

相对论性量子场论

吴晋渊，郭家祺

2021 年 12 月 28 日

如无特殊说明，本文所谓的本征态指的都是归一化。希腊字母的指标跑遍所有时空维度，而拉丁字母的指标仅仅跑遍空间维度，也就是 $\mu, \nu, \dots = 0, 1, 2, 3$ 而 $i, j, \dots = 1, 2, 3$ 。常规斜体字母 x, y, p 等若经说明为多分量对象，默认为四维矢量，相应的， $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{p}$ 等为它们的空间部分。常规斜体字母的点乘表示四维矢量乘法，如

$$p^2 = \omega^2 - |\boldsymbol{p}|^2, \quad x \cdot y = x^0 y^0 - \boldsymbol{x} \cdot \boldsymbol{y} = t_x t_y - \boldsymbol{x} \cdot \boldsymbol{y}.$$

指标 a, b, \dots 也有可能指各种多分量对象的指标，未必正好取 $1, 2, 3$ 。 T 表示编时算符， N 表示正规序算符。

张量的分量矩阵默认以排在前面的指标为行指标，以排在后面的指标为列指标。

本文将略去所有最为基础的formalism：量子力学的基本框架——包括路径积分和正则形式，无论是关于单粒子的还是关于场的；如何将多粒子态写成场的激发态（即所谓二次量子化）；微扰论的形式理论等等。这些东西可以在任何好的量子场论教科书中找到。

第一部分

相对论性量子场论

第一章 量子场论的形式理论框架

物理实验表明存在一个时间轴，且存在粒子，它的自由度在不考虑引力导致的空间翘曲时包括一个三维欧氏空间，可能还有一些内禀自由度。狭义相对论还告诉我们，移动参考系，时间和空间的变换是洛伦兹变换，这对应一个3 + 1维闵可夫斯基时空，即度规可以化为

$$\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$$

的四维几何。通常使用 t, x, y, z 或者 x^0, x^1, x^2, x^3 来依次标记这4个坐标。容易看出 x, y, z 或者说 x^1, x^2, x^3 就构成一个三维欧氏几何，它们是空间维。 x^0 则是时间维。

在狭义相对论中难以使用哈密顿动力学，因为此时“时间”的概念是不清楚的：应该使用固有时还是坐标时？因此，简单地将单粒子量子力学移植到相对论时空中不是好的选择。除此之外单粒子量子力学还存在一个很大的问题。为了让粒子有非平凡的运动，通常会往粒子的哈密顿量中引入一个势能，然而势能必须由另一些东西提供。如果这个“东西”的参数发生了变化，似乎应该认为势能发生了瞬时的变化，但这样就有信号的瞬时传递了。这和狭义相对论当然是矛盾的。

场论是解决这个问题的一种方案。就刚才势场的问题，我们知道，实际上能够看到的大部分势场都是电磁相互作用产生的，的确，通过推迟势能够解决信号瞬时传输的问题，但是更加自然的做法是引入电磁场的概念：物质激发出电磁场，电磁场反作用于物质。大量的场论能够很自然地服从狭义相对论——例如，电动力学在洛伦兹变换下不变。因此，将已知在相对论情况下运作良好的场论量子化似乎是一个相当吸引人的做法。

这个想法导致了我們称为量子场论的一整套理论。在比较系统性的教科书上，我们会看到以下形式理论：

1. 给定一个场论拉氏量 \mathcal{L} 和它做勒让德变换得到的哈密顿量密度 \mathcal{H} 。我们下面将单粒子量子力学中使用过的方案移植到场论中。识别出场 ϕ 和它的正则动量 π 之后，我们施加正则对易关系

$$[\phi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}), \quad (1.1)$$

或者，为了后面我们将会看到的导出费米子的目的，施加反对易关系

$$\{\phi(\mathbf{x}, t), \pi(\mathbf{y}, t)\} = i\delta^{(3)}(\mathbf{x} - \mathbf{y}). \quad (1.2)$$

在这里， ϕ 的地位相当于量子力学中的 x ，而 \mathbf{x} 则不是量子力学中的算符，而是区分不同的场自由度的标签，对应不同坐标 x_i 的标签 i 。在这个步骤之后我们其实已经得到了最为广义的“量子场论”：场的量子力学。

2. 场算符的线性叠加自然给出了一组满足产生湮灭算符对易关系的 a 和 a^\dagger ；通常它们就是动量表象下的 $a_{\mathbf{p}\sigma}$ 和 $a_{\mathbf{p}\sigma}^\dagger$ ，其中 σ 是另一些标签，比如说标记不同场分量（“偏振”）

的标签。于是，一个二次型的哈密顿量用这一组产生湮灭算符写出来就形如

$$H \propto \sum_{\mathbf{p}, \sigma} \omega_{\mathbf{p}\sigma} a_{\mathbf{p}\sigma}^\dagger a_{\mathbf{p}\sigma}. \quad (1.3)$$

这正是谐振子的哈密顿量。经典场论中谐振子可以连续变化，而量子的谐振子的能级则是分立的；此时哈密顿量的本征态可以用 $\{n_{\mathbf{p}\sigma}\}$ 标记。我们认为这些量实际上就是粒子数——实际上，容易验证一个单粒子态服从的薛定谔方程正是将 \mathcal{L} 对应的波动方程中的场诠释为波函数而得到的方程。于是我们发现，在二次型哈密顿量下（称为自由场），场的希尔伯特空间实际上是多粒子态的Fock空间，产生湮灭算符的标签——如动量、偏振等——正好也是单粒子的标签。

3. 以上我们从场得到了多粒子Fock空间，当然也可以从多粒子Fock空间构造产生湮灭算符而得到场。
4. 加入相互作用之后，可能出现粒子数生灭或者。因此我们发现“场的量子力学”实际上就是满足“场的哈密顿量是谐振子哈密顿量加上相互作用项之后得到的，场是粒子的产生湮灭算符，从而多粒子态可以看成场的激发态”的一个理论，也即，一种非常自然的量子多体理论。这是“量子场论”一词的更加通常的含义：通过场展开讨论的关于多粒子的量子多体理论。
5. 最后，我们可以使用量子力学中就已经做过的相干态路径积分方法，

实际上，通常所说的“量子场论”的意义更加局限，指的是满足相对论协变性的量子场论，即高能物理的量子场论；物理中如果需要分析非相对论协变性的量子场论，一般都是在分析凝聚态系统，此时基本的场就是电子场、声子场而相互作用是库仑相互作用，称为凝聚态场论。凝聚态场论实际上已经非常复杂；在高能物理的量子场论中我们通常只分析少数几个粒子的散射过程，基态为真空态，而凝聚态场论中讨论含有大量粒子的基态、讨论束缚态问题都是非常常见的。另一方面，凝聚态场论不受到相对论协变性约束，因此其中有更多可能性。虽说低速下的时空对称性是伽利略对称性，但是凝聚态场论很多是伽利略对称性也不遵守的，例如晶格的存在本身就破坏了连续平移对称性和旋转对称性。因此，很多时候高能物理的量子场论实际上反而更加简单，虽然凝聚态场论背后的基本物理机制实际上只有库仑相互作用。实际上，一些人甚至认为依靠凝聚态物理中复杂的演化行为能够完全复刻出整个高能物理。

以上形式理论看起来非常优美：它将正则量子化和路径积分量子化都推广到了一般人能够想象的最广义的程度，成功统一了看起来完全不相干的粒子和场的理论。然而我们必须承认这种优美实际上是一种假象。具体来说，以上形式理论面临如下挑战：

- 首先，它难以数学上良定义。“场自由度”显然是无穷维的数学对象，一般都极其难以处理。此外，一般的量子场论——如QED或是QCD——中普遍存在的发散意味着实际上如上形式理论允许我们定义紫外不完备的理论，即依照它自身的结构就不能够适用于任意高能量的理论。因此，如果我们要将以上最为一般的形式理论数学化，重整化、有效理论、紫外截断等信息必须被纳入其中。不言而喻，这非常具有挑战性。
- 应当指出对场作用正则量子化是一个物理意义非常不明确的操作（这也是很多初学者无法接受量子场论的原因：完全无法理解将场量子化的动机）。在单粒子量子力学中，将物理量提升为算符，施加正则对易关系的目的是非常明确的：我们需要一个允许物

理量不确定的理论，但是它又不是简单的经典概率理论，那么很快会发现单粒子量子力学是一种很好的备选理论；正则量子化之后粒子的波动性与被实验非常好地展现的经验事实完全一致。问题是，在量子场论中，我们几乎从来不讨论“波泛函” $\langle\phi|\Psi\rangle$ ——我们从来只考虑粒子的行为。因此，量子场论中的场似乎只是起了一个辅助作用：我们用它来帮助构造一个多体理论而已。

如果事情只是这样，那么还不算太糟糕。我们可以将“场的量子化”当成一个陈旧的历史术语，把它当成施加产生湮灭算符对易或反对易关系的一个简便写法。然而应当注意，量子场论取得的所有成功全部是关于粒子散射实验和 S 矩阵的——看着拉氏量写下费曼规则计算 S 矩阵的步骤是确定无疑非常有用的，即使它们可能未必是终极理论的形式理论。现在问题来了：从“场的量子力学”——以下将 \mathcal{L} 对应的场的量子力学记作 $\text{QMFT}(\mathcal{L})$ ——能够推导出“看着拉氏量写下 S 矩阵”需要的各种规则——以下记作 $\text{SMat}(\mathcal{L})$ ，但是真的被实验证实可靠的只有 $\text{SMat}(\mathcal{L})$ ，而它实际上不需要 $\text{QMFT}(\mathcal{L})$ 就能够被定义。除了 $\text{SMat}(\mathcal{L})$ 以外的 $\text{QMFT}(\mathcal{L})$ 的形式理论反而变成了累赘——甚至于“ ϕ 和 π 构成的哈密顿量”是否真的有物理意义我们都是不知道的。

- 例如，我们实际上不知道 $\text{QMFT}(\mathcal{L})$ 的束缚态预言是不是和实验符合。量子力学能够正确处理束缚态，但是没有人能够完整地证明，比如说， $\text{QMFT}(\mathcal{L}_{\text{QED}})$ 和“彼此库伦排斥的电子”的束缚态量子力学理论是一致的。我们只能证明 $\text{SMat}(\mathcal{L}_{\text{QED}})$ 中电子-电子散射的概率振幅在非相对论极限下和库伦散射一致，从而，勉强算是（说“勉强”见下一条疑难）证明了 $\text{QMFT}(\mathcal{L}_{\text{QED}})$ 的散射态行为和库伦散射一致，但是它的束缚态部分到底如何是不好说的。按理说，两个量子理论有一样的束缚态行为和不同的散射态行为是完全可能的事情，所以 $\text{QMFT}(\mathcal{L}_{\text{QED}})$ 对束缚态电子的预言居然不对——或者它对腔体中的光子的预言居然不对——并不是一个可以轻易排除的可能。
- 最后，实际上 $\text{QMFT}(\mathcal{L})$ 和 $\text{SMat}(\mathcal{L})$ 的散射态是不是能够说一样其实也成问题。我们知道一个一般的场论 \mathcal{L} 的 $\text{SMat}(\mathcal{L})$ 给出的微扰级数一般来说是不收敛的，是一个渐进级数，我们需要按照一定的准则来决定保留前几阶；然而如果 $\text{QMFT}(\mathcal{L})$ 是良定义的，至少我们得想出一种自圆其说的办法让整个微扰级数都有明确意义。

基于以上原因，最好还是将 $\text{QMFT}(\mathcal{L})$ 当成一种推导 $\text{SMat}(\mathcal{L})$ 的启发式方法，而不是什么巍峨堂皇的理论体系。毕竟严格来说既然我们对其束缚态行为知之甚少，眼下它甚至不能用于从QED推导出凝聚态场论和量子力学。它的高能物理中的唯一作用就是让人们能够接受 SMat 的规则。另一方面，凝聚态场论倒的确是定义良好的一个 QMFT 意义下的场论：直接将多电子哈密顿量用二次量子化形式理论写出即可：

$$H = \int d^3\mathbf{r} \psi_{\sigma}^{\dagger} \left(-\frac{\nabla^2}{2m} + V_{\text{ion}}(\mathbf{r}) \right) \psi_{\sigma} + \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \psi_{\sigma'}^{\dagger}(\mathbf{r}') \psi_{\sigma}^{\dagger}(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_{\sigma}(\mathbf{r}) \psi_{\sigma'}(\mathbf{r}), \quad (1.4)$$

但是凝聚态场论中 SMat 没有太大意义；因此，直接从凝聚态场论的第一性原理哈密顿量其实计算不出太多东西。

第二章 从对称性出发得到的物理

2.1 闵可夫斯基时空和狭义相对论回顾

基本上整个高能物理中的场论都以四维闵可夫斯基时空为底流形。四维闵可夫斯基时空有三个空间维，一个时间维。

2.1.1 闵可夫斯基时空的导出

光速不变原理的推论

我们假定存在一些参考系，在其中不受外力的粒子做匀速直线运动，并将这样的参考系称为惯性参考系。惯性参考系的存在性是直觉上非常合理的。考虑两个惯性参考系 S 和 S' ，分别用 (t, x, y, z) 和 (t', x', y', z') 标记其中的事件。我们假定有某种运动的速度在不同的惯性参考系中都是一样的，这个速度称为光速，这个假定称为光速不变原理。“光速”一词来自波动方程中的光速在参考系变换下不变这一事实，不过在这里我们暂时将它作为一个无意义的名词使用；实际上，狭义相对论中光速的作用在于在时间和空间之间建立关系。

设一个信号¹以光速运动，它某时刻在某地出现，记作事件 P_1 ，另一时刻在另一地点出现，记作事件 P_2 。于是我们有

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = c^2(t_1 - t_2)^2,$$

以及

$$(x'_1 - x'_2)^2 + (y'_1 - y'_2)^2 + (z'_1 - z'_2)^2 = c^2(t'_1 - t'_2)^2.$$

在这里，光速不变性的不同寻常之处已经可以体现出来了：设 S' 系相对 S 以速度 u 沿着 x 轴运动，则

$$x'_1 = x_1 + ut_1, \quad x'_2 = x_2 + ut_2,$$

于是

$$(x_1 - x_2 + ut_1 - ut_2)^2 - (x_1 - x_2)^2 = c^2(t'_1 - t'_2)^2 - c^2(t_1 - t_2)^2,$$

因此

$$(t'_1 - t'_2)^2 \neq (t_1 - t_2)^2.$$

这意味着两个参考系中时间流动的速率不一样。因此满足光速不变性的参考系变换肯定不是伽利略变换。

¹什么是“信号”其实是需要手动指定的。可以想象非常远的距离的两端依次发生了两个事件，但这并不代表有什么东西从这段距离的一端传到了另一端。一个粒子显然是一个信号，空间中的场的可以连续传播的构型的“位置”或许也可以看成一个信号。

简而言之，如果一个信号以光速运动，那么在任何参考系中都有

$$c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2 = 0.$$

反之，如果某个参考系中有上式成立，那么信号一定以光速运动。于是我们定义两个事件的间隔为

$$s^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2, \quad (2.1)$$

并且如果一个参考系中的间隔为零，那么别的参考系中的间隔也为零。

现在我们考虑靠得很近的两个事件，即考虑 ds^2 。坐标变换是连续的，即 ds^2 和 ds'^2 是同阶小量，并且 $ds^2 = 0$ 时 $ds'^2 = 0$ 。因此我们设

$$ds^2 = a ds'^2.$$

现在再引入一个惯性参考系 S'' 。由空间的均匀性和各向同性， a 不应该显含任何坐标，无论是时间还是空间，因此它只应该依赖于两个惯性参考系之间的相对速度（我们不能预期 S 在 S' 中的运动速度和 S' 在 S 中的运动速度只差一个负号，但是它们之间显然是有关系的），而且不能依赖于相对速度的方向。设 S' 在 S 中的运动速度大小为 V_1 ， S'' 在 S 中的运动速度大小为 V_2 ， S'' 在 S' 中的运动速度为 V_3 ，则显然

$$ds'^2 = a(V_1) ds^2, \quad ds''^2 = a(V_2) ds^2, \quad ds''^2 = a(V_3) ds'^2,$$

于是

$$a(V_3) = \frac{a(V_1)}{a(V_2)}.$$

但是 V_3 不仅依赖于 V_1 和 V_2 ，肯定还依赖于它们的相对角度，而上式右边却没有出现任何相对角度，因此仅有的可能是 a 根本就是一个常数，代入上式就有 $a = 1$ ，于是

$$ds^2 = ds'^2, \quad (2.2)$$

即时空间隔在不同的惯性参考系中完全一样。在重新定义时间的单位，从而用 t 代替 ct 之后，不同惯性参考系实际上成了一个四维闵可夫斯基时空的坐标系，惯性参考系之间的坐标变换就是参考系变换，需要满足的条件就是间隔微元相等。

不同惯性系之间相对做匀速直线运动，即 (x', y', z') 可以写成 (x, y, z) 加上某个三维矢量乘以 t ，或许加上一个常数，即从 (t, x, y, z) 到 (x', y', z') 的变换是仿射的。代入间隔微元相等的条件中可以发现到 t' 的变换也是仿射的。反之，如果两个参考系之间的坐标变换是仿射的，且其中一个是惯性系，那么另一个当然也是惯性系。因此，惯性系之间的坐标变换总是仿射的，并且保持时空间隔不变，即为等度规变换。惯性系之间的坐标变换在忽略了平移之后就是一个矩阵群，称为洛伦兹变换。

本文将一个任意的洛伦兹变换记作 Λ_ν^μ ，实际上应该记作 $\Lambda^\mu{}_\nu$ 。

2.1.2 闵可夫斯基时空中的几何对象

四维矢量的分量

本节暂时不区分时间和空间的单位，具体的讨论见第2.2.1节。为了避免不必要的麻烦，我们在直角坐标系加时间维下讨论问题。设时间维为第0维，闵可夫斯基时空的度规为

$$g_{\mu\nu} = g^{\mu\nu} = \text{diag}(1, -1, -1, -1). \quad (2.3)$$

对一个四维矢量 A^μ ，我们设

$$A^\mu = (A^0, \mathbf{A}), \quad (2.4)$$

则由指标升降关系自然得到

$$A_\mu = (A_0, -\mathbf{A}), \quad A_0 = A^0. \quad (2.5)$$

虽然梯度算符的行为和矢量非常相似，由于按照定义

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad \partial^\mu = \frac{\partial}{\partial x_\mu},$$

而坐标显然是货真价实的矢量，我们有

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = (\partial^0, \partial^i) = (\partial_t, -\nabla), \quad \partial_t = \partial^t, \quad (2.6)$$

对应的

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_t, \nabla). \quad (2.7)$$

这里有一个看起来比较奇怪的地方，就是

$$[\partial^i]_i = -\nabla,$$

但是上式中的 ∂^i 算符定义在闵可夫斯基时空中，而闵可夫斯基时空中空间维的度规为 -1 。如果 ∂^i 是指三维欧氏空间中的那个梯度，那么就有

$$[\partial^i]_i = \nabla,$$

也就是说闵可夫斯基时空和欧氏空间中的空间梯度算符差一个负号。为了避免混乱，之后我们将 ∂^i 局限为欧氏空间中的梯度算符。

闵可夫斯基时空版本的拉普拉斯算符——也就是达朗贝尔算符——定义为

$$\square^2 = \partial_\mu \partial^\mu = \partial_t^2 - \nabla^2. \quad (2.8)$$

2.2 单位制，指标记号和度规选取

2.2.1 单位制

到现在为止我们的方程都是自然单位制的。实际计算时会使用一些不同的单位制，会让方程变得复杂一些。

2.2.2 度规的圣战

时间在整个四维矢量中的位置，一些人取为 x^0 ，一些人取为 x^4 。

— + + + 和 + - - -

在以上所有的讨论中，我们都在使用度规+ - - -。度规- + + +在一些文献中也是常用的，并且在做Wick转动时更加方便（见下一节）。

两种度规的协变矢量保持一致，点乘差一个负号。

Wick转动

为了免去洛伦兹度规的麻烦，一些人会做Wick转动，即令 $\tau = it$ ，这样就不需要区分逆变协变了，并且很多积分的性质会变得良好。为了尽可能减少需要改动的地方，最方便的做法是在Wick转动后的理论（“欧氏空间度规理论”）和 $-++++$ 度规之间切换，需要时再切换到 $+- - -$ 度规。

如果单纯是做替换 $\tau = it$ ，那么没有太多可说，但实际上其它物理量也需要变化，而这些变化有很多自由发挥的空间。例如，原本在相对论性量子场论中含有傅里叶变换的表达式在做完Wick转动之后是否需要修改为拉普拉斯变换？格林函数的定义是否需要更动？这些都是需要指定的。

我们于是施加以下条件：

- Wick转动后的理论，如果将时间的积分区域设为实数，应该给出原理论对应的有限温度场论。这就是说，对应关系 $\tau \leftrightarrow it$ 和 $i\omega_n \leftrightarrow \omega$ 应当可以在Wick转动的过程中找到，虽然由于 t 和 ω 区分逆变和协变，有待进一步澄清以上关系式中的 t 和 ω 指的是什么。
- Wick转动前后标量尽可能不变。
- Wick转动将原本的傅里叶变换映射为 $e^{ik_i x_i}$ 形式的欧氏空间傅里叶变换。
- 闵氏时空下，矢量的欧氏部分的协变分量和欧氏四维空间中对应的分量需要完全一样，从而我们无需在 $(x^M)^i$ 和 x_i^E 之间做任何区分。这里上标M和E分别表示闵氏时空和欧氏四维空间，下同。

我们来看一下这些条件意味着什么。首先，第二个和第三个条件意味着应有

$$(x^M)^0 p_0^M + (x^M)^1 p_1^M + (x^M)^2 p_2^M + (x^M)^3 p_3^M = x_0^E p_0^E + x_1^E p_1^E + x_2^E p_2^E + x_3^E p_3^E,$$

因为我们有

$$e^{ik^M \cdot x^M} = e^{ik^E \cdot x^E}.$$

请注意在 $-++++$ 度规下 p_i 和 p^i 并无差别，于是按照第四个条件，我们有

$$(x^M)^1 p_1^M + (x^M)^2 p_2^M + (x^M)^3 p_3^M = x_1^E p_1^E + x_2^E p_2^E + x_3^E p_3^E,$$

即

$$(x^M)^0 p_0^M = x_0^E p_0^E.$$

在相对论性量子场论中我们通常认为 t 就是 x^0 而 ω 就是 p^0 ，它和 p_0 正好差了一个负号，而显然我们应该指定 $\tau = it$ 为 x_0^E ，因此就有

$$p_0^E = i\omega,$$

即应有

$$p_0^E = -\omega_n = i\omega, \quad x_0^E = \tau = it. \quad (2.9)$$

换句话说，四维欧氏空间中的 p^E 的时间分量和松原频率差了一个负号。这其实是正确的，因为在凝聚态场论中有

$$\phi(\tau) \propto \sum_n e^{-i\omega_n \tau} \phi_n,$$

加入动量之后就是

$$\phi(\tau, \mathbf{x}) \propto \sum_n \int d^3\mathbf{p} \phi_n(\mathbf{p}) e^{-i\omega_n \tau + i\mathbf{p} \cdot \mathbf{x}},$$

考虑到 $\omega_n \tau = \omega t$ ，做了反Wick转动之后这正好就是 $-+++$ 度规，即使用 ω_n 表示的傅里叶变换的 e 指数本身遵循 $-+++$ 度规而不是 $++++$ 度规，自然会导致 ω_n 和 p_0^E 差一个负号。

任何使用爱因斯坦求和得到的洛伦兹标量在Wick转动后形式均不变，因为Wick转动相当于做了一次坐标变换，而满足上下同指标求和规则的量在坐标变换下形式不变。由于Wick转动后是欧氏度规，我们不必再区分逆变和协变。需要修改的主要是积分测度，即需要加上或减少一个因子 i 。因此以上给出的关于坐标和频率的Wick转动足够让我们完成标量场论的配分函数的Wick转动。对矢量场只需要对 A^0 做代换即可，即取

$$A_0^E = (A^E)^0 = i(A^M)^0. \quad (2.10)$$

在完成配分函数的Wick转动之后还需要注意格林函数也需要做一些调整。从闵可夫斯基时空的理论中得到关联函数的方式是做变分导数 $\delta Z / \delta(iJ)$ ， J 是在以

$$\exp\left(i \int d^4x J\phi\right)$$

形式引入的。在Wick转动后，以上激励项变成了

$$\exp\left(\int d^4x J\phi\right).$$

在此过程中场 J 和 ϕ 没有发生任何变化，所以用四维欧氏时空计算闵氏时空中的格林函数，只需要先计算 $\delta Z^E / \delta(iJ)$ 然后做反Wick转动即可。然而，需要注意一件事：四维欧氏时空中本身也定义有格林函数，即所谓虚时间格林函数，而当激励项以

$$\exp\left(\int d^4x J\phi\right)$$

形式给出时，计算虚时间格林函数使用的泛函导数是 $\delta Z^E / \delta J$ ，没有 i 。（这和松原格林函数又差了一个负号）在四维欧氏时空中计算虚时间格林函数，使用Wick定理，画费曼图等使用的都是 $\delta Z^E / \delta J$ 方法求出的格林函数，但是最后切换回 $-+++$ 度规的闵氏时空时需要把 i 加回去，格林函数中涉及几个场加几个。例如，对二体格林函数，四维欧氏时空中的虚时间格林函数计算出来之后，需要做反Wick转动（在此过程中由于 $\tau = it$ ，闵氏时空格林函数的分子上会多出 $-i$ ），然后加上一个负号（由于泛函导数而导致的 i 有两个，因为有两个场，而 $i^2 = -1$ ），才能得到 $-+++$ 度规的闵氏时空中的格林函数。

对旋量场，事情略微复杂一些，因为 γ 矩阵也需要做同样的变换——当然其实可以不做任何变换，但是这样很多公式会看起来很奇怪。考虑配分函数：

$$\begin{aligned} Z &= \int \mathcal{D}\psi \int \mathcal{D}\bar{\psi} \exp\left(i \int d^4x \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi\right) \\ &= \int \mathcal{D}\psi \int \mathcal{D}\bar{\psi} \exp\left(\int d\tau \int d^3\mathbf{x} \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi\right) \\ &= \int \mathcal{D}\psi \int \mathcal{D}\bar{\psi} \exp\left(\int d\tau \int d^3\mathbf{x} \bar{\psi}\left(-\gamma^0 \frac{\partial}{\partial \tau} + i\gamma^i \partial_i - m\right)\psi\right). \end{aligned}$$

如果我们做变换

$$(\gamma^M)^0 = (\gamma^E)^0, \quad -i(\gamma^M)^i = (\gamma^E)^i, \quad (2.11)$$

并且，由于到了欧氏空间中，不再区分上下标，就得到

$$Z = \int \mathcal{D}\psi \int \mathcal{D}\bar{\psi} \exp\left(-\int d^4x^E \bar{\psi}(\gamma_\mu^E \partial_\mu^E + m)\psi\right) \quad (2.12)$$

这就是旋量场的Wick转动。相应的 γ 矩阵的代数需要做调整，但是旋量本身并不需要做调整（除了用 $-i\tau$ 代替 t 以外）。 γ 矩阵的变动意味着， $\gamma^\mu a_\mu$ 形式的量即 \not{a} 虽然看起来像是点乘，但是在Wick转动下会发生变化： a_0^E 相比于 $(a^M)^0$ 多出来了一个 i ，而 γ_i^E 相比于 $(\gamma^M)^i$ 多出来了一个 $-i$ ，再加上度规从 $-+++$ 变成了 $++++$ ，我们就得到

$$\gamma_\mu^E a_\mu^E = \not{a}^E = -i\not{a}^M = -i\gamma^{M\mu} a_\mu^M. \quad (2.13)$$

以上推导都是针对相对论情况下的。非相对论情况下的理论大多都是一个相对论情况下的理论的低能有效理论，因此所有的矢量分量的Wick转动规则仍然适用。例如，电磁场和非相对论性电子的耦合给出的拉氏量不具有洛伦兹协变性，但是电磁分量

2.3 洛伦兹群及其表示

粒子物理中底流形是给定的：四维闵可夫斯基时空，因此讨论时空的拓扑没有什么意义；我们将直接开始讨论闵可夫斯基时空的对称性。对称性可以大致分为两类：一类来自时空坐标变换，可能是平移，也可能是旋转或者推动，此时 δx 不为零，且常常 $\delta\phi$ 也不为零（注意即使平移时 $\delta\phi$ 也不为零，这是为了保持 $\delta\phi$ 始终为零）；另一类为内禀对称性，即多分量场内部不同分量之间的“转动”。无论是时空坐标变换还是内禀对称性实际上都是某些对称群——在本文中通常是一些李群——的表示。李群的连通分支可以完全用它的李代数加上指数映射

$$g = \exp(i\theta^i \sigma_i)$$

描述，其中 g 是任意的李群群元而 θ^i 为群参数，而不同分支之间的切换则对应某个离散的群，因此我们只需要讨论离散群的表示和李代数的表示就可以。

李代数的表示可以分成有限维表示，即矩阵表示，和无限维表示，通常是把李代数中的生成元写成一些导数算符。从李代数的有限维表示可以导出其无限维表示以及场表示，设 $\phi(x)$ 是一个场， $R(\Lambda)$ 是李群成员 Λ 在坐标上的有限维表示， $M(\Lambda)$ 是 Λ 在 $\phi(x)$ 的场值上的有限维表示，则将 Λ 作用在 $\phi(x)$ 上，就有

$$\phi'(R(\Lambda)x) = M(\Lambda)\phi(x),$$

现在让 Λ 为无穷小操作， θ 为很小的群参数取值， σ 是全体生成元，就得到

$$\bar{\delta}\phi = \underbrace{\theta^i M(\sigma_i)\phi}_{\delta\phi} - (\theta^i R(\sigma_i))^\mu \partial_\mu \phi, \quad (2.14)$$

第一项就是李代数的有限维表示作用在了 ϕ 的场值上，第二项是有限维表示作用在坐标上而诱导出的无限维表示，两者共同组成了所谓的场表示。

我们通常要求李群的有限维表示是么正的，因为如前所述场经过量子化后被提升为场算符，而场算符作用在真空态上得到多粒子态，既然多粒子态的变换是么正的，场算符的变换也必须是么正的。这里还有一个对应关系：设 M_{state} 是某个群在态空间上的表示（而与此同时不改变与之配套的算符），而 M_{field} 是某个群在算符上的场表示（而与此同时不改变量子态），则

$$[M_{\text{state}}, \psi] = M_{\text{field}}\psi, \quad (2.15)$$

其中 ψ 为任意的场。

2.3.1 平移

平移可能是我们能想到的最简单的变换，但实际上它相当特殊——在实际的物理问题中平移群通常不使用矩阵群表示，因为它无疑不是线性的。²

生成元 P^μ

$$[P_\mu, P_\nu] = 0. \quad (2.16)$$

接下来讨论平移群的线性表示。非Cartan元素的缺乏意味着平移群的有限维表示中不能够构造其Cartan元素（也就是全体平移生成元）的升降算符，因此平移群的有限维表示一定是平凡的。因此转而观察其无限维表示。我们考虑

$$\phi'(x) = \phi(x - a) = \phi(x) - a^\mu \partial_\mu \phi,$$

其中 a 是一个小量。于是，平移群的李代数的无穷维表示形如

$$P_\mu \propto -\partial_\mu.$$

不失一般性地我们取么正表示，并适当选取群参数，那么就有

$$P_0 = i\partial_0, \quad P_i = -i\partial_i. \quad (2.17)$$

我们特意让时间平移变换的方向和空间平移的方向反过来了。这是为了和物理中通常的时间演化方程的形式匹配。

2.3.2 旋转

三维欧氏空间的旋转矩阵

本节讨论 \mathbb{R}^3 中的旋转。所谓旋转，指的是一个可微的（从而可以通过一个物理上的微分方程实现的）等距同构变换。显然， \mathbb{R}^3 中的等距同构变换的全体就是 $O(3)$ ，其中任何一个矩阵的行列式都是 ± 1 。另一方面，变换可微意味着，可以找到形如下式的无穷小变换：

$$\mathbf{r} \longrightarrow \mathbf{r}' = \mathbf{r} + d\mathbf{r}, \quad d\mathbf{r} = dt \mathbf{A} \cdot \mathbf{r},$$

从而

$$\frac{d\mathbf{r}'}{d\mathbf{r}} = \mathbf{I} + dt \mathbf{A},$$

$$\det \left(\frac{d\mathbf{r}'}{d\mathbf{r}} \right) = 1 + dt \operatorname{tr} \mathbf{A} \approx 1.$$

因此 \mathbb{R}^3 中的旋转变换的全体就是 $SO(3)$ ，也就是行列式为1的全体 3×3 矩阵。

我们来分析 $SO(3)$ 的结构。矩阵 A 在 $SO(3)$ 中，当且仅当

$$AA^\top = I, \quad \det A = 1. \quad (2.18)$$

矩阵 A 含有9个分量； $AA^\top = I$ 是对称的，因此它等价于6个独立的纯数量方程； $\det A = \pm 1$ 可以直接从 $AA^\top = I$ 推出，则 $\det A = 1$ 的要求仅仅是去掉了其中的一支，因此对维数没

²可以在仿射空间中使用矩阵表示平移，但是这对本文没有太大意义。

有影响。这样， $SO(3)$ 就是 $9 - 6 = 3$ 维的，因此它有三个彼此独立的生成元。注意到绕 x 轴旋转、绕 y 轴旋转、绕 z 轴旋转都是 $SO(3)$ 的子群，这三者的表达式分别为

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, R_y = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}, R_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.19)$$

(2.19)自然导出三个生成元：

$$J_1 = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, J_2 = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J_3 = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (2.20)$$

注意到各个生成元都是厄米的，因为 $SO(3)$ 在此处的矩阵表示是幺正的。(2.20)也可以直接通过分析 $SO(3)$ 的抽象性质得到。对无穷小变换 $A = I + \theta J$ (θ 是小量)，(2.18)中的两个方程分别代表

$$(I + \theta J)(I + \theta J^\top) = I, \quad \det(I + \theta J) = 1 + \theta \operatorname{tr} J = 1,$$

也就是

$$J + J^\top = 0, \quad \operatorname{tr} J = 0.$$

上式描述了一个三阶方阵构成的向量空间，容易看出(2.20)正是它的一组基。通过显式表达式(2.20)可以导出

$$[J_i, J_j] = i\epsilon_{ijk}J_k. \quad (2.21)$$

这就得到了旋转群的李代数的对易关系，从而我们刻画了 $SO(3)$ 的李代数的结构。

$SU(2)$ 群

然而，能够实施空间旋转的并不只有 $SO(3)$ 。实际上， $SU(2)$ 也能做到这件事。下面我们引入 $SU(2)$ 群。 $SU(2)$ 是由二阶复方阵组成的矩阵群，由

$$A^\dagger A = I, \quad \det A = 1 \quad (2.22)$$

定义。二阶复数方阵一共有 $4 \times 2 = 8$ 个自由度。矩阵方程 $A^\dagger A = I$ 含有4个复数方程，但是因为它是厄米的，因此它只含有2个独立的复数方程，从而它含有4个独立的实数方程，也即它将总自由度降到了4。从 $A^\dagger A = I$ 只能够推出 $|\det A| = 1$ ， $\det A$ 仍然可以连续变化； $\det A = 1$ 则将 $\det A$ 确定到一个点上面，因此它将总自由度降到了3。因此 $SU(2)$ 的维度为3。可以证明， $SU(2)$ 中的每一个元素均形如

$$A = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}, \quad a^2 + b^2 + c^2 + d^2 = 1, \quad (2.23)$$

其中

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \mathbf{j} = \begin{pmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{pmatrix}, \mathbf{k} = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{pmatrix}, \quad (2.24)$$

它们实际上就是四元数。可以直接使用以上两式计算出 $SU(2)$ 的生成元，但是这样比较繁琐。使用 $SU(2)$ 的一般定义(2.22)，套用到无穷小变换

$$A = I + i\epsilon\sigma$$

上，可以发现 σ 是么正、无迹的矩阵；么正、无迹的矩阵的一组基（通常称为泡利矩阵）为

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.25)$$

它们也是厄米的。这正好是三个线性独立的矩阵，从而它们就是 $SU(2)$ 的一组彼此独立的生成元。我们有

$$\left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i\epsilon_{ijk} \frac{\sigma_k}{2}, \quad (2.26)$$

这表明 $SU(2)$ 和 $SO(3)$ 的李代数是—致的。

然而， $SU(2)$ 和 $SO(3)$ 是两个不同的群。³要看出这是为什么，考虑 $SU(2)$ 在由(2.24)的所有线性组合形成的空间上的表示。记

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}, \quad a^2 + b^2 + c^2 + d^2 = 1,$$

显然它是 $SU(2)$ 在二维复数空间上的矩阵表示。可以验证

$$q^{-1} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k},$$

则 $SU(2)$ 在(2.24)的所有线性组合形成的空间上的作用可写成

$$x \longrightarrow qxq^{-1}. \quad (2.27)$$

注意我们把 x 当成了算符来看待。 \mathbb{R}^3 中的向量 \mathbf{v} 与

$$x = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \quad (2.28)$$

一一对应，且容易证明

$$\det x = |\mathbf{v}|.$$

我们注意到

$$\det(qxq^{-1}) = (\det q)(\det x)(\det q^{-1}) = \det x,$$

因此(2.27)是等距同构。由于 $SU(2)$ 是李群，其表示(2.27)也必然是可微的。这表明变换（这是将(2.27)和(2.28)写在一起的结果）

$$\mathbf{v} \longrightarrow x \longrightarrow x' = qxq^{-1} \longrightarrow \mathbf{v}'$$

给出了 $SO(3)$ 中全部的成员，而且也仅仅给出这么多成员。然而，同一个变换(2.27)实际上对应着两个 q 。这表明 $SU(2)$ 实际上是 $SO(3)$ 的双覆盖。(2.23)意味着 $SU(2)$ 实际上就是四维球面 S^4 ，因此它是单连通群，因此它是李代数 $\mathfrak{so}(3)$ ——也就是 $\mathfrak{su}(2)$ ——的覆盖群。所有以 $\mathfrak{so}(3)$ 为李代数的李群中 $SU(2)$ 是最大的。

顺带提一句：以上的推导也表明，一个群在一个特定空间上的表示有时并不能完整地展现这个群的结构。 $\mathfrak{su}(2)$ 在 \mathbb{R}^3 上的表示，也就是 $\mathfrak{so}(3)$ 在 \mathbb{R}^3 上的表示，就是(2.20)，把它放进指数映射中得到的就是 $SO(3)$ 的三阶方阵形式。因此 $SU(2)$ 和 $SO(3)$ 在 \mathbb{R}^3 上的表示完全一样。换言之， $SU(2)$ 在 \mathbb{R}^3 上的表示不是忠实的。我们需要(2.24)这样更大的空间才能完全展示 $SU(2)$ 的结构。⁴

³当我们说两个群相同时我们是指它们在群论中同构，也就是说，我们所谓的“群”指的是抽象的群结构而不是具体的矩阵集合。如果 $SU(2)$ 和 $SO(3)$ 中的元素可以一一对应，且在这个对应下相应的乘法关系不变，那么这两个群——虽然使用了不同阶数的矩阵来表述——还是相同的。然而正如我们马上要看到的那样， $SO(3)$ 中的一个元素可以和 $SU(2)$ 中的两个元素相对应，且这种对应保持相应的乘法关系不变。因此这两个群不同。

⁴严格来说本节中我们使用了两种 $SU(2)$ 的表示。其一是 q ，也就是 \mathbb{C}^2 上的表示，其二是(2.27)，也就是 $GL(\mathbb{C}^2)$ 上的表示。

有限维不可约表示

本节分析 $SU(2)$ 的有限维不可约表示。注意到, J_1, J_2, J_3 互不对易。因此 J_3 就是 $\mathfrak{su}(2)$ 的一个Cartan子代数。使用本征值标记这些本征矢为 $|J_3^{(1)}\rangle, |J_3^{(2)}\rangle, \dots$ 。由于我们仅仅讨论有限维不可约表示, 本征值序列一定有上下限。记最大的本征值为 j 。现在寻找升降算符。我们将使用寻找李代数中升降算符的一般方法, 设升降算符形如

$$\hat{a}^\dagger = \lambda^j J_j, \quad [\hat{a}^\dagger, J_i] = c \hat{a}^\dagger,$$

并设 f_{ij}^k 为李代数的结构常数, 则方程

$$(f_{ji}^k - c \delta_j^k) \lambda^j = 0$$

的非零解给出了全部 c 和 λ 。求解得到 $c = \pm 1$, $c = 1$ 对应着 $\lambda^2 = i\lambda^1$, $\lambda^3 = 0$; $c = -1$ 对应着 $\lambda^2 = -i\lambda^1$ 。这表明有限维表示中仅有的升降算符为

$$J_+ \propto \lambda^1 J_1 + i\lambda^1 J_2, \quad J_- \propto \lambda^1 J_1 - i\lambda^1 J_2.$$

习惯上我们取

$$J_+ = \frac{1}{\sqrt{2}}(J_1 + iJ_2), \quad J_- = \frac{1}{\sqrt{2}}(J_1 - iJ_2). \quad (2.29)$$

J_+ 让本征值加一, J_- 让本征值减一。两个算符采取同样的系数是为了让 J_1, J_2, J_3 厄米时, J_+ 和 J_- 互为共轭转置。容易得到下面的对易关系:

$$[J_3, J_\pm] = \pm J_\pm, \quad [J_+, J_-] = J_3. \quad (2.30)$$

从升降算符的定义可以得到

$$J_+ |k\rangle = \alpha_k |k+1\rangle, \quad J_- |k+1\rangle = \alpha_k^* |k\rangle,$$

由于 $|j\rangle$ 是本征值最大的本征态,

$$J_+ |j\rangle = 0,$$

从而

$$\langle j | J_- = 0.$$

我们有

$$J_- |j\rangle = \alpha_{j-1}^* |j-1\rangle,$$

两边求模长,

$$\begin{aligned} |\alpha_{j-1}|^2 &= |J_- |j\rangle|^2 = \langle j | J_+ J_- |j\rangle \\ &= \langle j | J_+ J_- |j\rangle - \langle j | J_- J_+ |j\rangle \\ &= \langle j | [J_+, J_-] |j\rangle \\ &= \langle j | J_3 |j\rangle = \langle j | j |j\rangle = j. \end{aligned}$$

没有其它条件能够确定 α_{j-1} 具体应该取什么值。这是因为仅仅靠对易关系并不能够确定 J_1 和 J_2 (从而 J_+ 和 J_-) 作用在 $|k\rangle$ 上面会得到什么样的结果。然而, 注意到只要是同维度的非奇异

后者和 $SO(3)$ 在 $GL(\mathbb{C}^2)$ 上的表示完全一样, 但后者中的每一个都对应两个 q , 两者间的对应就是

$$\hat{B} \longrightarrow \hat{A} \hat{B} \hat{A}^{-1},$$

它并非一一对应。

有限维表示之间可以通过相似变换相互转换，不失一般性地我们可以认为所有的 α 都是正实数。于是

$$\alpha_{j-1} = \sqrt{j}.$$

另一方面，

$$\begin{aligned} |\alpha_k|^2 &= |J_- |k+1\rangle|^2 = \langle k+1 | J_+ J_- |k+1\rangle \\ &= \langle k+1 | [J_+, J_-] |k+1\rangle + \langle k+1 | J_- J_+ |k+1\rangle \\ &= \langle k+1 | J_3 |k+1\rangle + |J_+ |k+1\rangle|^2 \\ &= k+1 + |\alpha_{k+1}|^2, \end{aligned}$$

于是可以递推得到

$$|\alpha_k| = \frac{(j+k+1)(j-k)}{2},$$

得到显式表达式

$$J_+ |k\rangle = \sqrt{\frac{(j+k+1)(j-k)}{2}} |k+1\rangle, \quad J_- |k\rangle = \sqrt{\frac{(j+k)(j-k+1)}{2}} |k-1\rangle. \quad (2.31)$$

由于是有限维表示，反复作用 J_- 在 $|j\rangle$ 上最后一定会得到0。由(2.31)，得到零的唯一一种可能就是 $k = -j$ 。这表明从 $|j\rangle$ 出发不断作用 J_- 能够得到的全部非零本征向量为

$$|j\rangle, |j-1\rangle, \dots, |-j+1\rangle, |-j\rangle,$$

而由于 J_+ 和 J_- 是仅有的升降算符，它们就是 J_3 仅有的本征向量。这等价于 j 是半整数，且整个向量空间的维度为 $2j+1$ 。于是我们找到了 $SU(2)$ 的所有不可约有限维表示。

作为最简单的两个例子： $j = 0$ 时向量空间维数为1，所有李代数中的生成元都是0，而李群中的元素的表示为恒等运算；而在 $j = 1/2$ 时向量空间维度为2，相应的表示，使用 $|\frac{1}{2}\rangle$ 和 $|\frac{1}{2}\rangle$ 为基底，就是 $\sigma_1/2$ ， $\sigma_2/2$ 和 $\sigma_3/2$ 。

舒尔引理说明，在这种有限维不可约表示中卡西米尔算符的表示一定是单位矩阵的某个倍数，因此可以使用这个倍数来标记有限维不可约表示。就旋转群而言，

$$J^2 = (J_1)^2 + (J_2)^2 + (J_3)^2 \quad (2.32)$$

足够起到这个作用了。容易验证这的确是一个卡西米尔算符，且

$$J^2 |k\rangle = j(j+1) |k\rangle. \quad (2.33)$$

$\mathfrak{su}(2)$ 的任何一个可约表示是若干不可约表示的直和。我们使用 j 来标记各不可约表示，则 $\mathfrak{su}(2)$ 的任何一个可约表示的基矢量均可以写成 $\{|jm\rangle\}$ ，其中对于标记同一个矢量的 j 和 m ，有

$$m = -j, -j+1, \dots, 0, \dots, j,$$

且 j 是半整数。

无限维表示

旋转群的李代数的无穷维表示为

$$(J_i)_{\text{inf}} = -((J_i)_{\text{fin}} x) \cdot \nabla.$$

由于旋转群和时间维无关，梯度算符也可以看成是仅仅作用在空间维上的。代入具体的值就得到

$$J_1 = i(x^3 \partial_2 - x^2 \partial_3), \quad J_2 = i(x^1 \partial_3 - x^3 \partial_1), \quad J_3 = i(x^2 \partial_1 - x^1 \partial_2). \quad (2.34)$$

很容易验证，以上三个算符确实满足李代数(2.21)。

2.3.3 洛伦兹群

四维闵可夫斯基时空中的洛伦兹矩阵

首先讨论洛伦兹群在四维闵可夫斯基空间 $\mathbb{R}^{3,1}$ 上的表示，也就是保持闵可夫斯基度规⁵

$$\eta \equiv [\eta_{\mu\nu}]_{\mu\nu} = \text{diag}(1, -1, -1, -1) \quad (2.35)$$

不变的全体坐标变换矩阵 $O(1,3)$ 。记这些矩阵中的一个为 Λ ，容易看出保持度规不变等价于

$$\Lambda^\mu_\sigma \Lambda^\nu_\rho \eta_{\mu\nu} = \eta_{\sigma\rho}, \quad (2.36)$$

或者如果将 Λ 的矩阵形式看成是 $[\Lambda^\mu_\nu]_{\mu\nu}$ ，⁶那么就是

$$\Lambda^\top \eta \Lambda = \eta. \quad (2.37)$$

从(2.37)可以看出

$$\det \Lambda = \pm 1. \quad (2.38)$$

(2.37)意味着

$$(\Lambda_0^0)^2 - (\Lambda_0^1)^2 - (\Lambda_0^2)^2 - (\Lambda_0^3)^2 = 1,$$

从而

$$\Lambda_0^0 = \pm \sqrt{1 + (\Lambda_0^1)^2 + (\Lambda_0^2)^2 + (\Lambda_0^3)^2}. \quad (2.39)$$

现在我们将满足 $\det \Lambda = 1 > 0$ ——也就是说，坐标系的手性不改变——以及 $\Lambda_0^0 > 0$ ——也就是说，变换前的坐标时增加则变换后的坐标时也增加——的变换称为正规洛伦兹群。容易验证这确实是一个群。记之为 $SO(1,3)^\dagger$ 。正规洛伦兹群由于不改变时间维的指向，设 a^μ 是一个四维矢量，则 $\theta(a^0)$ 在正规洛伦兹群下不变，其中 θ 为阶跃函数。正规洛伦兹群中的成员称为洛伦兹变换。我们不认为接下来要谈到的含有宇称算符和时间反演算符的洛伦兹群成员为洛伦兹变换，因为它们无法通过可微的物理过程实现。

按照 $\det \Lambda$ 和 Λ_0^0 的正负，可以将洛伦兹群分成四支。其中两者皆为正的那一支就是 $SO(1,3)^\dagger$ 。考虑矩阵

$$\Lambda_P = \text{diag}(1, -1, -1, -1), \quad (2.40)$$

容易看出， $\Lambda_P \Lambda$ 还是洛伦兹群的成员，并且

$$\det \Lambda = -\det(\Lambda_P \Lambda).$$

它就是宇称算符。同样，时间反演算符

$$\Lambda_T = \text{diag}(-1, 1, 1, 1) \quad (2.41)$$

也将一个洛伦兹变换转化为另一个洛伦兹变换，且

$$\Lambda_0^0 = -(\Lambda_T \Lambda)_0^0.$$

⁵当然，正如我们在欧氏空间的张量分析当中能够看到的那样，如果基矢量彼此不正交，那么度规就不能够写成对角形式。这里我们实际上已经做了一个不失一般性的设定，要求度规一定是对角的。这总是可以做到，因为不对角的度规可以通过一个合同变换（这个变换对应着一个坐标变换）变成对角的。

⁶更加规范的写法是 Λ^μ_ν ，利用“第一个指标是行号、第二个指标是列号”的习惯。相应的也有 Λ_ν^μ 。使用这种记号，

$$\Lambda^\mu_\sigma \Lambda^\nu_\rho \eta_{\mu\nu} = \Lambda^\mu_\sigma \eta_{\mu\nu} \Lambda^\nu_\rho = (\Lambda^T)_\sigma^\mu \eta_{\mu\nu} \Lambda^\nu_\rho = \Lambda^T \eta \Lambda.$$

由于这两个算符都是可逆的，且可以作用在任何洛伦兹群的成员上，实际上洛伦兹群的四支之间只相差一个宇称算符和/或一个时间反演算符，也就是

$$O(1,3) = \{SO(1,3)^\uparrow, \Lambda_P SO(1,3)^\uparrow, \Lambda_T SO(1,3)^\uparrow, \Lambda_T \Lambda_P SO(1,3)^\uparrow\}. \quad (2.42)$$

需注意除了 $SO(1,3)^\uparrow$ 以外的部分只是陪集，并不能构成群，因为它们没有单位元。

洛伦兹群的李代数

洛伦兹群的四支之间不能通过可微的变换相互联系。因此，所谓洛伦兹群的李代数就是 $SO(1,3)^\uparrow$ 的李代数。洛伦兹群的定义(2.37)含有10个彼此独立的实数方程，因此留下6个自由度。因此只需要寻找6个生成元就可以得到洛伦兹群的李代数。由(2.37)可以得到无穷小生成元需要满足的关系为

$$\eta K + K^\top \eta = 0. \quad (2.43)$$

容易验证，设(2.20)中的 J 为 $J^{3\text{dim}}$ ，并定义

$$J_i = \begin{pmatrix} 0 & \\ & J_i^{3\text{dim}} \end{pmatrix}, \quad (2.44)$$

则 $J_i, i = 1, 2, 3$ 满足(2.43)，这样我们就得到了洛伦兹群的三个生成元。当然， $SO(3)$ 实际上是 $SO(1,3)^\uparrow$ 的一部分，因此这是合理的。通过考虑涉及 x^0 的矩阵，我们可以得到另外三个生成元：

$$K_1 = i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_2 = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_3 = i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (2.45)$$

这些生成元对应的李群元素称为推动，物理上它们涉及到时间，因此实际上是参考系变换。每个矩阵前面都加上了 i 是为了保持指数映射的形式成立，而我们现在讨论的洛伦兹群中的矩阵都是实数矩阵，因此 $i\theta\sigma$ 必须是实数，而由于参数 θ 是实数，生成元前面就应该多一个纯虚数，否则 $i\theta\sigma$ 不会是实数。请注意(2.45)不是厄米的。这可以预期，因为洛伦兹群涉及推动的部分是无界的，从而不是紧致的——这是闵可夫斯基时空的性质决定的，而非紧致的李群的有限维表示是平凡的。这也就是我们讨论的洛伦兹群的场表示中，场的分量的变换都是 $SU(2)$ 的表示，因为场的分量的变换对应场的内禀自由度的变换而这应该是么正的；推动和平移都没有非平凡的有限维么正表示，前者有有限维非么正表示，后者干脆没有有限维表示。容易验证，有下面的对易关系：

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \quad [J_i, K_j] = i\epsilon_{ijk} K_k, \quad [K_i, K_j] = -i\epsilon_{ijk} J_k. \quad (2.46)$$

这就得到了洛伦兹群的李代数。(2.46)中 J 之间的运算是封闭的，而 K 之间的运算不是封闭的，而且 J 和 K 之间不对易。但如果定义

$$N_i^\pm = \frac{1}{2}(J_i \pm iK_i), \quad (2.47)$$

就有

$$[N_i^+, N_j^+] = i\epsilon_{ijk} N_k^+, \quad [N_i^-, N_j^-] = i\epsilon_{ijk} N_k^-, \quad [N_i^+, N_j^-] = 0. \quad (2.48)$$

这表明 $SO(1, 3)^\dagger$ 的李代数是两个 $\mathfrak{su}(2)$ 的直和，或者说是 $\mathfrak{sl}(2, \mathbb{C})$ 。但可以证明， $SO(1, 3)^\dagger$ 并不是 $SL(2, \mathbb{C})$ ，事实上后者是前者的覆盖群，而且是双覆盖。

现在我们考虑洛伦兹群中的其它三支。事实上，由于(2.42)，讨论洛伦兹群的作用等价于讨论正规洛伦兹群在四维向量空间以及作用了宇称变换和/或时间反演变换的四维向量空间上的作用。⁷因此只需要讨论 $SO(1, 3)^\dagger$ 在宇称变换和时间反演变换之下会怎么变化，从而只需要讨论生成元 J_i 和 K_i 在宇称变换和时间反演变换之下的变化。宇称变换，写成分量矩阵的形式，是（取 μ 为行指标）

$$(\Lambda_P)^\mu_\nu = \Lambda_P = \text{diag}(1, -1, -1, -1).$$

我们没有区分行指标和列指标，因为宇称变换无非是洛伦兹变换的一种，而洛伦兹变换的应用场景中不会出现需要区分行指标和列指标的情况。⁸若对一个四维矢量做宇称变换，那么作用在其上的算符——也就是 4×4 矩阵——会发生如下所示的变换：⁹

$$A^\mu_\nu \longrightarrow (A')^\mu_\nu = (\Lambda_P)^\mu_\rho (\Lambda_P)^\sigma_\nu A^\rho_\sigma.$$

然后简单的计算就表明：

$$J_i \xrightarrow{\Lambda_P} J_i, \quad J_i \xrightarrow{\Lambda_T} J_i, \quad K_i \xrightarrow{\Lambda_P} -K_i, \quad K_i \xrightarrow{\Lambda_T} -K_i. \quad (2.49)$$

$SO(1, 3)^\dagger$ 的魏尔旋量表示

为了方便起见，我们把洛伦兹群的表示以及它的双覆盖的表示统称为洛伦兹群的表示。我们先只讨论 $SO(1, 3)^\dagger$ 的表示。完整的洛伦兹群的表示——也就是说考虑了宇称变换和时间反演变换——只需要在 $SO(1, 3)^\dagger$ 的表示上额外增加宇称变换和时间反演变换的规则即可。由于(2.48)， $SO(1, 3)^\dagger$ 的有限维表示全部可以看成是两个 $SU(2)$ 的有限维不可约表示的直积，或者若干个这样的直积的直和。我们记 $SO(1, 3)^\dagger$ 的有限维表示为 (j_1, j_2) ， j_1 和 j_2 分别代表其中一个 $SU(2)$ 的有限维表示的 j （见第2.3.2节）。通常用 j_1 表示 N_i^+ 那部分李代数的 j ， j_2 代表 N_i^- 那一部分李代数的 j 。

首先是 $(0, 0)$ 表示。这个表示作用在一个 $1 \times 1 = 1$ 维向量空间上。由之前的讨论， N_i^+ 和 N_i^- 都是零，也就是说， $(0, 0)$ 维表示是一个平凡的、只有恒等运算的表示。这个向量空间当然就是**标量**，这个表示称为**标量表示**，其中的对象是单分量的，在洛伦兹变换下不变。可以证明只有在这种表示下任何对象都不变。也就是说，只有标量在洛伦兹变换下完全不变；多分量对象不可能在洛伦兹变换下完全不变。

⁷数学上说，完整的洛伦兹群 $O(1, 3)$ 是正规洛伦兹群 $SO(1, 3)^\dagger$ 和 $\{I, \Lambda_P, \Lambda_T\}$ 的半直积。 $SO(1, 3)^\dagger$ 是一个正规子群，因此只需要知道 Λ_P 和 Λ_T 下 $SO(1, 3)^\dagger$ 的变动即可完全刻画 $O(1, 3)$ 。

⁸洛伦兹变换可以看成一种坐标变换，因此可以把 Λ 看成指标变换符号的另一种写法：

$$\Lambda^\nu_\mu = \beta^{\nu'}_\mu.$$

由于本文是从洛伦兹群的表示出发讨论问题而淡化“带坐标的流形”这一几何概念，有时对不同坐标系下同一矢量或旋量的各分量我们不使用 x^μ ， x^μ 这样明确区分坐标系的写法；相反，洛伦兹群的作用被认为是作用在同一个向量空间上的算符。

⁹你可能会疑惑为什么这个变换的形式看起来不是

$$\hat{B} \longrightarrow \hat{A} \hat{B} \hat{A}^{-1}.$$

实际上，如果我们将变换后的指标加上撇号，那么此变换形如

$$A^\mu_\nu \longrightarrow A^{\mu'}_{\nu'} = (\Lambda_P)^{\mu'}_\rho (\Lambda_P)^\sigma_{\nu'} A^\rho_\sigma.$$

由于 Λ_P 实际上是坐标变换符号，分量矩阵 $[(\Lambda_P)^{\mu'}_\rho]_{\mu', \rho}$ 和 $[(\Lambda_P)^\sigma_{\nu'}]_{\sigma, \nu'}$ 的确确互为逆矩阵。由于我们使用的度规是正规型，因此无需在意 Λ_P 、它的转置或者它的逆的区别——三者实际上是一样的。

接着是 $(\frac{1}{2}, 0)$ 表示。这是一个二维表示，这个表示中，

$$N_i^+ = \frac{\sigma_i}{2}, \quad N_i^- = 0. \quad (2.50)$$

使用(2.47)可以推导出

$$J_i = \frac{1}{2}\sigma_i, \quad K_i = -\frac{i}{2}\sigma_i. \quad (2.51)$$

于是使用指数映射就能够得到洛伦兹变换下这种二分量对象的变化方式，也就是

$$R_\theta = \exp\left(\frac{1}{2}i\theta_i\sigma_i\right), \quad B_\phi = \exp\left(\frac{1}{2}\phi_i\sigma_i\right). \quad (2.52)$$

