

第十章 电子自旋

§10.1 电子自旋及其描述

Zeeman效应

10.1.1 电子自旋的发现

在分光计的分辨率提高以后，发现碱金属原子的特征光谱是彼此靠得很近的两条线，例如钠(Na)黄线的二重分裂，还发现了原子能谱在外加磁场中发生分裂的塞曼(Zeeman)效应，尤其是反常塞曼效应。这些现象都暗示电子可能有内禀的磁矩。

更直接的实验是施特恩-格拉赫(Stern-Gerlach)实验。这个实验的目标是测量银原子的磁矩（实际上是磁矩在选定方向上的投影）。让银原子束流通过一个不均匀磁场

$$\vec{B} \approx B(z)\vec{e}_z,$$

那么银原子的磁矩 \vec{M} 在这个磁场中就有“势能”

$$U = -\vec{M} \cdot \vec{B} = -M_z B(z),$$

所以它将受到力

$$\vec{F} = -\nabla U = M_z \frac{\partial B(z)}{\partial z} \vec{e}_z,$$

这个力使它的飞行轨道发生偏转，偏转的大小显然和原子磁矩在磁场方向上的投影 M_z 成正比。测出这个偏转（再结合其他一些实验参数）就可以算出 M_z 。

经典理论的预言是 $-|\vec{M}| \leq M_z \leq |\vec{M}|$ ，连续变化，而实验的结果却是

$$M_z = \pm M_B, \quad M_B = \frac{e\hbar}{2m_e}. \quad (\text{Bohr 磁子})$$

这个结果不仅从经典理论的角度无法解释，而且从 Schrödinger 量子力学的角度也无法解释，因为基态银原子的轨道角动量是零，所以基态银原子的轨道磁矩也是零（参见第六章）。我们只能把实验测得的这个磁矩归因于电子本身的“内禀”性质。由于电子是带电的，我们还可以进一步推论说：电子还有“内禀”的角动量，也就是有自旋(spin)。这就引出了如下的假设。内禀性质

乌伦贝克-高斯密特(Uhlenbeck-Goudsmit)电子自旋假设 (1925): 电子有自旋角动量，其投影只能取两个值：

$$S_z = \pm \frac{\hbar}{2},$$

这又导致电子有自旋磁矩，其投影为

$$M_z = -\frac{e}{m_e} S_z = \mp \frac{e\hbar}{2m_e} = \mp M_B. \quad (\text{SI 制})$$

写成矢量关系，自旋角动量算符记为 \hat{S} ，自旋磁矩算符记为 \hat{M}_s ，则

$$\hat{M}_s = -\frac{e}{m_e} \hat{S}.$$

电磁场：角动量与磁矩的关系

10.1.2 电子自旋的描述 Pauli 矩阵

自旋有纯量子力学的起源，只能用矩阵描写。电子自旋的分量只有两个可能的测量值，所以算符 $\hat{S}_x, \hat{S}_y, \hat{S}_z$ 都是 2×2 矩阵。对应到角动量的一般理论，电子自旋属于 $j = 1/2, m = \pm 1/2$ 的情形。在 \hat{S}_z 对角的表象中，利用一般公式（见第九章）

$$\begin{aligned} \langle j, m+1 | \hat{J}_x | j, m \rangle &= \frac{1}{2} \sqrt{(j+m+1)(j-m)} \hbar, \\ \langle j, m-1 | \hat{J}_x | j, m \rangle &= \frac{1}{2} \sqrt{(j-m+1)(j+m)} \hbar, \\ \langle j, m+1 | \hat{J}_y | j, m \rangle &= \frac{-i}{2} \sqrt{(j+m+1)(j-m)} \hbar, \\ \langle j, m-1 | \hat{J}_y | j, m \rangle &= \frac{i}{2} \sqrt{(j-m+1)(j+m)} \hbar. \end{aligned}$$

可知 $\hat{S}_x, \hat{S}_y, \hat{S}_z$ 的矩阵是:

$$\hat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad ??$$

从 \hat{S}_x, \hat{S}_y 还可以构造 \hat{S}_z 的升级算符和降级算符

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{S}_- = \hat{S}_x - i\hat{S}_y = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

引入 **泡利(Pauli)矩阵**

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

则可记为

$$\hat{\vec{S}} = \frac{\hbar}{2} \vec{\sigma}.$$

Pauli 矩阵的应用极为广泛, 读者应该牢记。

10.1.3 Pauli 矩阵的主要性质

(1) Pauli 矩阵是 Hermitian 矩阵。

(2)

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z, \quad \sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x, \quad \sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y,$$

所以 Pauli 矩阵是彼此**反对易的**。上式也可以合在一起写为

$$\sigma_i \sigma_j = i\epsilon_{ijk} \sigma_k, \quad (i, j, k = 1, 2, 3 = x, y, z; i \neq j)$$

其中 ϵ_{ijk} 是三阶完全反对称单位张量。由此不难验证 $\hat{S}_x, \hat{S}_y, \hat{S}_z$ 确实满足角动量算符的普遍对易关系

$$[\hat{S}_i, \hat{S}_j] = i\hbar \epsilon_{ijk} \hat{S}_k.$$

(3)

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I,$$

其中 I 是 2×2 单位矩阵。所以 Pauli 矩阵又是么正矩阵。总起来说, Pauli 矩阵满足

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk} \sigma_k. \quad (i, j, k = 1, 2, 3)$$

具有这类关系的代数系称为 Clifford 代数。

从上面这些关系出发还可以导出 Pauli 矩阵的其它一些有用的关系, 例如

(4)

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma},$$

其中 \vec{a}, \vec{b} 是任何三维矢量(纯数或者算符)。请同学自己证明一下这个等式。

10.1.4 二分量波函数 矩阵算符

显然, \hat{S}_z 的对应于本征值 $\pm\hbar/2$ 的本征矢量分别是:

$$\text{对于 } S_z = +\frac{\hbar}{2}, \quad \alpha \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\text{对于 } S_z = -\frac{\hbar}{2}, \quad \beta \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

现在电子的波函数还应该同时描写它的自旋状态。由叠加原理可知应有,

$$\Psi(\vec{r}, t) = \Psi_1(\vec{r}, t)\alpha + \Psi_2(\vec{r}, t)\beta = \begin{pmatrix} \Psi_1(\vec{r}, t) \\ \Psi_2(\vec{r}, t) \end{pmatrix},$$

这称为电子的**二分量波函数**, 又称为**二分量旋量**。它的 Hermitian 共轭波函数是

$$\Psi^\dagger(\vec{r}, t) = (\Psi_1^*(\vec{r}, t), \Psi_2^*(\vec{r}, t)).$$

如果 $\Psi_1(\vec{r}, t)$ 和 $\Psi_2(\vec{r}, t)$ 不是彼此成正比的, 那么电子就处于轨道和自旋的耦合态, 或者用另一种术语说, 叫做轨道自由度和自旋自由度的**纠缠态**。

与此同时，现在的算符也成为 2×2 矩阵算符，其一般形式为：

$$\hat{G} = \begin{pmatrix} \hat{G}_{11}(\vec{r}, -i\hbar\nabla) & \hat{G}_{12}(\vec{r}, -i\hbar\nabla) \\ \hat{G}_{21}(\vec{r}, -i\hbar\nabla) & \hat{G}_{22}(\vec{r}, -i\hbar\nabla) \end{pmatrix}.$$

如果一个力学量实际上与自旋无关，例如电子的动量，那么它在自旋空间中就简单地是一个单位矩阵，也就是说，它的矩阵算符是

$$\hat{p} = \begin{pmatrix} -i\hbar\nabla & 0 \\ 0 & -i\hbar\nabla \end{pmatrix} = -i\hbar\nabla \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

而电子自旋与它的轨道运动是互相独立的运动，所以自旋角动量算符 $\hat{S}_x, \hat{S}_y, \hat{S}_z$ 是常数矩阵，不包含对坐标的任何运算（见前）。与二者都有关的一个常见例子是**自旋-轨道耦合**算符

$$\hat{L} \cdot \hat{S} \equiv \hat{L}_x \hat{S}_x + \hat{L}_y \hat{S}_y + \hat{L}_z \hat{S}_z,$$

其中 $\hat{L}_x, \hat{L}_y, \hat{L}_z$ 是与 $\vec{r}, -i\hbar\nabla$ 有关的算符（参见第九章），而 $\hat{S}_x, \hat{S}_y, \hat{S}_z$ 是常数矩阵，把它们代入就得到

$$\hat{L} \cdot \hat{S} = \frac{\hbar}{2} \begin{pmatrix} \hat{L}_z & \hat{L}_x - i\hat{L}_y \\ \hat{L}_x + i\hat{L}_y & -\hat{L}_z \end{pmatrix}.$$

