

量子力学初步

Phys1501

1.11 版

曲豆豆 整理

2022 年 5 月 10 日



$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle$$

这份笔记整理了曲豆豆与 WZ, WYW, CZL 私下组织的物理学读书会的量子力学部分的讨论内容.

目录

1	数学背景知识	4
1.1	矢量空间与内积空间	4
1.2	线性算子与矩阵	14
1.3	伴随算子, 厄米特算子与酉算子	20
1.4	本征值与本征矢量	26
1.5	指数映射与算子求导	34
1.6	狄拉克 δ -函数	37
1.7	无穷维空间中的线性算子	44
2	经典力学背景知识	50
2.1	力学系统的自由度与广义坐标	50
2.2	最小作用量原理与拉格朗日力学	54
2.3	哈密顿力学	59
2.4	正则变换	63
3	量子力学基本原理	67
3.1	导引: 量子世界中的“豌豆”	67
3.2	量子态与观测测量	73
3.3	本征态, 期望与方差, 不确定原理	82
3.4	量子态随时间的演化: 薛定谔方程	87
3.5	数学补充: 二维复矢量空间的代数与几何	97

1 数学背景知识

物理学家擅长发明有用的记号.

— 沃·兹基·硕德

各位好奇的小朋友们好, 本趟列车开往神奇的量子世界. 在上车之前, 我们需要一些准备, 想必各位小朋友们都了解一些微积分, 线性代数与经典力学.

众所周知, 物理学家擅长发明有用的记号. 在本章, 我们来回忆一些对量子力学有用的必要的数学知识, 并把这些知识用物理学家们习惯的记号来重写一遍, 以便我们提前适应量子力学的说话方式.

1.1 矢量空间与内积空间

量子力学严重依赖线性代数, 尤其是复数域 \mathbb{C} 上的线性代数, 比如**酉空间**, **厄米特矩阵** 等概念; 想必绝大多数小朋友们听说过这些名词, 但当初觉得没用于是没认真学, 或者你们的老师根本没认真讲. 此外, 有些工科的小朋友没学过**线性空间**, **子空间**, **对偶空间** 等基本概念, 这锅笔者不敢乱甩, 背后水挺深. 当然这份笔记不准备给这部分小朋友补课, 而是假定大家都熟悉这些基本概念, 然后我们来一起回忆, 并用量子力学的符号习惯来重新表示它们.

首先我们需要**矢量空间**, 也就是数学上所谓的“线性空间”或者“向量空间”.

矢量空间.

定义. **矢量空间** \mathbb{V} 首先是一个集合, 该集合中的元素常记作 $|1\rangle, |2\rangle, \dots, |V\rangle, \dots, |W\rangle$ 等, 称为**矢量**. 并且 \mathbb{V} 中有加法运算与数乘运算, 分别称为**矢量加法** 与**矢量数乘**, 即

1. 对于任意两个矢量 $|V\rangle, |W\rangle$, 这两个矢量能够相加: $|V\rangle + |W\rangle$ 仍为 \mathbb{V} .
2. 矢量 $|V\rangle$ 与标量 a 可以数乘: $a|V\rangle$.

并且矢量加法与数乘需要满足以下性质：

1. 数乘对矢量加法的分配律: $a(|V\rangle + |W\rangle) = a|V\rangle + a|W\rangle$.
2. 数乘对标量加法的分配律: $(a + b)|V\rangle = a|V\rangle + b|V\rangle$.
3. 数乘的结合律: $(ab)|V\rangle = a(b|V\rangle)$.
4. 矢量加法的交换律: $|V\rangle + |W\rangle = |W\rangle + |V\rangle$.
5. 矢量加法的结合律: $(|V\rangle + |W\rangle) + |Z\rangle = |V\rangle + (|W\rangle + |Z\rangle)$.
6. 存在某个特殊的矢量 $|0\rangle$, 称为**零矢量**, 使得对任何矢量 $|V\rangle$ 都成立 $|V\rangle + |0\rangle = |V\rangle$.
7. 任何矢量 $|V\rangle$ 都存在关于矢量加法的逆元 $|-V\rangle$, 使得 $|V\rangle + |-V\rangle = 0$.

总之, 矢量加法与乘法需要满足的上述 7 条性质都是“自然而然的”. 特别注意, 在上述定义中用于数乘的**标量**之全体构成一个**域**. 常见的域包括**实数域** \mathbb{R} 与**复数域** \mathbb{C} . 如果我们把域取为实数域 \mathbb{R} , 也就是允许实数与矢量作数乘, 则这样的矢量空间称为**实矢量空间**; 而如果把域取为复数域 \mathbb{C} , 即允许复数和矢量作数乘, 这种矢量空间称为**复矢量空间**.

以后, 若不特别说明, 我们提到的矢量空间都默认是复矢量空间.

我们依然可以像以前一样, 把矢量想象成有大小有方向的量. 常见的物理量, 比如位置, 速度, 力, 动量等等都是矢量的例子. 它们与实数作数乘可以直观理解为“沿着矢量的方向把矢量的长度作伸缩”, 而它们与复数作数乘似乎没有实际意义. 以后会知道, 量子力学中的矢量需要与复数作数乘, 也就是说量子力学中常把域取为复数域 \mathbb{C} .

我们以前习惯用字母上方加箭头 \vec{v} 或者粗体 \mathbf{v} 来表示矢量, 但在量子力学中这个习惯要改, 要习惯用 $|V\rangle$ 来表示矢量. 符号 “ $|V\rangle$ ” 读作 “ket V ”.¹

要注意, 我们不要将矢量局限于中学数学所学“平面向量”与“空间向量”. 事实上, 根据定义可知, 只要能作加法, 能作数乘, 并且加法和数乘满足那些自然的性质, 就叫做**矢量**.² 比如, 下面这些都是矢量:

¹来自英文单词 bracket. 脑洞大开的物理学家把此单词拆成 bra 与 ket 两部分. 以后我们将见到 bra.

²数学家有云: “什么是鸭子? 只要嘴巴像鸭子, 羽毛像鸭子, 游泳姿势像鸭子, 那它就是鸭子! 嗯,

1. 中学所学“平面向量”与“空间向量”.
2. 考虑所有 2×2 矩阵构成的集合. 其加法为通常的矩阵加法, 数乘为通常的矩阵数乘. 这样就“能作加法能作数乘”, 并且容易验证如此加法与数乘的确满足矢量空间定义里的 7 条性质, 从而构成矢量空间. 这个矢量空间中的零矢量为 $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
3. 给定某个区间 $[0, L]$, 考虑定义在这个区间上的所有的函数 $f(x)$ 构成的集合. 两个函数可以相加 $(f+g)(x) = f(x) + g(x)$, 函数可以和标量作数乘: $(af)(x) = af(x)$. 容易验证此加法和数乘的确满足矢量空间定义中的 7 条性质. 因此定义在区间 $[0, L]$ 上的函数构成的集合在此意义下是矢量空间. [该集合中的元素, 也就是定义在 $[0, L]$ 上的函数, 称为矢量.] 这个矢量空间里的零矢量是恒为零的常值函数.

定义. 对于矢量空间 \mathbb{V} ,

1. 称 \mathbb{V} 的子集 $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$ **线性相关**, 如果存在不全为零的标量 a_1, \dots, a_n , 使得

$$\sum_{i=1}^n a_i |i\rangle = |0\rangle.$$

若不然, 则称 $\{|1\rangle, \dots, |n\rangle\}$ **线性无关**.

2. 如果 \mathbb{V} 存在 n 元线性无关子集, 并且任何 $n+1$ 元子集都线性相关, 则称矢量空间 \mathbb{V} 的**维数**是 n ; 如果对任意 (充分大的) 正整数 n , \mathbb{V} 都存在 n 元线性无关子集, 则称 \mathbb{V} 是**无穷维空间**.
3. 若 \mathbb{V} 的维数是 n , 则 \mathbb{V} 的 n 元线性无关子集称为 \mathbb{V} 的一组**基**.

例如, 2×2 矩阵构成的矢量空间的维数是 4,

$$|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad |3\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad |4\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

构成它的一组基. 再比如, 定义 $[0, L]$ 上的全体函数构成的矢量空间是无穷维的,

就算它三个头五条腿, 只要嘴巴, 羽毛, 游泳姿势像鸭子, 它也是鸭子.”

这是因为对任意正整数 n , 考虑函数 f_1, f_2, \dots, f_n 如下:

$$f_k(x) = \begin{cases} 1 & \text{若 } \frac{k-1}{n}L < x < \frac{k}{n}L, \\ 0 & \text{其余情况} \end{cases}, \quad k = 1, 2, \dots, n,$$

则容易验证集合 $\{|f_1\rangle, \dots, |f_n\rangle\}$ 线性无关.

以下性质众所周知:

性质 1.1. 设 $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$ 是矢量空间 \mathbb{V} 的一组基, 则该矢量空间中的任意矢量 $|V\rangle$ 都可唯一地表示为

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle,$$

标量 v_i 称为矢量 $|V\rangle$ 在这组基下的第 i 个分量.

若 $|V\rangle = \sum_{i=1}^n v_i |i\rangle, |W\rangle = \sum_{i=1}^n w_i |i\rangle$, 则显然有

$$|V\rangle + |W\rangle = \sum_{i=1}^n (v_i + w_i) |i\rangle.$$

$$a|V\rangle = \sum_{i=1}^n av_i |i\rangle.$$

粗俗地说, 在一组基下, 两个矢量相加其实就是“各分量相加”, 矢量的数乘其实就是“各分量同时乘”.

内积空间.

我们在中学数学中所学“向量”是可以谈论长度和夹角的. 而现在, 对于一般的矢量空间, 我们如何谈论长度和夹角呢? 比如 $[0, L]$ 上的函数构成的矢量空间, 随便取这里面的两个函数 $f(x) = x, g(x) = x^2$, 能否谈论矢量 $|f\rangle, |g\rangle$ 的夹角呢?

在中学, 我们认为长度与夹角不言自明. 有了长度和夹角, 我们可以定义两个向量的内积:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta.$$

在直角坐标系下, $\mathbf{a} = (a_x, a_y, a_z)$, $\mathbf{b} = (b_x, b_y, b_z)$, 则内积可以通过下式计算:

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z.$$

并且我们早在中学就知道 (平面, 空间) 向量的内积 (点乘) 满足以下三条性质:

1. 对称性: $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.
2. 严格正定性: $\mathbf{a} \cdot \mathbf{a} \geq 0$, 并且等号成立当且仅当 $\mathbf{a} = \mathbf{0}$.
3. 线性性: $\mathbf{a} \cdot (\lambda \mathbf{b} + \mu \mathbf{c}) = \lambda \mathbf{a} \cdot \mathbf{b} + \mu \mathbf{a} \cdot \mathbf{c}$.

而现在, 对于一般的矢量空间, 并且特别注意在量子力学中我们需要考虑复数域 \mathbb{C} , 即需要谈论矢量与复数作数乘, 于是情况有点变化.

定义. 对于矢量空间 \mathbb{V} , 如果 \mathbb{V} 上有如下的二元运算: 输入两个矢量 $|V\rangle, |W\rangle$, 则输出一个复数, 记作 $\langle V|W\rangle$, 并且这个二元运算满足以下三条性质:

1. 共轭对称性: $\langle V|W\rangle = \langle W|V\rangle^*$,³
2. 严格正定性: $\langle V|V\rangle \geq 0$, 且等号成立当且仅当 $|V\rangle = |0\rangle$.
3. 关于 ket 的线性性: $\langle V|aW + bZ\rangle = a \langle V|W\rangle + b \langle V|Z\rangle$,

则称 \mathbb{V} 配以该二元运算为**内积空间**,⁴ 该二元运算称为 \mathbb{V} 上的**内积**.⁵ 这里的 $|aW + bZ\rangle := a|W\rangle + b|Z\rangle$.

在量子力学中我们的确习惯考虑复数域 \mathbb{C} 上的情形. 上述定义在 \mathbb{C} 上都能适用, 那它更适用于实数域 \mathbb{R} . 若只考虑实数域 \mathbb{R} , 则上述定义中所要求的三条性质与我们中学所学无异.

我们常把 $a|V\rangle + b|W\rangle$ 简记为 $|aV + bW\rangle$, “ $\langle |$ ” 的情形也类似. 设 \mathbb{V} 为内积空间, 则由内积空间的定义, 容易证明:

$$\langle aW + bZ|V\rangle = a^* \langle W|V\rangle + b^* \langle Z|V\rangle. \quad (1.1)$$

³这里的星号 $*$ 是指**复数共轭**. 即, 若 $z = a + bi$, 其中 a, b 为实数, 则 $z^* = a - bi$

⁴也就是线性代数教科书里所谓的**酉空间**.

⁵人类的本质是复读机. 数学家有云: “什么是鸭子? 只要嘴巴像鸭子, 羽毛像鸭子, 游泳姿势像鸭子, 那它就是鸭子! 嗯, 就算它三个头五条腿, 只要嘴巴, 羽毛, 游泳姿势像鸭子, 它也是鸭子.” 聪明的小朋友们想一想, 到底什么是内积呢?

定义. 设 \mathbb{V} 为内积空间, $|V\rangle, |W\rangle$ 是 \mathbb{V} 中的矢量, 则

1. $\|V\| := \sqrt{\langle V|V\rangle}$ 称为矢量 V 的模长 (长度).
2. 如果 $\langle V|W\rangle = 0$, 则称矢量 $|V\rangle$ 与 $|W\rangle$ 正交 (垂直).
3. 设 $\{|1\rangle, \dots, |n\rangle\}$ 是 \mathbb{V} 中的一组基, 如果这组基中的矢量的模长都是 1, 并且两两正交, 则称这组基为么正基.

以前我们认为, “先有长度和夹角, 再由长度和夹角定义内积”. 现在时代变了, 数学家的思维方式是反人类的, 如今我们先有内积, 再由内积来定义什么是长度, 什么是垂直.

在实数域 \mathbb{R} 的情形, 两个矢量的内积是实数, 于是我们可以通过中学数学熟知的公式

$$\cos \theta = \frac{\langle V|W\rangle}{\|V\|\|W\|}$$

来定义 $|V\rangle$ 与 $|W\rangle$ 的夹角. 而在复数域 \mathbb{C} 的情形, 也就是量子力学常用的情形, 两个矢量的内积 $\langle V|W\rangle$ 可以不是实数, 而是一般的复数. 我们无法模仿实数域 \mathbb{R} 的情况来谈论任意两个矢量的夹角.

现在, 设 $\{|1\rangle, \dots, |n\rangle\}$ 是内积空间 V 的一组么正基, 则

$$\langle i|j\rangle = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad 1 \leq i, j \leq n.$$

对于矢量 $|V\rangle = \sum_{i=1}^n v_i |i\rangle$, $|W\rangle = \sum_{i=1}^n w_i |i\rangle$, 则容易验证

$$\langle V|W\rangle = \sum_{i=1}^n \sum_{j=1}^n v_i^* w_j \langle i|j\rangle = \sum_{i=1}^n v_i^* w_i, \quad (1.2)$$

特别地,

$$\|V\| = \sqrt{\langle V|V\rangle} = \left(\sum_{i=1}^n |v_i|^2 \right)^{\frac{1}{2}}.$$

在么正基 $\{|1\rangle, \dots, |n\rangle\}$ 下, 我们常常把矢量 $|V\rangle$ 等同于 $|V\rangle$ 在此基下的各个分量所排成的数组 (列向量):

$$|V\rangle \mapsto \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

$|W\rangle$ 也类似. 则内积 $\langle V|W\rangle$ 可以用矩阵乘法表示为:

$$\langle V|W\rangle = (v_1^*, \dots, v_n^*) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

注意等号右边的行向量为 $|V\rangle$ 所对应的矩阵的**共轭转置**.

希尔伯特空间.

设 \mathbb{V} 是内积空间, 则我们可以通过矢量的模长来自然定义两个矢量之间的距离, 进而定义收敛性. 具体地, 设 $|V_n\rangle, n = 1, 2, \dots$ 与 $|V\rangle$ 是内积空间 \mathbb{V} 中的矢量, 如果

$$\lim_{n \rightarrow \infty} \|V_n - V\| = 0,$$

则称矢量序列 $|V_n\rangle$ 收敛于 $|V\rangle$, 记作 $\lim_{n \rightarrow \infty} |V_n\rangle = |V\rangle$.

定义. 设 \mathbb{V} 是内积空间, $|V_n\rangle, n = 1, 2, \dots$ 是 \mathbb{V} 中的矢量序列.

1. 如果

$$\forall \varepsilon > 0, \exists N > 0, \forall n_1, n_2 \geq N, \|V_{n_1} - V_{n_2}\| < \varepsilon,$$

则称序列 $\{|V_n\rangle\}_{n \geq 1}$ 为**柯西序列**.

2. 如果 \mathbb{V} 中的任何柯西序列都收敛, 则称 \mathbb{V} 的内积是**完备**的, 并且称内积空间 \mathbb{V} 为**希尔伯特空间**(Hilbert space).

上述定义是数学家们搞的, 不熟悉的读者可以先不深究. 事实上, 可以证明:

性质 1.2. 有限维内积空间都是希尔伯特空间.

也就是说, 所谓内积的“完备性”其实是针对无穷维空间的. 大致来说, 增加所谓完备性条件后, (无穷维) 希尔伯特空间的性质与有限维空间非常类似. 有限维空间的**绝大多数**常见结论都能类似照搬的无穷维希尔伯特空间上. 等我们以后躲不开无穷维空间的时候, 再进一步介绍.

以后我们会知道, 量子力学的理论建立在希尔伯特空间之上.

左矢与右矢.

设 \mathbb{V} 为 n 维内积空间, $\{|1\rangle, \dots, |n\rangle\}$ 是 \mathbb{V} 的一组幺正基. 证如上文所说, 在这组给定的幺正基下, \mathbb{V} 中的任何矢量 $|V\rangle$ 都自然视为列向量 ($n \times 1$ 矩阵):

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle \mapsto \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

事实上, \mathbb{V} 中的矢量也可通过某种方式视为**行向量**. 对于矢量 $|V\rangle$, 我们按如下规则引入符号 $\langle V|$, 读作“bra V ”,⁶ 称为矢量 $|V\rangle$ 的**伴随** (或者叫**共轭**).

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle \implies \langle V| = \sum_{i=1}^n \langle i| v_i^*,$$

并且 $\langle V|$ 自然视为行向量 ($1 \times n$ 矩阵)

$$(v_1^*, \dots, v_n^*),$$

这个行向量是 $|V\rangle$ 所对应的列向量的**共轭转置**. 在此记号下, 两个矢量 $|V\rangle$ 与 $|W\rangle$ 的内积 $\langle V|W\rangle$ 恰好是 $\langle V|$ 与 $|W\rangle$ 所对应的矩阵的矩阵乘法.

形如 $|V\rangle$ 的矢量称为**右矢**(ket), 形如 $\langle V|$ 的符号称为**左矢**(bra).

⁶人类的本质是复读机. 脑洞大开的物理学家把英文单词 bracket 拆成 bra 与 ket 两部分. 我们之前见过 ket, 这里出现 bra.

我们知道右矢 $|V\rangle$ 可以与标量作数乘, 而左矢也类似. 事实上, 对于左矢 $\langle V|$ 与复数 a , 我们记

$$\langle aV| = \langle V|a^*,$$

注意等号右边是 a^* , 即 a 的复数共轭.

矢量在么正基下的展开.

我们知道, 若 $\{|1\rangle, \dots, |n\rangle\}$ 是矢量空间 \mathbb{V} 的一组基, 则 \mathbb{V} 中的任何矢量 $|V\rangle$ 在这组基下都能唯一地表示为

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle. \quad (1.3)$$

现在, 若 \mathbb{V} 是内积空间, 并且 $\{|1\rangle, \dots, |n\rangle\}$ 是么正基, 则矢量 $|V\rangle$ 在此基下的各个分量 v_i 可以用内积来表示. 我们用左矢右矢的记号来推导之. 对每个 $1 \leq j \leq n$, 将等式(1.3)两边与 $|j\rangle$ 作内积, 得

$$\langle j|V\rangle = \sum_{i=1}^n v_i \langle j|i\rangle = \sum_{i=1}^n v_i \delta_{ji} = v_j.$$

也就是说, $|V\rangle$ 的第 j 分量 v_j 其实就是 $|V\rangle$ 与基矢量 $|j\rangle$ 的内积 $\langle j|V\rangle$. 于是, $|V\rangle$ 可写为

$$|V\rangle = \sum_{i=1}^n \langle i|V\rangle |i\rangle = \sum_{i=1}^n |i\rangle \langle i|V\rangle. \quad (1.4)$$

类似地, 对于左矢 $\langle V|$, $\langle V| = \sum_{i=1}^n \langle i|v_i^*$, 展开系数 $v_i^* = \langle V|i\rangle = (\langle i|V\rangle)^*$, 从而

$$\langle V| = \sum_{i=1}^n \langle V|i\rangle \langle i|. \quad (1.5)$$

施瓦茨不等式与三角不等式.

以下是内积空间中基本且重要的性质:

性质 1.3. 设 V 是内积空间, $|V\rangle, |W\rangle$ 是 V 中的任意矢量. 则成立:

1. 施瓦茨不等式:

$$\|\langle V|W\rangle\| \leq \|V\|\|W\|. \quad (1.6)$$

2. 三角不等式:

$$\|V + W\| \leq \|V\| + \|W\|. \quad (1.7)$$

注意上述性质同样适用于复数域 \mathbb{C} 的情形. 其证明与实数域 \mathbb{R} 的情形类似, 不算困难. 留做习题或者参考任何一本线性代数教材.

子空间

定义. 设 V 是矢量空间, V_i 是 V 的一个子集. 如果 V_i 在 V 的矢量加法, 数乘运算下也构成矢量空间, 则称 V_i 是 V 的**子空间**(subspace).

我们来看一些简单例子.

1. 考虑二维实矢量空间 $V^2(\mathbb{R})$, 即中学所学“平面向量”之全体. 则所有与 x 轴平行的向量构成 $V^2(\mathbb{R})$ 的一个子空间, 这个子空间的维数是 1.
2. 考虑三维实矢量空间 $V^3(\mathbb{R})$, 即中学所学“空间向量”之全体. 则容易验证子集 $X = \{(x, y, z) \in V^3(\mathbb{R}) \mid x + 2y - z = 0\}$ 构成 $V^3(\mathbb{R})$ 的一个子空间. 事实上 X 是三维空间 $V^3(\mathbb{R})$ 中的一个平面.
3. 设 V 是全体 2×2 矩阵构成的矢量空间. 则全体 2×2 对称矩阵之全体构成 V 的子空间, 这个子空间的维数是 3; 全体 2×2 反对称矩阵之全体构成 V 的子空间, 这个子空间的维数是 1.
4. 设 V 为定义在区间 $[0, L]$ 上的所有函数所构成的矢量空间, 则 $[0, L]$ 上的连续函数之全体构成 V 的一个子空间.

看完简单例子, 再看一些重要例子.

定义. 设 V 是矢量空间, V_1 与 V_2 是 V 的两个子空间,

1. 定义两个子空间的和

$$\mathbb{V}_1 + \mathbb{V}_2 = \left\{ |aV_1 + bV_2\rangle \mid |V_1\rangle \in \mathbb{V}_1, |V_2\rangle \in \mathbb{V}_2, a, b \in \mathbb{C} \right\}.$$

容易验证这的确也构成 \mathbb{V} 的子空间.

2. 如果 $\mathbb{V}_1 \cap \mathbb{V}_2 = \{|0\rangle\}$, 则称 $\mathbb{V}_1 + \mathbb{V}_2$ 是 \mathbb{V}_1 与 \mathbb{V}_2 的直和, 记作

$$\mathbb{V}_1 \oplus \mathbb{V}_2.$$

而对于内积空间, 我们还有:

定义. 设 \mathbb{V} 为 n 维内积空间, $\mathbb{V}_1, \mathbb{V}_2$ 是 \mathbb{V} 的子空间,

1. 如果对 \mathbb{V}_1 中的任何矢量 $|V\rangle$, 以及 \mathbb{V}_2 中的任何矢量 $|W\rangle$, 都有 $\langle V|W\rangle = 0$, 则称子空间 \mathbb{V}_1 与 \mathbb{V}_2 正交.
2. 容易验证与 \mathbb{V}_1 中的左右矢量都正交的矢量之全体

$$\mathbb{V}_1^\perp := \left\{ |W\rangle \in \mathbb{V} \mid \langle W|V\rangle = 0, \forall |V\rangle \in \mathbb{V}_1 \right\}$$

构成 \mathbb{V} 的子空间, 该子空间称为 \mathbb{V}_1 的正交补.

容易验证, 如果子空间 \mathbb{V}_1 与 \mathbb{V}_2 正交, 那么 $\mathbb{V}_1 + \mathbb{V}_2$ 必为直和, 此时直和 $\mathbb{V}_1 \oplus \mathbb{V}_2$ 称为正交直和.

1.2 线性算子与矩阵

线性算子的定义.

定义. 设 \mathbb{V} 为矢量空间, 则 \mathbb{V} 上的线性算子 Ω (也就是线性代数教材里所谓线性变换) 是指这样的映射: 输入 \mathbb{V} 中的矢量 $|V\rangle$, 则输出一个 \mathbb{V} 中的矢量, 记作 $\Omega|V\rangle$; 并且此映射满足:

$$\Omega|aV + bW\rangle = a\Omega|V\rangle + b\Omega|W\rangle \quad (1.8)$$

对 \mathbb{V} 中任意矢量 $|V\rangle, |W\rangle$ 以及任意标量 a, b 都成立.

也就是说, 线性算子 Ω 的“功能”是“线性地”作用于右矢 $|V\rangle$. 而特别地, 对于内积空间, 线性算子还有另外的“功能”, 那就是, 它也能够作用在左矢上.

定义. 设 Ω 为内积空间 \mathbb{V} 上的线性算子, $\langle V|$ 为 \mathbb{V} 中的左矢, 则定义 $\langle V|\Omega$ 是由如下性质所唯一确定的左矢: 对任意右矢 $|W\rangle$,

$$(\langle V|\Omega)|W\rangle = \langle V|(\Omega|W\rangle) =: \langle V|\Omega|W\rangle. \quad (1.9)$$

容易验证, 线性算子在左矢上的作用也满足线性性:

$$\langle aV + bW|\Omega = a\langle V|\Omega + b\langle W|\Omega \quad (1.10)$$

对任意左矢 $\langle V|, \langle W|$ 与标量 a, b 都成立.

定义. 设 Ω 与 Λ 是矢量空间 \mathbb{V} 上的两个线性算子, 则定义一个新的线性算子 $\Omega\Lambda$, 它满足: 对任意矢量 $|V\rangle$,

$$(\Omega\Lambda)|V\rangle := \Omega(\Lambda|V\rangle) =: \Omega|\Lambda V\rangle.$$

线性算子 $\Omega\Lambda$ 称为 Ω 与 Λ 的**乘积**.

用数学的语言, 线性算子首先是集合 \mathbb{V} 到 \mathbb{V} 的**映射**, 而算子乘积 $\Omega\Lambda$ 无非是两个映射的**复合**.

一般来说, $\Omega\Lambda$ 与 $\Lambda\Omega$ 未必相等, 即线性算子的乘积不满足交换律. 我们记

$$[\Omega, \Lambda] := \Omega\Lambda - \Lambda\Omega, \quad (1.11)$$

容易验证 $[\Omega, \Lambda]$ 也是线性算子, 它称为 Ω 与 Λ 的**对易子**(commutator). [数学书里也把它叫做**交换子**]. 容易验证交换子满足如下等式: 对任意线性算子 Ω, Λ, Θ , 以及标量 a, b , 成立

$$[\Omega, \Lambda] = -[\Lambda, \Omega] \quad (1.12)$$

$$[\Omega, a\Lambda + b\Theta] = a[\Omega, \Lambda] + b[\Omega, \Theta] \quad (1.13)$$

$$[\Omega, [\Lambda, \Theta]] = [[\Omega, \Lambda], \Theta] + [\Lambda, [\Omega, \Theta]] \quad (1.14)$$

$$[\Omega, \Lambda\Theta] = \Lambda[\Omega, \Theta] + [\Omega, \Lambda]\Theta \quad (1.15)$$

对于矢量空间 \mathbb{V} , 我们考虑这样的映射: 把 \mathbb{V} 中的每个矢量 $|V\rangle$ 都映到 $|V\rangle$ 自己. 显然该映射是线性算子, 我们把它记为 I , 即 $I|V\rangle = |V\rangle$. 称这个线性算子 I 为**恒等算子**. 显然, 对任何线性算子 Ω , 都成立 $\Omega = I\Omega = \Omega I$.

定义. 对于矢量空间 \mathbb{V} 上的线性算子 Ω , 如果存在线性算子 Ω^{-1} 使得

$$\Omega\Omega^{-1} = \Omega^{-1}\Omega = I,$$

则称线性算子 Ω 可逆, Ω^{-1} 是 Ω 的逆算子, 简称逆.

数学上可以证明, 如果 Ω 可逆, 则 Ω 的逆算子是唯一的. 特别注意, 并非所有的线性算子都可逆. 此外容易证明, 如果线性算子 Ω, Λ 都可逆, 则乘积 $\Omega\Lambda$ 也可逆, 并且

$$(\Omega\Lambda)^{-1} = \Lambda^{-1}\Omega^{-1}. \quad (1.16)$$

线性算子的矩阵表示.

我们不妨考虑 n 维内积空间 \mathbb{V} , 取定 \mathbb{V} 的一组幺正基 $\{|1\rangle, \dots, |n\rangle\}$. 我们之前已经知道, 可以不妨粗暴地将右矢 $|V\rangle$ 视为列向量 ($n \times 1$ 矩阵), 把左矢 $\langle W|$ 视为行向量 ($1 \times n$ 矩阵). 这样子, 矢量内积 $\langle W|V\rangle$ 相应地就是矩阵乘法.

我们做了什么呢? 矢量空间是抽象定义的, 内积也是抽象定义的.⁷ 而行向量, 列向量, 矩阵乘法相比之下更接地气, 看得见摸得着, 容易实际操作. 现在, 线性算子从定义上看也是“抽象的东西”, 我们企图给它对应某个“具体的东西”.

设 Ω 为 \mathbb{V} 上的一个线性算子, 我们记

$$\Omega_{ij} := \langle i|\Omega|j\rangle, \quad 1 \leq i, j \leq n. \quad (1.17)$$

现在考虑 Ω 作用于基矢量 $|i\rangle$, 由公式(1.4)可知,

$$\Omega|i\rangle = \sum_{j=1}^n |j\rangle \langle j|\Omega|i\rangle = \sum_{j=1}^n \Omega_{ji}|j\rangle. \quad (1.18)$$

于是, 对于 \mathbb{V} 中的一般的矢量 $|V\rangle = \sum_{i=1}^n v_i|i\rangle$, 记 $|V'\rangle := \Omega|V\rangle = \sum_{i=1}^n v'_i|i\rangle$, 我

⁷人类的本质是复读机. 吾日三省吾身: “什么是鸭子? 什么是鸭子? 什么是鸭子?” 如果忘了什么是鸭子, 就往前翻, 找找脚注.

们来推导系数 v'_i 与 v_i 之间的关系. 其实直接计算就行:

$$\begin{aligned}\Omega|V\rangle &= \sum_{i=1}^n v_i \Omega|i\rangle = \sum_{i=1}^n \sum_{j=1}^n v_i \Omega_{ji} |j\rangle \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n \Omega_{ji} v_i \right) |j\rangle.\end{aligned}\tag{1.19}$$

与 $\Omega|V\rangle = \sum_{i=1}^n v'_i |i\rangle$ 比较系数, 得

$$v'_i = \sum_{j=1}^n \Omega_{ij} v_j.\tag{1.20}$$

我们学过矩阵乘法, 上式用矩阵的语言可重写为:

$$\begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \cdots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \cdots & \Omega_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \cdots & \Omega_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.\tag{1.21}$$

这启发我们, 在给定的么正基下, 线性算子 Ω 可自然地视为 $n \times n$ 矩阵:

$$\Omega \mapsto \begin{pmatrix} \Omega_{11} & \Omega_{12} & \cdots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \cdots & \Omega_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \cdots & \Omega_{nn} \end{pmatrix},$$

其中矩阵元 $\Omega_{ij} = \langle i|\Omega|j\rangle$. 再按之前的方式将矢量 $|V\rangle$ 视为列向量, 那么线性算子 Ω 在 $|V\rangle$ 的作用 $\Omega|V\rangle$ 在此意义下其实就是通常的矩阵乘法——把抽象的算子作用看成具体的矩阵运算!

