

图 9-7 氢分子跃迁概率与频率的函数关系

$1/f_0 T$, 这个数值为 5×10^{-8} 时, 跃迁概率下降为零。显然频率必须非常接近于 ω_0 才能得到可观的跃迁概率。这一效应是能够用“原子”钟获得极高精度的基础, 而原子钟就是根据微波激射器原理工作的。

§ 9-6 光 的 吸 收

上面的处理方法适用于比氨微波激射器更普遍的情况。我们处理过分子在电场的影响下的行为, 不论场是否局限于腔内。所以我们可以简单地把一束“光”——微波频率——照射到分子上, 并求发射或吸收的概率。我们的方程同样适用于这种情况, 但我们用辐射的强度而不是用电场作为参数来重新写出那些方程。如果定义强度 \mathcal{I} 为每秒钟通过单位面积的平均能量, 则由第 2 卷第 27 章, 我们可以写出

$$\mathcal{I} = \epsilon_0 c^2 |\mathbf{E} \times \mathbf{B}|_{\text{平均}} = \frac{1}{2} \epsilon_0 c^2 |\mathbf{E} \times \mathbf{B}|_{\text{最大}} = 2 \epsilon_0 c \mathcal{E}_0^2.$$

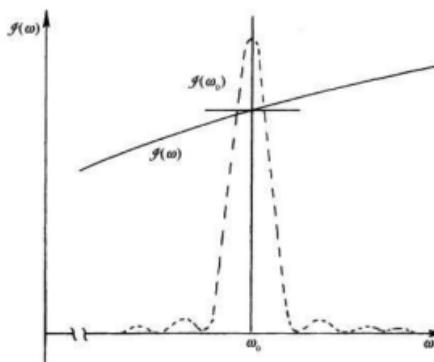
(\mathcal{E} 的最大值是 $2\mathcal{E}_0$) 于是跃迁概率为

$$P(\text{I} \rightarrow \text{II}) = 2\pi \left[\frac{\mu^2}{4\pi\epsilon_0 \hbar^2 c} \right] \mathcal{I} T^2 \frac{\sin^2[(\omega - \omega_0)T/2]}{[(\omega - \omega_0)T/2]^2}. \quad (9.53)$$

通常照射在这样一种系统上的光并不是严格单色的。因此, 值得再求解一个问题——就是计算光的频率分布在包括 ω_0 在内的一定宽度区间上的跃迁概率, 令每单位频率间隔的光强为 $\mathcal{I}(\omega)$ 。这时, 从 $|\text{I}\rangle$ 至 $|\text{II}\rangle$ 跃迁的概率成为一个积分:

$$P(\text{I} \rightarrow \text{II}) = 2\pi \left[\frac{\mu^2}{4\pi\epsilon_0 \hbar^2 c} \right] T^2 \int_0^\infty \mathcal{I}(\omega) \frac{\sin^2[(\omega - \omega_0)T/2]}{[(\omega - \omega_0)T/2]^2} d\omega. \quad (9.54)$$

一般来说, $\mathcal{I}(\omega)$ 随 ω 的变化将远比尖锐的共振项来得缓慢。两个函数也许就像图 9-8 所示那种样子。在这种情况下, 我们就可以将 $\mathcal{I}(\omega)$ 用它在尖锐的共振曲线中央位置处的值

图 9-8 光谱强度 $J(\omega)$ 可以近似地用它在 ω_0 处的值表示

$J(\omega_0)$ 来代替，并将它提出积分号外。所留下的正是图 9-7 的曲线下的积分，如我们所知此积分正好等于 $2\pi/T$ 。我们得到如下结果

$$P(I \rightarrow II) = 4\pi^2 \left[\frac{\mu^2}{4\pi\epsilon_0\hbar^2 c} \right] J(\omega_0) T. \quad (9.55)$$

这是个重要的结果，因为它是光被任何分子或原子系统吸收的普遍理论。尽管我们是从考虑态 $|I\rangle$ 比态 $|II\rangle$ 能量高的情况开始的，但我们的讨论中没有任何依赖这个事实的地方。如果态 $|I\rangle$ 具有的能量比态 $|II\rangle$ 低，式(9.55)仍然成立；这时 $P(I \rightarrow II)$ 就表示从入射电磁波吸收能量发生跃迁的概率。任何原子系统对光的吸收总是与能量差为 $E = \hbar\omega_0$ 的两个状态之间在振荡电场中发生跃迁的振幅有关。所以，对任何特定情况，此振幅总是用我们这里所用的方法计算，并可得到如式(9.55)那样的表达式。因此，我们要强调指出这个结果的下列几个特点。首先，概率正比于 T 。换句话说，单位时间发生跃迁的概率是常数。其次，这个概率正比于入射到系统上的光强。最后，跃迁概率正比于 μ^2 。你们该记得， $\mu\epsilon$ 决定了由电场 ϵ 所引起能量变化。正因为这一点， $\mu\epsilon$ 作为耦合项也出现在式(9.38)及(9.39)中，就是它引起原来是定态的两个状态 $|I\rangle$ 与 $|II\rangle$ 之间的跃迁。换句话说，对于我们所考虑的小的 ϵ ， $\mu\epsilon$ 就是哈密顿矩阵元中联系着 $|I\rangle$ 与 $|II\rangle$ 的所谓“微扰项”。在一般情况下，我们要用矩阵元 $\langle II | H | I \rangle$ 代替 $\mu\epsilon$ (见 § 5-6)。

在第 1 卷 § 42-5 中我们讲过用爱因斯坦系数 A 与 B 表示的光的吸收、受激发射，及自发发射之间的关系。现在，我们终于有了计算这些系数的量子力学程序了。我们所说的双态氮分子的 $P(I \rightarrow II)$ 正好对应于爱因斯坦辐射理论的吸收系数 B_{mn} 。对于复杂的氮分子——无论什么人要对它进行计算都太困难——我们已将矩阵元 $\langle II | H | I \rangle$ 取作为 $\mu\epsilon$ ，并讲明 μ 要由实验得到。对简单的原子系统，属于任何特定跃迁的 μ_{mn} 可由以下定义计算出：

$$\mu_{mn}\epsilon = \langle m | H | n \rangle = H_{mn}, \quad (9.56)$$

这里 H_{mn} 是包括弱电场效应的哈密顿矩阵元。这样计算出的 μ_{mn} 称为电偶极子矩阵元。所以，光的吸收与发射的量子力学理论就归结为对特定原子系统计算这些矩阵元。

我们对简单的双态系统的研究就这样使我们得以理解光的吸收与发射的普遍问题。

第 10 章 其他双态系统

§ 10-1 氢分子离子

上一章我们在把氨分子看作双态系统的近似下讨论了氨分子的一些特性。当然，氨分子实际上不是双态系统——还存在着转动、平动、振动等等许多状态，但由于氮原子的上下翻转，这些运动状态的每一个都要分解成两种内部状态。这里，我们要考虑在某种近似下可视为双态系统的另一些例子。有许多问题都是近似的，因为总是有许多其他状态存在，在更精确的分析中应当把它们考虑进去。但在所举的每个例子中，仅仅考虑两个状态我们还是能学到许多东西。

因为我们只处理双态系统，所以所需要的哈密顿就与上一章所用过的一样。当哈密顿与时间无关时，我们知道有两个具有确定——通常不相同的——能量的定态。我们的分析所用的一组基础态通常不是这些定态，而是可能具有其他某种简单物理意义的态。于是，系统的定态将用这些基础态的线性组合来表示。

为方便起见，我们总结一下第 9 章中重要的方程式。设原来选取的基础态是 $|1\rangle$ 及 $|2\rangle$ 。那么任何态 $|\psi\rangle$ 可以表示为如下的线性组合

$$|\psi\rangle = |1\rangle\langle 1| \psi\rangle + |2\rangle\langle 2| \psi\rangle = |1\rangle C_1 + |2\rangle C_2. \quad (10.1)$$

振幅 C_i （就是指 C_1 或 C_2 ）满足两个线性微分方程

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j, \quad (10.2)$$

这里 i 与 j 都取值 1 及 2。

当哈密顿的各项 H_{ij} 不依赖于 t 时，具有确定能量的两种状态（定态），即我们称为

$$|\psi_I\rangle = |I\rangle e^{-(i/\hbar)E_I t} \text{ 和 } |\psi_{II}\rangle = |II\rangle e^{-(i/\hbar)E_{II} t}$$

的状态具有能量

$$\begin{aligned} E_I &= \frac{H_{11} + H_{22}}{2} + \sqrt{\left(\frac{H_{11} - H_{22}}{2}\right)^2 + H_{12}H_{21}}, \\ E_{II} &= \frac{H_{11} + H_{22}}{2} - \sqrt{\left(\frac{H_{11} - H_{22}}{2}\right)^2 + H_{12}H_{21}}. \end{aligned} \quad (10.3)$$

这两种状态的两个振幅 C 具有相同的时间的依赖关系。与定态相联系的态矢量 $|I\rangle$ 及 $|II\rangle$ 与原来的基础态 $|1\rangle$ 及 $|2\rangle$ 之间有以下关系

$$\begin{aligned} |I\rangle &= |1\rangle a_1 + |2\rangle a_2, \\ |II\rangle &= |1\rangle a'_1 + |2\rangle a'_2. \end{aligned} \quad (10.4)$$

这些 a 都是复数, 它们满足

$$\begin{aligned} |a_1|^2 + |a_2|^2 &= 1, \\ \frac{a_1}{a_2} &= \frac{H_{12}}{E_1 - H_{11}}, \end{aligned} \quad (10.5)$$

$$\begin{aligned} |a'_1|^2 + |a'_2|^2 &= 1, \\ \frac{a'_1}{a'_2} &= \frac{H_{12}}{E_2 - H_{11}}. \end{aligned} \quad (10.6)$$

如 H_{11} 与 H_{22} 相等——比方说都等于 E_0 ——而 $H_{12} = H_{21} = -A$, 那么 $E_1 = E_0 + A$, $E_2 = E_0 - A$, 从而态 $|I\rangle$ 及 $|II\rangle$ 就特别简单:

$$|I\rangle = \frac{1}{\sqrt{2}}[|1\rangle - |2\rangle], \quad |II\rangle = \frac{1}{\sqrt{2}}[|1\rangle + |2\rangle]. \quad (10.7)$$

现在我们要利用这些结果来讨论一些取自化学和物理学领域中的有兴趣例子。第一个例子是氢分子的离子, 一个带正电的电离氢分子由两个质子及一个电子组成, 这个电子在两个质子周围运行。如果两个质子相离很远, 我们预期这样的系统会有什么状态呢? 答案很清楚: 电子将靠近一个质子而形成一个处于最低能态的氢原子, 而另一个质子则独自成为一个正离子。所以, 如果两个质子离得很远, 我们可想象这样一种物理状态, 即电子“附着”在其中一个质子上。显然, 还有另一种与之对称的状态, 即电子靠近另一个质子, 而第一个质子就成为一个离子。我们将取这两种态作为基础态, 而称它们为 $|1\rangle$ 及 $|2\rangle$ 。两种态画在图 10-1 中。当然, 一个电子靠近一个质子的状态实际上有许多, 因为这种组合可以以任何一种氢原子的激发态存在。目前我们对这些各种各样的态并不感兴趣; 我们只考虑氢原子处在最低能态——基态——的情况, 而且我们暂时不考虑电子的自旋。我们可以假设在所有状态中电子都具有沿 z 轴*“朝上”的自旋。

从氢原子中移去一个电子需要 13.6 eV 能量。只要氢分子离子的两个质子相距很远, 将电子移至靠近两质子中点某处也差不多需要这么多能量——在我们现在的问题中这是很大的能量。所以从经典物理的观点来看, 电子从一个质子跳往另一个质子是不可能的。然而, 在量子力学中这是可能的——尽管可能性不很大。仍有一定的小的振幅使电子从一个质子跑向另一个质子。作为第一级近似, 我们令每个基础态 $|1\rangle$ 与 $|2\rangle$ 都有能量 E_0 , 这正是一个氢原子加一个质子的能量。我们可以取哈密顿矩阵元 H_{11} 与 H_{22} 都近似等于 E_0 。对另外的矩阵元 H_{12} 与 H_{21} (它是电子往来跳动的振幅)我们仍记为 $-A$ 。

你们会发现这里的事情与上两章所玩的是同样的把戏。如果不考虑电子能够来回跳动

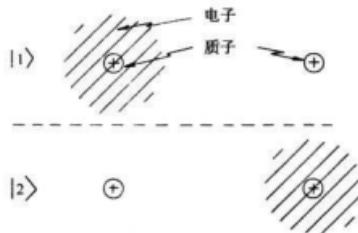


图 10-1 两个质子及一个电子的一组基础态

* 只要磁场不起作用, 这种情况就能满足, 我们将在本章后面讨论磁场对电子的效应, 第 12 章中讨论自旋在氢原子中的极小效应。

的事实，我们就有能量完全相同的两个态。然而，由于电子有来回跳动的概率，这个能级就分裂为两个能级，跃迁的概率越大，分裂就越大。这样，系统的两个能级就是 $E_0 + A$ 及 $E_0 - A$ ，带有这两个确定能量的状态就由式(10.7)给出。

从解答中我们看出，如果使质子与氢原子靠近，电子就不会总待在其中一个质子一边，而要在两个质子间来回跳动。如果它开始处在其中一个质子旁，那它就将在态 $|1\rangle$ 及 $|2\rangle$ 之间来回振荡，即给出一个随时间变化的解。为了得到最低能量的解（这解不随时间变化），开始时此系统中的电子围绕各个质子的振幅就必须相等。记住，这里不是两个电子——我们并没有说围绕每个质子各有一个电子。总共只有一个电子，它在任一个质子旁的位置上都具有相同的振幅，其大小为 $1/\sqrt{2}$ 。

本来靠近一个质子的电子跑到另一个质子旁的振幅 A 依赖于质子之间的距离。质子靠得越近，振幅就越大。你们记得我们在第7章中谈到过在经典（观念）上不可能有电子“穿透势垒”的振幅。这里我们有同样的情况，电子穿透的振幅在距离大时大致上随距离增大而指数式减少。既然质子靠近时跃迁概率（从而 A ）变大，能级的分离也就增大。如果系统在态 $|I\rangle$ ，能量 $E_0 + A$ 随着距离的减小而增大，所以这样的量子力学效应将产生一种排斥力使质子分离。相反，若系统处在态 $|II\rangle$ ，那么如让质子靠近，总能量将减少；有个吸引力将质子拉到一起。两种能量随质子之间距离而变化的情况大致如图 10-2 所示。于是我们对使 H_2^+ 离子结合在一起的结合力就有了一个量子力学解释。

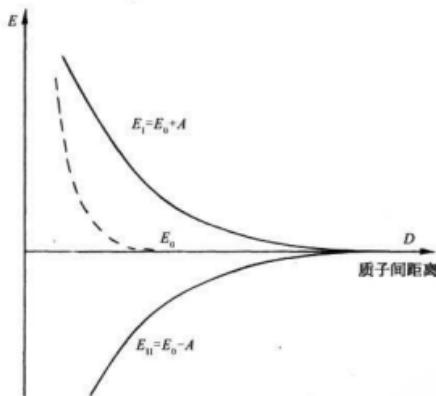


图 10-2 H_2^+ 离子两种定态能量作为两个质子间距离的函数

但是，我们忘了一件事。除了我们刚才所描写的力外，在两个质子间还存在着静电斥力。当两个质子相隔较远时——如图 10-1 中的那种情况——“裸”质子只见到一个中性原子，因此静电力可以忽略不计。但当距离十分靠近时，“裸”质子开始进入电子的分布区“内部”，也就是说，平均而言，氢原子核离另一个质子比离电子为近。这样便开始出现一些额外的静电能，当然是正的。这个能量——也随距离而变化——应包括在 E_0 之中。所以我们应当取图 10-2 中的虚曲线表示 E_0 ，当距离小于氢原子半径时，它迅速上升。我们应当由 E_0

加上及减去翻转能量 A 。这样做以后，能量 E_1 及 E_2 就将如图 10-3 所示那样随着质子间距离而变。[在这个图上我们画的是经过更详细计算的结果。质子间距离以 $1 \text{ nm} (10^{-9} \text{ m})$ 为单位标出。超过一个质子加上一个氢原子的能量 ΔE ，则以氢原子束缚能——所谓“里德伯”能量 E_H 13.6 eV ——为单位标出。]我们看到态 $|1\rangle$ 有一个能量极小点，这就是 H_2^+ 离子的平衡位形，即最低能量态。这点的能量低于一个分离的质子与一个氢离子的能量，所以系统是束缚着的。单个电子起着使两个质子结合在一起的作用。化学家称之为“单电子键”。

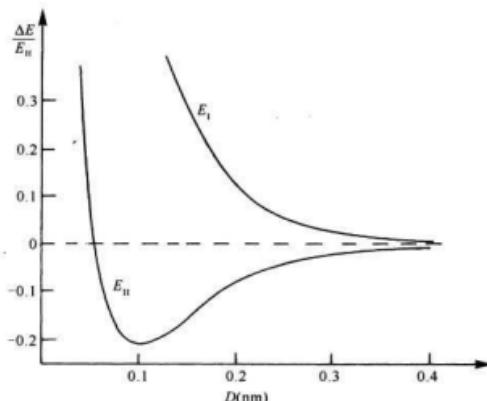


图 10-3 H_2^+ 离子能级作为质子间距离 D 的函数 ($E_H = 13.6 \text{ eV}$)

这类化学键联也常称为“量子力学共振”(用我们先前描述过的两个耦合着的单摆作类比)。但这种称呼听起来比事情本身更神秘些。其实，只有当你从对基础态的不恰当选择出发——正如我们也做过的那样，它才“共振”！假如你选了态 $|1\rangle$ ，你就会有最低的能量——情况就是如此。

我们可以用另一种方式来看一下为什么这样一个状态具有比一个质子与一个氢原子低的能量。让我们设想一个电子靠近彼此相距某个固定但不太远的距离的两个质子。你们还记得，当只有一个质子时，由于不确定性原理，电子是“弥散”的。电子在具有低的库仑势能和不要被束缚在太小的空间两者之间寻求一种平衡。因为由于不确定性关系 ($\Delta p \Delta x \approx \hbar$)，电子被束缚在小的空间内则它的动能就要增高。如果有两个质子，就有更多的空间使电子具有较低的势能。它可以在不加大势能的条件下弥散开来——从而降低动能。净结果就是得到低于一个质子和一个氢原子的能量。那么为什么另一种态 $|1\rangle$ 具有较高能量？请注意，这个态是态 $|1\rangle$ 及 $|2\rangle$ 的差。由于 $|1\rangle$ 及 $|2\rangle$ 的对称性，它们的差使得在两个质子的半当中找到电子的振幅必定为零。这意味着电子在某种程度上被束缚得更紧，这就导致较大的能量。

应该指出，一旦两个质子之间的距离近到图 10-3 中曲线的极小值时，把 H_2^+ 离子视为双态系统的近似处理就完全失效了，因而这时不能得到实际结合能的精确值。对于小间距，我们在图 10-1 中所设想的两种“态”的能量不再恰好等于 E_0 ，需要一种更精确的量子力学处理方法。

假定我们现在问,如果不是两个质子,而是两个不同的客体,比方说一个质子与一个正的锂离子(两个粒子仍都带有单位正电荷),情况又会怎样?在这种情况下,两项哈密顿量 H_{11} 与 H_{22} 不再相等;事实上,它们将很不一样。如果出现了差 $(H_{11} - H_{22})$ 的绝对值远大于 $A = -H_{12}$ 的情况,吸引力就变得非常微弱,这可由下面看出。

如果将 $H_{12} H_{21} = A^2$ 代入式(10.3),就有

$$E = \frac{H_{11} + H_{22}}{2} \pm \frac{H_{11} - H_{22}}{2} \sqrt{1 + \frac{4A^2}{(H_{11} - H_{22})^2}}.$$

当 $(H_{11} - H_{22})^2$ 远大于 A^2 时,平方根就近似地等于

$$1 + \frac{2A^2}{(H_{11} - H_{22})^2},$$

于是两个能量就是

$$\begin{aligned} E_1 &= H_{11} + \frac{A^2}{H_{11} - H_{22}}, \\ E_2 &= H_{22} - \frac{A^2}{H_{11} - H_{22}}. \end{aligned} \tag{10.8}$$

它们非常接近于孤立原子的能量 H_{11} 与 H_{22} ,只是由于翻转振幅 A 而略有差别。

能量差 $E_1 - E_2$ 是

$$(H_{11} - H_{22}) + \frac{2A^2}{H_{11} - H_{22}}.$$

来自电子振荡的附加能级分裂不再等于 $2A$,而小了一个因子 $A / (H_{11} - H_{22})$,该因子我们现在假设为远小于 1。此外, $E_1 - E_2$ 对两个原子核间距的依赖关系也比 H_2^+ 离子小得多——也减小因子 $A / (H_{11} - H_{22})$ 。现在我们明白了为什么不对称双原子分子的键联一般很微弱。

在我们的 H_2^+ 离子理论中我们已经发现了对两个质子共有的一个电子实际上提供了两质子间的吸引力这种机制的解释,这种吸引力即使在质子相距较远时也能存在。此吸引力来自电子从一个质子跳到另一个质子的可能性所造成的系统能量的减少。在这种跳跃过程中,系统由一种位形(氢原子,质子)变到另一种位形(质子,氢原子),或者反过来。我们可以用符号来表示这种过程



由于这个过程产生的能量偏移正比于振幅 A , A 即能量为 $-W_H$ (电子在氢原子中的结合能)的电子可从一个质子跑到另一个质子的振幅。

当两个质子的间距离 R 较大时,电子在其跳跃过程中必须通过的大部分空间中静电势能都接近于零。因而在这些地方,电子就像自由粒子在真空中那样运动——但带有负的能量!我们在第3章式(3.7)中已知道,具有确定能量的粒子从一个地点跑到距离 r 的另一地点的振幅正比于

$$\frac{e^{(1/m)r}}{r},$$

这里 p 是与确定能量对应的动量。在现在的情况下(用非相对论公式), p 由下式给定

$$\frac{p^2}{2m} = -W_H, \quad (10.9)$$

这意味着 p 是虚数,

$$p = i\sqrt{2mW_H}.$$

(带负号根式在这里没有意义。)

于是我们预期, 当两个质子间距离 R 大时, H_2^+ 离子的振幅 A 将按下式变化

$$A \propto \frac{e^{-i(\sqrt{2mW_H/\hbar})R}}{R}. \quad (10.10)$$

由电子键联而产生的能量偏移正比于 A , 于是就有一个力将两个质子拉近, 此力在 R 大的情况下正比于式(10.10)对 R 的导数。

最后, 为完备起见, 我们应当指出对双质子和单电子的系统还有另一种给出能量对 R 相关的效应。迄今为止我们都忽略了这点, 因为通常它是不很重要的——例外只在距离很大时出现, 这时交换项 A 的能量已指数地衰减为很小的值。我们所考虑的新效应是质子对氢原子的静电吸引, 这种吸引的产生方式与任何带电体吸引一个中性物体的情况一样。裸质子在中性氢原子处产生一个电场 ϵ (随 $1/R^2$ 变化)。原子便极化而形成一个正比于 ϵ 的感生偶极矩 μ 。偶极子的能量是 $\mu\epsilon$, 它正比于 ϵ^2 或 $1/R^4$ 。因而在系统的能量中有随距离的 4 次方而衰减的一项。(它是对 E_0 的校正。)此能量随距离的减少比式(10.10)给出的 A 的改变来得慢; 在大距离 R 处它就成为仅剩的给出能量随 R 而变的重要项, 因此也就是仅剩的力。注意对两个基础态来说静电项都有相同的符号(力是引力, 故能量为负), 对两个定态也是如此, 而电子交换项 A 对两种定态则给出了相反的符号。

§ 10-2 核 力

我们已经看到, 一个氢原子与一个质子的系统具有由交换单个电子引起的相互作用能, 它在距离 R 大时的变化是

$$\frac{e^{-\alpha R}}{R}, \quad (10.11)$$

这里 $\alpha = \sqrt{2mW_H}/\hbar$ 。(人们常说当电子——就像在这里——跃过它在其中具有负能量的空间时就是一次“虚”电子的交换。更明确地说, “虚交换”的意思是这种现象包含了在交换态与非交换态之间的量子力学干涉。)

现在我们可以问以下问题: 在其他种类粒子之间的作用力是否可能也有类似的起源? 例如, 一个中子与一个质子, 或两个质子之间的核力是否也相类似? 在试图解释核力的本性时, 汤川提出两个核子间的力就起因于类似的交换效应, 只是在这种情况下, 不是起因于电子的虚交换, 而是起因于他称为“介子”的新粒子的虚交换。今天, 我们认为质子或其他粒子在高能碰撞中所产生的 π 介子与汤川介子是同一种粒子。

作为一个例子, 让我们来看看, 在质子与中子之间交换一个质量为 m_π 的正 π 介子(π^+)时, 预期会出现哪一种力。就像氢原子 H^0 可以放弃一个电子 e^- 而成为质子 p^+ :



质子可以释放一个 π^+ 介子而变成一个中子 n^0 :

$$p^+ \longrightarrow n^0 + \pi^+. \quad (10.13)$$

所以,如果a点有一个质子,b点有一个中子,相距为R。在质子可以通过发射一个 π^+ 成为中子,而b处中子吸收这个 π^+ 后就成为一个质子。在双核子(加上 π 介子)系统中存在一种相互作用能,它依赖于 π 介子交换的振幅A——正像我们在 H_2^+ 离子中看到的电子交换情况一样。

在式(10.12)的过程中, H^0 原子的能量比质子能量小 W_H (作非相对论计算,并忽略电子的静能 mc^2),于是电子具有负动能——或虚动量——如式(10.9)所示。在核过程(10.13)式中,质子与中子具有几乎相等的质量,所以 π^+ 的总能量为零。对质量为 m_π 的 π 介子,其总能量E与动量p之间的关系是

$$E^2 = p^2 c^2 + m_\pi^2 c^4,$$

因为E为零(或至少与 m_π 相比可忽略),动量又是虚的了:

$$p = i m_\pi c.$$

利用我们对束缚电子穿透两质子间的空间中的势垒之振幅所作的同样论证,得到在核情况下的交换振幅A,在R大时,它应为

$$\frac{e^{-(m_\pi c/R)}}{R}. \quad (10.14)$$

相互作用能正比于A,所以也以同样方式变化。我们得到了以所谓汤川势形式表示的两核子间的能量变化关系。顺便提一下,我们早先曾直接从自由空间中 π 介子的运动微分方程得到过这同一公式[见第2卷第28章式(28.18)]。

按同样的思路,我们可以讨论由交换一个中性 π 介子(π^0)而产生的两个质子(或两个中子)间的相互作用。现在基本过程是

$$p^+ \longrightarrow p^+ + \pi^0. \quad (10.15)$$

一个质子可以发射一个虚 π^0 介子,而仍旧是一个质子。如果我们有两个质子,1号质子可以发射一个虚 π^0 介子,这个 π^0 介子被2号质子吸收。结果,我们仍然有两个质子。这多少与 H_2^+ 离子有些不同。在那里 H^0 在发射一个电子后变为另一种状态——质子。而现在我们假设质子可以发射一个 π^0 而不改变它的性质。事实上在高能碰撞中可观察到这种过程。它类似于一个电子发射一个光子后仍然是电子的情况:

$$e \longrightarrow e + \text{光子}. \quad (10.16)$$

在光子被发射前或被吸收后我们并没有在电子内部“见到”它们,它们的发射也不改变电子的“性质”。

现在回到两个质子的问题上,存在一种相互作用能,它是由下述过程的振幅A引起的:一个质子发射一个中性 π 介子,该介子(带有虚动量)抵达另一个质子并被其吸收。A也正比于式(10.14),式中 m_π 是中性 π 介子的质量。完全同样的论证给出两个中子之间同样的相互作用能。由于中子与质子、质子与质子、中子与中子之间的核力(与电效应无关)相同,我们可以推断带电的及中性 π 介子的质量应当相同。实验表明,它们的质量确实很接近于相等,微小的差别可以预料是来自电自能的校正(见第2卷第28章)。

还有其他种类的粒子——如 K 介子——也可以在两个核子间交换。两个 π 介子同时交换也是可能的。但所有这其他的交换“物体”都具有比 π 介子质量 m_π 大的静止质量 m_x ，在交换振幅中都出现依下式变化的项

$$\frac{e^{-(m_x c^2 / \hbar) R}}{R}.$$

随着 R 的增加，这些项衰减得比单介子项快。今天，还没有人知道怎么去计算这些高质量项，但对于足够大的 R 值而言，只有单 π 介子项才保留下来。而实际上，那些核相互作用的实验确实表明相互作用能只在距离大时才如单 π 介子交换理论所预言的那样。

在电磁学的经典理论中，库仑静电相互作用与加速电荷产生的光辐射密切相关，都可用麦克斯韦方程组说明。我们在量子论中已看到光可以用箱子中经典电磁场的简谐振动的量子激发来描述。另一方面，可以用服从玻色统计的粒子——光子来描写光，由此建立起量子理论，我们在 § 4-5 中曾强调指出这两种可供选择的观点总是给出同样的预言。那么，第二种观点是否可以始终贯彻包括所有的电磁效应？特别是，如果我们想纯粹用玻色粒子——即用光子——来描写电磁场的话，那么库仑力由什么而产生？

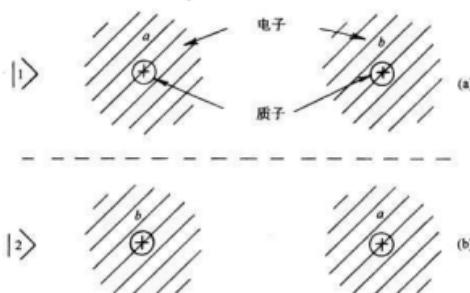
从“粒子”的观点看，两个电子之间的库仑相互作用来自于虚光子的交换。一个电子发射一个光子——就像式(10.16)中的反应——这个光子跑到第二个电子那里，在同样反应的逆过程中被吸收。于是相互作用又能用式(10.14)那样的公式给出，但现在 m_γ 由光子的静止质量——等于零——代替。所以两个电子间光子的虚交换给出与电子间距 R 成简单反比关系的相互作用能——这正是通常的库仑势能！在电磁学的“粒子”理论中，虚光子交换过程引起所有的静电现象。

§ 10-3 氢 分 子

作为下一个双态系统的例子，我们来考察一下中性氢分子 H_2 。当然，由于它有两个电子，理解起来更复杂些。我们还是从考虑两个质子分得相当开时所发生的情况开始。只是现在要加上两个电子。为了追踪它们，称其中一个为“电子 a ”，另一个为“电子 b ”。我们仍旧设想两种可能的状态。一种可能性是“电子 a ”围绕第一个质子，“电子 b ”围绕第二个质子，如图 10-4(a)所示。这时就只有两个氢原子。我们称此为状态 $|1\rangle$ 。另外还有一种可能性：“电子 b ”围绕第一个质子，而“电子 a ”则围绕第二个质子。称此为状态 $|2\rangle$ 。由情况的对称性，这两种可能性能量应该相等，只是（如我们将见到的）系统的能量不正好就是两个氢原子的能量。我们应当指出还存在其他许多可能性。例如，“电子 a ”可能靠近第一个质子，而“电子 b ”则可能在围绕同一个质子另一种状态中。但我们不去考虑这种情况，因为，它无疑具有较高的能量（由于两个电子间的库仑斥力）。如要求更高的精确性，就得包括这些态，但是只要考虑图 10-4 中的两种状态，我们就能懂得分子键联的本质。在这种近似下，我们可以通过给出处在态 $|1\rangle$ 的振幅 $(1|\psi)$ 及处在态 $|2\rangle$ 的振幅 $(2|\psi)$ 来描写任何状态 ψ 。换句话说，态矢量 $|\psi\rangle$ 可以写成线性组合

$$|\psi\rangle = \sum_i |i\rangle \langle i| \psi\rangle.$$

接下去，像往常那样，我们假设有电子通过两个质子之间的空间交换位置的一定振幅 A 。

图 10-4 H₂ 分子的一组基础态

这种交换的可能性意味着系统的能量要分裂，就像我们在别的双态系统中已经见到的那样。像氢分子离子的情况一样，当质子间距较大时，分裂很小。当质子彼此靠近时，电子来回跑的

振幅增大，因此分裂加大。低能态的降低意味着存在将原子拉在一起的吸引力。当两个质子靠得非常近时，由于库仑斥力，能级又将升高。最终净结果是两种定态具有如图 10-5 所示那样的随着分开的距离而变的能量。在距离约为 0.74 Å (1 Å = 0.1 nm) 时，低能级达极小值，这就是真实氢分子的质子间的距离。

此刻你也许正想到一个反对意见，怎么处理两个电子是全同粒子这件事呢？我们一直称它们为“电子 a”与“电子 b”，但实际上根本无法说出谁是谁。在第 4 章中我们曾说过，对于电子——它们是费米子——来说，如果通过交换电子引起的事情有两种可能的方式，两者的振幅将以负号相互干涉。这意味着如果我们把这个电子与那个电子互换，振幅的符号必须反过来。但是，我们刚才已得出结论，氢分子的束缚态应该是（当 $t = 0$ 时）

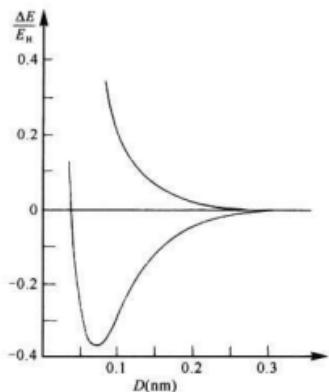


图 10-5 对不同质子间距离 D 的 H₂ 分子的能级 ($E_h = 13.6$ eV)

$$| \text{II} \rangle = \frac{1}{\sqrt{2}}(| 1 \rangle + | 2 \rangle).$$

而根据第 4 章的法则，这种状态是不允许的。如果我们交换两个电子，就得到状态

$$\frac{1}{\sqrt{2}}(| 2 \rangle + | 1 \rangle),$$

这一来，符号就相同而不是相反。

如果两个电子具有相同的自旋，上述论证是正确的。确实，如两个电子自旋都朝上（或二者自旋都朝下），唯一允许的态就是

$$| \text{I} \rangle = \frac{1}{\sqrt{2}}(| 1 \rangle - | 2 \rangle).$$

对这种态来说,两个电子的交换得到

$$\frac{1}{\sqrt{2}}(|2\rangle - |1\rangle),$$

这正是 $-|I\rangle$,符合要求。所以如果我们让两个氢原子彼此靠近而它们电子自旋的方向相同,它们就会进入态 $|I\rangle$ 而不是 $|II\rangle$ 。但请注意态 $|I\rangle$ 是个较高能态。它的能量-间距曲线没有极小值。这两个氢原子总是互相排斥而不能形成一个分子。所以我们得出结论,氢分子不可能以两个电子的自旋平行的方式存在。这是符合事实的。

另一方面,态 $|II\rangle$ 对两个电子来说是完全对称的。事实上,如果互换我们称为 a 的电子与我们称为 b 的电子,就恰好回到同一状态。在§4-7中我们曾看到,如果两个费米子处在同一状态,它们必定有相反的自旋。所以束缚而成的氢分子必定有一个自旋朝上的电子和一个自旋朝下的电子。

如果我们将质子的自旋包括在内的话,氢分子的整个描述就真的更复杂了。那样把分子视为双态系统就不再正确了。事实上应当将它看作为八态系统——对于态 $|1\rangle$ 和 $|2\rangle$ 的每一个状态各有4种可能的自旋配置。由此可见我们略去自旋使事情简单一些。但最终的结论还是正确的。

我们发现 H_2 分子的最低能态——唯一的束缚态——有自旋相反的两个电子。电子的总自旋角动量是零。另一方面,带有平行自旋——因而具有总角动量 \hbar ——的两个靠近的氢原子必定处在较高(非束缚)能量的状态;两个原子相互排斥。在自旋与能量间有一种有趣的相关。这里为以前讨论过的情况提供了另一个实例,由于自旋平行的情况具有比自旋相反的情况更高的能量,由此看来在两个自旋之间存在一种“相互作用”能。在某种意义上你可以说自旋试图达到一种反平行状况,而在这样做的过程中它具有释放出能量的潜力——这不是因为存在着较大的磁力,而是因为不相容原理。

我们在§10-1中看到,单个电子所造成的两个不同离子的键联很可能是十分微弱的。若用两个电子键联的话,情况就不是如此了。假设图10-4中的一对质子被任意两个离子(闭合的内电子壳层和一个离子电荷)所代替,而一个电子在两个离子上的结合能是各不相同的。态 $|1\rangle$ 和 $|2\rangle$ 的能量仍然相等,因为在每个这样的态中,都是一个电子与一个离子相结合。因此,我们总是有与 A 成比例的能量分裂。两个电子结合是普遍存在的,它是最常见的价键。化学键通常就是这种两个电子玩的翻转游戏。虽然也可以只用一个电子来键合两个原子,但那是相对少见的,因为它需要恰恰正好的条件。

最后我们想谈一下,如果电子与一个原子核的吸引能量远比与另一个核的能量高,那么早先所说的忽略其他可能的状态的讲法就不再正确了。假设核 a (也可以是一个正离子)对电子的吸引力远大于核 b 对电子的吸引力。那就可能发生即使两个电子都在核 a 旁而没有电子在核 b 旁时,总能量仍然相当低的情况。强吸引作用可能比补偿两个电子的相互排斥作用所需的还多。如果是这样,在最低的能态中,就可能有较大的振幅在核 a 旁找到两个电子(形成一个负离子),而只有很小的振幅在核 b 旁找到任何电子。这种情况就好像一个负离子和一个正离子。事实上,这正是 $NaCl$ 这样的“离子”分子中所发生的情况。你可以看出,在共价键与离子键之间各种渐变的键合形式都是可能的。

现在你已开始看到,许多化学事实怎样借助于量子力学的描述而获得最清楚的理解。

§ 10-4 苯 分 子

化学家们发明了一些精美的图式来表示复杂的有机分子。现在我们要来讨论其中最有趣的图式之一——如图 10-6 所示的苯分子。

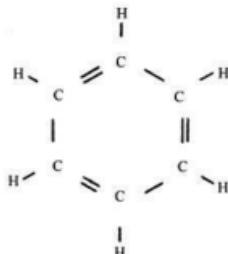


图 10-6 苯分子 C_6H_6

它由对称排列的 6 个碳原子与 6 个氢原子组成。图中的每根短线表示一对自旋相反的电子，它们起着共价键作用。每个氢原子提供 1 个电子而每个碳原子提供 4 个电子，总共包含有 30 个电子。(每个碳原子核近旁还有组成第一壳层或 K 壳层的两个电子。由于这些电子紧紧地受到束缚以至于在共价键形成中并没有起明显作用，故在图中没有画出。)所以图上的每条短线表示一个键，或一对电子，而双键则意味着每隔一对碳原子间有两对电子。

这种苯分子有个奥妙的地方。由于化学家们已经测量了各种含有几个苯环的化合物的能量(例如，他们通过研究乙烯而得知双键的能量，等等)，所以我们可以计算为形成这样的化合物所需要的能量。这样我们就能计算我们预期的苯分子的总能量。然而，苯环的实际能量要远小于由这种计算所得的值；比起根据所谓“未饱和双键系统”而预期的情况而言，苯环的结合要紧密得多。通常不处在这样一个环上的双键系统因其有相对较高的能量而在化学上容易被破坏，这种双键很容易被外加的其他氢原子断开。但苯中的环则十分稳固而难以打断。换句话说，苯所具有的能量要比由键的图像所算出的数值低得多。

另外还有一个奥妙。假设我们用两个溴原子代替两个相邻的氢原子以形成邻二溴苯分子。有两种方式组成这种分子，如图 10-7 所示。溴原子可以像图(a)中的那样联在双键的两端，也可像图(b)中那样联在单键的两端。人们会以为邻二溴苯分子应有两种不同的形式，但事实并非如此。只有一种这样的化合物*。

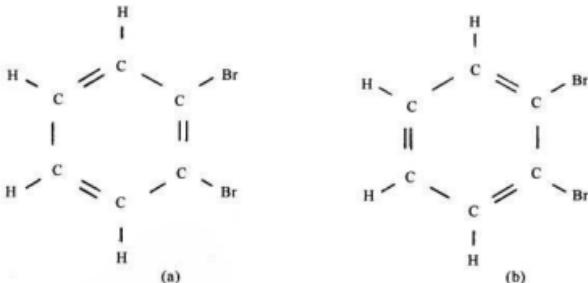


图 10-7 邻二溴苯的两种可能组态。两个溴原子可相隔一个单键或一个双键

* 我们将事情稍微简单化了。原来，化学家认为应当有 4 种形式的二溴苯：两种形式是溴原子连在相邻的碳原子上(邻二溴苯)，第三种形式是溴原子相间地连在碳原子上(间二溴苯)，而第四种形式是两个溴原子彼此相对(对二溴苯)。然而，他们只找到 3 种形式——只有 1 种形式的邻位溴苯分子的形式。

现在我们来解答这些奥秘——或许你们已经猜到如何来做了,当然,要注意到苯环的“基态”实际上是个双态系统。我们可以设想苯中的键可以取图 10-8 所示两种配置的任一形式。你会说,“它们实际上是相同的,它们应当有相同的能力。”确实应当如此,正因为如此,必须将它们作为双态系统来分析。每个态表示全体电子的不同的位形,整个结构有某个振幅 A 从这种配置转换到另一种配置——电子有机会由一种位形翻转到另一种位形。

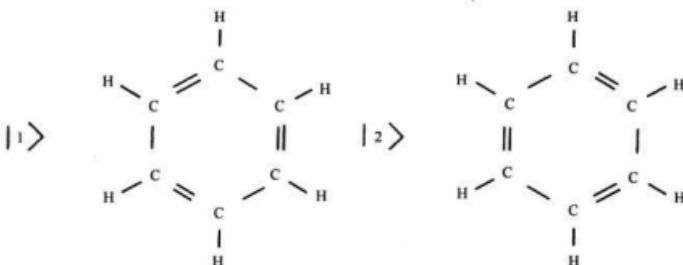


图 10-8 苯分子的一组基础态

我们已知道,这种翻转的可能性就造成了一个混合态,它的能量比分别按图 10-8 中的两幅图案所作的计算结果来得低。这一来,就有两种定态——一种的能量值大于预期值,另一种则小于预期值。因此,实际上图 10-8 所示的两种可能状态都不是苯的真正的正常状态(最低能量),而是它处于图示的每种态的振幅都是 $1/\sqrt{2}$ 。这是常温下苯化学中所涉及到的唯一状态。附带说一下,较高能态也是存在的,我们知道它的存在是因为苯对于频率为 $\omega = (E_1 - E_2)/\hbar$ 的紫外线有强烈的吸收。你们记得,在氨分子中,来回翻转的是 3 个质子,而其能量的差在微波区域。对苯分子,翻转的是电子,由于它们轻得多,就更容易来回翻转,这就使系数 A 大得多。结果能量差也大得多,约为 1.5 eV,这相当于紫外光子的能量*。

如果用溴原子代替会发生什么情况?图 10-7 中的(a)与(b)两种“可能性”仍表示两种不同的电子位形。唯一的差别在于开始时所取的两种基础态所具有的能量略有不同。能量最低的定态仍是两种态的线性组合,但具有不相等的振幅。比方说,处在态 |1> 的振幅可能会取 $\sqrt{2/3}$ 这样的值,而处在态 |2> 的振幅的大小就可能是 $\sqrt{1/3}$ 。如没有更多的信息,我们就不能说得十分确切,但只要两种能量 H_{11} 与 H_{22} 不再相等,那么振幅 C_1 与 C_2 就不会再有相同的数值。当然,这意味着,图中的两种可能状态中有一种可能性大些,但由于电子的活动性足够大,因而两种情况都具有一定振幅。不过另一种状态具有不同的振幅(如 $\sqrt{1/3}$ 及 $-\sqrt{2/3}$),但处于较高能量上。只有一个最低能态,不像旧的固定化学键的理论所假设的有两个。

* 我们的说法会造成一点误解。把苯作为双态系统时,紫外线的吸收很微弱,因为两种态之间的偶极矩的矩阵元是零[两种状态是电学对称的,所以在我们的跃迁概率公式(9.55)中,偶极矩 μ 为零,因而不吸收光。]如这些态是仅有的状态,较高能态的存在就应当用其他方式来证实。然而,取较多基础态(诸如具有相邻双键的态)的苯分子的更完全理论表明,真正的苯分子定态比我们所谈到的略有变形。最后所得的偶极矩将允许发生课文中所说的由于吸收紫外线而造成的跃迁。

§ 10-5 染 料

我们再举一个化学中双态现象的例子——这次是在较大的分子尺度上。它与染料的理

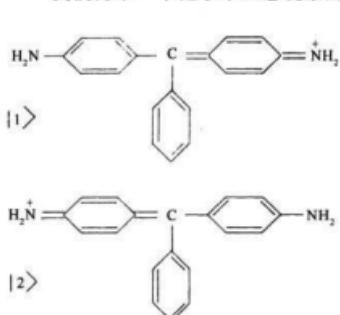


图 10-9 品红染料分子的两个基础态

论有关。许多染料——事实上是大多数人造染料——的特性十分有趣；它们具有某种对称性。图 10-9 表示一种称为品红的特殊染料的离子，这染料呈现紫红色。它的分子是三个环状结构，其中两个是苯环。第三个环跟苯环不完全相同，因为环中只有两个双键。图上画出了两幅同样圆满的图案，我们会猜想它们具有相同的能量。但存在所有的电子都从一种状态翻转至另一种状态一定的振幅，而将“空”位挪到另一端。由于包括了这么多电子，翻转振幅要比苯的情况略低些，故两种定态的能量差就要小些。然而，仍存在着通常的两种定态

$|I\rangle$ 及 $|II\rangle$ ，它们是图中两种基础态的和与差。得

到的 $|I\rangle$ 与 $|II\rangle$ 的能量差值等于可见光区光子的能量。如果将光照射在分子上，在某一频率上就会出现很强的吸收，从而显示出明亮的色彩。这就是染料的成因！

这种染料分子的另一个有趣特性是，在图示的两个基础态中，电荷中心处于不同的位置。结果，分子将受外电场的强烈影响。在氨分子中我们就见过类似的效果。显然，只要知道 E_0 与 A 的数值，就可以用完全相同的数学方法对它进行分析。一般来说， E_0 与 A 的数值是通过汇集实验数据得到的。如果对许多染料进行测量，就往往可能猜出某个相关的染料分子的情况。由于电荷中心的大位移，式(9.55)中的 μ 值就较大，从而材料吸收特征频率为 $2A/\hbar$ 的光的概率就大。因此，染料不仅有颜色，而且颜色非常强——少量的材料就能吸收大量的光。

翻转率——从而 A ——对整个分子的结构非常敏感。改变 A ，能量分裂以及与此相关的染料颜色就要变化。还有，分子也不必是完全对称的。即使存在着微小的不对称，我们仍看到同样的基本现象，只是稍有不同。所以，可以通过在分子中造成一点非对称性来改变颜色。例如，另一种重要染料孔雀绿就与品红十分相似，只是其中两个氢原子被 CH_3 所代替。由于 A 改变了，翻转率也改变了，所以它就有另一种颜色。

§ 10-6 磁场中自旋 $1/2$ 粒子的哈密顿

现在我们讨论包含自旋 $1/2$ 粒子的双态系统。我们所要讲的有些内容已在以前的几章中讨论过，但再讲一次有助于将某些不明白的地方弄得更清楚些。我们可以把一个静止电子视为双态系统。尽管本节谈论“一个电子”，但所得到的东西对任何自旋 $1/2$ 粒子都是正确的。假设我们选 $|1\rangle$ 与 $|2\rangle$ 作为基础态，它们分别为电子自旋的 z 分量是 $+\hbar/2$ 及 $-\hbar/2$ 。

当然，这些态与我们在前几章中称为(+)态及(-)态的是同样的态。尽管如此，为使本章的记号前后一致，我们称“正”自旋态为 $|1\rangle$ ，“负”自旋态为 $|2\rangle$ ，这里“正”与“负”指的是沿

z 轴的角动量。

电子的任何可能状态 ψ 可以像式(10.1)那样通过给出电子处在态 $|1\rangle$ 的振幅 C_1 及处在态 $|2\rangle$ 的振幅 C_2 来描写。为了处理这个问题, 我们需要知道这个双态系统——即处在磁场内的电子的哈密顿。我们从磁场沿 z 方向这种特殊情况开始讨论。

假设矢量 \mathbf{B} 只有 z 分量 B_z 。由两个基础态的定义(即自旋平行于和反平行于 \mathbf{B})知道它们已经是磁场中有确定能量的定态了。态 $|1\rangle$ 对应于 $-\mu B_z$ 的能量*, 而态 $|2\rangle$ 则对应于 $+\mu B_z$ 的能量。在这种情况下, 由于处在态 $|1\rangle$ 的振幅 C_1 不受 C_2 的影响, 反之亦然, 所以哈密顿必定十分简单:

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= E_1 C_1 = -\mu B_z C_1, \\ i\hbar \frac{dC_2}{dt} &= E_2 C_2 = +\mu B_z C_2. \end{aligned} \quad (10.17)$$

对于这种特殊情况, 哈密顿是

$$\begin{aligned} H_{11} &= -\mu B_z, & H_{12} &= 0, \\ H_{21} &= 0, & H_{22} &= +\mu B_z. \end{aligned} \quad (10.18)$$

所以我们知道对于沿 z 方向磁场的哈密顿是什么, 并且也知道定态的能量。

现在假定磁场不在 z 方向上, 那么哈密顿是什么呢? 如场不沿 z 方向, 矩阵元变成什么样子? 我们要提出一个假设, 即哈密顿各项服从一种叠加原理。更具体地说, 我们要假设: 如果两个磁场叠加在一起, 那么, 哈密顿的各项只要相加——如果知道仅有 B_z 时的 H_z , 也知道仅有 B_x 时的 H_x , 那么, 当 B_x 及 B_z 两者都一同存在, H_z 就只是两者分别存在情况下之和。如果我们考虑的只是沿 z 方向的场, 上述结论肯定正确——因为若使 B_z 加倍, 所有的 H_y 都加倍。所以我们假设在场 \mathbf{B} 中 H 是线性的。这就是为了对任何磁场都能求出 H_y 所需的一切了。

假定有个恒定均匀磁场 \mathbf{B} , 我们完全可以选取 z 轴沿着磁场方向, 从而就会找到能量为 $\mp\mu B$ 的两个定态。但是仅仅沿不同方向选取坐标轴并不会改变物理实质。这时我们定态的描述将会不同, 但它们的能量将仍为 $\mp\mu B$, 即

$$\begin{aligned} E_1 &= -\mu\sqrt{B_x^2 + B_y^2 + B_z^2} \\ \text{和} \quad E_2 &= +\mu\sqrt{B_x^2 + B_y^2 + B_z^2}. \end{aligned} \quad (10.19)$$

剩下的事情是容易的。这里已有了能量公式。我们需要一个与 B_x , B_y , B_z 成线性关系的哈密顿, 将它代入一般公式(10.3)中就给出这些能量。问题是: 找到哈密顿。首先, 注意到能级分裂是对称的, 其平均值为零。从式(10.3)我们可以直接看出, 这要求

$$H_{22} = -H_{11}.$$

(注意这与我们已经知道的当 B_x , B_y 都为零时的情况相符, 在该情况下 $H_{11} = -\mu B_z$, 及 $H_{22} = +\mu B_z$ 。) 现在, 如果使式(10.3)的能量与从式(10.19)得出的能量相等, 就有

* 我们将静止能量 $m_e c^2$ 取作我们能量的“零点”, 并且将电子的磁矩 μ 取作负数, 因为 μ 的指向与自旋相反。

$$\frac{(H_{11} - H_{22})^2}{4} + |H_{12}|^2 = \mu^2(B_x^2 + B_y^2 + B_z^2). \quad (10.20)$$

(我们还利用了 $H_{21} = H_{12}^*$ 的事实, 所以 $H_{12}H_{21}$ 也可写成 $|H_{12}|^2$ 。) 再回到对于场沿 z 方向的特殊情况, 上式给出

$$\mu^2 B_z^2 + |H_{12}|^2 = \mu^2 B_z^2.$$

显见在这种特殊情况下 $|H_{12}|$ 必须为零, 这意味着 H_{12} 中不可能有任何 B_z 的项。(记住, 我们曾说过, 所有的项必须与 B_x , B_y , B_z 成线性关系。)

至此我们已发现 H_{11} 与 H_{22} 中有含 B_z 的项, 而 H_{12} 及 H_{21} 则没有。我们可以作一个满足式(10.20)的简单猜测, 即只要设

$$\begin{aligned} H_{11} &= -\mu B_z, \\ H_{22} &= \mu B_z, \end{aligned} \quad (10.21)$$

及

$$|H_{12}|^2 = \mu^2(B_x^2 + B_y^2).$$

结果发现这是唯一可行的办法!

“等一等”——你们会说——“ H_{12} 与 B 并不成线性关系, 式(10.21)给出 $H_{12} = \mu\sqrt{B_x^2 + B_y^2}$ 。”未必。还有另一种确实是线性关系的可能形式, 那就是

$$H_{12} = \mu(B_x + iB_y).$$

实际上, 这样的可能性有好几种, 最一般地, 我们可以写为

$$H_{12} = \mu(B_x \pm iB_y)e^{i\delta},$$

这里 δ 是某个任意的相位。我们应当用哪个符号及什么相位? 结果发现你可以任选一种符号, 任选相位, 而物理结果总是相同的。所以符号及相位的选择只是一种约定。在我们之前已有人选择了负号, 并取 $e^\delta = -1$ 。我们可以照着做, 并写下

$$H_{12} = -\mu(B_x - iB_y), \quad H_{21} = -\mu(B_x + iB_y).$$

(附带地说一下, 这些约定与第6章中所作的几种随意选择有关系, 并且与它们相一致。)

这样, 处在任意磁场中的电子的完整哈密顿是

$$\begin{aligned} H_{11} &= -\mu B_z, \quad H_{12} = -\mu(B_x - iB_y), \\ H_{21} &= -\mu(B_x + iB_y), \quad H_{22} = +\mu B_z. \end{aligned} \quad (10.22)$$

而振幅 C_1 与 C_2 的方程组就是

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= -\mu[B_z C_1 + (B_x - iB_y)C_2], \\ i\hbar \frac{dC_2}{dt} &= -\mu[(B_x + iB_y)C_1 - B_z C_2]. \end{aligned} \quad (10.23)$$

这样我们就找到了在磁场中电子的“自旋态的运动方程”。利用一些物理论证我们猜到了它们, 但任何哈密顿的真正检验在于它应当作出与实验吻合的预言。至今所做过的任何检验都表明, 这些方程是正确的。事实上, 尽管我们的讨论只是恒定场, 但所写出的哈密顿

对于随时间变化的磁场也一样正确。所以我们现在可以用式(10.23)来考察种种有趣问题了。

§ 10-7 磁场中自旋的电子

第一个例子：我们从一个沿 z 方向的恒定磁场开始。这里只有能量为 $\pm \mu B_z$ 的两个定态。假定我们在 x 方向上加上弱磁场。于是方程组看起来与原来的双态问题相像。我们又一次碰到翻转问题，而能级又稍稍分开一些。现在使场的 x 分量随时间变化，譬如按 $\cos \omega t$ 变化。于是方程就与第9章中我们在氨分子上施加一个振荡电场后所得到的方程相同。你们可以用同样方法求出详细解答。你将得到这样的结果：当水平电场在共振频率 $\omega_0 = 2\mu B_z / \hbar$ 附近振荡时，振荡电场会引起从 $+z$ 态至 $-z$ 态——或者相反——的跃迁。这就是我们在第2卷第35章中所描写的磁共振现象的量子力学理论。

也可以利用自旋 $1/2$ 的系统造成一种微波激射器。利用施特恩-格拉赫装置产生一束沿(比方说) $+z$ 方向极化的粒子，再将其送入在恒定磁场中的空腔。腔内的振荡电场与磁矩耦合而诱发跃迁，从而将能量递交给空腔。

现在让我们来考察下面的问题。假设有磁场 B ，它指向极角为 θ 和方位角为 ϕ 的方向，如图 10-10 所示。此外还假定有一个电子，而且我们已使它的自旋方向与磁场方向相同。这一电子的振幅 C_1 与 C_2 是什么？换句话说，令电子状态为 $|\psi\rangle$ ，我们可以写出

$$|\psi\rangle = |1\rangle C_1 + |2\rangle C_2,$$

这里 C_1 与 C_2 为

$$C_1 = \langle 1 | \psi \rangle, \quad C_2 = \langle 2 | \psi \rangle,$$

这里用 $|1\rangle$ 与 $|2\rangle$ 表示与我们习惯称为 $|+\rangle$ 及 $|-\rangle$ (相对于所选的 z 轴而言) 的相同的状态。

这个问题的答案也包括在双态系统的一般方程中。首先我们知道，由于电子自旋平行于 B ，它处在能量 $E_1 = -\mu B$ 的定态中。因此 C_1 与 C_2 两者都必定像式(9.18)那样，按 $e^{-iE_1 t/\hbar}$ 变化，而它们的系数 a_1 与 a_2 则由式(10.5)给定，即

$$\frac{a_1}{a_2} = \frac{H_{12}}{E_1 - H_{11}}, \quad (10.24)$$

一个附加条件是 a_1 与 a_2 应当归一化，使 $|a_1|^2 + |a_2|^2 = 1$ 。从式(10.22)中取出 H_{11} 及 H_{12} ，并利用

$$B_z = B \cos \theta, \quad B_x = B \sin \theta \cos \phi, \quad B_y = B \sin \theta \sin \phi.$$

就有

$$\begin{aligned} H_{11} &= -\mu B \cos \theta, \\ H_{12} &= -\mu B \sin \theta (\cos \phi - i \sin \phi). \end{aligned} \quad (10.25)$$

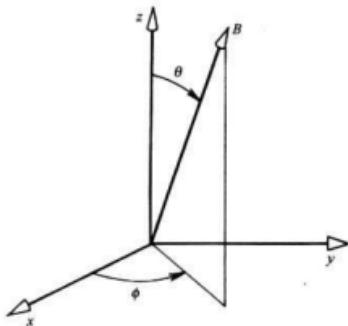


图 10-10 B 的方向用极角 θ 和方位角 ϕ 确定

附带提一下,第二个式子中的最后一个因子是 $e^{-\theta}$,所以这样写更简单些:

$$H_{12} = -\mu B \sin \theta e^{-\theta}. \quad (10.26)$$

在式(10.24)中代入这些矩阵元,并从分子和分母中消去 $-\mu B$,我们发现

$$\frac{a_1}{a_2} = \frac{\sin \theta e^{-\theta}}{1 - \cos \theta}. \quad (10.27)$$

利用这个比值及归一化条件,就可求得 a_1 与 a_2 。这并不困难,但我们可以利用小小的技巧来走一下捷径。注意到:

$$1 - \cos \theta = 2 \sin^2 \frac{\theta}{2},$$

$$\sin \theta = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}.$$

于是式(10.27)等同于

$$\frac{a_1}{a_2} = \frac{\cos \frac{\theta}{2} e^{-\theta}}{\sin \frac{\theta}{2}}. \quad (10.28)$$

一个可能的答案是

$$a_1 = \cos \frac{\theta}{2} e^{-\theta}, \quad a_2 = \sin \frac{\theta}{2}, \quad (10.29)$$

因为它符合式(10.28)。同时也满足

$$|a_1|^2 + |a_2|^2 = 1.$$

正像你们知道的, a_1 与 a_2 都乘以一个任意的相位因子不会改变任何东西。人们一般喜欢将式(10.29)的两者都乘上 $e^{i\phi/2}$ 而使它们更对称些,所以常用的形式就是

$$a_1 = \cos \frac{\theta}{2} e^{-\theta/2}, \quad a_2 = \sin \frac{\theta}{2} e^{i\phi/2}. \quad (10.30)$$

这就是我们问题的答案。当我们知道电子的自旋沿着极角为 θ 、方位角为 ϕ 的方向时, a_1 与 a_2 的数值就是电子自旋沿着 z 轴朝上或朝下的振幅。(振幅 C_1 与 C_2 只是 a_1 与 a_2 乘以 $e^{-i\phi/2}$ 。)

现在我们注意到一件有趣的事。在式(10.30)中任何地方都不出现磁感应强度 B 。显然,在 B 趋向于零的极限情况下,结果也一样。这意味着我们已经一般地回答了怎么来表示一个自旋沿任意方向的粒子的问题了。式(10.30)的振幅是自旋 $1/2$ 粒子的投影振幅,此投影振幅相当于我们在第 5 章中[式(5.38)]所给出的自旋 1 粒子的投影振幅。现在我们能够求得自旋 $1/2$ 的已过滤的粒子束经过任何特定的施特恩-格拉赫装置的振幅了。

设 $|+z\rangle$ 表示自旋沿 z 轴朝上的态, $| -z\rangle$ 则表示自旋朝下的态。如 $|+z'\rangle$ 表示自旋沿 z' 轴朝上的态,而 z' 轴与 z 轴成极角 θ 和方位角 ϕ ,那么按第 5 章的记号,我们有

$$\langle +z | +z' \rangle = \cos \frac{\theta}{2} e^{-\theta/2}, \quad \langle -z | +z' \rangle = \sin \frac{\theta}{2} e^{i\phi/2}. \quad (10.31)$$

这些结果与第 6 章中我们利用纯几何论证求得的式子(6.36)等价。(所以如果你已经决定

跳过第6章的话,现在也得到了主要结果。)

作为最后一个例子,我们再来考察一件多次提到过的事情。我们考虑以下问题。我们从某个给定自旋方向的电子开始,接着加上沿 z 方向的磁场25分钟,然后撤去磁场。最后的状态是什么?我们还是用线性组合来表示该状态 $|\psi\rangle = |1\rangle C_1 + |2\rangle C_2$ 。然而,对这个问题,具有确定能量的态也就是我们的基础态 $|1\rangle$ 及 $|2\rangle$ 。所以 C_1 与 C_2 只在相位上变化。我们知道

$$C_1(t) = C_1(0)e^{-iE_1 t/\hbar} = C_1(0)e^{+i\mu B_z t/\hbar},$$

$$C_2(t) = C_2(0)e^{-iE_2 t/\hbar} = C_2(0)e^{-i\mu B_z t/\hbar}.$$

起初我们就已说过电子自旋沿着给定方向。这意味着原先的 C_1 与 C_2 是式(10.30)所给出的两个数。当我们等待了一段时间 T 之后,新的 C_1 与 C_2 就是原先两个数分别乘以 $e^{i\mu B_z T/\hbar}$ 和 $e^{-i\mu B_z T/\hbar}$ 。那是什么态?这不难回答。这正是方位角 ϕ 减少了 $2\mu B_z T/\hbar$,而极角 θ 保持不变的态。那意味着在时间 T 的终了,态 $|\psi\rangle$ 表示自旋的方向与原来方向的不同只是绕 z 轴转动了一个角度 $\Delta\phi = 2\mu B_z T/\hbar$ 。因为这个角度正比于 T ,我们也可说自旋的方向以角速度 $2\mu B_z/\hbar$ 绕 z 轴作进动。这个结果我们以前曾以不那么完全和严格的方式讨论过好几次。现在,我们已对原子磁体的进动得到了一个完整与精确的量子力学描述。

有趣的是,刚才对磁场中自旋的电子所用过的数学概念可适用于任何双态系统。这就是说,通过与自旋电子作数学类比,有关双态系统的任何问题都可用纯几何的方法加以解决。做法是,首先,移动能量的零点使得 $(H_{11} + H_{22})$ 等于零,因而有 $H_{11} = -H_{22}$ 。于是任何双态问题在形式上与处在磁场中的电子的问题相同。你必须做的一切就是把 $-\mu B_z$ 等同于 H_{11} ,把 $-\mu(B_x - iB_y)$ 等于 H_{12} 。不管原先的物理内容如何——氨分子或其他什么东西——你都可以将它转换为相应的电子的问题。

所以,只要我们能够一般地解决了电子的问题,我们就能解决所有的双态问题。

而我们已经有了对电子问题的一般解答!

