

Quantum Mechanics

大白菜*

起草于2018-12-13

*✉ tlwyxzcx@sina.cn, QQ:243114781, QQ群:234674270. 作者是群主,本书需要继续完善欢迎提出本书的错误欢迎交流,本书尚未完成,请群友不要外传,以免丢人现眼.

目录

I	非相对论量子力学	1
1	线性代数	1
1.1	简介	1
1.2	向量空间和子空间	1
1.3	赋范空间	6
1.4	Banach空间	13
1.5	线性映射	15
1.6	线性空间的同构	20
1.7	线性变换	21
1.8	Hilbert空间	23
1.9	正交归一系统	33
2	基本物理图像	44
2.1	经典力学	44
2.1.1	基本力学概念	44
2.1.2	运动积分与诺特定理	47
2.1.3	哈密顿力学	52
2.1.4	泊松括号	55
2.1.5	可积系统	61
2.1.6	哈密顿雅克比方程	62
2.2	量子力学中的概率	65
2.2.1	双缝干涉	65
2.2.2	概率幅	67
2.2.3	测量问题	68
2.3	量子力学基础概念	70
2.3.1	量子态	70
2.3.2	左矢量	72
3	算符理论	76
3.1	线性算符	76
3.1.1	线性算符	76

3.1.2	有界算符和逆算符	78
3.1.3	练习题	80
3.2	自伴算符(厄米算符)和投影算符	80
3.2.1	自伴算符	80
3.2.2	投影算符	82
3.2.3	练习题	84
3.3	谱理论	85
3.3.1	算符的谱	85
3.3.2	算符函数与谱映射定理	85
3.4	可观测力学量	85
4	量子化	87
4.1	基本物理学量	87
4.2	表象变换与本征态	87
4.3	李乘法	87
4.4	基本量子化条件	87
4.5	不确定关系	87
5	量子动力学	87
5.1	薛定谔方程	87
5.1.1	能量本征方程	87
5.1.2	概率流守恒	87
5.1.3	一维well问题举例	87
5.2	单参数么模群——时间演化算符	87
5.3	传播子与格林函数	87
5.4	动力学图像——薛定谔图像、海森堡图像、狄拉克图像	87
6	量子力学对称性	87
6.1	平移对称性与动量	87
6.2	转动对称性与角动量	87
6.3	李群李代数及其表示	87
6.4	角动量及其耦合	87
6.5	不可约张量算符	87

7 三维束缚态	87
7.1 渐进分析	87
7.2 无限深球well	87
7.3 有限深球well	87
7.4 三维谐振子	87
7.5 氢原子严格解	87
8 三维散射态	87
8.1 经典散射与微分截面	87
8.2 含时散射	87
8.3 $S - matrix$	87
8.4 Dyson公式	87
8.5 Born近似	87
9 多粒子量子系统	87
9.1 全同粒子与交换对称性	87
9.2 置换群与杨图、杨算子	87
9.3 二次量子化	87
9.4 电磁场正则量子化	87
9.5 Casmir效应	87
II 相对论量子力学	87
10 连续场基本方程	87
10.1 欧拉-拉格朗日方程	87
10.2 诺特原理与诺特流	87
10.3 克莱因-高登方程	87
10.4 克莱因-高登方程正则量子化	87
11 洛伦兹群	87
11.1 洛伦兹变换	87
11.2 Clifford代数的表示—— γ 矩阵	87
11.3 狄拉克方程	87
11.4 狄拉克方程正则量子化	87

11.5 洛伦兹群的分离对称性与电荷对称性—— CPT 联合不变性	87
12 微扰理论	87
12.1 Gellman-Low公式与真空	87
12.2 Wick定理	87
12.3 Feynman图与费曼规则	87
12.3.1 ϕ^4 的费曼规则	87
12.3.2 QED的费曼规则	87
12.4 截面与S矩阵元	87
13 树图计算	87
13.1 $e^+e^- \rightarrow \mu^+\mu^-$	87
13.2 康普顿散射	87
13.3 库仑散射	87
14 圈图计算	87
14.1 顶角近似	87
14.2 辐射修正,正规化方案——PV方案	87
14.3 场强重整化	87
14.4 LSZ约化定理	87
14.5 光学定理	87
14.6 Ward-Takahashi恒等式	87
14.7 电荷重整化	87
III 规范理论	87
15 $U(1)$规范理论	87
16 $SU(n)$规范理论	87
16.1 规范理论的内禀空间与底流形	87
16.2 规范自由与纤维丛	87
16.3 规范固定与截面	87
16.4 Yang-Mills方程	87
16.5 Wilson圈	87

17 路径积分量子化	87
17.1 量子力学的路径积分	87
17.2 标量场的量子化	87
17.3 QED的量子化	87
17.4 旋量场的量子化	87
18 <i>Non - Abelian</i>规范场量子化	87
18.1 Faddeev-Popov路径积分量子化	87
18.2 鬼	87
18.3 BRST对称性	87
18.4 <i>Non - Abelian</i> 场的单圈修正	87
18.5 渐进自由	87
19 量子色动力学	87
19.1 夸克理论	87
19.2 胶子辐射	87
19.3 强子散射:轻子对	87
20 微扰理论	87
20.1 二维手征流	87
20.2 四维手征流	87
20.3 Goldstone定理	87
20.4 手征流反常	87
21 自发对称性破缺	87
21.1 Higgs机制	87
21.2 弱相互作用的Glashow-Weinberg-Salam理论	87

Part I

非相对论量子力学

1 线性代数

1.1 简介

量子力学最主要的数学工具是线性代数. 要学好量子力学, 必须非常熟悉线性代数中的基本概念和运算方法. 虽然我们假定读者已经学习过线性代数, 但考虑到不同读者采用的符号标记不同, 练习不够, 甚至接受了某些错误的概念, 有时会对量子力学学习造成一些不必要的麻烦. 因此本书之初, 我们结合物理学, 复习线性代数中的基本概念和运算方法, 统一符号, 强调某些容易混淆的概念, 我们愿意提醒读者, 理解这些概念和运算方法, 并不等于能熟练使用它们.

1.2 向量空间和子空间

向量空间和向量空间的基

如果数集 F 中任意两个数作某一运算后的结果仍然在 F 中, 我们就称数集 F 对这个运算是封闭的, 对加减乘除四则运算封闭的数集 F 称为数域.

最常见的数域是有理数域 \mathbf{Q} 、实数域 \mathbf{R} 、复数域 \mathbf{C} . 此外还有很多其他数域, 比如 $\mathbf{Q}(\sqrt{2})$:

$$\mathbf{Q}(\sqrt{2}) = \{a + b\sqrt{2} | a, b \in \mathbf{Q}\}$$

定义 1.1 (向量空间) 设 E 是一个非空集合, F 是一个数域, 在集合 E 的元素之间定义了加法. 在数域 F 与集合 E 的元素之间还定义了数量乘法, 如果加法与数量乘法满足下述规则:

(1) 加法交换律

$$a + b = b + a$$

(2) 加法结合律

$$(a + b) + c = a + (b + c)$$

(3) 在 E 中有单位元 0

$$a + 0 = a$$

(4) 对于 E 中每一个元素 a 都存在逆元素 b

$$a + b = \mathbf{0}$$

(5) $1a = a$

(6) $k(la) = (kl)a$

(7) $(k + l)a = ka + la$

(8) $k(a + b) = ka + kb$

则称 E 为数域 F 上的向量空间, 其中 $k, l \in F$, 而 $a, b \in E$, 其中的元素也常称为向量, 数域 F 中的元素也常称为标量.

例如矩阵元是 F 中标量构成的矩阵, 对于矩阵加法和矩阵的数乘, 构成数域 F 上的向量空间. 全体实函数按照函数加法与乘法也构成实数域上的向量空间

定义 1.2 (向量空间的基) 若向量空间 E 中, 有 n 个向量 $a_1, a_2, a_3, \dots, a_n$ 线性无关, 而 E 中任意 $n + 1$ 个向量线性相关, 则把这 a_1, a_2, \dots, a_n 称为向量空间 E 的一组基底, 由于向量空间中的所有基底都含有相同数量的向量, 称 n 为向量空间 E 的维数. 记作: $\dim E = n$

向量空间的子空间

定义 1.3 (子空间) 如果数域 F 上的向量空间 E 的非空子集 E_1, E_2 对于 E 的两种运算也构成向量空间, 则称 E_1, E_2 为 E 的一个子空间

例如 $W = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in F \right\}$ 是 $F^{2 \times 2}$ 的一个子空间.

显然向量空间中, 由单个的零向量所构成的子集合是一个向量空间, 它叫做**零子空间**.

向量空间 E 自己本身也是 E 的一个子空间.

在向量空间总, 零子空间和向量空间本身这两个子空间有时候叫做**平凡子空间**, 而其他线性子空间叫做**非平凡子空间**.

定义 1.4 (子空间的和) 设 E_1, E_2 是向量空间 E 的两个子空间. 把所有能表示成 $a_1 + a_2$ ($a_1 \in E_1, a_2 \in E_2$) 的向量组成的子集合叫做 E_1 和 E_2 的和, 记作: $E_1 + E_2$

很容易看出, 子空间的和 $E_1 + E_2$ 与子空间的交 $E_1 \cap E_2$ 也是都是子空间

定义 1.5 设 E_1, E_2 是向量空间 E 的两个子空间.

$$\dim(E_1) + \dim(E_2) = \dim(E_1 + E_2) + \dim(E_1 \cap E_2)$$

子空间和的维数一般比子空间维数之和要小.

定义 1.6 (子空间的直和) 设 E_1, E_2 是向量空间 E 的两个子空间. 如果 $E = E_1 + E_2$ 并且 $E_1 \cap E_2 = \{0\}$. 则把 E 叫做 E_1, E_2 的直和, 记作 $E = E_1 \oplus E_2$. 其中我们把 E_1, E_2 叫做 E 的互补子空间.

直和有三条等价条件

- (1) $E_1 \cap E_2 = \{0\}$
- (2) $\dim(E) = \dim(E_1) + \dim(E_2)$
- (3) E 中任意向量都可以唯一地分解成分别属于 E_1, E_2 中的向量之和

下面是几个向量空间的例子:

例 1.1 (函数空间) 若 X 是一个非空集合, E 是一个向量空间. 记 F 是 $X \mapsto E$ 的所有函数构成的空间. 如果加法和标量乘法满足下面的规则:

- (1) $(f + g)(x) = f(x) + g(x)$
- (2) $(\lambda f)(x) = \lambda f(x)$

则 F 构成向量空间, 其中 0 矢量是对于 X 中每一个元素 x 都满足 $f(x) = 0$, 映射到 \mathbb{E} 中的 0 向量所对应的映射.

例 1.2 Ω 是 R^N 的一个开子集, 所有 $\Omega \mapsto \mathbb{C}$ 的函数, 满足下列条件构成子空间:

- (1) $\mathcal{C}(\Omega)$ = 定义在 Ω 上的所有连续复函数的空间
- (2) $\mathcal{C}^k(\Omega)$ = 定义在 Ω 上的所有 k 阶连续可偏导复函数的空间
- (3) $\mathcal{C}^\infty(\Omega)$ = 定义在 Ω 上的所有无穷阶可微复函数的空间
- (4) $\mathcal{P}(\Omega)$ = 定义在 Ω 上的所有 N 元多项式的空间

例 1.3 (序列空间) 如果在1.1中的非空集合 X 是正整数集, 那么对应的函数空间叫做**序列空间**. 加法和标量乘法如下定义:

- (1) $(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots)$
- (2) $\lambda(x_1, x_2, \dots) = (\lambda x_1, \lambda x_2, \dots)$

所有复序列构成的空间是一个向量空间.有界复序列的空间是这个空间的真子空间.收敛复序列的空间是有界序列空间的真子空间

我们用 (x_n) 或者 (x_1, x_2, \dots) 来表示序列的第 n 项 x_n ,用 $\{x_n : n \in \mathbb{N}\}$ 去表示序列元素的集合.值得注意的是, 尽管 (x_n) 是无穷序列, 但 $\{x_n : n \in \mathbb{N}\}$ 可以是有限集合. 在大多数情况下, 证明一个空间是向量空间是容易的或者平凡的.下面给两个例子.

例 1.4 (l^p 空间) 无穷复序列 (z_n) 满足 $\sum_{n=1}^{\infty} |z_n|^p < \infty$ 的空间记作 **l^p 空间** ($p \geq 1$)

证明 1.1 l^p 空间是复序列空间的子空间, 如果 $(x_n), (y_n) \in l^p$, 很容易得到 $(x_n + y_n) \in l^p$ 和 $(\lambda x_n) \in l^p$

首先显然

$$\sum_{n=1}^{\infty} |\lambda x_n|^p = |\lambda|^p \sum_{n=1}^{\infty} |x_n|^p < \infty$$

而条件 $\sum_{n=1}^{\infty} |x_n + y_n|^p < \infty$ 则可以用Minkowski不等式得到:

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{\frac{1}{p}} \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{1}{p}}$$

Minkowski不等式建立于Hölder不等式.读者可以自证,以下是参考证明:

定理 1.1 (Hölder不等式) 设 $p > 1, q > 1$ 并且有 $\frac{1}{p} + \frac{1}{q} = 1$, 如果 $(x_n) \in l^p$ 以及 $(y_n) \in l^q$, 那么有:

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} |y_n|^q\right)^{\frac{1}{q}}$$

证明 1.2 不失一般性, 我们假设 $\sum_{n=1}^{\infty} |x_n| \neq 0, \sum_{n=1}^{\infty} |y_n| \neq 0$. 首先我们观察:

$$x^{\frac{1}{p}} \leq \frac{1}{p}x + \frac{1}{q}$$

对于 $0 < x < 1$ 成立. 设 a, b 是非负数, 且有 $a^p \leq b^q$. 这样就有 $0 \leq \frac{a^p}{b^q} \leq 1$. 因此我们有:

$$ab^{-\frac{q}{p}} \leq \frac{1}{p} \frac{a^p}{b^q} + \frac{1}{q}$$

因为 $-\frac{q}{p} = 1 - q$.

$$ab^{1-q} \leq \frac{1}{p} \frac{a^p}{b^q} + \frac{1}{q}$$

两边同时乘以 b^q , 得到

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad (1)$$

证明1用了假设条件 $a^p \leq b^q$. 同理可以证明1同样对于 $a^p \geq b^q$ 也成立. 对于1我们假设:

$$a = \frac{x_j}{(\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}} \quad b = \frac{y_j}{(\sum_{k=1}^n |y_k|^q)^{\frac{1}{q}}}$$

其中 $n \in \mathbb{N}$, 我们得到

$$\frac{x_j}{(\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}} \frac{y_j}{(\sum_{k=1}^n |y_k|^q)^{\frac{1}{q}}} \leq \frac{1}{p} \frac{|x_j|^p}{\sum_{k=1}^n |x_k|^p} + \frac{1}{q} \frac{|y_j|^q}{\sum_{k=1}^n |y_k|^q}$$

对于 $1 \leq j \leq n$. 叠加上式, 得到

$$\frac{\sum_{k=1}^n |x_j| |y_j|}{(\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}} (\sum_{k=1}^n |y_k|^q)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

让 $n \rightarrow \infty$ 得到Hölder不等式

定理 1.2 (Minkowski不等式) 设 $p \geq 1$. 如果 $(x_n), (y_n) \in l^p$, 那么

$$(\sum_{n=1}^{\infty} |x_n + y_n|^p)^{\frac{1}{p}} \leq (\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}} + (\sum_{n=1}^{\infty} |y_n|^p)^{\frac{1}{p}}$$

证明 1.3 对于 $p = 1$ 很显然退化三角不等式. 如果 $p > 1$, 那么存在 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$. 根据Hölder不等式, 有:

$$\begin{aligned} \sum_{n=1}^{\infty} |x_n + y_n|^p &\leq \sum_{n=1}^{\infty} |x_n + y_n| |x_n + y_n|^{p-1} \\ &\leq \sum_{n=1}^{\infty} |x_n| |x_n + y_n|^{p-1} + \sum_{n=1}^{\infty} |y_n| |x_n + y_n|^{p-1} \\ &\leq (\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} |x_n + y_n|^{q(p-1)})^{\frac{1}{q}} \\ &\quad + (\sum_{n=1}^{\infty} |y_n|^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} |x_n + y_n|^{q(p-1)})^{\frac{1}{q}} \end{aligned}$$

因为 $q(p-1)=p$,所以有:

$$\sum_{n=1}^{\infty} |x_n + y_n|^p \leq ((\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}} + (\sum_{n=1}^{\infty} |y_n|^p)^{\frac{1}{p}})(\sum_{n=1}^{\infty} |x_n + y_n|^p)^{\frac{1}{q}}$$

两边同时除以 $(\sum_{n=1}^{\infty} |x_n + y_n|^p)^{\frac{1}{q}}$ 后, *Minkowski*不等式得证.

定义 1.7 (向量空间的笛卡尔乘积) 设 E_1, \dots, E_n 是在相同数域 \mathbf{F} 上的向量空间, 定义:

$$E = (x_1, \dots, x_n) : x_1 \in E_1, x_2 \in E_2, \dots, x_n \in E_n$$

并且有:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

E 是一个向量空间, 叫做 E_1, \dots, E_n 的笛卡尔乘积. 记作 $E = E_1 \times E_2 \times \dots \times E_n$

设 x_1, \dots, x_n 是向量空间 E 的元素. 如果存在标量 $\alpha_1, \dots, \alpha_k$ 满足:

$$x = \alpha_1 x_1 + \dots + \alpha_k x_k$$

则向量 $x \in E$ 叫做向量 x_1, \dots, x_n 的线性组合. 例如 \mathbb{R}^n 中的元素是

$$e_1 = (1, 0, \dots, 0, 0), \quad e_2 = (0, 1, \dots, 0, 0), \dots, \quad e_n = (0, 0, \dots, 0, 1)$$

的线性组合.

类似的, k 阶多项式是 $1, x, x^2, \dots, x^k$ 的线性组合

定义 1.8 (线性无关) 如果 $\alpha_1 x_1 + \dots + \alpha_k x_k = 0$ 当且仅当 $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, 那么向量 x_1, x_2, \dots, x_k 线性无关

1.3 赋范空间

范数

定义 1.9 (范数) 向量空间 $E \mapsto \mathbb{R}$ 的函数 $x \mapsto \|x\|$ 被称为范数, 如果满足下列条件:

(a) $\|x\| = 0$, 当且仅当 $x = 0$

(b) 对于任意 $x \in E, \lambda \in F$, 有 $\|\lambda x\| = |\lambda| \|x\|$

(c) 对于任意 $x, y \in E$, 有 $\|x + y\| \leq \|x\| + \|y\|$

条件(c)叫做三角不等式

例 1.5 函数

$$\|z\| = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}, \quad z = (z_1, \dots, z_n) \in \mathbb{C}^n$$

定义了 \mathbb{C}^n 上的一个范数. 这个范数叫做 *Euclidean* 范数. 下列同样也是 \mathbb{C}^n 的范数:

$$\|z\| = |z_1| + |z_2| + \cdots + |z_n|$$

$$\|z\| = \max\{|z_1|, \dots, |z_n|\}$$

例 1.6 设 Ω 是 \mathbb{R}^n 上的有界闭子集. 函数 $\|f\| = \max_{x \in \Omega} |f(x)|$ 定义了 $\mathcal{C}(\Omega)$ 上的范数

例 1.7 设 $z = (z_n) \in l^p$. $\|z\| = (\sum_{n=1}^{\infty} |z_n|^p)^{\frac{1}{p}}$ 定义了 l^p 上的一个范数 ($p > 1$), *Minkowski* 不等式保证了这一点.

赋范空间

定义 1.10 具有范数的向量空间叫做赋范空间

在同一个向量空间上可能定义不同的范数. 因此定义赋范空间需要具体的向量空间和范数定义. 我们也常说赋范空间指的是一对 $(E, \|\cdot\|)$, 其中 E 是指向量空间, $\|\cdot\|$ 是指定义在 E 上的范数. 一些向量空间也具有标准的范数. 但我们谈论到 \mathbb{R}^n 的时候, 我们立即意识到 *Euclidean* 范数:

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

同样类似的, 1.6、1.7 里定义的范数也是标准的. 如果我们想要考虑这些空间上的不同范数, 我们就必须说清楚 “考虑...上具有...范数的空间”.

指的注意, 赋范空间的子空间也要有相同的范数定义. 绝对值是 \mathbb{R} 和 \mathbb{C} 上的一个范数. 因为两个数差的绝对值表示是这些数的距离, 并且收敛是指 “与极限点的接近”, 所以绝对值通常用于定义收敛. 范数也扮演同样的角色. 当 $\|x\|$ 被解释为 x 的大小时, $\|x - y\|$ 提供了 x 和 y 之间的距离尺度.

赋范空间中的收敛

定义 1.11 设 $(E, \|\cdot\|)$ 是一个赋范空间, 如果对于任意 $\varepsilon > 0$, 存在一个数 M , 使得对于每一个 $n \geq M$ 都有 $\|x_n - x\| < \varepsilon$, 我们说 E 中元素序列 (x_n) 收敛于 $x \in E$. 在这样的情况下, 可以写作 $\lim_{n \rightarrow \infty} x_n = x$ 或者简单写作 $x_n \rightarrow x$

赋范空间上的收敛有和 \mathbb{R} 上收敛的相同性质:

1. 如果 $x_n \rightarrow x$ 和 $\lambda_n \rightarrow \lambda$ (λ_n, λ 是数域里的标量), 那么有 $\lambda_n x_n \rightarrow \lambda x$
2. 如果 $x_n \rightarrow x$ 和 $y_n \rightarrow y$, 那么有 $x_n + y_n \rightarrow x + y$

证明方法与 \mathbb{R} 中收敛的方法一样.

向量空间 E 中的范数诱导向量空间 E 的收敛. 换句话说, 如果我们有一个赋范空间 E , 我们自然地就在 E 中定义了收敛. 很自然的我们会想到, 反过来, 如果给定了 E 上的收敛, 是否可以找到 E 上的范数, 这个范数可以定义这个收敛. 答案是不一定.

例 1.8 考虑定义在有界闭集 $\Omega \subset \mathbb{R}^n$ 上的所有连续函数空间 $C(\Omega)$. 若对于任意 $\varepsilon > 0$, 存在一个常数 n_0 , 使得对于所有 $x \in \Omega$ 和 $n \geq n_0$ 都有 $|f(x) - f_n(x)| < \varepsilon$ 成立, 我们说序列 $f_1, f_2, \dots, f_n \in C(\Omega)$ 一致收敛于 f . 显然 1.6 中的范数定义了序列 f_n 一致收敛于 f , 当且仅当 $\|f_n - f\| = \max_{x \in \Omega} |f_n(x) - f(x)| \rightarrow 0$ ($n \rightarrow \infty$). 这个范数叫做一致收敛范数

范数的等价

定义 1.12 定义在相同向量空间的两个范数, 如果他们定义相同的收敛, 那么被称为等价范数. 更详细的说, 定义在 E 上的范数 $\|\cdot\|_1$ 和 $\|\cdot\|_2$, 如果对于在 E 上的任意序列 (x_n) 和 $x \in E$ 都有,

$$\|x_n - x\|_1 \rightarrow 0 \quad \text{当且仅当} \quad \|x_n - x\|_2 \rightarrow 0$$

那么被称为等价的.

例 1.9 下列定义在 \mathbb{R}^2 上的范数是等价的

$$\begin{aligned} \|(x, y)\|_1 &= \sqrt{x^2 + y^2} & \|(x, y)\|_2 &= |x| + |y| \\ \|(x, y)\|_3 &= \max |x|, |y| \end{aligned}$$

接下来将会证明定义在任意有限维向量空间的范数都是等价的. 证明需要一些准备.

定理 1.3 设 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 是向量空间的两个范数. $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 被称为等价的, 当且仅当存在正数 α 和 β , 使得

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1 \quad x \in E$$

证明 1.4 假设两个范数 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 等价, 也就是说 $\|x_n - x\|_1 \rightarrow 0$ 当且仅当 $\|x_n - x\|_2 \rightarrow 0$. 假设不存在 $\alpha > 0$ 使得对于任意 $x \in E$ 都有 $\alpha\|x\|_1 \leq \|x\|_2$. 那么对于任意 $n \in \mathbb{N}$ 都存在 $x_n \in E$ 使得

$$\frac{1}{n}\|x_n\|_1 > \|x_n\|_2$$

定义

$$y_n = \frac{1}{\sqrt{n}} \frac{x_n}{\|x_n\|_2}$$

显然 $\|y_n\|_2 = \frac{1}{\sqrt{n}} \rightarrow 0$. 另一方面, $\|y_n\|_1 \geq n\|y_n\|_2 \geq \sqrt{n}$. 这个矛盾表明了必须存在 $\alpha > 0$ 满足定理左半边. 同样的可以证明 β 的存在.

引理 1.1 如果 x_1, \dots, x_n 是赋范空间中线性无关的元素, 则存在常数 $c > 0$ 使得

$$\|\alpha_1 x_1 + \dots + \alpha_n x_n\| \geq c(|\alpha_1| + \dots + |\alpha_n|) \quad (2)$$

对于所有的 $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ 都成立

证明 1.5 当 $|\alpha_1| + \dots + |\alpha_n| = 0$ 时, (2) 对与任意 c 都成立. 所以 (2) 等价于

$$\|\beta_1 x_1 + \dots + \beta_n x_n\| \geq c, \quad |\beta_1| + \dots + |\beta_n| = 1$$

连续函数 $f: \mathbb{R}^n \mapsto \mathbb{R}$ 定义为:

$$f(\beta_1, \dots, \beta_n) = \|\beta_1 x_1 + \dots + \beta_n x_n\|$$

因为集合 $B = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n: |\beta_1| + \dots + |\beta_n| = 1$ 是有界闭集. f 在 B 上取得最小值. 值得注意最小值非零, 因为 $\beta_1 x_1 + \dots + \beta_n x_n = 0, |\beta_1| + \dots + |\beta_n| = 1$ 与 x_1, \dots, x_n 线性无关矛盾. 因此

$$c = \min_{(\beta_1, \dots, \beta_n) \in B} f(\beta_1, \dots, \beta_n) = \min_{(\beta_1, \dots, \beta_n) \in B} \|\beta_1 x_1 + \dots + \beta_n x_n\| > 0$$

证明完毕

定理 1.4 如果 E 是有限维向量空间, 则该空间任意两个范数都等价.