对于恒等算子 I , 由 $I_{ij} = \langle i|I|j\rangle = \langle i|j\rangle = \delta_{ij}$, 从而恒等算子对应的矩阵为

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, \text{ 即单位矩阵. 回忆(1.4)式,}$$

$$|V\rangle = \sum_{i=1}^n |i\rangle \langle i|V\rangle = \left(\sum_{i=1}^n |i\rangle \langle i| \right) |V\rangle.$$

于是我们自然把符号 “ $\sum_{i=1}^n |i\rangle \langle i|$ ” 想象成恒等算子 I . 一般地, 注意(1.19)式, 并且注意系数 $v_i = \langle i|V\rangle$, 从而

$$\begin{aligned} \Omega|V\rangle &= \sum_{i,j=1}^n v_i \Omega_{ji}|j\rangle = \sum_{i,j=1}^n \langle i|V\rangle \Omega_{ji}|j\rangle \\ &= \sum_{i,j=1}^n \Omega_{ij}|i\rangle \langle j|V\rangle = \left(\sum_{i,j=1}^n |i\rangle \langle j| \right) |V\rangle. \end{aligned}$$

于是有:

$$\Omega = \sum_{i,j=1}^n \Omega_{ij}|i\rangle \langle j|. \quad (1.22)$$

其中 “ $|i\rangle \langle j|$ ” 是指这样的映射: 它把基向量 $|k\rangle$ 映为 $|i\rangle \langle j|k\rangle = \delta_{jk}|i\rangle$. 容易验证这是线性算子, 并且线性算子 $|i\rangle \langle j|$ 所对应的矩阵满足: 第 i 行 j 列等于 1, 其余矩阵元全为 0.

我们也可从另一个角度来看 $|i\rangle \langle j|$. 我们把 $|i\rangle$ 视为列向量 $(0, \dots, 1, \dots, 0)^\top$ (第 i 分量为 1, 其余分量为 0), 把 $\langle j|$ 视为行向量 $(0, \dots, 1, \dots, 0)$ (第 j 分量为 1, 其余分量为 0). 则容易验证算子 $|i\rangle \langle j|$ 所对应的矩阵

$$|i\rangle \langle j| \mapsto \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} (0, \dots, 1, \dots, 0),$$

右边按通常的矩阵乘法计算, 得到的 $n \times n$ 矩阵恰为第 i 行 j 列为 1, 其余矩阵元都为 0 的矩阵.

一般地, 设 $|V\rangle, |W\rangle$ 是 \mathbb{V} 的两个矢量, 则 $|V\rangle\langle W|$ 是线性算子: 它把矢量 $|Z\rangle$ 映为 $|V\rangle\langle W|Z\rangle = \langle W|Z\rangle|V\rangle$.

例子: 投影算子.

设 \mathbb{V} 为 n 维内积空间, $\{|1\rangle, \dots, |n\rangle\}$ 是它的一组么正基. 我们记

$$\mathbb{P}_i := |i\rangle\langle i|, \quad i = 1, \dots, n. \quad (1.23)$$

称这样的算子为**投影算子**. 对任何矢量 $|V\rangle = \sum_{i=1}^n v_i|i\rangle$, 容易验证

$$\begin{aligned} \mathbb{P}_i|V\rangle &= \sum_{j=1}^n v_j\mathbb{P}_i|j\rangle \\ &= \sum_{j=1}^n v_j|i\rangle\langle i|j\rangle = \sum_{j=1}^n v_j\delta_{ij}|i\rangle = v_i|i\rangle. \end{aligned}$$

可见 $\mathbb{P}_i|V\rangle$ 恰为矢量 $|V\rangle$ 沿 $|i\rangle$ 方向的部分. 此外注意

$$I = \sum_{i=1}^n |i\rangle\langle i| = \sum_{i=1}^n \mathbb{P}_i, \quad (1.24)$$

即恒等算子可分解为沿各分项的投影算子之和. [如果读者们觉得我们花太多笔墨在恒等算子上, 甚至觉得无聊没意思, 那就大错特错了. 我们走着瞧.]

投影算子在左矢上的作用效果也类似. 对于 $\langle V| = \sum_{i=1}^n v_i\langle i|$, 则 $\langle V| = \sum_{i=1}^n \langle i|v_i^*$, 于是

$$\langle V|\mathbb{P}_i = \sum_{j=1}^n \langle j|i\rangle\langle i|v_j^* = \sum_{j=1}^n \delta_{ji}\langle i|v_j^* = \langle i|v_i^*.$$

容易验证, 投影算子 \mathbb{P}_i 所对应的矩阵为对角阵 $\text{diag}(0, \dots, 1, \dots, 0)$, 第 i 个主对角元为 1, 其余元素全为零. 此外, 两个投影算子的乘积满足

$$\mathbb{P}_i\mathbb{P}_j = |i\rangle\langle i|j\rangle\langle j| = \delta_{ij}|i\rangle\langle j| = \delta_{ij}\mathbb{P}_i = \delta_{ij}\mathbb{P}_j.$$

算子乘积与矩阵乘法.

设 \mathbb{V} 为 n 维内积空间, $\{|1\rangle, \dots, |n\rangle\}$ 是它的一组幺正基. 在此基下, 我们把矢量 $|V\rangle$ 看成列向量, 把线性算子 Ω 看成 $n \times n$ 矩阵, 则线性算子在矢量上的作用 $\Omega|V\rangle$ 自然就相当于矩阵与列向量相乘.

设 Ω, Λ 为两个线性算子, 考虑线性算子的乘积 $\Omega\Lambda$, 这也是一个线性算子. 线性算子 $\Omega\Lambda$ 所对应的矩阵是什么? 事实上, 注意 $\Omega\Lambda = \Omega I \Lambda$, 于是

$$\begin{aligned} (\Omega\Lambda)_{ij} &= \langle i|\Omega\Lambda|j\rangle = \langle i|\Omega I \Lambda|j\rangle \\ &= \sum_{k=1}^n \langle i|\Omega|k\rangle \langle k|\Lambda|j\rangle = \sum_{k=1}^n \Omega_{ik} \Lambda_{kj}. \end{aligned}$$

与我们熟悉的**矩阵乘法**法则比较, 上式可重写为:

$$\begin{pmatrix} (\Omega\Lambda)_{11} & (\Omega\Lambda)_{12} & \cdots & (\Omega\Lambda)_{1n} \\ (\Omega\Lambda)_{21} & (\Omega\Lambda)_{22} & \cdots & (\Omega\Lambda)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (\Omega\Lambda)_{n1} & (\Omega\Lambda)_{n2} & \cdots & (\Omega\Lambda)_{nn} \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \cdots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \cdots & \Omega_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \cdots & \Omega_{nn} \end{pmatrix} \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1n} \\ \Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{n1} & \Lambda_{n2} & \cdots & \Lambda_{nn} \end{pmatrix}$$

也就是说: 两个算子的乘积所对应的矩阵, 等于两个算子所对应的矩阵的矩阵相乘. [两个算子的乘积的定义是“自然而然”的, 因此相应地, 矩阵乘法的法则也应该是自然而然的. 很多初学线性代数的小朋友第一次接触矩阵乘法时往往迷惑于矩阵乘法的神奇法则, 感觉仿佛从天而降. 其实不然.]

1.3 伴随算子, 厄米特算子与酉算子.

在本节, 依旧考虑 n 维内积空间 \mathbb{V} , 并给定它的一组幺正基 $\{|1\rangle, \dots, |n\rangle\}$.

伴随算子.

我们知道, \mathbb{V} 中的每个矢量 $|V\rangle$ 都对应于某个左矢 $\langle V|$: 若 $|V\rangle = \sum_{i=1}^n v_i |i\rangle$, 则 $\langle V| = \sum_{i=1}^n \langle i| v_i^*$. [注意这里的复数共轭!] 若把 $|V\rangle$ 视为列向量, $\langle V|$ 视为行向量. 则在此意义下 $\langle V|$ 其实就是 $|V\rangle$ 的**共轭转置**.

我们回忆: 对于标量 a 与矢量 $|V\rangle$, 矢量数乘 $a|V\rangle = |aV\rangle$ 所对应的左矢 $\langle aV|$ 满足: [注意复数共轭, 真的是 a^* , 不是 a .]

$$\langle aV| = \langle V|a^*.$$

现在, 把标量 a 换成一般的线性算子 Ω . 那自然要问: 矢量 $\Omega|V\rangle = |\Omega V\rangle$ 所对应的左矢 $\langle \Omega V|$ 与 $\langle V|$ 之间有何关系? 莫非也是 Ω 的某种“共轭”?

性质 1.4. 给定 n 维内积空间 \mathbb{V} 的一组么正基 $\{|1\rangle, \dots, |n\rangle\}$. 设 Ω 为 \mathbb{V} 上的线性算子, 那么存在线性算子 Ω^\dagger , 满足: 对 \mathbb{V} 中任何矢量 $|V\rangle$, 都成立

$$\langle \Omega V| = \langle V|\Omega^\dagger. \quad (1.25)$$

线性算子 Ω^\dagger 称为 Ω 的**伴随算子**.

证明. 依然是直接验证, 对任意矢量 $|V\rangle = \sum_{i=1}^n v_i |i\rangle$, 注意 $v_i = \langle i|V\rangle$, $v_i^* = \langle V|i\rangle$. 从而

$$\begin{aligned} \langle \Omega V| &= \langle \Omega V|I = \sum_{i=1}^n \langle \Omega V|i\rangle \langle i| \\ &= \sum_{i=1}^n \langle i|\Omega V\rangle^* \langle i| = \sum_{i=1}^n \langle i|\Omega|V\rangle^* \langle i| = \sum_{i,j=1}^n (v_j \langle i|\Omega|j\rangle)^* \langle i| \\ &= \sum_{i,j=1}^n (\Omega_{ij} v_j)^* \langle i| = \sum_{i,j=1}^n \Omega_{ij}^* v_j^* \langle i| = \sum_{i,j=1}^n \Omega_{ij}^* \langle V|j\rangle \langle i| \\ &= \langle V| \left(\sum_{i,j=1}^n \Omega_{ij}^* |j\rangle \langle i| \right). \end{aligned}$$

因此, 只需取

$$\Omega^\dagger = \sum_{i,j=1}^n \Omega_{ij}^* |j\rangle \langle i| = \sum_{i,j=1}^n \Omega_{ji}^* |i\rangle \langle j| \quad (1.26)$$

即可. □

我们考虑伴随算子所对应的矩阵. 由(1.26)立刻得到

$$(\Omega^\dagger)_{ij} = \Omega_{ji}^*.$$

也就是说, Ω^\dagger 所对应的矩阵其实是 Ω 所对应的矩阵的**共轭转置**. 由伴随算子的定义 [或者暴力使用矩阵共轭转置的性质] 容易证明: 对于线性算子 Ω, Λ , 成立

$$\Omega^{\dagger\dagger} = \Omega, \quad (1.27)$$

$$(\Omega\Lambda)^\dagger = \Lambda^\dagger\Omega^\dagger. \quad (1.28)$$

厄米特算子与酉算子.

我们来介绍在量子力学中非常重要的几种线性算子.

定义. 设 Ω 是 n 维内积空间 \mathbb{V} 上的线性算子.

1. 如果 $\Omega^\dagger = \Omega$, 则称 Ω 是**厄米特算子**.
2. 如果 $\Omega^\dagger = -\Omega$, 则称 Ω 是**反厄米特算子**.
3. 如果 $\Omega^\dagger = \Omega^{-1}$, 则称 Ω 是**酉算子**.

对任何线性算子 Ω , 容易验证 $\frac{\Omega + \Omega^\dagger}{2}$ 是厄米特算子, 而 $\frac{\Omega - \Omega^\dagger}{2}$ 是反厄米特算子, 并且

$$\Omega = \frac{\Omega + \Omega^\dagger}{2} + \frac{\Omega - \Omega^\dagger}{2}.$$

如果把算子伴随 \dagger 想象为取复数共轭, 则厄米特算子可以类比为实数: $a^* = a$. 反厄米特算子可以类比为纯虚数: $(ia)^* = -ia$. 而酉算子则类比为模长为 1 的复数 $e^{i\theta}$: $e^{i\theta}e^{-i\theta} = 1$.

厄米特算子, 反厄米特算子, 酉算子 [在给定幺正基下的] 矩阵分别是线性代数课本里所谓 **厄米特矩阵**, **反厄米特矩阵**, **酉矩阵**. [它们分别是对称矩阵, 反对称矩阵, 正交矩阵在复数域 \mathbb{C} 上的推广.]

性质 1.5. 酉算子保持矢量内积. 具体地说, 若 U 为酉算子, 对任何矢量 $|V\rangle, |W\rangle$, 成立

$$\langle V|W\rangle = \langle UV|UW\rangle.$$

证明. 这是因为 $\langle UV|UW\rangle = \langle V|U^\dagger U|W\rangle = \langle V|I|W\rangle = \langle V|W\rangle$. \square

由此容易推出: 酉算子把一组么正基变成另一组么正基, 确切地说, 若 U 是酉算子, $\{|1\rangle, \dots, |n\rangle\}$ 是 \mathbb{V} 的一组么正基, 则 $\{|i'\rangle := U|i\rangle \mid 1 \leq i \leq n\}$ 也是 \mathbb{V} 的一组么正基. [反过来, 我们在下一段将看到, “任何两组么正基都能通过酉算子联系起来”.]

基变换与坐标变换.

我们在一组给定的么正基下, 把矢量 $|V\rangle$ 视为列向量, 把线性算子 Ω 视为 $n \times n$ 矩阵. 但要注意, 把抽象的矢量, 线性算子对应为具体的列向量, 矩阵, 这种对应方式与么正基的选取有关. 同一个矢量, 线性算子, 在不同么正基下所对应的列向量, 矩阵会有所不同.

现在设 $\{|1\rangle, \dots, |n\rangle\}$ 是内积空间 \mathbb{V} 的一组么正基, 而 $\{|1'\rangle, \dots, |n'\rangle\}$ 是 \mathbb{V} 的另一组么正基. 则对 $i = 1, \dots, n$, 注意到:

$$\begin{aligned} |i'\rangle &= \sum_{k=1}^n |k\rangle \langle k|i'\rangle = \sum_{k,j=1}^n |k\rangle \langle k|j'\rangle \delta_{ji} \\ &= \sum_{k,j=1}^n |k\rangle \langle k|j'\rangle \langle j|i\rangle = \left(\sum_{k,j=1}^n \langle k|j'\rangle |k\rangle \langle j| \right) |i\rangle. \end{aligned}$$

我们记线性算子

$$U := \sum_{k,j=1}^n \langle k|j'\rangle |k\rangle \langle j|,$$

则 $U|i\rangle = |i'\rangle$, 也就是说 U 依次将 $|1\rangle, \dots, |n\rangle$ 映为 $|1'\rangle, \dots, |n'\rangle$. 注意 $\{|1'\rangle, \dots, |n'\rangle\}$ 也是么正基, $\langle i'|j'\rangle = \delta_{ij}$, 于是

$$\begin{aligned} U^\dagger U &= \sum_{i,j=1}^n |i\rangle \langle i| U^\dagger U |j\rangle \langle j| = \sum_{i,j=1}^n |i\rangle \langle i'|j'\rangle \langle j| \\ &= \sum_{i,j=1}^n \delta_{ij} |i\rangle \langle j| = \sum_{i=1}^n |i\rangle \langle i| = I, \end{aligned}$$

从而 U 是酉算子. 算子 U 在么正基 $\{|1\rangle, \dots, |n\rangle\}$ 下的矩阵的矩阵元为:

$$U_{ij} := \langle i|U|j\rangle = \langle i|j'\rangle. \quad (1.29)$$

现在, 考虑矢量 $|V\rangle$ 分别在在这两个么正基下的展开:

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle = \sum_{i=1}^n v'_i |i'\rangle,$$

在么正基 $\{|1\rangle, \dots, |n\rangle\}$ 下, $|V\rangle$ 所对应的列向量为 $(v_1, \dots, v_n)^\top$; 而在另一组么正基 $\{|1'\rangle, \dots, |n'\rangle\}$ 下, $|V\rangle$ 对应于另一个列向量 $(v'_1, \dots, v'_n)^\top$. 注意

$$v'_i = \langle i'|V\rangle = \sum_{j=1}^n \langle i'|j\rangle \langle j|V\rangle = \sum_{j=1}^n \langle i'|j\rangle v_j,$$

上式容易改写为如下矩阵形式:

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = \begin{pmatrix} \langle 1'|1\rangle & \cdots & \langle 1'|n\rangle \\ \vdots & \ddots & \vdots \\ \langle n'|1\rangle & \cdots & \langle n'|n\rangle \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.30)$$

正所谓不同么正基下的坐标变换公式.

对于线性算子 Ω , 我们把 Ω 在么正基 $\{|1\rangle, \dots, |n\rangle\}$ 下的矩阵记作 $[\Omega]$, 而把 Ω 在么正基 $|1'\rangle, \dots, |n'\rangle$ 下的矩阵记作 $[\Omega]'$. 那么(1.29)中的酉算子 U 满足

$$[U] = \begin{pmatrix} \langle 1|1'\rangle & \cdots & \langle 1|n'\rangle \\ \vdots & \ddots & \vdots \\ \langle n|1'\rangle & \cdots & \langle n|n'\rangle \end{pmatrix},$$

因此坐标变换公式(1.30)自然改写为

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = [U]^\dagger \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad (1.31)$$

这里的 $[U]^\dagger$ 是矩阵 $[U]$ 的共轭转置.

接下来, 我们考虑同一个算子在不同么正基下的矩阵之间的关系. 设 Ω 为内积空间 \mathbb{V} 上的线性算子, 它在两组么正基下的矩阵分别为 $[\Omega], [\Omega]'$. 则

$$[\Omega]'_{ij} = \langle i'|\Omega|j'\rangle = \sum_{k,l=1}^n \langle i'|k\rangle \langle k|\Omega|l\rangle \langle l|j'\rangle$$

$$= \sum_{k,l=1}^n \langle i'|k\rangle \Omega_{kl} \langle l|j'\rangle$$

写成矩阵形式, 容易验证

$$[\Omega]' = [U]^\dagger [\Omega] [U]. \quad (1.32)$$

也就是说, 矩阵 $[\Omega]'$ 与 $[\Omega]$ 酉相似.

线性算子的迹.

定义. 设 Ω 是 n 维内积空间 \mathbb{V} 上的线性算子, $\{|1\rangle, \dots, |n\rangle\}$ 是 \mathbb{V} 的一组么正基. 我们称

$$\text{tr } \Omega := \sum_{i=1}^n \langle i|\Omega|i\rangle \quad (1.33)$$

为线性算子 Ω 的迹(trace).

从定义里显然可见, $\text{tr } \Omega$ 其实就是矩阵 $[\Omega]$ 的各个对角元之和.

但问题来了, 注意看定义, 先验地看, $\text{tr } \Omega$ 似乎与么正基 $\{|i\rangle\}$ 的选取有关. 如果我们换另一组么正基, 得到的“迹”会不会变化? 也就是说, 算子的迹是否是“算子本身固有的性质”, 是否受“外界环境”影响? 或者用数学家的语言来问, 这个定义是否良定(well-defined)?

假设我们再取另一组么正基 $\{|1'\rangle, \dots, |n'\rangle\}$, 则恒等算子 $I = \sum_{i=1}^n |i'\rangle\langle i'|$ 也成立. 于是

$$\begin{aligned} \sum_{i=1}^n \langle i|\Omega|i\rangle &= \sum_{i,j,k=1}^n \langle i|j'\rangle \langle j'|\Omega|k'\rangle \langle k'|i\rangle \\ &= \sum_{j,k=1}^n \left(\sum_{i=1}^n \langle k'|i\rangle \langle i|j'\rangle \right) \langle j'|\Omega|k'\rangle \\ &= \sum_{j,k=1}^n \langle k'|j'\rangle \langle j'|\Omega|k'\rangle = \sum_{j,k=1}^n \delta_{kj} \langle j'|\Omega|k'\rangle \\ &= \sum_{i=1}^n \langle i'|\Omega|i'\rangle. \end{aligned}$$

可见, 算子 Ω 的迹 $\text{tr } \Omega$ 与么正基的选取无关.

此外, 容易验证迹满足如下性质: 对任意线性算子 Ω, Λ , 都成立

$$\text{tr}(\Omega\Lambda) = \text{tr}(\Lambda\Omega), \quad (1.34)$$

证明留给读者.

1.4 本征值与本征矢量

基本概念.

定义. 设 Ω 是矢量空间 \mathbb{V} 上的线性算子. 如果非零矢量 $|V\rangle$ 与标量 ω 满足

$$\Omega|V\rangle = \omega|V\rangle,$$

则称 ω 是 Ω 的一个**本征值**, $|V\rangle$ 是 Ω 的属于本征值 ω 的**本征矢量**.

以后我们会知道, 这是量子物理当中无比重要的概念! 现在我们令 \mathbb{V} 为内积空间, 先来看一些简单例子:

1. 恒等算子 $I = \sum_{i=1}^n |i\rangle\langle i|$ 具有本征值 1. \mathbb{V} 中的任何非零矢量 $|V\rangle$ 都是 I 的属于该本征值的本征矢量: $I|V\rangle = |V\rangle$.
2. 投影算子 $\mathbb{P}_i = |i\rangle\langle i|$ 具有本征值 0, 1. $|i\rangle$ 是它的属于本征值 1 的本征矢量, 而 $|j\rangle$, ($j \neq i$) 是它的属于本征值 0 的本征矢量.
3. 若 $|V\rangle$ 是 Ω 的属于本征值 ω 的本征矢量, 则对任何非零常数 a , $a|V\rangle$ 也是属于 ω 的本征矢量.
4. 设 \mathbb{V} 是 2 维内积空间, 取么正基 $\{|1\rangle, |2\rangle\}$. 取实数 θ , 考虑算子

$$R(\theta) = \cos\theta|1\rangle\langle 1| + \sin\theta|1\rangle\langle 2| - \sin\theta|2\rangle\langle 1| + \cos\theta|2\rangle\langle 2|,$$

则 $R(\theta)$ 在基 $\{|1\rangle, |2\rangle\}$ 下的矩阵为

$$[R(\theta)] = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

则容易验证 $R(\theta)$ 具有本征值 $\omega_1 = e^{i\theta}$, $\omega_2 = e^{-i\theta}$, 而矢量 $|V_1\rangle = |1\rangle + i|2\rangle$, $|V_2\rangle = |1\rangle - i|2\rangle$ 分别是属于 ω_1, ω_2 的本征矢量.

一般地, 对于线性算子 Ω , 如何求 Ω 的本征值呢? 如果 ω 是 Ω 的一个本征值, 则存在非零矢量 $|V\rangle$ 使得 $\Omega|V\rangle = \omega|V\rangle$, 于是

$$(\Omega - \omega I)|V\rangle = |0\rangle.$$

也就是说, 算子 $\Omega - \omega I$ 把某个非零矢量 $|V\rangle$ 映为零矢量 $|0\rangle$. 从而容易证明算子 $\Omega - \omega I$ 不可逆. 于是由矩阵理论 [见任何一本线性代数教材], 必有

$$\det(\Omega - \omega I) = 0, \quad (1.35)$$

上式等号左边是算子 $\Omega - \omega I$ 在任何一组基下的矩阵的行列式, 它与基的选取无关. 反之, 关于 ω 的方程(1.35)的任何解都是 Ω 的本征值. 方程(1.35)称为线性算子 Ω 的**特征方程**. 注意 $\det(\Omega - \omega I)$ 是关于变量 ω 的 n 次多项式函数, 即形如

$$\det(\Omega - \omega I) = \sum_{k=0}^n c_k \omega^k.$$

于是称 n 次多项式 $\det(\Omega - \omega I)$ 为线性算子 Ω 的**特征多项式**.

定义. 设 Ω 是 n 维矢量空间 \mathbb{V} 上的线性算子. 如果存在 \mathbb{V} 的一组基 $\{|1\rangle, \dots, |n\rangle\}$, 使得每个基矢量 $|i\rangle$ 都是 Ω 属于某个本征值 ω_i 的本征矢量, 则称线性算子 Ω **可对角化**.

注意这里并不要求 \mathbb{V} 具有内积, 从而无法谈论模长与正交, 于是并不要求这组基是么正基.

若 Ω 可对角化, 则 Ω 在基 $|1\rangle, \dots, |n\rangle$ 下的矩阵为对角阵

$$\begin{pmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{pmatrix}.$$

现在, 任取 \mathbb{V} 的一组基, 考虑 Ω 在该基下的矩阵 $[\Omega]$. 则容易验证: 线性算子 Ω 可对角化, 当且仅当存在可逆矩阵 P , 使得 $P^{-1}[\Omega]P$ 是对角矩阵. 也就是线性代数中所谓的“相似于对角矩阵”.

对角矩阵的好处是方便计算, 于是深受物理学家喜爱, 我们也想方设法把算子, 矩阵搞成对角的. 事实上, 并非所有的线性算子都可对角化, 不过“几乎所有

的”线性算子都可对角化.⁸至于到底什么样的算子才可对角化, 详见任何一本数学专业的线性代数教材⁹, 或者去查**若尔当标准形**理论, 这里不深入讨论.

但至少我们知道:

性质 1.6. 如果线性算子 Ω 的特征方程(1.35)无重根, 即 Ω 具有 n 个两两互异的**本征值**, 则 Ω 可对角化.

证明. 这个证明很简单, 留给读者. □

注意这只是可对角化的一个充分条件, 而非必要条件. 恒等算子 I 显然可对角化, 但它的特征方程为 $(1 - \omega)^n = 0$, $\omega = 1$ 是它的 n 重根.

一般来说, 随便的一个 n 次多项式“基本上都”没有重根, 有重根的多项式“非常少乃至可以忽略不计”. 若线性算子 Ω 的特征多项式有重根, 则称 Ω 是**退化的**或者**共振的**, 或者**简并的**.¹⁰

厄米特算子与酉算子的对角化.

在量子力学中我们总是考虑内积空间, 并且特别关心厄米特算子与酉算子. 我们不加证明地叙述几个对量子力学有用的基本定理, 这些定理的证明见任何一本线性代数或者矩阵论教材.

性质 1.7. 设 H 是 n 维内积空间 \mathbb{V} 上的厄米特算子, 则:

1. H 的本征值都是实数.
2. 存在 \mathbb{V} 的一组**么正基** $\{|1\rangle, \dots, |n\rangle\}$, 使得 H 在该基下的矩阵是对角阵, 且该矩阵对角元是 H 的全部本征值. 特别地, H 可对角化.
3. 若 $|V\rangle, |W\rangle$ 是属于 H 的不同本征值的本征矢量, 则 $\langle V|W\rangle = 0$.

⁸数学家鄙视物理学家的段子: 你们搞物理的见到矩阵就想对角化, 这不体面.

⁹别看工科的线性代数教材.

¹⁰这都是物理学家喜爱的叫法.

对于反厄米特算子, 完全类似. 只需注意: 若 Ω 是反厄米特算子, 则 $i\Omega$ 是厄米特算子 [这里的 i 是虚数单位], 然后对 $i\Omega$ 使用上述定理.

而对于酉算子, 也有类似结果:

性质 1.8. 设 U 是 n 维内积空间 \mathbb{V} 上的酉算子, 则:

1. U 的本征值都是模长为 1 的复数, 即形如 $e^{i\theta}$.
2. 存在 \mathbb{V} 的一组幺正基 $\{|1\rangle, \dots, |n\rangle\}$, 使得 U 在该基下的矩阵是对角阵, 且该矩阵对角元是 U 的全部本征值. 特别地, U 可对角化.
3. 若 $|V\rangle, |W\rangle$ 是属于 U 的不同本征值的本征矢量, 则 $\langle V|W\rangle = 0$.

此外, 下述“同时对角化定理”也格外重要:

性质 1.9. 设 H_1, H_2 是 n 维内积空间 \mathbb{V} 上的两个厄米特算子. 如果 H_1 与 H_2 对易, 即

$$[H_1, H_2] = 0,$$

则存在 \mathbb{V} 的幺正基 $\{|1\rangle, \dots, |n\rangle\}$, 使得 H_1, H_2 在这组基下的矩阵都是对角阵, 即 H_1, H_2 能够“同时对角化”.

我们以后会知道这个数学定理的量子物理学意义.

例子: 二体耦合振动问题与传播子.

我们接下来介绍经典力学中的一个典型例子, 来体会目前所介绍的线性代数知识的应用, 并从中感受一些量子力学的大概.

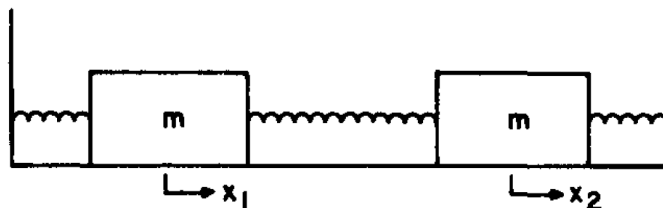


图: 二体耦合振动问题

如图, 水平面上两个滑块 m_1, m_2 的质量都为 m , 它们通过弹簧连接, 弹簧的弹性系数均为 k . 不计一切摩擦. 记 x_i 为滑块 m_i 距离其平衡位置的距离, 规定向右为正方向, $i = 1, 2$. 那么 $x_i = x_i(t)$ 是关于时间 t 的函数. 设两个滑块在初始时刻 $t = 0$ 的位置分别为 $x_1(0)$ 与 $x_2(0)$, 不妨两个滑块的初速度都为 0. 试求两个滑块的运动, 即函数 $x_1(t), x_2(t)$ 的表达式.

我们来求解这个问题. 由中学物理知识 (牛顿第二定律), 容易列出如下方程:

$$\ddot{x}_1 = -\frac{2k}{m}x_1 + \frac{k}{m}x_2 \quad (1.36)$$

$$\ddot{x}_2 = \frac{k}{m}x_1 - \frac{2k}{m}x_2 \quad (1.37)$$

这里沿用牛顿的记号, \ddot{x}_i 表示 $x_i(t)$ 关于时间 t 的二阶导数, 即滑块 m_i 的**加速度**. 我们要做的事情其实就是把上面那个微分方程组解出来.

我们观察一下这个方程组. 就拿(1.36)来说, 等号左边是 x_1 的二阶导数, 而等号右边既有 x_1 又有 x_2 , 也就是说 m_1 的加速度会受滑块 m_2 位置的影响. 同样, m_2 的加速度也会受 m_1 影响. 总之, 这两个滑块互相影响对方, 情况比较复杂. 物理学家或者工程师把这种“多个因素互相影响”的现象称为**耦合**.

耦合的情形往往比较麻烦. 我们更希望这样: \ddot{x}_1 只受 x_1 影响而不受 x_2 影响, 同样 \ddot{x}_2 只受 x_2 影响而不受 x_1 影响, 也就是说两个滑块“互不影响”, 这种“非耦合”的情形更方便处理. 事实上, 我们企图把这个耦合问题转化成非耦合情形.

我们把滑块的运动方程改写为矩阵形式:

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{k}{m} \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

注意矩阵 $\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} := \frac{k}{m} \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ 是厄米特矩阵. 我们记

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

则 $\{|1\rangle, |2\rangle\}$ 构成平面上的一组么正基. 注意

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

于是

$$|x\rangle = x_1|1\rangle + x_2|2\rangle,$$

这里的 $|x\rangle = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ 是关于时间 t 的函数, 也是我们需要求解的未知函数. $|x\rangle$ 所满足方程自然改写为

$$|\ddot{x}\rangle = \Omega|x\rangle. \quad (1.38)$$

将这个耦合问题转化为非耦合问题的关键在于把 Ω 对角化. 取

$$|1'\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |2'\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

则容易验证 $\{|1'\rangle, |2'\rangle\}$ 也是一组么正基, 并且

$$\begin{aligned} \Omega|1'\rangle &= -\omega_1^2|1'\rangle \\ \Omega|2'\rangle &= -\omega_2^2|2'\rangle \end{aligned}$$

其中

$$\omega_1 := \sqrt{\frac{k}{m}}, \quad \omega_2 := \sqrt{\frac{3k}{m}}. \quad (1.39)$$

[至于 $|1'\rangle$, ω_1 是怎么找到的, 这其实就是求解矩阵 Ω 的本征值与本征矢量. 具体计算方法见前文, 或见任何一本线性代数教材. 总之计算细节留给读者.]

考虑 $|x(t)\rangle$ 在新的基 $\{|1'\rangle, |2'\rangle\}$ 下的分解:

$$|x(t)\rangle = x_{1'}(t)|1'\rangle + x_{2'}(t)|2'\rangle,$$

这里的 $x_{1'}, x_{2'}$ 是新的未知函数, 它与原来的 x_1, x_2 有联系 [同一个矢量在不同基下的表示而已, 回忆坐标变换公式(1.30)]. 于是, 我们就把问题“求解 x_1, x_2 ”转化为“求解 $x_{1'}, x_{2'}$ ”. 而

$$|\ddot{x}\rangle = \Omega|x\rangle = \Omega(x_{1'}|1'\rangle + x_{2'}|2'\rangle) = -\omega_1^2 x_{1'}|1'\rangle - \omega_2^2 x_{2'}|2'\rangle,$$

于是有

$$\ddot{x}_{1'} = -\omega_1^2 x_{1'} \quad (1.40)$$

$$\ddot{x}_{2'} = -\omega_2^2 x_{2'}. \quad (1.41)$$

这就是我们想要的“非耦合”情形. 由微分方程知识, 轻易解出

$$\begin{aligned}x_{1'}(t) &= C_{11} \cos(\omega_1 t) + C_{12} \sin(\omega_1 t) \\x_{2'}(t) &= C_{21} \cos(\omega_2 t) + C_{22} \sin(\omega_2 t)\end{aligned}$$

其中 C_{11}, \dots, C_{22} 是待定的常数. 之后由 $x_{1'}, x_{2'}$ 与 x_1, x_2 之间的转换关系, 就可得到 $x_1(t), x_2(t)$ 的表达式. 最后再由已知的初始位置与初速度为 0 的条件, 就能将常数 C_{11}, \dots, C_{22} 确定下来.

