量子力学Note

李成蹊[1]

起草于2018-12-13

[1]✉ 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, 请不要外传, 版权所属, 违者必究.

李成蹊 中国科学院物理研究所博士在读, 主要方向是超冷原子. 目前研究兴趣人造规范场、拓扑绝缘体、SOC、超流体量子化涡旋、光晶格. 本Note主要是闲暇时间所写, 参考了一部分数学物理教材, 有些是直接挪用. Note中若有何处不正确, 欢迎指出.联系方式:QQ243114781, tlwyxzcx@sina.cn

感谢: 感谢高等数学学习群(234674270)里的网友指出了大量的错误. 特别感谢”光神”(461241165)仔细阅读后对本note指出了大量错误.

作者一些想说的话

本来刚开始只是打算简单写一下, 没想到写到后来越写越多. 其实量子力学一点也不难, 只要认真去学习. 我读过的第一本量子力学教材是曾谨言的《量子力学(第四版)》, 时至今日我清晰地记得扉页上写的第一段前言: “**他山之石可以攻玉**.” 或许是我智慧尚未修炼到位, 我读了三遍, 所有的习题我都破解掉了, 但我感觉自己对量子力学的理解依旧没有足够深刻. 直到我接触的第二本量子力学教材狄拉克写《量子力学原理》, 我才发现量子化的核心就在于从经典的力学量Poison括号过渡到量子力学里的对易括号. 实际上我们引入辛流形后会发现从经典力学过渡到量子力学是非常自然地一件事.

读过的第三本书是顾莱纳著的《量子力学:对称性》, 这本书从量子力学入手讲述对称性与守恒量的关系, 在这本书里我才明白了经典力学里的诺特原理怎么对应到量子力学版本的. 这本书大量篇幅讲述的李群表示, 特别是对角动量的讲述上非常清晰, 这点对于作者后面学习量子场论时起到了非常重要的帮助. 作者读过的第四本书是樱井纯先生写的《现代量子力学》. 时至今日我还记得当时上微分几何课程时, 杨老师调笑我拿了一本现代量子力学, 问我现代在哪里? 可惜我资质愚钝, 直到现在我也没弄明白什么是“现代”, 更可惜杨老师后来身体不适, 住院休假, 希望一切都会好起来吧.

图 1: 士郎, 我会一直爱着你

起草于:
北京市保福寺桥南
中国科学院物理研究所
桃李不言下自成蹊

目录

I 非相对论量子力学 1

1 线性代数 3

1.1 简介 . . . 3
1.2 向量空间和子空间 . . . 3
1.3 赋范空间 . . . 9
1.4 Banach空间 . . . 15
1.5 线性映射 . . . 18
1.6 线性空间的同构 . . . 23
1.7 线性变换 . . . 24
1.8 Hilbert空间 . . . 26
1.9 正交归一系统 . . . 36

2 基本物理图像 47

2.1 经典力学 . . . 47
2.1.1 基本力学概念 . . . 47
2.1.2 运动积分与诺特定理 . . . 50
2.1.3 哈密顿力学 . . . 56
2.1.4 泊松括号 . . . 58
2.1.5 可积系统 . . . 64
2.1.6 哈密顿雅克比方程 . . . 66
2.2 量子力学中的概率 . . . 68
2.2.1 双缝干涉 . . . 69
2.2.2 概率幅 . . . 70
2.2.3 测量问题 . . . 71
2.3 量子力学基础概念 . . . 73

2.3.1 量子态 . . . 73
2.3.2 左矢量 . . . 75

3 算符理论 81
3.1 线性算符 . . . 81
3.1.1 线性算符 . . . 81
3.1.2 有界算符和逆算符 . . . 83
3.1.3 练习题 . . . 85
3.2 自伴算符(厄米算符)和投影算符 . . . 86
3.2.1 自伴算符 . . . 86
3.2.2 投影算符 . . . 89
3.2.3 练习题 . . . 91
3.3 谱理论 . . . 91
3.3.1 算符的谱 . . . 92
3.3.2 算符函数与谱映射定理 . . . 93
3.3.3 习题 . . . 97
3.4 可观测力学量 . . . 97
3.4.1 力学量假设 . . . 97
3.4.2 测量假设 . . . 99
3.4.3 统计学解释 . . . 102
3.4.4 力学量的相容性 . . . 104
3.4.5 不确定关系 . . . 107
3.4.6 习题 . . . 109
3.5 算符对易关系 . . . 109
3.5.1 习题 . . . 113

4 量子化与表象 115
4.1 基本量子化条件 . . . 115
4.2 一次量子化 . . . 118
4.2.1 习题 . . . 120
4.3 表象理论 . . . 120
4.3.1 量子态的表象 . . . 121
4.3.2 算符的表象 . . . 124
4.3.3 表象变换 . . . 125

4.3.4 坐标表象 . . . 130
4.3.5 动量表象 . . . 133

5 量子动力学 135
5.1 薛定谔方程 . . . 135
5.1.1 能量本征方程 . . . 135
5.1.2 概率流守恒 . . . 135
5.1.3 一维well问题举例 . . . 135
5.2 单参数幺模群——时间演化算符 . . . 135
5.3 传播子与格林函数 . . . 135
5.4 动力学图像——薛定谔图像、海森堡图像、狄拉克图像 . . . 135

6 量子力学对称性 137
6.1 平移对称性与动量 . . . 137
6.2 转动对称性与角动量 . . . 137
6.3 李群李代数及其表示 . . . 137
6.4 角动量及其耦合 . . . 137
6.5 不可约张量算符 . . . 137

7 三维束缚态 139
7.1 渐进分析 . . . 139
7.2 无限深球well . . . 139
7.3 有限深球well . . . 139
7.4 三维谐振子 . . . 139
7.5 氢原子严格解 . . . 139

8 三维散射态 141
8.1 介绍 . . . 141
8.1.1 经典散射与微分截面 . . . 141
8.1.2 量子散射波 . . . 143
8.2 含时散射 . . . 145
8.2.1 入射出射态的近似 . . . 145
8.2.2 Moller算子 . . . 146
8.2.3 Moller算子的性质 . . . 148
8.2.4 S矩阵 . . . 149

8.2.5 散射振幅 . . . 151
8.2.6 量子散射截面 . . . 152
8.2.7 量子散射截面计算 . . . 154
8.2.8 光学定理 . . . 157
8.3 无自旋粒子两体散射 . . . 158
8.3.1 两体波函数 . . . 159
8.3.2 两体S矩阵 . . . 161
8.3.3 T矩阵 . . . 163
8.3.4 各种参考系下的截面 . . . 164
8.4 自旋粒子两体散射 . . . 165
8.4.1 自旋粒子的Hilbert空间 . . . 165
8.4.2 自旋粒子的S算子 . . . 165
8.4.3 散射振幅和散射矩阵 . . . 165
8.4.4 自旋求和 . . . 165
8.4.5 旋量的入射出射 . . . 165
8.5 Dyson公式 . . . 165
8.6 Born近似 . . . 165

9 多粒子量子系统 167
9.1 全同粒子与交换对称性 . . . 167
9.2 置换群与杨图、杨算子 . . . 167
9.3 二次量子化 . . . 167
9.4 电磁场正则量子化 . . . 167
9.5 Casmir效应 . . . 167

II 相对论量子力学 169

10 连续场基本方程 171
10.1 欧拉-拉格朗日方程 . . . 171
10.2 诺特原理与诺特流 . . . 171
10.3 克莱因-高登方程 . . . 171
10.4 克莱因-高登方程正则量子化 . . . 171

11 洛伦兹群 **173**
11.1 洛伦兹变换 . . . 173
11.2 Clifford代数的表示——γ矩阵 . . . 173
11.3 狄拉克方程 . . . 173
11.4 狄拉克方程正则量子化 . . . 173
11.5 洛伦兹群的分离对称性与电荷对称性——$\mathcal{CPT}$联合不变性 . . . 173

12 微扰理论 **175**
12.1 Gellman-Low公式与真空 . . . 175
12.2 Wick定理 . . . 175
12.3 Feynman图与费曼规则 . . . 175
12.3.1 ϕ^4的费曼规则 . . . 175
12.3.2 QED的费曼规则 . . . 175
12.4 截面与S矩阵元 . . . 175

13 树图计算 **177**
13.1 $e^+e^- \to \mu^+\mu^-$ . . . 177
13.2 康普顿散射 . . . 177
13.3 库仑散射 . . . 177

14 圈图计算 **179**
14.1 顶角近似 . . . 179
14.2 辐射修正,正规化方案——PV方案 . . . 179
14.3 场强重整化 . . . 179
14.4 LSZ约化定理 . . . 179
14.5 光学定理 . . . 179
14.6 Ward-Takahashi恒等式 . . . 179
14.7 电荷重整化 . . . 179

III 规范理论 **181**

15 $U(1)$规范理论 **183**

16 $SU(n)$规范理论 **185**
16.1 规范理论的内禀空间与底流形 . . . 185

16.2 规范自由与纤维丛 . . . 185
16.3 规范固定与截面 . . . 185
16.4 Yang-Mills方程 . . . 185
16.5 Wilson圈 . . . 185

17 路径积分量子化 187
17.1 量子力学的路径积分 . . . 187
17.2 标量场的量子化 . . . 187
17.3 QED的量子化 . . . 187
17.4 旋量场的量子化 . . . 187

18 $Non-Abel$规范场量子化 189
18.1 Faddeev-Poppov路径积分量子化 . . . 189
18.2 鬼 . . . 189
18.3 BRST对称性 . . . 189
18.4 $Non-Abel$场的单圈修正 . . . 189
18.5 渐进自由 . . . 189

19 量子色动力学 191
19.1 夸克理论 . . . 191
19.2 胶子辐射 . . . 191
19.3 强子散射:轻子对 . . . 191

20 微扰理论 193
20.1 二维手征流 . . . 193
20.2 四维手征流 . . . 193
20.3 Goldstone定理 . . . 193
20.4 手征流反常 . . . 193

21 自发对称性破缺 195
21.1 Higgs机制 . . . 195
21.2 弱相互作用的Glashow-Weinberg-Salam理论 . . . 195

Part I

非相对论量子力学

Chapter 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**Q**、实数域**R**、复数域**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 1.2.1 (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$k,l\in\mathbf{F}$,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 1.2.2 (向量空间的基) 若向量空间E中, 有n个向量$a_1,a_2,a_3,\cdots,a_n$线性无关,而E中任意$n+1$个向量线性相关, 则把这$a_1,a_2,\cdots,a_n$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:$dimE=n$

向量空间的子空间

定义 1.2.3 (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$\mathbb{W}=\{\begin{pmatrix}a&b\\0&0\end{pmatrix}|a,b\in\mathbf{F}\}$是$F^{2\times2}$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 1.2.4 (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$E_1 \bigcap E_2$也是都是子空间

定义 1.2.5 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 1.2.6 (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 1.2.1 (函数空间) 若X是一个非空集合, E是一个向量空间.记F是$X \longmapsto E$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 1.2.2 Ω是R^N的一个开子集, 所有$\Omega \longmapsto \mathbb{C}$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 1.2.3 (序列空间) 如果在1.2.1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 1.2.4 (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 1.2.1 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$\sum_{n=1}^{\infty} |x_n + y_n| < \infty$则可以用*Minkowski*不等式得到:

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

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 1.2.1 (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$(y_n) \in l^p$, 那么有:

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

证明 1.2.2 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

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \tag{1.1}$$

证明1.1用了假设条件$a^p \leq b^q$.同理可以证明1.1同样对于$a^p \geq b^q$也成立. 对于1.1我们假设:

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

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$n \to \infty$得到*Hölder*不等式

定理 1.2.2 (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 1.2.3 对于$p = 1$很显然退化成三角不等式. 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:

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

因为$q(p-1)=p$,所以有:

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

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 1.2.7 (向量空间的笛卡尔乘积) 设$E_1,\ldots,E_n$是在相同数域**F**上的向量空间, 定义:

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

并且有:

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

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

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

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

$$x=\alpha_1x_1+\cdots+\alpha_kx_k$$

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

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

的线性组合.

类似的, k阶多项式是$1,x,x^2,\ldots,x^k$的线性组合

定义 1.2.8 (线性无关) 如果$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$当且仅当$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, 那么矢量$x_1, x_2, \cdots, x_k$线性无关

1.3 赋范空间

范数

定义 1.3.1 (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 1.3.1 函数

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

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:

$$\|z\| = |z_1| + |z_2| + \cdots + |z_n|$$
$$\|z\| = max|z_1|, \ldots, |z_n|$$

例 1.3.2 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 1.3.3 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 1.3.2 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, 1.3.2、1.3.3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R和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 1.3.3 *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$\lim_{n\to\infty} x_n = x$或者简单写作$x_n \to x$*

赋范空间上的收敛有和$\mathbb{R}$上收敛的相同性质:

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

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 1.3.4 *考虑定义在有界闭集$\Omega \subset \mathbb{R}^n$上的所有连续函数空间$\mathcal{C}(\Omega)$. 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$f_1, f_2, \ldots, f_n \in \mathcal{C}(\Omega)$一致收敛于$f$. 显然1.3.2中的范数定义了序列$f_n$一致收敛于$f$, 当且仅当$\|f_n - f\| = \max_{x\in\Omega} |f_n(x) - f(x)| \to 0 (n \to 0)$. 这个范数叫做一致收敛范数*

范数的等价

定义 1.3.4 *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,

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

那么被称为等价的.

例 1.3.5 下列定义在R^2上的范数是等价的

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

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 1.3.1 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

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

证明 1.3.1 假设两个范数$\|\cdot\|_1$和$\|\cdot\|_2$等价, 也就是说$\|x_n - x\|_1 \to 0$当且仅当$\|x_n - x\|_2 \to 0$. 假设不存在$\alpha > 0$使得对于任意$x \in E$都有$\alpha\|x\|_1 \le \|x\|_2$. 那么对于任意$n \in \mathbb{N}$都存在$x_n \in E$使得

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

定义

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

显然$\|y_n\|_2 = \frac{1}{\sqrt{n}} \to 0$. 另一方面, $\|y_n\|_1 \ge n\|y_n\|_2 \ge \sqrt{n}$. 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$\beta$的存在.

引理 1.3.1 如果$x_1, \ldots, x_n$是赋范空间中线性无关的元素, 则存在常数$c > 0$使得

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \ge c(|\alpha_1| + \cdots + |\alpha_n|) \tag{1.2}$$

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

证明 1.3.2 当$|\alpha_1| + \cdots + |\alpha_n| = 0$时,(1.2)对与任意$c$都成立. 所以(1.2)等价于

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \ge c, \quad |\beta_1| + \cdots + |\beta_n| = 1$$

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

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

因为集合$B = (\beta_1,\ldots,\beta_n) \in \mathbb{R}^n : |\beta_1| + \cdots + |\beta_n| = 1$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$\beta_1 x_1 + \cdots + \beta_n x_n = 0, |\beta_1| + \cdots + |\beta_n| = 1$与$x_1,\ldots,x_n$线性无关矛盾.因此

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

证明完毕

定理 1.3.2 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 1.3.3 E是有限维向量空间,$\{e_1,\ldots,e_n\}$是E的基.定义$\|\cdot\|_0$范数如下:

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

利用*1.3.1*

$$\begin{aligned}\|\alpha_1 e_1 + \cdots + \alpha_n e_n\| &\le |\alpha_1|\|e_1\| + \cdots + |\alpha_n|\|e_n\| \\ &\le \max\|e_1\|,\ldots,\|e_n\|(|\alpha_1| + \cdots + |\alpha_n|)\end{aligned}$$

所以对于任意$x \in E$都有

$$\|x\| \le \beta\|x\|_0$$

其中$\beta = \max\|e_1\|,\ldots,\|e_n\|$. 由于引理*1.3.1*立即可以得到, 对于任意$x \in E$

$$\alpha\|x\|_0 \le \|x\|$$

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 1.3.5 *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 1.3.6 *下面是几个R^2中球的例子.*

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

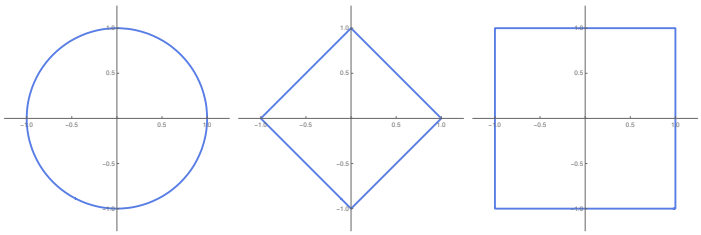


图 1.1: $\mathbb{R}^2$中球的例子, $\|(x,y)\|_1 \leq 1, \|(x,y)\|_2 \leq 1, \|(x,y)\|_3 \leq 1$

开集和闭集

定义 1.3.6 *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 1.3.7 *开球是开集. 闭球和球面是闭集.*

定理 1.3.3

(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 1.3.4 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.

$$x_1, x_2, \ldots, x_n \in S, \quad x_n \to x, \quad x \in S$$

证明 1.3.4 假设S是E的闭子集, $x_1, x_2, \cdots \in S$, $x_n \to x$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$\|x - x_n\| \to 0$, 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$x_1, x_2, \cdots \in S$, 并且有$x_n \to x$,$x \in S$. 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$x_1, x_2, \cdots \in S$使得$x_n \in B(x, \frac{1}{n})$. 但另一方面$x_n \to x$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 1.3.7 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 1.3.5 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.

$$clS = \{x \in E : \exists x_1, x_2, \cdots \in S s.t. x_n \to x\}$$

读者自证

稠密

定义 1.3.8 赋范空间E有子集S, 如果$clS = E$, 则S被称为稠密的,

例 1.3.8 $[a, b]$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 1.3.6 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$x_1, x_2, \cdots \in S$,使得$x_n \to x$

(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 1.3.9 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 1.3.9 R^n, C^n中的有界闭集紧致

定理 1.3.7 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 1.3.10 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 1.3.8 紧致集合是有界闭的.

定理 1.3.9 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 1.4.1 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 1.4.1 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$\|x_{p_n} - x_{q_n}\| \to 0, n \to \infty$成立

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

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

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

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 1.4.1 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$n = 1, 2, \ldots$, (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 1.4.1 若(x_n)是赋范空间的一个*Cauchy*序列, 则序列$(\|x_n\|)$收敛.

证明 1.4.1 因为$|\|x\| - \|y\|| \leq \|x - y\|$, 所以有$|\|x_m\| - \|x_n\|| \leq \|x_m - x_n\| \to 0, (m, n \to \infty)$. 范数序列是实柯西序列, 因此它收敛.

Banach空间

定义 1.4.2 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 1.4.2 l^2是完备空间. 设(a_n)是l^2的柯西序列, 如果

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

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$mn, \geq n_0$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

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

进一步证明a是l^2的元素, 序列(a_n)收敛于a. 根据(1.4.2), 令$m \to \infty$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

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

序列(a_n)收敛于a,且$a \in l^2$

收敛与绝对收敛

定义 1.4.3 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$n \to \infty$时$\|x_1 + x_2 + \cdots + x_n - x\| \to 0$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 1.4.2 *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 1.5.1 如果映射$L : E_1 \mapsto E_2$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 1.5.2 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$\|x_n - x_0\| \to 0$意味着$\|F(x_n) - F(x_0)\| \to 0$, 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 1.5.1 赋范空间E的范数是$E \mapsto \mathbb{R}$的连续映射.

证明 *1.5.1* 如果有$\|x_n - x\| \to 0$,则$|\|x_n\| - \|x\|| \leq \|x_n - x\| \to 0$

定理 1.5.1 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 1.5.2 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 1.5.2 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$\|Lx_n - Lx\| = \|L(x_n - x + x_0) - Lx_0\| \to 0$. 得证

有界线性映射

定义 1.5.3 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 1.5.3 线性映射是连续的, 当且仅当它是有界的.

证明 1.5.3 如果L是有界线性映射,且有$\|x_n\| \to 0$, 则$\|Lx_n\| \leq \alpha\|x_n\| \to 0$. 即线性映射$L$在0连续, 根据定理1.5.2, L是连续映射. 如果L不是有界的, 那么对于任意$n \in N$, 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$y_n \to 0$,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$E_1 \mapsto E_2$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 1.5.4 *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 1.5.4 *如果$L_1, L_2 \in \mathcal{B}(E_1, E_2)$有$x \in E_1$使得$\|x\| = 1$, 则有*

$$\|L_1 x + L_2 x\| \leq \|L_1 x\| + \|L_2 x\|$$

这意味着

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

因此有,

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

所以上述范数定义的确满足三角不等式.

(1.5.4)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 1.5.5 *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 1.5.5 *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:*

$$\|L_m x - L_n x\| \leq \|L_m - L_n\| \|x\| \to 0 \quad m, n \to \infty$$

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$L_n x \to y$. 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:

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

接下来再证明$L \in \mathcal{B}(E_1, E_2)$和$\|L_n - L\| \to 0$. 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

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

因此L也是有界的, 这样$L \in \mathcal{B}(E_1, E_2)$. 接下来还有$\|L_n - L\| \to 0$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$n \to \infty$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 1.5.6 (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$\lim_{i\to\infty} x_{ij} = 0$

(b) 对于每一个指标列p_n有子列q_n使得

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

则$\lim_{i\to\infty} x_{ii} = 0$

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 1.5.4 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 1.5.2 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 1.5.3 令$E = \mathcal{C}([0,1])$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 1.5.7 *(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$f(z) = z$

证明 1.5.6 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, $x_n = f(x_{n-1}), n = 1, 2, \ldots$. 我们将证明$(x_n)$是*Cauchy*序列. 首先观察到, 对于任意$n \in \mathbb{N}$

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

因此对于任意$m, n \in \mathbb{N}$使得$m < n$有

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

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$n \to \infty$时$x_n \to z$.接下来证明唯一存在z使得$f(z) = z$. 首先因为

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

$\|f(z)-z\|=0$, 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$a_1,a_2,\cdots,a_n$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

$$\mathbf{x}=x_1a_1+x_2a_2+\cdots+x_na_n=\sum_{i=1}^{n}x_ia_i$$
$$\mathbf{y}=y_1a_1+y_2a_2+\cdots+y_na_n=\sum_{j=1}^{n}y_ja_j$$

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

$$\mathbf{x}+\mathbf{y}=(x_1+y_1)a_1+\cdots+(x_n+y_n)a_n$$
$$k\mathbf{x}=kx_1a_1+\cdots+kx_na_n$$

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

$$(x_1+y_1,x_2+y_2,\cdots,x_n+y_n)=(x_1,\cdots,x_n)+(y_1,\cdots,y_n)$$
$$(kx_1,kx_2,\cdots,kx_n)=k(x_1,x_2,\cdots,x_n)$$

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 1.6.1 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$\mathbb{F}$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) $\sigma(\sum_{i=1}^{n}k_ia_i)=\sum_{i=1}^{n}k_i\sigma(a_i)$

(3) E中向量组$a_1,\cdots,a_n$线性相关的充分必要条件是,它们的像$\sigma(a_1),\sigma(a_2),\cdots,\sigma(a_n)$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 1.7.1 如果对于线性空间E中任意元素a,b和数域$\mathbf{F}$中任意数k,都有这样的线性变换A,满足:

$$A(a+b)=A(a)+A(b)$$
$$A(ka)=kA(a)$$

则这个变换A称为线性变换,以后使用大写$A,B,\cdots$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 1.7.1 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 1.7.2 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 1.7.2 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 1.7.3 实数域R上的线性空间$R[x]$中, 线性变换

$$D(f(x)) = f'(x)$$
$$F(f(x)) = \int_0^x f(t)dt$$

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

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

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

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

由于线性映射保证线性结构不变.

$$Lx = x_1 Le_1 + x_2 Le_2 + \cdots + x_n Le_n, \quad x_i \in \mathbb{R}, n = 1, 2, \ldots$$

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

$$L_i = Le_i \in \mathbb{R}$$

称为对偶向量L的分量, 可将$L \in E^*$记作

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

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$e_1' e_2^*, \ldots, e_n^*$线性无关. 它们形成了$E^*$中的一组基, 称为与基$(e_1, e_2, \ldots, e_n)$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$(e_1, e_2, \ldots, e_n)$可以用$(|1\rangle, |2\rangle, \ldots, |n\rangle)$去表示.

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

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

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

基矢$\langle i|, i = 1, \ldots, n$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 1.8.1 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) $\langle x,\alpha y+\beta z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$

(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 1.8.1 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 1.8.2 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 1.8.3 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 1.8.4 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*1.8.3*.

例 1.8.5 区间$[a,b]$上的连续复函数的空间$\mathcal{C}([a,b])$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 1.8.6 $L^2(\mathbb{R})$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 1.8.7 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 1.8.1 对于内积空间里的任意两个元素x,y, 有

$$|\langle x,y\rangle| \le \|x\|\|y\| \tag{1.3}$$

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

证明 1.8.1 若$y = 0$, 则(1.3)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(1.8.1)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.

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

证毕.

推论 1.8.1 *(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 1.8.2

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

其中Rez表示$z \in \mathbb{C}$的实部.

内积空间的范数

定义 1.8.2 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 1.8.2 *(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 1.8.3 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$y \to -y$, 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 1.8.3 *(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, $\perp$关系是对称的.

定理 1.8.3 *(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 1.8.4 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1.8.1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$\theta$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 1.8.4 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 1.8.8 由于$\mathbb{C}$与$\mathbb{C}^n$是完备的, 它们都是*Hilbert*空间.

例 1.8.9 l^2是*Hilbert*空间, 其完备性在之前例子1.4.2已经证明过了, 这里不再复述.

例 1.8.10 例子1.8.4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

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

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 1.8.11 例子1.8.5中的空间是另一个不完备的内积空间例子. 事实上,考虑$\mathcal{C}([0,1])$中的如下函数序列:

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

显然f_n是连续的. 进一步可知

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

因此(f_n)是柯西序列. 显然这个序列是逐点收敛于函数

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

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

例 1.8.12 $L^2(R)$和$L^2([a,b])$是*Hilbert*空间

例 1.8.13 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).

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

定义

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

因为

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

(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

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

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

例 1.8.14 *(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$\tilde{H}^m(\Omega), m = 1,2,\ldots$, 使得对于所有的$|\alpha| \leq m$有$D^\alpha f \in L^2(\Omega)$成立, 其中$\alpha = (\alpha_1,\ldots,\alpha_n), \alpha_1,\ldots,\alpha_n$是非负整数, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, 以及

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

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$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, 我们有

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

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

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

$\tilde{H}^m(\Omega)$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 1.8.5 内积空间E中的序列(x_n), 如果当$n \to \infty$时有$\|x_n - x\| \to 0$, 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 1.8.6 内积空间E中有序列(x_n), 如果当$n \to \infty$时对于任意$y \in E$有$\langle x_n, y\rangle \to \langle x, y\rangle$, 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$n \to \infty$时, 对于任意$y \in E$有$\langle x_n - x, y\rangle \to 0$. 我们通常用$x_n \to x$表示强收敛, $x_n \xrightarrow{w} x$表示弱收敛

定理 1.8.4 强收敛必定弱收敛, 也就是说, $x_n \to x$意味着$x_n \xrightarrow{w} x$

证明 1.8.5 设序列(x_n)强收敛于x. 也就是说当$n \to \infty$时, 有$\|x_n - x\| \to 0$. 根据*Cauchy-Schwarz*不等式

$$|\langle x_n - x, y\rangle| \le \|x_n - x\|\|y\| \to 0 当 n \to \infty$$

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

对于内积空间E中任意的固定点y, 映射$\langle \cdot, y\rangle : E \mapsto \mathbf{C}$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$\langle x, \cdot\rangle \mapsto \mathbf{C}$也连续.

定理 1.8.5 如果$x_n \to x, y_n \to y$, 则$\langle x_n, y_n\rangle \to \langle x, y\rangle$

证明 1.8.6 若$x_n \to x, y_n \to y$, 则有

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

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:

$$x_n \to x意味着\|x_n\| \to \|x\|$$

定理 1.8.6 *若$x_n \xrightarrow{w} x$且$\|x_n\| \to \|x\|$, 则$x_n \to x$.*

证明 1.8.7 *若对于所有的y,有$x_n \xrightarrow{w} x$, 我们有:*

$$\langle x_n, x\rangle \to \langle x, y\rangle, \quad n \to \infty$$

因此

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

现在, 当$n \to \infty$时

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

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 1.9.1 *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$y_1, y_2, \ldots, y_n$都正交, 则x与$y_1, y_2, \ldots, y_n$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 1.9.1 正交系统是线性无关的系统

证明 1.9.1 令S是一个正交系统. 假设存在$x_1, x_2 \ldots, x_n \in S$和$\alpha_1, \alpha_2, \ldots, \alpha_n$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$x_1, \ldots, x_n$是线性无关的.

正交序列

定义 1.9.2 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 1.9.1 对于向量$e_n = (0, 0, \ldots, 1, 0, \ldots, \ldots)$, 其中1在第$n$个位置.集合$S = e_1, e_2, \ldots, e_n, \ldots$构成$l^2$中的正交归一系统.

例 1.9.2 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

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

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 1.9.3 *Legendre*多项式如下定义

$$P_0(x) = 1$$
$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \quad n = 1, 2, 3, \ldots$$

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

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

类似的过程

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

其中我们有$p_n(x)$的$2n$次导数, 这样就只剩下$2n$次项. 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 1.9.4 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$(-1)^{n+m}$

$$(-1)^{n+m}\langle\varphi_n,\varphi_m\rangle=\left[e^{x^2}\frac{d^n}{dx^n}e^{-x^2}\frac{d^m}{dx^m}e^{-x^2}\right]_{-\infty}^{\infty}-\int_{-\infty}^{\infty}\frac{d}{dx}\left[e^{x^2}\frac{d^n}{dx^n}e^{-x^2}\right]\frac{d^{m-1}}{dx^{m-1}}e^{-x^2}dx \tag{1.4}$$

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

$$x^ke^{-x^2}\to 0当x\to\infty$$

所以(1.4)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

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

因为$H_n(x)$是n次多项式, 直接求导得到

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

然后再求导

$$\frac{d^n}{dx^n}\left[e^{x^2}\frac{d^n}{dx^n}e^{-x^2}\right]=\frac{d^n}{dx^n}((-2x)^n+\cdots)=(-1)^n2^nn!$$

所以得到

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

因此函数

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

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$(w_n),(x_n)$

$$w_1 = y_1 \qquad\qquad x_1 = \frac{w_1}{\|w_1\|}$$
$$w_k = y_k - \sum_{n=1}^{k} \langle x_n, y_k\rangle x_n, \quad x_k = \frac{w_k}{\|w_k\|}$$

序列(w_n)是正交的. 我们注意到

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

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

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

因此向量$w_1, w_2, \ldots, w_k$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$x_1, x_2, \ldots, x_n$的线性组合也是$y)1, \ldots, y_n$的线性组合, 反之亦然.换句话说对于任意$n \in \mathbb{N}$有$span x_1, \ldots, x_n = span y_1, \ldots, y_n$.