证明 1.6 E 是有限维向量空间, $\{e_1, \dots, e_n\}$ 是 E 的基. 定义 $\|\cdot\|_0$ 范数如下:

$$\|\alpha_1 e_1 + \dots + \alpha_n e_n\|_0 = |\alpha_1| + \dots + |\alpha_n|$$

利用 1.3

$$\begin{aligned} \|\alpha_1 e_1 + \dots + \alpha_n e_n\| &\leq |\alpha_1| \|e_1\| + \dots + |\alpha_n| \|e_n\| \\ &\leq \max \|e_1\|, \dots, \|e_n\| (|\alpha_1| + \dots + |\alpha_n|) \end{aligned}$$

所以对于任意 $x \in E$ 都有

$$\|x\| \leq \beta \|x\|_0$$

其中 $\beta = \max \|e_1\|, \dots, \|e_n\|$. 由于引理 1.1 立即可以得到, 对于任意 $x \in E$

$$\alpha \|x\|_0 \leq \|x\|$$

证明完毕

如果定义距离函数 $d(x, y) = \|x - y\|$, 则赋范空间 $(E, \|\cdot\|)$ 成为度规空间. 由范数定义的收敛和由这个度规定义的收敛是一样的. E 中的度规定义了 E 中的拓扑. 拓扑概念可以通过开集去定义, 没有必要先定义度规, 然后用度规去定义拓扑. 下面我介绍一些赋范空间的拓扑性质.

开球、闭球和球面

定义 1.13 x 是度量空间 E 里的元素, r 是正数, 则有如下概念:

1. $B(x, y) = y \in E : \|y - x\| < r$ (开球)
2. $B(x, y) = y \in E : \|y - x\| \leq r$ (闭球)
3. $B(x, y) = y \in E : \|y - x\| = r$ (球面)

x 叫做球心, r 叫做半径

例 1.10 下面是几个 R^2 中球的例子.

$$\begin{aligned} \|(x, y)\|_1 &= \sqrt{x^2 + y^2}, \quad \|(x, y)\|_2 = |x| + |y| \\ \|(x, y)\|_3 &= \max |x|, |y| \end{aligned}$$

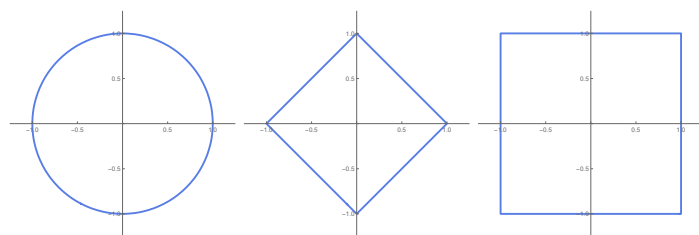


图 1: \mathbb{R}^2 中球的例子, $\|(x,y)\|_1 \leq 1, \|(x,y)\|_2 \leq 1, \|(x,y)\|_3 \leq 1$

开集和闭集

定义 1.14 S 是赋范空间 E 的子集, 若对于任意 $x \in S$ 都存在 $\varepsilon > 0$, 使得 $B(x, \varepsilon) \subseteq S$, 则称 S 为开集. 如果子集 S 的补集 $(E \setminus S)$ 是开集, 则称 S 为闭集.

很容易意识到, 尽管球的定义不同, 但范数的等价定义了相同的开集. 开集同样的定义了闭集、稠集和紧集以及其他的拓扑概念.

例 1.11 开球是开集. 闭球和球面是闭集.

定理 1.5

- (a) 开集的任意并集是开集
- (b) 开集的有限交集是开集
- (c) 闭集的有限并集是闭集
- (d) 闭集的任意交集是闭集
- (e) 空集和全集既是开也是闭的

定理 1.6 赋范空间 E 的子集 S 是闭的, 当且仅当 S 的元素序列在 E 中收敛, 并且极限在 S 中.

$$x_1, x_2, \dots, x_n \in S, \quad x_n \rightarrow x, \quad x \in S$$

证明 1.7 假设 S 是 E 的闭子集, $x_1, x_2, \dots \in S, x_n \rightarrow x$ 且 $x \notin S$. 因为 S 是闭集, $E \setminus S$ 是开集. 所以存在 $\varepsilon > 0$ 使得 $B(x, \varepsilon) \subseteq E \setminus S$. 换句话说, 因为 $\|x - x_n\| \rightarrow 0$, 对于充分大的 $n \in \mathbb{N}$, 我们有 $\|x - x_n\| < \varepsilon$. 这个矛盾表明了 $x \in S$.

现在假设 $x_1, x_2, \dots \in S$, 并且有 $x_n \rightarrow x, x \in S$. 如果 S 不是闭集, 则 $E \setminus S$ 非开. 所以存在 $x \in E \setminus S$ 使得任意球 $B(x, \varepsilon)$ 都包含 S 中的元素. 这样我们能找到 $x_1, x_2, \dots \in S$ 使得 $x_n \in B(x, \frac{1}{n})$. 但另一方面 $x_n \rightarrow x$ 以及假设条件 $x \in S$. 这与 $x \in E \setminus S$ 假设相互矛盾. 因此 S 必须是闭集.

闭包

定义 1.15 S 是赋范空间 E 的子集, S 的闭包定义为所有闭集的交, 记作 clS . 表示最小的闭集.

定理 1.7 S 是赋范空间 E 的子集, 闭包 clS 是指 S 中所有收敛列极限的集合.

$$clS = \{x \in E : \exists x_1, x_2, \dots \in S \text{ s.t. } x_n \rightarrow x\}$$

读者自证

稠密

定义 1.16 赋范空间 E 有子集 S , 如果 $clS = E$, 则 S 被称为稠密的,

例 1.12 $[a, b]$ 上所有多项式的集合在 $C([a, b])$ 上稠密. 有限项非零项的所有复数列的集合在 $l^p (p > 1)$ 上稠密.

定理 1.8 S 是赋范空间 E 的子集, 以下条件相互等价:

- (a) S 在 E 中稠密
- (b) 对于每一个 $x \in E$, 都存在 $x_1, x_2, \dots \in S$, 使得 $x_n \rightarrow x$
- (c) 对于 E 的每一个非空子集至少包含 S 中的一个元素

紧致

定义 1.17 赋范空间 E 有子集 S , 如果 S 中的每一个序列 (x_n) 都包含极限在 S 中的收敛子序列, 则称 S 为紧致集合

例 1.13 R^n, C^n 中的有界闭集紧致

定理 1.9 当且仅当闭单位球是紧致的时候, 赋范空间 E 是有限维的

有界

定义 1.18 赋范空间 E 中有子集 S , 如果存在 $r > 0, S \subseteq B(0, r)$, 则称 S 是有界子集

定理 1.10 紧致集合是有界闭的.

定理 1.11 当且仅当 E 中的单位球是紧致的, 赋范空间 E 是有限维的.

读者自证

1.4 Banach空间

Cauchy序列

定义 1.19 如果对于任意 $\varepsilon > 0$ 存在数 M 使得 $\|x_m - x_n\| < \varepsilon$ 对于 $m, n > M$ 都成立, 则称赋范空间中的向量序列 (x_n) 为 *Cauchy* 序列.

Augustin Louis Cauchy(1789-1857)

定理 1.12 下列条件相互等价

(a) (x_n) 是 *Cauchy* 序列.

(b) 对于任意一对递增正整数序列 $(p_n), (q_n)$, 都有 $\|x_{p_n} - x_{q_n}\| \rightarrow 0, n \rightarrow \infty$ 成立

(c) 对于任意递增正整数序列 (p_n) , 都有 $\|x_{p_{n+1}} - x_{p_n}\| \rightarrow 0, n \rightarrow \infty$ 成立

显然每一个收敛序列都是 *Cauchy* 序列, 事实上若 $\|x_n - x\| \rightarrow 0$, 则

$$\|x_{p_n} - x_{q_n}\| \leq \|x_{p_n} - x\| + \|x_{q_n} - x\| \rightarrow 0$$

对于每一个递增序列 $(p_n), (q_n)$ 都成立. 收敛序列是 *Cauchy* 序列, 但 *Cauchy* 序列不一定是收敛序列.

例 1.14 设 $\mathcal{P}([0, 1])$ 是 $[0, 1]$ 上的多项式空间, 并且有一致收敛的范数 $\|P\| = \max_{[0, 1]} |P(x)|$. 定义

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

其中 $n = 1, 2, \dots$, (P_n) 是一个 *Cauchy* 序列, 但显然并不收敛于 $\mathcal{P}([0, 1])$, 因为极限不是一个多项式.

引理 1.2 若 (x_n) 是赋范空间的一个 *Cauchy* 序列, 则序列 $(\|x_n\|)$ 收敛.

证明 1.8 因为 $|||x|| - ||y||| \leq \|x - y\|$, 所以有 $|||x_m|| - ||x_n||| \leq \|x_m - x_n\| \rightarrow 0, (m, n \rightarrow \infty)$. 范数序列是实柯西序列, 因此它收敛.

Banach空间

定义 1.20 如果每一个 *Cauchy* 序列都收敛于 E , 则称 E 是完备的. 完备的赋范空间称为 *Banach* 空间

例 1.15 l^2 是完备空间. 设 (a_n) 是 l^2 的柯西序列, 如果

$$a_n = (\alpha_{n,1}, \alpha_{n,2}, \dots)$$

对于给定的任意 $\varepsilon > 0$ 都存在数 n_0 使得

$$\sum_{k=1}^{\infty} |\alpha_{m,k} - \alpha_{n,k}|^2 < \varepsilon^2$$

对于所有的 $m, n \geq n_0$ 都成立. 值得注意, 这实际上是对于每一个固定的 $k \in \mathbb{N}$ 和每一个 ε 都存在一个数 n_0 使得

$$|\alpha_{m,k} - \alpha_{n,k}| < \varepsilon$$

对于任意 $m, n \geq n_0$ 都成立. 这也意味着对于任意 k , 序列 $(\alpha_{n,k})$ 是 \mathbb{C} 里的柯西序列, 所以是收敛序列. 记

$$\alpha_k = \lim_{n \rightarrow \infty} \alpha_{n,k}, \quad k = 1, 2, \dots \quad a = (\alpha_n)$$

进一步证明 a 是 l^2 的元素, 序列 (a_n) 收敛于 a . 根据 (1.15), 令 $m \rightarrow \infty$ 得到

$$\sum_{k=1}^{\infty} |\alpha_k - \alpha_{n,k}|^2 \leq \varepsilon^2$$

对于任意 $n \geq n_0$ 都成立. 因为

$$\sum_{k=1}^{\infty} |\alpha_{n_0,k}|^2 < \infty$$

再利用 Minkowski 不等式

$$\begin{aligned} \sqrt{\sum_{k=1}^{\infty} |\alpha_k|^2} &= \sqrt{\sum_{k=1}^{\infty} (|\alpha_k| - |\alpha_{n_0,k}| + |\alpha_{n_0,k}|)^2} \\ &\leq \sqrt{\sum_{k=1}^{\infty} (|\alpha_k| - |\alpha_{n_0,k}|)^2} + \sqrt{\sum_{k=1}^{\infty} |\alpha_{n_0,k}|^2} \\ &\leq \sqrt{\sum_{k=1}^{\infty} |\alpha_k - \alpha_{n_0,k}|^2} + \sqrt{\sum_{k=1}^{\infty} |\alpha_{n_0,k}|^2} < \infty \end{aligned}$$

这就证明了序列 $a = (\alpha_n)$ 是 l^2 的元素. 因为 ε 是任意小的, 所以有

$$\lim_{n \rightarrow \infty} \|a - a_n\| = \lim_{n \rightarrow \infty} \sqrt{\sum_{k=1}^{\infty} (|\alpha_k| - |\alpha_{n,k}|)^2} = 0$$

序列 (a_n) 收敛于 a ,且 $a \in l^2$

收敛与绝对收敛

定义 1.21 赋范空间 E 中有级数 $\sum_{n=1}^{\infty} x_n$, 如果序列的部分和收敛于 E , 也就是说, 存在 $x \in E$ 使得当 $n \rightarrow \infty$ 时 $\|x_1 + x_2 + \cdots + x_n - x\| \rightarrow 0$ 成立, 则叫做收敛级数. 这种情况下记作 $\sum_{n=1}^{\infty} x_n = x$. 如果 $\sum_{n=1}^{\infty} \|x_n\| < \infty$, 则称级数为绝对收敛.

定理 1.13 Banach空间的闭子空间是Banach空间自己

1.5 线性映射

我们先介绍一些概念. 设 E_1 和 E_2 是两个向量空间, L 是 E_1 到 E_2 的映射. 如果 $y = L(x)$, 则称 y 是 x 的像.

如果 A 是 E_1 的子集, 则 $L(A)$ 表示 A 的像集, 也就是说, $L(A)$ 是 A 中元素的像的集合, 并且像在 E_2 里. 如果 B 是 E_2 的子集, 则 $L^{-1}(B)$ 表示 B 的原像.

$$L(A) = \{L(x) : x \in A\}, \quad L^{-1}(B) = \{x \in E_1 : L(x) \in B\}$$

值得注意的是 L^{-1} 并不意味着 L 是可逆的.

我们通常考虑的映射是定义在向量空间的真子空间上的. 所以提出 L 的定义域这个概念是很重要的, 我们用 $\mathcal{D}(L)$ 表示 L 的定义域. 集合 $L(\mathcal{D}(L))$ 叫做 L 的值域, 用 $\mathcal{R}(L)$ 表示.

$$\mathcal{R}(L) = \{y \in E_2 : L(x) = y, x \in \mathcal{D}(L)\}$$

L 的零空间记作 $\mathcal{N}(L)$, 这指的是所有的 $x \in \mathcal{D}(L)$ 使得 $L(x) = 0$. 在最后简单介绍一下映射的图, 用 $\mathcal{G}(L)$ 表示映射 L 的图, 指的是如下定义的 $E_1 \times E_2$ 的子集

$$\mathcal{G}(L) = \{(x, y) : x \in \mathcal{D}(L), y = L(x)\}$$

线性映射

定义 1.22 如果映射 $L : E_1 \rightarrow E_2$ 对于任意的 $x, y \in E_1, \alpha, \beta \in \mathbf{F}$ 满足 $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$, 则称 L 为线性映射

在向量空间中, 我们常用 Lx 的写法替代 $L(x)$.

连续映射

定义 1.23 设 E_1 和 E_2 是赋范空间, F 是 E_1 到 E_2 的映射, 如果对于所有 E_1 中的序列 (x_n) 都收敛于 x_0 , 有序列 $(F(x_n))$ 收敛于 $F(x_0)$, 即如果 $\|x_n - x_0\| \rightarrow 0$ 意味着 $\|F(x_n) - F(x_0)\| \rightarrow 0$, 那么就说 F 在 x_0 连续. 如果 F 在任意 $x \in E_1$ 都连续, 则称 F 是连续的.

我们将会在量子力学里面遇到很多连续性问题.

例 1.16 赋范空间 E 的范数是 $E \mapsto \mathbb{R}$ 的连续映射.

证明 1.9 如果有 $\|x_n - x\| \rightarrow 0$, 则 $|\|x_n\| - \|x\|| \leq \|x_n - x\| \rightarrow 0$

定理 1.14 设 $F: E_1 \mapsto E_2$. 下列条件相互等价.

(a) F 是连续的

(b) E_2 的任意开子集 U 的原象 $F^{-1}(U)$ 在 E_1 中仍然是开的.

(c) E_2 的任意闭子集 S 的原象 $F^{-1}(S)$ 在 E_1 中仍然是闭的.

定理 1.15 如果存在 $x_0 \in E_1$, 有线性映射 $L: E_1 \mapsto E_2$ 在 x_0 连续, 则线性映射 L 是连续映射的.

证明 1.10 假设 L 在 $x_0 \in E_1$ 连续. 设 x 是 E_1 的任意一个元素 (x_n) 是收敛于 x 的序列. 自然地有序列 $(x_n - x + x_0)$ 收敛于 x_0 , 这样我们有 $\|Lx_n - Lx\| = \|L(x_n - x + x_0) - Lx_0\| \rightarrow 0$. 得证

有界线性映射

定义 1.24 线性映射 $L: E_1 \mapsto E_2$, 如果存在数 $\alpha > 0$ 使得对于所有的 $x \in E_1$ 都有 $\|Lx\| \leq \|x\|$, 则 L 叫做有界线性映射.

值得注意的是, 这个定义相当于说 L 在 E_1 的单位球面上以 α 为界, 也就是说对于所有的 $x \in E_1$ 有 $\|Lx\| \leq \alpha$ 使得 $\|x\| = 1$

定理 1.16 线性映射是连续的, 当且仅当它是有界的.

证明 1.11 如果 L 是有界线性映射, 且有 $\|x_n\| \rightarrow 0$, 则 $\|Lx_n\| \leq \alpha\|x_n\| \rightarrow 0$. 即线性映射 L 在0连续, 根据定理1.15, L 是连续映射. 如果 L 不是有界的, 那么对于任意 $n \in \mathbb{N}$, 存在 $x_n \in E_1$ 使得 $\|Lx_n\| > n\|x_n\|$. 定义序列:

$$y_n = \frac{x_n}{n\|x_n\|}$$

显然 $y_n \rightarrow 0$, 而对于任意 $n \in N$ 有 $\|Ly_n\| > 1$, 所以 L 不是连续映射

有限维空间的线性映射是有界的, 上述定理也意味着对于线性映射连续和一致连续是等价的.

如果加法和数乘如下定义, 向量空间 E_1 到 E_2 的所有线性映射也构成向量空间:

$$(L_1 + L_2)x = L_1x + L_2x, \quad (\lambda L)x = \lambda(Lx)$$

如果 E_1 和 E_2 是赋范空间, 则所有 $E_1 \rightarrow E_2$ 的有界线性映射的集合是上面定义的空间的线性子空间, 记作 $\mathcal{B}(E_1, E_2)$

定理 1.17 如果 E_1 和 E_2 是赋范空间, 则 $\mathcal{B}(E_1, E_2)$ 也是赋范空间, 范数如下定义:

$$\|L\| = \sup_{\|x\|=1} \|Lx\|$$

证明 1.12 如果 $L_1, L_2 \in \mathcal{B}(E_1, E_2)$ 有 $x \in E_1$ 使得 $\|x\| = 1$, 则有

$$\|L_1x + L_2x\| \leq \|L_1x\| + \|L_2x\|$$

这意味着

$$\|L_1x + L_2x\| \leq \sup_{\|x\|=1} \|L_1x\| + \sup_{\|x\|=1} \|L_2x\| = \|L_1\| + \|L_2\|$$

因此有,

$$\|L_1 + L_2\| = \sup_{\|x\|=1} \|L_1x + L_2x\| \leq \|L_1\| + \|L_2\|$$

所以上述范数定义的确满足三角不等式.

(1.17) 定义的范数是 $\mathcal{B}(E_1, E_2)$ 的标准范数, 当我们谈论起 $\mathcal{B}(E_1, E_2)$ 的范数时, 指的就是这个范数.

定理 1.18 若 E_1 是赋范空间 E_2 是 Banach 空间, 则 $\mathcal{B}(E_1, E_2)$ 是 Banach 空间

证明 1.13 首先说明 $\mathcal{B}(E_1, E_2)$ 是完备的. (L_n) 是 $\mathcal{B}(E_1, E_2)$ 中的柯西序列, x 是 E_1 的任意元素. 则有:

$$\|L_mx - L_nx\| \leq \|L_m - L_n\| \|x\| \rightarrow 0 \quad m, n \rightarrow \infty$$

这表明了 $(L_n x)$ 是 E_2 中的Cauchy序列. 由于 E_2 的完备性, 存在唯一的元素 $y \in E_2$ 使得 $L_n x \rightarrow y$. 因为 x 是 E_1 的任意元素, 这定义了一个从 E_1 到 E_2 的映射 L :

$$Lx = \lim_{n \rightarrow \infty} L_n x$$

接下来再证明 $L \in \mathcal{B}(E_1, E_2)$ 和 $\|L_n - L\| \rightarrow 0$. 显然, L 是一个线性映射. 因为Cauchy序列有界, 存在一个常数 α 使得对于所有的 $n \in \mathbb{N}$ 有 $\|L_n\| \leq \alpha$ 成立. 所以

$$\|Lx\| = \|\lim_{n \rightarrow \infty} L_n x\| = \lim_{n \rightarrow \infty} \|L_n x\| \leq \alpha \|x\|$$

因此 L 也是有界的, 这样 $L \in \mathcal{B}(E_1, E_2)$. 接下来还有 $\|L_n - L\| \rightarrow 0$ 需要证明. 令 $\varepsilon > 0$, 存在 k 使得对于所有的 $m, n \geq k$ 都有 $\|L_m - L_n\| < \varepsilon$. 如果 $\|x\| = 1, m, n \geq k$, 则

$$\|L_m x - L_n x\| \leq \|L_m - L_n\| < \varepsilon$$

令 $n \rightarrow \infty$ 得到对于任意 $m \geq k, x \in E_1, \|x\| = 1$ 有 $\|L_m x - Lx\| \leq \varepsilon$. 所以对于所有的 $m > k, \|L_m - L\| \leq \varepsilon$. 得证

定理 1.19 (对角定理) E 是一个赋范空间, $(x_{ij}), (i, j \in \mathbb{N})$ 是 E 中的无限维矩阵的矩阵元. 若有:

(a) 对于任意 $j \in \mathbb{N}$ 有 $\lim_{i \rightarrow \infty} x_{ij} = 0$

(b) 对于每一个指标列 p_n 有子列 q_n 使得

$$\lim_{i \rightarrow \infty} \sum_{j=1}^{\infty} x_{q_i q_j} = 0$$

则 $\lim_{i \rightarrow \infty} x_{ii} = 0$

留作练习, 读者自证. 下面简单介绍一下压缩映射和不动点定理

压缩映射

定义 1.25 赋范空间 E 的子集 A 到 E 之间有一映射 f , 如果存在正数 $\alpha < 1$ 使得对于所有的 $x, y \in A$ 都有:

$$\|f(x) - f(y)\| \leq \alpha \|x - y\|$$

则把 f 叫做压缩映射.

例 1.17 考虑一个非线性方程 $x^3 - x - 1 = 0$. 方程有三个根. 把方程写成 $Tx = x$ 形式有多种方式. 例如

$$Tx = (1+x)^{\frac{1}{3}}, \quad Tx = x^3 - 1, \quad Tx = \frac{1}{x^2 - 1}$$

原方程的根在 $[1, 2]$. 由 $T(x) = (1+x)^{\frac{1}{3}}$ 定义的映射 T 是一个压缩映射. 由均值定理可以得到

$$|Tx - Ty| = |(1+x)^{\frac{1}{3}} - (1+y)^{\frac{1}{3}}| \leq \frac{2^{\frac{1}{3}}}{6} |x - y|$$

其余两个不是压缩映射

Banach不动点定理

首先要介绍一下不动点. 如果映射 f 满足存在点 z 使得 $f(z) = z$, 则称 z 为 f 的不动点. 这个定理在后面的微分积分方程里有很大的用处.

例 1.18 令 $E = C([0, 1])$ 是定义在闭区间 $[0, 1]$ 上的连续复函数空间. 再令 T 如下定义

$$(Tx)(t) = x(0) + \int_0^t x(\tau) d\tau$$

显然, 对于任意 $a \in \mathbb{C}$, 函数 $x(t) = ae^t$ 是 T 的不动点.

定理 1.20 (不动点定理) 令 F 是赋范空间 E 的闭子集, f 是从 F 到 F 的压缩映射, 则存在唯一的 $x \in F$ 使得 $f(x) = x$

证明 1.14 令 $0 < \alpha < 1$ 使得对于所有的 $x, y \in F$ 有

$$\|f(x) - f(y)\| \leq \alpha \|x - y\|$$

令 x_0 是 F 中的任意点, $x_n = f(x_{n-1})$, $n = 1, 2, \dots$. 我们将证明 (x_n) 是 *Cauchy* 序列. 首先观察到, 对于任意 $n \in \mathbb{N}$

$$\|x_{n+1} - x_n\| \leq \alpha \|x_n - x_{n-1}\| \leq \alpha^2 \|x_{n-1} - x_{n-2}\| \leq \dots \leq \alpha^n \|x_1 - x_0\|$$

因此对于任意 $m, n \in \mathbb{N}$ 使得 $m < n$ 有

$$\begin{aligned}\|x_n - x_m\| &\leq \|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\| + \cdots + \|x_{m+1} - x_m\| \\ &\leq (\alpha^{n-1} + \alpha^{n-2} + \cdots + \alpha^m) \|x_1 - x_0\| \\ &= \frac{\|x_1 - x_0\|}{1 - \alpha} \alpha^m \rightarrow 0, \quad m \rightarrow \infty\end{aligned}$$

这样 (x_n) 是 *Cauchy* 序列. 因为 F 是完备空间的闭子集, 所以存在 $z \in F$ 使得当 $n \rightarrow \infty$ 时 $x_n \rightarrow z$. 接下来证明唯一存在 z 使得 $f(z) = z$. 首先因为

$$\begin{aligned}\|f(z) - z\| &\leq \|f(z) - x_n\| + \|x_n - z\| \\ &= \|f(z) - f(x_{n-1})\| + \|x_n - z\| \\ &\leq \alpha \|x - z_{n-1}\| + \|x_n - z\| \rightarrow 0, \quad n \rightarrow \infty\end{aligned}$$

$\|f(z) - z\| = 0$, 这样就有 $f(z) = z$. 假设有另一个点 $w \in F, f(w) = w$, 则

$$\|z - w\| = \|f(z) - f(w)\| \leq \alpha \|z - w\|$$

因为 $0 < \alpha < 1$, 必有 $\|z - w\| = 0$, 这意味着 $z = w$

1.6 线性空间的同构

设 a_1, a_2, \dots, a_n 是线性空间 E 的一组基, 在这组基下, E 中每一个向量都有确定的坐标, 而向量的坐标可以看成 \mathbf{F}^n 的元素. 因此向量与它的坐标之间的对应实质上就是 E 到 \mathbf{F} 的一个映射, 这个映射是单射且满射的. 坐标给出了线性空间 E 与 \mathbf{F}^n 的双射. 这个对应最重要体现在运算关系的对应上.

