

量子化学

Ira N. Levine

第七版

目录

1 薛定谔方程	1
1.1 量子化学	1
1.2 量子力学的历史背景	2
1.3 不确定性原理	5
1.4 含时薛定谔方程	6
1.5 定态薛定谔方程	9
1.6 概率	11
1.7 复数	12
1.8 单位	14
1.9 微积分学	14
2 箱中粒子	17
2.1 微分方程	17
2.2 一维盒子中的粒子	18
2.3 一维自由粒子	23
2.4 矩形势阱中的粒子	24
2.5 隧穿	25
3 算符	27
3.1 算符	27
3.2 本征函数和本征值	30
3.3 算符和量子力学	32
3.4 三维多粒子的薛定谔方程	36
3.5 三维盒子中的粒子	38
3.6 简并	41
3.7 平均值	42
3.8 波函数的约束条件	45
4 谐振子	47
4.1 微分方程的幂级数解	47
4.2 一维谐振子	49
4.3 双原子分子的振动	57
4.4 一维定态薛定谔方程的数值解法	60

5	角动量	65
5.1	同时确定多个属性	65
5.2	矢量	68
5.3	单粒子系统的角动量	72
5.4	角动量阶梯算符法	83
6	氢原子	87
6.1	单粒子中心力问题	87
6.2	非相互作用粒子和变量分离	89
6.3	将双粒子问题简化为两个单粒子问题	90
6.4	双粒子的刚性转子模型	92
6.5	氢原子	96
6.6	束缚态氢原子波函数	102
6.7	类氢轨道	109
6.8	塞曼效应	113
6.9	径向薛定谔方程的数值解法	114
7	量子力学理论	117
7.1	符号	117
7.2	厄米算符	118
7.3	用本征函数展开	122
7.4	对易算符的本征函数	127
7.5	宇称	129
7.6	测量和态叠加	132
7.7	位置本征函数	136
7.8	量子力学的基本假设	138
7.9	测量和量子力学解释	142
7.10	矩阵	144
8	变分法	149
8.1	变分理论	149
8.2	变分法的推广	153
8.3	行列式	154
8.4	联立线性方程组	156
8.5	线性变分函数	159
8.6	矩阵、本征值和本征向量	165
9	微扰理论	173
9.1	微扰理论	173
9.2	非简并微扰理论	174
9.3	微扰理论求解基态氦原子	180
9.4	变分法求解基态氦	180
9.5	能量简并的微扰理论	180
9.6	久期方程的简化	180

9.7 微扰理论求解第一激发态的氢原子	180
9.8 含时微扰理论	180
9.9 辐射和物质的相互作用	180
10 电子自旋和自旋-统计定理	181
10.1 电子自旋	181
10.2 自旋和氢原子	181
10.3 自旋-统计定理	181
10.4 氢原子	181
10.5 Pauli 不相容原理	181
10.6 Slater 行列式	181
10.7 微扰理论求解基态锂原子	181
10.8 变分法求解基态锂原子	181
10.9 自旋磁矩	181
10.10 电子自旋的梯度算符	181
11 多电子原子	183
11.1 Hartree-Fock 自洽场方法	183
11.2 轨道和元素周期表	183
11.3 电子相关	183
11.4 角动量的叠加	183
11.5 多电子原子的角动量	183
11.6 自旋-轨道耦合	183
11.7 原子哈密顿算符	183
11.8 Condon-Slater 规则	183
12 分子对称性	185
12.1 对称元素和对称操作	185
12.2 分子点群	185
13 双原子分子的电子结构	187
13.1 Born-Oppenheimer 近似	188
13.2 双原子分子核的运动	188
13.3 原子单位制	188
13.4 氢分子离子	188
13.5 H^{2+} 离子基态的近似解	188
13.6 H^{2+} 离子激发态的分子轨道	188
13.7 同核双原子分子的分子轨道构型	188
13.8 双原子分子的分子光谱项	188
13.9 氢分子	188
13.10 价键理论求解 H_2	188
13.11 分子轨道理论和价键理论比较	188
13.12 同核双原子分子的分子轨道理论和价键理论波函数	188
13.13 H_2 分子的激发态	188

13.14双原子分子的自洽场方法波函数	188
13.15分子轨道理论处理异核双原子分子	188
13.16价键理论处理异核双原子分子	188
13.17价电子近似	188
14 分子量子力学理论	189
14.1 电子概率密度	189
14.2 偶极矩	189
14.3 分子的 Hartree-Fock 方法	189
14.4 Virial 理论	189
14.5 Virial 理论和化学键	189
14.6 Hellmann-Feynman 理论	189
14.7 静电力定理	189
15 分子的电子结构	191
15.1 从头计算、密度泛函、半经验和分子力学方法	192
15.2 多原子分子的分子光谱项	192
15.3 自洽场方法、分子轨道理论求解多原子分子	192
15.4 基函数	192
15.5 自洽场方法、分子轨道理论求解H ₂ O 分子	192
15.6 布居分析和键级	192
15.7 分子静电势、分子表面和原子电荷	192
15.8 定域分子轨道	192
15.9 自洽场方法、分子轨道理论求解甲烷、乙烷和乙烯	192
15.10分子结构学	192
15.11构象搜寻	192
15.12分子振动频率	192
15.13热力学定律	192
15.14从头计算方法的量子化学程序	192
15.15进行从头计算	192
15.16快速 Hartree-Fock 计算方法	192
15.17溶剂效应	192
16 电子关联方法	193
16.1 关联能	194
16.2 组态相互作用	194
16.3 多体微扰理论	194
16.4 耦合簇方法	194
16.5 密度泛函理论	194
16.6 能量计算的复合方法	194
16.7 扩散量子 Monte Carlo 法	194
16.8 非共价相互作用	194
16.9 NMR 屏蔽常数	194
16.10碎裂法	194

16.11 相对论效应	194
16.12 价键理论求解多原子分子	194
16.13 GVB (广义价键波函数)、VBSCF (价键自洽场)、BOVB (呼吸轨道价键) 方法	194
16.14 化学反应	194
17 半经验和分子力学方法求解分子	195
17.1 半经验分子轨道求解平面共轭分子	195
17.2 Hückel 分子轨道理论	195
17.3 PPP 方法	195
17.4 一般的半经验分子轨道和密度泛函方法	195
17.5 分子力学方法	195
17.6 经验和半经验方法处理溶剂效应	195
17.7 化学反应	195
17.8 量子化学的未来	195
A 附录	197
A.1 Complex Numbers	197
A.2 Calculus Formulas	197
B 参考文献	199
C 部分习题答案	203
D 索引	205

Chapter 1

薛定谔方程

1.1 量子化学

十七世纪末，艾萨克·牛顿（Isaac Newton）创立了**经典力学**，即宏观物体的运动规律。二十世纪初，物理学家发现经典力学无法正确描述原子和分子的电子和原子核等极小粒子的行为。这类粒子的行为由一组称为**量子力学**的定律来描述。

量子化学将量子力学应用于化学问题。量子化学的影响在化学的各个分支中都很明显。物理化学家利用量子力学（借助统计力学）计算气体的热力学性质（如熵、热容量）；解释分子光谱，从而通过实验确定分子性质（如分子几何形状、偶极矩、内旋转能垒、同分异构体之间的能量差异）；从理论上计算分子性质；计算化学反应中过渡态的性质，从而估算速率常数；理解分子间作用力；以及处理固体中的成键问题。

有机化学家利用量子力学估算分子的相对稳定性，计算反应中间产物的性质，研究化学反应的机理，以及分析和预测核磁共振谱。

分析化学家广泛使用光谱法。只有利用量子力学，才能正确理解和解释光谱中的谱线频率和强度。

无机化学家使用配体场理论（一种近似量子力学的方法）来预测和解释过渡金属络离子的性质。

虽然重要生物分子的巨大尺寸使得对它们进行量子力学计算极其困难，但生物化学家们正开始从对生物分子构象、酶与底物结合以及生物分子溶解的量子力学研究中获益。

量子力学决定了纳米材料（至少有一个维度在 1 nm 到 100 nm 之间的物体）的特性，目前正在开发处理纳米材料的计算方法。当材料的一个或多个尺寸小于 100 nm（尤其是小于 20 nm）时，其光学、电子、化学和其它特性就会发生与宏观材料显著不同的巨大变化。一个维度在 1 nm 到 100 nm 之间的半导体或金属物体称为量子阱；两个维度在此范围内的称为量子线；三个维度都在此范围内的称为量子点。这些名称中的“量子”一词表明量子力学在决定此类材料特性方面发挥了关键作用。许多人猜测，纳米科学和纳米技术将带来“下一次工业革命”。

计算机计算速度的迅速提高和分子计算新方法（如密度泛函理论-参见第 16.4 节）的发展，使量子化学成为化学各个领域的实用工具。现在，有几家公司出售用于分子量子化学计算的量子化学软件。这些软件不仅适用于量子化学家，也适用于其他各类化学家。由于量子化学及相关理论和计算方法的作用日益显著，美国化学学会于 2005 年开始出版新的期刊《化学理论与计算杂志》（Journal of Chemical Theory and Computation）。

“量子力学……几乎是所有现代科学技术的基础。它支配着晶体管和集成电路的行为……并且是……现代化学和生物学的基础。” [斯蒂芬·霍金（Stephen Hawking），《时间简史》，1988 年，

Bantam 出版社, 第 4 章)]

1.2 量子力学的历史背景

量子力学的发展起源于 1900 年普朗克 (Planck) 对加热固体发出的光的研究, 因此我们首先要讨论光的本质。

1803 年, 托马斯·杨 (Thomas Young) 通过观察光穿过两个相邻针孔时的衍射和干涉, 为光的波动性提供了令人信服的证据。(衍射是波绕障碍物弯曲。干涉是将两个频率相同的波结合在一起, 产生一个波, 该波在空间每一点的扰动是每个干涉波在该点产生的扰动的代数和或矢量和。参见任何一年级物理课本)。

1864 年, 詹姆斯·克拉克·麦克斯韦 (James Clerk Maxwell) 建立了四个方程, 即麦克斯韦方程组 (Maxwell's equations), 统一了电学和磁学定律。麦克斯韦方程预言: 加速的电荷会以电磁波的形式辐射能量, 电磁波由振荡的电场和磁场组成。麦克斯韦方程预测的这些波的速度与实验测得的光速相同。

1888 年, 海因里希·赫兹 (Heinrich Hertz) 根据麦克斯韦方程组的预测, 探测到了电火花中加速电荷产生的无线电波。这使物理学家确信, 光确实是一种电磁波。

所有电磁波在真空中都以 $c = 2.998 \times 10^8$ m/s 的速度传播。电磁波的频率 ν 和波长 λ 的关系是

$$\boxed{\lambda\nu = c} \quad (1.1)$$

(方框内的等式应牢记; 附录中给出了希腊字母表)。根据电磁波的频率, 人们给电磁波进行了分类。按照频率递增的顺序依次为无线电波、微波、红外线辐射、可见光、紫外线辐射、X 射线和 γ 射线。我们用光来表示任何一种电磁辐射。可见光和紫外线辐射的波长以前用埃 (\AA) 来表示, 现在用纳米 (nm) 来表示:

$$\boxed{1 \text{ nm} = 10^{-9} \text{ nm}, \quad 1 \text{ \AA} = 10^{-10} \text{ m} = 0.1 \text{ nm}} \quad (1.2)$$

19 世纪 90 年代, 物理学家测量了在固定温度下加热黑体发出的各种频率的光强度, 并在多个温度下进行了这些测量。黑体是一种能吸收落在其上的所有光线的物体。一种对黑体对良好近似是一个带有小孔的空腔。1896 年, 物理学家维恩 (Wien) 提出了黑体辐射与光频和黑体温度的关系式: $I = a\nu^3/e^{b\nu/T}$, 其中 a 与 b 是经验常数, $I d\nu$ 是黑体在单位时间和单位表面积内辐射的频率在 ν 到 $\nu + d\nu$ 范围内的能量, 其中 $d\nu$ 是一个无穷小的频率范围。维恩的公式很好地拟合了 1896 年的黑体辐射数据, 但他对该公式的理论论证却不令人满意。

1899-1900 年, 对黑体辐射的测量扩展到了比以前更低的频率, 维恩公式在低频下显示出较大的偏差。这些偏差促使物理学家马克斯·普朗克 (Max Planck) 于 1900 年 10 月提出了以下公式 $I = a\nu^3/(e^{b\nu/T} - 1)$, 并发现该公式适用于所有频率的数据。

提出这个公式后, 普朗克一直在寻找理论依据。1900 年 12 月, 他向德国物理学会提交了他的公式的理论推导。普朗克假定黑体中的辐射发射器和吸收器是与空腔中的电磁辐射处于平衡状态的谐振电荷 (谐振子)。他假设频率为 ν 的谐振子的总能量由 N 个不可分割的“能量元素”组成, 每个能量元素的大小为 $h\nu$, 其中 N 为整数, h (普朗克常数) 是物理学中的一个新常数。普朗克将这些能量元素分布在谐振子中。实际上, 每个谐振子的能量必须是 $h\nu$ 的整数倍 (尽管普朗克没有明说)。因此, 每个谐振子的能量都被量子化了, 这意味着谐振子的能量只允许是某些分立的值。普朗克的理论导出了 $a = 2\pi h/c^2$ 和 $b = h/k$, 其中 k 是玻尔兹曼 (Boltzmann) 常量。通过拟合实验中得到的黑体辐射曲线, 普朗克常数为 $h = 6.6 \times 10^{-34}$ J·s。

普朗克的工作通常被认为是量子力学的开端。然而，物理学史学家们一直在争论：1900 年的普朗克究竟是将能量量子化视为对物理事实的描述，还是仅仅将其视为一种数学近似，从而使他能够获得正确的黑体辐射公式。[见 O. Darrigol, *Centaurus*, 43, 219 (2001); C. A. Gearhart, *Phys. Perspect.*, 4, 170 (2002) (可在线查阅: employees.csbsju.edu/cgearhart/Planck/PQH.pdf; S. G. Brush, *Am. J. Phys.*, 70, 119 (2002) (www.punsterproductions.com/~sciencehistory/cautious.htm)]。物理历史学家克拉格 (Kragh) 指出：“如果说 1900 年 12 月物理学发生了一场革命，似乎没有人注意到它。普朗克也不例外，赋予他的工作的重要性在很大程度上是一种历史重构。” (H. Kragh, 《物理学世界》，2000 年 12 月，第 31 页)

能量量子化的概念与以往物理学的所有理念直接相悖。经典力学认为：物质体的能量可以连续变化。然而，只有在能量量子化的假设下，才能得到正确的黑体辐射曲线。

能量量子化的第二个应用是光电效应。在光电效应中，光照射到金属上会导致电子被激发。波的能量与其强度成正比，而与其频率无关。因此光的电磁波图象会让人想到：发射的光电子的动能会随着光强度的增加而增加，但不会随着光频率的变化而变化。

1905 年，爱因斯坦 (Einstein) 指出：将光视为由粒子状实体 (称为光子) 组成的集合体，每个光子都具有能量

$$E = h\nu \quad (1.3)$$

就可以解释这些观测结果。当金属中的电子吸收光子时，所吸收的光子能量的一部分被用来克服将电子固定在金属中的力，剩余的能量则转化为电子逸出后的动能。由能量守恒，有 $h\nu = \phi + T$ ，其中 ϕ 是电子逸出金属表面所需的最小能量 (金属的功函数)， T 是被激发电子的最大动能。提高光的频率 ν 会增加光子能量，从而增加被激发电子的动能。增加固定频率下的光强度会增加光子撞击金属的速度，从而增加电子的发射速度，但不会改变每个发射电子的动能。(克拉格认为：“可以有力地证明，是爱因斯坦首先认识到了量子理论的本质。”；H. Kragh, 《物理学世界》，2000 年 12 月，第 31 页)

光电效应表明：除了在衍射实验中表现出波动性外，光还可以表现出粒子性。

1907 年，爱因斯坦将能量量子化应用于固体元素中原子的振动。假设每个原子在每个方向 (x, y, z) 上的振动能量被限制为 $h\nu_{vib}$ 的整数倍，其中振动频率 ν_{vib} 是元素的性质。爱因斯坦利用统计力学推导出固体恒容热容 C_V 的表达式。他的公式与已知的钻石 C_V 与温度的关系数据相当吻合。

现在我们来考虑物质的结构。

19 世纪末，对放电管和自然放射性的研究表明：原子和分子是由带电粒子组成的。电子带有负电荷，而质子带有与电子电荷大小相等但符号相反的正电荷，其质量是电子的 1836 倍。原子的第三种成分-中子 (1932 年发现) 不带电荷，比质子稍重。

从 1909 年开始，卢瑟福 (Rutherford)、盖革 (Geige) 和马斯登 (Marsden) 反复将一束 α 粒子穿过薄金属箔，并让它们落在荧光屏上，从而观察到粒子的偏转。 α 粒子是从天然放射性衰变中获得的带正电荷的氦核。大多数 α 粒子通过金属箔时基本上没有发生偏转，但令人惊讶的是，少数粒子发生了很大的偏转，有些粒子向后偏转。如果正电荷遍布整个原子 [正如 J. J. 汤姆逊 (Thomson) 在 1904 年提出的那样]，那么一旦高能 α 粒子穿透原子，斥力就会减弱，根据经典静力学，在原子中心的斥力为零。因此，卢瑟福得出结论：只有当正电荷集中在一个微小而沉重的原子核中时，才会发生如此大的偏转。

原子包含一个极小 (半径为 10^{-13} 到 10^{-12} cm) 的重核，该核由中子和 Z 个质子组成，其中 Z 是原子序数，原子核外有 Z 个电子，带电粒子的相互作用遵循库仑定律。(核子在原子核内通过强相互作用结合在一起，这与我们无关) 原子的半径约为 1 \AA ，气体动力学理论的结果就表明了这一点。

原子和分子的化学性质是由其电子结构决定的，因此就产生了电子运动和能量的性质问题。由于原子核的质量远大于电子，我们认为原子核的运动与电子的运动相比是微小的。

1911 年，卢瑟福提出了原子的行星模型，其中电子以不同的轨道围绕原子核旋转，就像行星围绕太阳旋转一样。然而，这一模型存在一个根本性的难题。根据经典电磁理论，加速的带电粒子会以电磁波（光波）的形式辐射能量。以恒定速度围绕原子核旋转的电子正在被加速，因为其速度矢量的方向在不断变化。因此，卢瑟福模型中的电子应该不断地通过辐射损失能量，从而向原子核螺旋运动。因此，根据经典（19 世纪）的物理学，卢瑟福的原子结构是不稳定的，会发生坍缩。

1913 年，尼尔斯·玻尔（Niels Bohr）将能量量子化的概念应用于氢原子，提出了解决这一难题的可行方法。玻尔假定氢原子中电子的能量是量子化的，电子只能在一系列允许的轨道中的一个轨道上运动。当电子从一个玻尔轨道过渡到另一个玻尔轨道时，一个频率 ν 满足

$$E_{upper} - E_{lower} = h\nu \quad (1.4)$$

的光子被吸收或激发。其中 E_{upper} 和 E_{lower} 是高低两个状态的能量（能量守恒）。玻尔假定电子从自由（电离）态过渡到束缚轨道时会发射出一个光子，其频率是电子在束缚轨道中经典旋转频率的二分之一的整数倍，他利用经典力学推导出了氢原子能级的计算公式。利用 (1.4)，他得到了与观测到的氢谱一致的结果。然而，用玻尔理论拟合氦光谱的尝试却失败了。此外，该理论无法解释分子中的化学键。

玻尔模型的失败源于使用经典力学来描述原子中的电子运动。原子光谱显示出离散的频率，这表明只有某些能量的运动是允许的；以及电子能量是量子化的。然而，经典力学允许连续的能量范围。因此，路易斯·德布罗意（Louis de Broglie）在 1923 年提出：电子的运动可能具有波的一面。质量为 m 、速度为 v 的电子将具有波长

$$\lambda = \frac{h}{mv} = \frac{h}{p} \quad (1.5)$$

其中 p 是线动量。德布罗意通过与光子的类比推理得出公式 (1.5)。根据爱因斯坦的狭义相对论，光子的能量可以用 $E = pc$ 表示，其中 c 是光速， p 是光子的动量。利用 $E_{photon} = h\nu$ 可以得到光子以 c 的速度运动时的 $pc = h\nu = hc/\lambda$ 和 $\lambda = h/p$ 。方程式 (1.5) 是电子相应的方程式。

1927 年，戴维森（Davisson）和格尔默（Germer）通过实验证实了德布罗意的假设，他们从金属中反射电子并观察到衍射效应。1932 年，斯特恩（Stern）在氢原子和氢分子中观察到了同样的效应，从而验证了波效应并非电子所特有，而是微观粒子运动的某种普遍规律的结果。通过衍射光栅可以观察到大至 $C_{48}H_{26}F_{24}N_8O_8$ 的分子的衍射和干涉现象。[T. Juffmann 等, Nat. Nanotechnol., 7,297 (2012)]。有关分子干涉图案形成的影片，请访问 www.youtube.com/watch?v=vCiOMQIRU7I。

因此，电子的行为在某些方面像粒子，而在另一些方面则像波。我们面临着物质（以及光）表面上自相矛盾的“波粒二象性”。电子怎么可能既是粒子（局部实体），又是波（非局部实体）呢？答案是电子既不是波也不是粒子，而是另一种东西。用经典物理学的波或粒子概念来准确描述电子的行为是不可能的。经典物理学的概念是根据宏观世界的经验发展而来的，并不能正确描述微观世界。进化塑造了人类大脑，使其能够理解并有效处理宏观现象。人类神经系统的发展并不是为了处理原子和分子层面的现象，因此我们无法完全理解这些现象也就不足为奇了。

虽然光子和电子都表现出明显的二元性，但它们并不是同一类实体。光子在真空中以 c 的速度运动，其静止质量为零；而电子始终具有 $v < c$ 和非零静止质量。光子必须始终按相对论处理，但速度远小于 c 的电子可以按非相对论处理。

1.3 不确定性原理

让我们来看看波粒二象性对同时测量微观粒子的 x 坐标和线动量 x 分量的尝试有什么影响。我们从一束动量为 p 、沿 y 方向运动的粒子束开始，让粒子束落在一个窄缝上。狭缝后面是一块照相板。见图1.1。

粒子通过宽度为 w 的狭缝时，其 x 坐标的不确定性为 w 。我们将这种 x 值的分布称为 Δx ，即有 $\Delta x = w$ 。

由于微观粒子具有波的特性，它们会被狭缝衍射，在平板上产生衍射图样（就像光束一样）。图1.1中图形的高度表示到达给定点的粒子数量。衍射图样显示：当粒子被狭缝衍射时，它们的运动方向发生了改变，因此部分动量被转移到了 x 方向。动量在 x 方向上的分量 p_x 等于动量矢量 \vec{p} 在 x 方向上的投影。向上偏转 α 角的粒子的动量满足 $p_x = p \sin \alpha$ ；向下偏转 α 角的粒子的动量满足 $p_x = -p \sin \alpha$ 。由于大多数粒子的偏转范围在 $-\alpha$ 到 α 之间，其中 α 是与衍射图样中第一个最小值的夹角，因此我们将以中心衍射峰中动量值的一般作为动量 x 分量不确定性 Δp_x 的量度： $\Delta p_x = p \sin \alpha$ 。

因此，在狭缝处进行测量，有

$$\Delta x \Delta p_x = pw \sin \alpha \quad (1.6)$$



图 1.1: 狭缝的电子衍射

第一个衍射最小值出现的角度 α 很容易计算。出现第一个最小值的条件是：通过狭缝上边缘的粒子和通过狭缝中心的粒子所经过的距离之差应等于 $\frac{1}{2}\lambda$ ，其中 λ 是相关波的波长。这样，从狭缝顶部发出的波与从狭缝中心发出的波正好相位相反，它们相互抵消。从狭缝中点以下距离 d 处发出的波与从狭缝顶部以下距离 d 处发出的波相互抵消。在图1.2中画出 AC ，使 $AD = CD$ ，则路径长度之差为 BC 。狭缝到屏幕的距离比狭缝宽度大。因此，由于 AD 和 BD 几乎平行，则 $\angle ACB$ 基本上是直角，因此 $\angle BAC = \alpha$ ，路径差 BC 就是 $\frac{1}{2}w \sin \alpha$ 。设 BC 等于 $\frac{1}{2}\lambda$ ，我们就得到了 $w \sin \alpha = \lambda$ ，公式(1.6)就变成了 $\Delta x \Delta p_x = p\lambda$ ，其中波长 λ 由德布罗意关系式 $\lambda = h/p$ 给出，因此 $\Delta x \Delta p_x = h$ 。由于不确定度尚未精确定义，因此使用等号并不合适。我们可以用

$$\Delta x \Delta p_x \approx h \quad (1.7)$$

来表示 x 方向上的不确定性，同时 p_x 具有与普朗克常数相当的数量级。



图 1.2: 计算一级衍射的最小值

尽管我们只用一种特例证明了 (1.7)，但其有效性是普遍的。无论如何尝试，微观“粒子”的波粒二象性都限制了我们的同时测量这种粒子的位置和动量的能力。我们对位置的测定越精确，对动量的测定就越不准确。（在图 1.1 中， $\sin \alpha = \lambda/w$ ，因此缩小狭缝会增加衍射图样的扩散）这种限制就是沃纳·海森堡（Werner Heisenberg）于 1927 年发现的**不确定性原理**（uncertainty principle）。

由于波粒二象性，测量行为会给被测系统带来不可控制的干扰。我们从粒子具有精确的 p_x 值（零）开始。通过施加狭缝，我们测量了粒子的 x 坐标，精确度为 w ，但这次测量给粒子的 p_x 值带来了不确定性。测量改变了系统的状态。

1.4 含时薛定谔方程

经典力学只适用于宏观粒子。对于微观“粒子”，我们需要一种新的力学形式，即**量子力学**。为简单起见，我们将讨论一个单粒子的一维系统。

在经典力学中，粒子的运动遵循牛顿第二定律：

$$F = ma = m \frac{d^2 x}{dt^2} \quad (1.8)$$

其中 F 是作用在粒子上的合外力， m 是粒子的质量， t 是时间； a 是加速度，由 $a = dv/dt = (d/dt)(dx/dt) = d^2 x/dt^2$ 给出，其中 v 是速度。方程 (1.8) 包含坐标 x 对时间的二阶导数。为了求解这个方程，我们必须进行两次积分。这就引入了两个积分常量 c_1 和 c_2 ，有

$$x = g(t, c_1, c_2) \quad (1.9)$$

其中 g 是时间的某个函数。我们现在要问：在某一特定时间 t_0 ，我们必须掌握哪些信息才能预测粒子未来的运动？如果我们知道在 t_0 时刻，粒子位于 x_0 处，则有

$$x_0 = g(t_0, c_1, c_2) \quad (1.10)$$

由于我们需要确定两个常数，因此需要更多信息。将 (1.9) 左右两边对时间求导，有

$$\frac{dx}{dt} = v = \frac{d}{dt} g(t, c_1, c_2)$$

如果我们也知道在 t_0 时刻，粒子具有速度 v_0 ，那么有进一步的关系

$$v_0 = \left. \frac{d}{dt} g(t, c_1, c_2) \right|_{t=t_0} \quad (1.11)$$

知道 x_0 和 v_0 后，我们可以用方程 (1.10) 和 (1.11) 求出 c_1 和 c_2 。知道 c_1 和 c_2 后，就可以用方程 (1.9) 来精确预测粒子在未来的运动。

作为式 (1.8) 至 (1.11) 的例子，考虑一个粒子在地球重力场中的运动。取垂直向上为 x 轴正方向，则作用在粒子上的合外力方向垂直向下，大小为 $-mg$ ，其中 g 是重力加速度。则牛顿第二定律 (1.8) 形式为 $-mg = m d^2x/dt^2$ ，所以 $d^2x/dt^2 = -g$ 。积分一次，有 $dx/dt = -gt + c_1$ 。如果我们知道粒子在 t_0 时刻的速度为 v_0 ，就可以求出积分常量 c_1 。又由于 $v = dx/dt$ ，将 $v_0 = -gt_0 + c_1$ 即 $c_1 = v_0 + gt_0$ 带入，有 $dx/dt = -gt + gt_0 + v_0$ 。再次积分，我们引入了另一个积分常量 c_2 。如果我们知道粒子在 t_0 时刻的位置为 x_0 ，则可以求出 c_2 。因此，我们有 $x = x_0 - \frac{1}{2}g(t - t_0)^2 + v_0(t - t_0)$ 。知道了在 t_0 时刻粒子的 x_0 和 v_0 ，我们就能预测粒子未来的运动轨迹。

在一维范围内运动的粒子的经典力学势能 V 的定义为满足

$$\boxed{\frac{\partial V(x, t)}{\partial x} = -F(x, t)} \quad (1.12)$$

例如，对于一个在地球重力场中运动的粒子，有 $\partial V/\partial x = -F = mg$ ，积分一次，有 $V = mgx + c$ ，其中 c 为积分常量。我们可以任意选择我们想要的零势能面。令 $c = 0$ ，我们有势能函数 $V = mgx$ 。

经典力学中的“状态”一词是指系统中每个粒子在某一时刻的位置和速度，以及作用在粒子上的力。根据牛顿第二定律，给定一个系统在任何时候的状态，它的未来状态和未来运动都是完全确定的，如式 (1.9)-(1.11) 所示。牛顿定律在解释行星运动方面取得的巨大成功让许多哲学家将牛顿定律作为哲学决定论的论据。数学家和天文学家拉普拉斯 (Laplace) (1749-1827) 假定宇宙是由服从牛顿定律的粒子组成的。因此，给定宇宙在某一时刻的状态，宇宙万物未来的运动是完全确定的。原则上，一个能知道宇宙任何瞬间状态的超级星体可以计算出所有未来的运动。

尽管经典力学是确定性的，但许多经典力学系统（例如，在重力、摩擦力和周期性变化的驱动力影响下摆动的摆）在系统参数的一定范围内表现出混沌行为。在混沌系统中，运动对粒子位置和速度的初始值以及作用力异常敏感，两个初始状态相差一个实验无法检测的量，最终会导致系统的未来行为大相径庭。因此，由于测量初始状态的精度有限，即使系统服从确定性方程，实际上也不可能预测混沌经典力学系统的长期行为。对太阳系行星数千万年轨道的计算机计算表明，行星的运动是混沌的。[I. Peterson, 《牛顿的钟：混沌的太阳系》，Freeman, 1993 年；J. J. Lissauer, Rev. Mod. 71, 835 (1999)]

只要精确了解经典力学系统的当前状态，我们就能预测其未来状态。然而，海森堡不确定性原理表明：我们无法同时确定微观粒子的精确位置和速度，因此无法获得经典力学所需的数据来预测系统的未来运动。在量子力学中，我们必须满足于对未来精确运动的不完全预测。

我们研究量子力学的方法是假设基本原理，然后利用这些假设推导出实验可检验的结果，例如原子的能级。为了描述量子力学中一个系统的状态，我们假设存在一个粒子坐标的函数 Ψ ，称为**状态函数** (state function) 或**波函数** (wave function) (通常写作**波函数** (wavefunction))。由于状态一般会随时间变化，因此 Ψ 也是时间的函数。对于一个单粒子的一维系统，我们有 $\Psi = \Psi(x, t)$ 。波函数包含了一个系统所有可能的信息，因此我们不说“波函数 Ψ 所描述的状态”，而只说“状态 Ψ ”。牛顿第二定律告诉我们，如何通过了解经典力学系统当前状态的了解来找到它的未来状态。而要根据量子力学系统当前的状态找到它未来的状态，我们需要一个方程来告诉我们波函数是如何随时间变化的。对于单粒子的一维系统，将这个方程假定为

$$-\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x, t) \Psi(x, t) \quad (1.13)$$

其中常数 \hbar (约化普朗克常数) 定义为

$$\boxed{\hbar \equiv \frac{h}{2\pi}} \quad (1.14)$$

1926 年，奥地利物理学家埃尔温·薛定谔（Erwin Schrödinger）（1887-1961）建立了波函数的概念及其随时间变化的方程。这个被称为含时薛定谔方程（或薛定谔波方程）（time-dependent Schrödinger Equation）中， i 是虚数单位， m 是粒子的质量， $V(x, t)$ 是系统的势能函数。（量子力学中许多具有重要历史意义的论文可在 dieumsnh.qfb.umich.mx/archivoshistoricosmq 网站上查阅）

含时薛定谔方程包含波函数对时间的一阶导数，如果我们知道 t_0 时刻的波函数，就可以计算出未来任意时刻的波函数（状态）。

波函数包含了我们可能知道的关于它所描述的系统的所有信息。关于粒子 x 坐标的测量结果， Ψ 给了我们什么信息？我们不能指望 Ψ 会像经典力学系统的状态那样具有涉及位置的明确说明。这个问题的正确答案是在薛定谔发现薛定谔方程后不久，由马克斯·玻恩（Max Born）给出的。玻恩推测：对于单粒子的一维系统，

$$|\Psi(x, t)|^2 dx \quad (1.15)$$

给出了在 t 时刻，在 x 到 $x+dx$ 范围内找到该粒子的概率。在 (1.15) 中，两条竖线表示绝对值， dx 表示 x 轴上一个无限小的长度。函数 $|\Psi(x, t)|^2$ 是在 x 轴上不同位置找到粒子的概率密度（probability density）。（第 1.6 节回顾了概率）例如，假设在某个特定时间 t_0 ，粒子处于由波函数 $a^2 e^{-2bx^2}$ 描述的特征状态，其中 a 和 b 是实常数。如果我们在 t_0 时刻观测该粒子的位置，会得到任意的 x 值，因为概率密度 $a^2 e^{-2bx^2}$ 处处非零。由于 $|\Psi|^2$ 在原点取到最大值，因此在 $x=0$ 附近的 x 值比其他地方的更有可能被观测到。

为了将 Ψ 与实验测量联系起来，我们需要许多完全相同的非相互作用系统，每个系统都处于相同的状态 Ψ 。如果我们对 n 个系统进行了 n 次测量，令 dn_x 表示在 x 和 $x+dx$ 之间找到粒子的测量次数，那么 dn_x/n 就是在 x 和 $x+dx$ 之间找到粒子的概率。因此，有

$$\frac{dn_x}{n} = |\Psi|^2 dx$$

而 $(1/n)dn_x/dx$ 与 x 的关系图则给出了概率密度 $|\Psi|^2$ 与 x 的函数关系。也许有人认为，我们可以通过一个处于 Ψ 状态的系统，反复测量粒子的位置来找到概率密度函数。这种方法是错误的，因为测量过程通常会改变系统的状态。我们在讨论不确定性原理（第 1.3 节）时看到过一个这方面的例子。

量子力学是统计性质的。在知道状态的情况下，我们无法准确预测位置测量的结果；我们只能预测各种可能结果的概率。氢原子的玻尔理论规定了电子的精确路径，因此不是正确的量子力学图景。

量子力学并没有说电子像波一样分布在空间的很大区域。相反，用于描述电子运动的概率模式（波函数）表现得像波，并满足波方程。

波函数如何为我们提供位置以外的其他属性信息，将在后面的章节中讨论。

热力学公设（热力学第一、第二和第三定律）是由宏观经验提出的，因此相当容易理解。量子力学的公设是从微观世界的角度阐述的，显得相当抽象。不要指望第一次接触就能完全理解量子力学公设。随着我们对各种实例的处理，对这些公设的理解会加深。

我们引入了薛定谔方程，却没有试图证明它的合理性，这可能会让读者感到困惑。通过几何光学和经典力学之间的类比，以及波光学和量子力学之间的类比，我们可以证明薛定谔方程的合理性。几何光学是波光学的近似，在光的波长远小于仪器尺寸时有效。（回忆一下它在处理镜片和镜子时的用途）。同样地，经典力学是波动力学的近似，当粒子的波长远小于仪器的尺寸时有效。我们可以根据已知的几何光学方程和波光学方程之间的关系，猜测如何从经典力学中得到量子力学的正确方程。由于许多化学家对光学并不特别熟悉，因此省略了这些论证。无论如何，这些类比只能使薛定谔方程看起来可信。它们不能用来推导或证明这个方程。薛定谔方程是理论的一个假设，要通过其预测

与实验的一致性来检验。(关于薛定谔得出方程的推理细节, 请参阅 Jammer, 第 5.3 节。参考文献中的作者姓名用斜体标出)

量子力学为微观粒子提供了运动规律。从实验上看, 宏观物体遵从经典力学。因此, 要使量子力学成为有效的理论, 当我们从微观粒子过渡到宏观粒子时, 量子力学应还原为经典力学。量子效应与德布罗意波长 $\lambda = h/mv$ 有关。由于 h 非常小, 宏观物体的德布罗意波长近似为零。因此, 在极限 $\lambda \rightarrow 0$ 中, 我们期望含时薛定谔方程能够还原为牛顿第二定律。我们可以证明这一点 (见问题 7.59)。

狭义相对论与经典力学的关系也存在类似的情况。在以 c 为光速的极限 $v/c \rightarrow 0$ 中, 狭义相对论可还原为经典力学。我们将要发展的量子力学形式将是非相对论的。相对论与量子力学的完全融合尚未实现。

从历史上看, 量子力学是由海森堡、玻恩和乔丹 (Jordan) 于 1925 年首次使用矩阵提出的, 比薛定谔于 1926 年使用微分方程提出的量子力学早几个月。薛定谔证明: 海森堡公式 (称为**矩阵力学**) 等同于薛定谔公式 (称为**波动力学**)。1926 年, 狄拉克和乔丹分别独立地用一种称为**变换理论**的抽象版本提出了量子力学, 它是矩阵力学和波动力学的概括 (见《狄拉克》)。1948 年, 费曼设计出了量子力学的路径积分公式 [R. P. Feynman, Rev. Mod. Phys., 20, 367 (1948); R. P. Feynman, Rev. Mod. Phys., 20, 367(1948); R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, 1965]。

1.5 定态薛定谔方程

含时薛定谔方程 (1.13) 看上去很可怕。幸运的是, 量子力学在化学中的许多应用都不使用这个方程。取而代之的是更简单的定态薛定谔方程。现在, 我们从一维单粒子的含时薛定谔方程导出定态薛定谔方程。

首先, 我们只讨论一种特殊情况, 即势能 V 不是时间的函数, 而只取决于 x 。系统没有受到与时间相关的外力作用时会出现此情况。此时, 含时薛定谔方程为

$$-\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x, t) \Psi(x, t) \quad (1.16)$$

将式 (1.16) 的解假设为时间函数和坐标 x 函数的乘积:

$$\boxed{\Psi(x, t) = f(t) \psi(x)} \quad (1.17)$$

大写的 Ψ 用来表示含时的波函数, 而小写的 ψ 表示只与 x 坐标有关的因子。与 (1.17) 形式的波函数相对应的状态具有某些特性 (稍后讨论), 这些特性使它们极具吸引力。[并非 (1.16) 的所有解都具有 (1.17) 的形式; 见证明 3.51。] 取 (1.17) 的偏导数, 我们有

$$\frac{\partial \Psi(x, t)}{\partial t} = \frac{df(t)}{dt} \psi(x), \quad \frac{\partial^2 \Psi(x, t)}{\partial x^2} = f(t) \frac{d^2 \psi(x)}{dx^2}$$

代入 (1.16), 有

$$\begin{aligned} -\frac{\hbar}{i} \frac{df(t)}{dt} \psi(x) &= -\frac{\hbar^2}{2m} f(t) \frac{d^2 \psi(x)}{dx^2} + V(x) f(t) \psi(x) \\ -\frac{\hbar}{i} \frac{1}{f(t)} \frac{df(t)}{dt} &= -\frac{\hbar^2}{2m} \frac{1}{\psi(x)} \frac{d^2 \psi(x)}{dx^2} + V(x) \end{aligned} \quad (1.18)$$

此时方程左右两边同除 $f\psi$ 。一般来说, 我们希望 (1.18) 两边各相等的量是 x 和 t 的某个函数。然而, (1.18) 的右边不依赖于 t , 所以 (1.18) 两边相等的函数必须与 t 无关。我们称这个常数为 E 。

令式 (1.18) 的左边等于 E ，我们有

$$\frac{df(t)}{dt} = -\frac{iE}{\hbar} dt$$

左右对时间求积分，有

$$\ln f(t) = -iEt/\hbar + C$$

其中 C 是任意常数。因此，

$$f(t) = e^C e^{-iEt/\hbar} = A e^{-iEt/\hbar}$$

其中 A 是代替了 e^C 的任意常数。由于 A 可以作为因子包含在与 $f(t)$ 相乘的函数 $\psi(x)$ 中，因此 A 可以从 $f(t)$ 中略去。所以，

$$f(t) = e^{-iEt/\hbar}$$

令 (1.18) 的右边等于 E ，我们有

$$\boxed{-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)} \quad (1.19)$$

方程 (1.19) 就是质量为 m 的单粒子一维运动的**定态薛定谔方程** (time-independent Schrödinger Equation)。常数 E 有什么意义？由于 E 在 (1.19) 中作为 $[E - V(x)]$ 出现， E 的维数与 V 相同，因此 E 具有能量的量纲。事实上，我们假设 E 就是系统的能量。(这是后一章将讨论的一个更普遍假设的特例) 因此，在势能仅是 x 的函数的情况下，存在如下形式的波函数

$$\Psi(x, t) = e^{-iEt/\hbar} \psi(x) \quad (1.20)$$

这些波函数对应于恒定能量 E 的状态。在接下来的几章中，我们将主要关注如何找到各种系统的方程 (1.19) 的解。

式 (1.20) 中的波函数是复数函数，但在实验中可观测到的量是概率密度 $|\Psi(x, t)|^2$ 。复数绝对值的平方由复数与其复共轭的乘积给出，复共轭则是在出现 i 的地方用 $-i$ 代替 i (见第 1.7 节)。因此，

$$\boxed{|\Psi|^2 = \Psi^* \Psi} \quad (1.21)$$

其中， $*$ 号表示复共轭。对于波函数 (1.20)，有

$$\begin{aligned} |\Psi(x, t)|^2 &= [e^{-iEt/\hbar} \psi(x)]^* e^{-iEt/\hbar} \psi(x) \\ &= e^{iEt/\hbar} \psi^*(x) e^{-iEt/\hbar} \psi(x) \\ &= e^0 \psi^*(x) \psi(x) = \psi^*(x) \psi(x) \\ |\Psi(x, t)|^2 &= |\psi(x)|^2 \end{aligned} \quad (1.22)$$

在推导 (1.22) 时，我们假设 E 是实数，即 $E = E^*$ 。这将会在第 7.2 节给出证明。

因此，对于 (1.20) 形式的状态，概率密度由 $|\Psi(x)|^2$ 给出，且不随时间变化。这种状态称为**定态** (stationary states)。由于具有物理意义的量是 $|\Psi(x, t)|^2$ ，而且对于定态有 $|\Psi(x, t)|^2 = |\psi(x)|^2$ ，函数 $\psi(x)$ 通常被称为**波函数**，尽管定态的完整波函数是 $\psi(x)$ 与 $e^{-iEt/\hbar}$ 的乘积。定态一词不应误导读者，让他们以为处于定态的粒子就是静止的。确定的是概率密度 $|\Psi|^2$ ，而不是粒子本身。

我们将主要关注能量恒定的状态 (定态)，因此通常会处理定态薛定谔方程 (1.19)。为简单起见，我们将此方程称为“薛定谔方程”。请注意，薛定谔方程包含两个未知数：允许的能量 E 和允许的波函数 ψ 。为了求解这两个未知数，我们除了要求它满足 (1.19) 以外，还需要有额外的条件 (称为边界条件)。边界条件决定了允许的能量，因为只有特定的 E 值才能满足边界条件。这一点在后面章节讨论具体例子时会更加清楚。

1.6 概率

概率在量子力学中起着基础性的作用。本节主要回顾数学中的概率。

关于概率的正确定义一直存在很多争议。其中一个定义如下：如果一个实验有 n 个同样可能的结果，其中 m 个结果有利于某个事件 A 的发生，那么 A 发生的概率就是 m/n 。请注意：这个定义是循环论证，因为它指定了同样可能的结果，而概率才是我们要定义的。它只是假设我们能够识别同样可能发生的结果。另一个定义是基于多次实际操作实验：假设我们进行了 N 次实验，在其中的 M 次实验中，事件 A 发生了。那么将 A 发生的概率定义为

$$\lim_{N \rightarrow \infty} \frac{M}{N}$$

因此，如果我们重复抛一枚硬币，随着次数的增加，正面的比例会逐渐接近 $1/2$ 。

例如，假设我们想知道从一副包含 13 张红心的 52 张标准牌中随机抽取一张牌时，抽到红心的概率。有 52 张牌，因此有 52 种同样可能的结果。有 13 张红心，因此抽出红心有 13 种可能的结果。所以， $m/n = 13/52 = 1/4$ 。即抽到红心的概率为 $1/4$ 。

有时我们会关心两个相关事件同时发生的概率。例如，我们可能会问从一副 52 张牌中抽出两张红心的概率，假设抽出第一张牌后我们不放回去。第一次抽牌有 52 种可能的结果，而每一种可能都对应 51 种可能的第二次抽牌。我们有 $52 \cdot 51$ 种可能的结果。因为有 13 张红心，所以有 $13 \cdot 12$ 种不同的方法抽出两张红心。期望概率为 $(13 \cdot 12) / (52 \cdot 51) = 1/17$ 。这个计算说明了定理：两个事件 A 和 B 同时发生的概率是 A 发生的概率乘以 B 随后发生的条件概率（计算时假设 A 发生）。因此，如果 A 是第一次抽中红心的概率，那么 A 的概率就是 $13/52$ 。如果第一次抽到红心，那么第二次抽到红心的概率是 $12/51$ ，因为牌中还有 12 张红心。那么抽到两张红心的概率是 $(13/52)(12/51) = 1/17$ ，如前所述。

在量子力学中，我们必须处理涉及连续变量（例如 x 坐标）的概率。谈论在 $x = 0.5000\dots$ 这样一个特定点上发现粒子的概率是没有多大意义的，因为 x 轴上有无数个点，而对于我们进行的任何有限次数的测量来说，精确得到 $0.5000\dots$ 的概率都是微乎其微的。相反，我们谈论的是在 x 轴上位于 x 和 $x + dx$ 之间的微小区间内找到粒子的概率， dx 是长度的一个无穷小量。这个概率自然与区间的长度 dx 成正比，并且会随着 x 轴的不同区域而变化。因此，粒子在 x 和 $x + dx$ 之间被发现的概率等于 $g(x)dx$ ，其中 $g(x)$ 是某个函数，表示概率在 x 轴上的变化情况。函数 $g(x)$ 称为**概率密度**，因为它是单位长度上的概率。由于概率是实数、非负数，因此 $g(x)$ 必须是一个处处非负的实函数。波函数 Ψ 可以取负值和复值，不是概率密度。因此，量子力学假设概率密度为 $|\Psi|^2$ [公式 (1.15)]。

粒子位于空间的某个有限区域 $a \leq x \leq b$ 的概率是多少？为了求出这个概率，我们将在位于 a 和 b 之间的所有无限小区域中找到粒子的概率 $|\Psi|^2 dx$ 相加，这就是如下定积分的定义：

$$\int_a^b |\Psi|^2 dx = \text{Pr}(a \leq x \leq b) \quad (1.23)$$

其中， Pr 表示概率。概率为 1 则代表必然性。由于粒子肯定位于 x 轴上的某处，因此我们要求

$$\int_{-\infty}^{\infty} |\Psi|^2 dx = 1 \quad (1.24)$$

若波函数 Ψ 满足式 (1.24)，则我们说 Ψ 是归一化的 (normalized)。对于定态波函数，有 $|\Psi|^2 = |\psi|^2$ 且 $\int_{-\infty}^{\infty} |\psi|^2 dx = 1$ 。

例题：

一个单粒子一维系统在 $t = 0$ 时刻满足 $\Psi = a^{-1/2}e^{-|x|/a}$ ，其中 $a = 1.0000 \text{ nm}$ 。在 $t = 0$ 时对粒子的位置进行一次测量。

(a) 求测量值介于 $x = 1.5000 \text{ nm}$ 和 $x = 1.5001 \text{ nm}$ 之间的概率；

(b) 求测量值介于 $x = 0 \text{ nm}$ 和 $x = 2 \text{ nm}$ 之间的概率；

(c) 求证： Ψ 是归一化的。

(a) 在这个极小的区间内， x 只变化了 0.0001 nm ，而 Ψ 从 $e^{-1.5000 \text{ nm}} \text{ nm}^{-1/2} = 0.22313 \text{ nm}^{-1/2}$ 到 $e^{-1.5001 \text{ nm}} \text{ nm}^{-1/2} = 0.22311 \text{ nm}^{-1/2}$ ，所以 Ψ 在这个区间可以近似为常数，将这一区间视为无穷小区间是一个很好的近似方法。带入 (1.15) 给出的期望概率公式，有

$$\begin{aligned} |\Psi|^2 dx &= a^{-1/2} e^{-2|x|/a} dx = (1 \text{ nm})^{-1} e^{-2(1.5 \text{ nm})/(1 \text{ nm})} (0.0001 \text{ nm}) \\ &= 4.979 \times 10^{-6} \end{aligned}$$

(也可见证明 1.14)

(b) 使用式 (1.23) 以及 $x \geq 0$ 时 $|x| = x$ ，有

$$\begin{aligned} \text{Pr}(0 \leq x \leq 2 \text{ nm}) &= \int_0^{2 \text{ nm}} |\Psi|^2 dx = a^{-1} \int_0^{2 \text{ nm}} e^{-2x/a} dx \\ &= -\frac{1}{2} e^{-2x/a} \Big|_0^{2 \text{ nm}} = -\frac{1}{2} (e^{-4} - 1) = 0.4908 \end{aligned}$$

(c) 使用 $\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^{\infty} f(x) dx$ ，以及 $x \geq 0$ 时 $|x| = x$ 和 $x < 0$ 时 $|x| = -x$ ，有

$$\begin{aligned} \int_{-\infty}^{\infty} |\Psi|^2 dx &= a^{-1} \int_{-\infty}^0 e^{2x/a} dx + a^{-1} \int_0^{\infty} e^{-2x/a} dx \\ &= a^{-1} \left(\frac{1}{2} a e^{2x/a} \Big|_{-\infty}^0 \right) + a^{-1} \left(-\frac{1}{2} a e^{-2x/a} \Big|_0^{\infty} \right) = \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

得证。

练习：

对于一个测量时态函数为 $\Psi = (32a^3/\pi)^{1/4} x e^{-ax^2}$ 的系统（其中 $a = 1.0000 \text{ nm}^{-2}$ ），求测量值在 $x = 1.2000 \text{ nm}$ 和 1.2001 nm 之间的概率。将区间视为无穷小区间。

(答案：0.0000258。)

1.7 复数

我们已经知道波函数可以是复数，所以我们来回顾一下复数的一些性质。

复数 z 具有如下的形式：

$$z = x + iy, \quad \text{其中 } i \equiv \sqrt{-1} \quad (1.25)$$

且 x 和 y 是**实数**（不涉及负数平方根的数），如果在 (1.25) 中 $y = 0$ ，则 z 是**实数**。如果 $y \neq 0$ ，则 z 是**虚数**。如果 $x = 0$ 且 $y \neq 0$ ，则 z 是**纯虚数**。例如：6.83 是实数， $5.4 - 3i$ 是虚数，而 $0.60i$ 是纯虚数。实数和纯虚数是复数的特例，在 (1.25) 中， x 和 y 分别是 z 的实部和虚部，即 $x = \text{Re}(z)$ ； $y = \text{Im}(z)$ 。

复数可以代表复平面 (图 1.3) 上的一个点, 其中, z 的实部绘制在横轴上, 虚部绘制在纵轴上。这幅图显然提示我们定义两个表征复数 z 的量: 点 z 与原点的距离 r 称为 z 的绝对值或模; 点 z 的半径矢量与正横轴所成的角度 θ 则称为 z 的相位或辐角。我们有

$$|z| = r = (x^2 + y^2)^{1/2}, \quad \tan \theta = y/x \quad (1.26)$$

$$x = r \cos \theta, \quad y = r \sin \theta$$

所以我们可以把 $z = x + iy$ 写作

$$z = r \cos \theta + i \sin \theta = r e^{i\theta} \quad (1.27)$$

由欧拉公式带入。(见证明 4.3)

$$e^{i\theta} = \cos \theta + i \sin \theta \quad (1.28)$$

以上方程中的 θ 角均采用弧度制。

如果 $z = x + iy$, 则它的复共轭 z^* 定义为

$$z^* \equiv x - iy = r e^{-i\theta} \quad (1.29)$$



图 1.3: (a) 复数 $z = x + iy$ 的图象; (b) 复数 $-2 + i$ 的图象

如果 z 是实数, 那么它的虚部为 0。因此, 当且仅当 $z = z^*$ 时有 z 是实数。对复数 z 作两次复共轭, 我们再次得到了 z , 所以 $(z^*)^* = z$ 。将 z 与其复共轭 z^* 作乘积, 我们有

$$zz^* = (x + iy)(x - iy) = x^2 + iyx + ixy - i^2y^2$$

$$zz^* = x^2 + y^2 = r^2 = |z|^2 \quad (1.30)$$

对于两个复数 $z_1 = r_1 e^{i\theta_1}$ 与 $z_2 = r_2 e^{i\theta_2}$ 的积和商, 我们有

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}, \quad \frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)} \quad (1.31)$$

根据复共轭的定义或 (1.31), 我们很容易证明

$$(z_1 z_2)^* = z_1^* z_2^* \quad (1.32)$$

同样地,

$$(z_1 / z_2)^* = z_1^* / z_2^*, \quad (z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 - z_2)^* = z_1^* - z_2^* \quad (1.33)$$

对于积和商的绝对值, 由 (1.31) 可知

$$|z_1 z_2| = |z_1| |z_2|, \quad \left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|} \quad (1.34)$$

因此, 如果 ψ 是复波函数, 我们有

$$|\psi^2| = |\psi|^2 = \psi^* \psi \quad (1.35)$$

现在我们得到了数字 1 的 n 次根式。我们可以将数字 1 的相位取为 0 或 2π 或 4π , 以此类推。因此 $1 = e^{i2\pi k}$, 其中 k 是任何整数, 0、负数或正数。现在考虑数字 ω , 其中 $\omega \equiv e^{i2\pi k/n}$, n 为正整数。使用式 (1.31) n 次, 我们有 $\omega^n = e^{i2\pi k} = 1$ 。因此, ω 是 n 次单位根。有 n 个不同的复数 n 次单位根, 连续取 n 个整数 k 的值, 就可以得到所有这些根:

$$\omega = e^{i2\pi k/n}, \quad k = 0, 1, 2, \dots, n-1 \quad (1.36)$$

除了 (1.36) 中的 k 值之外, 任何其他 k 值所得到的数的相位都与 (1.36) 中的一个数相差 2 的整数倍, 因此不是不同的根。对于 (1.36) 中的 $n = 2$, 我们得到 1 的两个平方根; 对于 $n = 3$, 得到 1 的三个立方根; 以此类推。

1.8 单位

本书使用国际单位制。在国际单位制 (SI) 中, 长度、质量、时间的单位分别是米 (m)、千克 (kg) 和秒 (s)。力的单位是牛顿 (N), 能量的单位是焦耳 (J)。描述真空中相距 r 的两个电荷 Q_1 和 Q_2 之间作用力大小的库仑定律, 用国际单位制单位表示为

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 r^2} \quad (1.37)$$

其中电荷量 Q_1 和 Q_2 的单位是库仑 (C), ϵ_0 是常数 (真空介电常数), 其值为 $8.854 \times 10^{-12} \text{C}^2 \text{N}^{-1} \text{m}^{-2}$ 。(物理常数的精确值见附录)

1.9 微积分学

量子化学中大量使用微积分, 因此应牢记以下公式。其中 c 、 n 和 b 为常数, f 和 g 为 x 的函数。

$$\frac{dc}{dx} = 0, \quad \frac{d(cf)}{dx} = c \frac{df}{dx}, \quad \frac{dx^n}{dx} = nx^{n-1}, \quad \frac{de^{cx}}{dx} = ce^{cx}$$

$$\frac{d(\sin cx)}{dx} = c \cos cx, \quad \frac{d(\cos cx)}{dx} = -c \sin cx, \quad \frac{d \ln cx}{dx} = \frac{1}{x}$$

$$\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}, \quad \frac{d(fg)}{dx} = f \frac{dg}{dx} + g \frac{df}{dx}$$

$$\frac{d(f/g)}{dx} = \frac{d(fg^{-1})}{dx} = -fg^{-2} \frac{dg}{dx} + g^{-1} \frac{df}{dx}$$

$$\frac{d}{dx} f(g(x)) = \frac{df}{dg} \frac{dg}{dx}$$

最后一个公式有一个例子： $d[\sin(cx^2)]/dx = 2cx \cos(cx^2)$ 。这里 $g(x) = cx^2$ 而 $f = \sin$ 。

$$\int cf(x) dx = c \int f(x) dx, \quad \int [f(x) + g(x)] dx = \int f(x) dx + \int g(x) dx$$

$$\int dx = x, \quad \int x^n dx = \frac{x^{n+1}}{n+1} \text{ 其中 } n \neq -1, \quad \int \frac{1}{x} dx = \ln x$$

$$\int e^{cx} dx = \frac{e^{cx}}{c}, \quad \int \sin cx dx = -\frac{\cos cx}{c}, \quad \int \cos cx dx = \frac{\sin cx}{c}$$

总结

习题

Chapter 2

箱中粒子

单粒子一维系统的定态波函数和能级是通过求解定态薛定谔方程 (1.19) 得到的。在本章中，我们将求解一个简单系统的定态薛定谔方程，即一维盒子中的粒子（第 2.2 节）。由于薛定谔方程是微分方程，我们首先讨论微分方程。

2.1 微分方程

本节只讨论常微分方程，即只有一个自变量的方程。[偏微分方程有不只一个自变量。含时薛定谔方程 (1.16) 就是一个例子，其中 t 和 x 都是自变量。] 常微分方程是自变量 x 、因变量 $y(x)$ 和 y 的 1、2、...、 n 阶导数 $y(y', y'', \dots, y^{(n)})$ 之间的关系式。例如

$$y''' + 2x(y')^2 + y^2 \sin x = 3e^x \quad (2.1)$$

微分方程的阶数与方程中最高阶导数的阶数相同。因此，方程 (2.1) 为三阶微分方程。

线性微分方程是一种特殊的微分方程，其形式为

$$A_n(x)y^{(n)} + A_{n-1}(x)y^{(n-1)} + A_{n-2}(x)y^{(n-2)} + \dots + A_1(x)y' + A_0(x)y = g(x) \quad (2.2)$$

其中所有的 A_i 和 g （有些可能为零）只是 x 的函数。在 n 阶线性微分方程 (2.2) 中， y 及其各阶导数的次幂均为一次。不满足式 (2.2) 的微分方程为非线性微分方程。如果 (2.2) 中的 $g(x) = 0$ ，则线性微分方程是齐次的，否则是非齐次的。一维定态薛定谔方程 (1.19) 就是二阶线性齐次微分方程。

通过除以 y'' 的系数，我们可以将每一个二阶线性齐次微分方程化为以下形式

$$y'' + P(x)y' + Q(x)y = 0 \quad (2.3)$$

设 y_1 和 y_2 是两个满足方程 (2.3) 的独立函数。“独立”的意思是 y_2 并不是 y_1 的简单倍数。因此，线性齐次微分方程 (2.3) 的通解为

$$y = c_1 y_1 + c_2 y_2 \quad (2.4)$$

其中 c_1 和 c_2 是任意常数。将式 (2.4) 代入 (2.3) 的左边即可轻松证明：

$$\begin{aligned} & c_1 y_1'' + c_2 y_2'' + P(x)c_1 y_1' + P(x)c_2 y_2' + Q(x)c_1 y_1 + Q(x)c_2 y_2 \\ &= c_1 [y_1'' + P(x)y_1' + Q(x)y_1] + c_2 [y_2'' + P(x)y_2' + Q(x)y_2] \\ &= c_1 \cdot 0 + c_2 \cdot 0 = 0 \end{aligned} \quad (2.5)$$

其中我们用到了 y_1 和 y_2 是方程 (2.3) 的解。

n 次微分方程的通解通常有 n 个任意常数。为了确定这些常数，我们可能需要定解条件，即规定 y 或其各阶导数在某一点或多点的值的条件。例如：如果 y 是固定在两点上的振动弦的位移，我们就知道 y 在这两点上必须为零。

一个重要的特例是二阶常系数线性齐次微分方程：

$$y'' + py' + qy = 0 \quad (2.6)$$

其中 p 与 q 是常数。为了解这个方程 (2.6)，我们暂且假定解的形式是 $y = e^{sx}$ 。我们正在寻找一个导数与常数相乘会与原函数相互抵消的函数。指数函数在微分时会重复，因此是正确的选择。将其代入 (2.6) 即可得出

$$s^2 e^{sx} + p s e^{sx} + q e^{sx} = 0$$

$$\boxed{s^2 + ps + q = 0} \quad (2.7)$$

方程 (2.7) 称为特征方程。它是一个有两个根 s_1 和 s_2 的一元二次方程，只要 s_1 和 s_2 不相等，就能给出 (2.6) 的两个独立解。因此，(2.6) 的通解为

$$\boxed{y = c_1 e^{s_1 x} + c_2 e^{s_2 x}} \quad (2.8)$$

例如，对于方程 $y'' + 6y' - 7y = 0$ ，其特征方程为 $s^2 + 6s - 7 = 0$ 。解这个特征方程，有 $s_1 = 1$ ， $s_2 = -7$ ，所以通解为 $y = c_1 e^x + c_2 e^{-7x}$ 。

2.2 一维盒子中的粒子

本节求解一维盒子中粒子的定态薛定谔方程。我们所说的粒子是指受到特定势能函数作用的粒子，该势能函数沿 x 轴的任何位置都为无穷大，只有长度为 l 的线段除外，在该线段上势能为零。这样的系统在物理上似乎是不真实的，但这种模型可以成功地应用于某些共轭分子；见问题 2.17。我们把原点放在线段的左端（图 2.1）。



图 2.1: 一维盒子中粒子的势函数 $V(x)$

我们需要考虑三个区域。在区域 I 和 III 中，势能 V 为无穷大，此处的定态薛定谔方程 (1.19) 为

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} = (E - \infty) \psi$$

与 ∞ 相比， E 的大小可以忽略不计。我们略去 E ，有

$$\frac{d^2 \psi}{dx^2} = \infty \psi, \quad \psi = \frac{1}{\infty} \frac{d^2 \psi}{dx^2}$$

因此，在盒子的外侧， ψ 为零：

$$\psi_I = 0, \quad \psi_{II} = 0 \quad (2.9)$$

对于区域 II， x 在 0 和 l 之间时，势能 V 为零，定态薛定谔方程 (1.19) 为

$$\frac{d^2\psi_{II}}{dx^2} + \frac{2m}{\hbar^2} E \psi_{II} = 0 \quad (2.10)$$

其中 m 是粒子的质量， E 是能量。我们将方程 (2.10) 视为一个二阶常系数线性齐次微分方程。其特征方程由 (2.7) 给出：

$$s^2 + 2mE\hbar^{-2} = 0$$

$$s = \pm (-2mE)^{1/2} \hbar^{-1} \quad (2.11)$$

$$s = \pm i (2mE)^{1/2} \hbar^{-1} \quad (2.12)$$

其中 i 是虚数单位。使用 (2.8)，我们有

$$\psi_{II} = c_1 e^{i(2mE)^{1/2} x / \hbar} + c_2 e^{-i(2mE)^{1/2} x / \hbar} \quad (2.13)$$

我们姑且令

$$\theta \equiv (2mE)^{1/2} x / \hbar$$

$$\psi_{II} = c_1 e^{i\theta} + c_2 e^{-i\theta}$$

由 $e^{i\theta} = \cos \theta + i \sin \theta$ [1.28] 以及

$$\boxed{\cos(-\theta) = \cos \theta \text{ and } \sin(-\theta) = -\sin \theta} \quad (2.14)$$

我们有 $e^{-i\theta} = \cos(-\theta) + i \sin(-\theta) = \cos \theta - i \sin \theta$ 。因此，

$$\begin{aligned} \psi_{II} &= c_1 \cos \theta + i c_1 \sin \theta + c_2 \cos \theta - i c_2 \sin \theta \\ &= (c_1 + c_2) \cos \theta + (i c_1 - i c_2) \sin \theta \\ &= A \cos \theta + B \sin \theta \end{aligned}$$

其中， A 和 B 是新的任意常数。则

$$\psi_{II} = A \cos \left[\hbar^{-1} (2mE)^{1/2} x \right] + B \sin \left[\hbar^{-1} (2mE)^{1/2} x \right] \quad (2.15)$$

现在通过定解条件，我们来求出 A 和 B 。假设波函数是连续的，也就是说，它的值不会突然跳变，这似乎是合理的（见图 3.4）。如果 ψ 在 $x = 0$ 处连续，那么有 ψ_I 和 ψ_{II} 在 $x = 0$ 处的极限相同：

$$\begin{aligned} \lim_{x \rightarrow 0} \psi_I &= \lim_{x \rightarrow 0} \psi_{II} \\ 0 &= \lim_{x \rightarrow 0} \left\{ A \cos \left[\hbar^{-1} (2mE)^{1/2} x \right] + B \sin \left[\hbar^{-1} (2mE)^{1/2} x \right] \right\} \\ 0 &= A \end{aligned}$$

由于

$$\boxed{\sin 0 = 0, \quad \cos 0 = 1} \quad (2.16)$$

将 $A = 0$ 带入 (2.15)，有

$$\psi_{II} = B \sin \left[(2\pi/\hbar) (2mE)^{1/2} x \right] \quad (2.17)$$

又由函数在 $x = l$ 处连续，我们有

$$B \sin \left[(2\pi/\hbar) (2mE)^{1/2} l \right] = 0 \quad (2.18)$$

因为我们不能令波函数处处为零-得到一个空箱子, 所以 B 不能为零。有

$$\sin \left[(2\pi/h) (2mE)^{1/2} l \right] = 0$$

\sin 函数的零点位于 $0, \pm\pi, \pm2\pi, \dots, \pm n\pi$ 。则

$$(2\pi/h) (2mE)^{1/2} l = \pm n\pi \quad (2.19)$$

$n = 0$ 是一种特殊情况。由 (2.19), $n = 0$ 意味着 $E = 0$ 。若 $E = 0$ 则特征方程 (2.12) 有两个相等的根, 且式 (2.13) 不是薛定谔方程的完全解。为了找到完全解, 我们回到 (2.10), 其中 $E = 0$ 有 $\frac{d^2\psi_{II}}{dx^2} = 0$ 。积分一次, 有 $d\psi_{II}/dx = c$; 再积分一次, 有 $\psi_{II} = cx + d$, 其中 c 和 d 都是常数。由于 $x = 0$ 时有 $\psi_{II} = 0$, 所以 $d = 0$; 又由于 $x = l$ 时也有 $\psi_{II} = 0$, 所以 $c = 0$ 。因此, $\psi_{II} = 0$ 推出 $E = 0$, 则 $E = 0$ 不是允许的能量值。即 $n = 0$ 是不允许的。

从方程 (2.19) 中解出 E , 我们有

$$E = \frac{n^2 h^2}{8ml^2}, \quad n = 1, 2, 3 \dots \quad (2.20)$$

只有能量值满足式 (2.20) 的波函数 ψ , 才能符合在 $x = l$ 处连续的定解条件。定解条件的应用迫使我们得出能量值是量子化的结论 (图 2.2)。这与经典结果形成了鲜明对比, 经典结果认为盒子中的粒子可以具有任何非负能量。请注意: 粒子的能量有一个大于零的最小值。能量最小的状态称为**基态**。能量高于基态能量的状态为**激发态**。(在经典力学中, 粒子在盒子中的最低能量为零。经典粒子在盒内静止不动时, 动能为零, 势能为零)。

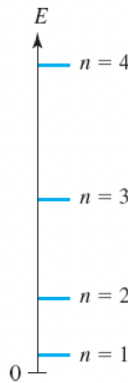


图 2.2: 一维盒子中粒子能量最低四个能级

例题:

有一质量为 2.00×10^{-26} g 的粒子在长度 4.00 nm 的势箱中运动。求该粒子从 $n = 3$ 能级进入 $n = 2$ 能级时发射光子的频率和波长。

根据能量守恒, 发射光子的能量等于两个定态之间的能量差 [公式 (1.4); 另见第 9.9 节]:

$$h\nu = E_{upper} - E_{lower} = \frac{n_u^2 h^2}{8ml^2} - \frac{n_l^2 h^2}{8ml^2}$$

$$\nu = \frac{(n_u^2 - n_l^2) h}{8ml^2} = \frac{(3^2 - 2^2) (6.626 \times 10^{-34} \text{ J} \cdot \text{s})}{8 (2.00 \times 10^{-29} \text{ kg}) (4.00 \times 10^{-9} \text{ m})^2} = 1.29 \times 10^{12} \text{ s}^{-1}$$

其中 u 和 l 分别代表高能级和低能级。由 $\lambda\nu = c$, 有 $\lambda = 2.32 \times 10^{-4} \text{ m}$ 。(学生常见的错误是将 $h\nu$ 设为其中一个状态的能量, 而不是状态间的能量差)

练习:

对于某个在一维盒子中运动的电子, 其能级跃迁的最长波长为 400 nm。求盒子的长度。

答案: 0.603 nm

将 (2.19) 带入 (2.17), 解出波函数为

$$\psi_{II} = B \sin\left(\frac{n\pi x}{l}\right), \quad n = 1, 2, 3 \dots \quad (2.21)$$

在 $n\pi$ 前面使用负号并不能得到另一个独立的解。因为 $\sin(-\theta) = -\sin\theta$, 我们只会得到一个常数 -1, 乘以带正号的解。

公式 (2.21) 中的常数 B 仍然是任意的。为了将其确定下来, 我们使用公式 (1.24) 和 (1.22) 对 ψ_{II} 进行归一化:

$$\begin{aligned} \int_{-\infty}^{\infty} |\Psi|^2 dx &= \int_{-\infty}^{\infty} |\psi|^2 dx = 1 \\ \int_{-\infty}^0 |\psi_I|^2 dx + \int_0^l |\psi_{II}|^2 dx + \int_l^{\infty} |\psi_{III}|^2 dx &= 1 \\ |B|^2 \int_0^l \sin^2\left(\frac{n\pi x}{l}\right) dx &= 1 = |B|^2 \frac{l}{2} \end{aligned} \quad (2.22)$$

其中的积分是用附录中的公式 (A.2) 求得的。我们有

$$|B|^2 = (2/l)^{1/2}$$

注意我们只求出了 B 的绝对值, 但 B 可以取 $\pm(2/l)^{1/2}$ 。此外, B 也不必是实数, 我们可以取任何模为 $(2/l)^{1/2}$ 的复数。所以, 我们可以说 $B = (2/l)^{1/2} e^{i\alpha}$, 其中 α 是 B 的辐角, 可以取 0 到 2π 之间的任何值 (第 1.7 节)。取辐角为 0, 则箱中粒子的定态波函数为

$$\psi_{II} = \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n\pi x}{l}\right), \quad n = 1, 2, 3 \dots \quad (2.23)$$

波函数和概率密度的图如图 2.3 和 2.4 所示。

能量 (2.20) 和波函数 (2.23) 中的数字 n 称为量子数。量子数 n 的每个不同值都会产生不同的波函数和不同的状态。

波函数在某些点为零, 这些点被称为节点。量子数 n 每增加 1, 波函数 ψ 就多一个节点。 ψ 和 $|\psi|^2$ 中节点的存在似乎令人惊讶。因此, 图 (2.4) 表明: 对于 $n = 2$, 在箱的中点 $x = l/2$ 处找到粒子的概率为零。粒子怎么可能从盒子的一侧运动到另一侧, 而在任何时候都不会出现在盒子的中心呢? 这个明显的悖论产生于我们试图用宏观粒子运动的日常经验来理解微观粒子的运动。然而, 正如第 1 章所指出的: 电子和其他微观“粒子”无法用宏观世界的经典物理学概念来完整、正确地描述。

图 2.4 显示, 在盒子的不同位置找到粒子的概率与经典结果截然不同。从经典角度看, 一个固定能量的粒子在盒子中以恒定的速度在两壁之间弹性地来回弹跳。因此, 在盒子中的任何一点发现它的可能性都是相同的; 从量子力学角度看, 我们会发现在盒子中心的最低能级出现概率最大。随着能级越来越高, 节点越来越多, 概率的最大值和最小值就会越来越接近, 而沿着盒子长度方向的

图 2.3: 三种能量最低的箱中粒子状态的 ψ 曲线图图 2.4: 三种能量最低的箱中粒子状态的 $|\psi|^2$ 曲线图

概率变化最终会变得难以察觉。当量子数非常大时，我们就获得了接近于经典的均匀概率密度结果。

这一结果，即在大量子数极限下量子力学与经典力学的关系，被称为玻尔对应原理。由于牛顿力学适用于宏观物体（运动速度远小于光速），我们期望非相对论量子力学能给出与经典力学对宏观物体相同的答案。由于普朗克常数极其微小，能量的量子化对于宏观物体来说是不可观测的。由于粒子的质量和盒子长度的平方出现在式 (2.20) 的分母中，因此在宏观盒子中具有宏观运动能量的宏观物体将具有巨大的 n 值，那么根据对应原理，宏观物体将显示出经典行为。

现在，我们有了一整套波函数，每个波函数对应不同的能量，并以量子数 n （正整数）为特征。让下标 i 表示量子数为 n_i 的特定波函数：

$$\psi_i = \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n_i \pi x}{l}\right), \quad 0 < x < l$$

$$\psi_i = 0, \quad \text{其他区域}$$

由于波函数已归一化，我们有

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = 1, \quad \text{若 } i = j \quad (2.24)$$

现在我们想知道，当我们使用对应于不同能级的波函数时，这个积分的值是多少：

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = \int_0^l \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n_i \pi x}{l}\right) \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n_j \pi x}{l}\right) dx, \quad n_i \neq n_j$$

使用附录中的公式 (A.5)，我们有

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = \frac{2}{l} \left[\frac{\sin[(n_i - n_j)\pi]}{2(n_i - n_j)\pi/l} - \frac{\sin[(n_i + n_j)\pi]}{2(n_i + n_j)\pi/l} \right] = 0 \quad (2.25)$$

由于当 m 是整数时有 $\sin m\pi = 0$ ，则

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = 0, \quad i \neq j \quad (2.26)$$

当 (2.26) 成立时, 我们说对 $i \neq j$ 的波函数 ψ_i 和 ψ_j 是互相正交的。将 (2.24) 与 (2.26) 相结合, 我们有

$$\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = \delta_{ij} \quad (2.27)$$

其中 δ_{ij} 是 **Kronecker 符号** (由数学家的名字命名)。当 i 和 j 相等时, 它等于 1; 当 i 和 j 不相等时, 它等于 0:

$$\delta_{ij} \equiv \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases} \quad (2.28)$$

波函数的特性 (2.27) 称为正交归一性。我们只证明了粒子盒内波函数的正交归一性。我们将在第 7.2 节更广泛地证明它。

您可能会对式 (2.26) 感到困惑, 不明白为什么我们要将一个状态的波函数乘以另一个状态的波函数。我们稍后会看到 (例如第 7.3 节), 使用包含一个系统所有波函数之和的方程通常很有帮助, 而这样的方程会导致像 (2.26) 这样的积分。

研究无限深势阱中粒子的一个更严谨的方法是: 首先处理粒子在壁箱中的势能有限跃迁, 然后当 V 的跃迁变得无限大时求取极限。取极限时的结果与 (2.20) 和 (2.23) 相同 (见问题 2.22)。

我们只考虑了粒子在一维盒子中的定态。该系统的非稳态示例, 请参见第 7.8 节末尾的示例。

关于粒子在盒子中的在线计算机模拟, 请访问 www.chem.uci.edu/undergraduate/applets/dwell/dwell.htm (显示了在盒子中间引入高度和宽度可变的屏障时对波函数和能级的影响); web.williams.edu/wpetc/chemistry/dbingemann/Chem153/particle.html (通过绘制薛定谔方程在能量变化和盒子长度变化时的解来显示量子化); 以及 falstad.com/qm1d/ (同时显示含时薛定谔方程和定态薛定谔方程; 见 7.47)。

2.3 一维自由粒子

我们所说的自由粒子是指不受任何力作用的粒子。对于自由粒子, 对 (1.12) 的积分表明, 无论 x 的值是多少, 势能都保持不变。由于基态能量的选择是任意的, 我们可以设定 $V(x) = 0$ 。则薛定谔方程 (1.19) 变为

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} E\psi = 0 \quad (2.29)$$

式 (2.29) 与式 (2.10) 相同 (定解条件除外)。因此, (2.29) 的通解为 (2.13):

$$\psi = c_1 e^{i(2mE)^{1/2} x/\hbar} + c_2 e^{-i(2mE)^{1/2} x/\hbar} \quad (2.30)$$

我们可以使用什么边界条件呢? 似乎可以合理地假设 (因为 $\psi^* \psi dx$ 代表一种概率), 当 x 趋于 $\pm\infty$ 时, ψ 将保持有限值。那么如果能量 E 小于零, 那么就违反了边界条件, 因为对于 $E < 0$, 我们有

$$i(2mE)^{1/2} = i(-2m|E|)^{1/2} = i \cdot i \cdot (2m|E|)^{1/2} = -(2m|E|)^{1/2}$$

因此, 当 x 接近负无穷大时, (2.30) 中的第一项将变为无穷大。同样, 如果 E 为负值, 当 x 接近正无穷大时, (2.30) 中的第二项也将变为无穷大。因此, 对于一维自由粒子, 定解条件要求

$$E \geq 0 \quad (2.31)$$

波函数是振荡的, 是一个正弦项和一个余弦项的线性组合 [公式 (2.15)]。对于自由粒子, 能量没有量子化, 所有非负能量都是允许的。在这种情况下, 由于我们规定 $V = 0$, 能量 E 全为动能。如

果我们想通过归一化的方式求出常数 c_1 和 c_2 ，会发现积分 $\int_{-\infty}^{\infty} \psi^*(x) \psi(x) dx$ 的值为无穷大。换句话说，自由粒子的波函数在通常意义上是不可归一化的。这在物理上是意料之中的，因为当 x 变为 $\pm\infty$ 时，找到自由粒子的概率没有理由趋近于零。

自由粒子问题是一种不真实的情况，因为我们实际上不可能有一种粒子与宇宙中的任何其他粒子没有相互作用。

2.4 矩形势阱中的粒子

考虑在一个具有有限势垒的一维盒子中的粒子（图 2.5 a）。势能函数为当 $x < 0$ 和 $x > l$ 时 $V = V_0$ ，而 $0 \leq x \leq l$ 时有 $V_0 = 0$ 。需要考虑两种情况，即粒子的能量高于或低于 V_0 。

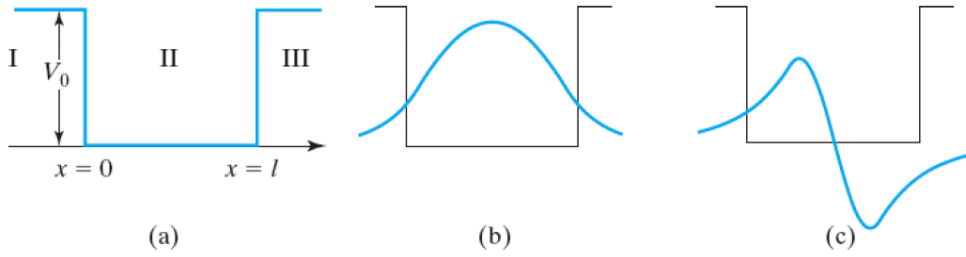


图 2.5: (a) 一维矩形势阱中粒子的势能。(b) 该势能的基态波函数。(c) 第一激发态的波函数。

我们首先考虑 $E < V_0$ 的情况。区域 I 和区域 III 的薛定谔方程为 $d^2\psi/dx^2 + (2m/\hbar^2)(E - V_0)\psi = 0$ 。这是一个常系数齐次线性微分方程，其特征方程 (2.7) 为 $s^2 + (2m/\hbar^2)(E - V_0) = 0$ ，特征根为 $s = \pm (2m/\hbar^2)^{1/2}(V_0 - E)^{1/2}$ 。因此，

$$\begin{aligned}\psi_I &= C \exp \left[(2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right] + D \exp \left[- (2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right] \\ \psi_{III} &= F \exp \left[(2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right] + G \exp \left[- (2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right]\end{aligned}$$

