

目录

第五章 定态微扰论和变分法	2
5.1 非简并定态微扰论	3
5.1.1 未扰动的系统以及投影算符	4
5.1.2 布里渊-维格纳微扰论	5
5.1.3 范德瓦尔斯力、万有引力、以及射电天文学	8
5.1.4 习题	15
5.2 *有效哈密顿量、简并微扰论以及不稳定态	15
5.2.1 有效哈密顿量	16
5.2.2 简并微扰论	18
5.2.3 不稳定离散态	21
5.2.4 习题	25
5.3 变分法在量子力学中的应用	25
5.3.1 变分法求基态能量	25
5.3.2 定态作为能量泛函的极值状态	27
5.3.3 有关基态的几个定理	29
5.3.4 相变与对称性自发破缺	36
5.3.5 习题	41

第五章 定态微扰论和变分法

陈童

像上一章那样可以精确求解而又重要的量子系统并不是随处可见。在量子力学中，更多实际的系统都不可精确求解。这时候就需要发展近似计算的方法。本章讨论两种重要的求解量子力学系统的方法，定态微扰论和变分法。值得强调的是，它们不仅仅是数学上的方法，同时也是我们建立物理图像的重要方法。而且，通过应用变分法，我们还介绍了相变和对称性自发破缺的物理机制。

此外，本章还介绍了一种有效哈密顿量的方法，并用它讨论了与连续谱耦合的离散态的衰变问题。

5.1 非简并定态微扰论

对于哈密顿量不显含时间的体系，最重要的就是求解定态薛定谔方程，进而得到系统的能谱和能量本征态。但是，像线性谐振子和氢原子这样可以精确求解的例子其实并不多见。大量的量子力学体系可能都要求助于数值计算，比方说即使像氦原子这样的两个核外电子的体系都没有精确解，而需要求助于计算机的帮助。但是数值求解常常有一个缺陷，那就是虽然我们能求出一些可以和实验数据进行对比的值，但是背后隐藏的物理图像却常常被掩盖了，这会使得我们对出现这些数值背后的物理机制缺乏了解，这样我们就没有真正地解释实验现象，也没有真正地理解物理现象的本质。当然，数值计算有时候也可以引导我们发现物理机制。但是在量子力学上，还存在一大类物理图像比较清晰的求解定态薛定谔方程的方法。这就是我们这里要介绍的定态微扰理论。

定态微扰理论之所以有比较清晰的物理图像，是因为它用来处理的是这样一些问题。在这类问题中，我们可以把系统看成是一个相对简单的主要部分加上一个比较复杂的微扰。其中这个相对简单的部分提供物理图像的基础，也就是说，我们对于这个简单部分的了解是清楚的，我们已经通过各种办法得到了这个部分的能谱和定态波函数，并在这个基础上发展了我们的物理图像。复杂部分往往是阻碍我们发展物理图像的原因，但现在，它只是对主要部分的一个微扰，因此我们就可以用微扰展开的办法逐级分析这个微扰如何修正原来的能谱和定态波函数。由此我们也能进一步搞清楚这个微扰的引入在微扰展开的每一级上给我们原来的物理图像带来了些什么新的东西。这样我们就能以简单的主要部分的物理图像为基础得到整个问题一个相对清晰的物理图像。正因为如此，在量子力学中，定态微扰论不仅仅是一个近似计算的方法，它同时也是一个建立物理图像的方法。幸运的是，自然界中有大量物理体系都可以用类似这样的微扰论方法来得到物理解。不幸的是，微扰论的方法并不能处理所有的物理体系，的确有一些系统是强耦合强关联的，这时候非微扰的方法将会变得很重要。然而，在量子的世界中，目前人们还没有发展出一套像定态微扰论这样系统的非微扰理论框架。可以说如何处理非微扰系统是很多物理学领域共同面临的难题。

5.1.1 未扰动的系统以及投影算符

假设我们有一个了解得很清楚的量子系统, 它的哈密顿算符我们记作 H_0 。所谓了解得很清楚, 我们指的是, 我们已经求出了 H_0 的所有本征值 ε_m 和相应的本征态 $|m\rangle$,

$$H_0|m\rangle = \varepsilon_m|m\rangle, \quad (5.1)$$

不妨假定 $m = 1, 2, 3, \dots$ 。当然, 我们可以将这些本征态取成是正交归一的, 即满足

$$\langle m|n\rangle = \delta_{mn}. \quad (5.2)$$

用这些正交归一本征态我们可以构造一些被称为投影算符的特殊算符, 比方说我们可以定义

$$P_m = |m\rangle\langle m|. \quad (5.3)$$

很显然, P_m 是一个厄米算符, 而且 $P_m^2 = |m\rangle\langle m|m\rangle\langle m| = |m\rangle\langle m| = P_m$, 即 P_m 满足

$$P_m^2 = P_m, \quad (5.4)$$

满足这种幂等关系的厄米算符就是投影算符。由于 $P_m|\phi\rangle = |m\rangle\langle m|\phi\rangle = \langle m|\phi\rangle|m\rangle$, 所以投影算符 P_m 的作用是, 当它作用在任意态 $|\phi\rangle$ 上时, 都会将这个态投影到 $|m\rangle$ 态上。由于正交性, 我们很容易看出, 当 $m \neq n$ 时, $P_m P_n = P_n P_m = 0$, 这时候我们称这两个投影算符正交, 并且这时候很容易验证 $P_m + P_n$ 也是一个投影算符。

一般的, 对于正交归一本征态的任何一个子集 \mathcal{S} , 我们可以定义

$$P_{\mathcal{S}} = \sum_{m \in \mathcal{S}} P_m = \sum_{m \in \mathcal{S}} |m\rangle\langle m|, \quad (5.5)$$

很容易验证 $P_{\mathcal{S}}$ 是一个投影算符(即满足 $P_{\mathcal{S}}^2 = P_{\mathcal{S}}$), 它的作用就是将任意量子态投影到由集合 \mathcal{S} 里的所有本征态所张成的希尔伯特子空间里面, 我们常常简称这个子空间为 $P_{\mathcal{S}}$ 子空间。假设我们记集合 \mathcal{S}^\perp 为集合 \mathcal{S} 在所有正交归一本征态全集中的补集, 则很显然 $P_{\mathcal{S}^\perp}$ 也是一个投影算符, 而且由于两个集合中的态相互正交, 所以 $P_{\mathcal{S}^\perp} P_{\mathcal{S}} = P_{\mathcal{S}} P_{\mathcal{S}^\perp} = 0$, 并且由于 $\mathcal{S} \cup \mathcal{S}^\perp$ 就是所有本征态的全集, 所以

$$P_{\mathcal{S}^\perp} + P_{\mathcal{S}} = 1, \quad (5.6)$$

这里的 $1 = \sum_m |m\rangle\langle m|$ 表示单位算符。正因为如此，我们常常将 P_{S^\perp} 记作 P_S^\perp ，即 $P_S^\perp = P_{S^\perp}$ ，它的作用就是将量子态投影到 P_S 希尔伯特子空间的正交补空间。特别的，我们有投影算符 P_n^\perp ，

$$P_n^\perp = \sum_{m \neq n} |m\rangle\langle m|, \quad (5.7)$$

它就是投影到除 $|n\rangle$ 态之外所有其余本征态所张成的希尔伯特子空间的投影算符。由于涉及到的这些态都是 H_0 的本征态，所以很显然

$$P_n H_0 = H_0 P_n, \quad P_n^\perp H_0 = H_0 P_n^\perp. \quad (5.8)$$

5.1.2 布里渊-维格纳微扰论

现在，设想用一个所有矩阵元的取值都比较小的算符 V 扰动原来那个系统 H_0 ，进而得到一个哈密顿算符为 $H = H_0 + V$ 的新系统。由于在很多情况下 V 虽然是一个微扰，但是它可能是一个很复杂的算符，会在原来的系统中引入很复杂的相互作用，因此精确求解 H 的本征谱和本征态可能就会变得很困难。但鉴于我们已经了解清楚了原来的系统 H_0 ，同时也鉴于 V 是一个微扰，这时候我们往往可以用微扰展开的办法来处理新的系统 H 。

具体来说，假设在扰动之后，原来 H_0 的本征态 $|n\rangle$ 变成了新系统的某个相应本征态 $|\psi_n\rangle$ ，相应的本征值也变成 E_n ，即

$$H|\psi_n\rangle = E_n|\psi_n\rangle. \quad (5.9)$$

假定原来的能量本征值 ε_n 和 H_0 的其余本征谱之间存在着一个有限的谱隙，即对于任何 $m \neq n$ ， $|\varepsilon_n - \varepsilon_m| \geq \Delta > 0$ 。则，我们总是能够将 $|\psi_n\rangle$ 在分别由 P_n 和 P_n^\perp 投影出来的两个正交且互补的空间中进行正交分解，通过合适地调整量子态整体的未定系数，我们可以设

$$|\psi_n\rangle = |n\rangle + P_n^\perp |\psi_n\rangle. \quad (5.10)$$

这个式子有很清楚的物理解释， $|n\rangle$ 就是原来 H_0 的第 n 个本征态， $P_n^\perp |\psi_n\rangle$ 表示由于 V 的引入而对这个态产生的扰动， $|\psi_n\rangle$ 就是扰动之后新的 H 的相应本征态，它由原来的 $|n\rangle$ 态加上扰动部分组成，当然，这种形式的 $|\psi_n\rangle$ 态是没有归一化的。

为了从(5.10)式出发得到关于 $|\psi_n\rangle$ 态的微扰展开式，我们将本征态 $|\psi_n\rangle$ 所满足的本征方程(5.9)重写为 $V|\psi_n\rangle = (E_n - H_0)|\psi_n\rangle$ ，则 $P_n^\perp V|\psi_n\rangle = P_n^\perp (E_n -$

$H_0|\psi_n\rangle = (E_n - H_0)P_n^\perp|\psi_n\rangle$, 由于 H_0 在 P_n^\perp 子空间上的谱和 E_n 之间有一个有限的谱隙, 因此我们可以在这个子空间上求算符 $E_n - H_0$ 的逆算符¹, 则

$$P_n^\perp|\psi_n\rangle = \frac{1}{E_n - H_0}P_n^\perp V|\psi_n\rangle. \quad (5.11)$$

定义 $G_0^\perp(E_n) = \frac{1}{E_n - H_0}P_n^\perp = \sum_{m \neq n} \frac{|m\rangle\langle m|}{E_n - \varepsilon_m} = P_n^\perp \frac{1}{E_n - H_0} P_n^\perp$, 则将(5.11)代入(5.10)式, 我们可以得到关于 $|\psi_n\rangle$ 的一个方程

$$|\psi_n\rangle = |n\rangle + G_0^\perp(E_n)V|\psi_n\rangle. \quad (5.12)$$

将这个方程进行迭代², 就可以得到

$$\begin{aligned} |\psi_n\rangle &= |n\rangle + G_0^\perp(E_n)V|n\rangle + G_0^\perp(E_n)V G_0^\perp(E_n)V|n\rangle \\ &+ G_0^\perp(E_n)V G_0^\perp(E_n)V G_0^\perp(E_n)V|n\rangle + \dots \end{aligned} \quad (5.13)$$

这就是关于 $|\psi_n\rangle$ 按照微扰 V 进行级数展开的展开式。但是这个展开式依赖于 E_n , 而到目前为止 E_n 的值还是未知的, 所以下一步我们就是要给出计算 E_n 的方程。

为此, 将 $V|\psi_n\rangle = (E_n - H_0)|\psi_n\rangle$ 左右两边分别和 $|n\rangle$ 作内积, 利用 $\langle n|\psi_n\rangle = \langle n|n\rangle = 1$, 就可以得到 $E_n - \varepsilon_n = \langle n|V|\psi_n\rangle$, 进而代入 $|\psi_n\rangle$ 的微扰展开式(5.13)就可以得到

$$E_n - \varepsilon_n = \langle n|V|n\rangle + \langle n|V G_0^\perp(E_n)V|n\rangle + \langle n|V G_0^\perp(E_n)V G_0^\perp(E_n)V|n\rangle + \dots \quad (5.14)$$

这当然是一个关于 E_n 的方程, 通常可以通过将这个方程截断到有限项进而求出 E_n 的一个近似解。

但是, 我们也可以按照微扰迭代的办法不断逼近方程(5.14)的精确解, 具体来说即是: 首先, 每个 V 都是一个一阶小量。其次, 我们设想可以按照微扰小量的阶数逐渐逼近 E_n 的精确值。也就是, 我们先算出 E_n 的0阶近似 $E_n^{(0)}$, 当然 $E_n^{(0)} = \varepsilon_n$, 由 $E_n^{(0)}$ 加上一阶微扰, 我们就能算出 E_n 的包含一阶微扰的近似 $E_n^{(1)}$, 忽略掉(5.14)式中的高阶微扰, 很容易看出 $E_n^{(1)} = \varepsilon_n + \langle n|V|n\rangle$ 。由 $E_n^{(0)}$ 和 $E_n^{(1)}$ 可以算出 E_n 的包含前二阶微扰的近似 $E_n^{(2)}$, 这时候我们要计算(5.14)式的二阶项 $\langle n|V G_0^\perp(E_n)V|n\rangle$, 但由于这一项本身已经是微扰的二阶小量了, 所以其中的 E_n 只需用0阶近似 $E_n^{(0)}$ 代入

¹在整个希尔伯特空间上, 算符 $E_n - H_0$ 可能有0本征值, 从而不一定有逆算符

²即将公式(5.12)理解成 $|\psi_n\rangle_{i+1} = |n\rangle + G_0^\perp(E_n)V|\psi_n\rangle_i$, $i = 0, 1, 2, 3, \dots$, 并令 $|\psi_n\rangle_0 = |n\rangle$

即可, 因此由(5.14)式忽略掉三阶以上的微扰, 就有 $E_n^{(2)} = \varepsilon_n + \langle n|V|n\rangle + \langle n|VG_0^\perp(E_n^{(0)})V|n\rangle$ 。同样的, 由 $E_n^{(0)}, E_n^{(1)}, E_n^{(2)}$ 又可以进一步算出 E_n 的包含前三阶微扰的近似 $E_n^{(3)}$, 为 $E_n^{(3)} = \varepsilon_n + \langle n|V|n\rangle + \langle n|VG_0^\perp(E_n^{(1)})V|n\rangle + \langle n|VG_0^\perp(E_n^{(0)})VG_0^\perp(E_n^{(0)})V|n\rangle$, 等等³, 依此类推。很显然, 假设这个过程收敛的话, 那么必有

$$\lim_{i \rightarrow +\infty} E_n^{(i)} = E_n, \quad (5.15)$$

E_n 就是能量本征值的精确值。

即使没有算出 E_n 的精确值, 得出这些 E_n 的近似值 $E_n^{(i)}$ 以后, 代入 $|\psi_n\rangle$ 的微扰展开式(5.13), 我们也可以算出 $|\psi_n\rangle$ 的近似解, 比如说 $|\psi_n\rangle$ 的包含前二阶微扰的近似解 $|\psi_n^{(2)}\rangle$ 为, $|\psi_n^{(2)}\rangle = |n\rangle + G_0^\perp(E_n^{(1)})V|n\rangle + G_0^\perp(E_n^{(0)})VG_0^\perp(E_n^{(0)})V|n\rangle$ 。

