

基于对称性的现代物理学

原著：Jakob Schwichtenberg 翻译：超理汉化组

Translate Version: 2017 年 8 月 31 日

本文档由热心网友翻译制作, 原著书名为: Physics from Symmetry, 属 Springer 公司 Undergraduate Lecture Notes in Physics(ULNP) 丛书, ISBN 978-3-319-19200-0, DOI 10.1007/978-3-319-19201-7。 仅限教学科研使用, 使用与传播应严格遵守相关法律规定。

《中华人民共和国著作权法》第四节第二十二条:

在下列情况下使用作品, 可以不经著作权人许可, 不向其支付报酬, 但应当指明作者姓名、作品名称, 并且不得侵犯著作权人依照本法享有的其他权利:

...

(六) 为学校课堂教学或者科学研究, 翻译或者少量复制已经发表的作品, 供教学或者科研人员使用, 但不得出版发行;

...

© Springer International Publishing Switzerland 2015

© 2016 Chaoli-Translating-Group (超理汉化组)

NOT PUBLISHED

Preface

世界上最无法理解的事情是这个世界竟然是可以理解的。

– Albert Einstein¹

¹ 引自

Jon Fripp, Deborah Fripp, and Michael Fripp. *Speaking of Science*. Newnes, 1st edition, 4 2000. ISBN 9781878707512

在物理课上，我像任何一个物理系学生那样熟悉许多基本方程与它们的解，但我不是很清楚它们之间的联系。

当我理解它们中的大多数具有共同的起源 — **对称性 (Symmetry)** 时我十分激动。对我而言搞物理最美的经历莫过于原本费解的东西经过深入探究之后豁然开朗，我因此深爱着对称性。

例如，有一段时间我不能真正理解自旋 — 几乎所有基本粒子都具有的奇特的内禀角动量，后来我学到了原来自旋是一种对称性 (称为 Lorentz 对称性) 的直接结果，于是有关自旋的内容开始充满意义。

本书的目的就是为读者提供这样的经历，在某种意义上，当我开始学物理的时候我就有写这本书的愿望了。¹ 对称性漂亮地解释了其它方面的许多复杂物理现象，这让我们认为可以从对称导出物理学的基本理论，本书正是基于这个锐利而光芒耀眼的信念。²

¹同感! 这就是物理系学生的中二病吗 2333333 — 译者 (SI)

²锐 (するど) くギラつかせた希望 (きぼう), もっと本気 (ほんき) 出 (だ) していいよ, あ~, あ~, あ~, あ~

可以说本书的写作顺序是倒过来的：在我们讨论经典力学与非相对论量子力学之前，我们将利用大自然（据我们所知）精确的对称性导出量子场论的基本方程。不过尽管途径不同，本书的内容仍为标准物理学，我们不涉及仍有争议未经实验验证的内容，而是通过物理学的标准假设导出物理学的标准理论。

根据读者的物理水平，本书可以有两种用途。

- 物理学的初学者³可以把这本书当成入门教材。它包含了经典力学、电动力学量子力学、狭义相对论和量子场论的基本理论，在阅读之后，读者可再深入学习各部分内容。各部分有许多更加深入的优秀书籍，在各章的末尾列出了一些延伸阅读的推荐表。如果你觉得你属于这一种，我们建议你在阅读正文之前先从附录的数学补充²开始。
- 另一种情况，身经百战见得多的学生可以通过本书将松散的各物理领域紧密联系起来。回顾历史的进程我们会发现许多物理思想可能看起来随意甚至乱来。但从对称性的观点看，它们往往就变得必然而直接。

とにかく，不管哪种情况都应该按顺序阅读本书，因为各章之间是递进的。

开始的短章节是关于狭义相对论的，它是之后讨论的所有内容的基础。我们会看到对物理理论最重要的限制之一是它们必须符合狭义相对论。

本书的第二部分引入了描述对称性的数学工具（在物理味道的表述下），大多数工具来自数学的重要分支——群论。之后我们介绍拉格朗日形式 (Lagrangian formalism)，它使我们能够在物理体系中直接利用对称性。

第五章和第六章利用之前引入的拉格朗日形式和群论导出了现代物理学的基本方程。

在最后一部分，我们把之前的基本方程加以应用：应用于粒子理论我们可以导出量子力学，应用于场理论我们可以导出量子场论。然后我们研究这些理论在非相对论极限与经典极限的变体，这样又导出了经典力学与电动力学。

每章的开始是本章内容的摘要。如果你发现你在思考‘这是在讲些啥’，不妨回到章节开头看看摘要以明白某一部分的目的是什

² 开始自附录 A。此外，当正文使用一个新的数学概念时，边注中会说明相应的附录目录。

³ 这个‘初学者’是‘相对而言的’(relatively)... — 译者注

么。书页留有巨大的页边空白⁴以防止想出 Fermat 大定理的证明
却没地方写 你可以一边阅读一边在页边记下笔记与灵感³。

我希望你读这本书能够像我写这本书的时候一样愉悦。

³ 许多页边注的内容是拓展信息与图景。

Karlsruhe, 2015.01

Jakob Schwichtenberg

DRAFT TRANSLATION

⁴给译文排版造成了小小的麻烦 — 译者

Acknowledgments

感谢所有帮我编写这本书的人。我特别感激 Fritz Waitz, 他的评论、想法与纠正让本书质量大大改善。我十分感谢 Arne Becker 和 Daniel Hilpert, 感谢他们无价的建议、意见与细致的校对。感谢 Robert Sadlier 对我英文的帮助以及 Jakob Karalus 的解释。

我还想感谢与我有许多见解深刻的讨论的 Marcel Köpke, 感谢 Silvia Schwichtenberg 和 Christian Nawroth 的支持。

最后, 我亏欠最多的是我的父母, 他们支持着我, 教导我知识高于一切。

如果发现文中的错误, 我非常希望你能够寄一封短邮件到 errors@jakobschwichtenberg.com。勘误表的地址是 <http://physicsfromsymmetry.com/errata>。

Contents

I	Foundations 基础	
1	简介	5
1.1	我们所不知的事情	5
1.2	全书概览	7
1.3	基本粒子和基本相互作用力	9
2	狭义相对论	11
2.1	狭义相对论的中的不变量	12
2.2	固有时	15
2.3	速度上限	16
2.4	Minkowski 记法	17
2.5	Lorentz 变换	20
2.6	不变性，对称性，协变性	21

3	Lie 群	27
3.1	群	28
3.2	二维旋转	31
3.2.1	用单位复数表示旋转变换	33
3.3	三维旋转	35
3.3.1	四元数	36
3.4	Lie 代数	40
3.4.1	$SO(3)$ 群的生成元与 Lie 代数	44
3.4.2	Lie 代数的抽象定义	47
3.4.3	$SU(2)$ 的生成元与 Lie 代数	48
3.4.4	Lie 群的抽象定义	50
3.5	表示论	52
3.6	$SU(2)$	55
3.6.1	$SU(2)$ 群的有限维不可约表示	56
3.6.2	$SU(2)$ 的 Casimir 算符	58
3.6.3	$SU(2)$ 的一维表示	59
3.6.4	$SU(2)$ 的二维表示	59
3.6.5	$SU(2)$ 的三维表示	60
3.7	Lorentz 群	61
3.7.1	Lorentz 群的 1 表示	63
3.7.2	Lorentz 群其他部分的生成元	67
3.7.3	$SO(1, 3)^\uparrow$ 群的 Lie 代数	69
3.7.4	$(0, 0)$ 表示	71
3.7.5	$(\frac{1}{2}, 0)$ 表示	71
3.7.6	$(0, \frac{1}{2})$ 表示	73
3.7.7	Van der Waerden 符号	74
3.7.8	$(\frac{1}{2}, \frac{1}{2})$ 表示	79
3.7.9	旋量和宇称	82
3.7.10	旋量与电荷共轭	84
3.7.11	无限维表示	85
3.8	Poincaré 群	87
3.9	基本粒子	89
3.10	附录：复向量空间中的旋转	90

	11
3.11 附录：流形	90
4 理论框架	93
4.1 拉格朗日形式	93
4.1.1 Fermat 原理	94
4.1.2 变分法：基本思想	94
4.2 限制	95
4.3 粒子理论与场论	96
4.4 欧拉-拉格朗日方程	97
4.5 Noether 定理	99
4.5.1 粒子理论的 Noether 定理	99
4.5.2 场论中的 Noether 定理——时空对称性	104
4.5.3 旋转和推动	107
4.5.4 自旋	108
4.5.5 场论中的 Noether 定理——内禀对称性	109
4.6 附录：粒子理论中的推动不变性与其守恒量	111
4.7 附录：场论中的推动不变性与其守恒量	112
<div>III The Equations of Nature 自然界的方程</div>	
5 测量	119
5.1 量子力学中的算符	119
5.1.1 自旋与角动量	120
5.2 量子场论的算符	121
6 自由场理论	123
6.1 Lorentz 协变性与不变量	123
6.2 Klein-Gordon 方程	124
6.2.1 复 Klein-Gordon 场	126
6.3 Dirac 方程	126
6.4 Proca 方程	130