假设电子起初在自旋对某方向“朝上”的状态,而你们指向另外某个方向施加指向的磁场 \mathbf{B} 。你只要将自旋方向绕 \mathbf{B} 的方向以某个角速度矢量 $\omega(t)$ 旋转, $\omega(t)$ 等于一个常数乘以矢量 \mathbf{B} (即 $\omega = 2\mu\mathbf{B}/\hbar$)。若 \mathbf{B} 随时间变化,你要始终使转轴

平行于 \mathbf{B} ,并改变转动速率以使它总是与 \mathbf{B} 的强度成正比。参见图10-11。如果你坚持这样做,最后就会得到自旋的最终确定指向,而振幅 C_1 与 C_2 就可以由将式(10.30)在你所用坐标系内投影求得。你看,这正好是个几何学问题:跟踪着所有的转动之后到达的最终位置。尽管弄清所涉及的内容并不难,但在一般情况下直接而明了地解出这个几何问题(求出以变化的角速度矢量进行转动的最终结果)不容易。但是,无论如何,在原则上我们已知道了任何双态问题的一般解。下一章里,我们将进一步考察用来处理自旋 $1/2$ 粒子这一重要情况——也可用来处理一般的双态系统——的数学技巧。

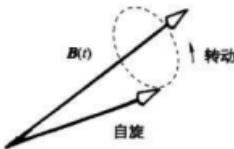


图10-11 在变化磁场 $\mathbf{B}(t)$ 中电子自旋的方向以频率 $\omega(t)$ 绕平行于 \mathbf{B} 的轴进动

第 11 章 再论双态系统

§ 11-1 泡利自旋矩阵*

我们现在继续讨论双态系统。上一章末了我们讨论了处在磁场中的自旋 $1/2$ 粒子。我们用自旋角动量的 z 分量为 $+\hbar/2$ 的振幅 C_+ 和自旋分量为 $-\hbar/2$ 的振幅 C_- 描述粒子的自旋态。在前几章中我们曾称这些基础态为 $|+\rangle$ 及 $|-\rangle$ ，现在我们再回过来用这种记号法。虽然有时我们会觉得交替使用 $|+\rangle$ 或 $|1\rangle$, $|-\rangle$ 或 $|2\rangle$ 较为方便。

在上一章中我们看到, 当磁矩为 μ 的自旋 $1/2$ 粒子在磁场 $\mathbf{B} = (B_x, B_y, B_z)$ 中, 振幅 $C_+ (= C_1)$ 和 $C_- (= C_2)$ 由下列微分方程联系起来:

$$\begin{aligned} i\hbar \frac{dC_+}{dt} &= -\mu [B_z C_+ + (B_x - iB_y) C_-], \\ i\hbar \frac{dC_-}{dt} &= -\mu [(B_x + iB_y) C_+ - B_z C_-]. \end{aligned} \quad (11.1)$$

换言之, 哈密顿矩阵 H_0 是

$$\begin{aligned} H_{11} &= -\mu B_z, & H_{12} &= -\mu (B_x - iB_y), \\ H_{21} &= -\mu (B_x + iB_y), & H_{22} &= +\mu B_z. \end{aligned} \quad (11.2)$$

因而, (11.1)式就等同于

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j, \quad (11.3)$$

这里 i 与 j 取十与一(或 1 与 2)。

电子自旋这一双态系统是如此重要, 因而书写上使用更简洁的方法是非常有用的。我们现在要离题讲一点数学, 告诉你人们通常是怎么写双态系统方程的。可以这样做: 首先, 注意哈密顿矩阵中的每一项都正比于 μ 以及 \mathbf{B} 的某些分量, 于是我们可以——纯粹形式地——写出

$$H_0 = -\mu [\sigma_y^x B_x + \sigma_y^y B_y + \sigma_y^z B_z], \quad (11.4)$$

这里毫无新的物理内容, 上式只意味着, 系数 σ_y^x , σ_y^y , σ_y^z (它们共有 $4 \times 3 = 12$ 个)都可以求出来, 所以式(11.4)和式(11.2)完全相同。

让我们看看它们必须取什么值。从 B_z 开始, 由于 B_z 只出现在 H_{11} 与 H_{22} 中, 只要取

$$\begin{aligned} \sigma_{11}^z &= 1, & \sigma_{12}^z &= 0, \\ \sigma_{21}^z &= 0, & \sigma_{22}^z &= -1. \end{aligned}$$

* 在第一次阅读这本书的时候这一节可以跳过。这一节的内容比适合于第一级课程的材料要更深一些。

那就一切都解决了。我们常常将矩阵 H_θ 写成这样一个小表格：

$$H_\theta = \begin{array}{c} j \rightarrow \\ i \downarrow \end{array} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}.$$

对一个处在磁场 B_z 中的自旋 $1/2$ 粒子的哈密顿来说，上式就和

$$H_\theta = \begin{array}{c} j \rightarrow \\ i \downarrow \end{array} \begin{bmatrix} -\mu B_z & -\mu(B_x - iB_y) \\ -\mu(B_x + iB_y) & +\mu B_z \end{bmatrix}.$$

一样。用同样方法，我们可以把系数 σ_i^x 写成矩阵

$$\sigma_\theta^x = \begin{array}{c} j \rightarrow \\ i \downarrow \end{array} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (11.5)$$

看一下 B_x 的系数，我们得到 σ_x 的各项必须是

$$\begin{aligned} \sigma_{11}^x &= 0, & \sigma_{12}^x &= 1, \\ \sigma_{21}^x &= 1, & \sigma_{22}^x &= 0; \end{aligned}$$

或简写为

$$\sigma_\theta^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (11.6)$$

最后，看一看 B_y ，我们得到

$$\begin{aligned} \sigma_{11}^y &= 0, & \sigma_{12}^y &= -i, \\ \sigma_{21}^y &= i, & \sigma_{22}^y &= 0; \end{aligned}$$

或

$$\sigma_\theta^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (11.7)$$

有了这 3 个 σ 矩阵，式(11.2)就与式(11.4)恒等。我们将 x, y, z 作为 σ 的上标，指明哪个 σ 与哪个 \mathbf{B} 的分量相对应，而将下标留给 i 与 j 。但通常 i, j 略去不写——因为不难想象它们是在哪儿——而把 x, y, z 写作下标。这样式(11.4)就写成

$$H = -\mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z). \quad (11.8)$$

因为 σ 矩阵非常重要（它们一直被专家们使用着），我们在表 11-1 中将它们汇总在一起。（任何打算从事量子物理学工作的人都一定得记住它们。）人们也以发明这些矩阵的物理学家的名字将它们命名为泡利自旋矩阵。

在这张表上我们还多加了一个 2×2 矩阵，如果我们要处理一个含有带同样能量的两个自旋态的系统，或者想选择一个不同零点能量的话，就需要那样一个矩阵。在这些情况下我们必须在式(11.1)的第一个方程中加上 $E_0 C_+$ ，而在第二个方程中加上 $E_0 C_-$ 。如果定义单位矩阵“1”为 δ_0 ，

表 11-1 泡利自旋矩阵

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$I = \delta_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (11.9)$$

并将其改写为

$$H_0 = E_0 \delta_{ij} - \mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z), \quad (11.10)$$

就能把以上两项包括在我们的新记法中。通常不用说明就会明白任何像 E_0 那样的常数都自动地与单位矩阵相乘，于是上式简写为

$$H = E_0 - \mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z). \quad (11.11)$$

自旋矩阵之所以有用，其中一个理由是，任何 2×2 矩阵都能用它们表示。任何一个你能够写出的这种矩阵中都有 4 个数，比方说

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

它们总可以写为 4 个矩阵的线性组合。例如：

$$M = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

有许多这种组合法，但有一种特殊方法是将 M 看作一定量的 σ_x 加一定量的 σ_y ，等等，形如

$$M = \alpha I + \beta \sigma_x + \gamma \sigma_y + \delta \sigma_z,$$

这里“数量” α, β, γ 与 δ 一般来说可以是复数。

既然任何 2×2 矩阵都可用单位矩阵与 σ 矩阵表示，我们就有了处理任何双态系统所需要的一切了。不管双态系统是什么——氦分子、品红染料或者其他任何东西——哈密顿方程都可用 σ 矩阵写出。虽然在电子处于磁场中这种物理条件下可以看出 σ 矩阵具有几何意义，但也可以将它们看作只是适用于任何双态问题的有用矩阵而已。

例如，在某种看法下，可以把质子与中子看作为处于两种状态之一的同一种粒子，我们说核子（质子或中子）是双态系统——在这种情况下，两个态是就关系到它们是否带电荷。照这种观点，态 $|1\rangle$ 可表示质子，而态 $|2\rangle$ 可表示中子。人们便说核子有两个“同位旋”状态。

由于我们要使用 σ 矩阵作为双态系统量子力学的“算术”，所以让我们很快地复习一下矩阵代数的一些规则。所谓两个或几个矩阵的“和”的意义在式(11.4)中很明显。一般说，如果将两个矩阵 A 与 B “相加”，那么它们的“和” C 就意味着其每一项 C_{ij} 由下式给出：

$$C_{ij} = A_{ij} + B_{ij}.$$

C 的每一项是 A 与 B 中同样位置的两项之和。

在 § 5-6 中我们已接触到矩阵“积”的概念。这个概念也同样用于处理 σ 矩阵。一般而言，两个矩阵 A 和 B （按此顺序）的“积”定义为矩阵 C ，它的元素是

$$C_{ij} = \sum_k A_{ik} B_{kj}. \quad (11.12)$$

这是从 A 的第 i 行及 B 的第 j 列取出相应一对元素相乘然后求和。如果矩阵写成图 11-1 那样的表格形式，就有一个求出积矩阵各项很好的“系统”。假定你要计算 C_{23} ，将你的左手食指顺着 A 的第二行依次移动，右手食指顺着 B 的第三列向下依次移动，将移动时所遇到的每一对相乘，再相加。我们已试着在图中说明具体做法。

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{pmatrix}$$

$$C_{ij} = \sum_k A_{ik} B_{kj}$$

例

$$C_{23} = A_{21} B_{13} + A_{22} B_{23} + A_{23} B_{33} + A_{24} B_{43}$$

图 11-1 两个矩阵相乘

当然，对 2×2 矩阵来说特别简单。例如，如果我们将 σ_x 乘 σ_x ，就得到

$$\sigma_x^2 = \sigma_x \cdot \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

这正是单位矩阵 1。或举另一个例子，我们来算出 $\sigma_x \sigma_y$ ：

$$\sigma_x \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

参照表 11-1，可看出乘积正是 i 乘上矩阵 σ_y 。（请记住一个数与矩阵相乘，就是该数与矩阵的每一项相乘。）因为一次求两个 σ 矩阵的积很重要，也相当有趣，所以我们已将这些积全部列在表 11-2 内。你可以像求 σ_x^2 及 $\sigma_x \sigma_y$ 那样求出它们来。

表 11-2 自旋矩阵的积

$\sigma_x^2 = 1$	$\sigma_y^2 = 1$	$\sigma_z^2 = 1$
$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$		
$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x$		
$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y$		

关于这些 σ 矩阵，还有一个十分重要而有趣之点。如果愿意，我们可以设想 3 个矩阵 σ_x ， σ_y 及 σ_z 类似于某个矢量的 3 个分量——有时把它称为“ σ 矢量”——而记为 σ 。它确实是个“矩阵矢量”或“矢量矩阵”。它是 3 个不同的矩阵，每一个矩阵分别和 x ， y 或 z 轴中的一个相联系。由此，我们可把系统的哈密顿以在任何坐标系内都成立的简洁形式写出：

$$H = -\mu \sigma \cdot \mathbf{B}. \quad (11.13)$$

虽然这3个矩阵是在这样的表示中写下的，在该表示中“朝上”与“朝下”是对z轴而言的（因而 σ_z 特别简单）。但我们也就可以求出在其他某个表示中这些矩阵的形式。尽管需要作许多代数运算，你能够证明这些矩阵之间的变换就像一个矢量的分量的变换一样。（然而，我们此刻不想去操心证明这点，如果你愿意，你可以验证这一点。）你们可以在不同的坐标系下应用 σ ，就仿佛它是个矢量一样。

你们记得在量子力学中 H 与能量有关。事实上，在只有一个态的简单状况下， H 正好等于能量。即使对电子自旋的双态系统，当我们将哈密顿写成式(11.13)那样时，它非常像一个磁矩为 μ 的小磁体处在磁场 \mathbf{B} 中的能量的经典公式。在经典物理学中，我们有

$$U = -\mu \cdot \mathbf{B}, \quad (11.14)$$

这里 μ 是物体的性质， \mathbf{B} 是外磁场。如果用哈密顿代替经典的能量，用矩阵 $\mu \sigma$ 代替经典的一个 μ ，我们可以看出式(11.14)能够转换为式(11.13)。于是，根据这种纯粹形式的代换，我们将结果解释为矩阵方程。有时人们说，对经典物理中的每个量，量子力学中都有一个矩阵与之对应，实际上更确切的说法是哈密顿矩阵对应着能量，而任何可以通过能量来定义的量都有着相应的矩阵。

例如，磁矩可以通过能量来定义，只要指出它在外场 \mathbf{B} 中的能量是 $-\mu \cdot \mathbf{B}$ 。这就定义了磁矩矢量 μ 。然后我们考察处在磁场中的真实（量子）客体的哈密顿的公式，并试着去辨别这些矩阵与经典公式中相对应的各种物理量。这就是有时可以找到和经典物理量对应的量子力学物理量的技巧。

如果愿意的话，你们可以试试看去弄清一个经典矢量怎么会跟一个矩阵 $\mu \sigma$ 相等，或许你们会发现一些东西——但切勿为之太伤脑筋。那个想法不妥当——它们并不相等。量子力学是另一种类型的描述世界的理论。只是碰巧存在着一定的对应关系，但这至多是记忆的工具——用它来帮助记忆。就是说，当你学习经典物理时，你记住了式(11.14)；因而如果你记得对应关系 $\mu \rightarrow \mu \sigma$ ，你就容易记住式(11.13)了。当然，自然界遵循量子力学，而经典力学只是近似；因此，毫不奇怪在经典力学中会有量子力学定律的某种影子——量子力学定律正是经典力学的基础。用任何直接方式从影子重建原物是不可能的，但影子确实会帮助你记住原物像什么样子。式(11.13)是真理，而式(11.14)是影子。因为我们先学习经典力学，所以希望能由它得到量子力学公式，但这样做时根本不存在什么肯定成功的方案。我们必须一再回到真实世界以发现正确的量子力学方程。如得到的方程跟经典物理中的某个东西相像，就是我们的幸运。

如果上述有关经典物理与量子物理之间关系的告诫显得太啰嗦，并且在你们看来都是毋庸待言的自明之理，那就请你们原谅这位教授的条件反射吧，他通常是对进研究生院前并没听说过泡利自旋矩阵的学生讲授量子力学。这些学生总像是抱着某种希望，希望量子力学是他们几年前已透彻学过的经典力学逻辑发展的结果。（或许他们想避免不得不学习新东西这件事吧。）你们只是在几个月前学习了经典公式(11.14)——并且接着又被提醒说它并不是合适的公式——所以你们大概不会很不愿意将量子力学公式(11.13)来作为基本的真理吧。

§ 11-2 作为算符的自旋矩阵

当我们正讨论数学记号这一题目时, 我们想再描写一下另一种书写方法, 这种写法很常用, 因为它非常简洁。它直接由第 8 章所引进的记法得到。假定我们有个处在随时间变化的状态 $|\psi(t)\rangle$ 的系统, 我们可像在式(8.34)中所做的那样把系统于时刻 $t + \Delta t$ 在态 $|i\rangle$ 的振幅写成

$$\langle i | \psi(t + \Delta t) \rangle = \sum_j \langle i | U(t + \Delta t, t) | j \rangle \langle j | \psi(t) \rangle.$$

矩阵元 $\langle i | U(t + \Delta t, t) | j \rangle$ 是在时间间隔 Δt 内基础态 $|j\rangle$ 转变为基础态 $|i\rangle$ 的振幅。于是我们可以写出下式来定义 H_{ij} :

$$\langle i | U(t + \Delta t, t) | j \rangle = \delta_{ij} - \frac{i}{\hbar} H_{ij}(t) \Delta t,$$

我们已经证明各振幅 $C_i(t) = \langle i | \psi(t) \rangle$ 之间由下列微分方程组相联系:

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j, \quad (11.15)$$

假如明确写出振幅 C_i , 那么上式成为

$$i\hbar \frac{d}{dt} \langle i | \psi \rangle = \sum_j H_{ij} \langle j | \psi \rangle. \quad (11.16)$$

但矩阵元 H_{ij} 也是振幅, 可以将它写为 $\langle i | H | j \rangle$; 所以我们的微分方程就变为这样:

$$i\hbar \frac{d}{dt} \langle i | \psi \rangle = \sum_j \langle i | H | j \rangle \langle j | \psi \rangle. \quad (11.17)$$

我们看到 $(-i/\hbar) \langle i | H | j \rangle dt$ 就是在 H 所描写的物理条件下态 $|j\rangle$ 在时间 dt 内将“产生”出态 $|i\rangle$ 的振幅。(所有这些都已隐含在 § 8-4 的讨论之中。)

现在依照 § 8-2 的思路, 我们可以丢掉式(11.17)中的公共项 $\langle i |$ ——因为此式对任何态 $|i\rangle$ 都成立——而把该方程简写为

$$i\hbar \frac{d}{dt} |\psi\rangle = \sum_j H |j\rangle \langle j | \psi \rangle. \quad (11.18)$$

或再进一步, 我们还可去掉 j 而写为

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle. \quad (11.19)$$

在第 8 章里我们曾指出, 表式写成这种形式时, 在 $H|j\rangle$ 或 $H|\psi\rangle$ 中的 H 称为算符。从现在开始我们要给算符戴上一顶小帽子($\hat{\cdot}$)来提醒你这是个算符而不是个数字。我们以后写成 $\hat{H}|\psi\rangle$ 。虽然两个方程式(11.18)及(11.19)和式(11.17)或(11.15)的意义完全相同, 我们却可以用不同的方式看待它们。例如, 我们可以这样来描写式(11.18):“态矢量 $|\psi\rangle$ 对时间的导数乘以 $i\hbar$ 等于将哈密顿算符 \hat{H} 作用于每个基础态后, 乘上 ψ 在某一态 j 中的振幅 $\langle j | \psi \rangle$, 然后再对所有 j 求和所得的结果。”而式(11.19)可以这样来描写:“态 $|\psi\rangle$ 对时间的导

数(乘上 $i\hbar$)等于用哈密顿算符 \hat{H} 作用在态矢量 $|\psi\rangle$ 上所得到的结果。”这只是对式(11.17)中内容的一种简短说法而已,但是,你们会看到,那是十分便利的。

如果我们高兴,还可以把“抽象”的思想再往前推进一步。方程式(11.19)对任何状态 $|\psi\rangle$ 都正确。等式左边的 $i\hbar d/dt$ 也是个算符——它是“对 t 求导再乘上 $i\hbar$ ”的运算。所以式(11.19)也可以认为是算符之间的一个方程,即算符方程

$$i\hbar \frac{d}{dt} = \hat{H}.$$

哈密顿算符(除一个常数因子之外)作用到任何态上所得的结果与 d/dt 的作用相同。请记住这个方程——像式(11.19)一样——并非 \hat{H} 算符恰与 $i\hbar d/dt$ 是恒等运算的表述。这些方程是自然界量子体系的动力学定律,即运动定律。

为用这些概念做些练习,我们给你证明可以换一种方法得到式(11.18)。你们知道,我们可以用一个态 $|\psi\rangle$ 在某一组基础态上的投影来表示这个态[参见式(8.8)],

$$|\psi\rangle = \sum_i |i\rangle \langle i| \psi\rangle. \quad (11.20)$$

态 $|\psi\rangle$ 怎样随时间变化呢? 只要求它的导数:

$$\frac{d}{dt} |\psi\rangle = \frac{d}{dt} \sum_i |i\rangle \langle i| \psi\rangle. \quad (11.21)$$

设基础态 $|i\rangle$ 不随时间变化(至少我们总是把它们取为确定不变的态),但振幅 $\langle i|\psi\rangle$ 是个可能变化的数。于是式(11.21)变为

$$\frac{d}{dt} |\psi\rangle = \sum_i |i\rangle \frac{d}{dt} \langle i| \psi\rangle. \quad (11.22)$$

从式(11.16)可以知道 $d\langle i|\psi\rangle/dt$,所以我们得到

$$\begin{aligned} \frac{d}{dt} |\psi\rangle &= -\frac{i}{\hbar} \sum_i |i\rangle \sum_j H_{ij} \langle j| \psi\rangle \\ &= -\frac{i}{\hbar} \sum_j |i\rangle \langle i| H |j\rangle \langle j| \psi\rangle \\ &= -\frac{i}{\hbar} \sum_j H |j\rangle \langle j| \psi\rangle. \end{aligned}$$

这就是式(11.18)。

所以我们有多种看待哈密顿的方法。可以把一组系数 H_{ij} 看作为只是一组数字,或把哈密顿算符看作“振幅” $\langle i|H|j\rangle$,或看作“矩阵” H_{ij} ,或看作“算符” \hat{H} 。它们全都表示同一事物。

现在让我们回到双态系统上来。假如我们用 σ 矩阵(带有适当的数值系数,如 B_z 等等)表示哈密顿,我们也能把 σ_0^x 明确地看作振幅 $\langle i|\sigma_x|j\rangle$,或者,简言之就是算符 $\hat{\sigma}_x$ 。如果使用算符的概念,我们可以把磁场中的态 $|\psi\rangle$ 的运动方程写成

$$i\hbar \frac{d}{dt} |\psi\rangle = -\mu(B_x \hat{\sigma}_x + B_y \hat{\sigma}_y + B_z \hat{\sigma}_z) |\psi\rangle. \quad (11.23)$$

当我们想要“使用”这样一个方程时,通常必须用基矢表示态 $|\psi\rangle$ (这就像要用特定的数值来表示空间矢量就一定要求出空间矢量的分量一样)。所以我们通常把式(11.23)写成展开式:

$$i\hbar \frac{d}{dt} |\psi\rangle = -\mu \sum_i (B_x \hat{\sigma}_x + B_y \hat{\sigma}_y + B_z \hat{\sigma}_z) |i\rangle \langle i| \psi\rangle. \quad (11.24)$$

现在你们会明白为什么算符概念是如此简洁了。为了利用式(11.24),我们要知道 $\hat{\sigma}$ 算符作用在各个基础态上会得到什么结果。让我们求出它们来。假定我们求 $\hat{\sigma}_z |+\rangle$, 它是某个矢量 $|?\rangle$, 但是, 是什么矢量呢? 别急, 让我们在 $\hat{\sigma}_z |+\rangle$ 上左乘 $\langle +|$, 那就有

$$\langle +| \hat{\sigma}_z |+\rangle = \sigma_{z+}^* = 1.$$

(利用表 11-1), 所以我们知道

$$\langle +| ?\rangle = 1. \quad (11.25)$$

现在, 让我们在 $\hat{\sigma}_z |+\rangle$ 上左乘 $\langle -|$, 我们得到

$$\langle -| \hat{\sigma}_z |+\rangle = \sigma_{z+}^* = 0,$$

所以

$$\langle -| ?\rangle = 0. \quad (11.26)$$

只有一个态矢量同时满足(11.25)和(11.26)两式, 这就是 $|+\rangle$ 。于是我们发现

$$\hat{\sigma}_z |+\rangle = |+\rangle. \quad (11.27)$$

利用这样的论证, 你们很容易证明 σ 矩阵的所有性质可以用表 11-3 所列的一组法则以算符记号来描写。

表 11-3

$\hat{\sigma}$ 算符的性质		
$\hat{\sigma}_z +\rangle = +\rangle$	$\hat{\sigma}_x +\rangle = -\rangle$	$\hat{\sigma}_y +\rangle = i -\rangle$
$\hat{\sigma}_z -\rangle = - -\rangle$	$\hat{\sigma}_x -\rangle = +\rangle$	$\hat{\sigma}_y -\rangle = -i +\rangle$

如果我们求 σ 矩阵的乘积, 它们就变为求算符的乘积。当两个算符作为乘积一起出现时, 你们要先将最靠右的算符进行运算。例如, 对 $\hat{\sigma}_x \hat{\sigma}_y |+\rangle$, 我们必须理解为 $\hat{\sigma}_x (\hat{\sigma}_y |+\rangle)$, 由表 11-3, 求得 $\hat{\sigma}_x \hat{\sigma}_y |+\rangle = i|-\rangle$, 所以

$$\hat{\sigma}_x \hat{\sigma}_y |+\rangle = \hat{\sigma}_x (i|-\rangle). \quad (11.28)$$

而对任何数——譬如 i ——只要从算符中移出就可以了(算符只作用在态矢量上), 所以式(11.28)与下式相同:

$$\hat{\sigma}_x \hat{\sigma}_y |+\rangle = i \hat{\sigma}_x |-\rangle = i|+\rangle.$$

如果对 $\hat{\sigma}_x \hat{\sigma}_y |-\rangle$ 作同样运算, 你们就得到

$$\hat{\sigma}_x \hat{\sigma}_y |-\rangle = -i |-\rangle.$$

查一下表 11-3, 你们会看出 $\hat{\sigma}_x \hat{\sigma}_y$ 作用在 $|+\rangle$ 或 $|-\rangle$ 上正好得到 $\hat{\sigma}_z$ 作用其上的结果再乘以 $-i$ 。所以我们可以说 $\hat{\sigma}_x \hat{\sigma}_y$ 的作用与 $i\hat{\sigma}_z$ 的作用相同, 并把这个表述写为算符方程:

$$\hat{\sigma}_x \hat{\sigma}_y = i\hat{\sigma}_z. \quad (11.29)$$

注意这个方程与表 11-2 中所列的矩阵方程之一相同。所以我们又一次见到了矩阵观点和算符观点之间的对应性。因而表 11-2 中的每个方程都可以看作是关于 σ 算符的方程。你们可以验证一下它们确实可由表 11-3 得到。在作这些验证时, 最好不要去管 σ 或 H 这些量究竟是算符还是矩阵。不论从哪种观点来看, 所有的方程都相同, 所以表 11-2 既适用于 σ 算符, 也适用于 σ 矩阵, 随你的便。

§ 11-3 双态方程的解

现在我们可用多种形式写出双态方程。例如可写为

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j,$$

也可写为

$$i\hbar \frac{d|\psi\rangle}{dt} = \hat{H} |\psi\rangle. \quad (11.30)$$

它们表示同一回事。对处在磁场中的一个自旋 $1/2$ 粒子, 哈密顿 H 由式(11.8)或式(11.13)给出。

如果磁场沿 z 方向, 那么——正像到现在为止我们已屡次见到的那样——解就是态 $|\psi\rangle$ (不管它是什么) 绕着 z 轴进动(就好像你取一个物体, 使它整个地绕着 z 轴旋转一样), 它的进动角速度等于场强乘以 μ/\hbar 的两倍。当然, 当磁场沿其他任何方向时, 情况同样如此, 因为物理规律与坐标系无关。假如有这样一种状况, 磁场以复杂的方式随着时间变化, 那么我们可以用下述方式加以分析。假定开始时自旋沿着 $+z$ 方向, 而磁场沿 x 方向。自旋就开始转向。接着如果 x 方向磁场撤消, 自旋就停止转向。如果这时加上一个 z 方向磁场, 自旋就绕着 z 轴进动, 等等。所以根据磁场按时间变化的方式, 你可以描绘出终态是什么——它将指向哪个轴的方向。然后你就可用第 10 章(或第 6 章)的投影公式变换回到原先的对于 z 轴的 $|+\rangle$ 态与 $|-\rangle$ 态来表示这个态。如果终态的自旋的指向沿着 (θ, ϕ) 方向, 它就会有一个朝上的振幅 $\cos(\theta/2)e^{-i\phi/\hbar}$ 和一个朝下的振幅 $\sin(\theta/2)e^{+i\phi/\hbar}$ 。这就解决了任何问题。这是微分方程解法的语言描述。

刚才所述的求解方法相当普遍, 可以处理任何双态系统。我们不妨取氢分子为例子——其中还包括电场的效应。假如我们用态 $|I\rangle$ 及 $|II\rangle$ 描写系统, 方程式(9.38)和(9.39)就写成:

$$\begin{aligned} i\hbar \frac{dC_1}{dt} &= +AC_1 + \mu e C_{\parallel}, \\ i\hbar \frac{dC_{\parallel}}{dt} &= -AC_{\parallel} + \mu e C_1. \end{aligned} \quad (11.31)$$

你们会说：“不对，我记得里面还有个电场 E_0 。”不错，但我们也已移动了能量的原点，使 E_0 为 0。（通过使两个振幅改变同样的因子 $e^{iE_0 T/\hbar}$ 总可做到这一点——为了去掉任何常数能量。）既然相应的方程总有相同的解，那我们实在不必再去解一次方程。假如我们看一下这些方程，再看下方程式(11.1)，那我们就可进行下述判断。我们可称 $|I\rangle$ 为态 $|+\rangle$ ， $|II\rangle$ 为态 $|-\rangle$ 。这并不意味着我们把氨分子在空间排列起来，也不是 $|+\rangle$ 及 $|-\rangle$ 与 z 轴有任何关系。这纯粹是人为的。我们有一个人造的空间，可以称为“氨分子表示空间”或别的什么名称——一个三维的“构图”，其中的“朝上”对应于分子处在 $|I\rangle$ 态，而沿着假想的 z 轴“朝下”则表示分子处在 $|II\rangle$ 态。这样，对这些方程就可作如下的理解。首先，你知道哈密顿可用 σ 矩阵表示为

$$H = +A\sigma_z + \mu\epsilon\sigma_x. \quad (11.32)$$

或者，换一种方式说，式(11.1)中的 μB_z 对应于式(11.32)中的 $-A$ ，而 μB_x 对应于 $-\mu\epsilon$ 。于是，在我们的“模型”空间中，有一个沿 z 方向的恒定 B 场。如果我们有一个随时间而变的电场 ϵ ，那我们就有—个沿 x 方向按比例变化的 B 场。所以，处在一个 z 方向分量保持不变而 x 方向振荡分量的磁场中的电子的行为，数学上类似于并且完全对应于一个处在振动电场中的氨分子的行为。遗憾的是，我们没有时间进一步讨论这种对应的详情，也没有时间讨论任何技术细节。我们只想指出，所有双态系统，都与一个在磁场中进动的自旋 $1/2$ 粒子相类似。

§ 11-4 光子的偏振态

还有另外许多双态系统研究起来很有趣，我们要谈的第一个新系统是光子。为了描写一个光子，我们必须首先给出它的动量矢量。对自由光子，频率由动量决定，所以我们无须再说它的频率是多少。然而，除动量外我们还须考虑一个称为偏振的性质。设想有个确定单色频率的光子向着你射来（频率在这一讨论中始终保持不变，所以没有许多种不同的动量态）。就有两个偏振方向。在经典理论中，光可以描写为具有（譬如说）一个水平振动的电场和一个垂直振动的电场，这两类光称为 x 偏振光及 y 偏振光。光也可能沿其他某个方向偏振，这可以由沿 x 方向的场与沿 y 方向的场的叠加构成。或者，如果你使 x 分量与 y 分量之间的相位相差 90° ，就会得到一个旋转的电场——这样的光是椭圆偏振光。（这只不过是我们在第 1 卷第 38 章所学过的偏振光经典理论的简单回顾。）

现在假设我们有单个光子——只是一个。这里不存在可以用同样方法讨论的电场，我们所有的只是一个光子。但是光子必定有经典偏振现象的类似的性质。必定存在至少两种不同的光子。起先你可能会以为应当有无限多种——因为电矢量可以指向各种方向。但我们可以把光子的偏振作为双态系统来描写。光子可以处在 $|x\rangle$ 态或处在 $|y\rangle$ 态。 $|x\rangle$ 态指的是在经典物理学中是沿 x 方向振动的一束光线里的每个光子的偏振态。而 $|y\rangle$ 态指的是沿 y 方向振动的光束中的每个光子的偏振态。^{*} 我们可以把 $|x\rangle$ 与 $|y\rangle$ 作为具有指向你的给定

* 为明确起见，并且为符合现在国内偏振光的通用命名法则，我们将原文 x -polarized light 译成沿 x 方向振动的光，振动方向就是电磁波电场的方向。——译者注

动量——我们将称该指向为 z 方向——的光子的基础态。所以有 $|x\rangle$ 与 $|y\rangle$ 两个基础态，它们就是描写任何光子所需要的一切。

举例来说，如果我们有块偏振片，其取向能让沿 x 轴振动的偏振光通过，我们往这偏振片射去一个我们已知道是 $|y\rangle$ 态的光子，它将被偏振片吸收。假如射去一个我们知道是 $|x\rangle$ 态的光子，它将径直通过偏振片后仍旧是 $|x\rangle$ 。如果我们取一块方解石，它让一束偏振光分解为一个 $|x\rangle$ 光束及一个 $|y\rangle$ 光束，这块方解石完全类似于把一束银原子分裂为 $|+\rangle$ 及 $|-\rangle$ 两种状态的施特恩-格拉赫装置。所以先前我们对粒子和施特恩-格拉赫装置所做的每一件事都能重新用在光和方解石上。

那么，当光透过一块透振方向角度为 θ 的偏振片时，会出现什么情况呢？是否变成另一种状态？是的，它的确是另一种状态。我们称偏振片的透振轴为 x' ，以便跟我们的基础态的轴区分开来。参见图 11-2。射来的光子将处在 $|x'\rangle$ 态。但任何态都可以用基础态的线性组合来表示，这里，组合的公式是

$$|x'\rangle = \cos \theta |x\rangle + \sin \theta |y\rangle. \quad (11.33)$$

这就是说，如果光子通过一块透振方向（和 x 轴的）夹角为 θ 的偏振片后，它仍可以（比方说，用一块方解石）分解为 $|x\rangle$ 束与 $|y\rangle$ 束。或者，如果愿意，你可以只在想象中把它分解为 x 分量与 y 分量。无论用哪种方法，你都将得到处在 $|x\rangle$ 态的振幅是 $\cos \theta$ ，处在 $|y\rangle$ 态的振幅是 $\sin \theta$ 。

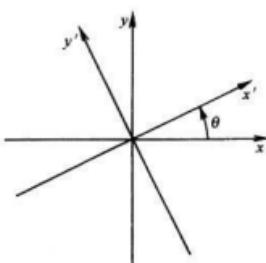


图 11-2 与光子的动量
矢成直角的坐标

现在我们问这一问题。假设一个光子通过角度为 θ 的偏振片后得到沿着 x' 方向振动的偏振光，而后到达一块角度为零的偏振片——如图 11-3 所示，那会发生什么情况呢？它将有多大的概率通过？答案如下：当它通过第一片偏振片后，它一定处在 $|x'\rangle$ 态，第二片偏振片只让 $|x\rangle$ 态的光子通过（吸收 $|y\rangle$ 态的光子）。所以我们要问，出现处在 $|x\rangle$ 态光子的概率有多大？我们可以从 $|x'\rangle$ 态的光子在 $|x\rangle$ 态的振幅 $\langle x|x'\rangle$ 的绝对值的平方得到此概率。 $\langle x|x'\rangle$ 是什么？只要把式(11.33)左乘以 $\langle x|$ 就得到

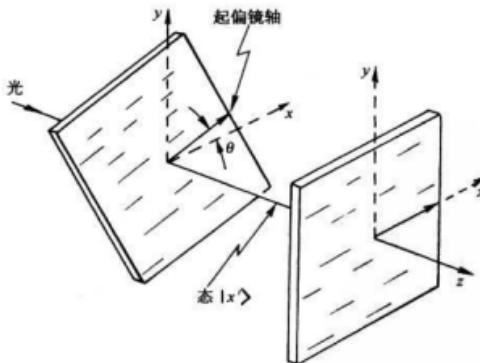


图 11-3 透振方向间成 θ 角的两个偏振片

$$\langle x | x' \rangle = \cos \theta \langle x | x \rangle + \sin \theta \langle x | y \rangle.$$

而从物理意义得知, $\langle x | y \rangle = 0$ ——如果 $|x\rangle$ 和 $|y\rangle$ 是基础态的话就必须如此——而 $\langle x | x \rangle = 1$ 。所以得到

$$\langle x | x' \rangle = \cos \theta,$$

概率就是 $\cos^2 \theta$ 。举例来说, 如果第一片偏振片角度为 30° , 那么一个光子有 $3/4$ 的机会得以穿过, 而有 $1/4$ 的机会光子被偏振片所吸收而将它加热。

现在我们来看一下在同样情况下按经典理论会发生什么。我们有一束光, 它的电场以不定的方式变化着——我们就说这束光是“非偏振”的。当它通过第一片偏振片后, 电场沿 x' 方向振动, 而大小为 ϵ , 我们可以在如图 11-4 那样的图上把电场画成峰值为 ϵ_0 的振动矢量。当光到达第二片偏振片时, 只有电场的 x 分量 $\epsilon_0 \cos \theta$ 得以通过。光强与电场的平方成正比, 因此与 $\epsilon_0^2 \cos^2 \theta$ 成正比。所以从第二片偏振片出射时能量比进入时的能量减弱一个 $\cos^2 \theta$ 因子。

经典图像和量子图像给出了相同的结果。如果你往第二片偏振片注射 100 亿个光子, 每个光子通过的平均概率比方说是 $3/4$, 那你就可以指望 100 亿光子中有 $3/4$ 得以通过。类似地, 它们所携带的能量也只是你注入能量的 $3/4$ 。

经典理论不说事件的统计意义——它只是说通过的能量正好是你送入的能量的 $3/4$ 。当然, 如果只有一个光子的话, 这就不可能了, 不存在 $3/4$ 个光子那样的东西。它要么整个在那里, 要么根本不在那里。量子力学则告诉我们, 有 $3/4$ 的机会它整个在那里。两种理论的关系是清楚的。

那么, 对另一类偏振呢? 比方说右旋圆偏振又如何? 在经典理论中, 右旋圆偏振有着大小相等而相位差为 90° 的 x 振动分量及 y 振动分量。在量子理论中, 右旋圆偏振(RHC)光子在偏振态 $|x\rangle$ 及 $|y\rangle$ 有相等的振幅, 而振幅间相位差为 90° 。称 RHC 光子为 $|R\rangle$ 态, 左旋圆偏振(LHC)光子为 $|L\rangle$ 态, 我们可以写出(参见第 1 卷 § 33-1)

$$\begin{aligned} |R\rangle &= \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle), \\ |L\rangle &= \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle). \end{aligned} \tag{11.34}$$

因子 $1/\sqrt{2}$ 是为了使态归一化而引进的。用这些态并利用量子理论的定律, 你可以计算任何你想研究的滤光或干涉效应。如果愿意你也可以选择 $|R\rangle$ 与 $|L\rangle$ 为基础态, 而用它们表示任何状态。只要先证明 $\langle R | L \rangle = 0$ ——取上式中第一个方程的共轭式[参见式(8.13)], 然后乘以第二个方程, 就可证明这一点。你可以将光分解为 x 向振动和 y 向振动, 或分解为 x' 向振动和 y' 向振动, 也可以分解为右旋圆偏振及左旋圆偏振等作为基础态。

作为一个例子, 我们来试一下把公式掉转过来。能不能把态 $|x\rangle$ 表示为右旋态与左旋态的线性组合呢? 可以, 这就是

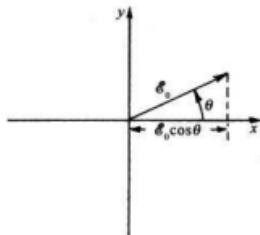


图 11-4 电矢量 ϵ 的经典图像

$$\begin{aligned} |\mathbf{x}\rangle &= \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle), \\ |\mathbf{y}\rangle &= -\frac{i}{\sqrt{2}}(|R\rangle - |L\rangle). \end{aligned} \quad (11.35)$$

证明：将式(11.34)中的两式相加与相减即可。很容易从一种基础态变到另一种基础态。

然而，必定会出现一个奇特的情况。如果光子是右旋圆偏振的，它不应跟 x 及 y 轴有任何关系。如果我们从一个绕光子飞行方向转过某一角度的坐标系来看同一个现象的话，光应当仍然是右旋圆偏振的，对左旋圆偏振光，情况也一样。右旋和左旋圆偏振光对任何转动都不变，它们的定义与 x 方向怎么选择无关（只是光子的方向已给定）。定义左、右旋不需要定任何坐标轴，这岂非好事，这要比取定 x 和 y 方便多了。但另一方面，当你把右旋与左旋加在一起时，可以求得 x 方向，这不很奇怪吗？如果“右”旋和“左”旋丝毫不依赖于 x ，那我们把它们重新放到一起时又怎么能得到 x ？在某种程度上我们可以这样来回答上述问题：在 x' ， y' 坐标系中写出表示 RHC 偏振光子的态矢量 $|R'\rangle$ ，在这个坐标系中，你可以写出：

$$|R'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle + i|y'\rangle).$$

这样一个态在 x ， y 坐标系中又显得怎样呢？只要用式(11.33)代 $|x'\rangle$ 并用相应的表达式代 $|y'\rangle$ ——我们未曾将它写出来，但它就是 $(-\sin\theta)|x\rangle + (\cos\theta)|y\rangle$ 。于是就有

$$\begin{aligned} |R'\rangle &= \frac{1}{\sqrt{2}}[\cos\theta|x\rangle + \sin\theta|y\rangle - i\sin\theta|x\rangle + i\cos\theta|y\rangle] \\ &= \frac{1}{\sqrt{2}}[(\cos\theta - i\sin\theta)|x\rangle + i(\cos\theta - i\sin\theta)|y\rangle] \\ &= \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle)(\cos\theta - i\sin\theta). \end{aligned}$$

第一项正是 $|R\rangle$ ，而第二项是 $e^{-\#}$ ；我们的结果就是

$$|R'\rangle = e^{-\#}|R\rangle. \quad (11.36)$$

态 $|R'\rangle$ 与 $|R\rangle$ 除了一个相位因子 $e^{-\#}$ 外完全相同。如果你对 $|L'\rangle$ 作同样的计算，就会得到*

$$|L'\rangle = e^{+\#}|L\rangle. \quad (11.37)$$

现在你看到发生了什么情况。如果我们将 $|R\rangle$ 与 $|L\rangle$ 相加，就得到跟 $|R'\rangle$ 与 $|L'\rangle$ 相加不同的结果。譬如说，一个 x 偏振光子是[式(11.35)] $|R\rangle$ 和 $|L\rangle$ 的和，但一个 y 偏振光子则是把 $|R\rangle$ 的相位后移 90° ，把 $|L\rangle$ 的相位前移 90° 后的两者之和。这正是在特殊的角度 $\theta = 90^\circ$ 的情况下对 $|R'\rangle$ 与 $|L'\rangle$ 求和所得的结果。这是正确的。在带“撇”的参考系中的 x 偏振跟

* 这类似于我们在(第6章)对自旋 $1/2$ 粒子所得到的结论。当我们把坐标轴绕 z 轴转动，就得到相位因子 $e^{\pm\#\pi/2}$ 。事实上，这正是我们在 § 5-7 对自旋 1 粒子的 $|+\rangle$ 态与 $|-\rangle$ 态所写下的结果——这并非巧合。光子是个自旋为 1 的粒子，但它没有“零”态。

原来参考系的 y 偏振相同，所以，认为圆偏振光子在任何坐标系中看来都相同这种说法不完全正确。它的相位(右旋与左旋圆偏振态的相位关系)始终保持着 x 方向的关系。

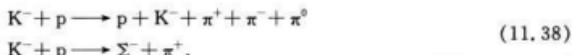
§ 11-5 中性 K 介子*

我们现在要来描写奇异粒子世界中的一个双态系统——一个量子力学对之作出了最令人惊异的预言的系统。要完全地描述这一系统，我们就得涉及到许多有关奇异粒子的知识，所以，很遗憾，我们将不得不说得简短些。我们只能对某个发现的获得过程作出大致的描绘——把所包括的推理方式告诉你们。这是从盖尔曼(Gell-Mann)和西岛(Nishijima)发现奇异性概念以及新的奇异性守恒定律开始的。正是当盖尔曼和佩斯(Pais)分析这些新概念所得的结果时，他们提出了我们正要描写的一个最令人惊异的现象的预言。然而，首先我们需要谈一点“奇异性”。

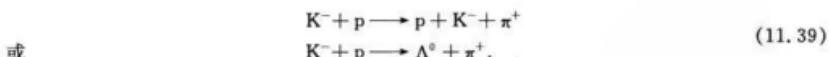
我们必须从核粒子之间的所谓强相互作用开始。这种作用是强核力相互作用的根源，它和相对较弱的电磁相互作用不同。作用“强”指的是，如果两个粒子靠得足够近而产生相互作用，它们的作用非常强烈，并且很容易产生其他粒子。核粒子之间也存在着所谓的“弱相互作用”，由此也会发生某些现象，诸如 β 衰变，但是用核的时间尺度来衡量，发生作用总是非常缓慢——弱作用的大小要比强作用弱许多数量级，甚至比电磁相互作用还弱得多。

当人们用大型加速器对强相互作用进行研究时，惊奇地发现有些“应当”发生的事——预期会发生的事——却没有发生。譬如，在某些相互作用中，某种类型的粒子没有像原先预料的那样出现。盖尔曼和西岛注意到，如果提出一条新的守恒定律——奇异性守恒，这许多奇异现象就能马上得到解释。他们提出，每个粒子都具有一种新的属性——他们称之为“奇异”数，而在任何强相互作用过程中，“奇异性的数量”是守恒的。

举例来说，假定一个高能负 K 介子(譬如说带有几十亿电子伏特能量)与一个质子相碰撞。这样的相互作用会产生许多别的粒子： π 介子，K 介子， Λ 粒子， Σ 粒子，即列于第 1 卷表 2-2 中的各种介子或重子。人们观察到只出现某些组合，另外一些组合则从不出现。已经知道某些守恒定律在这里是起作用的。首先，能量和动量总是守恒的，事件后的总能量和总动量必定与事件前的相同。其次，有电荷守恒，这就是说，所产生的粒子的总电荷必定等于原有粒子所带的总电荷。在我们这个 K 介子与质子相碰撞的例子中，以下的反应确实发生了：

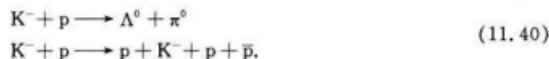


由于电荷守恒，我们永远不会得到

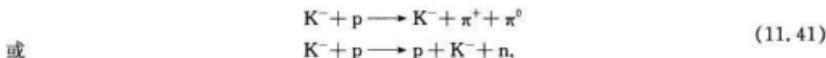


* 我们现在感到这节的材料对眼下的课程来说是过于冗长和困难了。我们建议你们跳过这节继续阅读 § 11-6 节。如果你们有雄心又有时间的话，以后也许要再回到这一节来。我们把这一节留在这里，因为这是取自高能物理新近成果的一个美妙例子，说明用我们关于双态系统的量子力学公式能够做出些什么结果来。

我们也知道，重子数是守恒的。反应后的重子数必定等于反应前的重子数。在这条定律中，重子的反粒子作为一个负重子来计算。这意味着我们可以——也确实——见到这样的反应：

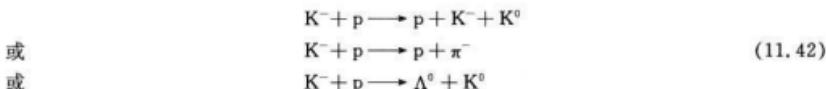


(这里 \bar{p} 是反质子，它带一个负电荷。)但是我们从没有见到



(即使能量很大也是如此)，因为上式中重子数不守恒。

然而，这些定律不能解释这样一件奇特的事实，即下列反应：



也从未观察到过，而这些反应初看起来跟式(11.38)或式(11.40)的某些反应并无多大的差别。对此的解释是奇异性守恒。每个粒子都有一个数，即奇异数 S ，而在任何强相互作用中，都存在着这样一条定律：反应后的总奇异数必定等于反应前的总奇异数。质子和反质子(p, \bar{p})，中子和反中子(n, \bar{n})，以及 π 介子(π^+, π^0, π^-)的奇异数全都为0， K^+ 与 \bar{K}^0 (K^0 的反粒子)， Λ^0 与 Σ 粒子(+, 0, -)奇异数为+1。 K^- 与 \bar{K}^0 (K^0 的反粒子)， Λ^0 与 Σ 粒子(+, 0, -)奇异数为-1。还有一种粒子 Ξ 粒子(“ ξ ”的大写)的奇异数为-2，或许还有其他奇异数尚不知道的粒子。我们已把这些奇异数列在表 11-4 中。

表 11-4 强相互作用粒子的奇异数

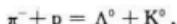
	S			
	-2	-1	0	+1
重子	Ξ^0	Σ^+	p	
	Ξ^-	Λ^0, Σ^0	n	
		Σ^-		
介子			π^+	K^+
		\bar{K}^0	π^0	K^0
		K^-	π^-	

注： π^- 为 π^+ 的反粒子(反之亦然)。

让我们看看在已写出的某些反应中，奇异性守恒是怎么起作用的。如果开始时是一个 K^- 和一个质子，总的奇异数为 $(-1 + 0) = -1$ 。奇异性守恒要求反应后产物的奇异数的总和也必须是-1。你们看到式(11.38)及(11.40)满足这一要求。但在式(11.42)的反应中，右边的奇异性在三种情况中都是0。这些反应前后奇异性不保持守恒，因而不会发生。为什么？没人知道，没人知道得比我们刚才告诉你的有关情况更多一些。自然界正是以这种方式运行着。

现在我们来看一看下列反应：一个 π^- 撞击一个质子。你也许(比方说)得到一个 Λ^0 粒子加一个中性 K 粒子——两个中性粒子。但会得到哪一种中性 K ？因为 Δ 粒子有奇异数-1，而 π 与 p^+ 的奇异数为0，并且由于这是个快速产生的反应，奇异性必定不变。 K 粒子

一定要有奇异数+1，所以它必须是 K^0 。这反应是



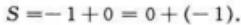
而

$$S = 0 + 0 = (-1) + (+1) \text{ (守恒).}$$

如果在这里用 \bar{K}^0 代替 K^0 ，右边的奇异数就会是 -2，而这是自然界所不允许的，因为左边的奇异数为 0。不过， \bar{K}^0 可以在别的反应中产生，诸如



或



你们或许会想，“这完全是些废话，因为你怎么知道它是 \bar{K}^0 还是 K^0 ？它们看起来完全相同。它们互为反粒子，它们有完全相等的质量，而电荷又都是 0。你们怎么来区分它们？”回答是通过它们所发生的反应。例如，一个 \bar{K}^0 可以跟物质作用产生一个 Δ 粒子，如



但 K^0 却不行。当 K^0 与普通的物质（质子和中子）相互作用时，不会使它产生 Δ 粒子*。所以 K^0 和 \bar{K}^0 之间的实验鉴别就是它们之中有一个能产生 Δ 粒子，另一个却不能。

这样，奇异性理论的预言之一就是，如果在一个高能 π 介子的实验中，产生了一个 Δ 粒子与一个中性 K 介子，然后这个中性 K 介子撞到另一些物质时决不会产生 Δ 粒子。实验可以这样来进行，把一束 π^- 介子射入一个巨大的氢气泡室， π^- 的径迹消失了，但另外某个地方出现了一对径迹（一个质子和一个 π^- ），这表明一个 Δ 粒子衰变了**（见图 11-5）。这样你就知道在某个地方有个你不能看到的 K^0 。

但你可以利用动量和能量守恒求出它往哪里去。[它后来会衰变为两个带电粒子而显现出来，如图 11-5(a)所示。]当 K^0 向前飞行时，可能跟一个氢核（质子）发生相互作用，也许产生一些别的粒子。而奇异性理论的预言是，它绝不会在下述类型的简单反应中产生一个 Δ 粒子：



而 \bar{K}^0 是可以发生这种反应的。这就是说，在气泡室中，一个 \bar{K}^0 可以产生如图 11-5(b)所描画的事件——其中的 Δ^0 由于衰变而被看到，但 K^0 却不能产生这种过程。这就是我们故事的第一部分，这就是奇异性守恒。

不过，奇异性守恒不是完美无缺的。奇异粒子可以发生极缓慢的衰变——衰变时间长到约 10^{-10} s***，这时奇异性不再守恒，这些衰变称为“弱”衰变。例如， K^0 衰变为一对寿命

* 当然，除非它还产生两个 K^+ 或总奇异数为 +2 的其他一些粒子。我们可以认为在上述反应中没有足够能量来产生这些更多的奇异粒子。

** 自由 Δ 粒子经过一弱相互作用而缓慢衰变（故奇异性不必守恒）。衰变产物是一个 p 及一个 π^- ，或一个 n 及一个 π^0 。寿命为 2.2×10^{-10} s。

*** 强相互作用的典型时间差不多是 10^{-23} s。

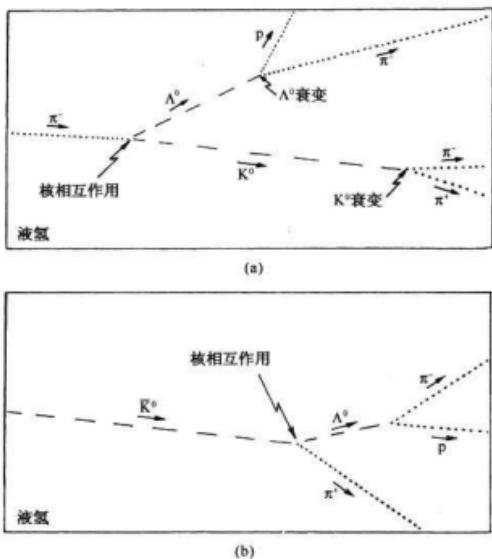


图 11-5 在氢气泡室观察到的高能事件:(a)一个 π^- 介子与一个氢核(质子)相互作用产生一个 Λ^0 粒子和一个 K^0 介子。两个粒子都在气泡室中衰变。(b)一个 \bar{K}^0 介子和一个质子相互作用产生一个 π^+ 介子和一个 Λ^0 粒子,然后 Λ^0 衰变(中性粒子没有留下径迹,它们推断出的轨迹如图中断续线所示)

为 10^{-19} s 的 π 介子(+及-)。事实上,这是首次见到 K 粒子的方法。注意衰变反应



中奇异性并不守恒,所以它不可能通过强相互作用而“很快”发生,只可能通过弱衰变过程进行。

注意 \bar{K}^0 也以同样方式衰变为一个 π^+ 及一个 π^- 而且也有同样的寿命,



这也是一个弱衰变,因为它奇异性并不保持守恒。有一条原理:对任何反应,总对应地存在一个用“反物质”代替“物质”的反应,反之亦然。由于 \bar{K}^0 是 K^0 的反粒子,它应当衰变为 π^+ 与 π^- 的反粒子,但 π^+ 的反粒子是 π^- 。(或者,如果你乐意,反过来也可以。结果发现对 π 介子来说,你们称哪一个为“物质”都无所谓。)所以作为弱衰变的一个结果, K^0 和 \bar{K}^0 可以变为同样的最终产物。当我们通过衰变来“见到”它们时——如同气泡室中那样——它们看上去是同一种粒子。只是它们的强相互作用不同。

我们终于可以来描述盖尔曼和佩斯的工作了。他们首先注意到,既然 K^0 及 \bar{K}^0 都能转变成两个 π 介子的状态,必定存在某个 K^0 转变为 \bar{K}^0 的振幅,并且也有一定 \bar{K}^0 转变为 K^0

的振幅。像在化学中所做那样写出反应方程式，有



这些反应意味着，每单位时间内， K^0 通过弱相互作用衰变成两个 π 介子再转变为 \bar{K}^0 有一定的振幅，这振幅写成 $-i/\hbar$ 乘以 $\langle \bar{K}^0 | W | K^0 \rangle$ 。对于逆反应来说，也有相应的振幅 $\langle K^0 | W | \bar{K}^0 \rangle$ 。由于物质与反物质的行为完全相同，这两个振幅在数值上相等，我们把它们都称为 A ，

$$\langle \bar{K}^0 | W | K^0 \rangle = \langle K^0 | W | \bar{K}^0 \rangle = A. \quad (11.44)$$

盖尔曼和佩斯说，这里出现一个有趣的情况。人们一向所说的世界的两个不同的状态—— K^0 与 \bar{K}^0 ——实际上应当被看作为是一个双态系统，因为从一个态到另一个态有一定振幅。当然，做完全的处理时，必须处理两个以上的态，因为也还存在着一些像两个 π 介子的态，等等；但是既然他们主要对 K^0 与 \bar{K}^0 的关系感兴趣，就不必把事情搞得太复杂，而可以作双态系统的近似。其他态在某种程度上已经考虑到，它们的效应已隐含在式(11.44)的振幅中。

因而盖尔曼和佩斯把这个中性粒子作为一个双态系统来进行分析。他们首先把 $|K^0\rangle$ 和 $|\bar{K}^0\rangle$ 态取为两个基础态。(这样，事情非常像氦分子的情况。)这样，中性 K 粒子的任何态 $|\psi\rangle$ 就可用它处在两个基础态的振幅来描写。我们将称这两个振幅为

$$C_+ = \langle K^0 | \psi \rangle, C_- = \langle \bar{K}^0 | \psi \rangle. \quad (11.45)$$

下一步就是要写下这个双态系统的哈密顿方程。如果 K^0 与 \bar{K}^0 之间没有耦合，方程就只是

$$\begin{aligned} i\hbar \frac{dC_+}{dt} &= E_0 C_+, \\ i\hbar \frac{dC_-}{dt} &= E_0 C_-. \end{aligned} \quad (11.46)$$

但由于有 \bar{K}^0 转变为 K^0 的振幅 $\langle K^0 | W | \bar{K}^0 \rangle$ 存在，所以应当还有一项

$$\langle K^0 | W | \bar{K}^0 \rangle C_- = AC_-$$

加在第一个方程的右边。类似地， AC_+ 项应该加到 C_- 随时间变化的方程中去。

但还不止这些。当考虑到双 π 介子效应时，还存在 K^0 通过以下过程变为自身的附加振幅，



我们将用 $\langle K^0 | W | K^0 \rangle$ 来表示的这个附加振幅，它正好等于振幅 $\langle \bar{K}^0 | W | K^0 \rangle$ ，因为对于 K^0 和 \bar{K}^0 来说，它们变为一对 π 介子与由一对 π 介子变回自身的振幅是完全相同的。如果需要，可以详细写出这个论证。首先，我们写下*

$$\text{和 } \langle \bar{K}^0 | W | K^0 \rangle = \langle \bar{K}^0 | W | 2\pi \rangle \langle 2\pi | W | K^0 \rangle$$

$$\langle K^0 | W | K^0 \rangle = \langle K^0 | W | 2\pi \rangle \langle 2\pi | W | K^0 \rangle.$$

由于物质和反物质的对称性，

* 这里我们作了点简化。 2π 系统可以具有对应于不同的 π 介子动量的许多状态，所以我们应当将方程的右边写成对 π 介子的不同基态求和。完整的处理仍导致同样的结论。

以及

$$\begin{aligned}\langle 2\pi | W | K^0 \rangle &= \langle 2\pi | W | \bar{K}^0 \rangle, \\ \langle K^0 | W | 2\pi \rangle &= \langle \bar{K}^0 | W | 2\pi \rangle.\end{aligned}$$

由此得到 $\langle K^0 | W | K^0 \rangle = \langle \bar{K}^0 | W | K^0 \rangle$, 以及 $\langle \bar{K}^0 | W | K^0 \rangle = \langle K^0 | W | \bar{K}^0 \rangle$, 这就是我们前面所讲过的。不管怎么说, 两个都等于 A 的附加振幅 $\langle K^0 | W | K^0 \rangle$ 和 $\langle \bar{K}^0 | W | \bar{K}^0 \rangle$, 应当包括在哈密顿方程中。第一个给出的一项 AC_+ 加在 dC_+/dt 的方程的右边, 第二个给出的是加在 dC_-/dt 的方程右边新的一项 AC_- 。通过这样论证, 盖尔曼与佩斯推断 $K^0\bar{K}^0$ 系统的哈密顿方程应当是

$$\begin{aligned}i\hbar \frac{dC_+}{dt} &= E_0 C_+ + AC_- + AC_+, \\ i\hbar \frac{dC_-}{dt} &= E_0 C_- + AC_+ + AC_-. \quad (11.47)\end{aligned}$$

现在我们必须对前几章中所说的作一些纠正: 像 $\langle K^0 | W | \bar{K}^0 \rangle$ 和 $\langle \bar{K}^0 | W | K^0 \rangle$ 这样两个互相逆转的振幅总是互为复共轭。这只有在讨论中不发生衰变的粒子时才成立。而如果粒子可以衰变——因此, 也就可能“消失”——那么两个振幅不一定是复共轭。所以等式(11.44)并不意味着振幅是实数, 事实上, 它们是复数。所以, 系数 A 是复数, 这样我们不能把它归到能量 E_0 中去。

在经常跟电子自旋以及诸如此类的事情打交道后, 我们的主人公看出哈密顿方程组(11.47)意味着存在另外一对基础态, 它们也可用来表示 K 粒子系统, 而且具有特别简单的性质。他们说: “让我们将两个方程相加与相减, 并且以 E_0 为零点测量所有有关能量, 并使用使 $\hbar = 1$ 的能量与时间的单位。”(这是近代理论物理学家常做的事, 这样并不改变物理实质, 但使方程的形式变得简单些。)他们的结果是:

$$i \frac{d}{dt}(C_+ + C_-) = 2A(C_+ + C_-), \quad i \frac{d}{dt}(C_+ - C_-) = 0. \quad (11.48)$$

显然, 振幅的组合 $(C_+ + C_-)$ 与 $(C_+ - C_-)$ 互相独立地起着作用。(显然, 它们相当于我们早就研究过的定态。)所以, 他们断定, 用 K 粒子的另一种不同的表示会更为方便。他们定义两个态

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle), \quad |K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle). \quad (11.49)$$

他们说, 用 K_1 与 K_2 两个“粒子”(就是“态”)来代替 K^0 和 \bar{K}^0 两个介子也同样可以。(当然, 这两个态对应着我们通常称为 |I⟩ 与 |II⟩ 的态, 我们不用老的记号是因为我们现在要用原作者的记号——你们将在物理讨论会上见到的记号。)

但盖尔曼和佩斯做所有这些并不只是为了给粒子起个不同的名称, 这里还存在某种令人惊奇的新物理内容。假设 C_1 及 C_2 是某个态 $|\psi\rangle$ 表现为 K_1 或 K_2 介子的振幅:

$$C_1 = \langle K_1 | \psi \rangle, \quad C_2 = \langle K_2 | \psi \rangle.$$

由式(11.49),

$$C_1 = \frac{1}{\sqrt{2}}(C_+ + C_-), \quad C_2 = \frac{1}{\sqrt{2}}(C_+ - C_-). \quad (11.50)$$

于是式(11.48)变为

$$i \frac{dC_1}{dt} = 2AC_1, i \frac{dC_2}{dt} = 0, \quad (11.51)$$

其解为

$$C_1(t) = C_1(0)e^{-iAt}, C_2(t) = C_2(0). \quad (11.52)$$

显然, 这里 $C_1(0)$ 和 $C_2(0)$ 为 $t = 0$ 时的振幅。

这些式子表明如果一个中性 K 粒子在时刻 $t = 0$ 时从态 $|K_1\rangle$ [这样 $C_1(0) = 1$ 而 $C_2(0) = 0$] 出发, 那么, 在时刻 t 的振幅是

$$C_1(t) = e^{-iAt}, C_2(t) = 0.$$

注意到 A 是复数, 不妨取 $2A = \alpha - i\beta$, (由于发现 $2A$ 的虚部是个负值, 我们把它写成负 $i\beta$)。代入 $C_1(t)$ 后, 可以写成

$$C_1(t) = C_1(0)e^{-\beta t}e^{-i\alpha t}. \quad (11.53)$$

在时刻 t 找到 K_1 粒子的概率是这个振幅的绝对值平方, 即 $e^{-2\beta t}$ 。而由式(11.52), 在任何时候找到 K_2 态的概率为零。这就是说, 如果你让一个 K 粒子起先处在 $|K_1\rangle$ 态, 在同一态中找到它的概率将随时间指数下降——但你永远不会在态 $|K_2\rangle$ 中找到它。那么它跑到哪里去了呢? 它以平均寿命 $\tau = 1/2\beta$ 衰变为两个 π 介子了, 实验测得 τ 为 10^{-10} s。我们前面说 A 是复数时就是为此作了准备。

另一方面, 式(11.52)说明, 如果使一个 K 粒子完全处在 K_2 态, 它就将永远这样待下去。当然, 这实际上是不对的。实验上观察到这个 K 粒子衰变为 3 个 π 介子, 不过其速率要比刚才描写过的双 π 介子衰变慢 600 倍。可见, 在我们的近似中扔掉了其他一些小项。但只要我们考虑的只是双 π 介子衰变, K_2 就将“永远”存在下去。

现在继续讲完盖尔曼和佩斯的故事。他们继续考虑在强相互作用下产生一个 K 粒子与一个 Λ^0 粒子时会发生什么事。由于 K 粒子必须具有奇异数 +1, 所以它产生时必须处在 K^0 态。这样, 在 $t = 0$ 时, 它既不是 K_1 也不是 K_2 , 而是一个混合态。初始条件是

$$C_+(0) = 1, C_-(0) = 0.$$

但是根据式(11.50), 这意味着

$$C_1(0) = \frac{1}{\sqrt{2}}, C_2(0) = \frac{1}{\sqrt{2}},$$

而由式(11.52)和(11.53), 有

$$C_1(t) = \frac{1}{\sqrt{2}}e^{-\beta t}e^{-i\alpha t}, C_2(t) = \frac{1}{\sqrt{2}}. \quad (11.54)$$

注意 K^0 和 \bar{K}^0 都是 K_1 和 K_2 的线性组合。在式(11.54)中, 已选定振幅, 使 $t = 0$ 时 \bar{K}^0 部分由于干涉而相互抵消, 只留下 K^0 态。但 $|K_1\rangle$ 态随时间而变, 而 $|K_2\rangle$ 态则不随时间而变。在 $t = 0$ 以后, C_1 与 C_2 的干涉将使 K^0 与 \bar{K}^0 都具有一定的振幅。

所有这些意味着什么? 让我们回过去想一下如图 11-5 所示的一个实验。一个 π^- 介子产生一个 Λ^0 粒子和一个 K^0 介子, 这个 K^0 介子嘟嘟地穿过气泡室中的氢。当它向前跑的

时候,它有某种很小的但一定的机会与一个氢核碰撞。起先我们认为奇异性守恒将阻止 K 粒子在这样的相互作用中产生一个 Λ^0 粒子。然而,现在我们看到这是不对的。虽然 K 粒子开始时为 K^0 ——它不可能产生一个 Λ^0 ——但它不会永远这样。过一会儿,就有一一定的振幅使它跳变到 \bar{K}^0 态。所以我们可以预期沿着 K 粒子的踪迹有时会看到产生一个 Λ^0 。这件事发生的机会由振幅 C_- 给定,而 C_- 可以与 C_1 和 C_2 联系起来[倒过来利用式(11.50)]。它们的关系是

$$C_- = \frac{1}{\sqrt{2}}(C_1 - C_2) = \frac{1}{2}(e^{-\beta t} e^{-i\omega t} - 1). \quad (11.55)$$

当 K 粒子向前飞行时,它“出现像 \bar{K}^0 那样行为”的概率等于 $|C_-|^2$,其值为

$$|C_-|^2 = \frac{1}{4}(1 + e^{-2\beta t} - 2e^{-\beta t} \cos \alpha t). \quad (11.56)$$

真是个复杂而奇特的结果!

