

# 目录

第十三章 *狄拉克方程介绍	2
13.1 狄拉克方程	3
13.1.1 狄拉克方程的引入	3
13.1.2 矩阵表示	5
13.1.3 反粒子	7
13.1.4 协变形式	8
13.2 从狄拉克方程到自旋磁矩以及LS耦合	9
13.2.1 有效哈密顿量方法回顾	9
13.2.2 泡利哈密顿量	10
13.2.3 LS耦合	11

## 第十三章 \*狄拉克方程介绍

陈童

本章讲述狄拉克方程的最基本知识，包括狄拉克方程是如何引入的，以及狄拉克是如何提出正电子的。

之后我们着重讲述了如何通过取狄拉克方程的非相对论近似得出正确的电子自旋磁矩，以及导出正确的自旋轨道耦合项。这体现了复杂的物理效应如何可以从简洁的方程中推导出来，充分体现了狄拉克方程的威力。

我们知道电子有自旋，也许你常将自旋想象成一个小球绕着自身的转轴转动。这是自旋的经典图像，然而电子的自旋是相对论量子力学的，经典的自旋图像并不能真正解释电子的自旋，比方说，电子的自旋磁矩比按照经典图像算出来的磁矩大2倍。既然经典图像不对，那么，电子的自旋是从哪儿来的呢？

电子自旋的起源，以及自旋磁矩的那个神秘的2倍都曾经让物理学家伤透了脑筋。令人意想不到的，纯粹从方程的数学之美出发，狄拉克得到了一个很简单的方程，从中自动就能得出自旋，尤其是能够得出那个神秘的2。这个方程就是著名的狄拉克方程。本章是关于狄拉克方程的一个介绍，在本章中我们取 $\hbar = 1, c = 1$ 的自然单位制。本章我们使用求和约定，即对方程中重复出现的指标默认求和。

## 13.1 狄拉克方程

### 13.1.1 狄拉克方程的引入

让我们从单粒子的薛定谔方程开始，也就是

$$i\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = H\psi(\mathbf{x}, t). \quad (13.1)$$

这个方程有两个最基本的特征，第一，这是一个线性偏微分方程，因此满足量子力学最基本的原理，态的线性叠加原理。第二，这个方程里只出现时间的一阶导数，这使得我们可以定义一个正定的守恒概率 $\int d^3\mathbf{x}|\psi(\mathbf{x}, t)|^2$ 。

薛定谔在发表这个方程时取的是非相对论形式，也就是将哈密顿算符取成 $H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$ ，其中 $\mathbf{p} = -i\nabla$ 是粒子的动量算符。哈密顿算符对应的当然是粒子的能量，然而，考虑到相对论以后，自由粒子的能量与动量之间的关系应该是

$$E^2 = \mathbf{p}^2 + m^2. \quad (13.2)$$

因此，考虑到相对论，我们就应该将自由粒子哈密顿算符取成 $H = \sqrt{\mathbf{p}^2 + m^2}$ 。这里涉及到算符 $\mathbf{p}^2 + m^2$ 的开方，一个可行的定义是利用泰勒展开，但是这一方面会使得方程变得丑陋，另一方面，在一般性的相对论情形下，算符 $\mathbf{p}^2$ 与 $m^2$ 的大小可能相当，这样泰勒展开就没有良好的数学定义。

看起来绕过这个困难的唯一办法是，放弃开方，直接使用相对论关系(13.2)，也就是说，将自由粒子薛定谔方程推广成 $-(\partial^2/\partial t^2)\psi(\mathbf{x}, t) = (\mathbf{p}^2 + m^2)\psi(\mathbf{x}, t)$ 。这样一来我们得到的就是一个关于时间的二阶微分方程。这就引起了新的问题，因为这样就破坏了原来薛定谔方程所满足的第二条特征，这使得我们再也不能定义一个守恒并且正定的概率。而一个正定的守恒概率，是整个量子力学理论的基石之一。

