

目录

第一章 数学导论	1
1.1 线性向量空间：基底	1
1.2 内积空间	8
1.3 对偶空间与狄拉克符号	12
1.3.1 向量的在规范正交基上的展开	13
1.3.2 伴随运算	14
1.4 子空间	18
1.5 线性算符	19
1.6 线性算子的矩阵元	22
1.7 主动和被动转换	30
1.8 特征值问题	31
1.9 算符函数及相关概念	56
1.10 无限维的推广	59
第二章 回顾经典力学	75
2.1 最小作用原理和拉格朗日力学	75
第三章 自旋	83
3.1 导言	83
3.2 自旋的本质是什么？	83
附录 A 附录	85
A.1 矩阵求逆	85

第一章 数学导论

本书的目的是从量子力学的公理开始为您提供量子力学的入门知识。本章的目的是给你配备必要的数学工具。你所需要的所有数学都是在这里发展的，从一些假定你已经知道的基本向量的和矩阵开始。本书提供了许多与经典力学相关的例子和练习，二者既帮助你数学中解脱出来，又展示了这里数学思想的广泛适用性。你在这一章中所付出的努力是值得的：它不仅会让你为这门课程做好准备，还会将你可能零碎学到的许多想法统一起来。要真正学好这一章，你必须像其他任何一章节一样解决一些问题。

1.1 线性向量空间：基底

在本节中，您将会被介绍一些线性向量空间的概念。您肯定熟悉物理中这些带有长度、速度方向、力、位移、力矩等等的箭头。你知道如何将它们相加并乘以标量，以及这些运算所遵循的规则。例如，你知道标量乘法可分配的：两个向量之和的倍数就是倍数之和。我们想做的是从这个简单的例子中抽象出一组基本性质或公理，并且可以说任何服从这组相同条件的对象都形成一个线性向量空间。聪明之处在于在概括中决定保留哪些性质。如果你保留太多，将就没有例子能够存在，如果您保留太少，将不会从公理中推导出一些有意思的东西。

接下来是数学家们所明智选择后作为线性空间的必要条件的性质列表。当您阅读它们时，请将它们与所有箭头进行比较，并确保这些确实是这些熟悉的向量所拥有的性质。但是，还要注意，这里明显缺少每个矢量都有大小和方向的要求，这是我们脑海中第一个也是最显著的特征。因此，您可能会认为，删除这一要求时，就像婴儿被连同洗澡水一起倒掉了。然而，随着您继续前进并在向量空间的标题下看到了多种思想的统一和综合，您将有足够的时间来欣赏这种选择背后的智慧。您将看到涉及实体的矢量空间的例

子，你无法直接地他们将视为具有大小或方向的实体。虽然您应该对所有这一切都留有印象，但请记住，根据箭头去思考这些概括推广并使用直觉来证明定理或至少预测它们没有一点儿坏处。

定义 1 一个线性空间 \mathbb{V} 是存在且称作向量 $|1\rangle, |2\rangle, \dots, |v\rangle, \dots, |w\rangle$ 的一个集合

1. 向量求和的一个明确性规定，记作 $|V\rangle + |W\rangle$
2. 标量乘法的一个明确性规定，记作 $a|V\rangle$ 且伴随如下性质
 - 这些操作的结果是该线性空间的另一个基础元素，一个叫做封闭性的特征 $|V\rangle + |W\rangle \in \mathbb{V}$
 - 标量乘法在向量中是可分配的： $a(|V\rangle + |W\rangle) = a|V\rangle + a|W\rangle$
 - 标量乘法在标量中是可分配的： $(a + b)|V\rangle = a|V\rangle + b|V\rangle$
 - 加法服从交换律： $|V\rangle + |W\rangle = |W\rangle + |V\rangle$
 - 加法服从结合律： $|V\rangle + (|W\rangle + |Z\rangle) = (|V\rangle + |W\rangle) + |Z\rangle$
 - 存在一个零向量 $|0\rangle$ 使得 $|V\rangle + |0\rangle = |V\rangle$
 - 对于每个向量 $|V\rangle$ 来说存在一个加法下的负元 $|-V\rangle$ 使得 $|V\rangle + |-V\rangle = 0$

有一个很好的方法可以记住所有这些性质；那就是顺其自然。

定义 2 数字 a, b, \dots 被称为定义在向量空间上的域

如果该域由所有实数组成，则我们有一个实向量空间，如果它们是复数，我们有一个复数向量空间。矢量本身既不是实数的也不是复数的；这个形容词只适用于标量。

现在让我们指出上述公理所暗指性质

- $|0\rangle$ 是唯一的等等，如果 $|0'\rangle$ 具有 $|0\rangle$ 所有性质，那么 $|0\rangle = |0'\rangle$
- $0|V\rangle = |0\rangle$
- $-|V\rangle = |-V\rangle$

- $|-V\rangle$ 是 $|V\rangle$ 唯一的可加负元

这些证明将会被留下作为随后的练习，你可以不必知道怎么证明但是你必须清楚以上陈述的这些性质。

练习 1.1.1 证明这些所声称的性质，首先考虑 $|0\rangle + |0'\rangle$ 并且依次使用两个零向量展现的性质。对于第二项从 $|0\rangle = (0+1)|V\rangle + |-V\rangle$ 。对于第三项，从 $|V\rangle + (-|V\rangle) = 0 = |0\rangle$ 。对最后一项，让 $|W\rangle$ 也满足 $|V\rangle + |W\rangle = |0\rangle$ 。因为 $|0\rangle$ 具有唯一性，这意味着 $|V\rangle + |W\rangle = |V\rangle + |-V\rangle$

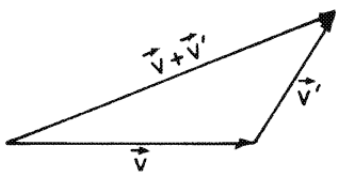


图 1.1: 向量加法规则，注意它遵循公理 (i)-(iii)

练习 1.1.2 考虑到 (a,b,c) 形式的所有数组集合，其中这些数组元素都是实数，加法和标量乘法的定义如下：

$$(a,b,c) + (d,e,f) = (a+d, b+e, c+f)$$

$$a(a,b,c) = (aa, ab, ac).$$

写出这些零向量和 (a,b,c) 的负元，证明 $(a,b,1)$ 形式的向量不能形成一个向量空间

注意我们正在使用一个新的符号 $|V\rangle$ 来表示一个普通向量。这个符号叫做右矢 $V(\text{ket } v)$ 这个命名法是由狄拉克命名的，稍后详细讨论。我们第一步并没有故意使用符号 \vec{V} 来表示向量目的是让你摆脱把向量作为箭头的概念。然而直到你看到足够多的向量不是箭头时，你会放弃把向量和 $|V\rangle$ 联系起来，并准备抛弃这个概念。

您应该去核对这些箭头的集合是否符合你读过公理向量空间的条件。以下是一些你应该仔细研究的关键想法。向量空间由箭头组成，典型一个例子是 \vec{V} 和 \vec{V}' 。我们对加法规则非常熟悉：取第二个箭头的尾端，将其放在第一个箭头的始端，依此类推如图 (1.1) 所示。

标量乘法相当于通过一个因子 a 拉伸向量。这是一个实的向量空间因为复数拉伸是没有意义的。（如果 a 是负的，我们将其解释为改变箭头的方

向以及重新缩放按 $|a|$ 倍) 因为这些作用于箭头的操作运算产生了更多的箭头, 所以我们有了封闭性的限定。加法和标量乘法显然具有所有所需的结合律和分配律特征。零向量是长度为零的箭头, 然而一个向量的负元是在方向上相反的向量。

因此所有箭头的集合都满足向量空间的条件。但我们不能篡改. 例如, z 分量为正的所有箭头的集合不形成一个向量空间: 没有负元。

到目前为止, 请注意还没有提到幅度或方向。虽然箭头具有这些性质, 但向量空间的元素并不需要。除非我能给你举个例子, 否则这种说法毫无意义, 所以这里有两个例子。

考虑到所有 2×2 的矩阵集合。我们知道怎么把他们相加和进行标量乘法 (将矩阵所有元素乘以标量) 他们相应地遵循封闭性、结合性和分配性的要求。零矩阵里面所有元素都是零一个矩阵与其负元相加所有矩阵元都会消失。你必须承认这里有一个真正的向量空间由没有明显长度或方向相关的东西组成。当我们想强调矩阵 M 是向量空间的一个元素这一事实时, 我们可能想把它称为右矢 $|4\rangle$ 或者 $|4\rangle$ 。

第二个例子, 我们考虑被定义在区间 $0 \leq x \leq L$ 的所有函数 $f(x)$ 。我们简单将 a 的标量乘法定义为 $af(x)$, 将加法定义为逐点相加: 两个函数的和 f 和 g 在点 x 的值为 $f(x) + g(x)$ 。零函数处处为零, f 的相反函数为 $-f$

练习 1.1.3 在端点 $x = 0$ 和 $x = L$ 处消失的函数是否形成一个向量空间服从 $f(0)=f(L)$ 的周期函数怎么样呢? $f(0) = 4$ 的函数怎么样? 如果这些函数都不行, 请列出它们出错的地方。

接下来是向量 $|1\rangle, |2\rangle \dots |n\rangle$ 线性独立的概念首先考虑一个线性相关的形式

$$\sum_{i=1}^n a_i |i\rangle = |0\rangle \quad (1.1.1)$$

在不失一般性的情况下我们可以假设等式左侧不包含任意倍数的 $|0\rangle$ 。如果是那样, 那么我们可以向右边移项并与那里的 $|0\rangle$ 结合组合出更多的 $|0\rangle$ 。(我们使用的性质是任何数乘上 $|0\rangle$ 都等于 $|0\rangle$)

定义 3 向量集被认为是线性独立的当且仅当方程(1.1.1)中所有系数满足 $a_i = 0$, 如果向量集不是线性独立的, 那我们说它们是线性相关的。

方程(1.1.1)告诉我们在加入任何其他线性无关的集合, 另一方面, 如果向量集是线性相关的, 则会存在这样的关系, 并且它必须包含至少两个非零

系数。让我们令 $a_3 \neq 0$ 。然后我们可以写

$$|3\rangle = \sum_{i=1, \neq 3}^n \frac{-a_i}{a_3} |i\rangle \quad (1.1.2)$$

因此通过其他项表示 $|3\rangle$

作为一个具体的例子，考虑一个平面中的两个非平行向量 $|1\rangle$ 和 $|2\rangle$ 。这些形成了一个线性独立的集合。没有办法将一个向量写为其他向量的倍数或者无法同样将它们组合成零向量。另一方面，如果向量是平行的，我们可以清楚地将其中一个向量写为其他向量的倍数或者同样的将它们相互组合成 0。

注意到我们这里所说的是 0 而不是 $|0\rangle$ 。严格的来说这是不正确的因为一个向量的集合仅仅只能加到一个向量上去而不是一个数字。然而用 0 来表示零向量是非常普遍常见的。

假定我们引入第三个向量而且也在这个平面上。如果它与前两个向量中任何一个向量平行，那么我们已经有了一个线性相关集。所以让我们假设事实并非如此但即使是现在，它们三者也是线性相关的。这是因为我们可以将其中一个写为其他两个的线性组合为了找到线性组合向量，从 $|3\rangle$ 尾部沿着 $|1\rangle$ 方向画一条线，接着从 $|3\rangle$ 顶端画一条沿着 $|2\rangle$ 的反平行线。这些线将会相交因为假设 $|1\rangle$ 和 $|2\rangle$ 是线性独立的相交点 P 将会确定我们多长的 $|1\rangle$ 和 $|2\rangle$ ：我们从 $|3\rangle$ 的尾部出发到 P 点将会消耗适当倍数的 $|1\rangle$ 同理我们从 P 点到 $|3\rangle$ 的顶端将会消耗适当倍数的 $|2\rangle$ 。

练习 1.1.4 考虑来自 2×2 的实矩阵的向量空间的三个矩阵元

$$|1\rangle = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad |3\rangle = \begin{bmatrix} -2 & -1 \\ 0 & -2 \end{bmatrix}$$

它们是线性独立的吗？请详细的证明你的答案。（请注意，我们将这些矩阵称为向量并使用右矢来表示它们以强调它们作为向量空间矩阵元的作用）

练习 1.1.5 证明接下来的行向量是线性独立的 $(1, 1, 0)$, $(1, 0, 1)$ 和 $(3, 2, 1)$ 。
证明 $(1, 1, 0)$, $(1, 0, 1)$ $(0, 1, 1)$ 线性相关

定义 4 如果一个向量空间最大能容纳 n 个线性独立向量，那么它是 n 维度的。如果是实数域的将记作 $\mathbb{V}^n(R)$ ，如果是复数域将记作 $\mathbb{V}^n(C)$

鉴于前面的讨论，平面是二维的并且所有箭头集合构成了一个三维向量空间而不被该平面限制。那么 2×2 矩阵怎么样呢？它们形成了一个四维向量空间。这里有个证明。接下来向量是线性独立的：

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

因为它们中的任何三个向量的不可能线性组合来给出第四个向量，所以它们中的任意三个将在摸个地方变成零而第四个不会。所以这个空间至少是四维的。它会更大吗？不，因为任何任意的 2×2 矩阵都可以用它们来写：

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a|1\rangle + b|2\rangle + c|3\rangle + d|4\rangle$$

如果标量 a, b, c 是实数的，我们有一个实四维空间，如果它们复数的我们有一个复四维空间。

定理 1 任意在 n 维空间的向量 $|V\rangle$ 能够别写成 n 个线性独立 $|1\rangle \dots |n\rangle$ 向量的线性组合

证明如下：如果这里存在一个本来不可能存在的向量 $|V\rangle$ ，那么它可以加入已经给出的线性无关向量组并且形成一组 $n+1$ 的线性无关向量组，根据定义这在 n 维空间中是不可能的。

定义 5 n 维度空间中一组 n 个线性独立的向量组叫做基底

因此我们能在上述基础上写下

$$|V\rangle = \sum_{i=1}^n v_i |i\rangle \quad (1.1.3)$$

向量 $|i\rangle$ 形成一组基底

定义 6 向量在一个线性无关的基底上 ($|i\rangle$) 的的膨胀系数 v_i 被称为该向量在这个基上的的分量。

定理 2 在方程(1.1.3)的的膨胀系数是唯一的

假定膨胀系数不唯一，则我们有可以有第二个膨胀系数

$$|V\rangle = \sum_{i=1}^n v'_i |i\rangle \quad (1.1.4)$$

方程(1.1.3)与方程(1.1.4)相减（也就是说，第二个方程乘以-1 然后两个方程相加）我们得到

$$|0\rangle = \sum_{i=1}^n (v_i - v'_i) |i\rangle \quad (1.1.5)$$

这暗示着

$$v_i = v'_i \quad (1.1.6)$$

因为基底向量是线性独立的并且不存在线性关系，请注意给出的基底分量是唯一的，但是如果我们改变基底，那么相应地分量也会发生改变。我们把 $|V\rangle$ 看做为抽象的向量，它有自己的存在，并满足涉及其他向量的各种关系。当我们选择一个基时，假定根据在它们分量向量上具体的形式和向量之间的关系被分量所满足。想象一个平面三个箭头的例子，根据箭头的加法的法则 $\vec{A}, \vec{B}, \vec{C}$ 满足 $\vec{A} + \vec{B} = \vec{C}$ 。目前为止，我们还没有选择任何基也不需要一个基底来描述向量三角形法则。现在我们选择一组基并根据分量写出每个向量。他们之间的分量满足 $C_i = A_i + B_i, i = 1, 2$ 。如果我们选择不同的基底，分量的数值将会改变，但它们之间的关系（表示 C 与其他两个的和相等）在新的分量集之间仍然有效。

在非箭头向量的情况下，由于公理的存在，按照分量将它们相加的过程与初等情况相同。如果

$$|V\rangle = \sum_i v_i |i\rangle \quad (1.1.7)$$

$$|W\rangle = \sum_i w_i |i\rangle \quad (1.1.8)$$

$$|V\rangle + |W\rangle = \sum_i (v_i + w_i) |i\rangle \quad (1.1.9)$$

其中我们使用了公理来重整化术语。这里有一些结论：

向量相加等于他们之间分量相加

这里我们没有提到把一个向量的尾不放在其他向量的首部等等，因为通常向量没有头部或尾部。当然，如果我们正在处理箭头，我们可以使用向量尾部和尖端添加它们，也可以简单地在基础上添加它们的分量。同样，我们有：

$$a|V\rangle = a \sum_i v_i |i\rangle = \sum_i av_i |i\rangle \quad (1.1.10)$$

换句话说

要将标量乘以向量等于请将其所有分量乘以该标量

1.2 内积空间

矩阵和函数的例子一定让你相信，我们可以有一个向量空间，没有预先指定的元素长度或方向的定义。然而，我们可以构造出与箭头的长度和角度具有相同财产的量。第一步是定义点积的可感测模拟，在箭头的情况下，从点积开始

$$\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}| \cos \theta \quad (1.2.1)$$

我们可以将 \vec{A} 的长度读作 $\sqrt{|\vec{A}| \cdot |\vec{A}|}$ 以及两个向量的余弦读作 $\vec{A} \cdot \vec{B} / |\vec{A}| |\vec{B}|$ 。现在你可能会理所当然地反对：如果点积本身需要长度和角度的知识，你怎么能使用点积来定义长度和角度？答案是这样的。回想一下，点积在分量方面有第二个等效表达式：

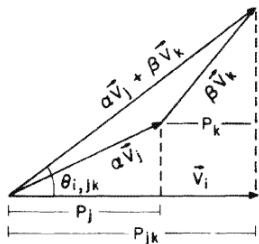


图 1.2: 点积服从内积公理 (3) 的几何证明。该公理要求投影服从 $P_k + P_j = P_{jk}$ 。

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z \quad (1.2.2)$$

我们的目标是在普遍情况下定义一个类似的公式，在这些情况下我们确实有着一些关于在一组基底下在有着分量的观念的。为此，我们回顾了上述点

积的主要特征：

1. $\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}$ (symmetry)
2. $\vec{A} \cdot \vec{A} \geq 0$ 0 iff $\vec{A} = 0$ (positive semidefiniteness)
3. $\vec{A} \cdot (b\vec{B} + c\vec{C}) = b\vec{A} \cdot \vec{B} + c\vec{A} \cdot \vec{C}$ (linearity)

点积的线性如图 (1.2) 所示。

我们想将任意两个向量 $|V\rangle$ 和 $|W\rangle$ 之间的内积或者向量积更一般化。我们用符号 $\langle V|W\rangle$ 来表示它。它又是取决于于两个向量的一个数字（通常是复数）。我们要求它遵守以下公理：

- $\langle V|W\rangle = \langle W|V\rangle^*$ (斜对称)
- $\langle V|V\rangle \geq 0$ iff $|V\rangle = |0\rangle$ (半正定)
- $\langle V|(a|W\rangle + b|Z\rangle) \equiv \langle V|aW + bZ\rangle = a\langle V|W\rangle + b\langle V|Z\rangle$ (右矢线性
linearity in ket)

定义 7 一个带有内积的向量空间叫做内积空间

注意，我们还没有给出一个明确的规则来实际评估标量乘积，我们只是要求我们提出的任何规则都必须具有这些性质。为了找到这样一条规则，让我们回顾熟悉我们自己的公理。第一条公理不同于对应的点积，它使内积对两个因子的顺序感到敏感，这两种选择导致复杂的共轭。在实向量空间中该公理规定了在两个向量交换下点积的对称性。就目前而言，我们注意到这一公理确保了 $\langle V|V\rangle$ 是实数的。

第二个公理说 $\langle V|V\rangle$ 不仅是实的而且也是半正定的，只有当向量本身为 0 时内积才为 0。如果我们将向量的长度定义为其与自身的内积的平方根（如点积中的），那么对于所有非零向量，这个量最好是实数和正数。

最后一个公理表示当线性叠加 $a|W\rangle + b|Z\rangle \equiv |aW + bZ\rangle$ 作为标量乘积中的第二个向量出现时内积的线性。我们已经讨论了它在箭头情况下的有效性（图 (1.2)）。

如果乘积中的第一个因子是线性叠加的，即 $|aW + bZ\rangle|V\rangle$ 是什么？这是由第一条公理决定的：

$$\begin{aligned}
\langle aW + bZ|V \rangle &= \langle V|aW + bZ \rangle^* \\
&= (a\langle V|W \rangle + b\langle V|Z \rangle) \\
&= a^*\langle V|W \rangle^* + b^*\langle V|Z \rangle^* \\
&= a^*\langle W|V \rangle + b^*\langle Z|V \rangle
\end{aligned} \tag{1.2.3}$$

其表示关于内积中的第一个因子的内积反线性。换言之，如果叠加发生在第二个因子中，一个线性叠加的向量的内积是相应内积的叠加，而如果叠加出现在第一个因子中则是所有共轭系数的叠加。这种在实向量空间中不熟悉的不对称性会一直存在，随着时间的推移，你会习惯的。

让我们继续使用内积的概念。尽管我们试图摆脱向量作为箭头的有限的概念，并寻求点积的相应推广，但我们仍然使用了一些相同的术语。

定义 8 如果两个向量的内积为 0。我们说他们是正交的或垂直的，

定义 9 我们将把 $\sqrt{\langle V|V \rangle} \equiv |V|$ 称为向量的范数或长度。归一化向量具有单位范数。

定义 10 一组正交单位范数的基向量被称为正交基

我们也经常将内积或标量积称为点积。

我们现在准备好根据投影分量给出内积的具体公式了。给定

$$\begin{aligned}
|V\rangle &= \sum_i v_i |i\rangle \\
|W\rangle &= \sum_j w_j |j\rangle
\end{aligned}$$

我们遵循内积服从的公理得到：

$$\langle V|W \rangle = \sum_i \sum_j v_i^* w_j \langle i|j \rangle \tag{1.2.4}$$

为了更进一步了解，我们必须知道 $\langle i|j \rangle$ 是什么，即基向量之间的内积。这取决于基向量的细节，并且我们所知道的是它们是线性无关的。箭头也存在这种情况。考虑一个二维问题，其中基向量是两个线性独立但非正交的向量。如果我们根据这组基来写所有向量，那么任意两个向量的点积也将是具有四项的二重和 (double sum) (由基向量之间的四个可能的点积确定) 以及向量分量也一样。然而，如果我们用一组规范正交基例如 \vec{i}, \vec{j} ，仅有像 $\langle i|i \rangle$

对角项能被保留下来并且们将得到熟悉的结果 $\vec{A} \cdot \vec{B} = \vec{A}_x \vec{B}_x + \vec{A}_y \vec{B}_y$ 仅取决于投影分量。

对于更一般的非箭头情况，我们引用定理3。

定理 3 (格拉姆—施密特) 给定一组线性独立的基，我们可以将基向量的线性组合以获得一组规范正交基。

证明推迟一会儿。让我们假设该程序已经开始运行并且当前的基底是规范正交的：

$$\langle i|j \rangle = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \equiv \delta_{ij}$$

其中 δ_{ij} 被称为克罗内克德尔塔符号。将其代入等式(1.2.4) 我们发现由于 Kronecker delta 二重和将会消失成为一个，从而得出

$$\langle V|W \rangle = \sum_i v_i^* w_i \quad (1.2.5)$$

这就是我们从现在开始使用的内积的形式。

你现在可以理解第一条公理了；第一个向量的投影分量的要不是复共轭的话， $\langle V|V \rangle$ 甚至不会是实数的，更不用说正数的了。但现在它被给定

$$\langle V|V \rangle = \sum_i |v_i|^2 \geq 0 \quad (1.2.6)$$

并且当且仅当为零向量时才为 0。这使得将 $\langle V|V \rangle$ 称为向量的长度或范数平方是合理的。

考虑等式(1.2.5)。由于向量 $|V\rangle$ 在给定的基底上是由其投影分量唯一的，我们可以在这个基底上将其写成列向量：

$$|V\rangle \rightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ 在这组基底上} \quad (1.2.7)$$

