

初识量子力学

陈童

2024 年 8 月 22 日

目录

第四章 电子自旋以及相关模型	2
4.1 电子自旋	2
4.1.1 电子自旋的发现	2
4.1.2 泡利算符与自旋算符	4
4.1.3 泡利矩阵	7
4.2 两自旋耦合系统	10
4.2.1 海森堡模型	10
4.2.2 求解与讨论	11

第四章 电子自旋以及相关模型

前面几章着重在建立现代量子力学的理论体系，本章将研究两个具体例子。首先就是研究电子自旋，电子自旋可以说是最简单，同时又最具有量子力学特征的例子了，而且上一章中我们已经涉及到电子自旋了，但是上一章并没有系统研究电子自旋的描述，本章将从最基本的实验事实出发，利用量子力学的基本理论，建立起对电子自旋的完整描述。其次，本章还要研究一个和电子自旋相关的物理模型，具体来说就是两自旋耦合的模型。这依然是用来作为上一章理论体系的一个具体例子。

4.1 电子自旋

4.1.1 电子自旋的发现

这一小节让我们更仔细地研究一下施特恩和格拉赫(Stern and Gerlach)的原版实验。

让我们回到历史。1922年施特恩和格拉赫(Stern and Gerlach)做了一个著名的实验，这个实验最终表明电子有一个内禀的角动量，也就是电子自旋。在经典物理中，当一个粒子围绕另一个粒子作轨道运动时，它就有一个轨道角动量 $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ 。而且我们知道，对于一个带电量为 q 的带电粒子，它的轨道角动量会产生一个磁矩，我们可以记作 $\vec{\mu}_L$ ，电动力学的知识(不知道也没关系，先接受这个结果就可以了)告诉我们

$$\vec{\mu}_L = \frac{q}{2m} \mathbf{L}. \quad (4.1)$$

当然，我们也知道，磁矩会和磁场耦合，如果存在一个磁场 \mathbf{B} ，那磁矩 $\vec{\mu}$ 和

磁场相互作用的势能 V 将是

$$V = -\vec{\mu} \cdot \mathbf{B}. \quad (4.2)$$

但是电子的自旋角动量完全是内禀的，它和轨道运动没有任何关系。人们可以粗略地将它想象成是电子围绕着自身的转轴自转，但是严格来说，这幅经典物理的图像是错的，电子的自旋纯粹是一个相对论量子力学的效应。实际上，由于电子有自旋，所以一个电子要转720度(而不是360度)才能回到原来的状态。所以在我们经典世界的人看来，电子自旋的存在是不可思议的，在理论上，要等到狄拉克提出著名的相对论量子力学方程，也就是狄拉克方程，人们才算对电子自旋有了比较深入的理解。

施特恩和格拉赫的实验装置如图(4.1)a所示。在这个实验中，施特

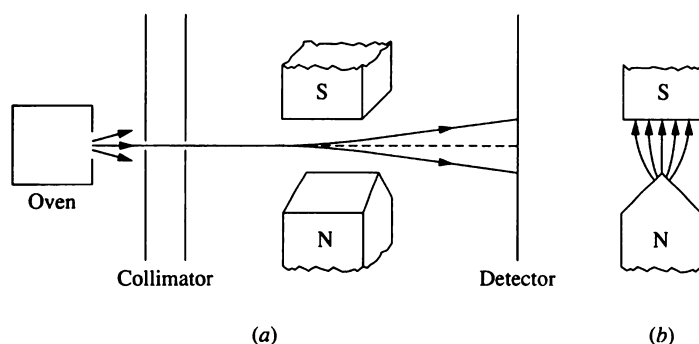


图 4.1: 施特恩-格拉赫实验。

恩和格拉赫让一束经过准直的银原子束通过一个非均匀磁场，然后再打在一个屏幕上。这个磁场如图(4.1)b所示，它在竖直的 z 方向有一个梯度 $\frac{\partial B_z}{\partial z}$ 。因此如果银原子有一个磁矩 μ 的话，它在 z 方向将受到一个力 F_z ， $F_z = -\frac{\partial V}{\partial z} = \vec{\mu} \cdot \frac{\partial \mathbf{B}}{\partial z} \simeq \mu_z \frac{\partial B_z}{\partial z}$ 。从图(4.1)b可以看到，在这里磁场的梯度 $\frac{\partial B_z}{\partial z}$ 其实是负的(下方的磁场更强，上方的磁场更弱)，因此如果 μ_z 为负的话，那银原子的受力将沿着 z 轴的正方向。

按照经典物理， $\mu_z = |\mu| \cos \theta$ ， θ 是银原子的磁矩与 z 方向的夹角。在斯特恩-格拉赫实验中，入射的银原子束是完全非极化的，也就是说，要么它的磁矩为0，要么它磁矩的方向就是随机的，即经典物理告诉我们的 θ 将有

连续的不同取值。如果银原子束的磁矩为0，那银原子将不受磁场的作用，因此最终打在屏幕上的银原子将会是一束。如果银原子有磁矩，那按照经典物理， θ 连续可变，因此银原子的受力 $F_z \simeq \mu_z \frac{\partial B_z}{\partial z} = \frac{\partial B_z}{\partial z} |\mu| \cos \theta$ 将在一个范围内连续可变，磁矩取向不同的银原子将会受到不同的力，从而最终会打在屏幕的不同位置，那最终打在屏幕上的银原子将展宽成连续的一片。