7	相互作用理论	133
7.1	$U(1)$ 相互作用	135
7.1.1	自旋 $\frac{1}{2}$ 自由场的内禀对称性	135
7.1.2	自旋为 1 自由场的内禀对称性	137
7.1.3	把奇怪的东西堆一起	138
7.1.4	非齐次麦克斯韦方程和最小耦合	141
7.1.5	再一次电荷共轭变换	142
7.1.6	内禀 $U(1)$ 对称性的 Noether 定理	143
7.1.7	零自旋有质量场的相互作用	145
7.1.8	自旋为 1 有质量场的相互作用	146
7.2	$SU(2)$ 相互作用	146
7.3	质量项、$U(1)$ 和 $SU(2)$ 的统一	153
7.4	宇称破坏	162
7.5	轻子质量项	166
7.6	夸克质量项	170
7.7	同位旋	171
7.7.1	Labelling 态	173
7.8	$SU(3)$ 相互作用	174
7.8.1	色	176
7.8.2	夸克描述	178
7.9	Bosons 和 Fermions 间的游戏	179
IV	Applications 应用	
8	量子力学	185
8.1	对应到粒子理论	186
8.2	相对论性能量-动量关系	186
8.3	量子力学的数学形式	187
8.3.1	期望值	190
8.4	Schrödinger 方程	191
8.4.1	有外场的 Schrödinger 方程	193
8.5	从波动方程到粒子的运动	194
8.5.1	例：自由粒子	194

8.5.2	例：盒子中的粒子	195
8.5.3	Dirac 符号	198
8.5.4	例：盒子中的粒子，续	201
8.5.5	自旋	201
8.6	Heisenberg 不确定性原理	204
8.7	对几种诠释的评议	205
8.8	附录：Dirac 旋量分量的诠释	207
8.9	附录：解 Dirac 方程	211
8.10	附录：Dirac 旋量的不同的基	213
8.10.1	质量基下 Dirac 方程的解	215
9	量子场论	219
9.1	对应到场理论	220
9.2	自旋为 0 的自由场理论	221
9.3	自由的自旋 $\frac{1}{2}$ 理论	227
9.4	自旋 1 的自由场理论	230
9.5	相互作用场论	230
9.5.1	散射振幅	231
9.5.2	态的时间演化	231
9.5.3	Dyson 级数	234
9.5.4	计算级数	236
9.6	Klein-Gordon 方程的一般解	239
10	经典力学	243
10.1	相对论力学	245
10.2	非相对论力学的拉格朗日量	246
11	电动力学	249
11.1	齐次 Maxwell 方程组	250
11.2	Lorentz 力	252
11.3	Coulomb 势	253
12	引力	255
13	结束语	261

A	矢量代数	265
A.1	基矢	266
A.2	坐标系变换	268
A.3	矩阵乘法	269
A.4	标量	270
A.5	左手/右手坐标系	271
B	微积分	273
B.1	Product Rule	273
B.2	分部积分	273
B.3	Taylor 级数	274
B.4	级数	276
B.4.1	几个重要的级数	276
B.4.2	分裂求和	278
B.4.3	Einstein 求和约定	279
B.5	指标记号	280
B.5.1	哑指标	280
B.5.2	带多个指标的对象	280
B.5.3	对称/反对称指标	281
B.5.4	对称和反对称求和	281
B.5.5	两个重要的符号	282
C	线性代数	285
C.1	基本变换	285
C.2	矩阵指数函数	286
C.3	行列式	286
C.4	本征值与本征向量	287
C.5	对角化	287
D	其他数学概念	289
D.1	Fourier 变换	289
D.2	δ 分布函数	290



Foundations 基础

1	简介	5
1.1	我们所不知的事情	
1.2	全书概览	
1.3	基本粒子和基本相互作用力	
2	狭义相对论	11
2.1	狭义相对论的中的不变量	
2.2	固有时	
2.3	速度上限	
2.4	Minkowski 记法	
2.5	Lorentz 变换	
2.6	不变性，对称性，协变性	

“真相总比你想的简单”

– Richard P. Feynman

as quoted by

K. C. Cole. *Sympathetic Vibrations*.

"The truth always turns out to be simpler than you thought."

K. C. Cole.

Sympathetic Vibrations.

Bantam, reprint edition,
10 1985.

ISBN 9780553342345

DRAFT TRANSLATION

1. Introduction 简介

1.1 What we Cannot Derive 我们所不知的事情

在讨论对称性能告诉我们什么之前，我们得先说说需要人为的在理论中添加些什么。首先，目前没有任何理论可以给出自然界的常数。这些常数需要实验测定，比如各种相互作用的耦合常数啊，基本粒子的质量啊这种的。这就是需要人为加入的参数。

除此之外，我们还有很多解释不了的：数字 3。这不是术数的那种神秘主义的东西，而是我们不能解释所有的直接与数字 3 相联系的限制。比如：

- 对应三种标准模型描述的基本作用力有三种规范理论¹。这些力是由分别对应于对称群 $U(1)$, $SU(2)$ 和 $SU(3)$ 的规范理论描述的。为什么没有对应 $SU(4)$ 带来的基本作用力？没人知道！
- 轻子有三代，夸克也有三代。为什么没有第四代？我们只能从实验²中知道没有第四代。
- 我们只在拉格朗日量里面包含 Φ 的最低三阶 (Φ^0, Φ^1, Φ^2)，其中 Φ 指代一些描述我们的物理系统的东西，是个通称，而我们用拉格朗日量描述自由（即无相互作用）的场/粒子。
- Poincaré 群双覆盖的基本表示中我们只用到三个，分别对应自旋 $0, \frac{1}{2}$ 和 1 。没有基本粒子的自旋是 $\frac{3}{2}$ 。

这是目前理论中必须人为引入的假设。我们从实验上知道这

¹ 如果你不理解这个简介中的某些名词，比如规范理论或者双覆盖，不需要太过担心。本书将会详尽的解释，在这里提到只是为了完整。

² 比如，现在宇宙中元素的丰度是依赖于代的数量的。更进一步，对撞机实验中有对此的很强的证据。（见 Phys. Rev. Lett. 109, 241802）

些假设是正确的，但是目前为止没有更深刻的原理告诉我们为什么一到 3 就得停。

除此之外，还有两件事情没法从对称性中得到，但却是一个严谨的理论所必须考虑的：

- 我们只允许在拉格朗日量中引入尽可能低阶的非平庸微分算符 ∂_μ 。一些理论使用一阶的微分算符 ∂_μ ，而另一些理论中，Lorentz 不变性禁止了一阶导数，从而二阶导数 $\partial_\mu \partial^\mu$ 是最低阶的可能的非平庸项。除此之外我们就再也得不到一个合理的理论了。存在高阶导数项的理论，能量没有下界，这导致它可以是一个任意大小的负值；因此，这些理论中的态总可以跃迁到能量更低的态，从而永远不会稳定。
- 类似的，如果半整数自旋的粒子和整数自旋的粒子拥有完全一样的行为的话，宇宙中就不会有稳定的物质。因此，这两者必然有某些东西不一样，而我们没得选，只有一种可能而且合理的选择³是正确的。这引出了半整数自旋粒子的 Fermi-Dirac 统计的概念和整数自旋粒子的 Bose-Einstein 统计的概念。半整数自旋的粒子从而通常被称为 Fermion¹，它们中永远不存在两个粒子处在完全一样的态上。相反，这种情况对整数自旋的粒子 — 通常被称为 Boson² — 是可能的。

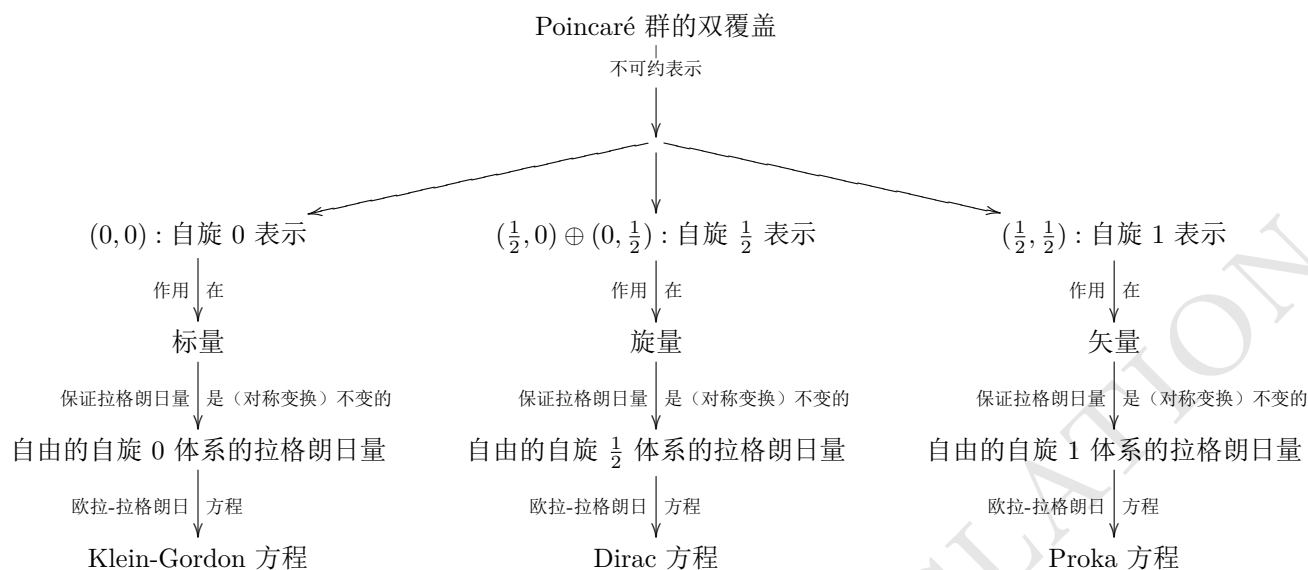
最后呢，本书不能导出引力理论。大名鼎鼎的广义相对论就是优美而准确的描述引力的理论；然而这个理论与其他理论完全不一样，而导出它的框架则超出了本书的研究范围。尝试将引力问题划入相同框架下的量子引力理论仍待完善，目前没有人能够成功得出一个完整的形式。但最后一章我们会对引力做一些评述。