同样的

$$|W\rangle \rightarrow \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \text{ 在这组基底上} \quad (1.2.8)$$

$\langle V|W\rangle$ 的内积由列向量 $|V\rangle$ 的的转置共轭与列向量 $|W\rangle$ 的矩阵乘积表示：

$$\langle V|W\rangle = \begin{bmatrix} v^*, v_2^*, \dots, v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad (1.2.9)$$

1.3 对偶空间与狄拉克符号

这里有一个小技巧。内积是一个数字，我们试图从两个右矢 $|V\rangle$ 和 $|W\rangle$ 中得到，并且这两个右矢在基底上都用列向量表示。现在没有办法直接通过矩阵乘法从两列中得出一个数，但有一种方法可以通过一行乘以一列的矩阵乘法得出一个数字。我们从两列中产生一个数的技巧是将一个唯一的行向量与一列（其转置共轭）相关联，并与表示另一列的列形成其矩阵乘积。这具有的特征是，答案取决于我们要转换为行的两个向量中的哪一个，这两个选择 $\langle V|W\rangle$ 和 $\langle W|V\rangle$ 决定通过复数共轭相关的答案。

但人们也可以考虑以下的其他观点。列向量是一个抽象向量 $|V\rangle$ 或 ket 在一组基底的具体表现。我们也可以反向操作，从列向量变成到抽象的右矢。但是同样可以反向操作，并将把它每个行向量与一个抽象对象 $\langle W|$ ，称为左矢-W 相联系。现在我们可以随心所欲地命名左矢，但让我们做以下操作。每个右矢 $|V\rangle$ 与之相关联的是一个列向量。让我们取它的伴随，或者转置共轭，形成一个行向量。与此相关抽象的左矢将带有相同的标签，即，它将被称为 $\langle V|$ 。因此，存在两个向量空间，右矢的空间和左矢的对偶空间，每个左矢都有一个右矢，反之亦然（投影分量通过伴随运算相关联）。内积实际上只定义在左矢和右矢上，因此内积来源于是由两个截然不同但相关的向量空间的元素。这里存在一组基底向量 $|i\rangle$ 用来展开右矢，以及一个类似于展开右矢的基底向量 $\langle i|$ 用，基右矢 $|i\rangle$ 通过一组基底中除了第 i 行是 1 其余带有全为零的列向量被表示而基左矢 $\langle i|$ 被第 i 列是 1 的一组基底表示所有这些可以总结如下：

$$|V\rangle \leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \leftrightarrow [v_1^*, v_2^*, \dots, v_n^*] \leftrightarrow \langle V| \quad (1.3.1)$$

其中 \leftrightarrow 表示“在一组基里”

然而，第一种观点将标量乘积与一对列向量或右矢相关联（不参考另一个对偶空间），并接受内积中第一个和第二个向量之间的不对称性（哪一个是转置共轭？），这并没有错。如果你发现上面的讨论很激烈，你可以暂时忽略它。你唯一必须记住的是，在一般非箭头向量空间的情况下：

- 向量在某些规范正交基上可以被分配指定多少投影分量，就像箭头向量一样但这些分量可能很复数的
- 任何两个向量的内积通过方程(1.2.5)根据这些分量给出的。其内积服从所有公理。

1.3.1 向量的在规范正交基上的展开

假设我们希望在在一组规范正交基上展开一个向量 $|V\rangle$ 。为了找到展开所需的投影分量，我们按照以下步骤进行。我们两边都与假定展开的 $|j\rangle$ 做点积：（或是 $\langle j|$ 如果你是纯粹主义者）

$$|V\rangle = \sum_i v_i |i\rangle \quad (1.3.2)$$

$$\langle j|V\rangle = \sum_i v_i \langle j|i\rangle \quad (1.3.3)$$

$$= v_j \quad (1.3.4)$$

即为了找到向量的第 j 个分量，我们取第 j 个单位向量做点积，就像箭头一样。利用这个结果，我们可以写下

$$|V\rangle = \sum_i |i\rangle \langle i|V\rangle \quad (1.3.5)$$

让我们确保基向量的形式看起来正确。如果我们在方程(1.3.5)中设置 $|V\rangle = |j\rangle$ ，我们会得到正确的答案：第 j 个基向量的第 i 个分量为 δ_{ij} 。因此，例如，

表示第 4 个基向量的列在第 4 行为 1，其他位置均为 0。抽象关系式：

$$|V\rangle = \sum_i v_i |i\rangle \quad (1.3.6)$$

在这组基底上变成

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots v_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (1.3.7)$$

1.3.2 伴随运算

我们已经看到通过共轭转置操作即转置共轭，我们可以从表示一个右矢的列向量转换为表示相应左矢的行向量。现在让我们问一个问题：如果 $\langle V|$ 是对应于右矢 $|V\rangle$ 的右矢，那么对于标量 a 的右矢 $a|V\rangle$ 对应于什么左矢呢？通过转换到任意基底很容易找到答案：

$$a|V\rangle \rightarrow \begin{bmatrix} av_1 \\ av_2 \\ \vdots \\ av_n \end{bmatrix} \rightarrow [a^* v_1^*, a^* v_2^*, \dots, a^* v_n^*] \rightarrow \langle V|a^* \quad (1.3.8)$$

习惯上将 $a|V\rangle$ 写作 $|aV\rangle$ 且相应的右矢写作 $\langle aV|$ 。我们所发现：

$$\langle aV| = \langle V|a^* \quad (1.3.9)$$

由于左矢和右矢之间的关系是线性的，我们可以说如果我们有一个关于右矢的方程，例如：

$$a|V\rangle = b|W\rangle + c|Z\rangle + \cdots \quad (1.3.10)$$

这意味着对应的左矢之间也存在另一个方程：

$$\langle V|a^* = \langle W|b^* + \langle Z|c^* + \cdots \quad (1.3.11)$$

上述两个方程被称为彼此的共轭。就像涉及复数的任何方程都意味着通过对两边取复共轭得到另一个方程一样，一个关于（左矢）右矢的方程意味着

另一个关于（右矢）左矢的方程。如果你考虑一个基底，你会发现这只是因为如果两列相等，它们的转置共轭也相等。

以下是求共轭伴随的规则：

求取关于右矢（左矢）的线性方程的共轭伴随，需将每个右矢（左矢）替换为其左矢（右矢），并对所有系数进行复共轭。

我们可以按照以下方式扩展这个规则。假设我们有一个向量的展开式：

$$|V\rangle = \sum_{i=1} v_i |i\rangle \quad (1.3.12)$$

根据基向量。那么其共轭伴随为：

$$\langle V| = \sum_{i=1} \langle i| v_i^*$$

回顾 $v_i = \langle i|V\rangle$ 和 $v_i^* = \langle V|i\rangle$ 由此可得共轭伴随为：

$$|V\rangle = \sum_{i=1} |i\rangle \langle i|V\rangle \quad (1.3.13)$$

是

$$\langle V| = \sum_{i=1} \langle V|i\rangle \langle i| \quad (1.3.14)$$

对于涉及左矢、右矢和系数的方程取共轭伴随的方法是将所有因子的顺序颠倒，交换左矢和右矢并对所有系数进行复共轭。

Gram-Schmidt Theorem

现在让我们来开始讨论 Gram-Schmidt 过程，将一个线性独立的基转换为正交归一的基。现在通过一个简单的例子可以看出这个基本想法。想象一个平面上的二维箭头空间。我们取两个非平行的向量作为基。为了从这些向量中得到一个正交归一的基，我们进行以下操作：

- 将第一个向量按照其长度进行缩放，使其成为一个单位向量。这将成为第一个基向量。
- 从第二个向量中减去其在第一个向量上的投影，只留下垂直于第一个向量的部分。（由于假设向量是非平行的，所以垂直的部分将保留下来。）

- 将剩下的部分按照其长度进行缩放。现在我们得到了第二个基向量：它与第一个基向量正交且长度为 1。

这个简单的例子告诉了这个过程背后的全部故事，现在将在 Dirac 符号中以一般术语进行讨论。

让 $|I\rangle, |II\rangle$

$rangle$ 是一组局线性独立的基底。第一个向量的正交基为

$$|1\rangle = \frac{|I\rangle}{|I|} \text{quad where } |I| = \sqrt{\langle I | I \rangle}$$

很明显

$$\langle 1 | 1 \rangle = \frac{\langle I | I \rangle}{|I|^2} = 1$$

至于基底中的第二个向量，考虑到

$$|2'\rangle = |II\rangle - |1\rangle\langle 1 | II\rangle$$

这是 $|II\rangle$ 减去沿着第一单位向量指向的部分。(当你继续阅读时，想想箭头的例子。) 毫不奇怪它与后者正交。

$$\langle 1 | 2' \rangle = \langle 1 | II \rangle - \langle 1 | 1 \rangle \langle 1 | II \rangle = 0$$

我们现在将 $|2'\rangle$ 除以它的范数来得到 $|2\rangle$ 它将与第一个向量正交并被归一化为单位长度。最后，考虑

$$|3'\rangle = |III\rangle - |1\rangle\langle 1 | III \rangle - |2\rangle\langle 2 | III \rangle$$

这个向量将与 $|1\rangle$ 和 $|2\rangle$ 都正交。通过除以其范数，我们得到 $|3\rangle$ ，即正交基的第三个向量。对于其余基向量的生成没有什么新的内容。

我们在哪里使用了原始基向量的线性独立性呢？如果我们从一个线性相关的基开始，那么在某个时刻类似 $|2'\rangle$ 或 $|3\rangle$ 的向量将会消失，使整个过程停止。另一方面，线性独立性将确保我们永远不会发生这种情况，因为这相当于具有线性独立的向量的非平凡线性组合 (a nontrivial linear combination) 使得它们的和为零向量。(回顾一下 $|2'\rangle$ 或 $|3\rangle$ 的方程，并自己验证它们是旧基向量的线性组合。)

练习 1.3.1 从 $\hat{A} = 3\hat{i} + 4\hat{j}$ 和 $\hat{B} = 2\hat{i} - 6\hat{j}$ 开始形成一组二维正交基。你能从这两个向量开始给出另一个规范正交基吗？如果可以这样，就给出一个。

练习 1.3.2 展示如何从基底开始

$$|I\rangle = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \quad |II\rangle = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad |III\rangle = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$$

到规范正交基

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \quad |3\rangle = \begin{bmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

当我们初学维度的概念时，我们通常将其与垂直方向向量的数量相关联。在本章中，我们根据线性独立向量的最大数量来定义维度。以下定理将这两个定义联系起来。

定理 4 一个空间的维度等于 n_{\perp} ，即其中互相正交的向量的最大数量。

为了证明这个，首先要注意任何相互正交的集合也是线性无关的。假设我们有一组正交向量的线性组合等于零。通过将其中基底任意一个向量的两边做点积，并利用正交性，我们可以证明乘以该向量的系数必须为零。这显然可以对所有系数进行，从而表明线性组合是平凡的。

现在， n_{\perp} 只能等于、大于或小于空间的维度 n 。Gram-Schmidt 过程通过显式构造消除了最后一种情况，而垂直向量的线性独立性排除了倒数第二种选择。

施瓦茨不等式和三角不等式

有两个强大的定理适用于满足我们公理的任何内积空间：

定理 5 施瓦茨不等式

$$|\langle V|W\rangle| \leq |V||W| \quad (1.3.15)$$

定理 6 三角不等式

$$|V + W| \leq |V| + |W| \quad (1.3.16)$$

将提供第一个定理的证明，这样你就能熟悉使用 bra 和 ket 符号。第二个定理将作为一个练习留给你。

在证明任何定理之前，需要注意的是这些结果对于箭头（向量）显然是成立的：Schwarz 不等式表明两个向量的点积不会超过它们长度的乘积，而三角不等式表明和的长度不会超过各部分长度的和。这个例子说明了将抽象向量视为箭头并猜测它们可能与箭头共享的属性的优点。当然，证明必须仅依赖于公理。

为了证明 Schwarz 不等式，考虑公理 $\langle Z|Z \rangle \geq 0$ 适用于

$$|Z\rangle = |V\rangle - \frac{\langle W|V\rangle}{|W|^2}|W\rangle \quad (1.3.17)$$

我们得到

$$\begin{aligned} \langle Z|Z \rangle &= \langle V - \frac{\langle W|V\rangle}{|W|^2}W | V - \frac{\langle W|V\rangle}{|W|^2}W \rangle \\ &= \langle V|V \rangle - \frac{\langle W|V\rangle\langle V|W \rangle}{|W|^2} - \frac{\langle W|V\rangle^* \langle W|V \rangle}{|W|^2} \\ &\quad + \frac{\langle W|V\rangle^* \langle W|V \rangle \langle W|W \rangle}{|W|^4} \\ &\geq 0 \end{aligned} \quad (1.3.18)$$

我们在这里利用了内积相对于 bra 的反线性性质。用

$$\langle W|V \rangle^* = \langle V|W \rangle$$

我们发现

$$\langle V|V \rangle \geq \frac{\langle W|V \rangle \langle V|W \rangle}{|W|^2} \quad (1.3.19)$$

乘以 $|W|^2$ 并取平方根得到结果。

练习 1.3.3 这个等式何时成立？这是否与您在处理箭头时的经验相符？

练习 1.3.4 以 $|V + W|^2$ 开始证明三角不等式。你必须使用 $\text{Re}\langle V|W \rangle \leq |\langle V|W \rangle|$ 和 Schwarz 不等式。证明不等式最终变成等式当且仅当只有在 $|V\rangle = a|W\rangle$ 其中 a 是一个正实数标量的情况下才成立。

1.4 子空间

定义 11 给定一个向量空间 \mathbb{V} ，它的元素的子集在它们中形成一个向量空间[‡]，则称之为子空间。我们将用 $\mathbb{V}_i^{n_i}$ 来表示维度为 n_i 的特定子空间 i 。

[‡]在子空间中，向量的加法和标量乘法的定义与在 \mathbb{V} 中相同。

例 1.4.1 在三维实数向量空间 $\mathbb{V}^3(R)$ 中，以下是一些子空间的例子：(a) 所有沿 x 轴的向量，即空间 \mathbb{V}_x^1 ；(b) 所有沿 y 轴的向量，即空间 \mathbb{V}_y^1 ；(c) 所有在 x - y 平面上的向量，即空间 \mathbb{V}_{xy}^2 。注意，所有子空间都包含零向量，并且每个向量都有其相反向量以满足向量空间的公理。因此，仅由所有沿着正 x 轴的向量组成的集合并不构成一个向量空间。

定义 12 给定两个子空间 $\mathbb{V}_i^{n_i}$ 和 $\mathbb{V}_j^{m_j}$ ，我们将它们的和 $\mathbb{V}_i^{n_i} \oplus \mathbb{V}_j^{m_j} = \mathbb{V}_k^{m_k}$ 定义为一个集合将包含 (1) $\mathbb{V}_i^{n_i}$ 的所有元素，(2) $\mathbb{V}_j^{m_j}$ 的所有元素，(3) 上述元素的所有可能的线性组合。但对于元素 (3)，将会丢失封闭性。

例 1.4.2 如果，例如 $\mathbb{V}_1^x \oplus \mathbb{V}_1^y$ 只包含沿 x 轴和 y 轴的向量，那么通过添加这两个向量元素，每个方向一个，我们可以生成一个既不沿 x 轴也不沿 y 轴的向量。然而，如果我们还包括所有的线性组合，我们将得到正确的结果，即 $\mathbb{V}_1^x \oplus \mathbb{V}_1^y = \mathbb{V}_2^{xy}$

例 1.4.3 * 在一个 \mathbb{V}_n 空间，我们需要证明满足对任意非零向量 $|V\rangle \neq |0\rangle$ 正交的所有向量集合 $\{|V_\perp^1\rangle, |V_\perp^2\rangle, \dots\}$ ，构成一个子空间 \mathbb{V}^{n-1} 。

例 1.4.4 假设 $\mathbb{V}_1^{n_1}$ 和 $\mathbb{V}_2^{n_2}$ 是两个子空间，使得 \mathbb{V}_1 的任意元素与 \mathbb{V}_2 的任意元素正交。证明 $\mathbb{V}_1 \oplus \mathbb{V}_2$ 的维度为 $n_1 + n_2$ 。(提示：使用定理 4。)

1.5 线性算符

一个算符 Ω 是将给定向量 $|V\rangle$ 转换为另一个向量 $|V'\rangle$ 的指令。算符的作用可以被如下表示：

$$\Omega|V\rangle = |V'\rangle \quad (1.5.1)$$

一个说法是算符 Ω 将 $\text{ket } |V\rangle$ 转换为 $\text{ket } |v'\rangle$ 。我们将我们注意力自始至终集中在算符 Ω 上致使使我们不会脱离向量空间，也就是说，如果 $|V\rangle$ 是向量空间 \mathbb{V} 的元素，那么 $|V'\rangle = \Omega|V\rangle$ 也是

算符也能作用在 bra 上

$$\langle V'| = \langle V''| \quad (1.5.2)$$

我们将只关注线性算符也就是说即遵循以下规则的算符：

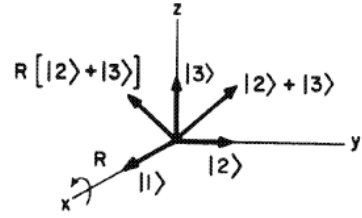
$$\Omega \alpha |V_i\rangle = \alpha \Omega |V_i\rangle \quad (1.5.3a)$$

$$\Omega\{\alpha|V_i\rangle + \beta|V_j\rangle\} = \alpha\Omega|V_i\rangle + \beta\Omega|V_j\rangle \quad (1.5.3b)$$

$$\langle V_i|\alpha\Omega = \langle V_i|\Omega\alpha \quad (1.5.4a)$$

$$(\langle V_i|\alpha + \langle V_j|\beta)\Omega = \alpha\langle V_i|\Omega + \beta\langle V_j|\Omega \quad (1.5.4b)$$

图 1.3: 算符 $R(\frac{1}{2}\pi i)$ 的作用。注意到 $R[|2\rangle + |3\rangle] = R|2\rangle + R|3\rangle$ 这符合线性算符的预期性质。(如果不会引起混淆, 我们通常将 $R(\frac{1}{2}\pi i)$ 简写为 R 。)



例 1.5.1 最简单的运算符是单位运算符 I , 它表示以下指令:

: $I \rightarrow$ 将向量保持不变!

所以

$$I|V\rangle = |V\rangle \quad \text{for all kets } |V\rangle \quad (1.5.5)$$

且

$$\langle V|I = \langle V| \quad \text{for all bras } \langle V| \quad (1.5.6)$$

我们接下来将转向一个更有趣的作用于 $\mathbb{V}^3(\mathbf{R})$ 的运算符:

$R(\frac{1}{2}\pi i) \rightarrow$ 将向量绕单位向量 \mathbf{i} 旋转 $\frac{1}{2}\pi$ 。

更一般地, $R(\theta)$ 代表绕平行于单位向量 $\hat{\theta} = \frac{\theta}{|\theta|}$ 的轴旋转角度 $\theta = |\theta|$ 的旋转。让我们考虑该运算符在三个单位向量 \mathbf{i} , \mathbf{j} 和 \mathbf{k} 的作用, 我们用 $|1\rangle$, $|2\rangle$ 和 $|3\rangle$ 的符号来表示 (见图 (1.3))。从图中可以明确看出

$$R(\frac{1}{2}\pi \mathbf{i})|1\rangle = |1\rangle \quad (1.5.7a)$$

$$R(\frac{1}{2}\pi \mathbf{i})|3\rangle = -|2\rangle \quad (1.5.7b)$$

$$R(\frac{1}{2}\pi \mathbf{i})|2\rangle = |3\rangle \quad (1.5.7c)$$

显然 $R(\frac{1}{2}\pi i)$ 是线性的。例如，从相同的图中可以明确看出 $R[|2\rangle + |3\rangle] = R|2\rangle + R|3\rangle$ 。

线性算子的一个好处是，一旦它们在基向量上的作用已知，它们在空间中任意向量上的作用就确定了。如果对于在 \mathbb{V}^n 的一组基向量有 $|1\rangle, |2\rangle, \dots, |n\rangle$

$$\Omega|i\rangle = |i'\rangle$$

那么对于任意向量 $|V\rangle = \sum_i v_i$ 有

$$\Omega|V\rangle = \sum_i \Omega v_i |i\rangle = \sum_i v_i \Omega|i\rangle = \sum_i v_i |i'\rangle \quad (1.5.8)$$

这就是例子 $\Omega = R(\frac{1}{2}\pi i)$ 中的情况。如果

$$|V\rangle = v_1|1\rangle + v_2|2\rangle + v_3|3\rangle$$

是任何向量，那么

$$R|V\rangle = v_1 R|1\rangle + v_2 R|2\rangle + v_3 R|3\rangle = v_1|1\rangle + v_2|3\rangle - v_3|2\rangle$$

两个运算符的乘积表示这样的指令：按顺序执行这两个运算符所对应的指令。

$$\Lambda\Omega|V\rangle = \Lambda(\Omega|V\rangle) = \Lambda|\Omega V\rangle \quad (1.5.9)$$

其中 $|\Lambda V\rangle$ 表示 Λ 作用于 V 后得到的 ket。在一个乘积中运算符的顺序非常重要：一般来说，

$$\Omega\Lambda - \Lambda\Omega \equiv [\Omega, \Lambda]$$

称为 Ω 和 Λ 的对易子 (commutator 数学中交换子) 不为 0。例如 $R(\frac{1}{2}\pi i)$ 和 $R(\frac{1}{2}\pi j)$ 不可交换。也就是说，他们的对易算符是非零。

Ω 的逆记作 Ω^{-1} ，满足[‡]

$$\Omega\Omega^{-1} = \Omega^{-1}\Omega = I \quad (1.5.10)$$

并非每个运算符都具有逆运算。存在逆运算的条件在附录 A.1. 中给出。算符 $R(\frac{1}{2}\pi i)$ 具有逆运算，即 $R(-\frac{1}{2}\pi i)$ 。运算符的积的逆运算是逆序进行的逆运算的乘积。

$$(\Omega\Lambda)^{-1} = \Lambda^{-1}\Omega^{-1} \quad (1.5.11)$$

只有在这种情况下我们才有：

$$(\Omega\Lambda)(\Omega\Lambda)^{-1} = (\Omega\Lambda)(\Lambda^{-1}\Omega^{-1}) = \Omega\Lambda\Lambda^{-1}\Omega^{-1} = \Omega\Omega^{-1} = I \quad (1.5.12)$$

[‡]在有限维 $\mathbb{V}^n(C)$ 中， $\Omega^{-1}\Omega = I \Leftrightarrow \Omega\Omega^{-1} = I$ 。证明这个结论。这个结论会在定理 A.1.1.，附录 A.1 中被介绍

1.6 线性算子的矩阵元

我们现在习惯于将抽象向量在基底中用一个 n 元组的数表示，这个 n 元组被称为向量的分量，通过这些分量，我们可以进行所有向量运算。现在我们将看到线性算符以同样的方式可以通过一组 n^2 个数字在基底中表示，写成一个 $n \times n$ 的矩阵，并称之为该基底下的矩阵元素。尽管矩阵元素和向量分量一样，是基底相关的，但它们有助于计算所有基底无关的量，通过使抽象算符更加具体化。