看起来往前往后都走不通，我们已经走投无路了！

然而，狄拉克意识到还有一条新的道路，一方面相对论当然要满足，但在另一方面方程应该仅仅只含时间的一阶导数这个特征也应该保持。由于在相对论中时间和空间可以相互转换，因此这两点结合起来的逻辑推论就是，我们应该让方程对于动量算符(即对空间的导数)也是一阶的。而要保持一阶就必须将下面的开方开出来

$$H = \sqrt{\mathbf{p}^2 + m^2}. \quad (13.3)$$

这怎么可能呢？然而狄拉克说这是可能的，只要你放弃对普通的数的坚持，转入到奇妙的反对易数世界里去，也就是像我们在讨论泡利算符时所做的那样。

现在，假定你已经将开方(13.3)开出来了，写成

$$H = \vec{\alpha} \cdot \mathbf{p} + \beta m, \quad (13.4)$$

这里的 $\vec{\alpha}$ 和 $\beta$ 就像四元数里的 $i, j, k$ 一样，是抽象的代数。为了满足相对论的关系式(13.2)，我们再将(13.4)两边平方，就得到

$$\begin{aligned} E^2 &= \alpha_i \alpha_j p_i p_j + m(\alpha_i \beta + \beta \alpha_i) p_i + m^2 \beta^2, \\ &= \frac{1}{2}(\alpha_i \alpha_j + \alpha_j \alpha_i) p_i p_j + m(\alpha_i \beta + \beta \alpha_i) p_i + m^2 \beta^2. \end{aligned} \quad (13.5)$$

式中指标 $i, j = 1, 2, 3$ 。与(13.2)比较，我们发现 $\vec{\alpha}$ ， $\beta$ 必须满足下面的抽象代数，

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{ij} \quad (13.6)$$

$$\alpha_i \beta + \beta \alpha_i = 0,$$

$$\beta^2 = 1. \quad (13.7)$$

$\beta$ 和 $\vec{\alpha}$ 的这种反对易代数很像四元数代数，两者同样奇妙，然而狄拉克得到的这个代数并不是四元数代数，而是数学家研究过的另一种称

为Clifford代数的东西。现在，利用 $\vec{\alpha}$ 和 $\beta$ 的Clifford代数，并注意到(13.4)，我们就可以将自由电子的相对论波动方程写成，

$$i\partial_t\Psi = (-i\vec{\alpha}\cdot\vec{\partial} + \beta m)\Psi, \quad (13.8)$$

这就是著名的狄拉克方程，它所描述的粒子我们称为狄拉克粒子，电子就是一个狄拉克粒子，式中 $\vec{\partial} = \nabla$ 。由于相互不对易，式中的 $\vec{\alpha}$ ,  $\beta$ 应该理解为与坐标和动量无关的算符，而波函数 $\Psi(\mathbf{x}, t)$ 关于坐标变量是一个通常的波函数，但对于算符 $\vec{\alpha}$ ,  $\beta$ 来说却应该理解成一个狄拉克右矢。

如果我们定义概率密度 $\rho(\mathbf{x}, t) = \Psi^\dagger\Psi$ ，定义概率流密度

$$\mathbf{j} = \Psi^\dagger\vec{\alpha}\Psi. \quad (13.9)$$

则利用(13.8)式容易验证下面的概率守恒方程

$$\frac{\partial\rho}{\partial t} + \vec{\partial}\cdot\mathbf{j} = 0. \quad (13.10)$$

### 13.1.2 矩阵表示

从(13.4)式容易看出，为了保证哈密顿算符的厄密性， $\vec{\alpha}$ 和 $\beta$ 都应该是厄密算符，不妨令 $\alpha_4 = \beta$ ，则对于任何 $a = 1, 2, 3, 4$ ，应该有

$$\alpha_a^\dagger = \alpha_a. \quad (13.11)$$

另外， $\vec{\alpha}$ 和 $\beta$ 的代数关系可以概括成

$$\alpha_a\alpha_b + \alpha_b\alpha_a = 2\delta_{ab}, \quad a, b = 1, 2, 3, 4. \quad (13.12)$$