于是,盖尔曼和佩斯就作了这样一个惊人的预言:当一个 K^0 产生后,它转变为 \bar{K}^0 ——这可由能够产生一个 Λ^0 来显示——的概率按式(11.56)随时间变化。得出这个预言所用的只是纯粹的逻辑推理和量子力学的基本原理——而根本不知道 K 粒子的内部机理。由于没有人知道任何有关内部机理的情况,盖尔曼和佩斯最多也只能走到这一步。他们不能给出 α 与 β 的任何理论值。直到目前也没有能给出这两个值。他们由实验观察到的衰变为两个 π 介子的速率能得出一个 β 值 ($2\beta = 10^{10} \text{ s}^{-1}$),但关于 α ,他们什么也说不出。

在图 11-6 中我们对于两个 α 值画出了式(11.56)的函数的曲线图。你们可以看到图形和 α 与 β 的比值有很大关系。起先 \bar{K}^0 的概率为 0,然后它增大。如果 α 大,概率将具有大的振荡。如果 α 小,振荡就小,或者没有振荡——概率只平滑地上升至 $1/4$ 。

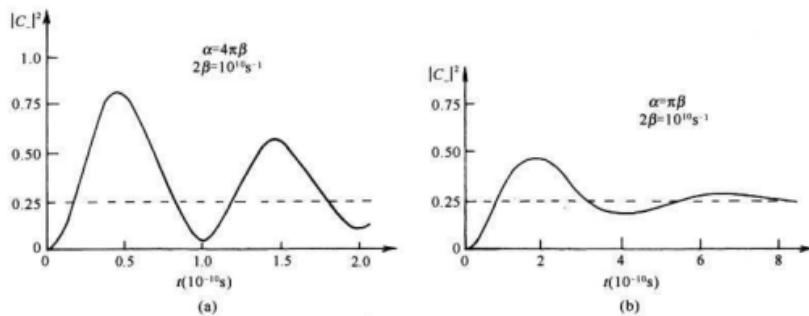


图 11-6 式(11.56)的函数图像:(a)设 $\alpha = 4\pi\beta$, (b) 设 $\alpha = \pi\beta$ ($2\beta = 10^{10} \text{ s}^{-1}$)

通常,K 粒子以接近于光速的恒定速度飞行。图 11-6 的曲线因而也表示沿着轨迹观察到 \bar{K}^0 的概率,其典型距离为几个厘米。你们可以看出为什么这个预言这样离奇。在产生一个粒子后,它不仅会衰变,也会发生别的事情。有时它衰变,有时它又转变为另一种粒子。它产生某种效应的特征概率随着它的行进而以一种奇特的方式变化着。在自然界里没有别

的东西与它相像。而作出这个极为令人惊奇的预言所依据的只是关于振幅干涉的论证。

如果有一个地方，我们有机会以最纯粹的方式检验量子力学的主要原理——振幅的叠加原理是否成立？——这个地方就是这里。尽管这个效应的预言至今已有好几年了，还没有非常清楚的实验判决。只有一些粗略的结果表明 α 不是零，不过这种效应确实出现了，这些结果表明 α 在 2β 和 4β 之间。这些都是实验得到的。如能精确地检验一下曲线，看一看在奇异粒子这样的神秘世界中，叠加原理是否确实仍然成立，这将是件非常美妙的事情——对奇异粒子来说，我们不知它们为什么衰变，也不知它们为什么有奇异性。

我们刚才所做的分析方法表明了目前在探索解释奇异粒子的研究中所用的量子力学方法的特点。所有你们可能听到的有关的复杂理论无非都是这类利用叠加原理和同一水平的其他量子力学原理来玩的骗人把戏而已。有些人声称他们已拥有能够计算 α 和 β ，或者至少在给定 β 时能算出 α 的理论，但这些理论都是完全无用的。比方说，有个在给出 β 后能预言 α 值的理论告诉我们 α 值应是无穷大。他们开始所用的方程组包括两个 π 介子，而后又从两个 π 介子回到了一个 K^0 ，等等。在计算全部完成后，确实得到一对类似于我们这里所写的方程，但是由于两个 π 介子的动量有无穷多个状态，对所有这些可能性求积分后得出 α 值是无穷大。但是，自然界的 α 并不是无穷大。所以这个动力学理论是错的。在奇异粒子世界中完全能被预言的现象竟能从你们现在所学到的水平的量子力学原理得出，这实在是非常惊奇的。

§ 11-6 对 N 态系统的推广

我们已经讲完了所有想谈的关于双态系统的问题。在下面几章里我们将继续研究具有更多状态的系统。把我们就双态所得到的概念推广到 N 态系统去是直截了当的事，可以按下面这样方式来做。

如果一个系统有 N 个不同的态，我们可把任何态 $|\psi(t)\rangle$ 表示成任何一组基础态 $|i\rangle$ ($i=1, 2, \dots, N$) 的线性组合，

$$|\psi(t)\rangle = \sum_{\text{所有} i} |i\rangle C_i(t). \quad (11.57)$$

系数 $C_i(t)$ 就是振幅 $\langle i|\psi(t)\rangle$ 。振幅 C_i 随时间的变化方式由下列方程组决定

$$i\hbar \frac{dC_i(t)}{dt} = \sum_j H_{ij} C_j, \quad (11.58)$$

这里能量矩阵 H_{ij} 描写问题的物理条件。这看上去和双态系统相同。只是现在 i 和 j 都必须遍及所有 N 个基础态，而能量矩阵 H_{ij} ——或者，你愿意称之为哈密顿也可——是带有 N^2 个数的 $N \times N$ 矩阵。像以前一样， $H_{ij}^* = H_{ji}$ ——只要粒子数守恒——而对角元 H_{ii} 是实数。

当能量矩阵为常数(不依赖于 t)时，我们已经求得双态系统的系数 C 的一般解。当 H 与时间无关时，也不难解出 N 态系统的方程组(11.58)。我们仍从寻找振幅全都有相同时间依赖性的可能解开始。试解

$$C_i = \alpha_i e^{-(i/\hbar)E_i t}, \quad (11.59)$$

在把这些 C_i 代入式(11.58)后，导数 $dC_i(t)/dt$ 就是 $(-i/\hbar)EC_i$ 。消掉所有各项的公共指数

因子, 得到

$$Ea_i = \sum_j H_{ij} a_j, \quad (11.60)$$

这是有 N 个未知数 a_1, a_2, \dots, a_N 的 N 元线性代数方程组, 只有在碰巧时, 即所有 a_i 的系数行列式为零时, 才存在一个解。但是并不一定要用这种复杂的方法求解, 你们可以随便用什么方法来着手解方程组, 你们会发现只有对一定的 E 值才能解出它们。(注意 E 是我们方程组中唯一可调节的量。)

但如果你们想用规范的方法求解, 可以把式(11.60)写为

$$\sum_j (H_{ij} - \delta_{ij} E) a_j = 0. \quad (11.61)$$

接下来你们可以利用这条规则(如果你们知道的话), 即只有对那些 E 值满足下式的方程组才有解,

$$\text{Det}(H_{ij} - \delta_{ij} E) = 0. \quad (11.62)$$

行列式的每一项正好是 H_{ij} , 只是对每个对角元要减去 E 。式(11.62)就是

$$\text{Det} \begin{bmatrix} H_{11} - E & H_{12} & H_{13} & \cdots \\ H_{21} & H_{22} - E & H_{23} & \cdots \\ H_{31} & H_{32} & H_{33} - E & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{bmatrix} = 0, \quad (11.63)$$

当然, 这只是 E 的代数方程的一种特殊写法, 这个方程是行列式所有各项按一定方式乘积之和。这些乘积将给出 E 的所有幂次, 直到 E^N 。

所以我们就有一个等于 0 的 N 次多项式, 一般说来它有 N 个根。(但要记住其中有些可以是重根, 就是说两个或更多的根彼此相等。) 不妨称 N 个根为

$$E_I, E_{II}, E_{III}, \dots, E_n, \dots, E_N, \quad (11.64)$$

(我们将利用 n 表示第 n 个罗马数字, 这样 n 取值为 I, II, III, ..., N。) 其中有些能量可能相等, 比方说 $E_{II} = E_{III}$, 但我们仍用不同的名字称呼它们。

方程组(11.60)或(11.61)对每个 E 值有一个解。假如将任何一个 E 值(譬如 E_n)代入式(11.60)并求解 a_i , 就得到一组属于能量 E_n 的解。我们将称这组解为 $a_i(n)$ 的振幅。

在式(11.59)中代入这些 $a_i(n)$ 值, 我们就有了某确定能态中处在基础态 $|i\rangle$ 的振幅 $C_i(n)$ 。设 $|n\rangle$ 表示 $t=0$ 时这个确定能态的态矢量, 我们可写出

$$C_i(n) = \langle i | n \rangle e^{-(i/\hbar)E_n t},$$

其中

$$\langle i | n \rangle = a_i(n). \quad (11.65)$$

于是有完全确定能量的状态 $|\psi_n(t)\rangle$ 可以写成

$$|\psi_n(t)\rangle = \sum_i |i\rangle a_i(n) e^{-(i/\hbar)E_n t}$$

或

$$|\psi_n(t)\rangle = |\mathbf{n}\rangle e^{-(i/\hbar)E_n t}. \quad (11.66)$$

态矢量 $|\mathbf{n}\rangle$ 描述确定能量状态的位形,而时间相依性被提出在外。这样它们就是常矢量,如果愿意的话,它们可以用来作一组新的基。

每个态 $|\mathbf{n}\rangle$ 都有个性质,——你能很容易地予以证明——当它们被哈密顿算符 \hat{H} 作用后,就正好得到 E_n 乘以同一个态:

$$\hat{H} |\mathbf{n}\rangle = E_n |\mathbf{n}\rangle. \quad (11.67)$$

这样,能量 E_n 就是描写哈密顿算符 \hat{H} 的特征的一个数值。我们已经见到,一般说,一个哈密顿有几个特征能量。在数学家的术语中,这些数称为矩阵 H_n 的“特征值”。物理学家常称它们为 \hat{H} 的“本征值”。“Eigen”是德文的“特征”或“本征”。对 \hat{H} 的每个本征值——换句话说对每个能量——都存在着确定能量的状态,我们已称它们为“定态”。物理学家通常称 $|\mathbf{n}\rangle$ 为“ \hat{H} 的本征态”。每个本征态对应着一个特定的本征值 E_n 。

一般来说,态 $|\mathbf{n}\rangle$ (有N个)也可用来作为一组基础态。要做到这点,所有的态必须相互正交,意即对其中任何两个态, $|\mathbf{n}\rangle$ 和 $|\mathbf{m}\rangle$,有

$$\langle \mathbf{n} | \mathbf{m} \rangle = 0. \quad (11.68)$$

如果所有能量都互不相同,那么上式将会自动满足。我们也可以将所有的 $a_i(\mathbf{n})$ 乘以一个适当的因子,而使所有的态归一化。归一化后对所有的 \mathbf{n} 就有

$$\langle \mathbf{n} | \mathbf{n} \rangle = 1. \quad (11.69)$$

当式(11.63)恰巧有两个(或更多)具有相同能量的根时,问题就稍有点复杂。首先,存在两组不同的 a_i 具有相同的能量,但是它们所给出的态可能不正交。假设你们通过归一化手续找到了两个具有同样能量的定态——称它们为 $|\mu\rangle$ 及 $|\nu\rangle$ 态。不幸它们不正交,就是说

$$\langle \mu | \nu \rangle \neq 0.$$

但我们总可以造出两个新的态,我们称为 $|\mu'\rangle$ 和 $|\nu'\rangle$,它们具有相同的能量,而且也正交,即

$$\langle \mu' | \nu' \rangle = 0. \quad (11.70)$$

$|\mu\rangle$ 和 $|\nu\rangle$ 适当的线性组合可以作出 $|\mu'\rangle$ 和 $|\nu'\rangle$,并适当选择系数,使式(11.70)成立。做到这点总是不难的。我们一般假定这点已经做到,因而总认为我们的本征能态 $|\mathbf{n}\rangle$ 全都相互正交。

为提高兴趣,我们来证明当两个定态有不同的能量时它们的确正交。对具有能量 E_n 的态 $|\mathbf{n}\rangle$,我们有

$$\hat{H} |\mathbf{n}\rangle = E_n |\mathbf{n}\rangle. \quad (11.71)$$

这个算符方程实际上表示有个数字方程。补上没写出的部分,它的意义与下式相同

$$\sum_j \langle i | \hat{H} | j \rangle \langle j | \mathbf{n} \rangle = E_n \langle i | \mathbf{n} \rangle. \quad (11.72)$$

如果取这式的复共轭,就得

$$\sum_j \langle i | \hat{H} | j \rangle^* \langle j | \mathbf{n} \rangle^* = E_n^* \langle i | \mathbf{n} \rangle^*. \quad (11.73)$$

注意振幅的复共轭是逆振幅, 所以式(11.73)可以重新写为

$$\sum_j \langle n | j \rangle \langle j | \hat{H} | i \rangle = E_n^* \langle n | i \rangle. \quad (11.74)$$

因为这个方程对任何 i 成立, 它的“简短形式”是

$$\langle n | \hat{H} = E_n^* \langle n |, \quad (11.75)$$

这称为式(11.71)的伴随式。

现在很容易证明 E_n 是个实数。将式(11.71)乘 $\langle n |$ 就得到

$$\langle n | \hat{H} | n \rangle = E_n^*, \quad (11.76)$$

因为 $\langle n | n \rangle = 1$ 。然后我们将式(11.75)右乘 $|n\rangle$, 得

$$\langle n | \hat{H} | n \rangle = E_n. \quad (11.77)$$

比较式(11.76)和(11.77), 显然可见

$$E_n = E_n^*, \quad (11.78)$$

这意味着 E_n 是个实数。我们可以擦掉式(11.75)中 E 上的星号。

我们终于可以证明不同的能量状态是正交的了。设 $|n\rangle$ 和 $|m\rangle$ 是任意的两个有确定能量的基础态。对 $|m\rangle$ 态应用式(11.75), 并乘以 $|n\rangle$, 我们得到

$$\langle m | \hat{H} | n \rangle = E_n \langle m | n \rangle.$$

但如果用 $\langle m |$ 乘以式(11.71), 得到

$$\langle m | \hat{H} | n \rangle = E_n^* \langle m | n \rangle.$$

因为这两个等式左边相等, 所以右边也相等,

$$E_n \langle m | n \rangle = E_n^* \langle m | n \rangle. \quad (11.79)$$

如果 $E_n = E_n^*$, 上式没告诉我们什么。但如果两个态 $|m\rangle$ 和 $|n\rangle$ 的能量不同 ($E_m \neq E_n$), 式(11.79)就表明 $\langle m | n \rangle$ 必定为零, 这正是我们要证明的。只要 E_n 和 E_m 在数值上不同, 两个态必然正交。

第 12 章 氢的超精细分裂

§12-1 由两个自旋 $1/2$ 粒子组成的系统的基础态

在这一章中，我们来着手讨论氢的“超精细分裂”问题，因为这是一个我们已经能够用量子力学处理的、物理学中有兴趣的例子。这是一个具有两个以上的态的例子，可以用它说明量子力学应用到稍为复杂一些问题上的方法。这个问题也足够复杂，你懂得了处理它的方法后就能立即推广到各种问题上去。

如你们所知，氢原子包含有一个位于质子附近的电子，它可以处于许多分立的能量状态中的任何一个状态，在每一个能量状态中电子的运动图式都不相同。例如，第一激发态位于基态之上 $3/4$ 里德伯能量，或者大约 10 eV 处。但是由于电子和质子都具有自旋，即使是氢的所谓基态，其实也不是具有单一的确定能量的态。正是它们的自旋造成了能级的“超精细结构”，将所有的能级都分裂成几个靠近的能级。

电子可以具有或者“朝上”或者“朝下”的自旋，质子的自旋也可以“朝上”或者“朝下”。因此，原子的每一种动力学条件下都存在着 4 种可能的自旋态。这就是说，当人们谈到氢原子的“基态”时，实际上是指这“4 个基态”，而不是指能量最低的态。这 4 种自旋态并不都具有完全相同的能量，它们对于无自旋时的能量稍有移动。但是与基态到第一激发态之间 10 eV 左右的能量差相比这种能量的移动是非常非常小的。因此，每个动力学状态的能级都分裂成一组彼此非常靠近的能级，这就是所谓的超精细分裂。

这 4 个自旋态之间的能量差就是我们在本章中要计算的。超精细分裂是由电子和质子磁矩之间的相互作用引起，这种相互作用对于各个自旋态给出稍微不同的磁能。这些能量移动大约只有 10^{-7} eV ——与 10 eV 相比实在是太小了！正因为能级之间有很大的差距，所以我们把氢原子的基态看作是一个“四态”系统，而不必为实际上有很多更高能量的状态这一事实操心。这里我们只限于研究氢原子基态的超精细结构。

就我们的目的而言，我们对电子和质子空间位置的细节丝毫不感兴趣，因为它们的位置由原子结构决定——原子进入基态本身就决定了它们的位置。我们需要知道的只是电子和质子以某种确定的空间关系彼此靠近。此外，它们的自旋可以有多种不同的相对取向。我们要研究的只是自旋的效应。

我们一定要回答的第一个问题是：这系统的基础态是什么？但这个问题表述不正确。没有规定的基础态，因为你们可选取的基础态组不是唯一的。将老的一组线性组合就可以构成新的一组。基础态总是有多种选择，这些选择都同样的合法。所以问题并不是这组基础态是什么，而是什么可以作为基础态。我们可以按我们的方便选取想要的任何一组基础态。通常最好以一组物理意义最明确的基础态作为开始。它可能不是任何问题的解，也可能不具有任何直接的意义，但它一般会使产生的物理现象易于理解。

我们选择下列4个基础态：

态1:电子和质子的自旋都“朝上”。

态2:电子自旋“朝上”，质子自旋“朝下”，

态3:电子自旋“朝下”，质子自旋“朝上”，

态4:电子和质子的自旋都“朝下”。

对于这4个态，要有一种简便的记法，为此，我们将用下面这种方式来表示它们：

态1: $|++\rangle$, 电子朝上, 质子朝上,

态2: $|+-\rangle$, 电子朝上, 质子朝下,

态3: $|--\rangle$, 电子朝下, 质子朝上,

态4: $|--\rangle$, 电子朝下, 质子朝下。

(12.1)

你一定得记住，第一个+或-号表示电子，而第二个则表示质子。为便于参考，我们已把这些记法总结在图12-1中。有时也把这些态方便地记为 $|1\rangle$, $|2\rangle$, $|3\rangle$ 和 $|4\rangle$ 。

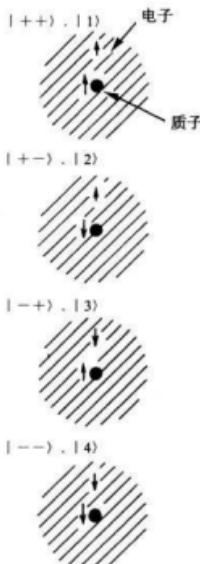


图 12-1 氢原子基态的一组基础态

你或许会说：“但是这些粒子有相互作用，可能这些态不是恰当的基础态。听起来好像你把这两个粒子看作是相互独立的。”确实如此！相互作用产生了问题：系统的哈密顿是什么？但如何描述这个系统并不涉及到相互作用。我们选取什么作为基础态可以不涉及到以后发生的情况。一个原子即使它一开始处于这些基础态中的一个态，它也不可能永远保持这个态。这是另外一个问题。问题是：在一个特殊（固定）基础态中的振幅如何随时间变化？选取基础态只是为我们的描述选取“单位矢量”而已。

趁讨论这个课题时，让我们来看一下在多于一个粒子的情况下如何寻求一组基础态这样一个普遍性的问题。你们已知道关于单个粒子的基础态。例如，对现实生活——不是我们这种简化的情况，而是现实生活——中的一个电子的完整描述是给出它在下列各个态中的振幅：

|动量为 \mathbf{p} 的自旋“朝上”的电子〉

或 |动量为 \mathbf{p} 的自旋“朝下”的电子〉。

实际上存在两组无限多态，每个 \mathbf{p} 值一个态。这就是说，如果你们知道了所有的振幅

$\langle +, \mathbf{p} | \psi \rangle$ 及 $\langle -, \mathbf{p} | \psi \rangle$,

则一个电子的状态 $|\psi\rangle$ 就被完全描述了。式中+和-表示角动量沿某个轴——通常指 z 轴——的分量， \mathbf{p} 为动量矢量。因此，对每个可能的动量，必定存在两个振幅（一组多重无限基础态）。对单个粒子的描述就是这些。

当存在一个以上粒子时，其基础态可用类似的方式写出。例如，若有一个电子和一个质子处在比我们考虑的更为复杂的情况下，其基础态可能是如下的类型：

|一个自旋“朝上”、以动量 \mathbf{p}_1 运动的电子和一个自旋“朝下”、以动量 \mathbf{p}_2 运动的质子〉。其他自旋组合以此类推。如果存在两个以上的粒子——同样的思路。所以，你们看到，要写出可能的基础态实际上是很容易的。唯一的问题是：哈密顿是什么？

就研究氢原子基态而言,我们不需要用到全部不同动量的基础态。当我们说“基态”时,我们是指特定的质子和电子的动量状态。组态——所有动量基础态的振幅——的细节是可以计算的,但这是另一个问题。现在所关心的只是自旋效应,因此可只取式(12.1)的4个基础态。接下来的问题是:对于这组基础态的哈密顿是什么?

§12-2 氢原子基态的哈密顿

我们马上就来回答这个问题。但首先应提醒你们一件事:任何态总能够写成基础态的线性组合。对于任何一个态 $|\psi\rangle$ 我们可以写成

$$|\psi\rangle = |++\rangle\langle++| + |+-\rangle\langle+-| + |-+\rangle\langle-+| + |--\rangle\langle--| \quad (12.2)$$

请记住,完整的括号只不过是复数,所以我们也可以把它们写成常用的形式 C_i ,其中*i*=1,2,3或4,因而式(12.2)可以写成

$$|\psi\rangle = |++\rangle C_1 + |+-\rangle C_2 + |-+\rangle C_3 + |--\rangle C_4. \quad (12.3)$$

只要给出了4个振幅 C_i ,我们就能完整地描述自旋态 $|\psi\rangle$ 。如果这4个振幅随时间而变化,事实上也是如此,则它们的时间变化率由算符 \hat{H} 给出。于是,问题在于找出 \hat{H} 。

如何写出一个原子体系的哈密顿并没有一个普遍的规则可循,而找到一个正确的公式比起找到一组基础态来更需要技巧。我们可以告诉你们一个普遍规则,据此可写出任何关于一个电子和一个质子问题的一组基础态,但是在目前的水平上要去描述这种组合的一般的哈密顿却是太难了。因此,我们将通过某种启发式的论证给你们导出一个哈密顿,而你们一定要承认它是正确的哈密顿,因为用它所得出的结果与实验观察一致。

你们应记得,在上一章中我们可以用 σ 矩阵——或与之完全等价的 σ 算符——来表示单个自旋1/2粒子的哈密顿。这些算符的性质概括在表12-1中。这些算符——它们只是表述 $|+\sigma_z|+$ 这种类型的矩阵元的一种简略而方便的方式——适用于描述自旋1/2的单个粒子的行为。问题是:我们是否能找到一种类似的记号来描述具有两个自旋粒子的体系?答案是肯定的,而且非常简单,具体如下。我们发明一种称之为“ σ 电子”的东西,用矢量算符 σ^e 来表示,它具有x,y和z的分量 σ_x^e , σ_y^e , σ_z^e 。现在我们作出一个约定,当这些算符中任何一个作用到氢原子的4个基础态之一时,它只作用于电子的自旋,而且其作用完全像电子单独存在时一样。例如, $\sigma_z^e|+-\rangle$ 是什么?由于 σ_z 作用于“朝下”电子得出-i乘以相应的电子“朝上”的态,即

$$\sigma_z^e|+-\rangle = -i|++\rangle.$$

(当 σ^e 作用于组合态时,它使电子的自旋倒转,而对质子不起作用,再对结果乘以-i。) σ_z^e 作

表 12-1

$\sigma_z +\rangle = + +\rangle$	$\sigma_z - \rangle = - -\rangle$
$\sigma_x +\rangle = + -\rangle$	$\sigma_x - \rangle = + +\rangle$
$\sigma_y +\rangle = +i -\rangle$	$\sigma_y - \rangle = -i +\rangle$

用于其他的态给出

$$\begin{aligned}\sigma_y^z |++\rangle &= i |-+\rangle, \\ \sigma_y^z |+-\rangle &= i |--\rangle, \\ \sigma_y^z |--\rangle &= -i |+-\rangle.\end{aligned}$$

要记住算符 σ^z 只作用于第一个自旋符号——即作用于电子的自旋。

接着, 我们定义质子的自旋相应的“ σ 质子”算符。它的 3 个分量 σ_x^p , σ_y^p , σ_z^p 的作用方式与 σ^z 相似, 只作用在质子自旋上。例如, 如果我们把 σ_z^p 作用于 4 个基础态的每一个上, 我们得到——仍旧利用表 12-1——

$$\begin{aligned}\sigma_z^p |++\rangle &= |+-\rangle, \\ \sigma_z^p |+-\rangle &= |++\rangle, \\ \sigma_z^p |-+\rangle &= |--\rangle, \\ \sigma_z^p |--\rangle &= |+-\rangle.\end{aligned}$$

你可以看到, 这并不很困难。

在最一般的情况下, 事情可能更复杂一些。例如, 我们可能有像 $\sigma_x^z \sigma_z^p$ 这样的两个算符的乘积。当我们碰到这种乘积的运算时, 我们先作右边算符的运算, 然后再作另一个算符的运算*。例如, 我们有

$$\sigma_x^z \sigma_z^p |+-\rangle = \sigma_x^z (\sigma_z^p |+-\rangle) = \sigma_x^z (-|+-\rangle) = -\sigma_x^z |+-\rangle = -|--\rangle.$$

注意这些算符对纯数没有任何作用——我们在写出 $\sigma_x^z (-1) = (-1)\sigma_x^z$ 时就利用了这一事实。我们说这些算符与纯数“对易”, 或者说一个数“可以移到算符另一边”。作为练习。你可以证明一下乘积 $\sigma_x^z \sigma_z^p$ 对 4 个基础态给出下列结果:

$$\begin{aligned}\sigma_x^z \sigma_z^p |++\rangle &= +|-+\rangle, \\ \sigma_x^z \sigma_z^p |+-\rangle &= -|--\rangle, \\ \sigma_x^z \sigma_z^p |-+\rangle &= +|++\rangle, \\ \sigma_x^z \sigma_z^p |--\rangle &= -|+-\rangle.\end{aligned}$$

如果我们取所有可用的算符, 每一种算符仅用一次, 则有 16 种可能性。是的, 16 种——倘若我们把“幺正算符” i 也包括进去。首先是 3 个: σ_x^z , σ_y^z , σ_z^z 。然后是 σ_x^p , σ_y^p 和 σ_z^p 3 个, 总共有 6 个。此外, 还有 9 个 $\sigma_x^z \sigma_z^p$ 这种形式的乘积, 使总数达到 15 个。另外, 还有一个使任何态不变的“幺正算符”。总共 16 个。

现在注意, 一个 4 个态的系统的哈密顿矩阵必定是一个系数的 4×4 矩阵——它有 16 项。不难证明, 任何一个 4×4 的矩阵——特别是哈密顿矩阵——都可以写成与我们刚才写出的一组算符相对应的 16 个双自旋矩阵的线性组合。因此, 对于一个电子与一个质子之间仅仅涉及自旋的相互作用来说, 我们能预期其哈密顿算符可以写成同样的 16 个算符的线性组合。问题仅在于如何写。

首先, 我们知道相互作用不依赖于坐标轴的选取。如果没有外界扰动——如磁场之

* 对于这几个特定的算符, 你会注意到这些算符的次序是无所谓的。

类——来确定空间的独特方向，哈密顿就不可能与我们选取的 x 、 y 和 z 轴的方向有关。这意味着从哈密顿不应具有孤零零的 σ_z^2 这样的项。否则用不同的坐标系将得出不同的结果，这显然是荒谬的。

唯一的可能性是含么正矩阵的项，例如常数 a （乘上 \hat{I} ），以及不依赖于坐标的 σ 算符的某种组合——某种“不变”组合。两个矢量的唯一标量不变组合是标积，对我们的 σ 来说就是

$$\sigma^x \cdot \sigma^y = \sigma_x^z \sigma_y^z + \sigma_x^y \sigma_y^x + \sigma_x^z \sigma_y^x. \quad (12.4)$$

这个算符对坐标系的任何转动都是不变式。所以具有特定的空间对称性的哈密顿，只可能是一常数乘以么正矩阵加上一个常数乘以上述标积，即

$$\hat{H} = E_0 + A\sigma^x \cdot \sigma^y. \quad (12.5)$$

这就是我们所要求的哈密顿。只要不存在外场，根据空间对称性它是唯一的可能。常数项没有多大意义，它仅取决于我们选取来测量能量的起始能级。我们完全可以取 $E_0 = 0$ 。第二项才告诉我们求氢原子能级分裂所需知道的一切。

要是你愿意的话，你可以不同的方式来考虑这哈密顿。如果有两个磁矩为 μ_e 和 μ_p 且彼此靠得很近的磁体，相互作用能量依赖于 $\mu_e \cdot \mu_p$ ——包括在其他各种能量中。而且你记住，我们曾发现经典的 μ_e 在量子力学中以 $\mu_e \sigma_e$ 的形式出现。同样，经典理论中的 μ_p 通常在量子力学中表示为 $\mu_p \sigma_p$ （这里 μ_p 是质子的磁矩，它比 μ_e 约大 1 000 倍，且符号相反）。所以式(12.5)表明相互作用能和两个磁体之间的相互作用能相同——不过不是完全相同，因为两磁体之间的相互作用能取决于它们之间的径向距离。但是式(12.5)是——并且事实上也确实是——某种平均相互作用能。电子在原子内部各处运动，我们的哈密顿只给出平均相互作用能。总之，按照经典物理学，对于在空间以一定方式排列的电子和质子之间有一个正比于这两个磁矩之间夹角余弦的能量。这种经典的定性图像也许会帮助你们理解它的来源，但是重要的是，式(12.5)是正确的量子力学公式。

两磁体间经典相互作用能的数量级是两磁矩的乘积除以它们之间距离的立方。氢原子内电子与质子间的距离，粗略地说，为原子半径的一半，即 0.5 \AA 。因此我们可以作一个粗略估计：式中的常数 A 大致应等于两个磁矩 μ_e 和 μ_p 的乘积除以 0.5 \AA 的立方。这种估计所给出的数值还是可以的。当你们了解了氢原子全部量子理论——至今我们还没有讲到——就知道可以准确地算出 A 。实际上计算精确度已达到大约百万分之三十。所以，与氦分子的翻转常数 A 不同，它还不能由理论很好地计算出来，而氢的常数 A 可以由更为精确的理论算得。不过没关系，为了目前的目的，我们把 A 看作是一个能由实验确定的数，然后分析这种情况的物理意义。

取式(12.5)的哈密顿，我们可以将它和方程

$$i\hbar \dot{C}_i = \sum_j H_{ij} C_j \quad (12.6)$$

一同用来求出自旋相互作用对能级的影响。为此，我们需要找出与式(12.1)的 4 个基础态中每一对相对应的 16 个矩阵元 $H_{ij} = \langle i | H | j \rangle$ 。

我们先对 4 个基础态中的每一个求出 $\hat{H} | j \rangle$ 。例如，

$$\hat{H} |++\rangle = A \sigma^z \cdot \sigma^y |++\rangle = A |\sigma_x^z \sigma_x^y + \sigma_y^z \sigma_x^y + \sigma_z^z \sigma_x^y| |++\rangle. \quad (12.7)$$

利用稍早一些时候讲过的方法——如果已记住表 12-1 的话, 就很容易了——我们求出每对 σ 作用于 $|++\rangle$ 的结果。答案为

$$\begin{aligned}\sigma_x^z \sigma_x^y |++\rangle &= +|--\rangle, \\ \sigma_y^z \sigma_x^y |++\rangle &= -|--\rangle, \\ \sigma_z^z \sigma_x^y |++\rangle &= +|++\rangle.\end{aligned}\quad (12.8)$$

所以式(12.7)变为

$$\hat{H} |++\rangle = A (|--\rangle - |--\rangle + |++\rangle) = A |++\rangle. \quad (12.9)$$

因为我们的 4 个基础态都是互相正交的, 立即可得

$$\begin{aligned}\langle ++| H |++\rangle &= A \langle ++|++\rangle = A, \\ \langle +-| H |++\rangle &= A \langle +-|++\rangle = 0, \\ \langle -+| H |++\rangle &= A \langle -+|++\rangle = 0, \\ \langle --| H |++\rangle &= A \langle --|++\rangle = 0.\end{aligned}\quad (12.10)$$

由于 $\langle j | H | i \rangle = \langle i | H | j \rangle^*$, 我们已经能够写出振幅 C_1 的微分方程:

$$i\hbar \dot{C}_1 = H_{11} C_1 + H_{12} C_2 + H_{13} C_3 + H_{14} C_4$$

或

$$i\hbar \dot{C}_1 = AC_1. \quad (12.11)$$

就这些! 我们仅得到一项。

表 12-2 氢原子的自旋算符

$\sigma_x^z \sigma_x^y ++\rangle = + --\rangle$	$\sigma_y^z \sigma_x^y ++\rangle = - --\rangle$	$\sigma_z^z \sigma_x^y ++\rangle = + ++\rangle$
$\sigma_x^z \sigma_x^y +-\rangle = + -+\rangle$	$\sigma_y^z \sigma_x^y +-\rangle = + -+\rangle$	$\sigma_z^z \sigma_x^y +-\rangle = - +-\rangle$
$\sigma_x^z \sigma_x^y -+\rangle = + +-\rangle$	$\sigma_y^z \sigma_x^y -+\rangle = + +-\rangle$	$\sigma_z^z \sigma_x^y -+\rangle = - +-\rangle$
$\sigma_x^z \sigma_x^y --\rangle = + ++\rangle$	$\sigma_y^z \sigma_x^y --\rangle = - ++\rangle$	$\sigma_z^z \sigma_x^y --\rangle = + --\rangle$

为了获得哈密顿方程中的其余项, 我们必须用相同的步骤把 \hat{H} 作用到其他态上去。首先, 作为练习, 请你核对一下我们在表 12-2 中写出的所有 σ 的乘积, 然后我们可以利用它们得出

$$\begin{aligned}\hat{H} |+-\rangle &= A \{2 |+-\rangle - |+-\rangle\}, \\ \hat{H} |-+\rangle &= A \{2 |-+\rangle - |-+\rangle\}, \\ \hat{H} |--\rangle &= A |--\rangle.\end{aligned}\quad (12.12)$$

然后依次对每一项在左边乘上所有其他的态矢量, 我们就得到如下的哈密顿矩阵 H_y

$$H_y = \overset{i}{\downarrow} \overset{j}{\rightarrow} \begin{bmatrix} A & 0 & 0 & 0 \\ 0 & -A & 2A & 0 \\ 0 & 2A & -A & 0 \\ 0 & 0 & 0 & A \end{bmatrix}. \quad (12.13)$$

当然,这意思仅仅是表明4个振幅 C_i 的微分方程为:

$$\begin{aligned} i\hbar C_1 &= AC_1, \\ i\hbar C_2 &= -AC_2 + 2AC_3, \\ i\hbar C_3 &= 2AC_2 - AC_3, \\ i\hbar C_4 &= AC_4. \end{aligned} \quad (12.14)$$

在解这些方程之前,我们忍不住要告诉你们一个由狄拉克得出的巧妙规则——它将使你们感到自己真正获得了很大的进步——虽然我们现在的工作并不需要它。由方程式(12.9)及(12.12),我们得到

$$\begin{aligned} \sigma^x \cdot \sigma^y |++\rangle &= |++\rangle, \\ \sigma^x \cdot \sigma^y |+-\rangle &= 2 |+-\rangle - |+-\rangle, \\ \sigma^x \cdot \sigma^y |-+\rangle &= 2 |-+\rangle - |-+\rangle, \\ \sigma^x \cdot \sigma^y |--\rangle &= |--\rangle. \end{aligned} \quad (12.15)$$

请看,狄拉克说:我也可以把第一和最后的方程写成

$$\begin{aligned} \sigma^x \cdot \sigma^y |++\rangle &= 2 |++\rangle - |++\rangle, \\ \sigma^x \cdot \sigma^y |--\rangle &= 2 |--\rangle - |--\rangle; \end{aligned}$$

于是这4个方程就十分相像了。现在我发明一个新的算符,我称它为 $P_{\text{自旋交换}}$,并且定义它有如下性质*:

$$\begin{aligned} P_{\text{自旋交换}} |++\rangle &= |++\rangle, \\ P_{\text{自旋交换}} |+-\rangle &= |-+\rangle, \\ P_{\text{自旋交换}} |-+\rangle &= |+-\rangle, \\ P_{\text{自旋交换}} |--\rangle &= |--\rangle. \end{aligned}$$

该算符的作用是交换两个粒子的自旋方向。这样,我就可把式(12.15)中的全部方程写成一个简单的算符方程:

$$\sigma^x \cdot \sigma^y = 2P_{\text{自旋交换}} - 1. \quad (12.16)$$

这就是狄拉克的公式。他的“自旋交换算符”提供了一个演算 $\sigma^x \cdot \sigma^y$ 的简便规则。(你们看现在你们什么都可以做了。大门打开了。)

§12-3 能 级

现在我们已准备就绪,可以通过解哈密顿方程式(12.14)来算出氢原子基态的能量级了。我们要求定态的能量。这就是说我们要找到那些特定的状态 $|\psi\rangle$,对它们来说,属于 $|\psi\rangle$ 的一组振幅 $C_i = \langle i | \psi \rangle$ 的每一个均有相同的时间依赖关系——即 $e^{-i\omega t}$ 。这个态具有能量 $E = \hbar\omega$ 。所以我们所要的一组振幅为

* 这算符现在称为“泡利自旋交换算符”。

$$C_i = a_i e^{(-i/\hbar)E}, \quad (12.17)$$

式中4个系数 a_i 与时间无关。为了看出能否获得这组振幅，我们把式(12.17)代入方程式(12.14)，看看有何结果。方程式(12.14)中的每个 $i\hbar dC/dt$ 变为 EC ，并且——消去共同的指数因子后——各个 C 都变成了 a ，我们得到：

$$\begin{aligned} Ea_1 &= Aa_1, \\ Ea_2 &= -Aa_2 + 2Aa_3, \\ Ea_3 &= 2Aa_2 - Aa_3, \\ Ea_4 &= Aa_4, \end{aligned} \quad (12.18)$$

我们从这些方程解出 a_1, a_2, a_3 和 a_4 。正好第一个方程与其他方程无关——这意味着我们立即可以看出一个解。若我们取 $E = A$ ，则

$$a_1 = 1, a_2 = a_3 = a_4 = 0,$$

是一个解。(当然，若取所有 a 的值均为零也是一个解，但这表示根本没有任何态存在。)让我们把这第一个解称为态 $|I\rangle^*$ ：

$$|I\rangle = |1\rangle = |++\rangle, \quad (12.19)$$

它的能量为

$$E_I = A.$$

有了这一线索，你们立即可从式(12.18)中最后一个方程看出另一个解：

$$\begin{aligned} a_1 &= a_2 = a_3 = 0, a_4 = 1, \\ E &= A. \end{aligned}$$

我们称该解为态 $|II\rangle$ ：

$$\begin{aligned} |II\rangle &= |4\rangle = |--\rangle, \\ E_{II} &= A. \end{aligned} \quad (12.20)$$

现在要稍微困难些了，式(12.18)中余下的两个方程是混合的。但是以前我们也碰到同样的情况。将此两式相加，我们得到

$$E(a_2 + a_3) = A(a_2 + a_3), \quad (12.21)$$

将两式相减，得

$$E(a_2 - a_3) = -3A(a_2 - a_3). \quad (12.22)$$

通过观察——并回忆氯的情况——我们看到存在两组解：

$$\begin{aligned} a_2 &= a_3, E = A; \\ a_2 &= -a_3, E = -3A. \end{aligned} \quad (12.23)$$

它们是 $|2\rangle$ 和 $|3\rangle$ 的混合物。把这两个态称为 $|III\rangle$ 和 $|IV\rangle$ ，并乘上一个因子 $1/\sqrt{2}$ ，使这两个态归一化，我们得到

$$|III\rangle = \frac{1}{\sqrt{2}}(|2\rangle + |3\rangle) = \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle),$$

* 此态实际上为 $|I\rangle e^{-(i/\hbar)E_I t}$ ，但我们通常用常数矢量来表示，该矢量等于 $t = 0$ 时的完全矢量。

$$E_{\pm} = A, \quad (12.24)$$

及

$$|\text{IV}\rangle = \frac{1}{\sqrt{2}}(|2\rangle - |3\rangle) = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle),$$

$$E_{\mp} = -3A, \quad (12.25)$$

我们得出了4个定态及其能量。顺便提一下，注意到这4个态是相互正交的，所以如果我们愿意，也可以把它们作为基础态。我们的问题完全解决了。

上述4个态中3个态的能量为A，而第4个态的能量为-3A。其平均能量为0——这意味着如果我们在式(12.5)中取 $E_0=0$ ，就是我们选择以平均能量为基准来量度所有能量。我们可以把氢原子基态能级图画在图12-2中。

态|IV>与其他任何一个态之间的能量差为4A。一个处于态|I>的原子，可能从该态落至态|IV>并发光。但不是光学光子，因为能量太小了——它将发射出一个微波量子。或者，若用微波照射氢气，当处在态|IV>的原子获得能量，并跃迁至上能态时，我们将发现能量的吸收——但只吸收频率 $\omega = 4A / \hbar$ 。这个频率已经由实验测得，最近所得到的最好结果*为：

$$f = \frac{\omega}{2\pi} = (1420405751.800 \pm 0.028) \text{ Hz.} \quad (12.26)$$

其误差仅为一千亿分之二！大概没有一个基本物理量测得比它还精确的了——它是物理学中最最精确的测量之一。理论家对于他们能把能量计算到 $3/10^5$ 的准确度感到非常高兴，而此时它却被测量至 $2/10^{11}$ 的准确度——比理论准确100万倍。所以实验物理学家大大领先于理论家。在氢原子基态理论方面，你已不比别人差。你也完全可以从实验中求得自己的A值，这也是每个人最后必须做的工作。

你以前大概已听说过氢的“21 cm 谱线”这件事吧，这就是氢的两个超精细态之间1420兆周谱线的波长。星系中的氢原子气体发射或吸收这一波长的辐射。所以用调谐至21 cm波（或约1420 MHz）的射电望远镜，我们就能观察氢原子气体浓集处的位置和速度。通过强度的测量，我们可以估计氢的数量；从测量多普勒效应引起的频移，我们可以求出星系中气体的运动情况。这是射电天文学的宏大计划之一。所以我们现在谈论的是完全真实的事情——它并不是一个臆造的问题。

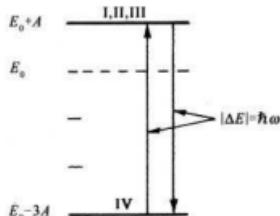


图12-2 氢原子基态的能级图

§12-4 塞曼分裂

虽然我们已经完成了求氢原子基态能级的问题，但是我们还要再进一步研究这一有趣的系统。为了对此系统有更多的了解——例如，为了计算氢原子吸收和发射21 cm的射电波的速率——我们必须知道原子受到扰动时发生的情况。我们要照我们求解氨分子那样去

* Crampton, Kleppner, and Ramsey; *Physical Review Letters*, Vol. 11, 338(1963).

做——在求出能级后进一步研究氨分子在电场中的情形。于是我们就能够计算出射电波的电场产生的效应。对氢原子来说,电场除了使所有的能级都移动一个与电场的平方成正比的常量外,别无作用——由于它并不改变能量差,故不必注意。现在重要的是磁场。所以接下来就是要写出原子处在外磁场中这种更为复杂的情况的哈密顿。

那么哈密顿是什么呢?这里我们只告诉你们答案,除了说原子的行为就是这样外不再给你们任何“证明”。

哈密顿为

$$\hat{H} = A(\boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^p) - \mu_e \boldsymbol{\sigma}^e \cdot \mathbf{B} - \mu_p \boldsymbol{\sigma}^p \cdot \mathbf{B}. \quad (12.27)$$

它由三部分组成。第一项 $A\boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^p$ 表示电子和质子间的磁相互作用——即使没有磁场也同样有这一项。这是我们原来就有的项;磁场对常数 A 的影响可忽略不计。外磁场的影响反映在最后两项中。第二项 $-\mu_e \boldsymbol{\sigma}^e \cdot \mathbf{B}$,是电子单独处于外磁场中具有的能量*,同样,最后一项 $-\mu_p \boldsymbol{\sigma}^p \cdot \mathbf{B}$ 是质子单独存在于磁场中具有的能量。在经典理论中两个粒子共同的总能量是它们两个的能量之和,在量子力学中也同样如此。在磁场中,由于磁场所产生的相互作用能,就是电子同外磁场的相互作用能与质子同外磁场的相互作用能之和——两者都用 σ 算符来表示。在量子力学中,这些项并不真正是能量,但是按经典的能量公式去想象它们是记住写出哈密顿规则的一种方法。不管怎样,式(12.27)是正确的哈密顿。

现在我们必须回到起点,再从头来解这一问题。但是有许多工作已经完成了——我们只需把新的项产生的效应加进去,我们令恒定磁场 \mathbf{B} 沿 z 方向,于是必须在我们的哈密顿算符中加进两个新的项——我们可以称之为 \hat{H}' :

$$\hat{H}' = -(\mu_e \sigma_z^e + \mu_p \sigma_z^p) B.$$

利用表 12-1,我们立即得到

$$\begin{aligned}\hat{H}' |++\rangle &= -(\mu_e + \mu_p) B |++\rangle, \\ \hat{H}' |+-\rangle &= -(\mu_e - \mu_p) B |+-\rangle, \\ \hat{H}' |-+\rangle &= -(-\mu_e + \mu_p) B |-+\rangle, \\ \hat{H}' |--\rangle &= (\mu_e + \mu_p) B |--\rangle.\end{aligned} \quad (12.28)$$

多么方便! \hat{H}' 作用于每个态只是得到一个数乘以这个态。因此,矩阵 $\langle i | H' | j \rangle$ 只有对角元——我们只要直接把式(12.28)中的系数加到式(12.13)中相应的对角项上去,于是哈密顿方程式(12.14)变为:

$$\begin{aligned}i\hbar \frac{dC_1}{dt} &= |A - (\mu_e + \mu_p) B| C_1, \\ i\hbar \frac{dC_2}{dt} &= -|A + (\mu_e - \mu_p) B| C_2 + 2AC_3, \\ i\hbar \frac{dC_3}{dt} &= 2AC_2 - |A - (\mu_e - \mu_p) B| C_3, \\ i\hbar \frac{dC_4}{dt} &= |A + (\mu_e + \mu_p) B| C_4.\end{aligned} \quad (12.29)$$

* 请记住,在经典理论中 $U = -\boldsymbol{\mu} \cdot \mathbf{B}$, 所以当磁矩沿着磁场方向时能量最低。对于带正电荷的粒子,磁矩平行于自旋,带负电荷的粒子则相反。所以在式(12.27)中 μ_p 是正数,而 μ_e 是负数。

方程的形式并没有什么不同——只是系数不同而已。只要 B 不随时间变化, 我们就可以继续照以前那样去做。将 $C_i = a_i e^{-(i/\hbar)B}$ 代入, 我们得到——式(12.18)的变形——

$$\begin{aligned} Ea_1 &= A\{-(\mu_e + \mu_p)B\}a_1, \\ Ea_2 &= -[A + (\mu_e - \mu_p)B]a_2 + 2Aa_3, \\ Ea_3 &= 2Aa_2 - [A - (\mu_e - \mu_p)B]a_3, \\ Ea_4 &= [A + (\mu_e + \mu_p)B]a_4. \end{aligned} \quad (12.30)$$

很巧, 上式中第一和第四个方程仍旧与其他方程无关, 所以我们仍可用同样的技巧求解。

态 $|I\rangle$ 是一个解, 对于这个解 $a_1 = 1, a_2 = a_3 = a_4 = 0$, 或

以及

$$|I\rangle = |1\rangle = |+\rangle, \quad (12.31)$$

$$E_I = A - (\mu_e + \mu_p)B.$$

另一个解是

以及

$$|II\rangle = |4\rangle = |-\rangle, \quad (12.32)$$

$$E_{II} = A + (\mu_e + \mu_p)B.$$

在剩下的两个方程中, 因为系数 a_2 和 a_3 不再相等, 所以稍微复杂些。但是它们和我们曾得到过的氨分子的一对方程很像。回顾一下式(9.20), 我们就能作出如下类比(记住那里的指标 1 和 2 相当于这里的 2 和 3)

$$\begin{aligned} H_{11} &\longrightarrow -A - (\mu_e - \mu_p)B, \\ H_{12} &\longrightarrow 2A, \\ H_{21} &\longrightarrow 2A, \\ H_{22} &\longrightarrow -A + (\mu_e - \mu_p)B. \end{aligned} \quad (12.33)$$

于是能量由式(9.25)给出, 它们是

$$E = \frac{H_{11} + H_{22}}{2} \pm \sqrt{\frac{(H_{11} - H_{22})^2}{4} + H_{12}H_{21}}, \quad (12.34)$$

将式(12.33)代入上式, 能量表达式就成为

$$E = -A \pm \sqrt{(\mu_e - \mu_p)^2 B^2 + 4A^2}.$$

虽然我们在第 9 章中把这两个能量称为 E_I 和 E_{II} , 而在目前的问题中我们把它们称为 E_{\pm} 和 E_{\mp} ,

$$\begin{aligned} E_{\pm} &= A \left\{ -1 + 2\sqrt{1 + \frac{(\mu_e - \mu_p)^2 B^2}{4A^2}} \right\}, \\ E_{\mp} &= -A \left\{ 1 + 2\sqrt{1 + \frac{(\mu_e - \mu_p)^2 B^2}{4A^2}} \right\}. \end{aligned} \quad (12.35)$$

所以我们就找到了氢原子处在恒定磁场中的 4 个定态的能量。让我们令 B 趋向于零以核对上述结果是否与上节所得的能量相同。你们就可以看到我们会得到预期结果。当 $B = 0$ 时, 能量 E_I, E_{II} 和 E_{\pm} 等于 $+A$, 而 E_{\mp} 等于 $-3A$ 。甚至态的标号也和以前的一致, 可是当加上磁场时, 各项能量以不同的方式改变。让我们来看看它们如何变化。

首先,我们必须记住,电子的 μ_e 是负值,质子的 μ_p 为正值,而且 μ_e 约比 μ_p 大 1 000 倍。所以 $\mu_e + \mu_p$ 和 $\mu_e - \mu_p$ 都是负数,且接近相等。让我们把它们称为 $-\mu$ 和 $-\mu'$:

$$\mu = -(\mu_e + \mu_p), \quad \mu' = -(\mu_e - \mu_p), \quad (12.36)$$

(这里 μ 和 μ' 两者都是正数,差不多等于 μ_e 的大小——约为一个玻尔磁子。)于是我们的 4 个能量分别为

$$\begin{aligned} E_1 &= A + \mu B, \\ E_{\parallel} &= A - \mu B, \\ E_{\perp} &= A \left(-1 + 2 \sqrt{1 + \frac{\mu'^2 B^2}{4A^2}} \right), \\ E_{\nu} &= -A \left(1 + 2 \sqrt{1 + \frac{\mu'^2 B^2}{4A^2}} \right). \end{aligned} \quad (12.37)$$

能量 E_1 从 A 开始,随 B 线性增加——斜率为 μ 。能量 E_{\parallel} 也是从 A 开始,但是它随 B 的增加而线性地减少——其斜率为 $-\mu$ 。这两个能级随 B 的变化情况如图 12-3 所示。我们还在图中画出了能量 E_{\perp} 和 E_{ν} 。但它们与 B 的依赖关系不同。当 B 很小时,它们依赖于 B 的平方,所以开始时其斜率是水平的,然后变成曲线,当 B 很大时,它们趋近于斜率为 $\pm \mu'$ 的直线,该直线的斜率与 E_1 和 E_{\parallel} 的斜率几乎相等。

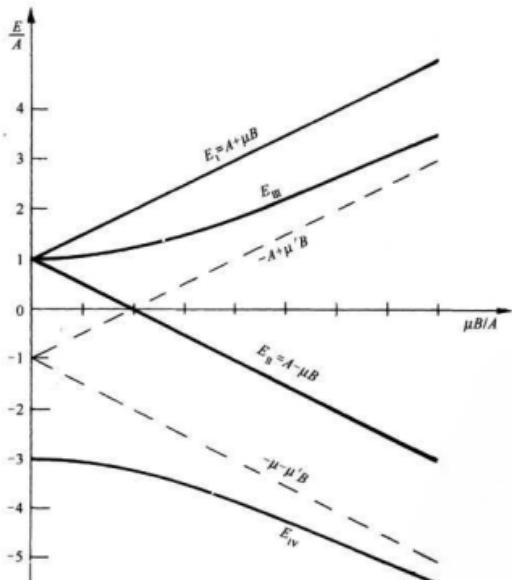


图 12-3 在磁场 B 中氢的基态能级图

由于磁场而引起的原子能级的移动称为塞曼效应。我们说,图 12-3 中的曲线表示氢原子基态的塞曼分裂。当不存在磁场时,我们仅得到氢超精细结构的一条谱线。态 $|IV\rangle$ 与其他任何一个态之间的跃迁,会吸收或发射一个光子,其频率 1420 兆周为 $1/h$ 乘以能量差 $4A$ 。但是,当原子处在磁场 B 中时,则有更多的谱线。4 个态中任何两个态之间的跃迁都是可能的。所以如果在所有 4 个态上都有原子,则在图 12-4 中 6 个垂直的箭头所示的任何一个吸收或发射能量的跃迁都可能发生。这其中的许多跃迁都可以用我们在第 2 卷 §35-3 所叙述的拉比(Rabi)分子束技术观察到。

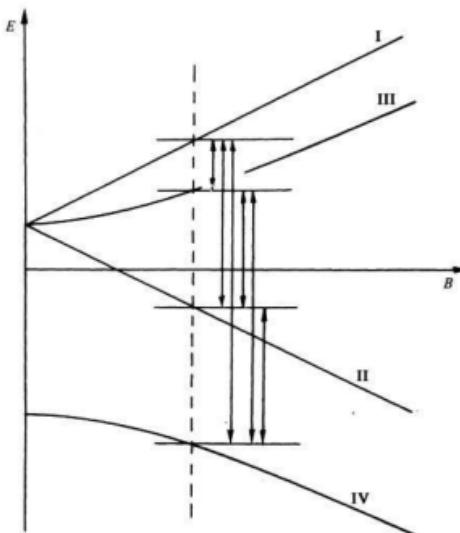


图 12-4 在特定磁场 B 中氢的基态能级之间的跃迁

引起这种跃迁的原因是什么呢? 如果我们在稳定的强磁场 B 中再加上一个随着时间变化的小的扰动磁场,就会发生这种跃迁。就像我们把变动着的电场加到氨分子上所看到的那样。只是这里是磁场与磁矩相耦合而变的把戏。但是理论与我们计算氨时所用的方法相同。如果我们在 xy 平面内转动微扰磁场则理论将最简单——虽然任何水平的振动磁场都行。你把这个微扰磁场作为附加项加到哈密顿中去,就得到振幅随时间变化的解——正如我们在讨论氨分子时所求得的那样。所以你可以很容易地精确计算出从一个态到另一个态的跃迁概率。你会发现它们与实验完全一致!