³ 量子场论里一开始使用反对易子而不是对易子，从而防止理论没有下界。

¹常译为费米子

²常译为玻色子

1.2 Book Overview 全书概览



这本书使用**自然单位制**，也就是说 Planck 常数 $\hbar = 1$ ，光速 $c = 1$ 。这是基本理论的惯例，它免除了很多不必要的笔墨。而对于应用来说呢，这些常数需要被再一次的加上去从而回到标准的 SI 单位制。

本书从**狭义相对论**的基本假设出发：一方面，在所有的惯性参考系 — 一些相互之间的速度保持恒定的参考系 — 中，光的速度不变，为 c ；而另一方面物理规律在所有惯性系中相同。

满足这些对称性的所有的变换构成的集合叫做 **Poincaré 群**。我们接下来介绍一些数学知识从而使得我们能够利用好对称性和这些对称变换。数学上这一分支叫群论。我们会得到 Poincaré 群⁴的不可约表示——你可以当成是组成其它所有表示的“地基”。这些表示就是后面用来描述粒子和场的不同自旋的表示。自旋一方面是对不同种类的粒子和场的标记，而另一方面也可以视为内禀角动量。

之后介绍力学问题的**拉格朗日形式**，它使得我们可以方便的在物理问题里面使用对称性。这里面的核心对象是**拉格朗日量**，我们通过考虑对称性来给出不同物理系统的拉格朗日量³。而在此之上我们就得到了**欧拉-拉格朗日方程**，从而我们可以得到给定拉格朗日量下，系统的运动方程。利用 Poincaré 群的不可约表示，就可以导出带有不同自旋的场和粒子的基本运动方程。

⁴ 严格来说，我们会得到 Poincaré 群的双覆盖而不是 Poincaré 群本身。“双覆盖”这个词是由于一个群的双覆盖和这个群的关系是前者的两个元素与后者的一个元素相对应。这在后面的节 3.3.1 中会讲到。

³ 实际上，拉格朗日量是人为写出来的，而不是能够推导出来的。它是我们对某类体系『唯象』且经验的刻画 — 译者注

这里面的中心思想是在 Poincaré 群的作用下，拉格朗日量必须是不变的。这使得在任意的参考系中得到的运动方程都一样，就像之前说的那样，“物理规律在所有惯性系中相同”。

之后我们就会发现对于自旋 $\frac{1}{2}$ 场的拉格朗日量的另一个对称性：在 $U(1)$ 变换下的对称性。类似的，对于自旋 1 的场的内禀对称性也可以找到。局域的 $U(1)$ 对称性可以使得我们得到自旋 $\frac{1}{2}$ 场和自旋 1 场之间的耦合项。带这样的耦合项的拉格朗日量是量子电动力学的拉格朗日量的正确形式。类似的局域 $SU(2)$ 和 $SU(3)$ 变换会得到弱和强相互作用的拉格朗日量的正确形式。

作为补充，我们会讨论自发对称破缺和 Higgs 机制。这些让我们能够去描述有质量的粒子⁵。

之后就可以导出 Noether 定理，它给我们展示了对称性和守恒量之间深刻的联系。在物理量和相应对称性的生成元之间划等号，就能导出量子力学中最重要的方程⁴

$$[\hat{x}_i, \hat{p}_j] = i\delta_{ij} \quad (1.1)$$

和量子场论中

$$[\hat{\Phi}(x), \hat{\pi}(y)] = i\delta(x - y) \quad (1.2)$$

接下来可以通过对自旋 0 粒子的运动方程，即 Klein-Gordon 方程，取非相对论性⁶极限，来得到了著名的 Schrödinger 方程。而与此同时，我们又认识到了物理量与对称性生成元之间的联系，这些一起就构成了量子力学的基石。

接下来我们从不同的运动方程⁷的解，和方程(1.2)出发来考察自由场的量子场论。这之后我们通过仔细审视包含不同自旋场之间的耦合的拉格朗日量来考虑相互作用。这使得我们可以讨论散射过程的概率振幅是如何得到的。

我们用 Ehrenfest 定理来演示量子 and 经典力学的联系。更进一步，我们够得了经典电动力学的基本方程，包括 Maxwell 方程组和 Lorentz 力定律。

最后简要的介绍现代引力理论 — 广义相对论的基本结构，并点出一些构建引力的量子理论的困难。

这本书的主要精力会花在处理对称性的数学工具，以及导出

⁵ 在对称性自发破缺前，拉格朗日量中的描述质量的项会因破坏对称性而被禁止

⁶ 非相对论性指所有物体相比光速来说都运动的十分缓慢，从而狭义相对论的古怪的特性很小不会被观察到。

⁷ Klein-Gordon, Dirac, Proca 和 Maxwell 方程

⁴原书有 typo

被称为**标准模型**的理论上。标准模型是用量子场论为语言来描述所有已知的基本粒子的行为的。直到现在，所有标准模型的实验预言都是正确的。这里引入的其他任何理论都可以看作标准模型下的特殊情况，比如对于宏观物体：经典力学，或者对于低能的基本粒子：量子力学。对于现今已知的基本粒子和它们之间相互作用的情况，我们会在下一小节进行一个快速的概述，以方便那些不熟悉相关内容的读者。

1.3 Elementary Particles and Fundamental Forces 基本粒子和基本相互作用力

基本粒子分成两个主要的大类：**bosons** 和 **fermions**。**Pauli 不相容原理**说明不能有两个 **fermions** 处在完全相同的态上，而 **boson** 则可以有任意多个粒子处在同一个态上。这种自然界奇怪的事实导致了这些粒子截然不同的性质：

- Fermions 构成物质⁵
- Bosons 构成自然界的力

这说明，比如，原子是由 **fermions**⁸组成的，但是电磁相互作用力是通过被称为光子的 **bosons** 传播的。这带来的一个最令人震惊的结果是竟然能形成稳定的物质。如果允许有无限多的 **fermions** 处在相同的状态，根本就不可能有稳定的物质存在，我们会在第6章中讨论。

目前我们知道四种基本作用力

- 电磁相互作用力，通过无质量的光子传播
- 弱相互作用力，通过有质量的 W^+ , W^- 和 Z - **bosons** 传播。
- 强相互作用力通过无质量的胶子传播。
- 引力，（或许）通过引力子传播。

这里面，一些相互作用对应的 **bosons** 是无质量的，而另一些不是，这揭露了一些自然界的深刻的事情。我们在做好足够的准备之后会去仔细理解这一点，现在只需要记住每一种相互作用力都与一个对称性相关。弱相互作用力传递的 **bosons** 是有质量的说明其对应的对称性被破坏了。这种**自发对称破缺**是所有基本粒子质量的起源。我们后面会看到这实际上是通过其与另一种基本

⁸ 原子有电子，质子和中子组成，它们都是 **fermions**。然而注意质子和中子都不是基本的粒子，它们有夸克组成，而夸克也是 **fermions**。

⁵ 这里，『构成』实际上不是很合适的翻译。直译的话应该叫 **Fermions 对物质『负责』**。不过大体上就是那么个意思吧。---译者注

⁹ 在第7章中我们会讲，所有的荷都有着相同的美丽起源

¹⁰ 通常弱相互作用力的荷会伴随着“弱”这个前缀，所以也被称为弱同位旋；而且还存在一种其它的叫同位旋的概念适用于强相互作用里面的复合物。不仅如此，这不是一个基本的荷，而本书中会省略“弱”这个前缀。

boson, **Higgs boson** 的耦合来实现的。
基本粒子如果携带某种荷⁹的时候它可以通过某种力来相互作用。

- 对于电磁力来说，这种荷就是**电荷**，而相应的只有带电粒子能够通过电磁力相互作用。
- 对于弱相互作用力来说，这种荷称为¹⁰**同位旋**。所有已知的 fermions 都携带同位旋，因此它们之间通过弱作用力来相互作用。
- 强作用力的荷叫**色**，因为它与人对色彩的知觉有某种奇妙的共性。不要被这个名字搞糊涂，这个荷实际上与日常生活中的色彩没有任何关系。