即这四个厄密算符两两反对易，且每个厄密算符的平方都等于1。我们可以将这些厄密算符表示成矩阵，这一小节就是要具体解决什么样的矩阵能够表示它们？

首先，根据(13.12)的代数关系，这些厄密算符的平方都等于1，因此它们只有 $\pm 1$ 的本征值，一般来说，每个本征值都可能有多重简并，我们首先要解决+1本征值和-1本征值分别有几重简并。为此我们定义算符 $\alpha_{ij} = \alpha_i\alpha_j$ ，根据(13.12)式，我们有

$$[\alpha_1, \alpha_{12}] = 2\alpha_2, \quad [\alpha_2, \alpha_{12}] = -2\alpha_1. \quad (13.13)$$

由此我们容易有  $\text{Tr}(\alpha_1) = -\frac{1}{2}\text{Tr}([\alpha_2, \alpha_{12}]) = 0$ ,  $\text{Tr}(\alpha_2) = \frac{1}{2}\text{Tr}([\alpha_1, \alpha_{12}]) = 0$ 。类似的, 利用  $\alpha_{34}$ , 我们也可以证明  $\text{Tr}(\alpha_3) = \text{Tr}(\alpha_4) = 0$ 。概括起来即有

$$\text{Tr}(\alpha_a) = 0, \quad a = 1, 2, 3, 4. \quad (13.14)$$

因此每个  $\alpha_a$  的 +1 本征值的重数必定与 -1 本征值的重数相等, 不妨设这个重数为  $n$ , 则  $\alpha_a$  的矩阵表示必定为  $2n \times 2n$  的厄密矩阵。

最简单的情形是  $n = 1$ , 即表示为  $2 \times 2$  的厄密矩阵。我们知道, 独立的  $2 \times 2$  厄密矩阵刚好有四个, 3 个泡利矩阵, 再加上单位矩阵, 3 个泡利矩阵当然两两反对易, 但是它们都和第四个矩阵, 即单位矩阵, 对易。因此这就无法满足 4 个厄密矩阵两两反对易的要求。所以  $n = 1$  不行。

实际上, 能满足要求的最小  $n$  是  $n = 2$ , 相应的即是把  $\alpha_a$  表示成  $4 \times 4$  厄密矩阵。不妨让我们选取  $\beta = \alpha_4$  对角的表象, 取前两个分量对应 +1 本征空间, 后两个分量对应 -1 本征空间, 即把  $\beta$  表示成

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (13.15)$$

式中 1 表示  $2 \times 2$  的单位矩阵。由于  $\alpha_i$ ,  $i = 1, 2, 3$  均与  $\beta$  反对易, 所以它们的表示矩阵必定为如下形式

$$\alpha_i = \begin{pmatrix} 0 & v_i^\dagger \\ v_i & 0 \end{pmatrix}, \quad (13.16)$$

式中  $v_i$  均为  $2 \times 2$  矩阵。由于  $\alpha_i$  各自的平方等于 1, 且两两反对易, 我们容易有

$$v_i^\dagger v_j + v_j^\dagger v_i = v_i v_j^\dagger + v_j v_i^\dagger = 2\delta_{ij}. \quad (13.17)$$

很显然, 只要取  $v_i = v_i^\dagger = \sigma_i$  就能满足这些要求, 这里  $\sigma_i$  表示 3 个泡利矩阵。所以我们有

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}. \quad (13.18)$$

式(13.15)和式(13.18)合起来就构成厄密算符  $\alpha_a$  最简单的矩阵表示。在这个表示下自由粒子哈密顿算符(13.4)就是  $4 \times 4$  的算符矩阵,

$$H = \begin{pmatrix} m & \vec{\sigma} \cdot \mathbf{p} \\ \vec{\sigma} \cdot \mathbf{p} & -m \end{pmatrix}. \quad (13.19)$$

而相应的波函数 $\Psi(\mathbf{x})$ 就应该表示成4分量的所谓旋量波函数。实际上， $\alpha_a$ 的更高 $n$ 值的表示总可以约化成多个上述 $4 \times 4$ 矩阵的直和。所以我们只需考虑这种四分量表示就足够了。

