} e^{-iE_1 t/\hbar}\right|^2 = \sin^2 \frac{\theta}{2}. \end{aligned}$$

根据(2.11)，平均值为

$$\left(\frac{1}{2}\hbar\right) \Pr\left\{\frac{1}{2}\hbar\right\} + \left(\frac{3}{2}\hbar\right) \Pr\left\{\frac{3}{2}\hbar\right\} = \left(1 - \frac{1}{2} \cos \theta\right) \hbar\omega.$$

5.4 谐振子

一个粒子在势能为 $V(x) = \frac{1}{2}m\omega^2 x^2 + Cx$ 的势场中运动, 其中 $C \neq 0$ 是一个常数, 请计算粒子能量的本征值。

求解: 此时定态薛定谔方程为

$$-\frac{\hbar^2}{2m}\Psi''(x) + \left(\frac{1}{2}m\omega^2 x^2 + Cx\right)\Psi(x) = E\Psi(x),$$

记 $\{\Psi_n(x), E_n\}_{n=0,1,2,\dots}$ 为满足该方程的一组解, 即对于 $n = 0, 1, 2, \dots$, 都具有

$$-\frac{\hbar^2}{2m}\Psi_n''(x) + \left(\frac{1}{2}m\omega^2 x^2 + Cx\right)\Psi_n(x) = E_n\Psi_n(x),$$

令 $\tilde{x} = x + \frac{C}{m\omega^2}$, $\Phi_n(x) = \Psi_n\left(x - \frac{C}{m\omega^2}\right)$, 可得

$$-\frac{\hbar^2}{2m}\Phi_n''(\tilde{x}) + \frac{1}{2}m\omega^2 \tilde{x}^2 \Phi_n(\tilde{x}) = \left(E_n + \frac{C^2}{2m\omega^2}\right)\Phi_n(\tilde{x}),$$

根据(1.24)可得

$$\begin{aligned} E_n + \frac{C^2}{2m\omega^2} &= \left(n + \frac{1}{2}\right)\hbar\omega, \\ E_n &= \left(n + \frac{1}{2}\right)\hbar\omega - \frac{C^2}{2m\omega^2}. \end{aligned}$$

5.5 谐振子

对于定态薛定谔方程, 如果存在2个本征态, 它们对应的能量是相同的, 则称这2个本征态**简并**。给定能量 E , 如果最多只能找出 n 个本征态 Ψ_1, \dots, Ψ_n 使得它们对应的能量为 E , 则称能量 E 的**简并度**为 n 。假设粒子在2维势场 $V(x, y) = \frac{1}{2}m\omega^2(4x^2 + y^2)$ 中运动, 对于任意正整数 N , 能量 $(N + \frac{1}{2})\hbar\omega$ 的简并度为多少?

求解: 此时定态薛定谔方程为

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2}\right) + V(x, y)\Psi(x, y) = E\Psi(x, y).$$

令 $\Psi(x, y) = F(x)G(y)$, 可得

$$-\frac{\hbar^2}{2m}(FG'' + F''G) + VFG = EFG,$$

两边除以 $F(x)G(y)$ 可得

$$-\frac{\hbar^2}{2m}\left(\frac{F''}{F} + \frac{G''}{G}\right) + V = E,$$

代入势能函数可得

$$\left(-\frac{\hbar^2}{2m} \frac{F''}{F} + \frac{1}{2}m(2\omega)^2 x^2\right) + \left(-\frac{\hbar^2}{2m} \frac{G''}{G} + \frac{1}{2}m\omega^2 y^2\right) = E,$$

一个自变量为 x 与一个自变量为 y 的函数之和是一个常数，这意味着这2个函数取值只能为常数。记它们为 $C, D > 0$ (如果小于等于0, 那么方程无解), 即

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{F''}{F} + \frac{1}{2}m(2\omega)^2 x^2 &= C, \\ -\frac{\hbar^2}{2m} \frac{G''}{G} + \frac{1}{2}m\omega^2 y^2 &= D. \end{aligned}$$

其中 $C + D = E$ 。根据(1.24)可得, 这2个方程的解为

$$\begin{aligned} C_m &= \left(m + \frac{1}{2}\right) \hbar(2\omega), \\ D_n &= \left(n + \frac{1}{2}\right) \hbar\omega, \end{aligned}$$

其中 $m, n = 0, 1, 2, \dots$ 。当 $E = (N + \frac{1}{2}) \hbar\omega$ 时

$$\begin{aligned} 2\left(m + \frac{1}{2}\right) + \left(n + \frac{1}{2}\right) &= N + \frac{1}{2}, \\ 2m + n &= N - 1. \end{aligned}$$