当然, 在实际中我们用得最多的是, 能量本征值的直到二阶微扰为止的近似 $E_n^{(2)}$, 以及量子态的到一阶微扰为止的近似 $|\psi_n^{(1)}\rangle$, 由上面给出的微扰展开方程以及 $G_0^\perp(E_n)$ 的定义, 我们可以得到

$$\begin{aligned} E_n^{(2)} &= \varepsilon_n + \langle n|V|n\rangle + \sum_{m \neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{E_n^{(0)} - \varepsilon_m} \\ &= \varepsilon_n + \langle n|V|n\rangle + \sum_{m \neq n} \frac{|\langle n|V|m\rangle|^2}{\varepsilon_n - \varepsilon_m}. \end{aligned} \quad (5.16)$$

$$|\psi_n^{(1)}\rangle = |n\rangle + \sum_{m \neq n} |m\rangle \frac{\langle m|V|n\rangle}{\varepsilon_n - \varepsilon_m}. \quad (5.17)$$

人们通常用瑞利-薛定谔(Rayleigh - Schrodinger)微扰理论得到这两个结论, 但是瑞利-薛定谔微扰理论推广到更高阶会比较复杂, 这里我们采用布里渊-维格纳(Brillouin Wigner)的办法来处理, 正如我们已经看到的, 这种处理办法的好处是, 可以很容易地得到任意阶微扰论的一般公式。

为了进一步理解二阶微扰修正, 我们注意到 ε_n 和其余所有能级之间存在一个有限的能隙 Δ , 即对于 $m \neq n$, $|\varepsilon_n - \varepsilon_m| \geq \Delta$ 。因此 $|\sum_{m \neq n} \frac{|\langle n|V|m\rangle|^2}{\varepsilon_n - \varepsilon_m}| \leq$

³注意 $E^{(1)}$ 中包含了一个因为微扰而来的一阶小量, 但是 $E^{(1)}$ 出现在表达式 $G_0^\perp(E_n^{(1)})$ 的分母上, 因此式中的 $G_0^\perp(E_n^{(1)})$ 也应该按照这个微扰一阶小量的级数展开来理解, 而且如果我们还要继续计算更高阶的 $E_n^{(i)}, i > 3$ 的话, 那 $G_0^\perp(E_n^{(1)})$ 只能取到微扰级数的1阶为止, 更高阶小量应该舍去, 因为否则的话我们就将本应归入更高阶微扰的贡献算进 $E_n^{(3)}$ 中了。类似的, 所有 $G_0^\perp(E_n^{(i)})$ 这样的表达式都应该理解成只取到微扰级数展开的第 i 阶小量为止, 更高阶小量要舍去。

$$\sum_{m \neq n} \frac{|\langle n|V|m\rangle|^2}{|\varepsilon_n - \varepsilon_m|} \leq \frac{1}{\Delta} (\sum_{m \neq n} \langle n|V|m\rangle \langle m|V|n\rangle) = \frac{1}{\Delta} [\langle n|V|(1 - |n\rangle\langle n|)|V|n\rangle] = \frac{1}{\Delta} [\langle n|V^2|n\rangle - (\langle n|V|n\rangle)^2] = \frac{1}{\Delta} \langle n|(\Delta V)^2|n\rangle$$
 (注意分母中的 Δ 表示能隙, 而分子中的 $\Delta V = V - \langle n|V|n\rangle$)。从上面推导过程最后的式子 $\frac{1}{\Delta} \langle n|(\Delta V)^2|n\rangle$ 可以看出, 二阶微扰修正实际上和 V 的涨落 $\langle n|(\Delta V)^2|n\rangle$ 密切相关, 如果没有这个 V 的涨落那就不会有二阶微扰修正。

另外, 从上面关于波函数的一阶微扰修正中我们很容易看出来微扰论适用的条件, 那就是

$$\left| \frac{\langle m|V|n\rangle}{\varepsilon_n - \varepsilon_m} \right| \ll 1. \quad (5.18)$$

5.1.3 范德瓦尔斯力、万有引力、以及射电天文学

这一节我们举三个应用二阶微扰论的例子。第一个例子是用二阶微扰论解释非极性分子之间的范德瓦尔斯力, 顺带也解释了一下为什么万有引力是一个吸引力, 其道理其实和非极性分子之间的范德瓦尔斯力是一个吸引力(当分子间的距离足够大时)有相通之处。在这个例子之后, 利用一个类似的偶极相互作用微扰项, 我们还用一阶微扰论估算了氢原子超精细能级分裂, 进而简单解释了射电天文学中著名的21cm波长谱线。最后, 我们要举的第三个例子是有关非线性振子能级的一个二阶微扰论计算。

非极性分子之间的范德瓦尔斯力

当分子间距离合适时, 范德瓦尔斯力是两个中性分子之间的吸引力。两个中性分子之间存在吸引力有时候并不难理解, 比方说, 如果这两个分子是极性分子, 那么它们的电偶极矩之间就可能存在相互作用, 这时候这两个分子之间就会存在范德瓦尔斯力。不过极性分子之间相互作用势能和它们的电偶极矩的相对取向有关系, 而由于热运动, 分子电偶极矩的空间取向有一定的随机性, 特别的, 当温度无限高时, 分子电偶极矩的空间取向将会完全随机, 因此两个分子各自电偶极矩的空间取向将完全无关, 这时候, 这种由非0的电偶极矩间的相互作用所产生的范德瓦尔斯力就会趋于0。

但是, 令人觉得不可思议的是, 即使是非极性分子之间也存在范德瓦尔斯力, 甚至大小并不比极性分子之间的范德瓦尔斯力弱。非极性分子之间的范德瓦尔斯力是怎么来的呢? 为什么在分子间的距离足够大时它是一个吸引力呢? 甚至为什么在距离足够大时, 这个力也和距离的7次方成反比

呢? 后面这两个问题可以合并成一个问题, 即为什么非极性分子之间在距离足够大时, 它们的相互作用能趋向于 $-\frac{A}{R^6}$ 呢? 这里 R 是两个分子之间的距离, $A > 0$ 是一个常数。要解释这里的物理规律就需要用到量子力学二阶微扰论。

首先, 在量子力学里面, 电偶极矩当然也要表示成算符, 对于任何多个带电粒子的体系, 这个算符总是能够定义的, 所有的分子(即使是非极性分子)都是有多个带电粒子的体系, 因此当然都有一个电偶极矩算符。当我们说一个分子是非极性分子时, 指的是在分子所处的量子态上, 其电偶极矩算符的期望值等于0。但是这个算符总是存在的。因此在算符的意义上, 任何两个分子的电偶极矩之间都会有相互作用, 并且给系统贡献一个电偶极相互作用项 V , 很显然, V 应该和每个分子的电偶极矩算符成正比。假设我们记所考察的这两个分子的电偶极矩算符分别为 \mathbf{d}^A 和 \mathbf{d}^B , 则它们的电偶极相互作用算符 V 显然将具有 $V = C_{ij}d_i^A d_j^B$ 这样的形式, 这里 C_{ij} 依赖于两个分子之间的距离, 并且显然只要分子间的距离足够大, 它应该就是随着距离衰减的, 同时这里的 $i, j = 1, 2, 3$ 表示三个直角坐标分量, 并且我们默认对重复的指标进行求和。

这样一来, 当分子间的距离足够大时(也就是当两个分子间的距离大于每个分子内部电子的德布罗意波长时), 它们的电偶极相互作用 V 就可以看成是对这个双分子系统的微扰。如果忽略掉这个微扰相互作用, 那么两个分子将相互独立(注意我们假设了分子间距离足够大, 从而两个分子内部电子的波函数不会出现交叠), 各自有其哈密顿量 H_0^A 和 H_0^B 。由于能量最低原理, 这两个分子将分别处在 H_0^A 的基态 $|\psi_0^A\rangle$, 和 H_0^B 的基态 $|\psi_0^B\rangle$, 它们的基态能量分别记作 ε_0^A 和 ε_0^B 。如果忽略 V 的扰动, 这个双分子系统的哈密顿量将是 $H_0 = H_0^A + H_0^B$ 。很显然 H_0 的基态将是 $|\psi_0\rangle = |\psi_0^A, \psi_0^B\rangle = |\psi_0^A\rangle|\psi_0^B\rangle$, 基态能量 ε_0 将是 $\varepsilon_0 = \varepsilon_0^A + \varepsilon_0^B$ 。

考虑到微扰 V , 那这个双分子系统的完整哈密顿量应该是 $H = H_0 + V$ 。对基态能量的一阶微扰修正将是 $\langle\psi_0|V|\psi_0\rangle = \langle\psi_0^B|\langle\psi_0^A|C_{ij}d_i^A d_j^B|\psi_0^A\rangle|\psi_0^B\rangle = C_{ij}\langle\psi_0^A|d_i^A|\psi_0^A\rangle\langle\psi_0^B|d_j^B|\psi_0^B\rangle$ 。但是我们前面已经说过, 非极性分子电偶极矩算符的期望值为0, 因此 $\langle\psi_0^A|d_i^A|\psi_0^A\rangle = \langle\psi_0^B|d_j^B|\psi_0^B\rangle = 0$, 从而 $\langle\psi_0|V|\psi_0\rangle = 0$ 。因此对于非极性分子, V 的一阶微扰对系统的能量没有贡献, 要解释范德瓦尔斯力我们需要进一步考虑二阶微扰。

根据二阶微扰论, 对基态能量的二阶微扰将是(我们将把二阶微扰对基

态能量的修正记为 ε')

$$\varepsilon' = \sum_{n \neq 0, m \neq 0} \frac{|\langle \psi_0^A, \psi_0^B | V | \psi_n^A, \psi_m^B \rangle|^2}{\varepsilon_0 - (\varepsilon_n^A + \varepsilon_m^B)}, \quad (5.19)$$

式中对于 $n \neq 0, m \neq 0$, ε_n^A 表示 H_0^A 的激发态能级, $|\psi_n^A\rangle$ 是相应的激发态, ε_m^B 表示 H_0^B 的激发态能级, $|\psi_m^B\rangle$ 同样是相应的激发态。因此 $|\psi_n^A, \psi_m^B\rangle = |\psi_n^A\rangle|\psi_m^B\rangle$ 就是整个 H_0 的激发态, $\varepsilon_n^A + \varepsilon_m^B$ 是相应的激发态能量本征值。

很显然 $\langle \psi_0^A, \psi_0^B | V | \psi_n^A, \psi_m^B \rangle = C_{ij} \langle \psi_0^A | d_i^A | \psi_n^A \rangle \langle \psi_0^B | d_j^B | \psi_m^B \rangle$ 。对于非极性分子, 虽然 d_i^A 和 d_j^B 在各自基态上的期望值都等于0, 但是, 由于每个分子都有可能形变, 从而都存在正负电荷中心不完全重合的激发态, 因此, 一般来说, 总存在一些激发态 $|\psi_n^A\rangle$ 和 $|\psi_m^B\rangle$ 使得 $\langle \psi_0^A | d_i^A | \psi_n^A \rangle \neq 0$, 同时 $\langle \psi_0^B | d_j^B | \psi_m^B \rangle \neq 0$, 这时候就会出现 $\langle \psi_0^A, \psi_0^B | V | \psi_n^A, \psi_m^B \rangle \neq 0$ 。这也就是说, 二阶微扰修正总是非0的。并且由于 $|\langle \psi_0^A, \psi_0^B | V | \psi_n^A, \psi_m^B \rangle|^2 \propto |C_{ij}|^2$, 所以很显然 $\varepsilon' \propto |C_{ij}|^2$ 。在物理上, 这是因为非极性分子的偶极矩在基态上的期望值虽然为0, 但是根据不确定原理, 偶极矩的涨落依然存在, 由于这样的涨落, 每个分子可能发生瞬时的变形, 产生瞬时的偶极矩, 这些瞬时的偶极矩会产生一个瞬时的电场作用在另一个分子上, 并使得另一个分子出现瞬时的极化, 而极化以后的另一个分子的瞬时偶极矩又会产生一个瞬时的电场反作用在第一个分子上。整个过程可以形象地说成是两个分子交换了一个瞬时的偶极子电场, 最后就表现为一个非0的二阶微扰相互作用。并且正是因为这种相互作用涉及到偶极电场的交换, 所以它的效应不是正比于与偶极子电场密切相关的 C_{ij} , 而是正比于它的模方。

但是, 公式(5.19)最重要的一个结论是, $\varepsilon' < 0$, 即对这个两分子系统基态能量的二阶微扰修正是负的。原因其实很简单, 因为在公式(5.19)中, 分母上的 $\varepsilon_n^A + \varepsilon_m^B$ 代表的是激发态的能量, 而 $\varepsilon_0 = \varepsilon_0^A + \varepsilon_0^B$ 代表的是未微扰的基态能量, 因此很显然 $\varepsilon_n^A + \varepsilon_m^B > \varepsilon_0$, 从而公式(5.19)的分母一定是负的, 这显然意味着整个表达式必然是负的, 即 $\varepsilon' < 0$ 。实际上, 这是一个普遍的结果, 在量子力学中, 不管什么系统, 对基态的二阶微扰修正总是小于0的, 人们可以很容易地用二阶微扰论的公式一般性地证明这一点。

这里我们插入一个简短的其它话题。正因为二阶微扰对基态能量的修正是负的, 所以在粒子物理中任何两个粒子如果交换一个标量粒子, 所产生的相互作用力将一定是一个吸引力。这是因为, 两个粒子交换一个标量粒子的过程可以看成是对这个两粒子系统的一个二阶微扰, 微扰之前系统处在这个被交换的标量粒子的真空态(也就是基态)上, 因此根据刚才的

推导，交换标量粒子的这个二阶微扰给系统能量带来的修正 ε' 一定小于0，即 $\varepsilon' < 0$ 。而另一方面，交换标量粒子所产生的相互作用强度当然随着两个粒子的距离 R 单调递减，即相互作用能的绝对值 $|\varepsilon'|$ 必然是 R 的减函数，因此 $\frac{\partial}{\partial R}|\varepsilon'(R)| < 0$ 。同时，由于 $\varepsilon' < 0$ ，所以 $-\varepsilon' = |\varepsilon'|$ ，因此根据保守力的公式，交换标量粒子所产生的相互作用力 $F = -\frac{\partial}{\partial R}\varepsilon'(R) = \frac{\partial}{\partial R}|\varepsilon'(R)| < 0$ ，从而必然是一个吸引力。在一定的意义上，万有引力之所以是一个吸引力也是因为类似的原因，当然引力子并不是一个标量粒子，但是物质通过交换引力子进行相互作用其道理和交换标量粒子其实有些类似。因此你也可以说，二阶微扰修正的负号同样可以解释为什么万有引力势能是负的。

上一段的讨论当然也完全适用于范德瓦尔斯力的情形。我们已经看到，这时候 $\varepsilon' < 0$ ，并且由于在分子间距离足够大时， C_{ij} 随着距离衰减，所以二阶微扰带来的非极性分子之间的相互作用能 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2$ 必然也随着距离衰减，因此根据上一段的讨论，这时候范德瓦尔斯力必定是一个吸引力。

下面我们进一步来理解为什么 ε' 反比于分子间距离的6次方。为此我们只需要研究 ε' 随着分子间距离是如何变化的(R 足够大时)，由于 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2$ ，所以我们又只需要确定 $C_{ij}(R)$ 的函数形式。下面的步骤其实就和量子力学没有什么关系了，完全是电动力学的内容。如果你还记得两个偶极子之间的相互作用势能反比于距离3次方的话，那你马上就知道， $C_{ij}(R) \propto 1/R^3$ ，从而 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2 \propto 1/R^6$ ，注意到 $\varepsilon' < 0$ ，那你马上就能得到答案

$$\varepsilon'(R) = -\frac{A}{R^6}, \quad (5.20)$$

在原则上， A 是某个可以根据上面描述的二阶微扰论算出来的大于0的常数，不过，对于两个比较复杂的分子，具体计算出 A 当然并不是一件容易的事情，但这是原子分子物理学家要做的事情，我们这里暂时忽略它。