设

$$\begin{aligned}\mathbf{x} &= x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = \sum_{i=1}^n x_i a_i \\ \mathbf{y} &= y_1 a_1 + y_2 a_2 + \cdots + y_n a_n = \sum_{j=1}^n y_j a_j\end{aligned}$$

即向量 \mathbf{x}, \mathbf{y} 的坐标分别是 $(x_1, \dots, x_n), (y_1, \dots, y_n)$, 那么

$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1 + y_1) a_1 + \cdots + (x_n + y_n) a_n \\ k\mathbf{x} &= kx_1 a_1 + \cdots + kx_n a_n\end{aligned}$$

向量 $\mathbf{x} + \mathbf{y}, k\mathbf{x}$ 的坐标分别是

$$\begin{aligned}(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) &= (x_1, \dots, x_n) + (y_1, \dots, y_n) \\ (kx_1, kx_2, \dots, kx_n) &= k(x_1, x_2, \dots, x_n)\end{aligned}$$

向量的运算可以归结成它们坐标的运算, 因而线性空间 E 的讨论可以归结于 \mathbf{F}^n 的讨论

定义 1.26 数域 \mathbf{F} 上两个线性空间 E 与 E' 称为同构的, 如果从 E 到 E' 有一个双射 σ . 具有以下性质:

$$(1) \sigma(a + b) = \sigma(a) + \sigma(b)$$

$$(2) \sigma(ka) = k\sigma(a)$$

其中 a, b 是 E 中任意向量, k 是 \mathbf{F} 中任意数, 这样的映射称为同构映射.

同构映射基本性质:

$$(1) \sigma(0) = 0, \sigma(-a) = -\sigma(a)$$

$$(2) \sigma(\sum_{i=1}^n k_i a_i) = \sum_{i=1}^n k_i \sigma(a_i)$$

(3) E 中向量组 a_1, \dots, a_n 线性相关的充分必要条件是, 它们的像 $\sigma(a_1), \sigma(a_2), \dots, \sigma(a_n)$ 线性相关

(4) 如果 E_1 是 E 的一个线性子空间, 那么, E_1 在 σ 下像集合

$$\sigma(E_1) = \{\sigma(a) | a \in E_1\}$$

是 $\sigma(E)$ 的子空间, 并且 $\dim(E_1) = \dim(\sigma(E_1))$

(5) 同构映射的逆以及两个同构映射的乘积还是同构映射

1.7 线性变换

线性变换基本概念

线性空间中, 事物之间的联系反映为线性空间的映射, 线性空间到自身的映射通常称为 E 的一个变换. 线性变换是最简单最基本的一种变换, 线性变换是量子力学里最常用的概念.

定义 1.27 如果对于线性空间 E 中任意元素 a, b 和数域 F 中任意数 k , 都有这样的线性变换 A , 满足:

$$\begin{aligned} A(a + b) &= A(a) + A(b) \\ A(ka) &= kA(a) \end{aligned}$$

则这个变换 A 称为线性变换, 以后使用大写 A, B, \dots 代表 E 的变换, $A(a)$ 或 Aa 代表元素 a 在变换 A 下的像. **线性变换保持加法与数量乘法.**

例 1.19 平面上的向量构成实数域上的二维线性空间. 把平面围绕坐标原点绕逆时针方向转动 θ 角, 就是一个线性变换, 用 A_θ 表示. 如果平面上一个向量 a 在直角坐标系下的坐标是 (x, y) , 那么像 $A_\theta(a)$ 的坐标 (x', y') 为:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

同样的三维空间中绕轴有限转动也是一个线性变换

例 1.20 线性空间 E 中的恒等变换(单位变换)满足:

$$I(a) = a \quad (a \in E)$$

以及零变换 $0: 0(a) = 0 \quad (a \in E)$

线性变换的运算

线性变换作为映射, 利用映射的复合可以定义线性变换的乘法

定义 1.28 设 A, B 是线性空间 E 上的两个线性变换, 定义他们乘积 AB 为

$$(AB)(a) = A(B(a)) \quad (a \in E)$$

容易得到, 线性变换的乘积也是线性变换(读者自证)

通过以上定义可以看出线性变换满足乘法结合律

$$(AB)C = A(BC)$$

但线性变换的乘法一般是不可交换的.

例 1.21 实数域 \mathbb{R} 上的线性空间 $\mathbb{R}[x]$ 中, 线性变换

$$\begin{aligned} D(f(x)) &= f'(x) \\ F(f(x)) &= \int_0^x f(t)dt \end{aligned}$$

的乘积 $DF = I$, 但一般 $FD \neq I$

线性变换还可以定义加法. 设 A, B 是线性空间 E 上的两个线性变换, 定义它们的和 $A + B$ 为

$$(A + B)(a) = A(a) + B(a) \quad a \in E$$

线性变换的和还是线性变换, 不难证明线性变换的加法满足加法结合律与交换律. 对于加法, 零变换 0 有特殊地位, 它与所有线性变换 A 的和仍为 A :

$$A + 0 = A$$

同样可以定义负变换

$$(-A)(a) = -A(a) \quad a \in E$$

显然: $A + (-A) = 0$.

1.8 Hilbert空间

在复习Hilbert空间之前, 我们先复习对偶空间的概念. 设 E 是实数域 \mathbb{R} 上的 n 维向量空间, 可定义 E 上的线性映射 L . 即向量空间 E 与实数域 \mathbb{R} 间的线性映射

$$\begin{aligned} L : E &\mapsto \mathbb{R} \\ x &\mapsto L(x) \in \mathbb{R}, \quad x \in E \end{aligned}$$

线性映射保持线性结构.

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y), \quad x, y \in E, \alpha, \beta \in \mathbb{R}$$

在之前的线性映射章节里, 我们讨论了 $E_1 \mapsto E_2$ 的线性映射的集合也构成向量空间 $\mathcal{B}(E_1, E_2)$. 这里我们把这个向量空间 $\mathcal{B}(E, \mathbb{R})$ 称为向量空间 E 的对偶空间, 记作 E^*

$$L \in E^* = \mathcal{B}(E, \mathbb{R})$$

如果 (e_1, e_2, \dots, e_n) 为向量空间的 E 一组基底, 任意向量 x 可用这组基展开

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n, \quad x_i \in \mathbb{R}, n = 1, 2, \dots$$

由于线性映射保证线性结构不变.

$$Lx = x_1 L e_1 + x_2 L e_2 + \dots + x_n L e_n, \quad x_i \in \mathbb{R}, n = 1, 2, \dots$$

即所有的线性映射 L 都可以由它们在基底 e_i 上取的值决定. 记

$$L_i = L e_i \in \mathbb{R}$$

称为对偶向量 L 的分量, 可将 $L \in E^*$ 记作

$$f = f_1 e_1^* + f_2 e_2^* + \dots + f_n e_n^*$$

其中 $e_i^* \in E^*$, 为向量空间中满足下列条件的线性映射.

$$e_i^*(e_k) = \delta_{ik}$$

显然这组线性映射 $e_1^*, e_2^*, \dots, e_n^*$ 线性无关. 它们形成了 E^* 中的一组基, 称为与基 (e_1, e_2, \dots, e_n) 对偶的基. 对偶空间 E^* 也为 n 维向量空间, 且空间 E 与空间 E^* 相互对偶.

向量空间 E 和 E 上的线性映射空间 E^* 是两个重要概念, 为帮助熟悉他的含义, 我们采用熟悉的Dirac符号去表示. 向量空间 E 上的元素 x 可以用 $|x\rangle$ 去表示. 基组 (e_1, e_2, \dots, e_n) 可以用 $(|1\rangle, |2\rangle, \dots, |n\rangle)$ 去表示.

$$|x\rangle = x_1 |1\rangle + x_2 |2\rangle + \dots + x_n |n\rangle$$

对偶空间 E^* 上的元素 L 记作 $\langle L|$, 基组 $(e_1^*, e_2^*, \dots, e_n^*)$ 记作 $(\langle 1|, \langle 2|, \dots, \langle n|)$

$$\langle L| = L_1 \langle 1| + L_2 \langle 2| + \dots + L_n \langle n|$$

基矢 $\langle i|, i = 1, \dots, n$ 由它对整个向量空间 E 的作用决定

$$\langle i|j\rangle = \delta_{ij}$$

值得注意, E 与 E^* 均为线性空间, 在 E 与 E^* 之间可以定义内积

$$\langle L|x \rangle = \sum_{k=1}^{k=n} L_k x_k$$

内积空间

我们现在开始讨论定义了内积的复向量空间, 内积的定义如下. 这里 $\mathbf{F} = \mathbb{C}$, 对于 $x \in \mathbb{C}$, 复共轭的标记我们用 x^* 表示.

定义 1.29 令 E 是一个复向量空间, 有映射 $\langle \cdot, \cdot \rangle \mapsto \mathbb{C}$, 如果对于任意 $x, y, z \in E$ 和 $\alpha, \beta \in \mathbb{C}$ 满足下列条件

- (a) $\langle x, y \rangle = \langle y, x \rangle^*$
- (b) $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$
- (c) $\langle x, x \rangle \geq 0$
- (d) $\langle x, x \rangle = 0$ 当且仅当 $x = 0$

则把这个映射叫做内积. 具有内积的向量空间叫做内积空间

根据定义, 两个向量的内积是复数. 根据条件 a, $\langle x, x \rangle = \langle x, x \rangle^*$, 对于任意 $x \in E$, $\langle x, x \rangle$ 是实数. 根据 b 可以得到关系式

$$\langle \alpha x + \beta y, z \rangle = \langle z, \alpha x + \beta y \rangle^* = (\alpha \langle z, x \rangle + \beta \langle z, y \rangle)^* = \alpha^* \langle z, x \rangle^* + \beta^* \langle z, y \rangle^* = \alpha^* \langle x, z \rangle + \beta^* \langle y, z \rangle$$

特别的

$$\langle \alpha x, y \rangle = \alpha^* \langle x, y \rangle, \quad \langle x, \alpha y \rangle = \alpha \langle x, y \rangle$$

因此如果 $\alpha = 0$ 我们有 $\langle 0, y \rangle = \langle x, 0 \rangle = 0$

例 1.22 最简单最重要的内积空间例子是复数空间 \mathbb{C} . 其中内积的定义是 $\langle x, y \rangle = x^* y$

例 1.23 n 元复数数组 $x = (x_1, \dots, x_n)$ 的空间 \mathbb{C}^n , 构成内积空间, 其内积的定义是

$$\langle x, y \rangle = \sum_{k=1}^n x_k^* y_k, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$$

例 1.24 所有复数序列 (x_1, x_2, x_3, \dots) 构成的 l^2 空间 $(\sum_{k=1}^{\infty} |x_k|^2 < \infty)$ 是一个内积空间, 其内积定义是

$$\langle x, y \rangle = \sum_{k=1}^{\infty} x_k^* y_k, \quad x = (x_1, x_2, x_3, \dots), y = (y_1, y_2, y_3, \dots)$$

之后我们将会见到, 这个空间是内积空间里最重要的例子.

例 1.25 仅有限非零项的复数序列 (x_1, x_2, x_3, \dots) 构成的空间是一个内积空间, 内积的定义如1.24.

例 1.26 区间 $[a, b]$ 上的连续复函数的空间 $C([a, b])$ 是一个内积空间, 内积定义如下

$$\langle f, g \rangle = \int_a^b f^*(x)g(x)dx$$

例 1.27 $L^2(\mathbb{R})$ 是内积空间, 内积定义如下:

$$\int_{-\infty}^{+\infty} f^*(x)g(x)dx$$

进一步说, $L^2(\mathbb{R}^n)$ 也是内积空间, 其内积定义是

$$\int_{\mathbb{R}^n} f^*(x)g(x)dx$$

例 1.28 令 E 是内积空间 E_1 和 E_2 的Cartesian乘积, $E = E_1 \times E_2 = (x, y) : x \in E_1, y \in E_2$. 空间 E 是一个内积空间, 其内积定义如下:

$$\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle$$

值得注意 E_1 和 E_2 可以分别的看成子空间 $E \times 0$ 和 $0 \times E_2$

内积空间是具有内积的向量空间, 每一个内积空间也是赋范空间, 范数如下定义

$$\|x\| = \sqrt{\langle x, x \rangle}$$

首先注意函数是良定义的, 因为 $\langle x, x \rangle$ 总是非负的实数, 并且 $\|x\| = 0$ 当且仅当 $x = 0$. 进一步说

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^* \lambda \langle x, x \rangle} = |\lambda| \|x\|.$$

还剩下三角不等式有待证明, 这不像前两个条件这样好证明. 首先我们介绍一下Cauchy-Schwarz不等式.

定理 1.21 对于内积空间里的任意两个元素 x, y , 有

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (3)$$

当且仅当 x, y 线性相关时, 等式 $|\langle x, y \rangle| = \|x\| \|y\|$ 成立.

证明 1.15 若 $y = 0$, 则(3)刚好满足, 左右两边均是0. 假设 $y \neq 0$, 我们有

$$0 \leq \langle x + \alpha y, x + \alpha y \rangle = \langle x, x \rangle + \alpha \langle x, y \rangle + \alpha^* \langle y, x \rangle + |\alpha|^2 \langle y, y \rangle$$

在(1.15)中令 $\alpha = -\frac{\langle y, x \rangle}{\langle y, y \rangle}$. 两边同时乘以 $\langle y, y \rangle$.

$$0 \leq \langle x, x \rangle \langle y, y \rangle - |\langle x, y \rangle|^2$$

这就给出了Cauchy-Schwarz不等式. 如果 x, y 线性相关, 存在 $\alpha \in \mathbb{C}$ 有 $x = \alpha y$. 因此

$$|\langle x, y \rangle| = |\langle x, \alpha x \rangle| = |\alpha| \langle x, x \rangle = |\alpha| \|x\| \|x\| = \|x\| \|\alpha x\| = \|x\| \|\alpha y\|$$

令 x, y 是两个向量, 使得 $|\langle x, y \rangle| = \|x\| \|y\|$, 也就是说

$$\langle x, y \rangle \langle y, x \rangle = \langle x, x \rangle \langle y, y \rangle$$

接下来利用Cauchy-Schwarz不等式等号成立条件去证明 $\langle y, y \rangle x - \langle y, x \rangle y = 0$, 这表明了 x, y 线性相关.

$$\begin{aligned} & \langle \langle y, y \rangle x - \langle y, x \rangle y, \langle y, y \rangle x - \langle y, x \rangle y \rangle \\ &= |\langle y, y \rangle|^2 \langle x, x \rangle - \langle y, y \rangle \langle y, x \rangle \langle x, y \rangle - \langle y, x \rangle \langle y, y \rangle \langle y, x \rangle + |\langle x, y \rangle|^2 \langle y, y \rangle \\ &= 0 \end{aligned}$$

证毕.

推论 1.1 (三角不等式) 对于内积空间中的任意两个元素 x, y 有:

$$\|x + y\| \leq \|x\| + \|y\|$$

证明 1.16

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

其中 $\operatorname{Re} z$ 表示 $z \in \mathbb{C}$ 的实部.

内积空间的范数

定义 1.30 根据 $\|x\| = \sqrt{\langle x, x \rangle}$ 可以定义内积空间 E 的范数.

我们已经证明了每一个内积空间都是赋范空间. 很自然的我们会问到, 是否每一个赋范空间是内积空间呢? 更精确地表述是: 是否可能在赋范空间 $(E, \|\cdot\|)$ 中定义一个内积 $\langle \cdot, \cdot \rangle$ 使得对于每一个 $x \in E$, 有 $\|x\| = \sqrt{\langle x, x \rangle}$? 一般来说, 答案是否定的. 接下来我们将会给出内积空间范数的性质, 这些性质对于赋范空间称为内积空间是充分必要的.

定理 1.22 (平行四边形法则) 对于内积空间中任意两个元素, 有:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

证明 1.17 我们有

$$\|x+y\|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$$

因此

$$\|x+y\|^2 = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2$$

现在令 $y \rightarrow -y$, 上式得到

$$\|x-y\|^2 = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2$$

两式叠加可得平行四边形法则

内积空间里最重要的是可以定义正交向量. 这使得 Hilbert 空间与 Banach 空间非常的不同.

定义 1.31 (正交向量) 内积空间里有两个向量 x, y , 如果 $\langle x, y \rangle = 0$, 则称这两个向量正交, 记作 $x \perp y$

如果 $x \perp y$, 则有 $\langle y, x \rangle = \langle x, y \rangle^* = 0$, 因此 $y \perp x$. 换句话说, \perp 关系式对称的.

定理 1.23 (勾股定理) 对于任意一对正交向量 x, y , 我们有

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

证明 1.18 如果 $x \perp y$, 则 $\langle x, y \rangle = 0$, 根据平行四边形法则, 立即可以得到上式.

在之前我们定义的内积空间都是在 \mathbb{C} 中定义的. 我们也可以在 \mathbb{R} 上定义内积空间, 也就是说任意两个向量的内积是一个实数. 内积的定义1.29中的**b**要修改成 $\langle x, y \rangle = \langle y, x \rangle$, 而之前所有的定理和例子都要修改成 \mathbb{R} 上成立. 有限维实内积空间也叫做Euclidean空间

如果 \mathbb{R}^n 上有 $x = (x_1, x_2, \dots), y = (y_1, y_2, \dots)$, 则内积的定义 $\langle x, y \rangle = \sum_{k=1}^n x_k y_k$ 等价于定义 $\langle x, y \rangle = \|x\| \|y\| \cos \theta$, 其中 θ 是向量 x, y 的夹角. 在这种情况下, Cauchy-Schwarz不等式成为

$$\frac{|\langle x, y \rangle|}{\|x\| \|y\|} = \cos \theta \leq 1, \quad x \neq 0, y \neq 0$$

希尔伯特空间

定义 1.32 完备内积空间叫做Hilbert空间

这里 E 的完备性是指用 E 的内积定义的范数的完备性. 我们现在给出一些内积空间和Hilbert空间的例子

例 1.29 由于 \mathbb{C} 与 \mathbb{C}^n 是完备的, 它们都是Hilbert空间.

例 1.30 l^2 是Hilbert空间, 其完备性在之前例子1.15已经证明过了, 这里不再复述.

例 1.31 例子1.25里的空间 E 不是Hilbert空间, 考虑柯西序列序列

$$x_n = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, 0, 0, \dots)$$

序列满足

$$\lim_{m, n \rightarrow \infty} \|x_m - x_n\| = \lim_{n, m \rightarrow \infty} \left[\sum_{k=\min\{m, n\}+1}^{\max\{m, n\}} \frac{1}{k^2} \right]^{\frac{1}{2}} = 0$$

然而这个序列并不收敛于 E 中. 因为极限 $(1, \frac{1}{2}, \frac{1}{3}, \dots)$ 不在 E 中.

例 1.32 例子 1.26 中的空间是另一个不完备的内积空间例子. 事实上, 考虑 $C([0, 1])$ 中的如下函数序列:

$$f_n(x) = \begin{cases} 1 & \text{如果 } 0 \leq x \leq \frac{1}{2} \\ 1 - 2n(x - \frac{1}{2}) & \text{如果 } \frac{1}{2} \leq x \leq \frac{1}{2n} + \frac{1}{2} \\ 0 & \text{如果 } \frac{1}{2n} + \frac{1}{2} \leq x \leq 1 \end{cases}$$

显然 f_n 是连续的. 进一步可知

$$\|f_n - f_m\| \leq \left(\frac{1}{n} + \frac{1}{m}\right)^{\frac{1}{2}} \rightarrow 0, \quad \text{当 } m, n \rightarrow \infty$$

因此 (f_n) 是柯西序列. 显然这个序列是逐点收敛于函数

$$f(x) = \begin{cases} 1 & \text{如果 } 0 \leq x \leq \frac{1}{2} \\ 0 & \text{如果 } \frac{1}{2} \leq x \leq 1 \end{cases}$$

函数极限不连续, 因此它不属于 $C([0, 1])$, 序列 (f_n) 不收敛于 $C([0, 1])$. $C([0, 1])$ 不是 Hilbert 空间.

例 1.33 $L^2(R)$ 和 $L^2([a, b])$ 是 Hilbert 空间

例 1.34 ρ 是区间 $[a, b]$ 上的可测函数且在 $[a, b]$ 上 $\rho(x) > 0$. 记 $L^{2,\rho}([a, b])$ 为 $[a, b]$ 上复可测函数构成的空间

$$\int_a^b |f(x)|^2 \rho(x) dx < \infty$$

这是定义了内积

$$\langle f, g \rangle = \int_a^b f^*(x) g(x) \rho(x) dx$$

的一个 Hilbert 空间. 为了证明完备性, 考虑 $L^{2,\rho}([a, b])$ 上的柯西序列 (f_n) .

$$\|f_m - f_n\|_{L^{2,\rho}([a, b])}^2 = \int_a^b |f_m(x) - f_n(x)|^2 \rho(x) dx \rightarrow 0 \quad \text{当 } n \rightarrow \infty$$

定义

$$F_n = f_n \sqrt{\rho}, \quad n \in \mathbb{N}$$

因为

$$\begin{aligned}
\|F_m - F_n\|_{L^2([a,b])}^2 &= \int_a^b |F_m(x) - F_n(x)|^2 dx \\
&= \int_a^b |f_m(x)\sqrt{\rho(x)} - f_n(x)\sqrt{\rho(x)}|^2 dx \\
&= \int_a^b |f_m(x) - f_n(x)|^2 \rho(x) dx \\
&= \|f_m - f_n\|_{L^{2,\rho}([a,b])}^2
\end{aligned}$$

(F_n) 是 $L^2([a,b])$ 中的柯西序列. 因此存在 $F \in L^2([a,b])$ 使得

$$\|F_n - F_m\|_{L^2([a,b])}^2 = \int_a^b |F_n(x) - F_m(x)|^2 dx \rightarrow 0$$

我们很容易看到 $f_n \rightarrow \frac{F}{\sqrt{\rho}} \in L^{2,\rho}([a,b])$, 证明完毕

例 1.35 (Sobolev 空间) Ω 是 \mathbb{R}^n 的一个开集. 所有复函数 $f \in C^m(\Omega)$ 构成的空间记作 $\tilde{H}^m(\Omega)$, $m = 1, 2, \dots$, 使得对于所有的 $|\alpha| \leq m$ 有 $D^\alpha f \in L^2(\Omega)$ 成立, 其中 $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_1, \dots, \alpha_n$ 是非负整数, $|\alpha| = \alpha_1 + \dots + \alpha_n$, 以及

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$$

例如, 如果 $n = 2, \alpha = (2, 1)$, 我们有

$$D^\alpha f = \frac{\partial^3 f}{\partial x_1^2 \partial x_2}$$

对于 $f \in \tilde{H}^m(\Omega)$, 对于每一个多元指标 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 我们有

$$\int_{\Omega} \left| \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} \right|^2 < \infty$$

使得 $|\alpha| \leq m$. $\tilde{H}^m(\Omega)$ 中的内积如下定义:

$$\langle f, g \rangle = \int_{\Omega} \sum_{|\alpha| \leq m} (D^\alpha f)^* D^\alpha g$$

如果 $\Omega \in \mathbb{R}^2$, $\tilde{H}^2(\Omega)$ 的内积就是

$$\langle f, g \rangle = \int_{\Omega} (f^*g + f_x^*g_x + f_y^*g_y + f_{xx}^*g_{xx} + f_{yy}^*g_{yy} + f_{xy}^*g_{xy})$$

如果 $\Omega = (a, b) \subset \mathbb{R}$, $\tilde{H}^m(a, b)$ 的内积就是

$$\langle f, g \rangle = \int_a^b \sum_{n=0}^m \left(\frac{d^n f}{dx^n} \right)^* \frac{d^n g}{dx^n}$$

$\tilde{H}^m(\Omega)$ 是一个内积空间, 但不是 Hilbert 空间, 因为它不完备. 完备的 $\tilde{H}^m(\Omega)$ 记作 $H^m(\Omega)$, 它是希尔伯特空间. $H^m(\Omega)$ 是更一般空间 $W^{m,p}(\Omega)$ (Sobolev 空间) 的特殊情况. 我们有 $H^m(\Omega) = W^{m,2}(\Omega)$. 由于偏微分方程的广泛应用, $H^m(\Omega)$ 是最重要的一个 Hilbert 空间.

因为每一个内积空间都有范数, 所以它也有由范数定义的收敛. 这个收敛叫做强收敛

强收敛

定义 1.33 内积空间 E 中的序列 (x_n) , 如果当 $n \rightarrow \infty$ 时有 $\|x_n - x\| \rightarrow 0$, 则称序列 (x_n) 强收敛于 E 中的向量 x .

强收敛这个词是为了区别于弱收敛.

弱收敛

定义 1.34 内积空间 E 中有序列 (x_n) , 如果当 $n \rightarrow \infty$ 时对于任意 $y \in E$ 有 $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$, 则称序列 (x_n) 弱收敛于 E 中的向量 x

上述定义可以写成当 $n \rightarrow \infty$ 时, 对于任意 $y \in E$ 有 $\langle x_n - x, y \rangle \rightarrow 0$. 我们通常用 $x_n \rightarrow x$ 表示强收敛, $x_n \xrightarrow{w} x$ 表示弱收敛

定理 1.24 强收敛必定弱收敛, 也就是说, $x_n \rightarrow x$ 意味着 $x_n \xrightarrow{w} x$

证明 1.19 设序列 (x_n) 强收敛于 x . 也就是说当 $n \rightarrow \infty$ 时, 有 $\|x_n - x\| \rightarrow 0$. 根据 Cauchy-Schwarz 不等式

$$|\langle x_n - x, y \rangle| \leq \|x_n - x\| \|y\| \rightarrow 0 \text{ 当 } n \rightarrow \infty$$

这样, 当 $n \rightarrow \infty$ 时, 对于任意 $y \in E$, 有 $\langle x_n - x, y \rangle \rightarrow 0$.

对于内积空间 E 中任意的固定点 y , 映射 $\langle \cdot, y \rangle : E \rightarrow \mathbb{C}$ 是 E 上的一个线性泛函. 上述定理告诉了我们对于任意 $y \in E$, 这个线性泛函连续. 同样的线性泛函 $\langle x, \cdot \rangle \mapsto \mathbb{C}$ 也连续.

定理 1.25 如果 $x_n \rightarrow x, y_n \rightarrow y$, 则 $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$

证明 1.20 若 $x_n \rightarrow x, y_n \rightarrow y$, 则有

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &\leq |\langle x_n, y_n \rangle - \langle x, y_n \rangle| + |\langle x, y_n \rangle - \langle x, y \rangle| \\ &= |\langle x_n - x, y_n \rangle| + |\langle x, y_n - y \rangle| \\ &\leq \|x_n - x\| \|y_n\| + \|x\| \|y_n - y\| \rightarrow 0 \end{aligned}$$

其中序列 (y_n) 有界.