其中 C 、 D 、 F 和 G 都是常数。

如第 2.3 节所述，当 $x \rightarrow -\infty$ 时，我们必须阻止 ψ_I 趋于无穷。由于我们假设 $E < V_0$ ，则 $(V_0 - E)^{1/2}$ 是正实数，当 $x \rightarrow -\infty$ 时，为了满足 ψ_I 的值为有限值，我们必须有 $D = 0$ 。同样地，当 $x \rightarrow +\infty$ 时，为了满足 ψ_{III} 的值为有限值，我们有 $F = 0$ 。因此，

$$\psi_I = C \exp \left[(2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right], \quad \psi_{III} = G \exp \left[- (2m/\hbar^2)^{1/2} (V_0 - E)^{1/2} x \right]$$

在区域 II 中， $V = 0$ ，薛定谔方程与 (2.10) 一致，其通解为 (2.15)

$$\psi_{II} = A \cos \left[(2m\hbar^2)^{1/2} E^{1/2} x \right] + B \sin \left[(2m\hbar^2)^{1/2} E^{1/2} x \right] \quad (2.32)$$

为了完整解决这个问题，我们需要用到定解条件。与无限深势阱中粒子的情况类似，我们需要波函数在 $x = 0$ 和 $x = l$ 处连续，则 $\psi_I(0) = \psi_{II}(0)$ ， $\psi_{II}(l) = \psi_{III}(l)$ 。波函数有四个任意常数，因此需要的定解条件不止这两个。除了需要 ψ 连续以外，我们还需要其一阶导数 $d\psi/dx$ 处处连续。为了证明这一要求的合理性，我们假设如果 $d\psi/dx$ 在某点处不连续，则其导数（ ψ 的二阶导数的瞬时变化率）在该点处会趋于无穷。然而，对于矩形势阱中的粒子，其薛定谔方程 $d^2\psi/dx^2 = (2m/\hbar^2)(V - E)\psi$ 的右侧不包含任何趋于无穷的因子。所以 $d^2\psi/dx^2$ 不可能趋于无穷 [更严谨的论证，见 D. Branson, Am. J. Phys., 47, 1000 (1979)]。因此，我们在 $x = 0$ 时有 $d\psi_I/dx = d\psi_{II}/dx$ ，在 $x = l$ 时有 $d\psi_{II}/dx = d\psi_{III}/dx$ 。

由 $\psi_I(0) = \psi_{II}(0)$, 我们有 $C = A$ 。由 $\psi'_I(0) = \psi'_{II}(0)$, 我们有 $B = (V_0 - E)^{1/2} A / E^{1/2}$ 。由 $\psi_{II}(l) = \psi_{III}(l)$, 我们得到了一个复杂的等式, 可以根据 A 求出 G 。常数 A 由归一化计算求出。

用 $\psi'_{II}(l) = \psi'_{III}(l)$ 除以 $\psi_{II}(l) = \psi_{III}(l)$, 再用 A 表示 B , 就得到下面的能级方程 (问题 2.21b):

$$(2E - V_0) \sin \left[(2mE)^{1/2} l / \hbar \right] = 2(V_0 E - E^2)^{1/2} \cos \left[(2mE)^{1/2} l / \hbar \right] \quad (2.33)$$

[虽然 $E = 0$ 也满足 (2.33), 但会导致 $\psi = 0$, 所以它不是允许的能量。(见问题 2.21 b) 定义无量纲的常数 ε 和 b 满足

$$\varepsilon \equiv E/V_0, \quad b \equiv (2mV_0)^{1/2} l / \hbar \quad (2.34)$$

用 V_0 除以 (2.33), 我们有

$$(2\varepsilon - 1) \sin(b\varepsilon^{1/2}) - 2(\varepsilon - \varepsilon^2)^{1/2} \cos(b\varepsilon^{1/2}) = 0 \quad (2.35)$$

只有满足 (2.33) 的粒子的能量才有连续和连续导数的波函数, 所以能级量子化的条件为 $E < V_0$, 为了找到允许的能级, 我们可以在 $0 < \varepsilon < 1$ 的条件下绘制 (2.35) 左侧表达式与 ε 的关系图, 并找出曲线与横轴交叉的点 (另见问题 4.31c)。详细研究 (Merzbacher, 第 6.8 节) 表明: $E < V_0$ 的允许能级数为 N , 其中 N 满足以下条件

$$N - 1 < b/\pi \leq N, \quad \text{其中 } b \equiv (2mV_0)^{1/2} l / \hbar \quad (2.36)$$

例如, 如果 $V_0 = \hbar^2/ml^2$, 那么 $b/\pi = 2(2^{1/2}) = 2.83$, 即 $N = 3$ 。图 (2.5) 显示了最低两个能级的 ψ 。波函数在盒内是振荡的, 而在盒外则呈指数衰减。事实证明: 每提高一个能级, 节点数就增加一个。

到目前为止, 我们只考虑了 $E < V_0$ 的状态。对于 $V_0 < E$, $(V_0 - E)^{1/2}$ 是虚数, 并且波函数 ψ_I 和 ψ_{III} 在 $x \rightarrow \pm\infty$ 时不会趋于零, 而是会振荡 (类似于自由粒子)。我们也再也没有理由将 ψ_I 中的 D 和 ψ_{III} 中的 F 设为零了, 这些额外的常数可以满足 ψ 和 ψ' 的定解条件, 我们会发现不需要限制 E 就可以得到符合条件的波函数。因此, 上述所有能量都是允许的。

$x \rightarrow \pm\infty$ 时 $\psi \rightarrow 0$ 的状态称为**束缚态**。对于束缚态, 只有在有限的空间区域内才有可能找到粒子。对于**非束缚态**, 当 $x \rightarrow \pm\infty$ 时 ψ 不会趋于零, 且无法归一化。对于在有限深矩形势阱中的粒子, $E < V_0$ 的状态是束缚态, 而 $E > V_0$ 的状态是非束缚态。而在无限深势阱中的粒子则所有的状态都为束缚态, 自由粒子则全为非束缚态。

要在线模拟粒子在势阱中的情况, 请访问 www.falstad.com/qm1d, 并在“设置”框中选择“有限势阱”。您可以改变势阱的宽度和深度, 查看对能级和波函数的影响。

2.5 隧穿

对于有限深矩形势阱中的粒子 (第 2.4 节) 图 (2.5), 波函数 ψ_I 和 ψ_{III} 表明: 就束缚态而言, 在粒子总能量 E 小于其势能 $V = V_0$ 的区域 I 和 III 中找到粒子的概率不为零。经典上, 这是不允许的。经典方程 $E = T + V$ 且 $T \geq 0$, 其中 T 是动能, 则经典力学中 E 是不可能比 V 小的。

考虑一个在有限宽度和深度盒子中作一维运动的粒子, 从经典角度看, 除非粒子的能量大于势能势垒 V_0 , 否则粒子无法逸出。然而, 量子力学的处理方法 (此处省略) 表明, 在盒子外发现总能量小于 V 的粒子的概率是不为零的。

隧穿 (tunneling) 一词表示粒子穿透经典禁区 (如图 2.5 所示), 或粒子穿过高度超过粒子能量的能垒。由于隧穿是一种量子效应, 因此粒子的行为越背离经典力学, 其发生的概率就越大。因此, 质量小的粒子最容易发生隧穿现象 (注意: 质量 m 越大, 第 2.4 节中的 ψ_I 和 ψ_{II} 函数趋于零的速度

就越快)。电子很容易发生隧穿。氢原子和离子比更重的原子更容易发生隧穿。放射性原子核发射 α 粒子时, α 粒子要通过子核与 α 粒子之间的短程核吸引力和库仑斥力所产生的势能势垒才能发生隧穿。 NH_3 是金字塔形分子, 分子反转存在势能障碍, 平面构型时势能最大。氢原子可以穿透这个势垒, 从而使分子倒转。在乙烷分子中, 内旋转存在障碍, 在氢的黯淡位置存在势能最大值。氢原子可以从一个交错位置隧穿到下一个位置。电子的隧穿在氧化还原反应和电极过程中非常重要。隧穿通常对涉及氢原子转移的化学反应的速率有很大贡献。参见 R. P. Bell, *The Tunnel Effect in Chemistry*, Chapman & Hall, 1980。



图 2.6: 有限高度和宽度的一维盒子中粒子的势能函数

某些酶催化反应中会出现 H 原子的隧穿现象; 见《酶催化反应中的量子隧穿》, R. Allemann 和 N. Scrutton (编), RSCP Publishing 出版社, 2009 年。

扫描隧穿显微镜发明于 1981 年, 它利用电子在极细的金属丝尖端与导电固体表面之间的空间隧穿, 生成固体表面单个原子的图像。在固体和金属丝之间施加一个微小的电压, 当针尖以几 \AA 的高度在表面上移动时, 针尖的高度会进行调整, 以保持电流恒定。根据针尖高度与位置的关系绘制出表面图像。

总结

习题

Chapter 3

算符

3.1 算符

现在，我们以一种比以前更一般的方式来发展量子力学理论。首先，我们将单粒子一维定态薛定谔方程 (1.19) 写成以下形式

$$\left[-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x) \right] \psi(x) = E\psi(x) \quad (3.1)$$

(3.1) 中括号内的实体是一个算符 (operator)。方程 (3.1) 表明我们有一个能量算符，它对波函数起作用后，波函数又回来了，不过是乘以一个允许的能量值。因此，接下来我们讨论算符。

算符是一种将给定函数转换为另一个函数的规则。例如，让 \hat{D} 算符来表示对关于 x 的函数求导。我们用一个小尖号来表示算符。给定一个可导函数 $f(x)$ ，则将 \hat{D} 作用到 $f(x)$ 上即可表示为 $\hat{D}f(x) = f'(x)$ 。例如， $\hat{D}(x^2 + 3e^{2x}) = 2x + 6e^{2x}$ 。如果 $\hat{3}$ 是将函数乘以 3 的算符，那么 $\hat{3}(x^2 + 3e^x) = 3x^2 + 9e^x$ 。如果 \tan 是取函数正切值的算符，那么将 \tan 应用于函数 $x^2 + 1$ ，我们得到了 $\tan(x^2 + 1)$ 。如果算符 \hat{A} 将函数 $f(x)$ 转换为另一个函数 $g(x)$ ，我们将其写作 $\hat{A}f(x) = g(x)$ 。

我们将两个算符 \hat{A} 和 \hat{B} 的**和与差**定义为

$$\begin{aligned} (\hat{A} + \hat{B})f(x) &\equiv \hat{A}f(x) + \hat{B}f(x) \\ (\hat{A} - \hat{B})f(x) &\equiv \hat{A}f(x) - \hat{B}f(x) \end{aligned} \quad (3.2)$$

例如，若 $\hat{D} \equiv d/dx$ ，那么

$$(\hat{D} + \hat{3})(x^3 - 5) = \hat{D}(x^3 - 5) + \hat{3}(x^3 - 5) = 3x^2 + (3x^3 - 15) = 3x^3 + 3x^2 - 15$$

一个算符也可以涉及多个变量。例如，算符 $\partial^2/\partial x^2 + \partial^2/\partial y^2$ 具有如下性质：

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2)g(x, y) = \partial^2 g/\partial x^2 + \partial^2 g/\partial y^2$$

两个算符的**积**定义为

$$\hat{A}\hat{B}f(x) = \hat{A}[\hat{B}f(x)] \quad (3.3)$$

换句话说，我们先将乘积右边的算符作用到函数 $f(x)$ 上，然后将左边的算符作用到得到的函数上。例如， $\hat{3}\hat{D}f(x) = \hat{3}[\hat{D}f(x)] = \hat{3}f'(x) = 3f'(x)$

算符 $\hat{A}\hat{B}$ 和 $\hat{B}\hat{A}$ 可能不会有同样的效果。例如，考虑算符 d/dx 和 \hat{x} (\hat{x} 表示将函数乘以 x)：

$$\hat{D}\hat{x}f(x) = \frac{d}{dx}[xf(x)] = f(x) + f'(x) = (\hat{1} + \hat{x}\hat{D})f(x) \quad (3.4)$$

$$\hat{x}\hat{D}f(x) = \hat{x}\left[\frac{d}{dx}f(x)\right] = xf'(x)$$

在这个例子中, $\hat{A}\hat{B}$ 和 $\hat{B}\hat{A}$ 是不同的算符。

我们可以建立如下的**算符代数**。若对任意的函数 f , 有 $\hat{A}f = \hat{B}f$, 则我们说算符 \hat{A} 和算符 \hat{B} 是**相等的**。相等的算符作用到给定的函数上时, 会有相同的结果。例如, 式 (3.4) 表明

$$\hat{D}\hat{x} = 1 + \hat{x}\hat{D} \quad (3.5)$$

算符 $\hat{1}$ (乘以 1) 称作**单位算符**。算符 $\hat{0}$ (乘以 0) 称作**空算符**。我们通常省略简单常数算符顶上的小尖号。我们还可以将算符从算符方程的一边转移到另一边 (问题 3.7)。因此, 式 (3.5) 与 $\hat{D}\hat{x} - \hat{x}\hat{D} - 1 = 0$ 等价, 其中省略了空算符和单位算符上的小尖号。

算符运算遵循结合律:

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \quad (3.6)$$

式 (3.6) 的证明见问题 3.10。例如, 令 $\hat{A} = d/dx$, $\hat{B} = \hat{x}$, 而 $\hat{C} = 3$ 。使用式 (3.5), 我们有

$$\begin{aligned} (\hat{A}\hat{B}) &= \hat{D}\hat{x} = 1 + \hat{x}\hat{D}, & [(\hat{A}\hat{B})\hat{C}]f &= (1 + \hat{x}\hat{D})3f = 3f + 3xf' \\ (\hat{B}\hat{C}) &= 3\hat{x}, & [\hat{A}(\hat{B}\hat{C})]f &= \hat{D}(3xf) = 3f + 3xf' \end{aligned}$$

算符代数与普通代数的一个主要区别是: 数的运算遵守乘法的交换律, 但算符不一定。例如: 若 a 和 b 都是数, 那么有 $ab = ba$; 但是 $\hat{A}\hat{B}$ 和 $\hat{B}\hat{A}$ 不一定相等。我们定义算符 \hat{A} 和 \hat{B} 的**交换子** (commutator) $[\hat{A}, \hat{B}]$ 为 $\hat{A}\hat{B} - \hat{B}\hat{A}$:

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \quad (3.7)$$

如果有 $\hat{A}\hat{B} = \hat{B}\hat{A}$, 那么 $[\hat{A}, \hat{B}] = 0$, 则我们说算符 \hat{A} 和 \hat{B} 是**可对易的**。如果 $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, 那么 \hat{A} 和 \hat{B} 是**不可对易的**。注意 $[\hat{A}, \hat{B}]f = \hat{A}\hat{B}f - \hat{B}\hat{A}f$ 。由于我们作用算符 3 和 d/dx 的顺序没有区别, 因此我们有

$$\left[\hat{3}, \frac{d}{dx}\right] = \hat{3}\frac{d}{dx} - \frac{d}{dx}\hat{3} = 0$$

从式 (3.5), 我们有

$$\left[\frac{d}{dx}, \hat{x}\right] = \hat{D}\hat{x} - \hat{x}\hat{D} = 1 \quad (3.8)$$

那么算符 d/dx 和 \hat{x} 是**不可对易的**。

例题: 求 $[z^3, d/dz]$ 。

为了求出 $[z^3, d/dz]$, 我们需要将其作用到任意一个函数 $g(x)$ 上。由式 (3.7) 交换子的定义及算符差和乘积的定义, 我们有

$$\begin{aligned} [z^3, d/dz]g &= [z^3(d/dz) - (d/dz)z^3]g = z^3(d/dz)g - (d/dz)(z^3g) \\ &= z^3g' - 3z^2g - z^3g' = -3z^2g \end{aligned}$$

删去任意的函数 g , 我们有算符方程 $[z^3, d/dz] = -3z^2$ 。

练习:

求 $[d/dx, 5x^2 + 3x + 4]$ 。(答案: $10x + 3$)

算符的平方定义为算符和它本身的乘积： $\hat{B}^2 = \hat{B}\hat{B}$ 。我们来求求导算符的平方：

$$\begin{aligned}\hat{D}^2 f(x) &= \hat{D}(\hat{D}f) = \hat{D}f' = f'' \\ \hat{D}^2 &= d^2/dx^2\end{aligned}$$

再比如：取函数复共轭的算符的平方等于单位算符，因为取两次复共轭可以得到原始函数。算符 \hat{B}^n ($n = 1, 2, 3 \dots$) 定义为连续应用 n 次算符 \hat{B} 。

事实表明，量子力学中出现的算符都是线性的。当且仅当算符 \hat{A} 具有以下两个性质时，它是线性算符：

$$\hat{A}[f(x) + g(x)] = \hat{A}f(x) + \hat{A}g(x) \quad (3.9)$$

$$\hat{A}[cf(x)] = c\hat{A}f(x) \quad (3.10)$$

其中 f 和 g 是任意函数， c 是任意常数（可以不必是实数）。线性算符的例子有 \hat{x}^2 、 d/dx 和 d^2/dx^2 。非线性算符有 \cos 和 $(\)^2$ ，其中 $(\)^2$ 是对作用的函数求平方。

例题：

d/dx 是线性算符吗？ $\sqrt{\quad}$ 是线性算符吗？

我们有

$$(d/dx)[f(x) + g(x)] = df/dx + dg/dx = (d/dx)f(x) + (d/dx)g(x)$$

$$(d/dx)[cf(x)] = cdf/dx = c(d/dx)f(x)$$

d/dx 符合式 (3.9) 和 (3.10)，是线性算符。然而，

$$\sqrt{f(x) + g(x)} \neq \sqrt{f(x)} + \sqrt{g(x)}$$

所以 $\sqrt{\quad}$ 不是线性算符。

练习：算符 $x^2 \times$ （乘以 x^2 ）是线性算符吗？（答案：是）

线性算符运算中有用的算符有

$$(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C} \quad (3.11)$$

$$\hat{A}(\hat{B} + \hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C} \quad (3.12)$$

例题：证明线性算符满足分配律 (3.11)。

开始证明的好方法是先写下给出的内容和要证明的内容。已知 \hat{A} 、 \hat{B} 和 \hat{C} 是线性算符，求证 $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$ 。

为了证明算符 $(\hat{A} + \hat{B})\hat{C}$ 等于算符 $\hat{A}\hat{C} + \hat{B}\hat{C}$ ，我们必须证明这两个算符作用到任意函数 f 上的结果是相等的。即

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

我们从左边 $[(\hat{A} + \hat{B})\hat{C}]f$ 开始。这个表达式包含了算符 \hat{A} 与 \hat{B} 的乘积及算符 \hat{C} 。将算符乘积的定义 (3.3) 中的 \hat{A} 用 $\hat{A} + \hat{B}$ 替代， \hat{B} 用 \hat{C} 替代，得出 $[(\hat{A} + \hat{B})\hat{C}]f = (\hat{A} + \hat{B})(\hat{C}f)$ 。将 $\hat{C}f$ 整体视为一个函数，使用式 (3.2) 关于两个算符 \hat{A} 与 \hat{B} 的和 $\hat{A} + \hat{B}$ 的定义，得出

$(\hat{A} + \hat{B})(\hat{C}f) = \hat{A}(\hat{C}f) + \hat{B}(\hat{C}f)$ 。因此，

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

使用式 (3.3) 给出的算符积的定义，有 $\hat{A}(\hat{C}f) = \hat{A}\hat{C}f$ 及 $\hat{B}(\hat{C}f) = \hat{B}\hat{C}f$ 。则

$$[(\hat{A} + \hat{B})\hat{C}]f = \hat{A}\hat{C}f + \hat{B}\hat{C}f \quad (3.13)$$

将算符和的定义 (3.2) 中的 \hat{A} 用 $\hat{A}\hat{C}$ 替代， \hat{B} 用 $\hat{B}\hat{C}$ 替代，得出 $(\hat{A}\hat{C} + \hat{B}\hat{C})f = \hat{A}\hat{C}f + \hat{B}\hat{C}f$ ，则式 (3.13) 变成了

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

这就是我们要证明的。因此， $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$ 。

注意我们没有用到 \hat{A} 、 \hat{B} 和 \hat{C} 是线性算符这一条件，事实上，式 (3.11) 对任意算符都成立。但式 (3.12) 只在 \hat{A} 是线性算符时成立。（问题 3.17）

例题：求算符 $d/dx + \hat{x}$ 的平方。

为了求 $(d/dx + \hat{x})^2$ ，我们将其作用到任意一个函数 $f(x)$ 上。令 $\hat{D} \equiv d/dx$ ，我们有

$$\begin{aligned} (\hat{D} + \hat{x})^2 f(x) &= (\hat{D} + \hat{x}) [(\hat{D} + \hat{x})f] = (\hat{D} + \hat{x})(f' + xf) \\ &= f'' + f + xf' + xf' + x^2 f = (\hat{D}^2 + 2\hat{x}\hat{D} + \hat{x}^2 + 1)f(x) \end{aligned}$$

$$(\hat{D} + \hat{x})^2 = \hat{D}^2 + 2\hat{x}\hat{D} + \hat{x}^2 + 1$$

让我们只使用算符方程重复这一计算：

$$\begin{aligned} (\hat{D} + \hat{x})^2 &= (\hat{D} + \hat{x})(\hat{D} + \hat{x}) = \hat{D}(\hat{D} + \hat{x}) + \hat{x}(\hat{D} + \hat{x}) \\ &= \hat{D}^2 + \hat{D}\hat{x} + \hat{x}\hat{D} + \hat{x}^2 = \hat{D}^2 + \hat{x}\hat{D} + 1 + \hat{x}\hat{D} + \hat{x}^2 \\ &= \hat{D}^2 + 2\hat{x}\hat{D} + \hat{x}^2 + 1 \end{aligned}$$

其中用到了式 (3.11)、(3.12) 和 (3.5)，并省略了“乘以 x ”算符上的小尖号。在彻底掌握算符之前，最安全的做法是在进行算符操作时，始终保持算符对任意函数 f 进行操作，然后在最后删除 f 。

练习：

求 $(d^2/dx^2)^2$ 。（答案： $d^4/dx^4 + 2xd^2/dx^2 + 2d/dx + x^2$ ）

3.2 本征函数和本征值

假设某线性算符 \hat{A} 作用在某个函数 $f(x)$ 上的结果是对 $f(x)$ 乘上一个常数 k 。我们说 $f(x)$ 是 \hat{A} 的一个**本征函数** (eigenfunction)， k 是**本征值** (eigenvalue)。作为定义的一部分，我们将要求本征函数 $f(x)$ 不等于零。我们的意思是说，尽管 $f(x)$ 可能在不同点上消失，但它并非处处为零。我们有

$$\boxed{\hat{A}f(x) = kf(x)} \quad (3.14)$$

例如, 函数 e^{2x} 是算符 d/dx 的一个本征函数, 其本征值为 2:

$$(d/dx) e^{2x} = 2e^{2x}$$

然而 $\sin 2x$ 不是 d/dx 的本征函数, 因为 $(d/dx)(\sin 2x) = 2 \cos 2x$, 不等于一个常数乘以 $\sin 2x$ 。

例题: 如果函数 $f(x)$ 是线性算符 \hat{A} 的一个本征函数, c 是任意常数。求证: $cf(x)$ 也是算符 \hat{A} 的本征函数, 并与 $f(x)$ 有相同的本征值。

了解如何进行证明的一个好方法是按以下步骤进行:

1. 写下给定的信息, 并将这些信息从文字转化为等式。
2. 用一个或多个方程的形式写下要证明的内容。
3. (a) 处理步骤 1 中的给定方程, 将其转化为步骤 2 中的所需方程。(b) 或者, 从我们要证明的方程的一边开始, 使用步骤 1 中的给定方程来处理这一边, 直到它转化为要证明的方程的另一边。

我们三个条件: $f(x)$ 是算符 \hat{A} 的本征函数; \hat{A} 是线性算符; c 是常数。将这些表述转化为方程 [见方程 (3.14)、(3.9) 和 (3.10)], 我们有

$$\hat{A}f = kf \quad (3.15)$$

$$\hat{A}(f+g) = \hat{A}f + \hat{A}g, \quad \text{及} \quad \hat{A}(bf) = b\hat{A}f \quad (3.16)$$

$$c = a \text{ (常数)}$$

其中 k 和 b 是常数, f 和 g 是函数。

我们想要证明 $cf(x)$ 也是算符 \hat{A} 的本征函数且有相同的本征值, 将其写成方程, 即

$$\hat{A}(cf) = k(cf)$$

使用策略 (3) 的 (b), 我们从待证明方程的左侧 $\hat{A}(cf)$ 出发, 尝试证明它等于 $k(cf)$ 。由线性算符定义 (3.16) 的第二个方程, 我们有 $\hat{A}(cf) = c\hat{A}f$ 。使用本征方程 (3.15), 有 $c\hat{A}f = ckf$ 。因此,

$$\hat{A}(cf) = c\hat{A}f = ckf = k(cf)$$

得证。

例题: (a) 求算符 d/dx 的本征函数及对应的本征值; (b) 如果我们假设定解条件为本征函数在 $x \rightarrow \pm\infty$ 时为有限值, 求对应的本征值。

(a) 将算符 $\hat{A} = d/dx$ 带入方程 (3.14), 有

$$\frac{df(x)}{dx} = kf(x) \quad (3.17)$$

$$\frac{1}{f}df = kdx$$

求积分, 有

$$\ln f = kx + \text{常数}$$

$$f = e^{\text{常数} \cdot kx}$$

$$f = ce^{kx} \quad (3.18)$$

算符 d/dx 的本征函数由 (3.18), 本征值为 k , 它可以是任何数字, 而 (3.17) 仍然成立。本征函数包含一个任意的常数 c 。这对每个线性算符的特征函数都是正确的, 正如前面的例子所证明的。(3.18) 中 k 的每个不同值都会产生不同的特征函数。然而, k 值相同而 c 值不同的特征函数并不是相互独立的。

(b) 由于 k 可以是复数, 我们可以写作 $k = a + ib$, 其中 a 和 b 是实数。我们有 $f(x) = ce^{ax}e^{ibx}$ 。若 $a > 0$, 则当 x 趋于无穷时因子 e^{ax} 也趋于无穷。若 $a < 0$, 则 $x \rightarrow -\infty$ 时 $e^{ax} \rightarrow \infty$ 。因此, 这两个定解条件要求 $a = 0$, 所以本征值 $k = ib$, 其中 b 是实数。

在第 3.1 节的第一个例题中, 我们知道对任意的函数 g , 有 $[z^3, d/dz]g(z) = -3z^2g(z)$, 所以 $[z^3, d/dz] = -3z^2$ 。相反地, 本征方程 $\hat{A}f(x) = kf(x)$ [式 (3.14)] 却不一定对任意函数 $f(x)$ 成立, 从这个方程中我们并不能得出 $\hat{A} = k$ 。因此, 结论 $(d/dx)e^{2x} = 2e^{2x}$ 并不意味着 d/dx 与乘以 2 相等。

3.3 算符和量子力学

我们现在考虑算符和量子力学之间的关系。将式 (3.1) 与 (3.14) 比较, 我们可以发现薛定谔方程是一个本征值问题。能量 E 的值就是本征值。本征函数就是定态薛定谔方程 ψ 。所需的本征函数和本征值的算符是 $-(\hbar^2/2m)d^2/dx^2 + V(x)$ 。这个算符称为系统的**哈密顿算符** (Hamiltonian operator)。

威廉·罗文·哈密顿爵士 (Sir William Rowan Hamilton) (1805-1865) 设计了牛顿运动方程的另一种形式, 其中涉及一个函数 H , 即系统的**哈密顿函数** (Hamiltonian function)。对于势能仅是坐标函数的系统, 总能量随时间保持不变; 也就是说, E 是守恒的。我们将只讨论这种保守系统。对于保守系统, 经典力学哈密顿函数可以简单地用坐标和共轭动量来表示总能量。对于笛卡尔坐标 x 、 y 、 z , **共轭动量**是线动量在 x 、 y 、 z 方向上的分量 p_x 、 p_y 、 p_z :

$$p_x \equiv mv_x, \quad p_y \equiv mv_y, \quad p_z \equiv mv_z \quad (3.19)$$

其中 v_x 、 v_y 、 v_z 分别是粒子的速度矢量在 x 、 y 、 z 方向上的分量。

让我们找出质量为 m 的粒子在一维中运动并受到势能 $V(x)$ 作用的经典力学哈密顿函数。哈密顿函数等于系统的能量, 而能量包括动能和势能。然而, 我们熟悉的动能形式 $\frac{1}{2}mv_x^2$ 并不适用, 因为我们必须将哈密顿表示为坐标和动量的函数, 而不是速度的函数。由于 $v_x = p_x/m$, 我们想要的动能表示形式为 $p_x^2/2m$ 。那么哈密顿函数为

$$H = \frac{p_x^2}{2m} + V(x) \quad (3.20)$$

定态薛定谔方程 (3.1) 表明: 与哈密顿函数 (3.20) 相对应, 我们有一个量子力学算符

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$$

其本征值为系统能量的允许值。经典力学中的物理量与量子力学中的算符之间的这种对应关系是普遍的。量子力学的一个基本假设是, 每一个物理量 (例如能量、 x 坐标、动量) 都有一个对应的量子力学算符。我们进一步假设, 将 B 的经典力学表达式写成笛卡尔坐标和相应动量的函数, 然后进行如下替换, 就能找到与属性 B 相对应的算符。每个笛卡尔坐标 q 都用与该坐标相乘的算符来代替:

$$\hat{q} = q \times$$

线动量 p_q 的每个笛卡尔分量都由以下算符代替：

$$\hat{p}_q = \frac{\hbar}{i} \frac{\partial}{\partial q} = -i\hbar \frac{\partial}{\partial q}$$

其中 i 是虚数单位， $\partial/\partial q$ 是 p 对坐标 q 的偏导数算符。

考虑几个例子。与 x 坐标相对应的算符是 x 的乘法运算：

$$\hat{x} = x \times \quad (3.21)$$

同样地，

$$\hat{y} = y \times, \quad \hat{z} = z \times \quad (3.22)$$

线动量各分量的算符为

$$\hat{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}, \quad \hat{p}_y = \frac{\hbar}{i} \frac{\partial}{\partial y}, \quad \hat{p}_z = \frac{\hbar}{i} \frac{\partial}{\partial z} \quad (3.23)$$

与 p_x^2 相对应的算符为

$$\hat{p}_x^2 = \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right)^2 = \frac{\hbar}{i} \frac{\partial}{\partial x} \frac{\hbar}{i} \frac{\partial}{\partial x} = -\hbar^2 \frac{\partial^2}{\partial x^2} \quad (3.24)$$

p_y 与 p_z 的表示类似。

一维的势能算符和动能算符是什么？设一个系统的势能函数为 $V(x) = ax^2$ ，其中 a 是常数。将 x 用 $x \times$ 替代，我们可以看到，势能算符只是乘以 ax^2 。即， $\hat{V}(x) = ax^2 \times$ 。一般来说，对于任何势能函数

$$\hat{V}(x) = V(x) \times \quad (3.25)$$

经典力学中动能 T 的表达式 (3.20) 为

$$T = p_x^2/2m \quad (3.26)$$

将 p_x 用对应的算符 (3.23) 代替，我们有

$$\hat{T} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \quad (3.27)$$

其中用到了式 (3.24)，偏导数就变成了单变量的常导数。经典力学的哈密顿量 (3.20) 为

$$H = T + V = p_x^2/2m + V(x) \quad (3.28)$$

相应的量子力学哈密顿（或能量）算符为

$$\hat{H} = \hat{T} + \hat{V} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \quad (3.29)$$

与薛定谔方程 (3.1) 中的算符一致。请注意，所有这些算符都是线性的。

量子力学算符如何与系统对应的属性相关联？每个这样的算符都有自己的一组本征函数和本征值。令 \hat{B} 是与物理量 B 相对应的量子力学算符，令 f_i 和 b_i 分别代表算符 \hat{B} 的本征函数和本征值。带入式 (3.14)，我们有

$$\hat{B}f_i = b_i f_i, \quad i = 1, 2, 3 \dots \quad (3.30)$$

算符 \hat{B} 有许多本征函数和本征值，用下标 i 相区分。 \hat{B} 通常是线性算符，且式 (3.30) 是微分方程，对应的解可以给出本征函数和本征值。量子力学假设（无论系统的状态函数是什么）对属性 B 的测量必须得到算符 \hat{B} 的一个本征值 b_i 。例如，一个系统的能量只能通过能量（哈密顿）算符 \hat{H} 的本征值求出。使用 ψ_i 来表示 \hat{H} 的本征函数，我们有本征方程 (3.30)

$$\hat{H}\psi_i = E_i \psi_i \quad (3.31)$$

将式 (3.29) 带入式 (3.31)，我们可以得到：对于单粒子的一维系统，有

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi_i = E_i \psi_i \quad (3.32)$$

就是式 (3.1) 的定态薛定谔方程。因此，我们关于算符的假设与我们之前的工作是一致的。稍后，我们将进一步证明动量算符 (3.23) 的选择是正确的，因为在向经典力学的极限过渡中，这一选择会导出 $p_x = m(dx/dt)$ ，这是理所应当的。（见问题 7.59）

在第一章中，我们假设量子力学系统的状态是由一个状态函数 $\Psi(x, t)$ 指定的，这个状态函数包含了我们可能知道的系统所有的信息。 Ψ 是如何告诉我们关于性质 B 的信息的？我们假设：若 Ψ 是 \hat{B} 的本征函数，本征值为 b_k ，则对性质 B 的测量可得到值为 b_k 的结果。例如，考虑系统的能量。能量算符的本征函数为定态薛定谔方程 (3.32) 的解 $\psi(x)$ 。假设系统处于定态，有态函数 (1.20)

$$\Psi(x, t) = e^{-iEt/\hbar} \psi(x) \quad (3.33)$$

$\Psi(x, t)$ 是能量算符 \hat{H} 的本征函数吗？我们有

$$\hat{H}\Psi(x, t) = \hat{H}e^{-iEt/\hbar} \psi(x)$$

\hat{H} 不包含对时间的导数，因此不会影响指数因子 $e^{-iEt/\hbar}$ ，我们有

$$\begin{aligned} \hat{H}\Psi(x, t) &= e^{-iEt/\hbar} \hat{H}\psi(x) = Ee^{-iEt/\hbar} \psi(x) = E\Psi(x, t) \\ \hat{H}\Psi &= E\Psi \end{aligned} \quad (3.34)$$

其中，我们用到了式 (3.31)。因此，对于定态系统， $\Psi(x, t)$ 是 \hat{H} 的本征函数，我们在测量能量时肯定会得到值为 E 的结果。

考虑另一个动量的例子。算符 \hat{p}_x 的本征函数 g 由以下方程求出：

$$\begin{aligned} \hat{p}_x g &= \hbar k g \\ \frac{\hbar}{i} \frac{dg}{dx} &= \hbar k g \end{aligned} \quad (3.35)$$

我们有（问题 3.29）

$$g = Ae^{ikx/\hbar} \quad (3.36)$$

其中 A 是任意常数。当 $|x|$ 充分大时，为了确保 g 不趋于无穷，本征值 k 必须是实数。因此， \hat{p}_x 算符的本征值一定为实数

$$-\infty < k < \infty \quad (3.37)$$

这是很合理的。任何对 p_x 的测量都会有符合式 (3.37) 的 \hat{p}_x 的本征值。(3.36) 中每一个不同的 k 值都给出了不同的本征函数 g 。物理量动量的算符涉及虚数 i ，这似乎令人惊讶。事实上， i 的出现确保了本征值 k 一定为实数。回忆算符 d/dx 的本征值为虚数（第 3.2 节）。比较自由粒子波函数 (2.30) 和 \hat{p}_x 的本征函数 (3.36)，我们注意到以下物理解释：(2.30) 中的第一项相当于正动量，代表 $+x$ 方向的运动；(2.30) 中的第二项相当于负动量，代表 $-x$ 方向的运动。

现在来考虑一个盒子中粒子的动量。在一维盒子中处于静止状态的粒子的状态函数为 [式 (3.33), (2.20), 和 (2.23)]。

$$\Psi(x, t) = e^{-iEt/\hbar} \left(\frac{2}{l} \right)^{1/2} \sin \left(\frac{n\pi x}{l} \right) \quad (3.38)$$

其中 $E = n^2 h^2 / 8ml^2$ 。粒子的 p_x 值是确定的吗？即 $\Psi(x, t)$ 是算符 \hat{p}_x 的本征函数吗？注意 \hat{p}_x 的本征函数，我们可以看到：没有任何实常数 k 的数值可以使 (3.36) 中的指数函数变成正弦函数，如 (3.38) 中的正弦函数。因此 Ψ 不是 \hat{p}_x 的本征函数。我们可以直接验证这一点：我们有

$$\hat{p}_x \Psi = \frac{\hbar}{i} \frac{\partial}{\partial x} e^{-iEt/\hbar} \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n\pi x}{l}\right) = \frac{n\pi\hbar}{il} e^{-iEt/\hbar} \left(\frac{2}{l}\right)^{1/2} \cos\left(\frac{n\pi x}{l}\right)$$

由于 $\hat{p}_x \Psi \neq \text{常数} \cdot \Psi$ ，态函数 Ψ 不是算符 \hat{p}_x 的本征函数。

注意：系统的态函数 Ψ 不一定是与系统物理量 B 相对应的算符 \hat{B} 的本征函数 f_i 。因此，箱中粒子定态波函数不是 \hat{p}_x 的本征函数。尽管如此，当我们对箱中粒子定态的动量值进行观测时，我们也必须得到 \hat{p}_x 的一个本征值 (3.37)。

箱中粒子定态波函数是 \hat{p}_x 的本征函数吗？由 (3.24)，我们有

$$\begin{aligned} \hat{p}_x^2 \Psi &= -\hbar^2 \frac{\partial^2}{\partial x^2} e^{-iEt/\hbar} \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n\pi x}{l}\right) = \frac{n^2 \pi^2 \hbar^2}{l^2} e^{-iEt/\hbar} \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n\pi x}{l}\right) \\ \hat{p}_x^2 \Psi &= \frac{n^2 h^2}{4l^2} \Psi \end{aligned} \quad (3.39)$$

因此，对处于量子数为 n 的定态粒子，对 \hat{p}_x^2 的观测始终得到值为 $n^2 h^2 / 4l^2$ 的结果。这一点不足为奇：盒子中的势能为零，哈密顿算符为

$$\hat{H} = \hat{T} + \hat{V} = \hat{T} = \hat{p}_x^2 / 2m$$

则有式 (3.34)：

$$\begin{aligned} \hat{H} \Psi &= E \Psi = \frac{\hat{p}_x^2}{2m} \Psi \\ \hat{p}_x^2 \Psi &= 2mE \Psi = 2m \frac{n^2 h^2}{8ml^2} \Psi = \frac{n^2 h^2}{4l^2} \Psi \end{aligned} \quad (3.40)$$

与 (3.39) 一致。 p_x 的唯一可能取值是

$$p_x^2 = n^2 h^2 / 4l^2 \quad (3.41)$$

等式 (3.41) 表明，测量 p_x 必然会得到以下两个值中的一个，即 $p_x = \pm \frac{1}{2} nh/l$ ，对应于粒子在盒中向右或向左移动。这种似是而非的说法并不准确。使用第 7 章的方法进行的分析表明：测量值很有可能接近以下两个值之一： $p_x = \pm \frac{1}{2} nh/l$ 。但任何符合 (3.37) 的值都可以通过测量盒子中粒子的 p_x 得出；见问题 7.41。

我们假设对性质 B 的观测一定会给出 \hat{B} 的其中一个本征值。如果态函数 Ψ 恰好是 B 的本征函数，本征值为 b ，当我们观测 B 时就一定会得到 b 。然而，假设 Ψ 不是 \hat{B} 的本征函数，会发生什么呢？我们仍然断言：当我们测量 B 时，会得到 \hat{B} 的一个本征值，但我们无法预测会得到哪个本征值。我们将在第 7 章中看到，得到 \hat{B} 的各种本征值的概率是可以预测的。

例题：测量长度为 l 的一维盒子中质量为 m 的粒子的能量。

如果开始测量时，粒子的状态函数为

$$(a) \Psi = (30/l^5)^{1/2} x(l-x), \quad 0 \leq x \leq l;$$

$$(b) \Psi = (2/l)^{1/2} \sin(3\pi x/l), \quad 0 \leq x \leq l;$$

测量结果可能是哪些值？

- (a) 对性质 E 进行测量的可能结果是系统能量（哈密顿）算符 \hat{H} 的本征值。因此，测量值一定是 $n^2\hbar^2/8ml^2$, $n = 1, 2, 3 \dots$ 中的一个值。由于 Ψ 不是 \hat{H} 的本征函数 $[(2/l)^{1/2} \sin(n\pi x/l)]$ 式 (2.23)，我们无法预测这个非稳态会得到其中哪一个本征值。（获得这些本征值的概率可在第 7.6 节的最后一个例子中找到）
- (b) 由于 Ψ 是 \hat{H} 的本征函数，本征值为 $3^2\hbar^2/8ml^2$ ，测量值为 $9^2\hbar^2/8ml^2$ 。

3.4 三维多粒子的薛定谔方程

到目前为止，我们的研究仅限于一维单粒子系统。上一节中发展的算符形式主义允许我们将工作扩展到三维多粒子系统。含时薛定谔方程的状态函数的时间发展假定为公式 (1.13) 的形式：

$$\boxed{i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi} \quad (3.42)$$

能量本征函数和本征值的定态薛定谔方程为

$$\boxed{\hat{H}\psi = E\psi} \quad (3.43)$$

将势能视为与时间无关，并应用从 (1.13) 得出 (1.19) 所用的变量分离方法，即可从 (3.42) 导出该式。

对于单粒子三维系统，经典力学下的哈密顿量为

$$H = T + V = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \quad (3.44)$$

引入量子力学算符，我们有哈密顿算符

$$\hat{H} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(x, y, z) \quad (3.45)$$

(3.45) 中括号内的算符称为拉普拉斯算符（Laplacian operator, ∇^2 ）：

$$\boxed{\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}} \quad (3.46)$$

则单粒子三维系统的定态薛定谔方程为

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi = E\psi \quad (3.47)$$

现在考虑一个有 n 个粒子的三维系统。设粒子 i 具有质量 m_i ，坐标 (x_i, y_i, z_i) ，其中 $i = 1, 2, 3, \dots, n$ 。系统的总动能为动能是各个粒子动能的总和：

$$T = \frac{1}{2m_1} (p_{x_1}^2 + p_{y_1}^2 + p_{z_1}^2) + \frac{1}{2m_2} (p_{x_2}^2 + p_{y_2}^2 + p_{z_2}^2) + \dots + \frac{1}{2m_n} (p_{x_n}^2 + p_{y_n}^2 + p_{z_n}^2)$$

其中 p_{x_i} 是第 i 个粒子动量在 x 方向上的分量。动能算符则为

$$\hat{T} = -\frac{\hbar^2}{2m_1} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \right) - \dots - \frac{\hbar^2}{2m_n} \left(\frac{\partial^2}{\partial x_n^2} + \frac{\partial^2}{\partial y_n^2} + \frac{\partial^2}{\partial z_n^2} \right)$$

$$\boxed{\hat{T} = -\sum_{i=1}^n \frac{\hbar^2}{2m_i} \nabla_i^2} \quad (3.48)$$

$$\nabla_i^2 \equiv \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \quad (3.49)$$

我们通常只讨论势能只取决于 n 个粒子的 $3n$ 个坐标的情况：

$$V = V(x_1, y_1, z_1, \dots, x_n, y_n, z_n)$$

则有 n 粒子的三维系统的哈密顿算符为

$$\hat{H} = - \sum_{i=1}^n \frac{\hbar^2}{2m} \nabla_i^2 + V(x_1, \dots, z_n) \quad (3.50)$$

，对应的定态薛定谔方程为

$$\left[- \sum_{i=1}^n \frac{\hbar^2}{2m} \nabla_i^2 + V(x_1, \dots, z_n) \right] \psi = E\psi \quad (3.51)$$

其中定态波函数是 n 个粒子的 $3n$ 个坐标的函数：

$$\psi = \psi(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \quad (3.52)$$

薛定谔方程 (3.51) 是一个线性偏微分方程。

例如，考虑一个由两个相互作用的粒子组成的系统，其势能与它们之间的距离成反比， c 为比例常数。薛定谔方程 (3.51) 变为

$$\left[- \frac{\hbar^2}{2m_1} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \right) - \frac{\hbar^2}{2m_2} \left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2} \right) + \frac{c}{[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]^{1/2}} \right] \psi = E\psi \quad (3.53)$$

虽然 (3.53) 看起来很困难，但我们将在第 6 章中解决它。



图 3.1: 位于 x', y', z' 的一个无穷小箱形区域

对于单粒子一维系统，玻恩假设 [式 (1.15)] 表明： $|\Psi(x', t)|^2 dx$ 是在 t 时刻在 x' 到 $x' + dx$ 区域内找到粒子的概率，其中 x' 是 x 的一个特定值。我们对这一假设作如下扩展：对于三维单粒子系统，表达式

$$|\Psi(x', y', z', t)|^2 dx dy dz \quad (3.54)$$

是在 x' 到 $x' + dx$ 、 y' 到 $y' + dy$ 、 z' 到 $z' + dz$ 这一无限小空间内找到粒子的概率。由于找到粒子的总概率为 1，因此归一化条件为

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Psi(x', y', z', t)|^2 dx dy dz = 1 \quad (3.55)$$

对于 n 粒子的三维系统，我们假设

$$\boxed{|\Psi(x'_1, y'_1, z'_1, x'_2, y'_2, z'_2, \dots, x'_n, y'_n, z'_n, t)|^2 dx_1 dy_1 dz_1 dx_2 dy_2 dz_2 \dots dx_n dy_n dz_n} \quad (3.56)$$

是在 (x'_1, y'_1, z'_1) 处带边 dx_1, dy_1, dz_1 的无限小矩形盒状区域内发现粒子 1，在 (x'_2, y'_2, z'_2) 处带边 dx_2, dy_2, dz_2 的无限小矩形盒状区域内发现粒子 2， \dots ，在 (x'_n, y'_n, z'_n) 处带边 dx_n, dy_n, dz_n 的无限小矩形盒状区域内发现粒子 n 的概率。由于找到粒子的总概率为 1，因此归一化条件为

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Psi|^2 dx_1 dy_1 dz_1 \dots dx_n dy_n dz_n = 1 \quad (3.57)$$

在量子力学中，对一个系统的全空间进行积分的习惯做法是 $\int d\tau$ 。简写式 (3.55) 和 (3.57)，有

$$\boxed{\int |\Psi|^2 d\tau = 1} \quad (3.58)$$

虽然 (3.58) 看起来像是一个不定积分，但我们可以理解为一个定积分。根据上下文可以理解积分变量及其范围。

对于定态系统， $|\Psi|^2 = |\psi|^2$ ，及

$$\boxed{\int |\psi|^2 d\tau = 1} \quad (3.59)$$

3.5 三维盒子中的粒子

目前，我们只讨论单粒子问题。在本节中，我们将考虑第 2.2 节中解决的问题的三维情况，即粒子在盒子中的情况。



图 3.2: 在这个箱形区域内， $V = 0$ 。

三维盒子有多种可能的形状。我们选择的坐标系使盒子的一个角位于原点，盒子位于空间的第

一卦限（图 3.2）。在盒子内，势能为零；而在盒子外，势能为无穷大：

$$V(x, y, z) = 0, \quad \text{其中} \begin{cases} 0 < x < a \\ 0 < y < b \\ 0 < z < c \end{cases} \quad (3.60)$$

$$V = \infty, \quad \text{elsewhere}$$

由于粒子具有无限能量的概率为零，因此在盒子外波函数一定为零。在盒内，势能算符为零，薛定谔方程（3.47）为

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) = E\psi \quad (3.61)$$

为了求解（3.61），我们假设解可以写成三个分别只与 x 、 y 和 z 有关的单独函数的乘积：

$$\psi(x, y, z) = f(x)g(y)h(z) \quad (3.62)$$

我们可能会认为，这一假设抛弃了不满足（3.62）形式的解。然而，可以证明的是：如果我们能找到满足边界条件的（3.62）形式的解，那么就不存在满足边界条件的薛定谔方程的其他解了（证明见 G. F. D. Duff 和 D. Naylor, *Differential Equations of Applied Mathematics*, Wiley, 1966, 第 257-258 页）。我们用来求解（3.62）的方法称为**变量分离**（separation of variables）。

由（3.62），我们有

$$\frac{\partial^2 \psi}{\partial x^2} = f''(x)g(y)h(z), \quad \frac{\partial^2 \psi}{\partial y^2} = f(x)g''(y)h(z) \quad (3.63)$$

将（3.63）和（3.62）代入式（3.61），有

$$-(\hbar^2/2m)f''gh - (\hbar^2/2m)fg''h - (\hbar^2/2m)fgh'' - E fgh = 0 \quad (3.64)$$

将上式左右两边同除以 fgh ，有

$$-\frac{\hbar^2 f''}{2mf} - \frac{\hbar^2 g''}{2mg} - \frac{\hbar^2 h''}{2mh} - E = 0 \quad (3.65)$$

$$-\frac{\hbar^2 f''(x)}{2mf(x)} = \frac{\hbar^2 g''(y)}{2mg(y)} + \frac{\hbar^2 h''(z)}{2mh(z)} + E \quad (3.66)$$

令 E_x 与上式的左边相等，有

$$E_x \equiv -\hbar^2 f''(x)/2mf(x) \quad (3.67)$$

定义（3.67）表明 E_x 与 y 和 z 无关，方程（3.66）表明 E_x 等于 $\hbar^2 g''(y)/2mg(y) + \hbar^2 h''(z)/2mh(z) + E$ ；因此， E_x 也与 x 无关。综上所述， E_x 一定是与 x 、 y 和 z 都无关的常数。

与（3.67）类似，我们定义 E_y 与 E_z ：

$$E_y \equiv -\hbar^2 g''(y)/2mg(y), \quad E_z \equiv -\hbar^2 h''(z)/2mh(z) \quad (3.68)$$

由于 x 、 y 、 z 在（3.65）中对称出现，因此与证明 E_x 是常数的推理相同， E_y 和 E_z 也是常数。将定义（3.67）和（3.68）代入（3.65）可得

$$E_x + E_y + E_z = E \quad (3.69)$$

方程（3.67）和（3.68）为

$$\frac{d^2 f(x)}{dx^2} + \frac{2m}{\hbar^2} E_x f(x) = 0 \quad (3.70)$$

$$\frac{d^2 g(y)}{dy^2} + \frac{2m}{\hbar^2} E_y g(y) = 0, \quad \frac{d^2 h(z)}{dz^2} + \frac{2m}{\hbar^2} E_z h(z) = 0 \quad (3.71)$$

我们已将三变量的偏微分方程转换为三个常微分方程。(3.70) 的边界条件是什么？由于波函数在盒外消失，因此连续性要求波函数 ψ 在盒壁处消失。特别是，在 yOz 平面上的盒壁上必须为零，此处 $x = 0$ ；此外，在箱体的平行壁上必须为零，其中 $x = a$ 。因此， $f(0) = f(a) = 0$ 。

现在将公式 (3.70) 与一维盒子中粒子的薛定谔方程 [公式 (2.10)] 进行比较。方程的形式相同，(3.70) 中的 E_x 与 (2.10) 中的 E 相对应。边界条件是否相同？是的，只是自变量消失的第二点是 $x = a$ ，而不是 $x = l$ 。因此，我们可以利用第 2.2 节中的工作来写出解 [见式 (2.23) 和 (2.20)]。

$$f(x) = \left(\frac{2}{a}\right)^{1/2} \sin\left(\frac{n_x \pi x}{a}\right)$$

$$E_x = \frac{n_x^2 \hbar^2}{8ma^2}, \quad n_x = 1, 2, 3 \dots$$

将同样的推理应用于 y 和 z 方程，可以得出

$$g(y) = \left(\frac{2}{b}\right)^{1/2} \sin\left(\frac{n_y \pi y}{b}\right), \quad h(z) = \left(\frac{2}{c}\right)^{1/2} \sin\left(\frac{n_z \pi z}{c}\right)$$

$$E_y = \frac{n_y^2 \hbar^2}{8mb^2}, \quad n_y = 1, 2, 3 \dots, \quad E_z = \frac{n_z^2 \hbar^2}{8mc^2}, \quad n_z = 1, 2, 3 \dots$$

由 (3.69)，能量为

$$E = \frac{\hbar^2}{8m} \left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right) \quad (3.72)$$

与一维盒子中的粒子一样，基态能量大于经典力学的最低能量值 0。

由式 (3.62)，箱中粒子的波函数为

$$\psi(x, y, z) = \left(\frac{8}{abc}\right)^{1/2} \sin\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{b}\right) \sin\left(\frac{n_z \pi z}{c}\right) \quad (3.73)$$

波函数有三个量子数 n_x 、 n_y 和 n_z ，我们将其归因于问题的三维性质。三个量子数彼此独立变化。

在一维单粒子问题中，节点满足 $\psi(x) = 0$ ，解这个方程，我们就得到了 $\psi = 0$ 的 x 值。而在三维单粒子问题中，节点满足 $\psi(x, y, z) = 0$ ，从方程中解出 z ，我们有 $z = f(x, y)$ ，每个这样的解都是三维空间中一个节面 (nodal surface) 的方程。例如，对于 $n_x = 1$ 、 $n_y = 1$ 、 $n_z = 2$ 的定态，令波函数 (3.73) 为 0，解得节面为 $z = c/2$ ，这是一个平面的方程，该平面平行于盒子的顶面和底面，并且位于这两个面的中间。同样地，对于状态 $n_x = 2$ 、 $n_y = 1$ 、 $n_z = 1$ ，节面为 $x = a/2$ 的平面。

由于波函数中的 x 、 y 和 z 因子可以独立归一化，将波函数归一化，有

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi|^2 dx dy dz = \int_0^a |f(x)|^2 dx \int_0^b |g(y)|^2 dy \int_0^c |h(z)|^2 dz = 1$$

其中我们用到了

$$\boxed{\int \int \int F(x) G(y) H(z) dx dy dz = \int F(x) dx \int G(y) dy \int H(z) dz} \quad (3.74)$$

式 (3.73) 中 $\psi(x, y, z)$ 的量纲是什么？对于单粒子三维系统， $|\psi|^2 dx dy dz$ 是概率，概率是没有量纲的。由于 $dx dy dz$ 的量纲为长度³，则为了使 $|\psi|^2 dx dy dz$ 无量纲， $\psi(x, y, z)$ 一定具有长度^{-3/2} 的量纲。

假设 $a = b = c$ ，则我们得到了一个立方体。能级为

$$E = (h^2/8ma^2) (n_x^2 + n_y^2 + n_z^2) \quad (3.75)$$

让我们用表格列出被限制在立方体势箱中粒子的一些允许能量：

$n_x n_y n_z$	111	211	121	112	122	212	221	113	131	311	222
$E (8ma^2/h^2)$	3	6	6	6	9	9	9	11	11	11	12



图 3.3: 立方体盒子中粒子最低几个状态的能量。

请注意：量子数不同的状态可能具有相同的能量（图3.3）。例如，状态 ψ_{211} 、 ψ_{121} 和 ψ_{112} （下标表示量子数）都具有相同的能量。然而，公式 (3.73) 表明这三组量子数给出了三个不同的独立波函数，因此确实代表了系统的不同状态。当两个或两个以上的独立波函数对应于具有相同能量本征值的状态时，该本征值是简并（degenerate）的。一个能级的简并度（degree of degeneracy or simply degeneracy）就是具有该能量的状态的数量。因此，立方体中粒子的次低能级是三重简并的。当我们令盒子的边长相等时，就得到了简并的现象。简并度通常与系统的对称性有关。注意波函数态 ψ_{211} 、 ψ_{121} 和 ψ_{112} 可以通过旋转立方体盒相互转化。通常，一维问题中的束缚态能级是非简并的。

在对理想气体的分子分配函数进行统计力学评估时，将每个气体分子的平移能级视为三维矩形盒子（盒子是容纳气体的容器）中的粒子能级；见 Levine，《物理化学》，第 21.6 和 21.7 节。

在金属的自由电子理论中，非过渡金属的价电子被视为一个盒子中的非相互作用粒子，盒子的两侧是金属的表面。这种近似方法虽然粗糙，但对金属的某些性质给出了相当不错的结果。

3.6 简并

对于一个 n 重简并的能级，有 n 个独立的波函数 $\psi_1, \psi_2, \psi_3, \dots, \psi_n$ 每个波函数均有本征值 w ：

$$\hat{H}\psi_1 = w\psi_1, \quad \hat{H}\psi_2 = w\psi_2, \dots, \quad \hat{H}\psi_n = w\psi_n \quad (3.76)$$

我们想要证明以下重要理论：具有能量本征值 w 的简并波函数的每一个线性组合

$$\phi \equiv c_1\psi_1 + c_2\psi_2 + c_3\psi_3 + \dots + c_n\psi_n \quad (3.77)$$

也是能量本征函数，且具有相同的能量本征值 w 。[式 (3.77) 是函数 $\psi_1, \psi_2, \psi_3, \dots, \psi_n$ 线性组合的定义。其中 c 是常数，有些可能为零。] 为了证明这个理论，我们必须指明 $\hat{H}\phi = w\phi$ ，或者

$$\hat{H}(c_1\psi_1 + c_2\psi_2 + \dots + c_n\psi_n) = w(c_1\psi_1 + c_2\psi_2 + \dots + c_n\psi_n) \quad (3.78)$$

由于 \hat{H} 是线性算符, 我们可以对式 (3.78) 的左边应用式 (3.9) $n-1$ 次, 有

$$\hat{H}(c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n) = \hat{H}(c_1\psi_1) + \hat{H}(c_2\psi_2) + \cdots + \hat{H}(c_n\psi_n)$$

再使用式 (3.10), 则 (3.79) 变为

$$\begin{aligned}\hat{H}(c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n) &= c_1\hat{H}\psi_1 + c_2\hat{H}\psi_2 + \cdots + c_n\hat{H}\psi_n \\ &= c_1w\psi_1 + c_2w\psi_2 + \cdots + c_nw\psi_n \\ \hat{H}(c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n) &= w(c_1\psi_1 + c_2\psi_2 + \cdots + c_n\psi_n)\end{aligned}$$

这就证明了我们的理论。

例如, 在立方势箱中粒子的定态波函数 ψ_{211} 、 ψ_{121} 和 ψ_{112} 是简并的, 则它们的线性组合 $c_1\psi_{211} + c_2\psi_{121} + c_3\psi_{112}$ 立方势箱中粒子哈密顿算符的本征函数, 且具有相同的能量本征值 $6h^2/8ma^2$, 与 ψ_{211} 、 ψ_{121} 和 ψ_{112} 相同。

注意: 若 ψ_1 和 ψ_2 对应于不同的能量本征值 ($\hat{H}\psi_1 = E_1\psi_1$, $\hat{H}\psi_2 = E_2\psi_2$, 且 $E_1 \neq E_2$), 则它们的线性组合 $c_1\psi_1 + c_2\psi_2$ 不是能量算符 \hat{H} 的本征函数。

由于与简并能级相对应的波函数的任何线性组合都是具有相同本征值的 \hat{H} 的本征函数, 因此我们可以为任何简并能级构造出无数个不同的波函数。事实上, 我们只对线性无关的本征函数感兴趣。若 f_1, f_2, \dots, f_n 当且仅当所有常数 c_1, c_2, \dots, c_n 都为零时, 满足关系式 $c_1f_1 + c_2f_2 + \cdots + c_nf_n = 0$, 则称 f_1, f_2, \dots, f_n 是线性无关的。这就意味着在一组线性独立函数中, 没有一个函数可以用其余函数的线性组合来表示。例如, 函数 $f_1 = 3x$, $f_2 = 5x^2 - x$, $f_3 = x^2$ 不是线性无关的, 因为 $f_2 = 5f_3 - \frac{1}{3}f_1$ 。而函数 $g_1 = 1$, $g_2 = x$, $g_3 = x^2$ 是线性无关的, 因为没有函数可以用其余函数的线性组合来表示。

一个能级的简并度等于与该能值相对应的线性独立波函数的数目。一维自由粒子波函数 (2.30) 是两个线性独立函数的线性组合, 这两个函数分别是具有相同能量本征值 E 的本征函数。因此每个这样的能级 (除了 $E = 0$) 都是二重简并的 (简并度为 2)。

3.7 平均值

第 3.3 节指出: 若态函数 Ψ 不是算符 \hat{B} 的本征函数, 则对 B 的观测将得到一组可能值 (\hat{B} 的本征值的其中一个)。我们现在考虑具有状态 Ψ 的系统的性质 B 的平均值。为了在实验中找到 B 的平均值, 我们需要许多相同的、不相互影响的系统 Ψ , 每个系统都处于相同的状态, 然后测量每个系统中的 B 。 B 的平均值 (average value), 记作 $\langle B \rangle$, 定义为所有测量值 b_1, b_2, \dots, b_N 的算术平均:

$$\langle B \rangle = \frac{1}{N} \sum_{j=1}^N B_j \quad (3.79)$$

其中系统的数目 N 是很大的一个数。

我们可以对 B 的所有可能值求和, 而不是对 B 的观测值求和, 将每个可能值乘以观测到的次数, 得到等价表达式

$$\langle B \rangle = \frac{1}{N} \sum_b n_b b \quad (3.80)$$

其中 n_b 是观测到的值 b 的次数。举一个例子就会很清楚。假设一个班级的五名学生进行了五道题的测试, 取得的分数为: 0, 20, 20, 60, 60, 80, 80, 80, 100。则用式 (3.79) 计算平均分, 我们有

$$\frac{1}{N} \sum_{j=2}^N B_j = \frac{0 + 20 + 20 + 60 + 60 + 80 + 80 + 80 + 100}{9} = \frac{500}{9} = 56$$

用式 (3.80) 计算平均分, 对所有可能的分数: 0, 20, 40, 60, 80, 100 求和, 我们有

$$\frac{1}{N} \sum_b n_b b = \frac{1(0) + 2(20) + 0(40) + 2(60) + 3(80) + 1(100)}{9} = \frac{500}{9} = 56$$

方程 (3.80) 可以写作

$$\langle B \rangle = \sum_b \left(\frac{n_b}{N} \right) b$$

由于 N 很大, n_b/N 是观测到值 b 的概率 P_b , 因此

$$\langle B \rangle = \sum_b P_b b \quad (3.81)$$

现在考虑态函数为 $\Psi(x, t)$ 的一维单粒子系统中 x 坐标的平均值。 x 坐标的取值范围是连续的, 在 x 到 $x + dx$ 的无限小区间内找到粒子的概率为 $|\Psi|^2 dx$ 。对无限小区的求和等同于对整个 x 范围的积分, 因此 (3.81) 变为

$$\langle x \rangle = \int_{-\infty}^{\infty} x |\Psi(x, t)|^2 dx \quad (3.82)$$

对于单粒子三维系统, 在位于 (x, y, z) , 具有边长 dx, dy, dz 的体积微元内找到粒子的概率为

$$|\Psi(x, y, z, t)|^2 dx dy dz \quad (3.83)$$

如果我们想知道在 x 到 $x + dx$ 的范围内找到粒子的概率, 我们必须用 y 和 z 的所有可能取值对式 (3.83) 进行积分, 由于当 x 在 $x + dx$ 范围内时, 粒子的 y 和 z 坐标可以任意取值。因此, 对于三维的例子, 式 (3.82) 变为

$$\langle x \rangle = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Psi(x, y, z, t)|^2 dy dz \right] x dx \quad (3.84)$$

$$\langle x \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x |\Psi(x, y, z, t)|^2 dx dy dz$$

现在, 我们来看看某种物理特性 $B(x, y, z)$ 的平均值, 它是粒子坐标的函数。一个例子就是势能函数 $V(x, y, z)$ 。根据式 (3.84) 的相同推理, 可以得出

$$\langle B(x, y, z) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Psi(x, y, z, t)|^2 B(x, y, z) dx dy dz \quad (3.85)$$

$$\langle B(x, y, z) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi^* B \Psi dx dy dz \quad (3.86)$$

式 (3.86) 的形式似乎有点异想天开, 因为它与 (3.85) 没有什么不同。但稍后我们将看到它的重要意义。

一般来说, 对于单粒子三维系统, 属性 B 既取决于坐标, 也取决于动量:

$$B = B(x, y, z, p_x, p_y, p_z)$$

我们如何找到 B 的平均值? 我们假设处于 Ψ 状态的系统的 $\langle B \rangle$ 为

$$\begin{aligned} \langle B \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi^* B \left(x, y, z, \frac{\hbar^2}{i} \frac{\partial}{\partial x}, \frac{\hbar^2}{i} \frac{\partial}{\partial y}, \frac{\hbar^2}{i} \frac{\partial}{\partial z} \right) \Psi dx dy dz \\ \langle B \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi^* \hat{B} \Psi dx dy dz \end{aligned} \quad (3.87)$$

其中 \hat{B} 是性质 B 的量子力学算符。[稍后，我们将利用 (3.87) 来证明含时薛定谔方程在从量子力学到经典力学的转变过程中还原为牛顿第二定律，从而为这一假设提供一些理由；见问题 7.59。]，对于 n 粒子体系，我们假设

$$\langle B \rangle = \int \Psi^* \hat{B} \Psi d\tau \quad (3.88)$$

其中 $\int d\tau$ 表示对全空间的 $3n$ 坐标的积分。由于我们将 $\Psi^* \Psi$ 视为概率密度，式 (3.88) 中的态函数必须先归一化。重要的是，要将算符放在在 Ψ^* 和 Ψ 中间。表达式 $\hat{B} \Psi^* \Psi$ 、 $\Psi^* \Psi \hat{B}$ 和 $\Psi^* \hat{B} \Psi$ 的含义是不同的，除非 B 只是坐标的函数。在 $\int \Psi^* \hat{B} \Psi d\tau$ 中，第一个操作是将算符 \hat{B} 应用于态函数 Ψ ，然后将结果与 Ψ^* 相乘并对全空间进行积分，就得到了 $\langle B \rangle$ 。

对于定态，我们有 [式 1.20]

$$\Psi^* \hat{B} \Psi = e^{iEt/\hbar} \psi^* \hat{B} e^{-iEt/\hbar} \psi = e^0 \psi^* \hat{B} \psi = \psi^* \hat{B} \psi$$

由于 \hat{B} 不包含对时间的导数，也不会影响 Ψ 中的时间因子，因此，对于定态，有

$$\langle B \rangle = \int \psi^* \hat{B} \psi d\tau \quad (3.89)$$

因此，在定态系统中，若 \hat{B} 与时间无关，则 $\langle B \rangle$ 也是与时间无关的。

考虑一个特殊的例子，若 Ψ 是 \hat{B} 的本征函数。当 $\hat{B} \Psi = k \Psi$ 时，式 (3.88) 变为

$$\langle B \rangle = \int \Psi^* \hat{B} \Psi d\tau = \int \Psi^* k \Psi d\tau = k \int \Psi^* \Psi d\tau = k$$

其中 Ψ 已归一化。由于当 $\hat{B} \Psi = k \Psi$ 时， k 是我们在测量时能找到的唯一可能的 B 值，因此这个结果是合理的（第 3.3 节）。

由式 (3.88) 很容易证明以下关于平均值的性质（见问题 3.49）：

$$\langle A + B \rangle = \langle A \rangle + \langle B \rangle \quad \langle cB \rangle = c \langle B \rangle \quad (3.90)$$

其中 A 和 B 是系统的物理量， c 是常数。但是，平均值的乘积不一定等于乘积的平均值： $\langle AB \rangle \neq \langle A \rangle \langle B \rangle$ 。

通常使用**期望值**（expectation value）一词来代替平均值。期望值并不一定是我们观察到的可能值之一。

例题：对于处于基态的三维单粒子系统，求 $\langle x \rangle$ 和 $\langle p_x \rangle$ 的值。

将基态波函数 $\psi = f(x)g(y)h(z)$ [式 (3.62)] 代入式 (3.89)，有

$$\langle x \rangle = \int \psi^* \hat{x} \psi d\tau = \int_0^c \int_0^b \int_0^a f^* g^* h^* x f g h dx dy dz$$

而在势箱外有 $\psi = 0$ 。由于 $g(y)$ 和 $h(z)$ 已归一化，使用式 (3.74)，我们有

$$\langle x \rangle = \int_0^a x |f(x)|^2 dx \int_0^b |g(y)|^2 dy \int_0^c |h(z)|^2 dz = \int_0^a x |f(x)|^2 dx$$

由于系统处于基态，有 $n_x = 1$ ，则 $f(x) = (2/a)^{1/2} \sin(\pi x/a)$ ，因此

$$\langle x \rangle = \frac{2}{a} \int_0^a x \sin^2 \left(\frac{\pi x}{a} \right) dx = \frac{a}{2} \quad (3.91)$$

其中我们用到了附录中的 (A.3)。根据图 2.4，这个结果是合理的。

同样地,

$$\begin{aligned}\langle p_x \rangle &= \int \psi^* \hat{p}_x \psi d\tau = \int_0^c \int_0^b \int_0^a f^* g^* h^* \frac{\hbar}{i} \frac{\partial}{\partial x} [f(x) g(y) h(z)] dx dy dz \\ \langle p_x \rangle &= \frac{\hbar}{i} \int_0^a f^*(x) f'(x) dx \int_0^b |g(y)|^2 dy \int_0^c |h(z)|^2 dz \\ \langle p_x \rangle &= \frac{\hbar}{i} \int_0^a f(x) f'(x) dx = \frac{\hbar}{2i} f^2(x) \Big|_0^a = 0\end{aligned}\quad (3.92)$$

其中, 我们用到了定解条件 $f(0) = f(a) = 0$ 。由于粒子向 $+x$ 和 $-x$ 方向的移动是等概率的, 因此这个结果是合理的。

练习: 求处于基态的三维单粒子系统的 $\langle p_x^2 \rangle$ 的值。

(答案: $\hbar^2/4l^2$ 。)

3.8 波函数的约束条件

在求解箱中粒子的问题时, 我们要求 ψ 是连续的。我们现在讨论波函数需要满足的其他约束条件。

由于 $|\psi|^2 d\tau$ 是概率, 我们希望通过选择一个合适的归一化常数 (normalization constant) 与波函数相乘, 使得波函数可以归一化。若 ψ 未归一化而 $N\psi$ 已归一化, 则根据归一化条件, 有

$$\begin{aligned}1 &= \int |N\psi|^2 d\tau = N^2 \int |\psi|^2 d\tau \\ N &= \left(\int |\psi|^2 d\tau \right)^{-1/2}\end{aligned}\quad (3.93)$$

只有当 ψ 处处为零时, 定积分 $\int |\psi|^2 d\tau$ 才为零。然而, ψ 不可能处处为零 (这意味着没有粒子存



图 3.4:

函数 (a) 是连续函数, 其一阶导数也连续;

函数 (b) 是连续函数, 但其一阶导数不连续;

函数 (c) 不是连续函数

在)。因此, 这个积分一定不会为零。若 $\int |\psi|^2 d\tau$ 为无穷大, 则 $|N|$ 为零, ψ 就无法归一化。当且仅当 $\int |\psi|^2 d\tau$ 具有有限而非无限值时, 我们才能对 ψ 进行归一化。如果对全空间的积分 $\int |\psi|^2 d\tau$ 为有限值, 我们说 ψ 是平方可积的 (quadratically integrable)。因此, 我们一般要求 ψ 是平方可积的。一个最重要的例外不处于束缚态的粒子。因此, 粒子在势阱中 (第 2.4 节) 的非束缚态和自由粒子的波函数都不是平方可积的。

由于 $\psi^* \psi$ 是概率密度, 因此波函数一定是单值的。如果我们的理论对在某一点发现粒子的概率给出了两个不同的值, 那就太尴尬了。如果我们要求 ψ 是单值的, 那么显然 $\psi^* \psi$ 也是单值的。 ψ 可能会有多个值 [例如, $\psi(q) = -1, 1, i$], 但 $\psi^* \psi$ 只能有一个值。然而, 我们仍要求 ψ 是单值的。

除了要求 ψ 是连续, 我们也通常要求其偏导数 $\partial\psi/\partial x$, $\partial\psi/\partial y$ 等也是连续的 (见图3.4a)。不过, 回过头来看第 2.2 节, 我们会发现: 对于箱中粒子 $\partial\psi/\partial x$ 在箱壁上不连续, 以及 $\partial\psi/\partial x$ 在箱外为零, 但是由方程 (2.23) 可知, $\partial\psi/\partial x$ 在箱壁上又不为零。 ψ' 的不连续性是由在箱壁处势能突然变为无穷大造成的。对于具有有限高度的势垒来说, ψ' 在箱壁处是连续的。

与平方可积的要求一致, 有时也会说波函数必须在任何地方都是有限的, 包括无穷大。然而, 这通常是比平方可积性更强的要求。事实上, 氢原子的一些相对论波函数在原点处是无限的, 但却是平方可积的。偶尔, 我们也会遇到在原点处无限大的非相对论波函数 [L. D. D. 兰道和 E. M. 利夫希茨, 《量子力学》, 第三版 (1977 年), 第 35 节]。因此, 基本要求是平方可积性, 而不是有限性。

我们要求任何代表物理量的算符的本征函数都满足上述要求。符合这些要求的函数被称为品优函数 (well-behaved function)。

总结

习题

Chapter 4

谐振子

我们将在第 13.2 节中看到，气相分子的能量可以近似为平动、转动、振动和电子能量之和。电子能量的计算将在第 13 章至第 17 章中讨论。平动能是分子作为一个整体在容纳物质的容器空间中的运动动能。平动能级可视为三维空间中粒子的能级（第 3.5 节）。二原子（双原子）分子的转动能级可以用刚性转子的转动能级来近似表示，第 6.4 节将对此进行讨论。双原子分子的最低几个振动能级可以用谐振子的能级来近似表示（第 4.3 节），多原子分子的振动能可以用线性谐振子振动能之和来近似表示（第 15.12 节）。本章第 4.2 节将求解谐振子的薛定谔方程。作为初步介绍，第 4.1 节讨论了求解微分方程的幂级数解法，该方法用于求解谐振子的薛定谔方程。第 4.3 节讨论分子振动。第 4.4 节介绍了求一维薛定谔方程特征值和特征函数的数值方法。

4.1 微分方程的幂级数解

到目前为止，我们只考虑了势能 $V(x)$ 是常数的情况。这使得薛定谔方程成为一个具有常数系数的二阶线性齐次微分方程，我们知道如何求解这个方程。对于 V 随 x 变化的情况，一种有用的方法是尝试薛定谔方程的幂级数解法。

为了说明这种方法，考虑微分方程

$$y''(x) + c^2 y(x) = 0 \quad (4.1)$$

其中 $c^2 > 0$ 。当然，这个微分方程有常数系数，但我们可以用幂级数法求解。让我们先用特征方程 $s^2 + c^2 = 0$ 来求解这个方程。特征方程的根是 $s = \pm ic$ 。回忆第 2.2 节中的工作 [公式 (2.10) 和 (4.1) 是相同的]，当特征方程的根是纯虚数时，我们会得到三角函数形式的通解：

$$y(x) = A \cos cx + B \sin cx \quad (4.2)$$

其中 A 和 B 是积分常量。式 (4.2) 的另一种形式是

$$y = D \sin(cx + e) \quad (4.3)$$

其中 D 和 e 是任意常数。利用两角之和的正弦公式，我们可以证明 (4.3) 等价于 (4.2)。

现在让我们用幂级数法来求解 (4.1)。我们假设方程的解可以在 $x = 0$ 处展开成 Taylor 级数的形式 (见问题：4.1)，也就是说，我们假设

$$y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \quad (4.4)$$

其中 a 为常数系数, 有待确定, 以使式 (4.1) 成立。对式 (4.4) 求导, 有

$$y'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots = \sum_{n=1}^{\infty} na_nx^{n-1} \quad (4.5)$$

其中, 我们假设该数列可以逐项求导 (对于无穷级数来说, 这并不总是正确的)。对于 y'' , 我们有

$$y''(x) = 2a_2 + 3(2)a_3x + \cdots = \sum_{n=2}^{\infty} n(n-1)a_nx^{n-2} \quad (4.6)$$

将 (4.6) 和 (4.4) 代入 (4.1), 得到

$$\sum_{n=2}^{\infty} n(n-1)a_nx^{n-2} + c^2 \sum_{n=0}^{\infty} a_nx^n = 0 \quad (4.7)$$

我们想把 (4.7) 中的两个和合并起来。只要满足一定条件, 我们就可以将两个无穷级数逐项相加, 得到它们的和:

$$\sum_{j=0}^{\infty} b_jx^j + \sum_{j=0}^{\infty} c_jx^j = \sum_{j=0}^{\infty} (b_j + c_j)x^j \quad (4.8)$$

要将 (4.8) 应用于 (4.7) 中的两个和, 我们希望每个和的极限相同, x 的幂相同。因此, 我们改变了 (4.7) 中第一个和的下标, 定义 $k \equiv n-2$ 。极限 $n=2$ 到无穷大与 $k=0$ 到无穷大相对应, 将 $n = k+2$ 代入第一个和, 有

$$\sum_{n=2}^{\infty} n(n-1)a_nx^{n-2} = \sum_{k=0}^{\infty} (k+2)(k+1)a_{k+2}x^k = \sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2}x^n \quad (4.9)$$

(4.9) 中的最后一个等式是有效的, 因为求和指数是一个**虚拟变量** (dummy variable); 我们用什么字母来表示这个变量没有区别。例如, 求和 $\sum_{i=1}^3 c_ix^i$ 和 $\sum_{m=1}^3 c_mx^m$ 是相等的, 因为它们只有求和指数不同。如果我们将这个求和展开就会更好理解:

$$\sum_{i=1}^3 c_ix^i = c_1x^1 + c_2x^2 + c_3x^3, \quad \sum_{m=1}^3 c_mx^m = c_1x^1 + c_2x^2 + c_3x^3$$

在 (4.9) 的最后一个等式中, 我们只是将表示求和指数的符号从 k 改为 n 。定积分中的积分变量也是一个虚拟变量, 因为定积分的值不受变量字母的影响:

$$\int_a^b f(x) dx = \int_a^b f(t) dt \quad (4.10)$$

将式 (4.9) 代入 (4.7), 我们得到

$$\sum_{n=0}^{\infty} [(n+2)(n+1)a_{n+2} + c^2a_n]x^n = 0 \quad (4.11)$$

如果 (4.11) 对所有 x 值都成立, 那么 x 的每个幂的系数都必须为零。为了说明这一点, 请看方程

$$\sum_{j=0}^{\infty} b_jx^j = 0 \quad (4.12)$$

在式 (4.12), 令 $x=0$, 我们有 $b_0=0$ 。将式 (4.12) 左右两侧对 x 求导, 则 $b_1=0$ 。求 n 阶导数, 令 $x=0$, 我们得到 $b_n=0$ 。因此, 从式 (4.11) 中, 我们有:

$$(n+2)(n+1)a_{n+2} + c^2a_n = 0 \quad (4.13)$$

$$a_{n+2} = -\frac{c^2}{(n+1)(n+2)}a_n \quad (4.14)$$

方程 (4.14) 是一个**递推关系式** (recursion relation), 如果我们知道 a_0 的值, 就能用式 (4.14) 计算出 a_2 、 a_4 、 a_6 等的值。如果我们知道 a_1 的值, 就能用式 (4.14) 计算出 a_3 、 a_5 、 a_7 等的值。由于对 a_0 和 a_1 的选择没有限制, 它们是任意常数, 我们分别用 A 和 Bc 来表示:

$$a_0 = A, \quad a_1 = Bc \quad (4.15)$$

则利用式 (4.14), 我们可以计算系数:

$$a_0 = A, \quad a_2 = -\frac{c^2}{1 \cdot 2}A, \quad a_4 = \frac{c^4}{1 \cdot 2 \cdot 3}A, \quad a_6 = -\frac{c^6}{6!}A, \quad \dots$$

$$a_{2k} = (-1)^k \frac{c^{2k}}{(2k)!}A, \quad k = 0, 1, 2, \dots \quad (4.16)$$

$$a_1 = Bc, \quad a_3 = -\frac{c^3}{2 \cdot 3}Bc, \quad a_5 = \frac{c^5}{2 \cdot 3 \cdot 4 \cdot 5}Bc, \quad a_7 = -\frac{c^7}{7!}Bc, \quad \dots$$

$$a_{2k+1} = (-1)^k \frac{c^{2k+1}}{(2k+1)!}B, \quad k = 0, 1, 2, \dots \quad (4.17)$$