假设你想知道得更具体一些，那么我们可以将两个分子的偶极相互作用描述成，一个分子的偶极子产生一个电场，这个电场作用在另一个分子的偶极矩上。电动力学里是这样描述的：首先，一个偶极子 d_i^A 在距离为 R 处(R 足够大)产生的电势为

$$\varphi = -\frac{1}{4\pi\epsilon_0}d_i^A\partial_i\left(\frac{1}{R}\right), \quad (5.21)$$

这里 $\partial_i = \frac{\partial}{\partial x_i}$ (x_i 是离分子 A 的位置为 \mathbf{x} 处的坐标分量，因此 $R^2 = x_i x_i$)，并且我们默认对重复的指标进行求和。其次， \mathbf{x} 位置的偶极子 d_j^B 处在 A 的电

场 φ 中，与电场 φ 的相互作用势能为

$$V = d_j^B \partial_j \varphi. \quad (5.22)$$

将结果(5.21)和结果(5.22)结合起来，我们就能够算出偶极子A与偶极子B之间的偶极相互作用势能为

$$V = -\frac{1}{4\pi\epsilon_0} d_i^A d_j^B \partial_i \partial_j \left(\frac{1}{R} \right) = \frac{1}{4\pi\epsilon_0} d_i^A d_j^B \left(\frac{\delta_{ij} - 3e_i e_j}{R^3} \right), \quad (5.23)$$

式中 $e_i = \frac{x_i}{R}$ 是从分子A指向分子B的单位矢量。从这个偶极相互作用表达式，我们可以读出

$$C_{ij} = \frac{\delta_{ij} - 3e_i e_j}{4\pi\epsilon_0 R^3}, \quad (5.24)$$

显然，它的确反比于距离3次方。(这一段的讨论请读者参考电动力学里关于电多极矩展开的相关内容)

超精细能级

完全类似的，读者很容易明白，两个磁偶极子也会有非常类似的偶极相互作用，它也可以写成 $V = C_{ij} \mu_i^A \mu_j^B$ 的形式，式中 μ_i^A 和 μ_j^B 分别是A, B两个磁偶极子的磁矩，不过由于是磁场，所以这时候

$$C_{ij} = \frac{\mu_0}{4\pi} \frac{\delta_{ij} - 3e_i e_j}{R^3}. \quad (5.25)$$

比如对于氢原子来说，电子是在质子所产生的场中运动，但是电子本身有自旋从而有磁矩，同样质子也有自旋因此也有磁矩，所以电子和质子的磁矩之间就会发生这种偶极相互作用(实际上电子与质子的磁矩相互作用比这要略微复杂一些，由于电子可能有非0的概率出现在质子所在的坐标原点，所以两者的磁矩相互作用中其实还要包含一个接触项)。不过这时候这种偶极相互作用常常会有非0的一阶微扰效应，实际上，它是造成氢原子能级超精细分裂的原因之一。比方说，它会引起氢原子的基态产生超精细分裂。注意到这是一个一阶微扰效应，所以我们很容易估算这种超精细分裂的大小。首先，它正比于电子磁矩和质子磁矩的乘积，根据玻尔磁子的数值我们很容易估算，电子磁矩大约是 10^{-23} (单位 $A \cdot m^2$ ，即安培乘以米的平方)，由于质子质量比电子重1000多倍，所以质子磁矩大约是电子磁矩的1/1000，也就是大约 $10^{-26} A \cdot m^2$ ，将这两者乘起来再乘以 $\mu_0/(4\pi) = 10^{-7} N/A^2$ ，就

得到 $10^{-56} N \cdot m^4$ 。由于 C_{ij} 反比于电子与质子距离的3次方，而对于基态上的氢原子 $R \sim a_0 \sim 0.5 \times 10^{-10} m$ (a_0 是玻尔半径)，所以 $1/R^3 \sim 10^{31} m^{-3}$ 。将 $10^{-56} N \cdot m^4$ 和 $1/R^3 \sim 10^{31} m^{-3}$ 乘起来就是能级超精细分裂的量级，显然结果是

$$10^{-25} N \cdot m \sim 10^{-6} eV. \quad (5.26)$$

将这个结果除以普朗克常数 $h \sim 6.62607015 \times 10^{-34} J \cdot s$ ，得到 $10^9 Hz = 1000 MHz$ ，这就是对氢原子超精细能级跃迁发出来的光频率的估算。更仔细的计算可以得出氢原子基态的超精细分裂所发出来的光频率为 $1420 MHz$ ，波长在 $21 cm$ 。

这个 $21 cm$ 谱线在射电天文学中有很重要的应用。因为星系中的氢原子会发射或吸收这一波长的辐射，所以用调谐到 $1420 MHz$ 的射电望远镜，我们就能观察星系中氢原子气体密集处的位置和速度，进而就能了解星系的位置和运动情况。射电望远镜利用的这个 $21 cm$ 谱线，波长在微波波段，这个波长的好处在于，它远比无线电波要短，因此有很好的定向性。同时它的波长又远比可见光的几百个纳米要长，因此也不像可见光那样容易被散射，因此射电望远镜能够看到很多光学望远镜看不到的天文事物。例如，通过射电望远镜，人们发现了天鹅座A的射电星系，它每秒钟发出的射电能量要比太阳每秒钟发出的能量强1亿亿倍以上，是迄今发现的最大射电星系，而光学望远镜对它却是一无所知。尤其是，宇宙微波背景辐射也是用射电望远镜发现的，因为微波背景辐射的波长当然在微波波段。目前世界上最大的射电望远镜，就是中国的位于贵州平塘的500米口径球面射电望远镜，简称FAST，俗称中国天眼。

非线性振子的能级

我们要考察的第二个例子是一个非线性振子。具体来说就是给单自由度线性谐振子 $H_0 = \frac{P^2}{2m} + \frac{1}{2} m \omega^2 X^2$ 加上一个非线性微扰

$$V = \lambda \hbar \omega \left(\frac{\hbar}{m \omega} \right)^{-2} X^4, \quad (5.27)$$

式中我们已经注意到 $\left(\frac{\hbar}{m \omega} \right)^{\frac{1}{2}}$ 具有长度量纲， $\hbar \omega$ 具有能量量纲，因此将 V 写成这种形式是为了使得式中的 $\lambda \ll 1$ 是一个无量纲的常数。因此，在作了 $\hbar = m = \omega = 1$ 的无量纲化处理以后，我们的微扰相互作用就是简单的

$$V = \lambda X^4. \quad (5.28)$$

下面我们就是要用二阶微扰论来计算这个微扰给谐振子的能级造成的影响。

为此我们引入产生湮灭算符 $a = \frac{1}{\sqrt{2}}(X + iP)$, $a^\dagger = \frac{1}{\sqrt{2}}(X - iP)$ 。则无量纲化的 H_0 就可以表示成 $H_0 = a^\dagger a + \frac{1}{2}$, 其本征值为 $n + \frac{1}{2}$ ($n \geq 0$), 相应的本征态我们记作 $|n\rangle$ 。并且我们有代数关系 $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, $a|n\rangle = \sqrt{n}|n-1\rangle$, 以及 $[a, a^\dagger] = 1$ 。

很显然, 由产生湮灭算符的定义我们有 $X = \frac{1}{\sqrt{2}}(a + a^\dagger)$, 代入(5.28)并不断地利用产生湮灭算符的代数关系 $[a, a^\dagger] = 1$, 我们就可以将微扰相互作用重写成

$$\begin{aligned} V &= \frac{\lambda}{4}(a + a^\dagger)^4 = \lambda[a^\dagger a^\dagger(a^\dagger a) + (a^\dagger a)aa] + \frac{3\lambda}{2}[a^\dagger a^\dagger + aa] \\ &+ \frac{\lambda}{4}[a^4 + (a^\dagger)^4] + \frac{3\lambda}{2}[(a^\dagger a)^2 + (a^\dagger a) + \frac{1}{2}]. \end{aligned} \quad (5.29)$$

从这个表达式我们很容易得到 V 的非0矩阵元,

$$\begin{aligned} \langle n+4|V|n\rangle &= \langle n|V|n+4\rangle = \lambda \frac{1}{4} \sqrt{(n+1)(n+2)(n+3)(n+4)} \\ \langle n+2|V|n\rangle &= \langle n|V|n+2\rangle = \lambda \left(n + \frac{3}{2}\right) \sqrt{(n+1)(n+2)} \\ \langle n|V|n\rangle &= \lambda \frac{3}{2} \left(n^2 + n + \frac{1}{2}\right). \end{aligned} \quad (5.30)$$

为了得到这些结果, 我们要不断地使用 $(a^\dagger a)|n\rangle = n|n\rangle$ 以及 $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, $a|n\rangle = \sqrt{n}|n-1\rangle$ 这样的关系式。除了这些矩阵元之外, V 的其它矩阵元都是0。

利用二阶微扰论的公式(5.16), 并代入(5.30)式给出来的 V 的矩阵元, 注意到对于第 n 能级的二阶微扰修正只有 $n \pm 2$ 能级和 $n \pm 4$ 能级有贡献, 进而我们就能够得到能级 E_n 直到二阶微扰的结果,

$$\begin{aligned} E_n/(\hbar\omega) &= \lambda^2 \frac{1}{64} [(n-3)(n-2)(n-1)n - (n+1)(n+2)(n+3)(n+4)] \\ &+ \lambda^2 \frac{1}{2} \left[\left(n - \frac{1}{2}\right)^2 (n-1)n - \left(n + \frac{3}{2}\right)^2 (n+1)(n+2) \right] \\ &+ n + \frac{1}{2} + \lambda \frac{3}{2} \left(n^2 + n + \frac{1}{2}\right) \\ &= n + \frac{1}{2} + \lambda \frac{3}{2} \left(n(n+1) + \frac{1}{2}\right) - \lambda^2 \frac{1}{8} (2n+1)(17n(n+1) + 21). \end{aligned}$$

这里我们已经恢复了量纲。很显然, 这个结果有其适用范围, 它只对比较低的能级适用, 具体来说就是要求 $n^2\lambda \ll 1$ 。也即是说, 对于非常高的激

发态，微扰论的计算其实是失效的。原因在于对于高激发态，振子的能量很高，因此振幅就会比较大，也即是说波函数在 X 取值很大的地方也会有非0值。而当 X 很大的时候， λX^4 就会变得很重要，这时候它就不再能当作微扰来看了，事实上，这时候这一项比标准的 X^2 项更重要。因此，前面的微扰论处理当然就失效了。

5.1.4 习题

1. 假设某两态系统的哈密顿量可以写成

$$H = \begin{pmatrix} \varepsilon_1 + a\epsilon & b\epsilon \\ b\epsilon & \varepsilon_2 \end{pmatrix}. \quad (5.31)$$

式中 $\varepsilon_1 \neq \varepsilon_2$ ， a, b 为实数， ϵ 非常小。(1). 请用二阶微扰算出系统的两个能量本征值。(2). 请求出以上哈密顿量的精确本征值，然后将它们按照 ϵ 进行级数展开，并将结果与二阶微扰的结果进行比较。

2. 两个全同的零自旋玻色子放在一个无限深的方势阱中，阱内势能为零，阱外势能为正无穷，势阱的宽度为 a 。假设这两个粒子之间存在相互作用 $V(x_1, x_2) = -V_0 a \delta(x_1 - x_2)$ 。(1). 忽略两粒子间的相互作用，请求出系统的基态和第一激发态(包括能量本征值和相应的本征波函数)。(2). 请用一阶微扰论计算相互作用对基态和对第一激发态的能量修正。

3. 一个刚性转子绕着固定的 z 轴转动，转动惯量为 I 。现在用一个 x 方向的均匀电场 E 来扰动此系统，假设转子的电偶极矩大小为 D ，请用二阶微扰论计算转动的能量本征值和相应的本征波函数。

5.2 *有效哈密顿量、简并微扰论以及不稳定态

上一节我们推导定态微扰论的相关公式时有一个关键性的假设，即假设我们考察的能级 n 非简并，我们假设 $|n\rangle$ 态和其余所有能态之间有一个有限的能隙 Δ 。但其实有时候我们也会碰到这种情况，即作为 H_0 的本征态，我们所关心的这个能态 $|n\rangle$ 其实是简并的，那这个时候微扰论该怎么处理呢？这就是我们这一节要探讨的内容之一，通常称作简并定态微扰理论。

简并定态微扰理论的核心思想是，消去量子态在简并子空间之外的分量，将问题约化到简并子空间里来处理。这种降低希尔伯特空间的维数，

将物理问题约化到物理上更相关的希尔伯特子空间里来处理的思想其实很重要，它本身有很广泛的用途。在这一节中我们将首先用有效哈密顿量的概念来介绍这一思想。然后将简并定态微扰理论作为有效哈密顿量方法的一个应用。最后，作为有效哈密顿量方法的另一个应用，我们还将用它来讨论与连续谱耦合的态的不稳定性。

5.2.1 有效哈密顿量

在讨论简并微扰论之前，我们先来讨论一下所谓的有效哈密顿量方法。所谓的有效哈密顿量就是在求解量子力学问题时，我们可以先把一部分我们不关心的量子态信息消去，进而将系统约化到某个希尔伯特子空间上所得到的对系统的有效描述。

具体来说，假设我们要求解定态薛定谔方程

$$H|\psi\rangle = E|\psi\rangle. \quad (5.32)$$

假设系统的整个希尔伯特空间可以分解成两个正交补空间，分别对应投影算子 P 和 P^\perp ，它们满足

$$P + P^\perp = 1, \quad (5.33)$$

此外还满足 $PP^\perp = P^\perp P = 0$ 。因此我们可以将 $|\psi\rangle$ 态正交分解成 $|\psi\rangle = P|\psi\rangle + P^\perp|\psi\rangle$ ，其中 $P^\perp|\psi\rangle$ 的信息是我们不关心的，我们想推导的就是，消去 $P^\perp|\psi\rangle$ 以后， $P|\psi\rangle$ 满足的方程是什么？

为此，我们注意到，我们所求解的定态薛定谔方程(5.32)可以重写为 $H(P|\psi\rangle) + H(P^\perp|\psi\rangle) = E|\psi\rangle$ ，分别用投影算符 P 和 P^\perp 作用在这个方程上，就可以得到

$$\begin{aligned} H_{PP}(P|\psi\rangle) + H_{PP^\perp}(P^\perp|\psi\rangle) &= E(P|\psi\rangle) \\ , \quad H_{P^\perp P}(P|\psi\rangle) + H_{P^\perp P^\perp}(P^\perp|\psi\rangle) &= E(P^\perp|\psi\rangle), \end{aligned} \quad (5.34)$$

式中 $H_{PP} = PHP$ 为哈密顿算符在 P 子空间里的作用， $H_{P^\perp P^\perp} = P^\perp H P^\perp$ 为哈密顿算符在 P^\perp 子空间里的作用，而 $H_{PP^\perp} = (H_{P^\perp P})^\dagger = PHP^\perp$ 将两个正交的子空间耦合了起来。

由(5.34)的第二个方程可以得到 $H_{P^\perp P}(P|\psi\rangle) = (E - H_{P^\perp P^\perp})(P^\perp|\psi\rangle)$ ，由于我们关心的是 P 子空间里的物理，所以可以假设能量 E 和 $H_{P^\perp P^\perp}$ 的所