值得注意的是，洛伦兹变换下各个分量混合起来的系数一般来说含有虚部。类似的， $(0, \frac{1}{2})$ 表示也是一个二维表示，其中

$$N_i^+ = 0, \quad N_i^- = \frac{\sigma_i}{2}, \quad (2.53)$$

这又等价于

$$J_i = \frac{1}{2}\sigma_i, \quad K_i = \frac{i}{2}\sigma_i. \quad (2.54)$$

其变换方式为

$$R_\theta = \exp\left(\frac{1}{2}i\theta_i\sigma_i\right), \quad B_\phi = \exp\left(-\frac{1}{2}\phi_i\sigma_i\right). \quad (2.55)$$

我们称 $(\frac{1}{2}, 0)$ 表示为**左手旋量**， $(0, \frac{1}{2})$ 表示为**右手旋量**。两者统称为**魏尔旋量**。容易看出，两种旋量在旋转下的变化相同，在推动下的变化差一个负号。

以下，我们仿照向量的指标升降、坐标变换等规则，定义一套旋量的指标升降、坐标变换规则，称为**范德瓦尔登符号**。

旋量度规和指标升降 首先定义**旋量度规**

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad (2.56)$$

为什么叫做这个名字马上可以看到。很容易看出，

$$(-\epsilon)\epsilon = I, \quad \epsilon\sigma_i^*(-\epsilon) = -\sigma_i. \quad (2.57)$$

使用这两个关系式可以证明：若 χ_L 是一个左手旋量，那么

$$\chi_L^C = \epsilon\chi_L^* \quad (2.58)$$

就是一个右手旋量；若 χ_R 是一个右手旋量，则

$$\chi_R^C = -\epsilon\chi_R^* \quad (2.59)$$

就是一个左手旋量。（方法是，将 R_θ 或者 B_ϕ 作用到 χ_L 上得到 χ_L' ，从而可以计算出 $(\chi_L^C)'$ ，然后使用(2.57)凑出 $(\chi_L^C)'$ 和 χ_L 之间的关系）在(2.59)中我们特意加了一个负号，这样

$$(\chi_L^C)^C = \chi_L, \quad (\chi_R^C)^C = \chi_R.$$

请注意 ϵ 是可逆的，因此，(2.58)和(2.59)表明有一样多的左手旋量和右手旋量，它们通过(2.58)和(2.59)一一对应。一对通过(2.58)和(2.59)相对应的左手旋量和右手旋量就可以看成一个抽象的魏尔旋量 χ 分别在 $(\frac{1}{2}, 0)$ 和 $(0, \frac{1}{2})$ 中的表示。我们使用 χ_a 表示 χ_L 的第 a 个分量， $\chi^{\dot{a}}$ 表示 χ_R 的第 a 个分量，那么由于

$$\chi_R = \chi_L^C, \quad \chi_L = \chi_R^C,$$

有

$$\chi^{\dot{a}} = \sum_b (\text{the } (a, b)\text{-element of } \epsilon) \cdot \chi_b^*, \quad \chi_a = \sum_b (-\text{the } (a, b)\text{-element of } \epsilon) \cdot (\chi^{\dot{b}})^*.$$

于是定义¹⁰

$$\epsilon^{ab} = \epsilon^{\dot{a}\dot{b}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \epsilon_{ab} = \epsilon_{\dot{a}\dot{b}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (2.60)$$

以及

$$\chi^{\dot{a}} = (\chi^a)^*, \quad \chi_{\dot{a}} = (\chi_a)^*, \quad (2.61)$$

我们得到了 χ 在左右手旋量空间中的表示相互切换的公式

$$\chi^a = \epsilon^{ab} \chi_b, \quad \chi^{\dot{a}} = \epsilon^{\dot{a}\dot{b}} \chi_{\dot{b}}, \quad \chi_a = \epsilon_{ab} \chi^b, \quad \chi_{\dot{a}} = \epsilon_{\dot{a}\dot{b}} \chi^{\dot{b}}. \quad (2.62)$$

其中一上一下两个相同指标要求和。我们看到了 ϵ 的地位正是矢量分析中度规的地位，因此称它为旋量度规。

使用指标升降的一般理论，我们发现，若

$$\psi_b = A_b^a \chi_a, \quad \psi_{\dot{b}} = A_{\dot{b}}^{\dot{a}} \chi_{\dot{a}}, \quad \psi^b = A^b_a \chi^a, \quad \psi^{\dot{b}} = A^{\dot{b}}_{\dot{a}} \chi^{\dot{a}}, \quad (2.63)$$

则

$$A^c_d = \epsilon^{cb} A_b^a \epsilon_{ad}, \quad A^{\dot{c}}_{\dot{d}} = (A^c_d)^*, \quad A^{\dot{c}}_{\dot{d}} = \epsilon^{\dot{c}\dot{b}} A_{\dot{b}}^{\dot{a}} \epsilon_{\dot{a}\dot{d}}. \quad (2.64)$$

通常对作用在左手旋量上的矩阵 A ，规定 A_a^b 就是 A ，这样(2.63)中的四个式子全部等价于

$$\psi_L = A \chi_L.$$

坐标变换 得到了指标升降关系，再来看坐标变换关系。按照前述规定，

$$(\sigma_i)_b^a = \sigma_i,$$

对左手旋量我们有

$$\chi'_a = \Lambda_a^b \chi_b = \exp \left(\frac{1}{2} i \theta_i \sigma_i + \frac{1}{2} \phi_i \sigma_i \right)_a^b \chi_b.$$

与之对应的右手旋量会怎样变换？很容易想到，应该使用(2.64)来得到对应的作用在右手旋量上的变换矩阵。但实际上右手旋量的变换方式在定义时就已经确定了（因为 $(0, \frac{1}{2})$ 表示本来就是 $SO(1, 3)^\dagger$ 的某种表示）。我们要验证这两种变换方式是不是一致。从右手旋量的定义出发我们有（下面的公式是在处理分量，具体指标 a 和 b 是上标还是下标由于没有微分几何的阐释，尚不确定）

$$\chi'^{\dot{a}} = \sum_b \text{the } (a, b)\text{-element of } \exp \left(\frac{1}{2} i \theta_i \sigma_i - \frac{1}{2} \phi_i \sigma_i \right) \cdot \chi^{\dot{b}}.$$

而如果右手旋量的洛伦兹变换服从(2.64)，那么就有

$$\chi'^{\dot{a}} = \Lambda^{\dot{a}}_{\dot{b}} \chi^{\dot{b}}.$$

这两种变换方式是一致的，当且仅当

$$\text{the } (a, b)\text{-element of } \exp \left(\frac{1}{2} i \theta_i \sigma_i - \frac{1}{2} \phi_i \sigma_i \right) = \sum_{c, d} \epsilon^{\dot{a}\dot{c}} \left(\exp \left(\frac{1}{2} i \theta_i \sigma_i + \frac{1}{2} \phi_i \sigma_i \right)_c^d \right)^* \epsilon_{\dot{b}\dot{d}}$$

¹⁰到目前为止我们还没有赋予上下指标任何意义，所以我们可以任意地规定涉及它们的表达式。此处上下指标看起来和逆变-协变关系很像，虽然它们可能未必可以赋予微分几何中的意义。

使用(2.57)很容易证明这确实是对的。类似的,可以表明 $\chi_{\dot{a}}$ 和 χ^a 的变换矩阵正是 $\Lambda_a^{\dot{b}}$ 通过(2.64)变换得到的。于是我们就得到了 $\Lambda_a^{\dot{b}}$ 和 $\Lambda_{\dot{b}}^a$ 两个分量矩阵。

朴素地使用指标升降,可以得到

$$\begin{cases} \Lambda_a^{\dot{b}} = \exp\left(\frac{1}{2}i\theta_i\sigma_i + \frac{1}{2}\phi_i\sigma_i\right)_a^{\dot{b}}, & \Lambda_{\dot{b}}^a = \exp\left(\frac{1}{2}i\theta_i\sigma_i - \frac{1}{2}\phi_i\sigma_i\right)_a^{\dot{b}}, \\ \Lambda_{\dot{a}}^{\dot{b}} = \exp\left(-\frac{1}{2}i\theta_i\sigma_i^* + \frac{1}{2}\phi_i\sigma_i^*\right)_a^{\dot{b}}, & \Lambda_{\dot{b}}^a = \exp\left(-\frac{1}{2}i\theta_i\sigma_i^* - \frac{1}{2}\phi_i\sigma_i^*\right)_a^{\dot{b}}, \end{cases} \quad (2.65)$$

且它们满足(2.64)。通过 σ 矩阵的定义¹¹,会发现 $\sigma_i^* = \sigma_i^\top$,于是我们还可以得到

$$\Lambda_{\dot{a}}^{\dot{b}} = \exp\left(-\frac{1}{2}i\theta_i\sigma_i + \frac{1}{2}\phi_i\sigma_i\right)_a^{\dot{b}}, \quad \Lambda_{\dot{b}}^a = \exp\left(-\frac{1}{2}i\theta_i\sigma_i - \frac{1}{2}\phi_i\sigma_i\right)_b^a. \quad (2.66)$$

可以看到 $\Lambda_a^{\dot{b}}$ 和 $\Lambda_{\dot{a}}^{\dot{b}}$ 是不同的,实际上,它们正好互为逆。这和坐标变换矩阵应该具有的对称性有冲突。当然,只要采用主动观点,把 $\Lambda_a^{\dot{b}}$ 之类的东西看成保持坐标系不变、改变系统中所有旋量的操作,就不会有任何矛盾:所谓的 $\Lambda_a^{\dot{b}}\chi_b$ 实际上是 $\chi_b(\Lambda^{-1})_a^{\dot{b}}$,然后我们容易验证后面将要提到的 $\chi_a\psi^a$ 的洛伦兹不变性。另一方面,如果把洛伦兹变换看成坐标变换,那么我们是在用 $\Lambda_a^{\dot{b}}$ 表示 $\Lambda_{a'}^{\dot{b}}$,而在用 $\Lambda_{\dot{a}}^{\dot{b}}$ 表示 $\Lambda_{\dot{a}'}^{\dot{b}}$ 。我们知道 $\Lambda_a^{\dot{b}}$ 确实是 $\Lambda_{a'}^{\dot{b}}$ (定义如此),那么对它做指标升降会得到

$$\epsilon^{ac}\Lambda_c^{\dot{d}}\epsilon_{db} = \epsilon^{a'c'}\Lambda_{c'}^{\dot{d}}\epsilon_{db} = \Lambda_b^{a'},$$

因此实际上, $\Lambda_a^{\dot{b}}$ 中下标 a 对应一个洛伦兹变换之后的旋量,上标 b 对应洛伦兹变换之前的旋量;而 $\Lambda_{\dot{a}}^{\dot{b}}$ 则正好相反, b 对应变换前的旋量而 a 对应变换后的旋量,因此如果不考虑 Λ 的两个指标实际上是不同基底下的指标而朴素地做升降指标,就会让 $\Lambda_a^{\dot{b}}$ 和 $\Lambda_{\dot{a}}^{\dot{b}}$ 不同。实际上这个情况在普通的张量分析中也会出现。这样,使用 $\Lambda_a^{\dot{b}}$ 的记号会写成

$$\chi'^b = \Lambda_a^{\dot{b}}\chi^a$$

的洛伦兹变换现在就写成

$$\chi'^b = \Lambda_a^{\dot{b}}\chi^a,$$

即 χ^b 和 χ_b 的变换可以统一地用简洁的爱因斯坦求和写成

$$\chi'^b = \Lambda_a^{\dot{b}}\chi^a, \quad \chi'_b = \Lambda_{\dot{b}}^a\chi_a.$$

带点的指标同理。 $\Lambda_a^{\dot{b}}$ 和 $\Lambda_{\dot{a}}^{\dot{b}}$ 正好互为逆的事实给出了下面直觉上非常合理的式子:

$$\Lambda_{a'}^{\dot{b}}\Lambda_c^{a'} = \Lambda_a^{\dot{b}}\Lambda_c^a = \delta_b^a, \quad \Lambda_{\dot{a}'}^{\dot{b}}\Lambda_{\dot{c}}^{\dot{a}'} = \Lambda_{\dot{a}}^{\dot{b}}\Lambda_{\dot{c}}^{\dot{a}} = \delta_{\dot{c}}^{\dot{b}}. \quad (2.67)$$

¹¹这里有一个可能引起困惑的细节。在张量代数中,我们有

$$(T^\top)_a^{\dot{b}} = T_a^{\dot{b}},$$

但是在此处我们却似乎写出了这样的表达式:

$$(\sigma_a^{\dot{b}})^\top = \sigma_b^{\dot{a}},$$

两者相差一个指标升降。产生这样的现象的原因在于,当我们通过(2.25)定义

$$\sigma_a^{\dot{b}} = \sigma$$

时,左右两边的 σ 实际上有微妙的差异——左边的 σ 是某种旋量张量,右边的 σ 只是一个矩阵。这就意味着,左边的 σ 的转置运算并不是简单的“把矩阵翻转过来”(因为转置之后的结果必须是协变的), $(\sigma^\top)_a^{\dot{b}} = \sigma_b^{\dot{a}}$ 关于左边的 σ (不是分量矩阵 $\sigma_a^{\dot{b}}$!)成立。然而,(2.25)——从而由它导出的 σ 矩阵的厄米性——是关于右边的 σ 的,也就是说它仅仅关于左边的分量矩阵 $\sigma_a^{\dot{b}}$ 。因此,通过(2.25)导出的 $\sigma^* = \sigma^\top$ 中的转置就是简单的将矩阵翻转过来。

不变量 因此我们已经验证了旋量满足指标升降关系和坐标变换关系，尽管指标升降关系中用到的“度规”并不是对称的。因此我们可以使用 $\chi^{a'}$ 这样的记号代替 χ'^a ，也就是说我们把 χ 看成某种抽象的实体，它在坐标变换之下是不变的。¹²

下面我们要讨论使用旋量以及它们的一阶导数能够构造出怎样的标量。这是很重要的，因为如果需要使用旋量场来描述某种物理过程，那么对应的拉氏量应该是标量，或者至少是协变的，也就是说在洛伦兹变换下的变化量能够写成一个散度项。注意到

$$\chi_a \xi^a \longrightarrow \chi'_a \xi'^a = \Lambda_a^b \chi_b \Lambda^a_c \xi^c = \delta_c^b \chi_b \xi^c = \chi_a \xi^a,$$

同样的有

$$\chi_{\dot{a}} \xi^{\dot{a}} \longrightarrow \chi'_{\dot{a}} \xi'^{\dot{a}} = \chi_{\dot{a}} \xi^{\dot{a}}.$$

但 $\chi_a \xi^{\dot{a}}$ 或者 $\chi^{\dot{a}} \xi_a$ 这种量却没有不变性。这表明，只有“同类”——也就是都带点或者都不带点——的指标才能够一上一下地求和而得到一个标量。

四维矢量表示

我们接着讨论 $(\frac{1}{2}, \frac{1}{2})$ 表示。由于 N_i^- 和 N_i^+ 对易，这个表示实际上就是 $(\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$ ，于是这个表示可以使用一个左手旋量和一个右手旋量的张量积表示，记作 $v_a^{\dot{b}}$ 。其中每个指标都独立地以(2.65)变换。当然，可以使用指标升降关系把 $v_a^{\dot{b}}$ 转化为 v_{ab} 。这么做的好处在于，我们可以发现 $(\frac{1}{2}, \frac{1}{2})$ 表示实际上可以约化为一个厄米的不可约表示和一个反厄米的不可约表示的直和。¹³ 注意到

$$v'_{cd} = \Lambda_c^a \Lambda_d^{\dot{b}} v_{ab},$$

可以得到

$$v'_{dc} = \Lambda_d^a \Lambda_c^{\dot{b}} v_{ab} = \Lambda_d^b \Lambda_c^{\dot{a}} v_{ba},$$

从而

$$(v'_{dc})^* = \Lambda_d^{\dot{b}} \Lambda_c^a (v_{ba})^*.$$

显然，如果分量矩阵 v_{ab} 是厄米的，那么它经过洛伦兹变换之后还是厄米的；如果它是反厄米的，那么经过洛伦兹变换之后它还是反厄米的。因此 $(\frac{1}{2}, \frac{1}{2})$ 可以分解成其厄米子代数和反厄米子代数的直和。

我们来仔细分析其厄米子代数。二阶厄米方阵组成的向量空间有 $2^2 \times 2/2 = 4$ 维。因此只需要找到四个独立的厄米方阵即可。容易看出，三个泡利矩阵连同单位矩阵构成了这样的一组基。为便于书写下标，使用 $\sigma^i, i = 1, 2, 3$ 表示三个泡利矩阵，又使用 σ^0 表示单位矩阵。我们记 σ_{ab}^μ 为 σ^μ 的第 a 行 b 列。（这个记号和第2.3.3节中的 σ_a^b 是不一样的！）于是就可以把 $(\frac{1}{2}, \frac{1}{2})$ 表示的厄米子代数的成员统一地写成

$$v_{ab} = v_\nu \sigma_{ab}^\nu, \quad v_a^{\dot{b}} = \epsilon^{\dot{b}c} v_\nu \sigma_{ac}^\nu. \quad (2.68)$$

于是就可以使用 v_ν 来代替整个 v_{ab} ，并且 v_ν 的变换方式正是洛伦兹变换下的矢量。总之， $(\frac{1}{2}, \frac{1}{2})$ 表示的厄米子代数就是四维矢量。因此，正如矢量可以看成二阶张量的平方根那样，魏尔旋量也可以看成矢量的平方根。可以预期，四维矢量不见得能够描述所有的物理系统，因为它们不够基本。

¹²在微分几何中矢量、张量等可以处理为一个抽象的几何实体，然后定义矢量和张量的分量，然后导出指标升降规则和坐标变换规则；我们这里的步骤则正好相反，我们是首先通过洛伦兹群的表示得到指标升降规则和坐标变换规则，然后发现这两个规则允许我们把 χ 看成某种抽象的实体。

¹³同样，这里所谓的厄米和反厄米也是就分量矩阵 v_{ab} 而论的，并不涉及旋量张量 v 的（协变的）转置。见注11。

同样我们考虑四维矢量能够构造出来的二阶的不变量。通过坐标变换关系容易看出，这样的不变量一定具有形式 $A^\mu B_\mu$ 。

狄拉克旋量

在第2.3.3节和第2.3.3节中我们只讨论正规洛伦兹变换的表示。当然，完整的洛伦兹群的表示肯定还是旋量，只不过我们还需要指定时间反演变换和宇称变换的表示。注意到(2.49)，并考虑 N_i^\pm 的定义(2.47)，在宇称变换下

$$N_i^+ \xrightarrow{\Lambda_P} N_i^-, \quad N_i^- \xrightarrow{\Lambda_P} N_i^+. \quad (2.69)$$

这意味着经过宇称变换，原本是 $(\frac{1}{2}, 0)$ 表示的变换矩阵现在变成了 $(0, \frac{1}{2})$ 表示的变换矩阵，原本是 $(0, \frac{1}{2})$ 表示的变换矩阵则变成了 $(\frac{1}{2}, 0)$ 表示的变换矩阵。也就是说宇称变换下左手旋量按照右手旋量的方式变换，右手旋量按照左手旋量的方式变换。¹⁴这就是“左手”、“右手”名称的来历：它们之间的变换和左右手坐标系之间的变换是完全一样的。需要注意的是， χ_L 经过宇称变换之后未必变成 χ_R ，同理 χ_R 经过宇称变换之后也未必变成 χ_L 。

如果一个体系在宇称变换之下不变，那么不可能仅仅使用一个左手旋量或者一个右手旋量描述它，因为宇称变换会改变旋量的手征。因此，完整描述一个体系的物理量一定能够写成一对对相互匹配的左手旋量和右手旋量。我们称像这样由一个左手旋量和一个右手旋量组合而成的场

$$\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (2.70)$$

为狄拉克旋量。它是洛伦兹群的一个可约表示，因为其变换矩阵

$$\Lambda_{\text{Dirac}} = \begin{pmatrix} \Lambda_{(\frac{1}{2}, 0)} & \\ & \Lambda_{(0, \frac{1}{2})} \end{pmatrix} \quad (2.71)$$

是左手旋量和右手旋量的变换矩阵直和起来得到的结果。现在我们做一个宇称变换，那么变换矩阵就成为

$$\Lambda'_{\text{Dirac}} = \begin{pmatrix} \Lambda_{(0, \frac{1}{2})} & \\ & \Lambda_{(\frac{1}{2}, 0)} \end{pmatrix}.$$

因此，考虑宇称变换在狄拉克旋量上的表示 $(\Lambda_P)_{\text{Dirac}}$ ，我们有

$$(\Lambda_P)_{\text{Dirac}} \Lambda_{\text{Dirac}} (\Lambda_P)_{\text{Dirac}}^{-1} = \Lambda'_{\text{Dirac}},$$

于是

$$(\Lambda_P)_{\text{Dirac}} = a \begin{pmatrix} 0 & I_{2 \times 2} \\ I_{2 \times 2} & 0 \end{pmatrix}.$$

由于我们考虑的是么正表示， a 的模长为1。请注意宇称变换、时间反演变换和恒等变换自成一个群，因此从有关正规洛伦兹群的表示的任何知识——它们是通过考虑正规洛伦兹群的李代数得到的，和宇称变换没有任何关系——都不可能把 a 的值确定下来。不失一般性地通常取 $a = 1$ ，于是就有

$$(\Lambda_P)_{\text{Dirac}} = \begin{pmatrix} 0 & I_{2 \times 2} \\ I_{2 \times 2} & 0 \end{pmatrix}. \quad (2.72)$$

¹⁴我们并没有给出宇称变换的变换矩阵到底是什么。实际上因为宇称变换不在正规洛伦兹群中，只靠旋量的定义是不可能完全确定下宇称变换的矩阵表示是什么的。

那么狄拉克旋量在宇称变换下的变化就是

$$\psi \xrightarrow{\Lambda_P} \psi^P = \begin{pmatrix} 0 & I_{2 \times 2} \\ I_{2 \times 2} & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix}. \quad (2.73)$$

也就是说宇称变换把狄拉克旋量变成了另一个同样具有 χ_L 和 ξ_R 的对象，但是上下位置发生了变化。

狄拉克旋量上可以定义一种重要的变换。它不是洛伦兹变换的一种，其形式为

$$\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \longrightarrow \psi^C = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}. \quad (2.74)$$

容易看出变换之后得到的结果还是狄拉克旋量。

无限维表示

洛伦兹群的无限维表示完全由旋转生成元的无限维表示和推动生成元的无限维表示确定。前者已经由(2.34)给出了。虽然(2.34)是在三维空间中推导出来的，但是因为推导它时各个坐标的标号和本节一致，都是 x^1, x^2, x^3 ，它也适用于四维闵可夫斯基时空。使用基本上一样的方法，从(2.45)可以导出

$$K_1 = -i(x^1 \partial_0 + x^0 \partial_1), \quad K_2 = -i(x^2 \partial_0 + x^0 \partial_2), \quad K_3 = -i(x^3 \partial_0 + x^0 \partial_3). \quad (2.75)$$

2.3.4 庞加莱群

正规子群平移群半直积上洛伦兹群就得到了庞加莱群。

现在我们分析庞加莱群的李代数。独立的生成元总共有10个，4个是平移群生成元，3个是旋转生成元，3个是推动生成元。平移群的李代数为(2.16)，而洛伦兹群的李代数为(2.46)，因此只需要 P_μ 和 J_i, K_i 的对易关系就能够完全确定庞加莱群的李代数。首先由于旋转操作不涉及时间维，显然我们有

$$[J_i, P_0] = 0. \quad (2.76)$$

使用(2.17)、(2.34)和(2.75)三式，可以推导出以下关系：

$$[J_i, P_j] = i\epsilon_{ijk} P_k, \quad [K_i, P_j] = i\delta_{ij} P_0, \quad [K_i, P_0] = -iP_i. \quad (2.77)$$

虽然使用的是特殊的表示，但由于推导出来的都是对易关系，因此它们普遍成立。(2.77)和(2.46)给出了庞加莱群的李代数。

为了简化记号，我们设 $M_{\mu\nu}$ 是一个反对称的矩阵，且

$$J_i = \frac{1}{2}\epsilon_{ijk} M_{jk}, \quad K_i = M_{0i}, \quad (2.78)$$

那么(2.77)完全等价于

$$[M_{\mu\nu}, P_\rho] = i(\eta_{\mu\rho} P_\nu - \eta_{\nu\rho} P_\mu), \quad (2.79)$$

而(2.46)等价于

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho} M_{\nu\sigma} + \eta_{\nu\sigma} M_{\mu\rho} - \eta_{\mu\sigma} M_{\nu\rho} - \eta_{\nu\rho} M_{\mu\sigma}). \quad (2.80)$$

(2.16)，(2.79)和(2.80)共同描述了庞加莱群的李代数。

庞加莱群的群参数一共有10个，在定义了 M 之后我们可以把群参数写成 a^μ 和 $\omega^{\mu\nu}$ ，其中 ω 是反对称矩阵，从而庞加莱群就可以写成

$$\Lambda = \exp(i a^\mu P_\mu + i \omega^{\mu\nu} M_{\mu\nu}).$$

实际上，即使 ω 不是反对称的，上式照样给出庞加莱群，因为 ω 的对称部分和 M 相乘得到的一定是零。反对称的要求仅仅是让每个变换对应唯一一个群参数。

庞加莱群的卡西米尔元有两个，它们分别是

$$P_\mu P^\mu = m^2, \quad (2.81)$$

和

$$W_\mu W^\mu = j_1 + j_2 = j, \quad (2.82)$$

其中

$$W^\mu = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} P_\nu M_{\rho\sigma}. \quad (2.83)$$

我们使用 $P_\mu P^\mu$ 给出的 m^2 实际上正是该表示下场的质量的平方。详情见第3.2.1节。

2.4 对称性操作和守恒量

2.4.1 连续对称性对应的守恒量

平移和能动张量

首先是最简单的平移。处理平移时假定场是标量场，这无损一般性，因为平移没有有限维表示，因此不会导致场分量发生混合。平移变换作用于场上得到的结果是：

$$x^\mu \longrightarrow x'^\mu = x^\mu + a^\mu, \\ \bar{\delta}\phi = \phi'(x') - \phi(x) = 0.$$

平移变换导致的场的变化（就是没有变化）的扣除了坐标变化导致的影响的场的变化就是

$$\delta\phi = \bar{\delta}\phi - \partial_\mu \phi \delta x^\mu = -\partial_\mu \phi \delta a^\mu,$$

或者，由于场在坐标拖曳下的变动实际上就是平移变换的无限维表示，可以直接使用(2.17)得到上式。于是对应的守恒流为

$$0 = \partial_\mu \left(-\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial_\nu \phi \delta a^\nu + \mathcal{L} \delta a^\mu \right) = \partial_\mu \left(-\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial_\nu \phi + \mathcal{L} \delta_\nu^\mu \right) \delta a^\nu,$$

考虑到 δa^μ 的任意性，我们有

$$T_\mu^\nu = \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \partial_\mu \phi - \mathcal{L} \delta_\mu^\nu, \quad \partial_\nu T_\mu^\nu = 0. \quad (2.84)$$

我们称 T_μ^ν 为能动张量。它给出了4个守恒荷，其中一个是来自时间平移不变性的能量

$$E = \int d^3x T_0^0 = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \partial_0 \phi - \mathcal{L} \right), \quad (2.85)$$

很容易看出 $\mu = 0$ 时这就是在做从拉氏量到哈密顿量的勒让德变换，即上式确实就是能量。另外三个是来自空间平移不变性的动量

$$P_i = \int d^3x T_i^0 = \int d^3x \frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \partial_i \phi. \quad (2.86)$$

能动张量的纯空间部分是能量和动量的输运流，因此就是应力张量。¹⁵ 相应的，

$$\mathcal{P}_\mu = \frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \partial_\mu \phi - g_\mu^0 \mathcal{L} \quad (2.87)$$

为四维动量(E, \mathbf{p})的密度。在计算场的三维动量时要注意一点：由于闵可夫斯基度规为 $(+, -, -, -)$ ，闵可夫斯基时空中空间部分的基矢量实际上是指向空间坐标减少的方向的。从而，

$$\begin{aligned} \mathbf{P} &= \int d^3x \frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \partial_i \phi \mathbf{g}^i \\ &= - \int d^3x \frac{\partial \mathcal{L}}{\partial \partial_0 \phi} \partial_i \phi \mathbf{g}_{3\text{dim}}^i, \end{aligned}$$

也即

$$\mathbf{P} = - \int d^3x \pi \nabla \phi. \quad (2.88)$$

旋转和角动量

接下来是旋转对称性。¹⁶ 旋转对称性不涉及时间维，于是我们有

$$\delta x^i = \epsilon^i_{jk} x^j \theta^k,$$

相应的

$$\delta \phi^a = i(J_i)^a_b \theta^i \phi^b - \epsilon^i_{jk} x^j \theta^k \partial_i \phi^a.$$

其中指标 a, b 跑遍 ϕ 的所有分量， J 指的是旋转生成元在 ϕ 所属的向量空间上的表示，具体表达式见第2.3.2节。我们使用了 $\epsilon^i_{jk} x^j \theta^k$ 这样的记号是因为要与(2.20)保持一致，因为若取 J 为(2.20)，正好就有

$$\epsilon^i_{jk} x^j \theta^k = (i\theta^k J_k x)^i.$$

则诺特定理导致的守恒流为

$$\begin{aligned} 0 &= \partial_0 \left(\frac{\partial \mathcal{L}}{\partial \partial_0 \phi^a} (i(J_k)^a_b \theta^k \phi^b - \epsilon^i_{jk} x^j \theta^k \partial_i \phi^a) \right) + \partial_i (\text{something}) \\ &= \theta^k \partial_0 \left(i(J_k)^a_b \phi^b \frac{\partial \mathcal{L}}{\partial \partial_0 \phi^a} - \epsilon^i_{jk} x^j \mathcal{P}_i \right) \\ &= \theta^k \partial_0 \left(i(J_k)^a_b \phi^b \frac{\partial \mathcal{L}}{\partial \partial_0 \phi^a} + \epsilon^i_{jk} x^j \mathcal{P}_i \right), \end{aligned}$$

于是我们就得到了一个守恒量

$$L_k = \int d^3x \left(i(J_k)^a_b \phi^b \frac{\partial \mathcal{L}}{\partial \partial_0 \phi^a} + \epsilon^i_{jk} x^j \mathcal{P}_i \right). \quad (2.89)$$

其中的

$$\mathcal{M}_k = \epsilon^i_{jk} x^j \mathcal{P}_i \quad (2.90)$$

部分和质点的角动量形式一致，称为轨道角动量，另一部分

$$\mathcal{S}_k = i(J_k)^a_b \phi^b \frac{\partial \mathcal{L}}{\partial \partial_0 \phi^a} \quad (2.91)$$

则不会出现在没有内部结构的质点的角动量内部。由于其形式和有自转的粒子除去轨道角动量以外的角动量一致，称其为自旋角动量。正如其名称暗示的那样，自旋角动量源自场

¹⁵ 在非相对论连续介质力学中这些结果也是成立的，因为时间和空间平移同时出现在伽利略群和庞加莱群中。

¹⁶ 同样，这个对称性无论是在相对论性场论还是非相对论性场论中都是成立的。

的内禀旋转自由度，也就是说，旋转群在场上的表示不仅仅含有无穷维的表示，还含有一个有限维的表示的时候就会出现自旋角动量。如果场是标量场，那么旋转群在其上的有限维表示就是平凡的，此时 J_i 全部为零，就没有自旋角动量。

以上推导在有相互作用时可能是不安全的，因为此时涉及多个场，且相互作用会引入额外的能量，从而难以良定义“一个场的能量”或是“一个场的动量”。

推动

推动操作同样对应着守恒流。

2.4.2 宇称、正反粒子变换和时间反演

时间反演变换造成的结果和宇称变换完全一样，除了具体的变换矩阵差一个负号。

2.5 单粒子态

2.5.1 单粒子态与庞加莱群的表示

在实际的物理系统中，如果我们能通过给一些可观测量安排一些实观测值从而完全确定某些系统，且这些系统直观上应该可以使用坐标（当然也可以是动量）加上一些比较简单的、通常是离散的标签标记，那么我们把把这些基础的系统称为基本粒子。

我们可能常常听到这样的说法：基本粒子是庞加莱群的不可约表示。通过上述讨论，我们知道对于一个量子系统来说，各种对称性变换构成的李群是以么正表示的方式（连续地、有时投影地）作用在系统的希尔伯特空间上的。具体来说，例如李群 G 中的一个元素 $g \in G$ ，考虑其么正表示 $g \rightarrow U(g)$ ，此群元同样可以自伴算子的方式作用在系统的一个可观测量 A 上： $A \rightarrow U(g)AU^\dagger(g)$ ，后者表示一个对称性变换 g 作用在系统上后，原来的可观测量 A 的形式。此变换既可看成 g 作用在系统本身上，也可看做作用在参考系上，但其视角并不妨碍我们的讨论。因此，虽然本文主要是关于相对论性量子场论的，绕过场的概念直接分析单粒子态仍然是有益的。

例如我们考虑最简单的体系，我们通过实验得到一些质量 m 的观测值，它们为非负实数；同时得到一系列自旋 s 的观测值，它们的可取值有 $\{1/2, 1, 3/2, \dots\}$ ，而不同的数组 (m, s) 意味着不同的基本粒子（ m 与 s 恰好为庞加莱群 \mathcal{P} 的卡西米尔元）。这些数字有如下性质：它们在（正时正规）庞加莱群的群作用下是不变的，这是显然而且必须的，因为同一类粒子的相同可观测量的观测值必然不依赖于参考系，而参考系之间的变换正是由庞加莱群所连接的。考虑量子情形，我们因此必然有相应的可观测量 M 表示质量与 S 表示自旋，并且它们在庞加莱群的群作用下是不变的，即：

$$U(g)MU^\dagger(g) = M, U(g)SU^\dagger(g) = S, \forall g \in \mathcal{P} \quad (2.92)$$

同时也必须满足有固定值，即

$$M = m\mathbf{I}, S = s\mathbf{I} \quad (2.93)$$

而Wigner证明了如果满足以上两个条件，那么表示 $g \rightarrow U(g)$ 是不可约的。¹⁷实际上，如果我们反过来看，由于 M, S 为 \mathcal{P} 的卡西米尔元，也即 \mathcal{P} 的李代数的泛包络代数中的元素，他们

¹⁷同时Wigner也证明了另一个非常重要的定理，即任何对称变换可以表示成希尔伯特空间上的算符，这个算符要么是线性且么正的，要么是反线性且反么正的，见Weinberg。

可以用群表示的自伴的生成元来定义，那么我们应有： $U(g)MU^\dagger(g) = M, U(g)SU^\dagger(g) = S, \forall g \in \mathcal{P}$ 。如果 U 是不可约表示，那么 $U(g)M = MU(g), U(g)S = SU(g)$ ，根据舒尔引理，我们自然有 $M = m\mathbf{I}, S = s\mathbf{I}$ ，其中 $s, m \in \mathbb{R}$ 。

Wigner根据 m 的本征值的取值将庞加莱群的非负能量不可约表示分为了三类（即Wigner's classification）：

- $m > 0$ ，即 $P = (m, 0, 0, 0)$ 被称为有质量态。有质量态被Spin(3)的不可约表示进一步分类（即自旋）。即：

$$P^2\Psi = m^2\Psi, W^2\Psi = m^2s(s+1)\Psi \quad (2.94)$$

- $m = 0, P_0 > 0$ 。这些态可以被进一步分为两类，其中一种表示被整数倍的 $1/2$ 的数字所分类，它们被称为螺旋度；而另一种则被称为“连续自旋”表示（continuous spin representation）。即：

$$P^2\Psi = 0, W^2\Psi = \mu^2\Psi \quad (2.95)$$

实际上，连续自旋表示虽然符合Wigner分类的所有要求，但是事实上我们并没有观测到连续自旋的无质量粒子¹⁸。我们也可以通过要求因果性来排除掉连续自旋表示作为实际粒子，但其证明都非常冗长¹⁹。

对于无质量粒子，其螺旋度等于手性(chirality)（有些约定中差一个 $\hbar/2$ ），也常常被称为自旋。但对于有质量粒子，其螺旋度被定义为自旋在动量方向上的投影。如果粒子的自旋和运动方向相同，那么其螺旋度被定义为右手的（正的），反之则是左手的。

但值得注意的是，如果对于有质量粒子，观测者有可能在某些参考系下运动地比粒子“快”，那么原本的螺旋度就会反向。因此我们可以看出，有质量粒子的螺旋度并非洛伦兹协变的，虽然它是一个运动常数，而对于有质量粒子，相应的洛伦兹协变的量为手性，但它并非一个运动常数。之前我们用自旋分类了有质量粒子，但其实可以用手性进一步分类。例如，对自旋 $1/2$ 的粒子来说，在旋转 2π 后，波函数会获得一个负号（旋转给波函数带上一个复相位），而粒子的手性所区分的就是在复平面上，这个复相位是如何到达 -1 的，即是“左手地”从 1 转到 -1 ，还是“右手地”从 1 转到 -1 。

如果我们考虑狄拉克方程(3.41)在手性基下的拉格朗日量：

$$\mathcal{L}_D = i\psi_L^\dagger \sigma^\mu \partial_\mu \psi_L + i\psi_R^\dagger \bar{\sigma}^\mu \partial_\mu \psi_R - m\psi_L^\dagger \psi_R - m\psi_R^\dagger \psi_L \quad (2.96)$$

其中 ψ_L 为手性为左的费米子（后简称左手费米子）， ψ_R 为手性为右的费米子，它们在质量项中可以相互转换。

- $m = 0, P^\mu = 0$ ，这种情况被称为真空，而并不被看成一种基本粒子。

注意，并不是所有基本粒子都可以被Wigner的分类法所分类，例如中微子的质量 m 是“震荡”的；对于夸克，其内在的对称性为 $SU(3)$ 对称性，此时庞加莱群就不能用来定义它们了，在描述它们时我们需要引入相应的规范群。

如果你有很多基本粒子，那么系统的希尔伯特空间自然是它们单独的希尔伯特空间的张量积，那么对应的变换也是相应的不可约表示的张量积，不过总表示就不再是不可约表示了。因此后面我们会着重讨论单粒子态，随后根据基本粒子的分类讨论相应的场，如 $s = 0$ 的标量场， $s = 1/2$ 旋量场， $s = 1$ 的矢量场等等。

¹⁸温伯格言：“are not observed to have”，见其量子场论P74。

¹⁹参考文献可见网页：<https://www.physicsoverflow.org/24751>中的推荐列表。

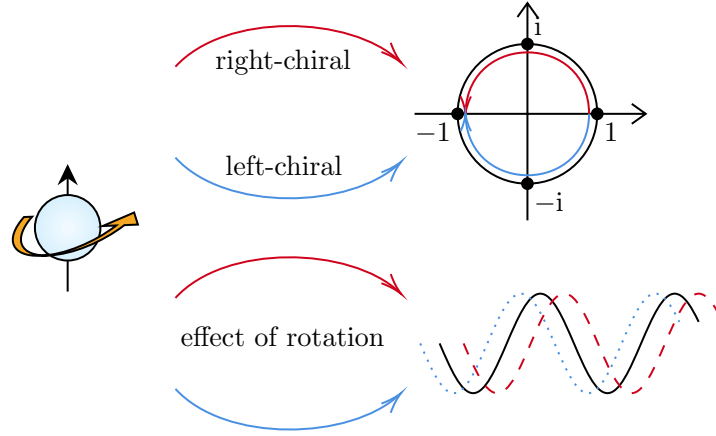


图 2.1: 当你旋转一个自旋为1/2的费米子时, 左手与右手费米子的相位变化表现不同, 反映在波函数的变化上就是偏置不同。

2.5.2 粒子的自由度

位置算符与动量算符

先考虑一维的位置算符 \hat{x} 。由于我们认为 \hat{x} 对应位置, 其谱为连续谱, 本征值没有上下界, 而是跑遍整个实数轴²⁰。于是, 我们考虑一个可以完全由 \hat{x} 描述的希尔伯特空间 \mathcal{H}_{1d} , 在其上我们可以写出

$$\hat{x} = \int dx x |x\rangle\langle x|, \quad (2.97)$$

有

$$\hat{x} |x'\rangle = x' |x'\rangle, \quad (2.98)$$

其中 $|x'\rangle$ 代表位置在 x' 的本征态。

很自然地, 我们考虑空间平移群导致的物理量。空间平移群是李群, 它在 \mathcal{H}_{1d} 上有么正表示, 则其生成元是厄米的, 从而是一个可观察量。设²¹

$$\hat{Q}(dx) = \hat{I} + \frac{1}{i} dx \hat{p}, \quad (2.99)$$

其中 \hat{p} 是一个不显含任何参量的厄米算符。容易看出它具有动量量纲。注意到

$$\begin{aligned} \hat{x} \hat{Q}(dx') |x'\rangle &= \hat{x} |x' + dx'\rangle = (x' + dx') |x' + dx'\rangle, \\ \hat{Q}(dx') \hat{x} |x'\rangle &= \hat{Q}(dx') x' |x'\rangle = x' \hat{Q}(dx') |x'\rangle = x' |x' + dx'\rangle, \end{aligned}$$

²⁰注意这是定义: 我们单纯构造了一个李代数, 仅此而已。这一步实际上并没有用到任何物理概念。

²¹容易看出这代表

$$\hat{p} = i \frac{\partial \hat{Q}}{\partial x},$$

刚好和一般的定义差了一个负号。这是因为平移群对“场算符”位置 \hat{x} 的作用是

$$\hat{x} \xrightarrow{Q_{\text{operator}}} \hat{x} + a,$$

使用同一个群参数 a , 上式按照(2.15)诱导出的在态矢量上的作用就是

$$|x\rangle \xrightarrow{Q_{\text{operator}}} |x - a\rangle.$$

然而, 我们通常希望空间平移群在态矢量上的作用是

$$|x\rangle \xrightarrow{Q_{\text{state}}} |x + a\rangle,$$

因此为了让 Q_{operator} 按(2.15)诱导出的在态矢量上的作用的生成元和 Q_{state} 的生成元一致, 在定义 Q_{state} 的生成元时我们特意加了一个负号。正文中的 \hat{Q} 指的都是 Q_{state} 。

就有

$$[\hat{x}, \hat{Q}(\mathrm{d}x')] |x'\rangle = \mathrm{d}x' |x' + \mathrm{d}x'\rangle \approx \mathrm{d}x' |x'\rangle.$$

考虑到 $|x'\rangle$ 的任意性, 我们得到

$$[\hat{x}, \hat{Q}(\mathrm{d}x')] = \left[\hat{x}, \hat{I} + \frac{1}{i} \mathrm{d}x \hat{p} \right] = \hat{I},$$

从而

$$[\hat{x}, \hat{p}] = i. \quad (2.100)$$

对易关系(2.100)完全确定了 \hat{x} 和 \hat{p} 的李代数结构。实际上, 完全可以使用更加简单的方法获得(2.100)。设一维空间平移群在 \mathcal{H}_{1d} 上的表示为 \hat{p} 。一维空间平移群对算符 \hat{x} 的作用的微分为

$$\frac{1}{i} \frac{\mathrm{d}}{\mathrm{d}a} ((x+a) - x) = -i,$$

从而由算符上的场表示和态空间上的表示之间的关系(2.15), 我们有

$$[\hat{p}, \hat{x}] = -i,$$

就得到了(2.100)。

我们来分析动量算符在坐标表象下的表示。我们有

$$\begin{aligned} (1 - i\hat{p}\mathrm{d}x) |\psi\rangle &= \hat{Q}(\mathrm{d}x') |\psi\rangle \\ &= \int \mathrm{d}x' \hat{Q}(\mathrm{d}x) |x'\rangle \langle x'|\psi\rangle \\ &= \int \mathrm{d}x' |x' + \mathrm{d}x\rangle \langle x'|\psi\rangle \\ &= \int \mathrm{d}x' |x'\rangle \langle x' - \mathrm{d}x|\psi\rangle \\ &= \int \mathrm{d}x' |x'\rangle \left(\langle x'|\psi\rangle - \mathrm{d}x \frac{\partial}{\partial x'} \langle x'|\psi\rangle \right) \\ &= |\psi\rangle - \mathrm{d}x \int \mathrm{d}x' |x'\rangle \frac{\partial}{\partial x'} \langle x'|\psi\rangle, \end{aligned}$$

从而

$$\hat{p} |\psi\rangle = -i \int \mathrm{d}x' |x'\rangle \frac{\partial}{\partial x'} \langle x'|\psi\rangle, \quad (2.101)$$

或者等价的,

$$\langle x|\hat{p}|\psi\rangle = -i \frac{\partial}{\partial x} \langle x|\psi\rangle. \quad (2.102)$$

当然, 完全可以使用空间平移群在无限维空间上的表示得到这个结果; 由注21, 虽然使用同一个群参数的空间平移群在算符 \hat{x} 和态 $|x\rangle$ 上的作用正好是相反的, 但由于定义 \hat{p} 时已经考虑了这一点, 从(2.17)就能够得到(2.102)。

从(2.102)可以导出位置的本征态和动量本征态之间的切换关系。取 $|\psi\rangle$ 为 $|p\rangle$, 我们有

$$\langle x|\hat{p}|p\rangle = -i \frac{\partial}{\partial x} \langle x|p\rangle = \langle x|p|p\rangle = p \langle x|p\rangle,$$

即微分方程

$$\frac{\partial}{\partial x} \langle x|p\rangle = ip \langle x|p\rangle,$$

从而

$$\langle x|p\rangle = C e^{ipx}.$$

由于总是可以在所有 $|p\rangle$ 上一起乘上一个模长为1的常数而不产生任何影响, 不失一般性地我们取 C 为实数。归一化条件为

$$\delta(p - p') = \langle p|p'\rangle = \int dx \langle p|x\rangle \langle x|p'\rangle,$$

从而计算得到

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi}} e^{ipx}. \quad (2.103)$$

这表明坐标表象和动量表象之间的变换是傅里叶变换。

下面转而讨论三维的情况。三维位置算符 $\hat{\mathbf{x}}$ 指的是

$$\hat{\mathbf{x}} = \hat{x}^1 \mathbf{e}_1 + \hat{x}^2 \mathbf{e}_2 + \hat{x}^3 \mathbf{e}_3, \quad (2.104)$$

其中为方便起见选取 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 为一组规范正交基。 $\hat{x}^1, \hat{x}^2, \hat{x}^3$ 的本征值均跑遍整条实数轴, 因此 $\hat{\mathbf{x}}$ 的本征值是 \mathbb{R}^3 中全体矢量。我们还需要加入另一个假设: $\hat{x}^1, \hat{x}^2, \hat{x}^3$ 彼此对易。这个假设要求这三个算符满足某种“独立性”。在一个完全能够由 $\hat{\mathbf{x}}$ 描述的希尔伯特空间 \mathcal{H}_{3d} 中我们写出其形式

$$\hat{\mathbf{x}} = \int d^3\mathbf{x} \mathbf{x} |\mathbf{x}\rangle \langle \mathbf{x}| = \int d^3\mathbf{x} (\hat{x}^1 \mathbf{e}_1 + \hat{x}^2 \mathbf{e}_2 + \hat{x}^3 \mathbf{e}_3) |x^1, x^2, x^3\rangle \langle x^1, x^2, x^3|. \quad (2.105)$$

沿着 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 的平移操作是对易的, 这就意味着

$$[\hat{p}_1, \hat{p}_2] = [\hat{p}_2, \hat{p}_3] = [\hat{p}_3, \hat{p}_1] = 0. \quad (2.106)$$

所有生成元彼此对易。这样三维空间平移群的李代数就被拆分成了三个一维空间平移群的李代数的直和, 三维空间平移群就被拆分成了三个一维空间平移群的直积。这三个一维空间平移群分别是 x^1 方向上的平移群 $\hat{Q}_1(a)$, x^2 方向上的平移群 $\hat{Q}_2(a)$, 以及 x^3 方向上的平移群 $\hat{Q}_3(a)$ 。另一方面, 注意到三维位置算符对应的本征态张成的空间 $\{|x^1, x^2, x^3\rangle\}_{x^1, x^2, x^3}$ 实际上是三个一维位置算符对应的本征态张成的空间的直积, 这是因为 $\hat{x}^1, \hat{x}^2, \hat{x}^3$ 彼此对易。于是我们做拆分

$$\mathcal{H}_{3d} = \mathcal{H}_{1d1} \otimes \mathcal{H}_{1d2} \otimes \mathcal{H}_{1d3},$$

并且用 \hat{x}^1 完全描述 \mathcal{H}_{1d1} , 用 \hat{x}^2 完全描述 \mathcal{H}_{1d2} , 用 \hat{x}^3 完全描述 \mathcal{H}_{1d3} 。由于 $\hat{Q}_1(a)$ 不改变 x^2, x^3 , 它在 $\mathcal{H}_{1d2}, \mathcal{H}_{1d3}$ 上没有作用。同样也可以对 $\hat{Q}_2(a), \hat{Q}_3(a)$ 做同样的论证。从而, 下标不一样的 \hat{x}^i 和 \hat{p}_i 彼此对易。这样关于诸 \hat{x} 和诸 \hat{p} 的李代数就可以拆分成 $\{\hat{x}^1, \hat{p}_1\}$ 、 $\{\hat{x}^2, \hat{p}_2\}$ 和 $\{\hat{x}^3, \hat{p}_3\}$ 三对量的李代数的直和。最后注意到在每个空间 \mathcal{H}_{1di} 中,

$$\hat{Q}_i(a) |x^i\rangle = |x^i + a\rangle,$$

于是我们可以原封不动地套用对一维动量和位置的论证, 得到

$$[\hat{x}^i, \hat{p}_i] = i\delta_j^i.$$

于是, 在没有做任何计算, 而只是观念性地拆分了态空间之后, 我们得到三维情况下的动量-位置对易关系:

$$[\hat{x}^i, \hat{x}^j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}^i, \hat{p}_j] = i\delta_j^i. \quad (2.107)$$

同样, 也可以通过(2.15)得到上式。同样, 套用一维动量和位置的论证, 我们得到动量算符在坐标表象下的形式

$$\langle \mathbf{x}|\hat{\mathbf{p}}|\psi\rangle = -i\nabla \langle \mathbf{x}|\psi\rangle, \quad (2.108)$$

以及相应的表象变换矩阵

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{p} \cdot \mathbf{x}}. \quad (2.109)$$

以上我们讨论了三维动量表象。但实际上，空间平移群是在闵可夫斯基时空中的，因此我们还可以讨论四维动量表象。所谓四维动量就是四维空间平移群的场表示按照(2.15)产生的、作用在态矢量上的生成元。它的空间部分就是我们已知的 $\hat{\mathbf{p}}$ ，它的时间部分 \hat{E} 满足

$$-i \frac{\partial \psi}{\partial t} = [\hat{E}, \psi],$$

其中 ψ 是某个场。因此 \hat{E} 就是能量。我们不说这是哈密顿量是因为它仅仅关于被讨论的单个粒子，而真正的哈密顿量应该关于整个体系。按照第2.3.1节中的描述，四维动量的四个分量彼此对易。因此，四维矢量算符 \hat{p}^μ 似乎应该是态空间

$$\{|p\rangle\}_p = \{|\mathbf{p}\rangle | E\rangle\}$$

的CSCO。然而，完整地描述单粒子的坐标空间实际上并不需要全部的 $\{|\mathbf{p}\rangle | E\rangle\}$ ——这是当然的，既然使用 $\hat{\mathbf{p}}$ 已经能够完整描述态空间了。事实上，考虑到(2.81)，我们有

$$\hat{E}^2 - \hat{\mathbf{p}}^2 = m^2, \quad (2.110)$$

如果态 $|p\rangle$ 描述的是坐标空间中的某个态，就可以把上式作用在 $|p\rangle$ 上：

$$\begin{aligned} m^2 |p\rangle &= \hat{E}^2 |p\rangle - \hat{\mathbf{p}}^2 |p\rangle \\ &= (E^2 - \mathbf{p}^2) |p\rangle, \end{aligned}$$

从而

$$m^2 = E^2 - \mathbf{p}^2 = p_\mu p^\mu.$$

为了保证 E 的单值性我们要求 $E > 0$ ，于是

$$m^2 = E^2 - \mathbf{p}^2 = p_\mu p^\mu, \quad E > 0. \quad (2.111)$$

我们看到，(2.111)在四维闵可夫斯基时空中选出了一个三维的子流形。我们称这个子流形为**质壳**。态 $|p\rangle$ 描述了一个实际的单粒子的坐标空间，当且仅当， p 在质壳(2.111)上，或者说 **p 在壳**。从而，给定一个 \mathbf{p} ，我们把使用质壳方程写出的对应能量记作

$$\omega_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}. \quad (2.112)$$

到现在为止这里的 m 都完全是一个来自群表示论的参数。然而，我们将在第3.2.1节中看到，这里的 m 和场的克莱因-高登方程中的 m 是同一个东西，而随后退化到非相对论情况后，我们会发现这个 m 实际上根本就是粒子质量。

由于(2.111)让我们能够从 \mathbf{p} 直接导出 p_0 也就是 E ，实际上完全可以使用 \mathbf{p} 来标记在质壳上的 $|p\rangle$ 。但我们并没有这么做，原因马上可以看到。现在我们来归一化 $|p\rangle$ 。因为我们在四维闵可夫斯基时空中工作，不再能够使用 $\int d^3\mathbf{p}$ 来做归一化了，因为它不满足洛伦兹协变性：做一个洛伦兹变换，就有可能把一部分 \mathbf{p} 弄到 p_0 中。因此我们需要在四维矢量空间上定义一个积分测度，它只能在质壳上给出非零值，且这个积分测度必须是洛伦兹标量。容易看出，

$$\int d^4p \delta(p^2 - m^2) \theta(p_0)$$

正是一个满足这种条件的积分测度。实际上，质壳是两个对顶的四维锥面，以上积分相当于对其中一个超锥面做积分。因此归一化条件为

$$1 = \int d^4p \delta(p^2 - m^2) \theta(p_0) |p\rangle\langle p|.$$

由于

$$\begin{aligned} \int d^4p \delta(p^2 - m^2) \theta(p_0) &= \int d^3\mathbf{p} \int dp_0 \delta(p_0^2 - \mathbf{p}^2 - m^2) \theta(p_0) \\ &= \int d^3\mathbf{p} \int dp_0 \left(\frac{\delta(p_0 - \omega_{\mathbf{p}})}{2\omega_{\mathbf{p}}} + \frac{\delta(p_0 + \omega_{\mathbf{p}})}{-2\omega_{\mathbf{p}}} \right) \theta(p_0) \\ &= \int \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}}, \end{aligned}$$

我们得到

$$\int \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}} |p\rangle\langle p| = 1.$$

另一方面，单粒子量子力学中动量本征态通常不是用上式归一化的。我们于是需要在这两种动量本征态之间建立一个关系。我们不失一般性地认为在壳的 $|p\rangle$ 就在单粒子量子力学的坐标空间中，且由于 $|p\rangle$ 与以它的空间部分 \mathbf{p} 为标记的态 $|\mathbf{p}\rangle$ 一一对应，不失一般性地要求 $|p\rangle$ 和 $|\mathbf{p}\rangle$ 差一个实数。例如，下式是一种常用的归一化方案：

$$\int d^3\mathbf{p} |\mathbf{p}\rangle\langle \mathbf{p}| = 1, \quad \langle \mathbf{q} | \mathbf{p} \rangle = \delta^3(\mathbf{q} - \mathbf{p}),$$

在此假设下就有

$$|p\rangle = \sqrt{2\omega_{\mathbf{p}}} |\mathbf{p}\rangle. \quad (2.113)$$

在三维动量空间中， $|p\rangle$ 是没有被归一化的；但在质壳中， $|\mathbf{p}\rangle$ 反而是没有被归一化的那个。为了保证洛伦兹协变性，在讨论相对论性量子场论时多用 $|p\rangle$ 而不是 $|\mathbf{p}\rangle$ ，虽然它们表示的物理状态完全是一样的。在坐标表象和 p 表象之间切换只需要使用

$$\langle \mathbf{x} | p \rangle = \sqrt{\frac{2\omega_{\mathbf{p}}}{(2\pi)^3}} e^{i\mathbf{p} \cdot \mathbf{x}}. \quad (2.114)$$

但实际上这个公式很少真的被使用，因为一来相对论情况下很少需要讨论粒子位置，二来从相对论动量表象也就是 p 表象切换回 \mathbf{x} 表象时需要使用积分测度 $\int d^3\mathbf{x}$ ，而这不是洛伦兹协变的。由于我们在实际计算时有时需要将时间单独拿出来当成一个“演化参数”来看待而不是把它当成坐标，通常不讨论 $\int d^3\mathbf{x}$ 的洛伦兹协变版本，从而从 p 表象切换回 \mathbf{x} 表象会有困难，我们也不去理会它。

在本文中，我们还需要对上式做一个小小的调整。如第3.1.1节所示，我们希望动量本征态可以直接由产生算符作用在真空态上产生，而产生算符是（坐标表象下的）场算符的傅里叶分量。将坐标表象下的场算符展开成傅里叶分量时需要对所有动量取值求和，而通常我们使用积分测度 $\int d^3\mathbf{p} / (2\pi)^3$ ，于是我们转而使用归一化方案

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}\rangle\langle \mathbf{p}| = 1, \quad \langle \mathbf{q} | \mathbf{p} \rangle = (2\pi)^3 \delta^3(\mathbf{q} - \mathbf{p}), \quad (2.115)$$

但是为了让

$$|p\rangle = \sqrt{2\omega_{\mathbf{p}}} |\mathbf{p}\rangle \quad (2.116)$$

仍然成立， $|p\rangle$ 的归一化可能看起来比较奇怪：

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} |p\rangle\langle p| = 1. \quad (2.117)$$

角动量

角动量算符就是旋转生成元在态矢量上的表示。有限维的旋转生成元是一些矩阵，它们实际上就是三个方向上的角动量算符在 \hat{J}_z 的本征态上的分量矩阵（为什么是 \hat{J}_z 是因为在导出具体的矩阵表达式时是使用 J_x 和 J_y 构造 J_z 的本征态的产生湮灭算符的）。与注21中提到的类似，如果旋转群 $\hat{R}(\phi)$ 是直接作用在态上的，那么定义

$$\hat{R}(\phi) = \hat{I} + \frac{1}{i}(\phi_1 \hat{J}^1 + \phi_2 \hat{J}^2 + \phi_3 \hat{J}^3), \quad (2.118)$$

而如果它是作用在算符上，那么定义

$$\hat{R}(\phi) = \hat{I} + i(\phi_1 \hat{J}^1 + \phi_2 \hat{J}^2 + \phi_3 \hat{J}^3).$$

两种定义给出的 \hat{J}^i 是一样的，只是群参数差了一个负号。这就是注21中提到的情况。

J^i 可以分解成两部分：一部分改变 \mathbf{x} 的取值，一部分不改变。改变 \mathbf{x} 取值的那一部分记作 L^i ，称为**轨道角动量**，它与粒子的位置自由度有关，因此有这个名称；不改变 \mathbf{x} 取值的那一部分记作 S^i ，称为**自旋角动量**，因为它和粒子的内禀自由度有关。 J^i 服从(2.21)，因此我们也称(2.21)为**角动量代数**。通常设

$$\mathbf{J} = J_1 \mathbf{e}_1 + J_2 \mathbf{e}_2 + J_3 \mathbf{e}_3, \quad (2.119)$$