根据式 (4.4)、(4.16) 和 (4.17) 我们有

$$y = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0,2,4,\dots}^{\infty} a_n x^n + \sum_{n=1,3,5,\dots}^{\infty} a_n x^n \quad (4.18)$$

$$y = A \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k} x^{2k}}{(2k)!} + B \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1} x^{2k+1}}{(2k+1)!} \quad (4.19)$$

(4.19) 中的两个级数是 $\cos cx$ 和 $\sin cx$ 的 Taylor 级数 (问题: 4.2)。因此, 与式 (4.2) 相同, 我们有 $y = \cos cx + \sin cx$ 。

4.2 一维谐振子

在本节中, 我们将通过求解一维谐振子的薛定谔方程来增加我们的量子力学知识。该系统是分子振动的重要模型。

经典力学的处理

在研究谐振子的波动力学之前, 我们先回顾一下经典的处理方法。我们有一个质量为 m 的粒子, 它被一个与粒子离原点的位移成正比的力吸引向原点:

$$F_x = -kx \quad (4.20)$$

比例常数 k 称为**力常数** (force constant)。 F_x 是作用在粒子上的力在 x 方向上的分量。这也是这个一维问题中的总力。附着在弹簧上的质点遵守公式 (4.20), 前提是弹簧没有从平衡位置大幅拉伸。

由牛顿第二定律 $F = ma$, 我们有

$$-kx = m \frac{d^2 x}{dt^2} \quad (4.21)$$

其中 t 是时间。若令 $c^2 = k/m$, 则式 (4.21) 与式 (4.1) 相同。因此, 其解为 [将 $c = (k/m)^{1/2}$ 代入式 (4.3)], 有

$$x = A \sin(2\pi\nu t + b) \quad (4.22)$$

其中 A 是**振幅** (the amplitude of vibration), b 是积分常数, **振动频率** (vibration frequency) ν 是

$$\nu = \frac{1}{2\pi} \left(\frac{k}{m} \right)^{1/2} \quad (4.23)$$

由于正弦函数的最大最小值分别为 1 和 -1, 相应地, 满足式 (4.22) 的 x 值在 A 和 $-A$ 之间振荡。正弦函数的最小正周期为 2π , 则完成一次完整振荡所需的时间 (称为**周期** (period)) 为正弦函数参数增加 2π 的时间。在 $t + 1/\nu$ 时刻, 正弦函数的参数为 $2\pi\nu(t + 1/\nu) + b = 2\pi\nu t + 2\pi + b$, 比 t 时刻的参数多了 2π 。因此, 周期为 $1/\nu$ 。周期的倒数就是单位时间内的振动次数 (振动频率), 因此频率为 ν 。

现在我们来考虑能量。在三维系统中, 势能 V 与力的分量的关系为

$$F_x = -\frac{dV}{dx}, \quad F_y = -\frac{dV}{dy}, \quad F_z = -\frac{dV}{dz} \quad (4.24)$$

方程 (4.24) 是势能的定义。由于这是一个一维系统, 我们有 [式 (1.12)]

$$F_x = -\frac{dV}{dx} = -kx \quad (4.25)$$

对 (4.25) 两边积分, 我们得到 $V = \int kx dx = \frac{1}{2}kx^2 + C$, 其中 C 是积分常数。势能总是有一个任意的附加常数。我们令 $C = 0$, 则

$$V = \frac{1}{2}kx^2 \quad (4.26)$$

将式 (4.23) 带入, 有

$$V = 2\pi^2\nu^2mx^2 \quad (4.27)$$

势能 V 的图象是一个抛物线。动能 T 可以由式 (4.22) 对时间 t 求导得到:

$$T = \frac{1}{2}m \left(\frac{dx}{dt} \right)^2 \quad (4.28)$$

将 T 和 V 相加, 我们得到了总能量:

$$E = T + V = \frac{1}{2}kA^2 = 2\pi^2\nu^2mA^2 \quad (4.29)$$

其中我们用到了 $\sin^2\theta + \cos^2\theta = 1$ 。

根据式 (4.22), 经典力学中的谐振子在 $x = A$ 和 $x = -A$ 之间振荡运动。这两点是运动的转折点。在这两点, 粒子的速度为零, 而在 $x = 0$ 处速度达到最大值, 同时势能为零而动能最大。经典的谐振子在 $x = A$ 和 $x = -A$ 附近的运动会比在 $x = 0$ 附近花费更多的时间 (速度最慢)。问题 4.18 计算了在不同位置找到经典谐振子的概率密度。(有趣的是, 这个概率密度在转折点处变得无限大)

量子力学的处理

由式 (3.27) 和 (4.27), 谐振子的哈密顿算符为

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + 2\pi^2\nu^2mx^2 = -\frac{\hbar^2}{2m} \left(\frac{d^2}{dx^2} - \alpha^2x^2 \right) \quad (4.30)$$

其中, 为了书写方便, 我们定义了 α 为

$$\alpha \equiv 2\pi\nu m/\hbar \quad (4.31)$$

乘上 $2m/\hbar^2$ 后, 薛定谔方程 $\hat{H}\psi = E\psi$ 变为

$$\frac{d^2\psi}{dx^2} + (2mE\hbar^{-2} - \alpha^2 x^2)\psi = 0 \quad (4.32)$$

现在我们可以尝试求得 (4.32) 的幂级数解。如果我们现在考虑具有式 (4.4) 形式的 ψ 幂级数, 我们会发现它推出了一个三项递推关系, 这比式 (4.14) 这样的两项递推关系更难处理。因此, 我们修改 (4.32) 的形式, 以便在尝试幂级数解时能得到两项的递推关系。能实现这一目的的替代方法 (见问题 4.22) 是 $f(x) \equiv e^{\alpha x^2/2}\psi(x)$ 。因此,

$$\psi = e^{-\alpha x^2/2} f(x) \quad (4.33)$$

这个方程只是定义了一个新函数 $f(x)$, 取代 $\psi(x)$ 作为未知函数来求解。(我们可以在微分方程中进行任何替换) 对式 (4.33) 两边求导两次, 得到

$$\psi'' = e^{-\alpha x^2/2} (f'' - 2\alpha x f' - \alpha f + \alpha^2 x^2 f) \quad (4.34)$$

将 (4.34) 和 (4.33) 代入 (4.32), 我们得到

$$f''(x) - 2\alpha x f'(x) + (2mE\hbar^{-2} - \alpha) f(x) = 0 \quad (4.35)$$

现在我们可以尝试幂级数解。假设

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \quad (4.36)$$

假设可以对式 (4.36) 逐项求导, 我们有

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1} = \sum_{n=0}^{\infty} n c_n x^{n-1} \quad (4.37)$$

[式 (4.37) 中第二个加和的第一项为零]。同样地,

$$f''(x) = \sum_{n=2}^{\infty} n(n-1) c_n x^{n-2} = \sum_{j=0}^{\infty} (j+2)(j+1) c_{j+2} x^j = \sum_{n=0}^{\infty} (n+2)(n+1) c_{n+2} x^n$$

其中, 我们令 $j = n - 2$, 然后将求和指数从 j 改为 n [见公式 (4.9)]。代入 (4.35), 我们得到

$$\begin{aligned} \sum_{n=0}^{\infty} (n+2)(n+1) c_{n+2} x^n - 2\alpha \sum_{n=0}^{\infty} n c_n x^n + (2mE\hbar^{-2} - \alpha) \sum_{n=0}^{\infty} c_n x^n &= 0 \\ \sum_{n=0}^{\infty} [(n+2)(n+1) c_{n+2} - 2\alpha n c_n + (2mE\hbar^{-2} - \alpha) c_n] x^n &= 0 \end{aligned} \quad (4.38)$$

令 x^n 的系数为零 [与式 (4.11) 一致], 我们得到

$$c_{n+2} = \frac{\alpha + 2\alpha n - 2mE\hbar^{-2}}{(n+1)(n+2)} c_n \quad (4.39)$$

这正是我们想要的两项递推关系。式 (4.39) 与式 (4.14) 类似, 知道了 c_n 的值, 就可以计算出 c_{n+2} 的值。因此, 我们有了两个任意常数 c_0 和 c_1 。如果我们令 $c_1 = 0$, 那么我们将得到一个只包含 x 的偶数次幂乘以指数因子的幂级数作为解:

$$\psi = e^{-\alpha x^2/2} f(x) = e^{-\alpha x^2/2} \sum_{n=0,2,4,\dots}^{\infty} c_n x^n = e^{-\alpha x^2/2} \sum_{l=0}^{\infty} c_{2l} x^{2l} \quad (4.40)$$

如果我们令 $c_0 = 0$ ，那么我们将得到另一个独立的解：

$$\psi = e^{-\alpha x^2/2} f(x) = e^{-\alpha x^2/2} \sum_{n=1,3,5,\dots}^{\infty} c_n x^n = e^{-\alpha x^2/2} \sum_{l=0}^{\infty} c_{2l+1} x^{2l+1} \quad (4.41)$$

薛定谔方程的通解为上述两个独立解的线性组合：

$$\psi = A e^{-\alpha x^2/2} \sum_{l=0}^{\infty} c_{2l+1} x^{2l+1} + B e^{-\alpha x^2/2} \sum_{l=0}^{\infty} c_{2l} x^{2l} \quad (4.42)$$

其中 A 和 B 是任意常数。

现在我们必须看看波函数的边界条件是否会对解法产生任何限制。为了了解这两个无穷级数在 x 充分大时的表现，我们要考察每个级数中连续系数的比值。在第二个级数中， x^{2l+2} 与 x^{2l} 的比值为 [在 (4.39) 中令 $n = 2l$]

$$\frac{c_{2l+2}}{c_{2l}} = \frac{\alpha + 4\alpha l - 2mE\hbar^2}{(2l+1)(2l+2)}$$

假设对于较大的 x 值，级数中的后项是主要项，那么我们就来看看对于较大的 l 值，这一比值是多少：

$$\frac{c_{2l+2}}{c_{2l}} \sim \frac{4\alpha l}{4l^2} = \frac{\alpha}{l} \quad \text{当 } l \text{ 很大时} \quad (4.43)$$

在 (4.39) 中令 $n = 2l + 1$ ，我们发现，当 l 较大时，第一个序列中连续系数的比值也是 α/l 。现在考虑函数 $e^{\alpha x^2}$ 的幂级数展开。（问题 4.3）

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \dots \quad (4.44)$$

我们有

$$e^{\alpha x^2} = 1 + \alpha x^2 + \dots + \frac{\alpha^l x^{2l}}{l!} + \frac{\alpha^{l+1} x^{2l+2}}{(l+1)!} + \dots$$

在该级数中， x^{2l+2} 和 x^{2l} 的系数比为

$$\frac{\alpha^{l+1}}{(l+1)!} \div \frac{\alpha^l}{l!} = \frac{\alpha}{l+1} \sim \frac{\alpha}{l} \quad \text{当 } l \text{ 很大时}$$

因此，解 (4.42) 中每个无穷级数的连续系数比与 l 很大时的级数 $e^{\alpha x^2}$ 相同。因此，我们可以得出结论：当 x 很大时，每个级数的行为都与 $e^{\alpha x^2}$ 相同。[这不是一个严格的证明。H. A. Buchdahl, Am. J. Phys. J. Phys., 42, 47 (1974); see also M. Bowen and J. Coster, Am. J. Phys., 48, 307 (1980)]

如果每个级数的行为都与 $e^{\alpha x^2}$ 相同，那么当 x 很大时， ψ 的行为也将与 $e^{\alpha x^2}$ 相同。当 x 趋于无穷时，波函数的值将趋于无穷大，就不满足波函数平方可积的要求。如果我们可以将级数在某一项打断，使得级数的后续项为零，那么因子 $e^{\alpha x^2}$ 就能保证在 x 趋于无穷时， ψ 的值趋于零。（使用 L' Hopital 法则，显然当 $x \rightarrow \infty$ 时，有 $x^p e^{-\alpha x^2} \rightarrow 0$ ， p 为任意有限次幂）若想让级数在某一项后被打断，则当 n 为某个值时必须要有递推公式 (4.39) 中 c_n 的系数为零，记 $n = v$ 。这就使得 c_{v+2} 、 c_{v+4} ... 等全为零。而式 (4.42) 中的一个级数将具有有限项。在递推关系 (4.39) 中，有一个量的值尚未确定，但可以通过调整，使 c_v 消失。这个量就是能量 E 。在 (4.39) 中令 c_v 的系数为零，对 α 使用式 (4.31)，我们有

$$\begin{aligned} \alpha + 2\alpha v - 2mE\hbar^2 &= 0 \\ 2mE\hbar^2 &= (2v+1) 2\pi\nu m/\hbar^{-1} \\ E &= \left(v + \frac{1}{2}\right) h\nu, \quad v = 0, 1, 2, \dots \end{aligned} \quad (4.45)$$



图 4.1: 一维谐振子最低的五个能级图

谐振子定态能级 (4.45) 间距相等 (图4.1)。不要混淆量子数 v (vee) 和振动频率 ν (nu)。

将 (4.45) 代入 (4.39), 我们得到

$$c_{n+2} = \frac{2\alpha(n-v)}{(n+1)(n+2)}c_n \quad (4.46)$$

根据 (4.45) 对能量进行量子化, 我们已经使其中一个级数在有限项后被打断。为了消除 (4.42) 中的另一个无穷级数, 我们必须将与之相乘的任意常数设为零。这样, 根据 v 是偶数还是奇数, 波函数就是只包含 x 的偶次或奇次幂的有限幂级数的 $e^{-\alpha x^2/2}$ 倍。在该级数中, 由于我们选择的能量 E 使得 c_{v+2} 、 c_{v+4} 、 \dots 全消失, 则 x 的最高次幂为 v 。因此, 波函数 (4.42) 的形式为

$$\psi_v = \begin{cases} e^{-\alpha x^2/2} (c_0 + c_2 x^2 + \dots + c_v x^v), & \text{当 } v \text{ 为偶数时} \\ e^{-\alpha x^2/2} (c_1 x + c_3 x^3 + \dots + c_v x^v), & \text{当 } v \text{ 为奇数时} \end{cases} \quad (4.47)$$

其中任意常数 A 与 B 可被包含在 c_0 和 c_1 中, 因此可以略去。 c_0 和 c_1 之后的系数可以通过递推关系式 (4.46) 计算得到。由于量子数 v 出现在了递推关系中, 对于不同的 v 值, 我们可以得到一组不同的系数 c_i 。例如, ψ_4 中的 c_2 与 ψ_2 中的 c_2 是不同的。

与盒中粒子一样, 波函数必须是品优波函数的要求迫使我们能量量子化。对于不满足式 (4.45) 的能量, ψ 不是平方可积的。例如, 图4.2显示了能量分别满足 $E = 0.499h\nu$ 、 $E = 0.500h\nu$ 和 $E = 0.501h\nu$ 。当 $E = 0.500h\nu$ 时, 式 (4.40) 中的 ψ , 其中我们用到了递推关系 (4.39) 来计算系数 c_n (另见问题 4.23)。图4.2显示了这些曲线在 $\alpha^{1/2}x$ 附近区域的放大图。

谐振子的基态能量不为零。这个能量 $\frac{1}{2}h\nu$ 被称为**零点能量** (zero-point energy)。这是在绝对零度的温度下, 谐振子集合中谐振子的振动能量。零点能可以由不确定性原理来理解。如果最低态的能量为零, 那么它的势能和动能 (均为非负值) 都必须为零。动能为零则意味着动量也为零, 那么 $\Delta p_x = 0$ 。势能为零则意味着粒子始终处于原点, 则 $\Delta x = 0$ 。但我们不能同时让 $\Delta p_x = 0$ 和 $\Delta x = 0$, 因此, 基态能量必须大于零。零点能 (ZPE) 的定义为 $E_{ZPE} = E_{gs} - V_{min}$, 其中 E_{gs} 是基态能量, V_{min} 是势能的最小值。

偶函数和奇函数

在详细考察波函数之前, 我们先定义奇函数和偶函数。如果函数 $f(x)$

$$f(-x) = f(x) \quad (4.48)$$



图 4.2:

当 $E = 0.499h\nu$ 、 $E = 0.500h\nu$ 和 $E = 0.501h\nu$ 时，
 谐振子薛定谔方程的解只包含 x 的偶次幂。
 在 $x = 0$ 附近的区域，三条曲线几乎重合。
 当 $|\alpha^{1/2}x| > 3$ 时， $E = 0.500h\nu$ 的曲线几乎与 x 轴重合。

图 4.3: 图4.2在 $\alpha^{1/2}x$ 附近区域的放大图。

则称 $f(x)$ 为**偶函数** (even function)。因为 $(-x)^2 = x^2$ 和 $e^{-b(x^2)} = e^{-b(-x^2)}$ ，所以 x^2 和 e^{bx^2} 均是偶函数。偶函数的图象关于 y 轴对称 (例如，见图4.4a)。因此，对于偶函数，有

$$\int_{-a}^{+a} f(x) dx = 2 \int_0^a f(x) dx \quad (4.49)$$

如果 $g(x)$ 满足

$$g(-x) = -g(x) \quad (4.50)$$

则称 $g(x)$ 为**奇函数** (odd function)。例如 x 、 $1/x$ 和 $x^3e^{x^2}$ 。在式 (4.50) 中，令 $x = 0$ ，若 $g(0)$ 存在且为单值，则 $g(0) = 0$ 。奇函数的图形大致如图4.4 b 所示。由于 y 轴一侧的正贡献会被另一侧

相应的负贡献抵消，我们可以得出：对于奇函数，有

$$\int_{-a}^{+a} g(x) dx = 0 \quad (4.51)$$

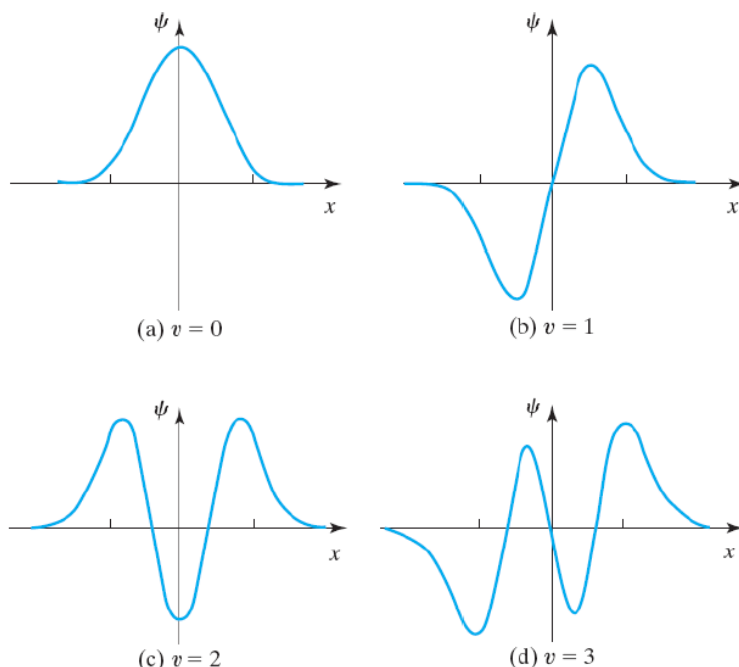


图 4.4:
谐振子波函数。

所有图形使用相同的比例尺。 x 轴上标记的点为 $\alpha^{1/2}x = \pm 2$ 。

很容易证明，两个偶函数或两个奇函数的乘积是偶函数，而一个偶函数和一个奇函数的乘积是奇函数。

谐振子波函数

式 (4.47) 中的指数部分 $e^{-\alpha x^2/2}$ 是偶函数。若 v 是偶数，则多项式因子只包含 x 的偶次幂，使得 ψ_v 是偶函数。若 v 是奇数，则多项式因子只包含 x 的奇次幂，使得 ψ_v 是一个奇函数和一个偶函数的乘积，为奇函数。每一个谐振子状态方程 ψ 都是一个奇函数或偶函数，取决于 v 的奇偶性。在第 7.5 节中，我们将看到：当势能函数 V 是偶函数时，非简并能级的波函数一定是奇函数或偶函数。

现在，我们找到了最低三个能级波函数的显式。对于 $v=0$ 的基态，式 (4.47) 给出

$$\psi_0 = c_0 e^{-\alpha x^2/2} \quad (4.52)$$

其中 ψ 的下标表示量子数 v 。通过归一化，我们可以求出 c_0 ：

$$1 = \int_{-\infty}^{\infty} |c_0| e^{-\alpha x^2} dx = 2 |c_0|^2 \int_0^{\infty} e^{-\alpha x^2} dx$$

其中，我们用到了式 (4.49)。使用附录中的积分 (A.9)，我们有 $|c_0| = (\alpha/\pi)^{1/4}$ 。因此，如果我们选择令归一化常数的相位为零，则基态波函数为

$$\psi_0 = (\alpha/\pi)^{1/4} e^{-\alpha x^2/2} \quad (4.53)$$

波函数 (4.53) 是一个高斯函数 (图4.4a)。

对于状态 $v = 1$, 式 (4.47) 给出

$$\psi_1 = c_1 x e^{-\alpha x^2/2} \quad (4.54)$$

使用 (A.10) 进行归一化, 我们有

$$\psi_1 = (4\alpha^3/\pi)^{1/4} x e^{-\alpha x^2/2} \quad (4.55)$$

图4.4b 显示了 ψ_1 的图象。

对于状态 $v = 2$, 式 (4.47) 给出

$$\psi_2 = (c_0 + c_2 x^2) e^{-\alpha x^2/2}$$

将 $v = 2$ 代入 (4.46), 我们有

$$c_2 = \frac{2\alpha(-2)}{1 \cdot 2} c_0 = -2\alpha c_0$$

因此,

$$\psi_2 = c_0 (1 - 2\alpha x^2) e^{-\alpha x^2/2} \quad (4.56)$$

使用归一化解出 c_0 , 我们有 (问题 4.10)

$$\psi_2 = (4\alpha^3/\pi)^{1/4} (2\alpha x^2 - 1) e^{-\alpha x^2/2} \quad (4.57)$$

注意: ψ_2 中的 c_0 和 ψ_0 中的 c_0 是不同的。

波函数节点的个数与量子数 v 相同。可以证明 (见《Messiah》, 第 109-110 页): 对于一维定态束缚态问题, 若 ψ 处于基态, 则边界点内部的节点数为零; 而对于每一个连续的激发态, 节点数都会增加一个。



图 4.5: 谐振子的经典允许区域 $-a \leq x \leq a$ 和禁阻区域 $x > a, x < -a$

谐振子波函数中的多项式因子在数学中是众所周知的, 被称为厄米多项式 (Hermite polynomials), 是以一位法国数学家的名字命名的。(见问题 4.21)

根据量子力学解法, 在 x 轴上的任何一点 (节点除外) 都有一定概率找到粒子。从经典角度讲, $E = T + V$, 动能 T 不能为负: $T \geq 0$ 。因此, $E - V = T \geq 0$ 则 $V \leq E$ 。势能 V 是位置的函数, 经典粒子被限制在 $V < E$ 的空间区域内。也就是说, 势能不超过总能量。在图 (4.5) 中, 标有 E 的水平线表示谐振子的能量, 抛物线表示势能 $\frac{1}{2}kx^2$ 。对于区域 $x < -a$ 和 $x > a$, 我们有 $V > E$, 这些

区域是经典禁阻 (classical forbidden) 的。经典允许的区域 (classical allowed region) $-a \leq x \leq a$ 在图 (4.5) 中满足 $V \leq E$ 。

在量子力学中, 定态波函数不是算符 \hat{T} 或 \hat{V} 的本征函数, 因此我们无法确定定态系统的动能或势能值。与经典力学的方程 $E = T + V$ 和 $T \geq 0$ 不同, 在量子力学中我们有 $\langle E \rangle = \langle T \rangle + \langle V \rangle$ (问题 6.35) 和 $\langle T \rangle \geq 0$ (问题 7.7), 所以量子力学中满足 $\langle V \rangle \leq E$, 但是我们不能写作 $V \leq E$, 因此, 粒子有一定概率在经典禁区 ($V > E$ 的区域) 中被发现。

如果说粒子可以在经典允许的区域之外被发现, 我们似乎就允许它具有负动能。实际上, 量子力学的观点并不存在悖论。为了验证粒子是否处于经典禁止区域, 我们必须测量它的位置。这种测量会改变系统的状态 (第 1.3 和 1.4 节)。振子与测量仪器的相互作用会将足够的能量传递给振子, 使其处于经典禁区。对 x 的精确测量会给动量带来很大的不确定性, 因此也会给动能带来很大的不确定性。2.4 节和 2.5 节讨论了经典禁区的穿透问题。

一个定态谐振子满足 $E = (v + \frac{1}{2})h\nu$ 和 $V = \frac{1}{2}kx^2 = 2\pi^2\nu^2mx^2$ 。因此, 在经典允许的区域 ($V \leq E$) 有 $2\pi^2\nu^2mx^2 \leq (v + \frac{1}{2})h\nu$, 即 $x^2 \leq (v + \frac{1}{2})h/2\pi^2\nu^2m = (2v + 1)\alpha$, 其中 $\alpha \equiv 2\pi\nu m/\hbar$ [式 (4.31)]。因此, 谐振子的经典允许区域为 $-(2v + 1)^{1/2}\alpha^{1/2} \leq x \leq (2v + 1)^{1/2}\alpha^{1/2}$ 。

从图 (4.4) 中可以看出: ψ 在经典允许区域中振荡, 而在经典禁止区域则以指数形式下降到零。我们之前在矩形势阱中看到过粒子的这种行为 (第 2.4 节)。

图 (4.4) 显示, 当谐振子的能量逐渐升高时, ψ 和 $|\psi|^2$ 的最大值往往离原点越来越远。由于离原点越远, 势能 $V = \frac{1}{2}kx^2$ 越大, 平均势能 $\langle V \rangle = \int_{-\infty}^{\infty} |\psi|^2 V dx$ 也会随着量子数的增加而增大。平均动能由 $\langle T \rangle = -(\hbar^2/2m) \int_{-\infty}^{\infty} \psi^* \psi'' dx$ 给出。应用分部积分 (问题 7.7b), 有 $\langle T \rangle = (\hbar^2/2m) \int_{-\infty}^{\infty} |d\psi/dx|^2 dx$ 。量子数越大的状态中节点的数量越多, ψ 的变化速度就越快, 因此随着量子数的增加, $\langle T \rangle$ 也会增加。

经典谐振子最有可能出现在运动转折点附近的区域, 即振子运动速度最慢、 V 值较大的区域。相反, 对于量子谐振子的基态, 最有可能出现的区域是原点附近的区域。对于高振动量子数, 人们会发现的外峰值大于原点附近的峰值, 最可能的区域变成了经典转折点附近的区域, 此时 V 很大 (见问题 4.18)。这是对对应原理 (第 2.2 节) 的一个例子。

量子谐振子的一些在线模拟可从以下网站获得: www.phy.davidson.edu/StuHome/cabell_f/energy.html (显示能级和波函数, 并显示当能量从允许值改变时波函数如何发散); www.falstad.com/qm1d/ (从顶部的下拉菜单中选择谐振子; 双击底部的小圆圈以显示静止状态; 显示能量、波函数、概率密度; m 和 k 可以改变); demonstrations.wolfram.com/HarmonicOscillatorEigenfunctions (展示了 $|\psi|^2$)。

4.3 双原子分子的振动

我们将在第 13.1 节中看到, 在极好的近似条件下, 我们可以分别处理分子中电子的运动和原子核的运动 (这是由于原子核的质量要大得多)。首先想象原子核静止不动, 然后求解电子能量 U 的薛定谔方程 (U 还包括核排斥能量)。对于二原子 (双原子) 分子, 电子能量 U 取决于原子核之间的距离 R , $U = U(R)$, 而 U 与 R 的关系曲线则如图 13.1 所示。

找到 $U(R)$ 后, 可以先解决核运动的薛定谔方程, 使用 $U(R)$ 作为核运动的势能函数。对于双原子分子, 核薛定谔方程是一个双粒子方程。我们将在第 6.3 节中看到, 当双粒子系统的势能只取决于粒子之间的距离时, 系统的能量是 (a) 整个系统在空间平动的动能和 (b) 粒子相对于彼此的内部运动能量之和。双粒子内部运动能量的经典表达式是粒子间相互作用的势能与假设粒子的动能之和, 其中假设粒子的质量为 $m_1 m_2 / (m_1 + m_2)$ (m_1 和 m_2 分别是两个原子的质量), 该粒子的坐标为一个粒子相对于另一个粒子的坐标。 $m_1 m_2 / (m_1 + m_2)$ 称为约化质量 (reduced mass) μ 。

二原子分子的内部运动包括**振动**和**转动**，前者与两个原子核之间距离 R 的变化相对应，后者与原子核连接线空间方向的变化相对应。通常可以近似地将振动和转动分开处理。转动能级见第 6.4 节。此处我们考虑振动能级。

双原子分子振动的薛定谔方程有一个质量为 $m_1 m_2 / (m_1 + m_2)$ 假设粒子的动能算符和一个由 $U(R)$ 给出的势能项。如果我们将原点与图 13.1 中 U 曲线的最低点重合，并将该最低点的能量取为势能的零点，那么 $U(R)$ 曲线的下半部分将与具有适当力常数 k 的谐振子的势能曲线几乎重合（见图 4.6 和问题 4.28）。 $U(R)$ 曲线的最小值出现在**平衡核间距**（equilibrium distance） R_e 的位置。在图 4.6 中， x 是核间距与平衡核间距的偏差： $x \equiv R - R_e$ 。



图 4.6:
二原子分子振动的势能（实线）和谐振子的势能（虚线）。

图中还显示了二原子分子的束缚态振动能级。

与谐振子相比，二原子分子只有有限数量的束缚振动能级。

谐振子的力常数 k 在式 (4.26) 中可以由 $k = d^2V/dx^2$ 计算得到，在 $R = R_e$ 处，谐振子曲线与 $U(R)$ 曲线的下半部分几乎重合，所以分子的力常数 $k = d^2U/dR^2|_{R=R_e}$ 。（另见问题 4.28）。核质量的差异对电子能量曲线 $U(R)$ 几乎没有影响，因此同一分子但由不同同位素取代的种类具有基本相同的力常数 k 。

因此，我们预计二原子分子振动能级 E_{vib} 的一个合理近似是谐振子振动能级；由式 (4.45) 和式 (4.23)，有

$$E_{vib} \approx \left(v + \frac{1}{2}\right) h\nu_e, \quad v = 0, 1, 2, \dots \quad (4.58)$$

$$\nu_e = \frac{1}{2\pi} \left(\frac{k}{\mu}\right)^{1/2}, \quad \mu = \frac{m_1 m_2}{m_1 + m_2}, \quad k = \left.\frac{d^2U}{dR^2}\right|_{R=R_e} \quad (4.59)$$

其中 ν_e 称为**简谐振动频率**（equilibrium or harmonic vibrational frequency）。这种近似方法最适用于较低的振动能级。随着 v 的增加，原子核位于远离其平衡间隔区域的时间更长。对于这些区域，势能大大偏离谐振子的势能，谐振子近似效果不佳。我们会发现：随着 v 的增大，双原子分子的振动能级不是等距排列，而是越靠越近（图 4.6）。最终，振动能量大到足以将双原子分子解离成互不结合的原子。与谐振子不同，双原子分子只有有限数量的束缚态振动能级。考虑到振动的非谐性，分

子振动能的更精确表达式是

$$E_{vib} = \left(v + \frac{1}{2}\right) h\nu_e - \left(v + \frac{1}{2}\right)^2 h\nu_e \chi_e \quad (4.60)$$

其中非谐性常数 (anharmonicity constant) $\nu_e \chi_e$ 几乎在所有情况下都大于零。(译者注: 一般认定 χ_e 为非谐性常数?)

使用含时薛定谔方程, 我们可以发现 (见第9.9节): 当双原子分子暴露在电磁辐射下时, 最有可能发生的振动跃迁是那些 v 变化了 ± 1 的跃迁。此外, 要吸收或发射电磁辐射, 振动必须改变分子的偶极矩。因此, 同核双原子分子 (例如 H_2 、 N_2) 不能通过吸收或发射辐射实现振动能级之间的转换 (这种转变可以发生在分子间碰撞中)。关系 $E_{upper} - E_{lower} = h\nu$ (其中 ν 为辐射频率)、近似方程 (4.58)、吸收和发射辐射的选律 (selection rule) $\Delta v = \pm 1$ 表明: 一个振动频率为 ν_e 的异核双原子分子将对近似为下式给出频率的辐射 ν_{light} 有强烈吸收:

$$\nu_{light} = (E_2 - E_1) / h \approx \left[\left(v_2 + \frac{1}{2}\right) h\nu_e - \left(v_1 + \frac{1}{2}\right) h\nu_e \right] / h = (v_2 - v_1) \nu_e = \nu_E \quad (4.61)$$

双原子分子 k 和 μ 值的特性使得 ν_{light} 通常位于光谱的红外区域。 $\Delta v = 2, 3, \dots$ 的跃迁也会发生, 但这些 (称为泛频 (overtones)) 的吸收要比 $\Delta v = 1$ 的跃迁弱得多。

使用更准确的方程 (4.60), 有 (问题 4.27)

$$\nu_{light} = \nu_e - 2\nu_e \chi_e (v_1 + 1) \quad (4.62)$$

其中 v_1 是较低能级的振动量子数, 选律为 $\Delta v = \pm 1$ 。

热平衡系统中分子两个能级的布居数由玻尔兹曼分布律 (Boltzmann distribution law) 给出 (见物理化学教材):

$$\frac{N_i}{N_j} = \frac{g_i}{g_j} e^{-(E_i - E_j)/kT} \quad (4.63)$$

其中能级 i 和 j 的能量分别为 E_i 和 E_j , g_i 和 g_j 是对应能级的简并度, k 是玻尔兹曼常数, T 是热力学温度。对于非简并能级, 有 $g_i = 1$ 。

$\nu = (1/2\pi)(k/\mu)^{1/2}$ 的数量级使得在室温下, 对于轻分子 (例如 H_2 、 HCl 、 CO), 只有 $v = 0$ 的振动能级是被显著占据的。对于重分子 (例如 I_2), 在室温下有一个或多个激发态振动能级被显著占据。

极性双原子分子的振动吸收光谱包含 $v = 0 \rightarrow 1$ 吸收带, 很弱的泛频吸收带 ($v = 0 \rightarrow 2, 0 \rightarrow 3, \dots$), 以及若 $v > 0$ 的能级也被显著占据, 也有热带 (hot bands), 例如 $v = 1 \rightarrow 2$ 、 $v = 1 \rightarrow 3$ 等。与特定振动跃迁相对应的每个波段都由几条紧密相连的线组成。每一条这样的线都对应于与振动状态变化同时发生的不同的转动状态变化。每一条线都是振动-转动跃迁的结果。

光谱频率的国际单位为赫兹 (hertz, Hz), 定义为 $1 \text{ Hz} \equiv 1 \text{ s}^{-1}$ 。通常使用的倍数有等于 10^6 赫兹的兆赫兹 (MHz) 和等于 10^9 赫兹的千兆赫兹 (GHz)。红外 (IR) 吸收峰通常由波数 (wavenumber) $\tilde{\nu}$ 给出, 定义为

$$\tilde{\nu} \equiv \frac{\nu}{c} = \frac{1}{\lambda} \quad (4.64)$$

其中 λ 是真空中辐射的波长。

在谐振子近似中, 多原子分子的量子力学能级为 $E_{vib} = \sum i (v_i + \frac{1}{2}) h\nu_i$, 其中 ν_i 是分子每个正则振动模式的频率, v_i 是对应第 i 个正则振动的量子数。每个振动量子数 v_i 的取值 $0, 1, 2, \dots$ 都与其他振动量子数的取值无关。一个具有 n 个原子的线性分子具有 $3n - 6$ 个正则振动模式, 而一个具有 n 个原子的非线性分子具有 $3n - 5$ 个正则振动模式。

为了由式 (4.59) 计算折合质量 μ ，我们需要知道每种同位素的质量。附录表 A.3 列出了一些常见同位素的相对原子质量。

例题： $^{12}\text{C}^{16}\text{O}$ 分子的最强红外吸收峰位于 $\tilde{\nu} = 2140 \text{ cm}^{-1}$ 。计算该分子的力常数 k 。请说明所做的近似计算。

最强的红外吸收峰对应于 $v = 0 \rightarrow 1$ 的跃迁。我们将分子振动近似为谐振子模型。由式 (4.61)，分子的平衡振动频率近似为

$$\nu_e \approx \nu_{\text{light}} = \tilde{\nu}c = (2140 \text{ cm}^{-1})(2.9979 \times 10^{10} \text{ cm/s}) = 6.424 \times 10^{13} \text{ s}^{-1}$$

使用式 (4.59)，将 k 与 ν_e 联系起来，我们需要折合质量 $\mu = m_1 m_2 / (m_1 + m_2)$ 。1 mol ^{12}C 的质量为 12 g，包含阿伏伽德罗常数 (N_A) 个分子，则一个 ^{12}C 原子的质量为 $(12 \text{ g}) / (6.02214 \times 10^{23})$ 。折合质量和力常数分别为

$$\mu = \frac{12(15.9949) \text{ g}}{27.9949} \frac{1}{6.02214 \times 10^{23}} = 1.1385 \times 10^{-23} \text{ g}$$

$$k = 4\pi^2 \nu_e^2 \mu = 4\pi^2 (6.424 \times 10^{13} \text{ s}^{-1})^2 (1.1385 \times 10^{-23} \text{ g}) = 1855 \text{ N/m}$$

练习： (a) 求分子 $^{12}\text{C}^{16}\text{O}$ 零点能的近似值；(答案： $2.1 \times 10^{-20} \text{ J}$)；(b) 求分子 $^{13}\text{C}^{16}\text{O}$ 的 ν_e 。(答案： $6.28 \times 10^{13} \text{ s}^{-1}$)

4.4 一维定态薛定谔方程的数值解法

Numerov 方法

我们已经精确求解了箱中粒子和谐振子的薛定谔方程。对于许多势能函数 $V(x)$ ，单粒子一维薛定谔方程是无法精确求解的。本节将介绍一种用于计算机求解单粒子一维薛定谔方程的数值方法 (**Numerov 方法**)，该方法可以精确求出任意势能函数 $V(x)$ 的薛定谔方程的束缚态本征值和本征函数。

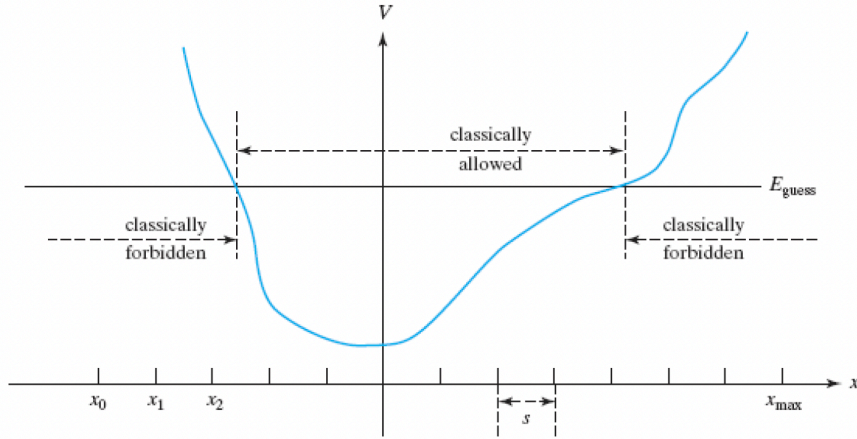
为了使用数值法求解薛定谔方程，我们需要处理 x 轴的一部分，其中包括经典允许区域，以及在经典允许区域两端延伸到经典禁止区域的部分。我们将 x 轴的这一部分划分为若干小区间，每个区间的长度为 s (图 4.7)。点 x_0 和 x_{max} 是这一部分的边界点，而 x_n 是第 n 个区间的右端点。令 ψ_{n-1} 、 ψ_n 和 ψ_{n+1} 分别为点 x_{n-s} 、 x_n 和 x_{n+s} 处波函数 ψ 的值，分别地 (这些点是相邻区间的端点)：

$$\psi_{n-1} \equiv \psi(x_n - s), \quad \psi_n \equiv \psi(x_n), \quad \psi_{n+1} \equiv \psi(x_n + s) \quad (4.65)$$

不要被这些符号迷惑了。下标 $n-1$ 、 n 和 $n+1$ 并代表不同的状态，而是表示 x 轴上以 s 为间隔的各点上一个特定波函数 ψ 的值。下标 n 表示 ψ 在 x_n 点求值。我们将薛定谔方程 $-(\hbar^2/m)\psi'' + V\psi = E\psi$ 写成

$$\psi'' = G\psi, \quad \text{其中} \quad G \equiv m\hbar^{-2} [2V(x) - 2E] \quad (4.66)$$

用涉及 s 的幂的 Taylor 级数展开 $\psi(x_n + s)$ 和 $\psi(x_n - s)$ ，将两式相加即可消去 s 的奇次幂项，并用薛定谔方程来表示 ψ'' 和 $\psi^{(\text{iv})}$ ，以 ψ 表示，并忽略 s^6 和 s 的高次幂项 (如果 s 较小，近

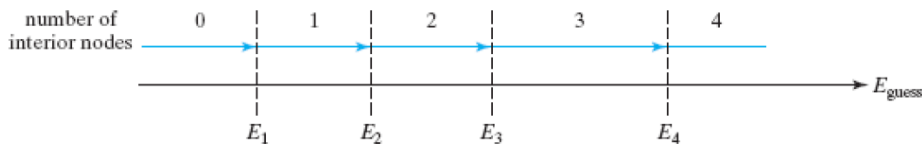
图 4.7: 一维单粒子系统的 V 关于 x 的图象

似值将是准确的), 我们会发现 (问题 4.43)

$$\psi_{n+1} \equiv \frac{2\psi_n - \psi_{n-1} + 5G_n\psi_n s^2/6 + G_{n-1}\psi_{n-1}s^2/12}{1 - G_{n+1}s^2/12} \quad (4.67)$$

其中 $G_n \equiv G(x_n) \equiv m\hbar^{-2}[2V(x_n) - 2E]$ [式 (4.66) 和 4.65]。如果知道 ψ_n 和 ψ_{n-1} , 即 ψ 在 x_n 和 $x_n - s$ 处的值, 方程 (4.67) 允许我们计算 ψ_{n+1} , 即 ψ 在点 x_{n+s} 处的值。

我们如何使用式 (4.67) 来求解薛定谔方程? 我们首先猜测一个能量本征值 E_{guess} 。我们从左侧经典禁区 (图 4.7) 中的一点 x_0 开始, 此处 ψ 的值很小, 我们假设 ψ 在该点处的值为零: $\psi_0 \equiv \psi(x_0) = 0$ 。同样地, 我们也在经典禁区 (此处 ψ 的值也很小) 的右侧点 x_{max} 处假设 $\psi(x_{max})$ 的值为零。我们为连续点之间的间隔 s 选取一个较小的数值, 在 $x_0 + s$ 处, 我们将 ψ 的值设为一个较小的数值, 例如 0.0001: $\psi_1 \equiv \psi(x_1) \equiv \psi(x_0 + s) = 0.0001$ 。 ψ_1 的值对找到的本征值不会有任何影响。如果我们令 ψ_1 为 0.001, 而不是 0.0001, 公式 (4.67) 表明, 这只需将后续点上的所有 ψ 值乘以 10 即可 (问题 4.41)。这不会影响本征值 [见公式 (3.14) 后的示例]。在找到每个本征值后, 可以对波函数进行归一化处理。

图 4.8: Numerov 方法求解的节点数与能量 E_{guess} 的函数关系。

现在我们有 ψ_0 和 ψ_1 的值, 可以使用式 (4.67) 计算 $\psi_2 \equiv \psi(x_2) \equiv \psi(x_1 + 2s)$ 的值, 其中我们用 E_{guess} 计算得到 G 的值。接着, 在式 (4.67) 令 $n = 2$, 我们可以计算出 ψ_3 的值, 以此类推, 直到达到 x_{max} 。如果 E_{guess} 不等于或不非常接近于本征值, 那么 ψ 将不满足平方可积的要求, 且 $|\psi(x_{max})|$ 的值会变得非常大。如果发现 $|\psi(x_{max})|$ 不接近于零, 那我们重新选一个 E_{guess} , 并从 x_0 开始重复上述过程。这个过程将会循环往复, 直到找到能使 $|\psi(x_{max})|$ 非常接近于零的 E_{guess} 的值。此时的 E_{guess} 基本上等同于一个本征值。找到本征值的系统方法是计算由 E_{guess} 生成的 ψ 中的节点。回忆在一维问题中 (第 4.2 节), 基态的内部节点数为 0, 第一激发态的内部节点数为 1, 以此类推。令 E_1, E_2, E_3, \dots 分别表示基态、第一激发态和第二激发态等的能量值。如果 ψ_{guess} 在 x_0 和 x_{max} 之间没有节点, 那么 E_{guess} 应该小于或等于 E_1 。如果 ψ_{guess} 在 x_0 和 x_{max} 之间有一个

节点, 那么 E_{guess} 应该位于 E_1 和 E_2 之间。随后将给出例子。

无量纲变量

Numerov 方法需要我们猜测 E 的值。但是我们猜测的数量级: 10^{-20} J, 10^{-15} J, ... 应该是多少? 为了回答这个问题, 我们以谐振子为例, 用无量纲变量重表示薛定谔方程。

谐振子的势能函数为 $V = \frac{1}{2}kx^2$, 其薛定谔方程包含三个常数, k 、 m 和 \hbar 。我们试图找到无量纲约化能量 E_r 和无量纲约化 x 坐标 x_r , 它们的定义是

$$E_r \equiv E/A, \quad x_r \equiv x/B \quad (4.68)$$

其中常数 A 是 k 、 m 和 \hbar 的组合, 具有能量的量纲, 而 B 是具有长度量纲的组合。能量的量纲为 质量 \times 长度² \times 时间⁻², 我们写作

$$[E] = \text{ML}^2\text{T}^{-2} \quad (4.69)$$

其中, E 的括号表示其维度, M、L 和 T 分别表示质量、长度和时间。方程 $V = \frac{1}{2}kx^2$ 表明 k 的量纲为 能量 \times 长度⁻², 由式 (4.69), 有 $[k] = \text{MT}^{-2}$ 。 \hbar 的量纲为 能量 \times 时间。因此,

$$[m] = \text{M}, \quad [k] = \text{MT}^{-2}, \quad [\hbar] = \text{ML}^2\text{T}^{-1} \quad (4.70)$$

式 (4.68) 中的 A 和 B 的量纲分别为能量和长度, 则

$$[A] = \text{ML}^2\text{T}^{-2}, \quad [B] = \text{L} \quad (4.71)$$

令 $A = m^a k^b \hbar^c$, 其中 a 、 b 和 c 是待定的指数, 由 A 的量纲一定为 ML^2T^{-2} 所确定。我们有

$$[A] = [m^a k^b \hbar^c] = \text{M}^a (\text{MT}^{-2})^b (\text{ML}^2\text{T}^{-1})^c = \text{M}^{a+b+c} \text{L}^{2c} \text{T}^{-2b-c} \quad (4.72)$$

令 (4.71) 和 (4.72) 中 M、L 和 T 的指数相等, 得到

$$a + b + c = 1, \quad 2c = 2, \quad -2b - c = -2$$

解这个方程组, 得到 $c = 1$, $b = \frac{1}{2}$, $a = -\frac{1}{2}$ 。因此

$$A = m^{-1/2} k^{1/2} \hbar \quad (4.73)$$

同样地, 令 $B = m^d k^e \hbar^f$ 。用得出式 (4.73) 相同的量纲分析法, 有 (问题 4.44)

$$B = m^{-1/4} k^{-1/4} \hbar^{1/2} \quad (4.74)$$

根据式 (4.68)、(4.73) 和 (4.74), 谐振子的约化变量为

$$E_r = E/m^{-1/2} k^{1/2} \hbar, \quad x_r = x/m^{-1/4} k^{-1/4} \hbar^{1/2} \quad (4.75)$$

使用 $k^{1/2} = 2\pi\nu m^{1/2}$ [式 (4.23)] 将 k 从 (4.75) 中消去, 并回忆 $\alpha \equiv 2\pi\nu m/\hbar$ [式 (4.31)], 我们得到以下备选的表达式

$$E_r = E/h\nu, \quad x_r = \alpha^{1/2} x \quad (4.76)$$

与方程 $E_r \equiv E/A$ [式 (4.68)] 类似, 我们可以定义约化势能函数 V_r :

$$V_r \equiv V/A \quad (4.77)$$

由于 $|\psi(x)|^2 dx$ 是概率，概率是没有量纲的，归一化后的 $\psi(x)$ 一定具有长度^{-1/2} 的量纲。因此，我们可以定义一个没有量纲的约化波函数 ψ_r 。由式 (4.71)， B 具有长度的量纲，则 $B^{-1/2}$ 具有长度^{-1/2} 的量纲。因此，

$$\psi_r = \psi/B^{-1/2} \quad (4.78)$$

ψ_r 满足 $\int_{-\infty}^{\infty} |\psi_r|^2 dx_r = 1$ ；这是由式 (4.76)、(4.78) 和 $\int_{-\infty}^{\infty} |\psi|^2 dx = 1$ 得出的。

我们现在使用无量纲变量 E_r 、 x_r 和 ψ_r 来重写薛定谔方程。有

$$\begin{aligned} \frac{d^2\psi}{dx^2} &= \frac{d^2}{dx^2} B^{-1/2} \psi_r = B^{-1/2} \frac{d}{dx} \frac{d\psi_r}{dx} = B^{-1/2} \frac{d}{dx} \frac{d\psi_r}{dx_r} \frac{dx_r}{dx} = B^{-1/2} \frac{d}{dx_r} \frac{d\psi_r}{dx_r} \frac{dx_r}{dx} \frac{dx_r}{dx} \\ &= B^{-5/2} \frac{d^2\psi_r}{dx_r^2} \end{aligned} \quad (4.79)$$

其中，我们用到了 $dx_r/dx = B^{-1/2}$ 。将式 (4.68)、(4.77) 和 (4.79) 代入薛定谔方程 $-(\hbar^2/m)(d^2\psi/dx^2) + V\psi = E\psi$ ，得到

$$\begin{aligned} -\frac{\hbar^2}{m} B^{-5/2} \frac{d^2\psi_r}{dx_r^2} + AV_r B^{-1/2} \psi_r &= AE_r B^{-1/2} \psi_r \\ -\frac{\hbar^2}{2m} \frac{1}{AB^2} \frac{d^2\psi_r}{dx_r^2} + V_r \psi_r &= E_r \psi_r \end{aligned} \quad (4.80)$$

根据式 (4.73) 和 (4.74)，有 $AB^2 = \hbar^2/m$ ，所以对于谐振子，有 $\hbar^2/mAB^2 = 1$ 。

更一般地说，让 V 包含一个非无量纲参数 c 。例如，我们有 $V = cx^4$ 或 $V = cx^2(1 + 0.05m^{1/2}c^{1/2}\hbar^{-1}x^2)$ （注意 $m^{1/2}c^{1/2}\hbar^{-1}x^2$ 是无量纲的，因为 1 是无量纲的，所以它必须是无量纲的）。式 (4.80) 中的 AB 一定可以写成 $\hbar^r m^s c^t$ 的形式，其中 r 、 s 和 t 是待定的指数。由于 (4.80) 中的 $V_r \psi_r$ 是无量纲的，因此第一项也是无量纲的。那么 \hbar^2/mAB^2 是无量纲的， AB^2 与 \hbar^2/m 具有相同的量纲。所以 $r = 2$ ， $s = 1$ ， $t = 0$ 。因此， $AB^2 = \hbar^2/m$ ，带入 (4.80)，有无量纲的薛定谔方程

$$\frac{d^2\psi}{dx_r^2} = (2V_r - 2E_r) \psi_r \quad (4.81)$$

$$\psi'' = G_r \psi_r, \quad \text{其中 } G_r \equiv 2V_r - 2E_r \quad (4.82)$$

对于谐振子， $V_r \equiv V/A = \frac{1}{2}kx^2/m^{-1/2}\hbar k^{1/2} = \frac{1}{2}x_r^2$ ，因此 [式 (4.73) 和式 (4.75)]:

$$V_r = \frac{1}{2}x_r^2 \quad (4.83)$$

将谐振子薛定谔方程约化为只涉及无量纲量的形式 (4.81) 后，我们可以预期最低能量本征值的数量级为 1。

约化谐振子薛定谔方程 (4.82) 与 (4.66) 具有相同的形式，所以我们可以用 Numerov 公式 (4.67) 来求解。使用 ψ_r 的值代替 ψ ， G_r 的值代替 G ，与 (4.68) 类似， s_r 的值代替 s 。

一旦找到约化能量 E_r 的数值，就可以根据 (4.75) 或 (4.76) 求出能量 E 。

x_r 、 x_0 、 x_{max} 和 s_r 的选择

我们需要选择 x_r 的起始值、结束值，以及相邻点之间的间隔值 s_r 。假设我们希望找到满足 $E_r \leq 5$ 的谐振子的所有本征函数和本征值。我们从左侧经典禁区开始求解，因此首先要找到 $E_r = 5$ 的经典禁区。经典允许和经典禁止区域的边界满足 $E_r = V_r$ 。由式 (4.83)，有 $V_r = \frac{1}{2}x_r^2$ ，因此 $E_r = V_r$ 变成了 $5 = \frac{1}{2}x_r^2$ 所以满足 $E_r \leq 5$ 的经典允许区域为 $x_r = -(10)^{1/2} = -3.16$ 到 $x_r = (10)^{1/2} = 3.16$ 。若 $E_r < 5$ ，则经典允许区域将更小。我们希望从左侧经典禁区 (ψ 非常小) 内的某一点开始求解，

我们希望在右侧经典禁区内的某一点结束求解。对于 $E_r = 5$ ，左侧的经典禁区在 $x_r = -3.16$ 处结束，，所以一个合理的选择为 $x_r = -5$ [太深入经典禁区有时会带来麻烦（见下文），因此在选择起点时可能需要试错]。由于 V 是对称的，因此我们将在 $x_r = 5$ 处结束求解。

为了达到合理的精确度，通常至少需要 100 个点，因此我们取 $s_r = 0.1$ 以得到 100 个点。从 Numerov 方法的推导中可以看出， s_r 必须很小。合理的规则可能是 s_r 不大于 0.1。

如果像通常情况那样，当 $x \rightarrow \pm\infty$ 时有 $V \rightarrow \infty$ ，那么在经典禁区内起步太远会使 Numerov 公式 (4.67) 中的分母 $1 - G_{n+1}s^2/12$ 变成负数。我们有 $G_r = 2V_r - 2E_r$ ，如果我们从 V_r 非常大的点 x_0 开始，该点的 G_r 可能大到足以使 Numerov 公式的分母为负。这样，该方法就会失效。我们将 ψ_0 取为零，将 ψ_1 取为正数。Numerov 公式 (4.67) 表明，如果分母是负数，那么 ψ_2 也将是负数，我们将在 x_1 和 x_2 之间产生一个虚假的 ψ 节点。为了避免这个问题，我们可以减小步长 s_r 或 x_r 、 x_{max} 、 $-x_r$ 和 $-x_0$ 。（见问题 4.46）。

Numerov 方法的计算机程序

使用电子表格求解一维薛定谔方程

Numerov 方法步骤摘要

总结

习题

Chapter 5

角动量

5.1 同时确定多个属性

在本章中，我们将讨论角动量，在下一章中，我们将证明对于氢原子的静止态，电子角动量的大小是恒定的。首先，我们要考虑用什么标准来决定一个系统的哪些性质可以同时赋予定值。

在第3.3节中，我们假设：若态函数 Ψ 是算符 \hat{A} 的本征函数，本征值为 s ，则对物理量 A 的测量结果一定为 s 。若 Ψ 同时是 \hat{A} 和 \hat{B} 的本征函数，即 $\hat{A}\Psi = s\Psi$ ， $\hat{B}\Psi = t\Psi$ ，那么我们就可以同时为物理量 A 和 B 赋以确定的值。什么时候 Ψ 会同时是两个算符的本征函数呢？在第7章中，我们将证明如下两个理论。首先，两个算符的同时本征函数的完整集合存在的必要条件是算符之间可对易（这里使用的“完整”一词有一定的技术含义，我们将在第7章再讨论这个问题）。反之亦然，若对应于物理量的两个算符 \hat{A} 和 \hat{B} 可对易，则存在一组完整的函数，它们同时是 \hat{A} 和 \hat{B} 的本征函数。因此，若 $[\hat{A}, \hat{B}] = 0$ ，那么 Ψ 可以同时是 \hat{A} 和 \hat{B} 的本征函数。

回忆算符 \hat{A} 和 \hat{B} 的交换子为 $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$ [式 (3.7)]。下列等式有助于计算交换子。通过详细写出交换子可以证明这些等式（问题 5.2）：

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \quad (5.1)$$

$$[\hat{A}, \hat{A}^n] = 0, \quad n = 1, 2, 3, \dots \quad (5.2)$$

$$[k\hat{A}, \hat{B}] = [\hat{A}, k\hat{B}] = k[\hat{A}, \hat{B}] \quad (5.3)$$

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \quad [\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}] \quad (5.4)$$

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}] \quad [\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \quad (5.5)$$

其中 k 是常数，所有算符都假设为线性算符。

例题：从 $[\partial/\partial x, x] = 1$ [式 (3.8)] 出发，使用对易子的性质（式 (5.1) 至 (5.5)）求以下对易子：

(a) $[\hat{x}, \hat{p}_x]$ ；(b) $[\hat{x}, \hat{p}_x^2]$ ；(c) 三维单粒子系统的 $[\hat{x}, \hat{H}]$ 。

(a) 使用式 (5.3) 和 (5.1)，因为 $[\partial/\partial x, x] = 1$ ，我们有

$$\begin{aligned} [\hat{x}, \hat{p}_x] &= \left[\hat{x}, \frac{\hbar}{i} \frac{\partial}{\partial x} \right] = \frac{\hbar}{i} \left[\hat{x}, \frac{\partial}{\partial x} \right] = -\frac{\hbar}{i} \left[\frac{\partial}{\partial x}, \hat{x} \right] = -\frac{\hbar}{i} \\ &[\hat{x}, \hat{p}_x] = i\hbar \end{aligned} \quad (5.6)$$

(b) 由式 (5.5) 和 (5.6), 有

$$\begin{aligned} [\hat{x}, \hat{p}_x^2] &= [\hat{x}, \hat{p}_x] \hat{p}_x + \hat{p}_x [\hat{x}, \hat{p}_x] = i\hbar \cdot \frac{\hbar}{i} \frac{\partial}{\partial x} + \frac{\hbar}{i} \frac{\partial}{\partial x} \cdot i\hbar \\ &= 2\hbar^2 \frac{\partial}{\partial x} \end{aligned} \quad (5.7)$$

(c) 使用式 (5.4)、(5.3) 和 (5.7), 有

$$\begin{aligned} [\hat{x}, \hat{H}] &= [\hat{x}, \hat{T} + \hat{V}] = [\hat{x}, \hat{T}] + [\hat{x}, \hat{V}(x, y, z)] = [\hat{x}, \hat{T}] \\ &= [\hat{x}, (1/2m) (\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2)] \\ &= (1/2m) [\hat{x}, \hat{p}_x^2] + (1/2m) [\hat{x}, \hat{p}_y^2] + (1/2m) [\hat{x}, \hat{p}_z^2] \\ &= \frac{1}{2m} 2\hbar^2 \frac{\partial}{\partial x} + 0 + 0 \\ &= \frac{\hbar^2}{m} \frac{\partial}{\partial x} = \frac{i\hbar}{m} \hat{p}_x \end{aligned} \quad (5.8)$$

练习: 求证: 对于单粒子三维系统, 有

$$[\hat{p}_x^2, \hat{H}] = -i\hbar \partial V(x, y, z) / \partial x \quad (5.9)$$

这些对易子有重要的物理影响。由于 $[\hat{x}, \hat{p}_x] \neq 0$, 我们不能希望态函数同时是 \hat{x} 和 \hat{p}_x 的本征函数。因此, 与不确定性原理相符, 我们不能同时测定 x 和 p_x 的精确值。由于 \hat{x} 与 \hat{H} 不可对易, 我们不能指望同时为能量和 x 坐标赋予确定的值。定态 (具有确定的能量) 会显示出 x 的各种可能值, 观察到各种 x 值的概率由玻恩公设给出。

对于不是 \hat{A} 的本征函数的态函数 Ψ , 在完全相同的系统中对 A 的测量将得到各种可能的结果。我们需要对观测值 A_i 的分布或离散性进行某种测量。如果 $\langle A \rangle$ 是这些值的平均值, 那么每个测量值与平均值的偏差为 $A_i - \langle A \rangle$ 。如果我们对所有偏差求平均值, 就会得到零, 因为正负偏差会抵消。因此, 为了使所有偏差都为正值, 我们要将它们平方。偏差平方的平均值称为 A 的方差 (variance), 在统计学中用 $\sigma^2(A)$ 表示, 在量子力学中用 $(\Delta A)^2$ 表示:

$$(\Delta A)^2 \equiv \sigma^2(A) = \langle (A - \langle A \rangle)^2 \rangle = \int \Psi^* (\hat{A} - \langle A \rangle)^2 \Psi d\tau \quad (5.10)$$

其中我们用到了平均值的表达式 (3.88)。定义 (5.10) 与以下表达式等价 (问题 5.7):

$$(\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2 \quad (5.11)$$

方差的正平方根称为**标准差** (standard deviation), 用 ΔA 或 $\sigma(A)$ 表示。标准差是最常用的价差度量, 我们将用它来度量属性 A 的“不确定性”。

罗伯逊 (Robertson) 于 1929 年证明, 状态函数为 Ψ 的量子力学系统的两个属性的标准偏差的乘积必须满足不等式

$$\sigma(A) \sigma(B) = \Delta A \Delta B \geq \frac{1}{2} \left| \int \Psi^* [\hat{A}, \hat{B}] \Psi d\tau \right| \quad (5.12)$$

(5.12) 的证明源于量子力学公设, 见问题 7.60。如果 \hat{A} 和 \hat{B} 可对易, 那么式 (5.12) 中的积分为零, 则 ΔA 和 ΔB 可能都为零, 与上文的讨论相符。

作为式 (5.12) 的一个例子, 我们使用式 (5.6)、 $|z_1 z_2| = |z_1| |z_2|$ [式 (1.34)], 以及归一化方法:

$$\Delta x \Delta p_x \geq \frac{1}{2} \left| \int \Psi^* [\hat{x}, \hat{p}_x] \Psi d\tau \right| = \frac{1}{2} \left| \int \Psi^* i\hbar \Psi d\tau \right| = \frac{1}{2} \hbar \left| \int \Psi^* \Psi d\tau \right|$$

$$\sigma(x)\sigma(p_x) \equiv \Delta x \Delta p_x \geq \frac{1}{2}\hbar \quad (5.13)$$

式 (5.13) 通常被认为是海森堡不确定性原理 (Heisenberg Uncertainty Principle) (第 1.3 节) 的定量表述。然而, 公式 (5.12) 和 (5.13) 中标准偏差的含义与第 1.3 节中不确定度的含义截然不同。为了在式 (5.13) 中找到 Δx , 我们选取大量的系统, 每个系统都有相同的状态函数 ψ , 在每个系统中对 x 进行一次测量。将测得的数据标记为 x_i , 计算 $\langle x \rangle$ 和偏差的平方 $(x_i - \langle x \rangle)^2$ 。我们求出偏差平方的平均值, 得到方差, 再取平方根, 得到标准偏差 $\sigma(x) \equiv \Delta x$ 。然后我们再取大量的系统, 它们具有与得到 Δx 时一致的状态 Ψ , 对每个系统测量一次 p_x , 由这些数据计算出 Δp_x 。因此, (5.13) 中的统计量 Δx 和 Δp_x 不是单个测量的误差, 也不是通过同时测量 x 和 p_x 得出的。

令 $\varepsilon(x)$ 是对 x 的单次测量的典型误差, 令 $\eta(p_x)$ 是测量 x 时对 p_x 造成的典型干扰。1927 年, 海森堡分析了进行位置测量的具体思想实验, 得出结论: 乘积 $\varepsilon(x)\eta(x)$ 的数量级为 \hbar 或更大。海森堡没有给出这些量的精确定义。小泽正直 (Masanao Ozawa) 将海森堡的关系改写为

$$\varepsilon(x)\eta(p_x) \geq \frac{1}{2}\hbar$$

其中 $\varepsilon(x)$ 定义为 x 的测量值与理论值的方均根偏差, $\eta(p_x)$ 定义为测量 x 所产生的 p_x 变化的方均根偏差。更一般地说, 对于任意两个性质, 小泽将海森堡不确定性原理写成

$$\varepsilon(A)\eta(B) \geq \frac{1}{2} \left| \int \Psi^* [\hat{A}, \hat{B}] \Psi d\tau \right| \quad (5.14)$$

小泽提出的论点是: 在某些情况下, 海森堡不等式 (5.14) 可能被违反。小泽推导出以下关系来取代 (5.14) [M. Ozawa, Phys. Rev. A, 67, 042105 (2003); available at arxiv.org/abs/quant-ph/0207121]:

$$\varepsilon(A)\eta(B) + \varepsilon(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{1}{2} \left| \int \Psi^* [\hat{A}, \hat{B}] \Psi d\tau \right|$$

其中 $\sigma(A)$ 和 $\sigma(B)$ 由式 (5.10) 给出。2012 年, 一项测量中子自旋分量的实验发现: 自旋分量不服从海森堡误差-扰动不等式 (5.14), 但服从小泽不等式 [J. Erhart 等, 《自然物理学》, 8, 185 (2012); arxiv.org/abs/1201.1833]

另一个不等式是海森堡不确定关系, 即用一台仪器同时测量 A 和 B 的两种性质:

$$\varepsilon(A)\varepsilon(B) \geq \frac{1}{2} \left| \int \Psi^* [\hat{A}, \hat{B}] \Psi d\tau \right|$$

其中 $\varepsilon(A)$ 和 $\varepsilon(B)$ 分别是测量 A 与 B 时的实验误差。这一关系已被证明是成立的, 前提是某个可信的假设 (相信对目前所有可用的测量设备都是成立的) 成立 (见上文引用的小泽论文参考文献 6-12)。

例题: 式 (3.91)、(3.92)、(3.39)、(3.89) 之后的等式和问题 3.48 指出: 对于单粒子三维系统的基态, 有

$$\langle x \rangle = \frac{a}{2}, \quad \langle x^2 \rangle = \left(\frac{1}{3} - \frac{1}{2\pi^2} \right) a^2, \quad \langle p_x \rangle = 0, \quad \langle p_x^2 \rangle = \frac{h^2}{4a^2}$$

使用这些结果证明该系统是否遵守不确定性原理 (5.13)。

我们有

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \left(\frac{1}{3} - \frac{1}{2\pi^2} \right) a^2 - \frac{a^2}{4} = \frac{\pi^2 - 6}{12\pi^2} a^2$$

$$\Delta x = \left(\frac{\pi^2 - 6}{12} \right)^{1/2} \frac{a}{\pi}$$

$$\Delta p_x = \langle p_x^2 \rangle - \langle p_x \rangle^2 = \frac{h^2}{4a^2}, \quad \Delta p_x = \frac{h}{2a}$$

$$\Delta x \Delta p_x = \left(\frac{\pi^2 - 6}{12} \right)^{1/2} \frac{h}{2\pi} = 0.568\hbar > \frac{1}{2}\hbar$$

还有一种涉及能量和时间的不确定关系：

$$\Delta E \Delta t \geq \frac{1}{2}\hbar \quad (5.15)$$

一些文章指出：(5.15) 是通过将 $i\hbar \partial/\partial t$ 作为能量算符，并乘以 t 作为时间算符而从 (5.12) 推导出来的。然而，能量算符是哈密顿 \hat{H} ，而不是 $i\hbar \partial/\partial t$ 。此外，时间不是一种可观测的参数，而是量子力学中的一个参数。因此，不存在量子力学的时间算符（量子力学中的名词 **可观测** (observable) 指的是系统的物理可测量属性）。方程 (5.15) 必须通过特殊处理才能得出，我们略去不表（见 Ballentine, 第 12.3 节）。式 (5.15) 的推导表明： Δt 被解释为能量不确定度为 ΔE 的状态的时间寿命。通常情况下，(5.15) 中的 Δt 是能量测量的持续时间。然而，Aharonov 和 Bohm 已经证明：“能量可以在任意短的时间内重复测量” [Y. Aharonov and D. Bohm, *Phys. Rev.*, 122, 1649 (1961); 134, B1417 (1964); see also S. Massar and S. Popescu, *Phys. Rev. A*, 71, 042106 (2005); P. Busch, *The Time-Energy Uncertainty Relation*, arxiv.org/abs/quant-ph/0105049].