### 13.1.3 反粒子

假定我们考虑一个动量确定的自由狄拉克粒子，即考察动量本征态，这时候动量算符 $\mathbf{p}$ 其实就取其本征值，我们依然记作 $\mathbf{p}$ ，所以现在 $\mathbf{p}$ 不再是算符了，而是普通的数。但是， $\alpha, \beta$ 依然是算符，所以这时候狄拉克粒子的哈密顿量(13.4)依然是算符。而由于 $H^2 = \mathbf{p}^2 + m^2$ 为普通的数，所以，这样的动量本征态同时也是 $H^2$ 的本征态。很显然，在这样的态空间上， $H = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m$ 可以取 $\pm\sqrt{\mathbf{p}^2 + m^2}$ 这两个不同的本征值。注意，这也就是说，狄拉克粒子可以处于本征值为 $-\sqrt{\mathbf{p}^2 + m^2}$ 这样的负能态。

所以狄拉克方程存在负能量态，它的能量本征值没有下界。通常来说，这是一个很大的问题，因为能量最低原理告诉我们，粒子总是更喜欢待在能量更低的态上。如果一个系统的能量本征值没有下界的话，那粒子就会持续不断地往更低的能态跃迁，这样的系统就不可能稳定存在。

为了解决这个问题，狄拉克注意到电子是一个费米子，它的每一个能态上只能占据一个电子，狄拉克想象所有的负能态都已经被电子占据满了，即是说真空其实是一片填满负能电子的海洋，称为狄拉克海。由于负能电子都被占满了，所以真空中的电子就不再能够向负能态跃迁了，它只能待在正能态上，这样就解决了稳定性问题。

奇妙的是，狄拉克进一步想象负能海中的某个电子被激发到了正能态。这时候真空中就产生了两种结果，一是多出了一个正能电子，二是负能海空出了一个未被占据的态，称之为空穴。负能海填满了电子当然带负电，但作为一种时空背景，我们无法感受到这个负能海的电荷。但是，如果负能海空出了一个态，那在我们的感受中，这个空穴就应该表现为正电荷！而且由于空穴是空出来的一个电子态，所以它在负能海中移动时其惯性应该和电子一样。即是说空穴也会表现出一个质量，并且和电子的质量一样。这就是说，负能海中的空穴在所有方面的表现都和电子一样，除了电荷相反，是正电荷。因此我们完全可以把负能海中的空穴当成是一种基本粒子，由于电荷相反，狄拉克称之为正电子。所以激发负能海的结果就是在真空中产生了一对粒子，电子和正电子。反过来，假设我们把正电子和电子放在一起，则由于正电子是空穴，所以这时候正能的电子就会向下跃迁到这

个空穴上，留下一个填满的负能海真空，同时辐射出光子。我们称这个过程为电子和正电子的湮灭。

所以，电子和正电子可以成对产生，也可以成对湮灭！正电子除了电荷相反之外，其它表现都和电子相同，通常我们称之为电子的反粒子。正电子这种反粒子的存在是狄拉克的一个伟大预言，而它很快就在实验中观测到了！不仅电子有反粒子，实际上一切微观粒子都有反粒子，不过有些粒子，比如光子，由于完全中性，不带任何荷，所以它的反粒子其实就是它本身。不过，系统地处理反粒子的问题需要用到量子场论。

### 13.1.4 协变形式

狄拉克方程更常见的写法是这样的，我们定义 $\gamma^0 = \beta$ ,  $\gamma^i = \beta\alpha_i$ , 它们组成一个四维时空的矢量 $\gamma^\mu$ ,  $\mu = 0, 1, 2, 3$ , 这样(13.8)就可以重写成

$$(i\gamma^\mu \partial_\mu - m)\Psi = 0. \quad (13.20)$$

式中 $\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_t, \vec{\partial})$ ,  $x^\mu = (t, \mathbf{x})$ 为坐标四矢量。现在 $\gamma^\mu$ 满足如下Clifford代数，