§12-5 在磁场中的态

现在我们要讨论图 12-3 中曲线的形状。首先,对于强磁场情况下的能量是易于理解

的,而且相当有趣。当磁场 B 足够强时(即 $\mu B/A \gg 1$),我们可以忽略式(12.37)中的 1。这 4 个能量变为

$$\begin{aligned} E_I &= A + \mu B, \quad E_{\text{II}} = A - \mu B, \\ E_{\text{III}} &= -A + \mu' B, \quad E_{\text{IV}} = -A - \mu' B. \end{aligned} \quad (12.38)$$

这就是图 12-3 中 4 条直线的方程式。我们可以用下述方式去理解这些能量的物理意义。磁场为零时定态的性质完全由两个磁矩的相互作用确定。在定态 $| \text{III} \rangle$ 和 $| \text{IV} \rangle$ 中,基础态 $| + \rangle$ 和 $| - \rangle$ 的混合就是由于这种相互作用。但是在强的外磁场中,质子和电子几乎一点也不受对方的场影响;它们各自都像独自处在外场中一样。于是——正如我们多次看到的那样——电子的自旋不是平行于外磁场就是与外磁场相反。

假定电子的自旋“朝上”——也就是沿着磁场方向,它的能量为 $-\mu_e B$ 。而质子仍旧可以有两种不同的方向。如果质子的自旋也是“朝上”,它的能量为 $-\mu_p B$ 。两者之和为 $-(\mu_e + \mu_p)B = \mu B$ 。这正是我们所得出的 E_I ——很好,因为我们正在描述态 $| ++ \rangle = | \text{I} \rangle$ 。仍存在一个表示电子和质子自旋平行时的相互作用能的微小附加项 A (现在 $\mu B \gg A$)。(原先我们取 A 为正,因为我们所用到的理论表明它应该是正的,而实验证明它确是这样。)另一方面,质子的自旋也可以朝下,于是它在外磁场中的能量为 $+\mu_p B$,因此它与电子能量之和为 $-(\mu_e - \mu_p)B = \mu' B$ 。而它们的相互作用能变为 $-A$ 。两者之和正好就是式(12.38)中的能量 E_{II} 。所以在强磁场中态 $| \text{III} \rangle$ 必定变成态 $| + \rangle$ 。

现在假定电子自旋“朝下”,它在外磁场中的能量为 $\mu_e B$ 。如果质子自旋也“朝下”,两者合在一起具有能量 $(\mu_e + \mu_p)B = -\mu B$,加上它们的相互作用能 A ——因为它们的自旋平行。这正好就是式(12.38)中的能量 E_{III} ,且相应于态 $| -- \rangle = | \text{II} \rangle$ ——这很妙。最后,若电子自旋“朝下”而质子自旋“朝上”,我们得到能量 $(\mu_e - \mu_p)B - A$ (相互作用能取为负 A 是因为两个自旋相反),它刚好就是 E_{IV} ,而相应的态为 $| - \rangle$ 。

“但是,等一下!”你或许会说,“态 $| \text{III} \rangle$ 和 $| \text{IV} \rangle$ 并非是态 $| + \rangle$ 和 $| - \rangle$;它们是后两者的混合。”有道理,但不全面。当 $B = 0$ 时它们的确是两者的混合态,但是我们还没有计算出它们在强磁场 B 时的情况。我们在第 9 章中用推导(12.33)式类似的方法计算了定态能量,现在也可以用第 9 章的公式求相应的振幅。从式(9.24)得到

$$\frac{a_2}{a_3} = \frac{E - H_{21}}{H_{21}}.$$

当然,比值 a_2/a_3 就是 C_2/C_3 。代入式(12.33)给出的相应的量,得

$$\begin{aligned} \text{或} \quad \frac{C_2}{C_3} &= \frac{E + A - (\mu_e - \mu_p)B}{2A} \\ \frac{C_2}{C_3} &= \frac{E + A + \mu' B}{2A}, \end{aligned} \quad (12.39)$$

式中的 E 可以采用适当的能量—— E_{II} 或 E_{IV} 。例如,对态 $| \text{III} \rangle$ 来说,我们有

$$\left(\frac{C_2}{C_3}\right)_{\text{II}} \approx \frac{\mu' B}{A}. \quad (12.40)$$

对于强磁场 B , 态 $|III\rangle$ 的 $C_2 \gg C_1$, 该态几乎完全成了态 $|2\rangle = |+-\rangle$ 。同样, 如果在式 (12.39) 中代入 E_B , 则得 $(C_2/C_1)_B \ll 1$; 强场中态 $|IV\rangle$ 正好成为态 $|3\rangle = |-+\rangle$ 。你们看到构成定态的基础态的线性组合的系数与 B 有关。我们称为 $|III\rangle$ 的态在极弱磁场的情况下是 $|+-\rangle$ 和 $|-+\rangle$ 的 50-50 的混合态, 但是在强场中就完全转移到 $|+-\rangle$ 。同样, 态 $|IV\rangle$ 在弱场中也是 $|+-\rangle$ 和 $|-+\rangle$ (具有相反的符号) 的 50-50 的混合态, 但当自旋在强的外场中解耦合后态 $|IV\rangle$ 就趋向态 $|-\rangle$ 。

我们还特别提请你们注意在极弱磁场下所发生的情况。有一个能级——在 $-3A$ 处——在加上弱磁场时并不发生变化。而另一个能级——在 $+A$ 处——在加上弱磁场时分裂成 3 个不同的能级。对于弱场, 能量随 B 的变化如图 12-5 所示。假定我们用某种方法选出一束氢原子, 它们全都具有 $-3A$ 的能量。如果我们用它们做施特恩-格拉赫实验——所用的场不太强——我们将发现它们沿直线通过实验装置。(因为它们的能量与 B 无关, 根据虚功原理——它们在磁场梯度中不受力的作用。) 另外, 假定我们选取具有能量 $+A$ 的原子束, 并使它们通过施特恩-格拉赫装置——譬如说 S 装置。(同样装置中的磁场也不应太强, 不影响原子的内部。也就是说磁场应足够小, 使能量随 B 线性变化。) 我们将会发现出来的原子分成 3 束。态 $|I\rangle$ 和 $|II\rangle$ 受到相反的作用力——它们的能量随 B 以斜率土 μ 作线性变化, 所以该力就与作用于磁矩为 $\mu_s = \pm \mu$ 的磁偶极子的力一样。但是态 $|III\rangle$ 沿直线通过。所以我们正好回到了第 5 章。一个能量为 $+A$ 的氢原子是一个自旋 1 粒子。这个能量状态属于一个 $j = 1$ 的“粒子”, 它可以用第 5 章中用的 3 个基础态 $|+S\rangle$, $|0S\rangle$, $| -S\rangle$ 来描写——对于空间的一组轴。另一个方面, 当氢原子具有能量 $-3A$ 时, 它是一个自旋零的粒子。(记住, 上面所说的情况仅对无限小的磁场才严格成立。) 所以我们可以把零磁场下氢原子的态以下列方法归类:

$$\begin{aligned} |I\rangle &= |++\rangle \\ |III\rangle &= \frac{|+-\rangle + |-+\rangle}{\sqrt{2}} \\ |II\rangle &= |--\rangle \end{aligned} \quad \left. \begin{array}{l} \text{自旋 } 1 \\ |+S\rangle \\ |0S\rangle \\ |-S\rangle \end{array} \right\} \quad (12.41)$$

$$|IV\rangle = \frac{|+-\rangle - |-+\rangle}{\sqrt{2}} \quad \text{自旋 } 0. \quad (12.42)$$

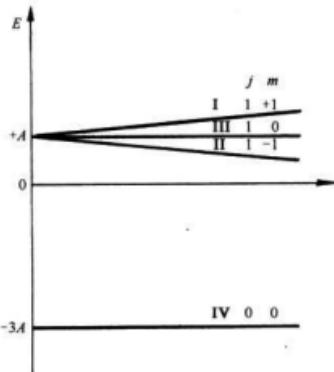


图 12-5 氢原子在弱磁场中的态

我们在第 2 卷第 35 章中曾说过, 任何一个粒子沿任何轴的角动量分量只可能具有相差为 \hbar 的某些值。角动量的 z 分量 J_z 可以是 $j\hbar$, $(j-1)\hbar$, $(j-2)\hbar$, …, $(-j)\hbar$, 这里 j 为粒子的自旋(可取整数或半整数)。虽然我们在这里不这么说, 但人们通常写成

$$J_z = m\hbar, \quad (12.43)$$

式中 m 代表 $j, j-1, j-2, \dots, -j$ 等数中的一个。所以，你们会看到，人们在书中用所谓量子数 j 和 m [j 常称为“总角动量量子数”， m 常称为“磁量子数”] 来标记氢原子的 4 个基态。于是他们把状态表示成 $|j, m\rangle$ 。而不是写成我们的态符号 $|I\rangle, |II\rangle$ 等等。所以他们把磁场为零时式(12.41)和(12.42)表示的态写用表 12-3 中所示的符号标记。这并不是什么新的物理，它只是记号的问题。

表 12-3 在零场时氢原子的态

态 $ j, m\rangle$	j	m	我们的记号
$ 1, +1\rangle$	1	+1	$ I\rangle = +S\rangle$
$ 1, 0\rangle$	1	0	$ III\rangle = 0S\rangle$
$ 1, -1\rangle$	1	-1	$ II\rangle = -S\rangle$
$ 0, 0\rangle$	0	0	$ IV\rangle$

§12-6 自旋 1 粒子的投影矩阵*

我们现在要用我们的氢原子的知识去解决一些特殊问题。在第 5 章中我们曾讨论过一个自旋 1 粒子对于一特定取向的施特恩-格拉赫装置——譬如说一个 S 装置——处于基础态 (+, 0 或 -) 之一，则该粒子对于不同空间取向的 T 装置具有一定的振幅处于其 3 个态的各个态中。这样的振幅 $\langle jT|is\rangle$ 有 9 个，它们组成投影矩阵。在 §5-7 中我们未加证明地给出了 T 相对 S 具有不同取向时投影矩阵中的各项。现在我们将告诉你们一种能够推导出这些项的方法。

在氢原子中我们曾经有一个由两个自旋 $1/2$ 粒子组成的自旋 1 的系统。我们在第 6 章中就已经解决了如何变换自旋 $1/2$ 的振幅。我们可以利用这种知识去计算自旋 1 的变换。方法是这样的：我们有一个系统——一个能量为 $+A$ 的氢原子——其自旋为 1。假定我们让它通过一施特恩-格拉赫过滤器 S ，因此我们知道它处于 S 的基础态中的一个，譬如说 $|+S\rangle$ 。而它对 T 装置中的某一基础态，譬如 $|+T\rangle$ 的振幅将是什么呢？如果我们把 S 装置的坐标系称为 x, y, z 系统， $|+S\rangle$ 态就是我们曾称之为 $|++\rangle$ 的态。但是，假定另一个人把他的 z 轴取在 T 的轴上。他以我们称为 x', y', z' 的参考系来表示他的态。他的电子和质子的“朝上”和“朝下”的态将与我们的不同。他的“正-正”态——我们写成 $|+'+'\rangle$ ，相对于带撇参考系——是自旋 1 粒子的 $|+T\rangle$ 态。我们想要的是 $\langle +T|+S\rangle$ ，这只是振幅 $\langle +'+'|++\rangle$ 的另一种写法。

我们可用下述方法求振幅 $\langle +'+'|++\rangle$ 。在我们的参考系中，处在 $|++\rangle$ 态的电子具有“朝上”的自旋。这意味着在他的参考系中具有“朝上”的振幅 $\langle +'|+\rangle_e$ ，以及“朝下”的振幅 $\langle -'|+\rangle_e$ 。同样，处在 $|++\rangle$ 态的质子在我们的参考系中自旋“朝上”，而在带撇的参考系中，具有自旋“朝上”和“朝下”的振幅分别为 $\langle +'|+\rangle_p$ 和 $\langle -'|+\rangle_p$ 。既然我们现在谈论的是两个不同的粒子，因此这两个粒子在他的参考系中同一“朝上”的振幅为上述两个振幅的乘积，即

$$\langle +'+'|++\rangle = \langle +'|+\rangle_e \cdot \langle +'|+\rangle_p. \quad (12.44)$$

为明确起见，我们在振幅 $\langle +'|+\rangle$ 后面加上下标 e 和 p 。但是它们都只是自旋 $1/2$ 粒子的变

* 凡是没有读过第 6 章的人也应略过本节。

换振幅，所以实际上它们是完全相同的数。事实上，它们就是第6章里称为 $\langle +T|+S \rangle$ 的振幅，也就是该章末尾表中所列的。

但是，现在我们在记号方面即将遇到麻烦了。我们必须能把对自旋1/2粒子的振幅 $\langle +T|+S \rangle$ 和对自旋1粒子的，我们也称为 $\langle +T|+S \rangle$ 的振幅区分开来——这两种振幅是完全不同的！我们希望它不至于过分混淆，但是至少在目前，对自旋1/2振幅，我们不得不用一种不同的符号。为简捷起见，我们将新的记号概括于表12-4中。对自旋1粒子的态，我们将继续使用记号 $|+S\rangle$, $|0S\rangle$ 和 $| -S\rangle$ 。

表 12-4 自旋1/2粒子的振幅

本 章	第6章
$a = \langle +' + \rangle$	$\langle +T +S \rangle$
$b = \langle -' + \rangle$	$\langle -T +S \rangle$
$c = \langle +' - \rangle$	$\langle +T -S \rangle$
$d = \langle -' - \rangle$	$\langle -T -S \rangle$

用我们现在的新记号，式(12.44)就变成

$$\langle +'+'|++\rangle = a^2,$$

而这正好就是自旋1的振幅 $\langle +T|+S \rangle$ 。现在让我们假定，另一个人的坐标系——即T或带“撇”的装置——相对于我们的z轴旋转了角度 ϕ ，那么从表6-2得：

$$a = \langle +'|+\rangle = e^{i\phi/2}.$$

所以由式(12.44)，我们得到自旋1粒子的振幅为

$$\langle +T|+S \rangle = \langle +'+'|++\rangle = (e^{i\phi/2})^2 = e^{i\phi}. \quad (12.45)$$

你们可以看出它是怎样得出的。

现在我们将求出对所有态的普遍情况。如果质子和电子在我们的参考系——S系——中两者都“朝上”，则在另一个人的参考系——T系——中，它处于4个可能态中任一个态的振幅为：

$$\begin{aligned} \langle +'+'|++\rangle &= \langle +'|+\rangle, \langle +'|+\rangle_p = a^2, \\ \langle +'-'|++\rangle &= \langle +'|+\rangle, \langle -'|+\rangle_p = ab, \\ \langle -'+'|++\rangle &= \langle -'|+\rangle, \langle +'|+\rangle_p = ba, \\ \langle -'-'|++\rangle &= \langle -'|+\rangle, \langle -'|+\rangle_p = b^2. \end{aligned} \quad (12.46)$$

于是我们可以把态 $|++\rangle$ 写成如下的线性组合：

$$|++\rangle = a^2 | +'+' \rangle + ab | +'-' \rangle + | -'+' \rangle + b^2 | -'-' \rangle. \quad (12.47)$$

我们注意到： $| +'+' \rangle$ 是态 $|+T\rangle$, $| +'-' \rangle + | -'+' \rangle$ 就是 $\sqrt{2}$ 乘以态 $|0T\rangle$ ——参看式(12.41)——而 $| -'-' \rangle = | -T\rangle$ 。换句话说，式(12.47)可以改写成

$$|+S\rangle = a^2 |+T\rangle + \sqrt{2}ab |0T\rangle + b^2 |-T\rangle. \quad (12.48)$$

用类似的方法你们可以很容易地证明：

$$|-S\rangle = c^2 |+T\rangle + \sqrt{2}cd |0T\rangle + d^2 |-T\rangle. \quad (12.49)$$

对于 $|0S\rangle$,要稍为复杂些,因为

$$|0S\rangle = \frac{1}{\sqrt{2}} (|+-\rangle + |-+\rangle).$$

但是,我们可以用带“撇”的态表示态 $|+-\rangle$ 和 $|-+\rangle$,然后取其和。这就是

$$|+-\rangle = ac |+'+'\rangle + ad |+'-' \rangle + bc |-'+'\rangle + bd |-'-' \rangle \quad (12.50)$$

及

$$|-\rangle = ac |+'+'\rangle + bc |+'-' \rangle + ad |-'+'\rangle + bd |-'-' \rangle. \quad (12.51)$$

以 $1/\sqrt{2}$ 乘以这两者之和,得

$$|0S\rangle = \frac{2}{\sqrt{2}} ac |+'+'\rangle + \frac{ad+bc}{\sqrt{2}} (|+'-' \rangle + |-'+'\rangle) + \frac{2}{\sqrt{2}} bd |-'-' \rangle.$$

由之得:

$$|0S\rangle = \sqrt{2}ac |+T\rangle + (ad+bc) |0T\rangle + \sqrt{2}bd |-T\rangle. \quad (12.52)$$

现在我们已经得到了所要的全部振幅。式(12.48), (12.49)及(12.52)中的系数就是矩阵元 $\langle jT | iS \rangle$ 。让我们把它们都放在一起:

$$\begin{array}{c} iS \\ \rightarrow \\ \langle jT | iS \rangle = \begin{bmatrix} a^2 & \sqrt{2}ac & c^2 \\ \sqrt{2}ab & ad+bc & \sqrt{2}cd \\ b^2 & \sqrt{2}bd & d^2 \end{bmatrix}. \end{array} \quad (12.53)$$

我们已经用自旋 $1/2$ 的振幅 a, b, c 和 d 表示了自旋 1 的振幅变换。

例如,如果 T 参考系相对于 S 系统 y 轴旋转 α 角——如图5-6所示——表12-4中的各个振幅正好就是表6-2中 $R_y(\alpha)$ 的矩阵元。

$$a = \cos \frac{\alpha}{2}, b = -\sin \frac{\alpha}{2},$$

$$c = \sin \frac{\alpha}{2}, d = \cos \frac{\alpha}{2}. \quad (12.54)$$

把这些代入式(12.53),我们就得到(5.38)的公式,我们在那里未加证明给出了该式。

态 $|IV\rangle$ 又是怎样呢?! 它是一个自旋零的系统,所以只有一个态——它在所有坐标系中都是相同的。我们可以通过取式(12.50)和(12.51)的差来核对每一个结果,我们得到:

$$|+-\rangle - |-\rangle = (ad - bc) (|+'-' \rangle - |-'+'\rangle).$$

但是, $(ad - bc)$ 是自旋 $1/2$ 矩阵的行列式,所以它等于 1 。对于两个具有任意相对取向的坐标系来说,我们得到

$$|IV'\rangle = |IV\rangle.$$

第13章 在晶格中的传播

§13-1 电子在一维晶格中的状态

乍看起来，你可能以为低能电子很难穿过固态晶体。晶体中的原子挤在一起，中心相距只有几个 Å ($1 \text{ Å} = 10^{-10} \text{ m}$)，而且原子对电子散射的有效直径大致也是 1 Å 左右。这就是说，比起它们之间的间距来，原子是很大的。因此你可以估计出两次碰撞之间的平均自由程是几个 Å 的数量级——实际上这是微不足道的，你可能会料到电子几乎动一动就撞在这个或那个原子上。然而，自然界中普遍存在着的现象却是：假如晶格是完美的，电子能轻而易举顺畅地通过晶体——几乎和在真空中的情形一样。正是这个奇妙的事实使得金属导电如此容易，也使得许多实用器件的开发成为可能。例如，它使晶体管模拟无线电电子管成为可能。在无线电电子管中电子自由地穿过真空，而在晶体管中电子自由地穿过晶格。晶体管行为的内部机理将在这一章中叙述；而这些原理在各种实用器件中的应用则放到下一章去讲解。

电子在晶体中的传导只是一种非常普遍的现象的一个例子。不仅电子可以穿过晶体，其他的“东西”像原子的激发也能够以同样的方式运动。所以我们要讨论的现象以多种形式出现在固态物理学的研究中。

你们还记得我们曾经讨论过双态系统的许多例子。现在让我们想象一个电子可以处在两个位置中的任一位置上，在每个位置上它都是处在相同的环境中。我们还假定电子具有一定的振幅从一个位置跑到另一个位置，当然，它也具有同样的振幅跑回来，就像我们在 §10-1 中对氢分子离子的讨论那样。量子力学定律给出下面的结果，对电子来说有两个可能的具有确定能量的状态。每一个状态可以用电子在这两个基础位置的振幅来描写。对于每一个能量确定的状态，振幅的大小都不随时间变化，而两者的相位以同一频率随时间变化。另一方面，如果开始时电子在一个位置，以后它会跑到另一个位置，再过一会它又会回到第一个位置。其振幅和两个耦合摆的运动相似。

现在来考虑完美晶体，我们想象其中一个电子处于某一特定原子的某种“陷阱”中，并具有某个特定的能量。还假定电子具有一定振幅运动到邻近另一个原子陷阱中。这有点像双态系统——但是更复杂些。当电子到达邻近的原子后，还可以继续运动到下一个新的位置，也可以回到原来的位置。现在的情形并不是类似两个耦合的摆，而是类似于无限个耦合在一起的单摆。这有点像你们在第一学年物理课上看到过的那种用来演示波的传播的机械——在扭转金属线上装有一长串杆棒。

假如你有一个谐振子，它被耦合到另一个谐振子上，而后者再耦合到下一个谐振子上，等等……，假如你在某一个位置上扰动一下，这个扰动会像波一样沿线传播。假如你把一个电子放在一长串原子中的一个原子上也会出现同样的情况。

通常，分析这种力学问题最简单的方法不是去考虑在某个地点发出一个脉冲后会发生

什么，而是去考虑稳态波的解。存在着某种位移的图样，它作为单一的不变频率的波在晶体中传播。基于同一理由，对电子也发生同样的情况——因为在量子力学中两者用相似的方程描述。

然而，你必须懂一件事：电子在一个地点的振幅是振幅而不是概率。如果电子像水流过小孔那样只是简单地从一处渗漏到另一处，它的行为就会完全不同。譬如说，假如有两个水槽，我们用管子将其连通，使水能够从一个水槽流到另一个水槽中去，那么两个水槽中的水面将会按指数规律相互趋近。但是对于电子，所发生的是振幅的流动而不是单纯的概率流动。这是虚数项——量子力学微分方程中的 i ——的特点，它使指数解变成振荡解。于是所发生的过程就完全不同于两个连通的水槽之间的渗漏。

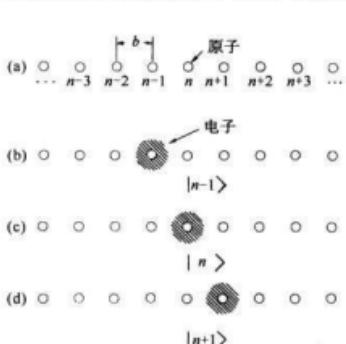


图 13-1 一维晶体中电子的基础态

现在我们来定量地分析这量子力学情况。设想由一长串原子组成的一维体系如图 13-1(a)所示。(当然，晶体是三维的，但它们的物理原理完全相同，你一旦理解了一维的情形，就能理解三维中所发生的过程。)下一步，我们来看一下如果把一个电子放到这一串原子上将会发生什么情况。当然，在真实的晶体中原来就已经有亿万个电子了。但是它们中的大多数(对于绝缘的晶体则几乎是所有的)都以某种方式围绕着它自己的原子运动——一切都十分稳定。然而，我们现在要考虑的是放进一个额外的电子后会发生些什么。我们不考虑其他的电子正在做什么，因为我们假定改变它们的运动状态需要很高的激发能量。加上一个电子就好像

制造出一个受到微弱束缚的负离子。在观察这一个额外的电子行为的时候，我们事实上取一个近似，即忽略原子内部的结构。

当然电子可以运动到另一个原子上，从而使负离子转移到另一个地点。就像电子可以在两个质子之间跳跃一样。我们假设电子具有从一个原子跳跃到在它任何一边的相邻原子上的一定的振幅。

我们现在怎样描写这样的体系呢？合理的基础态是什么？假如你还记得当我们只有两个可能的位置时我们曾经是怎样做的，你就能猜出我们将会怎么办。假设这一原子串中各个的间距都相等，我们把原子依次编号如图 13-1(a)所示。一个基础态是电子在第 6 号原子上，另一基础态是电子在第 7 号原子上或在第 8 号原子上，依此类推。我们可以说电子位于第 n 个原子处就表示第 n 基础态。让我们称它为基础态 $|n\rangle$ 。图 13-1 表示

$$|n-1\rangle, |n\rangle, \text{以及} |n+1\rangle$$

这 3 个基础态指的是什么。我们可以用这些基础态来描述我们的一维晶体中电子的任何状态 $|\psi\rangle$ ，只要给出 $|\psi\rangle$ 在每一基础态的振幅 $\langle n|\psi\rangle$ ——即电子位于某一特定原子的振幅。于是我们可以将状态 $|\psi\rangle$ 写成所有基础态的叠加：

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle. \quad (13.1)$$

下面我们假设电子位于某一原子时具有一定的振幅渗透到任何一边的原子中。我们将取最简单的情况，电子只能渗透到紧邻的原子中——要到达次邻近的原子，它必须走两步。我们令电子从一个原子跳跃到下一个原子的振幅为 iA/\hbar (每单位时间)。

我们目前把电子位于第 n 个原子的振幅 $|n\rangle$ 写成 C_n 。于是式(13.1)就要写成：

$$|\psi\rangle = \sum_n |n\rangle C_n. \quad (13.2)$$

假如我们已经在给定时刻的各个振幅 C_n 取它们的绝对值的平方，就能得到在该时刻观察第 n 个原子时发现电子的概率。

在以后的某个时刻情况将会怎样呢？与我们已经研究过的双态系统相类似，我们认为这种系统的哈密顿方程组应当由下面这样的方程式组成：

$$i\hbar \frac{dC_n(t)}{dt} = E_0 C_n(t) - AC_{n+1}(t) - AC_{n-1}(t). \quad (13.3)$$

右边的第一个系数 E_0 的物理意义是电子如果不能离开一个原子所具有的能量。(我们把什么叫作 E_0 是无所谓的，正像我们已经见到过多次，它只不过表示我们对零点能的选择。)第二项表示电子在单位时间从第 $(n+1)$ 个陷阱漏到第 n 个陷阱的振幅，最后一项是从 $(n-1)$ 个陷阱漏入的振幅。像往常那样，我们假设 A 是常数(不依赖于 t)。

为了完全描写任意状态 $|\psi\rangle$ 的行为，对每一个振幅 C_n 都要有像式(13.3)那样的一个方程。因为我们要考虑具有大量原子的晶体，我们假设有无限多数目状态——原子在两个方向上都延伸到无限。(对于有限的情况，我们必须特别注意在端点所发生的过程。)如果基础态的数目 N 是无限大，那么全部哈密顿方程的数目也是无限大！我们只写下典型的例子：

$$\begin{array}{c} \vdots & \vdots \\ i\hbar \frac{dC_{n-1}}{dt} & = E_0 C_{n-1} - AC_{n-2} - AC_n, \\ i\hbar \frac{dC_n}{dt} & = E_0 C_n - AC_{n-1} - AC_{n+1}, \\ i\hbar \frac{dC_{n+1}}{dt} & = E_0 C_{n+1} - AC_n - AC_{n+2}, \\ \vdots & \vdots \end{array} \quad (13.4)$$

§13-2 确定能量的状态

我们可以研究许多有关晶格中电子的问题，但我们首先来试试求解有确定的能量的状态。正像我们在早先几章中所看到的，这意味着我们必须找到一种情况，如果振幅要随时间改变的话全都要以同样的频率改变。我们预料解的形式为：

$$C_n = a_n e^{-i\omega t/\hbar}, \quad (13.5)$$

复数 a_n 是发现电子在第 n 个原子的振幅不随时间改变的部分。假如我们把这个试解代入方程式(13.4)中去试一下，所得结果为

$$Ea_n = E_0 a_n - Aa_{n+1} - Aa_{n-1}. \quad (13.6)$$

对于无限个未知数 a_n ，我们有无限个这样的方程式——吓坏人的。

所有我们必须做的是取行列式……可是且慢！当有 2, 3 或 4 个方程的时候采用行列式是个好方法。但如果有大量——或无限多——的方程式，行列式就不很方便了。我们最好试试看直接解这些方程式。首先让我们按照其位置来标记这些原子，我们说原子 n 在 x_n 处，原子 $n+1$ 在 x_{n+1} 处。假如原子的间距等于 b ——如图 13-1 中那样——我们就有 $x_{n+1} = x_n + b$ 。把原点选在第零号原子上，我们有 $x_0 = nb$ 。我们可以把式(13.5)重新写成：

$$C_n = a(x_n) e^{-ikx_n}, \quad (13.7)$$

方程式(13.6)就变成

$$Ea(x_n) = E_0 a(x_n) - Aa(x_{n+1}) - Aa(x_{n-1}). \quad (13.8)$$

或者，利用 $x_{n+1} = x_n + b$ 这个事实，我们也可写下

$$Ea(x_n) = E_0 a(x_n) - Aa(x_n + b) - Aa(x_n - b). \quad (13.9)$$

这个方程与微分方程有点类似。它告诉我们一个量 $a(x)$ 在某一点 (x_n) 的数值和某些相邻点 $(x_n \pm b)$ 处的同一物理量的关系。(微分方程把一个函数在某一点上的数值和无限靠近的点上的数值联系了起来。)或许我们经常用来解微分方程的方法在这里也有效，让我们来试试看。

常系数线性微分方程总能以指数函数作为解。这里我们可以试试同样的解法，我们取试解

$$a(x_n) = e^{ikx_n}. \quad (13.10)$$

于是方程式(13.9)变成

$$Ee^{ikx_n} = E_0 e^{ikx_n} - Ae^{ik(x_n+b)} - Ae^{ik(x_n-b)}. \quad (13.11)$$

我们现在可以消去公因子 e^{ikx_n} ，得到

$$E = E_0 - Ae^{ikb} - Ae^{-ikb}. \quad (13.12)$$

最后两项正好等于 $(2A\cos kb)$ ，所以

$$E = E_0 - 2A\cos kb. \quad (13.13)$$

我们发现任意选择一个常数 k 都可以得到一个解，其能量由上式决定。有许多依赖于 k 的可能的能量，并且每一个 k 对应于一个不同的解。一共有无穷数目的解——这不足为奇，因为我们开始时就有无穷数目的基础态。

我们来看一看这些解意味着什么。对于每一个 k ， a 由式(13.10)给出。于是振幅 C_n 为

$$C_n = e^{ikx_n} e^{-(i/E)b}, \quad (13.14)$$

这里你们应当记住按照式(13.13)能量 E 也依赖于 k 。振幅与空间有关的部分是 e^{ikx_n} 。当我们从一个原子走到下一个原子时振幅随之振荡。

我们说在空间中，振幅就像复数振荡——在每一个原子处，振幅的大小都相同，但是在某一给定的时刻的相位则是一个原子比下一个原子超前一个量(ikb)。像我们在图 13-2 中所画的那样，用竖线来表示每一个原子处振幅的实部，这样我们就可以形象地描绘发生的过

程。显然,这些竖线的包络线(如虚线所示)是余弦曲线。 C_n 的虚部也是振荡函数,但相位移动了 90° ,所以绝对值的平方(就是实数部分和虚数部分的平方和)对所有的 C 都相等。

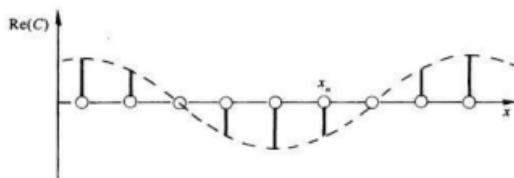


图 13-2 C_n 实部随 x_n 的变化

这样,如果我们选定一个 k ,我们就得到一个具有特定能量 E 的定态。而对于任一个这样的状态,电子在每一个原子处都同样可能被发现——对于任何一个原子都不偏爱。对于不同的原子只有相位的不同。随着时间的流逝相位也在改变。由式(13.14),实数部分和虚数部分以波的形式沿着晶体传播——即表达式

$$e^{[ikx_n - (E/\hbar)t]} \quad (13.15)$$

的实数和虚数部分。此波可向正 x 或负 x 传播,取决于我们所选择的 k 的符号。

要注意,我们一直把试解式(13.10)中的数字 k 假设为一个实数。现在我们可以看出,为什么对于无限长的一串原子必须如此。假定 k 是虚数,譬如说 ik' 。那么振幅 C_n 就成为 e^{ix_n} ,这意味着当我们向正 x 方向——或者负 x 方向,如果 k' 为负数——前进时振幅会变得越来越大。假如我们处理的是一串有终点的原子链,这样的解倒是很好,但对于无限的原子链,它就不是有物理意义的解了。它会给出无限大的振幅——因此就是无限大的概率——它不可能是实际情况的描述,以后我们将会看到虚数 k 真的是有意义的例子。

式(13.13)所给出的能量 E 和波数 k 之间的关系画在图 13-3 上。从图上可以看出,能量可从 $k=0$ 处的 $(E_0 - 2A)$ 变化到 $k=\pm\pi/b$ 处的 $(E_0 + 2A)$ 。这个图是对正的 A 画的,如果 A 是负数,曲线只要颠倒过来,但范围仍相同。意味深长的结果是,在一定范围或能“带”内的任何能量都是可能的,但是不可能具有其他能量。按照我们的假设,如果电子在晶体内处于定态,除了这个带内的数值外不可能具有其他的能量值。

按照式(13.13),最小的 k 对应于低能态—— $E \approx (E_0 - 2A)$ 。当 k 的数值(向正值或负值方向)增加,能量先是增加,然后在 $k = \pm\pi/b$ 处达到最大值,如图 13-3 所示。当 k 大于 π/b 时,能量开始减小。但我们并不真正需要考虑 k 的这些数值,因为它们并不给出新的状态——它们只是重复我们在较小的 k 值已经得出的那些状态。用下面的方法我们可以看出这一点。考虑 $k=0$ 的最低能量状态。系数 $a(x_n)$ 对所有的 x_n 都是相同的。现在对 $k=2\pi/b$ 我们将得到同

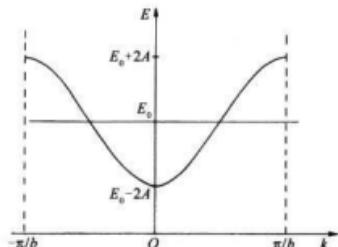


图 13-3 作为参数 k 的函数的定态能量

样的能量。由方程式(13.10)我们有

$$a(x_n) = e^{i(2\pi/b)x_n}.$$

然而,取 x_0 为原点,我们令 $x_n = nb$,于是 $a(x_n)$ 变成

$$a(x_n) = e^{inx} = 1.$$

用这些 $a(x_n)$ 描写的状态在物理上和 $k = 0$ 状态完全相同。它并不代表不同的解。

作为另一个例子,假定 k 为 $-\pi/4b$ 。 $a(x_n)$ 的实部如图 13-4 中曲线 1 所表示的那样变化。如果 k 增大为 7 倍 ($k = 7\pi/4b$), $a(x_n)$ 的实部就像图中曲线 2 那样变化。(当然,整个余弦曲线并无任何意义,有意义的是它们在 x_n 点的数值。曲线只是帮助你了解事情是怎样的。)你看到了 k 的两个值在所有的 x_n 处都给出同样的振幅。

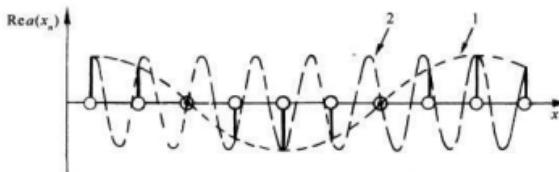


图 13-4 描写同一物理情况的 k 的两个数值,曲线 1 是 $k = -\pi/4b$,曲线 2 是 $k = 7\pi/4b$

结论是,我们只要取某个有限范围内的 k 就得到问题的所有可能的解。我们取 $-\pi/b$ 到 $+\pi/b$ 这一范围——图 13-3 所示的范围。在这个范围内定态能量随 k 的大小的增加而增加。

你们会注意到一个附带的问题。假如电子不是只能以振幅 iA/\hbar 跳到最靠近的原子,并且还可能以另外某个振幅 iB/\hbar 直接跳到下一个近邻原子。你们会发现也可以把解写成 $a_n = e^{ikx_n}$ 的形式——这种形式的解是普遍的。你们还会发现具有波数 k 的定态具有能量 ($E_0 - 2A \cos kb - 2B \cos 2kb$)。这表明 E 对 k 的曲线的形状不是普遍的,而取决于问题的具体假设。它并不一定是余弦曲线——甚至不一定对某一水平线对称。然而,曲线在一 π/b 到 π/b 的区间外一定重复它在这一区间内的形状,所以你不必为其他的 k 值操心。

让我们更为仔细地考察一下 k 值很小的情形——即当振幅从一个 x_n 到下一个的变化十分缓慢的情形。假定我们通过定义 $E_0 = 2A$ 来选择能量的零点,那么图 13-3 中曲线的最小值就在能量为零处。对于足够小的 k ,我们可以把 $\cos kb$ 写成

$$\cos kb \approx 1 - \frac{k^2 b^2}{2},$$

式(13.13)的能量就变成

$$E = Ak^2 b^2. \quad (13.16)$$

我们得到状态的能量正比于波数的平方,这波数描写振幅 C_n 的空间变化。

§13-3 与时间有关的状态

在这一节中我们想较为详细地讨论一下一维晶格中的状态的行为。假如一个电子在

x_* 处的振幅是 C_* , 找到它在那儿的概率是 $|C_*|^2$ 。对于式(13.14)描写的定态, 这个概率对所有的 x_* 都相同并且不随时间变化。我们怎样来表示一个具有确定能量的电子处在某一定的区域这样的状态呢? 这个电子在某个地方比在另外的地方更容易被找到。我们可以用好几个像式(13.14)那样但 k 值略微不同——因而能量略微不同——的解的叠加来表示。那么由于各项之间的干涉至少在 $t = 0$ 时振幅 C_* 将随位置而变。就像不同波长的波叠加时产生拍(我们在第1卷第48章已讨论过这一情形)。所以我们可以用中心占优势的 k_0 , 以及 k_0 附近的其他波数构成一个“波包”*。

在我们的定态叠加中, 不同 k 的振幅代表能量稍微不同的状态, 也就是频率稍有不同的状态, 因此总的 C_* 的干涉图样将随时间而变——就会出现“拍”的图样。我们在第1卷第48章中已经看到, 拍的峰 [$|C(x_*)|^2$ 大的地方] 将随时间沿着 x 运动, 它们以我们称之为“群速度”的速率运动。我们发现这个群速度与 k 随频率变化的关系为:

$$v_{\text{eff}} = \frac{d\omega}{dk}, \quad (13.17)$$

这一关系在此同样有效。成为一“团”的电子态——即 C_* 在空间的变化像图 13-5 中的波包那样——沿着我们的一维“晶体”以等于 $d\omega/dk$ 的速率 v 运动, 其中 $\omega = E/\hbar$ 。对能量 E 应用式(13.16), 我们得到(略去 v_{eff} 的下标写成 v)

$$v = \frac{2Ab^2 k}{\hbar}. \quad (13.18)$$

换句话说, 电子以正比于中心 k 值的速率运动。式(13.16)表明这种电子的能量正比于它的速度的平方——它表现像个经典粒子。只要我们在足够大的尺度上进行考察而不去计较精细结构, 我们的量子力学图像就开始给出经典物理那样的结果。事实上, 假如我们由方程式(13.18)解出 k 并把它代入式(13.16), 我们可将能量 E 写成:

$$E = \frac{1}{2}m_{\text{有效}} v^2, \quad (13.19)$$

其中 $m_{\text{有效}}$ 是一个常数。就像经典粒子那样波包中电子的额外“运动能量”取决于速度。常数 $m_{\text{有效}}$ ——称作“有效质量”——由

$$m_{\text{有效}} = \frac{\hbar^2}{2Ab^2} \quad (13.20)$$

给出。并且注意, 我们可以写出

$$m_{\text{有效}} v = \hbar k. \quad (13.21)$$

如果我们把 $m_{\text{有效}} v$ 叫做“动量”, 它和波数 k 的关系就和我们以前讲过的自由粒子的情况一样。

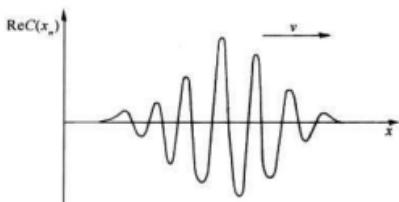


图 13-5 对于能量相近的几个状态的叠加的 $C(x_*)$ 的实部作为 x 的函数
(在图中 x 的标度上间隔 b 非常小。)

* 倘若我们不想使波包太窄。

不要忘记 $m_{\text{有效}}$ 和电子的实际质量并没有什么关系, 它们可以很不相同——虽然在实际晶体中, 它们往往同数量级, 有效质量大约是自由空间电子质量的 2 到 20 倍。

我们现在已经说明了一个奇怪而不可思议的事件——晶体中的电子(例如放进锗里的一个额外电子)怎么能一直穿过晶格运动, 即使它与所有的原子碰撞也能完全自由地流动。它是这样的: 电子的振幅噼叭-噼叭-噼叭地从一个原子跳到下一个原子, 奋力挤过晶体。这就是固体能导电的缘故。

§13-4 三维晶格中的电子

让我们花一点时间来考察一下怎样能把同样的概念应用到三维晶格中电子的运动上。我们发现其结果与一维情况十分类似。假设我们有一原子的长方格, 在 3 个方向上的格点间隔分别为 a, b, c 。(如果你要一个立方格就取 3 个间隔相等。)假设在 x 方向跳到相邻原子上的振幅是 (iA_x/\hbar) , 在 y 方向跳跃是 (iA_y/\hbar) , 在 z 方向跳跃是 (iA_z/\hbar) , 现在我们应该如何描写基础态呢? 就像一维情况那样, 一个基础态代表电子在位于 x, y, z 处的原子上, 这里 (x, y, z) 是晶格的一个格点。把原点选在一个原子上, 所有这些点都位于

$$x = n_x a, \quad y = n_y b, \quad \text{以及} \quad z = n_z c.$$

其中 n_x, n_y, n_z 是任意 3 个整数。现在我们用 x, y 和 z 而不用下标来表示这种点, 把它们理解为只取晶格上的数值。于是, 基础态由符号 | 电子在 x, y, z 来表示, 在某一状态 $|\psi\rangle$ 的电子处在此基础态中的振幅 $C(x, y, z) = \langle \text{电子在 } x, y, z | \psi \rangle$ 。

像以前一样, 振幅 $C(x, y, z)$ 可以随时间改变。按照我们的假设, 哈密顿方程应为:

$$\begin{aligned} i\hbar \frac{dC(x, y, z)}{dt} &= E_0 C(x, y, z) - A_x C(x+a, y, z) - A_z C(x-a, y, z) \\ &- A_y C(x, y+b, z) - A_z C(x, y-b, z) \\ &- A_x C(x, y, z+c) - A_z C(x, y, z-c). \end{aligned} \quad (13.22)$$

它看起来相当长, 但你能理解其中各项是怎样来的。

我们可以再来试求这样的定态, 其中所有的 C 都以同样方式随时间变化。其解仍是指数式:

$$C(x, y, z) = e^{-iE_0 t/\hbar} e^{ik_x x + ik_y y + ik_z z}. \quad (13.23)$$

如果把这个式子代入式(13.22), 你就会看出它是合适的解, 如果能量 E 以下面的方式和 k_x, k_y 和 k_z 相联系就可以了:

$$E = E_0 - 2A_x \cos k_x a - 2A_y \cos k_y b - 2A_z \cos k_z c. \quad (13.24)$$

现在能量依赖于 3 个波数 k_x, k_y, k_z 。顺便提一下, 它是三维矢量 \mathbf{k} 的分量。事实上, 我们可以用矢量记法来表示式(13.23):

$$C(x, y, z) = e^{-iE_0 t/\hbar} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (13.25)$$

振幅的变化就像在 \mathbf{k} 的方向上运动的三维复平面波, 具有波数 $k = (k_x^2 + k_y^2 + k_z^2)^{1/2}$ 。

与这些定态相联系的能量按式(13.24)给出的复杂方式依赖于 \mathbf{k} 的 3 个分量。 E 随 \mathbf{k}

变化的性质依赖于 A_x , A_y 和 A_z 的相对的符号和大小。如果这 3 个数都是正的, 并且我们只对小的 k 值感兴趣, 它们的关系就比较简单。

像以前得出式(13.16)时所做的那样, 我们把余弦展开, 就能得到

$$E = E_{\min} + A_x a^2 k_x^2 + A_y b^2 k_y^2 + A_z c^2 k_z^2. \quad (13.26)$$

对于格点间隔为 a 的简单立方格, 我们认为 A_x , A_y 和 A_z 都相等——譬如说正好都是 A ——我们就有

$$E = E_{\min} + A a^2 (k_x^2 + k_y^2 + k_z^2)$$

或

$$E = E_{\min} + A a^2 k^2, \quad (13.27)$$

这恰巧和式(13.16)相同。按照这里所采用的论据, 我们断定: 三维的电子波包(由近乎相等的许多能量的状态叠加而成)也像具有某一有效质量的经典粒子一样运动。

在对称性较立方面为低的晶体中(或者即使在立方晶体中但每个原子上的电子状态不对称)3 个系数 A_x , A_y 和 A_z 是不同的。那末, 电子的局限在一个小范围内的“有效质量”依赖于它的运动方向。例如, 它在 x 方向运动和在 y 方向运动就可能有不同的惯性。(这种情形的细节常用所谓“有效质量张量”来描写。)

§13-5 晶格中的其他状态

按照式(13.24), 我们所讨论的电子状态只能具有某一能“带”中的能量, 其能量范围从最小能量

$$E_0 - 2(A_x + A_y + A_z)$$

到最大能量

$$E_0 + 2(A_x + A_y + A_z).$$

其他能量也是可能的, 但它们属于另一级电子状态。对于我们已经描述过的状态, 我们设想的基础态是电子位于在某种特定状态, 例如最低能量状态的晶体中的一个原子上。

假定在虚空空间中有一个原子, 加上一个电子就构成一个离子, 这个离子可以按多种方式形成。电子的加入可以构成最低能量的状态, 也可以构成离子的其他各种可能“激发态”, 每一个激发态的能量都高于最低能量。同样的情形在晶体中也会发生。让我们假定上面选定的能量 E_0 相当于离子处于最低的可能能量的基础态的能量。我们也可以设想一组新的基础态, 其中电子以一种不同的方式位于第 n 个原子附近——即离子的一个激发态——所以现在能量 E_0 较前高得多。像以前一样, 电子有某一振幅 A (与前面的不同)从一个原子的激发态跳到相邻原子的同一激发态。整个分析过程和以前一样, 我们找到中心能量较高的可能的能带。一般地说, 可以有许多这样的能带, 每一能带相当于一个不同的激发能级。

也有另外一些可能性。电子也可能具有一些振幅从一个原子的激发态跳到相邻原子的非激发态。(这称为能带间的相互作用。)当你计入越来越多的能带, 加进越来越多的可能的状态之间的漏泄系数, 其数学理论就变得越来越复杂。然而, 这里并没有提出新的概念, 方程式的建立仍和我们在简单的例子中所做的一样。

我们还应当指出,关于出现在上述理论中的各个系数,诸如振幅 A 等,没有更多的可说了。一般说来,这些系数是很难计算的,所以在实际情况中理论上关于这些参数所知极少,对于任何特定的实际情况,我们只能通过实验测定数值。

还有另一些情况,其中的物理和数学与我们对晶体中运动的电子所得出的几乎完全一样,但其中运动的“客体”却完全不同。例如,假定我们开始时讨论的晶体——或者更确切地说线型晶格——是排成一直线的中性原子,每一个原子都有一个束缚得很松的外层电子。设想我们去掉一个电子,哪一个原子失去了它的电子呢?用 C_n 表示从位于 x_n 的原子上失去电子的振幅。一般说来,相邻原子——譬如说第 $(n-1)$ 个原子——上的电子具有振幅 iA/\hbar 跳到第 n 个原子上而留下一个失去电子的第 $(n-1)$ 个原子。这等于说“失去的电子”具有振幅 A 从第 n 个原子跳到第 $(n-1)$ 个原子。你们可以看到方程式将会完全相同——当然, A 的数值不必要和我们前面的相同。我们又会得到能级的、有关式(13.18)表示的群速度通过晶体运动的概率“波”,以及关于有效质量等等同样的公式。只不过现在的波描写失去的电子——就是所谓的“空穴”——的行为。所以“空穴”就像具有确定质量 $m_{\text{有效}}$ 的粒子那样行动。你们可以看出这种粒子表现出带有正电荷。关于这种空穴,在下一章我们还将进一步讨论。

作为另一个例子,我们考虑在排成一线的相同中性原子,其中有一个原子已经处在激发态——就是说它具有比正常的基态更高的能量。令 C_n 是第 n 个原子激发的振幅。它能和邻近的原子作用,把过多的能量移交给邻近的原子而回到基态。把这种过程的振幅叫做 iA/\hbar 。你可以看出有关的数学和以前的完全相同。现在运动的客体被称为“激子”。它的行为就像一个中性“粒子”,带着激发能量穿过晶体运动。这类运动可以发生在某些生物学的过程中,像视觉或光合作用。我们猜想在视网膜中吸收了一个光子产生一个“激子”,它穿过某种周期性构造(如我们在第 1 卷第 36 章中描写的视杆细胞中的层状结构图 36-5)运动,并且被积聚到某个特殊位置,能量在此地被用来引起化学反应。

§13-6 在不完整的晶格上的散射

我们现来考虑在不完美的晶体中的单个电子。我们以前的分析表明,完美晶体具有极好的导电性——电子可以无摩擦地滑过晶体,就像在真空中一样。能使不停地运动着的电子停下来最重要的因素是晶体中的不完整性或不规则性。作为一个例子,如果晶体中某一个地方少掉一个原子,或者如果某一个人在某个原子的位置上摆错了一个原子,从而比之于其他的原子位置来这里的情况就不同了。譬如说能量 E_0 或振幅 A 就会不同。那么我们怎样来描写所发生的事呢?