现在，我们来考虑同时为三个物理量赋予确定值的可能性：假设

$$[\hat{A}, \hat{B}] = 0, \quad [\hat{A}, \hat{C}] = 0 \quad (5.16)$$

这是否足以确保存在三个算符的同时本征函数？由于 $[\hat{A}, \hat{B}] = 0$ ，我们可以为 \hat{A} 和 \hat{B} 构造一组共同的本征函数。由于 $[\hat{A}, \hat{C}] = 0$ ，我们可以为 \hat{A} 和 \hat{C} 构造一组共同的本征函数。如果这两组本征函数是相同的，那么我们会为所有三个算符找到一组共同的本征函数。因此我们要问：线性算符 \hat{A} 的本征函数集是唯一确定的吗（除了乘以任意常数）？一般来说，答案都是否。如果与 \hat{A} 的本征值相对应的独立本征函数不止一个（即简并），那么简并本征值的本征函数的任何线性组合都是 \hat{A} 的本征函数（第 3.6 节）。给出 \hat{B} 本征函数所需的适当线性组合很可能与给出 \hat{C} 本征函数的线性组合不同。事实上，要想拥有三个算符的共同完整的本征函数集，除了满足式 (5.16) 外，还需要满足 $[\hat{B}, \hat{C}] = 0$ 。要有一组完整的函数同时是几个算符的本征函数，每个算符必须与其他每个算符可对易。

5.2 矢量

下一节我们将求解角动量的本征值问题，角动量是一种矢量。因此，我们首先回顾一下矢量。

完全由大小指定的物理属性（如质量、长度、能量）称为**标量** (scalar)。需要同时指定大小和方向的物理属性（如力、速度、动量）称为**矢量** (vector)。矢量由一条有方向的线段表示，其长度和方向表示属性的大小和方向。

两个矢量 \mathbf{A} 和 \mathbf{B} 的和由如下过程定义：滑动第一个矢量，使其终点与第二个矢量的起点重合，同时保持其方向不变。然后从第二个矢量的尾部到第一个矢量的头部绘制一个新的矢量。参见图 5.1。矢量与标量乘积 $c\mathbf{A}$ 定义为：长度为 $|c|$ 的矢量长度乘以 \mathbf{A} 的长度，若 c 为正，则方向与 \mathbf{A} 相同；若 c 为负，则方向与 \mathbf{A} 相反。

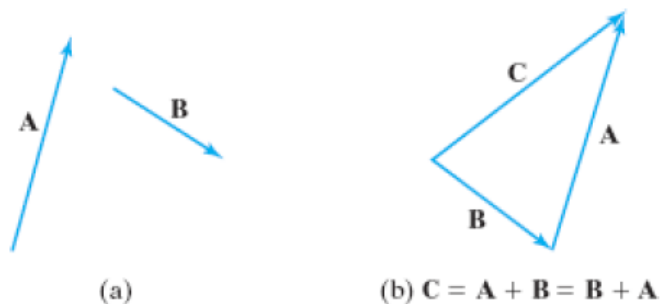


图 5.1: 两个矢量的和

图 5.2: 单位矢量 \mathbf{i} 、 \mathbf{j} 和 \mathbf{k} , \mathbf{A} 的三个分量

为了获得表示矢量的代数（以及几何）方法，我们在空间中建立了笛卡尔坐标。我们在 x 轴正方向画一个单位矢量，称为 \mathbf{i} （与虚数单位没有关系）。 y 轴和 z 轴的单位矢量分别称为 \mathbf{j} 和 \mathbf{k} （图 5.2）。为了用这三个单位矢量表示空间中的任意矢量 \mathbf{A} ，我们首先将 \mathbf{A} 的起点与原点重合，保持其方向不变。然后，我们找到 \mathbf{A} 在 x 轴、 y 轴和 z 轴上的投影，分别为 A_x 、 A_y 和 A_z 。由矢量加和的定义，我们有（图 5.2）：

$$\boxed{\mathbf{A} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}} \quad (5.17)$$

我们可以通过指定 \mathbf{A} 的三个分量（ A_x 、 A_y 和 A_z ）来指定矢量 \mathbf{A} 。因此，三维空间中的矢量可以定义为由三个数组成的有序集合。

当且仅当两个矢量所有对应的分量都相等时： $A_x = B_x$ 、 $A_y = B_y$ 和 $A_z = B_z$ ，我们才说这两个向量 \mathbf{A} 和 \mathbf{B} 相等。因此，一个矢量方程等价于三个标量方程。

为了将两个矢量相加，我们将它们对应的分量相加：

$$\mathbf{A} + \mathbf{B} = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k} + B_x \mathbf{i} + B_y \mathbf{j} + B_z \mathbf{k}$$

$$\boxed{\mathbf{A} + \mathbf{B} = (A_x + B_x) \mathbf{i} + (A_y + B_y) \mathbf{j} + (A_z + B_z) \mathbf{k}} \quad (5.18)$$

若 c 是标量，则

$$\boxed{c\mathbf{A} = cA_x \mathbf{i} + cA_y \mathbf{j} + cA_z \mathbf{k}} \quad (5.19)$$

矢量 \mathbf{A} 的模 (magnitude) 为它的长度, 由 A 或 $|\mathbf{A}|$ 表示。模 A 是一个标量。

两个矢量 $\mathbf{A} \cdot \mathbf{B}$ 的点积 (dot product) 或数量积 (scalar product) 定义为

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}| |\mathbf{B}| \cos \theta = \mathbf{B} \cdot \mathbf{A} \quad (5.20)$$

其中 θ 是 \mathbf{A} 和 \mathbf{B} 之间的夹角。点积是三个标量的乘积, 因此是一个标量。注意 $|\mathbf{A}| \cos \theta$ 是 \mathbf{A} 在 \mathbf{B} 方向上的投影。根据矢量加法的定义, 可以得出矢量 $\mathbf{A} + \mathbf{B}$ 在某个矢量 \mathbf{C} 上的投影是 \mathbf{A} 和 \mathbf{B} 在 \mathbf{C} 上的投影之和。因此,

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = (\mathbf{A} \cdot \mathbf{C}) + (\mathbf{B} \cdot \mathbf{C}) \quad (5.21)$$

由于三个单位矢量 \mathbf{i} 、 \mathbf{j} 和 \mathbf{k} 的长度均为单位长度且互相垂直, 我们有

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = \cos 0 = 1, \quad \mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k} = \mathbf{j} \cdot \mathbf{k} = \cos(\pi/2) = 0 \quad (5.22)$$

使用 (5.22) 和分配律 (5.21), 我们有

$$\mathbf{A} \cdot \mathbf{B} = (A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}) \cdot (B_x \mathbf{i} + B_y \mathbf{j} + B_z \mathbf{k})$$

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \quad (5.23)$$

其中点积的 9 项中有 6 项为零。

由式 (5.20) 有

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2 \quad (5.24)$$

使用式 (5.23), 我们有

$$|\mathbf{A}| = (A_x^2 + A_y^2 + A_z^2)^{1/2} \quad (5.25)$$

对于三维矢量, 还有另一种乘积方式。 $\mathbf{A} \times \mathbf{B}$ 的叉积 (cross product) 或向量积 (vector product) 是一个新的矢量, 其模为

$$\mathbf{A} \times \mathbf{B} = |\mathbf{A}| |\mathbf{B}| \sin \theta \quad (5.26)$$

其线段垂直于 \mathbf{A} 和 \mathbf{B} 所在的平面, 其方向与 \mathbf{A} 和 \mathbf{B} 形成一个右手定则的系统 (与 x 轴、 y 轴和 z 轴的右手定则相同)。见图 5.3。因此, 由定义, 我们有

$$\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B}$$

同样地 (见 *Taylor and Mann*, Section 10.2)

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C} \quad (5.27)$$

对于三个单位向量, 我们有



图 5.3: 两个矢量的叉积

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \sin 0 = 0$$

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{i} = -\mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}, \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}$$

使用这些等式及分配律 (5.27), 我们有

$$\begin{aligned} \mathbf{A} \times \mathbf{B} &= (A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}) \times (B_x \mathbf{i} + B_y \mathbf{j} + B_z \mathbf{k}) \\ \mathbf{A} \times \mathbf{B} &= (A_y B_z - A_z B_y) \mathbf{i} + (A_z B_x - A_x B_z) \mathbf{j} + (A_x B_y - A_y B_x) \mathbf{k} \end{aligned}$$

为了便于记忆, 我们可以将叉积表示为行列式 (见第 8.3 节):

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = \mathbf{i} \begin{vmatrix} A_y & A_z \\ B_y & B_z \end{vmatrix} - \mathbf{j} \begin{vmatrix} A_x & A_z \\ B_x & B_z \end{vmatrix} + \mathbf{k} \begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix} \quad (5.28)$$

我们定义如下的矢量微分 (del) 算符:

$$\nabla \equiv \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \quad (5.29)$$

则由式 (3.23), 线动量的矢量算符为 $\hat{\mathbf{p}} = -i\hbar \nabla$ 。

函数 $g(x, y, z)$ 的梯度 (gradient) 定义为将微分算符作用到函数 g 上的结果:

$$\mathbf{grad} g(x, y, z) \equiv \nabla g(x, y, z) \equiv \mathbf{i} \frac{\partial g}{\partial x} + \mathbf{j} \frac{\partial g}{\partial y} + \mathbf{k} \frac{\partial g}{\partial z} \quad (5.30)$$

一个标量函数的梯度为一个矢量函数。矢量 $\nabla g(x, y, z)$ 代表函数 g 的空间变化率: ∇g 在 x 方向的分量表示函数 g 在 x 方向的变化率, 以此类推。可以证明, 矢量 ∇g 指向 g 变化率最大的方向。由式 (4.24), 势能和力之间的关系为

$$\mathbf{F} = -\nabla V(x, y, z) = -\mathbf{i} \frac{\partial V}{\partial x} - \mathbf{j} \frac{\partial V}{\partial y} - \mathbf{k} \frac{\partial V}{\partial z} \quad (5.31)$$

假设某矢量 \mathbf{A} 的三个分量是某参数 t 的函数, 即 $A_x = A_x(t)$ 、 $A_y = A_y(t)$ 和 $A_z = A_z(t)$ 。则我们将 \mathbf{A} 对 t 的导数 (derivative) 定义为

$$\frac{d\mathbf{A}}{dt} = \mathbf{i} \frac{dA_x}{dt} + \mathbf{j} \frac{dA_y}{dt} + \mathbf{k} \frac{dA_z}{dt} \quad (5.32)$$

矢量符号是表示函数变量的一种便捷方法。双粒子系统的波函数可以写作 $\psi(x_1, y_1, z_1, x_2, y_2, z_2)$, 如果矢量 \mathbf{r}_1 从原点指向粒子 1, 那么 \mathbf{r}_1 就具有 x, y, z 三个方向的分量, 指定 \mathbf{r}_1 等同于指定三个坐标 x, y, z 。 \mathbf{r}_2 也是如此。因此, 我们可以将波函数写成 $\psi(\mathbf{r}_1, \mathbf{r}_2)$ 。矢量符号也可以用在积分中, 例如式 (3.57) 对全空间的积分经常写成 $\int \cdots \int |\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_n)|^2 d\mathbf{r}_1 \cdots d\mathbf{r}_n$ 。

n 维空间的矢量

矢量的定义可以推广到三维以上的空间。三维空间中的矢量 \mathbf{A} 可以由其模 $|\mathbf{A}|$ 和方向来定义; 若在笛卡尔坐标系中, 还可以用其三个分量 (A_x, A_y, A_z) 来定义。因此, 我们可以把三维矢量定义为按特定顺序排列的三个实数 (A_x, A_y, A_z) 的集合。 n 维实数矢量“空间” (有时称为超空间) 中的矢量 \mathbf{B} 定义为 n 个实数的有序集合 (B_1, B_2, \cdots, B_n) , 其中 B_1, B_2, \cdots, B_n 是 \mathbf{B} 的分量 (components)。不要担心无法将 n 维空间中的矢量可视化。

函数的变量通常使用 n 维矢量符号表示。例如, 在双粒子系统中, 与 $\psi(x_1, y_1, z_1, x_2, y_2, z_2)$ 不同, 我们可以定义一个六维矢量 \mathbf{q} , 其分量满足 $q_1 = x_1, q_2 = y_1, q_3 = z_1, q_4 = x_2, q_5 = y_2, q_6 = z_2$,

且将波函数写成 $\psi(\mathbf{q})$ 。对于有 n 个粒子的系统，我们可以定义一个有 $3n$ 个分量的矢量 \mathbf{q} ，将波函数写成 $\psi(\mathbf{q})$ ，使用 $\int |\psi(\mathbf{q})|^2 d\mathbf{q}$ 来进行归一化处理。

寻找分子平衡几何的理论使用 n 维矢量（第 15.10 节）。第 5.2 节的其余部分与第 15.10 节相关，在学习第 15.10 节之前不必阅读。

当且仅当 n 维矢量 \mathbf{B} 和 \mathbf{C} 的所有分量都相等时： $B_1 = C_1, B_2 = C_2, \dots, B_n = C_n$ ，我们才说这两个向量相等（equal）。因此，在 n 维空间中，一个矢量方程等价于 n 个标量方程。两个 n 维矢量 \mathbf{B} 与 \mathbf{D} 和（sum）的定义为矢量 $(B_1 + D_1, B_2 + D_2, \dots, B_n + D_n)$ 。矢量的差定义类似。矢量 $k\mathbf{B}$ 定义为矢量 $(kB_1, kB_2, \dots, kB_n)$ ，其中 k 是标量。在三维空间中，若 $k > 0$ ，则矢量 $k\mathbf{A}$ 的方向一致。在 n 维空间中亦然。与实数对 (A_x, A_y, A_z) 定义三维空间中的一个点类似， (B_1, B_2, \dots, B_n) 定义 n 维空间中的一个点（point）。

n 维实数矢量的长度（length）（或大小（magnitude）或欧几里得范数（Euclidean norm）） $|\mathbf{B}|$ （有时表示为 $\|\mathbf{B}\|$ ）定义为

$$|\mathbf{B}| \equiv (\mathbf{B} \cdot \mathbf{B})^{1/2} = (B_1^2 + B_2^2 + \dots + B_n^2)^{1/2}$$

长度为 1 的矢量称为归一化的（normalized）。

两个 n 维矢量 \mathbf{B} 和 \mathbf{G} 的内积（inner product）或数量积（scalar product）定义为标量

$$\mathbf{B} \cdot \mathbf{G} \equiv B_1 G_1 + B_2 G_2 + \dots + B_n G_n$$

若 $\mathbf{B} \cdot \mathbf{G} = 0$ ，则称 \mathbf{B} 和 \mathbf{G} 正交（orthogonal）。由式 (5.20)，两个 n 维矢量夹角 θ 的余弦值定义为 $\cos \theta \equiv (\mathbf{B} \cdot \mathbf{C}) / |\mathbf{B}| |\mathbf{C}|$ 。我们可以证明，这个定义使得 θ 位于 -1 到 1 的范围内。

在三维空间内，单位矢量 $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$ 互相垂直。因此，任意矢量都可以表示为这三个矢量的线性组合 [式 (5.17)]。在 n 维空间中，单位矢量 $\mathbf{e}_1 \equiv (1, 0, \dots, 0)$, $\mathbf{e}_2 \equiv (0, 1, \dots, 0)$, \dots , $\mathbf{e}_n \equiv (0, 0, \dots, 1)$ 互相正交。由于任意 n 维矢量 \mathbf{B} 等于 $B_1 \mathbf{e}_1 + B_2 \mathbf{e}_2 + \dots + B_n \mathbf{e}_n$ ，任意 n 维矢量都可以表示为这 n 个单位矢量的线性组合。因此，这 n 个矢量可以说是 n 维实数矢量空间的基（basis）。由于矢量 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 是正交且归一的，它们是这个实矢量空间的正交基（orthonormal basis）。点积 $\mathbf{B} \cdot \mathbf{e}_i$ 给出了矢量 \mathbf{B} 在 \mathbf{e}_i 方向上的分量。矢量空间有许多可能的基集。任何由 n 个线性无关的实矢量组成的集合都可以作为 n 维实矢量空间的基。

一个三维矢量可以用它的三个分量来指定，也可以用它的长度和方向来指定。可以通过给出矢量与 x 、 y 和 z 轴正半轴所成的三个角度来指定方向。这些角度称为矢量的方向角（direction angles），其值在 0 到 π 之间。但是，一旦给出了其他两个方向角， z 轴的方向角就会固定下来，因此只有两个方向角是独立的。因此，一个三维矢量可以用长度和两个方向角来表示。同样，在 n 维空间中，一个矢量和每个单位矢量 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 之间的方向角也可以根据上述两个矢量之间夹角的余弦公式求出。因此，一个 n 维矢量可以用它的长度和 $n - 1$ 个方向角来指定。

三变量函数的梯度由 (5.30) 定义。 n 变量函数 $f(q_1, q_2, \dots, q_n)$ 的梯度（gradient）定义为 n 维矢量，其分量分别为 f 对每个变量的一阶偏导数：

$$\nabla f \equiv (\partial f / \partial q_1) \mathbf{e}_1 + (\partial f / \partial q_2) \mathbf{e}_2 + \dots + (\partial f / \partial q_n) \mathbf{e}_n$$

我们考虑的是实数 n 维矢量空间。狄拉克（Dirac）对量子力学的表述使用的是复数、无穷维的矢量空间，对它的讨论从略。

5.3 单粒子系统的角动量

在第 3.3 节中，我们找到了线动量算符 \hat{p}_x 的本征函数和对应的本征值。本节中我们对粒子的角动量进行类似的分析。角动量在原子结构的量子力学中扮演重要的角色。我们从回顾经典力学的角

动量开始。

单粒子的经典力学角动量

考虑一个质量为 m 的运动粒子。我们建立一个固定在空间的直角坐标系。设 \mathbf{r} 为从原点指向粒子瞬时位置的矢量，我们有

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad (5.33)$$

其中 x 、 y 和 z 是粒子在给定瞬间的坐标。这些坐标是时间的函数。将速度矢量 \mathbf{v} 定义为粒子位置矢量 \mathbf{r} 对时间的导数，我们有 [式 (5.32)]

$$\mathbf{v} \equiv \frac{d\mathbf{r}}{dt} = \mathbf{i} \frac{dx}{dt} + \mathbf{j} \frac{dy}{dt} + \mathbf{k} \frac{dz}{dt} \quad (5.34)$$

$$v_x = dx/dt, \quad v_y = dy/dt, \quad v_z = dz/dt$$

我们定义粒子的**线动量** (linear momentum) 的矢量 \mathbf{p} 为

$$\boxed{\mathbf{p} \equiv m\mathbf{v}} \quad (5.35)$$

$$p_x = mv_x, \quad p_y = mv_y, \quad p_z = mv_z \quad (5.36)$$

粒子相对于坐标原点的**角动量** (angular momentum) \mathbf{L} 在经典力学中定义为

$$\boxed{\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}} \quad (5.37)$$

$$\mathbf{L} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} \quad (5.38)$$

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \quad (5.39)$$

其中，我们用到了式 (5.28)。 L_x 、 L_y 和 L_z 分别是角动量矢量 \mathbf{L} 在 x 、 y 和 z 方向上的分量。角动量向量的方向垂直于由粒子的位置矢量 \mathbf{r} 和速度矢量 \mathbf{v} 所确定的平面 (图5.4)。



图 5.4: $\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$

作用在粒子上的**力矩** (torque) $\boldsymbol{\tau}$ 定义为力 \mathbf{F} 和位置矢量 \mathbf{r} 的叉积: $\boldsymbol{\tau} \equiv \mathbf{r} \times \mathbf{F}$ 。可以证明: $\boldsymbol{\tau} = d\mathbf{L}/dt$ 。当没有力矩作用于粒子时, 其角动量的变化率为零; 也就是说, 其角动量是恒定的 (或守恒的)。对于绕太阳运行的行星来说, 引力是沿径向分布的。由于两个平行矢量的叉积为零, 因此行星上没有力矩, 其角动量是守恒的。

单粒子轨道角动量算符

现在, 让我们来看看量子力学的处理方法。量子力学中有两种角动量。**轨道角动量** (orbital angular momentum) 是粒子在空间运动的结果, 与经典力学量 \mathbf{L} 类似。**自旋角动量** (spin angular momentum) (第10章) 是许多微观粒子的固有属性, 与经典力学没有类似之处。我们现在只考虑轨道角动量。将经典方程 (5.39) 中的坐标和角动量替换为相应的算符 [式 (3.21)-(3.23)], 我们就得到了粒子轨道角动量分量的量子力学算符。有

$$\hat{L}_x = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \quad (5.40)$$

$$\hat{L}_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \quad (5.41)$$

$$\hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \quad (5.42)$$

(由于 $\hat{y}\hat{p}_z = \hat{p}_z\hat{y}$ 等, 我们在构造这些算符时不会遇到不可对易的问题)。使用

$$\hat{L}^2 = |\hat{\mathbf{L}}|^2 = \hat{\mathbf{L}} \cdot \hat{\mathbf{L}} = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \quad (5.43)$$

我们可以根据 (5.40)-(5.42) 中的算符构建角动量大小平方的算符。

由于对易关系决定了哪些物理量可以同时被赋予定值, 我们对角动量的对易关系进行了研究。将算符 \hat{L}_y 作用到函数 $f(x, y, z)$ 上, 我们有

$$\hat{L}_y f(x, y, z) = -i\hbar \left(z \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial z} \right)$$

将 \hat{L}_x 作用到上一个方程上, 我们有

$$\hat{L}_x \hat{L}_y f = -\hbar^2 \left(y \frac{\partial f}{\partial x} + yz \frac{\partial^2 f}{\partial z \partial x} - yx \frac{\partial^2 f}{\partial z^2} - z^2 \frac{\partial^2 f}{\partial y \partial x} + zx \frac{\partial^2 f}{\partial y \partial z} \right) \quad (5.44)$$

同样地,

$$\hat{L}_x f = -i\hbar \left(y \frac{\partial f}{\partial z} - z \frac{\partial f}{\partial y} \right)$$

将 \hat{L}_y 作用到上一个方程上, 我们有

$$\hat{L}_y \hat{L}_x f = -\hbar^2 \left(zy \frac{\partial^2 f}{\partial x \partial z} - z^2 \frac{\partial^2 f}{\partial x \partial y} - xy \frac{\partial^2 f}{\partial z^2} + x \frac{\partial f}{\partial y} + xz \frac{\partial^2 f}{\partial z \partial y} \right) \quad (5.45)$$

(5.44) 减去 (5.45), 我们得到

$$\begin{aligned} \hat{L}_x \hat{L}_y f - \hat{L}_y \hat{L}_x f &= -\hbar^2 \left(y \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial y} \right) \\ [\hat{L}_x, \hat{L}_y] &= i\hbar \hat{L}_z \end{aligned} \quad (5.46)$$

其中我们用到了类似以下的关系:

$$\boxed{\frac{\partial^2 f}{\partial z \partial x} = \frac{\partial^2 f}{\partial x \partial z}} \quad (5.47)$$

对于品优函数是正确的。我们可以用相同的过程求出 $[\hat{L}_y, \hat{L}_z]$ 和 $[\hat{L}_z, \hat{L}_x]$ 。但我们可以通过注意 (5.40)-(5.42) 中的某种对称性来节省时间。所谓 x 、 y 和 z 的**循环排列** (cyclic permutation) 是指用 y 替换 x , 用 z 替换 y , 用 x 替换 z 。如果我们在 \hat{L}_x 中进行循环排列, 就得到了 \hat{L}_y , 在 \hat{L}_y 中



图 5.5: 球面坐标

进行循环排列就得到了 \hat{L}_z ，在 \hat{L}_z 中进行循环排列就得到了 \hat{L}_x 。因此，对 (5.46) 连续进行两次循环置换，就可以得到

$$[\hat{L}_y, \hat{L}_z] = i\hbar \hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar \hat{L}_y \quad (5.48)$$

接下来，我们利用第5.1节中对易子的性质，求 \hat{L}^2 和它对应分量的对易子。

$$\begin{aligned} [\hat{L}^2, \hat{L}_x] &= [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x] \\ &= [\hat{L}_x^2, \hat{L}_x] + [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\ &= [\hat{L}_y^2, \hat{L}_x] + [\hat{L}_z^2, \hat{L}_x] \\ &= \hat{L}_y [\hat{L}_y, \hat{L}_x] + [\hat{L}_y, \hat{L}_x] \hat{L}_y + [\hat{L}_z, \hat{L}_x] \hat{L}_z + \hat{L}_z [\hat{L}_z, \hat{L}_x] \\ &= -i\hbar \hat{L}_z \hat{L}_y - i\hbar \hat{L}_y \hat{L}_z + i\hbar \hat{L}_y \hat{L}_z + i\hbar \hat{L}_z \hat{L}_y \\ &= 0 \end{aligned} \quad (5.49)$$

对 x 、 y 和 z 进行循环替换后， $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ 没有变化，如果对式 (5.49) 进行循环替换，我们就可以得到

$$[\hat{L}^2, \hat{L}_y] = 0, \quad [\hat{L}^2, \hat{L}_z] = 0 \quad (5.50)$$

我们可以给 L, L_x, L_y, L_z 中的哪些量赋定值？由于 \hat{L}^2 与其所有分量对应算符均可对易，我们可以为 L^2 和其任意一个分量指定一个精确的值。然而， $\hat{\mathbf{L}}$ 的任意两个分量之间都不存在对易关系，因此我们不能同时确定多个分量（这个说法有一个例外，稍后讨论）。传统的做法是将 L_z 作为角动量的分量，与 L^2 一起指定。注意：在指定 $L^2 = |\mathbf{L}|^2$ 的值时，我们并没有指定矢量 \mathbf{L} 的方向。我们只指定了其大小。完整指定 \mathbf{L} 需要同时指定其三个分量，这是我们做不到的。在经典力学中，当角动量守恒时，其三个分量均具有确定的值。而在量子力学中，当角动量守恒时，只有其大小和其中一个分量是确定的。

现在，使用这些算符在空间直角坐标系中的形式，我们尝试求出 \hat{L}^2 和 \hat{L}_z 的本征函数和本征值。然而，我们会发现得到的偏微分方程无法分离变量。因此，我们将这些算符转换为球面坐标 (spherical coordinate)。坐标 r 表示点 (x, y, z) 到原点的距离。而角度 θ 是矢量 \mathbf{r} 与 z 轴正半轴之间的夹角， ϕ 是矢量 \mathbf{r} 在 xOy 平面上的投影与 x 轴正半轴之间的夹角（数学书通常将 θ 与 ϕ 互换）。使用三角法计算，有

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta \quad (5.51)$$

$$r^2 = x^2 + y^2 + z^2, \quad \cos \theta = \frac{z}{(x^2 + y^2 + z^2)^{1/2}}, \quad \tan \phi = \frac{y}{x} \quad (5.52)$$

为了将角动量算符转换为球坐标, 我们需要对 $\partial/\partial x$ 、 $\partial/\partial y$ 和 $\partial/\partial z$ 进行球坐标变换 [如果需要, 可以略读这一变换。在 (5.64) 之后重新开始阅读]。

要进行这种变换, 我们需要使用链式法则 (chain rule)。假设我们有一个关于 r 、 θ 和 ϕ 的函数 $f(r, \theta, \phi)$ 。如果我们改变自变量, 将

$$r = r(x, y, z), \quad \theta = \theta(x, y, z), \quad \phi = \phi(x, y, z)$$

带入 f , 我们就将其变换为 x 、 y 和 z 的函数

$$f[r(x, y, z), \theta(x, y, z), \phi(x, y, z)] = g(x, y, z)$$

例如, 假设 $f(r, \theta, \phi) = 3r \cos \theta + 2 \tan^2 \phi$ 。使用式 (5.52), 我们有 $g(x, y, z) = 3z + 2y^2/x^2$ 。

链式法则告诉我们 $g(x, y, z)$ 的偏导数是如何与 $f(r, \theta, \phi)$ 的偏导数联系起来的。事实上,

$$\left(\frac{\partial g}{\partial x}\right)_{y,z} = \left(\frac{\partial f}{\partial r}\right)_{\theta,\phi} \left(\frac{\partial r}{\partial x}\right)_{y,z} + \left(\frac{\partial f}{\partial \theta}\right)_{r,\phi} \left(\frac{\partial \theta}{\partial x}\right)_{y,z} + \left(\frac{\partial f}{\partial \phi}\right)_{r,\theta} \left(\frac{\partial \phi}{\partial x}\right)_{y,z} \quad (5.53)$$

$$\left(\frac{\partial g}{\partial y}\right)_{x,z} = \left(\frac{\partial f}{\partial r}\right)_{\theta,\phi} \left(\frac{\partial r}{\partial y}\right)_{x,z} + \left(\frac{\partial f}{\partial \theta}\right)_{r,\phi} \left(\frac{\partial \theta}{\partial y}\right)_{x,z} + \left(\frac{\partial f}{\partial \phi}\right)_{r,\theta} \left(\frac{\partial \phi}{\partial y}\right)_{x,z} \quad (5.54)$$

$$\left(\frac{\partial g}{\partial z}\right)_{x,y} = \left(\frac{\partial f}{\partial r}\right)_{\theta,\phi} \left(\frac{\partial r}{\partial z}\right)_{x,y} + \left(\frac{\partial f}{\partial \theta}\right)_{r,\phi} \left(\frac{\partial \theta}{\partial z}\right)_{x,y} + \left(\frac{\partial f}{\partial \phi}\right)_{r,\theta} \left(\frac{\partial \phi}{\partial z}\right)_{x,y} \quad (5.55)$$

为了将这些方程转换为算符方程, 我们删去 f 和 g , 得出

$$\frac{\partial}{\partial x} = \left(\frac{\partial r}{\partial x}\right)_{y,z} \frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial x}\right)_{y,z} \frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial x}\right)_{y,z} \frac{\partial}{\partial \phi} \quad (5.56)$$

$\partial/\partial y$ 和 $\partial/\partial z$ 的方程类似。现在的任务是计算类似于 $(\partial r/\partial x)_{y,z}$ 的偏导数。将 (5.52) 中第一个方程在 y 和 z 恒定的条件下对 x 求偏导数, 我们有

$$\begin{aligned} 2r \left(\frac{\partial r}{\partial x}\right)_{y,z} &= 2x = 2r \sin \theta \cos \phi \\ \left(\frac{\partial r}{\partial x}\right)_{y,z} &= \sin \theta \cos \phi \end{aligned} \quad (5.57)$$

将式 $r^2 = x^2 + y^2 + z^2$ 对 y 和 z 求偏导数, 我们有

$$\left(\frac{\partial r}{\partial y}\right)_{x,z} = \sin \theta \cos \phi, \quad \left(\frac{\partial r}{\partial z}\right)_{x,y} = \cos \theta \quad (5.58)$$

由 (5.52) 中第二个方程, 我们有

$$\begin{aligned} -\sin \theta \left(\frac{\partial \theta}{\partial x}\right)_{y,z} &= -\frac{xz}{r^3} \\ \left(\frac{\partial \theta}{\partial x}\right)_{y,z} &= \frac{\cos \theta \cos \phi}{r} \end{aligned} \quad (5.59)$$

同样地,

$$\left(\frac{\partial \theta}{\partial y}\right)_{x,z} = \frac{\cos \theta \sin \phi}{r}, \quad \left(\frac{\partial \theta}{\partial z}\right)_{x,y} = -\frac{\sin \theta}{r} \quad (5.60)$$

由 $\tan \phi = y/x$, 我们有

$$\left(\frac{\partial \phi}{\partial x}\right)_{y,z} = -\frac{\sin \phi}{r \sin \theta}, \quad \left(\frac{\partial \phi}{\partial y}\right)_{x,z} = \frac{\cos \phi}{r \sin \theta}, \quad \left(\frac{\partial \phi}{\partial z}\right)_{x,y} = 0 \quad (5.61)$$

将 (5.57)、(5.59) 和 (5.61) 代入 (5.56), 我们得到

$$\frac{\partial}{\partial x} = \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \quad (5.62)$$

类似地, 我们可以得到

$$\frac{\partial}{\partial y} = \sin \theta \sin \phi \frac{\partial}{\partial r} + \frac{\cos \theta \sin \phi}{r} \frac{\partial}{\partial \theta} + \frac{\cos \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \quad (5.63)$$

$$\frac{\partial}{\partial z} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \quad (5.64)$$

最后，我们可以用球面坐标来表示角动量分量了。将 (5.51)、(5.63) 和 (5.64) 带入 (5.40)，得到

$$\begin{aligned}\hat{L}_x = -i\hbar \left[r \sin \theta \sin \phi \left(\cos \theta \frac{\partial}{\partial r} \right) - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right. \\ \left. - r \cos \theta \left(\sin \theta \sin \phi \frac{\partial}{\partial r} + \frac{\cos \theta \sin \phi}{r} \frac{\partial}{\partial \phi} + \frac{\cos \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \right) \right] \\ \hat{L}_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} - \cos \theta \cos \phi \frac{\partial}{\partial \phi} \right)\end{aligned}\quad (5.65)$$

同样，我们有

$$\hat{L}_y = -i\hbar \left(\cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \quad (5.66)$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \quad (5.67)$$

对 \hat{L}_x 、 \hat{L}_y 和 \hat{L}_z 进行平方并相加，我们可以构造出 $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ [式 (5.43)]。结果为（问题 5.17）

$$\hat{L}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \quad (5.68)$$

虽然角动量算符取决于所有三个直角坐标 x 、 y 和 z ，但它们只涉及两个球面坐标 θ 和 ϕ 。

单粒子轨道角动量的本征函数和本征值

我们现在希望找到 \hat{L}^2 和 \hat{L}_z 的共同本征函数，记作 Y 。由于这些算符只涉及 θ 和 ϕ ，那么 Y 就是这两个变量的函数： $Y = Y(\theta, \phi)$ （当然，由于这些算符都是线性算符，我们可以用任意一个关于 r 的函数乘以 Y 而不改变 Y 是 \hat{L}^2 和 \hat{L}_z 的本征函数的事实）。我们必须解以下方程：

$$\hat{L}_z Y(\theta, \phi) = bY(\theta, \phi) \quad (5.69)$$

$$\hat{L}^2 Y(\theta, \phi) = cY(\theta, \phi) \quad (5.70)$$

其中 b 和 c 分别是 \hat{L}_z 和 \hat{L}^2 的本征值。

使用算符 \hat{L}_z ，我们有

$$-i\hbar \frac{\partial}{\partial \phi} Y(\theta, \phi) = bY(\theta, \phi) \quad (5.71)$$

由于 (5.71) 中的算符不涉及 θ ，我们可以尝试分离变量。将 Y 写成

$$Y(\theta, \phi) = S(\theta)T(\phi) \quad (5.72)$$

则方程 (5.71) 变为

$$\begin{aligned}-i\hbar \frac{\partial}{\partial \phi} [S(\theta)T(\phi)] &= bS(\theta)T(\phi) \\ -i\hbar S(\theta) \frac{dT(\phi)}{d\phi} &= bS(\theta)T(\phi) \\ \frac{dT(\phi)}{T(\phi)} &= \frac{ib}{\hbar} d\phi \\ T(\phi) &= Ae^{ib\phi/\hbar}\end{aligned}\quad (5.73)$$

其中 A 是任意常数。

T 适合作为本征函数吗。答案是否定的, 因为它一般不是单值函数。如果我们将 ϕ 增加 2π , 我们仍将处于空间中的同一点, 因此我们希望这样做时 T 保持不变。为了使 T 成为单值函数, 我们需要做如下限制:

$$\begin{aligned} T(\phi + 2\pi) &= T(\phi) \\ Ae^{ib\phi/\hbar} e^{ib2\pi/\hbar} &= Ae^{ib\phi/\hbar} \\ e^{ib2\pi/\hbar} &= 1 \end{aligned} \quad (5.74)$$

为了满足 $e^{i\alpha} = \cos \alpha + i \sin \alpha = 1$, 我们需要 $\alpha = 2\pi m$, 其中

$$m = 0, \pm 1, \pm 2, \pm \dots$$

因此, 由式 (5.74) 我们得到

$$\begin{aligned} 2\pi b/\hbar &= 2\pi m \\ b &= m\hbar, \quad m = \dots, -2, -1, 0, 1, 2, \dots \end{aligned} \quad (5.75)$$

式 (5.73) 变为

$$T(\phi) = Ae^{im\phi}, \quad m = 0, \pm 1, \pm 2, \dots \quad (5.76)$$

角动量在 z 方向分量的本征值是量子化的。

通过归一化 T , 我们可以确定常数 A 。首先, 让我们考虑对关于 r 、 θ 和 ϕ 的函数 F 进行归一化。三个独立变量的取值范围是 (见图 5.5)

$$0 \leq r \leq \infty, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi \quad (5.77)$$

球坐标系下的体积微元为 (*Taylor and Mann*, 第 13.9 节)

$$d\tau = r^2 \sin \theta \, dr \, d\theta \, d\phi \quad (5.78)$$

(5.78) 是一个无限小区域空间的体积, 其球面坐标位于 r 到 $r + dr$ 、 θ 到 $\theta + d\theta$ 和 ϕ 到 $\phi + d\phi$ 之间。因此, 球坐标下 F 的归一化条件为

$$\int_0^\infty \left[\int_0^\pi \left[\int_0^{2\pi} |F(r, \theta, \phi)|^2 d\phi \right] \sin \theta d\theta \right] r^2 dr = 1 \quad (5.79)$$

如果 F 恰好可以写成

$$F(r, \theta, \phi) = R(r) S(\theta) T(\phi)$$

则对 (5.79) 使用积分 (3.74), 有

$$\int_0^\infty |R(r)|^2 r^2 dr \int_0^\pi |S(\theta)|^2 \sin \theta d\theta \int_0^{2\pi} |T(\phi)|^2 d\phi = 1$$

方便的做法是将 F 的每个因子都作归一化处理:

$$\int_0^\infty |R(r)|^2 r^2 dr = 1, \quad \int_0^\pi |S(\theta)|^2 \sin \theta d\theta = 1, \quad \int_0^{2\pi} |T(\phi)|^2 d\phi = 1 \quad (5.80)$$

因此,

$$\begin{aligned} \int_0^{2\pi} (Ae^{im\phi})^* (Ae^{im\phi}) d\phi &= 1 = |A|^2 \int_0^{2\pi} d\phi \\ |A| &= (2\pi)^{-1/2} \end{aligned}$$

$$T(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}, \quad m = 0, \pm 1, \pm 2, \dots \quad (5.81)$$

我们现在尝试使用式 (5.70) 来求解 \hat{L}^2 的本征值 c 。对 \hat{L}^2 使用式 (5.68)、对 Y 使用式 (5.72)，由式 (5.81)，我们有

$$\begin{aligned} -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \left(S(\theta) \frac{1}{\sqrt{2\pi}} e^{im\phi} \right) &= c S(\theta) \frac{1}{\sqrt{2\pi}} e^{im\phi} \\ \frac{d^2 S}{d\theta^2} + \cot \theta \frac{dS}{d\theta} - \frac{m^2}{\sin^2 \theta} S &= -\frac{c}{\hbar^2} S \end{aligned} \quad (5.82)$$

为了解方程 (5.82)，我们需要进行一些繁琐的操作，如果需要，可以略过不读。在式 (5.91) 处重新开始阅读。首先，为了方便起见，我们将自变量代替为

$$w = \cos \theta \quad (5.83)$$

将 S 变换为一个新的关于 w 的函数：

$$S(\theta) = G(w) \quad (5.84)$$

由链式法则，我们有

$$\frac{dS}{d\theta} = \frac{dG}{dw} \frac{dw}{d\theta} = -\sin \theta \frac{dG}{dw} = -(1-w^2)^{1/2} \frac{dG}{dw} \quad (5.85)$$

同样地，我们得出（问题 5.25）

$$\frac{d^2 S}{d\theta^2} = (1-w^2) \frac{d^2 G}{dw^2} - w \frac{dG}{dw} \quad (5.86)$$

由式 (5.86)、(5.85) 和 $\cot \theta = \cos \theta / \sin \theta = w / (1-w^2)^{1/2}$ ，我们可以将 (5.82) 重写为

$$(1-w^2) \frac{d^2 G}{dw^2} - 2w \frac{dG}{dw} + \left[\frac{c}{\hbar^2} - \frac{m^2}{1-w^2} \right] G(w) = 0 \quad (5.87)$$

w 的取值范围是 $-1 \leq w \leq 1$ 。

在尝试幂级数求解时，为了得到两项的递推关系，我们需要对因变量作如下改变：

$$G(w) = (1-w^2)^{|m|/2} H(w) \quad (5.88)$$

对 (5.88) 求导，我们得到了 G' 和 G'' ，在除以 $(1-w^2)^{|m|/2}$ 后，我们可以将 (5.87) 变为

$$(1-w^2) H'' - 2(|m|+1)wH' + [c\hbar^{-2} - |m|(|m|+1)]H = 0 \quad (5.89)$$

现在，我们尝试 H 的幂级数解

$$H(w) = \sum_{n=0}^{\infty} a_n w^n \quad (5.90)$$

对其求导 [与式 (4.36)-(4.38) 类似]，我们有

$$H'(w) = \sum_{j=0}^{\infty} j a_j w^{j-1}$$

$$H''(w) = \sum_{j=0}^{\infty} j(j-1) a_j w^{j-2} = \sum_{j=0}^{\infty} (j+2)(j+1) a_{j+2} w^j$$

将这些幂级数带入式 (5.89)，合并求和后，有

$$\sum_{j=0}^{\infty} \left[(j+2)(j+1) a_{j+2} + \left(-j^2 - j - 2|m|j + \frac{c}{\hbar^2} - |m|^2 - |m| \right) a_j \right] w^j = 0$$

令 w^j 的系数为零，我们得到

$$a_{j+2} = \frac{(j+|m|)(j+|m|+1) - c/\hbar^2}{(j+1)(j+2)} a_j \quad (5.91)$$

与谐振子的例子一样, 式 (5.89) 的通解也是偶次幂级数 (系数由 a_0 决定) 和奇次幂级数 (系数由 a_1 决定) 的任意线性组合。可以证明, 由式 (5.91) 定义的无穷级数不是品优函数。[许多文献都指出: 该级数在 $w = \pm 1$ 处发散。然而, 这并不足以否定该无穷级数, 因为本征函数是平方可积的, 即使某两点上是无穷的。仔细讨论见 M. Whippman, Am. Phys., 34, 656 (1966)。] 因此, 与谐振子的情况一致, 我们必须让级数在某一项 $a_k w^k$ 处打断。根据 k 是奇数还是偶数, 我们设 a_0 或 a_1 为零, 从而消除其他级数。

令式 (5.91) 中 a_k 的系数为零, 得到

$$c = \hbar^2 (k + |m|) (k + |m| + 1), \quad k = 0, 1, 2, \dots \quad (5.92)$$

由于 $|m|$ 的值为 $0, 1, 2, \dots$, $k + |m|$ 的值为 $0, 1, 2, \dots$ 。因此, 我们可以定义量子数 l 为

$$l = k + |m| \quad (5.93)$$

角动量大小平方的本征值为

$$c = l(l+1)\hbar^2, \quad l = 0, 1, 2, \dots \quad (5.94)$$

单粒子轨道角动量的大小为

$$|\mathbf{L}| = [l(l+1)]^{1/2} \hbar \quad (5.95)$$

由式 (5.93) 有 $|m| = l$ 。因此, m 的取值范围为

$$m = -l, -l+1, -l+2, \dots, -1, 0, 1, \dots, l-2, l-1, l \quad (5.96)$$

现在让我们检查角动量的本征函数。由式 (5.83)、(5.84)、(5.88)、(5.90) 和 (5.93), 本征函数中的 θ 因子为

$$S_{l,m}(\theta) = \sin^{|m|} \theta \sum_{\substack{j=1,3,\dots \\ j=0,2,\dots}}^{l-|m|} a_j \cos^j \theta \quad (5.97)$$

其中, $l - |m|$ 的奇偶性决定求和从 $j = 1$ 还是 $j = 0$ 开始。使用式 (5.94), a_j 的系数满足的递推关系 (5.91) 可以写成

$$a_{j+2} = \frac{(j+|m|)(j+|m|+1) - l(l+1)}{(j+1)(j+2)} a_j \quad (5.98)$$

\hat{L}^2 和 \hat{L}_z 的本征函数由式 (5.72) 和式 (5.81) 给出:

$$Y_l^m(\theta, \phi) = S_{l,m}(\theta) T(\phi) = \frac{1}{\sqrt{2\pi}} S_{l,m}(\theta) e^{im\phi} \quad (5.99)$$

Y_l^m 中的 m 是一个标签, 而不是指数。

例题: 分别求当 (a) $l = 0$; (b) $l = 1$ 时 $Y_l^m(\theta, \phi)$ 、 \hat{L}^2 和 \hat{L}_z 的本征值。

(a) 对于 $l = 0$, 由式 (5.96), 有 $m = 0$ 。则式 (5.97) 变为

$$S_{0,0}(\theta) = a_0 \quad (5.100)$$

由归一化条件 (5.80), 我们有

$$\int_0^\pi |a_0|^2 \sin \theta d\theta = 1 = 2|a_0|^2$$

$$|a_0| = 2^{-1/2}$$

则式 (5.99) 变为

$$Y_0^0(\theta, \phi) = \frac{1}{\sqrt{4\pi}} \quad (5.101)$$

[显然, 式 (5.101) 是 \hat{L}^2 、 \hat{L}_x 、 \hat{L}_y 和 \hat{L}_z 的本征函数, 方程 (5.65)-(5.68)。] 对于 $l=0$, 本征函数不存在角度依赖关系, 我们说本征函数是球对称的 (spherically symmetric)。

对于 $l=0$ 和 $m=0$, 由式 (5.69)、(5.70)、(5.75) 和 (5.94), 我们有 \hat{L}^2 的本征值为 $c=0$, \hat{L}_z 的本征值为 $b=0$ 。

(b) 对于 $l=1$, 由式 (5.96), 有 $m=-1, 0, 1$ 。当 $|m|=1$ 时, 式 (5.97) 变为

$$S_{1,\pm 1}(\theta) = a_0 \sin \theta \quad (5.102)$$

(5.102) 中的 a_0 和 (5.100) 中的 a_0 不一定相等。由归一化条件, 有

$$1 = |a_0^2| \int_0^\pi \sin^2 \theta \sin \theta d\theta = |a_0^2| \int_{-1}^1 (1-w^2) dw$$

$$|a_0| = \sqrt{3}/2$$

其中, 我们将 w 用 $\cos \theta$ 代替。因此, 我们有 $S_{1,\pm 1}(3^{1/2}/2) \sin \theta$, 式 (5.99) 变为

$$Y_1^1 = (3/8\pi)^{1/2} \sin \theta e^{i\phi}, \quad Y_1^{-1} = (3/8\pi)^{1/2} \sin \theta e^{-i\phi} \quad (5.103)$$

当 $l=1$ 而 $m=0$ 时, 我们得出 (见下文的练习题) $S_{1,0} = (3/2)^{1/2} \cos \theta$, $Y_1^0 = (3/4\pi)^{1/2} \cos \theta$ 。因此, 当 $l=1$ 时, 由式 (5.94) 可得 \hat{L}^2 的本征值为 $2\hbar^2$; 由式 (5.75), \hat{L}_z 的本征值为 \hbar 、0 和 $-\hbar$, 分别与 $m=1, 0, -1$ 对应。

练习: 验证 $S_{1,0}$ 和 Y_1^0 的表达式。

函数 $S_{l,m}(\theta)$ 在数学中是众所周知的, 是乘以一个归一化系数后的连带勒让德函数 (associated Legendre function)。连带勒让德函数的定义见问题 5.34。表 (5.1) 列出了 $l \leq 3$ 的 $S_{l,m}(\theta)$ 表达式。

表 5.1: $S_{l,m}(\theta)$

$l=0$:	$S_{0,0} = \frac{1}{2}\sqrt{2}$
$l=1$:	$S_{1,0} = \frac{1}{2}\sqrt{6} \cos \theta$ $S_{1,\pm 1} = \frac{1}{2}\sqrt{3} \sin \theta$
$l=2$:	$S_{2,0} = \frac{1}{4}\sqrt{10} (3 \cos^2 \theta - 1)$ $S_{2,\pm 1} = \frac{1}{2}\sqrt{15} \sin \theta \cos \theta$ $S_{2,\pm 2} = \frac{1}{4}\sqrt{15} \sin^2 \theta$
$l=3$:	$S_{3,0} = \frac{3}{4}\sqrt{14} (\frac{5}{3} \cos^3 \theta - \cos \theta)$ $S_{3,\pm 1} = \frac{1}{8}\sqrt{42} \sin \theta (5 \cos^2 \theta - 1)$ $S_{3,\pm 2} = \frac{1}{4}\sqrt{105} \sin^2 \theta \cos \theta$ $S_{3,\pm 3} = \frac{1}{8}\sqrt{70} \sin^3 \theta$

$$\hat{L}^2 Y_l^m(\theta, \phi) = l(l+1) \hbar^2 Y_l^m(\theta, \phi), \quad l=0, 1, 2, \dots \quad (5.104)$$

$$\hat{L}_z Y_l^m(\theta, \phi) = m\hbar Y_l^m(\theta, \phi), \quad m = -l, -l+1, \dots, l-1, l \quad (5.105)$$

其中，本征函数由式 (5.99) 给出。我们通常使用符号 m_l 代替 m 来表示 L_z 量子数。我们稍后会看到球谐函数是正交函数（方程 7.27）。

由于 $l \geq |m|$ ，因此除了 $l = 0$ 以外，轨道角动量 \mathbf{L} 的大小 $[l(l+1)]^{1/2} \hbar^2$ 大于其 z 方向的分量 $|m|\hbar$ 。如果可以让角动量的大小等于其 z 分量，这就意味着 x 和 y 分量为零，我们就可以指定 \mathbf{L} 的所有三个分量了。然而这是做不到的，因为角动量三个分量的算符不可对易。只有一种例外，就是 l 为零时。在该情况中， $|\mathbf{L}|^2 = L_x^2 + L_y^2 + L_z^2 = 0$ ，因此所有三个分量都为零。由式 (5.12)，角动量分量的不确定性满足

$$\Delta L_x \Delta L_y \geq \left| \frac{1}{2} \Psi^* [\hat{L}_x, \hat{L}_y] \Psi \right| = \frac{\hbar}{2} \left| \int \Psi^* \hat{L}_z \Psi d\tau \right| \quad (5.106)$$

通过循环排列能得出其余两个相似的方程。当 $\hat{L}_x \Psi = 0$ ， $\hat{L}_y \Psi = 0$ ， $\hat{L}_z \Psi = 0$ ，式 (5.106) 的右侧和另两个相似的方程也为零，因此 $\Delta L_x = \Delta L_y = \Delta L_z = 0$ 被允许。然而，根据第 5.1 节，要同时得到两个算符的本征函数，这两个算符必须可对易，这又是怎么回事呢？答案是，这个定理指的是一个算符的整套本征函数是另一个算符的本征函数的可能性。因此，尽管 \hat{L}_x 和 \hat{L}_z 不可对易，但仍有可能使得 \hat{L}_x 的本征函数同时也是 \hat{L}_z 的本征函数（ $l = 0 = m$ 时）。然而，使得 \hat{L}_x 的所有本征函数都是 \hat{L}_z 的本征函数是不可能的。



图 5.6: \mathbf{L} 的方向

由于我们无法指定 L_x 和 L_y ，所以矢量 \mathbf{L} 可以位于以 z 轴为轴，高度为 $m\hbar$ 的圆锥面上的任何位置。当 $l = 1$ 时， \mathbf{L} 相对于 z 轴的可能方向如图 5.7 所示。对于每一个 \hat{L}^2 的本征值，都有 $2l+1$ 个不同的本征函数 Y_l^m ，与 m 的 $2l+1$ 个取值相对应。我们说 \hat{L}^2 的本征值是 $(2l+1)$ 重简并的，简并（degeneracy）一词适用于任何算符的本征值，而不仅仅是哈密顿算符。



图 5.7: $l = 1$ 时 \mathbf{L} 相对于 z 轴的方向

当然, z 轴并没有什么特别之处。空间中的所有方向都是等价的。如果我们选择确定 L^2 和 L_x (而不是 L_z), 我们会得到与 L_z 相同的 L_x 的本征值。然而, 解 \hat{L}_z 的本征方程更为简单, 因为在球坐标系中, \hat{L}_z 的形式更简单, 其中涉及到围绕 z 轴的旋转角度 ϕ 。

5.4 角动量阶梯算符法

通过将轨道角动量的算符表示为微分算符, 并求解由此产生的微分方程, 我们得到了 \hat{L}^2 和 \hat{L}_z 的本征值。我们现在证明: 只需要利用算符之间的对易关系就能找到这些本征值。本节的工作适用于任何满足角动量对易关系的算符。它尤其适用于自旋角动量 (第 10 章) 和轨道角动量。

我们使用字母 L 来表示轨道角动量。在这里, 我们将使用字母 M 来表示我们正在处理的任何一种角动量。我们三个线性算符 \hat{M}_x 、 \hat{M}_y 和 \hat{M}_z , 我们只知道它们符合对易关系 [与式 (5.46) 和 (5.48) 相似]

$$[\hat{M}_x, \hat{M}_y] = i\hbar\hat{M}_z, \quad [\hat{M}_y, \hat{M}_z] = i\hbar\hat{M}_x, \quad [\hat{M}_z, \hat{M}_x] = i\hbar\hat{M}_y \quad (5.107)$$

我们定义算符 \hat{M}^2 为

$$\hat{M}^2 = \hat{M}_x^2 + \hat{M}_y^2 + \hat{M}_z^2 \quad (5.108)$$

我们的问题是如何找到 \hat{M}^2 和 \hat{M}_z 的本征值。

使用式 (5.107) 和 (5.108), 我们从对 \hat{M}^2 及其分量的对易子进行评估开始。与推导 (5.49) 和 (5.50) 时的推导类似, 我们有

$$[\hat{M}^2, \hat{M}_x] = [\hat{M}^2, \hat{M}_y] = [\hat{M}^2, \hat{M}_z] = 0 \quad (5.109)$$

因此, 我们可以得到 \hat{M}^2 和 \hat{M}_z 的同时本征函数。

现在我们定义两个新算符, **升算符** (raising operator) \hat{M}_+ 和 **降算符** (lowering operator) \hat{M}_- , 它们的定义为

$$\hat{M}_+ \equiv \hat{M}_x + i\hat{M}_y \quad (5.110)$$

$$\hat{M}_- \equiv \hat{M}_x - i\hat{M}_y \quad (5.111)$$

这些是**阶梯算符** (ladder operator) 的例子。使用这些术语的原因很快就会清楚。我们有

$$\begin{aligned} \hat{M}_+ \hat{M}_- &= (\hat{M}_x + i\hat{M}_y)(\hat{M}_x - i\hat{M}_y) = \hat{M}_x(\hat{M}_x - i\hat{M}_y) + i\hat{M}_y(\hat{M}_x - i\hat{M}_y) \\ &= \hat{M}_x^2 - i\hat{M}_x\hat{M}_y + i\hat{M}_y\hat{M}_x + \hat{M}_y^2 = \hat{M}^2 - \hat{M}_z^2 + i[\hat{M}_y, \hat{M}_x] \\ \hat{M}_+ \hat{M}_- &= \hat{M}^2 - \hat{M}_z^2 + \hbar\hat{M}_z \end{aligned} \quad (5.112)$$

同样地, 我们有

$$\hat{M}_- \hat{M}_+ = \hat{M}^2 - \hat{M}_z^2 - \hbar\hat{M}_z \quad (5.113)$$

对于这些算符与 \hat{M}_z 的对易子, 我们有

$$\begin{aligned} [\hat{M}_+, \hat{M}_z] &= [\hat{M}_x + i\hat{M}_y, \hat{M}_z] = [\hat{M}_x, \hat{M}_z] + i[\hat{M}_y, \hat{M}_z] = -i\hbar\hat{M}_y - \hbar\hat{M}_x \\ [\hat{M}_+, \hat{M}_z] &= -\hbar\hat{M}_+ \\ \hat{M}_+ \hat{M}_z &= \hat{M}_z \hat{M}_+ - \hbar\hat{M}_+ \end{aligned} \quad (5.114)$$

其中我们用到了式 (5.107)。同样地, 我们有

$$\hat{M}_- \hat{M}_z = \hat{M}_z \hat{M}_- + \hbar\hat{M}_- \quad (5.115)$$

设 Y 是 \hat{M}^2 和 \hat{M}_z 的共同本征函数, 即

$$\hat{M}^2 Y = cY \quad (5.116)$$

$$\hat{M}_z Y = bY \quad (5.117)$$

其中 c 和 b 是本征值。将算符 \hat{M}_+ 作用于式 (5.117), 我们有

$$\hat{M}_+ \hat{M}_z Y = \hat{M}_+ bY$$

使用式 (5.114), 及 \hat{M}_+ 是线性算符, 我们有

$$\begin{aligned} (\hat{M}_z \hat{M}_+ - \hbar \hat{M}_+) Y &= b \hat{M}_+ Y \\ \hat{M}_z (\hat{M}_+ Y) &= (b + \hbar) (\hat{M}_+ Y) \end{aligned} \quad (5.118)$$

最后一个方程说明: $\hat{M}_+ Y$ 是算符 \hat{M}_z 的本征函数, 其本征值为 $b + \hbar$ 。换句话说, 将升算符 \hat{M}_+ 作用到本征函数 Y 上后, 可以将 Y 转换成另一个 \hat{M}_z 的本征函数, 其本征值比 Y 的大 \hbar 。如果我们现在对式 (5.118) 使用升算符, 并再次使用式 (5.114), 我们得到

$$\hat{M}_z (\hat{M}_+^2 Y) = (b + 2\hbar) (\hat{M}_+^2 Y)$$

重复应用升算符, 有

$$\hat{M}_z (\hat{M}_+^k Y) = (b + k\hbar) (\hat{M}_+^k Y), \quad k = 0, 1, 2, \dots \quad (5.119)$$

如果我们对式 (5.117) 使用降算符 \hat{M}_- 并应用式 (5.115), 我们已同样的方式发现:

$$\hat{M}_z (\hat{M}_- Y) = (b - \hbar) (\hat{M}_- Y) \quad (5.120)$$

$$\hat{M}_z (\hat{M}_-^k Y) = (b - k\hbar) (\hat{M}_-^k Y) \quad (5.121)$$

这样, 通过对本征值为 b 的本征函数使用升降算符, 我们得到了一个本征值阶梯, 阶梯之间的差值为 \hbar :

$$\dots, \quad b - 2\hbar, \quad b - \hbar, \quad b, \quad b + \hbar, \quad b + 2\hbar, \quad \dots$$

函数 $\hat{M}_\pm^k Y$ 是 \hat{M}_z 的本征函数, 其本征值为 $b \pm k\hbar$ [式 (5.119) 和 (5.121)]。我们现在表明这些函数也是 \hat{M}^2 的本征函数, 其本征值为 c :

$$\hat{M}_z \hat{M}_\pm^k Y = (b \pm k\hbar) \hat{M}_\pm^k Y \quad (5.122)$$

$$\hat{M}^2 \hat{M}_\pm^k Y = c \hat{M}_\pm^k Y \quad (5.123)$$

为了证明 (5.123), 我们首先表明 \hat{M}^2 与 \hat{M}_+ 和 \hat{M}_- 对易。我们有

$$[\hat{M}^2, \hat{M}_\pm] = [\hat{M}^2, \hat{M}_x \pm i\hat{M}_y] = [\hat{M}^2, \hat{M}_x] \pm i[\hat{M}^2, \hat{M}_y] = 0 \pm 0 = 0$$

也有

$$[\hat{M}^2, \hat{M}_\pm^2] = [\hat{M}^2, \hat{M}_\pm] \hat{M}_\pm + \hat{M}_\pm [\hat{M}^2, \hat{M}_\pm] = 0 + 0 = 0$$

通过归纳法可以得出

$$[\hat{M}^2, \hat{M}_\pm^k] = 0, \quad \text{或} \quad \hat{M}^2 \hat{M}_\pm^k = \hat{M}_\pm^k \hat{M}^2, \quad k = 0, 1, 2, \dots \quad (5.124)$$

如果我们对式 (5.116) 使用升降算符 \hat{M}_\pm , 并使用 (5.124), 我们有

$$\hat{M}_\pm^k \hat{M}^2 Y = \hat{M}_\pm^k cY$$

$$\hat{M}^2 (\hat{M}_{\pm}^k Y) = c (\hat{M}_{\pm}^k Y) \quad (5.125)$$

这正是我们要证明的。

接下来我们想证明：使用阶梯算符生成的 \hat{M}_z 的本征值集合是有界的。对于特定的，本征值为 b 的算符 \hat{M}_z 的本征函数 Y ，我们有

$$\hat{M}_z Y = bY$$

若考虑由阶梯算符生成的本征函数与本征值的集合，我们有

$$\hat{M}_z Y_k = b_k Y_k \quad (5.126)$$

其中

$$Y_k = \hat{M}_{\pm}^k Y \quad (5.127)$$

$$b_k = b \pm k\hbar \quad (5.128)$$

(\hat{M}_+ 和 \hat{M}_- 的使用破坏了 Y 的归一性，所以 Y_k 不是归一化的。若想找到归一化常数，见问题 10.27。)

对式 (5.126) 使用 \hat{M}_z ，我们有

$$\begin{aligned} \hat{M}_z^2 Y_k &= b_k \hat{M}_z Y_k \\ \hat{M}_z^2 Y_k &= b_k^2 Y_k \end{aligned} \quad (5.129)$$

现在用式 (5.123) 减去式 (5.129)，使用式 (5.127) 和 (5.108)，有

$$\begin{aligned} \hat{M}^2 Y_k - \hat{M}_z^2 Y_k &= c Y_k - b_k^2 Y_k \\ (\hat{M}_x^2 + \hat{M}_y^2) Y_k &= (c - b_k^2) Y_k \end{aligned} \quad (5.130)$$

算符 $\hat{M}_x^2 + \hat{M}_y^2$ 对应于一个非负的物理量，因此其本征值应该是非负的（在问题 7.11 中得到证明）。因此，式 (5.130) 表明 $c - b_k^2 \geq 0$ 即 $c^{1/2} \geq |b_k|$ 。因此，

$$c^{1/2} \geq b_k \geq c^{1/2}, \quad k = 0, \pm 1, \pm 2, \dots \quad (5.131)$$

由于 k 变化时 c 保持不变，因此式 (5.131) 表明本征值 b 同时具有上界和下界。令 b_{\max} 和 b_{\min} 分别为 b_k 的最大值和最小值， Y_{\max} 和 Y_{\min} 分别为对应的本征函数：

$$\hat{M}_z Y_{\max} = b_{\max} Y_{\max} \quad (5.132)$$

$$\hat{M}_z Y_{\min} = b_{\min} Y_{\min} \quad (5.133)$$

现在，将升算符作用到式 (5.132) 上，并使用式 (5.114)，我们有

$$\begin{aligned} \hat{M}_z \hat{M}_+ Y_{\max} &= b_{\max} \hat{M}_+ Y_{\max} \\ \hat{M}_z (\hat{M}_+ Y_{\max}) &= (b_{\max} + \hbar) (\hat{M}_+ Y_{\max}) \end{aligned} \quad (5.134)$$

由于最后一个方程声明 $\hat{M}_+ Y_{\max}$ 是 \hat{M}_z 的本征函数，其本征值为 $b_{\max} + \hbar$ ，因此似乎与 b_{\max} 是 b_k 的最大值矛盾。解决这一矛盾的唯一办法就是让 $\hat{M}_+ Y_{\max} = 0$ 消失（基于物理原因，我们总是拒绝将零作为本征函数）。因此，

$$\hat{M}_+ Y_{\max} = 0 \quad (5.135)$$

将算符作用到式 (5.133) 上，并使用式 (5.113)、(5.132) 和 (5.116)，我们有

$$0 = \hat{M}_- \hat{M}_+ Y_{\max} = (\hat{M}^2 \hat{M}_z^2 - \hbar \hat{M}_z) Y_{\max} = (c - b_{\max}^2 + \hbar b_{\max}) Y_{\max}$$

$$\begin{aligned}
c - b_{\max}^2 - \hbar b_{\max} &= 0 \\
c &= b_{\max}^2 + \hbar b_{\max}
\end{aligned} \tag{5.136}$$

类似的论证表明

$$\hat{M}_- Y_{\min} = 0 \tag{5.137}$$

通过将升算符作用于这个方程并使用式 (5.112)，我们有

$$c = b_{\min}^2 - \hbar b_{\min}$$

从式 (5.136) 中减去最后一个等式，得出

$$b_{\max}^2 + \hbar^2 b_{\max} + (\hbar^2 b_{\min} - b_{\min}^2) = 0$$

这是一个关于 b_{\max} 的二次方程，使用通常的公式（在量子力学中仍然有效），有

$$b_{\max} = -b_{\min}, \quad b_{\max} = b_{\min} - \hbar$$

由于第二个根表明 b_{\max} 小于 b_{\min} ，应舍去，所以

$$b_{\min} = -b_{\max} \tag{5.138}$$

除此以外，式 (5.128) 表明 b_{\max} 和 b_{\min} 之差是 \hbar 的整数倍：

$$b_{\max} - b_{\min} = \hbar n, \quad n = 0, 1, 2, \dots \tag{5.139}$$

将 (5.138) 代入 (5.139)，我们得到了 \hat{M}_z 的本征值：

$$\begin{aligned}
b_{\max} &= \frac{1}{2} n \hbar \\
b_{\max} &= j \hbar \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \\
b_{\min} &= -j \hbar
\end{aligned} \tag{5.140}$$

$$b = -j \hbar, (-j+1) \hbar, (-j+2) \hbar, \dots, (j-2) \hbar, (j-1) \hbar, j \hbar \tag{5.141}$$

由式 (5.136)，我们找到了 \hat{M}^2 的本征值：

$$c = j(j+1) \hbar^2, \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \tag{5.142}$$

因此，

$$\hat{M}^2 Y = j(j+1) \hbar^2 Y, \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \tag{5.143}$$

$$\hat{M}_z Y = m_j \hbar Y, \quad m_j = -j, -j+1, \dots, j-1, j \tag{5.144}$$

通过使用算符间的对易关系，我们已经找到了 \hat{M}^2 和 \hat{M}_z 的本征值。然而，将式 (5.143) 和 (5.144) 与式 (5.104) 和 (5.105) 进行比较，我们发现：角量子数除了可以取整数以外，还可以取半整数值。这或许表明：除了轨道角动量以外，可能还有另一种角动量。在第10章中，我们将看到自旋角动量可以具有半整数和整数量子数。对于轨道角动量， $T(\phi)$ 本征函数的单值定解条件 [见 (5.73) 之后的等式] 消除了角动量量子数的半整数值。[并不是每个人都接受单值性作为波函数的有效边界条件，而且还有许多其他拒绝半整轨道-角动量量子数的理由；见 C. G. Gray, Am. J. Phys., 37, 559 (1969); M. L. Whippman, Am. J. Phys., 34, 656 (1966)。]

阶梯算符法也可用于求解其他本征值问题，见问题 5.36。

总结

习题

Chapter 6

氢原子

6.1 单粒子中心力问题

在研究氢原子之前，我们先来看看单个粒子在中心力作用下运动这一更为普遍的问题。本节的结果适用于任何中心力问题。例如氢原子（第 6.5 节）和各向同性的三维谐振子（第 6.3 节）。

中心力（central force）是由球面对称的势能函数导出的，这意味着中心力只是粒子与原点距离的函数 $V = V(r)$ 。力和势能的关系由式 (5.31) 给出：

$$\mathbf{F} = -\nabla V(x, y, z) = -\mathbf{i}(\partial V/\partial x) - \mathbf{j}(\partial V/\partial y) - \mathbf{k}(\partial V/\partial z) \quad (6.1)$$

(6.1) 中的偏导数可以用链式法则求得 [式 (5.53)- (5.55)]。在这个例子中，由于 V 只是 r 的函数，所以我们有 $(\partial V/\partial \theta)_{r,\phi} = 0$ 和 $(\partial V/\partial \phi)_{r,\theta} = 0$ 。因此，

$$\left(\frac{\partial V}{\partial x}\right)_{y,z} = \frac{dV}{dr} \left(\frac{\partial r}{\partial x}\right)_{y,z} = \frac{x}{r} \frac{dV}{dr} \quad (6.2)$$

$$\left(\frac{\partial V}{\partial y}\right)_{x,z} = \frac{y}{r} \frac{dV}{dr}, \quad \left(\frac{\partial V}{\partial z}\right)_{x,y} = \frac{z}{r} \frac{dV}{dr} \quad (6.3)$$

其中我们用到了式 (5.57) 和 (5.58)。则方程 (6.1) 可以写成

$$\mathbf{F} = -\frac{1}{r} \frac{dV}{dr} (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = -\frac{dV(r)}{dr} \frac{\mathbf{r}}{r} \quad (6.4)$$

其中 \mathbf{r} 的计算使用了式 (5.33)。式 (6.4) 中的 \mathbf{r}/r 是径向上的单位矢量。因此中心力是径向的。

现在，让我们来看看受中心力作用的单粒子量子力学。其哈密顿算符为

$$\hat{H} = \hat{T} + \hat{V} = -(\hbar^2/2m) \nabla^2 + V(r) \quad (6.5)$$

其中 $\nabla^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ [式 (3.46)]。由于 V 是球对称的，我们将在球坐标系中进行计算。因此，我们希望将拉普拉斯算符转换为球坐标的形式。我们已经有了 $\partial/\partial x$ 、 $\partial/\partial y$ 和 $\partial/\partial z$ 的表达式 [式 (5.62)-(5.65)]，通过将这些算符平方并相加，我们就得到了拉普拉斯算符。该计算留作课后习题。结果为（问题 6.4）：

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2} \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \quad (6.6)$$

回过头来看式 (5.68)，其中给出了单粒子轨道角动量大小平方的算符 \hat{L}^2 ，我们可以看到

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{1}{r^2 \hbar^2} \hat{L}^2 \quad (6.7)$$

哈密顿算符 (6.5) 变为

$$\hat{H} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \hat{L}^2 + V(r) \quad (6.8)$$

在经典力学中, 受中心力作用的粒子其角动量是守恒的 (第5.3节)。在量子力学中, 我们可能会问, 我们是否可以拥有能量和角动量都有确定值的状态? 为了使 \hat{H} 的本征函数组也是 \hat{L}^2 的本征函数, 对易子 $[\hat{H}, \hat{L}^2]$ 必须为零。我们有

$$\begin{aligned} [\hat{H}, \hat{L}^2] &= [\hat{T}, \hat{L}^2] + [\hat{V}, \hat{L}^2] \\ [\hat{T}, \hat{L}^2] &= \left[-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \hat{L}^2, \hat{L}^2 \right] \\ [\hat{T}, \hat{L}^2] &= -\frac{\hbar^2}{2m} \left[\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r}, \hat{L}^2 \right] + \frac{1}{2m} \left[\frac{1}{r^2} \hat{L}^2, \hat{L}^2 \right] \end{aligned} \quad (6.9)$$

回忆 \hat{L}^2 只是 θ 和 ϕ 的函数, 而不是 r 的函数 [式 (5.68)]。因此它与任何包含 r 的算符可对易。[为了证明这一结论, 我们必须使用 (5.47) 类似的关系, 但将 x 和 z 分别换成 r 和 θ 。] 因此, 式 (6.9) 中第一项的对易子为零。此外, 由于任何算符都与它自己对易, 式 (6.9) 第二项的对易子也为零。因此, $[\hat{T}, \hat{L}^2] = 0$ 。同样地, 由于 \hat{L}^2 不包含 r , V 只是 r 的函数, 我们有 $[\hat{V}, \hat{L}^2] = 0$ 。因此,

$$[\hat{H}, \hat{L}^2] = 0, \quad \text{若 } V = V(r) \quad (6.10)$$

当势能函数与 θ 和 ϕ 都无关时, \hat{H} 和 \hat{L}^2 可对易。

现在, 我们来考虑算符 $\hat{L}_z = -i\hbar\partial/\partial\phi$ [式 (5.67)]。由于 \hat{L}_z 不包含 r , 因此它与 \hat{L}^2 可对易 [式 (5.50)], 因此 \hat{H} (6.8) 和 \hat{L}_z 也可对易。我们有

$$[\hat{H}, \hat{L}_z] = 0, \quad \text{若 } V = V(r) \quad (6.11)$$

因此, 我们现在有了中心力问题中, 一组同时是 \hat{H} 、 \hat{L}^2 和 \hat{L}_z 的本征函数的函数。令 ψ 表示这些共同本征函数:

$$\boxed{\hat{H}\psi = E\psi} \quad (6.12)$$

$$\boxed{\hat{L}^2\psi = l(l+1)\hbar^2\psi, \quad l = 0, 1, 2, \dots} \quad (6.13)$$

$$\boxed{\hat{L}_z\psi = m\hbar\psi, \quad m = -l, -l+1, \dots, l-1, l} \quad (6.14)$$

其中, 我们用到了式 (5.104) 和式 (5.105)。

根据式 (6.8) 和 (6.13), 对薛定谔方程 (6.12), 有

$$\begin{aligned} -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \hat{L}^2\psi + V(r)\psi &= E\psi \\ -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{l(l+1)\hbar^2}{2mr^2}\psi + V(r)\psi &= E\psi \end{aligned} \quad (6.15)$$

\hat{L}^2 的本征函数是球谐函数 $Y_l^m(\theta, \phi)$, 由于 \hat{L}^2 不包含 r , 我们可以将任意一个关于 r 的函数与球谐函数相乘, 新函数依然是 \hat{L}^2 和 \hat{L}_z 的本征函数。因此,

$$\boxed{\psi = R(r)Y_l^m(\theta, \phi)} \quad (6.16)$$

将 (6.16) 代入 (6.15), 然后两边同除以 Y_l^m , 得到关于未知函数 $R(r)$ 的微分方程:

$$-\frac{\hbar^2}{2m} \left(R'' + \frac{2}{r} R' \right) + \frac{l(l+1)\hbar^2}{2mr^2} R + V(r)R = ER \quad (6.17)$$

我们已经证明: 对于任意势能函数 $V(r)$ 是球对称的单粒子问题, 定态波函数为 $\psi = R(r)Y_l^m(\theta, \phi)$, 其中 $R(r)$ 满足微分方程 (6.17)。通过指定式 (6.17) 中势能函数 $V(r)$ 的形式, 我们可以针对某个特定问题解决它。

6.2 非相互作用粒子和变量分离

到目前为止，我们只解决了单粒子量子力学问题。氢原子是一个双粒子系统，作为处理氢原子的第一步，我们首先考虑一个更简单的情况，即两个不相相互作用的粒子。

假设一个系统由两个无相互作用的粒子 1 和 2 组成。令 q_1 和 q_2 分别表示粒子 1 的 (x_1, y_1, z_1) 坐标和粒子 2 的 (x_2, y_2, z_2) 坐标。由于粒子之间没有相互作用力，系统的经典机械能就是两个粒子的能量之和： $E = E_1 + E_2 = T_1 + V_1 + T_2 + V_2$ ，经典哈密顿量是两个粒子哈密顿量之和： $H = H_1 + H_2$ 。因此，系统的哈密顿算符为

$$\hat{H} = \hat{H}_1 + \hat{H}_2$$

其中 \hat{H}_1 只和坐标 q_1 有关，动量算符 \hat{p}_1 对应于 q_1 。系统的薛定谔方程为

$$(\hat{H}_1 + \hat{H}_2) \psi(q_1, q_2) = E \psi(q_1, q_2) \quad (6.18)$$

我们尝试用分离变量法来求解 (6.18)，令

$$\psi(q_1, q_2) = G_1(q_1) G_2(q_2) \quad (6.19)$$

我们有

$$\hat{H}_1 G_1(q_1) G_2(q_2) + \hat{H}_2 G_1(q_1) G_2(q_2) = E G_1(q_1) G_2(q_2) \quad (6.20)$$

由于 \hat{H}_1 只包含粒子 1 的位置和动量算符，我们有 $\hat{H}_1 [G_1(q_1) G_2(q_2)] = G_2(q_2) \hat{H}_1 G_1(q_1)$ ，因此，就 \hat{H}_1 而言， $G_2(q_2)$ 是常数。使用这个方程以及类似的关于 \hat{H}_2 的方程，我们发现 (6.20) 变为

$$G_2(q_2) \hat{H}_1 G_1(q_1) + G_1(q_1) \hat{H}_2 G_2(q_2) = E G_1(q_1) G_2(q_2) \quad (6.21)$$

$$\frac{\hat{H}_1 G_1(q_1)}{G_1(q_1)} + \frac{\hat{H}_2 G_2(q_2)}{G_2(q_2)} = E \quad (6.22)$$

现在，通过与式 (3.65) 相同的论证，我们得出结论：在 (6.22) 中，左边的每项必须是常数。令 E_1 和 E_2 分别表示这些常数，我们有

$$\frac{\hat{H}_1 G_1(q_1)}{G_1(q_1)} = E_1, \quad \frac{\hat{H}_2 G_2(q_2)}{G_2(q_2)} = E_2$$

$$E = E_1 + E_2 \quad (6.23)$$

因此，当系统由两个无相互作用的粒子组成时，通过求解下式，我们可以将双粒子问题简化为两个单粒子问题：

$$\hat{H}_1 G_1(q_1) = E_1 G_1(q_1), \quad \hat{H}_2 G_2(q_2) = E_2 G_2(q_2) \quad (6.24)$$

它们是每个粒子的独立薛定谔方程。

将该结果推广到 n 个无相互作用的粒子，我们有

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \dots + \hat{H}_n$$

$$\psi(q_1, q_2, \dots, q_n) = G_1(q_1) G_2(q_2) \dots G_n(q_n) \quad (6.25)$$

$$E = E_1 + E_2 + \dots + E_n \quad (6.26)$$

$$\hat{H}_i G_i = E_i G_i, \quad i = 1, 2, \dots, n \quad (6.27)$$

对于一个无相互作用粒子系统，能量是每个粒子各自能量的总和，波函数是每个粒子波函数的乘积。利用哈密顿算符 \hat{H}_i ，求解粒子 i 的薛定谔方程，即可求得粒子 i 的波函数 G_i 。

这些结果也适用于单个粒子，其哈密顿算符是每个坐标的独立项之和：

$$\hat{H} = \hat{H}_x(\hat{x}, \hat{p}_x) + \hat{H}_y(\hat{y}, \hat{p}_y) + \hat{H}_z(\hat{z}, \hat{p}_z)$$

在这种情况下，我们得出结论：波函数和能量分别为

$$\psi(x, y, z) = F(x)G(y)K(z), E = E_x + E_y + E_z$$

$$\hat{H}_x F(x) = E_x F(x), \quad \hat{H}_y G(y) = E_y G(y), \quad \hat{H}_z K(z) = E_z K(z)$$

例如三维盒子中的质点（第 3.5 节）、三维自由质点（问题 3.42）和三维谐振子（问题 4.20）。

6.3 将双粒子问题简化为两个单粒子问题

氢原子包含两个粒子——质子和电子。对于一个双粒子系统（粒子 1 的坐标为 (x_1, y_1, z_1) ，粒子 2 的坐标为 (x_2, y_2, z_2) ），粒子间相互作用的势能通常只是粒子相对坐标 $x_2 - x_1$ 、 $y_2 - y_1$ 和 $z_2 - z_1$ 的函数。在这种情况下，双粒子问题可以简化为两个单粒子问题，正如我们现在所要证明的。

先考虑经典力学处理两个相互作用粒子的情况，令其质量分别为 m_1 和 m_2 。我们用从直角坐标系原点出发的半径矢量 \mathbf{r}_1 和 \mathbf{r}_2 来指定它们的位置（图 6.1）。粒子 1 和 2 的坐标分别为 (x_1, y_1, z_1) 和 (x_2, y_2, z_2) 。我们从粒子 1 到粒子 2 画出矢量 $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ ，并用 x 、 y 和 z 表示其分量：

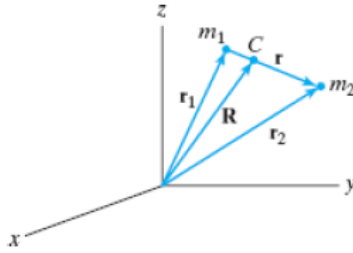


图 6.1: 质心位于 C 的双粒子系统

$$\boxed{x = x_2 - x_1, \quad y = y_2 - y_1, \quad z = z_2 - z_1} \quad (6.28)$$

x 、 y 和 z 称为相对坐标或内坐标（relative or internal coordinates）。

我们现在从原点指向系统的质心 C 画出矢量 \mathbf{R} ，其分量用 X 、 Y 和 Z 表示：

$$\mathbf{R} = \mathbf{i}X + \mathbf{j}Y + \mathbf{k}Z \quad (6.29)$$

根据双粒子系统质心的定义，有

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}, \quad Y = \frac{m_1 y_1 + m_2 y_2}{m_1 + m_2}, \quad Z = \frac{m_1 z_1 + m_2 z_2}{m_1 + m_2} \quad (6.30)$$

这三个方程与矢量方程等价：

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} \quad (6.31)$$

我们也有

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \quad (6.32)$$

我们将式 (6.31) 和 (6.32) 联立为关于 \mathbf{r}_1 和 \mathbf{r}_2 的方程组，并求解得到

$$\mathbf{r}_1 = \mathbf{R} - \frac{m_2}{m_1 + m_2} \mathbf{r}, \quad \mathbf{r}_2 = \mathbf{R} + \frac{m_1}{m_1 + m_2} \mathbf{r} \quad (6.33)$$

方程 (6.31) 和 (6.32) 代表了一种将 x_1, y_1, z_1 和 x_2, y_2, z_2 坐标转换为 X, Y, Z 和 x, y, z 坐标的变换。考虑一下在这种变化下，系统的哈密顿量会发生什么变化。用字母顶上的点表示对时间的导数。粒子 1 的速度为 [式 (5.34)] $\mathbf{v}_1 = d\mathbf{r}_1/dt = \dot{\mathbf{r}}_1$ 。系统的动能为两个粒子的动能之和：

$$T = \frac{1}{2}m_1 |\dot{\mathbf{r}}_1|^2 + \frac{1}{2}m_2 |\dot{\mathbf{r}}_2|^2 \quad (6.34)$$

将 (6.33) 的时间导数代入 (6.34)，我们得到

$$T = \frac{1}{2}m_1 \left(\dot{\mathbf{R}} - \frac{m_2}{m_1 + m_2} \dot{\mathbf{r}} \right) \cdot \left(\dot{\mathbf{R}} - \frac{m_2}{m_1 + m_2} \dot{\mathbf{r}} \right) + \frac{1}{2}m_2 \left(\dot{\mathbf{R}} + \frac{m_1}{m_1 + m_2} \dot{\mathbf{r}} \right) \cdot \left(\dot{\mathbf{R}} + \frac{m_1}{m_1 + m_2} \dot{\mathbf{r}} \right)$$

其中，我们用到了 $|\mathbf{A}|^2 = \mathbf{A} \cdot \mathbf{A}$ [式 (5.24)]。使用矢量数量积的分配律，化简后，我们得到

$$T = \frac{1}{2}(m_1 + m_2) |\dot{\mathbf{R}}|^2 + \frac{1}{2} \left(\frac{m_1 m_2}{m_1 + m_2} \right) |\dot{\mathbf{r}}|^2 \quad (6.35)$$

令 M 为系统的总质量：

$$M \equiv m_1 + m_2 \quad (6.36)$$

我们定义双粒子系统的折合质量 (reduced mass) 为

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2} \quad (6.37)$$

那么

$$T = \frac{1}{2}M |\dot{\mathbf{R}}|^2 + \frac{1}{2}\mu |\dot{\mathbf{r}}|^2 \quad (6.38)$$

(6.38) 的第一项是质量为 M 的整个系统平动产生的动能。平动 (translational motion) 是指每个质点都经历相同位移的运动。 $\frac{1}{2}M |\dot{\mathbf{R}}|^2$ 是位于质心，具有质量 M 的假想粒子的动能。(6.38) 的第二项是两个粒子内运动 (相对运动) 产生的动能。内运动有两种形式。两个粒子间的距离 r 可以变化 (振动)，以及矢量 \mathbf{r} 的方向可以发生变化 (转动)。注意： $|\dot{\mathbf{r}}| = |d\mathbf{r}/dt| \neq d|\mathbf{r}|/dt$ 。

与六个原始坐标 $x_1, y_1, z_1, x_2, y_2, z_2$ 相对应，我们有六个线动量：

$$p_{x_1} = m_1 \dot{x}_1, \quad \dots, \quad p_{z_2} = m_2 \dot{z}_2 \quad (6.39)$$

将 (6.34) 和 (6.38) 相比较，我们将新坐标 X, Y, Z, x, y, z 的动量定义为

$$p_X \equiv M \dot{X}, \quad p_Y \equiv M \dot{Y}, \quad p_Z \equiv M \dot{Z} \\ p_x \equiv \mu \dot{x}, \quad p_y \equiv \mu \dot{y}, \quad p_z \equiv \mu \dot{z}$$

我们定义两个新动量矢量：

$$\mathbf{p}_M = \mathbf{i}M\dot{X} + \mathbf{j}M\dot{Y} + \mathbf{k}M\dot{Z}, \quad \text{和} \quad \mathbf{p}_\mu = \mathbf{i}\mu\dot{x} + \mathbf{j}\mu\dot{y} + \mathbf{k}\mu\dot{z}$$

将这些动量代入式 (6.38)，我们得到

$$T = \frac{|\mathbf{p}_M|^2}{2M} + \frac{|\mathbf{p}_\mu|^2}{2\mu} \quad (6.40)$$

现在，我们来考虑势能。我们限制 V 只是两个粒子的相对坐标 x, y 和 z 的函数：

$$V = V(x, y, z) \quad (6.41)$$

(6.41) 的一个例子是相互作用遵守库仑定律的两个带电粒子 [见式 (3.53)]。根据对 V 的这一限制, 哈密顿函数为

$$H = \frac{p_M^2}{2M} + \left[\frac{p_\mu^2}{2\mu} + V(x, y, z) \right] \quad (6.42)$$

现在, 假设我们有一个由质量为 M 的粒子和质量为 μ 的粒子组成的系统, 前者不受力, 后者受势能函数 $V(x, y, z)$ 的作用。再假设这些粒子之间没有相互作用。如果 (X, Y, Z) 表示质量为 M 的粒子的坐标, 而 (x, y, z) 表示质量为 μ 的粒子的坐标, 那么这个假设系统的哈密顿量是什么? 显然, 它与式 (6.42) 完全相同。

哈密顿量 (6.42) 可以看作是两个无相互作用假想粒子的哈密顿量之和: 质量为 M 的粒子其哈密顿量为 $p_M^2/2M$, 质量为 μ 的粒子其哈密顿量为 $p_\mu^2/2\mu + V(x, y, z)$ 。因此, 第6.2的结果表明: 系统的量子力学能量是两个假想粒子的能量之和 [式 (6.23)]: $E = E_M + E_\mu$ 。由式 (6.24) 和 (6.42) 可知, 平动能 E_M 可以通过求解薛定谔方程 $(\hat{p}_M^2/2M)\psi_M = E_M\psi_M$ 得到。这是质量为 M 的自由粒子的薛定谔方程, 因此其可能的本征值全是非负数 [式 (2.31)]: $E_m \geq 0$ 。由式 (6.42) 和 (6.24) 可知, 内运动能 E_μ 可以通过求解以下薛定谔方程得到:

$$\left[\frac{\hat{p}_\mu^2}{2\mu} + V(x, y, z) \right] \psi_\mu(x, y, z) = E_\mu \psi_\mu(x, y, z) \quad (6.43)$$

这样, 我们就把两个粒子根据仅取决于相对坐标 x, y, z 的势能函数 $V(x, y, z)$ 进行相互作用的问题分离成了两个独立的单粒子问题: (1) 质量为 M 的整体系统平动, 它只是在系统能量的基础上增加了一个恒非负的能量 E_M ; (2) 通过求解质量为 μ 的假想粒子的薛定谔方程 (6.43) 来处理相对运动或内部运动, 该粒子的坐标为相对坐标 x, y, z , 并在势能 $V(x, y, z)$ 的作用下运动。

例如, 对于包含一个电子 (e) 和一个质子 (p) 的氢原子, 原子的总能量为 $E = E_M + E_\mu$, 其中的 E_M 是质量为 $M = m_e + m_p$ 的整个原子在空间中的平动能, 其中的 E_μ 是通过将 $\mu = m_e m_p / (m_e + m_p)$ 代入 (6.43) 得到的, V 是电子和质子相互作用的库仑定律势能; 见第 6.5 节。

6.4 双粒子的刚性转子模型

在求解氢原子的薛定谔方程之前, 我们先来处理双粒子的刚性转子。这是一个双粒子系统, 粒子之间的距离固定, 由一根长度为 d 的刚性无质量杆连接。对于这个问题, 图6.1中的矢量 \mathbf{r} 的长度 $|\mathbf{r}|$ 是常数 d 。因此 (见第6.3节), 内运动的动能完全是转动能。转子的能量完全由动能提供, 以及

$$V = 0 \quad (6.44)$$

方程 (6.44) 是 (6.41) 的一个特例, 因此, 我们可以利用上一节得到的结果, 将系统的平动分离出来。我们将只关注转动能量。旋转运动的哈密顿算符由 (6.43) 中括号内的项给出:

$$\hat{H} = \frac{\hat{p}_\mu^2}{2\mu} = -\frac{\hbar^2}{2\mu} \nabla^2, \quad \mu = \frac{m_1 m_2}{m_1 + m_2} \quad (6.45)$$

其中 m_1 和 m_2 是两个粒子的质量。质量为 μ 的假想粒子的坐标为 m_1 和 m_2 之间的相对坐标。[式 (6.28)]。

与使用相对笛卡尔坐标 x, y 和 z 不同, 使用相对球坐标 r, θ 和 ϕ 来描述转子会更有成效。 r 坐标与图6.1中的矢量 \mathbf{r} 的长度相同, 由于 m_1 和 m_2 受限于保持固定的距离, 我们有 $r = d$ 。因此, 这个问题等价于一个质量为 μ 的粒子受限在半径为 r 的球面上运动。由于径向坐标是恒定的, 波函数将只是 θ 和 ϕ 的函数。因此, 对波函数进行运算时, 式 (6.8) 中拉普拉斯算符的前两项为零, 可

以省略。换个角度看问题，我们注意到式 (6.8) 中涉及 r 导数的算符对应于径向运动的动能，而由于没有径向运动， \hat{H} 中不包含 r 的导数。

由于 $V = 0$ 是 $V = V(r)$ 的特例，第 6.1 节的结论表明：本征函数由式 (6.16) 给出，其中 r 的因子被省略：

$$\psi = Y_l^m(\theta, \phi) \quad (6.46)$$

其中，转动角量子数用 J 而不是 l 来表示。

哈密顿算符由式 (6.8) 给出，其中省略了 r 的导数及 $V(r) = 0$ 。因此，

$$\hat{H} = (2\mu d^2)^{-1} \hat{L}^2$$

使用式 (6.13)，有

$$\begin{aligned} \hat{H}\psi &= E\psi \\ (2\mu d^2)^{-1} \hat{L}^2 Y_J^m(\theta, \phi) &= E Y_J^m(\theta, \phi) \\ (2\mu d^2)^{-1} J(J+1) \hbar^2 Y_J^m(\theta, \phi) &= E Y_J^m(\theta, \phi) \\ E &= \frac{J(J+1) \hbar^2}{2\mu d^2} \end{aligned} \quad (6.47)$$

由 n 个粒子组成的系统绕空间某一特定轴旋转产生的**转动惯量**（moment of inertia） I 定义为

$$I \equiv \sum_{i=1}^n m_i \rho_i^2 \quad (6.48)$$

其中 m_i 是粒子 i 的质量， ρ_i 是粒子 i 到旋转轴的垂直距离。 I 的值与选择的旋转轴有关。对于双粒子的刚性转子，我们选择通过质心并垂直于连接 m_1 和 m_2 的直线作为我们的轴线（图 6.2）。如果我们将转子置于坐标系的 x 轴上，使得质心 C 位于原点，连接 m_1 和 m_2 的直线与 x 轴重合，那么 C 的坐标为 $(0, 0, 0)$ ， m_1 的坐标为 $(-\rho_1, 0, 0)$ ， m_2 的坐标为 $(\rho_2, 0, 0)$ 。将这些坐标代入式 (6.30)，我们得到

$$m_1 \rho_1 = m_2 \rho_2 \quad (6.49)$$

转子绕我们选择的轴旋转的转动惯量为

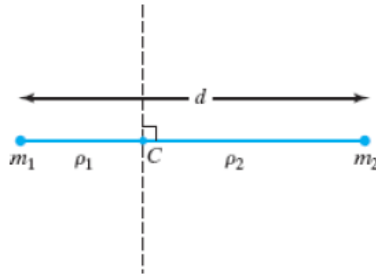


图 6.2:
用于计算双粒子刚性转子
转动惯量的轴线（虚线）。
 C 是质心。

$$I = m_1 \rho_1^2 + m_2 \rho_2^2 \quad (6.50)$$

使用式 (6.49)，我们可以将 (6.50) 写成（问题 6.14）

$$I = \mu d^2 \quad (6.51)$$

其中 $\mu \equiv m_1 m_2 / (m_1 + m_2)$ 是系统的折合质量， $d \equiv \rho_1 + \rho_2$ 是 m_1 和 m_2 之间的距离。双粒子刚性转子允许的能级由式 (6.47) 给出：

$$E = \frac{J(J+1)\hbar^2}{2I}, \quad J = 0, 1, 2, \dots \quad (6.52)$$

最低的能级满足 $E = 0$ ，所以没有零点转动能。转动能为零，因此转子的角动量为零，这并不违反不确定性原理；回顾公式 (5.105) 后面的讨论。请注意： E 随着 $J^2 + J$ 的增大而增大，因此相邻旋转能级的间隔随着 J 的增大而增大。

转子的能级 (6.52) 是简并的吗？能量只取决于 J ，而波函数 (6.46) 同时取决于 J 和 m ，其中 $m\hbar$ 是转子角动量在 z 方向上的分量。对于每个 J 的值，都有 $2J+1$ 个 m 的可能值，从 $-J$ 到 J 。因此，能级是 $(2J+1)$ -重简并的。简并能级的状态具有不同的转子绕空间固定轴方向的角动量矢量方向。

波函数 (6.46) 中的 θ 和 ϕ 角是两个质点的相对坐标。如果我们建立一个笛卡尔坐标系，令转子的质心与原点重合，那么 θ 和 ϕ 将如图 6.3 所示。该坐标系与转子质心的平移运动相同，但不在空间中转动。

转动角动量 $[J(J+1)\hbar^2]^{1/2}$ 是两个粒子相对于系统质心 C 处原点的角动量。

双原子分子的旋转能级可以用双粒子刚性转子模型 (6.52) 来很好地近似。研究发现 (Levine, *Molecular Spectroscopy*, Section 4.4)：当双原子分子吸收或发射辐射时，允许的纯转动跃迁由以下选律给出：

$$\Delta J = \pm 1 \quad (6.53)$$

此外，分子必须具有非零偶极矩，才能显示出纯转动光谱。纯转动跃迁 (pure rotational transition) 是指只有转动量子数发生变化的跃迁。[振动-转动跃迁（第 4.3 节）涉及振动量子数和转动量子数的变化。] 相邻较低旋转能级之间的能级间隔明显小于振动能级之间的能级间隔，纯转动光谱位于微波（或远红外）区域。因此，双原子分子纯转动光谱线的频率（近似值）为：

$$\nu = \frac{E_{J+1} - E_J}{h} = \frac{[(J+1)(J+2) - J(J+1)]\hbar^2}{8\pi^2 I} = 2(J+1)B \quad (6.54)$$

$$B \equiv h/8\pi^2 I, \quad J = 0, 1, 2, \dots \quad (6.55)$$

B 称为分子的转动常数 (rotational constant)。

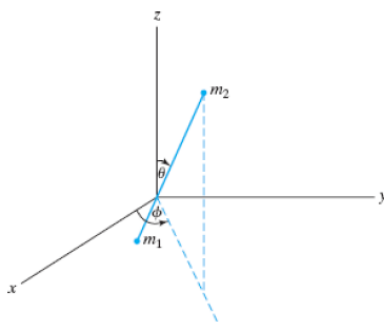


图 6.3: 双粒子刚性转子的坐标系

在室温下，具有较低 J 值的双粒子旋转能级间隔通常小于 kT 或与 kT 处于同一数量级，因此，玻尔兹曼分布律 (4.63) 表明：室温下许多转动能级可以被显著占据。 $J = 0$ 的双原子分子吸收辐射后会产生一条频率为 $2B$ 的谱线（对应 $J = 0 \rightarrow 1$ 的跃迁）， $J = 1$ 的分子吸收辐射后会产生一条频率为 $4B$ 的谱线（对应 $J = 1 \rightarrow 2$ 的跃迁）， $J = 2$ 的分子吸收辐射后会产生一条频率为 $6B$ 的谱线（对应 $J = 2 \rightarrow 3$ 的跃迁），依此类推。见图6.4。

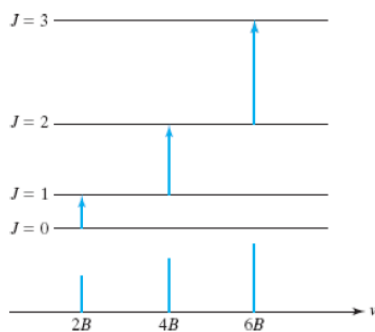


图 6.4: 双粒子刚性转子吸收跃迁

通过测量转动吸收频率可以得到 B 的值。根据 B ，我们可以得到分子的转动惯量 I ，根据 I 我们可以得到键长 d 。找到的 d 值是 $v = 0$ 振动运动的平均值。由于图 4.6 和图 13.1 中的势能曲线不对称， d 比图 13.1 中的平衡键长略长。

如第4.3节所述， $^1\text{H}^{35}\text{Cl}$ 和 $^1\text{H}^{37}\text{Cl}$ 等同位素分子的电子能量曲线 $U(R)$ 几乎相同，因此其平衡核间距也几乎相同。然而，不同的同位素质量会产生不同的转动惯量，从而产生不同的转动吸收频率。

由于分子不是刚性的，因此双原子分子的旋转能级与刚性转子的旋转能级略有不同。由式 (6.52) 和 (6.55)，双粒子刚性转子的能级为 $E_{rot} = BhJ(J+1)$ 。由于分子振动的非谐性（图4.6、4.5），当振动量子数 v 增加时，核间距平均值也会增加，所以 v 增加时，转动惯量 I 也会增加，则转动常数 B 减小。为了考虑 B 对 v 的依赖性，在 E_{rot} 中用 B_v 代替 B 。振动能级 v 的平均转动常数（mean rotational constant） B_v 满足 $B_v = B_e - \alpha_e(v + 1/2)$ ，其中 B_e 是利用图4.5和4.5势能曲线底部的平衡核间距 R_e 计算得出的，振转耦合常数（vibrational-rotational coupling constant） α_e 是一个大于零的常数（因分子的不同而异），且远小于 B_e 。此外，随着转动能量的增加，平衡核间距有非常轻微的增加 [这种现象称为离心形变（centrifugal distortion）]。这就在 E_{rot} 中引入了一项 $-hDJ^2(J+1)^2$ ，其中离心形变常数 D 是一个极小的正数，随着分子的不同而异。例如，对 $^{12}\text{C}^{15}\text{O}$ 分子， $B_0 = 57636 \text{ MHz}$ ， $\alpha_e = 540 \text{ MHz}$ ， $D = 0.18 \text{ MHz}$ 。如第4.3节所述，对于轻双原子分子，几乎所有分子在室温下都处在基态振动能级，观测到的转动常数为 B_0 。

有关双原子分子核运动的更多讨论，请参见第13.2节。关于多原子分子的转动能，请参见 Townes and Schawlow, chaps. 2-4。

例题：分子 $^{12}\text{C}^{32}\text{S}$ 的最低频率纯转动吸收线位于 48991.0 MHz 。计算分子的键长。

最低频率的转动吸收为 $J = 0 \rightarrow 1$ 线。根据式 (1.4)、(6.52) 和 (6.51)，有

$$h\nu = E_{upper} - E_{lower} = \frac{1(2)\hbar^2}{2\mu d^2} - \frac{0(1)\hbar^2}{2\mu d^2}$$

其中 $d = (h/4\pi^2\nu\mu)^{1/2}$ 。使用附录中的表 A.3, 有

$$\mu = \frac{m_1 m_2}{m_1 + m_2} = \frac{12 (31.97207)}{(12 + 31.97207)} \frac{1}{6.02214 \times 10^{23}} \text{ g} = 1.44885 \times 10^{-23} \text{ kg}$$

质量的国际单位为千克, 则

$$d = \frac{1}{2\pi} \left(\frac{h}{\nu_{0 \rightarrow 1} \mu} \right)^{1/2} = \frac{1}{2\pi} \left[\frac{6.62607 \times 10^{-34} \text{ J} \cdot \text{s}}{(48991.0 \times 10^6 \text{ s}^{-1}) (1.44885 \times 10^{-26} \text{ kg})} \right]^{1/2} \\ = 1.5377 \times 10^{-10} \text{ m} = 1.5377 \text{ \AA}$$

练习: 分子 $^{12}\text{C}^{16}\text{O}$ 的 $J = 1 \rightarrow 2$ 的纯转动跃迁位于 230.538 GHz。 (1 GHz = 10^9 Hz。) 计算分子的键长。(答案: $1.1309 \times 10^{-10} \text{ m}$ 。)

6.5 氢原子

氢原子包含一个质子和一个电子, 若用 e 表示质子的电荷量 ($e = +1.6 \times 10^{-19} \text{ C}$), 则电子的电荷量为 $-e$ 。

一些科学家推测, 质子和电子的电荷量并不完全相等。实验表明, 电子和质子的电荷量差距在 10^{-21} 的数量级。见 G. Bressi 等., *Phys. Rev. A*, **83**, 052101 (2011) (available online at arxiv.org/abs/1102.2766).