具体计算细节留给读者, 最终我们解得

$$\begin{aligned}x_1(t) &= \frac{x_1(0) + x_2(0)}{2} \cos(\omega_1 t) + \frac{x_1(0) - x_2(0)}{2} \cos(\omega_2 t) \\x_2(t) &= \frac{x_1(0) + x_2(0)}{2} \cos(\omega_1 t) - \frac{x_1(0) - x_2(0)}{2} \cos(\omega_2 t)\end{aligned}$$

其中 ω_1, ω_2 见(1.39).

这个题目已经做出来了, 但还远远没完, 有很多东西值得回味. 首先, 我们用**对角化**的技巧讲耦合问题化为非耦合问题, 其实是对角化技巧的常见应用, 也是物理学家, 工程师们喜欢对角阵的原因之一.

但真正重要的是下面的内容: 注意我们最终解得的 x_1, x_2 能够写成矩阵形式

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \frac{\cos(\omega_1 t) + \cos(\omega_2 t)}{2} & \frac{\cos(\omega_1 t) - \cos(\omega_2 t)}{2} \\ \frac{\cos(\omega_1 t) - \cos(\omega_2 t)}{2} & \frac{\cos(\omega_1 t) + \cos(\omega_2 t)}{2} \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} \quad (1.42)$$

一般地, 一个物理系统的状态可以用矢量 $|x\rangle$ 来表示, $|x\rangle = |x(t)\rangle$ 随时间变化, 且随时间变化的规律形如

$$|x(t)\rangle = U(t)|x(0)\rangle. \quad (1.43)$$

算子 $U(t)$ 与时间有关, 而与系统的初始状态 $|x(0)\rangle$ 无关. 算子 $U(t)$ 称为该系统的**传播子**(propagator). 在二体耦合振动问题中, 传播子 $U(t)$ 在 $\{|1\rangle, |2\rangle\}$ 下的矩阵为

$$\begin{pmatrix} \frac{\cos(\omega_1 t) + \cos(\omega_2 t)}{2} & \frac{\cos(\omega_1 t) - \cos(\omega_2 t)}{2} \\ \frac{\cos(\omega_1 t) - \cos(\omega_2 t)}{2} & \frac{\cos(\omega_1 t) + \cos(\omega_2 t)}{2} \end{pmatrix}.$$

留给读者验证, 在此例中, 传播子具有紧凑的表达式

$$\begin{aligned} U(t) &= |1'\rangle\langle 1'| \cos(\omega_1 t) + |2'\rangle\langle 2'| \cos(\omega_2 t) \\ &= \sum_{i=1}^2 |i'\rangle\langle i'| \cos(\omega_i t). \end{aligned} \quad (1.44)$$

换言之, 传播子 $U(t)$ 在基 $\{|1'\rangle, |2'\rangle\}$ 下的矩阵为

$$\begin{pmatrix} \cos(\omega_1 t) & \\ & \cos(\omega_2 t) \end{pmatrix}.$$

此外还注意, 如果我们把滑块初始状态 $|x(0)\rangle$ 取为 $|i'\rangle, (i = 1, 2)$, 则

$$|x(t)\rangle = U(t)|x(0)\rangle = \sum_{j=1}^2 \cos(\omega_j t) |j'\rangle \langle j'|i'\rangle = |i'\rangle \cos(\omega_i t).$$

也就是说此时系统作频率为 ω_i 的振动. 可见 $|1'\rangle, |2'\rangle$ 是两种特殊的初始状态: 若初始状态为 $|i'\rangle$, 则系统作频率为 ω_i 的振动. 一般地, 这个系统的演化 $|x(t)\rangle$ 是频率为 ω_1, ω_2 的这两种振动模式的叠加.

注记

而量子力学大概在搞这种事情^a: 某个“量子系统”的“状态”= 某个内积空间 \mathbb{V} 中的矢量 $|\psi(t)\rangle$, $|\psi(t)\rangle$ 与时间 t 有关, 满足如下薛定谔方程

$$i\hbar \dot{|\psi\rangle} = H|\psi\rangle,$$

其中 \hbar 是所谓普朗克常数; 而 H 是某个不随时间变化的厄米特算子, 称为这个量子系统的哈密顿量.

^a量子力学的基本框架的确差不多就这样.

上述薛定谔方程的求解方式与我们刚才介绍的“两个滑块”完全类似. 性质1.7保证了 H 总可对角化. 取 \mathbb{V} 的一组幺正基 $\{|1\rangle, \dots, |n\rangle\}$, 使得 $H|i\rangle = \lambda_i|i\rangle$, 其中 $\lambda_1, \dots, \lambda_n$ 是 H 的全部本征值. 则 $|\psi\rangle$ 在这组幺正基下的分量会“好看些”. 最终我们求解薛定谔方程, 解出来的 $|\psi(t)\rangle$ 也形如

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle,$$

$U(t)$ 是所谓“传播子”. 在量子力学中, 传播子是**酉算子**. 如果我们把初始状态 $|\psi(0)\rangle$ 取为某个 $|i\rangle$, 则 $|\psi(t)\rangle$ 的表达式也会简洁一些. 在量子力学中, 量子态 $|1\rangle, \dots, |n\rangle$ 称为哈密顿量 H 的**本征态**.

至此, 我们对量子力学已经有了大致了解. 量子力学并不神秘.¹¹

1.5 指数映射与算子求导

我们知道, 标量与线性算子都可作用于矢量. 对于标量 x , 我们能谈论关于 x 的函数, 例如 $e^x, \sin x$ 等. 而对于线性算子 Ω , 其实也有办法谈论诸如 $\sin \Omega, e^\Omega$ 等关于 Ω 的函数.

线性算子的指数映射.

我们考虑幂级数

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

其中 a_k 为常数. 常见的“好的”函数 (**解析函数**) 都可通过**泰勒展开**来表示成这样的幂级数.¹²

而对于线性算子 Ω , 我们定义

$$f(\Omega) := \sum_{k=0}^{\infty} a_k \Omega^k.$$

一般来说, 需要考虑上述级数的收敛性. 不过, 我们目前只关心一类特殊的幂级数:

$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!}, \text{ 以及相应的}$$

$$e^\Omega := \sum_{k=0}^{\infty} \frac{\Omega^k}{k!}. \quad (1.45)$$

数学上可以证明, 对任何 Ω , 幂级数 e^Ω 都收敛, 从而 e^Ω 良定. 于是, 我们输入一个线性算子 Ω , 通过上述方式输出了另一个线性算子 e^Ω . 这个映射记作 \exp , 即 $\exp: \Omega \mapsto e^\Omega$. 映射 \exp 称为线性算子的**指数映射**.

¹¹其实挺神秘的, 这里是作者的无知者无畏.

¹²数学家对物理学家的三大刻板印象: 见到矩阵就对角化, 见到函数就幂级数展开, 见到两重极限就随便换序. 不过, 物理上常见的函数往往性质都很好, 还真能幂级数展开.

如何具体计算 e^Ω ? 在量子力学中, 被作指数映射的往往是 (反) 厄米特算子. 对于厄米特算子 Ω , 性质 1.7 表明 Ω 可对角化, 于是能够选取 \mathbb{V} 的一组幺正基 $\{|1\rangle, \dots, |n\rangle\}$, 使得 Ω 在该基下的矩阵是对角阵: ¹³

$$\Omega = \begin{pmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{pmatrix},$$

于是对任意 $k \geq 0$,

$$\Omega^k = \begin{pmatrix} \omega_1^k & & \\ & \ddots & \\ & & \omega_n^k \end{pmatrix},$$

因此在这组幺正基下,

$$e^\Omega = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\omega_1^k}{k!} & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty} \frac{\omega_n^k}{k!} \end{pmatrix} = \begin{pmatrix} e^{\omega_1} & & \\ & \ddots & \\ & & e^{\omega_n} \end{pmatrix}. \quad (1.46)$$

也就是说, 每个对角元都分别取指数.

容易验证, 指数映射有如下简单性质:

性质 1.10. 设 Ω 为线性算子, 则

1. e^Ω 可逆, 且 $(e^\Omega)^{-1} = e^{-\Omega}$.
2. $(e^\Omega)^\dagger = e^{\Omega^\dagger}$.
3. 若 Λ 是可逆的线性算子, 则 $\Lambda^{-1}e^\Omega\Lambda = e^{\Lambda^{-1}\Omega\Lambda}$.
4. 若 Ω 是厄米特算子, 则 $e^{i\Omega}$ 是酉算子, 并且 $\det(e^{i\Omega}) = e^{i\text{tr}\Omega}$.

证明. 证明留作习题, 都很简单. □

¹³在靠上下文不产生歧义的情况下, 我们为了省事, 以后把线性算子与它在某组幺正基下的矩阵等同.

此外, 还有如下重要性质:

性质 1.11. 设 Ω, Λ 为线性算子, 且 $[\Omega, \Lambda] = 0$, 则

$$e^{\Omega+\Lambda} = e^{\Omega}e^{\Lambda} = e^{\Lambda}e^{\Omega}.$$

证明也不难, 留给读者. 但要特别注意, 条件 $[\Omega, \Lambda] = 0$ 一般不能去掉. 一般来说, $e^{\Omega}e^{\Lambda}$, $e^{\Lambda}e^{\Omega}$ 与 $e^{\Lambda+\Omega}$ 往往互不相等.

线性算子的求导.

我们考虑线性算子值函数 $\Theta(t)$, 即: 输入标量 t , 输出线性算子 $\Theta(t)$. 与通常的函数求导定义类似, 我们定义

$$\frac{d\Theta(t)}{dt} := \lim_{\Delta t \rightarrow 0} \frac{\Theta(t + \Delta t) - \Theta(t)}{\Delta t}.$$

考虑算子 $\Theta(t)$ 在某组基下的矩阵, 则矩阵元 $\Theta_{ij}(t)$ 是关于 t 的函数. 显然 $\Theta(t)$ 的导数的矩阵元恰为相应矩阵元的导数.

我们特别关心一类线性算子值函数. 取定线性算子 Ω , 考虑

$$\Theta(t) = e^{t\Omega}$$

容易验证

$$\frac{d}{dt}e^{t\Omega} = \Omega e^{t\Omega} = e^{t\Omega}\Omega,$$

即 $\Theta(t)$ 满足微分方程

$$\frac{d}{dt}\Theta(t) = \Omega\Theta(t) = \Theta(t)\Omega.$$

设 $|V(t)\rangle$ 是关于 t 的矢量值函数, 且满足微分方程

$$\frac{d}{dt}|V(t)\rangle = \Omega|V(t)\rangle,$$

则容易验证

$$|V(t)\rangle = e^{t\Omega}|V(0)\rangle \tag{1.47}$$

是此方程的一个解.¹⁴

¹⁴上一节末的某注记悄悄提了一下薛定谔方程. 而这里悄悄告诉大家如何解这种薛定谔方程.

1.6 狄拉克 δ -函数

仅仅有限维空间的线性代数明显不够用，量子力学当中的内积空间往往是无穷维的。研究无穷维矢量空间的数学学科分支是所谓**泛函分析**，这也是量子力学的重要数学基础。

我们这里不打算严格，深入地讲太多数学，而是以物理学家的视角，“不严谨地”介绍一下无穷维空间的基本想法。

内积空间 $\mathbb{V}_{[0,L]}$.

无穷维矢量空间的典型例子是闭区间 $[0, L]$ 上的全体**连续函数**构成的空间 [物理上方便起见，我们只考虑连续的函数]，在本节我们把这个矢量空间暂时记作 $\mathbb{V}_{[0,L]}$. 设 $f(x)$ 是一个定义在 $[0, L]$ 上的连续函数，则 $|f\rangle$ 是 $\mathbb{V}_{[0,L]}$ 中的一个矢量. 注意这里的矢量空间，矢量是抽象的概念¹⁵，我们希望把它们类比于“具体的向量”. 一个自然的想法是：取一个（很大的）正整数 n ，把区间 $[0, L]$ 作 n 等分，分割点 $x_k = \frac{k}{n}L$, $1 \leq k \leq n$. 考虑 n 维列向量

$$|f_n\rangle = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix},$$

则 $|f_n\rangle$ 可以认为是 $|f\rangle$ 的某种“离散化”. 凭感觉，当 n 越大， n 维列向量 $|f_n\rangle$ 就越能“近似代替” $|f\rangle$ ，即 “ $\lim_{n \rightarrow +\infty} |f_n\rangle = |f\rangle$ ”.

现在取定 n ，我们研究 $|f_n\rangle$. $|f_n\rangle$ 所在的 n 维矢量空间具有标准基

$$|x_i\rangle = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{第 } i \text{ 分量为 } 1), \quad 1 \leq i \leq n,$$

¹⁵对，这是“抽象的鸭子”，而不是“具体的鸭子”. 试回忆：**什么是鸭子**？若是想不起来，就往前翻，找找脚注.

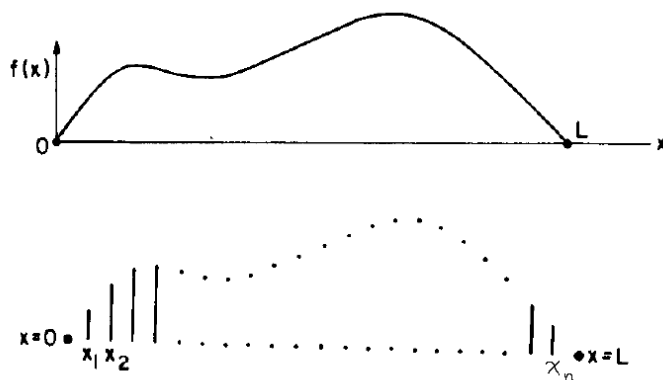


示意图: $|f\rangle$ 与 $|f_n\rangle$.

在此基下,

$$|f_n\rangle = \sum_{i=1}^n f(x_i)|x_i\rangle.$$

注意 n 维矢量空间上有标准内积, 就是我们通常熟悉的那种内积:

$$\langle x_i | x_j \rangle = \delta_{ij}.$$

于是, $\mathbb{V}_{[0,L]}$ 中的任意两个矢量 $|f\rangle, |g\rangle$, 考虑它们的“ n 维近似” $|f_n\rangle, |g_n\rangle$, 有

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f^*(x_i)g(x_i).$$

特别地, 如果取 $f(x) = g(x) \equiv 1$ 是恒为 1 的常函数, 则容易验证 $\langle f_n | g_n \rangle = n$.

我们认为“ $\lim_{n \rightarrow \infty} |f_n\rangle = |f\rangle$ ”, 从而自然希望“ $\lim_{n \rightarrow \infty} \langle f_n | g_n \rangle = \langle f | g \rangle$ ”, 从而引入无穷维空间 $\mathbb{V}_{[0,L]}$ 上的内积. 但是, 刚才说过, 若 $f(x) = g(x) \equiv 1$, 则 $\langle f_n | g_n \rangle = n$, 此时令 $n \rightarrow \infty$ 就发散了. 这种问题普遍存在. 解决这种问题的方法是适当修改 $\langle f_n | g_n \rangle$ 的定义: 我们令

$$\langle f_n | g_n \rangle := \sum_{i=1}^n f^*(x_i)g(x_i)\Delta_n,$$

其中 $\Delta_n := \frac{L}{n}$ 为 n 等分后的每个小区间的长度. 此时令 $n \rightarrow \infty$, 由定积分的定义, 可知

$$\lim_{n \rightarrow \infty} \langle f_n | g_n \rangle = \lim_{n \rightarrow \infty} \sum_{i=1}^n f^*(x_i) g(x_i) \Delta_n = \int_0^L f^*(x) g(x) dx.$$

于是, 我们引入:

定义. 对于 $\mathbb{V}_{[0,L]}$ 中的矢量 $|f\rangle, |g\rangle$, 定义

$$\langle f | g \rangle := \int_0^L f^*(x) g(x) dx. \quad (1.48)$$

容易验证, 这的确是矢量空间 $\mathbb{V}_{[0,L]}$ 的一个内积, $\mathbb{V}_{[0,L]}$ 在此内积下构成内积空间.

狄拉克 δ -函数: 引入.

我们现在有了内积空间 $\mathbb{V}_{[0,L]}$, 于是自然要问这个空间的么正基是什么. 我们看看能不能类比有限维空间. n 维时, $|f_n\rangle$ 的第 i 个分量是函数 $f(x)$ 在区间 $[0, L]$ 的等分点 $x_i = \frac{i}{n}L$ 处的取值. 现在 $n \rightarrow \infty$, 自然会认为 $|f\rangle$ 有无穷多个分量: 对于区间 $[0, L]$ 中的每个点 x , 函数 f 在 x 处的取值 “应该是” $|f\rangle$ 的某个分量.

于是我们就猜测, 对每个 $x \in [0, L]$, 都会有某个矢量 $|x\rangle$, 使得集合 $\{|x\rangle | x \in [0, L]\}$ 构成 $\mathbb{V}_{[0,L]}$ 的一组么正基. 这要求:

$$\langle x | x' \rangle = 0, \text{ 若 } x \neq x'.$$

那我们是否也要求 $\langle x | x \rangle = 1$ 呢? 事实上我们不可贸然这样. 对于矢量 $|f\rangle$, f 在 $x \in [0, L]$ 处的取值被认为是 $|f\rangle$ 在 $|x\rangle$ 处的分量, 类比有限维的有关公式:

$$f(x) = \langle x | f \rangle.$$

此外, 我们希望(1.24)在无穷维也成立:

$$\int_0^L |x'\rangle \langle x'| dx' = I.$$

于是, 我们就有:

$$f(x) = \langle x | f \rangle = \langle x | I | f \rangle = \int_0^L \langle x | x' \rangle \langle x' | f \rangle dx' = \int_0^L \langle x | x' \rangle f(x') dx'. \quad (1.49)$$

从这里, 我们看看 $\langle x|x' \rangle$ 到底是什么东西. 我们说过, 当 $x' \neq x$ 时, 正交性表明 $\langle x|x' \rangle = 0$. 所以(1.49)最右边的关于 x' 的被积函数当中的 $\langle x|x' \rangle$ “几乎处处为零”? 所以等号右边的积分为零? 但它明明是 $f(x)$, 而 f 与 x 都是任意的.

这是非常诡异. 固定 $x \in [0, L]$, 对任意给定 (充分小) 的正数 $\varepsilon > 0$, 我们考虑函数

$$f(x') = \begin{cases} 1 & |x' - x| \leq \varepsilon \\ 0 & \text{其余情况} \end{cases},$$

[这个 f 不是连续函数, 但这其实无妨, 我们不妨用某一族连续函数无限逼近它], 将此 $|f\rangle$ 代入(1.49), 则得到

$$\int_{x-\varepsilon}^{x+\varepsilon} \langle x|x' \rangle dx' = 1. \quad (1.50)$$

注意上式对任意 (充分小的) $\varepsilon > 0$ 都成立. 这明显不是普通的函数! 当 $x' \neq x$ 时 $\langle x|x' \rangle = 0$, 所以 $x' = x$ 时 $\langle x|x' \rangle$ 到底等于几? 为了使得(1.50)成立, 一个恐怖而邪恶的想法是:

$$\langle x|x' \rangle = \begin{cases} 0 & x' \neq x \\ +\infty & x' = x \end{cases}$$

可见, 这并不是通常意义下的函数; $x' = x$ 处的取值为无穷大, 而无穷大是多大呢: 它被下式所确定

$$\int_0^L \langle x|x' \rangle dx' = 1.$$

$\langle x|x' \rangle$ 鬼魅如同幽灵般的行为曾经令数学家们迷惑不解, 长得像函数但又不是函数.¹⁶ 我们更习惯把这个 $\langle x|x' \rangle$ 记作 $\delta(x - x')$. 这个 $\delta(x)$ 称为**狄拉克 delta 函数**(Dirac delta function), 它由著名的天才量子物理学家**狄拉克**引入.

$\delta(x)$ 并非通常的函数, 我们这里不打算介绍它严格的数学定义, 只需要先掌握它的运算性质. 至于严格的数学定义, 那是数学家们多年之后才做的事情. 在数学上, 狄拉克 δ -函数是一种**分布** (也叫**广义函数**). 有关数学理论详见任何一本进阶水平的实分析或泛函分析教材.

¹⁶嘴巴像鸭子, 羽毛像鸭子, 但是游泳姿势不像鸭子. 这种酷似鸭子的妖怪叫做 “变异的鸭子”.

狄拉克 δ -函数: 运算性质.

现在, 我们已引入狄拉克 δ -函数 $\delta(x)$. 由(1.49), 它满足: 对任意函数 f ,

$$\int \delta(x - x') f(x') dx' = f(x). \quad (1.51)$$

[为了方便, 我们省略积分上下限.] 注意, 上式是 $\delta(x - x')$ 真正的定义.

狄拉克 δ -函数并不是通常的函数, 但它能够被通常的函数逼近. 数学上常用的方法是考虑如下高斯函数:

$$g_{\Delta}(x - x') = \frac{1}{\Delta\sqrt{\pi}} e^{-\frac{(x-x')^2}{\Delta^2}},$$

其中 $\Delta > 0$ 为某个参数. 这个函数在概率论中被大家所熟知 (所谓正态分布), 在古典物理学里也众所周知 (例如热传导方程的解, 热核).

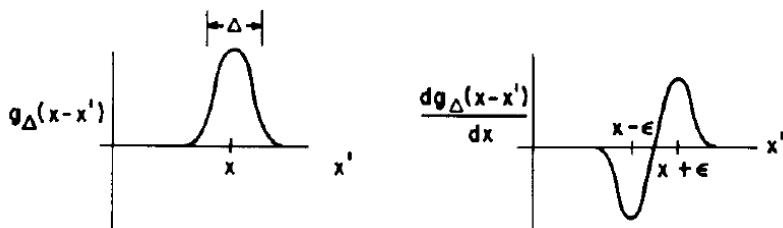


图: $g_{\Delta}(x - x')$ 及其导函数的大致图像.

容易验证 $\int g_{\Delta}(x - x') dx' \equiv 1$. 注意当 $\Delta \rightarrow 0$ 时, 函数 g_{Δ} 的图像越来越“瘦高”, 函数 g_{Δ} 的值越来越“向 x 点处集中”. 事实上, 在某种意义上完全可以认为 $\delta(x - x') = \lim_{\Delta \rightarrow 0} g_{\Delta}(x - x')$. 这样就用通常的函数来“逼近”狄拉克 δ -函数.

狄拉克 δ -函数是偶函数:

$$\delta(x - x') = \delta(x' - x). \quad (1.52)$$

这是因为对任意 $f(x)$, 记 $g(x) = f(-x)$, 则由定义式(1.51), 利用换元积分 $x' \mapsto -y$, 有

$$\int \delta(x' - x) f(x') dx'$$

$$\begin{aligned}
 &= \int \delta((-x) - y) f(-y) dy \\
 &= \int \delta((-x) - y) g(y) dy = g(-x) = f(x)
 \end{aligned}$$

[第一步换元, 也许有人认为少写了负号, 实则不然. 我们省略积分上下限, 但注意: 我们总约定积分下限 \leq 积分上限.] 因此, 对任意函数 f 都有

$$\int \delta(x - x') f(x') dx' = \int \delta(x' - x) f(x') dx',$$

因此 $\delta(x - x') = \delta(x' - x)$. 当然, 我们也可这样理解: 显然 $g_{\Delta}(x - x') = g_{\Delta}(x' - x)$, 令 $\Delta \rightarrow 0$ 即可.

我们甚至还能谈论狄拉克 δ -函数的导函数:

$$\delta'(x - x') = \frac{d}{dx} \delta(x - x') = -\frac{d}{dx'} \delta(x - x').$$

它可以这样引入: 先考虑 $g_{\Delta}(x - x')$ 的导函数, 然后令 $\Delta \rightarrow 0$,

$$\frac{d}{dx} \delta(x - x') \text{“}=\text{”} \lim_{\Delta \rightarrow 0} \frac{d}{dx} g_{\Delta}(x - x')$$

g_{Δ} 的导函数图像如图, Δ 越接近 0, 其图像越“瘦高”, 且在点 x 附近的“正负震荡越剧烈”. 令 $\Delta \rightarrow 0$, 可见 $\delta'(x - x')$ 在 $x' = x$ 处的“取值”实在一言难尽. 对于“广义函数”, 一般来说不能像通常函数那样谈论在某点处的取值. 但我们可以这样来理解 $\delta'(x - x')$: 对任意函数 f ,

$$\begin{aligned}
 \int \delta'(x - x') f(x') dx' &= \lim_{\varepsilon \rightarrow 0} \int \frac{\delta(x - x' + \varepsilon) - \delta(x - x')}{\varepsilon} f(x') dx' \\
 &= \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon) - f(x)}{\varepsilon} = f'(x).
 \end{aligned}$$

或者用分部积分法:

$$\begin{aligned}
 \int \delta'(x - x') f(x') dx' &= \int f(x') d\delta(x - x') \\
 &= f(x') \delta(x - x') \Big|_0^L + \int \delta(x - x') f'(x') dx' = f'(x).
 \end{aligned}$$

[我们假装 $\delta(x - x')$ 在除了 $x' = x$ 之外的点处都等于 0.] 总之, 我们引入:

定义. 狄拉克 δ -函数的导函数 $\delta'(x - x')$ 是指满足如下性质的广义函数:

$$\int \delta'(x - x') f(x') dx' = f'(x). \quad (1.53)$$

对任意函数 f 都成立.

我们至今没有给出, 也不想在这给出“广义函数”的严格定义. 但至少要知道, 我们不能像描述通常的函数那样, 通过确定函数在定义域内每一点处的取值来描述广义函数. 广义函数的“功能”是作用在通常的函数上. 确定一个广义函数, 只需确定它每个通常的函数的作用, 就像(1.51)与(1.53)那样.

为理解上面这段话, 我们看一个简单性质:

性质 1.12. 设 a 为非零实数, 则

$$\delta(ax) = \frac{\delta(x)}{|a|}.$$

证明. 由于 $\delta(x)$ 是偶函数, 不妨 $a > 0$. 从而对任意函数 f , 注意换元积分, 有

$$\begin{aligned} \int \delta(ax) f(x) dx &= \frac{1}{a} \int \delta(y) f\left(\frac{y}{a}\right) dy = \frac{1}{a} f(0) \\ &= \frac{1}{a} \int \delta(x) f(x) dx, \end{aligned}$$

因此 $\delta(ax) = \frac{\delta(x)}{a}$. 而 $a < 0$ 时类似. 从而证毕. \square

我们当然还能谈论狄拉克 δ -函数的高阶导数

$$\delta^{(n)}(x - x') = \frac{d^n}{dx^n} \delta(x - x') = (-1)^n \frac{d^n}{d(x')^n} \delta(x - x'),$$

$\delta^{(n)}(x - x')$ 由下式所确定:

$$\int \delta^{(n)}(x - x') f(x') dx' = f^{(n)}(x). \quad (1.54)$$

最后, 狄拉克 δ -函数也能作傅立叶变换. 现在我们把函数定义域 $[0, L]$ 换成 $\mathbb{R} = (-\infty, \infty)$. 我们回忆, 对于通常的“好的”函数 $f(x)$, f 的傅立叶变换与傅立

叶逆变换 分别为

$$\mathcal{F}[f](k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx} f(x) dx \quad (1.55)$$

$$\mathcal{F}^{-1}[f](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx} f(k) dk. \quad (1.56)$$

并且傅立叶变换与傅立叶逆变换满足: $\mathcal{F}^{-1}[\mathcal{F}[f]] = f$. 事实上, 上述傅立叶变换公式也适用于广义函数. 容易验证狄拉克 δ -函数的傅立叶变换为

$$\mathcal{F}[\delta](k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx} \delta(x) dx = \frac{1}{\sqrt{2\pi}}.$$

即, $\delta(x)$ 的傅立叶变换是取值恒为 $\frac{1}{\sqrt{2\pi}}$ 的常函数. 我们再作傅立叶逆变换, 就容易得到 δ -函数的表达式

$$\delta(x' - x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x' - x)} dk. \quad (1.57)$$

上式右边也不是通常高等数学教材里的反常积分. 注意当 $x = x'$ 时等号右边 $= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk = +\infty$ 积分发散; 而 $x \neq x'$ 时, 上式右边的反常积分依然发散, 被积函数是震荡的.¹⁷

1.7 无穷维空间中的线性算子

我们考虑定义在 $\mathbb{R} = (-\infty, +\infty)$ 上的所有“好的”函数构成的矢量空间 $\mathbb{V}_{\mathbb{R}}$. 至于什么样的函数是“好的”, 我们到后面再说. 比如, 我们希望函数是连续甚至可导. 定义域 \mathbb{R} 上的每个点 x 都对应于某个“矢量” $|x\rangle$ [这仅仅是个形式上的记号, 并非真的是 $\mathbb{V}_{\mathbb{R}}$ 中的矢量], 所有的这些 $|x\rangle$ 构成 $\mathbb{V}_{\mathbb{R}}$ 的一组标准正交基:

$$\langle x|x'\rangle = \delta(x - x').$$

对于每个函数 $|f\rangle$, $|f\rangle$ 在基矢量 $|x\rangle$ 下的分量

$$\langle x|f\rangle = f(x),$$

¹⁷总之, 这个式子等号两边都是诡异的东西, 这两个诡异的东西居然还相等. 我们这里的数学推导不严谨, 但应该足够“让读者相信是正确的”. 事实上我们这里介绍的内容都有严格的数学理论保证, 数学上能说清楚, 但限于篇幅我们不方便细说. 读者完全可以放心接受这些.

$|x\rangle$ 在这组基下的展开式为

$$|f\rangle = \int_{-\infty}^{+\infty} f(x)|x\rangle dx = \int_{-\infty}^{+\infty} \langle x|f\rangle |x\rangle dx = \left(\int_{-\infty}^{+\infty} |x\rangle \langle x| dx \right) |f\rangle.$$

于是, 与有限维空间情形类似, 恒等算子 I 具有展开式

$$I = \int_{-\infty}^{+\infty} |x\rangle \langle x| dx.$$

微分算子 D 与 K .

我们当然可以考虑无穷维空间 $\mathbb{V}_{\mathbb{R}}$ 上的线性算子. 线性算子的定义对有限维空间, 无穷维空间都适用.

设 Ω 为线性算子. 在有限维空间的时候, 我们可以谈论算子 Ω 在某组么正基下的矩阵; 而在无穷维空间, 也可类似谈论 Ω 在基 $\{|x\rangle | x \in \mathbb{R}\}$ 下的“矩阵”. 其“矩阵元”

$$\Omega_{xx'} := \langle x|\Omega|x'\rangle$$

是关于 x, x' 的广义函数 [马上就看到例子].