有本征值都不同, 则我们就有 $\frac{1}{E - H_{P^\perp P^\perp}} H_{P^\perp P}(P|\psi) = (P^\perp|\psi)$ 。将这个结果代入(5.34)的第一个方程消去 $(P^\perp|\psi)$ 就可以得到

$$\left(H_{PP} + H_{PP^\perp} \frac{1}{E - H_{P^\perp P^\perp}} H_{P^\perp P} \right) (P|\psi) = E(P|\psi). \quad (5.35)$$

换言之, 假设对于任意探测能量 $z = \hbar\omega$ 我们定义作用在 P 子空间里的有效哈密顿量 $H_{\text{eff}}(z)$

$$H_{\text{eff}}(z) = H_{PP} + H_{PP^\perp} \frac{1}{z - H_{P^\perp P^\perp}} H_{P^\perp P}. \quad (5.36)$$

那么满足这个有效哈密顿量本征方程 $H_{\text{eff}}(E)|\varphi\rangle = E|\varphi\rangle$ 的本征值 E 就是原哈密顿量 H 的本征值, 而 $|\varphi\rangle$ 就是 H 的相应本征态 $|\psi\rangle$ 在 P 子空间里的投影, 即 $|\varphi\rangle = P|\psi\rangle$ 。换言之, 对于探测能量 $z = \hbar\omega$, P 子空间里所有量子态的动力学都完全由有效哈密顿量 $H_{\text{eff}}(z)$ 决定。

实际上, 可以证明,

$$P \frac{1}{z - H} P = \frac{1}{z - H_{\text{eff}}(z)}. \quad (5.37)$$

证明方法是: 在 P 和 P^\perp 两个正交子空间中, 分别将哈密顿量 H 和投影算符 P 表示成如下分块矩阵

$$\begin{pmatrix} H_{PP} & H_{PP^\perp} \\ H_{P^\perp P} & H_{P^\perp P^\perp} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (5.38)$$

则 $P \frac{1}{z - H} P$ 就相当于

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z - H_{PP} & -H_{PP^\perp} \\ -H_{P^\perp P} & z - H_{P^\perp P^\perp} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (5.39)$$

最后, 利用矩阵代数里分块矩阵求逆矩阵的方法就可以得到上式的结果是

$$\begin{pmatrix} \frac{1}{z - H_{\text{eff}}(z)} & 0 \\ 0 & 0 \end{pmatrix}. \quad (5.40)$$

这就已经证明了我们需要的结果(5.37)。

一个简单的例子

让我们来考察一个最简单的例子，假设某系统的哈密顿量 H 可以写成一个 2×2 的矩阵，

$$H = \begin{pmatrix} U & t \\ t & 0 \end{pmatrix}. \quad (5.41)$$

我们可以消去这个系统的第1行和第1列，从而得到一个作用在第2行第2列的子空间里的有效哈密顿量 $H_{\text{eff}}(z)$ 。由(5.36)式，我们有

$$H_{\text{eff}}(z) = \frac{t^2}{z - U}. \quad (5.42)$$

如果我们要求原哈密顿量 H 的本征值 E ，那根据有效哈密顿量方法，我们可以在约化后的子空间里解方程 $H_{\text{eff}}(E)|\psi\rangle = E|\psi\rangle$ ，由于现在只有一行一列，所以本征值 E 必定满足

$$E = H_{\text{eff}}(E) = \frac{t^2}{E - U}, \quad (5.43)$$

很显然，这个方程的确是原哈密顿量 H 的本征值方程。

就我们这个简单例子来说，方程(5.43)很容易精确求解。但这里我们想描述一种更具一般性的求解方法，为此我们假设 $t \ll U$ 。这时候我们可以这样来求解(5.43)，简单来说即是迭代，首先，我们把 t 看成是一个微扰，很显然如果忽略这个微扰，那第2行第2列子空间里的本征值将是 $E = 0$ 。因此我们将 $E = 0$ 代入方程(5.43)右边，这样就会得到 $E = -\frac{t^2}{U}$ ，接着再将 $E = -\frac{t^2}{U}$ 代入方程(5.43)右边，近似就会得到 $E = -\frac{t^2}{U} + \frac{t^4}{U^3}$ ，如此不断迭代下去，就可以得到一个关于 t/U 的无穷级数，如果这个级数收敛，那它最终就会收敛到真正的能量本征值。

5.2.2 简并微扰论

下面我们将上一小节的有效哈密顿量理论用于简并微扰论。为此我们取投影算符 P 对应于 H_0 的某个简并度为 d 的简并能级所对应的简并子空间。不妨将这 d 个简并的能态记为 $|1\rangle, |2\rangle, \dots, |d\rangle$ ，我们以 $\mathcal{D} = \{i = 1, 2, 3, \dots, d\}$ 来表示这 d 个简并态的集合，我们记这 d 个态所对应的 H_0 本征值为 ε 。那么投影算符 P 就可以写成

$$P = \sum_{i=1}^d |i\rangle\langle i|. \quad (5.44)$$

同样，我们以 P^\perp 来表示 H_0 的所有正交归一本征态中除了我们正在关心的这 d 个简并态之外，其余所有本征态(即 $\{|m\rangle, m \in \mathcal{D}^\perp\}$)共同张成的希尔伯特子空间的投影算符，我们记这些非简并本征态的 H_0 本征值为 ε_m 。如果按照我们在正文中所引入的记号，那显然有 $P = P_{\mathcal{D}}$, $P^\perp = P_{\mathcal{D}^\perp}$ 。当然， P^\perp 和 P 之间就不再有任何简并了，而是有有限的能隙 Δ 。另外，由于 P 和 P^\perp 都是用 H_0 的本征态定义的，所以很显然， P 和 P^\perp 均与 H_0 对易。

现在我们加入微扰 V ，因此完整的哈密顿量应该是 $H = H_0 + V$ 。假设考虑到微扰以后，原来 H_0 的那个 d 重简并子空间分裂成了 H 的 d 个不同本征态 $\{|\psi_l\rangle, l = 1, 2, 3, \dots, d\}$ ，它们分别满足本征方程

$$H|\psi_l\rangle = E_l|\psi_l\rangle. \quad (5.45)$$

我们将要做的，就是要研究如何用微扰展开的办法计算 E_l 和 $|\psi_l\rangle$ 。

很显然，我们可以利用上一小节的有效哈密顿量方法先求解出 $P|\psi_l\rangle$ ，为此我们需要先计算有效哈密顿量 $H_{\text{eff}}(z)$ 。由于 $H_{PP} = PH_0P + PVP = \varepsilon + V_{PP}$, $H_{PP^\perp} = PH_0P^\perp + PVP^\perp = H_0PP^\perp + V_{PP^\perp} = V_{PP^\perp}$ ，同理 $H_{P^\perp P} = P^\perp VP = V_{P^\perp P}$ 。因此我们很容易得到有效哈密顿量 $H_{\text{eff}}(z)$ 为

$$H_{\text{eff}}(z) = \varepsilon + V_{PP} + V_{PP^\perp} \frac{1}{z - H_{P^\perp P^\perp}} V_{P^\perp P}. \quad (5.46)$$

式中 $H_{P^\perp P^\perp} = H_0^\perp + V_{P^\perp P^\perp}$ ，其中 $H_0^\perp = P^\perp H_0 P^\perp$ 。从而 $P|\psi_l\rangle$ 满足的方程即是

$$H_{\text{eff}}(E_l)(P|\psi_l\rangle) = E_l(P|\psi_l\rangle). \quad (5.47)$$

假设我们定义一个有效的“未微扰”哈密顿量 $H_{\text{eff},0}$ ，和一个有效的微扰 $V_{\text{eff}}(E_l)$ ，它们的定义分别是

$$H_{\text{eff},0} = \varepsilon + V_{PP}, \quad V_{\text{eff}}(E_l) = V_{PP^\perp} \frac{1}{E_l - H_{P^\perp P^\perp}} V_{P^\perp P}. \quad (5.48)$$

从而， $H_{\text{eff}}(E_l) = H_{\text{eff},0} + V_{\text{eff}}(E_l)$ 。那么关于 $P|\psi_l\rangle$ 的方程(5.47)显然就可以根据标准的非简并定态微扰论来求解。也即是说，我们可以首先在 P 子空间中求出 $H_{\text{eff},0}$ 的 d 个本征态 $|\phi_l\rangle$ ，

$$H_{\text{eff},0}|\phi_l\rangle = \tilde{\varepsilon}_l|\phi_l\rangle. \quad (5.49)$$

一般来说，因为定义中包含了 V ，所以 $H_{\text{eff},0}$ 的 d 个本征态通常就不再简并了，那么下面我们就可以将 $V_{\text{eff}}(E_l)$ 当成微扰，然后照搬上一节中发展的非简并微扰论的办法来求解方程(5.47)了。

比方说, 如果保留到微扰小量 V 的二阶(由于 $V_{\text{eff}}(E_l)$ 本身是 V 的二阶项, 因此当然就只需要算到 $V_{\text{eff}}(E_l)$ 的一阶微扰修正 $\langle\phi_l|V_{\text{eff}}(E_l)|\phi_l\rangle$), 按照上面的方法我们可以得到

$$E_l = \tilde{\varepsilon}_l + \sum_{i,j \in \mathcal{D}, m \in \mathcal{D}^\perp} \frac{\langle\phi_l|i\rangle\langle i|V|m\rangle\langle m|V|j\rangle\langle j|\phi_l\rangle}{\varepsilon - \varepsilon_m}. \quad (5.50)$$

注意, 上式右边的第二项我们代入了 E_l 的零阶近似, 而 $\tilde{\varepsilon}_l$ 中是包含了 V 的一阶修正的, E_l 的0阶近似应该是 ε 。另外, 这个式子中的 $\langle i|\phi_l\rangle$ 是 $|\phi_l\rangle$ 在原来的简并基 $\{|i\rangle, i = 1, 2, 3, \dots, d\}$ 中的分量形式。其实在关于简并微扰论的实际应用中, 最有用也是用得最多的就是这个结果。

以上就是一般性的简并微扰理论, 我们最终是把它转变为非简并微扰的情形来处理的。上面讨论的简并微扰理论有一种特别简单的情形, 那就是一开始的时候 H_0 的所有本征态都简并, 那这种情况很简单, 即有 $P = 1$, $P^\perp = 0$, 由(5.48)式可以看出, 这时候 $V_{\text{eff}}(E_l) = 0$, $H_{\text{eff},0} = \varepsilon + V$, 所以这时候只需要直接求 V 的本征态和本征值就可以了。

当然, 前面还留有一个小问题, 那就是逆算符 $\frac{1}{E_l - H_0^\perp - V_{P^\perp P^\perp}}$ 如何处理。一般来说, 我们可以将它按照 $V_{P^\perp P^\perp}$ 进行级数展开, 从而得到

$$\begin{aligned} & \frac{1}{E_l - H_0^\perp - V_{P^\perp P^\perp}} \\ &= \frac{1}{E_l - H_0^\perp} + \frac{1}{E_l - H_0^\perp} V_{P^\perp P^\perp} \frac{1}{E_l - H_0^\perp} \\ &+ \frac{1}{E_l - H_0^\perp} V_{P^\perp P^\perp} \frac{1}{E_l - H_0^\perp} V_{P^\perp P^\perp} \frac{1}{E_l - H_0^\perp} + \dots \end{aligned}$$

这里我们利用了关于可逆算符 A 的恒等式

$$\begin{aligned} & \left(\frac{1}{A} + \frac{1}{A} B \frac{1}{A} + \frac{1}{A} B \frac{1}{A} B \frac{1}{A} + \dots \right) (A - B) = 1 \\ & (A - B) \left(\frac{1}{A} + \frac{1}{A} B \frac{1}{A} + \frac{1}{A} B \frac{1}{A} B \frac{1}{A} + \dots \right) = 1. \end{aligned}$$

在形式上, 读者可以通过直接将这两个等式左边乘出来进而验证它们。不过, 严格一点来说, 这两个等式仅在括号里的那个无穷级数收敛时才有意义, 因此通常只在 B 是一个微扰时才能这么做。很显然, 这两个恒等式意味着

$$\frac{1}{A - B} = \frac{1}{A} + \frac{1}{A} B \frac{1}{A} + \frac{1}{A} B \frac{1}{A} B \frac{1}{A} + \dots \quad (5.51)$$

5.2.3 不稳定离散态

这一节我们将利用有效哈密顿量的方法处理一类普遍存在而又重要的物理现象，那就是不稳定离散态。这类现象之所以普遍，首先是因为，由于和环境耦合，基本上所有束缚系统的激发态都不稳定，都会衰变，从而都属于不稳定离散态。比方说，原子的激发态能级就不稳定，会向低能级跃迁。其次，无论是粒子物理中，还是凝聚态物理中，都有大量不稳定的粒子，比方说著名的上帝粒子(Higgs)粒子就是不稳定粒子，比方说凝聚态物理里面大量的准粒子其实也都是不稳定粒子。这些不稳定的，会衰变成其它粒子的粒子其实也都属于不稳定离散态。不过，对于这种不稳定粒子的情形，我们得稍微解释一下离散的含义。虽然不稳定粒子的动量是连续的，但是，在单个这样的不稳定粒子的衰变过程中，动量是守恒的，因此我们完全可以在一个总动量固定不变的子空间里来考察问题，而在保持动量不变的前提下，单个不稳定粒子当然也是离散态。

离散态的衰变不稳定性其实有一个共同的本质，那就是它们和一个较宽的连续谱有耦合。比方说原子向低能级跃迁放出光子，由于能量守恒，放出的这个光子的频率当然是分立的，但是它从属于一个连续谱。一个静止的不稳定粒子衰变成多个粒子，比方说衰变成两个粒子，这两个粒子的总动量保持是0，但是它们各自的动量和能量分配都连续可变，因此衰变的末态也是一个连续谱。总之，离散态之所以不稳定，都是因为它们可以衰变到一个较宽的连续谱。

为了用有效哈密顿量的方法分析与连续谱耦合的离散态的不稳定性，我们记这些正交归一离散态为 $\{|i\rangle, i = 1, 2, 3, \dots, N\}$ ，记这些离散态张成的希尔伯特子空间为投影算符 P ， $P = \sum_{i=1}^N |i\rangle\langle i|$ 。假设与这些离散态耦合的连续谱态张成了 P 的正交补空间 P^\perp 。假设在 P^\perp 子空间上，有效哈密顿量中的 $H_{P^\perp P^\perp}$ 可以表示成

$$H_{P^\perp P^\perp} = \int d\alpha E_\alpha |\alpha\rangle\langle\alpha|. \quad (5.52)$$

从而 $|\alpha\rangle$ 构成 P^\perp 的正交基，满足正交归一关系

$$\langle\alpha|\alpha'\rangle = \delta(\alpha - \alpha'). \quad (5.53)$$

这里指标 α 是用来标记这组连续谱态的抽象记号，它必须包含一个或多个连续指标，当然， α 里面也可以包含离散指标，对于 α 里面有离散指标的情形，人们只需将正交归一关系中相应的狄拉克 δ 函数理解成克龙内克 δ 符号，同时将相应的指标积分理解成离散求和就可以了！

由于 $|\alpha\rangle$ 是连续谱态， E_α 当然是连续变量。很多时候，我们常常选能量本身作为 $|\alpha\rangle$ 态的指标之一，这时候我们重新标记 $|\alpha\rangle$ 态为， $|\alpha\rangle = |E, \chi\rangle$ ， E 在数值上就等于 E_α ， χ 是用来进一步区分 $|\alpha\rangle$ 态的其它指标。作了这样的指标变换以后，态空间微元 $d\alpha$ 一般就会变换成

$$d\alpha = \rho(E, \chi) dE d\chi, \quad (5.54)$$

式中 $\rho(E, \chi)$ 是指标变换的雅可比行列式，也叫做态密度，因为它表示单位能量区间单位 χ 微元之内 $|\alpha\rangle$ 态的数目。