我们的起点是之前提到的观察结果，即线性算符的作用完全由其对基向量的作用确定。如果基向量发生变化：

$$\Omega|i\rangle = |i'\rangle$$

(其中 $|i'\rangle$ 是已知的)，那么该空间中的任何向量都会发生可轻易计算的变化：

$$\Omega|V\rangle = \Omega \sum_i v_i |i\rangle = \sum_i v_i \Omega|i\rangle = \sum_i v_i |i'\rangle$$

当我们说 $|i'\rangle$ 是已知，我们的意思指的是它在原始基底下的分量已知。

$$\langle j|i'\rangle = \langle j|\Omega|i\rangle \equiv \Omega_{ji} \quad (1.6.1)$$

这 n^2 个数值 Ω_{ij} 是 Ω 在这个基底下的矩阵元素。如果

$$\Omega|V\rangle = |V'\rangle$$

那么转换后的 $\text{ket}|V'\rangle$ 的分量可以用 Ω_{ij} 和 $|V\rangle$ 的分量来表示：

$$\begin{aligned} v'_i &= \langle i|V'\rangle = \langle i|\Omega|V\rangle = \langle i|\Omega \left(\sum_j v_j |j\rangle \right) \\ &= \sum_j v_j \langle i|\Omega|j\rangle \\ &= \sum_j \Omega_{ij} v_j \end{aligned} \quad (1.6.2)$$

方程(1.6.2)可以用矩阵形式描述为：

$$\begin{bmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{bmatrix} = \begin{bmatrix} \langle 1|\Omega|1\rangle & \langle 1|\Omega|2\rangle & \cdots & \langle 1|\Omega|n\rangle \\ \langle 2|\Omega|1\rangle & & & \vdots \\ \vdots & & & \vdots \\ \langle n|\Omega|1\rangle & \cdots & \langle n|\Omega|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad (1.6.3)$$

一个助记符：第一列的元素简单地就是给定基底中第一个变换后基向量 $|1'\rangle = \Omega|1\rangle$ 的分量。同样，第 j 列的元素表示第 j 个基向量在 Ω 作用后的象。

让自己相信，对于与任何 $\langle V'|$ 相对应的行向量作用在其左边的同一个矩阵 Ω 会给出与 $\langle V''|$ 相对应的行向量 $\langle V''| = \langle V'|\Omega$ 。

例 1.6.1 结合我们的助记符和算符 $R(\frac{1}{2}\pi i)$ 对基向量的作用效果如下：

$$\begin{aligned} R(\frac{1}{2}\pi i)|1\rangle &= |1\rangle \\ R(\frac{1}{2}\pi i)|2\rangle &= |3\rangle \\ R(\frac{1}{2}\pi i)|3\rangle &= -|2\rangle \end{aligned}$$

我们可以写出在基向量 $|1\rangle, |2\rangle, |3\rangle$ 下表示它的矩阵：

$$R(\frac{1}{2}\pi i) \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (1.6.4)$$

例如，第三列中的-1 告诉我们 R 将 $|3\rangle$ 旋转为 $-|2\rangle$ 。也可以完全忽略助记符直接使用定义 $R_{ij} = \langle i|R|j\rangle$ 来计算矩阵。

练习 1.6.1 一个算符 Ω 由以下矩阵给出：

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

它的作用是什么？

现在让我们考虑一些特定的算符，并看看它们在矩阵形式中的表示。

(1) 单位算符 I

$$I_{ij} = \langle i|I|j\rangle = \langle i|j\rangle = \delta_{ij} \quad (1.6.5)$$

因此单位算符 I 由一个沿对角线是 1 的对角矩阵表示。你应该验证一下我们的助记符是否得出了相同的结果。

(2) 投影算符 (Projection Operators)。首先让我们熟悉投影算符。考虑将任意 $\text{ket}|V\rangle$ 在一个基底下展开的表达式：

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

从 $|i\rangle\langle i|$ 方面说来, 它们是线性算子, 根据定义, 它们对 $|V\rangle$ 起作用并给出 $|ket i\langle i|V\rangle$ 的方式, 我们可以将上述表达式写为:

$$|V\rangle = \left(\sum_{i=1}^n |i\rangle\langle i| \right) |V\rangle \quad (1.6.6)$$

由于方程(1.6.6)对所有 $|V\rangle$ 成立, 括号中的对象必须等同单位算符。

$$I = \sum_{i=1}^n |i\rangle\langle i| = \sum_{i=1}^n \mathbb{P}_i \quad (1.6.7)$$

对象 $\mathbb{P} = |i\rangle\langle i|$ 被称为 $\text{ket}|i\rangle$ 的投影算符。方程(1.6.7)被称为完备关系, 它将单位算符表示为投影算符的求和形式, 并且对我们来说非常宝贵。(如果你认为花在单位算符上的时间似乎没有任何作用, 是浪费时间, 那么请稍等片刻, 你会看到它的价值。)

考虑到

$$\mathbb{P}_i |V\rangle = |i\rangle\langle i|V\rangle = |i\rangle v_i \quad (1.6.8)$$

很明显 \mathbb{P}_i 是线性的。注意无论 $|V\rangle$ 是什么, $\mathbb{P}|V\rangle$ 都是与 $|i\rangle$ 成比例的, 其中系数 v_i 是 $|V\rangle$ 沿着 $|i\rangle$ 方向的分量。由于 \mathbb{P}_i 可以投影出任何态 $|V\rangle$ 沿着 $|i\rangle$ 方向的分量, 它被称为投影算符。完备关系, 方程(1.6.7), 表示一个向量沿着所有 n 个方向的投影之和等于向量本身。投影算符也可以以相同的方式作用于 bra:

$$\langle V|\mathbb{P}_i = \langle V|i\rangle\langle i| = v_i^* \langle i| \quad (1.6.9)$$

对应于基向量的投影算符遵守以下关系:

$$\mathbb{P}_i \mathbb{P}_j = |i\rangle\langle i|j\rangle\langle j| = \delta_{ij} \mathbb{P}_j \quad (1.6.10)$$

这个方程告诉我们以下几点: (1) 一旦 \mathbb{P}_i 投影出 $|V\rangle$ 沿着 $|i\rangle$ 的部分, 进一步应用 \mathbb{P}_i 不会产生任何变化; (2) 连续应用 $\mathbb{P}_j (j \neq i)$ 将得到零, 因为完全沿着 $|i\rangle$ 方向的矢量无法在垂直方向 $|j\rangle$ 上有投影。

光学中的以下示例可能会对讨论有所启发。考虑沿 z 轴行进且在与 y 轴成角度 θ 的 $x-y$ 平面上偏振的光束 (见图 (1.4))。如果放置一个只允许沿 y 轴偏振的偏振片 P_y , 那么投影 $E \cos \theta$ 将沿 y 轴的被传输。在其路径上再放置一个偏振片 P_y 对光束没有进一步的影响。我们可以将偏振片的作用等同于对电场矢量 E 起作用的投影算符 \mathbb{V}_y 。如果 P_y 后面跟着一个偏振片 P_x , 光束将完全被阻挡。因此, 偏振片遵循投影算符的预期方程 $P_i P_j = \delta_{ij} P_j$ 。

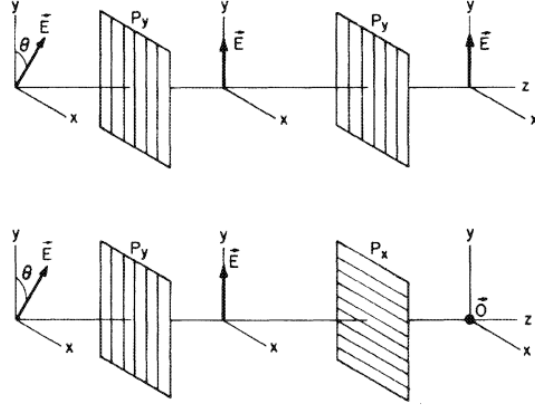


图 1.4: P_x 和 P_y 是放置在沿着 z 轴行进的光束路径上的偏振片。偏振片对电场 E 的作用遵循投影算符的组合规律: $P_i P_j = \delta_{ij} P_j$ 。

接下来, 我们将转向 \mathbb{V}_i 的矩阵元素。有两种方法。第一种方法有点间接, 但可以让我们对对象 $|i\rangle\langle i|$ 是什么有所感觉。我们知道

$$|i\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}$$

并且

$$\langle i| \leftrightarrow [0, 0, \dots, 1, 0, 0, \dots, 0]$$

因此

$$|i\rangle\langle i| \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [0, 0, \dots, 1, 0, \dots, 0] \equiv \begin{bmatrix} 0 & & \cdots & & 0 \\ & \ddots & & & \\ & & 0 & & \\ \vdots & & & 1 & \\ 0 & & & & 0 & \ddots & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{bmatrix} \quad (1.6.11)$$

根据矩阵乘法规则, 然而 $\langle V|V'\rangle = (1 \times n) \times (n \times 1) = 1 \times 1$ 是一个标量, 而 $|V\rangle\langle V'| = (n \times 1) \times (1 \times n) = (n \times n)$ 是一个算符。内积 $\langle V|V'\rangle$ 表示一个 bra 和一个 ket 相互匹配, 而 $|V\rangle\langle V'|$, 有时称为外积, 是两个因子朝着另一边寻找一个 bra 或 ket 进行点积

对于矩阵元素, 更直接的方法给出了

$$(\mathbb{P}_i)_{kl} = \langle k | i \rangle \langle i | l \rangle = \delta_{ki} \delta_{il} = \delta_{kl} \delta_{li} \quad (1.6.12)$$

这当然与方程(1.6.11)是相同的。同样的结果也可以从助记法推导出来。每个投影算子只有一个非零的矩阵元素, 即对角线上的第 i 个元素为 1。完备关系, 方程(1.6.7), 表明当所有的 \mathbb{P}_i 被加起来时, 对角线填满以得到单位矩阵。如果我们只对一些投影算子进行求和, 我们得到的算子将把给定的向量投影到由相应的基向量张成的子空间中。

算符乘积对应的矩阵

接下来考虑表示运算符乘积的矩阵。这些矩阵与表示单个运算符的矩阵相关联, 通过应用方程(1.6.7):

$$\begin{aligned} (\Omega\Lambda)_{ij} &= \langle i | \Omega\Lambda | j \rangle = \langle i | \Omega I \Lambda | j \rangle \\ &= \sum_k \langle i | \Omega | k \rangle \langle k | \Lambda | j \rangle = \sum_k \Omega_{ik} \Lambda_{kj} \end{aligned} \quad (1.6.13)$$

因此表示运算符乘积的矩阵就是表示因子的矩阵的乘积。

运算符的伴随

回顾一下给定一个 ket $\alpha|V\rangle \equiv |\alpha V\rangle$ 相应的 bra 是

$$\langle \alpha V | = \langle V | \alpha^*$$

同样的方式, 给定一个 ket

$$\Omega|V\rangle = |\Omega V\rangle$$

相应地 bra 是

$$\langle \Omega V | = \langle V | \Omega^\dagger \quad (1.6.14)$$

这定义了算符 Ω^\dagger 。我们可以用文字来表达这个方程: 如果 Ω 将一个 ket $|V\rangle$ 转换为 $|\Omega V\rangle$, 那么 Ω^\dagger 将 bra $\langle V|$ 转换为 $\langle \Omega V|$ 。就像 α 和 α^* 、 $|V\rangle$ 和 $\langle V|$

相关但不同， Ω 和 Ω^\dagger 也是相关但不同的对象。 Ω 和 Ω^\dagger 之间的关系，称为 Ω 的伴随或 “omega dagger,” 在一组基底中最容易理解：

$$\begin{aligned}(\Omega^\dagger)_{ij} &= \langle i|\Omega^\dagger|j\rangle = \langle \Omega i|j\rangle \\ &= \langle j|\Omega i\rangle^* = \langle j|\Omega|i\rangle^*\end{aligned}$$

因此

$$\Omega_{ij}^\dagger = \Omega_{ji}^* \quad (1.6.15)$$

换句话说，表示的 Ω^\dagger 矩阵是表示 Ω 的矩阵的转置共轭。（回想一下，表示 $\langle V|$ 的行向量是表示 $|V\rangle$ 的列向量的转置共轭）。在给定的基底中，伴随操作等同于取转置共轭。一个乘积的伴随是逆序伴随的乘积：

$$(\Omega\Lambda)^\dagger = \Lambda^\dagger\Omega^\dagger \quad (1.6.16)$$

为了证明这一点，我们考虑 $\langle \Omega\Lambda V|$ 。首先我们将 $\Omega\Lambda$ 视为一个运算符然后我们得到：

$$\langle \Omega\Lambda V| = \langle (\Omega\Lambda)V| = \langle V|(\Omega\Lambda)^\dagger$$

接下来，我们将 (ΛV) 视为另一个向量，并写作：

$$\langle \Omega\Lambda V| = \langle \Omega(\Lambda V)| = \langle \Lambda V|\Omega^\dagger$$

接下来，我们将 Λ 提出来，将 Ω^\dagger 推到更远的位置：

$$\langle \Lambda V|\Omega^\dagger = \langle V|\Lambda^\dagger\Omega^\dagger$$

将这个结果与前面几行得到的结果进行比较，我们得到了所期望的结果。

现在考虑一个由 ket、标量和运算符组成的方程，如下所示：

$$\alpha_1|V_1\rangle = \alpha_2|V_2\rangle + \alpha_3|V_3\rangle\langle V_4|V_5\rangle + \alpha_4\Omega\Lambda|V_6\rangle \quad (1.6.17a)$$

它的伴随是什么？根据我们之前的规则，它是：

$$\langle V_1|\alpha_1^* = \langle V_2|\alpha_2^* + \langle V_5|V_4\rangle\langle V_3|\alpha_3^* + \langle \Omega\Lambda V_6|\alpha_4^*$$

在最后一项中，我们可以将 $\langle \Omega\Lambda V_6|$ 替换为：

$$\langle V_6|(\Omega\Lambda)^\dagger = \langle V_6|\Lambda^\dagger\Omega^\dagger$$

因此，最终我们得到方程 (1.6.17a) 的伴随形式：

$$\langle V_1|\alpha_1^* = \langle V_2|\alpha_2^* + \langle V_5|V_4\rangle\langle V_3|\alpha_3^* + \langle V_6|\Lambda^\dagger\Omega^\dagger\alpha_4^* \quad (1.6.17b)$$

对于我们可能遇到的最一般方程，求其伴随的最终规则是：

当遇到运算符、bra、ket 和明确的数值系数的乘积时，颠倒所有因子的顺序，并进行以下替换： $\Omega \leftrightarrow \Omega^\dagger, |\rangle \leftrightarrow \langle|, \alpha \leftrightarrow \alpha^*$ 。

(当然，实际上没有必要颠倒标量 a 的位置，只是为了统一性考虑。)

厄米 (Hermitian)、反厄米 (Anti-Hermitian) 和么正 (Unitary) 运算符

现在我们将注意力转向在量子力学中起重要作用的某些特殊类别的运算符。

定义 13 如果 $\Omega^\dagger = \Omega$ ，则运算符 Ω 是厄米 (Hermitian) 的。

定义 14 如果 $\Omega^\dagger = -\Omega$ ，则运算符 Ω 是反厄米 (Anti-Hermitian) 的。

伴随运算符对应于数的复共轭。厄米和反厄米运算符类似于纯实数和纯虚数。就像每个数可以分解为纯实部分和纯虚部分的和一样

$$\alpha = \frac{\alpha + \alpha^*}{2} + \frac{\alpha - \alpha^*}{2}$$

我们可以将每个运算符分解为其厄米和反厄米部分：

$$\Omega = \frac{\Omega + \Omega^\dagger}{2} + \frac{\Omega - \Omega^\dagger}{2} \quad (1.6.18)$$

练习 1.6.2 * 假设 Ω 和 Λ 都是厄米运算符，那么关于以下表达式：(1) $\Omega\Lambda$ (2) $\Omega\Lambda + \Lambda\Omega$ (3) $[\Omega, \Lambda]$ (4) $i[\Omega, \Lambda]$

定义 15 如果运算符 U 是么正 (Unitary) 的，那么

$$UU^\dagger = I \quad (1.6.19)$$

该方程告诉我们 U 和 U^\dagger 是彼此的逆运算符。因此，根据方程(1.5.12)我们有：

$$U^\dagger U = I \quad (1.6.20)$$

根据运算符和数字之间的类比，么正运算符类似于单位模数的复数，即 $u = e^{i\theta}$ 。就像 $u * u = 1$ 一样， $U^\dagger U = I$ 。

练习 1.6.3 * 证明么正运算符的乘积仍然是么正的。

定理 7 么正运算符保持内积在它们作用的向量上。

证明 设

$$|V'_1\rangle = U|V_1\rangle$$

且

$$|V'_2\rangle = U|V_2\rangle$$

然后

$$\begin{aligned}\langle V'_2|V'_1\rangle &= \langle UV_2|UV_1\rangle \\ &= \langle V_2|U^\dagger U|V_1\rangle = \langle V_2|V_1\rangle\end{aligned}\quad (1.6.21)$$

(Q.E.D)

么正算子是从 $\mathbb{V}_3(C)$ 到 $\mathbb{V}_n(C)$ 的旋转算子的推广，就像三维空间中的旋转算子一样，它们保持向量的长度和点积。实际上，在实数向量空间中，么正条件变为 $U^{-1} = U^T$ (其中 T 表示转置)，这定义了一个正交或旋转矩阵 [$R(\frac{1}{2})\pi i$ 是一个例子]。

定理 8 如果将 $N \times n$ 么正矩阵的列视为 n 个向量的分量，则这些向量是正交的。同样地，行也可以被解释为 n 个正交向量的分量。

证明 1. 根据我们的助记符，表示 U 的矩阵的第 j 列是 U 作用于第 j 个基向量后的象。由于 U 保持内积，旋转后的向量集也是正交的。接下来考虑行。我们现在使用 U^T 也是一个旋转的事实 (否则它怎么能中和 U 得到 $U^\dagger U = I$?)。由于 U 的行是 U^T 的列 (除了整体的复共轭，这不影响正交性的问题)，我们已经对么正矩阵的列得到的结果告诉我们 U 的行是正交的。

证明 2. 因为 $U^\dagger U = I$,

$$\begin{aligned}\delta_{ij} &= \langle i|I|j\rangle = \langle i|U^\dagger U|j\rangle \\ &= \sum_k \langle i|U^\dagger|k\rangle \langle k|U|j\rangle \\ &= \sum_k U_{ik}^\dagger U_{kj} = \sum_k U_{ki}^* U_{kj}\end{aligned}\quad (1.6.22)$$

这证明了列的情况下的定理。如果我们从方程 $UU^\dagger = I$ 开始，就可以得到行的类似结果。QED。

需要注意的是, $U^\dagger U = I$ 和 $UU^\dagger = I$ 这两个条件并不是独立的。

练习 1.6.4 * 假设您已经了解以下内容: (1) 行列式的概念, (2) $\det \Omega^T = \det \Omega$ (其中 T 表示转置), (3) 矩阵乘积的行列式等于行列式的乘积。如果您对这些性质不熟悉, 请通过二维情况下的例子进行验证,

$$\Omega = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

其中 $\det \Omega = (\alpha\delta - \beta\gamma)$ 。证明么正矩阵的行列式是单位模的复数。

练习 1.6.5 * 通过检查其矩阵, 验证 $R(\frac{1}{2}\pi i)$ 是否是么正的 (正交的)。

练习 1.6.6 验证以下矩阵是否为么正矩阵:

$$\frac{1}{2^{1/2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \quad \frac{1}{2} \begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$$

验证每个矩阵的行列式是否具有 $e^{i\theta}$ 的形式。上面的矩阵中有厄密矩阵吗?

1.7 主动和被动转换

假设我们对空间中的所有向量 $|V\rangle$ 进行一个么正变换。

$$|V\rangle = U|V\rangle \quad (1.7.1)$$

在这个变换下, 任意算符 Ω 的矩阵元素会被修改如下:

$$\langle V'|\Omega|V\rangle \rightarrow \langle UV'|\Omega|UV\rangle = \langle V'|U^\dagger\Omega U|V\rangle \quad (1.7.2)$$

如果我们保持向量不变, 而对所有运算符进行变换, 那么显然会产生相同的变化效果。

$$\Omega \rightarrow U^\dagger\Omega U \quad (1.7.3)$$

第一种情况称为主动变换, 第二种情况称为被动变换。当前的命名方式与向量有关: 它们在主动变换中受到影响, 在被动情况下则保持不变。从算符的角度来看, 情况恰恰相反。

后面我们将看到, 在量子理论中, 物理过程存在于算符的矩阵元中, 并且主动和被动变换为我们提供了两种等效的描述相同物理变换的方式。

练习 1.7.1 * 矩阵的迹被定义为其对角线元素的和。

$$\text{Tr}\Omega = \sum_i \Omega_{ii} \quad (1.7.4)$$

证明

1. $\text{Tr}(\Omega\Lambda) = \text{Tr}(\Lambda\Omega)$
2. $\text{Tr}(\Omega\Lambda\theta) = \text{Tr}(\Lambda\theta\Omega) = \text{Tr}(\theta\Omega\Lambda)$ (排列是周期的)
3. 一个算子的迹在基矢变换 $|i\rangle \rightarrow U|i\rangle$ 不受影响。[等价地, 证明 $\text{Tr}\Omega = \text{Tr}U^\dagger\Omega U$ 。]

练习 1.7.2 证明矩阵的行列式基矢变换不受影响。[等价地, 证明 $\det\Omega = \det(U^\dagger\Omega U)$ 。]

1.8 特征值问题

现在考虑一个作用在任意非零 $\text{ket}|V\rangle$ 的线性算符 Ω :

$$\Omega|V\rangle = |V'\rangle \quad (1.8.1)$$

除非算符恰好是 trivial 的, 例如单位算符或其倍数, 否则 ket 将发生 non-trivial 的变化, 即 $|V'\rangle$ 不能简单地与 $|V\rangle$ 相关。这是对于任意的 ket 而言。然而, 每个算符都有其特定的 ket , 称为它的本征矢, 其中它的作用仅仅是进行重新缩放的作用:

$$\Omega|V\rangle = \omega|V\rangle \quad (1.8.2)$$

方程(1.8.2)是一个本征值方程: $|V\rangle$ 是具有本征值 ω 的算符 Ω 的本征矢。在本章中, 给定一个算符 Ω , 我们将看到如何系统地确定它的所有本征值和本征向量。这样的方程怎么进入物理学的将在本节末尾的几个力学示例中进行说明, 并且一旦我们进入到量子力学的正式内容, 本征、本征、本征将贯穿始终。

例 1.8.1 为了说明特征值问题到底有多容易, 我们将从一个将被完全解决的情况开始: 案例 $\Omega = I$ 。因为

$$I|V\rangle = |V\rangle$$

对于所有的 $|V\rangle$, 我们得出结论

(1) I 的唯一本征值是 1。

(2) 所有向量都是它的本征向量, 其对应的本征值为 1。 \square

例 1.8.2 在这一无条件的成功之后, 我们被鼓励着尝试稍微困难一些的情况: $\Omega = \mathbb{P}_V$, 这个投影算符与一个已归一化的 $ket|V\rangle$ 相关联。很明显