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}, \quad (13.21)$$

式中 $\eta^{\mu\nu} = \text{diag}\{1, -1, -1, -1\}$ 为四维闵可夫斯基时空的度规张量，而任何形如 $a^\mu b_\mu$ 的式子其实都可以理解成 $\eta_{\mu\nu} a^\mu b^\nu$ 。很容易看出来，如果在方程(13.20)的左边乘上 $(i\gamma^\mu \partial_\mu + m)$ , 那么得到的就是克莱因-戈登方程 $(\partial^\mu \partial_\mu + m^2)\Psi = 0$ , 因此，狄拉克方程可以看成是克莱因-戈登方程的开方。

习惯上常常定义 $\bar{\Psi} = \Psi^\dagger \gamma^0$ , 则很显然概率密度 $j^0 = \rho$ 可以重写成

$$j^0 = \Psi^\dagger \Psi = \bar{\Psi} \gamma^0 \Psi. \quad (13.22)$$

而概率流密度则可以重写成

$$\mathbf{j} = \Psi^\dagger \vec{\alpha} \Psi = \bar{\Psi} \vec{\gamma} \Psi. \quad (13.23)$$

概率密度和概率流密度合起来构成一个洛伦兹四矢量 $j^\mu$ ,

$$j^\mu = \bar{\Psi} \gamma^\mu \Psi. \quad (13.24)$$

而概率守恒方程则可以写成如下协变形式

$$\partial_\mu j^\mu = 0. \quad (13.25)$$



注意,  $\gamma^0$ 是厄密的, 而 $\gamma^i$ 则反厄密, 即 $\vec{\gamma}^\dagger = (\beta\vec{\alpha})^\dagger = \vec{\alpha}^\dagger\beta^\dagger = \vec{\alpha}\beta = -\beta\vec{\alpha} = -\vec{\gamma}$ 。关于这些gamma矩阵, 除了上一段所用的表象之外, 还有一个常用的表象是取

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}. \quad (13.26)$$

读者很容易验证这样的选取方式的确满足代数关系(13.21)。为了看清楚这是一个什么表象, 我们不妨计算 $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$ , 很容易得到

$$\gamma^5 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad (13.27)$$

式中 $i$ 表示 $2 \times 2$ 的常数矩阵。很显然, 这个表象是 $\gamma^5$ 的对角表象。 $\gamma^5$ 是一个反厄密算符, 其平方为 $-1$ ,  $\pm i$ 是它的本征值。我们可以将四分量的旋量波函数按照 $\gamma^5$ 的本征值进行分解, 对应本征值 $i$ 的旋量叫左手旋量, 对应本征值 $-i$ 的旋量叫右手旋量。

## 13.2 从狄拉克方程到自旋磁矩以及LS耦合

这一节我们将通过取狄拉克方程的非相对论近似导出正确的电子自旋磁矩, 以及正确的自旋轨道耦合。为此, 请读者回想一下我们在第三章第3.4节中的有关讨论。在那里我们看到, 从自旋1/2带电粒子的泡利哈密顿量出发我们能得到正确的电子自旋磁矩, 但问题是, 为什么有泡利哈密顿量呢? 本节我们将看到, 泡利哈密顿量其实是狄拉克方程的非相对论近似结果。不仅如此, 我们还将看到, 进一步研究这个非相对论近似还能导出正确的自旋轨道耦合。

### 13.2.1 有效哈密顿量方法回顾

为了讲清楚狄拉克方程的非相对论近似, 我们先来回顾一下第五章第5.2节中讨论过的有效哈密顿量方法。根据5.2节的讨论, 如果我们能将矩阵按照1和2两个子空间进行正交分解, 则对于一个如下算符矩阵形式的哈密顿量 $H$ ,

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \quad (13.28)$$

(式中子矩阵 $H_{11}, H_{12}, H_{21}, H_{22}$ 同时也是算符), 我们可以将其本征矢量 $\Psi$ 按照两个子空间分解成

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \quad (13.29)$$

通过消去子空间2里的分量 $\psi_2$ , 我们可以得到子空间1里的有效哈密顿量 $H_{\text{eff}}(E)$

$$H_{\text{eff}}(E) = H_{11} + H_{12} \frac{1}{E - H_{22}} H_{21}, \quad (13.30)$$