由于离散态和连续谱态之间有耦合，所以 $(H_{P^\perp P})^\dagger = H_{PP^\perp} \neq 0$ 。不妨记 $\langle i|H_{PP^\perp}|\alpha\rangle = \langle i|H|\alpha\rangle = A_{i\alpha}$ ，式中我们用到了 $|\alpha\rangle$ 在 P^\perp 中的投影是它本身，同样 $|i\rangle$ 在 P 中的投影也是它本身的事实。并且，我们常常记 $\sum_{i=1}^N |i\rangle A_{i\alpha} = A_\alpha$ 。因此 $\sum_{i=1}^N \langle \alpha|H_{P^\perp P}|i\rangle \langle i| = \sum_{i=1}^N A_{i\alpha}^* \langle i| = A_\alpha^\dagger$ 。如此一来，根据(5.36)式，离散态子空间 P 上的有效哈密顿量 $H_{\text{eff}}(z)$ 就可以写成

$$\begin{aligned} H_{\text{eff}}(z) &= H_{PP} + \int d\alpha \frac{A_\alpha A_\alpha^\dagger}{z - E_\alpha} \\ &= H_{PP} + \int dE d\chi \rho(E, \chi) \frac{A_{E,\chi} A_{E,\chi}^\dagger}{z - E}. \end{aligned} \quad (5.55)$$

很显然，如果探测能量 $z = \hbar\omega$ 落在连续能谱之外，那上面的(5.55)式定义良好。但是，如果探测能量 $z = \hbar\omega$ 落在连续能谱的区间之内，那在 $E = z$ 处，上面(5.55)的两个式子右边的第二项都将没有定义。为了避免这种没有定义的困难，我们可以给探测能量 $\hbar\omega$ 加一个无穷小的正虚部，即取 $z = \hbar\omega^+ = \hbar\omega + i\epsilon$ （至于为什么要这样处理我们稍后再解释），最后再取 $\epsilon \rightarrow 0$ 。利用数学公式(式中 \mathcal{P} 表示取主值部分)

$$\lim_{\epsilon \rightarrow 0} \frac{1}{x + i\epsilon} = \mathcal{P} \frac{1}{x} - i\pi\delta(x), \quad (5.56)$$

我们就可以将 $H_{\text{eff}}(\hbar\omega)$ 重写成

$$H_{\text{eff}}(\hbar\omega) = H_{PP} + \Delta(\hbar\omega) - \hbar \frac{i}{2} \Gamma(\hbar\omega), \quad (5.57)$$

式中， $\Delta(\hbar\omega)$ 和 $\Gamma(\hbar\omega)$ 均为作用在离散态子空间 P 上的厄米算符，它们分别是

$$\Delta(\hbar\omega) = \mathcal{P} \int dE d\chi \rho(E, \chi) \frac{A_{E,\chi} A_{E,\chi}^\dagger}{\hbar\omega - E} = \mathcal{P} \int d\alpha \frac{A_\alpha A_\alpha^\dagger}{\hbar\omega - E_\alpha}, \quad (5.58)$$

$$\Gamma(\hbar\omega) = \frac{2\pi}{\hbar} \int d\chi A_{E,\chi} A_{E,\chi}^\dagger \rho(E, \chi) |_{E=\hbar\omega}. \quad (5.59)$$

可见，由于和连续谱的耦合，有效哈密顿量 $H_{\text{eff}}(\hbar\omega)$ 不再是一个厄米算符！尤其是，它有了一个负虚部 $-\hbar\frac{i}{2}\Gamma(\hbar\omega)$ ，下面我们将看到，这个负虚部的存在将意味着 P 子空间里的态是不稳定的，会随着时间指数衰减，从而表现出衰变行为。

值得提及的是，在一些场合中，人们也常常将结果(5.57)中的 $\Delta(\hbar\omega) - \hbar\frac{i}{2}\Gamma(\hbar\omega)$ 记作 $\Sigma(\hbar\omega)$ ，即 $\Sigma(\hbar\omega) = \Delta(\hbar\omega) - \hbar\frac{i}{2}\Gamma(\hbar\omega)$ ，称为自能。

在清楚地看到自能的负虚部会导致量子态的指数衰变之前，让我们先来理解前面碰到(5.55)式时，为什么要作 $z = \hbar\omega + i\epsilon$ 的处理，这个无穷小的正虚部从何而来。为此，我们注意到，对于任何不显含时间的哈密顿量 H ，当时间 $t > 0$ 时，时间演化算符 $e^{-iHt/\hbar}$ 可以表示成，

$$e^{-iHt/\hbar} = \frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega + i\epsilon - H}. \quad (5.60)$$

为了证明这个式子，我们将它作用在 H 的任意一个本征值为 E 的本征态 $|E\rangle$ 上，这时候方程(5.60)左边将给出时间演化因子 $e^{-iEt/\hbar}$ ，而方程(5.60)右边的时间演化因子是积分 $\frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega + i\epsilon - E}$ 。我们将这个积分的被积函数延拓到整个 ω 复平面上，注意到 $t > 0$ ，所以被积函数在 ω 的下半平面指数衰减，这是因为在复下半平面 ω 有一个负虚数，从而指数因子 $e^{-i\omega t}$ 呈现出指数衰减，特别是，在下半平面无穷远处的半圆上， $e^{-i\omega t}$ 将衰减为0。因此，沿着实轴的积分 $\frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega + i\epsilon - E}$ 其实等于沿着实轴然后再加下半平面无穷远处的半圆这样一个闭合围道 C 上的积分，如图(5.1)所示。

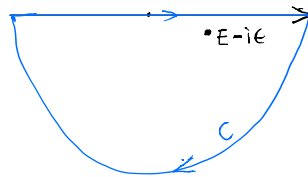


图 5.1: ω 复平面上的围道积分，最终要取极点 $E - i\epsilon$ 的虚部 $\epsilon \rightarrow 0$ 。

从图(5.1)也可以清楚地看到，当将 ω 加上一个无穷小正虚部以后，围道积分 $\frac{i}{2\pi} \oint_C \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega + i\epsilon - E}$ 将包含极点 $E - i\epsilon$ ，从而根据留数定理我们可以得到 $\frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega + i\epsilon - E} = e^{-iEt/\hbar}$ ，刚好与方程(5.60)左边的结果相同。由于方程(5.60)作用在任意一个能量本征态 $|E\rangle$ 上都成立，而任意量子态一定能够写成能量本征态的线性叠加，从而方程(5.60)作用在任意量子态上都成立，从而方程本身成立。

从上面对方程(5.60)的证明过程中我们清楚地看到, 为了从 $\frac{1}{\hbar\omega-E}$ 这样的表达式中得到正确的量子态随时间的演化, 我们必须给 $\hbar\omega$ 加上一个无穷小的正虚部。这也就是我们前面碰到(5.55)式时那样处理的原因。

现在我们来考察离散态子空间 P 之内的量子态如何随时间演化, 这时候决定时间演化的当然是有效哈密顿量 $H_{\text{eff}}(\hbar\omega)$ 。实际上, 将量子态的时间演化方程(5.60)投影到 P 子空间上, 那么等式(5.60)左边就变成 $P e^{-iHt/\hbar} P$, 而等式右边就是要算 $\frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega P \frac{e^{-i\omega t}}{\hbar\omega+i\epsilon-H} P$, 利用算符恒等式(5.37), 这个式子的最终结果就是将哈密顿量 H 替换成有效哈密顿量 $H_{\text{eff}}(\hbar\omega+i\epsilon)$, 从而也就是计算积分

$$\begin{aligned} & \frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega - H_{\text{eff}}(\hbar\omega+i\epsilon) + i\epsilon} \\ = & \frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega - H_{PP} - \Sigma(\hbar\omega) + i\epsilon} \\ = & \frac{i}{2\pi} \int_{-\infty}^{+\infty} \hbar d\omega \frac{e^{-i\omega t}}{\hbar\omega - (H_{PP} + \Delta(\hbar\omega)) + \hbar\frac{i}{2}\Gamma(\hbar\omega) + i\epsilon}. \end{aligned} \quad (5.61)$$

对于这个积分, 我们可以采用和上面一样的围道来处理。但是现在, 由于自能 $\Sigma(\hbar\omega)$ 有了一个负虚部 $-\hbar\frac{i}{2}\Gamma(\hbar\omega)$, 情况就有了根本性的不同。为了看清楚这个不同, 我们不妨假设 $\Gamma(\hbar\omega)$ 随着探测能量 $\hbar\omega$ 的变化足够平缓, 以致于近似可以看成与 ω 无关, 记作 Γ 。这时候对 ω 的围道积分所包围的极点将有一个有限大的负虚部 $-\frac{i}{2}\Gamma$, 根据留数定理, 这个负虚部最终将在积分(5.61)中贡献出如下因子

$$e^{-\frac{1}{2}\Gamma t}. \quad (5.62)$$

很显然, 这个因子就意味着, P 子空间里的这些态将随着时间指数衰减, 典型的衰减时间(称之为态的寿命) τ 为

$$\tau = \frac{1}{|\Gamma|}. \quad (5.63)$$

式中 $|\Gamma|$ 表示算符 Γ 的某个本征值的大小。

以上就清楚地说明了, 与连续谱的耦合的确将导致离散态发生衰变。在上面结果的实际应用中, 系统 H 往往可以分解成微扰之前的部分 H_0 和微扰部分 V 之和。这时候我们可以选 P 由 H_0 的离散本征态张成, 而 P^\perp 由 H_0 的连续谱本征态张成, 因此 $H_{PP^\perp} = PH_0P^\perp + PV P^\perp = H_0PP^\perp + PV P^\perp = V_{PP^\perp}$, 即 $H_{PP^\perp} = V_{PP^\perp}$ 。因此, 正是微扰 V 产生了 H_0 的两种不同谱态之间

的耦合。同时，在这种情况下，为了保证(5.52)式依然成立，我们还常常假定 $V_{P^\perp P^\perp} = 0$ 。

5.2.4 习题

1. 假设某系统的哈密顿量 H 可以写成如下矩阵，

$$H = \begin{pmatrix} U & 0 & t & t \\ 0 & U & t & t \\ t & t & 0 & 0 \\ t & t & 0 & 0 \end{pmatrix}. \quad (5.64)$$

- (1). 请求出这个系统约化到右下角的 2×2 子空间里的有效哈密顿量 $H_{\text{eff}}(z)$ 。
- (2). 假设 $t \ll U$ ，请求出系统在0附近的两个本征值(计算到 t^2/U^2 阶为止)。(这个题目来自Physics StackExchange 上的提问。提示：不能直接代简并微扰论一节最终给出的二阶微扰公式，得回到有效哈密顿量上想办法。)

2. 一个质量为 m 的粒子在势场 $V(x, y) = \frac{1}{2}m\omega^2(x^2 + y^2)$ 中运动，其中 ω 是常数。假设给粒子加上一个小的微扰 $V' = \lambda xy$ 。请证明这个微扰会使得系统第一激发态的两个简并能态产生能级劈裂，请用二阶微扰论计算劈裂以后的能级分别为多少。

5.3 变分法在量子力学中的应用

5.3.1 变分法求基态能量

在量子力学中，求解一个系统的基态尤其是求解基态能量，有最为重要的作用。因为一个系统的基态和低激发态常常决定了系统的宏观行为。但是，能够精确求解的量子力学系统少之又少。因此人们发展了一些对基态能量和基态波函数的近似计算方法，变分法就是这些近似计算方法中最为重要的方法之一。

用变分法计算基态能量基于下面这个简单的定理。即对于任意量子态 $|\psi\rangle$ ，我们有

$$E[\psi] = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \geq E_0, \quad (5.65)$$

式中, H 是系统的哈密顿量(要求这个哈密顿量有下界, 否则将不存在基态), E_0 就是系统的基态能量, $E[\psi]$ 表示依赖于量子态 $|\psi\rangle$ 的一个能量泛函, 式中的等号当且仅当 $|\psi\rangle$ 是系统基态时才能取到。

这个定理的证明非常简单。假设我们记 $|n\rangle$ ($n \geq 0$)为 H 的本征态, 本征值为 E_n , $n = 0$ 对应的就是基态。则我们可以将 $|\psi\rangle$ 展开为 $|\psi\rangle = \sum_n \psi_n |n\rangle$, ψ_n 为展开系数。很显然 $E[\psi] = \frac{\sum_n E_n |\psi_n|^2}{\sum_m |\psi_m|^2}$, 注意到 $E_n \geq E_0$ 就有, $E[\psi] \geq E_0 \frac{\sum_n |\psi_n|^2}{\sum_m |\psi_m|^2} = E_0$, 定理因此成立。很明显, 等于号成立的充要条件是 $\psi_n = 0, n \neq 0$, 这时候 $|\psi\rangle$ 本身就是基态。

变分法的基本思想是, 合理地猜测一组试探波函数, 然后选取这组试探波函数中使得能量泛函 $E[\psi]$ 取值最低的一个, 将之作为对基态波函数的近似, 相应的能量泛函的值就作为对基态能量的近似。很显然, 根据刚才的定理, 这样算出来的基态能量一定是偏高的, 因此为了得到更好的近似结果, 人们通常是在一个逐步扩大的试探波函数集合中选取对基态的近似, 因此得到的基态能量近似值也将是逐步降低的, 当这种降低趋向于收敛的时候, 我们就可以认为已经得到对基态和基态能量的很好近似了。在实际操作中, 我们常常是让试探波函数依赖于一些参数 $\lambda_1, \lambda_2, \dots$, 用这个带参数的试探波函数计算能量泛函 $E[\psi]$, 算出来的当然就是一个依赖于这些参数的能量表达式 $E(\lambda_1, \lambda_2, \dots)$, 求这个表达式的最小值, 并得出最小值所对应的参数, 则与最小值参数相应的试探波函数就可以看成是近似的基态波函数, 能量泛函的最小值就可以看成是近似的基态能量。同样, 为了保证得到足够好的近似, 我们可以不断地增加参数, 然后接着求新参数所对应的能量泛函最小值, 直到这些最小值趋于收敛。

求出足够好的近似基态以后, 我们可以接着在与这个近似基态正交的子空间里按照同样的办法求第一激发态波函数以及能量的近似。因为很显然, 在与基态正交的子空间里有一个和上面的定理完全类似的定理, 只不过这时候不等号右边的下界是第一激发态的能量 E_1 。当然, 按照这样的办法人们还可以逐步地对各个激发态都进行近似求解。不过, 相比较来说, 近似求解基态是变分法用得最多的地方。

下面我们用变分法来近似求解氢原子的基态能级。我们假设氢原子核是不动的, 位于坐标原点, 因此氢原子的哈密顿量是

$$H = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}}, \quad (5.66)$$

式中 $\mathbf{x}_1, \mathbf{x}_2$ 是两个电子的位置矢量, $r_1 = |\mathbf{x}_1|, r_2 = |\mathbf{x}_2|, r_{12} = |\mathbf{x}_1 - \mathbf{x}_2|$ 。很显然氢原子的哈密顿量可以看成是在两个相互独立的 He^+ 离子的哈密顿量

的基础上加上一个电子与电子的相互作用项 $\frac{e^2}{4\pi\epsilon_0 r_{12}}$ 。 He^+ 离子是一个类氢系统，其基态波函数在归一化以后是 $\left(\frac{Z^3}{\pi a_0^3}\right)^{1/2} e^{-Zr/a_0}$ ，式中 Z 就是原子序数，对于 He^+ 离子当然 $Z = 2$ ， a_0 就是玻尔半径。因此，两个 He^+ 离子(共用同一个核)的基态波函数当然就是

$$\psi(r_1, r_2) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0}. \quad (5.67)$$