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 1.9.2 (勾股定理) *如果$x_1, x_2, \ldots, x_n$是内积空间中的正交向量, 则有*

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

证明 1.9.2 *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(1.9.2)对于$n-1$都成立, 也就是说*

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

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

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

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 1.9.3 (Bessel等式和不等式) 令$x_1, \ldots, x_n$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

$$\left\| x - \sum_{k=1}^{n} \langle x_k, x \rangle x_k \right\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x_k, x \rangle|^2 \tag{1.5}$$

并且有

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

证明 1.9.3 由于有勾股定理(1.9.2), 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$\alpha_1 \ldots, \alpha_n$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(1.5).根据(1.5)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$n \to \infty$我们可以得到

$$\sum_{k=1}^{\infty} |\langle x_k, x\rangle|^2 \leq \|x\|^2 \tag{1.6}$$

所以

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

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1.6暗含了这样的一个关系, 对于任意$x \in E$都有$\sum_{n=1}^{\infty} |\langle x_k, x\rangle|^2$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:

$$x \sim \sum_{n=1}^{\infty} \langle x_n, x\rangle x_n \tag{1.7}$$

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

$$\left\| \sum_{n=1}^{\infty} \alpha_n x_n \right\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2 \tag{1.8}$$

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

证明 1.9.4 *对于任意$m > k > 0$, 利用勾股定理我们可以得到,*

$$\| \sum_{n=k}^{m} \alpha_n x_n \|^2 = \sum_{n=k}^{m} |\alpha_n|^2 \tag{1.9}$$

若$\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$, 则根据方程(1.9)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(1.9)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(1.8)只需要令$k = 1, m \to \infty$即可得.

这个定理和等式(1.6)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 1.9.5 令$H = L^2([-\pi,\pi]), x_n(t) = \frac{1}{\sqrt{\pi}}\sin nt, n = 1,2,\ldots$. 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 1.9.3 (完备正交归一序列) 若内积空间E中, 对于任意$x \in E$有

$$x = \sum_{n=1}^{\infty}\langle x_n, x\rangle x_n \tag{1.10}$$

则正交归一序列(x_n)被叫做完备序列

方程(1.10)的右边是无穷级数, 这意味着

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

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

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

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 1.9.4 (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 1.9.5 *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 *1.9.5* 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(1.6)和定理1.9.4, y在H中存在. 因此对于任意$n \in \mathbb{N}$

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

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 1.9.6 (*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$x \in H$

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, x_n \rangle|^2 \tag{1.11}$$

证明 1.9.6 令$x \in H$, 由方程(1.5)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$n \to \infty$收敛, 因此

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

(1.11)成立.

相反地, 若(1.11)成立, 则(1.11)右边表达式在$n \to \infty$收敛, 这样

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

这就证明了序列(x_n)是完备的.

例 1.9.6 正交系

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

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 1.9.7 序列函数

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

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.

$$\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}} \cos x, \sqrt{\frac{2}{\pi}} \cos 2x, \sqrt{\frac{2}{\pi}} \cos 3x, \ldots$$
$$\sqrt{\frac{2}{\pi}} \sin x, \sqrt{\frac{2}{\pi}} \sin 2x, \sqrt{\frac{2}{\pi}} \sin 3x$$

Chapter 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$\mathbf{r}_a(a = 1, 2, \ldots, N)$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 2.1.1 (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$q_1, q_2, \ldots, q_s$, 其导数$\dot{q}_i$和$\ddot{q}_i(i = 1, 2, \ldots, s)$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.

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

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 2.1.2 (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

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

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

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

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 2.1.3 (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 2.1.1 *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*

$$\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = 0, (i = 1, 2, \ldots, s)$$

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 2.1.1 *根据最小作用量原理可以知道*

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

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

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

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

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

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 2.1.4 (运动积分) *在运动过程中, 变量$q_i, \dot{q}_i (i = 1,2,\ldots,s)$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

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

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$2s$维相空间中去看运动积分, 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$(q_i(t), \dot{q}_i)(i = 1,2,\ldots,s)$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 2.1.5 (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

$$E = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \tag{2.1}$$

证明 2.1.2 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

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

上式两边乘以$\ldots q_i$, 然后对i求和. 得到:

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

等式左边第一项可以写成:

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

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

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

由于

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

所以有

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

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

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{\partial L}{\partial t}$$

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

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

在这种情况下, E表示系统的机械能.

定义 2.1.6 (广义动量) 当我们用广义坐标去描述系统时, 定义广义动量

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.2}$$

广义动量包含了通常的动量、角动量等.

例 2.1.1 广义动量并不是唯一的

证明 2.1.3 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

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

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(2.2), 相应于L'有

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

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 2.1.2 (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 2.1.4 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

$$\begin{aligned} t \mapsto t' &= t + \epsilon_\alpha \xi_\alpha(q(t),t) \\ q_i \mapsto q_i' &= q_i + \epsilon_\alpha \eta_{i\alpha}(q(t),t) \end{aligned} \tag{2.3}$$

其中$\epsilon_\alpha(\alpha = 1,2,\ldots,r)$是无穷小实数, $\xi_\alpha(q,t), \eta_{i\alpha}(q,t)$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(2.3)下, 作用量

$$S = \int_{t_1}^{t_1} L(q(t),\dot{q}(t),t) \tag{2.4}$$

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(2.4)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(2.4)表明了变换前后拉格朗日函数之间存在这样的关系

$$L'(q'(t'),\dot{q}'(t'),t')\frac{dt'}{dt} = L(q(t),\dot{q}(t),t) \tag{2.5}$$

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

$$L'(q,\dot{q},t) = L(q,\dot{q},t) \tag{2.6}$$

则称L在变换(2.3)下具有形式不变性

(2) 协变性:

如果变换前后满足

$$L'(q,\dot{q},t) = L(q,\dot{q},t) + \frac{d}{dt}\Lambda(q,t) \tag{2.7}$$

则称L在变换(2.3)下具有形式协变性.

(2.7)是(2.6)的推广. 这里$\Lambda(q,t)$是q,t的连续可微函数. 具有性质(2.6)和性质(2.7)的变换(2.3)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(2.7)中, 有

$$\left(L(q'(t'),\dot{q}'(t'),t') + \frac{d}{dt'}\Lambda(q'(t'),t')\right)\frac{dt'}{dt} = L(q(t),\dot{q}(t),t) \tag{2.8}$$

由(2.3)可以得到

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

其中均保留到ϵ的一阶项. 有

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

其中

$$\Pi_\alpha(q(t),t)=\frac{\partial\Lambda}{\partial t}\xi_\alpha(q(t),t)+\frac{\partial\Lambda}{\partial q_i}\eta_{i\alpha}(q(t),t) \tag{2.9}$$

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(2.3)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(2.8)保留一阶项可以得到

$$\begin{aligned}\xi_\alpha(q(t),t)\frac{\partial L}{\partial t}+\left[\eta_{i\alpha}(q(t),t)\frac{\partial L}{\partial q_i}+\left(\frac{d\eta_{i\alpha}(q(t),t)}{dt}-\dot{q}_i\frac{d\xi_\alpha(q(t),t)}{dt}\right)\frac{\partial L}{\partial\dot{q}_i}\right]\\+L\frac{d\xi_\alpha(q(t),t)}{dt}=-\frac{d}{dt}\Pi_\alpha(q(t),t)\end{aligned} \tag{2.10}$$

方程(2.10)可以用来判断变换(2.3)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(2.10)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

$$\begin{aligned}I_\alpha(q,t)&=(\eta_{i\alpha}-\dot{q}_i\xi_\alpha)\frac{\partial L}{\partial\dot{q}_i}+L\xi_\alpha+\Pi_\alpha\\&=\eta_{i\alpha}p_i-E\xi_\alpha+\Pi_\alpha,\quad(\alpha=1,2,\cdots,r)\end{aligned} \tag{2.11}$$

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(2.11)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(2.11)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 2.1.7 (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

$$H(p,q,t)=\sum_{i=1}^{s}p_i\dot{q}_i-L \tag{2.12}$$

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

$$\left(\frac{\partial H}{\partial\lambda}\right)_{p,q}=-\left(\frac{\partial L}{\partial\lambda}\right)_{\dot{q},q} \tag{2.13}$$

一个特例是对时间的偏导

$$\left(\frac{\partial H}{\partial t}\right)_{p,q}=-\left(\frac{\partial L}{\partial t}\right)_{\dot{q},q} \tag{2.14}$$

特别是

$$\frac{dH}{dt}=\frac{\partial H}{\partial t} \tag{2.15}$$

定义 2.1.8 (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

$$\begin{aligned}\dot{q}_i&=\frac{\partial H}{\partial p_i}\\ \dot{p}_i&=-\frac{\partial H}{\partial q_i},\quad(i=1,2,\ldots,s)\end{aligned} \tag{2.16}$$

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 2.1.5 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$p_i = \frac{\partial L}{\partial q_i}$同时我们对哈密顿函数$H = H(p_i, q_i)$进行全微分展开

$$dH = \sum_{i=1}^{s} \frac{\partial H}{\partial q_i} dq_i + \sum_{i=1}^{s} \frac{\partial H}{\partial p_1} dp_i$$

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(2.14)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(2.15)可以知道此时有能量守恒. 当L中有循环坐标时, 根据(2.13)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$q_i (i = s - m + 1, \ldots, s)$, 则对应的广义动量为

$$p_i = \beta_i$$

将广义动量代入H中有

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

非循环坐标满足哈密顿正则方程

$$\dot{q}_k = \frac{\partial H}{\partial p_k}$$
$$\dot{p}_k = -\frac{\partial H}{\partial q_k}$$

由此解出$2(s-m)$个q_k和p_k. 将其代入H中, 并通过对$\dot{q}_k = \frac{\partial H}{\partial p_i}$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 2.1.9 (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

$$\{f, g\} = \sum_k \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \tag{2.17}$$

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

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

(6) 雅克比恒等式(Jaccobi 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 2.1.10 (正则变量) 如果一组变量q_k, p_k满足下列关系

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

则叫做正则变量.

定理 2.1.3 (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

$$\frac{df}{dt}=\{f,H\}+\frac{\partial f}{\partial t} \tag{2.18}$$

证明 2.1.6 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 2.1.4 (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

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

证明 2.1.7 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(2.18)可以得到

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

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 2.1.8

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

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 2.1.9

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

’ 性质4的证明

证明 2.1.10

$$\begin{aligned}\{f_1f_2,g\} &= \sum_k \left[\frac{\partial(f_1f_2)}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial(f_1f_2)}{\partial p_k}\frac{\partial g}{\partial q_k}\right]\\ &= \sum_k \left[f_1\left(\frac{\partial f_2}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f_2}{\partial p_k}\frac{\partial g}{\partial q_k}\right) + f_2\left(\frac{\partial f_1}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k}\frac{\partial g}{\partial q_k}\right)\right]\\ &= f_1\sum_k \left(\frac{\partial f_2}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f_2}{\partial p_k}\frac{\partial g}{\partial q_k}\right) + f_2\sum_k \left(\frac{\partial f_1}{\partial q_k}\frac{\partial g}{\partial p_k} - \frac{\partial f_1}{\partial p_k}\frac{\partial g}{\partial q_k}\right)\\ &= f_1\{f_2,g\}+f_2\{f_1,g\}\end{aligned}$$

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 2.1.11 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

$$\begin{aligned}\{g,\{h,f\}\}+\{h,\{f,g\}\} &= \{g,\{h,f\}\}-\{h,\{g,f\}\}\\ &= D_1(D_2(f))-D_2(D_1(f))\\ &= (D_1D_2-D_2D_1)f\\ &= [D_1,D_2]f\end{aligned}$$

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$x_1,x_2,\ldots$的任意函数. 于是

$$\begin{aligned}D_1D_2 &= \sum_{k,l}\xi_k\eta_l\frac{\partial^2}{\partial x_k\partial x_l}+\sum_{k,l}\xi_k\frac{\partial\eta_l}{\partial x_k}\frac{\partial}{\partial x_l}\\ D_2D_1 &= \sum_{k,l}\eta_k\xi_l\frac{\partial^2}{\partial x_k\partial x_l}+\sum_{k,l}\eta_k\frac{\partial\xi_l}{\partial x_k}\frac{\partial}{\partial x_l}\end{aligned}$$

它们之差

$$D_1D_2-D_2D_1=\sum_{k,l}\left(\xi_k\frac{\partial\eta_l}{\partial x_k}-\eta_k\frac{\partial\xi_l}{\partial x_k}\right)\frac{\partial}{\partial x_l}$$

是一个一阶微分算子. 注意到我们现在的系统有

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

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 2.1.2 质点动量**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 2.1.12 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 2.1.3 角动量**J**各个分量组成的泊松括号满足

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

证明 2.1.13

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

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 2.1.11 (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$\phi_1, \phi_2, \ldots, \phi_s$, 且满足

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

这样的系统称为可积系统, 此时称$\phi_1, \phi_2, \ldots, \phi_s$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

$$\frac{dA}{dt} = \mathbf{AB} - \mathbf{BA} = [\mathbf{A}, \mathbf{B}]$$

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$\frac{d\mathbf{A}^n}{dt}$

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

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$tr(A^n)$是系统的运动积分

例 2.1.4 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

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

$$S = \int_{t_1}^{t_2} L dt \tag{2.19}$$

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 2.1.5 作用量对坐标的偏导为对应的动量:

$$p_i = \frac{\partial S}{\partial q_i} \tag{2.20}$$

证明 2.1.14 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

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

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 2.1.6 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 2.1.15 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2.20和定理2.1.6可以得到作用量全微分表达式

$$dS = \sum_i p_i dq_i - Hdt \tag{2.21}$$

其中作用量看做(2.19)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(2.21)写成积分形式

$$S = \int \sum_i (p_i dq_i - H dt) \tag{2.22}$$

则可以由最小作用量原理推导出哈密顿方程

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

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2.1.6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

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

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的尺寸增大时, 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2.1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.

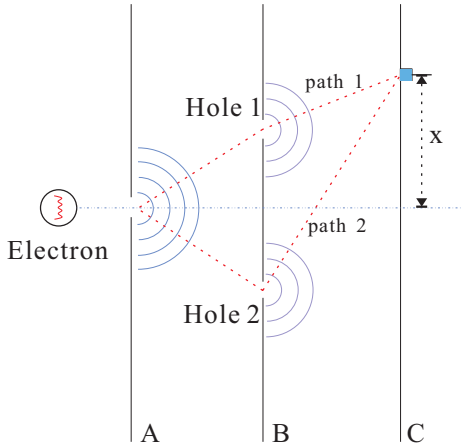


图 2.1: 实验装置A处射出的电子前进到屏C处的探测器, 其间插入一个带有两个孔的屏B. 对于到达的每一个电子, 探测器记录下一个计数; 当探测器距离屏幕中心为x时, 测量电子到达的比例, 并以x进行标定, 如图2.2所示.

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2.2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$P \neq P_1 + P_2$.

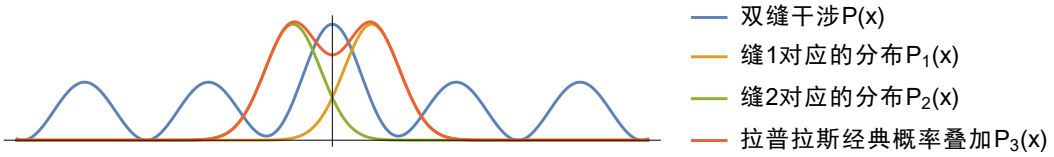


图 2.2: 实验结果. 电子到达x的概率是相对探测器位置x画出的. 图2.1的实验结果画在蓝色线中; 若仅仅只有孔1开启, 则电子仅仅能通过孔1, 其结果是绿色线所代表; 若仅仅孔2开启, 则是黄色线所代表的的结果. 如果我们设想每个电子恰好通过这个空或那个孔, 则当两孔都开启时, 预期得到的曲线应该是绿色线+黄色线=红色线. 这与实际所得的蓝色线大不相同.

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.

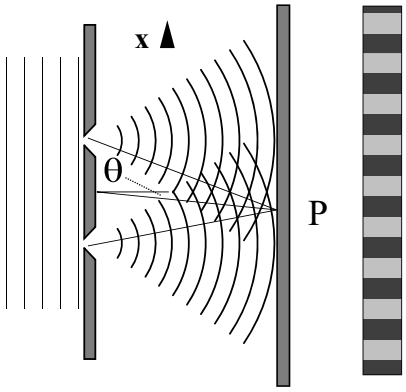


图 2.3: 相干干涉叠加出的电子分布曲线有极大值也有极小值, 在一些点发自两个不同孔的电子相干相消, 在另一些点上发自不同孔的两个电子相干相长, 由此产生了电子分布曲线$I(x)$的条纹

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2.4). 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.

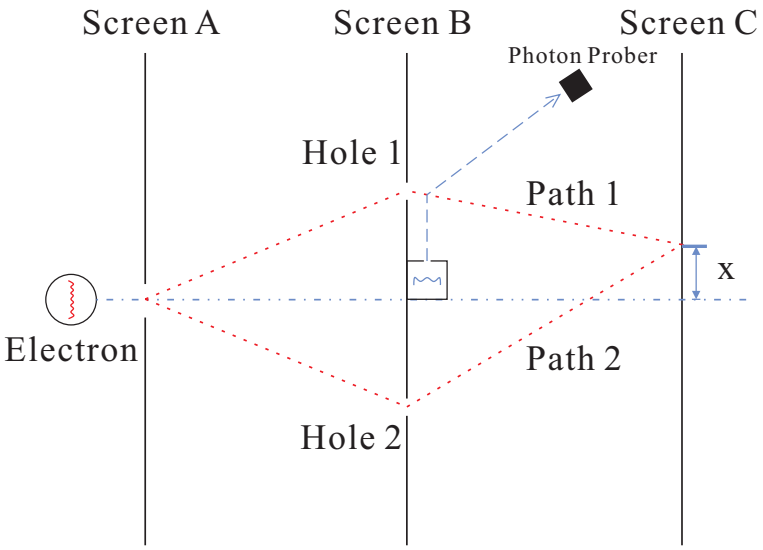


图 2.4: 实验2.1的一个修正. 我们在屏B后面放一个光源L, 并寻找通过孔1或者孔2的电子散射的光. 用一个很强的光源, 确实可以发现每个电子通过这个或那个孔. 但现在到达x的概率不再是由图2.2中的蓝线给出而是由图2.2中的红线给出.

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2.2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2.2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2.2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: 当时崭新的力学自洽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. 量子力学自洽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 2.3.1 (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$|\xi_1\xi_2\ldots\xi_s\rangle$, 这里$\xi_1,\xi_2,\ldots,\xi_s$是$s$个独立不同的量子数, 有些地方也用记号$|\psi_{\xi_1,\xi_2,\ldots,\xi_s}\rangle$.

例 2.3.1 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2.1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

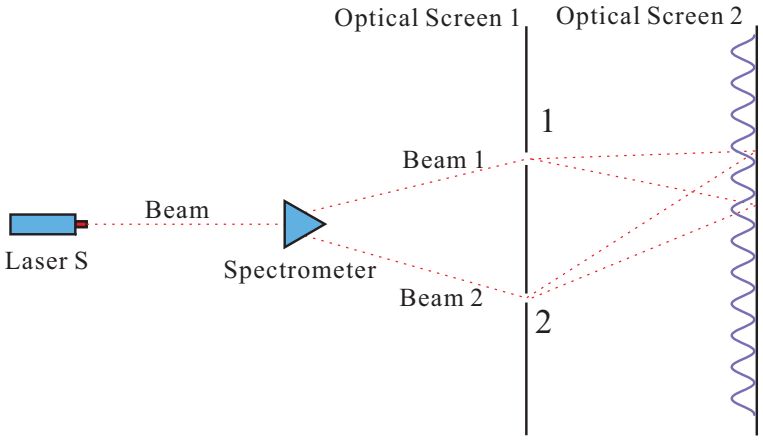


图 2.5: 单光子干涉实验: 激光器发射一束激光通过分光器分为两束, 束流1通过孔1, 束流2通过孔2, 然后在光屏2上发生干涉. 若减弱激光器S, 以至于一束激光在足够长的时间内只射出一个光子, 该光子通过分光器后处于束流1的状态称作$|1\rangle$,处于束流2的状态称作状态$|2\rangle$

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. 用光学的语言来说就是, 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 2.3.1 (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 2.3.1 *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 2.3.1 *根据量子态线性叠加的假设2.3.1我们可以得到*

$$c_1|\phi\rangle + c_2|\phi\rangle = (c_1 + c_2)|\phi\rangle$$
$$c|\phi_1\rangle + c|\phi_2\rangle = c(|\phi_1\rangle + |\phi_2\rangle)$$

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

$$|\phi_1\rangle + |\phi_2\rangle = |\phi_2\rangle + |\phi_1\rangle$$
$$(|\phi_1\rangle + |\phi_2\rangle) + |\phi_3\rangle = |\phi_1\rangle + (|\phi_2\rangle + |\phi_3\rangle)$$

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**F**(量子力学里$\mathbf{F} = \mathbb{C}$)是量子态之间的内积

定义 2.3.2 (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$\langle\varphi|$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$\langle\varphi|\phi\rangle \in \mathbb{C}$, 所以左矢量$\langle\phi|$是一个线性映射$L : |\phi\rangle \longmapsto \mathbb{C}$. 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 2.3.2 (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:*

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

证明 2.3.2 *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ψ与之对应. 我们假定这个函数数是线性的, 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$\phi$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$\langle\varphi|$. 它们之间的线性条件可以写成:

$$\begin{aligned} \langle\varphi|(|\psi\rangle + |\psi'\rangle) &= \langle\varphi|\psi\rangle + \langle\varphi|\psi'\rangle \\ \langle\varphi|(c|\psi\rangle) &= c\langle\varphi|\psi\rangle \end{aligned} \tag{2.23}$$

其中$c \in \mathbb{C}$.

性质 2.3.1 (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$\langle\varphi|\psi\rangle = 0$, 则$\langle\varphi| = 0$

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

$$(\langle\varphi| + \langle\varphi'|)|\psi\rangle = \langle\varphi|\psi\rangle + \langle\varphi'|\psi\rangle \tag{2.24}$$

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

方程(2.23)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 2.3.2 *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 2.3.3 (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ϕ, 我们有它们的内积$\langle\varphi|c\psi\rangle = c\langle\varphi|\psi\rangle$与$\langle c\psi|\varphi\rangle = c^*\langle\psi|\varphi\rangle$.这使得我们进行进一步的假设:

假设 2.3.3 (态矢量内积Hermite共轭) *给定量子态φ与ψ, 我们假定内积$\langle\varphi|\psi\rangle$与$\langle\psi|\varphi\rangle$之间满足Hermite共轭*

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

在这个假设下, 令$\langle\varphi| = \langle\psi|$, 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

$$\langle\psi|\psi\rangle > 0 \tag{2.25}$$

[1]这是为了用内积去定义范数而做的假设.

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$|\phi_1 + \phi_2| < |\phi_1| + |\phi_2|$. 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.

定理 2.3.4 (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 2.3.3 *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$\||\varphi\rangle\| + \||\psi\rangle\| > \||\varphi + \psi\rangle\|$

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

定义 2.3.3 *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2.25和归一化条件2.3.3, 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2.5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:

[2]测量问题会在算符谱理论中进行解释

定理 2.3.5 (完备性) 若$\{|\psi_n\rangle\} \subset \mathcal{H}$是一组完备正交基. 则对于任意$|\varphi\rangle \subset \mathcal{H}$.

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

该定理实质上就是Parseval关系. 证明过程在1.11. 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.

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

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.

Chapter 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$\{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_n\rangle\}$), 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 3.1.1 (线性算符) $\mathcal{H}$中的线性算符$\hat{A}$是指映射$A : \mathcal{H} \longmapsto \mathcal{H}$, $\hat{A}$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:

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

下面是一些$\mathcal{H}$上线性算符的例子.

例 3.1.1

1 恒等映射

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

2 坐标算符

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

3 动量算符

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

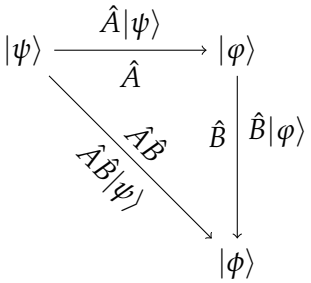
4 两个不同量子态$\psi \longmapsto \varphi = \hat{A}\psi$

$$\hat{A}:|\psi\rangle \longmapsto |\varphi\rangle = |\hat{A}\psi\rangle = \hat{A}|\psi\rangle$$

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$|\psi\rangle \overset{Dual}{\longmapsto} \langle\psi|$. 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.

$$\begin{aligned}
\text{量子态:}&\psi \longmapsto \varphi = \hat{A}\psi \\
\text{右矢量:}&|\psi\rangle \longmapsto |\varphi\rangle = |\hat{A}\psi\rangle = \hat{A}|\psi\rangle \\
\text{左矢量:}&\langle\psi| \longmapsto \langle\varphi| = \langle\hat{A}\psi|
\end{aligned}$$

两个不同线性算符的乘积也可以利用线性映射定义



定义 3.1.2 (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2.3.5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 3.1.1 (完备性条件) 若$\{|\psi_n\rangle : |\psi_n \in \mathcal{H}\}$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 3.1.1 根据上一节的*Parseval*公式2.3.5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 3.1.3 (有界算符) 算符$\hat{A}$是$\mathcal{H}$上的有界算符定义为

$$\|\hat{A}\| := \sup_{\{|\psi\rangle\in\mathcal{H}|\||\psi\rangle\|=1\}} \|\hat{A}|\psi\rangle\| < \infty \tag{3.1}$$

事实上, 表达式3.1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 3.1.1 如果算符A在定义域$\mathcal{D} \subset \mathcal{H}$上稠密, 对于$\psi$满足$\|\hat{A}|\psi\rangle\| \leq C\||\psi\rangle\|$($C$与$\psi$无关), 则它可以扩展到$\mathcal{H}$上的有界算符, 满足相同的界:

$$\|\hat{A}|\psi\rangle\| \leq C\||\psi\rangle\|, \quad |\psi\rangle \in \mathcal{H}$$

证明 3.1.2 对于任意$|\psi_n\rangle \in \mathcal{H}$, 有序列$\{|\psi_n\rangle\} \subset \mathcal{D}$使得当$n \to \infty$时$|\psi_n\rangle \to |\psi\rangle$($\mathcal{D}$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$\{A|\psi_n\rangle\}$是一个*Cauchy*序列, 所以对于一些$|\varphi\rangle \in \mathcal{H}$, $\hat{A}|\psi_n\rangle \to |\varphi\rangle$, 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 3.1.4 (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$j \to \infty$时,$|\psi_j\rangle \to |\psi\rangle, A|\psi_j\rangle \to |\varphi\rangle$, 可以的得到$|\psi\rangle \in \mathcal{D}(\hat{A})$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 3.1.5 (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 3.1.2 定义在全体$\mathcal{H}$上的闭算子或对称算子是有界的.

证明 3.1.3 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 3.1.6 (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$\mathcal{D}(\hat{B}) = \mathcal{R}(\hat{A}), \mathcal{D}(\hat{A}) = \mathcal{R}(\hat{B})$,

$$\hat{B}\hat{A} = \mathbf{1}|_{\mathcal{R}(\hat{A})} \quad \hat{A}\hat{B} = \mathbf{1}_{\mathcal{R}(B)}$$

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$|\varphi\rangle \in \mathcal{R}(A)$.

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, $\hat{A}|\psi\rangle = 0 \Longrightarrow |\psi\rangle = 0$, 或者等价于它是平凡核或者零空间:

$$\mathrm{Ker}\,(\hat{A}) := \left\{|\psi\rangle \in \mathcal{D}(\hat{A}) : \hat{A}|\psi\rangle = 0\right\} = \{0\} \tag{3.2}$$

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:

$$\mathcal{R}(\hat{A}) = \mathcal{H} \tag{3.3}$$

条件3.2和3.3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3.1.4, A^{-1}也是, 因此根据定理3.1.2, A^{-1}是有界的.

2. 如果$\hat{A}$是对称算符, 那么根据定理3.1.2A^{-1}也是对称算符.

算符的逆运算有许多常用性质, 这里稍微列举一下

性质 3.1.1 (若算符$\hat{A}, \hat{B}, \hat{C}$是可逆算符)

1. $(\hat{A}\hat{B}\hat{C})^{-1} = \hat{C}^{-1}\hat{B}^{-1}\hat{A}^{-1}$

2. 若$c \in \mathbb{C}$, 则有$(c\hat{A})^{-1} = \frac{1}{c}\hat{A}^{-1}$

除了这两个关系外, 下列定理在量子散射里比较常用.

定理 3.1.3 *$\hat{A}$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 3.1.1 *证明:$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}$*

问题 3.1.2 *若线性算符$\hat{A}$可逆, $\{|\psi_k\rangle\}(k = 1, 2, 3, \cdots, n)$是$\mathcal{D}(\hat{A})$的一组完备集, 证明$\hat{A}|\psi_k\rangle(k = 1, 2, \cdots, n)$在$\mathcal{R}(\hat{A})$中也是一组完备集.*

问题 3.1.3 (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 3.1.4 (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$1+\hat{A}$.

$$(1+\hat{A})^{-1} = \sum_{n=0}^{\infty}(-\hat{A})^n$$

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 3.2.1 (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

$$(\hat{A}+i\mathbf{1})|\psi\rangle = |\phi\rangle$$

对于所有的$|\phi\rangle \in \mathcal{H}$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 3.2.2 (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:

$$\langle\hat{A}\varphi|\psi\rangle = \langle\varphi|\hat{A}^\dagger\psi\rangle \tag{3.4}$$

对于所有的$|\psi\rangle \in \mathcal{D}(\hat{A})$以及对于定义域

$$\mathcal{D}(\hat{A}^\dagger) := \{\varphi \in \mathcal{H} : \text{对于某些与}|\psi\rangle\text{无关的常数}C_\varphi, |\langle\varphi|\hat{A}\psi\rangle| \le C_{|\varphi\rangle}\||\psi\rangle\|, \forall|\psi\rangle \in \mathcal{D}(\hat{A})\}$$

中的$|\varphi\rangle$都成立

之前我们定义了对称算符3.1.5,这里我们引入一种特殊的对称算符,叫做自伴算符.