对于这样的波函数和算符，原先给出的公式需要稍加修正。

(1) 波函数的归一化是：

$$\int \Psi^\dagger \Psi d^3\vec{r} = \int (|\Psi_1|^2 + |\Psi_2|^2) d^3\vec{r} = 1.$$

(2) 电子的空间几率密度是：

$$\rho(\vec{r}, t) = \Psi^\dagger \Psi = |\Psi_1|^2 + |\Psi_2|^2.$$

(3) 电子的两种自旋状态的几率是：

$$W\left(+\frac{\hbar}{2}\right) = \int |\Psi_1|^2 d^3\vec{r}, \quad W\left(-\frac{\hbar}{2}\right) = \int |\Psi_2|^2 d^3\vec{r}.$$

(4) 算符的平均值是：

$$\bar{G} = \int \Psi^\dagger \hat{G} \Psi d^3\vec{r},$$

其中的 $\Psi^\dagger \hat{G} \Psi$ 在一般情况下既包括对坐标函数的运算（例如微分和乘法）又包括矩阵运算。

如果电子处于自旋和轨道不耦合（或者说不纠缠）的状态，那么它的波函数就简单地是

$$\Psi(\vec{r}, t) = \Psi_0(\vec{r}, t) \begin{pmatrix} a \\ b \end{pmatrix},$$

其中 Ψ_0 是复函数，满足

$$\int |\Psi_0|^2 d^3\vec{r} = 1,$$

而 a, b 是复常数，满足

$$|a|^2 + |b|^2 = 1.$$

自然界的电子当然是带有自旋的，但是我们以前不考虑电子自旋（如 Schrödinger 的氢原子理论）也做了许多计算，在实质上，那等于是假设了电子是处在没有自旋-轨道耦合的状态，所以只有 $\Psi_0(\vec{r}, t)$ 需要考虑，而自旋自由度不带来可观察的影响。

§10.2 电子总角动量和自旋-轨道耦合

10.2.1 轨道角动量和自旋角动量的合成

实际的电子既有轨道角动量 $\hat{\mathbf{L}}$ 也有自旋角动量 $\hat{\mathbf{S}}$ ，这二者的矢量和称为它的总角动量 $\hat{\mathbf{J}}$ ：

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$$

对照一般的规则，现在 $j_1 = l = 0, 1, 2, \dots$, $j_2 = 1/2$ ，所以

$$j = l + \frac{1}{2} \text{ 或 } l - \frac{1}{2}, \text{ 但是对于 } l = 0 \text{ 只有 } j = \frac{1}{2}.$$

注意： j 一定是半整数。

那么总角动量的本征态如何构成？这就要用到 $j_1 = l$, $j_2 = 1/2$ 的 CG 系数 $C(j_1 = l, j_2 = (1/2), j; m_1 = m - m_2, m_2, m)$ ，它们列在下表中。这里有 4 个 CG 系数，即 $j = j_1 \pm (1/2)$, $m_2 = \pm 1/2$ ，其值被表为 $j_1 (= l)$ 和 $m (= m_j)$ 的函数。

	$m_2 = +1/2$	$m_2 = -1/2$
$j = j_1 + \frac{1}{2}$	$\sqrt{\frac{j_1 + m + (1/2)}{2j_1 + 1}}$	$\sqrt{\frac{j_1 - m + (1/2)}{2j_1 + 1}}$
$j = j_1 - \frac{1}{2}$	$-\sqrt{\frac{j_1 - m + (1/2)}{2j_1 + 1}}$	$\sqrt{\frac{j_1 + m + (1/2)}{2j_1 + 1}}$

所以电子的总角动量为 $j = l + (1/2)$ 的二分量波函数是

$$\phi_{l+(1/2), m_j} = \begin{pmatrix} \sqrt{\frac{l + m_j + (1/2)}{2l + 1}} Y_{l, m_j - (1/2)}(\theta, \varphi) \\ \sqrt{\frac{l - m_j + (1/2)}{2l + 1}} Y_{l, m_j + (1/2)}(\theta, \varphi) \end{pmatrix},$$

其中 $m_j = l + (1/2), l - (1/2), \dots, -l - (1/2)$ 是总角动量的投影量子数。它也可以用 j, m_j 表为

$$\phi_{jm_j} = \frac{1}{\sqrt{2j}} \begin{pmatrix} \sqrt{j + m_j} Y_{j-(1/2), m_j - (1/2)}(\theta, \varphi) \\ \sqrt{j - m_j} Y_{j-(1/2), m_j + (1/2)}(\theta, \varphi) \end{pmatrix}. \quad \left(j = l + \frac{1}{2} \right)$$

当 $l = 0$ 的时候，就只有

$$\phi_{1/2, +1/2} = \begin{pmatrix} Y_{00}(\theta, \varphi) \\ 0 \end{pmatrix} = \frac{1}{\sqrt{4\pi}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \phi_{1/2, -1/2} = \begin{pmatrix} 0 \\ Y_{00}(\theta, \varphi) \end{pmatrix} = \frac{1}{\sqrt{4\pi}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

类似地， $j = l - (1/2)$ 的二分量波函数是

$$\phi_{l-(1/2), m_j} = \begin{pmatrix} -\sqrt{\frac{l - m_j + (1/2)}{2l + 1}} Y_{l, m_j - (1/2)}(\theta, \varphi) \\ \sqrt{\frac{l + m_j + (1/2)}{2l + 1}} Y_{l, m_j + (1/2)}(\theta, \varphi) \end{pmatrix},$$

或者写为

$$\phi_{jm_j} = \frac{1}{\sqrt{2j+2}} \begin{pmatrix} -\sqrt{j - m_j + 1} Y_{j+(1/2), m_j - (1/2)}(\theta, \varphi) \\ \sqrt{j + m_j + 1} Y_{j+(1/2), m_j + (1/2)}(\theta, \varphi) \end{pmatrix}. \quad \left(j = l - \frac{1}{2} \right)$$

但是当 $l = 0$ 的时候这些波函数 $\equiv 0$ 。

10.2.2 电子的自旋-轨道耦合

由于电子带有自旋磁矩，所以它在电场中运动时会产生一种新的相互作用。这种相互作用的基本机理可以这样来理解：以氢原子为例，在与电子一起运动的参照系中来看，带电的原子核是在围绕电子运

动，因而要在电子处产生磁场，这个磁场就和电子的自旋磁矩发生了相互作用。这种相互作用称为电子的**自旋-轨道耦合**。它在多电子系统中的推广称为 LS 耦合或 Russell-Saunders 耦合。由于磁矩 \vec{M} 与磁场 \vec{B} 的相互作用能是 $-\vec{B} \cdot \vec{M}$ ，容易想见自旋-轨道耦合的大小应该和 $\vec{L} \cdot \vec{S}$ 成正比。

严格处理这个问题需要用电子的相对论性量子力学方程，即 Dirac 方程，不再详细介绍。对 Dirac 方程做非相对论 ($v \ll c$) 近似，给出了电子自旋-轨道耦合 Hamiltonian (算符) 的如下表达式：

$$\hat{H}_{LS} = \frac{1}{2m_e^2 c^2} (\nabla V \times \hat{\vec{p}}) \cdot \hat{\vec{S}},$$

其中 V 是电子的电势能， $\hat{\vec{p}} = -i\hbar \nabla$ 是电子的动量算符， $\hat{\vec{S}} = (\hbar/2)\vec{\sigma}$ 是电子的自旋算符。如果电子是处在中心力场中，那么 $V = V(r)$ ， $\nabla V = (dV/dr)(\vec{r}/r)$ ，代入上式就得到

$$\hat{H}_{LS} = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV}{dr} \hat{\vec{L}} \cdot \hat{\vec{S}},$$

通常记为

$$\hat{H}_{LS} = \xi(r) \hat{\vec{L}} \cdot \hat{\vec{S}}, \quad \xi(r) \equiv \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV}{dr}.$$