根据这个定理, 我们立马可以获得强收敛的一个性质:

$$x_n \rightarrow x \text{ 意味着 } \|x_n\| \rightarrow \|x\|$$

定理 1.26 若 $x_n \xrightarrow{w} x$ 且 $\|x_n\| \rightarrow \|x\|$, 则 $x_n \rightarrow x$.

证明 1.21 若对于所有的 y , 有 $x_n \xrightarrow{w} x$, 我们有:

$$\langle x_n, x \rangle \rightarrow \langle x, x \rangle, \quad n \rightarrow \infty$$

因此

$$\langle x_n, x \rangle \rightarrow \langle x, x \rangle = \|x\|^2$$

现在, 当 $n \rightarrow \infty$ 时

$$\begin{aligned} \|x_n - x\|^2 &= \langle x_n - x, x_n - x \rangle \\ &= \langle x_n, x_n \rangle - \langle x_n, x \rangle - \langle x, x_n \rangle + \langle x, x \rangle \\ &= \|x_n\|^2 - 2\operatorname{Re}\langle x_n, x \rangle + \|x\|^2 \rightarrow \|x\|^2 - 2\|x\|^2 + \|x\|^2 = 0 \end{aligned}$$

因此序列 (x_n) 强收敛于 x

1.9 正交归一系统

向量空间中可以给定一族基 \mathcal{B} , 使得任意向量 $x \in E$ 能被写成 $\sum_{n=1}^m \lambda_n x_n$, 其中 $x_n \in \mathcal{B}$, λ_n 是标量. 在内积空间里正交基非常重要. 无限求和可以取代有限求和 $\sum_{n=1}^m \lambda_n x_n$, 并且线性无关的条件可以替换成正交条件.

正交和正交归一系统

定义 1.35 内积空间 E 中的一族非零向量族 S , 如果对于 S 中任何不同的两个元素 x, y 都有 $x \perp y$, 则叫做正交系统. 除此之外, 如果对于任意 $x \in S$ 都有 $\|x\| = 1$, 则叫 S 正交归一系统.

每一个正交非零向量的集合都可以归一化: 如果 S 是一个正交系统, 则 $S = \frac{x}{\|x\|} : x \in S$ 是一个正交归一系统.

指的注意如果 x 于每一个 y_1, y_2, \dots, y_n 都正交, 则 x 与 y_1, y_2, \dots, y_n 的任意线性组合都正交. 事实上如果 $y = \sum_{k=1}^n \lambda_k y_k$, 我们有:

$$\langle x, y \rangle = \langle x, \sum_{k=1}^n \lambda_k y_k \rangle = \sum_{k=1}^n \lambda_k \langle x, y_k \rangle = 0$$

定理 1.27 正交系统是线性无关的系统

证明 1.22 令 S 是一个正交系统. 假设存在 $x_1, x_2, \dots, x_n \in S$ 和 $\alpha_1, \alpha_2, \dots, \alpha_n$ 有 $\sum_{k=1}^n \alpha_k x_k = 0$, 则

$$0 = \sum_{m=1}^n \langle 0, \alpha_m x_m \rangle = \sum_{m=1}^n \langle \sum_{k=1}^n \alpha_k x_k, \alpha_m x_m \rangle = \sum_{m=1}^n |\alpha_m|^2 \|x_m\|^2$$

这表明了对于任意 $m \in \mathbb{N}$ 都有 $\alpha_m = 0$, 因此 x_1, \dots, x_n 是线性无关的.

正交序列

定义 1.36 组成正交归一系统的向量序列叫做正交归一序列.

在实际中, 我们经常用全体整数 \mathbb{Z} 来表示序列指标. 正交归一序列 (x_n) 的条件可以用Kronecker符号去表示.

$$\langle x_m, x_n \rangle = \delta_{mn} = \begin{cases} 0 & \text{如果 } m \neq n \\ 1 & \text{如果 } m = n \end{cases}$$

例 1.36 对于向量 $e_n = (0, 0, \dots, 1, 0, \dots, \dots)$, 其中1在第 n 个位置. 集合 $S = e_1, e_2, \dots, e_n, \dots$ 构成 l^2 中的正交归一系统.

例 1.37 令 $\varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}, n \in \mathbb{Z}$. 集合 $\varphi_n : n \in \mathbb{Z}$ 是 $L^2([-\pi, \pi])$ 中的一个正交归一系统. 的确, 对于 $m \neq n$, 我们有

$$\langle \varphi_n, \varphi_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x} dx = \frac{e^{i\pi(m-n)} - e^{-i\pi(m-n)}}{2\pi i(m-n)} = 0$$

另一方面

$$\langle \varphi_n, \varphi_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-n)x} dx = 1$$

因此 $\langle \varphi_m, \varphi_n \rangle = \delta_{mn}, m, n \in \mathbb{Z}$.

例 1.38 Legendre 多项式如下定义

$$P_0(x) = 1$$

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n = 1, 2, 3, \dots$$

构成 $L^2([-1, 1])$ 上的正交系统. 为了方便, 我们记 $(x^2 - 1)^n = p_n(x)$, 这样

$$\int_{-1}^1 P_n(x) x^m dx = \frac{1}{2^n n!} \int_{-1}^1 p_n^{(n)}(x) x^m dx$$

对于 $m < n$, 我们通过递推计算这个积分. 首先我们注意到

$$\int_{-1}^1 p_n^{(n)}(x) x^m dx = -m \int_{-1}^1 p_n^{(n-1)}(x) x^{m-1} dx$$

重复这个操作直到最后推导出

$$(-1)^m m! \int_{-1}^1 p_n^{n-m}(x) dx = (-1)^m m! [p_n^{(n-m-1)}(x)]|_{-1}^1 = 0$$

所以

$$\int_{-1}^1 P_n(x) x^m dx = 0, \quad m < n$$

因为 $P_m(x)$ 是 m 阶多项式, 这显然可以导出

$$\langle P_n, P_m \rangle = \int_{-1}^1 P_n(x) P_m(x) dx = 0, \quad m \neq n$$

这样我们证明了 Legendre 多项式的正交性. 为了获得 Legendre 多项式的正交归一系统. 我

们还需要计算 $L^2([-1, 1])$ 中 P_n 的范数:

$$\|P_n\| = \sqrt{\int_{-1}^1 (P_n(x))^2 dx}$$

首先我们得到

$$\begin{aligned} \int_{-1}^1 (1-x^2)^n dx &= \int_{-1}^1 (1-x)^n (1+x)^n dx \\ &= \frac{n}{n+1} \int_{-1}^1 (1-x)^{n-1} (1+x)^{n+1} dx = \dots \\ &= \frac{n(n-1)\dots 2 \cdot 1}{(n+1)(n+2)\dots 2n} \int_{-1}^1 (1+x)^{2n} dx \\ &= \frac{(n!)^2 2^{2n+1}}{(2n)!(2n+1)} \end{aligned}$$

类似的过程

$$\begin{aligned} \int_{-1}^1 (p_n^{(n)}(x))^2 dx &= 0 - \int_{-1}^1 p_n^{(n-1)}(x) p_n^{(n+1)}(x) dx \\ &= \dots \\ &= (-1)^n \int_{-1}^1 p_n(x) p_n^{(2n)}(x) dx \\ &= (2n)! \int_{-1}^1 (1-x)^n (1+x)^n dx \end{aligned}$$

其中我们有 $p_n(x)$ 的 $2n$ 次导数, 这样就只剩下 $2n$ 项. 根据以上各个式子可以得到:

$$\int_{-1}^1 (P_n(x))^2 dx = \frac{1}{(2^n n!)^2} (2n)! \frac{(n!)^2 2^{2n+1}}{(2n)!(2n+1)} = \frac{2}{2n+1}$$

这样多项式 $\sqrt{n + \frac{1}{2}} P_n(x)$ 构成 $L^2([-1, 1])$ 中的正交归一系统.

例 1.39 我们用 H_n 表示 n 阶Hermite多项式.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$

函数 $\varphi_n(x) = e^{-\frac{x^2}{2}} H_n(x)$ 构成 $L^2(\mathbb{R})$ 上的正交系统. 内积为

$$\langle \varphi_n, \varphi_m \rangle = (-1)^{n+m} \int_{-\infty}^{\infty} e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \frac{d^m}{dx^m} e^{-x^2} dx$$

两边乘以 $(-1)^{n+m}$

$$(-1)^{n+m} \langle \varphi_n, \varphi_m \rangle = \left[e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \frac{d^m}{dx^m} e^{-x^2} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d}{dx} \left[e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \right] \frac{d^{m-1}}{dx^{m-1}} e^{-x^2} dx \quad (4)$$

在微分符号里的所有项都包含因子 e^{-x^2} . 因此对于所有的 $k \in \mathbb{N}$ 我们有

$$x^k e^{-x^2} \rightarrow 0 \text{ 当 } x \rightarrow \infty$$

所以(4)第一项收敛. 不断重复分部积分最后就能得到

$$\langle \varphi_n, \varphi_m \rangle = 0, n \neq m$$

为了获得正交归一系统, 我们计算范数

$$\|\varphi_n\|^2 = \int_{-\infty}^{\infty} e^{-x^2} (H_n(x))^2 dx = \int_{-\infty}^{\infty} e^{-x^2} \left[e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \right]^2 dx$$

因为 $H_n(x)$ 是 n 次多项式, 直接求导得到

$$e^{x^2} \frac{d^n}{dx^n} e^{-x^2} = (-2x)^n + \dots$$

然后在求导

$$\frac{d^n}{dx^n} \left[e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \right] = \frac{d^n}{dx^n} ((-2x)^n + \dots) = (-1)^n 2^n n!$$

所以得到

$$\|\varphi_n\|^2 = 2^n n! \int_{-\infty}^{\infty} e^{-x^2} dx = 2^n n! \sqrt{\pi}$$

因此函数

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} e^{-\frac{x^2}{2}} H_n(x)$$

构成 $L^2(\mathbb{R})$ 中的正交归一系统.

在之前的例子里, 正交函数序列是正交的但不是归一的. 尽管计算比较复杂, 但我们总是

可以对函数归一化获得正交归一的函数序列. 同样的对于线性无关的函数序列, 我们也可以通过Schmidt正交化的方案获得正交归一序列. 过程如下

给定内积空间中线性无关的序列 (y_n) , 定义序列 $(w_n), (x_n)$

$$\begin{aligned} w_1 &= y_1 & x_1 &= \frac{w_1}{\|w_1\|} \\ w_k &= y_k - \sum_{n=1}^k \langle x_n, y_k \rangle x_n, & x_k &= \frac{w_k}{\|w_k\|} \end{aligned}$$

序列 (w_n) 是正交的. 我们注意到

$$\begin{aligned} \langle w_1, w_2 \rangle &= \langle y_2 - \langle x_1, y_2 \rangle x_1, y_1 \rangle = \langle y_2, y_1 \rangle - \langle y_2, x_1 \rangle \langle x_1, y_1 \rangle \\ &= \langle y_2, y_1 \rangle - \frac{\langle y_2, y_1 \rangle \langle y_1, y_1 \rangle}{\|y_1\|^2} = 0 \end{aligned}$$

假设 w_2, \dots, w_{k-1} 都是正交的, 对于任意 $m < k$ 都有

$$\begin{aligned} \langle w_k, w_m \rangle &= \langle y_k, w_m \rangle - \frac{\sum_{n=1}^{k-1} \langle y_k, w_n \rangle \langle w_n, w_m \rangle}{\|w_m\|^2} \\ &= \langle y_k, w_m \rangle - \frac{\langle y_k, w_m \rangle \langle w_m, w_m \rangle}{\|w_m\|^2} = 0 \end{aligned}$$

因此向量 w_1, w_2, \dots, w_k 是正交的. 根据数学归纳法, 序列 (w_n) 是正交序列, 因此 (x_n) 是正交归一序列. 很容易检验向量 x_1, x_2, \dots, x_n 的线性组合也是 y_1, \dots, y_n 的线性组合, 反之亦然. 换句话说对于任意 $n \in \mathbb{N}$ 有 $\text{span} x_1, \dots, x_n = \text{span} y_1, \dots, y_n$.

在上一章节我们证明了在内积空间中任意正交向量都满足勾股定理, 显然这可以推广到任意有限个正交向量.

定理 1.28 (勾股定理) 如果 x_1, x_2, \dots, x_n 是内积空间中的正交向量, 则有

$$\left\| \sum_{k=1}^n x_k \right\|^2 = \sum_{k=1}^n \|x_k\|^2$$

证明 1.23 如果 $x_1 \perp x_2$ 则 $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2$. 显然这个定理对于 $n = 2$ 成立. 现在假设(1.28)对于 $n - 1$ 都成立, 也就是说

$$\left\| \sum_{k=1}^{n-1} x_k \right\|^2 = \sum_{k=1}^{n-1} \|x_k\|^2$$

令 $x = \sum_{k=1}^{n-1} x_k$, 并因为 $x \perp y$, 我们有

$$\left\| \sum_{k=1}^n x_k \right\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2 = \sum_{k=1}^{n-1} \|x_k\|^2 + \|x_n\|^2 = \sum_{k=1}^n \|x_k\|^2$$

得证

勾股定理将会帮我们证明正交归一集合更重要的性质.

定理 1.29 (Bessel等式和不等式) 令 x_1, \dots, x_n 是内积空间 E 中的正交归一集合的向量. 对于每一个 $x \in E$, 我们有

$$\left\| x - \sum_{k=1}^n \langle x_k, x \rangle x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x_k, x \rangle|^2 \quad (5)$$

并且有

$$\sum_{k=1}^n |\langle x_k, x \rangle|^2 \leq \|x\|^2$$

证明 1.24 由于有勾股定理(1.28), 我们可以得到

$$\left\| \sum_{k=1}^n \alpha_k x_k \right\|^2 = \sum_{k=1}^n \|\alpha_k x_k\|^2 = \sum_{k=1}^n |\alpha_k|^2$$

对于任意复数 $\alpha_1, \dots, \alpha_n$ 都成立. 因此

$$\begin{aligned} \left\| x - \sum_{k=1}^n \alpha_k x_k \right\|^2 &= \left\langle x - \sum_{k=1}^n \alpha_k x_k, x - \sum_{k=1}^n \alpha_k x_k \right\rangle \\ &= \|x\|^2 - \left\langle x, \sum_{k=1}^n \alpha_k x_k \right\rangle - \left\langle \sum_{k=1}^n \alpha_k x_k, x \right\rangle + \sum_{k=1}^n |\alpha_k|^2 \|x_k\|^2 \\ &= \|x\|^2 - \sum_{k=1}^n \alpha_k \langle x_k, x \rangle^* - \sum_{k=1}^n \alpha_k^* \langle x_k, x \rangle \\ &= \|x\|^2 - \sum_{k=1}^n |\langle x_k, x \rangle|^2 + \sum_{k=1}^n |\langle x_k, x \rangle - \alpha_k|^2 \end{aligned}$$

若 $\alpha_k = \langle x_k, x \rangle$, 这将得到(5). 根据(5)由可以得到

$$0 \leq \|x\|^2 - \sum_{k=1}^n |\langle x_k, x \rangle|^2$$

Q.E.D.

若 (x_n) 是正交归一序列, 则令 $n \rightarrow \infty$ 我们可以得到

$$\sum_{k=1}^{\infty} |\langle x_k, x \rangle|^2 \leq \|x\|^2 \quad (6)$$

所以

$$\lim_{n \rightarrow \infty} \langle x_n, x \rangle = 0$$

因此正交归一序列是弱收敛于0的. 另一方面来说, 因为对于所有的 $n \in \mathbb{N}$ 都有 $\|x\| = 1$, 正交归一序列不是强收敛的.

方程6暗含了这样的一个关系, 对于任意 $x \in E$ 都有 $\sum_{n=1}^{\infty} |\langle x_n, x \rangle|^2$ 收敛. 换句话说, 序列 $(\langle x_n, x \rangle)$ 是 l^2 中的元素. 我们可以说 E 中的正交归一序列诱导 E 到 l^2 的映射. 展开式:

$$x \sim \sum_{n=1}^{\infty} \langle x_n, x \rangle x_n \quad (7)$$

叫做 x 的广义傅里叶级数. 标量 $\langle x_n, x \rangle$ 叫做 x 对于正交归一序列 (x_n) 的广义傅里叶系数.

定理 1.30 令 (x_n) 是Hilbert空间中得正交归一序列, (α_n) 是一个复序列. 则当且仅当满足 $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$ 且

$$\left\| \sum_{n=1}^{\infty} \alpha_n x_n \right\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2 \quad (8)$$

级数 $\sum_{n=1}^{\infty} \alpha_n x_n$ 收敛.

证明 1.25 对于任意 $m > k > 0$, 利用勾股定理我们可以得到,

$$\left\| \sum_{n=k}^m \alpha_n x_n \right\|^2 = \sum_{n=k}^m |\alpha_n|^2 \quad (9)$$

若 $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$, 则根据方程(9)可以得到序列 $s_m = \sum_{n=1}^m \alpha_n x_n$ 是一个柯西序列. 这意味着由于 H 的完备性序列 $\sum_{n=1}^{\infty} \alpha_n x_n$ 是收敛的. 所以, 若级数 $\sum_{n=1}^{\infty} \alpha_n x_n$ 收敛, 则方程(9)意味着 $\sum_{n=1}^{\infty} |\alpha_n|^2$ 的收敛. 因为序列 $\sigma_m = \sum_{n=1}^m |\alpha_n|^2$ 是 \mathbb{R} 的柯西序列. 为了获得方程(8)只需要令 $k = 1, m \rightarrow \infty$ 即可得.

这个定理和等式(6)意味着在Hilbert空间 H 中, 对于任意 $x \in H$ 都有级数 $\sum_{n=1}^{\infty} \langle x_n, x \rangle x_n$ 收敛

例 1.40 令 $H = L^2([-\pi, \pi])$, $x_n(t) = \frac{1}{\sqrt{\pi}} \sin nt$, $n = 1, 2, \dots$. 序列 (x_n) 是 H 中的正交归一序列. 换句话说, 对于 $x(t) = \cos t$, 我们有

$$\begin{aligned} \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n(t) &= \sum_{n=1}^{\infty} \left[\frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} \cos t \sin nt dt \right] \frac{\sin nt}{\sqrt{\pi}} \\ &= \sum_{n=1}^{\infty} 0 \cdot \sin nt = 0 \neq \cos t \end{aligned}$$

定义 1.37 (完备正交归一序列) 若内积空间 E 中, 对于任意 $x \in E$ 有

$$x = \sum_{n=1}^{\infty} \langle x_n, x \rangle x_n \quad (10)$$

则正交归一序列 (x_n) 被叫做完备序列

方程(10)的右边是无穷级数, 这意味着

$$\lim_{n \rightarrow \infty} \|x - \sum_{k=1}^n \langle x_k, x \rangle x_k\| = 0$$

其中 $\|\cdot\|$ 是 E 中的范数. 例如 $E = ([-\pi, \pi])$ 并且 (f_n) 是 E 中的正交归一序列, 则有

$$f = \langle f_m, f \rangle f_n$$

同时也有

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \left| f(t) - \sum_{k=1}^n \alpha_k f_k(t) \right|^2 dt = 0, \quad \text{其中 } \alpha_k = \int_{-\pi}^{\pi} f_k^*(t) f(t) dt$$

一般来说, 这并不意味逐点收敛, 所以我们不能说 $f(x) = \sum_{n=1}^{\infty} \alpha_n f_n(x)$

定义 1.38 (正交归一基) 内积空间 E 中有一个正交归一系统 B , 若任意 $x \in E$ 有唯一的表示

$$x = \sum_{n=1}^{\infty} \alpha_n x_n$$

则该系统叫做一组正交归一基, 其中 $\alpha_n \in \mathbb{C}$.

值得注意的是, 完备正交归一序列 (x_n) 是正交基, 并且是唯一的.

$$x = \sum_{n=1}^{\infty} \alpha_n x_n, \quad x = \sum_{n=1}^{\infty} \beta_n x_n$$

则有

$$0 = \|x - x\|^2 = \left\| \sum_{n=1}^{\infty} \alpha_n x_n - \sum_{n=1}^{\infty} \beta_n x_n \right\|^2 = \left\| \sum_{n=1}^{\infty} (\alpha_n - \beta_n) x_n \right\|^2 = \sum_{n=1}^{\infty} |\alpha_n - \beta_n|^2$$

这意味着对于所有 $n \in \mathbb{N}$ 都有 $\alpha_n = \beta_n$. 唯一性得证.

定理 1.31 Hilbert空间 H 中有正交归一序列 (x_n) 是完备的条件为, 对于所有的 $n \in \mathbb{N}$ 有 $\langle x_n, x \rangle = 0$ 成立当且仅当 $x = 0$.

证明 1.26 假设 (x_n) 是 H 中的完备正交归一序列, 则对于任意 $x \in H$ 有

$$x = \sum_{n=1}^{\infty} \langle x_n, x \rangle x_n$$

这样, 若对于所有的 $n \in \mathbb{N}$ 有 $\langle x_n, x \rangle = 0$, 则 $x = 0$.

反过来, 假设对于所有的 $n \in \mathbb{N}$ 都有 $\langle x_n, x \rangle = 0$, 这意味着 $x = 0$. 令 x 是 H 中的元素. 定义

$$y = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$$

根据式子(6)和定理1.30, y 在 H 中存在. 因此对于任意 $n \in \mathbb{N}$

$$\begin{aligned} \langle x_n, x - y \rangle &= \langle x_n, x \rangle - \left\langle x_n, \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k \right\rangle \\ &= \langle x_n, x \rangle - \sum_{k=1}^{\infty} \langle x_n, x_k \rangle \langle x_k, x \rangle \\ &= \langle x_n, x \rangle - \langle x_n, x \rangle = 0 \end{aligned}$$

我们有 $x - y = 0$, 因此 $x = \sum_{n=1}^{\infty} \langle x_n, x \rangle x_n$

这个定理的一个推论叫做Parseval公式.

定理 1.32 (Parseval公式) Hilbert空间 H 中的正交归一序列 (x_n) 是完备的, 当且仅当

对于任意 $x \in H$

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \quad (11)$$

证明 1.27 令 $x \in H$, 由方程(5)对于任意 $n \in \mathbb{N}$ 我们有

$$\left\| x - \sum_{k=1}^n \langle x_k, x \rangle x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x_k, x \rangle|^2$$

若 (x_n) 是完备序列, 则上式左边在 $n \rightarrow \infty$ 收敛, 因此

$$\lim_{n \rightarrow \infty} \left[\|x\|^2 - \sum_{k=1}^n |\langle x_k, x \rangle|^2 \right] = 0$$

(11) 成立.

相反地, 若(11)成立, 则(11)右边表达式在 $n \rightarrow \infty$ 收敛, 这样

$$\lim_{n \rightarrow \infty} \left\| x - \sum_{k=1}^n \langle x_k, x \rangle x_k \right\|^2 = 0$$

这就证明了序列 (x_n) 是完备的.

例 1.41 正交系

$$\varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}, \quad n = 0, \pm 1, \pm 2, \dots,$$

在 $L^2([-\pi, \pi])$ 中是完备的, 这个系统的正交性并不是那么容易证明, 证明将会用到我们后面要讲的 δ 函数.

例 1.42 序列函数

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos 2x, \frac{1}{\sqrt{\pi}} \sin 2x, \dots$$

是 $L^2([-\pi, \pi])$ 中的完备正交归一系统.

下列两个序列函数分别构成 $L^2([0, \pi])$ 的正交归一系统.

$$\begin{aligned} & \frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}} \cos x, \sqrt{\frac{2}{\pi}} \cos 2x, \sqrt{\frac{2}{\pi}} \cos 3x, \dots \\ & \sqrt{\frac{2}{\pi}} \sin x, \sqrt{\frac{2}{\pi}} \sin 2x, \sqrt{\frac{2}{\pi}} \sin 3x \end{aligned}$$

2 基本物理图像

自牛顿力学以来, 经典力学不断发展, 并广泛应用于动力学系统, 包括与物质相互作用的电磁场. 基本思想与支配它们应用的规律形成了一个简单而优美的方案, 人们不禁认为, 如果对这种方案作重大修改, 必然破坏其所有吸引人的特点. 尽管如此, 现在发展了一种新的方案, 称为量子力学, 它不仅更适合描述原子尺度的现象, 而且在某些方面, 它比经典方案更优美, 更令人满意. 这是由于新方案所包含的改变具有十分深刻的特征, 而且这一改变与那些使得经典力学如此吸引人的特点并不冲突, 因而新方案能够兼容经典理论学的所有这些特点. 在这一节, 我们对于量子力学基本架构将会做一个简短的描述. 量子力学里详细的数学讨论将会在后面章节里陆续介绍. 首先我们很有必要回顾一下旧的经典力学.

2.1 经典力学

2.1.1 基本力学概念

对于力学系统, 设有 N 个质点, 通常用位矢 $\mathbf{r}_a (a = 1, 2, \dots, N)$ 表示其中各个质点的位置, 也就是所谓的系统位形, 系统位形一共有 $3N$ 个坐标. 但是实际系统会受到各种约束的限制, $3N$ 个坐标并不都是独立的, 其中有可能存在各种各样的约束条件.

定义 2.1 (广义坐标) 能唯一确定系统位形所需要的独立变量数目叫做系统的自由度(设为 s), 所对应的独立变量称为广义坐标, 记作 q_1, q_2, \dots, q_s , 其导数 \dot{q}_i 和 $\ddot{q}_i (i = 1, 2, \dots, s)$ 分别称为广义速度和广义加速度. 广义坐标不一定是Cartesian坐标系, 广义坐标与Cartesian坐标之间的变换关系称为坐标变换.

$$\mathbf{r}_a = \mathbf{r}_a(q_1, q_2, \dots, q_s, t), \quad (a = 1, 2, \dots, N)$$

系统的状态由广义坐标和广义速度确定, 原则上它们可以确定系统任意时刻的位形和位形变化. 值得注意的是如果系统具有一定的对称性, 选取与对称变换相关联的参量作为广义坐标可以为运动的求解带来一定的便利, 因为这种坐标将不会出现在系统的Lagrangian中, 这样的坐标称为循环坐标, 将会存在对应的运动积分

在力学中处理问题时, 第一步要做的就是明确系统, 这是讨论所有问题的基本出发点. 系统各质点或者质量微元的位置集合称为系统的位形.