定义 3.2.3 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 3.2.1 设$\hat{A}$是对称算符. 若存在一些$z(Imz > 0)$有$\mathcal{R}(\hat{A} - z\mathbf{1}) = \mathcal{H}$, 则对于$Imz > 0$的每一个$z$都成立.同样的, 对于$Imz < 0$也有相同的结论. 进一步说, 若$\hat{A}$是自伴算符, 则对于每个$Imz \neq 0$的$z$, $\hat{A} - z\mathbf{1}$是可逆的, 并且满足估计

$$\|(\hat{A} - z\mathbf{1})^{-1}\| \leq \frac{1}{|Imz|} \tag{3.5}$$

证明 3.2.1 记$z = \lambda + i\mu \quad \lambda, \mu \in R$.由于$\hat{A}$是对称算符, 我们有

$$\begin{aligned}
\|(\hat{A} - z\mathbf{1})|\psi\rangle\|^2 &= \langle(\hat{A} - z\mathbf{1})\psi|(\hat{A} - z\mathbf{1})\psi\rangle \\
&= \langle(\hat{A} - (\lambda + i\mu)\mathbf{1})\psi|(\hat{A} - (\lambda + i\mu)\mathbf{1})\psi\rangle \\
&= \langle(\hat{A} - \lambda\mathbf{1})\psi|(\hat{A} - \lambda\mathbf{1})\psi\rangle - i\mu\langle(\hat{A} - \lambda\mathbf{1})\psi|\psi\rangle + i\mu\langle\psi|(\hat{A} - \lambda\mathbf{1})\psi\rangle + \mu^2\langle\psi|\psi\rangle \\
&= \|(\hat{A} - \lambda\mathbf{1})|\psi\rangle\|^2 + \||\mu\psi\rangle\|^2 \geq |\mu|^2\||\psi\rangle\|^2
\end{aligned} \tag{3.6}$$

因此, 若$Imz > 0$(或者$Imz < 0$), 则$\mathrm{Ker}\,(A - z) = 0$. 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(3.6)意味着$(\hat{A} - z\mathbf{1})^{-1}$是有界的, 界由(3.5)确定(如果定义$|\varphi\rangle := (\hat{A} - z\mathbf{1})|\psi\rangle$), 所以特别是$(\hat{A} - z\mathbf{1})$是可逆的. 那么根据定理3.1.3, 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$|z' - z| < |Imz|$, $A - z'\mathbf{1}$是可逆的, 所以我们可以拓展$\hat{A} - z\mathbf{1}$的可逆性到全体$Imz' > 0$(或者$Imz' < 0$)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3.5.

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:

$$\hat{A}\text{是自伴算符} \Rightarrow \text{对于}\forall Imz \neq 0,\ (\hat{A} - z\mathbf{1})|\psi\rangle = |\phi\rangle\text{有唯一的解}$$

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 3.2.1 若$\hat{A}$是有界对称算符, 则$\hat{A}$是自伴算符.

证明 3.2.2 根据引理3.2.1, 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

$$(\mathbf{1} + \hat{K}(\lambda))|\psi\rangle = |\varphi\rangle$$

其中$\hat{K}(\lambda) = (i\lambda)^{-1}\hat{A}, |\varphi\rangle = (i\lambda)^{-1}|\phi\rangle$. 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3.1.4).

下面列举一些自伴算符满足的性质:

(1) $(\hat{A}^\dagger)^\dagger = \hat{A}$

(2) $(\hat{A}\hat{B}\hat{C})^\dagger = \hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger$

(3) $(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$

(4) $(\hat{A}^{-1})^\dagger = (\hat{A}^\dagger)^{-1}$

我们给出上述几条性质的证明

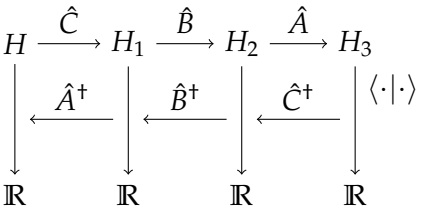
证明 3.2.3

首先是性质1

$$\begin{aligned}\langle\varphi|\hat{A}|\psi\rangle &= \langle\varphi|\hat{A}\psi\rangle \\ &= \langle\hat{A}\psi|\varphi\rangle^\dagger \\ &= \langle\psi|\hat{A}^\dagger\varphi\rangle^\dagger \\ &= \langle\hat{A}^\dagger\varphi|\psi\rangle = \langle\psi|(\hat{A}^\dagger)^\dagger|\psi\rangle\end{aligned}$$

性质2, 为了证明这个性质, 我们首先需要解释一下诱导伴随映射的概念

不失一般性, 我们考虑三个一般的线性映射$\hat{A}, \hat{B}, \hat{C}$, 这三个线性映射分别依次把$|\psi\rangle \in \mathcal{H}$映射到$|\psi'\rangle = \hat{C}|\psi\rangle \in \mathcal{H}_1, |\psi''\rangle = \hat{B}|\psi'\rangle \in \mathcal{H}_2, |\psi'''\rangle = \hat{A}|\psi\rangle \in \mathcal{H}_3$. 在之前左矢量的定义里, 我们知道左矢量构成矢量空间是由映射$\langle\cdot|\psi\rangle, |\psi\rangle \in \mathcal{H}$诱导出来的. 自然地



右矢量空间的映射L也可以由内积诱导左矢量空间的伴随映射$L^\dagger$. 内积的不变性就要求$\langle\varphi|\hat{A}\hat{B}\hat{C}\psi\rangle$诱导$\langle\hat{A}^\dagger\varphi|\hat{B}\hat{C}\psi\rangle$. 依次的, 可以得到诱导伴随映射的反顺序关系:

$$(\hat{A}\hat{B}\hat{C})^\dagger = \hat{C}^\dagger\hat{B}^\dagger\hat{A}^\dagger$$

性质3的证明非常简单. 利用$|\psi\rangle$的线性展开即可证明.

最后一个性质4的证明也很显然若$(\hat{A}^{-1})^\dagger \neq (\hat{A}^\dagger)^{-1}$, 则有

$$(\hat{A}^{-1})^\dagger \neq (\hat{A}^\dagger)^{-1}|\psi\rangle$$

这样就得到

$$\hat{A}^\dagger(\hat{A}^{-1})^\dagger|\psi\rangle \neq |\psi\rangle$$

两边用任意左矢量$\langle\varphi|$作用

$$\langle\varphi|\hat{A}^\dagger(\hat{A}^{-1})^\dagger\psi\rangle \neq \langle\varphi|\psi\rangle$$

利用一下伴算符的定义

$$\langle\hat{A}^{-1}\hat{A}\varphi|\psi\rangle = \langle\varphi|\psi\rangle \neq \langle\varphi|\psi\rangle$$

矛盾. 所以$\hat{(A^{-1})}^\dagger = (\hat{A}^\dagger)^{-1}$

3.2.2 投影算符

定义 3.2.4 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

$$|\phi\rangle \in \mathcal{R}(\hat{P}) \Longleftrightarrow \hat{P}|\phi\rangle = |\phi\rangle \quad \text{和} \quad |\phi\rangle \in (\mathcal{R}(\hat{P}))^\perp \Longleftrightarrow \hat{P}^\dagger|\phi\rangle = 0 \tag{3.7}$$

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 3.2.1 (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[1]$\{\psi_i\}_{i=1}^N$. 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$\{|\psi_k\}$. 则

$$|\psi\rangle \mapsto \sum_k \langle \psi_k | \psi \rangle |\psi_k\rangle \tag{3.8}$$

是一个投影.

由于$\langle \psi_k | \psi \rangle \in \mathbb{C}$, 式子(3.8)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

$$|\psi_k\rangle\langle\psi_k| : |\psi\rangle \longmapsto |\psi_k\rangle$$

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 3.2.5 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(3.7)意味着

$$|\psi\rangle \perp \mathcal{R}(\hat{P}) \Longleftrightarrow \hat{P}|\psi\rangle = 0$$

[1] $\langle \psi_i, \psi_j \rangle = \delta_{ij}$

即:$\mathrm{Ker}(\hat{P}) = \mathcal{R}(\hat{P})^{\perp}$[2] 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[3]. 而第二个例子是正交的当且仅当$\varphi = \psi$

3.2.3 练习题

问题 3.2.1 证明方程3.4定义了$D(\hat{A}^{\dagger})$上唯一的线性算符

问题 3.2.2 证明自伴算符两种定义3.2.1与3.2.2是等价的.

提示: 利用

引理 3.2.2 *若$\hat{A}$是有界对称算符, 则$\hat{A}$是自伴算符.*

问题 3.2.3 证明式子(3.7).

问题 3.2.4 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 3.2.5 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$\mathcal{R}(1-\hat{P}) \perp \mathcal{R}(\hat{P}), \mathrm{Ker}\,(1-\hat{P}) = \mathcal{R}(\hat{P})$;
3. $\mathcal{H} = \mathcal{R}(\hat{P}) \oplus \mathrm{Ker}\,\hat{P}$

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

[2]集合$S^{\perp}$表示与集合S正交的集合.

[3]特别应该注意$|\psi_k\rangle\langle\psi_k|$是正交投影算符

3.3.1 算符的谱

定义 3.3.1 *HIlbert*空间$\mathcal{H}$中算符$\hat{A}$的谱是由下列给定$\mathbb{C}$的子集

$$\sigma(\hat{A}) := \{\lambda \in \mathbb{C} : \hat{A} - \lambda\mathbf{1}\text{是不可逆的(无有界逆)}\}$$

$\mathbb{C}$中算符$\hat{A}$的谱的补集叫做$\hat{A}$的预解集:$\rho(\hat{A}) := \mathbb{C} \setminus \sigma(\hat{A})$. 对于$\lambda \in \rho(\hat{A})$,算符$(\hat{A} - \lambda\mathbf{1})^{-1}$是良定义的, 叫做$\hat{A}$的预解.

关于算符谱$\sigma(\hat{A})$有两个重要的定理

定理 3.3.1 谱$\sigma(\hat{A})$是闭集

证明 3.3.1 我们证明了谱的补集$\rho(\hat{A}) := \mathbb{C}/\sigma(\hat{A})$, 称为预解集, 是一个开集. 的确, 设$z \in \rho(\hat{A})$. 则$\hat{A} - z\mathbf{1}$是可逆的, 因此由于定理3.1.3, 若$|z' - z| < \|(\hat{A} - z\mathbf{1})^{-1}\|^{-1}$, 则$\hat{A} - z'\mathbf{1} = (\hat{A} - z\mathbf{1})[\mathbf{1} + (z - z')(\hat{A} - z\mathbf{1})^{-1}]$

我们看到自伴算符有实数谱.

定理 3.3.2 若$\hat{A} = \hat{A}^\dagger$, 则$\sigma(\hat{A}) \subset \mathbb{R}$.

,

证明 3.3.2 由3.2.1显然可知

$\hat{A} - \lambda\mathbf{1}$不可逆的一个熟悉原因是$(\hat{A} - \lambda\mathbf{1})|\psi\rangle = 0$有非零解$\psi \in \mathcal{D}(\hat{A}) \subset \mathcal{H}$. 这种情况下, 我们称$\lambda$是$\hat{A}$的本征值, $|\psi\rangle$称为本征值对应的本征矢量.

定义 3.3.2 算符$\hat{A}$的离散谱是

$$\sigma_d(\hat{A}) = \{\lambda \in \mathbb{C} : \lambda\text{是}\hat{A}\text{的有限重孤立本征值}\}$$

这里, 特征值λ的多重性是指本征空间的维数

$$\text{Ker}\,(\hat{A} - \lambda\mathbf{1}) := \{|\psi\rangle \in \mathcal{H} : (\hat{A} - \lambda\mathbf{1})|\psi\rangle = 0\}$$

孤立意思是λ的一些邻域与$\sigma(\hat{A})$的剩余部分不相交. 其余的谱叫做算符$\hat{A}$的连续谱:

$$\sigma_e(\hat{A}) := \sigma(\hat{A}) \setminus \sigma_d\hat{A}$$

定理 3.3.3 (正交性定理) 对应于自伴算符$\hat{A}$不同本征值λ,λ'的两个本征矢相互正交. 也就是说, 若$\hat{A}|\psi_\lambda\rangle = \lambda|\psi_\lambda\rangle$, $\hat{A}|\psi_{\lambda'}\rangle = \lambda'|\psi_{\lambda'}\rangle$, 则有

$$\langle\psi_\lambda|\psi_{\lambda'}\rangle = 0 \tag{3.9}$$

对于自伴算符$\hat{A}$, 集合$\{\hat{A}$的特征向量生成空间$\}$和$\{\hat{A}$的特征向量生成空间$\}^\perp$在定义的意义上的$\hat{A}$下是不变的, 其中

$$W^\perp := \{|\psi\rangle \in \mathcal{H} : \langle\psi|w\rangle = 0 \forall |w\rangle \in W\}$$

定义 3.3.3 *Hilbert*空间$\mathcal{H}$的子空间$W \subset \mathcal{H}$, 若无论何时$|w\rangle \in W \cap \mathcal{D}(\hat{A})$有$\hat{A}|w\rangle \in W$, 则$W$在算符$\hat{A}$下是不变的

3.3.2 算符函数与谱映射定理

我们考虑函数f在$\sigma(\hat{A})$的邻域内解析下的算符函数$f(\hat{A})$, 其中$\hat{A}$是有界算符.

算符函数

定理 3.3.4 设$\hat{A}$是有界自伴算符, 则$\sigma(\hat{A}) \subset [-\|\hat{A}\|, \|\hat{A}\|]$

证明 3.3.3 若有$|z| > \|\hat{A}\|$, 则$\|(\hat{A} - z)|\psi\rangle\| \geq (|z| - \|\hat{A}\|)\||\psi\rangle\| > 0 \neq 0$

设$f(\lambda)$在半径为R的圆域里解析, $\{\lambda \in \mathbb{C} : |\lambda| < R\}$, 其中$R > \|\hat{A}\|$. 所以$f$有泰勒级数

$$f(\lambda) = \sum_{n=0}^{\infty} a_n \lambda^n$$

在$|\lambda| < R$内收敛. 利用收敛级数定义算符$f(\hat{A})$

$$f(\hat{A}) := \sum_{n=0}^{\infty} a_n \hat{A}^n \tag{3.10}$$

在这样的算符函数定义下, 举几个常用例子

例 3.3.1 (有界自伴算符的指数形式$e^{\hat{A}}$) 有界算符的指数形式定义为:

$$e^{\hat{A}} = \sum_{n=0}^{\infty} \frac{\hat{A}^n}{n!}$$

这是绝对收敛的, 由于

$$\sum_{n=0}^{\infty}\frac{\|\hat{A}^n\|}{n!}\leq\sum_{n=0}^{\infty}\frac{\|\hat{A}\|^n}{n!}=e^{\|A\|}<\infty$$

例 3.3.2 (Neumann级数) 考虑函数$f(\lambda)=(\lambda-z)^{-1},|z|>\|\hat{A}\|$, 在圆域$\|\hat{A}\|<R<|z|$($R$是圆域半径)内解析, 有幂级数展开式

$$f(\lambda)=-\frac{1}{z}\frac{1}{1-\frac{\lambda}{z}}=-\frac{1}{z}\sum_{k=0}^{\infty}(\frac{\lambda}{z})^k$$

由式子(3.10)对应的算符定义式是预解式

$$(\hat{A}-z\mathbf{1})^{-1}=-\frac{1}{z}\sum_{k=0}^{\infty}$$

预解式对于预解集里$\forall z\in\rho(\hat{A})\mathbb{C}\setminus\sigma(\hat{A})$是有定义的. 事实上, $(\hat{A}-z\mathbf{1})^{-1}$是$z\in\rho(\hat{A})$的算符值解析函数. 为了见到这一点, 我们从下列关系式开始

$$(\hat{A}-z\mathbf{1})^{-1}=(\hat{A}-z_0\mathbf{1})^{-1}-(z_0-z)(\hat{A}-z_0)^{-1}(\hat{A}-z)^{-1}$$

其中$z,z_0\in\rho(\hat{A})$, 这个式子很容易证明. 这样就有

$$(\hat{A}-z\mathbf{1})^{-1}=[\mathbf{1}-(z-z_0)(\hat{A}-z_0\mathbf{1})^{-1}]^{-1}(\hat{A}-z_0\mathbf{1})^{-1}$$

若$|z-z_0|<(\|(\hat{A}-z_0)^{-1}\|)^{-1}$, 上式右边第一个逆可以用*Neumann*级数展开,

$$(\hat{A}-z\mathbf{1})^{-1}=\sum_{k=0}^{\infty}(z-z_0)^k(\hat{A}-z_0\mathbf{1})^{-k-1}$$

这样$(\hat{A}-z\mathbf{1})^{-1}$在$\forall z_0\in\rho(\hat{A})$的邻域内解析.

谱映射定理

定理 3.3.5 (谱映射定理) 设$\hat{A}$是一个有界算符, f是在圆域半径$R>\|\hat{A}\|$上的解析函数. 若$\hat{A}|\phi\rangle=\lambda|\phi\rangle$, 则$f(\hat{A})|\phi\rangle=f(\lambda)|\phi\rangle$

证明 3.3.4 若$\{a_k\}$是f幂级数的系数, 则

$$f(\hat{A})|\phi\rangle=\sum_{k=0}^{\infty}a_k\hat{A}^k|\phi\rangle=(\sum_{k=0}^{\infty}a_k\lambda^k)|\phi\rangle=f(\lambda)|\phi\rangle$$

假设f在$\sigma(\hat{A})$邻域内解析. 式子3.10可以替换成Riez围道积分形式

$$f(\hat{A}) := \frac{1}{2\pi i}\oint_\Gamma f(z)(\hat{A}-z\mathbf{1})^{-1}dz \tag{3.11}$$

其中Γ是$\mathbb{C}$中包围$\sigma(\hat{A})$的围道. 这里积分可以被理解成下列形式

$$\langle\psi|f(\hat{A})|\phi\rangle := \frac{1}{2\pi i}\oint_\Gamma f(z)\langle\psi|(\hat{A}-z\mathbf{1})^{-1}|\phi\rangle dz$$

其中$|\psi\rangle, |\phi\rangle \in \mathcal{H}$

对于$\forall|\psi\rangle, |\phi\rangle \in \mathcal{H}$, $\langle\psi|f(\hat{A})|\phi\rangle$唯一的决定算符$f(\hat{A})$.

类似的公式可以用于无界算符来定义算符函数$f(\hat{A})$.

若$\hat{A}$是无界自伴算符, 且f是连续有界函数, 则有界算符$f(\hat{A})$仍然可以有定义.

例 3.3.3 (无界自伴算符$\hat{A}$的指数形式$e^{i\hat{A}}$) 对于无界自伴算符我们定义指数形式我们通过有界算符逼近$\hat{A}$来定义指数形式$e^{i\hat{A}}$, 利用有界算符的幂级数展开

$$e^{\hat{A}} = \sum_{n=0}^{\infty}\frac{\hat{A}^n}{n!} \tag{3.12}$$

由于

$$\sum_{n=0}^{\infty}\frac{\|\hat{A}^n\|}{n!} \le \sum_{n=0}^{\infty}\frac{\|\hat{A}\|^n}{n!} = e^{\|\hat{A}\|} < \infty$$

式子(3.12)绝对收敛.

在之前例子3.3.1里我们知道式子3.12对于有界自伴算符$\hat{A}$成立. 我们现在把它推广到无界算符$\hat{A}$. 由于引理3.2.1, 算符

$$\hat{A}_\lambda := \frac{1}{2}\lambda^2[(\hat{A}+i\lambda\mathbf{1})^{-1}+(\hat{A}-\imath\lambda\mathbf{1})^{-1}]$$

是良定义的, 并且对于$\lambda > 0$是有界的. 在$\lambda\to\infty, \hat{A}_\lambda|\psi\rangle \to \hat{A}|\psi\rangle$的意义上算符$\hat{A}_\lambda$逼近于$\hat{A}$. 为了证明这点, 首先记

$$\hat{A}_\lambda = \hat{B}_\lambda\hat{A}, \quad \hat{B}_\lambda := \frac{1}{2}i\lambda[(\hat{A}+i\lambda\mathbf{1})^{-1}-(\hat{A}-i\lambda\mathbf{1})^{-1}] \tag{3.13}$$

以及

$$1-\hat{B}_\lambda = \frac{1}{2}[(\hat{A}+i\lambda\mathbf{1})^{-1}+(\hat{A}-i\lambda\mathbf{1})^{-1}]\hat{A}$$

利用引理3.2.1, 对于$\forall|\phi\rangle \in \mathcal{D}(\hat{A})$, 我们有

$$\begin{aligned}\|(\mathbf{1}-\hat{B}_\lambda)|\phi\rangle\| &= \|\frac{1}{2}[(\hat{A}+i\lambda\mathbf{1})^{-1}+(\hat{A}-i\lambda\mathbf{1})^{-1}]\hat{A}|\phi\rangle\| \\ &\leq \frac{1}{\lambda}\|\hat{A}|\phi\rangle\| \to 0 \quad 当 \quad \lambda\to\infty\end{aligned}$$

由于$\mathcal{D}(\hat{A})$是稠密的, $\|\hat{B}\| \leq 1$, 当$\lambda \to \infty$时, 对于$\forall|\phi\rangle \in \mathcal{H}$我们有$\hat{B}_\lambda \to |\phi\rangle$. 最后, 对于$\forall|\psi\rangle \in \mathcal{D}(\hat{A})$, 令$|\phi\rangle = \hat{A}|\psi\rangle$, 根据式子(3.13), 可以得到

$$\hat{A}_\lambda|\psi\rangle \to \hat{A}|\psi\rangle \quad 其中 \quad \lambda\to\infty \quad 对于\forall|\psi\rangle\in\mathcal{D}(\hat{A}) \tag{3.14}$$

由于$\hat{A}_\lambda$是有界的, 我们能利用幂级数来定义指数形式$e^{i\hat{A}_\lambda}$. 接下来证明集合簇$\{e^{i\hat{A}_\lambda}, \lambda > 0\}$在下列条件的意义下是柯西序列, 对于$\forall|\psi\rangle \in \mathcal{D}(\hat{A})$, 当$\lambda, \lambda' \to \infty$时

$$\|(e^{i\hat{\lambda'}} - e^{i\hat{A}_\lambda})|\psi\rangle\| \to 0$$

为了证明这点, 我们将范数内的算符表示为积分导数:

$$e^{i\hat{A}_{\lambda'}} - e^{i\hat{A}_\lambda} = \int_0^1 \frac{\partial}{\partial s} e^{is\hat{A}_{\lambda'}} e^{i(1-s)\hat{A}_\lambda} ds$$

由于$\hat{A}_\lambda$是对称有界的, 所以它是自伴算符. 计算上式的范数

$$\begin{aligned}\|(e^{i\hat{A}_{\lambda'}} - e^{i\hat{A}_\lambda})|\psi\rangle\| &= \|\int_0^1 e^{is\hat{A}_{\lambda'}} e^{i(1-s)\hat{A}_\lambda} i(\hat{A}_{\lambda'} - \hat{A}_\lambda)|\psi\rangle ds\| \\ &\leq \int_0^1 \|e^{is\hat{A}_{\lambda'}} e^{i(1-s)\hat{A}_\lambda} i(\hat{A}_{\lambda'} - \hat{A}_\lambda)|\psi\rangle\| ds \\ &= \int_0^1 \|(\hat{A}_\lambda - \hat{A}_{\lambda'})|\psi\rangle\| ds = \|(\hat{A}_{\lambda'} - \hat{A}_\lambda)|\psi\rangle\|\end{aligned}$$

第一步是*Minkowski*不等式. 由于(3.14), 关系式(3.3.3)如下

*Cauchy*序列(3.3.3)表明了, 对于$\forall|\psi\rangle \in \mathcal{D}(\hat{A})$, 当$\lambda \to \infty$时, 矢量$e^{i\hat{A}_\lambda}|\psi\rangle$收敛于*Hilbert*空间中的某些元素. 对于$|\psi\rangle \in \mathcal{D}(\hat{A})$我们能定义

$$e^{i\hat{A}}|\psi\rangle := \lim_{\lambda\to\infty} e^{i\hat{A}_\lambda}|\psi\rangle$$

我们已经知道对于有界算符$\hat{A}$是显然成立的. 所以对于$\forall|\psi\rangle \in \mathcal{D}(\hat{A})$有$\|e^{i\hat{A}}|\psi\rangle\| \leq \||\psi\rangle\|$, 其中$\mathcal{D}(\hat{A})$在$\mathcal{H}$中稠密. 这样由于引理3.2.1, 我们能扩展$e^{i\hat{A}}$的定义到所有的$|\psi\rangle \in \mathcal{H}$. 这就

对任意自伴算符$\hat{A}$定义了指数函数$e^{\hat{A}}$.

我们可以把算符推广到傅里叶变换

定义 3.3.4 (算符的傅里叶变换) 假设$\hat{A}$是自伴算符, $\tilde{f}(\lambda)$是傅里叶逆变换函数, f可积, $\int |f(t)|dt < \infty$. 若f是实的, 则算符

$$\tilde{f}(\hat{A}) := \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} f(t) e^{i\frac{\hat{A}t}{\hbar}} dt$$

是有界自伴良定义的.

对于上式我们利用Riez积分(3.11)可以得到

$$e^{-i\hat{A}t} f(\hat{A}) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\lambda f(\lambda) e^{-i\lambda t} \mathrm{Im}(\hat{A} - \lambda \mathbf{1} - i0)^{-1}$$

有些书上也叫这个公式为谱定理. 最后介绍一个量子力学里最常用的公式, 谱分解公式.

公式 3.3.1 (谱分解) 记量子力学算符量为$\hat{A}$, 则有

$$\hat{A} = \sum_k |a_k\rangle a_k \langle a_k| + \int_{\sigma_e} |a\rangle a \langle a| da$$

3.3.3 习题

问题 3.3.1 证明$\mathrm{Null}\,(\hat{A} - \lambda \mathbf{1})$是线性空间.

问题 3.3.2 若$\hat{A}$是自伴算符, 证明$\hat{A}$不同的本征值的本征矢量相互正交.

问题 3.3.3 设$\hat{A}$是$\mathcal{H}$上的算符, 证明:若λ是$\sigma(\hat{A})$的一个聚点, 则$\lambda \in \sigma_e(\hat{A})$

问题 3.3.4 设$\hat{A}$是自伴算符. 证明

(1) 若W在$\hat{A}$下是不变的, 则$W^{\perp}$也是

(2) $\hat{A}$的本征矢量的生成空间和正交补在$\hat{A}$下是不变的

3.4 可观测力学量

3.4.1 力学量假设

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. 在之前的章节里, 自伴算符满足$\hat{A} = \hat{A}^{\dagger}$,

也就是$\langle\hat{A}\varphi|\psi\rangle = \langle\varphi|\hat{A}^{\dagger}\psi\rangle = \langle\varphi|\hat{A}\psi\rangle$, 并且在定理3.3.2中我们知道自伴算符的谱是实数, 而我们实际物理测量得到的值也是实数. 我们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. 这一系列对应性不得不使我们猜想, 量子力学里的力学量可能是自伴算符.

定义 3.4.1 *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[4]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[5].

假设 3.4.1 *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$\hat{A} = A^{\dagger}$.

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. 统一记作$\langle\psi|\hat{A}|\psi\rangle$, 力学量$\hat{A}$可以向后乘$\langle\psi|\overrightarrow{\hat{A}}|\psi\rangle$也可以向左乘$\langle\psi|\overleftarrow{\hat{A}}|\psi\rangle$. 在自伴算符$\hat{A}$的左右矢量对称关系下, 很自然的得到左右本征矢量$\langle a|, |a\rangle$的关系

定理 3.4.1 *若$|a\rangle$满足$\hat{A}|a\rangle = a|a\rangle$, 则$\langle a|\hat{A} = a\langle a|$*

证明 3.4.1 *假设$\hat{A}|a\rangle = a|a\rangle$, $\langle a|\hat{A} = a'\langle a|$*

$$\langle a|\hat{A}|a\rangle = a'\langle a|a\rangle$$
$$\langle a|\hat{A}|a\rangle = a\langle a|a\rangle$$

上下两式作差

$$(a - a')\langle a|a\rangle = 0$$

由于$\langle a|a\rangle > 0$, 得到$a = a'$, 得证

[4]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

[5]例如幺正幺模算符$\hat{U}$, 宇称算符$\hat{P}$, 时间反演算符$\hat{T}$等等, 都不可观测.

在定理3.4.2中我们知道, 自伴算符$\hat{A}$的本征矢量是Hilbert空间$\mathcal{H}$里的一组完备基. 若算符本征值全部在离散谱$\sigma_d(\hat{A})$里, 本征矢量满足Parseval等式.

$$\sum_a |a\rangle\langle a| = \mathbf{1}$$

其中$|a\rangle$是算符$\hat{A}$对应本征值a的本征态, 满足$\hat{A}|a\rangle = a|a\rangle$. 实际物理系统中并不是所有的力学量一定是离散谱,谱的形式实际上取决于系统的边界条件. 对于连续谱而言, 与离散谱不同的是把求和变成积分$\sum_a \mapsto \int_\Omega da$. 在之前(1.11)里, 推导了离散形式的Parseval等式. 这里只需要做上述替换就能得到连续形式的Parseval等式.

$$\int_\omega da|a\rangle\langle a| = \mathbf{1}$$

对于既有连续谱又有离散谱的力学量, 有如下定理

定理 3.4.2 (自伴算符完备性) *设$\hat{A}$是定义在Hilbert空间$\mathcal{H}$稠密定义域上的自伴算符. 若$\hat{A}$有离散谱σ_d和连续谱σ_e, 则$\hat{A}$存在可以生成Hilbert空间的$\mathcal{H}$正交归一本征基矢组.*

$$\sum_{a\in\sigma_d} |a\rangle\langle a| + \int_{\sigma_e} da'|a'\rangle\langle a'| = \mathbf{1}$$

3.4.2 测量假设

现在我们为自伴算符与可观测量做一些假设. 若果量子系统处于力学量$\hat{A}$的本征态$|a\rangle$, 对应的本征值为a, 在这个态$|a\rangle$下测量$\hat{A}$得到的结果将会是a. 反之, 如果测量力学量$\hat{A}$得到的结果总是a, 那么这个系统所处的态是$\hat{A}$的本征态$|a\rangle$.

这里指的注意的是对于力学量$\hat{A}$包含简并的情况, 也就是说对于自伴算符$\hat{A}$的某个本征值a对应有多个本征态$|a_1\rangle, |a_2\rangle, \cdots, |a_k\rangle$. 它们叠加所形成的任意叠加态也是$\hat{A}$的本征态, 并且对应的本征值是$a$. 假设$|\psi\rangle = \sum_k c_k|a_k\rangle$

$$\hat{A}|\psi\rangle = \hat{A}\sum_k c_k|a_k\rangle = \sum_k c_k\hat{A}|a_k\rangle = a\sum_k c_k|a_k\rangle = a|\psi\rangle$$

由此可以推断, 如果有多个态, 对于它们测量$\hat{A}$得到的结果总是a, 那么对于它们叠加形成的任何态测量$\hat{A}$任会得到a. 此外由于正交性定理, 如果对于两个态测量力学量$\hat{A}$总是分别得到两个不同的结果a, a', 那么这两个态$|a\rangle, |a'\rangle$相互正交.

当测量某个力学量$\hat{A}$时, 测量动作总是会引起量子系统的突变. 例如在实验2.4中我们为了测量电子不得不引入了光, 电子与光相互作用破坏了原来的相干性. 物理系统具有连

续重复性, 如果测量一次力学量$\hat{A}$后立即再次测量一次力学量$\hat{A}$, 第二次测量得到的结果必定与第一次测量得到的结果一样. 所以在第一次测量完成后第二次测量没有引入不确定性. 因此第一次测量结束后, 系统处于力学量$\hat{A}$的本征态, 他所属的本征值等于第一次测量的结果a. 这样我们不禁可以作出假设

假设 3.4.2 (测量) *力学量是自伴算符, 测量力学量$\hat{A}$总是导致系统突变到所测量力学量的本征态$|a\rangle$, 这个本征态$|a\rangle$所属的本征值a是测量的结果*

在之前的定理3.4.2中我们知道力学量$\hat{A}$的本征态$|a\rangle$可以构成Hilbert空间的一组完备基底$\{|a\rangle\}$.

$$\sum_{a\in\sigma_d}|a\rangle\langle a|+\int_{\sigma_e}da'|a'\rangle\langle a'|=\mathbf{1}$$

这样对于任意的态$|\psi\rangle$都可以用这一组基底线性表示.