基本的 fermions 被分成两个种类：**夸克**，组成质子和中子的基本组成；与**轻子**，如电子和中微子。其中的区别体现为夸克可以通过强作用力来相互作用，这就是说他们带色，而轻子则不是这样。夸克和轻子有三代，每一代有两种粒子：

	第一代	第二代	第三代	电荷	同位旋	色
夸克:	上 (u)	粲 (c)	顶 (t)	$+\frac{2}{3}e$	$\frac{1}{2}$	✓
	下 (d)	奇异 (s)	底 (b)	$-\frac{1}{3}e$	$-\frac{1}{2}$	✓
轻子:	电子中微子	μ 子中微子	τ 子中微子	0	$+\frac{1}{2}$	-
	电子	μ 子	τ 子	$-e$	$-\frac{1}{2}$	-

一般的来说，不同的粒子可以通过**标签**来识别。除了荷和质量之外，它们还有很多其他的十分重要的标签，称为**自旋**，你可以把它理解为内禀的角动量，我们会在第4.5.4节中得到这一点。Bosons 携带整数自旋，fermions 携带半整数自旋。我们上面列出来的基本的 fermions 都只有 $\frac{1}{2}$ 的自旋。基本的 bosons 有着 1 的自旋。除此之外，只有一种基本粒子有 0 自旋：Higgs boson。

每种粒子都有其反粒子，带着大小完全一样的相反¹¹的标签。比如电子的反粒子就是正电子，但是一般的话我们不会人为的说一个新的名字，只是在原来的粒子前面加上前缀“反”。比如说，顶夸克的反粒子被称为反顶夸克。有些诸如光子¹²一类的粒子自己就是自己的反粒子。

上面提到的这些想法之后都会更具体的解释。现在是时候去尝试推导理解这门能够正确描述粒子大观园中的不同特性是如何相互影响在一起的理论了。而第一步，也就是我们下一章的主题，是 Einstein 著名的狭义相对论。

¹¹ 除了质量标签。这点目前正在接受实验的检验，比如瑞士日内瓦的 CERN 进行的 ARGIS, ATRAP 和 ALPHA 实验。

¹² 或许还有中微子，这一点目前正在接受各种关于搜索无中微子的双光子通道衰变实验研究。

2. Special Relativity 狭义相对论

著名的 Michelson-Morley 实验告诉我们，光在任何参考系中都具有相同的速度¹。Albert Einstein 首先意识到这个结果所蕴含的深刻意义，从而在此基础上建立了狭义相对论。从光速不变原理出发，Einstein 预言了许多有趣而又悖于常理的现象，最后都被实验证实是正确的。在这里我们首先阐明狭义相对论的基础，然后再体会这个原理的强大之处。狭义相对论有两个基本假设：

- **相对性原理：** 任意惯性系，即任意两个相对速度恒定的参考系的物理规律相同。
- **光速不变原理：** 在任意惯性系中的光速均为常数 c 。

除此之外，我们还假设时空均匀且各向同性。这意味着不论在哪儿（均匀），不论朝着什么方向（各向同性）做实验，物理规律都是一样的。例如两个物理学家，一个在纽约，一个在东京，他们做完全一样的两个实验，会得到相同²的物理规律，就算是在火星也一样。

正确的物理定律不应该随看实验的角度或是进行实验的时间点而变。¹此外，狭义相对论的第一条假设告诉我们，不管是在匀速运动的马车上，还是在静止的实验室中，做同一物理实验都会得到相同的结果。以上论述均与我们的生活经验相符。举个例

¹ 日常生活中所观测到的物体的速度取决于选定的参考系。如果一个观察者站在火车站上，测得的火车速度为 50km h^{-1} ，另一个观察者以 15km h^{-1} 的速度与火车一同运动，那么测得的火车速度就应是 35km h^{-1} 。与此不同的是，光始终以 $1.08 \times 10^9\text{km h}^{-1}$ 运动，不论如何相对于光运动。

² 这当然要排除某些参数的影响，例如重力加速度。

¹ 原文为：The laws of physics, formulated correctly, shouldn't change if you look at the experiment from a different perspective or repeat it tomorrow.

子，如果你闭着眼睛坐在匀速运动的汽车上，你是没有办法分辨出你是真的在运动还是静止的。

如果没有了各向同性和均匀性，物理学就会遇到很大的麻烦：从实验中得到的物理定律如果仅仅在空间中的某一点对于确定的某一方向成立，显然是毫无用处的。

唯一有些不直观的就是狭义相对论的第二条假设，毕竟它违背了我们所有的日常经验。虽然如此，至今为此所有的实验都表明这个假设是正确的。

2.1 The Invariant of Special Relativity 狭义相对论中的不变量

在接下来的几节中，我们将使用狭义相对论的两个基本假设推导出 Minkowski 度规，有了它，我们就能计算两个物理事件的“距离”。物理事件在这里指在 Minkowski 时空中的点，而整个狭义相对论都建立在 Minkowski 时空之上。我们从而得知任意两个不同惯性参考系之间的变换必须保证 Minkowski 度规不变，通过这一条件我们能够找到连接两个允许存在的惯性系（也就是光速为常数的参考系）之间所有的变换²。在本书其余的部分我们将会用到这些关于变换的知识，来寻找在这些变换下不变的方程。让我们从一个能导出一条狭义相对论假设的最重要的推论之一的思想实验开始。设想坐标系中一观察者，在原点处向上发出一个光脉冲，经过一段时间后被垂直镜面反射回原点。如图2.1所示

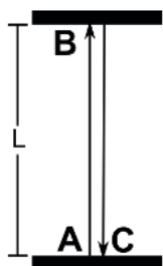


图 2.1: 思想实验图示

³ 对于恒定速度 v 而言，有 $v = \frac{\Delta s}{\Delta t}$ ， Δs 表示经过的距离， Δt 代表所需时间，因此 $\Delta t = \frac{\Delta s}{v}$

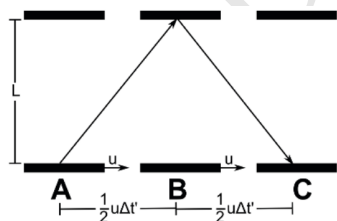


图 2.2: 思想实验图示. 第二位移动观察者向左移动，因此第一位观察者相对他向右移动

⁴ 那些能将两相对速度恒定的惯性参考系中的物理量相互转化的变换称为**推动 (Boost)** 变换，后面我们会对此进行详述。

有 3 个重要的事件：

- A: 光从原点发出
- B: 光在镜面上反射
- C: 光回到原点

事件 AC 之间的时间间隔为³

$$\Delta t = t_C - t_A = \frac{2L}{c} \quad (2.1)$$

式中 L 表示原点与反射点之间的距离。

接下来考虑第二位观察者，在 t_A 时刻处于他所在的坐标系的原点，并以恒定速度 u 相对于第一个观察者⁴向左运动。为简便起见，我们假设第二个参考系的原点在 t_A 时刻与第一位观察者的坐

²非常类似我们要找旋转变换对应的各个参数的变换规律，而旋转变换保证了长度的守恒

标系原点重合。第二位观察者所见到的现象就与第一位不一样了。在他的参考系中，光脉冲的起点和终点并不在同一位置 (见图2.2)。

用数学语言表示：

$$x'_A = 0 \neq x'_C = u\Delta t' \quad \rightarrow \quad \Delta x' = u\Delta t' \quad (2.2)$$

带撇物理量代表第二位观察者所测量的量。对于第一位静止系的观察者而言有

$$x_A = x_C \quad \rightarrow \quad \Delta x = 0 \quad (2.3)$$

假定了第二位观察者运动沿 x 轴，因此

$$y'_A = y'_C \quad z'_A = z'_C \quad \rightarrow \quad \Delta y' = 0 \quad \Delta z' = 0 \quad (2.4)$$

那么同样也有

$$y_A = y_C \quad z_A = z_C \quad \rightarrow \quad \Delta y = 0 \quad \Delta z = 0 \quad (2.5)$$

接下来的问题是：第二位观察者所测得的时间间隔是多少？因为已经假定了光速为常数，那么事件 **AC** 对于第二位观察者而言将具有不同的间隔！时间间隔 $\Delta t' = t'_C - t'_A$ ，等于光在第二位观察者的参考系中走过的距离 l 除以光速 c 。

$$\Delta t' = \frac{l}{c} \quad (2.6)$$

我们可以利用古老的 Pythagoras³定理 (见图2.2) 计算光传播的距离

$$l = 2\sqrt{\left(\frac{1}{2}u\Delta t'\right)^2 + L^2} \quad (2.7)$$

利用式(2.6)可以得到

$$c\Delta t' = 2\sqrt{\left(\frac{1}{2}u\Delta t'\right)^2 + L^2} \quad (2.8)$$

³通译为毕达哥拉斯。

再利用式(2.2)中的 $\Delta x' = u\Delta t'$ 可得

$$\begin{aligned} c\Delta t' &= 2\sqrt{\left(\frac{1}{2}u\Delta t'\right)^2 + L^2} \\ \rightarrow (c\Delta t')^2 &= 4\left(\left(\frac{1}{2}u\Delta t'\right)^2 + L^2\right) \\ \rightarrow (c\Delta t')^2 - (\Delta x')^2 &= 4\left(\left(\frac{1}{2}u\Delta t'\right)^2 + L^2\right) - (\Delta x')^2 = 4L^2 \quad (2.9) \end{aligned}$$