则角动量代数就是

$$\mathbf{J} \times \mathbf{J} = i\mathbf{J}. \quad (2.120)$$

不过，实际计算时很少用到这个矢量，因为由于 \hat{J}^i 是三维旋转群的李代数的无穷维厄米表示，它们遵循角动量代数，因此彼此间不对易。 $\hat{\mathbf{J}}$ 没有本征态。但由角动量代数， $\hat{\mathbf{J}}^2$ 却和三个 \hat{J}^i 都对易。

需要注意的是，以上定义不能够一般地将轨道角动量和自旋角动量完全分开，因为可以任意地将自旋角动量划归一部分给轨道角动量。因此我们人为要求轨道角动量仅仅作用在坐标空间上，而自旋角动量仅仅作用在内禀自由度上。

按照(2.34)和(2.108)，我们有

$$\begin{aligned} \langle \mathbf{x} | \hat{L}^1 | \psi \rangle &= x^2 (-i\partial_3) \langle \mathbf{x} | \psi \rangle - x^3 (-i\partial_2) \langle \mathbf{x} | \psi \rangle \\ &= \langle \mathbf{x} | (\hat{x}^2 \hat{p}_3 - \hat{x}^3 \hat{p}_2) | \psi \rangle, \\ \hat{L}^1 &= \hat{x}^2 \hat{p}_3 - \hat{x}^3 \hat{p}_2. \end{aligned}$$

同样也可以这样计算出 \hat{L}^2 和 \hat{L}^3 。从而我们得到

$$\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}, \quad (2.121)$$

其中 $\hat{\mathbf{L}}$ 就是

$$\hat{L} = \hat{L}^i \mathbf{e}_i. \quad (2.122)$$

$$\hat{L}^2 = \hat{x}^2 \hat{p}^2 - (\hat{\mathbf{x}} \cdot \hat{\mathbf{p}})^2 + i\hat{\mathbf{x}} \cdot \hat{\mathbf{p}} = \hat{x}^2 \hat{p}^2 - (\hat{\mathbf{x}} \cdot \hat{\mathbf{p}})(\hat{\mathbf{p}} \cdot \hat{\mathbf{x}}). \quad (2.123)$$

总角动量 \hat{J}^i 和轨道角动量 \hat{L} 都遵循角动量代数，也即

$$\begin{aligned} [\hat{L}^i + \hat{S}^i, \hat{J}^j + \hat{S}^j] &= i\epsilon_{ijk}(\hat{J}^k + \hat{S}^k), \\ [\hat{L}^i, \hat{L}^j] &= i\epsilon_{ijk}\hat{L}^k, \end{aligned}$$

而由于轨道角动量和自旋角动量作用在态空间的不同自由度上，它们是彼此对易的，从而我们推导出

$$[\hat{S}^i, \hat{S}^j] = i\epsilon_{ijk}\hat{S}^k,$$

因此自旋角动量确实是一种角动量：它也满足角动量代数。

我们也可以定义

$$H = \hat{\mathbf{S}} \cdot \frac{\hat{\mathbf{P}}}{|\mathbf{P}|} \quad (2.124)$$

如前所述， J^2 和 \hat{J}_i 中的任何一个都对易。习惯上考虑 \hat{J}_3 ，我们会发现 \hat{J}^2 和 J_3 实际上构成了角动量代数中的一组CSCO，因为 \hat{J}^2 是角动量代数的卡西米尔元，它确定了 \hat{J}_3 的取值的上下限，再加上 \hat{J}_3 的取值，系统的角动量状况就完全确定了。通常称 \hat{L}^2 的本征值 $\sqrt{j(j+1)}$ 中的 j 为角量子数， \hat{L}_3 的本征值为磁量子数，而 \hat{S}_3 称为自旋量子数。 \hat{S}^2 没有专门的称呼，因为对种类确定的粒子它是固定的。

容易验证，若 \mathbf{J}_1 和 \mathbf{J}_2 都是角动量算符（也即，它们构成角动量代数），则它们之和 $\mathbf{J}_1 + \mathbf{J}_2$ 也是角动量算符。我们会看到将 \mathbf{J}_1 和 \mathbf{J}_2 叠加会得到一个可约表示， $\mathbf{J}_1 + \mathbf{J}_2$ 作用在 \mathbf{J}_1 和 \mathbf{J}_2 各自作用的空间的直积上，而这个直积实际上是一系列空间的直和。我们设

$$\hat{J}^2 |jm\rangle = j(j+1) |jm\rangle, \quad \hat{J} |jm\rangle = m |jm\rangle,$$

特意加上一个 j 的标记是因为这个表示可能是可约的。到现在为止 j 可以取哪些值都是完全不清楚的，不过无论如何，上式构成了 \hat{J} 的表示空间的一组基。另一方面，容易看出

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle$$

实际上也是 \hat{J} 的表示空间的一组基。容易验证 $|j_1 j_2 m_1 m_2\rangle$ 是 J_z 的一组本征态，且

$$m = m_1 + m_2, \quad (2.125)$$

但是并不是 \mathbf{J}^2 的一组本征态。两组基之间的变换系数 $\langle j_1 j_2 m_1 m_2 | jm \rangle$ 称为CG系数。 $\{|jm\rangle\}$ 实际上是由若干不可约表示直和而成的基矢量（由不同的 j 对应的表示直和而成），而 $\{|j_1 j_2 m_1 m_2\rangle\}$ 则是两个不可约表示直积而成的基矢量，CG系数用于在这两组基矢量来回切换。可以预期，CG系数应该能够反映(2.125)这一事实。的确如此：将用CG系数表示的 $|jm\rangle$ 代入本征方程得到

$$\sum_{m_1, m_2} (m - m_1 - m_2) \langle j_1 j_2 m_1 m_2 | jm \rangle |j_1 j_2 m_1 m_2\rangle = 0,$$

于是

$$(m - m_1 - m_2) \langle j_1 j_2 m_1 m_2 | jm \rangle = 0.$$

这就意味着

$$|jm\rangle = \sum_{m_1} |j_1 j_2 m_1 (m - m_1)\rangle \langle j_1 j_2 m_1 (m - m_1) | jm \rangle.$$

由于 m 是 m_1 和 m_2 之和， m 的最大值是 $j_1 + j_2$ ，最小值是 $-j_1 - j_2$ ，在此之间的所有值都可以取到。 j 的取值为

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2|.$$

在半经典的图景中，这意味着叠加两个态的角动量时，它们的夹角是量子化的。计算每个 j 对应的不可约表示的维度，把它们加起来，得到的结果正是 $(2j_1 + 1)(2j_2 + 1)$ 。

CG系数的具体计算需要用到 \mathbf{J} 的升降算符, 注意到

$$J_{\pm} |jm\rangle = \sum_{m_1+m_2=m} (J_{1\pm} |j_1 m_1\rangle |j_2 m_2\rangle + |j_1 m_1\rangle J_{2\pm} |j_2 m_2\rangle) \langle j_1 j_2 m_1 m_2 | jm \rangle, \quad (2.126)$$

首先可以用这个关系式计算出 $j = m$ 的CG系数, 注意到

$$J_+ |jm\rangle = 0,$$

就有可以递推得到

类似的方法还可以计算出三个角动量算符叠加而得到的结果, 乃至更多。

需要注意的是按照以上步骤将角动量算符叠加之后得到的角动量基矢量并没有对称化/反对称化。反对称化可能会让一些状态实际上取不到。

单粒子态与洛伦兹变换

前文已经证明过, 四动量 \hat{p}^μ 的分量全部对易, 因此可以构成空间的一个CSCO, 下面我们尝试区分除了动量之外的其他自由度。记其他自由度为 σ , 那么我们有

$$P^\mu |\psi, p, \sigma\rangle = p^\mu |\psi, p, \sigma\rangle,$$

由于我们只考虑单粒子态, 我们可以要求 σ 是离散指标²²。

下面我们需要考察态 $|\psi, p, \sigma\rangle$ 在一般的齐次洛伦兹变换下如何变换。根据

$$U(\Lambda, a) P^\rho U^{-1}(\Lambda, a) = \Lambda_\mu{}^\rho P^\mu,$$

我们知道在齐次洛伦兹变换 $U(\Lambda, 0) \equiv U(\Lambda)$ 作用下, 原本动量为 p 的本征态 $|\psi, p, \sigma\rangle$ 变为动量为 Λp 的本征态:

$$\begin{aligned} P^\mu U(\Lambda) |\psi, p, \sigma\rangle &= U(\Lambda) [U^{-1}(\Lambda) P^\mu U(\Lambda)] |\psi, p, \sigma\rangle = U(\Lambda) \Lambda^\mu{}_\rho P^\rho |\psi, p, \sigma\rangle \\ &= \Lambda^\mu{}_\rho p^\rho U(\Lambda) |\psi, p, \sigma\rangle \end{aligned}$$

因此, 我们可以以 $|\psi, \Lambda p, \sigma'\rangle$ 作为基, 展开原来的态 $U(\Lambda) |\psi, p, \sigma\rangle$:

$$U(\Lambda) |\psi, p, \sigma\rangle = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda, p) |\psi, \Lambda p, \sigma'\rangle.$$

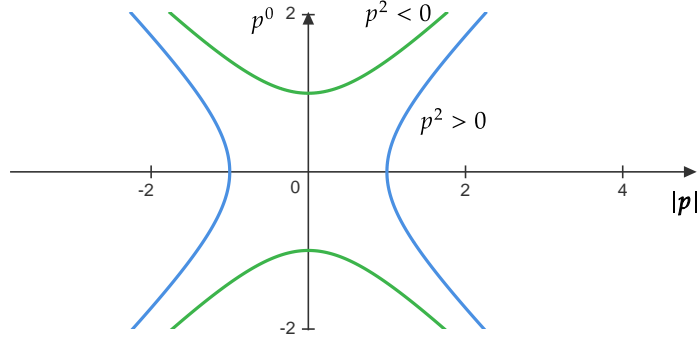
如果我们选择合适的指标 σ , 有可能使得 $C_{\sigma'\sigma}(\Lambda, p)$ 为分块对角阵, 即对不同的 σ , 每一个块构成非齐次洛伦兹群的不可约表示。因此我们再一次看到, 我们可以将粒子态与非齐次洛伦兹群的不可约表示的分量等同起来。接下来, 我们自然需要解出 $C_{\sigma'\sigma}(\Lambda, p)$ 在非齐次洛伦兹群不可约表示中的结构。

如果要解出洛伦兹群不可约表示的结构, 我们首先需要对洛伦兹变换根据 p 进行初步分类。考虑在任何固有正时洛伦兹变换 $\Lambda^\mu{}_\nu$ 下保持不变的 p 的函数, 显然 p^2 保持不变:

$$p^2 = p^\mu p_\mu \rightarrow \Lambda^\mu{}_\nu p^\nu \Lambda_\mu{}^\sigma p_\sigma = p^\nu p_\nu = p^2.$$

同时还有 p^0 的符号, 注意这个量并非在什么情况下都不变, 只有对于类时或者类光的 p 才成立, 对于类空的 p , p^0 的符号有可能改变。考虑 $p^2 = -(p^0)^2 + |\mathbf{p}|^2 = \text{const.}$ 的曲线 $\gamma(|\mathbf{p}|, p^0)$, 显然为两支互不联通的双曲线, 如图2.5.2所示。

²²对于一般的态, 如几个非束缚态粒子构成的态, σ 可以是连续的。



由于固有正时洛伦兹变换与单位元连通，因此只能从在图2.5.2上可以看出对于 $p^2 < 0$ 的曲线，两支不能通过 Λ^μ_ν 互相转换，即 p^0 不能改变符号；而对于 $p^2 > 0$ 的曲线，由于在 $|\mathbf{p}| > 0$ 的部分是连通的，因此 p^0 的符号可以改变。

从而我们可以利用 p^2 的值以及（ $p^2 \leq 0$ 时）的 p^0 的符号来选择一个“标准动量” k^μ ，使得

$$p^\mu = L^\mu_\nu(p) k^\nu,$$

其中 $L^\mu_\nu(p)$ 为依赖于 p^μ 的洛伦兹变换。从而我们可以将动量为 p 的态 $|\psi, p, \sigma\rangle$ 定义为

$$|\psi, p, \sigma\rangle \equiv N(p) U(L(p)) |\psi, k, \sigma'\rangle$$

其中 $N(p)$ 为归一化系数，稍后我们会证明 $N(p)$ 可以取

$$N(p) = \sqrt{\frac{k^0}{p^0}},$$

使得

$$\langle \psi, p', \sigma' | \psi, p, \sigma \rangle = \delta_{\sigma'\sigma} \delta^3(\mathbf{p}' - \mathbf{p}).$$

如果我们将齐次洛伦兹变换 $U(\Lambda)$ 作用在态 $|\psi, p, \sigma\rangle$ 上，有

$$U(\Lambda) |\psi, p, \sigma\rangle = N(p) U(\Lambda L(p)) |\psi, k, \sigma\rangle = N(p) U(L(\Lambda p)) U(L^{-1}(\Lambda p) \Lambda L(p)) |\psi, k, \sigma\rangle$$

注意到变换 $L^{-1}(\Lambda p) \Lambda L(p)$ 先将 k 变为 p ，再变为 Λp ，最后再变回 k ，因此保持 k^μ 不变。这样的保持 k^μ 不变的洛伦兹变换 W^μ_ν 构成齐次洛伦兹群的子群，被称为小群（little group）。对于这样的 W ，我们有

$$U(W) |\psi, k, \sigma\rangle = \sum_{\sigma'} D_{\sigma'\sigma}(W) |\psi, k, \sigma'\rangle, \quad (2.127)$$

这样的 $D(W)$ 给出了小群的一个表示。那么洛伦兹变换作用在单粒子态上的效果就变成了

$$U(\Lambda) |\psi, p, \sigma\rangle = N(p) U(L(p)) \sum_{\sigma, \sigma'} D_{\sigma, \sigma'}(W(\Lambda, p)) |\psi, k, \sigma'\rangle.$$

根据 $|\psi, k, \sigma\rangle$ 的定义，我们有

$$\begin{aligned} U(\Lambda) |\psi, p, \sigma\rangle &= \frac{N(p)}{N(\Lambda p)} \sum_{\sigma, \sigma'} D_{\sigma, \sigma'}(W(\Lambda, p)) |\psi, \Lambda p, \sigma'\rangle \\ &= \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma, \sigma'} D_{\sigma, \sigma'}(W(\Lambda, p)) |\psi, \Lambda p, \sigma'\rangle, \end{aligned} \quad (2.128)$$

表 2.1: 不同 p^2 以及 p^0 的符号给出标准 k^μ 及其对应的小群

	standard k^μ	little group
(a) $p^2 = -M^2 < 0, p^0 > 0$	$(M, 0, 0, 0)$	$SO(3)$
(b) $p^2 = -M^2 < 0, p^0 < 0$	$(-M, 0, 0, 0)$	$SO(3)$
(c) $p^2 = 0, p^0 > 0$	$(\kappa, 0, 0, \kappa)$	$E(2)$
(d) $p^2 = 0, p^0 < 0$	$(-\kappa, 0, 0, \kappa)$	$E(2)$
(e) $p^2 = N^2 > 0$	$(0, 0, 0, N)$	$SO(2, 1)$
(f) $p^\mu = 0$	$(0, 0, 0, 0)$	$SO(3, 1)$

这意味着原先的解出 $C_{\sigma'\sigma}(\Lambda, p)$ 在非齐次洛伦兹群不可约表示中的结构的问题现在变成了寻找小群表示 $D(W(\Lambda, p))$ 的问题。在表2.1中, 我们给出了一种标准 k^μ 的选择以及不同 p^2, p^0 的值给出的小群。我们的任务就是针对不同的情况, 找出相应的小群表示。这被称为诱导表示法。

这几种情况中, 只有(a),(c),(f)已知有物理解释, 而(f)则是真空, 因此我们只需要考虑(a), 即正质量粒子的情况以及(c), 即零质量粒子的情况。

下面我们考虑归一化系数的问题。首先, 我们可以很自然地令 k^μ 标记的态是正交的, 即

$$\langle \psi, k', \sigma' | \psi, k, \sigma \rangle = \delta^3(\mathbf{k}' - \mathbf{k}) \delta_{\sigma'\sigma}.$$

从这里我们可以看出来, 小群表示 $D(W)$ 一定是幺正的, 因为

$$\begin{aligned} \langle \psi, k', \sigma' | \psi, k, \sigma \rangle &= \langle \psi, k', \sigma' | U^\dagger(\Lambda) U(\Lambda) | \psi, k, \sigma \rangle \\ &= \sum_{\rho, \rho'} \left\langle \psi, k', \rho' \left| D_{\rho'\sigma'}^\dagger(W) D_{\rho\sigma}(W) \right| \psi, k, \rho \right\rangle \\ &= \sum_{\rho, \rho'} D_{\rho'\sigma'}^\dagger(W) D_{\rho\sigma}(W) \delta^3(\mathbf{k}' - \mathbf{k}) \delta_{\sigma'\sigma} \\ &= \sum_{\rho, \rho'} D_{\rho'\sigma}^\dagger(W) D_{\rho\sigma}(W) \delta^3(\mathbf{k}' - \mathbf{k}) \stackrel{!}{=} \delta^3(\mathbf{k}' - \mathbf{k}), \end{aligned}$$

即这意味着

$$D^\dagger(W) = D^{-1}(W).$$

如果我们计算上述内积, 可以得到:

$$\langle \psi, p', \sigma' | \psi, p, \sigma \rangle = N(p) \langle U^{-1}(L(p)) \psi, p', \sigma' | \psi, k, \sigma \rangle = N(p) N^*(p') D(W(L^{-1}(p), p'))_{\sigma\sigma'}^* \delta^3(\mathbf{k}' - \mathbf{k})$$

这里 $k' \equiv L^{-1}(p)p'$ 。由于 $k = L^{-1}(p)p$, 根据 δ 函数的性质

$$\delta(F(x)) = \sum_{i=1}^m \frac{\delta(x - a_i)}{|F'(a_i)|}, F(a_i) = 0,$$

我们有

$$\delta^3(\mathbf{k} - \mathbf{k}') \propto \delta^3(\mathbf{p} - \mathbf{p}'),$$

因此随后我们需要解出比例因子。注意到对于满足 $p^2 = -M^2 \geq 0, p^0 > 0$ 的区域上, 标量函

数 $f(p)$ 可以写为

$$\begin{aligned}\int d^4p \delta(p^2 + M^2) \theta(p^0) f(p) &= \int d^3\mathbf{p} dp^0 \delta((p^0)^2 - \mathbf{p}^2 - M^2) \theta(p^0) f(\mathbf{p}, p^0) \\ &= \int d^3\mathbf{p} \frac{f(\mathbf{p}, \sqrt{\mathbf{p}^2 + M^2})}{2\sqrt{\mathbf{p}^2 + M^2}},\end{aligned}$$

即在 $p^2 = -M^2$ 上积分时，不变体元实际上为

$$d^3\mathbf{p} \frac{f(\mathbf{p}, \sqrt{\mathbf{p}^2 + M^2})}{\sqrt{\mathbf{p}^2 + M^2}}.$$

根据 δ 函数的定义：

$$\begin{aligned}F(\mathbf{p}) &= \int F(\mathbf{p}') \delta^3(\mathbf{p} - \mathbf{p}') d^3\mathbf{p}' \\ &= \int F(\mathbf{p}') \left[\sqrt{\mathbf{p}'^2 + M^2} \delta^3(\mathbf{p}' - \mathbf{p}) \right] \frac{d^3\mathbf{p}'}{\sqrt{\mathbf{p}'^2 + M^2}} \\ &= \int F(\mathbf{p}') [p^0 \delta^3(\mathbf{p}' - \mathbf{p})] \frac{d^3\mathbf{p}'}{\sqrt{\mathbf{p}'^2 + M^2}},\end{aligned}$$

因而我们有

$$p^0 \delta^3(\mathbf{p}' - \mathbf{p}) = k^0 \delta^3(\mathbf{k}' - \mathbf{k}),$$

从而我们选择归一化因子

$$N(p) = \sqrt{\frac{k^0}{p^0}},$$

从而我们有

$$\langle \psi, \mathbf{p}', \sigma' | \psi, \mathbf{p}, \sigma \rangle = |N(p)|^2 \delta_{\sigma'\sigma} \left(\frac{p^0}{k^0} \right) \delta^3(\mathbf{p}' - \mathbf{p}) = \delta_{\sigma'\sigma} \delta^3(\mathbf{p}' - \mathbf{p}).$$

下面我们可以考虑正质量粒子和零质量粒子给出的小群表示。

正质量粒子

正质量粒子满足的情况即 $p^2 = -M^2$ ，其标准 $k^\mu = (M, 0, 0, 0)$ ，保持 k^μ 不变的洛伦兹变换构成的小群为 $SO(3)$ 。我们知道 $SO(3)$ 的表示可以用 $D_{\sigma'\sigma}^{(j)}(R)$ 来分类，其中 j 为正整数或正半整数为表示的维数。具体的表示可以用无穷小旋转 $R_{ik} = \delta_{ik} + \omega_{ik}$ 进行构造，其中 $\omega_{ik} = -\omega_{ki}$ 为无穷小量：

$$\begin{aligned}D_{\sigma'\sigma}^{(j)}(1 + \omega) &= \delta_{\sigma'\sigma} + \frac{i}{2} \omega_{ik} (J_{ik}^{(j)})_{\sigma'\sigma} \\ (J_{23}^{(j)} \pm i J_{31}^{(j)})_{\sigma\beta\sigma} &= (J_1^{(j)} \pm i J_2^{(j)}) = \delta_{\sigma', \sigma \pm 1} \sqrt{(j \mp \sigma)(j \mp \sigma) + 1} \\ (J_{12}^{(j)})_{\sigma\beta\sigma} &= (J_3^{(j)})_{\sigma'\sigma} = \sigma \delta_{\sigma'\sigma},\end{aligned}$$

其中 $\sigma = -j, -j+1, \dots, j$ 。对于 $M > 0$ ，自旋为 j 的粒子，齐次洛伦兹变换作用在此单粒子态上的效果为

$$U(\Lambda) |\psi, p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D_{\sigma'\sigma}^{(j)}(W(\Lambda, p)) |\psi, \Lambda p, \sigma'\rangle, \quad (2.129)$$

其小群元 $W(\Lambda, p)$ 即为Wigner转动:

$$W(\Lambda, p) = L^{-1}(\Lambda p) \Lambda L(p). \quad (2.130)$$

如果我们能证明对于任意的三维旋转 \mathcal{R} , $W(\mathcal{R}, p) = \mathcal{R}$, 就是说对于一个运动的有质量粒子的态, 它在旋转下的变换与非相对论量子力学中相同, 那么我们可以将幼儿园学过的球谐函数, CG系数等工具直接照搬过来。

下面我们来证明这件事。首先我们需要选择一个“标准推动” $L(p)$, 使得 k^μ 变为 p^μ 。我们直接给出这样一个构造:

$$\begin{aligned} L^i_k(p) &= \delta_{ik} + (\gamma - 1)\hat{p}_i\hat{p}_k, \\ L^i_0(p) &= L^0_i(p) = \hat{p}_i\sqrt{\gamma^2 - 1}, L^0_0(p) = \gamma, \end{aligned} \quad (2.131)$$

其中

$$\hat{p}_i \equiv \frac{p_i}{|\mathbf{p}|}, \gamma \equiv \frac{\sqrt{\mathbf{p}^2 + M^2}}{M}.$$

我们可以验证对于 $k^\mu = (M, 0, 0, 0)$, 我们有:

$$\begin{aligned} p^0 &= L^0_\nu(p)k^\nu = \gamma M + \hat{p}_i\sqrt{\gamma^2 - 1} \cdot 0 = \sqrt{\mathbf{p}^2 + M^2}, \\ p^i &= L^i_\nu(p)k^\nu = \hat{p}^i\sqrt{\gamma^2 - 1} \cdot M = \frac{p^i}{|\mathbf{p}|} \sqrt{\frac{\mathbf{p}^2}{M^2}} \cdot M = p^i, \end{aligned}$$

即 $p^\mu = (\sqrt{\mathbf{p}^2 + M^2}, p^i)$, 满足 $p^\mu p_\mu = -\mathbf{p}^2 - M^2 + \mathbf{p}^2 = -M^2$ 。为什么 $L(p)$ 是一个推动? 我们实际上可以发现下列关系:

$$L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}}), \quad (2.132)$$

其中 $R(\hat{\mathbf{p}})$ 为令 z 轴指向 \mathbf{p} 方向的旋转, $B(|\mathbf{p}|)$ 为我们熟悉的推动的标准形式:

$$B(|\mathbf{p}|) = \begin{pmatrix} \gamma & 0 & 0 & \sqrt{\gamma^2 - 1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sqrt{\gamma^2 - 1} & 0 & 0 & \gamma \end{pmatrix}.$$

这个关系是显然的, 因为在令 \mathbf{p} 只有 z 轴分量后, $\hat{p}_1 = \hat{p}_2 = 0, \hat{p}_3\hat{p}_3 = 1$, 将其带入 $L(p)$ 的表达式就可以获得 $B(|\mathbf{p}|)$ 的形式。那么对于任意旋转 \mathcal{R} , 我们有

$$W(\mathcal{R}, p) = R(\mathcal{R}\hat{\mathbf{p}})B^{-1}(|\mathbf{p}|)R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}})$$

但由于 $R(\hat{\mathbf{p}})$ 将 z 轴转到 $\hat{\mathbf{p}}$ 方向, 再转到 $\mathcal{R}\hat{\mathbf{p}}$ 方向, 最后再转回原来的方向, 因此 $R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})$ 一定是朝第三轴某个方向的旋转, 可以写为:

$$R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}}) = R(\theta) \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

而 $R(\theta)$ 与 $B(|\mathbf{p}|)$ 对易, 因此

$$W(\mathcal{R}, p) = R(\mathcal{R}\hat{\mathbf{p}})B^{-1}(|\mathbf{p}|)R(\theta)B(|\mathbf{p}|)R^{-1}(\hat{\mathbf{p}}) = R(\mathcal{R}\hat{\mathbf{p}})R(\theta)R^{-1}(\hat{\mathbf{p}}),$$

即

$$W(\mathcal{R}, p) = \mathcal{R}.$$

从而我们可以和非相对论量子力学一样对待有质量粒子的旋转。

零质量例子

对于无质量粒子，我们取标准的4-动量为 $k^\mu = (1, 0, 0, 1)$ ，小群元 $W^\mu{}_\nu$ 满足 $W^\mu{}_\nu k^\nu = k^\mu$ 。下面我们需要解 W 的具体形式，因此考虑 W 作用在类时矢量 $t^\mu = (1, 0, 0, 0)$ 上，由于 W 作为洛伦兹变换满足 $|W|^2 = 1$ ，因此我们有

$$\begin{aligned} t^\mu t_\mu &= |t|^2 = |Wt|^2 = (Wt)^\mu (Wt)_\mu = -1 \\ t^\mu k_\mu &= (Wt)^\mu (Wk)_\mu = (Wt)^\mu k_\mu = -1. \end{aligned}$$

满足第二个条件的4-矢可写成

$$(Wt)^\mu = (1 + \zeta, \alpha, \beta, \zeta),$$

因为显然满足

$$(Wt)^\mu k_\mu = -1 - \zeta + \zeta = -1.$$

第一个条件给出约束

$$\zeta = \frac{\alpha^2 + \beta^2}{2},$$

因此可以知道 $W^\mu{}_\nu$ 作用在 t^ν 上的效果与 $S^\mu{}_\nu$ 作用在 t^ν 上的效果相同：

$$W^\mu{}_\nu t^\nu = \begin{pmatrix} 1 + \zeta \\ \alpha \\ \beta \\ \zeta \end{pmatrix} = \begin{pmatrix} 1 + \zeta & \alpha & \beta & -\zeta \\ \alpha & 1 & 0 & -\alpha \\ \beta & 0 & 1 & -\beta \\ \zeta & \alpha & \beta & 1 - \zeta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \equiv S^\mu{}_\nu(\alpha, \beta) t^\nu$$

显然， $S^\mu{}_\nu k^\nu = k^\mu$ ， $|S| = 1$ ，且满足 $S^\mu{}_\sigma S^\nu{}_\rho \eta^{\sigma\rho} = \eta^{\mu\nu}$ 。但这并不意味着 $W^\mu{}_\nu = S^\mu{}_\nu$ ，不过至少我们有

$$(S^{-1}W)t = t,$$

即保时间分量不变，因此 $S^{-1}W$ 为一个纯旋转；同时由于 $S^{-1}W$ 保 $k^\mu = (1, 0, 0, 1)$ 不变，因此为绕 z 轴方向的纯旋转，可以写成

$$S^{-1}(\alpha, \beta)W = R(\theta), R(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

从而我们可以写出小群元 W 的一般形式：

$$W(\theta, \alpha, \beta) = S(\alpha, \beta)R(\theta).$$

下面我们看看这个群的结构：注意对子群 $\alpha = \beta = 0$ 或者 $\theta = 0$ 都是阿贝尔的，同时 $\theta = 0$ 的子群为不变子群，即

$$R(\theta)S(\alpha, \beta)R^{-1}(\theta) = S(\alpha \cos \theta + \beta \sin \theta, -\alpha \sin \theta + \beta \cos \theta).$$

因此我们可以看出这其实是 $E(2)$ 的乘法规则，即平移（ α, β 做参数）和旋转（ θ 做参数）构成的群。下面我们考虑 $E(2)$ 的李代数：

$$W(\theta, \alpha, \beta)^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, \omega_{\mu\nu} = \begin{pmatrix} 0 & -\alpha & -\beta & 0 \\ \alpha & 0 & \theta & -\alpha \\ \beta & -\theta & 0 & -\beta \\ 0 & \alpha & \beta & 0 \end{pmatrix}.$$

那么相应的希尔伯特空间的算符是

$$U(W(\theta, \alpha, \beta)) = 1 + i\alpha A + i\beta B + i\theta J_3,$$

其中

$$\begin{aligned} A &= -J^{13} + J^{10} = J_2 + K_1, \\ B &= -J^{23} + J^{20} = -J_1 + K_2. \end{aligned}$$

其对易关系易得

$$\begin{aligned} [J_3, A] &= +iB \\ [J_3, B] &= -iA \\ [A, B] &= 0. \end{aligned}$$

由于 A, B 对易, 可以被同时对角化, 因此我们有

$$\begin{aligned} A|\psi, a, b\rangle &= a|\psi, a, b\rangle \\ B|\psi, a, b\rangle &= b|\psi, a, b\rangle. \end{aligned}$$

容易计算出

$$\begin{aligned} U[R(\theta)]AU^{-1}[R(\theta)] &= A\cos\theta - B\sin\theta, \\ U[R(\theta)]BU^{-1}[R(\theta)] &= A\cos\theta + B\sin\theta, \end{aligned}$$

因而我们考虑 A, B 作用在 $|\psi, k, \alpha, b, \theta\rangle \equiv U^{-1}(R(\theta))|\psi, k, a, b\rangle$ 上, 有

$$\begin{aligned} A|\psi, a, b, \theta\rangle &= (a\cos\theta - b\sin\theta)|\psi, k, a, b, \theta\rangle \\ B|\psi, a, b, \theta\rangle &= (a\cos\theta + b\sin\theta)|\psi, k, a, b, \theta\rangle. \end{aligned}$$

如果 a, b 不为0, 那么我们会得到一个完全连续的谱, 这被称为连续自旋表示。但实验上我们并没有发现像 θ 这样的自由度, 因此我们要求 $a = b = 0$, 即

$$A|\psi, k, \sigma\rangle = B|\psi, k, \sigma\rangle = 0.$$

那么这些态可以通过剩下的唯一的自由度进行区分:

$$J_3|\psi, k, \sigma\rangle = \sigma|\psi, k, \sigma\rangle,$$

其中 σ 给出角动量在运动方向上的分量, 或者称其为螺旋度。

那么随后我们就可以考察一般无质量态在洛伦兹变换下的变换性质。由于 A, B 对易, 因此任意一个小群群元可以写成

$$U(W)|\psi, k, \sigma\rangle = \exp(i\alpha A + i\beta B)\exp(i\theta J_3)|\psi, k, \sigma\rangle = \exp(i\theta\sigma)|\psi, k, \sigma\rangle.$$

从而, 根据方程2.127, 我们可以写出

$$D_{\sigma'\sigma}(W) = \exp(i\theta\sigma)\delta_{\sigma'\sigma}.$$

根据之前推出的洛伦兹变换作用在态上以小群表示作为系数的方程2.128, 我们有

$$U(\Lambda)|\psi, p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}}\exp(i\sigma\theta(\Lambda, p))|\psi, \Lambda p, \sigma\rangle, \quad (2.133)$$

其中 $\theta(\Lambda, p)$ 定义为

$$W(\Lambda, p) \equiv L^{-1}(\Lambda p)\Lambda L(p) \equiv S(\alpha(\Lambda, p), \beta(\Lambda, p))R(\theta(\Lambda, p)).$$

后面我们会看到，电磁理论的规范不变性来自于小群中被 α 和 β 参数化的部分。目前 σ 还可以取任意实数，但是某些拓扑上的原因让 σ 与有质量粒子相同，只能取整数或者半整数。

我们定义

$$P_{\pm} = A \pm iB,$$

那么根据之前的对易关系，我们可以推出

$$[J_3, P_{\pm}] = P_{\pm}, [P_{\pm}, P_{\pm}] = 0.$$

由于 J_3 为 $SO(2)$ 的生成元，而 $SO(2)$ 的表示可以用一个整数 m 来刻画：

$$D^m(R(\theta)) = e^{im\theta}.$$

记表示 D^m 的表示空间的基为 $|m\rangle$ ，那么群元素对态的作用自然可以表示为

$$e^{i\theta J_3} |m\rangle = D^m(e^{i\theta J_3}) |m\rangle = e^{im\theta} |m\rangle \Rightarrow J_3 |m\rangle = m |m\rangle,$$

即 $|m\rangle$ 为 J_3 的本征态，也因此 m 刻画了 $SO(2)$ 的不可约表示。对于 $E(2)$ ，我们可以用 $SO(2)$ 的不可约表示生成 $E(2)$ 的忠实不可约表示。考虑 P_{\pm} 作用在 $|m\rangle$ 上（此时 $|m\rangle$ 被看做 $E(2)$ 的表示空间的一个基）：

$$J_3(P_{\pm} |m\rangle) = ([J_3, P_{\pm}] + P_{\pm} J_3) |m\rangle = \pm P_{\pm} |m\rangle + P_{\pm} m |m\rangle = (m \pm 1) P_{\pm} |m\rangle,$$

即 $P_{\pm} |m\rangle$ 为 J_3 的本征值为 $m \pm 1$ 的本征态，即 $P_{\pm} |m\rangle = |m \pm 1\rangle$ 。因此，当我们把 D^m 当做 $E(2)$ 的表示时，他们不再是单个独立的表示，而通过升降算符 P_{\pm} 所连接，即所有可能的 D^m 都被单个纳入单个的 $E(2)$ 的不可约表示，并且其基为 $\{|m\rangle | m \in \mathbb{Z}\}$ 。

因此我们用 J_3 作为小群的生成元后，其作用在单粒子态上的作用就是绕动量方向旋转，而由于在拓扑上，洛伦兹变换的底流形——非齐次洛伦兹群同构于 $\mathbb{R}_4 \times \mathbb{R}_3 \times S_3/\mathbb{Z}_2$ ，其并非单连通群，而绕动量方向角度为 4π 的旋转与单位元相同，因此 $\exp(4\pi i\sigma) = 1$ ，从而 σ 必须是整数或者半整数。

下面我们计算给定 Λ, p 后的小群群元 $W(\Lambda, p)$ 。首先我们要选一个能将 $k^{\mu} = (\kappa, 0, 0, \kappa)$ 变为原始 p^{μ} 的标准洛伦兹变换：

$$L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|/\kappa),$$

其中 $B(u)$ 为沿 z 方向的推动：

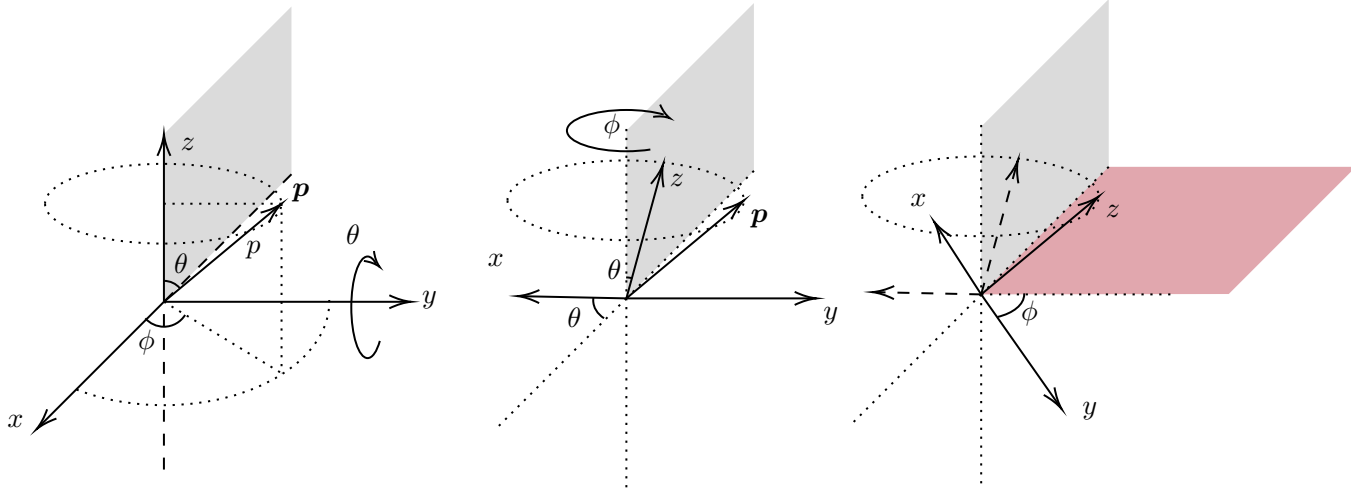
$$B(u) = \begin{pmatrix} (u^2 + 1)/2u & 0 & 0 & (u^2 - 1)/2u \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ (u^2 - 1)/2u & 0 & 0 & (u^2 + 1)/2u \end{pmatrix},$$

这里 $u = \gamma \pm \sqrt{\gamma^2 - 1} = \gamma \pm \beta$ 。而 $R(\hat{\mathbf{p}})$ 是使 z 轴转向 $\hat{\mathbf{p}}$ 方向的转动，例如 $\hat{\mathbf{p}} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ ，那么我们可以将 $R(\hat{\mathbf{p}})$ 取成绕第二轴旋转 θ ，从而 $(0, 0, 1)$ 变为 $(\sin \theta, 0, \cos \theta)$ ，然后再绕第三轴旋转角度 ϕ 得到 $\hat{\mathbf{p}}$ ，如图2.2所示

于是我们有

$$U(R(\hat{\mathbf{p}})) = \exp(-i\phi J_3) \exp(-i\theta J_2).$$

当我们选择了 $L(p)$ 后，根据公式 $W(\Lambda, p) \equiv L^{-1}(\Lambda p)\Lambda L(p)$ 我们便可直接得到 $W(\Lambda, p)$ 的表达式。

图 2.2: 将 z 轴转到 p 方向的标准旋转

注意到螺旋度是洛伦兹不变的，从而给定了 σ 的无质量粒子在所有惯性系下看起来是相同的，而我们认为螺旋度不同的无质量粒子是不同种类的粒子。但实际上，螺旋度相反的粒子通过空间反演对称性相关联，而电磁力与引力都是空间反演对称的，从而我们称螺旋度为 ± 1 的粒子为光子，而螺旋度为 ± 2 的粒子为引力子。在标准模型框架下， β 衰变中发出的螺旋度为 $1/2$ 的无质量粒子²³并不参与引力以外的具有空间反演对称性的相互作用，因此我们称 $1/2$ 螺旋度的为中微子，而 $-1/2$ 螺旋度的为反中微子，它们被认为是不同种粒子。我们一般认为光子和引力子为螺旋度为 $\pm 1, 2$ 的态的线性叠加。

虽然无质量粒子的螺旋度是洛伦兹不变的，但态本身却不是，根据2.133，我们可以看出对于混合了正负螺旋度的光子和引力子来说，它们在一个洛伦兹变换后会变为另一个不同的态，例如4-动量为 p 的一般光子态可以被写成

$$|\psi, p, \alpha\rangle = \alpha_+ |\psi, p, +1\rangle + \alpha_- |\psi, p, -1\rangle,$$

其中 $|\alpha_+|^2 + |\alpha_-|^2 = 1$ 。我们一般称这种现象为偏振。 $\alpha_- = \alpha_+^*$ 时，我们称其为线偏振，而 α_+ 的相位可以视为偏振面与某个垂直于 p 的固定参考系之间的夹角，洛伦兹变换 Λ 会让这个角度旋转 $\theta(\Lambda, p)$ 。注意，对于引力子的偏振，方程2.133中系数 $\exp(i\sigma\theta)$ 说明洛伦兹变换会将偏振面旋转角度 $2\theta(\Lambda, p)$ 。

例1 假设观测者 \mathcal{O} 观测到一个 W 玻色子（自旋为1且质量 $m \neq 0$ ），动量为 p ，动量方向为 y 方向且自旋 z 分量为 σ 。第二个观测者 \mathcal{O}' 相对第一个观测者以速度 v 沿着 z 方向运动，那么 \mathcal{O}' 会怎么描述 W 的这个态？

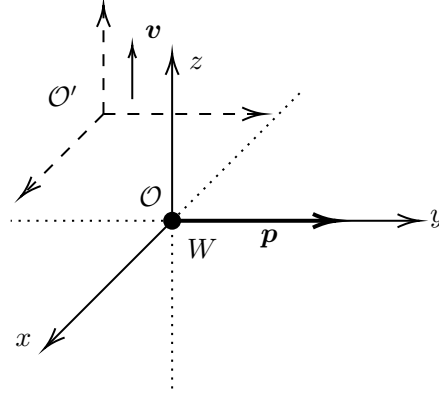
在 \mathcal{O} 系下，我们标记这个本征态为 $|\psi, p, \sigma\rangle$ ，其中 $p = (p^0, 0, \rho, 0)$, $p^0 = \sqrt{\rho^2 + m^2}$, $\rho \equiv |p|$ 。我们选择标准 $k^\mu = (m, 0, 0, 0)$ ，那么根据之前推出的变换规则(3)

$$U(\Lambda) |\psi, p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D_{\sigma'\sigma}^{(j)}(W(\Lambda, p)) |\psi, \Lambda p, \sigma'\rangle,$$

便可计算。首先我们需要知道小群群元(4)：

$$W(\Lambda, p) = L^{-1}(\Lambda p) \Lambda L(p),$$

²³ 对中微子质量的描述模型超出标准模型的范畴。

图 2.3: 以 O 为静止系观察动量为 y 方向的粒子。

这里我们的洛伦兹变换为

$$\Lambda(v)^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & \sqrt{\gamma^2 - 1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sqrt{\gamma^2 - 1} & 0 & 0 & \gamma \end{pmatrix}, \gamma = \frac{1}{\sqrt{1 - v^2}}.$$

而我们选取的将 $k^\mu{}_\nu$ 变回 $p^\mu{}_\nu$ 的洛伦兹变换为:

$$L(p)^\mu{}_\nu = R(\hat{\mathbf{p}}) \begin{pmatrix} \gamma' & 0 & 0 & \sqrt{\gamma'^2 - 1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sqrt{\gamma'^2 - 1} & 0 & 0 & \gamma' \end{pmatrix} R^{-1}(\hat{\mathbf{p}}), \gamma' \equiv \frac{\sqrt{\mathbf{p}^2 + m^2}}{m}.$$

注意这里, γ 给出 O 和 O' 之间的洛伦兹变换, 而 γ' 则是我们选择的“标准推动”的参数, 依赖于四矢量 p 。那么我们可以计算出 $W(\Lambda, p) = L^{-1}(\Lambda p)\Lambda L(p)$ 的具体值为

$$W(\Lambda, p)^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{pmatrix} = \mathcal{R}(\theta), \quad (2.134)$$

其中

$$\cos \theta = \frac{\gamma \left(\gamma m \sqrt{m^2 + \rho^2} + \rho^2 \right) - m \sqrt{m^2 + \rho^2}}{\gamma^2 \rho^2 + (\gamma^2 - 1) m^2}, \sin \theta = \frac{\rho \left(-\sqrt{\gamma^2 - 1} m + \gamma \sqrt{(\gamma^2 - 1)(m^2 + \rho^2)} \right)}{\gamma^2 \rho^2 + (\gamma^2 - 1) m^2}.$$

注意这里在计算时, $L^{-1}(\Lambda p)$ 中需要计算 $R(\Lambda \mathbf{p})$, 而非简单的 $R(\mathbf{p})$ 。我们可以发现, W 确实是小群 $SO(3)$ 的群元:

$$W^\mu{}_\nu k^\nu = k^\mu, W^\mu{}_\nu W^\nu{}_\rho = \delta^\mu{}_\rho.$$

那么由于 W 玻色子的自旋 $j = 1$, 我们需要求出 $D^1(W)$, 即Wigner旋转。考虑对 $SO(3)$ 的一个元素 R 做 $z - y - z$ 分解, 即将一个旋转写成

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma) = e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3},$$

其中 α, β, γ 为该旋转的欧拉角。那么表示 $D^{(1)}(R)$ 为:

$$D^{(1)}(\alpha, \beta, \gamma) = \begin{pmatrix} e^{i(\alpha+\gamma)} \cos^2 \frac{\beta}{2} & -\frac{e^{i\alpha} \sin \beta}{\sqrt{2}} & e^{i(\alpha-\gamma)} \sin^2 \frac{\beta}{2} \\ \frac{e^{i\gamma} \sin \beta}{\sqrt{2}} & \cos \beta & -\frac{e^{-i\gamma} \sin \beta}{\sqrt{2}} \\ e^{-i(\alpha-\gamma)} \sin^2 \frac{\beta}{2} & \frac{e^{-i\alpha} \sin \beta}{\sqrt{2}} & e^{-i(\alpha+\gamma)} \cos^2 \frac{\beta}{2} \end{pmatrix}.$$

对于 $W(\Lambda, p)$, 这是一个绕 x 轴的旋转, 那么我们只需先将绕 z 轴旋转使得 x 轴与原来的 y 轴重合, 再沿着 y 轴旋转原来的角度, 最后转回去即可, 那么我们得到

$$\mathcal{R}(\alpha, \beta, \gamma) = R_z\left(-\frac{\pi}{2}\right) R_y(\theta) R_z\left(\frac{\pi}{2}\right) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

于是 $D^{(1)}(W)$ 就为

$$D^{(1)}(W) = \begin{pmatrix} \cos^2 \frac{\theta}{2} & \frac{i \sin \theta}{\sqrt{2}} & -\sin^2 \frac{\theta}{2} \\ \frac{i \sin \theta}{\sqrt{2}} & \cos \theta & \frac{i \sin \theta}{\sqrt{2}} \\ -\sin^2 \frac{\theta}{2} & \frac{i \sin \theta}{\sqrt{2}} & \cos^2 \frac{\theta}{2} \end{pmatrix}.$$

那么, 作用了该洛伦兹变换后, 该单粒子态变为

$$U(\Lambda) |\psi, p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D_{\sigma'\sigma}^{(j)}(W(\Lambda, p)) |\psi, \Lambda p, \sigma'\rangle = \sqrt{\gamma} \sum_{\sigma' \in \{-1, 0, 1\}} D_{\sigma'\sigma}^{(1)}(W) |\psi, \Lambda p, \sigma'\rangle.$$

例如, 如果 $\sigma = -1$, 那么我们有

$$\begin{aligned} U(\Lambda) |\psi, p, -1\rangle &= \sqrt{\gamma} \left(D_{-1,-1}^{(1)}(W) |\psi, \Lambda p, -1\rangle + D_{0,-1}^{(1)}(W) |\psi, \Lambda p, 0\rangle + D_{1,-1}^{(1)}(W) |\psi, \Lambda p, 1\rangle \right) \\ &= \sqrt{\gamma} \left(\cos^2 \frac{\theta}{2} |\psi, \Lambda p, -1\rangle + \frac{i \sin \theta}{\sqrt{2}} |\psi, \Lambda p, 0\rangle - \sin^2 \frac{\theta}{2} |\psi, \Lambda p, 1\rangle \right). \end{aligned}$$

例2 假设观测者 \mathcal{O} 观测到一个光子, 动量为 \mathbf{p} , 动量方向为 y 方向且极化矢量在 z 方向。第二个观测者 \mathcal{O}' 相对第一个观测者以速度 \mathbf{v} 沿着 z 方向运动, 那么 \mathcal{O}' 会怎么描述这个光子?

首先我们需要将该光子的态写出来:

$$|\psi, p, \alpha\rangle = \alpha_+ |\psi, p, 1\rangle + \alpha_- |\psi, p, -1\rangle,$$

随后根据之前给出的式2.133

$$U(\Lambda) |\psi, p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \exp(i\sigma\theta(\Lambda, p)) |\psi, \Lambda p, \sigma\rangle$$

计算即可。与有质量态的情况相似, 区别在于我们选择标准4动量为 $k^\mu = (1, 0, 0, 1)$ 。我们首先计算小群元 $W(\Lambda, p)$ 。取 $L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|)$, 我们有

$$W(\Lambda, p) = \begin{pmatrix} 1 - \frac{1-\gamma^2}{2\rho^2\gamma^2} & -\frac{\sqrt{\gamma^2-1}}{\gamma\rho} & 0 & -\frac{\gamma^2-1}{2\gamma^2\rho^2} \\ -\frac{\sqrt{\gamma^2-1}}{\gamma\rho} & 1 & 0 & \frac{\sqrt{\gamma^2-1}}{\gamma\rho} \\ 0 & 0 & 1 & 0 \\ \frac{\gamma^2-1}{2\gamma^2\rho^2} & -\frac{\sqrt{\gamma^2-1}}{\gamma\rho} & 0 & \frac{1-\gamma^2}{2\rho^2\gamma^2} + 1 \end{pmatrix} = S(\alpha, \beta)R(\theta).$$

下面我们需要将 W 分解, 求出参数 θ 。由于

$$S(\alpha, \beta)R(\theta) = \begin{pmatrix} \zeta + 1 & \alpha \cos \theta - \beta \sin \theta & \alpha \sin \theta + \beta \cos \theta & -\zeta \\ \alpha & \cos \theta & \sin \theta & -\alpha \\ \beta & -\sin \theta & \cos \theta & -\beta \\ \zeta & \alpha \cos \theta - \beta \sin \theta & \alpha \sin \theta + \beta \cos \theta & 1 - \zeta \end{pmatrix},$$

对比系数可得

$$\zeta = \frac{\gamma^2 - 1}{2\gamma^2 \rho^2}, \alpha = -\frac{\sqrt{\gamma^2 - 1}}{\gamma \rho}, \beta = 0, \theta = 0.$$

从而作用了该洛伦兹变换后，该单粒子态变为

$$U(\Lambda) |\psi, p, \alpha\rangle = \sqrt{\gamma} (\alpha_+ |\psi, p, 1\rangle + \alpha_- |\psi, p, -1\rangle).$$

时间反演与空间反演

我们已经知道，任何齐次洛伦兹变换都可以由固有正时洛伦兹变换（ $\det \Lambda = 1, \Lambda^0_0 \geq 0$ ）以及时间反演 \mathcal{T} 和空间反演 \mathcal{P} 生成，这里 \mathcal{T} 和 \mathcal{P} 为

$$\mathcal{P}^\mu{}_\nu = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \mathcal{T}^\mu{}_\nu = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

曾经人们相信对于包含了 \mathcal{P}, \mathcal{T} 的 Λ ，庞加莱群的乘积规则

$$U(\Lambda_1, a_1)U(\Lambda_2, a_2) = U(\Lambda_1\Lambda_2, \Lambda_2 a_2 + a_1)$$

是不证自明的，同时人们也相信对应于 \mathcal{P}, \mathcal{T} 的算符 $\mathbf{P} \equiv U(\mathcal{P}, 0), \mathbf{T} \equiv U(\mathcal{T}, 0)$ ，以及任何固有正时洛伦兹变换 $\Lambda^\mu{}_\nu$ 和平移 a^μ ，都有

$$\begin{aligned} \mathbf{P}U(\Lambda, a)\mathbf{P}^{-1} &= U(\mathcal{P}\Lambda\mathcal{P}^{-1}, \mathcal{P}a), \\ \mathbf{T}U(\Lambda, a)\mathbf{T}^{-1} &= U(\mathcal{T}\Lambda\mathcal{T}^{-1}, \mathcal{T}a). \end{aligned} \quad (2.135)$$

这就是我们通常所说的“宇称守恒”的含义。但后来人们才意识到，这些仅在弱相互作用可以被忽略的近似下，对于 \mathbf{P}, \mathbf{T} 才是近似正确的。在后文中，我们先假设满足方程2.135的算符 \mathbf{P}, \mathbf{T} 真的存在。

下面考虑无穷小洛伦兹变换，即

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, a^\mu = \epsilon^\mu,$$

那么

$$\begin{aligned} \Lambda'^\mu{}_\nu &= \mathcal{P}^\mu{}_\rho (\delta^\rho{}_\sigma + \omega^\rho{}_\sigma) \mathcal{P}^\sigma{}_\nu = \delta^\mu{}_\nu + \mathcal{P}^\mu{}_\rho \omega^\rho{}_\sigma \mathcal{P}^\sigma{}_\nu, \\ a'^\mu &= \mathcal{P}^\mu{}_\nu \epsilon^\nu. \end{aligned}$$

带入其表示

$$U(1 + \omega, \epsilon) = 1 + \frac{i}{2} \omega_{\rho\sigma} J^{\rho\sigma} - i \epsilon_\rho P^\rho,$$

我们可以得到 \mathbf{P} 和 \mathbf{T} 的变换性质：

$$\begin{aligned} \mathbf{P}(1 + \frac{i}{2} \omega_{\rho\sigma} J^{\rho\sigma} - i \epsilon_\rho P^\rho) \mathbf{P}^{-1} &= 1 + \mathbf{P} \frac{i}{2} \omega_{\rho\sigma} J^{\rho\sigma} \mathbf{P}^{-1} - \mathbf{P} i \epsilon_\rho P^\rho \mathbf{P}^{-1} \\ &\stackrel{!}{=} 1 + \frac{i}{2} \mathcal{P}_{\rho\alpha} \omega^\alpha{}_\beta \mathcal{P}^\beta{}_\sigma J^{\rho\sigma} - i \mathcal{P}_{\rho\sigma} \epsilon^\sigma P^\rho \end{aligned}$$

令 $\omega_{\rho\sigma}, \epsilon_\rho$ 的系数相等，我们有

$$\mathbf{P} i J^{\rho\sigma} \mathbf{P}^{-1} = i \mathcal{P}_\mu{}^\rho \mathcal{P}_\nu{}^\sigma J^{\mu\nu}, \mathbf{P} i P^\rho \mathbf{P}^{-1} = i \mathcal{P}_\mu{}^\rho P^\mu.$$

对于 T ，我们可以同样计算得到

$$TiJ^{\rho\sigma}T^{-1} = i\mathcal{T}_\mu{}^\rho \mathcal{T}_\nu{}^\sigma J^\mu, TiP^\rho T^{-1} = i\mathcal{T}_\mu{}^\rho P^\mu.$$

注意，这里我们没有消掉两边的 i 是因为这时我们还没决定 P, T 是线性还是反线性（即是否与 i 反对易）的。但如果我们令 $\rho = 0$ ，便可以得到

$$PiHP^{-1} = iH,$$

其中 $H = P^0$ 为能量。如果 P 反线性且反么正，那么它与 i 反对易，从而有 $PHP^{-1} = -H$ ，但这意味着对于一个能量为 $E > 0$ 的态 $|\psi\rangle$ ，我们有

$$HP^{-1}|\psi\rangle = -P^{-1}H|\psi\rangle = -P^{-1}E|\psi\rangle = -EP^{-1}|\psi\rangle,$$

即存在一个能量为 $-E$ 的态 $P^{-1}|\psi\rangle$ ，但由于不可能存在比真空能量还低的态，因此 P 必须是线性且么正的，并且与 H 对易。

同理，对于 T ，我们同样有

$$TiHT^{-1} = -iH,$$

若 T 是么正且线性的，我们也可以得出同样的矛盾，因此 T 必须是反线性且反么正的。从而我们有

$$PHP^{-1} = THT^{-1} = H$$

在加上其他几个分量：

$$\begin{aligned} PJP^{-1} &= +J, & TJT^{-1} &= -J, \\ PKP^{-1} &= -K, & TKT^{-1} &= +K, \\ PPP^{-1} &= -P, & TPT^{-1} &= -P. \end{aligned}$$

记得其中

$$\mathbf{J} = \{J^{23}, J^{31}, J^{12}\}, \mathbf{K} = \{J^{01}, J^{02}, J^{03}\}, \mathbf{P} = \{P^3, P^2, P^1\}.$$

我们可以轻松证明上述关系和庞加莱群的对易关系是相容的，例如对于

$$[J_i, J_j] = i\epsilon_{ijk}J_k,$$

我们有

$$\begin{aligned} [J_i, J_j] &= [PJ_iP^{-1}, PJ_jP^{-1}] = PJ_iP^{-1}PJ_jP^{-1} - PJ_jP^{-1}PJ_iP^{-1} \\ &= P[J_i, J_j]P^{-1} \\ &= Pi\epsilon_{ijk}J_kP^{-1} \\ &= i\epsilon_{ijk}PJ_kP^{-1} = i\epsilon_{ijk}J_k. \end{aligned}$$

下面我们考虑 T, P 对单粒子态的作用。

正质量态，空间反演 我们将单粒子态 $|\psi, k, \sigma\rangle$ 定义为 \mathbf{P}, H, J_3 的本征矢，本征值分别为 $0, M, \sigma$ 。根据我们得到的对易关系：

$$PJ_3|\psi, k, \sigma\rangle = P\sigma|\psi, k, \sigma\rangle = \sigma P|\psi, k, \sigma\rangle = J_3P|\psi, k, \sigma\rangle,$$

因此我们知道 $P|\psi, k, \sigma\rangle$ 对 \mathbf{P}, H, J_3 的本征值同样为 $0, M, \sigma$ ，且与本来的态只差一个相位

$$P|\psi, k, \sigma\rangle = \eta_\sigma|\psi, k, \sigma\rangle.$$

实际上, η_σ 独立于 σ , 因为由

$$J_1 \pm iJ_2 |\psi, k, \sigma\rangle = \sqrt{(j \mp \sigma)(j \pm \sigma + 1)} |\psi, k, \sigma \pm 1\rangle,$$

两边作用 \mathbf{P} , 我们有

$$\eta_\sigma = \eta_{\sigma \pm 1}.$$

从而

$$\mathbf{P} |\psi, k, \sigma\rangle = \eta |\psi, k, \sigma\rangle,$$

我们称相位 η 为内禀宇称, 因为它仅依赖于 \mathbf{P} 所作用的粒子种类。对于一般的动量 p 态, 我们有

$$|\psi, p, \sigma\rangle = \sqrt{\frac{M}{p^0}} U(L(p)) |\psi, k, \sigma\rangle,$$

其中对于 $L(p)$, 我们选择式2.131所给出的标准推动。注意到(写成矩阵形式则是显然的):

$$\mathcal{P} L(p) \mathcal{P}^{-1} = L(\mathcal{P} p), \mathcal{P} p = (\sqrt{\mathbf{p}^2 + M^2}, -\mathbf{p}),$$