(1) 任何形式为 $\alpha|V\rangle = |\alpha V\rangle$ 与 $|V\rangle$ 平行的 ket , 都是具有本征值为 1 的本征矢:

$$\mathbb{P}_V|\alpha V\rangle = |V\rangle\langle V|\alpha V\rangle = \alpha|V\rangle|V|^2 = 1 \cdot |\alpha V\rangle$$

(2) 任何垂直于 $|V\rangle$ 的 ket , 即 V_\perp , 都是具有本征值为 0 的本征矢:

$$\mathbb{P}_V|V_\perp\rangle = |V\rangle\langle V|V_\perp\rangle = 0 = 0|V_\perp\rangle$$

(3) 一些 ket 既平行又不垂直, $\alpha|V\rangle + \beta|V_\perp\rangle$ 也不是本征矢:

$$\mathbb{P}_V(\alpha|V\rangle + \beta|V_\perp\rangle) = |\alpha V\rangle \neq \gamma(\alpha|V\rangle + \beta|V_\perp\rangle)$$

由于空间中的每个 ket 都属于上述类别之一, 我们已经找到了所有的本征值和本征向量。 \square

例 1.8.3 现在考虑算符 $R(\frac{1}{2}\pi i)$ 。我们已经知道它有一个本征 ket , 即沿着 x 轴的基向量 $|1\rangle$:

$$R(\frac{1}{2}\pi i)|1\rangle = |1\rangle$$

还有其他的本征矢吗? 当然, 沿着 x 轴的任何向量 $\alpha|1\rangle$ 也不受 x 轴旋转的影响。这是本征值方程的一般特征反映了算符的线性性质:

如果

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

那么

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

对于任何倍数 α 来说。因为本征值方程固定了本征向量仅取决于整体缩放因子, 所以我们不会将本征向量的倍数视为不同的本征向量。在这种理解下, 让我们问一下除了 $|1\rangle$ 之外, $R(\frac{1}{2}\pi i)$ 是否还有其他本征向量。我们的直

觉告诉我们没有，因为不沿着 x 轴的任何向量都会被 $R(\frac{1}{2}\pi i)$ 旋转，不可能转化为自身的倍数。由于每个向量要么平行于 $|1\rangle$ ，要么不平行于，我们已经完全解决了本征值问题。

然而这个结论是错误的！ $R(\frac{1}{2}\pi i)$ 除了 $|1\rangle$ 之外还有另外两个本征向量。但是我们的直觉并不应受到责备，因为这些向量位于 $\mathbb{V}^3(C)$ 而不是 $\mathbb{V}^3(R)$ 中。从这个例子中可以清楚地看出，我们需要一种可靠而系统的方法来解决在 $\mathbb{V}^n(R)$ 中本征值问题。现在我们将把注意力转向这个问题。

特征方程与特征值问题的解

我们开始将方程(1.8.2)重写为：

$$(\Omega - \omega I)|V\rangle = |0\rangle \quad (1.8.3)$$

在两边都左乘 $(\Omega - \omega I)^{-1}$ ，假设它存在，我们得到：

$$|V\rangle = (\Omega - \omega I)^{-1}|0\rangle \quad (1.8.4)$$

现在，任何作用于零向量的有限算符（具有有限矩阵元素的算符）只能给出零向量。因此，在要求非零特征向量 $|V\rangle$ 时，我们似乎试图从方程(1.8.4)中无中生有。这是不可能的。由此可见，我们假设算符 $(\Omega - \omega I)^{-1}$ 存在（作为有限算符）是错误的。因此，我们要问的是这种情况何时发生。基本矩阵理论告诉我们（见附录 A.1）矩阵 M 的逆是由给定的

$$M^{-1} = \frac{\text{cofactor } M^T}{\det M} \quad (1.8.5)$$

现在如果矩阵 M 是有限的，那么它的余子式也是有限的。因此，我们需要的是行列式的消失。因此，非零特征向量的条件是

$$\det(\Omega - \omega I) = 0 \quad (1.8.6)$$

这个方程将会确定本征值 ω ，为了找到它，我们将方程(1.8.3)投影到一组基上。方程两边点成一组基 $\text{bra}\langle i|$ ，我们得到

$$\langle i|\Omega - \omega I|V\rangle = 0$$

通过引入单位算符的表示【方程(1.6.7)】，在 $|V\rangle$ 的左边，我们得到方程(1.8.3)的以下形式：

$$\sum_j (\Omega_j - \omega \delta_{ij}) v_j = 0 \quad (1.8.7)$$

将行列式设置为零将给出以下形式的表达式:

$$\sum c_m \omega^m = 0 \quad (1.8.8)$$

方程(1.8.8)被称为特征方程并且

$$P^n(\omega) = \sum_{m=0}^n c_m \omega^m \quad (1.8.9)$$

被称为特征多项式。虽然多项式是在特定基础上确定的,但其多项式的根即特征值是基础无关的,因为它们由抽象方程(1.8.3)定义,该方程与任何基础无关。

现在在分析中的一个基本结果是每个 n 次多项式有 n 个根,这些根不一定是不同的,也不一定是实数。因此,在 $\mathbb{V}_n(C)$ 中的每个算符都有 n 个特征值。一旦知道了特征值,可以找到特征向量,至少对于厄米和么正算符,可以使用以下示例所示的过程来找到特征向量。【在 $\mathbb{V}_n(C)$ 上的非上述类型的算子可能没有 n 个特征向量—参见习题 1.8.4。定理 10 和 12 证明了 \mathbb{V}_n 上的厄米和么正算子将具有 n 个特征向量。】

例 1.8.4 让我们使用上面开发的一般技巧来找到 $R(\frac{1}{2}\pi i)$ 的所有特征向量和特征值。回想一下,表示它的矩阵是:

$$R(\frac{1}{2}\pi i) \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

因此特征方程为:

$$\det(R - \omega I) = \begin{vmatrix} 1 - \omega & 0 & 0 \\ 0 & -\omega & -1 \\ 0 & 1 & -\omega \end{vmatrix} = 0$$

也就是说

$$(1 + \omega)(\omega^2 + 1) = 0 \quad (1.8.10)$$

具有根 $\omega = 1, \pm i$ 。我们知道 ω 对应于 $|1\rangle$ 。让我们看看这个公式的结果。将 $\omega = 1$ 代入公式(1.8.7)中,我们发现对应特征向量的组成部分 x_1 、 x_2 和 x_3 必须遵守以下方程:

$$\begin{bmatrix} 1-1 & 0 & 0 \\ 0 & 0-1 & -1 \\ 0 & 1 & 0-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \left. \begin{array}{l} 0 = 0 \\ -x_2 - x_3 = 0 \\ x_2 - x_3 = 0 \end{array} \right\} \rightarrow x_2 = x_3 = 0$$

因此任何形式的向量都是可以接受的

$$x_1|1\rangle \leftrightarrow \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix}$$

这是预料之中的。惯例上会利用自由度来规范化特征向量。因此，在这种情况下选择

$$|\omega = 1\rangle = |1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

我说的是一种选择而不是所有选择，因为向量可以乘以模为 1 的数字而不改变范数。除非可以选择具有实部分量的向量，否则没有普遍接受的约定来消除这种自由。

请注意，上面的三个同时方程中，第一个不是实方程。通常，只有 $(n-1)$ 个方程。这就是向量范数不固定的原因，正如附录 A.1 所示，行列式为零的原因。

接下来考虑与 $\omega = i$ 对应的方程。特征向量的分量遵从以下方程：

$$\begin{aligned} (1-i)x_1 &= 0 & (i.e., x_1 &= 0) \\ -ix_2 - x_3 &= 0 & (i.e., x_2 &= ix_3) \\ x_2 - ix_3 &= 0 & (i.e., x_2 &= ix_3) \end{aligned}$$

再次注意，我们只有 $n-1$ 个有效方程。上述方程的适当归一化解为：

$$|\omega = i\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix}$$

相同的步骤去处理第三个向量

$$|\omega = -i\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 0 \\ -i \\ 1 \end{bmatrix} \quad \square$$

在上面的例子中，我们引入了一种常用的约定：按照特征值给特征向量标记。例如，与 $\omega = \omega_1$ 对应的 *ket* 被标记为 $|\omega = \omega_i\rangle$ ，或者简单地称为 ω_i 。

这种表示法前提假设是对于每个 ω_i 只有一个向量与之对应。虽然这不总是情况，但只需要稍作修改就可以涵盖一般情况。

当单个特征值表示多于一个特征向量时，这种现象称为简并，并对应于特征多项式的重根。面对简并，我们不仅需要修改标记，还需要修改上述示例中用于找到特征向量的过程。想象一下，如果我们处理的不是 $R(\frac{1}{2}\pi i)$ ，而是另一个在 $\mathbb{V}_3(R)$ 根为 ω_1 和 ω_2 的算符 Ω ，似乎可以通过上述方法获得两个特征向量，每个不同的 ω 对应一个。那么我们怎么得到第三个呢？还是没有第三个吗？当我们详细讨论简并问题时，所有这些问题都将得到一般解答。现在我们把注意力转向两个关于厄米算符的核心定理，它们在量子力学中起着至关重要的作用。

定理 9 厄米算符的特征值是实数。

证明 令

$$\Omega|\omega\rangle = \omega|\omega\rangle$$

两边都左乘 $\langle\omega|$

$$\langle\omega|\Omega|\omega\rangle = \omega\langle\omega|\omega\rangle \quad (1.8.11)$$

取伴随得到

$$\langle\omega|\Omega^\dagger|\omega\rangle = \omega^*\langle\omega|\omega\rangle$$

因为 $\Omega = \Omega^\dagger$ ，方程变成

$$\langle\omega|\Omega|\omega\rangle = \omega^*\langle\omega|\omega\rangle$$

从等式(1.8.11)中减去

$$\begin{aligned} 0 &= (\omega - \omega^*)\langle\omega|\omega\rangle \\ \omega &= \omega^* \quad \text{Q.E.D.} \end{aligned}$$

定理 10 对于每个厄米算符 Ω ，(至少) 存在一个由其正交归一特征向量构成的基。在这个特征基中它是对角化的，其特征值作为对角线上的数字。

证明 让我们从特征方程开始。它必须至少有一个根，记为 ω_1 。对应于 ω_1 ，必须存在至少一个非零特征向量 $|\omega_1\rangle$ [如果不存在，定理 (A.1.1) 将意味着 $(\Omega - \omega_1 I)$ 是可逆的]。考虑到所有与 $|\omega_1\rangle$ 正交的向量构成的子空间

$\mathbb{V}_{n-1}^{\perp 1}$ 。我们选择 $|\omega\rangle$ (归一化为单位向量) 以及任意的 $n-1$ 个正交向量 $\{\mathbb{V}_1^{\perp 1}, \mathbb{V}_{\perp 1}^1, \mathbb{V}_{\perp 1}^2, \dots, \mathbb{V}_{\perp 1}^{n-1}\}$ 作为我们的基向量。在这个基 Ω 有以下形式:

$$\Omega \leftrightarrow \begin{bmatrix} \omega_1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \boxed{} \\ 0 & & & & & & \\ \vdots & & & & & & \\ 0 & & & & & & \end{bmatrix} \quad (1.8.12)$$

第一列只是 Ω 作用在 $|\omega_1\rangle$ 上得到的象。根据 Ω 的厄米性质, 第一行可以由第一列推导出来。

现在特征方程的形式为:

$$(\omega_1 - \omega) \cdot (\text{方块内子矩阵的行列式}) = 0$$

$$(\omega_1 - \omega) \sum_0^{n-1} c_m \omega^m = (\omega_1 - \omega) P^{n-1}(\omega) = 0$$

现在多项式 P^{n-1} 也必须生成一个根 ω_2 , 并且有一个归一化的特征向量 $|\omega_2\rangle$ 。定义子空间 $\mathbb{V}_{\perp 1,2}^{n-2}$, 其中的向量也在 $\mathbb{V}_{\perp 1}^{n-1}$ 中与 $|\omega_2\rangle$ 正交 (自动也与 ω_1 正交), 并重复之前的步骤。最终, 在基向量 $|\omega_1\rangle, |\omega_1\rangle, \dots, |\omega_n\rangle$ 下, 矩阵 Ω 变为:

$$\Omega \leftrightarrow \begin{bmatrix} \omega_1 & 0 & 0 & \cdots & 0 \\ 0 & \omega_2 & 0 & & 0 \\ 0 & 0 & \omega_3 & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & & \omega_n \end{bmatrix}$$

由于每个 $|\omega_i\rangle$ 都是从与先前 $|\omega_1\rangle, |\omega_1\rangle, \dots, |\omega_{i-1}\rangle$ 正交的空间中选择的, 所以特征向量的基是正交的 (请注意, 我们从未假设特征值都是不同的)。Q.E.D.

【实数和厄米算子之间的类比进一步加强了, 原因是在某个特定的基 (特征向量的基) 上, 厄米算子可以用一个所有元素都是实数的矩阵表示。】

在定理 10 的陈述中, 指出可能存在多个能使 n 对角化的特征向量基。如果存在简并性, 这种情况就会发生。假设 $\omega_1 = \omega_2 = \omega$ 。那么我们就有两

个满足正交条件的单位向量，满足

$$\begin{aligned}\Omega|\omega_1\rangle &= \omega|\omega_1\rangle \\ \Omega|\omega_2\rangle &= \omega|\omega_2\rangle\end{aligned}$$

由此可得，

$$\Omega[\alpha|\omega_1\rangle + \beta|\omega_2\rangle] = \alpha\omega|\omega_1\rangle + \beta\omega|\omega_2\rangle = \omega[\alpha|\omega_1\rangle + \beta|\omega_2\rangle]$$

对于任意的 α 和 β 。由于向量 $|\omega_1\rangle$ 和 $|\omega_2\rangle$ 是正交的（因此线性无关），我们可以发现存在一个由 ω_1 和 ω_2 张成的整个二维子空间，其中的元素是 Ω 的特征向量，其特征值为 ω 。这个空间被称为 Ω 的特征空间，其特征值为 ω 。除了向量 $|\omega_1\rangle$ 和 $|\omega_2\rangle$ 之外，还存在无限多个正交的向量对 $|\omega'_1\rangle, |\omega'_2\rangle$ ，通过对 ω 和 ω_2 进行刚性旋转获得，我们可以从中选择任意一对来构建 n 的特征基。一般来说，如果一个特征值出现 m_i 次，也就是说，如果特征方程有 m_i 个根等于某个 ω_i ，那么将存在一个特征空间 $\mathbb{V}_{\omega_i}^{m_i}$ ，我们可以从中选择任意 m_i 个正交向量来构成定理 10 所指的基。

在不存在简并性的情况下，我们可以很容易地证明定理 9 和 10。让我们从两个特征向量开始：

$$\Omega|\omega_i\rangle = \omega_i|\omega_i\rangle \quad (1.8.13a)$$

$$\Omega|\omega_j\rangle = \omega_j|\omega_j\rangle \quad (1.8.13b)$$

第一项与 $\langle\omega_j|$ 点成和第二项与 $\langle\omega_i|$ 点成，我们得到

$$\langle\omega_j|\Omega|\omega_i\rangle = \omega_i\langle\omega_j|\omega_i\rangle \quad (1.8.14a)$$

$$\langle\omega_i|\Omega|\omega_j\rangle = \omega_j\langle\omega_i|\omega_j\rangle \quad (1.8.14b)$$

最后一个等式取伴随并且运用 Ω 的厄米性质，我们得到

$$\langle\omega_i|\Omega|\omega_j\rangle = \omega_j^*\langle\omega_j|\omega_i\rangle$$

从方程(1.8.14a)中减去这个方程，我们得到

$$0 = (\omega_i - \omega_j^*)\langle\omega_j|\omega_i\rangle \quad (1.8.15)$$

如果 $i = j$ ，我们得到，由于 $\langle\omega_i|\omega_j\rangle \neq 0$ ，

$$\omega_i = \omega_i^* \quad (1.8.16)$$

如果 $i \neq j$, 我们得到

$$\langle \omega_i | \omega_j \rangle = 0 \quad (1.8.17)$$

由于假设 $\omega_i - \omega_j^* = \omega_i - \omega_j \neq 0$, 正交性的证明在 $\omega_i = \omega_j$ 的情况下失效并不奇怪, 因为由简并特征值标记的两个向量可以是简并空间中的任意两个成员, 它们不一定是正交的。在这种简并情况下修改该证明需要使用基本上与证明定理 10 相同的论证。定理 10 首先被证明的优点是, 在简并情况下它不需要进行修改。

简并

现在, 我们如承诺的那样, 讨论简并问题。我们对定理 10 的一般分析表明, 面对简并, 我们不仅有一个, 而是无限个单位正交本征基。让我们通过一个例子来看看当我们寻找本征向量时, 这种多样性如何表现出来以及如何处理它。

例 1.8.5 考虑一个具有矩阵元为

$$\Omega \leftrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

在某个基下的算符 Ω 。这个特征方程是

$$(\omega - 2)^2 \omega = 0$$

即

$$\omega = 0, 2, 2$$

对应于 $\Omega = 0$ 的特征向量可以通过通常的方法求得为

$$|\omega = 0\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

对于 $\omega = 2$ 的情况, 得到了以下关于特征向量分量的方程:

$$-x_1 + x_3 = 0$$

$$0 = 0$$

$$x_1 - x_3 = 0$$

现在我们只有一个方程，而不是我们已经习惯了的两个 $(n-1)!$ 这是简并的反映。对于每个除第一个外的根的额外出现，它会消除一个方程。因此，除了通常的规范化自由度外，简并允许我们拥有额外的自由度。条件

$$x_1 = x_3$$

$$x_2 \text{ 任意}$$

定义了一组与第一个向量 $|\omega = 0\rangle$ 垂直的向量，即位于与 $|\omega = 0\rangle$ 垂直的平面上。这与我们对二重简并性应导致二维特征空间的预期一致。 x_2 （或更准确地说， x_2/x_3 的比值）的自由度对应于该平面中的取向自由度。让我们任意选择 $x_2 = 1$ ，得到一个对应于 $\omega = 2$ 的归一化特征向量：

$$|\omega = 2\rangle \leftrightarrow \frac{1}{3^{1/2}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

现在选择第三个向量位于这个平面上，并且与第二个向量正交（位于这个平面上自动使它垂直于第一个向量 $|\omega = 0\rangle$ ）：

$$|\omega = 2, \text{second one}\rangle \leftrightarrow \frac{1}{6^{1/2}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

显然，对于每个不同的比值 x_2/x_3 的选择，我们得到一个不同的具有特征值 2 的正交特征向量双重态。□

注意在面对简并性时， ω_i 不再指代一个特定的态，而是指代特征空间 $\mathbb{V}_{\omega_i}^{m_i}$ 的一个通用元素。要指代一个特定的元素，我们必须使用符号 $|\omega_i, \alpha\rangle$ ，其中 α 标记了特征空间中的态。关于如何选择标记 α 的方法将很快讨论。

现在我们考虑单位算子的类似于定理 9 和定理 10 的推论。

定理 11 单位算子的特征值是具有单位模的复数。

定理 12 单位算子的特征向量是相互正交的。（我们假设这里没有简并性。）

证明这两个定理（假设不存在退化情形），令

$$U|u_i\rangle = u_i|u_i\rangle \quad (1.8.18a)$$

且

$$U|u_j\rangle = u_j|u_j\rangle \quad (1.8.18b)$$

将第二个方程的伴随取出并将其分别与第一个方程对应的一边点乘，我们得到

$$\langle u_j|U^\dagger U|u_i\rangle = u_i u_j^* \langle u_j|u_i\rangle$$

因此

$$(1 - u_i u_j^*) \langle u_j|u_i\rangle = 0 \quad (1.8.19)$$

如果 $i = j$ ，我们得到，由于 $\langle u_i|u_j\rangle \neq 0$ ，

$$u_i u_j^* = 1 \quad (1.8.20a)$$

而如果 $i \neq j$

$$\langle u_i|u_j\rangle = 0 \quad (1.8.20b)$$

因为 $|u_i\rangle \neq |u_j\rangle \Rightarrow u_i \neq u_j \Rightarrow u_i u_j^* \neq u_i u_i^* \Rightarrow u_i u_j^* \neq 1$. (Q.E.D.)

如果 U 是简并的，我们可以进行与厄米算符 Ω 类似的分析，只有一个区别。而在方程(1.8.2)中，第一行的零元素是由第一列的零元素和 $\Omega^\dagger = \Omega$ 所决定的，这里的零元素则来自于每行中元素模的平方之和等于 1 的要求。由于 $|u_1| = 1$ ，第一行的所有其他元素必须为零。

厄米矩阵的对角化

考虑一个在 $\mathbb{V}^n(C)$ 上表示为某个正交归一基底 $|1\rangle, \dots, |i\rangle, \dots, |n\rangle$ 的矩阵的厄米算符 Ω 。如果我们将这个基底转换为特征基底 $|\omega_1\rangle, |\omega_j\rangle, |\omega_n\rangle$ ，表示 Ω 的矩阵将变为对角矩阵。现在，引入的改变基底的算符 U

$$|\omega_i\rangle = U|i\rangle \quad (1.8.21)$$

很明显算符 U 是幺正的，因为它将一个正交归一基底“旋转”成另一个基底。(如果你愿意，可以应用我们的记忆法来验证 U 的幺正性：它的列包含正交归一的特征向量 $|\omega_i\rangle$ 的分量)。这个结果通常可以概括为以下陈述：

任何在 $\mathbb{V}^n(C)$ 上的厄米矩阵都可以通过一个幺正基底变换对角化。

我们可以用被动变换的术语重新陈述这个结果，如下所示：

如果 Ω 是一个厄米矩阵, 那么存在一个由 ω 的特征向量构成的幺正矩阵 U , 使得 $U^\dagger \Omega U$ 是对角矩阵。

因此找到一个能对角化 Ω 的基底的问题等价于解它的特征值问题。

练习 1.8.1 (1) 求该矩阵的所有特征值和规范化特征向量。

$$\Omega = \begin{bmatrix} 1 & 3 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 4 \end{bmatrix}$$

(2) 该矩阵是否为厄米矩阵? 特征向量是否正交?