$N = 0$ 显然无解。令 $N \geq 1$, 如Figure 5.1所示, 可以发现:

(A) 当 N 是奇数时, $2m + n$ 是偶数, 对应图像中的红线, 可以发现红线经过 $\frac{N+1}{2}$ 个格点。

(B) 当 N 是偶数时, $2m + n$ 是奇数, 对应图像中的蓝线, 可以发现红线经过 $\frac{N}{2}$ 个格点。

因此, 能级的简并度为

$$\left\lceil \frac{N}{2} \right\rceil,$$

其中该符号表示向上取整, 例如 $\lceil 0.9 \rceil = 1$ 。

5.6 动量表象、线性算子的期望值、测不准原理

一个粒子在初始时刻的波函数为

$$\Psi(x) = \begin{cases} 2C^{3/2}xe^{-Cx}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

其中 $C > 0$ 是一个常数。

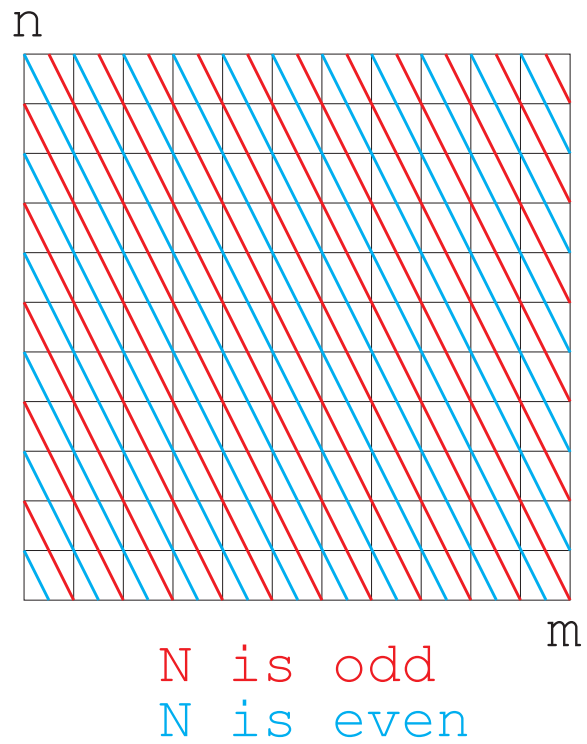


Figure 5.1: 简并度的计算

- (1) 计算粒子动量表象的波函数 $\Phi(p)$ 。
- (2) 计算粒子位置的平均值 $\langle X \rangle$ 。
- (3) 计算粒子动量的平均值 $\langle P \rangle$ 。
- (4) 计算粒子位置的不确定度 ΔX 。
- (5) 计算粒子动量的不确定度 ΔP 。
- (6) 验证测不准原理 $\Delta X \Delta P \geq \frac{\hbar}{2}$ 。

求解： 利用(1.7)的一维形式：

$$\begin{aligned}\Phi(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x) dx \\ &= \sqrt{\frac{2C^3}{\pi\hbar}} \int_0^{\infty} x e^{-(ip/\hbar + C)x} dx \\ &= \sqrt{\frac{2C^3}{\pi\hbar}} \frac{1}{(C + ip/\hbar)^2}.\end{aligned}$$

根据(2.15)可得

$$\begin{aligned}\langle X \rangle &= \int_{-\infty}^{\infty} x |\Psi(x)|^2 dx \\ &= 4C^3 \int_0^{\infty} x^3 e^{-2Cx} dx \\ &= \frac{3}{2C}.\end{aligned}$$

根据(2.16)可得

$$\begin{aligned}\langle P \rangle &= \int_{-\infty}^{\infty} p |\Phi(p)|^2 dp \\ &= \frac{2C^3}{\pi\hbar} \int_{-\infty}^{\infty} \frac{p}{(C^2 + p^2/\hbar^2)^2} dp \\ &= 0.\end{aligned}$$

根据(2.15)可得

$$\begin{aligned}\langle X^2 \rangle &= \int_{-\infty}^{\infty} x^2 |\Psi(x)|^2 dx \\ &= 4C^3 \int_0^{\infty} x^4 e^{-2Cx} dx \\ &= \frac{3}{C^2}.\end{aligned}$$

根据(2.16)可得

$$\begin{aligned}\langle P^2 \rangle &= \int_{-\infty}^{\infty} p^2 |\Phi(p)|^2 dp \\ &= \frac{2C^3}{\pi\hbar} \int_{-\infty}^{\infty} \frac{p^2}{(C^2 + p^2/\hbar^2)^2} dp \\ &= C^2 \hbar^2.\end{aligned}$$