但是，实验发现，通过非均匀磁场以后，银原子束分裂了，但是仅仅分裂成了上下两束！也就是说，银原子的确有磁矩，但是 μ_z 的取值不是连续可变的，而是量子化的， μ_z 只有两个不同的分立取值。

实际上，银原子有47个核外电子和一个很重的原子核。斯特恩-格拉赫实验发现的银原子磁矩不可能来自于银原子核，因为磁矩的大小反比于粒子的质量，银原子核很重，其磁矩将会很小从而可以忽略。另外，我们知道带电粒子的磁矩总是来源于其角动量，而银原子核外的47个电子中，46个电子的状态完全球对称，因此总角动量是0，因此银原子的磁矩必定来源于额外的那第47个电子。但是最后的这个电子的空间波函数也是球对称的，因此其轨道角动量也是0。因此，斯特恩-格拉赫的实验结果就意味着，电子除了轨道角动量以外，必定还有一个内禀的角动量，也就是自旋！它也是一种角动量，通常记作 \mathbf{S} ，银原子的磁矩就来自于由它产生的自旋磁矩。而 μ_z 的量子化就意味着，电子自旋的 z 分量 S_z 必定是量子化的，而且只能有两个不同的分立取值。当然， z 分量相对于 x 分量和 y 分量来说并没有任何特殊的地方，因此 μ_x, S_x 以及 μ_y, S_y 必定也都是量子化的，也都只能有两个不同的分立取值。

4.1.2 泡利算符与自旋算符

根据施特恩-格拉赫实验，通过施特恩-格拉赫装置(假设磁场沿着 z 方向)的电子会分裂成两束，因此电子磁矩的 z 分量 μ_z 必然取量子化的两个分立值，不妨记为 $\pm\mu_B$ 。当然， z 方向并没有任何特殊，我们同样可以将施特恩-格拉赫装置转向 x 方向和 y 方向(即转动转置使得磁场分别沿着 x 方向和 y 方向)，进而可知，电子磁矩的 x 分量 μ_x 以及 y 分量 μ_y 也都只能取 $\pm\mu_B$ 两个量子化的分立值。同样，可以把施特恩-格拉赫装置转向任意 \mathbf{n} 方向(\mathbf{n} 为单位矢量)，进而即知 $\mu_{\mathbf{n}} = \vec{\mu} \cdot \mathbf{n}$ 也只能取 $\pm\mu_B$ 两个分立值。

另一方面，在量子力学中，电子的磁矩当然也要用厄米算符来描述，

仍然记电子的磁矩算符为 $\vec{\mu}$, 则根据上一段的论述可知, 对于任意单位矢量 \mathbf{n} , 算符 $\mu_{\mathbf{n}} = \vec{\mu} \cdot \mathbf{n}$ 都只有 $\pm\mu_B$ 两个本征值, 换言之

$$(\vec{\mu} \cdot \mathbf{n})^2 = \mu_B^2. \quad (4.3)$$

不妨引入一个不带量纲的算符 $\vec{\sigma}$ (即所谓的泡利算符), 它和电子磁矩算符的联系是

$$\vec{\mu} = \mu_B \vec{\sigma}. \quad (4.4)$$

将这个式子代入上面的(4.3), 即有

$$(\vec{\sigma} \cdot \mathbf{n})^2 = 1. \quad (4.5)$$

或者说, 对于任意 \mathbf{n} , $\vec{\sigma} \cdot \mathbf{n}$ 的本征值总是 ± 1 。

下面记 $\vec{\sigma}$ 的三个分量为 $\sigma_a, a = x, y, z$ (它们的本征值当然也都是 ± 1), 记单位矢量 \mathbf{n} 的三个分量为 $n_a, a = x, y, z$, 并记

$$\vec{\sigma} \cdot \mathbf{n} = \sigma_a n_a, \quad (4.6)$$

式中我们使用了所谓的求和约定, 即对表达式中重复出现两次的指标默认对它的所有可能性求和, 式中也就是默认对 a 求和, 后文也采用同样约定。根据这些约定, 我们就能把(4.5)式重写为

$$(\sigma_a n_a)(\sigma_b n_b) = 1. \quad (4.7)$$

式中 $a, b = x, y, z$ 。由这个式子即有

$$\sigma_a \sigma_b n_a n_b = 1 \Rightarrow \frac{1}{2}(\sigma_a \sigma_b + \sigma_b \sigma_a + [\sigma_a, \sigma_b]) n_a n_b = 1. \quad (4.8)$$