练习 1.8.2 * 考虑矩阵:

$$\Omega = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

(1) 判断它是否为厄米矩阵。

(2) 找出它的所有特征值和特征向量。

(3) 验证 $U^\dagger \Omega U$ 是否为对角矩阵, 其中 U 是 Ω 特征向量矩阵。

练习 1.8.3 * 考虑矩阵:

$$\Omega = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

(1) 证明 $\omega_1 = \omega_2 = 1; \omega = 2$

(2) 证明 $|\omega = 2\rangle$ 是以下形式的任意向量。

$$\frac{1}{(2a^2)^{1/2}} \begin{bmatrix} 0 \\ a \\ -a \end{bmatrix}$$

(3) 证明 $\omega = 1$ 特征空间包含所有以下形式的向量。

$$\frac{1}{(b^2 + 2c^2)^{1/2}} \begin{bmatrix} b \\ c \\ c \end{bmatrix}$$

可以通过将 $\omega = 1$ 代入方程或要求 $\omega = 1$ 的特征空间与 $|\omega = 2\rangle$ 正交来实现。

练习 1.8.4 任意 $n \times n$ 矩阵不一定有 n 个特征向量。以以下矩阵为例：

$$\Omega = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$$

(1) 证明 $\omega_1 = \omega_2 = 3$ 。

(2) 通过代入该特征值，证明我们只能得到一个形式为的特征向量，即

$$\frac{1}{(2a^2)^{1/2}} \begin{bmatrix} +a \\ -a \end{bmatrix}$$

我们无法找到另一个线性独立的特征向量。

练习 1.8.5 * 考虑矩阵

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

(1) 证明它是酉矩阵。

(2) 证明其特征值为 $e^{i\theta}$ 和 $e^{-i\theta}$ 。

(3) 找到相应的特征向量，并证明它们是正交的。

(4) 验证 $U^\dagger \Omega U = (\text{对角矩阵})$ ，其中 U 是 Ω 的特征向量矩阵。

练习 1.8.6 * (1) 我们已经看到，矩阵在酉变换下的行列式保持不变。现在证明对于一个 Hermitian 或酉矩阵 Ω 有

$$\det \Omega = \text{product of eigenvalues of } \Omega = \prod_{i=1}^n \omega_i$$

(2) 相同变换下迹的不变性，证明

$$\text{Tr } \Omega = \sum_{i=1}^n \omega_i$$

练习 1.8.7 利用上题中关于迹和行列式的结论，证明矩阵

$$\Omega = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

的特征值为 3 和 -1。通过显式计算进行验证。注意，此处矩阵的 Hermitian 性质是必不可少的。

练习 1.8.8 * 考虑满足下列条件的 *Hermitian* 矩阵 M^1, M^2, M^3, M^4 :

$$M^i M^j + M^j M^i = 2\delta^{ij} I, \quad i, j = 1, \dots, 4$$

(1) 证明 M^i 的特征值为 ± 1 。(提示: 转换到 M^i 的特征基上, 利用 $i=j$ 时的方程。)

(2) 通过考虑关系式

$$M^i M^j = -M^j M^i \quad \text{for } i \neq j$$

证明 M^i 是无迹的。(提示: $\text{Tr}(ACB) = \text{Tr}(CBA)$ 。)

(3) 证明它们不能是奇数维的矩阵。

练习 1.8.9 给定一组质量为 m_a 的物体, 它们位于 \mathbf{r}_a 处, 并绕一个共同轴以角速度 $\boldsymbol{\omega}$ 旋转。该系统的角动量可以表示为

$$\mathbf{l} = \sum_a m_a (\mathbf{r}_a \times \mathbf{v}_a)$$

其中 $\mathbf{v}_a = \boldsymbol{\omega} \times \mathbf{r}_a$ 是 m_a 的速度。通过使用恒等式

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

证明 I 的每个笛卡尔分量 l_i 可以表示为

$$l_i = \sum M_{ij} \omega_j$$

其中

$$M_{ij} = \sum_a m_a [r_a^2 \delta_{ij} - (\mathbf{r}_a)_i (\mathbf{r}_a)_j]$$

或者在 *Dirac* 符号中表示为

$$|l\rangle = M|\omega\rangle$$

(1) 角动量与角速度是否总是平行的?

(2) 证明惯性矩阵 M_{ij} 是 *Hermitian* 的。

(3) 现在证明存在三个方向, 使得 I 和 $\boldsymbol{\omega}$ 相平行。这些方向如何找到?

(4) 考虑一个球体的惯性矩阵。由于球体的完全对称性, 很清楚每个方向都是它的旋转本征方向。这能说明矩阵 M 的三个特征值吗?

同时对角化两个厄米算符

接下来我们考虑同时对角化两个厄米算符的问题。

定理 13 如果 Ω 和 Λ 是两个可交换的厄米算符，那么存在（至少）一组共同特征向量的基，可以同时对角化它们。

证明 首先考虑至少一个算符是非简并的情况，即对于给定的特征值，存在唯一的特征向量（除了一个比例因子）。让我们假设 Ω 是非简并的。考虑其中的一个特征向量：

$$\begin{aligned}\Omega|\omega_i\rangle &= \omega_i|\omega_i\rangle \\ \Lambda\Omega|\omega_i\rangle &= \omega_i\Lambda|\omega_i\rangle\end{aligned}$$

因为 $[\Lambda, \Omega] = 0$

$$\Omega\Lambda|\omega_i\rangle = \omega_i\Lambda|\omega_i\rangle \quad (1.8.22)$$

即 $\Lambda|\omega_i\rangle$ 是 Ω 特征值为 ω_i 的特征向量。由于这个向量只有一个唯一的特征向量（除了一个比例因子），

$$\Lambda|\omega_i\rangle = \lambda_i|\omega_i\rangle \quad (1.8.23)$$

因此， ω_i 也是 Λ 的特征向量，其特征值为 λ_i 。由于 Ω 的每个特征向量都是 Λ 的特征向量，因此，基 ω_i 将同时将两个算符对角化。由于 A 是非简并的，因此只有这样的一组基满足这个性质。

如果两个算符都是简并的怎么办呢？通过对基向量进行排序，使每个特征空间中的元素相邻，我们可以将它们之中的一个（比如 Ω ）表示为如下形式（定理 10）：

$$\Omega \leftrightarrow \begin{bmatrix} \omega_1 & & & & & \\ & \omega_1 & & & & \\ & & \ddots & & & \\ & & & \omega_1 & & \\ & & & & \omega_2 & \\ & & & & & \ddots \\ & & & & & & \omega_m \\ & & & & & & & \omega_n \end{bmatrix}$$

现在这个基是不唯一的：在每个特征值为 λ_i 对应的特征空间 $\mathbb{V}_{\omega_i}^{m_i} \equiv \mathbb{V}_i^{m_i}$ 中，存在无穷多个基。让我们在每个特征空间 $\mathbb{V}_{\omega_i}^{m_i}$ 中任意选择一个集合 $|\omega_i, \alpha\rangle$ ，其中额外的标签 α 从 1 取到 m_i 。

在这个基中 Λ 是如何呈现的？虽然我们没有特别努力将 Λ 变成一个简单的形式，但它已经是一个简单的形式因为它与 Ω 对易的。让我们开始模仿在非简并情况下的证明：

$$\Omega\Lambda|\omega_i, \alpha\rangle = \Lambda\Omega|\omega_i, \alpha\rangle = \omega_i\Lambda|\omega_i, \alpha\rangle$$

然而，由于 Ω 的简并性，我们仅能推断出：

$$\Lambda|\omega_i, \alpha\rangle \text{ lies in } \mathbb{V}_i^{m_i}$$

现在，由于不同特征空间的向量是正交的【方程式(1.8.15)】，

$$\langle\omega_j, \beta|\Lambda|\omega_i, \alpha\rangle = 0$$

如果 $|\omega_i, \alpha\rangle$ 和 $|\omega_j, \beta\rangle$ 是基向量，其中 $\omega_i \neq \omega_j$ 。因此，在这个基中，

$$\Lambda \leftrightarrow \begin{bmatrix} \boxed{\Lambda_1} & & & 0 \\ & \boxed{\Lambda_2} & & \\ & & \ddots & \\ 0 & & & \boxed{\Lambda_k} \end{bmatrix}$$

顾名思义这被称为分块对角矩阵是显而易见的。 Λ 的分块对角形式反映了以下事实：当 Λ 作用于特征空间 $\mathbb{V}_i^{m_i}$ 的某个元素 $|\omega_i, \alpha\rangle$ 时，它将其转化为 $\mathbb{V}_i^{m_i}$ 中的另一个元素。在每个子空间 i 中， Λ 由矩阵 Λ_i 给出，它在上面的方程中出现为一个块。考虑 $\mathbb{V}_i^{m_i}$ 中的矩阵 Λ_i 。它是厄米的，因为 Λ 是厄米的。显然，通过将我们开始时在 $\mathbb{V}_i^{m_i}$ 中的基 $|\omega_i, 1\rangle, |\omega_i, 2\rangle \cdots |\omega_i, m_i\rangle$ 替换为 Λ_i 的特征基，我们可以将其对角化。让我们在每个特征空间中进行这样的基变换，从而使 Λ 成为对角矩阵。与此同时 Ω 呢？它当然保持对角形式，因为它对于每个简并特征空间中正交归一基的选择是无关紧要的。如

果 Λ_i 的特征值是 $\lambda_i^1, \lambda_i^2, \dots, \lambda_i^{m_i}$, 则我们最终得到

$$\Lambda \leftrightarrow \begin{bmatrix} \lambda_1^{(1)} & & & & \\ & \lambda_1^{(2)} & & & \\ & & \ddots & & \\ & & & \lambda_1^{(m_1)} & \\ & & & & \lambda_1^{(m_1)} \\ & & & & & \ddots \\ & & & & & & \lambda_k^{(m_k)} \end{bmatrix}$$

$$\Omega \leftrightarrow \begin{bmatrix} \omega_1 & & & & \\ & \omega_1 & & & \\ & & \ddots & & \\ & & & \omega_1 & \\ & & & & \omega_2 \\ & & & & & \ddots \\ & & & & & & \omega_n \end{bmatrix}$$

Q.E.D.

如果 Λ 在给定的任何子空间中都没有简并性, 即对于任意的 k, l 和 i , $\lambda_i^k \neq \lambda_i^l$, 那么我们得到的基是唯一的: 在每个特征子空间中, 由 Ω 给出的自由度完全被 Λ 消除。该基的元素可以通过索引对 ω 和 λ 唯一地命名为 $|\omega, \lambda\rangle$, 其中 λ 扮演额外的标签 α 的角色。如果 Λ 在 Ω 的特征空间内是简并的, 如果说 $\lambda_1^{(1)} = \lambda_1^{(2)}$, 则存在一个二维特征空间, 我们可以从中选择任意两个正交向量作为共同基向量。然后需要引入第三个算符与 Ω 和 Λ 都对易且在此子空间内非简并的算符 Γ 。通常情况下, 对于有限的 n , 我们总是可以找到一组相互对易的算符 $\{\Omega, \Lambda, \Gamma, \dots\}$, 从而确定一个唯一的共同特征基, 其元素可以无歧义地标记为 $|\omega, \lambda, \gamma, \dots\rangle$ 。在我们对量子力学的研究中, 假设如果 n 是无穷的则存在这样一组完备的对易算符。

练习 1.8.10 * 通过考虑对易子, 证明下列 Hermitian 矩阵可以同时对角化。找到两者共同的特征向量, 并验证在变换到这个基向量的么正变换下, 这两

个矩阵都会被对角化。

$$\Omega = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

由于 Ω 是简并的，而 Λ 不是，因此在决定基的选择时必须谨慎考虑哪个矩阵起主导作用。

例 1.8.6 我们现在将详细讨论力学问题的完整解决方案。重要的是，你要彻底理解这个例子，因为它不仅说明了本章中开发的数学技术的使用，还包含了量子力学中的核心问题的主要特征。

所讨论的力学系统如图 (1.5) 所示。两个质量为 m 的物体通过弹性系数为 k 的弹簧相互耦合，并连接到墙壁上。如果 x_1 和 x_2 分别表示物体相对于它们的平衡位置的位移，这些坐标遵循以下方程，通过对牛顿定律的简单应用得出：

$$\dot{x}_1 = -\frac{2k}{m}x_1 + \frac{k}{m}x_2 \quad (1.8.24a)$$

$$\ddot{x}_2 = \frac{k}{m}x_1 - \frac{2k}{m}x_2 \quad (1.8.24b)$$

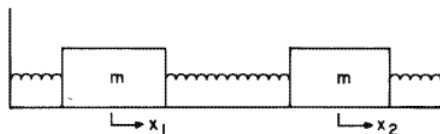


图 1.5: 相互耦合的质量问题。所有的质量都是 m ，所有的弹性系数都是 k ，物体相对于平衡位置的位移分别为 x_1 和 x_2 。

问题是在给定初始值数据的情况下找到 $x_1(t)$ 和 $x_2(t)$ ，其中初始值数据包括初始位置和速度。如果我们限制在零初始速度的情况下，我们的问题是在给定 $x_1(0)$ 和 $x_2(0)$ 的情况下找到 $x_1(t)$ 和 $x_2(t)$ 。

接下来，我们将用线性向量空间的术语来表述问题，并利用本章中介绍的方法来解决它。作为第一步，我们将方程(1.8.24)以矩阵形式重写：

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (1.8.25a)$$

其中厄米矩阵的矩阵元 Ω_{ij} 是

$$\Omega_{11} = \Omega_{22} = -2k/m, \quad \Omega_{12} = \Omega_{21} = k/m \quad (1.8.25b)$$

现在我们将 x_1 和 x_2 视为抽象向量 $|x\rangle$ 的分量, Ω_{ij} 视为一个厄米算符 Ω 的矩阵元素。由于向量 $|x\rangle$ 具有两个实分量, 它是 $\mathbb{V}^2(R)$ 的一个元素, 而 Ω 是 $\mathbb{V}^2(R)$ 上的一个厄米算符。方程(1.8.25a)的抽象形式是:

$$|\ddot{x}(t)\rangle = \Omega|x(t)\rangle \quad (1.8.26)$$

通过方程(1.8.25b)到基向量 $|1\rangle$ 和 $|2\rangle$ 上投影得到方程(1.8.25a), 其具有以下物理意义:

$$|1\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \leftrightarrow \begin{bmatrix} \text{first mass displaced by unity} \\ \text{second mass undisplaced} \end{bmatrix} \quad (1.8.27a)$$

$$|2\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} \text{first mass undisplaced} \\ \text{second mass displaced by unity} \end{bmatrix} \quad (1.8.27b)$$

表示物体被位移 x_1 和 x_2 的任意状态, 由下式这组基给出

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_2 \quad (1.8.28)$$

上述方程的抽象对应式为:

$$|x\rangle = |1\rangle x_1 + |2\rangle x_2 \quad (1.8.29)$$

在这个 $|1\rangle, |2\rangle$ 基础上, Ω 由出现在方程(1.8.25)中的矩阵表示, 其矩阵元为 $-2k/m, k/m$ 等。

基向量 $|1\rangle, |2\rangle$ 在物理上非常有意义, 因为在这个基础上 $|x\rangle$ 的分量 (x_1 和 x_2) 可以简单地解释为质量的位移。然而, 从寻找初始值问题的数学解的角度来看, 它并不是很理想, 因为分量 x_1 和 x_2 满足耦合满足耦合的微分方程(1.8.24a)和(1.8.24b)。耦合是通过非对角矩阵元 $\Omega_{12} = \Omega_{21} = k/m$ 来调节的。

在确定了 $|1\rangle, |2\rangle$ 基的问题之后, 我们现在可以看到如何解决它: 我们必须切换到一个 Ω 对角化的基础上。在这个基上 $|x\rangle$ 的分量将满足另一个解耦的微分方程, 可以很容易地求解。在找到解之后, 我们可以回到物理上更可取的 $|1\rangle, |2\rangle$ 基。这就是我们的大致策略, 现在我们转向细节。

通过对厄米算符的研究, 我们知道对角化 Ω 的基础是其归一化特征向量的基础。让 $|I\rangle$ 和 $|II\rangle$ 为其特征向量, 定义为

$$\Omega|I\rangle = -\omega_I^2|I\rangle \quad (1.8.30a)$$

$$\Omega|II\rangle = -\omega_{II}^2|II\rangle \quad (1.8.30b)$$

我们在这里改变了通常的表示法： Ω 的特征值写为 $-\omega^2$ 而不是 ω ，这是因为我们预期 Ω 具有 $-\omega^2$ 的形式的特征值，其中 ω 是实数。我们还使用符号 $|I\rangle$ 和 $|II\rangle$ 来表示我们惯常称为 $|- \omega_I^2\rangle$ 和 $|- \omega_{II}^2\rangle$ 的符号。

解决 Ω 在 $|1\rangle, |2\rangle$ 基底中（其中 Ω 的矩阵元素已知）的特征值问题是一个简单的练习（你应该完成），并且得到

$$\omega_{II} = \left(\frac{k}{m}\right)^{1/2}, \quad |I\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (1.8.31a)$$

$$\omega_{II} = \left(\frac{k}{m}\right)^{1/2}, \quad |II\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (1.8.31b)$$

如果我们现在将向量 $|x(t)\rangle$ 在这个新基础上展开，如下所示：

$$|x(t)\rangle = |I\rangle x_I(t) + |II\rangle x_{II}(t) \quad (1.8.32)$$

【类比于公式(1.8.29)】，分量 x_I 和 x_{II} 将按以下方式演化：

$$\begin{aligned} \begin{bmatrix} \ddot{x}_I \\ \ddot{x}_{II} \end{bmatrix} &= \begin{bmatrix} -\omega_I^2 & 0 \\ 0 & -\omega_{II}^2 \end{bmatrix} \begin{bmatrix} x_I \\ x_{II} \end{bmatrix} \\ &= \begin{bmatrix} -\omega_I^2 x_I \\ -\omega_{II}^2 x_{II} \end{bmatrix} \end{aligned} \quad (1.8.33)$$

我们将公式(1.8.26)重写为 $|I\rangle, |II\rangle$ 基础上来获得这个方程，其中 Ω 的特征值成为对角元， $|x\rangle$ 的分量为 x_I 和 x_{II} 。我们可以将算符应用于

$$\frac{d^2}{dt^2} - \Omega$$

公式(1.8.32)的两边，得到：

$$|0\rangle = |I\rangle(\ddot{x}_I + \omega_I^2 x_I) + |II\rangle(\ddot{x}_{II} + \omega_{II}^2 x_{II}) \quad (1.8.34)$$

由于 $|I\rangle$ 和 $|II\rangle$ 是正交的，所以每个系数都为零。

解耦方程组的解为

$$\ddot{x}_i + \omega_i^2 x_i = 0, \quad i = I, II \quad (1.8.35)$$

在初始速度为零的条件下，解为

$$x_i(t) = x_i(0) \cos \omega_i t, \quad i = I, II \quad (1.8.36)$$

正如预期的那样, $|x\rangle$ 在 $|\text{I}\rangle, |\text{II}\rangle$ 基上的分量遵循可以轻松被求解的解耦方程。将公式(1.8.36)代入公式(1.8.32)我们得到

$$|x(t)\rangle = |\text{I}\rangle x_{\text{I}}(0) \cos \omega_{\text{I}} t + |\text{II}\rangle x_{\text{II}}(0) \cos \omega_{\text{II}} t \quad (1.8.37a)$$

$$= |\text{I}\rangle \langle \text{I} | x(0) \rangle \cos \omega_{\text{I}} t + |\text{II}\rangle \langle \text{II} | x(0) \rangle \cos \omega_{\text{II}} \quad (1.8.37b)$$

方程(1.8.37)为初始值问题提供了明确的解。它对应于以下算法, 用于给定 $|x(0)\rangle$ 求解 $|x(t)\rangle$ 。

步骤 (1) 解决 Ω 的特征值问题。

步骤(2)在下面展开式中找到系数 $x_{\text{I}}(0) = \langle \text{I} | x(0) \rangle$ 和 $x_{\text{II}}(0) = \langle \text{II} | x(0) \rangle$ 。

$$|x(0)\rangle = |\text{I}\rangle x_{\text{I}}(0) + |\text{II}\rangle x_{\text{II}}(0)$$

步骤 (3) 为每个系数 $x_i(0) (i = \text{I}, \text{II})$ 添加时间依赖性 $\cos \omega_i t$ 以获得 $|x(t)\rangle$ 展开中的系数

现在让我通过解决以下 (一般) 的初值问题来说明这个算法: 在 $t = 0$ 时, 给定质量被 $x_1(0)$ 和 $x_2(0)$ 位移的情况下, 找到系统的未来状态。

步骤 (1)。我们可以忽略这一步, 因为特征值问题已经被解决【公式(1.8.31)】。

步骤 (2)

$$x_{\text{I}}(0) = \langle \text{I} | x(0) \rangle = \frac{1}{2^{1/2}}(1, 1) \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \frac{x_1(0) + x_2(0)}{2^{1/2}}$$

$$x_{\text{II}}(0) = \langle \text{II} | x(0) \rangle = \frac{1}{2^{1/2}}(1, -1) \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \frac{x_1(0) - x_2(0)}{2^{1/2}}$$

步骤 (3)

$$|x(t)\rangle = |\text{I}\rangle \frac{x_1(0) + x_2(0)}{2^{1/2}} \cos \omega_{\text{I}} t + |\text{II}\rangle \frac{x_1(0) - x_2(0)}{2^{1/2}} \cos \omega_{\text{II}} t$$

通过将 $|x(t)\rangle$ 投影到 $|1\rangle, |2\rangle$ 基底上找到物块的位移 $x_1(t)$ 和 $x_2(t)$ 可以得

到上式的显式解。我们可以得到（输入 ω_I 和 ω_{II} 的显式公式）

$$\begin{aligned}
 x_1(t) &= \langle 1|x(t) \rangle \\
 &= \langle 1|I \rangle \frac{x_1(0) + x_2(0)}{2^{1/2}} \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] + \langle 1|II \rangle \frac{x_1(0) - x_2(0)}{2^{1/2}} \cos \left[\left(\frac{3k}{m} \right)^{1/2} t \right] \\
 &= \frac{1}{2} [x_1(0) + x_2(0)] \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] + \frac{1}{2} [x_1(0) - x_2(0)] \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right]
 \end{aligned} \tag{1.8.38a}$$

利用以下事实：

$$\langle 1|I \rangle = \langle 1|II \rangle = 1/2^{1/2}$$

同样可以证明：

$$x_2(t) = \frac{1}{2} [x_1(0) + x_3(0)] \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] - \frac{1}{2} [x_1(0) - x_2(0)] \cos \left[\left(\frac{3k}{m} \right)^{1/2} t \right] \tag{1.8.38b}$$

我们可以将公式(1.8.38)重新以矩阵形式表示为：

$$\begin{aligned}
 \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= \begin{bmatrix} \frac{\cos [(k/m)^{1/2}t] + \cos [(3k/m)^{1/2}t]}{2} & \frac{\cos [(k/m)^{1/2}t] - \cos [(3k/m)^{1/2}t]}{2} \\ \frac{\cos [(k/m)^{1/2}t] - \cos [(3k/m)^{1/2}t]}{2} & \frac{\cos [(k/m)^{1/2}t] + \cos [(3k/m)^{1/2}t]}{2} \end{bmatrix} \\
 &\times \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}
 \end{aligned} \tag{1.8.39}$$

这就完成了我们在给定初始状态的情况下对系统未来状态的确定。

传播子

公式(1.8.39)中有两个显著的特点：

- (1) 最终态矢量是通过将初始态矢量与一个矩阵相乘得到的。
- (2) 这个矩阵与初始态无关。我们称这个矩阵为传播子。找到传播子相当于找到问题的完全解，对于任何其他具有位移 $\tilde{x}_1(t)$ 和 $\tilde{x}_2(t)$ 的初始态，我们可以通过将相同的矩阵应用于初始态矢量来得到 $\tilde{x}_1(t)$ 和 $\tilde{x}_2(t)$ 。

我们可以将公式(1.8.39)视为在抽象关系的基础上用 $|1\rangle, |2\rangle$ 表示的图像。

$$|x(t)\rangle = U(t)|x(0)\rangle \quad (1.8.40)$$

通过将这个方程与公式(1.8.37b)进行比较, 我们可以找到 U 的抽象表示:

$$U(t) = |I\rangle\langle I| \cos \omega_I t + |II\rangle\langle II| \cos \omega_{II} t \quad (1.8.41a)$$

$$= \sum_{i=1}^{II} |i\rangle\langle i| \cos \omega_i t \quad (1.8.41b)$$

你可以很容易地说服自己, 如果我们在 $|1\rangle, |2\rangle$ 基础上取这个算符的矩阵元素, 我们会重新得到出现在公式(1.8.39)中的矩阵。例如:

$$\begin{aligned} U_{11} &= \langle 1|U|1\rangle \\ &= \langle 1| \left\{ |I\rangle\langle I| \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] + |II\rangle\langle II| \cos \left[\left(\frac{3k}{m} \right)^{1/2} t \right] \right\} |1\rangle \\ &= \langle 1|I\rangle\langle I|1\rangle \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] + \langle 1|II\rangle\langle II|1\rangle \cos \left[\left(\frac{3k}{m} \right)^{1/2} t \right] \\ &= \frac{1}{2} \left\{ \cos \left[\left(\frac{k}{m} \right)^{1/2} t \right] + \cos \left[\left(\frac{3k}{m} \right)^{1/2} t \right] \right\} \end{aligned}$$