根据(2.17)可得

$$\Delta X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} = \frac{\sqrt{3}}{2C},$$

$$\Delta P = \sqrt{\langle P^2 \rangle - \langle P \rangle^2} = C\hbar,$$

因此

$$\Delta X \Delta P = \frac{\sqrt{3}}{2} \hbar.$$

5.7 狄拉克符号

2个量子态 $|\Psi_1\rangle$ 与 $|\Psi_2\rangle$ 的保真度定义为 $F = |\langle \Psi_1 | \Psi_2 \rangle|^2$, 计算

$$|\Psi_1\rangle = \cos \frac{\theta_1}{2} |0\rangle - \sin \frac{\theta_1}{2} |1\rangle,$$

$$|\Psi_2\rangle = \cos \frac{\theta_2}{2} |0\rangle + i \sin \frac{\theta_2}{2} |1\rangle,$$

的保真度。

求解: 根据(2.45)可得

$$\langle \Psi_1 | = \cos \frac{\theta_1}{2} \langle 0 | - \sin \frac{\theta_1}{2} \langle 1 |,$$

根据(2.47)可得

$$\begin{aligned} \langle \Psi_1 | \Psi_2 \rangle &= \left(\cos \frac{\theta_1}{2} \langle 0 | - \sin \frac{\theta_1}{2} \langle 1 | \right) \left(\cos \frac{\theta_2}{2} |0\rangle + i \sin \frac{\theta_2}{2} |1\rangle \right) \\ &= \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \langle 0|0\rangle + i \cos \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \langle 0|1\rangle - \sin \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \langle 1|0\rangle - i \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \langle 1|1\rangle \\ &= \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} - i \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2}. \end{aligned}$$

因此

$$F = |\langle \Psi_1 | \Psi_2 \rangle|^2 = \cos^2 \frac{\theta_1}{2} \cos^2 \frac{\theta_2}{2} + \sin^2 \frac{\theta_1}{2} \sin^2 \frac{\theta_2}{2}.$$

5.8 氢原子能级

已知氢原子处于能级 $n = 2$, 写出所有可能的 (n, l, m) 。

求解: 根据(3.3)可得

$$l = 0, 1, \dots, n - 1,$$

所以 $l = 0, 1$ 。而

$$m = -l, -l + 1, \dots, l - 1, l,$$

所以当 $l = 0$ 时 $m = 0$, 当 $l = 1$ 时 $m = -1, 0, 1$ 。

综上所述, 所有可能的 (n, l, m) 为 $(2, 0, 0), (2, 1, -1), (2, 1, 0), (2, 1, 1)$ 。

5.9 氢原子波函数

已知氢原子初始时刻的波函数为

$$\Psi(\mathbf{x}, 0) = \frac{1}{\sqrt{10}} \left(2\Psi_{1,0,0} + \Psi_{2,1,0} + \sqrt{2}\Psi_{2,1,1} + \sqrt{3}\Psi_{2,1,-1} \right).$$

- (1) 计算在任意时刻 t 氢原子的波函数。
- (2) 在任意时刻 t 测量氢原子的角动量, 列出所有可能得到的结果, 以及对应的概率。
- (3) 在任意时刻 t 测量氢原子的角动量 z 分量, 列出所有可能得到的结果, 以及对应的概率。
- (4) 在任意时刻 t 测量氢原子的能量, 列出所有可能得到的结果, 以及对应的概率。

求解: 根据(1.16)与(3.4)可得

$$\Psi(\mathbf{x}, t) = \frac{1}{\sqrt{10}} \left(2\Psi_{1,0,0}e^{iE_0t/\hbar} + \Psi_{2,1,0}e^{iE_0t/4\hbar} + \sqrt{2}\Psi_{2,1,1}e^{iE_0t/4\hbar} + \sqrt{3}\Psi_{2,1,-1}e^{iE_0t/4\hbar} \right).$$

根据(3.4), $\Psi_{1,0,0}$ 对应 L^2 本征值为 $0(0+1)\hbar^2 = 0$, $\Psi_{2,1,0}, \Psi_{2,1,1}, \Psi_{2,1,-1}$ 对应 L^2 本征值为 $1(1+1)\hbar^2 = 2\hbar^2$, 所以根据(2.10), 得到结果为 $0, 2\hbar^2$ 的概率分别为

$$\begin{aligned} \Pr\{0\} &= \left| \frac{2}{\sqrt{10}}e^{iE_0t/\hbar} \right|^2 = \frac{2}{5}, \\ \Pr\{2\hbar^2\} &= \left| \frac{1}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 + \left| \frac{\sqrt{2}}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 + \left| \frac{\sqrt{3}}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 = \frac{3}{5}. \end{aligned}$$