为明确起见,我们回到一维的情况,并且假定第“零”号原子是一个“杂质”原子。它具有和其他原子不同的 E_0 值,我们令这个能量是 $(E_0 + F)$ 。会发生些什么呢?当电子到达“零”号原子处时,电子有向后散射的概率。假设一个波包向前运动,当它来到一个情况稍有不同的地方,波包的一部分继续前进,而另一部分则被反弹回去。这一情形很难用波包来分析,因为每样东西都随时间变化。用定态解来处理就方便多了。所以我们用定态来处理,我们将发现定态可以由透射和反射两部分连续波构成。在三维空间中,我们把反射部分称为散射波,因为它可向各个方向散开。

我们从和式(13.6)类似的一组方程式开始,只是其中 $n = 0$ 的方程式和所有其余的方程都不同。对于 $n = -2, -1, 0, +1$ 和 $+2$ 的 5 个方程式是这样的:

$$\begin{array}{ll} \vdots & \vdots \\ Ea_{-2} = E_0a_{-2} - Aa_{-1} - Aa_{-1}, & \\ Ea_{-1} = E_0a_{-1} - Aa_0 - Aa_{-2}, & \\ Ea_0 = (E_0 + F)a_0 - Aa_1 - Aa_{-1}, & (13.28) \\ Ea_1 = E_0a_1 - Aa_2 - Aa_0, & \\ Ea_2 = E_0a_2 - Aa_3 - Aa_1, & \\ \vdots & \vdots \end{array}$$

当然,所有其他方程式的 $|n|$ 都大于 2。它们看起来都与式(13.6)相同。

对于一般的情况,我们对电子跳向和跳离“零”号原子的振幅实际上应该用不同的 A ,但从所有的 A 都相等的简化例子中还是能够看到所发生过程的主要面貌。

式(13.10)仍可作为所有的方程式的解,只是“零”号原子的方程式除外——它对这个方程不适用。我们需要一个不同的解,我们能用下面的方法把它打造出来。式(13.10)表示沿正 x 方向进行的波。沿负 x 方向进行的波也是同样好的解,它可以写成:

$$a(x_n) = e^{-ikx_n}.$$

式(13.6)的最普遍的解将是向前和向后的波的组合,即

$$a_n = \alpha e^{ikx_n} + \beta e^{-ikx_n}, \quad (13.29)$$

这个解表示一振幅为 α 沿 $+x$ 方向传播的复波和一振幅为 β 沿 $-x$ 方向传播的波。

现在看一看我们新的问题的一组方程式——在式(13.28)中的以及所有其他原子的那些方程。包含 $n \leq 1$ 的 a_n 的方程式都能被式(13.29)满足,只要 k 和 E 以及和晶格间隔 b 的关系满足下述条件:

$$E = E_0 - 2A \cos kb, \quad (13.30)$$

其物理意义是,振幅为 α 的“入射”波从左边趋向“零”号原子(“散射原子”),振幅为 β 的“散射”或“反射”波向左边退回去。假如我们令入射波的振幅 α 等于 1 并不会失去任何普遍性。而一般说来 β 是一个复数。

关于 $n \geq 1$ 的 a_n 的解我们也可同样处理。但系数可能是不同的,所以我们有

$$a_n = \gamma e^{ikx_n} + \delta e^{-ikx_n}, \text{ 对于 } n \geq 1, \quad (13.31)$$

其中 γ 是向右传播的波的振幅, δ 是从右边来的波。我们考虑这种物理情况,波起初只从左边发出,并且从散射原子——或者说杂质原子——后面出射的只有“透射”波。我们试求 $\delta = 0$ 的解。我们肯定能用下面的试解使除了式

(13.28)中间的 3 个式子以外的所有 a_n 的方程式都满足

$$a_n (\text{对于 } n < 0) = e^{ikx_n} + \beta e^{-ikx_n}, \quad (13.32)$$

$$a_n (\text{对于 } n > 0) = \gamma e^{ikx_n}.$$

我们所谈论的情形画在图 13-6 中。

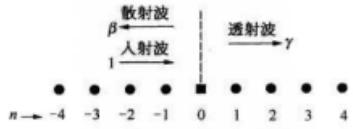


图 13-6 在 $n = 0$ 处有一个“杂质”原子的一维晶格中的波

将式(13.32)中的公式用于 a_{-1} 和 a_{+1} , 我们可以从式(13.28)中间的 3 个方程解出 a_0 和两个系数 β 和 γ 。从而我们就得到完整的解。设 $x_n = nb$, 我们需要解 3 个方程式:

$$\begin{aligned}(E - E_0) |e^{ik(-b)} + \beta e^{-ik(-b)}| &= -A |a_0 + e^{ik(-2b)} + \beta e^{-ik(-2b)}|, \\ (E - E_0 - F)a_0 &= -A |\gamma e^{ibb} + e^{ik(-b)} + \beta e^{-ik(-b)}|, \\ (E - E_0)\gamma e^{ibb} &= -A |\gamma e^{ik(2b)} + a_0|.\end{aligned}\quad (13.33)$$

记住 E 是通过式(13.30)用 k 来表示的。如果你把 E 的这个值代入方程式, 并且记住 $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$, 从第一个方程式得到

$$a_0 = 1 + \beta, \quad (13.34)$$

以及从第三个方程式得到

$$a_0 = \gamma, \quad (13.35)$$

要上面两个式子一致, 必须:

$$\gamma = 1 + \beta. \quad (13.36)$$

这个式子表明, 透射波(γ)正好等于原来的入射波(1), 加上反射波(β)。这并不总是正确的, 只是对一个原子的散射碰巧如此。如果有一群杂质原子, 加到向前传播的波上的数量就不一定和反射波相同。

我们可以从式(13.33)中间的方程式得到反射波的振幅 β , 我们求得

$$\beta = \frac{-F}{F - 2iA \sin kb}. \quad (13.37)$$

我们得到了具有一个异常原子的晶格的完全解。

你可能会感到奇怪, 从式(13.34)中表示出, 透射波怎么会比入射波“更多”。但是记住 β 和 γ 是复数, 并且波中的粒子数目(更确切地说是找到粒子的概率)正比于振幅绝对值的平方。事实上, 仅当

$$|\beta|^2 + |\gamma|^2 = 1 \quad (13.38)$$

时才会有“电子守恒”。你可以证明我们的解满足这一关系。

§13-7 被晶格的不完整性陷俘

假如 F 是负数就会出现另一个有趣的情况。如果电子的能量在杂质原子(在 $n = 0$)中比在其他任何地方都低, 那么电子可能被这个原子捕获。这就是说, 如果 $(E_0 + F)$ 低于能带底 $(E_0 - 2A)$, 那么电子可能被“陷俘”在 $E < E_0 - 2A$ 的状态下。按照我们现在为止所讨论过的内容是得不出这样的解的。然而, 如果我们在所取的试解式(13.10)中允许 k 为虚数, 我们就能求得这一解。设 $k = ix$ 。同样, 对于 $n < 0$ 和 $n > 0$ 仍可有不同的解。对 $n < 0$ 的一个可能解是

$$a_n(\text{对于 } n < 0) = ce^{+inx}. \quad (13.39)$$

我们应当在指数上取正号, 否则 n 为大的负值时, 振幅将会变为无限大。同样, 对于 $n > 0$

的一个可能解将是

$$a_s(\text{对于 } n > 0) = c' e^{-\kappa x_s}. \quad (13.40)$$

如果我们把这些试解代入式(13.28), 假如

$$E = E_0 - A(e^{\kappa b} + e^{-\kappa b}), \quad (13.41)$$

则除了中间的3个式子, 其余的式子都可满足。因为两个指数项的和总是大于2, 这个能量在正常能带的下面, 而这正是我们所要求的。如果 $a_0 = c = c'$, 并按下式来选定 κ :

$$A(e^{\kappa b} - e^{-\kappa b}) = -F, \quad (13.42)$$

式(13.28)中余下的3个方程式皆可满足。把这个方程式和式(13.41)结合起来, 我们就能得到被陷俘的电子的能量, 我们得到

$$E = E_0 - \sqrt{4A^2 + F^2}, \quad (13.43)$$

陷俘电子有单值的能量——在稍低于导带的地方。

注意, 式(13.39)和式(13.40)所给出的振幅并不表示陷俘电子正好位于杂质原子上。在附近的原子中找到电子的概率正比于这些振幅的平方。对于特别选择的一组参数, 它可能像图13-7中的长条图那样变化。在杂质原子上找到电子的概率最大。对于附近的原子, 概率随着离开杂质原子的距离的增加指数式地下降。这是“势垒穿透”的又一个例子。从经典物理学的观点来看, 电子并没有足够的能量能从陷俘中心的能量“空穴”中跑出来。但是按照量子力学它却可以泄漏跑出一小段距离来。

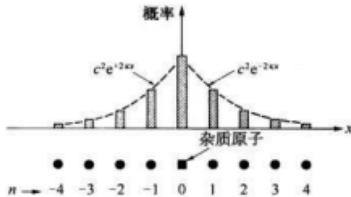


图13-7 在陷俘杂质原子附近的原子上找到陷俘电子的相对概率

§13-8 散射振幅和束缚态

最后, 我们的例子可以用来说明当前在高能粒子物理学中一个很有用的观点。这牵涉到散射振幅和束缚态之间的关系问题。假定我们已经发现了——通过实验和理论分析—— π 介子被质子散射的方式。于是一个新的粒子被发现了, 并且有人怀疑这是否只是 π 介子和质子结合在一起形成的某种束缚态(和电子束缚在质子周围组成氢原子的方式相似)。所谓束缚态我们指的是一种组合, 具有比两个自由粒子更低的能量。

有一个普遍的理论, 它告诉我们具有这样的能量将存在束缚态: 如果用代数方法外推(数学名称叫“解析延拓”)到所允许能带的能量范围以外时散射振幅变为无穷大。

这一理论的物理思路如下: 绑缚态是这样一种状态, 它只有束缚在一定位置上的波而不是由外来的波激起的, 它只是自身存在着。所谓“散射”或产生的波和被“送入”的波的相对比例是无限大。我们可以在我门这个例子中检验这个观念。让我们直接用被散射粒子的能量 E (而不用 k)来写出表示散射振幅的表达式(13.37)。因为式(13.30)可以写成:

$$2A \sin kb = \sqrt{4A^2 - (E - E_0)^2},$$

散射振幅为

$$\beta = \frac{-F}{F - i\sqrt{4A^2 - (E - E_0)^2}}, \quad (13.44)$$

根据我们推导过程的要求,这个方程式只适用于真实状态——那些能量处在能带 $E = E_0 \pm 2A$ 之中的状态。但是假定我们忘记了这个事实并且把公式扩展到 $|E - E_0| > 2A$ 的“非物理的”能量区域。对于这些非物理的区域我们可以写出*:

$$\sqrt{4A^2 - (E - E_0)^2} = i\sqrt{(E - E_0)^2 - 4A^2}.$$

于是“散射振幅”(不管它意味着什么)就是:

$$\beta = \frac{-F}{F + \sqrt{(E - E_0)^2 - 4A^2}}. \quad (13.45)$$

现在我们要问:有没有使得 β 变成无限大的能量(即在这个能量对 β 的表达式有一个“极点”)? 有的,只要 F 是负数,当

$$(E - E_0)^2 - 4A^2 = F^2$$

时,或者

$$E = E_0 \pm \sqrt{4A^2 + F^2},$$

式(13.45)的分母就是零。由上式的负号所给出的能量正是我们在式(13.43)中得出的陷阱能量。

正号表示什么呢? 由它给出的能量高于允许的能带。的确,那里有另一个束缚态,我们在求解式(13.28)时没有考虑到它。我们把求这个束缚态的能量和振幅 a_s 作为一个智力测验留给你们。

对于探求当前新的奇异粒子的实验观察的解释,散射和束缚态之间的关系提供了一个最有用的线索。

* 这里根号前符号的选择是与式(13.39)和式(13.40)中允许的 κ 的符号有关的技术问题。我们在这儿不作讨论。

第 14 章 半 导 体

§ 14-1 半导体中的电子和空穴*

近年来一个引人注目的发展是把固体科学应用于像晶体管之类的电器件的技术发展。对半导体的研究导致它们的有用性质以及大量的实际应用的发现。这个领域的变化如此之快以至于我们今天告诉你们的东西到了明年就可能不再正确了。它肯定是不完善的。非常清楚,由于对这些材料的不断研究,随着时间的推移将会出现许多新的和更惊人的事物。为了学习这一卷中以后各章的内容,你们不一定非要弄懂这一章不可。但是,看到所学的内容中至少有一部分与现实世界有关时你们会感到兴趣的。

已知的半导体种类很多,但我们将着重于现在有最大技术应用价值的半导体。这些也是了解得最清楚的,了解了它们,对其他许多半导体也将得到一定的了解。今天最常用的半导体物质是硅和锗。这些元素结晶形成金刚石晶格,这是一种立方结构,其中一个原子和最靠近的 4 个原子形成四面体键。虽然在室温下它们多少有点导电,但在很低的温度时——近于绝对零度——它们是绝缘体,它们不是金属;它们被称作半导体。

假如我们以某种方法在处于低温的硅或锗晶体中放入一个额外的电子,我们就遇到上一章中所描写的那种情况。这个电子会在晶体中游荡,从一个原子跳到下一个原子。实际上我们只讨论过长方格中的电子,而对于真实的硅或锗的晶格,方程式多少有点不同。不过,所有的基本特征都可用长方格的结果来说明。

正如我们在第 13 章中曾经看到的,这些电子只能具有某一定能带——叫做导带——中的能量。在这个带中,能量和概率幅 C 的波数 k 的关系(见式 13.24)是

$$E = E_0 - 2A_x \cos k_x a - 2A_y \cos k_y b - 2A_z \cos k_z c. \quad (14.1)$$

式中的 A 是在 x , y 和 z 方向上跳跃的振幅, a , b 和 c 是这些方向上的晶格间隔。

当能量近于能带底时,可以取式(14.1)的近似(参见 § 13-4):

$$E = E_{\min} + A_x a^2 k_x^2 + A_y b^2 k_y^2 + A_z c^2 k_z^2. \quad (14.2)$$

如果我们考虑电子在某个特殊方向上运动,使得 k 的 3 个分量总有相同的比例。能量是波数的二次函数——和电子的动量的关系一样。我们可以写成:

$$E = E_{\min} + \alpha k^2, \quad (14.3)$$

其中 α 是某一常数。我们可以画出 E 对 k 的函数图。如图 14-1 所示。我们把这种图称为

* 参考: C. Kittel, Introduction to Solid State Physics, John Wiley and Sons. Inc., New York. 2nd ed., 1956 chapters 13, 14 and 18. ——译者注

“能量图”。一个处于特定能量和动量状态的电子可以用图上的一点(如 S 点)来表示。

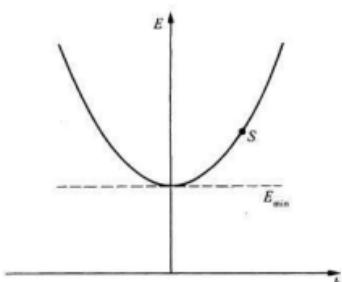


图 14-1 电子在绝缘晶体中的能量图

正像我们在第 13 章中也曾讲过的,如果我们从中性绝缘体上取走一个电子,我们就有类似的情况。于是,一个电子可以从附近的原子上跳过来填充这个“空穴”,而在原来的原子处留下了另一个“空穴”。我们可以通过写出在任一特定原子处找到空穴的振幅,以及说明空穴可以从一个原子跳到下一个原子来描写它的行为。(显然,空穴从原子 a 跳到原子 b 的振幅 A 正好和原子 b 上的一个电子跳进原子 a 的空穴的振幅相同。)对于空穴和对于额外的电子,它们的数学式子完全相同。我们再次得出空穴的能量和它的波数的关系由和式(14.1)或式(14.2)相同的方程式表示出来。当然振幅 A_x , A_y 和 A_z 的数值是不同的。空穴的能量和它的概率幅的波数有关。它的能量处在有限能带中。在能带底附近,它的能量随波数——或动量——的平方变化,如图 14-1 所示。按照 § 13-3 的论证,我们将发现空穴的行为也像具有一定的有效质量的经典粒子一样,——只是在非立方晶体中质量依赖于运动的方向。所以空穴就像一个在晶体中运动的带正电的粒子。空穴粒子的电荷是正的。因为它在于失去电子的位置;当它在一个方向上运动时,实际上是电子在相反的方向上运动。

假如我们在中性晶体中放进几个电子,它们将像低气压气体中的原子那样四处活动。假如电子不很多,它们之间的相互作用不十分重要。如果我们在晶体上加上电场,电子就开始运动,于是有电流流过。最终电子都被拉向晶体的一边,如果那里有一个金属电极,电子都要被电极收集,而晶体又变成中性。

同样我们可以把很多空穴放进晶体。在没有外加电场时,它们将随机地到处活动。在电场作用下它们要向负端流动并被“收集”——实际上发生的是它们被自金属电极来的电子中和。

晶体中也可以同时存在空穴和电子。如果它们的数目不多,它们将各自独立地运动。加上电场,它们都对电流作出贡献。由于明显的理由,电子被称为负载流子,空穴被称作正载流子。

到现在为止,我们考虑的电子是从外面放进晶体的,或把电子取走造成空穴。我们也可以这样来“创造”一个电子-空穴对:从一个中性原子上取走一个被束缚的电子并把它放到同一晶体中一定距离外的地方。于是我们就要一个自由电子和一个自由空穴,二者都能像我们已经描写过的那样运动。

把一个电子放进状态 S ——我们说“产生”状态 S ——所需的能量就是图 14-2 中所标出的能量 E^- 。它是高于 E_{\min} 的某个能量。“产生”处于某一状态 S' 的一个空穴所需的能量是图 14-3 中的能量 E^+ ,它是比 E_{\min}^+ 大的某个能量,如果现在我们要产生在状态 S 和 S' 的电子空穴对,所需的能量就等于 $E^- + E^+$ 。

电子空穴对的产生是一个共生的过程(我们以后将会看到),所以很多人喜欢把图 14-2 和图 14-3 放在同一张图上——把空穴能量向下画,虽然它实际上是正能量。我们用这种方式把两个图合起来画在图 14-4 中。这种图的好处是产生处于 S 中的电子和处于 S' 中的空穴这一电子-空穴对所需的能量 $E_{\text{gen}} = E^- + E^+$ 正好等于 S 和 S' 间的垂直距离,如图 14-4 所示。产生一个电子-空穴对所需的最小能量称为“隙”能,它等于 $E_{\min}^- + E_{\min}^+$ 。

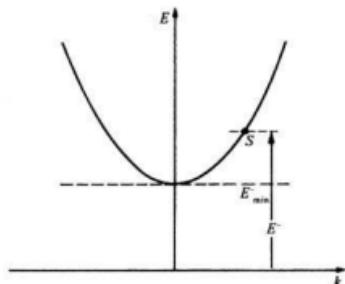
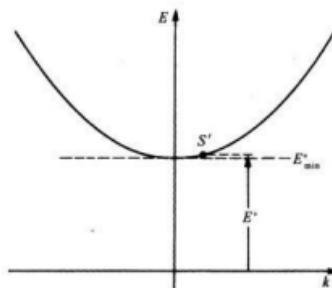
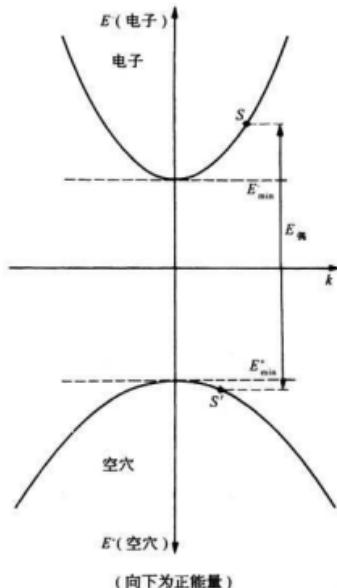
图 14-2 “产生”一个自由电子需要能量 E^- 图 14-3 “产生”一个在 S' 态的空穴需要能量 E^+ 

图 14-4 电子和空穴画在一起的能量图

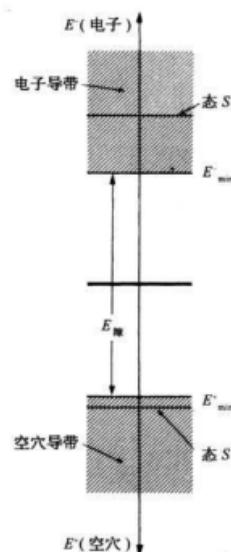


图 14-5 电子和空穴的能级图

有时你会看到一种比较简单的图，称作能级图。当人们对变量 k 不感兴趣时就画这种图。这样的图——如图 14-5 所示——只表示电子和空穴的可能的能量^{*}。

- 在很多书上以不同的方式解释这种能量图。能量标度只是对电子的能量而言。他们想象占据着空穴的电子应具有的能量而不是空穴的能量。这个能量比自由电子的能量低——事实上，所低的数值正好是你在图 14-5 中所看到的。按照这种对能量标度的解释，隙能是使电子从束缚态跃迁到导带所必须给予电子的最低能量。

怎样创造出电子-空穴对呢？有几种方法。例如，吸收光或X射线的光子，如果光子能量高于隙的能量就会产生电子-空穴对。电子-空穴对的产生率正比于光的强度。如果在晶片两边镀上两片电极并加上“偏置”电压，电子和空穴就会被拉向电极。电路中的电流将正比于光的强度。这就是产生光电导现象和光电导管作用的机理。

电子-空穴对也可用高能粒子来产生。当快速运动的带电粒子——例如能量为几十或几百MeV的质子或 π 介子——穿过晶体时，它的电场会把电子从它的束缚态撞击出来产生电子-空穴对。在径迹的每1毫米路程上会发生几十万次这样的事件。粒子通过以后，载流子能够被收集起来，这样就得到一个电脉冲。这就是近来在核物理实验中使用的半导体计数器的作用机理。这类计数器并不一定需要半导体，也可以用晶态绝缘体制造。事实上，第一个这种计数器是用金刚石晶体制造的，它在室温下是绝缘体。如果要使电子和空穴能够自由地运动到电极上而不被陷俘，就需要非常纯的晶体。采用半导体硅和锗是因为能将它们制造成高纯度的适当大小(厘米的线度)的晶体。

迄今为止我们只涉及温度接近绝对零度的半导体晶体。在任何有限的温度下，还有另一种产生电子-空穴对的机理。产生电子-空穴对的能量可由晶体的热能提供，晶体热振动可以把它们的能量转移给电子-空穴对——引起“自发”产生。

单位时间内，像隙能 E_g 那样大的能量聚集到一个原子的位置上的概率正比于 $e^{-E_g/kT}$ ，其中 T 是温度， k 是玻尔兹曼常数(参见第1卷第40章)。在绝对零度附近没有可以觉察的概率，但随着温度的增加，产生电子-空穴对的概率就增加。在任何有限的温度下，对的产生以恒定的速率不断继续下去，负的和正的载流子越来越多。当然，这是不会发生的，因为过了一会儿，电子和空穴会偶然地相遇——电子落进空穴并把多余的能量交给晶格。我们说电子和空穴“湮没”了。在单位时间内容空穴和电子有一定的概率相遇，于是它们互相湮没。

假定单位体积的电子数为 N_e (n 表示负载流子)，正载流子的密度是 N_p 。单位时间内一个电子和一个空穴相遇并湮没的机会正比于乘积 $N_e N_p$ 。在平衡时，这个速率必定等于电子-空穴对的产生速率。你们知道，在平衡时 N_e 和 N_p 的乘积应等于某一常数乘上玻尔兹曼因子：

$$N_e N_p = \text{常数} \cdot e^{-E_g/kT}. \quad (14.4)$$

当我们说常数时，我们指的是近似于常数。更完整的理论——包含有关空穴和电子彼此如何“相遇”更多的细节——表明“常数”稍稍依赖于温度，但对温度的主要依赖是在指数上。

作为一个例子，我们考虑原来是中性的纯净材料。在有限的温度下，你会料到正的和负的载流子的数目相同， $N_e = N_p$ 。它们各自都按 $e^{-E_g/kT}$ 随温度变化。半导体的性质——例如电导率——的许多变化主要决定于指数因子，因为所有其他因子随温度的变化要慢得多。锗的隙能大约是0.72eV，硅是1.1eV。

在室温下 kT 大约是1eV的1/40。在这样的温度下有足够的空穴和电子给出可观的电导率，而在譬如说30K——室温的十分之一——电导率是难以觉察的。金刚石的隙能是6~7eV，在室温下金刚石是好的绝缘体。

§ 14-2 掺杂的半导体

到现在为止我们已经谈论过将额外电子放入理想完整的晶体的晶格中的两种方法。一

种方法是从外源注入电子,另一种方法从一个中性原子上敲出一个束缚电子同时产生一个电子和一个空穴。还可以用别的方法把电子放到晶体的导带里面。假定我们想象锗晶体中一个锗原子被一个砷原子所代替。锗原子是4价,晶体的结构由4个价电子控制。另一方面,砷是5价。我们发现一个砷原子能占据锗晶格中的一个位置(因为它的大小大致合适),但这样一来它必须表现得和4价原子一样——用它的4个价电子形成晶体键,于是还有一个电子剩了下来。这个多余的电子只是很松地束缚着——结合能大约只有一个电子伏特的 $1/100$ 。在室温下电子很容易从晶体的热能中获得这样多的能量,并且独自离开——像自由电子那样在晶格中运动。像砷那样的杂质原子被称作施主位,因为它能放出一个负载流子到晶体中。假如锗晶体是从加进非常少量的砷的熔融金属中生长出来的,砷施主位将分布在整个晶体中,于是在晶体中将建立起一定的负载流子密度。

你们也许认为只要对晶体加上任意小的电场,就会将这些载流子扫出去。然而,事实并非如此,因为晶体内的每个砷原子都带有正电荷。如果晶体要保持中性,负载流子电子的平均密度必须等于施主位的密度。如果你在晶体的两边加上电极,并把它们联结到电池上,就会有电流流动。但当载流电子从一端被扫出,新的传导电子必定从另一端的电极进入,结果传导电子的平均密度保持和施主位的密度差不多相等。

因为施主位带正电荷,当导电电子在晶体中扩散时有被它们俘获的趋势。所以,施主位相当于我们在上一节中讨论过的陷阱。但如果陷阱能量足够小——砷就是这样——在任一时刻被陷阱的载流子数目只是总数的一小部分。要完全理解半导体的行为,人们必须把这种陷阱作用考虑进去。然而,在我们讨论的其余部分中,我们将假定陷阱能量足够低,并且温度足够高,从而所有的施主位都已失去了它们的电子。当然这只是一种近似。

也可以在锗的晶体中掺入一些3价的杂质原子,例如铝。铝原子试图窃取一个额外的电子并表现得像4价原子那样。它从某个邻近的锗原子窃得一个电子,结果成为一个有效价数为4的带负电的原子。当然,当它从锗原子窃得一个电子后,就在那里留下一个空穴,这个空穴能作为正载流子在晶体中游荡。能用这种方法产生空穴的杂质原子称为受主,因为它“接受”一个电子。假如锗或硅晶体是从加进少量杂质铝的熔融材料中生长的,晶体内部就具有一定的内建空穴密度,空穴就相当于正载流子。

当施主或受主杂质加进半导体中后,我们说材料被“掺杂”了。

当具有内建施主杂质的锗晶体处于室温时,由激发感应产生的电子-空穴对以及施主位都贡献出一些传导电子。自然,从这两个来源产生的电子是相等的,在达到平衡的统计过程中起作用是总数 N_s 。如果温度不太低,施主杂质原子提供的负载流子的数目大致等于晶体中的杂质原子的数目。平衡时式(14.4)必定仍旧有效;给定温度条件下,乘积 $N_s N_n$ 是一定的。这意味着,假如我们加入一些使 N_s 增加的施主杂质,正载流子的数目 N_n 必定减少一定数量以使 $N_s N_n$ 不变。如果杂质浓度足够高,负载流子数目 N_n 由施主位的数目决定并且几乎不依赖于温度——所有指数因子上的变化由 N_s 提供,虽然它比 N_n 少得多。在另一些有少量施主杂质的纯净的晶体中,大多数载流子是负载流子,这种材料叫做“n型”半导体。

假如在晶格中加入受主型杂质,一些新的空穴将到处漂移,并与一些由热起伏产生的自由电子发生湮没。这一过程将一直继续到式(14.4)被满足。在平衡条件下,正载流子数目将增加而负载流子数目将减少以保持它们的乘积为常数。正载流子较多的材料称为“p型”半导体。

假如我们把两个电极放在一块半导体晶体上并把它们联结到电势差的电源上去。晶体中将建立起电场。电场会使正的和负的载流子运动,于是有电流流动。我们首先来考虑在 n 型材料中会发生什么过程,这种材料中绝大多数是负载流子。对于这类材料,我们可以不理会其中的空穴,因为它们是如此之少,对电流几乎没有贡献。在理想的晶体中载流子将不受阻碍地运动。然而,在实际晶体中,在有限的温度下——特别是在有一些杂质的晶体中——电子的运动并不是完全自由的,它们不断地发生碰撞,被撞离原来的轨道,即改变它们的动量。这些碰撞就是我们在上一章中谈到的散射,在晶格中任何具有不规则性的地方都会发生这种散射。在 n 型材料中,散射主要就是由产生载流子的施主位引起的。因为传导电子在施主位上有略微不同的能量,概率波就从这些点上散射。然而,即使在完全纯净的晶体中,(在有限的温度下)热振动也要在晶格中引起不规则性。按照经典的观点,我们可以说原子并不是准确地排列在规则的晶格上,在任何瞬间由于热振动都要稍稍离开原位。在第 13 章所介绍的理论中,和每一格点相联系的能量 E 随位置变化而略有变化,所以概率振幅波并不是完全透射而是以不规则的方式被散射。在非常高的温度下,对非常纯净的材料,这种散射可能变得很重要。但在实际的器件所用的大多数掺杂材料中,在杂质原子上的散射起主要作用。我们现在要对这种材料的导电率作一个估算。

当外电场加到 n 型半导体上,每一个负载流子在电场中被加速,它的速度不断增加直到在一个施主位上被散射。这意味着平时具有热能,以随机方式运动着的载流子将沿着电场力线的方向获得一个平均漂移速度,形成电流流过晶体的。与典型的热运动速度相比,漂移速度一般是很小的,所以我们可以通过假定载流子在两次散射之间的平均时间是常数来估算电流。我们设负载流子具有有效电荷 q_n 。在电场 ϵ 中,作用于载流子的力是 $q_n \epsilon$ 。在第 1 卷 § 43-3 中,我们曾计算了在这种情形下的平均漂移速度,并得到漂移速度由 $F\tau/m$ 给出,其中 F 是作用于电荷上的力, τ 是两次碰撞间的平均自由时间, m 是质量。我们应当用上一章算出的有效质量,但因为只要作粗略的估算,我们假定有效质量在各个方向上都是相等的。这里我们把它称作 m_n 。在这个近似下,平均漂移速度是:

$$v_{\text{漂移}} = \frac{q_n \epsilon \tau_n}{m_n} \quad (14.5)$$

知道了漂移速度我们就能求出电流。电流密度 j 等于单位体积内载流子的数目 N_n 乘以平均漂移速度,再乘以每个载流子的电荷。因此这里电流密度是

$$j = N_n v_{\text{漂移}} q_n = \frac{N_n q_n^2 \tau_n \epsilon}{m_n} \quad (14.6)$$

我们看到电流密度正比于电场,这样的半导体材料遵从欧姆定律。 j 和 ϵ 的比例系数,即电导率 σ 是

$$\sigma = \frac{N_n q_n^2 \tau_n}{m_n} \quad (14.7)$$

对于 n 型半导体,电导率相对地不依赖于温度。首先,多数载流子的数目 N_n 主要决定于晶体中施主的密度(只要温度不是那么低使得太多的载流子被陷俘)。其次,碰撞之间的平均时间 τ_n 主要由杂质原子的密度控制,当然它不依赖于温度。

我们可以把完全相同的论证应用于 p 型材料,只要改变一下式(14.7)中的参数的数

值。假如同时存在的负的和正的两种载流子的数目可以相比较，我们必须把两种载流子的贡献加起来。总电导率由下式给出

$$\sigma = \frac{N_s q_s^2 \tau_s}{m_s} + \frac{N_p q_p^2 \tau_p}{m_p}. \quad (14.8)$$

对于非常纯净的材料， N_p 和 N_s 近于相等，它们比掺杂材料中的要小，所以电导率要小一些。可是它们随温度变化得很快(正像我们已经看到过的按照 $e^{-E_{\text{g}}/kT}$ 变化)，所以电导率随温度而极快地变化。

§ 14-3 霍尔效应

只有电子是相对自由的导体的材料中，也会有行为像正粒子的空穴所运载的电流，这确实是一件难以理解的事。因此，我们要描写一个实验，这实验能十分明确地证明电流的载流子绝对肯定时正的。假设我们有一块半导体材料——也可以是金属——我们在上面加上电场，从而在某个方向引起电流，例如水平方向，如图 14-6 所示。现在假设我们在这块材料上再加上磁场，其指向和电流方向成直角，譬如说指向图的里面。运动着的载流子会感受到磁力 $q(v \times B)$ 的作用。由于平均漂移速度不是向右就是向左——取决于载流子上电荷的符号——作用于载流子上的平均磁力不是向上就是向下。不，这是不对的！对于我们所假定的电流和磁场方向，作用于运动电荷上的磁力总是向上的。正电荷在 j 的方向上运动(向右)会受到向上的力。如果电流由负电荷运载，它们向左运动(对于同样符号的传导电流)，于是它们也会受到向上的力。然而，在稳定情况下，载流子并没有向上的运动，因为电流只能从左边流到右边。所发生的过程是，开始时有一些电荷向上流动，在半导体的上表面形成面电荷密度——在晶体的下表面上留下相等但相反的面电荷密度。电荷聚集在上表面和下表面上，直到它们作用于运动着的电荷上的电力正好和(平均的)磁力完全相抵消，因此稳恒电流照旧水平地流动。上表面和下表面上的电荷将产生一个垂直跨过晶体的电势差，可以用高电阻伏特计来测量这电势差，如图 14-7 所示。伏特计记录的电势差的符号取决于形成电流的载流子的电荷的符号。

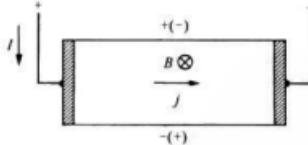


图 14-6 霍尔效应来自作用于载流子的磁力

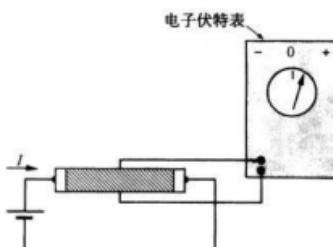


图 14-7 测量霍尔效应

在第一次做此实验时，正像对负的传导电子所预期的那样，人们预期电势差的符号将是负的。因此，当人们发现对于某些材料电势差的符号与预期的相反就感到十分奇怪。这表

明载流子是带正电荷的粒子。从我们关于掺杂半导体的讨论中可以理解, n 型半导体应当产生与负的载流子相应的电势差符号, 而 p 型半导体应当给出相反的电势差, 因为电流是由带正电的空穴负载的。

霍尔效应中电势差的反常符号最初是在金属中而不是在半导体中发现的。曾经假定金属都是电子导电的, 然而, 对于铍却发现其电势差的符号恰好相反。现在了解了, 在金属中和在半导体中一样, 这也是可能的, 在一定的条件下, 起传导作用的“客体”是空穴。虽然, 归根结蒂运动的是晶体中的电子, 然而动量和能量的关系以及对外场的反应却正是我们对电流由正粒子载运时所预期的那样。

我们来看一看是否能对霍尔效应预言的电势差的大小作一个定量的估计。假如图14-7的伏特计中的电流可以忽略, 那么半导体内的电荷一定从左向右运动, 并且垂直的磁力一定精确地被垂直的电场力抵消, 我们把这垂直的电场记作 \mathcal{E}_u (“tr”表“横向”)。假如此电场抵消磁力, 必须:

$$\mathcal{E}_u = -v_{\text{漂移}} \times \mathbf{B}, \quad (14.9)$$

应用式(14.6)给出的漂移速度和电流密度的关系, 我们得到

$$\mathcal{E}_u = -\frac{1}{qN} j B.$$

晶体上面和下面的电势差当然等于电场强度乘以晶体的高度。晶体中的电场强度 \mathcal{E}_u 正比于电流密度和磁场强度。比例常数 $1/qN$ 称为霍尔系数, 通常用符号 R_H 表示。霍尔系数只取决于载流子的密度——倘若某一符号的载流子是绝大多数, 因此, 霍尔效应的测量是一种确定半导体中载流子密度的方便的实验方法。

§ 14-4 半 导 体 结

我们现在来讨论如果取两块具有不同内部特性——例如掺有不同种类或不同数量的杂质——的锗或硅, 并把它们放在一起做成一个“结”, 这时将发生什么情形。让我们从所谓的

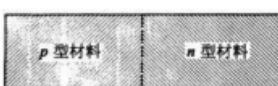


图 14-8 p - n 结

p - n 结开始, 在结上边界的一边是 p 型锗, 边界的另一边是 n 型锗——如图 14-8 所示。实际上, 把分开的两块晶体放在一起并使它们在原子的尺度上均匀接触是不实际的。实际上, 结是在一块单晶上做出来的, 将单晶做成两个分隔的区域。一个方法是当晶体生长到一半的时候在

“熔体”中掺入一些适当的杂质。另一种方法是在表面上涂一点杂质元素, 然后加热晶体使一些杂质原子扩散到晶体的内部。用这些方法制成的结没有明显的边界, 但是边界可以做到 10^{-4} cm 左右那样薄。对于下面的讨论, 我们将考虑理想的情况, 即晶体这两个具有不同性质的区域有明显的边界隔开。

在 p - n 结的 n 型一边有可以运动的自由电子, 还有使总电荷平衡的固定的施主位。在 p 型的一边有自由空穴运动着, 并有等量的受主位使电荷平衡。实际上这描写的是我们使两种材料相互接触之前的状况。一旦它们联结在一起, 靠近交界处的情况就改变了。当 n 型材料中的电子到达边界时, 它们并不像在自由表面上那样被反射回去, 而可以一直进入 p

型材料。因此, n 型材料中的一些电子会扩散到电子比较少的 p 型材料中。这过程不会一直进行下去, 因为当 n 型的一边失去电子后净的正电荷就要增加, 直到最后建立起一个电压以阻止电子扩散到 p 型一边去。同样, p 型材料中的正载流子通过结扩散进入 n 型材料中。当它们这样做时, 在后面留下了过量的负电荷。在平衡条件下, 净扩散电流必定等于零。这是由电场造成的, 因为所建立的电场要把正载流子拉回 p 型材料。

我们这里描写的两种扩散过程是同时进行的。你们要注意, 这两个过程都使 n 型材料带正电, 使 p 型材料带负电。由于半导体材料的有限的电导率, 从 p 的一边到 n 的一边电势的变化只在靠近边界比较窄的区域内发生, 在每块材料的主体部分中, 电势仍然是均匀的。我们设 x 轴的方向垂直于边界表面。那么电势将随 x 而变化, 如图 14-9(b) 所示。我们还在图(c) 中画出了预计的 n 载流子密度 N_n 和 p 载流子密度 N_p 的变化。在离结远的地方, 载流子密度 N_p 和 N_n 应当正好等于同样温度下两块材料各自的平衡密度。(图中所画的结的两边 p 型材料比 n 型材料掺杂更重。) 由于结处的电势梯度, 正载流子必须爬过一电势坡才能到达 n 型的一边。这意味着在平衡条件下在 n 型材料中比在 p 型材料中有较少的正载流子。回忆一下统计力学的定律, 我们预期两边的 p 型载流子数目的比由下面的方程给出:

$$\frac{N_p(\text{n 边})}{N_p(\text{p 边})} = e^{-q_p V / kT}, \quad (14.10)$$

指数的分子中乘积 $q_p V$ 就是使电荷 q_p 通过电势差 V 所需的能量。

对 n 型载流子密度我们有完全相同的方程式:

$$\frac{N_n(\text{n 边})}{N_n(\text{p 边})} = e^{-q_n V / kT}. \quad (14.11)$$

假如我们已知在两种材料内各自的平衡密度, 我们可以用上面两个方程式的任何一个来测定给两边的电势差。

注意, 假如方程式(14.10)和(14.11)给出同样的电势差 V 数值, 乘积 $N_p N_n$ 无论在 p 的一边还是在 n 一边必定相同。(记住 $q_n = -q_p$ 。) 然而, 我们前面已经看到, 这一乘积只依赖于温度和晶体的隙能。假定晶体的两边都处于同样的温度, 这两个方程式中的电势差具有同一数值。

因为从结的一边到另一边有一个电势差, 它看上去有些像一个电池。假如我们从 n 型的一边到 p 型一边联结一根导线或许会获得电流。真是这样的话, 确实十分美妙, 因为如果真是这样的话电流就会一直流下去而不会消耗掉任何材料, 于是我们就有一个人违背热力

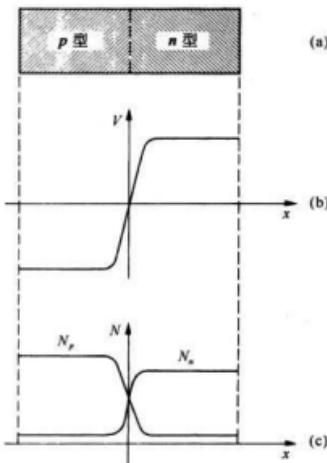


图 14-9 在未加偏压的半导体结上的电势和载流子密度

学第二定律的无穷无尽的能源！然而，如果你从 p 的一边接一根导线到 n 的一边，实际上不会有电流。其理由很容易看出来的。假定我们先设想一根由未掺杂的材料制成的导线。当我们把这根导线联结到 n 型的一端时，我们就有了一个结。在结的两边就会产生电势差。我们假定这正好等于从 p 型材料到 n 型材料的电势差的一半。当我们把未掺杂的导线联结到结的 p 型一边，在这个结上也有电势差——也等于 $p-n$ 结上电势降落的一半。在所有结上，电势差会自动调整到使得电路中没有净电流。不论你用哪一种导线联结 $n-p$ 结的两边，你都造出两个新的结，只要所有的结都在同样的温度下，在结上的电势跃变都互相补偿，从而电路中没有电流。不过——假如你仔细研究一下——如果有一些结和另一些结的温度不同，结果就会有电流流动。其中有一些结会被这电流加热而另一些将被冷却，于是热能就转化为电能。测量温度的热电偶以及温差发电机就是利用这一效应工作的。同一效应也用于制造小型制冷机。

如果我们不能测量 $n-p$ 结两边的电势差，我们怎样才能肯定图 14-9 所示的电势梯度确实存在呢？一个方法是用光来照射结。光子被吸收后会产生电子-空穴对。在结处存在的强电场（等于图 14-9 的电势曲线的斜率）中，空穴会被驱赶到 p 型区域，而电子会被驱赶到 n 型区域。现在如果把结的两边联结到外电路中，这些额外的电荷会产生电流。在结上，光能会转换为电能。使某些人造卫星运转的产生电功率的太阳能电池就是按照这个原理工作的。

在关于半导体结的作用的讨论中，我们一直假定空穴和电子的行为或多或少是相互独立的——除了它们以某种方式达到适当的统计平衡。当我们描写光照在结上产生电流时，我们假定在结区产生的电子或空穴在被相反极性的载流子湮没之前就进入了晶体的主体部分。在紧靠结的区域内，两种符号的载流子的密度近似地相等，电子-空穴湮没效应（也常称之为“复合”）是重要的效应，在对半导体结的详尽分析中必须认真考虑这个效应。我们还假定在结区内产生的电子或空穴在复合之前有很多的机会进入晶体体内。对于典型的半导体材料，电子或空穴找到其异号的伴侣并湮没的典型时间在 $10^{-3} \sim 10^{-7}$ s 范围之内。顺便说说，这个时间远远长于我们分析电导率时所采用的与晶体中的散射中心的两次碰撞之间的平均自由时间。在典型的 $n-p$ 结中，在结区中形成的电子或空穴被扫至晶体内部所需的时间一般大大短于复合时间。因此，大多数电子-空穴对都将对外电流作出贡献。

§ 14-5 半导体结的整流

下面我们要说明 $p-n$ 结为何可以用作整流器。如果我们在结两边加上电压，当极性为某一方向时会有很大的电流流过，但在相反的方向上加上同样的电压时，却只有很小的电流。假如在结上加的是交流电压，净电流只沿一个方向流动——电流被“整流”。让我们再来看一看图 14-9 所描绘的平衡条件下发生些什么。在 p 型材料中有很大的正载流子浓度 N_p ，这些载流子向各处扩散并且每秒内有一定数量到达结处。到达结的正载流子电流正比于 N_p 。然而，其中的大多数被结处高的电势坡挡了回去，只有约为 $e^{-qV/kT}$ 的一部分能通过。从另一边也有正载流子电流来到结处。这个电流也正比于 n 型区域的正载流子密度，但这里的载流子密度大大低于 p 型一边的密度。当正的载流子从 n 型一边来到结处时，它们遇到的是具有负的坡度的电势坡并立即滑下斜坡到结的 p 型一边。我们把这电流称为 I_0 。在平衡时从两个方向来的电流相等。我们认为下面的关系式成立：

$$I_0 \propto N_p (n\text{边}) = N_p (p\text{边}) e^{-qV/kT}. \quad (14.12)$$

你们会注意到这个方程实际上和式(14.10)完全一样。我们只是用了不同的方法导出它。

不过,假定我们使结的 n 边的电压降低一个数量 ΔV ——我们可以通过在结上加上外电势差来做到这一点。现在电势坡两面的电势差不再是 V 而是 $V - \Delta V$ 。从 p 边到 n 边的正载流子电流的表达式中这个电势差在指数因子中出现。令这个电流为 I_1 , 我们有

$$I_1 \propto N_p (p\text{边}) e^{-q(V-\Delta V)/kT}.$$

这个电流比 I_0 正好大了一个因子 $e^{q\Delta V/kT}$ 。所以 I_1 和 I_0 之间有如下的关系:

$$I_1 = I_0 e^{+q\Delta V/kT}. \quad (14.13)$$

来自 p 边的电流随着外加电压 ΔV 指数式地增加。然而, 只要 ΔV 不太大从 n 边来的正载流子的电流保持不变。当它们来到势垒, 这些载流子还是会发现向下的电势坡并且全部降落到 p 边。(如果 ΔV 大于自然电势差 V , 情形就不同了, 但是我们不去考虑在这样高的电压下会发生些什么。) 穿过结的正载流子的净电流 I 就是两边来的电流之差:

$$I = I_0 (e^{+q\Delta V/kT} - 1). \quad (14.14)$$

空穴的净电流 I 流进 n 型区域。从那里空穴扩散到 n 型区域的体内, 最终被多数 n 型载流子——电子——湮没。在这湮没过程中损失的电子将由 n 型材料外端来的电子电流补偿。

当 ΔV 为零, 式(14.14)中的净电流为零。对于正的 ΔV , 电流随外加电压很快增加。对于负的 ΔV , 电流改变符号, 但是指数项很快变为可忽略, 并且负的电流永远不会超过 I_0 ——按照我们的假设这 I_0 是相当小的。这个反向电流 I_0 被结的 n 边少数 p 型载流子很小的密度所限制。

如果你对流过结的负载流子的电流作完全相同分析, 首先对没有电势差的情形, 然后对小的外加电势差 ΔV , 你会再次得到完全和式(14.14)一样的净电子电流的方程式。因为总电流是两种载流子贡献的电流之和, 倘若我们认为 I_0 是加上反向电压时可能流过的最大电流, 那么式(14.14)完全可用总全电流。

式(14.14)的电压-电流特性曲线如图 14-10 所示。它表示固体二极管(诸如现代的计算机中用的那种二极管)的典型性质。我们应当看到式(14.14)只对小的电压是正确的。当电压和自然的内部电势差 V 不相上下或者比它更大时, 其他的效应就要起作用了, 电流不再遵从这一简单的式子。

顺便说说, 你们还记得在第 1 卷第 46 章讨论“力学整流器”——棘轮和掣爪——的时候, 我们曾得到和这里得出的式(14.14)完全相同的式子。我们之所以在这两种情况中得到同样的式子是因为两者的基本物理过程十分相似。

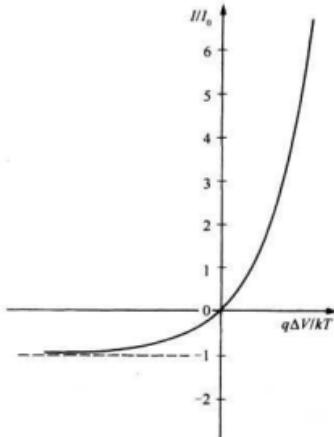


图 14-10 通过结的电流作为
结两端电压的函数

§ 14-6 晶体管

或许半导体最重要的应用是晶体管。晶体管由两个互相靠得非常近的半导体结组成。

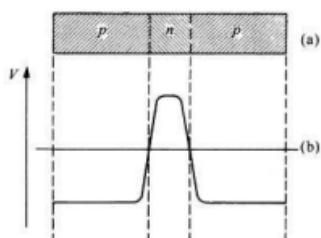


图 14-11 没有加外电压时晶体管中的电势分布

它的工作原理有一部分与我们刚才描写的关于半导体二极管即整流结的原理相同。假设我们制造一小块锗，上面有 3 个不同的区域，一个 p 型区域，一个 n 型区域，以及另一个 p 型区域，如图 14-11(a) 所示。这一组合称作 $p-n-p$ 晶体管。晶体管中的这两个结的行为都和我们在上一节中所描写的一样。特别是每一个结上都有一个电势梯度，从 n 型区域到每一个 p 型区域都有一定的电势降落。假如两个 p 型区域有同样的内部性质，我们沿着晶体观察，测量到的电势变化就如图 14-11(b) 所示。

现在让我们设想，把 3 个区域的每一个都加上外电压，如图 14-12(a) 所示。我们以接在左方 p 区的端点作为所有电压的参考点，所以按照定义这点为零电势。我们把这一端称为发射极。 n 型区域称为基极，它联接到小的负电势上。右方的 p 型区域称为集电极，并联接到较大的负电势上。在这样的情形下晶体上的电势变化就如图 14-12(b) 所示。

我们先来看一看正载流子的情形，因为主要是它的行为控制着 $p-n-p$ 晶体管的工作。既然发射极相对于基极是处在正电势，正载流子的电流就会从发射极区流入基极区。由于结在“正向电压”作用下——相应于图 14-10 中图的右半部分——流过的电流比较大。在这样的条件下，正载流子或空穴从 p 型区域被“发射”到 n 型区域。你可能想象这个电流会通过基极的端点 b 流出 n 型区域。然而，现在接触到了晶体管的奥秘了。 n 型区域做得非常薄——典型的厚度是 10^{-3} cm 或更小，大大小于它们的横向线度。这意味着当空穴进入 n 型区域后，在和 n 型区域中的电子湮没之前有很大的机会扩散到另一个结。当它们到达 n 型区域的右方边界时，它们遇到急剧下降的电势坡并立即落入右方的 p 型区域。晶体的这一边称为集电极，因为它“收集”扩散穿过 n 型区域的空穴。在一个典型的晶体管中，除了很小的百分比以外，离开发射极并进入基极的空穴电流大多在集电极区域被收集起来，其余只有很少一部分贡献净的基极电流。基极和集电极电流之和显然等于发射极电流。

现在想象一下，假如我们稍稍改变一下基极端的电势 V_b 会发生些什么。因为我们是在图 14-10 的曲线的相对陡峭的部分，电势 V_b 很小的变化会引起发射极电流 I_e 很大的变化。

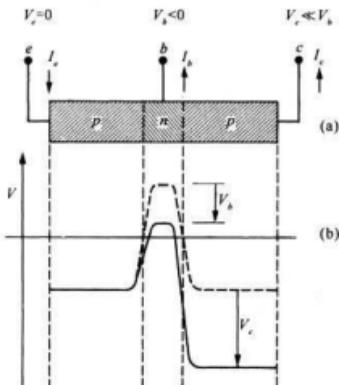


图 14-12 晶体管工作时的电势分布

因为集电极电压 V_c 大大负于基极电压，电势的微小变化不会显著地影响基极和集电极之间陡峭的电势坡。发射到 n 区域的大多数正载流子仍旧被集电极俘获。于是当我们改变基极电势时，集电极的电流 I_c 就会有相应的变化。不过，根本的一点是基极电流 I_b 始终只是集电极电流的一小部分。晶体管是一个放大器，基极上引入小电流 I_b 在集电极上给出大电流——大 100 倍左右。

迄今为止我们一直忽略了的负载流子——电子——又怎样呢？首先，我们并不认为任何值得注意的电子电流在基极和集电极之间流动。集电极上加有大的负电压时，基极中的电子不得不爬上很高的势能高坡，而这样做的概率是很小的。只有很小的电子电流流向集电极。

在另一方面，基极中的电子确实能够进入发射极区域。事实上，你可能会料到在这个方向的电子电流比得上从发射极到基极的空穴电流。但是这种电子电流不仅没有用，而且还有害，因为对于一定的集电极空穴电流，它使所需的基极总电流增大。所以，设计晶体管时要把发射极的电子电流减到最小。电子电流正比于 N_a (基极)，即基极材料中的负载流子密度，而从发射极来的空穴电流取决于 N_p (发射极)，即发射极区中的正载流子密度。对 n 型材料进行低掺杂， N_a (基极)可以做得比 N_p (发射极)小得多。(非常薄的基区也有很大的帮助，因为集电极扫除了这个区域内的空穴，从而大大地增加了从发射极进入基极的平均空穴电流，而电子电流却保持不变。)最后结果是穿过发射极-基极结的电子电流比空穴电流小得多，所以在 $p-n-p$ 晶体管的运作中电子并没有起任何重要的作用。电流受空穴运动所支配，晶体管则如前所述表现得像个放大器。

将图 14-11 中的 p 型和 n 型材料对换也可以做成晶体管。于是我们得到所谓的 $n-p-n$ 晶体管。在 $n-p-n$ 晶体管中，大部分电流由电子携带，这些电子从发射极流入基极并由此流到集电极。显然，如果将电极的电势选取相反的符号，我们关于 $p-n-p$ 晶体管的所有论证同样适用于 $n-p-n$ 晶体管。

第 15 章 独立粒子近似

§ 15-1 自 旋 波

在第 13 章中，我们已经求出了关于一个电子或某些其他“粒子”，譬如一个原子的激发，穿过晶格传播的理论。上一章，我们把这个理论应用于半导体。但是每当我们讨论有许多电子存在的情形时，我们总是忽略它们之间的相互作用，当然这样做只是一种近似。在本章里我们将进一步讨论电子间的相互作用可以忽略这一概念。我们还将趁此机会向读者进一步介绍粒子传播理论的某些应用。既然我们一般仍旧忽略粒子间的相互作用，所以在本章中除了一些新的应用以外，实际上没有什么新的内容。我们首先要考虑的例子是，当存在着一个以上的“粒子”时，仍能相当精确地写出其正确方程的情形。虽然我们不去仔细分析这个问题，但是通过这些例子我们能看出这种忽略相互作用的近似方法是如何做出的。

作为第一个例子我们考虑铁磁晶体中的自旋波。在第 2 卷第 36 章中我们讨论过铁磁性理论。在绝对零度时所有铁磁晶体内的磁性有贡献的电子自旋都是平行排列的。这些自旋之间存在着相互作用能，当所有自旋都向下时，这一相互作用能最低。但是当温度不为零时，其中一部分自旋就有机会转向。在第 2 卷第 36 章中我们近似计算了这一概率。现在我们将叙述其量子力学理论，这样你可以看出，如果要更严格地求解这个问题，你应该做些什么。（这里我们仍旧要作一些理想化，假设这些电子被限定在各自的原子中，而且只有相邻的自旋之间才有相互作用。）

我们考虑这样一种模型，该模型中每个原子的电子除一个以外其余都是成对的，因此所有的磁效应都来自每个原子中的一个自旋为 $1/2$ 的电子。我们进一步设想这些电子被定域在晶格的原子座上，这个模型大致对应于金属镍。

我们再假定任何两个相邻的自旋电子间有相互作用存在，这个相互作用对整个体系的能量给出了这样一项

$$E = - \sum_{i,j} K \sigma_i \cdot \sigma_j, \quad (15.1)$$

式中 σ 表示电子的自旋，而求和遍及所有相邻的电子对。当我们考虑氢原子中由于电子和质子磁矩的相互作用而引起的超精细分裂时，就曾经讨论过这种类型的相互作用能。那时我们把它表示为 $A \sigma_e \cdot \sigma_p$ 。现在，对于给定的一对电子，譬如说位于原子 4 和原子 5 的一对电子，其哈密顿为 $-K \sigma_4 \cdot \sigma_5$ 。对每一对这样的电子都有这样的一项，而哈密顿（正如对经典能量所预期的）就是各对相互作用项之和。由于能量含有一 K 的因子，所以正 K 将相当于铁磁性，也就是说，当相邻自旋平行时总能量最低。在实际晶体中，可能还有次近邻的相互作用项等等，但在现阶段我们没有必要考虑这种复杂的情形。

有了式(15.1)的哈密顿我们就有了一——在我们的近似条件下——对铁磁体的完整描

述，并且应可由之导出磁化的性质，我们也应能计算由于磁化的热力学性质。如果我们能够求出全部能级，则在温度 T 时晶体的各种性质可以由下述统计力学原理求得，即系统处在给定能量为 E 的状态的概率与 $e^{-E/kT}$ 成正比。这个问题至今尚未被完全解出。

我们举一个所有原子都排成一条直线即一维晶格的简单例子来说明一些问题。你们很容易将这些概念推广到三维的情形。在每一个原子的位置上都有一个电子，每个电子都有两个可能的状态，不是自旋朝上就是自旋朝下，整个系统可通过说明所有自旋是如何排列的来描述。我们取体系的哈密顿作为相互作用的能量算符。将式(15.1)中的自旋矢量解释为 σ 算符(或 σ 矩阵)，对于线性晶格我们可写出：

$$\hat{H} = \sum_n -\frac{A}{2} \hat{\sigma}_n \cdot \hat{\sigma}_{n+1}. \quad (15.2)$$

为方便起见，在上式中我们将常数写成 $A/2$ 。(这样一来，以后的一些式子将和第13章的式子完全相同。)

现在我们要问，什么是这个体系的最低能量状态呢？体系的最低能量状态是所有自旋都平行的状态，譬如说所有自旋都向上*的态。我们可以把这个态写成 $| \cdots + + + \cdots \rangle$ ，或 $| \text{gnd} \rangle$ ，就是“基态”或最低能量态。我们不难算出这个态的能量。一种办法是用 $\hat{\sigma}_z$ ， $\hat{\sigma}_x$ 和 $\hat{\sigma}_y$ 表示出所有 σ 矢量，再仔细算出哈密顿的每一项在基态时的值，然后把结果相加。其实我们也可使用一个好而简捷的方法。在 § 12-2 中我们已经知道， $\hat{\sigma}_i \cdot \hat{\sigma}_j$ 可以用泡利自旋交换算符来表示，如

$$\hat{\sigma}_i \cdot \hat{\sigma}_j = (2 \hat{P}_{ij}^{\text{自旋交换}} - 1), \quad (15.3)$$

式中算符 $\hat{P}_{ij}^{\text{自旋交换}}$ 的作用是交换第 i 个电子与第 j 个电子的自旋。把这个表示式代入式(15.2)，哈密顿就成为：

$$\hat{H} = -A \sum_n \left(\hat{P}_{n, n+1}^{\text{自旋交换}} - \frac{1}{2} \right). \quad (15.4)$$

现在要计算不同态的情形就容易了。例如，若 i 和 j 这两个电子的自旋同时向上，则交换自旋并不会引起任何变化，那么 \hat{P}_{ij} 作用于该态上又回到同一态，这与乘以 +1 等价。因此表示式 $(\hat{P}_{ij} - 1/2)$ 正好等于 $1/2$ 。(从现在开始我们省去 \hat{P} 的上标。)

对于所有自旋都向上的基态；如果你们交换一对特定的自旋，你们仍旧回到原来的态。基态是一个定态。如果把哈密顿作用在这个态上，得到的又是这个态再乘上一个求和项，每一对自旋态提供 $-A/2$ 。这就是说，在体系的基态能量中每个原子提供 $-A/2$ 。

下面我们考虑某些激发态的能量。相对于基态来测量能量是很方便的，这就是说选择基态作为我们的能量零点。为了做到这一点，我们可以对哈密顿中的每一项加上能量 $A/2$ 。这样正好把式(15.4)中的“ $1/2$ ”变成“1”，因此新的哈密顿为

$$\hat{H} = -A \sum_n (\hat{P}_{n, n+1} - 1). \quad (15.5)$$

* 这里的基态实际上是“简并”的，还有其他的状态也具有同样的能量，例如所有自旋都朝下，或者都指向任意别的方向。在 z 方向上加一很小的外磁场后将使所有这些状态具有不同的能量，那末我们这里所选取的状态就是真正的基态。

使用这个哈密顿，体系最低态的能量为零，(对于基态来说)自旋交换算符相当于乘以1，正好把每一项中的“1”消去。

为了描述基态之外的其他态，我们需要一组适当的基础态。一种较方便的处理方法是这样的，按照一个电子自旋向下，或两个电子自旋向下……等等对态进行分组。当然，一个

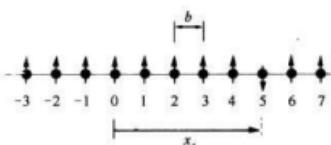


图 15-1 一线性排列的自旋体系的基础态 $|x_s\rangle$ 。其中除位于 x_5 的一个自旋向下外，其余的自旋都向上

该向下自旋位于第 n 个原子的坐标 x_n 处。

(15.5) 式的哈密顿对态 $|x_5\rangle$ 的作用结果是什么呢？譬如说哈密顿算符中 $-A(\hat{P}_{7,8}-1)$ 这一项，其中算符 $\hat{P}_{7,8}$ 是交换 7、8 两个相邻原子的自旋，而在态 $|x_5\rangle$ 中这两个原子的自旋都是向上的，所以它的作用并没有改变什么， $P_{7,8}$ 只相当于乘以 1：

$$\hat{P}_{7,8} |x_5\rangle = |x_5\rangle.$$

由此得到

$$(\hat{P}_{7,8}-1) |x_5\rangle = 0.$$

于是，哈密顿中除了包含原子 5 的那些项以外，其余所有的项作用于 $|x_5\rangle$ 后都为 0。算符 $\hat{P}_{4,5}$ 对态 $|x_5\rangle$ 的作用使原子 4 的自旋(向上)和原子 5 的自旋(向下)作了交换，结果得到除了原子 4 的自旋向下外，其余所有原子的自旋都向上的态，即

$$\hat{P}_{4,5} |x_5\rangle = |x_4\rangle.$$

同理

$$\hat{P}_{5,6} |x_5\rangle = |x_6\rangle.$$

因此，哈密顿中只留下 $-A(\hat{P}_{4,5}-1)$ 和 $-A(\hat{P}_{5,6}-1)$ 两项，它们对 $|x_5\rangle$ 的作用，分别得到 $-A|x_4\rangle + A|x_5\rangle$ 和 $-A|x_6\rangle + A|x_5\rangle$ ，最后的结果为

$$\hat{H}|x_5\rangle = -A \sum_n (\hat{P}_{n,n+1} - 1) |x_5\rangle = -A(|x_6\rangle + |x_4\rangle - 2|x_5\rangle). \quad (15.6)$$

当哈密顿作用在态 $|x_5\rangle$ 上时，它给出了一些处在态 $|x_4\rangle$ 和 $|x_6\rangle$ 的振幅，这就意味着向下的自旋跳跃到相邻原子上具有一定振幅。由于自旋之间的相互作用，如果开始时某一个自旋向下，那么在以后的时刻，另一个原子代替它变成自旋向下就有一定的概率。将哈密顿作用在一般态 $|x_n\rangle$ 上就给出：

$$\hat{H}|x_n\rangle = -A(|x_{n+1}\rangle + |x_{n-1}\rangle - 2|x_n\rangle). \quad (15.7)$$

这里要特别注意，如果我们所取的是只有一个自旋向下的态的完全集，则它们将只在它们本身之间混合。哈密顿绝不会把这些态同其他具有多于一个自旋向下的态混合。你们只能交

换自旋，你们就绝不要改变总的向下自旋的数目。

使用哈密顿的矩阵记法是方便的，譬如说 $H_{n,n} = \langle x_n | \hat{H} | x_n \rangle$ ；方程式(15.7)等效于

$$\begin{aligned} H_{n,n} &= A, \\ H_{n,n+1} &= H_{n,n-1} = -A, \\ H_{n,m} &= 0, \text{ 当 } |n-m| > 1. \end{aligned} \quad (15.8)$$

现在我们要问，对于具有一个自旋向下的那些态的能级是怎样的呢？和往常一样，我们令 C_n 为某个态 $|\psi\rangle$ 处于态 $|x_n\rangle$ 的振幅。如果 $|\psi\rangle$ 是一个具有确定能量的态，则所有的 C_n 必定以相同的方式随时间变化，即

$$C_n = a_n e^{-iE_n t/\hbar}. \quad (15.9)$$

我们可以把这个试解代入通常的哈密顿方程中

$$i\hbar \frac{dC_n}{dt} = \sum_m H_{nm} C_m. \quad (15.10)$$

同时使用式(15.8)的矩阵元。当然我们得到无限多个方程，但可以吧它们都写成如下形式：

$$E a_n = 2A a_n - A a_{n-1} - A a_{n+1}. \quad (15.11)$$

这样我们又得到了与在第 13 章中已解出的完全相同的问题，不过那里是 E_0 而现在是 $2A$ 。相应于振幅 C_n 的解（自旋向下的振幅）是以传播常量 k 和能量为

$$E = 2A(1 - \cos kb) \quad (15.12)$$

的沿晶格传播的波。式中 b 是晶格常数。

具有确定能量的解与自旋向下的“波”——称为“自旋波”——相对应，对于每个波长，有一个相应的能量。对于长波长（小的 k ），其能量按下式变化

$$E = Ab^2 k^2. \quad (15.13)$$

和以前一样，我们可以考虑一个定域的波包（不过其中只包含长波长），它相当于在晶格中某一部分的一个自旋向下的电子。这个向下自旋表现得像一个“粒子”。因为它的能量与 k 的关系由式(15.13)描写，这个“粒子”具有有效质量：

$$m_{\text{有效}} = \frac{\hbar^2}{2Ab^2}. \quad (15.14)$$

有时我们把这些“粒子”称为“磁波子”。

§ 15-2 双自旋波

现在我们讨论存在两个向下自旋的情形。我们还是要选取一组基础态，选择在两个原子的位置上有向下自旋的那些态，如图 15-2 所示。我们可以用两个向下自旋的位置的坐标 x 来表示这个态，图 15-2 中所示的态可写成 $|x_s, x_s\rangle$ 。一般说来，基础态为 $|x_e, x_m\rangle$ ，这是

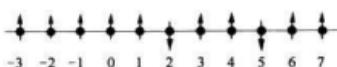


图 15-2 具有两个向下自旋的状态

一个二重无限集！在这套描述体系中，态 $|x_1, x_2\rangle$ 和态 $|x_2, x_1\rangle$ 是完全相同的态，因为这两个态都只是说在位置 4 和位置 9 处各有一个向下的自旋，它们的次序并没有什么意义。此外，态 $|x_1, x_2\rangle$ 是无意义的，并不存在这种情况。我们可以这样来

描述任何一个态 $|\psi\rangle$ ：即给出它在每个基础态的振幅。如此， $C_{m,n} = \langle x_m, x_n | \psi \rangle$ 表示处于态 $|\psi\rangle$ 的系统在第 m 个原子和第 n 个原子同时具有向下自旋这种态中的振幅。现在所出现的这种复杂性并不是概念上的复杂，仅仅是簿记的复杂性而已。（量子力学的一个复杂性就在于登录记号法。随着向下自旋越来越多，我们所用的记号也因含有大量标记而变得越来越复杂，方程式也总是令人望而生畏，其实概念并不比最简单的情形复杂多少。）

这个自旋系统的运动方程是一组关于 $C_{m,n}$ 的微分方程。它们是

$$i\hbar \frac{dC_{n,m}}{dt} = \sum_{\bar{q}} (H_{nm, \bar{q}}) C_{\bar{q}}. \quad (15.15)$$

假定我们希望求出该系统的定态，和往常一样，把对时间的微商变成 E 乘上振幅，而 $C_{m,n}$ 可以用系数 $a_{m,n}$ 代替。接下来我们必须仔细地算出 H 对自旋 m 和 n 向下的这个态作用的结果。这种计算并不困难。暂时假定 m 和 n 相距甚远，这样我们就不必为显而易见的麻烦而操心。对位于 x_n 处的向下自旋的交换操作是将向下自旋或者转移到第 $n+1$ 个原子或者移到第 $n-1$ 个原子上去。所以有一个振幅是现在的态来自态 $|x_m, x_{n+1}\rangle$ ，也有一个振幅来自态 $|x_m, x_{n-1}\rangle$ 。或者也可能是另一个自旋在交换，就是 $C_{n,m}$ 来自 $C_{m+1,n}$ 或 $C_{m-1,n}$ ，都有一定的振幅。所有这些效应都应是相等的。最后关于 $C_{m,n}$ 的哈密顿方程为

$$Ea_{m,n} = -A(a_{m+1,n} + a_{m-1,n} + a_{m,n+1} + a_{m,n-1}) + 4Aa_{m,n}. \quad (15.16)$$