现在回到式(2.1)，即 $\Delta t = \frac{2L}{c}$ ，那么

$$(c\Delta t')^2 - (\Delta x')^2 = 4L^2 = (c\Delta t)^2 = (c\Delta t)^2 - \underbrace{(\Delta x)^2}_{=0 \text{ 由式(2.3)知}} \quad (2.10)$$

终于，我们得到⁵

$$(c\Delta t')^2 - (\Delta x')^2 - \underbrace{(\Delta y')^2}_{=0} - \underbrace{(\Delta z')^2}_{=0} = (c\Delta t)^2 - \underbrace{(\Delta x)^2}_{=0} - \underbrace{(\Delta y)^2}_{=0} - \underbrace{(\Delta z)^2}_{=0} \quad (2.11)$$

考虑第三个观察者，相对于第一个观察者以不同的速度运动，用同样的推理可以得到

$$(c\Delta t'')^2 - (\Delta x'')^2 - (\Delta y'')^2 - (\Delta z'')^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \quad (2.12)$$

因此，我们得到了一些对于所有观察者都相同的量：即二次型

$$(\Delta s)^2 \equiv (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 \quad (2.13)$$

此外，我们还能看出对于不同观察者， $(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$ 或者说 $(c\Delta t)^2$ 是不同的。我们将在下一节讨论不变量 Δs^2 所蕴含的物理意义。

⁵ 注意到我们在此处采用的是得到这个结果的最简方法，因为我们假定 t_A 时刻两坐标的原点是重合的。尽管如此，就算我们任意选择两惯性参考系原点的关系，仍然可以证明这个结论，只是过程会复杂一些。因为物理定律在任意惯性系中都是一样的，这给了我们任意选择便于计算的参考系的自由。如果第二位观察者的参考系运动方向是任意的，那就不再有 $\Delta y' = 0$ 和 $\Delta z' = 0$ ，但虽然如此，我们能够证明方程仍然是成立的，因为物理定律在任意惯性系中都是一样的。⁴

⁴ 译者注：原文的确把这句话说了两遍

2.2 Proper Time 固有时

我们在上一节推导出了狭义相对论中的不变量 Δs^2 ，其在所有观察者眼中都相同。接下来我们要讨论这个量的物理意义。为简便起见，我们将问题限制在一维空间中。对于一个相对于观察者静止的物体，我们能作出它的时空图 (见图2.3)。相应的，一个匀速运动的物体能作出如图2.2所示的时空图。

我们画来用于确定物体在时空中位置的线称为**世界线 (world line)**。世界线总是依赖于观察者的。两不同观察者对于同一物体可能会画出完全不同的世界线。若一观察者眼中的物体时空图为图2.2，那么对于速度与物体相同的观察者，其时空图将为图2.5，即对于此观察者物体静止。为了解释两位观察者给出的不同描述，我们引入 x' 和 t' 表示第二位观察者的参考系中的时间和位置。⁴

我们可以看到，两个观察者将对事件 AB 之间的变化持不同观点。对于第一位观察者， $\Delta x \neq 0$ ，但对于第二位观察者， $\Delta x' = 0$ 。两观察者都认为事件 A 和 B 之间的时间间隔非零： $\Delta t \neq 0$ 和 $\Delta t' \neq 0$ ，且认为 Δs^2 也相同 (见上节推导的结论，任意观察者都有相同的 Δs^2)。这将导出一个令人惊讶的结论：事件 A 到 B 经历的时间对于两观察者不同

$$(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 \quad (2.14)$$

$$(\Delta s')^2 = (c\Delta t')^2 - \underbrace{(\Delta x')^2}_{=0} = (c\Delta t')^2 \quad (2.15)$$

$$(\Delta s)^2 = (\Delta s')^2 \rightarrow (\Delta t')^2 \neq (\Delta t)^2 \quad \text{因为 } (\Delta x)^2 \neq 0 \quad (2.16)$$

这是有关狭义相对论最著名的一个效应，通常称为**钟慢效应 (time-dilation)**。时间间隔依赖于观察者，正如空间距离一样。不同观察者时间流逝速度不同，因此两事件经历的时间也就自然不同了。

现在时间将变为一个相对的概念，如果我们能找到一个对于所有观察者都相同的时间量，这将会非常有用。在上述例子中，第

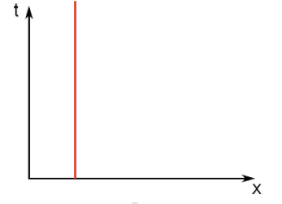


图 2.3: 静止物体的世界线。物体位置不随时间流逝而变化。

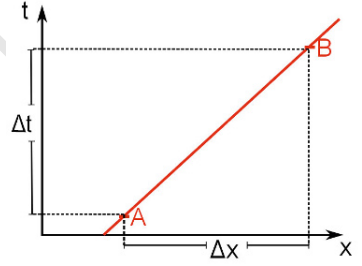


图 2.4: 某匀速直线运动物体的世界线。物体前后经过某两点记为事件 A, B 。它们间的空间距离为 Δx ，时间间隔为 Δt

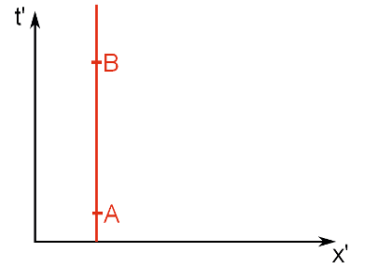


图 2.5: 同一个运动物体的世界线，但观察者与该物体速度相同。该观察者观测到的事件 AB 之间的空间距离为 $\Delta x' = 0$

⁴注意：我们这里使用的例子与节2.1不同

二位观察者与物体有相同的速度，有

$$(\Delta s')^2 = (c\Delta t')^2 \quad (2.17)$$

上式意味着狭义相对论中的不变量恰为观察者所观测到的时间间隔乘上光速 c 。这让我们有机会能用 $(\Delta s)^2$ 定义一个对于所有观察者相同的时间量。我们定义

$$(\Delta s)^2 = (c\Delta\tau)^2 \quad (2.18)$$

τ 称为**固有时 (proper time)**。固有时是所有相对于该物体静止的观察者所测量的时间。

当然，现实生活中的物体绝非只能做匀速运动，但我们可以取足够短的时间 (无穷小)，使得运动近似为匀速运动，这样固有时的概念就说得通了。因此数学上我们需将概念过渡到无穷小，即 $\Delta \rightarrow d$:

$$(ds)^2 = (cd\tau)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \quad (2.19)$$

因此，就算一个物体到处乱跑，我们也能假设一个与物体一起运动的自带时钟的观察者，这样他总与该物体相对静止。对于这个特定的观察者而言，他所测量的时间间隔就是固有时，这一数值对其余观察者也是相同的，因为 $(ds)^2 = (cd\tau)^2$ 对于任意观察者成立。再次强调，这并不意味着对所有观察者都有同样的时间间隔，只是这些观察者承认该特殊观察者所测量的时间间隔为一不变量罢了。

2.3 Upper Speed Limit 速度上限

在上节中我们对狭义相对论这一不变量有了解释，现在可以更进一步，导出另一个狭义相对论的惊人结论。

由于不变量的定义中带有负号，这意味着对于时空中的两个事件， $(\Delta s)^2$ 的值可能为 0。 $(\Delta s)^2$ 的值甚至可能是负的，但这样我们将得到一个复的固有时⁶，通常复的固有时没有物理意义。因此， $(\Delta s)^2 = 0$ 时固有时最小， $\tau = 0$ ，令

$$\Delta s_{min}^2 = 0 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

⁶ 按固有时定义, $(ds)^2 = (cd\tau)^2$, $(ds)^2 < 0 \rightarrow d\tau$ 为复数。

$$\rightarrow (c\Delta t)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

$$\rightarrow c^2 = \frac{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}{(\Delta t)^2} \quad (2.20)$$

在上式右侧我们有速度平方 v^2 ，也就是距离除上时间，把这个式子写成极限，则

$$\rightarrow c^2 = \frac{(dx)^2 + (dy)^2 + (dz)^2}{(dt)^2} \quad (2.21)$$

函数 $x(t)$, $y(t)$, $z(t)$ 是描述两事件之间路径的参数方程，这样式子右侧就是两事件间的运动速率。

因此，当某一观察者测得的固有时为 0 时其速度将满足下式

$$\rightarrow c^2 = v^2 \quad (2.22)$$

这意味着没有物体能以超过光速 c 的速度运动！⁵ 我们给物理学中的一切事物找到了一个速度上限。时空中的两事件相互影响的传播速度不能超光速。

这一点满足物理学上的局域性原理 (principle of locality)，即物理学中的对象只能被它邻近的对象影响。相互作用都是局域的，不存在超距作用，物理效应的传播需要时间。