我们将假定质子和电子是质点, 其相互作用遵守库仑定律。在讨论原子和分子时, 我们通常会考虑孤立的系统, 而忽略原子和分子间的相互作用。

我们不只讨论氢原子, 而是稍微考虑一个更为普遍的问题: **类氢原子** (hydrogenlike atom), 它由一个电子和一个电荷为 Ze 的原子核组成。若 $Z = 1$, 则类氢原子就是氢原子; 若 $Z = 2$, 则类氢原子是 He^+ ; 若 $Z = 3$, 则类氢原子是 Li^{2+} ; 依此类推。类氢原子是量子化学中最重要的系统。由于存在电子间斥力, 因此无法获得具有一个以上电子的原子的薛定谔方程的精确解。如果作为粗略的第一近似值, 我们忽略这些斥力, 那么电子就可以独立处理 (参见第 6.2 节)。原子波函数将由单电子函数的乘积近似得出, 而单电子函数将是类氢波函数。单电子波函数 (无论是否为类氢波函数) 称为**轨道** (orbital)。(更确切地说, **轨道**是一个单电子空间波函数, 空间一词的意思是波函数取决于电子的三个空间坐标 x, y 和 z 或 r, θ 和 ϕ 。我们将在第 10 章中看到, 电子自旋的存在为单电子波函数增加了第四个坐标, 即所谓的自旋轨道。) 原子中电子的轨道称为**原子轨道** (atomic orbital)。我们将使用原子轨道来构建多电子原子的近似波函数 (第 11 章)。轨道还可用于构建分子的近似波函数。

对于类氢原子, 令 (x, y, z) 为电子相对于原子核的坐标, 再令 $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ 。电子在类氢原子中受到的库仑力为 [见公式 (1.37)]

$$\mathbf{F} = -\frac{Ze^2}{4\pi\epsilon_0 r^2} \frac{\mathbf{r}}{r} \quad (6.56)$$

其中 \mathbf{r}/r 是 \mathbf{r} 方向的单位矢量。负号表示为吸引力。

我们考虑了库仑定律出现微小偏差的可能性。实验表明, 如果将库仑定律力写成与 r^{-2+s} 成正比, 那么 $|s| < 10^{-16}$ 。库仑定律的偏差可以证明光子静止质量不为零。没有证据表明光子的静止质量不为零, 而且数据表明, 任何这种质量都必须小于 10^{-51} g ; A. S. Goldhaber and M. M. Nieto, *Rev. Mod. Phys.*, **82**, 939 (2010) (arxiv.org/abs/0809.1003); G. Spavieri et al., *Eur. Phys. J. D*, **61**, 531 (2011) (link.springer.com/content/pdf/10.1140/epjd/e2011-10508-7).

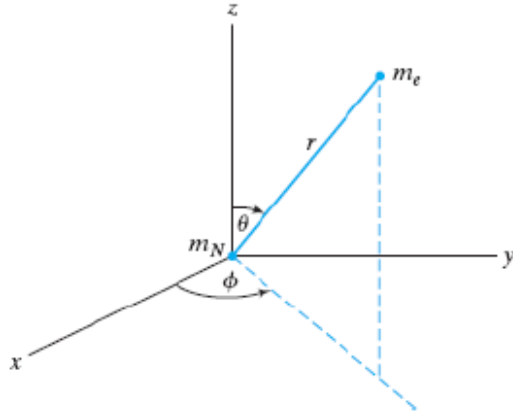


图 6.5: 相对球坐标

式 (6.56) 中的力是中心力，将其与式 (6.4) 相比较，有 $d(r)/dr = Ze^2/4\pi\epsilon_0 r^2$ 。积分，得出

$$V = \frac{Ze^2}{4\pi\epsilon_0} \int \frac{1}{r^2} dr = -\frac{Ze^2}{4\pi\epsilon_0 r} \quad (6.57)$$

其中积分常数取 0，以便在电荷间距无限大时能满足 $V = 0$ 。对于任意两个相距 r_{12} ，电荷量分别为 Q_1 和 Q_2 的粒子，式 (6.57) 可以写成

$$V = -\frac{Q_1 Q_2}{4\pi\epsilon_0 r_{12}} \quad (6.58)$$

由于这个双粒子系统的势能只取决于粒子的相对坐标，我们可以应用第 6.3 节的结果将问题简化为两个单粒子问题。原子的整体平移运动只是在总能量中增加了一个常数，我们将不再关注它。为了处理系统的内部运动，我们引入一个质量由下式定义的假想粒子

$$\mu = \frac{m_e m_N}{m_e + m_N} \quad (6.59)$$

其中 m_e 为电子的质量， m_N 为原子核的质量。质量为 μ 的粒子在势能函数 (6.57) 的作用下运动，其坐标 (r, θ, ϕ) 是一个粒子相对于另一个粒子的球面坐标（图6.5）。

粒子内部运动的哈密顿算符为 [式 (6.43)]

$$\hat{H} = -\frac{\hbar^2}{2\mu} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \quad (6.60)$$

由于 V 只是坐标 r 的函数，这是一个单粒子中心力问题，所以我们可以使用第6.1节的结果来求解。使用方程 (6.16) 和 (6.17)，我们可以写出波函数

$$\psi(r, \theta, \phi) = R(r) Y_l^m(\theta, \phi), \quad l = 0, 1, 2, \dots, \quad |m| \leq l \quad (6.61)$$

其中 Y_l^m 是球谐函数，径向函数 $R(r)$ 满足

$$-\frac{\hbar^2}{2\mu} \left(R'' + \frac{2}{r} R' \right) + \frac{l(l+1)\hbar^2}{2\mu r^2} R - \frac{Ze^2}{4\pi\epsilon_0 r} R = ER(r) \quad (6.62)$$

为了节省时间，我们定义常数 a 满足

$$a \equiv \frac{4\pi\epsilon_0 \hbar^2}{\mu e^2} \quad (6.63)$$

则方程 (6.62) 可以写成

$$R'' + \frac{2}{r} R' + \left[\frac{8\pi\epsilon_0 E}{ae^2} + \frac{2Z}{ar} - \frac{l(l+1)}{r^2} \right] R = 0 \quad (6.64)$$

径向方程的求解

现在我们可以尝试 (6.64) 的幂级数解法，但得到的是三项递推关系，而不是两项递推关系。因此，我们要寻找一种能得到两项递推关系的替代方法。事实证明，通过研究大 r 值时解的行为，可以找到合适的替代方法。当 r 值很大时，方程 (6.64) 变为

$$R'' + \frac{8\pi\epsilon_0 E}{ae^2} R = 0, \quad \text{当 } r \text{ 很大时} \quad (6.65)$$

该微分方程可以使用特征方程 (2.7) 求解。其解为

$$\exp \left[\pm (-8\pi\epsilon_0 E/ae^2)^{1/2} r \right] \quad (6.66)$$

假设 E 为正数。(6.66) 中平方根符号下的量是负数，乘以 r 的因子是虚数：

$$R(r) \sim e^{\pm i\sqrt{2\mu E}r/\hbar}, \quad E \geq 0 \quad (6.67)$$

其中我们用到了 (6.63)。(6.67) 中的符号 \sim 表示我们给出的是在 r 值较大时 $R(r)$ 的行为，这就是函数的渐近 (asymptotic) 行为。请注意 (6.67) 与自由粒子波函数公式 (2.30) 的相似性。(6.67) 并没有给出正能量波函数中的完整径向因子。进一步的研究 (Bethe and Salpeter, 第 21-24 页) 表明，无论 E 的值如何， $E \geq 0$ 的径向函数在所有 r 值上都是有限的。因此，正如自由粒子一样，氢原子的所有非负能量都是允许的。在物理学上，这些本征函数对应于电子不与原子核结合的状态，即原子电离时的状态。(一个经典的机械比喻是彗星围绕太阳在双曲线轨道上运动。彗星不受约束，只访问太阳系一次。) 由于我们得到的 $E \geq 0$ 的允许值是连续的而不是离散的，因此正能量本征函数又称为连续谱本征函数 (continuum eigenfunctions)。连续谱波函数的角度部分是球面谐波。与自由粒子波函数一样，连续谱本征函数也不是通常意义上可归一化的。

现在我们考虑氢原子的束缚态，即 $E < 0$ 。(对于一个束缚态，当 $x \rightarrow \pm\infty$ 时， $\psi \rightarrow 0$ 。) 在这种情况下，(6.66) 中括号内的量为正数。由于我们希望波函数在 r 变为无穷大时保持有限，因此我们更倾向于使用 (6.66) 中的负号，为了得到两项递推关系，我们将其替换为

$$R(r) = e^{-Cr} K(r) \quad (6.68)$$

$$C \equiv \left(-\frac{8\pi\epsilon_0 E}{ae^2} \right)^{1/2} \quad (6.69)$$

其中，(6.68) 中的 e 是自然对数的底数，而不是电子的电荷量。使用代换 (6.68) 无法保证当 r 很大时的波函数行为。这样代换得到的微分方程仍然有两个线性独立的解。我们可以在微分方程中进行任意代换；事实上，我们可以进行如下代换 $R(r) = e^{+Cr} J(r)$ 并最终得到正确的本征函数和本征值。 J 和 K 的关系自然是 $J(r) = e^{-2Cr} K(r)$ 。

根据 (6.68)，我们来计算 R' 和 R'' ，带入 (6.64)，乘以 $r^2 e^{Cr}$ 并利用 (6.69) 得到 $K(r)$ 的微分方程：

$$r^2 K'' + (2r - 2Cr^2) K' + [(2Za^{-1} - 2C)r - l(l+1)] K = 0 \quad (6.70)$$

现在，我们可以用以下形式的幂级数来代替 K ：

$$K = \sum_{k=0}^{\infty} c_k r^k \quad (6.71)$$

如果我们这样做，就会发现：一般来说，(6.71) 中的前几项的系数都是零。如果 c_s 是第一个不为零的系数，那么 (6.71) 可以写为

$$K = \sum_s c_k r^k, \quad c_s \neq 0 \quad (6.72)$$

令 $j \equiv k - s$ ，再定义 $b_j \equiv c_{j+s}$ ，我们有

$$K = \sum_{j=0}^{\infty} c_{j+s} r^{j+s} = r^s \sum_{j=0}^{\infty} b_j r^j, \quad b_0 \neq 0 \quad (6.73)$$

(虽然我们所做的各种替换看似随意，但它们却是用幂级数求解微分方程的标准程序。) 将整数 s 代入微分方程即可求出结果。方程 (6.73) 变为

$$K(r) = r^s M(r) \quad (6.74)$$

$$M(r) = \sum_{j=0}^{\infty} b_j r^j, \quad b_0 \neq 0 \quad (6.75)$$

求出 (6.74) 中的 K' 和 K'' ，将其代入 (6.70)，得到

$$r^2 M'' + [(2s+2)r - 2Cr^2] M' + [s^2 + s + (2Za^{-1} - 2C - 2Cs)r - l(l+1)] M = 0 \quad (6.76)$$

为了求出 s ，我们先看 $r=0$ 时的 (6.76)。由 (6.75)，我们得到

$$M(0) = b_0, \quad M'(0) = b_1, \quad M''(0) = 2b_2 \quad (6.77)$$

在 (6.76) 中使用 (6.77)，得到

$$b_0 (s^2 + s - l^2 - l) = 0 \quad (6.78)$$

由于 b_0 不为零，括号中式子就必须为零： $s^2 + s - l^2 - l = 0$ 。这是一个关于未知数 s 的二次方程。它的根为

$$s = l, \quad s = -l - 1 \quad (6.79)$$

这些根对应于微分方程的两个线性独立解。让我们从波函数正确行为的角度来研究它们。由 (6.68)、(6.74) 和 (6.75)，我们有

$$R(r) = e^{-Cr} r^s \sum_{j=0}^{\infty} b_j r^j \quad (6.80)$$

由于 $e^{-Cr} = 1 - Cr + \dots$ ，当 r 很小时， $R(r)$ 的行为与 $b_0 r^s$ 相同。当 $s = l$ 时， $R(r)$ 在原点处表现正常。然而，当 $s = -l - 1$ 时， $R(r)$ 与下式在 r 很小时成正比：

$$\frac{1}{r^{l+1}} \quad (6.81)$$

由于 $l = 0, 1, 2, \dots$ ，则根 $s = -l - 1$ 使得波函数中径向部分在原点处变得无限大。许多文献都将此作为拒绝接受这一根的充分理由。然而，这并不是一个很好的论据，因为对于相对论 (relativistic) 氢原子来说， $l = 0$ 的本征函数在 $r = 0$ 处是无限的。因此，让我们从平方可积性的角度来研究 (6.81)，因为我们当然要求束缚态本征函数是可归一化的。

与 (6.81) 类似的径向函数的归一化积分 [式 (5.80)] (当 r 不是很大时) 如下

$$\int_0^\infty |R|^2 r^2 dr \approx \int_0^\infty \frac{1}{r^{2l}} dr \quad (6.82)$$

积分下限处的积分行为是

$$\left. \frac{1}{r^{2l}} \right|_{r=0} \quad (6.83)$$

对于 $l = 1, 2, 3, \dots$ ，(6.83) 是无限的，那么归一化积分也是无限的。因此我们舍去当 $l \geq 1$ 时的根 $s = -l - 1$ 。不过，对于 $l = 0$ ，(6.83) 是有限的，平方可积性也没有问题。因此，径向方程存在一个平方可积的解，在小 r 时表现为 r^{-1} 。

对这一解决方案的进一步研究表明，它所对应的能量值是实验氢原子谱所不存在的。因此， r^{-1} 的解必须被舍去，但对这样做的原因存在一些争议。一种观点认为：除原点外， $1/r$ 的在空间的任何地方都满足薛定谔方程，因此必须予以舍去 [Dirac, page 156; B. H. Armstrong and E. A. Power, *Am. J. Phys.*, **31**, 262 (1963)]。第二种观点认为，必须舍去 $1/r$ 的解，因为哈密顿算符对它来说不是厄米的 (Merzbacher, Section 10.5)。(在第 7 章中，我们将定义厄米算符，并说明量子力学算符必须是厄米的。) 进一步讨论见 A. A. Khelashvili 和 T. P. Nadareishvili, *Am. J. Phys.*, **79**, 668 (2011) (see arxiv.org/abs/1102.1185) and in Y. C. Cantelaube, arxiv.org/abs/1203.0551.

取 (6.79) 的第一个根，我们得到了径向部分的因子 (6.80)

$$R(r) = e^{-Cr} r^l M(r) \quad (6.84)$$

由于 $s = l$, (6.75) 变为

$$rM'' + (2l + 1 - 2Cr) M' + (2Za^{-1} - 2C - 2Cl) M = 0 \quad (6.85)$$

由 (6.75)，我们有

$$\begin{aligned} M(r) &= \sum_{j=0}^{\infty} b_j r^j \quad (6.86) \\ M' &= \sum_{j=0}^{\infty} b_j j r^{j-1} = \sum_{j=1}^{\infty} j b_j r^{j-1} = \sum_{k=0}^{\infty} (k+1) b_{k+1} r^k = \sum_{j=0}^{\infty} (j+1) b_{j+1} r^j \\ M'' &= \sum_{j=0}^{\infty} j(j-1) b_j r^{j-2} = \sum_{j=1}^{\infty} b_j j(j-1) r^{j-2} = \sum_{k=0}^{\infty} (k+1) k b_{k+1} r^{k-1} \\ M'' &= \sum_{j=0}^{\infty} (j+1) j b_{j+1} r^{j-1} \quad (6.87) \end{aligned}$$

将这些式子代入 (6.85) 并求和，得到

$$\sum_{j=0}^{\infty} \left[j(j+1) b_{j+1} + 2(l+1)(j+1) b_{j+1} + \left(\frac{2Z}{a} - 2C - 2Cl - 2Cj \right) b_j \right] r^j = 0$$

令 r^j 的系数为零，得到递推关系式

$$b_{j+1} = \frac{2C + 2Cl + 2Cj - 2Za^{-1}}{j(j+1) + 2(l+1)(j+1)} b_j \quad (6.88)$$

现在我们必须研究大 r 时无穷级数 (6.86) 的行为。根据对 (4.42) 中谐振子幂级数的相同检验方法得出的结果，对于大 r ，无穷级数 (6.86) 的表现与 e^{2Cr} 一致 (见问题 6.20)。当 r 很大时，径向函数 (6.84) 的表现

$$R(r) \sim e^{-Cr} r^l e^{2Cr} = r^l e^{Cr} \quad (6.89)$$

因此，当 r 变得无穷大时， $R(r)$ 将变为无穷大，并且不具有平方可积性。避免这种“无穷大灾难” (如谐振子情况) 的唯一方法是让级数在有限的项数后被打断，在这种情况下， e^{-Cr} 因子将确保波函数在 r 变为无穷大时趋于零。令该级数的最后一项为 $b_k r^k$ 。那么为了使 b_{k+1}, b_{k+2}, \dots 都为零，我们必须让 $j = k$ 时，递推关系 (6.88) 中与 b_j 相乘的系数消失。因此，

$$2C(k+l+1) = 2Za^{-1}, \quad k = 0, 1, 2, \dots \quad (6.90)$$

k 和 l 都是整数，现在我们定义一个新的整数 n :

$$n \equiv k + l + 1, \quad n = 1, 2, 3, \dots \quad (6.91)$$

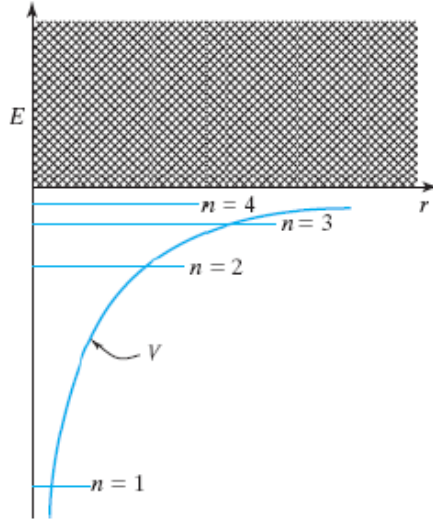


图 6.6: 氢原子的能级

根据 (6.91), 量子数 l 必须满足

$$l \leq n - 1 \quad (6.92)$$

因此, l 的取值范围为 0 到 $n - 1$ 。

能级

将 (6.91) 代入 (6.90), 得到

$$Cn = Za^{-1} \quad (6.93)$$

将 $C \equiv (-8\pi\epsilon_0 E/ae^2)^{1/2}$ [式 (6.69)] 代入 (6.93), 解出 E , 得到

$$E = -\frac{Z^2}{n^2} \frac{e^2}{8\pi\epsilon_0 a} = -\frac{Z^2 \mu e^4}{8\epsilon_0^2 n^2 h^2} \quad (6.94)$$

其中 $a \equiv 4\pi\epsilon_0 \hbar^2 / \mu e^2$ [式 (6.63)]。这些是类氢原子的束缚态能级, 它们是离散的。图 6.6 显示了氢原子 ($Z = 1$) 的势能曲线 [公式 (6.57)] 和一些允许能级。横线表示允许使用所有正能量。

事实证明, 在光的吸收和发射过程中, n 的所有变化都是允许的。因此, 氢原子谱线的波长 [公式 (4.64)] 为

$$\tilde{\nu} = \frac{1}{\lambda} = \frac{\nu}{c} = \frac{E_2 - E_1}{hc} = \frac{e^2}{8\pi\epsilon_0 a h c} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (6.95)$$

其中 $R_H = 109677.6 \text{ cm}^{-1}$ 是氢的里德伯常数 (Rydberg constant)。

简并

氢原子的能级是简并的吗? 对于束缚态, 能量 (6.94) 只取决于 n 。然而, 波函数 (6.61) 取决于所有三个量子数 n 、 l 和 m , 它们的允许值分别为 [式 (6.91)、(6.92)、(5.104) 和 (5.105)]

$$\boxed{n = 1, 2, 3, \dots} \quad (6.96)$$

$$l = 0, 1, 2, \dots, n-1 \quad (6.97)$$

$$m = -l, -l+1, \dots, 0, \dots, l-1, l \quad (6.98)$$

l 或 m 值不同但 n 值相同的氢原子态具有相同的能量。能级是简并的，只有 $n=1$ 除外，此时 l 和 m 必须都为 0。对于给定的 n 值，我们可以有 n 个不同的 l 值。对于每个 l 值，我们都可以有 $2l+1$ 个 m 值。H 原子束缚态能级的简并度等于 n^2 （省略了旋转方面的考虑），见问题 6.16。对于连续能级，事实证明，对于给定的能量， l 的最大值没有限制；因此，这些能级是无穷重简并的。

氢原子的径向方程也可以通过梯形算符（也称为因式分解（factorization））来求解；参见 Z. W. Salsburg, *Am. J. Phys.*, **33**, 36 (1965)。

6.6 束缚态氢原子波函数

径向因子

使用 (6.93)，我们有递推关系 (6.88)

$$b_{j+1} = \frac{2Z}{na} = \frac{j+l+1-n}{(j+1)(j+2l+2)} b_j \quad (6.99)$$

(6.91) 前面的讨论表明：多项式 $M(r) = \sum_j b_j r^j$ 中 r 的最高次幂为 $k = n-l-1$ 。因此，将 $C = Z/na$ (6.93) 代入 $R(r) e^{-Cr} M(r)$ (6.84)，我们得到

$$R_{nl}(r) = r^l e^{-Zr/na} \sum_{j=0}^{n-l-1} b_j r^j \quad (6.100)$$

其中 $a \equiv 4\pi\epsilon_0\hbar^2/\mu e^2$ [式 (6.63)]。完整的类氢束缚态波函数为 [式 (6.61)]

$$\psi_{nlm} = R_{nl}(r) Y_l^m(\theta, \phi) = R_{nl}(r) S_{lm}(\theta) \frac{1}{\sqrt{2\pi}} e^{im\phi} \quad (6.101)$$

其中前几个 θ 函数见表 5.1。

$R(r)$ 有多少个节点？径向函数在 $r = \infty$ 、 $l \neq 0$ 时 $r=0$ 和使 $M(r)$ 消失的 r 值处为零。 $M(r)$ 是一个 $n-l-1$ 次多项式，可以证明， $M(r)=0$ 的所有根均为正实数。因此，除了原点和无穷远之外， $R(r)$ 还有 $n-l-1$ 个节点。球谐函数的节点将在问题 6.41 中讨论。

基态波函数和能量

对于类氢原子的基态，我们有 $n=1$ ， $l=0$ ， $m=0$ 。因此，径向因子 (6.100) 为

$$R_{10}(r) = b_0 e^{-Zr/a} \quad (6.102)$$

常数 b_0 可以通过归一化条件 [式 (5.80)] 来确定：

$$|b_0|^2 \int_0^\infty e^{-2Zr/a} r^2 dr = 1$$

使用附录中的积分 (A.8)，我们得到

$$R_{10}(r) = 2 \left(\frac{Z}{a} \right)^{3/2} e^{-Zr/a} \quad (6.103)$$

乘上 $Y_0^0 = 1/(\sqrt{4\pi})$ ，我们得到基态波函数：

$$\psi_{100} = \frac{1}{\pi^{1/2}} \left(\frac{Z}{a} \right)^{3/2} e^{-Zr/a} \quad (6.104)$$

氢原子的能量和波函数涉及约化质量，由式 (6.59) 给出

$$\mu_H = \frac{m_e m_p}{m_e + m_p} = \frac{m_e}{1 + m_e/m_p} = \frac{m_e}{1 + 0.000544617} = 0.9994557 m_e \quad (6.105)$$

其中 m_p 是质子的质量， m_e/m_p 的值从表 A.1 中可以查到。约化质量与电子的质量非常接近。正因为如此，一些教科书在 H 原子薛定谔方程中使用电子的质量而不是约化质量。这相当于假设质子的质量与 (6.105) 中电子的质量相比是无限大的，并且所有的内部运动都是电子的运动。对于氢原子来说，用电子质量代替约化质量所带来的误差约为 1/2000。对于较重的原子，假设原子核无限重所带来的误差甚至比这还要小。此外，对于多电子原子，核运动的修正形式也相当复杂。基于这些原因，我们今后将假设原子核无限重，并在书写原子的薛定谔方程时简单地使用电子质量。

如果我们用电子质量代替氢原子的约化质量，则由 (6.63) 定义的量 a 变为

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.529177 \text{ \AA} \quad (6.106)$$

其中下标 0 表示使用电子质量而不是约化质量。因为 a_0 是玻尔理论中氢原子基态中电子移动的圆半径，所以它被称为玻尔半径 (Bohr radius)。当然，由于基态波函数 (6.104) 对所有有限的 r 值都不为零，因此在离原子核的任何距离上都有一定概率找到电子。电子当然不会被限制在一个圆内。

电子能量的方便单位是**电子伏特** (electronvolt) (eV)，定义为电子在 1 伏特 (V) 电位差中加速所获得的动能。电位差的定义是单位电荷的能量。由于 $e = 1.6021766 \times 10^{-19} \text{ C}$ 以及 $1 \text{ V C} = 1 \text{ J}$ ，因此

$$1 \text{ eV} = 1.6021766 \times 10^{-19} \text{ J} \quad (6.107)$$

例题：使用国际单位制计算氢原子的基态能量，并将结果转换为电子伏特。

氢原子基态能量由式 (6.94) 给出，其中 $n = 1$ ， $Z = 1$ ，能量为 $E = -\mu e^4 / 8\hbar^2 \epsilon_0^2$ 。将 μ 用 (6.105) 代替，得到

$$\begin{aligned} E &= - \frac{0.9994557 (9.109383 \times 10^{-31} \text{ kg}) (1.6021766 \times 10^{-19} \text{ C})^4}{8 (6.626070 \times 10^{-34} \text{ J} \cdot \text{s})^2 (8.8541878 \times 10^{-12} \text{ C}^2/\text{N} \cdot \text{m}^2)^2} \frac{Z^2}{n^2} \\ E &= - (2.178686 \times 10^{-18} \text{ J}) (Z^2/n^2) [(1 \text{ eV}) / (1.6021766 \times 10^{-19} \text{ J})] \\ E &= - (13.598 \text{ eV}) (Z^2/n^2) = -13.598 \text{ eV} \end{aligned} \quad (6.108)$$

这是一个值得记住的数字。基态氢原子电离所需的最小能量为 13.598 eV。

练习：求 Li^{2+} 的 $n = 2$ 能量，用电子伏特表示；尽可能减少计算量。（答案：-30.60 eV。）

例题：求氢原子基态的动能平均期望 $\langle T \rangle$ 。

对 $\langle T \rangle$ 的计算, 我们使用 (3.89); 而对 $\nabla^2\psi$ 的计算, 我们使用 (6.7)。因此,

$$\langle T \rangle = \int \psi^* \hat{T} \psi d\tau = -\frac{\hbar^2}{2\mu} \int \psi^* \nabla^2 \psi d\tau$$

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r} \frac{\partial \psi}{\partial r} - \frac{1}{r^2 \hbar^2} \hat{L}^2 \psi = \frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r} \frac{\partial \psi}{\partial r}$$

其中, 我们用到了 $\hat{L}^2 \psi = l(l+1)\hbar^2 \psi$ 以及对于 s 态有 $l=0$ 。将 $Z=1$ 带入 (6.104), 我们有 $\psi = \pi^{-1/2} a^{-3/2} e^{-r/a}$ 。因此, $\partial\psi/\partial r = -\pi^{-1/2} a^{-5/2} e^{-r/a}$, $\partial^2\psi/\partial r^2 = \pi^{-1/2} a^{-7/2} e^{-r/a}$ 。使用 $d\tau = r^2 \sin\theta dr d\theta d\phi$ (5.78), 我们有

$$\langle T \rangle = -\frac{\hbar^2}{2\mu} \frac{1}{\pi a^4} \int_0^{2\pi} \int_0^\pi \int_0^\infty \left(\frac{1}{a} e^{-2r/a} - \frac{2}{r} e^{-2r/a} \right) r^2 \sin\theta dr d\theta d\phi$$

$$= -\frac{\hbar^2}{2\mu\pi a^4} \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta \int_0^\infty \left(\frac{r^2}{a} e^{-2r/a} - 2r e^{-2r/a} \right) dr = \frac{\hbar^2}{2\mu a^2} = \frac{e^2}{8\pi\epsilon_0 a}$$

其中, 我们用到了附录中的积分 A.8 以及 $a = 4\pi\epsilon_0 \hbar^2 / \mu e^2$ 。根据 (6.94), $e^2/8\pi\epsilon_0 a$ 为基态氢原子能量的负值, $\langle T \rangle$ 由 (6.108) 给出 $\langle T \rangle = 13.598 \text{ eV}$ 。(另见第14.4节。)

练习: 使用 (6.113), 求氢原子 $2p_0$ 的动能平均期望 $\langle T \rangle$ 。(答案: $e^2/32\pi\epsilon_0 a = (13.598 \text{ eV})/4 = 3.40 \text{ eV}$ 。)

让我们来看看基态波函数 (6.104) 的一个重要性质。我们有 $r = (x^2 + y^2 + z^2)^{1/2}$ 。对于在 x 轴上的点, $y=0, z=0$, 则 $r = (x^2)^{1/2} = |x|$ 。即

$$\psi_{100}(x, 0, 0) = \pi^{-1/2} (Z/a)^{-3/2} e^{-Z|x|/a} \quad (6.109)$$

图 6.7 显示了 (6.109) 沿 x 轴的变化情况。尽管 ψ_{100} 在原点处连续, 但曲线切线的斜率在原点左侧为正, 而在原点右侧为负。因此 $\partial\psi/\partial x$ 在原点处不连续。我们说波函数在原点处有一个尖点 (cusp)。尖顶的出现是因为势能 $V = -Ze^2/4\pi\epsilon_0 r$ 在原点处变得无限大。回想一下, 箱中粒子波函数在箱壁上的斜率是不连续的。

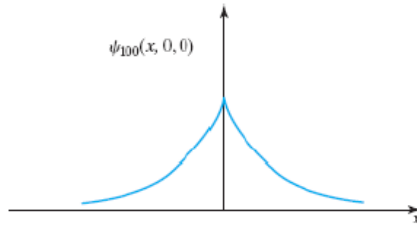


图 6.7: 氢原子基态波函数中的尖点

我们用三个下标来表示氢原子束缚态波函数, 分别表示 n 、 l 和 m 的值。在另一种表示中, l 的值用字母表示:

字母	s	p	d	f	g	h	i	k	...
l	0	1	2	3	4	5	6	7	...

(6.110)

字母 s 、 p 、 d 、 f 源于光谱学, 分别代表尖锐 (sharp)、主要 (principal)、扩散 (diffuse) 和基本

(fundamental)。之后，我们按字母顺序排列，只是省略了 j 。在 l 的字母前，我们写上 n 的值。因此，基态波函数 ψ_{100} 可以写成 ψ_{1s} 或更简单的 $1s$ 。

$n = 2$ 的波函数

对于 $n = 2$ ，我们有状态 ψ_{200} 、 ψ_{21-1} 、 ψ_{210} 、 ψ_{211} 。我们将 ψ_{200} 表示为 ψ_{2s} 或简单的 $2s$ 。为了区分三个 $2p$ 函数，我们使用下标来表示 m 值，并将它们表示为 $2p_1$ 、 $2p_0$ 和 $2p_{-1}$ 。从 (6.100) 可以看出，波函数中的径向因子取决于 n 和 l ，但不取决于 m 。因此，三个 $2p$ 波函数都具有相同的径向因子。根据 (6.100) 和 (6.99) 可以用通常的方法求出 $2s$ 和 $2p$ 的径向因子，然后进行归一化处理。结果见表 6.1。请注意， $n = 2$ 径向函数中的指数因子与 R_{1s} 函数中的指数因子不同。将径向因子乘以适当的球谐函数，就能得到完整的波函数。使用 (6.101)、表 6.1 和表 5.1，我们有

$$2s = \frac{1}{2\pi^{1/2}} \left(\frac{Z}{a}\right)^{3/2} \left(1 - \frac{Zr}{2a}\right) e^{-Zr/2a} \quad (6.111)$$

$$2p_{-1} = \frac{1}{8\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \sin\theta e^{-i\phi} \quad (6.112)$$

$$2p_0 = \frac{1}{2\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \cos\theta \quad (6.113)$$

$$2p_1 = \frac{1}{8\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \sin\theta e^{i\phi} \quad (6.114)$$

表 6.1: 类氢原子波函数的径向因子

$R_{1s} = 2 \left(\frac{Z}{a}\right)^{3/2} e^{-Zr/a}$
$R_{2s} = \frac{1}{\sqrt{2}} \left(\frac{Z}{a}\right)^{3/2} \left(1 - \frac{Zr}{2a}\right) e^{-Zr/2a}$
$R_{2p} = \frac{1}{2\sqrt{6}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a}$
$R_{3s} = \frac{2}{3\sqrt{3}} \left(\frac{Z}{a}\right)^{3/2} \left(1 - \frac{Zr}{3a} + \frac{2Z^2 r^2}{27a^2}\right) e^{-Zr/3a}$
$R_{3p} = \frac{8}{27\sqrt{6}} \left(\frac{Z}{a}\right)^{3/2} \left(\frac{Zr}{a} - \frac{Z^2 r^2}{6a^2}\right) e^{-Zr/3a}$
$R_{3d} = \frac{4}{81\sqrt{30}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a}$

表 6.1 列出了类氢波函数中的一些归一化径向因子。图 6.8 是一些径向函数的示意图。除了 s 状态以外，因子 r^l 使得径向函数在 $r = 0$ 处为零。

径向分布函数

在空间中位于 r 到 $r + dr$ 、 θ 到 $\theta + d\theta$ 、 ϕ 到 $\phi + d\phi$ 的体积微元内找到电子的概率为 [式 (5.78)]

$$|\psi|^2 d\tau = [R_{nl}(r)]^2 |Y_l^m(\theta, \phi)|^2 r^2 \sin\theta dr d\theta d\phi \quad (6.115)$$

我们现在想知道：在不限 θ 和 ϕ 值的情况下，电子的径向坐标介于 r 和 $r + dr$ 之间的概率是多少？我们想知道在以原点为中心的薄球壳中找到电子的概率，其内半径为 r ，外半径为 $r + dr$ 。因

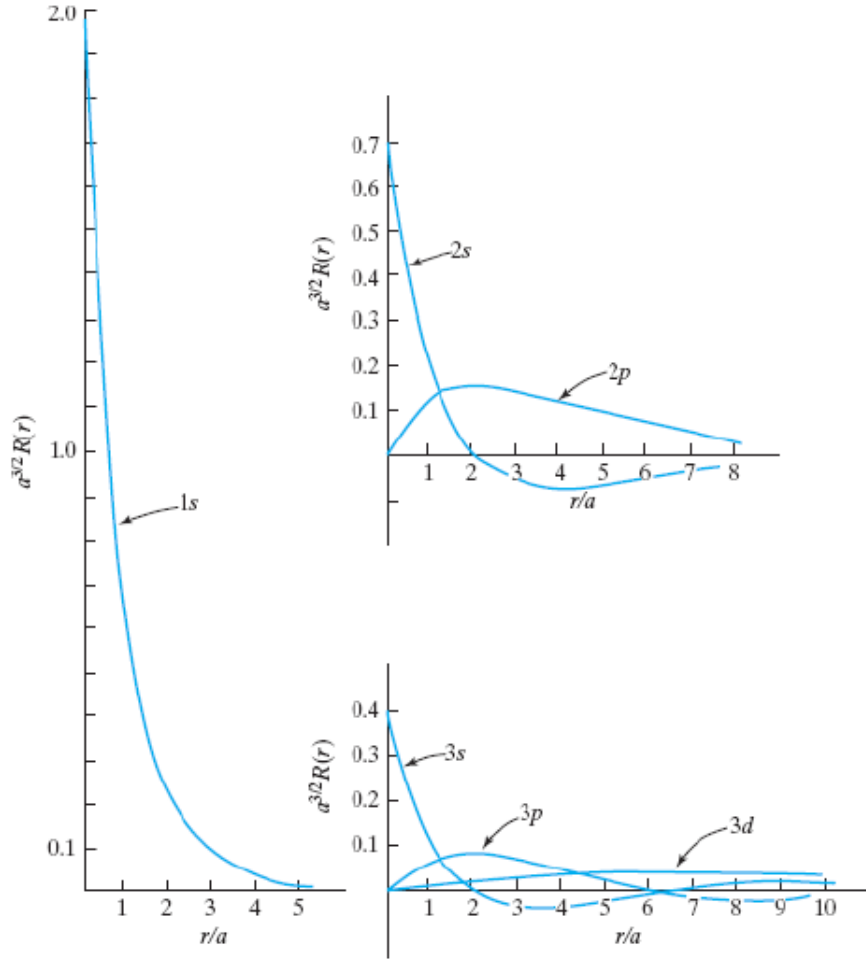


图 6.8:
氢原子 ($Z = 1$) 波函数
径向因子 $R_{nl}(r)$ 的示意图。
所有图表都使用了相同的比例尺。
(在某些课本中, 这些函数没有按比例绘制)。

此, 我们必须在保持 r 不变的情况下, 将 θ 和 ϕ 所有可能值的无穷小概率 (6.115) 相加。这相当于在 θ 和 ϕ 上对 (6.115) 进行积分。因此, 在 r 到 $r + dr$ 的区间内找到电子的概率为

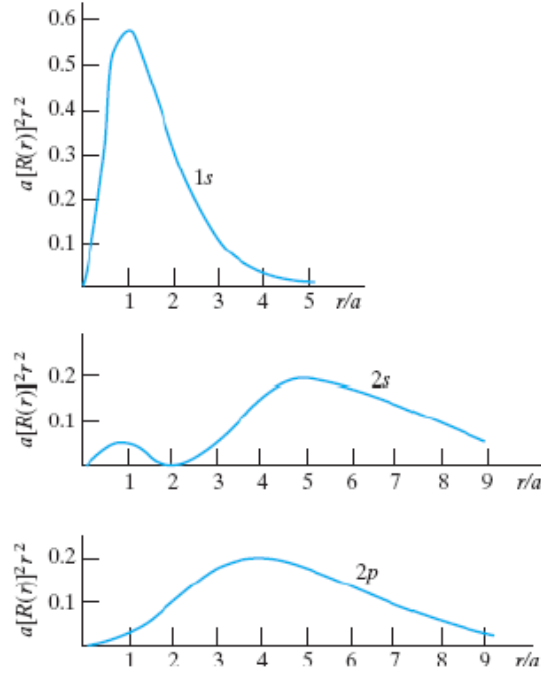
$$[R_{nl}(r)]^2 r^2 dr \int_0^{2\pi} \int_0^\pi |Y_l^m(\theta, \phi)|^2 \sin \theta d\theta d\phi = [R_{nl}(r)]^2 r^2 dr \quad (6.116)$$

其中, 我们用到了球谐函数是归一化的:

$$\int_0^{2\pi} \int_0^\pi |Y_l^m(\theta, \phi)|^2 \sin \theta d\theta d\phi = 1 \quad (6.117)$$

从 (5.72) 和 (5.80) 可以看出。函数 $R^2(r) r^2$ 决定了在距离原子核 r 处找到电子的概率, 称为**径向分布函数** (radial distribution function); 见图 6.9。

对于氢原子的 $1s$ 基态, 概率密度 $|\psi|^2$ 由式 (6.104) 给出, 等于 $e^{-2r/a}$ 乘上一个常数, 因此 $|\psi_{1s}|^2$ 在 $r = 0$ 处达到最大值 (见图 6.14)。然而, 径向分布函数 $[R_{1s}(r)]^2 r^2$ 在原点处为零, 而在 $r = a$ 处达到最大值 (见图 6.9)。这两个事实并不矛盾。概率密度 $|\psi|^2$ 与在体积为 $dx dy dz$ 的无限小盒子中找到电子的概率成正比, 而这一概率在原子核处为最大值。径向分布函数与在内外半径分别为 r 和

图 6.9: 氢原子径向分布函数 $[R_{nl}(r)]^2 r^2$ 的图象。

$r + dr$ 的薄球壳中找到电子的概率成正比，这一概率在 $r = a$ 处达到最大值。由于 ψ_{1s} 只与 r 有关，1s 的概率密度在薄球壳中基本恒定。如果我们把薄壳想象成无数个无穷小的盒子，每个盒子的体积都是 $dxdydz$ ，那么我们就可以把在薄壳中每个小盒子中的概率 $|\psi_{1s}|^2 dxdydz$ 相加，得到在薄壳中找到电子的概率是 $|\psi_{1s}|^2 V_{\text{shell}}$ 。薄壳的体积 V_{shell} 为

$$\frac{4}{3}\pi(r+dr)^3 - \frac{4}{3}\pi r^3 = 4\pi r^2 dr$$

其中 $(dr)^2$ 和 $(dr)^3$ 与 dr 相比可以忽略。因此，位于薄壳中的概率是

$$|\psi_{1s}|^2 V_{\text{shell}} = R_{1s}^2 (Y_0^0)^2 4\pi r^2 dr = R_{1s}^2 \left[(4\pi)^{-1/2}\right]^2 4\pi r^2 dr = R_{1s}^2 r^2 dr$$

与 (6.116) 一致。1s 径向分布函数在 $r = 0$ 时为零，因为当 r 变为零时，薄球壳的体积 $4\pi r^2 dr$ 变为零。它们的乘积 $|\psi_{1s}|^2 4\pi r^2 dr$ 在 $r = a$ 处达到最大值。

例题：求基态 H 原子中电子与原子核的距离小于 a 的概率。

我们希望找到径向坐标介于 0 到 a 之间的概率。将处于 r 和 $r + dr$ 之间的无穷小概率 (6.116) 在 0 到 a 的范围内求和，就可以得出这个结果。这个无穷小量之和就是定积分

$$\begin{aligned} \int_0^a R_{nl}^2 r^2 dr &= \frac{4}{a^3} \int_0^a e^{-2r/a} r^2 dr = -\frac{4}{a^3} e^{-2r/a} \left(-\frac{r^a}{2} - \frac{2ra^2}{4} - \frac{2a^3}{8} \right) \Big|_0^a \\ &= 4 [e^{-2} (-5/4) - (-1/4)] = 0.323 \end{aligned}$$

其中 R_{10} 来自表 6.1，用到了附录中的积分 A.7。

练习：求 $2p_1$ H 原子中电子与原子核的距离小于 a 的概率。使用积分表或网站 integrals.wolfram.com。（答案：0.00366。）

实际类氢函数

除了 $m = 0$ 以外, 因子 $e^{im\phi}$ 使得球谐函数为复数。化学家通常不使用 (6.112) 和 (6.114) 等复波函数, 而是使用由复函数的线性组合形成的实类氢波函数。第 3.6 节的定理给出了这一步骤的理由: 简并能级的任何本征函数的线性组合都是具有相同本征值的哈密顿算符的本征函数。由于氢原子的能量与 m 无关, 状态 $2p_1$ 和 $2p_{-1}$ 属于一个简并的能级。它们的任何线性组合都是哈密顿算符的本征函数, 具有相同的能量本征值。

将这两个函数组合起来得到一个实函数的方法是

$$2p_x \equiv \frac{1}{\sqrt{2}} (2p_1 + 2p_{-1}) = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \sin \theta \cos \phi \quad (6.118)$$

其中, 我们用到了 (6.112)、(6.114) 和 $e^{\pm i\phi} = \cos \phi \pm i \sin \phi$ 。因子 $1/\sqrt{2}$ 使得波函数 $2p_x$ 归一化:

$$\begin{aligned} \int |2p_x|^2 d\tau &= \frac{1}{2} \left(\int |2p_1|^2 d\tau + \int |2p_{-1}|^2 d\tau + \int (2p_{-1})^* 2p_1 d\tau + \int (2p_1)^* 2p_{-1} d\tau \right) \\ &= \frac{1}{2} (1 + 1 + 0 + 0) = 1 \end{aligned}$$

其中我们用到了 $2p_1$ 和 $2p_{-1}$ 是归一化且互相正交的性质, 因为

$$\int_0^{2\pi} (e^{i\phi})^* e^{i\phi} d\phi = \int_0^{2\pi} e^{2i\phi} d\phi = 0$$

如果我们注意到 (5.51) 给出了下式, 那么 (6.118) 的命名 $2p_x$ 就会更加清晰:

$$2p_x = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a}\right)^{5/2} x e^{-Zr/2a} \quad (6.119)$$

第二种组合函数的方法是

$$2p_y \equiv \frac{1}{i\sqrt{2}} (2p_1 - 2p_{-1}) = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a}\right)^{5/2} r \sin \theta \sin \phi e^{-Zr/2a} \quad (6.120)$$

$$2p_y = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a}\right)^{5/2} y e^{-Zr/2a} \quad (6.121)$$

函数 $2p_0$ 是实数, 通常用下式表示:

$$2p_0 = 2p_z = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a}\right)^{5/2} z e^{-Zr/2a} \quad (6.122)$$

其中大写的 Z 表示核中的质子数, 小写的 z 为电子的 z 坐标。函数 $2p_x$ 、 $2p_y$ 和 $2p_z$ 是互相正交的。注意 $2p_z$ 在 xOy 平面内为零, 在该平面上方为正, 在下方为负。

函数 $2p_{-1}$ 和 $2p_1$ 都是算符 \hat{L}^2 的本征函数, 其本征值也相同, 为 $2\hbar^2$ 。第 3.6 节的推理表明, 线性组合 (6.118) 和 (6.120) 也是 \hat{L}^2 的本征函数, 其本征值为 $2\hbar^2$ 。然而, $2p_{-1}$ 和 $2p_1$ 是 \hat{L}_z 的本征函数, 但本征值分别为 $-\hbar$ 和 \hbar 。因此, $2p_x$ 和 $2p_y$ 不是 \hat{L}_z 的本征函数。

我们可以推广这一过程, 为更高的状态构建实波函数。由于 m 的范围是从 $-l$ 到 $+l$, 对于每个包含因子 $e^{-i|m|\phi}$ 的复函数, 都有一个 n 和 l 值相同但包含因子 $e^{+i|m|\phi}$ 的函数, 将这些函数相减, 得到两个实函数, 其中一个具有因子 $\cos(|m|\phi)$, 另一个具有因子 $\sin(|m|\phi)$ 。表 6.2 列出了类氢原子的实波函数。这些函数的下标来自与 $2p_x$ 、 $2p_y$ 和 $2p_z$ 函数类似的考虑。例如, $3d_x$ 函数与 xy 成正比 (问题 6.37)。

实类氢波函数由复函数导出, 方法如下: 当 $m \neq 0$ 时, 将 $e^{im\phi}/(2\pi)^{1/2}$ 替换为 $\pi^{-1/2} \cos(|m|\phi)$ 或 $\pi^{-1/2} \sin(|m|\phi)$; 当 $m = 0$ 时, 对于实函数和复函数, ϕ 因子都为 $1/(2\pi)^{1/2}$ 。

在处理分子时，实类氢轨道比复类氢轨道更有用。例如，我们将在第节中看到，氧原子的实原子轨道 $2p_x$ 、 $2p_y$ 和 $2p_z$ 具有适当的对称性，可用于构建 H_2O 分子的波函数，而复数 $2p$ 轨道则不具备这种对称性。

表 6.2: 类氢原子的实波函数

$1s = \frac{1}{\pi^{1/2}} \left(\frac{Z}{a}\right)^{3/2} e^{-Zr/a}$
$2s = \frac{1}{4(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{3/2} \left(2 - \frac{Zr}{2a}\right) e^{-Zr/2a}$
$2p_z = \frac{1}{4(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \cos \theta$
$2p_x = \frac{1}{4(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \sin \theta \cos \phi$
$2p_y = \frac{1}{4(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{5/2} r e^{-Zr/2a} \sin \theta \sin \phi$
$3s = \frac{1}{81(3\pi)^{1/2}} \left(\frac{Z}{a}\right)^{3/2} \left(27 - 18\frac{Zr}{a} + 2\frac{Z^2 r^2}{a^2}\right) e^{-Zr/3a}$
$3p_z = \frac{2^{1/2}}{81\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} \left(6 - \frac{Zr}{a}\right) r e^{-Zr/3a} \cos \theta$
$3p_x = \frac{2^{1/2}}{81\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} \left(6 - \frac{Zr}{a}\right) r e^{-Zr/3a} \sin \theta \cos \phi$
$3p_y = \frac{2^{1/2}}{81\pi^{1/2}} \left(\frac{Z}{a}\right)^{5/2} \left(6 - \frac{Zr}{a}\right) r e^{-Zr/3a} \sin \theta \sin \phi$
$3d_{z^2} = \frac{1}{81(6\pi)^{1/2}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a} (3 \cos^2 \theta - 1)$
$3d_{xz} = \frac{2^{1/2}}{81(6\pi)^{1/2}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a} \sin \theta \cos \theta \cos \phi$
$3d_{yz} = \frac{2^{1/2}}{81(6\pi)^{1/2}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a} \sin \theta \cos \theta \sin \phi$
$3d_{x^2-y^2} = \frac{1}{81(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a} \sin^2 \theta \cos 2\phi$
$3d_{xy} = \frac{1}{81(2\pi)^{1/2}} \left(\frac{Z}{a}\right)^{7/2} r^2 e^{-Zr/3a} \sin^2 \theta \sin 2\phi$

6.7 类氢轨道

类氢波函数是单电子空间波函数，类氢轨道也是单电子空间波函数（第 6.5 节）。这些函数是针对单电子原子得出的，我们不能指望用它们来真正准确地表示多电子原子的波函数。第 11 章将讨论如何使用轨道概念来近似多电子原子的波函数。现在我们只讨论单电子原子。

描绘轨道有两种完全不同的方法。一种方法是绘制函数图象；第二种方法是绘制概率密度恒定的等值面。

首先考虑绘制函数图象。要将 ψ 的变化绘制成三个自变量 r 、 θ 和 ϕ 的函数，我们需要四个维度。我们世界的三维特性使我们无法绘制这样的图表。因此，我们绘制 ψ 中某个因子的图象。绘制 $R(r)$ 与 r 的关系图，我们得到图 6.8 的曲线，其中不包含 ψ 角度变化的任何信息。

现在来考虑 $S(\theta)$ 的图象，由表 (5.1)，我们有

$$S_{0,0} = 1/\sqrt{2}, \quad S_{1,0} = \frac{1}{2}\sqrt{6} \cos \theta$$

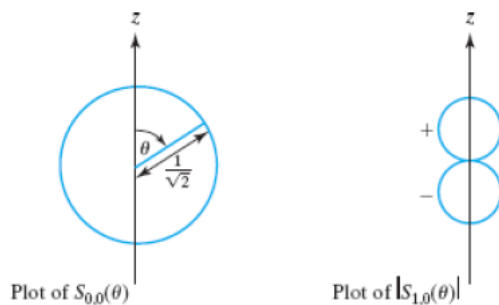


图 6.10: 氢原子波函数中 s 和 p_z 的 θ 因子的极坐标图。

我们可以使用二维直角坐标绘制这些函数的图象，在纵轴上绘制 S ，在横轴上绘制 θ 。 $S_{0,0}$ 给出了一条水平直线，而 $S_{1,0}$ 给出了一条余弦曲线。更常见的情况是使用平面极坐标系绘制 S 的图象。变量 θ 是与 z 轴正半轴的夹角，而 $S(\theta)$ 是从原点到图上点的距离。对于 $S_{0,0}$ 我们得到了一个圆；对于 $S_{1,0}$ ，我们得到了两个相切的圆（图 6.10）。 $S_{1,0}$ 图像中较低的圆旁边的负号，表示当 $\frac{1}{2}\pi < \theta \leq \pi$ 时， $S_{1,0}$ 为负值。严格来说，在绘制 $\cos \theta$ 时，我们只得到了较高的圆，它被绘制了两次；要得到两个相切的圆，我们必须绘制 $|\cos \theta|$ 。

我们可以绘制一张 $|S(\theta)T(\phi)|$ 与 θ 和 ϕ 的函数关系图，而不是分别绘制角向因子的图象。我们将使用球坐标，从原点到图上点的距离将为 $|S(\theta)T(\phi)|$ 。对于 s 状态， ST 与角度无关，因此我们得到了一个半径为 $1/(4\pi)^{1/2}$ 的球面。而对于 p_z 状态， $ST = \frac{1}{2}(3/\pi)^{1/2} \cos \theta$ ， $|ST|$ 的图象由两个球体组成，球体中心位于 z 轴，并与原点相切（图 6.11）。毫无疑问，图 6.11 是大家都很熟悉的。有些书上说这给出了 p_z 轨道的形状，这是错误的。图 6.11 只是一张波函数 p_z 中的角向因子图。角向因子 p_x 和 p_y 的图形分别给出了位于 x 和 y 轴上的切线球。如果我们以球面坐标绘制 S^2T^2 ，就会得到具有我们熟悉的八字形截面的表面；重复一遍，这些是图形，而不是轨道形状。

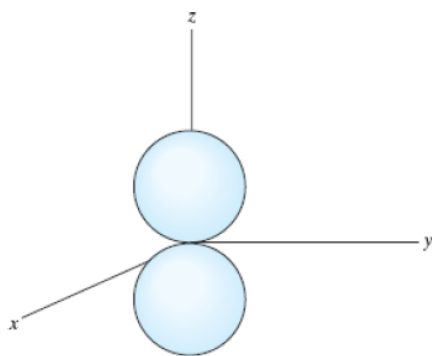


图 6.11: p_z 波函数中角向因子 $|Y_1^0(\theta, \phi)|$ 的图象。

现在，考虑绘制概率密度恒定的等值面。我们将在空间中画出曲面，在每个曲面上，概率密度 $|\psi|^2$ 的值都是恒定的。自然，如果 $|\psi|^2$ 在给定的曲面上是常数，那么 $|\psi|$ 在该曲面上也是常数。 $|\psi|^2$ 和 $|\psi|$ 的等值面是相同的。

对于 s 轨道， ψ 只与 r 有关，等值面就是一个恒定 r 的曲面，即以原点为中心的球面。为了确定一个轨道的大小，我们取一个等值面，在这个等值面内找到电子的概率比如说是 95%，因此我们希望 $\int_V |\psi|^2 d\tau = 0.95$ ，其中 V 是轨道等值面所包围的体积。

让我们来计算类氢轨道 $2p_y$ 在 yOz 平面上的横截面。在该平面上，有 $\phi = \pi/2$ （图 6.5）以及

$\sin \phi = 1$ ；因此表 6.2 给出了该轨道在 yOz 平面上的情况

$$|2p_y| = k^{5/2} \pi^{-1/2} r e^{-kr} |\sin \theta| \quad (6.123)$$

其中 $k = Z/2a$ 。为了找到轨道横截面，我们使用平面极坐标来绘制 ψ 固定值的 (6.123) 图； r 是到原点的距离，而 θ 是与 z 轴的夹角。图 6.12 显示了典型等值线的结果（问题 6.44）。由于 $ye^{-kr} = y \exp \left[-k(x^2 + y^2 + z^2)^{1/2} \right]$ ，我们可以看到 $2p_y$ 轨道是 y 和 $(x^2 + z^2)$ 的函数。因此，在以 y 轴为中心、平行于 xOz 平面的圆上， $2p_y$ 是常数。因此，将图 6.12 中的横截面绕 y 轴旋转，可以得到一对扭曲的椭圆体，从而形成三维等值面。实际 $2p$ 轨道的形状是两个分离的扭曲椭圆体，而不是两个相切的球体。

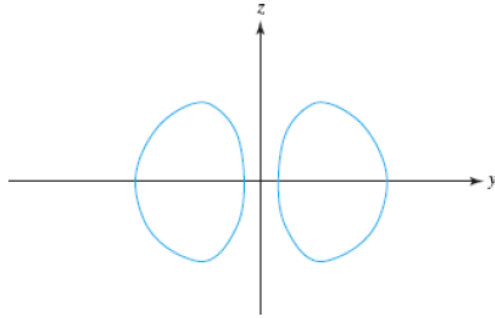


图 6.12: $2p_y$ 轨道等值线。

现在来考虑两个复轨道 $2p_{\pm 1}$ 的形状。我们有

$$2p_{\pm 1} = k^{5/2} \pi^{-1/2} r e^{-kr} \sin \theta e^{\pm i\phi}$$

$$|2p_{\pm 1}| = k^{5/2} \pi^{-1/2} e^{-kr} r |\sin \theta| \quad (6.124)$$

这两个轨道有相同的形状。由于 (6.124) 和 (6.123) 的右边相同，我们得出结论：图 6.12 也给出了 yOz 平面上 $2p_{\pm 1}$ 轨道的横截面。由于 [式 (5.51)]

$$e^{-kr} r |\sin \theta| = \exp \left[-k(x^2 + y^2 + z^2)^{1/2} \right] (x^2 + y^2)^{1/2}$$

我们看到 $2p_{\pm 1}$ 轨道是 z 和 $(x^2 + y^2)$ 的函数，因此，我们将图 6.12 绕 z 轴旋转，就能得到三维轨道形状。这样就形成了一个甜甜圈状的表面。

图 6.13 显示了一些类氢轨道面。 $2s$ 轨道有一个看不到的球形节点； $3s$ 轨道有两个这样的节点。 $3p_z$ 轨道有一个球形节点（虚线表示）和一个节点平面（ xOy 平面）。 $3d_{z^2}$ 轨道有两个节点锥。 $3d_{x^2-y^2}$ 轨道有两个节平面。请注意，不同轨道的视图并不相同。波函数的相对符号已标出。表 6.2 中的其他三个实 $3d$ 轨道与 $3d_{x^2-y^2}$ 轨道形状相同，但方向不同。 $3d_{xy}$ 轨道的裂片位于 x 轴和 y 轴之间，它是通过将 $3d_{x^2-y^2}$ 轨道绕 z 轴旋转 45° 而得到的。 $3d_{yz}$ 和 $3d_{xz}$ 轨道的裂片分别位于 y 轴和 z 轴之间以及 x 轴和 z 轴之间。（真实类氢轨道的在线三维视图见 www.falstad.com/qmatom；可使用鼠标旋转）

图 6.14 表示各种轨道在 yOz 平面上的概率密度。特定区域中点的数量与该区域中的 $|\psi|^2$ 值成正比。将这些图表绕垂直（ z ）轴旋转，就得到了三维概率密度。 $2s$ 轨道的角度系数是一个常数，因此没有角向节点；对于这个轨道， $n - l - 1 = 1$ ，表明有一个径向节点。图 6.14 显示了 $\psi_{2s} = 0$ 的球面。

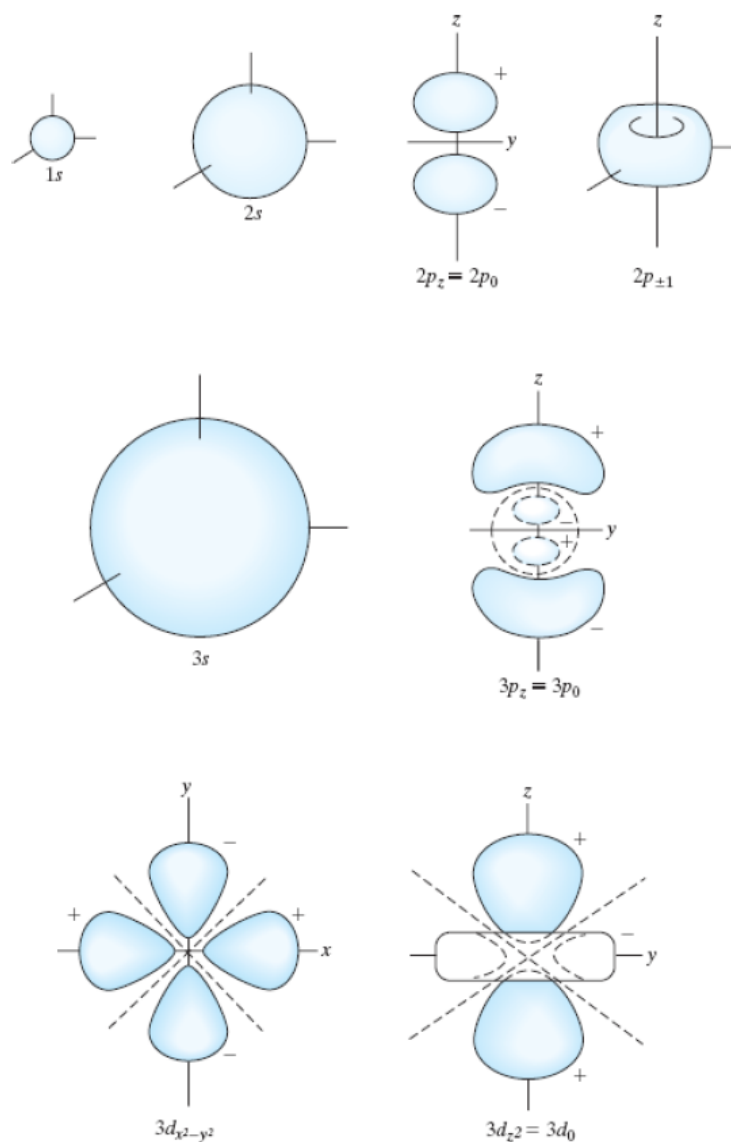
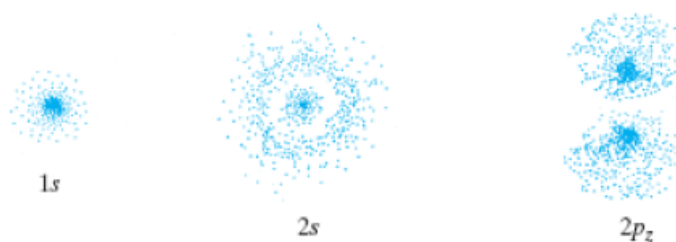


图 6.13: 一些氢原子轨道的形状。

图 6.14:
一些氢原子状态的概率密度。

有关精确的立体图，请参见 D. T. Cromer, *J. Chem. Educ.*, **45**, 626 (1968).