并将原来的本征方程 $H\Psi = E\Psi$ 转换成子空间1里的有效本征方程

$$H_{\text{eff}}(E)\psi_1 = E\psi_1. \quad (13.31)$$

由于狄拉克方程中 $\beta$ 乘在静止能量 $m$ 上, 它的本征值反映的是粒子的静止能量本征值, 而非相对论极限下静止能量远大于动能。因此下面我们的基本思路就是取 $\beta$ 的本征值为1的本征空间为子空间1, 本征值为 $-1$ 的本征空间为子空间2, 然后将狄拉克方程的哈密顿算符分解成(13.19)这种形式, 将四分量狄拉克旋量相应分解成如下形式

$$\Psi = \begin{pmatrix} u \\ v \end{pmatrix}, \quad (13.32)$$

$u, v$ 均为两分量旋量波函数。然后再通过消去静止能量本征值为负的 $v(\mathbf{x})$ 分量推导出通常的正静止能量分量 $u(\mathbf{x})$ 所满足的有效本征方程。

### 13.2.2 泡利哈密顿量

为了得到电子磁矩, 我们需要加上磁场, 然后考察电子自旋与磁场的耦合。电子与磁场的耦合可以通过将动量算符 $\mathbf{p}$ 替换成力学动量算符 $\vec{\pi} = \mathbf{p} - q\mathbf{A}$ 来得到, 这里 $\mathbf{A}$ 为电磁场矢量势,  $q = -e$ 为电子电量。这样替换以后, 相对论电子的哈密顿量(13.19)就变成了,

$$H = \begin{pmatrix} m & \vec{\sigma} \cdot \vec{\pi} \\ \vec{\sigma} \cdot \vec{\pi} & -m \end{pmatrix}. \quad (13.33)$$

按照(13.30)式,  $\beta = 1$ 的本征子空间里的有效哈密顿量 $H_{\text{eff}}(E)$ 为

$$H_{\text{eff}}(E) = m + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{E + m}. \quad (13.34)$$

相应的有效本征方程为

$$\left[ m + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{E + m} \right] u(\mathbf{x}) = Eu(\mathbf{x}). \quad (13.35)$$

在非相对论近似下, 能量本征值的零级近似为静止能量 $m$ , 记相应的非相对论修正为 $\varepsilon$ , 将能量的零级近似 $m$ 代入方程(13.35)左边, 则右边就可以给出非相对论修正 $E = m + \varepsilon$ , 很显然

$$\left[ \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} \right] u(\mathbf{x}) = \varepsilon u(\mathbf{x}). \quad (13.36)$$

因此

$$\frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} = \frac{[\vec{\sigma} \cdot (\mathbf{p} - q\mathbf{A})]^2}{2m} \quad (13.37)$$

就是电子在磁场中的非相对论哈密顿量。很显然, 这正是泡利给出来的泡利哈密顿量, 前面第三章3.4节中我们已经看到, 这个哈密顿量能够给出正确的电子自旋磁矩。所以, 对电子自旋磁矩那个2倍因子的正确解释最终需要用到狄拉克方程, 它是狄拉克方程自动的一个结果。

### 13.2.3 LS耦合

下面我们进一步假设有一个电势 $\phi(\mathbf{x})$ 作用在电子上, 则相对论电子的哈密顿量就应该写成

$$H = \begin{pmatrix} m & \vec{\sigma} \cdot \vec{\pi} \\ \vec{\sigma} \cdot \vec{\pi} & -m \end{pmatrix} + q\phi(\mathbf{x}) \cdot 1 = \begin{pmatrix} m + q\phi & \vec{\sigma} \cdot \vec{\pi} \\ \vec{\sigma} \cdot \vec{\pi} & -m + q\phi \end{pmatrix}. \quad (13.38)$$