对于式中的 $[\sigma_a, \sigma_b] n_a n_b$, 通过交换指标 a, b (反正它们都是要求和的指标), 即有 $[\sigma_a, \sigma_b] n_a n_b = [\sigma_b, \sigma_a] n_b n_a = -[\sigma_a, \sigma_b] n_a n_b$, 一个式子等于负的它本身, 那这个式子必定等于零, 从而 $[\sigma_a, \sigma_b] n_a n_b = 0$ 。从而(4.8)式可以简化成

$$\frac{1}{2}(\sigma_a \sigma_b + \sigma_b \sigma_a) n_a n_b = 1. \quad (4.9)$$

与单位矢量的归一条件 $\mathbf{n}^2 = n_a n_a = \delta_{ab} n_a n_b = 1$ 比较, 并由于单位矢量 n_a 的任意性, 即知

$$\sigma_a \sigma_b + \sigma_b \sigma_a = 2\delta_{ab}. \quad (4.10)$$

满足类似这组式子的抽象代数就叫Clifford代数，数学家已经系统地研究过这类代数的矩阵表示了，后文我们将演示给大家这样的表示可以如何进行。

(4.10)式也可以更清楚地写成

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \quad (4.11)$$

以及

$$\sigma_x \sigma_y + \sigma_y \sigma_x = 0, \quad \sigma_y \sigma_z + \sigma_z \sigma_y = 0, \quad \sigma_z \sigma_x + \sigma_x \sigma_z = 0. \quad (4.12)$$

也就是说，这三个泡利算符是两两反交换的，或者说反对易的，所谓两个算符反对易即是指 $AB = -BA$ 。

根据上面的反对易性，不难验证算符 $\sigma_x \sigma_y \sigma_z$ 和三个泡利算符 $\sigma_x, \sigma_y, \sigma_z$ 均对易！但是只有常数才可能和所有的 σ_a 均对易，所以必定有 $\sigma_x \sigma_y \sigma_z$ 为常数。另一方面，根据上面反对易关系(4.12)式，通过不断将两个相同的 σ_a 分量对易到一起，进而利用(4.11)式，从而不难得到

$$(\sigma_x \sigma_y \sigma_z)^2 = -1. \quad (4.13)$$

从而 $\sigma_x \sigma_y \sigma_z = \pm i$ ，由于只要将 $-\sigma_a$ 重新定义成 σ_a 就能交换式中的正负号，所以这两种可能性本质上是等价的，通常取正号，所以有

$$\sigma_x \sigma_y \sigma_z = i. \quad (4.14)$$

将这个式子两边分别乘以 $\sigma_x, \sigma_y, \sigma_z$ ，并利用(4.11)式和(4.12)式，即能得到

$$\sigma_x \sigma_y = i \sigma_z, \quad \sigma_y \sigma_z = i \sigma_x, \quad \sigma_z \sigma_x = i \sigma_y, \quad (4.15)$$

这三个式子涉及的三个指标刚好构成轮换关系。后文我们将尤其要用到

$$\sigma_x \sigma_y = i \sigma_z. \quad (4.16)$$

电子磁矩正比于电子自旋，同时也正比于泡利算符，所以我们可以利用泡利算符来定义电子的自旋算符 \mathbf{S} ，它的三个分量分别定义为

$$S_x = (\hbar/2)\sigma_x, \quad S_y = (\hbar/2)\sigma_y, \quad S_z = (\hbar/2)\sigma_z. \quad (4.17)$$

从这个定义可以知道, S_a 的本征值必定为 $\pm\hbar/2$ 。

利用泡利算符的反对易性以及(4.15)式, 不难验证泡利算符满足如下代数关系

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (4.18)$$

进而根据上面的定义, 可得自旋算符的代数关系

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y. \quad (4.19)$$

式中涉及的三个指标刚好构成轮换关系。实际上, 这组代数关系不仅适用于电子自旋算符, 也可以推广到任何角动量算符。

由于自旋算符任何两个分量均不对易, 所以根据上一章讲述的定理, 电子自旋的任何两个分量都不可能同时确定, 比方说, 电子自旋的 S_x 和 S_y 不可能同时有确定值。这和经典物体的自转是完全不同的, 经典的角动量三个分量都可以同时有确定值。

4.1.3 泡利矩阵

因此, 要将电子自旋算符表示成矩阵, 我们只需要等价地将泡利算符表示成矩阵就可以了。而这可以按照Clifford代数的标准表示理论进行。

下面的推导在数学上是完全严格的。它完全是从泡利算符的抽象代数出发, 也即是从方程(4.11), (4.12)以及方程(4.16)出发, 然后构造出一切。尤其是, 我们没有隐含地假设任何算符的本征态的存在性, 而是直接构造出了这些本征态。

下面开始推导。首先我们注意到 σ_z 可以由 σ_x 和 σ_y 的乘积给出, 所以我们的注意力将主要放在 σ_x 和 σ_y 上。我们定义两个新的算符 σ 和 σ^\dagger ,

$$\sigma = (\sigma_x - i\sigma_y)/2, \quad \sigma^\dagger = (\sigma_x + i\sigma_y)/2. \quad (4.20)$$

则利用方程(4.11)和(4.12), 很容易得到

$$\sigma^2 = \sigma^{\dagger 2} = 0. \quad (4.21)$$

而且, 由于 $\sigma^\dagger \sigma = (\sigma_x + i\sigma_y)(\sigma_x - i\sigma_y)/4 = (\sigma_x^2 + \sigma_y^2 + i\sigma_y\sigma_x - i\sigma_x\sigma_y)/4 = (1 + \sigma_z)/2$, 这里最后一个等号我们利用了 σ_x 和 σ_y 的反对易以及 $\sigma_x\sigma_y = i\sigma_z$,

类似的，也可以算得 $\sigma\sigma^\dagger = (1 - \sigma_z)/2$ ，归纳一下即有

$$\sigma^\dagger\sigma = (1 + \sigma_z)/2, \sigma\sigma^\dagger = (1 - \sigma_z)/2. \quad (4.22)$$