函数空间 $\mathbb{V}_{\mathbb{R}}$ 上的一个典型的线性算子是**微分算子** $D = \frac{d}{dx}$, 它把函数 $f(x)$ 变成 f 的导函数 $f'(x)$, 即

$$D|f\rangle = |f'\rangle.$$

根据线性算子的定义 [再次强调, 线性算子的定义并没有对矢量空间的维数有要求, 于是同样适用于无穷维空间!] 容易验证 D 的确是线性算子. 我们来看看 D 在“标准基” $\{|x\rangle | x \in \mathbb{R}\}$ 下的矩阵元是什么, 也就是确定 $\langle x|D|x'\rangle$ 的取值. 但是, 基矢量 $|x\rangle$ 并非普通的函数, 仅仅是形式上的记号, 我们无法直接谈论 $D|x\rangle$. 事实上, 我们间接的看, 对任意函数 $f(x)$, 注意

$$\begin{aligned} f'(x) &= \langle x|f'\rangle = \langle x|D|f\rangle = \langle x|DI|f\rangle \\ &= \int_{-\infty}^{+\infty} \langle x|D|x'\rangle \langle x'|f\rangle dx' = \int_{-\infty}^{+\infty} D_{xx'} f(x') dx', \end{aligned}$$

把此式与(1.53)比较, 立刻得到

$$D_{xx'} = \langle x|D|x'\rangle = \delta'(x - x'). \quad (1.58)$$

微分算子 D 是厄米特算子吗? 在有限维, 我们知道, 算子 Ω 是厄米特算子当且仅当 Ω 在某组么正基下的矩阵是厄米特矩阵, 即矩阵元满足 $\Omega_{ij} = \Omega_{ji}^*$. 而无限维情形应该也类似. 考察微分算子 D 的矩阵元

$$D_{xx'} = \delta'(x - x') = -\delta'(x' - x) = -\delta'(x' - x)^* = -D_{x'x}^*$$

其中注意狄拉克 δ -函数的导函数是奇函数, 并且假装它的函数值都是实数¹⁸. 可见, 从矩阵元来看, D 更像是反厄米特算子. 我们记线性算子

$$K := -iD, \quad (1.59)$$

其中 i 是虚数单位, 则容易验证 $K_{xx'} = K_{x'x}^*$. 从而在矩阵上, K 看起来像厄米特算子. 我们希望它就是厄米特算子. 无穷维空间比较反直觉, 我们还是小心一些, 万无一失的方法是直接从原始定义出发. 为此, 需要证明: 对任意函数 f, g , 成立 $\langle f | K^\dagger | g \rangle = \langle f | K | g \rangle$. 而

$$\begin{aligned} \langle f | K^\dagger | g \rangle &= \langle Kf | g \rangle = \langle g | Kf \rangle^* = [-i \langle g | f' \rangle]^* \\ &= i \left(\int_{-\infty}^{+\infty} g^*(x) f'(x) dx \right)^* = i \int_{-\infty}^{+\infty} f'^*(x) g(x) dx, \\ \langle f | K | g \rangle &= -i \langle f | g' \rangle = -i \int_{-\infty}^{+\infty} f^*(x) g'(x) dx \\ &= -i \int_{-\infty}^{+\infty} f^*(x) dg(x) \\ &= -i f^*(x) g(x) \Big|_{-\infty}^{+\infty} + i \int_{-\infty}^{+\infty} f'^*(x) g(x) dx. \end{aligned}$$

可见, K 是厄米特算子当且仅当对 $\mathbb{V}_{\mathbb{R}}$ 中的任意函数 f, g 都成立

$$f^*(x) g(x) \Big|_{-\infty}^{+\infty} = 0, \quad (1.60)$$

而这一般来说未必成立. 解决此问题的办法是, 重新定义“好的函数”构成的空间 $\mathbb{V}_{\mathbb{R}}$, 说清楚什么是“好的函数”. 直接的办法是, 我们要求 $\mathbb{V}_{\mathbb{R}}$ 中的函数都满足

$$\lim_{|x| \rightarrow \infty} f(x) = 0,$$

¹⁸这一段推导“连猜带蒙”, 属于物理学家们的脑洞.

这样子, (1.60)总成立. 除了在无穷远处趋于 0 的函数是“好的函数”, 出于物理上的某些考虑, 我们还认为在无穷远处“均匀震荡”的函数

$$g(x) = e^{ikx}, \quad k \text{ 为某个非零实数}$$

也是好的函数. 虽然函数 $g(x) = e^{ikx}$ 在 $|x| \rightarrow \infty$ 的通常意义下的极限不存在, 但它的函数值总是在 0 周围震荡, “平均一下”就是 0. 于是我们认为在某种意义上 $\lim_{|x| \rightarrow 0} e^{ikx} = 0$. 在数学上, 我们其实就要把函数极限的概念推广: $f(x)$ 在 $+\infty$ 的“广义极限”是指 f 在离原点“充分远”的“充分长”的区间上的“平均值”的极限, 即

$$\lim_{x \rightarrow +\infty} f(x) := \lim_{\substack{L \rightarrow +\infty \\ \Delta \rightarrow +\infty}} \frac{1}{\Delta} \int_L^{L+\Delta} f(x) dx.$$

$x \rightarrow -\infty$ 的广义极限也类似定义. 在此意义下, 很多极限原本不存在的函数有了极限, 比如 e^{ikx} .

综上, 我们约定: 在 $|x| \rightarrow \infty$ 处的广义极限为 0 的可微函数是“好的函数”, $\mathbb{V}_{\mathbb{R}}$ 是全体“好的函数”构成的空间.

算子 K 的本征值问题与傅立叶变换.

我们来考虑厄米特算子 $K = -iD$ 的本征值问题. 设 k 是 K 的一个本征值, $|\psi_k\rangle$ 是相应的本征矢量, 则 $K|\psi_k\rangle = k|\psi_k\rangle$, 从而对任意 $x \in \mathbb{R}$, 都有 $\langle x|K|\psi_k\rangle = k\langle x|\psi_k\rangle = k\psi_k(x)$. 另一方面, $\langle x|K|\psi_k\rangle = -i\langle x|\psi_k'\rangle = -i\psi_k'(x)$, 从而本征矢量 $|\psi_k\rangle$ 满足

$$\psi_k'(x) = ik\psi_k(x),$$

容易解得

$$\psi_k(x) = Ce^{ikx}.$$

注意这里的本征值 k 是实数. 如果 k 的虚部非零, 那么 e^{ikx} 在 $|x| \rightarrow \infty$ 时无界, 甚至就连广义极限也不存在, 这不是“好的函数”.

综上所述, K 的本征值构成的集合为 \mathbb{R} ; 对每个实数 k , K 的属于本征值 k 的本征矢量 $|\psi_k\rangle$ 为

$$\psi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad (1.61)$$

这里我们把常数 C 特别取为 $\frac{1}{\sqrt{2\pi}}$. 于是, 根据(1.57), 有

$$\langle \psi_k | \psi_{k'} \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi_k^*(x) \psi_{k'}(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i(k'-k)x} dx = \delta(k - k').$$

可见, $\{|\psi_k\rangle | k \in \mathbb{R}\}$ 也构成 $\mathbb{V}_{\mathbb{R}}$ 的一组么正基. 这是除了我们已有的 $\{|x\rangle\}$ 之外的另一组么正基.

对于函数 $f(x)$, 自然就考虑 $|f\rangle$ 在基 $\{|\psi_k\rangle\}$ 下的分量是什么. 直接计算得

$$\langle \psi_k | f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx} f(x) dx = \mathcal{F}[f](k),$$

恰为 f 的傅立叶变换在 k 处的取值. 此外, 由于 $\{|\psi_k\rangle\}$ 也是么正基, 从而 $I = \int_{-\infty}^{+\infty} |\psi_k\rangle \langle \psi_k| dk = I$ 也成立, 从而

$$\begin{aligned} f(x) &= \langle x | f \rangle = \int_{-\infty}^{+\infty} \langle x | \psi_k \rangle \langle \psi_k | f \rangle dk \\ &= \int_{-\infty}^{+\infty} \psi_k(x) \mathcal{F}[f](k) dk \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx} \mathcal{F}[f](k) dk, \end{aligned}$$

恰为众所周知的傅立叶逆变换公式.

位置算子 X .

现在, 我们介绍了一个算子 K 与两组基 $\{|x\rangle\}, \{|\psi_k\rangle\}$. 算子 K 在这两组基下的矩阵元分别为

$$\begin{aligned} \langle x | K | x' \rangle &= -i\delta'(x - x') \\ \langle \psi_k | K | \psi_{k'} \rangle &= k' \langle \psi_k | \psi_{k'} \rangle = k' \delta(k - k'). \end{aligned}$$

现在, 引入算子 X , 使得

$$X|x\rangle = x|x\rangle. \quad (1.62)$$

则对于函数 $f(x)$, 有

$$(Xf)(x) = \langle x | Xf \rangle = \langle x | XI | f \rangle = \int_{-\infty}^{+\infty} \langle x | X | x' \rangle \langle x' | f \rangle dx'$$

$$\begin{aligned}
 &= \int_{-\infty}^{+\infty} x' \langle x|x' \rangle f(x') dx' = \int_{-\infty}^{+\infty} x' f(x') \delta(x-x') dx' \\
 &= x f(x).
 \end{aligned}$$

容易验证 X 是厄米特算子. 至此, 我们有两个算子 X, K 以及两组基 $\{|x\rangle\}, \{|\psi_k\rangle\}$. 读者试自行总结这两个算子分别在这两组基下的矩阵元. 总之:

$$\begin{aligned}
 X &: f(x) \mapsto x f(x) \\
 K &: f(x) \mapsto -i f'(x).
 \end{aligned}$$

性质 1.13.

$$[X, K] = iI.$$

证明. 对任意函数 $f(x)$, 只需注意

$$\begin{aligned}
 XKf(x) &= -iXf'(x) = -ixf'(x) \\
 KXf(x) &= K(xf(x)) = -i(xf(x)) = -if(x) - ix f'(x),
 \end{aligned}$$

从而

$$(XK - KX)f(x) = -if(x),$$

因此

$$[X, K] = -iI.$$

□

2 经典力学背景知识

经典力学博大精深, 我们都该会一点点.

— 沃·兹基·硕德

量子力学并非从天而降的神学, 其在物理上有众多实验现象支撑, 并且数学理论框架与经典力学一脉相承. 为理解量子力学, 我们必须先回顾经典力学.

众所周知, 经典力学起源于牛顿, 牛顿三大运动定律流传千古, 能解释并预测“几乎所有”的力学现象, 被写入中学物理课本. 牛顿的力学理论是经典力学的一部分, 称为**牛顿力学**.

然而, 经典力学 [也就是不考虑相对论与量子现象的“通常”力学] 有三套理论体系: 除了牛顿力学, 还有**拉格朗日力学**与**哈密顿力学**. 这三套理论的研究对象相同, 而对同样的力学现象有不同的解释, 它们各自的基本原理的数学表述也各不相同.¹⁹

经典力学与量子力学联系密切, 一脉相承. 量子力学中的**费曼路径积分理论**是拉格朗日力学的“量子版本”; 而薛定谔, 海森堡, 狄拉克, 波尔等人最初建立的“入门量子力学”则是哈密顿力学的“量子版本”. 从而可以说, 量子力学深受经典力学影响.

本章带大家快速回忆牛顿力学, 拉格朗日力学与哈密顿力学这三套经典力学理论的基本原理.

2.1 力学系统的自由度与广义坐标.

无论学习哪种力学, 首先要学会描述一个力学系统在某时刻的状态. 经典力学的三套理论对于如何描述力学系统的状态给出了不同回答.

¹⁹ 打一个非常不恰当的比方: 中医和西医的理论体系迥然不同, 各自有各自的术语, 思维方式; 但是, 只要能把病治好, 我们就认为它是对的.

第一个例子：单质点系统.

先考虑最简单的情形：3 维空间 \mathbb{R}^3 中的一个质点 m . 这个力学系统的“状态”无非就是指质点 m 的位置. 而质点 m 的位置可以用**矢量** $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ 来描述, 这里的 x, y, z 是质点 m 在某给定的空间直角坐标系下的各坐标分量. 换言之, 这个物理系统的状态被 x, y, z 这三个参数所唯一确定. 我们说这个力学系统随时间演化, 无非就是指这个质点 m 的位置随时间 t 变化, 即“运动”. 描述质点的运动的方法自然是把 x, y, z 视为关于时间 t 的函数, 即 $t \mapsto \mathbf{x}(t) := \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$. 函数 $\mathbf{x}(t)$ 自然给出空间 \mathbb{R}^3 中的一条曲线, 这恰为质点 m 的运动轨迹.

牛顿力学认为, 质点所受的**力** \mathbf{F} 是改变质点运动的原因, 具体地说, 成立

$$\mathbf{F} = m\ddot{\mathbf{x}}(t), \quad (2.1)$$

其中 $\ddot{\mathbf{x}}(t)$ 是函数 $\mathbf{x}(t)$ 关于时间 t 的二阶导数, 即质点 m 的**加速度**. 在本章, 我们用在函数名称上方的“点”来表示对时间 t 求导, 这沿用了牛顿当年的记号. (2.1)正是众所周知的**牛顿第二定律**. 利用牛顿第二定律, 只要我们知道质点的初始位置, 初始速度, 以及质点在任意时刻所受的力, 那么这个质点的运动就唯一确定下来了. 具体这样操作: 记初始时刻 t_0 的位置 $\mathbf{x}(t_0)$ 与速度 $\dot{\mathbf{x}}(t_0)$, 时刻 t 所受力记为 $\mathbf{F}(t)$. 则经过一段微小的时间间隔 Δt 后, 粒子的位置变成

$$\mathbf{x}(t_0 + \Delta t) \approx \mathbf{x}(t_0) + \dot{\mathbf{x}}(t_0)\Delta t$$

而由牛顿第二定律, 粒子在 $t_0 + \Delta t$ 时刻的速度变为

$$\dot{\mathbf{x}}(t_0 + \Delta t) \approx \dot{\mathbf{x}}(t_0) + \frac{\mathbf{F}(t_0)}{m}\Delta t.$$

之后可近似计算粒子在“下一个”时刻 $t_0 + 2\Delta t$ 的位置与速度. 不断这样做下去, 并令时间间隔 $\Delta t \rightarrow 0$, 就能得到粒子运动轨迹.

第二个例子：N 自由质点系统.

言归正传. 我们再看稍微复杂的力学系统: 3 维空间中的 N 个自由质点 m_1, \dots, m_N . 这个系统的“状态”无非是指每一个粒子的位置. 换言之, 欲确定此系

统在某时刻的状态, 只需确定在该时刻每个粒子的位置. 数学上, 我们自然考虑用 $3N$ 个参数 $(x_1, \dots, x_N; y_1, \dots, y_N; z_1, \dots, z_N)$ 来描述此系统的状态. 其中 (x_k, y_k, z_k) 是第 k 个质点 m_k 的空间位置坐标. 我们自然把 $(x_1, \dots, x_N; y_1, \dots, y_N; z_1, \dots, z_N)$ 视为 $3N$ 维空间 \mathbb{R}^{3N} 中的一个点 (或者矢量)! \mathbb{R}^{3N} 中的每一个点, 都对应于此系统的一个状态; 反之此力学系统的每一个状态都对应于 \mathbb{R}^{3N} 中的一个点. 我们认为 \mathbb{R}^{3N} 是此力学系统所有可能的状态构成的集合. 同样, 我们考虑这个系统随时间的演化. 确定该系统随时间的演化, 只需确定其中每个质点的位置如何随时间变化, 即确定描述该系统的 $3N$ 个参数 $x_1, \dots, x_N; \dots; z_1, \dots, z_N$ 如何随时间变化, 数学上自然把它们都视为关于时间 t 的函数. 于是, $t \mapsto (x_1(t), \dots, x_N(t); \dots; z_1(t), \dots, z_N(t))$ 自然给出了空间 \mathbb{R}^{3N} 中的一条曲线.

我们总结如下:

3 维空间 \mathbb{R}^3 的 N 个点	自然视为 $\xrightarrow{\quad}$	$3N$ 维空间 \mathbb{R}^{3N} 的 一个点
N 个质点的力学系统 随时间 t 演化	自然视为 $\xrightarrow{\quad}$	$3N$ 维空间 \mathbb{R}^{3N} 中的 一个点的运动

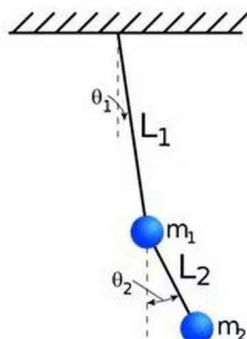
上述表格中, 左边是具体直观的物理对象, 而右边则相对抽象. 用牛顿力学来求解此系统, 只需要对其中每个质点作受力分析, 并使用牛顿第二定律.

这个例子给我们如下启示: 无论现实空间中的力学系统当中有多少个质点, 这个系统的状态总可以用某个“维数足够高”的抽象空间中的一个点来表示.

第三个例子: 双摆.

我们考虑以下力学系统. 如图: 两个小球 (视为质点) 被理想的细绳按图中方式连接. 并假设小球 m_1 与天花板的悬挂点之间的距离始终为绳长 L_1 , 而两小球之间的距离始终为绳长 L_2 .

如何确定此系统的状态? 最直接的方法是, 考虑小球 m_1 的横纵坐标 (x_1, y_1) , 以及小球 m_2 的横纵坐标 (x_2, y_2) , 这 4 个参数 $(x_1, x_2; y_1, y_2)$ 唯一确定此系统的状态, 此系统的状态自然视为四维空间 \mathbb{R}^4 中的点 $(x_1, x_2; y_1, y_2)$. 这当然没问题, 但如果仅仅就这, 只是第二个例子 (N 质点系统) 的特殊情况, 那意思就不大了.



图：双摆

我们注意, 这两个小球并不“自由”, 它们被绳子所约束:

- m_1 与天花板悬挂点的距离为 L_1 ,
- m_1 与 m_2 的距离为 L_2 ,

共两个约束条件. 数学上说, 描述此系统状态的四个参数 $x_1, x_2; y_1, y_2$ 需要满足两个方程. 根据解方程的经验来看, 两个方程能消掉两个未知数, 从而“真正自由”的参数只剩下 $4 - 2 = 2$ 个.

事实上, 考虑两根绳子与竖直方向的夹角 θ_1, θ_2 (如图所示). 直观上显然地看, 此力学系统的状态被 θ_1, θ_2 这两个参数所唯一确定. 我们完全可以用 \mathbb{R}^2 中的点 (θ_1, θ_2) 来描述此系统的状态. 这里的 $\mathbb{R}^2 = \{(\theta_1, \theta_2)\}$ 并非真实具体的物理空间, 而是数学上的抽象空间 [甚至比第二个例子中的 \mathbb{R}^{3N} 还抽象].

位形空间与 jet 空间.

一般来说, 如果某个力学系统能被 n 个**独立**的参数 (q_1, \dots, q_n) 所确定, 则称该系统的**自由度**为 n .²⁰ 比如, 在第二个例子里, 3 维空间中的 N 自由质点系统有 $3N$ 个自由度. 而在第三个例子中, $(x_1, x_2; y_1, y_2)$ 四个参数虽然能描述系统的

²⁰本讲义只考虑有限多个自由度的情况. 一般来说, 力学系统往往可以有无穷多自由度, 比如流体力学. 本讲义所介绍的理论在无穷多自由度情况下也有相应推广版本. 大致来说, 有限多自由度对应于有限维空间的理论, 而处理无穷多自由度就需要无穷维空间的数学理论, 我们在上一章也稍微介绍了一些.

状态, 但这四个参数并不独立, 被两个方程约束; 而 (θ_1, θ_2) 这两个参数是独立的, 因此双摆系统的自由度为 2.

对于一般的 n 自由度系统, 其状态可用 \mathbb{R}^n 中的点 (q_1, \dots, q_n) 来描述. 空间 $\mathbb{R}^n = \{(q_1, \dots, q_n) \mid q_i \in \mathbb{R}\}$ 称为该力学系统的**位形空间**(configuration space). 参数 q_1, \dots, q_n 称为该力学系统的 n 个**广义坐标**. 在第三个例子中, (θ_1, θ_2) 是双摆系统的一组广义坐标. 于是

力学系统随时间演化 \implies 位形空间中的点的运动

描述力学系统随时间的演化, 只需把各个广义坐标视为关于时间 t 的函数: $\mathbf{q}(t) = (q_1(t), \dots, q_n(t))$. 同时, $\mathbf{q}(t)$ 也自然视为位形空间中的一条曲线.

设 $\mathbf{q}(t) = (q_1(t), \dots, q_n(t))$ 为某力学系统随时间的演化, 与通常牛顿力学中“速度”的概念类似, 广义坐标对时间的导数 $\dot{\mathbf{q}}(t) = (\dot{q}_1(t), \dots, \dot{q}_n(t))$ 称为此力学系统随时间演化的**广义速度**.

位形空间 \mathbb{R}^n 中的点描述了力学系统的状态. 求解力学系统随时间的演化, 一般来说不仅要知道初始位置, 也要知道初始速度. 我们考虑构造一个更抽象的空间, 使得该空间中的点包含力学系统此时的“位置”与“速度”的信息. 这其实不难, 记某力学系统的位形空间为 \mathbb{R}^n , 则定义

$$J^1(\mathbb{R}^n) = \{(q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_n) \in \mathbb{R}^{2n} \mid q_i, \dot{q}_i \in \mathbb{R}\} = \mathbb{R}^{2n},$$

称为位形空间 \mathbb{R}^n 的 **jet 空间**. jet 空间中的点 $(\mathbf{q}, \dot{\mathbf{q}})$ 表示: 此力学系统某时刻的广义坐标为 \mathbf{q} , 广义速度为 $\dot{\mathbf{q}}$. 可见, jet 空间中的点也表示力学系统的“状态”, 不仅包含某时刻的“位置”信息, 还包含该时刻的“速度”信息.

特别注意: 位形空间 \mathbb{R}^n 中的(可微)曲线 $\mathbf{q}(t)$ 自然视为 jet 空间 $J^1(\mathbb{R}^n)$ 中的曲线 $(\mathbf{q}(t); \dot{\mathbf{q}}(t))$.

2.2 最小作用量原理与拉格朗日力学

我们开始学习拉格朗日力学.

最小作用量原理.

与牛顿力学不同, 拉格朗日力学的基本原理如下:

1. 力学系统的在某时刻的状态对应于位形空间 \mathbb{R}^n 中的点 $\mathbf{q} = (q_1, \dots, q_n)$.
2. 存在 $J^1(\mathbb{R}^n)$ 上的实值函数 $\mathcal{L} = \mathcal{L}(\mathbf{q}; \dot{\mathbf{q}})$, 这个函数由力学系统本身的性质决定, 称为该力学系统的拉格朗日量(Lagrangian).
3. 给定时刻 $t_0 < t_1$, 则对位形空间中的任何(可微)曲线 $\mathbf{q}(t)$, $t \in [t_0, t_1]$, 记

$$S[\mathbf{q}(t)] = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) dt,$$

$S[\mathbf{q}(t)]$ 称为函数 $\mathbf{q}(t)$ 的作用量.²¹

4. 若已知此力学系统在初时刻 t_0 的广义坐标为 $\mathbf{q}(t_0)$, 在末时刻 t_1 的广义坐标为 $\mathbf{q}(t_1)$, 则此系统随时间的演化 $\mathbf{q}_{cl}(t)$ 满足: $\mathbf{q}_{cl}(t)$ 是所有的连接 $(t_0, \mathbf{q}(t_0))$ 与 $(t_1, \mathbf{q}(t_1))$ 两点的曲线 $\mathbf{q}(t)$ 当中使得 $S[\mathbf{q}(t)]$ 取到极值的那条曲线. [通常是极小值]²²

上述原理称为最小作用量原理.²³

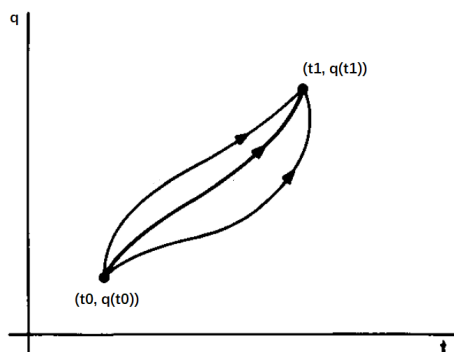


图: 从 $(t_0, \mathbf{q}(t_0))$ 到 $(t_1, \mathbf{q}(t_1))$ 有无穷多种可能的路径

粗俗地说, 对于每条路径 $\mathbf{q}(t)$, 作用量 $S[\mathbf{q}(t)]$ 是沿这条路径所付出的“代价”, 或者“走这条路所花的路费”. 拉格朗日力学的原理表明, 系统总会走“路费最低

²¹可见 S 是这样的映射: 输入一个函数 $\mathbf{q}(t)$, 则输出一个实数. 即, S 是“函数的函数”, 这种东西俗称泛函.

²²这里 \mathbf{q}_{cl} 的下标“cl”指 classical 的, 意为“经典的”, 用于强调这是经典力学系统.

²³作用量其实是取极值, 甚至也可以是极大值. 但我们还是习惯说“最小作用量”.

的那条路”(其实是取极值,也可以是最高).与牛顿力学不同,拉格朗日力学抛弃了“力”的概念:牛顿力学认为“力”驱使物体改变运动,拉格朗日力学认为“拉格朗日量”像一只“看不见的手”影响系统的演化.在拉格朗日力学看来,世间万物仿佛有了智慧,在抵达目标的无数种可能的道路中总选择“最有利”的那条.以此衍生出来的哲学探讨就耐人寻味了,这里不敢深入.²⁴

事实上,上述最小作用量原理确实堪称当今物理学中的哲学.不仅是经典力学,甚至是电磁学,相对论系统的演化,也都遵循最小作用量原理.在这种“哲学思想”的指导下,对于任何一个新发现的物理规律,物理学家都会想方设法凑一个拉格朗日量,使得这个物理规律可以解释为“系统遵循关于该拉格朗日量的最小作用量原理”.拉格朗日力学为物理学提供了一个强大的理论框架.

而在量子力学中,与经典力学不同,系统会同时走所有可能的路径,而不是“只选择其中某一条”.我们在后文将深入探讨.

欧拉-拉格朗日方程.

拉格朗日力学的“最小作用量原理”听起来像玄学.显然物理学不能止步于此,否则与就神学家的辩经无异.现在,若已知某力学系统的拉格朗日量 $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$, 我们来推导系统演化路径 $\mathbf{q}_{\text{cl}}(t)$ 所满足的方程.

根据最小作用量原理,系统的真实演化路径 $\mathbf{q}_{\text{cl}}(t)$ 是作用量取极值的那条.对任意函数 $\boldsymbol{\eta}(t) = (\eta_1(t), \dots, \eta_n(t))$, 如果 $\boldsymbol{\eta}(t_0) = \boldsymbol{\eta}(t_1) = \mathbf{0}$, 则对任何 (充分小的) 实数 ε , $t \mapsto \mathbf{q}_{\text{cl}}(t) + \varepsilon \boldsymbol{\eta}(t)$ 也是连接 $(t_0, \mathbf{q}(t_0))$ 与 $(t_1, \mathbf{q}(t_1))$ 的路径.因此,最小作用量原理表明

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} S[\mathbf{q}_{\text{cl}} + \varepsilon \boldsymbol{\eta}] = 0$$

对任何满足 $\boldsymbol{\eta}(t_0) = \boldsymbol{\eta}(t_1) = \mathbf{0}$ 的 (可微) 函数 $\boldsymbol{\eta}(t)$ 都成立.注意当 $\varepsilon \rightarrow 0$ 时

$$\begin{aligned} S[\mathbf{q}_{\text{cl}} + \varepsilon \boldsymbol{\eta}] &= \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}_{\text{cl}}(t) + \varepsilon \boldsymbol{\eta}(t); \dot{\mathbf{q}}_{\text{cl}}(t) + \varepsilon \dot{\boldsymbol{\eta}}(t)) dt \\ &= \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}_{\text{cl}}(t); \dot{\mathbf{q}}_{\text{cl}}(t)) dt + \varepsilon \int_{t_0}^{t_1} \left(\sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial q_i}(\mathbf{q}_{\text{cl}}(t); \dot{\mathbf{q}}_{\text{cl}}(t)) \eta_i(t) \right. \end{aligned}$$

²⁴如今绝大多数哲学系文科生受数理知识水平限制,对数学的了解仅限于四百年前的微积分,对近代物理的了解仅限于科普层次的相对论与量子力学.幸好他们不了解拉格朗日力学,不知道经典力学还有除了牛顿力学之外的理论,否则他们说不定能从拉格朗日力学里面深挖出些什么.

$$+ \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_i}(\mathbf{q}_{\text{cl}}(t); \dot{\mathbf{q}}_{\text{cl}}(t)) \dot{\eta}_i(t) \Big) dt + O(\varepsilon^2),$$

因此

$$\begin{aligned} 0 &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{S}[\mathbf{q}_{\text{cl}} + \varepsilon \boldsymbol{\eta}] = \int_{t_0}^{t_1} \left(\sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial q_i} \eta_i(t) + \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{\eta}_i(t) \right) dt \\ &= \int_{t_0}^{t_1} \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial q_i} \eta_i(t) dt + \int_{t_0}^{t_1} \sum_{i=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\eta_i(t) \\ &= \sum_{i=1}^n \left. \frac{\partial \mathcal{L}}{\partial q_i} \eta_i(t) \right|_{t_0}^{t_1} + \int_{t_0}^{t_1} \sum_{i=1}^n \left(\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \eta_i(t) dt \\ &= \int_{t_0}^{t_1} \sum_{i=1}^n \left(\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \eta_i(t) dt. \end{aligned}$$

于是, 由 $\boldsymbol{\eta}(t)$ 的任意性可知, 必有

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = 0, \quad 1 \leq i \leq n. \quad (2.2)$$

上式(2.2)称为**欧拉-拉格朗日方程**, 它是拉格朗日力学的基本方程. 一般来说, 这是关于未知函数 $\mathbf{q}(t)$ 的二阶常微分方程组, 含有 $\mathbf{q}(t)$ 关于时间 t 的二阶导数.

于是, 用拉格朗日力学求解力学问题只需以下步骤:

1. 选取适当的广义坐标 (q_1, \dots, q_n) 描述力学系统的状态.
2. 写出该力学系统的拉格朗日量 $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$.
3. 代入欧拉-拉格朗日方程(2.2), 然后解微分方程即可.

工程实践中的力学问题往往十分复杂, 若用牛顿力学求解, 会在**受力分析**上花费大量精力; 而拉格朗日力学则无需画图, 受力分析. 多数情况下, 拉格朗日力学的计算更加简便.

拉格朗日力学与牛顿力学.

拉格朗日力学的理论框架不仅适用于经典力学, 同样适用于相对论, 电磁学等. 而在经典力学中, 拉格朗日力学其实与牛顿力学等价.

考虑最简单的例子：质点在有势场中的一维自由运动。考虑一维空间中的质点 m ，其位置 $x = x(t)$ 随时间变化。该质点的势能为 $V = V(x)$ ，于是质点受力

$$F = -\frac{\partial V}{\partial x}.$$

根据牛顿力学，该质点的运动服从牛顿第二定律：

$$m\ddot{x}(t) = -\frac{\partial V}{\partial x}(t). \quad (2.3)$$

我们再用拉格朗日力学求解。已知该系统的拉格朗日量为

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x)$$

注意 $T := \frac{1}{2}m\dot{x}^2$ 为质点的**动能**。从而 $\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial V}{\partial x}$ ，以及 $\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$ 。代入欧拉-拉格朗日方程，也能得到(2.3)式，殊途同归！

经典力学系统的拉格朗日量一般为

$$\mathcal{L} = T - V, \quad (2.4)$$

其中 T 为系统的总动能， V 系统的势能。

广义动量与广义力。

我们注意，在简单例子“质点在有势场中的一维自由运动”中， $\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$ ，而 $m\dot{x}$ 恰为该质点的**动量**。这可推广到一般情形：

定义。 设某力学系统的拉格朗日量为 $\mathcal{L}(q, \dot{q})$ ，则称

$$\mathbf{p} := \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_1}, \dots, \frac{\partial \mathcal{L}}{\partial \dot{q}_n} \right)^\top \quad (2.5)$$

为该系统的**广义动量**。

广义动量的第 i 分量 $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ ， $1 \leq i \leq n$ 。从而欧拉-拉格朗日方程可改写为

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \dot{p}_i, \quad (2.6)$$

也就是说, $\frac{\partial \mathcal{L}}{\partial q_i}$ 是广义动量 p_i 关于时间的变化率. 我们回忆, 在牛顿力学中, 动量随时间的变化率恰为质点所受的力. 于是, 我们把

$$F_i := \frac{\partial \mathcal{L}}{\partial q_i} \quad (2.7)$$

称为**广义力**.

2.3 哈密顿力学

将拉格朗日力学的理论稍作数学变形, 就得到数学表达更漂亮的哈密顿力学. 此外, 本讲义将要介绍的“入门”量子力学的理论框架与哈密顿力学非常像.

勒让德变换与哈密顿正则方程.

我们回忆, 在拉格朗日力学中, 拉格朗日量 $\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ 是关于广义坐标与广义速度的函数, 广义动量 $\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_1}, \dots, \frac{\partial \mathcal{L}}{\partial \dot{q}_n} \right)^\top$. 在物理学中, **动量**往往被认为是基本的物理量, 比**速度**更加基本.²⁵ 甚至在相对论中, 很多经典物理规律不再适用, 而依然可以谈论动量. 于是, 我们希望找一个关于广义坐标与广义动量的某函数 $\mathcal{H}(\mathbf{q}, \mathbf{p})$ 来取代拉格朗日量 $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$, 并且使得 $\dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}$. 即

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \rightsquigarrow \dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}. \quad (2.8)$$

为此, 我们需要一个数学技巧. 设函数 $f(x)$ 满足

$$\frac{df(x)}{dx} = u(x).$$

记 $u(x)$ 的**反函数**为 $x(u)$. 我们令函数

$$g(u) = ux(u) - f(x(u)), \quad (2.9)$$

则容易验证

$$\frac{dg}{du} = x + u \frac{dx}{du} - \frac{df}{dx} \frac{dx}{du} = x.$$

²⁵当然这属于“马后炮”. 最初发明哈密顿力学的时候显然还没有相对论, 当时的人更不会意识到动量在近代物理中的重要地位.

总结: 对于函数 f , 我们构造了一个新函数 g , 使得 f' 与 g' 互为反函数. 我们称(2.9)中的 f 与 g 互为**勒让德变换**(Legendre transformation).