这一项应该加在原先的 (未考虑电子自旋的) Hamiltonian 中。

这里出现的问题是：由于 $\hat{\vec{L}}$, $\hat{\vec{S}}$ 分别地都不和 $\hat{\vec{L}} \cdot \hat{\vec{S}}$ 对易，所以电子的轨道角动量和自旋角动量都不再是守恒的了，那么应该如何表征电子的状态？可以证明：电子的总角动量 $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ 和 $\hat{\vec{L}} \cdot \hat{\vec{S}}$ 是对易的，所以它是守恒量。证明如下。

$$\begin{aligned} [\hat{L}_i, \hat{\vec{L}} \cdot \hat{\vec{S}}] &= \sum_j [\hat{L}_i, \hat{L}_j \hat{S}_j] = \sum_j [\hat{L}_i, \hat{L}_j] \hat{S}_j = i \sum_j \epsilon_{ijk} \hat{L}_k \hat{S}_j, \\ [\hat{S}_i, \hat{\vec{L}} \cdot \hat{\vec{S}}] &= \sum_j [\hat{S}_i, \hat{L}_j \hat{S}_j] = \sum_j \hat{L}_j [\hat{S}_i, \hat{S}_j] = i \sum_j \epsilon_{ijk} \hat{L}_j \hat{S}_k, \end{aligned}$$

所以

$$[\hat{L}_i + \hat{S}_i, \hat{\vec{L}} \cdot \hat{\vec{S}}] = i \sum_j \epsilon_{ijk} (\hat{L}_k \hat{S}_j + \hat{L}_j \hat{S}_k) = 0. \quad \blacksquare$$

实际上，电子的总角动量本征态同时也是 $\hat{\vec{L}} \cdot \hat{\vec{S}}$ 的本征态，这是因为总角动量本征态是 L^2, J^2, \hat{J}_z 的同时本征态，而

$$J^2 = (\hat{\vec{L}} + \hat{\vec{S}})^2 = L^2 + S^2 + 2 \hat{\vec{L}} \cdot \hat{\vec{S}},$$

所以

$$\hat{\vec{L}} \cdot \hat{\vec{S}} = \frac{J^2 - L^2 - S^2}{2},$$

因此

$$(\hat{\vec{L}} \cdot \hat{\vec{S}}) \phi_{lm_j} = \left(\frac{J^2 - L^2 - S^2}{2} \right) \phi_{lm_j} = \frac{j(j+1) - l(l+1) - (3/4)}{2} \hbar^2 \phi_{lm_j}.$$

更具体地说，当 $j = l + (1/2)$ 时，

$$(\hat{\vec{L}} \cdot \hat{\vec{S}}) \phi_{lm_j} = \frac{l}{2} \hbar^2 \phi_{lm_j}, \quad \left(j = l + \frac{1}{2} \right)$$

当 $j = l - (1/2)$ 时，

$$(\hat{\vec{L}} \cdot \hat{\vec{S}}) \phi_{lm_j} = -\frac{l+1}{2} \hbar^2 \phi_{lm_j}. \quad \left(j = l - \frac{1}{2} \right)$$

这些结果是很有用的。

10.3.1 碱金属原子的 Hamiltonian

碱金属原子有一个价电子，它在原子实（原子核加上内层满壳电子）的影响下运动，所受到的势场可以用屏蔽的 Coulomb 场来描写。比如一种可能的假设是

$$V(r) = -\frac{k_1 e^2}{r} \left(1 + (Z-1)e^{-\kappa r}\right),$$

其中 Z 是原子序数，它在 $r \rightarrow \infty$ 的时候 $\rightarrow -k_1 e^2 / r$ ，也就是中心电荷为 $+e$ 的 Coulomb 场，而在 $r \rightarrow 0$ 的时候 $\rightarrow -k_1 Z e^2 / r$ ，也就是中心电荷为 $+Ze$ 的 Coulomb 场。但是下面的分析并不涉及 $V(r)$ 的具体形式。所以，在不考虑电子自旋的情况下，碱金属原子的 Hamiltonian 是

$$\hat{H} = \frac{1}{2m_e} \hat{p}^2 + V(r).$$

与纯 Coulomb 场（氢原子）的情况不同的是，屏蔽 Coulomb 场的电子能级与 n, l ($l = 0, 1, \dots, n-1$) 都有关，记为 $E_{n,l}$ ，但是仍然和 m_l 无关，所以能级的简并度是 $2l+1$ 。一般地说， n 相同而 l 不同的能级之间的距离小于 n 不同的能级之间的距离。再把电子自旋考虑进来，应该再加上自旋-轨道耦合，最后的 Hamiltonian 算符是

$$\hat{H} = \frac{1}{2m_e} \hat{p}^2 + V(r) + \xi(r) \hat{L} \cdot \hat{S}. \quad \left(\xi(r) = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV(r)}{dr} \right)$$

现在系统的守恒量完备集是 $\{\hat{H}, L^2, J^2, \hat{J}_z\}$ ，所以可设

$$\psi(r, \theta, \varphi, s_z) = R(r) \phi_{lm_j}(\theta, \varphi, s_z),$$

代入能量本征方程

$$\left[\frac{\hbar^2}{2m_e} \left(-\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{L^2}{r^2} \right) + V(r) + \xi(r) \hat{L} \cdot \hat{S} \right] \psi = E \psi,$$

就可得出径向波函数 $R(r)$ 满足的方程。我们在上一节里已经计算过 $\hat{L} \cdot \hat{S}$ 的值，所以对于 $j = l + (1/2)$ 的情形，径向方程是

$$\left[\frac{\hbar^2}{2m_e} \left(-\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{l(l+1)}{r^2} \right) + V(r) + \frac{l\hbar^2}{2} \xi(r) \right] R = ER,$$

而对于 $j = l - (1/2)$ 的情形则是

$$\left[\frac{\hbar^2}{2m_e} \left(-\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{l(l+1)}{r^2} \right) + V(r) - \frac{(l+1)\hbar^2}{2} \xi(r) \right] R = ER.$$

从这里就可以解出能级 E （它仍然是这些方程的本征值）和径向波函数 $R(r)$ 。

10.3.2 碱金属原子的能级分裂和光谱的精细结构

从上面的方程可以看出，在考虑了电子自旋以后，碱金属原子的能级变得和 n, l, j 都有关了，能级只对 m_j 是简并的，简并度是 $2j+1$ 。这些能量本征态在光谱学上用 nL_j 来标记，例如（略去主量子数）： $S_{1/2}$, $P_{1/2}$, $P_{3/2}$, $D_{3/2}$, $D_{5/2}$, $F_{5/2}$, $F_{7/2}$, 等等。所以，原先不考虑电子自旋时候的能级 $E_{n,l}$ 现在分裂成了两个能级，分别对应于 $j = l + (1/2)$ 和 $j = l - (1/2)$ 。由于 $dV/dr > 0$ ，所以 $\xi(r) > 0$ ，这就使得 $j = l + (1/2)$ 的能级高于 $j = l - (1/2)$ 的能级：

$$E_{n,l,j=l+(1/2)} > E_{n,l,j=l-(1/2)},$$

也就是说，例如

$$3P_{3/2} > 3P_{1/2}.$$

分裂的间距随原子序数 Z 的增加而增加。以 Na 原子 ($Z=11$) 为例, 它的 $3P_{3/2}$ 和 $3P_{1/2}$ 的能级间距已经比较明显, 电子分别从它们向 $3S_{1/2}$ 能级跃迁就造成 Na 的特征光谱的 D 黄线分裂为 D_1 和 D_2 两条线, 波长分别为 $\lambda = 589.5930 \text{ nm}$ 和 $\lambda = 588.9963 \text{ nm}$, 二者的差别大约是平均波长的千分之一。这种分裂称为 Na 原子特征光谱的精细结构。实际上, 各种原子的特征光谱普遍都有精细结构。在精密光学技术十分发达的今天, 分辨这些精细结构已经不是什么困难的事情。

*10.3.3 氢原子光谱的精细结构, 超精细结构和 Lamb 移动

前面给出的自旋-轨道耦合 Hamiltonian 虽然能够定性地说出原子特征光谱精细结构的出现, 但是在定量的水平上是不够令人满意的 (给出的能级裂距偏小)。这其中的一个重要原因是自旋-轨道耦合在本质上相对论效应, 而前面给出的表达式只是非相对论近似, 它要求粒子的运动速度要远小于光速。如果电子的运动速度和光速相比不是很小, 那么“质量变大”这样的相对论效应也就应该考虑, 这时就必须用严格的相对论性量子力学来处理。在这样的构架下, “自旋-轨道耦合”和“质量变大”这两种相对论效应实际上是无法区分的。对于氢原子的情形, 当电子处在基态的时候, 电子的平均速度与光速之比是精细结构常数

$$\alpha = \frac{k_1 e^2}{\hbar c} \approx \frac{1}{137.037},$$

它不能说很小。所以对氢原子进行完全的相对论处理是必要的。我们在这里不可能详细介绍相对论性量子力学, 只是指出: 对于氢原子, 完全的相对论处理也是可以严格进行的。下面我们就简单介绍一下它的结果。首先我们回忆一下不考虑电子自旋和相对论效应时的氢原子能级, 它们是 (见第六章)

$$E_n = -\frac{mk_1^2 e^4}{2\hbar^2} \frac{1}{n^2}, \quad (n=1, 2, 3, \dots)$$

(这里我们暂时不考虑有限核质量的影响, 所以上式中的 m 就是电子的静止质量 m_e)。有趣的是, 它也可以写成下面的样子:

$$E_n = -mc^2 \frac{\alpha^2}{2n^2},$$

其中的 mc^2 是电子的静止能。用相对论性量子力学处理氢原子, 得出它的能级的精确表达式为

$$\varepsilon = \frac{mc^2}{\sqrt{1 + \frac{\alpha^2}{(\sqrt{\kappa^2 - \alpha^2} + n - |\kappa|)^2}}}, \quad (\kappa = \pm[j + (1/2)]).$$