定理 3.4.3 *对于任意态$|\psi\rangle$, 总可以用力学量$\hat{A}$的本征值去展开, 用$|a_k\rangle$标记力学量$\hat{A}$的离散本征态, 用$|a\rangle$标记力学量的连续本征态, 那么我们有*

$$|\psi\rangle=\sum_k|a_k\rangle\langle a_k|\psi\rangle+\int_{a\in\sigma_e}|a\rangle\langle a|\psi\rangle$$

如果记$c_k=\langle a_k|\psi\rangle, c(a)=\langle a|\psi\rangle$, 则可把上式写成

$$|\psi\rangle=\sum_k c_k|a_k\rangle+\int_{a\in\sigma_e}c(a)|a\rangle da$$

先考虑离散情况, 若算符$\hat{A}$的谱是离散谱, 对应的本征态也是离散的. 此时任意态用力学量$\hat{A}$的本征态展开可以写成

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

在之前的假设里, 对于态$|\psi\rangle$测量一次力学量$\hat{A}$得到的是$\hat{A}$的本征值a_k, 对应的本征态是$|a_k\rangle$. 很自然的就会问到:对于任意一个态, 测量完之后是如何从任意态$|\psi\rangle$突变到某个本征态$|a_k\rangle$里的. 任意的一个量子态$|\psi\rangle$测量完后就被破坏到力学量的某个本征态, 对此可以制备一个很多个$|\psi\rangle$态组成的系综(纯态). 对于这个系综的每一个$|\psi\rangle$可以进行一次力学量$\hat{A}$的测量, 我们将会得到不同个本征值. 对这些本征值进行统计, 可以得到不同本征值出现的频率次数. 若我们制备的系综足够大($n\to\infty$)不同本征值a_k ($k=1,2,\cdots$)出现的频率将会成为态$|\psi\rangle$突变到$|a_k\rangle$的概率. 类比之前的实验2.1, 很自然的会在假设3.4.2基础上进行进一步假设

假设 3.4.3 (测量概率) 对于任意给定的态$|\psi\rangle$, 测量力学量算符$\hat{A}$得到本征值a_k的概率为a_k对应本征态$|a_k\rangle$系数的模方$|c_k|^2$

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

测量力学量$\hat{A}$得到a_k的概率:

$$|\langle a_k|\psi\rangle|^2 = |c_k|^2$$

$$|\psi\rangle \longmapsto \begin{cases} |a_1\rangle \longmapsto a_1 \longmapsto \text{概率}|c_1|^2 \\ \cdots \\ |a_k\rangle \longmapsto a_k \longmapsto \text{概率}|c_k|^2 \\ \cdots \\ |a_m\rangle \longmapsto a_m \longmapsto \text{概率}|c_m|^2 \end{cases}$$

根据归一化条件:$\langle\psi|\psi\rangle = 1$

$$|c_1|^2 + |c_2|^2 + \cdots + |c_k|^2 + \cdots = 1$$

在本节最后一点里, 我们讨论一下本征值连续的情况. 任意选取一个态$|\psi\rangle$, 在这个态下测量力学量$\hat{A}$, $\hat{A}$具有连续的本征值和本征矢量, $\hat{A}|a\rangle = a|a\rangle$, 并且有Parseval等式:

$$\int_{a\in\sigma_e} |a\rangle\langle a| da = \mathbf{1}$$

这样任意的态$|\psi\rangle$可以展开为

$$|\psi\rangle = \int_{a\in\sigma_e} |a\rangle\langle a|\psi\rangle = \int_{a\in\sigma_e} c(a)|a\rangle$$

同样的我们可以取另一个任意态.

$$|\varphi\rangle = \int_{a'\in\sigma_e} d(a')|a'\rangle da'$$

上面两式内积可以得到

$$\langle\psi|\varphi\rangle = \iint_{a,a'\in\sigma_e} c^*(a)d(a')\langle a'|a\rangle \mathrm{d}a\mathrm{d}a'$$

我们考虑这样的一个单重积分$\int_{a'\in\sigma_e} d(a')\langle a'|a\rangle da'$, 这个积分只有在$a \neq a'$时不为零, 其余的情况全部是0. 而当$a = a'$时, 如果$\langle a'|a\rangle$有限, 则由于有限非零孤立奇点的积分是0, 最终的积分$\int_{a\in\sigma_e} d(a')\langle a'|a\rangle da' = 0$, $\langle\psi|\varphi\rangle = 0$. 此时任意两个态都是正交的, 这显然是荒谬的. 为了避免这点, 我们只能让$\langle a'|a\rangle$是一个无穷大来保证这个积分$\int_{a\in\sigma_e}\langle a'|a\rangle da'$有限.

类比δ函数的性质

$$\int_{x_0\in\Omega} f(x)\delta(x - x_0)dx = f(x_0)$$

似乎要求$\langle a'|a\rangle$正比与δ函数是合理的. 若$\langle a'|a\rangle = \delta(a-a')$, 则积分值

$$\begin{aligned}\langle\psi|\varphi\rangle &= \iint_{a,a'\in\sigma_e} c^*(a)d(a')\langle a'|a\rangle \mathrm{d}a\mathrm{d}a' \\ &= \iint_{a,a'\in\sigma_e} c^*(a)d(a')\delta(a-a')\mathrm{d}a\mathrm{d}a' \\ &= \int_{a\in\sigma_e} c^*(a)d(a)\mathrm{a}\end{aligned}$$

是非零的值. 而要求$\langle a'|a\rangle = \delta(a-a')$, 与我们之前所要求的正交归一化条件相矛盾. 我们在离散谱里要求的正交归一化条件是$\langle a_i|a_j\rangle = \delta_{ij}$. 这不得不使我们去扩充我们对正交归一化的定义

定义 3.4.2 (连续谱的归一化) *自伴算符$\hat{A}$的连续谱σ_e所对应本征矢量$|a\rangle$的正交归一化条件为*

$$\langle a'|a\rangle = \delta(a-a')$$

其中$\delta(a-a')$是δ函数.

连续谱的正交归一化条件意味着连续谱的本征值a对应的本征态$|a\rangle$的内积范数$\||a\rangle\| = \sqrt{\langle a|a\rangle}$是无穷大的. 为了避免这个无穷大矢量长度, 我们不得不放宽之前2.3.3里对归一化的定义. 连续谱所对应的态矢量$|a\rangle$归一化条件被扩充成

$$\int_{a'\in\sigma_e}\langle a'|a\rangle \mathrm{d}a' = 1$$

这样就建立起了力学量与物理测量之间的联系

3.4.3 统计学解释

在经典力学中, 对于任意特定状态$|\psi\rangle$的系统E, 测量一次力学量总能得到一个值, 并且这个值是实的. 在之前构建的数学里, 任意选取一个力学量和两个态$|\psi\rangle, |\varphi\rangle$, 我们知道$\langle\varphi|\hat{A}|\psi\rangle$是一个复数, 这个数与实际测量值并不是十分类似:

1. $\langle\varphi|\hat{A}|\psi\rangle$与两个态有关, 但实际测量值只与一个态有关

2. 不是实数

3. $\langle\varphi|\hat{A}|\psi\rangle$包含一个模为1的不确定因子

$$\begin{aligned}|\psi\rangle &\longmapsto e^{i\phi_1}|\psi\rangle\\ |\varphi\rangle &\longmapsto e^{i\phi_2}|\varphi\rangle\\ \langle\varphi|\hat{A}|\psi\rangle &\longmapsto e^{i(\phi_1-\phi_2)}\langle\varphi|\hat{A}|\psi\rangle\end{aligned}$$

若我们取$\langle\varphi| = \langle\psi|$, 上面这三个问题自动消除掉. 此时得到的数$\langle\psi|\hat{A}|\psi\rangle$是实数, 并且若$|\psi\rangle$满足归一化条件, 它唯一的被确定下来.

在章节3.4.2里, 我们知道对于$|\psi\rangle$态, 测量力学量$\hat{A}$得到的结果必定是$\hat{A}$的一系列本征值中的某一个, 至于是哪一个则是以一定概率出现的, 概率值则是$|\psi\rangle$展开系数的模方$|c_k|^2$(这里$|\psi\rangle$已经归一化了). 利用这一点我们可以计算得到

$$\begin{aligned}\langle\psi|\hat{A}|\psi\rangle &= \sum_{i,j}\langle\psi_i|c_i^*\hat{A}c_j|\psi_j\rangle\\ &= \sum_{i,j}c_i^*c_ja_j\langle\psi_i|\psi_j\rangle\\ &= \sum_{i,j}c_i^*c_ja_j\delta_{ij} = \sum_i|c_i|^2a_i\end{aligned}$$

从这一点可以看到$\langle\psi|\hat{A}|\psi\rangle$的是算符$\hat{A}$的各个本征值的加权平均, 其权重正好是$|\psi\rangle$态的展开系数模方$|c_i|^2$. 对于单个$|\psi\rangle$单纯谈论$\langle\psi|\hat{A}|\psi\rangle$是没有意义的, 因为测量后的结果必定是$\hat{A}$的某个本征值, 测量后的态一定会被破坏到本征值对应的本征态. 为了解释$\langle\psi|\hat{A}|\psi\rangle$这个值, 不得不引入系综, 这样得到一个自然的结论

定理 3.4.4 (平均值) *对于处于$|\psi\rangle$态的纯态系综($|\psi\rangle$满足归一化条件), 测量力学量$\hat{A}$足够多次, 那么所得全部结果的平均值是$\langle\psi|\hat{A}|\psi\rangle$. 若$|\psi\rangle$不是归一化的, 则测量结果的平均值正比于$\langle\psi|\hat{A}|\psi\rangle$*

我们接下来讨论两个力学量$\hat{A},\hat{B}$在平均值意义下的关系.

预先给定的两个力学量$\hat{A},\hat{B}$, 我们用符号$\langle\hat{A}\rangle,\langle\hat{B}\rangle$分别去表示$\hat{A},\hat{B}$的平均值. 对于$\hat{A}+\hat{B}$, 很容易得到$\langle\hat{A}+\hat{B}\rangle = \langle\hat{A}\rangle+\langle\hat{B}\rangle$; 而对于$\hat{A}\hat{B}$, 一般而言$\langle\hat{A}\hat{B}\rangle \neq \langle\hat{A}\rangle\langle\hat{B}\rangle$.

$$\begin{aligned}\langle\hat{A}\hat{B}\rangle &= \sum_{i,j}\langle\psi|\hat{A}|a_i\rangle\langle b_j|\hat{B}|\psi\rangle\\ &= \sum_{ij}a_ib_j\langle\psi|a_i\rangle\langle b_j|\psi\rangle\end{aligned} \tag{3.15}$$

$$\begin{aligned}\langle\hat{A}\rangle\langle\hat{B}\rangle &= \sum_{ij}\langle\psi|\hat{A}|a_i\rangle\langle a_i|b_j\rangle\langle b_j|\hat{B}|\psi\rangle\\ &= \sum_{ij}a_ib_j\langle\psi|a_i\rangle\langle a_i|b_j\rangle\langle b_j|\psi\rangle\end{aligned} \tag{3.16}$$

式(3.15)与(3.16)作差[6], 合并指标可以得到

$$\langle \hat{A}\hat{B}\rangle - \langle \hat{A}\rangle\langle \hat{B}\rangle = \sum_{ij} a_i b_j \langle\psi|a_i\rangle\langle b_j|\psi\rangle(\langle a_i|b_j\rangle - \langle a_i|\psi\rangle\langle\psi|b_j\rangle) \neq 0$$

可以看出$\langle \hat{A}\hat{B}\rangle \neq \langle \hat{A}\rangle\langle \hat{B}\rangle$. 类比数学统计里$\overline{ab} \neq \bar{a}\bar{b}$, 把$\langle\psi|\hat{A}|\psi\rangle$定义成算符$\hat{A}$的平均值在数学上也是合理的

3.4.4 力学量的相容性

在讨论两个力学量相容性之前, 我们先考虑一个新的概念, 叫做共同本征态.

一个物理测量对于某一特定的态有**一个特定的值**的说法, 在量子力学里只有在本征态的情况下才是允许的. 对于两个力学量$\hat{A}, \hat{B}$, 也是可以存在这样的一个特定态, 这个特定态使得这两个力学量总是有特定的值.

定义 3.4.3 *若对于态$|ab\rangle$, 力学量$\hat{A}, \hat{B}$满足*

$$\hat{A}|ab\rangle = a|ab\rangle$$
$$\hat{B}|ab\rangle = b|ab\rangle$$

则称$|ab\rangle$为力学量$\hat{A}, \hat{B}$的共同本征态

这意味着, 对于共同本征态而言, 测量力学量$\hat{A}, \hat{B}$无论是谁先测量亦或是谁后测量, 最终得到的值必定是两个力学量对应的本征值的乘积. 本征值的乘积与先后顺序无关, 也就是说在共同本征态的条件下, 力学量$\hat{A}, \hat{B}$之间的测量顺序没有引入不确定性.

我们定义对易子$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$. 很显然有如下结论

推论 3.4.1 *若两个力学量$\hat{A}, \hat{B}$满足$[\hat{A}, \hat{B}] = 0$, 则共同本征态一定存在*

证明 3.4.2 *假设自伴算符$\hat{A}, \hat{B}$存在共同本征态$|ab\rangle$, 使得*

$$\hat{A}|ab\rangle = a|ab\rangle$$
$$\hat{B}|ab\rangle = b|ab\rangle$$

则有

$$\hat{A}\hat{B}|ab\rangle = \hat{A}b|ab\rangle = ab|ab\rangle = ba|ab\rangle = \hat{B}\hat{A}|ab\rangle$$

[6]上述方程一步使用了Parseval等式$\sum_i |a_i\rangle\langle a_i| = \mathbf{1}, \sum j|b_j\rangle\langle b_j| = \mathbf{1}$

或者写成

$$(\hat{A}\hat{B}-\hat{B}\hat{A})|ab\rangle=[\hat{A},\hat{B}]|ab\rangle=0$$

若$[\hat{A},\hat{B}]\neq 0$, 则上述方程必定只能存在零解, 而我们要求量子态的范数$\||ab\rangle\|>0$. 所以只有$[\hat{A},\hat{B}]=0$上述方程才存在非零解. 得证.

值得注意, 若两个力学量$\hat{A},\hat{B}$不对易, 则共同本征态也可能存在, 但这是非常少见的情况[7]. 对于一般的态$|\psi\rangle$而言$\hat{A}\hat{B}|\psi\rangle\neq\hat{B}\hat{A}|\psi\rangle$, 原因在于我们对$\hat{A}$进行一次测量时破坏原来的态$|\psi\rangle$至$\hat{A}$的某个本征态$|a_i\rangle$, 而$|a_i\rangle$不属于$\hat{B}$的本征态集$\{|b_j\rangle\}$. 当我们再一次测量时, 将会破坏$|a_i\rangle$至某一个$|b_j\rangle$. 而这个过程存在不确定性.

这就是说, 若两个力学量$\hat{A},\hat{B}$相互对易, 则这两个力学量可以同时具有确定值(本征值). 若$\hat{A},\hat{B}$相互不对易, 则一般来说这两个力学量不能同时有确定值[8].

我们现在考虑一个实际的问题, 如何求解共同本征态. 这个问题分为两个情况考虑, 先设$[\hat{A},\hat{B}]=0$且有

$$\hat{A}|a_n\rangle=a_n|a_n\rangle$$

(1) 先设$\hat{A}_n$的本征态不存在简并. 由于$[\hat{A},\hat{B}]=0$

$$\hat{A}(|\hat{B}|a_n\rangle)=\hat{B}\hat{A}|a_n\rangle=a_n\hat{B}|a_n\rangle$$

也就是说$\hat{B}|a_n\rangle$也是$\hat{A}|a\rangle$的本征态, 本征值为a_n, 但是$\hat{A}$不简并, 所以$\hat{B}|a_n\rangle$与$|a_n\rangle$必须是线性相关的, 最多只能差一个常数因子, 记作b_n

$$\hat{B}|a_n\rangle=b_n|a_n\rangle$$

这样$|a_n\rangle$就已经是$\hat{A}$和$\hat{B}$的共同本征态, 为了方便记作$|a_nb_n\rangle$, 本征值分别是a_n和b_n.

(2) 再设$\hat{A}_n$有f_n重简并, f_n重简并态的集合构成的f_n维子空间记作E'

$$\hat{A}|a_{n\alpha}\rangle=a_n|a_{n\alpha}\rangle\quad(\alpha=1,2,\ldots,f_n)$$

假设$\{|a_{n\alpha}\rangle\}$已经正交归一化, 则$\{|a_{n\alpha}\rangle\},(\alpha=1,2,\ldots,f_n)$构成$E'$的一组基底. 但$|a_{n\alpha}\rangle$并不一定就是$\hat{B}$的本征态. 考虑到

$$\hat{A}\hat{B}|a_{n\alpha}\rangle=\hat{B}\hat{A}|a_{n\alpha}\rangle=a_n\hat{B}|a_{n\alpha}\rangle$$

[7]例如角动量算符$\hat{l}_x$与$\hat{l}_y$不对易, $[\hat{l}_x,\hat{l}_y]=i\hbar\hat{l}_z$, 但$|0\rangle$是$\hat{l}_x,\hat{l}_y$的共同本征态, 本征值为0

[8]对于一些特例$\langle[\hat{A},\hat{B}]\rangle=0$的情况可能存在个别的共同本征态, 但这些共同的本征态并不完备, 这点将在下节解释

也就是说$\hat{B}|a_{n\alpha}\rangle$仍为$\hat{A}$的本征态, 本征值为$a_n$. 因此$\hat{B}|a_{n\alpha}\rangle$可表示为

$$\hat{B}|a_{n\alpha}\rangle = \sum_{\alpha'=1}^{f_n} b_{\alpha'\alpha}|a_{n\alpha'}\rangle$$

其中$b_{\alpha'\alpha} = \langle a_{n\alpha'}|a_{n\alpha}\rangle$ 可以看出, 一般来说, $|a_{n\alpha}\rangle$还不是$\hat{B}$的本征态, 但我们可以对$\alpha = 0, 1, \ldots, f_n$这些简并本征态进行叠加, 叠加后的矢量$|ab\rangle$表示为

$$|ab\rangle = \sum_{\alpha=1}^{f_n} c_\alpha |a_{n\alpha}\rangle$$

显然

$$\hat{A}|ab\rangle = \sum_{\alpha=1}^{f_n} c_\alpha \hat{A}|a_{n\alpha}\rangle = \sum_{\alpha=1}^{f_n} c_\alpha a_n |a_{n\alpha}\rangle = a_n|ab\rangle$$

即$|ab\rangle$仍然是$\hat{A}$的本征态, 而且对应的本征值为a_n. 我们现在假设$|ab\rangle$是$\hat{B}$的本征态

$$\hat{B}|ab\rangle = \sum_{\alpha=1}^{f_n} c_\alpha \hat{B}|a_{n\alpha}\rangle = \sum_{\alpha',\alpha=1}^{f_n} c_n b_{\alpha'\alpha}|a_{n\alpha'}\rangle$$

而我们假设了

$$\hat{B}|ab\rangle = b|ab\rangle = b\sum_{\alpha'=1}^{f_n} c_{\alpha'}|a_{n\alpha'}\rangle$$

比较上式得

$$\sum_{\alpha=1}^{f_n} c_\alpha b_{\alpha'\alpha} = b c_{\alpha'}$$

利用Kroneck记号可以写成

$$\sum_{\alpha=1}^{f_n} (b_{\alpha'\alpha} - b\delta_{\alpha'\alpha})c_\alpha = 0 \tag{3.17}$$

左边是$f_n \times f_n$的行列式, 这是关于b的f_n次代数方程, 只需要求解这f_n次代数方程的根即可[9]. 这f_n个根分别记作$b_\beta(\beta = 1, 2, \ldots, f_n)$.

(i) 若b_β无重根, 把每个根b_β代入(3.17)里可以计算出一组$c_{\beta\alpha}, (\alpha =, 1, 2, \ldots, f_n)$. 这

[9]由于$\hat{B}^\dagger = \hat{B}$, 所以$B_{\alpha'\alpha} = B^*_{\alpha\alpha'}$, 再利用有限群表示, 可以证明这$f_n$次代数方程的根是存在的

样就可以得到组合共同本征态

$$|a_n b_\beta\rangle = \sum_{\alpha=1}^{f_n} c_{\beta\alpha}|a_{n\alpha}\rangle$$

这样的共同本征态一共有f_n个$(\beta = 1, 2, \ldots, f_n)$

$$\begin{aligned}\hat{A}|a_n b_\beta\rangle &= a_n|a_n b_\beta\rangle \\ \hat{B}|a_n b_\beta\rangle &= b_n|a_n b_\beta\rangle \quad (\beta = 1, 2, \ldots, f_n)\end{aligned}$$

此时共同本征态将$\hat{A}$的简并完全解除.

(ii) 若b_β有重根, 则简并还没有完全解除. 此时可以寻找与$\hat{A}, \hat{B}$都对易的另外力学量$\hat{C}, \cdots$, 求这一组力学量$(\hat{A}, \hat{B}, \hat{C})$的共同本征态, 直到简并态完全可以标记清楚为止.

定义 3.4.4 (完全集) 设有一组彼此独立而且相互对易的自伴算符集合$X = \{\hat{A}, \hat{B}, \ldots\}$, 它们的共同本征态记作$|\alpha\rangle$, ξ表示一组共同的量子数$(a, b, \ldots)$. 若给定一组量子数$\alpha = a, b, \ldots$后, 就能确定体系的唯一一个可能状态, 则我们称力学量集合X为对易力学量完全集

力学量的完全集的共同本征态完全的解除掉了简并, 任何一个共同本征态都被标定上了不同的量子数. 力学量本征态的集合构成$\mathcal{H}$的一组完备基底. 按照态叠加原理, 系统的任何一个状态$|\psi\rangle$均可以用系统的任何一个力学量完全集$|\alpha\rangle$来展开

$$|\psi\rangle = \sum_\alpha c_\alpha|\alpha\rangle$$

利用正交性, 很容易得到展开系数$c_\alpha = \langle\psi|\alpha\rangle$. $|c_\alpha|^2$表示在$|\psi\rangle$态下, 测量力学量$\hat{A}$得到a_α的概率. 这点在上一节统计解释里表述了. 对于连续情况, 则$\sum_\alpha \longmapsto \int d\alpha$, 相应的展开系数模方代表概率密度.

3.4.5 不确定关系

假设系统出于力学量$\hat{A}$的本征态, 则对它测量$\hat{A}$时, 将得到一个确切值, 并不会出现任何不确定性(涨落). 很自然的就会思考, 在这个量子态下去测量另一个力学量$\hat{B}$, 是否也能得到一个确定的值?

定理 3.4.5 (海森堡不确定性)
记$\Delta A = \sqrt{\langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle}, \Delta B = \sqrt{\langle(\hat{B} - \langle\hat{B}\rangle)^2\rangle}$, 则有下列不等式

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\hat{A}, \hat{B}]\rangle|$$

证明 3.4.3 (方法一) 记$\Delta\hat{A} = \hat{A} - \langle\hat{A}\rangle, \Delta\hat{B} = \hat{B} - \langle\hat{B}\rangle$

则有$[\Delta\hat{A}, \Delta\hat{B}] = [\hat{A} - \langle\hat{A}\rangle, \hat{B} - \langle\hat{B}\rangle] = [\hat{A}, \hat{B}]$

$$\begin{aligned} \Delta A &= \sqrt{\langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle} \\ &= \sqrt{\langle\psi|(\Delta\hat{A})^2|\psi\rangle} \\ &= \sqrt{\langle\Delta\hat{A}\psi|\Delta\hat{A}\psi\rangle} \\ &= \||\Delta\hat{A}\psi\rangle\| \end{aligned} \tag{3.18}$$

$$\begin{aligned} \Delta B &= \sqrt{\langle(\hat{B} - \langle\hat{B}\rangle)^2\rangle} \\ &= \sqrt{\langle\psi|(\Delta\hat{B})^2|\psi\rangle} \\ &= \sqrt{\langle\Delta\hat{B}\psi|\Delta\hat{B}\psi\rangle} \\ &= \||\Delta\hat{B}\psi\rangle\| \end{aligned} \tag{3.19}$$

上式(3.18)与(3.19)相乘

$$\begin{aligned} \Delta A \Delta B &= \||\Delta\hat{A}\psi\rangle\|\||\Delta\hat{B}\psi\rangle\| \\ &\geq |\langle\Delta\hat{A}\psi|\Delta\hat{B}\psi\rangle| \geq |\mathrm{Im}\langle\Delta\hat{A}\psi|\Delta\hat{B}\psi\rangle| \\ &= |\frac{\langle\Delta\hat{A}\psi|\Delta\hat{B}\psi\rangle - \langle\Delta\hat{B}\psi|\Delta\hat{A}\psi\rangle}{2i}| = \frac{|\langle\psi|[\Delta\hat{A}, \Delta\hat{B}]|\psi\rangle|}{2} = \frac{|\langle[\hat{A}, \hat{B}]\rangle|}{2} \end{aligned}$$

海森堡不确定性关系:$\Delta A \Delta B \geq \frac{|\langle[\hat{A},\hat{B}]\rangle|}{2}$, 得证

证明 3.4.4 (方法二) 考虑这样一个关于$\xi \in \mathbb{R}$的函数$I(\xi)$:

$$\begin{aligned} I(\xi) &= \langle(\xi\hat{A} + i\hat{B})\psi|(\xi\hat{A} + i\hat{B})\psi\rangle \\ &= \xi^2\langle\hat{A}\psi|\hat{A}\psi\rangle + i\xi\langle\hat{A}\psi|\hat{B}\psi\rangle - i\xi\langle\hat{B}\psi|\hat{A}\psi\rangle + \langle\hat{B}\psi|\hat{B}\psi\rangle \\ &= \xi^2\langle\psi|\hat{A}^2|\psi\rangle + i\xi\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle + \langle\psi|\hat{B}^2|\psi\rangle \end{aligned}$$

我们引入新的自伴算符$\hat{C}$

$$\hat{C} = i[\hat{B}, \hat{A}]$$

显然这个算符的平均值$\langle\hat{C}\rangle$在任何情况下都是实数

$$\langle\hat{C}\rangle^* = -i\langle[\hat{B}, \hat{A}]\rangle^* = -i\langle\hat{B}\psi|\hat{A}\psi\rangle^* + i\langle\hat{A}\psi|\hat{B}\psi\rangle^* = i(\langle\hat{B}\psi|\hat{A}\psi\rangle - \langle\hat{A}\psi|\hat{B}\psi\rangle) = \langle\hat{C}\rangle$$

这样$I(\xi)$可以表示为

$$I(\xi) = \xi^2\langle\hat{A}^2\rangle - \xi\langle\hat{C}\rangle + \langle\hat{B}^2\rangle \tag{3.20}$$

在之前的态矢量的正定性假设(2.25)里可以得到, 对于任意ξ, $I(\xi) = \langle(\xi\hat{A} + i\hat{B})\psi|(\xi\hat{A} + i\hat{B})\psi\rangle \geq 0$. 所以在(3.20)里有判别式$\Delta \leq 0$

$$\langle\hat{C}\rangle^2 - 4\langle\hat{A}^2\rangle\langle\hat{B}^2\rangle \leq 0$$

所以得到

$$\langle\hat{A}^2\rangle\langle\hat{B}^2\rangle \geq \frac{\langle\hat{C}\rangle^2}{4}$$

由于这个等式对于任意自伴算符$\hat{A}, \hat{B}$都成立, 取$\hat{A}, \hat{B}$为$\Delta\hat{A}, \Delta\hat{B}$. 得到

$$\sqrt{\langle\Delta\hat{A}^2\rangle\langle\Delta\hat{B}^2\rangle} \geq \frac{|\langle i[\Delta\hat{A}, \Delta\hat{B}]\rangle|}{2} = \frac{|\langle[\hat{A}, \hat{B}]\rangle|}{2}$$

这就是两个力学量涨落应该满足的关系. 从不确定关系3.4.5可以看出, 若两力学量$\hat{A}, \hat{B}$不对易, 则一般说来, $\Delta\hat{A}$与$\Delta\hat{B}$不能同时为零, 即$\hat{A}$与$\hat{B}$不能同时有确定的测量值($\langle[\hat{A}, \hat{B}]\rangle = 0$可能是例外), 或者说它们不能有共同的本征态.

反之, 若两个自伴算符$\hat{A}, \hat{B}$对易, 即$[\hat{A}, \hat{B}] = 0$, 则可以找到这样的态, 使得$\Delta\hat{A} = 0$和$\Delta\hat{B} = 0$同时满足, 也就是可以求出它们的共同本征态.

3.4.6 习题

问题 3.4.1 如果两个力学量有共同本征态, 则它们是否对易?

问题 3.4.2 若两个力学量不对易, 是否它们一定没有共同本征态?

问题 3.4.3 若两个自伴算符对易, 是否在所有态下它们都同时具有确定的值?

3.5 算符对易关系

这里我们先对算符的运算做一个总结

1. 线性算符的和算符$\hat{A}, \hat{B}$的和记作为$(\hat{A} + \hat{B})$, 定义为

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle$$

$|\psi\rangle$是任意态矢量.

算符加法满足交换律和结合律

$$\hat{A}+\hat{B}=\hat{B}+\hat{A}$$
$$\hat{A}+(\hat{B}+\hat{C})=(\hat{A}+\hat{B})+\hat{C}$$

算符之和仍为线性算符.

2. 算符的乘积算符$\hat{A},\hat{B}$的乘积记作$\hat{A}\hat{B}$, 定义为

$$(\hat{A}\hat{B})|\psi\rangle=\hat{A}(\hat{B}|\psi\rangle)$$

即$\hat{A}\hat{B}$对任意态的运算结果等于先用$\hat{B}$运算得$\hat{B}|\psi\rangle$, 然后用$\hat{A}$对$(\hat{B}|\psi\rangle)$运算得到的结果. 一般来说算符的乘积不满足交换律, 即

$$\hat{A}\hat{B}\neq\hat{B}\hat{A}$$

这里是算符运算规则与数的运算规则唯一不同之处.

在实际量子力学计算中经常会遇到$\hat{A}\hat{B}-\hat{B}\hat{A}$的形式, 为了方便起见, 我们引入对易符号$[\cdot,\cdot]$. $[\cdot,\cdot]$表示$\mathfrak{g}\times\mathfrak{g}\mapsto\mathfrak{g}$的映射, 这里$\mathfrak{g}$表示算符构成的线性空间$\mathfrak{g}$.

量子力学里对于算符$\hat{A},\hat{B}$我们定义

$$[\hat{A},\hat{B}]=\hat{A}\hat{B}-\hat{B}\hat{A}$$

这样定义后的量子力学算符实际上构成了李代数.