注意到 $U(t)$ 【公式(1.8.41b)】的确定完全取决于 Ω 的特征向量和特征值。我们可以将我们之前的算法重新表述如下。要解决方程:

$$|\ddot{x}\rangle = \Omega|x\rangle$$

- (1) 解 Ω 的特征值问题。
- (2) 利用特征值和特征向量构造传播子 U 。
- (3) $|x(t)\rangle = U(t)|x(0)\rangle$ 。

简正模式

有两个初始状态 $|x(0)\rangle$, 其时间演化特别简单。毫不奇怪, 这些是特征态 $|I\rangle$ 和 $|II\rangle$ 。假设我们有 $|x(0)\rangle = |I\rangle$ 。那么在时间 t 处的状态是:

$$\begin{aligned} |I(t)\rangle &\equiv U(t)|I\rangle \\ &= (|I\rangle\langle I| \cos \omega_I t + |II\rangle\langle II| \cos \omega_{II} t)|I\rangle \\ &= |I\rangle \cos \omega_I t \end{aligned} \quad (1.8.42)$$

因此, 以 $|I\rangle$ 为初始状态的系统只受到整体因子 $\cos \omega_I t$ 的修正。类似的情况也适用于 $I \rightarrow II$ 。这两种振动模式中, 向量的所有 (两个) 分量都同步振荡, 称为简正模式。

简正模式的物理意义在 $|I, II\rangle$ 基下是很容易理解的。在这组基中

$$|I\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

并且对应于两个物块都平移相同距离的状态。此时, 中间的弹簧只是旁观者, 每个质量分别振动, 其频率为 $\omega_I = (k/m)^{1/2}$, 是对其最近的末端弹簧做出的响应。因此,

$$|I(t)\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} \cos[(k/m)^{1/2}t] \\ \cos[(k/m)^{1/2}t] \end{bmatrix}$$

另一方面, 如果我们从以下状态开始:

$$|II\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

物块以相等但相反的位移偏离平衡位置。在这种情况下, 中间的弹簧会被扭曲, 位移是每个物块位移的两倍。如果分别用 Δ 和 $-\Delta$ 来表示物块的调整量, 那么每个物块都会感受到一个回复力 $3k\Delta$ (其中 $2k\Delta$ 来自中间的弹簧, $k\Delta$ 来自最靠近它的末端弹簧)。由于有效力常数是 $k_{\text{eff}} = 3k\Delta/\Delta = 3k$, 振动频率为 $(3k/m)^{1/2}$, 并且

$$|II(t)\rangle \leftrightarrow \frac{1}{2^{1/2}} \begin{bmatrix} \cos[(3k/m)^{1/2}t] \\ -\cos[(3k/m)^{1/2}t] \end{bmatrix}$$

如果系统起始于 $|I\rangle$ 和 $|II\rangle$ 的线性组合, 它会演化为相应的正则模式 $|I(t)\rangle$ 和 $|II(t)\rangle$ 的线性组合。这就是传播子方程的内容。

$$\begin{aligned} |x(t)\rangle &= U(t)|x(0)\rangle \\ &= |I\rangle\langle I|x(0)\rangle \cos \omega_I t + |II\rangle\langle II|x(0)\rangle \cos \omega_{II} t \\ &= |I(t)\rangle\langle I|x(0)\rangle + |II(t)\rangle\langle II|x(0)\rangle \end{aligned}$$

另一种看待初始态 $|I\rangle$ 和 $|II\rangle$ 的演化的简单方法是确定在 $|I\rangle, |II\rangle$ 基下代表 U 的矩阵:

$$U \xleftrightarrow[\substack{\text{I, II} \\ \text{basis}}]{\substack{\text{I, II} \\ \text{basis}}} \begin{bmatrix} \cos \omega_{\text{I}} t & 0 \\ 0 & \cos \omega_{\text{II}} t \end{bmatrix} \quad (1.8.43)$$

你应该通过取方程(1.8.41b)中 $U(t)$ 的适当矩阵元来验证这个结果。由于上面的每一列都是相应基向量 ($|\text{I}\rangle$ 或者 $|\text{II}\rangle$) 在作用 $U(t)$ (即时间演化) 后的象, 我们看到初始态 $|\text{I}\rangle$ 和 $|\text{II}\rangle$ 简单地随时间演化。

量子力学中的中心问题与我们刚才讨论的简单例子非常相似。一个系统的状态在量子理论中用符号为 $|\psi\rangle$ 的矢量来描述, 并且遵守 Schrödinger 方程:

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

其中 \hbar 是一个与普朗克常数 h 相关的常数, 由 $\hbar = h/2\pi$ 计算得出。 H 是一个厄米算符, 称为哈密顿算符。问题是在给定 $|\psi(0)\rangle$ 情况下找到 $|\psi(t)\rangle$ 。【由于方程是 t 的一阶方程, 不需要对 $|\dot{\psi}(0)\rangle$ 做任何假设, 它由 Schrödinger 方程确定为 $(-i/\hbar)H|\psi(0)\rangle$ 。】

在大多数情况下, H 是一个时间无关的算符, 解决这个初值问题的算法与我们刚才看到的算法完全类似:

步骤 (1): 解 H 的本征值问题。

步骤 (2): 根据 H 的本征向量和本征值找到传播子 $U(t)$ 。

步骤 (3): $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ 。

当然, 你必须等到第四章才能找到 $|\psi\rangle$ 的物理解释、算符 H 的真正形式以及 $U(t)$ 与 H 的本征值和本征向量之间的精确关系。□

练习 1.8.11 考虑上述耦合质量问题。

(1) 假设初始状态是 $|1\rangle$, 其中第一个物块被位移单位长度而第二个质量保持不变, 按照算法计算出 $|1(t)\rangle$ 。

(2) 将你的结果与从等式(1.8.39)得到的结果进行比较。

练习 1.8.12 再次考虑上一个例子中的问题。

(1) 假设

$$|\ddot{x}\rangle = \Omega|x\rangle$$

有一个解

$$|x(t)\rangle = U(t)|x(0)\rangle$$

求出 $U(t)$ 满足的微分方程。利用 $x(0)$ 是任意的这一事实。

(2) 假设 Ω 和 U 可以同时对角化, 用这个共同的基解决矩阵 U 的元素并重新得到等式(1.8.43)。假设 $|\dot{x}(0)\rangle = 0$ 。

1.9 算符函数及相关概念

我们已经遇到了两种作用于向量的对象: 标量, 它们互相之间和与所有算符之间都可以交换; 算符则通常不能相互交换。习惯上, 我们把前者称作 c 数, 后者称作 q 数。我们经常用 c 数的函数, 如 $\sin(x)$ 、 $\log(x)$ 等, 那么问题来了: 是否可以将 q 数的函数也赋予一个有意义的含义? 在这里, 我们仅考虑那些可以表示为幂级数的函数。考虑一个级数:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad (1.9.1)$$

其中 x 是一个 c 数。我们定义一个算符函数或者 q 数为

$$f(\Omega) = \sum_{n=0}^{\infty} a_n \Omega^n \quad (1.9.2)$$

这个定义只有在总和收敛到一个明确的极限时才有意义。为了看清楚这是什么意思, 我们来考虑一个常见的例子:

$$e^{\Omega} = \sum_{n=0}^{\infty} \frac{\Omega^n}{n!} \quad (1.9.3)$$

让我们将注意力限制在厄米矩阵上。通过厄米矩阵 Ω 的本征态基, 我们可以轻松地进行等式(1.9.3)的求和。因为

$$\Omega = \begin{bmatrix} \omega_1 & & & \\ & \omega_2 & & \\ & & \ddots & \\ & & & \omega_n \end{bmatrix} \quad (1.9.4)$$

和

$$\Omega^m = \begin{bmatrix} \omega_1^m & & & \\ & \omega_2^m & & \\ & & \ddots & \\ & & & \omega_n^m \end{bmatrix} \quad (1.9.5)$$

$$e^\alpha = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{\omega_1^m}{m!} & & & \\ & \ddots & & \\ & & \sum_{m=0}^{\infty} \frac{\omega_n^m}{m!} & \end{bmatrix} \quad (1.9.6)$$

由于每个求和都收敛到熟悉的极限 e^{ω_i} ，因此在这个基础上（因此在任何其他基础上）用幂级数定义了算符 e^Ω 。

练习 1.9.1 * 我们知道级数

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

可以等同于函数 $f(x) = (1-x)^{-1}$ ，如果 $|x| < 1$ 。通过进入本征基研究当 q 数幂级数

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

属于厄米算符 Ω 时，它可以被认同为 $(1-\Omega)^{-1}$ 。

练习 1.9.2 * 如果 H 是一个埃尔米特算符，证明 $U = e^{iH}$ 是么正的。（注意与 c 数的类比：如果 θ 是实数， $u = e^{i\theta}$ 是模长为 1 的数。）

练习 1.9.3 对于上述情况，证明 $\det U = e^{i\text{Tr}H}$ 。

算符参数的导数

考虑下一个依赖于参数 λ 的算子 $\theta(\lambda)$ 。它对 λ 的导数被定义为：

$$\frac{d\theta(\lambda)}{d\lambda} = \lim_{\Delta\lambda \rightarrow 0} \left[\frac{\theta(\lambda + \Delta\lambda) - \theta(\lambda)}{\Delta\lambda} \right]$$

如果将 $\theta(\lambda)$ 在某个基底下表示为矩阵, 则表示 $d\theta(\lambda)/d\lambda$ 的矩阵是通过对 $\theta(\lambda)$ 的矩阵元素进行微分得到的。我们感兴趣的是 $\theta(\lambda)$ 的一个特殊情况。

$$\theta(\lambda) = e^{\lambda\Omega}$$

其中 Ω 是厄米的, 我们能够通过 Ω 的本征基来证明

$$\frac{d\theta(\lambda)}{d\lambda} = \Omega e^{\lambda\Omega} = e^{\lambda\Omega} \Omega = \theta(\lambda) \Omega \quad (1.9.7)$$

即使 Ω 不是厄米的, 也可以通过使用幂级数来得到相同的结果, 前提是幂级数存在。

$$\frac{d}{d\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n \Omega^n}{n!} = \sum_{n=1}^{\infty} \frac{n \lambda^{n-1} \Omega^n}{n!} = \Omega \sum_{n=1}^{\infty} \frac{\lambda^{n-1} \Omega^{n-1}}{(n-1)!} = \Omega \sum_{m=0}^{\infty} \frac{\lambda^m \Omega^m}{m!} = \Omega e^{\lambda\Omega}$$

相反地, 我们可以说如果我们面对微分方程(1.9.7), 其解由以下给出:

$$\theta(\lambda) = c \exp \left(\int_0^\lambda \Omega d\lambda' \right) = c \exp(\Omega \lambda)$$

(在此假设指数函数存在。)在上述方程中, c 是积分的常数(算子)。解 $\theta = e^{\lambda\Omega}$ 对应于选择 $c = I$ (单位算子)。

在上述所有运算中, 我们可以看到 Ω 的行为就像它是一个复数一样。现在, c 数和 q 数之间的真正区别在于后者通常不对易。然而, 如果只有一个 q 数(或其幂)参与到问题中, 所有物理量都对易, 我们可以将它们视为 c 数。如果记住这个助记符, 可以节省很多时间。

另一方面, 如果涉及到多个 q 数, 则因子的顺序非常重要。例如, 以下关系式成立:

$$e^{\alpha\Omega} e^{\beta\Omega} = e^{(\alpha+\beta)\Omega}$$

可通过幂级数展开验证, 而以下关系式则不成立:

$$e^{\alpha\Omega} e^{\beta\theta} = e^{\alpha\Omega + \beta\theta}$$

或者

$$e^{\alpha\Omega} e^{\beta\theta} e^{-\alpha\Omega} = e^{\beta\theta}$$

除非 $[\Omega, \theta] = 0$ 。同样, 在对乘积求导数时, 我们需要使用链式法则:

$$\frac{d}{d\lambda} e^{\lambda\Omega} e^{\lambda\theta} = \Omega e^{\lambda\Omega} e^{\lambda\theta} + e^{\lambda\Omega} e^{\lambda\theta} \theta \quad (1.9.8)$$

我们可以自由地通过 $e^{\lambda\Omega}$ 移动 Ω 并将第一项写成

$$e^{\lambda\Omega}\Omega e^{\lambda\theta}$$

而不是

$$e^{\lambda\Omega}e^{\lambda\theta}\Omega$$

除非 $[\Omega, \theta] = 0$ 。

1.10 无限维的推广

在前面的讨论中，空间的维度 (n) 没有具体说明，但假定为某个有限数。现在我们考虑将前面的概念推广到无限维度的情况。

让我们首先熟悉无限维向量。考虑在某个区间 $[a, b]$ 上定义的函数。一个具体的例子是固定在 $x = 0$ 和 $x = L$ 处的绳子的位移 $f(x, t)$ (图 (1.6))。

假设我们想要将绳子在某个时间 t 的位移 $f(x)$ 传达给月球上的一个人。一种简单的方法是将区间 $0 - L$ 分为 20 等份，在 19 个点 $x = L/20, 2L/20, \dots, 19L/20$ 处测量位移 $f(x_i)$ ，然后通过无线传输这 19 个值。有了这些 $f(x_i)$ ，我们在月球上的朋友将能够重构出的绳子的近似图像如图 (1.7) 所示。

如果我们希望图像更准确，我们可以在更多的点上指定 $f(x)$ 的值。让我们用 $f_n(x)$ 表示与 $f(x)$ 在 n 个点上重合且在这些点之间为零的离散近似值。现在让我们将有序 n 元组 $\{f_n(x_1), f_n(x_2), \dots, f_n(x_n)\}$ 解释为向量空间 $\mathbb{V}^n(R)$ 中一个 ket $|f_n\rangle$ 的分量。

$$|f_n\rangle \leftrightarrow \begin{bmatrix} f_n(x_1) \\ f_n(x_2) \\ \vdots \\ f_n(x_n) \end{bmatrix} \quad (1.10.1)$$

图 1.6: 这根绳子固定在 $x = 0$ 和 $x = L$ 处。在其平面内自由振荡

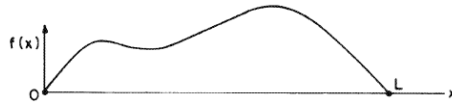


图 1.7: 被月球上的人重构出的绳子的图像。



在这个空间中，基向量是

$$|x_i\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{th place} \quad (1.10.2)$$

对应于在 $x = x_i$ 处为 1 其他地方为 0 的离散函数。这些基向量满足：

$$\langle x_i | x_j \rangle = \delta_{ij} \text{ (orthogonality)} \quad (1.10.3)$$

$$\sum_{i=1}^n |x_i\rangle \langle x_i| = I \text{ (completeness)} \quad (1.10.4)$$

试着想象一个包含 n 个相互垂直轴的空间，每个轴对应一个点 x_i 。沿着每个轴都有一个单位向量 $|x_i\rangle$ 。函数 $f_n(x)$ 由一个向量表示，它在第 i 个方向上的投影为 $f_n(x_i)$ 。

$$|f_n\rangle = \sum_{i=1}^n f_n(x_i) |x_i\rangle \quad (1.10.5)$$

对于每个可能的离散逼近 $g_n(x), f_n(x)$ 等，都存在对应的 ket $|g_n\rangle, |h_n\rangle$ 等，反之亦然。你应该确信，如果我们将向量的相加定义为分量的相加，将标量乘法定义为每个分量乘以标量，那么表示在 $x = 0, L$ 处消失且在中间的 n 个点上指定的离散函数的所有 kets 构成一个向量空间。

接下来，我们在这个空间中定义内积：

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f_n(x_i) g_n(x_i) \quad (1.10.6)$$

如果函数 $f_n(x)$ 和 $g_n(x)$ 的内积 $\langle f_n | g_n \rangle = 0$ 则它们被称为正交的。

现在我们暂时不考虑月球上的人，而是考虑绳子位移的最大具体化，通过在区间 $0 \sim L$ 的每个点上给出其值。在这种情况下， $f_\infty(x) \equiv f(x)$ 由一

系列有序的无穷多个数字来指定：每个点 x 上的 $f(x)$ 。现在每个函数都由一个无限维向量空间中的 $\text{ket } |f_\infty(x)\rangle$ 来表示，反之亦然。向量的加法和标量的乘法仍然与以前一样定义。然而，考虑内积。对于有限的 n ，内积被定义为

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

特别的

$$\langle f_n | f_n \rangle = \sum_{i=1}^n [f_n(x_i)]^2$$

如果我们现在让 n 趋向于无穷大，那么对于几乎任何函数来说，求和也会趋向于无穷大。我们需要的是将有限 n 的内积重新定义，以使得当 n 趋向于无穷大时，获得一个平滑的极限。自然的选择当然是：

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f_n(x_i) g_n(x_i) \Delta, \quad \Delta = L/(n+1) \quad (1.10.6')$$

这样，在 n 趋向于无穷大的极限情况下，根据积分的常规定义，我们得到：

$$\langle f | g \rangle = \int_0^L f(x) g(x) dx \quad (1.10.7)$$

$$\langle f | f \rangle = \int_0^L f^2(x) dx \quad (1.10.8)$$

如果我们希望不局限于该绳子的情况并且考虑在某个区间 $a \leq x \leq b$ 对 x 的复函数进行处理，我们唯一需要修改的是内积：在某个区间 $[a, b]$ 内

$$\langle f | g \rangle = \int_a^b f^*(x) g(x) dx \quad (1.10.9)$$

在这个空间中，基矢量是什么，它们如何归一化？我们知道每个点 x 都对应一个基矢量 $|x\rangle$ 。两个不同轴的正交性要求：

$$\langle x | x' \rangle = 0, \quad x \neq x' \quad (1.10.10)$$

那么 $x = x'$ 时怎么办？我们是否也应该像在有限维情况下一样要求 $\langle x | x \rangle = 1$ ？答案是否定的，最好的方法是推导出正确的归一化条件。我们从标记为连续量 x' 的 ket 的完备性关系公式(1.10.4)自然的推广开始：

$$\int_a^b |x'\rangle \langle x'| dx' = I \quad (1.10.11)$$

如往常一样, 等式要求单位算符保持每个 ket 不变。将等式(1.10.11)两边与任意 ket $|f\rangle$ 右乘, 与基态 bra $\langle x|$ 左乘:

$$\int_a^b \langle x|x'\rangle \langle x'|f\rangle dx' = \langle x|I|f\rangle = \langle x|f\rangle \quad (1.10.12)$$

现在 $\langle x|f\rangle$ 是 $|f\rangle$ 沿着基 ket $|x\rangle$ 的投影, 即 $f(x)$ 。同样的 $\langle x'|f\rangle$ 。令 $\langle x|x'\rangle$ 为某个未知函数 $\delta(x, x')$ 。因为 $\delta(x, x')$ 在 $x \neq x'$ 为零。所以可以将公式(1.10.12)中的积分限制在接近 $x = x'$ 的一个无穷小区域内:

$$\int_{x-\epsilon}^{x+\epsilon} \delta(x, x') f(x') dx = f(x) \quad (1.10.13)$$

在这个无穷小区域内, 对于任何合理光滑的 f , $f(x')$ 都可以近似为其在 $x' = x$ 时的值, 并从积分中提取出来:

$$f(x) \int_{x-\epsilon}^{x+\epsilon} \delta(x, x') dx' = f(x) \quad (1.10.14)$$

因此,

$$\int_{x-\epsilon}^{x+\epsilon} \delta(x, x') dx' = 1 \quad (1.10.15)$$

显然, 当 $x' = x$ 时, $\delta(x, x')$ 不能是有限的, 否则它在一个无穷小区域内的积分也将是无穷小的。事实上, $\delta(x, x')$ 应该是一种使其积分为 1 的方式的无穷大。由于 $\delta(x, x')$ 仅取决于 $x - x'$ 的差异, 让我们将其写为 $\delta(x - x')$ 。函数 $\delta(x - x')$ 具有以下性质

$$\begin{aligned} \delta(x - x') &= 0, \quad x \neq x' \\ \int_a^b \delta(x - x') dx' &= 1, \quad a < x < b \end{aligned} \quad (1.10.16)$$

称为 Dirac delta 函数, 并确定了基向量的归一化:

$$\langle x|x'\rangle = \delta(x - x') \quad (1.10.17)$$

每当基矢量由一个连续指标如 x 标记时, 它就会被用到。请注意, 它仅在积分的背景下被定义: 将 delta 函数 $\delta(x - x')$ 与任何平滑函数 $f(x')$ 进行积分得到的结果是 $f(x)$ 。有时人们将 delta 函数称为采样函数, 因为它在一个点上对函数 $f(x')$ 的值进行采样[‡]。

$$\int \delta(x - x') f(x') dx' = f(x) \quad (1.10.18)$$

[‡]如果积分的上下限不重要, 我们通常会省略它们。



图 1.8: (a) 当 $\Delta \rightarrow 0$ 时, 高斯函数 $g(x)$ 趋近于 delta 函数 (b) 当 $\Delta \rightarrow 0$ 时, 它的导数 $(dg/dx)(x-x')$ 趋近于 $\delta(x-x')$

Delta 函数不像我们之前见过的任何函数, 它的值要么是无穷大, 要么是零。因此, 将其视为更传统函数的极限是有用的。考虑一个高斯函数:

$$g_{\Delta}(x-x') = \frac{1}{(\pi\Delta^2)^{1/2}} \exp\left[-\frac{(x-x')^2}{\Delta^2}\right] \quad (1.10.19)$$

如图1.8a所示, 高斯函数以 $x' = x$ 为中心, 宽度为 Δ , 最大高度为 $(\pi\Delta^2)^{1/2}$, 并且单位面积, 与 Δ 无关。当 Δ 趋近于零时, g_{Δ} 越来越接近于 delta 函数的近似[§]。

从高斯模型可以明显看出, 函数是偶函数。可以通过以下方式验证:

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

因为 δ 函数是实函数。

接下来考虑一个比 delta 函数更奇特的对象: 它相对于第一个参数 x 的导数:

$$\delta'(x-x') = \frac{d}{dx}\delta(x-x') = -\frac{d}{dx'}\delta(x-x') \quad (1.10.20)$$

这个函数在积分下的作用是什么? 线索来自于高斯模型。将 $dg_{\Delta}(x-x')/dx = -dg_{\Delta}(x-x')/dx'$ 视为关于 x' 的函数。当 g_{Δ} 缩小时, 每个 $\pm\epsilon$ 处的凸起将在缩放因子的作用下变成 δ 函数。第一个凸起将取样 $-f(x-\epsilon)$, 第二个凸起将取样 $+f(x+\epsilon)$, 再次在缩放因子的作用下进行缩放使得

$$\int \delta'(x-x')f(x')dx' \propto f(x+\epsilon) - f(x-\epsilon) = 2\epsilon \left. \frac{df}{dx'} \right|_{x=x'}$$