2.4 The Minkowski Notation Minkowski 记法

从现在起，孤立的空间和孤立的时间注定要消失成为影子，只有两者的统一才能保持独立的存在。

-Hermann Minkowski⁷

写出狭义相对论中不变量

$$ds^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \quad (2.23)$$

在此处将采用一种新的记法，这种记法初学时可能稍显复杂，不过后文中将经常用到：

⁵译者注：否则其固有时为复

⁷ 出自 Hermann Minkowski 在 Assembly of German Natural Scientists and Physicians(1908.9.21) 上的演讲

$$\begin{aligned}
ds^2 &= \eta^{\mu\nu} dx_\mu dx_\nu \\
&= \eta^{00}(dx_0)^2 + \eta^{11}(dx_1)^2 + \eta^{22}(dx_2)^2 + \eta^{33}(dx_3)^2 \\
&= dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2 \\
&= (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2
\end{aligned} \tag{2.24}$$

在这里我们用了一些新的约定和记号，这些东西在近代物理中应用很广泛，所以越早熟悉越好：

- Einstein 求和约定：某指标在同一项内出现两次则代表遍历该指标并求和。例如： $\sum_{i=1}^3 a_i b_i = a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3$ ，但 $\sum_{i=1}^3 a_i b_j = a_1 b_j + a_2 b_j + a_3 b_j \neq a_i b_j$
- 像 μ, ν, σ 这样的希腊字母⁸作指标时，一般代表从 0 到 3 求和： $x_\mu y_\mu = \sum_{\mu=0}^3 x_\mu y_\mu$
- 我们约定变量 $x_0 \equiv ct$, $x_1 \equiv x$, $x_2 \equiv y$, $x_3 \equiv z$ 。这样时间空间地位等同，也便于我们使用 Einstein 求和约定
- 引入 Minkowski 度规 (η 看作一矩阵, η^{ij} 为其矩阵元): $\eta^{00} = 1$, $\eta^{11} = -1$, $\eta^{22} = -1$, $\eta^{33} = -1$, 当 $\mu \neq \nu$ 时 $\eta^{\mu\nu} = 0$ (也可以等价的写为⁹ $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$)

⁸ 像 i, j, k 这样的罗马字母作指标时一般代表从 1 到 3 求和： $x_i x_i \equiv \sum_i x_i x_i$ 。后面我们还会用到大写的罗马字母如 A, B, C ，它们作指标时代表从 1 到 8 求和。

$$\begin{aligned}
&^9 \eta^{\mu\nu} = \\
&\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}
\end{aligned}$$

另外按常规我们还要引入四维矢量 (four-vector)，简称 4

矢。

$$dx_\mu = \begin{pmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} \tag{2.25}$$

式(2.24)可用 4 矢和 Minkowski 度规表出

$$\begin{aligned}
(ds)^2 &= dx_\mu \eta^{\mu\nu} dx_\nu \\
&= \begin{pmatrix} dx_0 & dx_1 & dx_2 & dx_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} dx_0 \\ dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} \\
&= dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2
\end{aligned} \tag{2.26}$$

采用这种记法能让我们省不少事。对于 ds ，我们给出的解释是时空 (Minkowski 时空) 中两事件的“距离”，这个“距离”并不只是空间距离，它还将时间间隔考虑在内。如果我们考虑 3 维 Euclidean 空间¹⁰中两点间最短距离的平方¹¹

$$\begin{aligned}
 (ds)^2 &= dx_i \delta^{ij} dx_j \\
 &= \begin{pmatrix} dx_1 & dx_2 & dx_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} \\
 &= (ds)^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 \quad (2.27)
 \end{aligned}$$

这种数学工具叫**度规 (metric)**，它能够告诉我们无限临近的两点之间的距离。在 Euclidean 空间中度规就是单位矩阵 δ_{ij} 。在广义相对论的弯曲时空中将会有更复杂的度规，但在狭义相对论中我们采用的是相对简单的 Minkowski 度规 $\eta^{\mu\nu}$ 。度规是计算长度的工具，我们可以通过度规定义 **4 矢的长度 (length of a four-vector)**，即 4 矢与自身的标量积¹²

$$x^2 = x \cdot x \equiv x_\mu x_\nu \eta^{\mu\nu}$$

类似地，两任意 4 矢的标量积定义为

$$x \cdot y \equiv x_\mu y_\nu \eta^{\mu\nu} \quad (2.28)$$

对于上下标，也有一些约定能使计算过程更清晰。我们定义带有上指标的 4 矢¹³

$$x^\mu \equiv \eta^{\mu\nu} x_\nu \quad (2.29)$$

或者

$$y^\nu \equiv \eta^{\mu\nu} y_\mu \quad (2.30)$$

Minkowski 度规是对称的 $\eta^{\mu\nu} = \eta^{\nu\mu}$

因此，标量积可以写为¹⁴

$$x \cdot y \equiv x_\mu y_\nu \eta^{\mu\nu} = x_\mu y^\mu = x^\nu y_\nu \quad (2.31)$$

¹⁰ 3 维 Euclidean 空间就是经典物理中一般的空间，在此空间中我们将时间和空间分别处理，也就是未将时空考虑成一个整体的几何结构来考虑。但一旦将时间作为新坐标，结合出时空的观念，我们便能将时间空间放在同一坐标下考虑。

¹¹ Kronecker 函数 δ_{ij} 即是单位矩阵在指标记法下的表示，详见附录 B.5.5

¹² 此定义在 Euclidean 空间中同样适用：由于度规为 $\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ，易得矢量 \vec{v} 的长度 $= \vec{v} \cdot \vec{v} = v_1^2 + v_2^2 + v_3^2$ 。

¹³ 带有下指标的 4 矢通常叫做协变 4 矢，带有上指标的 4 矢通常叫做逆变 4 矢。

¹⁴ 指标的名称如何选取并不影响最后的值，详见附录 B.5.1。

上式中, 变换的上指标是可以任意选择的, 这只是为了避免公式中总是出现 $\eta^{\mu\nu}$ 的一种方法, 正如引入 Einstein 求和约定是为了避免总是出现求和号一样。

2.5 Lorentz Transformations Lorentz 变换

下一步, 我们将尝试找出两参考系之间不违背狭义相对论基本假设的变换方式。由上文可知, 从狭义相对论的两个基本假设可以推出对于所有惯性参考系均有 $ds^2 = \eta^{\mu\nu} dx_\mu dx_\nu$, 即

$$ds'^2 = dx'_\mu dx'_\nu \eta^{\mu\nu} = ds^2 = dx_\mu dx_\nu \eta^{\mu\nu} \quad (2.32)$$

因此, 两参考系间所允许的变换要能保证这个二次式的形式, 和 Minkowski 时空中的标量积在变换下不变。设变换为 Λ , 那么变换后

$$dx_\mu \rightarrow dx'_\mu = \Lambda_\mu^\sigma dx_\sigma \quad (2.33)$$

由于 $(ds)^2$ 在变换下不变

$$\begin{aligned} (ds)^2 &= (ds')^2 \\ &\rightarrow dx \cdot dx \stackrel{!}{=} dx' \cdot dx' \\ &\rightarrow dx_\mu dx_\nu \eta^{\mu\nu} \stackrel{!}{=} dx'_\mu dx'_\nu \eta^{\mu\nu} \stackrel{(2.33) \text{式}}{=} \underbrace{\Lambda_\mu^\sigma dx_\sigma \Lambda_\nu^\gamma dx_\gamma}_{\text{重命名哑指标}} \eta^{\mu\nu} \\ &\stackrel{!}{=} \underbrace{\Lambda_\mu^\sigma dx_\sigma \Lambda_\nu^\gamma dx_\gamma}_{\text{重命名哑指标}} \eta^{\mu\nu} \\ &\stackrel{!}{=} \underbrace{\Lambda_\mu^\sigma \Lambda_\nu^\gamma}_{dx_\nu \text{ 任意}} dx_\sigma dx_\gamma \eta^{\mu\nu} \end{aligned} \quad (2.34)$$

¹⁵ 详见附录 C.1。

或者用矩阵形式来写¹⁵

$$\eta = \Lambda^T \eta \Lambda \quad (2.35)$$

此即变换 Λ_μ^ν 所需满足条件。

这个条件现在看起来可能有些奇怪, 但在后文中我们会用更自然的方式导出这个条件。在下一章里, 我们会看到 Euclidean 空间中的旋转所导出的变换¹⁶ O 能够保证 Euclidean 空间的标量积不变¹⁷

¹⁶ 稍后会讲明 O 的意义。

¹⁷ “ \cdot ” 表示的是矢量的点乘, 如果我们将矢量写为列向量, 那么按矩阵相乘的定义有 $\vec{a} \cdot \vec{b} = \vec{a}^T \cdot \vec{b}$ 。有关 $(Oa)^T = a^T O^T$ 这一点详见附录 C.1, 式 C.3。

$$\vec{a} \cdot \vec{b} \stackrel{!}{=} \vec{a}' \cdot \vec{b}' = \underbrace{\vec{a}'^T O^T O \vec{b}}_{\text{注意 } (Oa)^T = a^T O^T} \quad (2.36)$$