定义 3.5.1 (有限维实(复)李代数) 具有映射$[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\mapsto\mathfrak{g}$的有限维实(复)线性空间$\mathfrak{g}$叫做有限维实(复)李代数, 其中映射$[\cdot,\cdot]$对于线性空间的元素$\hat{A}_1,\hat{A}_2,\cdots,\hat{A}_n$满足下列三个性质:

1. $[\cdot,\cdot]$是双线性映射
2. $[\cdot,\cdot]$满足反对称性:$[\hat{A}_i,\hat{A}_j]=-[\hat{A}_j,\hat{A}_i]$对于所有的$\hat{A}_i,\hat{A}_j\in\mathfrak{g}$都成立.
3. 对于所有的$\hat{A}_i,\hat{A}_j,\hat{A}_k\in\mathfrak{g}$满足雅克比恒等式:

$$[\hat{A}_i,[\hat{A}_j,\hat{A}_k]]+[\hat{A}_j,[\hat{A}_k,\hat{A}_i]]+[\hat{A}_k,[\hat{A}_i,\hat{A}_j]]=0$$

量子力学列定义的$[\hat{A},\hat{B}]=\hat{A}\hat{B}-\hat{B}\hat{A}$不难验证有下列一系列的性质:

1. $[\hat{A},\hat{B}]=-[\hat{B},\hat{A}]$
2. $[\hat{A},\hat{A}]=0$
3. $[\hat{A},c]=0 \quad (c\in\mathbb{C})$
4. $[\hat{A},\hat{B}+\hat{C}]=[\hat{A},\hat{B}]+[\hat{A},\hat{C}]$
5. $[\hat{A},\hat{B}\hat{C}]=\hat{B}[\hat{A},\hat{C}]+[\hat{A},\hat{B}]\hat{C}$
6. $[\hat{A}\hat{B},\hat{C}]=\hat{A}[\hat{B},\hat{C}]+[\hat{A},\hat{C}]\hat{B}$
7. $[\hat{A},[\hat{B},\hat{C}]]+[\hat{B},[\hat{C},\hat{A}]]+[\hat{C},[\hat{A},\hat{B}]]=0$

量子力学里的算符实质上构成了李代数, 而量子对易关系正是一种李乘法$[\cdot,\cdot]$. 这里我们介绍一个李代数里常用的公式

公式 3.5.1 (Baker-Campbell-Hausdorff公式) 如果线性算符$\hat{A},\hat{B}$以及它们的对易$\hat{C}=[\hat{A},\hat{B}]$满足

$$[\hat{A},\hat{C}]=[\hat{B},\hat{C}]=0$$

则有

$$e^{\hat{A}+\hat{B}}=e^{\hat{A}}e^{\hat{B}}e^{-\frac{1}{2}\hat{C}}$$

证明 3.5.1 由于$[\hat{A},\hat{B}]=\hat{C},\ [\hat{A},\hat{C}]=[\hat{B},\hat{C}]=0$

所以$\hat{A},\hat{B},\hat{C}$构成的一个3维李代数.

定义单参数t的算符函数$\hat{F}(t)$

$$\hat{F}(t)=e^{t(\hat{A}+\hat{B})}$$

由于$\hat{A},\hat{B},\hat{C}$构成完备的代数空间, 所以$\hat{F}(t)$一定可以表示为

$$\hat{F}(t)=e^{x(t)\hat{A}}e^{y(t)\hat{B}}e^{z(t)\hat{C}} \tag{3.21}$$

其中$x(t),y(t),z(t)$满足初始条件

$$x(0)=y(0)=z(0)=0$$

我们有

$$\frac{\mathrm{d}\hat{F}(t)}{\mathrm{d}t}=(\hat{A}+\hat{B})\hat{F}(t)$$

由于方程(3.21), 又有

$$\begin{aligned}\frac{\mathrm{d}\hat{F}(t)}{\mathrm{d}t} &= (\hat{A}+\hat{B})\hat{F}(t)\\ &= \left(\frac{\mathrm{d}x}{\mathrm{d}t}\hat{A}+\frac{\mathrm{d}y}{\mathrm{d}t}e^{x(t)\hat{A}}\hat{B}e^{-x(t)\hat{A}}+\frac{\mathrm{d}z}{\mathrm{d}t}[\hat{A},\hat{B}]\right)\hat{F}(t)\\ &= \left(\frac{\mathrm{d}x}{\mathrm{d}t}\hat{A}+\frac{\mathrm{d}y}{\mathrm{d}t}(\hat{B}+x(t)\hat{C})+\frac{\mathrm{d}z}{\mathrm{d}t}\hat{C}\right)\hat{F}(t)\\ &= \left(\frac{\mathrm{d}x}{\mathrm{d}t}\hat{A}+\frac{\mathrm{d}y}{\mathrm{d}t}\hat{B}+(\frac{\mathrm{d}z}{\mathrm{d}t}+\frac{\mathrm{d}y}{\mathrm{d}t}x)\hat{C}\right)\hat{F}(t)\end{aligned}$$

比较两式得到

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t}&=1\\ \frac{\mathrm{d}y}{\mathrm{d}t}&=1\\ \frac{\mathrm{d}z}{\mathrm{d}t}+\frac{\mathrm{d}y}{\mathrm{d}t}x&=0\end{aligned}$$

求得

$$\begin{aligned}x(t)&=t\\ y(t)&=t\\ z(t)&=-\frac{t^2}{2}\end{aligned}$$

令$t=1$即得所证

除了构造算符函数以外, 还可以直接利用算符函数的泰勒展开去证明, 留作练习

除了Baker-Campbell-Hausdorff公式外量子力学里还常用这样的一个展开式

公式 3.5.2 *$\hat{A},\hat{B}$是量子力学算符, 则有下列关系式*

$$e^{\hat{A}}\hat{B}e^{-\hat{A}}=\hat{B}+[\hat{A},\hat{B}]+\frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]]+\frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]]+\cdots \tag{3.22}$$

证明 3.5.2 *构造单参数算符函数$\hat{F}(t)=e^{t\hat{A}}\hat{B}e^{-t\hat{A}}$*

则有

$$\begin{aligned}\frac{\mathrm{d}\hat{F}(t)}{\mathrm{d}t} &= e^{t\hat{A}}[\hat{A},\hat{B}]e^{-t\hat{A}}\\ \frac{\mathrm{d}^2\hat{F}(t)}{\mathrm{d}t^2} &= e^{t\hat{A}}[\hat{A},[\hat{A},\hat{B}]]e^{-t\hat{A}}\\ &\cdots\cdots\end{aligned}$$

所以得到$\hat{F}(t)$的*Taylor*展开:

$$\hat{F}(t)=e^{t\hat{A}}\hat{B}e^{-t\hat{A}}=\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{\mathrm{d}^k\hat{F}(t)}{\mathrm{d}t^k}\right)\Bigg|_{t=0}t^k=\sum_{k=0}^{\infty}\frac{1}{k!}[\hat{A},[\hat{A},\cdots,[\hat{A},\hat{B}]]]t^k$$

取$t=1$即得证

最后介绍一个在散射问题里求解格林函数常用的公式:

公式 3.5.3

$$(\hat{A}+\hat{B})^{-1}=\hat{A}^{-1}-\hat{A}^{-1}\hat{B}(\hat{A}+\hat{B})^{-1}$$

证明留给读者

3.5.1 习题

问题 3.5.1 利用泰勒展开证明*Baker-Campbell-Hausdorff*公式3.5.1.

问题 3.5.2 证明下列公式

$$\begin{aligned}[\hat{A},\hat{B}\hat{C}] &= \hat{B}[\hat{A},\hat{C}]+[\hat{A},\hat{B}]\hat{C}\\ [\hat{A}\hat{B},\hat{C}] &= \hat{A}[\hat{B},\hat{C}]+[\hat{A},\hat{C}]\hat{B}\end{aligned}$$

问题 3.5.3 若算符$\hat{B}$与算符$[\hat{A},\hat{B}]$对易, 证明:

$$[\hat{A},\hat{B}^n]=n\hat{B}^{n-1}[\hat{A},\hat{B}]$$

问题 3.5.4 已知$[\hat{A},\hat{B}]=c, c\in\mathbb{C}$, 证明

$$[\hat{A}^m,\hat{B}^n]=\sum_{k=1}^{\infty}\frac{m!n!}{k!(m-k)!(n-k)!}c^k\hat{A}^{m-k}\hat{B}^{n-k}$$

Chapter 4

量子化与表象

到现在为止我们建立起了一套完整的描述量子力学的数学工具, 描述系统状态的是态矢量, 描述系统力学量的是自伴算符, 测量对应的是力学量的本征值和本征态, 测量概率对应的是叠加态展开系数模方. 自然地就会好奇为什么要去建立这样的一套数学, 从经典力学是如何过渡到量子力学的. 这节主要讲述经典力学与量子力学的联系.

4.1 基本量子化条件

在经典力学里描述s自由度系统的力学量是哈密顿$H = H(q_\alpha, p_\alpha, t)$, 任意力学量都是广义坐标$q_\alpha$和广义动量$p_\alpha$的函数, 其中$\alpha = 1, 2, \ldots, s$. 在经典力学里两个力学量$f, g$之间满足的Poisson括号2.18被定义为

$$\{f, g\} = \sum_{k=1}^{s} \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{q_k} \right)$$

这样就得到广义动量与广义坐标之间的Poisson括号表示为

$$\{q_i, p_j\} = \delta_{ij}$$

在经典力学里, 任意力学量f都可以写成广义坐标和广义动量的函数形式

$$f = f(q_i, p_j, t)$$

因此, 任意在经典力学里两个力学量的Poisson括号最终都可以化简成广义坐标q_i与广义动量p_j的Poisson括号. 此外Poisson括号满足的代数关系与量子对易括号的代数关系一

样, 实质上是李乘法的一种形式. 这使得我们去猜想如果把经典力学量的Poisson括号改写成量子对易括号, 是否可以从经典力学过度到量子力学呢? 答案是肯定的. 这个过程叫做量子化过程, 但是如何量子化, 以及采取哪种量子化方案则是复杂的.

我们现在尝试引入量子Poisson括号, 从经典Poisson括号可以类比, 我们假设量子Poisson括号满足经典力学Poisson括号的一系列条件:

性质 4.1.1 *(1) 反对称性*

$$\{u,v\} = -\{v,u\}$$

(2) 与常数的括号为0

$$\{u,c\} = 0, \quad c\text{为常数}$$

(3) 分配律(线性性)

$$\{u_1+u_2,v\} = \{u_1,v\} + \{u_2,v\}$$
$$\{u,v_1+v_2\} = \{u,v_1\} + \{u,v_2\}$$

(4) 乘积法则

$$\{u_1u_2,v\} = u_1\{u_2,v\} + \{u_1,v\}u_2$$
$$\{v,u_1u_2\} = u_1\{v,u_2\} + \{v,u_1\}u_2$$

(5) 微分法则

$$\frac{\partial}{\partial t}\{u,v\} = \{\frac{\partial u}{\partial t},v\} + \{u,\frac{\partial v}{\partial t}\}$$

(6) 雅克比恒等式(Jaccobi identity)

$$\{u,\{v,w\}\} + \{v,\{w,u\}\} + \{w,\{u,v\}\} = 0$$

我们要求方程组4第一式中的因子u_1和u_2的次序在整个方程中保持不变,方程组4第二式中的v_1, v_2也有同样的要求. 这些条件已经足以唯一的确定量子Poisson括号形式. 我们用两种不同的顺序去计算$[u_1u_2, v_1v_2]$

$$\begin{aligned}\{u_1u_2,v_1v_2\} &= \{u_1,v_1v_2\}u_2 + u_1\{u_2,v_1v_2\} \\ &= (\{u_1,v_1\}v_2 + v_1\{u_1,v_2\})\,u_2 + u_1\,(\{u_2,v_1\}v_2 + v_1\{u_2,v_2\}) \\ &= \{u_1,v_1\}v_2u_2 + v_1\{u_1,v_2\}u_2 + u_1\{u_2,v_1\}v_2 + u_1v_1\{u_2,v_2\}\end{aligned}$$

以及

$$\begin{aligned}\{u_1u_2, v_1v_2\} &= \{u_1u_2, v_1\}v_2 + v_1\{u_1u_2, v_2\} \\ &= \{u_1, v_1\}u_2v_2 + u_1\{u_2, v_1\}v_2 + v_1\{u_1, v_2\}u_2 + v_1u_1\{u_2, v_2\}\end{aligned}$$

令上面两式子相等, 得到

$$\{u_1, v_1\}(u_2v_2 - v_2u_2) = (u_1v_1 - v_1u_1)\{u_2, v_2\}$$

由于u_1, v_1独立于u_2, v_2. 所以必然有

$$\begin{aligned}u_1v_1 - v_1u_1 &= i\hbar\{u_1, v_1\} \\ u_2v_2 - v_2u_2 &= i\hbar\{u_2, v_2\}\end{aligned}$$

$\hbar$必然与u_1, v_1, u_2, v_2都无关, 并且与它们都对易, 这就表明了$\hbar$仅仅只是一个常数. 由于我们知道在经典力学中我们知道两个力学量的Poisson括号的运算结果必定是一个实力学量, 所以对上式两边同时取Hermite共轭可知$\hbar$之前添加一个纯虚单位i是物理上所要求的.

这样我们就可以定义量子Poisson括号:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

由上面的推导可以知道, 量子Poisson括号正是我们之前为了简化运算所定义的对易括号. 同时我们得到量子对易括号与Poisson括号的关系:

$$[\hat{A}, \hat{B}] = \imath\hbar\{A, B\} \tag{4.1}$$

在之前的篇幅里我们已经验证了上述对易括号确实与经典Poisson括号满足完全一样的代数关系. 现在寻找量子化条件的问题, 简化成了在量子力学中确定Poisson括号的问题. 在提出一个新理论的时候, 我们往往优先关注的是与经典理论有对应的物理量. 最简单的量子对易括号就是正则坐标与正则动量.

$$[\hat{q}_\alpha, \hat{p}_\beta] = i\hbar\{q_\alpha, p_\beta\} = i\hbar\delta_{\alpha,\beta}$$

依次可以计算出下列关系

$$[\hat{q}_\alpha, \hat{p}_\beta] = i\hbar\delta_{\alpha,\beta} \quad [\hat{q}_\alpha, \hat{q}_\beta] = 0 \quad [\hat{p}_\alpha, \hat{p}_\beta] = 0 \tag{4.2}$$

这个一组关系叫做**基本量子条件**. 它表示了正则坐标与正则动量的非对易性, 由于任何与经典有对应的力学量都可以写成q_α, p_α的函数, 所以只要知道正则坐标与正则动量之间的对易关系即可得到任何有经典对应力学量之间的对易关系. 但有些没有非经典力学对应的量则需要从另外的角度去寻找量子化条件, 例如自旋.

经典力学里利用Poisson括号定义两个正则力学量ξ, η. 与经典力学类似, 量子力学里我们也利用量子Poisson括号来定义正则力学量$\hat{\xi}, \hat{\eta}$.

定义 4.1.1 (正则力学量) *若两个力学量ξ, η满足$[\hat{\xi}, \hat{\eta}] = i\hbar$, 则称$\hat{\xi}, \hat{\eta}$为正则力学量.*

从经典Poisson括号与量子Poisson括号的关系(4.1)和基本量子条件(4.2)的角度来看这个定义是非常自然地. 首先我们知道两个经典正则力学量ξ, η之间满足

$$\xi, \eta = 1$$

而正则量子力学量$\hat{\xi}, \hat{\eta}$与经典力学量ξ, η之间的对应关系

$$[\hat{\xi}, \hat{\eta}] = i\hbar\{\xi, \eta\} = i\hbar$$

这样就很自然的得到上面这个量子正则力学量的定义.

4.2 一次量子化

从上一节的描述中可以看出, 对于有经典对应的系统, 若是想找到量子化表述, 只要写出这个系统的经典表述, 例如哈密顿函数$H(p_\alpha, q_\alpha, t)$, 然后将正则坐标与正则动量写成正则量子力学对易算符$\hat{q}_\alpha \hat{p}_\alpha$:

$$q_\alpha, p_\alpha \longmapsto \hat{q}_\alpha, \hat{p}_\alpha$$

其中经典正则量与量子力学正则量之间满足:

$$\{q_\alpha, p_\beta\} = \delta_{\alpha\beta} \longmapsto [\hat{q}_\alpha, \hat{p}_\beta] = \imath\hbar\delta_{\alpha\beta}$$

我们把这个过程叫做**一次量子化**. 总结一下, 可以列成下表 在经典力学中, 系统所有其他的物理量A都可以表示为正则坐标q_α和正则坐标q_α的函数$A(q_\alpha, p_\alpha)$, 所以很容易写出这些

对象	经典力学	量子力学
状态空间	$\mathbb{R}^3 \times \mathbb{R}^3$, Poisson括号	$\mathcal{H}$, 对易括号
力学量	实变函数	自伴算符
测量	决定性	概率性
特征函数	哈密顿函数$H(q_\alpha, p_\alpha, t)$	哈密顿算符$\hat{H}(\hat{q}_\alpha, \hat{p}_\alpha, t)$
正则坐标	q_α, p_α的函数	$\hat{q}_\alpha, \hat{p}_\alpha$的函数

表 4.1: 经典力学与量子力学

物理量在量子力学中的算符. 只要将q_α, p_α换成$\hat{q}_\alpha, \hat{p}_\alpha$即可得到量子力学算符$\hat{A}(\hat{q}_\alpha, \hat{p}_\alpha)$. 例如角动量算符$\hat{\mathbf{l}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$

例 4.2.1 (角动量算符) 在笛卡尔系[1]中角动量算符定义为$\hat{\mathbf{l}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. 我们知道笛卡尔系中位置矢量[2]可以展开成$\hat{\mathbf{r}} = \hat{x}\mathbf{e}_x + \hat{y}\mathbf{e}_y + \hat{z}\mathbf{e}_z$, 同样的动量可以展开成$\hat{\mathbf{p}} = \hat{p}_x\mathbf{e}_x + \hat{p}_y\mathbf{e}_y + \hat{p}_z\mathbf{e}_z$.

所以角动量算符可以展开[3]成

$$\hat{\mathbf{l}} = \hat{l}_x\mathbf{e}_x + \hat{l}_y\mathbf{e}_y + \hat{l}_z\mathbf{e}_z$$

其中$\hat{l}_x, \hat{l}_y, \hat{l}_z$分别为

$$\hat{l}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \hat{l}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z, \quad \hat{l}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

利用*Levi-Civita*张量可以统一写成:

$$\hat{l}_k = \varepsilon_{ijk}\hat{x}_i\hat{p}_j$$

很自然我们也能计算出角动量之间满足的对易关系:

$$[\hat{l}_i, \hat{l}_j] = i\hbar\varepsilon_{ijk}\hat{l}_k$$

然而这样得出的算符不一定对. 其一, 这样直接代换构造的量子力学量算符不一定是自伴算符, 其二, 量子力学量算符是非对易性的, 不同的次序会造成不同的结果. 为此我们需要采取一种标准化量子化规则. 目前有两种对应的量子化规则, 一种叫做Bohm规则, 一种叫

[1]本书规定, 在笛卡尔系中的坐标我们一律用x, y, z表示, 而普遍的坐标我们规定用$q_i, (i = 1, 2, 3)$表示

[2]我们把形如$\hat{\mathbf{A}} = \hat{A}_x\mathbf{e}_x + \hat{A}_y\mathbf{e}_y + \hat{A}_z\mathbf{e}_z$这样的算符叫做矢量算符. 矢量算符满足线性展开性:$\hat{\mathbf{A}}|\psi\rangle = \hat{A}_x|\psi\rangle\mathbf{e}_x + \hat{A}_y|\psi\rangle\mathbf{e}_y + \hat{A}_z|\psi\rangle\mathbf{e}_z$

[3]展开时注意算符顺序

做Weyl规则. 这两种规则正确性目前尚无定论, 因为判断所得到的力学量是否正确, 需要用实验来判断.

定理 4.2.1 (Bohm规则) 将力学量的经典形式$A(q,p)$写成q和p的多项式, 每一项写成$f(q)g(p)$形式, 则这一项相应算符表示规定为

$$\frac{1}{2}\left(f(\hat{q})g(\hat{p})+g(\hat{p})f(\hat{q})\right)$$

例如经典力学量$\hat{x}\hat{p}$对应的量子算符为$\frac{1}{2}(\hat{x}\hat{p}+\hat{p}\hat{x})$. 这一规则所得结果的唯一性和厄米性是明显的.

定理 4.2.2 (Weyl规则) 将力学量经典表示$A(q,p)$写成以ξ,η为变量的傅里叶积分

$$A(q,p)=\iint_{\mathbb{R}^2} a(\xi,\eta)e^{i\xi q+i\eta p}d\xi d\eta$$

然后将积分中指数上的q,p改写成对应算符$\hat{q},\hat{p}$. 所得结果即为与$A(q,p)$对应的力学量算符$\hat{A}(\hat{q},\hat{p})$.

$$\hat{A}(\hat{q},\hat{p})=\iint_{\mathbb{R}^2} a(\xi,\eta)e^{i\xi\hat{q}+i\eta\hat{p}}d\xi d\eta$$

4.2.1 习题

问题 4.2.1 利用*Levi-Civita*张量计算

1. $\hat{p}\cdot\hat{l}=0 \quad \hat{q}\cdot\hat{l}=0$
2. $[\hat{l},\hat{q}\cdot\hat{p}]=0$
3. $\hat{l}^2=\hat{q}^2\hat{p}^2-(\hat{q}\cdot\hat{p})(\hat{p}\cdot\hat{q})-2i\hbar\hat{q}\cdot\hat{p}$

问题 4.2.2 利用*Bohm*规则和*Weyl*规则计算q^3p的量子算符表示形式.

4.3 表象理论

在之前的理论里, 我们进行的所有计算都在抽象的代数空间里. 但是物理需要的并不是这种抽象代数关系, 而是具体的计算值. 比如计算两个力学量算符$\hat{x}\hat{p}$的乘积, 到目前为止, 我们还无能为力. 因为我们尚不知$\hat{x}$和$\hat{p}$算符具体的形式, 也不知道算符的乘法具体是如何进行的, 我们只知道一个抽象的乘法定义:

$$\hat{A}\hat{B}|\psi\rangle=\hat{A}(\hat{B}|\psi\rangle)$$

对于矩阵乘法我们是非常熟悉的, 而矩阵正好满足量子力学所需要的线性关系. 这样我们企图构建态矢量、力学量与向量、矩阵之间的同构关系. 通过这种同构关系, 将抽象的算符与态矢量转化成我们熟悉的矩阵运算. 本节主要目的就是解决如何找到矩阵表示以及不同矩阵表示之间联系的问题

4.3.1 量子态的表象

我们首先从坐标变换开始介绍起, 如图在二维直角坐标系O里有矢量$\mathbf{A}$, 直角坐标系O的两个基矢量$\mathbf{e}_1, \mathbf{e}_2$相互正交.

$$\langle \mathbf{e}_1 | \mathbf{e}_2 \rangle = \delta_{ij}$$

这组基矢是完备的, 二维平面上任何一组基矢都可以用它展开

$$\mathbf{A} = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2$$

其中

$$A_i = \langle \mathbf{e}_i | \mathbf{A} \rangle$$

分别代表矢量$\mathbf{A}$的两个分量. 当A_1, A_2确定后, 就完全确定了平面上一个矢量$\mathbf{A}$. 因此可以认为(A_1, A_2)就是矢量$\mathbf{A}$在坐标系O中的表示.

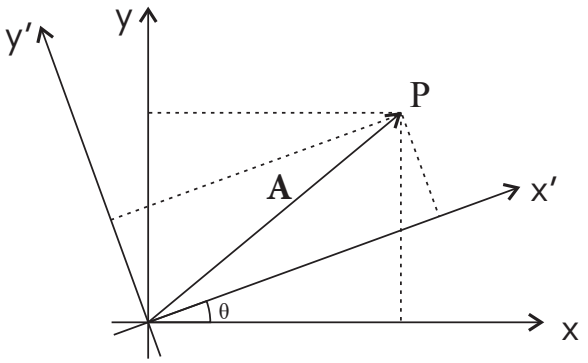


图 4.1: 坐标变换

现在取另一个坐标系O', 新坐标系相当于原坐标系绕O点逆时针转动θ角度, 其基矢分别用$\mathbf{e}_1', \mathbf{e}_2'$表示. 其中

$$\langle \mathbf{e}_i' | \mathbf{e}_j' \rangle = \delta_{ij}, \quad i, j = 1, 2$$

同样的矢量$\mathbf{A}$可以在新坐标系中展开

$$\mathbf{A} = A_1' | \mathbf{e}_1' \rangle + A_2' | \mathbf{e}_2' \rangle + A_3' | \mathbf{e}_3' \rangle$$

其中

$$A_i = \langle \mathbf{e}_i | \mathbf{A} \rangle$$

我们同样叫(A_1', A_2')为$\mathbf{A}$在坐标系O'的表示.

现在很自然会好奇, 同一个矢量在不同坐标系有不同表示, 不同表示之间的联系是什

么. 对于矢量**A**我们有

$$\mathbf{A} = A_1|\mathbf{e}_1\rangle + A_2|\mathbf{e}_2\rangle = A_1'|\mathbf{e}_1'\rangle + A_2'|\mathbf{e}_2'\rangle$$

上面等式同时分别用$\langle\mathbf{e}_1'|, \langle\mathbf{e}_2'|$作用,得到

$$A_1' = A_1\langle\mathbf{e}_1'|\mathbf{e}_1\rangle + A_2\langle\mathbf{e}_1'|\mathbf{e}_2\rangle$$
$$A_2' = A_1\langle\mathbf{e}_2'|\mathbf{e}_1\rangle + A_2\langle\mathbf{e}_2'|\mathbf{e}_2\rangle$$

化简得到

$$\begin{bmatrix} A_1' \\ A_2' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

记$R(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, 上述矩阵方程可以写成

$$\begin{bmatrix} A_1' \\ A_2' \end{bmatrix} = R(\theta) \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

这说明同样的二维矢量在不同二维坐标系的表示可以不一样, 但这些表示之间用矩阵$R(\theta)$联系起来. 其中$R(\theta)$满足关系式

$$\det R = \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = 1$$
$$R^T R = RR^T = 1 \quad (R^T\text{是}R\text{的转置矩阵})$$

满足第一行的正交矩阵叫做正当正交矩阵. 显然考虑R的Hermite共轭形式[4]$R^\dagger$为

$$R(\theta)^\dagger = (R(\theta)^T)* = R(-\theta)$$

显然$R(\theta)R(-\theta) = \mathbf{1}$, $R(-\theta)$是$R(\theta)$的逆矩阵. 而$R^{-1} = R^\dagger$, 所以

$$RR^\dagger = R^\dagger R = \mathbf{1}$$

这样的矩阵叫做Unitary(酉)矩阵. $R(\theta)$矩阵的集合U满足乘法封闭性, 乘法结合律, 存在

[4]在一些数学教材里习惯于把矩阵A的Hermite共轭形式写成A^H, 其中$A^H = (A^T)* = (A^*)^H$

逆元$R(-\theta)$, 存在单位元$\mathbf{1}$. 这个集合构成群, 并且满足酉性$R^\dagger R = RR^\dagger = \mathbf{1}$, 叫做$U(1)$群. 类似地, 在量子力学里, 相容力学量完全集$F = \{\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_s\}$的共同本征态集合$\{||f_k\rangle\}, (k = 1, 2, \cdots, s)$构成$\mathcal{H}$的一组完备正交基矢. 我们为了简单起见, 先考虑离散谱对应的本征基矢. 对于离散谱对应的本征基矢, 很容易可以得到完备正交归一的基矢组$\{|F_k\rangle\}$.

$$\langle F_i|F_j\rangle = \delta_{ij}$$

以$\{|F_k\rangle\}$为基底的表示称为F表象. 量子系统中任何一个量子态都可以用这组表象基底展开

$$|\psi\rangle = \sum_{k=1}^{s} a_k|F_k\rangle$$

其中

$$a_k = \langle F_k|\psi\rangle$$

这一组数$(a_1, a_2, \cdots, a_s)$就构成量子态$|\psi\rangle$在F表象中的表示. 当我们确定这一组数$(a_1, a_2, \cdots, a_s)$后, 量子态就完全确定下来了. 用矩阵去表示量子态为

$$|\psi\rangle \longleftrightarrow \begin{bmatrix} \langle F_1|\psi\rangle \\ \cdots \\ \langle F_n|\psi\rangle \end{bmatrix}$$

对于连续谱的情况, 则需要做替换$\sum \to \int d\lambda$. 若相容力学量$F = \{\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_s\}$的连续本征值为$\lambda$, 对应的归一化本征态集合为$\{|\lambda\rangle\}$, 则任意态矢量可以表示为

$$|\psi\rangle = \int_{\sigma_e} a(\lambda)|\lambda\rangle d\lambda$$

其中

$$a(\lambda) = \langle\lambda|\psi\rangle$$

这一组连续变量$a(\lambda)$构成量子态$|\psi\rangle$在F表象中的表示(此时是连续谱的表示). 同样的, 确定这组变量后, 量子态即可完全确定. 连续谱的表象基矢是不可数的, 这种情况在量子力学中非常普遍.