由于 $\sigma^2 = 0$ ，所以必定可以在电子的所有自旋量子态中找到某个量子态，记作 $|\downarrow\rangle$ ，它满足

$$\sigma|\downarrow\rangle = 0. \quad (4.23)$$

这是因为，任取一个自旋量子态 $|\psi\rangle$ ，如果它满足 $\sigma|\psi\rangle = 0$ ，那这个 $|\psi\rangle$ 就是我们要找的态，如果 $\sigma|\psi\rangle \neq 0$ ，那我们就可以令 $\sigma|\psi\rangle = |\phi\rangle$ ，这时候将有 $\sigma|\phi\rangle = \sigma^2|\psi\rangle = 0$ （后一个等号是由于 $\sigma^2 = 0$ ），因此 $|\phi\rangle$ 就将是我们要找的态。总之，满足方程(4.23)的量子态 $|\downarrow\rangle$ 总能找到。对于这个态我们必有 $0 = \langle\downarrow|\sigma^\dagger\sigma|\downarrow\rangle = \langle\downarrow|(1 + \sigma_z)|\downarrow\rangle/2$ ，即有 $\langle\downarrow|(1 + \sigma_z)|\downarrow\rangle = 0$ ，由于 σ_z 的本征值为 ± 1 ，因此这个结果就告诉我们 $|\downarrow\rangle$ 必定是 σ_z 的本征值为 -1 的本征态，即

$$\sigma_z|\downarrow\rangle = -|\downarrow\rangle. \quad (4.24)$$

这就是为什么记这个态为自旋向下态的原因，因为 σ_z 在这个态上的本征值为 -1 ，这就意味着它是 S_z 的本征值为 $-\hbar/2$ 的本征态，我们称这样的量子态为自旋向下态。

下面引入一个新的自旋量子态，记作 $|\uparrow\rangle$ ，它的定义是

$$|\uparrow\rangle = \sigma^\dagger|\downarrow\rangle. \quad (4.25)$$

很显然 $\langle\uparrow|\sigma\sigma^\dagger|\uparrow\rangle = \langle\uparrow|\sigma\sigma^{\dagger 2}|\downarrow\rangle = 0$ ，利用 $\langle\uparrow|\sigma\sigma^\dagger|\uparrow\rangle = \langle\uparrow|(1 - \sigma_z)/2|\uparrow\rangle$ ，即有 $\langle\uparrow|(1 - \sigma_z)|\uparrow\rangle = 0$ ，从而 $|\uparrow\rangle$ 必为 σ_z 的本征值为 $+1$ 的本征态，即

$$\sigma_z|\uparrow\rangle = |\uparrow\rangle. \quad (4.26)$$

这就是为什么将这个态记作自旋向上态的基本原因，因为它作为 σ_z 的本征值为 $+1$ 的本征态就必然也是 S_z 的本征值为 $+\hbar/2$ 的本征态。

将方程(4.25)共轭转置，就有 $\langle\uparrow| = \langle\downarrow|\sigma$ ，因此， $\langle\uparrow|\downarrow\rangle = \langle\downarrow|\sigma|\downarrow\rangle = 0$ （后一个等号是由于方程(4.23)），也即

$$\langle\uparrow|\downarrow\rangle = 0. \quad (4.27)$$

因此自旋向上态和自旋向下态必定是两个正交的量子态，这正符合这两个态可以确定地区分的物理要求。而且也可以算得 $\langle \uparrow | \uparrow \rangle = \langle \downarrow | \sigma \sigma^\dagger | \downarrow \rangle = \langle \downarrow | (1 - \sigma_z)/2 | \downarrow \rangle = \langle \downarrow | \downarrow \rangle$, 因此如果我们将最开始时选的 $|\downarrow\rangle$ 态归一化, 那由方程(4.25)定义的 $|\uparrow\rangle$ 也将自动是归一化的, 即

$$\langle \downarrow | \downarrow \rangle = \langle \uparrow | \uparrow \rangle = 1. \quad (4.28)$$

由于同时满足正交归一性, 因此我们构造出来的 $\{|\downarrow\rangle, |\uparrow\rangle\}$ 可以作为电子自旋量子态的两个基矢量。

从物理上来说, 这两个基矢量当然应该是完备的。但是, 我们能够从泡利算符的代数出发从数学上证明这一点吗? 回答很简单, 可以。比方说我们观察刚才的两个基矢量 $|\downarrow\rangle, |\uparrow\rangle = \sigma^\dagger |\downarrow\rangle$, 我们发现 $|\uparrow\rangle$ 是用 σ^\dagger 作用在 $|\downarrow\rangle$ 上得到的, 那我们能用 σ^\dagger 再重复作用一次从而得到第三个态吗? 答案是不能, 因为 $\sigma^{\dagger 2} = 0$, 因此 σ^\dagger 的重复作用只能得到 0, 即

$$\sigma^\dagger |\uparrow\rangle = 0. \quad (4.29)$$

那如果用 σ 作用在 $|\uparrow\rangle$ 上能得到新的态吗? 答案是, 也不能, 因为 $\sigma |\uparrow\rangle = \sigma \sigma^\dagger |\downarrow\rangle = (1 - \sigma_z)/2 |\downarrow\rangle = |\downarrow\rangle$, 也即

$$\sigma |\uparrow\rangle = |\downarrow\rangle. \quad (4.30)$$

因此我们就证明了电子自旋的基矢量 $\{|\downarrow\rangle, |\uparrow\rangle\}$ 不能再进一步扩大, 从而它们就是完备的。因此它们就可以作为电子自旋态空间的矢量基。