根据(3.4), $\Psi_{1,0,0}, \Psi_{2,1,0}$ 对应 L_z 本征值为 $0\hbar = 0$, $\Psi_{2,1,1}$ 对应 L_z 本征值为 $1\hbar = \hbar$, $\Psi_{2,1,-1}$ 对应 L_z 本征值为 $-1\hbar = -\hbar$, 所以根据(2.10), 得到结果为 $0, \hbar, -\hbar$ 的概率分别为

$$\begin{aligned} \Pr\{0\} &= \left| \frac{2}{\sqrt{10}}e^{iE_0t/\hbar} \right|^2 + \left| \frac{1}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 = \frac{1}{2}, \\ \Pr\{\hbar\} &= \left| \frac{\sqrt{2}}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 = \frac{1}{5}, \\ \Pr\{-\hbar\} &= \left| \frac{\sqrt{3}}{\sqrt{10}}e^{iE_0t/4\hbar} \right|^2 = \frac{3}{10}. \end{aligned}$$

根据(3.4), $\Psi_{1,0,0}$ 对应 H 本征值为 $-E_0$, $\Psi_{2,1,0}, \Psi_{2,1,1}, \Psi_{2,1,-1}$ 对应 L^2 本征值为 $-E_0/2^2 = -E_0/4$, 所以根据(2.10), 得到结果为 $-E_0, -E_0/4$ 的概率分别为

$$\begin{aligned}\Pr\{-E_0\} &= \left| \frac{2}{\sqrt{10}} e^{iE_0 t/\hbar} \right|^2 = \frac{2}{5}, \\ \Pr\left\{-\frac{E_0}{4}\right\} &= \left| \frac{1}{\sqrt{10}} e^{iE_0 t/4\hbar} \right|^2 + \left| \frac{\sqrt{2}}{\sqrt{10}} e^{iE_0 t/4\hbar} \right|^2 + \left| \frac{\sqrt{3}}{\sqrt{10}} e^{iE_0 t/4\hbar} \right|^2 = \frac{3}{5}.\end{aligned}$$

5.10 氢原子跃迁

某一个类氢离子的光谱, 已知同一线系(从任意能级 $n > n_0$ 跃迁到固定的能级 n_0)得到的其中3条谱线波长为99.2 nm, 108.5 nm, 121.5 nm。那么类氢离子的电荷数 Z , 以及对应固定的能级 n_0 是多少?

求解: 记3条谱线的波长分别为 $\lambda_1, \lambda_2, \lambda_3$ (从小到大排序), 分别对应能级 $n_1, n_2, n_3 > n_0$ 跃迁到能级 n_0 , 其中 $n_1 < n_2 < n_3$ 。因此, 根据(3.5), 谱线波长 λ 满足

$$\begin{aligned}\frac{1}{\lambda_1} &= Z^2 R \left(\frac{1}{n_0^2} - \frac{1}{n_1^2} \right), \\ \frac{1}{\lambda_2} &= Z^2 R \left(\frac{1}{n_0^2} - \frac{1}{n_2^2} \right), \\ \frac{1}{\lambda_3} &= Z^2 R \left(\frac{1}{n_0^2} - \frac{1}{n_3^2} \right).\end{aligned}$$

我们将它进行恒等变换可得

$$\begin{aligned}\frac{1}{\lambda_1} &= R \left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_1/Z)^2} \right), \\ \frac{1}{\lambda_2} &= R \left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_2/Z)^2} \right), \\ \frac{1}{\lambda_3} &= R \left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_3/Z)^2} \right).\end{aligned}$$

里德伯常数为 $R = 1.10 \times 10^7 \text{ m}^{-1}$, 代入数据可得

$$\begin{aligned}\left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_1/Z)^2} \right) &= 0.750, \\ \left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_2/Z)^2} \right) &= 0.840, \\ \left(\frac{1}{(n_0/Z)^2} - \frac{1}{(n_3/Z)^2} \right) &= 0.919,\end{aligned}$$

根据第1行的表达式, 可以猜测

$$\frac{n_0}{Z} = 1,$$

$$\frac{n_1}{Z} = 2,$$

代入第2, 3行的表达式可得

$$\frac{n_2}{Z} = 2.5,$$

$$\frac{n_3}{Z} = 3.5.$$

根据 n_0, n_1, n_2, n_3 均为正整数的性质, 可得

$$Z = 2,$$

$$n_0 = 2.$$

5.11 泡利矩阵

令 $\mathbf{A} = (A^1, A^2, A^3), \mathbf{B} = (B^1, B^2, B^3) \in \mathbb{R}^3$, 证明等式

$$(\mathbf{A} \cdot \boldsymbol{\sigma})(\mathbf{B} \cdot \boldsymbol{\sigma}) = (\mathbf{A} \cdot \mathbf{B})I + i(\mathbf{A} \times \mathbf{B}) \cdot \boldsymbol{\sigma},$$