4.3.2 算符的表象

算符$\hat{A}$把态$|\psi\rangle$映射到了$|\varphi\rangle$.

$$\hat{A} : |\psi\rangle \longmapsto |\varphi\langle\hat{A}|\psi\rangle$$

在上一节里, 我们通过相容力学量完全集本征态集找到了态矢量与向量之间的同构关系, 而向量的线性变换可以用矩阵描述, 此外矩阵的乘法也满足非对易性$AB - BA \neq 0$. 这样促使我们寻找量子力学算符与矩阵之间的对应关系, 我们希望利用我们所熟悉的矩阵乘法去计算量子力学中抽象的算符乘法. 首先为了简单起见, 若相容力学量完全集F的本征值是离散的, 则本征态矢量集$|f_i\rangle, (i = 1, 2, \cdots)$有Parseval等式

$$|f_i\rangle\langle f_i| = \mathbf{1}$$

把Parseval等式代入$\hat{A}|\psi\rangle = |\varphi\rangle$并用$\langle f_i|$作用, 得到

$$\sum_j \langle f_i|\hat{A}|f_j\rangle\langle f_j|\psi\rangle = \langle f_i|\varphi\rangle$$

我们记$A_{ij} = \langle f_i|\hat{A}|f_j\rangle$, 写成矩阵形式

$$\begin{bmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \langle f_1|\psi\rangle \\ \langle f_2|\psi\rangle \\ \cdots \end{bmatrix} = \begin{bmatrix} \langle f_1|\varphi_1\rangle \\ \langle f_2|\varphi_1\rangle \\ \cdots \end{bmatrix}$$

由上节内容可知列向量$\langle f_i|\psi\rangle, (i = 1, 2, \cdots)$正好是态$|\psi\rangle$在对易力学量集合本征态集$F$表象的表示. 此外自然地, 我们可以看到, 矩阵$\langle f_i|\hat{A}|f_j\rangle, (i, j = 1, 2, \cdots)$是算符$\hat{A}$的矩阵表示. 两个力学量算符的乘法对应于矩阵乘法:

$$\sum_k \langle f_i|\hat{A}\hat{B}|f_k\rangle f_k|\psi\rangle\langle f_k|\varphi\rangle$$

$$\sum_{j,k} \langle f_i|\hat{A}|f_j\rangle\langle f_j|\hat{B}|f_k\rangle\langle f_k|\psi\rangle = \langle f_i|\varphi\rangle$$

写成矩阵形式:

$$\begin{bmatrix} \langle f_1|\hat{A}|f_1\rangle & \langle f_1|\hat{A}|f_2\rangle & \cdots \\ \langle f_2|\hat{A}|f_1\rangle & \langle f_2|\hat{A}|f_2\rangle & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \langle f_1|\hat{B}|f_1\rangle & \langle f_1|\hat{B}|f_2\rangle & \cdots \\ \langle f_2|\hat{B}|f_1\rangle & \langle f_2|\hat{B}|f_2\rangle & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \langle f_1|\psi\rangle \\ \langle f_2|\psi\rangle \\ \cdots \end{bmatrix} = \begin{bmatrix} \langle f_1|\varphi\rangle \\ \langle f_2|\varphi\rangle \\ \cdots \end{bmatrix}$$

这样就找到了离散谱力学量的矩阵表示, 叫做F表象的表示. 同样的写成矩阵为

$$\hat{A} \longleftrightarrow \begin{bmatrix} \langle f_1|\hat{A}|f_1\rangle & \cdots & \langle f_1|\hat{A}|f_n\rangle \\ \vdots & \ddots & \vdots \\ \langle f_n|\hat{A}|f_1\rangle & \cdots & \langle f_n|\hat{A}|f_n\rangle \end{bmatrix}$$

对于力学量完全集的连续谱来说, 同样地, 我们只需要利用Parseval等式.

$$\int_{\sigma_e} \langle f|\hat{A}|f'\rangle\langle f'|\psi\rangle df' = \langle f|\psi\rangle$$

同样的, 力学量的乘积可以写成

$$\int_{\sigma_e} \langle f|\hat{A}\hat{B}|f'\rangle\langle f'|\psi\rangle df' = \langle f|\varphi\rangle$$
$$\int_{\sigma_e} \langle f|\hat{A}|f''\rangle\langle f''|\hat{B}|f'\rangle\langle f'|\psi\rangle df'df'' = \langle f|\varphi\rangle$$

这样就得到力学量$\hat{A}$在F表象的表示. 值得注意, 对于连续谱而言, 力学量和算符矩阵表示的矩阵元已经不再可数, 但是习惯上我们依旧称呼$\langle f|\hat{A}|f'\rangle$为力学量$\hat{A}$在完全集$F$表象中的矩阵, 只是此时的矩阵求和$\sum$替换成$\int df$.

4.3.3 表象变换

上一节里定义了态矢量和算符在某个表象下的表示, 算符和态矢量的乘法与矩阵和列向量的乘法一一对应, 算符的乘法与矩阵的乘法一一对应. 若我们用$D(\hat{A}), D(\hat{B})$表示力学量$\hat{A}, \hat{B}$的矩阵表示. 则有下列关系图

$$\begin{array}{ccccc} \hat{A} & \cdot & \hat{B} & = & \hat{A}\hat{B} \\ \downarrow & & \downarrow & & \downarrow \\ D(\hat{A}) & & D(\hat{A}) & = & D(\hat{A})D(\hat{B}) \end{array}$$

正如本节第一部分所说, 取不同完全集F, G的本征矢集作为表象, 得到态矢量和算符的矩阵表示也是不同的. 事实上只要选取一组完备的基矢, 我们就可以建立起一个对应的表象, 选取不同的基矢组, 所得到的表示也是不同的. 很自然的会思考, 两个不同表象之间的联系是什么.

当我们取两组不同的本征基矢$\{|a_i\rangle\}, \{|b_j\rangle\}$作为表象时, 实际上找到的是两组坐标基矢, 量子态和力学量在两组基的表示不一样. 我们感兴趣的是找到这两种表示是如何关联的. 首先我们认为新的坐标基$\{|b_j\rangle\}$是老的坐标基$\{|a_i\rangle\}$变换得到的, 记这个变换为U

$$|b_1\rangle = U|a_1\rangle, |b_2\rangle = U|a_2\rangle, \cdots, |b_i\rangle = U|a_i\rangle, \cdots, |b_n\rangle = U|a_n\rangle$$

当我们对一个物理系统取两个分别不同的坐标系时, 实际对应的物理规律不会发生任何改变. 例如我们在北京海淀取一个坐标系去计算矢量$\vec{a}$的长度, 在浙江取一个坐标系取计算矢量$\vec{a}$的长度, 这两个长度不会随着我们选取的坐标系不同而发生改变. 也就是说, 在量子系统中, 我们要求不同表象中矢量的长度不会发生任何改变.

$$\langle a_i|a_i\rangle = \langle b_i|b_i\rangle = \langle Ua_i|Ua_i\rangle = \langle a_i|U^\dagger U|a_i\rangle$$

也就是说

$$U^\dagger U = UU^\dagger = \mathbf{1}$$

现在我们寻找这样的一个U算符的具体形式. 首先我们有

$$|b_i\rangle = U|a_i\rangle \tag{4.3}$$

利用正交性$\langle a_k|a_i\rangle = \delta_{ki}$

$$\begin{aligned} |b_i\rangle &= U|a_i\rangle \\ &= \sum_k |b_k\rangle\delta_{ki} \\ &= \sum_k |b_k\rangle\langle a_k|a_i\rangle = (\sum_k |b_k\rangle\langle a_k|)|a_i\rangle \end{aligned}$$

这样就得到$U = \sum_k |b_k\rangle\langle a_k|$. 由于$U^\dagger U = \sum_{kk'} |a_k\rangle\langle b_k|b_{k'}\rangle\langle a_{k'}| = \sum_k |a_k\rangle\langle a_k| = \mathbf{1}$, 可知$U$是幺正算符, 满足$U^\dagger U = UU^\dagger = \mathbf{1}$.

在之前的章节里, 我们知道量子态$|\psi\rangle$在F表象里的矩阵形式, 实际上就是$|\psi\rangle$在基底$|f_i\rangle$的

投影$\langle f_i|\psi\rangle$. 给定任意一个右矢量$|\psi\rangle$, 它在旧基$\{|f_k\rangle\}$的展开形式为

$$|\psi\rangle = \sum_k |f_k\rangle\langle f_k|\psi\rangle \tag{4.4}$$

展开系数$\langle f_k|\psi\rangle$是已知的. 我们希望能找到右矢量$|\psi\rangle$在新基$\{|g_k\rangle\}$中的展开系数$\langle g_k|\psi\rangle$. 做法很简单, 只需要利用Parseval等式即可.

首先在式子((4.4))中两边左乘$\langle g_k|$

$$\langle g_k|\psi\rangle = \sum_{k'}\langle g_k|f_{k'}\rangle\langle f_{k'}|\psi\rangle = \sum_{k'}\langle Uf_k|f_{k'}\rangle\langle f_{k'}|\psi\rangle = \sum_{k'}\langle f_k|U^\dagger|f_{k'}\rangle\langle f_{k'}|\psi\rangle$$

而$\langle f_k|U^\dagger|f_{k'}\rangle$是表象变换伴算符$U^\dagger$的矩阵表示形式. 我们知道

$$\langle f_k|U^\dagger|f_{k'}\rangle = \sum_{k''}\langle f_k|g_{k''}\rangle\langle f_{k''}|f_{k'}\rangle = \langle f_k|g_{k'}\rangle$$

这样就得到了$U^\dagger$在旧基$\{|f_k\rangle\}$中的矩阵表示

$$D(U^\dagger) = \begin{bmatrix} \langle f_1|g_{2'}\rangle & \cdots & \langle f_1|g_{k'}\rangle & \cdots \\ \vdots & \ddots & \vdots & \vdots \\ \langle f_k|g_1\rangle & \cdots & \langle f_k|g_{k'}\rangle & \cdots \\ \vdots & \cdots & \cdots & \ddots \end{bmatrix}$$

这样我们就得到了量子态在新表象下和旧表象下的关系

$$\langle g_k|\psi\rangle = \sum_{k'} D_{kk'}(U^\dagger)\langle f'_k|\psi\rangle$$

其中$D_{kk'}(U^\dagger)$是矩阵$D(U^\dagger)$的第k,k'个矩阵元. 总结一下,只要矩阵$D(U^\dagger)$作用在旧表象下量子态的列向量上, 即可得到新表象下的列矩阵:

(新列向量)=$D(U^\dagger)$(旧列向量)

由于算符与算符对应的矩阵表示同态, 我们知道$D(AB) = D(A)D(B)$. 由于U的幺正性$U^\dagger U = \mathbf{1}$, 所以有

$$D(U^\dagger)D(U) = \mathbf{1}_{n\times n}, \quad D(U^\dagger) = D(U)^\dagger$$

在这里为了避免混淆, 我们统一规定一下. 任意力学量$\hat{A}$在旧基$\{|f_i\rangle\}$的矩阵表示记作$D(\hat{A})$,

而在新基$|g_j\rangle$构成的表象下的矩阵表示记作$\bar{D}(\hat{A})$. 这样对于任意算符$\hat{A}$, 取F表象可以得到线性变换$A|\psi\rangle = |\varphi\rangle$的矩阵表示:

$$\sum_k \langle f_k|\hat{A}|\psi\rangle = \langle f_k|\varphi\rangle$$

利用Parseval等式$\sum_i |g_i\rangle\langle g_i| = \mathbf{1}$可以得到

$$\sum_k \sum_{i,j} \langle f_k|g_i\rangle\langle g_i|\hat{A}|g_j\rangle\langle g_j|f_k\rangle\langle f_k|\psi\rangle = \langle f_k|\varphi\rangle$$

写成矩阵形式为

$$D_{ki}(U)\bar{D}_{ij}(\hat{A})D^{\dagger}_{jk}(U)\langle f_k|\psi\rangle = \langle f_k|\varphi\rangle$$

比较上面两式得到

$$D(U)\bar{D}(\hat{A})D^{\dagger}(U) = D(A), \quad \bar{D}(U) = D^{\dagger}(U)D(\hat{A})D(U)$$

算符在新旧表象基底下的矩阵变换正是我们以前所熟悉的相似变换!

定义算符A的迹为

定义 4.3.1 一个算符$\hat{A}$的迹定义为对角矩阵元之和:

$$tr(\hat{A}) = \sum_k \langle f_k|\hat{A}|f_k\rangle$$

尽管在迹的定义中使用了一组表象, 但是$tr(\hat{A})$是不依赖于表象的选择:

$$\begin{aligned} tr(\hat{A}) &= \sum_i \langle f_i|\hat{A}|f_i\rangle \\ &= \sum_{i,j,k} \langle f_i|g_j\rangle\langle g_j|\hat{A}|g_k\rangle\langle g_k|f_i\rangle \\ &= \sum_{i,j,k} \langle g_k|f_i\rangle\langle f_i|g_j\rangle\langle g_j|\hat{A}|g_k\rangle = \sum_{j,k} \langle g_k|g_j\rangle\langle g_j|\hat{A}|g_k\rangle \\ &= \sum_j \langle g_j|\hat{A}|g_j\rangle \end{aligned}$$

同样的技术, 可以容易证明

(1) $tr(\hat{A}\hat{B}\hat{C}) = tr(\hat{B}\hat{C}\hat{A}) = tr(\hat{C}\hat{A}\hat{B})$

(2) $tr(U^{\dagger}\hat{A}U) = tr(\hat{A})$

(3) $tr(|f_i\rangle\langle f_j|) = \delta_{ij}$

(4) $tr(|g_i\rangle\langle f_i) = \langle f_i|g_i\rangle$

至此我们已经介绍了一般性的力学量和量子态右矢量如何在给定表象中得到矩阵表示, 以及在不同表象中, 两种不同的矩阵表示的联系是如何联系起来的. 我们发现力学量不同表示之间的联系实质上就是相似变换, 这样我们就可以把我们之前求解力学量本征值的问题化归为矩阵对角化的问题, 而这个问题在线性代数中我们早已熟悉.

在给定表象中我们得到力学量$\hat{A}$的矩阵表示$D(A)$, 求解$\hat{A}|\psi\rangle = |\varphi\rangle$的本征值问题现在变成求解关于$\lambda$的$N$次代数方程:

$$\det(\hat{A} - \lambda\mathbf{1}) = 0$$

正如我们从线性代数中所知道的, 解得的N个根就等同于我们试图找到的各个本征值a_i, 知道了a_i我们就能求出a_i对应的特征向量解.

对于这种矩阵做法, 算符A是否可**幺正**对角化问题变得非常重要. 在群论里我们会知道, $\hat{A}$是自伴算符的时候一定可以幺正对角化. 而对于其它的算符则不能保证. 例如

$$S_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

不可能用任何幺正矩阵对角化. 在后面谐振子相干态里我们会遇到非自伴算符的本征右矢, 但是这样的本征右矢不能构成一个完备的正交归一基. 对于相干态而言, 我们需要引入Cartan外积和外微分来描述, 最终能得到一个超完备的相干态空间.

对于两个相容算符$\hat{A}, \hat{B}$之间满足相容性条件

$$\left[\hat{A}, \hat{B}\right] = 0$$

这意味着, 若我们取$\hat{A}$的本征态作为表象基矢量, 显然$\hat{A}$在$\hat{A}$的本征态表象下的矩阵表示是对角化的.

$$D(\hat{A}) = \begin{bmatrix} \langle a_1|a_1\rangle & 0 & \cdots & \cdots \\ 0 & \langle a_2|a_2\rangle & 0 & \cdots \\ 0 & 0 & \langle a_3|a_3\rangle & 0 \\ \cdots & \cdots & \cdots & \ddots \end{bmatrix}$$

把$\hat{A}$的本征态基矢作用在$[\hat{A},\hat{B}]=0$两端, 得到

$$a_i\langle a_i|\hat{B}|a_j\rangle = a_j\langle a_i|\hat{B}|a_j\rangle$$

$$(a_i-a_j)\langle a_i|\hat{B}|a_j\rangle = 0$$

只有当$i=j$时$\langle a_i|\hat{B}|a_j\rangle \neq 0$, 若$i\neq j$, 则$\langle a_i|\hat{B}|a_j\rangle = 0$. 意味着, $\hat{B}$在A表象下也是对角化的.

$$D(\hat{B}) = \mathrm{diag}(\langle a_1|\hat{B}|a_1\rangle, \cdots, \langle a_i|\hat{B}|a_j\rangle, \cdots)$$

这与我们所学过的矩阵理论一样, 一般来说两个矩阵M_1和M_2是不能对易的, 但若M_1, M_2是两个对角矩阵, 则是可以对易的.

4.3.4 坐标表象

对于一般笛卡尔系, 正则坐标一般是直角坐标$\mathbf{x}=(x_1,x_2,x_3)$. 这里为了简化问题, 我们先考虑一维坐标. 坐标算符满足本征方程

$$\hat{x}|x'\rangle = x'|x'\rangle$$

坐标算符$\hat{x}$是自伴算符, $\hat{x}$所有的本征态对应的本征矢量构成一个完备的基底, 满足Parseval恒等式.

$$\int_{\mathbb{R}^3}|x'\rangle\langle x'|\mathrm{d}x' = \mathbf{1}$$

利用Parseval等式可以得到坐标表象下的量子态和力学量矩阵表示. 在坐标表象下

$$|\psi\rangle \longmapsto \langle x|\psi\rangle = \psi(x)$$

我们定义的$\psi(x)=\langle x|\psi\rangle$是坐标算符本征值$x$的函数, 叫做量子态$|\psi\rangle$的**坐标波函数**. 这样量子态归一化条件为

$$\langle\psi|\psi\rangle = \int_{\mathbb{R}^3}\langle\psi|x\rangle\langle x|\psi\rangle\mathrm{d}x = \int_{\mathbb{R}^3}\psi^*(x)\psi(x)\mathrm{d}x = \int_{\mathbb{R}^3}|\psi(x)|^2\,\mathrm{d}x = 1$$

在量子力学里任意一个量子态可以用一个自伴算符的本征态来展开

$$|\psi\rangle = \int_{R^3}|x\rangle\langle x|\psi\rangle = \int_{\mathbb{R}^3}\psi(x)|x\rangle$$

根据量子态叠加原理和测量原理, $\psi(x)$是测量量子态$|\psi\rangle$得到坐标算符本征态$|x\rangle$的概率. 而$|x\rangle$是坐标算符$\hat{x}$在空间x点处的本征态, 满足本征方程

$$\hat{x}|x\rangle = x|x\rangle$$

$\psi(x)$就被解释成在量子态的位置概率分布. 现在我们暂时不考虑含时演化的问题, 本章我们所有的讨论都认为时间t没有任何变化, 含时问题我们下一章再讨论.

在量子力学里, 我们描述一个粒子是用量子态去描述的, 描述粒子的量子态里包含了有关这个粒子所有的必要信息. 所以对于波函数$\psi(x)$, 必须要有三个性质

1. $|\psi(x)|^2$是粒子位置的概率分布. 也就是说, 粒子在空间$\Omega \subset \mathbb{R}^3$的概率是$\int_\Omega |\psi(x)|^2 \mathrm{d}x$.
2. 对于束缚态[5], $\psi(x)$必须是平方可积的, 也就是说
$$L^2(\mathbb{R}^3) := \{\psi : \mathbb{R}^3 \to \mathbb{C} \int_{\mathbb{R}^3} |\psi(x)|^2 \mathrm{d}x < \infty\}$$
3. 根据我们之前双缝干涉选择测量实验, 在物理上$\psi(x)$必须对应双缝干涉实验中的相干波函数. 也就是说$\psi(x)$必须满足某一类波动方程, 只是到本节为止, 尚未介绍.

为了进一步解释算符的坐标表象形式, 我们先引入积分算符的定义

定义 4.3.2 (积分算符) *定义积分算符:*

$$\mathcal{K} : \psi \mapsto \int K(\cdot, y)\psi(y)\mathrm{d}y$$

例如$(\mathcal{K}\psi)(x) = \int K(x,y)\psi(y)\mathrm{d}y$. *函数*$K : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{C}$*叫做积分算符*$\mathcal{K}$*的积分核.*

有了积分算符的定义, 我们很容易描述算符在坐标表象的形式. 首先对于任意算符$\hat{A}$作用在量子态右矢量$|\psi\rangle$有

$$\hat{A}|\psi\rangle = |\varphi\rangle \xrightarrow{\text{两边作用}\langle x'|} \langle x'|\hat{A}|\psi\rangle = \langle x'|\varphi\rangle$$

利用坐标算符$\hat{x}$的Parseval等式, 可以得到

$$\int \mathrm{d}x''\langle x'|\hat{A}|x''\rangle\langle x''|\psi\rangle = \langle x'|\varphi\rangle$$

[5]满足边条件$x \to \infty, \psi(x) \to 0$条件叫束缚态波函数$\psi(x)$

显然, 算符$\hat{A}$在坐标表象下是积分算符$\mathcal{A}:\psi(x')\mapsto\varphi(x')=\int A(x',x'')\psi(x'')\mathrm{d}x''$的核. 而原来的表达式在坐标表象下写成

$$\hat{A}|\psi\rangle=|\varphi\rangle\xrightarrow{\text{坐标表象}}\int A(x',x'')\psi(x'')\mathrm{d}x''=\varphi(x')$$

其中$A(x',x'')=\langle x'|\hat{A}|x''\rangle$. 对于坐标算符$\hat{x}$的函数$A(\hat{x})$, 算符在坐标表象中的矩阵形式为:

$$\langle x'|A(\hat{x})|x''\rangle=\langle x'|A(x'')|x''\rangle=A(x'')\delta(x'-x'')$$

力学量作用在量子态$A(\hat{x})|\psi\rangle=|\varphi\rangle$的表示可以写成

$$\int_{\mathbb{R}^3}dx''\langle A(x'')\delta(x'-x'')\psi(x'')=\varphi(x')$$

化简得到

$$A(x')\psi(x')=\varphi(x')$$

在坐标表象中, 原来抽象的线性映射现在变成具体的函数方程. 在上一章里我们介绍了基本量子化条件

$$[\hat{q},\hat{p}]=i\hbar$$

这意味着正则坐标$\hat{q}$与正则动量$\hat{p}$之间没有共同的本征态, 也就是说, 我们找不到一个表象, 使得$\hat{q},\hat{p}$同时对角化. 我们试着先确定坐标表象, 然后在坐标表象里计算出$\hat{p}$的表示.

首先我们先计算矩阵元$\langle x|p\rangle$

公式 4.3.1 ($\langle x|p\rangle$) 矩阵元$\langle x|p\rangle=\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}px}$

证明 4.3.1 由方程((3.22))我们可以得到

$$e^{\frac{i}{\hbar}\hat{p}x'}\hat{x}e^{-\frac{i}{\hbar}\hat{p}x'}=\hat{x}+x'$$

所以有

$$\begin{aligned}\hat{x}e^{-\frac{i}{\hbar}\hat{p}x'}|x\rangle&=e^{-\frac{i}{\hbar}\hat{p}x'}e^{\frac{i}{\hbar}\hat{p}x'}\hat{x}e^{-\frac{i}{\hbar}\hat{p}x'}|x\rangle\\&=e^{-\frac{i}{\hbar}\hat{p}x'}(\hat{x}+x')|x\rangle\\&=(x+x')e^{-\frac{i}{\hbar}\hat{p}x'}|x\rangle\end{aligned}$$

这意味着:$e^{-\frac{i}{\hbar}\hat{p}x'}|x\rangle=|x+x'\rangle$

计算如下矩阵元:

$$\langle x|e^{\frac{i}{\hbar}\hat{p}x'}|p\rangle = \langle x+x'|p\rangle = e^{\frac{i}{\hbar}\hat{p}x'}\langle x|p\rangle$$

令$x=0$得到

$$\langle x'|p\rangle = e^{\frac{i}{\hbar}px'}\langle 0|p\rangle$$

由于$\langle p'|p\rangle = \delta(p-p')$

$$\begin{aligned}\langle p'|p\rangle &= \int_{\mathbb{R}^3}\langle p'|x'\rangle\langle x'|p\rangle dx' \\ &= \int_{\mathbb{R}^3} dx'\langle 0|p\rangle\langle p'|0\rangle e^{-\frac{i}{\hbar}(p'-p)x'} \\ &= |\langle 0|p\rangle|^2 2\pi\hbar\delta(p-p') = \delta(p-p')\end{aligned}$$

所以$\langle 0|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}$

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}px}$$

得证

4.3.5 动量表象

Chapter 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 单参数幺模群——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

Chapter 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

Chapter 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

Chapter 8

三维散射态

几种典型的散射包括:

1. 弹性散射$a + b \to a + b$散射粒子的内部状态没有改变

2. 非弹性散射$a + b \to a' + b'$

3. 重排散射$a + b \to c + d + e$

4. 共振散射

我们主要关注弹性散射. 散射问题是一个微妙的问题. 一般来说, 散射问题被看成量子力学的一种重要应用. 用简明自洽的方式去解释这类问题是比较困难的. 散射态是一个对应于连续谱的非束缚态, 与束缚态的离散分离能级相反, 散射态能级是连续的. 一般来说, 我们关心散射过程中入射粒子的各种属性和散射角分布, 而不是如我们之前一般求解薛定谔方程所关心的能级分布.

8.1 介绍

8.1.1 经典散射与微分截面

我们首先简单的讨论经典散射过程

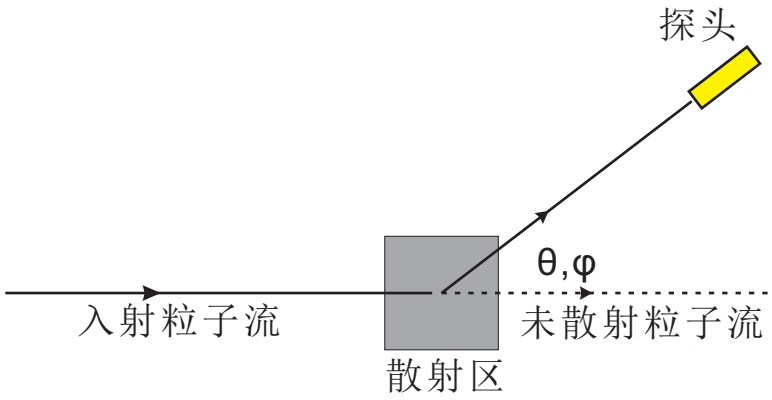


图 8.1: 散射示意图

假设入射粒子靶距参数为b, 射靶方向角为φ, 散射粒子散射方向角为(θ, φ). 经典散射理论的关键问题是对于一个给定的打靶参数b计算散射角θ. 一般来说, 我们感兴趣的是研究入射粒子流的分布. 我们首先给出流密度**J**的定义.

定义 8.1.1 (流密度) 单位时间单位面积通过的粒子数定义为粒子流密度.

$$J = \frac{dN}{d\sigma dt} \tag{8.1}$$

所以单位时间入射的粒子数为$n_{in} = J_{in} d\sigma$, 这里J_{in}表示入射粒子流密度. 我们考虑经历散射$n_{in} \to n_{out}$后单位时间出射粒子数为$D(d\sigma \to d\Omega) J_{in} d\Omega$. 散射前后粒子数不变, $n_{in} = n_{out}$. 所以我们有

$$n_{out} = \frac{d\sigma}{d\Omega} J_{in} \tag{8.2}$$

这样我们就找到了$d\sigma$和方向角$d\Omega$之间的对应关系. $d\sigma$正是被靶散射向$d\Omega$那部分粒子对应的入射面积$d\sigma$. 我们称其为**截面**.

利用$d\sigma$与$d\Omega$的对应关系, 当$\Omega \to 0$时, $d\sigma = \frac{d\sigma}{d\Omega} d\Omega$. 我们得到微分截面的定义

定义 8.1.2 (微分截面) 被散射到单位立体角的粒子所对应的入射截面

$$D(\theta, \varphi) = \frac{d\sigma}{d\Omega} \tag{8.3}$$

叫做微分截面, 这里θ, φ分别是球坐标的经度和纬度.

根据打靶参数b与方向角φ, 我们有

$$\begin{aligned} \mathrm{d}\sigma &= \mathrm{d}b \times b \sin d\varphi \approx b\mathrm{d}b\mathrm{d}\varphi \\ \mathrm{d}\Omega &= \sin\theta \mathrm{d}\theta \mathrm{d}\varphi \end{aligned} \tag{8.4}$$

上下两式相除得到

$$D(\theta) = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin\theta}\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right| \tag{8.5}$$

根据微分截面可以定义总截面

定义 8.1.3 (总截面) 全体立体角对应的入射截面

$$\sigma_{tot} = \int \mathrm{d}\Omega D(\theta,\varphi) = \int_0^{2\pi} d\varphi \int_0^{\pi} \mathrm{d}\theta \sin\theta D(\theta,\varphi) \tag{8.6}$$

叫做总截面

例 8.1.1 (硬球散射) 考虑一个经典散射过程, 假设靶是一个半径为R的刚性球, 入射粒子是无质量的刚性球, 在表面法神弹性散射. 如果打靶距离$b < R$, 我们有

$$b = R\sin\alpha = R\sin\frac{\pi}{2} - \frac{\theta}{2} = R\cos\frac{\theta}{2} \tag{8.7}$$

所以我们有$\theta = 2\arccos\frac{b}{R}, \quad b < R$

若$b > R, \theta = 0$

这样就有微分截面

$$D(\theta) = \frac{b}{\sin\theta}\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right| = \frac{R^2}{4} \tag{8.8}$$

总截面$\sigma_{tot} = \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\theta D(\theta) = \pi R^2$

8.1.2 量子散射波

在量子散射问题中, 我们需要研究量子散射波函数. 这里为了方便起见, 我们讨论坐标表象下的粒子波函数$\psi(\mathbf{r},t) = \langle \mathbf{r}|\psi\rangle$. 坐标表象下的波函数满足坐标表象的薛定谔方程

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r},t) \tag{8.9}$$

定理 8.1.1 (定域概率流密度守恒) 定义概率密度$\rho(\mathbf{r},t)=|\psi(\mathbf{r},t)|^2$与流密度

$$J=-\frac{i\hbar}{2m}\left[\psi^*(\mathbf{r},t)\nabla\psi(\mathbf{r},t)-\psi(\mathbf{r},t)\nabla\psi^*(\mathbf{r},t)\right] \tag{8.10}$$

则有

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t}+\nabla\cdot J=0 \tag{8.11}$$

我们假设入射波函数为$\psi(z)=Ae^{ikz}$, 沿着z方向入射, 在散射区与球对称的势发生散射. 由于球对称的势, 出射波函数必须是球对称的. 所以距离散射中心比较远距离$(r\to\infty)$的波函数遵循一般形式

$$\psi(r,\theta)=A\left[e^{ikz}+f(\theta)\frac{e^{ikr}}{r}\right] \tag{8.12}$$

这里$f(\theta)$是散射振幅. 由于入射粒子是渐进自由的粒子, 无穷远处的能量为$E=\frac{k^2}{2m}$. 所以这里

$$k=\frac{\sqrt{2mE}}{\hbar} \tag{8.13}$$

值得注意, 散射波函数第一项e^{ikz}表示向前散射组分, 这部分组分实质上并未发生散射. 第二项$\frac{e^{ikr}}{r}$表示球对称部分, 这部分是真正发生散射的组分.

利用流密度的定义, 可以得到入射粒子流密度

$$J_i=-\frac{i\hbar}{2m}(e^{-ikz}\nabla e^{ikz}-e^{ikz}\nabla e^{-ikz})=\frac{\hbar k}{m} \tag{8.14}$$

另一方面, 类似的计算得到散射部分的粒子流密度

$$J_s=-\frac{i\hbar}{2m}\left\{f^*(\theta)\frac{e^{-ikr}}{r}\frac{\partial}{\partial r}[f(\theta)\frac{e^{ikr}}{r}]-f(\theta)\frac{e^{ikr}}{r}\frac{\partial}{\partial r}[f^*(\theta)\frac{e^{-ikr}}{r}]\right\}=\frac{\hbar k}{m}\frac{|f(\theta)|^2}{r^2} \tag{8.15}$$

所以沿着散射角θ, 我们有

$$\mathrm{d}n=J_s r^2\mathrm{d}\Omega=\frac{\hbar k}{m}|f(\theta)|^2\mathrm{d}\Omega \tag{8.16}$$

根据微分截面的定义, 得到

$$D(\theta)=\frac{1}{J_i}\frac{\mathrm{d}n}{\mathrm{d}\Omega}=|f(\theta)|^2 \tag{8.17}$$

总截面为

$$\sigma_{tot}=\int\mathrm{d}\Omega D(\theta,\varphi)=\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\mathrm{d}\theta\sin\theta|f(\theta)|^2=2\pi\int_0^{\pi}\sin\theta|f(\theta)|^2 \tag{8.18}$$

原则上, 我们可以通过求解薛定谔方程得到散射振幅$f(\theta)$.

$$[\hat{H}_0 + \hat{V}(r)]\psi(r) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right]\psi(r) = E\psi(r) \tag{8.19}$$

其中所要求的边界条件为, 当$r \to \infty$时

$$\psi(r,\theta) \to e^{ikz} + f(\theta)\frac{e^{ikr}}{r} \tag{8.20}$$

现在看来, 量子散射问题可以简化为求解静态特征值问题. 但是在实际散射历程中, 散射问题时一个依赖于时间的过程. 但在实际散射实验中, 散射问题是随时间变化的. 为了理解为什么我们可以把含时散射问题当作定态散射问题来研究, 我们接下来研究含时散射理论.