多元函数也有类似的勒让德变换. 现在, 我们考虑拉格朗日量 $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}})$ 的勒让德变换

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) := \sum_{i=1}^n p_i \dot{q}_i - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}), \quad (2.10)$$

注意此时 \mathbf{p}, \mathbf{q} 是独立变量, $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p})$, 并且 $\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}}$, 于是容易验证

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i + \sum_{j=1}^n p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_{j=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial p_i} = \dot{q}_i,$$

即 $\frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{q}}$. 这里的 $\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p})$ 称为该力学系统的**哈密顿量**(Hamiltonian). 由欧拉-拉格朗日方程, 容易得到

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial q_i} &= \sum_{j=1}^n p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial \mathcal{L}}{\partial q_i} - \sum_{j=1}^n \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial q_i} \\ &= -\frac{\partial \mathcal{L}}{\partial q_i} = -\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = -\dot{p}_i, \end{aligned}$$

即 $\frac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\dot{\mathbf{p}}$. 综上所述, 我们得到

$$\begin{cases} \dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \\ \dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} \end{cases} \quad (2.11)$$

此方程组称为**哈密顿正则方程**(Hamilton's canonical equation), 这是哈密顿力学的基本方程, 它是未知函数 $\mathbf{p}(t), \mathbf{q}(t)$ 的关于 t 的一阶常微分方程组. 哈密顿正则方程也可更加紧凑地写成如下形式:

$$\frac{d}{dt} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} & I \\ -I & \end{pmatrix} \begin{pmatrix} \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \\ \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \end{pmatrix} \quad (2.12)$$

在哈密顿力学中, 力学系统的状态用 \mathbb{R}^{2n} 中的点 (\mathbf{q}, \mathbf{p}) 来描述, 代表该系统的广义坐标为 \mathbf{q} , 广义动量为 \mathbf{p} . 空间 $\mathbb{R}^{2n} = \{(\mathbf{q}, \mathbf{p}) \mid \mathbf{q}, \mathbf{p} \in \mathbb{R}^n\}$ 称为该力学系统的**相空间**(phase space). 力学系统随时间的演化自然视为相空间中的曲线.

综上, 我们总结哈密顿力学的基本原理如下:

1. 力学系统在某时刻的状态对应于相空间 \mathbb{R}^{2n} 中的点 (\mathbf{q}, \mathbf{p}) .
2. 存在相空间 \mathbb{R}^{2n} 上的函数 $\mathcal{H}(\mathbf{q}, \mathbf{p})$, 它由力学系统本身的性质决定, 称为该系统的哈密顿量.
3. 力学系统随时间的演化 $(\mathbf{q}(t), \mathbf{p}(t))$ 满足哈密顿正则方程(2.11).

如果哈密顿量 $\mathcal{H}(\mathbf{q}, \mathbf{p})$ 不显含自变量 q_i , 即 $\frac{\partial \mathcal{H}}{\partial q_i} = 0$, 那么由哈密顿正则方程,

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} = 0,$$

即广义动量 p_i 不随时间变化 (守恒). 此时, 广义坐标 q_i 称为 **循环坐标**(cyclic coordinates).

我们回忆, 经典力学系统的拉格朗日量一般为 $\mathcal{L} = T - V$. 而哈密顿量通常为 $\mathcal{H} = T + V$, 即体系的总能量. 例如, 势场中的一维粒子的哈密顿量为

$$\mathcal{H}(x, p) = \frac{p^2}{2m} + V(x).$$

再比如, 考虑平面上的中心对称势场 $V(x, y) = V(\sqrt{x^2 + y^2})$, 即势能只与到原点的距离有关, 而与方位角无关 [比如质点的引力场, 点电荷的静电场等]. 则该势场中的粒子的哈密顿量

$$\mathcal{H}(x, y; p_x, p_y) = \frac{p_x^2 + p_y^2}{2m} + V(x, y),$$

注意到势场 V 的中心对称性, 我们选取极坐标 (r, θ) 作为系统的广义坐标将更方便. 此时 $V = V(r)$, 并且易知粒子动能 $T = \frac{m(\dot{x}^2 + \dot{y}^2)}{2} = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$. 于是, 拉格朗日量

$$\mathcal{L}(r, \theta; \dot{r}, \dot{\theta}) = T - V = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - V(r),$$

从而广义动量

$$\begin{aligned} p_r &= \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \\ p_\theta &= \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta}. \end{aligned}$$

可见广义动量 p_r 是动量的径向分量, 而 $p_\theta = mr^2\dot{\theta}$ 为通常的**角动量**. 用此广义坐标与广义动量, 对拉格朗日量作勒让德变换(2.10), 得到哈密顿量

$$\mathcal{H} = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r).$$

注意 $\frac{\partial \mathcal{H}}{\partial \theta} = 0$, 从而广义坐标 θ 是循环坐标, 于是广义动量 p_θ 守恒, 这恰为众所周知的**角动量守恒**.

泊松括号.

现在考虑相空间上的函数 $\omega = \omega(\mathbf{q}, \mathbf{p})$. 力学系统随时间演化 $t \mapsto (\mathbf{q}(t), \mathbf{p}(t))$, 则相应的 $\omega(t) = \omega(\mathbf{q}(t), \mathbf{p}(t))$ 也随时间变化. 由哈密顿正则方程, 容易得到 $\omega(t)$ 关于时间的导数

$$\frac{d\omega}{dt} = \sum_{i=1}^n \left(\frac{\partial \omega}{\partial q_i} \dot{q}_i + \frac{\partial \omega}{\partial p_i} \dot{p}_i \right) = \sum_{i=1}^n \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right). \quad (2.13)$$

一般地, 对于相空间上的函数 $f(\mathbf{q}, \mathbf{p})$ 与 $g(\mathbf{q}, \mathbf{p})$, 我们定义相空间上的一个新的函数 $\{f, g\} = \{f, g\}(\mathbf{q}, \mathbf{p})$ 如下:

$$\{f, g\} = \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right), \quad (2.14)$$

称为 f 与 g 的**泊松括号**(Poisson bracket). 于是, (2.13)改写为

$$\frac{d\omega}{dt} = \{\omega, \mathcal{H}\}, \quad (2.15)$$

特别地, ω 是守恒量 (不随时间变化) 当且仅当 $\{\omega, \mathcal{H}\} = 0$.

容易验证, 对相空间上的任意函数 f, g, h , 以及任意常数 a, b , 成立

$$\{f, g\} = -\{g, f\} \quad (2.16)$$

$$\{f, ag + bh\} = a\{f, g\} + b\{f, h\} \quad (2.17)$$

$$\{f, \{g, h\}\} = \{\{f, g\}, h\} + \{g, \{f, h\}\} \quad (2.18)$$

$$\{f, gh\} = g\{f, h\} + \{f, g\}h. \quad (2.19)$$

[注意这与矩阵对易子的运算律(1.12)-(1.15)相同], 留作简单的微积分习题.

特别地, 考虑相空间的坐标函数 $q_i: (\mathbf{q}, \mathbf{p}) \mapsto q_i$ 与 $p_i: (\mathbf{q}, \mathbf{p}) \mapsto p_i$, 容易验证

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad (2.20)$$

$$\{q_i, p_j\} = \delta_{ij}. \quad (2.21)$$

最后注意, 哈密顿正则方程(2.11)可用泊松括号改写为

$$\frac{dq_i}{dt} = \{q_i, \mathcal{H}\} \quad (2.22)$$

$$\frac{dp_i}{dt} = \{p_i, \mathcal{H}\}. \quad (2.23)$$

2.4 正则变换

在哈密顿力学中, n 自由度的力学系统的状态可由相空间 \mathbb{R}^{2n} 中的点来描述; 我们取定 \mathbb{R}^{2n} 的一组坐标 $(\mathbf{q}, \mathbf{p}) = (q_1, \dots, q_n; p_1, \dots, p_n)$. 将有关的力学量视为相空间 \mathbb{R}^{2n} 上的函数. \mathbb{R}^{2n} 上的两个函数可以作泊松括号运算: 对于函数 f 与 g , 其泊松括号 $\{f, g\}$ 仍是 \mathbb{R}^{2n} 上的函数. 泊松括号 $\{f, g\}$ 在给定的坐标 (\mathbf{q}, \mathbf{p}) 下的表达式见(2.14).

正则变换.

我们在实际计算中往往会另选新的坐标系, 使得新坐标系中的计算更简便, 例如某些情况下用平面极坐标系比用平面直角坐标系更简便. 一般地, 考虑相空间 \mathbb{R}^{2n} 的一组新坐标, 记作 $(\bar{\mathbf{q}}, \bar{\mathbf{p}}) = (\bar{q}_1, \dots, \bar{q}_n; \bar{p}_1, \dots, \bar{p}_n)$, 这里的 $\bar{q}_i = \bar{q}_i(\mathbf{q}, \mathbf{p})$, $\bar{p}_i = \bar{p}_i(\mathbf{q}, \mathbf{p})$, $1 \leq i \leq n$, 都是关于旧坐标 \mathbf{p}, \mathbf{q} 的函数. 此时, 相空间 \mathbb{R}^{2n} 上的任何函数 $f(\mathbf{q}, \mathbf{p})$ 都可视为关于新变量 $\bar{\mathbf{q}}, \bar{\mathbf{p}}$ 的函数 $f(\bar{\mathbf{q}}, \bar{\mathbf{p}})$.

在新坐标 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 下, 泊松括号与哈密顿正则方程的样子会如何? 首先注意泊松括号的如下性质:

性质 2.1. 设 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 是相空间 \mathbb{R}^{2n} 的任何一组坐标, f, g 是 \mathbb{R}^{2n} 上的函数, 则泊松括号满足

$$\{f, g\} = \sum_{j=1}^n \left(\frac{\partial g}{\partial \bar{q}_j} \{f, \bar{q}_j\} + \frac{\partial g}{\partial \bar{p}_j} \{f, \bar{p}_j\} \right). \quad (2.24)$$

在证明之前,我们先来看看这个引理是什么意思. 我们固定 f , 则自然得到这样一个映射: 将函数 g 变成函数 $\{f, g\}$, 即 $\{f, \cdot\}: g \mapsto \{f, g\}$. 而(2.24)的等号右边非常像复合函数求导的链式法则, 于是此性质表明, 我们可以把运算 $g \mapsto \{f, g\}$ 想象成对 g “求某种导数”. [其实, 上一节的(2.19)式也可以想象成这种求导的 “乘积求导法则”. 种种迹象表明, 泊松括号的运算满足求导运算的运算律, 于是我们不妨就把它当成求导运算.] ²⁶

证明. 直接验证即可, 只需注意求导的链式法则

$$\frac{\partial g}{\partial p_i} = \sum_{j=1}^n \left(\frac{\partial g}{\partial \bar{q}_j} \frac{\partial \bar{q}_j}{\partial p_i} + \frac{\partial g}{\partial \bar{p}_j} \frac{\partial \bar{p}_j}{\partial p_i} \right),$$

$\frac{\partial g}{\partial q_i}$ 也类似. 代入泊松括号定义式(2.14)硬算如下:

$$\begin{aligned} \{f, g\} &= \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right) \\ &= \sum_{i,j=1}^n \left[\frac{\partial f}{\partial q_i} \left(\frac{\partial g}{\partial \bar{q}_j} \frac{\partial \bar{q}_j}{\partial p_i} + \frac{\partial g}{\partial \bar{p}_j} \frac{\partial \bar{p}_j}{\partial p_i} \right) - \frac{\partial f}{\partial p_i} \left(\frac{\partial g}{\partial \bar{q}_j} \frac{\partial \bar{q}_j}{\partial q_i} + \frac{\partial g}{\partial \bar{p}_j} \frac{\partial \bar{p}_j}{\partial q_i} \right) \right] \\ &= \sum_{j=1}^n \left[\frac{\partial g}{\partial \bar{q}_j} \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \frac{\partial \bar{q}_j}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial \bar{q}_j}{\partial q_i} \right) + \frac{\partial g}{\partial \bar{p}_j} \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \frac{\partial \bar{p}_j}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial \bar{p}_j}{\partial q_i} \right) \right] \\ &= \sum_{j=1}^n \left(\frac{\partial g}{\partial \bar{q}_j} \{f, \bar{q}_j\} + \frac{\partial g}{\partial \bar{p}_j} \{f, \bar{p}_j\} \right), \end{aligned}$$

从而得证. □

现在我们研究新坐标 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 下的哈密顿正则方程. 设相空间 \mathbb{R}^{2n} 上的函数 \mathcal{H} 是该系统的哈密顿量. 首先注意(2.15)式与坐标选取无关. 特别地, 在(2.15)中取 $\omega = \bar{q}_i$, 并利用刚证明的性质(2.24), 可得

$$\begin{aligned} \frac{d\bar{q}_i}{dt} &= \{\bar{q}_i, \mathcal{H}\} \\ &= \sum_{j=1}^n \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_j} \{\bar{q}_i, \bar{q}_j\} + \frac{\partial \mathcal{H}}{\partial \bar{p}_j} \{\bar{q}_i, \bar{p}_j\} \right). \end{aligned}$$

²⁶再次重复: “只要嘴巴像鸭子, 羽毛像鸭子, 游泳姿势像鸭子, 那它就是鸭子!”

类似地, 也有

$$\frac{d\bar{p}_i}{dt} = \sum_{j=1}^n \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_j} \{\bar{p}_i, \bar{q}_j\} + \frac{\partial \mathcal{H}}{\partial \bar{p}_j} \{\bar{p}_i, \bar{p}_j\} \right).$$

如果我们希望新坐标 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 下的哈密顿正则方程也有 (2.11) 的样子, 即我们希望成立

$$\begin{cases} \frac{d\bar{q}_i}{dt} = \frac{\partial \mathcal{H}}{\partial \bar{p}_i} \\ \frac{d\bar{p}_i}{dt} = -\frac{\partial \mathcal{H}}{\partial \bar{q}_i}, \end{cases}$$

那么新坐标 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 必须满足: 对任意 $1 \leq i, j \leq n$, 成立

$$\{\bar{q}_i, \bar{q}_j\} = \{\bar{p}_i, \bar{p}_j\} = 0, \quad \{\bar{q}_i, \bar{p}_j\} = \delta_{ij}. \quad (2.25)$$

我们总结如下:

定义. 若相空间 \mathbb{R}^{2n} 的坐标 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 满足关系 (2.25), 则称 $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ 是相空间 \mathbb{R}^{2n} 的一组 **正则坐标** (canonical coordinate). 若坐标变换 $(\mathbf{q}, \mathbf{p}) \mapsto (\bar{\mathbf{q}}(\mathbf{q}, \mathbf{p}), \bar{\mathbf{p}}(\mathbf{q}, \mathbf{p}))$ 将正则坐标变为正则坐标, 则称该变换为 **正则变换** (canonical transformation).²⁷

容易验证 [留给读者练习], 若 $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ 是正则坐标, 则在此正则坐标下, 泊松括号的表达式为:

$$\{f, g\} = \sum_{i=1}^n \left(\frac{\partial f}{\partial \bar{q}_i} \frac{\partial g}{\partial \bar{p}_i} - \frac{\partial f}{\partial \bar{p}_i} \frac{\partial g}{\partial \bar{q}_i} \right), \quad (2.26)$$

与上一节我们所熟悉的样子一模一样. 而在非正则坐标下, 泊松括号与哈密顿正则方程的表达式会更加复杂.

正则变换的例子.

考虑在二维平面上运动的粒子, 其相空间为 $\mathbb{R}^{2 \times 2} = \{(x, y; p_x, p_y)\}$, 这里的 x, y 是通常的平面直角坐标, p_x, p_y 是相应方向上的动量.

1. 考虑旋转变换

$$\bar{x} = x \cos \theta - y \sin \theta$$

²⁷在数学中, canonical 一词常翻译为“典范”或“典则”; 而这里译为“正则”是沿袭国内老一辈物理学家的习惯.

$$\begin{aligned}\bar{y} &= x \sin \theta + y \cos \theta \\ \bar{p}_x &= p_x \cos \theta - p_y \sin \theta \\ \bar{p}_y &= p_x \sin \theta + p_y \cos \theta,\end{aligned}$$

则 $(\bar{x}, \bar{y}; \bar{p}_x, \bar{p}_y)$ 是正则坐标.

2. 考虑极坐标变换

$$\begin{aligned}\rho &= \sqrt{x^2 + y^2}, & \phi &= \arctan \frac{y}{x}, \\ p_\rho &= \frac{x p_x - y p_y}{\sqrt{x^2 + y^2}}, & p_\phi &= x p_y - y p_x,\end{aligned}$$

则 $(\rho, \phi; p_\rho, p_\phi)$ 是正则坐标.

以上两例的验证留给读者练习.

3 量子力学基本原理

欧罗巴人天文推算之密，工匠制作之巧，实逾前古；
其议论夸诈迂怪，亦为异端之尤。
国朝节取其技能，禁传其学术，具存深意矣。

——《四库全书总目》

前两章介绍了线性代数与经典力学的一小部分必要预备知识，以帮助我们进入量子世界。预备知识就先暂时介绍这些；前两章限于篇幅没介绍的内容会等以后需要用的时候再介绍。

是时候进入量子世界了！

3.1 导引：量子世界中的“豌豆”

想必读者早已从各种科普途径听说过量子现象的神奇之处，比如：

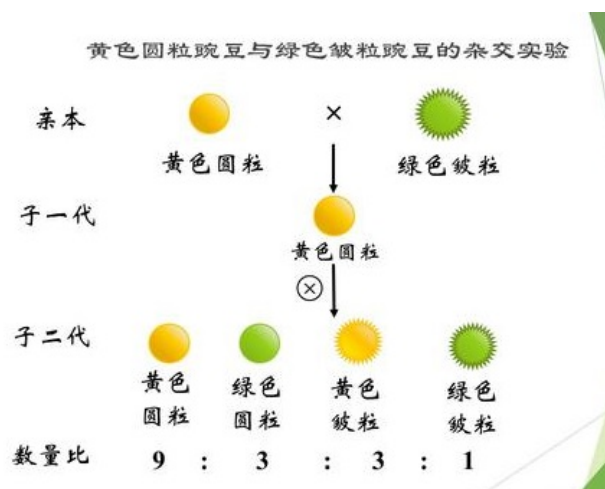
双缝干涉实验，薛定谔的猫...
量子通信，量子计算...

在本节，我们讲一个量子世界里的故事。这个故事是真实发生的，但我们先不透露主人公的真实身份，暂时把主角称为**豌豆**。

豌豆可以有很多种不同性状，我们只关心其中的两个：**颜色**与**形状**。我们在中学生物课上学过，豌豆的颜色有**黄**、**绿**两种，形状有**圆**、**皱**两种。当然，我们要讲的故事不再是杂交实验，性状分离的老生常谈，而比这更加幼稚。

经典世界里的“豌豆”

首先我们看经典世界，也就是我们经验感知范围内的世界里发生的故事。



图：豌豆杂交实验 [图文无关]

- 有一大麻袋好多好多的豌豆，里面有黄有绿，有圆有皱。
- 有两个喜欢豌豆小朋友，我们暂时给他们起名“颜色”与“形状”。
- 名字叫“颜色”的小朋友对豌豆的颜色很敏感，喜欢按颜色将豌豆分类；名字叫“形状”的小朋友则更在意豌豆的形状。

有一天，这两个小朋友和那一麻袋豌豆发生了这样的事情：

1. 首先，颜色 观察每一粒豌豆的颜色，发现黄绿各占一半；他把绿豌豆都扔掉，把黄豌豆都给形状。
2. 然后，形状 观察他收到的每一粒黄豌豆的形状，发现圆皱各占一半，并把圆的那一堆返还给颜色。
3. 颜色 收到这些圆豌豆之后立刻开心地检查这些豌豆的颜色，请问颜色 会发现什么呢？

显然，颜色 会发现，这些圆豌豆全都是黄的！我们再想象，如果在第 (2) 步的时候，形状 把皱的那一堆返还给颜色，那么 颜色 会发现什么？他会发现这些皱豌豆也居然全是黄的！

这是经典世界中的“正常现象”，成年人只会觉得无聊幼稚，而好奇的小朋友则能体会到“物外之趣”。

量子世界里的“豌豆”

我们再来看另一个世界里发生的故事.

同样有一大堆豌豆, 颜色有黄有绿, 形状有圆有皱; 不过这些豌豆是“量子豌豆”, 其形状和颜色无法用肉眼直接观测, 需要借助仪器. 这个世界也有两个小朋友: 颜色 与 形状, 他们各有一台高科技仪器, 分别能够检测豌豆的颜色与形状.

故事剧本仿照经典世界:

1. 首先, 颜色 利用手中的高科技仪器依次观测每一粒豌豆的颜色, 发现黄绿各占一半, 并把所有 观测结果为黄 的豌豆都给 形状, 把 观测结果为绿 的豌豆扔掉.
2. 然后, 形状 用手中的高科技仪器依次观测他收到的每一粒豌豆的形状, 发现圆皱各占一半, 并把 观测结果为圆 的豌豆返还给 颜色.
3. 此时, 颜色 用手中的高科技仪器观测他收到的豌豆的颜色, 会有什么观测结果呢?

故事的结局与经典世界的情形很不一样: 经 颜色 观测, 这些豌豆有一半的观测结果为黄, 另一半的观测结果为绿! 我们回忆这个故事的前两步, 在第 1 步中 颜色 将其中的黄色豌豆筛选出来, 然后在第 2 步, 形状 又从中筛选出圆豌豆. 所以, 按照经典世界里正常人的思维方式, 颜色 在第 3 步所收到的豌豆应该都是“又黄又圆”的, 再观测它们的颜色应该全部检出黄色; 但真实的观测结果 [一半黄一半绿] 并非如此.

我们继续讲量子豌豆的故事. 故事剧本稍微改一下 [修改第 2 步]:

1. 首先, 颜色 利用手中的高科技仪器依次观测每一粒豌豆的颜色, 发现黄绿各占一半, 并把所有 观测结果为黄 的豌豆都给 形状, 把 观测结果为绿 的豌豆扔掉.
2. 然后, 形状 用手中的高科技仪器依次观测他收到的每一粒豌豆的形状, 发现圆皱各占一半. 然后, 形状 把观测之后的所有豌豆 [无论圆皱] 都返还给 颜色.
3. 此时, 颜色 用手中的高科技仪器观测他收到的豌豆的颜色, 会有什么观测结果呢?

故事结局依然是：经颜色 观测，这些豌豆有一半的观测结果为黄，另一半的观测结果为绿，这种事情不会在经典世界里发生。在第 1 步里明明已经筛选出黄豌豆了，仅仅是拿去给形状 玩了一会，这些豌豆的颜色就变了？

颜色 觉得其中必有蹊跷，一定是形状 不守规矩，玩豌豆的时候偷偷做手脚。于是，接下来颜色 变得谨慎多疑起来。我们的第三个故事如下：

1. 首先，颜色 利用手中的高科技仪器依次观测每一粒豌豆的颜色，发现黄绿各占一半，并把所有观测结果为黄 的豌豆放到一堆，把观测结果为绿 放到另一堆。
2. 颜色 为了保险起见，把刚才挑选出的那堆黄豌豆重新逐一观测一遍，观测结果表明它们的确都是黄色；同样，又把刚才挑选出来的那堆绿豌豆重新逐一观测一遍，观测结果表明它们的确都是绿色。[这符合经典世界里的经验.]
3. 颜色 反复执行第 2 步，反复确认豌豆颜色，发现黄的就是黄的，绿的就是绿的，并没有像前两个故事里“从之前确认的黄豌豆里发现出绿的”。

不仅颜色 有疑惑，形状 的内心也同样疑惑，因为还可以有这个故事：

1. 首先，形状 利用手中的高科技仪器依次观测每一粒豌豆的形状，发现圆皱各占一半，并把所有观测结果为圆 的豌豆都给形状，把观测结果为皱 的豌豆扔掉。
2. 然后，颜色 用手中的高科技仪器依次观测他收到的每一粒豌豆的颜色，发现黄绿各占一半。然后，颜色 把观测之后的所有豌豆[无论黄绿]都返还给形状。
3. 此时，形状 用手中的高科技仪器观测他收到的豌豆的形状，会有什么观测结果呢？

答案是：在“黄豌豆”中，观测结果为圆和皱的各占一半；在“绿豌豆”中，观测结果也是圆皱各占一半。经典世界不可能发生这种事，因为在第 1 步就已经筛选出圆豌豆，淘汰掉皱豌豆了！

以上便是量子世界的神秘现象之一。

一些哲学讨论.

我们如何描述豌豆的属性? 先回忆经典世界里的事情.

“经典豌豆”的属性可以用二元组来表示:

(黄, 圆), (黄, 皱), (绿, 圆), (绿, 皱).

这个二元数组就是豌豆在经典意义下的“**态**”. 我们可以说:

“这是一粒黄的圆豌豆”,

或者说, 豌豆处于状态 (黄, 圆). 这句话包含了两个**表象**: 颜色表象和形状表象. 更细致地说, 我们首先在颜色表象中对小球进行观测, 然后在形状表象中观测. 我们还可以说:

“这是一粒圆的黄豌豆”,

先说形状表象, 再是颜色表象. 这两种说法等价. 于是我们说, 这两个表象**可对易**. 进而, 在颜色表象和在形状表象中看到的现象“彼此相容”, “可以共存”. 在经典世界中, 不同的表象都是彼此相容的. 在一个表象中进行的测量不会影响到在另一个表象中的测量.

而“量子豌豆”所表现出的种种迹象表明, 其颜色表象和形状表象并不相容. 我们不能用经典世界里习以为常的思维来思考量子世界.

斯特恩-格拉赫实验.

以上“量子豌豆”是其实是根据真实事件改编, 这就是著名的**斯特恩-格拉赫实验**(Stern-Gerlach experiment).²⁸ 这个实验将大量**银原子**射入磁场, 射入磁场的银原子受磁场影响, 其运动轨迹发生偏转; 射出磁场的银原子打在屏上, 屏可以显示银原子的位置. 如图所示:

²⁸若读者读过刘慈欣的小说《三体》第一部, 则不难发现笔者在此致敬“三体游戏”. “三体游戏”借助我们熟悉的伏羲、秦始皇、牛顿等地球人, 将三体世界的真实历史事件娓娓道来.

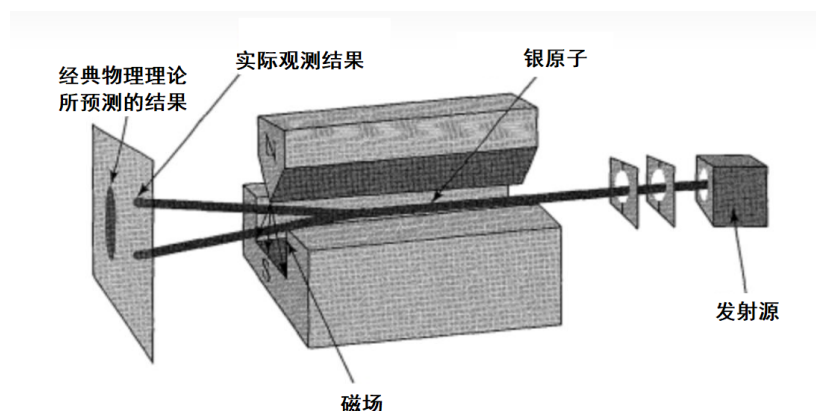


图: 斯特恩-格拉赫实验

银原子穿过磁场的过程中, 受磁场作用, 其运动轨迹会向上或向下偏转, 偏转的角度取决于银原子本身的性质 [最外层电子的自旋所对应的磁矩]。根据经典物理理论, 大量银原子偏转的角度分布应该是连续不断的, 屏幕上应该观察到一条从上到下的连续的线。

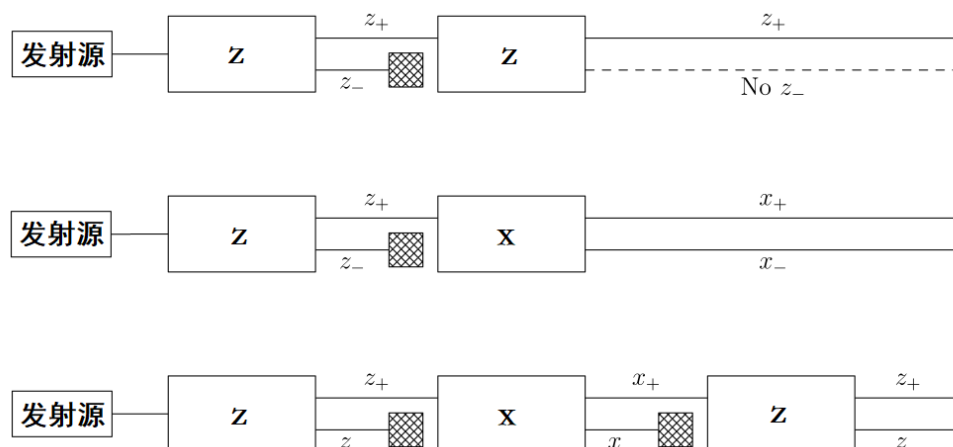
可实际观测结果是: 屏幕上观察到一上一下**两个亮斑**! 这个现象俗称“**空间量子化**”, 说明银原子最外层电子自旋的磁矩的 z -分量的取值是**离散的**, 只有**上**和**下**两种取值。[这对应于豌豆的“黄”与“绿”]。

这个磁场是一台高科技仪器, 记作“ z ”, 用来检测银原子是“上”还是“下”。

除了这个磁场, 还有另一个方向的磁场, 使得银原子从中穿过之后会向左偏或者向右偏。实验结果类似: 屏幕上出现的是左右两个斑点, 而不是经典物理学预测的从左到右的连续不断的一条线。可见银原子还有另一种属性, 其取值为**左**和**右**。[这对应于豌豆的“圆”与“皱”]。该属性同样可以用“高科技仪器”来检测, 此仪器记作“ x ”。

实验并没有结束, 它还有后续, 如下图所示:

- 第一个实验: 先让银原子通过 z , 挑选出向上偏转的; 然后让这些向上偏转的原子再次通过 z , 检测结果依然是向上偏转。
- 第二个实验: 先让银原子通过 z , 挑选出向上偏转的; 然后让这些向上偏转的



图：相继的斯特恩-格拉赫实验

原子通过 x , 检测结果为一半向左一半向右.

- 第三个实验: 先让银原子通过 z , 挑选出向上偏转的; 然后让这些向上偏转的原子通过 x , 检测结果为一半向左一半向右, 然后再用 z 来检测向左偏转的原子, 检测结果为一半向上一半向下.

这里第三个实验的现象正是前文“量子豌豆”的故事.

3.2 量子态与观测量

“量子豌豆”的诡异行为与经典世界截然不同. 除了斯特恩-格拉赫实验, 还有很多其他著名物理实验, 这些实验从各种方面表明, 我们所生活的真实世界里存在很多违反日常生活直觉与经典物理理论的诡异现象. 限于篇幅, 这里不再介绍更多的物理实验.

而量子力学要做的, 就是编出一套逻辑自洽, 能自圆其说的理论来解释这些诡异现象²⁹. 我们回忆, 在经典力学的理论里, 首先要回答两个问题:

- 如何描述力学系统的状态? 如何描述力学系统的某个力学量?

²⁹经典力学, 电磁学, 相对论, 乃至所有的科学理论, 不也是这样吗? 都是为了解释实验现象而编造出来的. 所有的理论都是对真实世界的近似描述.

三种经典力学 (牛顿, 拉格朗日, 哈密顿) 给出的回答大致相同: 力学系统的状态是某个 (抽象的) 空间里的点, 力学量是指那个空间上的函数. 而 (古典) 量子力学³⁰对此问题的回答如下:

- **量子公理 1.** 所谓量子系统, 是指 (复数域 \mathbb{C} 上的) 希尔伯特空间 \mathbb{V} (有限维或无限维). 量子系统的状态, 即量子态, 是指 \mathbb{V} 中的非零矢量 $|\psi\rangle$.
- **量子公理 2.** 量子系统 \mathbb{V} 上的观测测量是指 \mathbb{V} 上的厄米特算子.

[不同于量子科普爽文, 量子力学理论严重依赖复数域 \mathbb{C} 上的线性代数. 一些必要预备知识和记号说明可见本书第 1 章.]

我们以“量子豌豆”为例来解释上述两条公理.

对于一个“量子豌豆”, 颜色和形状是这个量子系统的两个观测量. 颜色的取值有两种可能: 黄, 绿; 而形状的取值也有两种: 圆, 皱.

在进一步解释这两条公理之前, 我们再来看看“量子豌豆”的诡异之处. 实验表明: 如果豌豆的颜色已被完全确定为黄色, 那这个豌豆的形状就“完全不确定”, 它有一半的可能是圆, 有一半可能是皱; 同样, 若豌豆的形状完全确定, 则它的颜色就会随机.

复盘一下量子豌豆的故事. 当量子豌豆的颜色被确认为黄色时, 它的形状就随机. 此时, 若检测这颗豌豆的形状, 则检测结果为圆, 皱的概率都是 $\frac{1}{2}$. 一旦形状的检测结果显示出来, 比如检测结果为圆, 它的颜色就又不确定了 [即使之前检测结果为黄], 再次检测颜色则会有 $\frac{1}{2}$ 的概率检出绿色.