定义 2.2 (约束) 约束是对系统位形和位形变化的一种限制.

按照约束是否对位形以及位形变化有所约束可以分为完整约束(几何约束)和非完整约束两类. 前者仅仅对位形有限制, 约束表达式通常可以写成

$$f_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = 0, \quad (i = 1, 2, \dots, l)$$

其中 l 为完整约束个数. 非完整约束对位形以及位形变化均有限制, 约束表达式通常可以写成

$$f_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, \dots, \dot{\mathbf{r}}_N, t) = 0, \quad (i = 1, 2, \dots, l')$$

根据约束方程是否显含时间, 可以把约束方程分为定常约束和非定常约束. 前者显含时间, 后者则否. 根据约束是否可以解除, 又可以把约束分为单侧约束和双侧约束. 单侧约束是可以解除的约束, 表达式通常为不等式. 在实际情况中约束是否可以解除通常取决于约束力的情况. 约束方程为等式的约束叫做双侧约束.

根据约束力虚功的代数和是否为0可以把约束分为理想约束和非理想约束. 设各个质点的虚位移为 $\delta \mathbf{r}_a$, 受到的约束力为 \mathbf{F}'_a , 如果有

$$\sum_{a=1}^N \mathbf{F}'_a \cdot \delta \mathbf{r}_a = 0$$

则约束是理想约束.

最小作用量原理与拉格朗日方程

定义 2.3 (最小作用量原理) 力学系统的运动规律是作用量泛函

$$S = \int_{t_1}^{t_2} L dt$$

取极值的时候, 也就是 $\delta S = 0$ 时广义坐标所满足的微分方程所描述的. 其中 L 是系统的拉格朗日函数(lagrangian). 这个原理称为最小作用量原理

定理 2.1 在固定边界条件下, 即 $\delta q_i(t_1) = \delta q_i(t_2) = 0$ 时, 哈密顿原理给出拉格朗日方程

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0, \quad (i = 1, 2, \dots, s)$$

这是关于 q_i 的二阶微分方程, 也就是系统的运动方程.

证明 2.1 根据最小作用量原理2.3可以知道

$$\begin{aligned}
 0 = \delta S &= \delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i) dt, \quad (i = 1, 2, \dots, s) \\
 &= \int_{t_1}^{t_2} \sum_{i=1}^s \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] dt \\
 &= \sum_{i=1}^s \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] dt \\
 &= \sum_{i=1}^s \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] dt \\
 &= \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t_1}^{t_2} + \sum_{i=1}^s \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \delta q_i - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] dt
 \end{aligned}$$

上式最后一行第一项根据边界条件 $\delta q_i(t_1) = \delta q_i(t_2) = 0$ 可知为零.所以进一步可以推导出

$$\sum_{i=1}^s \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} \delta q_i - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] dt = 0$$

即, 得证

值得注意的是, 上述变分是指等时变分, 而不是指全变分. 全变分与时间求导算子是不可以交换的, 但是等时变分可以. 此外根据上面的推导过程我们可以知道拉格朗日函数具有如下三条性质:

- (1) 若系统可以拆分成几个独立系统, 则系统总的拉格朗日函数为所有独立子系统拉格朗日函数之和;
- (2) 拉格朗日函数乘以任意常数不会改变运动方程;
- (3) 两个Lagrangian, $L'(q, \dot{q}, t)$ 和 $L(q, \dot{q}, t)$, 如果他们相差某个坐标和时间的函数 $f(q, t)$ 对时间的全导数

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q, t)$$

则 L', L 将给出完全一样的动力学方程.

通常3这种关系称为规范变换. 在规范变换下系统的运动方程不发生变化, 这种性质通常称为系统具有规范不变性.

2.1.2 运动积分与诺特定理

当我们求解一个力学系统的时候, 我们总是优先从守恒力学量去简化方程, 而如何找到非平凡守恒量则是物理学中至关重要的问题. 诺特原理告诉了我们, 系统的对称性和守恒量是密不可分的, 每一种对称性都对应了一个守恒量. 这样我们把寻找对称性的问题归咎于如何找到系统的对称性. 具体而言, 动量守恒与空间的均匀性以及平移变换下的不变性相联系, 角动量守恒与空间的各向同性以及在转动变换下的不变性相联系, 而机械能守恒与时间的均匀性, 即时间平移下的不变性相联系.

定义 2.4 (运动积分) 在运动过程中, 变量 $q_i, \dot{q}_i (i = 1, 2, \dots, s)$ 的某些函数如果不随时间变换则称为运动积分. 对于 s 个自由度的封闭力学系统, 独立的运动积分的个数最多为 $2s - 1$. 运动积分的值完全由初始条件决定.

我们考虑 $2s - 1$ 个运动积分与系统运动的关系. 首先 $2s - 1$ 个运动积分可以表示为

$$C_\alpha = C_\alpha(q_1, q_2, \dots, q_s, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_s, t), \quad \alpha = 1, 2, \dots, 2s - 1$$

换一个视角, 我们在广义坐标 q_i 和广义动量 \dot{q}_i 构成的 $2s$ 维相空间中去考察运动积分, 则每一个运动积分都表示该空间的一个 $2s - 1$ 维子空间(超曲面). 因为轨迹 $(q_i(t), \dot{q}_i(t)) (i = 1, 2, \dots, s)$ 必须同时位于这 $2s - 1$ 个 $2s - 1$ 维子空间(超曲面)上, 这意味着轨迹必须位于这些子空间(超曲面)的交集上. 并且交集在 t 时刻是 $2s$ 维相空间中的一条曲线.

值得注意的是, 如若 C_α 显含时间, 则该交集并不是相空间的一条固定曲面. 若 C_α 中的某些运动积分不显含时, 这样的积分将会对运动给出一些限制, 有助于简化系统的求解, 这种运动积分叫做整体运动积分. 整体运动积分常常与系统的时空对称性有关.

定义 2.5 (广义能量) 对于封闭系统, 时间具有均匀性. 拉格朗日函数在时间平移变换下不变. 也就是说拉格朗日函数不显含时间 t , 系统的能量 E 是一个运动积分, 定义为

$$E = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \quad (12)$$

证明 2.2 系统的自由度为 s , 拉格朗日方程一共有 s 个二阶微分方程

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, 2, \dots, s$$

上式两边乘以 $\dots q_i$, 然后对 i 求和. 得到:

$$\sum_{i=1}^s \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i - \sum_{i=1}^s \frac{\partial L}{\partial q_i} \dot{q}_i = 0$$

等式左边第一项可以写成:

$$\sum_{i=1}^s \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i = \sum_{i=1}^s \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i$$

交换求和 $\sum_{i=1}^s$ 和求导 $\frac{d}{dt}$ 的次序

$$\frac{d}{dt} \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i = 0$$

由于

$$\frac{dL}{dt} = \sum_{i=1}^s \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \frac{\partial L}{\partial t}$$

所以有

$$\frac{d}{dt} \left(\sum_{i=1}^s \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) = - \frac{\partial L}{\partial t}$$

令 $E = \sum_{i=1}^s p_i \dot{q}_i - L$. 这样就得到了广义能量. 并且有

$$\frac{dE}{dt} = - \frac{\partial L}{\partial t}$$

这表明了一件事情, 若 L 不显含时间, 则 $\frac{\partial L}{\partial t} = 0$, 此时 E 是常量. 它是拉格朗日方程的一个运动积分, 称为广义能量积分.

值得注意的是, 广义能量具有能量的量纲, 但不一定就是相对惯性系的机械能. 对于完整有势系统, $L = T(q, \dot{q}) - V(q)$. 在动能 $T(q, \dot{q})$ 是广义速度 \dot{q}_i 的二次齐次函数时, 能量 E 的形式可以写成

$$E = T(q, \dot{q}) + V(q) = \sum_{a=1}^N \frac{1}{2} m_a v_a^2 + V(\mathbf{r}_1, \mathbf{r}_2, \dots)$$

在这种情况下, E 表示系统的机械能.

定义 2.6 (广义动量) 当我们用广义坐标去描述系统时, 定义广义动量

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (13)$$

广义动量包含了通常的动量、角动量等.

例 2.1 广义动量并不是唯一的

证明 2.3 相差广义坐标 q 和时间 t 的某一函数 $f(q, t)$ 对时间的全微分的两个拉格朗日函数有相同的动力学方程, 也就是说

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt}f(q, t)$$

L, L' 通过拉格朗日方程可以得到相同的微分方程. 根据广义动量的定义(13), 相应于 L' 有

$$\begin{aligned} p'_i &= \frac{\partial L}{\partial \dot{q}_i} \\ &= p_i + \frac{\partial}{\partial \dot{q}_i} \left(\frac{d}{dt}f(q, t) \right) \\ &= p_i + \frac{\partial}{\partial \dot{q}_i} \left(\sum_{j=1}^s \frac{\partial f}{\partial \dot{q}_j} \dot{q}_j + \frac{\partial f}{\partial t} \right) \\ &= p_i + \frac{\partial f}{\partial \dot{q}_i} \end{aligned}$$

对称性和守恒律之间是存在对应关系的, 对应关系由Noether定理所描述.

定理 2.2 (Noether定理) 任何一种在时间坐标连续变换下系统作用量保持不变的对称性都存在相对应的运动积分

证明 2.4 假设 $L(q, \dot{q}, t)$ 是描述一个封闭系统的拉格朗日函数. 设无穷小变换为

$$\begin{aligned} t &\mapsto t' = t + \epsilon_\alpha \zeta_\alpha(q(t), t) \\ q_i &\mapsto q'_i = q_i + \epsilon_\alpha \eta_{i\alpha}(q(t), t) \end{aligned} \tag{14}$$

其中 $\epsilon_\alpha (\alpha = 1, 2, \dots, r)$ 是无穷小实数, $\zeta_\alpha(q, t), \eta_{i\alpha}(q, t)$ 是 q, t 的连续函数. 这里我们为了简化表达式, 采用了Einstein双重指标求和规则($a_i b_i = \sum_i a_i b_i$). 当无穷小实数 $\epsilon_\alpha = 0$ 时, $t' = t, q'_i = q_i$, 此时为恒等变换. 如果系统在变换(14)下, 作用量

$$S = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt \tag{15}$$

不变, 也就是有

$$\begin{aligned} S' &= \int_{t'_1}^{t'_2} L'(q'(t'), \dot{q}'(t'), t') dt' \\ &= \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt = S \end{aligned}$$

则存在相应的运动积分. 注意上式中 $\dot{q}'_i = \frac{dq'_i}{dt'}$. 我们继续往下面推导出具体的运动积分形式.

注意到方程(15)的第一行可以写成

$$\int_{t'_1}^{t'_2} L'(q'(t'), \dot{q}'(t'), t') dt' = \int_{t_1}^{t_2} L'(q'(t'), \dot{q}'(t'), t') \frac{dt'}{dt} dt$$

则方程(15)表明了变换前后拉格朗日函数之间存在这样的关系

$$L'(q'(t'), \dot{q}'(t'), t') \frac{dt'}{dt} = L(q(t), \dot{q}(t), t) \quad (16)$$

值得注意的是变换前后拉格朗日函数的不变性并没有确定 $L'(q'(t'), \dot{q}'(t'), t')$ 的形式. 换句话说, 它现在还是不确定的. 当我们要求无穷小变换前后拉格朗日函数具有相同的形式, 可以按照下面两种方式指定 L' 的形式

(1) 不变性:

如果变换前后满足

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) \quad (17)$$

则称 L 在变换(14)下具有形式不变性

(2) 协变性:

如果变换前后满足

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} \Lambda(q, t) \quad (18)$$

则称 L 在变换(14)下具有形式协变性.

(18)是(17)的推广. 这里 $\Lambda(q, t)$ 是 q, t 的连续可微函数. 具有性质(17)和性质(18)的变换(14)称为系统的对称变换, 相应地称系统在该变换下是对称的.

把式子(3)代入式子(18)中, 有

$$\left(L(q'(t'), \dot{q}'(t'), t') + \frac{d}{dt'} \Lambda(q'(t'), t') \right) \frac{dt'}{dt} = L(q(t), \dot{q}(t), t) \quad (19)$$

由(14)可以得到

$$\begin{aligned}
\frac{dt'}{dt} &= 1 + \epsilon_\alpha \frac{d\zeta_\alpha(q(t), t)}{dt}, \\
\dot{q}'_i &= \frac{d\dot{q}'_i}{dt} = \frac{\dot{q}'_i}{dt} \frac{dt}{dt'} = \frac{\left(\dot{q}_i + \epsilon_\alpha \frac{d\eta_{i\alpha}(q(t), t)}{dt} \right)}{\left(1 + \epsilon_\alpha \frac{d\zeta_\alpha(q(t), t)}{dt} \right)} \\
&= \left(\dot{q}_i + \epsilon_\alpha \frac{d\eta_{i\alpha}(q(t), t)}{dt} \right) \left(1 - \epsilon_\alpha \frac{d\zeta_\alpha(q(t), t)}{dt} \right) \\
&= \dot{q}_i + \epsilon_\alpha \left(\frac{\eta_{i\alpha}(q(t), t)}{dt} - \dot{q}_i \frac{d\zeta_\alpha(q(t), t)}{dt} \right)
\end{aligned}$$

其中均保留到 ϵ 的一阶项. 有

$$\begin{aligned}
L(q'(t'), \dot{q}'(t'), t') &= L(q(t), \dot{q}(t), t) + \frac{\partial L}{\partial t}(t' - t) + \left[\frac{\partial L}{\partial \dot{q}_i}(q'_i - q_i) + \frac{\partial L}{\partial \dot{q}_i}(\dot{q}'_i - \dot{q}_i) \right] \\
&= L(q(t), \dot{q}(t), t) + \frac{\partial L}{\partial t} \epsilon_\alpha \zeta_\alpha(q(t), t) \\
&\quad + \epsilon_\alpha \left[\eta_{i\alpha}(q(t), t) \frac{\partial L}{\partial q_i} + \left(\frac{d\eta_{i\alpha}(q(t), t)}{dt} - \dot{q}_i \frac{d\zeta_\alpha(q(t), t)}{dt} \right) \frac{\partial L}{\partial \dot{q}_i} \right] \\
\Lambda(q'(t'), t') &= \Lambda(q(t), t) + \frac{\partial \Lambda}{\partial t}(t' - t) + \frac{\partial \Lambda}{\partial q_i}(q'_i - q_i) \\
&= \Lambda(q(t), t) + \epsilon_\alpha \left[\frac{\partial \Lambda}{\partial t} \zeta_\alpha(q(t), t) + \frac{\partial \Lambda}{\partial q_i} \eta_{i\alpha}(q(t), t) \right] \\
&= \Lambda(q(t), t) + \epsilon_\alpha \Pi_\alpha(q(t), t)
\end{aligned}$$

其中

$$\Pi_\alpha(q(t), t) = \frac{\partial \Lambda}{\partial t} \zeta_\alpha(q(t), t) + \frac{\partial \Lambda}{\partial q_i} \eta_{i\alpha}(q(t), t) \quad (20)$$

它是 q, t 的函数. 注意到 $L(q'(t'), \dot{q}'(t'), t') - L(q(t), \dot{q}(t), t)$ 在变换(14)下是一阶无穷小, 这意味着 $\frac{d\Lambda(q'(t'), t')}{dt'}$ 是同阶无穷小. 将以上各式代入(19)保留一阶项可以得到

$$\begin{aligned}
&\zeta_\alpha(q(t), t) \frac{\partial L}{\partial t} + \left[\eta_{i\alpha}(q(t), t) \frac{\partial L}{\partial q_i} + \left(\frac{d\eta_{i\alpha}(q(t), t)}{dt} - \dot{q}_i \frac{d\zeta_\alpha(q(t), t)}{dt} \right) \frac{\partial L}{\partial \dot{q}_i} \right] \\
&\quad + L \frac{d\zeta_\alpha(q(t), t)}{dt} = - \frac{d}{dt} \Pi_\alpha(q(t), t)
\end{aligned} \quad (21)$$

方程(21)可以用来判断变换(14)是否为对称变换. 对于给定的 L 和变换, 计算上式左边, 若可以表示成某一函数对时间的全微分, 则变换是对称变换, 同时 $\Pi_\alpha(q(t), t)$ 也就确定下来

了. 注意到

$$\frac{d}{dt}L(q(t), \dot{q}(t), t) = \frac{\partial L}{\partial t} + \left(\frac{\partial L}{\partial q_i} + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right)$$

可以将式子(21)改写成

$$\frac{d}{dt} \left[(\eta_{i\alpha} - \dot{q}_i \xi_\alpha) \frac{\partial L}{\partial \dot{q}_i} + L \xi_\alpha + \Pi_\alpha \right] + \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) (\eta_{i\alpha} - \dot{q}_i \xi_\alpha) = 0$$

所以当系统运动满足拉格朗日方程时, 存在运动积分

$$\begin{aligned} I_\alpha(q, t) &= (\eta_{i\alpha} - \dot{q}_i \xi_\alpha) \frac{\partial L}{\partial \dot{q}_i} + L \xi_\alpha + \Pi_\alpha \\ &= \eta_{i\alpha} p_i - E \xi_\alpha + \Pi_\alpha, \quad (\alpha = 1, 2, \dots, r) \end{aligned} \quad (22)$$

其中第二等号用了广义动量和广义能量的定义, 得证

举几个特殊的例子.

1. 当拉格朗日函数不显含时间 t 的时候, $L = L(q, \dot{q})$, 如果仅作时间平移变换, 拉格朗日函数在时间平移变换下保持不变, 也就是说 $\Lambda = 0$, 取 $\xi_\alpha = 1, \eta_{i\alpha} = 0$, 根据(22)可知运动积分为 $I = -E$. 也就是说时间平移不变性的条件会得到广义能量积分.
2. 如果仅有空间变换, 即 $\xi_\alpha = 0$, 假设Lagrangian在这个变换下保持不变, 即 $\Lambda = 0$, 根据(22)可以得到运动积分为 $I = \eta_{i\alpha} p_i$, 这就是动量守恒与角动量守恒.

2.1.3 哈密顿力学

定义 2.7 (拉格朗日函数的勒让德变换) 定义Hamiltonian H 是拉格朗日函数 L 的勒让德变换

$$H(p, q, t) = \sum_{i=1}^s p_i \dot{q}_i - L \quad (23)$$

值得注意, Hamiltonian是 p, q, t 的函数, 而 L 是 \dot{q}, q, t 的函数. 勒让德变换表明, 拉格朗日函数和哈密顿函数对参数 λ 的偏导数之间存在关系

$$\left(\frac{\partial H}{\partial \lambda} \right)_{p, q} = - \left(\frac{\partial L}{\partial \lambda} \right)_{\dot{q}, q} \quad (24)$$

一个特例是对时间的偏导

$$\left(\frac{\partial H}{\partial t} \right)_{p, q} = - \left(\frac{\partial L}{\partial t} \right)_{\dot{q}, q} \quad (25)$$

特别是

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \quad (26)$$

定义 2.8 (哈密顿正则方程) 用变量 p 和 q 描述系统的状态时, 对于自由度为 s 的完整保守系统, 其动力学方程为

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \quad (i = 1, 2, \dots, s) \end{aligned} \quad (27)$$

称为Hamilton方程, 它们构成 $2s$ 个位置函数的 $2s$ 个一阶微分方程组. H 称为哈密顿函数, 它有一个特别的性质

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}$$

证明 2.5 根据拉格朗日函数的勒让德变换

$$H = \sum_{i=1}^s p_i \dot{q}_i - L$$

对上式两边做全微分

$$\begin{aligned} dH &= d \left(\sum_{i=1}^s p_i \dot{q}_i - L \right) \\ &= \left(\sum_{i=1}^s p_i d\dot{q}_i + \sum_{i=1}^s \dot{q}_i dp_i \right) - \left(\sum_{i=1}^s \dot{p}_i dq_i + \sum_{i=1}^s p_i d\dot{q}_i \right) \\ &= \sum_{i=1}^s \dot{q}_i dp_i - \sum_{i=1}^s \dot{p}_i dq_i \end{aligned}$$

值得注意, 上面第二行我们使用了拉格朗日方程 $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i}$ 和广义动量的定义 $p_i = \frac{\partial L}{\partial \dot{q}_i}$ 同时我们对哈密顿函数 $H = H(p_i, q_i)$ 进行全微分展开

$$dH = \sum_{i=1}^s \frac{\partial H}{\partial q_i} dq_i + \sum_{i=1}^s \frac{\partial H}{\partial p_i} dp_i$$

上述两式比较可以得到哈密顿正则方程

$$\begin{aligned}\dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}\end{aligned}$$

哈密顿运动积分

关系式(25)表明拉格朗日函数不显含时间时, 相应的哈密顿函数也不显含时间. 又根据(26)可以知道此时有能量守恒. 当 L 中有循环坐标时, 根据(24)可以知道它们也是 H 的循环坐标. 再根据哈密顿方程可以知道相应的运动积分是广义动量 p_i .

假设有 m 个循环坐标 $q_i (i = s - m + 1, \dots, s)$, 则对应的广义动量为

$$p_i = \beta_i$$

将广义动量代入 H 中有

$$H = H(q_k, p_k, \beta_i, t), \quad (k = 1, 2, \dots, s - m, \quad i = s - m + 1, \dots, s)$$

非循环坐标满足哈密顿正则方程

$$\begin{aligned}\dot{q}_k &= \frac{\partial H}{\partial p_k} \\ \dot{p}_k &= -\frac{\partial H}{\partial q_k}\end{aligned}$$

由此解出 $2(s - m)$ 个 q_k 和 p_k . 将其代入 H 中, 并通过 $\dot{q}_k = \frac{\partial H}{\partial p_i}$ 积分可以得到

$$q_i = \int \frac{\partial H}{\partial p_i} dt$$

可以求出循环坐标. 哈密顿力学里, 与循环坐标有关的自由度是完全可以分离的. 从而对有 m 个循环坐标的情况可化为自由度为 $s - m$ 的情况, 简化问题的求解.

2.1.4 泊松括号

定义 2.9 (泊松括号) 任意一函数 f, g , 它们是 q, p, t 的函数, 泊松括号定义为

$$\{f, g\} = \sum_k \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \quad (28)$$

泊松括号有一系列基本性质:

(1) 反对称性

$$\{f, g\} = -\{g, f\}$$

(2) 与常数的括号为0

$$\{f, c\} = 0, \quad c \text{ 为常数}$$

(3) 分配律(线性性)

$$\{f_1 + f_2, g\} = \{f_1, g\} + \{f_2, g\}$$

(4) 乘积法则

$$\{f_1 f_2, g\} = f_1 \{f_2, g\} + \{f_1, g\} f_2$$

(5) 微分法则

$$\frac{\partial}{\partial t} \{f, g\} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\}$$

(6) 雅克比恒等式(Jacobi identity)

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$$

有关这些性质的证明, 我们放在后面去证明.

定义 2.10 (正则变量) 如果一组变量 q_k, p_k 满足下列关系

$$\begin{aligned} \{q_k, f\} &= \frac{\partial f}{\partial p_k} \\ \{p_k, f\} &= -\frac{\partial f}{\partial q_k} \\ \{q_i, q_j\} &= 0, \quad \{p_i, p_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij} \end{aligned}$$

则叫做正则变量.

定理 2.3 (泊松公式) 正则变量的函数 $f(q, p, t)$ 对时间的全微分可以表示为

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \quad (29)$$

证明 2.6 首先 $f = f(q, p, t)$, 全微分展开

$$\frac{df}{dt} = \sum_k \frac{\partial f}{\partial q_k} \dot{q}_k + \frac{\partial f}{\partial p_k} \dot{p}_k + \frac{\partial f}{\partial t}$$

将哈密顿正则方程 $\dot{q}_k = \frac{\partial H}{\partial p_k}, \dot{p}_k = -\frac{\partial H}{\partial q_k}$ 代入上式, 得到

$$\frac{df}{dt} = \sum_k \frac{\partial f}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial H}{\partial q_k} + \frac{\partial f}{\partial t} = \{f, H\} + \frac{\partial f}{\partial t}$$

得证

根据泊松公式, 哈密顿正则方程可以改写成

$$\dot{q}_i = \{q_i, H\}, \quad \dot{p}_i = \{p_i, H\}$$

其中 H 为系统的哈密顿函数. 值得注意的是, 若 $\frac{df}{dt} = 0$, 则 f 是一个运动积分, 此时 f 满足运动方程:

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} = 0$$

若 f 不显含时间 t , 则有 $\{f, H\} = 0$. 这里值得注意, 一个力学量如果不显含时间, 这个力学量存在运动积分的条件就是 $\{f, H\} = 0$. 这个条件在后面的章节里是判断守恒量的关键条件.

定理 2.4 (泊松定理) 如果 f 和 g 是两个运动积分, 则它们构成的泊松括号也是运动积分

$$\{f, g\} = \text{const}$$

证明 2.7 如果 f, g 不显含时间, 根据雅克比恒等式, 令 $h = H$ 可以得到

$$\{f, \{g, H\}\} + \{g, \{H, f\}\} + \{H, \{f, g\}\} = 0$$

由此可见, 如果有 $\{g, H\} = 0, \{f, H\} = 0$, 则 $\{\{f, g\}, H\} = 0$, 于是不显含时的情况得证.

如果运动积分 f, g 显含时间 t . 根据泊松公式(29)可以得到

$$\begin{aligned}\frac{d}{dt}\{f, g\} &= \{\{f, g\}, H\} + \frac{\partial}{\partial t}\{f, g\} \\ &= -\{\{g, H\}, f\} - \{\{H, f\}, g\} + \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} \\ &= \left\{f, \frac{dg}{dt}\right\} + \left\{\frac{df}{dt}, g\right\}\end{aligned}$$

若 f, g 都是运动积分, 则有 $\frac{df}{dt} = 0, \frac{dg}{dt} = 0$, 显然上式为0, 泊松定理得证.

当然, 应用泊松定理, 我们不能总是得到新的运动积分, 因为仅有 $2s - 1$ 个运动积分, s 是自由度. 在一些情况下我们得到的是一些平凡的结果, 泊松括号为常数. 在另一些情况下, 新得到的运动积分只不过是原来的运动积分 f 和 g 的函数. 然而如果不是上面两种情况, 则泊松括号给出新的运动积分.