我们有

$$\mathbf{P} |\psi, p, \sigma\rangle = \sqrt{\frac{M}{p^0}} \mathbf{P} U(L(p)) |\psi, k, \sigma\rangle = \sqrt{\frac{M}{p^0}} \mathbf{P} U(L(p)) \mathbf{P}^{-1} \mathbf{P} |\psi, k, \sigma\rangle = \sqrt{\frac{M}{p^0}} U(L(\mathcal{P} p)) \eta |\psi, k, \sigma\rangle,$$

即

$$\mathbf{P} |\psi, p, \sigma\rangle = \eta |\psi, \mathcal{P} p, \sigma\rangle.$$

正质量态, 时间反演 与之前相同, 我们可以直接看出 \mathbf{T} 作用在 $|\psi, k, \sigma\rangle$ 上的效果是

$$\mathbf{P}(\mathbf{T} |\psi, k, \sigma\rangle) = 0, H(\mathbf{T} |\psi, k, \sigma\rangle) = M(\mathbf{T} |\psi, k, \sigma\rangle), J_3(\mathbf{T} |\psi, k, \sigma\rangle) = -\sigma(\mathbf{T} |\psi, k, \sigma\rangle).$$

从而我们有

$$\mathbf{T} |\psi, k, \sigma\rangle = \zeta_\sigma |\psi, k, -\sigma\rangle.$$

这里相因子 ζ_σ 与 σ 相关, 因为由于 \mathbf{T} 与 \mathbf{J}, i 反对易, 我们知道

$$(-J_1 \pm iJ_2) \zeta_\sigma |\psi, k, -\sigma\rangle = \sqrt{(j \mp \sigma)(j \pm \sigma + 1)} \zeta_{\sigma \pm 1} |\psi, k, -\sigma \mp 1\rangle,$$

从而有

$$-\zeta_\sigma = \zeta_{\sigma \pm 1}.$$

这里我们将相位重新写成

$$\begin{aligned} \zeta_\sigma &= \zeta(-)^{j-\sigma} \\ \Rightarrow \mathbf{T} |\psi, k, \sigma\rangle &= \zeta(-)^{j-\sigma} |\psi, k, -\sigma\rangle, \end{aligned}$$

这里 ζ 只依赖于粒子的种类。注意这里相位 ζ 并没有物理意义, 因为如果我们重新定义粒子态为

$$|\psi, k, \sigma\rangle \rightarrow |\psi', k, \sigma\rangle = \zeta^{1/2} |\psi, k, \sigma\rangle,$$

这样在变换中,

$$\mathbf{T} |\psi', k, \sigma\rangle = \zeta^{*1/2} \mathbf{T} |\psi, k, \sigma\rangle = \zeta^{*1/2} \zeta(-)^{j-\sigma} |\psi, k, -\sigma\rangle = (-)^{j-\sigma} |\psi', k, -\sigma\rangle,$$

即与 ζ 无关。

对于一般的动量态，推导过程与之前完全一样，注意到

$$\mathcal{T}L(p)\mathcal{T}^{-1} = L(\mathcal{P}p),$$

我们有：

$$\mathbb{T}|\psi, p, \sigma\rangle = \zeta(-)^{j-\sigma}|\psi, \mathcal{P}p, -\sigma\rangle.$$

零质量态，空间反演 下面我们来证明前文提到过的重要结论，即空间反演对称性要求任何种类的螺旋度非零的无质量粒子都必须伴随着另一个螺旋度相反的粒子。

我们继续取 $|\psi, k, \sigma\rangle$ 为 P^μ 和 J_3 的本征矢，其本征值为 $k^\mu = (\kappa, 0, 0, \kappa)$ ， σ ，那么当 \mathbf{P} 作用在这个态上时，它会产生4-动量为 $(\mathcal{P}k)^\mu = (\kappa, 0, 0, -\kappa)$ 而 J_3 本征值仍为 σ 的态，这意味着如果存在空间反演对称性，那么这意味着螺旋度为 σ 的态和 $-\sigma$ 的态实际上是一样的，这正是我们所要证明的。

下面我们考虑 \mathbf{P} 作用在一般的动量态 p 上。与有质量态不同，对于有质量态， \mathbf{P} 保持 $|\psi, k, \sigma\rangle$ 不变，而对于无质量态则不是，因此我们考虑构造一个保持标准动量不变的算符，即 $U(R_2^{-1})\mathbf{P}$ ，这里的 R_2 是一个能让 k 变为 $\mathcal{P}k$ 的旋转，我们将其取为绕 y 轴旋转 π ：

$$U(R_2) = \exp(i\pi J_2).$$

注意到 $U(R_2^{-1})$ 改变 J_3 的符号，我们可以写出：

$$U(R_2^{-1})\mathbf{P}|\psi, k, \sigma\rangle = \eta_\sigma|\psi, k, -\sigma\rangle.$$

为了让 k 变回到一般的 p ，我们取与之前相同的标准推动 $L(p) = R(\hat{\mathbf{p}})B(|\mathbf{p}|/\kappa)$ ，注意到 \mathcal{P} 与 $R(\hat{\mathbf{p}})$ 对易，同时 $R_2^{-1}\mathcal{P}$ 与 B 对易，我们有

$$\begin{aligned} \mathbf{P}|\psi, p, \sigma\rangle &= \sqrt{\frac{\kappa^0}{p^0}}U(R(\hat{\mathbf{p}})R_2B(|\mathbf{p}|/\kappa))U(R_2^{-1})\mathbf{P}|\psi, k, \sigma\rangle \\ &= \sqrt{\frac{\kappa^0}{p^0}}U(R(\hat{\mathbf{p}})R_2B(|\mathbf{p}|/\kappa))\eta_\sigma|\psi, k, -\sigma\rangle. \end{aligned}$$

这里注意， $R(\hat{\mathbf{p}})R_2$ 虽然使第三轴转向 $-\hat{\mathbf{p}}$ ，但是 $U(R(\hat{\mathbf{p}})R_2)$ 并不等于 $U(R(-\hat{\mathbf{p}}))$ ，而是多了一个沿着第三轴方向的旋转：

$$U(R(\hat{\mathbf{p}})R_2) = U(R(-\hat{\mathbf{p}}))\exp(\pm i\pi J_3).$$

因为我们考虑 $U(R(-\hat{\mathbf{p}}))$ ：

$$U(R(-\hat{\mathbf{p}})) = \exp(-i(\phi \pm \pi)J_3)\exp(-i(\pi - \theta)J_2),$$

这里选择绕第二轴旋转的角度为 $\pi - \theta$ ，是为了保证 $\pi - \theta$ 也在 $[0, \pi]$ 的范围内。而根据 ϕ 的范围，再让其保持在 $[0, 2\pi]$ 的范围内后选择 \pm 号。从而我们有

$$\begin{aligned} U^{-1}(R(-\hat{\mathbf{p}}))U(R(\hat{\mathbf{p}})R_2) &= e^{i(\pi-\theta)J_2}e^{i(\phi\pm\pi)J_3}e^{-i\phi J_3}e^{-i\theta J_2}e^{i\pi J_2} \\ &= e^{i(\pi-\theta)J_2}e^{\pm i\pi J_3}e^{i(\pi-\theta)J_2} = e^{\pm i\pi J_3}, \end{aligned}$$

最后一个等号是由于绕第三轴的 180° 旋转会改变 J_2 的符号。从而我们有

$$U(R(\hat{\mathbf{p}})R_2) = U(R(-\hat{\mathbf{p}}))\exp(\pm i\pi J_3).$$

但这意味着 $R(-\hat{\mathbf{p}})B(|\mathbf{p}|/\kappa)$ 实际上为 $\mathcal{P}p$ 方向上的标准增速 $L(\mathcal{P}p)$ ，从而对于一个一般的 p 态来说：

$$P|\psi, p, \sigma\rangle = \eta_\sigma \exp(\mp i\pi\sigma) |\psi, \mathcal{P}p, -\sigma\rangle.$$

其中正负号的选择取决于 \mathbf{p} 的第二个分量是正或负。

零质量态，时间反演 同样取 $|\psi, k, \sigma\rangle$ 为 P^μ 和 J_3 的本征矢，其本征值为 $k^\mu = (\kappa, 0, 0, \kappa), \sigma$ ，那么当 T 作用在这个态上时，它会产生4-动量为 $(\mathcal{P}k)^\mu = (\kappa, 0, 0, -\kappa)$ 而 J_3 本征值为 $-\sigma$ 的态，即 T 不改变螺旋度 $\mathbf{J} \cdot \hat{\mathbf{k}}$ 。

同理考虑作用在一般的态上，与之前相同，我们考虑 $U(R_2^{-1})T$ ，有

$$U(R_2^{-1})T|\psi, k, \sigma\rangle = \zeta_\sigma |\psi, k, \sigma\rangle.$$

同样的过程给出

$$T|\psi, p, \sigma\rangle = \sqrt{\frac{\kappa^0}{p^0}} U(R(\hat{\mathbf{p}})R_2B(|\mathbf{p}|/\kappa))\zeta_\sigma |\psi, k, \sigma\rangle = \zeta_\sigma \exp(\pm i\pi\sigma) |\psi, \mathcal{P}p, \sigma\rangle.$$

Kramers简并 所谓的Kramers简并是指对于满足条件 $T^2|\psi\rangle = -|\psi\rangle$ 的态， $T|\psi\rangle$ 必然与 $|\psi\rangle$ 能量相同且不是同一个态（或相差一个相位）。

我们首先考虑 T^2 作用在单粒子态上，首先考虑正质量态：

$$T^2|\psi, p, \sigma\rangle = T\zeta(-)^{j-\sigma} |\psi, \mathcal{P}p, -\sigma\rangle = (-)^{2j} |\psi, p, \sigma\rangle.$$

而对于零质量态：

$$T^2|\psi, p, \sigma\rangle = T\zeta_\sigma \exp(\pm i\pi\sigma) |\psi, \mathcal{P}p, \sigma\rangle = \zeta_\sigma^* \exp(\mp i\pi\sigma) \zeta_\sigma \exp(\mp i\pi\sigma) |\psi, p, \sigma\rangle = \exp(\mp 2i\pi\sigma) |\psi, p, \sigma\rangle.$$

由于 σ 为整数或者半整数，我们有

$$T^2|\psi, p, \sigma\rangle = (-)^{2|\sigma|} |\psi, p, \sigma\rangle.$$

如果态满足 $T^2|\psi\rangle = -|\psi\rangle$ ，即“自旋”为半整数，考虑 $T|\psi\rangle$ 。如果 $T|\psi\rangle$ 与 $|\psi\rangle$ 是同一个态，那么

$$T|\psi\rangle = \zeta |\psi\rangle,$$

但这样就有

$$T^2|\psi\rangle = |\zeta|^2 |\psi\rangle = |\psi\rangle,$$

这与假设矛盾。从而 $T|\psi\rangle$ 与 $|\psi\rangle$ 简并。

电荷

粒子-反粒子共轭

第三章 三种常见的场和它们的量子化

3.1 概述

3.1.1 正则量子化的大致手续

在相对论性量子场论中我们仍然要求自由粒子的拉氏量具有最高的对称性，也就是说，在庞加莱群作用下没有变化，且拉氏量只含有二次项（从而给出线性的运动方程）。¹ 空间平移不变性意味着拉氏量不能显含坐标；空间各向同性意味着拉氏量中的参数必须是标量，不能出现多分量的参数。

在已经写出拉氏量之后，我们可以计算三种场的哈密顿量，为正则量子化做好准备。由于我们通常在欧氏空间中写出哈密顿量而在闵可夫斯基时空中讨论拉氏量，需要格外注意一点：闵可夫斯基时空的度规为 $(+, -, -, -)$ 而欧几里得空间的度规为 $(+, +, +)$ ，因此

$$A_\mu A^\mu = (\dot{A}^0)^2 - \mathbf{A}^2,$$

上式左边为闵可夫斯基时空中的表达式，右边为欧几里得空间中的表达式。换言之，闵可夫斯基时空中的 $A_i A^i$ 和欧几里得空间中的 $A_i A^i$ 差一个负号。

在得到了场的哈密顿量之后，我们据此执行正则量子化。我们接下来需要对场算符施加正则对易或反对易关系。由于空间平移不变性，我们将在动表象下工作。此外，为了让公式看起来好看一些，我们忽略算符的帽子。在处理单粒子量子力学坐标表象下的问题时我们同时需要讨论算符 \hat{x} 和可以随意变动的空间坐标 \mathbf{x} ，但是在量子场论中几乎从来不需要讨论场变量的取值，因此忽略帽子并不会造成任何问题。在对自由场论做量子化时为了突出场的作用，我们采用海森堡绘景，即让场发生时间演化。

场算符做傅里叶变换之后的产生湮灭算符的归一化涉及一个微妙的、和相对论特性有关的地方。考虑到(2.116)，我们有

$$|p, \sigma\rangle = \sqrt{2\omega_p} |p, \sigma\rangle,$$

于是设 $a_{p,\sigma}$ 和 $a_{p,\sigma}^\dagger$ 为单粒子态 $|p, \sigma\rangle$ 对应的产生湮灭算符，则有

$$(\alpha_{p,\sigma})^\dagger = \sqrt{2\omega_p} a_{p,\sigma}^\dagger,$$

这样场算符的展开式就是

$$\varphi(x) \propto \sum_\sigma \int \frac{d^3 p}{\sqrt{(2\pi)^3 2\omega_p}} (a_{p,\sigma}^\dagger e^{ip_\mu x^\mu} + a_{p,\sigma} e^{-ip_\mu x^\mu}) e_\sigma,$$

¹虽然本文主要分析正则量子化，但写出运动方程还是用的是拉氏量。这是更加方便的做法，因为正则表述在理论框架上将时间和空间分开对待了，因此不容易观察哈密顿量在洛伦兹变化之下是不是给出恒定不变的动力学。

其中 e_σ 标记场 ϕ 属于的场表示的有限维部分的基底。或者，既然我们已经转而在三维空间中讨论问题，即已经不再要求洛伦兹协变性了，可以使用三维矢量更加清晰地写出

$$\varphi(\mathbf{x}, t) \propto \sum_{\sigma} \int \frac{d^3\mathbf{p}}{\sqrt{(2\pi)^3 2\omega_{\mathbf{p}}}} (a_{\mathbf{p},\sigma}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x} + i\omega_{\mathbf{p}}t} + a_{\mathbf{p},\sigma} e^{i\mathbf{p}\cdot\mathbf{x} - i\omega_{\mathbf{p}}t}) e_{\sigma}. \quad (3.1)$$

由于三种场都服从克莱因-高登方程，将(3.1)代入(3.10)会发现 $\omega_{\mathbf{p}}$ 和 \mathbf{p} 正好服从质壳关系(2.111)。另一方面，也可以使用相对论协变的积分测度，写出

$$\varphi(\mathbf{x}, t) \propto \sum_{\sigma} \int \frac{d^3\mathbf{p}}{2\omega_{\mathbf{p}}(2\pi)^{3/2}} (\alpha_{\mathbf{p},\sigma}^\dagger e^{ip_{\mu}x^{\mu}} + \alpha_{\mathbf{p},\sigma} e^{-ip_{\mu}x^{\mu}}) e_{\sigma}. \quad (3.2)$$

展开式(3.2)和(3.1)有各自的好处。(3.2)给出的产生湮灭算符以及它们产生的单粒子态是洛伦兹协变的，但是在处理对易关系的时候会略有复杂，因为此时产生湮灭算符的对易关系必定也是协变的，因此必须指定

$$[\alpha_{\mathbf{p}}, \alpha_{\mathbf{p}'}^\dagger] \sim \omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{p}')$$

这样的对易关系，或者类似的反对易关系；当然，因为此时使用的积分测度是 $\int d^3\mathbf{p} / (2\omega_{\mathbf{p}})$ ，这样的对易关系是正确的——无非是 δ 函数要修改为某种相对论形式而已。(3.1)给出的产生湮灭算符以及它们产生的单粒子态不是洛伦兹协变的，但是可以简化对易关系以及归一化时使用的积分测度。例如在对所有的动量模式求和时如果我们选取积分测度为 $\int d^3\mathbf{p}$ ，那么就需要指定

$$[a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger] = \delta^3(\mathbf{p} - \mathbf{p}'),$$

或者类似的反对易关系。而如果指定

$$[a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger] = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'),$$

此时只需要始终使用积分测度

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3}$$

对所有动量模式求和即可。两种展开式都是傅里叶变换，因此都能够消除哈密顿量中的导数。

为了避免引起混乱，这里我们统一指定下面使用的归一化常数和记号。本文采取Peskin和Quantum Field Theory for the Gifted Amateur的归一化常数和记号的一种混合。前面的 $|p\rangle$ 和 $|\mathbf{p}\rangle$ 的区别是Quantum Field Theory for the Gifted Amateur中使用的记号（Peskin中没有引入 $|p\rangle$ ，Peskin中的 $|\mathbf{p}\rangle$ 就是本文的 $|p\rangle$ ），而本文中的归一化常数则保持和Peskin一致。对傅里叶变换，规定对动量求和时积分测度带因子 $1/(2\pi)^3$ 而对坐标求和时积分测度就是 $d^3\mathbf{x}$ 。这样，产生湮灭算符的对易关系就是

$$[a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger] = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \quad (3.3)$$

或者将相对论协变的 α 算符转化为差一个常数的 a 算符，即令

$$\alpha_{\mathbf{p},\sigma}^\dagger = \sqrt{2\omega_{\mathbf{p}}} a_{\mathbf{p},\sigma}^\dagger, \quad (3.4)$$

就有

$$[\alpha_{\mathbf{p}}, \alpha_{\mathbf{p}'}^\dagger] = (2\pi)^3 2\omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{p}'). \quad (3.5)$$

反对易关系只需要将 $[\cdot, \cdot]$ 改成 $\{\cdot, \cdot\}$ 即可。场算符的展开和(2.117)遵从一样的习惯, 为

$$\varphi(\mathbf{x}, t) = \sum_{\sigma} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} (\alpha_{\mathbf{p}, \sigma}^{\dagger} e^{-i\mathbf{p} \cdot \mathbf{x} + i\omega_{\mathbf{p}} t} + \alpha_{\mathbf{p}, \sigma} e^{i\mathbf{p} \cdot \mathbf{x} - i\omega_{\mathbf{p}} t}) e_{\sigma}, \quad (3.6)$$

或者

$$\varphi(\mathbf{x}, t) = \sum_{\sigma} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}, \sigma}^{\dagger} e^{-i\mathbf{p} \cdot \mathbf{x} + i\omega_{\mathbf{p}} t} + a_{\mathbf{p}, \sigma} e^{i\mathbf{p} \cdot \mathbf{x} - i\omega_{\mathbf{p}} t}) e_{\sigma}. \quad (3.7)$$

e指数中的 $i\omega t - i\mathbf{p} \cdot \mathbf{x}$ 也可以写成相对论协变形式 $i\mathbf{p} \cdot x$ 。

对单粒子态, 有

$$|p, \sigma\rangle = \sqrt{2\omega_{\mathbf{p}}} a_{\mathbf{p}, \sigma}^{\dagger} |0\rangle,$$

从而有

$$\langle p, \alpha | q, \beta \rangle = (2\pi)^3 2\omega_{\mathbf{p}} \delta_{\alpha\beta} \delta^3(\mathbf{p} - \mathbf{q}). \quad (3.8)$$

在某一个给定的时间, 将 $\phi(\mathbf{x}, t)$ 作用在真空态上得到

$$\begin{aligned} \varphi(\mathbf{x}, t) |0\rangle &= \sum_{\sigma} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a_{\mathbf{p}, \sigma}^{\dagger} e^{-i\mathbf{p} \cdot \mathbf{x} + i\omega_{\mathbf{p}} t} e_{\sigma} |0\rangle \\ &= \sum_{\sigma} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{-i\mathbf{p} \cdot \mathbf{x} + i\omega_{\mathbf{p}} t} e_{\sigma} |p, \sigma\rangle. \end{aligned}$$

这里说明一下为什么我们采取了上面所说的傅里叶变换, 而没有, 比如说, 改变一下正负号选取。我们知道

$$\varphi_{\sigma}^{\dagger}(\mathbf{x}, t) |0\rangle \sim |\mathbf{x}(t), \sigma\rangle,$$

从而

$$\langle 0 | \varphi_{\sigma}(\mathbf{x}, t) | \psi \rangle \sim \text{single particle wavefunction}.$$

我们希望(2.109)成立。可以验证, 取上述傅里叶变换, 可以得到

$$\langle 0 | \varphi_{\sigma}(\mathbf{x}, t) | \psi \rangle \sim e^{-i\mathbf{p} \cdot \mathbf{x}} \propto e^{i\mathbf{p} \cdot \mathbf{x}},$$

这正是(2.109)。因此本节给出的傅里叶变换是正确的。表面上看这会导致一个疑难, 就是

$$\varphi_{\sigma}^{\dagger}(\mathbf{x}, t) |0\rangle \propto e^{i\omega t},$$

但是一个单粒子态的演化应该以 $e^{-i\omega t}$ 为时间演化因子, 但是其实这里没有任何问题: 此处的 φ 是海森堡绘景中的, $|\mathbf{x}\rangle$ 以 $e^{i\omega t}$ 时间演化, 而实际的单粒子态是没有时间演化的。那么, 现在我们做一个绘景变换, 在所有态上乘以一个因子 $e^{-i\omega t}$, 那么 $|\mathbf{x}\rangle$ 就没有时间演化了, 而实际的单粒子态以 $e^{-i\omega t}$ 时间演化, 这正是正确的薛定谔绘景。

场算符 (对应于湮灭算符的场算符) 有时候会被粗略地当成 “相对论量子力学中的波函数”。

判断应该使用对易关系来量子化场还是应该使用反对易关系来量子化场应当遵守几个条件:

- 非平凡性。哈密顿量不应该给出平凡的结果。
- 因果性。在某一个时空点施加相互作用只应该产生局域的影响。特别的, 在一个时空点做测量不应该对与之间隔 (指的是闵可夫斯基时空中的 “距离”) 为正的时空点产生影响。

- 能量正定性。哈密顿量应该可以写成产生湮灭算符的正定二次型，以避免能量无限下降。

以上给出的步骤完全描述了场算符的量子化过程。这种使用傅里叶变换得到对角化的哈密顿量的方式有时也称为正则量子化，因为它是算符的正则量子化（即施加对易或反对易关系）之后立刻可以完成的。

关于与场算符配套的真空态要说一句：在自由场下，无论采取哪种绘景，真空态 $|0\rangle$ 或者没有时间演化，或者时间演化只是乘上一个复数因子。这是因为真空态一定是哈密顿算符的本征态。此外，本节采用的量子化方案也体现出了一个重要的物理图像：具有确定能量 E 的粒子在经典极限下就对应着以圆频率 E 振荡的场。

最后，以上将场做傅里叶变换以消去运动方程中的导数的做法在经典情况下当然也适用。容易看出，动量为 \mathbf{p} 的粒子模式的经典极限就是波矢为 \mathbf{p} 的平面波；相应的，位置为 \mathbf{r} 的粒子模式的经典极限就是

$$\phi(\mathbf{r}') = \delta(\mathbf{r}' - \mathbf{r}).$$

以上两个模式的振幅均不确定；量子情况下振幅是分立的而经典情况下则不是。

即使在经典场中，也存在动量和位置不能同时确定的现象。场的量子化带来的不是动量和位置不能同时确定，而是场的振幅是离散化的——经典情况下，平面波的振幅可以任意变化，而量子情况下， ϕ （或者别的场算符）的本征值是离散的，此时才能够良定义“粒子”。

在本节剩下的部分中，我们将会看到，标量场和矢量场只能被量子化为玻色场，而旋量场只能被量子化为费米场。这是自旋统计定理的特例，这个定理说，对满足洛伦兹对称性的系统，半整数自旋对应着费米子，而整数自旋对应着玻色子。在没有洛伦兹对称性时这个结论不一定成立，实际上，对很多具有实际意义的系统——如格点系统——我们甚至没有旋转对称性，所以也无从讨论自旋。

3.1.2 经典近似

一个相对论性量子场论的退化：

- 量子性的单个粒子，即系统的基本自由度是坐标和一些额外的离散标签（自旋等），在这里，近似体现在缺乏粒子生灭，对那些本身缺乏粒子生灭的理论，从量子场论退化到单粒子理论没有任何近似。
 - 进一步，相对论性单粒子理论通常并没有什么意义，因此非相对论性近似往往也是必要的。
 - 退化得到的非相对论性单粒子理论也可以做二次量子化，得到一个非相对论性量子场论。这个理论当然可以通过原本的相对论性量子场论做非相对论近似一口气得到。
 - 还可以完全忽略量子涨落得到经典单粒子理论。
- 经典场论。如果原本的相对论性量子场论能够退化成单粒子理论，则它退化成的经典场论和单粒子理论的波动力学具有一样的形式。这就是“把波函数量子化”能够行得通的原因，也是麦克斯韦方程有时候看起来就好像一个波函数的原因——光子没有特别良定义的坐标表象下的单光子量子力学，但是如果硬是要定义一个，那么大体上就

是麦克斯韦方程的样子；经典的麦克斯韦方程中的平面波在量子化之后真的就是动量本征态，可以得到坐标-动量不确定性关系，等等。

– 可以做WKB近似，或者说程函近似，得到的 \mathbf{k} 和经典单粒子理论是一致的。

3.2 标量场

3.2.1 实标量场的克莱因-高登方程

拉氏量和运动方程

标量场的拉氏量中只能够出现 ϕ 和 $\partial_\mu\phi$ 构成的一次或二次不变量。 ϕ 构造出的一次不变量有 ϕ ，二次不变量有 ϕ^2 ， $\partial_\mu\phi$ 是矢量， $\partial_\mu\phi$ 不可能和 ϕ 缩并，而由于拉氏量中的参数都是标量它也不可能和参数缩并，因此它只能和自己缩并，得到 $\partial_\mu\phi\partial^\mu\phi$ 。这样我们得到

$$\mathcal{L} = A\phi + B\phi^2 + C\partial_\mu\phi\partial^\mu\phi.$$

拉氏量中的 ϕ 项实际上无关紧要，因为完全可以通过重新定义一个 $\phi' = \phi + \text{const}$ 来把一次项消除掉，于是我们略去这一项。最后，通过重新定义长度单位和 m ，可以得到

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2). \quad (3.9)$$

这个拉氏量导致下面的运动方程：

$$(\partial_\mu\partial^\mu + m^2)\phi = 0. \quad (3.10)$$

这就是克莱因-高登方程，标量场或者说自旋0场的基本运动方程。可以证明，为了让(3.10)给出有物理意义的预言（例如不出现无限下降的能量，等等），应当取 $m \geq 0$ 。

实际上，所有场的运动方程均满足克莱因-高登方程。我们将在推导其它场的运动方程之后证明这一点。这就导致了一个重要的结果。平移生成元在场表示中为(2.17)，从而

$$P_\mu P^\mu = -\partial_\mu\partial^\mu,$$

于是代入(3.10)，得到

$$P_\mu P^\mu\phi = -\partial_\mu\partial^\mu\phi = m^2\phi,$$

其中 m 为克莱因-高方程中出现的那个 m 。也即，通过 $P_\mu P^\mu$ 的表示的本征值（实际上就是它和恒等变换之间差的倍数，因为 $P_\mu P^\mu$ 是卡西米尔元）得到的 m 和场的运动方程得到的 m 是一样的。这个 m 实际上就是粒子的静质量。在第3.2.4节中我们会看到它就对应着薛定谔方程中的质量。

现在导出标量场的哈密顿表述。计算得到

$$\pi = \partial_0\phi = \dot{\phi}, \quad (3.11)$$

相应的

$$\mathcal{H} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2. \quad (3.12)$$

如我们希望的那样，哈氏量密度是正定的。这当然是因为我们适当地选择了 \mathcal{L} 的正负号。

另外，注意到

$$P_0 = E, \quad P_i e^i = \mathbf{p},$$

在场表示中我们可以写出

$$E^2 - \mathbf{p}^2 = m^2$$

或者说

$$E^2 = m^2 + \mathbf{p}^2.$$

这正是质壳关系(2.111)。这就提示我们，还有另一种量子化方式：做替换

$$E \longrightarrow \hat{E} = i\partial_0, \quad \mathbf{p} \longrightarrow \hat{\mathbf{p}} = -i\nabla,$$

则从能量-动量关系就可以得到克莱因-高登方程。我们不采用这种方案，因为它隐含地引入了太多的假设：算符 E, \mathbf{p} 是作用在一个算符场上而不是态矢量上； E, \mathbf{p} 是厄米算符，也即，平移群在此算符场上取么正表示（注意这一点并不一般成立！例如，场表示中的有限维表示就常常不是么正的），等等。为了和经典场论中的记号保持一致，我们后面将用 ω 代替 E 。 ω 或者说场的傅里叶分量的频率在经典场论下有明确意义，而在量子场论的语境下，它实际上对应了这个场的傅里叶分量（粒子产生算符或者湮灭算符）产生/湮灭的粒子的能量 E 。

守恒量

在第2.4.1节中我们泛泛地谈论过对称性带来的守恒量，现在我们对自由标量场重复这些计算。

首先考虑平移——无论是时间还是空间——不变性带来的守恒量。一个一般的无穷小四维平移变换形如

$$x^\mu \longrightarrow x^\mu + \epsilon^\mu, \quad (3.13)$$

3.2.2 实标量场的正则量子化

平面波模式

无需额外考虑标量场的基，于是对实标量场，可以直接取

$$\phi(\mathbf{x}, t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x} + i\omega_{\mathbf{p}}t} + a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x} - i\omega_{\mathbf{p}}t}), \quad (3.14)$$

显然它是(3.10)的一个解。相应的使用(3.11)，有

$$\pi(\mathbf{x}, t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} i\sqrt{\frac{\omega_{\mathbf{p}}}{2}} (a_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x} + i\omega_{\mathbf{p}}t} - a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x} - i\omega_{\mathbf{p}}t}) \quad (3.15)$$

共轭动量不是洛伦兹协变的。这并不让人意外，因为其定义和时间维的选取有关。计算得到

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2} \omega_{\mathbf{p}} (a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + a_{\mathbf{p}} a_{\mathbf{p}}^\dagger).$$

下面把正则对易关系施加到标量场 ϕ 上。我们知道，大体上，对 ϕ, π 施加正则对易关系基本上等价于要求 a, a^\dagger 满足产生湮灭算符的对易关系，然后。不过为了保证没有多余的归一化常数我们还是需要做一些具体计算。实际上，更加方便的做法是首先假定 a 和 a^\dagger 满足正则对易关系

$$[a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger] = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \quad [a_{\mathbf{p}}, a_{\mathbf{p}'}] = 0, \quad (3.16)$$

然后计算得到 ϕ 和 π 的对易关系。能否将标量场量子化为费米子呢？反对易关系等价于

$$\{a_{\mathbf{p}}, a_{\mathbf{p}'}^\dagger\} = (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \quad \{a_{\mathbf{p}}, a_{\mathbf{p}'}\} = 0.$$

将反对易关系代入哈密顿量表达式会导致哈密顿量变成常数，因此这是平凡解，舍去。将对易关系代入哈密顿量的表达式，得到

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} \left(a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2} [a_{\mathbf{p}}, a_{\mathbf{p}}^\dagger] \right). \quad (3.17)$$

容易看出第二项实际上是发散的。产生这种发散的原因在于，相对论性量子场论不会被用于处理动量特别高的问题（在那里需要新的物理，通常称为“紫外端的物理”），因此所谓的对整个动量空间的积分实际上只是对动量空间中一块很大的区域的积分。在这种意义下，(3.17)中的第二项是一个很大的常数，称为真空零点能。因此在讨论全空间内的问题时，可以丢弃它得到等效的哈密顿量（注意此时哈密顿量的正定性实际上被破坏了）

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}}. \quad (3.18)$$

这是一个福克空间上的1粒子算符。它表明自由场情况下单粒子携带能量为 $\omega_{\mathbf{p}}$ 。通过反复使用对易关系(3.16)以及真空态被湮灭算符作用后得到0这一事实，可以计算出

$$H a_{\mathbf{p}}^\dagger |0\rangle = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger |0\rangle. \quad (3.19)$$

因此正如我们预期的那样，单粒子态 $|\mathbf{p}\rangle$ 是哈密顿量的本征态。

真空零点能的出现实际上意味着原来的哈密顿量中的各个项是不对易的，因此真空态的能量不能是零，如果它是零，那么由哈密顿量的正定性，哈密顿量中的每一项作用在真空态上都会得到零，于是真空态是哈密顿量的每一项的本征态，这就产生了矛盾。不对易性是纯粹的量子概念，因此真空零点能只有在量子场论中才能够得到良好的定义。如果哈密顿量中所有的项都是彼此对易的，就不会有真空零点能。有时真空零点能的存在也称为量子涨落，因为即使在真空态，也不是所有的物理量都有完全确定的值。

需要注意的是如果我们讨论的问题不是定义在全空间上的，可能不能直接把真空零点能丢弃。例如，设有两块无穷大的金属板，它们施加的边界条件会让(3.14)中的一些模式为零，通过计算可以发现板间的真空零点能小于板外，从而产生一个板之间的吸引力。

总之，标量场需要使用正则对易关系来量子化，不能用反对易关系。因此标量场描述0自旋玻色子。

守恒量

场的动量为

$$P_i = \int d^3\mathbf{x} \pi \partial_i \phi,$$

从而

$$\mathbf{P} = - \int d^3\mathbf{x} \pi \nabla \phi = \int d^3\mathbf{x} \mathbf{p} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \text{vacuum zero-point item}, \quad (3.20)$$

正如我们预期的，场的动量也是单粒子算符，并且正好就是所有粒子的动量之和。

标量场没有内禀自由度，因此也不携带自旋角动量。我们无需讨论其自旋角动量。

实标量场的传播子

通过(3.16)中的归一化常数我们可以写出动量空间中的同时间两点关联函数:

$$\langle 0|a_{\mathbf{p}}a_{\mathbf{q}}^{\dagger}|0\rangle = (2\pi)^3\delta(\mathbf{p}-\mathbf{q}), \quad \langle 0|a_{\mathbf{p}}^{\dagger}a_{\mathbf{q}}|0\rangle = 0. \quad (3.21)$$

代入展开式(3.14), 就得到实空间下的两点关联函数 (这两个点可以不等时):

$$D(x-y) = \langle 0|\phi(x)\phi(y)|0\rangle = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{-ip\cdot(x-y)}|_{p^0=E_{\mathbf{p}}}. \quad (3.22)$$

被积分的e指数是洛伦兹不变的, 积分测度也是洛伦兹不变的, 等式左边也是洛伦兹不变的——本该如此, 既然标量场是洛伦兹不变的。这个式子可以写成更加明显的洛伦兹不变的形式。容易看出

$$\int d\omega \frac{1}{(\omega - \omega_{\mathbf{p}})(\omega + \omega_{\mathbf{p}} - i0^+)} e^{-ip\cdot(x-y)} = -\pi i \frac{1}{2\omega_{\mathbf{p}}} e^{-ip\cdot(x-y)}|_{p^0=E_{\mathbf{p}}},$$

这里积分号表示计算积分主值, 负号是由于 $e^{-i\omega t}$ 在下半平面衰减而不是上半平面。于是就有

$$\begin{aligned} D(x-y) &= 2 \int \frac{d^4p}{(2\pi)^4} \frac{i}{(\omega - \omega_{\mathbf{p}})(\omega + \omega_{\mathbf{p}} - i0^+)} e^{-ip\cdot(x-y)} \\ &= 2 \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 - i0^+(\omega - \omega_{\mathbf{p}})} e^{-ip\cdot(x-y)}. \end{aligned} \quad (3.23)$$

这个积分看起来非常不自然: 我们需要忽略其中一个极点, 并且取另一个极点的积分主值, 然后再将结果乘以2。这种不自然性暗示 $\langle 0|\phi(x)\phi(y)|0\rangle$ 可能并不是非常好用的表示两点关联的方式——它的确不是, 因为当 $x^0 < y^0$ 时它还是会给出一个值, 并没有很好地反映因果性。对这个量的有意义的使用局限于 $x^0 > y^0$ 时。

在等时情况下上式实际上可以进一步往下计算。设 $x^0 - y^0 = 0$, $\mathbf{x} - \mathbf{y} = \mathbf{r}$, 采用球坐标系, 有

$$\begin{aligned} D(x-y) &= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{i\mathbf{p}\cdot\mathbf{r}} \\ &= \frac{2\pi}{(2\pi)^3} \int \frac{p^2 dp}{2\omega_{\mathbf{p}}} \frac{e^{ipr} - e^{-ipr}}{ipr} \\ &= -\frac{i}{2(2\pi)^2 r} \int_{-\infty}^{\infty} \frac{p dp}{\sqrt{p^2 + m^2}} e^{ipr}, \end{aligned}$$

被积函数是一个多值函数, 从 im 到 $i\infty$, $-im$ 到 $-i\infty$ 作割线, 即确定一个单值分支。将积分路径换成围绕从 im 到 $i\infty$ 的割线, 并作变量代换 $p = i\rho$, 可将上式化为

$$D(x-y) = \frac{1}{4\pi^2 r} \int_m^{\infty} d\rho \frac{\rho e^{-\rho r}}{\sqrt{\rho^2 - m^2}} \stackrel{r \rightarrow \infty}{\sim} e^{-mr}.$$

这就是说, 即使是彼此之间有类空间隔的两点之间实际上也是有关联的, 虽然随着举例增加关联函数快速下降; 或者说, 在光锥之外, 跃迁振幅会指数下降但是不会一开始就降到零。

这实际上并没有直接破坏因果律, 因为量子态并不是直接可观测的, 真正有意义的是测量, 如果具有类空间隔的两点之间的同种可观测测量是对易的, 因果律还是能够保持。计算 $\phi(x)$ 和 $\phi(y)$ 的对易子:

$$\begin{aligned} [\phi(x), \phi(y)] &= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} (e^{ip\cdot(y-x)} - e^{ip\cdot(x-y)}) \\ &= D(x-y) - D(y-x), \end{aligned} \quad (3.24)$$

在 x 和 y 之间为类空间隔时，洛伦兹变换可以让 $x-y$ 变成 $y-x$ ，由于 $D(x-y)$ 和 $D(y-x)$ 各自是洛伦兹标量， $[\phi(x), \phi(y)]$ 就是零。否则，洛伦兹变换不能让 $x-y$ 变成 $y-x$ ，因为此时时间的先后顺序是绝对的，从而 $\phi(x)$ 和 $\phi(y)$ 不对易。因此因果律在这里确实是能够保持的。

在实际计算中通常计算编时格林函数，因为其携带了足够多的信息，并且在自由场情况下有Wick定理。仿照前述引入 ω 积分的方法，可以验证我们有

$$D_F(x-y) = \langle 0|T\phi(x)\phi(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i0^+} e^{-ip \cdot (x-y)}. \quad (3.25)$$

这称为费曼传播子，可以看到它在动量空间中的形式是非常简单的。

3.2.3 实标量场的路径积分量子化

旋量场的路径积分需要额外提几句，因为此时所谓的“经典场”实际上是一些彼此反对易的算符（所谓格拉斯曼数），而不是正常的数。格拉斯曼数的积分基本上只用在写出配分函数上，

$$S = \int d^4x \bar{\psi}(iG^{-1})\psi$$

3.2.4 非相对论极限

克莱因-高登方程的退化形式

我们将讨论克莱因-高登方程的退化形式。旋量场和标量场由于也服从克莱因-高登方程，没有必要单独考虑——它们多出来的自由度可以使用其它方式，如自旋等，引入。实际上我们讨论的应该是复的克莱因-高登方程，因为旋量场是复的，但本节的讨论并不会用到场是不是复的这个信息。

首先我们注意到一个事实：时谐波

$$\phi = e^{-imt} \quad (3.26)$$

是(3.10)的解，并且它的能量最低，就是零。（代入哈氏量可得）因此，能量不高的场只是微微偏离(3.26)，我们设其为

$$\phi(\mathbf{x}, t) = \psi(\mathbf{x}, t)e^{-imt}, \quad (3.27)$$

则

$$(\partial_\mu \partial^\mu + m^2)\phi = e^{-imt}(-2im\partial_t\psi + \partial_t^2\psi - \nabla^2\psi).$$

由于 ϕ 只是略微偏离(3.26)²，可以预期 ψ 的时间部分振荡不会特别明显，于是取近似

$$\partial_t^2\psi \ll m\partial_t\psi,$$

就得到

$$i\partial_t\psi + \frac{1}{2m}\nabla^2\psi = 0. \quad (3.28)$$

(3.28)称为薛定谔场的运动方程。容易看出它不是洛伦兹协变的，这是理所当然的，因为它描述的现象发生在低能近似下，此时伽利略对称性就足够了。薛定谔场是复的，无论 ϕ 是不是复场，因为拟设(3.27)引入了一个复数因子。

²需要注意的是这个说法字面上实际上是不严谨的。 ϕ 是一个算符，它包含了所有可能的 ϕ 的取值，不应该“只是略微偏离(3.26)”。然而， ϕ 的本征态中非常偏离(3.26)的那部分模式在我们的低能有效理论中并不会被涉及到。换言之，我们关心的那部分 $|\phi\rangle$ 只是略微偏离(3.26)，因此认为 ϕ 只是略微偏离(3.26)并不会显著改变我们的理论的行为。

方程(3.28)是以下拉氏量³

$$\mathcal{L} = \frac{i}{2} (\psi^\dagger \dot{\psi} - \dot{\psi} \psi^\dagger) - \frac{1}{2m} \nabla \psi^\dagger \cdot \nabla \psi \quad (3.29)$$

的运动方程。把 ψ 和 ψ^\dagger 看成两个独立的场，分别应用欧拉-拉格朗日方程，就能够得到(3.28)和其共轭转置。

容易看出，

$$\pi(\psi) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = \frac{i}{2} \psi^\dagger, \quad \pi(\psi^\dagger) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^\dagger} = -\frac{i}{2} \psi^\dagger,$$

从而可以计算出

$$\mathcal{H} = \frac{1}{2m} \nabla \psi^\dagger \cdot \nabla \psi. \quad (3.30)$$

这个哈氏量中出现了 π 的导数，处理起来会比较麻烦。为了规避这些麻烦，我们将不再讨论经典的哈密顿动力学，而直接开始做量子化。

电荷密度为

$$\rho(\mathbf{r}) = q\psi^\dagger(\mathbf{r})\psi(\mathbf{r}), \quad (3.31)$$

而且电流密度为

$$\mathbf{j}(\mathbf{r}) = \frac{1}{2mi} (\psi^\dagger(\mathbf{r}) \nabla \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla \psi^\dagger(\mathbf{r})) \quad (3.32)$$

薛定谔场

实际上也可以通过量子化薛定谔场来得到非相对论性量子场论。薛定谔场实际上是标量场、旋量场、矢量场退化而来的场，因此它也有内禀自由度。使用自旋（或者螺旋度）标记这些内禀自由度。由于薛定谔场不是实场，考虑对易关系

$$[\psi^i(\mathbf{x}, t), \pi_j(\mathbf{y}, t)] = i\delta_j^i \delta^3(\mathbf{x} - \mathbf{y}),$$

即

$$[\psi^i(\mathbf{x}, t), (\psi^j)^\dagger(\mathbf{y}, t)] = 2\delta^3(\mathbf{x} - \mathbf{y}),$$

这表明

总之，在不涉及粒子相互作用时，单粒子量子力学足以覆盖薛定谔场的情况，即“非相对论量子场论”就是量子力学。⁴我们再一次看到场自由度和数量可变的粒子自由度实际上就是一回事。初等量子力学中可以直接定义S算符、单粒子费曼图（“原子吸收一个光子、释放一个光子”，等等），你可能会问为什么这些本来用于场论的概念也可以被用在单粒子量子力学上，毕竟前者是3+1维理论而后者是0+1维理论。但实际上，这些用在量子场论上的概念完全可以被应用在薛定谔场上，而由于薛定谔场不涉及粒子数变化，这些概念就可以被套用在单粒子态量子力学上。

³ $\nabla \psi^\dagger \cdot \nabla \psi$ 代表将两个梯度算符做缩并，行向量 ψ^\dagger 和列向量 ψ 相乘，即 $\partial_i \psi^\dagger \partial^i \psi$ 。混合使用不变量记号和矩阵记号是因为我们并不知道 ψ 的内部结构，只知道 $\psi^\dagger \psi$ 是标量，因此把 ψ 当成一个整体，好像一个标量一样，来做计算。

⁴ 量子场论和量子力学的对应实际上有两方面：首先，量子场论和量子力学都可以写成哈密顿动力学的形式，当然前者各个物理量可以使用空间位置作为标签而后者物理量的标签都是离散的；其次，量子场论和量子力学都能够描述多粒子态。当我们说非相对论性量子场论就是量子力学时我们是在说后者，当我们说量子场论是3+1维量子力学时我们是在说前者。

关于归一化的注记

在将相对论性的相互作用顶角移植到非相对论极限下时需要注意归一化问题。在非相对论性极限下我们更喜欢使用 $\{|p\rangle\}$ 表象，而不是 $\{|p\rangle\}$ 表象。我们知道

$$|p\rangle = \sqrt{2\omega_p} |p\rangle,$$

在非相对论极限下就有

$$|p\rangle = \frac{1}{\sqrt{2m}} |p\rangle, \quad (3.33)$$

于是应有

$$\langle p_1, \dots, p_m | S | q_1, \dots, q_n \rangle = \frac{1}{(2m)^{(m+n)/2}} \langle p_1, \dots, p_m | S | q_1, \dots, q_n \rangle. \quad (3.34)$$

3.3 旋量场

3.3.1 旋量场的狄拉克方程

旋量场实际上几乎从来不会在经典情况下遇到，因为它们的场值是复数，因此不具有直接的物理意义。

本节讨论旋量的运动方程。使用凑拉氏量的方法处理旋量会比较困难，因为旋量的指标分带点的和不带点的，因此会频繁地涉及求共轭等运算，在拉格朗日动力学中讨论这些问题并不方便。因此接下来我们尝试直接构造旋量的运动方程，从这些运动方程反推对应的拉氏量。我们将尝试构造一阶运动方程。如果对魏尔旋量和狄拉克旋量都能够构造出一阶运动方程，那就没有必要考虑更高阶的运动方程。

魏尔旋量的运动方程

首先讨论魏尔旋量的运动方程。满足平移不变性的方程形如

$$\partial_0 \psi = b^i \partial_i \psi + C \psi,$$

其中 b^i 和 C 是常数， C 可以是一个旋量矩阵。由于我们同时还要求旋转不变性， C 只能是一个标量。显然，这个方程中所有含有导数的项加在一起必然得到一个旋量，即

$$\partial_0 \psi - b^i \partial_i \psi = \text{a covariant term} = C \psi.$$

梯度算符是矢量，按照(2.68)，我们可以写出作用在魏尔旋量上的导数算符

$$\partial_{ab} = \partial_\nu \sigma_{ab}^\nu. \quad (3.35)$$

从而魏尔旋量的梯度就是

$$\partial_{ab} \psi^b = \partial_\nu \sigma_{ab}^\nu \psi^b$$

和

$$\partial^{\dot{a}b} \psi_b = \partial_\mu (\sigma^\mu)^{\dot{a}b} \psi_b.$$

两个表达式中， σ 的指标都一个带点一个不带点，这是为了保证梯度算符的协变性，因为矢量是一个左手旋量和一个右手旋量直积的结果。 $\partial_0 \psi - b^i \partial_i \psi$ 应该能够写成以上两种旋量梯度的函数。由于以上两种旋量梯度带一个指标，而 $\partial_0 \psi - b^i \partial_i \psi$ 也是单指标对象，显然两者

只应该差一个倍数。(当然,也可以将一个旋量张量参数和旋量梯度做缩并,但这样就没有旋转不变性了)这个倍数可以被吸收到 C 中。从而运动方程形如

$$\text{grad}\psi = C\psi.$$

然而,注意到左手旋量的梯度是一个右手旋量,右手旋量的梯度是一个左手旋量,因此以上的方程会让一个右手旋量的各个分量等于一个左手旋量的各个分量,从而破坏了洛伦兹协变性。消除这个矛盾的唯一可能就是让 $C = 0$,于是左手旋量的运动方程为

$$\partial_\mu(\sigma^\mu)^{\dot{a}b}\psi_b = 0,$$

右手旋量的运动方程为

$$\partial_\nu\sigma^\nu_{\dot{a}b}\psi^{\dot{b}} = 0.$$

由定义, $(\sigma^\mu)_{ab}$ 的分量矩阵就是 σ 矩阵,而 $(\sigma^\mu)^{\dot{a}b}$ 的分量矩阵则需要通过指标升降关系

$$(\sigma^\mu)^{\dot{a}b} = (\epsilon^{ac}(\sigma^\mu)_{cd}\epsilon^{\dot{b}d})^*$$

得到。定义 $(\sigma^\mu)^{\dot{a}b}$ 的分量矩阵为 $\bar{\sigma}^\mu$,通过计算可以发现

$$\bar{\sigma}^0 = \sigma^0 = I, \quad \bar{\sigma}^i = -\sigma^i, \quad (3.36)$$

于是可以使用矩阵形式写出运动方程:

$$\partial_0\psi \pm \sigma^i\partial_i\psi = 0. \quad (3.37)$$

负号为左手旋量,正号为右手旋量。从(3.37)立刻可以得到

$$\partial_\mu\partial^\mu\psi = 0,$$

无论是右手旋量还是左手旋量。因此,单独一个旋量场一定是没有质量的。

然而,拉氏量中的质量项实际上是自由理论的一部分,因此一种非常简单的让旋量场获得质量而与此同时理论保持为自由理论的方法是让一对旋量耦合起来。两个旋量场的耦合意味着拉氏量中要出现一个二次项,它由两个旋量的乘积构成,而由于这个二次项必须是标量,它应该包含一个左手旋量和一个右手旋量。考虑到我们可以把一个左手旋量和一个右手旋量打包成一个狄拉克旋量,有必要分析狄拉克旋量的运动方程。

狄拉克旋量的运动方程

魏尔表象 我们现在让狄拉克旋量中的两个旋量之间有线性的相互作用(从而,关于狄拉克旋量的方程仍然是线性的)。我们将狄拉克旋量写成一个左手旋量和一个右手旋量拼合的形式,称为**魏尔表象**或**手性基**,虽然实际上有其它方式可以表示狄拉克旋量。洛伦兹不变性的要求意味着,唯一可能的方程形式如下:

$$\begin{aligned} (\partial_0 + \sigma^i\partial_i)\psi_R &= -im_1\psi_L, \\ (\partial_0 - \sigma^i\partial_i)\psi_L &= -im_2\psi_R. \end{aligned} \quad (3.38)$$

其中我们为了节省符号,使用 ψ_L 和 ψ_R 分别代表狄拉克旋量 ψ 的左手部分和右手部分。当然, $m = 0$ 时就狄拉克旋量的运动方程就退化为一对完全无关的左手旋量和右手旋量。这也

就是实际计算时没有必要单独讨论魏尔旋量的原因。在(3.38)中各个方程的两边作用适当的算符来让两个场解耦，就得到

$$\partial_\mu \partial^\mu \psi_L = -m_1 m_2 \psi_L, \quad \partial_\mu \partial^\mu \psi_R = -m_1 m_2 \psi_R,$$

也就是说狄拉克旋量也满足克莱因-高登方程，只需要我们令 $m^2 = m_1 m_2$ 。从而为了得到物理解，我们要求 $m \geq 0$ 。可以看到 m_1 和 m_2 的具体取值实际上是无关紧要的，重要的是它们的乘积，因此以下让它们都取 m 。

为了将(3.38)写成更加紧凑的形式，引入 γ 矩阵⁵

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}, \quad (3.39)$$

从而

$$\gamma_\mu = \eta_{\mu\nu} \gamma^\nu = \begin{pmatrix} 0 & \bar{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix}, \quad (3.40)$$

则得到

$$(i\gamma^\mu \partial_\mu - m)\psi = 0. \quad (3.41)$$

这就是**狄拉克方程**。如前所述，它能够推导出克莱因-高登方程，并且在 m 取零时退化为一个左手旋量场和一个右手旋量场的简单组合。

现在我们尝试拼凑一个拉氏量出来。由于狄拉克场的运动方程是一阶的而它又是一个复场，需要通过 ψ 的复共轭拼凑出一个在运动方程意义下“独立”的场，然后构造一个同时包含 ψ 及其复共轭的拉氏量，由这个拉氏量给出关于 ψ 和它的复共轭的两个方程，并且这两个方程必须等价。现在我们尝试寻找和(3.41)等价，但是仅仅包含其复共轭的方程。由(3.41)取共轭转置⁶，得到

$$(-i)\partial_\mu \psi^\dagger (\gamma^\mu)^\dagger - m\gamma^\dagger = 0.$$

容易验证 γ 矩阵具有下面的性质：

$$(\gamma^0)^\dagger = \gamma^0, \quad (\gamma^i)^\dagger = -\gamma^i,$$

以及

$$\gamma^i \gamma^0 = -\gamma^0 \gamma^i,$$

我们发现

$$i\partial_0 \psi^\dagger \gamma^0 \gamma^0 + i\partial_i \psi^\dagger \gamma^0 \gamma^i + m\psi^\dagger \gamma^0 = 0.$$

定义

$$\bar{\psi} = \psi^\dagger \gamma^0, \quad (3.42)$$

则其运动方程为

$$i\partial_\mu \bar{\psi} \gamma^\mu + m\bar{\psi} = 0. \quad (3.43)$$

这正是我们需要的另一个运动方程。我们会发现，拉氏量

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (3.44)$$

⁵这里给出的 γ 矩阵的形式实际上只是一种可能性。我们称这种将狄拉克旋量的左手部分和右手部分分开处理（或者等价地说，狄拉克旋量的基或者只含有左手旋量，或者只含有右手旋量），并且按照(3.39)引入 γ 矩阵的方式为**魏尔表象**。也可以取其它的旋量基，从而获得其它表象。

⁶这里的共轭转置是指场的共轭转置，不需要对作用在场上的算符 ∂_μ 取共轭转置。

分别对 ψ 和 $\bar{\psi}$ 应用欧拉-拉格朗日方程, 就得到(3.41)和(3.43)。同时容易验证这是一个洛伦兹标量。这表明(3.44)确实就是狄拉克场的拉氏量。我们之后还会非常频繁地写 $\gamma^\mu a_\mu$ 这样的量, 可以将其简记为 \not{a} 。

从(3.44)可以推导出对应的哈氏量。计算共轭动量可以得到

$$\pi = i\bar{\psi}\gamma^0 = i\psi^\dagger, \quad (3.45)$$

从而能够得到哈氏量密度

$$\mathcal{H} = -i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi = -\pi\gamma^0\gamma^i\partial_i\psi - im\pi\gamma^0\psi. \quad (3.46)$$

有时我们需要让一个狄拉克旋量中的左手部分或是右手部分单独拿出来与其它场耦合。容易看出, 矩阵

$$P_L = \frac{1-\gamma^5}{2}, \quad P_R = \frac{1+\gamma^5}{2} \quad (3.47)$$

分别是左手旋量和右手旋量的投影算符。

狄拉克表象 以上我们都设 ψ 的各分量由一个左手旋量和一个右手旋量拼凑而成。这称为手性基或者魏尔表象。⁷在手性基当中, 拉氏量的质量项为

$$-\bar{\psi}m\psi = -m(\chi_L^\dagger\xi_R + \xi_R^\dagger\chi_L),$$

这不是一个对角化的二次型。若做分量变换

$$\psi' = \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_U \psi, \quad \gamma'^\mu = U\gamma^\mu U^\dagger, \quad (3.48)$$

质量项就被对角化了。我们称这种分量选取为质量基或者狄拉克表象。在后面做量子化时会看到, 在魏尔表象中, ψ 的 $p^0 > 0$ 的解包含权重相等的一个左手旋量分量和一个右手旋量分量, $p^0 < 0$ 的解包含权重正好差一个负号的一个左手旋量分量和一个右手旋量分量, 因此狄拉克表象下, ψ 的前两个分量代表 $p^0 > 0$ 的模式而后两个分量代表 $p^0 < 0$ 的模式。实际上, 这就意味着狄拉克表象的前两个分量代表自旋一上一下的两种粒子, 而后两个分量分别代表自旋一上一下的两种反粒子。

容易计算出质量基下

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \quad (3.49)$$

在质量基下, 左手投影算符和右手投影算符分别是

γ 矩阵和Clifford代数

γ 矩阵的乘法规则

γ 矩阵的洛伦兹变换 γ 矩阵的洛伦兹变换需要一些特别的注记。如果我们做像本文定义的那样的洛伦兹变换, 即让旋量场变换, 那么 γ 矩阵在洛伦兹变换下根本就不应该发生变化。但是, 另一方面, $\bar{\psi}\gamma^\mu\psi$ 可以验证的确是矢量。这种看似冲突的情况—— γ 矩阵同时看起来像标量和矢量——当然来自它们代表的是可以作用在单粒子态上的算符, 或者作用在场算

⁷ 虽然我们在讨论场算符, 但是这些场算符可以直接激发出单粒子态, 后者和前者之间的关系是线性的, 所以“基底”或“表象”的术语确实是正确的: 它们的确给出了单费米子的一组基底。

符上的“元算符”，而非场算符本身这一事实。回顾引入狄拉克旋量的过程， γ 矩阵的作用其实是“正确地将左手部分的魏尔旋量和右手部分的魏尔旋量混合起来”，它实际上提供了拉氏量关于旋量场的形式，自然在洛伦兹变换下可以没有变化。

这样在实际计算中我们其实可以在很多时候将旋量有关的量——旋量场和 γ 矩阵——当成没有内部结构的对象，即假装系统中只有标量和矢量，而通过正确地排列旋量场和 γ 矩阵的顺序来完成本应通过旋量指标完成的工作。对矢量当然也可以这么做，实际上研究数值计算时我们将一切都写成矩阵就是在做这件事；在量子场论中不这么做的主要原因是，在量子场论中我们需要频繁讨论矢量的任意张量积，从而保留上下指标会让计算变得直观，而对旋量做的操作则相当有限。

γ 矩阵的物理意义 我们已经看到 $(1-\gamma^5)/2$ 和 $(1+\gamma^5)/2$ 分别是左手和右手的手征投影算符。还可以验证，

$$\gamma^5 P_L \psi = -P_L \psi, \quad \gamma^5 P_R \psi = P_R \psi, \quad (3.50)$$

因此 γ^5 是**手征算符**，其负本征值对应左旋态而正本征值对应右旋态。另一方面， γ^0 则交换一个狄拉克旋量的左手和右手部分。

3.3.2 旋量场的正则量子化

平面波模式

我们还是例行公事地观察狄拉克方程的平面波解，并将其中的场变量用算符代替。接下来将使用手性基。平面波旋量场的分量应该包括这些标记：首先是动量。三维动量给定之后，对时谐场有 $p^0 = \pm\omega_{\mathbf{p}}$ ，但由于旋量场不是实场，因此其正频率部分和负频率部分之间没有简单的关系，即不能统一用 a 和 a^\dagger 表示而需要引入两种模式。在动量给定后，只需要求解

$$(\gamma^\mu p_\mu - m)\psi = 0.$$

洛伦兹变换不会改变这个方程的解的结构，所以为了弄清楚方程的解大致是什么样的我们可以先求解

$$(\gamma^0 p_0 - m)\psi = 0.$$

我们会发现 $p^0 = \pm\omega_{\mathbf{p}}$ 各自对应两个独立的解，分别是

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

容易验证自旋算符和哈密顿量对易，且魏尔旋量的两个分量正好就代表两种不同的自旋。因此，标记平面波旋量场的分量的独立标签包括：三维动量，频率的正负，以及自旋1/2。这样，我们设

$$\psi(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} (a_{\mathbf{p},s} u^s(p) e^{-ip \cdot x} + b_{\mathbf{p},s} v^s(p) e^{ip \cdot x}), \quad (3.51)$$