这个方程是正确的，但下面两种情况除外：如果 $m = n$ ，则方程根本不存在；如果 $m = n \pm 1$ ，则方程式(15.16)中有两项应去掉。我们将不去考虑这些例外情形。我们暂不理会这些方程中有少数要稍加改动这一事实。毕竟我们假定晶体是无限的，而且我们有无限多项，所以忽略很少几个不会有大关系。那么作为粗略的一级近似，让我们忘掉那些改动过的方程，换句话说，我们假定方程式(15.16)对所有 m 和 n ，即使它们相差 1 也无所谓。这就是我们的近似方法的本质部分。

方程的求解并不难，我们立即得到

$$C_{m,n} = a_{m,n} e^{-iEt/\hbar}, \quad (15.17)$$

而

$$a_{m,n} = (\text{常数}) e^{ik_1 x_m} e^{ik_2 x_n}, \quad (15.18)$$

式中

$$E = 4A - 2A\cos k_1 b - 2A\cos k_2 b. \quad (15.19)$$

至此我们可以想一想，如果我们有两个独立的、分别对应于 $k = k_1$ 和 $k = k_2$ 的单自旋波（如上一节中的情况），将会出现什么情况呢？根据式(15.12)，它们具有能量

$$\epsilon_1 = (2A - 2A\cos k_1 b)$$

及

$$\epsilon_2 = (2A - 2A\cos k_2 b).$$

注意,式(15.19)中的能量 E 只是它们之和,即

$$E = \epsilon(k_1) + \epsilon(k_2). \quad (15.20)$$

换句话说,我们可以这样来考虑我们所得到的解:有两个粒子,也就是两个自旋波,一个粒子具有以 k_1 描述的动量,另一个粒子具有用 k_2 表示的动量,体系的能量就是这两个粒子能量之和,这两个粒子的作用是完全独立的。对于这个体系就只有这些。

当然,我们已经做了某些近似,但是我们现在并不想讨论答案的精确性。不过,你们也许猜想到,在一个适当大小的晶体中包含着数十亿个原子,因此在哈密顿中就有数十亿项,丢掉几项并不会造成很大的误差。要是我们有很多向下的自旋,以致向下自旋的密度相当大,那么我们就得考虑修正了。

[有趣的是:如果只有两个向下自旋,就可以写出精确的解。这结果并不特别重要,但是引起兴趣的是对于这种情况的方程可以精确求解。这一解为:

$$a_{m,n} = \exp[ik_c(x_m + x_n)] \sin(k|x_m - x_n|), \quad (15.21)$$

而能量为: $E = 4A - 2\text{Acos } k_1 b - 2\text{Acos } k_2 b,$

波数 k_c , k 与 k_1 , k_2 的关系为:

$$k_1 = k_c - k, \quad k_2 = k_c + k. \quad (15.22)$$

这个解包含了这两个自旋的“相互作用”,它说明当两个自旋互相靠近时就有一定的散射机会,自旋表现得很像具有相互作用的粒子。但是,有关它们散射的详细理论超出了我们想要在这里讨论的范围。]

§ 15-3 独立粒子

在上节中,我们已经写出了两个粒子系统的哈密顿式(15.15)。然后取近似,这近似相当于忽略两个粒子间任何“相互作用”,我们求得了由式(15.17)和(15.18)所描述的定态,这个态正好就是两个单粒子态的乘积。然而,我们对于式(15.18)中 $a_{m,n}$ 的解实在并不满意。我们早就谨慎地指出,态 $|x_1, x_2\rangle$ 与 $|x_1, x_2\rangle$ 并无区别—— x_m 和 x_n 的次序没有什么意义。一般来说,如果我们交换 x_m 和 x_n 的值,振幅 $C_{m,n}$ 的代数表达式必须不变。因为这种交换并不改变状态,不管怎么说,它应代表在 x_m 处及 x_n 处各找到一个向下自旋的态的振幅。但是要注意,由于 k_1 和 k_2 通常是不同的,式(15.18)对 x_m 和 x_n 并不对称。

问题在于我们没有使式(15.15)的解一定要满足这个附加条件。幸好这个问题很容易解决。首先我们注意到,该哈密顿方程的一个解

$$a_{m,n} = K e^{ik_2 x_m} e^{ik_1 x_n} \quad (15.23)$$

与式(15.18)一样好,甚至其能量也与式(15.18)所得的能量相同。式(15.18)和(15.23)的任何线性组合也是一个很好的解,其能量仍由式(15.19)给出。由于对称性要求,我们所应选取的解就是式(15.18)和式(15.23)之和:

$$a_{m,n} = K [e^{ik_1 x_m} e^{ik_2 x_n} + e^{ik_2 x_m} e^{ik_1 x_n}]. \quad (15.24)$$

现在,给定任何 k_1 和 k_2 ,振幅 $C_{m,n}$ 和我们安置 x_m 和 x_n 的方式无关——如果我们恰巧把 x_m 和 x_n 规定反了,我们仍然得到相同的振幅。我们用“磁波子”解释方程式(15.24)结果就不同,我们不能再说明该方程表示波数为 k_1 的一个粒子和波数为 k_2 的第二个粒子。振幅式(15.24)表示具有两个粒子(磁波子)的一个态。这个态的特征由两个波数 k_1 和 k_2 来表征。我们的解看上去很像是由一个动量为 $p_1 = \hbar k_1$ 的粒子和另一个动量为 $p_2 = \hbar k_2$ 的粒子组成的复合态,但是我们不能在这个态中指出哪个粒子是哪一个。

至此,读者应该从这些讨论中回忆起第4章以及我们关于全同粒子的描述。我们刚才正好证明自旋波粒子——磁波子——的行为像全同玻色子。所有的振幅对两个粒子的坐标来说必须是对称的——这等于说,如果我们“交换两个粒子”,我们将回到同样的振幅并且具有相同的符号。但是,你们可能会想,为什么我们在构成式(15.24)时选取两项相加的方式?为什么不把这两项相减?采用减号时,交换 x_m 和 x_n 只不过改变 $a_{m,n}$ 的符号,这并没有什么关系。但是交换 x_m 和 x_n 并没有引起任何改变——晶体中的所有电子仍旧位于原来的地方,甚至振幅的符号也没有理由要改变。由此可见,磁波子的行为像玻色子*。

上面的讨论主要有两点:第一,告诉你们一些有关自旋波之事;第二,向读者揭示一种态,它的振幅是两个振幅的乘积,它的能量是相当于这两个振幅的能量的和。对于独立粒子来讲,振幅相乘,能量相加。你们不难理解为什么能量是求和。能量是虚指数中 t 的系数,它与频率成正比。如果有两个粒子正在做某些事情,一个具有振幅 $e^{-iE_1 t/\hbar}$,而另一个具有振幅 $e^{-iE_2 t/\hbar}$,如果这两件事情一起发生的振幅是各个振幅之积,在这乘积中就出现一个频率,它是两个频率之和,与振幅乘积相对应的能量是两个能量之和。

我们通过相当冗长的论证来向你们说明一件简单的事情,当你们不考虑粒子间的任何作用时,就可以把每个粒子看成是独立的。这些粒子可以各自存在于它们单独存在时所具有的各种不同的状态中,每个粒子都贡献出它们单独存在时所具有的能量。然而你们必须记住,如果它们是全同粒子,则它们会表现为不是玻色子就是费米子,这由具体问题来定。例如,把两个额外的电子加到晶体里去,它们就会表现得像费米子。当交换这两个电子的位置时,其振幅必定改变符号。在相应于式(15.24)的方程中,在右边的两项中间就应该有一个负号。因而两个费米子不可能处在完全相同的情况下——具有相同的自旋和相同的 k 。这种态的振幅为零。

§ 15-4 芳 分 子

虽然量子力学提供了确定分子结构的基本定律,但是这些定律只能精确地用在最简单的化合物上。因此化学家们设计出各种近似方法来计算复杂分子的某些性质。现在我们要对你们说明的是有机化学家们是怎样应用独立粒子近似的。我们从讨论苯分子开始。

在第10章中我们从另一种观点讨论苯分子,在那里我们对分子采用的是双态系统的近似图像,它具有如图15-3所示的两个基础态。苯分子有一个由6个碳构成的环,每个碳原

* 一般说来,我们所讨论的这类准粒子,其行为或者像玻色子,或者像费米子,而对自由粒子来说,具有整数自旋的粒子是玻色子,具有半整数自旋的粒子是费米子。“磁波子”代表一个自旋向上的电子转成自旋向下,其自旋的变化为1。磁波子具有整数自旋,它是玻色子。

子上键联着一个氢原子。按照惯用的价键图像,我们必须假设碳原子之间半数是双键,而在最低能量的条件下,双键的分布有如图所示的两种可能形式。当然还有别的能量较高的态。当我们在第 10 章讨论苯分子时,我们只采用这两个态,并没有考虑别的态。我们发现分子的基态能量并不是图中所示的任一态的能量,而是比这个态的能量低了一个量值,该量值与从其中一个态跃迁至另一态的振幅成正比。

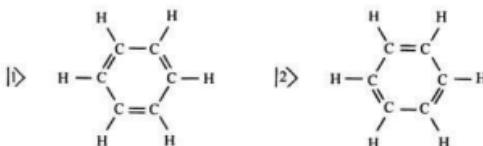


图 15-3 第 10 章中采用的苯分子的两个基础态

现在我们从完全不同的观点——用不同的近似方法来考察这种分子。这两种观点将会给我们不同的答案,但是如果我们改进任何一种近似方法,我们都应得到正确的结果,即对苯分子的有效描述。然而,如果我们不肯费心去改进它们,当然通常往往如此,那么你们就不应该为这两种描述不完全一致而感到惊奇。我们至少将用新的观点来证明,苯分子的最低能量比如图 15-3 所示的任一个三键结构要低。

现在我们采用下面的图像:我们设想苯分子的 6 个碳原子仅以单键相联,如图 15-4 所示。由于一个键代表一对电子,在上述图像中我们已去掉了 6 个电子,它是一个六重电离的苯分子。现在我们来考虑把这 6 个电子放回去会发生什么情况,我们一次放回 1 个电子,并想象每个电子都可以环绕这个环自由运动。我们还假设图 15-4 所示的所有键都满足了,没有必要对之作进一步的考虑。

当我们把 1 个电子放回到这个分子离子时会发生些什么呢?当然它可能落在环的 6 个位置——相当于 6 个基础态——中的任何一个上。它从一个位置跑到相邻的位置上去应该有一定的振幅,比方说为 A。如果我们分析这些定态,就会有一些可能的能级。这仅仅是对一个电子而言。

接下来把第二个电子放进去,现在我们作一个你所能想象到的最荒谬的近似——一个电子的行为并不影响另一个电子的行为。当然它们实际上是有相互作用的,它们通过库仑力相互排斥,此外当这两个电子在同一位置时,它们所具有的能量必定与只有一个电子在那里时的能量的两倍有很大的差别。当只有 6 个位置,特别是当我们要放进第 6 个电子的时候,独立粒子的近似确实是不合理的。但是,有机化学家们已经能从这种近似中获得许多知识。

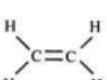


图 15-5 乙烯分子

在我们对苯分子作出详细计算之前,让我们考虑一个更简单的例子——乙烯分子,它只有两个碳原子,在每边都各有两个氢原子,如图 15-5 所示。这分子有一个含有两个电子的“额外”的价键,处在两个碳原子之间。现在去掉这两个电子中的一个,我们得到什么呢?我们可以把它看成是一个两态系统——留下的一个电子可以位于这个碳原子或者另一个碳原子旁。我们可

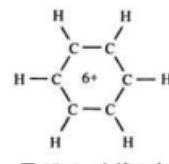


图 15-4 去掉 6 个电子的苯环

以把它作为一个双态系统来分析。这单个电子可能具有的能量是 $(E_0 - A)$, 或 $(E_0 + A)$, 如图 15-6 所示。

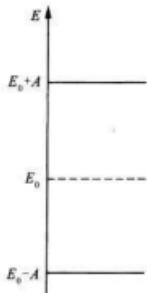


图 15-6 乙烯分子中“额外”电子可能的能级

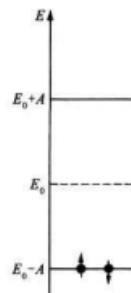


图 15-7 在乙烯分子的额外价键中,可以有两个电子
(一个自旋向上,一个自旋向下)占据最低能级

现在我们加进第二个电子。好,如果我们有两个电子,我们可以把第一个电子放在较低的状态,而把第二个电子放在较高的状态。不太对,我们忘记了某些事情。每一个状态实际上是双重的。当我们说存在一个能量为 $(E_0 - A)$ 的可能状态时,实际上有两个。如果一个电子的自旋向上而另一个电子的自旋向下的话,则这两个电子可以进入同一个态。(由于不相容原理,不能再放入更多的电子。)所以实际上能量为 $(E_0 - A)$ 的可能状态有两个。我们可以画一个图,如图 15-7 所示,它既表明能级,又表明这些能级被占据的情况。在最低能量条件下,两个电子以相反的自旋一同处在最低状态。如果我们忽略两个电子的相互作用,那么乙烯分子的额外价键的能量是 $2(E_0 - A)$ 。

现在让我们回过头来讨论苯。图 15-3 中的两个态的每一个都有 3 个双键。每个双键都像乙烯中的键,每个双键贡献 $2(E_0 - A)$ 的能量。其中 E_0 是把一个电子放入苯中某个位置的能量,A 是这个电子跃迁到相邻位置上的能量。所以能量大致应为 $6(E_0 - A)$ 。但是我们在以前研究苯分子时,得出的能量比有 3 个额外价键结构的能量更低。让我们来看看,根据我们的新观点,是否苯分子的能量比 3 个价键的更低。

我们从六重电离的苯环开始讨论,并加进一个电子。现在我们有一个 6 个态的系统。虽然我们还没有解过这类系统的问题,但是我们知道怎样去解,我们可以写出含有 6 个振幅的 6 个方程等等。但是让我们省一些劳动——我们注意到:当我们算出无限原子线列中的一个电子问题时,我们已经解决了这个问题。当然,苯并不是无限长的直线,它有 6 个原子组成一个环。但是我们设想打开这个环,把它拉成一直线,并且将沿此直线的原子从 1 到 6 予以编号。在无限长直线的情况下,下一个位置应是 7。但是如果我们将第七个位置和第一个位置全同,并以此类推,那么情况就和苯环完全相同了。换句话说,我们可以使用无限长直线的解,但有一个附加的要求是这个解具有 6 个原子长的周期性。在第 13 章中我们已知道,处于一直线上的电子,当它在每个位置上的振幅为 $e^{ikx_k} = e^{ikn}$ 时,它就具有确定能量的状态。对于每个 k ,能量为

$$E = E_0 - 2A \cos kb. \quad (15.25)$$

我们现在只要那些每 6 个原子就重复的解。让我们先讨论 N 个原子的环这种一般的情况。如果这个解具有 N 个原子长度的周期性，则 $e^{i\theta N}$ 必定为 1，或者， kbN 必定是 2π 的整数倍。令 s 表示任一整数，于是上述条件为

$$kbN = 2\pi s. \quad (15.26)$$

以前我们就已知道，取土 π/b 范围以外的 k 值是毫无意义的。这意味着：我们取土 $N/2$ 范围内的 s 值就得到所有可能的状态。

我们发现，对于这 N 个原子的环来说，存在 N 个确定能量的状态*，这些态的波数由下式给出

$$k_s = \frac{2\pi}{Nb} s, \quad (15.27)$$

每个态的能量由式(15.25)给定。我们得到了所有可能能级的一组线状谱。苯 ($N = 6$) 的能谱如图 15-8(b) 所示。(括号里的数字表示具有相同能量的不同状态数目。)

有一个使这 6 个能级形象化的好方法，如图 15-8(a) 所示。想象一个圆，其圆心在能量为 E_0 的能级上，半径为 $2A$ 。如果从这圆的底部开始量出 6 个相等的弧（在从底点量起角度为 $kb = 2\pi s/N$ ，对苯来说为 $2\pi s/6$ ），那么图上这些点的垂直高度就是式(15.25)的解。6 个点代表 6 个可能的状态。最低能级为 $(E_0 - 2A)$ ，有两个态具有相同的能量 $(E_0 - A)$ ，等等**。这些都是一个电子可能具有的状态。如果电子数超过 1 个，那么每个态可容纳具有相反自旋方向的两个电子。

对于苯分子来说，我们必须放进 6 个电子。对基态来说，它们要进入可能最低的能量状态——两个电子处在 $s = 0$ ，两个电子处在 $s = +1$ ，另两个电子处在 $s = -1$ 。按照独立粒子近似，该基态能量为

$$E_{\text{基}} = 2(E_0 - 2A) + 4(E_0 - A) = 6E_0 - 8A. \quad (15.28)$$

这个能量的确比 3 个分开的双键能量要少 $2A$ 。

将苯和乙烯的能量加以比较，就可确定 A 。结果为 0.8 eV，或者用化学家们喜欢的单位，每摩尔 18 千卡（1 千卡 = 4 186.8 J）。

我们可以用上面的叙述来计算或了解苯的其他性质。例如，利用图 15-8，我们就可以讨论苯被光激发的情形。如果我们试着去激发苯中的一个电子，那会出现什么情况呢？这个电子可以跳到一个空着的较高的能量状态上去。最低激发能量应是从最高的被占据能级

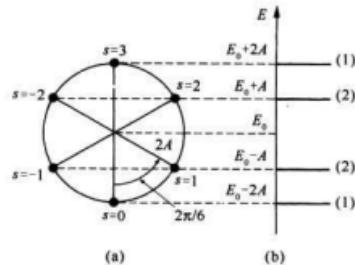


图 15-8 具有 6 个电子位置的环的能级(例如: 苯环)

* 你们可能以为当 N 是偶数时会有 $N + 1$ 个态。其实并非如此，因为 $s = \pm N/2$ 给出相同的态。

** 如果有两个态（它们具有不同的振幅分布）具有相同的能量，我们称这两个态是“简并”的。要注意，可以有 4 个电子具有能量 $E_0 - A$ 。

跃迁到最低空能级的能量,这需要能量 $2A$ 。当 $\hbar\nu = 2A$ 时,苯将吸收频率为 ν 的光。苯也会吸收能量为 $3A$ 和 $4A$ 的光子。不用说,苯的吸收光谱已经测量过了,谱线图多少与我们的分析相符,只是最低的跃迁在紫外区。为了和实验数据相吻合, A 的值应选取在 $1.4 \sim 2.4$ eV 之间。这就是说, A 的数值比根据化学结合能预料的要大两到三倍。

对于这种情况,化学家的处理方法是:分析许多同类的分子,从中得出一些经验规则。例如,他知道:为了计算结合能,要用某一个 A 值,但为了得出大致正确的吸收光谱要用另外一个 A 值。你们可能感到这听起来有点荒谬。从试图根据第一原理来理解自然界的物理学家的观点来说,这种方法不是很令人满意的。但是化学家面临的问题却不同。对于目前尚未制造出来、或者尚未完全了解的分子,化学家在事前必须试着对它们的性质进行猜测。他所需要的是一系列经验规则,至于这些规则是哪里来的则无关紧要。所以他应用理论的方式与物理学家很不相同。他所选用的方程式中含有真理的影子,但是他必须改变其中的常数——作出经验修正。

就苯来说,与实验不一致的主要原因是假设电子是独立的——我们作为出发点的理论实际上并不合理。然而,它包含着某些真理的影子,因为它的结果似乎是对路的。有机化学家在选作研究对象的那些复杂事物的困境中,就是利用这样一些方程再加上某些经验规则(包括各种例外情况)奋力进行的。(不要忘记,一个物理学家之所以能够从第一原理出发真正计算一些东西,是因为选择的都只是一些简单问题,他从来没有解决过那些具有 42 个或者即使是只有 6 个电子的问题。至今,物理学家所能作出精确、合理的计算的仅仅是氢原子和氦原子。)

§ 15-5 其他有机化学分子

现在让我们来看看,怎样才能用这些概念来研究其他分子。考虑像丁二烯($1,3$)这样的分子,在图 15-9 中按照通常的价键图画出了这个分子。

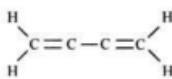


图 15-9 丁二烯($1,3$)分子的价键图



图 15-10 N 个分子排成一直线

对于与两个双键相对应的 4 个额外电子,我们可以用同样的处理方法。如果我们去掉这 4 个电子,就剩下位于一直线上的 4 个碳原子。你们已经知道如何解决一条直线的问题了。你们会说,“啊,不,我只知道如何解无限长直线的问题。”但是无限长线的解也包括有限长线段的解。注意,设在一直线上的原子数目为 N ,并将原子按 1 到 N 编号如图 15-10 所示。在写关于位置 1 的振幅方程中你可以不写位置 0 提供的项。同样,有关位置 N 的振幅方程也和无限长情况的方程不同,因为现在不存在由位置 $N+1$ 提供的任何项。但是,这要假定我们能够得到具有下述性质的无限长线的解:位于原子 0 处的振幅为零,位于原子 $N+1$ 处的振幅也为零。那么这些解就能满足有限长线上从 1 到 N 的所有位置上的方程组。你

们可能认为对无限长线不存在这样的解,因为我们的解都是 e^{ikx} 那种样式,它在每个地方都有相同的振幅绝对值。但是你得记住:能量只和 k 的绝对值有关。因此对于相同的能量, e^{-ikx} 是另一个同样合理的解。这两个解的任何叠加也具有相同的能量。把上面的两个解相减,我们可以得到解 $\sin kx_n$, 它满足在 $x = 0$ 处振幅为 0 这一要求。这个解仍对应于能量 ($E_0 - 2A \cos kb$)。现在只要适当选取 k 值, 我们就可使振幅在 x_{N+1} 处也为零。这要求 $(N+1)kb$ 是 π 的整数倍, 或者

$$kb = \frac{\pi}{(N+1)} s, \quad (15.29)$$

式中 s 是从 1 到 N 的整数。(这里我们只选取正的 k 值, 因为每个解都包括 $+k$ 和 $-k$; 改变 k 的符号又给出同样的态。) 对于丁二烯分子来讲,

$N = 4$, 所以有 4 个态, 其 k 值由下式决定

$$kb = \frac{\pi}{5}, \frac{2\pi}{5}, \frac{3\pi}{5} \text{ 和 } \frac{4\pi}{5}. \quad (15.30)$$

我们也可用一个与苯的情形相似的圆来表示能级。这次我们用分成 5 个相等部分的半圆, 如图 15-11 所示。位于底部的点对应于 $s = 0$, 它根本没有给出什么态, 顶上对应于 $s = N + 1$ 的一点也是如此。剩下的 4 点给出了 4 个允许的能量。一共有 4 个定态, 我们从 4 个基础态出发所期望的正是这一情况, 在圆图中, 各个角度间隔为 $\pi/5$ 或 36° 。最低能量为 $(E_0 - 1.618A)$ 。(啊, 多奇妙的数学方法! 按照这个理论, 希腊人*的黄金分割给了我们丁二烯分子的最低能量态!)

现在我们就可以计算出放进 4 个电子后丁二烯分子的能量。我们用 4 个电子填满最低的两个能级, 每个能级上有两个自旋相反的电子。总能量为

$$E = 2(E_0 - 1.618A) + 2(E_0 - 0.618A) = 4(E_0 - A) - 0.472A. \quad (15.31)$$

这个结果看来是合理的。这能量比两个简单双键的能量略低一些, 但是结合力没有苯那样强。不管怎么说, 这是化学家分析某些有机分子的方法。

化学家不但能够利用能量, 而且能够利用概率幅来分析问题。知道了每个态的振幅以及哪些态被占据, 他们就能说出在分子中任何地方找到一个电子的概率。电子最可能出现的那些地方, 是适合于需要和其他原子团共有一个电子的化学取代中发生反应的地方, 别的位置则更可能在取代反应中有产生额外电子给系统的倾向。

上面的这些概念也能使我们对叶绿素这样复杂的分子也有所了解, 图 15-12 所示为叶绿素的一种结构。注意, 我们用粗线画出的双键和单键构成了具有 20 个间隔的大闭合环, 双键上的额外电子可以沿着这个环运动。用独立粒子方法, 我们可以得到一整套能级, 在光

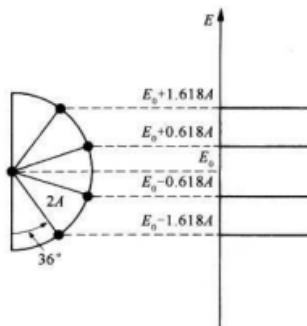


图 15-11 丁二烯的能级

* 如果一个长方形可以分割成一个正方形再加上一个和原形相似的长方形, 那么这个长方形的两个邻边之比就称为黄金分割。

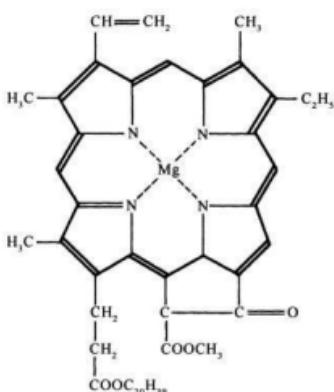
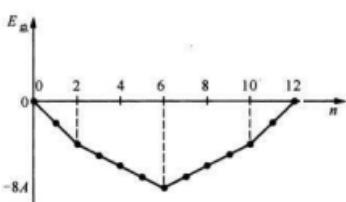
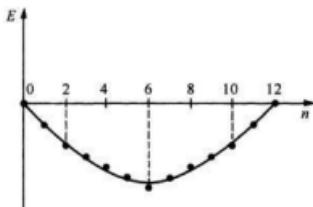


图 15-12 一个叶绿素分子

是一条直线，对于相继放入的每一组电子，斜率增加，而各组电子之间的斜率是不连续变化的。当我们刚好填满了一组能量相同的能级，而必须把电子填到下一组能量较高的能级上去的时候，斜率就要改变。

实际上苯离子的能量与图 15-13 的曲线有很大的差异，这是因为有被我们忽略掉的电子间的相互作用和静电能存在。然而，这些修正将以颇为平稳的方式随 n 而变化。即使我们对所有这些都进行了修正，所得出的能量曲线在正好填满某一能级的 n 值处仍有拐折。

图 15-13 当图 15-8 中的最低能态被 n 个电子占有时，所有电子能量的和(取 $E_0 = 0$)图 15-14 用一条光滑曲线联结图 15-13 的点。具有 $n = 2, 6, 10$ 的分子比别的态更稳定

现在考虑适合于各点平均的一条非常光滑的曲线，如图 15-14 所示。我们可以说，这曲线上面的各点具有“比正常值高”的能量，而曲线下面各点具有“低于正常值”的能量。通常，我们可以预料，低于正常能量的那些组态有较高的平均稳定性——从化学上说。注意，那些远低于曲线的组态总是出现在直线线段的端点，换句话说，总是出现在有足够的电子正好填满所谓的“能量壳层”的情况下。这是理论很准确的预言。当分子或离子中可供填充的电子刚好填满能量壳层时，这个分子或离子(与其他类似的组态相比较)就显得特别稳定。

这个理论解释和预言了一些非常奇特的化学事实。作为一个非常简单的例子，我们考

虑一个由3个碳原子构成的环。几乎很难令人相信，化学家能够制造出三环并且它是稳定的，但是的确已经制成了。3个电子的能量圆画在15-15中。现在如果你们把两个电子放至较低能态中，这只是3个所需的电子中的两个。第三个电子必须放在高得多的能级上，根据我们的论证，这个分子不应该特别稳定，而双电子结构应该是稳定的。事实上：中性三苯基环丙烷基分子确实很难制成，但是图15-16所示的正离子就相对地比较容易制造出来。实际上制造三环绝非易事，这是因为当一个有机分子中的键构成一个等边三角形时，总是有很大的应力。要制造一个稳定的化合物，就必须用某种方法使其结构稳定。不管怎样如果在3个角上加上3个苯环，那么就能够构成一个正离子。（加苯环的理由至今还没有真正搞懂。）

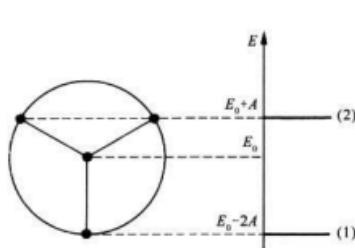


图 15-15 三个原子环的能量图

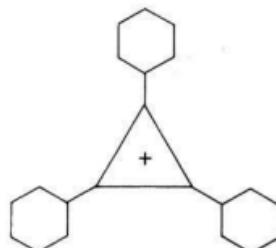


图 15-16 三苯基环丙烷基正离子

具有5个边的环也可以用类似的方法来分析。如果你们画出它的能级图，就可以定性地看出6个电子的结构应是一种特别稳定的结构，所以这样一种分子应该在成为负离子时最为稳定。现在5环已为大家熟知，也容易制造，并且总是作为负离子存在。同样，你们可以很容易地证实：4环或8环并不令人很感兴趣，但是，14环或10环——类似6环——的中性分子也应该是特别稳定的。

§ 15-6 近似方法的其他应用

我们将对另外两种类似的情况作些简单的描述。在考虑原子的结构时，我们认为电子依次填充各个壳层。电子运动的薛定谔理论只有对单个电子在有心力场中运动的情形才能较容易地求解，所谓有心力场就是随到某一点的距离变化而变化的场。那么我们怎样才能知道在一个具有22个电子的原子中所发生的情况呢？！一种办法就是利用一种独立粒子近似法。首先你们计算只有一个电子时的情形。你会得到若干能级。把一个电子放到最低的能量状态。作为一个粗糙的模型，继续忽略电子间的相互作用，而把电子依次填入相继的壳层。但是有一种获得更好的答案的方法是计入——至少以一种近似的方式——电子所带电荷的影响。你们每加一个电子就计算该电子处在各个不同位置的振幅，然后用此振幅去估算出一种球对称的电荷分布。用这种分布的场——加上带正电的原子核以及前面加进去的那些电子所产生的场——来计算适合下一个电子的态。这样，你就可以对中性原子以及各种离子态的能量作出合理而正确的估计。你们会发现这里也有能量壳层，就像我们曾看到的环形分子中的电子那样。当原子具有未全部填满的壳层时，该原子就表现出倾向于要获

取一个或更多的额外电子,或者倾向于释出一些电子,以使它进入满壳层的最稳定的状态。

这个理论解释了元素周期表中所反映出来的隐含在基本化学性质中的机理。惰性气体就是具有刚好填满的壳层的元素,特别难以使它们发生化学反应。(当然有些惰性气体还是会同例如氟和氯等发生反应,但是这些化合物的结合非常弱,所谓的惰性气体只是近乎惰性而已,)比惰性气体原子多一个或少一个电子的原子很容易失去或者得到一个电子而进入特别稳定的(低能)状态,稳定状态来源于原子具有完全填满的壳层,这些原子是非常活泼的+1价或-1价元素。

另一种情况出现在原子核物理学中。在原子核中,质子和中子之间的相互作用相当强。即使如此,在分析原子核的结构时仍然可以用独立粒子模型。首先在实验上发现:如果原子核含有某些特殊数目——2, 8, 20, 28, 50, 82——的中子时,这些核特别稳定。含有这些数目质子的核也特别稳定。由于起先对这些数字无法作出解释,因此把它们称为原子核物理学中的“幻数”。众所周知,中子和质子间存在着很强的相互作用,因此,当人们发现用独立粒子模型预言的壳层结构得出了前面几个幻数时,感到非常惊奇。这个模型假定每个核子(质子和中子)在一个有心势场中运动,这个有心势场是由所有其他核子的平均效应产生的。但是这个模型不能给出较高幻数的正确值。后来迈耶(Maria Mayer),以及詹森(Jensen)和他的同事们各自独立地发现,只要将独立粒子模型加上所谓“自旋-轨道相互作用”,就成为一个能给出所有幻数的改进模型。(如果一个核子的自旋与它在核内运动的轨道角动量的方向相同,则自旋-轨道相互作用将使这个核子的能量降低。)该理论甚至能给出更多的东西——这个所谓原子核的“壳层结构”图像能使我们预言原子核及核反应的某些特性。

独立粒子近似在广泛的学科领域中——固态物理学,化学,生物学,原子核物理学——都得到了应用。虽然它往往只是一种粗略的近似,但是它能使我们了解为什么在壳层里会存在特别稳定的条件。由于其忽略了各单个粒子间相互作用的所有复杂性,因此,它往往完全不能正确地给出许多重要的细节,这是不足为奇的。

第 16 章 振幅对位置的依赖关系

§ 16-1 一维情形的振幅

现在我们将议论量子力学的概率幅在空间是如何变化的。在前面几章中，你们可能对有些事情被忽略感到不太舒服。例如，当我们谈论氯分子时，我们选取了两个基础态来描述它。我们选择氯原子处于 3 个氢原子所构成的平面之“上”的情形作为一个基础态，氯原子处于 3 个氢原子所构成的平面之“下”的情形作为另一个基础态。为什么我们仅仅挑选这样两个基础态呢？为什么氯原子就不可能处于 3 个氢原子所构成平面之上 2\AA ($1\text{\AA} = 10^{-10}\text{ m}$) 处，或者 3\AA 或 4\AA 处呢？无疑有许多位置可供氯原子占据。又如当我们谈到氢分子离子时（在此离子中两个质子共有一个电子），我们想象了两个基础态：一个态是电子在第 1 号质子附近，另一个态是电子在第 2 号质子附近。显然我们略去了许多细节。这个电子并不是正好位于第 2 号质子处，而只是在它附近，它可能在这个质子上面的某个地方，或者下面、左面、右面的某个地方。

以前我们故意避免讨论这些细节。我们说，我们只是对问题的某些特征感兴趣，所以我们想象当该电子在第 1 号质子附近时，它就处于某种相当确定的状态。电子处于这个状态时，在质子周围找到它的概率应有某种相当确定的分布，但是我们对此细节不感兴趣。

我们也可以用另一种方法处理这个问题。在我们关于氢分子离子的讨论中，当我们用两个基础态来描述这种情况时，我们采用的是一种近似描述法。实际上存在着许许多多这样的状态。处于质子周围的电子可以占据最低的能量状态，即基态，但是也还存在着许多激发态。对于各个激发态来说，电子在质子周围的分布各不相同。过去我们忽略了这些激发态，说是感兴趣的只是低能量的情形。然而正是这些激发态提供了电子在质子周围各种不同分布的概率。如果我们希望详细地描述氢分子离子，我们就必须把这些可能的基础态也考虑进去，要做到这一点可以有几种方法，一种方法就是去详细地考虑那些较仔细地描述电子在空间位置的状态。

现在我们准备考虑一个比较复杂的程序，通过给出在一定条件下在任何地方找到电子的概率幅，我们就能够详细地讨论电子的位置。这个比较完整的理论为我们以前讨论中所用的近似方法提供了基础。在某种意义上，我们可以把前面的方程作为这种更完整的理论的一种近似而推导出来。

你们可能会感到奇怪，为什么我们不一开始就采用较完整的理论，在讨论过程中再作近似处理。我们感到，从双态近似开始，逐步建立更完善的理论，比起用相反的方式来处理这一问题，更容易使你们理解量子力学的基本机理。正是由于这一原因，我们处理问题的次序和你们在许多书上看到的相反。

当我们深入研究本章的课题时，你们将会注意到我们正在打破我们以往一直遵循的一

个规则。无论我们在研究哪一个课题时,总是试图或多或少给出完善的物理描述——尽可能详细地告诉你们这观念通向何处。我们一直既试图去描述理论的一般结果,又试图描述某些具体的细节,从而使你们能看出理论会把我们引向何方。我们现在要打破这个规则,将叙述怎样才能谈论空间的概率幅,并向你们介绍它们所满足的微分方程,而没有时间去继续讨论由这个理论导出的许多明显包含的内容。的确,我们甚至不可能更进一步将这个理论与我们早先使用的某些近似公式(例如关于氢分子和氦分子的近似公式)联系起来。就这一次,我们只好不把讨论进行到底。我们的课程即将结束,只能满足于试图向你们介绍一般的概念以及指出我们至今一直在描述的方法和研究量子力学课题的其他方法之间的关系。我们希望给你们足够的观念,使你们能够自己进行学习,并通过阅读各种书籍学到与我们即将描述的问题有关的许多含义。我们毕竟得留些内容将来再学。

让我们再来复习一下已求出的关于一个电子如何沿原子线列运动的问题。当一个电子具有从一个原子跃迁到相邻原子的振幅时,存在着具有确定能量的状态,在这些态中找到电子的概率幅是以行波的形式沿晶格分布的。对于长波长——即对于波数 k 之值小的情况——态的能量与波数的平方成正比。对于间隔为 b 的晶格,其中电子在单位时间内从一个原子跃迁到相邻原子的振幅为 iA/\hbar , 态的能量与 k (对于小的 kb)的关系为

$$E = Ak^2 b^2. \quad (16.1)$$

(参见 § 13-2。)我们也已知道,具有相似能量的这样一群波会形成一个波包,这个波包的行为与一个具质量 $m_{\text{有效}}$ 的经典粒子很相似。 $m_{\text{有效}}$ 为

$$m_{\text{有效}} = \frac{\hbar^2}{2Ab^2}. \quad (16.2)$$

既然晶体中的概率幅波行为像一个粒子,我们完全可以期望:一个粒子的一般量子力学描述应显现出我们对晶格观察到的相同类型的波的行为。假定我们考虑的晶格在直线上,并且想象晶格间隔 b 越来越小。在极限情况下我们会想到电子可以处在该直线的任何地方。我们就过渡到概率幅连续分布的情况。这样我们就有沿此直线在任何地方找到一个电子的振幅。这是描述一个电子在真空中运动情况的一种方法。换句话说,如果设想空间能够用无限个非常靠近的点标记出来,并能求得表示一点的振幅与相邻点振幅之间关系的方程,我们就得到了一个电子在空间运动的量子力学定律。

让我们先回想一下某些量子力学的一般原理。假定一个粒子可以存在于一个量子力学系统各种不同的条件下,我们把任何一个可以在其中找到电子的特殊情况称为“态”,并用态矢量例如 $|i\rangle$ 来表示。其他一些态可以用别的态矢量,例如 $|\phi\rangle$ 来表示。我们然后引进基础态的概念,说有一组态 $|1\rangle, |2\rangle, |3\rangle, |4\rangle$ 等等,它们具有如下性质:

第一,所有这些态都是完全不同的——我们说它们是正交的。所谓正交意思是指对于任意两个基础态 $|i\rangle$ 和 $|j\rangle$, 表示一个已知处在态 $|i\rangle$ 的电子同时又处在态 $|j\rangle$ 的振幅 $\langle i|j\rangle$ 等于零——当然除非 $|i\rangle$ 和 $|j\rangle$ 代表同一个态。我们将此符号表示为

$$\langle i|j\rangle = \delta_{ij}. \quad (16.3)$$

你们应记住,若 i 与 j 不等,则 $\delta_{ij} = 0$; 若 i 和 j 相等,则 $\delta_{ij} = 1$ 。

第二,这些基础态 $|i\rangle$ 必定是一个完全集,所以任何态都能用它们表述。这就是说,任何态 $|\psi\rangle$ 都可以由给定的所有振幅 $\langle i|\psi\rangle$ 完整地描述。 $\langle i|\psi\rangle$ 是处在态 $|i\rangle$ 的一个粒子也会在态

$|i\rangle$ 找到的振幅。事实上,态矢量 $|\phi\rangle$ 等于各基础态乘上一个系数后的和,这系数就是态 $|\phi\rangle$ 也处在态 $|i\rangle$ 的振幅,即

$$|\phi\rangle = \sum_i |i\rangle \langle i | \phi \rangle. \quad (16.4)$$

最后,如果我们考虑任何两个态 $|\phi\rangle$ 和 $|\psi\rangle$,那么要求态 $|\psi\rangle$ 也同时处在态 $|\phi\rangle$ 的振幅,可以先将态 $|\phi\rangle$ 投影到各基础态上,然后再将每个基础态投影到态 $|\psi\rangle$ 上而求得。我们把这写成如下形式:

$$\langle \phi | \psi \rangle = \sum_i \langle \phi | i \rangle \langle i | \psi \rangle. \quad (16.5)$$

式中的求和当然是对整个基础态 $|i\rangle$ 的集进行的。

在第13章中,当我们计算位于一直线排列的原子中的电子情况时,我们曾选取过一组基础态,在这些态中电子总被定域于直线上的这个或那个原子上。基础态 $|n\rangle$ 代表电子被定域在第“ n ”个原子上的情形。(当然,我们把基础态称为 $|n\rangle$ 而不称为 $|i\rangle$ 并没有什么特别的意义。)接着,我们发现用原子的坐标 x_n 要比用原子在排列中的数字编号表示基础态来得方便。态 $|x_n\rangle$ 只是书写态 $|n\rangle$ 的另一种方法。然后,遵循一般规则,任何态,比方说 $|\phi\rangle$,处于 $|\psi\rangle$ 态的电子也可以用在态 $|x_n\rangle$ 中的一个的振幅来描述。为方便起见,我们选用符号 C_n 代表这些振幅

$$C_n = \langle x_n | \phi \rangle. \quad (16.6)$$

既然,这些基础态与直线上的位置有关,我们可把振幅 C_n 看作是坐标 x 的函数,并把它写成 $C(x_n)$ 。一般来说,振幅 $C(x_n)$ 将随时间而改变,因此它也是 t 的函数。通常我们并不费力气去把这种对时间的依赖关系明显地表示出来。

于是我们在第13章中提出:振幅 $C(x_n)$ 将以哈密顿方程[式(13.3)]所描述的方式随时间变化。用我们的新符号,该方程为

$$i\hbar \frac{\partial C(x_n)}{\partial t} = E_0 C(x_n) - AC(x_n + b) - AC(x_n - b). \quad (16.7)$$

上式右边最后两项代表位于第 $(n+1)$ 个原子或第 $(n-1)$ 个原子的电子可以输送至第 n 个原子的过程。

我们发现式(16.7)具有与确定能量的状态相对应的解,我们把这些解写成

$$C(x_n) = e^{-iE_0 n / \hbar} \cdot e^{ikx_n}. \quad (16.8)$$

对于那些低能量的状态来说,波长较长(k 较小),且能量与 k 的关系为

$$E = (E_0 - 2A) + Ak^2 b^2, \quad (16.9)$$

或者,我们这样选取能量的零点,使 $E_0 - 2A = 0$ 。那么能量就由式(16.1)给出。

现在我们来看看,要是让晶格间隔 b 趋于零,但保持波数 k 不变,可能会发生些什么。如果所发生的只是上述情况,则式(16.9)中的最后一项正好为零,那就没有物理意义了。但是假如 A 和 b 一起变化,以至当 b 趋于零时乘积 Ab^2 保持恒定*——应用式(16.2),我们将

* 你们可以想象,当这些点 x_n 相互趋近时,从 x_{n+1} 跃迁到 x_n 的振幅 A 将增加。

把 Ab^2 写成常数 $\hbar^2/2m_{\text{有效}}$ 。在这些情况下, 方程式(16.9)将不发生变化, 但是微分方程式(16.7)会出现什么情况呢?

首先, 我们把方程式(16.7)改写成

$$i\hbar \frac{\partial C(x_n)}{\partial t} = (E_0 - 2A)C(x_n) + A[2C(x_n) - C(x_n + b) - C(x_n - b)]. \quad (16.10)$$

对于我们所选取的 E_0 , 上式第一项为零。其次, 我们可以想象一个连续函数 $C(x)$, 它平滑地联接每个 x_n 处 $C(x_n)$ 的固有值。当间隔 b 趋于零时, 这些点 x_n 越来越紧密地靠在一起, 而(如果我们使 $C(x)$ 的变化相当平滑的话)在括号中的量正好同 $C(x)$ 的两阶微商成正比, 将每一项作泰勒展开后, 我们可以写出下列等式

$$2C(x) - C(x + b) - C(x - b) \approx -b^2 \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.11)$$

然后, 在取 b 趋于零的极限条件下, 保持 $b^2 A$ 等于 k , 则式(16.7)变成

$$i\hbar \frac{\partial C(x)}{\partial t} = -\frac{\hbar^2}{2m_{\text{有效}}} \cdot \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.12)$$

我们就得到了一个方程, 它告诉我们: $C(x)$ ——在 x 处找到电子的振幅——的时间变化率与在相邻点找到电子的振幅有关, 它正比于该振幅对位置的两阶微商。

关于电子在自由空间运动的正确的量子力学方程最早是由薛定谔发现的。对于沿一直线运动的电子来说, 如果用该电子在自由空间的质量 m 代替 $m_{\text{有效}}$, 则正确的量子力学方程就和式(16.12)的形式完全一样。在自由空间中沿一直线运动的电子, 薛定谔方程为:

$$i\hbar \frac{\partial C(x)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 C(x)}{\partial x^2}. \quad (16.13)$$

我们无意使你们认为我们已经导出了薛定谔方程, 而只是希望告诉你们一种思考这个问题的方法。当薛定谔第一次写出这个方程式的时候, 他的推导是建立在启发性的论证和卓越的直觉猜测上的。他所使用的某些论证甚至是错误的, 但是这并没有什么关系, 唯一重要的是这个最终的方程给出了对自然界的正确描述。我们讨论这问题的目的, 只是想告诉你们, 正确的、基本的量子力学方程式(16.13)和你们在电子沿着原子线列运动的极限情况下所得到的方程具有相同的形式。这意味着我们可以把式(16.13)中的微分方程看作是描述概率幅沿此直线从一点到相邻一点的扩散。也就是说, 如果一个电子在某一点具有某个振幅, 则过一会, 它会在邻近一点具有某个振幅。事实上, 这个方程看起来有点像我们在第1卷中所用的扩散方程。但是存在着一个主要的区别: 时间微商前的虚系数使得此方程的性质与普通的扩散(诸如气体沿一细管扩散的情况)完全不同。普通扩散给出实指数的解, 而式(16.13)的解却是复波。

§ 16-2 波 函 数

现在你们对要讨论的将是些什么问题已有了些概念, 我们想从头再来研究描述一个电子沿一直线运动的问题, 而不去考虑与晶格上原子有关的态。我们要从头开始来看看: 如

果要描述一个自由粒子在空间的运动,我们必须用些什么概念。既然我们对沿一连续直线运动的粒子的行为感兴趣,那就要处理无数个可能的态,你们将会看到,对于我们为处理有限数目的态而发展起来的概念需要作出某些技术性的修正。

我们从设态矢量 $|x\rangle$ 代表一个粒子精确地定位在坐标 x 处的态着手。对于沿直线的每一个 x 值,例如 1.73 或 9.67 或 10.00,都有一个相应的态,我们将取这些态 $|x\rangle$ 作为基础态,而且,要是包括直线上的所有点,我们就有了一维运动的一个完全集。现在假定我们有一个不同类型的态,譬如 $|\psi\rangle$,在该态中一个电子以某种方式沿此直线分布。描述这种态的一个办法是:给出在每一个基础态 $|x\rangle$ 找到电子的所有振幅。每个 x 值一个振幅,我们必须给出这些振幅的一组无限集合。我们把这些振幅写成 $\langle x|\psi\rangle$ 。每个这样的振幅都是一个复数,由于对每个 x 值都有这样一个复数,所以振幅 $\langle x|\psi\rangle$ 必定是 x 的函数,我们也把它写成 $C(x)$,

$$C(x) \equiv \langle x|\psi\rangle. \quad (16.14)$$

当我们在第 7 章中谈到振幅随时间的变化时,我们已经考虑过这种以连续的方式随坐标变化的振幅。例如,我们在那里曾证明,具有确定动量的一个粒子,它在空间的振幅必定具有特定的变化形式。如果一个粒子具有确定的动量 p 及相应确定的能量 E ,则在任何位置 x 找到该粒子的振幅就像

$$\langle x|\psi\rangle = C(x) \propto e^{ipx/\hbar}. \quad (16.15)$$

这个方程式表达了量子力学的一个重要的普通原理,它把与空间不同位置相对应的基础态与另一个基础态系统,即所有具有确定动量的态联系了起来。对于某些类型的问题来说,使用具有确定动量的态往往比用 x 表示的态更加方便。当然,对于描述量子力学情况,这两组基础态都同样适用。我们在后面还要回过头来讨论这两种态之间的关系。目前我们希望把讨论限于用态 $|x\rangle$ 来描述的方法。

在作进一步讨论之前,我们要稍微改变一下符号,并希望不至于引起太大的混淆。式(16.14)定义的函数 $C(x)$ 的表示式当然与我们所考虑的那个态 $|\psi\rangle$ 有关,我们应该用某种方式把这种关系表示出来。我们可以用一个下标,例如 $C_\psi(x)$,来表明我们所讨论的是哪一个函数 $C(x)$ 。虽然这将是一个完全令人满意的符号,但是它有点麻烦,并且在大多数书中都不采用它。大多数人就直接省去字母 C ,而只用符号 ψ 来定义这个函数

$$\psi(x) \equiv C_\psi(x) = \langle x|\psi\rangle. \quad (16.16)$$

既然世界上其他人都使用这个符号,你们也会习惯它,你们在别的地方遇到它时也不会被吓住。然而要记住,我们现在有两种不同的使用 ψ 的方式。在式(16.14)中 ψ 代表我们给予电子的某一特定物理状态的标记。另一方面,在式(16.16)的左边,符号 ψ 被用来定义一个 x 的数学函数,它等于与沿直线上每一点 x 相联系的振幅。我们希望你们一旦习惯了这个概念后不会引起太多的混淆。顺便指出:函数 $\psi(x)$ 通常称为“波函数”——因为它多半具有其变量的复波形式。

既然我们已经把 $\psi(x)$ 定义为处在态 ψ 的电子在位置 x 处的振幅,那么就希望把 ψ 的绝对值平方解释为在位置 x 处找到一个电子的概率。遗憾的是,严格说来在任何特定位置上找到一个电子的概率是零。一般说来,电子将“弥散”在线上的某个区域中,而且,由于任何一小线段上存在无数个点,因此电子在任何一点的概率不可能是有限数值。我们

只能用概率分布^{*}来描述找到一个电子的概率，概率分布给出了在直线上各个位置附近找到这个电子的相对概率。让我们用 $\text{prob}(x, \Delta x)$ 这个符号来代表位于 x 附近很小间隔 Δx 中找到电子的机会。如果我们采用在任何物理情况下都足够小的尺度，则概率就会随位置平滑地变化，而在任何小的有限线段 Δx 中找到电子的概率将与 Δx 成正比，我们可以按此考虑来修改我们的定义。

我们可以把振幅 $\langle x | \psi \rangle$ 看作是表示所有基础态 $|x\rangle$ 在一小范围内的一种“振幅密度”。既然在 x 处的小间隔 Δx 中找到一个电子的概率应该与间隔 Δx 成正比，我们可以这样来选取 $\langle x | \psi \rangle$ 的定义，使下述关系成立：

$$\text{prob}(x, \Delta x) = |\langle x | \psi \rangle|^2 \Delta x.$$

因此振幅 $\langle x | \psi \rangle$ 就与处在态 $|\psi\rangle$ 的一个电子在基础态 $|x\rangle$ 中被找到的振幅成正比，而比例常数这样选取，使得振幅 $\langle x | \psi \rangle$ 的绝对值平方给出在任何小范围内找到一个电子的概率密度。我们可以等效地写成

$$\text{prob}(x, \Delta x) = |\psi(x)|^2 \Delta x. \quad (16.17)$$

现在我们必须对前面的某些方程作些修改，使它们与这个新定义相协调。假定一个电子处在 $|\psi\rangle$ 态，而我们希望知道在另一个不同态 $|\phi\rangle$ 中找到此电子的振幅，态 $|\phi\rangle$ 可能对应于该电子的另一个不同分布状况。以前当我们谈论一组有限的分立状态时，我们就用式 (16.5)。在修改我们的振幅定义前我们把这振幅写成

$$\langle \phi | \psi \rangle = \sum_{\text{所有} x} \langle \phi | x \rangle \langle x | \psi \rangle. \quad (16.18)$$

现在如果这些振幅按我们以前所叙述的方法来归一化，那么在 x 处小范围内所有状态的和就相当于乘以 Δx ，而对所有 x 值求和就变为对 x 的积分。用我们修改过的定义，这个振幅的正确形式为

$$\langle \phi | \psi \rangle = \int_{\text{所有} x} \langle \phi | x \rangle \langle x | \psi \rangle dx. \quad (16.19)$$

振幅 $\langle x | \psi \rangle$ 就是我们现在所说的 $\psi(x)$ ，同样地，我们也选取 $\phi(x)$ 来表示振幅 $\langle x | \phi \rangle$ 。请记住， $\langle \phi | x \rangle$ 是 $\langle x | \phi \rangle$ 的复共轭，我们可以把式 (16.19) 写成

$$\langle \phi | \psi \rangle = \int \phi^*(x) \psi(x) dx. \quad (16.20)$$

按照我们的新定义，如果我们用对 x 的积分来代替对 x 的求和，则所有一切都遵从与以前相同的公式。

对上面我们所说的还应提出一个限制。如果我们要用任何一组合适的基础态来充分地描述所发生的情况，则这组基础态必须是完全的。对于一个一维电子来说，仅仅用基础态 $|x\rangle$ 来描述其实是不够完善的，因为对于每个这样的态中的电子，其自旋可以向上或向下，获得一个完全集的一种办法是在 x 处取两组态，一组对应于向上的自旋，一组对应于向下的自旋。不过，目前我们不必为这种复杂的情形而操心。

* 关于概率分布的讨论请参看第1卷 § 6-4。

§ 16-3 具有确定动量的态

假定我们有一个电子, 它处在用概率幅 $\langle x | \psi \rangle = \psi(x)$ 描述的态 $|\psi\rangle$ 中。我们知道电子处于这个态时以一定的分布展开在直线上, 从而在位置 x 处小间隔 Δx 中找到电子的概率正好为

$$\text{prob}(x, dx) = |\psi(x)|^2 dx.$$

关于这个电子的动量, 我们能讲些什么呢? 我们要问, 这个电子具有动量 p 的概率是多大? 让我们从计算态 $|\psi\rangle$ 在另一个态 $|\text{动量 } p\rangle$ 中的振幅开始, 我们定义态 $|\text{动量 } p\rangle$ 为具有确定动量 p 的态。利用关于振幅分解的基本方程式(16.19), 我们可以求得这个振幅。按态 $|\text{动量 } p\rangle$ 的定义, 我们得

$$\langle \text{动量 } p | \psi \rangle = \int_{-\infty}^{\infty} \langle \text{动量 } p | x \rangle \langle x | \psi \rangle dx. \quad (16.21)$$

该电子具有动量 p 的概率应由这个振幅绝对值的平方给出。然而, 我们又遇到了归一化的小问题。一般说来, 我们所能问的只是在动量 p 附近 dp 的小范围内找到具有该动量的电子的概率。动量严格为某值 p 的概率必定为零(除非这个态 $|\psi\rangle$ 恰巧是一个具有确定动量的态)。只有当我们问到在动量 p 附近的小范围 dp 内找到这电子的概率时, 我们才会得到一个确定的概率。有好几种办法适用于使它归一化, 我们选一种我们认为最方便的办法, 不过这在目前对你们来说可能不是显而易见的。

我们所作的归一化使概率与振幅的关系满足

$$\text{prob}(p, dp) = |\langle \text{动量 } p | \psi \rangle|^2 \frac{dp}{2\pi\hbar}, \quad (16.22)$$

振幅 $\langle \text{动量 } p | x \rangle$ 的归一化就按这个定义来确定。当然, 振幅 $\langle \text{动量 } p | x \rangle$ 正好是振幅 $\langle x | \text{动量 } p \rangle$ 的复共轭, 而 $\langle x | \text{动量 } p \rangle$ 又正是我们在式(16.15)中所写的式子。采用这种归一化, 指数前面那个适当的比例常数正好为 1, 即

$$\langle \text{动量 } p | x \rangle = \langle x | \text{动量 } p \rangle^* = e^{-ipx/\hbar}. \quad (16.23)$$

于是式(16.21)变为

$$\langle \text{动量 } p | \psi \rangle = \int_{-\infty}^{\infty} e^{-ipx/\hbar} \langle x | \psi \rangle dx. \quad (16.24)$$

这个方程加上式(16.22)就能使我们对任何态 $|\psi\rangle$ 求出动量分布。

让我们来看一个特殊的例子——例如一个电子定域在 $x = 0$ 附近某一区域内。假定我们采用如下形式的波函数:

$$\psi(x) = K e^{-x^2/\alpha^2}. \quad (16.25)$$

那么这个波函数在 x 处的概率分布是上式绝对值的平方, 即

$$\text{prob}(x, dx) = P(x)dx = K^2 e^{-x^2/\alpha^2} dx. \quad (16.26)$$

式中的概率密度函数 $P(x)$ 是如图 16-1 所示的高斯曲线。大部分概率集中在 $x = +\sigma$

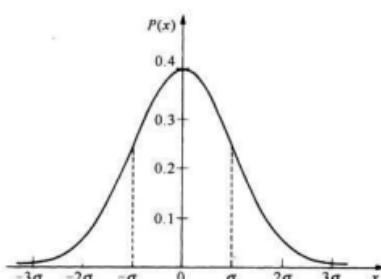


图 16-1 式(16.25)中波函数的概率密度

和 $x = -\sigma$ 之间。我们说这个曲线的“半宽度”为 σ 。(更精确地说, σ 等于某事物按此分布弥散时其坐标 x 的方均根值。)通常, 我们这样选取常数 K , 使概率密度 $P(x)$ 不仅与在 x 处单位长度内找到电子的概率成正比, 而且具有这样的标度, 使 $P(x)\Delta x$ 等于在 x 附近 Δx 范围内找到电子的概率。可以令 $\int_{-\infty}^{+\infty} P(x)dx = 1$ 来确定常数 K , 因为在整个区域内找到电子的概率必定为 1。这里, 我们得到 $K = (2\pi\sigma^2)^{-1/2}$ 。(我们用了公式 $\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$, 参见第 1 卷 § 40-4。)

现在我们来求动量分布。令 $\phi(p)$ 代表找到具有动量 p 的电子的振幅,

$$\phi(p) = \langle \text{动量 } p \mid \psi \rangle. \quad (16.27)$$

将式(16.25)代入式(16.24), 得:

$$\phi(p) = \int_{-\infty}^{+\infty} e^{-ipx/\hbar} \cdot K e^{-x^2/4\sigma^2} dx. \quad (16.28)$$

上述积分也可改写成

$$K e^{-p^2\sigma^2/\hbar^2} \int_{-\infty}^{+\infty} e^{-\frac{1}{4\sigma^2}(x+2ip\sigma^2/\hbar)^2} dx. \quad (16.29)$$

现在作代换 $u = x + 2ip\sigma^2/\hbar$, 这个积分就为

$$\int_{-\infty}^{+\infty} e^{-u^2/4\sigma^2} du = 2\sigma\sqrt{\pi}. \quad (16.30)$$

(数学家们可能会反对我们在里面的推导方式, 但不管怎么说, 结果是正确的。)

$$\phi(p) = (8\pi\sigma^2)^{1/4} e^{-p^2\sigma^2/\hbar^2}. \quad (16.31)$$

我们得到了一个有趣的结果, 即对 p 的振幅函数与 x 的振幅函数具有相同的数学形式, 只是高斯分布的宽度不同。我们可以把上式写成

$$\phi(p) = (\eta^2/2\pi\hbar^2)^{-1/4} e^{-p^2/\hbar^2\eta^2}, \quad (16.32)$$

式中 p 分布函数的半宽度 η 与 x 分布函数的半宽度 σ 有下列关系

$$\eta = \frac{\hbar}{2\sigma}. \quad (16.33)$$

我们所得结果表明: 如果使 σ 变小, 从而使对 x 的分布宽度变得很小, 那么 η 就变得很大, 对 p 的分布就散布得非常宽。或者反过来说, 如果我们有一个对 p 的狭窄的分布, 它必然对应着一个对 x 的散开的分布。要是我们愿意, 可以把 η 和 σ 看作是处在我们所研究状态中的电子之位置和动量区域的不确定性的某种量度。如果我们把 η 和 σ 分别称为 Δp 和 Δx , 则式(16.33)变为:

$$\Delta p \Delta x = \frac{\hbar}{2}. \quad (16.34)$$

有意思的是,可以证明对于 x 或 p 的任何其他形式的分布来说,乘积 $\Delta p \Delta x$ 不会比我们这里得到的小。高斯分布给出了均方根宽度乘积的最小可能值。一般我们可以说

$$\Delta p \Delta x \geq \frac{\hbar}{2}. \quad (16.35)$$

这就是海森伯的不确定性原理的定量表述,对此,我们以前已多次作过定性讨论。通常我们作这种近似表述:乘积 $\Delta p \Delta x$ 的最小值与 \hbar 有相同的数量级。

§ 16-4 对 x 的态的归一化

当涉及基础态的连续问题时,就需要对我们的基本方程进行修改,现在就来讨论这个问题。当我们所讨论的是有限数目的分立态时,这一组基础态必须满足的基本条件为:

$$\langle i | j \rangle = \delta_{ij}. \quad (16.36)$$

如果一个粒子处在一个基础态中,那么它处在另一个基础态中的振幅就为零。通过选取适当的归一化条件,我们定义振幅 $\langle i | i \rangle$ 为 1。这两个条件由式(16.36)描述。现在来看看,当用 $|x\rangle$ 表示在一条直线上的粒子的基础态时,这个关系应如何修改。如果已知这个粒子处在基础态之一 $|x\rangle$ 中,则它同时处在另一个基础态 $|x'\rangle$ 的振幅是多少呢?如果 x 和 x' 是沿直线上两个不同的位置,那么振幅 $\langle x | x' \rangle$ 肯定为 0,从而与式(16.36)一致。但是如果 x 和 x' 相同,而振幅 $\langle x | x' \rangle$ 将不为 1,则仍旧是因为老的归一化问题。为了看清楚我们必须怎样来补救这个问题,回到式(16.19),并把此式用到态 $|\psi\rangle$ 正好是基础态 $|x'\rangle$ 的特殊情况中去。于是我们得

$$\langle x' | \psi \rangle = \int \langle x' | x \rangle \psi(x) dx. \quad (16.37)$$