式中 $q\phi$ 为电势能。则 $\beta = 1$ 的本征子空间里的有效哈密顿量 $H_{\text{eff}}(E)$ 现在就应该为

$$H_{\text{eff}}(E) = m + (\vec{\sigma} \cdot \vec{\pi}) \frac{1}{E + m - q\phi} (\vec{\sigma} \cdot \vec{\pi}). \quad (13.39)$$

相应的有效本征方程就是

$$\left[ m + (\vec{\sigma} \cdot \vec{\pi}) \frac{1}{E + m - q\phi} (\vec{\sigma} \cdot \vec{\pi}) \right] u(\mathbf{x}) = Eu(\mathbf{x}). \quad (13.40)$$

为了取非相对论近似, 令 $E = m + \varepsilon$ , 对静止能量的修正 $\varepsilon$ 来自于动量 $\vec{\pi}$ 和电势能 $q\phi$ 。注意到

$$\frac{1}{E + m - q\phi} = \frac{1}{2m(1 + \frac{\varepsilon - q\phi}{2m})} \simeq \frac{1}{2m} \left( 1 - \frac{\varepsilon - q\phi}{2m} \right), \quad (13.41)$$

代入上面的(13.40)式, 并将所有的 $\varepsilon$ 项移到方程右边, 则有

$$\left[ q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{q}{4m^2}(\vec{\sigma} \cdot \vec{\pi})\phi(\vec{\sigma} \cdot \vec{\pi}) \right] u(\mathbf{x}) = \varepsilon \left[ 1 + \frac{1}{4m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] u(\mathbf{x}) \quad (13.42)$$

我们想把修正能量 $\varepsilon$ 看成是某个厄密哈密顿量 $H'$ 的本征值, 但是上面这个方程显然不具有这种定态薛定谔方程的形式, 为了将之改写成这种形式, 我们重新定义一个两分量旋量波函数 $\psi(\mathbf{x})$ ,

$$\psi(\mathbf{x}) = \left[ 1 + \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] u(\mathbf{x}). \quad (13.43)$$

并利用(忽略 $\vec{\pi}/m$ 的四次方阶)

$$\begin{aligned} \left[ 1 + \frac{1}{4m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] &\simeq \left[ 1 + \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right]^2 \\ \left[ 1 - \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] \left[ 1 + \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] &\simeq 1. \end{aligned} \quad (13.44)$$

则我们就能把方程(13.42)改写成标准的

$$H'\psi(\mathbf{x}) = \varepsilon\psi(\mathbf{x}), \quad (13.45)$$

式中厄密哈密顿量 $H'$ 为

$$H' \simeq \left[ 1 - \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right] \left[ q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{q}{4m^2}(\vec{\sigma} \cdot \vec{\pi})\phi(\vec{\sigma} \cdot \vec{\pi}) \right] \left[ 1 - \frac{1}{8m^2}(\vec{\sigma} \cdot \vec{\pi})^2 \right].$$

展开上面 $H'$ 的表达式就能得到

$$H' = q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{q}{8m^2}(\vec{\sigma} \cdot \vec{\pi})[\phi, \vec{\sigma} \cdot \vec{\pi}] + \frac{q}{8m^2}[\vec{\sigma} \cdot \vec{\pi}, \phi](\vec{\sigma} \cdot \vec{\pi}) + \dots \quad (13.46)$$

省略号表示 $\vec{\pi}/m$ 的四次方项和更高次项。利用

$$[\vec{\pi}, \phi] = [\mathbf{p}, \phi] = -i\nabla\phi, \quad (13.47)$$

可以进一步得到

$$H' = q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{iq}{8m^2} \left[ (\vec{\sigma} \cdot \vec{\pi})(\vec{\sigma} \cdot \nabla\phi) - (\vec{\sigma} \cdot \nabla\phi)(\vec{\sigma} \cdot \vec{\pi}) \right] + \dots \quad (13.48)$$

利用恒等式

$$(\vec{\sigma} \cdot \vec{F})(\vec{\sigma} \cdot \vec{G}) = (\vec{F} \cdot \vec{G}) + i\vec{\sigma} \cdot (\vec{F} \times \vec{G}), \quad (13.49)$$