其中 $s = \pm 1/2$ ， u^s 和 v^s 分别满足（这里 p_i 仍然和 p^i 差一个负号，以和闵可夫斯基时空度规一致）

$$(\gamma^0 \omega_{\mathbf{p}} + \gamma^i p_i) u_s(p) = m u_s(p), \quad (-\gamma^0 \omega_{\mathbf{p}} + \gamma^i p_i) v_s(p) = m v_s(p), \quad (3.52)$$

并且被归一化为

$$(u^r(p))^\dagger u^s(p) = 2\omega_p \delta^{rs}, \quad (v^r(p))^\dagger v^s(p) = 2\omega_p \delta^{rs}, \quad (u^r(p))^\dagger v^s(p) = 0. \quad (3.53)$$

这里 $2\omega_p$ 因子是为了和后面的正则量子化中的动量积分中的 $1/2\omega_p$ 因子抵消, 从而保证正确的归一化, 让动量产生湮灭算符的对易子为 $\delta^3(\mathbf{p}-\mathbf{p})$ 的同时, ψ 和 $i\psi^\dagger$ 的对易子为 $i\delta(\mathbf{x}-\mathbf{y})$ 。

代入哈密顿量, 可以得到(我们会发现哈密顿量(3.46)中的第一项和(3.52)左边的第二项抵消了)

$$H = \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_p (a_{\mathbf{p},s}^\dagger a_{\mathbf{p},s} - b_{\mathbf{p},s}^\dagger b_{\mathbf{p},s}).$$

看起来我们有麻烦了—— b 模式上的粒子产生得越多, 能量越低, 似乎出现了能量的无限下降。然而, 如果对 ψ 场施加反对易关系, 就会得到

$$H = \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_p (a_{\mathbf{p},s}^\dagger a_{\mathbf{p},s} + b_{\mathbf{p},s} b_{\mathbf{p},s}^\dagger - 1),$$

在略去发散但是是常数的真空零点能之后就得到

$$H = \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_p (a_{\mathbf{p},s}^\dagger a_{\mathbf{p},s} + b_{\mathbf{p},s} b_{\mathbf{p},s}^\dagger). \quad (3.54)$$

现在如果我们将 b 看成某种粒子的产生算符, 而将 b^\dagger 看成湮灭算符, 那么我们就完成了量子化——哈密顿量变成了占据数的一次函数, 并且有正确的反对易关系。另一方面, 如果施加对易关系, 虽然也可以得到不会无限下降的哈密顿量, 但是不能将 b 看成某种粒子的产生算符, 而将 b^\dagger 看成湮灭算符, 因为设 $\tilde{b} = b^\dagger$, 就有 $[\tilde{b}, \tilde{b}^\dagger] = -1$, 而不是我们需要的 $+1$ 。

于是, 我们交换 b 和 b^\dagger 的位置, 而要求

$$\psi(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} (a_{\mathbf{p},s} u^s(p) e^{-ip \cdot x} + b_{\mathbf{p},s}^\dagger v^s(p) e^{ip \cdot x}), \quad (3.55)$$

并施加反对易关系

$$\{\psi_a(\mathbf{x}), \psi_b^\dagger(\mathbf{y})\} = \delta^3(\mathbf{x}-\mathbf{y}) \delta_{ab}, \quad (3.56)$$

则哈密顿量就是

$$H = \sum_s \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_p (a_{\mathbf{p},s}^\dagger a_{\mathbf{p},s} + b_{\mathbf{p},s}^\dagger b_{\mathbf{p},s}). \quad (3.57)$$

b 模式和 a 模式互为反粒子。通过将 ψ 场的傅里叶展开中的积分变量换成 $-\mathbf{p}$ (不改变积分值), 可以发现

$$a_{\mathbf{p}}^\dagger e^{-i\omega_p t} = b_{-\mathbf{p}} e^{i\omega_p t},$$

即它们的确互为反粒子。我们通常以 b^\dagger 为反粒子, 虽然这只是习惯问题。在绘制含有反粒子的费曼图时应当令传播子上的箭头和与传播子平行的动量箭头方向相反, 这样顶角的形式无需因为传入的是粒子还是反粒子而改变 (例如, $\gamma^\mu A_\mu \bar{\psi} \psi$ 的形式在这种情况下永远是“一条玻色子线连接一条方向指向顶角的费米子线和一条方向远离顶角的费米子线”), 而又提醒了我们一个传播子是反粒子, 同时顶角的动量守恒关系也非常直观。这可以直观地展示为

$$\begin{array}{c} \xrightarrow{-k} \quad \xleftarrow{k} \quad \xrightarrow{-k} \\ \bullet \quad \bullet \quad \bullet \\ \xrightarrow{\quad} a(-k) = \xleftarrow{\quad} = b^\dagger(k) \xrightarrow{\quad} \end{array}$$

总之, 对旋量场, 由于负能量部分 (即反粒子) 的存在, 对易关系是不适用的, 因此必须选择反对易关系, 这意味着旋量场一定是费米场。

旋量场的偏振

现在我们只是形式地写出了 u 和 v ，但是并没有真的计算出它们具体是多少。计算它们具体是多少是很重要的，因为显然易见，计算偏振方向任意的散射振幅时要用到它们。

首先，对 $\mathbf{p} = 0$ 的情况，有

$$u(p) = \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix}, \quad \xi^\dagger \xi = 1, \quad (3.58)$$

这样和归一化条件(3.53)一致。

不失一般性地，认为 \mathbf{p} 指向 z 轴。我们的策略是这样的：首先写出动量和推动参数 η 之间的关系⁸，然后将 η 代入旋量场的洛伦兹变换计算出动量为 \mathbf{p} 的模式。

设我们做了一个洛伦兹变换，空间动量方向指向 z 轴，即有

$$\begin{pmatrix} E \\ p^3 \end{pmatrix} = \exp \left(\eta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \begin{pmatrix} m \\ 0 \end{pmatrix} = \begin{pmatrix} m \cosh \eta \\ m \sinh \eta \end{pmatrix},$$

另一方面，在 $p = (E, 0, 0, p^3)$ 时

$$u(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix}, \quad (3.59)$$

类似地可以得到

$$v(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ -\sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix}. \quad (3.60)$$

$$\bar{u}^r(p) u^s(p) = 2m \delta^{rs}, \quad u^{r\dagger}(p) u^s(p) = 2\omega_{\mathbf{p}} \delta^{rs}, \quad (3.61)$$

$$\bar{v}^r(p) v^s(p) = -2m \delta^{rs}, \quad v^{r\dagger}(p) v^s(p) = 2\omega_{\mathbf{p}} \delta^{rs}. \quad (3.62)$$

$$\sum_s u^s(p) \bar{u}^s(p) = \not{p} + m, \quad \sum_s v^s(p) \bar{v}^s(p) = \not{p} - m. \quad (3.63)$$

后面我们将会用 ξ 表示 (ξ^1, ξ^2) ，从而 $\xi^\dagger \xi = \sigma^0$ 。

Gordon恒等式

$$\bar{u}(q_2)(q_1^\mu) \quad (3.64)$$

守恒量和洛伦兹不变量

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad (3.65)$$

⁸ 这里有一个可能的问题，虽然我们确定与能量匹配、共同组成四维动量的那个 \mathbf{p} 和标记了自由旋量场的稳定模式的那个 \mathbf{p} 都是空间平移的生成元，万一它们没有被一起“定标”怎么办？比如说，如果正好差一个常数怎么办？自旋就有这样的情况。好在我们已经证明，对后者也有 $E^2 = |\mathbf{p}|^2 + m^2$ ，因此这两个 \mathbf{p} 的确是同一个东西。

旋量场的传播子

现在计算旋量场的各种传播子。旋量场的传播子定义需要特别注意，因为一个 ψ 中含有一种粒子和一种反粒子。所幸，一个只含有空间定域的正费米子的态还是可以通过 $\psi^\dagger|0\rangle$ 创建，因为反费米子的湮灭算符作用在真空态上给出零，没有任何贡献。看起来，用 $\langle 0|\psi\psi^\dagger|0\rangle$ 做正费米子的传播子而用 $\langle 0|\psi^\dagger\psi|0\rangle$ 做反费米子的传播子应该是正确的。然而，可以发现，这样的传播子不是洛伦兹协变的。能够从相对论性量子场论的路径积分计算出来的关联函数肯定是洛伦兹协变的，既然 $\langle 0|T\psi\psi^\dagger|0\rangle$ 不是洛伦兹协变，那么可能并不能很容易地计算它，可能Wick定理都不一定成立，等等。因此我们应该计算和 $\langle 0|T\psi\psi^\dagger|0\rangle$ 只差一个可逆矩阵变换但是是洛伦兹协变的 $\langle 0|T\psi\bar{\psi}|0\rangle$ 。⁹

旋量场的路径积分量子化

其实用路径积分算这个更快

$$D_F(x-y) = \langle 0|T\psi(x)\bar{\psi}(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i0^+} e^{-ip \cdot (x-y)}. \quad (3.66)$$

一个一般的场论中的费米子的费曼规则比较难写，因为其传播子在坐标或是动量固定后仍然是矩阵（含有 γ 矩阵），相乘和缩并的细节处理起来很棘手。不过我们会发现，实用的高能物理理论都是规范场论，其中费米子的费曼规则会很容易写出，因为一张图中不存在涉及一个以上费米子的顶角，即每个顶角中都只有一个费米子进去，一个费米子出来，矩阵 γ^μ ， u 和 v （列矢量按照 s 指标排成矩阵）只需要按照这个费米子的“世界线”依次累乘即可。

3.3.3 非相对论极限

在非相对论极限下旋量场当然也会退化成薛定谔场。

3.4 矢量场

3.4.1 矢量场的麦克斯韦方程和布洛卡方程

对称性分析

由于自由场导数阶数的限制，出现在拉氏量中的只能是 A^μ 和 $\partial^\nu A^\mu$ 构成的一次或二次不变量。当然，实际上也可以出现 $\partial_\mu A^\nu$ 或者 $\partial_\mu A_\mu$ 这种，但因为它们都可以使用 $\partial^\nu A^\mu$ 表示出来，故没有必要考虑它们。只含有 A^μ 二次不变量为 $A^\mu A_\mu$ ，没有一次不变量；只含有 $\partial^\mu A^\nu$ 的一次不变量是它自我缩并得到的 $\partial^\mu A_\mu$ ，二次的不变量是两个 $\partial^\mu A^\nu$ 缩并得到的 $\partial^\mu A^\nu \partial_\mu A_\nu$ 和 $\partial^\mu A^\nu \partial_\nu A_\mu$ 。由于参数都是标量， $\partial^\mu A^\nu$ 不能和参数缩并，也不能和 A^μ 缩并（ $C^\nu A^\mu \partial_\nu A_\mu$ 要求参数是矢量， $A^\mu A^\nu \partial_\mu A_\nu$ 是三次项），因此我们得到了所有可能的不变量。从而拉氏量形如

$$\mathcal{L} = C_1 A^\mu A_\mu + C_2 \partial^\mu A_\mu + C_3 \partial^\mu A^\nu \partial_\mu A_\nu + C_4 \partial^\mu A^\nu \partial_\nu A_\mu.$$

⁹ 没有什么规定了传播子一定要定义成 $\langle 0|\psi\psi^\dagger|0\rangle$ 的形式；计算传播子的目的是获得一种简便易行的方式，可用于从一个拉氏量中提取时间演化算符在多粒子态（可以使用场的连乘构造）下的矩阵元，因此传播子只需要大体上长成 $\langle 0|\psi\psi^\dagger|0\rangle$ 的形式即可，差一个线性变换是完全可以的，只要所有重要的信息都能够恢复出来。

代入欧拉-拉格朗日方程可以看出, C_2 项在运动方程中不会引入任何项, 故略去。于是

$$\mathcal{L} = C_1 A^\mu A_\mu + C_3 \partial^\mu A^\nu \partial_\mu A_\nu + C_4 \partial^\mu A^\nu \partial_\nu A_\mu.$$

代入欧拉-拉格朗日方程, 得到

$$\partial_\mu (C_3 \partial^\mu A^\nu + C_4 \partial^\nu A^\mu) = C_1 A^\nu. \quad (3.67)$$

我们首先考虑 $C_3 = -C_4$ 时的特殊情况。重新定义各系数, 使得

$$\mathcal{L} = -\frac{1}{2} \partial^\mu A^\nu \partial_\mu A_\nu + \frac{1}{2} \partial^\mu A^\nu \partial_\nu A_\mu + \frac{m^2}{2} A_\mu A^\mu, \quad (3.68)$$

对应的,

$$\partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) + m^2 A^\nu = 0. \quad (3.69)$$

常定义

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (3.70)$$

于是就有

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu. \quad (3.71)$$

现在回到一般情况。我们指出这样一个结论: 无论 C_3, C_4 取什么值, 对应的场 A^μ 都可以和 $C_3 = -C_4$ 时的某个场 A'^μ 建立一一对应。因此布洛卡方程(3.69)就不失一般性地描写了所有的矢量场的运动方程。

(3.69)在 $m \neq 0$ 时可以推导出克莱因-高登方程。注意到

$$m^2 \partial_\nu A^\nu = \partial_\nu \partial_\mu \partial^\nu A^\mu - \partial_\mu \partial^\mu \partial_\nu A^\nu = 0,$$

于是

$$\partial_\mu A^\mu = 0. \quad (3.72)$$

回代入(3.69), 发现其左边第二项为零, 于是

$$\partial_\mu \partial^\mu A^\nu + m^2 A^\nu = 0.$$

于是(3.69)就约化成了(3.72)和四个克莱因-高登方程。而当 $m = 0$ 时, 运动方程在规范变换

$$A^\mu \longrightarrow A'^\mu = A^\mu + \partial^\mu \varphi \quad (3.73)$$

下不变。这意味着矢量场 A^μ 的四个自由度实际上是多余的。¹⁰换言之, 存在规范冗余性。

¹⁰显然, 只要选定了一个 φ , 同一个时间点上的 A^μ 和 A'^μ 之间必定可以建立起一一对应关系。形象地说, 不同 φ 对应的 A' 的运行轨迹相互平行, 因此只需要其中一条轨迹就能够确定所有轨迹。选取特定的一条轨迹就是选取一个规范。规范自由度——也就是决定“实际的轨道是哪一条”的自由度——是一个隐藏的额外自由度。这里的情况和对称性自发破缺有点类似, 在后者中, 隐藏的自由度是序参量。不同的隐藏的额外自由度取值将系统的态空间分成了互不相交的分支。可以认为规范自由度不是物理的自由度, 也就是说它仅仅出现在拉氏量中, 而规范自由度取值不同的状态在希尔伯特空间中应该被认为是同样的状态。选取一个规范意味着先假定规范自由度取值不同的状态真的是不一样的, 然后取状态空间中的一个分支。

重矢量场的哈密顿量

矢量场的共轭动量为

$$\pi_\mu = \partial_\mu A^0 - \partial^0 A_\mu,$$

或者写成

$$\pi^\mu = \partial^\mu A^0 - \partial^0 A^\mu. \quad (3.74)$$

注意到 $\pi^0 = 0$ ，因此可以只讨论其空间部分 $\boldsymbol{\pi}$ 。在质量 m 不为零时场没有规范不变性，可以直接做计算得到

$$A^0 = -\frac{1}{m^2} \nabla \cdot \boldsymbol{\pi}, \quad (3.75)$$

以及

$$\partial_0 A^0 = -\partial_i A^i = -\nabla \cdot \mathbf{A}, \quad (3.76)$$

哈氏量为

$$\mathcal{H} = \frac{1}{2} \boldsymbol{\pi}^2 + \frac{1}{2m^2} (\nabla \cdot \boldsymbol{\pi})^2 + \frac{1}{2} (\nabla \times \mathbf{A})^2 + \frac{1}{2} m^2 \mathbf{A}^2. \quad (3.77)$$

m 出现在了分母中，这意味着无质量的场需要额外处理。

麦克斯韦理论的经典规范选取

现在来处理无质量的场。其运动方程为

$$\partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) = 0.$$

无论 $\partial_\mu A^\mu$ 是什么，总可以找到一个 φ 使得

$$\partial_\mu \partial^\mu \varphi = -\partial_\mu A^\mu,$$

从而对应的有

$$\partial_\mu A'^\mu = 0.$$

于是不失一般性地强行要求(3.72)对 $m = 0$ 时的矢量场成立。这称为选取了洛伦兹规范。选取了洛伦兹规范意味着，实际的场自由度只有三个。知道了 A 的三个分量就可以计算出第四个。当然，这不是唯一的规范选取方式。例如可以直接要求 $A^0 = 0$ ，称为辐射规范。选取洛伦兹规范的好处在于，方程(3.72)是洛伦兹协变的，因此在做量子化时能够直接套用正则量子化关系而不必担心场方程不是洛伦兹协变而产生的修正。

我们施加洛伦兹规范。当然也可以选取别的规范，但这可能会破坏洛伦兹协变性，从而导致我们得到的哈密顿动力学实际上是带有约束的，从而给之后做量子化带来麻烦。此时运动方程为

$$\partial_\mu \partial^\mu A^\nu = 0. \quad (3.78)$$

拉氏量(3.71)直接导出的不是这个方程，于是我们使用另一个能够直接导出(3.78)的拉氏量

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2. \quad (3.79)$$

可以看到这个拉氏量多出了一项，这个实际上就是所谓的规范固定项。在给定了洛伦兹规范的前提下，这个拉氏量实际上就和 $m = 0$ 的(3.71)是等价的。此时

$$\pi^0 = -\partial_\mu A^\mu, \quad \pi^i = \partial^i A^0 - \partial^0 A^i. \quad (3.80)$$

当然，由洛伦兹规范， π^0 就是零，不过我们完全可以算出哈密顿量之后再施加洛伦兹规范。哈密顿量为

$$\mathcal{H} = \quad (3.81)$$

3.4.2 无质量矢量场的正则量子化

无质量矢量场的正则量子化涉及很多棘手的细节。

横场量子化

使用(3.7)展开一个无质量矢量场为

$$A_\mu(\mathbf{x}, t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \sum_{r=0}^3 \epsilon_\mu^r(\mathbf{p}) (a_{\mathbf{p},r}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega_{\mathbf{p}}t} + a_{\mathbf{p},r} e^{i\mathbf{p}\cdot\mathbf{x}-i\omega_{\mathbf{p}}t}), \quad (3.82)$$

由于没有质量,

$$\omega_{\mathbf{p}} = |\mathbf{p}|. \quad (3.83)$$

ϵ^r 为一组闵可夫斯基时空的基矢量, 称它们为偏振矢量, 也即,

$$(\epsilon^r)_\mu (\epsilon^{r'})^\mu = \eta^{rr'}. \quad (3.84)$$

为了确定偏振矢量, 通常要求

$$\epsilon^1 \cdot \mathbf{p} = \epsilon^2 \cdot \mathbf{p} = 0, \quad (3.85)$$

并认为 ϵ^0 是类时的, 而 $\epsilon^{1,2,3}$ 是类空的。这样, 当 $\mathbf{p}^\mu \propto (1, 0, 0, 1)$, 即 \mathbf{p} 指向 z 轴时, 我们有

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.86)$$

\mathbf{p} 取其它值时只需要对(3.86)做洛伦兹变换即可, 因为 ϵ 的定义完全是洛伦兹协变的。

可以计算出

$$\pi^\mu(\mathbf{x}, t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{p}}}{2}} i \sum_{r=0}^3 (\epsilon^r)^\mu(\mathbf{p}) (a_{\mathbf{p},r} e^{i\mathbf{p}\cdot\mathbf{x}-i\omega_{\mathbf{p}}t} - a_{\mathbf{p},r}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega_{\mathbf{p}}t}), \quad (3.87)$$

施加正则对易关系, 通过计算得到

$$[a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}^\dagger] = -\eta_{\lambda\lambda'} (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \quad [a_{\mathbf{p},\lambda}^\dagger, a_{\mathbf{p}',\lambda'}^\dagger] = [a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}] = 0. \quad (3.88)$$

$\lambda = 1, 2, 3$ 时对易关系是正确的, 但是 $\lambda = 0$ 给出了一个不正常的对易关系

$$[a_{\mathbf{p},0}, a_{\mathbf{p}',0}^\dagger] = -(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}').$$

例如, 它产生的同样的单粒子态的内积将会是一个负数, 这和我们通常对单粒子态的认识不符。此外, 哈密顿量成为

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} \left(-a_{\mathbf{p},0}^\dagger a_{\mathbf{p},0} + \sum_{i=1}^3 a_{\mathbf{p},i}^\dagger a_{\mathbf{p},i} \right), \quad (3.89)$$

因此能量非正定。显然这些问题都和 $a_{\mathbf{p},0}^\dagger$ 有关, 也就是说来自一个非物理的自由度。会有非物理的自由度是显然的, 因为我们在处理一个有规范不变性的场却从来没有选取过一个规范。现在我们处理的是量子场, 因此既可以直接对场做约束, 也可以缩小态空间的范围。

我们先尝试直接将洛伦兹规范作用在场上，即要求对场算符有

$$\partial^\mu A_\mu = 0.$$

然而，这是不可能的，经过检验，Gupia-Blenler量子化条件

$$\partial^\mu A_\mu^{(+)} |\psi\rangle = 0 \quad (3.90)$$

是一个可行的方案。它实际上约束了态空间的范围。代入(3.82)，并注意到 ϵ^1 与 ϵ^2 和四维动量做内积得到零，我们发现

$$(a_{\mathbf{p},0} - a_{\mathbf{p},3}) |\psi\rangle = 0. \quad (3.91)$$

这意味着在无质量矢量场的态空间中哈密顿量实际上是

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} (a_{\mathbf{p},1}^\dagger a_{\mathbf{p},1} + a_{\mathbf{p},2}^\dagger a_{\mathbf{p},2}). \quad (3.92)$$

于是负能量问题也就解决了。哈密顿量中没有出现的量可以直接被略去，因为它们对系统的动力学不产生任何影响。于是取

$$A_\mu(\mathbf{x}, t) = \int \frac{d^3\mathbf{x}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \sum_{r=1}^2 \epsilon_\mu^r(\mathbf{p}) (a_{\mathbf{p},r}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega_{\mathbf{p}}t} + a_{\mathbf{p},r} e^{i\mathbf{p}\cdot\mathbf{x}-i\omega_{\mathbf{p}}t}), \quad (3.93)$$

以及

$$\pi^\mu(\mathbf{x}, t) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \sqrt{\frac{\omega_{\mathbf{p}}}{2}} i \sum_{r=1}^2 (\epsilon^r)^\mu(\mathbf{p}) (a_{\mathbf{p},r} e^{i\mathbf{p}\cdot\mathbf{x}-i\omega_{\mathbf{p}}t} - a_{\mathbf{p},r}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega_{\mathbf{p}}t}), \quad (3.94)$$

重新计算对易关系得到

$$[a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}^\dagger] = \delta_{\lambda\lambda'} (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \quad [a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}^\dagger] = [a_{\mathbf{p},\lambda}, a_{\mathbf{p}',\lambda'}] = 0, \quad \lambda = 1, 2. \quad (3.95)$$

守恒量

下面我们推导动量和自旋角动量的公式。轨道角动量的由于是动量衍生出来的量，我们暂不考虑。首先假设 $p^\mu \propto (1, 0, 0, 1)$ 。按照(2.91)可以计算得到

$$S_3 = \int d^3\mathbf{x} \mathcal{S}_3 = i \int \frac{d^3\mathbf{p}}{(2\pi)^3} (-a_{\mathbf{p},1} a_{\mathbf{p},2}^\dagger + a_{\mathbf{p},1}^\dagger a_{\mathbf{p},2} + a_{\mathbf{p},2} a_{\mathbf{p},1}^\dagger - a_{\mathbf{p},2}^\dagger a_{\mathbf{p},1}),$$

另外两个方向上的自旋角动量都是零。

我们原本预期矢量场会有三个自由度（因为(3.72)消除掉了一个自由度），但是实际上无质量矢量场只有两个自由度。导致这一切的原因当然是无质量这个事实——它使得四维动量 p 不再能够写成 $(1, 0, 0, 0)$ 这样的形式，而只能写成 $(1, 0, 0, 1)$ 这样，从而让 A^0 和 A^3 相互抵消了。从洛伦兹群在态空间上的表示出发可以更好地看待这个问题： $m = 0$ 时洛伦兹群保持动量不变的小群不再是旋转群。以一种更加物理的视角，无质量矢量场对应的粒子一直在以光速运动，不能找到一个相对它静止的参考系，因此对一个这样的粒子，实际上总是有一个特定的空间方向即它的运动方向，为了保持协变性，其自旋只能够沿着这个方向。换言之，此时有意义的实际上是螺旋度而不是三维的角动量 \mathbf{S} ，即其内禀自由度是平面旋转群（以运动方向为轴旋转）的表示而不是三维旋转群的表示。而对有质量的粒子，总是可以找到一个相对它静止的参考系，在这个参考系中空间是各向同性的，因此可以应用 $SO(3)$ 的表示。这和经典电磁场的偏振只有两个方向是对应的。

传播子

由于洛伦兹度规的空间部分全部都是负的，我们有

$$D_F = \langle 0|TA_\mu(x)A_\nu(y)|0\rangle = -\int \frac{d^4p}{(2\pi)^4} \frac{i\eta_{\mu\nu}}{p^2 + i0^+} e^{-ip\cdot(x-y)} = \int \frac{d^4p}{(2\pi)^4} \frac{i\delta_{\mu\nu}}{p^2 + i0^+} e^{-ip\cdot(x-y)}. \quad (3.96)$$

有一点应该指出：规范选取(3.90)是作用在态上的。这就是说，

3.4.3 重矢量场的正则量子化

3.4.4 路径积分量子化和规范固定

我们现在从路径积分的角度量子化无质量矢量场。我们首先尝试朴素地直接将 $m = 0$ 的(3.71)放入 e 指数，通过分部积分，得到

$$S = -\frac{1}{4} \int d^4x F_{\mu\nu} F^{\mu\nu} = \frac{1}{2} \int d^4x A^\mu (\eta_{\mu\nu} \partial^2 - \partial_\mu \partial_\nu) A^\nu, \quad (3.97)$$

那么，合乎情理的下一步就是取积分核 $(\eta_{\mu\nu} \partial^2 - \partial_\mu \partial_\nu)$ 的逆。现在问题出现了：这个积分核根本就没有逆——其行列式为零。因此，直接将 $m = 0$ 的(3.71)放入 e 指数得到的结果是发散的。这当然是因为我们没有做规范固定。对无相互作用的无质量矢量场，局域对称操作

$$A^\mu \longrightarrow A^\mu + \partial^\mu \alpha$$

不改变哈密顿量。实际上这是一个规范对称性，即这个局域操作不会改变量子态。然而，朴素地做路径积分不会考虑这一点，从而路径积分会对本该只求一遍的场构型求和无数多遍。

本节将使用**Faddeev-Popov量子化**方法，大致的思路是通过适当手段插入一个 δ 函数，对一族规范等价的场构型，只挑选其中一个纳入路径积分。这一方法的不足之处在于可能还不能消除所有的规范冗余性，但是对微扰计算来说，由于场构型的变动总是充分小的，Faddeev-Popov量子化足够给出可靠的答案。我们在朴素版本的配分函数的 A 的积分后面插入

$$1 = \int \mathcal{D}\alpha \delta(G(A^\alpha)) \det \left(\frac{\delta G(A^\alpha)}{\delta \alpha} \right),$$

其中 A^α 就是 $A + \partial\alpha$ 。这样就有

$$\begin{aligned} Z &= \int \mathcal{D}A \int \mathcal{D}\alpha \delta(G(A^\alpha)) \det \left(\frac{\delta G(A^\alpha)}{\delta \alpha} \right) e^{iS[A]} \\ &= \int \mathcal{D}\alpha \int \mathcal{D}A^\alpha \delta(G(A^\alpha)) \det \left(\frac{\delta G(A^\alpha)}{\delta \alpha} \right) e^{iS[A^\alpha]}, \end{aligned}$$

其中 A^α 的积分测度和 A 完全一样，且规范对称性意味着 $S[A^\alpha]$ 和 $S[A]$ 也一样。为了保持洛伦兹协变性， $G(A)$ 可以选择为这样：

$$G(A^\alpha) = \partial^\mu A_\mu^\alpha - \omega(x) = \partial^\mu A_\mu + \partial^2 \alpha - \omega(x), \quad (3.98)$$

其中 ω 是一个任意的标量函数。这样，雅可比行列式因子就可以提出积分号外，得到

$$\begin{aligned} Z &= (\det \partial^2) \int \mathcal{D}\alpha \int \mathcal{D}A^\alpha \delta(\partial^\mu A_\mu^\alpha - \omega) e^{iS[A^\alpha]} \\ &= (\det \partial^2) \int \mathcal{D}\alpha \int \mathcal{D}A \delta(\partial^\mu A_\mu - \omega) e^{iS[A]}, \end{aligned}$$

这里我们已经重新标记 A^α 为 A 了。既然 ω 可以随意变动，我们不妨对所有的 ω 做一次求和，得到

$$\begin{aligned} Z' &= \int \mathcal{D}\omega e^{-i \int d^4x \frac{\omega^2}{2\xi}} (\det \partial^2) \int \mathcal{D}\alpha \int \mathcal{D}A \delta(\partial^\mu A_\mu - \omega) e^{iS[A]} \\ &= (\det \partial^2) \int \mathcal{D}\alpha \int \mathcal{D}A e^{iS[A]} \exp\left(-i \int d^4x \frac{1}{2\xi} (\partial^\mu A_\mu)^2\right). \end{aligned}$$

现在可以把前面的一大堆因子全部扔掉，得到最后的配分函数：

$$Z = \int \mathcal{D}A \exp\left(i \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial^\mu A_\mu)^2\right)\right). \quad (3.99)$$

可以看到，这实际上就是向拉氏量加了一个**规范固定项**，破坏了拉氏量的规范对称性，但是得到的物理是完全一样的。 $\xi = 1$ 称为**费曼规范**，和(3.79)是一致的； $\xi = 0$ 称为**朗道规范**。

做完规范固定之后，可以按照正常的手续求出传播子。通过分部积分将作用量写成

$$S = \int d^4x (A^\mu (\eta_{\mu\nu} \partial^2 - (1 - \frac{1}{\xi}) \partial_\mu \partial_\nu) A^\nu),$$

于是

$$\langle 0 | T A_\mu(x) A_\nu(y) | 0 \rangle = \int \frac{d^4x}{(2\pi)^4} \frac{i}{-k^2 \eta^{\mu\nu} + (1 - \frac{1}{\xi}) k^\mu k^\nu},$$

并需要加上一个无穷小虚部。我们会发现

$$\left(-k^2 \eta^{\mu\nu} + (1 - \frac{1}{\xi}) k^\mu k^\nu\right) \left(\eta_{\nu\rho} - (1 - \xi) \frac{k_\nu k_\rho}{k^2}\right) = -k^2 \delta^\mu_\rho,$$

于是

$$D_F(x-y)_{\mu\nu} = \langle 0 | T A_\mu(x) A_\nu(y) | 0 \rangle = \int \frac{d^4k}{(2\pi)^4} \frac{-i}{k^2 + i0^+} \left(\eta_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2}\right). \quad (3.100)$$

这里多出来的一个负号是因为，在标量场和旋量场中，拉氏量动能项的两个导数算符分别作用在两个场上，将一个场视为入射，一个视为出射，切换到动量空间，得到的是 $-ik \cdot ik = k^2$ ，而在这里，两个导数算符都作用在右边的场上，得到的是 $(ik)^2 = -k^2$ 。请注意上式已经是做完规范固定之后的结果，无需再额外做任何规范固定，这就是说，对一张费曼图的中间态求和时， μ 和 ν 的确要取遍0到4。只有入射光子外线和出射光子外线需要特别选择。

第四章 用量子场论计算可观察量

4.1 散射

高能物理实验中涉及的物理过程的时间尺度通常远远小于我们能够观察的时间尺度，并且一般很难研究束缚态（部分是因为束缚态问题基本上是凝聚态研究的范畴而此时除了库仑定律以外也不需要太多物理，部分是因为微扰计算束缚态问题非常困难）。因此，大部分情况下我们都只需要考虑初末态均在无穷远处的散射问题即可。因此，需要计算的主要就是 S 矩阵——实际上是动量表象下的 S 矩阵，以下如无特殊说明， S 矩阵指的就是动量表象下的 S 矩阵。然而，实验中能够制备的含有一定量的动量确定的粒子的态是完整的哈密顿量的本征态（而不是自由理论的本征态），此时的多粒子态的能量不是其单粒子能量的简单相加。原则上我们同样可以微扰计算完整的哈密顿量的本征态，但是这将是非常费力并且困难的。

然而，在实际的散射问题中，无论是入射态还是出射态的结构都相对简单，因为实际的物理粒子的波函数的分布不可能是遍布全空间的平面波，而只能是动量大体上确定的波包，而将时间推到 ∞ 或是 $-\infty$ 时这些波包会相隔得足够远，从而它们之间没有相互作用，整个量子态的能量也就是各个粒子的能量之和。（或者等价地说，入射和出射时相互作用可能认为是“关闭”的，而只有粒子相隔得足够近才被打开）因此，入射态和出射态实际上仍然能够写成

$$|p_1, p_2, \dots, p_n\rangle = \sqrt{2\omega_{p_1}} \cdots \sqrt{2\omega_{p_n}} a_{p_1}^\dagger a_{p_2}^\dagger \cdots a_{p_n}^\dagger |\Omega\rangle$$

的形式，但是此时的 a 算符——可以记为 a_{in} 或 a_{out} ——已经不再是自由场（“裸的”场）的 a 算符了，因为它创建的单粒子态是实际的、加入了相互作用的哈密顿量的本征态，即它创建的单粒子态是物理粒子，而不是没有相互作用时的裸粒子；同样 ω 也不再是裸的单粒子能量。¹在海森堡绘景下， S 矩阵的矩阵元基本上就具有

$$\langle p_1, p_2 | k_1, k_2 \rangle = \sqrt{2\omega_{p_1}} \sqrt{2\omega_{p_2}} \sqrt{2\omega_{k_1}} \sqrt{2\omega_{k_2}} \langle \Omega | a_{p_2} a_{p_1} a_{k_1}^\dagger a_{k_2}^\dagger | \Omega \rangle$$

这样的形式，这意味着只要能够找到 a_{in} ， a_{out} 和裸场 a 之间的关系，即可确定关联函数和 S 矩阵的关系。为了看出 a_{in} ， a_{out} 和裸场 a 之间的关系，我们可以尝试为这种近乎独立的 a_{in} 和 a_{out} 所描述的“自由粒子”写下一个有效理论，显然这个有效理论是完整的、带有相互作用的理论不断重整化的结果，相互作用会修正“自由理论”的参数。由于只考虑单粒子，相互作

¹ 我们对入射态和出射态的要求是非常严格的：一方面，出现在这些态中的粒子一定是能够稳定存在的粒子，寿命有限的准粒子（对应完整的哈密顿量的展宽的能级，可以看成是近似的本征态，但是因为各种扰动而不是严格的本征态）不能出现在入射态和出射态中；我们还进一步要求这些粒子之间的距离足够远，以至于可以近似看成自由的；最后我们还要求这些态被正确地归一化了，从而引入场强重整化因子。

凝聚态理论考虑的系统状态比这多得多：寿命有限的准粒子时常需要被讨论，并且通常不会一次重整化就把所有相互作用都去除而只留下带自能修正的粒子。例如，费米液体理论中的“电子”实际上是已经经过相互作用修正的能带电子，但是即使在系统基态中这些准粒子的相互碰撞仍然需要纳入考虑。

用带来的修正实际上就是所谓的自能修正：能够调整的参数包括质量项和 $\partial_\mu \phi \partial^\mu \phi$ 项，对后者的调节等价于对场本身的变换。（此外，由于只考虑单粒子，且四维动量 p 的安排非常接近在壳，粒子无法衰变，因此自能修正没有虚部）因此我们得出结论： a_{in} ， a_{out} 和裸场 a 之间应该差一个常数，这个常数是场强重整化引入的因子。这就确定了关联函数和 S 矩阵之间的关系：在统一到动量空间中之后，两者首先由于相互作用修正的原因会差一个场强重整化因子；动量空间中的关联函数的解析性质和自由场的关联函数非常相似，只不过前者的极点给出的“质量”会出现跑动；相比之下， S 矩阵要光滑很多，因为没有极点。可以从关联函数计算 S 矩阵，但是不能反过来，因为 S 矩阵仅仅包含初态位于 $-\infty$ 而末态位于 ∞ 的过程。因此应当有：

$$\prod_i \frac{\text{renormalization factors}}{\omega - \omega_k + i0^+} \times \langle p|S|k \rangle \propto \text{Fourier transformation of } \langle \Omega | \phi(x_1) \cdots \phi(x_n) | \Omega \rangle.$$

虽然我们做出了以上形式的论证，完整地做Wilson重整化群计算而得到场强重整化因子当然还是非常繁琐的，因此本节尝试以一种更加简单的方法建立微扰计算 S 矩阵的方法。本节将始终在以下对高能物理来说非常一般的假设下工作：系统具有完整的洛伦兹对称性；入射和出射态中各个粒子相隔足够远；只有接近关联函数极点时的动量-能量安排才是值得分析的，因为其它时候实验现象也不明显。

本节将首先介绍一些可以使用 S 矩阵计算的物理量，以展示计算 S 矩阵的必要性，然后给出微扰计算 S 矩阵的方法。

4.1.1 可观察物理量

散射截面

$$\langle p_1, p_2, \dots, p_m | iT | k_1, k_2, \dots, k_n \rangle = i(2\pi)^4 \delta^4 \left(\sum_i p_i - \sum_j k_j \right) \mathcal{M}(k_1, k_2, \dots, k_n \rightarrow p_1, p_2, \dots, p_m). \quad (4.1)$$

跃迁率

非相对论极限

我们知道 S 矩阵可以用所谓李普曼-施温格方程写出。非相对论性单粒子量子力学中通常采取这样的归一化方案：

$$\hat{T} = \hat{H}' + \hat{H}' \frac{1}{E - \hat{H}_0} \hat{H}' + \cdots, \quad \hat{S} = 1 - 2\pi i \delta(E_f - E_i) \hat{T},$$

4.1.2 S 矩阵和关联函数的关系

标量场编时格林函数的解析结构和LSZ约化公式

考虑一个 n 点关联函数 $\langle \Omega | T \phi(x_1) \cdots \phi(x_n) | \Omega \rangle$ 。我们对 x_1 做傅里叶变换，这是因为实际上用于计算 S 矩阵矩阵元的本征态都是动量本征态，因此我们首先应该在 n 点关联函数中引入动量的概念。我们没有直接使用用动量标记的产生湮灭算符，因为我们还希望用频率代替时间，而“频率”并不是量子数，因此不是任何一套产生湮灭算符的标签的一部分，如果使用用动量标记的产生湮灭算符就还需要对时间做一次傅里叶变换，表达式看起来不

协变了。反之，对四维坐标 x_1 做完傅里叶变换之后，我们就可以用四维动量 k_1 代替 x_1 作为标记了，洛伦兹协变性就很明显。这样得到的动量空间关联函数不保证入射和出射的四维动量是在壳的。（另一方面， S 矩阵肯定是在壳的）

我们有

$$\begin{aligned}
 & \int d^4 x_1 e^{ik_1 \cdot x_1} \langle \Omega | T \phi(x_1) \phi(x_2) \cdots \phi(x_n) | \Omega \rangle \\
 &= \left(\int_{-\infty}^{T_-} + \int_{T_-}^{T_+} + \int_{T_+}^{\infty} \right) dt_1 \int d^3 \mathbf{x}_1 e^{i\omega_1 t_1 - i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | T \phi(x_1) \phi(x_2) \cdots \phi(x_n) | \Omega \rangle \\
 &= \int_{-\infty}^{T_-} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | T[\phi(x_2) \cdots \phi(x_n)] \phi(x_1) | \Omega \rangle \\
 &+ \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
 &+ \int_{T_-}^{T_+} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | T[\phi(x_1) \phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
 \end{aligned} \tag{4.2}$$

最后一个等号只需要令 T^+ 充分大， T_- 充分小即可成立。做完傅里叶变换之后的结果肯定具有奇异性：在相互作用下稳定的单粒子态有确定的动量和能量，因此在 ω_1 和 \mathbf{k}_1 之间满足正确的色散关系时，上式发散。（多粒子态由于能量给定时其中某个动量可以连续变化，不会产生奇异性，但是会产生有限高度的峰）

高能物理实验中实际会测量的区域通常就在这种会导致奇异性的 ω 和 \mathbf{k} 的取值附近，因为这里效应最明显。在(4.2)的极点附近， T_- 到 T_+ 区段的积分基本上没有什么作用，因为其奇异性弱于 T_+ 到 ∞ 的积分和 $-\infty$ 到 T_- 的积分。因此之后我们不再考虑 T_- 到 T_+ 区段的积分。

现在使用常用的技巧：在算符连乘积序列中插入一个完备关系，从而将 n 点关联函数转化为 $n-1$ 点关联函数。即使在加入相互作用之后，动量仍然是好量子数，因此可以有完备关系

$$1 = \sum_{\Lambda} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} |\Lambda_{\mathbf{p}} \rangle \langle \Lambda_{\mathbf{p}}|,$$

其中 Λ 标记了冗余的量子数。对(4.2)最后一步的第二项，插入完备关系得到

$$\begin{aligned}
 & \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
 &= \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) \sum_{\Lambda} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} |\Lambda_{\mathbf{p}} \rangle \langle \Lambda_{\mathbf{p}}| T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
 &= \sum_{\Lambda} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) | \Lambda_{\mathbf{p}} \rangle \langle \Lambda_{\mathbf{p}} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
 \end{aligned}$$

因子 $\langle \Omega | \phi(x_1) | \Lambda_{\mathbf{p}} \rangle$ 可以进一步展开。首先，因子 $\langle \Omega | \phi(0) | \Lambda_{\mathbf{k}_1} \rangle$ 实际上只在 $|\Lambda_{\mathbf{k}_1} \rangle$ 是单粒子态时才能够有非零值（因为它可以写成一个经过场强重整化的产生算符作用在真空态上，而入射态和出射态可以近似看成自由理论，从而被夹在 $\langle \Omega | \cdot | \Omega \rangle$ 中的产生算符和湮灭算符数目必须一致）。其次，我们注意到，实际上可以将 $\phi(x_1)$ 写成 $e^{iP \cdot x} \phi(0) e^{-iP \cdot x}$ ，其中 P 是全空间的动量算符，或者说平移群的生成元。在有相互作用的情况下，这样作用的结果并不方便计算，因为其中包含一个时间演化算符，但是我们总是可以形式地写出这样的式子。于是

$$\langle \Omega | \phi(x) | \Lambda_{\mathbf{p}} \rangle = \langle \Omega | \phi(x) | \Lambda_{\mathbf{p}} \rangle e^{-iP \cdot x} |_{p^0=E_{\mathbf{p}}}.$$

于是就有

$$\begin{aligned}
& \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
&= \sum_{\Lambda} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \int_{T_+}^{\infty} dt_1 e^{i(\omega_1 - E_{\mathbf{p}})t_1} \int d^3 \mathbf{x}_1 e^{-i(\mathbf{k}_1 - \mathbf{p}) \cdot \mathbf{x}_1} \langle \Omega | \phi(0) | \Lambda_{\mathbf{p}} \rangle \langle \Lambda_{\mathbf{p}} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
&= \sum_{\Lambda} \int \frac{d^3 \mathbf{p}}{(2\pi)^3} \left(-\frac{1}{i(\omega_1 - E_{\mathbf{p}} + i0^+)} e^{i(\omega_1 - E_{\mathbf{p}})T_+} \right) (2\pi)^3 \delta(\mathbf{k}_1 - \mathbf{p}) \\
&\quad \times \langle \Omega | \phi(0) | \Lambda_{\mathbf{p}} \rangle \langle \Lambda_{\mathbf{p}} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
&= \sum_{\Lambda} \frac{i}{\omega_1 - E_{\mathbf{k}_1}(\Lambda) + i0^+} e^{i(\omega_1 - E_{\mathbf{p}})T_+} \langle \Omega | \phi(0) | \Lambda_{\mathbf{k}_1} \rangle \langle \Lambda_{\mathbf{k}_1} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
\end{aligned}$$

上式中（以及之后）， $E_{\mathbf{k}}$ 代表的都是 $|\Lambda_{\mathbf{k}_1}\rangle$ 的能量，不是裸粒子的能量，已经加入了自能修正。在这里我们使用标准的在对时间的积分中引入无穷小虚部，让 ω_1 的奇点出现在下半平面的方法；实际上，如果不这样，积分也无法收敛。我们现在稍微改变一下 $|\Lambda_{\mathbf{p}}\rangle$ 的归一化方式。目前使用的归一化方案是不协变的，对应于 $|\mathbf{p}\rangle$ ，现在我们转而使用

$$|\lambda_{\mathbf{p}}\rangle = \sqrt{2E_{\mathbf{p}}} |\Lambda_{\mathbf{p}}\rangle,$$

于是就有

$$\begin{aligned}
& \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
&= \sum_{\lambda} \frac{1}{2E_{\mathbf{k}_1}(\lambda)} \frac{i}{\omega_1 - E_{\mathbf{k}_1}(\lambda) + i0^+} e^{i(\omega_1 - E_{\mathbf{p}})T_+} \langle \Omega | \phi(0) | \lambda_{\mathbf{k}_1} \rangle \langle \lambda_{\mathbf{k}_1} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
\end{aligned}$$

在接近(4.2)的极点时，就有

$$\begin{aligned}
& \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
&= \sum_{\lambda} \frac{i}{(\omega_1)^2 - (E_{\mathbf{k}_1}(\lambda))^2 + i0^+} \langle \Omega | \phi(0) | \lambda_{\mathbf{k}_1} \rangle \langle \lambda_{\mathbf{k}_1} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
\end{aligned}$$

类似地可以得到

$$\begin{aligned}
& \int_{-\infty}^{T_-} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | T[\phi(x_2) \cdots \phi(x_n)] \phi(x_1) | \Omega \rangle \\
&= - \sum_{\lambda} \frac{1}{2E_{\mathbf{k}_1}(\lambda)} \frac{i}{\omega_1 + E_{\mathbf{k}_1}(\lambda) + i0^+} e^{i(\omega_1 + E_{\mathbf{p}})T_-} \langle \Omega | T[\phi(x_2) \cdots \phi(x_n)] | \lambda_{\mathbf{k}_1} \rangle \langle \lambda_{\mathbf{k}_1} | \phi(0) | \Omega \rangle,
\end{aligned}$$

其中由于 $\phi(x_1)$ 的位置发生了变化，一些量的正负号和左右矢的顺序发生了变化。不过，这一项并不产生极点。因此最终我们得到

$$\begin{aligned}
& \int_{T_+}^{\infty} dt_1 e^{i\omega_1 t_1} \int d^3 \mathbf{x}_1 e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\
& \stackrel{\omega_1 \rightarrow E_{\mathbf{k}_1}(\lambda)}{\sim} \sum_{\lambda} \frac{i}{(\omega_1)^2 - (E_{\mathbf{k}_1}(\lambda))^2 + i0^+} \langle \Omega | \phi(0) | \lambda_{\mathbf{k}_1} \rangle \langle \lambda_{\mathbf{k}_1} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.
\end{aligned}$$

此外，设 U 是一个让三维动量减小 \mathbf{k}_1 的洛伦兹变换，则

$$\begin{aligned}
\langle \Omega | \phi(0) | \lambda_{\mathbf{k}_1} \rangle &= \langle \Omega | \phi(0) U^{-1} U | \lambda_{\mathbf{k}_1} \rangle \\
&= \langle \Omega | U \phi(0) U^{-1} U | \lambda_{\mathbf{k}_1} \rangle \\
&= \langle \Omega | \phi(0) | \lambda_0 \rangle,
\end{aligned}$$

第二个等号是因为真空态在洛伦兹变换下显然是不变的，第三个等号用到了 $U\phi(0)U^{-1} = \phi(0)$ 这一事实（将 $\phi(0)$ 展开为傅里叶级数就能看出为什么）。于是就得到

$$\int d^4x_1 e^{ik_1 \cdot x_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\ \stackrel{\omega_1 \rightarrow E_{k_1}(\lambda)}{\sim} \sum_{\lambda} \frac{i}{(\omega_1)^2 - (E_{k_1}(\lambda))^2 + i0^+} \langle \Omega | \phi(0) | \lambda_0 \rangle \langle \lambda_{k_1} | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.$$

上面的结果的意义非常明显了： x_1 变量做了傅里叶变换的关联函数会有一系列极点，这些极点的具体位置由在带相互作用的场论下的本征态的能谱（而不是自由粒子的能谱）决定；并且，会多出来一个因子 $\langle \Omega | \phi(0) | \lambda_0 \rangle$ 。由于 $|\lambda_{k_1}\rangle$ 是经过相互作用修正的单粒子态，实际上只需要一个动量参数就足够标记它，于是我们去掉对 λ 的求和（因为显然只有一个可能的 λ ），并且用 ω 代替 E （再次提醒：这是已经经过相互作用修正的单粒子能量），就得到

$$\int d^4x_1 e^{ik_1 \cdot x_1} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\ \stackrel{\omega_1 \rightarrow \omega_{k_1}}{\sim} \frac{i}{\omega_1^2 - \omega_{k_1}^2 + i0^+} \langle \Omega | \phi(0) | p=0 \rangle \langle k_1 | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle. \quad (4.3)$$

这里我们已经用 $|k_1\rangle$ 代替了 $|\lambda_{k_1}\rangle$ ，前者就表示无穷远处的单粒子态。

现在设想我们对 $\langle \Omega | T\phi(x_1)\phi(x_n) | \Omega \rangle$ 中的每一个位置变量都做傅里叶变换。由(4.3)，会发现每个 k_i 实际上都有极点。我们实际上可以在一个公式内把所有这些极点都反映出来。考虑对(4.3)中的 x_2 做傅里叶变换，由于 x_2 仅仅包含在最后一个因子中，只需要计算

$$\int d^4x_2 e^{ik_2 \cdot x_2} \langle k_1 | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\ = \left(\int_{-\infty}^{T_-} + \int_{T_-}^{T_+} + \int_{T_+}^{\infty} \right) dt_2 e^{i\omega_2 t_2} \int d^3\mathbf{x}_2 e^{-i\mathbf{k}_2 \cdot \mathbf{x}_2} \langle k_1 | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle.$$

仿照我们先前做的操作，只需要计算 T_+ 到 ∞ 的积分即可得到最为奇异的部分，于是可以将 $\phi(x_2)$ 提出编时算符的作用域内，放在最左边，然后仿照前面的操作，在 $\phi(x_2)$ 和编时算符序列之间插入完备性关系，得到

$$\int_{T_+}^{\infty} dt_2 e^{i\omega_2 t_2} \int d^3\mathbf{x}_2 e^{-i\mathbf{k}_2 \cdot \mathbf{x}_2} \langle k_1 | \phi(x_2) T[\phi(x_3) \cdots \phi(x_n)] | \Omega \rangle \\ = \int \frac{d^3\mathbf{p}_1}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}_1}} \int \frac{d^3\mathbf{p}_2}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}_2}} \int_{T_+}^{\infty} dt_2 e^{i\omega_2 t_2} \int d^3\mathbf{x}_2 e^{-i\mathbf{k}_2 \cdot \mathbf{x}_2} \\ \times \langle k_1 | \phi(x_2) | p_1, p_2 \rangle \langle p_1, p_2 | T[\phi(x_3) \cdots \phi(x_n)] | \Omega \rangle \\ = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} \int_{T_+}^{\infty} dt_2 e^{i\omega_2 t_2} \int d^3\mathbf{x}_2 e^{-i\mathbf{k}_2 \cdot \mathbf{x}_2} \langle \Omega | \phi(x_2) | p \rangle \langle k_1, p | T[\phi(x_3) \cdots \phi(x_n)] | \Omega \rangle$$

这一次只有 $\langle k_1 | \phi(x_2) | p_1, p_2 \rangle$ 有非零值（注意 $|p_1, p_2\rangle$ 是入射/反射态，可以用经过场重整化的产生算符作用两次产生），于是我们引入了两个动量积分；第二个等号还是因为 $|p_1, p_2\rangle$ 可以拆分，从而其中一个必须和 k_1 相等。对因子 $\langle \Omega | \phi(x_2) | p \rangle$ 施加如前所述的插入四维平移算符的操作，得到

$$\langle \Omega | \phi(x_2) | p \rangle = \langle \Omega | \phi(0) | p=0 \rangle e^{-ip \cdot x_2} |_{p^0=E_{\mathbf{p}}},$$

然后积掉 t_2 和 \mathbf{x}_2 ，就得到

$$\int d^4x_1 e^{ik_1 \cdot x_1} \int d^4x_2 e^{ik_2 \cdot x_2} \langle \Omega | \phi(x_1) T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle \\ \stackrel{\omega_1 \rightarrow \omega_{k_1}, \omega_2 \rightarrow \omega_{k_2}}{\sim} \frac{i \langle \Omega | \phi(0) | p=0 \rangle}{\omega_1^2 - \omega_{k_1}^2 + i0^+} \frac{i \langle \Omega | \phi(0) | p=0 \rangle}{\omega_2^2 - \omega_{k_2}^2 + i0^+} \langle k_1, k_2 | T[\phi(x_2) \cdots \phi(x_n)] | \Omega \rangle. \quad (4.4)$$

如此重复——实际上，我们还可以反过来做以上步骤，此时为了让极点出现，傅里叶变换的 e 指数要加上一个减号，在本节处理的标量场中这无所谓，但是对有粒子-准粒子区别的场，代表入射粒子的场算符和代表出射粒子的场算符差一个负号，从而一些要正着做傅里叶变换一些要反着做傅里叶变换——就得到

$$\prod_{i=1}^m \int d^4 x_i e^{ip_i \cdot x_i} \prod_{j=1}^n \int d^4 y_j e^{-ik_j \cdot y_j} \langle \Omega | T[\phi(x_1) \cdots \phi(x_n) \phi(y_1) \cdots \phi(y_n)] | \Omega \rangle$$

$$\stackrel{p_i^0 \rightarrow \omega_{\mathbf{p}_i}, k_j^0 \rightarrow \omega_{\mathbf{k}_j}}{\sim} \prod_{i=1}^m \frac{i\sqrt{Z}}{\omega_i^2 - \omega_{\mathbf{p}_i}^2 + i0^+} \prod_{j=1}^n \frac{i\sqrt{Z}}{\omega_j^2 - \omega_{\mathbf{k}_j}^2 + i0^+} \langle p_1, p_2, \dots, p_m | k_1, k_2, \dots, k_n \rangle, \quad (4.5)$$

其中

$$Z = |\langle \Omega | \phi(0) | p=0 \rangle|^2. \quad (4.6)$$

以上所有的推导都是在海森堡绘景下完成的，因此上式右边的因子就是 S 矩阵。于是我们就得到了联系 S 矩阵和关联函数的公式，其形式和之前的分析完全一样， Z 正是场强重整化因子。(4.5)就是标量场的**LSZ约化公式**，其形式和之前我们预期的完全一样。实际上，LSZ约化公式说明重整化后的格林函数是重整化前的 $1/Z^{(m+n)/2}$ 倍。

S 矩阵的微扰计算

在得到了(4.5)之后就可以微扰计算 S 矩阵了，因为可以微扰计算关联函数。实际上，计算关联函数比计算 S 矩阵更加困难。这本质上是因为 S 矩阵丢弃了初末态均为有限时间的信息（ S 矩阵仅仅保留了动量空间关联函数那些四维动量均在壳的那部分初末态），并且在费曼图的语言下有非常显然的解释。下面我们来分析关联函数的费曼图，并给出直接微扰计算 S 矩阵的方法。

计算关联函数的任何一张费曼图都具有这样的形式：一个amputated diagram居于中间，外线和它之间连接有自能修正图。在动量空间下，一张自能修正图可以等价地看成

$$\int d^4 x e^{ip \cdot x} \int d^4 y e^{-ik \cdot y} \langle \Omega | \phi(x) \phi(y) | \Omega \rangle,$$

按照(4.6)， \sqrt{Z} 对应 $\langle \Omega | \phi(x_1) a_{\text{in}}^\dagger | \Omega \rangle$ ，这个关联函数中的两个场中，一个做了场强重整化而另一个没有。既然 $\langle \Omega | a_{\text{out}} a_{\text{in}}^\dagger | \Omega \rangle$ 就是1， $\langle \Omega | \phi(x_1) \phi(x_2) | \Omega \rangle$ 的场强重整化因子为 Z 。

关联函数 $\langle \Omega | \phi(x_1) \phi(x_2) | \Omega \rangle$ 可以直接微扰计算，从而在理论已经给定的情况下，我们可以把 Z 到底是什么写出来。设 $-iM^2(p^2)$ 是单粒子不可约图（对应自能修正），那么就有

$$\begin{aligned} & \int d^4 x e^{ip \cdot x} \int d^4 y e^{-ik \cdot y} \langle \Omega | \phi(x_1) \phi(x_2) | \Omega \rangle \\ &= \frac{i}{p^2 - m_0^2} + \frac{i}{p^2 - m_0^2} (-iM^2(p^2)) \frac{i}{p^2 - m_0^2} + \cdots \\ &= \frac{i}{p^2 - m_0^2 - M^2(p^2)}, \end{aligned}$$

其中 m_0 为粒子裸质量。自能 M^2 关于 p^2 的最低阶项当然是 p^2 自己，即

$$M^2 = c_0 + c_1 p^2 + \cdots$$

在低阶近似下 $c_1 = 1$ ，高阶下则会有可见的修正，实际上这就对应着对 $(\partial_\mu \phi)^2$ 的修正。因此，在极点附近，我们有

$$\int d^4 x e^{ip \cdot x} \int d^4 y e^{-ik \cdot y} \langle \Omega | \phi(x_1) \phi(x_2) | \Omega \rangle \stackrel{p^0 \rightarrow E_{\mathbf{p}}}{\sim} \frac{iZ}{p^2 - m^2} + \text{regular}. \quad (4.7)$$

因此这给出了计算 Z 和有效质量 m 的方法：微扰计算自能修正，然后按照上式化简，观察极点位置就得到了 m ，在极点附近比较关联函数和 $i/(p^2 - m^2)$ 就得到了 Z 。将上式代入(4.5)，就发现

$$\langle p_1, p_2, \dots, p_m | k_1, k_2, \dots, k_n \rangle_{p_i^0 \rightarrow \omega_{p_i}, k_j^0 \rightarrow \omega_{k_j}} (\sqrt{Z})^{m+n} \times \text{amputated diagrams}. \quad (4.8)$$

因此，计算自能修正并得到 Z 之后，只需要计算amputated diagrams就能够微扰计算得到 S 矩阵。

我们还可以更加简化一些。注意到，所有只含有单粒子自能修正的图——即有 n 个入射粒子， n 个出射粒子，总共有 n 个连通子图，不同粒子对应不同连通子图的图——在amputate之后就是简单地将入射端和出射端连接起来，而这些图正好对应 $S = 1 + iT$ 中的1，因此如果只计算 \mathcal{M} ，无需计算这些图。