以 $\{|\uparrow\rangle, |\downarrow\rangle\}$ 为矢量基(现在要注意这两个基矢量的排列顺序), 就能够分别计算出泡利算符 $\sigma_z, \sigma, \sigma^\dagger$ 在这个基中的表示矩阵仍然记作 $\sigma_z, \sigma, \sigma^\dagger$ (也就是说, 我们使用了相同的符号来表示算符以及它们的表示矩阵, 读者不难根据上下文进行区分)。

$$\sigma_z = \begin{bmatrix} \langle \uparrow | \sigma_z | \uparrow \rangle & \langle \uparrow | \sigma_z | \downarrow \rangle \\ \langle \downarrow | \sigma_z | \uparrow \rangle & \langle \downarrow | \sigma_z | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (4.31)$$

$$\sigma = \begin{bmatrix} \langle \uparrow | \sigma | \uparrow \rangle & \langle \uparrow | \sigma | \downarrow \rangle \\ \langle \downarrow | \sigma | \uparrow \rangle & \langle \downarrow | \sigma | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad (4.32)$$

$$\sigma^\dagger = \begin{bmatrix} \langle \uparrow | \sigma^\dagger | \uparrow \rangle & \langle \uparrow | \sigma^\dagger | \downarrow \rangle \\ \langle \downarrow | \sigma^\dagger | \uparrow \rangle & \langle \downarrow | \sigma^\dagger | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (4.33)$$

在计算中, 需要使用方程(4.23), (4.24), (4.25), (4.26), (4.27), (4.28), (4.29), (4.30), 当然, 这些方程中的每一个其实都很简单。

而由 σ, σ^\dagger 的定义式(4.20), 我们又可以进一步得到 σ_x 的表示矩阵 $\sigma_x = \sigma + \sigma^\dagger$ 以及 σ_y 的表示矩阵 $\sigma_y = i(\sigma - \sigma^\dagger)$ 。这三个矩阵 $\sigma_x, \sigma_y, \sigma_z$ 就是著名的泡利矩阵, 我们将它们归纳在下面,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (4.34)$$

实际上, 泡利算符 σ_x 就是上一章定义过的算符 X , 泡利算符 σ_z 就是上一章定义的算符 Z . X, Z 这样的记号在量子信息和量子计算中用得多一些, 泡利算符的记号则在其它物理领域用得多一些。

4.2 两自旋耦合系统

4.2.1 海森堡模型

铁, 钴, 镍等物质在很小的外磁场影响下, 就能产生远大于其他物质的磁化效应, 这就是铁磁性。这些物质之所以有铁磁性, 是因为在它们的每一个小区域之内都会自发磁化形成磁畴。而自发磁化的原因, 是因为铁磁物质不同原子的磁矩产生了相互平行的指向, 就像这样($\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$)。这些原子的磁矩从哪儿来呢? 来自于铁磁原子未配对的核外电子的自旋, 已经配对的那些电子由于泡利不相容原理, 它们的自旋一定是相反的, 因而磁矩总是抵消为0。那么不同原子的这些未配对电子的自旋为什么会平行地指向呢? 为什么只有铁磁性物质会这样平行指向呢? 为了对这一问题进行理论研究, 海森堡提出, 这是因为铁磁性物质邻近原子的未配对电子存在自旋与自旋间的相互作用, 为了反映这种自旋相互作用是如何导致自发磁化的, 海森堡提出了一个简单的模型, 这就是今天依然有大量研究者在研究的著名的海森堡模型, 这个模型对于铁磁性的研究, 甚至对于整个凝聚态物理来说都非常重要, 因此我们想介绍一下。海森堡模型的哈密顿算符是下面这样的

$$H = J \sum_{\langle i, j \rangle} \vec{S}_i \cdot \vec{S}_j, \quad (4.35)$$

式中的 $\langle i, j \rangle$ 表示两个邻近格点位置 i 和 j , \vec{S}_i, \vec{S}_j 分别表示格点位置 i, j 上的原子的自旋, 式中的求和表示将每一对相互邻近的原子的贡献都加起来, 另外, J 的大小表示两自旋之间的耦合强度, 更常见的是将之写成 $-J$, 不过后面我们对 J 取正和取负两种情况都会进行讨论。