下面我们证明一下泊松括号性质的证明, 首先是性质1的证明

证明 2.8

$$\begin{aligned}\{f, g\} &= \sum_k \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \\ &= - \sum_k \left(\frac{\partial g}{\partial q_k} \frac{\partial f}{\partial p_k} - \frac{\partial g}{\partial p_k} \frac{\partial f}{\partial q_k} \right) \\ &= -\{g, f\}\end{aligned}$$

此为性质1的结果

性质2根据常数的偏导数为0可以直接得到, 性质3的证明如下:

证明 2.9

$$\begin{aligned}\{f_1 + f_2, g\} &= \sum_k \left(\frac{\partial(f_1 + f_2)}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial(f_1 + f_2)}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \\ &= \sum_k \left(\frac{\partial f_1}{\partial q_k} \frac{\partial g}{\partial p_k} + \frac{\partial f_2}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k} \frac{\partial g}{\partial q_k} - \frac{\partial f_2}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \\ &= \sum_k \left[\left(\frac{\partial f_1}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k} \frac{\partial g}{\partial q_k} \right) + \left(\frac{\partial f_2}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_2}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \right] \\ &= \{f_1, g\} + \{f_2, g\}\end{aligned}$$

性质4的证明

证明 2.10

$$\begin{aligned}
 \{f_1 f_2, g\} &= \sum_k \left[\frac{\partial(f_1 f_2)}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial(f_1 f_2)}{\partial p_k} \frac{\partial g}{\partial q_k} \right] \\
 &= \sum_k \left[f_1 \left(\frac{\partial f_2}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_2}{\partial p_k} \frac{\partial g}{\partial q_k} \right) + f_2 \left(\frac{\partial f_1}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \right] \\
 &= f_1 \sum_k \left(\frac{\partial f_2}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_2}{\partial p_k} \frac{\partial g}{\partial q_k} \right) + f_2 \sum_k \left(\frac{\partial f_1}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \\
 &= f_1 \{f_2, g\} + f_2 \{f_1, g\}
 \end{aligned}$$

此即为所要证明的结果

性质5留给读者自证, 接下来我们证明雅克比恒等式

证明 2.11 为了证明这个恒等式, 我们注意到这样的结果. 根据泊松括号的定义, 泊松括号是 f 和 g 一阶导数的双线性齐次函数. 所以很自然的 $\{h, \{f, g\}\}$ 是 f 和 g 二阶导数的线性齐次函数. 因此6中的等式左端是所有三个函数 f, g, h 的二阶导数的线性齐次函数. 我们将含有 f 的二阶导数合并同类项. 第一个括号 $\{f, \{g, h\}\}$ 不含这样的项, 因为他只涉及 f 的一阶导数. 定义

$$D_1(\varphi) = \{g, \varphi\}, \quad D_2(\varphi) = \{h, \varphi\}$$

这样可以将性质6中的符号写成

$$\begin{aligned}
 \{g, \{h, f\}\} + \{h, \{f, g\}\} &= \{g, \{h, f\}\} - \{h, \{g, f\}\} \\
 &= D_1(D_2(f)) - D_2(D_1(f)) \\
 &= (D_1 D_2 - D_2 D_1)f \\
 &= [D_1, D_2]f
 \end{aligned}$$

显然 $[D_1, D_2]$ 是一个一阶微分算子. 事实上, 线性微分算子的一般形式为

$$D_1 = \sum_k \xi_k \frac{\partial}{\partial x_k}, \quad D_2 = \sum_k \eta_k \frac{\partial}{\partial x_k}$$

其中 ξ_k, η_k 是 x_1, x_2, \dots 的任意函数. 于是

$$\begin{aligned} D_1 D_2 &= \sum_{k,l} \xi_k \eta_l \frac{\partial^2}{\partial x_k \partial x_l} + \sum_{k,l} \xi_k \frac{\partial \eta_l}{\partial x_k} \frac{\partial}{\partial x_l} \\ D_2 D_1 &= \sum_{k,l} \eta_k \xi_l \frac{\partial^2}{\partial x_k \partial x_l} + \sum_{k,l} \eta_k \frac{\partial \xi_l}{\partial x_k} \frac{\partial}{\partial x_l} \end{aligned}$$

它们之差

$$D_1 D_2 - D_2 D_1 = \sum_{k,l} \left(\xi_k \frac{\partial \eta_l}{\partial x_k} - \eta_k \frac{\partial \xi_l}{\partial x_k} \right) \frac{\partial}{\partial x_l}$$

是一个一阶微分算子. 注意到我们现在的系统有

$$x_k = \begin{cases} q_k, & (k = 1, 2, \dots, s); \\ p_{k-s}, & (k = s+1, s+2, \dots, 2s). \end{cases}$$

于是有

$$\{g, \{h, f\}\} + \{h, \{f, g\}\} = \sum_{k,l} \left(\xi_k \frac{\partial \eta_l}{\partial x_k} - \eta_k \frac{\partial \xi_l}{\partial x_k} \right) \frac{\partial f}{\partial x_l} = \sum_k \left(A_k \frac{\partial f}{\partial q_k} + B_k \frac{\partial f}{\partial p_k} \right)$$

注意 A_k 和 B_k 是 g 和 h 的函数, 但不是 f 的函数, 所以如果令 $f = p_i$, 得到

$$\{g, \{h, p_i\}\} + \{h, \{p_i, g\}\} = \left\{ g, \frac{\partial h}{\partial q_i} \right\} - \left\{ h, \frac{\partial g}{\partial q_i} \right\} = \frac{\partial}{\partial q_i} \{g, h\} = B_i$$

同理, 令 $f = q_i$

$$\{g, \{h, q_i\}\} + \{h, \{q_i, g\}\} = - \left\{ g, \frac{\partial h}{\partial p_i} \right\} + \left\{ h, \frac{\partial g}{\partial p_i} \right\} = - \frac{\partial}{\partial p_i} \{g, h\} = A_i$$

由此有

$$\begin{aligned} \{g, \{h, f\}\} + \{h, \{f, g\}\} &= \sum_k \left(\frac{\partial}{\partial q_k} \{g, h\} \frac{\partial f}{\partial p_k} - \frac{\partial}{\partial p_k} \{g, h\} \frac{\partial f}{\partial q_k} \right) \\ &= \{\{g, h\}, f\} = -\{f, \{g, h\}\} \end{aligned}$$

所以雅克比恒等式得证

从代数结构来看, 泊松括号是李代数的乘法, 不同函数之间的泊松括号给出的雅克比恒等式是李代数的基本定义关系.

例 2.2 质点动量 \mathbf{p} 和角动量 $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ 的笛卡尔坐标分量组成的泊松括号满足

$$\{J_i, p_i\} = \varepsilon_{ijk} p_k$$

其中 $i, j, k = x, y, z$, 这里为了简化书写采用了Einstein双重指标求和规则, 而 ε_{ijk} 是Levi-Civita张量, 其定义为

$$\varepsilon_{ijk} = \begin{cases} 1, & \text{当 } ijk \text{ 为偶置换;} \\ -1, & \text{当 } ijk \text{ 为奇置换;} \\ 0, & \text{当 } ijk \text{ 有重复指标.} \end{cases}$$

证明 2.12 根据角动量的定义 $\mathbf{J} = \mathbf{r} \times \mathbf{p}$, 可以得到笛卡尔系下各个坐标分量 J_x, J_y, J_z .

$$J_i = \varepsilon_{ijk} x_j p_k$$

x_i, p_i 是正则力学量

$$\{p_k, f\} = -\frac{\partial f}{\partial x_k}$$

所以得到

$$\begin{aligned} \{J_i, p_j\} &= \frac{\partial J_i}{\partial x_j} \\ &= \frac{\varepsilon_{ij'k} x_{j'} p_k}{x_j} \\ &= \varepsilon_{ij'k} \frac{x_{j'} p_k}{x_j} = \varepsilon_{ij'k} \delta_{jj'} p_k = \varepsilon_{ijk} p_k \end{aligned}$$

得证

例 2.3 角动量 \mathbf{J} 各个分量组成的泊松括号满足

$$\{J_i, J_j\} = \varepsilon_{ijk} J_k$$

证明 2.13

$$\begin{aligned}
\{J_i, J_j\} &= \{J_i, \varepsilon_{i'jk'} x_{k'} p_{i'}\} \\
&= \varepsilon_{i'jk'} \{J_i, x_{k'} p_{i'}\} \\
&= \varepsilon_{i'jk'} [\{J_i, x_{k'}\} p_{i'} + x_{k'} \{J_i, p_{i'}\}] \\
&= \varepsilon_{i'jk'} (\varepsilon_{ik'\alpha} x_\alpha p_{i'} + x_{k'} \varepsilon_{ii'\beta} p_\beta) \\
&= \varepsilon_{i'jk'} \varepsilon_{ik'\alpha} x_\alpha p_{i'} + \varepsilon_{i'jk'} \varepsilon_{ii'\beta} x_{k'} p_\beta \\
&= (\delta_{i'\alpha} \delta_{ji} - \delta_{i'i} \delta_{j\alpha}) x_\alpha p_{i'} + (\delta_{j\beta} \delta_{k'i} - \delta_{ji} \delta_{k'\beta}) x_{k'} p_\beta \\
&= \delta_{ji} x_{i'} p_{i'} - x_j p_i + x_i p_j - \delta_{ij} x_{k'} p_{k'} \\
&= \varepsilon_{ijk} x_i p_j
\end{aligned}$$

得证

2.1.5 可积系统

泊松定理表明, 有可能通过已有的运动积分构造新的运动积分. 更多的运动积分有助于解析地描述系统的运动情况.

定义 2.11 (可积系统) 对于有 s 个自由度的力学系统, 如果可以找到 s 个相互独立的运动积分 $\phi_1, \phi_2, \dots, \phi_s$, 且满足

$$[\phi_i, \phi_j] = 0, \quad (i, j = 1, 2, \dots, s)$$

这样的系统称为可积系统, 此时称 $\phi_1, \phi_2, \dots, \phi_s$ 是相互正合的.

这个结论是由刘维尔首先证明的, 所以也称为刘维尔可积性定理. 系统可积是, 其哈密顿函数属于运动积分之中一个, 或者可以用它表示出来, 而系统运动方程可以通过积分的方法求解. 这个定理的证明需要利用辛流形的内容, 我们暂时放在后面的补充内容里.

除了刘维尔可积性外, 还存在其它类型的可积性. 例如对于一个系统, 如果可以构造两个矩阵 $\mathbf{A} = (a_{ij})$ 和 $\mathbf{B} = (b_{ij})$, 使得他们满足关系

$$\frac{d\mathbf{A}}{dt} = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = [\mathbf{A}, \mathbf{B}]$$

同时上述关系完全等同于系统的动力学方程, 则该系统也是可积的. 这里矩阵 \mathbf{A} 和 \mathbf{B} 称作 Lax pair. 上述关系称作 Lax 方程.

利用 \mathbf{A} 可以构造运动积分, 首先我们计算 $\frac{d\mathbf{A}^n}{dt}$

$$\begin{aligned}\frac{d\mathbf{A}^n}{dt} &= \sum_{k=0}^n \mathbf{A}^k \frac{d\mathbf{A}}{dt} \mathbf{A}^{n-1-k} = \sum_{k=0}^n \mathbf{A}^k [\mathbf{A}, \mathbf{B}] \mathbf{A}^{n-1-k} \\ &= \sum_{k=0}^n (\mathbf{A}^{k+1} \mathbf{B} \mathbf{A}^{n-1-k} - \mathbf{A}^k \mathbf{B} \mathbf{A}^{n-k})\end{aligned}$$

两边求trace

$$\frac{dtr(\mathbf{A}^n)}{dt} = \sum_{k=0}^n tr(\mathbf{A}^{k+1} \mathbf{B} \mathbf{A}^{n-1-k} - \mathbf{A}^k \mathbf{B} \mathbf{A}^{n-k}) = \sum_{k=0}^n [tr(\mathbf{B} \mathbf{A}^n) - tr(\mathbf{B} \mathbf{A}^n)] = 0$$

所以 $tr(\mathbf{A}^n)$ 是系统的运动积分

例 2.4 对于一位谐振子系统, 拉格朗日函数为:

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2$$

可以构造矩阵

$$\mathbf{A} = \sqrt{m} \begin{bmatrix} \dot{x} & \omega x \\ \omega x & -\dot{x} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & \frac{\omega}{2} \\ -\frac{\omega}{2} & 0 \end{bmatrix}$$

这里 \mathbf{A} 具有性质 $\mathbf{A}^2 = E$, E 是系统具有的机械能, $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2$, 相应地有

$$tr(\mathbf{A}^{2n+1}) = 0, \quad tr(\mathbf{A}^{2n}) = 2(2E)^n$$

由此可以看出仅仅只有机械能 E 是系统的运动积分. 一个自由度的系统有一个运动积分, 一维谐振子系统是可积系统.

值得注意, 对于多体系统, 可积的系统是非常少的, 绝大多数是不可积的. 最具代表性的例子是所谓的三体问题, 仅当有比较强的限制时才可以精确求解. 通常必须借助数值计算确定多体运动行为以及相关性质.

2.1.6 哈密顿雅克比方程

作为坐标时间函数的作用量

之前的章节2.1.1里我们讨论过在两个固定位置 $q^{(1)}$ 和 $q^{(2)}$ 之间轨道的积分

$$S = \int_{t_1}^{t_2} L dt \quad (30)$$

其中 $q^{(1)}$ 和 $q^{(2)}$ 是在给定时刻 t_1 和 t_2 系统占据的位置. 在对作用量变分时, 我们比较有相同值 $q(t_1)$ 和 $q(t_2)$ 的临近轨道的这个积分值. 这些轨道中只有一条对应这使得运动, 这就是积分 S 取极值的轨道.

我们换一个角度去看待作用量 S . 我们将 S 看做表征沿着真实轨道运动的量, 并比较有相同初始位置 $q(t_1) = q^{(1)}$ 但 t_2 时刻通过不同位置的那些轨道对应的作用量的值. 也就是说, 真实轨道作用量积分看做积分上限对应坐标的函数

定理 2.5 作用量对坐标的偏导为对应的动量:

$$p_i = \frac{\partial S}{\partial q_i} \quad (31)$$

证明 2.14 首先我们计算从一条轨道变到邻近轨道时, 作用量产生的变化量为

$$\begin{aligned} \delta S &= \sum_i \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \\ &= \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t_1}^{t_2} + \sum_i \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta q_i dt \end{aligned}$$

实际运动满足拉格朗日方程, 所以第二项积分为0. 第一项积分下限对应的 $\delta q(t_1) = 0$, 将 $\delta q(t_2)$ 简记为 δq , 而 $p_i = \frac{\partial L}{\partial \dot{q}_i}$. 最后可得:

$$\delta = \sum_i p_i \delta q_i$$

所以作用量对坐标的偏导数为

$$\frac{\partial S}{\partial q_i} = p_i$$

得证

定理 2.6 作用量对时间的偏导为哈密顿函数的负数

$$\frac{\partial S}{\partial t} = -H$$

证明 2.15 我们考虑给定时刻 t_1 从给定位置 $q^{(1)}$ 出发, 在不同时刻 $t_2 = t$ 终结于给定位置 $q^{(2)}$ 的轨道, 可以将作用量看成时间的显函数. 根据作用量定义, 沿轨道对时间的全导数为

$$\frac{dS}{dt} = L$$

将 S 看成坐标和时间的函数, 全微分展开

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i = \frac{\partial S}{\partial t} + \sum_i p_i \dot{q}_i$$

比较上面两个式子, 可以得到

$$\frac{\partial S}{\partial t} = L - \sum_i p_i \dot{q}_i = -H$$

得证

根据上面定理31和定理2.6可以得到作用量全微分表达式

$$dS = \sum_i p_i dq_i - H dt \quad (32)$$

其中作用量看做(30)中积分上限坐标和时间的函数. 假设运动的初始点和终点的坐标与时间都发生变化. 显然, S 为两端的差

$$dS = \sum_i p_i^{(2)} dq_i^{(2)} - H^{(2)} dt^{(2)} - \sum_i p_i^{(1)} dq_i^{(1)} + H^{(1)} dt^{(1)}$$

这个关系表明, 在运动过程中, 无论外部对系统有什么样的影响, 运动的末态都不会是初态的任意函数, 只有上式右端表达式能写成全微分的运动才是可能的. 所以存在与拉格朗日函数具体形式无关的最小作用量原理, 这给了可能运动的范围一定限制. 如果将坐标和动量看成独立变量, 并且将方程(32)写成积分形式

$$S = \int \sum_i (p_i dq_i - H dt) \quad (33)$$

则可以由最小作用量原理推导出哈密顿方程

$$\begin{aligned} \delta S &= \sum_i \int_{t_1}^{t_2} \left(\delta p_i dq_i + p_i \delta dq_i - \frac{\partial H}{\partial q_i} \delta q_i dt - \frac{\partial H}{\partial p_i} \delta p_i dt \right) \\ &= \sum_i \int_{t_1}^{t_2} \delta p_i \left(dq_i - \frac{\partial H}{\partial p_i} dt \right) + \sum_i p_i \delta q_i \Big|_{t_1}^{t_2} - \sum_i \int_{t_1}^{t_2} \delta q_i \left(dp_i + \frac{\partial H}{\partial q_i} dt \right) = 0 \end{aligned}$$

积分限两端有 $\delta q = 0$, 这样积分出来的项为0, 只剩下第一项和第三项两个积分项, 而剩下

两项的 p, q 是独立的, 只有被积函数为0才能满足整个表达式为0. 于是有

$$dq_i = \frac{\partial H}{\partial p_i} dt, \quad dp_i = -\frac{\partial H}{\partial q_i} dt$$

这便是哈密顿正则方程

哈密顿雅克比方程

根据定理2.6我们可以知道函数 $S(q, t)$ 对时间的偏导数与哈密顿函数的关系为

$$\frac{\partial S}{\partial t} + H(q, p, t) = 0$$

我们将 $p = \frac{\partial S}{\partial q}$ 代入上式, 得到 $S(q, t)$ 应该满足

$$\frac{\partial S}{\partial t} + H(q_1, \dots, q_s; \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_s}; t) = 0$$

这是一阶偏微分方程, 叫做哈密顿-雅克比方程.

2.2 量子力学中的概率

20世纪初开始, 实验物理积累了大量令人费解的奇怪现象, 这些现象表明, 经典物理似乎不是那么适用了. 光和电子表现的有时像波, 有时像粒子. 为了解决这一种矛盾, 我们似乎要对理论进行一定的妥协. 在这种妥协的方案下, 我们退而求次, 从拉普拉斯的决定论到概率. 但更为重要的发现是, 量子力学的概率叠加并不是拉普拉斯经典概率理论的叠加定律. 但当实验所涉及的对象尺寸增大时, 物理世界的量子力学概率与经典概率行为十分相近. 所以经典的概率只能描述大量多粒子系统的整体性质, 而在单粒子的概率行为上是无能为力的.

2.2.1 双缝干涉

概率的数学概念在量子力学中并没有什么改变. 当我们说一个实验结果的某个结果概率为 p 时, 我们就是以传统意义讲的, 也就是这个实验重复多次, 出现这个结果的频率大致是 p . 这里就不在赘述.

我们在量子力学里所需要改变的观点并非数学的观点, 而是物理上量子力学计算概率的方法. 在微观尺度非常小的情况下, 这个量子力学上的概率影响最大. 首先我们从一个实验开始讲起.

一个假想实验如图2所示. . A 处有一个电子源 S . S 处的所有电子以相同的能量向各个方向射出, 撞击在屏 B 上, 屏 B 上有1和2两个小孔, 电子可以通过它们. 最后在 B 后面的屏 C 处放置一台电子探测器, 该探测器距离屏幕中心为 x .

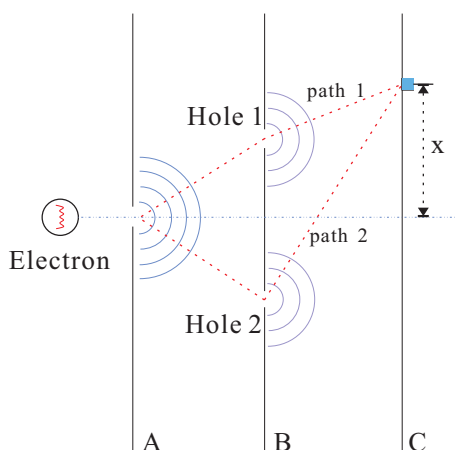


图 2: 实验装置 A 处射出的电子前进到屏 C 处的探测器, 其间插入一个带有两个孔的屏 B . 对于到达的每一个电子, 探测器记录下一个计数; 当探测器距离屏幕中心为 x 时, 测量电子到达的比例, 并以 x 进行标定, 如图3所示.

如果探测器足够灵敏, 则会发现, 到达 x 的电流不是连续的, 而是一个个电子相继到达的. 若电子源 S 的强度足够弱, 探测器将会记录下一个个的脉冲. 脉冲时间间隔内没有电子到达, 因为这个原因, 我们说电子是粒子. 如果将探测器分布在屏 C 上的各个位置, 并且电子源 S 足够微弱, 那么一次只有一个探测器响应, 过间隔时间后另一个探测器才响应. 如此不断地连续下去. 探测器绝对不会出现半个响应或者同时多个响应. 也就是说, 要么一次一个电子的完整响应, 要么什么也没有. 换句话说, 探测器记录了电子从源 S 到 x 的经历.

我们需要做的是对于不同位置 x 的探测器, 测量每秒钟脉冲的平均个数, 最后得到电子从 S 到 x 相对概率 P 作为 x 的函数. 概率 P 作为 x 函数的图像是一条复杂的曲线, 定性的画在下图, 它有极大值极小值, 在中心的附近一些区域几乎没有电子到达, 我们需要处理的物理问题就是找到这条曲线的规律.

由于电子是粒子, 我们必须做出如下假定

I 从 S 到达 x 的电子必须经过孔1或者2

II 到达 x 的概率应该是两部分之和, 也就是说通过孔1的概率 P_1 加上通过孔2的概率 P_2 .

III 存在复数 ϕ_1, ϕ_2 使得

$$P = |\phi|^2$$

$$\phi = \phi_1 + \phi_2$$

$$P_1 = |\phi_1|^2, P_2 = |\phi_2|^2$$

我们可以用实验来检验这是否正确. 每一组分的概率是容易测量的. 只需要分别遮住不同孔就能得到单个孔对应的电子概率分布. 如图3所示.

根据图中的结果, 蓝色曲线完全与两单孔对应的两条曲线叠加之和不同, 说明 $P \neq P_1 + P_2$.

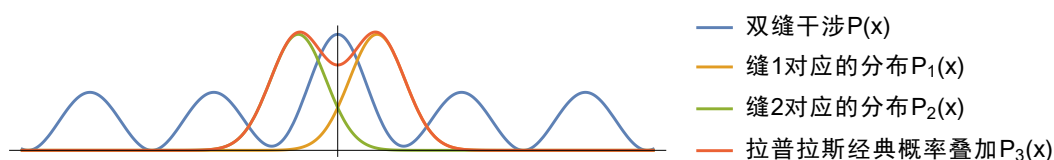


图 3: 实验结果. 电子到达 x 的概率是相对探测器位置 x 画出的. 图2的实验结果画在蓝色线中; 若仅仅只有孔1开启, 则电子仅仅能通过孔1, 其结果是绿色线所代表; 若仅仅孔2开启, 则是黄色线所代表的的结果. 如果我们设想每个电子恰好通过这个空或那个孔, 则当两孔都开启时, 预期得到的曲线应该是绿色线+黄色线=红色线. 这与实际所得的蓝色线大不相同.

2.2.2 概率幅

两孔都开启时, 电子到达 x 的概率并不是单独开启孔1的概率与单独开启孔2的概率之和. 实际上, 复杂曲线 $P(x)$ 是我们熟悉的电子通过两个孔后干涉叠加的分布强度曲线. 在波动力学里, 描述波是用复变函数 $\phi(x)$ 描述的, 而 $|\phi(x)|^2$ 正好描述了这个强度分布曲线 $P(x)$, 而 $\phi(x)$ 叫做电子到达 x 的概率幅. 这样我们就能用数学来描述 $P(x)$ 的规律. 并且 $\phi(x)$ 是两个孔对应 $\phi_1(x), \phi_2(x)$ 之和.

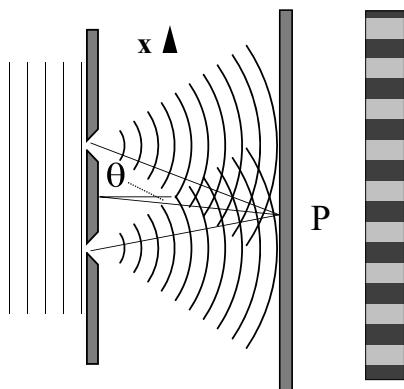


图 4: 相干干涉叠加出的电子分布曲线有极大值也有极小值, 在一些点发自两个不同孔的电子相干相消, 在另一些点上发自不同孔的两个电子相干相长, 由此产生了电子分布曲线 $I(x)$ 的条纹

这样我们就能得到III的推断.以后我们会介绍如何计算 ϕ_1, ϕ_2 . 总之, 我们能计算出到达 x 处探测器的波强度, 并且把这个强度解释为粒子到达 x 处的概率.

2.2.3 测量问题

值得注意的是, 这里我们描述电子的运动使用了波动力学里描述波用的复变函数 $\phi(x)$. 这意味着我们应该抛弃经典力学里粒子与波的概念. 在量子力学里同时使用波的概念和粒子的概念并不产生矛盾.

为了更详细地讨论这一点, 首先考虑观测结果是干涉条纹的情况. 新的概率叠加规律意味着 $P_1 + P_2 = P$ 是不正确的. 这就导致了一个结论: 当两个孔开启时, 粒子通过这个孔或那个孔的说法是不对的. 因为如果粒子通过这个孔或那个孔, 我们可以根据通过的孔将粒子分为不相干的两类, 并且到达 x 的次数 P 一定是粒子经过孔1和孔2次数之和

这与我们经典的理解完全不一样, 为了理解这点, 历史上曾经有人思考过电子也许在一条复杂的轨道上运动, 先通过孔1再以某种复杂的方式通过孔2; 或许电子以某种方式散开, 部分地通过两个孔, 导致最终的干涉结果; 或许由于遮住孔2有可能影响孔1附近的电子运动, 从而电子穿过孔1的概率没有正确的被确定, 人们曾经试图用许多这种经典力学模型去解释这个结果. 当我们使用光子时, 可以使干涉路径1和2在空间上相距几个厘米, 因此两个交叉轨道几乎肯定是独立的. 下面的实验可以证明, 实际情况比开始所设想的具有更深远的意义.