如果只计算树图，那么显然 $Z = 1$ ，因为此时没有任何自能修正，从而无论是物理质量还是场强都没有做重整化。因此，计算树图时只将amputated tree diagram求和即可。圈图的计算一般要用到重整化，此时为了方便看出场强重整化因子，并不会使用(4.8)。但是，在圈图计算中，场强重整化的那些 \sqrt{Z} 通常是作为抵消项被引入了，在施加重整化条件之后，它们和圈图计算中的发散抵消了，从而其实也无需显式计算 Z 。总之我们实际上不会直接使用LSZ约化公式计算散射振幅：如果只计算树图，那么 $Z = 1$ ；而如果计算圈图，那么 Z 是作为抵消项被考虑进去的，圈图计算的最终结果是对树图中的传播子的物理参数（基本上是质量）和顶角函数做了一定修正，因此 Z 同样不会出现在最终的计算结果中。

4.2 正规化和重整化

如果朴素地做圈图计算，通常会得到紫外发散。²这并不特别令人意外，因为没有什么保证了我们的理论在任意能标下都一定成立，从而，将动量积分的上限推到无穷大大约是不合适的。要想计算出有意义的结果通常要求我们知道更高能标处的物理的细节；然而，如果我们的理论实际上具有低能标下的一个不动点，那么更高能标处的物理实际上无关紧要，只要我们只关心低能的现象。

在这种情况下，可以做下面的操作来消除发散：

1. 在圈图计算中寻找一个能够标记重整化群流的参数，可能是和Wilson重整化群流一致的动量积分上限 Λ ，也可以将维数延拓到实数中，从而用维数偏离4的程度 ϵ 做这个参数。此时积分的发散部分可以被分离出来，这就是正规化。
2. 令理论中的各个参数（所谓“裸”参数）跑动起来，包括场强，即引入抵消项，其大小暂时未知。
3. 计算若干个可以实际观察到的物理量（通常具有和裸参数类似的物理意义，从而它们可以称为物理参数，如自由场的关联函数的极点给出裸质量，而有相互作用的关联函数的极点给出物理质量），将它们写成裸参数、参数跑动（即抵消项中的参数）和重整化群流参数的形式。

² 相对论性量子场论中的红外发散一般是因为少考虑了一些过程，如由于光子无能隙，可以任意地产生和消灭，从而一个过程的概率分散在很多个有入射和出射“软光子”（能量很低并无可观测效应的光子）。当求和所有这些过程后，发散一般就消失了，并不具有特别的意义。

凝聚态场论中的情况正好相反：紫外发散是没有关系的，因为凝聚态系统中有最小的特征长度，如果出现了发散，可以引入一个物理可观测的紫外截断消除这个发散（如BCS理论中的 ω_D ）。红外发散反而不那么好处理。

4. 在低能不动点处，物理参数应当和重整化群流参数无关，因此可以反过来将参数跑动写成重整化群流参数和物理参数的函数，从而求解出所有抵消项。用于确定抵消项的条件即为重整化条件。如果理论可重整，此时所有其它可以实际观察到的物理量中的发散都会相互抵消，从而我们成功地将一些可以实际观察到的物理量写成了另一些可以实际观察到的物理量（即物理参数）的函数，即给出了实验预言。

在最终的计算结果中裸参数都没有出现：这是正确的，因为实际上并没有什么能够真的“观察”到裸参数——无法确定实际测到的物理参数有多少来自裸参数，多少来自相互作用修正。实际上，为了和发散抵消，裸参数一般都是反向发散的。

在实际的计算中，还有以下技巧：

- 由于圈图数目非常多，通常我们会以需要计算的散射振幅的树图当成骨架，将圈图当成对骨架图中各个成分的修正。这样的好处是，做完全部修正的骨架图的各个成分通常足够给出物理参数了，如做完自能修正的传播子可以给出物理质量，做完顶角修正的amputated vertex diagram可以给出有效相互作用强度。
- 由于实际有意义的关联函数、散射振幅等均已做过场强重整化，可以用场强重整化之后的场来做拉氏量中的基本自由度；这样与场强重整化有关的抵消项会自动出现在拉氏量中，并且做完重整化之后直接计算amputated diagram即可得到散射振幅，没有必要再显式计算 Z 。
- 抵消项可以被显式给出。如果不希望引入太多顶角，也可以首先形式地写出含有未知的裸参数的散射振幅的形式，然后用未知的裸参数去拼凑出物理参数。如果显式地使用抵消项，由于我们在重整化不动点附近工作，其实可以将抵消项设置为“裸参数偏离物理参数的多少”，而直接将物理参数放进拉氏量中。

表面上，任何一个理论都可以做这样的操作——对称性允许的拉氏量中的项是无限多的，我们可以引入任意多的抵消项来消除发散。但是，如果需要引入无数多的抵消项，那么重整化操作就是无法完成的。

为了估计发散的度，我们可以引入一些指标。原始发散图指的是只要切断一根内线（即不计算这根内线的积分），就能够收敛的图。发散的图是用原始发散图组装起来的。如果一个理论中的原始发散图的个数有限，

由于紫外发散来自动量积分有太多重，一张图 Γ 中的动量的幂次——所谓表观发散度——为

$$D(\Gamma) = \sum_i n_i d_i + 2I_B + 3I_F - 4\left(\sum_i n_i - 1\right), \quad (4.9)$$

其中 n_i 指的是某一类型的顶角的个数， d_i 是类型 i 的顶角中的动量幂次， I_B 和 I_F 分别表示玻色子和费米子线的个数，因为费米子传播子的分母中只有一个 k ，做完四维动量积分之后动量幂次为3，而玻色子传播子的分母中有两个 k ，做完四维动量积分之后动量幂次为2。最后一项是因为顶角会引入一个动量守恒条件；我们故意减去了1，因为费曼图最终的计算结果也肯定满足动量守恒条件，即有一个 $\delta(\sum \mathbf{k})$ 并没有被积分掉，而是留到了计算结果中。现在我们进一步设 i 类型顶角中有 b_i 个玻色子线， f_i 个费米子线，并设有 E_F 条费米子外线， E_B 条玻色子外线，则

$$E_F + 2I_F = \sum_i n_i f_i, \quad E_B + 2I_B = \sum_i n_i b_i,$$

因为一条内线连接两个顶角。这样就有

$$D(\Gamma) = \sum_i n_i \left(d_i + b_i + \frac{3}{2} f_i - 4 \right) + 4 - E_B - \frac{3}{2} E_F. \quad (4.10)$$

表观发散度实际上就是在做量纲分析，而且做的是朴素的工程量纲分析，因此是不尽然可靠的。大体上说，如果表观发散度大于零，那么这张图发散，如果小于零，那么这张图收敛，如果等于零，那么这张图应该对数发散，但是这只是一个非常粗略，可能不准确的估计。不过，**Weinberg power counting theorem**保证了，当且仅当一张图及其子图的表观发散度都是小于零，它收敛。

为了保证尽可能多的图的表观发散度小于等于零，我们会要求

$$d_i + b_i + \frac{3}{2} f_i - 4 \leq 0, \quad (4.11)$$

因此比较安全的安排是，一个顶角最多有四条玻色子线、两条费米子线，否则有可能产生无穷多种发散的图，理论可能不能重整化。

4.2.1 维数正规化

注：经常用Wick转动来化简此处的积分，但是由于 $p^0 > 0$ 时极点在下半平面而 $p^0 < 0$ 时极点在上半平面，为了避免撞上奇点，Wick转动应该将积分路径顺时针旋转 90° ，从而可以设 $l = i l^E$ 。TODO：这和我们后面做的将整个理论做Wick转动时用的记号似乎不一样？

将内线动量从4维扩充为 n 维度，外线动量保持不变。

让维数变化时，没有必要让 γ 矩阵的维数发生变化，因此对 γ 矩阵的乘积的迹计算无需做任何调整，即自旋指标不需要做任何调整，而对坐标指标（ μ 这种）的迹计算（如 $\gamma^\mu \gamma_\mu$ ）则需要调整。

4.2.2 骨架图的修正

重整化条件：

- 在修正后的单粒子格林函数的极点处，有质量粒子的自能修正对 p^2 （玻色子）或者 p （费米子）的一阶导数为零；无质量粒子的自能修正为零。这是为了确保没有场强重整化。动量远离单粒子格林函数极点时它们当然可以不是零；本应如此，否则圈图修正无法体现。
- 在修正后的单粒子格林函数的极点处，极点给出的质量（通过 $p^2 = m^2$ 解出）就是我们设定的物理质量；在显式引入自能修正时，有质量粒子的自能修正为零。这是为了确保质量的修正为零。
- 顶角函数和物理相互作用强度相同，这是为了确保顶角修正为零。具体什么是“物理相互作用强度”取决于探测方式，如量子电动力学中通常是使用静电学方法测定电磁相互作用的强度，于是我们要求顶角函数在光子动量为零时和静电学方法测得的电磁相互作用强度（其实就是元电荷）相同。其它时候顶角函数可以有偏离，以展现高阶过程的修正。

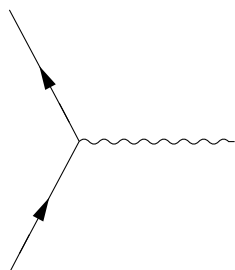
第二部分

规范场论

在单粒子量子力学和前面作为例子计算过的一些场论中，哈密顿量或是拉格朗日量中出现了些动力学变量，我们从这些动力学变量中挑选出一些来，它们不多也不少地可以标记希尔伯特空间的一组基矢量。略微推广一下，我们其实可以研究这样的理论：从其中的彼此对易的动力学变量中挑选出一组，它们标记的基矢量张成的空间要大于我们要研究的希尔伯特空间。这允许在理论中引入更加丰富的行为，而与此同时保持物理的自由度数正确。一种获得这样的理论的构造方式是通过所谓的“规范对称性”。物理学中的对称性通常包括时空对称性（即将物理事件的时空坐标做一个变换，一般来说，是洛伦兹变换）和内部对称性（即某个参数空间中的变换，通常是各点上场的变换）。规范对称性指的则是变换参数依赖局域时空坐标的对称性，即与定域的变换相关的对称性；通常我们在每一个空间点都放置某个群 G 的副本，即在每个空间点都放置一个 G 的群元 $g(x)$ ，我们要求理论在任何一个 $g(x)$ 场的变换（即所谓规范变换）下都保持不变，或者说，它具有局域的 G 对称性。选取规范变换作用在希尔伯特空间上的轨道的一个截面就称为选取一个规范：选取了一个规范后的一个波函数做时间演化绝对不会演化到另一个规范下的某个波函数中。这暗示我们，规范对称性实际上并不是真正的对称性：到头来，我们只不过是（不必要地）让希尔伯特空间扩张了若干倍的代价下，让理论可以具有的形式更加多样而已：一个规范理论总是可以使用一个非规范理论重写，但是如果前者是局域的，那么后者一般不是局域的。这和普通的场被引入的动机类似：不入场，就会出现粒子间的超距作用，则必须适当地说明粒子间的哪种超距作用是允许的。实际上我们总是在已经选取了一个规范以后的希尔伯特空间中工作：如果算符 U 是一个真正的对称性操作，那么 $|\psi\rangle$ 和 $U|\psi\rangle$ 是不同的波函数，而如果 U 是一个规范对称性，那么如果 $|\psi\rangle$ 在希尔伯特空间中， $U|\psi\rangle$ 就不在希尔伯特空间中；或者，我们可以要求 $|\psi\rangle$ 和 $U|\psi\rangle$ 认同。

这里其实有一个含混之处：如果一个理论具有规范结构，那么它肯定具有表面上的局域对称性；但是如果一个理论具有局域对称性，它是否只能是一个规范理论？Elitzur定理 [1]表明，一个理论的局域对称性无法被破缺：一个普通的对称性总是可以通过改变拉氏量而被自发破缺，但是局域对称性没有这种现象。

显然规范理论的方法是有好处的：它能够构造出原本我们想象不到的理论，还能够让我们批量构造这种理论。我们可以强迫一个自由理论具有规范对称性，或者说给它赋予一个规范结构，然后引入适当的规范场。这样的做法是非常套路性的，以至于我们可以发明出若干套方案，对每套方案，给定一个规范群就能够自动产生一个规范场论。此外，如果我们总是通过最小耦合将物质和规范玻色子耦合在一起，那么规范玻色场的运动方程的外源总是物质场的守恒流。这和我们在电动力学中的直观——物质场的“密度”驱动了规范玻色场——是非常一致的。我们可以将费米型的玻色场看成经典电动力学中的物质密度场开平方根得到的，做了这个开平方操作后费曼图能够直观地反映“物质粒子吸收或发射一个规范玻色子之后进入另一个模式”：



只有规范玻色子和密度场的理论是没有这种图的；这正是经典电动力学中的情况，在那里我们只能讨论电磁场如何激发出密度模式，密度模式再如何激发出电磁场。出于某些原因，

大自然为基本粒子赋予的物理是非常节俭的。我们将看到，目前已知的除了引力以外的所有相互作用都是通过一套这样的方案——杨-米尔斯理论——产生的。这一事实——即所谓规范原理——是量子场论历史上所谓“改变人心的转换”，它被系统应用之前，各个场的相互作用基本上只能唯象确定，它被系统应用之后，只需要写出规范群（即局域对称性的对称群）就能够确定相互作用。

本文将首先介绍电动力学，分析其性质，然后通过考虑其自然推广而得到杨-米尔斯理论。

第五章 量子电动力学

5.1 电动力学的拉氏量

5.1.1 规范场和狄拉克旋量场的最小耦合

在相对论性量子场论中——也即，在以闵可夫斯基时空为底流形的量子场论中，我们尝试将一个自旋1/2的狄拉克旋量场和一个无质量矢量场耦合起来。旋量场的拉氏量为

$$\mathcal{L}_{\text{spin}} = -m\bar{\psi}\psi + i\bar{\psi}\gamma_{\mu}\partial^{\mu}\psi, \quad (5.1)$$

而矢量场的拉氏量为

$$\mathcal{L}_{\text{vec}} = -\frac{1}{2}(\partial^{\mu}A^{\nu}\partial_{\mu}A_{\nu} - \partial^{\mu}A^{\nu}\partial_{\nu}A_{\mu}). \quad (5.2)$$

很容易看出(5.1)具有全局 $U(1)$ 对称性：它在变换

$$\psi \longrightarrow \psi' = \psi e^{i\alpha}$$

下保持不变。同样，(5.2)具有场的全局平移不变性（自由无质量矢量场的规范对称性），它在变换

$$A^{\mu} \longrightarrow A'^{\mu} = A^{\mu} + a^{\mu}$$

下保持不变。这两个对称性都是全局的：如果 α 或 a^{μ} 依赖于坐标，由于导数的链式法则，会多出一些项。具体来说，我们有

$$\mathcal{L}_{\text{spin}} \longrightarrow \mathcal{L}'_{\text{spin}} = \mathcal{L}_{\text{spin}} - \bar{\psi}\gamma_{\mu}\psi\partial^{\mu}\alpha. \quad (5.3)$$

对于矢量场，在 a^{μ} 的形式任意的情况下， \mathcal{L}_{vec} 的变换无规律可循，但是如果我们用某个标量的梯度 $\partial^{\mu}a$ 代替 a^{μ} ，那么有

$$A^{\mu} \longrightarrow A'^{\mu} = A^{\mu} + \partial^{\mu}a, \quad \mathcal{L}_{\text{vec}} \longrightarrow \mathcal{L}'_{\text{vec}} = \mathcal{L}_{\text{vec}}.$$

也就是说，矢量场的场的平移对称性实际上可以稍加推广而仍然成立。

ψ 的变换的相位因子是一个标量； A^{μ} 的场的平移量也是一个标量的梯度。很容易想到的尝试是，我们是否可以将两个场耦合起来，并要求整个系统在局域变换（ e 是常数而 $a(\mathbf{x})$ 依赖于坐标）

$$\psi \longrightarrow \psi' = \psi e^{iea}, \quad A^{\mu} \longrightarrow A'^{\mu} = A^{\mu} + \partial^{\mu}a \quad (5.4)$$

下保持不变？¹自由矢量场部分肯定是不变的，那么就要适当设计相互作用项的形式，把 ψ 做

¹早期的物理学家会认为，一个变换应该是物理上可行的，因此它不应该是全局的，而只能是局域的（例如我们可以让 a 在很小的范围内才不为零）。但是实际上这种观点是错误的——使得我们想要从头写下一个拉氏量的原因实际上是我们希望从一个非常简洁的源头推导出麦克斯韦方程，但是后面会看到，在经典情况下描述了一切电磁现象的麦克斯韦方程本身在 $U(1)$ 规范变换下不变，这意味着(5.4)展示的对对称性实际上是一种冗余，即系统中存在非物理、可以略去的自由度。这些自由度不参与和实际观测值有关的任何相互作用，作用在它们上面的变换完全没有必要是局域的。我们要求系统的动力学在局域 $U(1)$ 变换下不变，归根到底还是满足实验观测结论的需要。

局域 $U(1)$ 变换之后拉氏量多出来的一项吸收掉。当然，如果相互作用项是 $-eA^\mu\bar{\psi}\gamma_\mu\psi$ ，那就正好，因为

$$-eA^\mu\bar{\psi}\gamma_\mu\psi - \bar{\psi}\psi\gamma_\mu\partial^\mu(ea) = -eA'^\mu\bar{\psi}'\gamma_\mu\psi'.$$

于是我们得出结论：拉氏量

$$\mathcal{L} = \underbrace{-m\bar{\psi}\psi + i\bar{\psi}\gamma_\mu\partial^\mu\psi}_{\mathcal{L}_{\text{spin}}} - \underbrace{\frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu)}_{\mathcal{L}_{\text{vec}}} - \underbrace{eA^\mu\bar{\psi}\gamma_\mu\psi}_{\text{interaction}} \quad (5.5)$$

具有局域 $U(1)$ 不变性。推导出(5.5)的方法就是最小耦合。²

于是，我们得出结论：一个自旋1/2狄拉克旋量场和一个无质量矢量场耦合，并要求理论具有局域 $U(1)$ 对称性，那么就会得到我们将会看到，这个理论实际上就是电动力学。

还可以引入一些记号来简化(5.5)。首先引入反对称张量电磁张量

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (5.6)$$

则自由矢量场拉氏量为

$$\mathcal{L}_{\text{vec}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

另一方面，相互作用项和含有旋量场的导数的项形式非常接近，因此可以定义协变导数

$$iD^\mu = i\partial^\mu - eA^\mu, \quad (5.7)$$

最后将(5.5)写成

$$\begin{aligned} \mathcal{L} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \end{aligned} \quad (5.8)$$

这里我们用斜杠记号表示 $\gamma_\mu A^\mu$ 。

5.1.2 运动方程和守恒量

从(5.5)马上可以使用欧拉-拉格朗日方程写出运动方程。对 ψ 我们有

$$i\partial_\mu\bar{\psi}\gamma^\mu + e\bar{\psi}\gamma_\mu A^\mu + m\bar{\psi} = 0,$$

对其取共轭，或者对 $\bar{\psi}$ 应用欧拉-拉格朗日方程，就得到

$$i\gamma^\mu\partial_\mu\psi - m\psi = e\gamma_\mu A^\mu\psi. \quad (5.9)$$

对 A^μ 应用欧拉-拉格朗日方程，则有

$$\partial_\mu F^{\mu\nu} = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = e\bar{\psi}\gamma^\nu\psi. \quad (5.10)$$

以上两个方程给出了(5.5)的运动方程。(如前所述， ψ 和 $\bar{\psi}$ 虽然是独立的场，但它们的运动方程并不独立，因为运动方程是一阶的)

现在我们分析局域 $U(1)$ 对称性带来的守恒量。在局域 $U(1)$ 变换下，我们有

$$\delta\psi = ie\psi\delta a, \quad \delta A^\mu = \partial^\mu\delta a,$$

²需要注意的是最小耦合实际上并不是唯一的能够让理论满足局域 $U(1)$ 对称性的方案。理论中有哪些规范场、相互作用的形式如何，归根到底都需要实验上的提示。例如，电磁理论中最小耦合适用是因为它能够导出麦克斯韦方程，而麦克斯韦方程是已经验证了的在经典情况下正确的电磁定律。

则守恒流为

$$\begin{aligned} J^\mu \delta a &= -e\bar{\psi}\gamma^\mu\psi \delta a + (-\partial^\mu A^\nu + \partial^\nu A^\mu)\partial_\nu \delta a \\ &= -e\bar{\psi}\gamma^\mu\psi \delta a - \partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu)\delta a, \end{aligned}$$

第二个等号实际上是忽略了一个边界项后得到的结果。³无论如何，这个守恒流的第二项是平凡的，因为它就是电磁张量的一个指标求散度之后的结果，它的散度当然是零。那么，我们就有以下守恒荷：

$$J^\mu = e\bar{\psi}\gamma^\mu\psi, \quad \partial_\mu J^\mu = 0. \quad (5.11)$$

回过头看，实际上这是**全局 $U(1)$ 对称性**的守恒荷——全局 $U(1)$ 对称性中 a 在时空上是均匀的，那么 $\partial_\mu \delta a$ 就是零，正好让含有 A 的那个平凡的项消失。实际上从(5.10)中我们也可以得到这个守恒流。由于电磁张量是反对称的，我们有：

$$0 = \partial_\mu \partial_\nu F^{\mu\nu} = \partial_\mu (e\bar{\psi}\gamma^\mu\psi).$$

这就导出了(5.11)。使用(5.11)可以将(5.10)写成

$$\partial_\mu F^{\mu\nu} = J^\nu. \quad (5.12)$$

我们最后评论一下以上守恒量和运动方程的意义。(5.12)实际上就是麦克斯韦方程的一部分。构成麦克斯韦方程另外一部分的是以下恒等式

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0, \quad (5.13)$$

它是 $F_{\mu\nu}$ 定义为 A^μ 的梯度的反对称化导致的结果。 J^μ 给出了麦克斯韦方程中的 ρ 和 \mathbf{j} ，即它正是**电荷**的守恒流，电荷是全局 $U(1)$ 对称性对应的守恒荷。(5.11)就是 e 乘以粒子数密度，因此 e 就是 ψ 场激发的粒子携带的电荷量。对电子，它是 $e = -|e|$ ， $|e|$ 是元电荷。

5.1.3 规范

电动力学在局域 $U(1)$ 变换下的对称性实际上是一个**规范对称性**，也就是说，做任意的局域 $U(1)$ 变换，不会有任何可以观察到的变化，也就是说(5.5)中实际上有多余的自由度。我们需要对 A 和 ψ 施加适当的约束，以确保满足这个约束的 A 和 ψ 取值可以覆盖所有物理上可能产生的状态，同时不含有任何冗余的自由度，也即要**选取一个规范**。既然局域 $U(1)$ 对称性是规范对称性，任何反映系统实际状态的物理量都应该在 $U(1)$ 规范变换下不变。

形式最漂亮的应该是**洛伦兹规范**，也就是

$$\partial_\mu A^\mu = 0, \quad (5.14)$$

在这个规范下(5.12)转化为

$$\square^2 A = J, \quad \square^2 = \partial_\mu \partial^\mu. \quad (5.15)$$

我们得到了一个四维波动方程，当然，这就是**电磁波**。一个很自然的问题是，洛伦兹规范是否不失一般性？是否存在一组 A^μ 不能够通过一个规范变换变换为一组满足洛伦兹规范的 A'^μ ？实际上洛伦兹规范确实是不失一般性的，因为波动方程的性质很良好，对一个给定的标量场 C ，总是可以找到一个标量场 a ，使得

$$\partial_\mu \partial^\mu a = C,$$

³考虑到 δa 在每一点都可以独立地变化， $\int A \delta a = \int B \delta a$ 意味着 $A = B$ 。

这样不论原本 A^μ 取什么值，只需要解出一个 a^μ 使得

$$\partial_\mu \partial^\mu a = \partial_\mu A^\mu,$$

然后做规范变换

$$A'^\mu = A^\mu - \partial^\mu a, \quad \psi' = \psi e^{-iea},$$

得到的 A'^μ 就是服从洛伦兹规范的——并且表示和 A^μ 完全一样的物理状态。因此洛伦兹规范确实是不失一般性的。

容易看出 $F^{\mu\nu}$ 是一个规范不变量。实际上，在选定了规范之后，可以从它恢复出 A^μ 。不失一般性地选择洛伦兹规范，则我们有

$$\partial_\mu F^{\mu\nu} = \partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = \partial_\mu \partial^\mu A^\nu,$$

由于波动方程的良好性质，我们就从上式反解出 A^μ 了。如果是别的规范，就按照它转换到洛伦兹规范的方式，从洛伦兹规范转换到原有规范即可。总之，原则上任何规范不变量都可以通过 $F^{\mu\nu}$ 求导、积分等得到。

5.2 量子化

本节我们讨论量子化之后的电动力学，即量子电动力学，或者简称QED。自由无质量矢量场本身体现出规范对称性，因此对它的量子化和QED的量子化是紧密相连的。实际上，这让一些量子场论教科书——如Peskin——甚至直接将矢量场的有关内容作为QED的一部分介绍。

5.2.1 正则量子化

5.2.2 路径积分量子化

5.2.3 可重整性

Ward-Takahashi恒等式

考虑仅仅对 ψ 和 $\bar{\psi}$ 做局域 $U(1)$ 变换而不对矢量场做变换的一个情况。此时根据(5.3)，有

$$\delta\psi = ie\alpha\psi, \quad \delta\bar{\psi} = -ie\alpha\bar{\psi}, \quad \delta\mathcal{L} = -e\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha. \quad (5.16)$$

这不是一个经典意义下的对称操作，因为作用量会发生变化，但是我们知道量子情况下只要保持路径积分测度不变，有时仍然能够得到类似于诺特定理的结果。此外如果没有适当的规范对称性，也无法写出如此漂亮的“拉氏量的变化就是 $\partial_\mu\alpha$ 乘以电流”的变换。因此，应当记住(5.16)实际上是规范对称性的产物。Ward-Takahashi恒等式是量子版本的诺特定理，而这里我们要导出它关于某些散射振幅的特定形式。对关联函数 $\langle\Omega|T\psi(x_1)\bar{\psi}(x_2)|\Omega\rangle$ 做以上变换，会发现

$$\begin{aligned} 0 = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A e^{iS} & \left(-i \left(\int d^4x \partial_\mu\alpha(x) \right) (j^\mu(x)\psi(x_1)\bar{\psi}(x_2)) \right. \\ & \left. + (ie\alpha(x_1)\psi(x_1))\bar{\psi}(x_2) + \psi(x_1)(-ie\alpha(x_2)\bar{\psi}(x_2)) \right). \end{aligned}$$

其中 j^μ 就是 $e\bar{\psi}\gamma^\mu\psi$ 。对第一项做分部积分，并且将后两项写成 δ 函数的形式，就得到

$$\partial_\mu \langle \Omega | T j^\mu(x) \psi(x_1) \bar{\psi}(x_2) | \Omega \rangle = -e\delta(x-x_1) \langle \Omega | T \psi(x_1) \bar{\psi}(x_2) | \Omega \rangle + e\delta(x-x_2) \langle \Omega | T \psi(x_1) \bar{\psi}(x_2) | \Omega \rangle. \quad (5.17)$$

这就是量子版本的电荷守恒方程。

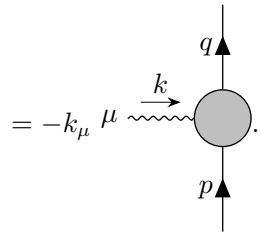
现在我们做傅里叶变换

$$\int d^4x e^{-ik \cdot x} \int d^4x_1 e^{iq \cdot x_1} \int d^4x_2 e^{-ip \cdot x_2},$$

等式右边没有太多可说，等式左边的 $\partial_\mu j^\mu(x)$ 变成了

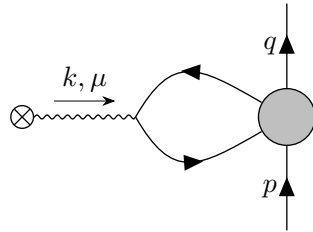
$$\int d^4x e^{-ik \cdot x} \partial_\mu j^\mu(x) = \int \frac{d^4k_1}{(2\pi)^4} i k_\mu e \bar{\psi}(k_1+k) \gamma^\mu \psi(k_1).$$

值得注意的是，组成 $j^\mu(x)$ 的两个费米子算符现在都具有不确定的动量，即它们对应内线或者说传播子。此外应当注意 $-ie\gamma^\mu$ 正是光子-电子相互作用顶角，因此上式在去掉 k_μ 之后乘上因子 ϵ_μ 就得到子图“一个光子入射，打出一对电子和正电子”。这样我们就有

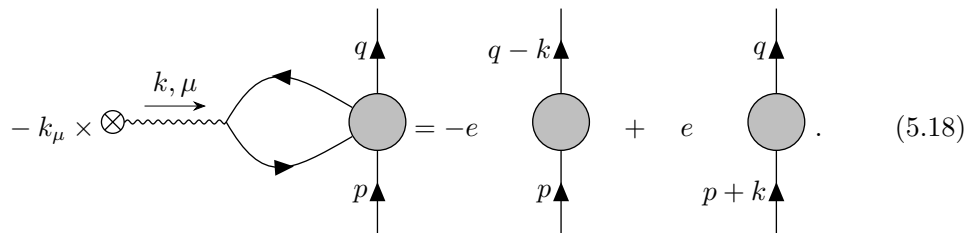
$$\int d^4x e^{-ik \cdot x} \int d^4x_1 e^{iq \cdot x_1} \int d^4x_2 e^{-ip \cdot x_2} \partial_\mu \langle \Omega | T j^\mu(x) \psi(x_1) \bar{\psi}(x_2) | \Omega \rangle$$


$$= -k_\mu \epsilon_\mu \text{ (diagram)}.$$

这里有一个微妙的地方：费曼图实际上就是微扰计算某种“格林矩阵”或是“跃迁矩阵”的矩阵元的图形，其外线就好像矩阵元的指标；通常，外线要么代表 S 矩阵的入射态和出射态（此时外线没有传播子），要么代表关联函数中出现的场（此时外线有传播子，并且通常会有自能修正）。然而，上式中的图形中的外线实际上不止一种：电子线是关联函数中的外线，有传播子，但光子线实际上是（给定偏振，从而不和 ϵ_μ 点乘的）外场引入的。很容易验证，上式左边没有光子传播子，那么右边当然也没有，因此光子线不是关联函数中的外线；但是这里的光子线同样不是 S 矩阵的外线，因为光子不必在壳（物理地说，外场意味着有一个非常重的系统在和我们讨论的系统交换光子，这些光子当然无需在壳）。实际上，下图



是更加合理的画法，光子线与 \otimes 相连代表它没有传播子这一事实，虽然一些教科书（如Peskin并没有采取这种画法）。所以最后我们得到

$$-k_\mu \times \otimes \text{ (diagram)} = -e \text{ (diagram)} + e \text{ (diagram)}. \quad (5.18)$$


这就是量子电动力学中通常所说的**Ward-Takahashi恒等式**。实际上，由一般的Ward-Takahashi恒等式（即“量子诺特定理”）可以看出，在变换(5.16)下，我们在上式两边的关联函数中添加更多的场（即加入更多入射和出射电子线），上式也是满足的。注意以上公式中从空间做傅里叶变换得到的动量空间的关联函数并不是在壳的，且(5.18)中的 e 实际上是裸参数 e_0 。

(5.18)看起来非常不自然，但应当注意到它的左边包括两个电子传播子而右边有一个电子传播子，从而在左边留下因子 Z_2 ；此外，它显然含有顶角函数，即含有 Z_1 。光子没有传播子，但是仍然有自能修正（因为项链图还是存在的；在算 S 矩阵矩阵元时无需考虑项链图，因为只需要算amputated的图，在算关联函数时项链图自动被纳入了光子传播子的自能修正；这里没有光子传播子，但是项链图还在）。但是，在 $k \rightarrow 0$ 时光子不再有自能修正，因为电子有质量，一个动量几乎为零的光子无法激发出正负电子对，因此可以忽略因子 Z_3 。因此，等式(5.18)意味着在量子电动力学中 Z_1 和 Z_2 之间有某种约束关系。在 $k \rightarrow 0$ 时，由重整化条件，顶角函数退化为“物理电荷”乘以 γ^μ 。因此(5.18)左边在 $k \rightarrow 0$ 时可以写成

$$= -k_\mu \frac{iZ_2}{\not{p} - m + i0^+} \frac{iZ_2}{\not{q} - m + i0^+} (-ie\gamma^\mu) (2\pi)^4 \delta^4(q - k - p), \quad \text{as } k \rightarrow 0.$$

$\sqrt{Z_3}$ 因子就是1， e 是重整化之后的电荷。另一方面，(5.18)的右边则是（请注意推导(5.18)时用的电荷就是裸电荷，这里记作 e_0 ）

$$= -e_0 (2\pi)^4 \delta(q - k - p) \left(\frac{iZ_2}{\not{p} - m + i0^+} - \frac{iZ_2}{\not{p} + \not{k} - m + i0^+} \right).$$

考虑到电荷的重整化为 $e_0 Z_2 Z_3^{1/2} = e Z_1$ 而 Z_3 在 $k \rightarrow 0$ 时为1，有

$$-k_\mu \frac{iZ_2}{\not{p} - m + i0^+} \frac{iZ_2}{\not{p} + \not{k} - m + i0^+} (-ie_0 Z_2 Z_1^{-1} \gamma^\mu) = -e_0 \left(\frac{iZ_2}{\not{p} - m + i0^+} - \frac{iZ_2}{\not{p} + \not{k} - m + i0^+} \right),$$

从而最终得到

$$Z_1 = Z_2. \quad (5.19)$$

这个结论是严格成立的；换句话说，计算无穷阶微扰之后，重整化因子 Z_1 和 Z_2 一定是一样的。

(5.18)还有另一个用处。(5.18)中的电子外线全部是关联函数中的外线，带有传播子，四维动量可以不在壳。然而，如果我们考虑四维动量在壳的那些情况，那么(5.18)右边的两个关联函数中均有四维动量是离壳的（例如容易验证，如果 q 和 k 在壳那么 $q - k$ 肯定是离壳的），因此这些关联函数对 S 矩阵没有贡献。因此，对那些所有入射和出射外线——无论是电子还是光子——都是 S 矩阵型外线的单光子图——其实就是单光子过程的 \mathcal{M} ——我们有

$$k_\mu \mathcal{M}^\mu = 0, \quad (5.20)$$

其中 $\mathcal{M} = \epsilon_\mu \mathcal{M}^\mu$ 。这称为**Ward恒等式**。这是Ward-Takahashi恒等式的在壳情况。

总之，规范对称性、所之而来的电荷守恒、Ward恒等式基本上具有同样的来源。

5.3 非相对论极限

5.3.1 电子，光子和电场

光子无法做非相对论近似，因为无论如何，麦克斯韦方程都应该成立，而这个方程就是洛伦兹协变的。需要做非相对论近似的只有电子。在非相对论近似下，一切有质量的场都退化为薛定谔场，电子也不例外。因此在QED的非相对论极限下，基本的粒子包括电子和光子，光子无任何变化，电子场则是动能为 $\mathbf{k}^2/2m$ ，不再满足相对论协变性，由动量和自旋标记的场。

电子的薛定谔-泡利方程

我们现在推导非相对论性电子场遵循的方程，然后得到非相对论极限下的电子哈密顿量。对称性告诉我们这个方程和狄拉克方程和电磁场耦合后取非相对论性近似的结果肯定是一样的。但是，应该注意，前者中的电磁场都是算符场，而后者中的电磁场是一个经典外场，即后者忽略了电磁场的量子涨落。电磁场的量子涨落会产生可观测的效应（即所谓**辐射修正**）。实际上，这是量子场论的理论框架适用于电动力学的实验证据：一个相对论性的关于电子和光的理论未必要采取量子场论的形式，将电磁场量子化之后，圈图修正就会预言一些经典的电磁场不会产生的现象。如果这些现象实际上没有观测到，那么量子场论的理论框架就是没有用的或者说错误的，但是实际上我们观测到了这些现象，那么量子场论很可能就是对的。

QED中，电子的狄拉克方程为

$$(i\gamma^\mu \partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0.$$

使用狄拉克表象以区分电子和正电子，设旋量场为

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix},$$

我们将要分析 ϕ 遵循的运动方程。为此，根据

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix},$$

写出狄拉克方程的分量表达式：

$$(i\partial_t - e\varphi - m)\phi + (e(\boldsymbol{\sigma} \cdot \mathbf{A}) + i(\boldsymbol{\sigma} \cdot \nabla))\chi = 0,$$

$$(i\partial_t - e\varphi + m)\chi + (e(\boldsymbol{\sigma} \cdot \mathbf{A}) + i(\boldsymbol{\sigma} \cdot \nabla))\phi = 0,$$

这里我们需要指出一个容易弄混的地方：正确的分量是

$$\partial_\mu = (\partial_t, \nabla), \quad A_\mu = (\varphi, -\mathbf{A}),$$

而贸然地很容易认为前者是 $(\partial_t, -\nabla)$ 而后者是 (φ, \mathbf{A}) 。我们消去 χ ，就得到

$$(i\partial_t - e\varphi - m)\phi - (e(\boldsymbol{\sigma} \cdot \mathbf{A}) + i(\boldsymbol{\sigma} \cdot \nabla)) \frac{1}{i\partial_t - e\varphi + m} (e(\boldsymbol{\sigma} \cdot \mathbf{A}) + i(\boldsymbol{\sigma} \cdot \nabla))\phi = 0. \quad (5.21)$$

到目前为止我们没有做任何近似。现在我们做最强的非相对论近似。我们按照通常的从克莱因-高登方程获得薛定谔方程的方法，设

$$\phi = \psi e^{-imt}.$$

这里我们重复使用了符号 ψ ，用它表示一个旋量场，以及从这个旋量场得到的一个薛定谔场；通常这不会导致混淆，因为前者只出现在高能物理中而后者只出现在凝聚态物理中。我们发现

$$(i\partial_t - e\varphi - m)\phi = e^{-imt}\partial_t\psi - e\varphi\psi e^{-imt},$$

而

$$(i\partial_t - e\varphi + m)\phi = e^{-imt}\partial_t\psi - e\varphi\psi e^{-imt} + 2me^{-imt}\psi.$$

在非相对论极限下粒子的动能相较于其静能（请注意这里都是自然单位制， $c = 1$ ， m 就是静能）是非常低的，于是

$$\partial_t\psi \ll 2m\psi,$$

因此我们有

$$i\partial_t\psi - e\varphi\psi - \boldsymbol{\sigma} \cdot (e\mathbf{A} + i\nabla) \frac{1}{2m - e\varphi} \boldsymbol{\sigma} \cdot (e\mathbf{A} + i\nabla)\psi = 0.$$

进一步，认为电场不很强（否则电子会轻易被加速到很高的速度，非相对论极限不正确），从而

$$e\varphi \ll m,$$

那么就得到

$$i\partial_t\psi - e\varphi\psi - \frac{(\boldsymbol{\sigma} \cdot (e\mathbf{A} + i\nabla))^2}{2m}\psi = 0. \quad (5.22)$$

做了如此的近似之后，我们相当于忽略了所有涉及正电子的辐射修正。这个方程可以化成更加清晰的一个形式。我们知道

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

而由于 $(e\mathbf{A} - i\nabla)$ 是算符，它自乘并不是零，而是

$$(e\mathbf{A} + i\nabla) \times (e\mathbf{A} + i\nabla)\psi = ie\mathbf{A} \times (\nabla\psi) + ie\nabla \times (\mathbf{A}\psi) = ie(\nabla \times \mathbf{A})\psi = ie\mathbf{B}\psi,$$

于是我们就有

$$i\partial_t\psi = \frac{(e\mathbf{A} + i\nabla)^2}{2m}\psi + e\varphi\psi - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m}\psi. \quad (5.23)$$

场论哈密顿量为

$$\begin{aligned} H &= \int d^3\mathbf{x} \psi^\dagger \left(\frac{(-i\nabla - e\mathbf{A})^2}{2m} + e\varphi\psi - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m} \right) \psi \\ &= \int d^3\mathbf{x} \psi^\dagger \left(-\frac{(\nabla - ie\mathbf{A})^2}{2m} + e\varphi\psi - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m} \right) \psi, \end{aligned} \quad (5.24)$$

其中

$$\int d^3\mathbf{x} \psi^\dagger \frac{(-i\nabla - e\mathbf{A})^2}{2m} \psi = \sum_{\sigma} \int \frac{d^3\mathbf{p}}{(2\pi)^3} a_{\mathbf{p}\sigma}^\dagger \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} a_{\mathbf{p}\sigma}. \quad (5.25)$$

ψ 是电子的湮灭算符，其运动方程就是薛定谔绘景下单电子波函数的运动方程。于是，坐标表象下单电子的波函数（当然，此时已经隐含地认为电磁场是经典场，否则会有等效电子-电子相互作用，单电子图像不再适用）满足的方程就是

$$i\partial_t\psi = \frac{(e\mathbf{A} + i\nabla)^2}{2m}\psi - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m}\psi + e\varphi\psi = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m}\psi - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m}\psi + e\varphi\psi. \quad (5.26)$$

这个方程的形式和自由电子薛定谔方程很像，但是哈密顿量为

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} - \frac{e\boldsymbol{\sigma} \cdot \mathbf{B}}{2m} + e\varphi = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} - \boldsymbol{\mu} \cdot \mathbf{B} + e\varphi. \quad (5.27)$$

我们称此方程为泡利方程或是薛定谔-泡利方程。这里的 \mathbf{p} 的意义实际上尚不明确，它的定义是“在坐标表象下是 $-i\nabla$ 的那个算符”。大体上它是“动量”，但是实际上它不是 $m\mathbf{v}$ 的那个机械动量。如果我们计算(5.27)作为经典哈密顿量给出的运动方程，会得到

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p} - e\mathbf{A}}{m}, \quad \frac{d\mathbf{p}}{dt} = -\nabla_{\mathbf{r}} H, \quad (5.28)$$

展开计算，并注意到

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{A},$$

最终可以计算得到

$$\begin{aligned} m \frac{d\mathbf{v}}{dt} &= \frac{d(\mathbf{p} - e\mathbf{A})}{dt} = -e\nabla\varphi - e \frac{\partial \mathbf{A}}{\partial t} + e\mathbf{v} \times (\nabla \times \mathbf{A}) + \nabla(\boldsymbol{\mu} \cdot \mathbf{B}) \\ &= e\mathbf{E} + e\mathbf{v} \times \mathbf{E} + \nabla(\boldsymbol{\mu} \cdot \mathbf{B}). \end{aligned} \quad (5.29)$$

因此这里的 \mathbf{p} 不是机械动量。

无论电磁场是经典场还是考虑量子涨落，(5.24)都是成立的。这些式子可以看成对(5.27)中的电子做二次量子化得到的；这些式子也意味着，如果我们要将相对论性的、量子化的电磁场和非相对论性的电子耦合起来，只需要量子化自由电磁场即可，因为电磁场和非相对论性的电子耦合的哈密顿量总是(5.24)，无论电磁场是经典的还是量子的。(5.24)也可以看成对非相对论性自由电子做 $U(1)$ 最小耦合得到的结果。

(5.24)中， \mathbf{A} 出现二阶项这件事值得分析，因为在QED中不存在这个情况：电磁场和电子旋量场之间的耦合中电磁场是一阶的。回顾以上求解过程，这个 \mathbf{A}^2 项实际上来自积掉 χ 场的过程：我们只希望保留电子模式 ϕ 而不想保留正电子模式 χ ，于是就有(5.21)，实际上就是积掉了 χ 场。分析 $i\bar{\psi}\gamma^\mu D_\mu\psi$ 项的形式可以知道，与 χ 场相关的过程包括： χ 模式吸收一个 φ 光子； χ 模式和 ϕ 模式互换，这个顶角的耦合常数含有动量的一次方； χ 模式和 ϕ 模式互换，吸收一个 \mathbf{A} 光子。因此，积掉 χ 场引入了这样一个等效过程： ϕ 模式或是吸收一个 \mathbf{A} 光子或是不吸收，变成 χ 模式，然后吸收若干个 φ 光子，最后或是吸收一个 \mathbf{A} 光子或是不吸收而重新变成 ϕ 模式。这个过程涉及 \mathbf{A} 的零次方、一次方和平方，以及任意阶次的 φ 。然而， φ 光子实际上是在给 χ 场提供自能修正：

$$\frac{1}{i\partial_t + m} + \frac{1}{i\partial_t + m} e\varphi \frac{1}{i\partial_t + m} + \cdots = \frac{1}{i\partial_t + m - e\varphi},$$

而在非相对论极限下 $e\varphi$ 相比于 m 肯定是很小的，从而可以忽略。另一方面， \mathbf{A} 的大小和 \mathbf{p} 的大小却不好比较：两者都不能太大。因此在(5.24)中出现了非线性的 \mathbf{A} 耦合，但是没有出现非线性的 φ 耦合。

应当指出这个非线性的 \mathbf{A} 耦合并不意味着以机械动量标记的电子与光子的耦合中，电磁场就是非线性的，因为(5.27)中的 \mathbf{p} 是正则动量而不是机械动量。(5.27)中的 \mathbf{p} 不是实际可以直接观察到的机械动量，并且存在 \mathbf{A}^2 项，这两者彼此抵消，正好导致了关于机械动量，并且电磁场耦合项完全线性的(5.29)。回顾(5.29)的推导，会发现其中 \mathbf{A}^2 项在运动方程中产生一些类似于 $\mathbf{A} \times (\nabla \times \mathbf{A})$ 的项，而 $\mathbf{p} \cdot \mathbf{A}$ 项产生了一些类似于 $\mathbf{p} \times (\nabla \times \mathbf{A})$ 的项，由于在(5.27)中 \mathbf{p} 是正则动量，这两项彼此抵消，正好产生了 $\mathbf{v} \times (\nabla \times \mathbf{A})$ 这样的项。可以使用与(5.27)等价的拉氏量

$$L = \frac{1}{2}m\mathbf{v}^2 - e\varphi + e\mathbf{v} \cdot \mathbf{A} + \boldsymbol{\mu} \cdot \mathbf{B} \quad (5.30)$$

来更加清楚地看到这一点。我们不可能只使用机械动量就得到完整的非相对论性粒子在电磁场中的运动，因为这相当于做了一个规范变换让 $\mathbf{A} = 0$ ，而这是不可能的——能够产生磁场的那部分 \mathbf{A} 无法转化为 φ 。

相对论修正

相对论修正指的是当电子的能量仍然不是特别高（从而没有必要使用完整的狄拉克方程）但是已经比较高时，相对论效应造成的修正。它和量子涨落造成的修正是基本上无关的，因为即使不考虑量子涨落，狄拉克方程仍然会导致偏离(5.27)的结果。

腔体中的光场

5.3.2 树图阶有效相互作用

原则上(5.24)和量子化后的电磁场给出了非相对论性电子和电磁辐射的全部理论，包括了全部电子和光的相互作用。然而，实际上我们还可以做进一步的近似，将一些过程用等效相互作用替代。做完这种近似之后需要特别小心，以避免将同样的过程重复计数。

一种在凝聚态物理中特别常见的做法是，将那些不满足横波条件的光子积掉，因为凝聚态物理中的光子基本上能够直接对应为能够长距离传播的经典电磁波，它能够长距离地“逃出”介质。能够通过大片真空的电磁波模式只能是满足横波条件的电磁波——不满足横波条件的电磁波模式只能出现在电荷附近。因此，不满足横波条件的光子应当被积掉，因为它们代表“介质内部发生的、无法直接观察的事情”，而满足横波条件的光子应该被保留，介质的光学性质就是在描述介质和它们的相互作用。这样不会有重复计数。

还有一些时候，不满足横波条件的光子实际上是我们讨论的那部分系统和一个非常“重”的系统（它基本上不会受到我们讨论的那部分系统的影响，可以看成背景）交换的，即是外场的一部分。

积掉光子的过程实际上就是计算所有外线都是电子的散射振幅的过程。我们下面要做的计算都是在高能场论中完成的；(5.27)只是关于电子旋量场的一部分分量的，从而，我们下面获得的等效相互作用在加入(5.27)时可能需要适当地作用投影算符。凝聚态物理能标下基本上只需要计算树图，本节也不计算圈图。

库伦相互作用

在非相对论情况下，光子的能量不足以激发出电子-正电子对，真空极化并不重要。因此，光子虽然是无能隙的，仍然可以积掉光子而得到电子-电子等效相互作用。我们将其它场看成背景场，因为QED中没有光子-光子相互作用顶角，这个等效相互作用背后的实际的QED过程只包含树图。

我们首先直接在QED中计算电子-电子等效相互作用，即使用旋量场 ψ 做计算，这相当于将所有的“电子发出一个光子，这个光子被另一个电子吸收”的子图用一个电子-电子等效相互作用替代，而不去碰“电子发出一个光子，这个光子被一个不是电子的东西吸收”，或是“一个不知道哪儿来的光子被电子吸收”的过程（这些过程可能又被做了别的近似，比如说如果系统中充满了“不知道哪儿来的光子”，那就可以把电磁场当成经典场，等等）。⁴这么做了之后，不计入电子-电子等效相互作用的那些过程仍然是电子与电磁场耦合

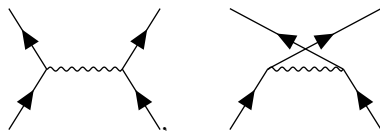
⁴ 这里说的“电子”指的都是我们的理论考虑的那部分电子。“无穷远处射来一束光，被电子吸收”这个过程显然在我们的理论中不是电子-电子相互作用，但是把视角调远一些，很可能这个“无穷远处射来的光”本身是另一些电子辐射产生的，例如

的过程，从而仍然可以得到(5.24)。于是最终我们就得到一个由(5.24)，一个电子-电子相互作用，和电磁场自身的哈密顿量三部分拼在一起的模型，其中(5.24)中的 φ 和 \mathbf{A} 由电子激发出来的部分不能作用在电子上（从而避免了重复计数）。这实际上就是凝聚态场论涉及电子和光子的部分（涉及声子的部分还未加入，因为此时尚无晶格）。

我们也可以换一个等价的思路，首先直接做非相对论极限得到(5.24)，其中的 \mathbf{A} 和 φ 由电子激发出来的部分可以作用在电子上。然后我们将所有的“电子发出一个光子，这个光子被另一个电子吸收”的子图用一个电子-电子等效相互作用替代，而“电子发出一个光子，这个光子被一个不是电子的东西吸收”，或是“一个不知道哪儿来的光子被电子吸收”的过程则不纳入等效相互作用，而是保留下来。

这两种思路会得到一样的理论，在其中一方面光子场仍然是实际存在的场而没有被真的积掉，一方面纯粹的电子-光子-电子相互作用（即光子只是内线）却又不用考虑，因为它们已经被等效的库仑相互作用考虑进去了。这看起来很不自然，但在很多情况下——例如，在凝聚态物理，包括量子光学中——光子可以分成截然不同的两部分，一部分是物质之间交换的、基本上导致一个库仑相互作用的光子，一部分是遥远处的光源产生的“入射光”；前者是有源无旋的，后者是无源有旋的。这是两种光子模式，彼此没有任何相互作用（能标足够低，不会有QED的四光子等效相互作用），并且前者从来不会出现在外线中。这种情况下使用等效的电子-电子库仑相互作用就是非常合理的：我们马上要做的计算表明电子之间由光子传递的相互作用的头阶项是 φ 或者说 A^0 传递的库仑相互作用，这已经将有源无旋的光子产生的相互作用完全囊括了；然后我们就可以假定，在我们的模型中显式出现的光子都是无源有旋的、和我们通常对“电磁波”的想象一致的那部分光子。当然我们也总是可以从头做计算，不引入等效的电子-电子库仑相互作用，计算每个电子受外场驱动而产生的辐射（包括近场和远场）。等效电子-电子库仑相互作用实际上就是这里的近场辐射的无推迟时间近似而已。

给出等效电子-电子相互作用的图有两张：



如果参与散射的两个粒子是可以分辨的（即除了动量和自旋以外还有别的标签可以区分它们），那么第二张图和第一张图不在一个相互作用通道中。低能过程由于动量低，相应的特征尺度很大，即粒子不会离得很近，这种情况下粒子的“位置”近似起到了区分两个粒子的标签的作用。这意味着第二张图可以忽略。

于是我们计算第一张图，它给出

$$\begin{array}{c} \text{Diagram: } p' \text{ and } k' \text{ incoming, } p \text{ and } k \text{ outgoing, connected by a wavy line.} \end{array} = (-ie)^2 \bar{u}(p') \gamma^\mu u(p) \frac{-i\eta_{\mu\nu}}{(p' - p)^2 + i0^+} \bar{u}(k') \gamma^\nu u(k).$$

我们考虑 p, p', k, k' 都几乎是零的情况，并且只对 $\mu = \nu = 0, 3$ 的情况求和——其实我们会看到， $\mu = \nu = 1, 2$ 两种情况并不会贡献。此时计算会发现

$$\bar{u}(p') \gamma^0 u(p) = u^\dagger(p') u(p) \approx m \begin{pmatrix} \xi^\dagger & \xi^\dagger \end{pmatrix} \begin{pmatrix} \xi \\ \xi \end{pmatrix} = 2m\sigma^0,$$

它可能来自原子能级跃迁产生的辐射，但是能级跃迁产生的辐射说到底还是电子产生的。然而，也许在我们讨论的问题的能标下，原子可以看成是一个整体，那么，这个过程在我们的低能有效理论中就不是电子-电子相互作用。

而

$$\bar{u}(p')\gamma^i u(p) = u^\dagger(p') \begin{pmatrix} -\sigma^i & \\ & \sigma^i \end{pmatrix} u(p) \approx m \begin{pmatrix} \xi^\dagger & \xi^\dagger \\ & \end{pmatrix} \begin{pmatrix} -\sigma^i & \\ & \sigma^i \end{pmatrix} \begin{pmatrix} \xi \\ \xi \end{pmatrix} = 0.$$

于是我们就得到

$$\begin{aligned} \text{Diagram} &= (-ie)^2 \bar{u}(p')\gamma^\mu u(p) \frac{-i\eta_{\mu\nu}}{(p' - p)^2 + i0^+} \bar{u}(k')\gamma^\nu u(k) \\ &= ie^2 2m(\sigma^0)_p \frac{1}{(p' - p)^2 + i0^+} 2m(\sigma^0)_k \\ &= -\frac{ie^2(2m)^2\sigma^0}{|\mathbf{p}' - \mathbf{p}|^2 - i0^+}. \end{aligned}$$

这里由于我们取了非相对论极限，有

$$p^\mu = (m, \mathbf{p}),$$

于是

$$(p' - p)^2 = -|\mathbf{p} - \mathbf{p}'|^2.$$

上式给出的是 S 矩阵的矩阵元，考虑到外线全部是电子，它是 $\{|p\rangle\}$ 表象下的等效相互作用哈密顿量矩阵元；下标 p 和 k 用于区分作用在不同单粒子态上的矩阵。矩阵 σ^0 给出了自旋的变化情况，可以看到以上相互作用通道不挑选入射自旋，也不改变入射自旋。我们要做非相对论近似，所以要转换到非洛伦兹协变的、在单粒子量子力学中使用的 $\{|\mathbf{p}\rangle\}$ 表象下，由于有四条外线，要除以因子 $(\sqrt{2m})^4$ 。于是非相对论极限下，我们获得相互作用顶角

$$\text{Diagram} = -i \frac{e^2}{|\mathbf{p} - \mathbf{p}'|^2} (2\pi)^4 \delta^4(k' + p' - p - k). \quad (5.31)$$

这个相互作用顶角的形式实际上正是动量空间中的库伦定律。为了更加清晰地看出库伦定律，我们将上式切换回实空间，做傅里叶变换

$$\int \frac{d^4 p'}{(2\pi)^4} e^{ip' \cdot x_1} \int \frac{d^4 k'}{(2\pi)^4} e^{ik' \cdot x_2} \int \frac{d^4 p}{(2\pi)^4} e^{-ip \cdot x_3} \int \frac{d^4 p}{(2\pi)^4} e^{-ip \cdot x_4},$$

计算发现

$$\begin{aligned} \text{Diagram} &= -ie^2 \delta(t_4 - t_1) \delta^4(x_1 - x_3) \delta^4(x_2 - x_4) \int \frac{d^3 \mathbf{q}}{(2\pi)^3} \frac{e^{-i\mathbf{q} \cdot (\mathbf{x}_4 - \mathbf{x}_1)}}{|\mathbf{q}|^2 - i0^+} \\ &= -i\delta(t_4 - t_1) \delta^4(x_1 - x_3) \delta^4(x_2 - x_4) \frac{e^2}{4\pi|\mathbf{x}_4 - \mathbf{x}_1|}. \end{aligned} \quad (5.32)$$

因此我们的确得到了库伦相互作用。在计算时有一个细节：计算(5.32)的第一个等号右边的积分时，我们有

$$\begin{aligned} \int \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot(\mathbf{x}_4-\mathbf{x}_1)}}{|\mathbf{q}|^2 - i0^+} &= \frac{2\pi}{(2\pi)^3} \int_0^\pi \sin\theta d\theta \int_0^\infty q^2 dq \frac{e^{-iq|\mathbf{x}_4-\mathbf{x}_1|\cos\theta}}{|\mathbf{q}|^2 - i0^+} \\ &= \frac{1}{4\pi^2} \int_0^\infty \frac{q^2}{q^2 - i0^+} dq \frac{e^{-iq|\mathbf{x}_4-\mathbf{x}_1|} - e^{iq|\mathbf{x}_4-\mathbf{x}_1|}}{-iq|\mathbf{x}_4-\mathbf{x}_1|} \\ &= \frac{1}{4\pi^2 i} \int_{-\infty}^\infty \frac{q^2}{q^2 - i0^+} dq \frac{e^{iq|\mathbf{x}_4-\mathbf{x}_1|}}{q|\mathbf{x}_4-\mathbf{x}_1|}. \end{aligned}$$

如果我们将因子 $q^2/(q^2 - i0^+)$ 直接当成1，上式就没有确定的值了，因为极点直接出现在了积分路径上。不过，我们有

$$\frac{q^2}{q^2 - i\epsilon} = \frac{1}{1 - \frac{i\epsilon}{q^2}} \rightarrow 0 \quad \text{as } q \rightarrow 0,$$

因此我们应该取积分主值，即取

$$\int \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{e^{-i\mathbf{q}\cdot(\mathbf{x}_4-\mathbf{x}_1)}}{|\mathbf{q}|^2 - i0^+} = \frac{1}{4\pi^2 i} \text{P} \int_{-\infty}^\infty dq \frac{e^{iq|\mathbf{x}_4-\mathbf{x}_1|}}{q|\mathbf{x}_4-\mathbf{x}_1|} = \frac{1}{4\pi^2 i} \frac{\pi i}{|\mathbf{x}_4-\mathbf{x}_1|} = \frac{1}{4\pi|\mathbf{x}_4-\mathbf{x}_1|}.$$

单粒子量子力学中的散射理论相当于梯形图近似。

从上面的计算可以看出，光子传播子和电子-电子等效相互作用势能是等效的；之后考虑真空极化时，我们可以认为库伦势能被修正了——这正是“真空极化”的说的来源，因为好像真空自己作为一种介质而导致了纯库伦相互作用的修正一样。

自旋磁矩

在(5.27)和(5.24)中我们看到，由于算符不对易这一特点，哈密顿量中除了经典的机械动能项（正则动量 \mathbf{p} 减去电磁动量 $e\mathbf{A}$ 得到机械动量）以外还多出来一项。电子自旋算符为

$$\mathbf{S} = \frac{\boldsymbol{\sigma}}{2},$$

于是这一项就是

$$H_{\text{spin}} = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (5.33)$$

其中

$$\boldsymbol{\mu} = \frac{e\boldsymbol{\sigma}}{2m} = \frac{e}{m} \mathbf{S} = -\frac{|e|\hbar}{m} \mathbf{S} = -g \frac{|e|\hbar}{2m} \mathbf{S}. \quad (5.34)$$