现在电子的状态由 (n, j, m_j, κ) 这 4 个量子数来表征。 $j=1/2, 3/2, \dots$ 仍然是总角动量, m_j 是总角动量的 Z 轴投影, $\kappa = \pm[j + (1/2)]$ 代替了原先量子数 l 的位置, $\kappa > 0$ 相当于 $l = j - (1/2)$, $\kappa < 0$ 相当于 $l = j + (1/2)$ 。主量子数 n 现在是径向量子数 $n_r = (0, 1, 2, \dots)$ 与 $|\kappa|$ 的和, 即 $n = n_r + |\kappa|$, 而 $|\kappa|$ 的最小值是 1, 所以 n 的取值范围是 $n=1, 2, 3, \dots$ 。反过来说, 当 n 给定以后 $|\kappa|$ 的取值范围是 $|\kappa|=1, \dots, n$ 。从上式看, 给定 n 和 $|\kappa|$ 以后能量就完全定了, 而 $j = |\kappa| - (1/2)$, 所以也相当于 n 和 j 完全决定了能量, 能级对于 $m_j = j, \dots, -j$ 和 κ 的正或负是简并的, 简并度是 $2(2j+1)$, 但在 $j=1/2$ 时只有 $2j+1=2$ 。但是要注意, 这个 ε 是“总能量”, 即包括了电子的静止能、动能和在 Coulomb 场中的势能。为了和上面的结果进行对比, 我们要从它当中减去电子的静止能, 所以

$$\frac{E}{mc^2} = \frac{1}{\sqrt{1 + \frac{\alpha^2}{(\sqrt{\kappa^2 - \alpha^2} + n - |\kappa|)^2}}} - 1,$$

这个 E 显然 < 0 。这个结果虽然精确却比较复杂。由于 α^2 是一个小量，所以可以把这个表达式对 α^2 做展开。准确到 α^4 阶得到如下的近似式：

$$E_{nlj} = -mc^2 \frac{\alpha^2}{2n^2} \left\{ 1 + \frac{\alpha^2}{n^2} \left(\frac{n}{j + (1/2)} - \frac{3}{4} \right) \right\},$$

其中大括号里的第二项就代表了电子自旋和相对论效应带来的修正，它给出了氢原子光谱的**精细结构**。这个结果和实验观察完全一致。

事情还不止于此。在谱线分辨精度进一步提高以后，发现氢原子的能谱还有**超精细结构**和**Lamb 移动**。从现象上说，超精细结构是氢原子**基态能级**的分裂，它给出了氢原子特征光谱中的**21 cm** 线。超精细结构的物理起因是电子的自旋磁矩和质子（原子核）的自旋磁矩之间的相互作用。Lamb 移动是这样现象：根据上面的能级公式， E_{nlj} 实际上是和 l 无关的，这是在 Coulomb 势场的特殊情况下产生的简并，这样就应该有

$$2S_{1/2} = 2P_{1/2}.$$

但是实验发现 $2S_{1/2}$ 的能量**略高于** $2P_{1/2}$ 。这个裂距非常之小，所以直到射频技术取得重大进步（这得益于第二次世界大战中关于雷达的研究）以后，Lamb 和 Retherford 才在 1947 年用射频波谱学的方法准确地测出了二者的裂距所对应的频率为 1057.8 ± 0.1 MHz。理论的分析表明，Lamb 移动来自量子场论的高阶修正（所谓的“真空极化”和“辐射反冲”效应）。在实验的挑战下，量子理论又向前发展了自己（主要是解决了量子场论中的“发散困难”），并且再次获得了与实验观察完全一致的理论结果。

总而言之，氢原子光谱的实验观察对于量子理论（包括量子力学，相对论性量子力学，量子场论）的正确性是一个理想的试金石，因为在这里一方面实验可以测得很准，另一方面理论也可以算得很准。我们可以说，直到目前为止，量子理论在这个试金石面前的表现是非常令人满意的，也就是说，理论计算的结果与实验观察的结果是符合得非常好的。这使得我们相信，**量子理论对我们的自然界的描写是相当正确的和可靠的。**

§10.4 Zeeman 效应

10.4.1 有自旋的电子在电磁场中的 Hamiltonian

实验证明，在外磁场中，原子的能级会发生分裂，结果是原子的特征谱线也发生分裂，这称为**塞曼 (Zeeman) 效应**。理论的解释是：电子的磁矩和外磁场产生了附加的相互作用能。

先不考虑自旋和原子实电场的作用，电子($q = -e$)在外磁场中的 Hamiltonian 是（见第七章）

$$\hat{H} = \frac{1}{2m_e} (\hat{\vec{P}} + e \vec{A})^2,$$

在原子的范围里，外磁场可以认为是均匀的，所以上式中的矢量势是 $\vec{A} = \vec{B} \times \vec{r} / 2$ ，如果 $\vec{B} = B \vec{e}_z$ ，那么 $\vec{A} = (-By/2, Bx/2, 0)$ ，代入得

$$\hat{H} = \frac{1}{2m_e} \left[\left(\hat{P}_x - \frac{eB}{2} y \right)^2 + \left(\hat{P}_y + \frac{eB}{2} x \right)^2 + \hat{P}_z^2 \right] = \frac{1}{2m_e} \hat{P}^2 + \frac{Be}{2m_e} \hat{L}_z + \frac{B^2 e^2}{8m_e} (x^2 + y^2),$$

其中

$$\hat{P}^2 = \hat{P}_x^2 + \hat{P}_y^2 + \hat{P}_z^2.$$

再考虑电子带有自旋，如果没有外磁场，那么注意到有公式（见本章第 1 节）

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma},$$

而在 $\vec{a} = \vec{b}$ 的时候它化为 $(\vec{a} \cdot \vec{\sigma})^2 = a^2$ ，所以原来的 \hat{P}^2 可以修改为 $(\hat{\vec{P}} \cdot \vec{\sigma})^2$ 。加了外磁场以后，要再用 $\hat{\vec{P}} + e \vec{A}$ 代替 $\hat{\vec{P}}$ ，所以 Hamiltonian 最后修改为

$$\begin{aligned} \hat{H} &= \frac{1}{2m_e} \left((\hat{\vec{P}} + e \vec{A}) \cdot \vec{\sigma} \right)^2 \\ &= \frac{1}{2m_e} \left((\hat{\vec{P}} + e \vec{A})^2 + i e (\hat{\vec{P}} \times \vec{A} + \vec{A} \times \hat{\vec{P}}) \cdot \vec{\sigma} \right) = \frac{1}{2m_e} \left((\hat{\vec{P}} + e \vec{A})^2 + e \hbar (\nabla \times \vec{A}) \cdot \vec{\sigma} \right) \\ &= \frac{1}{2m_e} (\hat{\vec{P}} + e \vec{A})^2 + \frac{e \hbar}{2m_e} \vec{B} \cdot \vec{\sigma}. \end{aligned}$$

再注意到 $\vec{S} = (\hbar/2) \vec{\sigma}$ ， $\vec{B} = B \vec{e}_z$ 以及上面 $(\hat{\vec{P}} + e \vec{A})^2$ 展开的结果，它又可以写为

$$\hat{H} = \frac{1}{2m_e} \hat{P}^2 + \frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z) + \frac{B^2 e^2}{8m_e} (x^2 + y^2).$$

它与无自旋电子的 Hamiltonian 的不同之处只是用 $\hat{L}_z + 2\hat{S}_z$ 代替了 \hat{L}_z 。再把原子实电场的作用以及自旋-轨道耦合加上，最后得到

$$\hat{H} = \frac{1}{2m_e} \hat{P}^2 + V(r) + \frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z) + \xi(r) \hat{\vec{L}} \cdot \hat{\vec{S}} + \frac{B^2 e^2}{8m_e} (x^2 + y^2).$$