对于一个任意的两维或三维格点, 求解海森堡模型不是一件容易的事情, 因此我们当然不可能在这里讨论这样的课题。以上所说主要是给我们提供一个物理背景, 我们真正将要研究的, 是只考虑两个原子, 每个原子只有一个未配对电子, 这两个电子的自旋之间按照海森堡模型的形式进行耦合。这时候, 由于电子的自旋算符可以用泡利算符来进行表达, 所以我们可以将相应的模型写成下面的形式,

$$H = J(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z). \quad (4.36)$$

(4.36)才是我们这一节将要进行求解的问题。我们将会看到, 即使这样简单的一个模型也能告诉我们, 什么时候系统会出现铁磁性, 也就是两电子的自旋相平行, 什么时候系统又会出现反铁磁性, 也就是两电子的自旋反平行。不过, 首先我们想说一下, 什么是求解出这个系统? 我们这里所谓的求解出哈密顿算符(4.36), 实际上指的是求解出它的本征值, 也就是系统的本征能量 $\{E_n\}$, 以及相应的本征态 $\{|E_n\rangle\}$

$$H|E_n\rangle = E_n|E_n\rangle. \quad (4.37)$$

正如我们已经知道的, 由于在哈密顿算符的本征态 $|E_n\rangle$ 上, 系统有确定的能量取值 E_n , 所以这样的本征态也叫做定态, 相应的本征能量的集合也称作系统的能谱或者能级。为了求解这个定态薛定谔方程, 我们都需要取定一个表象(也就是选取一组正交归一的矢量基), 取定表象以后, 这个方程当然就变成了相应的哈密顿矩阵的本征方程。

4.2.2 求解与讨论

先总结一下上一节得到的关于泡利算符的一些有用结果。首先, 通常称上一节定义的 σ 为自旋降算符, 称 σ^\dagger 为自旋升算符, 这是因为 σ 的作用会把自旋向上态降为自旋向下, 而 σ^\dagger 的作用则会把自旋向下降为自旋向上,

如下

$$\sigma^\dagger |\downarrow\rangle = |\uparrow\rangle, \quad \sigma |\uparrow\rangle = |\downarrow\rangle. \quad (4.38)$$

其次, $|\downarrow\rangle$ 无法进一步再降了, 所以在 σ 的作用下必定为零, 同样 $|\uparrow\rangle$ 无法再升了, 所以在 σ^\dagger 的作用下也必定为零, 即

$$\sigma |\downarrow\rangle = 0. \quad (4.39)$$

$$\sigma^\dagger |\uparrow\rangle = 0. \quad (4.40)$$

公式(4.38)以及公式(4.39)、(4.40)这几个结果再加上下面的两个本征方程, 这些都是上一节得到的基本结论, 现在归纳在这里

$$\sigma^z |\uparrow\rangle = |\uparrow\rangle, \quad \sigma^z |\downarrow\rangle = -|\downarrow\rangle. \quad (4.41)$$

下面我们就可以开始求解(4.36)给出来的两自旋耦合系统了。首先, 两个不同电子的自旋当然是相互独立的, 因此任何两个属于不同电子的泡利算符都必定相互对易。其次, 类似于上一节的做法, 我们分别引入两个不同电子的自旋降算符和自旋升算符, $\sigma_1 = \frac{1}{2}(\sigma_1^x - i\sigma_1^y), \sigma_2 = \frac{1}{2}(\sigma_2^x - i\sigma_2^y), \sigma_1^\dagger = \frac{1}{2}(\sigma_1^x + i\sigma_1^y), \sigma_2^\dagger = \frac{1}{2}(\sigma_2^x + i\sigma_2^y)$ 。则很容易验证, 原来的哈密顿量(4.36)可以重写成

$$H = 2J(\sigma_1\sigma_2^\dagger + \sigma_1^\dagger\sigma_2) + J(\sigma_1^z\sigma_2^z). \quad (4.42)$$

现在, 我们给系统选定四个正交归一基矢量 $\{|\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\downarrow\rangle\}$ (请注意这四个基矢量的排列顺序, 后面的处理要与这个顺序一致), 以基矢量 $|\downarrow\uparrow\rangle$ 为例, 它两个标记符号的前一个 \downarrow 表示第1个电子的自旋状态, 后一个符号 \uparrow 表示第2个电子的自旋状态, 其余三个基矢量也是用类似办法进行标记的。

为了求解我们的系统, 我们需要将哈密顿算符在这四个基矢量构成的矢量基中表示出来。为了做到这一点, 让我们首先注意一个简单的数学事实, 即, 如果有归一化的量子态 $|u\rangle, |v_1\rangle, |v_2\rangle$, 并且 $|v_1\rangle$ 和 $|v_2\rangle$ 正交, 若某算符 A 将 $|u\rangle$ 映射到 $|v_1\rangle$ 和 $|v_2\rangle$ 的某个叠加态, 比如 $A|u\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle$ (α_1, α_2 为叠加系数), 那么利用 $|v_1\rangle$ 和 $|v_2\rangle$ 的正交关系可知必有 $\langle v_1|A|u\rangle = \alpha_1, \langle v_2|A|u\rangle = \alpha_2$ 。