我们无法同时确定“量子豌豆”的颜色和形状, 不能像经典世界里那样说豌豆“又黄又圆”. 这种随机性被认为是量子系统的本质性质, 绝不是测量误差导致. 爱因斯坦曾非常抵触随机性, 并说出如下名言:

上帝不会掷骰子.³¹

³⁰量子力学也有多种不同理论, 本讲义目前只介绍诞生于 100 年前, 最容易入门的版本, 姑且称为“古典量子力学”. 我们将看到, 这套理论是对哈密顿力学的“模仿”.

³¹这算是爱因斯坦的黑历史了.

量子态

量子力学理论的关键之一就在于如何描述这种“观测结果不确定性”. 一般来说, “量子豌豆”的“状态”是“又黄又绿”, “又圆又皱”的“叠加态”.

这与希尔伯特空间有何关系? 事实上, “量子豌豆”系统所对应的希尔伯特空间 $\mathbb{V} = \mathbb{C}^2$, 即 2 维复线性空间. 这个空间 \mathbb{V} 有如下么正基:

$$\{|黄\rangle, |绿\rangle\}.$$

也就是说, $|黄\rangle$ 与 $|绿\rangle$ 是 \mathbb{V} 中的两个矢量, 它们模长都为 1, 并且两两正交:

$$\langle黄|黄\rangle = \langle绿|绿\rangle = 1, \quad \langle黄|绿\rangle = \langle绿|黄\rangle = 0. \quad (3.1)$$

这一组基也称作 \mathbb{V} 的“颜色表象”.

除了颜色表象, 空间 \mathbb{V} 还有另一组么正基 $\{|圆\rangle, |皱\rangle\}$:

$$\langle圆|圆\rangle = \langle皱|皱\rangle = 1, \quad \langle圆|皱\rangle = \langle皱|圆\rangle = 0. \quad (3.2)$$

这组基称为 \mathbb{V} 的“形状表象”. 这两组基的转换关系如下:

$$|黄\rangle = \frac{1}{\sqrt{2}}|圆\rangle + \frac{1}{\sqrt{2}}|皱\rangle \quad (3.3)$$

$$|绿\rangle = \frac{1}{\sqrt{2}}|圆\rangle - \frac{1}{\sqrt{2}}|皱\rangle \quad (3.4)$$

一般地, 量子豌豆所处的状态, 即“量子态”, 由 \mathbb{V} 中的非零矢量 $|\psi\rangle$ 所描述. $|\psi\rangle$ 在颜色表象下形如

$$|\psi\rangle = a_1|黄\rangle + a_2|绿\rangle,$$

这里的 a_1, a_2 是不全为 0 的复数. 同样 $|\psi\rangle$ 也可以在形状表象下展开.

我们总结“经典豌豆”与“量子豌豆”的状态描述如下:

- “经典豌豆”的态只有 4 种: (黄, 圆), (黄, 皱), (绿, 圆), (绿, 皱).
- “量子豌豆”的态有无数种: $|\psi\rangle = a_1|黄\rangle + a_2|绿\rangle$.

至此, 我们已举例说明量子公理 1, 即如何描述量子态. 我们接下来说明如何把观测量视为 \mathbb{V} 上的厄米特算子.

厄米特算子 (II).

为理解量子公理 2, 我们需要继续学习³² 厄米特算子的有关知识.

先以有限维空间 \mathbb{V} 为例, 注意有限维内积空间都是希尔伯特空间. 我们回忆厄米特算子满足性质 1.7. 对于 n 维希尔伯特空间 \mathbb{V} 上的厄米特算子 H , 记 $\sigma(H) := \{\lambda_1, \dots, \lambda_m\}$ 是 H 的全部本征值构成的集合. 注意厄米特算子的本征值 $\lambda_1, \dots, \lambda_m$ 都是实数. 对于 $k = 1, \dots, m$, 记

$$\mathbb{V}_{H, \lambda_k} := \{|\psi\rangle \in \mathbb{V} \mid H|\psi\rangle = \lambda_k \psi\},$$

即 H 的属于本征值 λ_k 的本征矢量之全体. 容易验证 $\mathbb{V}_{H, \lambda_k}$ 是 \mathbb{V} 的子空间, 它称为 H 的关于本征值 λ_k 的本征子空间. 这些本征子空间两两正交, 并且满足直和分解

$$\mathbb{V} = \mathbb{V}_{H, \lambda_1} \oplus \mathbb{V}_{H, \lambda_2} \oplus \cdots \oplus \mathbb{V}_{H, \lambda_m}. \quad (3.5)$$

一般来说, 若 H 具有 n 个两两互异的本征值, 即 $m = n$ 时, 每个本征值所对应的本征子空间 $\mathbb{V}_{H, \lambda_k}$ 都是 1 维的. 但也可能发生例外, 即本征值的个数小于空间 \mathbb{V} 的维数, 此时必有某个本征值 λ_k , 其对应的本征子空间 $\mathbb{V}_{H, \lambda_k}$ 的维数大于 1. 若 $\dim \mathbb{V}_{H, \lambda_k} > 1$, 则称 λ_k 是**简并**的; 若 $\dim \mathbb{V}_{H, \lambda_k} = 1$, 则称 λ_k **无简并**.

现在, 设 H 是 n 维希尔伯特空间 \mathbb{V} 上的厄米特算子, 则在空间分解(3.5)下, \mathbb{V} 中的每个矢量 $|\psi\rangle$ 都能唯一写成

$$|\psi\rangle = |\psi_1\rangle + |\psi_2\rangle + \cdots + |\psi_m\rangle,$$

使得 $|\psi_k\rangle \in \mathbb{V}_{H, \lambda_k}$, $\forall k = 1, \dots, m$. 注意 $|\psi_k\rangle$ 是 $|\psi\rangle$ 在子空间 $\mathbb{V}_{H, \lambda_k}$ 上的**正交投影**. 于是, 对于 H 的每个本征值 λ_k , 自然引入 \mathbb{V} 上的线性算子 $\mathbb{P}_{H, \lambda_k}$, 使得

$$\mathbb{P}_{H, \lambda_k} |\psi\rangle = |\psi_k\rangle, \quad 1 \leq k \leq m, \quad (3.6)$$

其中 $|\psi_k\rangle$ 满足(3.5). 算子 $\mathbb{P}_{H, \lambda_k}$ 是沿子空间 $\mathbb{V}_{H, \lambda_k}$ 的正交投影算子. 容易验证 $\mathbb{P}_{H, \lambda_k}$ 是厄米特算子, 并且

$$\sum_{k=1}^m \mathbb{P}_{H, \lambda_k} = I, \quad \mathbb{P}_{H, \lambda_k} \mathbb{P}_{H, \lambda_{k'}} = \delta_{k, k'} \mathbb{P}_{H, \lambda_k}.$$

³²根本不是所谓“继续学习”. 笔者假定大家都已经掌握线性代数, 这里仅仅是为了澄清, 统一符号约定.

对每个子空间 \mathbb{V}_{H,λ_k} , 取定 \mathbb{V}_{H,λ_k} 的一组么正基 $\{|h_{k,1}\rangle, |h_{k,2}\rangle, \dots, |h_{k,d_k}\rangle\}$, 其中 $d_k := \dim \mathbb{V}_{H,\lambda_k}$, 满足 $d_1 + d_2 + \dots + d_m = n$. 于是, $\{|h_{k,l}\rangle \mid 1 \leq k \leq m, 1 \leq l \leq d_k\}$ 构成 \mathbb{V} 的一组么正基. 在这组基下, (3.6)式中的算子 \mathbb{P}_{H,λ_k} 可表示为

$$\mathbb{P}_{H,\lambda_k} = \sum_{l=1}^{d_k} |h_{k,l}\rangle \langle h_{k,l}|. \quad (3.7)$$

上述表示线性算子的方法已在第 1 章介绍, 读者可复习(1.23)附近的内容.

观测量与概率.

我们回到量子公理 2. 到底什么是观测量? 为了把这个问题说清楚, 不妨把此公理与下述第 3 条公理一起理解:

• 量子公理 3. 设 H 是量子系统 \mathbb{V} 中的观测量, $|\psi\rangle$ 是量子态.

1. 观测量 H 的值, 是指 $\sigma(H) = \{\lambda_1, \lambda_2, \dots\}$ 中的元素, 即 H 的本征值.
2. 量子态 $|\psi\rangle$ 被 H 观测, 观测结果为 λ_k 的概率 p_{H,λ_k} 等于

$$p_{H,\lambda_k} = \frac{\langle \psi | \mathbb{P}_{H,\lambda_k} | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (3.8)$$

3. 若观测结果为 λ_k , 则观测后 的量子态从之前的 $|\psi\rangle$ 变化为 $\mathbb{P}_{H,\lambda_k} |\psi\rangle$.

这里需要说明几点.

1. 量子态 $|\psi\rangle$ 被 H 观测, 测得结果为 λ_k 的概率 p_{H,λ_k} 的表达式(3.8)的含义如下: 注意 \mathbb{P}_{H,λ_k} 是幂等算子 (即 $\mathbb{P}_{H,\lambda_k}^2 = \mathbb{P}_{H,\lambda_k}$), 并且是厄米特算子, 从而

$$p_{H,\lambda_k} = \frac{\langle \mathbb{P}_{H,\lambda_k} \psi | \mathbb{P}_{H,\lambda_k} \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\|\mathbb{P}_{H,\lambda_k} \psi\|^2}{\|\psi\|^2}; \quad (3.9)$$

特别地, 若 $|\psi\rangle$ 的模长为 1, 则 $p_{H,\lambda_k} = \|\mathbb{P}_{H,\lambda_k}\psi\|^2$, 为 $|\psi\rangle$ 在相应的本征子空间 \mathbb{V}_{H,λ_k} 的投影的模长平方.

2. 对于量子态 $|\psi\rangle$ 与观测量 H , “用 H 测量 $|\psi\rangle$ ” 所得的结果实际上是一个**概率分布**: 测量结果为 λ_k 的概率为 $p_{H,\lambda_k} = \frac{\|\mathbb{P}_{H,\lambda_k}\psi\|^2}{\|\psi\|^2}$. 于是, 所谓“观测量”, 也可以私下里悄悄理解为这样的函数: 输入 \mathbb{V} 中的非零向量, 则输出一个相应的概率分布.
3. 对于量子态 $|\psi\rangle$, 给零任意非零常数 $c \in \mathbb{C}$, 记 $|\psi'\rangle := c|\psi\rangle$. 也就是说, 矢量 $|\psi\rangle$ 与 $|\psi'\rangle$ 相差 (非零) 常数倍. 对于这样两个相差常数倍的量子态, **我们无法通过观测来区分它们**. 具体地说, 对任何观测量 H , $|\psi\rangle$ 与 $|\psi'\rangle$ 被 H 测量都表现出相同的概率分布 [留给读者验证].
4. 接上条, 于是, 我们把相差非零常数倍的量子态认为是同一个量子态. 对于量子态 $|\psi\rangle$, 我们考虑量子态 $|\tilde{\psi}\rangle := \frac{1}{\|\psi\|}|\psi\rangle$, 则 $|\tilde{\psi}\rangle$ 与 $|\psi\rangle$ 相差常数倍, 从而是与 $|\psi\rangle$ 相同的量子态. 注意 $\|\tilde{\psi}\| = 1$. 矢量 $|\tilde{\psi}\rangle$ 称为 $|\psi\rangle$ 的**归一化**. 综上所述, 对于量子态 $|\psi\rangle$, 我们总可以不妨假设其模长等于 1, 即 $\|\psi\| = 1$.
5. 注意量子态 $|\psi\rangle$ 生活在**复线性空间**. 满足归一化条件 $\|\psi\| = 1$ 且与 $|\psi\rangle$ 表示相同量子态的矢量仍有很多. 事实上, 对任意实数 $\theta \in \mathbb{R}$, 考虑 $|\psi'\rangle := e^{i\theta}|\psi\rangle$, 则 $|\psi'\rangle$ 与 $|\psi\rangle$ 代表相同的量子态, 并且它们的模长相等. 注意 $e^{i\theta}$ 是模长为 1 的复数, 通常称为**相因子**. 以后会知道, 相因子在量子力学中会起到很微妙的作用.
6. 最后注意此公理的第 3 条. 被 H 观测之后, 该系统的量子态从 $|\psi\rangle$ “变化” 为它到与观测结果相应的子空间上的投影. 可见, 量子力学中的“观测” 与经典的观测很不一样. 在经典世界里, 黄色豌豆永远是黄豌豆, 无论怎么观测, 观测后依然是黄豌豆; 而量子世界里, “观测” 是有关仪器与量子系统的相互作用, 会对量子系统产生影响. 量子态从 $|\psi\rangle$ 到 $\mathbb{P}_{H,\lambda_k}|\psi\rangle$ 的变化过程称为**坍缩**. 在“古典量子力学” 里, 我们姑且认为坍缩过程是瞬间完成的³³.

我们以“量子豌豆” 为例, 更加具体地说明上述公理. 我们回忆, 量子豌豆系

³³然而, 任何物理过程都需要经过一定的时间. 从 $|\psi\rangle$ 坍缩为 $\mathbb{P}_{H,\lambda_k}|\psi\rangle$ 中间经历了怎样的过程, 这超出了本书范围.

统的希尔伯特空间 $\mathbb{V} = \mathbb{C}^2$, \mathbb{V} 的两组么正基满足转换关系(3.3)-(3.4). 我们约定,

$$\text{黄} = 1, \quad \text{绿} = -1;$$

$$\text{圆} = 1, \quad \text{皱} = -1,$$

即“豌豆颜色的观测结果为 1”意为“豌豆是黄色的”, 其余同理.

我们关心的两个观测量, 颜色与形状, 分别简记为色, 形; 它们都是希尔伯特空间 $\mathbb{V} = \mathbb{C}^2$ 上的厄米特算子:

$$\begin{aligned} \text{色} &= |\text{黄}\rangle\langle\text{黄}| - |\text{绿}\rangle\langle\text{绿}|, \\ \text{形} &= |\text{圆}\rangle\langle\text{圆}| - |\text{皱}\rangle\langle\text{皱}|. \end{aligned} \tag{3.10}$$

我们先看色. 由正交关系(3.1)容易验证,

$$\begin{aligned} \text{色}|\text{黄}\rangle &= |\text{黄}\rangle\langle\text{黄}|\text{黄}\rangle - |\text{绿}\rangle\langle\text{绿}|\text{黄}\rangle \\ &= |\text{黄}\rangle. \end{aligned}$$

同理, 容易验证

$$\text{色}|\text{绿}\rangle = -|\text{绿}\rangle.$$

也就是说, 厄米特算子 色 具有本征值 1 与 -1 , $|\text{黄}\rangle$ 是 色 的属于本征值 1 的本征矢量, $|\text{绿}\rangle$ 是 色 的属于本征值 -1 的本征矢量. 本征值 ± 1 都是无简并的. 相应的本征子空间

$$\mathbb{V}_{\text{色},1} = \mathbb{C}|\text{黄}\rangle, \quad \mathbb{V}_{\text{色},-1} = \mathbb{C}|\text{绿}\rangle,$$

满足

$$\mathbb{V} = \mathbb{V}_{\text{色},1} \oplus \mathbb{V}_{\text{色},-1}.$$

容易验证, 相应的投影算子

$$\mathbb{P}_{\text{色},1} = |\text{黄}\rangle\langle\text{黄}|, \quad \mathbb{P}_{\text{色},-1} = |\text{绿}\rangle\langle\text{绿}|. \tag{3.11}$$

同样地, 容易验证 形 具有本征值 ± 1 , 相应的本征子空间分解为

$$\mathbb{V} = \mathbb{V}_{\text{形},1} \oplus \mathbb{V}_{\text{形},-1},$$

其中

$$\mathbb{V}_{\text{形},1} = \mathbb{C}|\text{圆}\rangle, \quad \mathbb{V}_{\text{形},-1} = \mathbb{C}|\text{皱}\rangle;$$

相应的投影算子

$$\mathbb{P}_{\text{形},1} = |\text{圆}\rangle\langle\text{圆}|, \quad \mathbb{P}_{\text{形},-1} = |\text{皱}\rangle\langle\text{皱}|.$$

现在, 假设“量子豌豆”的态 $|\psi\rangle = |\text{黄}\rangle$.

- 根据有关公理, 豌豆颜色的观测结果为 1, -1 [即黄色, 绿色] 的概率分别为

$$p_{\text{色},1} = \|\mathbb{P}_{\text{色},1}|\text{黄}\rangle\|^2 = 1, \quad p_{\text{色},-1} = \|\mathbb{P}_{\text{色},-1}|\text{黄}\rangle\|^2 = 0.$$

也就是说, 颜色的观测结果是确定的, 就是 1(黄色). 观测后豌豆的状态坍缩为

$$\mathbb{P}_{\text{色},1}|\text{黄}\rangle = |\text{黄}\rangle,$$

与观测前没有变化.

- 有意思的问题是, 对于量子豌豆 $|\text{黄}\rangle$, 观测其形状会得到怎样的结果呢? 注意转换关系(3.3), 以及正交关系(3.2), 可知

$$\begin{aligned} \mathbb{P}_{\text{形},1}|\text{黄}\rangle &= |\text{圆}\rangle\langle\text{圆}|\text{黄}\rangle \\ &= \frac{1}{\sqrt{2}}|\text{圆}\rangle\langle\text{圆}|\text{圆}\rangle + \frac{1}{\sqrt{2}}|\text{圆}\rangle\langle\text{圆}|\text{皱}\rangle \\ &= \frac{1}{\sqrt{2}}|\text{圆}\rangle. \end{aligned}$$

同理, 也有 $\mathbb{P}_{\text{形},-1}|\text{黄}\rangle = \frac{1}{\sqrt{2}}|\text{皱}\rangle$. 于是, 形状的观测结果为 1 [即圆形] 的概率为

$$p_{\text{形},1} = \|\mathbb{P}_{\text{形},1}|\text{黄}\rangle\|^2 = \left\|\frac{1}{\sqrt{2}}|\text{圆}\rangle\right\|^2 = \frac{1}{2};$$

若形状的观测结果为 1 [即圆形], 则观测后量子豌豆的态坍缩为 $\mathbb{P}_{\text{形},1}|\text{黄}\rangle = \frac{1}{\sqrt{2}}|\text{圆}\rangle$. 注意我们把相差常数倍的矢量认为是同一个量子态, 于是不妨将坍缩后的态归一化. 于是观测后的态为 $|\text{圆}\rangle$. 同样方法计算可知, 形状的观测结果为 -1 [皱] 的概率, $p_{\text{形},-1} = \frac{1}{2}$; 若观测结果为皱, 则豌豆的态坍缩为 $|\text{皱}\rangle$ [同样做了归一化处理].

可见, 若豌豆的量子态为 $|\text{黄}\rangle$, 则豌豆的颜色完全确定, 100% 的概率检测结果为黄, 而且“经得起反复检测”. 但这个豌豆的形状就“完全不确定”了:

$$|\text{黄}\rangle = \frac{1}{\sqrt{2}}|\text{圆}\rangle + \frac{1}{\sqrt{2}}|\text{皱}\rangle,$$

量子态 $|\text{黄}\rangle$ 其实是 $|\text{圆}\rangle$ 与 $|\text{皱}\rangle$ 的线性叠加, 即所谓**叠加态**. 常见科普书里会说, 在观测其形状之前, 这个豌豆是“又圆又皱”的叠加态³⁴.

- 回忆量子豌豆故事, 我们观测“黄豌豆”(量子态 $|\text{黄}\rangle$) 的形状, 挑选出观测结果为圆的豌豆, 然后再重新观测其颜色. 观测结果是, 一半黄一半绿. 请读者用量子力学公理来解释此现象. [提示: 我们挑选出来的观测结果为圆的豌豆, 其量子态从观测前的 $|\text{黄}\rangle$ 坍缩为 $|\text{圆}\rangle$. 接下来“重新观测颜色”, 观测的其实是量子态为 $|\text{圆}\rangle$ 的豌豆的颜色.]

至此, 我们用目前 3 条量子力学公理解释了“量子豌豆”现象.

例题 3.1. 沿用本节“量子豌豆”的记号. 已知“量子豌豆”的量子态

$$|\psi\rangle = \frac{1}{\sqrt{2}}|\text{圆}\rangle + \frac{e^{i\theta}}{\sqrt{2}}|\text{皱}\rangle,$$

其中 θ 为给定实数. 观测其颜色, 试分别求观测结果为黄, 绿的概率.

解. 容易验证 $|\psi\rangle$ 的模长等于 1, 从而观测结果为黄, 绿的概率为 $|\psi\rangle$ 被相应投影算子(3.11) 作用所得矢量的模长平方. 结合(3.2)-(3.4), 简单计算可得

$$\begin{aligned}\mathbb{P}_{\text{色},1}|\psi\rangle &= |\text{黄}\rangle\langle\text{黄}|\left(\frac{1}{\sqrt{2}}|\text{圆}\rangle + \frac{e^{i\theta}}{\sqrt{2}}|\text{皱}\rangle\right) \\ &= \frac{1+e^{i\theta}}{2}|\text{黄}\rangle \\ \mathbb{P}_{\text{色},-1}|\psi\rangle &= |\text{绿}\rangle\langle\text{绿}|\left(\frac{1}{\sqrt{2}}|\text{圆}\rangle + \frac{e^{i\theta}}{\sqrt{2}}|\text{皱}\rangle\right) \\ &= \frac{1-e^{i\theta}}{2}|\text{绿}\rangle.\end{aligned}$$

从而观测结果为黄色的概率 $p_{\text{色},1}$ 为:

$$p_{\text{色},1} = \|\mathbb{P}_{\text{色},1}|\psi\rangle\|^2 = \left|\frac{1+e^{i\theta}}{2}\right|^2 = \frac{1+\cos\theta}{2}.$$

类似计算可知观测结果为绿色的概率 $p_{\text{色},-1} = \left|\frac{1-e^{i\theta}}{2}\right|^2 = \frac{1-\cos\theta}{2}$. □

³⁴更“哗众取宠”的说法是众所周知的所谓**薛定谔的猫**: 在观测猫是死是活之前, 这只猫是“又死又活”的叠加态. 喵喵喵 ~

3.3 本征态, 期望与方差, 不确定原理

本节我们更加深入地探讨量子观测. 给定量子系统 \mathbb{V} (希尔伯特空间), 以及观测 H (希尔伯特空间 V 上的厄米特算子). 方便起见, 暂时只考虑 \mathbb{V} 是有限维空间的情况, 事实上本节的结论都能够类似照搬到无限维空间³⁵. 记观测 H 的本征值之全体 $\sigma(H) = \{\lambda_1, \dots, \lambda_m\}$, 相应的本征子空间记为 $\mathbb{V}_{H, \lambda_k}$, $1 \leq k \leq m$.

本征态与叠加态.

如果量子态 $|\psi_k\rangle$ 满足 $H|\psi_k\rangle = \lambda_k|\psi_k\rangle$, 换言之 $|\psi_k\rangle$ 是 H 的属于本征值 λ_k 的本征矢量, 则由厄米特算子的性质 [不同本征值的本征子空间两两正交] 以及量子公理 3, 容易验证:

- $|\psi_k\rangle$ 被 H 观测的结果为 λ_k 的概率是 1,
- $|\psi_k\rangle$ 被 H 观测的结果为 λ_l , ($l \neq k$) 的概率是 0.

也就是说, 这样的量子态被 H 观测的结果是“完全确定”的; 此外还容易验证, 被 H 观测后 [即, 坍缩后] 该系统的状态仍为 $|\psi_k\rangle$. 我们把这样的量子态称为关于 H 的**本征态**. 总结如下:

定义. 给定量子系统 \mathbb{V} 的观测 H . 如果量子态 $|\psi\rangle$ 是厄米特算子 H 的属于本征值 λ_k 的本征矢量, 则称 $|\psi\rangle$ 是 H 的关于 λ_k 的**本征态**.

注意 \mathbb{V} 可分解为 H 的本征子空间的正交直和 [回忆(3.5)式]:

$$\mathbb{V} = \mathbb{V}_{H, \lambda_1} \oplus \mathbb{V}_{H, \lambda_2} \oplus \cdots \oplus \mathbb{V}_{H, \lambda_m},$$

在此分解下, 任何量子态 $|\psi\rangle$ 都能够唯一地分解为

$$|\psi\rangle = |\psi_1\rangle + |\psi_2\rangle + \cdots + |\psi_m\rangle,$$

使得 $|\psi_k\rangle \in \mathbb{V}_{H, \lambda_k}$, 即 $|\psi_k\rangle$ 是 H 的关于 λ_k 的本征态. 换言之, 任何量子态都必形如 H 的属于不同本征值的本征态的线性组合. 正因为如此, 如果量子态 $|\psi\rangle$ 不是关于 H 的本征态, 则称 $|\psi\rangle$ 是关于 H 的**叠加态**.

³⁵当然会略微牺牲一些数学上的严谨性. 严格地处理无限维希尔伯特空间需要用到较多**泛函分析**知识, 这超出本书范围.

考虑 H 无简并的特殊情形³⁶. 此时, H 具有 $m = n := \dim \mathbb{V}$ 个两两不同的本征值 $\lambda_1, \dots, \lambda_n$, 每个本征子空间 $\mathbb{V}_{H, \lambda_k}$ 的维数都是 1. 对每个 $1 \leq k \leq n$, 取定 $\mathbb{V}_{H, \lambda_k}$ 中的一个模长为 1 的矢量, 记作 $|k\rangle$; 则 $H|k\rangle = \lambda_k|k\rangle$, $\mathbb{V}_{H, \lambda_k} = \mathbb{C}|k\rangle$; 并且

$$\{|1\rangle, |2\rangle, \dots, |n\rangle\}$$

构成 \mathbb{V} 的一组么正基. 任何量子态 $|\psi\rangle$ 都可唯一写成

$$|\psi\rangle = c_1|1\rangle + c_2|2\rangle + \dots + c_n|n\rangle,$$

其中 c_1, \dots, c_n 是不全为零的复常数; 不妨对 $|\psi\rangle$ 作归一化处理, 要求 $\|\psi\|^2 = |c_1|^2 + |c_2|^2 + \dots + |c_n|^2 = 1$. 此时, 观测结果为 λ_k 的概率

$$p_{H, \lambda_k} = |c_k|^2. \quad (3.12)$$

例如, 在上一节“量子豌豆”例子里, $|\text{黄}\rangle, |\text{绿}\rangle$ 是观测量 **色** 的本征态; 而注意 $|\text{黄}\rangle = \frac{1}{\sqrt{2}}|\text{圆}\rangle + \frac{1}{\sqrt{2}}|\text{皱}\rangle$, 从而 $|\text{黄}\rangle$ 是观测量 **形** 的叠加态. 一般地, 设 H_1, H_2 是量子系统 \mathbb{V} 的两个观测量, 则 H_1 的本征态未必是 H_2 的本征态, 同样地, H_2 的本征态在 H_1 看来也可能是叠加态.

期望与方差

设 H 是量子系统 \mathbb{V} 的观测量, $|\psi\rangle$ 为某个量子态, 不妨 $\|\psi\|^2 = 1$. 则 $|\psi\rangle$ 被 H 测量得到结果 λ_k , ($k = 1, \dots, m$) 的概率

$$p_{H, \lambda_k} = \|\mathbb{P}_{H, \lambda_k}|\psi\rangle\|^2 = \langle\psi|\mathbb{P}_{H, \lambda_k}|\psi\rangle,$$

特别注意投影算子 $\mathbb{P}_{H, \lambda_k}$ 满足

$$\sum_{k=1}^m \mathbb{P}_{H, \lambda_k} = I.$$

一个自然的问题是, $|\psi\rangle$ 被 H 观测的结果的**平均值**是多少? 回忆初等概率论的知识, 我们不妨把 $|\psi\rangle$ 视为离散型随机变量: 取到 λ_k 的概率为 p_{H, λ_k} ³⁷; 于是, 所谓“平均值”, 其实就是随机变量的期望. 事实上, 容易证明:

³⁶其实是“常见”情形

³⁷在无穷维希尔伯特空间的情形, 看情况视为连续性随机变量.

性质 3.2. 归一化量子态 $|\psi\rangle$ 被观测量 H 观测的结果的期望为

$$\langle H \rangle := \langle \psi | H | \psi \rangle.$$

这里“归一化量子态”即模长等于 1 的矢量.

证明. 注意 $\mathbb{P}_{H,\lambda_k}|\psi\rangle$ 是 H 的属于本征值 λ_k 的本征矢量, 即 $H\mathbb{P}_{H,\lambda_k}|\psi\rangle = \lambda_k\mathbb{P}_{H,\lambda_k}|\psi\rangle$, 从而 $|\psi\rangle$ 被 H 观测的期望值

$$\begin{aligned} \langle H \rangle &= \sum_{k=1}^m \lambda_k p_{H,\lambda_k} = \sum_{k=1}^m \lambda_k \langle \psi | \mathbb{P}_{H,\lambda_k} | \psi \rangle \\ &= \langle \psi | \sum_{k=1}^m \lambda_k \mathbb{P}_{H,\lambda_k} | \psi \rangle = \langle \psi | \sum_{k=1}^m H \mathbb{P}_{H,\lambda_k} | \psi \rangle \\ &= \langle \psi | H \left(\sum_{k=1}^m \mathbb{P}_{H,\lambda_k} \right) | \psi \rangle = \langle \psi | H | \psi \rangle. \end{aligned}$$

□

为了说话省事, 我们常把“ $|\psi\rangle$ 被 H 观测的结果的期望”简单说成“观测量 H 的期望”, 并记为 $\langle H \rangle$. 特别地, 观测结果为 λ_k 的概率 $p_{H,\lambda_k} = \langle \psi | \mathbb{P}_{H,\lambda_k} | \psi \rangle$ 恰为投影算子 \mathbb{P}_{H,λ_k} 的期望.

在初等概率论中我们还关心随机变量的**方差**, 即随机变量与其期望的差的平方的平均值, 它反映了随机变量的“不确定程度”.

同样地, 对于量子态 $|\psi\rangle$, 我们也用方差来描述其被 H 观测的结果的不确定程度. 被 H 观测的结果的方差 $(\Delta H)^2$ 满足

$$\begin{aligned} (\Delta H)^2 &= \sum_{k=1}^m p_{H,\lambda_k} (\lambda_k - \langle H \rangle)^2 \\ &= \langle \psi | (H - \langle H \rangle)^2 | \psi \rangle \\ &= \langle H^2 \rangle - \langle H \rangle^2. \end{aligned}$$

我们回忆, 方差 $(\Delta H)^2$ 的算术平方根 ΔH 称为**标准差**.

海森堡不确定原理.

我们早在“量子豌豆”中就意识到, 我们无法同时确定量子豌豆的颜色与形状: 当颜色确定, 其形状就会“很随机”, 当形状确定, 颜色则会“很随机”. 事实上, 这是量子力学中的普遍规律, 一般来说不能使得两个不同的观测量 H_1, H_2 的观测结果“都非常确定”.

设 H_1, H_2 是量子系统 \mathbb{V} 的两个观测量, 给定量子态 $|\psi\rangle$. 我们记

$$\hat{H}_1 := H_1 - \langle H_1 \rangle, \quad \hat{H}_2 := H_2 - \langle H_2 \rangle, \quad (3.13)$$

注意 \hat{H}_1, \hat{H}_2 也是厄米特算子. 则 $|\psi\rangle$ 被 H_i , ($i = 1, 2$) 观测的结果的方差

$$\begin{aligned} (\Delta H_i)^2 &= \langle \hat{H}_i^2 \rangle = \langle \psi | \hat{H}_i^2 | \psi \rangle = \langle \psi | \hat{H}_i^\dagger \hat{H}_i | \psi \rangle \\ &= \langle \hat{H}_i \psi | \hat{H}_i \psi \rangle = \|\hat{H}_i \psi\|^2. \end{aligned}$$

于是, 利用施瓦茨不等式(1.6)可得

$$(\Delta H_1)^2 (\Delta H_2)^2 = \|\hat{H}_1 \psi\|^2 \|\hat{H}_2 \psi\|^2 \geq \left| \langle \hat{H}_1 \psi | \hat{H}_2 \psi \rangle \right|^2 = \left| \langle \psi | \hat{H}_1 \hat{H}_2 | \psi \rangle \right|^2, \quad (3.14)$$

取到等号当且仅当 $\hat{H}_1 |\psi\rangle$ 与 $\hat{H}_2 |\psi\rangle$ 线性相关.

特别注意, 两个厄米特算子的乘积 $\hat{H}_1 \hat{H}_2$ 未必是厄米特算子, 相应地 $\langle \psi | \hat{H}_1 \hat{H}_2 | \psi \rangle$ 未必是实数, 计算其模长平方需要分别考虑它的实部与虚部. 自然考虑如下分解:

$$\begin{aligned} \hat{H}_1 \hat{H}_2 &= \frac{\hat{H}_1 \hat{H}_2 + \hat{H}_2 \hat{H}_1}{2} + \frac{\hat{H}_1 \hat{H}_2 - \hat{H}_2 \hat{H}_1}{2} \\ &= \frac{1}{2} [\hat{H}_1, \hat{H}_2]_+ + \frac{1}{2} [\hat{H}_1, \hat{H}_2]. \end{aligned} \quad (3.15)$$

这里的 $[\hat{H}_1, \hat{H}_2]_+ := \hat{H}_1 \hat{H}_2 + \hat{H}_2 \hat{H}_1$, 称为 \hat{H}_1 与 \hat{H}_2 的**反对易子**; 而 $[\hat{H}_1, \hat{H}_2]$ 恰为通常的对易子 [见前文(1.11)式].