这就是有自旋的电子在电磁场中的完整的 Hamiltonian。

在这个 Hamiltonian 中， B 的线性项可以做这样的物理解释：第六章中证明了 $-e\vec{L}/2m_e = \vec{M}_L$ 是电子的轨道磁矩，本章第 1 节中又说明了 $-e\vec{S}/m_e = \vec{M}_S$ 是电子的自旋磁矩，所以实际上

$$\frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z) = -\vec{B} \cdot (\vec{M}_L + \vec{M}_S).$$

这正是磁矩与磁场的相互作用能量的经典表达式。此外，现在的 ψ 当然是二分量的波函数。

10.4.2 正常 Zeeman 效应

现在估计一下 \hat{H} 中增加的那些项的大小。设原子的尺度为 a （Bohr 半径的量级），那么

$$\left| \frac{e \vec{A}}{\vec{P}} \right| \square \frac{Ba^2}{(h/e)} \square \frac{Ba^2}{\Phi_0},$$

其中 $\Phi_0 = h/e \approx 4 \times 10^{-15} \text{ V s}$ 是磁通的量子单位。由于 $a \ll 10^{-10} \text{ m}$ ，而且通常来说 $B < 10 \text{ T}$ ，所以 $Ba^2/\Phi_0 < 10^{-4}$ 。这样一来在 \hat{H} 中的 B^2 项可以略去。再进一步，如果磁场比较强，那么自旋-轨道耦合相对于磁矩-磁场相互作用也可以先不考虑。在这样的两个近似之下，Hamiltonian 就化为

$$\hat{H} = \frac{1}{2m_e} \hat{P}^2 + V(r) + \frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z),$$

因而 Schrödinger 方程成为：

$$-\frac{\hbar^2}{2m_e} \nabla^2 \psi + V(r) \psi + \frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z) \psi = E \psi.$$

先考虑磁场 $B=0$ 的情形。这时候守恒量的完备集可以选为 $\{\hat{H}, L^2, \hat{L}_z, \hat{S}_z\}$ ，对应的能量本征态用 n, l, m_l, m_s 四个量子数来表征，记为 $\psi_{nlm_l m_s}$ ，能量本征值和 n, l 有关，所以我们有以下的 4 个同时本征方程：

$$\begin{aligned} \left(-\frac{\hbar^2}{2m_e} \nabla^2 + V(r) \right) \psi_{nlm_l m_s} &= E_{nl} \psi_{nlm_l m_s}, \\ L^2 \psi_{nlm_l m_s} &= l(l+1) \hbar^2 \psi_{nlm_l m_s}, \\ \hat{L}_z \psi_{nlm_l m_s} &= m_l \hbar \psi_{nlm_l m_s}, \\ \hat{S}_z \psi_{nlm_l m_s} &= m_s \hbar \psi_{nlm_l m_s}, \end{aligned}$$

其中 $l = 0, \dots, n-1$; $m_l = l, \dots, -l$; $m_s = 1/2, -1/2$ 。我们发现，即使把磁矩-磁场相互作用项加进来， $\psi_{nlm_l m_s}$ 仍然可以使上面的 Schrödinger 方程得到满足，只不过其中的能量本征值现在变成

$$E_{nlm_l m_s} = E_{nl} + \frac{Be\hbar}{2m_e} (m_l + 2m_s).$$

所以能级发生了分裂和移动。注意到 m_l 和 m_s 的取值范围，能级的改变是（以 $Be\hbar/2m_e$ 为单位）：

$$\Delta E_{nl} = l+1, \quad l, \quad l-1, \quad l-2, \quad \text{等等, 直到 } -(l+1),$$

单态, 单态, 双态, 双态, 单态。

但是谱线的分裂没有这么多。应用选择定则（见第十一章）

$$\Delta l = \pm 1, \quad \Delta m_l = 0, \pm 1, \quad \Delta m_s = 0,$$

原来频率为 ω_0 的一条谱线现在变成了 3 条，其频率分别为：

$$\omega_0, \quad \omega_0 + \frac{Be}{2m_e}, \quad \omega_0 - \frac{Be}{2m_e}.$$

以钠 Na 的 D 黄线为例，在不考虑自旋-轨道耦合的情况下，它对应着从 3P 到 3S 的跃迁。加了强磁场以后，3P 分裂为 5 个能级（其中有 1 个能级是 2 重简并的），3S 分裂为 2 个能级。再考虑到能级跃迁时的选择定则，从 3P 到 3S 可以发生 6 种不同的跃迁，然而只有 3 种不同的频率，因此一条谱线变成了 3 条谱线。

但是，在光谱技术已经能够很容易地分辨精细结构的情况下，完全忽略自旋-轨道耦合就过于粗糙了，所以我们应该随后把自旋-轨道耦合再加进来。利用微扰论（见后面的第十一章）可知，这时 3P 仍然分裂为 5 个能级（其中有 1 个能级 2 重简并），3S 仍然分裂为 2 个能级，然而裂距有了变动，所以 3P 到 3S 的 6 种跃迁具有了 5 种不同的频率。这使频率为 ω_0 的那条谱线没有分裂，而频率为 $\omega_0 \pm (Be/2m_e)$ 的那 2 条谱线再各分裂为 2 条谱线，因此 D 黄线分裂为 5 条谱线。

这就是强磁场下叠加了自旋-轨道耦合的 Zeeman 效应，也就是实验观察到的正常 Zeeman 效应，见图 10.4.1。

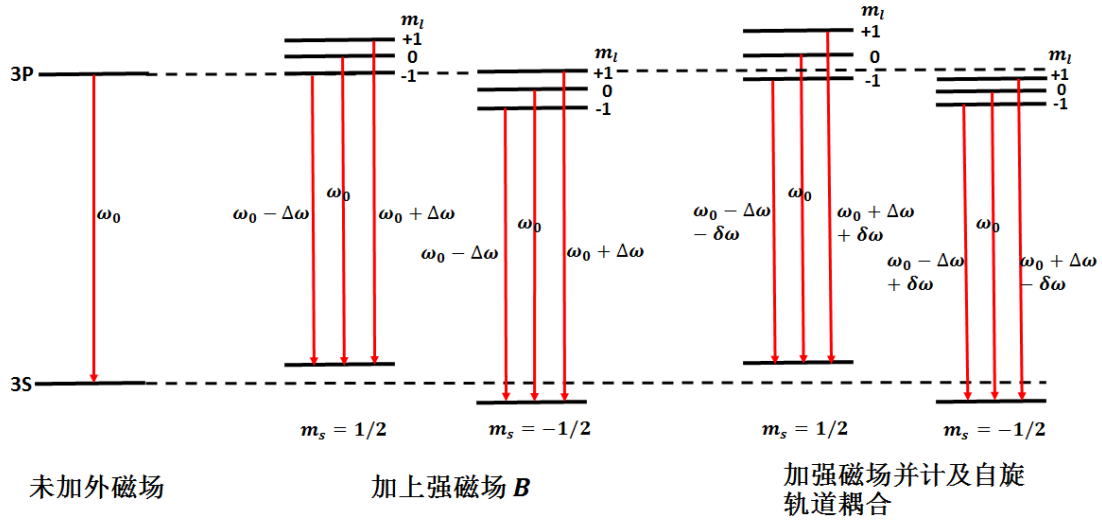


图 10.4.1 正常 Zeeman 效应的能级分裂和谱线分裂

*10.4.3 反常 Zeeman 效应

假如磁场不是很强，那么自旋-轨道耦合在一开始就不可忽略，所以 Hamiltonian 成为

$$\hat{H} = \frac{1}{2m_e} \hat{P}^2 + V(r) + \frac{Be}{2m_e} (\hat{L}_z + 2\hat{S}_z) + \xi(r) \hat{L} \cdot \hat{S}.$$

在这种情况下， $\{\hat{H}, L^2, \hat{L}_z, \hat{S}_z\}$ 和 $\{\hat{H}, L^2, J^2, \hat{J}_z\}$ 都不再是守恒量完备集，但注意到 $\hat{J}_z = \hat{L}_z + \hat{S}_z$ ，可以把 \hat{H} 重写为

$$\hat{H} = \frac{1}{2m_e} \hat{P}^2 + V(r) + \xi(r) \hat{L} \cdot \hat{S} + \frac{Be}{2m_e} \hat{J}_z + \frac{Be}{2m_e} \hat{S}_z,$$