薛定谔最初对 $|\psi|^2$ 的解释是，电子被“涂抹”成电子云。如果我们考虑一个电子从一种介质传递到另一种介质，我们会发现 $|\psi|^2$ 在两种介质中都不为零。根据电子云的解释，这意味着电子的一部分被反射，一部分被透射。然而，在实验中，我们永远无法检测到电子的一小部分；电子是不可

分割的实体。根据玻恩解释，两种介质中的 $|\psi|^2$ 值给出了反射和透射的概率，从而解决了这一难题。我们绘制的轨道形状给出了找到电子的总概率为 95% 的空间区域。

6.8 塞曼效应

1896 年，塞曼 (Zeeman) 观察到外加磁场会导致原子光谱线分裂。我们将研究氢原子的塞曼效应 (Zeeman Effect)。我们首先回顾一下磁学。

磁场产生于移动的电荷。速度为 \mathbf{v} 的电荷 Q 在空间 P 点产生磁场 \mathbf{B} ，其中

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{Q\mathbf{v} \times \mathbf{r}}{r^3} \quad (6.125)$$

其中 \mathbf{r} 是从电荷 Q 到点 P 的位置矢量， μ_0 [称为**真空磁导率** (permeability of vacuum) 或**磁常数** (magnetic constant)] 定义为 $4\pi \times 10^{-7} \text{ N C}^{-2} \text{ s}^2$ 。[公式 (6.125) 只适用于速度远小于光速的非加速电荷。] 矢量 \mathbf{B} 称为**磁感应强度** (magnetic induction) 或**磁通量** (magnetic flux density)。(以前认为矢量 \mathbf{H} 是基磁场矢量，所以称 \mathbf{H} 为**磁场强度**。现在人们知道 \mathbf{B} 是基磁场矢量) 公式 (6.125) 采用国际单位制， Q 单位为库仑， \mathbf{B} 单位为特斯拉 (Tesla) (T)，其中 $1 \text{ T} = 1 \text{ N C}^{-1} \text{ m}^{-1} \text{ s}$ 。

两个电荷 $+Q$ 和 $-Q$ 相隔很小的距离 b 就构成了一个电偶极子。**电偶极矩** (electric dipole moment) 被定义为从 $-Q$ 到 $+Q$ 的矢量，大小为 Qb 。对于一个小的平面电流环，电流中的移动电荷所产生的磁场与电偶极子所产生的电场的数学表达式相同，只是电偶极矩被**磁偶极矩** (magnetic dipole moment) \mathbf{m} 所取代； \mathbf{m} 是一个大小为 IA 的矢量，其中 I 是在面积为 A 的环中流动的电流。 \mathbf{m} 的方向与电流回路的平面垂直。

考虑与以速度 v 在半径为 r 的圆内运动的电荷 Q 有关的磁 (偶极) 子。电流是单位时间内的电荷流量。圆的周长为 $2\pi r$ ，旋转一周的时间为 $2\pi r/v$ 。因此，电流为 $I = Qv/2\pi r$ 。因此，磁偶极矩 \mathbf{m} 的大小为

$$|\mathbf{m}| = IA = (Qv/2\pi r) \pi r^2 = Qvr/2 = Qrp/2m \quad (6.126)$$

其中 m 是带电粒子的质量， p 为其动量。由于径向矢量 \mathbf{r} 与 \mathbf{p} 垂直，我们有

$$\mathbf{m}_L = \frac{Q\mathbf{r} \times \mathbf{p}}{2m} = \frac{Q}{2m} \mathbf{L} \quad (6.127)$$

其中我们用到了轨道角动量 \mathbf{L} 的定义， \mathbf{m} 的下标表示它来源于粒子的轨道运动。虽然我们用圆周运动的特殊情况得出了 (6.127)，但其有效性是普遍的。对于一个电子， $Q = -e$ ，因其轨道运动而产生的磁矩为

$$\mathbf{m}_L = -\frac{e}{2m_e} \mathbf{L} \quad (6.128)$$

矢量 \mathbf{L} 的大小由 (5.95) 给出，而具有轨道角动量子数 l 的电子的轨道磁矩大小为

$$|\mathbf{m}_L| = \frac{e\hbar}{2m_e} [l(l+1)]^{1/2} = \mu_B [l(l+1)]^{1/2} \quad (6.129)$$

常量 $e\hbar/2m_e$ 称为**玻尔磁子** (Bohr magneton)，记作 μ_B ：

$$\mu_B \equiv e\hbar/2m_e = 9.2740 \times 10^{-24} \text{ J/T} \quad (6.130)$$

现在我们来考虑给氢原子施加外部磁场。磁偶极子 \mathbf{m} 与外部磁场 \mathbf{B} 之间的相互作用能量可表示为

$$E_B = -\mathbf{m} \cdot \mathbf{B} \quad (6.131)$$

使用式 (6.128)，我们有

$$E_B = \frac{e}{2m_e} \mathbf{L} \cdot \mathbf{B} \quad (6.132)$$

我们沿外加磁场的方向确定 z 轴的位置： $\mathbf{B} = B\mathbf{k}$ ，其中 \mathbf{k} 是 z 方向的单位矢量。我们有

$$E_B = \frac{e}{2m_e} B (L_x \mathbf{i} + L_y \mathbf{j} + L_z \mathbf{k}) \cdot \mathbf{k} = \frac{e}{2m_e} B L_z = \frac{\mu_B}{\hbar} B L_z$$

其中 L_z 是轨道角动量在 z 方向上的分量。我们现在将 L_z 替换为算符 \hat{L}_z ，在哈密顿算符中得到以下由外部磁场产生的附加项：

$$\hat{H}_B = \mu_B B \hbar^{-1} \hat{L}_z \quad (6.133)$$

外加磁场中氢原子的薛定谔方程为

$$(\hat{H} + \hat{H}_B) \psi = E \psi \quad (6.134)$$

其中， \hat{H} 是在无外场作用下氢原子的哈密顿算符。我们很容易验证公式 (6.134) 的解就是复类氢波函数 (6.61)：

$$(\hat{H} + \hat{H}_B) R(r) Y_l^m(\theta, \phi) = \hat{H} R Y_l^m + \mu_B \hbar^{-1} B \hat{L}_z R Y_l^m = \left(-\frac{Z^2}{n^2} \frac{e^2}{8\pi\epsilon_0 a} + \mu_B B m \right) R Y_l^m \quad (6.135)$$

其中用到了式 (6.94) 和 (5.105)。因此，能量中存在一个附加项，而外部磁场消除了 m 简并。由于显而易见的原因， m 通常被称为磁量子数 (magnetic quantum number)。实际上，由于电子自旋磁矩的存在，观测到的能量偏移与公式 (6.135) 的预测值并不一致（第 10 章和第 11.7 节。）

在第 5 章中，我们发现在量子力学中， \mathbf{L} 位于一个圆锥面上。对 \mathbf{L} 在外加磁场中运动的经典力学处理表明：场对 \mathbf{m}_L 产生力矩，使 \mathbf{L} 以 $|\mathbf{m}_L| B/2\pi |L|$ 给定的恒定频率绕 \mathbf{B} 的方向旋转，同时与 \mathbf{B} 保持恒定角度。这种脱落运动称为进动 (precession)。在量子力学中，完整指定 \mathbf{L} 是不可能的。然而，我们会发现 $\langle \mathbf{L} \rangle$ 绕着场的方向进动。(Dicke and Wittke, Section 12-3)

6.9 径向薛定谔方程的数值解法

对于单粒子中心力问题，波函数 $\psi = R(r) Y_l^m(\theta, \phi)$ 由 (6.16) 给出，径向因子 $R(r)$ 通过求解径向方程 (6.17) 得到。第 4.4 节中的 Numerov 方法适用于形式为 $\psi'' = G(x) \psi(x)$ 的微分方程 [式 (4.66)]，为了应用该方法，我们需要消去 (6.17) 中的一阶导数 R' 。让我们定义 $F(r) \equiv rR(r)$ ，则

$$R(r) = r^{-1} F(r) \quad (6.136)$$

因此， $R' = -r^{-2} F + r^{-1} F'$ ， $R'' = -2r^{-3} F + 3r^{-2} F' - r^{-1} F''$ 。将这些表达式代入 (6.17)，将径向方程变换为

$$-\frac{\hbar^2}{2m} F''(r) + \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right] F(r) = E F(r) \quad (6.137)$$

$$F''(r) = G(r) F(r), \quad \text{其中 } G(r) \equiv \frac{m}{\hbar^2} (2V - 2E) + \frac{l(l+1)}{r^2} \quad (6.138)$$

其形式符合了 Numerov 方法的要求。为了使用数值法解 (6.137)，我们需要分别处理每个 l 的值。方程 (6.137) 类似于二维薛定谔方程—— $(\hbar^2/2m) \psi''(x) + V(x) \psi(x) = E \psi(x)$ ，只不过用 r (范围是 $0 \rightarrow \infty$) 代替了 x (范围是 $-\infty$ 到 $+\infty$)， $F(r) \equiv rR(r)$ 代替了 ψ ， $V(r) + l(l+1)\hbar^2/2mr^2$ 代替了 $V(x)$ 。我们可以预期，对于每个 l 值，能量最低的解将有 0 个内部节点 (即 $0 < r < \infty$ 的节点)，下一个最低的解将有 1 个内部节点，依此类推。

回顾式 (6.81) 之后的讨论, 若 $R(r)$ 在原点附近表现与 $1/r^b$ 相似, 那么如果 $b > 1$, 则 $R(r)$ 不是平方可积的; 值 $b = 1$ 是不允许的, 如 (6.83) 之后所述。因此, 当 $r = 0$ 时, 必须有 $F(r) \equiv rR(r)$ 为零。

对于 $l \neq 0$ 的情况, (6.138) 中的 $G(r)$ 在 $r = 0$ 处发散, 这会让大多数电脑感到不爽。为了避免这个问题, 我们可以从极小的 r 值开始求解 (例如, 10^{-15} 为无量纲的 r_r) 并近似 $F(r)$ 在该点为零。

例如, 我们将使用 Numerov 方法求解最低束缚态 H 原子能量。有 $V = -e^2/4\pi\epsilon_0 r = -e'^2/r$, 其中 $e' \equiv e/(4\pi\epsilon_0\hbar^2/m)^{1/2}$ 。径向方程 (6.62) 包含三个常数 e' 、 μ 和 \hbar , 其中 $e' \equiv e/(4\pi\epsilon_0\hbar^2/m)^{1/2}$ 具有国际单位 $\text{m N}^{1/2}$ (见附录的表 A.1), 因此, 有量纲的 $[e'] = \text{L}^{3/2}\text{M}^{1/2}\text{T}^{-1}$ 。按照公式 (4.73) 的推导过程, 我们发现 H 原子的约化能量和约化径向坐标为 (问题 6.47)

$$E_r = E/\mu e'^4 \hbar^{-2}, \quad r_r = r/B = r/\hbar^2 \mu^{-1} e'^{-2} \quad (6.139)$$

使用 (6.139)、(4.76) 和 (4.77) 将 ψ 替换为 F , 使用 $B = \hbar^2 \mu^{-1} e'^{-2}$ 将氢原子的式 (6.137) 变换为 (问题 6.47)

$$F_r'' = G_r F_r, \quad \text{其中 } G_r = l(l+1)/r_r^2 - 2/r_r - 2E_r \quad (6.140)$$

其中 $F_r = F/B^{-1/2}$ 。

氢原子束缚态能量均小于零。假设我们想知道 $E_r \leq -0.04$ 的氢原子束缚态本征值。令该能量等于 V_r , 我们有 (问题 6.47) $-0.04 = -1/r_r$, 而这个能量值的经典允许区域从 $r_r = 0$ 一直延申到 $r_r = 25$ 。进入经典禁区 2 个单位长度, 我们有 $r_{r,\max} = 27$ 并要求 $F_r(27) = 0$ 。我们将 s_r 设为 0.1, 将 0 到 27 均分为 270 个点 (更准确地说, 是 10^{-15} 到 $27 + 10^{-15}$)。

(6.140) 中的 G_r 包含参数 l , 所以必须修改表 4.1 中的程序, 以输入 l 的值。设置电子表格时, 请在某个单元格中输入 l 值, 并在输入定义 G_r 的单元格 B7 的公式时参考该单元格 (图 4.9 4.8)。A 列的起始位置为 $r_r = 1 \times 10^{-15}$ 。电子表格的 C 列将包含 F_r 值, 而不是 ψ_r 值, F_r 在 $r_r = 1 \times 10^{-15}$ 处与零相差很小, 可以忽略不计, 因此在该点将取为零。

根据上述选择, 我们可以找到 (问题 6.48a) 的最低三个 $l = 0$ 的 H 原子本征值为 $E_r = -4970, -0.1246, -0.0205499$, 最低两个 $l = 1$ 本征值为 -0.1250 和 -0.05526 。真实值 [式 (6.94) 和 (6.139)] 为 $-0.5000, -0.1250, -0.05555$ 。精度一般的主要原因是 $G(r)$ 在 $r = 0$ 附近的快速变化。如果 s_r 取为 0.025 而不是 1 (得到 1080 个点), $l = 0$ 的本征值将精确到 $-0.4998, -0.12497$ 和 -0.05510 。另见问题 6.48b。

总结

习题

Chapter 7

量子力学理论

7.1 符号

单电子原子的薛定谔方程（第6章）是可以精确求解的。然而，由于哈密顿算符中的电子间斥力项，多电子原子和分子的薛定谔方程无法精确求解。因此，我们必须寻求近似的求解方法。第8章和第9章将介绍两种主要的近似方法，即变分法和微扰理论。要推导出这些方法，我们必须进一步发展量子力学理论，这就是本章所要做的。

在开始之前，我们先介绍一些符号。夹在两个函数之间的算符在全空间上的定积分经常出现，并使用各种缩写：

$$\int f_m^* \hat{A} f_n d\tau \equiv \langle f_m | \hat{A} | f_n \rangle \equiv \langle m | \hat{A} | n \rangle \quad (7.1)$$

其中 f_m 和 f_n 是两个函数。如式 (7.1) 所示，若两个函数的含义是明确的，我们就可以只用下标来区分。狄拉克引入的上述符号称为**括号符号**（bracket notation）（此处为直译，多用狄拉克符号，下同）。另一种符号是

$$\int f_m^* \hat{A} f_n d\tau \equiv A_{mn} \quad (7.2)$$

符号 A_{mn} 和 $\langle m | \hat{A} | n \rangle$ 都表示我们使用字母在前的复函数的复共轭。定积分 $\langle m | \hat{A} | n \rangle$ 称为算符 \hat{A} 的一个**矩阵元**（matrix element）。矩阵是数字的矩形数组，遵守一定的组合规则（见第7.10节）。

我们将对两个函数全空间的定积分写作

$$\int f_m^* f_n d\tau \equiv \langle f_m | f_n \rangle \equiv \langle m | n \rangle \quad (7.3)$$

注意：

$$\langle f | \hat{B} | g \rangle = \langle f | \hat{B} g \rangle$$

其中 f 和 g 是函数。由于 $(\int f_m^* f_n d\tau)^* = \int f_n^* f_m d\tau$ ，我们可以得到

$$\boxed{\langle m | n \rangle^* = \langle n | m \rangle} \quad (7.4)$$

由于取了 (7.1) 中 f_m 的复共轭，因此可以得出

$$\boxed{\langle cf | \hat{B} | g \rangle = c^* \langle f | \hat{B} | g \rangle \quad \text{和} \quad \langle f | \hat{B} | cg \rangle = c \langle f | \hat{B} | g \rangle} \quad (7.5)$$

其中 \hat{B} 是线性算符， c 是常数。

7.2 厄米算符

代表物理量的量子力学算符都是线性算符（第3.1节）。这些算符必须满足一个额外的要求，我们现在就来讨论这个问题。

厄米算符的定义

设 \hat{A} 是对应某个物理量 A 的线性算符。则 A 的平均期望为 [式 (3.88)]

$$\langle A \rangle = \int \Psi^* \hat{A} \Psi d\tau$$

其中 Ψ 是系统的态函数。由于物理量的平均值一定是一个实数，我们要求

$$\begin{aligned} \langle A \rangle &= \langle A \rangle^* \\ \int \Psi^* \hat{A} \Psi d\tau &= \left[\int \Psi^* \hat{A} \Psi d\tau \right]^* = \int (\Psi^*)^* (\hat{A} \Psi)^* d\tau \\ \int \Psi^* \hat{A} \Psi d\tau &= \int \Psi (\hat{A} \Psi)^* d\tau \end{aligned} \quad (7.6)$$

式 (7.6) 必须对每个代表系统可能状态的波函数 Ψ 都成立，也就是说，所有品优函数 Ψ 都必须满足式 (7.6)。对所有品优函数都满足式 (7.6) 的算符称为**厄米算符** (Hermitian operator) (由数学家厄米的名字命名——Charles Hermite)。

许多教科书将厄米算符定义为对所有品优函数 f 和 g 都满足下式的线性算符：

$$\int f^* \hat{A} g d\tau = \int g (\hat{A} f)^* d\tau \quad (7.7)$$

注意：在 (7.7) 的左侧，算符 \hat{A} 作用在函数 g 上；而在右侧，算符 \hat{A} 作用在函数 f 上。对于特例 $f = g$ ，式 (7.7) 退化为式 (7.6)。等式 (7.7) 显然是比 (7.6) 更强的要求，但我们将证明 (7.7) 是 (7.6) 的结果。因此，两种厄米算符的定义是等价的。

我们首先在 (7.6) 中令 $\Psi = f + cg$ ，其中 c 是一个任意常数。则

$$\begin{aligned} \int (f + cg)^* \hat{A} (f + cg) d\tau &= \int (f + cg) [\hat{A} (f + cg)]^* d\tau \\ &= \int (f^* + c^* g^*) \hat{A} f d\tau + \int (f^* + c^* g^*) \hat{A} c g d\tau \\ &= \int (f + cg) (\hat{A} f)^* d\tau + \int (f + cg) (\hat{A} c g)^* d\tau \\ &= \int f^* \hat{A} f d\tau + c^* \int g^* \hat{A} f d\tau + c \int f^* \hat{A} g d\tau + c^* c \int g^* \hat{A} g d\tau \\ &= \int f (\hat{A} f)^* d\tau + c \int g (\hat{A} f)^* d\tau + c^* \int f (\hat{A} g)^* d\tau + c c^* \int g (\hat{A} g)^* d\tau \end{aligned}$$

根据 (7.6)，最后一个等式两边的首项相等；同样，两边的尾项也相等。因此，

$$c^* \int g^* \hat{A} f d\tau + c \int f^* \hat{A} g d\tau = c \int g (\hat{A} f)^* d\tau + c^* \int f (\hat{A} g)^* d\tau \quad (7.8)$$

在 (7.8) 中令 $c = 1$ ，我们有

$$\int g^* \hat{A} f d\tau + \int f^* \hat{A} g d\tau = \int g (\hat{A} f)^* d\tau + \int f (\hat{A} g)^* d\tau \quad (7.9)$$

在 (7.8) 中令 $c = i$, 两边同除以 i , 我们有

$$-\int g^* \hat{A} f d\tau + \int f^* \hat{A} g d\tau = \int g (\hat{A} f)^* d\tau - \int f (\hat{A} g)^* d\tau \quad (7.10)$$

将 (7.9) 和 (7.10) 相加, 我们得到了 (7.7)。结论得到了证明。

因此, 厄米算符 \hat{A} 定义为满足如下条件的线性算符:

$$\boxed{\int f_m^* \hat{A} f_n d\tau = \int f_n (\hat{A} f_m)^* d\tau} \quad (7.11)$$

其中 f_m 和 f_n 是任意的品优函数, 积分为对全空间的定积分。使用狄拉克符号和矩阵元符号, 我们将其记作

$$\boxed{\langle f_m | \hat{A} | f_n \rangle = \langle f_n | \hat{A} | f_m \rangle^*} \quad (7.12)$$

$$\boxed{\langle m | \hat{A} | n \rangle = \langle n | \hat{A} | m \rangle^*} \quad (7.13)$$

$$\boxed{A_{mn} = (A_{nm})^*} \quad (7.14)$$

(7.12) 的两边不同之处在于交换了函数并取了复共轭。

厄米算符的例子

让我们证明, 我们一直在使用的一些算符确实是厄米的。简单起见, 我们将从一个维度进行研究。要证明一个算符是厄米的, 只需证明它对所有品优函数都满足式 (7.6) 即可。不过, 我们将通过证明满足 (7.11) 来增加一点难度。

首先, 来考虑单粒子一维势能算符。式 (7.11) 的右侧为

$$\int_{-\infty}^{\infty} f_n(x) [V(x) f_m(x)]^* dx \quad (7.15)$$

由于势能是实函数, 我们有 $V^* = V$ 。此外, 式 (7.15) 中因子的顺序无关紧要。因此,

$$\int_{-\infty}^{\infty} f_n (V f_m)^* dx = \int_{-\infty}^{\infty} f_n V^* f_m^* dx = \int_{-\infty}^{\infty} f_m^* V f_n dx$$

这就证明了 V 是厄米的。

再考虑动量在 x 方向分量的算符 $\hat{p}_x = -i\hbar d/dx$ [式 (3.23)]。对于这个算符, 式 (7.11) 的左侧为

$$-i\hbar \int_{-\infty}^{\infty} f_m^*(x) \frac{df_n(x)}{dx} dx$$

现在我们使用分部积分法:

$$\int_a^b u(x) \frac{dv(x)}{dx} dx = u(x) v(x) \Big|_a^b - \int_a^b v(x) \frac{du(x)}{dx} dx \quad (7.16)$$

令

$$u(x) \equiv -i\hbar f_m^*(x), \quad v(x) \equiv f_n(x)$$

那么

$$-i\hbar f_m^* \frac{df_n}{dx} dx = -i\hbar f_m^* f_n \Big|_{-\infty}^{\infty} + i\hbar \int_{-\infty}^{\infty} f_n(x) \frac{df_m^*(x)}{dx} dx \quad (7.17)$$

由于 f_m 和 f_n 是品优函数，因此在无穷远处它们的值为（如果它们在无穷远处不为零，则不符合平方可积的要求）。于是，(7.17) 变为

$$\int_{-\infty}^{\infty} f_m^* \left(-i\hbar \frac{df_n}{dx} \right) dx = \int_{-\infty}^{\infty} f_n \left(-i\hbar \frac{df_m}{dx} \right)^* dx$$

与 (7.11) 一致，证明了 \hat{p}_x 是厄米的。证明动能算符是厄米的留给读者自行考虑。可以证明，两个厄米算符的和也是厄米的。因此，哈密顿算符 $\hat{H} = \hat{T} + \hat{V}$ 也是厄米的。

厄米算符定理

现在我们证明一些有关厄米算符的本征值和本征函数的重要定理。

由于与物理量 A 相对应的算符 \hat{A} 的本征值是测量 A 的可能结果（第 3.3 节），因此这些特征值都应该是实数。我们现在证明厄米算符的本征值是实数。

我们现在知道 \hat{A} 是厄米算符。将这个条件转换为方程，对于所有品优函数 f_m 和 f_n ，我们有 [式 (7.11)]

$$\int f_m^* \hat{A} f_n d\tau = \int f_n (\hat{A} f_m)^* d\tau \quad (7.18)$$

我们希望证明 \hat{A} 的每个本征值都是实数。将其转换为方程，我们希望证明 $a_i = a_i^*$ ，其中本征值 a_i 满足 $\hat{A} g_i = a_i g_i$ ，其中 g_i 是 \hat{A} 的本征函数。

为了将本征值 a_i 引入式 (7.18)，我们将 (7.18) 写成一个特例，其中 $f_m = g_i$ 和 $f_n = g_i$ ：

$$\int g_i^* \hat{A} g_i d\tau = \int g_i (\hat{A} g_i)^* d\tau$$

使用 $\hat{A} g_i = a_i g_i$ ，我们有

$$\begin{aligned} a_i \int g_i^* g_i d\tau &= \int g_i (a_i g_i)^* d\tau = a_i^* \int g_i g_i^* d\tau \\ (a_i - a_i^*) \int |g_i|^2 d\tau &= 0 \end{aligned} \quad (7.19)$$

由于对 $|g_i|^2$ 的积分永远不会为负，因此 (7.19) 中积分为零的唯一可能情况是 g_i 在所有坐标值下都为零。然而，基于物理原因，我们总是拒绝将 $g_i = 0$ 作为本征函数。因此，式 (7.19) 中的积分不能为零。那么只有 $(a_i - a_i^*) = 0$ ，即 $a_i = a_i^*$ 。得证。

定理 1： 厄米算符的本征值为实数。

为了帮助读者熟悉狄拉克符号，我们将使用狄拉克符号重新证明定理 1。在式 (7.13) 中，令 $m = i$ 以及 $n = i$ ，我们得到了 $\langle i | \hat{A} | i \rangle = \langle i | \hat{A} | i \rangle^*$ 。选择下标为 i 的函数作为 \hat{A} 的本征函数，使用本征方程 $\hat{A} g_i = a_i g_i$ ，我们有 $\langle i | a_i | i \rangle = \langle i | a_i | i \rangle^*$ 。因此， $a_i \langle i | i \rangle = a_i^* \langle i | i \rangle^* = a_i^* \langle i | i \rangle$ ，则 $(a_i - a_i^*) \langle i | i \rangle = 0$ 。所以 $a_i = a_i^*$ ，其中我们用到了 (7.4) 中 $m = n$ 。

我们已经证明：两个不同的箱中粒子能量本征函数 ψ_i 和 ψ_j 是正交的，也就是说对 $i \neq j$ ，有 $\int_{-\infty}^{\infty} \psi_i^* \psi_j dx = 0$ [式 (2.26)]。若两个坐标相同的函数 f_1 和 f_2 满足下式。则它们是正交的 (orthogonal)：

$$\boxed{\int f_1^* f_2 d\tau = 0} \quad (7.20)$$

其中的积分是对坐标的全空间进行定积分。我们现在证明更一般的结论：厄米算符的本征函数是或可以被选择为相互正交的。有

$$\hat{B}F = sF, \quad \hat{B}G = tG \quad (7.21)$$

其中 F 和 G 是厄米算符 \hat{B} 的两个线性无关的本征函数，我们希望证明

$$\int F^* G \, d\tau \equiv \langle F|G \rangle = 0$$

我们从 (7.12) 开始，它表达了 \hat{B} 的厄米性质：

$$\langle F|\hat{B}|G \rangle = \langle G|\hat{B}|F \rangle^*$$

使用式 (7.21)，我们有

$$\begin{aligned} \langle F|t|G \rangle &= \langle G|s|F \rangle^* \\ t \langle F|G \rangle &= s^* \langle G|F \rangle \end{aligned}$$

由于厄米算符的本征值为实数（定理 1），我们有 $s^* = s$ 。使用 $\langle G|F \rangle = \langle F|G \rangle^*$ （式 (7.4)），我们得到

$$\begin{aligned} t \langle F|G \rangle &= s \langle F|G \rangle \\ (t - s) \langle F|G \rangle &= 0 \end{aligned}$$

若 $s \neq t$ ，则

$$\langle F|G \rangle = 0 \quad (7.22)$$

我们已经证明，对应于不同本征值的厄米算符的两个本征函数是正交的。现在的问题是：我们能否有两个具有相同本征值的独立本征函数？答案是肯定的。在简并的情况下，我们有一个以上的本征函数有相同的本征值。因此，只有当厄米算符的两个独立本征函数不对应于一个简并本征值时，我们才能确定它们是正交的。现在我们要证明，在简并的情况下，我们可以构造出彼此正交的本征函数。我们将使用第 3.6 节中证明的定理，即任何与简并本征值对应的本征函数线性组合都是具有相同本征值的本征函数。因此，让我们假设 F 和 G 是具有相同本征值的独立本征函数：

$$\hat{B}F = sF, \quad \hat{B}G = sG$$

我们对 F 和 G 进行线性组合，形成两个新的本征函数 g_1 和 g_2 ，它们将相互正交。我们选择

$$g_1 \equiv F, \quad g_2 \equiv G + cF$$

其中常数 c 的选择是为了确保正交性。我们希望

$$\begin{aligned} \int g_1^* g_2 \, d\tau &= 0 \\ \int F^* (G + cF) \, d\tau &= \int F^* G \, d\tau + c \int F^* F \, d\tau = 0 \end{aligned}$$

因此，我们选择

$$c = - \int F^* G \, d\tau / \int F^* F \, d\tau \quad (7.23)$$

我们有两个与简并本征值相对应的正交本征函数 g_1 和 g_2 。这个过程 [称为**施密特正交化**或**格拉姆-施密特正交化** (Schmidt orthogonalization or Gram-Schmidt orthogonalization)] 可以推广到 n 重简并的情况，从而得到与简并本征值相对应的 n 个线性无关的正交本征函数。

因此，虽然不能保证简并本征值的本征函数是正交的，但如果我们愿意，总是可以通过施密特（或其他）正交化方法来选择它们是正交的。事实上，除非另有说明，我们将始终假定本征函数是正交的：

$$\int g_i^* g_k \, d\tau = 0, \quad i \neq k \quad (7.24)$$

其中 g_i 和 g_k 是厄米算符的独立本征函数。我们证明了：

定理 2: 厄米算符 \hat{B} 的两个对应不同本征值的本征函数是正交的。属于简并本征值的 \hat{B} 的本征函数总是可以选择为正交的。

本征函数通常可以乘以一个常数来归一化，除非另有说明，我们将假定所有本征函数都归一化了：

$$\int g_i^* g_i d\tau = 1 \quad (7.25)$$

例外情况是本征值形成一个连续体，而不是一组离散的值。在这种情况下，本征函数不满足平方可积性。例子有动量本征函数、自由粒子能量本征函数和氢原子连续能量本征函数。

使用 Kronecker 三角，定义为 $\delta_{jk} \equiv 1$ 当 $j = k$ 时， $\delta_{jk} \equiv 0$ 当 $j \neq k$ [式 (2.28)]，我们可以将 (7.24) 和 (7.25) 组合成一个方程：

$$\int g_i^* g_k d\tau = \langle i|k \rangle = \delta_{ik} \quad (7.26)$$

其中 g_i 和 g_k 是某厄米算符的本征函数。

例如，考虑球谐函数。我们希望证明

$$\int_0^{2\pi} \int_0^\pi [Y_l^m(\theta, \phi)]^* Y_{l'}^{m'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{l,l'} \delta_{m,m'} \quad (7.27)$$

其中 θ 因子来自球面坐标中的体积元素，(5.78)。球谐函数是厄米算符 \hat{L}^2 的本征函数 [式 (5.104)]。由于属于不同本征值的厄米算符的本征函数是正交的，我们得出：除非 $l = l'$ ，否则 (7.27) 中的积分为零。同样地，由于 Y_l^m 系列函数是 \hat{L}_z 的本征函数 [式 (5.105)]，因此除非 $m = m'$ ，否则 (7.27) 中的积分为零。此外， Y_l^m 中的乘常数 [问题 5.34 的式 (5.147)] 可以选择使得球谐函数得到归一化 [式 (6.117)]。因此，(7.27) 成立。

若 f 或 g 是某厄米算符的本征函数，则积分 $\langle f|\hat{B}|g \rangle$ 可以得到简化。若 $\hat{B}g = cg$ ，其中 c 是常数，则

$$\langle f|\hat{B}|g \rangle = \langle f|\hat{B}g \rangle = \langle f|cg \rangle = c \langle f|g \rangle$$

若 $\hat{B}f = kf$ ，其中 k 是常数，那么利用 \hat{B} 的厄米特性就可以得到

$$\langle f|\hat{B}|g \rangle = \langle g|\hat{B}|f \rangle^* = \langle g|\hat{B}f \rangle^* = \langle g|kf \rangle^* = k^* \langle g|f \rangle^* = k \langle f|g \rangle$$

其中用到了本征值 k 是实数。关系 $\langle f|\hat{B}|g \rangle = k \langle f|g \rangle$ 表明厄米算符 \hat{B} 可以在 $\langle f|\hat{B}|g \rangle$ 向左作用。

不确定性原理的证明见问题 7.60。

7.3 用本征函数展开

在上一节中，我们证明了厄米算符本征函数的正交性。现在我们讨论这些函数的另一个重要性质；这一性质使我们可以用这些本征函数来展开一个任意的品优函数。

我们使用了 Taylor 级数展开 (问题 4.1) 来将函数表示为 $(x - a)$ 的非负整数幂的线性组合。除了 $1, (x - a), (x - a)^2, \dots$ 以外，我们还能将函数展开为其他函数集的线性组合吗？答案是肯定的，正如傅里叶 (Fourier) 在 1807 年所发现的那样。Fourier 级数是将一个函数展开为无数个正弦和余弦函数的线性组合。我们将不再详细介绍 Fourier 级数，而只是举例说明。

使用箱中粒子的波函数展开函数

让我们考虑用箱中粒子的定态波函数来展开一个函数，即 [式 (2.23)]

$$\psi_n = \left(\frac{2}{l}\right)^{1/2} \sin\left(\frac{n\pi x}{l}\right), \quad n = 1, 2, 3, \dots \quad (7.28)$$

其中 x 介于 0 和 l 之间。用如下形式的级数表示区间 $0 \leq x \leq l$ 内的任意函数 $f(x)$ 的可能性有多大?

$$f(x) = \sum_{n=1}^{\infty} a_n \psi_n = \left(\frac{2}{l}\right)^{1/2} \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{l}\right), \quad 0 \leq x \leq l \quad (7.29)$$

其中 a_n 等为常数。将 $x=0$ 和 $x=l$ 代入 (7.29)，我们得到限制条件 $f(0)=0$ 和 $f(l)=0$ 。也就是说， $f(x)$ 必须满足和 ψ_n 相同的边界条件。我们同样假设 $f(x)$ 的有限、单值以及连续，但不一定可导。根据上述假设，可以证明展开式 (7.29) 是成立的。我们将不证明 (7.29)，而只是说明它在表示函数时的用法。

在我们对某一特定函数 $f(x)$ 应用式 (7.29) 之前，我们必须推导出展开系数 a_n 的表达式。我们先将 (7.29) 两边乘以 ψ_m^* ：

$$\psi_m^* f(x) = \sum_{n=1}^{\infty} a_n \psi_m^* \psi_n = \left(\frac{2}{l}\right) \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{l}\right) \sin\left(\frac{m\pi x}{l}\right) \quad (7.30)$$

现在我们将该方程从 0 到 l 进行积分。假设将积分和无限求和互换是成立的，我们可以得出

$$\int_0^l \psi_m^* f(x) dx = \sum_{n=1}^{\infty} a_n \int_0^l \psi_m^* \psi_n dx = \sum_{n=1}^{\infty} a_n \left(\frac{2}{l}\right) \int_0^l \sin\left(\frac{n\pi x}{l}\right) \sin\left(\frac{m\pi x}{l}\right) dx$$

我们已经证明了箱中粒子波函数的正交性 [式 (2.27)]。因此，最后一个方程变为

$$\int_0^l \psi_m^* f(x) dx = \sum_{n=1}^{\infty} a_n \delta_{mn} \quad (7.31)$$

(7.31) 中的求和方式经常出现。将其展开，我们得到

$$\begin{aligned} \sum_{n=1}^{\infty} a_n \delta_{mn} &= a_1 \delta_{m,1} + a_2 \delta_{m,2} + \dots + a_{m-1} \delta_{m,m-1} + a_m \delta_{m,m} + a_{m+1} \delta_{m,m+1} + \dots \\ &= 0 + 0 + \dots + a_m + 0 + \dots \\ \sum_{n=1}^{\infty} a_n \delta_{mn} &= a_m \end{aligned} \quad (7.32)$$

因此，除了 $n=m$ 的项以外， δ_{mn} 都为零，那么只有一项是保留的，其他项都消失，式 (7.31) 变为

$$a_m = \int_0^l \psi_m^* f(x) dx \quad (7.33)$$

这正是希望的系数表达式。

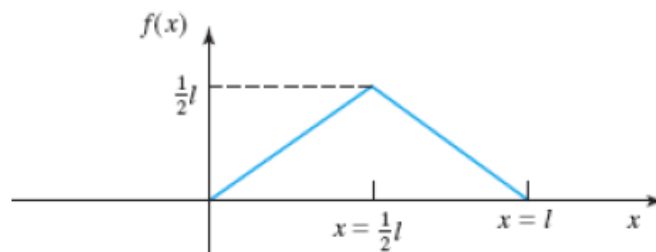


图 7.1: 按箱中粒子函数展开的函数。

在 (7.33) 中将 m 替换为 n 并带入到 (7.29) 中，我们得到

$$f(x) = \sum_{n=1}^{\infty} \left[\int_0^l \psi_n^* f(x) dx \right] \psi_n(x) \quad (7.34)$$

这是任意函数 $f(x)$ 作为箱中粒子波函数 ψ_n 的线性组合的展开所需的表达式。注意：定积分 $\int_0^l \psi_n^* f(x) dx$ 是一个数，而不是一个关于 x 的函数。

我们现在使用 (7.29) 来代表一个特定的函数，图 (7.1) 所示的函数由下式定义：

$$\begin{cases} f(x) = 0, & 0 \leq x < \frac{l}{2} \\ f(x) = 1, & \frac{l}{2} \leq x < l \end{cases} \quad (7.35)$$

为了找到展开系数 a_n ，我们将 (7.35) 和 (7.28) 代入 (7.33) 中：

$$\begin{aligned} a_n &= \int_0^l \psi_n^* f(x) dx = \left(\frac{2}{l}\right)^{1/2} \int_0^l \sin\left(\frac{n\pi x}{l}\right) f(x) dx \\ a_n &= \left(\frac{2}{l}\right)^{1/2} \int_0^{l/2} x \sin\left(\frac{n\pi x}{l}\right) dx + \left(\frac{2}{l}\right)^{1/2} \int_{l/2}^l (l-x) \sin\left(\frac{n\pi x}{l}\right) dx \end{aligned}$$

使用附录中的积分 A.1，我们得到

$$a_n = \frac{(2l)^{3/2}}{n^2 \pi^2} \sin\left(\frac{n\pi}{2}\right) \quad (7.36)$$

将 (7.36) 代入 (7.29)，我们得到 [由于 $\sin(n\pi/2)$ 在 n 为奇数时为 1 或 -1，在 n 为偶数时为 0]：

$$\begin{aligned} f(x) &= \frac{4l}{\pi^2} \left[\sin\left(\frac{\pi x}{l}\right) - \frac{1}{3^2} \sin\left(\frac{3\pi x}{l}\right) + \frac{1}{5^2} \sin\left(\frac{5\pi x}{l}\right) - \dots \right] \\ f(x) &= \frac{4l}{\pi^2} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{(2n-1)^2} \sin\left[(2n-1) \frac{\pi x}{l}\right] \end{aligned} \quad (7.37)$$

其中 $f(x)$ 由 (7.35) 给出。让我们检查 $x = \frac{l}{2}$ 时 (7.37) 的值：

$$f\left(\frac{l}{2}\right) = \frac{4l}{\pi^2} \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots\right) \quad (7.38)$$

将 (7.38) 的右边作为我们在无穷级数中所取项数的函数列表，我们可以得到

项数	1	2	3	4	5	20	100
式 (7.38) 的右侧	0.405l	0.450l	0.467l	0.475l	0.480l	0.495l	0.499l

如果我们取无穷多项，那么级数的和应为 $\frac{1}{2}l$ ，与 $f(\frac{l}{2})$ 的值一致。假设级数成立，我们会得到一个有趣的结果，即 (7.38) 中右侧括号的求和等于 $\pi^2/8$ 。图 7.2 显示了 $f(x) - \sum_{n=1}^k a_n \psi_n$ 的图形 [其中 f 、 a_n 和 ψ_n 的定义见 (7.35)、(7.36) 和 (7.28)]，其中 k 是取的项数，分别为 1 和 5。随着 k （即展开式中的项数）的增加，级数逐渐趋近于 $f(x)$ ，而 f 与级数之间的差值逐渐趋于零。

用本征函数展开函数

我们已经看到了用一组函数——箱中粒子能量本征函数——来展开一个函数的例子。我们可以使用许多不同的函数集来展开任意函数。如果每个品优函数 f 都服从与 g_i 函数相同的边界条件，并且可以根据以下条件展开为 g_i 的线性组合，则称函数集 $g_1, g_2, \dots, g_i, \dots$ 为**完全集** (complete set)：

$$f = \sum_i a_i g_i \quad (7.39)$$

其中 a_i 是常数。当然，我们很容易理解 f 和 g_i 都是同一组自变量的函数。(7.39) 中的和省略了极限。不言而喻，这个和包含了完整集合的所有成员。根据 Fourier 分析定理（我们尚未证明），可以证明粒子箱中的能量本征函数是一个完全集。

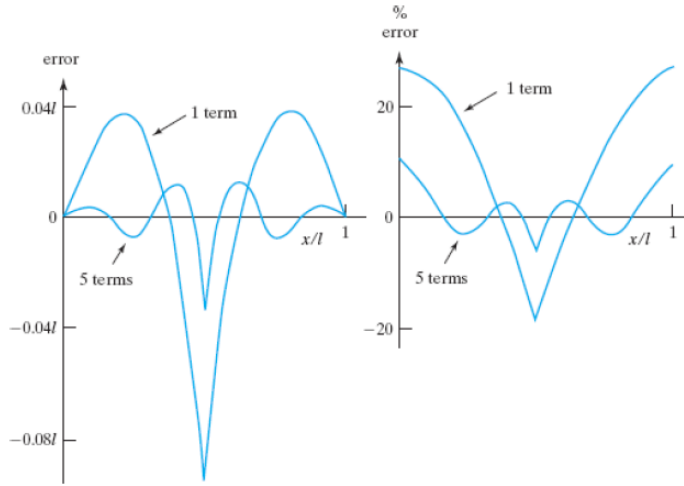


图 7.2:
图 7.1 中以粒子箱中波函数展开的函数, 展开式分别
取 1 项和 5 项时, (a) 误差和 (b) 百分比误差图。

我们现在假设: 表示物理量的每个厄米算符的本征函数集是一个完全集。(本征函数的完备性在很多情况下都可以证明, 但在一般情况下必须假设。)因此, 每个满足与本征函数集合相同边界条件的品优函数都可以根据 (7.39) 展开。方程 (7.29) 就是 (7.39) 的一个例子。

谐振子波函数由厄米多项式 H_n 乘以一个指数因子给出 [问题 4.21c 的式 (4.86)]。根据展开公设, 任何品优函数 $f(x)$ 都可以展开为谐振子能量本征函数的线性组合:

$$f(x) = \sum_{n=0}^{\infty} a_n (2^n n!)^{-1/2} (\alpha/\pi)^{1/4} H_n(\alpha^{1/2} x) e^{-\alpha x^2/2}$$

使用氢原子束缚态波函数来展开任意函数 $f(r, \theta, \phi)$ 如何? 答案是这些函数并不构成一个完全集, 我们无法用它们来展开 f 。要有一个完全集, 我们必须使用特定厄米算符的所有本征函数。除了氢原子哈密顿的束缚态本征函数外, 我们还有与电离态相对应的连续态本征函数。如果将连续本征函数与束缚态本征函数一起包含在内, 那么我们就得到了一个完全集。(对于箱中粒子和谐振子来说, 连续函数不存在。)等式 (7.39) 意味着对连续本征函数 (如果有的话) 进行积分。因此, 若 $\psi_{nlm}(r, \theta, \phi)$ 是氢原子束缚态函数, $\psi_{Elm}(r, \theta, \phi)$ 是连续本征函数, 则式 (7.39) 变为

$$f(r, \theta, \phi) = \sum_{n=1}^{\infty} \sum_{l=0}^{n-1} \sum_{m=-l}^l a_{nlm} \psi_{nlm}(r, \theta, \phi) + \sum_{l=0}^{\infty} \sum_{m=-l}^l \int_0^{\infty} a_{lm}(E) \psi_{Elm}(r, \theta, \phi) dE$$

另一个例子: 考虑算符 \hat{p}_x 的本征函数 [式 (3.36)]:

$$g_k = e^{ikx/\hbar}, \quad -\infty < k < \infty$$

这里的本征值都是连续的, 任意函数 f 的本征函数展开式 (7.39) 变为

$$f(x) = \int_{-\infty}^{\infty} a(k) e^{ikx/\hbar} dk$$

具有良好数学基础的读者可能会发现, 这个积分几乎就是 $a(k)$ 的 Fourier 变换。

让我们对 $f = \sum_i a_i g_i$ [式 (7.39)] 中的展开系数进行计算, 其中的 g_i 函数是厄米算符的本征函数的完全集。这一过程与推导 (7.33) 所用的过程相同。我们将 $f = \sum_i a_i g_i$ 与 g_k^* 相乘并对全空间积分:

$$g_k^* f = \sum_i a_i g_k^* g_i$$

$$\int g_k^* f \, d\tau = \sum_i a_i \int g_k^* g_i \, d\tau = \sum_i a_i \delta_{ik} = a_k$$

$$a_k = \int g_k^* f \, d\tau \quad (7.40)$$

其中我们用到了厄米算符本征函数之间的正交性： $\int g_k^* g_i \, d\tau = \delta_{ik}$ [式 (7.26)]。得出 (7.40) 的过程会经常用到，值得记住。将 (7.40) 中的 a_i 带入 $f = \sum_i a_i g_i$ 中，我们得到

$$f = \sum_i \left[\int g_i^* f \, d\tau \right] g_i = \sum_i \langle g_i | f \rangle g_i \quad (7.41)$$

例题： 令

$$F(x) = \begin{cases} x(l-x), & 0 \leq x \leq l \\ 0, & \text{elsewhere} \end{cases}$$

将 F 用箱中粒子能量本征函数 $\psi_n = (2/l)^{1/2} \sin(n\pi x/l)$, $0 \leq x \leq l$ 展开。

注意到 $F(0) = 0$ 以及 $F(l) = 0$ ，因此满足与 ψ_n 相同的边界条件。我们可以使用 ψ_n 来展开 F 。展开式为 $F = \sum_{n=1}^{\infty} a_n \psi_n$ ，其中系数 $a_n = \int \psi_n^* F \, dx$ [式 (7.40) 和 (7.39)]。因此，

$$a_n = \int \psi_n^* F \, d\tau = \left(\frac{2}{l}\right)^{1/2} \int_0^l \left(\sin \frac{n\pi x}{l}\right) x(l-x) \, dx = \frac{2^{3/2} l^{5/2}}{n^3 \pi^3} [1 - (-1)^n]$$

其中积分计算的具体细节留作习题（问题 7.18）。展开式 $F = \sum_{n=1}^{\infty} a_n \psi_n$ 为

$$x(l-x) = \frac{4l^2}{\pi^3} \sum_{n=1}^{\infty} \frac{1 - (-1)^n}{n^3} \sin\left(\frac{n\pi x}{l}\right), \quad 0 \leq x \leq l$$

练习： 令

$$G(x) = \begin{cases} 1, & 0 \leq x \leq l \\ 0, & \text{elsewhere} \end{cases}$$

将 G 用箱中粒子能量本征函数展开。由于 G 在原点和 l 点不为零，因此展开式在这些点上不表示 G ，而在其他地方表示 G 。利用展开式的前 7 项非零项计算 G 在 $x = \frac{1}{4}l$ 处的值。使用前 70 个非零项重复上述步骤（使用可编程计算器）。（答案：0.1219, 0.9977。）

以下是一条有用的定理：

定理 3： 令 g_1, g_2, \dots 是厄米算符 \hat{A} 的本征函数完全集，令 F 是 \hat{A} 的本征函数，本征值为 k （也就是说， $\hat{A}F = kF$ ）。若 F 展开为 $F = \sum_i a_i g_i$ ，则展开式中的非零系数 a_i 仅对应于与 k 相同的本征值的本征函数 g_i 。（由于简并的存在，可能有多本征函数 g_i 对应于同一本征值 k 。）

因此，在 F 的展开式中，我们只包括那些与 F 具有相同本征值的本征函数。根据 $a_k = \int g_k^* f \, d\tau$ [公式 (7.40)] 即可得出定理 3 的证明；若 F 和 g_i 对应于同一厄米算符 \hat{A} 的不同本征值，那它们是正交的 [式 (7.22)]，则 a_k 消失。

我们偶尔会使用一种符号（称为 **ket** 符号），在这种符号中，函数 f 用以下符号表示 $|f\rangle$ 。这个符号似乎没有任何意义，但在量子力学的高级公式中，它具有特殊的意义。在 ket 符号中，式 (7.41)

可以写成

$$|f\rangle = \sum_i |g_i\rangle \langle g_i|f\rangle = \sum_i |i\rangle \langle i|f\rangle \quad (7.42)$$

Ket 符号方便地用于通过列出本征值来指定本征函数。例如，量子数分别为 n 、 l 和 m 的氢原子波函数可以写成 $\psi_{nlm} = |n, l, m\rangle$ 。

第7.2和7.3节的内容可以总结为厄米算符的本征函数构成一个完整的正交集，且本征值为实数。

7.4 对易算符的本征函数

如果态函数 Ψ 同时是两个算符 \hat{A} 和 \hat{B} 的本征函数，本征值分别为 a_j 和 b_j ，那么对物理量 A 的观测将得到值为 a_j 的结果，对物理量 B 的观测将得到值为 b_j 的结果。因此，当 Ψ 同时是 \hat{A} 和 \hat{B} 的本征函数时，物理量 A 和 B 都具有确定的值。

在第 5.1 节中，我们对两个算符的同时本征函数做了一些说明。现在我们证明这些陈述。

首先，我们证明：如果存在两个线性算符本征函数的共同完全集，那么这两个算符可对易。令 \hat{A} 和 \hat{B} 是两个线性算符， g_1, g_2, \dots 是共同的本征函数完全集：

$$\hat{A}g_i = a_i g_i, \quad \hat{B}g_i = b_i g_i \quad (7.43)$$

其中 a_i 和 b_i 是本征值。我们需要证明

$$[\hat{A}, \hat{B}] = 0 \quad (7.44)$$

方程 (7.44) 是算符方程。要使两个算符相等，则用其中任何一个算符对任意品优函数 f 运算的结果必须相同。因此我们必须证明

$$(\hat{A}\hat{B} - \hat{B}\hat{A})f = \hat{0}f = 0$$

其中 f 是任意函数。首先，我们根据完整的本征函数集 g_i 对 f 进行展开（假设它遵守适当的边界条件）：

$$f = \sum_i c_i g_i$$

将算符 $\hat{A}\hat{B} - \hat{B}\hat{A}$ 作用于上述方程，我们得到

$$(\hat{A}\hat{B} - \hat{B}\hat{A})f = (\hat{A}\hat{B} - \hat{B}\hat{A}) \sum_i c_i g_i$$

由于算符的积 $\hat{A}\hat{B}$ 和 $\hat{B}\hat{A}$ 是线性算符（问题 3.16），我们有

$$(\hat{A}\hat{B} - \hat{B}\hat{A})f = \sum_i c_i (\hat{A}\hat{B} - \hat{B}\hat{A})g_i = \sum_i [\hat{A}(\hat{B}g_i) - \hat{B}(\hat{A}g_i)]$$

其中用到了算符和与积的定义。使用本征值方程 (7.43)，我们得到

$$(\hat{A}\hat{B} - \hat{B}\hat{A})f = \sum_i [\hat{A}(b_i g_i) - \hat{B}(a_i g_i)] = \sum_i c_i (b_i a_i g_i - a_i b_i g_i) = 0$$

这就完成了如下定理的证明：

定理 4： 如果两个线性算符的本征函数完全集是相同的，那么这两个算符可对易。

有时有人错误地认为：如果 \hat{A} 和 \hat{B} 存在一个共同的本征函数，那么它们就可对易。举一个例子可以说明这种说法是错误的，考虑球谐函数 Y_0^0 ，它同时是 \hat{L}_z 和 \hat{L}_x 的本征函数，但显然这两个算符不可对易（第5.3节）。研究一下为这一错误说法提供的所谓证据是很有启发性的。令 g 为共同的本征函数： $\hat{A}g = ag$ 、 $\hat{B}g = bg$ 。我们有

$$\begin{aligned}\hat{A}\hat{B}g &= \hat{A}bg = abg \quad \text{和} \quad \hat{B}\hat{A}g = \hat{B}ag = bag = abg \\ \hat{A}\hat{B}g &= \hat{B}\hat{A}g\end{aligned}\tag{7.45}$$

通过从式 (7.45) 的两边消去 g ，我们完成了所谓的“证明”：

$$\hat{A}\hat{B} = \hat{B}\hat{A} (?) \tag{7.46}$$

错误出现在从 (7.45) 导出 (7.46) 的过程中。 $\hat{A}\hat{B}$ 和 $\hat{B}\hat{A}$ 对单个函数 g 的作用结果是无法得出 $\hat{A}\hat{B} = \hat{B}\hat{A}$ 的。（例如，当 d/dx 和 d^2/dx^2 作用于函数 e^x 时，结果是相同的，但这并不意味着 d/dx 和 d^2/dx^2 是相等的。）这两个算符在作用于每个品优函数时必须得到相同的结果，我们才能得出它们相等的结论。因此，尽管 \hat{A} 和 \hat{B} 不可对易，但可能存在一个或更多的共同本征函数。然而，我们不可能有两个非对易算符的共同完整本征函数集，正如我们在本节前面所证明的那样。

我们已经证明了：若两个算符 \hat{A} 和 \hat{B} 的本征函数完全集是相同的，那么这两个算符可对易。现在我们证明：

定理 5： 如果两个厄米算符 \hat{A} 和 \hat{B} 可对易，我们可以为它们选择一组共同的完整本征函数。

证明如下。令 g_i 和 a_i 分别是算符 \hat{A} 的本征函数和本征值：

$$\hat{A}g_i = a_i g_i$$

将 \hat{B} 作用于上述方程，我们得到

$$\hat{B}\hat{A}g_i = \hat{B}(a_i g_i)$$

由于 \hat{A} 和 \hat{B} 可对易，以及 \hat{B} 是线性算符，我们有

$$\hat{A}(\hat{B}g_i) = a_i (\hat{B}g_i) \tag{7.47}$$

该方程表明：函数 $\hat{B}g_i$ 也是算符 \hat{A} 的本征函数，本征值为 a_i ，与 g_i 一致。假设算符 \hat{A} 的本征值都是非简并的，那么对于任何给定的本征值 a_i ，存在唯一的线性无关的本征函数。如果是这样，那么两个对应于同一个本征值的本征函数 $\hat{B}g_i$ 和 g_i 是线性相关的，也就是说，其中一个一定是另一个的简单乘积：

$$\hat{B}g_i = k_i g_i \tag{7.48}$$

其中 k_i 是常数。这个方程说明了 g_i 是 \hat{B} 的本征函数，因此得证。在第 7.3 节中，我们假设任何表示物理量的算符的本征函数都构成一个完全集。因此， g_i 构成一个完全集。

我们刚才证明了非简并情况下的定理，那么简并情况下呢？设本征值 a_i 是 n 重简并的。由式 (7.47)， $\hat{B}g_i$ 也是 \hat{A} 的本征函数，本征值为 a_i 。因此，第 7.3 节中的定理 3 表明：若函数 $\hat{B}g_i$ 可以用算符 \hat{A} 的本征函数完全集展开，那么除了本征值为 a_i 的项以外，其他项的系数均为零。也就是说， $\hat{B}g_i$ 一定是 n 个本征值为 a_i 的线性无关本征函数的线性组合：

$$\hat{B}g_i = \sum_{k=1}^n c_k g_k, \quad \text{其中} \quad \hat{A}g_k = a_i g_k \quad k = 1, 2, \dots, n \tag{7.49}$$

其中 g_1, \dots, g_n 表示 n 个本征值为 a_i 的算符 \hat{A} 的本征函数。式 (7.49) 表明 g_i 不一定是 \hat{B} 的本征函数。然而, 通过对 n 个对应于简并本征值 a_i 的线性无关的 \hat{A} 本征函数进行适当的线性组合, 我们能构造出一组新的 n 个线性无关的本征函数, 他们也将是 \hat{B} 的本征函数。该说法的证明见 Merzbacher, 第 8.5 节。

因此, 当 \hat{A} 和 \hat{B} 可对易时, 我们可以为它们选择一组共同的本征函数完全集。例如, 考虑一个氢原子, 我们已经证明了算符 \hat{L}_z 和 \hat{H} 是可对易的。如果我们愿意, 我们可以将 \hat{B} 的本征函数中的 ϕ 因子写成 $\sin m\phi$ 和 $\cos m\phi$ 的形式 (第 6.6 节)。如果这样做, 除了 $m = 0$ 的情况以外, 我们就不会得到 \hat{L}_z 的本征函数了。然而, 根据 3.6 节的定理, 以下线性组合

$$R(r) S(\theta) (\cos m\phi + i \sin m\phi) = R S e^{im\phi}, \quad m = -l, \dots, l$$

给出的 \hat{L}_z 的本征函数仍然是 \hat{H} 的本征函数。

将上述证明推广到两个以上算符的情况表明, 当且仅当每个算符都与其他算符可对易时, 对于一组厄米算符 $\hat{A}, \hat{B}, \hat{C}, \dots$, 存在一个共同的完整本征函数集。

与定理 5 相关的一个有用定理为:

定理 6: 若 g_m 和 g_n 是厄米算符 \hat{A} 的本征函数, 但其本征值不同 (即 $\hat{A}g_m = a_m g_m$ 和 $\hat{A}g_n = a_n g_n$, 其中 $a_m \neq a_n$), 若线性算符 \hat{B} 与 \hat{A} 可对易, 那么

$$\langle g_n | \hat{B} | g_m \rangle = 0, \quad a_n \neq a_m \quad (7.50)$$

为了证明 (7.50), 我们从以下式子入手:

$$\langle g_n | \hat{A} \hat{B} | g_m \rangle = \langle g_n | \hat{B} \hat{A} | g_m \rangle \quad (7.51)$$

对上式的左侧使用算符 \hat{A} 的厄米性质, 有

$$\begin{aligned} \langle g_n | \hat{A} \hat{B} | g_m \rangle &= \langle g_n | \hat{A} | \hat{B} g_m \rangle = \langle \hat{B} g_m | \hat{A} | g_n \rangle^* = \langle \hat{B} g_m | a_n | g_n \rangle^* \\ &= a_n^* \langle g_n | \hat{B} g_m \rangle = a_n \langle g_n | \hat{B} | g_m \rangle \end{aligned}$$

式 (7.51) 的右侧可以写成

$$\langle g_n | \hat{B} \hat{A} | g_m \rangle = \langle g_n | \hat{B} | \hat{A} g_m \rangle = a_m \langle g_n | \hat{B} | g_m \rangle$$

令 (7.51) 左右两边的最终表达式相等, 我们得到

$$\begin{aligned} a_n \langle g_n | \hat{B} | g_m \rangle &= a_m \langle g_n | \hat{B} | g_m \rangle \\ (a_n - a_m) \langle g_n | \hat{B} | g_m \rangle &= 0 \end{aligned}$$

由于 $a_n \neq a_m$, 因此 $\langle g_n | \hat{B} | g_m \rangle = 0$, 这就证明了 (7.50)。

7.5 宇称

某些量子力学算符没有相对应的经典类似量。宇称算符就是一个例子。回想一下, 谐振子的波函数要么是奇函数, 要么是偶函数。我们将展示这一性质与宇称算符的关系。

宇称算符 (parity operator) $\hat{\Pi}$ 定义为其对任意函数 f 的以下作用结果:

$$\hat{\Pi} f(x, y, z) = f(-x, -y, -z) \quad (7.52)$$

宇称算符将每个直角坐标替换为其负数。例如： $\hat{\Pi}(x^2 - ze^{ay}) = x^2 + ze^{-ay}$ 。

与其他任意量子力学算符一样，我们对宇称算符的本征函数 g_i 和本征值 c_i 很感兴趣：

$$\hat{\Pi}g_i = c_i g_i \quad (7.53)$$

问题的关键是求出 $\hat{\Pi}$ 的平方：

$$\hat{\Pi}^2 f(x, y, z) = \hat{\Pi} [\hat{\Pi} f(x, y, z)] = \hat{\Pi} f(-x, -y, -z) = f(x, y, z)$$

由于 f 是任意给定的函数，我们得到 $\hat{\Pi}^2$ 等于单位算符：

$$\hat{\Pi}^2 = \hat{1} \quad (7.54)$$

我们现在将 $\hat{\Pi}$ 作用到 (7.53) 的两边，得到 $\hat{\Pi}\hat{\Pi}g_i = \hat{\Pi}c_i g_i$ 。由于 $\hat{\Pi}$ 是线性算符（问题 7.26），我们有 $\hat{\Pi}^2 g_i = c_i \hat{\Pi}g_i$ ，则

$$\hat{\Pi}^2 g_i = c_i^2 g_i \quad (7.55)$$

其中用到了本征方程 (7.53)。由于 $\hat{\Pi}^2$ 是单位算符，式 (7.55) 的左侧是 g_i ，那么

$$g_i = c_i^2 g_i$$

函数 g_i 不可能处处为零（根据物理原理，零总是被拒绝作为本征函数）。因此，两边除以 g_i ，我们得到 $c_i^2 = 1$ ，即

$$c_i = \pm 1 \quad (7.56)$$

算符 $\hat{\Pi}$ 的本征值为 +1 和 -1。请注意，这一推导适用于其平方为单位算符的任何算符。

本征方程 g_i 等于什么？本征方程 (7.53) 为

$$\hat{\Pi}g_i(x, y, z) = \pm g_i(x, y, z)$$

$$g_i(-x, -y, -z) = \pm g_i(x, y, z)$$

若本征值为 +1，则 $g_i(-x, -y, -z) = g_i(x, y, z)$ ，这表明 g_i 是偶函数；若本征值为 -1，则 $g_i(-x, -y, -z) = -g_i(x, y, z)$ ，这表明 g_i 是奇函数。因此，宇称算符 $\hat{\Pi}$ 的本征函数为可能的品优奇函数和品优偶函数。

当宇称算符与哈密顿算符 \hat{H} 可对易时，根据第 7.4 节的定理，我们可以为这两个算符选择一组共同的本征函数完全集。 \hat{H} 的本征函数是定态波函数 ψ_i 。因此，若

$$[\hat{\Pi}, \hat{H}] = 0 \quad (7.57)$$

波函数 ψ_i 可以被选择成为 $\hat{\Pi}$ 的本征函数。我们刚刚证明了 $\hat{\Pi}$ 的本征函数是奇函数或偶函数。因此，当 (7.57) 成立时，每个波函数可以被选择为奇函数或偶函数。让我们看看宇称算符和哈密顿算符在什么情况下可对易。

对于单粒子系统，我们有

$$[\hat{H}, \hat{\Pi}] = [\hat{T}, \hat{\Pi}] + [\hat{V}, \hat{\Pi}] = -\frac{\hbar^2}{2m} \left(\left[\frac{\partial^2}{\partial x^2}, \hat{\Pi} \right] + \left[\frac{\partial^2}{\partial y^2}, \hat{\Pi} \right] + \left[\frac{\partial^2}{\partial z^2}, \hat{\Pi} \right] \right) + [\hat{V}, \hat{\Pi}]$$

由于

$$\begin{aligned} \hat{\Pi} \left[\frac{\partial^2}{\partial x^2} (x, y, z) \right] &= \frac{\partial}{\partial(-x)} \frac{\partial}{\partial(-x)} f(-x, -y, -z) = \frac{\partial^2}{\partial x^2} f(-x, -y, -z) \\ &= \frac{\partial^2}{\partial x^2} \hat{\Pi}(x, y, z) \end{aligned}$$

其中 f 是任意函数, 我们得到

$$\left[\frac{\partial^2}{\partial x^2}, \hat{\Pi} \right] = 0$$

y 和 z 坐标有类似的方程, 则 $[\hat{H}, \hat{\Pi}]$ 变为

$$[\hat{H}, \hat{\Pi}] = [\hat{V}, \hat{\Pi}] \quad (7.58)$$

现在

$$\hat{\Pi}[V(x, y, z)f(x, y, z)] = V(-x, -y, -z)f(-x, -y, -z) \quad (7.59)$$

若势能函数为偶函数, 即 $V(-x, -y, -z) = V(x, y, z)$, 则 (7.59) 变为

$$\hat{\Pi}[V(x, y, z)f(x, y, z)] = V(x, y, z)f(-x, -y, -z) = V(x, y, z)\hat{\Pi}f(x, y, z)$$

所以 $[\hat{V}, \hat{\Pi}] = 0$ 。因此, 当势能函数为偶函数时, 宇称算符和哈密顿算符可对易:

$$[\hat{H}, \hat{\Pi}] = 0 \quad \text{若 } V \text{ 为偶函数} \quad (7.60)$$

这些结果很容易推广到 n 粒子的情况。对于一个 n 粒子系统, 宇称算符定义为

$$\hat{\Pi}f(x_1, y_1, z_1, \dots, x_n, y_n, z_n) = f(-x_1, -y_1, -z_1, \dots, -x_n, -y_n, -z_n) \quad (7.61)$$

很容易发现, 若 V 满足下式

$$V(x_1, y_1, z_1, \dots, x_n, y_n, z_n) = V(-x_1, -y_1, -z_1, \dots, -x_n, -y_n, -z_n) \quad (7.62)$$

则 (7.57) 成立。若 V 满足方程 (7.62), 则称其为 $3n$ 坐标的偶函数 (even function of the $3n$ coordinates)。总而言之, 我们有

定理 7: 当势能 V 是偶函数时, 我们可以选择定态波函数, 使每个波函数要么是偶函数, 要么是奇函数。

要么是偶数要么是奇数的函数被称为具有确定的宇称 (definite parity)。

如果能级都是非简并能级 (在一维问题中通常如此), 那么每个能量本征值只对应一个独立的波函数, 波函数中没有可选择的元素 (除了任意的乘常数)。因此, 对于非简并情况, 当 V 是偶函数时, 定态波函数必须具有确定的宇称。例如, 一维谐振子的势能函数为 $V = \frac{1}{2}kx^2$, 是偶函数, 则波函数具有确定的宇称。

氢原子势能函数是偶函数, 类氢轨道可以选择具有确定的宇称 (问题 7.22 和 7.29)。

对于简并的情况, 我们在波函数中有一个选择元素, 因为与简并能级相对应的函数的任意线性组合是 \hat{H} 的本征函数。对于简并能级, 通过适当的线性组合, 我们可以选择具有确定奇偶性的波函数, 但并不要求它们必须具有确定的奇偶性。

宇称有助于积分的计算。我们已经证明了当 $f(x)$ 为奇函数时有 $\int_{-\infty}^{\infty} f(x) dx = 0$ 。让我们将这一结果推广到 $3n$ 维情况。 $3n$ 个变量的奇函数 (odd function) 满足

$$g(-x_1, -y_1, -z_1, \dots, -x_n, -y_n, -z_n) = -g(x_1, y_1, z_1, \dots, x_n, y_n, z_n) \quad (7.63)$$

若 g 是 $3n$ 个变量的奇函数, 则

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, \dots, z_n) dx_1 \dots dz_n = 0 \quad (7.64)$$

其中的积分是对 $3n$ 个坐标的积分。这个等式成立的原因为: $(x_1, y_1, z_1, \dots, x_n, y_n, z_n)$ 处 g 值对积分的贡献被 $(-x_1, -y_1, -z_1, \dots, -x_n, -y_n, -z_n)$ 处的 g 值抵消了。当积分是部分 (但不一定是全部) 变量的奇函数时, 等式 (7.64) 也成立。参见问题 7.30。

7.6 测量和态叠加

量子力学可被视为一种计算测量各种可能结果概率的方法。例如，如果我们知道态函数 $\Psi(x, t)$ ，那么在时刻 t 时测量粒子的位置位于 x 到 $x + dx$ 的无穷小区间内的概率为 $|\Psi(x, t)|^2 dx$ 。现在，我们考虑对一般性质 B 进行测量。我们的目的是找出如何使用 Ψ 来计算 B 的测量结果中每种可能结果的概率。本节的结果告诉我们状态函数 Ψ 中包含哪些信息，是量子力学的核心所在。

我们处理一个 n 粒子系统，使用 q 来表示 $3n$ 个坐标。我们推测，算符 \hat{B} 的本征值 b_i 是对属性 B 进行测量的唯一可能结果。使用 g_i 来表示 \hat{B} 的本征函数，我们有

$$\hat{B}g_i(q) = b_i g_i(q) \quad (7.65)$$

我们在第 7.3 节中假设，任何代表物理可观测属性的厄米算符的本征函数都构成一个完全集。由于 g_i 构成一个完全集，我们可以将态函数 Ψ 展开为

$$\Psi(q, t) = \sum_i c_i(t) g_i(q) \quad (7.66)$$

为了允许 Ψ 随时间变化，我们在系数 c_i 中引入时间依赖性。

由于 $|\Psi|^2$ 表示概率密度，我们要求

$$\int \Psi^* \Psi d\tau = 1 \quad (7.67)$$

将 (7.66) 代入归一化条件，并使用式 (1.33) 和 (1.32)，我们得到

$$1 = \int \sum_i c_i^* g_i^* \sum_j c_j g_j d\tau = \int \sum_i c_i^* g_i^* \sum_k c_k g_k d\tau = \int \sum_i \sum_k c_i^* c_k g_i^* g_k d\tau \quad (7.68)$$

由于 (7.68) 中两个和的求和下标不一定具有相同的值，因此这两个下标必须使用不同的符号。例如，考虑如下两个求和的积：

$$\sum_{i=1}^2 s_i \sum_{i=1}^2 t_i = (s_1 + s_2)(t_1 + t_2) = s_1 t_1 + s_1 t_2 + s_2 t_1 + s_2 t_2$$

如果我们不小心写成了

$$\sum_{i=1}^2 s_i \sum_{i=1}^2 t_i = (\text{错误的}) \sum_{i=1}^2 \sum_{i=1}^2 s_i t_i = \sum_{i=1}^2 (s_1 t_1 + s_2 t_2) = 2(s_1 t_1 + s_2 t_2)$$

我们就得到了错误的答案。写乘积的正确方式是

$$\sum_{i=1}^2 s_i \sum_{i=1}^2 t_i = \sum_{i=1}^2 s_i \sum_{k=1}^2 t_k = \sum_{i=1}^2 \sum_{k=1}^2 s_i t_k = \sum_{i=1}^2 (s_i t_1 + s_i t_2) = s_1 t_1 + s_1 t_2 + s_2 t_1 + s_2 t_2$$

这就得到了正确的结果。

假定交换 (7.68) 中的无限求和与积分是成立的，我们得到

$$\sum_i \sum_k c_i^* c_k \int g_i^* g_k d\tau = 1$$

由于 \hat{B} 是厄米算符，它的本征函数之间应该是正交的 [式 (7.26)]；因此

$$\sum_i \sum_k c_i^* c_k \delta_{ik} = 1$$

$$\sum_i |c_i|^2 = 1 \quad (7.69)$$

我们很快就会指出 (7.69) 的重要性。

回顾假设 (第 3.7 节), 如果 Ψ 是系统的归一化状态函数, 那么属性 B 的平均值为

$$\langle B \rangle = \int \Psi^*(q, t) \hat{B} \Psi(q, t) d\tau$$

将展开式 (7.66) 代入上式, 我们得到

$$\langle B \rangle = \int \sum_i c_i^* g_i^* \hat{B} \sum_k c_k g_k d\tau = \sum_i \sum_k c_i^* c_k \int g_i^* \hat{B} g_k d\tau$$

其中我们用到了算符 \hat{B} 的线性性质。使用 $\hat{B}g_k = b_k g_k$ [式 (7.65)], 我们有

$$\begin{aligned} \langle B \rangle &= \sum_i \sum_k c_i^* c_k b_k \int g_i^* g_k d\tau = \sum_i \sum_k c_i^* c_k b_k \delta_{ik} \\ \langle B \rangle &= \sum_i |c_i|^2 b_i \end{aligned} \quad (7.70)$$

我们如何解释 (7.70)? 我们在第 3.39 节中假设, 当我们测量算符所代表的物理性质时, 算符的本征值是我们能得到的唯一可能的数字。在对 B 的任何测量中, 我们都会得到其中一个值 b_i (假设没有实验误差)。现在回忆式 (3.81):

$$\boxed{\langle B \rangle = \sum_{b_i} P(b_i) b_i} \quad (7.71)$$

其中 $P(b_i)$ 是对 B 进行测量, 得到测量结果为 b_i 的概率。(7.71) 中的和涉及不同的本征值 b_i , 而 (7.70) 中的和涉及不同的本征函数 g_i , 因为展开式 (7.66) 涉及的是 g_i 。如果每个本征值只有一个独立的本征函数, 那么本征函数的总和就等于本征值的总和, 比较 (7.71) 和 (7.70) 可以看出, 当 \hat{B} 本征值没有简并时, $|c_i|^2$ 就是在测量属性 B 时得到值 b_i 的概率。请注意, 这些值的总和为 1, 因为概率的总和必定为 1 [式 (7.69)]。假设本征值 b_i 是简并的。根据 (7.71), $P(b_i)$ 由与 b_i 相乘的量给出。在简并的情况下, (7.70) 中有不止一项包含 b_i , 所以在测量中获得 b_i 的概率 $P(b_i)$ 是所有本征值为 b_i 的本征函数对应的 $|c_i|^2$ 的总和。我们已经证明了如下定理:

定理 8: 若 b_m 是算符 \hat{B} 的一个非简并本征值, g_m 是对应的归一化本征函数 ($\hat{B}g_m = b_m g_m$), 那么在量子力学系统中, 测量时状态函数为 Ψ 时, 对属性 B 的测量得到结果为 b_m 的概率由 $|c_m|^2$ 给出, 其中 c_m 是 g_m 在展开式 $\Psi = \sum_i c_i g_i$ 中的系数。若 b_m 是简并本征值, 则测量得到结果为 b_m 的概率由所有对应于 b_m 的本征函数的 $|c_i|^2$ 的总和给出。

什么时候可以准确预测 B 的测量结果? 若展开式 $\Psi = \sum_i c_i g_i$ 中的系数 c_i 全为零, 除了 c_i 等于零, 以及对于 $i \neq k$, $c_k \neq 0$ 的情况外, 这是可以做到的。对于这种情况, 式 (7.69) 表明 $|c_k|^2 = 1$, 我们一定会得到值为 b_k 的结果。在该情况中, 态函数 $\Psi = \sum_i c_i g_i$ 简化为 $\Psi = g_k$ 。当 Ψ 是 \hat{B} 的本征函数, 本征值为 b_k 时, 若我们对 B 进行测量, 我们一定会得到值为 b_k 的结果。

因此, 我们可以将展开式 $\Psi = \sum_i c_i g_i$ [式 (7.66)] 看作是将一般状态 Ψ 表达为算符 \hat{B} 本征态 g_i 的态叠加 (superposition)。每个本征态 g_i 对应于物理量 B 的一个可能测量结果 b_i 。任何本征函数 g_i 出现在 Ψ 展开式中的程度由 $|c_i|^2$ 给出, 它决定了在对 B 进行测量时得到结果 b_i 的概率。

为了得到概率 $|c_i|^2$, 我们如何计算出展开系数 c_i ? 我们将 g_j^* 与 $\Psi = \sum_i c_i g_i$ 相乘, 对全空间进行积分并使用厄米算符 \hat{B} 本征函数的正交性:

$$\int g_j^* \Psi d\tau = \sum_i c_i \int g_j^* g_i d\tau = \sum_i c_i \delta_{ij}$$

$$c_j = \int g_j^* \Psi \, d\tau = \langle g_j | \Psi \rangle \quad (7.72)$$

在对性质 B 的测量中, 发现非简并本征值 b_j 的概率为

$$|c_j|^2 = \left| \int g_j^* \Psi \, d\tau \right|^2 = |\langle g_j | \Psi \rangle|^2 \quad (7.73)$$

其中 $\hat{B}g_j = b_j g_j$ 。量 $\langle g_j | \Psi \rangle$ 称为**概率幅** (probability amplitude)。

因此, 如果我们知道由状态函数 Ψ 确定的系统状态, 我们就可以使用 (7.73) 来预测对任何属性 B 进行测量的各种可能结果的概率。确定 \hat{B} 的本征函数 g_j 和本征值 b_j 是一个数学问题。

为了在实验中确定测量 B 时发现 g_j 的概率, 我们需要大量 n 个相同的、无相互作用的系统, 每个系统都处于相同的状态 Ψ , 并测量每个系统中的属性 B 。如果第 n_j 次测量得到的结果为 b_j , 则 $P(b_j) = n_j/n = |\langle g_j | \Psi \rangle|^2$ 。

我们可以对定理 8 的第一部分进行如下重述:

定理 9: 若在某量子力学系统中对性质 B 进行测量, 测量时的状态为 Ψ , 则观测到算符 \hat{B} 的非简并本征值 b_j 的概率为 $|\langle g_j | \Psi \rangle|^2$, 其中 g_j 是已归一化的, 对应于本征值 b_j 的本征函数。

如果归一化函数 g_j 和 Ψ 非常相似, 并且在空间的每个区域都有相似的大小, 那么积分 $\langle g_j | \Psi \rangle = \int g_j^* \Psi \, d\tau$ 的绝对值就会很大。若 g_j 和 Ψ 不相似, 那么在 g_i 很大而 Ψ 很小的区域中 (反之亦然), 乘积 $g_j^* \Psi$ 就会很小, 那么积分 $\int g_j^* \Psi \, d\tau$ 的绝对值也会很小, 即找到 b_j 的概率 $\langle g_j | \Psi \rangle$ 也会很小。

例题: 假设我们对氢原子核外电子的 L_z 分量进行测量, 开始测量时系统的状态为 $2p_x$ 。给出测量的可能结果, 并给出每种结果的概率。使用这些概率计算 $2p_x$ 状态的 $\langle L_z \rangle$ 。

由 (6.118):

$$\Psi = 2p_x = 2^{-1/2} (2p_1) + 2^{-1/2} (2p_{-1})$$

这个方程将 Ψ 展开为 \hat{L}_z 的本征函数的线性组合。其中唯一两个具有非零系数的项为 $2p_1$ 和 $2p_{-1}$, 它们的本征值分别为 \hbar 和 $-\hbar$ 。(回想一下, 1 和 -1 的下标给出了量子数 m , \hat{L}_z 的本征值为 $m\hbar$ 。)使用定理 8, 我们取 Ψ 展开式中系数绝对值的平方来得到概率。因此测量 L_z 时得到 \hbar 的概率为 $|2^{-1/2}|^2 = 0.5$, 得到 $-\hbar$ 的概率也是 0.5。为了从这些概率计算 $\langle L_z \rangle$, 我们使用 (7.71):

$$\langle L_z \rangle = \sum_{L_{z,i}} P(L_{z,i}) L_{z,i} = 0.5\hbar + 0.5(-\hbar) = 0$$

其中 $P(L_{z,i})$ 是测量时得到结果 $L_{z,i}$ 的概率, 除了两项以外, 其他概率都为零。当然, $\langle L_z \rangle$ 也可以用平均值方程 (3.88) 计算: $\langle 2p_x | \hat{L}_z | 2p_x \rangle$ 。

练习: 写出满足如下条件的氢原子波函数: 对能量 E 测量时得到 $n=2$ 能级的概率为 1, 对 L^2 测量时得到 $2\hbar^2$ 的概率为 1, 对 L_z 测量时得到 $-\hbar, 0, \hbar$ 的概率相等。是否只有一种可能的答案? (部分答案: 否。)

例题: 现有一个箱中粒子的系统, 箱子长为 l 。对系统的能量进行一次测量, 测量时粒子处于非定态 $\Psi = 30^{1/2} l^{-5/2} x(l-x)$, $0 \leq x \leq l$ 。给出测量的可能结果, 并给出每种结果的概率。

第 3.9 节的假设 (c) 给出了可能的结果, 即能量算符 \hat{H} 的本征值。箱中粒子哈密顿算符的本征

值为 $E = n^2 h^2 / 8ml^2$ ，其中 $n = 1, 2, \dots$ ，并且它们是非简并的。概率由 Ψ 对 \hat{H} 的本征函数的展开式给出。即 $\Psi = \sum_{n=1}^{\infty} c_n \psi_n$ ，其中 $\psi_n = (2/l)^{1/2} \sin(n\pi x/l)$ 。在 (7.41) 之后的例子中，函数 $x(l-x)$ 是根据箱中粒子能量本征函数展开的。状态函数 $30^{1/2} l^{-5/2} x(l-x)$ 等于 $x(l-x)$ 乘以归一化常数 $30^{1/2} l^{-5/2}$ 。因此，将前面例子中的系数乘以 $30^{1/2} l^{-5/2}$ ，就可以得到展开系数 c_n

$$c_n = \frac{(240)^{1/2}}{n^3 \pi^3} [1 - (-1)^n]$$

观测到能量值 $E_n = n^2 h^2 / 8ml^2$ 的概率 $P(E_n)$ 等于 $|c_n|^2$ ：

$$P(E_n) = \frac{240}{n^6 \pi^6} [1 - (-1)^n]^2 \quad (7.74)$$

前几个可能值的概率为

n	1	2	3	4	5
E_n	$h^2/8ml^2$	$4h^2/8ml^2$	$9h^2/8ml^2$	$16h^2/8ml^2$	$25h^2/8ml^2$
$P(E_n)$	0.99855	0	0.001370	0	0.000064

找到 $n = 1$ 能级的概率很高，这对应于抛物线态函数 $30^{1/2} l^{-5/2} x(l-x)$ 与 $n = 1$ 的箱中粒子波函数 $(2/l)^{1/2} \sin(\pi x/l)$ 非常相似（图7.3）。零概率的能级为 $n = 2, 4, 6, \dots$ ，这对应于以下事实：若原点与箱子的中心重合，则态函数 $\Psi = 30^{1/2} l^{-5/2} x(l-x)$ 是偶函数，而 $n = 2, 4, 6, \dots$ 的箱中粒子波函数是奇函数（图2.3），对 Ψ 的展开式没有贡献。当被积函数为奇函数时积分 $\langle g_n | \Psi \rangle$ 消失。

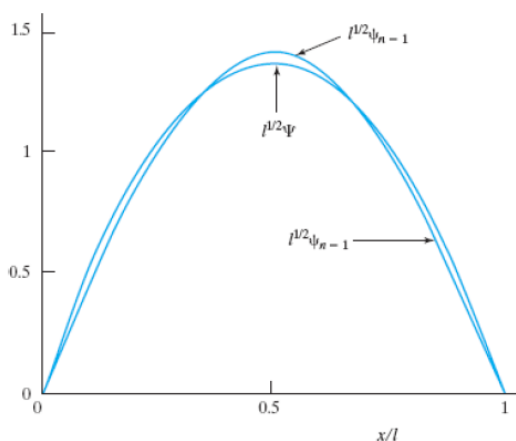


图 7.3: $\Psi = 30^{1/2} l^{-5/2} x(l-x)$ 和 $n = 1$ 箱中粒子波函数的图象

若性质 B 的本征值可以连续变化（例如：位置，第7.7节），则对 Ψ 展开式的求和 (7.66) 变为对 b 所有可能值的积分：

$$\Psi = \int c_b g_b(q) db \quad (7.75)$$

$|\langle g_b(q)|\Psi\rangle|^2$ 依然是概率密度, 也就是说, 在位于 b 到 $b + db$ 的无穷小区间内测量属性 B 时, 得到结果 b 的概率为

$$|\langle g_b(q)|\Psi(q,t)\rangle|^2 db \quad (7.76)$$

7.7 位置本征函数

我们推导出了线动量和角动量算符的本征函数。现在我们想知道: 位置算符 \hat{x} 的本征函数是什么?

算符 \hat{x} 表示对函数乘以 x 。将位置本征函数记作 $g_a(x)$, 写作

$$xg_a(x) = ag_a(x) \quad (7.77)$$

其中的 a 代表可能的本征值。由此可见,

$$(x - a)g_a(x) = 0 \quad (7.78)$$

由 (7.78), 我们得知:

$$g_a(x) = 0, \quad \text{若 } x \neq a \quad (7.79)$$

此外, 由于本征函数处处为零是不可接受的, 我们有

$$g_a(x) \neq 0, \quad \text{若 } x = a \quad (7.80)$$

这些结论是有道理的。如果某状态函数是 \hat{x} 的本征函数, 本征值为 a , $\Psi = g_a(x)$, 我们知道(第7.6节): 对 x 的测量一定得到值为 a 的结果。这只在 $x \neq a$ 时有概率密度 $|\Psi|^2$ 为零时才成立, 与 (7.79) 一致。

在更进一步考虑 $g_a(x)$ 的性质之前, 我们先定义海维赛德阶梯函数 (Heaviside step function) (见图7.4):

$$\begin{aligned} H(x) &= 1, & x > 0 \\ H(x) &= 0, & x < 0 \\ H(x) &= \frac{1}{2}, & x = 0 \end{aligned} \quad (7.81)$$

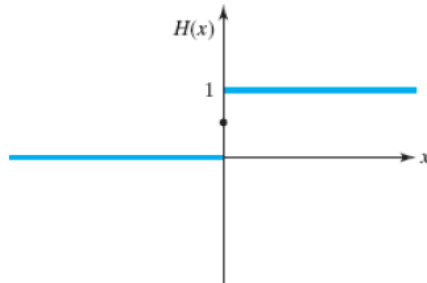


图 7.4: 海维赛德阶梯函数

接下来, 我们再定义狄拉克 δ 函数 (Dirac δ function) $\delta(x)$ 为海维赛德阶梯函数的导数:

$$\delta(x) \equiv \frac{dH(x)}{dx} \quad (7.82)$$

根据 (7.81) 和 (7.82), 我们立即得出 (另见图7.4):

$$\delta(x) = 0, \quad x \neq 0 \quad (7.83)$$

由于 $H(x)$ 在 $x = 0$ 处的值突然变化, 其导数在原点为无穷大:

$$\delta(x) = \infty, \quad x = 0 \quad (7.84)$$

我们可以通过设置 $x = t - a$, 然后将符号 t 改为 x , 将这些方程稍作归纳。方程 (7.81) 至 (7.84) 可以写成

$$H(x - a) = 1, \quad x > a \quad (7.85)$$

$$H(x - a) = 0, \quad x < a \quad (7.86)$$

$$H(x - a) = \frac{1}{2}, \quad x = a \quad (7.87)$$

$$\delta(x - a) = dH(x - a)/dx \quad (7.88)$$

$$\delta(x - a) = 0, \quad x \neq a \quad \text{以及} \quad \delta(x - a) = \infty, \quad x = a \quad (7.89)$$

现在, 考虑如下的积分:

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx$$

我们使用分部积分法计算:

$$\begin{aligned} \int u dv &= uv - \int v du \\ u &= f(x), \quad dv = \delta(x - a) dx \end{aligned}$$

使用 (7.88), 我们有

$$\begin{aligned} du &= f'(x) dx, \quad v = H(x - a) \\ \int_{-\infty}^{\infty} f(x) \delta(x - a) dx &= f(x) H(x - a) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H(x - a) f'(x) dx \\ \int_{-\infty}^{\infty} f(x) \delta(x - a) dx &= f(\infty) - \int_{-\infty}^{\infty} H(x - a) f'(x) dx \end{aligned} \quad (7.90)$$

其中, 我们用到了 (7.85) 和 (7.90)。由于 $H(x - a)$ 在 $x < a$ 时为零, (7.90) 变为

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta(x - a) dx &= f(\infty) - \int_{-\infty}^a H(x - a) f'(x) dx \\ &= f(\infty) - \int_{-\infty}^a f'(x) dx = f(\infty) - f(x) \Big|_a^{\infty} \\ \int_{-\infty}^{\infty} f(x) \delta(x - a) dx &= f(a) \end{aligned} \quad (7.91)$$

将 (7.91) 与方程 $\sum_k c_k \delta_{ik} = c_i$ 比较, 我们可以看到, 狄拉克 δ 函数在积分中所起的作用与 Kronecker 三角在求和中所起的作用相同。(7.91) 的一个特例为 $f(x) = 1$:

$$\int_{-\infty}^{\infty} \delta(x - a) dx = 1$$

狄拉克 δ 函数的性质 (7.89) 与位置本征函数 $g_a(x)$ 的性质 (7.79) 和 (7.80) 一致。因此, 我们暂定

$$g_a(x) = \delta(x - a) \quad (7.92)$$

为了证明 (7.92), 我们现在需要证明其符合玻恩假设, 即在 a 到 $a + da$ 的无穷小区间内观测到 x 值的概率为 $|\Psi(a, t)|^2 da$ 。根据 (7.76), 这个概率由下式给出:

$$|\langle g_a(x) | \Psi(x, t) \rangle|^2 da = \left| \int_{-\infty}^{\infty} g_a^*(x) \Psi(x, t) dx \right|^2 da \quad (7.93)$$

先使用 (7.92) 后使用 (7.91), 我们将 (7.93) 变为

$$\left| \int_{-\infty}^{\infty} \delta(x - a) \Psi(x, t) dx \right|^2 da = |\Psi(a, t)|^2 da$$

得证。

由于 $\delta(x - a)$ 中的 a 可以去任何实数值, 那么 \hat{x} 的本征值构成一个连续体: $-\infty < a < \infty$ 。与连续本征函数一致, $\delta(x - a)$ 不是平方可积的 (问题 7.43)。

总而言之, 位置的本征函数和本征值为

$$\hat{x} \delta(x - a) = a \delta(x - a) \quad (7.94)$$

其中 a 是任意实数。

δ 函数不是品优函数, 因此我们所做的处理缺乏严谨性, 会让数学家不寒而栗。不过, 我们可以将 δ 函数视为在零点处峰值逐渐增大的函数的极限情况, 从而对其进行严格的表述 (图 7.5)。 δ 函数其实不是函数, 而是数学家所说的分布 (distribution)。(见 en.wikipedia.org/wiki/Dirac_delta_function)

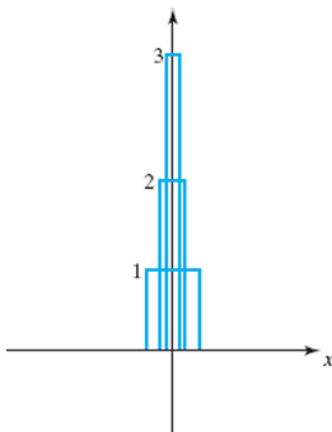


图 7.5: 近似 $\delta(x)$ 函数, 精确度依次提高。每条曲线下的面积为 1。

7.8 量子力学的基本假设

本节总结了前几章介绍的量子力学假设。

假设 1: 一个系统的状态由一个关于粒子坐标和时间的函数 Ψ 给出。这个函数被称为状态函数或波函数, 它包含了可以确定的关于系统的所有信息。我们进一步假设 Ψ 是单值、连续和平方可积的。对于连续态, 平方可积性的要求可以省略。

用“波函数”来称呼 Ψ 也许不是最好的选择。在三维空间中运动的物理波是三个空间坐标和时间的函数。然而，对于一个 n 粒子系统，函数 Ψ 是 $3n$ 空间坐标和时间的函数。因此，对于一个多粒子系统，我们不能把 Ψ 解释为任何一种物理波。状态函数最好被视为一个函数，我们可以从中计算出系统的各种属性。 Ψ 所包含信息的性质是假设 5 及其结果的主题。

假设 2: 每个物理可观测量都对应一个线性厄米算符。要找到这个算符，可以用笛卡尔坐标和相应的线动量分量写出可观测特性的经典力学表达式，然后用算符 x 替换每个坐标 x ，用算符 $-i\hbar \partial/\partial x$ 替换每个动量分量 p_x 。

我们在第 7.2 节中看到，对厄米算符的限制源于物理量的平均值必须是实数的要求。线性的要求与第 7.6 节讨论的态叠加密切相关。在我们对用 \hat{B} 本征函数的叠加展开的状态其属性平均值 B 的 (7.70) 的推导中， \hat{B} 的线性性质起了关键作用。

当经典量包含笛卡尔坐标与其共轭动量的乘积时，我们在构建正确的量子力学算符时就会遇到非对易的问题。已经提出了几种不同的规则来处理这种情况。见 J. R. Shewell, *Am. J. Phys.*, **27**, 16 (1959); E. H. Kerner and W. G. Sutcliffe, *J. Math. Phys.*, **11**, 391 (1970); A. de Souza Dutra, *J. Phys. A: Math. Gen.*, **39**, 203 (2006) (arxiv.org/abs/0705.3247).