其中 I 是单位矩阵, $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, $\sigma_1, \sigma_2, \sigma_3$ 是泡利矩阵, 如3.10所示。记向量 $\mathbf{N} = (N^1, N^2, N^3)$ 与 $\boldsymbol{\sigma}$ 的内积为 $N^1\sigma_1 + N^2\sigma_2 + N^3\sigma_3$ 。

求解: 利用 $\mathbf{A} \times \mathbf{B} = \sum_{i,j,k=1}^3 (\epsilon_{ijk} A^i B^j) e_k$ (可以自行验证它正确与否), 其中 $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$, ϵ_{ijk} 的取值如Figure 3.1所示。因此

$$(\mathbf{A} \cdot \boldsymbol{\sigma})(\mathbf{B} \cdot \boldsymbol{\sigma}) = \sum_{i,j=1}^3 A^i B^j \sigma_i \sigma_j,$$

$$(\mathbf{A} \cdot \mathbf{B})I + i(\mathbf{A} \times \mathbf{B}) \cdot \boldsymbol{\sigma} = \sum_{i,j=1}^3 \left((\delta_{ij} A^i B^j)I + i \sum_{k=1}^3 \epsilon_{ijk} A^i B^j \sigma_k \right).$$

根据(3.12)与(3.13)可得

$$(\mathbf{A} \cdot \boldsymbol{\sigma})(\mathbf{B} \cdot \boldsymbol{\sigma}) - (\mathbf{A} \cdot \mathbf{B})I - i(\mathbf{A} \times \mathbf{B}) \cdot \boldsymbol{\sigma} = \sum_{i,j=1}^3 A^i B^j \left(\sigma_i \sigma_j - \delta_{ij} I - i \sum_{k=1}^3 \epsilon_{ijk} \sigma_k \right)$$

$$= \sum_{i,j=1}^3 A^i B^j \left(\sigma_i \sigma_j - \frac{1}{2} \{ \sigma_i, \sigma_j \} - \frac{1}{2} [\sigma_i, \sigma_j] \right)$$

$$= 0.$$

5.12 自旋的测量

根据(3.15), 假设粒子的波函数可以表示为向量

$$\Psi \longleftrightarrow \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

分别测量粒子自旋在 x, y, z 方向的分量, 会得到什么结果, 并且相应的概率是多少? 最后, 平均值是多少?

求解:

(A) 测量粒子自旋在 x 方向的分量。根据(3.18)可得, 我们可以将粒子的波函数表示为

$$\Psi = \frac{1}{\sqrt{2}} (\cos \theta + \sin \theta) \chi_0(1) + \frac{1}{\sqrt{2}} (\cos \theta - \sin \theta) \chi_1(1),$$

因此, 根据(2.10), 得到结果为 $\hbar/2, -\hbar/2$ 的概率分别为

$$\begin{aligned} \Pr \left\{ \frac{\hbar}{2} \right\} &= \left| \frac{1}{\sqrt{2}} (\cos \theta + \sin \theta) \right|^2 = \frac{1}{2} (1 + \sin 2\theta), \\ \Pr \left\{ -\frac{\hbar}{2} \right\} &= \left| \frac{1}{\sqrt{2}} (\cos \theta - \sin \theta) \right|^2 = \frac{1}{2} (1 - \sin 2\theta). \end{aligned}$$

根据(2.11), 平均值为

$$\left(\frac{\hbar}{2} \right) \Pr \left\{ \frac{\hbar}{2} \right\} + \left(-\frac{\hbar}{2} \right) \Pr \left\{ -\frac{\hbar}{2} \right\} = \frac{\hbar}{2} \sin 2\theta.$$

(B) 测量粒子自旋在 y 方向的分量。根据(3.18)可得, 我们可以将粒子的波函数表示为

$$\Psi = \frac{1}{\sqrt{2}} e^{-i\theta} \chi_0(2) + \frac{1}{\sqrt{2}} e^{i\theta} \chi_1(2),$$

因此, 根据(2.10), 得到结果为 $\hbar/2, -\hbar/2$ 的概率分别为

$$\begin{aligned} \Pr \left\{ \frac{\hbar}{2} \right\} &= \left| \frac{1}{\sqrt{2}} e^{-i\theta} \right|^2 = \frac{1}{2}, \\ \Pr \left\{ -\frac{\hbar}{2} \right\} &= \left| \frac{1}{\sqrt{2}} e^{i\theta} \right|^2 = \frac{1}{2}. \end{aligned}$$