但是现在，由于电子与电子之间的库伦排斥相互作用，上面的 $\psi(r_1, r_2)$ 当然不可能是氦原子真正的基态波函数，但是我们可以不妨将它作为对氦原子基态的一个试探波函数。同时，考虑到由于第二个电子的屏蔽效应，每个电子不再能感受到一个裸的氦原子核的库伦场了，我们可以设想每个电子将感受到一个有效的原子序数 $Z \neq 2$ ，并通过变分法来确定这个有效的 Z 值。也就是说，我们不妨将公式(5.67)中的 Z 变成一个参数，从而得到一组试探波函数 $\psi(r_1, r_2, Z)$ 。计算相应的能量泛函

$$\begin{aligned} E(Z) &= \int d\tau_1 d\tau_2 \psi(r_1, r_2, Z) \left[-\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} \right] \psi(r_1, r_2, Z) \\ &+ \int d\tau_1 d\tau_2 \psi(r_1, r_2, Z) \left[\frac{e^2}{4\pi\epsilon_0 r_{12}} \right] \psi(r_1, r_2, Z) \\ &= -2\left(\frac{1}{2} m \alpha^2 c^2\right) [4Z - Z^2 - \frac{5}{8} Z]. \end{aligned} \quad (5.68)$$

式中 $d\tau$ 表示三维体积元， $\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \simeq 1/137$ 是精细结构常数，因此 $\frac{1}{2} m \alpha^2 c^2 = 13.6 eV$ 是里德堡能量。式中的计算只有 $\int d\tau_1 d\tau_2 \psi(r_1, r_2, Z) \left[\frac{e^2}{4\pi\epsilon_0 r_{12}} \right] \psi(r_1, r_2, Z)$ 稍微难一点点，不过也无需用到任何特殊函数之类的东西，算是一道高数习题吧。

求表达式 $E(Z)$ 的最小值得到 $Z = \frac{27}{16}$ ，代入得到 $E(Z)$ 的最小值约为 $-77.5 eV$ ，这就是我们用变分法求出来的氦原子基态能量。这个值比真实值 $-78.6 eV$ 略大，但是误差很小。考虑到氦原子基态并没有精确解，因此我们能如此迅速地得到这样的近似结果，这已经足以说明变分法的威力了。

5.3.2 定态作为能量泛函的极值状态

从上一节关于变分法的讨论中其实我们已经能够想见，能量本征态一定是能量泛函 $E[\psi]$ 的极值状态，基态更是整体的极小值状态。这里我们首先用数学推导来确认一下这一点。假设 $|\psi\rangle$ 态发生了一个无穷小变动，

$|\psi\rangle \rightarrow |\psi\rangle + |\delta\psi\rangle$, 式中 $|\delta\psi\rangle$ 表示一个正比于无穷小量的任意状态。则能量泛函的改变量为

$$\begin{aligned}\delta(E[\psi]) &= \frac{\langle\psi|H|\psi\rangle + \langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle + O((\delta\psi)^2)}{\langle\psi|\psi\rangle + \langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle + O((\delta\psi)^2)} - \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \\ &= \frac{\langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle - E[\psi](\langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle)}{\langle\psi|\psi\rangle} + O((\delta\psi)^2) \\ &= \frac{\langle\delta\psi|(H - E[\psi])|\psi\rangle + \langle\psi|(H - E[\psi])|\delta\psi\rangle}{\langle\psi|\psi\rangle} + O((\delta\psi)^2). \quad (5.69)\end{aligned}$$

式中 $O((\delta\psi)^2)$ 代表高阶无穷小量。能量泛函的极值条件等价于 $\delta(E[\psi])$ 表达式中的一阶小量等于0, 从上面的表达式容易看出这又等价于 $\langle\delta\psi|(H - E[\psi])|\psi\rangle + \langle\psi|(H - E[\psi])|\delta\psi\rangle = 0$ 对于任意的无穷小变动 $|\delta\psi\rangle$ 都成立。不妨令 $|\varphi\rangle = (H - E[\psi])|\psi\rangle$, 因此一阶小量等于0的充要条件就变成 $\langle\delta\psi|\varphi\rangle + \langle\varphi|\delta\psi\rangle = 0$ 对于任意的 $|\delta\psi\rangle$ 成立。取 $|\delta\psi\rangle = \epsilon|\varphi\rangle$, 其中 ϵ 为一个实无穷小量, 则这个一阶小量等于0的条件又意味着, $\epsilon\langle\varphi|\varphi\rangle = 0$, 而这又意味着 $\langle\varphi|\varphi\rangle = 0$, 从而 $|\varphi\rangle = 0$ 。也即是说, $|\psi\rangle$ 是能量泛函的极值状态的充要条件是下式成立

$$(H - E[\psi])|\psi\rangle = 0, \quad (5.70)$$

由于 $E[\psi]$ 是一个数, 所以这一方程其实就是哈密顿算符的本征方程。因此, 能量泛函取极值的充要条件是 $|\psi\rangle$ 是哈密顿算符的本征态, 这时候能量泛函的值 $E[\psi]$ 显然就等于相应的本征值。

上面的这个结论通常也称之为里兹定理。它当然是对变分法的一个确认和补充。有时候用拉格朗日乘子法来表述里兹定理也是很方便的, 即能量本征态是下面的能量泛函 $E[\psi]$ 在归一化约束 $\langle\psi|\psi\rangle = 1$ 下的极值状态,

$$E[\psi] = \langle\psi|H|\psi\rangle. \quad (5.71)$$

这样表述以后人们就可以用标准的约束下取极值的拉格朗日乘子法来处理这一极值问题。当然, 这样的处理和我们之前的处理完全等价。

有时候将能量泛函的表达式在坐标表象中用波函数的形式具体写出来也非常有用。我们以哈密顿量 $H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})$ 为例来进行说明。很显然这时候 $E[\psi] = \int d\tau \psi^*(\mathbf{x}) \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x}) \right] \psi(\mathbf{x}) / (\int d\tau |\psi(\mathbf{x})|^2)$ 。对这个式子进行分部积分就可以得到 $E[\psi] = \left\{ \int d\tau \left[\frac{\hbar^2}{2m} |\nabla\psi(\mathbf{x})|^2 + V(\mathbf{x}) |\psi(\mathbf{x})|^2 \right] - \frac{\hbar^2}{2m} \int_{S(\infty)} d\vec{S} \cdot \psi^*(\mathbf{x}) \nabla\psi(\mathbf{x}) \right\} / (\int d\tau |\psi(\mathbf{x})|^2)$ (式中 $S(\infty)$ 表示空间的边界)。对于

表达式中的边界项 $\frac{\hbar^2}{2m} \int_{S(\infty)} d\vec{S} \cdot \psi^*(\mathbf{x}) \nabla \psi(\mathbf{x})$, 假如我们考虑的是束缚态情形, 或者是周期性边界条件的情形(这相当于没有边界), 那这个边界项其实就等于0。在这种情况下我们就有

$$E[\psi] = \frac{\int d\tau \left[\frac{\hbar^2}{2m} |\nabla \psi(\mathbf{x})|^2 + V(\mathbf{x}) |\psi(\mathbf{x})|^2 \right]}{\int d\tau |\psi(\mathbf{x})|^2}. \quad (5.72)$$

当然, 人们也可以将这个能量泛函用拉格朗日乘子法来表述, 这时候根据里兹定理, 将能量泛函在归一化约束之下对波函数 $\psi(\mathbf{x})$ 进行变分求极值, 我们会反过来得到坐标表象下的定态薛定谔方程。实际上, 在某些比较复杂的坐标系中(比如球坐标), 人们常常可以先在这种坐标系下写出能量泛函(5.72)(取波函数归一化约束), 然后再用拉格朗日乘子法对波函数变分求出这种坐标下相应的定态薛定谔方程。相比于直接写这样的复杂坐标下的定态薛定谔方程, 这样处理的简捷之处在于, 能量泛函(5.72)中只涉及到一阶偏导, 而一阶偏导在复杂坐标系中的表达形式远比直接写定态薛定谔方程时涉及的拉普拉斯算符 ∇^2 这样的二阶偏导要好处理。

5.3.3 有关基态的几个定理

这一节, 我们利用变分原理讨论几个有关于系统基态的一般性定理。首先我们证明对于一大类量子力学系统, 定态波函数总能取成实函数。这些系统就是具有如下形式定态薛定谔方程的系统,

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x) = E\psi(x). \quad (5.73)$$

很显然对于任意满足上面方程的 $\psi(x)$, 其复共轭 $\psi^*(x)$ 也满足同样的方程, 所以对于这样的系统我们总是能重新将定态波函数选成 $\psi(x) + \psi^*(x)$ 或者 $-i(\psi(x) - \psi^*(x))$, 显然, 这样重新选择以后的定态波函数是实函数。特别的, 这一类系统的基态波函数总能取成实函数。

因此结合变分原理, 系统(5.73)的基态就是使得下面的能量泛函取最小值的态。

$$E[\phi] = \frac{\int dx \left[\frac{\hbar^2}{2m} \left(\frac{\partial \phi}{\partial x} \right)^2 + V\phi^2 \right]}{\int \phi^2 dx}, \quad (5.74)$$

其中 $\phi(x)$ 是一个任意的实波函数。

从这里我们能够得出一个定理，那就是系统(5.73)的基态波函数不能有节点，也就是说，除了在空间边界处之外，基态波函数必定恒大于0。因为假设基态波函数有像图(5.2)所示的节点的话，那就必定有一部分波函数的值是负的，根据能量泛函的表达式，这时候我们完全可以把波函数的负值部分沿着自变量的 x 轴对折为正，就像图(5.3)所示的那样，由于能量泛函 $E[\phi]$ 对波函数以及波函数的导数的依赖都是平方型的，所以很显然这样做对 $E[\phi]$ 的值不会有任何影响。但是，一个类似于图(5.3)那样的波函数不可能对应能量泛函的最小值，因为很显然图(5.3)中波函数的导数在节点处有突变，而这就意味着我们可以将波函数作一个微小的改变，从而使得它在节点附近变得光滑，就像图(5.4)所示的那样，很显然这种光滑化的小手续会明显地降低能量泛函中与导数相关的部分 $(\frac{\partial\phi}{\partial x})^2$ 在节点附近的值，并保持对能量泛函的所有其它贡献不变，因此通过这样的光滑化小手续我们就可以进一步降低能量泛函的值。这也即是说，图(5.3)那样的波函数不可能对应能量泛函的最小值，根据变分原理，从而它也就不可能是系统的基态，当然这也意味着，任何像图(5.2)那样的波函数也不可能是系统的基态。因此，系统(5.73)的基态波函数必定不能有节点。

上一段的定理有一个重要的推论。那就是系统(5.73)的基态不可能简并。因为否则的话，假设 $\phi_0(x)$ 和 $\phi'_0(x)$ 都是系统(5.73)的基态，则根据线性叠加原理 $a\phi_0(x) + b\phi'_0(x)$ 也将是系统(5.73)的基态，这里 a, b 是两个任意的非0实数。很显然，由于 $\phi_0(x)$ 和 $\phi'_0(x)$ 是两个不同的态，因此我们只要合适地选取叠加系数 a, b 就能让 $a\phi_0(x) + b\phi'_0(x)$ 出现节点。但根据上一段的定理，这将与 $a\phi_0(x) + b\phi'_0(x)$ 是系统(5.73)的基态相矛盾。从而这就证明了系统(5.73)的基态不可能简并。

另外，假设 $\phi_n(x)$ 是系统(5.73)的某个取成实函数的激发态波函数。则由于激发态和基态 $\phi_0(x)$ 必定要正交，从而我们可以知道，激发态波函数 $\phi_n(x)$ 必定存在节点。因为否则的话 $\int dx \phi_n(x)\phi_0(x)$ 的被积函数将恒大于0或者恒小于0，那积分结果将不可能为0，从而与正交性相矛盾。

此外还有一个推论也值得一提，那就是假设我们所考虑的系统(5.73)是宇称守恒的，即假设在空间反射 $x \rightarrow -x$ 之下有 $V(-x) = V(x)$ ，那系统的基态波函数 $\phi_0(x)$ 将必定是偶宇称的，即必定满足 $\phi_0(-x) = \phi_0(x)$ 。这是因为，如果 $V(-x) = V(x)$ ，那 $\phi_0(-x)$ 也必定满足基态的本征方程，但由于基态不简并，所以必有 $\phi_0(-x) = \alpha\phi_0(x)$ 。由于空间反射两次等于不做任何操作，所以 $\alpha^2 = 1$ ， $\alpha = \pm 1$ 。但是由于 $\phi_0(x)$ 没有节点，所以它在空间反射之下也不能改变符号，因此只可能是 $\alpha = 1$ ，也即是说， $\phi_0(-x) = \phi_0(x)$ 。

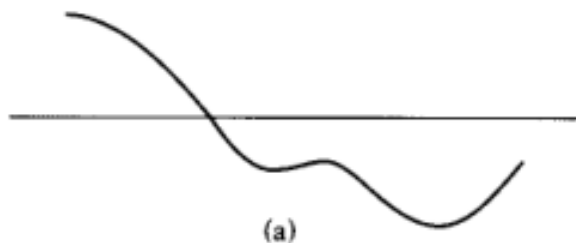


图 5.2: 假设波函数有节点. 图片来源: 费曼统计力学

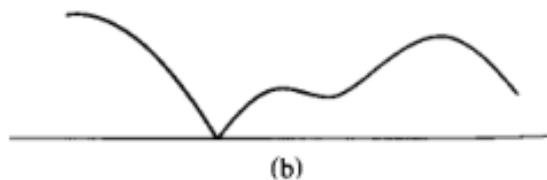


图 5.3: 把波函数的负部分反射成正. 图片来源: 费曼统计力学

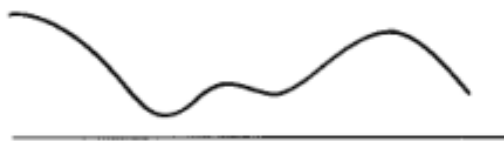


图 5.4: 把反射后的波函数作微小的光滑化. 图片来源: 费曼统计力学

以上所有的结论我们都是以单自由度系统为例来得到的, 但显然这些结论全都可以推广到类似下面那样的多自由度多粒子体系。

$$\left[-\sum_{i=1}^N \frac{\hbar^2}{2m_i} \nabla_i^2 + \sum_{i<j} V(r_{ij}) \right] \psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = E\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N), \quad (5.75)$$

这里我们假设的是一个 N 粒子系统, \mathbf{x}_i 是其中第 i 个粒子的位置矢量, $V(r_{ij})$ 是两体相互作用势能, 其中 $r_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ 是两个粒子的空间距离。

另外, 如果我们考虑的是 N 个完全相同粒子的全同多粒子体系, 如果这时候系统的哈密顿量依然具有(5.75)这样的形式的话(当然这时候所有全同粒子的质量都是相同的, 即 $m_i = m$), 我们还可以得出这种哈密顿量的