以下性质留给读者验证, 作为简单习题:

- $[\hat{H}_1, \hat{H}_2]_+$ 是厄米特算子, $[\hat{H}_1, \hat{H}_2]$ 是反厄米特算子.
- $[\hat{H}_1, \hat{H}_2] = [H_1, H_2]$.

记 $[H_1, H_2] =: i\Gamma$, $[i$ 为虚数单位], 则 Γ 是厄米特算子. 我们继续化简(3.14)的右端. 易知

$$\langle \psi | \hat{H}_1 \hat{H}_2 | \psi \rangle = \frac{1}{2} \langle \psi | [\hat{H}_1, \hat{H}_2]_+ | \psi \rangle + \frac{i}{2} \langle \psi | \Gamma | \psi \rangle.$$

注意 $[\hat{H}_1, \hat{H}_2]_+$ 与 Γ 都是厄米特算子, 所以 $\langle \psi | [\hat{H}_1, \hat{H}_2]_+ | \psi \rangle$ 与 $\langle \psi | \Gamma | \psi \rangle$ 都是实数, 因此复数 $\langle \psi | \hat{H}_1 \hat{H}_2 | \psi \rangle$ 的实部和虚部分别为 $\frac{1}{2} \langle \psi | [\hat{H}_1, \hat{H}_2]_+ | \psi \rangle$ 与 $\frac{1}{2} \langle \psi | \Gamma | \psi \rangle$. 因此,

$$\begin{aligned} (\Delta H_1)^2 (\Delta H_2)^2 &\geq \left| \langle \psi | \hat{H}_1 \hat{H}_2 | \psi \rangle \right|^2 \\ &= \frac{1}{4} \langle \psi | [\hat{H}_1, \hat{H}_2]_+ | \psi \rangle^2 + \frac{1}{4} \langle \psi | \Gamma | \psi \rangle^2 \\ &= \frac{1}{4} \langle [\hat{H}_1, \hat{H}_2]_+ \rangle^2 + \frac{1}{4} \langle \Gamma \rangle^2. \end{aligned}$$

综上所述, 我们立刻得到著名的**海森堡不确定原理**:

定理 3.3. 设 H_1, H_2 是量子系统 \mathbb{V} 的两个观测量, 则对该系统的任何量子态 $|\psi\rangle$, 成立

$$\Delta H_1 \cdot \Delta H_2 \geq \frac{1}{2} |\langle \Gamma \rangle|, \quad (3.16)$$

其中 $[H_1, H_2] = i\Gamma$; 上述不等式取到等号当且仅当以下同时成立:

1. $\hat{H}_1 |\psi\rangle$ 与 $\hat{H}_2 |\psi\rangle$ 线性相关;
2. $\langle [\hat{H}_1, \hat{H}_2]_+ \rangle = 0$.

若 $i\Gamma = [H_1, H_2] \neq 0$, 则 $|\langle \Gamma \rangle|$ 通常是正实数, 于是由此定理可知, ΔH_1 与 ΔH_2 不可能“都很小”. 这个定理常用于对某些物理量作估计, 我们以后会介绍它的各种应用.

若 $[H_1, H_2] = 0$, 即观测量 H_1, H_2 **对易**, 则由厄米特算子的性质 [性质1.9] 可知, 存在 \mathbb{V} 的一组基 $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$, 使得每个 $|k\rangle$, $1 \leq k \leq n$ 都满足: $|k\rangle$ 是 H_1 的本征态, 且 $|k\rangle$ 是 H_2 的本征态. 也就是说, 对于量子态 $|k\rangle$, H_1, H_2 的观测结果能够同时确定, $\Delta H_1 = \Delta H_2 = 0$.

3.4 量子态随时间的演化：薛定谔方程

我们已经学习了如何描述, 观测量子态; 但最核心的问题尚未回答:

一个系统的状态如何随时间流逝而变化?

无论是经典力学还是量子力学, 这都是关键的问题. 哈密顿力学认为, **哈密顿量** \mathcal{H} [即相空间 \mathbb{R}^{2n} 上的某个给定的函数] 决定力学系统的演化, 见(2.22),(2.23)式:

$$\begin{aligned}\frac{dq_i}{dt} &= \{q_i, \mathcal{H}\} \\ \frac{dp_i}{dt} &= \{p_i, \mathcal{H}\}.\end{aligned}$$

此外, 对相空间上的任意函数 $\omega(\mathbf{q}, \mathbf{p})$, 当 (\mathbf{q}, \mathbf{p}) 随时间 t 变化时, $\omega(t) := \omega(\mathbf{q}(t), \mathbf{p}(t))$ 自然视为关于时间 t 的函数, 即“力学量 ω 随时间演化”. 我们回忆, $\omega(t)$ 满足(2.15)式:

$$\frac{d\omega}{dt} = \{\omega, \mathcal{H}\}.$$

而量子力学³⁸ 的相应理论可以看成对哈密顿力学的模仿类比.

普朗克常数与薛定谔方程.

哈密顿力学中, 系统的演化是被某个特定的力学量 \mathcal{H} , 即哈密顿量, 所支配. 类比到量子力学中, 量子系统 \mathbb{V} 中的量子态随时间的演化的也应该被某个特定的观测量 H 所支配, 这个 H 是希尔伯特空间 \mathbb{V} 上的厄米特算子, 称为量子系统 \mathbb{V} 的**哈密顿量**.

再回看哈密顿力学. 对于相空间 \mathbb{R}^{2n} 上的函数 ω , 映射 $\omega \mapsto \{\omega, \mathcal{H}\}$ 是关于 ω 的线性映射. 换言之, 哈密顿量 \mathcal{H} 在演化方程中扮演了线性算子 $\{-, \mathcal{H}\}$ 的角色. 于是比较自然的类比是, 在量子力学的演化方程里, 哈密顿量 H 也线性地作用于量子态上.

当然以上“类比”都是在瞎猜, 量子态随时间演化的方程到底是什么样, 最终由实验结果说得算. 事实上, 经过大量实验验证, 我们相信以下成立:

³⁸本讲义的“量子力学”仅仅是其中一种对初学者友好的入门理论. 量子力学与经典力学一样, 有多种不同理论.

- **量子公理 4.** 设 H 是量子系统 \mathbb{V} 的哈密顿量, 则该系统的量子态 $|\psi(t)\rangle$ 随时间 t 的演化满足方程

$$i\hbar \frac{d}{dt}|\psi\rangle = H|\psi\rangle. \quad (3.17)$$

方程(3.17)即著名的**薛定谔方程**, 其中 $\hbar \approx 6.62 \times 10^{-34} \text{J} \cdot \text{s}$, 称为**普朗克常数**; \hbar 与光速 c , 万有引力常数 G 一样, 被认为是最基本的物理学常数. [其实暂时先不用太纠结这个 \hbar 的物理含义, 我们只需要把他当成一个很小的正数, 甚至把它当成一个形式上的符号就可以.]

有意思的是, 薛定谔方程等号左边出现了虚数单位 i , 这是量子力学中最令人困惑的地方之一³⁹. 我们习惯把 “ $i\hbar \frac{d}{dt}$ ” 看作一个整体.

由薛定谔方程可知, 经过一段很短的时间 Δt , 则量子态的微小变化 $|\Delta\psi\rangle := |\psi(t + \Delta t)\rangle - |\psi(t)\rangle$ 满足

$$|\Delta\psi\rangle \approx \left(-\frac{i}{\hbar}\Delta t\right) H|\psi(t)\rangle, \quad \Delta t \rightarrow 0,$$

也就是说, 量子态的微小变化正比于量子态被线性算子 H 的作用. 我们说, 哈密顿量 H 是量子态随时间演化的**无穷小生成元**.

注记

据说, 薛定谔方程(3.17) 最初就是在经典力学的启发类比之下被 “连猜带凑” 蒙出来的.

求解薛定谔方程

设 H 是量子系统 \mathbb{V} 的哈密顿量, 我们来求解薛定谔方程(3.17). 回忆(1.47)式附近, 可知

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}tH}|\psi(0)\rangle \quad (3.18)$$

³⁹毕竟绝大多数正常人类都非常抵触虚数 i , 认为这东西没有实际意义.

是薛定谔方程 (3.17) 的满足初始条件 $|\psi\rangle|_{t=0} = |\psi(0)\rangle$ 的解.

我们将上式中出现的关于算子 H 的指数映射 $\exp\left(-\frac{i}{\hbar}tH\right)$ 记为 $U(t)$, 即

$$U(t) := \exp\left(-\frac{i}{\hbar}tH\right), \quad (3.19)$$

我们称 $U(t)$ 为哈密顿量 H 的 **传播子**(propogator). 于是, 薛定谔方程的解可改写为

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle. \quad (3.20)$$

注意, 由性质 1.10 的第 4 条可知, 传播子 $U(t)$ 是希尔伯特空间 \mathbb{V} 上的**酉算子**.

注记

我们曾在(1.43)式附近, 对传播子做了一个粗俗暴力且牵强的经典力学类比.

在量子力学中, 由于传播子 $U(t)$ 是酉算子, 从而矢量 $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ 与 $|\psi(0)\rangle$ 的模长始终相等, 换言之, 量子态 $|\psi\rangle$ 在随时间演化过程中始终保持模长不变. 特别地, 若初始态 $|\psi\rangle$ 是归一化量子态 (即模长等于 1), 则该量子态在演化过程中始终都是归一化的. [我们常把复线性空间里的酉算子类比为实线性空间里的正交算子, 而正交算子的几何意义是“(高维) 空间里的旋转”; 于是我们也偷偷地把酉算子想象成复线性空间里的“旋转”. 于是, 量子态随时间演化 $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ 的过程可以偷偷想成“矢量 $|\psi\rangle$ 在旋转”]

注记

薛定谔方程里的虚数单位 i 起到了什么作用? 正是有了 i , 才保证了传播子 $U(t)$ 为酉算子, 进而量子态随时间的演化具有“看起来比较舒服”的几何解释 (貌似是旋转).

为方便讨论, 先考虑 \mathbb{V} 是有限维空间的情况. 注意哈密顿量 H 是厄米特算子, 从而存在 \mathbb{V} 的由 H 的本征矢量构成的一组幺正基 $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$, 其中

$H|k\rangle = \lambda_k|k\rangle$, 本征值 λ_k 都是实数. 则算子 H 在该基下的矩阵为对角阵:

$$H = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

相应的传播子在该基下的矩阵为

$$U(t) = e^{-\frac{i}{\hbar}tH} = \begin{pmatrix} e^{-\frac{i}{\hbar}\lambda_1 t} & & \\ & \ddots & \\ & & e^{-\frac{i}{\hbar}\lambda_n t} \end{pmatrix}.$$

若在该量子系统中, $t = 0$ 时刻的量子态在该基下展开为

$$|\psi(0)\rangle = \sum_{k=1}^n c_k |k\rangle,$$

则经过时间 t 后的量子态为

$$|\psi(t)\rangle = \sum_{k=1}^n c_k e^{-\frac{i}{\hbar}\lambda_k t} |k\rangle. \quad (3.21)$$

而对于 \mathbb{V} 是无限维空间的情形, 上述结果也都有类似版本. 等我们遇到无限维空间的时候再说.

薛定谔图景与海森堡图景

我们要注意, 希尔伯特空间 \mathbb{V} , 量子态矢量 $|\psi\rangle$ 等东西都是为了解释现实世界中的实验现象而引入的抽象概念. 物理学以实验为基础, 我们只相信我们能观测到的实验现象. 换言之, 我们如何确定量子系统处于什么状态? 我们是通过观测该系统的各种不同的物理量 [例如观察豌豆的颜色, 形状, 以及其他可能的别的性质] 来确定.

我们用抽象的希尔伯特空间那套语言来企图解释量子现象, 把量子态视为希尔伯特空间 \mathbb{V} 中的矢量. 而量子态随时间的演化 $|\psi(t)\rangle$, 也就是“矢量 $|\psi\rangle$ 在 \mathbb{V} 中随时间而运动”. 换言之, 量子系统随时间演化, 是指量子态 $|\psi\rangle$ 随时间的变化 $t \mapsto |\psi(t)\rangle$.

以上是关于量子态随时间演化的观点之一. 除此之外, 还有另一种观点. 我们是通过量子系统进行**观测**并查看观测结果, 来确定量子系统处于什么状态. 量子系统随时间变化, 则对该系统的观测结果相应也随时间变化. 对于给定的观测量 Ω [注意 Ω 是希尔伯特空间 \mathcal{V} 上的厄米特算子], 我们回忆观测量 Ω 的期望为

$$\langle \Omega \rangle = \langle \psi | \Omega | \psi \rangle, \quad (3.22)$$

(不妨 $|\psi\rangle$ 是归一化量子态). 观测量的期望 $\langle \Omega \rangle$ 是我们真正“看得见摸得着”, 能通过实验来直观感受的对象.

现在, $|\psi\rangle = |\psi(t)\rangle$ 随时间演化, 则观测量 Ω 的期望也随时间变化, 满足:

$$\langle \Omega \rangle(t) := \langle \psi(t) | \Omega | \psi(t) \rangle = \left\langle \psi \left| U(t)^\dagger \Omega U(t) \right| \psi \right\rangle \quad (3.23)$$

我们把等号最右边出现的算子记作 $\Omega(t)$, 即:

$$\Omega(t) := U(t)^\dagger \Omega U(t), \quad (3.24)$$

注意 $\Omega(t)$ 依然是厄米特算子; 由于 $U(t)$ 为酉算子, $U(t)^\dagger = U(t)^{-1}$, 从而 $\Omega(t)$ 与 Ω 相似, 故具有相同的本征值.

注意, (3.23)等号最右边可以视为算子 $\Omega(t)$ 的期望. 于是, 我们得到另一种观点: 量子系统随时间演化, 是指观测量 Ω 随时间的变化 $t \mapsto \Omega(t)$.

以上两种观点分别称为 **薛定谔图景** 与 **海森堡图景**, 我们粗糙地总结如下:

$$\begin{cases} \text{薛定谔图景:} & \text{观测量 } \Omega \text{ 不变, 量子态 } |\psi\rangle \text{ 变化.} \\ \text{海森堡图景:} & \text{观测量 } \Omega \text{ 变化, 量子态 } |\psi\rangle \text{ 不变.} \end{cases}$$

上述两种看待量子系统演化的观点其实等价, 在计算具体问题时我们随机应变选择其中一种观点 [通常来说, 薛定谔图景的计算往往更简单些, 我们常取薛定谔图景; 但海森堡图景也有其可取之处].

回归经典: Ehrenfest 定理

经典物理规律适用于我们日常生活的空间尺度, 在我们无法直接感知的微观世界中发生的量子现象不适用于经典力学规律, 而是服从量子物理学的规律 (量子公理 1-4).

我们不希望经典物理理论与量子物理理论是两套截然不同的理论 [否则试想, 在“日常生活的世界”与“微观世界”的交界处的物理规律将暧昧不清]; 事实上, 经典力学理论是对真实物理规律的一种近似描述⁴⁰, 仅在“日常生活的世界”范围内才能较好地近似; 而量子理论的适用范围更广, 而且比经典理论更接近真实物理规律: 不仅适用于“微观世界”, 也应该能兼容“日常生活的世界”的经典理论, 即, 用量子理论研究“日常生活世界”所得出的结论应该与用经典物理理论所得出的结论非常接近, 不能相差太离谱.

也许有读者会疑惑: 量子理论中观测一个量子态的结果是概率分布, 甚至还有海森堡不确定原理, 量子理论的这种**随机性**与经典物理理论所描述的**确定性**有本质上的, 颠覆三观的不同呀? 这个嘛, 不能像文科生一样只会定性地看, 更要定量地看. 还是用豌豆讲个故事: 我们来研究经典世界的豌豆, 对于某个经典豌豆, 假如用经典物理理论研究它能得出结论

这个豌豆是黄色的;

那么用量子理论研究这个经典豌豆, 得出的结论可能会是

这个豌豆有 99.999 999 999 999 999 999 999% 的概率是黄色,
有 0.000 000 000 000 000 000 001% 的概率是绿色.

可见经典物理理论与量子理论所断言的结果不能说相同, 但足够非常接近. 这就是所谓量子理论“兼容”经典物理理论.

总之, 我们不应该认为量子理论与经典物理理论矛盾, 将经典物理理论推翻; 而可以认为量子理论是经典物理理论的“升级版”: 适用范围更广, 准确程度更高.

我们来考察量子理论与经典物理理论的联系. 设某个量子系统随时间演化, 则某个观测量 Ω 的观测结果也随时间变化. 而 Ω 的观测结果是一个概率分布, 为了与经典物理理论中的确定的观测结果比较, 我们考虑观测量 Ω 的期望 $\langle \Omega \rangle$.

⁴⁰人类提出的所有的高端的物理理论都是对真实物理规律的近似描述

定理 3.4. (*Ehrenfest 定理*) 设 H 是量子系统 \mathbb{V} 的哈密顿量, $|\psi(0)\rangle = |\psi\rangle$ 是该系统在 $t=0$ 时刻的量子态, Ω 是该系统的任意给定的观测量. 则观测量 Ω 在 t 时刻的期望 $\langle\Omega\rangle(t)$ 满足如下方程:

$$i\hbar \frac{d}{dt} \langle\Omega\rangle(t) = \langle[\Omega, H]\rangle(t). \quad (3.25)$$

上式右边的 $[\cdot, \cdot]$ 是算子的对易子, $\langle[\Omega, H]\rangle(t) := \langle\psi(t)|\Omega H - H\Omega|\psi(t)\rangle$.

证明. 对薛定谔方程(3.17)两边取伴随, 注意 $i^\dagger = i^* = -i$, 可得

$$-i\hbar \frac{d}{dt} \langle\psi(t)| = \langle\psi(t)|H^\dagger = \langle\psi(t)|H.$$

从而直接求导验证如下:

$$\begin{aligned} i\hbar \frac{d}{dt} \langle\Omega\rangle(t) &= i\hbar \frac{d}{dt} \left(\langle\psi(t)|\Omega|\psi(t)\rangle \right) \\ &= \left(i\hbar \frac{d}{dt} \langle\psi(t)| \right) \Omega|\psi(t)\rangle + \langle\psi(t)|\Omega \left(i\hbar \frac{d}{dt} |\psi(t)\rangle \right) \\ &= \langle\psi(t)|(-H)\Omega|\psi(t)\rangle + \langle\psi(t)|\Omega H|\psi(t)\rangle \\ &= \langle\psi(t)|[\Omega, H]|\psi(t)\rangle \\ &= \langle[\Omega, H]\rangle(t). \end{aligned}$$

□

上述 **Ehrenfest 定理** 中的公式(3.25)与哈密顿力学中的公式(2.15) 非常像! 只需要把哈密顿力学里的 $\frac{d}{dt}$ 换成 $i\hbar \frac{d}{dt}$, 然后把泊松括号 $\{\cdot, \cdot\}$ 换成对易子 $[\cdot, \cdot]$, 最后再加上求期望, 就能得到量子力学版本的方程.

注记

表面上看, 泊松括号 $\{\cdot, \cdot\}$ 与对易子 $[\cdot, \cdot]$ 一个来自微积分, 一个来自线性代数, 但它们具有相同的运算性质: 详见(2.16)-(2.19), 以及(1.12)-(1.15).

事实上, 我们希望把哈密顿力学里的每一个概念都找到其在量子力学里的某种“对应”, 大致来说, 这个过程叫做**量子化**. 我们将会在后文继续探讨量子化.

我们将目前所学内容对比总结如下:

哈密顿力学	量子力学
相空间 \mathbb{R}^{2n}	希尔伯特空间 \mathbb{V}
\mathbb{R}^{2n} 中的点 (\mathbf{q}, \mathbf{p})	\mathbb{V} 中的矢量 $ \psi\rangle$
\mathbb{R}^{2n} 上的函数 ω	\mathbb{V} 上的厄米特算子 Ω
哈密顿量 \mathcal{H}	哈密顿量 H
泊松括号 $\{\omega_1, \omega_2\}$	对易子 $[\Omega_1, \Omega_2]$
$\frac{d\omega}{dt} = \{\omega, \mathcal{H}\}$	$i\hbar \frac{d\langle\Omega\rangle}{dt} = \langle[\Omega, H]\rangle.$

数学补充: Baker-Hausdorff 公式

本节最后, 我们再稍微提一下海森堡图景下的计算. 在此观点下, 随时间变化的是观测量 $\Omega = \Omega(t)$, 它满足(3.24)式:

$$\Omega(t) := U(t)^\dagger \Omega U(t),$$

其中 $U(t) = e^{-\frac{i}{\hbar}tH}$ 为传播子. 注意 $U(t)$ 满足 $i\hbar \frac{d}{dt}U(t) = HU(t) = U(t)H$, 从而用与证明 Ehrenfest 定理类似的方法, 容易验证 $\Omega(t)$ 满足微分方程

$$\frac{d}{dt}\Omega(t) = [\Omega(t), H]. \quad (3.26)$$

[留给读者练习.]

将 $U(t) = e^{-\frac{i}{\hbar}tH}$ 代入(3.24)式, 可得

$$\Omega(t) = e^{\frac{i}{\hbar}tH} \Omega e^{-\frac{i}{\hbar}tH}. \quad (3.27)$$

而另一方面, 我们当然也可以通过求解微分方程(3.26)来得到 $\Omega(t)$ 的表达式. 为此, 我们不妨再强调一些线性代数的概念.

设 \mathbb{V} 为矢量空间, 记 $\mathcal{L}(\mathbb{V})$ 是 \mathbb{V} 上的所有的线性算子构成的集合. 注意两个线性算子可以相加, 线性算子可以和复数作数乘, 并且加法, 数乘显然都满足我们希望的好的性质, 从而 $\mathcal{L}(\mathbb{V})$ 也构成矢量空间, $\mathcal{L}(\mathbb{V})$ 中的矢量是 \mathbb{V} 上的线性算子. 对于 \mathbb{V} 上的线性算子 X (随便一个线性算子都行, 可以不是厄米特算子), 我们定

义 $\text{ad}_X \in \mathcal{L}(\mathcal{L}(\mathbb{V}))$, [即 ad_X 是矢量空间 $\mathcal{L}(\mathbb{V})$ 上的线性算子] 如下:

$$\begin{aligned}\text{ad}_X: \mathcal{L}(\mathbb{V}) &\rightarrow \mathcal{L}(\mathbb{V}) \\ Y &\mapsto [X, Y],\end{aligned}$$

容易验证 ad_X 确实是线性算子 [留做习题].

也就是说, 对于 \mathbb{V} 上的线性算子 X, Y ,

$$\text{ad}_X Y = [X, Y], \quad (3.28)$$

我们把对易子 $[X, Y]$ 看成是 ad_X 作用于 Y [当然也可以看成 $-\text{ad}_Y$ 作用于 X]. 引入 $\mathcal{L}(\mathbb{V})$, ad_X 这些记号并非在玩弄文字游戏, 而是为了更清楚地把握其中的代数结构. 注意到, 用此记号, 方程(3.26)可被改写为

$$i\hbar \frac{d}{dt} \Omega(t) = -\text{ad}_H \Omega(t). \quad (3.29)$$

我们待求解的 $\Omega(t)$ 是 $\mathcal{L}(\mathbb{V})$ 中的矢量, 而 $-\text{ad}_H$ 是 $\mathcal{L}(\mathbb{V})$ 上的线性算子.

注记

我们始终牢记, 对于任意的矢量空间 \mathbb{W} , 以及 \mathbb{W} 上的线性算子 A , 关于 \mathbb{W} 中的未知矢量 $|\psi(t)\rangle$ 的微分方程 $\frac{d}{dt}|\psi(t)\rangle = A|\psi(t)\rangle$ 的通解为 $|\psi(t)\rangle = e^{At}|\psi(0)\rangle$.

在上述注记里, 特别地, 取 $\mathbb{W} = \mathcal{L}(\mathbb{V})$, $A = \frac{i}{\hbar} \text{ad}_H$, $|\psi(t)\rangle = \Omega(t)$, 就得到: 方程(3.26)的通解为

$$\Omega(t) = e^{\frac{i}{\hbar} t \text{ad}_H} \Omega. \quad (3.30)$$

上式中的 $e^{\frac{i}{\hbar} t \text{ad}_H} = 1 + \frac{i}{\hbar} t \text{ad}_H + \frac{1}{2!} \left(\frac{i}{\hbar} t \text{ad}_H \right)^2 + \frac{1}{3!} \left(\frac{i}{\hbar} t \text{ad}_H \right)^3 + \dots$, 于是 $\Omega(t)$ 在 $t = 0$ 附近的泰勒展开为

$$\begin{aligned}\Omega(t) &= \Omega + \left(\frac{it}{\hbar} \text{ad}_H \right) \Omega + \frac{1}{2} \left(\frac{it}{\hbar} \text{ad}_H \right)^2 \Omega + \frac{1}{6} \left(\frac{it}{\hbar} \text{ad}_H \right)^3 \Omega + \dots \\ &= \Omega + \frac{it}{\hbar} [H, \Omega] - \frac{t^2}{2\hbar^2} [H, [H, \Omega]] - \frac{it^3}{6\hbar^3} [H, [H, [H, \Omega]]] + \dots\end{aligned}$$

至此, 我们通过两种方法得到了 $\Omega(t)$ 的表达式, 即(3.27), (3.30)两式. 两种方法应该殊途同归, 我们自然猜测 [其实是 “证明了”] 以下等式成立:

$$e^{\frac{i}{\hbar}tH}\Omega e^{-\frac{i}{\hbar}tH} = e^{\frac{i}{\hbar}t\text{ad}_H}\Omega. \quad (3.31)$$

事实上, 数学家早就注意到了:

定理 3.5. (*Baker-Hausdorff 公式*).

设 A, B 是矢量空间 \mathbb{V} 上的任意两个线性算子, 则成立:

$$e^A B e^{-A} = e^{\text{ad}_A} B. \quad (3.32)$$

证明其实很简单, 不必依赖微分方程那些东西:

证明. 定义 $l_A \in \mathcal{L}(\mathcal{L}(\mathbb{V}))$ 如下: $l_A X = AX$, 即 l_A 在线性算子 $X \in \mathcal{L}(\mathbb{V})$ 上的作用效果为左乘 A . 容易验证 l_A 确实是 $\mathcal{L}(\mathbb{V})$ 上的线性算子. 类似定义右乘算子 r_A 如下: $r_A X = XA$. 则显然有如下恒等式:

$$\text{ad}_A = l_A - r_A.$$

此外, 还要注意到, 显然成立

$$[l_A, r_A] = 0,$$

即 $l_A r_A = r_A l_A$ [留给读者验证].

接下来, 再注意

$$\begin{aligned} e^A B &= B + AB + \frac{1}{2}A^2 B + \frac{1}{6}A^3 B + \cdots \\ &= B + l_A B + \frac{1}{2}l_A^2 B + \frac{1}{6}l_A^3 B + \cdots \\ &= \left(1 + l_A + \frac{1}{2}l_A^2 + \frac{1}{6}l_A^3 + \cdots\right) B \\ &= e^{l_A} B. \end{aligned}$$

同理也有 $B e^{-A} = e^{-r_A} B$.

注记

我们还要牢记线性算子指数映射的运算性质: 对任意矢量空间 \mathbb{W} , 以及 \mathbb{W} 上的线性算子 X, Y , 如果 $[X, Y] = 0$, 则

$$e^X e^Y = e^Y e^X = e^{X+Y}.$$

综上所述, 我们立刻得到:

$$e^A B e^{-A} = e^{l_A} e^{-r_A} B = e^{l_A - r_A} B = e^{\text{ad}_A} B.$$

证完了. □

上述证明过程充分发挥了抽象思维的威力. Baker-Hausdorff 公式是常用的基本结论, 我们以后还会用到.

3.5 数学补充: 二维复矢量空间的代数与几何

设 Ω 是量子系统 \mathbb{V} 的一个观测量. 我们回忆, 如果 Ω 有 m 种可能的不同观测结果 $\omega_1, \dots, \omega_m$ (换言之, Ω 有 m 个不同的本征值 $\omega_1, \dots, \omega_m$), 那么希尔伯特空间 \mathbb{V} 的维数 $\dim \mathbb{V} \geq m$. 并且, $\dim \mathbb{V} = m$ 当且仅当 Ω 的所有的本征值都无简并 (重数为 1). 注意到, 在希尔伯特空间 \mathbb{V} 上的所有的厄米特算子之中, “几乎所有的” 厄米特算子都是本征值无简并的. 于是, 大致上来说, 描述量子系统的希尔伯特空间 \mathbb{V} 的维数 = 该量子系统的 “一般的” 观测量的所有可能的观测结果数.

“量子豌豆”⁴¹ 是量子力学种的重要研究对象, 观测量 **色**, **形** 的所有可能观测结果分别为 {黄, 绿} 与 {圆, 皱}, 观测结果数都为 2. 描述 “量子豌豆” 的希尔伯特空间是二维复线性空间 $\mathbb{V} = \mathbb{C}^2$.

本节将介绍二维复矢量空间 $\mathbb{V} = \mathbb{C}^2$ 的代数与几何. 相比于矢量空间的一般理论, 二维复矢量空间 \mathbb{C}^2 这个特殊且相对简单的例子有其独到之处.

\mathbb{C}^2 中的矢量与线性算子

对于二维 (复) 复矢量空间 $\mathbb{V} = \mathbb{C}^2$, 取定它的一组基 $\{|1\rangle, |2\rangle\}$, 则 \mathbb{V} 中任意矢量都形如 $|\psi\rangle = z_1|1\rangle + z_2|2\rangle$. 在此基下, 我们把矢量 $|\psi\rangle$ 表示为通常的列向量,

⁴¹ 其实是**电子**的**自旋**. “抛开事实不谈”, 为了说话方便我们依然假装说成豌豆.

即

$$|\psi\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \quad (3.33)$$

众所周知, 对于复数 $z \in \mathbb{C}$, z 可唯一地表示为 $z = a + bi$, 其中 a, b 为实数. 通过考虑复数的实部与虚部, 我们自然建立了 \mathbb{C} 与 \mathbb{R}^2 的一一对应. 此外, \mathbb{C} 也自然视为实数域 \mathbb{R} 上的矢量空间, $\{1, i\}$ 构成 \mathbb{C} 的一组 \mathbb{R} -基. \mathbb{C} 作为 \mathbb{R} -矢量空间, 其维数为 2, 即

$$\dim_{\mathbb{R}} \mathbb{C} = 2.$$

注记

\dim 右下角标的 \mathbb{R} 用来强调: 被求维数的矢量空间是实数域 \mathbb{R} [而非复数域 \mathbb{C}] 上的矢量空间.

而二维复矢量空间 $\mathbb{V} = \mathbb{C}^2$ 也自然视为实矢量空间: 将 \mathbb{C} 视为 \mathbb{R}^2 , 则 \mathbb{C}^2 自然视为 \mathbb{R}^4 . 在(3.33)式中, 记 $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$, 其中 a_1, b_1, a_2, b_2 均为实数, 则

$$|\psi\rangle = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b_1 \begin{pmatrix} i \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + b_2 \begin{pmatrix} 0 \\ i \end{pmatrix}.$$

容易验证 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ i \end{pmatrix}$ 这 4 个矢量构成 \mathbb{C}^2 的一组 \mathbb{R} -基, $\dim_{\mathbb{R}} \mathbb{C}^2 = 4$.

注记

在本书中, 若不特别说明, 我们总是默认矢量空间都是复矢量空间, 谈论基与维数的时候也都默认是在复矢量空间意义下; 而谈论实矢量空间时都会格外强调.

对于 $\mathbb{V} = \mathbb{C}^2$, 我们回忆, $\mathcal{L}(\mathbb{C}^2)$ 是 \mathbb{C}^2 上的全体 (复) 线性算子构成的集合, 并且具有自然的 (复) 矢量空间结构. 在 $\mathbb{V} = \mathbb{C}^2$ 的给定基下, \mathbb{V} 上的 (复) 线性算子 $\Omega \in \mathcal{L}(\mathbb{V})$ 具有矩阵表示

$$\Omega = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix},$$

上式中的四个矩阵元 z_1, \dots, z_4 均为复数. 可见, 描述 $\mathbb{V} = \mathbb{C}^2$ 上的 (复) 线性算子需要用到 4 个复参数. 事实上, 容易验证

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

构成复矢量空间 $\mathcal{L}(\mathbb{C}^2)$ 的一组基, $\dim \mathcal{L}(\mathbb{C}^2) = 4$.

例题 3.6. 设 $\mathbb{V} = \mathbb{C}^2$.