根据(2.11), 平均值为

$$\left(\frac{\hbar}{2} \right) \Pr \left\{ \frac{\hbar}{2} \right\} + \left(-\frac{\hbar}{2} \right) \Pr \left\{ -\frac{\hbar}{2} \right\} = \frac{\hbar}{2}.$$

(C) 测量粒子自旋在 z 方向的分量。根据(3.18)可得, 我们可以将粒子的波函数表示为

$$\Psi = \cos \theta \chi_0(3) + \sin \theta \chi_1(3),$$

因此, 根据(2.10), 得到结果为 $\hbar/2, -\hbar/2$ 的概率分别为

$$\begin{aligned} \Pr \left\{ \frac{\hbar}{2} \right\} &= |\cos \theta|^2 = \cos^2 \theta, \\ \Pr \left\{ -\frac{\hbar}{2} \right\} &= |\sin \theta|^2 = \sin^2 \theta. \end{aligned}$$

根据(2.11), 平均值为

$$\left(\frac{\hbar}{2} \right) \Pr \left\{ \frac{\hbar}{2} \right\} + \left(-\frac{\hbar}{2} \right) \Pr \left\{ -\frac{\hbar}{2} \right\} = \frac{\hbar}{2} \cos 2\theta.$$

5.13 自旋体系的能量

根据(3.16), 假设哈密顿算子的矩阵表示为

$$H \longleftrightarrow \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}.$$

其中 $E_0, A > 0$ 是常数。

(1) 计算哈密顿算子的本征值与本征态的向量表示。

(2) 假设粒子的波函数可以表示为向量

$$\Psi \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

测量粒子的能量, 会得到什么结果, 并且相应的概率是多少? 最后, 平均值是多少?

求解: 假设本征态的向量表示为

$$\begin{pmatrix} c_0 \\ c_1 \end{pmatrix},$$

其中 $|c_0|^2 + |c_1|^2 = 1$ 。根据(3.17)可得, 我们需要求解方程

$$\begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \lambda \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}.$$

该方程具有2个解，分别为

$$\begin{aligned} c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= \frac{1}{\sqrt{2}}, & \lambda &= E_0 - A, \\ c_0 &= \frac{1}{\sqrt{2}}, & c_1 &= -\frac{1}{\sqrt{2}}, & \lambda &= E_0 + A. \end{aligned}$$

因此，本征值分别为 $E_0 - A, E_0 + A$ ，对应的本征向量表示分别为

$$\begin{aligned} \Psi_0 &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \Psi_1 &\longleftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

而

$$\Psi = \frac{1}{\sqrt{2}} \Psi_0 + \frac{1}{\sqrt{2}} \Psi_1,$$

因此，根据(2.10)，得到结果为 $E_0 - A, E_0 + A$ 的概率分别为

$$\begin{aligned} \Pr\{E_0 - A\} &= \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}, \\ \Pr\{E_0 + A\} &= \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}. \end{aligned}$$

根据(2.11)，平均值为

$$(E_0 - A)\Pr\{E_0 - A\} + (E_0 + A)\Pr\{E_0 + A\} = E_0.$$

5.14 自旋动力学

假设电子在磁感应强度为 $\mathbf{B} = (B_0, 0, 0)$ 的磁场中运动，其哈密顿算子为

$$H = \frac{e}{m} \mathbf{B} \cdot \mathbf{S} = \frac{e}{m} B_0 S_1,$$

其中 $\mathbf{S} = (S_1, S_2, S_3)$ ， S_1, S_2, S_3 为自旋算子。记电子初始时刻波函数的向量表示为

$$\Psi(0) = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix},$$

其中 $|c_0|^2 + |c_1|^2 = 1$ 。

(1) 计算电子在任意时刻 t 波函数的向量表示。

(2) 在任意时刻 t 测量电子自旋在 z 方向的分量, 会得到什么结果, 并且相应的概率是多少? 最后, 平均值是多少?