观测的后果

之前的逻辑推导得到的结果 $P \neq P_1 + P_2$. 这个结果可以说明根据电子通过孔1或者

孔2来分析电子的运动是不正确的. 然而可以设计一个实验来验证这个结论, 在孔后面放一个光源, 就可以观看电子通过哪一个孔(如图5). 由于电子散射光, 因此如果光在孔1后面被散射, 我们立马可以判定电子通过的是孔1; 若光在孔2后面被散射, 我们立马可以判定电子通过的是孔2.

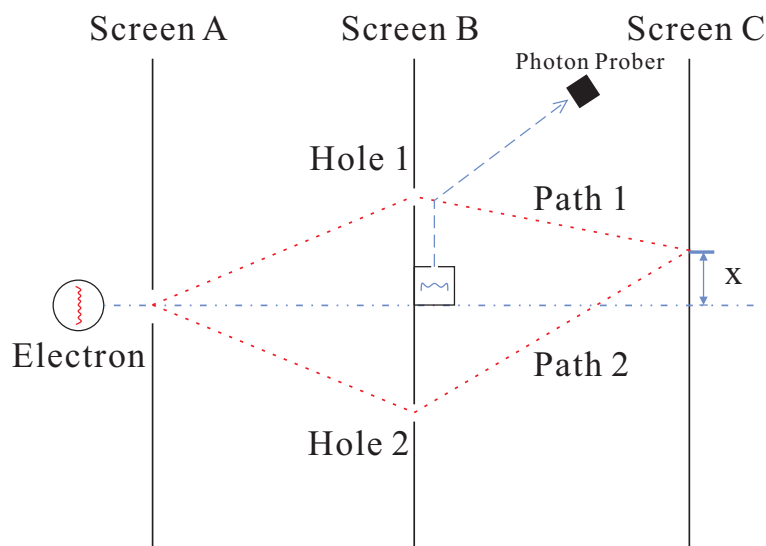


图 5: 实验2的一个修正. 我们在屏B后面放一个光源L, 并寻找通过孔1或者孔2的电子散射的光. 用一个很强的光源, 确实可以发现每个电子通过这个或那个孔. 但现在到达x的概率不再是由图3中的蓝线给出而是由图3中的红线给出.

这个实验结果表示, 电子确实通过的孔1或者孔2. 对于每一个到达屏幕C的电子, 光要么在孔1后面被散射, 要么在孔2后面被散射, 并且绝不会在两处同时被散射. (更精细的实验可以得到电子都是以一个完整电荷 e 出现的, 绝不会出现它的部分电荷.)

这似乎引起了一个荒谬的问题. 假设将两个实验联合起来查看电子通过了哪个孔, 并同时测量电子到达x的概率, 然后对于每一个到达x的电子, 我们可以从实验上来说明电子是否通过了孔1或者孔2. 首先我们知道 P_1 由图3中的橙色曲线给出, 因为我们若选择到达x的电子只是通过孔1的电子, 则我们得到的正是橙色曲线 P_1 (孔2是否遮住均可得到这个结果, 因此可以说明孔2是否开启对孔1附近电子运动毫无影响). 如果选取孔2出散射光的电子, 我们得到的则是绿色曲线 P_2 . 现在每个电子要么出现在孔1要么出现在孔2, 并且可以将它分成不相干的两类. 因此如果我们把这两类放在一起, 一定得到的是红色曲线 $P = P_1 + P_2$, 实验确实如此. 然而该分布没有出现蓝色干涉曲线!

这里发生的变化实际上是当我们观察电子通过哪个孔时, 就能得到结果 $P = P_1 + P_2$, 当我们观察时, 就得到不同的结果 $P = |\phi_1 + \phi_2|^2 \neq P_1 + P_2$.

实际上正是我们观察了电子才改变了它到达x的概率. 为了观察电子, 我们是用了光,

而光与电子的碰撞会改变电子的运动, 或者更确切的说, 改变了电子到达 x 的概率.

另一方面, 我们是否能使用较弱的光源来降低光对电子的影响呢? 事实上一个很弱甚至可以忽视的干扰对分布肯定不会产生从蓝色到红色曲线的有限小的改变. 但弱的光并不意味着弱的干扰. 光以能量 $h\nu$ 或者动量 h/λ 的量子形式出, 减弱光意味着使用较少的光子. 因此我们会漏看一些电子. 但是当我们确实看见一个电子的时候, 就意味着一个完整的光子被散射, 并且这个电子获得大约为 h/λ 的有限动量.

漏看的那些电子是根据干涉定律图3中蓝色曲线分布的, 而我们真正看到的是散射一个光子并到达 x 的那些电子, 这些电子的概率分布是图3中的红色曲线. 因此在这种情况下, 总的分布是蓝色曲线和红色曲线的加权平均. 在强的光线下, 几乎所有的电子都散射光, 分布近似红色曲线, 在弱的光线下, 散射极少, 它更相似于蓝色曲线.

我们仍然可以认为, 既然光携带的动量是 h/λ , 那么我们采用波长足够长的光就可以减弱影响效果. 但这是有限度的, 如果使用波长太长的光, 我们将不能区分电子到底从哪个孔后面被散射. 因为波长为 λ 的光源用于空间定位时, 其精度不超过 λ 的数量级.

因此我们看到, 为测定电子通过哪个孔而设计的任何物理实验, 一定会产生强的干扰, 足以使分布从蓝色曲线变到红色曲线. 海森堡第一个注意到这些, 并在他的不确定性原理中作了陈述: 当时崭新的力学自治性对所能进行的实验精度有一个限制. 在我们的情形中, 这个原理可以得出: 设计一种装置来测定电子通过什么样的孔, 并且足够精细, 使得电子的偏转不足以破坏干涉图形的企图一定不会成功. 量子力学自治性要求这必须是一个普遍性的陈述. 对于不确定性原理, 从未发现过任何例外情况. 在后面的章节里, 我们会进一步讨论不确定性原理, 量子力学的不确定性是一种本质上的不确定性, 而不是实验所导致的.

2.3 量子力学基础概念

2.3.1 量子态

我们考虑一个微观量子系统, 它由具有一些物理特征(例如质量、电荷、角动量等)的粒子组成, 这些粒子之间按照力学规律去运动, 我们把每一种符合运动规律的状态叫做一个量子态. 按照经典力学的做法, 我们如果确定了这个力学系统的坐标 q 和动量 p , 这个力学系统就确定下来了, 因为我们可以通过求解哈密顿正则方程去得到任何时刻的力学状态. 然而之前章节论证了, 对于一个量子系统, 存在不确定性, 我们无法的到这个力学系统所有的信息. 测量能力的局限性限制了一个量子态可以具有数据的数量. 这样就导致了量子系统的状态不能用某一个时刻的全部坐标与动量去确定, 而必须用比较少的数据, 或者一些具有概率性的数据来确定.

定义 2.12 (量子态的符号) 在量子力学里, 一般使用符号 $|\psi\rangle$ 表示一个量子态 ψ , 其中 ψ 是量子态的记号. 在不同的系统里, 根据不同的情况可以采用不同的量子数去标记量子态. 例如 $|\xi_1 \xi_2 \dots \xi_s\rangle$, 这里 $\xi_1, \xi_2, \dots, \xi_s$ 是 s 个独立不同的量子数, 有些地方也用记号 $|\psi_{\xi_1, \xi_2, \dots, \xi_s}\rangle$.

例 2.5 坐标为 x 的电子的量子态用符号 $|x\rangle$ 表示, 动量为 p 的电子用 $|p\rangle$ 表示.

在前面的干涉实验2里, 我们讨论了干涉测量问题. 为了简化问题, 我们这里暂不考虑电子, 现在我们考虑一个单光子干涉的实验问题

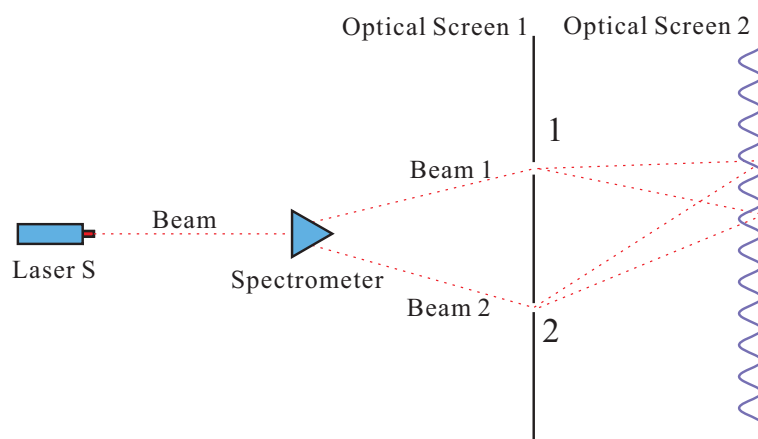


图 6: 单光子干涉实验: 激光器发射一束激光通过分光器分为两束, 束流1通过孔1, 束流2通过孔2, 然后在光屏2上发生干涉. 若减弱激光器 S , 以至於一束激光在足够长的时间内只射出一个光子, 该光子通过分光器后处于束流1的状态称作 $|1\rangle$, 处于束流2的状态称作状态 $|2\rangle$

我们把通过孔1的状态标记为 $|1\rangle$, 通过孔2的状态标记为 $|2\rangle$. 如果我们减弱激光器功率, 一直到一束激光在足够长的时间内只有一个光子射出, 很自然我们会问到: 单个光子通过分光器后会走哪条路径? 这个问题看起来很怪异, 但实际上却指向了量子力学的核心. 若我们认为光子通过分光器后走孔1或者孔2对应的路径, 类似前面电子测量问题的分析, 孔1和孔2是独立的, 这就不可避免的让我们又回到了拉普拉斯概率性叠加 $P = P_1 + P_2$. 由前面的分析, 这点与我们所需要的量子力学并不符合.

为了解决这个问题, 我们必须把光子描述成部分同时地进入由入射光分裂而成的两个组分($|1\rangle, |2\rangle$)中的每一个. 这样我们可以说光子处于一个量子态, 这个量子态是分别由两个组分对应的状态(量子态)叠加而成. 用光学语言来说就是, 每一个组分对应的量子态 $|i\rangle$ 与普通波动力学中的波函数 ψ_i 相联系. 每一个波函数可以描述一个组分的光子, 而总的波函数是由各个组分的波函数 ψ_i 叠加而成, 因此量子态 $|\cdot\rangle$ 可以像波函数一样线性叠加.

假设 2.1 (量子态的线性叠加原理) 在量子力学中量子态满足线性叠加原理

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle$$

其中 $c_n \in \mathbb{C}$.

我们暂不考虑为什么量子力学是定义在复数域 \mathbb{C} 上的, 该后续章节会解释. 量子态 $|\psi_n\rangle$ 满足线性叠加性, 这意味着 $|\psi_n\rangle$ 在数域 \mathbb{C} 上构成了线性空间 E . 我们把这个量子态 $|\psi_n\rangle$ 构成的线性空间叫做态空间 E , 态空间 E 的元素叫做态 $|\psi_n\rangle$.

定理 2.7 量子系统所有线性无关的量子态 $|\psi_n\rangle, (n = 1, 2, \dots)$ 构成一个 \mathbb{C} 上的线性空间 E , 量子力学里习惯叫做态空间. 这个系统任何时刻的量子态都处于态空间中, 并可以由不同量子态叠加得到

证明 2.16 根据量子态线性叠加的假设2.1我们可以得到

$$\begin{aligned} c_1|\phi\rangle + c_2|\phi\rangle &= (c_1 + c_2)|\phi\rangle \\ c|\phi_1\rangle + c|\phi_2\rangle &= c(|\phi_1\rangle + |\phi_2\rangle) \end{aligned}$$

若 $c_1 + c_2 = 0$, 则有 $|\phi\rangle + (-|\phi\rangle) = 0$

按照我们之前单光子实验的描述, 量子态的叠加态必须同时处于多个线性无关的组分中. 这意味着

$$\begin{aligned} |\phi_1\rangle + |\phi_2\rangle &= |\phi_2\rangle + |\phi_1\rangle \\ (|\phi_1\rangle + |\phi_2\rangle) + |\phi_3\rangle &= |\phi_1\rangle + (|\phi_2\rangle + |\phi_3\rangle) \end{aligned}$$

这样就证明了量子态构成一个线性空间, 有些地方把量子态矢量简称为态矢量.

2.3.2 左矢量

当我们建立起一个矢量集合之后, 我们总是能够通过矢量集合到数域 \mathbf{F} 之间的映射定义一个新的矢量集合, 我们叫做对偶矢量. 在物理系统里面最简单的到数域 \mathbf{F} (量子力学里 $\mathbf{F} = \mathbb{C}$)是量子态之间的内积

定义 2.13 (量子态的内积) 在态空间 E 中的两个态矢量 $|\varphi\rangle$ 与 $|\psi\rangle$ 的内积定义为:

$$\langle\varphi|\psi\rangle = \phi$$

其中 $\phi \in \mathbb{C}$, 这里 $\langle \phi|$ 称为左矢量, $|\psi\rangle$ 称为右矢量, ϕ 为复数.

在1.8章介绍了对偶空间的概念, 这里我们定义了量子态内积为 $\langle \phi|\phi\rangle \in \mathbb{C}$, 所以左矢量 $\langle \phi|$ 是一个线性映射 $L: |\phi\rangle \mapsto \mathbb{C}$. 而根据1.8章对偶空间的概念, 这意味着所有的左矢量也构成一个线性空间 E' .

定理 2.8 (态空间的对偶空间—左矢空间) 对于一个给定的量子系统, 量子态的内积诱导出左矢空间 E' , 左矢空间 E' 是右矢空间 E 的对偶空间, 其中两个态矢量的正交定义为:

$$\langle \phi|\psi\rangle = 0$$

证明 2.17 在第1.8章的对偶空间里, 我们已经证明过, 这里我们从物理的角度说明一下.

假定有一个数 ϕ , 它是右矢量 $|\psi\rangle$ 的函数. 也就是对于每一个右矢量 $|\psi\rangle$ 都有一个数 ϕ 与之对应. 我们假定这个函数是线性的, 意思是对应于 $|\psi\rangle + |\psi'\rangle$ 的数是对应于 $|\psi\rangle$ 的数与对应于 $|\psi'\rangle$ 的数之和, 而对应于 $c|\psi\rangle$ 的数等于 c 乘以对应于 $|\psi\rangle$ 的数. ($c \in \mathbb{C}$). 因此我们可以把数 ϕ 看成矢量 $|\psi\rangle$ 与某个新的矢量的标量积, 每一个右矢量 $|\psi\rangle$ 的线性函数都有一个这类新的矢量与之对应. 这个新的矢量就叫做左矢量. 我们假定对于线性函数 ϕ , $|\psi\rangle$ 对应的左矢量为 $\langle \phi|$. 它们之间的线性条件可以写成:

$$\begin{aligned}\langle \phi|(|\psi\rangle + |\psi'\rangle) &= \langle \phi|\psi\rangle + \langle \phi|\psi'\rangle \\ \langle \phi|(c|\psi\rangle) &= c\langle \phi|\psi\rangle\end{aligned}\tag{34}$$

其中 $c \in \mathbb{C}$.

性质 2.1 (左矢量的几个性质)

(1) 如果左矢量与任意右矢量的内积都完全确定下来, 那么左矢量就完全确定下来了. 特别是, 如果对所有的 $|\psi\rangle$ 都有 $\langle \phi|\psi\rangle = 0$, 则 $\langle \phi| = 0$

(2) 两个左矢量之和 $\langle \phi| + \langle \phi'|$ 定义为它与任意右矢量的内积等于 $\langle \phi|$ 和 $\langle \phi'|$ 分别与 $|\psi\rangle$ 的内积之和

$$(\langle \phi| + \langle \phi'|)|\psi\rangle = \langle \phi|\psi\rangle + \langle \phi'|\psi\rangle\tag{35}$$

(3) 左矢量 $\langle \phi|$ 与 \mathbb{C} 中的标量 c 的乘积为 $(c\langle \phi|)|\psi\rangle = c\langle \phi|\psi\rangle$

方程(34)和性质2与性质3表明, 左右矢量的乘法都满足分配律, 它们与复数域的标量 c 的乘积满足结合律.

这里所引入的左矢量是与右矢量完全不同的矢量, 到目前为止, 它们之间没有任何联系, 只是在左右矢量之间存在标量积(内积). 现在我们给定一个左右矢量之间一个新的假设, 以建立起左右矢量之间的一一对应关系.

假设 2.2 我们假定对应于 $|\psi\rangle + |\psi'\rangle$ 的左矢量等于对应于 $|\psi\rangle$ 的左矢量和对应于 $|\psi'\rangle$ 的左矢量之和, 而对应于 $c|\psi\rangle$ 的左矢量等于 c^* 乘对应于 $|\psi\rangle$ 的左矢量. 其中 c^* 是 c 的复共轭.

建立起这种对应关系后, 我们可以用相同的记号来表示与右矢量 $|\psi\rangle$ 对应的左矢量 $\langle\psi|$. 左右矢量之间的这种一一对应关系建立起来之后, 量子动力学系统在任意时刻 t 的任何状态既可以用右矢量 $|\psi(t)\rangle$ 表示也可以用左矢量 $\langle\psi(t)|$ 来表示. 事实上整个理论在左右矢量是对称的.

定理 2.9 (左右矢量与量子态的对应关系) 对应于两个独立量子态的叠加态 $c_1\phi_1 + c_2\phi_2$, 对应的右矢量为 $c_1|\phi_1\rangle + c_2|\phi_2\rangle$, 左矢量为 $c_1^*\langle\phi_1| + c_2^*\langle\phi_2|$.

$$\text{量子态 } \phi \longleftrightarrow \text{右矢量 } |\phi\rangle \longleftrightarrow \text{左矢量 } \langle\phi|$$

$$\text{量子态 } c_1\phi_1 + c_2\phi_2 \longleftrightarrow \text{右矢量 } c_1|\phi_1\rangle + c_2|\phi_2\rangle \longleftrightarrow \text{左矢量 } c_1^*\langle\phi_1| + c_2^*\langle\phi_2|$$

这样我们可以使用一个统一的符号来用左右矢量表示量子态:

$$\text{量子态 } c\phi \longmapsto |c\phi\rangle = c|\phi\rangle$$

$$\text{量子态 } c\phi \longmapsto \langle c\phi| = c^*\langle\phi|$$

给定两个量子态 ϕ 和 ψ , 我们有它们的内积 $\langle\phi|c\psi\rangle = c\langle\phi|\psi\rangle$ 与 $\langle c\psi|\phi\rangle = c^*\langle\psi|\phi\rangle$. 这使得我们进行进一步的假设:

假设 2.3 (态矢量内积复共轭) 给定量子态 ϕ 与 ψ , 我们假定它们的内积 $\langle\phi|\psi\rangle$ 与 $\langle\psi|\phi\rangle$ 之间满足复共轭

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$$

在这个假设下, 令 $\langle\phi| = \langle\psi|$, 很显然 $\langle\psi|\psi\rangle$ 是实数. 这样我们进一步假定¹除 $|\psi\rangle = 0$ 外

$$\langle\psi|\psi\rangle > 0 \tag{36}$$

在之前的干涉实验里, 概率叠加满足概率幅的线性叠加原理 $|\phi|^2 = |\psi_1 + \psi_2|^2$, 并且概率幅的模方满足正定性 $|\psi_i|^2 \geq 0$ 与三角不等式 $|\phi_1 + \phi_2| \leq |\phi_1| + |\phi_2|$. 根据量子态和右矢量之间的对应性, 再考虑到线性空间内积和范数的联系. 我们可以定义量子态的范数.

¹这是为了用内积去定义范数而做的假设.

定理 2.10 (范数) 在态空间 E 中, 任意量子态 ψ 对应的右矢量 $|\psi\rangle$ 的范数 $\| |\psi\rangle \|$ 为

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

证明 2.18 首先正定性显然满足 $\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle} > 0$

接下来证明三角不等式 $\| |\varphi\rangle \| + \| |\psi\rangle \| > \| |\varphi + \psi\rangle \|$

$$\begin{aligned} \| |\varphi\rangle + |\psi\rangle \|^2 &= (\langle \varphi | + \langle \psi |)(| \varphi \rangle + | \psi \rangle) \\ &= \langle \varphi | \varphi \rangle + \langle \varphi | \psi \rangle + \langle \psi | \varphi \rangle + \langle \psi | \psi \rangle \\ &= \langle \varphi | \varphi \rangle + 2\text{Re}(\langle \varphi | \psi \rangle) + \langle \psi | \psi \rangle \\ &< \langle \varphi | \varphi \rangle + 2|\langle \varphi | \psi \rangle| + \langle \psi | \psi \rangle \\ &< \langle \varphi | \varphi \rangle + 2\| |\varphi\rangle \| \| |\psi\rangle \| + \langle \psi | \psi \rangle \\ &= (\| |\varphi\rangle \| + \| |\psi\rangle \|)^2 \end{aligned}$$

定义 2.14 给定一个右矢量 $|\psi\rangle$, 我们可以构成一个归一的右矢量 $|\tilde{\psi}\rangle$, 记为:

$$|\tilde{\psi}\rangle = \frac{|\psi\rangle}{\langle \psi | \psi \rangle}$$

它有下列性质:

$$\langle \tilde{\psi} | \tilde{\psi} \rangle = 1$$

由于归一化, 在物理学中我们认为 $|\psi\rangle$ 和 $c|\psi\rangle$ 在 $c \neq 0$ 时表示同一个物理态. 换句话说, 态空间中只有方向是有意义的.

根据正定性假设³⁶和归一化条件^{2.14}, 很自然会做一个假设: 归一化态矢量 $|\psi\rangle$ 的自身内积 $\langle \psi | \psi \rangle$ 对应物理测量²得到的总概率¹.

实验⁶可以看出, 物理系统总是要求一个量子系统态矢量的集合 E 是完备的. 从物理的要求上来说, 态空间 E 必须是Hilbert空间 \mathcal{H} , 其中状态矢量是Hilbert空间 \mathcal{H} 的一个元素. 所有的量子力学演化过程都发生在Hilbert空间 \mathcal{H} 里, 这使得我们可以利用Hilbert空间 \mathcal{H} 来描述量子力学. 用数学的语言来说:

定理 2.11 (完备性) 若 $\{|\psi_n\rangle\} \subset \mathcal{H}$ 是一组完备正交基. 则对于任意 $|\varphi\rangle \in \mathcal{H}$.

$$\| \varphi \|^2 = \sum_n |\langle \varphi | \psi_n \rangle|^2$$

²测量问题会在算符谱理论中进行解释

该定理实质上就是Parseval关系. 证明过程在11. 在之前的章节里, 介绍了Hilbert空间里可以通过Schmidt正交化 \mathcal{H}^n 得到一组完备正交归一的基向量 $|\psi_n\rangle$. \mathcal{H} 中任意向量都可以用这组完备正交归一的基展开.

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle$$

其中 $c_k = \langle \psi_k | \psi \rangle$. 如果 \mathcal{H} 包含一组正交归一基, 则叫 \mathcal{H} 为可分的Hilbert空间.

3 算符理论

3.1 线性算符

在之前的章节里我们建立起了一套描述量子力学所需要的数学, 很自然地就会想到, 态空间中两个不同的右矢量之间有何联系. 实际上我们希望建立起两个不同量子态 $\psi \mapsto \varphi$ 之间的联系, 而这两个量子态之间的联系对应的物理实际意义正是这章我们要讨论的核心内容. 在实际物理实验中, 可以制备许多相同的量子态(纯态 $\{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_n\rangle\}$), 而在这些相同的量子态测量相同的力学量(例如角动量、动量、能量等)会得到不同的取值结果(返回值). 在经典力学里, n 个相同的状态分别测量相同力学量得到的 n 个结果必定一样, 这意味着量子力学里的力学量与经典力学里的力学量并不一样. 量子力学与经典力学之间的矛盾暗含了量子力学中的力学量应该对应着一个测量值的分布. 线性算符正好满足这个性质, 线性算符具有谱, 正好对应着量子力学的测量值分布. 线性算符的性质直接与实际物理相互对应.

3.1.1 线性算符

定义 3.1 (线性算符) \mathcal{H} 中的线性算符 \hat{A} 是指映射 $A : \mathcal{H} \mapsto \mathcal{H}$, \hat{A} 满足线性关系

$$\hat{A}(\alpha|\varphi\rangle + \beta|\psi\rangle) = \alpha\hat{A}|\varphi\rangle + \beta\hat{A}|\psi\rangle$$

其中 $|\varphi\rangle, |\psi\rangle \in \mathcal{H}, \alpha, \beta \in \mathbb{C}$.

实际上我们只需要算符 \hat{A} 定义在定义域 $\mathcal{D}(\hat{A}) \subset \mathcal{H}$, 其中定义域 $\mathcal{D}(\hat{A})$ 在 \mathcal{H} 中稠密:

$$A : \mathcal{D}(\hat{A}) \mapsto \mathcal{H}$$

下面是一些 \mathcal{H} 上线性算符的例子.

例 3.1

1 恒等映射

$$\mathbf{1} : |\psi\rangle \mapsto |\psi\rangle$$

2 坐标算符

$$\hat{x}_j : |\psi\rangle \mapsto \hat{x}_j |\psi\rangle$$

3 动量算符

$$\hat{p}_j : |\psi\rangle \mapsto \hat{p}_j |\psi\rangle$$

4 两个不同量子态 $\psi \mapsto \varphi = \hat{A}\psi$

$$\hat{A} : |\psi\rangle \mapsto |\varphi\rangle = |\hat{A}\psi\rangle = \hat{A}|\psi\rangle$$

我们是用右矢量 $|\psi\rangle$ 去标记量子态 ψ , 右矢量与量子态之间存在对应 $\psi \longleftrightarrow |\psi\rangle$, 右矢量与左矢量之间又存在线性对偶 $|\psi\rangle \xrightarrow{Dual} \langle\psi|$. 这就可以在量子态, 右矢量和左矢量之间建立起算符的对应.

$$\text{量子态: } \psi \mapsto \varphi = \hat{A}\psi$$

$$\text{右矢量: } |\psi\rangle \mapsto |\varphi\rangle = |\hat{A}\psi\rangle = \hat{A}|\psi\rangle$$

$$\text{左矢量: } \langle\psi| \mapsto \langle\varphi| = \langle\hat{A}\psi|$$

两个不同线性算符的乘积也可以利用线性映射定义

定义 3.2 (线性算符的乘积) \mathcal{H} 中两个线性算符 \hat{A}, \hat{B} 的乘积定义为

$$(\hat{A}\hat{B})|\psi\rangle = \hat{A}|\hat{B}\psi\rangle = \hat{A}(\hat{B}|\psi\rangle)$$

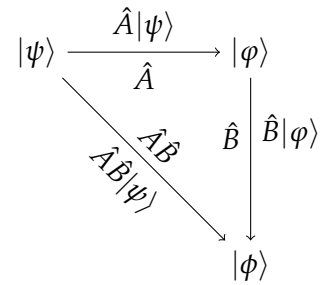
其中 $|\psi\rangle \in \mathcal{H}$.