因此¹⁸

$O^T \mathbf{1} O \stackrel{!}{=} \mathbf{1}$, 恰巧 Euclidean 空间的度规就是单位矩阵 $\mathbf{1}$, 其地位正如 Minkowski 度规在式(2.35)中一样。这个性质是旋转的定义中的一部分, 旋转不能改变矢量的长度, 即在数学上保证标量积不变^{19,6}此外我们还注意到旋转不改变坐标系的手向²⁰, 这表明 $\det(O) \stackrel{!}{=} 1$, 因为保证标量积不变的变换除了旋转变换还有空间反演²¹。

我们定义保 Minkowski 时空中标量积不变的变换为 **Lorentz 变换 (Lorentz transformations)**, 这也是为保证狭义相对论的假设所必需的。相应的, 每当我们得到在 Lorentz 变换下不变的项时, 都必须结合上下指标: $x_\mu y^\mu = x_\mu y_\nu \eta^{\mu\nu}$ 。在学会一些非常优雅的技术来处理这样的情况之后, 我们会在下一章构建这些变换具体的矩阵形式。

2.6 Invariance, Symmetry and Covariance 不变性, 对称性, 协变性

在继续之前, 我们有一些重要的概念要事先声明。首先, 一个量能被称为**不变量 (invariant)**, 这个量必须在变换下不变。比如说, 我们考虑一个依赖于不同参数 A, B, C, \dots 的任意函数 F , $F = F(A, B, C, \dots)$, 如果我们将 A, B, C, \dots 变换为 A', B', C', \dots , 有

$$F(A', B', C', \dots) = F(A, B, C, \dots) \quad (2.37)$$

那么 F 称为变换下的一个不变量。我们可以用对称性来描述。**对称性 (symmetry)** 是指在某一变换下或者某一系列变换下保持不变的性质。举个例子, 如果说一个物理系统在任意的旋转变换下不变, 那么该系统具有旋转对称性。再比如说, 一间室温为常温的房间, 房间内各点的温度与位置无关, 换言之, 如果把所有的点朝着某特定方向平移一段距离, 室温不变。因此我们说室温具有平移对称性。

协变性与不变性有共通之处而又不同。如果一个方程在变换

¹⁸ 这一条件一般称为**正交性 (orthogonality)** 条件, 因此用字母 O 表示。满足 $O^T O = \mathbf{1}$ 的矩阵称为正交矩阵, 因为它列与列之间是正交的。换句话说, 矩阵的每一列都可以看做一个矢量, 正交性条件就是说这些矢量相互正交。

¹⁹ 矢量的长度即是矢量与自身的标量积的平方根。

²⁰ 详见附录A.5。

²¹ 空间反演可简单地理解为映射 $\vec{x} \rightarrow -\vec{x}$ 。这种变换满足 $\det(O) \stackrel{!}{=} -1$ 及 $O^T O = \mathbf{1}$ 。因此, 如果我们要将变换限制为旋转变换, 就需要加上条件 $\det(O) \stackrel{!}{=} 1$ 。此外, 空间反演变换又称宇称变换。

⁶译者注: $\vec{a} \cdot \vec{b} = \frac{1}{2}[(\vec{a} + \vec{b})^2 - \vec{a}^2 - \vec{b}^2]$, 若保证模长不变, 标量积也不变。

下形式不变，那么称这个方程具有协变性。例如下面这个式子

$$E_1 = aA^2 + bBA' + cC^4$$

在变换之后这个方程写为

$$E'_1 = aA'^2 + bB'A' + cC'^4$$

那么这个方程具有协变性，因其在变换下形式不变。另一方程

$$E_2 = x^2 + 4axy + z$$

若在变换之后写为

$$E'_2 = y'^3 + 4az'y' + y'^2 + 8z'x'$$

那么它就不是协变的，因为其形式已经变化。

所有的物理规律都应具有 Lorentz 协变性⁷，因为只有这样的物理规律才能在任意参考系下成立。非协变的物理规律只能在某一特定参考系下成立，这会导致在东京和纽约具有不同的物理规律。显然这个主意糟透了，因为不应该存在一个特殊的参考系，我们得让物理规律在任意参考系中成立。后面我们将讲述如何用协变的方法来计算物理规律。

Further Reading Tips 阅读建议

- **E.Taylor and J.Wheeler - Spacetime Physics:Introduction to Special Relativity**²²适合用来入门。
- **D.Fleisch - A Student' s Guide to Vectors and Tensors**²³对狭义相对论中用到的张量范式有很有创意的解释。比如对于协变与逆变分量的差异。
- **N.Jeevanjee - An Introduction to Tensors and Group Theory for Physicists**²⁴是另一本把在狭义相对论中用到的数学讲的不错的书。
- **A.Zee - Einstein Gravity in a nutshell**²⁵是一本关于广义相对论的书，但是里面也有很多关于狭义相对论的很棒的解释。

⁷即在 Lorentz 变换下具有协变性---译者 (InSight)

²²Edwin F.Taylor and John Archibald Wheeler.

Spacetime Physics.

W.H.Freeman,2nd edition,

3 1992.ISBN 9780716723271

²³Daniel Fleisch.

A Student' s Guide

to Vectors and Tensors.

Cambridge University Press,

1st edition,11 2011.ISBN

9780521171908

²⁴Nadir Jeevanjee.

An Introduction to Tensors

and Group Theory for Physi-

cists.

Birkhaeuser,1st edition,August

2011.

ISBN 978-0817647148

²⁵Anthony Zee.

Einstein Gravity in a Nutshell.

Princeton University Press,

1st edition,5 2013.

ISBN 9780691145587

Symmetry Tools 对称性工具

3	Lie 群	27
3.1	群	
3.2	二维旋转	
3.3	三维旋转	
3.4	Lie 代数	
3.5	表示论	
3.6	$SU(2)$	
3.7	Lorentz 群	
3.8	Poincaré 群	
3.9	基本粒子	
3.10	附录：复向量空间中的旋转	
3.11	附录：流形	
4	理论框架	93
4.1	拉格朗日形式	
4.2	限制	
4.3	粒子理论与场论	
4.4	欧拉-拉格朗日方程	
4.5	Noether 定理	
4.6	附录：粒子理论中的推动不变性与其守恒量	
4.7	附录：场论中的推动不变性与其守恒量	

“数字是大小的度量，群是对称性的度量。”

– Mark A. Armstrong
in *Groups and Symmetry*.

"Numbers measure size,
groups measure symme-
try."

Groups and Symmetry.
Springer, 2nd edition, 1997. ISBN
9780387966755

DRAFT TRANSLATION

3. Lie Group Theory Lie 群

本章概述

本章的最终目的是导出 **Poincaré** 群双覆盖的基本表示，物理学现在认为 Poincaré 群是时空根本的对称性群。这些基本表示是描述所有基本粒子的必要工具，每一种表示对应一种基本粒子，它们揭示了自然界存在何种基本粒子。

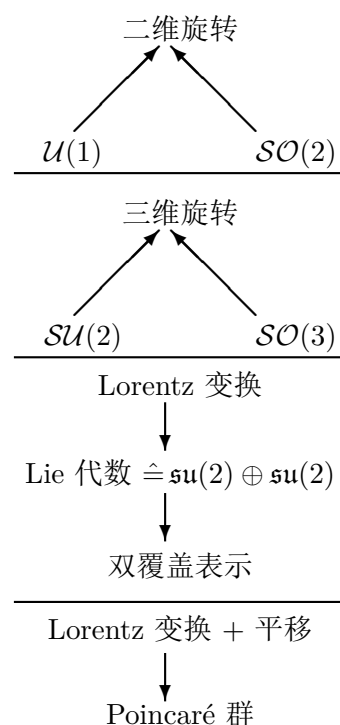
我们从两个简单例子引出群的定义，然后作为学习 Lie 群理论的第一步，我们讨论描述二维旋转变换的两种方式：

- 2×2 旋转矩阵。
- 单位复数¹。

接着我们尝试找出描述三维旋转的第二种方法（像复数那样，当然第一种方法是 3×3 矩阵），第二种方法法与一种超级重要的群 — $SU(2)$ ¹ 有关。之后我们研究 **Lie 代数**，使用相对简单的 Lie 代数就能深入研究复杂的 Lie 群。不同的群可以有相同的 Lie 代数，但只有其中的一部分是基本的。从上述基础出发就能准确揭示自然的基本对称性群 — Poincaré 群的双覆盖。我们将利用已知的变换操作导出 Lie 代数，并利用 Lie 代数得出对称变换的不同表示。这样就能看出开始时使用的表示其实只是一种特殊情况。于是又能研究 Poincaré 群的重要部分 — **Lorentz 群**，我们会看到 Lorentz 群双覆盖的 Lie 代数由两份 $\mathfrak{su}(2)$ Lie 代数所组成，因此可以直接

¹指模长为 1 的复数。

下面的图表是本章结构的示意图。当你迷失方向的时候记得回来看看，初学的时候不太需要看它。反正也看不懂



¹ S 表示特殊 (special)，它的含义为 $\det(M) = 1$ 。U 表示么正： $M^\dagger M = 1$ ，数字 2 表示这个群起初是用 2×2 矩阵定义的。