振幅 $\langle x | \psi \rangle$ 正好就是函数 $\psi(x)$ 。同样,既然振幅 $\langle x' | \psi \rangle$ 是指相同的态 ψ 而言的,因此它是变量 x' 的相同的函数,即 $\psi(x')$ 。于是我们可以把式(16.37)改写成

$$\psi(x') = \int \langle x' | x \rangle \psi(x) dx. \quad (16.38)$$

这个方程必须对任何态 $|\psi\rangle$ 都成立,因而对任意函数 $\psi(x)$ 也成立。这个条件将完全确定振幅 $\langle x | x' \rangle$ 的性质——当然, $\langle x | x' \rangle$ 只是与 x 和 x' 有关的函数。

现在我们的任务是找到一个函数 $f(x, x')$,当把它乘以 $\psi(x)$,并对所有 x 积分后正好给出量 $\psi(x')$ 。结果发现没有一个数学函数能做到这点!至少没有一个我们通常所说的“函数”可以满足这点。

假定我们把 x' 取为一个特殊数目 0,并定义振幅 $\langle 0 | x \rangle$ 为 x 的某个函数,譬如说 $f(x)$ 。于是式(16.36)就为:

$$\psi(0) = \int f(x) \psi(x) dx. \quad (16.39)$$

何种类型的函数 $f(x)$ 才能够满足这个方程呢?既然这个积分必须与 x 不为 0 时 $\psi(x)$ 的值

无关,所以很明显,对于 x 不为 0 的所有值, $f(x)$ 必定为 0。但是若 $f(x)$ 处处为 0, 则积分也将为 0, 因而式(16.39)不会被满足。所以我们遇到了一个不可能的情形: 我们希望有一个函数, 除了一点外处处为 0, 但是其积分仍为有限值。既然我们找不到一个具有这种性质的函数, 最方便的出路就是宣称函数 $f(x)$ 就是由式(16.39)定义的。也就是说 $f(x)$ 是能使式(16.39)正确成立的函数。这个函数最早是由狄拉克发明的, 所以就用他的名字来命名。我们把它写成 $\delta(x)$ 。所有我们说的是函数 $\delta(x)$ 具有这样的奇怪性质, 就是以它代替式(16.39)中的 $f(x)$, 则积分就取 $x = 0$ 时的 $\psi(x)$ 的值, 而且既然这积分与所有 x 不等于 0 的 $\psi(x)$ 值无关, 函数 $\delta(x)$ 必定除 $x = 0$ 的点外处处为 0。总括起来, 我们写成:

$$\langle 0 \mid x \rangle = \delta(x), \quad (16.40)$$

式中 $\delta(x)$ 由下式定义

$$\psi(0) = \int \delta(x) \psi(x) dx. \quad (16.41)$$

注意, 如果我们采用特殊的函数“1”作为式(16.41)中的函数 ψ , 那么我们就得到

$$1 = \int \delta(x) dx. \quad (16.42)$$

这就是说, 函数 $\delta(x)$ 具有这样的性质: 除了 $x = 0$ 外它处处为 0, 但是具有等于 1 的有限积分。我们必须把这函数 $\delta(x)$ 想象成在某一点为奇特的无穷, 使总面积等于 1。

想象狄拉克 δ 函数的一种方法是设想一系列的矩形——或者你愿意想象的任何峰函数——它变得越来越窄, 越来越高, 但始终保持面积为 1, 如图 16-2 所示。这函数从 $-\infty$ 到 $+\infty$ 的积分恒为 1。如果你们将它乘以任何函数 $\psi(x)$, 然后对这个乘积进行积分, 那么你们就会近似地得到这个函数 $\psi(x)$ 在 $x = 0$ 处的值, 当你使用的长方形越来越窄, 这一近似越来越好。如果你们愿意的话, 就可以用这种类型的极限过程来想象 δ 函数。但是, 唯一重要的事情是: δ 函数被定义为使得式(16.41)对每一个可能的函数 $\psi(x)$ 都成立。这就唯一地定义了 δ 函数。它的性质如上所述。

如果我们把 δ 函数的自变量从 x 变成 $x - x'$, 那么相应的关系式为:

$$\begin{aligned} \delta(x - x') &= 0, x' \neq x, \\ \int \delta(x - x') \psi(x) dx &= \psi(x'). \end{aligned} \quad (16.43)$$

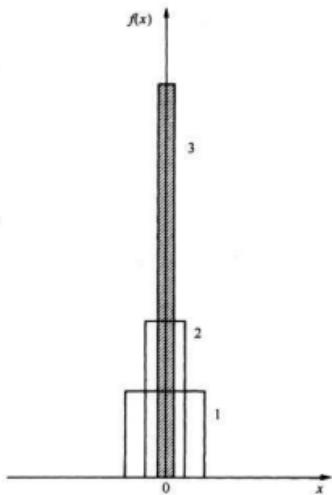


图 16-2 一组面积都为 1 的函数。

看上去越来越像 $\delta(x)$

如果我们用 $\delta(x - x')$ 来代替式(16.38)中的振幅 $\langle x|x'\rangle$, 那么该方程被满足。于是我们的结果是:

对 x 的基础态来说, 与式(16.36)相对应的条件为

$$\langle x' | x \rangle = \delta(x - x'). \quad (16.44)$$

现在我们已经完成了对我们的基本方程所作的必要的修改，在处理对应于直线上那些点的连续基础态问题时需要这些方程。可以直接把这种修改推广到三维的情况。首先我们用矢量 r 代替坐标 x ，然后把对 x 的积分变成对 x 、 y 和 z 的积分，换句话说，变成体积分。最后，一维 δ 函数必须被 3 个 δ 函数的乘积即 $\delta(x-x')\delta(y-y')\delta(z-z')$ 来代替，其中一个以 x 、一个以 y 、另一个以 z 为变量。把以上这些步合在一起，我们就得到在三维空间中粒子振幅的如下一组方程：

$$\langle \phi | \psi \rangle = \int \langle \phi | r \rangle \langle r | \psi \rangle dVol, \quad (16.45)$$

$$\langle r | \psi \rangle = \psi(r),$$

$$\langle r | \phi \rangle = \phi(r), \quad (16.46)$$

$$\langle \phi | \psi \rangle = \int \phi^*(r) \psi(r) dVol, \quad (16.47)$$

$$\langle r' | r \rangle = \delta(x-x')\delta(y-y')\delta(z-z'). \quad (16.48)$$

如果粒子多于一个将会发生什么情况呢？我们将告诉你们如何处理两个粒子系统的问题，你们就会很容易地看出，如果要处理粒子数目很多的系统时必须做些什么。假定有两个粒子，我们称之为粒子 1 和粒子 2。用什么作为基础态呢？一组能描述这一系统的完美的态为：粒子 1 处在 x_1 和粒子 2 处在 x_2 ，我们把这个态写成 $|x_1 x_2\rangle$ 。请注意，只对一个粒子位置的描述不能定义为基础态。每个基础态必须确定整个系统的情况。我们决不可认为每个粒子都是独立地像波一样在三维空间中运动。任何一个物理态 $|\psi\rangle$ ，可以通过给出在 x_1 和 x_2 处找到两个粒子的所有振幅 $\langle x_1, x_2 | \psi \rangle$ 而确定下来。因此这个广义振幅是两组坐标 x_1 和 x_2 的函数。你们可以看出这样的函数并不是在三维空间中传播的振动这种意义上的波。它一般也不只是两个单独的波——每个粒子一个波——的简单乘积。总的来说，它是由 x_1 和 x_2 所定义的六维空间中的某种波。如果在自然界中存在着两个具有相互作用的粒子，那么我们无法通过试图写出其中一个粒子单独的波函数来描述它发生的过程。我们在前面几章中所讨论的这个著名的佯谬——对一个粒子作测量就可以告知另一个粒子遭遇的事情，或者会破坏两粒子的干涉——已经给人们带来了种种麻烦，因为他们只愿意考虑单独一个粒子的波函数，而不愿去想象以两个粒子的坐标为变量的正确波函数。只有用两个粒子坐标的函数，才能正确地给出完整的描述。

§ 16-5薛定谔方程

到目前为止，我们只是为如何描述涉及一个可以位于空间各处的电子这样的态而操心，现在我们得为如何描述在各种情况下可能发生的物理问题而操心了。和前面一样，我们必须解决态如何随时间变化的问题。如果我们有一个态 $|\psi\rangle$ ，经一段时间后转变成了另一个态 $|\psi'\rangle$ ，那么我们只要使波函数——它就是振幅 $\langle r | \psi \rangle$ ——不仅是坐标的函数而且是时间的函数就可描述所有时刻的情况。于是我们可以用一个随时间变化的波函数 $\psi(r, t) = \psi(x, y,$

z, t) 来描写在给定情况下的一个粒子。这个随时间变化的波函数描述了随着时间的进展而出现的态的相继变化。这个所谓的“坐标表象”给出了态 $|\psi\rangle$ 在基础态 $|r\rangle$ 上的投影，它应用起来不一定总是最方便的，但是我们将首先讨论它。

在第8章中我们曾用哈密顿 H_u 描述态如何随时间变化。我们看到各种振幅的时间变化是用矩阵方程

$$i\hbar \frac{dC_i}{dt} = \sum_j H_{ij} C_j \quad (16.49)$$

给出的。这方程告诉我们：每个振幅 C_i 的时间变化与所有其他的振幅成正比，比例系数为 H_{ii} 。

当我们用连续基础态 $|x\rangle$ 时，式(16.49)是怎样的呢？首先让我们记住，式(16.49)也可写成

$$i\hbar \frac{d}{dt} \langle x | \psi \rangle = \sum_j \langle i | \hat{H} | j \rangle \langle j | \psi \rangle.$$

现在我们很清楚该怎么做。对 x 表象来说，我们认为

$$i\hbar \frac{\partial}{\partial t} \langle x | \psi \rangle = \int \langle x | \hat{H} | x' \rangle \langle x' | \psi \rangle dx'. \quad (16.50)$$

对基础态 $|j\rangle$ 的求和为对 x' 的积分所代替。既然 $\langle x | \hat{H} | x' \rangle$ 应是 x 和 x' 的某个函数，我们就可把它写成 $H(x, x')$ ——它相当于式(16.49)中的 H_{ii} 。于是式(16.50)与式

$$i\hbar \frac{\partial}{\partial t} \psi(x) = \int H(x, x') \psi(x') dx' \quad (16.51)$$

相同，其中

$$H(x, x') \equiv \langle x | \hat{H} | x' \rangle.$$

根据式(16.51)， ψ 在 x 处的变化率与 ψ 在所有各点 x' 处的 ψ 值有关；因子 $H(x, x')$ 是单位时间内电子从 x' 跃迁到 x 处的振幅。然而，结果是：在自然界中，除了非常靠近 x 的那些点 x' 以外，这个振幅总为 0。这就意味着——正如我们在本章开始所举的原子链的例子(式(16.12))中所看到的那样——式(16.51)的右边可以用在 x 处的 ψ 和 ψ 对 x 的微商完全表示出来。

对于一个不受力的作用和不受干扰而在空间自由运动的粒子来说，正确的物理定律是

$$\int H(x, x') \psi(x') dx' = -\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2} \psi(x).$$

这一式子我们是从什么地方得来的呢？没有地方。我们不可能从我们已知的任何事情中导出这个式子，它来源于薛定谔的头脑，是薛定谔在致力于了解对真实世界的各种实验观察的过程中发明的。你们想想我们从研究晶体中电子的传播而导出式(16.12)的过程，或许会获得上式为什么会是这种形式的某些线索。

当然，自由粒子并不十分令人兴奋。如果我们对粒子施以作用力，将会发生些什么呢？如果粒子所受的力可以用标势 $V(x)$ 来描写——这意味着我们考虑的是电力而不是磁力——如果我们只考虑低能的情况，从而可以忽略那些由相对论性的运动所引起的复杂性，于是适合真实世界的哈密顿给出

$$\int H(x, x') \psi(x') dx' = -\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2} \psi(x) + V(x) \psi(x). \quad (16.52)$$

同样,要是你们回顾一下电子在晶体中的运动,你们就可以对这个方程的来源获得一些线索;而且,如果从一个原子位置到另一个原子位置电子的能量变化很慢——如果对晶体加上电场,就会出现这种情况——那么你们知道方程要如何修改。这样,式(16.7)中的项 E_0 将随位置缓慢变化,并与式(16.52)中我们所加的新项相对应。

[你们可能感到奇怪,为什么我们直接从式(16.51)到式(16.52),而不是给出关于振幅 $H(x, x') = \langle x | \hat{H} | x' \rangle$ 的正确函数形式。我们这样做是因为:虽然式(16.51)右边的整个积分用我们习惯的方式表示,但是 $H(x, x')$ 只能写成奇特的代数函数。如果你们实在好奇的话,我们可以把 $H(x, x')$ 写成如下形式:

$$H(x, x') = -\frac{\hbar^2}{2m} \delta''(x - x') + V(x) \delta(x - x'),$$

式中 δ'' 表示 δ 函数的二阶微商。这个相当陌生的函数可以用一个稍微方便些的代数微分算符来代替,它是完全等效的:

$$H(x, x') = \left\{ -\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2} + V(x) \right\} \delta(x - x').$$

我们不用这些形式,而直接利用式(16.52)的形式。]

现在,如果我们对式(16.50)中的积分采用式(16.52)中的表示式,那么我们就对 $\psi(x) = \langle x | \psi \rangle$ 得到下列微分方程式:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \cdot \frac{\partial^2}{\partial x^2} \psi(x) + V(x) \psi(x). \quad (16.53)$$

如果我们对三维空间中的运动感兴趣,那么我们很清楚应该用什么方程来代替式(16.53)。唯一的改变是用

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

来代替 $\partial^2 / \partial x^2$,以及用 $V(x, y, z)$ 替代 $V(x)$ 。对于一个在势场 $V(x, y, z)$ 中运动的电子来说,其振幅 $\psi(x, y, z)$ 所遵循的微分方程为

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi. \quad (16.54)$$

上式称为薛定谔方程,是第一个为人所知的量子力学方程。在本书中我们提到的其他量子力学方程被发现之前,薛定谔就写下了这个方程。

虽然我们是沿着完全不同的途径来探讨这一问题的,但是标志物质量子力学描述诞生的伟大历史时刻是 1926 年薛定谔第一次写出他的方程之时。许多年中,物质内部的原子结构曾经是个巨大的谜。没有人能理解是什么原因使物质结合在一起,为什么会有化学键联,特别是原子怎么会是稳定的。虽然玻尔已经能够描述氢原子中电子的内部运动,而且看起来也可以解释所观察到的氢原子辐射出的光谱,但是电子为什么以这种方式运动仍然是一个谜。薛定谔所发现的电子在原子尺度上正确的运动方程提供了一种理论,可以定量、精确

并详细地计算原子现象。原则上,除了涉及磁和相对论的现象外,薛定谔方程能解释所有的原子现象。它解释了原子能级以及所有化学键联的事实。然而这在原则上成立——除了对最简单的问题,有关的数学立即变得非常复杂而无法精确求解。只对氢原子和氦原子进行了高度精确的计算。但是,用了各种近似方法(有些非常粗糙)后,许多有关更为复杂的原子以及分子化学键联的事实就可以理解了。在前面几章中我们已经给你们介绍过几种这样的近似方法。

我们所写出的薛定谔方程没有把任何磁效应考虑进去。只要在薛定谔方程中加进几项,就可以近似地把这种效应考虑进去。但是正如我们在第2卷中所看到的那样,磁性本质上是一种相对论性效应,所以只有用适当的相对论性方程才能正确地描述电子在任意电磁场中的运动,关于电子运动的正确的相对论方程是在薛定谔发明他的方程后一年由狄拉克发现的。它的形式与薛定谔方程有很大的不同。在这里我们根本无法去讨论它。

在我们进一步研究由薛定谔方程所得出的一些结果之前,我们愿向你们说明一下,对于具有大量粒子的系统,薛定谔方程是怎样的。我们并不打算应用这个方程,只是想让你们看一下这个方程,强调一下波函数 ψ 并不仅仅是空间的一个普通的波,而是一个具有许多变量的函数。如果存在许多粒子,薛定谔方程就变成

$$i\hbar \frac{\partial \psi(r_1, r_2, r_3 \dots)}{\partial t} = \sum_i -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x_i^2} + \frac{\partial^2 \psi}{\partial y_i^2} + \frac{\partial^2 \psi}{\partial z_i^2} \right) + V(r_1, r_2, r_3 \dots) \psi. \quad (16.55)$$

式中的势函数 V 相当于经典理论中所有粒子的总势能。如果没有外力作用在粒子上,函数 V 就是所有粒子相互作用的静电能。这就是说,如果第 i 个粒子带有电荷 $Z_i q_e$,则函数 V 就是*

$$V(r_1, r_2, r_3 \dots) = \sum_{\text{所有粒子对}} \frac{Z_i Z_j e^2}{r_{ij}}. \quad (16.56)$$

§ 16-6 量子化能级

后面有一章我们将对一个特殊例子详细讨论薛定谔方程的解。但是现在我们想告诉你们薛定谔方程的一个最不寻常的结果是怎样得出来的——即一个只包含空间连续变量的连

续函数的微分方程能给出诸如原子中分立能级这样的量子效应这种惊人的事实。我们要了解的基本要点是下述情形怎样发生的:当一个电子被某种类型的势“井”束缚在空间某区域时,它只能具有一组十分确定的分立能量中的这个或那个能量呢?

假定我们设想一个电子处在一维情况下,它的势能随 x 的变化方式如图 16-3 所示。我们设这个势是不变的——它不随时间变化。正如我们以前多次做过的那样,我们希望求出与确定能量、即确定频率的状态相对应

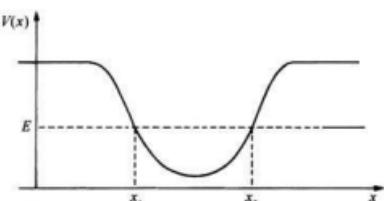


图 16-3 对于一个沿 x 轴运动的粒子的势阱

* 这里我们根据前面两卷所用的习惯,取 $e^2 \equiv q_e^2 / (4\pi\epsilon_0)$ 。

的解。让我们试一下这种形式的解

$$\psi = a(x) e^{-iE\tau/\hbar}. \quad (16.57)$$

如果我们把这个函数代入薛定谔方程，我们发现函数 $a(x)$ 必须满足下面的微分方程

$$\frac{d^2 a(x)}{dx^2} = \frac{2m}{\hbar^2} [V(x) - E] a(x). \quad (16.58)$$

这个方程告诉我们：在每个 x 处， $a(x)$ 对 x 的二次微商与 $a(x)$ 成正比，比例系数是量 $\frac{2m}{\hbar^2}(V - E)$ 。 $a(x)$ 的二次微商就是它的斜率的变化率。如果势 V 比粒子的能量 E 大，则 $a(x)$ 斜率的变化率与 $a(x)$ 具有相同的符号。这就意味着， $a(x)$ 这条曲线是凹的，随 x 增大，远离 x 轴而去。也就是说，它多少具有正或负指数函数 $e^{\pm x}$ 的性质。这意味着在图 16-3 中坐标轴上 x_1 左边 V 大于所设的能量 E 的区域，函数 $a(x)$ 的形状将像图 16-4(a) 中所示的某条曲线。

另一方面，如果势函数 V 小于能量 E ；则 $a(x)$ 对 x 的二次微商具有与 $a(x)$ 本身相反的符号， $a(x)$ 这条曲线将总是凹向 x 轴，与图 16-4(b) 中所示的那些曲线相似。在这样一个区域中的解逐渐接近于大致的正弦曲线的形状。

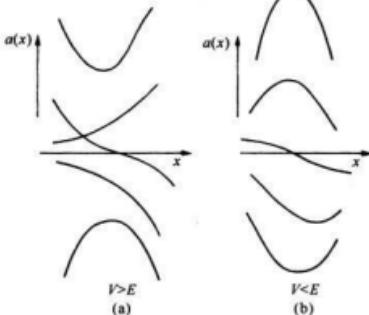


图 16-4 对于 $V > E$ 和 $V < E$ 的情形，波函数 $a(x)$ 的可能形状

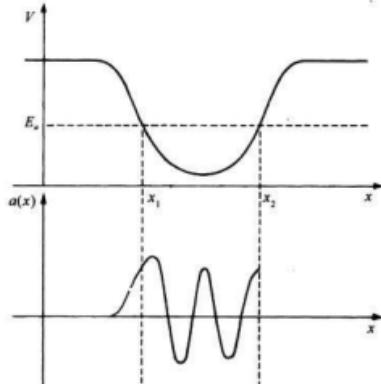


图 16-5 对于能量为 E_s 的波函数，在坐标轴 $x < x_1$ 的区域，它趋近于零

现在让我们来看看：我们是否能够用作图的办法求得函数 $a(x)$ 的解，它对应于能量为 E_s 的一个粒子处于如图 16-3 所示的势 V 中。既然我们试图描述一个粒子被束缚在势阱内的情形，因此我们所要找的解，其波振幅在势阱之外的 x 处具有很小的值。我们很容易想象出图 16-5 中所示的那样的曲线，对于大的负 x 值，它趋于零，当它趋于 x_1 的时候，它平滑地增长。既然在 x_1 处 V 等于 E_s ，则在该点此函数的曲率变为 0。在 x_1 和 x_2 之间的区域，量 $V - E_s$ 恒为负数，所以函数 $a(x)$ 总是凹向 x 轴，而且 E_s 和 V 之间的差值越大，曲率也越大。如果我们将曲线延长至 x_1 和 x_2 之间的区域内，它多少应如图 16-5 所示。

现在让我们把这条曲线延伸至 x_2 右边的区域，在那里它弯曲离开坐标轴并趋向大的正

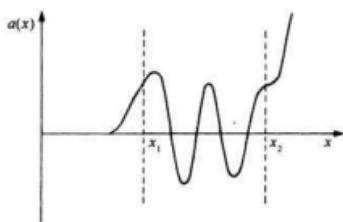


图 16-6 把图 16-5 的波函数 $a(x)$
继续延长到超过 x_2

值,如图 16-6 所示。对于我们所选取的 E_a , $a(x)$ 的解随着 x 的增加而变得越来越大。事实上,它的曲率也在增加(如果这个势函数继续保持平坦的话)。振幅以极大的比例急剧地增长。这意味着什么呢?这表示粒子没有被“束缚”在势阱里,它在阱外找到的可能性比阱内要大得多。对于我们所制造的这个解,在 $x=+\infty$ 处找到这个电子的可能性比其他任何地方大得多。我们没能找到一个束缚粒子的解。

让我们试试另一个能量,譬如一个稍微比 E_a 高些的能量 E_b ,如图 16-7 所示。如果我们在同样的情况下从左边开始,那么我们所得到的解就如图 16-7 的下半部分所示。初看起来,这解似乎比较好,其实它的结果与 E_a 的解一样糟——只不过现在当趋于大的 x 值时 $a(x)$ 变得越来越负。

可能这就是线索。既然,把能量 E_a 稍微改变一点到 E_b ,就使得这条曲线从坐标轴的一边跳到另一边,那么在 E_a 和 E_b 之间或许存在某个能量,对于这一能量当 x 为很大值时曲线将趋于零。的确有这么一个能量,我们在图 16-8 中粗略地画出了这个解的形式。

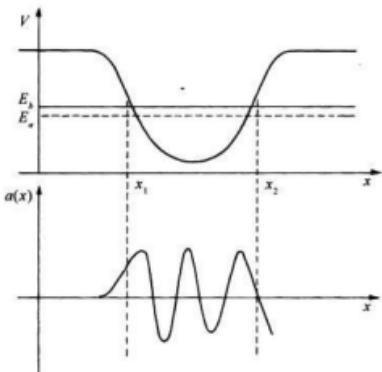


图 16-7 对于能量 E_b 大于 E_a 的波函数 $a(x)$

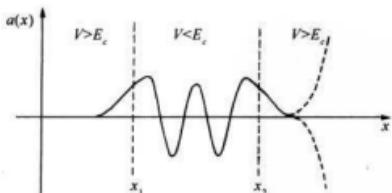


图 16-8 对于在 E_a 和 E_b 之间的能量 E_c 的波函数

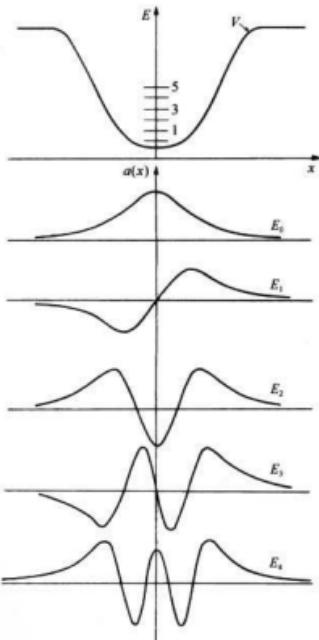


图 16-9 能量最低的 5 个束缚态的函数 $a(x)$

你们应该体会到我们在图中画出的解是一个非常特殊的解。如果我们将这个能稍微升高或降低一点，那么函数就会变成图 16-8 所示的两条虚线之一那样的曲线，而我们也就不会有关于一个束缚粒子的适当条件了。这样我们得到了一个结果：如果要使一个粒子被束缚在一个势阱中，只有当它具有一个完全确定的能量才行。

这是否意味着，对于一个束缚在势阱中的粒子，它只能具有一个能量呢？不，其他能量也可能，但不是非常靠近 E_c 的能量。注意，我们在图 16-8 中画出的波函数在 x_1 和 x_2 之间穿过坐标轴 4 次。如果我们所选的能量比 E_c 低相当数量的话，那么可以得到一个只与坐标轴相交 3 次、2 次、1 次或不相交的解。这些可能解画在图 16-9 中。（还可能有其他的解，对应于比图中所示的更高的能量值。）我们的结论是：如果一个粒子被束缚在势阱中，它的能量只能取一个分立能谱中的某些特定值。你们看到了，一个微分方程是怎样描述量子物理的基本事实的。

我们可以提醒你们注意另一种事。如果能量 E 在这个势阱顶部以上，那么就不再有分立的解，任何可能的能量都是允许的。这种解对应于自由粒子被势阱散射的情形，在讨论晶体中杂质原子的影响时，我们就已经看到过这种解的例子了。

第 17 章 对称性和守恒定律

§ 17-1 对 称 性

在经典物理学中,许多物理量是守恒的,如动量、能量和角动量。在量子力学中也存在着对应于这些量的守恒定理。从某种意义上来说,量子力学最美妙之处在于它可以从其他

一些东西导出这些守恒定理,而在经典力学中它们实际上是各定律的出发点。(在经典力学中,也有类似于我们在量子力学中用的这种处理方法,但这只有在很高深的水平上才能做到。)但是,在量子力学中,许多守恒定律与振幅的叠加原理及物理系统在各种变换下的对称性深刻关联。这就是本章的主题。虽然我们把这些概念主要应用到角动量的守恒上,但是最本质的一点在于:量子力学中所有物理量守恒的定理都与该系统的对称性有关。

因此,我们就从研究物理系统的对称性问题开始。氢分子离子是一个非常简单的例子——当然也可以氨分子为例——它们都有两个态。对氢分子离子来说,我们可把电子定域在第一个质子附近及定域在第二个质子附近这两种状态取作我们的基础态。这两个态——它们称为 $|1\rangle$ 和 $|2\rangle$ ——重画在图17-1(a)上。现在,只要这两个核是完全相同的,那么在这个物理系统中就存在某种对称性。这就是说,如果我们以一个位于两质子距离的一半的中间平面来反射这个系统(所谓反射是指把平面一边的一切东西都移到另一边的对称位置上),则我们得到图17-1(b)的情形。由于两个质子是完全相同的,因此,

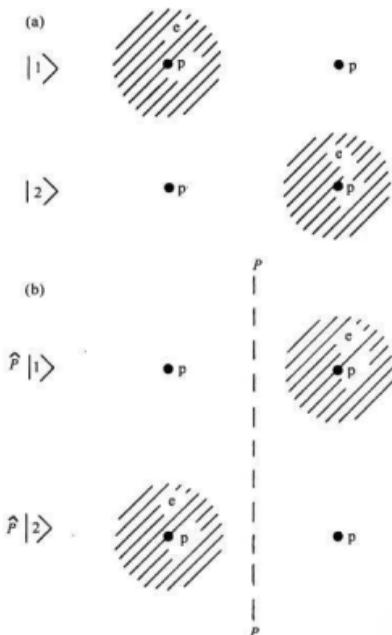


图 17-1 如果态 $|1\rangle$ 和 $|2\rangle$ 对平面 $P-P$ 反射,则它们分别变成态 $|2\rangle$ 和 $|1\rangle$

* 复习:第一卷 52 章,物理定律的对称性。

参考:A. R. Edmonds, Angular Momentum in Quantum Mechanics, Princeton University Press, 1957。——译者注

这种反射操作就使 $|1\rangle$ 变为 $|2\rangle$ 以及 $|2\rangle$ 变为 $|1\rangle$ 。我们把这个反射操作称为 \hat{P} ，并写为

$$\hat{P}|1\rangle = |2\rangle, \quad \hat{P}|2\rangle = |1\rangle. \quad (17.1)$$

所以在这种意义上 \hat{P} 是一个算符，它的“作用”是使一个态成为一个新的态。有趣的是 \hat{P} 作用于任何一个态所产生的是该系统的另外某个态。

与我们以前描述的任何算符一样， \hat{P} 也具有矩阵元，它可以用通常的符号来定义。如果在 $\hat{P}|1\rangle$ 和 $\hat{P}|2\rangle$ 的左边乘以 $\langle 1|$ ，就得到矩阵元

$$P_{11} = \langle 1 | \hat{P} | 1 \rangle \quad \text{以及} \quad P_{12} = \langle 1 | \hat{P} | 2 \rangle.$$

根据式(17.1)，它们为

$$\begin{aligned} \langle 1 | \hat{P} | 1 \rangle &= P_{11} = \langle 1 | 2 \rangle = 0, \\ \langle 1 | \hat{P} | 2 \rangle &= P_{12} = \langle 1 | 1 \rangle = 1. \end{aligned} \quad (17.2)$$

同样的方法可得 P_{21} 和 P_{22} 。 \hat{P} 对于基础态 $|1\rangle$ 和 $|2\rangle$ 的矩阵为

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (17.3)$$

我们再次看到在量子力学中算符和矩阵这两个词实际上是可以互相通用的。它们之间只有一些技术性的差别，就像“数字”和“数”的差别一样，但是这种差别有点学究式，我们不必为之操心。所以不论 \hat{P} 定义一个操作，还是实际上用来定义一个数字矩阵，我们可以随意把它称为算符或矩阵。

现在要指出，我们假定整个氯分子离子系统的物理性质是对称的。例如，它并不一定要与邻近的别的事物有关。然而，如果这个系统是对称的话，那么下面的概念肯定是正确的。假设系统在 $t = 0$ 时处于态 $|1\rangle$ ，经过一段时间 t 后，我们发现系统处于一个较为复杂的情况——两个基础态的某种线性组合。记得我们在第8章中用乘以算符 \hat{U} 来代表“经过一段时间”。这就是说，系统过一会儿——为明确起见，譬如说15 s——后处在另外某个态。例如，它可能是由 $\sqrt{2/3}$ 的态 $|1\rangle$ 和 $i\sqrt{1/3}$ 的态 $|2\rangle$ 组成的态，我们可写成：

$$|\text{第 } 15 \text{ s 时的 } \psi\rangle = \hat{U}(15, 0)|1\rangle = \sqrt{2/3}|1\rangle + i\sqrt{1/3}|2\rangle. \quad (17.4)$$

现在我们问，如果开始时系统处于对称态 $|2\rangle$ ，则在同样的条件下经15 s后系统会发生什么变化？显然，如果世界是对称的——正如我们所假定的那样——我们将得到与式(17.4)对称的态：

$$|\text{第 } 15 \text{ s 时的 } \psi\rangle = \hat{U}(15, 0)|2\rangle = \sqrt{2/3}|2\rangle + i\sqrt{1/3}|1\rangle. \quad (17.5)$$

图17-2表示两个同样的概念。所以如果一个系统的物理性质相对某个平面是对称的，并且我们求得了一个特定态的行为，也就知道了原来的态对此对称平面反射后所得到的态的行为。

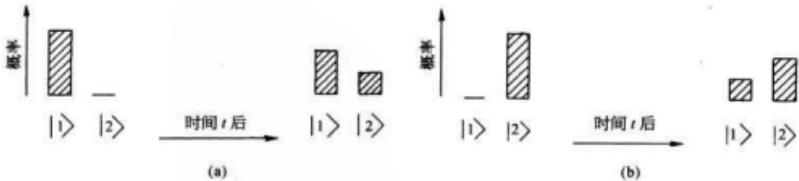


图 17-2 在一对称系统中，如果纯态 $|1\rangle$ 的发展如图(a)所示，则纯态 $|2\rangle$ 的发展必将如图(b)所示

现在我们希望把同样这些事情讲得更为一般化一些——这也就意味着稍微抽象一些。设 \hat{Q} 是任何一个不改变系统物理性质的操作。例如，我们可以把对氢分子的两个原子中间的平面上的反射操作 \hat{P} 取作为 \hat{Q} 。或者，在一个具有两个电子的系统中，我们可以把交换两个电子的操作看作 \hat{Q} 。另一个可能的操作是在一个球对称系统中，整个系统绕某个轴转动有限角度的操作，这种操作并不改变系统的物理性质。当然，我们总是对每种特定情况给 \hat{Q} 以某种特殊记号。明确地讲，我们通常把系统“绕 y 轴旋转 θ 角”的操作定义为 $\hat{R}_y(\theta)$ 。对于 \hat{Q} ，我们指的就是前面所描述过的，或者其他基本物理情况保持不变的任一操作。

让我们再考虑几个例子。如果一个原子没有加外磁场或外电场，并且如果使坐标轴绕任意一个轴旋转，则它仍然是相同的物理系统。再如氨分子，对于在平行于 3 个氢分子的平面上的反射是对称的——只要不存在电场。如果存在电场，并当我们作反射时，也不得不改变电场，这样就改变了物理问题。但是如果没外场，这个分子就是对称的。

现在我们来讨论一种一般情况。假设我们由态 $|\psi_1\rangle$ 开始，在给定的物理条件下经过某段时间后，它变成了态 $|\psi_2\rangle$ 。我们可以写成

$$|\psi_2\rangle = \hat{U} |\psi_1\rangle. \quad (17.6)$$

[你可以联想到式(17.4)。]现在设想我们对整个系统施以操作 \hat{Q} ，态 $|\psi_1\rangle$ 将转变成态 $|\psi'_1\rangle$ ，我们也可把 $|\psi'_1\rangle$ 写成 $\hat{Q} |\psi_1\rangle$ ，同样态 $|\psi_2\rangle$ 变成 $|\psi'_2\rangle = \hat{Q} |\psi_2\rangle$ 。现在如果在 \hat{Q} 作用下系统的物理性质是对称的（不要忘记如果，因为这并不是系统的一般性质），那么在同样的条件下经过相同的时间，我们应得

$$|\psi'_2\rangle = \hat{U} |\psi'_1\rangle. \quad (17.7)$$

[与式(17.5)相似。]但是我们可以把 $|\psi'_1\rangle$ 写成 $\hat{Q} |\psi_1\rangle$ 并把 $|\psi'_2\rangle$ 写成 $\hat{Q} |\psi_2\rangle$ ，因此式(17.7)也可写成

$$\hat{Q} |\psi_2\rangle = \hat{U} \hat{Q} |\psi_1\rangle. \quad (17.8)$$

现在如果用 $\hat{U} |\psi_1\rangle$ 代替 $|\psi_2\rangle$ ——式(17.6)——我们得到

$$\hat{Q} \hat{U} |\psi_1\rangle = \hat{U} \hat{Q} |\psi_1\rangle. \quad (17.9)$$

我们不难理解上面这个式子的意义。就氢离子而言，上式告诉我们：“先反射再等一段时间”[式(17.9)右边所示]与“先等一段时间然后再反射”[式(17.9)左边所示]是相同的。只要在反射的情况下 U 不变，这两种情形就相同。

由于式(17.9)对任何初始状态 $|\psi_1\rangle$ 均成立，所以实际上该式是一个关于算符的方程：

$$\hat{Q} \hat{U} = \hat{U} \hat{Q}. \quad (17.10)$$

这就是我们希望得出的结果——它是对称性的数学表述。当式(17.10)成立时，我们说算符 \hat{U} 和 \hat{Q} 对易。于是我们就可用下列方式来定义对称性：当 \hat{Q} 与 \hat{U} （时间的转移操作）相对易时，物理系统对于操作 \hat{Q} 是对称的。[用矩阵来表示，两个算符的乘积等价于矩阵的乘积。所以，对于在变换 Q 下是对称的物理系统，式(17.10)对矩阵 Q 和 U 也成立。]

顺便指出，因为对于无限小的时间 ϵ 而言， $\hat{U} = 1 - i \hat{H} \epsilon / \hbar$ ，式中 \hat{H} 为通常的哈密顿（参见第 8 章），所以你可以看出，若式(17.10)成立，则下式

$$\hat{Q} \hat{H} = \hat{H} \hat{Q} \quad (17.11)$$

也成立。因此式(17.11)是关于一个物理情况对操作 \hat{Q} 具有对称性的条件的数学表述。它定义了对称性。

§ 17-2 对称与守恒

在应用刚才所得到的结果之前，我们再讨论一些关于对称的概念。假设有一非常特殊的情况：当我们把 \hat{Q} 作用于某一态后，仍得到相同的态。这是一个非常特殊的情形，但是让我们假定碰巧对态 $|\psi\rangle$ 这种情形是正确的，即 $|\psi'\rangle = \hat{Q}|\psi\rangle$ 在物理上和 $|\psi\rangle$ 是相同的态。这意味着 $|\psi'\rangle$ 和 $|\psi\rangle$ 除相差某个相位因子^{*} 外是相等的。怎样才能发生这种情况呢？举例来说，假设我们有一个 H_2^+ 离子处于我们曾称为 $|I\rangle$ ^{**} 的态中。这个态在基础态 $|1\rangle$ 和 $|2\rangle$ 中具有相同的振幅。它们的概率如图 17-3(a)中的线条图所示。如果我们把反射算符 \hat{P} 作用于 $|I\rangle$ ，它将使态翻转， $|1\rangle$ 变成 $|2\rangle$ ， $|2\rangle$ 变成 $|1\rangle$ ，其概率如图 17-3(b)所示。但是这正好又是态 $|I\rangle$ 。如果我们从态 $|II\rangle$ 开始，反射前后的概率看上去完全相同。可是，如果考察其振幅，就有所差异了。对于态 $|I\rangle$ 来说，反射后的振幅没有改变，但对态 $|II\rangle$ 来说，振幅要反号。换句话说，

$$\begin{aligned}\hat{P}|I\rangle &= \hat{P}\left\{\frac{|1\rangle + |2\rangle}{\sqrt{2}}\right\} = \frac{|2\rangle + |1\rangle}{\sqrt{2}} = |I\rangle, \\ \hat{P}|II\rangle &= \hat{P}\left\{\frac{|1\rangle - |2\rangle}{\sqrt{2}}\right\} = \frac{|2\rangle - |1\rangle}{\sqrt{2}} = -|II\rangle.\end{aligned}\quad (17.12)$$

如果我们写成 $\hat{P}|\psi\rangle = e^\delta |\psi\rangle$ ，则对态 $|I\rangle$ 来说， $e^\delta = 1$ ，而对态 $|II\rangle$ 来说， $e^\delta = -1$ 。

让我们再看一下另一个例子。假设有一个右旋圆偏振光子沿 z 方向传播，如果我们进行绕 z 轴转动的操作，我们知道这只是对振幅乘以 e^δ ，其中 δ 是转动的角度。所以对于这种情况下转动操作来说， δ 就等于旋转角。

现在很清楚，如果碰巧算符 Q 在某个时候，譬如 $t = 0$ ，只是改变一个态的相位，那么它

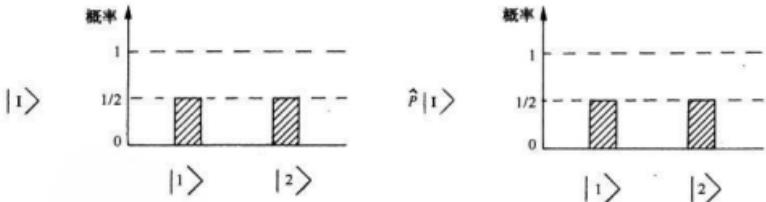


图 17-3 态 $|I\rangle$ 和由 $|I\rangle$ 经中心平面反射而得到的态 $\hat{P}|I\rangle$

* 顺便指出，你们可以证明 \hat{Q} 必定是一个幺正算符，这意味着如果它作用在 $|\psi\rangle$ 上而给出某个数乘以 $|\psi\rangle$ 的话，这个数一定是 e^δ 这种形式，这里 δ 是实数。这是个小小问题，其证明基于下述观察。诸如反射或转动这样的操作并不失去任何粒子，所以 $|\psi'\rangle$ 和 $|\psi\rangle$ 的归一化必定是相同的，它们只能相差一个纯虚数的相位因子。

** 参见 10-1 节，比之于前面的讨论，在本节中态 $|I\rangle$ 和 $|II\rangle$ 是反向的。

永远会改变态的相位。换言之,如果态 $|\psi_1\rangle$ 在一段时间 t 后变成态 $|\psi_2\rangle$,即

$$\hat{U}(t, 0)|\psi_1\rangle = |\psi_2\rangle, \quad (17.13)$$

并且如果情况的对称性使得

$$\hat{Q}|\psi_1\rangle = e^{i\theta}|\psi_1\rangle, \quad (17.14)$$

那么下式也成立

$$\hat{Q}|\psi_2\rangle = e^{i\theta}|\psi_2\rangle. \quad (17.15)$$

显然,由于

$$\hat{Q}|\psi_2\rangle = \hat{Q}\hat{U}|\psi_1\rangle = \hat{U}\hat{Q}|\psi_1\rangle,$$

而且如果 $\hat{Q}|\psi_1\rangle = e^{i\theta}|\psi_1\rangle$,则

$$\hat{Q}|\psi_2\rangle = \hat{U}e^{i\theta}|\psi_1\rangle = e^{i\theta}\hat{U}|\psi_1\rangle = e^{i\theta}|\psi_2\rangle.$$

[这一系列等式来自式(17.13)、(17.14)和关于对称系统的式(17.10),以及来自 $e^{i\theta}$ 这样的数与算符对易这一事实。]

所以由于某些对称性,某个在开始时是正确的事物,在其他任何时候也是正确的。而不就是守恒定律吗?是的!它表明如果你们考察初态,并且通过稍微作一点计算发现一个

操作(系统的对称操作)仅仅导致乘以一个相位因子,那么,你们就知道终态也具有相同的性质——相同的操作使终态乘上相同的相位因子。即使我们也许一点不知道使系统从初态变化到终态的宇宙内部机理,上述结论也总是正确的。即使我们无意考察系统从一个态变到另一个态的机构的具体细节,我们仍可断言:如果一个事物原来处于具有某种对称性的状态,而且如果这一事物的哈密顿在该对称操作下是对称的,那么,这个态在任何时候都具有相同的对称特征。这就是量子力学中所有守恒定律的基础。

看一个特殊例子。让我们回到算符 \hat{P} ,首先我们要稍微修正一下我们对 \hat{P} 的定义。应该认为 \hat{P} 不仅仅是镜面反射,因为那样需要先定义一个安放镜子的平面。有一种特殊的反射不需要这种明确的平面。假定我们这样来重新定义操作 \hat{P} :首先用在 xy 平面上的平面镜来反射,以使 z 变成 $-z$,而 x 仍为 x , y 仍为 y ;然后把该系统绕 z 轴旋转 180° ,于是 x 变为 $-x$, y 变为 $-y$ 。整个过程称为反演。每一点都通过原点而被投影到正好相对的位置上,每一点的所有坐标都改变了符号。我们仍用符号 \hat{P} 表示这种操作,如图17-4所示。

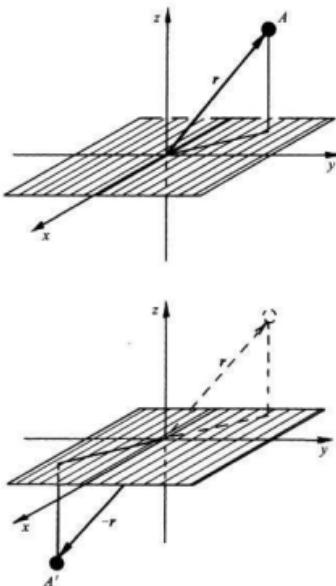


图 17-4 反演操作 \hat{P} ,位于 (x, y, z) 处的任意点 A 被移至位于 $(-x, -y, -z)$ 处的 A' 点

这比单纯反射要方便一点,因为它不要求你指明用哪个坐标平面来反射,而只要指明位于对称中心的哪个点就行了。

现在设态 $|\psi_0\rangle$ 经反演操作后变为 $e^{\theta}|\psi_0\rangle$ ——即

$$|\psi'_0\rangle = \hat{P} |\psi_0\rangle = e^{\theta} |\psi_0\rangle. \quad (17.16)$$

然后假设我们再反演一次。两次反演后我们正好又回到开始的状态,根本没有发生什么变化。我们必定得到

$$\hat{P} |\psi'_0\rangle = \hat{P} \cdot \hat{P} |\psi_0\rangle = |\psi_0\rangle.$$

但

$$\hat{P} \cdot \hat{P} |\psi_0\rangle = \hat{P} e^{\theta} |\psi_0\rangle = e^{\theta} \hat{P} |\psi_0\rangle = (e^{\theta})^2 |\psi_0\rangle.$$

由之得到

所以如果反演算符是态的一种对称运算,则 δ 只有两种可能:

$$e^{\theta} = \pm 1,$$

这就是说

$$\hat{P} |\psi_0\rangle = |\psi_0\rangle, \quad (17.17)$$

或

$$\hat{P} |\psi_0\rangle = -|\psi_0\rangle.$$

在经典物理学中,如果一个态在反演下是对称的,则反演操作给出原来的态。但是在量子力学中,却有两种可能性:我们得到的是原来的态或是负的原来态。当我们得到原来的态、即 $\hat{P} |\psi_0\rangle = |\psi_0\rangle$ 时,我们就说态 $|\psi_0\rangle$ 具有偶字称;当符号改变 $\hat{P} |\psi_0\rangle = -|\psi_0\rangle$ 时,我们就说该态具有奇字称。(反演算符 \hat{P} 也称为字称算符。) H_2^+ 的态 $|I\rangle$ 具有偶字称,而态 $|II\rangle$ 具有奇字称——参看式(17.12)。当然也有一些态,它们在 \hat{P} 的作用下没有对称性,这些都是没有确定字称的态。例如在 H_2^+ 的系统中,态 $|I\rangle$ 具有偶字称,态 $|II\rangle$ 具有奇字称,而态 $|1\rangle$ 的字称不确定。

当我们谈到像反演这样的操作作用于“物理系统”时,我们可以用两种方式来考虑这个问题。我们可以想象将位于 r 处的任何物理的东西都移动到 $-r$ 处,或者想象从新的参照系 x', y', z' 来考察同一系统,新参照系与原来的参照系 x, y, z 的关系为: $x' = -x$, $y' = -y$, $z' = -z$ 。同样,当我们考虑转动时,可以想象物理系统作实体转动,或者想象“系统”在空间固定不动,而转动我们对系统作测量的那个坐标系。一般来说,这两种看法实质上是等价的。对转动来说,如果不说系统转过 θ 角而说把参考系转过负 θ 角,这两种观点是等价的。在本讲义中我们通常讨论投影到一组新坐标轴上后发生些什么,这样得到的结果和你把坐标轴固定,而把系统反向转过同样大小的角度后得到的结果是相同的。当你这样做时,角度的符号反过来^{*}。

许多物理定律——但不是所有的定律——在坐标反射或反演下是不变的。它们对反演是对称的。例如,电动力学的定律,如果我们把所有方程中的 x 变成 $-x$, y 变成 $-y$, z 变成 $-z$,它们都是不变的。重力定律和核物理的强相互作用定律也是如此。只有弱相互作

* 在别的书中你可能发现具有不同符号的公式,他们可能采用不同的角度定义。

用——引起 β 衰变的相互作用——不具有这种对称性。(在第1卷第52章中我们曾详细地讨论过这个问题。)我们现在不考虑 β 衰变。那么在预料 β 衰变不会产生显著影响的任何物理系统中——原子的光发射就是一个例子——哈密顿算符 \hat{H} 将和算符 \hat{P} 对易。在此情形下,我们就有下述命题:如果一个态原来具有偶宇称,并且假如你在以后某个时刻观察其物理情况,它将仍然具有偶宇称。例如,假定一个即将发射光子的原子所处的状态具有偶宇称,在发射光子后,考察包括光子在内的整个系统,它仍然具有偶宇称(如果开始时系统具有奇宇称也一样。)这个原理称为宇称守恒。你可以看出,为什么在量子力学中“宇称守恒”和“反射对称”这两个名词是紧密地互相交织在一起的。虽然直到不多几年前,大家还认为在自然界中宇称总是守恒的,但现在我们知道这是不正确的。由于 β 衰变反应不具有其他物理定律所具有的反演对称性,才发现以前的看法是错误的。

现在我们可以来证明一个有趣的定理(只要我们能够忽略弱相互作用,这定律就是正确的)。任何具有确定能量的态,只要它不是简并的,它必定具有确定的宇称,它肯定不是具有偶宇称就是具有奇宇称。(记得我们曾看到过这种系统,其中的几个态具有相同的能量——我们称这些态是简并的。我们的定理对它们不适用。)

对于一个具有确定能量的态 $|\psi_0\rangle$ 来说,我们知道

$$\hat{H} |\psi_0\rangle = E |\psi_0\rangle, \quad (17.18)$$

式中 E 只是一个数,即这个态的能量。如果任意算符 \hat{Q} 是系统的一个对称算符,则只要 $|\psi_0\rangle$ 是确定能量的唯一一个态,就可以证明

$$\hat{Q} |\psi_0\rangle = e^{i\theta} |\psi_0\rangle. \quad (17.19)$$

考虑经 \hat{Q} 作用后所得到的新态 $|\psi'_0\rangle$,如果物理状态是对称的,则 $|\psi'_0\rangle$ 必定与 $|\psi_0\rangle$ 具有相同的能量。但是我们所考虑的是只有一个态,即 $|\psi_0\rangle$,具有这一能量的情形,所以 $|\psi'_0\rangle$ 必定是同一个态——它与 $|\psi_0\rangle$ 只可能相差一个相位。这就是在物理上的论证。

从数学上可以得到同样的结论。式(17.10)或式(17.11)是对称性的定义(这定义对任何态都成立),

$$\hat{H} \hat{Q} |\psi\rangle = \hat{Q} \hat{H} |\psi\rangle. \quad (17.20)$$

但是我们考虑的只是一个态 $|\psi_0\rangle$,它是一个具有确定能量的态,所以 $\hat{H} |\psi_0\rangle = E |\psi_0\rangle$ 。由于 E 只是一个数,如果需要,可将它移到 \hat{Q} 的前面,我们得

$$\hat{Q} \hat{H} |\psi_0\rangle = \hat{Q} E |\psi_0\rangle = E \hat{Q} |\psi_0\rangle.$$

因此

$$\hat{H} (\hat{Q} |\psi_0\rangle) = E (\hat{Q} |\psi_0\rangle). \quad (17.21)$$

故 $|\psi'_0\rangle = \hat{Q} |\psi_0\rangle$ 也是 \hat{H} 的一个具有确定能量的态——具有同样的 E 。但是根据我们的假设,只存在一个这样的态,因此,必定为 $|\psi'_0\rangle = e^{i\theta} |\psi_0\rangle$ 。

我们刚才所证明的结果,对于物理系统的任何对称算符 \hat{Q} 都是正确的。所以,在我们只考虑电力和强相互作用的情况下(没有 β 衰变),反演对称是一个许可的近似,我们得 $\hat{P} |\psi\rangle = e^{i\theta} |\psi\rangle$ 。但是我们也已看到, $e^{i\theta}$ 不是+1就是-1。所以任何具有确定能量的态(非简并态),不是具有偶宇称就是具有奇宇称。

§ 17-3 守 恒 定 律

现在我们转而讨论另一个操作的有趣例子：转动。考虑将原子系统绕 z 轴旋转角 ϕ 这样一种特殊的转动算符，我们称这种算符^{*}为 $\hat{R}_z(\phi)$ 。我们假定这样操作并不影响物理系统沿 x 轴和 y 轴的情况。任何电场或磁场都取成平行于 z 轴^{**}，从而，如果整个物理系统绕 z 轴旋转，外界条件不致改变。例如，如果在空间有一个原子，我们将它绕 z 轴转过 ϕ 角，仍然得出相同的物理系统。

有些特殊的状态具有这种性质：上述操作产生的新态是原来的态乘上某个相位因子。让我们顺便很快地证明一下，如果是这样，则相位的变化必定总是与角 ϕ 成正比的。假设你的角 ϕ 转动两次，这与以角 2ϕ 转动一次是一回事。如果转动 ϕ 具有态 $|\psi_0\rangle$ 乘以相位 $e^{i\phi}$ 的效果，所以

$$\hat{R}_z(\phi) |\psi_0\rangle = e^{i\phi} |\psi_0\rangle,$$

那么连续两次这样的转动应对该态乘以因子 $(e^{i\phi})^2 = e^{i2\phi}$ ，因为

$$\hat{R}_z(\phi) \cdot \hat{R}_z(\phi) |\psi_0\rangle = \hat{R}_z(\phi) e^{i\phi} |\psi_0\rangle = e^{i\phi} \hat{R}_z(\phi) |\psi_0\rangle = e^{i\phi} e^{i\phi} |\psi_0\rangle,$$

相位的变化 δ 必须正比于 ϕ^{***} 。我们接着来考虑这样的特殊状态。对该态来说

$$\hat{R}_z(\phi) |\psi_0\rangle = e^{im\phi} |\psi_0\rangle, \quad (17.22)$$

式中 m 是某个实数。

我们还知道值得注意的事实是：如果系统对于绕 z 轴的转动是对称的，并且如果原来的状态正好具有满足式(17.22)的性质，那么这个态以后也具有这一性质。所以 m 这个数很重要。如果我们开始时知道它的值，就知道了它结尾时的值。它是一个守恒的数—— m 是一个运动常数。我们特别提出 m 进行讨论是因为它与任何特殊角度 ϕ 没有任何关系，并且还因为它与经典力学中的某些量相对应。在量子力学中，对于 $|\psi_0\rangle$ 这样的态，我们决定把 $m\hbar$ 称为绕 z 轴的角动量。如果这样，我们发现在大系统的极限下，上述的量就等于经典力学中角动量的 z 分量。所以，如果有一个态，它绕 z 轴的转动正好产生了一个相位因子 $e^{im\phi}$ ，那么这个态绕该轴就有确定的角动量，而且角动量是守恒的，它现在是、并且永远是 $m\hbar$ 。当然，也可以绕任何轴转动，而且得出对不同的轴角动量守恒的结论。你们看到，角动量守恒是与下列事实密切相关的，即当你转动一下系统时你得到相同的态，该态仅具有一个新的相位因子。

我们要指出这个概念是相当普遍的。我们将把它应用到另外两个守恒定律上去，这两条守恒定律与角动量守恒定律在物理概念上是完全一致的。在经典物理中我们还有动量守恒和能量守恒，有趣的是这两个定律也同样与某种物理对称性有关。

* 我们将非常明确地把 $\hat{R}_z(\phi)$ 定义为物理系统绕 z 轴转动 $-\phi$ ，这和把坐标系转动 $+\phi$ 相同。

** 假如在某时刻仅有一个场，并且场的方向保持不变，我们总可以把 z 轴取为沿着场的方向。

*** 要做更严格的证明，我们应对很小的转动角 ϵ 作这种论证。因为任何角 ϕ 都是 n 次 ϵ 角度的转动之和， $\phi = n\epsilon$ ，则 $\hat{R}_z(\phi) = [\hat{R}_z(\epsilon)]^n$ ，而总相位的改变就是小角度 ϵ 所产生的相位改变的 n 倍，所以与 ϕ 成正比。

假定有一个物理系统——一个原子、某些复杂的原子核、或一个分子或某个事物——而且假定如果我们将整个系统移动到别的地方,不会对它产生任何影响,那么该系统的哈密顿就具有这种性质:在某种意义上它仅仅与系统的内部坐标有关,而与空间的绝对位置无关。在这种情况下存在一种我们可进行的特殊对称操作,它就是空间的平移。让我们定义 $\hat{D}_x(a)$ 为沿 x 轴移动距离 a 的位移算符,这样,对任何态我们可以进行这一操作而得到一个新的态。但是也可能有一些非常特殊的态,当它们沿 x 轴移动 a 时,我们仍然得到与原来相同的态,只是相差一个相位因子。我们还可以像上面那样来证明,对于这一情形相位必定正比于 a 。所以对这些特殊态 $|\psi_0\rangle$,我们可以写下

$$\hat{D}_x(a) |\psi_0\rangle = e^{ia} |\psi_0\rangle, \quad (17.23)$$

系数 k 乘以 \hbar 后称为动量的 x 分量。我们这样称呼的理由是:对于一个大的体系,它和经典动量 p_x 在数值上是相等的。普遍的表述是这样的:当一个系统位移时,若该系统的哈密顿不变,而且如果该系统开始时具有确定的 x 方向的动量,那么,随着时间的推移,系统在 x 方向的动量将一直保持不变。一个系统的总动量在碰撞或爆炸前后将是相同的。

另外还有一个与在空间的位移很相似的操作:即时间的延迟。假设有一个物理系统,不存在与时间有关的外部条件,我们使它在某一时刻从某一状态开始变化。现在在另一次实验中,例如我们在两秒钟以后,或者说延迟了时间 τ 以后开始使此物理系统作同样的变化,而且如果外界条件与绝对时间无关,其演变将相同,所得到的终态也与刚才得到的终态相同,只不过推迟了时间 τ 。在那种情形下,我们还可以发现一些特殊的状态,它们随时间的进展具有这种特性,即延迟后的态正好是原来的态乘上一个相位因子。对这些特殊状态来说其相位的变化显然又必定与 τ 成正比。我们可以写成

$$\hat{D}_t(\tau) |\psi_0\rangle = e^{-i\omega\tau} |\psi_0\rangle. \quad (17.24)$$

在定义 ω 的时候习惯上使用负号,采用这一习惯, $\omega \hbar$ 就是系统的能量,而且它是守恒的。所以具有确定能量的系统是这样的系统,它经时间位移 τ 后仍旧是它本身再乘以 $e^{-i\omega\tau}$ 。(这就是我们在前面定义具有确定能量的量子态时所说的,因此我们是前后一致的。)这意味着,如果一个系统处在确定能量的状态,而且如果其哈密顿与时间 t 无关,那么不论发生什么情况,它在以后所有各时刻都会具有相同的能量。

至此你一定看出了守恒定律与宇宙对称性之间的关系。有关时间位移的对称性意味着能量守恒;有关空间位置 x , y 或 z 的位移对称性意味着动量分量守恒;有关绕 x , y 和 z 轴旋转的对称性意味着角动量的 x , y 和 z 分量守恒;与反映有关的对称性意味着宇称守恒;与两个电子交换有关的对称性意味着某种我们还未命名的守恒,等等。这些原理中的一部分与经典物理中的类似,另一部分在经典物理中没有。量子力学中存在的守恒定律比经典力学中所用的要多,至少比经典力学中通常使用的要多。

为了使你们能够阅读其他量子力学书籍,我们有必要作些小的技术上的说明——介绍一下人们常用的符号。当然,与时间有关的位移操作,就是我们以前谈到的 \hat{U} :

$$\hat{D}_t(\tau) = \hat{U}(t + \tau, t). \quad (17.25)$$

大多数人都喜欢用无限小的时间位移、无限小的空间位移或无限小角度的转动来讨论每件事情。既然任何有限位移和角度都可以由一连串的无限小的位移和角度累积而成,所以首

先来分析无限小的情况往往就比较容易。无限小时间位移 Δt 的算符(正如我们在第8章中所定义的)为

$$\hat{D}_t(\Delta t) = 1 - \frac{i}{\hbar} \Delta t \hat{H}. \quad (17.26)$$

于是 \hat{H} 和称为能量的经典物理量相类似,因为如果 $\hat{H} |\psi\rangle$ 刚好是一个常数乘上 $|\psi\rangle$,即 $\hat{H} |\psi\rangle = E |\psi\rangle$,则该常数就是系统的能量。

对其他操作也作同样处理。如果我们在 x 方向作一小位移,譬如 Δx ,一般说来,态 $|\psi\rangle$ 将变成另一个态 $|\psi'\rangle$ 。我们可以写成

$$|\psi'\rangle = \hat{D}_x(\Delta x) |\psi\rangle = \left(1 + \frac{i}{\hbar} \hat{p}_x \Delta x\right) |\psi\rangle, \quad (17.27)$$

因为当 Δx 趋近零时, $|\psi'\rangle$ 应正好变为 $|\psi\rangle$,或 $\hat{D}_x(0) = 1$,并且对于小的 Δx 来说, $\hat{D}_x(\Delta x)$ 与 1 之差应该正比于 Δx 。按这种方式定义的算符 \hat{p}_x 称为动量算符——当然指的是 x 分量。