1. 复矢量空间 $\mathcal{L}(\mathbb{C}^2)$ 也能自然视为实矢量空间. 作为实矢量空间, 其维数 $\dim_{\mathbb{R}} \mathcal{L}(\mathbb{C}^2)$ 是多少?
2. 将 \mathbb{C}^2 视为四维实矢量空间 \mathbb{R}^4 . 作为实矢量空间, 我们可以谈论其上的 **实** 线性算子. 记 $\mathcal{L}_{\mathbb{R}}(\mathbb{C}^2)$ 是 \mathbb{C}^2 上的全体 **实** 线性算子构成的集合. 易知 $\mathcal{L}_{\mathbb{R}}(\mathbb{C}^2)$ 具有实矢量空间的结构, 试求其维数 $\dim_{\mathbb{R}} \mathcal{L}_{\mathbb{R}}(\mathbb{C}^2)$.

证明. 注意 $\mathcal{L}(\mathbb{C}^2)$ 是 4 维矢量空间, 同构于 \mathbb{C}^4 , 从而自然视为 \mathbb{R}^8 . 而 \mathbb{C}^2 作为实矢量空间, 同构于 \mathbb{R}^4 , 从而 $\mathbb{C}^2 = \mathbb{R}^4$ 上的 **实** 线性算子一一对应于 4×4 的实矩阵, 即 $\mathcal{L}_{\mathbb{R}}(\mathbb{C}^2) \cong \mathbb{R}^{4 \times 4} \cong \mathbb{R}^{16}$. 综上,

$$\dim_{\mathbb{R}} \mathcal{L}(\mathbb{C}^2) = 8, \quad \dim_{\mathbb{R}} \mathcal{L}_{\mathbb{R}}(\mathbb{C}^2) = 16.$$

细节留给读者, 读者可尝试具体地写出 $\mathcal{L}(\mathbb{C}^2)$, $\mathcal{L}_{\mathbb{R}}(\mathbb{C}^2)$ 的一组 \mathbb{R} -基. □

\mathbb{C}^2 的厄米特算子

我们知道, 厄米特算子在量子物理中的意义是量子系统的观测量. 每一个厄米特算子都代表一种观测视角, 正如“颜色”与“形状”是观测“量子豌豆”的两种不同视角. 除了颜色和形状, 我们还可以用其他可能的视角来观测“量子豌豆”. 自然要问, 到底有多少种不同的观测量, 即希尔伯特空间 \mathbb{V} 上的厄米特算子“有多少”, “有哪些”.

例题 3.7. 设 \mathbb{V} 为希尔伯特空间, 记 $\text{Herm}(\mathbb{V})$ 为 \mathbb{V} 上的厄米特算子之全体, 则显然 $\text{Herm}(\mathbb{V}) \subseteq \mathcal{L}(\mathbb{V})$. 注意 $\mathcal{L}(\mathbb{V})$ 具有复矢量空间的结构, 问: $\text{Herm}(\mathbb{V})$ 在通常的加法, 数乘下是否也构成复矢量空间?

解. 只需要检查 $\text{Herm}(\mathbb{V})$ 关于加法, 数乘运算是否封闭. 对任意的厄米特算子 $H_1, H_2 \in \text{Herm}(\mathbb{V})$, 显然 $H_1 + H_2$ 也是厄米特算子.

但不幸的是, 对于厄米特算子 $H \in \text{Herm}(\mathbb{V})$ 与复数 $z \in \mathbb{C}$, zH 未必是厄米特算子. 这是因为 $(zH)^\dagger = z^* H^\dagger = z^* H$, 一般来说不等于 zH . 所以 $\text{Herm}(\mathbb{V})$ 关于复数数乘运算不封闭, 从而不构成复矢量空间. \square

注记

虽说 $\text{Herm}(\mathbb{V})$ 关于复数数乘运算不封闭, 但从上述证明中容易看出它关于实数数乘运算是封闭的: 即, 对于实数 $r \in \mathbb{R}$ 以及厄米特算子 $H \in \text{Herm}(\mathbb{V})$, rH 一定是厄米特算子.

于是, $\text{Herm}(\mathbb{V})$ 具有实矢量空间结构.

作为实矢量空间, $\text{Herm}(\mathbb{V})$ 是 $\mathcal{L}(\mathbb{V})$ 的子空间.

我们暂时只关心 $\mathbb{V} = \mathbb{C}^2$ 的情形, 此时 $\mathbb{V} = \mathbb{C}^2$ 上具有内积 $\langle \cdot, \cdot \rangle$. 作为 \mathbb{R} -矢量空间, $\mathcal{L}(\mathbb{C}^2) \cong \mathbb{R}^8$. 而 $\text{Herm}(\mathbb{C}^2)$ 作为 8 维实空间 $\mathcal{L}(\mathbb{C}^2)$ 的 (实) 子空间, 其维数 $\dim_{\mathbb{R}} \text{Herm}(\mathbb{C}^2)$ 是多少? 可否具体写出它的一组 \mathbb{R} -基?

为研究此问题, 取定 $\mathbb{V} = \mathbb{C}^2$ 关于内积 $\langle \cdot, \cdot \rangle$ 的一组么正基 $\{|1\rangle, |2\rangle\}$, 即 $\langle i|j\rangle = \delta_{ij}$, $\forall i, j = 1, 2$. 在此基下, 我们不妨将算子与算子在此基下的矩阵等同. 于是, 对于算子 $\Omega = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \in \mathcal{L}(V)$, Ω 是厄米特算子当且仅当矩阵 Ω 满足

$$\begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}^\dagger,$$

上述条件等价于

$$z_1 = z_1^*, \quad z_2 = z_3^*, \quad z_3 = z_2^*, \quad z_4 = z_4^*,$$

也就是说, z_1, z_4 都是实数, z_2, z_3 彼此共轭. 因此, 我们证明了:

性质 3.8. 设 $\mathbb{V} = \mathbb{C}^2$ 为希尔伯特空间, $H \in \mathcal{L}(\mathbb{V})$. 则 H 是厄米特算子当且仅当它在 \mathbb{V} 的某组给定的么正基下的矩阵形如

$$H = \begin{pmatrix} a & b - ci \\ b + ci & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}. \quad (3.34)$$

可见, 描述 $\mathbb{V} = \mathbb{C}^2$ 的厄米特算子, 即 $\text{Herm}(\mathbb{C}^2)$ 中的元素, 需要 4 个 (实) 参数 a, b, c, d . 从而易知实矢量空间 $\text{Herm}(\mathbb{C}^2)$ 的维数 $\dim_{\mathbb{R}} \text{Herm}(\mathbb{C}^2) = 4$. 我们也能轻易地给出 $\text{Herm}(\mathbb{C}^2)$ 的一组基:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.35)$$

厄米特算子 $H = \begin{pmatrix} a & b - ci \\ b + ci & d \end{pmatrix} \in \text{Herm}(\mathbb{C}^2)$ 在此基下的坐标恰为 $(a, b, c, d)^\top$.

至此, 我们已经完全搞清楚 $\mathbb{V} = \mathbb{C}^2$ 的厄米特算子 “长什么样” [长得像(3.34)那样], 以及 “有多少” [构成 4 维空间]. 但事情还没做完, 因为刚才取的那组基(3.35)还不够漂亮, 出于某些几何, 物理背景 (将会慢慢介绍), 我们将(3.35)稍微线性组合一下, 另取一组 “更漂亮” 的基:

定义. 记 $\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 为 $\mathbb{V} = \mathbb{C}^2$ 的恒等算子, 再记

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.36)$$

矩阵 $\sigma_x, \sigma_y, \sigma_z$ 称为**泡利矩阵** (Pauli matrix).

容易验证, $\{\mathbf{1}, \sigma_x, \sigma_y, \sigma_z\}$ 构成 $\text{Herm}(\mathbb{C}^2)$ 的一组 \mathbb{R} -基.

泡利矩阵的运算性质

泡利矩阵 $\sigma_x, \sigma_y, \sigma_z$ 是具体研究 \mathbb{C}^2 的重要工具, 并且有丰富的量子物理含义 (以后会细说), 我们需要对其运算性质非常熟悉. 首先交代一些常用记号.

我们回忆, \mathbb{C}^2 上的厄米特算子 H [在某组么正基下的矩阵] 必形如

$$H = \gamma \mathbf{1} + k_x \sigma_x + k_y \sigma_y + k_z \sigma_z, \quad \gamma, k_x, k_y, k_z \in \mathbb{R},$$

按通常 3 维欧氏空间 \mathbb{R}^3 的解析几何中的记号, 记向量 $\mathbf{k} = (k_x, k_y, k_z)^\top \in \mathbb{R}^3$, 再形式地记 $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^\top$, 从而将 $k_x \sigma_x + k_y \sigma_y + k_z \sigma_z$ 简记为 $\mathbf{k} \cdot \boldsymbol{\sigma}$. 出于习惯, 记 $\mathbf{n} := \frac{\mathbf{k}}{|\mathbf{k}|}$ 为与 \mathbf{k} 同方向的单位向量, $k := |\mathbf{k}|$ 为向量 \mathbf{k} 的长度. 在此记号下,

$$H = \gamma \mathbf{1} + k \mathbf{n} \cdot \boldsymbol{\sigma}. \quad (3.37)$$

注记

“矢量”与“向量”在数学上是同一种东西；而考虑到物理含义，本书习惯把用于描述量子态的抽象的希尔伯特空间中的元素叫做“矢量”，把平面 \mathbb{R}^2 ，空间 \mathbb{R}^3 中的元素称为“向量”。

我们来考察泡利矩阵的运算性质，首先是矩阵乘法运算。

性质 3.9. 设 $\sigma_x, \sigma_y, \sigma_z$ 是(3.36)所定义的泡利矩阵，则

1. 泡利矩阵的乘法表如下：

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbf{1},$$

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$$

$$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x$$

$$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y$$

2. 设 $\mathbf{n} = (n_x, n_y, n_z) \in \mathbb{R}^3$ 为单位向量，则

$$(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = \mathbf{1}. \quad (3.38)$$

从而，矩阵 $\mathbf{n} \cdot \boldsymbol{\sigma}$ 的本征值为 ± 1 。

证明. 乘法表可直接计算验证. 只证 (2). 注意 $\sigma_x \sigma_y + \sigma_y \sigma_x = 0$ (以及另外两个类似的)，于是

$$\begin{aligned} (\mathbf{n} \cdot \boldsymbol{\sigma})^2 &= (n_x \sigma_x + n_y \sigma_y + n_z \sigma_z)^2 \\ &= n_x^2 \sigma_x^2 + n_y^2 \sigma_y^2 + n_z^2 \sigma_z^2 + n_x n_y (\sigma_x \sigma_y + \sigma_y \sigma_x) \\ &\quad + n_y n_z (\sigma_y \sigma_z + \sigma_z \sigma_y) + n_z n_x (\sigma_z \sigma_x + \sigma_x \sigma_z) \\ &= (n_x^2 + n_y^2 + n_z^2) \mathbf{1} = \mathbf{1}. \end{aligned}$$

从而 $(\mathbf{n} \cdot \boldsymbol{\sigma} + \mathbf{1})(\mathbf{n} \cdot \boldsymbol{\sigma} - \mathbf{1}) = 0$. 于是易知 $\mathbf{n} \cdot \boldsymbol{\sigma}$ 的本征值只可能是 ± 1 . 另一方面，考虑算子的迹(1.33)，即矩阵各对角元之和，易知 $\text{tr}(\mathbf{n} \cdot \boldsymbol{\sigma}) = 0$ ，从而 $\mathbf{n} \cdot \boldsymbol{\sigma} \neq \pm \mathbf{1}$. 所以 ± 1 都是 $\mathbf{n} \cdot \boldsymbol{\sigma}$ 的本征值. \square

注记

若记 $\mathcal{I} = -i\sigma_x$, $\mathcal{J} = -i\sigma_y$, $\mathcal{K} = -i\sigma_z$, 则 $\mathcal{I}, \mathcal{J}, \mathcal{K}$ 满足

$$\mathcal{I}^2 = \mathcal{J}^2 = \mathcal{K}^2 = -\mathbf{1}$$

$$\mathcal{IJ} = -\mathcal{JI} = \mathcal{K}, \quad \mathcal{JK} = -\mathcal{KJ} = \mathcal{I}, \quad \mathcal{KI} = -\mathcal{IK} = \mathcal{J}.$$

上述乘法表像什么呢? 众所周知, 复数域 \mathbb{C} 还能进一步扩充为四元数环 $\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$, 其中 i, j, k 为虚数单位, 满足 $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$.

\mathbb{C}^2 上的酉算子

在量子力学中, 求解量子系统随时间演化 (即薛定谔方程) 会解得 $|\psi(t)\rangle = e^{-\frac{i}{\hbar}tH}|\psi(0)\rangle =: U(t)|\psi(0)\rangle$, 其中 H 为该量子系统的哈密顿量, $U(t) := e^{-\frac{i}{\hbar}tH}$ 为传播子, 它是酉算子.

众所周知, 在任何希尔伯特空间 \mathbb{V} 上, 若 H 是厄米特算子, 则 e^{iH} 必然是酉算子. 传播子 $U(t)$ 就是形如这样的酉算子. 于是自然有个问题: 所有的酉算子都一定形如 e^{iH} 吗? 事实上, 对有限维希尔伯特空间, 上述问题的答案是肯定的, 其证明需要李群李代数的一般理论, 超出本书范围. 而 $\mathbb{V} = \mathbb{C}^2$ 的特殊情形则相对容易一些.

首先, 对于 $\mathbb{V} = \mathbb{C}^2$ 上的厄米特算子

$$H = \gamma \mathbf{1} + k \mathbf{n} \cdot \boldsymbol{\sigma},$$

我们先来计算 e^{iH} .

性质 3.10. 设 $\mathbf{n} \in \mathbb{R}^3$ 为单位向量, $k, \gamma \in \mathbb{R}$, 则成立

$$e^{i\gamma \mathbf{1} + ik \mathbf{n} \cdot \boldsymbol{\sigma}} = e^{i\gamma} (\mathbf{1} \cos k + i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin k). \quad (3.39)$$

证明. 注意单位矩阵 $\mathbf{1}$ 与 $\mathbf{n} \cdot \boldsymbol{\sigma}$ 可交换, 所以

$$e^{i\gamma \mathbf{1} + ik \mathbf{n} \cdot \boldsymbol{\sigma}} = e^{i\gamma \mathbf{1}} e^{ik \mathbf{n} \cdot \boldsymbol{\sigma}} = e^{i\gamma} e^{ik \mathbf{n} \cdot \boldsymbol{\sigma}}.$$

反复利用 $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = \mathbf{1}$, 可知

$$(\mathbf{n} \cdot \boldsymbol{\sigma})^n = \begin{cases} \mathbf{1} & \text{若 } n \text{ 为偶数} \\ \mathbf{n} \cdot \boldsymbol{\sigma} & \text{若 } n \text{ 为奇数.} \end{cases}$$

从而由指数映射的定义直接计算验证得

$$\begin{aligned} e^{ik\mathbf{n} \cdot \boldsymbol{\sigma}} &= \sum_{n \geq 0} \frac{(ik)^n}{n!} (\mathbf{n} \cdot \boldsymbol{\sigma})^n \\ &= \sum_{n \geq 0} \left(\frac{(ik)^{2n}}{(2n)!} \mathbf{1} + \frac{(ik)^{2n+1}}{(2n+1)!} (\mathbf{n} \cdot \boldsymbol{\sigma}) \right) \\ &= \mathbf{1} \cos k + i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin k. \end{aligned}$$

从而得证. □

注记

此性质可以看成欧拉公式 $e^{i\theta} = \cos \theta + i \sin \theta$ 的类比.

第二步, 我们来求出 \mathbb{C}^2 上的所有的酉算子. 设 U 为酉算子, 则 U 在么正基下的矩阵是酉矩阵, 即满足 $U^{-1} = U^\dagger$.

性质 3.11. 设 $U \in \mathcal{L}(\mathbb{C}^2)$, 则 U 是酉算子当且仅当 U 在么正基下的矩阵必形如

$$U = e^{i\gamma} \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}, \quad (3.40)$$

其中 $\gamma \in \mathbb{R}$, $z, w \in \mathbb{C}$ 满足 $|z|^2 + |w|^2 = 1$.

证明. 一方面, 直接验证可知形如(3.40)的矩阵必为酉矩阵. 只需再证明酉矩阵必形如(3.40).

若 U 为酉矩阵, 则 $|\det U| = 1$, 从而必存在 $\gamma \in \mathbb{R}$ 使得 $\det U = e^{2i\gamma}$. 注意 $V := e^{-i\gamma} U$ 依然是酉矩阵, 并且行列式 $\det(V) = (e^{-i\gamma})^2 \det U = 1$. 于是, V 是行列式为 1 的酉矩阵 [称为**特殊酉矩阵**], $U = e^{i\gamma} V$. 接下来, 只需搞清楚 2×2 的特殊酉矩阵长什么样.

记 $V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 其中 $a, b, c, d \in \mathbb{C}$. 注意 V 满足条件 $VV^\dagger = \mathbf{1}$, $\det V = 1$. 直接计算矩阵乘法, 比较矩阵元, 易知上述条件等价于

$$|a|^2 + |b|^2 = 1 \quad (3.41)$$

$$|c|^2 + |d|^2 = 1 \quad (3.42)$$

$$ac^* + bd^* = 0 \quad (3.43)$$

$$ca^* + db^* = 0 \quad (3.44)$$

$$ad - bc = 1 \quad (3.45)$$

将(3.43)两边同乘 d , 再注意(3.42), (3.45), 可得

$$0 = ac^*d + bdd^* = c^*(1 + bc) + bdd^* = c^* + b(cc^* + dd^*) = c^* + b,$$

从而 $c = -b^*$. 类似地, 用初等代数也能整理出 $d = a^*$ [留给读者]. 因此,

$$V = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad |a|^2 + |b|^2 = 1.$$

从而得证. □

注记

按照通用记号, 2×2 酉矩阵构成的集合记为 $U(2)$,
 2×2 特殊酉矩阵构成的集合记为 $SU(2)$.

于是, 我们证明了

$$U(2) := \left\{ e^{i\gamma} \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix} \mid |z|^2 + |w|^2 = 1, z, w \in \mathbb{C}, \gamma \in \mathbb{R} \right\},$$

$$SU(2) := \left\{ \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix} \mid |z|^2 + |w|^2 = 1, z, w \in \mathbb{C} \right\},$$

于是立即推出:

定理 3.12. 设 $U \in U(2)$, 则必存在厄米特矩阵 H , 使得 $U = e^{iH}$.

证明. 记 $U = e^{i\gamma} \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}$, 其中 $|z|^2 + |w|^2 = 1$. 分离 z, w 的实部与虚部, 记 $z = a + bi, w = c + di$, 则实数 a, b, c, d 满足 $a^2 + b^2 + c^2 + d^2 = 1$. 从而

$$\begin{aligned} U &= e^{i\gamma} \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} \\ &= e^{i\gamma} (a\mathbf{1} + i(d\sigma_x + c\sigma_y + b\sigma_z)). \end{aligned}$$

取 \mathbb{R}^3 的单位向量 $\mathbf{n} := \frac{1}{\sqrt{b^2 + c^2 + d^2}}(d, c, b)^\top$, 再取实数 k 使得 $\cos k = a, \sin k = \sqrt{b^2 + c^2 + d^2}$, 记

$$H = \gamma\mathbf{1} + k\mathbf{n} \cdot \boldsymbol{\sigma},$$

则 H 是厄米特算子, 且由公式(3.39)可知 $e^{iH} = U$, 得证. \square

本小节最后, 再额外提一下特殊酉矩阵构成的集合 $\text{SU}(2)$ 的几何意义. 注意 $\text{SU}(2)$ 中的元素形如 $\begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}$, 于是有如下显然的一一对应:

$$\begin{aligned} \text{SU}(2) &\cong \left\{ (z, w) \in \mathbb{C}^2 \mid |z|^2 + |w|^2 = 1 \right\} \\ \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix} &\mapsto (z, w). \end{aligned}$$

而 $\left\{ (z, w) \in \mathbb{C}^2 \mid |z|^2 + |w|^2 = 1 \right\}$ 作为 \mathbb{C}^2 的子集, 在 $\mathbb{C} \cong \mathbb{R}^2$ 的等同下, 自然看成 \mathbb{R}^4 的子集 $S^3 := \left\{ (a, b, c, d) \in \mathbb{R}^4 \mid a^2 + b^2 + c^2 + d^2 = 1 \right\}$, 而这正是 4 维欧氏空间的单位球面, 所谓的**三维球面**.

\mathbb{C}^2 的量子态: 复射影空间 \mathbb{CP}^1 与 Bloch 球面

我们回忆, 对于一般的量子系统 \mathbb{V} 中的两个非零矢量 $|\psi_1\rangle, |\psi_2\rangle$, 这两个矢量表示同一个量子态当且仅当存在常数 $c \in \mathbb{C}$ 使得 $|\psi_1\rangle = c|\psi_2\rangle$, 即, 这两个矢量“相差常数倍”. 于是, 出于量子观测的物理意义, 我们把相差常数倍的 (非零) 矢量认为是相同的矢量. 那么, 在此意义下, \mathbb{V} 中有多少“不同”的矢量?

对于非零矢量 $|\psi\rangle \in \mathbb{V}$, 我们称集合

$$l_{|\psi\rangle} := \mathbb{C}|\psi\rangle := \{c|\psi\rangle \mid c \in \mathbb{C}\}$$

为矢量 $|\psi\rangle$ 所在的**复直线**. 注意 $l_{|\psi\rangle}$ 其实是 \mathbb{V} 的由 $|\psi\rangle$ 张成的一维复子空间. 注意 $l_{|\psi\rangle}$ 中的所有的非零矢量显然都与 $|\psi\rangle$ “相同”, 另一方面所有与 $|\psi\rangle$ “相同” 的非零矢量都在 $l_{|\psi\rangle}$ 中. 从而, 易知有如下的一一对应:

$$\left\{ \begin{array}{l} \mathbb{V} \text{ 中的非零矢量} \\ \text{相差常数倍则认为相同} \end{array} \right\} \cong \{l \mid l \text{ 是 } \mathbb{V} \text{ 的一维 (复) 子空间}\}. \quad (3.46)$$

定义. 对于 (复) 矢量空间 \mathbb{V} ,

1. 记 $\mathbb{P}(\mathbb{V})$ 为 \mathbb{V} 的所有一维 (复) 子空间构成的集合, 称为 \mathbb{V} 的**射影化**.
2. 特别地, 对于 $\mathbb{V} = \mathbb{C}^{n+1}$ 是有限维复矢量空间的情形, 记 $\mathbb{CP}^n := \mathbb{P}(\mathbb{C}^{n+1})$. 集合 \mathbb{CP}^n 称为 n **维复射影空间**.

于是, “量子豌豆系统” $\mathbb{V} = \mathbb{C}^2$ 的所有不同量子态之全体为 \mathbb{CP}^1 . 然而 \mathbb{CP}^1 又是什么东西? 我们设法描述它.

对于 $l \in \mathbb{CP}^1$, 即 l 是 $\mathbb{V} = \mathbb{C}^2$ 的一维 (复) 子空间, 在给定的厄米特内积下, 考虑 l 的正交补 l^\perp , 则有正交直和 $\mathbb{V} = l \oplus l^\perp$. 在此直和分解下, 任何矢量 $|\psi\rangle$ 都可唯一地分解为 $|\psi\rangle = |\psi_l\rangle + |\psi_{l^\perp}\rangle$, 使得 $|\psi_l\rangle \in l$, $|\psi_{l^\perp}\rangle \in l^\perp$. 容易验证, 映射 $\mathbb{P}_l: |\psi\rangle \mapsto |\psi_l\rangle$ 定义了 \mathbb{V} 上的一个线性算子. 这个线性算子其实是关于子空间 l 的**投影算子**. 具体地, 任取单位矢量 $|\varphi\rangle \in l$, 则

$$\mathbb{P}_l = |\varphi\rangle\langle\varphi| \quad (3.47)$$

[若 $|\varphi\rangle$ 是 l 中的单位矢量, 则 $e^{i\theta}|\varphi\rangle$, $\theta \in \mathbb{R}$, 是 l 的全部单位矢量. 也容易直接验证 \mathbb{P}_l 的上述表达式与单位矢量 $|\varphi\rangle$ 的选取无关.] 注意投影算子一定是厄米特算子, $\mathbb{P}_l \in \text{Herm}(\mathbb{C}^2) \cong \mathbb{R}^4$.

于是, 我们把每个 $l \in \mathbb{CP}^1$ 都对应到 $\mathbb{P}_l \in \text{Herm}(\mathbb{C}^2)$. 显然, \mathbb{CP}^1 中的不同元素 l, l' 对应于不同的算子 $\mathbb{P}_l, \mathbb{P}_{l'}$ [这是显然的, 毕竟 \mathbb{P}_l 的几何意义是关于子空间的正交投影, 不同的子空间显然对应不同的投影.] 也就是说, 映射 $\mathbb{CP}^1 \rightarrow \text{Herm}(\mathbb{C}^2)$, $l \mapsto \mathbb{P}_l$ 是单射. 从而将 \mathbb{CP}^1 视为 $\text{Herm}(\mathbb{C}^2) \cong \mathbb{R}^4$ 的子集.

给定 $l \in \mathbb{CP}^1$, 取 l 中的单位矢量 $|\varphi\rangle = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, $|z_1|^2 + |z_2|^2 = 1$, 从而必存在 $\theta, \phi_1, \phi_2 \in \mathbb{R}$, 使得

$$z_1 = \cos \theta \cdot e^{i\phi_1}, \quad z_2 = \sin \theta \cdot e^{i\phi_2},$$

于是

$$|\varphi\rangle = \begin{pmatrix} \cos \theta \cdot e^{i\phi_1} \\ \cos \theta \cdot e^{i\phi_2} \end{pmatrix} = e^{i\frac{\phi_1+\phi_2}{2}} \begin{pmatrix} \cos \theta \cdot e^{i\frac{\phi_1-\phi_2}{2}} \\ \cos \theta \cdot e^{i\frac{\phi_2-\phi_1}{2}} \end{pmatrix}.$$

适当调整记号 [没别的意思, 就是为了好看], 将上式中的 $\frac{\phi_1+\phi_2}{2}$ 重新记为 γ , 将 θ 重新记为 $\frac{\theta}{2}$, 将 $\phi_1 - \phi_2$ 重新记为 $-\phi$, 则

$$|\varphi\rangle = e^{i\gamma} \begin{pmatrix} \cos \frac{\theta}{2} \cdot e^{-i\frac{\phi}{2}} \\ \sin \frac{\theta}{2} \cdot e^{i\frac{\phi}{2}} \end{pmatrix}, \quad (3.48)$$

于是关于子空间 l 的投影算子 \mathbb{P}_l 在么正基 $\{|1\rangle, |2\rangle\}$ 下的矩阵表达式为

$$\begin{aligned} \mathbb{P}_l = |\varphi\rangle\langle\varphi| &= \begin{pmatrix} \cos^2 \frac{\theta}{2} & \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cdot e^{-i\phi} \\ \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cdot e^{i\phi} & \sin^2 \frac{\theta}{2} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 + \cos \theta & \sin \theta \cdot e^{-i\phi} \\ \sin \theta \cdot e^{i\phi} & 1 - \cos \theta \end{pmatrix}. \\ &= \frac{1}{2} (1 + \sigma_x \sin \theta \cos \phi + \sigma_y \sin \theta \sin \phi + \sigma_z \cos \theta) \\ &= \frac{1 + \mathbf{n} \cdot \boldsymbol{\sigma}}{2}, \end{aligned} \quad (3.49)$$

其中 $\mathbf{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^\top \in \mathbb{R}^3$ 是 \mathbb{R}^3 中的单位向量. [注意 θ, ϕ 刚好是 \mathbf{n} 的球坐标的角向分量].

另一方面, 容易验证, 对于 \mathbb{R}^3 中的任意单位向量 \mathbf{n} , $\frac{1 + \mathbf{n} \cdot \boldsymbol{\sigma}}{2}$ 也一定是 \mathbb{C}^2 的关于某个一维子空间 l 的投影算子, 从而对应于 $l \in \mathbb{CP}^1$. 因此, 我们建立了如下的一一对应:

$$\begin{aligned} \mathbb{CP}^1 &\cong \left\{ \frac{1 + \mathbf{n} \cdot \boldsymbol{\sigma}}{2} \in \text{Herm}(\mathbb{C}^2) \cong \mathbb{R}^4 \mid \mathbf{n} \in \mathbb{R}^3, |\mathbf{n}| = 1 \right\} \\ &\cong \left\{ \mathbf{n} \in \mathbb{R}^3 \mid |\mathbf{n}| = 1 \right\} =: S^2. \end{aligned}$$

注意 $\left\{ \mathbf{n} \in \mathbb{R}^3 \mid |\mathbf{n}| = 1 \right\}$ 其实就是通常的三维欧氏空间 \mathbb{R}^3 的单位球面, 即二维球面, 记作 S^2 . 可见, 一维复射影空间 \mathbb{CP}^1 同构于二维球面 S^2 .

注记

综上所述, 2 维希尔伯特空间 \mathbb{C}^2 中的量子态 (相差常数倍的矢量认为是同一个) 与球面 S^2 上的点一一对应, 其对应关系为:

$$\begin{aligned} S^2 &\rightarrow \mathbb{CP}^1 \\ \begin{pmatrix} \sin \theta \sin \phi \\ \sin \theta \cos \phi \\ \cos \theta \end{pmatrix} &\mapsto \mathbb{C} \begin{pmatrix} \cos \frac{\theta}{2} e^{-i\frac{\phi}{2}} \\ \sin \frac{\theta}{2} e^{i\frac{\phi}{2}} \end{pmatrix} \end{aligned}$$

换言之, 我们可以用球面 S^2 上的点来描述 \mathbb{C}^2 的量子态. 此时, 球面 S^2 称为 **Bloch 球面**, 球面上的点 \mathbf{n} 称为 **Bloch 向量**.

本节小结

为方便查阅与对比, 将本节的部分内容整理如下:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

厄米特算子	$\text{Herm}(\mathbb{C}^2) \cong \mathbb{R}^4$	$\gamma \mathbf{1} + k \mathbf{n} \cdot \boldsymbol{\sigma}, \gamma \in \mathbb{R}, k \mathbf{n} \in \mathbb{R}^3.$
特殊酉算子	$\text{SU}(2) \cong S^3$	$\begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix}, (z, w) \in S^3$ 或者 $\mathbf{1} \cos k + i(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin k, k \in \mathbb{R}, \mathbf{n} \in S^2$
量子态	$\mathbb{CP}^1 \cong S^2$	$\mathbb{P}_l = \frac{\mathbf{1} + \mathbf{n} \cdot \boldsymbol{\sigma}}{2}, l \in \mathbb{CP}^1.$

而 $\text{U}(2)$ 中的元素则必形如 $e^{i\gamma}V$, 其中 $V \in \text{SU}(2)$.

- $U \in \text{U}(2)$ 当且仅当存在 $H \in \text{Herm}(\mathbb{C}^2)$, 使得 $U = e^{iH}$.
 - $U \in \text{SU}(2)$ 当且仅当存在 $H \in \text{Herm}(\mathbb{C}^2)$, 使得 $U = e^{iH}$, 且 $\text{tr}H = 0$.
- [留给读者].

参考文献

- [1] Albert, D. Z. (1992). *Quantum mechanics and experience*. Harvard University Press.
- [2] Hall, B. C. (2013). *Quantum theory for mathematicians* (Vol. 267). New York: Springer.
- [3] Miwa, T., Jimbo, M., Jimbo, M., & Date, E. (2000). *Solitons: Differential equations, symmetries and infinite dimensional algebras* (Vol. 135). Cambridge University Press.
- [4] Shankar, R. (2012). *Principles of quantum mechanics*. Springer Science & Business Media.
- [5] Sepanski, M. R. (Ed.). (2007). *Compact lie groups*. New York, NY: Springer New York.
- [6] 石名俊. (2018). 中国科学技术大学《量子力学 (A)》课程讲义. 网页链接: <http://staff.ustc.edu.cn/~shmj/>.
- [7] 知乎专栏. 现代量子力学阅读笔记. 网页链接: <https://zhuanlan.zhihu.com/p/33391591>

名词英汉对照

- bra 左矢, 11
- canonical coordinate 正则坐标, 65
- canonical transformation 正则变换, 65
- commutator 对易子, 15
- configuration space 位形空间, 54
- cyclic coordinates 循环坐标, 61
- Dirac delta function 狄拉克 delta 函数, 40
- Hamilton's canonical equation 哈密顿正则方程, 60
- Hamiltonian 哈密顿量, 60
- Hilbert space 希尔伯特空间, 10
- ket 右矢, 11
- Lagrangian 拉格朗日量, 55
- Legendre transformation 勒让德变换, 60
- Pauli matrix 泡利矩阵, 101
- phase space 相空间, 60
- Poisson bracket 泊松括号, 62
- propagator 传播子, 32
- propogator 传播子, 89
- Stern-Gerlach experiment 斯特恩-格拉赫实验, 71
- subspace 子空间, 13
- trace 迹, 25