这意味着电子即使在静止时也有磁矩，这个磁矩是来自其自旋而不是轨道运动的。 $g = 2$ 是自旋磁矩的朗德 g 因子。

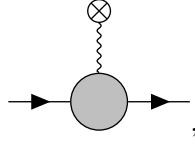
上面的推导实际上是将 \mathbf{B} 当成了一个完全确定的外场；然而，电磁场存在量子涨落。电子磁矩是实验可测的。粗略的实验确实指出 $g = 2$ ，更加精确的实验（基本上远远超过了凝聚态场论的能标）则显示 g 其实比2稍微大一些，并且计算QED圈图修正之后能够得到和实验测量结果非常接近的值，这是一个证明QED可靠性的有力论据。

5.4 低能一圈图辐射修正

并非所有辐射修正都需要完整地做重整化。可以设想，对那些入射动量真的能够保证接近于零的过程（如那些非相对论近似下就能出现的过程），应该想办法不做完整的重整化就得到有限的结果。

5.4.1 顶角函数和反常磁矩

QED会导致的一个能够观测的效应是**反常磁矩**，即自旋磁矩实际上不严格是2。我们此处先给出计算自旋磁矩的大致步骤。我们是要计算如下的图：

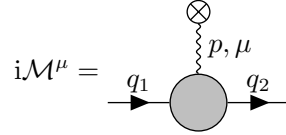


其中光子线代表一个比较强的、大小基本上确定的外场；这不代表不会有量子涨落。在上图中所有的量子涨落可以归结入顶角函数中；电子线由于amputating，无需自能修正，而光子线代表的外场也可以默认已经是amputated的。因此，严格计算电子自旋磁矩就归结为计算顶角函数。

在本节中，我们将计算一圈顶角函数修正，从中得到反常磁矩，并做重整化。

对称性分析和形状因子

S 矩阵矩阵元



是一个矢量，并且费曼图的形式显示它具有 $\bar{u} \times \cdots \times u$ 的形式，而系统中出现的矢量指标只有 γ^μ 以及三个动量，于是我们设

$$i\mathcal{M}^\mu = \text{diagram} = \bar{u}(q_2)(f_1\gamma^\mu + f_2p^\mu + f_3q_1^\mu + f_4q_2^\mu)u(q_1),$$

其中四个 f_i 因子是可以依赖于动量的复标量，此处复标量包括多分量对象 \not{p} 。这些因子必须是标量或是 \not{p} ，因为 \bar{u} 和 u 能够组合成的双线性形式的数目是有限的，以上四种情况穷尽了所有的 \bar{u} 和 u 能够组合成的双线性矢量。动量守恒意味着 f_2, f_3, f_4 不完全独立；不失一般性地我们不妨让 $f_2 = 0$ 。 f_i 中仅有的多分量对象就是 \not{q}_1 和 \not{q}_2 ，这两者都作用在 $u(q_1)$ 或是 $\bar{u}(q_2)$ 上，通过狄拉克方程

$$\not{q}_1 u(q_1) = m u(q_1), \quad \bar{u}(q_2) \not{q}_2 = m \bar{u}(q_2)$$

可以消去它们。因此 f_1, f_3, f_4 中只显含 $q_1^2, q_2^2, q_1 \cdot q_2$ 和 m 。前两者根据自由粒子的色散关系都是 m^2 ，因此最终 f_1, f_3, f_4 中只显含 m 和 $q_1 \cdot q_2$ ，或者等价地说 m 和 p^2 ；而量纲分析又告诉我们在这里真正重要的是 p^2/m^2 。使用Ward恒等式，我们有

$$\begin{aligned} 0 &= ip_\mu \mathcal{M}^\mu \\ &= \bar{u}(q_2)(f_1 p_\mu \gamma^\mu + f_3 p_\mu \not{q}_1^\mu + f_4 p_\mu \not{q}_2^\mu)u(q_1) \\ &= f_1 \bar{u}(q_2) \not{p} u(q_1) + f_3 \bar{u}(q_2) p \cdot q_1 u(q_1) + f_4 \bar{u}(q_2) p \cdot q_2 u(q_1), \end{aligned}$$

而

$$\begin{aligned} \bar{u}(q_2) \not{p} u(q_1) &= \bar{u}(q_2) \not{q}_2 u(q_1) - \bar{u}(q_2) \not{q}_1 u(q_1) \\ &= m \bar{u}(q_2) u(q_1) - m \bar{u}(q_2) u(q_1) = 0, \end{aligned}$$

且

$$\begin{aligned} p \cdot q_1 &= (q_2 - q_1) \cdot q_1 = q_2 \cdot q_1 - m^2, \\ p \cdot q_2 &= (q_2 - q_1) \cdot q_2 = m^2 - q_1 \cdot q_2, \end{aligned}$$

于是就能够证明 $f_3 = f_4$ 。因此独立的 f_i 只有两个。这样我们就有

$$\begin{aligned} i\mathcal{M}^\mu &= \bar{u}(q_2)(f_1\gamma^\mu + f_3q_1^\mu + f_3q_2^\mu)u(q_1) \\ &= (f_1 + 2mf_3)\bar{u}(q_2)\gamma^\mu u(q_1) + if_3\bar{u}(q_2)\sigma^{\mu\nu}(q_1^\nu - q_2^\nu)u(q_1), \end{aligned}$$

第二个等号用到了 Gordon 恒等式

$$\bar{u}(q_2)(q_1^\mu + q_2^\mu)u(q_1) = 2m\bar{u}(q_2)\gamma^\mu u(q_1) + i\bar{u}(q_2)\sigma^{\mu\nu}(q_1^\nu - q_2^\nu)u(q_1).$$

这样我们可以形式地写出

$$i\mathcal{M}^\mu = -ie\bar{u}(q_2)\left(F_1\left(\frac{p^2}{m^2}\right)\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m}p_\nu F_2\left(\frac{p^2}{m^2}\right)\right)u(q_1), \quad (5.35)$$

其中 F_1 和 F_2 被称为**形状因子**。通常将 F_2 限制为**圈图贡献的修正**。本节暂时将树图也纳入它们之中以便说明，最后再恢复正确的记号。

这个散射振幅提供了完整的外场和电子发生单次相互作用的方式。

一圈顶角函数

对自旋磁矩涉及的相互作用，我们考虑外场的 $\mu = 1, 2, 3$ 分量，因为 \mathbf{B} 和电势无关。(5.35)的第一项和自旋无关，因此我们将要考虑第二项。此外，在非相对论极限下， $q_1^0 = q_2^0 = m$ ，因此在这个极限下 $p_0 = 0$ 。因此我们只需要计算

$$i(\mathcal{M}^{\text{spin}})^i = -ie\bar{u}(q_2)\frac{i\sigma^{ij}}{2m}p_j F_2\left(\frac{p^2}{m^2}\right).$$

由于 $\mathbf{B} = \nabla \times \mathbf{A}$ ，在动量空间中 $\mathbf{B} = i\mathbf{p} \times \mathbf{A}$ ，即

$$B^i = i\epsilon^{ijk}p^j A^k,$$

因此我们有

$$\begin{aligned} i(\mathcal{M}^{\text{spin}})^i A_i &= -ie\bar{u}(q_2)\frac{i\sigma^{ij}}{2m}p_j F_2\left(\frac{p^2}{m^2}\right)u(q_1)A_i \\ &= -ie\bar{u}(q_2)\frac{i\sigma^{ij}}{2m}p^j A^i u(q_1)F_2\left(\frac{p^2}{m^2}\right) \\ &= -ie\bar{u}(q_2)\frac{i}{2m}\epsilon^{ijk}\begin{pmatrix} \sigma^k \\ \sigma^k \end{pmatrix}p_j A_i u(q_1)F_2\left(\frac{p^2}{m^2}\right) \\ &= -\frac{ie}{m}\bar{u}(q_2)\epsilon^{ijk}(ip^j)A^i S^k u(q_1)F_2\left(\frac{p^2}{m^2}\right) \\ &= -\frac{ie}{m}((-i\mathbf{p} \times \mathbf{A}) \cdot \mathbf{S})\bar{u}(q_2)u(q_1)F_2\left(\frac{p^2}{m^2}\right) \\ &= -\frac{ie}{m}(-\mathbf{B} \times \mathbf{S})\bar{u}(q_2)u(q_1)F_2\left(\frac{p^2}{m^2}\right). \end{aligned}$$

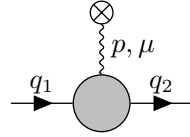
然后，在非相对论极限下 $u(q_1)$ 和 $\bar{u}(q_2)$ 可以写成 $(\xi \quad \xi)_{\mathbf{p}=0}$ 及其转置，如同在库伦相互作用中那样，我们有

$$\bar{u}(q_2)u(q_1) = 2m\sigma^0,$$

于是

$$i(\mathcal{M}^{\text{spin}})^i A_i = -\frac{ie}{m}(-\mathbf{B} \times \mathbf{S})2m\sigma^0 F_2(0).$$

这里我们已经将 F_2 的括号中的 p 取为零, 因为是非相对论极限。最后与推导库伦相互作用时类似, 切换到 $|\mathbf{p}\rangle$ 表象下, 由于有两条电子外线, 除以因子 $(\sqrt{2m})^2$, 就得到非相对论量子力学中的等效相互作用顶角



$$= -i \left(-F_2(0) \frac{e}{m} \mathbf{S} \cdot \mathbf{B} \right), \quad (5.36)$$

即, 电子的自旋和外磁场的相互作用哈密顿量为

$$H_{\text{spin}} = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad \boldsymbol{\mu} = \frac{e}{m} F_2(0) \mathbf{S}. \quad (5.37)$$

$F_2(0)$ 中树图阶的贡献我们已经在第5.3.2节中看到了, 就是1, 所以现在我们恢复惯常的记号, 让 F_2 只包括圈图部分的贡献, 就得到

$$H_{\text{spin}} = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad \boldsymbol{\mu} = \frac{e}{m} (1 + F_2(0)) \mathbf{S}, \quad (5.38)$$

或者用朗德 g 因子表示, 为

$$g = 2 + 2F_2(0), \quad (5.39)$$

其中 $F_2(0)$ 给出顶角函数圈图修正带来的修正。

下面我们要做的是计算 $F_2(0)$ 。本节只计算一圈图, 仅考虑顶角修正, 有

$$\begin{aligned} i\mathcal{M}^{(2)} &= \int \frac{d^4 k}{(2\pi)^4} \bar{u}(q_2) (-ie\gamma^\nu) \frac{i(\not{p} + \not{k} + m)}{(p+k)^2 - m^2 + i0^+} (-ie\gamma^\mu) \frac{i(\not{k} + m)}{k^2 - m^2 + i0^+} \\ &\quad \times (-ie\gamma^\sigma) u(q_1) \frac{-i\eta_{\sigma\nu}}{(k-q_1)^2 + i0^+} \\ &= -e^3 \bar{u}(q_2) \int \frac{d^4 k}{(2\pi)^4} \frac{\gamma^\nu (\not{p} + \not{k} + m) \gamma^\mu (\not{k} + m) \gamma_\nu}{(k^2 - m^2 + i0^+)((p+k)^2 - m^2 + i0^+)((k-q_1)^2 + i0^+)} u(q_1). \end{aligned} \quad (5.40)$$

为了计算上式我们做费曼参数化, 使用公式

$$\frac{1}{ABC} = 2 \int_0^1 dx dy dz \delta(x+y+z-1) \frac{1}{(xA+yB+zC)^3},$$

就得到

$$i\mathcal{M}^{(2)} = -2e^3 \int_0^1 dx dy dz \delta(x+y+z-1) \int \frac{d^4 k}{(2\pi)^4} \frac{N^\mu}{D}, \quad (5.41)$$

其中

$$N^\mu = \bar{u}(q_2) \gamma^\nu (\not{p} + \not{k} + m) \gamma^\mu (\not{k} + m) \gamma_\nu u(q_1), \quad (5.42)$$

而

$$D = (x(k^2 - m^2 + i0^+) + y((p+k)^2 - m^2 + i0^+) + z((k-q_1)^2 + i0^+))^3. \quad (5.43)$$

上式中, q_1 和 q_2 是在壳的, 而 p 是离壳的。我们有动量守恒关系

$$q_2 = q_1 + p. \quad (5.44)$$

我们首先化简 D 。将其展开，配方，并且利用在壳关系 $q_1^2 = m^2$ ，就得到

$$\begin{aligned} D &= (k^2 + 2yk \cdot p - 2zk \cdot q_1 + yp^2 + zq_1^2 - (x+y)m^2)^3 \\ &= ((k + yp - zq_1)^2 - y^2p^2 - z^2q_1^2 + 2yzp \cdot q_1 + yp^2 + zq_1^2 - (x+y)m^2)^3 \\ &= ((k + yp - zq_1)^2 - y^2p^2 + 2yzp \cdot q_1 + yp^2 - z^2m^2 + zm^2 - (x+y)m^2)^3. \end{aligned}$$

利用 $x + y + z = 1$ 得到

$$-z^2m^2 + zm^2 - (x+y)m^2 = -(1-z)^2m^2.$$

由于

$$p \cdot q_1 = \frac{1}{2}(q_1 + p)^2 - \frac{1}{2}q_1^2 - \frac{1}{2}p^2 = \frac{1}{2}(m^2 - m^2) - \frac{1}{2}p^2 = -\frac{1}{2}p^2,$$

我们有

$$-y^2p^2 + 2yzp \cdot q_1 + yp^2 = (-y^2 + y - yz)p^2 = y(1 - y - z)p^2 = xyp^2.$$

于是就得到

$$\begin{aligned} D &= ((k + yp - zq_1)^2 + xyp^2 - (1-z)^2m^2)^3 \\ &= (k'^2 - \Delta)^3, \end{aligned} \tag{5.45}$$

其中

$$k' = k + yp - zq_1, \quad \Delta = -xyp^2 + (1-z)^2m^2. \tag{5.46}$$

我们下面在对 k 做积分时可以将积分变量换成 k' 以简化计算。

然后我们分析 N^μ 。展开 \not{p} 的定义，并且使用 γ 矩阵的代数关系，可以得到

$$\begin{aligned} N^\mu &= \bar{u}(q_2)\gamma^\nu(p_\rho\gamma^\rho + k_\rho\gamma^\rho + m)\gamma^\mu(k_\sigma\gamma^\sigma + m)\gamma_\nu u(q_1) \\ &= \bar{u}(q_2)(4\eta^{\rho\mu}m(p_\rho + k_\rho) + (p_\rho + k_\rho)k_\sigma(-2\gamma^\sigma\gamma^\mu\gamma^\rho) + 4\eta^{\sigma\mu}mk_\sigma + m^2(-2\gamma^\mu))u(q_1) \\ &= \bar{u}(q_2)(4m(p^\mu + 2k^\mu) - 2\not{k}\gamma^\mu(\not{p} + \not{k}) - 2m^2\gamma^\mu)u(q_1). \end{aligned} \tag{5.47}$$

到了这一步，使用原本的 k 就很难计算下去了。我们现在将 k' 当成 k ，即在所有表达式中，做替换

$$k \longrightarrow k + zq_1 - yp, \tag{5.48}$$

就得到

$$D = (k^2 - \Delta)^3, \tag{5.49}$$

而

$$-\frac{1}{2}N^\mu = \bar{u}(q_2)(-2m(p^\mu + 2k^\mu + 2zq_1^\mu - 2yp^\mu) + (\not{k} + z\not{q}_1 - y\not{p})\gamma^\mu(\not{p} + \not{k} + z\not{q}_1 - y\not{p}) + m^2\gamma^\mu)u(q_1). \tag{5.50}$$

然后我们需要利用几个关系——Gordon恒等式，狄拉克方程的动量空间形式和 $x + y + z = 1$ ——来化简上式。此外，应注意到 D 对 k 是偶函数，从而 N^μ 相对 k 是奇函数的部分全部可以略去，既然 k 的取值范围遍布全空间。在上式第一项中， k^μ 项可以直接略去。在上式第二项中我们有

$$\begin{aligned} &\bar{u}(q_2)(\not{k} + z\not{q}_1 - y\not{p})\gamma^\mu(\not{p} + \not{k} + z\not{q}_1 - y\not{p})u(q_1) \\ &= \bar{u}(q_2)(\not{k} + z\not{q}_2 - (y+z)\not{p})\gamma^\mu(\not{k} + z\not{q}_1 + (1-y)\not{p})u(q_1) \\ &= \bar{u}(q_2)(\not{k} + zm - (y+z)\not{p})\gamma^\mu(\not{k} + zm + (1-y)\not{p})u(q_1). \end{aligned} \tag{5.51}$$

以上计算中用到了

$$\bar{u}(q_2)\not{q}_2 = m\bar{u}(q_2), \quad \not{q}_1 u(q_1) = mu(q_1).$$

(5.51)中又有一些 k 的奇函数项，可以直接略去。不能略去的项包括 $\not{k}\gamma^\mu\not{k}$ 项， $\not{p}\gamma^\mu$ 项及其转置，以及 $\not{p}\gamma^\mu\not{p}$ 项。在计算这些项时，我们有这样一个有用的公式：在所有量都是洛伦兹协变的时候，有

$$\int d^4k k_\mu k_\nu \times (\cdots) = \int d^4k \frac{1}{4}\eta_{\mu\nu}k^2 \times (\cdots), \quad (5.52)$$

其推导是，首先注意到 $\mu \neq \nu$ 时积分为零，而上式左边给出一个张量，满足这个条件的张量只有度规张量的某个倍数；系数 $1/4$ 是通过在等式两边同时乘上 $\eta^{\mu\nu}$ 并缩并而计算出来的。于是

$$\begin{aligned} \not{k}\gamma^\mu\not{k} &= k_\rho k_\sigma \gamma^\rho \gamma^\mu \gamma^\sigma = \frac{1}{4}\eta_{\rho\sigma}k^2 \gamma^\rho \gamma^\mu \gamma^\sigma \\ &= \frac{1}{4}k^2 \gamma^\sigma \gamma^\mu \gamma_\sigma = -\frac{1}{2}k^2 \gamma^\mu, \end{aligned}$$

而

$$\gamma^\mu \not{p} = p_\sigma \frac{\{\gamma^\sigma, \gamma^\mu\} - [\gamma^\sigma, \gamma^\mu]}{2} = p_\sigma \frac{2\eta^{\sigma\mu} - 2i\sigma^{\mu\sigma}}{2} = p^\mu - i\sigma^{\mu\sigma} p_\sigma,$$

同理

$$\not{p}\gamma^\mu = p^\mu + i\sigma^{\mu\sigma} p_\sigma.$$

对 $\not{p}\gamma^\mu\not{p}$ 项则有

$$\not{p}\gamma^\mu\not{p} = p_\rho p_\sigma \left(2\gamma^\rho \eta^{\mu\sigma} - \frac{\{\gamma^\rho, \gamma^\sigma\} + [\gamma^\rho, \gamma^\sigma]}{2} \gamma^\mu \right),$$

注意到 $p_\rho p_\sigma [\gamma^\rho, \gamma^\sigma]$ 是零，因为对易子反对称，于是

$$\not{p}\gamma^\mu\not{p} = 2\not{p}p^\mu - p^2 \gamma^\mu,$$

而这个式子中的第一项又应当略去，因为

$$\bar{u}(q_2)\not{p}u(q_1) = \bar{u}(q_2)(\not{q}_2 - \not{q}_1)u(q_1) = 0.$$

因此

$$\not{p}\gamma^\mu\not{p} = -p^2 \gamma^\mu.$$

综合以上各式，现在我们已经把 N^μ 写成不带有 \not{p} 的形式了：

$$\begin{aligned} -\frac{1}{2}N^\mu &= \bar{u}(q_2)(\gamma^\mu(-\frac{1}{2}k^2 + m^2 + z^2 m^2 + (y+z)(1-y)p^2) \\ &\quad - 2m((1-2y)p^\mu + 2zq_1^\mu) \\ &\quad + zm((1-y)(p^\mu - i\sigma^{\mu\sigma} p_\sigma) - (y+z)(p^\mu + i\sigma^{\mu\sigma} p_\sigma)))u(q_1). \end{aligned} \quad (5.53)$$

最后，将括号中的第二项和第三项放在一起，得到

$$\begin{aligned} &- 2m((1-2y)p^\mu + 2zq_1^\mu) + zm((1-y)(p^\mu - i\sigma^{\mu\sigma} p_\sigma) - (y+z)(p^\mu + i\sigma^{\mu\sigma} p_\sigma)) \\ &= m((4y-2+z-2yz-z^2)p^\mu - z(1+z)i\sigma^{\mu\nu} p_\nu - 4zq_1^\mu) \\ &= m((4y-2+3z-2yz-z^2)p^\mu - z(1+z)i\sigma^{\mu\nu} p_\nu - 2z(q_1^\mu + q_2^\mu)) \\ &= m(-(z-2)(z+2y-1)p^\mu - z(1+z)i\sigma^{\mu\nu} p_\nu - 2z(q_1^\mu + q_2^\mu)), \end{aligned}$$

使用Gordon恒等式, 将 $q_1 + q_2$ 用 $2m\gamma^\mu + i\sigma^{\mu\nu}(q_1^\nu - q_2^\nu)$ 代替, 得到

$$\begin{aligned} & -2m((1-2y)p^\mu + 2zq_1^\mu) + zm((1-y)(p^\mu - i\sigma^{\mu\sigma}p_\sigma) - (y+z)(p^\mu + i\sigma^{\mu\sigma}p_\sigma)) \\ & = m(-(z-2)(z+2y-1)p^\mu - z(1+z)i\sigma^{\mu\nu}p_\nu - 2z(2m\gamma^\mu - i\sigma^{\mu\nu}p_\nu)) \\ & = m((z-2)(x-y)p^\mu + z(1-z)i\sigma^{\mu\nu}p_\nu - 4mz\gamma^\mu), \end{aligned}$$

然后我们再将上式和(5.53)右边括号中的第一项放在一起, 得到

$$\begin{aligned} -\frac{1}{2}N^\mu & = \bar{u}(q_2)(\gamma^\mu(-\frac{1}{2}k^2 + (z^2 - 4z + 1)m^2 + (1-x)(1-y)p^2) \\ & \quad + imz(1-z)\sigma^{\mu\nu}p_\nu + m(z-2)(x-y)p^\mu)u(q_1). \end{aligned} \quad (5.54)$$

最后, 我们将(5.49)和(5.54)放在一起, 计算对 x, y, z, p 的积分。 N^μ 中的三项会让(5.41)中出现三项, 而按照(5.35), 只应该有两项。实际上我们会注意到 Δ 中交换 x 和 y 不变, 而(5.54)中, p^μ 这一项中 x 和 y 交换会多一个负号, 因此 p^μ 项对(5.41)其实是没有贡献的。实际上, 这个事实对应着Ward恒等式, 因为(5.54)中的另外两项中,

$$\bar{u}(q_2)p^\mu\gamma^\mu u(q_1) = \bar{u}(q_2)(\not{q}_2 - \not{q}_1)u(q_1) = \bar{u}(q_2)(m - m)u(q_1) = 0, \quad p_\mu\sigma^{\mu\nu}p_\nu = 0,$$

因此

$$p_\mu N^\mu = 0, \quad p_\mu \mathcal{M}^\mu = 0.$$

这就是Ward恒等式。与(5.35)对比, 我们能够得到

$$\begin{aligned} F_1\left(\frac{p^2}{m^2}\right) & = 4ie^2 \int_0^1 dx dy dz \delta(x+y+z-1) \\ & \quad \times \int \frac{d^4k}{(2\pi)^4} \frac{-\frac{1}{2}k^2 + (z^2 - 4z + 1)m^2 + (1-x)(1-y)p^2}{(k^2 - \Delta)^3}, \end{aligned} \quad (5.55)$$

而

$$F_2\left(\frac{p^2}{m^2}\right) = 8ie^2 m^2 \int_0^1 dx dy dz \delta(x+y+z-1) \int \frac{d^4k}{(2\pi)^4} \frac{z(1-z)}{(k^2 - \Delta)^3}. \quad (5.56)$$

要计算反常磁矩只需要计算 F_2 。由于

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2 - \Delta)^3} = -\frac{i}{32\pi^2\Delta},$$

我们有

$$\begin{aligned} F_2\left(\frac{p^2}{m^2}\right) & = \frac{e^2 m^2}{4\pi^2} \int_0^1 dx dy dz \delta(x+y+z-1) \frac{z(1-z)}{(1-z)^2 m^2 - xyp^2} \\ & = \frac{\alpha}{\pi} m^2 \int_0^1 dx dy dz \delta(x+y+z-1) \frac{z(1-z)}{(1-z)^2 m^2 - xyp^2}. \end{aligned}$$

最后, 我们可以计算出反常磁矩的朗德 g 因子了:

$$\begin{aligned} F_2(0) & = \frac{\alpha}{\pi} m^2 \int_0^1 dx dy dz \delta(x+y+z-1) \frac{z(1-z)}{(1-z)^2 m^2} \\ & = \frac{\alpha}{\pi} \int_0^1 dx dy dz \frac{z}{1-z} \\ & = \frac{\alpha}{2\pi}, \end{aligned}$$

于是

$$g = 2 + 2F_2(0) = 2 + \frac{\alpha}{\pi}. \quad (5.57)$$

这给出了电子的磁矩的正比于 α 的第一阶修正。

我们的运气是很好的, 在计算过程中没有遇到发散, 因为发散都在 F_1 里面了。

5.5 一圈图的正规化和重整化

本节更加系统地计算一圈图修正。我们现在要做的事情是：

1. 画出电子传播子、光子传播子和顶角函数的一圈图修正；
2. 根据费曼规则写出这些图（其中所有“外线”均不在壳）的表达式，然后依次做：
 - (a) γ 矩阵缩并；
 - (b) 将 $1/(k^2 - m^2)$ 连乘做费曼参数化或者说费曼折叠，引入一系列从0到1的积分；
 - (c) 做出动量积分，可能需要做适当的动量平移以便于套用一些公式；
 - (d) 完成费曼参数化中引入的积分。

在此过程中用维数正规化的 $d = 4 - \epsilon$ 将积分中的紫外发散部分提取出来。如果有红外发散，需要给光子加一个非常小的质量，以避免软光子散射。

3. 根据前述的紫外发散部分，适当引入对应于电子传播子、光子传播子和顶角函数的抵消项。
4. 计算电子传播子、光子传播子和顶角函数的一圈图修正和抵消项的贡献的和，消去所有的发散，得到修正后的有限的电子传播子、光子传播子和顶角函数。
5. 用这三个量组合得到需要的散射振幅。

5.5.1 三种紫外发散的一圈图的维数正规化

真空极化

极化指的是介质中光子的传播偏离经典电动力学中真空的情况，在QED中真空极化实际上就是光子自能，是光子传播子的一圈图修正。对称性分析表明光子自能的表达式一定是某个关于光子动能 p 的标量乘以某个正比于 p 的二次项的二阶张量。关于 p 的标量当然一定是 p^2 的函数，而用 p 能够构造出来的二阶张量无非是 $p^\mu p^\nu$ 和 $p^2 \eta^{\mu\nu}$ ，因此光子自能的一般形式是

$$\Pi^{\mu\nu}(p) = \Delta_1(p^2, m^2) p^2 \eta^{\mu\nu} + \Delta_2(p^2, m^2) p^\mu p^\nu. \quad (5.58)$$

然而，这第二项实际上没有太大作用，因为QED是规范场论，我们可以通过调整 ξ 来消掉这一项，而不产生其它任何影响。换言之，在之后计算光子自能时，但凡出现 $p^\mu p^\nu$ 项的，全部可以忽略——它们除了让 ξ 需要调整以下以外什么也不导致。

光子的正规自能中的一圈图是

$$\begin{aligned}
 i\Pi^{\mu\nu} &= \text{diagram} \\
 &= - \int \frac{d^4 k}{(2\pi)^4} \text{tr}(-ie\gamma^\mu) \frac{i(\not{k} - \not{p} + m)}{(k-p)^2 - m^2 + i0^+} (-ie\gamma^\nu) \frac{i(\not{k} + m)}{k^2 - m^2 + i0^+} \\
 &= -e^2 \int \frac{d^4 k}{(2\pi)^4} \frac{\text{tr}(\gamma^\mu (\not{k} - \not{p} + m) \gamma^\nu (\not{k} + m))}{((k-p)^2 - m^2 + i0^+)(k^2 - m^2 + i0^+)}.
 \end{aligned} \quad (5.59)$$

注意这里我们把光子自能图设成 $i\Pi$ 而不是 $-i\Sigma$ ，因为光子关联函数的形式相比电子关联函数多出来了一个负号。（但是一些书，如Schwartz，对电子用的也是 $i\Sigma$ ）图中所有的动量都不必在壳。

如果直接动手计算，马上就会发现这里有发散。我们此处做维数正规化，即计算

$$\Pi^{\mu\nu} = ie^2 \int \frac{d^d k}{(2\pi)^d} \frac{\text{tr}(\gamma^\mu (\not{k} - \not{p} + m) \gamma^\nu (\not{k} + m))}{((k-p)^2 - m^2 + i0^+)(k^2 - m^2 + i0^+)}. \quad (5.60)$$

首先做费曼参数化，有

$$\begin{aligned} \frac{1}{((k-p)^2 - m^2 + i0^+)(k^2 - m^2 + i0^+)} &= \int_0^1 dx \frac{1}{(x(k^2 - m^2) + (1-x)((k-p)^2 - m^2))^2} \\ &= \int_0^1 dx \frac{1}{(k^2 + (1-x)p^2 - 2(1-x)k \cdot p - m^2)^2}. \end{aligned}$$

然后计算分子，根据 γ 矩阵的迹公式（这些公式在维数正规化下是不变的），有

$$\text{tr}(\gamma^\mu (\not{k} - \not{p} + m) \gamma^\nu (\not{k} + m)) = 4(2k^\mu k^\nu - p^\mu k^\nu - k^\mu p^\nu + \eta^{\mu\nu}(m^2 - k^2 + k \cdot p)).$$

上式中的 $p^\mu p^\nu$ 项显然可以略去，它们只会让 ξ 做一个调整；由对称性， $p^\mu k^\nu$ 项和 $k^\mu p^\nu$ 项在对 k 积分后显然都会给出正比于 $p^\mu p^\nu$ 的项，从而也可以略去。最后，为了保持量纲——或者说为了显式地引入 ϵ 标记的重整化群流的标度变换部分——我们注意到在 d 维空间中 e 的量纲是 $2 - d/2$ ，因此如果我们固定 e 为四维空间中的电荷（从而其取值和 ϵ 无关），则在 d 维空间中应该用 $\mu^{2-d/2}e$ 代替 e ，以手动引入 ϵ 变动导致的跑动。因此我们得到

$$\Pi^{\mu\nu} = 4i\mu^{4-d}e^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{2k^\mu k^\nu + \eta^{\mu\nu}(m^2 + k \cdot p - k^2)}{(k^2 + (1-x)p^2 - 2(1-x)k \cdot p - m^2)^2}.$$

容易观察出，将分母配方的方法是做积分变量平移

$$k \longrightarrow k + (1-x)p,$$

就能够让分母对 k 的依赖只含有 k^2 。此时并再一次忽略所有 $p^\mu p^\nu$ 项，以及，由于此时分母对 k 是偶函数，可以忽略所有的 $k \cdot p$ 项，就有

$$\Pi^{\mu\nu} = 4i\mu^{4-d}e^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{2k^\mu k^\nu - \eta^{\mu\nu}(k^2 - x(1-x)p^2 - m^2)}{(k^2 + x(1-x)p^2 - m^2)^2}. \quad (5.61)$$

使用(5.52)，注意此时由于维数发生变化，应该使用其 d 维版本，即

$$\int d^d k k_\mu k_\nu \times (\cdots) = \int d^d k \frac{1}{d} \eta_{\mu\nu} k^2 \times (\cdots) \quad (5.62)$$

就得到

$$\Pi^{\mu\nu} = 4i\mu^{4-d}e^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{(\frac{2}{d} - 1)\eta^{\mu\nu}k^2 + \eta^{\mu\nu}(x(1-x)p^2 + m^2)}{(k^2 - \Delta)^2},$$

其中

$$\Delta = m^2 - x(1-x)p^2.$$

使用公式

$$\int \frac{d^d k}{(2\pi)^d} \frac{k^2}{(k^2 - \Delta)^2} = -\frac{i}{(4\pi)^{d/2}} \frac{d}{2} \Gamma\left(1 - \frac{d}{2}\right) \left(\frac{1}{\Delta}\right)^{1-\frac{d}{2}},$$

以及

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - \Delta)^2} = \frac{i}{(4\pi)^{d/2}} \Gamma\left(2 - \frac{d}{2}\right) \left(\frac{1}{\Delta}\right)^{2 - \frac{d}{2}},$$

并且在 $\Pi^{\mu\nu}$ 的第一项中做替换

$$\left(\frac{1}{\Delta}\right)^{1 - \frac{d}{2}} = \left(\frac{1}{\Delta}\right)^{2 - \frac{d}{2}} (m^2 - x(1-x)p^2), \quad \Gamma\left(1 - \frac{d}{2}\right) = \frac{\Gamma\left(2 - \frac{d}{2}\right)}{1 - \frac{d}{2}},$$

计算得到

$$\Pi^{\mu\nu} = -8p^2 \eta^{\mu\nu} \frac{e^2}{(4\pi)^{d/2}} \mu^{4-d} \Gamma\left(2 - \frac{d}{2}\right) \int_0^1 dx x(1-x) \left(\frac{1}{m^2 - x(1-x)p^2}\right)^{2 - \frac{d}{2}}.$$

现在引入 $d = 4 - \epsilon$ 展开, 我们得到

$$\begin{aligned} \Pi^{\mu\nu} &= -\frac{8p^2 \eta^{\mu\nu} e^2}{(4\pi)^2} \int_0^1 dx x(1-x) \left(1 + \frac{\epsilon}{2} \ln\left(\frac{4\pi\mu^2}{m^2 - x(1-x)p^2}\right) + \mathcal{O}(\epsilon^2)\right) \\ &\quad \times \left(\frac{2}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon)\right) \\ &= -\frac{p^2 \eta^{\mu\nu} e^2}{2\pi^2} \int_0^1 dx x(1-x) \left(\frac{2}{\epsilon} + \ln\left(\frac{\tilde{\mu}^2}{m^2 - x(1-x)p^2}\right)\right), \end{aligned} \quad (5.63)$$

其中我们定义

$$\tilde{\mu}^2 = 4\pi e^{-\gamma_E} \mu^2. \quad (5.64)$$

(5.63)给出正规化之后的光子自能。可以看到的确存在发散, 不过发散已经被提取出来了, 它是 ϵ 的一个一阶极点, 即

$$\Pi^{\mu\nu} = -\frac{p^2 e^2 \eta^{\mu\nu}}{6\pi^2 \epsilon} + \text{finite}. \quad (5.65)$$

电子自能

电子和光子的相互作用存在相互作用这一事实意味着可以有“电子吸收它自己产生的光子”这样的过程, 从而电子会带上“电磁质量”。这是在经典电动力学中也已经知道的一个现象, 但是在经典电动力学中不足以处理自能导致的发散。

我们现在在QED中处理电子自能。电子的正规自能中的一圈图是

$$\begin{aligned} -i\Sigma(\not{p}) &= \text{Diagram: a fermion line with momentum } p \text{ and } p, \text{ with a photon loop labeled } p-k \text{ and } k. \\ &= \int \frac{d^4 k}{(2\pi)^4} (-ie\gamma^\mu) \frac{i}{\not{k} - m + i0^+} (-ie\gamma^\nu) \frac{-i\eta_{\mu\nu}}{(p-k)^2 + i0^+} \\ &= -e^2 \int \frac{d^4 k}{(2\pi)^4} \gamma^\mu \frac{\not{k} + m}{k^2 - m^2 + i0^+} \gamma_\mu \frac{1}{(p-k)^2 + i0^+}. \end{aligned} \quad (5.66)$$

图中所有的动量都不必在壳。

仿照先前做光子自能计算时的步骤, 首先做维数正规化, 并引入因子 μ , 得到

$$\Sigma(\not{p}) = -i\mu^{4-d} e^2 \int \frac{d^d k}{(2\pi)^d} \gamma^\mu \frac{\not{k} + m}{k^2 - m^2 + i0^+} \gamma_\mu \frac{1}{(p-k)^2 + i0^+}. \quad (5.67)$$

首先处理分子, 根据 γ 矩阵的代数公式得到

$$\gamma^\mu (\not{k} + m) \gamma_\mu = md - (d-2)\not{k}.$$

对分母，做费曼参数化，得到

$$\begin{aligned}\frac{1}{(k^2 - m^2)(p - k)^2} &= \int_0^1 dx \frac{1}{(x(p - k)^2 + (1 - x)(k^2 - m^2))^2} \\ &= \int_0^1 dx \frac{1}{(k^2 - 2xp \cdot k + xp^2 + (x - 1)m^2)^2}.\end{aligned}$$

容易观察出正确的配方方式是做变量替换

$$k \longrightarrow k + xp,$$

此时分母变成 $(k^2 - \Delta)^2$ ，其中

$$\Delta = (1 - x)(m^2 - p^2 x),$$

而分子成为

$$md - (d - 2)(\not{k} + xp).$$

此时分母是 k 的偶函数，从而分子中是 k 的奇函数的项积分之后都是零，可以略去，因此最终得到

$$\begin{aligned}\Sigma(\not{p}) &= -i\mu^{4-d}e^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{md - x(d - 2)\not{p}}{(k^2 - \Delta)^2} \\ &= -i\mu^{4-d}e^2 \int_0^1 dx (md - x(d - 2)\not{p}) \frac{i}{(4\pi)^{d/2}} \Gamma\left(2 - \frac{d}{2}\right) \left(\frac{1}{\Delta}\right)^{2 - \frac{d}{2}}.\end{aligned}\quad (5.68)$$

做 $d = 4 - \epsilon$ 展开，得到

$$\begin{aligned}\Sigma(\not{p}) &= \frac{ie^2\mu^\epsilon}{(4\pi)^2} \int_0^1 dx \Gamma\left(\frac{\epsilon}{2}\right) (4\pi)^{\epsilon/2} \frac{(4 - \epsilon)m - (2 - \epsilon)x\not{p}}{((1 - x)(m^2 - xp^2))^{\epsilon/2}} \\ &= \frac{\alpha}{2\pi} \int_0^1 dx \left(\left(\frac{2}{\epsilon} + \ln\left(\frac{\tilde{\mu}^2}{(1 - x)(m^2 - xp^2)}\right)\right) (2m - x\not{p}) + (x\not{p} - m) \right),\end{aligned}\quad (5.69)$$

提取发散部分，为

$$\Sigma(\not{p}) = \frac{\alpha(4m - \not{p})}{2\pi\epsilon} + \text{finite}.\quad (5.70)$$

在计算电子自能时我们都采用了最为谨慎的做法，即写下费曼图对应的振幅后立刻将维数延拓到 $d = 4 - \epsilon$ ，然后做后面的计算。在计算光子自能时我们首先在 $d = 4$ 下消去了所有 γ 矩阵，然后再将维数延拓到 $d = 4 - \epsilon$ 。为什么计算电子自能时不能这么做？答案是，计算光子自能时我们在做的是对一系列 γ 矩阵的乘积求迹，这在任何维数下都给出一样的结果；换句话说，计算光子自能时我们是在对自旋和手性指标做缩并，这和空间维数无关。在计算电子自能时，我们是在缩并空间指标，例如我们会需要计算 $\gamma^\mu \gamma^\nu \gamma_\nu$ 这样的式子，而空间指标能够跑遍1到 d ，因此和 d 是有关的，因此需要先做维数延拓做 γ 矩阵缩并和动量积分。

不过，如果我们只是想计算出发散部分，那么无论先做 γ 矩阵缩并还是先做维数延拓都是一样的，因为 γ 矩阵缩并不导致任何发散。换言之，(5.70)和(5.65)这样的式子可以通过先做维数延拓再计算 γ 矩阵缩并的方法计算出来，也可以通过先计算 γ 矩阵缩并再做维数延拓的方法计算出来。使用后一种方法能够得到正确的发散项，但是不能得到正确的有限部分；使用这种方法认真计算(5.70)和(5.65)中的有限部分就错了，只计算发散部分却是对的。

顶角函数

我们现在重做第5.4.1节中的计算。按理说，此时两个电子也应该是离壳的，因为这是在修正顶角函数，本身不能保证在壳。但是我们马上会注意到QED中amputated one loop diagram中外线不可能是光子线，因此，在本节讨论的一圈图修正中，实际上只需要计算电子在壳的顶角函数的修正即可。

关于形状因子之类的讨论仍然是适用的。我们将一圈图导致的修正复述如下：

$$-ie\Gamma^\mu(p) = -e^3 \int \frac{d^4k}{(2\pi)^4} \frac{\bar{u}(q_2)\gamma^\nu(\not{p} + \not{k} + m)\gamma^\mu(\not{k} + m)\gamma_\nu u(q_1)}{(k^2 - m^2 + i0^+)((p+k)^2 - m^2 + i0^+)((k-q_1)^2 + i0^+)},$$

这里没有引入任何外线和 $u(q)$ 矩阵。一开始就引入维数延拓（用于消去紫外发散）和光子质量（用于消去红外发散），得到

$$\Gamma^\mu(p) = -i\mu^{4-d}e^2 \int \frac{d^dk}{(2\pi)^d} \frac{\bar{u}(q_2)\gamma^\nu(\not{p} + \not{k} + m)\gamma^\mu(\not{k} + m)\gamma_\nu u(q_1)}{(k^2 - m^2 + i0^+)((p+k)^2 - m^2 + i0^+)((k-q_1)^2 - m_\gamma^2 + i0^+)}, \quad (5.71)$$

并且对称性分析意味着

$$\Gamma^\mu(p) = F_1 \left(\frac{p^2}{m^2} \right) \gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m} p_\nu F_2 \left(\frac{p^2}{m^2} \right) \quad (5.72)$$

使用 γ 矩阵的代数关系可以将分子化为

$$\begin{aligned} & \gamma^\nu(\not{p} + \not{k} + m)\gamma^\mu(\not{k} + m)\gamma_\nu \\ &= \epsilon(\not{p} + \not{k} - m)\gamma^\mu \not{k} - 2\not{k}\gamma^\mu(\not{p} + \not{k}) - m^2(2 - \epsilon)\gamma^\mu + 4m(p^\mu + 2k^\mu). \end{aligned}$$

再对分母做费曼折叠，得到

$$\Gamma^\mu(p) = -2i\mu^{4-d}e^2 \int_0^1 dx dy dz \delta(x+y+z-1) \int \frac{d^dk}{(2\pi)^d} \frac{N^\mu}{D}, \quad (5.73)$$

其中

$$D = (x(k^2 - m^2 + i0^+) + y((p+k)^2 - m^2 + i0^+) + z((k-q_1)^2 + i0^+))^3 \quad (5.74)$$

而

$$N^\mu = \epsilon(\not{p} + \not{k} - m)\gamma^\mu \not{k} - 2\not{k}\gamma^\mu(\not{p} + \not{k}) - m^2(2 - \epsilon)\gamma^\mu + 4m(p^\mu + 2k^\mu). \quad (5.75)$$

之后的步骤和第5.4.1节中是完全一样的。一些观察可以大大简化计算：首先注意到 F_2 根本就不发散，我们实际上只需要计算 F_1 然后把 F_2 加上即可。然后我们会发现， $d = 4 - \epsilon$ 导致的额外的项可以归结为： $m^2\gamma^\mu$ 项和 $\not{k}\gamma^\mu \not{k}$ 项均需要乘上 $(2 - \epsilon)/2$ ； $\not{k}\gamma^\mu \not{p}$ 项也需要乘上 $(2 - \epsilon)/2$ （实际上是加上 $\epsilon \not{p}\gamma^\mu \not{k}$ ，但是交换 \not{p} 和 \not{k} 只会多出来一个 $\sigma^{\mu\nu}$ 项，从而归入 F_2 项中）；多出来一个 $-\epsilon m\gamma^\mu \not{k}$ 。如前所述， p^μ 最终会消失，这是Ward恒等式的一个推论。

实际上，如果只是希望得到发散部分，那么直接对(5.55)的动量积分做维数延拓即可。其发散部分由 k^2 项贡献，为

$$\begin{aligned} F_1 &= 4ie^2 \int_0^1 dx dy dz \int \frac{d^dk}{(2\pi)^d} \frac{-\frac{1}{2}k^2}{(k^2 - \Delta)^3} + \text{finite} \\ &= -2ie^2 \int_0^1 dx dy dz \delta(x+y+z-1) \frac{i}{(4\pi)^{2-\epsilon/2}} (2 - \epsilon/2) \frac{\Gamma(\frac{\epsilon}{2})}{2} \left(\frac{1}{\Delta} \right)^\epsilon + \text{finite} \\ &= 2e^2 \int_0^1 dx dy dz \delta(x+y+z-1) \frac{1}{(4\pi)^2} \frac{2}{\epsilon} + \text{finite} \\ &= \frac{e^2}{8\pi^2\epsilon} + \text{finite}, \end{aligned}$$

即

$$\Gamma^\mu(p^2) = \gamma^\mu \left(1 + \frac{e^2}{8\pi^2\epsilon} \right) + \text{finite} = \gamma^\mu + \gamma^\mu \frac{\alpha}{2\pi\epsilon} + \text{finite}. \quad (5.76)$$

5.5.2 重整化方案

抵消项

到目前为止，我们发现树图中需要修正的组件——电子传播子，光子传播子，相互作用顶角——的圈图修正均存在发散，并且已经通过维数正规化提取出了这些发散。然后就需要通过引入抵消项来消除这些发散。

本节采用一种比较简单的做法。我们要计算散射振幅，而LSZ约化公式为

$$\langle p_1, p_2, \dots, p_m | k_1, k_2, \dots, k_n \rangle_{p_i^0 \rightarrow \omega_{\mathbf{p}_i}, k_j^0 \rightarrow \omega_{\mathbf{k}_j}} (\sqrt{Z})^{m+n} \times \text{amputated diagrams}. \quad (5.77)$$

其中

$$Z = |\langle \Omega | \phi(0) | p = 0 \rangle|^2. \quad (5.78)$$

(5.77)右边的amputated diagram是使用裸的场和（会发散的）裸的参数计算出来的，左边则是关于（有限的）物理参数的散射振幅，它就是重整化后的场的关联函数扣除了入射、出射线的传播子得到的结果。右边的 Z 因子将右边的场强重整化前的场和左边的场强重整化后的场联系起来。将(5.77)右边的 Z 因子和amputated diagram合并起来，就相当于在计算外线是重整化后的场，但是费曼规则还是裸的的amputated diagram。

下面我们将拉氏量中的所有裸的东西都加上下标0，即将拉氏量写成

$$\mathcal{L} = -\frac{1}{4}(F_0^{\mu\nu})^2 + \bar{\psi}_0(i\gamma_\mu \partial^\mu - m_0)\psi_0 - e_0 A_0^\mu \bar{\psi}_0 \gamma_\mu \psi_0 =: \mathcal{L}_{\text{bare}}. \quad (5.79)$$

我们用下标R表示重整化之后的量，为了和(5.77)保持一致，我们设

$$\psi_0 = \sqrt{Z_2}\psi_R, \quad A_0^\mu = \sqrt{Z_3}A_R^\mu, \quad m_0 = Z_m m_R, \quad e_0 = Z_e e_R, \quad (5.80)$$

实际上，如果我们要求 e_R 是四维时空中的物理电荷，则我们应该设

$$\mu^{\frac{4-d}{2}} e_R = \frac{e_0}{Z_e} \quad (5.81)$$

以引入 ϵ 变化时引入的动量标度，但是我们在做重整化计算时暂时将 μ 吸收进 e_R ，而在分析重整化群流时再把它拿出来；或者说，在做重整化计算时我们设 e_R 为 d 维时空中的物理电荷。这样裸拉氏量就是

$$\begin{aligned} \mathcal{L}_{\text{bare}} &= -\frac{1}{4}Z_3(F_R^{\mu\nu})^2 + Z_2\bar{\psi}_R(i\gamma_\mu \partial^\mu - Z_m m_R)\psi_R - \sqrt{Z_3}Z_e Z_2 e_R A_R^\mu \bar{\psi}_R \gamma_\mu \psi_R \\ &= -\frac{1}{4}Z_3(F_R^{\mu\nu})^2 + Z_2\bar{\psi}_R(i\gamma_\mu \partial^\mu - Z_m m_R)\psi_R - Z_1 e_R A_R^\mu \bar{\psi}_R \gamma_\mu \psi_R, \end{aligned} \quad (5.82)$$

其中我们定义

$$Z_1 = \sqrt{Z_3}Z_e Z_2. \quad (5.83)$$

我们把(5.82)分离成物理部分 $\mathcal{L}_{\text{phys}}$ 和抵消项 $\delta\mathcal{L}$ ，就有

$$\begin{aligned} \mathcal{L} &= \underbrace{-\frac{1}{4}(F_R^{\mu\nu})^2 + \bar{\psi}_R(i\gamma_\mu \partial^\mu - m_R)\psi_R - e_R A_R^\mu \bar{\psi}_R \gamma_\mu \psi_R}_{\mathcal{L}_{\text{phys}}} \\ &\quad - \underbrace{\frac{1}{4}\delta Z_3 (F_R^{\mu\nu})^2 + \delta Z_2 \bar{\psi}_R i\gamma_\mu \partial^\mu \psi_R - (\delta Z_2 + \delta Z_m)m_R \bar{\psi}_R \psi_R - \delta Z_1 e_R A_R^\mu \bar{\psi}_R \gamma_\mu \psi_R}_{\delta\mathcal{L}}, \end{aligned} \quad (5.84)$$

其中

$$Z_1 = 1 + \delta Z_1, \quad Z_2 = 1 + \delta Z_2, \quad Z_3 = 1 + \delta Z_3. \quad (5.85)$$

严格来说按照(5.84)，应该有

$$\delta Z_m = Z_2(Z_m - 1), \quad (5.86)$$

但是在本节讨论的一圈图的重整化中 δZ_i 和 δZ_m 都被用于消去(5.65)，(5.70)和(5.76)这三个发散，而它们都正比于 e^2 ，从而可以认为

$$\delta Z_m = Z_m - 1. \quad (5.87)$$

我们计算(5.65)，(5.70)和(5.76)这三个发散时可以认为是在使用裸的拉氏量 $\mathcal{L}_{\text{bare}}$ ，也可以认为是在使用物理的拉氏量 $\mathcal{L}_{\text{phys}}$ ，既然这两个拉氏量的形式完全相同。本文采取的方法是，认为(5.65)，(5.70)和(5.76)这三个发散是在 $\mathcal{L}_{\text{phys}}$ 下计算的。在(5.84)中， $\mathcal{L}_{\text{phys}}$ 是关于物理参数的，但是会产生发散，而 $\delta\mathcal{L}$ 对应着裸的拉氏量中无法观测到的发散部分，正好能够消去用 $\mathcal{L}_{\text{phys}}$ 计算的图的分散。

切换到动量空间下，抵消项给出了三个可以用来抵消发散的相互作用顶角：

$$\text{---} \otimes \text{---} = i(\not{p} \delta Z_2 - (\delta Z_2 + \delta Z_m) m_R), \quad (5.88)$$

$$\text{~~~~~} \otimes \text{~~~~~} = -i \delta Z_3 (p^2 \eta^{\mu\nu} - p^\mu p^\nu), \quad (5.89)$$

或者拉氏量做了Faddeev-Popov量子化，那么费曼规范 $\frac{1}{2}(\partial_\mu A_\mu)^2$ 也会有一个 Z_3 因子，从而会消掉上式中第二项，给出

$$\text{~~~~~} \otimes \text{~~~~~} = -i \delta Z_3 p^2 \eta^{\mu\nu}. \quad (5.90)$$

还有一个顶角的抵消项

$$\text{---} \otimes \text{---} = -i \delta Z_1 e_R \gamma^\mu. \quad (5.91)$$

不言而喻，物理参数必须要能够通过一个实验过程提取出来，即它必须要和某个散射振幅建立联系，即我们需要指定“什么样的过程给出物理参数”。物理参数的具体定义、减除方案、抵消项这三者知道了一个就知道了剩下两个。指定这三者中的其中一个就称为指定了一个重整化方案。

我们现在有四个抵消项。但是，后面我们将看到，由于Ward恒等式，实际上只有三个——而且也只需要三个。

最小减除和修正的最小减除

所谓最小减除或者说MS指的是通过适当调整抵消项，让(5.65)，(5.70)和(5.76)这三个发散中的 $1/\epsilon$ 项全部被减除。此时考虑了抵消项的电子自能是

$$-i\Sigma = \text{---} \text{---} \text{---} + \text{---} \otimes \text{---},$$

即

$$\Sigma = \frac{\alpha(4m_R - \not{p})}{2\pi\epsilon} - (\not{p} \delta Z_2 - (\delta Z_2 + \delta Z_m) m_R) + \text{finite},$$

即

$$\delta Z_2 = -\frac{\alpha}{2\pi\epsilon}, \quad \delta Z_m = -\frac{3\alpha}{2\pi\epsilon}. \quad (5.92)$$

光子自能的做法类似，可以得到

$$\delta Z_3 = -\frac{2\alpha}{3\pi\epsilon}. \quad (5.93)$$

而顶角函数的最小减除则要求

$$\delta Z_1 = -\frac{\alpha}{2\pi\epsilon}. \quad (5.94)$$

因此MS方案要求的剪除项可以归结为：

$$\delta Z_1 = \delta Z_2 = -\frac{\alpha}{2\pi\epsilon} = -\frac{e_R^2}{8\pi^2\epsilon}, \quad \delta Z_3 = -\frac{e_R^2}{6\pi^2\epsilon}, \quad \delta Z_m = -\frac{3e_R^2}{8\pi^2\epsilon}. \quad (5.95)$$

虽然表面上有四个可以用来抵消发散的参数—— δZ_1 , δZ_2 , δZ_3 和 δZ_m ——但是实际上我们只用到了三个。当然，这是因为Ward恒等式要求

$$Z_1 = Z_2,$$

相应的

$$Z_e \sqrt{Z_3} = 1, \quad (5.96)$$

从而

$$\delta Z_e = \frac{e_R^2}{12\pi^2\epsilon}. \quad (5.97)$$

修正的最小减除或者说 $\overline{\text{MS}}$ 不仅减除 $1/\epsilon$ 项，还将 γ_E 和 $\ln 4\pi$ 项也减去，从而让所有的 $\tilde{\mu}$ 都变成 μ 。

这两种方案都会在取 $\epsilon \rightarrow 0$ 之后的结果中引入一个尚未确定的参数，MS会引入 $\tilde{\mu}$ 而 $\overline{\text{MS}}$ 会引入 μ 。这个参数是一个能量标度，或者也可以说是动量标度。

在壳重整化

在壳重整化的物理意义更加明显：它要求在物理参数附近不应该有任何自能修正，并且在某个动量尺度附近不应该有任何顶角修正，即要求如下重整化条件：

$$\Sigma(m_R) = 0, \quad \left. \frac{\partial \Sigma(p)}{\partial p} \right|_{p=m_R} = 0, \quad \Pi(0) = 0, \quad \Gamma^\mu(\mu) = \gamma^\mu. \quad (5.98)$$

μ 经常被选取为0，表示物理电荷是通过光子动量为零的过程确定的。

在壳重整化在更加现代的高能物理中使用得比较少，因为此时很难说“某个过程直接给出了物理参数”。例如，很多“粒子”实际上衰变非常厉害，从而很难直接通过关联函数的极点来定义“粒子的物理质量”，自然无所谓在壳。这种情况下我们实际上是在用MS或是 $\overline{\text{MS}}$ 来定义物理参数：定义使用这些减除方法计算得到的 m 和 e 就是物理质量和物理耦合常数。自然，此时没有什么办法能够直接“测”物理质量和物理耦合常数，但是我们总是可以先在某个实验中测量散射截面，用它反推出MS或是 $\overline{\text{MS}}$ 定义出来的物理参数，然后基于这些物理参数计算别的散射振幅，如果能够和这个别的散射振幅对应的实验结果对上，我们的理论就被验证了。

5.5.3 重整化群

有多种方式可以用来确定物理参数随能量标度的变化方式。

通过MS方案的抵消项计算重整化群流

本节通过抵消项计算重整化群流，具体来说是对 μ 求导。首先使用MS方案的抵消项。我们把之前消除发散时吸收进 e_R 的 μ 再拿出来，即

$$\mu^{\frac{4-d}{2}} e_R = \frac{1}{Z_e} e_0. \quad (5.99)$$

我们要写出一个含有 e_R 和 $\partial e_R / \partial \mu$ 的方程。裸参数和 μ 是一点关系都没有的，即

$$0 = \mu \frac{de_0}{d\mu} = \mu \frac{d}{d\mu} (Z_e \mu^{\epsilon/2} e_R) = \mu^{\epsilon/2} e_R Z_e \left(\frac{\epsilon}{2} + \frac{\mu}{e_R} \frac{de_R}{d\mu} + \frac{\mu}{Z_e} \frac{dZ_e}{d\mu} \right),$$

这其中的第一项最后会趋于零，第二项是我们需要的，于是可以拼凑出

$$\beta(e_R) = \mu \frac{de_R}{d\mu} = -\frac{\epsilon}{2} e_R - \frac{e_R}{Z_e} \mu \frac{dZ_e}{d\mu},$$

其中

$$\begin{aligned} \frac{\mu}{Z_e} \frac{dZ_e}{d\mu} &= \frac{\mu}{1 + \frac{e_R^2}{12\pi^2\epsilon}} \frac{1}{12\pi^2\epsilon} 2e_R \frac{de_R}{d\mu} \\ &= \left(1 - \frac{e_R^2}{12\pi^2\epsilon} + \dots \right) \frac{1}{12\pi^2\epsilon} 2e_R \mu \frac{de_R}{d\mu} \\ &= \left(1 - \frac{e_R^2}{12\pi^2\epsilon} + \dots \right) \frac{1}{12\pi^2\epsilon} 2e_R \left(-\frac{\epsilon}{2} e_R - \frac{e_R}{Z_e} \mu \frac{dZ_e}{d\mu} \right) \\ &= -\frac{e_R^2}{12\pi^2} + \mathcal{O}(e_R^4). \end{aligned}$$

因此我们就得到精确到 e_R^3 阶的 β 函数：

$$\beta(e_R) = \mu \frac{de_R}{d\mu} = -\frac{\epsilon}{2} e_R + \frac{e_R^3}{12\pi^2} + \mathcal{O}(e_R^4). \quad (5.100)$$

于是随着 μ 上升，我们就有如下参数跑动：

$$C - \frac{1}{2e_R^2} = \frac{\ln \mu}{12\pi^2},$$

或者，使用精细结构常数表示，有

$$e_R(\mu) = -\frac{3\pi}{2} \frac{1}{\ln \frac{\mu}{\Lambda_{\text{QED}}}}, \quad (5.101)$$

其中 Λ_{QED} 是一个能量标度。我们可以从中看到QED实际上不是一个紫外完备的理论：能标高到一定程度时，等效电荷将会发散。这个发散称为朗道极点。

在以上计算过程中可以看到几件事：首先，将原本吸收入 μ_R 中的 μ^ϵ 恢复出来是非常重要的，我们用它消去了一个发散；其次，我们可以用 μ 标记重整化群流，也可以用 e_R 标记重整化群流。实际上，我们可以用 μ 标记各个参数的跑动，这在参数空间中绘制出一条轨迹，可以通过实验探测获得这条轨迹，从而精密测定QED的准确性。最后一个应当注意的地方是，实际上影响 e_R 的只有 Z_3 ，因此使用真空极化的强度定义 e_R 其实也是可以的。

$$\frac{2}{\epsilon} \sim \ln \left(\frac{\Lambda^2}{\mu^2} \right) \quad (5.102)$$