那么 $\{\hat{H}, L^2, J^2, \hat{J}_z\}$ 对于除最后一项以外的部分构成守恒量完备集。所以我们可以这样来分析：当我们完全忽略磁场（也就是 \hat{H} 中的后两项都=0）的时候，它就是带有自旋-轨道耦合的 Hamiltonian，能谱就是带有精细结构的能谱，能级是 E_{nlj} ，简并度是 $2j+1$ (见本章的上一节)。加上 $(Be/2m_e) \hat{J}_z$ 这一项以后，能量本征态还是原来的，只不过能量变成了

$$E_{nljm_j} = E_{nlj} + \frac{Be\hbar}{2m_e} m_j,$$

所以原来的能级 E_{nlj} 分裂成 $2j+1$ 条，简并度完全被去除。然后再加上 $(Be/2m_e) \hat{S}_z$ 这一项，由于

$$\langle ljm_j | S_z | ljm_j \rangle = \begin{cases} m_j \hbar / 2j, & j = l + (1/2) \\ -m_j \hbar / 2(j+1), & j = l - (1/2) \quad (l \neq 0) \end{cases}$$

所以能级又成为（这里也利用了微扰论，见后面的第十一章）

$$E_{nljm_j} = E_{nlj} + \frac{Be\hbar}{2m_e} \begin{cases} m_j \left(1 + \frac{1}{2j}\right), & j = l + (1/2) \\ m_j \left(1 - \frac{1}{2(j+1)}\right), & j = l - (1/2) \quad (l \neq 0) \end{cases}$$

为了进一步得到光谱线的分裂情况，我们要利用选择定则（也见第十一章）

$$\Delta l = \pm 1, \quad \Delta j = 0, \pm 1, \quad \Delta m_j = 0, \pm 1.$$

由于现在能级分裂的大小和 l, j 都有关，而精细结构的谱线总是由 l, j 不同的能级间的跃迁所产生，因此谱线的分裂情形就比正常 Zeeman 效应复杂得多。仍然以钠 Na 的 D 黄线为例，上一节中已经指出：由于自旋-轨道耦合的缘故，它会分裂为 D₁ 和 D₂ 两条谱线（精细结构），而加了弱磁场以后，D₁ 又会再分裂为 4 条、D₂ 再分裂为 6 条谱线，所以一共分裂出 10 条谱线。这就是钠黄线的反常 Zeeman 效应，见图 10.4.2。

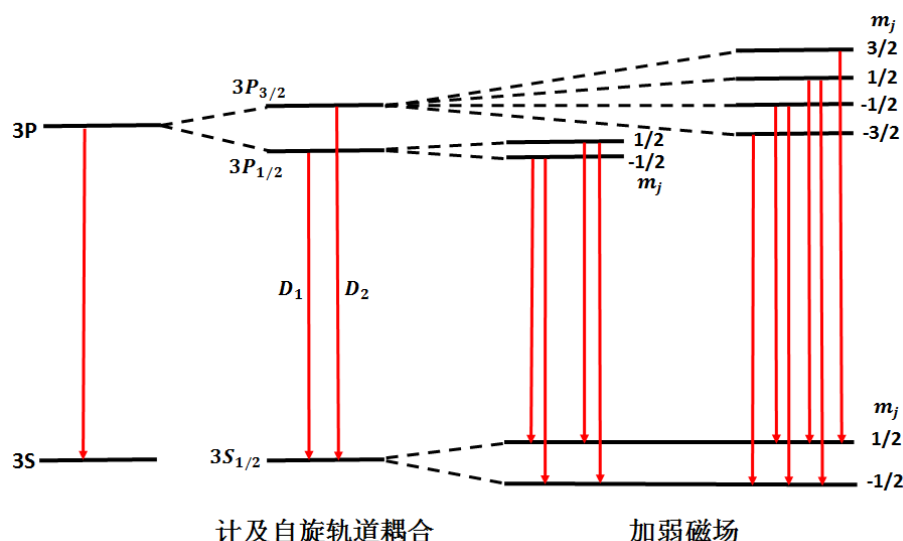


图 10.4.2 反常 Zeeman 效应的能级分裂和谱线分裂

*10.4.4 自旋电子学简介

现在我们知道了：电子是既带有电荷又带有自旋的、服从 Fermi-Dirac 统计的微观粒子。这是量子力学诞生以后人们对电子的不同于经典物理学的新认识。但是到目前为止，传统的电子学仍然只利用了电子带有电荷这个性质，而没有利用电子还带有自旋这个性质。

传统的电子学以超大规模集成电路为主要的支持硬件。如果我们想继续提高芯片的集成度，在原则上将面临两大困难。第一，当芯片上的线宽小到可以和电子的 de Broglie 波长相比的时候，电子的量子力学效应（波动效应）就不可以忽略了。这时候，集成电路已经不再能够简单地等效为分立（可隔离）元件构成的网络，而应当看作是一个量子力学系统。这样一来，对它进行分析的方法就要复杂多了。第二，芯片集成度的提高必然伴随着芯片单位面积上功耗的增加，也就是说芯片散热的问题会变得非常严重。在利用电子带有电荷这个特点对它进行操控的情况下，这个问题是无法避免的，因为电场力推动电荷运动的时候一定会做功，而这个功最后就会变成热量。

理论分析和工程实践的结果都表明，在目前的芯片制造的工艺水平上（线宽 $\square 0.1 \mu\text{m}$ ），前一个问题还不会产生明显的影响，而后一个问题倒是需要认真对待的。为了解决这个问题，必须有一个基本上不消耗能量就可以对电子进行操控的办法，而这正是控制电子的自旋。通过控制电子的自旋状态使之携带、加工和传输信号(信息)的方法称为自旋电子学 (Spintronics)。在自旋电子学里经常出现的术语是“自旋流”，“自旋池”，“自旋输运”，“自旋进动”，“自旋翻转”等等。我们在前面介绍的自旋-轨道耦合效应和 Zeeman 效应为自旋电子学的实现提供了物理的基础，因为前者通过电场、后者通过磁场影响电子的自旋状态。这些效应的最大特点就是基本上不需要付出能量的代价，而这正是我们要追求的目标。

自旋电子学是一门新兴的学科，目前仍然主要是在做基础研究的工作，但是它的发展前景是很好的。这里我们想强调一点：量子力学是研究自旋电子学的基本出发点。

§10.5 自旋纠缠态

10.5.1 两个电子自旋的合成 单态和三重态

设 \hat{S}_1 和 \hat{S}_2 代表两个电子的自旋，它们的总自旋是 $\hat{S} = \hat{S}_1 + \hat{S}_2$ ，对照角动量合成的一般规则，现在 $j_1 = j_2 = 1/2$ ，所以总自旋的大小可以取值

$$S = 1, 0.$$

形象地说，当两个电子的自旋互相平行的时候 $S = 1$ ，而当两个电子的自旋反平行的时候 $S = 0$ 。

我们还要解决总自旋的本征态如何用各电子的态矢量来表达的问题，换句话说，也就是要计算这个时候的 CG 系数。我们约定用 $|S, m_s\rangle$ 来表示总自旋本征态，而单个电子的状态分别用 α (自旋的 z 分量向上) 和 β (自旋的 z 分量向下) 来表示，并且在它们的右下角添上 1 和 2 来区分两个电子。合成后的状态 $|1, 1\rangle$ 和 $|1, -1\rangle$ 的分解是显然的：

$$|1, 1\rangle = \alpha_1 \alpha_2, \quad |1, -1\rangle = \beta_1 \beta_2.$$

但是 $|1, 0\rangle$ 和 $|0, 0\rangle$ 就比较复杂。我们必须假设，比如

$$|1, 0\rangle = c_1 \alpha_1 \beta_2 + c_2 \beta_1 \alpha_2,$$

因为这两项都给出 $S_z = 0$ 。为了决定 c_1, c_2 ，应该要求

$$S^2 |1, 0\rangle = 2\hbar^2 |1, 0\rangle.$$

注意到

$$S^2 = S_1^2 + S_2^2 + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z} = (3/2)\hbar^2 + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z},$$

以及

$$S_+\alpha = 0, \quad S_-\alpha = \hbar\beta,$$

$$S_+\beta = \hbar\alpha, \quad S_-\beta = 0,$$

$$S_z\alpha = (1/2)\hbar\alpha, \quad S_z\beta = -(1/2)\hbar\beta,$$

我们发现

$$\begin{aligned} S^2 |1, 0\rangle &= \left((3/2)\hbar^2 + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z} \right) (c_1 \alpha_1 \beta_2 + c_2 \beta_1 \alpha_2) \\ &= (3/2)\hbar^2 c_1 \alpha_1 \beta_2 + \hbar^2 c_1 \beta_1 \alpha_2 - (1/2)\hbar^2 c_1 \alpha_1 \beta_2 + (3/2)\hbar^2 c_2 \beta_1 \alpha_2 + \hbar^2 c_2 \alpha_1 \beta_2 - (1/2)\hbar^2 c_2 \beta_1 \alpha_2 \\ &= \hbar^2 (c_1 + c_2) \alpha_1 \beta_2 + \hbar^2 (c_1 + c_2) \beta_1 \alpha_2 \\ &= 2\hbar^2 |1, 0\rangle \\ &= 2\hbar^2 (c_1 \alpha_1 \beta_2 + c_2 \beta_1 \alpha_2), \end{aligned}$$