在非笛卡尔坐标中寻找量子力学算符的过程非常复杂。见 K. Simon, *Am. J. Phys.*, **33**, 60 (1965); G. R. Gruber, *Found. Phys.*, **1**, 227 (1971)。

假设 3: 对物理可观测量 B 的测量结果唯一可能的值是方程 $\hat{B}g_i = b_i g_i$ 中的本征值 b_i ，其中 \hat{B} 是与属性 B 相对应的算符。本征函数 g_i 必须是品优函数。

我们主要关注原子和分子的能级。它们由能量算符的本征值，即哈密顿量 \hat{H} 给出。 \hat{H} 的本征方程， $\hat{H}\psi = E\psi$ ，是含时薛定谔方程。然而，要找到任何属性的可能值，都需要求解本征值方程。

假设 4: 若算符 \hat{B} 代表某个物理可观测量的线性厄米算符，那么其本征函数 g_i 构成一个完全集。

有些厄米算符的本征函数并不构成一个完全集（见 *Griffiths*, pp. 99, 106; *Messiah*, p. 188; *Ballentine*, Sec. 1.3）。完备性要求对于发展量子力学理论至关重要，因此有必要假设所有与可观测性质相对应的厄米算符都有一组完备的本征函数。假设 4 使我们能够将任何状态的波函数展开为任何量子力学算符的正交本征函数的叠加：

$$\Psi = \sum_i c_i g_i = \sum_i |g_i\rangle \langle g_i|\Psi\rangle \quad (7.95)$$

假设 5: 若 $\Psi(q, t)$ 表示 t 时刻系统的已归一化状态函数，那么 t 时刻物理可观测测量 B 的平均值为

$$\langle B \rangle = \int \Psi^* \hat{B} \Psi \, d\tau \quad (7.96)$$

量子力学平均值的定义见第 3.7 节，不应与经典力学中使用的时间平均值相混淆。

根据假设 4 和 5，我们已经在第 7.6 节中得出了在对 B 的测量中观测到非简并本征值 b_i 的概率为 $P(b_i) = |\int g_i^* \Psi \, d\tau|^2 = |\langle g_i|\Psi\rangle|^2$ ，其中 $\hat{B}g_i = b_i g_i$ 。如果 Ψ 恰好是 \hat{B} 的本征函数直译，也就是说 $\Psi = g_k$ ，那么 $P(b_i)$ 变为 $P(b_i) = |\int g_i^* g_k \, d\tau|^2 = |\delta_{ik}|^2 = \delta_{ik}$ ，其中用到了厄米算符 \hat{B} 的本征函数之间互相正交的性质。当 $\Psi = g_k$ 时，我们一定会得到 b_k 的测量结果。

假设 6: 未受干扰的量子力学系统状态的时间发展由含时薛定谔方程给出

$$-\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = \hat{H} \Psi \quad (7.97)$$

其中 \hat{H} 是系统的哈密顿（也是能量）算符。

含时薛定谔方程关于时间的一阶微分方程，因此就像在经典力学中一样，未受干扰系统的当前状态决定了未来状态。然而，与经典力学中的状态知识不同，量子力学中的状态知识只涉及测量各种可能结果的概率。因此，假设我们有几个完全相同的非相互作用系统，每个系统在 t_0 时刻的状态函数均为 $\Psi(t_0)$ 。如果每个系统都不受干扰，那么每个系统的状态函数的变化都将遵循 (7.97)。由于每个系统哈密顿量都相同，在未来的任意时刻 t_1 ，每个系统的状态函数都将变为 $\Psi(t_1)$ 。然而，假设我们在 t_2 时刻对每个系统的属性 B 进行一次观测。虽然在测量开始时，每个系统具有相同的态函数 $\Psi(t_2)$ ，我们不会在每个系统中得到相同的结果。相反地，我们会得到一系列 b_i 的可能值，其中 b_i 是 \hat{B} 的本征值。我们得到每个 b_i 的相对次数可以通过下式计算： $|c_i|^2$ ，其中 $\Psi(t_2) = \sum_i c_i g_i$ ， g_i 是 \hat{B} 的本征函数。

如果哈密顿算符与时间无关，我们就可能观测到具有确定能量 E 的状态。对于这样的状态，态函数必须满足

$$\hat{H} \Psi = E \Psi \quad (7.98)$$

含时薛定谔方程变为

$$-\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = E \Psi$$

积分得到 $\Psi = A e^{-iEt/\hbar}$ ，其中 A 是积分“常量”，与时间无关。函数 Ψ 与坐标和时间都相关，所以 A 是坐标的函数，我们将其命名为 $\psi(q)$ 。因此，对于具有确定能量的状态，有

$$\Psi(q, t) = e^{-iEt/\hbar} \psi(q) \quad (7.99)$$

函数 $\psi(q)$ 满足定态薛定谔方程

$$\hat{H} \psi(q) = E \psi(q)$$

由 (7.98) 和 (7.99) 可得上式。因子 $e^{-iEt/\hbar}$ 只简单表明波函数 $\Psi(q, t)$ 的相位随时间的变化，没有直接的物理意义。因此我们通常将 $\psi(q)$ 视为“波函数”。哈密顿算符在量子力学中扮演着独特的角色，因为它出现在基本动力学方程——含时薛定谔方程中。 \hat{H} 的本征态（也成为定态）有特殊的性质，就是概率密度 $|\Psi|^2$ 不随时间变化。

含时薛定谔方程 (7.97) 为 $(i\hbar \partial/\partial t - \hat{H}) \Psi = 0$ 。由于算符 $i\hbar \partial/\partial t - \hat{H}$ 是线性算符，那么含时薛定谔方程 (7.97) 的解的任意线性组合也是 (7.97) 的解。例如，若 \hat{H} 与时间无关，那么有定态解 $\Psi_n = e^{-iE_n t/\hbar} \psi_n(q)$ [式 (7.99)]。任意线性组合

$$\Psi = \sum_n c_n \Psi_n = \sum_n c_n e^{-iE_n t/\hbar} \psi_n(q) \quad (7.100)$$

（其中 c_n 是与时间无关的常数）也是含时薛定谔方程的解，尽管它不是 \hat{H} 的本征函数。由于本征函数 Ψ_n 可以是复函数，若 \hat{H} 与时间无关，那么任何态函数都可以写成 (7.100) 的形式。（另见第 9.8 节。）状态函数 (7.100) 表示一种没有确定能量的状态。相反地，当我们测量能量时，得到能量值为 E_n 的概率为 $|c_n e^{-iE_n t/\hbar}|^2 = |c_n|^2$ 。

为了求出 (7.100) 中的系数 c_n ，我们将 (7.100) 在 t_0 时刻的表达式写成 $\Psi(t_0) = \sum_n c_n e^{-iE_n t_0/\hbar} \psi_n(q)$ 。将其与 $\psi_j^*(q)$ 相乘并对全空间进行积分，得到

$$\langle \psi_j(q) | \Psi(q, t_0) \rangle = \sum_n c_n e^{-iE_n t_0/\hbar} \langle \psi_j | \psi_n \rangle = \sum_n c_n e^{-iE_n t_0/\hbar} \delta_{jn} = c_j e^{-iE_j t_0/\hbar}$$

则 $c_n = \langle \psi_j | \Psi(q, t_0) \rangle e^{iE_j t_0 / \hbar}$ 。则 (7.100) 变为

$$\Psi(q, t) = \sum_n \langle \psi_j(q) | \Psi(q, t_0) \rangle e^{-iE_j(t-t_0)/\hbar} \psi_j(q), \quad \text{若 } \hat{H} \text{ 与时间无关} \quad (7.101)$$

其中 $\hat{H}\psi_j = E_j\psi_j$, ψ_j 被选择为正交的。方程 (7.101) 告诉我们如何从初始时间 t_0 时的 Ψ 求出时间 t 时的 Ψ , 当 \hat{H} 与 t 无关时, 它是含时薛定谔方程的通解。[方程 (7.101) 也可以直接从含时薛定谔方程导出, 见问题 7.46。]

例题: 长度为 l 的一维势箱中有一粒子, 其哈密顿量与时间无关, 其态函数为 $\Psi = 2^{-1/2}\psi_1 + 2^{-1/2}\psi_2$, 其中 ψ_1 和 ψ_2 分别是箱中粒子定态能量本征函数的 $n = 1$ 和 $n = 2$ 形式 [式 (2.23)]。(a) 求出概率密度, 表示为时间的函数; (b) 证明 $|\Psi|^2$ 振荡的周期为 $T = 8ml^2/3h$; (c) 使用 Excel 或 Mathcad 绘制 $l|\Psi|^2$ 在每个时刻 $jT/8$ 随 x/l 变化的图象, 其中 $j = 0, 1, 2, \dots, 8$ 。

(a) 由于 \hat{H} 与时间无关, 则在未来的任意时刻, Ψ 的形式由 (7.100) 给出, 其中 $c_1 = 2^{-1/2}$, $c_2 = 2^{-1/2}$, 其他的 c 值全为零。因此,

$$\begin{aligned} \Psi &= \frac{1}{\sqrt{2}} e^{-iE_1 t / \hbar} \left(\frac{2}{l} \right)^{1/2} \sin\left(\frac{\pi x}{l}\right) + \frac{1}{\sqrt{2}} e^{-iE_2 t / \hbar} \left(\frac{2}{l} \right)^{1/2} \sin\left(\frac{2\pi x}{l}\right) \\ &= \frac{1}{\sqrt{2}} e^{-iE_1 t / \hbar} \psi_1 + \frac{1}{\sqrt{2}} e^{-iE_2 t / \hbar} \psi_2 \end{aligned}$$

因此, 我们有概率密度 (见问题 7.47)

$$\Psi^* \Psi = \frac{1}{2} \psi_1^2 + \frac{1}{2} \psi_2^2 + \psi_1 \psi_2 \cos[(E_2 - E_1)t / \hbar] \quad (7.102)$$

(b) $|\Psi|^2$ 中与时间有关的因子为余弦因子 (7.102)。最小正周期为余弦函数的参数函数增加 2π 的时间间隔, 即 $(E_2 - E_1)T / \hbar = 2\pi$ 。由于 $E_n = n^2 h^2 / 8ml^2$, 因此周期为 $T = 2\pi \hbar / (E_2 - E_1) = 8ml^2 / 3h$ 。

(c) 使用 (7.102) 和 ψ_1 、 ψ_2 的表达式, 以及 $T = 2\pi \hbar / (E_2 - E_1)$, 我们有

$$l|\Psi|^2 = \sin^2(\pi x_r) + \sin^2(2\pi x_r) + \sin(\pi x_r) \sin(2\pi x_r) \cos(2\pi t/T) \quad (7.103)$$

其中 $x_r = x/l$ 。由于 $t = jT/8$, 对于每个 j 的值, 函数图象很容易得到。图中显示, 概率密度最大值在方框的左右两侧摆动。使用 Mathcad, 我们可以制作出 $|\Psi|^2$ 随着时间变化而变化的影片 (问题 7.47)。(www.falstad.com/qm1d 上有一个在线资源, 可以让人们跟踪 $|\Psi|^2$ 作为时间函数的系统, 例如箱中粒子或谐振子, 对于任意选定的静止状态的初始混合物; 人们可以通过点击底部的小圆圈并拖动圆圈内的每个旋转箭头来选择混合物。)

当 \hat{H} 与时间无关时, 方程 (7.100) 是含时薛定谔方程的通解, 其中的 c_n 是常数。对于一个受到随时间变化的外力作用的系统来说, 哈密顿算符包含一个与时间相关的项: $\hat{H} = \hat{H}^0 + \hat{H}'(t)$, 其中 \hat{H}^0 是系统在外力作用下的哈密顿算符, $\hat{H}'(t)$ 是系统与外力作用随时间变化的势能。在这种情况下, 我们使用 \hat{H}^0 的定态本征函数来展开 Ψ , 与 (7.100) 类似, 除了 c_n 与时间 t 有关。例如, 原子或分子暴露在电磁辐射 (光) 随时间变化的电场中; 见第 9.8 节。

是什么决定了系统处于 (7.99) 这样的定态还是 (7.100) 这样的非定态? 答案是系统的历史决定了它现在的状态。例如, 如果我们将一个处于定态的系统暴露在辐射下, 根据含时薛定谔方程, 辐射会导致其状态转变为非定态; 见第 9.8 节。

你可能会对玻恩假设的列表中没有 $|\Psi(x, t)|^2 dx$ 是在 x 和 $x + dx$ 之间找到粒子的概率感到奇

怪。它是假设 5 的结果，我们现在就来说明这一点。方程 (3.81) 为 $\langle B \rangle = \sum_b P_b b$ ，其中 P_b 是测量具有离散值的属性 B 时观测到本征值 b 的概率。连续变量 x 对应的方程为 $\langle x \rangle = \int_{-\infty}^{\infty} P(x) x dx$ ，其中 $P(x)$ 是测量连续变量 x 的概率密度。根据假设 5，我们有 $\langle x \rangle = \int_{-\infty}^{\infty} \Psi^* \hat{x} \Psi dx = \int_{-\infty}^{\infty} |\Psi|^2 x dx$ 。将两个关于 $\langle x \rangle$ 的表达式相比较，我们得到 $|\Psi|^2$ 就是概率密度 $P(x)$ 。

第 10 章给出了涉及自旋和自旋统计定理的两个量子力学假设。

7.9 测量和量子力学解释

在量子力学中，系统的状态函数有两种变化方式。[参见 E. P. Wigner, *Am. J. Phys.*, **31**, 6 (1963).] 首先，含时薛定谔方程 (7.97) 给出了随时间发生的连续因果变化。其次，在对系统进行测量时，会发生突然的、不连续的、概率性的变化。这种变化无法准确预测，因为测量结果无法准确预测，只有概率 (7.73) 可以预测。由测量引起的 Ψ 的突然变化被称为**波函数坍缩** (reduction or collapse of the wave function)。对性质 B 的测量有结果 b_k ，而同时将系统的态函数转变为 g_k ，其中 g_k 是与 b_k 对应的 \hat{B} 的本征函数。(若 b_k 是简并的，则 Ψ 变为对应于 b_k 的所有本征函数的线性组合。) 找到非简并本征值 b_k 的概率由 (7.73)，和定理 9 的 $|\langle g_k | \Psi \rangle|^2$ 给出，则当我们对性质 B 进行测量时， $|\langle g_k | \Psi \rangle|^2$ 是系统从状态 Ψ 转变为 g_k 的概率。

考虑一个例子。假设在 t 时刻，我们测量粒子的位置。假设 $\Psi(x, t_-)$ 是粒子测量前一瞬间的状态函数 (图 7.6a)。我们进一步假设，测量的结果是发现粒子位于以下空间的无限小区域内

$$a < x < a + da \quad (7.104)$$

我们现在想问：测量后一瞬间粒子的状态函数 $\Psi(x, t_+)$ 是什么？为了回答这个问题，假设我们在时间 t_+ 时对位置进行第二次测量。由于 t_+ 与第一次测量的时间 t 相差无穷小，我们仍然必须发现粒子被限制在 (7.104) 区域内。如果粒子在无限小的时间内移动了有限的距离，它就会有无限大的速度，这是不可接受的。由于 $|\Psi(x, t_+)|^2$ 表示找到连续变量 x 的概率密度，因此我们可以得出在 (7.104) 区域外， $|\Psi(x, t_+)|^2 = 0$ ，其图象一定如图 7.6b 所示。因此，在 t 时刻进行的位置测量将 Ψ 从一个遍布整个空间的函数坍缩为一个集中在 (7.104) 区域的函数。从 $\Psi(x, t_-)$ 到 $\Psi(x, t_+)$ 的变化是一种概率变化。

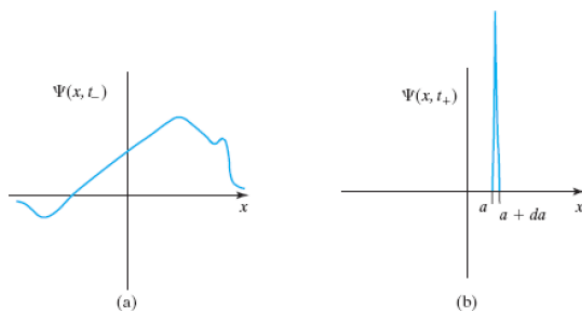


图 7.6: 对位置的测量导致波函数坍缩示意图

测量过程是量子力学中最具争议的领域之一。至于在测量过程中如何以及在哪个阶段出现坍缩，目前还不清楚。一些物理学家把 Ψ 的坍缩看作是一个额外的量子力学公设，而另一些物理学家则声称这是一个可以从其他公设中推导出来的定理。一些物理学家拒绝接受坍缩的观点 [见 M. Jammer, *The Philosophy of Quantum Mechanics*, Wiley, 1974, Section 11.4 ; L. E. Ballentine, *Am. J. Phys.*, **55**, 785 (1987)]。Ballentine 主张爱因斯坦对量子力学的统计集合解释，即波函数并不描述单个系统

的状态（如正统解释）(orthodox interpretation)，而是对以相同方式制备的大量系统的集合（集合）进行统计描述。在这种解释中，波函数的坍缩就不需要了。[见 L. E. Ballentine, *Am. J. Phys.*, **40**, 1763 (1972); *Rev. Mod. Phys.*, **42**, 358 (1970).] 统计集合解释存在许多严重问题 [见 Whitaker, pp. 213-217; D. Home and M. A. B. Whitaker, *Phys. Rep.*, **210**, 223 (1992); 问题 10.4]，这种解释在很大程度上已被否定。

“对大多数物理学家来说，找到一个一致的、可信的量子测量理论的问题仍未解决……。关于量子测量……众说纷纭……[是] 对整个量子力学解释的根本分歧的反映。” (M. Jammer, *The Philosophy of Quantum Mechanics*, pp. 519, 521).

量子力学的概率性质困扰了许多物理学家，包括爱因斯坦、德布罗意和薛定谔。这些物理学家和其他物理学家认为，量子力学可能无法完整描述物理现实。相反，量子力学的概率法则可能只是在亚量子力学层面上运行的、涉及“隐变量”的确定性法则的反映。物理学家博姆 (Bohm) 用尘埃粒子在空气中的布朗运动作了一个比喻。粒子的位置会发生随机波动，其运动并不完全由其位置和速度决定。当然，布朗运动是气体分子碰撞的结果，是由分子运动层面的变量决定的。同样，电子的运动可能是由存在于亚量子力学层面的隐变量决定的。由海森堡和玻尔提出的量子力学正统解释 [通常称为哥本哈根解释 (Copenhagen interpretation)] 否认隐变量的存在，并断言量子力学定律提供了对物理现实的完整描述。(F. J. Belinfante, *A Survey of Hidden-Variables Theories*, Pergamon, 1973, 中讨论了隐变量理论。)

1964 年，J. S. Bell 证明，在某些实验中，涉及对原本处于同一空间区域的两个相距甚远的粒子进行测量时，任何可能的局部隐变量理论都必须做出与量子力学不同的预测（见 Ballentine, Chapter 20）。在局部理论 (local theory) 中，两个相距甚远的系统彼此独立运行。这些实验的结果与量子力学的预言一致，从而为反对所有确定性的局部隐变量理论提供了非常有力的证据，但并不排除非局部隐变量理论。这些实验见 A. Aspect, in *The Wave-Particle Dualism*, S. Diner et al. (eds.), Reidel, 1984, pp. 377-390; A. Shimony, *Scientific American*, Jan. 1988, p. 46; A. Zeilinger, *Rev. Mod. Phys.*, **71**, S288 (1999).

量子理论预言并经实验证实，当对曾经相互作用但现在被无限距离分开的两个粒子进行测量时，对其中一个粒子的测量结果取决于对第二个粒子的测量结果，而且还取决于测量的是第二个粒子的哪种性质。（这种粒子被称为纠缠 (entangled) 粒子。有关纠缠的更多信息，请参阅 en.wikipedia.org/wiki/Quantum_entanglement; chaps.7-10 of J. Baggott, *Beyond Measure*, Oxford, 2004; L. Gilder, *The Age of Entanglement*, Vintage, 2009; Part II of A. Whitaker, *The New Quantum Age*, Oxford, 2012.) 这种瞬时的“超距作用”（爱因斯坦的说法）用一位物理学家的话说：“量子力学就是魔法” (D. Greenberger, quoted in N. D. Mermin, *Physics Today*, April 1985, p. 38)。

量子力学与心灵的关系一直是人们猜测的主题。Wigner 认为：当测量结果进入观察者的意识时，就会发生波函数的坍缩，因此“具有意识的存在物在量子力学中的作用一定不同于无生命的测量装置”。他认为，有意识的生命很可能遵守与无生命物体不同的自然法则，并建议科学家寻找意识作用于物质的异常效应。[E. P. Wigner, “Remarks on the Mind-Body Question,” in *The Scientist Speculates*, I. J. Good, ed., Capricorn, 1965, p. 284; *Proc. Amer. Phil. Soc.*, **113**, 95 (1969); *Found. Phys.*, **1**, 35 (1970).]

1952 年，戴维·玻姆 (David Bohm) (根据德布罗意在 1927 年提出的建议，即波函数可以作为引导粒子运动的先导波) 设计了一种非局部确定性隐变量理论，该理论预言了与量子力学相同的实验结果 [D. Bohm, *Phys. Rev.*, **85**, 166, 180 (1952)]。在玻姆的理论中，粒子在任何一瞬间都有一个确定的位置和动量（尽管这些量是无法观测到的），而且粒子的运动轨迹是确定的。粒子还拥有波函数 Ψ ，其时间变化服从含时薛定谔方程。在玻姆的理论中，波函数是决定粒子运动的真实物理实体。如果我们给定一个粒子在某一特定位置、某一特定波函数、某一特定时间 t ，那么玻姆理论

假定了一个特定的方程，可以让我们根据粒子的波函数和位置计算出粒子在该时间的速度；知道了 t 时的位置和速度，我们就可以求出时间 $t + dt$ 时的位置，并可以利用含时薛定谔方程求出 $t + dt$ 时的波函数，然后根据 $t + dt$ 时的位置和波函数计算出 $t + dt$ 时的速度，以此类推。因此，可以根据初始位置和波函数（假设我们知道势能）计算出路径。原来，在玻姆的理论中，粒子的位置服从类似牛顿第二定律的方程 $m d^2x/dt^2 = -\partial V/\partial x$ [式 (1.8) 以及 (1.12)]，只是把势能 V 换作 $V + Q$ ，其中 Q 是量子势，是由波函数以某种方式计算得出的。在玻姆理论中，波函数并没有发生坍缩。相反，系统与测量仪器的相互作用遵循了玻姆理论的方程，但这种相互作用导致系统在测量后以波函数坍缩时的方式演变。

多年来，博姆的研究在很大程度上被忽视了，但人们对他的理论的兴趣与日俱增。有关博姆理论的更多信息，参见 Whitaker, Chapter 7; D. Bohm and B.J. Hiley, *The Undivided Universe*, Routledge, 1992; D. Z. Albert, *Scientific American*, May 1994, p. 58; S. Goldstein, “Bohmian Mechanics,” plato.stanford.edu/entries/qm-bohm; en.wikipedia.org/wiki/De_Broglie-Bohm_theory。丽贝卡·戈德斯坦 (Rebecca Goldstein) 的小说《光的特性》(Properties of Light, Houghton Mifflin 出版社, 2000 年) 中的一个主要人物部分是以大卫·博姆 (David Bohm) 为原型创作的。

一些物理学家认为，波函数仅仅代表了我们对系统状态的认识 [哥本哈根观点采用了这种认识论 (epistemic 解释)，而另一些物理学家则认为，波函数直接对应于物理现实的一个要素本体论 (ontic 解释)]。2012 年发表的一篇论文利用某些温和的假设证明了一个结果，作者认为这个结果非常有利于本体论解释; M. F. Pusey, J. Barrett, and T. Rudolph, *Nature Phys.* **8**, 476 (2012) (arxiv.org/abs/1111.3328)。关于这一结果 (称为 PBR 定理) 的讨论，见 www.aps.org/units/gqi/newsletters/upload/vol6num3.pdf; mattleifer.info/2011/11/20/can-the-quantum-state-be-interpreted-statistically。

尽管量子力学的实验预测无可争议，但其概念解释仍是激烈争论的主题。附有相关评论的优秀文献有 B. S. DeWitt and R. N. Graham, *Am. J. Phys.*, **39**, 724 (1971); L. E. Ballentine, *Am. J. Phys.*, **55**, 785 (1987)。另见 B. d’Espagnat, *Conceptual Foundations of Quantum Mechanics*, 2nd ed., Benjamin, 1976; M. Jammer, *The Philosophy of Quantum Mechanics*, Wiley, 1974; Whitaker, Chapter 8; P. Yam, *Scientific American*, June 1997, p. 124。A. Cabello 关于量子力学基础的在线书目列出了对量子力学的 12 种不同解释 (arxiv.org/abs/quant-ph/0012089)；维基百科的一篇文章列出了对量子力学的 14 种解释 (en.wikipedia.org/wiki/Interpretations_of_quantum_mechanics)。

7.10 矩阵

矩阵代数是现代分子量子力学计算的重要数学工具。矩阵也为量子力学的许多理论提供了便捷的表述方法。矩阵方法将在后面一些章节中使用，但本书的编写方式是：如果时间不允许，可以省略有关矩阵的材料。

矩阵 (matrix) 是由数字组成的矩形数组。组成矩阵的数字称为**矩阵元素** (matrix elements)。设矩阵 \mathbf{A} 有 m 行 n 列，设 a_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) 表示第 i 行第 j 列的元素。那么

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

\mathbf{A} 称为 m 行 n 列矩阵。不要把 \mathbf{A} 和行列式 (第 8.3 节) 混淆，矩阵的行数和列数不一定相等，也不

等于一个值。

一个**行矩阵** (row matrix) [也称为**行向量** (row vector)] 是一个只有一行的矩阵。一个**列矩阵** (column matrix) [也称为**列向量** (column vector)] 是一个只有一列的矩阵。

当且仅当两个矩阵 \mathbf{R} 和 \mathbf{S} 的行数和列数相等, 且对应元素也相等时, 我们才说它们是**相等的** (equal)。若 $\mathbf{R} = \mathbf{S}$, 那么对 $j = 1, \dots, m$ 和 $k = 1, \dots, n$, 都有 $r_{jk} = s_{jk}$, 其中 m 和 n 是 \mathbf{R} 和 \mathbf{S} 的维数。因此, 一个矩阵方程等价于 nm 个标量方程组。

两个矩阵 \mathbf{A} 和 \mathbf{B} 的**和** (sum) 定义为一个新的矩阵, 其各元素为 \mathbf{A} 和 \mathbf{B} 对应元素的和, 只有 \mathbf{A} 和 \mathbf{B} 的行数和列数相等时, 我们才能定义它们的和。若 $\mathbf{P} = \mathbf{A} + \mathbf{B}$, 则我们有 mn 个标量方程 $p_{jk} = a_{jk} + b_{jk}$, 其中 $j = 1, \dots, m$ 和 $k = 1, \dots, n$ 。

$$\boxed{\text{若 } \mathbf{P} = \mathbf{A} + \mathbf{B}, \text{ 则 } p_{jk} = a_{jk} + b_{jk}} \quad (7.105)$$

标量 c 与矩阵 \mathbf{A} 的乘积定义为将 \mathbf{A} 的每个元素乘以 c :

$$\boxed{\text{若 } \mathbf{D} = c\mathbf{A}, \text{ 则 } d_{jk} = ca_{jk}} \quad (7.106)$$

若 \mathbf{A} 是 m 行 n 列的矩阵, \mathbf{B} 是 n 行 p 列的矩阵, 则**矩阵乘积** (matrix product) \mathbf{AB} 定义为一个 m 行 p 列的矩阵, 其元素为

$$\boxed{r_{jk} \equiv a_{j1}b_{1k} + a_{j2}b_{2k} + \dots + a_{jn}b_{nk} = \sum_{i=1}^n a_{ji}b_{ik}} \quad (7.107)$$

为了计算 r_{jk} , 我们取 \mathbf{A} 的第 j 行 (其元素为 $a_{j1}, a_{j2}, \dots, a_{jn}$), 将该行的每个元素与对应的 \mathbf{B} 的第 k 列 (其元素为 $b_{1k}, b_{2k}, \dots, b_{nk}$) 相乘并相加。例如, 我们假设

$$\mathbf{A} = \begin{pmatrix} -1 & 3 & \frac{1}{2} \\ 0 & 4 & 1 \end{pmatrix} \quad \text{和} \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 5 & 6 \\ -8 & 3 & 10 \end{pmatrix}$$

\mathbf{A} 的列数和 \mathbf{B} 的行数相等, 因此可以计算它们的乘积。 \mathbf{AB} 是 2×3 矩阵 \mathbf{A} 和 3×3 矩阵 \mathbf{B} 的乘积, 则 $\mathbf{R} \equiv \mathbf{AB}$ 是 2×3 矩阵。元素 r_{21} 由 \mathbf{A} 的第二行和 \mathbf{B} 的第一列计算得到: $r_{21} = 0(1) + 4(2) + 1(-8) = 0$ 。再计算其余的元素, 我们可以得出

$$\mathbf{R} = \begin{pmatrix} 1 & 16\frac{1}{2} & 25 \\ 0 & 23 & 34 \end{pmatrix}$$

矩阵乘积是不可对易的, \mathbf{AB} 和 \mathbf{BA} 不一定相等。(在上例中, 乘积 \mathbf{BA} 恰好没有定义。)可以证明, 矩阵乘法满足**结合律** (associative), 即 $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$, 也满足**分配律** (distributive), 即 $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ 和 $(\mathbf{B} + \mathbf{C})\mathbf{D} = \mathbf{BD} + \mathbf{CD}$ 。

行数和列数相等的矩阵称为**方阵** (square matrix)。方阵的**阶** (order) 等于其行数。

若 \mathbf{A} 是一个方阵, 则其平方和立方分别定义为 $\mathbf{A}^2 \equiv \mathbf{AA}$ 和 $\mathbf{A}^3 \equiv \mathbf{AAA}, \dots$ 。

n 阶方阵的元素 $a_{11}, a_{22}, \dots, a_{nn}$ 位于其**主对角线** (principal diagonal) 上。。若一个方阵的所有非对角元素都为零, 则称其为**对角矩阵** (diagonal matrix)。

方阵的**迹** (trace) 定义为主对角线元素的和。若 \mathbf{A} 是一个 n 阶方阵, 则其迹为 $\text{Tr}\mathbf{A} = \sum_{i=1}^n a_{ii}$ 。

对角元素都为 1 的对角矩阵称为**单位矩阵** (identity matrix or unit matrix)。单位矩阵的第 (j, k) 个元素为 Kronecker delta 函数: $(\mathbf{I})_{jk} = \delta_{jk}$, 其中 \mathbf{I} 是单位矩阵。例如, 一个三阶的单位阵为

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

令 \mathbf{B} 为一个与单位阵 \mathbf{I} 同阶的矩阵, 则矩阵积 \mathbf{IB} 的第 (j, k) 个元素由 (7.107) 给出: $(\mathbf{IB})_{jk} = \sum_i (\mathbf{I})_{ji} b_{ik} = \sum_i \delta_{ji} b_{ik} = b_{jk}$ 。由于对于所有的 j 和 k , \mathbf{IB} 和 \mathbf{B} 的第 (j, k) 个元素都相等, 我们有 $\mathbf{IB} = \mathbf{B}$ 。同理, $\mathbf{BI} = \mathbf{B}$ 。因此, 任意矩阵与单位矩阵相乘不会产生任何影响。

所有元素都为零的矩阵称为**零矩阵** (zero matrix), 表示为 $\mathbf{0}$ 。一个**非零矩阵** (nonzero matrix) 是一个至少有一个非零元素的矩阵。这些定义也适用于行向量和列向量。

量子化学中的大多数矩阵都是方阵或行矩阵或列矩阵。

矩阵和量子力学

在 7.1 节中, 积分 $\int f_m^* \hat{A} f_n d\tau$ 称为 \hat{A} 的一个矩阵元。现在, 我们通过证明此类积分遵守矩阵代数规则来证明这一名称的合理性。

令函数 f_1, f_2, \dots, f_n 是一个完全且正交的函数集, 令 $\{f_i\}$ 表示这个完全集。则数字 $A_{mn} \equiv \langle f_m | \hat{A} f_n \rangle \equiv \int f_m^* \hat{A} f_n d\tau$ 称为以 $\{f_i\}$ 为基的线性算符 \hat{A} 的**矩阵元** (matrix element of the linear operator \hat{A} in the basis $\{f_i\}$)。方阵

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} \langle f_1 | \hat{A} f_1 \rangle & \langle f_1 | \hat{A} f_2 \rangle & \dots \\ \langle f_2 | \hat{A} f_1 \rangle & \langle f_2 | \hat{A} f_2 \rangle & \dots \\ \dots & \dots & \dots \end{pmatrix} \quad (7.108)$$

称为线性算符 \hat{A} 以 $\{f_i\}$ 为基的**矩阵表示** (matrix representative)。因为 $\{f_i\}$ 通常包含无穷多个函数, 所以 \mathbf{A} 通常是一个无穷阶矩阵。

考虑矩阵元积分的加和。假设 $\hat{C} = \hat{A} + \hat{B}$ 。一个典型的以 $\{f_i\}$ 为基的 \hat{C} 矩阵元为

$$\begin{aligned} C_{mn} &= \langle f_m | \hat{C} | f_n \rangle = \langle f_m | \hat{A} + \hat{B} | f_n \rangle = \int f_m^* (\hat{A} + \hat{B}) f_n d\tau \\ &= \int f_m^* \hat{A} f_n d\tau + \int f_m^* \hat{B} f_n d\tau = A_{mn} + B_{mn} \end{aligned}$$

因此, 若 $\hat{C} = \hat{A} + \hat{B}$, 则 $C_{mn} = A_{mn} + B_{mn}$, 这是 (7.105) 的矩阵加和形式。因此, 若 $\hat{C} = \hat{A} + \hat{B}$, 则 $\mathbf{C} = \mathbf{A} + \mathbf{B}$, 其中 \mathbf{A} 、 \mathbf{B} 和 \mathbf{C} 是算符的 \hat{A} 、 \hat{B} 和 \hat{C} 的矩阵表示。

相似地, 若 $\hat{P} = c\hat{S}$, 其中 c 为常数, 则我们得到 (问题 7.52) $P_{jk} = cS_{jk}$, 这是矩阵与标量相乘的规则。

最后, 假设 $\hat{R} = \hat{S}\hat{T}$, 我们有

$$R_{mn} = \int f_m^* \hat{R} f_n d\tau = \int f_m^* \hat{S}\hat{T} f_n d\tau \quad (7.109)$$

函数 $\hat{T} f_n$ 可以用完全正交函数集 $\{f_i\}$ 展开 [式 (7.41)]:

$$\hat{T} f_n = \sum_i c_i f_i = \sum_i \langle f_i | \hat{T} f_n \rangle f_i = \sum_i \langle f_i | \hat{T} | f_n \rangle f_i = \sum_i T_{in} f_i$$

则 R_{mn} 变为

$$R_{mn} = \int f_m^* \hat{S} \sum_i T_{in} f_i d\tau = \sum_i \int f_m^* \hat{S} f_i d\tau T_{in} = \sum_i S_{mi} T_{in} \quad (7.110)$$

方程 $R_{mn} = \sum_i S_{mi} T_{in}$ 是 (7.107) 的矩阵乘积形式。因此, 若 $\hat{R} = \hat{S}\hat{T}$, 则 $\mathbf{R} = \mathbf{S}\mathbf{T}$ 。

我们已经证明: 我们已经证明, 在一个完整的正交基集中, 线性算符的矩阵表示所服从的方程与算符所服从的方程相同。将 (7.109) 与 (7.110) 结合起来, 我们可以得出有用的结论

$$\sum_i \langle m | \hat{S} | i \rangle \langle i | \hat{T} | n \rangle = \langle m | \hat{S}\hat{T} | n \rangle \quad (7.111)$$

假设选择的基函数 $\{f_i\}$ 为算符 \hat{A} 的本征函数 g_i 的完全正交集，其中 $\hat{A}g_i = a_i g_i$ 。则矩阵元 A_{mn} 表示为

$$A_{mn} = \langle g_m | \hat{A} | g_n \rangle = \langle g_m | \hat{A} g_n \rangle = \langle g_m | a_n g_n \rangle = a_n \langle g_m | g_n \rangle = a_n \delta_{mn}$$

以 \hat{A} 的本征函数为正交基表示 \hat{A} 的矩阵是一个对角矩阵，其对角元素为 \hat{A} 的本征值 a_n 。相反地，我们可以证明（问题 7.53），当某算符 \hat{A} 以其完整正交基表示为对角矩阵时，那么基函数为 \hat{A} 的本征函数，对角元素为 \hat{A} 的本征值。

我们使用了完全正交基 $\{f_i\}$ 来将算符 \hat{A} 表示为 (7.108) 的矩阵 \mathbf{A} 。这组基还可以用来表示任意函数 u ，如下所示。根据 $u = \sum_i u_i f_i$ ，用完全集 $\{f_i\}$ 展开函数 u ，其中展开系数 u_i 由 (7.40) 给出： $u_i = \langle f_i | u \rangle$ 。展开系数 u_1, u_2, \dots 构成的集合形成一个列矩阵（列向量），称为 \mathbf{u} ，则 \mathbf{u} 称为函数 u 以 $\{f_i\}$ 为基的**矩阵表示**（matrix representative）。若 $\hat{A}u = w$ ，其中 w 是另一个函数，那么我们可以证明（问题 7.54） $\mathbf{A}\mathbf{u} = \mathbf{w}$ ，其中 \mathbf{A} 、 \mathbf{u} 和 \mathbf{w} 分别是算符 \hat{A} 、函数 u 和函数 w 以 $\{f_i\}$ 为基的矩阵表示。因此，若我们知道算符 \hat{A} 的矩阵表示 \mathbf{A} ，那么我们就得知算符 \hat{A} 作用于任意函数 u 的结果。那么知道了矩阵表示 \mathbf{A} 等价于知道了算符 \hat{A} 。

总结

习题

Chapter 8

变分法

8.1 变分理论

为了处理包含相互作用粒子系统（如原子和分子）的定态薛定谔方程，我们必须使用近似方法。本章讨论变分法，它允许我们在不求解薛定谔方程的情况下近似得到系统的基态能量。变分法基于如下理论：

变分理论：

给定一个系统，其哈密顿算符 \hat{H} 与时间无关，其最低能量本征值为 E_1 ，如果 ϕ 是系统粒子坐标的任何满足问题定解条件的已归一化品优函数，那么

$$\int \phi^* \hat{H} \phi d\tau \geq E_1, \quad \phi \text{ 已归一化} \quad (8.1)$$

变分理论允许我们计算系统基态能量的最大值。

为了证明 (8.1)，我们将 ϕ 用 \hat{H} 的本征函数完全正交集，即定态本征函数 ψ_k 展开：

$$\phi = \sum_k a_k \psi_k \quad (8.2)$$

其中

$$\hat{H} \psi_k = E_k \psi_k \quad (8.3)$$

注意：展开式 (8.2) 要求 ϕ 满足与 ψ_k 相同的定解条件。将 (8.2) 代入 (8.1) 的左侧，我们得到

$$\int \phi^* \hat{H} \phi d\tau = \int \sum_k a_k^* \psi_k^* \hat{H} \sum_j a_j \psi_j d\tau = \int \sum_k a_k^* \psi_k^* \sum_j a_j \hat{H} \psi_j d\tau$$

利用本征方程 (8.3)，并假设交换积分和无限求和是成立的，我们可以得到

$$\begin{aligned} \int \phi^* \hat{H} \phi d\tau &= \int \sum_k a_k^* \psi_k^* \hat{H} \sum_j a_j \psi_j d\tau = \int \sum_k a_k^* \psi_k^* \sum_j a_j \hat{H} \psi_j d\tau \\ &= \sum_k \sum_j a_k^* a_j E_j \delta_{kj} \end{aligned}$$

其中用到了本征函数 ψ_k 的正交性。我们对 j 求和，像往常一样，除了 $j = k$ 的项之外，Kronecker δ 使所有项都为零，从而得到

$$\int \phi^* \hat{H} \phi d\tau = \sum_k a_k^* a_k E_k = \sum_k |a_k|^2 E_k \quad (8.4)$$

由于 E_1 是 \hat{H} 的最低能量本征值, 我们有 $E_k \geq E_1$, 又由于 $|a_k|^2$ 为非负数, 我们可以将不等式 $E_k \geq E_1$ 的左右两侧乘以 $|a_k|^2$ 而不改变不等号的方向, 从而得到 $|a_k|^2 E_k \geq |a_k|^2 E_1$ 。因此, $\sum_k |a_k|^2 E_k \geq \sum_k |a_k|^2 E_1$, 使用 (8.4), 我们得到

$$\int \phi^* \hat{H} \phi d\tau = \sum_k |a_k|^2 E_k \geq \sum_k |a_k|^2 E_1 = E_1 \sum_k |a_k|^2 \quad (8.5)$$

因为 ϕ 已归一化, 我们有 $\int \phi^* \phi d\tau = 1$ 。将展开式 (8.2) 代入归一化条件, 得到

$$\begin{aligned} 1 = \int \phi^* \phi d\tau &= \int \sum_k a_k^* \psi_k^* \sum_j a_j \psi_j d\tau = \sum_k \sum_j a_k^* a_j \int \psi_k^* \psi_j d\tau = \sum_k \sum_j a_k^* a_j \delta_{kj} \\ &= \sum_k |a_k|^2 \end{aligned} \quad (8.6)$$

[注意, 在推导公式 (8.4) 和 (8.6) 时, 我们基本上分别重复了公式 (7.70) 和 (7.69) 的推导。]

将 (8.6) 代入 (8.5) 就得到了变分理论 (8.1) 的证明:

$$\int \phi^* \hat{H} \phi d\tau \geq E_1, \quad \phi \text{ 已归一化} \quad (8.7)$$

假设我们有一个未归一化的函数 ϕ 。为了应用变分理论, 我们用一个归一化常数 N 乘以 ϕ 使得 $N\phi$ 满足归一化条件。在 (8.7) 中, 我们将 ϕ 替换为 $N\phi$, 得到

$$|N|^2 \int \phi^* \hat{H} \phi d\tau \geq E_1 \quad (8.8)$$

N 由下式导出: $\int (N\phi)^* (N\phi) d\tau = |N|^2 \int \phi^* \phi d\tau = 1$, 因此 $|N|^2 = 1 / \int \phi^* \phi d\tau$ 。将其代入 (8.8), 我们得到

$$\boxed{\frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} \geq E_1} \quad (8.9)$$

其中 ϕ 是满足问题定解条件的任意品优函数 (不必归一化)。

函数 ϕ 称为**试探变分函数** (trial variational function), (8.1) 中的积分 [或 (8.9) 中的积分之比] 称为**变分积分** (variational integral)。为了获取基态能量 E_1 的良好近似值, 我们会尝试许多试探变分函数, 并寻找能给出最小变分积分值的函数。由 (8.1) 可知, 变分积分的值越小, 我们对 E_1 的近似就越好。推翻量子力学的一种方法是找到一个试探变分函数, 使已知 E_1 的某个系统的变分积分小于 E_1 。

令 ψ_1 为系统基态能量的真实波函数:

$$\hat{H}\psi_1 = E_1\psi_1 \quad (8.10)$$

如果我们碰巧幸运地发现了一个等于 ψ_1 的变分函数, 那么在 (8.1) 中使用 (8.10), 我们可以看到变分积分将等于 E_1 。因此, 基态波函数给出了变分积分的最小值。因此, 我们预期变分积分值越小, 试探变分函数就越接近真实的基态波函数。然而, 事实证明, 变分积分接近 E_1 的速度要比试探变分函数接近 ψ_1 的速度快得多, 而且有可能用一个相当差的 ϕ 得到一个相当好的近似值 E_1 。

在实践中, 人们通常会在试探函数 ϕ 中引入几个参数, 然后对这些参数进行变分, 从而使变分积分最小化。能否成功使用变分法, 取决于能否对试探函数做出正确的选择。

让我们来看一些变分法的例子。虽然这种方法的真正用途是针对我们不知道确切解的问题, 但我们还是要考虑那些完全可解的问题, 以便判断我们结果的准确性。

例题：为长度为 l 的一维势箱中的粒子设计一个试探变分函数。

波函数在势箱外为零，定解条件要求在 $x = 0$ 和 $x = l$ 处 $\psi = 0$ 。变分函数 ϕ 必须满足这些在边界点为零的条件。如公式 (4.57) 后所述，基态 ψ 在边界点内部没有节点，因此希望 ϕ 也没有内部节点。具有这些性质的一个简单函数为抛物线函数

$$\phi = x(l - x), \quad 0 \leq x \leq l \quad (8.11)$$

以及在势箱外，有 $\phi = 0$ 。由于我们尚未归一化 ϕ ，我们使用 (8.9) 来计算。势箱内部的哈密顿算符为 $-(\hbar^2/2m) d^2/dx^2$ 。对于 (8.9) 的分子和分母，我们可以得出

$$\int \phi^* \hat{H} \phi d\tau = -\frac{\hbar^2}{2m} \int_0^l (lx - x^2) \frac{d^2}{dx^2} (lx - x^2) dx = \frac{\hbar^2}{m} \int_0^l (lx - x^2) dx = \frac{\hbar^2 l^3}{6m} \quad (8.12)$$

$$\int \phi^* \phi d\tau = \int_0^l x^2 (l - x)^2 dx = \frac{l^5}{30} \quad (8.13)$$

将它们带入变分理论 (8.9)，我们得到

$$E_1 \leq \frac{5\hbar^2}{4\pi^2 ml^2} = 0.1266515 \frac{\hbar^2}{ml^2}$$

根据 (2.20)， E_1 的准确值为 $\hbar^2/8ml^2 = 0.125\hbar^2/8ml^2$ ，能量的误差约为 1.3%。

由于 $\int |\phi|^2 d\tau = l^5/30$ ，将 (8.11) 归一化为 $(30/l^5)^{1/2} x(l - x)$ 。图 7.3 显示，该函数与真正的箱中粒子基态波函数非常相似。

练习：一个单粒子一维系统的势能函数满足

$$V = \begin{cases} V_0, & 0 \leq x \leq l \\ \infty, & \text{elsewhere} \end{cases}$$

(其中 V_0 是常数)。

(a) 使用试探函数

$$\phi = \begin{cases} \sin(\pi x/l), & 0 \leq x \leq l \\ 0, & \text{elsewhere} \end{cases}$$

来估算该系统的基态能量；

(b) 解释为什么 (a) 的结果是该系统的准确基态能量。提示：见第 4 章的问题之一。(答案：(a) $V_0 + \hbar^2/8ml^2$ 。)

上一个例题的试探函数中没有参数。下一个例题有。

例题：对于一维谐振子，设计一个带有参数的变分函数，并找出参数的最佳值。估计基态能量。

变分函数 ϕ 必须平方可积，即当 $x \rightarrow \pm\infty$ 时， $\phi \rightarrow 0$ 。函数 e^{-x} 在 $+\infty$ 时满足条件，但在 $-\infty$ 时变为无穷大。函数 e^{-x^2} 在 $\pm\infty$ 时都满足条件。然而，从量纲上看，这种做法并不令人满意，因为 e 的指数必须是无量纲的。这可以从 Taylor 级数中看出 $e^z = 1 + z + z^2/2! + z^3/3! + \dots$ [式 (4.44)]。由于该级数中的所有项都必须有相同的量纲，那么 z 的量纲就必须为 1，也就是说 e^z

中的 z 必须是无量纲的。因此，我们将 e^{-x^2} 替换为 e^{-cx^2} ，其中 c 的量纲为长度⁻²。我们将 c 作为变分参数。真实的基态 ψ 没有节点。此外，由于 $V = \frac{1}{2}kx^2$ 是偶函数，由于奇函数在原点有节点，那么基态的 ψ 一定具有确定的宇称，也一定是偶函数。试探函数 e^{-cx^2} 具有我们需要的性质：无节点，且为偶函数。

对 \hat{H} 使用 (4.30)，再使用附录中的积分，有（问题 8.3）

$$\begin{aligned}\int \phi^* \hat{H} \phi d\tau &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} e^{-2cx^2} \frac{d^2 e^{-cx^2}}{dx^2} dx + 2\pi^2 \nu^2 m \int_{-\infty}^{\infty} x^2 e^{-2cx^2} dx \\ &= \frac{\hbar^2}{m} \left(\frac{\pi c}{m}\right)^{1/2} + \nu^2 m \left(\frac{\pi^5}{8}\right)^{1/2} c^{-3/2} \\ \int \phi^* \phi d\tau &= \int_{-\infty}^{\infty} e^{-2cx^2} dx = 2 \int_0^{\infty} e^{-2cx^2} dx = \left(\frac{\pi}{2c}\right)^{1/2}\end{aligned}$$

试探积分 W 为

$$W \equiv \frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} = \frac{\hbar^2 c}{2m} + \frac{\pi^2 \nu^2 m}{2c} \quad (8.14)$$

现在，我们通过改变 c 来最小化 W (8.14)。 W 最小的必要条件为

$$\begin{aligned}\frac{dW}{dc} = 0 &= \frac{\hbar^2}{2m} - \frac{\pi^2 \nu^2 m}{2c^2} \\ c &= \pm \pi \nu m / \hbar\end{aligned} \quad (8.15)$$

其中，负根 $c = -\pi \nu m / \hbar$ 应舍去，因其会导致 $\phi = e^{-cx^2}$ 不满足平方可积的要求。将 $c = \pi \nu m / \hbar$ 代入 (8.14)，我们得到 $W = \frac{1}{2} \hbar \nu$ 。这是一维谐振子基态能量的准确值。当 $c = \pi \nu m / \hbar$ 时，变分函数 ϕ 除了未归一化以外，与谐振子基态波函数 (4.53) 和 (4.31) 完全一致。

对于归一化的谐振子变分函数 $\phi = (2c/\pi)^{1/4} e^{-cx^2}$ ，若 c 的值很大，会导致 ϕ 在 $x = 0$ 附近从其最大值快速下降。那么概率密度就只在 $x = 0$ 附近较大。势函数 $V = \frac{1}{2}kx^2$ 在 $x = 0$ 附近也很小，所以较大的 c 意味着较小的 $\langle V \rangle = \langle \phi | V | \phi \rangle$ 。[注意： $\langle V \rangle$ 等于 (8.14) 右侧的第二项。]然而，由于较大的 c 值得 ϕ 从其最大值快速减小，那么在 $x = 0$ 附近 $|d\phi/dx|$ 的值变得很大，根据问题 7.7b，大 $|d\phi/dx|$ 意味着大 $\langle T \rangle$ [等于 (8.14) 右侧的第一项。] c 的最佳值可以使总和 $\langle T \rangle + \langle V \rangle = W$ 最小化。在原子和分子中，真正的波函数是通过将电子限制在低 V 区域（靠近原子核）来最小化 $\langle V \rangle$ 的趋势与通过允许电子概率密度在大区域内扩散来最小化 $\langle T \rangle$ 的趋势之间的折衷。

练习：考虑一个单粒子一维系统，其势函数满足

$$V = \begin{cases} V = 0, & -\frac{1}{2}l \leq x \leq \frac{1}{2}l \\ V = b\hbar^2/ml^2, & \text{elsewhere} \end{cases}$$

其中 b 是正常数（图 2.5 中的 $V_0 = b\hbar^2/ml^2$ 和原点发生了偏移）。

(a) 对于变分函数

$$\phi = \begin{cases} (x-c)(x+c) = x^2 - c^2, & -c \leq x \leq c \\ 0, & \text{elsewhere} \end{cases}$$

其中变分参数 $c > \frac{1}{2}l$ 。我们发现变分积分 W 由下式给出：

$$W = \frac{\hbar^2}{ml^2} \left[\frac{5l^2}{4c^2} + b \left(1 - \frac{15l}{16c} + \frac{5l^3}{32c^3} - \frac{3l^5}{256c^5} \right) \right]$$

在同一幅图上绘制 ϕ 和 V 的图象。求使 W 最小的 c 值所满足的方程。

(b) 当 $V_0 = 20\hbar^2/ml^2$ 时, 求出 c 的最佳值和 W 的值, 将其与基态能量真实值 $2.814\hbar^2/ml^2$ 进行比较 (问题 4.31c)。(提示: 可以使用 Excel 表格求解器或可编程计算器来计算 c/l 的值。)

(答案: (a) $48t^4 - 24t^3 - 128t^3/b + 3 = 0$, 其中 $t \equiv c/l$; (b) $c = 0.6715l$, $W = 3.454\hbar^2/ml^2$ 。)

8.2 变分法的推广

上一节介绍的变分法只提供了基态能量和波函数的信息。现在我们讨论将变分法推广到激发态。(另请参见第 8.5 节)。

考虑一下我们可以如何推广变分法, 来估算第一激发态的能量。根据能量从低到高:

$$E_1 \leq E_2 \leq E_3 \leq \dots$$

我们将系统的状态标号为 $1, 2, 3, \dots$ 。我们已经证明: 对于一个已归一化的变分函数 ϕ , 有 [式 (8.4) 和 (8.6)]

$$\int \phi^* \hat{H} \phi d\tau = \sum_{k=1}^{\infty} |a_k|^2 E_k \quad \text{和} \quad \int \phi^* \phi d\tau = \sum_{k=1}^{\infty} |a_k|^2 = 1$$

其中的 a_k 是展开式 $\phi = \sum_k a_k \psi_k$ 中的系数 [式 (8.2)]。我们有 $a_k = \langle \psi_k | \phi \rangle$ [式 (7.40)]。让我们只讨论与真实基态波函数 ψ 正交的归一化函数 ϕ 。那么 $a_1 = \langle \psi_1 | \phi \rangle = 0$, 因此

$$\int \phi^* \hat{H} \phi d\tau = \sum_{k=2}^{\infty} |a_k|^2 E_k \quad \text{和} \quad \int \phi^* \phi d\tau = \sum_{k=2}^{\infty} |a_k|^2 = 1 \quad (8.16)$$

对于 $k \geq 2$, 我们有 $E_k \geq E_2$ 和 $|a_k|^2 E_k \geq |a_k|^2 E_2$ 。因此,

$$\sum_{k=2}^{\infty} |a_k|^2 E_k \geq \sum_{k=2}^{\infty} |a_k|^2 E_2 = E_2 \sum_{k=2}^{\infty} |a_k|^2 = E_2 \quad (8.17)$$

将 (8.16) 和 (8.17) 相结合, 我们得到了所需的结果:

$$\int \phi^* \hat{H} \phi d\tau \geq E_2, \quad \text{若} \quad \int \psi^* \phi d\tau = 0 \quad \text{且} \quad \int \phi^* \phi d\tau = 1 \quad (8.18)$$

通过不等式 (8.18) 我们可以得到第一激发态的能量 E_2 的上限。然而, 限制条件 $\langle \psi_1 | \phi \rangle = 0$ 使得这种方法在应用上会很麻烦。

对于特定的系统, 尽管我们不知道真实基态波函数, 但确定 $\langle \psi_1 | \phi \rangle = 0$ 是可能的。 V 是 x 的偶函数的一维问题就是一个例子。在这种情况下, 基态波函数总是偶函数, 而第一激发态波函数是奇函数。(所有波函数必须具有确定的宇称。基态波函数是无节点的, 由于奇函数在原点消失, 所以基态波函数必须是偶函数。第一个激发态波函数有一个节点, 一定是奇函数。) 因此, 对于奇函数, $\langle \psi_1 | \phi \rangle = 0$ 一定成立, 偶函数 ψ_1 乘以一个奇函数 ϕ 得到一个奇被积函数, 从 $-\infty$ 到 $+\infty$ 的积分为零。

另一个例子是在中心场中运动的粒子 (第 6.1 节)。势能的形式可能使我们无法求解本征函数中的径向因子 $R(r)$ 。然而, 其中的角度因子 ψ 是球谐函数 [式 (6.16)], 而不同 l 值的球谐函数是正交的。因此, 我们可以利用试探函数中的因子 Y_l^m , 求出任意给定角动量 l 的最低态能量上限。这一结果取决于将 (8.18) 推广到更高的激发态:

$$\frac{\int \phi^* \hat{H} \phi d\tau}{\int \phi^* \phi d\tau} \geq E_{k+1}, \quad \text{若} \quad \int \psi_1^* \phi d\tau = \int \psi_2^* \phi d\tau = \dots = \int \psi_k^* \phi d\tau = 0 \quad (8.19)$$

8.3 行列式

第 8.5 节讨论了一种变分函数，它产生了一个涉及行列式的方程。因此，我们现在讨论行列式。

行列式 (Determinant) 是一个包含 n^2 个量 (称为**元素** (elements)) 的正方形数组；行列式的值是根据其元素计算出来的，计算方法稍后给出。 n 称为行列式的**阶** (order)。使用 a_{ij} 来代表一个典型元素，我们有 n 阶行列式

$$\det(a_{ij}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \quad (8.20)$$

(8.20) 中的竖线与绝对值无关。在考虑 n 阶行列式值的定义方法之前，我们先考虑一阶、二阶和三阶行列式。

一阶行列式有一个元素，其值就是该元素本身。因此，

$$|a_{11}| = a_{11} \quad (8.21)$$

其中的竖线表示行列式而不是绝对值。

二阶行列式有四个元素，其值定义为

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (8.22)$$

三阶行列式的值定义为

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad (8.23)$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \quad (8.24)$$

三阶行列式的计算方法是：用正负号交替写下第一行的元素，然后将每个元素乘以某个二阶行列式；在三阶行列式中划掉某个元素出现的行和列，就能找到与该元素相乘的二阶行列式。将 n 阶行列式的第 i 行和第 j 列删去后得到的 $(n-1)$ 阶行列式称为元素 a_{ij} 的**余子式** (minor)。我们再定义 a_{ij} 的**代数余子式** (cofactor) 为 a_{ij} 的余子式乘以 $(-1)^{i+j}$ 。因此，(8.23) 说明三阶行列式的计算方法是：将顶行的每个元素乘以它的代数余子式，然后将三个乘积相加。[注意：式 (8.22) 通过代数余子式符合该方法。因为 a_{11} 的代数余子式为 a_{22} ， a_{12} 的代数余子式为 $-a_{21}$ 。] 一个数值的例子如下

$$\begin{vmatrix} 5 & 10 & 2 \\ 0.1 & 3 & 1 \\ 0 & 4 & 4 \end{vmatrix} = 5 \begin{vmatrix} 3 & 1 \\ 4 & 4 \end{vmatrix} - 10 \begin{vmatrix} 0.1 & 1 \\ 0 & 4 \end{vmatrix} + 2 \begin{vmatrix} 0.1 & 3 \\ 0 & 4 \end{vmatrix} \\ = 5(8) - 10(0.4) + 2(0.4) = 36.8$$

我们将 a_{ij} 的余子式记为 M_{ij} ，代数余子式记为 C_{ij} ，则

$$c_{ij} = (-1)^{i+j} M_{ij} \quad (8.25)$$

则三阶行列式的展开式 (8.23) 可以写为

$$\det(a_{ij}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \quad (8.26)$$

三阶行列式可以用任意一行的元素和相应的代数余子式展开。例如，利用第二行展开三阶行列式，我们可以得出

$$\det(a_{ij}) = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \quad (8.27)$$

$$\det(a_{ij}) = -a_{21} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \quad (8.28)$$

二阶行列式的展开式表明 (8.28) 等于 (8.24)。我们也可以使用任意一列的元素和相应的代数余子式来展开行列式，这一点很容易得到验证。因此，对于三阶行列式，我们可以写成

$$\det(a_{ij}) = a_{k1}C_{k1} + a_{k2}C_{k2} + a_{k3}C_{k3} = \sum_{l=1}^3 a_{kl}C_{kl}, \quad k = 1, 2, 3$$

$$\det(a_{ij}) = a_{1k}C_{1k} + a_{2k}C_{2k} + a_{3k}C_{3k} = \sum_{l=1}^3 a_{lk}C_{lk}, \quad k = 1, 2, 3$$

第一个展开式按某一行展开，第二个则按某一列展开。

我们通过类似的行（或列）展开来定义高阶行列式。对于 n 阶行列式，

$$\det(a_{ij}) = \sum_{l=1}^n a_{kl}C_{kl} = \sum_{l=1}^n a_{lk}C_{lk}, \quad k = 1, 2, \dots, n \quad (8.29)$$

一些有关行列式的性质如下所述（证明见 *Sokolnikoff and Redheffer*, pp. 702-707）:

1. 如果行列式某行（或列）的每个元素都为零，则行列式的值为零。
2. 交换行列式的两行（或两列）会使行列式的值改变符号。
3. 如果行列式的两行（或两列）相同，则该行列式的值为零。
4. 如果将行列式的一行（或一列）乘以一个常数 k ，则该行列式的值也乘以 k 。
5. 如在一行的每个元素上加上另一行相应元素的相同常数倍数，行列式的值不变。这个定理也适用于将一列的倍数加到另一列。
6. 互换所有相应的行和列，行列式的值保持不变。（这种互换意味着第一列变为第一行，第二列变为第二行，等等）

例题：使用性质 5 计算

$$B = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} \quad (8.30)$$

将第一行的元素乘以 -2 并加到第四行，将其变为 $2+(-2)1=0, 3+(-2)2=-1, 4+(-2)3=-2,$

$1 + (-2)4 = -7$ 。同理，将第一行乘以 -3 并加到第三行，乘以 -4 并加到第二行。这样，(8.30) 变为

$$B = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -7 & -10 & -13 \\ 0 & -2 & -8 & -10 \\ 0 & -1 & -2 & -7 \end{vmatrix} = 1 \begin{vmatrix} -7 & -10 & -13 \\ -2 & -8 & -10 \\ -1 & -2 & -7 \end{vmatrix} \quad (8.31)$$

其中，我们将 B 按第一列展开。第二行减去第三行的两倍，第一行再减去第三行的七倍，得出

$$B = \begin{vmatrix} 0 & 4 & 36 \\ 0 & -4 & 4 \\ -1 & -2 & -7 \end{vmatrix} = (-1) \begin{vmatrix} 4 & 36 \\ -4 & 4 \end{vmatrix} = -(16 + 144) = -160 \quad (8.32)$$

行列式中从左上方到右下方的对角线就是**主对角线** (principal diagonal)。**对角行列式** (diagonal determinant) 是指除主对角线上的元素外，其他元素均为零的行列式。对于一个对角行列式，

$$\begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & 0 & \cdots & 0 \\ 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} a_{22} \begin{vmatrix} a_{33} & 0 & \cdots & 0 \\ 0 & a_{44} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} \quad (8.33)$$

$$= \cdots = a_{11} a_{22} a_{33} \cdots a_{nn}$$

对角行列式的值等于其对角元素的乘积。

一个行列式的唯一非零元素出现在以主对角线为中心的方形块中，这个行列式就是**块对角形式** (block diagonal form)。如果我们把每个正方形块看作行列式，那么块对角行列式就等于块的乘积。例如

$$\begin{vmatrix} a & b & 0 & 0 & 0 & 0 \\ c & d & 0 & 0 & 0 & 0 \\ 0 & 0 & e & 0 & 0 & 0 \\ 0 & 0 & 0 & f & g & h \\ 0 & 0 & 0 & i & j & k \\ 0 & 0 & 0 & l & m & n \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} (e) \begin{vmatrix} f & g & h \\ i & j & k \\ l & m & n \end{vmatrix} \quad (8.34)$$

虚线（图中未画出）勾勒出区块的轮廓。用顶行元素展开左边的行列式，再用它们的顶行展开后面的几个行列式，就可以很容易地证明方程 (8.34) (问题 8.21)。

8.4 联立线性方程组

要处理下一节所要讨论的变分函数，我们需要知道联立线性方程组。

考虑以下关于 n 个未知数的 n 个线性方程系统：

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \quad (8.35)$$

其中的 a 和 b 是已知常数, x_1, x_2, \dots, x_n 是未知数。如果至少有一个 b 值为非零, 我们得到了一个非齐次 (inhomogeneous) 方程组。这样的系统可以用 Cramer 法则 (Cramer's rule) 求解。(Cramer 法则的证明见 *Sokolnikoff and Redheffer*, pp. 708。) 令 $\det(a_{ij})$ 表示 (8.35) 中未知数的系数行列式, 由 Cramer 法则, x_k ($k = 1, 2, \dots, n$) 的值由下式给出:

$$x_k = \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & b_1 & a_{1,k+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,k-1} & b_2 & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & b_n & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}}{\det(a_{ij})}, \quad k = 1, 2, \dots, n \quad (8.36)$$

其中的 $\det(a_{ij})$ 由 (8.20) 给出, 分子中的行列式通过将 $\det(a_{ij})$ 的第 k 列替换为 b_1, b_2, \dots, b_n 来获得。虽然 Cramer 法则具有理论意义, 但不应在数值计算中使用, 因为连续消除未知数的效率更高。

一种广泛使用的连续消元法称为高斯消元法 (Gaussian elimination), 其过程如下: 将第一个方程的两边同除以 x_1 的系数 a_{11} , 使得 x_1 的系数变为 1。然后用第二个方程减去第一个方程的 a_{21} 倍, 用第三个方程减去第一个方程的 a_{31} 倍, 以此类推, 直到用第 n 个方程减去第一个方程的 a_{n1} 倍。这就在所有方程中消去了 x_1 (除了第一个方程)。接下来, 将第二个方程的两边同除以 x_2 的系数 a_{22} , 使得 x_2 的系数变为 1。然后用第三, 第四 ..., 第 n 个方程减去第二个方程的适当倍数, 这就消去了 x_2 (除了第一和第二个方程)。重复这个过程, 直到第 n 个方程只包含 x_n , 第 $n-1$ 个方程只包含 x_{n-1} 和 x_n , 等等。通过第 n 个方程解得 x_n , 然后将其代入第 $n-1$ 个方程中解得 x_{n-1} , 将 x_{n-1} 和 x_n 代入第 $n-2$ 个方程中解得 x_{n-2} , 以此类推, 直到解出所有未知数。如果在任何阶段, 我们想要除以的系数恰好为零, 那么零系数的等式就会与后来在所需位置上具有非零系数的等式交换。(高斯消元法也是一种有效的行列式求值方法, 见问题 8.25。)

一种相关的方法是高斯-若尔当消元法 (Gauss-Jordan elimination), 其过程与高斯消元法类似, 但与高斯消元法将 x_2 从方程 3, 4, ..., n 中消去不同, 通过减去第二个方程的适当倍数, 我们将 x_2 从方程 1, 3, 4, ..., n 中消去; 与将 x_3 从方程 4, 5, ..., n 中消去不同, 我们将 x_3 从方程 1, 2, 4, 5, ..., n 中消去, 以此类推。最后, 方程 1 只包含 x_1 , 方程 2 只包含 x_2 , 方程 n 只包含 x_n 。高斯-若尔当消元法比高斯消元法需要更多的计算量。

若 (8.35) 中所有的 b 都为零, 我们得到了一个齐次线性方程组 (homogeneous linear equations):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned} \quad (8.37)$$

(8.37) 的一组显而易见的解为 $x_1 = x_2 = \dots = x_n = 0$, 称为平凡解 (trivial solution)。若 (8.37) 的系数行列式 $\det(a_{ij})$ 不为零, 我们可以使用 Cramer 法则 (8.36) 来求解未知数, 并发现 $x_k = 0, k = 1, 2, \dots, n$, 因为 (8.36) 分子中的行列式有一列所有元素都为零。因此, 当 $\det(a_{ij}) \neq 0$ 时, 唯一的解就是平凡解, 这没有任何意义。要使关于 n 个未知数的, 由 n 个齐次方程组成的线性方程组有一组非平凡解, 系数的行列式必须为零。而且, 这个条件足以确保非平凡解的存在。因此, 我们得到了极为重要的定理:

当且仅当系数行列式为零时, 关于 n 个未知数的, 由 n 个齐次方程组成的线性方程组有一组非平凡解。

假设 $\det(a_{ij}) = 0$, 则 (8.37) 有一组非平凡解。我们如何找到它呢? 由于 $\det(a_{ij}) = 0$, Cramer 法则 (8.36) 给出 $x_k = 0/0, k = 1, 2, \dots, n$, 这没有任何意义。因此, Cramer 法则并不能立即起到

帮助作用。注意到, 若 $x_1 = d_1, x_2 = d_2, \dots, x_n = d_n$ 是 (8.37) 的一组解, 那么 $x_1 = cd_1, x_2 = cd_2, \dots, x_n = cd_n$ 也是 (8.37) 的一组解, 其中 c 是任意常数。这一点很容易看出, 因为

$$a_{11}cd_1 + a_{12}cd_2 + \dots + a_{1n}cd_n = c(a_{11}d_1 + a_{12}d_2 + \dots + a_{1n}d_n) = c \cdot 0 = 0$$

以此类推。因此, 线性齐次方程组的解将包含一个任意常数, 我们无法为每一个未知数找到一个唯一的值。为了求解 (8.37), 我们给任意一个未知数赋以任意的值, 如 $x_n = c$, 其中 c 是任意常数。确定了 x_n 的值之后, 我们将 (8.37) 中每个方程的最后一项移到右边, 得到

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1,n-1}x_{n-1} &= -a_{1n}c \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2,n-1}x_{n-1} &= -a_{2n}c \\ &\vdots \\ a_{n-1,1}x_1 + a_{n-1,2}x_2 + \dots + a_{n-1,n-1}x_{n-1} &= -a_{n-1,n}c \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{n,n-1}x_{n-1} &= -a_{nn}c \end{aligned} \quad (8.38)$$

我们现在有 $n-1$ 个未知数和 n 个方程, 这比我们需要的多了一个。因此, 我们删去 (8.38) 中的任意一个方程, 以最后一个为例。这就得到了关于 $n-1$ 个未知数的, 由 $n-1$ 个非齐次方程组成的线性方程组。然后就可以用 Cramer 法则 (8.36) 来求解未知数 x_1, x_2, \dots, x_{n-1} 。因为 (8.38) 中, 右侧的常数都包含因子 c , 所以根据 8.3 节的定理 4, x_1, x_2, \dots, x_{n-1} 的值也包含因子 c 。因此, (8.37) 的解可以写成

$$x_1 = ce_1, \quad x_2 = ce_2, \quad \dots, \quad x_{n-1} = ce_{n-1}, \quad x_n = c \quad (8.39)$$

其中 c 是任意常数, e_1, e_2, \dots, e_{n-1} 是数字。

例题: 求解:

$$3x_1 + 4x_2 + x_3 = 0$$

$$x_1 + 3x_2 - 2x_3 = 0$$

$$x_1 - 2x_2 + 5x_3 = 0$$

这是线性齐次方程组, 我们先计算系数行列式。得出 (见练习)

$$\begin{vmatrix} 3 & 4 & 1 \\ 1 & 3 & -2 \\ 1 & -2 & 5 \end{vmatrix} = 0$$

因此, 存在一组非平凡解。我们给 x_3 赋值 c , 删去第三个方程, 得到

$$3x_1 + 4x_2 = -c$$

$$x_1 + 3x_2 = 2c$$

第一个方程减去第二个方程的三倍, 得到 $-5x_2 = -7c$, 因此 $x_2 = \frac{7}{5}c$ 。代入 $x_1 + 3x_2 = 2c$, 得到 $x_1 + \frac{21}{5}c = 2c$, 因此 $x_1 = -\frac{11}{5}c$ 。因此, 通解为 $x_1 = -\frac{11}{5}c, x_2 = \frac{7}{5}c, x_3 = c$ 。对于那些对分式过敏的人 (无意冒犯, 译者注), 我们定义一个新的任意常数 $s \equiv \frac{1}{5}c$, 则 $x_1 = -11s, x_2 = 7s, x_3 = 5s$ 。

练习:

(a) 证明本例中系数行列式的值为零;

(b) 证明: 将第二个方程乘以某个常数并与第一个方程相加, 就可以得到本例中的第三个方程。

如果 $n-1$ 个未知数中的 $n-1$ 个非均质方程组的行列式 [(8.38), 最后一个方程省略] 恰好为零, 则上述步骤将失效。这样 Cramer 法则的分母中就有了一个零, 也就失去了作用。为了解决这个问题, 我们可以先将任意值赋值给另一个未知数, 而不是 x_n 。我们还可以尝试删去 (8.38) 的其他方程, 而不是最后一个方程。我们要寻找的是一个不会消失的 $n-1$ 阶行列式, 它由系统 (8.37) 的系数行列式通过删除一行和一列而形成。如果存在这样的行列式, 那么按照给出的步骤, 正确选择要舍弃的方程, 正确选择要赋值的未知数, 我们就可以求解方程组, 并得到形式为 (8.39) 的解。如果不存在这样的行列式, 我们就必须给两个未知数任意赋值, 并尝试从这里开始求解。因此, (8.37) 的解可能包含两个 (甚至更多) 任意常数。

求解线性齐次方程组的有效方法是对方程进行高斯-若尔当消元。如果只存在平凡解, 最终得到的方程组将是 $x_1 = 0, x_2 = 0, \dots, x_n = 0$ 。如果存在非平凡解, 则至少有一个方程会简化为 $0 = 0$ 的形式; 如果得到 m 个 $0 = 0$ 形式的方程, 我们就为其中的 m 个未知数指定任意常数, 并用这 m 个未知数来表示其余的未知数。

例题: 使用高斯-若尔当消元法求解上例中的方程组。

在使用高斯消元法或高斯-若尔当消元法求解包含 n 个齐次或非齐次方程的方程组时, 通过略去变量 x_1, \dots, x_n , 只写出包含第 n 行和第 $(n+1)$ 列系数和常数 (包含任意零系数) 组成的阵列, 我们可以省去不必要的过程; 然后, 我们对每一行的数字进行运算, 就好像这一行就是它所代表的等式一样, 从而产生下一个阵列。

为了开始除法消元, 我们交换第一个方程和第二个方程, 从而以 $a_{11} = 1$ 开始。分离系数并进行高斯-若尔当消元, 我们得到

$$\begin{array}{cccc|cccc|cccc|cccc} 1 & 3 & -2 & 0 & & 1 & 3 & -2 & 0 & & 1 & 3 & -2 & 0 & & 1 & 0 & \frac{11}{5} & 0 \\ 3 & 4 & 1 & 0 & \rightarrow & 0 & -5 & 7 & 0 & \rightarrow & 0 & -1 & -\frac{7}{5} & 0 & \rightarrow & 0 & 1 & -\frac{7}{5} & 0 \\ 1 & -2 & 5 & 0 & & 0 & -5 & 7 & 0 & & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 \end{array}$$

第一个阵列是原方程组, 但第一个和第二个方程互换。为了将 x_1 从第二和第三个方程中消去, 我们将第二个方程乘以 -3 并加到第一个方程上, 将第一个方程乘以 -1 并加到第三个方程上, 这就得到了第二个阵列。对第二行除以 5 得到第三个阵列。为了将 x_2 从第一和第三个方程中消去, 我们将第二个方程乘以 -3 并加到第一个方程上, 将第二个方程乘以 $-7/5$ 并加到第三个方程上, 得到第四个阵列。由于第四个阵列中 x_3 的系数为零, 我们不能使用第三个方程从第一和第二个方程中消去 x_3 (这原本是高斯-若尔当消元法的下一步)。因此, 我们将 $0 = 0$ 的第三个方程删去, 设 $x_3 = k$, 其中 k 是任意常数。最后一个阵列中的第一和第二个方程有 $x_1 + \frac{11}{5}x_3 = 0$ 和 $x_2 - \frac{7}{5}x_3 = 0$, 或 $x_1 = -\frac{11}{5}x_3$ 和 $x_2 = \frac{7}{5}x_3$ 。因此, 通解为 $x_1 = -\frac{11}{5}k$, $x_2 = \frac{7}{5}k$, $x_3 = k$ 。

8.5 线性变分函数

在分子研究中被广泛使用的一种变分函数为线性变分函数。线性变分函数 (linear variational function) 是 n 个线性无关函数 f_1, f_2, \dots, f_n 的线性组合:

$$\phi = c_1 f_1 + c_2 f_2 + \dots + c_n f_n = \sum_{j=1}^n c_j f_j \quad (8.40)$$

其中 ϕ 是试探变分函数, 系数 c_j 是待确定使得变分积分最小的参数。函数 f_j [称为基函数 (basis functions)] 必须满足问题的定解条件。我们将只讨论实函数 ϕ , 则 c_j 和 f_j 分别是实数和实函数。在 (8.40) 中, f_j 是已知函数。

我们现在应用变分理论 (8.9)。对于线性变分函数，我们有

$$\int \phi^* \phi \, d\tau = \int \sum_{j=1}^n c_j f_j \sum_{k=1}^n c_k f_k \, d\tau = \sum_{j=1}^n \sum_{k=1}^n c_j c_k \int f_j f_k \, d\tau = \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} \quad (8.41)$$

我们定义**重叠积分** (overlap integral) S_{jk} 为

$$S_{jk} \equiv \int f_j^* f_k \, d\tau \quad (8.42)$$

注意： S_{jk} 不一定等于 δ_{jk} ，因此，没有理由认为函数组 f_j 是相互正交的。它们也不一定是任何一个算符的本征函数。(8.9) 的分母为

$$\begin{aligned} \int \phi^* \hat{H} \phi \, d\tau &= \int \sum_{j=1}^n c_j f_j \hat{H} \sum_{k=1}^n c_k f_k \, d\tau \\ &= \sum_{j=1}^n \sum_{k=1}^n c_j c_k \int f_j \hat{H} f_k \, d\tau = \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk} \end{aligned}$$

其中，我们定义 H_{jk} 为

$$H_{jk} \equiv \int f_j^* \hat{H} f_k \, d\tau \quad (8.43)$$

那么变分积分 W 为

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} = \frac{\sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk}}{\sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk}} \quad (8.44)$$

$$W \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk} \quad (8.45)$$

现在我们将 W 最小化，以尽可能接近 E_1 ($W \geq E_1$)。变分积分 W 是 n 个独立变量 c_1, c_2, \dots, c_n 的函数：

$$W = W(c_1, c_2, \dots, c_n)$$

多个变量的函数 W 的最小值的一个必要条件是：在最小值处，函数 W 对每个变量的偏导数必须为零：

$$\frac{\partial W}{\partial c_i} = 0, \quad i = 1, 2, \dots, n \quad (8.46)$$

我们现在将 (8.45) 对每个变量 c_i 求偏导数，获得 n 个方程

$$\frac{\partial W}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} + W \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k H_{jk}, \quad i = 1, 2, \dots, n \quad (8.47)$$