又可以得到(注意 $(\nabla\phi) \times \vec{\pi} = -\vec{\pi} \times (\nabla\phi)$ )

$$\begin{aligned} H' &= q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{iq}{8m^2} [\vec{\pi} \cdot (\nabla\phi) - (\nabla\phi) \cdot \vec{\pi}] - \frac{q}{4m^2} [\vec{\pi} \times (\nabla\phi)] \cdot \vec{\sigma} + \dots \\ &= q\phi + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{q}{8m^2} \nabla^2\phi - \frac{q}{4m^2} [\vec{\pi} \times (\nabla\phi)] \cdot \vec{\sigma} + \dots \end{aligned} \quad (13.50)$$

或者也可以记 $q\phi(\mathbf{x}) = V(\mathbf{x})$ 为电势能, 记 $\vec{\sigma}/2 = \mathbf{S}$ 为自旋算符, 进而将 $H'$ 写成

$$H' = V(\mathbf{x}) + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{1}{2m^2} [(\nabla V) \times \vec{\pi}] \cdot \mathbf{S} + \frac{1}{8m^2} \nabla^2 V + \dots \quad (13.51)$$

通过量纲分析恢复常数 $\hbar$ 和 $c$ , 则(13.51)就应该是

$$H' = V(\mathbf{x}) + \frac{(\vec{\sigma} \cdot \vec{\pi})^2}{2m} + \frac{1}{2m^2 c^2} [(\nabla V) \times \vec{\pi}] \cdot \mathbf{S} + \frac{\hbar^2}{8m^2 c^2} \nabla^2 V + \dots \quad (13.52)$$

相比于泡利哈密顿量, (13.52)右边多出两项, 其中 $\frac{\hbar^2}{8m^2 c^2} \nabla^2 V$ 是所谓的达尔文振颤项, 它来自于电子不是经典粒子, 它的位置会在康普顿波长 $\hbar/(mc)$ 的范围内随机振颤, 达尔文项就是这种振颤导致的能量修正。而(13.52)右边的

$$\frac{1}{2m^2 c^2} [(\nabla V) \times \vec{\pi}] \cdot \mathbf{S} \quad (13.53)$$

描写的就是自旋轨道耦合。特别的, 假设忽略磁场, 从而 $\vec{\pi} = \mathbf{p}$ , 并且考察中心力场, 从而 $V$ 仅仅依赖于径向半径 $r$ , 即有 $\nabla V = \mathbf{e}_r \frac{\partial V}{\partial r} = \frac{\mathbf{x}}{r} \frac{\partial V}{\partial r}$ , 则这一自旋轨道耦合项就可以重写成

$$\frac{1}{2m^2 c^2} \left[ \frac{\mathbf{x}}{r} \frac{\partial V}{\partial r} \times \mathbf{p} \right] \cdot \mathbf{S} = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{\partial V}{\partial r} \mathbf{L} \cdot \mathbf{S}, \quad (13.54)$$

式中 $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ 是电子的轨道角动量算符。对于氢原子 $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$ , 从而容易得到相应的自旋轨道耦合项为

$$\frac{e^2}{8\pi\epsilon_0} \frac{1}{m^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S}. \quad (13.55)$$

之所以有这样的自旋轨道耦合项, 是因为从电子的角度来看, 是氢原子核在围绕着它转动, 这种转动会在电子位置产生一个磁场, 这个磁场正比于轨道角动量 $\mathbf{L}$ , 电子的自旋磁矩会和这个磁场耦合, 这就产生了自旋轨道耦合, 简言之, 这一项来源于电子自旋磁矩与它感受到的磁场的耦合。

但是，想从这一物理图像得到系数完全正确的自旋轨道耦合项，人们必须考虑到一个相对论效应，即所谓的Thomas进动，否则结果会大了两倍。狄拉克方程由于本身是相对论的，将它取非相对论近似时，Thomas进动的相对论效应自动就包含进来了，所以我们上面才能导出正确的自旋轨道耦合。