在定理 2.11 里, 我们得到了量子力学完备性条件(Parseval公式). 这里可以进一步写成算符形式.

定理 3.1 (完备性条件) 若 $\{|\psi_n\rangle : |\psi_n\rangle \in \mathcal{H}\}$ 是态空间 \mathcal{H} 中的一组完备正交归一向量的集合, 则有完备性条件

$$\sum_n |\psi_n\rangle \langle\psi_n| = \mathbf{1}$$

其中 $\mathbf{1}$ 是单位算符 $\mathbf{1}_{n \times n}$.



证明 3.1 根据上一节的Parseval公式2.11有

$$\begin{aligned}
\| |\varphi\rangle \|^2 &= \sum_n |\langle \varphi | \psi_n \rangle|^2 \\
&= \sum_n \langle \varphi | \psi_n \rangle \langle \psi_n | \varphi \rangle \\
&= \langle \varphi | \left(\sum_n |\psi_n\rangle \langle \psi_n| \right) | \varphi \rangle \\
&= \langle \varphi | \varphi \rangle
\end{aligned}$$

所以

$$\sum_n |\psi_n\rangle \langle \psi_n| = \mathbf{1}$$

3.1.2 有界算符和逆算符

定义 3.3 (有界算符) 算符 \hat{A} 是 \mathcal{H} 上的有界算符定义为

$$\|\hat{A}\| := \sup_{\{|\psi\rangle \in \mathcal{H} \mid \|\psi\|=1\}} \|\hat{A}|\psi\rangle\| < \infty \quad (37)$$

事实上, 表达式37定义了一个范数, 它使 \mathcal{H} 上有界算符的空间 $B(\mathcal{H})$ 成为一个Banach空间. 我们将会看到在一些方面有界算符比无界算符要容易处理的多. 然而, 由于量子力学中一些最重要的算符是无界的, 我们需要研究它们的性质.

通常我们可以证明在定义域 \mathcal{D} 上稠密的算符 A 一致有界. 下面这个引理表示, 在这种情况下, A 可以推广到有界算符.

引理 3.1 如果算符 A 在定义域 $\mathcal{D} \subset \mathcal{H}$ 上稠密, 对于 ψ 满足 $\|\hat{A}|\psi\rangle\| \leq C\|\psi\|$ (C 与 ψ 无关), 则它可以扩展到 \mathcal{H} 上的有界算符, 满足相同的界:

$$\|\hat{A}|\psi\rangle\| \leq C\|\psi\|, \quad |\psi\rangle \in \mathcal{H}$$

证明 3.2 对于任意 $|\psi_n\rangle \in \mathcal{H}$, 有序列 $\{|\psi_n\rangle\} \subset \mathcal{D}$ 使得当 $n \rightarrow \infty$ 时 $|\psi_n\rangle \rightarrow |\psi\rangle$ (\mathcal{D} 的稠密性). 则关系

$$\|\hat{A}|\psi_n\rangle - \hat{A}|\psi_m\rangle\| \leq C\|\psi_n\rangle - |\psi_m\rangle\|$$

表明了 $\{\hat{A}|\psi_n\rangle\}$ 是一个Cauchy序列, 所以对于一些 $|\varphi\rangle \in \mathcal{H}$, $\hat{A}|\psi_n\rangle \rightarrow |\varphi\rangle$, 我们令 $\hat{A}|\psi\rangle := |\varphi\rangle$.

因为 $\|\hat{A}|\psi\rangle\| = \lim \|\hat{A}|\psi_n\rangle\| \leq C \lim \|\psi_n\rangle\| = C\|\psi\rangle\|$, 这使 \hat{A} 扩展到 \mathcal{H} 上的有界算符, 有相同的界 C .

对于特定重要的一类算符, 反过来说也成立: 定义在 \mathcal{H} 上的算符必须是有界的. 这样的一类算符是闭算符.

定义 3.4 (有界算符) \mathcal{H} 上有算符 \hat{A} . 若对于任意的序列 $|\psi_j\rangle_{j=1}^{\infty} \subset \mathcal{D}(\hat{A})$, 其中当 $j \rightarrow \infty$ 时, $|\psi_j\rangle \rightarrow |\psi\rangle$, $\hat{A}|\psi_j\rangle \rightarrow |\varphi\rangle$, 可以得到的 $|\psi\rangle \in \mathcal{D}(\hat{A})$ 以及 $\hat{A}|\psi\rangle = |\varphi\rangle$, 则 \hat{A} 被叫做闭算符. 换句话说, \hat{A} 的图 $\{(|\psi\rangle, \hat{A}|\psi\rangle) : |\psi\rangle \in \mathcal{D}(\hat{A})\} \subset \mathcal{H} \times \mathcal{H}$ 是封闭的.

另一类算符是对称算符.

定义 3.5 (对称算符) 若满足下列条件时, 算符 \hat{A} 是对称算符

$$\langle \varphi | \hat{A} \psi \rangle = \langle \hat{A} \varphi | \psi \rangle$$

其中 $\varphi, \psi \in \mathcal{D}(\hat{A})$.

量子力学里最重要的算符是对称算符, 特别是自伴算符, 这将会在下节介绍.

定理 3.2 定义在全体 \mathcal{H} 上的闭算子或对称算子是有界的.

证明 3.3 对于闭算符, 这是闭图象定理. 对于对称算符, 则是 *Hellinger-Toeplitz* 定理.

定义 3.6 (逆算符) 给定 *Hilbert* 空间 \mathcal{H} 上算符 \hat{A} , 满足下列条件的算符 \hat{B} 称为算符 \hat{A} 的逆算符. 如果 $\mathcal{D}(\hat{B}) = \mathcal{R}(\hat{A})$, $\mathcal{D}(\hat{A}) = \mathcal{R}(\hat{B})$,

$$\hat{B}\hat{A} = \mathbf{1}|_{\mathcal{R}(\hat{A})} \quad \hat{A}\hat{B} = \mathbf{1}|_{\mathcal{R}(\hat{B})}$$

这里

$$\mathcal{R}(\hat{A}) := \{\hat{A}|\psi\rangle : |\psi\rangle \in \mathcal{D}(\hat{A})\}$$

表示 \hat{A} 的值域

从该定义可以得出, 算符 \hat{A} 最多只能有一个逆. 算符 \hat{A} 的逆记作 \hat{A}^{-1} . 换句话说, 找到 \hat{A} 的逆等价于求解方程 $\hat{A}|\psi\rangle = |\varphi\rangle$, 其中 $|\varphi\rangle \in \mathcal{R}(\hat{A})$.

判断算符 \hat{A} 有逆的一个准则是: \hat{A} 是一一映射. 也就是说, $\hat{A}|\psi\rangle = 0 \implies |\psi\rangle = 0$, 或者等价于它是平凡核或者零空间:

$$\text{Ker}(\hat{A}) := \{|\psi\rangle \in \mathcal{D}(\hat{A}) : \hat{A}|\psi\rangle = 0\} = \{0\} \quad (38)$$

如果 \hat{A} 具有有界逆, 则称算符 \hat{A} 是可逆的. 由于有界算符是定义在 \mathcal{H} 上的, 所以除了一一映射之外, 可逆算符 \hat{A} 还必须有:

$$\mathcal{R}(\hat{A}) = \mathcal{H} \quad (39)$$

条件38和39确保了 A^{-1} 的存在, 并且 A^{-1} 定义在 \mathcal{H} 上. 事实上, 在一些重要的情况下, 它们足以确保 \hat{A} 确实是可逆的(即 A^{-1} 是有界的), 即:

1. 如果 \hat{A} 是闭算符, 根据定义3.4, A^{-1} 也是, 因此根据定理3.2, A^{-1} 是有界的.
2. 如果 \hat{A} 是对称算符, 那么根据定理3.2 A^{-1} 也是对称算符.

定理 3.3 \hat{A} 是可逆算符, \hat{B} 是有界算符, 并且满足 $\|\hat{B}\hat{A}^{-1}\| < 1$. 那么定义在 $\mathcal{D}(\hat{A} + \hat{B}) = \mathcal{D}(\hat{A})$ 算符 $\hat{A} + \hat{B}$ 是可逆的

这个关系式遵循关系式 $\hat{A} + \hat{B} = (\mathbf{1} + \hat{B}\hat{A}^{-1})\hat{A}$

3.1.3 练习题

问题 3.1 (可逆算符的乘积) 证明: 若算符 \hat{A} 和 \hat{B} 都是可逆的, 并且 \hat{B} 是闭算符, 则算符 $\hat{B}\hat{A}$ 定义在 $\mathcal{D}(\hat{B}\hat{A}) = \mathcal{D}(\hat{A})$ 并且是可逆的, 逆算符满足 $(\hat{B}\hat{A})^{-1} = \hat{A}^{-1}\hat{B}^{-1}$

问题 3.2 (Neumann级数) 证明满足 $\|\hat{A}\| < 1$ 的有界算符 \hat{A} , 级数 $\sum_{n=0}^{\infty} (-\hat{A})^n$ 是绝对收敛的, 并且它的逆算符是 $\mathbf{1} + \hat{A}$.

$$(\mathbf{1} + \hat{A})^{-1} = \sum_{n=0}^{\infty} (-\hat{A})^n$$

3.2 自伴算符(厄米算符)和投影算符

3.2.1 自伴算符

在量子力学里自伴算符扮演着核心角色. 本节专门讨论这类算符以及更广泛的对称算符. 所有的算符都假定定义在稠密的定义域 $\mathcal{D}(\hat{A})$ 上

定义 3.7 (自伴算符) Hilbert空间 \mathcal{H} 上有线性算符 \hat{A} , 若 \hat{A} 是对称的且 $\mathcal{R}(\hat{A} + i\mathbf{1}) = \mathcal{H}$, 则 \hat{A} 是自伴算符.

值得注意, 条件 $\mathcal{R}(\hat{A} + i\mathbf{1}) = \mathcal{H}$ 等价于方程

$$(\hat{A} + i\mathbf{1})\psi = f$$

对于所有的 $f \in \mathcal{H}$ 有解. 上述定义与通常用的定义不同, 但实质上是等价的. 这样定义自伴算符隐藏了我们真正需要的自伴算符性质, 并避免了与我们无关的冗长证明.

与上述定义等价的一个常用定义是利用伴算符来定义的.

定义 3.8 (伴算符) Hilbert 空间 \mathcal{H} 上的算符 \hat{A}^\dagger 满足下列关系叫做伴算符:

$$\langle \hat{A}\varphi | \psi \rangle = \langle \varphi | \hat{A}^\dagger \psi \rangle \quad (40)$$

对于所有的 $|\psi\rangle \in \mathcal{D}(\hat{A})$ 以及对于定义域

$$\mathcal{D}(\hat{A}^\dagger) := \{\varphi \in \mathcal{H} : \text{对于某些与 } |\psi\rangle \text{ 无关的常数 } C_\varphi, |\langle \varphi | \hat{A}\psi \rangle| \leq C_\varphi \|\psi\|, \forall |\psi\rangle \in \mathcal{D}(\hat{A})\}$$

中的 $|\varphi\rangle$ 都成立

之前我们定义了对称算符 3.5, 这里我们引入一种特殊的对称算符, 叫做自伴算符.

定义 3.9 若 $\hat{A}^\dagger = \hat{A}$, 则算符 \hat{A} 是自伴算符

根据自伴算符的定义, 每一个自伴算符都是对称算符. 事实上确实如此, 算符 \hat{A} 是自伴算符当且仅当它是对称的且 $\mathcal{D}(\hat{A}) = \mathcal{D}(\hat{A}^\dagger)$. 但是不是所有的对称算符都是自伴算符, 每一个自伴算符也不能唯一的扩展到更大的自伴定义域, 后面的章节里会接触这类问题.

接下来我们会说明有界对称算符是自伴算符. 为了说明这一点, 我们先引入一个引理.

引理 3.2 设 \hat{A} 是对称算符. 若存在一些 $z (\text{Im} z > 0)$ 有 $\mathcal{R}(\hat{A} - z\mathbf{1}) = \mathcal{H}$, 则对于 $\text{Im} z > 0$ 的每一个 z 都成立. 同样的, 对于 $\text{Im} z < 0$ 也有相同的结论. 进一步说, 若 \hat{A} 是自伴算符, 则对于每个 $\text{Im} z \neq 0$ 的 z , $\hat{A} - z\mathbf{1}$ 是可逆的, 并且满足估计

$$\|(\hat{A} - z\mathbf{1})^{-1}\| \leq \frac{1}{|\text{Im} z|} \quad (41)$$

证明 3.4 记 $z = \lambda + i\mu$ $\lambda, \mu \in \mathbb{R}$. 由于 \hat{A} 是对称算符, 我们有

$$\begin{aligned} \|(\hat{A} - z\mathbf{1})|\psi\rangle\|^2 &= \langle (\hat{A} - z\mathbf{1})\psi | (\hat{A} - z\mathbf{1})\psi \rangle \\ &= \langle (\hat{A} - (\lambda + i\mu)\mathbf{1})\psi | (\hat{A} - (\lambda + i\mu)\mathbf{1})\psi \rangle \\ &= \langle (\hat{A} - \lambda\mathbf{1})\psi | (\hat{A} - \lambda\mathbf{1})\psi \rangle - i\mu \langle (\hat{A} - \lambda\mathbf{1})\psi | \psi \rangle + i\mu \langle \psi | (\hat{A} - \lambda\mathbf{1})\psi \rangle + \mu^2 \langle \psi | \psi \rangle \\ &= \|(\hat{A} - \lambda\mathbf{1})|\psi\rangle\|^2 + \|\mu\psi\|^2 \geq |\mu|^2 \|\psi\|^2 \end{aligned} \quad (42)$$

因此, 若 $\text{Im}z > 0$ (或者 $\text{Im}z < 0$), 则 $\text{Ker}(A - z) = 0$. 若 $\mathcal{R}(\hat{A} - z\mathbf{1}) = \mathcal{H}$, 则 $(\hat{A} - z)^{-1}$ 定义在全体 \mathcal{H} . 进一步说, 式子(42)意味着 $(\hat{A} - z\mathbf{1})^{-1}$ 是有界的, 界由(41)确定(如果定义 $|\phi\rangle := (\hat{A} - z\mathbf{1})|\psi\rangle$), 所以特别是 $(\hat{A} - z\mathbf{1})$ 是可逆的. 那么根据定理3.3, 对于 $|z' - z| < \|(\hat{A} - z\mathbf{1})^{-1}\|^{-1}$, $\hat{A} - z'\mathbf{1} = (\hat{A} - z\mathbf{1}) + (z\mathbf{1} - z'\mathbf{1})$ 是可逆的. 因此如果 $|z' - z| < |\text{Im}z|$, $A - z'\mathbf{1}$ 是可逆的, 所以我们可以拓展 $\hat{A} - z\mathbf{1}$ 的可逆性到全体 $\text{Im}z' > 0$ (或者 $\text{Im}z' < 0$) 上. 若 \hat{A} 是自伴的, 则 $\mathcal{R}(\hat{A} \pm i = \mathcal{H})$, 因此 $\hat{A} - z\mathbf{1}$ 是可逆的且对于 $z \in \mathbb{C}$ 满足41.

这个引理表示了若 \hat{A} 是自伴的, 则对于任意 $\alpha \neq 0$ 的实数 α, β , $\alpha\hat{A} + \beta$ 也是自伴的以及:

$$\hat{A} \text{ 是自伴算符} \Rightarrow \text{对于 } \forall \text{Im}z \neq 0, (\hat{A} - z\mathbf{1})|\psi\rangle = |\phi\rangle \text{ 有唯一的解}$$

下面这个定理显示了对于有界算符, 自伴性很容易检验.

定理 3.4 若 \hat{A} 是有界对称算符, 则 \hat{A} 是自伴算符.

证明 3.5 根据引理3.2, 足以表明 $\mathcal{R}(\hat{A} + i\lambda\mathbf{1})$ 提供的 $|\lambda|$ 足够大. 这等价于对于 $|\phi\rangle \in \mathcal{H}$ 和这个 λ 求解方程

$$(\hat{A} + i\lambda\mathbf{1})|\psi\rangle = |\phi\rangle$$

两边除以 $i\lambda$ 得到

$$(\mathbf{1} + \hat{K}(\lambda))|\psi\rangle = |\phi\rangle$$

其中 $\hat{K}(\lambda) = (i\lambda)^{-1}\hat{A}$, $|\phi\rangle = (i\lambda)^{-1}|\phi\rangle$. 令 $|\lambda| > \|V\|$. 则 $\|\hat{K}(\lambda)\| = \frac{1}{\lambda}\|\hat{A}\| < 1$, 我们也推断出 $(\mathbf{1} + \hat{K}(\lambda))$ 是可逆的(习题3.2).

3.2.2 投影算符

定义 3.10 设 \mathcal{H} 是 Hilbert 空间. 如果有界算符 $\hat{P} \in \mathcal{H}$ 满足下面条件, 则叫做投影算符

$$\hat{P}^2 = \hat{P}$$

这个关系意味着 $\|\hat{P}\| \leq \|\hat{P}\|^2$, 所以 $\|\hat{P}\| \geq 1$ 规定了 $\hat{P} \neq 0$. 我们有

$$|\phi\rangle \in \mathcal{R}(\hat{P}) \iff \hat{P}|\phi\rangle = |\phi\rangle \quad \text{和} \quad |\phi\rangle \in (\mathcal{R}(\hat{P}))^\perp \iff \hat{P}^\dagger|\phi\rangle = 0 \quad (43)$$

确实, 如果 $|\phi\rangle \in \mathcal{R}(\hat{P})$, 则存在 $|\psi\rangle \in \mathcal{H}$ 使得 $|\phi\rangle = \hat{P}|\psi\rangle$, 所以 $\hat{P}|\phi\rangle = \hat{P}^2|\psi\rangle = \hat{P}|\psi\rangle = |\phi\rangle$.

例 3.2 (几个投影算符的例子) (1)

令 $\mathcal{H} = L^2(\mathbb{R}^n)$, E 是 \mathbb{R}^n 的子集. 则

$$\chi_{x \in E} : f(x) \mapsto \chi_E(x)f(x)$$

其中

$$\chi_E(x) := \begin{cases} 1 & x \in E; \\ 0 & x \notin E. \end{cases}$$

是投影

2. \mathcal{H} 是任意的 Hilbert 空间, 且 $\varphi, \psi \in \mathcal{H}$ 满足 $\langle \varphi, \psi \rangle = 1$. 则

$$\phi \mapsto \langle \varphi, \phi \rangle \psi$$

是一个投影

3. 如 2 里一样, 但现在有一个正交归一的基组³ $\{\psi_i\}_{i=1}^N$. 则有

$$\psi \mapsto \sum_i^N \langle \psi_i, \psi \rangle \psi_i$$

是一个投影

4. 如上例, 只是此时 Hilbert 空间是 \mathcal{H} 是态空间, 同样的设有一组正交归一的基组 $\{|\psi_k\rangle\}$. 则

$$|\psi\rangle \mapsto \sum_k \langle \psi_k | \psi \rangle |\psi_k\rangle \quad (44)$$

是一个投影.

由于 $\langle \psi_k | \psi \rangle \in \mathbb{C}$, 式子(44)可以写成

$$|\psi\rangle \mapsto \sum_k |\psi_k\rangle \langle \psi_k | \psi \rangle$$

若将 $|\psi_k\rangle \langle \psi_k|$ 看成算符映射

$$|\psi_k\rangle \langle \psi_k| : |\psi\rangle \mapsto |\psi_k\rangle$$

³ $\langle \psi_i, \psi_j \rangle = \delta_{ij}$

这表明了算符 $|\psi_k\rangle\langle\psi_k|$ 是一个投影算符 $\hat{P} = |\psi_k\rangle\langle\psi_k|$

定义 3.11 1. 如果投影算符 \hat{P} 有 $\dim \mathcal{R}(\hat{P}) = r$, 则称其秩 $rank\ r < \infty$

2. 如果投影算符 \hat{P} 是自伴的, 则被称为正交投影算符. 即 $\hat{P} = \hat{P}^\dagger$.

若 \hat{P} 是正交投影算子, 则式子(43)意味着

$$|\psi\rangle \perp \mathcal{R}(\hat{P}) \iff \hat{P}|\psi\rangle = 0$$

即: $\text{Ker}(\hat{P}) = \mathcal{R}(\hat{P})^\perp$ 上面的四个例子中除了第二个例子不是正交投影, 其余三个例子都是正交投影⁴. 而第二个例子是正交的当且仅当 $\varphi = \psi$

3.2.3 练习题

问题 3.3 证明方程40定义了 $D(\hat{A}^\dagger)$ 上唯一的线性算符

问题 3.4 证明自伴算符两种定义3.7与3.8是等价的.

提示: 利用

引理 3.3 若 \hat{A} 是有界对称算符, 则 \hat{A} 是自伴算符.

问题 3.5 证明式子(43).

问题 3.6 证明:

1. 当且仅当 $|\psi\rangle \perp \mathcal{R}(\hat{P})$ 时, $\hat{P}^\dagger|\psi\rangle = 0$
2. $\mathcal{R}(\hat{P})$ 是闭的
3. \hat{P}^\dagger 也是投影算符

问题 3.7 设 \hat{P} 是正交投影算符. 证明

1. $\|\hat{P}\| \leq 1$, 所以若 $\hat{P} \neq 0$, 则 $\|\hat{P}\| = 1$;
2. $1 - \hat{P}$ 也是正交算符, 并且 $\mathcal{R}(1 - \hat{P}) \perp \mathcal{R}(\hat{P})$, $\text{Ker}(1 - \hat{P}) = \mathcal{R}(\hat{P})$;
3. $\mathcal{H} = \mathcal{R}(\hat{P}) \oplus \text{Ker } \hat{P}$

⁴特别应该注意 $|\psi_k\rangle\langle\psi_k|$ 是正交投影算符

3.3 谱理论

迄今为止我们只在抽象的Hilbert空间中定义了算符, 而我们想要找到算符与实际物理对应的量是什么. 我们希望某些算符对应了某些物理量, 这些物理量正好是实验上可以观测的. 我们观测到的结果总是一个以某种概率分布出现的数值. 同时我们看到算符是具有谱结构的, 我们希望算符的谱能对应实际物理的结果. 这驱动我们去学习算符的一个重要特征——谱

3.3.1 算符的谱

3.3.2 算符函数与谱映射定理

3.4 可观测力学量

迄今为止我们已经基本建立起了Hilbert空间上右矢量、左矢量、线性算符. 很自然的会问到, 这样定义的数学对应的物理是什么. 在上一节里, 自伴算符满足 $\hat{A} = \hat{A}^\dagger$, 也就是 $\langle \hat{A}\varphi | \psi \rangle = \langle \varphi | \hat{A}^\dagger \psi \rangle = \langle \varphi | \hat{A} \psi \rangle$. 我们在上一章构建了右矢量的对偶, 也就是左矢量. 由于左矢量和右矢量之间存在一一映射, 左矢量空间实质上与右矢量空间同构. 我们所希望构建的一个新理论(量子力学)应该在左右矢量之间是对称的, 也就是左右矢量满足的运动方程实质上应该是一样的, 力学量在左右矢量之间应该是对称的. 这不得不使我们猜想, 量子力学里的力学量可能是自伴算符.

定义 3.12 具有可观测效应的物理量叫做力学量

所谓可观测效应是指在实验测量过程中总是能够被观测到的量. 比如动量 p , 角动量 J , 自旋⁵ s . 所谓不可观测的物理量是指那些仅仅具有数学意义, 而无论设计多么巧妙的实验都无法观测的量⁶.

假设 3.1 力学量是自伴算符, 用符号 \hat{A} 表示, 满足

$$\langle \hat{A}\psi | \psi \rangle = \langle \psi | \hat{A}^\dagger \psi \rangle = \langle \psi | \hat{A} \psi \rangle$$

即 $\hat{A} = \hat{A}^\dagger$.

由于这个假设, 我们通常在写力学量与态矢量乘积的时候并不区分力学量 \hat{A} 在左矢量上还是在右矢量上. 统一可以记作 $\langle \psi | \hat{A} | \psi \rangle$, 力学量 \hat{A} 可以向后乘 $\langle \psi | \vec{\hat{A}} | \psi \rangle$ 也可以向左乘 $\langle \psi | \overleftarrow{\hat{A}} | \psi \rangle$.

⁵粒子自旋是一个没有经典力学量对应的力学量, 但是我们可以通过观测粒子的磁矩得到自旋值

⁶例如么正么模算符 \hat{U} , 宇称算符 \hat{P} , 时间反演算符 \hat{T} 等等, 都不可观测.

4 量子化

4.1 基本物理学量

4.2 表象变换与本征态

4.3 李乘法

4.4 基本量子化条件

4.5 不确定关系

5 量子动力学

5.1 薛定谔方程

5.1.1 能量本征方程

5.1.2 概率流守恒

5.1.3 一维well问题举例

5.2 单参数么模群——时间演化算符

5.3 传播子与格林函数

5.4 动力学图像——薛定谔图像、海森堡图像、狄拉克图像

6 量子力学对称性

6.1 平移对称性与动量

6.2 转动对称性与角动量

6.3 李群李代数及其表示

6.4 角动量及其耦合

6.5 不可约张量算符

7 三维束缚态

7.1 渐进分析

7.2 无限深球well

7.3 有限深球well

7.4 三维谐振子

7.5 氢原子严格解

8 三维散射态