基态波函数必定是一个全对称波函数，即在不同粒子的位置作任意置换之下，波函数都保持不变。这是因为对于全同粒子，其哈密顿量在不同粒子位置矢量的任意置换 σ 下都必定是不变的。因此假设 $\psi_0(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ 是哈密顿量(5.75)的基态波函数，则 $\psi_0(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \dots, \mathbf{x}_{\sigma(N)})$ (式中 $\sigma(i)$ 表示 i 在置换 σ 作用之后的结果)必然也是哈密顿量(5.75)的一个基态波函数。但是，由于哈密顿量(5.75)的基态必定不简并，所以必有 $\psi_0(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \dots, \mathbf{x}_{\sigma(N)}) = \alpha_\sigma \psi_0(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ ，本来根据全同性原理⁴这样的 α_σ 可能等于 ± 1 ，但是由于基态波函数不能有节点，因此在坐标置换前后不能改变符号，从而必有 $\alpha_\sigma = 1$ (这也告诉我们，反对称的空间波函数必定是有节点的)。也即是说，对于任意置换 σ ，哈密顿量(5.75)的基态波函数必定满足

$$\psi_0(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \dots, \mathbf{x}_{\sigma(N)}) = \psi_0(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N), \quad (5.76)$$

从而必定是一个全对称波函数。

强调一下以上这些定理的适用范围：1. 它们只适用于空间波函数，因此不能包括自旋自由度。比方说，考虑到自旋自由度，一个理想的非相对论性氢原子的基态波函数也是二重简并的。2. 一般来说不能有磁场，因为否则哈密顿量(5.75)中出现的空间梯度 ∇ 就应该替换成 $\nabla - \frac{iq}{\hbar} \vec{A}(\mathbf{x})$ (式中 $\vec{A}(\mathbf{x})$ 是矢量势， q 是粒子的电荷)，因此相应的定态波函数就无法取成实函数。3. 不适用于粒子数目超过一定限度的全同费米子体系。比方说对于三个 $1/2$ 自旋的全同费米子体系，由于它们的自旋波函数无法完全反对称化，从而为了满足全同性原理，这种体系的空间波函数也不能做到全对称化，即使对于基态波函数也一样，因此这种体系当然是不满足以上定理的。一般的，对于多费米子体系，即使有时候可以完全将自旋自由度和空间自由度分离开来，但由于半整数自旋的自旋波函数一般来说不能做到完全反对称化(当粒子数目超过一定数量时)，因此它的空间波函数也不能是全对称的，从而这就违反了上面证明的定理。或者你也可以说，对于这样的体

⁴根据全同性原理，所有的电子都完全不可区分，因而是全同的。类似的，所有的光子也是全同的，所有的质子同样是全同的。全同粒子可以分成两类，像光子这样的自旋为整数的粒子是全同玻色子，像电子这样的自旋为半整数的粒子是全同费米子。对于全同玻色子体系，全同性原理要求系统的总波函数(同时包括空间变量和自旋变量的波函数)在这些全同粒子的任意置换下保持不变，人们称这样的波函数为全对称波函数。对于全同费米子体系，全同性原理要求系统的总波函数在这些全同费米子的奇置换下出负号，在偶置换下保持不变，人们称这样的波函数为全反对称波函数。有时候人们可以用全反对称的空间波函数乘上全对称的自旋波函数，或者用全对称的空间波函数乘上全反对称的自旋波函数来构造全同费米子的总波函数。

系，人们的确能求出哈密顿量(5.75)的一个全对称能量最低态波函数，但是由于这一波函数不能满足全同性原理，从而实际上是非物理的。4. 以上定理只适用于有限自由度的系统，在 $N \rightarrow +\infty$ 时可能不成立。5. 最后，严格来说，由于我们使用的能量泛函的表达形式没有包含边界项，因此这些定理只适用于束缚态，或者通过周期性边界条件将空间边界去除的情形。

两个例子

氢分子的基态波函数能很清楚地说明上面的这些定理。假设我们分别称氢分子的两个质子为 a 和 b ，由于质子比电子重很多，我们不妨假设这两个质子是静止的，两者之间的距离为 R 。两个电子，电子1和2，绕着这两个质子运动，这就构成了一个氢分子系统。很明显，这个氢分子系统的哈密顿量就属于(5.75)这样的类型，现在我们来分析这个系统的基态。由于质子对电子的库伦吸引力(质子与电子间的库伦势能为负)，为了保持最低的能量，电子1和2应该尽量靠近质子 a 和 b 。但是，如果两个电子同时靠近一个质子的话，那这两个电子之间就离得太近了，这时候由于电子与电子之间的库伦排斥力(电子与电子间的库伦势能为正)，系统的能量反而会比较。所以为了尽量降低系统的能量，最后的情况一定是，每一个电子分别围绕一个质子。但这样一来，系统好像就变成两个独立的氢原子放在一起了，问题就在于，为什么这两个氢原子能进一步降低能量，进而结合成氢分子呢？

原因在于，两个电子在两个质子上的配置有两个可能性：电子1围绕质子 a 电子2围绕质子 b ，以及电子1围绕质子 b 电子2围绕质子 a (如图(5.5)所示)，相应的波函数我们分别记为 $|A\rangle = \psi_{ab}(\mathbf{x}_1, \mathbf{x}_2)$ 和 $|B\rangle = \psi_{ab}(\mathbf{x}_2, \mathbf{x}_1)$ ，式中 \mathbf{x}_1 和 \mathbf{x}_2 分别是电子1和电子2的位置矢量。很显然，由于两个质子之间以



图 5.5: 氢分子两种可能性

及两个电子之间都没有什么区别，所以这两种可能性的地位完全平等，它们对应的系统能量可以记为 E_0 ，当两个质子之间的距离足够大时， E_0 就等于两个独立的氢原子的能量之和。但是， $|A\rangle$ 和 $|B\rangle$ 都是一种过于简化的描述，它们都不是氢分子系统的基态， E_0 也不是基态的能量，不过系统的

基态可以用一个以 $|A\rangle$ 和 $|B\rangle$ 为基的有效哈密顿量来描述。在这个有效描述中， $H_{AA} = H_{BB} = E_0$ 可以看作有效哈密顿量的对角项，但是，这个有效哈密顿量还有两个非0的非对角项 H_{AB} 和 H_{BA} 。这是由于，每个电子除了要受到其围绕的质子的库伦吸引之外，它还要受到另一个距离较远的质子的库伦吸引，因此尽管电子与电子之间的库伦排斥力造成了一个不小的势垒， $|A\rangle$ 和 $|B\rangle$ 这两种可能性之间依然有一定的概率幅通过量子隧穿效应相互转变。和任何典型的隧穿效应一样，这种隧穿效应也会产生一个负的非对角有效哈密顿量，这就是 H_{AB} 和 H_{BA} ，我们不妨记 $H_{AB} = H_{BA} = -J (J > 0)$ 。因此氢分子系统真正的基态可以通过求解下面的有效哈密顿量来得到，

$$\begin{pmatrix} E_0 & -J \\ -J & E_0 \end{pmatrix}, \quad (5.77)$$

这个哈密顿量的两个能量本征值分别为 $E_g = E_0 - J, E_e = E_0 + J$ ，相应的两个本征态分别为 $|g\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$ ， $|e\rangle = \frac{1}{\sqrt{2}}(|A\rangle - |B\rangle)$ 。或者也可以写成

$$|g\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1, \mathbf{x}_2) + \psi_{ab}(\mathbf{x}_2, \mathbf{x}_1)], |e\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1, \mathbf{x}_2) - \psi_{ab}(\mathbf{x}_2, \mathbf{x}_1)] \quad (5.78)$$

上面的分析告诉我们两个结论：第一，两个氢原子可以通过交换电子(即上面描述的量子隧穿)进一步将能量降低 $-J$ ，从而结合成一个氢分子。第二，氢分子的基态波函数 $|g\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1, \mathbf{x}_2) + \psi_{ab}(\mathbf{x}_2, \mathbf{x}_1)]$ ，很显然是一个对称波函数。正是在这里，这个结论和我们前面证明的一般性定理完全吻合。并且从这个基态波函数很容易看出，基态上的这两个电子实际上都是被两个质子共享的。

但是，由于电子是费米子，所以我们还得验证刚才得到的基态空间波函数 $|g\rangle$ 是否能满足全同性原理。答案是可以，因为现在只涉及到两个电子，所以对于对称的空间波函数，我们完全可以通过将自旋波函数构造成反对称的自旋单态来满足全同性原理。因此上面的分析其实是告诉我们，氢分子基态上的两个电子其实处于自旋单态，并且由于它们被两个质子共享，所以就在两个质子间形成了一个典型的共价键(如图(5.6)所示)，前面的 $-J$ 就是这个共价键的键能。而对于另外一个能量为 $E_0 + J$ 的态 $|e\rangle$ ，很显然它的空间波函数是反对称的，因此为了满足全同性原理，相应的电子自旋波函数就只能取自旋平行的对称三重态。因此，如果从电子自旋的视角来看待之前的有效哈密顿量，那它告诉我们的显然是，两电子自旋反平行(自旋单态)时能量较低，而自旋平行时能量较高。换言之，这个有效哈密顿量描述的是两电子自旋的反铁磁相互作用。

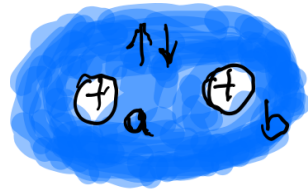


图 5.6: 氢分子

与氢分子这个例子略微有些相关的，我们举一个前面的那些定理不适用的例子，那就是铁磁体的基态。对于铁磁体而言，当它处于基态时，任意两个相邻格点上的电子自旋都是平行的(为简单起见我们假设每一个格点上只有一个电子围绕着格点上的这个原子运动)，但这就是铁磁体难以理解的地方。因为根据我们上面对氢分子的分析，通常而言相邻两个电子自旋反平行才会使得系统能量比较低，为什么铁磁体的自旋反而会平行呢?有一种物理上的解释是这样的：注意到铁磁体都是导体，所以除了围绕各自原子运动的电子之外，整个铁磁体中一定还存在传导电子，由于传导电子在整个铁磁体中运动，所以应该看成是被所有格点共享的。现在，由于泡利不相容原理，每个格点上的电子自旋会趋向于与传导电子的自旋反平行，但传导电子被所有格点共享了，因此这就意味着所有格点上的电子要自旋平行(如图(5.7)所示)，这就导致了铁磁体的出现。这种自旋平行化的出现显然意味着相邻格点上的电子的空间波函数不可能是对称的，这正好是不满足我们前面证明的那些定理的一个例子。但这完全是因为我们现在处理的体系涉及到的电子(费米子)数目太多(多于三个电子)，超出了定理的适用范围，属于我们前面已经注明过了的定理不适用的情形。

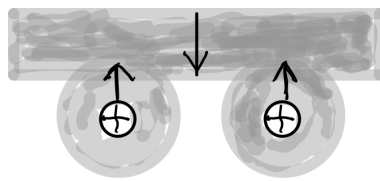


图 5.7: 铁磁体通过共享的传导电子而自旋平行

5.3.4 相变与对称性自发破缺

这一节我们将在变分法的近似下发展一个关于量子相变的物理图像。这一节的处理方式和例子基本上来自于文小刚老师关于拓扑序的一门在线课程的PPT。

通常来说，一个量子系统总会依赖于一些外界参数，不妨象征性地将这些参数记为 g ，比方说它可能是某个与系统耦合的外场等等。因此，这个量子系统的哈密顿量就将是 g 的光滑函数，不妨记为 H_g 。之所以强调是光滑函数是因为任何物理上合理的哈密顿量都会有比较好的数学性质，因此它们对物理参数的依赖都应该是光滑的。

现在假设我们用一组合理的试探波函数 $|\Psi_\phi\rangle$ 来近似求解 H_g 的基态能量，这里 ϕ 表示试探波函数里的未定参数。假设 $|\Psi_\phi\rangle$ 已经归一化好了，那么我们将得到一组能量泛函 $E_g(\phi)$ ，

$$E_g(\phi) = \langle \Psi_\phi | H_g | \Psi_\phi \rangle. \quad (5.79)$$

作为合理的试探波函数，它们的数学性质当然也应该比较好，因此 $|\Psi_\phi\rangle$ 对参数 ϕ 的依赖也将是光滑的。因此， $E_g(\phi)$ 将同时是系统参数 g 和试探参数 ϕ 的光滑函数。

但是， $E_g(\phi)$ 还不是系统的基态能量。按照变分法，为了得到系统基态能量的近似值，我们还需要求 $E_g(\phi)$ 关于 ϕ 的最小值，假设这个最小值在 $\phi = \phi_m(g)$ 时取到，那基态能量的近似值 $E(g)$ 实际上是

$$E(g) = E_g(\phi_m(g)). \quad (5.80)$$

注意到 $\phi_m(g)$ 一般来说也依赖于系统参数 g （在物理上这就相当于在不同物理条件 g 之下系统的基态会不一样），所以最后我们得到的基态能量近似值 $E(g)$ 作为 g 的函数就可能完全不同于原来的 $E_g(\phi)$ 。特别的，经过变分法的最终手续以后会产生这样一种可能性，即 $E(g)$ 作为 g 的函数不再是光滑的了，比方说它的某阶导数可能在某个参数点 g_c 变得不连续。在物理上，这就相当于说，随着外界物理参数的变化，系统的基态能量在 g_c 处出现了某种奇异性（即使系统的哈密顿量 H_g 关于 g 光滑），物理学家常常称这种奇异性的出现为系统发生了相变。

为了帮助读者看清这种可能性是如何发生的，我们不妨举一个简单的例子。假设有某个量子系统，我们用一组合理的试探波函数 $|\Psi_\phi\rangle$ 计算得出能量泛函为 $E_g(\phi) = (\phi^2 + g)^2$ 。很显然这个函数对于系统参数 g 和试探参

数 ϕ 来说都光滑。现在让我们求 $E_g(\phi)$ 关于 ϕ 的最小值,很显然当 $g \geq 0$ 时, $\phi_m = 0$,因此这时候有 $E(g) = g^2$,而当 $g < 0$ 时, $\phi_m = \pm\sqrt{-g}$,很明显这时候 $E(g) = 0$ 。因此最终我们得到的基态能量函数 $E(g)$ 是关于 g 的一个分段光滑函数,很显然这个分段函数整体上是不光滑的,它的二阶导数在 $g = g_c = 0$ 处不连续。对于这样一个系统,我们就说它在 $g_c = 0$ 处发生了相变,由于是基态能量的二阶导数出现了不连续,所以这样的相变也被称为二级相变。

仔细观察上面这个例子出现奇异性(即相变)的基本机制。我们就会发现,关键点就在于随着参数 g 由大于0变为小于0, $E_g(\phi)$ 关于 ϕ 的极小结构出现了突变。具体来说就是,当 $g \geq 0$ 时, $E_g(\phi)$ 关于 ϕ 只有一个极小点,即 $\phi_m = 0$,但是当 $g < 0$ 时, $E_g(\phi)$ 关于 ϕ 的极小点劈裂成了两个,即 $\phi_m = +\sqrt{-g}$ 和 $\phi_m = -\sqrt{-g}$ 。正是这种极小结构的突变导致了二级相变。