同理,人们通常把小的转动写成

$$\hat{R}_z(\Delta\phi) |\psi\rangle = \left(1 + \frac{i}{\hbar} \hat{j}_z \Delta\phi\right) |\psi\rangle, \quad (17.28)$$

并且称 \hat{j}_z 为角动量的 z 分量算符。对于那些 $\hat{R}_z(\phi) |\psi_0\rangle = e^{im\phi} |\psi_0\rangle$ 的特殊状态,我们可以对任何小角度 $\Delta\phi$,将该式右边展开至 $\Delta\phi$ 的第一级而得到

$$\hat{R}_z(\Delta\phi) |\psi_0\rangle = e^{im\Delta\phi} |\psi_0\rangle = (1 + im\Delta\phi) |\psi_0\rangle.$$

将此式与式(17.28)中 \hat{j}_z 的定义相比较,我们得到

$$\hat{j}_z |\psi_0\rangle = m\hbar |\psi_0\rangle. \quad (17.29)$$

换言之,如果你们把 \hat{j}_z 作用到对 z 轴具有确定角动量的态上,则得出 $m\hbar$ 乘以同样的态,其中 $m\hbar$ 为角动量的 z 分量的数值。这与把 \hat{H} 作用在具有确定能量的态上得到 $E|\psi\rangle$ 十分类似。

我们现在要对角动量守恒的概念举些应用的例子,以告诉你们这些概念是怎样起作用的。其意义在于它们实在非常简单。你们以前就知道角动量是守恒的,从本章中你们真正要记住的只是如果一个态 $|\psi_0\rangle$ 具有绕 z 轴转动一角度 ϕ 后变成 $e^{im\phi} |\psi_0\rangle$ 这一性质,它的角动量的 z 分量就等于 $m\hbar$ 。这就是以后我们要做许多有趣工作时所需要的全部知识。

§ 17-4 偏振光

首先我们希望检验一个概念。在 § 11-4 中我们已证明,当在绕 z 轴* 转过角 ϕ 的坐标系中来观看右旋圆偏振光时,它要乘上 $e^{i\phi}$ 。这是否意味着该右旋圆偏振的光子沿 z 轴具有一个单位**的角动量呢? 的确是这样。这也意味着,如果我们有一束光,它所含有的大量光子都具有同样的圆偏振——就像经典光束那样——那么这束光就具有角动量。如果在

* 请原谅,这个角与我们在 § 11-4 中所用的角差一负号。

** 通常用 \hbar 为单位来量度原子系统的角动量十分方便。这样,一个自旋 $1/2$ 的粒子对任意轴的角动量为 $\pm 1/2$ 。或者一般地讲,角动量的 z 分量为 m 。你不必老是去重复那个 \hbar 。

某时刻光束所具有的总能量为 W , 那么就有 $N = W / \hbar\omega$ 个光子, 每个光子具有角动量 \hbar , 所以总角动量为

$$J_z = N\hbar = \frac{W}{\omega}. \quad (17.30)$$

我们能否用经典的方法证明右旋圆偏振光带有与 W/ω 成正比的能量和角动量呢? 如果一切都正确的话, 这应该是一个经典的命题。这是一个我们可以从量子物理过渡到经典物理的例子。我们应该来看一下经典物理是否证实这一点。它将给予我们一个概念, 即是否有权把 m 称为角动量。回忆一下在经典物理中右旋圆偏振光是怎样的。它用一个电场来描写, 该电场具有振动的 x 分量和振动的 y 分量, 两者的相位差 90° , 所以合电场矢量 ϵ 在如图 17-5(a) 所示的圆上旋转。现在假设这种光照射到一个吸收(或者至少部分吸收)它的墙壁上, 并且按照经典物理的观点来考虑墙内的一个原子。我们以前通常把原子中电子的运动看成是一个可以用外场来驱使它振动的谐振子。我们假定该原子是各向同性的, 从而它可以在 x 方向上和 y 方向上同样地振动。另外在圆偏振光中, 其 x 位移和 y 位移相同, 但是一个比另一个落后 90° 。最后的结果是该电子也在一个圆上运动, 如图 17-5(b) 所示。该电子离开平衡位置原点有一个位移 r , 作圆运动, 相对于矢量 ϵ 它滞后某个相位。 ϵ 和 r 之间的关系如图 17-5(b) 所示。随着时间的推移, 电场不断旋转, 而位移则以相同的频率旋转, 结果它们的相对取向保持不变。现在让我们观察电场对该电子所作之功。给予这个电子的功率为电子的速度 v 乘上平行于这个速度的 $q\epsilon$ 的分量:

$$\frac{dW}{dt} = q\epsilon_i v. \quad (17.31)$$

但是注意, 有角动量注入该电子, 因为始终存在着对于原点的力矩。此力矩为 $q\epsilon_i r$, 它必定等于角动量的变化率 dJ_z/dt :

$$\frac{dJ_z}{dt} = q\epsilon_i r. \quad (17.32)$$

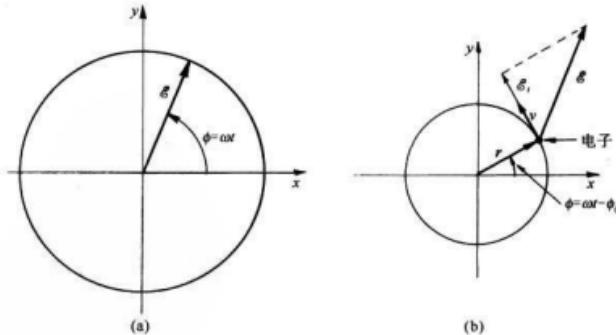


图 17-5 (a) 圆偏振光波的电场 ϵ ; (b) 圆偏振光驱动的电子的运动

因 $v = \omega r$, 所以我们得

$$\frac{dJ_z}{dW} = \frac{1}{\omega}.$$

因此, 如果我们对吸收的总角动量积分, 则它就正比于总能量——比例系数为 $1/\omega$, 与式(17.30)一致。光的确带有角动量——如果光沿 z 轴是右旋圆偏振的, 则为 +1 单位(乘以 \hbar), 如果是左旋圆偏振的, 则沿 z 轴为 -1 单位。

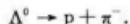
现在我们提出如下问题: 如果光是在 x 方向线偏振的, 那么它的角动量如何? x 方向的偏振光可以用右旋圆偏振光和左旋圆偏振光的叠加来表示。因此, 它对角动量为 $+\hbar$ 具有一定的振幅, 对角动量为 $-\hbar$ 具有另一振幅, 结果它没有确定的角动量。它既有一定的振幅以 $+\hbar$ 出现, 又具有相等的振幅以 $-\hbar$ 出现。这样两个振幅的干涉产生了线偏振, 但它具有相同的概率以 +1 或 -1 单位的角动量出现。对一束线偏振光的宏观测量表明: 它的总角动量为零。因为在大量光子中, 右旋圆偏振和左旋圆偏振光子的数目接近相等, 它们贡献的角动量相反——平均角动量为零。所以在经典理论中, 除非存在着一定的圆偏振, 否则你们就测不出角动量。

我们曾说过, 任何自旋 1 的粒子可以具有 3 个 J_z 值, 即 +1, 0, -1(即我们在施特恩-格拉赫实验中看到的 3 个态)。但光是螺旋型的, 它只有两个态, 而没有为零的态。这一奇特的缺漏与光不能静止这个事实有关。对于一个自旋 j 的静止的粒子来说, 必定有 $2j+1$ 个可能的态, 其 J_z 值以 1 为梯级从 $-j$ 增加到 $+j$ 。但是可以证明, 对某些自旋为 j 而质量为零的粒子来说, 只存在沿运动方向的分量为 $+j$ 和 $-j$ 的态。例如, 光没有 3 个态, 只有两个——虽然光子仍然是一个自旋 1 的客体。这和我们早先根据在空间旋转条件下发生的情况证明的自旋 1 的粒子必须具有 3 个态如何相一致呢? 对于一个静止的粒子来说, 可以绕任何轴旋转而不改变其动量状态。静止质量为零的粒子(如光子和中微子)不可能静止下来, 只有绕沿运动方向的轴旋转才不改变其动量状态。关于只绕一个轴的转动的论证不足以证明需要具有 3 个态, 而且其中一个态在旋转角 ϕ 的情况下系按照 $e^{i\phi}$ 变化的*。

还有一件值得注意的事。一般说来, 对于一个静止质量为零的粒子, 相对其运动方向的两个自旋态 $(+j, -j)$ 中只有一个才是真正必须的。例如中微子——它是自旋 $1/2$ 的粒子——在自然界中只有角动量分量与其运动方向相反的那个态 $(-\hbar/2)$ 才存在[而反中微子只有沿运动方向的态 $(+\hbar/2)$]。当一个系统具有反演对称性(从而宇称守恒, 光就是这样的)时, 才需要有两个分量 $(+j$ 和 $-j)$ 。

§ 17-5 Λ^0 的衰变

现在我们想举一个例子, 说明如何应用角动量守恒定理来处理一个特殊的量子物理问题。我们来看看 Λ^0 粒子的崩裂, 它通过“弱”相互作用衰变为一个质子和一个 π^- 介子:



* 对于一个质量为零的粒子, 我们曾试图至少找到一个证据, 以证明其沿运动方向的角动量必定是 $\hbar/2$ 的整数倍, 而不是像 $\hbar/3$ 这样的值。我们甚至用遍了洛伦兹变换的所有性质, 但还是失败了。或许这种想法本身就不正确。我们不得不去找维格纳(Wigner)教授谈谈, 他精通这方面的事情。

假定已知 π^- 介子的自旋为零,质子的自旋为 $1/2$, Λ^0 的自旋为 $1/2$ 。我们希望解决下列问题:假设 Λ^0 是以一种使其完全极化方式产生的——所谓极化指的是相对某个适当选取的 z 轴,它的自旋“向上”——参看图17-6(a)。我们的问题是,相对 z 轴以角度 θ 发射质子[如图17-6(b)所示]的衰变概率是多少?换言之,衰变的角分布如何?我们在 Λ^0 是静止的坐标系中来观察这种衰变,即在这个静止参照系中来测量角度,然后,如果需要的话,总可以把测得的结果变换到其他参照系中去。

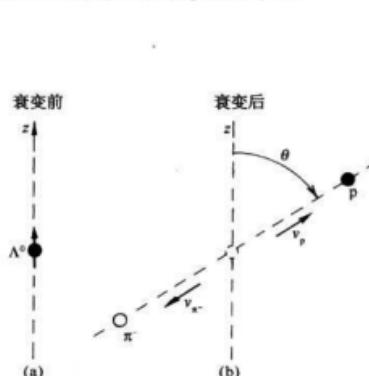


图17-6 一个自旋向上的 Λ^0 粒子衰变成一个质子和一个 π^- 介子(在质心参考系中)。在角 θ 方向发射质子的概率是多少呢?

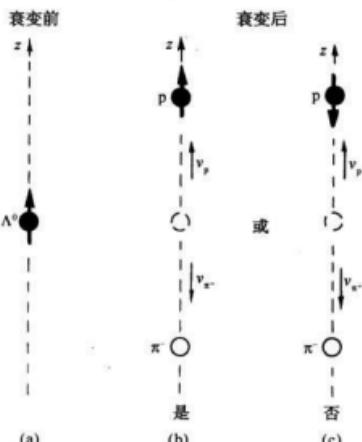


图17-7 一自旋“朝上”的 Λ^0 粒子,以质子沿 $+z$ 方向飞出的方式衰变的两种可能性。只有(b)角动量才守恒

我们从下面这个特殊情况着手,就是质子发射到沿 z 轴很小的一个立体角 $\Delta\Omega$ 中(如图17-7)。在衰变前 Λ^0 的自旋向上,如图17-7(a)所示,经过短时间后——由于至今还不知道的原因,除非与弱衰变有关,—— Λ^0 炸裂成一个质子和一个 π^- 介子。假定质子沿 $+z$ 轴向上运动。于是,根据动量守恒, π^- 介子必定向下运动。既然质子是自旋 $1/2$ 的粒子,其自旋必定不是“朝上”就是“朝下”——原则上还有两种可能,如图中(b)和(c)所示。但是角动量守恒要求质子具有“向上”的自旋。这从下面的论证中很容易明白。一个沿 z 轴运动的粒子,不可能因其运动而对沿这个轴的角动量有贡献,因此只有自旋才能对 J_z 有所贡献。衰变前对 z 轴的自旋角动量为 $+\hbar/2$,因此衰变后也必须是 $+\hbar/2$ 。我们可以说,既然 π^- 介子的自旋为零,那末质子的自旋必定“朝上”。

如果你担心这种论证在量子力学中可能不正确,那么我们可以花一点时间来证明一下这种论证是正确的。我们把初态(衰变前的态)称为 $|\Lambda^0, \text{自旋}+\rangle$,它具有如下性质:如果把该态绕 z 轴旋转角 ϕ ,则态矢量将乘上相位因子 $e^{i\phi\hat{z}}$ 。(在旋转后的参考系中态矢量为 $e^{i\phi\hat{z}} |\Lambda^0, \text{自旋}+\rangle$ 。)这就是我们所说的自旋 $1/2$ 粒子自旋“朝上”的意思。由于自然界的行為不取决于我们对坐标轴的选取,所以终态(质子加 π^- 介子)也必定具有相同的性质。譬如说,我们可以把终态写成

| 质子走向 $+z$, 自旋 $+z$; π 介子走向 $-z$).

但是我们实在并不需要标明 π 介子的运动, 因为在我们所选取的坐标系中它的运动方向总是与质子的相反, 所以我们可以把终态的描述简化为

| 质子走向 $+z$, 自旋 $+z$).

那么如果我们把坐标绕 z 轴旋转角 ϕ , 这个态矢量会发生什么变化呢?

由于质子和 π 介子都沿 z 轴运动, 它们的运动并不因这种旋转而改变。(这就是我们挑选这一特殊情况的原因, 否则我们就不可能作这种论证。) 再有, π 介子也不受影响, 因为它的自旋是 0。然后, 质子的自旋是 $1/2$ 。如果其自旋“朝上”, 它将贡献转动引起的 $e^{i\phi/2}$ 的相位变化。(若质子的自旋“朝下”, 则由于质子相位变化为 $e^{-i\phi/2}$ 。) 但是, 如果要角动量守恒, (它必须守恒, 因为没有外来因素影响系统的哈密顿。) 那么激发前后由于旋转而产生的相位变化必须相同。所以唯一的可能性是质子自旋“朝上”。如果质子向上运动, 它的自旋也必定“朝上”。

于是, 我们得出结论, 角动量守恒所允许的过程是图 17-7 中(b)所示的过程, 而不允许图中(c)所示过程。既然我们知道, 发生了衰变, 对于过程(b)——质子向上运动而自旋也“朝上”——就存在某种振幅, 让我们用 a 来代表在无限小的时间间隔内以这种方式发生衰变的振幅*。

现在让我们来看看, 如果开始时 Λ^0 的自旋是“朝下”的, 则会发生什么情况呢? 我们仍旧要问质子沿 z 轴向上运动的衰变情况, 如图 17-8 所示。你们一定知道, 如果角动量守恒, 则在这种情况下质子的自旋必定“朝下”。令这种衰变的振幅为 b 。

对于 a 和 b 这两个振幅, 我们再不能多说什么了。它们与 Λ^0 的内部机理以及弱衰变有关, 迄今还没有人知道如何计算这两个振幅。我们只能从实验中得到它们。但是只用这两个振幅就能够求出所有我们想要知道的衰变的角分布。不过必须始终仔细地、完整地定义我们所谈到的那些态。

我们想要知道质子沿与 z 轴成 θ 角的方向(在小立体角 $\Delta\Omega$ 中)飞出去的概率, 如图 17-6 所示。让我们在这个方向放置一个新的 z 轴, 称为 z' 轴。我们知道如何分析沿这个轴

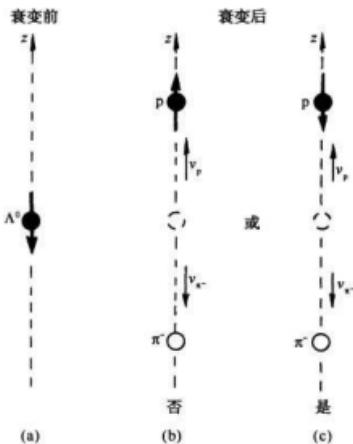


图 17-8 一个自旋“朝下”的 Λ^0
粒子沿 z 轴的衰变

* 我们现在假定你们对量子力学的机理已经相当熟悉, 因此我们可以用物理方式来讲述有关事情, 而不去花时间写出所有的数学细节。如果你们对这里所讲的不很清楚, 可以在本节末尾的注释中找到略去的细节。

发生的情况。相对于这个新轴, Λ^0 的自旋不再肯定“朝上”, 但它具有自旋“朝上”的某个振幅, 还有自旋“朝下”的另一个振幅。我们在第 6 章以及第 10 章式(10.30)早已讨论过这些情形。自旋“朝上”的振幅为 $\cos \theta/2$, 而自旋“朝下”的振幅为 $-\sin \theta/2$ 。当 Λ^0 的自旋沿 z' 轴为“朝上”时, 它将以振幅 a 在 $+z'$ 方向发射一个质子。所以沿 z' 方向找到一个自旋“朝上”的出射质子的振幅为

$$a \cos \frac{\theta}{2}. \quad (17.33)$$

同理, 沿正 z' 方向发现一个自旋“朝下”的出射质子的振幅为

$$-b \sin \frac{\theta}{2}. \quad (17.34)$$

这些振幅所代表的两个过程为图 17-9 所示。

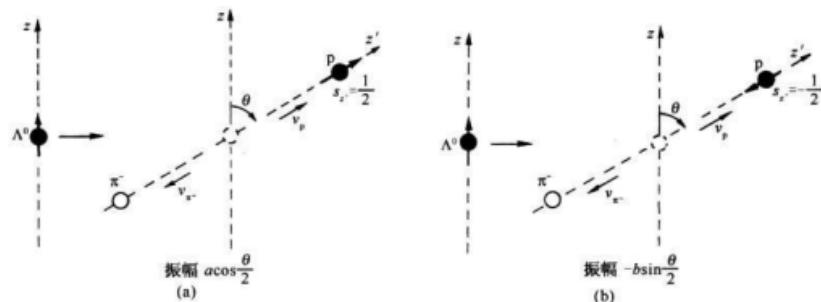


图 17-9 Λ^0 的两个可能的衰变态

现在提一个很容易的问题。假定 Λ^0 沿 z 轴自旋向上, 衰变质子在 θ 角方向离开的概率是多少? 这两个自旋态(沿 z' 方向“朝上”或“朝下”)是可以区分的, 虽然我们不打算去观察它们。所以为了求得这个概率, 我们将振幅平方, 并把它们相加。在 θ 附近很小的立体角 $d\Omega$ 中找到一个质子的概率 $f(\theta)$ 为

$$f(\theta) = |\alpha|^2 \cos^2 \frac{\theta}{2} + |\beta|^2 \sin^2 \frac{\theta}{2}. \quad (17.35)$$

用 $\sin^2 \frac{\theta}{2} = \frac{1}{2}(1 - \cos \theta)$, $\cos^2 \frac{\theta}{2} = \frac{1}{2}(1 + \cos \theta)$, 我们可将 $f(\theta)$ 写成

$$f(\theta) = \left(\frac{|\alpha|^2 + |\beta|^2}{2} \right) + \left(\frac{|\alpha|^2 - |\beta|^2}{2} \right) \cos \theta. \quad (17.36)$$

因此角分布的形式为

* 我们选取 z' 使其在 xx 平面内, 并利用有关 $R_x(\theta)$ 的矩阵元。对于任何别的选取, 也将得到同样的答案。

$$f(\theta) = \beta(1 + \alpha \cos \theta). \quad (17.37)$$

该概率有一部分与 θ 无关, 另一部分随 $\cos \theta$ 而线性变化。通过测量角分布, 我们可以求得 α 和 β , 从而求得 $|a|$ 和 $|b|$ 。

现在我们可以回答许多别的问题了。我们是否只对沿原来的 z 轴自旋“朝上”的质子感兴趣呢? 式(17.33)和(17.34)中的各项都给出沿 z' 轴 ($+z'$ 和 $-z'$) 发现自旋“朝上”和自旋“朝下”的质子的振幅。相对原来的轴自旋“朝上”的态 $|+z\rangle$ 可以用基础态 $|+z'\rangle$ 和 $| - z' \rangle$ 来表示。这样, 我们就可以把式(17.33)和(17.34)两个振幅用适当的系数 ($\cos \theta/2$ 和 $-\sin \theta/2$) 组合起来, 而得到总的振幅:

$$a \cos^2 \frac{\theta}{2} + b \sin^2 \frac{\theta}{2}.$$

它的平方就是质子以与 Δ° 相同的自旋 (沿 z 轴“朝上”)、在 θ 角方向出射的概率。

假如宇称是守恒的话, 我们就可以再说一件事。图 17-8 中的衰变正好就是图 17-7 中的衰变对 xy 平面的反射*。假如宇称是守恒的话, 则 b 就只能等于 a 或 $-a$ 。于是式(17.37)中的系数 a 就成为零, 在各个方向发生衰变的概率就要相等。

但是实验结果表明, 衰变中存在着不对称性。测得的角分布的确如我们所预言的那样按 $\cos \theta$ 变化, 而不是随 $\cos^2 \theta$ 或 $\cos \theta$ 的任何其他幂次变化。实际上, 既然角分布具有这个形式, 我们从这些测量中可以推得 Δ° 的自旋为 $1/2$ 。而且我们看到宇称并不守恒。事实上, 实验上得出的系数 a 为 -0.62 ± 0.05 , 所以 b 大约为 a 的两倍, 很明显, 对反射来说不存在对称性。

你们看我们从角动量守恒中能够得到多少东西。在下一章中我们将列举更多的例子。

附注 本节中的振幅 a 指的是: 在无限小的时间 dt 内从态 $|\Delta, \text{自旋}+z\rangle$ 产生态 $|\text{质子走向}+z, \text{自旋}+z\rangle$ 的振幅, 换言之, 就是

$$\langle \text{质子走向}+z, \text{自旋}+z | H | \Delta, \text{自旋}+z \rangle = i \hbar a, \quad (17.38)$$

这里的 H 是整个宇宙、至少是引起 Δ 衰变的所有事物所构成的那部分宇宙的哈密顿。角动量守恒意味着哈密顿必须具有如下性质

$$\langle \text{质子走向}+z, \text{自旋}-z | H | \Delta, \text{自旋}+z \rangle = 0. \quad (17.39)$$

对振幅 b 我们指的是

$$\langle \text{质子走向}+z, \text{自旋}+z | H | \Delta, \text{自旋}-z \rangle = i \hbar b. \quad (17.40)$$

角动量守恒意味着

$$\langle \text{质子走向}+z, \text{自旋}-z | H | \Delta, \text{自旋}-z \rangle = 0. \quad (17.41)$$

如果式(17.33)和(17.34)中所写出的振幅还不清楚的话, 那么我们可以用更为数学化的方式把它们表示如下。对式(17.33)我们指的是自旋沿 $+z$ 轴的 Δ 衰变成一个沿 $+z'$ 方向运动、其自旋也沿 $+z'$ 方向的质子的振幅, 即振幅

$$\langle \text{质子走向}+z', \text{自旋}+z' | H | \Delta, \text{自旋}+z \rangle. \quad (17.42)$$

根据量子力学的一般定理, 这个振幅可以写成

$$\sum_i \langle \text{质子走向}+z', \text{自旋}+z' | H | \Delta, i \rangle \langle i | \Delta, \text{自旋}+z \rangle, \quad (17.43)$$

这里的求和是对静止的 Δ 粒子的各个基础态 $|\Delta, i\rangle$ 进行的。由于 Δ 粒子为自旋 $1/2$, 因此在任何我们希望选取的参考基础中存在两个这样的基础态。如果我们使用相对 z' ($+z', -z'$) 自旋“朝上”和自旋“朝下”这两个态作为基础态, 式(17.43)的振幅就等于下列总和:

* 我们记得自旋是个轴矢量, 它在反射变换中要反向。

$$\begin{aligned} & \langle \text{质子走向} + z', \text{自旋} + z' | H | \Delta, + z' \rangle \langle \Delta, + z' | \Delta, + z \rangle \\ & + \langle \text{质子走向} + z', \text{自旋} + z' | H | \Delta, - z' \rangle \langle \Delta, - z' | \Delta, + z \rangle. \end{aligned} \quad (17.44)$$

根据式(17.38)的定义,以及根据角动量守恒得到的式(17.41),上式第一项中的第一个因子就是 a ,第二项的第一个因子为零。第二项中剩下的因子 $\langle \Delta, + z' | \Delta, + z \rangle$ 正好就是一个沿一轴自旋“朝上”的自旋 $1/2$ 粒子,它沿另一或 θ 角的轴也具有“朝上”的自旋的振幅,其值为 $\cos \theta/2$ ——参看表 6-2。所以式(17.44)刚好为 $a \cos \theta/2$,与我们在式(17.33)所写出的一样。对自旋“朝下”的 Δ 粒子作同样的论证,即得式(17.34)的振幅。

§ 17-6 转动矩阵概要

我们现在要把所学过的关于自旋 $1/2$ 和 1 的粒子转动的各种情况综合到一起,以便将来参考。下面你会看到有关自旋 $1/2$ 的粒子和自旋 1 的粒子以及光子(它是自旋 1 、静止质量为 0 的粒子)的两个转动矩阵 $R_z(\phi)$ 和 $R_y(\theta)$ 的表格。对每一个自旋,我们将给出绕 z 轴或 y 轴转动的矩阵项 $\langle j | R | i \rangle$ 。当然,它们和我们在前面几章中所用的 $\langle +T, | 0S \rangle$ 这类振幅是完全等价的。 $R_z(\phi)$ 的意思是把一个态投影到一个绕 z 轴旋转 ϕ 角的新坐标系中去——永远采用右手定则来定义转动的正指向。 $R_y(\theta)$ 表示参考系绕 y 轴旋转 θ 角。知道了这两种转动后,你当然可以作出任何的转动。我们按惯例这样来写矩阵元,使左边的态是新(转动过)的坐标系的一个基础态,而右边的态是老的(转动前)的坐标系中的基础态。你可以用多种方式来解释表中的各项。例如,表 17-1 中的项 $e^{-\#/\alpha}$ 表示矩阵元 $\langle - | R | - \rangle = e^{-\#/\alpha}$ 。它也表示 $\hat{R} | - \rangle = e^{-\#/\alpha} | - \rangle$ 或 $\langle - | \hat{R} = \langle - | e^{-\#/\alpha}$,它们全都是一回事。

表 17-1 自旋 $1/2$ 的转动矩阵

两个态: $| + \rangle$, 沿 z 轴“朝上”, $m = +1/2$
 $| - \rangle$, 沿 z 轴“朝下”, $m = -1/2$

$R_z(\phi)$	$ + \rangle$	$ - \rangle$	$R_y(\theta)$	$ + \rangle$	$ - \rangle$
$\langle + $	$e^{+\#/\alpha}$	0	$\langle + $	$\cos \frac{\theta}{2}$	$\sin \frac{\theta}{2}$
$\langle - $	0	$e^{-\#/\alpha}$	$\langle - $	$-\sin \frac{\theta}{2}$	$\cos \frac{\theta}{2}$

表 17-2 自旋 1 的转动矩阵

三个态: $| + \rangle$, $m = +1$
 $| 0 \rangle$, $m = 0$
 $| - \rangle$, $m = -1$

$R_z(\phi)$	$ + \rangle$	$ 0 \rangle$	$ - \rangle$
$\langle + $	$e^{+\#}$	0	0
$\langle 0 $	0	1	0
$\langle - $	0	0	$e^{-\#}$

续表

$R_x(\theta)$	$ +\rangle$	$ 0\rangle$	$ -\rangle$
$\langle + $	$\frac{1}{2}(1 + \cos \theta)$	$+\frac{1}{\sqrt{2}}\sin \theta$	$\frac{1}{2}(1 - \cos \theta)$
$\langle 0 $	$-\frac{1}{\sqrt{2}}\sin \theta$	$\cos \theta$	$+\frac{1}{\sqrt{2}}\sin \theta$
$\langle - $	$\frac{1}{2}(1 - \cos \theta)$	$-\frac{1}{\sqrt{2}}\sin \theta$	$\frac{1}{2}(1 + \cos \theta)$

表 17-3 光子

两个态: $|R\rangle = \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle)$, $m=+1$ (右旋圆偏振)

$|L\rangle = \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle)$, $m=-1$ (左旋圆偏振)

$R_z(\phi)$	$ R\rangle$	$ L\rangle$
$\langle R $	$e^{+i\phi}$	0
$\langle L $	0	$e^{-i\phi}$

第 18 章 角 动 量

§ 18-1 电偶极辐射

上一章，我们发展了量子力学中角动量守恒的概念，并且说明怎样用这些概念来预言 Δ 粒子衰变中质子的角分布。现在我们想给你们举一些其他类似的例子，来说明原子系统中角动量守恒的重要性。我们的第一个例子是原子的光辐射。角动量守恒（除了其他方面）将决定发射光子的偏振情况和角分布。

假设一个原子处于一个有确定角动量（譬如自旋为 1）的激发态，当它跃迁到一个角动量为零的低能态时，发射出一个光子。问题是如何计算光子的角分布和偏振情况。（除了现

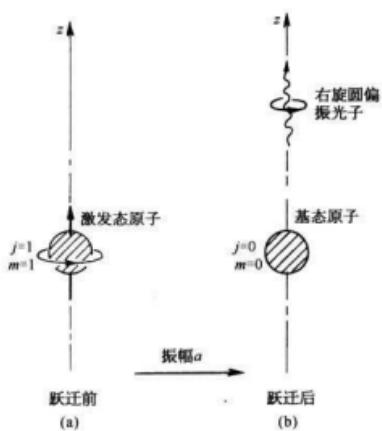


图 18-1 一个 $m = +1$ 的原子沿 $+z$ 轴方向发射一个 RHC 光子

立体制角 $d\Omega$ 内发射一光子的振幅。注意，在同一个方向发射 LHC（左旋圆偏振）光子的振幅为零。沿 z 轴的净角动量对于这种光子为 -1 ，而原子为 0 ，总的角动量为 -1 ，这将使角动量不守恒。

同样，如果原子的自旋起初是“朝下”的（沿 z 轴为 -1 ），那它只能在 z 轴正方向发射一个 LHC 光子，如图 18-2 所示。我们用 b 代表这一事件的振幅——意义同样是光子进入某一立体角 $\Delta\Omega$ 的振幅。另一方面，如果原子处于 $m = 0$ 的状态，则它根本不能在 $+z$ 方向发射光子，因为光子沿其运动方向的角动量，只能是 $+1$ 或 -1 。

在我们的粒子为自旋 1 而不是 $1/2$ 以外，这个问题和 Δ^0 的衰变几乎完全相同。) 因为原子的较高能态自旋为 1，所以其角动量的 z 分量有 3 种可能性。 m 的值可能为 $+1$ 、 0 或 -1 。在我们的例子中，取 $m = +1$ 。一旦你知道了如何处理这一情形，其他情形也就解决了。我们设想一个原子的角动量沿着 $+z$ 轴方向，如图 18-1(a) 所示；并且试问它沿 z 轴向上发射右旋圆偏振光，结果原子的角动量变为零的振幅是什么，如图 18-1(b) 所示，我们并不知道这个问题的答案，但是我们知道，对右旋圆偏振光，沿着它的传播方向具有一个单位角动量。所以发射光子后，情况将如图 18-1(b) 所示——留下的原子沿 z 轴的角动量为零，因为我们假定了原子在较低能态时自旋为零。我们用 a 代表这一事件的振幅。说得更明确些，我们令 a 为在 dt 时间内、向以 z 轴为中心的小

其次,我们可以证明 b 与 a 有关。假设我们把图18-1的情况作一反演,意思就是我们设想,如果把系统中每一部分移到原点对面的相应的点上去,该系统将会怎样。这并不意味着需要反射角动量矢量,因为它是人为的量。说得更确切些,我们应将对应于这种角动量的实际的运动特性倒转。在图18-3(a)和(b)中,我们显示了图18-1中的过程相对于原子中心反演前后的情况。注意原子转动的指向是不变的*,在图18-3(b)倒转的系统中, $m=+1$ 的原子向下发射一个LHC光子。

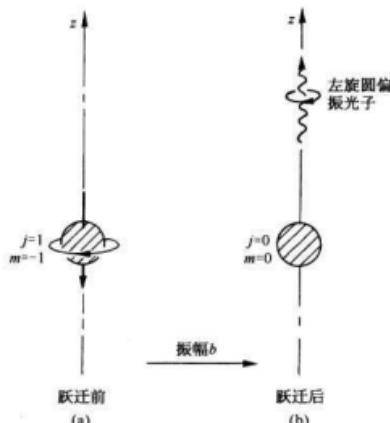


图18-2 一个 $m=-1$ 的原子沿 $+z$ 轴方向发射一个LHC光子

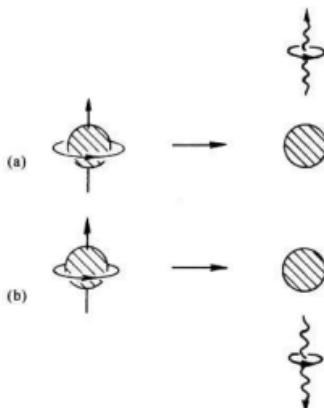


图18-3 如对(a)中的过程进行对原子中心的反演变换,就出现(b)所示的情形

如果现在我们将图18-3(b)的系统绕 x 轴或 y 轴旋转 180° ,那么它就变得和图18-2完全一样。反演和旋转的联合操作,使第二个过程成为第一个过程。由表17-2,我们知道对 y 轴旋转 180° ,恰好相当于把 $m=-1$ 的状态变成 $m=+1$ 的状态,所以除了由于反演可能造成符号改变外,振幅 b 必定和振幅 a 相等。反演中的符号改变将取决于原子的初态和终态的字称。

在原子的过程中,字称是守恒的,所以整个系统的字称在光子发射前后应该相同。所发生的情况取决于原子初态和终态的字称是偶还是奇——对于不同的情况,辐射的角分布将不一样。我们将选取一个一般的情况,初态为奇字称,而终态为偶字称,这将给出所谓的“电偶极辐射”。(如果初态和终态具有相同的字称,我们称其为“磁偶极辐射”,它具有回路中振荡电流的辐射特性。)如果初态的字称为奇,在使系统从图18-3的(a)到(b)的反演中,它的振幅改变符号。原子的终态具有偶字称,所以其振幅不改变符号。如果在反应过程中字称守恒,则振幅 b 和振幅 a 必然大小相等,符号相反。

* 当我们把 x, y, z 变成 $-x, -y, -z$ 时,你可能认为所有矢量都反了向。对于像位移和速度这样的极矢量,这是对的,但对于像角动量或任意由两个极矢量叉乘所得的轴矢量来说,这就不对了。轴矢量在反演后有相同的分量。

我们推断出:如果一个 $m=+1$ 的态向上发射一个光子的振幅是 a ,那么对我们所假设的初态和终态的字称而言, $m=-1$ 的态向上发射一个 LHC 光子的振幅为 $-a^*$ 。

我们已有了求出与 z 轴成任意角 θ 方向发射光子的振幅所需知道的一切。假设有一原子,原来具有 $m=+1$ 的偏振。我们可以把这个态分解为相对于新的 z' 轴的 $+1, 0, -1$ 等 3 个态,这新的 z' 轴在光子发射的方向上。对这 3 个态的振幅就是表 17-2 下半部分所给出的。于是在 θ 方向发射出一个 RHC(右旋圆偏振)光子的振幅,为 a 乘以在这个方向具有 $m=+1$ 的振幅,即

$$a\langle + | R_y(\theta) | + \rangle = \frac{a}{2}(1 + \cos \theta). \quad (18.1)$$

在同一方向发射一个 LHC 光子的振幅为 $-a$ 乘以此新方向具有 $m=-1$ 的振幅,利用表 17-2,即为

$$-a\langle - | R_y(\theta) | + \rangle = -\frac{a}{2}(1 - \cos \theta), \quad (18.2)$$

如果你对其他的偏振感兴趣,你可以将这两种振幅的叠加得出它们的振幅。当然,为了得到作为角度的函数的任何分量的强度,你必须取振幅绝对值的平方。

§ 18-2 光 散 射

让我们利用这些结果来解决一个稍微复杂、但也较为真实的问题。假设同样的原子处于其基态 ($j=0$) 上,并将一束入射光散射出去。设光起初沿 $+z$ 轴方向传播,也就是说光子从 $-z$ 方向射向原子,如图 18-4(a) 所示。我们可以认为光的散射分成两个阶段:光子先被吸收,然后再发射出去。如果我们从一个 RHC 光子出发[如图 18-4(a) 所示],并且角动量是守恒的,则原子在吸收光子后,就会处于 $m=+1$ 的态,如图 18-4(b) 所示。我们把这过程的振幅称为 c 。原子随后能够在 θ 方向发射出一个 RHC 光子,如图 18-4(c) 所示。一个 RHC 光子被散射到 θ 方向的总振幅就是 c 乘以式 (18.1)。如将此振幅称为 $\langle R' | S | R \rangle$,就有

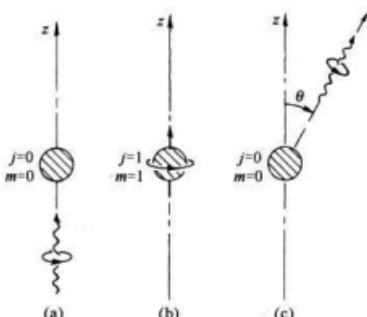


图 18-4 光被原子散射可看作一个两阶段过程

同样,吸收 RHC 光子而发射 LHC 光子也具有一定的振幅。这两个振幅的乘积就是 RHC 光子被散射成为 LHC 光子的振幅 $\langle L' | S | R \rangle$ 。由式 (18.2) 得

$$\langle L' | S | R \rangle = -\frac{ac}{2}(1 - \cos \theta). \quad (18.4)$$

* 你们有些人可能会根据我们所考虑的终态没有确定的字称,而反对我们刚才所做的论证。在本章末的附注 2 中,你将找到另一个你可能比较满意的证明。

现在我们要问,如果射进来的是一个左旋圆偏振光子,那将出现什么情况?当它被吸收后,原子将进入 $m = -1$ 的态,根据我们在上节中所用的同样的论证,可以证明这个振幅必然是 $-c$ 。一个处在 $m = -1$ 态的原子,在角 θ 方向发射一个 RHC 光子的振幅是 a 乘以振幅 $\langle + | R_s(\theta) | - \rangle$,它就等于 $\frac{1}{2}(1 - \cos \theta)$,所以我们有

$$\langle R' | S | L \rangle = -\frac{ac}{2}(1 - \cos \theta). \quad (18.5)$$

最后,对于一个 LHC 光子被散射为一个 LHC 光子的振幅为

$$\langle L' | S | L \rangle = \frac{ac}{2}(1 + \cos \theta). \quad (18.6)$$

(这里的两个负号相消了。)

如果我们测量关于圆偏振的任何给定组合的散射强度,它将和这 4 个振幅中一个振幅的平方成正比。例如,对于一束入射的 RHC 光,其散射辐射中 RHC 光的强度将随 $(1 + \cos \theta)^2$ 而变化。

所有这一切都不错,但是假如从线偏振光出发,那会怎样呢?如果是 x 向偏振光,则它可以用 RHC 光和 LHC 光的叠加来表示。我们写作(参见 § 1-4)

$$|x\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle). \quad (18.7)$$

如果是 y 向偏振光,则我们将有

$$|y\rangle = -\frac{i}{\sqrt{2}}(|R\rangle - |L\rangle). \quad (18.8)$$

现在你想知道什么呢?你想得到 x 向偏振的光在角 θ 方向被散射为 RHC 光子的振幅吗?你可以由组合振幅的一般规则得到它。首先用 $\langle R' | S$ 乘式(18.7)以得出

$$\langle R' | S | x \rangle = \frac{1}{\sqrt{2}}(\langle R' | S | R \rangle + \langle R' | S | L \rangle). \quad (18.9)$$

然后对两个振幅利用式(18.3)和(18.5),得到

$$\langle R' | S | x \rangle = \frac{ac}{\sqrt{2}} \cos \theta. \quad (18.10)$$

如果你要的是 x 向偏振的光子被散射成 LHC 光子的振幅,你将得到

$$\langle L' | S | x \rangle = \frac{ac}{\sqrt{2}} \cos \theta. \quad (18.11)$$

最后,假设你想知道一个 x 向偏振的光子被散射后仍保持其为 x 向偏振的振幅,你所要求的就是 $\langle x' | S | x \rangle$ 。它可以写为

$$\langle x' | S | x \rangle = \langle x' | R' \rangle \langle R' | S | x \rangle + \langle x' | L' \rangle \langle L' | S | x \rangle. \quad (18.12)$$

若再用下列关系

$$|R'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle + i|y'\rangle), \quad (18.13)$$

$$|L'\rangle = \frac{1}{\sqrt{2}}(|x'\rangle - i|y'\rangle), \quad (18.14)$$

由此得到

$$\langle x' | R' \rangle = \frac{1}{\sqrt{2}}, \quad (18.15)$$

$$\langle x' | L' \rangle = \frac{1}{\sqrt{2}}. \quad (18.16)$$

所以得到

$$\langle x' | S | x \rangle = a \cos \theta. \quad (18.17)$$

答案是,一束 x 向偏振光被散射到 θ 方向(在 xz 平面中的强度)和 $\cos^2 \theta$ 成正比。如果你要问关于 y 向偏振光的情况,你会发现

$$\langle y' | S | x \rangle = 0. \quad (18.18)$$

所以散射光在 x 方向是完全偏振的。

现在我们介绍一些有趣的事情。式(18.17)和(18.18)的结果完全和第1卷 § 32-6 给出的光散射的经典理论相一致。在那里我们想象电子被一线性恢复力束缚于原子中,电子就像一个经典振子。也许你在想:“这问题在经典理论中要容易得多,如果经典理论给出正确的答案,为什么要用量子理论来麻烦自己呢?”有一个理由是,我们至今所考虑的只是具有 $j=1$ 的激发态及 $j=0$ 的基态的原子这一特殊——虽然普通——的情况。如果激发态的自旋为 2,你将得到不同的结果。另外,我们并没有理由认为这种挂在弹簧上的电子并被振荡电场驱动的模型也应对单个光子适用。但是我们发现实际上它确实是适用的,所得出的光的偏振性和强度是正确的。所以从某种意义上讲,我们正在使整个课程接近真实情况。在第1卷中,我们曾用经典理论讨论过折射率和光散射理论,现在我们证明了量子理论在最一般的情况下给出了同样的结果。例如,我们现在实际上用量子力学论证了天上的光的偏振现象,这是唯一真正合理的方法。

当然,所有行得通的经典理论最终总得到合理的量子论据的支持。自然,我们花了很多时间来对你们解释的那些内容都是选自经典物理中对量子力学来说仍然正确的部分。你可能注意到,我们没有详细讨论过电子绕轨道运行的原子模型。这是因为这样的原子模型并不给出和量子力学相一致的结果。但是连在弹簧上的电子(原子“看上去”,根本不是这么回事)确实有用,所以我们在折射率的理论中采用了这一模型。

§ 18-3 电子偶素的湮没

下面我们想举一个非常漂亮的例子。它很有趣,虽然稍微复杂些,但我们希望不至于过分复杂。我们的例子是个称为电子偶素的系统,它是由一个电子和一个正电子组成的“原子”——一个 e^+ 和一个 e^- 组成的束缚态。它类似于一个氢原子,只是正电子代替了质子而已。和氢原子一样,它有许多状态。也像氢原子那样,其基态也由于磁矩的相互作用而分裂成“超精细结构”。电子和正电子的自旋都为 $1/2$,它们可以平行或反平行于任一给定的轴(在基

态不存在由于轨道运动而产生的角动量),所以它有4个状态:其中3个是自旋1的系统的支态,它们具有相同的能量,另一个是具有不同能量、自旋为零的状态。不过,能量的分裂比氢的1420兆周要大得多,这是因为正电子的磁矩比质子的磁矩大得多——大1000倍。

然而最重要的区别在于:电子偶素不能一直存在。正电子是电子的反粒子,它们可以互相湮没。这两个粒子完全消失——它们的静止能量转化为辐射能,后者表现为 γ 射线(光子)。在此衰变中,两个具有有限静止质量的粒子,变成了两个或多个静止质量为零的粒子*。

我们从分析处于自旋零的态的电子偶素的衰变开始。它衰变成两个寿命约为 10^{-10} s的 γ 射线(光子)。起初,正电子和电子靠近,它们的自旋反平行,构成了电子偶素系统。衰变后,有两个光子带着大小相等而方向相反的动量跑开(图18-5)。它们的动量必须大小相等而方向相反,因为,如果我们取的是处于静止状态的湮没这一情形,则衰变后的总角动量必然和衰变前一样为

零。如果电子偶素不是处于静止状态,则我们可以跟着它跑,解出这一问题,然后再把每件事变换到实验室坐标系中去。(看,我们现在已经工具齐全,可以做任何事情了。)

首先,我们注意角分布并不是很感兴趣。因为初态的自旋为零,它没有什么特殊的轴——它对所有的转动都是对称的。于是终态对所有的转动也应该是对称的。这意味着该衰变对每个角度都是同样可能的——光子走向任何方向的振幅都是一样的。当然,一旦我们在某个方向找到一个光子,另一个一定是在相反方向上。

剩下唯一我们要考察的问题是关于光子的偏振。设两个光子运动的方向为正z轴和负z轴。对于光子的偏振态,可以使用我们所需要的任何表象。我们选取右旋圆偏振(总是相对运动方向而言的)来描述它。我们可以立即看出,如果向上的光子是RHC,那么如果向下的光子也是RHC,则满足角动量守恒。每个光子相对于其动量方向来说都携带了+1单位的角动量,这就是说沿

z 轴**是+1和-1单位的角动量。两者的总和为零,衰变后的角动量和衰变前的相同。见图18-6。

* 今天在对世界的较深的了解中,我们还没有一种容易的方法,以区分光子的能量是否比电子的能量具有较少的物质性,因为正如你们所知道的,所有粒子的行为非常相似,唯一的区别是光子的静止质量为零。

** 注意,我们总是分析在粒子运动方向的角动量,如果要问及关于任何其他轴的角动量,那我们就要考虑到有“轨道”角动量——来自 $p \times r$ 项。例如,我们不能说光子正好是从电子偶素中心离去的,它们可能会像从旋转的车轮边缘抛射出去的两个物体那样离去。但是当我们把轴取在运动方向上时,我们就不必担心这种可能性了。

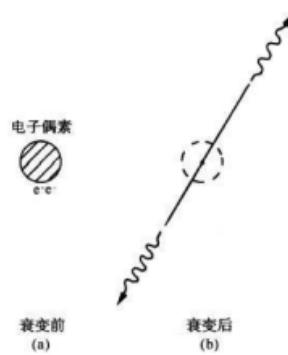


图18-5 电子偶素的双光子湮没

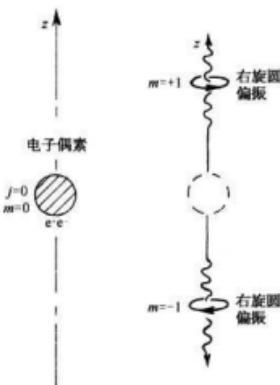


图18-6 电子偶素沿 z 轴湮没的一种可能性

同样的论证表明,如果向上的光子是 RHC,则向下的光子就不可能是 LHC,否则终态将具有两个单位的角动量。若初态自旋为零,这是不允许的。注意,对于其他自旋为1的电子偶素的基态,这样的终态也是不可能的,因为在任何方向角动量的极大值是一个单位。

现在我们要证明,对自旋1的态,双光子湮没是根本不可能的。你可能认为,如果取 $j=1, m=0$ 的态——它在z轴上的角动量为零——它应该与自旋零的态一样,可以衰变成两个RHC光子。当然,图18-7(a)所示的衰变对沿z轴的角动量守恒。但是来看一下,如果把此体系统y轴旋转180°,则会发生什么,我们将得到图18-7(b)所示的图像。它与图(a)完全相同。我们所做的一切只是把两个光子互换了一下。光子是玻色子,如果把它们互换一下,则振幅不变号,所以(b)部分衰变的振幅必定和(a)部分衰变的振幅相同,但是我们曾假设原来的物体自旋为1,如果我们把处于 $m=0$ 态、自旋1的物体绕y轴旋转180°,则它的振幅将改变符号(见表17-2, $\theta = \pi$)。所以图18-7中(a)和(b)的振幅应具有相反的符号,自旋1的态不能衰变为两个光子。

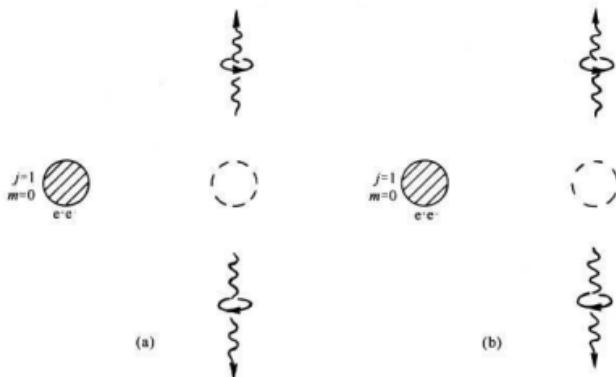


图18-7 对于电子偶素其 $j=1$ 的态,过程(a)和它绕y轴旋转180°后得到的过程(b)完全相同

当电子偶素形成时,你可能预期它有 $1/4$ 的时间停留在自旋0的态, $3/4$ 的时间留在自旋1的态(包括 $m=-1, 0$ 或 $+1$)。所以有 $1/4$ 的时间你将得到双光子的湮没;另外 $3/4$ 的时间可能没有双光子湮没;但湮没仍会发生,不过将产生3个光子。这种情形较难发生,因其寿命要长1000倍——约 10^{-7} s。这是实验观察到的结果。我们不再进一步详细讨论自旋1的粒子的湮没。

到现在为止,假如我们只考虑角动量,电子偶素自旋零的态可以变成两个RHC光子。另外还有一种可能性:它可以如图18-8所示的那样,变成两个LHC光子。接下来的问题是,这两种可能的衰变模式的振幅之间有什么关系?我们可以从宇称的守恒得出这种关系。

然而,要这样做,我们必须知道电子偶素的宇称。现在,理论物理学家在某种程度上已证明不容易解释为什么电子和正电子(电子的反粒子)的宇称必须相反,使得处于自旋零的基态的电子偶素的宇称必须为奇。我们只是假设它的宇称为奇,由此我们能得到与实验相

符合的结果，我们可以把这作为充分的证据。

让我们看看，如果把图 18-6 所示的过程作一反演，将会发生什么情况。当我们这样做后，两个光子的方向和偏振都反过来，反演后的情形正如图 18-8 所示。假设电子偶素的宇称为奇，则图 18-6 和图 18-8 中两个过程的振幅必须具有相反的符号。以 $|R_1 R_2\rangle$ 代表图 18-6 的终态，其中的两个光子都是 RHC，而以 $|L_1 L_2\rangle$ 代表图 18-8 的终态，其中的两个光子都是 LHC。真正的终态——让我们称它为 $|F\rangle$ ——必定是

$$|F\rangle = |R_1 R_2\rangle - |L_1 L_2\rangle. \quad (18.19)$$

于是，反演把 R 变成 L ，给出

$$P|F\rangle = |L_1 L_2\rangle - |R_1 R_2\rangle = -|F\rangle, \quad (18.20)$$

上式正好是式(18.19)加个负号。所以，终态 $|F\rangle$ 有负的宇称，和电子偶素自旋零的初态一样。这是角动量和宇称都守恒的唯一终态。衰变成这个态的振幅一定存在，但我们现在不必为此操心，因为我们只对偏振感兴趣。

式(18.19)中的终态有什么物理意义呢？一个意思是：如果我们用两个可以分别计数 RHC 和 LHC 光子探测器来观察发射出来的两种光子，我们将总是看到两个 RHC 光子在一起或两个 LHC 光子在一起。这就是说，如果你站在电子偶素的某一边，另一个人站在相反的一边，你可以测量偏振情况，并告诉另一个人他获得什么样的偏振。你有 50 对 50 的机会捕获一个 RHC 光子或一个 LHC 光子。不管你得到哪一种光子，你都能预料他也会得到同样的光子。

由于 RHC 或 LHC 偏振的机会是 50 对 50，听起来它可能像线偏振。我们要问：如果我们用只能接收线偏振光的计数器观察光子，将会怎样呢？测量 γ 射线的偏振不像测量光的偏振那样容易。对于这样短的波长，还没有很好的检偏振器。但是为使讨论方便起见，让我们假想有这种偏振器。假设你的计数器只接受 x 偏振的光，而在另一边的一个人只寻找具有 y 偏振的线偏振光。你检测到电子偶素湮没产生的两个光子的机会是多少呢？我们所要问的是 $|F\rangle$ 在态 $|x_1 y_2\rangle$ 的振幅是多少。换句话说，我们想求振幅

$$\langle x_1 y_2 | F \rangle,$$

当然，它就是

$$\langle x_1 y_2 | R_1 R_2 \rangle - \langle x_1 y_2 | L_1 L_2 \rangle. \quad (18.21)$$

虽然我们是在求两个光子的双粒子振幅，但由于每个粒子的行为是相互独立的，所以我们像以前处理单粒子振幅那样来处理。这意味着振幅 $\langle x_1 y_2 | R_1 R_2 \rangle$ 只是两个独立振幅 $\langle x_1 | R_1 \rangle$ 和 $\langle y_2 | R_2 \rangle$ 的乘积。由表 17-3，这两个振幅是 $1/\sqrt{2}$ 和 $i/\sqrt{2}$ ，所以

$$\langle x_1 y_2 | R_1 R_2 \rangle = +\frac{i}{2}.$$

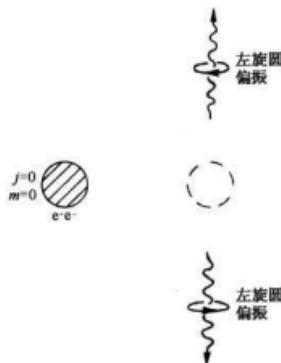


图 18-8 电子偶素湮没的另一种可能过程

同理得

$$\langle x_1 y_2 \mid L_1 L_2 \rangle = -\frac{i}{2}.$$

按照式(18.21), 将此两振幅相减得

$$\langle x_1 y_2 \mid F \rangle = +i. \quad (18.22)$$

所以得到一个单位概率^{*}, 如果你在 x 偏振的探测器中得到一个光子, 那么另外那个人将在他的 y 偏振的探测器中得到另一个光子。

现在假定另外那一个人也用和你一样的 x 偏振的计数器。当你得到一个光子时, 他决不会有计数。如果你从头到尾算一下, 就会得到

$$\langle x_1 x_2 \mid F \rangle = 0. \quad (18.23)$$

自然, 如果你用 y 偏振的计数器, 则只有当他用 x 偏振的计数器时, 他才会得到和你一致的符合计数。

现在所有这一切都导致一个有趣的情况。假定你用一块像方解石那样的东西, 把光分解为 x 偏振和 y 偏振的光束, 并在各光束中放一计数器, 我们把一个叫做 x 计数器, 另一个叫做 y 计数器。如果在另一边的那个人也这样做, 则你总是可以告诉他, 他的光子将跑进哪一束光去, 每当你和他同时记录到光子时, 你可以观察你自己这边哪个计数器捕获到光子, 然后告诉他, 他的哪一个计数器中有一个光子。我们说: 在某次衰变中, 你发现一个光子进入你的 x 计数器, 则你可以告诉他, 他也必定在他的 y 计数器中找到一个光子。

许多按通常(老式)的方式学习量子力学的人都会发现这个困惑, 他们总是认为一旦光子被发射出来, 它们就像一个具有特定性质的波一样。他们以为既然“任一给定的光子”具有 x 偏振或者 y 偏振的某些“振幅”, 那么就在 x 计数器或是在 y 计数器中找到它都应有某种机会, 并且这个机会应该和另外那个人找到另一个完全不同光子的机会无关。他们争辩说, “另一个人所做的测量, 不应该改变我找到什么的概率。”但是, 我们的量子力学告诉我们, 根据你对第一号光子所做的测量, 你能正确地料到, 当检测到第二号光子时它是什么样的偏振。这个观点从未为爱因斯坦所接受, 他对之极为烦恼——这成了著名的“爱因斯坦-波多尔斯基-罗森佯谬”。但是, 当像我们在这里所作的那样来描述这一情况时, 就似乎根本不存在什么佯谬, 并且很自然地得出在一个地方测得的结果与在另一个地方所测得的结果相关。关于该结果为佯谬的论证大致是这样的:

(1) 如果你有一个计数器告诉你, 你的光子是 RHC 还是 LHC, 则你可以精确地预言他将会发现哪一种类型的光子(RHC 或 LHC)。

(2) 因此他所接收到的光子的每一个都必定是纯 RHC 或纯 LHC, 有些是这一种, 有些是另一种。

(3) 你不能用改变你对你的光子观察的类型, 来改变他的光子的物理性质。不论你对你的光子作什么测量, 他的光子必定仍然不是 RHC 就是 LHC。

(4) 现在假定他用一块方解石改装的仪器, 把他的光子分成两种线性偏振的光束, 以至

* 我们还没有将我们的振幅归一化, 或乘上衰变成任一特殊终态的振幅, 但是我们可以知道这个结果是正确的, 因为当我们考察另一种选择时, 我们得到的概率为零(参见式 18.23)。

他的全部光子不是进入 x 偏振的光束，就是进入 y 偏振的光束。根据量子力学，绝对没有一个方法可以说出任一特定的 RHC 光子将进入哪个光束。它进入 x 光束的概率是 50%，进入 y 光束的概率也是 50%。对一个 LHC 光子来说情况也是这样。

(5) 既然每一个光子不是 RHC 就是 LHC——按照(2)和(3)——每个光子必然有 50 对 50 的机会进入 x 光束或 y 光束，那就无法预言它将走哪一条路。

(6) 然而理论却预言，如果你看见你的光子通过 x 起偏振器，则你能够肯定地预言，他的光子将通过他的 y 偏振光束。这与(5)矛盾，所以存在着佯谬。

然而大自然显然没有注意到这个“佯谬”。因为实验表明(6)中的预言实际上是对的。我们在第1卷第37章*关于量子力学的行为的第一次讲课中，已经讨论了这个“佯谬”的关键。在上面的论证中，(1)，(2)，(4)和(6)都是正确的，而(3)和它的推论(5)是错误的。它们不是自然界的正确描述。论证(3)表明，由你的测量（看见一个 RHC 或 LHC 光子），你能决定在他那里发生的是可选择的两个事件中的哪一个（看到一个 RHC 或 LHC 光子），而且，甚至不做你的测量，你仍然能讲出他那里所发生的事件将是这个或那个的两者之一。但是第1卷第37章中的要点恰恰是从一开始就指出在大自然中并非如此。大自然的方式需要有用干涉振幅来描述，每一个供选择的事件有一个振幅。测量实际上发生哪一个事件就破坏了这种干涉，但是如果^不进行测量，你就不可能再说：“这个或另一个事件仍要发生。”

如果你对你的每一个光子都能确定它是 RHC 还是 LHC，并且^也可以确定它是否为 x 偏振（对同一光子），那就真的有佯谬了。但你办不到——这是不确定性原理的一个例子。

你仍然认为有“佯谬”吗？事实上，要证实大自然的行为的佯谬，可以设计一个假想实验。对于这个假想实验，量子力学理论通过两种不同的论证会预言两个相互矛盾的结果。否则，这“佯谬”只不过是现实和你对“现实应该如何”的感觉之间的冲突。

你认为这虽^{不是}是佯谬而仍然是非常奇特的吗？我们完全同意你们的想法。这就是物理学迷人之处。

§ 18-4 任意自旋的转动矩阵

至此我们希望你们能看到，角动量概念在了解原子的过程中是多么重要。到目前为止，我们只考虑了自旋（或“总角动量”）为 0、 $1/2$ 或 1 的系统。当然还存在具有更高角动量的原子系统。为了分析这种系统，我们需要有 § 17-6 中那样的转动振幅表。这就是说，我们需要自旋为 $3/2$ 、 2 、 $5/2$ 、 3 等的振幅矩阵。虽然我们不去详细地计算出这些表，但我们希望告诉你们它是怎样算出的，以便你们一旦需要时能自己计算。

正如我们以前看到的，任何具有自旋或“总角动量” j 的系统能够存在于 $2j+1$ 个态中的任一个态，这些态的角动量在 z 轴上的分量，可以是 $j, j-1, j-2, \dots, -(j-1), -j$ （都以 \hbar 为单位）这一系列分立值中的任何一个值。把任何一个特殊态的角动量的 z 分量称为 $m\hbar$ ，通过给定两个“角动量量子数” j 和 m 的数值，我们就能定义一个角动量状态。我们以态矢量 $|j, m\rangle$ 来表示该态，在自旋 $1/2$ 粒子的情况下，其两个态就是 $|1/2, 1/2\rangle$ 和 $|1/2, -1/2\rangle$ ；而对于自旋 1 的系统，其状态用这种记号就被写成 $|1, +1\rangle$ 、 $|1, 0\rangle$ 及 $|1, -1\rangle$ ；当

* 也可参见本卷第一章。