[§]一个细致的观点, 在第 8 章之前不会涉及到您: 即使 Δ^2 是纯虚数, 例如等于 $i\beta^2$, 这个 delta function 的公式仍然是有效的。首先, 从公式(A.1.5)可以得知, g 具有单位面积。接下来考虑在包含 x 的 x' 的某个区域上, g 与 $f(x')$ 的乘积的积分。对大多数情况下, 由于 f 是平滑的, g 以 $\beta \rightarrow 0$ 的形式疯狂振荡, 因此积分大部分为零。然而, 在 $x = x'$ 时, g 的相位导数为零, 振荡被暂停。将 $f(x' = x)$ 移到积分外面, 我们得到了所期望的结果。

比例常数恰好是 $1/2\varepsilon$ ，因此

$$\int \delta'(x - x')f(x')dx' = \left. \frac{df}{dx'} \right|_{x'=x} = \frac{df(x)}{dx} \quad (1.10.21)$$

这个结果可以通过以下方式验证：

$$\begin{aligned} \int \delta'(x - x')f(x')dx &= \int \frac{d\delta(x - x')}{dx} f(x')dx' = \frac{d}{dx} \int \delta(x - x')f(x')dx' \\ &= \frac{df(x)}{dx} \end{aligned}$$

注意， $\delta(x - x')$ 是一个奇函数。这可以从图1.8b或等式(1.10.20)中清楚地看出。描述 δ' 函数作用的另一种等价方式是通过以下方程：

$$\delta'(x - x') = \delta(x - x') \frac{d}{dx'} \quad (1.10.22)$$

其中，我们明白两侧都是对 x' 积分的结果，微分算子作用于在被积函数中与 δ' 函数一起出现的函数上。在这种符号表示法下，我们可以描述 δ 函数的高阶导数所起的作用：

$$\frac{d''\delta(x - x')}{dx^n} = \delta(x - x') \frac{d''}{dx'^n} \quad (1.10.23)$$

现在，我们将发展出另一种 delta 函数的表示方法。根据基本的傅里叶分析，给定一个函数 $f(x)$ ，我们可以定义它的傅里叶变换：

$$f(k) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{-ikx} f(x)dx \quad (1.10.24)$$

以及它的反变换：

$$f(x') = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{ikx'} f(k)dk \quad (1.10.25)$$

将式(1.10.24)代入到式(1.10.25)中，我们得到：

$$f(x') = \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x'-x)} \right) f(x)dx$$

将这个结果与式(1.10.18)进行比较，我们可以看到：

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x'-x)} = \delta(x' - x) \quad (1.10.26)$$

练习 1.10.1 * 证明 $\delta(ax) = \delta(x)/|a|$ 。【考虑 $\int \delta(ax)d(ax)$ 。记住 $\delta(x) = \delta(-x)$ 。】

练习 1.10.2 * 证明

$$\delta(f(x)) = \sum_i \frac{\delta(x_i - x)}{|df/dx_i|}$$

其中 x_i 是 $f(x)$ 的零点。提示： $\delta(f(x))$ 在哪些地方会发散？在这些点附近用 Taylor 级数展开 $f(x)$ ，保留第一个非零项。

练习 1.10.3 * 考虑 theta 函数 $\theta(x - x')$ ，如果 $x - x'$ 是负值，则为零，如果 $x - x'$ 是正值，则为 1。证明 $\delta(x - x') = d/dx\theta(x - x')$ 。

无限维空间的算符

在熟悉了该函数空间的要素，即 kets $|f\rangle$ 和基向量 $|x\rangle$ 之后，让我们转向作用于它们的（线性）算符。考虑方程

$$\Omega|f\rangle = |\tilde{f}\rangle$$

由于 kets 与函数对应， Ω 将函数 $f(x)$ 转化为另一个函数 \tilde{f} 。现在，一个能够做到这件事的算符是我们熟悉的微分算符，对 $f(x)$ 作用，给出 $\tilde{f}(x) = df(x)/dx$ 。在函数空间中，我们可以描述该算符的作用为

$$D|f\rangle = |df/dx\rangle$$

其中， $|df/dx\rangle$ 是对应于函数 df/dx 的一个 ket。那么 D 在 $|x\rangle$ 基上的矩阵元是什么？为了找出，我们将上面方程的两边同时点积 $\langle x|$ ，

$$\langle x|D|f\rangle = \left\langle x \left| \frac{df}{dx} \right. \right\rangle = \frac{df(x)}{dx}$$

并在正确的位置插入单位矩阵的分解

$$\int \langle x|D|x'\rangle \langle x'|f\rangle dx' = \frac{df}{dx} \quad (1.10.27)$$

将此式与公式(1.10.21)进行比较，我们可以推断出

$$\langle x|D|x'\rangle = D_{xx'} = \delta'(x - x') = \delta(x - x') \frac{d}{dx'} \quad (1.10.28)$$

值得记住的是, $D_{xx} = \delta'(x - x')$ 应在第二个指标 (x') 上积分, 然后再从第一指标 (x) 上把导数提取出来。有些人更喜欢在第一指标上积分 $\delta'(x - x')$, 这样会提取出 $-df/dx'$ 。如果将 $D_{xx'}$ 看作是作用于向量 $|x\rangle$ 的分量 $f_{x'} = f(x')$ 的右矩阵, 那么我们的约定更为自然。因此熟悉的微分算符是一个具有上述元素的无限维矩阵。通常情况下, 我们不会将 D 看作矩阵, 因为当矩阵作用于向量时, 通常需要对一个公共指标进行求和。事实上, 公式 (1.10.27) 就包含了对指标 x_1 进行求和。然而, 如果我们将 $D_{xx'}$ 的值输入该方程, δ 函数将使积分变得平凡。

$$\int \delta(x - x') \frac{d}{dx'} f(x') dx' = \left. \frac{df}{dx'} \right|_{x'=x} = \frac{df}{dx}$$

因此 D 的作用就是在不需要对公共指标进行求和的情况下将 d/dx 应用于 $f(x)$ 。虽然我们最终也会放弃在公共指标上的积分, 但在一段时间内我们会继续使用它, 以提醒我们 D , 像所有的线性算符一样, 都是矩阵。

现在让我们来看看 D 是否是厄米算符并研究它的特征值问题。如果 D 是厄米的, 我们应有

$$D_{xx'} = D_{x'x}^*$$

但事实并非如此:

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

而

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

但我们可以通过与一个纯虚数相乘来轻松地将 D 转换为 Hermitian 矩阵。考虑

$$K = -iD$$

满足

$$K_{x'x}^* = [-i\delta'(x' - x)]^* = +i\delta'(x' - x) = -i\delta'(x - x') = K_{xx'}$$

结果证明尽管存在上述情况, 但算符 K 并不保证是 Hermitian 的, 如下分析所示。假设 $|k\rangle$ 和 $|g\rangle$ 是函数空间中的两个 kets, 在 X 基下它们的映像分别是区间 $a - b$ 中的两个函数 $f(x)$ 和 $g(x)$ 。如果 K 是 Hermitian 的, 那么它也必须满足

$$\langle g|K|f\rangle = \langle g|Kf\rangle = \langle Kf|g\rangle^* = \langle f|K^\dagger|g\rangle^* = \langle f|K|g\rangle^*$$

因此我们得到

$$\begin{aligned} & \int_a^b \int_a^b \langle g|x \rangle \langle x|K|x' \rangle \langle x'|f \rangle dx dx' \\ & \stackrel{?}{=} \left(\int_a^b \int_a^b \langle f|x \rangle \langle x|K|x' \rangle \langle x'|g \rangle dx dx' \right)^* \\ & \int_a^b g^*(x) \left[-\frac{idf(x)}{dx} \right] dx \stackrel{?}{=} \left\{ \int_a^b f^*(x) \left[-\frac{idg(x)}{dx} \right] dx \right\}^* = i \int_a^b \frac{dg^*}{dx} f(x) dx \end{aligned}$$

对左边进行分部积分得到

$$-ig^*(x)f(x) \Big|_a^b + i \int_a^b \frac{dg^*(x)}{dx} f(x) dx$$

所以当且仅当表面项 (surface term) 为零时 K 才是厄米的:

$$-ig^*(x)f(x) \Big|_a^b = 0 \quad (1.10.29)$$

与有限维情况不同, $K_{xx} = K_{x'x}^*$ 不是 K 是厄米的充分条件。我们还需要观察函数在端点 a 和 b 处的行为。因此, 如果空间由满足式(1.10.29)的函数组成的, 那么 K 才是厄米的。满足此条件的组函数是在 $x=0, L$ 处夹杂的弦的所有可能的形状 $f(x)$, 满足这个条件的一组函数是在 $x=0$ 和 $x=L$ 被固定的 $f(x)$, 因为 $f(x)$ 在端点处为零。但是, 条件 (1.10.29) 也可以通过另一种方式满足。考虑在我们自己的三维空间中以 r, θ 和 ϕ (ϕ 是绕 z 轴测量的角度) 为参数的函数。让我们要求这些函数是单值的。特别地, 如果我们从某一点开始, 沿着 z 轴旋转一周, 回到原点, 函数必须取回其原始值, 即

$$f(\phi) = f(\phi + 2\pi)$$

在这种周期函数的空间中, $K = -id/d\phi$ 是一个厄米算符。表面项消失是因为一个极限的贡献抵消了另一个极限的贡献:

$$-ig^*(\phi)f(\phi) \Big|_0^{2\pi} = -i[g^*(2\pi)f(2\pi) - g^*(0)f(0)] = 0$$

在量子力学的研究中, 我们对定义在整个区间 $-\infty \leq x \leq +\infty$ 上的函数感兴趣。它们可以分为两类, 一类是当 $|x| \rightarrow \infty$ 时趋于零的函数, 另一类是不趋于零的函数, 后者的行为类似于 e^{ikx} , 其中 k 是一个实参数, 用来标记这些函数。很明显, 当 $K = -id/dx$ 夹在两个第一类函数之间或夹

在一个第一类函数和一个第二类函数之间时，它是厄米的，因为在任一情况下，表面项都会消失。当 K 夹在两个第二类函数之间时，其厄米性取决于以下条件是否成立：

$$e^{ikx}e^{-ik'x} \Big|_{-\infty}^{\infty} \stackrel{?}{=} 0$$

如果 $k = k'$ ，则一个端点的贡献会抵消另一个端点的贡献。如果 $k \neq k'$ 那么答案是不明确的，因为 $e^{i(k-k')x}$ 是振荡的，而不是在 $|x| \rightarrow \infty$ 时趋于一个极限。现在，存在一种定义这种不能决定的函数的极限的方法： $|x| \rightarrow \infty$ 的极限被定义为在一个大区间上的平均值。根据这个规定，我们可以在 $|x| \rightarrow \infty$ 时得到：

$$\lim_{x \rightarrow \infty} e^{ikx}e^{-ik'x} = \lim_{\substack{L \rightarrow \infty \\ \Delta \rightarrow \infty}} \frac{1}{\Delta} \int_0^{L+\Delta} e^{i(k-k')x} dx = 0 \quad \text{if } k \neq k'$$

因此，在这个空间中， K 是厄米的。

现在我们来查看 K 的特征值问题。这个任务似乎非常艰巨，因为我们现在需要找到一个无穷阶特征多项式的根，并获得相应的特征向量。但事实上，这个问题非常简单，你可能以前已经做过几次而没有得到应有的赞赏。让我们从以下方程开始：

$$K|k\rangle = k|k\rangle \quad (1.10.30)$$

接下来标准步骤

$$\begin{aligned} \langle x|K|k\rangle &= k\langle x|k\rangle \\ \int \langle x|K|x'\rangle \langle x'|k\rangle dx' &= k\psi_k(x) \\ -i\frac{d}{dx}\psi_k(x) &= k\psi_k(x) \end{aligned} \quad (1.10.31)$$

其中根据定义， $\psi_k(x) = \langle x|k\rangle$ 。如果我们在 X 基矢中进行立即替换 $K = -id/dx$ ，上述方程可以直接写成这种形式。从现在开始，除非有充分的理由不这样做，否则我们将使用这种捷径。该方程的解简单地为：

$$\psi_k(x) = Ae^{ikx} \quad (1.10.32)$$

其中 A 是整体缩放因子，由特征值问题未指定。因此， K 的特征值问题完全解决：任何实数 k 都是一个特征值，相应的特征函数由 Ae^{ikx} 给出。通常情况下，缩放的自由度将用于归一化解。我们选择 A 为 $(1/2\pi)^{-1/2}$ ，以使得

$$|k\rangle \leftrightarrow \frac{1}{(2\pi)^{1/2}} e^{ikx}$$

和

$$\langle k|k'\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle x|k'\rangle dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k-k')x} dx = \delta(k-k') \quad (1.10.33)$$

(由于 $\langle k|k\rangle$ 是无穷大的, 任何 A 的选择都不能将 $|k\rangle$ 归一化为单位。当特征值谱是连续的时, 使用 delta 函数归一化是自然的。) 细心的读者可能会有一个问题: “为什么假设特征值 k 是实数? 很明显, 当 $k = k_1 + ik_2$ 时, 函数 Ae^{ikx} 也满足等式(1.10.31)。” 答案是, 是的, K 存在具有复特征值的特征函数。然而, 如果我们的空间包含这样的函数, K 必须被归类为非厄米算符。(由于 e^{ikx} 在 x 趋向于 $+\infty$ 或 $-\infty$ 时呈指数增长, 取决于虚部 k_2 的符号, 表面项不再消失。) 通过限制我们只考虑实数 k , 我们将自己限制在了我们所称的物理希尔伯特空间中, 这在量子力学中是很重要的。这个空间被定义为可以归一化为单位或狄拉克 δ 数的函数的空间, 在量子力学中起着核心作用。(我们使用“物理”这个修饰词来区分它与数学家定义的只包含可以归一化为单位的合适向量的希尔伯特空间。不合适向量在量子理论中的作用稍后将变得清晰。)

我们将假设一个在有限维情况下成立的定理, 即 Hermitian 算符的特征函数构成完备基, 在 Hilbert 空间中仍然成立[‡]。(无限维空间的问题在于, 即使你有无限个正交归一的特征向量, 你也不能确定是否已经拥有它们全部, 因为添加或减去一些仍然会使你拥有无限多个特征向量。)

由于 K 是一个 Hermitian 算符, 用 X 基展开的函数 $f(x) = |x\rangle$ 的分量也必须有一个 K 基的展开。为了找到这些分量, 我们从一个 ket $|f\rangle$ 开始, 进行以下操作:

$$f(k) = \langle k|f\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle x|f\rangle dx = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \quad (1.10.34)$$

返回到 X 基的过程如下:

$$f(x) = \langle x|f\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle k|f\rangle dk = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{ikx} f(k) dk \quad (1.10.35)$$

因此, 熟悉的傅里叶变换只是从一个完备基 $|x\rangle$ 到另一个基 $|k\rangle$ 的过渡。任何一个基都可以用来展开属于希尔伯特空间的函数。在 K 基下, K 的矩阵元非常简单:

$$\langle k|K|k'\rangle = k'\langle k|k'\rangle = k'\delta(k-k') \quad (1.10.36)$$

[‡]从现在开始, 我们将省略修饰语“物理的”。

现在, 我们知道 K 基是由 Hermitian 算符 K 生成的。那么哪个算符负责 X 基的正交归一性呢? 让我们称其为算符 X 。Ket $|x\rangle$ 是它的特征向量, 具有特征值 x :

$$X|x\rangle = x|x\rangle \quad (1.10.37)$$

在 X 基中, 它的矩阵元为:

$$\langle x'|X|x\rangle = x\delta(x' - x) \quad (1.10.38)$$

为了找到它在函数上的作用, 让我们从以下方程开始:

$$X|f\rangle = |\tilde{f}\rangle$$

并按照常规方法进行推导:

$$\begin{aligned} \langle x|X|f\rangle &= \int \langle x|X|x'\rangle \langle x'|f\rangle dx' = xf(x) = \langle x|\tilde{f}\rangle = \tilde{f}(x) \\ \therefore \quad \tilde{f}(x) &= xf(x) \end{aligned}$$

因此, X 算符的作用是将 $f(x)$ 乘以 x 。和 K 算符的情况类似, 我们通常省略共同指标的积分, 因为它被 delta 函数消除了。我们可以总结 X 在希尔伯特空间中的作用为:

$$X|f(x)\rangle = |xf(x)\rangle \quad (1.10.39)$$

其中, 像往常一样, $|xf(x)\rangle$ 是对应于函数 $xf(x)$ 的一个 ket。 X 算符和 K 算符之间存在着良好的互易性, 如果我们计算 X 在 K 基下的矩阵元:

$$\begin{aligned} \langle k|X|k'\rangle &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} x e^{ik'x} dx \\ &= i \frac{d}{dk} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k'-k)x} dx \right) = i\delta'(k - k')^\ddagger \end{aligned}$$

因此, 如果 $|g(k)\rangle$ 是一个 ket 在 k 基下的映像是 $g(k)$, 那么

$$X|g(k)\rangle = \left| \frac{dg(k)}{dk} \right\rangle \quad (1.10.40)$$

总结一下, 在 X 基下, X 作用为 x , K 作用为 $-id/dx$ (对于函数 $f(x)$), 而在 K 基下, K 作用为 k , X 作用为 id/dk (对于 $f(k)$)。具有这种相互关

[‡]在最后一步中我们利用了 $\delta(k' - k) = \delta(k - k')$ 这个事实。

系的算符被称为互为共轭。X 和 K 这两个共轭算符是不对易的。它们的对易子可以计算如下。让我们用两种可能的顺序在某个态矢 $|f\rangle$ 上作用 X 和 K，并跟踪在 X 基下的作用：

$$\begin{aligned} X|f\rangle &\rightarrow xf(x) \\ K|f\rangle &\rightarrow -i\frac{df(x)}{dx} \end{aligned}$$

所以

$$\begin{aligned} XK|f\rangle &\rightarrow -ix\frac{df(x)}{dx} \\ KX|f\rangle &\rightarrow -i\frac{d}{dx}xf(x) \end{aligned}$$

因此

$$[X, K]|f\rangle \rightarrow -ix\frac{df}{dx} + ix\frac{df}{dx} + if = if \rightarrow iI|f\rangle$$

由于 $|f\rangle$ 是任意的态矢量，我们现在得到了所期望的结果：

$$[X, K] = iI \quad (1.10.41)$$

这就结束了我们对 Hilbert 空间的讨论，除了最后一个例子。虽然在这个空间中还可以研究许多其他算符，但我们仅限于 X 和 K，因为几乎所有我们在量子力学中需要的算符都是 X 和 $P = n\hbar$ 的函数，其中 n 是待定的常数，稍后将定义。

例 1.10.1 在 Hilbert 空间中的一个正态模问题。考虑一个长度为 L 的绳子，在其两端 $x=0$ 和 $x=L$ 处被固定。位移 $\psi(x, t)$ 满足以下微分方程：

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial^2 \psi}{\partial x^2} \quad (1.10.42)$$

假设在 $t=0$ 时刻，位移为 $\psi(x, 0)$ ，速度为 $\dot{\psi}(x, 0) = 0$ ，我们希望确定绳子的时间演化。

除了维度的变化外，该问题与第 1.8 节末尾遇到的两个耦合质点问题完全相同（参见式(1.8.26)）。建议在继续之前，先回顾一下那个例子，以便刷新记忆。

我们首先将 $\psi(x, t)$ 识别为 Hilbert 空间中向量 $|\psi(t)\rangle$ 的分量，其元素与可能的位移 $|\psi\rangle$ 对应，即在区间 $0 \leq x \leq L$ 内连续且在端点处为零的函数。你可以验证这些函数确实构成一个向量空间。

在等式(1.8.26)中, 算符 n 的类比是算符 $\partial^2/\partial x^2$ 。我们认识到这实际上是算符 $K \leftrightarrow i\partial/\partial x$ 的平方的负数。由于 K 作用于满足边界条件 $\psi(0) = \psi(L) = 0$ 的空间, 它是厄米的, K^2 也是。等式(1.10.42)可以抽象表示为:

$$|\ddot{\psi}(t)\rangle = -K^2|\psi(t)\rangle \quad (1.10.43)$$

我们通过遵循在示例 (1.8.6) 中发展的算法来解决这个初值问题:

步骤 (1): 解 $-K^2$ 的特征值问题。

步骤 (2): 利用特征向量和特征值构造传播子 $U(t)$ 。

步骤 (3):

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \quad (1.10.44)$$

我们要解的方程是

$$K^2|\psi\rangle = k^2|\psi\rangle \quad (1.10.45)$$

在 X 基下, 这变为

$$-\frac{d^2}{dx^2}\psi_k(x) = k^2\psi_k(x) \quad (1.10.46)$$

其一般解为

$$\psi_k(x) = A \cos kx + B \sin kx \quad (1.10.47)$$

其中 A 和 B 是任意常数。然而, 并非所有这些解都属于我们考虑的 *Hilbert* 空间。我们只希望选择那些在 $x=0$ 和 $x=L$ 处为零的解。在 $x=0$ 处我们有

$$\psi_k(0) = 0 = A \quad (1.10.48a)$$

而在 $x=L$ 处我们有

$$0 = B \sin kL \quad (1.10.48b)$$

如果我们不想要平凡解 ($A = B = 0$), 我们必须要求

$$\sin kL = 0, \quad kL = m\pi, \quad m = 1, 2, 3, \dots \quad (1.10.49)$$

我们不考虑负的 m , 因为它不会导致任何其他的线性独立解 ($[\sin(-x) = -\sin x]$)。允许的特征向量因此形成一个由整数 m 标记的离散集合:

$$\psi_m(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{m\pi x}{L}\right) \quad (1.10.50)$$

其中我们选择了 $B(2/L)^{1/2}$ ，以使得

$$\int_0^L \psi_m(x) \psi_{m'}(x) dx = \delta_{mm'} \quad (1.10.51)$$

让我们将每个由整数 m 标记的解与一个抽象的态 $|m\rangle$ 关联起来：

$$|m\rangle \xrightarrow{x \text{ basis}} (2/L)^{1/2} \sin\left(\frac{m\pi x}{L}\right) \quad (1.10.52)$$

如果我们将 $|\psi(t)\rangle$ 投影到 $|m\rangle$ 基上，在这个基下 K 是对角的，具有特征值 $(m/\pi L)^2$ ，那么分量 $\langle m|\psi(t)\rangle$ 将满足解耦的方程

$$\frac{d^2}{dt^2} \langle m|\psi(t)\rangle = -\left(\frac{m^2\pi^2}{L^2}\right) \langle m|\psi(t)\rangle, \quad m = 1, 2, \dots \quad (1.10.53)$$

类比于等式(1.8.34)。这些方程可以很容易地求解（在初始速度为零的条件下），得到

$$\langle m|\psi(t)\rangle = \langle m|\psi(0)\rangle \cos\left(\frac{m\pi t}{L}\right) \quad (1.10.54)$$

因此

$$\begin{aligned} |\psi(t)\rangle &= \sum_{m=1}^{\infty} |m\rangle \langle m|\psi(t)\rangle \\ &= \sum_{m=1}^{\infty} |m\rangle \langle m|\psi(0)\rangle \cos \omega_m t, \quad \omega_m = \frac{m\pi}{L} \end{aligned} \quad (1.10.55)$$

或者

$$U(t) = \sum_{m=1}^{\infty} |m\rangle \langle m| \cos \omega_m t, \quad \omega_m = \frac{m\pi}{L} \quad (1.10.56)$$