不妨再举一个一级相变的例子,看看和二级相变有什么不同。比方说对于某个量子系统我们算出来 $E_g(\phi) = \phi^4/4 + g\phi^3/3 - \phi^2/2$,很显然这个函数既是 g 的光滑函数又是 ϕ 的光滑函数。但其实这个系统在 $g = g_c = 0$ 处也有相变。为了看清楚这一点,我们不妨考察 g 在0附近时的情况,这时候, $E_g(\phi)$ 的两个极小点近似为(保留到一阶小量) $\phi_{m1} \approx -\frac{g}{2} - 1$ 和 $\phi_{m2} \approx -\frac{g}{2} + 1$,相应的两个极小值约为(保留到一阶小量) $E_g(\phi_{m1}) \approx -\frac{1}{4} - \frac{1}{3}g$, $E_g(\phi_{m2}) \approx -\frac{1}{4} + \frac{1}{3}g$ 。显然,当 $g > 0$ 时, ϕ_{m1} 对应真正的最小值,因此这时 $E(g) \approx -\frac{1}{4} - \frac{1}{3}g$,反过来,当 $g < 0$ 时, ϕ_{m2} 对应真正的最小值,因此这时 $E(g) \approx -\frac{1}{4} + \frac{1}{3}g$ 。很显然,这样算出来的基态能量 $E(g)$ 也是一个分段光滑函数,它在 $g = 0$ 处不光滑,因为它的一阶导数在 $g = 0$ 处不连续。人们常常把这样的一阶导数不连续的相变称作一级相变。观察这种相变发生的原因,我们会发现是因为 $E_g(\phi)$ 关于 ϕ 有多个极小值点,而随着系统参数 g 的变化(由 $g > 0$ 变到 $g < 0$),整体的最小值点在不同的极小之间发生了切换(由 ϕ_{m1} 切换到了 ϕ_{m2})。这种极小切换就会导致一级相变。

但是,我们刚才给出来的两个例子都不是真正算出来的例子。也即是说,我们并没有给出实现这些例子的量子系统,而只是直接给出了两个能够说明相变现象的函数 $E_g(\phi)$ 。对真正量子系统的具体计算研究发现,对于有限的系统,比方说自由度数目有限的系统,只要试探波函数真正合理,最后得到的 $E(g)$ 实际上通常都是整体光滑的,通常不会有真正的量子相变。人们真正感兴趣的 $E(g)$ 的奇异性 and 量子相变其实都是发生在无限大的量子系统中。当然,对于无限大的量子系统,其基态能量 $E(g)$ 会正比于系统的

体积 V , 因此也是趋于无穷的, 所以这时候更合适的研究对象其实应该是能量密度, 因为即使系统体积趋于无穷, 其能量密度总还是有限的。也就是说, 我们要定义一个能量密度泛函 $\varepsilon_g(\phi)$,

$$\varepsilon_g(\phi) = E_g(\phi)/V, \quad (5.81)$$

然后用 $\varepsilon_g(\phi)$ 来进行我们刚才所做过的那些分析。相应的基态能量密度 $\varepsilon(g)$ 就是

$$\varepsilon(g) = \varepsilon_g(\phi_m(g)), \quad (5.82)$$

$\varepsilon(g)$ 关于参数 g 的奇异性就意味着相变。当然, 我们之前通过例子发现的相变机制依然可以照搬过来, 只要将能量换成能量密度就可以了。值得说明的是, 无限大系统并不是物理学家凭空构想出来的, 实际上, 任何一块宏观材料由于包含了巨量的原子分子, 因此都可以近似地看成是无限大系统, 要研究材料的宏观性质就必须研究无限大系统, 这其实就是物理学家为什么对无限大系统感兴趣的基本原因。

下面我们举一个真正量子系统的例子。假设有 N 个 $1/2$ 自旋等间距地排布在一个圆周上(因此第 $N+1$ 个自旋其实就是第1个自旋)。假设相邻两个自旋之间的相互作用为 $-\sigma_i^x \sigma_{i+1}^x$, 其中 $\sigma_i^x, \sigma_i^y, \sigma_i^z$ 是第 i 个自旋的三个泡利算符。假设整个系统处在一个 z 方向的磁场中, 因此这个磁场和每一个自旋的相互作用能可以写成 $-g\sigma_i^z$, 参数 g 正比于磁场强度, 不妨假设 $g > 0$ 。因此, 整个系统的哈密顿量 H_g 可以写成

$$H_g = - \sum_i (\sigma_i^x \sigma_{i+1}^x + g\sigma_i^z). \quad (5.83)$$

这样一个系统通常被称为横场伊辛模型, 它实际上是一个可以精确求解的模型, 不过我们这里是想用它来讨论变分法的应用, 而不是要讨论它的精确求解。注意到在绕着 z 轴的 180 度旋转之下, $\sigma_i^z \rightarrow \sigma_i^z, \sigma_i^x \rightarrow -\sigma_i^x, \sigma_i^y \rightarrow -\sigma_i^y$ 。很显然(5.83)式给出的哈密顿量在这个 180 度旋转之下是不变的, 因此这个操作是系统的一个对称性。实际上, 这个绕 z 轴 180 度旋转的对称性可以用么正变换 $U = \prod_i \exp(i\pi \frac{1}{2} \sigma_i^z)$ 来表示。

注意到 $g \rightarrow \infty$ 时系统的状态完全由 z 方向的磁场决定, 因此这时候的基态为 $\prod_i (|\uparrow_i\rangle)$, 而 $g = 0$ 时, 系统有两个简并的基态 $\prod_i (|\rightarrow_i\rangle)$ 和 $\prod_i (|\leftarrow_i\rangle)$ (这里 $|\rightarrow_i\rangle = \frac{1}{\sqrt{2}}(|\uparrow_i\rangle + |\downarrow_i\rangle)$, $|\leftarrow_i\rangle = \frac{1}{\sqrt{2}}(|\uparrow_i\rangle - |\downarrow_i\rangle)$ 为 σ_i^x 的两个本征态)。所以对于一般情形的基态, 我们可以取归一化的试探波函数 $|\Psi_\phi\rangle =$

$\prod_i |\psi_i\rangle$, $|\psi_i\rangle = \cos \frac{\phi}{2} |\uparrow_i\rangle + \sin \frac{\phi}{2} |\downarrow_i\rangle$ 。注意到 $\sigma^z |\uparrow\rangle = |\uparrow\rangle$, $\sigma^z |\downarrow\rangle = -|\downarrow\rangle$, $\sigma^x |\uparrow\rangle = |\downarrow\rangle$, $\sigma^x |\downarrow\rangle = |\uparrow\rangle$, 易有 $\langle \psi_i | \sigma_i^z | \psi_i \rangle = \cos^2 \frac{\phi}{2} - \sin^2 \frac{\phi}{2} = \cos \phi$, $\langle \psi_i | \sigma_i^x | \psi_i \rangle = 2 \sin \frac{\phi}{2} \cos \frac{\phi}{2} = \sin \phi$ 。进而可以得到能量密度泛函,

$$\begin{aligned} \varepsilon_g(\phi) &= \langle \Psi_\phi | H_g | \Psi_\phi \rangle / N \\ &= - \sum_i [\langle \psi_i | \sigma_i^x | \psi_i \rangle \langle \psi_{i+1} | \sigma_{i+1}^x | \psi_{i+1} \rangle + g \langle \psi_i | \sigma_i^z | \psi_i \rangle] / N \\ &= -(\sin^2 \phi + g \cos \phi). \end{aligned} \quad (5.84)$$

我们按照 g 从大到小的顺序画了几幅 $\varepsilon_g(\phi) = -(\sin^2 \phi + g \cos \phi)$ 作为 ϕ 的函数图像(如图(5.8)、(5.9)、(5.10)所示)。很显然, 当 $g \geq 2$ 时, $\varepsilon_g(\phi)$ 只

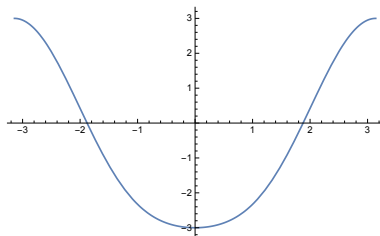


图 5.8: Fig of $-(\sin^2 \phi + 3 \cos \phi)$

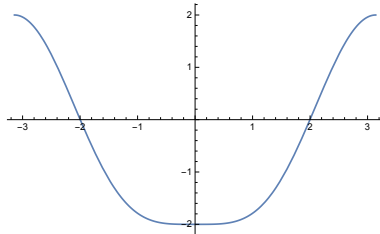


图 5.9: Fig of $-(\sin^2 \phi + 2 \cos \phi)$

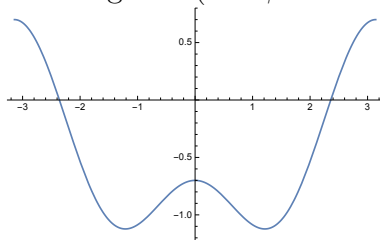


图 5.10: Fig of $-(\sin^2 \phi + 0.7 \cos \phi)$

有一个极小值点, 即 $\phi = 0$, 因此这时候基态能量密度 $\varepsilon(g) = -g$ 。但

是当 $g < 2$ 时, $\varepsilon_g(\phi)$ 的极小值点劈裂成了两个 $\phi = \phi_m = \arccos(\frac{g}{2})$, 以及 $\phi = -\phi_m = -\arccos(\frac{g}{2})$, 这时候 $\varepsilon(g) = -(1 + \frac{g^2}{4})$ 。这种极小劈裂的情况正符合我们前面关于二级相变机制的讨论。从基态能量密度的表达式也可以直接看出 $\varepsilon(g)$ 在 $g = 2$ 处的二阶导数不连续, 因此正是一个二级相变。

这个例子其实揭示了由朗道发现的关于连续相变(二级以上的相变)的一个一般性机制, 那就是, 连续相变所需要的能量泛函极小结构的突变常常都是由于对称性自发破缺带来的。在我们这个例子中, 在 $g > 2$ 的相, 变分法求出来的系统基态为 $|\Psi_0\rangle = \prod_i (|\uparrow_i\rangle)$, 这个态在绕着 z 轴的 180 度旋转下变换为 $|\Psi_0\rangle \rightarrow U|\Psi_0\rangle = \prod_i (i|\uparrow_i\rangle) = i^N |\Psi_0\rangle$ (非指标的那个 i 为虚数单位), 可见变换前后的基态只差一个相位, 在物理上它们描述的是同一个量子态, 因此这也就是说, 在 $g > 2$ 相中, 基态保持了绕 z 轴的 180 度旋转对称性。但是, 对于 $g < 2$ 的相, 变分法求出来的基态是简并的两个(这里严格来说要取 $N \rightarrow \infty$ 的无限大系统极限), 分别是 $|\Psi_{\phi_m}\rangle$ 和 $|\Psi_{-\phi_m}\rangle$, 而且人们很容易计算出这两个态在绕 z 轴的 180 度旋转之下是如何变换的, $|\Psi_{\phi_m}\rangle \rightarrow U|\Psi_{\phi_m}\rangle = i^N |\Psi_{-\phi_m}\rangle$, $|\Psi_{-\phi_m}\rangle \rightarrow U|\Psi_{-\phi_m}\rangle = i^N |\Psi_{\phi_m}\rangle$ 。可见, 这两个基态单独都不保持绕 z 轴的 180 度旋转对称性, 而是在这个旋转操作下互相变为对方。但是, 一个真实的系统只可能处在两个简并基态中的某一个, 但是不管真实系统是处在 $|\Psi_{\phi_m}\rangle$ 还是 $|\Psi_{-\phi_m}\rangle$, 它都破缺了这个绕 z 轴的 180 度旋转对称性。因此这就是一个典型的, 哈密顿量有一个对称性, 但是系统基态却破缺了这个对称性的例子, 这就叫做对称性自发破缺。我们这个例子正好演示了二级相变和对称性自发破缺的密切联系, 这种联系正是朗道首先指出的, 朗道的这个理论后来成为了理解凝聚态系统不同相和相变的基础, 因此成为凝聚态物理的一块基石, 并且还给粒子物理学家带来了极大的启发, 间接促成了基本粒子标准模型中的希格斯机制的提出。

上面例子中对称相和对称破缺相的不同还可以通过一个物理量直接显示出来, 那就是通过 σ_i^x 在基态上的期望值。利用 $\langle \Psi_\phi | \sigma_i^x | \Psi_\phi \rangle = \langle \psi_i | \sigma_i^x | \psi_i \rangle = \sin \phi$ 人们很容易得到, 在对称相, $\langle \Psi_0 | \sigma_i^x | \Psi_0 \rangle = 0$, 而在对称破缺相 $\langle \Psi_{\pm\phi_m} | \sigma_i^x | \Psi_{\pm\phi_m} \rangle = \pm \sin \phi_m = \pm \sqrt{1 - \frac{g^2}{4}} \neq 0$ 。在凝聚态物理中, 通常把这样的能够区分对称相和对称破缺相的物理量称之为序参量。

最后我们简单提一下, 横场伊辛模型的精确求解发现, 真正的相变点其实在 $g = 1$ 处, 而不是变分法发现的 $g = 2$ 处, 而且这实际上是一个量子临界点。很显然, 我们的变分法近似没有发现这一点。这主要是因为, 在我们用变分法求横场伊辛模型基态的时候, 我们是将所有的量子效应用一个整体的参数 ϕ 来平均了, 而并没有将各种量子涨落包括进来, 但是对

于 $g = 1$ 的量子临界点来说, 各个尺度上的量子涨落都很重要, 因此我们的变分法近似不能发现它也并不奇怪。这也清楚地说明了变分法毕竟只是一种近似方法, 它并不能解释所有的量子现象。

5.3.5 习题

1. 请用试探波函数 $\psi_\lambda = e^{-\lambda r}$ 估算氢原子基态的能量, 并和精确结果比较。

2. 假设在中心力场中运动的粒子受汤川势的相互作用, 势能为

$$V(r) = -A \frac{e^{-\mu r}}{r}, \quad (5.85)$$

式中 $A > 0$, 从而这是一个吸引势。请通过取试探波函数 $\psi_\lambda = e^{-\lambda r}$ 确定这系统存在束缚态的一个充分条件。

3. 两个离子可以通过共享一个电子进而降低能量形成共价键。最简单的例子是氢分子离子 H_2^+ , 它是两个质子共享一个电子。假设质子静止不动, 则系统的哈密顿量可以写成

$$H = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{4\pi\epsilon_0} \left[\frac{1}{|\mathbf{x}|} + \frac{1}{|\mathbf{x} - \mathbf{R}|} - \frac{1}{R} \right]. \quad (5.86)$$

式中 \mathbf{p} 是电子的动量, \mathbf{x} 是电子的位置矢量, \mathbf{R} 是两个质子之间的距离。假设取试探波函数

$$\Psi_{\mathbf{R}} = \psi(\mathbf{x}) + \psi(\mathbf{x} - \mathbf{R}), \quad \text{with} \quad \psi(\mathbf{x}) = \sqrt{\frac{1}{\pi a_0^3}} e^{-r/a_0}. \quad (5.87)$$

并注意到下面的积分

$$\begin{aligned} u(R) &= \int d^3\mathbf{x} \psi(\mathbf{x}) \psi(\mathbf{x} - \mathbf{R}) = \left(1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2}\right) e^{-R/a_0} \\ v(R) &= \int d^3\mathbf{x} \frac{\psi(\mathbf{x}) \psi(\mathbf{x} - \mathbf{R})}{|\mathbf{x}|} = \frac{1}{a_0} \left(1 + \frac{R}{a_0}\right) e^{-R/a_0} \\ w(R) &= \int d^3\mathbf{x} \frac{\psi^2(\mathbf{x})}{|\mathbf{x} - \mathbf{R}|} = \frac{1}{R} - \frac{1}{R} \left(1 + \frac{R}{a_0}\right) e^{-2R/a_0}. \end{aligned} \quad (5.88)$$

请利用这些结果证明, 哈密顿量在试探波函数上的平均能量 $E(R)$ 与氢原子基态的能量 E_0 之间的差可以写成

$$E(R) - E_0 = \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{R} - \frac{v(R) + w(R)}{1 + u(R)} \right). \quad (5.89)$$

请画出 $E(R) - E_0$ 作为横坐标 R 的函数的数值曲线，并以此估算氢分子离子的键能。[这个题目来自David Tong的课程习题]