所以必须有

$$c_1 + c_2 = 2c_1 = 2c_2,$$

这就给出了

$$c_1 = c_2 = \frac{1}{\sqrt{2}}.$$

其中已经考虑了归一化。所以

$$|1, 0\rangle = \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 + \beta_1 \alpha_2).$$

类似地可以得到

$$|0, 0\rangle = \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \beta_1 \alpha_2).$$

上面这种方法的物理意义比较清楚，但是演算比较繁琐。更简明计算的方法是利用阶梯算符（升级/降级算符）。具体地说，一个电子的自旋算符是 $\hat{S} = (\hbar/2)\vec{\sigma}$ ，所以自旋投影的上升和下降算符分别是

$$S_+ = S_x + iS_y = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S_- = S_x - iS_y = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

这直接验证了

$$\begin{cases} S_+ \alpha = 0, & S_+ \beta = \hbar \alpha, \\ S_- \alpha = \hbar \beta, & S_- \beta = 0. \end{cases}$$

两个电子的总自旋的升级/降级算符是单个电子的升级/降级算符的和，

$$S_{\pm} = S_{1\pm} + S_{2\pm},$$

第九章中已经给出

$$|1, 0\rangle = \frac{1}{\sqrt{2}\hbar} S_- |1, 1\rangle,$$

而 $|1, 1\rangle = \alpha_1 \alpha_2$ ，所以

$$|1, 0\rangle = \frac{1}{\sqrt{2}\hbar} (S_{1-} + S_{2-})(\alpha_1 \alpha_2) = \frac{1}{\sqrt{2}}(\alpha_1 \beta_2 + \beta_1 \alpha_2).$$

类似地，设

$$|0, 0\rangle = b_1 \alpha_1 \beta_2 + b_2 \beta_1 \alpha_2,$$

那么

$$S_- |0, 0\rangle = (S_{1-} + S_{2-})(b_1 \alpha_1 \beta_2 + b_2 \beta_1 \alpha_2) = b_1 \beta_1 \beta_2 + b_2 \beta_1 \beta_2 = (b_1 + b_2) \beta_1 \beta_2 = 0,$$

所以

$$b_1 = -b_2 = \frac{1}{\sqrt{2}}, \quad |0, 0\rangle = \frac{1}{\sqrt{2}}(\alpha_1 \beta_2 - \beta_1 \alpha_2).$$

总结一下。 $S = 1$ 是一个三重态， $m = 1, 0, -1$ 的态矢量分别为

$$\alpha_1 \alpha_2, \quad \frac{1}{\sqrt{2}}(\alpha_1 \beta_2 + \beta_1 \alpha_2), \quad \beta_1 \beta_2,$$

而 $S = 0$ 是一个单态（只有 $m = 0$ ），态矢量为

$$\frac{1}{\sqrt{2}}(\alpha_1 \beta_2 - \beta_1 \alpha_2).$$

它们有一个非常重要的特点： $S = 1$ 的三个状态对于两个电子的**交换是对称的**，而 $S = 0$ 的状态对于两个电子的交换是**反对称的**。有时候为了更形象地表达这些状态，改用 $|\uparrow\rangle$ 表示 α ， $|\downarrow\rangle$ 表示 β ，那么

$$|1, 1\rangle = |\uparrow, \uparrow\rangle,$$

$$|1, -1\rangle = |\downarrow, \downarrow\rangle,$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle),$$

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle).$$

两个电子的总自旋本征态的这种交换对称/反对称性有重要的意义。电子是费米子，系统的波函数应该是交换反对称的，而总的波函数是空间波函数和自旋波函数的乘积。所以，当两个电子的自旋合成为 $S = 1$ 的时候，它的空间波函数必须是交换反对称的，而当两个电子的自旋合成为 $S = 0$ 的时候，它的空间波函数必须是交换对称的。这些要求对于多电子原子的电子壳层的形成以及量子化学中化学键的形成等问题有重要的作用。

*10.5.2 两电子自旋纠缠态 Bell 基

观察上面的两个电子的总自旋的本征态，我们发现它们有两种不同的构成方式： $|1, 1\rangle$ 和 $|1, -1\rangle$ 是两个电子自旋本征态的简单直乘，而 $|1, 0\rangle$ 和 $|0, 0\rangle$ 是两个电子自旋本征态的不同直乘的线性组合。我们把前者称为可分离态 (separable state)，而后者称为纠缠态 (entangled state)。

一般地说，设一个量子系统由若干个子系统构成，如果该系统的量子态可以表示为各子系统量子态的直接乘积，则称为可分离态，否则就称为纠缠态。实际上，“纠缠 (entanglement)”，“耦合 (coupling)”，“关联 (correlation)”在量子力学里表达了相似的概念或状况，很难严格加以区分。

在两电子自旋这个系统中，我们也可以把系统的基底完全建立在纠缠态上，那就是

$$\begin{aligned}\psi^+ &= \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle), & \psi^- &= \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle), \\ \phi^+ &= \frac{1}{\sqrt{2}}(|\uparrow, \uparrow\rangle + |\downarrow, \downarrow\rangle), & \phi^- &= \frac{1}{\sqrt{2}}(|\uparrow, \uparrow\rangle - |\downarrow, \downarrow\rangle).\end{aligned}$$

它们不再是 $\{S^2, S_z\}$ 的共同本征态，但是不难验证它们是 $\{S_{1x}S_{2x}, S_{1y}S_{2y}, S_{1z}S_{2z}\}$ 中任何两个算符的共同本征态。具体地说，我们有

$$\begin{cases} S_{1x}S_{2x}\psi^\pm = \pm(\hbar^2/4)\psi^\pm, & S_{1x}S_{2x}\phi^\pm = \pm(\hbar^2/4)\phi^\pm, \\ S_{1y}S_{2y}\psi^\pm = \pm(\hbar^2/4)\psi^\pm, & S_{1y}S_{2y}\phi^\pm = \mp(\hbar^2/4)\phi^\pm, \\ S_{1z}S_{2z}\psi^\pm = -(\hbar^2/4)\psi^\pm, & S_{1z}S_{2z}\phi^\pm = +(\hbar^2/4)\phi^\pm. \end{cases}$$

这一套基底称为两电子自旋系统的贝尔 (Bell) 基。从量子纠缠的角度来说，Bell 基的优点是它们是最大纠缠态，而这对于量子信息学是很重要的。

量子纠缠这个概念最早是由 Schrödinger 提出来的(1935)，并因此产生了著名的 Schrödinger 猫 (Schrödinger's cat)。也在同一年，Einstein-Podolsky-Rosen 基于对纠缠态的测量行为的分析，向量子力学的“正统解释”(Copenhagen 解释)提出了质疑，被称为 EPR 佯谬(EPR paradox)。在实质上，它涉及到量子力学的非定域性(non-locality)与相对论的定域性(locality)之间的矛盾。后来有一些学者(如波姆 Bohm)提出了隐变量(hidden variables)假说，试图以定域实在论(local realism)的方式“解释”或“导出”量子力学，而 Bell 根据定域实在论提出了可以通过实验进行检验的 Bell 不等式。这些都属于向量子力学提出“挑战”的尝试。但是到目前为止，隐变量理论并未获得实验的支持，关于 Bell 不等式的实验检验结果与量子力学的预言相符，EPR 佯谬也并未“推翻”量子力学。所以我们应该说，到目前为止，尽管量子力学的非定域性与相对论的定域性之间的矛盾还没有解决，量子测量中的波包坍缩的实质（也就是量子测量过程的动力学）也还没搞清楚，量子力学和实验的符合程度却是非常令人满意的。所以，量子力学的“相对真理”的地位是无可置疑的。

另一方面，更值得一提的是，量子信息学(quantum information)在近年来的迅猛发展，把量子纠缠的研究提到了一个前所未有的高度，推到了一个崭新的阶段。概括地说，量子纠缠是实现量子信息学的各种功能（例如量子计算(quantum computing)，量子算法(quantum algorithm)，量子远程传态(quantum teleportation)，量子密码(quantum cryptogram) 等等）的必不可少的“资源”，换句话说，量子纠缠的存在，是量子信息学区别于经典信息学的根源。对这方面的发展给予适当的关心，对于更深入地理解量子力学是有益的。