因此为了求出哈密顿算符在我们的矢量基中的表示矩阵, 我们只需要求出它对四个基矢量的作用。为了求出这样的作用, 我们只需要反复利用上面归纳的简单公式(4.38)、(4.39)、(4.40)、以及(4.41)。比如, $H|\uparrow\uparrow\rangle = 2J(\sigma_1\sigma_2^\dagger + \sigma_1^\dagger\sigma_2)|\uparrow\uparrow\rangle + J(\sigma_1^z\sigma_2^z)|\uparrow\uparrow\rangle = J|\uparrow\uparrow\rangle$ (注意, $\sigma^\dagger|\uparrow\rangle = 0$), 即 $H|\uparrow\uparrow\rangle = J|\uparrow\uparrow\rangle$ 。类似的可以求出, $H|\downarrow\uparrow\rangle = 2J|\uparrow\downarrow\rangle - J|\downarrow\uparrow\rangle$, $H|\uparrow\downarrow\rangle = 2J|\downarrow\uparrow\rangle - J|\uparrow\downarrow\rangle$, 以及 $H|\downarrow\downarrow\rangle = J|\downarrow\downarrow\rangle$ 。因此就容易写出 H 在矢量基 $\{|\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\downarrow\rangle\}$ 中的表示矩阵 H (我们依然用 H 表示哈密顿算符的表示矩阵),

$$H = J \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4.43)$$

很容易求出这个矩阵的四个归一化本征矢量

$$\psi_{1,+1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \psi_{1,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \psi_{1,-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \psi_{0,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}. \quad (4.44)$$

对应的四个本征值分别为 $J, J, J, -3J$ 。没有学过线性代数的读者, 你只需要验证这四个本征矢量是上面给出的哈密顿矩阵 H 的本征矢量就可以了。

注意到, 在任意表象中态与表示它的列矢量的关系是 $|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$, 具体到我们这里即是

$$|\psi\rangle = \begin{bmatrix} |\uparrow\uparrow\rangle & |\downarrow\uparrow\rangle & |\uparrow\downarrow\rangle & |\downarrow\downarrow\rangle \end{bmatrix} \begin{bmatrix} \langle\uparrow\uparrow|\psi\rangle \\ \langle\downarrow\uparrow|\psi\rangle \\ \langle\uparrow\downarrow|\psi\rangle \\ \langle\downarrow\downarrow|\psi\rangle \end{bmatrix}. \quad (4.45)$$

因此, 分别以上面求出来的这四个本征矢量作为叠加系数, 相应的就有原来的哈密顿算符 H 的四个本征态

$$|1, +1\rangle = |\uparrow\uparrow\rangle, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \quad |1, -1\rangle = |\downarrow\downarrow\rangle, \quad (4.46)$$

以及

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (4.47)$$

其中前面三个本征态是简并的，对应的本征值都为 J ，这三个简并态称作两 $1/2$ 自旋耦合系统的自旋三重态，最后一个本征态 $|0,0\rangle$ 对应的本征能量是 $-3J$ ，它称作两自旋耦合系统的自旋单态。

人们很容易看到，自旋三重态的三个态有一个共同点，那就是它们对于两个电子来说是对称的，即如果把电子1和电子2进行交换，那么这三个态都将保持不变，而且后面我们将看到它们的总自旋大小是两电子贡献相加的，是1，因此我们常常称三重态为自旋平行态。相反，自旋单态关于两个电子是反对称的，如果将电子1和电子2进行交换，那相应的自旋单态就会出一个负号，而且它的总自旋两电子抵消为0了，因此我们常常称单态为自旋反平行态。

在量子力学中，一个系统的能量最低的本征态又称作系统的基态，基态的能量就是量子系统可能具有的最低能量，比基态能量更高的定态就称作激发态。基态(有时候也包括低激发态)对于研究一个系统的性质而言尤其重要，原因在于，对于一个宏观系统而言，只要环境的温度足够低，那么它就会处在基态附近，因此基态和低激发态往往决定了一个系统的宏观行为。

那么我们的两电子自旋耦合系统的基态是什么呢？很显然，答案依赖于 J 是正还是负。如果 J 大于0，那么由于自旋单态的能量为 $-3J$ ，为最低能量，因此系统的基态就将是这样一个自旋反平行态，这时我们称系统处在反铁磁相。想反，如果 $J < 0$ ，那么能量为 J 的自旋三重态就将是系统的基态，注意，这时候系统的基态是简并的， $|1,+1\rangle, |1,0\rangle, |1,-1\rangle$ 张成了一个简并子空间，任何一个给定的系统当然不可能同时处在这三个态，所以它就必须从这三个简并基态所张成的简并子空间中随机选取一个态。另一方面，从我们原来的哈密顿算符(4.36)来看，三个空间方向地位完全平等，因此我们的系统当然是空间旋转不变的。但是，这个简并子空间中的任意一个态都有一个特定的总自旋方向，比方说， $|1,+1\rangle$ 态总自旋沿着 z 轴向上， $|1,-1\rangle$ 态总自旋反着 z 轴向下，因此不管系统从简并子空间中选择了哪个态，都意味着它选定了一个特定的空间方向，因此就破坏了空间旋转不

变性, 由于这种对空间旋转不变性的破坏完全是系统自发的, 而不是因为我们在系统的哈密顿算符中引入了某个特定指向的磁场之类的东西, 所以人们常常称这种旋转对称性的自发破坏为自发破缺(注意我们不用破坏这个词, 而是用自发破缺)! 但是, 不管系统破缺到简并子空间中的哪一个态, 它都是一个自旋平行态(比如, $|1, 0\rangle$ 态的两电子也是自旋平行的, 只不过平行的方向不是在 z 轴上而已), 这时候我们就称系统处在铁磁相。