8.2 含时散射

8.2.1 入射出射态的近似

散射过程由于哈密顿$H = H_0 + V(r)$可以看成两个部分, 一部分是自由项$H_0 = \frac{p^2}{2m}$, 对应于自由粒子哈密顿, 另一部分$V(r)$对应于散射势. 一般而言, 我们考虑短程势. 所以我们可以将其看做

$$V(r > r_0) = 0 \tag{8.21}$$

进一步说, 为了将相互作用势看成局域势, 势函数$V(r)$必须满足下列条件

1. $V(r)$收敛的比r^{-3}要快
2. $V(r)$在原点处的奇性比$r^{\frac{3}{2}}$要小
3. 除了有限个有限大的不连续点外, $V(r)$在$0 < r < \infty$连续

含时量子态满足含时薛定谔方程

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{8.22}$$

由于哈密顿不含时, 可以定义时间演化算符$U(t) = exp(-\frac{i}{\hbar}Ht)$

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \tag{8.23}$$

这里为了方便, 我们记$|\psi\rangle = |\psi(0)\rangle$, 表示$t = 0$时刻的态矢量. 下面我们引入散射过程的渐进态. 在相互作用有效程范围之外, 我们有$H \to H_0$. 因此在相互作用范围外的时间演化算符可以替换成自由的时间演化算符

$$U_0(t) = e^{-\frac{i}{\hbar}H_0 t} \tag{8.24}$$

所以我们有理由说, 在远离相互作用区时, 散射态$|\psi\rangle$的演化行为$U(t)|\psi\rangle$类似于自由波函数. 这意味着存在态矢量$|\psi_{in}\rangle$和$|\psi_{out}$满足下列关系

$$\begin{aligned} \lim_{t\to-\infty}[U(t)|\psi\rangle - U_0(t)|\psi_{in}]\rangle = 0 \\ \lim_{t\to+\infty}[U(t)|\psi\rangle - U_0(t)|\psi_{out}]\rangle = 0 \end{aligned} \tag{8.25}$$

这里$|\psi_{in}\rangle$和$|\psi_{out}\rangle$分别是散射态的渐进入射态和出射态

8.2.2 Moller算子

我们首先不加证明地陈述一个渐近条件定理:

定理 8.2.1 (渐进条件) 如果相互作用势满足上节的三个条件, 则对于*Hilbert*空间$\mathcal{H}$中每一个渐进态$|\psi_{in}\rangle$和$|\psi_{out}\rangle$都有一个$|\psi\rangle$使得

$$\begin{aligned} \lim_{t\to-\infty}[U(t)|\psi\rangle - U_0(t)|\psi_{in}\rangle] = 0 \\ \lim_{t\to+\infty}[U(t)|\psi\rangle - U_0(t)|\psi_{out}\rangle] = 0 \end{aligned} \tag{8.26}$$

定义Moller算子

$$\begin{aligned} \Omega_+ = \lim_{t\to-\infty} U^\dagger(t)U_0(t) = \lim_{t\to-\infty} e^{iHt}e^{-iH_0 t} \\ \Omega_- = \lim_{t\to+\infty} U^\dagger(t)U_0(t) = \lim_{t\to+\infty} e^{iHt}e^{-iH_0 t} \end{aligned} \tag{8.27}$$

为了方便, 本章我们取自然单位$\hbar = 1$. 由Moller算子定义可以得到散射态和渐进态之间的关系

$$\begin{aligned} |\psi\rangle = \lim_{t\to-\infty} U^\dagger(t)U_0(t)|\psi_{in}\rangle = \Omega_+|\psi_{in}\rangle \\ |\psi\rangle = \lim_{t\to+\infty} U^\dagger(t)U_0(t)|\psi_{ont}\rangle = \Omega_-|\psi_{out}\rangle \end{aligned} \tag{8.28}$$

上述关系意味着Hilbert空间$\mathcal{H}$中的$|\psi_{in,out}\rangle$实际上都是某个真实演化的渐进.

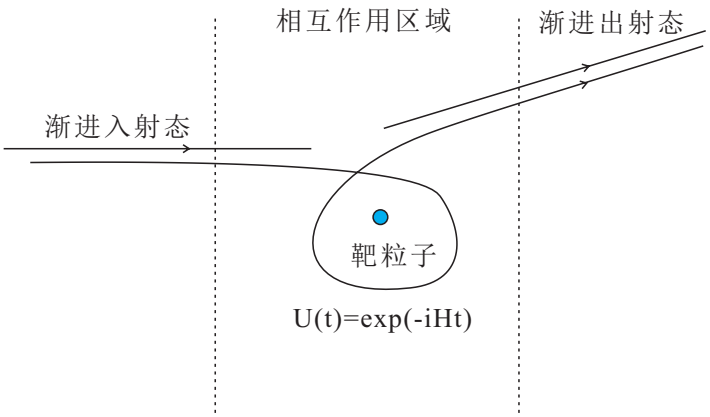


图 8.2: 散射过程图解

另一方面我们注意到, 并不是所有的$|\psi\rangle \in \mathcal{H}$都能定义一个有入射或者出射渐进的演化轨道$U(t)|\psi\rangle$. 如果$\mathcal{H}$有一部分束缚态$|n\rangle$, $U(t)|n\rangle$的行为表现为定态行为而不是自由行为. 因此Hilbert空间可以分成两个子空间$\mathcal{H} = \mathcal{R} \oplus \mathcal{B}$, 这里$\mathcal{R}$和$\mathcal{B}$分别对应于散射态和束缚态空间. 对于每一个$|\psi\rangle \in \mathcal{R}$我们有

$$\begin{aligned} |\psi_{in}\rangle &= \Omega_+^\dagger |\psi\rangle = \lim_{t\to-\infty} U_0^\dagger(t)U(t)|\psi\rangle \\ |\psi_{out}\rangle &= \Omega_-^\dagger |\psi\rangle = \lim_{t\to+\infty} U_0^\dagger(t)U(t)|\psi\rangle \end{aligned} \tag{8.29}$$

入射渐进态	$\xrightarrow{\Omega_+}$	$t=0$时刻散射态	$\xleftarrow{\Omega_-}$	出射渐进态
$\|\psi_{in}\rangle$	$\to$	$\|\psi\rangle$	$\leftarrow$	$\|\psi_{out}\rangle$
$\|\phi\rangle$	$\to$	$\|\phi+\rangle, \|\chi-\rangle$	$\leftarrow$	$\|\chi\rangle$

我们可以看到, 只要把Moller算子作用在入射态或者出射态上就能产生一个散射态. 那么问题来了: 对于每一个$|\psi\rangle \in \mathcal{H}$都能定义一个有入射渐进或者出射渐进对应的$U(t)|\psi\rangle$演化吗? 答案当然是否定的, 我们注意到$\mathcal{H}$中存在束缚态空间$\mathcal{B}$, 也就是说, 仅仅只有散射态空间里的$|\psi\rangle \in \mathcal{R}$才有渐进态对应.

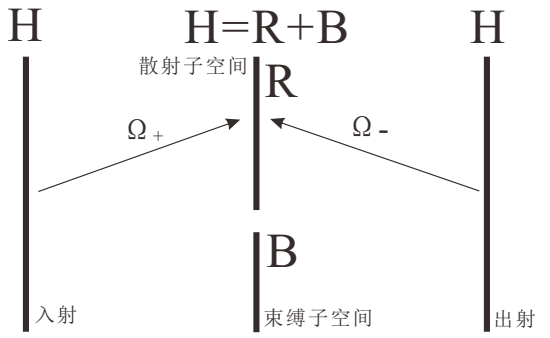


图 8.3: 渐进入射态和渐进出射态以及真实散射态之间的映射图

8.2.3 Moller算子的性质

公式 8.2.1 (纠缠关系)

$$H\Omega_{\pm} = \Omega_{\pm} H_0 \tag{8.30}$$

证明 8.2.1 对于给定的有限时间τ, 我们有

$$\begin{aligned} e^{iH\tau}\Omega_{\pm} &= \lim_{t\to\mp} e^{iH\tau} e^{-iH_0 t} = \lim_{t\to\mp} e^{iH(t+\tau)} e^{-iH_0 t} \\ &= \lim_{t\to\mp} e^{iH(t+\tau)} e^{-iH_0(t+\tau)} e^{iH_0\tau} = \Omega_{\pm} e^{iH_0\tau} \end{aligned} \tag{8.31}$$

上面两边同时对τ求导, 得到

$$iHe^{iH\tau}\Omega_{\pm} = \Omega_{\pm} iH_0 e^{iH_0\tau} \tag{8.32}$$

令$\tau \to 0$得到:$H\Omega_{\pm} = \Omega_{\pm} H_0$

公式 8.2.2 (正交关系)

$$\begin{aligned} \Omega_{\pm}^{\dagger}\Omega_{\pm} &= 1 \\ \Omega_{\pm}\Omega_{\pm}^{\dagger} &\neq 1 \end{aligned} \tag{8.33}$$

注意第二个式子, 这是由于$\mathcal{H}$中存在着束缚态空间, 并不是所有$\mathcal{H}$中的态都存在相应的渐进态.

为了推导出正交关系, 我们计算$\langle\psi_{in,out}|\Omega_{\pm}^{\dagger}\Omega_{\pm}|\psi_{in,out}\rangle = \langle\psi|\psi\rangle = 1$. 这里利用了归一化条件$\langle\psi|\psi\rangle = 1$. $|\psi_{in,out}\rangle$可以取到$\mathcal{H}$全空间, 因此$\Omega_{\pm}^{\dagger}\Omega_{\pm} = 1$. 类似的, 我们可以得到$\langle\psi|\Omega_{\pm}\Omega_{\pm}^{\dagger}|\psi\rangle = 1$, 这里$|\psi\rangle \in \mathcal{R}$. 由于$\mathcal{R}$不是完备的, 所以$\Omega_{\pm}\Omega_{\pm}^{\dagger} \neq 1$

利用正交关系, 纠缠关系可以写成

$$\Omega_{\pm}^{\dagger} H \Omega_{\pm} = H_0 \tag{8.34}$$

8.2.4 S矩阵

我们首先总结一下上节Moller算子的定义以及Moller算子在渐进态之间的基本映射关系

$$\Omega_{\pm} = \lim_{t \to \mp} U^{\dagger}(t) U_0(t) = \lim_{t \to \mp} e^{iHt} e^{-iH_0 t} \tag{8.35}$$

$$\begin{aligned} |\psi\rangle &= \Omega_{+} |\psi_{in}\rangle \\ |\psi\rangle &= \Omega_{-} |\psi_{out}\rangle \end{aligned} \tag{8.36}$$

这里$|\psi_{in,out}\rangle \in \mathcal{H}$

$$\begin{aligned} |\psi_{in}\rangle &= \Omega_{+}^{\dagger} |\psi\rangle \\ |\psi_{out}\rangle &= \Omega_{-}^{\dagger} |\psi\rangle \end{aligned} \tag{8.37}$$

这里$|\psi\rangle \in \mathcal{R}$

为了获得出射渐进态和入射渐进态之间的联系, 我们引入一个散射算子S. 定义如下

定义 8.2.1 (散射算子) 记入射渐进态$|\psi_{in}\rangle$和出射渐进态$|\psi_{out}\rangle$. 我们定义$|\psi_{in}\rangle$和$|\psi_{out}\rangle$的映射

$$S : |\psi_{in}\rangle \mapsto |\psi_{out}\rangle = S|\psi_{in}\rangle \tag{8.38}$$

叫做散射过程的散射算子

利用Moller算子在散射态和渐进态之间的映射关系(8.36)

$$\begin{aligned} |\psi_{out}\rangle &= \Omega_{-}^{\dagger} |\psi\rangle = \Omega_{-}^{\dagger} \Omega_{+} |\psi_{in}\rangle \\ |\psi_{in}\rangle &= \Omega_{+}^{\dagger} |\psi\rangle = \Omega_{+}^{\dagger} \Omega_{-} |\psi_{out}\rangle \end{aligned} \tag{8.39}$$

与散射算子S的定义相比较, 可以得知

$$S = \Omega_{-}^{\dagger} \Omega_{+}, \quad S^{\dagger} = \Omega_{+}^{\dagger} \Omega_{-} \tag{8.40}$$

$$|\psi_{out}\rangle = S|\psi_{in}\rangle, \quad |\psi_{in}\rangle = S^{\dagger} |\psi_{out}\rangle \tag{8.41}$$

显然散射算子S满足幺正性. 其幺正性可以通过$|\psi_{in,out}\rangle \in \mathcal{H}$是一组完备的态矢量来说明. 由于归一化条件$\langle\psi_{out}|\psi_{out}\rangle = \langle\psi_{in}|S^\dagger S^\dagger|\psi_{in}\rangle = 1$, 而$|\psi_{in}\rangle$是完备的, 所以得到$S^\dagger S = 1$. 同样的可以得到$SS^\dagger = 1$. 值得注意, 我们不能通过直接计算$S^\dagger S = \Omega_+^\dagger \Omega_- \Omega_-^\dagger \Omega_+ = \Omega_+^\dagger \Omega_+ = 1$得到, 因为Moller算子正交性条件$\Omega_- \Omega_-^\dagger \neq 1$.

现在我们进一步证明关系式:

$$[H_0, S] = 0 \tag{8.42}$$

证明 8.2.2 利用*Moller*算子的纠缠关系(8.30)$H\Omega_\pm = \Omega_\pm H_0$. 可以得到

$$SH_0 = \Omega_-^\dagger \Omega_+ H_0 = \Omega_-^\dagger H \Omega_+ = (H\Omega_-)^\dagger \Omega_+ = (\Omega_- H_0)^\dagger \Omega_+ = H_0 \Omega_-^\dagger \Omega_+ = H_0 S \tag{8.43}$$

这样就证明了$[S, H_0] = 0$

这个关系表明了散射前后能量守恒. 入射能量期望值

$$\text{入射能} = \langle\psi_{in}|H_0|\psi_{out}\rangle \tag{8.44}$$

同样的, 出射能量期望值

$$\text{出射能} = \langle\psi_{out}|H_0|\psi_{out}\rangle = \langle\psi_{in}|S^\dagger H_0 S|\psi_{in}\rangle \tag{8.45}$$

根据上式关系, 入射能和出射能相等.

进一步的, 由于H_0是自由哈密顿, 对应的有一簇相互正交的本征态矢. 由于$[H_0, S] = 0$, 我们可以取H_0的本征态矢作为基底来表示散射算子S. 一般而言$[H_0, p] = 0$, 所以可以取动量表象来表示散射算子S, 表示矩阵叫S矩阵.

$$S = \int dp dp' |p'\rangle\langle p'|S|p\rangle\langle p| \tag{8.46}$$

所以在动量表象中, 散射算子表示成$\langle p'|S|p\rangle$. 利用关系式(8.42), 可以得到

$$\langle p'|[H_0, S]|p\rangle = (E_{p'} - E_p)\langle p'|S|p\rangle = 0 \tag{8.47}$$

这表明, 若$E_{p'} \neq E_p$, 则$\langle p'|S|p\rangle = 0$, 若$E_{p'} = E_p$, 则$\langle p'|S|p\rangle$可以存在非0情况. 这个结论

告诉我们, S矩阵可以表示成下列形式

$$\langle p'|S|p\rangle = \delta(E_{p'} - E_p) \times (\text{残余项}) \tag{8.48}$$

定义 8.2.2 粒子从渐进入射态$|\psi_{in}\rangle$经历散射碰撞后跃迁到$|\psi_{out}\rangle$的跃迁概率定义为

$$\omega(\psi_{out} \leftarrow \psi_{in}) = |\langle\psi_{out}|S|\psi_{in}\rangle|^2 \tag{8.49}$$

其中$\langle\psi_{out}|S|\psi_{in}\rangle$是散射过程$(\psi_{out} \leftarrow \psi_{in})$的概率振幅.

8.2.5 散射振幅

接下来的工作, 我们证明微分散射截面可以用S矩阵元来表示. 为了消掉向前散射分量, 我们分解S算子为

$$S = 1 + iT \tag{8.50}$$

用来表示没有发生相互作用$(S = 1)$的部分和发生散射的部分iT. 类似S算子, T算子同样与H_0对易. T在H_0的动量表象叫做T矩阵, 表示成如下形式

$$\langle p'|iT|p\rangle = \delta(E_{p'} - E_p) \times (\text{残余项}) \tag{8.51}$$

为了方便起见, 我们用$\mathcal{M}(p' \leftarrow p)$表示上式残余部分

$$\langle p'|iT|p\rangle = 2\pi\delta(E_{p'} - E_p)i\mathcal{M}(p' \leftarrow p) \tag{8.52}$$

其中因子$2\pi i$是为了以后计算方便引入的. 稍后我们将会解释$\mathcal{M}(p' \leftarrow p)$. 利用$S = 1 + iT$, 我们可以得到如下关于$S$矩阵的表达式

$$\langle p'|S|p\rangle = \delta(p' - p) + 2\pi\delta(E_{p'} - E_p)i\mathcal{M}(p' \leftarrow p) \tag{8.53}$$

容易理解S矩阵分解成的两项. 第一项$\delta(p' - p)$显然是粒子没有经历散射的向前散射. 第二项对应的是真实经历的散射过程. 现在散射前后粒子的动量发生了变化, 但是能量水平保持不变. 这样(8.53)式应该表现出能量守恒. 这意味着这一项应该包含有与能量守恒有关的δ函数. 在非相对论情况下没有洛伦兹不变性$E^2 - p^2 = m^2$. 能量守恒并不意味着动量守恒, $\mathcal{M}$中也不需要包括动量之间的δ函数.

利用$E_p = \frac{p^2}{2m}$, 带入上式可以得到

$$\langle p'|S|p\rangle = \delta(p'-p) + 4m\pi\delta(p'^2 - p^2)i\mathcal{M}(p' \leftarrow p) \tag{8.54}$$

这里我们利用了公式$\delta(ax) = \frac{1}{a}\delta(x)$. 由于(8.53)式中$\mathcal{M}$仅仅在$E_{p'} = E_p$时才有定义. 也就是说, 在$p'$和$p$的变量空间中, 函数$\mathcal{M}$仅仅定义在能壳$p'^2 = p^2$上. 这种情况我们称$T$矩阵为On-shell T矩阵.

由于On-shell T矩阵的散射振幅$\mathcal{M}(p' \leftarrow p)$定义仅仅在$p'^2 = p^2$上. 很显然这不是一个很好的矩阵定义. 我们期望矩阵$\langle p'|iT|p\rangle$定义在所有的p'和p上. 在后续章节, 我们会看到定义这样的一个T矩阵$\langle p'|T|p\rangle$是很方便的. 当$p'^2 = p^2$时, $\langle p'|T|p\rangle$的散射振幅$\mathcal{M}(p' \leftarrow p)$正好与On-shell T矩阵的散射振幅$\mathcal{M}(p' \leftarrow p)$一致. 这个定义在所有的$p'$和$p$上的矩阵$\langle p'|T|p\rangle$叫做Off-shell T矩阵. 只要是可以观测到的散射都只与On-shell T有关. 另一方面来说, Off-shell T矩阵是一个方便计算的工具. 特别是在Lippmann-Schwinger方程中.

8.2.6 量子散射截面

在量子散射问题中, 入射渐进态定义为$|\psi_{in}\rangle$, 出射渐进态定义为$|\psi_{out}\rangle$. 利用动量表象可以得到, 入射态和出射态的动量概率分布

$$\begin{aligned}\langle\psi_{in}|\psi_{in}\rangle &= \int d^3p\langle\psi_{in}|p\rangle\langle p|\psi_{in}\rangle = \int d^3p|\psi_{in}(p)|^2 \\ \langle\psi_{out}|\psi_{out}\rangle &= \int d^3p\langle\psi_{out}|p\rangle\langle p|\psi_{out}\rangle = \int d^3p|\psi_{out}(p)|^2\end{aligned} \tag{8.55}$$

这里我们记$\psi_{in,out}(p) = \langle p|\psi_{in,out}\rangle$. 经历散射过程后观测到粒子动量为$p$的概率为$\omega(d^p \leftarrow \psi_{in})d^3p|\psi_{out}(p)|^2$. 利用球坐标, 粒子在单位动量$\mathbf{e}_p$对应的立体角$d\Omega$出现的概率为

$$\omega(d\Omega \leftarrow \psi_{in}) = d\Omega\int_0^{+\infty} p^2dp|\psi_{out}(p)|^2 \tag{8.56}$$

我们并不知道$|\psi_{in}\rangle$的具体细节, 这个结果对于单个散射事件并没有什么价值. 我们知道的全部仅仅是波函数ψ_p在一些值p_0附近有峰, 此外正如经典的情况, 我们必须根据此部分信息对大量实验取平均值.

粒子加速器并不会反复产生相同的波包. 但是我们可以对粒子进行调整, 使所产生的的入射波包仅仅是垂直于p_0的随机横向位移有所不同. 设一系列的实验均以入射态$|\psi_{in}\rangle =$

$|\phi_b\rangle$射入, 这个入射态是由某一个确定的态$|\phi\rangle$经过平移距离b得到, 利用平移算符

$$|\phi_p\rangle = e^{-ip\cdot b}|\phi\rangle, \quad \phi_b(p) = e^{-ip\cdot b}\phi(p) \tag{8.57}$$

得到入射态ϕ_b. 我们假设所有的情况下矢量$|\phi\rangle$都一样, 但位移b在垂直于p_0方向的分布是随机的. 这种对波包$\phi(p)$的随机横向调整与经典散射的打靶距离一致. 这里我们也称其为打靶距离b

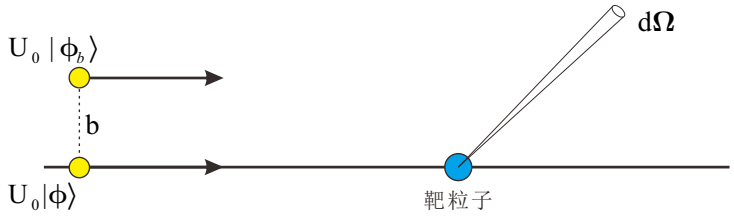


图 8.4: 垂直于平均动量方向的位移为b的两个入射波

如果重复试验次数足够多, 随机位移为$b_1, b_2, \cdots$, 则观测到散射在$d\Omega$内的粒子总数就是每一次散射概率$\omega(d\Omega \leftarrow \phi_{b,in})$的累加:

$$n_{out} = \sum_i \omega(d\Omega \leftarrow \phi_{b,in}) \tag{8.58}$$

重复次数足够多时求和可以写成积分

$$n_{out} = \int d^b J_{in}\omega(d\Omega \leftarrow \phi_b) \tag{8.59}$$

这里J_{in}是入射流密度. 实际上这个积分是对一个有限的面积求和, 但面积比远远比靶的直径要大. 因为对于足够大的b散射概率几乎为0, 所以积分可以取整个平面而不改变积分的值. 由于b是随机分布的, 入射流密度J_{in}是均匀的, 因而可以提出积分号外得到

$$n_{out} = J_{in}\int d^2b\omega(d\omega \leftarrow \phi_b) \tag{8.60}$$

与经典散射理论比较, 可以得知

$$d\sigma(d\Omega \leftarrow \phi) = \int d^2b\omega(d\Omega \leftarrow \phi_b) \tag{8.61}$$

这里$d\sigma(d\Omega \leftarrow \phi)$正是被散射到立体角$d\Omega$粒子的散射截面. 显然, 截面在量子散射里是对

于所有b的入射波ϕ_b被散射到$d\Omega$的总概率. 此外, 我们还可以把它看作是垂直于p_0平面的面积分, 每个面积元都带权重$\omega(d\Omega \leftarrow \phi_b)$. 可以看到$d\sigma(d\Omega \leftarrow \phi)$正是使入射波$\phi_b$被散射到$d\Omega$内的靶势场的有效散射截面. 在经典极限中, $d\sigma(d\Omega \leftarrow \phi)$的值要么是0要么是1, 因此积分在经典极限表示使粒子散射到$d\Omega$的真实面积.

8.2.7 量子散射截面计算

在上节我们得到量子散射跃迁概率和可观测截面之间的关系(8.61)

$$d\sigma(d\Omega \leftarrow \phi) = \int d^b\omega(d\Omega \leftarrow \phi_b) \tag{8.62}$$

我们知道散射跃迁概率$\omega(d\Omega \leftarrow \phi_b)$(8.56)为

$$\omega(d\Omega \leftarrow \phi_b) = d\Omega \int_0^{+\infty} p^2 dp |\psi_{out}(\mathbf{p})|^2 \tag{8.63}$$

式子中$\psi_{out}(\mathbf{p})$的$\mathbf{p}$方向指向$d\Omega$方向. $\psi_{out}(p)$是入射态$|\psi_{int}\rangle = |\phi_b\rangle$对应的出射渐进波函数. 我们知道入射渐进态和出射渐进态之间由S矩阵联系8.2.1. 对其两边取$\langle p|$

$$\langle p|\psi_{out}\rangle = \langle p|S|\psi_{in}\rangle = \int d^3p' \langle p|S|p'\rangle\langle p'|\psi_{in}\rangle \tag{8.64}$$

和上节一样, 为了避免计算向前散射部分, 带入$S = 1 + iT$

$$\langle p|S|p'\rangle = \delta(p - p') + 2\pi\delta(E_p - E_{p'})i\mathcal{M}(p \leftarrow p') \tag{8.65}$$

得到

$$\psi_{out}(p) = \psi_{in}(p) + 2\pi \int d^3p'\delta(E_p - E_{p'})i\mathcal{M}(p \leftarrow p')\psi_{in}(p') \tag{8.66}$$

这里$\langle p|\psi_{in,out}\rangle = \psi_{in,out}(p)$. 第一项为向前散射部分, 第二项为有效散射部分. 在我们之前的假定中, $|\psi_{in}\rangle = e^{-i\hat{p}\cdot b}|\phi\rangle$

我们观测的时候不在入射波前进方向观测, 我们要求观测方向避开p_0. 在此要求下(8.66)的第一项为0. 所以我们可以写成

$$\psi_{out}(p) = 2\pi \int d^3p'\delta(E_p - E_{p'})i\mathcal{M}(p \leftarrow p')e^{ip'\cdot b}\phi(p') \tag{8.67}$$

将上式带入(8.63)式得到

$$\omega(d\Omega \leftarrow \phi_b) = (2\pi)^2 d\Omega \int_0^{+\infty} p^2 dp \int d^3p' \delta(E_p - E_{p'}) \mathcal{M}(p \leftarrow p') e^{-ip'\cdot b} \phi(p') \\ \times \int d^3p'' \delta(E_p - E_{p''}) \mathcal{M}^*(p \leftarrow p'') e^{ip''\cdot b} \phi^*(p'') \tag{8.68}$$

所以截面$d\sigma(d\Omega \leftarrow \phi) = \int d^b \omega(d\Omega \leftarrow \phi_b)$为

$$d\sigma(d\Omega \leftarrow \phi) = (2\pi)^2 d\Omega \int d^2b \int_0^{+\infty} p^2 dp \int d^3p' \delta(E_p - E_{p'}) \mathcal{M}(p \leftarrow p') e^{-ip'\cdot b} \phi(p') \\ \times \int d^3p'' \delta(E_p - E_{p''}) \mathcal{M}^*(p \leftarrow p'') e^{ip''\cdot b} \phi^*(p'') \tag{8.69}$$

为了计算上述积分, 首先我们有

$$\int d^2b e^{i(p'-p'')\cdots b} = (2\pi)^2 \delta(p'_\perp - p \perp'') \tag{8.70}$$

式子中对b的积分是在垂直于p_0平面内的二维积分, $\delta(p'_\perp - p''_\perp)$表示在该平面的分量之间的$\delta$函数. 此外我们可以看到能量部分的$\delta$函数可以写成

$$\delta(E_{p'} - E_{p''}) = 2m\delta(p'^2 - p''^2) \tag{8.71}$$

由于$p^2 = p_\perp^2 + p_\parallel^2$, 所以$\delta$函数联合作用可以写成

$$\begin{aligned} \delta(p'_\perp - p''_\perp)\delta(E_{p'} - E_{p''}) &= 2m\delta(p'_\perp - p''_\perp)\delta(p'^2 - p''^2) \\ &= 2m\delta(p'_\perp - p''_\perp)\delta(p'^2_\perp - p''^2_\perp + p'^2_\parallel - p''^2_\parallel) \\ &= \frac{m}{p'_\parallel}\delta(p'_\perp - p''_\perp)\delta(p'_\parallel - p''_\parallel) = \frac{m}{p'_\parallel}\delta(\mathbf{p}' - \mathbf{p}'') \end{aligned} \tag{8.72}$$

值得注意, 这里由于(8.70)式保证了$p'_\perp = p''_\perp$, 所以我们将上式δ函数中的垂直分量消去. $p'_\parallel$是$\mathbf{p}'$沿$\mathbf{p}_0$方向的分量. (8.71)式中δ函数主要的作用是使$p'_\parallel = \pm p \parallel''$. 一般来说波包$\phi(p)$的长度比较窄, $p'_\parallel = -p''_\parallel$对积分无贡献.

将上式带入(8.69)式得到

$$
\begin{aligned}
D(d\Omega \leftarrow \phi) &= (2\pi)^4 d\Omega \int_0^{+\infty} p^2 dp \int d^3p' d^3p'' \delta(E_p - E'_p)\mathcal{M}(p \leftarrow p')\phi(p') \\
&\times \delta(E'_p - E''_p)\mathcal{M}^*(p \leftarrow p'')\phi^*(p'')\delta(p'_\perp - p''_\perp) \\
&= (2\pi)^4 d\Omega \int_0^{+\infty} p^2 dp \int d^3p' d^3p'' \delta(E_p - E_{p'})\mathcal{M}(p \leftarrow p')\phi(p')\mathcal{M}^*(p \leftarrow p'')\phi^*(p'')\frac{m}{p'_\parallel}\delta(\mathbf{p}' - \mathbf{p}'') \\
&= (2\pi)^4 m d\Omega \int_0^{+\infty} p^2 dp \int d^3p' \delta(E_p - E_{p'})\frac{1}{p'_\parallel}|\mathcal{M}(p \leftarrow p')\phi(p')|^2
\end{aligned}
$$

注意, 第一步利用了替换$\delta(x-y)\delta(x-z) = \delta(x-y)\delta(y-z)$. 这是由于对于$E_{p'} \neq E_{p''}$的情况, 积分值均是0. 利用关系式$\delta(E_p - E_{p'}) = 2m\delta(p^2 - p'^2) = \frac{m}{p'}\delta(p-p')$, 上式对$p$积分得到

$$
d\sigma(d\Omega \leftarrow \phi) = (2\pi)^4 m^2 d\Omega \int d^3p' \frac{p'}{p'_\parallel}|\mathcal{M}(p \leftarrow p')\phi(p')|^2
$$

式子中$|p| = |p'|$.