泡利矩陣

维基百科，自由的百科全书

在數學和數學物理中，**包立矩陣**是一組三個2×2的么正厄米複矩陣，^[1]一般都以希臘字母σ來表示，但有時當他們在和同位旋的對稱性做連結時，會被寫成τ。他們在包立表像（σ_z表像）可以寫成：

$$\begin{aligned}\sigma_1 &= \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \sigma_2 &= \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\ \sigma_3 &= \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\end{aligned}$$

這些矩陣是以物理學家沃爾夫岡·包立命名的。在量子力學中，它們出現在包立方程式中描述磁場和自旋之間交互作用的一項。所有的包立矩陣都是厄米矩陣，它們和單位矩陣*I*（有時候又被稱為第零號包立矩陣σ₀），的線性張成為2×2厄米矩陣的向量空間。

從量子力學的角度來看，埃爾米特矩陣（算符）代表可觀測的物理量，因此，σ_{*k*}, *k*= 0,1,2,3的線性張成代表所有作用在二維希爾伯特空間的物理量所形成的空間。從包立本人的的研究來看，σ_{*k*}, *k*=1,2,3所代表的物理量是自旋在三維歐幾里得空間ℝ³中第*k*個座標軸的投影分量。

目录
數學性質
 本徵值和本徵向量
 包立向量
 對易關係
 和內積、外積的關係
 包立向量的指數
 完備性關係
 和換位算符的關係
SU (2)
 四元數與包立矩陣
相關條目
參考文獻
 延伸閱讀

數學性質

三個包立矩陣可以共同用一種單一形式表達：

$$\sigma_a = \begin{bmatrix} \delta_{a3} & \delta_{a1} - i\delta_{a2} \\ \delta_{a1} + i\delta_{a2} & -\delta_{a3} \end{bmatrix}$$

其中δ_{*ab*}是克羅內克δ函數。當*a*=*b*時，其值為1；當*a*≠*b*時，其值為0。

本徵值和本徵向量

這些矩陣是對合的：

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

其中*I*是單位矩陣。

此外，包立矩陣的行列式和它們的跡分別為：

$$\begin{aligned}\det(\sigma_i) &= -1 \\ \mathrm{tr}(\sigma_i) &= 0\end{aligned}$$

故從上述關係可以推得每個包立矩陣σ_{*i*}的本徵值分別為±1。

每個包立矩陣有兩個本徵值，+1和-1，其對應的歸一化本徵向量為：

$$\begin{aligned}\psi_{x+} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \psi_{x-} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\ \psi_{y+} &= \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix} & \psi_{y-} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \\ \psi_{z+} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \psi_{z-} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}\end{aligned}$$

包立向量

包立向量定義為：

$$\vec{\sigma} = \sigma_1 \hat{x} + \sigma_2 \hat{y} + \sigma_3 \hat{z}$$

這個定義提供了將一般向量基底對應到包立矩陣的基底的機制

$$\begin{aligned}\vec{a} \cdot \vec{\sigma} &= (\alpha_i \hat{x}_i) \cdot (\sigma_j \hat{x}_j) \\ &= \alpha_i \sigma_j \hat{x}_i \cdot \hat{x}_j \\ &= \alpha_i \sigma_j \delta_{ij} \\ &= \alpha_i \sigma_i\end{aligned}$$

相同的下標是使用了愛因斯坦求和約定。此外：

$$\det \vec{a} \cdot \vec{\sigma} = -\vec{a} \cdot \vec{a} = -|\vec{a}|^2。$$

對易關係

包立矩陣有以下的對易關係：

$$[\sigma_a, \sigma_b] = 2i\varepsilon_{abc} \sigma_c,$$

以及以下的反對易關係。

$$\{\sigma_a, \sigma_b\} = 2\delta_{ab} I。$$

其中 ε_{abc} 是列維-奇維塔符號， δ_{ab} 是克羅內克函數，是 I 是2×2的單位矩陣。而一樣的，上面使用了愛因斯坦求和約定。

和內積、外積的關係

將包立矩陣的對易和反對易相加得：

$$\begin{aligned}[\sigma_a, \sigma_b] + \{\sigma_a, \sigma_b\} &= (\sigma_a \sigma_b - \sigma_b \sigma_a) + (\sigma_a \sigma_b + \sigma_b \sigma_a) \\ 2i \sum_c \varepsilon_{abc} \sigma_c + 2\delta_{ab} I &= 2\sigma_a \sigma_b\end{aligned}$$

因此可得：

$$\sigma_a \sigma_b = i \sum_c \varepsilon_{abc} \sigma_c + \delta_{ab} I$$

為了避免符號重複，將 a, b, c 改成 p, q, r ，然後把上式和三維向量 a_p 和 b_q 內積，可得：

$$\begin{aligned}a_p b_q \sigma_p \sigma_q &= a_p b_q \left(i \sum_r \varepsilon_{pqr} \sigma_r + \delta_{pq} I \right) \\ a_p \sigma_p b_q \sigma_q &= i \sum_r \varepsilon_{pqr} a_p b_q \sigma_r + a_p b_q \delta_{pq} I\end{aligned}$$

將它轉換成向量積的表達式：

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b}) I + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma}$$

包立向量的指數

令 $\vec{a} = a\hat{n}$ ，而且 $|\hat{n}| = 1$ 對於偶數 n 可得：

$$(\hat{n} \cdot \vec{\sigma})^{2n} = I$$

另外加上之前求得在 $n = 1$ 的情況可在 n 為奇數的情況：

$$(\hat{n} \cdot \vec{\sigma})^{2n+1} = \hat{n} \cdot \vec{\sigma}$$

利用矩陣指數的概念，加上正弦和餘弦的泰勒級數展開式，可得：

$$e^{ia(\hat{n} \cdot \vec{\sigma})} = I \cos a + i(\hat{n} \cdot \vec{\sigma}) \sin a \quad (2)$$

這可以看做是歐拉公式的類比。

完備性關係

另一個常用來區別包立矩陣的方法是用上標 i ，用不同的 i 來代表不同的包立矩陣，而下標則代表不同的矩陣元素。因此第 i 個包立矩陣的第 α 行第 β 列的元素可表示為 $\sigma_{\alpha\beta}^i$

利用這種表示方法，包立矩陣的完備性關係可寫作：

$$\vec{\sigma}_{\alpha\beta} \cdot \vec{\sigma}_{\gamma\delta} \equiv \sum_{i=1}^3 \sigma_{\alpha\beta}^i \sigma_{\gamma\delta}^i = 2\delta_{\alpha\delta} \delta_{\beta\gamma} - \delta_{\alpha\beta} \delta_{\gamma\delta}$$

證明

因為所有的包立矩陣，和 2×2 的單位矩陣可做為所有 2×2 矩陣在希爾伯特空間中的正交基底，表示任何一個複係數矩陣 M 皆可表示為：

$$M = cI + \sum_i a_i \sigma^i$$

其中 c 是一複數， a_i 是一複向量中的三個係數。

利用之前給的關係式，容易證明：

$$\text{tr } \sigma^i \sigma^j = 2\delta_{ij}$$

"tr"表示對該矩陣取其跡，因此， $c = \frac{1}{2} \text{tr } M$ 和 $a_i = \frac{1}{2} \text{tr } \sigma^i M$ 成立。

故，

$$2M = I \text{tr } M + \sum_i \sigma^i \text{tr } \sigma^i M$$

用矩陣的標號表示的話就成為：

$$2M_{\alpha\beta} = \delta_{\alpha\beta} M_{\gamma\gamma} + \sum_i \sigma_{\alpha\beta}^i \sigma_{\gamma\delta}^i M_{\delta\gamma}$$

在等號右邊，針對了兩個重複出現的標號 γ 和 δ ，使用了愛因斯坦求和約定。而因為這關係對所有矩陣 M 都成立，因此要證的完備性關係必然成立。

有時習慣上將 2×2 單位舉寫成 σ_0 ，也就是， $\sigma_{\alpha\beta}^0 = \delta_{\alpha\beta}$ 。如此一來完備性關係可以更為簡潔的表示成：

$$\sum_{i=0}^3 \sigma_{\alpha\beta}^i \sigma_{\gamma\delta}^i = 2\delta_{\alpha\delta} \delta_{\beta\gamma}$$

和換位算符的關係

令算符 P_{ij} 為換位算符（或稱為置換算符）。對於兩個在張量積空間 $\mathbb{C}^2 \otimes \mathbb{C}^2$ 中的自旋 σ_i 和 σ_j 該算符有：