求解1: 很显然, $\chi_0(1), \chi_1(1)$ 是哈密顿算子的本征态(因为它正比于 S_1)。因此, 我们将电子初始时刻的波函数表示为本征态的线性组合

$$\Psi(0) = \frac{1}{\sqrt{2}}(c_0 + c_1)\chi_0(1) + \frac{1}{\sqrt{2}}(c_0 - c_1)\chi_1(1).$$

而

$$\begin{aligned} H\chi_0(1) &= \frac{e}{m}B_0S_1\chi_0(1) = \frac{e\hbar B_0}{2m}\chi_0(1), \\ H\chi_1(1) &= \frac{e}{m}B_0S_1\chi_1(1) = -\frac{e\hbar B_0}{2m}\chi_1(1). \end{aligned}$$

因此, 对应的本征值(能量)分别为

$$\begin{aligned} E_0 &= \frac{e\hbar B_0}{2m}, \\ E_1 &= -\frac{e\hbar B_0}{2m}. \end{aligned}$$

因此, 根据(1.16)可得, 任意时刻 t 的波函数 $\Psi(t)$ 为

$$\begin{aligned} \Psi(t) &= \frac{1}{\sqrt{2}}(c_0 + c_1)e^{-iE_0t/\hbar}\chi_0(1) + \frac{1}{\sqrt{2}}(c_0 - c_1)e^{-iE_1t/\hbar}\chi_1(1) \\ &= \frac{1}{\sqrt{2}}(c_0 + c_1)e^{-ieB_0t/2m}\chi_0(1) + \frac{1}{\sqrt{2}}(c_0 - c_1)e^{ieB_0t/2m}\chi_1(1) \\ &\leftrightarrow \frac{1}{2}(c_0 + c_1)e^{-ieB_0t/2m} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}(c_0 - c_1)e^{ieB_0t/2m} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} c_0 \cos \frac{eB_0t}{2m} - ic_1 \sin \frac{eB_0t}{2m} \\ c_1 \cos \frac{eB_0t}{2m} - ic_0 \sin \frac{eB_0t}{2m} \end{pmatrix}. \end{aligned}$$

因此

$$\Psi(t) = \left(c_0 \cos \frac{eB_0t}{2m} - ic_1 \sin \frac{eB_0t}{2m} \right) \chi_0(3) + \left(c_1 \cos \frac{eB_0t}{2m} - ic_0 \sin \frac{eB_0t}{2m} \right) \chi_1(3).$$

在任意时刻 t 测量电子自旋在 z 方向, 得到 $\hbar/2, -\hbar/2$ 的概率分别为

$$\begin{aligned} \Pr \left\{ \frac{\hbar}{2} \right\} &= \left| c_0 \cos \frac{eB_0t}{2m} - ic_1 \sin \frac{eB_0t}{2m} \right|^2 = |c_0|^2 \cos^2 \frac{eB_0t}{2m} + |c_1|^2 \sin^2 \frac{eB_0t}{2m}, \\ \Pr \left\{ -\frac{\hbar}{2} \right\} &= \left| c_1 \cos \frac{eB_0t}{2m} - ic_0 \sin \frac{eB_0t}{2m} \right|^2 = |c_0|^2 \sin^2 \frac{eB_0t}{2m} + |c_1|^2 \cos^2 \frac{eB_0t}{2m}. \end{aligned}$$

很显然 $\Pr\left\{\frac{\hbar}{2}\right\} + \Pr\left\{-\frac{\hbar}{2}\right\} = 1$ 。平均值为

$$\left(\frac{\hbar}{2}\right) \Pr\left\{\frac{\hbar}{2}\right\} + \left(-\frac{\hbar}{2}\right) \Pr\left\{-\frac{\hbar}{2}\right\} = \frac{\hbar}{2} \cos \frac{eB_0 t}{m} (|c_0|^2 - |c_1|^2).$$

求解2: 根据求解1的步骤可得

$$\begin{aligned} H\chi_0(1) &= \frac{e\hbar B_0}{2m} \chi_0(1), \\ H\chi_1(1) &= -\frac{e\hbar B_0}{2m} \chi_1(1). \end{aligned}$$

根据(3.34)可得

$$\begin{aligned} \Psi_0^0 &= \frac{1}{\sqrt{2}}, & \Psi_0^1 &= \frac{1}{\sqrt{2}}, & E_0 &= \frac{e\hbar B_0}{2m}, \\ \Psi_1^0 &= \frac{1}{\sqrt{2}}, & \Psi_1^1 &= -\frac{1}{\sqrt{2}}, & E_1 &= -\frac{e\hbar B_0}{2m}. \end{aligned}$$

因此

$$\begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^{-ieB_0 t/2m} & 0 \\ 0 & e^{ieB_0 t/2m} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} c_0 \cos \frac{eB_0 t}{2m} - ic_1 \sin \frac{eB_0 t}{2m} \\ c_1 \cos \frac{eB_0 t}{2m} - ic_0 \sin \frac{eB_0 t}{2m} \end{pmatrix}.$$

之后步骤与求解1相同。

5.15 2个粒子的自旋态

对于2个粒子A,B, 假设初始时刻A,B的自旋状态分别为

$$\Psi^A \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Psi^B \longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(1) 测量系统的总角动量 $S^2(A, B)$, 会得到什么结果, 并且相应的概率是多少?