最后如果以p_0为波包中心的$\phi(p')$足够窄, 散射振幅$\mathcal{M}(p \leftarrow p')$可以认为在$p_0$附近近似不变, 这样被积函数里$\mathcal{M}(p \leftarrow p')$和$\frac{p'}{p'_\parallel}$均可以用$\sqrt{}_0$代替.

$$
\begin{aligned}
d\sigma(d\Omega \leftarrow \phi) &= (2\pi)^4 m^2 d\Omega |\mathcal{M}(\mathbf{p} \leftarrow \mathbf{p}_0)|^2 \int d^3p' |\phi(\mathbf{p}')|^2 \\
&= (2\pi)^4 m^2 d\Omega |\mathcal{M}(\mathbf{p} \leftarrow \mathbf{p}_0)|^2
\end{aligned}
\tag{8.73}
$$

最后一步利用了归一化条件$\int d^3p|\phi(\mathbf{p})|^2 = 1$

从截面和散射振幅的表达式可以看到, 只要波包$\phi(p)$在p_0附近足够狭窄, 计算结果就与ϕ的波包形状无关, 所以我们可以把散射截面$d\sigma(d\Omega \leftarrow \phi)$写成$d\sigma(d\Omega \leftarrow p_0)$. 与经典散射类似, 微分截面定义为

$$
D(\mathbf{p} \leftarrow \mathbf{p}_0)\frac{d\sigma}{d\Omega}
$$

式中$\mathbf{p}_0$方向为观测方向$d\Omega$对应的方向. 因此量子微分截面表示为

$$
D(p \leftarrow p_0) = \frac{d\sigma}{d\Omega} = (2\pi)^4 m^2 |\mathcal{M}(p \leftarrow p_0)|^2 \tag{8.74}
$$

这就是量子力学里微分截面和散射振幅之间的关系.

8.2.8 光学定理

我们之前建立了散射截面和散射振幅之间的关系(8.74). 现在我们进一步证明一个新的定理——光学定理. 这个定理可以通过S矩阵的幺正性和矩阵方程$S^\dagger S = 1$的对角元直接得到. 在之前, 我们将S算子拆成向前部分和散射部分$S = 1 + iT$. 带入$S^\dagger S = 1$得到$iT - iT^\dagger = -T^\dagger T$. 在动量表象中表示为

$$\langle p'|iT|p\rangle - \langle p'|iT^\dagger|p\rangle = -\langle p'|T^\dagger T|p\rangle$$

利用$\langle a|b\rangle = \langle b|a\rangle^*$得到

$$\begin{aligned} i(\langle p'|T|p\rangle - \langle p|T|p'\rangle^*) &= -\langle p'|T^\dagger T|\rangle p\rangle \\ &= -\int d^3p''\langle p'|T^\dagger|p''\rangle\langle p''|T|p\rangle \\ &= -\int d^3p''\langle p''|T|p'\rangle^*\langle p''|T|p\rangle \end{aligned}$$

带入$\langle p'|iT|p\rangle = 2\pi\delta(E_{p'} - E_p)i\mathcal{M}(p' \leftarrow p)$得到

$$\begin{aligned} &2\pi i\delta(E_{p'} - E_p)\left[\mathcal{M}(p' \leftarrow p) - \mathcal{M}^*(p \leftarrow p')\right] \\ &\quad = -\int d^3p''2\pi\delta(E_{p''} - E_{p'})\mathcal{M}^*(p'' \leftarrow p')2\pi\delta(E_{p''} - E_p)\mathcal{M}(p'' \leftarrow p) \end{aligned}$$

上式右边$\delta(E_{p''} - E_p)$提出积分外得到

$$\begin{aligned} &2\pi i\delta(E_{p'} - E_p)\left[\mathcal{M}(p' \leftarrow p) - \mathcal{M}^*(p \leftarrow p')\right] \\ &\quad = -\int d^3p''(2\pi)^2\delta(E_{p''} - E_{p'})\delta(E_{p''} - E_p)\mathcal{M}^*(p'' \leftarrow p')\mathcal{M}(p'' \leftarrow p) \end{aligned}$$

满足条件$E_{p'} = E_p$时, 上式两边可以消去$2\pi\delta(E_{p'} - E_p)$. 得到

$$i\left[\mathcal{M}(p' \leftarrow p) - \mathcal{M}^*(p \leftarrow p')\right] = -\int d^3p''2\pi\delta(E_{p'_,} - E_{p'})\mathcal{M}(p'' \leftarrow p)\mathcal{M}^*(p'' \leftarrow p')$$

最后令$\mathbf{p}' = \mathbf{p}$, 上式可以化简为

$$\begin{aligned} -2\mathrm{Im}\mathcal{M}(\mathbf{p} \leftarrow \mathbf{p}) &= -\int d^3p''2\pi\delta(E_{p''} - E_{p'})|\mathcal{M}(p'' \leftarrow p)|^2 \\ &= -2\pi\int \frac{m}{p}\delta(p'' - p)|\mathcal{M}(\mathbf{p}'' \leftarrow \mathbf{p})|^2 p''^2 dp'' d\Omega \\ &= -2\pi mp\int d\Omega|\mathcal{M}(\mathbf{p}'' \leftarrow \mathbf{p})|^2 \end{aligned} \tag{8.75}$$

值得注意$\mathcal{M}(\mathbf{p}' \leftarrow \mathbf{p})$并不包含$\mathbf{p}$大小有关的量, 与$\mathbf{p}$大小有关的量已经包含在了$\delta(E_{p'} - E_p)$中. 上面推导我们利用了$E_p = \frac{p^2}{2m}$和$\delta(f(x) - f(x_0)) = \frac{1}{|f(x_0)|}\delta(x - x_0)$. 由微分截面和散射振幅之间的关系(8.74)式

$$D(\mathbf{p}'' \leftarrow \mathbf{p}) = \frac{d\sigma}{d\Omega}(\mathbf{p}'' \leftarrow \mathbf{p}) = (2\pi)^4 m^2|\mathcal{M}(\mathbf{p}'' \leftarrow \mathbf{p})|^2$$

两边同时对立体角积分, 我们得到总截面$\sigma(\mathbf{p}) = (2\pi)^4 m^2 \int d\Omega|\mathcal{M}(\mathbf{p}'' \leftarrow \mathbf{p})|^2$, 带入(8.75)式得到

$$\mathrm{Im}\mathcal{M}(p \leftarrow p) = \frac{p}{2(2\pi)^3 m}\sigma(p)$$

光学定理表述了向前散射振幅的虚部正比于总截面$\sigma(p)$. 这可以推导出一系列有用的结论.

1. 振幅一般来说不是纯实数, 并且在向前散射方向有正定的虚部;

2. 利用光学定理, 可以在向前散射方向分别测量$\mathcal{M}(p' \leftarrow p)$的虚部和实部.

3. 光学定理可以用于色散关系, 色散关系可以把振幅的实部表示为其虚部的积分. 利用光学定理, 这意味着我们可以把向前散射振幅表示为总截面的积分, 因而得到一个能直接用实验检验的关系. 这种关系的重要性在于, 它的成立于相互作用准确细节无关, 于直接测量的量有关

8.3 无自旋粒子两体散射

之前的章节里, 我们已经用散射算符S和散射振幅$\mathcal{M}$完全的描述了单个无自旋粒子的散射. 现在我们进一步推广到两个无自旋粒子的弹性散射过程. 本节主要目的是证明一个直观上的结果: 质心系中两粒子的弹性散射和单个粒子被固定势的散射是等价的. 首先先引入二粒子态的Hilbert空间和两个单体粒子的Hilbert之间的关系.

8.3.1 两体波函数

两个不同性质的无自旋粒子态用与两个粒子位置$\mathbf{r}_1, \mathbf{r}_1$有关的波函数$\psi(\mathbf{r}_1, \mathbf{r}_2)$表示. 一个比较特殊的情况是$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)$的乘积形式. 相应的态矢量表示为$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$叫做直积态. 这里$|\psi\rangle$是二粒子希尔伯特空间$\mathcal{H}$的一个矢量, 而$|\psi_1\rangle$是第一个粒子单粒子空间$\mathcal{H}_1$的矢量, $|\psi_2\rangle$是第二个粒子单粒子空间$\mathcal{H}_2$的矢量. 直积态表示一个二粒子态, 第一个粒子在$|\psi_1\rangle$态, 第二个粒子在$|\psi_2\rangle$态.

直积态有两个比较重要的性质:

1. 若$\{|n_1\rangle\}$和$\{|\psi_2\rangle\}$是单粒子空间$\mathcal{H}_1$和$\mathcal{H}_2$的正交归一基, 则直积态

$$|n_1\rangle \otimes |n_2\rangle, \quad (n_1, n_2 = 1, 2, \cdots)$$

构成二粒子空间$\mathcal{H}$的正交归一基底

2. 内积满足关系

$$(\langle\psi_1| \otimes \langle\psi_2|)(\psi_1'\rangle \otimes |\psi_2'\rangle) = \langle\psi_1|\psi_1'\rangle\langle\psi_2|\psi_2'\rangle$$

利用这种方法由两个子空间$\mathcal{H}_1$和$\mathcal{H}_2$构造的二粒子空间$\mathcal{H}$叫做这两个子空间构造的直积空间$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. 值得注意这个等式并不意味着$\mathcal{H}$中每一个态矢都是直积态矢, 它只意味着$\mathcal{H}$是由直积态张开的空间, 每一个$|\phi\rangle \in \mathcal{H}$都可以由这些直积态基底线性表示.

张量积一般出现在多独立自由度的系统中. 例如有自旋s的单粒子空间, 对于这种系统

$$\mathcal{H} = \mathcal{H}_{space} \otimes \mathcal{H}_{spin}$$

这里$\mathcal{H}_{space} \in L^2(\mathbb{R}^3)$, 也就是不带自旋的波函数空间, 而$\mathcal{H}_{spin}$是$2s+1$维的自旋空间. 所以$\mathcal{H}$的态矢$\psi_m(\mathbf{r})$由坐标$\mathbf{r}$和自旋标记$m(S_z$本征值$)$标记, m标记自旋的某个基底. 通常将这些函数结合成由$2s+1$个分量的旋量波函数.

直积空间$\mathcal{H}_1 \otimes \mathcal{H}_2$可以看成是$\mathcal{H}$的分解. 但是空间$\mathcal{H}$的分解的方式有很多. 例如每个波函数$\psi(\mathbf{r}_1, \mathbf{r}_2)$可以表为质心坐标和相对坐标的函数:

$$\mathbf{r}_c = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}m, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$$

显然函数$\psi(r_c, r)$的空间可以用形为$\psi_1(r_c)\psi_2(r)$的直积函数张开

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathcal{H}_c \otimes \mathcal{H}_r$$

其中$\mathcal{H}_c$和$\mathcal{H}_r$分别是质心坐标和相对质心坐标的波函数空间. 对于许多散射问题来说质心分解更为有用.

我们现在定义算子的直积, 和$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$一样, 某些算子也是直积算子.

定义 8.3.1 *若A和B是分别作用于$\mathcal{H}_1$和$\mathcal{H}_2$的任意算子, 我们定义作用在$\mathcal{H}$上的算子$A \otimes B$:*

$$(A \otimes B)(|\psi_1\rangle \otimes |\psi_2\rangle) = A|\psi_1\rangle \otimes B|\psi_2\rangle$$

准确说, 一个两体系统的基本力学量形式上都可以写成$A \otimes 1$或者$1 \otimes B$. 例如二粒子空间$\mathcal{H}$上的第一个粒子的动量算子作用在$\mathcal{H}$上可以表示为

$$p_1(\mathcal{H}) = (p_1\mathcal{H}_1) \otimes (1\mathcal{H}_2)$$

这就是说p_1算符在二粒子空间中应该表示为$p_1 \otimes 1$, 其作用在二粒子态$|\psi_1\rangle \otimes |\psi_2\rangle$上为

$$(p_1 \otimes 1)(|\psi_1\rangle \otimes |\psi_2\rangle) = (p_1|\psi_1\rangle)(1|\psi_2\rangle)$$

类似地, 对于p_2而言可以写作$1 \otimes p_2$

正如所想, 任何两个不同自由度空间的算符$A \otimes 1$和$1 \otimes B$是对易的

$$A \otimes 1 \cdot 1 \otimes B(|\psi_1\rangle \otimes |\psi_2\rangle) - 1 \otimes B \cdot A \otimes 1(|\psi_1\rangle \otimes |\psi_2\rangle) = 0$$

一个众所周知的例子便是单粒子动量算符$\mathbf{p}_1 = -i\nabla_1$和$\mathbf{p}_2 = -i\nabla_2$. 一般来说, 我们在记号上为了方便并不区分$A \otimes 1$和$A$. 因此$A \otimes B$和$(A \otimes 1)(1 \otimes B)$是一样的, 简写成$AB$. 但当我们涉及到张量乘积算符的乘法结构时就需要恢复张量积的记号.

我们现在考虑二粒子系统的哈密顿算子和相应的时间演化算子. 哈密顿算子表示为

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V = H_0 + V$$

这里我们假定相互作用V只与相对坐标$\mathbf{r}$有关. 我们把H用总动量和相对动量表示

$$\mathbf{p}_c = \mathbf{p}_1 + \mathbf{p}_2, \quad \mathbf{p} - \frac{m_1\mathbf{p}_1 + m_2\mathbf{p}_2}{m_1 + m_2}$$

相应的共轭坐标和相对坐标

$$\mathbf{r}_c = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{\mathbf{r}_1 + \mathbf{r}_2}, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$$

利用链式法则$\mathbf{p}_1 = -i\frac{\partial}{\partial \mathbf{r}_1} = -i(\frac{\partial \mathbf{r}_1}{\mathbf{r}_c} + \frac{\partial \mathbf{r}_1}{\mathbf{r}})\frac{\partial}{\partial \mathbf{r}_1}$我们可以得到

$$\mathbf{p}_1 = \frac{m_1 \mathbf{p}_c}{m_1 + m_2} + \mathbf{p}, \quad \text{同样地} \quad \mathbf{p}_2 = \frac{m_2}{m_1 + m_2}\mathbf{p}_c - \mathbf{p}$$

带入H得到

$$H = \frac{\mathbf{p}_c^2}{2(m_1 + m_2)} + \frac{\mathbf{p}^2}{2m_1} + \frac{\mathbf{p}^2}{2m_2}$$

定义总质量$M = m_1 + m_2$和折合质量$m = \frac{m_1 m_2}{m_1 + m_2}$. 系统哈密顿可以简写为

$$H = \frac{\mathbf{p}_c^2}{2M} + \left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r})\right] = H_c + H_r$$

显然H是两项之和, 第一项只作用于$\mathcal{H}_c$第二项只作用于$\mathcal{H}_r$. 这两项分别属于不同的自由度空间分别相互对易. 所以时间演化算符也可以分解

$$U(t) = e^{iHt} = e^{-iH_c t}e^{-iH_r t} = e^{-iH_c t} \otimes e^{-iH_r t} \tag{8.76}$$

从表达式可以看出, 时间演化算符一部分属于$\mathcal{H}_c$一部分属于$\mathcal{H}_r$. 这说明质心运动与相对运动无关. 由于H_c就是$\frac{pp_c^2}{2M}$, 质心运动可以看做一个质量为M的自由粒子. 同样地H_r可以看成一个质量为m的单粒子处于固定势$V(\mathbf{r})$的运动. 所以$\mathcal{H}_r$中的演化过程应该与前几节讨论的固定势情况完全类似.

特别地, 自由两体演化算符也可以类似的分解

$$U_0(t) = e^{-iH_0 t} = e^{-iH_c t} \otimes e^{-iH_{0,r}}$$

同样的自由两体演化算符$U_0(t)$也可以分解为

$$U_0(t) = e^{-iH_1 t} \otimes e^{-H_2 t} \tag{8.77}$$

与$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$相对应. 这个结果表示了一个明显的事实, 两个无相互作用粒子的独立运动, 这个结果对于有相互作用的粒子并不适用.

8.3.2 两体S矩阵

两体散射问题与之前的固定势场散射问题看起来似乎不同. 两体散射问题在$L^2(\mathbb{R}^6)$表述的, 而固定势散射问题在$L^2(\mathbb{R}^3)$中表述的. 但是在上节, 我们利用质心分解$\mathcal{H} = \mathcal{H}_c \otimes \mathcal{H}_r$可以看到这两个问题实质上是等价的.

两体系统演化轨道$U(t)|\psi\rangle$, 这里$|\psi\rangle$是两体系统Hilbert空间$\mathcal{H}$中的任意态矢. 类似于单体固定势情况, 我们可以预测当$t \to \mp\infty$时, 两个粒子之间相互作用已经足够微弱, 以至于$U(t)|\psi\rangle$的演化近似为自由演化$U_0(t)|\psi\rangle$. 和单粒子情况一样, 这个预测毫无疑问是正确的, 并且可以用渐进条件和渐进完全性两种形式表述

定理 8.3.1 (渐进条件) *$\mathcal{H}$中每个矢量$|\psi_{in(out)}\rangle$是某个真实演化矢量$U(t)|\psi\rangle$的入(出)射渐进态*

$$\lim_{t\to-\infty(+\infty)}[U(t)|\psi\rangle - U_0(t)|\psi_{in}\rangle] = 0$$

为了证明上式, 我们需要证明对于每一个$|\psi_{in}\rangle$, 矢量$U^\dagger(t)U_0(t)|\psi_{in}\rangle$存在极限, 也就是说算子$U^\dagger(t)U_0(t)$收敛. 我们首先用(8.76)式和(8.77)式来表示$U^\dagger(t)U_0(t)$

$$U^\dagger(t)U_0(t) = (U_c(t)\otimes U_r(t))^\dagger U_c(t)\otimes U_{0,r}(t) = 1_c \otimes U_r^\dagger(t)U_{0,r}(t)$$

这里1_c表示空间$\mathcal{H}_c$的单位算子

这个算子只需要第二个因子存在极限, 它就存在极限. 而第二个因子对应的是一个单粒子系统, 单粒子系统的渐进极限在之前早已说明. 所以第二个因子确实存在极限, 我们记作$\Omega_\pm$, 入射渐进态和出射渐进态对应的实际演化轨道由下式给出

$$|\psi\rangle = \lim_{\mp} U^\dagger(t)U_0(t)|\psi_{in,out}\rangle = (1_c\otimes\Omega_+)|\psi_{in.out}\rangle$$

为了表示两体空间$\mathcal{H} = L^2(\mathbb{R}^6)$上的两体Moller算子, 我们引入记号$\Omega_\pm$它们形式为

$$\mathbf{\Omega}_\pm = 1_c \otimes \Omega_\pm \tag{8.78}$$

其中$\mathbf{\Omega}_\pm = \lim_{t\to\mp} e^{iH_r}e^{-iH_{0,r}t}$, 作用在空间$L^2(\mathbb{R}^3)$上, 并且具有固定势场Moller算子的结构性质. $\mathbf{\Omega}_\pm$中的单位算子1_c反映了质心运动类似于一个自由粒子的运动而不发生散射.

类似于之前固定势一样, 具有渐进入射态的轨道和出射渐进态的轨道是精确相同的, $\Omega_\pm$都把把$\mathcal{H}$映射到$\mathcal{R}$, 其次散射态空间$\mathcal{R}$和束缚态空间$\mathcal{B}$的直和是全空间$\mathcal{H} = \mathcal{R}\oplus\mathcal{B}$

同样的可以定义$\mathbf{S}$算子

$$\mathbf{S} = \mathbf{\Omega}_-^\dagger \mathbf{\Omega}_+$$

该算子是将$\mathcal{H}$任意态$|\psi_{in}\rangle$映射到相应的$|\psi_{out}\rangle = \mathbf{S}|\psi_{in}\rangle$的幺正算子. 由(8.78)式可知, 它具有简单的张量积结构

$$\mathbf{S} = 1_c \otimes S$$

其中$S = \Omega_-^\dagger\Omega_+$作用在$\mathcal{H}_r$上, 所以问题简化成在$\mathcal{H}_r$上计算单粒子$S$算子.

8.3.3 T矩阵

由于$\mathbf{S} = 1_c \otimes S$, $\mathbf{S}$与$\mathbf{P}_c = 1 \otimes P_c$对易. 所以总动量守恒. 和单粒子情况一样, $[\mathbf{S}_0, H_0] = 0$因而系统能量守恒. 同样的做法, 利用动量表象得出矩阵元

$$\langle p_1', p_2' | S | p_1, p_2 \rangle$$

含有因子$\delta(E_1' + E_2' - E_1 - E_2)\delta(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_1' - \mathbf{p}_2')$, 其中$E_i = \frac{\mathbf{p}_i^2}{2m_i}$.

同样我们在能壳上分解这个矩阵元. 注意到$\mathbf{p}_1$和$\mathbf{p}_2$的本征矢$|\mathbf{p}_1, \mathbf{p}_2\rangle$也是总动量$\mathbf{p}_c$和相对动量$\mathbf{p}$的本征矢量

$$|\mathbf{p}_1, \mathbf{p}_2\rangle = |\mathbf{p}_c\rangle \otimes |\mathbf{p}\rangle$$

这里$\mathbf{p}_c = \mathbf{p}_1 + \mathbf{p}_2$, $\mathbf{p} = \frac{m_2\mathbf{p}_1 - m_1\mathbf{p}_2}{m_1 + m_2}$. 利用这个关系可以分解

$$e^{i(\mathbf{p}_1 \cdot \mathbf{r}_1 + \mathbf{p}_2 \cdot \mathbf{r}_2)} = ei(\mathbf{p}_c \cdot \mathbf{r}_c + \mathbf{p} \cdot \mathbf{r}) = e^{i\mathbf{p}_c \cdot \mathbf{x}_c} e^{i\mathbf{p} \cdot \mathbf{r}}$$

同样的在质心系分解S算子$1_c \otimes S$, 可以看出

$$\begin{aligned} \langle \mathbf{p}_1', \mathbf{p}_2' | \mathbf{S} | \mathbf{p}_1, \mathbf{p}_2 \rangle &= \langle \mathbf{p}_c', \mathbf{p} | (1 \otimes \mathbf{S}) | \mathbf{p}_c, \mathbf{p} \rangle \\ &= \delta(\mathbf{p}_c' - \mathbf{p}_c) \langle \mathbf{p}' | S | \mathbf{p} \rangle \end{aligned}$$

S算子具有单粒子算子的结构, 矩阵元可以分解成

$$\langle \mathbf{p}' | S | \mathbf{p} \rangle = \delta(\mathbf{p}' - \mathbf{p}) + 2\pi\delta(E_{p'} - E_p) i\mathcal{M}(\mathbf{p}' \leftarrow \mathbf{p})$$

将上面两个方程联合起来. 得到

$$\langle \mathbf{p}_1', \mathbf{p}_2' | \mathbf{S} | \mathbf{p}_1, \mathbf{p}_2 \rangle = \delta(\mathbf{p}_1' - \mathbf{p}_1)\delta(\mathbf{p}_2' - \mathbf{p}_2) - 2\pi(\sum E_i' - \sum E_i)\delta(\sum \mathbf{p}_i' - \sum \mathbf{p}_i)\mathcal{M}(\mathbf{p}' \leftarrow \mathbf{p}) \tag{8.79}$$

这里第一项表示向前散射部分, 这部分对应的是自由哈密顿. 所以$|\mathbf{p}_1, \mathbf{p}_2\rangle = |\mathbf{p}_1\rangle \otimes |\mathbf{p}_2\rangle = |\mathbf{p}_c\rangle \otimes |\mathbf{p}\rangle$, 这也是第一项$\delta(\mathbf{p}_c' - \mathbf{p}_c)\delta(\mathbf{p}' - \mathbf{p}) = \delta(\mathbf{p}_1' - \mathbf{p}_1)\delta(\mathbf{p}_2' - \mathbf{p}_2)$的原因. 第二项是两体真正被散射的部分, 这一项里保持总动量和总能量守恒. 相对动量的各个分量并不守恒.

这里我们做个符号说明, 我们对于满足平移不变的两体问题算符$\mathbf{\Omega}_\pm$和$\mathbf{S}$用的是粗体字, 而描述相对运动的算符$\Omega_\pm$和S用的是普通字体. 算符$\mathbf{\Omega}_\pm$和$\mathbf{S}$作用于全部空间$L^2(\mathbb{R}^6)$, 矩阵元含有动量守恒的$\delta(\mathbf{p}_c' - \mathbf{p}_c)$. 算符$\Omega_\pm$和$S$作用于$L^2(\mathbb{R}^3)$并且精确地和单粒子算子

的结构一致.

整本书我们用粗体字表示平移不变系统的的Moller算子和S算子, 这些算子含有$\delta(\mathbf{p}_c' - \mathbf{p}_c)$保证了总动量的守恒. 普通字体$\Omega_\pm$和$S$适用于平移不变系统的相对运动, 或者固定势运动. 用这种方法我们保证$\Omega_\pm$和S总有相同的结构; 特别是它们矩阵元不包含所有动量守恒的δ函数. 在这两种情况中, On-shell T矩阵$\mathcal{M}(\mathbf{p}' \leftarrow \mathbf{p})$定义, 并且不包含$\delta$函数.

8.3.4 各种参考系下的截面

两体散射截面可以定义在各种参考系下. 大部分实验的靶粒子在初始时刻都是静止的. 因此由这个条件所定义的参考系叫做实验系. 为了理论上的目的, 最方便的参考系是静止的质心. 这种参考系叫做质心系. 在某些情况下, 有必要讨论两个任意动量运动粒子的参考系.

我们发现, 截面可以定义成被散射到立体角$d\Omega$内的粒子数在所有参考系中均有

$$n_{out} d\Omega(\mathbf{p}_1, \mathbf{p}_2) = J_{in}(\mathbf{p}_1, \mathbf{p}_2) d\sigma(\mathbf{p}_1, \mathbf{p}_2)$$

式中两粒子初始动量$\mathbf{p}_1$和$\mathbf{p}_2$必须保证在同一个参考系中, 所有的物理量都在此参考系中测量.

我们首先在质心系中考虑一个散射过程. 在上章我们知道散射带任意$d\Omega$的粒子流正比于入射粒子流密度J_{in}. 所以我们在质心系中可以定义截面$\sigma(d\Omega \leftarrow \mathbf{p}_1, \mathbf{p}_2)$.

这样定义的质心系截面保证了关系式$n_{out} d\Omega = J_{in} d\sigma$成立. 我们现在在另一个参考系里观测同样的一个碰撞过程. 为了确定我们观测到的物理量, 我们必须考虑如何将上式所含的物理量从一个参考系变换到另一个参考系. 质心系O观测的动量为$\mathbf{p}_1, \mathbf{p}_2$(显然质心系中必须有$\mathbf{p}_1 = -\mathbf{p}_2$). 另一个参考系$O'$中观测到的动量记作$\mathbf{p}_1', \mathbf{p}_2'$. 类似的在参考系$O$中的立体角元$d\Omega$在参考系$O'$中被称为$d\Omega'$. 一般来说两个不同的参考系观测到的立体角大小和方向都不相同. 但不管是哪个参考系观测者观测到的粒子数必定是相同的. 这就是说, 任意两个观测者O和O'

$$n_{out}(d\Omega' \leftarrow \mathbf{p}_1', \mathbf{p}_2') = n_{out}(d\Omega \leftarrow \mathbf{p}_1, \mathbf{p}_2)$$

n_{out}是不变的.

现在考虑入射流J_{in}的变换. 首先我们假设第二个参考系中的入射动量$\mathbf{p}_1'$和$\mathbf{p}_2'$在同一直线上(例如在实验系中). 在这种情况下, 穿过垂直于入射动量方向的单位面的入射粒子

总数在两个参考系中是相同的, 也就是说两个参考系中观测到的入射流是一样的

$$J_{in}(\mathbf{p}_1', \mathbf{p}_2') = J_{in}(\mathbf{p}_1, \mathbf{p}_2) \tag{8.80}$$

我们考虑某个参考系, 在此参考系中两个粒子倾斜的靠近. 但根据入射流的定义, 入射流是垂直于入射粒子和靶的相对运动放的入射流密度, 在任何参考系中$J_in(\mathbf{p}_1', \mathbf{p}_2')$和$J_{in}(\mathbf{p}_1, \mathbf{p}_2)$完全相同. 所以对于任何参考系(8.80)式总是成立的.

现在我们已经在质心系中定义了截面, 散射粒子数密度由关系$n_{out}d\Omega = J_{in}d\sigma$给出. 而物理量$n_{out}$和$J_{in}$在所有参考系中测量到的结果是不变的. 所以我们可以在任意参考系中定义截面, 使得$n_{out}d\Omega = J_{in}d\sigma$. 用

8.4 自旋粒子两体散射

8.4.1 自旋粒子的Hilbert空间

8.4.2 自旋粒子的S算子

8.4.3 散射振幅和散射矩阵

8.4.4 自旋求和

8.4.5 旋量的入射出射

8.5 Dyson公式

8.6 Born近似

Chapter 9

多粒子量子系统

9.1 全同粒子与交换对称性

9.2 置换群与杨图、杨算子

9.3 二次量子化

9.4 电磁场正则量子化

9.5 Casmir效应

Part II

相对论量子力学

Chapter 10

连续场基本方程

10.1 欧拉-拉格朗日方程

10.2 诺特原理与诺特流

10.3 克莱因-高登方程

10.4 克莱因-高登方程正则量子化

Chapter 11

洛伦兹群

11.1 洛伦兹变换

11.2 **Clifford**代数的表示——γ矩阵

11.3 狄拉克方程

11.4 狄拉克方程正则量子化

11.5 洛伦兹群的分离对称性与电荷对称性——$\mathcal{CPT}$联合不变性

Chapter 12

微扰理论

12.1 Gellman-Low公式与真空

12.2 Wick定理

12.3 Feynman图与费曼规则

12.3.1 ϕ^4的费曼规则

12.3.2 QED的费曼规则

12.4 截面与S矩阵元

Chapter 13

树图计算

13.1 $e^+e^- \to \mu^+\mu^-$

13.2 康普顿散射

13.3 库仑散射

Chapter 14

圈图计算

14.1 顶角近似

14.2 辐射修正,正规化方案——**PV**方案

14.3 场强重整化

14.4 **LSZ**约化定理

14.5 光学定理

14.6 **Ward-Takahashi**恒等式

14.7 电荷重整化

Part III

规范理论

Chapter 15

$U(1)$规范理论

Chapter 16

$SU(n)$规范理论

16.1 规范理论的内禀空间与底流形

16.2 规范自由与纤维丛

16.3 规范固定与截面

16.4 **Yang-Mills方程**

16.5 **Wilson圈**

Chapter 17

路径积分量子化

17.1 量子力学的路径积分

17.2 标量场的量子化

17.3 QED的量子化

17.4 旋量场的量子化

Chapter 18

$Non-Abel$规范场量子化

18.1 Faddeev-Poppov路径积分量子化

18.2 鬼

18.3 BRST对称性

18.4 $Non-Abel$场的单圈修正

18.5 渐进自由

Chapter 19

量子色动力学

19.1 夸克理论

19.2 胶子辐射

19.3 强子散射:轻子对

Chapter 20

微扰理论

20.1 二维手征流

20.2 四维手征流

20.3 Goldstone定理

20.4 手征流反常

Chapter 21

自发对称性破缺

21.1 Higgs机制

21.2 弱相互作用的Glashow-Weinberg-Salam理论