那么

$$\frac{\partial}{\partial c_i} \sum_{j=1}^n \sum_{k=1}^n c_j c_k S_{jk} = \sum_{j=1}^n \sum_{k=1}^n \left[\frac{\partial}{\partial c_i} (c_j c_k) \right] S_{jk} = \sum_{j=1}^n \sum_{k=1}^n \left(c_k \frac{\partial c_j}{\partial c_i} + c_j \frac{\partial c_k}{\partial c_i} \right) S_{jk}$$

c_j 是独立变量，因此

$$\begin{aligned} \frac{\partial c_j}{\partial c_i} &= 0 \quad \text{若 } j \neq i, \quad \frac{\partial c_i}{\partial c_i} = 1 \quad \text{若 } j = i \\ \frac{\partial c_j}{\partial c_i} &= \delta_{ij} \end{aligned} \quad (8.48)$$

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W & \cdots & H_{1n} - S_{1n}W \\ H_{21} - S_{21}W & H_{22} - S_{22}W & \cdots & H_{2n} - S_{2n}W \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} - S_{n1}W & H_{n2} - S_{n2}W & \cdots & H_{nn} - S_{nn}W \end{vmatrix} = 0 \quad (8.58)$$

将 (8.58) 中的行列式展开, 我们得到关于未知数 W 的 n 次代数方程。这个方程有 n 个根, 可以证明它们都是实数。将这些根按照能量依次递增的顺序排列, 我们将其表示为

$$W_1 \leq W_2 \leq \cdots \leq W_n \quad (8.59)$$

如果我们把系统的束缚态按能量递增的顺序排列, 就会得出

$$E_1 \leq E_2 \leq \cdots \leq E_n \leq E_{n+1} \leq \cdots \quad (8.60)$$

其中的 E 表示各种状态的真是能量值。根据变分原理, 我们知道 $E_1 \leq W_1$ 。此外, 可以证明 [J. K. L. MacDonald, *Phys. Rev.*, **43**, 830 (1933); R. H. Young, *Int. J. Quantum Chem.*, **6**, 596 (1972); 见问题. 8.40]:

$$E_1 \leq W_1, \quad E_2 \leq W_2, \quad E_3 \leq W_3, \dots, \quad E_n \leq W_n \quad (8.61)$$

因此, 线性变分法提供了系统最低 n 个束缚态的能量上限。我们使用 W_1, W_2, \dots, W_n 来表示最低态能量的近似值。如果需要更多状态的能量近似值, 我们可以在试探函数 ϕ 中添加更多的 f_k 。可以证明, 添加更多的函数 f_k 可以提高 (或不改变) 先前计算的能量的精确度。若 $\phi = \sum_k c_k f_k$ 中的 f_k 构成一个完全集, 那么我们就能够得到系统的真实波函数。不幸的是, 我们通常需要无数个函数才能得到一个完全集。

量子化学家可能会在线性变分函数中使用几十、几百、几千甚至几百万个项, 从而得到分子的精确结果。解 (8.58) [称为久期方程 (secular equation)] 及相关线性方程 (8.58) 的一种有效方法是矩阵法 (第 8.6 节)。

为了获得基态波函数的近似, 我们使用久期方程最小的根 W_1 , 将其带入方程组 (8.55), 解得系数 $c_1^{(1)}, c_2^{(1)}, \dots, c_n^{(1)}$, 其中我们添加了上标 (1) 来与 W_1 相对应。[如前所述, 我们只能确定系数之间的比值。根据 $c_1^{(1)}$ 的值, 我们可以得到 $c_2^{(1)}, c_3^{(1)}, \dots, c_n^{(1)}$ 的值, 再通过归一化条件确定 $c_1^{(1)}$ 。] 找到了 $c_k^{(1)}$ 之后, 我们将 $\phi_1 = \sum_{k=1}^n c_k^{(1)} f_k$ 作为基态波函数的近似。使用 (8.58) 中更高的根并带入 (8.55), 可以得到激发态波函数的近似。可以证明, 这些近似波函数是正交的 (问题 8.40)。

求解 (8.58) 和 (8.55) 时, 可通过使尽可能多的积分等于零来简化。我们可以通过将函数 f_k 选为与 \hat{H} 相乘的某个算符 \hat{A} 的本征函数, 使一些对角线外的 H_{ij} 消失。如果 f_i 和 f_j 对应于 \hat{A} 的不同本征值, 则 H_{ij} 消失 (第 7.4 节的定理 6)。如果函数 f_k 是正交的, 则对角线外的 S_{ij} 消失 ($S_{ij} = \delta_{ij}$)。如果最初选择的 f_k 不是正交的, 我们可以使用施密特 (或其他) 程序找出这些 f_k 的 n 个正交线性组合, 然后使用正交化的函数。

当取消变分函数为实函数的限制时, 方程 (8.55) 和 (8.58) 也是成立的 (问题 8.39)。

例题: 给第 8.1 节第一个例题中的函数 $x(l-x)$ 添加函数, 使其形成粒子在长度为 l 的一维势箱中的线性变分函数, 并求出最低四个状态的近似能量和波函数。

在变分函数 $\phi = \sum_{k=1}^n c_k f_k$ 中, 我们选择 $f_1 = x(l-x)$ 。由于我们希望知道最低四个状态的近似, n 应至少为 4。有无数种可能的品优函数可以用于 f_2, f_3 和 f_4 。函数 $x^2(l-x)^2$ 满足在 $x=0$ 和 $x=l$ 处消失的定解条件, 且对其积分较为简单, 所以我们选择 $f_2 = x^2(l-x)^2$ 。

如果把原点放在势箱的中心, 势能 (图 2.1) 就是偶函数, 正如第 8.2 节所述, 波函数在偶函数和奇函数之间交替变化 (另见图 2.3)。(在本例中, 偶函数和奇函数指的是原点位于势箱

中心。) 函数 $f_1 = x(l-x)$ 和 $f_2 = x^2(l-x)^2$ 都是偶函数 (见问题 8.34)。若我们选择 $\phi = c_1 x(l-x) + c_2 x^2(l-x)^2$, 我们最终会得到波函数为偶函数的最低两个态的能量上限 ($n=1$ 和 $n=3$) 并得到这两个状态波函数的近似。由于我们还需要 $n=2$ 和 $n=4$ 状态的近似, 我们需要增加两个奇函数。如方程 (4.50) 之后所述, 奇函数必须在原点为零, 所以我们必须要求函数在势箱中点 $x = \frac{1}{2}l$ 处为零, 与 $x=0$ 和 $x=l$ 一致。满足这些性质的一个简单函数为 $f_3 = x(l-x)(\frac{1}{2}l-x)$ 。为了得到 f_4 , 我们将 f_2 乘以 $(\frac{1}{2}l-x)$ 。因此, 由 $\phi = \sum_{k=1}^4 c_k f_k$ 组成的线性变分函数, 其中

$$f_1 = x(l-x), f_2 = x^2(l-x)^2, f_3 = x(l-x)\left(\frac{1}{2}l-x\right), f_4 = x^2(l-x)^2\left(\frac{1}{2}l-x\right) \quad (8.62)$$

如 (8.40) 所假设的那样, f_1 、 f_2 、 f_3 和 f_4 线性无关。

因为 f_1 和 f_2 是偶函数, f_3 和 f_4 是奇函数, 所以许多积分都为零。因此,

$$S_{13} = S_{31} = 0, \quad S_{14} = S_{41} = 0, \quad S_{23} = S_{32} = 0, \quad S_{24} = S_{42} = 0 \quad (8.63)$$

因为这些重叠积分中的每一个积分的被积函数都是相对于方框中心原点的奇函数。函数 f_1 、 f_2 、 f_3 和 f_4 都是宇称算符 $\hat{\Pi}$ 的本征函数 (第7.5节), 其中, 偶函数 f_1 和 f_2 对应于本征值 $+1$, 奇函数 f_3 和 f_4 对应于本征值 -1 。由于 V 是偶函数, 所以 \hat{H} 和 $\hat{\Pi}$ 对易, 根据 (7.4) 节的定理 6, 若 f_i 是奇函数而 f_j 是偶函数时, H_{ij} 为零, 反之亦然。因此,

$$H_{13} = H_{31} = 0, \quad H_{14} = H_{41} = 0, \quad H_{23} = H_{32} = 0, \quad H_{24} = H_{42} = 0 \quad (8.64)$$

根据 (8.63) 和 (8.64), $n=4$ 的久期方程 (8.58) 变为

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W & 0 & 0 \\ H_{21} - S_{21}W & H_{22} - S_{22}W & 0 & 0 \\ 0 & 0 & H_{33} - S_{33}W & H_{34} - S_{34}W \\ 0 & 0 & H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0 \quad (8.65)$$

久期行列式为块对角行列式, 其值等于各个块的乘积 [式 (8.34)]:

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W \\ H_{21} - S_{21}W & H_{22} - S_{22}W \end{vmatrix} \times \begin{vmatrix} H_{33} - S_{33}W & H_{34} - S_{34}W \\ H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0$$

该方程的四个根由以下两个方程给出:

$$\begin{vmatrix} H_{11} - S_{11}W & H_{12} - S_{12}W \\ H_{21} - S_{21}W & H_{22} - S_{22}W \end{vmatrix} = 0 \quad (8.66)$$

$$\begin{vmatrix} H_{33} - S_{33}W & H_{34} - S_{34}W \\ H_{43} - S_{43}W & H_{44} - S_{44}W \end{vmatrix} = 0 \quad (8.67)$$

令 (8.66) 的根为 W_1 和 W_2 , 它们分别代表 $n=1$ 和 $n=3$ 的近似能量, 同样地, 令 (8.67) 的根为 W_3 和 W_4 。在通过久期方程解出 W 后, 我们将其逐一带入方程组 (8.55) 以求得变分函数中的系数 c_k 。根据久期方程 (8.65), 带入 W_1 的方程组 (8.55) 为

$$\begin{cases} (H_{11} - S_{11}W_1)c^{(1)} + (H_{12} - S_{12}W_1)c^{(1)} = 0 \\ (H_{21} - S_{21}W_1)c^{(1)} + (H_{22} - S_{22}W_1)c^{(1)} = 0 \end{cases} \quad (8.68a)$$

$$\left. \begin{aligned} (H_{33} - S_{33}W_1)c^{(1)} + (H_{34} - S_{34}W_1)c^{(1)} &= 0 \\ (H_{43} - S_{43}W_1)c^{(1)} + (H_{44} - S_{44}W_1)c^{(1)} &= 0 \end{aligned} \right\} \quad (8.68b)$$

由于 W_1 是 (8.66) 的根, 方程组 (8.68a) 的系数行列式 [即 (8.66) 中的行列式] 为零。那么 (8.68a) 有非平凡解 $c_1^{(1)}$ 和 $c_2^{(1)}$, 然而, W_1 不是 (8.67) 的根, 所以方程组 (8.68b) 的系数行列式不为零。因此, (8.68b) 只有平凡解 $c_3^{(1)} = c_4^{(1)} = 0$ 。则对应于根 W_1 的试探函数 ϕ_1 的形式为 $\phi_1 = \sum_{k=1}^4 c_k^{(1)} f_k = c_1^{(1)} f_1 + c_2^{(1)} f_2$ 。根据同样的理由说明 ϕ_3 是 f_1 和 f_2 的线性组合, 而 ϕ_2 和 ϕ_4 分别是 f_3 和 f_4 的线性组合:

$$\begin{aligned} \phi_1 &= c_1^{(1)} f_1 + c_2^{(1)} f_2, & \phi_3 &= c_1^{(3)} f_1 + c_2^{(3)} f_2 \\ \phi_2 &= c_3^{(2)} f_3 + c_4^{(2)} f_4, & \phi_4 &= c_3^{(4)} f_3 + c_4^{(4)} f_4 \end{aligned} \quad (8.69)$$

偶波函数 ψ_1 和 ψ_3 使用偶函数 f_1 和 f_2 的线性组合来近似, 而奇波函数 ψ_2 和 ψ_4 使用奇函数 f_3 和 f_4 的线性组合来近似。

当久期方程为块对角形式时, 它可以分解成两个或多个较简单的久期方程, 联立方程组 (8.55) 也可以分解成两个或多个较简单的方程组。

为了从 (8.66) 和 (8.67) 中解出 W_1 、 W_2 、 W_3 和 W_4 , 我们现在必须计算 H_{ij} 和 S_{ij} 。我们有

$$\begin{aligned} H_{11} &= \langle f_1 | \hat{H} | f_1 \rangle = \int_0^l x(l-x) \left(\frac{-\hbar^2}{2m} \right) \frac{d^2}{dx^2} [x(l-x)] dx = \frac{\hbar^2 l^3}{6m} \\ S_{11} &= \langle f_1 | f_1 \rangle = \int_0^l x^2(l-x)^2 dx = \frac{l^5}{30} \end{aligned}$$

其中用到了 (8.12) 和 (8.13)。使用 (8.62)、(8.49) 和 (8.52), 我们可以计算出剩余的积分 (问题 8.35):

$$\begin{aligned} H_{12} &= H_{21} = \langle f_2 | \hat{H} | f_1 \rangle = \hbar^2 l^5 / 30m, & H_{22} &= \hbar^2 l^7 / 105m \\ H_{33} &= \hbar^2 l^5 / 40m, & H_{44} &= \hbar^2 l^9 / 1260m, & H_{34} &= H_{43} = \hbar^2 l^7 / 280m \\ S_{12} &= S_{21} = \langle f_1 | f_2 \rangle = l^7 / 140, & S_{22} &= l^9 / 630 \\ S_{33} &= l^7 / 840, & S_{44} &= l^{11} / 27720, & S_{34} &= S_{43} = l^9 / 5040 \end{aligned}$$

方程 (8.66) 变为

$$\begin{vmatrix} \frac{\hbar^2 l^3}{6m} - \frac{l^5}{30} W & \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140} W \\ \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140} W & \frac{\hbar^2 l^7}{105m} - \frac{l^9}{630} W \end{vmatrix} = 0 \quad (8.70)$$

使用第 8.3 节的定理 4, 为了消去分式, 我们将行列式的第一行乘以 $420m/l^3$, 第二行乘以 $1260m/l^5$, 右侧乘以相同的因子, 得到

$$\begin{vmatrix} 70\hbar^2 - 14ml^2 W & 14\hbar^2 l^2 - 3ml^4 W \\ 42\hbar^2 - 9ml^2 W & 12\hbar^2 l^2 - 2ml^4 W \end{vmatrix} = 0$$

$$m^2 l^4 W^2 - 56ml^2 \hbar^2 W + 252\hbar^4 = 0$$

$$W = (\hbar^2/ml^2) (28 \pm \sqrt{532}) = 0.1250018\hbar^2/ml^2, \quad 1.293495\hbar^2/ml^2 \quad (8.71)$$

将积分带入 (8.67), 得到根 (问题 8.36):

$$W = (\hbar^2/ml^2) (60 \pm \sqrt{1620}) = 0.5002930\hbar^2/ml^2, \quad 2.5393425\hbar^2/ml^2 \quad (8.72)$$

则 $(ml^2/h^2)W$ 的近似值为 0.1250018、1.293495、0.5002930 和 2.5393425，与最低四个能级的准确值 [方程 (2.20)] $(ml^2/h^2)E = 0.125, 0.5, 1.125$ 和 2 相比较。对于 $n = 1, 2, 3$ 和 4，百分误差分别为 0.0014%、0.059%、15.0% 和 27.0%。 $n = 1$ 和 2 的近似非常好，而 $n = 3$ 和 4 极差 (lousy)。

我们现在尝试求出对应于这些 W 值的近似波函数。将 $W_1 = 0.1250018\hbar^2/ml^2$ 代入对应于 (8.71) 的方程组 (8.68a)，我们得到 (同除 h^2 后)

$$\begin{aligned} 0.023095c_1^{(1)} - 0.020381c_2^{(1)}l^2 &= 0 \\ -0.061144c_1^{(1)} + 0.053960c_2^{(1)}l^2 &= 0 \end{aligned} \quad (8.73)$$

例如，其中的第一个系数由下式求出：

$$70\hbar^2 - 14ml^2W_1 = 70\hbar^2/4\pi^2 - 14(0.1250018) = 0.023095\hbar^2$$

要求解齐次方程 (8.73)，我们可以按照第 8.4 节末尾给出的步骤进行。将 (8.73) 的第二个方程删去，将 $c_2^{(1)}$ 项移到等号右边，解出系数的比值，我们得到

$$c_1^{(1)} = k, \quad c_2^{(1)} = 1.133k/l^2$$

其中的 k 是常数。我们通过归一化条件来求出 k ：

$$\begin{aligned} \langle \phi_1 | \phi_1 \rangle &= 1 = \langle kf_1 + 1.133kf_2/l^2 | kf_1 + 1.133kf_2/l^2 \rangle \\ &= k^2 (\langle f_1 | f_1 \rangle + 2.266 \langle f_1 | f_2 \rangle / l^2 + 1.284 \langle f_2 | f_2 \rangle / l^4) \\ &= k^2 (S_{11} + 2.266S_{12}/l^2 + 1.284S_{22}/l^4) = 0.05156k^2l^5 \end{aligned}$$

其中用到了之前求出的重叠积分值。因此 $k = 4.404/l^{5/2}$ ，

$$\begin{aligned} \phi_1 &= c_1^{(1)}f_1 + c_2^{(1)}f_2 = 4.404f_1/l^{5/2} + 4.990f_2/l^{9/2} \\ \phi_1 &= l^{-1/2} [4.404(x/l)(1-x/l) + 4.990(x/l)^2(1-x/l)^2] \end{aligned} \quad (8.74)$$

其中用到了 (8.62)。

在 (8.55) 中轮流使用 W_2 、 W_3 和 W_4 ，我们可以得到以下归一化线性变分函数 (问题 8.38)，其中 $X \equiv x/l$ ：

$$\begin{aligned} \phi_2 &= l^{-1/2} \left[16.78X(1-X) \left(\frac{1}{2} - X \right) + 71.85X^2(1-X^2) \left(\frac{1}{2} - X \right) \right] \\ \phi_3 &= l^{-1/2} [28.65X(1-X) - 132.7X^2(1-X)^2] \\ \phi_4 &= l^{-1/2} \left[98.99X(1-X) \left(\frac{1}{2} - X \right) - 572.3X^2(1-X)^2 \left(\frac{1}{2} - X \right) \right] \end{aligned} \quad (8.75)$$

8.6 矩阵、本征值和本征向量

矩阵代数 (第 7.10 节) 是英国数学家阿瑟·凯利 (Arthur Cayley) 在 1855-1858 年间发展起来的，是处理同时线性方程组和从一组变量到另一组变量的线性变换的简便方法。[在 1860 年获得教授职位之前的 14 年里，凯利一直靠当律师养活自己。在做律师期间，他发表了数百篇数学论文，并经常与他的律师同事约翰·约瑟夫·西尔维斯特 (John Joseph Sylvester) 讨论数学问题。西尔维斯特为矩形阵列创造了“矩阵”一词 (取自拉丁文中 “womb” 一词)，并于 1855 年获得数学教授职

位。] 多年来, 大多数物理学家都不知道矩阵代数, 当海森堡在 1925 年发现矩阵力学版本的量子力学时, 他并没有意识到他所处理的实体是矩阵。当他向玻恩展示自己的研究成果时, 受过良好数学训练的玻恩认出海森堡使用的是矩阵。如今, 矩阵代数已成为量子化学计算的重要工具, 并广泛应用于物理学的大多数分支。

线性非齐次方程组 (8.55) 可以写成矩阵形式:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (8.76)$$

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (8.77)$$

其中 \mathbf{A} 是系数矩阵, \mathbf{x} 和 \mathbf{b} 都是列向量。利用矩阵乘法法则 (7.107) 很容易验证 (8.35) 和 (8.76) 的等价性。

方阵 \mathbf{A} 的行列式 (determinant) 是一个所有元素与 \mathbf{A} 的元素完全一致的行列式。若 $\det \mathbf{A} \neq 0$, 则称矩阵 \mathbf{A} 是非奇异的 (nonsingular)。

n 阶方阵的逆矩阵 (inverse matrix) 定义为与 \mathbf{A} 相乘得到单位矩阵的矩阵。将逆矩阵记为 \mathbf{A}^{-1} , 那么

$$\boxed{\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}} \quad (8.78)$$

其中 \mathbf{I} 是单位阵。可以证明, 当且仅当 $\det \mathbf{A} \neq 0$ 时, \mathbf{A}^{-1} 才存在。(对于求 \mathbf{A}^{-1} 的高效方法, 见 *Press et al.*, Section 2.3 ; *Shoup*, Section 3.3; 问题 8.51。许多电子表格都内置了计算逆矩阵的功能。)

若 (8.76) 中系数矩阵 \mathbf{A} 的 $\det \mathbf{A} \neq 0$, 则可以通过将 \mathbf{A}^{-1} 乘以 (8.77) 的两边来得到 $\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ 。由于矩阵乘法满足结合律 (第 7.10 节), 我们得到了 $\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$ 。因此, 将 (8.77) 两边乘以 \mathbf{A}^{-1} , 我们得到 $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ 作为线性非齐次方程组的一组解。

线性变分法被广泛用于求近似分子波函数, 而矩阵代数给出了求解线性变分法方程的计算效率最高的方法。若线性变分函数 $\phi = \sum_{k=1}^n c_k f_k$ 中的 f_1, \dots, f_n 是正交的, 那么 $S_{ij} = \int f_i^* f_j d\tau = \delta_{ij}$, 则使变分积分最小的系数 c_k 构成的齐次方程组 (8.55) 变为

$$\begin{aligned} H_{11}c_1 + H_{12}c_2 + \cdots + H_{1n}c_n &= Wc_1 \\ H_{21}c_1 + H_{22}c_2 + \cdots + H_{2n}c_n &= Wc_2 \\ &\vdots \\ H_{n1}c_1 + H_{n2}c_2 + \cdots + H_{nn}c_n &= Wc_n \end{aligned} \quad (8.79a)$$

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = W \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad (8.79b)$$

$$\mathbf{H}\mathbf{c} = W\mathbf{c} \quad (8.79c)$$

其中 \mathbf{H} 是方阵, 其元素为 $H_{ij} = \langle f_i | \hat{H} | f_j \rangle$, \mathbf{c} 是系数 c_1, \dots, c_n 组成的列向量。

若

$$\boxed{\mathbf{A}\mathbf{c} = \lambda\mathbf{c}} \quad (8.80)$$

其中 \mathbf{A} 是方阵, \mathbf{c} 是至少有一个元素不为零的列向量, λ 是标量, 则我们称 \mathbf{c} 为本征向量 (eigenvector) 或特征向量 (characteristic vector), λ 为 \mathbf{A} 的本征值 (eigenvalue) 或特征值 (characteristic value)。

将 (8.80) 与 (8.79c) 相比较, 求解 $S_{ij} = \delta_{ij}$ 的线性变分问题等价于求解矩阵 \mathbf{H} 的本征值和本征向量。矩阵本征方程 $\mathbf{H}\mathbf{c} = W\mathbf{c}$ 等价于齐次方程组 (8.55), 当且仅当 $\det(H_{ij} - \delta_{ij}W) = 0$ 时的 c 有非平凡解 [$S_{ij} = \delta_{ij}$ 的方程 (8.57)]。对于一般的 n 阶方阵 \mathbf{A} , 本征值对应满足的方程为

$$\det(A_{ij} - \delta_{ij}\lambda) = 0 \quad (8.81)$$

方程 (8.81) 称为矩阵 \mathbf{A} 的特征方程 (characteristic equation)。当我们展开 (8.81) 中的 n 阶行列式时, 就得到了一个 n 次多项式 [称为特征多项式 (characteristic polynomial)], 其最高次项为 λ^n 。特征多项式对 λ 有 n 个根 (有些可能重合, 有些可能为虚数), 所以一个 n 阶方阵有 n 个本征值。[方阵 \mathbf{A} 的本征值组有时可以称为 \mathbf{A} 的谱 (spectrum)。]

\mathbf{H} 的矩阵方程 (8.79c) 与 \mathbf{A} 的矩阵方程 (8.80) 相对应。 \mathbf{A} 的特征向量元素满足以下方程组, 对应于 (8.79a):

$$\begin{aligned} A_{11}c_1 + A_{12}c_2 + \cdots + A_{1n}c_n &= \lambda c_1 \\ &\vdots \\ A_{n1}c_1 + A_{n2}c_2 + \cdots + A_{nn}c_n &= \lambda c_n \end{aligned} \quad (8.82)$$

对于每个不同的本征值, 我们都有一组不同的方程 (8.82) 和一组不同的数 c_1, c_2, \dots, c_n , 从而得到一个不同的本征向量。

如果矩阵的所有本征值都不同, 则可以证明求解独立本征向量 (见 *Strang*, Section 5.2), 这里的线性无关是指没有导致 n 个线性本征向量可以写成其他本征向量的线性组合。如果某些本征值相等, 那么矩阵的线性无关本征向量可能少于 n 个。量子力学中出现的矩阵通常是厄米矩阵 (本节稍后将对这一术语进行定义), 即使某些本征值相等, 阶数为 n 的厄米矩阵也总是具有 n 个线性无关的本征向量 (证明见 *Strang*, Section 5.6)。

若 \mathbf{A} 是对角矩阵 (对 $i \neq j$ 有 $a_{ij} = 0$), 则 (8.81) 中的行列式为对角行列式。对角行列式的值等于其对角元素的乘积 [式 (8.33)], 所以对角矩阵的特征方程为

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0$$

该方程的根为 $\lambda_1 = a_{11}$ 、 $\lambda_2 = a_{22}$ 、 \dots 、 $\lambda_n = a_{nn}$ 。对角矩阵的本征值就是其对角元素。(对于本征向量, 见问题 8.46。)

若 \mathbf{c} 是 \mathbf{A} 的本征向量, 很明显, $\mathbf{d} \equiv k\mathbf{c}$ 也是 \mathbf{A} 的本征向量, 其中 k 是任意常数。若我们选择的 k 满足

$$\sum_{i=1}^n |d_i|^2 = 1 \quad (8.83)$$

则列向量 \mathbf{d} 是归一化的 (normalized)。若两个分别有 n 个元素的列向量 \mathbf{b} 和 \mathbf{c} 满足下式, 则称它们是正交的 (orthogonal):

$$\sum_{i=1}^n b_i^* c_i = 0 \quad (8.84)$$

令变分方程 (8.79) 中 \mathbf{H} 的 n 个本征值为 W_1, W_2, \dots, W_n , 对应的本征向量为 $\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(n)}$ 。则

$$\mathbf{H}\mathbf{c}^{(i)} = W_i \mathbf{c}^{(i)}, \quad i = 1, 2, \dots, n \quad (8.85)$$

其中 $\mathbf{c}^{(i)}$ 是列向量, 其元素为 $c_1^{(i)}, c_2^{(i)}, \dots, c_n^{(i)}$, 基函数 f_1 是正交的。此外, 设 \mathbf{C} 为方阵, 其列为 \mathbf{H} 的本征向量; 设 \mathbf{W} 为对角矩阵, 其对角元素为 \mathbf{H} 的本征值:

$$\mathbf{C} = \begin{pmatrix} c_1^{(1)} & c_1^{(2)} & \cdots & c_1^{(n)} \\ c_2^{(1)} & c_2^{(2)} & \cdots & c_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ c_n^{(1)} & c_n^{(2)} & \cdots & c_n^{(n)} \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_n \end{pmatrix} \quad (8.86)$$

则 n 个本征方程组 (8.85) 可以写成单个方程:

$$\mathbf{H}\mathbf{C} = \mathbf{C}\mathbf{W} \quad (8.87)$$

为了验证矩阵方程 (8.87), 我们需要证明矩阵 $\mathbf{H}\mathbf{C}$ 的每个元素 $(\mathbf{H}\mathbf{C})_{ij}$ 等于矩阵 $\mathbf{C}\mathbf{W}$ 对应的元素 $(\mathbf{C}\mathbf{W})_{ij}$ 。根据矩阵乘法规则 (7.107), 我们有 $(\mathbf{H}\mathbf{C})_{ij} = \sum_k H_{ik} (\mathbf{C})_{kj} = \sum_k H_{ik} c_k^{(j)}$ 。考虑本征方程 $\mathbf{H}\mathbf{C}^{(j)} = W_j \mathbf{C}^{(j)}$ [式 (8.85)], $\mathbf{H}\mathbf{C}^{(j)}$ 和 $W_j \mathbf{C}^{(j)}$ 都是列矩阵。利用 (??) 令每个列矩阵第 i 行中的元素相等, 我们有 $\sum_k H_{ik} c_k^{(j)} = W_j c_i^{(j)}$ 。则

$$(\mathbf{H}\mathbf{C})_{ij} = \sum_k H_{ik} c_k^{(j)} = W_j c_i^{(j)}, \quad (\mathbf{C}\mathbf{W})_{ij} = \sum_k (\mathbf{C})_{ik} (\mathbf{W})_{kj} = \sum_k c_i^{(k)} \delta_{kj} W_k = c_i^{(j)} W_j$$

则 $(\mathbf{H}\mathbf{C})_{ij} = (\mathbf{C}\mathbf{W})_{ij}$, 从而验证了矩阵方程 (8.87)。

令矩阵 \mathbf{C} 有逆阵 (见下文), 我们可以将 (8.87) 两边乘以 \mathbf{C}^{-1} , 得到 $\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{C}^{-1}(\mathbf{C}\mathbf{W})$ 。[由于矩阵乘法不可对易, 当我们将 $\mathbf{H}\mathbf{C} = \mathbf{C}\mathbf{W}$ 两边乘以 \mathbf{C}^{-1} 时, 必须把因子 \mathbf{C}^{-1} 放在 $\mathbf{H}\mathbf{C}$ 和 $\mathbf{C}\mathbf{W}$ 的左边 (或右边)。] 得到了 $\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{C}^{-1}(\mathbf{C}\mathbf{W}) = (\mathbf{C}^{-1}\mathbf{C})\mathbf{W} = \mathbf{I}\mathbf{W} = \mathbf{W}$:

$$\mathbf{C}^{-1}\mathbf{H}\mathbf{C} = \mathbf{W} \quad (8.88)$$

为了化简 (8.88), 我们必须了解更多关于矩阵的知识。

若方阵 \mathbf{B} 的所有元素满足 $b_{mn} = b_{nm}$, 则 \mathbf{B} 为对称矩阵 (symmetric matrix)。对称矩阵的元素关于主对角线对称, 例如, $b_{12} = b_{21}$ 。若方阵 \mathbf{D} 的所有元素满足 $d_{mn} = d_{nm}^*$, 则 \mathbf{D} 为厄米矩阵 (Hermitian matrix)。例如, 若

$$\mathbf{M} = \begin{pmatrix} 2 & 5 & 0 \\ 5 & i & 2i \\ 0 & 2i & 4 \end{pmatrix}, \quad \mathbf{N} = \begin{pmatrix} 6 & 1+2i & 8 \\ 1-2i & -1 & -i \\ 8 & i & 0 \end{pmatrix} \quad (8.89)$$

则 \mathbf{M} 是对称矩阵, \mathbf{N} 是厄米矩阵。(注意: 厄米矩阵的对角元素一定是实数: $d_{mm} = d_{mm}^*$ 。) 实矩阵 (real matrix) 是指所有元素都是实数的矩阵。实厄米矩阵是对称矩阵。

矩阵的转置 (transpose) \mathbf{A}^T (也写作 $\tilde{\mathbf{A}}$) 是通过交换 \mathbf{A} 的行和列, 使第 1 列变为第 1 行, 第 2 列变为第 2 行, 以此类推而形成的矩阵。矩阵 \mathbf{A}^T 的元素 a_{mn}^T 与 \mathbf{A} 的元素关系为 $a_{mn}^T = a_{nm}$ 。对于方阵, 转置是通过将元素关于主对角线反映得到的。对称矩阵的转置等于其本身。因此, 对于 (8.89) 中的矩阵 \mathbf{M} , 我们有 $\mathbf{M}^T = \mathbf{M}$ 。

矩阵 \mathbf{A} 的复共轭 (complex conjugate) \mathbf{A}^* (国内主要写为 $\bar{\mathbf{A}}$, 译者注, 下同) 是通过将 \mathbf{A} 的每个元素取复共轭得到的矩阵。矩阵 \mathbf{A} 的共轭转置 (conjugate transpose) \mathbf{A}^\dagger (国内主要写为 \mathbf{A}^H) 是通过将 \mathbf{A}^* 转置得到的矩阵, 因此 $\mathbf{A}^\dagger = (\mathbf{A}^*)^T$ 以及

$$a_{mn}^\dagger = a_{nm}^* \quad (8.90)$$

[物理学家将 \mathbf{A}^\dagger 称为 \mathbf{A} 的伴随矩阵 (adjoint matrix), 是数学家用来表示一种完全不同的矩阵的名称。] 例如,

$$\mathbf{B} = \begin{pmatrix} 2 & 3+i \\ 0 & 4i \end{pmatrix}, \quad \mathbf{B}^T = \begin{pmatrix} 2 & 0 \\ 3+i & 4i \end{pmatrix}, \quad \mathbf{B}^\dagger = \begin{pmatrix} 2 & 0 \\ 3-i & -4i \end{pmatrix}$$

对于厄米矩阵 \mathbf{A} , 根据 (8.90), 我们有 $a_{mn}^\dagger = a_{nm}$, 因此 $(\mathbf{A}^\dagger)_{mn} = (\mathbf{A})_{mn}$ 以及 $\mathbf{A}^\dagger = \mathbf{A}$ 。因此, 厄米矩阵的共轭转置等于它本身。[物理学家通常使用术语自伴随 (self-adjoint) 来描述厄米矩阵。]

方阵的逆矩阵等于其转置的矩阵称为正交矩阵 (orthogonal matrix):

$$\mathbf{A}^{-1} = \mathbf{A}^T \quad \text{若 } \mathbf{A} \text{ 是正交矩阵} \quad (8.91)$$

方阵的逆矩阵等于其共轭转置的矩阵称为酉矩阵 (unitary matrix):

$$\mathbf{U}^{-1} = \mathbf{U}^\dagger \quad \text{若 } \mathbf{U} \text{ 是酉矩阵} \quad (8.92)$$

根据定义 (8.92), 我们得出: 若 \mathbf{U} 是酉矩阵, 则 $\mathbf{U}^\dagger \mathbf{U} = \mathbf{I}$ 。令 $(\mathbf{U}^\dagger \mathbf{U})_{mn} = (\mathbf{I})_{mn}$, 对于酉矩阵的 n 和 m 列, 我们得到 (问题 8.43):

$$\sum_k u_{km}^* u_{kn} = \delta_{mn} \quad (8.93)$$

因此, 根据 (8.84) 和 (8.83) 的定义, 酉矩阵的各列 (视为列向量) 是正交和归一化 (正交) 的。相反地, 若对所有的列都有 (8.93) 成立, 则 \mathbf{U} 是酉矩阵。若 \mathbf{U} 是实酉矩阵, 则 $\mathbf{U}^\dagger = \mathbf{U}^T$, 那么 \mathbf{U} 是正交矩阵。

我们可以证明, 厄米矩阵 \mathbf{H} 对应不同本征值的两个本征向量是正交的 (见 *Strang*, Section 5.5)。对于对应同一本征值的 \mathbf{H} 的本征向量, 可以取它们的线性组合, 这些线性组合将成为 \mathbf{H} 的正交本征向量。此外, 本征向量的元素可以乘以一个常数, 使本征向量归一化。因此, 可以选择厄米矩阵的本征向量为正交向量, 那么 (8.86) 中的本征向量矩阵 \mathbf{C} 是酉矩阵, 则 $\mathbf{C}^{-1} = \mathbf{C}^\dagger$ 。因此, (8.88) 可以写成

$$\mathbf{C}^\dagger \mathbf{H} \mathbf{C} = \mathbf{W} \quad \text{若 } \mathbf{H} \text{ 是厄米矩阵} \quad (8.94)$$

对于 \mathbf{H} 是实厄米矩阵 (\mathbf{H} 是实对称矩阵) 的常见情况, (8.79a) 中的 c 都是实数 (由于 W 和 H_{ij} 都是实数), \mathbf{C} 是实酉矩阵, 即 $\mathbf{C}^{-1} = \mathbf{C}^T$, 则 \mathbf{C} 是正交的。那么 (8.94) 可以写成

$$\mathbf{C}^T \mathbf{H} \mathbf{C} = \mathbf{W} \quad \text{若 } \mathbf{H} \text{ 是实对称矩阵} \quad (8.95)$$

可以证明, 厄米矩阵的本征值是实数 (*Strang*, Section 5.5)。

例题: 通过求解代数方程, 求以下厄米矩阵的本征值和归一化本征向量。并证明 $\mathbf{C}^\dagger \mathbf{A} \mathbf{C}$ 是对角矩阵, 其中 \mathbf{C} 是本征向量矩阵。

$$\mathbf{A} = \begin{pmatrix} 3 & 2i \\ -2i & 0 \end{pmatrix}$$

求本征值 λ 的特征方程 (8.81) 为 $\det(a_{ij} - \delta_{ij}\lambda) = 0$, 带入, 有

$$\begin{vmatrix} 3 - \lambda & 2i \\ -2i & 0 - \lambda \end{vmatrix} = 0$$

$$\lambda^2 - 3\lambda - 4 = 0$$

$$\lambda_1 = 4, \quad \lambda_2 = -1$$

检查本征值计算的一个有用定理如下 (*Strang*, Exercise 5.1.9): n 阶方阵 \mathbf{A} 的本征值 λ_i 之和等于对角元素之和, 即 $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}$ 。在本例中, $\sum_i a_{ii} = 3 + 0$ 等于本征值的和 $4 - 1 = 3$ 。

对于 $\lambda_1 = 4$, 联立方程组 (8.82), 得到

$$\begin{aligned}(3 - \lambda_1) c_1^{(1)} + 2ic_2^{(1)} &= 0 \\ -2ic_1^{(1)} - \lambda_1 c_2^{(1)} &= 0\end{aligned}$$

或

$$\begin{aligned}-c_1^{(1)} + 2ic_2^{(1)} &= 0 \\ -2ic_1^{(1)} - 4c_2^{(1)} &= 0\end{aligned}$$

将这些方程的任意一个删去, 我们得到

$$c_1^{(1)} = 2ic_2^{(1)}$$

根据归一化条件, 有

$$\begin{aligned}1 &= |c_1^{(1)}|^2 + |c_2^{(1)}|^2 = 4|c_2^{(1)}|^2 + |c_2^{(1)}|^2 \\ |c_2^{(1)}|^2 &= 1/\sqrt{5}, \quad c_2^{(1)} = 1/\sqrt{5} \\ c_1^{(1)} &= 2ic_2^{(1)} = 2i/\sqrt{5}\end{aligned}$$

其中我们选择 $c_2^{(1)}$ 的相位为零。

同样地, 对于 $\lambda_2 = -1$, 我们得到 (问题 8.49):

$$c_1^{(2)} = -i/\sqrt{5}, \quad c_2^{(2)} = 2/\sqrt{5}$$

归一化本征向量为

$$\mathbf{c}^{(1)} = \begin{pmatrix} 2i/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}, \quad \mathbf{c}^{(2)} = \begin{pmatrix} -i/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

由于厄米矩阵的两个本征值 λ_1 和 λ_2 不同, 因此它们的本征向量 $\mathbf{c}^{(1)}$ 和 $\mathbf{c}^{(2)}$ 是正交的 (需要你证明)。同样, $\mathbf{c}^{(1)}$ 和 $\mathbf{c}^{(2)}$ 是归一化的。因此, \mathbf{C} 是酉矩阵, $\mathbf{C}^{-1} = \mathbf{C}^\dagger$ 。构造 \mathbf{C} 及其共轭转置, 我们有

$$\begin{aligned}\mathbf{C}^{-1}\mathbf{A}\mathbf{C} &= \mathbf{C}^\dagger\mathbf{A}\mathbf{C} = \begin{pmatrix} 2i/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 3 & 2i \\ -2i & 0 \end{pmatrix} \begin{pmatrix} 2i/\sqrt{5} & -i/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} -2i/\sqrt{5} & 1/\sqrt{5} \\ i/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 8i/\sqrt{5} & i/\sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}\end{aligned}$$

这是本征向量的对角矩阵。

我们已经证明: 若 \mathbf{H} 是实对称矩阵, 本征值为 W_i , 正交本征向量为 $\mathbf{c}^{(i)}$ (即 $\mathbf{H}\mathbf{c}^{(i)} = W_i\mathbf{c}^{(i)}$, $i = 1, 2, \dots, n$), 则 $\mathbf{C}^T\mathbf{H}\mathbf{C} = \mathbf{W}$ [式 (8.95)], 其中 \mathbf{C} 是实正交矩阵, 其列为 $\mathbf{c}^{(i)}$, \mathbf{W} 是本征值 W_i 的对角矩阵。该定理的逆定理很容易证明: 若 \mathbf{H} 是实对称矩阵, \mathbf{B} 是实正交阵, 且 $\mathbf{B}^T\mathbf{H}\mathbf{B} = \mathbf{W}$, 则 \mathbf{B} 的列是 \mathbf{H} 的本征向量, \mathbf{W} 的对角元素是 \mathbf{H} 的本征值。

为了求 n 阶厄米矩阵的本征值和本征向量, 我们可以使用以下任一步骤: (1) 求特征方程 $\det(H_{ij} - \delta_{ij}W) = 0$ [式 (8.81)], 得到本征值 W_1, \dots, W_n 。然后将每个 W_k 带入代数方程组 (8.79a), 解出第 k 个本征向量的元素 $c_1^{(k)}, c_2^{(k)}, \dots, c_n^{(k)}$ 。(2) 寻找酉矩阵 \mathbf{C} , 使得 $\mathbf{C}^\dagger\mathbf{H}\mathbf{C}$ 成为对角矩阵。 $\mathbf{C}^\dagger\mathbf{H}\mathbf{C}$ 的对角元素是 \mathbf{H} 的本征值, \mathbf{C} 的列是 \mathbf{H} 的正交本征向量。对于量子化学中出现的大型矩阵, 步骤 (2) [称为矩阵对角化 (matrix diagonalization)] 的计算速度比 (1) 快得多。

展开特征行列式并求解特征方程不是求大矩阵本征值的好方法，原因之一是，对于大矩阵，特征多项式中一个系数的极小变化可能会导致本征值的巨大变化（见问题 8.54）。因此，我们可能需要将特征多项式中的系数计算到小数点后几百或几千位，才能得到精确到小数点后几位的本征值。虽然对于某些矩阵来说，矩阵元素值的微小变化可能会导致本征值的巨大变化，但我们可以证明：对于厄米矩阵来说，矩阵元素的微小变化始终只会导致本征值的微小变化。因此，上一段的方法（2）是获得准确本征值的正确方法。

实对称矩阵 \mathbf{H} 对角化的系统方法如下。构造一个正交阵 \mathbf{O}_1 使得矩阵 $\mathbf{H}_1 \equiv \mathbf{O}_1^T \mathbf{H} \mathbf{O}_1$ 在 \mathbf{H} 的 H_{12} 和 H_{21} 的对角元素为零。（由于 \mathbf{H} 是对称的，我们有 $H_{12} = H_{21}$ 。同样地，变换后的矩阵 $\mathbf{H}_1, \mathbf{H}_2, \dots$ 也是对称的。）然后构造一个正交阵 \mathbf{O}_2 使得矩阵 $\mathbf{H}_2 \equiv \mathbf{O}_2^T \mathbf{H}_1 \mathbf{O}_2$ 在 $\mathbf{H}_1 = \mathbf{O}_2^T \mathbf{O}_1^T \mathbf{H} \mathbf{O}_1 \mathbf{O}_2$ 的 \mathbf{H}_1 对角元素 $(\mathbf{H}_1)_{13}$ 和 $(\mathbf{H}_1)_{31}$ 为零，以此类推。不幸的是，当某对角元素在某一步中归零后，在前一步中归零的一些对角元素很可能会变成非零，因此我们必须返回并反复循环对角元素。一般来说，要使所有非对角元素都等于零，需要无限多的步骤。在实际操作中，如果在该步骤中归零的非对角元素的绝对值小于某个极小数，则跳过该步骤；如果所有非对角元素的绝对值都小于某个极小数，则停止该步骤。本征值就是变换后的矩阵 $\dots \mathbf{O}_3^T \mathbf{O}_2^T \mathbf{O}_1^T \mathbf{H} \mathbf{O}_1 \mathbf{O}_2 \mathbf{O}_3 \dots$ 的对角元素，本征向量矩阵是乘积 $\mathbf{O}_1 \mathbf{O}_2 \mathbf{O}_3 \dots$ 。该方法 [称为循环雅可比法 (cyclic Jacobi method)]，在串行计算机上运行时，它处理大型矩阵的效率不高，但在并行计算机上运行时效率很高。

与雅可比方法相比，更有效的实对称矩阵对角化方法首先是进行一系列正交变换，将原始矩阵 \mathbf{H} 变换为对称三对角矩阵 \mathbf{T} 。三对角矩阵是指除了主对角线上的元素（元素 t_{ii} ）和主对角线上下对角线上的元素（分别为元素 $t_{i-1,i}$ 和 $t_{i+1,i}$ ）外，其他元素均为零的矩阵。 \mathbf{T} 与 \mathbf{H} 之间的关系为 $\mathbf{T} = \mathbf{O}^T \mathbf{H} \mathbf{O}$ ，其中 \mathbf{O} 是实正交矩阵，是从 \mathbf{H} 到 \mathbf{T} 的各个步骤中使用的正交矩阵的乘积。求对称三对角矩阵本征值的有效方法是 **QR** 法。在这里， \mathbf{T} 表示为正交矩阵 \mathbf{Q} 和上三角矩阵 \mathbf{R} （主对角线以下元素均为零）的乘积。通过一系列迭代步骤，矩阵收敛到一个对角矩阵，其对角元素是 \mathbf{T} 的本征值，等于 \mathbf{H} 的本征值（问题 8.55）。经过一定的改进，**QR** 方法是一种非常有效的寻找本征值和本征向量的方法（详见 *Strang*, Section 5.3 and 7.3）。

有关矩阵对角化程序和计算机程序的详细信息，见 *Press et al.*, Chapter 11 ; *Acton*, Chapters 8 and 13 ; *Shoup*, Chapter 4 .

Press et al. 是科学和工程计算程序和计算机程序的主要汇编；该书旧版本的文本可在线免费阅读，网址是 www.nr.com。有关本书旧版本的评论，见 amath.colorado.edu/computing/Fortran/numrec.html。

数学和科学计算程序可在 www.netlib.org 和 gams.nist.gov 网站上找到。可在 archives.math.utk.edu 下载免费的个人电脑数学软件以及 Mathcad 和 Maple 等商业软件的演示软件。

使用矩阵代数求解非正交基函数时的线性变分方程的步骤见问题 8.57。

Excel 电子表格可用于查找本征值和本征向量；见问题 8.53。

计算机代数系统，如 Maple、Mathematica 和 Mathcad 以及一些电子计算器都有内置命令，可以轻松找到本征值和本征向量。

本节讨论的寻找矩阵本征值和本征向量的方法适用于阶次不超过 10^3 的矩阵。在某些量子化学计算中会出现阶数不超过 10^9 的矩阵，使用特殊方法可以求出这些矩阵的最低几个本征值和相应的本征向量（见第 16.2 节）。

如 (8.80) 之后所述，对于以正交函数 f_i 为基的线性变分函数 $\sum_{i=1}^n c_i f_i$ ，由矩阵元素 $\langle f_i | \hat{H} | f_j \rangle$ 组成的矩阵 \mathbf{H} 的本征值是久期方程的根，对应于本征值 W_m 的本征向量给出了对应于 W_m 的线性变分函数的系数。问题 8.60 至 8.65 将线性变分法应用于双势阱和谐振子等问题，使用箱中粒子波函数作为基函数，并使用 Mathcad 等计算机代数程序求出 \mathbf{H} 矩阵的本征值和本征向量。

我们已经在线性变分法中讨论了矩阵对角化的问题。然而，找到任意厄米算符的本征值 a_k 和本征函数 g_k ($\hat{A}g_k = a_k g_k$) 可以表述为一个矩阵对角化问题。如果我们选择一个完全正交基函数组

$\{f_i\}$, 将本征函数展开为 $g_k = \sum_i c_i^{(k)} f_i$, 则 (问题 8.59) 元素由 $a_{ij} = \langle f_i | \hat{A} | f_j \rangle$ 组成的矩阵 \mathbf{A} 的本征值 a_k 也是算符 \hat{A} 的本征值, \mathbf{A} 的本征向量 $\mathbf{c}^{(k)}$ 的元素 $c_i^{(k)}$ 是本征函数 g_k 在基函数 f_i 上的展开系数。

本节材料进一步强调了线性算符与矩阵之间的对应关系, 以及函数与列向量之间的对应关系 (第 7.10 节)。

谷歌收录的网页的页面排名 (PageRanks) 是由谷歌通过构建某个矩阵 (称为谷歌矩阵) 并找到其对应本征值 1 的本征向量计算得出的。谷歌矩阵的阶数等于被索引网页的数量, 在 2012 年介于 $10^{10.5}$ 和 10^{11} 之间。本征向量的分量为 PageRanks。见 www.ams.org/samplings/featurecolumn/fcarc-pagerank。

总结

习题

Chapter 9

微扰理论

9.1 微扰理论

本章讨论第二种主要的量子力学近似方法——微扰理论。

假设我们有一个系统，其定态哈密顿量为 \hat{H} ，我们无法通过求解薛定谔方程

$$\hat{H}\psi_n = E_n\psi_n \quad (9.1)$$

来获得其束缚态本征函数和本征值。再假设：哈密顿量 \hat{H} 与某一满足以下我们可解的薛定谔方程系统的哈密顿量 \hat{H}^0 之间只存在微小的差异：

$$\hat{H}^0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} \quad (9.2)$$

一维非谐振子的哈密顿量就是一个例子：

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}kx^2 + cx^3 + dx^4 \quad (9.3)$$

其哈密顿量 (9.3) 与以下谐振子的哈密顿量十分接近：

$$\hat{H}^0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}kx^2 \quad (9.4)$$

若 (9.3) 中的常数 c 和 d 很小，我们可以预测非谐振子的本征函数和本征值与谐振子的非常接近。

我们称具有哈密顿量 \hat{H}^0 的系统为**未扰动系统** (unperturbed system)，具有哈密顿量 \hat{H} 的系统为**扰动系统** (perturbed system)。二哈密顿量之间的不同 \hat{H}' 称为**微扰** (perturbation)，记为：

$$\hat{H}' \equiv \hat{H} - \hat{H}^0 \quad (9.5)$$

$$\boxed{\hat{H} = \hat{H}^0 + \hat{H}'} \quad (9.6)$$

(符号 ' 不表示微分。)对于具有哈密顿量 (9.3) 的非谐振子,对相关的谐振子的微扰为 $\hat{H}' = cx^3+dx^4$ 。

在 $\hat{H}^0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$ [式 (9.2)] 中, $E_n^{(0)}$ 和 $\psi_n^{(0)}$ 分别称为状态 n 的**未扰动能量** (unperturbed energy) 和**未扰动波函数** (unperturbed wave function)。若 \hat{H}^0 为谐振子的哈密顿量 (9.4), 则 $E_n^{(0)} = (n + \frac{1}{2})\hbar\nu$ [式 (4.45)], 其中的 n 是非负整数。(我们使用 n 代替 v 为了与微扰理论符号保持一致。)上标 (0) 不表示基态。微扰理论可以应用于任何状态。下标 n 表示我们正在处理的状态。而上标 (0) 表示未扰动系统的状态。

我们的目标是将扰动系统位置的本征函数和本征值与未扰动系统已知的本征函数和本征值联系起来。为此，我们假定扰动是逐渐增加的，从未扰动系统到扰动系统的变化是连续的。在数学上，这相当于在哈密顿量中插入了一个参数 λ ，使得

$$\hat{H} = \hat{H}^0 + \lambda \hat{H}' \quad (9.7)$$

当 λ 为零时，系统为未扰动系统。随着 λ 的增加，扰动逐渐变大，在 $\lambda = 1$ 时，扰动完全“被打开”。我们插入 λ 是为了将扰动本征函数和未扰动本征函数关联起来，最终设定 $\lambda = 1$ 并删去它。

第 9.1-9.7 节处理定态哈密顿量和定态系统，第 9.8 节处理含时微扰理论。

9.2 非简并微扰理论

微扰理论处理简并和非简并能级的方式不同。本节将研究微扰对非简并能级的影响。若某未扰动系统的能级是简并的，而另一些是非简并的，那么本节的处理方法将只适用于非简并能级。

非简并微扰理论

令 $\psi_n^{(0)}$ 为能量为 $E_n^{(0)}$ 的某个非简并未扰动系统的波函数。再令 ψ_n 为扰动系统的波函数，在施加扰动后， $\psi_n^{(0)}$ 变为 ψ_n 。由 (9.1) 和 (9.7)，扰动态的薛定谔方程为

$$\hat{H}\psi_n = (\hat{H}^0 + \lambda \hat{H}')\psi_n = E_n\psi_n \quad (9.8)$$

由于 (9.8) 中的哈密顿量与参数 λ 有关，本征函数 ψ_n 和本征值 E_n 也将与 λ 有关：

$$\psi_n = \psi_n(\lambda, q), \quad E_n = E_n(\lambda)$$

其中 q 表示粒子的坐标。我们现在使用 λ 的 Taylor 级数（问题 4.1）展开 ψ_n 和 E_n ：

$$\psi_n = \psi_n|_{\lambda=0} + \left. \frac{\partial \psi_n}{\partial \lambda} \right|_{\lambda=0} \lambda + \left. \frac{\partial^2 \psi_n}{\partial \lambda^2} \right|_{\lambda=0} \frac{\lambda^2}{2!} + \dots \quad (9.9)$$

$$E_n = E_n|_{\lambda=0} + \left. \frac{dE_n}{d\lambda} \right|_{\lambda=0} \lambda + \left. \frac{d^2 E_n}{d\lambda^2} \right|_{\lambda=0} \frac{\lambda^2}{2!} + \dots \quad (9.10)$$

根据假设，当 $\lambda \rightarrow 0$ 时， ψ_n 和 E_n 分别趋近于 $\psi_n^{(0)}$ 和 $E_n^{(0)}$ ：

$$\psi_n|_{\lambda=0} = \psi_n^{(0)}, \quad E_n|_{\lambda=0} = E_n^{(0)} \quad (9.11)$$

我们引入以下缩写：

$$\psi_n^{(k)} \equiv \left. \frac{1}{k!} \frac{\partial^k \psi_n}{\partial \lambda^k} \right|_{\lambda=0}, \quad E_n^{(k)} \equiv \left. \frac{1}{k!} \frac{d^k E_n}{d\lambda^k} \right|_{\lambda=0}, \quad k = 0, 1, 2, \dots \quad (9.12)$$

则 (9.9) 和 (9.10) 变为

$$\psi_n = \psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots + \lambda^k \psi_n^{(k)} + \dots \quad (9.13)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots + \lambda^k E_n^{(k)} + \dots \quad (9.14)$$

对于 $k = 1, 2, 3, \dots$ ，我们称 $\psi_n^{(k)}$ 和 $E_n^{(k)}$ 分别为波函数和能量的 k 级校正（ k th-order correction）。我们假设级数 (9.13) 和 (9.14) 在 $\lambda = 1$ 处收敛，我们希望，对于小扰动，只取级数的前几项就能很

好地逼近真实的能量和波函数。(通常, 微扰理论的级数并不收敛, 尽管如此, 不收敛级数的前几项也能给出有用的近似值。)

令 $\psi_n^{(0)}$ 归一化: $\langle \psi_n^{(0)} | \psi_n^{(0)} \rangle = 1$ 。与令 ψ_n 归一化不同, 我们要求 ψ_n 满足

$$\langle \psi_n^{(0)} | \psi_n \rangle = 1 \quad (9.15)$$

若 ψ_n 不满足该方程, 则将常数 $1/\langle \psi_n^{(0)} | \psi_n \rangle$ 乘到 ψ_n 上, 得到一个满足该性质的微扰波函数。条件 (9.15), 称为中间归一化 (intermediate normalization), 它简化了推导过程。注意: 将 ψ_n 与一个常数相乘并不会改变薛定谔方程 $\hat{H}\psi_n = E_n\psi_n$ 中的能量, 所以应用中间归一化不会影响能量校正的结果。如果需要, 在计算结束时, 可以将中间归一化的 ψ_n 乘以一个常数, 使其在通常意义上归一化。

将 (9.13) 代入 $1 = \langle \psi_n^{(0)} | \psi_n \rangle$ [式 (9.15)], 得到

$$1 = \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle + \lambda \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + \lambda^2 \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle + \dots$$

由于该方程对介于 0 和 1 之间的任意 λ 都成立, 如 (4.11) 之后所证明的, 方程两边 λ 同次幂的系数必须相等。令 λ^0 的系数相等, 我们有 $1 = \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle$, 由于 $\psi_n^{(0)}$ 归一化, 上式显然成立。令 λ^1 、 λ^2 等的系数相等, 我们有

$$\langle \psi_n^{(0)} | \psi_n^{(1)} \rangle = 0, \quad \langle \psi_n^{(0)} | \psi_n^{(2)} \rangle = 0, \quad \text{etc.} \quad (9.16)$$

使用中间归一化时, 波函数的校正与 $\psi_n^{(0)}$ 正交。

将 (9.13) 和 (9.14) 代入薛定谔方程 (9.8), 得到

$$\begin{aligned} (\hat{H}^0 + \lambda \hat{H}') (\psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots) \\ = (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) (\psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots) \end{aligned}$$

将 λ 的同次幂进行归并, 我们得到

$$\begin{aligned} \hat{H}^0 \psi_n^{(0)} + \lambda (\hat{H}' \psi_n^{(0)} + \hat{H}^0 \psi_n^{(1)}) + \lambda^2 (\hat{H}^0 \psi_n^{(2)} + \hat{H}' \psi_n^{(1)}) + \dots \\ = E_n^{(0)} \psi_n^{(0)} + \lambda (E_n^{(1)} \psi_n^{(0)} + E_n^{(0)} \psi_n^{(1)}) + \lambda^2 (E_n^{(2)} \psi_n^{(0)} + E_n^{(1)} \psi_n^{(1)} + E_n^{(0)} \psi_n^{(2)}) + \dots \end{aligned} \quad (9.17)$$

现在 (假设收敛性合适), 要使 (9.17) 两边的两个级数对所有 λ 值都相等, 两个级数中 λ 的同次幂的系数必须相等。

令 λ^0 项的系数相等, 我们得到 $\hat{H}^0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)}$, 这是未扰动系统的薛定谔方程 (9.2), 没有给我们有用的信息。

令 λ^1 项的系数相等, 我们得到

$$\begin{aligned} \hat{H}' \psi_n^{(0)} + \hat{H}^0 \psi_n^{(1)} &= E_n^{(1)} \psi_n^{(0)} + E_n^{(0)} \psi_n^{(1)} \\ \hat{H}^0 \psi_n^{(1)} - E_n^{(0)} \psi_n^{(1)} &= E_n^{(1)} \psi_n^{(0)} - \hat{H}' \psi_n^{(0)} \end{aligned} \quad (9.18)$$

一级能量校正

为了求出 $E_n^{(1)}$, 我们将 (9.18) 两边乘以 $\psi_n^{(0)*}$ 并对全空间积分, 有

$$\langle \psi_m^{(0)} | \hat{H} | \psi_n^{(1)} \rangle - E_n^{(0)} \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = E_n^{(1)} \langle \psi_m^{(0)} | \psi_n^{(0)} \rangle - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle \quad (9.19)$$

其中用到了狄拉克符号 [式 (7.1) 和 (7.3)]。\$\hat{H}^0\$ 是厄米算符，使用厄米算符的性质 (7.12)，(9.19) 左侧的第一项变为

$$\begin{aligned}\langle \psi_m^{(0)} | \hat{H}^0 | \psi_n^{(1)} \rangle &= \langle \psi_m^{(1)} | \hat{H}^0 | \psi_n^{(0)} \rangle^* = \langle \psi_m^{(1)} | \hat{H}^0 \psi_n^{(0)} \rangle^* \\ &= \langle \psi_m^{(1)} | E_n^{(0)} \psi_n^{(0)} \rangle^* = E_n^{(0)*} \langle \psi_n^{(1)} | \psi_m^{(0)} \rangle^* = E_m^{(0)} \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle\end{aligned}\quad (9.20)$$

其中我们用到了未扰动薛定谔方程 \$\hat{H}^0 \psi_m^{(0)} = E_m^{(0)} \psi_m^{(0)}\$、\$E_m^{(0)}\$ 是实数以及 (7.4)。将 (9.20) 代入 (9.19)，对未扰动本征函数使用正交方程 \$\langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = \delta_{mn}\$，我们得到

$$(E_m^{(0)} - E_n^{(0)}) \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = E_n^{(1)} \delta_{mn} - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle \quad (9.21)$$

若 \$m = n\$，则 (9.21) 的左侧为零，(9.21) 变为

$$E_n^{(1)} = \langle \psi_n^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle = \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau \quad (9.22)$$

通过对作用于适当未扰动波函数的扰动 \$\hat{H}'\$ 求平均值，可以求出能量的一级校正。

在 (9.14) 中令 \$\lambda = 1\$，我们得到

$$E_n \approx E_n^{(0)} + E_n^{(1)} = E_n^{(0)} + \int \psi_n^{(0)*} \hat{H}' \psi_n^{(0)} d\tau \quad (9.23)$$

例题：对于具有哈密顿量 (9.3) 的非谐振子，若未扰动系统取为谐振子，求基态的 \$E^{(1)}\$。

微扰由 (9.3) 至 (9.5) 给出：

$$\hat{H}' = \hat{H} - \hat{H}^0 = cx^3 + dx^4$$

具有量子数为 \$v\$ 状态的一级能量校正由 (9.22) 给出：\$E_v^{(1)} = \langle \psi_v^{(0)} | cx^3 + dx^4 | \psi_v^{(0)} \rangle\$，其中 \$\psi_v^{(0)}\$ 是状态 \$v\$ 的谐振子波函数。对于 \$v = 0\$ 的基态，使用式 \$\psi_0^{(0)} = (\alpha/\pi)^{1/4} e^{-\alpha x^2/2}\$ [式 (4.53)]，有

$$E_0^{(1)} = \langle \psi_0^{(0)} | cx^3 + dx^4 | \psi_0^{(0)} \rangle = \left(\frac{\alpha}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-\alpha x^2} (cx^3 + dx^4) dx$$

从 \$-\infty\$ 到 \$\infty\$ 对奇函数 \$cx^3 e^{-\alpha x^2}\$ 积分为零。使用附录 (A.10) 中的积分，令 \$n = 2\$；再对 \$\alpha\$ 使用 (4.31)，得到

$$E_0^{(1)} = 2d \left(\frac{\alpha}{\pi}\right)^{1/2} \int_0^{\infty} e^{-\alpha x^2} x^4 dx = \frac{3d}{4\alpha^2} = \frac{3dh^2}{64\pi^4 \nu^2 m^2}$$

未扰动基态能量为 \$E_0^{(0)} = \frac{1}{2}h\nu\$，\$E_0^{(0)} + E_0^{(1)} = \frac{1}{2}h\nu + \frac{3dh^2}{64\pi^4 \nu^2 m^2}\$。

练习：考虑一个单粒子一维系统，其势能函数满足

$$V = \begin{cases} \infty, & x < 0 \\ V = cx, & 0 \leq x < l \\ \infty, & x \geq l \end{cases}$$

其中 \$c\$ 是常数。

(a) 当 \$c > 0\$ 时，画出 \$V\$ 的图象。

(b) 将该系统看作箱中粒子系统的微扰，求量子数为 \$n\$ 状态的 \$E_n^{(1)}\$。使用式 (3.88) 说明得到的答案是符合预期的。

(部分答案：(b) \$\frac{1}{2}cl\$。)

一级波函数校正

对于 $m \neq n$, 式 (9.21) 为

$$(E_m^{(0)} - E_n^{(0)}) \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle, \quad m \neq n \quad (9.24)$$

为了求出 $\psi_n^{(1)}$, 我们将其用厄米算符 \hat{H}^0 的未扰动本征函数 $\psi_m^{(0)}$ 完整正交集展开:

$$\psi_n^{(1)} = \sum_m a_m \psi_m^{(0)}, \quad \text{其中} \quad a_m = \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle \quad (9.25)$$

其中 (7.41) 用于计算展开系数 a_m 。在 (9.24) 中使用 $a_m = \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle$, 我们得到

$$(E_m^{(0)} - E_n^{(0)}) a_m = - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle, \quad m \neq n$$

根据假设, 能级 $E_n^{(0)}$ 是非简并的, 因此若 $m \neq n$, 有 $E_m^{(0)} - E_n^{(0)} \neq 0$ 。两边同除以 $(E_m^{(0)} - E_n^{(0)})$, 我们得到

$$a_m = \frac{\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}, \quad m \neq n \quad (9.26)$$

除了 $\psi_n^{(0)}$ 的系数 a_n 外, $\psi_n^{(1)}$ 展开式 (9.25) 的系数 a_m 都可以用 (9.26) 求出。根据 (9.25) 的第二个方程, $a_n = \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle$ 。回忆我们为 ψ_n 选择的中间归一化条件使得 $\langle \psi_n^{(0)} | \psi_n^{(1)} \rangle = 0$ [式 (9.16)]。因此 $a_n = \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle = 0$, 方程 (9.25) 和 (9.26) 给出了波函数的一级校正:

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \quad (9.27)$$

符号 $\sum_{m \neq n}$ 表示对所有未扰动的状态求和, 除了状态 n 。

在 (9.13) 中令 $\lambda = 1$, 仅使用波函数的一级校正, 我们得到了微扰波函数的近似:

$$\psi_n \approx \psi_n^{(0)} + \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} \quad (9.28)$$

有关 $\psi_n^{(2)}$ 及 ψ 的归一化, 见 *Kemble*, Chapter XI。

二级能量校正

令 λ^2 项的系数相等, 我们得到

$$\hat{H}^0 \psi_n^{(2)} - E_n^{(0)} \psi_n^{(2)} = E_n^{(2)} \psi_n^{(0)} + E_n^{(1)} \psi_n^{(1)} - \hat{H}' \psi_n^{(1)} \quad (9.29)$$

乘以 $\psi_m^{(0)*}$ 并对全空间积分, 我们得到

$$\begin{aligned} \langle \psi_m^{(0)} | \hat{H}^0 | \psi_n^{(2)} \rangle - E_n^{(0)} \langle \psi_m^{(0)} | \psi_n^{(2)} \rangle \\ = E_n^{(2)} \langle \psi_m^{(0)} | \psi_n^{(0)} \rangle + E_n^{(1)} \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(1)} \rangle \end{aligned} \quad (9.30)$$

该方程中的积分 $\langle \psi_m^{(0)} | \hat{H}^0 | \psi_n^{(2)} \rangle$ 与 (9.20) 中的积分相同, 除了 $\psi_n^{(2)}$ 代替了 $\psi_n^{(1)}$ 。在 (9.20) 中用 $\psi_n^{(2)}$ 代替 $\psi_n^{(1)}$, 我们得到

$$\langle \psi_m^{(0)} | \hat{H}^0 | \psi_n^{(2)} \rangle = E_m^{(0)} \langle \psi_m^{(0)} | \psi_n^{(2)} \rangle \quad (9.31)$$

利用 (9.31) 和 (9.30) 中未扰动函数的正交性, 可以得出

$$(E_m^{(0)} - E_n^{(0)}) \langle \psi_m^{(0)} | \psi_n^{(2)} \rangle = E_n^{(2)} \delta_{mn} + E_n^{(1)} \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle - \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(1)} \rangle \quad (9.32)$$

若 $m = n$, 则 (9.32) 的左侧为零, 得到

$$\begin{aligned} E_n^{(2)} &= -E_n^{(1)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle + \langle \psi_n^{(0)} | \hat{H}' | \psi_n^{(1)} \rangle \\ E_n^{(2)} &= \langle \psi_n^{(0)} | \hat{H}' | \psi_n^{(1)} \rangle \end{aligned} \quad (9.33)$$

因为 $\langle \psi_n^{(0)} | \psi_n^{(1)} \rangle = 0$ [式 (9.16)]。从 (9.33) 可知, 要找到能量的二级校正, 我们只需要直到波函数的一级校正。事实上, 可以证明: 知道 $\psi_n^{(1)}$ 也足以确定 $E_n^{(3)}$ 。

一般地, 可以证明: 若我们知道波函数的 k 级校正, 则可以确定能量的 $2k+1$ 级校正。(见 *Bates*, Vol. I, p. 184。)

将 (9.27) 中的 $\psi_n^{(1)}$ 代入 (9.33), 我们得到

$$E_n^{(2)} = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \langle \psi_n^{(0)} | \hat{H}' | \psi_m^{(0)} \rangle \quad (9.34)$$

由于展开系数 a_m [式 (9.26)] 是常数, 可以提到积分外面。由于 \hat{H}' 是厄米算符, 我们有

$$\begin{aligned} \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle \langle \psi_n^{(0)} | \hat{H}' | \psi_m^{(0)} \rangle &= \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle \langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle^* \\ &= |\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle|^2 \end{aligned}$$

则 (9.34) 变为

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \quad (9.35)$$

这正是我们需要的, 用未扰动波函数和能量表示的 $E_n^{(2)}$ 。

将 $E_n^{(2)}$ 和 $\lambda = 1$ 代入 (9.14), 可以得到微扰态能量的近似值为

$$E_n \approx E_n^{(0)} + H'_{nn} + \sum_{m \neq n} \frac{|\langle \psi_m^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \quad (9.36)$$

其中的积分是对未扰动归一化波函数的积分。

有关高阶能量修正的公式, 请参见 *Bates*, Volume I, pages 181-185。本节提出的微扰理论形式称为瑞利-薛定谔微扰理论 (Rayleigh-Schrödinger perturbation theory); 还存在其他方法。

讨论

方程 (9.28) 表明: 作用在波函数 $\psi_n^{(0)}$ 上的微扰相当于“掺和”进了其他状态 $\psi_m^{(0)}$, $m \neq n$ 的贡献。由于因子 $1/(E_n^{(0)} - E_m^{(0)})$, 对微扰波函数最重要的贡献来源于 (除了 $\psi_n^{(0)}$ 之外) 能量最接近 n 的状态。

要计算能量的一级校正, 我们必须只计算单积分 H'_{nn} , 而要计算能量的二级校正, 我们必须计算第 n 个状态和所有其他状态 m 之间的 \hat{H}' 矩阵元素, 然后执行 (9.35) 中的无限求和。在很多情况下, 二级校正难以准确计算。处理三级和更高的就更难了。

(9.28) 和 (9.36) 中的和是不同状态的和, 而不是不同能量值的和。如果某些能级 (第 n 个能级除外) 是简并的, 我们必须在总和中为简并能级对应的每个线性无关波函数加入某个项。

我们在 (9.28) 和 (9.36) 中对状态求和, 是因为我们需要一组完整的函数来展开 (9.25), 因此我们必须在求和中包含所有线性无关的波函数。如果未扰动问题具有连续的波函数 (例如氢原子), 我

们还必须对连续波函数进行积分，这样才能得到一组完整的波函数。若 $\psi_E^{(0)}$ 是能量为 E 的未扰动连续波函数，则 (9.27) 和 (9.35) 变为

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)} + \int \frac{H'_{E,n}}{E_n^{(0)} - E^{(0)}} \psi_E^{(0)} dE^{(0)}$$

$$E_n^{(2)} = \sum_{m \neq n} \frac{|H'_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} + \int \frac{|H'_{E,n}|^2}{E_n^{(0)} - E^{(0)}} dE^{(0)} \quad (9.37)$$

其中 $H'_{E,n} = \langle \psi_E^{(0)} | \hat{H}' | \psi_n^{(0)} \rangle$ 。这些方程中的积分是在连续态能量范围内进行的（例如，氢原子的能量从零到无穷大）。未扰动问题中连续态的存在使得 $E_n^{(2)}$ 的计算变得更加困难。

变分法和微扰法的比较

微扰法适用于系统的所有束缚态。虽然第 8.1 节所述的变分原理只适用于给定对称性的最低态，但我们可以用线性变分法来处理束缚激发态。

由于需要计算二级和高级能量校正中出现的离散态的无限和以及连续态的积分，微扰计算通常很难进行。

在微扰法中，能量的计算（ $2k+1$ 级）比波函数的计算（ k 级）要精确得多。在变分法中也存在同样的情况，我们可以用一个相当不准确的波函数得到一个相当好的能量。如果计算能量以外的其他性质，结果通常不如计算出的能量可靠。

变分-微扰法

通过变分-微扰法，我们可以精确地估算出系统基态的 $E^{(2)}$ 和高级扰动理论能量校正，而无需计算 (9.36) 中的无限和。该方法基于以下不等式：

$$\langle u | \hat{H}^0 - E_g^{(0)} | u \rangle + \langle u | \hat{H}' - E_g^{(1)} | \psi_g^{(0)} \rangle + \langle \psi_g^{(0)} | \hat{H}' - E_g^{(1)} | u \rangle \geq E_g^{(2)} \quad (9.38)$$

其中 u 是任意满足定解条件的品优函数，下标 g 表示基态。(9.38) 的证明，参见 *Hameka*, Section 7-9。把 u 看作一个试探函数，我们改变参数，使 (9.38) 左侧最小化，就可以估算出 $E_g^{(2)}$ 。函数 u 原来就是 $\psi_g^{(1)}$ 的近似——基态波函数的一级校正， u 还可以用来估算 $E_g^{(3)}$ 。类似的变分法可用于找到对基态能量和波函数的高级校正。

9.3 微扰理论求解基态氦原子

9.4 变分法求解基态氦

9.5 能量简并的微扰理论

9.6 久期方程的简化

9.7 微扰理论求解第一激发态的氦原子

9.8 含时微扰理论

9.9 辐射和物质的相互作用

总结

习题

Chapter 10

电子自旋和自旋-统计定理

10.1 电子自旋

10.2 自旋和氢原子

10.3 自旋-统计定理

10.4 氦原子

10.5 Pauli 不相容原理

10.6 Slater 行列式

10.7 微扰理论求解基态锂原子

10.8 变分法求解基态锂原子

10.9 自旋磁矩

10.10 电子自旋的梯度算符

总结

习题

Chapter 11

多电子原子

11.1 Hartree-Fock 自洽场方法

11.2 轨道和元素周期表

11.3 电子相关

11.4 角动量的叠加

11.5 多电子原子的角动量

11.6 自旋-轨道耦合

11.7 原子哈密顿算符

11.8 Condon-Slater 规则

总结

习题

Chapter 12

分子对称性

12.1 对称元素和对称操作

12.2 分子点群

总结

习题

Chapter 13

双原子分子的电子结构

13.1 Born-Oppenheimer 近似

13.2 双原子分子核的运动

13.3 原子单位制

13.4 氢分子离子

13.5 H^{2+} 离子基态的近似解

13.6 H^{2+} 离子激发态的分子轨道

13.7 同核双原子分子的分子轨道构型

13.8 双原子分子的分子光谱项

13.9 氢分子

13.10 价键理论求解 H_2

13.11 分子轨道理论和价键理论的比较

13.12 同核双原子分子的分子轨道理论和价键理论波函数

13.13 H_2 分子的激发态

13.14 双原子分子的自洽场方法波函数

13.15 分子轨道理论处理异核双原子分子

13.16 价键理论处理异核双原子分子

13.17 价电子近似

Chapter 14

分子量子力学理论

14.1 电子概率密度

14.2 偶极矩

14.3 分子的 Hartree-Fock 方法

14.4 Virial 理论

14.5 Virial 理论和化学键

14.6 Hellmann-Feynman 理论

14.7 静电力定理

总结

习题

Chapter 15

分子的电子结构

- 15.1 从头计算、密度泛函、半经验和分子力学方法
- 15.2 多原子分子的分子光谱项
- 15.3 自洽场方法、分子轨道理论求解多原子分子
- 15.4 基函数
- 15.5 自洽场方法、分子轨道理论求解 H_2O 分子
- 15.6 布居分析和键级
- 15.7 分子静电势、分子表面和原子电荷
- 15.8 定域分子轨道
- 15.9 自洽场方法、分子轨道理论求解甲烷、乙烷和乙烯
- 15.10 分子结构学
- 15.11 构象搜寻
- 15.12 分子振动频率
- 15.13 热力学定律
- 15.14 从头计算方法的量子化学程序
- 15.15 进行从头计算
- 15.16 快速 Hartree-Fock 计算方法
- 15.17 溶剂效应

Chapter 16

电子关联方法

16.1 关联能

16.2 组态相互作用

16.3 多体微扰理论

16.4 耦合簇方法

16.5 密度泛函理论

16.6 能量计算的复合方法

16.7 扩散量子 Monte Carlo 法

16.8 非共价相互作用

16.9 NMR 屏蔽常数

16.10 碎裂法

16.11 相对论效应

16.12 价键理论求解多原子分子

16.13 GVB (广义价键波函数)、VBSCF (价键自洽场)、BOVB
(呼吸轨道价键) 方法

16.14 化学反应

习题

Chapter 17

半经验和分子力学方法求解分子

17.1 半经验分子轨道求解平面共轭分子

17.2 Hückel 分子轨道理论

17.3 PPP 方法

17.4 一般的半经验分子轨道和密度泛函方法

17.5 分子力学方法

17.6 经验和半经验方法处理溶剂效应

17.7 化学反应

17.8 量子化学的未来

20 世纪 50 年代, 人们普遍认为, 除了非常小的分子之外, 对所有分子性质进行有意义的从头计算法是不可能的。在这一时期撰写的量子化学书籍中, 有这样的表述: “我们永远无法指望 (对有机化合物) 进行令人满意的从头计算法”, “对于比氢分子离子更复杂的系统, 从一开始就放弃获得薛定谔方程精确解的尝试是明智的”。1959 年, Mulliken 和 Roothaan 发现了多原子分子量子力学精确计算的“瓶颈”, 即评估多中心积分的困难。现在, 这一瓶颈已被消除。

对中等尺寸分子进行 Hartree-Fock 从头计算法和几何优化已成为常规工作, 并可采用计算效率较高的方法 (如 DFT、MP2 和 SCS-MP2) 将电子相关性包括在内。各种量子力学方法和基集的可靠程度已通过大量计算得到证实。可以进行精确从头计算法或密度泛函计算的分子大小受到现有电子计算机的速度和存储容量的限制。随着更大、更快计算机的开发, 处理更大的分子将变得可行。

近年来, 量子化学取得了长足进步, 量子力学计算已成为帮助解决各种实际化学问题的重要工具。多年前, 有关分子的量子力学计算主要局限于理论化学家阅读的期刊, 而现在, 这类计算经常出现在各种化学家阅读的期刊上, 如《美国化学会志》。量子化学正被应用于溶液中离子的水合、表面催化、反应中间产物的结构和能量、生物分子的构象以及酶催化反应的研究等问题。在许多情况

下，理论计算可能无法给出明确的答案，但它们往往足以让理论与实验进行富有成效的互动。此外，伍德沃德-霍夫曼规则等定性概念也为化学反应过程和化学键的研究提供了重要启示。

1998 年诺贝尔化学奖由沃尔特·科恩 (Walter Kohn) (密度函数理论的开发者之一) 和约翰·波普尔 (John A. Pople) (高斯系列程序和广泛使用的高斯基集、PPP 方法、CNDO 和 INDO 方法的开发者之一，以及最早将 MP 和 CC 方法应用于分子计算的人之一) 分享。诺贝尔奖委员会指出：计算量子化学正在“彻底改变整个化学”。

习题

附录 A

附录

A.1 Complex Numbers

A complex number $z = x + iy$ can be expressed as:

$$z = re^{i\theta}, \quad r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}(y/x) \quad (\text{A.1})$$

A.2 Calculus Formulas

Key differentiation and integration formulas:

$$\frac{d}{dx}(e^{ax}) = ae^{ax} \quad (\text{A.2})$$

$$\int \sin(ax) \, dx = -\frac{1}{a} \cos(ax) + C \quad (\text{A.3})$$

附录 B

参考文献

参考文献

- [1] Levine, I. N. (2014). *Quantum Chemistry* (7th ed.). Pearson.

附录 C

部分习题答案

附录 D

索引