从这个例子我们可以想见, 更具一般性的海森堡模型(4.35)到底是铁磁的还是反铁磁的, 这取决于 J 的符号, 要从理论上解释某个材料的铁磁性, 那我们就必须从理论上计算出一个小于0的 J 。但这是一件相当困难的事情, 很少有人能从理论上计算出某个材料的 J 。因此人们通常采用的理论研究方案是, 从某一个对系统的更基本描述出发, 然后看看有没有一个机制能够让这个系统的有效描述可以是一个 J 为负的海森堡铁磁模型。如果理论上找到了一个这样的机制, 那么人们就会设计实验来实现这样的机制, 并看它是否能和理论推导一样使得系统处于铁磁相。总之, 即使从海森堡提出他的模型到现在已经有很长时间了, 关于铁磁性的研究依然一直是理论凝聚态的前沿领域之一, 就是因为这种研究并不容易。

两自旋耦合系统的总自旋

让我们再次回到两自旋耦合的哈密顿算符(4.36), 很显然, 我们可以用泡利算符的矢量形式将它重写成 $H = J\vec{\sigma}_1 \cdot \vec{\sigma}_2$ 。而且, 我们可以引入一个总泡利算符 $\vec{\sigma} = \vec{\sigma}_1 + \vec{\sigma}_2$, 它和系统总自旋 \vec{S} 之间的关系是 $\vec{S} = \hbar\vec{\sigma}/2$, 所以研究总泡利算符等价于研究系统总自旋。让我们先来考察一下 $(\vec{\sigma})^2 = (\vec{\sigma}_1)^2 + (\vec{\sigma}_2)^2 + 2\vec{\sigma}_1 \cdot \vec{\sigma}_2$, 而由于 $(\vec{\sigma}_1)^2 = (\sigma_1^x)^2 + (\sigma_1^y)^2 + (\sigma_1^z)^2 = 3$, 同样 $(\vec{\sigma}_2)^2 = 3$, 所以我们可以知道 $\vec{\sigma}_1 \cdot \vec{\sigma}_2 = (\vec{\sigma})^2/2 - 3$, 因此原来的哈密顿算符就可以重写成 $H = J[(\vec{\sigma})^2/2 - 3]$ 。

显然, 算符 $(\vec{\sigma})^2$ 和哈密顿算符 H 的本征态是一样的, 由前面求出来的 H 的本征值和本征态我们容易知道, 自旋三重态 $|1, +1\rangle, |1, 0\rangle, |1, -1\rangle$ 和自旋单态 $|0, 0\rangle$ 分别是 $(\vec{\sigma})^2$ 的本征值为8和本征值为0的本征态。如果用总自旋 $(\vec{S})^2 = (\vec{\sigma})^2\hbar^2/4$ 来考虑的话, 那么自旋三重态就是 $(\vec{S})^2$ 的本征值为 $2\hbar^2$ 的本征态, 通常我们说这种态的总自旋为1, 这就是三重态的记号中前一个1所代表的含义, 不过, 我们不用管为什么把这样的量子态称之为

自旋为1，我们只需要将它理解成一个约定，其含义就是，这种态对应于 $(\vec{S})^2 = 2\hbar^2$ ，也就是说，这三个态的总自旋大小是一样的，都是 $\sqrt{2}\hbar$ 。当然，用 $(\vec{S})^2$ 来考虑，自旋单态就对应于 $(\vec{S})^2 = 0$ ，也即是说，自旋单态是总自旋为0的态，这就是 $|0, 0\rangle$ 这个记号中前一个0的含义。

很明显， $(\vec{S})^2$ 的本征值不足以将两自旋耦合系统的四个本征态都区分出来，因为自旋三重态的那三个态对于 $(\vec{S})^2$ 来说是简并的。因此，人们通常进一步考察总自旋的 z 分量 $S_z = \sigma^z \hbar/2$ 。由于 $(\vec{\sigma})^2 = 6 + 2\vec{\sigma}_1 \cdot \vec{\sigma}_2 = 6 + 2(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z)$ ，我们很容易利用泡利算符的代数关系(4.18)验证 $[\sigma^z, (\vec{\sigma})^2] = 0$ ，这也即是说， S_z 和 $(\vec{S})^2$ 对易，

$$[(\vec{S})^2, S_z] = 0. \quad (4.48)$$

因此，根据上一章讲述的定理， S_z 和 $(\vec{S})^2$ 可以有共同的本征态，实际上，由于 $S_z = \frac{\hbar}{2}\sigma^z = \frac{\hbar}{2}(\sigma_1^z + \sigma_2^z)$ ，人们很容易验证： $S_z|1, +1\rangle = \hbar|1, +1\rangle$ （即 $|1, +1\rangle$ 是 S_z 的本征值为1倍 \hbar 的本征态，这就是 $|1, +1\rangle$ 记号中后一个+1的含义）， $S_z|1, 0\rangle = 0|1, 0\rangle$ （这就是本征态记号中的那个0的含义）， $S_z|1, -1\rangle = -\hbar|1, -1\rangle$ （即本征值为-1倍的 \hbar ，这就是记号 $|1, -1\rangle$ 中那个-1的含义），类似的， $S_z|0, 0\rangle = 0|0, 0\rangle$ （这就是本征态记号中第二个0的含义）。

简单归纳一下，即，自旋单态总自旋大小为0，当然 S_z 也为0，自旋三重态总自旋为1，其 S_z 的值分别为+1, 0, -1(以 \hbar 为单位)。