从以上计算过程我们可以看出，计算 e 的 β 函数实际上就是要使用 Z_e 求导，而 Z_e 单纯由系统中各个参量的紫外发散确定。因此，实际上无需对所有发散图做完整的重整化计算，只要分别提取出它们的紫外发散，就能够计算 β 函数了。

我们下面计算质量的重整化群流。

第六章 杨-米尔斯理论

6.1 经典杨-米尔斯理论

我们将经常用到李代数。可以采用下面的约定：

$$[T^a, T^b] = if^{abc}T^c. \quad (6.1)$$

对上式求迹，等式左边是零，从而等式右边也是零，即

$$\text{tr } T^a = 0. \quad (6.2)$$

$$\text{tr } (T^a T^b) = \frac{1}{2} \delta^{ab}, \quad (6.3)$$

对 $SU(N)$ 群及其李代数，我们有

$$T_{ij}^a T_{jk}^a = \frac{N^2 - 1}{2N} \delta_{ik}, \quad f^{acd} f^{bcd} = N \delta^{ab}. \quad (6.4)$$

6.1.1 杨-米尔斯理论的拉氏量

规范场的引入和协变导数

电动力学的Maxwell理论是一个比较简单的规范理论，其中旋量场可以做任意的局域 $U(1)$ 变换

$$\Omega(x) = e^{ia(x)},$$

为了让拉氏量在此变换下保持不变，一个额外的矢量场被耦合到旋量场上，当旋量场做 $U(1)$ 变换时矢量场的场值发生一个平移，从而拉氏量在局域 $U(1)$ 变换下不变。 $U(1)$ 群的特性决定了系统中存在矢量场这一事实，并且决定了相互作用的形式。以一种系统性的方式决定一个理论中应该有什么场，以及相互作用应该取什么形式，显然是非常有吸引力的。因此，非常自然地，我们希望用一个更加复杂的李群做规范变换，并且开发一套看着一个李群就能够写下一个具有规范对称性的理论的方式。对 $U(1)$ 规范理论的推广看起来似乎有非常多可能的选择，但所幸我们已经有一个成熟的、和微分几何紧密相关的、已经在描述强相互作用和弱相互作用的方面大获成功的理论框架：**杨-米尔斯理论**。

下面我们将用一种启发式的方法去导出杨-米尔斯理论，主要是通过模仿电动力学中的概念。我们只讨论紧致的李群 G ，此时其一定具有么正表示，这正是我们想要的。对标量场，看起来唯一能够作用在其上的操作就是乘以一个复因子。因此如果我们想要让 G 有一个 n 维么正矩阵表示就需要引入 n 个标量场。（ n 和李代数维数没有必然关系）此时，自由理论形如

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi^\dagger \partial^\mu \phi - \frac{1}{2} m^2 \phi^2,$$

其中 ϕ 是 n 个标量场排成的一个列矢量。我们用 t^a 标记 G 的 n 维么正表示的李代数成员。如果我们想让拉氏量规范不变，即在变换

$$\phi(x) \rightarrow \Omega(x)\phi(x), \quad \phi^\dagger(x) \rightarrow \phi^\dagger(x)\Omega^\dagger(x), \quad \Omega(x) = e^{-ig\theta_a(x)t^a},$$

下不变，那么无需调整质量项，因为显然

$$\phi^\dagger \Omega^\dagger \Omega \phi = \phi^2.$$

含有导数的项则需要修正，具体来说，我们需要找到某种协变导数，使得

$$D_\mu(\Omega(x)\phi(x)) = \Omega(x)D_\mu\phi(x). \quad (6.5)$$

∂_μ 肯定不满足这个条件，因为

$$\partial_\mu(\Omega(x)\phi(x)) = \Omega(x)\partial_\mu\phi + (\partial_\mu\Omega)\phi.$$

对狄拉克旋量，表面上看如果有 n 个旋量场， G 可以有 $4n$ 维表示，但是这实际上是行不通的，因为旋量场的拉氏量为

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi,$$

如果 Ω 是 $4n$ 维的，那么可以让 Ω 作用到旋量内部的各个分量上。然而，此时 $\bar{\psi}$ 的变换方式为

$$\bar{\psi} \rightarrow \psi^\dagger \Omega^\dagger \gamma^0,$$

没有什么能够保证 γ 和 Ω^\dagger 一定对易，但是 γ 和 Ω^\dagger 最好是对易的，否则简单地接受(6.5)并不能让拉氏量在规范变换下不变。一种比较方便的做法是在将 Ω 作用于旋量场上的时候将旋量（以及各个 γ 矩阵）看成一个整体，不允许对其分量单独进行操作，于是 Ω 应该是 n 维的，并且因为 γ^0 此时相当于一个标量，它和 Ω 肯定是对易的。同样，所有 n 个旋量场的质量必须一样，否则 m 将成为一个任意的对角矩阵，而未必和 Ω 对易，那么 $\bar{\psi}m\psi$ 就不是不变的了。接受这个做法之后，在旋量场的情况下同样只需要设法找到某种协变导数使得(6.5)成立即可。

现在我们分析协变导数的具体形式。在杨-米尔斯理论中，我们模仿电动力学，直接引入矢量场 A_μ 并要求

$$D_\mu = \partial_\mu - igA_\mu, \quad (6.6)$$

并以此为依据决定 A_μ 在规范变换下如何变动，即让 A_μ “吸收掉”多余的 $(\partial_\mu\Omega)\phi$ 项。这种协变导数的形式和微分几何中的协变导数非常一致， A 就是一个联络（我们将在第6.1.2节中看到它确实就是几何上那种“平移时矢量分量跟着转”的联络），即所谓规范联络或者说规范场。显然应有

$$\partial_\mu - igA'_\mu = \Omega(x)(\partial_\mu - igA_\mu)\Omega^{-1}(x),$$

容易看出 A_μ 的变换规则应为

$$A_\mu(x) \rightarrow \Omega(x)A_\mu(x)\Omega^{-1}(x) + \frac{i}{g}\Omega(\partial_\mu\Omega^{-1}(x)). \quad (6.7)$$

请注意 $\Omega(x)$ 是 n 维矩阵，因此为了避免得到平庸的结果，实际上我们也需要让 A_μ 变成一个 n 维矩阵，也就是除了时空指标 $\mu = 0, 1, 2, 3$ 以外还需要让 A 带两个从1跑到 n 的矩阵指标。在杨-米尔斯理论中我们实际上会将每个时空点、每个时空分量上都是 n 维矢量的 A 场限制为李代数 $\{t^a\}$ 的成员，因为 A_μ 的变换规则的无穷小版本为

$$\delta A_\mu = ig\theta_a[A_\mu, t^a] - t^a\partial_\mu\theta_a,$$

因此如果我们要求 $A_\mu(x)$ 是李代数 $\{t^a\}$ 的成员那么变换之后它还是李代数的成员。这样， A 可以用三个标签标记，一个是时空点 x ，一个是矢量指标 μ ，还有一个是规范指标 a ，写成

$$A_\mu(x) = A_\mu^a(x)t^a. \quad (6.8)$$

规范场和旋量场（或标量场）现在都带上了一个规范指标 a ，其于是就有

$$\begin{aligned} \delta A_\mu &= ig\theta_a[A_\mu^b t^b, t^a] - t^a \partial_\mu \theta_a \\ &= -t^a D_\mu \theta_a, \end{aligned} \quad (6.9)$$

其中

$$D_\mu \theta^a = \partial_\mu \theta^a + gf^{abc} A_\mu^b \theta^c. \quad (6.10)$$

这里需要解释一下记号 D_μ 。我们之前定义的 D_μ 作用在 G 的一个 n 维表示上；然而，规范场 A_μ^a 是将 A_μ 以 t^a 为基底展开得到的分量。作用在 A_μ 上的变换并不是 n 维表示空间上的线性变换，而是李代数的伴随表示的表示空间上的线性变换。此处的协变导数 D_μ 在后者中而不在前者中，它和作用在 ψ 或是 ϕ 上的 D_μ 的具体形式是不同的（虽然都代表“协变的平移”）。此外 $D_\mu \theta_a$ 实际上是 $(D_\mu \theta)_a$ ，即它会将 θ_a 的各个分量混合起来，正如微分几何中的情况那样，即我们有 $D_\mu \theta_a = D_{\mu \ ac} \theta_c$ 。

然后我们就会发现，如果李群是非阿贝尔的，那么 A_μ 的无穷小变换不仅仅是场值做一个平移，还需要加上一个对易子。电动力学仅讨论 $U(1)$ 变换，属于阿贝尔规范理论，杨-米尔斯理论则是非阿贝尔规范场论。今后我们称 A_μ 为**规范场**，而称 ψ 或是 ϕ 为**物质场**，因为和电动力学中的图像类似，似乎“物质场通过携带动量的规范玻色子发生相互作用”。当然，规范玻色子——如光子——其实也是一种物质，所以这个说法有不准确之处。

场强张量

规范场可以有它自己的动能项。在杨-米尔斯理论中，这个动能项大体上仍然应该和电动力学一致，即大体上仍有

$$\mathcal{L}_A = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

成立。在电动力学中我们有

$$[D_\mu, D_\nu] = ieF_{\mu\nu},$$

而在杨-米尔斯理论中， $[D_\mu, D_\nu]$ 在规范变换下为

$$[D_\mu, D_\nu] \rightarrow \Omega [D_\mu, D_\nu] \Omega^{-1},$$

因此可以定义

$$[D_\mu, D_\nu] = -igF_{\mu\nu} = -ig(\partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]), \quad (6.11)$$

作为电动力学中的场强张量的推广。由于 A 实际上是 n 维矩阵， $F_{\mu\nu} F^{\mu\nu}$ 也是 n 维矩阵，因此我们还需要加上一个求迹操作就能够得到规范不变而同时洛伦兹不变的拉氏量：

$$\mathcal{L}_A = -\frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) = -\frac{1}{4} F_{\mu\nu}^a F^{a \mu\nu}, \quad (6.12)$$

如果 G 是 $U(1)$ ，那么上式就自动退化为了电动力学。

可以看到在这种思路下面规范场本身是不能有质量的，因为质量项 $m^2 A_\mu A^\mu$ 无论如何没法变得规范不变。但是，通过希格斯机制，实际上可以给规范场引入一个等效的质量。本节暂时不讨论这些内容。

因此我们现在就得到了杨-米尔斯理论的拉氏量：如果规范场和旋量场耦合，那么拉氏量就是

$$\mathcal{L} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}. \quad (6.13)$$

$F_{\mu\nu}F^{\mu\nu}$ 项前面的系数本来可以有变化，但是我们完全可以将其吸收到 A 中，然后用调整 g 来保持协变导数不变。这个拉氏量看起来和电动力学基本上一样，但是因为 F 中的非线性部分，其经典行为实际上就非常有趣。

前面已经说明过，所有旋量场的质量都是一样的，并且从(6.13)也可以看出，规范玻色子没有质量。因此表面上，杨-米尔斯理论是非常局限的。但实际上并不是这样：通过希格斯机制是可以引入质量的。

我们之后将使用 a 来区分李代数的不同表示，而用指标 i 来区分 ψ 中含有的不同旋量场。 a 的取值范围是李代数的生成元个数，而 i 的取值范围是李代数的表示空间的维数。

关于李代数的限制

6.1.2 对称性和守恒量

平移

规范场的平移操作值得特别讨论。简单地令

$$x \longrightarrow x + \delta x, \quad \psi'(x') = \psi(x), \quad A'(x') = A(x) \quad (6.14)$$

的确可以保持拉氏量不变，也的确能够据此计算出一个诺特守恒流，但是这样的诺特守恒流不是规范不变的。这就是说，应该从中“删去”一些虽然平移不变，但是并非规范不变的东西，才能够得到定义良好的能量-动量张量。我们知道使用(6.14)计算能动张量实际上同时做了一个从拉氏量到哈密顿量的切换，其中我们取

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}$$

为正则动量，所以其实(6.14)得不到有意义的结果是非常正确的——规范场论本身含有冗余自由度，不能指望此时朴素的勒让德变换仍然成立。还可以从另一个角度出发看这个问题：变换(6.14)和规范变换不对易，其生成元自然和规范变换也不对易。

注意到，协变导数 D_μ 是和规范变换对易的，并且它的确代表某种平移，因此我们尝试以它做无穷小变换，或者等价地说，先做一个纯粹的平移变换，再做一个规范变换，即取

$$\begin{aligned} x^\mu &\longrightarrow x^\mu + \delta x^\mu, \\ \phi(x) &\longrightarrow \phi'(x') = \exp(i g \delta x^\mu A_\mu) \phi(x), \\ A_\mu(x) &\longrightarrow A'_\mu(x') = \exp(i g \delta x^\nu A_\nu) A_\mu(x) \exp(-i g \delta x^\rho A_\rho) \\ &\quad + \frac{i}{g} \exp(i g \delta x^\rho A_\rho) \partial_\mu \exp(-i g \delta x^\nu A_\nu), \end{aligned} \quad (6.15)$$

计算得到

$$\begin{aligned} \delta \phi &= -\delta x^\mu D_\mu \phi, \\ \delta A_\mu &= -\delta x^\nu F_{\nu\mu}, \end{aligned} \quad (6.16)$$

于是守恒流就是

规范对称性和规范荷

费米场和每一种规范玻色子之间的相互作用项都有一个不同的规范荷，这个规范荷由规范群的表示给出。

6.1.3 Wick转动

在做完Wick转动

6.1.4 微分几何的观点

6.2 杨-米尔斯理论的量子化

6.2.1 Faddeev–Popov量子化

规范固定和鬼场

规范对称性会导致正则量子化变得比较困难，因为需要做复杂的规范选取来消除多余的自由度，而在路径积分量子化中则可以通过Faddeev–Popov量子化比较容易地解决。我们已经在自由无质量矢量场的量子化中使用过了Faddeev–Popov量子化，这回我们如法炮制。和自由无质量矢量场的情况不同，此时不仅需要引入规范固定项，还需要引入鬼场。

我们通过在 $\int \mathcal{D}A_\mu$ 之后插入

$$1 = \int \mathcal{D}\alpha \delta(G(A^\alpha)) \det\left(\frac{\delta G(A^\alpha)}{\delta \alpha}\right)$$

来设法将对只相差一个规范变换的场重复计数导致的因子提取出来，其中 α 标记规范变换的参数，它带有一个规范指标；规范固定为 $G(A) = 0$ ， G 定义为洛伦兹协变的

$$G(A^\alpha) = \partial^\mu (A^\alpha)_\mu - \omega(x), \quad (6.17)$$

其中 $\omega(x)$ 是任意的标量场。我们让 α 取无穷小量，则有

$$G(A^\alpha) = \partial^\mu A_\mu + \frac{1}{g} \partial^\mu D_\mu \alpha^a - \omega(x),$$

于是

$$\begin{aligned} Z &= \int \mathcal{D}A \mathcal{D}\psi e^{iS[A, \psi]} \\ &= \int \mathcal{D}A \mathcal{D}\psi \int \mathcal{D}\alpha \delta(G(A^\alpha)) \det\left(\frac{\delta G(A^\alpha)}{\delta \alpha}\right) e^{iS[A, \psi]} \\ &= \frac{1}{g} \det(\partial^\mu D_\mu) \int \mathcal{D}A \int \mathcal{D}\psi \int \mathcal{D}\alpha \delta(G(A^\alpha)) e^{iS[A, \psi]} \\ &= \frac{1}{g} \det(\partial^\mu D_\mu) \int \mathcal{D}\psi \int \mathcal{D}\alpha \int \mathcal{D}A^\alpha \delta(G(A^\alpha)) e^{iS[A^\alpha, \psi]} \\ &= \frac{1}{g} \det(\partial^\mu D_\mu) \int \mathcal{D}\psi \int \mathcal{D}\alpha \int \mathcal{D}A \delta(G(A)) e^{iS[A, \psi]}, \end{aligned}$$

倒数第二个等号是因为 $\mathcal{D}A$ 和 $\mathcal{D}A^\alpha$ 相同，倒数第一个等号是我们重新标记了场。既然 $\omega(x)$ 可以任意取值，我们不妨重新定义配分函数，去掉无用的因子 g ，并对所有的 $\omega(x)$ 求和，得到

$$\begin{aligned} Z &= \det(\partial^\mu D_\mu) \int \mathcal{D}\omega e^{-i \int d^4x \frac{\omega^2}{2\xi}} \int \mathcal{D}A \int \mathcal{D}\psi \int \mathcal{D}\alpha \delta(\partial^\mu A_\mu - \omega(x)) e^{iS[A, \psi]} \\ &= \det(\partial^\mu D_\mu) \int \mathcal{D}\alpha \int \mathcal{D}A \int \mathcal{D}\psi \exp\left(-i \int d^4x \frac{(\partial^\mu A_\mu)^2}{2\xi}\right) e^{iS[A, \psi]}. \end{aligned}$$

无用的 $\int \mathcal{D}\alpha$ 因子可以略去。与自由场的情况不同,此时因子 $\det(\partial^\mu D_\mu)$ 中仍然含有场变量,不能直接丢弃。为此,可以引入一个鬼场 c ,它是一个复标量场,但是是格拉斯曼数,这样就能够满足

$$\int \mathcal{D}c \int \mathcal{D}\bar{c} \exp\left(i \int d^4x \bar{c}(-\partial^\mu D_\mu)c\right) = \det(\partial^\mu D_\mu),$$

如果 c 是普通标量场,那么行列式会出现在分母上。显然 c 并没有什么物理意义,在计算协变的物理量时也没有外线。

现在我们就完成了规范场论(6.13)的量子化:只需要用等效的拉氏量

$$\mathcal{L} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} - \underbrace{\frac{(\partial^\mu A_\mu^a)^2}{2\xi}}_{\text{gauge fixing}} + \underbrace{\bar{c}(-\partial^\mu D_\mu)c}_{\text{ghost}} \quad (6.18)$$

做路径积分即可。注意规范场和旋量场都有 n 个,因此协变导数 D_μ 是一个 n 维矩阵,鬼场 c 也是 n 维的,多出来一个规范指标标记这 n 个维度。

传播子和顶角

关于费米子交换加负号这件事:规范理论中相互作用项中费米场均呈现为 $\bar{\psi}\psi$ 形式,若干个相互作用项的乘积形如 $\bar{\psi}\psi \cdots \bar{\psi}\psi$ 。容易验证,这样一个算符序列和外部的算符缩并,缩并线的交叉次数一定是偶数,从而不会因为费米子算符的交换而产生负号。因此,负号只应该存在于这样一个算符序列内部的缩并,即出现在费米子线形成一个圈的时候。而容易验证,此时负号仅仅存在于 $\langle \bar{\psi}\psi \rangle$ 一个因子中。因此,费米子交换加负号这件事在规范场论的费曼图中体现为:但凡费米子线形成了一个圈,加负号,否则什么都不做。在单条费米子线形成一个圈时可以直接把这个圈算出来,此时负号已经体现在这个圈的值当中了,因此无需做任何额外处理。

从规范场论演生出的低能有效理论——如库伦相互作用——虽然不再有规范玻色子传播子,但是顶角的形式仍然保持不变,因此费米子交换加负号这一特点同样可以通过“但凡费米子线形成了圈,加负号”完全描述。

6.2.2 BRST对称性

引入鬼场之后,有效作用量就不再是规范不变的了,因为规范冗余性已经消除,规范不变性被规范固定项和鬼场去掉了。不过,BRST四人发现,之前定义的规范变换如果补充上一个鬼场的变换,能够有一个整体的对称性。这个对称性显然不是通常意义上的整体对称性,因为鬼场等都是非物理的,从而,毫不意外的,这个对称性对应的守恒荷是一个格拉斯曼数,其平方为零。这种变换——**BRST变换**——可以很好地描述规范场的拓扑性质。通过BRST变换还可以得到Ward-Takahashi恒等式和Slavnov-Taylor恒等式。

首先,我们引入一个玻色辅助场 B ——实际上是一系列玻色辅助场 B^a ,带有一个规范指标,其总数和李代数的维数一致,和 n 没有特别直接的关系——让规范固定项消失,得到

$$\mathcal{L} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \bar{c}^a(-\partial^\mu D_\mu^{ac})c^c + \frac{\xi}{2}(B^a)^2 + B^a \partial^\mu A_\mu^a. \quad (6.19)$$

现在考虑如下无穷小变换：

$$\begin{aligned}
 \delta\psi &= i g \epsilon c^a t^a \psi, \\
 \delta A_\mu^a &= \epsilon D_\mu^{ac} c^c, \\
 \delta c^a &= -\frac{1}{2} g \epsilon f^{abc} c^b c^c, \\
 \delta \bar{c}^a &= \epsilon B^a, \\
 \delta B^a &= 0,
 \end{aligned} \tag{6.20}$$

整个拉氏量(6.19)在变换(6.20)之下完全就是不变的。首先， A_μ 和 ψ 的变换就是以 ϵc^c 为位移的规范变换，因此(6.19)的头两项不变。第四项也不变，因为 B 根本就没有发生任何变化。很容易注意到

$$\delta B^a \partial^\mu A_\mu^a = \epsilon B^a \partial^\mu D_\mu^{ac} c^c = -(\delta \bar{c}^a)(-\partial^\mu D_\mu^{ac}) c^c,$$

因此只需要检验

$$\delta(D_\mu^{ac} c^c) = 0$$

即可。将上式展开，

$$\begin{aligned}
 \delta(D_\mu^{ac} c^c) &= D_\mu^{ac} \left(-\frac{1}{2} g \epsilon f^{cbd} c^b c^d \right) + g f^{abc} (\delta A_\mu^b) c^c \\
 &= -\frac{1}{2} g \epsilon \partial_\mu (f^{aed} c^e c^d) - \frac{1}{2} g^2 \epsilon f^{abc} f^{ced} A_\mu^b c^e c^d + \epsilon g f^{abc} (\partial_\mu c^b + g f^{bed} A_\mu^e c^d) c^c,
 \end{aligned}$$

展开第一项并使用鬼场的反交换性可以得到

$$-\frac{1}{2} g \epsilon \partial_\mu (f^{aed} c^e c^d) = g \epsilon f^{ade} c^e \partial_\mu c^d,$$

于是

$$\delta(D_\mu^{ac} c^c) = -\frac{1}{2} g^2 \epsilon f^{abc} f^{ced} A_\mu^b c^e c^d + \epsilon g^2 f^{abc} f^{bed} A_\mu^e c^d c^c.$$

我们将上式右边第二项写得更加对称一些（又一次用到了鬼场的反对称性）：

$$\begin{aligned}
 \delta(D_\mu^{ac} c^c) &= -\frac{1}{2} g^2 \epsilon f^{abc} f^{ced} A_\mu^b c^e c^d + \epsilon g^2 f^{abc} f^{bed} A_\mu^e c^d c^c \\
 &= -\frac{1}{2} g^2 \epsilon f^{abc} f^{ced} (A_\mu^b c^e c^d + A_\mu^e c^d c^b + A_\mu^d c^b c^e),
 \end{aligned}$$

重新排列指标，得到

$$\delta(D_\mu^{ac} c^c) = -\frac{1}{2} g^2 \epsilon (f^{abc} f^{ced} + f^{adc} f^{cbe} + f^{aec} f^{cdb}) A_\mu^b c^e c^d.$$

由雅可比恒等式发现上式为零。这就表明(6.19)在变换(6.20)之下不变。这称为**BRST**对称性。

设BRST变换的无穷小生成元为 Q ，则由于(6.20)中的每一条变换都正比于鬼场而鬼场是反对易的，我们立刻发现 $Q^2 = 0$ 。如果我们对(6.19)做正则量子化，那么它一定有一个幂零的守恒荷 Q 。一个幂零算符会给出希尔伯特空间的如下分割：

- 将被 Q 作用后不为零的那些态组成的子空间记作 \mathcal{H}_1 ；
- 将 Q 作用在 \mathcal{H}_1 上得到的子空间记作 \mathcal{H}_2 ；
- 将除此以外的部分记作 \mathcal{H}_0 。

由于 Q 和 H 对易，任何一个本征态被作用了 Q 之后得到的还是本征态。 \mathcal{H}_1 中的本征态被作用了 Q 之后就得到 \mathcal{H}_2 中的本征态。 \mathcal{H}_0 中的本征态被作用了 Q 之后得到的本征态还是在 \mathcal{H}_0 中。也即， \mathcal{H}_0 中有一个二重简并， \mathcal{H}_1 和 \mathcal{H}_2 放在一起构成二重简并。

为了获得定义良好的单粒子态，我们暂时令 $g = 0$ ，此时会发现，在 Q 的作用下反鬼场 \bar{c} 变为规范玻色子（注意根据(6.19)， B^a 和 A_μ^a 满足的算符方程是线性的，它们实际上对应同一种激发），

还有一个微妙的细节需要说明：BRST变换保持路径积分的积分测度不变，即这里没有量子反常。这件事的证明如下。首先，不难验证

$$\frac{\partial B^{a'}}{\partial B^a} = \frac{\partial \bar{c}^{a'}}{\partial \bar{c}^a} = \frac{\partial c^{a'}}{\partial c^a} = 0,$$

因此只需要计算 ψ 和 A 两组变量的雅可比行列式即可。我们有

$$\frac{\partial \psi'}{\partial \psi} = ig\epsilon c^a t^a, \quad \frac{\partial \bar{\psi}'}{\partial \bar{\psi}} = -ig\epsilon c^a t^a,$$

以及

$$\frac{\partial A_\mu^{a'}}{\partial A_\mu^b} = gf^{abc}c^c.$$

路径积分测度的变换因子为

$$\frac{\det(\partial A'/\partial A)}{\det(\partial \psi'/\partial \psi) \det(\partial \bar{\psi}'/\partial \bar{\psi})} = \frac{\prod_\mu (1 + gf^{abc}c^c)}{(\det(1 - ig\epsilon c^a t^a))(\det(1 + ig\epsilon c^a t^a))} = 1 + \mathcal{O}(\epsilon^2),$$

因此BRST变换下路径积分测度的确是几乎不变的。

6.2.3 微扰计算

微扰计算使用费曼图，费曼图中我们要对内线上的动量做四重积分。

6.3 正规化和重整化

本节将花大量篇幅分析微扰计算时如何做重整化。乍一看，基于费曼图的分析是不必要的，因为重整化群流具有紫外不动点这件事足够保证一个量子场论的可重整性。但是应当注意，重整化群流的计算本身常用微扰论，一阶或是二阶微扰论给出的重整化群流是不是可靠是不好说的；此外，一个理论可重整化和它的每个费曼图都能够消除所有发散并不是一回事——可能存在这样的情况，使得需要重求和一些图才能够消除所有发散。因此，基于费曼图和抵消项来分析理论的可重整性还是必要的。

6.3.1 圈图发散及其来源

树图计算通常不会发散。如果我们要考虑更高阶的微扰，即要计算圈图，则通常会导致发散，因为我们需要对一个传播子做四维动量积分，得到一个类似于

$$\int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2 - m^2} \sim \int dk k$$

的式子，就出现了发散，具体来说是紫外发散。红外发散可以通过给所有场加上一个小的质量而轻松解决，计算完成之后再让此质量趋于零，或者也可以通过重求和含有能量可以任意低、数量可以任意多的出射和入射无质量玻色子来解决，无需用正式的重整化消除它们。不是所有的圈图都发散，一个例子是梯形图，其中确实有一个圈，但是没有发散。因此，我们需要先判断什么图会发散。

用表观发散度做估计

从(6.13)可以看出, 各个物理量的量纲满足

$$[d^4x] + [F^2] = 0, \quad [F] = [\partial A] = 1 + [A],$$

于是

$$[A] = 1 = [\partial].$$

另一方面, 根据协变导数的定义, ∂ 和 gA 的量纲一致, 因此只有一种可能: g 根本就没有量纲。

量纲分析对表观发散度计算很重要。我们知道一张图的表观发散度是

$$D(\Gamma) = \sum_i n_i \left(d_i + b_i + \frac{3}{2} f_i - 4 \right) + 4 - E_B - \frac{3}{2} E_F, \quad (6.21)$$

其中 d_i , b_i 和 f_i 分别表示 i 类型顶角的动量幂次, 玻色子线数目(鬼场也算玻色子, 因为其动能项形式和玻色子一致; 量纲分析和对易还是反对易无关)和费米子线数目。玻色场的量纲是1(前面刚刚证明过), 费米场的量纲是3/2, 而如前所述耦合常数无量纲, 因此对来自协变导数的规范场-旋量场耦合项, 有

$$[d^4x] + d_i + b_i + \frac{3}{2} f_i = 0,$$

即

$$d_i + b_i + \frac{3}{2} f_i - 4 = 0.$$

其它的顶角来自 $F_{\mu\nu}F^{\mu\nu}$ 项, 其中的三玻色子顶角的耦合常数 $\sim g$ 而四玻色子顶角的耦合常数 $\sim g^2$, 因此同样耦合常数量纲为零, 同样有上式成立。这就意味着, 在杨-米尔斯理论中, 我们有

$$D(\Gamma) = 4 - E_B - \frac{3}{2} E_F. \quad (6.22)$$

考虑让 $D(\Gamma) \geq 0$ 的情况, 会发现以下几种图按照表观发散度的估计会发散:

- 四条玻色子外线, 表观发散度为0;
- 三条玻色子外线, 表观发散度为1;
- 两条玻色子外线, 表观发散度为2;
- 一条玻色子外线, 两条费米子外线, 表观发散度为0;
- 两条费米子外线, 表观发散度为1。

还有一些图, 如一条费米子外线和两条玻色子外线的图, 如果存在, 也会发散, 但是杨-米尔斯理论的顶角形式意味着这样的图是不可能出现的。

发散子图

表观发散度不能立刻用于判断一个图是不是发散。一个简单的费米子或是玻色子线按照上一节的说法可能会发散, 但是它们显然没有发散。不过, 这比较容易排除: 我们只需要只讨论amputated diagram即可, 此时表观发散度为零的图一定涉及对数发散。表观发散度为负数的图似乎不应该发散, 但是实际上也可能会发散, 例如我们在梯形图的任意一个

电子线上加一个自能修正，整张图立刻就发散了。换言之，一个表观发散度小于零的图仍然可能发散，此时发散来自其子图。

一些图如果切断任意一根内线（即：不对它求积分），那么就不发散。这样的图称为原始发散图。原始发散图的发散程度显然由表观发散度给定。如果一个图含有原始发散图那么它肯定发散。但是，在引入一个抵消项消除原始发散图之后，它可能还是发散，因为此时的那个原始发散子图虽然被重整化了，变成了一个内部结构无需讨论的点，但是留下来的骨架仍然可能含有另一个原始发散图。我们可以继续适当调整抵消项来让这个剩下来的原始发散图也被消除掉。但是，现在又有一个新的问题：一张图中可能有两个原始发散图，且它们共享一个传播子。此时的积分往往会给出一些难以处理的含有 $\log p$ 之类的因子的发散。正比于 p^n 的多项式型发散（称为局域发散，因为它在实空间是非常定域的）比较容易用抵消项消除，因为各个场的自能修正会给出 p^2 项，拼凑一下就可以得到一个多项式。正比于 $\log p$ 之类的因子的发散（称为非局域发散，因为它在动量空间是长程的）则无法消除，因为这要求形如 $\log \partial_\mu$ 的抵消项。

因此用抵消项减除发散是非常非平凡的。第6.3.3节给出了这么做的一种系统做法。

6.3.2 传统方法：维数正规化和抵消项

一种可行的重整化方法是，首先可以以树图为骨架而对传播子做自能修正而把顶角修正为顶角函数。这样，所有圈图修正都可以归入自能修正和顶角函数修正之一。我们要做的就是先做正规化，赋予发散的圈图积分一个定义，然后引入系数未知的抵消项，最后使用重整化条件引入物理参数，计算得到所有的东西。

使用这种办法不容易分析一个理论是否在每一阶都是可重整的。第6.3.3节给出了一种更加系统的方法。

重整化条件为

- 在修正后的单粒子格林函数的极点处，有质量粒子的自能修正对 p^2 （玻色子）或者 \not{p} （费米子）的一阶导数为零；无质量粒子的自能修正为零。这是为了确保没有场强重整化。动量远离单粒子格林函数极点时它们当然可以不是零；本应如此，否则圈图修正无法体现。
- 在修正后的单粒子格林函数的极点处，极点给出的质量（通过 $p^2 = m^2$ 解出）就是我们设定的物理质量；在显式引入自能修正时，有质量粒子的自能修正为零。这是为了确保质量的修正为零。
- 顶角函数和物理相互作用强度相同，这是为了确保顶角修正为零。具体什么是“物理相互作用强度”取决于探测方式，如量子电动力学中通常是使用静电学方法测定电磁相互作用的强度，于是我们要求顶角函数在光子动量为零时和静电学方法测得的电磁相互作用强度（其实就是元电荷）相同。当然，也可以在顶角函数在光子动量为一个有限非零值的时候做计算。这件事实际上是非常非平凡的，因为在不同能标下耦合常数实际上会有小的跑动。详见第6.3.4节。

6.3.3 BPHZ重整化

BPHZ重整化是一种系统的分析基于费曼图的可重整性的方法。前面看到，系统地做任意阶图的发散减除的主要问题在于交缠发散。BPHZ重整化中交缠发散可以自动地被消去。

下面我们引入一些记号： Γ 等表示一张费曼图，直接计算它，会得到

$$F_\Gamma = \int \prod_i d^4 k_i I_\Gamma, \quad (6.23)$$

I_Γ 是被积函数，整个积分发散。 I_Γ 具有这样的一般形式：

$$I_\Gamma = \prod_{ab} \Delta_{ab} \prod_c P_c, \quad (6.24)$$

其中 Δ 和 P 分别代表传播子和顶角的值，下标 a, b, c 代表顶角的编号。我们设

$$J_\Gamma = \int \prod_i d^4 k_i R_\Gamma \quad (6.25)$$

为 F_Γ 的有限部分。用 \bar{R}_Γ 表示 Γ 给出的，所有子图的发散均已经减除（从而子图可以用一个内部结构无需考虑的点代替）的被积函数，这个被积函数做了内线积分之后仍然可能发散。用 $t^\Gamma \Gamma$ 表示去除 Γ 的骨架（即其所有发散子图都被用一个点代替之后留下来东西）的发散的方法。显然如果 Γ 已经是原始发散图了，那么 $\bar{R}_\Gamma = R_\Gamma$ ，否则还需要做最后一步减除 $R_\Gamma = t^\Gamma \bar{R}_\Gamma$ 。

我们还将用 p 表示外动量，而用 k 表示内线动量。

大体上说在BPHZ重整化中我们要做这么几件事：

1. 首先，给出一种机械的，基于表观发散度的，拿到一张图立刻提取出其不发散的部分和发散的部分的方法，无论是整体的发散还是子图的分发。在这一步我们暂时不考虑抵消项的具体形式，也无需做传统意义上的带一个参数（ Λ 或者 ϵ ）正规化。具体来说，我们需要：
 - (a) 对每个原始发散图，指定一种方式减除其发散部分，即对每个原始发散图 Γ 都指定 $t^\Gamma \Gamma$ 。
 - (b) 对非原始发散图，指定一种方式系统地减除其所有子图的分发，这么做了之后，再减除其最外层骨架的分发。
 - (c) 说明交缠发散能够被消除。
2. 第一步中得到了一系列发散部分。这些发散部分的形式五花八门，但是还是可以分类的：一个 \bar{R}_γ 如果还是发散，那么它只能是一个原始发散图。因此每一个 $t^\gamma \bar{R}_\gamma$ 实际上都相当于用某个抵消项抵消了和它具有同样的外线的一个原始发散图产生的发散。

很显然如果理论中有一些原始发散图给出的发散需要形如 $\log \partial_\mu$ 之类的抵消项，或者原始发散图有无数多个，等等，那么这一步就是不现实的。因此这就给出了一种简便的方式，来看着一个理论的拉氏量，判断它是否可重整。
3. 引入重整化条件。如果我们只是要证明理论可以重整化，这一步可以省去。

原始发散图的去发散 对原始发散图 Γ ，设有 E 条外线，则有 $E - 1$ 个独立的外线动量。一个直截了当的去发散方案是在 $p = 0$ （其实不在 $p = 0$ 附近展开也是可以的，这基本上就是“在不同能标附近做计算”——见第6.3.4节）附近泰勒展开到 $D(\Gamma)$ 阶：

$$t^\Gamma I_\Gamma(p_1, p_2, \dots, p_E) = f(0, \dots, 0) + \dots + \frac{1}{D(\Gamma)!} \sum_{i_1, \dots, i_{D(\Gamma)}=1}^{E-1} p_{i_1} \cdots p_{i_{D(\Gamma)}} \frac{\partial^{D(\Gamma)} I_\Gamma}{\partial p_{i_1} \cdots \partial p_{i_{D(\Gamma)}}}. \quad (6.26)$$

也就是说我们丢弃 $D(\Gamma)$ 阶及以下的所有项。稍微考虑一下会发现这是合理的选择：(6.26)囊括了所有的发散， $(1 - t^\Gamma)I_\Gamma$ 应当是收敛的。一个发散的内线积分的积分变量或者和外线没有关系，或者是外线动量的某个线性组合（因为动量守恒关系）。如果是前者，它导致的发散和外线没有关系，那就会被收集到 $f(0, \dots, 0)$ 当中，于是就被减除了。如果是后者，那么对 p 求导实际上就是对 k 求导（因为 p 和 k 总是加在一起），那么 I_Γ 的一阶到 $D(\Gamma)$ 阶中， k 的次数加上4是大于等于零的（依照表观发散度即可看出这一点），更高阶项中， k 的次数加上4是小于零的。因此 I_Γ 的一阶到 $D(\Gamma)$ 阶做了内线积分会发散而其它阶则会收敛。

非原始发散图的去发散 消除原始发散图的发散时引入的那些抵消项当然也会出现在非原始发散图的微扰计算当中。现在如果用重整化之后的、包含抵消项的那一套费曼图微扰计算非原始发散图 Γ ，那么被积函数除了 I_Γ 以外还包括将 Γ 中的一些子图用抵消项代替而得到的一些图给出的被积函数——如果 Γ 有彼此不相交（无共同顶角和粒子线）的发散子图 $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ （无论它是否是原始发散图），那么使用重整化后的微扰论计算 Γ 时，计算 $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ 时做的减除也会被作用在 Γ 上。一些组合数学的论证让我们发现，用重整化之后的微扰论计算 Γ 时的被积函数是

$$I_\Gamma - \sum_{\{\gamma_1, \dots, \gamma_n\}} I_{\Gamma/\{\gamma_1, \dots, \gamma_n\}} \prod_{i=1}^n \bar{R}_{\gamma_i},$$

其中 $\Gamma/\{\gamma_1, \dots, \gamma_n\}$ 指的就是将所有被提到的子图都缩成一个点之后得到的图产生的被积函数，这些子图彼此不交（当然，也不彼此嵌套）。 n 大小不定，因为总是可以丢掉一些子图。很显然，这样由于不相交的子图导致的发散应该都消去了，尚待解决的是，是否将交缠发散也消去了？然而要注意到，将交缠在一起的两个圈图中的一个断开，得到的就是非常普通的原始发散图，计算这个图之后把断开的圈连上，当成内线积分，就得到了交缠发散。我们用抵消项取代其中一个圈图，两者具有正好相反的发散，因此用抵消项取代其中一个圈图之后算另一个圈图会得到一个反向的交缠发散！交缠发散就是如此被消去的。

这样我们就得到了一套系统地、递归减除所有发散子图的方法：

$$R_\Gamma = (1 - t^\Gamma)(I_\Gamma - \sum_{\{\gamma_1, \dots, \gamma_n\}} I_{\Gamma/\{\gamma_1, \dots, \gamma_n\}} \prod_{i=1}^n \bar{R}_{\gamma_i}) \quad (6.27)$$

最外层的 $(1 - t^\Gamma)$ 因子消去了 Γ 中所有发散子图都被重整化为一个点之后，整体作为原始发散图而产生的发散。或者，换句话说，我们有

$$\bar{R}_\Gamma = I_\Gamma - \sum_{\{\gamma_1, \dots, \gamma_n\}} I_{\Gamma/\{\gamma_1, \dots, \gamma_n\}} \prod_{i=1}^n \bar{R}_{\gamma_i}. \quad (6.28)$$

最后我们分析杨-米尔斯理论的可重整性。通过量纲分析得到的几种发散图中的原始发散图的外线和可能的顶角的外线完全一致（实际上，正好一一对应），因此只需要说明没有与 $\log p$ 之类的无法直接抵消的项相乘的发散即可。

6.3.4 重整化群

我们知道朴素地做圈图计算会出现发散的原因是杨-米尔斯理论并不是在任何能标下都完全成立的普适理论，而是某个自洽的万有理论的一个低能近似。不过，在我们能够探测的所有情况下，杨-米尔斯理论都是适用的，即动量截断 Λ 是非常、非常高的。这也就是我们没有必要知道高能标处的物理具体是什么原因：我们只需要知道它会让低能标处的参数发生跑动，从而抵消掉所有发散就可以了。所有远离 Λ 的地方的所有物理过程都可以通

过标准的微扰重整化的方法——无论是传统方法还是BPHZ——完全计算出来。不言而喻，相较于 Λ 的不同“低能标”之间仍然可以有很大的差异：几乎处于绝对零度的凝聚态系统和对撞机中的电子都服从QED，但是显然它们的能标非常不同。

微扰重整化方法能够处理的区域是非常宽广的。我们经常直觉性地說：“在这个能标下，某某效应是重要的，在那个能标下，另一些效应是重要的”。我们能否指出，什么叫做“某个能标下（不是说这个能标以下，而是说这个能标附近）重要的物理是什么”？我们需要给“能标”下一个操作性的定义。在凝聚态场论中直接使用 Λ 没有任何问题，因为它反比于可观测的晶格常数，而在高能物理中没有这种东西。在高能物理中有一个地方可以自然地引入“能标”的概念：顶角函数通常依赖于入射粒子的动量的某些线性组合，我们通常会有效耦合常数定义为特定入射粒子动量下的顶角函数大小。例如，QED中的有效耦合常数就定义为光子动量为零时的顶角函数。在规范场论中这个定义可以直接推广。我们把诸如此类的“电子-光子散射过程中的光子动量”或是“四玻色子过程中入射玻色子动量差”用符号 μ 表示。自能的重整化条件是确定的，顶角函数的重整化条件则依赖于 μ 的取值。因此我们也可以将 μ 称为减除点，即我们在这个能标附近减除发散。

不同 μ 下的物理质量或是物理耦合常数等显然存在不同。再仔细考虑一下就会发现这里实际上有一个重整化： $\mu = 0$ ——即通常选取的重整化条件下原则上是可以计算具有任意的入射动量的过程的，但是在 $\mu \neq 0$ 时同样原则上可以计算具有任意的入射动量的过程。这两个不同的 μ 对应的两套重整化条件——从而两套物理质量和耦合常数——实际上彼此互为有效理论。因此在高能物理中，虽然我们并不会去调整 Λ ，积掉一些自由度来导致参数跑动，但这不是说不存在参数跑动的现象。参数跑动是有的，只不过由于这实际上是由于“在一个 μ 下一些图被用于修正质量而在另一个 μ 下一些图被用于修正耦合常数”，对应的重整化群真的就是群，因为没有任何信息损失。两种不同的重整化条件之间，相差一个重整化群操作。¹

我们尝试写出这种重整化群的重整化群方程。用 g 泛泛地指代理论中的各种常数，用 p 泛泛地指代各个动量。设 G 是一个包含 f 种场的关联函数，其中类型为 i 的场出现了 n_i 次。用 G_0 指代用裸量计算的关联函数。 G 依赖于 μ （因为微扰计算 G 时用到了含有 μ 的重整化条件），也依赖于 g （因为用到了含有 μ 的重整化条件必然会用到这个条件下的物理参数），而 g 也是依赖于 μ 的（实际上，由于 g 无非可以由一些物理过程的振幅确定，我们知道了某个 μ 下的 g 之后就可以计算另一些 μ 下的 g ）。这样，就有

$$G^{n_1 \cdots n_f}(p, g, \mu) = G_0^{n_1 \cdots n_f}(p, g_0) \prod_{i=1}^f Z_i^{n_i/2}.$$

由于

$$\frac{dG_0}{d\mu} = 0$$

而

$$\mu \frac{d}{d\mu} = \mu \frac{\partial}{\partial \mu} + \mu \frac{\partial g}{\partial \mu} \frac{\partial}{\partial g} = \frac{\partial}{\partial \log \mu} + \frac{\partial g}{\partial \log \mu} \frac{\partial}{\partial g}, \quad (6.29)$$

我们有Callan–Symanzik方程

$$\left(\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} + \sum_{i=1}^f n_i \gamma_i \right) G^{n_1 \cdots n_f}(p, g, \mu) = 0, \quad (6.30)$$

¹ 这不是说高能物理中不存在需要真的积掉一些高动量自由度的情况。QCD在低能区域的研究基本上就是强子的凝聚态物理，低能有效理论是经常要用到的。

其中

$$\beta(g) = \frac{\partial g}{\partial \log \mu}, \quad \gamma_i = -\frac{\partial Z_i^{1/2}}{\partial \log \mu} \quad (6.31)$$

分别称为 β 函数和反常量纲。前者给出物理常数的跑动，后者给出经过场强重整化修正的场的量纲。

重整化群(6.30)允许我们在知道了一个能标下的关联函数之后去预测另一个能标下的关联函数。很容易产生一个问题：如果取 $\mu = 0$ 就足够计算所有问题的话，为什么还需要取不同的 μ ？最为简单的答案是，这样更加方便：例如，设我们需要计算某个能量尺度大体上是 μ 的过程，而手头边正好有一个能标 μ 下的物理常数，那么使用基于 μ 的重整化条件显然更加方便，并且可能只需要计算树图就足够得到非常精确的结果。相反，从 $\mu = 0$ 处的物理常数出发做计算就可能要计算一些圈图。因此某个能量尺度大体上是 μ 的过程中会发生什么，使用基于 μ 的重整化条件是最能够清楚地展现的。这就体现出了重整化群的一个作用：它可以用于定性（以及在需要的时候，随时可以定量）地、直观地分析“某个能标下哪些效应明显”：

1. 如果随着重整化群流，某个物理参数在高能标下发散（此时称为出现朗道极点），那么这个理论基本上不要想能够预测比这更高的能标处的物理，即一个自然的动量截断被给出了。QED是一个典型的例子。
2. 如果随着重整化群流，某个物理参数在高能标下变得很小，即出现渐进自由，那么它在高能标下就肯定是非常好处理的（只要使用了正确的重整化条件）。QCD是一个典型的例子。实际上到目前为止，只有非阿贝尔的杨-米尔斯理论在四维闵可夫斯基时空中具有渐进自由。实验观测到强子在高能标下的确有渐进自由，因此这可以
3. 随着重整化群流一些相互作用耦合常数可能会有时候正有时候负，即有时候有吸引有时候排斥，这通常意味着存在相变。
4. 如果随着重整化群流，某个吸引相互作用在低能标下变得很大，那就出现了禁闭。QCD也是这方面的一个例子。

在凝聚态物理中实际上也可以看到类似的处理，如我们分析Kondo效应时就对散射过程的耦合常数做了一个重整化，实际上就是把对散射贡献比较大的过程先求和到一起，那么看着散射过程的耦合常数随着能标的变化就能够大概知道不同能标下散射振幅大小的变化。温度对应能标，所以我们就看到了不同温度下散射振幅大小的变化。

在具体做计算时，经常使用维数正规化中引入的 μ 作为参数来计算重整化群流。这种方式得到的重整化群方程、用动量截断和Wilson重整化群计算得到的重整化群方程、用重整化条件中光子动能为能标计算得到的重整化群方程都是一样的。后两者是一样的已经说明了，第一种方法和Wilson重整化群的等价性是由于，我们关心的区域总的来说都是低能的，从而在两种方案下，我们能看到的只是“一个能标带着一些参数在跑动”，而由于 μ 和 Λ 在拉氏量中的地位是完全相同的，它们给出同样的重整化群方程。

第七章 量子色动力学

本节给出非阿贝尔杨-米尔斯理论的一个例子，著名的描述了强相互作用的量子理论，量子色动力学或者简称QCD。

QCD是通过 $SU(3)$ 规范对称性得到的理论。 $SU(3)$ 的李代数的结构常数为

$$f^{123} = 1, \quad f^{147} = f^{165} = f^{246} = f^{257} = f^{345} = f^{376} = \frac{1}{2}, \quad f^{458} = f^{678} = \frac{\sqrt{3}}{2}. \quad (7.1)$$

我们取 $SU(3)$ 的维数最小的么正表示，此时李代数的表示的基底是

$$T^a = \frac{1}{2}\lambda^a, \quad (7.2)$$

λ^a 是所谓的盖尔曼矩阵，它是泡利矩阵在 $SU(3)$ 中的对应物，定义为

$$\lambda^1 = \quad (7.3)$$

这是一个三维表示，从而我们需要三个放在一起的旋量场做表示空间，记作

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}, \quad (7.4)$$

称为三重态。

我们将 ψ 场给出的粒子称为夸克。QCD中夸克的标签包括旋量场一定有的动量、自旋、手性，以及 $i = 1, 2, 3$ 提供的色指标。在标准模型中还有更多夸克的标签，包括三代夸克，每代有两种。 A^μ 场给出的粒子称为胶子，QCD中胶子的标签包括无质量矢量场一定有的动量和偏振，以及 $a = 1$ 到8，这里 a 也称为色指标。

第三部分

粒子物理和标准模型

第四部分

共形场论

第八章 二维共形场论的基本结构

提出共形场论的一个动机是相变现象。凝聚态系统在相变点附近关联长度发散，从而可以预期，能够使用一个无能隙的理论描述它，并且很大一类系统的这种理论将会具有非常好的性质。如果这个理论正好就是一个自由理论，那么对任何的关联函数的计算都是显然的：我们总是可以找到一组场，记之为 $\phi(x)$ ，任何一个局域的算符 $A(x)$ 均可以展开为 ϕ^n 的正规序的线性组合：

$$A(x) = \sum_{n \geq 0} a_n : \phi(x)^n :, \quad (8.1)$$

其中

$$: \phi(x)^n := \phi(x)^n - \langle \phi(x)^n \rangle,$$

在这么定义会产生疑难的时候，只需要让各个 x 略微差一些，然后计算完成后让它们相等即可。由于自由理论中场的量纲就是工程量纲，无需做任何特殊的考虑，就能够得到

$$\langle \phi(x)\phi(y) \rangle \sim \frac{1}{r^{2d_\phi}}, \quad r = |x - y|. \quad (8.2)$$

于是我们可以据此估计出任何一个关联函数的衰减趋势，并实际上真的计算出它。

一个非常自然的问题是，能否将这种做法——仅仅通过一个理论的**尺度不变性**就确定重要的关联函数的形式而无需使用常规的费曼图等技术，甚至通过形式由对称性确定的若干关联函数定义一个理论，完全绕开拉氏量（这称为**bootstrap**）——推广到一些不那么平凡、带有相互作用的量子场论中。为了尽可能减少理论中其它的结构，从而让bootstrap可以通过非常清晰的方式实现（例如只通过对称性信息是无法确定下来QED的），我们下面将要研究一大类具有非常完美的对称性的场论——**共形场论**。

8.1 共形对称性和共形群

8.1.1 任意维度的共形群

简单地说，**共形群**为空间尺度变换和洛伦兹变换共同组成的群（允许局域的操作）。具有共形不变性的场论就是**共形场论**。更加清楚的定义是这样的：如果一个坐标变换 $x \rightarrow x'$ 会导致

$$g'(x') = \Lambda(x)g(x), \quad (8.3)$$

那么这就是一个共形变换；全体共形变换组成共形群。洛伦兹群显然是共形群的一个子群；局域的坐标缩放——让原本均匀的坐标网格变得一些区域大一些区域小——也是共形变换。显然，共形场论是相对论性量子场论的一个特例。

我们现在分析共形群里面有什么操作。无穷小坐标变换

$$x^\mu \rightarrow x^\mu + \epsilon^\mu$$

下度规变化为

$$g_{\mu\nu} \longrightarrow g_{\mu\nu} - (\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu), \quad (8.4)$$

共形变换的定义(8.3)等价于能够找到函数 $f(x)$ 使得

$$\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = g_{\mu\nu} f(x). \quad (8.5)$$

实际上我们还能够将 $f(x)$ 关于 ϵ 的形式确定下来——只需要在上式两边取迹即可，我们有

$$2\partial_\mu \epsilon^\mu = df(x),$$

即

$$f(x) = \frac{2}{d} \partial_\mu \epsilon^\mu, \quad (8.6)$$

其中 d 为时空总维数。这里要注意 $g_\mu{}^\nu$ 实际上就是 δ_μ^ν ，从而求迹之后一定是 d ，而不是 $\eta_{\mu\nu}$ 矩阵求迹后的 $2-d$ 。

下面为了方便起见我们只考虑 $g_{\mu\nu}$ 是欧氏空间度规的情况；也即，我们做一个Wick转动，将所有的相对论性量子场论切换到虚时间下。我们接着可以尝试确定可能的 ϵ 的形式。对(8.5)两边求导数，我们有

维数大于2的共形场论是一个有限维的普通李群，有有限个生成元。高维共性不变性是整体的，没有局域不变性，从而也不是规范理论。然而，对二维系统，共形变换是局域变换群，共性不变的场论也是一种规范理论。此时的共性群的生成元有无限多个，且量子化时会出现反常。

8.1.2 二维共形群和共形场论

二维共形场论的重要性一方面来自它自身的奇特性质——对称群有无穷多个生成元，具有很多非常有趣的性质——一方面也有很强的物理意义。

二维系统的 ϵ_1 和 ϵ_2 满足柯西-黎曼条件，从而二维的共性变换就是一个解析函数。

$$\partial_z = \frac{1}{2}(\partial_1 - i\partial_2), \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_1 + i\partial_2), \quad (8.7)$$

$$\partial_1 = \partial_z + \partial_{\bar{z}}, \quad \partial_2 = i(\partial_z - \partial_{\bar{z}}). \quad (8.8)$$

$$ds^2 = dx^2 + dy^2 = dz d\bar{z}. \quad (8.9)$$

$$g_{zz} = g_{\bar{z}\bar{z}} = 0, \quad g_{z\bar{z}} = g_{\bar{z}z} = \frac{1}{2}. \quad (8.10)$$

严格的共形不变性会对 $T_{\mu\nu}$ 做出非常强的限制：旋转对称性意味着

$$T_{\mu\nu} = T_{\nu\mu}, \quad (8.11)$$

标度不变性意味着

$$T_\mu{}^\mu = 0. \quad (8.12)$$

$x^\mu = (x^1, x^2)$, $x^{\mu'} = (z, \bar{z})$ ，则

$$\left[\frac{\partial x^\mu}{\partial x^{\mu'}} \right]_{\mu\mu'} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{i}{2} & \frac{i}{2} \end{pmatrix},$$

由于

$$T_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} T_{\mu\nu} \frac{\partial x^\nu}{\partial x^{\nu'}},$$

有

$$[T_{\mu'\nu'}]_{\mu'\nu'} = \left[\frac{\partial x^\mu}{\partial x^{\mu'}} \right]_{\mu\mu'}^\top [T_{\mu\nu}]_{\mu\nu} \left[\frac{\partial x^\mu}{\partial x^{\mu'}} \right]_{\mu\mu'},$$

计算得到

$$\begin{aligned} T_{zz} &= \frac{1}{4}(T_{11} - T_{22} + 2iT_{12}) =: T(z, \bar{z}), \\ T_{\bar{z}\bar{z}} &= \frac{1}{4}(T_{11} - T_{22} - 2iT_{12}) =: \bar{T}(z, \bar{z}) \\ T_{z\bar{z}} &= T_{\bar{z}z} = \frac{1}{4}(T_{11} + T_{22}) = \frac{1}{4}T_\mu^\mu =: \frac{1}{4}\Theta(z, \bar{z}). \end{aligned} \quad (8.13)$$

使用这些记号，守恒律 $\partial_\mu T^{\mu\nu} = 0$ 变成在共形对称性严格成立的时候，或者说在临界点上，有

$$\partial_z \bar{T}(z, \bar{z}) = \partial_{\bar{z}} T(z, \bar{z}) = 0. \quad (8.14)$$

8.1.3 二维无穷小共性变换

既然二维共形变换实际上就是一个解析函数，

无穷小生成元：

$$l_n = -z^{n+1}\partial, \quad \bar{l}_n = -\bar{z}^{n+1}\bar{\partial}. \quad (8.15)$$

8.2 共形场论中的关联函数和算符代数

8.2.1 局域场的算符代数

现在我们考虑相互作用系统。此时随意选择一个 $\varphi(x)$ ，一般来说是不能有以上操作了，因为此时 ϕ 未必有完全确定的反常量纲，例如场论中 ϕ^2 项的 β 函数可能同时显含其它很多项的参数。然而，假定我们确实找到了一组可数的场 $\{\varphi_i(x)\}$ ，使得 φ_i^2 项真的就是临界点附近重整化群流的本征方向，从而

$$\varphi_i(x) = \lambda^{d_i} \varphi_i(\lambda x), \quad (8.16)$$

并且，进一步，任何一个局域算符都可以写成这组场的线性组合（或者至少，在关联函数的括号 $\langle \cdot \rangle$ 中可以这么做——这种线性展开可能并不一般地成立；后文中很多类似的线性展开也需要如此理解），即

$$A(x) = \sum_i a_i \varphi_i(x). \quad (8.17)$$

对自由理论，显然

$$\varphi_i(x) =: \varphi(x)^i:, \quad (8.18)$$

对相互作用体系我们尚不清楚 $\varphi_i(x)$ 是什么。无论如何，在(8.16)严格成立时，我们有

$$\langle \varphi_n(x) \rangle = 0. \quad (8.19)$$

现在我们得到了任何局域的算符的代数：就是一个线性代数。 $\varphi_n(x)$ 是局域的算符的基底。现在考虑两个相隔了有限距离的点上的局域算符的乘积，即 $A(x_1)B(x_2)$ 。当 $x_2 \rightarrow x_1$ 时，这个形式会变成一个单一的局域算符，从而可以做算符展开，

$$A(x_1)B(x_2) = \sum_k \beta(x_1, x_2) \varphi_k(x_2). \quad (8.20)$$

特别的, 如果 $A(x)$ 和 $B(x)$ 实际上都是 $\phi(x)$ 场,

$$\varphi_p(x_1)\varphi_q(x_2) = \sum_r C_{pq}^r(x_1, x_2)\varphi_r(x_2). \quad (8.21)$$

$$C_{pq}^r(x_1, x_2) = c_{pq}^r \frac{1}{|x_1 - x_2|^{d_p+d_q-d_r}}. \quad (8.22)$$

其中 c_{pq}^r 是算符代数的结构常数。

似乎我们有两种方法定义共形场论, 其一是通过局部的、在尺度变换下协变的算符代数, 其二是假定度规在局域尺度变换下只差一个常数。

$$x' = x + \epsilon, \quad (8.23)$$

$$\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = \rho(x)g_{\mu\nu}, \quad (8.24)$$

对上式两边求迹, 得到

$$2\partial_\mu \epsilon^\mu = D\rho(x),$$

于是

$$\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu = \frac{2}{D}g_{\mu\nu}\partial \cdot \epsilon \quad (8.25)$$

描述它们的代数不再是普通的李代数, 而是Virasoro代数。

参考文献

- [1] S. Elitzur. Impossibility of spontaneously breaking local symmetries. *Physical Review D*, 12(12):3978–3982, dec 1975.