(2) 测量系统的总角动量 z 分量 $S_3(A, B)$, 会得到什么结果, 并且相应的概率是多少?

求解: 初始时刻, A,B的自旋态可以表示为

$$\Psi^A \otimes \Psi^B \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

根据(3.35), 可得

$$\Psi^A \otimes \Psi^B = \frac{1}{\sqrt{2}} (\chi_{1,0} + \chi_{0,0}).$$

根据(3.36), $\chi_{1,0}$ 对应 $S^2(A, B)$ 本征值为 $2\hbar^2$, $\chi_{0,0}$ 对应 $S^2(A, B)$ 本征值为0, 所以根据(2.10), 得到结果为 $2\hbar^2, 0$ 的概率分别为

$$\begin{aligned}\Pr\{2\hbar^2\} &= \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}, \\ \Pr\{0\} &= \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}.\end{aligned}$$

根据(3.36), $\chi_{1,0}, \chi_{0,0}$ 对应 $S_3(A, B)$ 本征值均为0, 所以根据(2.10), 得到结果为0的概率为

$$\Pr\{0\} = \left| \frac{1}{\sqrt{2}} \right|^2 + \left| \frac{1}{\sqrt{2}} \right|^2 = 1.$$

5.16 自旋-自旋耦合

假设氦原子核外2个电子处于2s3d组态, 它们在LS耦合下形成的原子态有几种? 用原子态的符号表示出来。

求解: 根据(3.39)可得, 2s组态对应 $l_1 = 0$, 3d组态对应 $l_2 = 2$ 。根据(3.40)可得, 原子的角动量取值为

$$L = l_1 + l_2, l_1 + l_2 - 1, \dots, |l_1 - l_2|,$$

因此 $L = 2$ 。根据(3.41)可得, 原子的自旋取值为 $S = 1, 0$, 根据(3.42)可得, 原子的总角动量取值为

$$J = L + S, L + S - 1, \dots, |L - S|,$$

因此, 根据(3.46), 总共有4个组态: 当 $L = 2, S = 1$ 时, $J = 3, 2, 1$, 对应组态为 $^3D_3, ^3D_2, ^3D_1$; 当 $L = 2, S = 0$ 时, $J = 2$, 对应组态为 1D_2 。

5.17 原子的磁矩

原子态 $^2D_{3/2}$ 对应的 $\mathbf{L} \cdot \mathbf{S}$ 的取值是多少? 计算相应的朗德因子以及实验测量得到的磁矩值。

求解: 根据(3.46)可得

$$\begin{aligned}J &= \frac{3}{2}, \\ S &= \frac{1}{2}, \\ L &= 2.\end{aligned}$$

根据(3.43)可得, $\mathbf{L} \cdot \mathbf{S}$ 的取值为

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} (J(J+1) - L(L+1) - S(S+1)) \hbar^2 = -\frac{3\hbar^2}{2}.$$

根据(3.44)可得

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} = \frac{4}{5}.$$

根据(3.45)可得

$$\mu_J = Jg\mu_B, (J-1)g\mu_B, \dots, (-J+1)g\mu_B, -Jg\mu_B = \frac{6}{5}\mu_B, \frac{2}{5}\mu_B, -\frac{2}{5}\mu_B, -\frac{6}{5}\mu_B.$$

5.18 纠缠态

通过CNOT门可以将4个贝尔态变为非纠缠态，这是为什么？

求解： 根据(4.24)与(4.29)可得

$$\text{CNOT}|\Phi^+\rangle \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

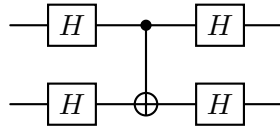
$$\text{CNOT}|\Phi^-\rangle \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\text{CNOT}|\Psi^+\rangle \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

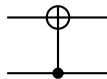
$$\text{CNOT}|\Psi^-\rangle \longleftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

5.19 量子门

证明



与CNOT₁₀门



是等价的。

求解： 根据(4.23)与(4.24)可得

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

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

类比(4.26)，第1个量子线路图等效于

$$(H \otimes H)(\text{CNOT})(H \otimes H),$$

而

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

因此

$$\begin{aligned} (H \otimes H)(\text{CNOT})(H \otimes H) &= \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \end{aligned}$$

根据(4.25)可得

$$\text{CNOT}_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

因此

$$(H \otimes H)(\text{CNOT})(H \otimes H) = \text{CNOT}_{10},$$

是等价的。