传播子方程

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

在 $|x\rangle$ 基中变为

$$\begin{aligned} \langle x|\psi(t)\rangle &= \psi(x, t) \\ &= \langle x|U(t)|\psi(0)\rangle \\ &= \int_0^L \langle x|U(t)|x'\rangle \langle x'|\psi(0)\rangle dx' \end{aligned} \quad (1.10.57)$$

由等式(1.10.56)可知

$$\begin{aligned} \langle x|U(t)|x'\rangle &= \sum_m \langle x|m\rangle \langle m|x'\rangle \cos \omega_m t \\ &= \sum_m \left(\frac{2}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{m\pi x'}{L}\right) \cos \omega_m t \end{aligned} \quad (1.10.58)$$

因此, 对于任意的 $\psi(x', 0)$, 我们可以通过在等式 (1.10.57) 中进行积分, 并使用等式 (1.10.58) 中的 $\langle x|U(t)|x'\rangle$, 得到 $\psi(x, t)$ 。如果传播子的语言对你来说太抽象, 我们可以从方程 (1.10.55) 开始。两边与 $\langle x|$ 做内积, 我们得到

$$\begin{aligned}\psi(x, t) &= \sum_{m=1}^{\infty} \langle x|m\rangle \langle m|\psi(0)\rangle \cos \omega_m t \\ &= \sum_{m=1}^{\infty} \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{m\pi x}{L}\right) \cos \omega_m t \langle m|\psi(0)\rangle\end{aligned}\quad (1.10.59)$$

给定 $|\psi(0)\rangle$, 我们需要计算

$$\langle m|\psi(0)\rangle = \left(\frac{2}{L}\right)^{1/2} \int_0^L \sin\left(\frac{m\pi x}{L}\right) \psi(x, 0) dx$$

通常我们会发现系数 $\langle m|\psi(0)\rangle$ 随着 m 的增大迅速减小, 因此只需取前几个主要项即可得到一个良好的近似。□

例 1.10.2 在 $t = 0$ 时一根绳子的位移如下所示:

$$\begin{aligned}\psi(x, 0) &= \frac{2xh}{L}, & 0 \leq x \leq \frac{L}{2} \\ &= \frac{2h}{L}(L - x), & \frac{L}{2} \leq x \leq L\end{aligned}$$

证明

$$\psi(x, t) = \sum_{m=1}^{\infty} \sin\left(\frac{m\pi x}{L}\right) \cos \omega_m t \cdot \left(\frac{8h}{\pi^2 m^2}\right) \sin\left(\frac{\pi m}{2}\right)$$

第二章 回顾经典力学

在本章中，我们将从牛顿定律出发，发展出拉格朗日和哈密顿力学的表述形式。这些随后的力学重新表述带来了许多优雅和计算上的便利。然而，我们对它们的主要兴趣源自于它们作为理解量子力学的理想跳板。从拉格朗日表述到量子力学的过渡是由费曼通过他的路径积分形式主义完成的。而更常见的通向量子力学的路径（我们大部分会遵循）是以哈密顿力学为起点，主要由薛定谔、海森堡、狄拉克和玻恩等人发现。

需要强调的是，这将会变得很明显，这三种力学表述本质上是相同的理论，因为它们的适用范围和预测是相同的。然而，在特定的背景下，根据概念、计算或者纯粹的美学原因，其中一种表述可能更具吸引力。

2.1 最小作用原理和拉格朗日力学

让我们以质量为 m 的质点沿着 x 轴在势能 $V(x)$ 下运动作为我们牛顿方案的原型。根据牛顿第二定律，

$$m \frac{d^2 x}{dt^2} = - \frac{dV}{dx} \quad (2.1.1)$$

如果我们已知初始状态变量，即位置 $x(t)$ 和速度 $\dot{x}(t_i)$ ，我们可以按照以下方法计算经典轨迹 $X_{cl}(t)$ 。利用初始速度和加速度（从方程 (2.1.1) 中获得），我们计算在时间 $t_i + \Delta t$ 处的位置和速度。例如，

$$x_{cl}(t_i + \Delta t) = x(t_i) + \dot{x}(t_i)\Delta t$$

在更新了时间 $t_i + \Delta t$ 的状态变量之后，我们可以再次重复这个过程，向前推进到 $t_i + 2\Delta t$ ，以此类推。

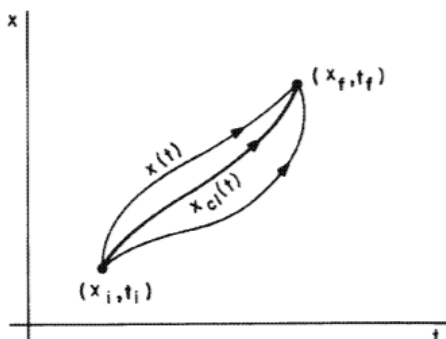


图 2.1: 拉格朗日的力学形式主要关注实际路径 $x_{cl}(t)$ 与连接两个端点 (x_i, t_i) 和 (x_f, t_f) 的所有可能路径之间的差异。

由于运动方程在时间上是二阶的，因此需要两个数据 $x(t_i)$ 和 $\dot{x}(t_i)$ 来确定唯一的 $x_{cl}(t)$ 。同样的，我们也可以指定轨迹上的两个时空点 (x_i, t_i) 和 (x_f, t_f) 来完成相同的任务，这种方式在某些情况下会被采用。

以上方法可以很容易地推广到包含多个粒子和多个维度的情况。如果使用 n 笛卡尔坐标 (x_1, x_2, \dots, x_n) 来指定粒子的位置，系统的空间构型可以被视为 n 维的构型空间中的一个点。（即使座标不是笛卡尔坐标，术语“位形空间”仍然适用。）代表点的运动由以下方程描述：

$$m_j \frac{d^2 x_j}{dt^2} = - \frac{\partial V}{\partial x_j} \quad (2.1.2)$$

其中 m_j 代表坐标为 x_j 的粒子的质量。这些方程可以像之前一样逐步积分，以确定轨迹。

在拉格朗日形式主义中，一个在势能 $V(x)$ 下的单粒子问题被以不同的方式提出：假设在时刻 t_i 和 t_f 时，粒子分别位于 x_i 和 x_f ，那么是什么使得实际轨迹 $x_{cl}(t)$ 与连接这些点的所有其他轨迹或路径不同呢？（见图（2.1））

因此，拉格朗日方法是一种整体性的方法，它试图一次性确定整个轨迹 $X_{cl}(t)$ ，与牛顿方案的局部方法形成对比，后者关注粒子在下一个无穷小时间间隔内会发生什么。

对上述问题的答案可以分为三个部分：

(1) 定义一个称为拉格朗日量的函数，由 $\mathcal{L} = T - V$ 给出，其中 T 和 V 分别是粒子的动能和势能。因此 $\mathcal{L} = \mathcal{L}(x, \dot{x}, t)$ 。如果粒子处于外部的时变场中，就可能出现明确的时间依赖性。然而，在这里我们假设不存在这种时间依赖性。

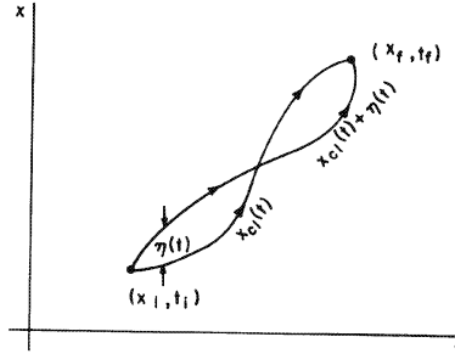
(2) 对于连接 (x_i, t_i) 和 (x_f, t_f) 的每条路径 $x(t)$ ，计算由以下定义的作

用量 $S[x(t)]$:

$$S[x(t)] = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x}) dt \quad (2.1.3)$$

我们使用方括号来包围 S 的参数, 以提醒我们 S 函数依赖于整个路径或函

图 2.2: 如果 $x_{cl}(t)$ 最小化了 S , 那么当我们考虑任意附近路径 $x_{cl}(t) + \eta(t)$ 时 $\delta S = 0$ 。



数 $x(t)$, 而不仅仅是某个时间 t 处的 x 值。我们称 S 为泛函, 表示它是一个函数的函数。

(3) 经典路径是使得 S 取得极小值的路径。(实际上, 我们只要求它是一个极值。然而, 通常将这个条件称为最小作用量原理。)

我们现在将验证这个原理是如何重新推导出牛顿第二定律的。

第一步是要意识到一个泛函 $S[x(t)]$ 实际上是一个随着 $n \rightarrow \infty$ 而变成 n 个变量的函数。换句话说, 函数 $x(t)$ 仅仅指定了一个无穷多个值 $x(t_i), \dots, x(t), \dots, x(t_f)$, 对于时间区间 $t_i \leq t \leq t_f$ 中的每个瞬时时刻 t , 而 S 是这些变量的函数。为了找到它的最小值, 我们只需要将有限 n 情况的过程推广即可。让我们回顾一下, 如果 $f = f(x_1, \dots, x_n) = f(x)$, 那么最小值 x^0 的特点是, 如果我们沿着任意方向以一个小量 η 远离它, 函数 f 的一阶变化 $\delta f^{(1)}$ 为零。也就是说, 如果我们进行泰勒展开:

$$f(\mathbf{x}^0 + \boldsymbol{\eta}) = f(\mathbf{x}^0) + \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}^0} \eta_i + \text{higher-order terms in } \boldsymbol{\eta} \quad (2.1.4)$$

那么

$$\delta f^{(1)} \equiv \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}^0} \eta_i = 0 \quad (2.1.5)$$

根据这个条件, 我们可以推导出一个等价且可能更为熟悉的最小值条件的表达式: 在 x^0 处, 每个一阶偏导数都为零。为了证明这一点, 对于偏

导数 $\partial f/\partial x$ ，我们只需要选择 η 沿着第 i 个方向。因此，我们有：

$$\left. \frac{\partial f}{\partial x_i} \right|_{x_0}, \quad i = 1, \dots, n \quad (2.1.6)$$

现在让我们模仿这个过程来处理作用量 S 。设 $x_{cl}(t)$ 为最小作用路径， $x_{cl}(t) + \eta(t)$ 为“附近”路径（见图2.2）。所有路径在 t_i 和 t_f 处重合的要求意味着：

$$\eta(t_i) = \eta(t_f) = 0 \quad (2.1.7)$$

现在

$$\begin{aligned} S[x_{cl}(t) + \eta(t)] &= \int_{t_i}^{t_f} \mathcal{L}(x_{cl}(t) + \eta(t); \dot{x}_{cl}(t) + \dot{\eta}(t)) dt \\ &= \int_{t_i}^{t_f} \left[\mathcal{L}(x_{cl}(t), \dot{x}_{cl}(t)) + \left. \frac{\partial \mathcal{L}}{\partial x(t)} \right|_{x_{cl}} \cdot \eta(t) \right. \\ &\quad \left. + \left. \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right|_{x_{cl}} \cdot \dot{\eta}(t) + \dots \right] dt \\ &= S[x_{cl}(t)] + \delta S^{(1)} + \text{higher-order terms} \end{aligned}$$

我们设定 $\delta S^{(1)} = 0$ ，类比于有限变量情况：

$$0 = \delta S^{(1)} = \int_{t_i}^{t_f} \left[\left. \frac{\partial \mathcal{L}}{\partial x(t)} \right|_{x_a} \cdot \eta(t) + \left. \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right|_{x_a} \cdot \dot{\eta}(t) \right] dt$$

如果我们对第二项进行分部积分，它将变为：

$$\left. \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right|_{x_{cl}} \cdot \eta(t) \Big|_{t_i}^{t_f} - \int_{t_i}^{t_f} \left[\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right]_{x_{cl}} \cdot \eta(t) dt$$

由于等式(2.1.7)第一项为零所以有

$$0 = \delta S^{(1)} = \int_{t_i}^{t_f} \left[\frac{\partial \mathcal{L}}{\partial x(t)} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right]_{x_{cl}} \cdot \eta(t) dt \quad (2.1.8)$$

注意条件 $\delta S^{(1)} = 0$ 意味着 S 被极值化，而不一定是最小化。然而，我们传统上将继续将这个极值称为最小值。这个等式类似于等式(2.1.5)：离散变量 η_i 被 $\eta(t)$ 取代，对 i 的求和被对 t 的积分取代，而 $\partial f/\partial x_i$ 被替换为

$$\frac{\partial \mathcal{L}}{\partial x(t)} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}(t)}$$

这里有两个项起到了 $\partial f/\partial x_i$ 的作用, 因为 \mathcal{L} (或等价地 S) 对 $x(t)$ 既有显式的依赖 (通过 \dot{x} 项), 也有隐式的依赖。由于 $\eta(t)$ 是任意的, 我们可以从中取出等式 (2.1.6) 的类似等式:

$$\left\{ \frac{\partial \mathcal{L}}{\partial x(t)} - \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \right] \right\}_{x_{cl}(t)} = 0 \quad \text{for } t_i \leq t \leq t_f \quad (2.1.9)$$

为了推导出这个结果对于某个特定的时间 t_0 , 我们只需选择一个在除了 t_0 附近的一个无穷小区域以外处处为零的 $\eta(t)$ 。

方程(2.1.9) 是著名的欧拉-拉格朗日方程。如果我们将 $\mathcal{L} = T - V$, $T = \frac{1}{2}m\dot{x}^2$, $V = V(x)$ 代入其中, 我们可以得到:

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial T}{\partial \dot{x}} = m\dot{x}$$

和

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial V}{\partial x}$$

因此, 欧拉-拉格朗日方程简化为:

$$\frac{d}{dt}(m\dot{x}) = -\frac{\partial V}{\partial x}$$

这正是牛顿第二定律, 即方程 (2.1.1)。

如果我们考虑由 n 个笛卡尔坐标描述的系统, 同样的推导过程会得到:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = \frac{\partial \mathcal{L}}{\partial x_i} \quad (i = 1, \dots, n) \quad (2.1.10)$$

现在,

$$T = \frac{1}{2} \sum_{i=1}^n m_i (\dot{x}_i)^2$$

和

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

因此, 等式(2.1.10)变为:

$$\frac{d}{dt}(m_i \dot{x}_i) = -\frac{\partial V}{\partial x_i} \quad (2.1.11)$$

这与方程(2.1.2)完全相同。因此, 如果我们选择 $\mathcal{L} = T - V$, 最小作用量原理确实能够重新推导出牛顿力学。

需要注意的是在上述证明中我们假设了 V 不依赖于速度。这个假设排除了磁场 \mathbf{B} 对运动电荷的力, 因为 $\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}$, 其中 q 是粒子的电荷,

$v = \dot{r}$ 是其速度。我们将很快展示，即使是这种力也可以纳入到拉格朗日形式中，也就是说，在使用方程 (2.1.10) 时我们可以找到一个 \mathcal{L} ，它产生正确的力学规律。但是，这个 \mathcal{L} 不再具有 $T - V$ 的形式。因此，我们摆脱了 $\mathcal{L} = T - V$ 的观念，并将 \mathcal{L} 视为一些函数 $\mathcal{L}(x_i, \dot{x}_i)$ ，当输入欧拉-拉格朗日方程时，它能够产生正确的牛顿动力学。对于那些想知道为什么要处理拉格朗日量，当它最终只能产生牛顿力学的读者，我给出一些它的主要优点，除了它与量子力学的紧密联系。这些优点将通过一个例子来说明。

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{\partial \mathcal{L}}{\partial q_i} \quad (2.1.12)$$

有两种方法可以验证在任意坐标变换下，欧拉-拉格朗日方程 (2.1.10) 保持形式不变。一种方法是通过在方程 (2.1.10) 中进行变量替换，观察到由新坐标取代 x 后得到的方程仍然具有相同的形式；另一种方法是回顾我们关于最小作用量条件的推导，可以看到在推导过程中并没有假设坐标为笛卡尔坐标。当然，在下一阶段证明欧拉-拉格朗日方程等效于牛顿方程时，使用了笛卡尔坐标，因为在这些坐标下，动能 T 和牛顿方程具有简单的形式。但是一旦我们发现最小作用量原理能够产生正确的动力学结果，我们就可以完全忽略牛顿定律，使用方程 (2.1.11) 作为运动方程。重点强调的是，这些方程表达了最小作用量条件，在任意坐标变换下保持形式不变。这种形式不变性与牛顿方程 (2.1.2) 形成对比，牛顿方程假设 X_i 是笛卡尔坐标。如果将 x_i 替换为另一组非笛卡尔坐标 q_i ，方程 (2.1.2) 将具有不同的形式（参见本节末尾的示例 2.1.1）。

方程 (2.1.11) 可以通过定义一个量来与牛顿第二定律相似：

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (2.1.13)$$

称为共轭于 q_i 的广义动量的正则动量，以及称为共轭于 q_i 的广义力的量

$$F_i = \frac{\partial \mathcal{L}}{\partial q_i} \quad (2.1.14)$$

虽然正则动量的变化率等于广义力，但必须记住， p_i 不总是线性动量（质量乘以速度或“ mv ”动量）， F_i 也不总是力（具有质量乘以加速度的维度）。例如，如果 q_i 是一个角度 θ ， p_i 将是角动量， F_i 将是扭矩。(3) 在这个形式体系中很容易得到守恒定律。假设拉格朗日量依赖于某个速度 q_i ，但不依赖于相应的坐标 q_i 。那么后者被称为循环坐标。据此可以得出相应的 p_i 是守恒

的:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{dp_i}{dt} = \frac{\partial \mathcal{L}}{\partial q_i} = 0 \quad (2.1.15)$$

虽然牛顿第二定律方程(2.1.2)也告诉我们, 如果笛卡尔坐标 x_i 是循环坐标, 相应的动量 $m_i \dot{x}_i$ 是守恒的, 但方程(2.1.14)更为一般。例如, 考虑一个仅依赖于二维空间中的 $\rho = (x^2 + y^2)^{1/2}$ 而不依赖于极角 ϕ 的势能 $V(x, y)$, 即 $V(\rho, \phi) = V(\rho)$ 。由此可知 ϕ 是循环坐标, 因为 T 仅依赖于 $\dot{\phi}$ (见下面的示例 2.1.1)。因此, $\partial \mathcal{L} / \partial \dot{\phi} = p_\phi$ 是守恒的。相比之下, 笛卡尔坐标方程(2.1.2)没有明显的守恒定律, 因为 x 和 y 都不是循环坐标。如果将牛顿定律改写为极坐标形式以利用 $\partial V / \partial \dot{\phi} = 0$, 相应的方程由于离心力和科里奥利力项变得复杂。正是拉格朗日形式体系允许我们选择最能反映势能对称性的坐标, 而不改变方程的简单形式。

$$x - 1/2^2 + y^2 = 1/4$$

第三章 自旋

3.1 引言

在本章中，我们将考虑了一类不能通过四个量子公设的直接应用来处理的量子现象。原因是这些现象涉及一种称为自旋的量子自由度，而这种量子自由度没有经典的对应物。因此，我们既不能通过量子公设 II 来获得自旋算符，也不能立即写下影响其时间演化的量子哈密顿量。对于大多数粒子来说，这个问题非常重要，电子、质子、中子、光子都有自旋自由度。幸运的是，这个问题可以通过经典直觉和类比推理的巧妙结合来解决。在这一章中，我们只研究电子自旋。对其他粒子自旋的处理非常相似，除了光子，它以 c 的速度移动，不能进行非相对论处理。光子自旋将在第 18 章中讨论。

在接下来三个章节我们将解决以下问题：

- (1) 这个新的自旋的自由度是什么？
- (2) 希伯尔特空间怎么样修正才能把这个自由度考虑在内？自旋算子在这个空间里看起来像什么（自旋运动学）
- (3) 自旋是如何随时间演化的，即它是如何进入哈密顿量的（旋转动力学）？

3.2 自旋的本质是什么？

在实验中，测量的不是角动量，而是一个称为磁矩的相关量。稍后将对此进行详细介绍。此外，自旋是在光谱证据的基础上首次发现的，而不是从上述类型的实验中发现的。描述自旋最好的方式是把它当成一种角动量形式。然而随着接下来的实验展示，他不是与算符 L 相关的角动量当一个电子处于零线性动量的状态下。即处于一个恒定（与空间无关）波函数的状态。由于算子 L_x, L_y 和 L_z 在作用于波函数上是时会变成零，我们现有的形式

表明如果沿着 z 方向的角动量去测量，就会得到零的结果。然而实际的实验表明这是错误的，结果是 $\pm \frac{\hbar}{2}$ *

*在实际操作中，测量的不是角动量，而是一个称为磁矩的相关量。稍后将对此进行详细介绍。此外，自旋是在光谱证据的基础上首次发现的，而不是从上述类型的实验中发现的

附录 A 附录

A.1 矩阵求逆

这个简短的部分小节仅用于帮助您理解主文中的公式(1.8.5)并不全面。

考虑一个 3×3 矩阵的求逆：

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \quad (\text{A.1.1})$$

在下面的讨论中，我们将 M 的元素命名为这种方式，而不是 M_{ij} ，因为我们将把行视为向量 \mathbf{A} 、 \mathbf{B} 和 \mathbf{C} 的分量，即在向量分析的符号表示法中（我们将在本节中遵循这种表示法）。

$$\mathbf{A} = a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k} \text{ and so on}$$

接下来考虑一个向量三元组。

$$\begin{aligned} \mathbf{A}_R &= \mathbf{B} \times \mathbf{C} \\ \mathbf{B}_R &= \mathbf{C} \times \mathbf{A} \\ \mathbf{C}_R &= \mathbf{A} \times \mathbf{B} \end{aligned} \quad (\text{A.1.2})$$

被称为与 \mathbf{A} 、 \mathbf{B} 和 \mathbf{C} 互逆的向量三元组。一般而言，

$$\mathbf{A} \cdot \mathbf{A}_R \neq 0, \quad \mathbf{A} \cdot \mathbf{B}_R = \mathbf{A} \cdot \mathbf{C}_R = 0 \quad \text{循环置换} \quad (\text{A.1.3})$$

如果我们现在构造一个矩阵 $\overline{\mathbf{M}}$ （称为 \mathbf{M} 的余子式转置），其列是互为倒数的向量，

$$\overline{\mathbf{M}} = \begin{bmatrix} (a_R)_1 & (b_R)_1 & (c_R)_1 \\ (a_R)_2 & (b_R)_2 & (c_R)_2 \\ (a_R)_3 & (b_R)_3 & (c_R)_3 \end{bmatrix}$$

那么

$$M \cdot \overline{M} = \begin{bmatrix} \mathbf{A} \cdot \mathbf{A}_R & \mathbf{A} \cdot \mathbf{B}_R & \mathbf{A} \cdot \mathbf{C}_R \\ \mathbf{B} \cdot \mathbf{A}_R & \mathbf{B} \cdot \mathbf{B}_R & \mathbf{B} \cdot \mathbf{C}_R \\ \mathbf{C} \cdot \mathbf{A}_R & \mathbf{C} \cdot \mathbf{B}_R & \mathbf{C} \cdot \mathbf{C}_R \end{bmatrix} = \begin{bmatrix} \mathbf{A} \cdot \mathbf{A}_R & 0 & 0 \\ 0 & \mathbf{B} \cdot \mathbf{B}_R & 0 \\ 0 & 0 & \mathbf{C} \cdot \mathbf{C} \end{bmatrix} \quad (\text{A.1.4})$$

$$I_0(\alpha, \beta) = e^{\beta^3/4\alpha} \int_{-\infty}^{\infty} e^{-\alpha(x-\beta/2\alpha)^2} dx = e^{\beta^2/4\alpha} \left(\frac{\pi}{\alpha}\right)^{1/2} \quad (\text{A.1.5})$$