利用我们熟悉的 $SU(2)$ 群的结论。最后将平移变换考虑进来，就得到了 Poincaré 群，Poincaré 群就是 Lorentz 群加上平移。完成上述所有之后我们终于可以将 Poincaré 群双覆盖的基本表示进行分类，这些在后面的章节中会大用特用，我们将从中导出物理学的基本定律。

3.1 Groups 群

我们需要合适的数学工具描述对称性以和民科(贬义的)区分开。描述对称的数学分支称为群论，而其中物理中经常遇到的连续对称性则需要群论的一个分支 Lie 理论谎言理论来进行描述。

我们把对称性定义为变换下的不变性，而描述对称的群就定义为某些变换的集合。让我们从两个简单例子开始体会群到底该怎么定义：

1. 正方形在数学上是一些点的集合 (例如四个顶点是该集合的一部分)，正方形的对称性是指在某些变换下 (变换: 将一个点映射到另一个点) 保持不变 (点集自身被映射回自身) 的性质。²

符合条件的变换有绕中心旋转 $90^\circ, 180^\circ, 270^\circ, 0^\circ$ 等等。这些旋转操作将正方形映射到它自身。这意味着把集合里的每个点都映射回了集合里。我们称这个集合 (正方形点集) 在这样的变换下具有不变性。

注意不是所有旋转变换对于正方形都是对称的。如果我们关注顶点的变化就容易看出来，比如绕中心顺时针旋转 5° ，这个变换将顶点映射到原来正方形点集之外，顶点 A 映射到了原来正方形集合之外的点 A' 。因此这个旋转变换对正方形不具有对称性。当然，变换后的点集仍然是个正方形，但却是另一个正方形 (即不同点的集合)。而绕中心转 90° 是对称的，如图 3.3，顶点 A 变换到点 B 等等，原来的正方形点集变换到相同的集合。

²译者注：原文里把对称性定义为了一种变换，而不是在某种变换下不变的性质(a symmetry of the square is a transformation that maps this set of points into itself.)，但是后面又说 ‘Symmetry means invariance under a transformation’，这里翻译时有些改变。另外后文中不少也是意译，不妨参考原文阅读。(QZR)

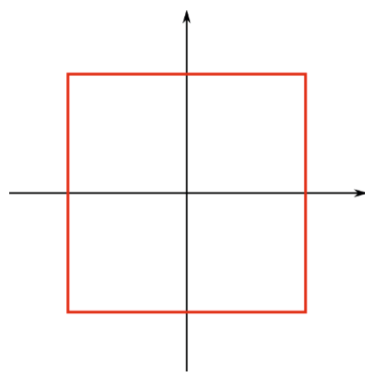


图 3.1: 正方形

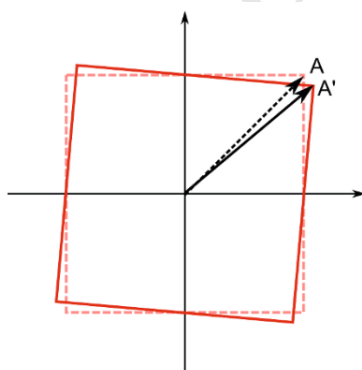
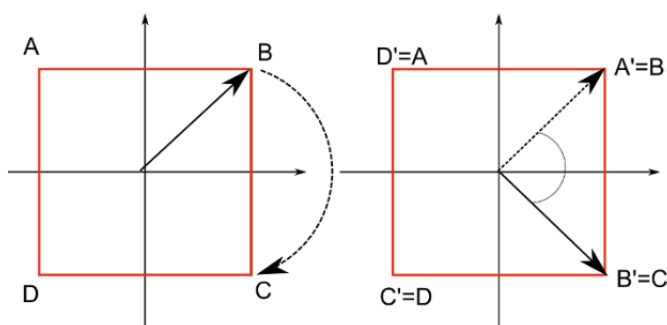


图 3.2: 正方形绕中心顺时针旋转 5°

图 3.3: 正方形绕中心旋转 90°

换个角度看：假如你看了一眼原来的正方形，然后闭上眼睛，这时有人可能把正方形做了一个变换。如果你睁眼以后不能分辨这个正方形是否被人变换过，那么这个变换就是个对称变换。

我们把所有使正方形不变的变换构成的集合称为群。变换参数（本例中就是旋转角度）不能任意取值（而是取分立的数），这样这个群称为离散群。

2. 另一个例子是使单位圆不变的变换构成的集合。类似地，单位圆还是一些点构成的集合，对称变换把这个集合映射到它自身。

单位圆绕圆心旋转任意角度都不变。换言之，变换参数（这里就是旋转角度）可以取任意值，因此这个群是连续群。

数学中除了几何图形之外还有好多别事物的都有对称性。例如，对于向量，我们可以考虑所有让向量长度不变的变换构成的集合。看出本节开头对称性定义的普遍性了没？对称性就是变换下的不变性。非常幸运，搞数学的早已建立了群论，它可以研究所有类型的对称性²。

为了让描述对称的工具——群——有确切的含义，我们把对称性的定义用数学语言精炼一下：

- 对某物什么也不做（比如转个 0° ）当然也是对称变换（按照我们的对称性定义），因此任何群都需要包含一个恒等变换（恒等元）。在上面的例子里，恒等元就是旋转 0° 。
- 对某物先做变换再做逆变换的结果必须等于啥也不做。因此对于群中的任意元素（简称群元），必须有相应的逆元素。按定义，先做变换再做逆变换等于恒等变换。例如先转 90° 再转 -90° （先逆针再顺时针转）与旋转 0° 相同。
- 先做一个对称变换再做一个对称变换，总体效果必须还是对

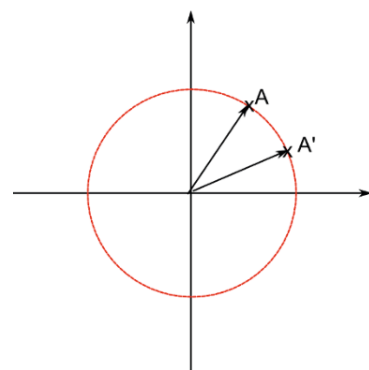


图 3.4: 单位圆旋转的示意图，单位圆绕圆心旋转任意角度都不变

² 历史上数学家们建立群论最初是为了描述方程的对称性

称变换。例如先旋转 90° 再转 180° 等于直接转 270° ，后者也是对称变换。对称变换的这个性质称为封闭性。

- 对称变换之间必须满足结合律³。例如转 90° 再转 40° 再转 110° 这种操作，无论是先转 130° 再转 110° ，还是先转 90° 再转 150° 都一样。用符号表示更加清楚：

$$R(110^\circ)R(40^\circ)R(90^\circ) = R(110^\circ)\left(R(40^\circ)R(90^\circ)\right) = R(110^\circ)R(130^\circ) \quad (3.1)$$

$$R(110^\circ)R(40^\circ)R(90^\circ) = \left(R(110^\circ)R(40^\circ)\right)R(90^\circ) = R(150^\circ)R(90^\circ) \quad (3.2)$$

这个性质称为**结合律**。

- 我们需要了解两个群元 (对称变换) 是怎样结合的。这种结合是个**二元运算**³，我们称它为**群乘法**。

在上面的例子里，旋转变换的标准表示方法是用旋转矩阵⁴，两个群元 (两个旋转矩阵，或两个旋转操作) 结合的规则是线性代数里的矩阵乘法。同样的变换经常有不同的表示方法⁵，群论非常系统性地概括了所有形式。群论的分支 — **表示论**就是研究相同变换的不同描述方式的，我们在3.5节学习表示论。

我们把上述群与对称变换的特征用严谨的数学语言表达，再将它们提升为公理，满足这些公理的一切数学结构就称为群。必须指出我们可以搞出超级抽象而又满足群公理的结构，但我们现在只关注上面例子里的旋转变换那样的群。因为我们是物理系的，而且这是本物理书

讨论完了这些，我们认识到群的抽象定义仅仅是（显然地）蕴含着这些 (我们通过对称变换的性质导出的) 群公理的东西⁶：

一个群是一个集合 G ，加上一个定义在 G 上的二元运算 (群乘法) \circ ，且要求 (G, \circ) 满足以下公理：

- 封闭性：对于任意 $g_1, g_2 \in G, g_1 \circ g_2 \in G$
- 单位元：存在单位元 $e \in G$ 使得对于所有 $g \in G, g \circ e = g =$

³两个群元结合成一个

³ 它们不一定满足交换律！例如绕不同轴的旋转变换之间不能交换顺序，比如 $R_x(\theta)R_z(\phi) \neq R_z(\phi)R_x(\theta)$

⁴ 旋转矩阵见附录A.2。

⁵ 例如，二维平面旋转还可以用单位复数描述，相应的群乘法是复数乘法。稍后就讨论这一点。

⁶ 别担心怎么才能根据这些公理凑出一个群来。搞物理的往往是从某个变换出发，考察它是否符合群公理，如果是经常是那么就应用群论来解决问题。

$$e \circ g$$

- 逆元: 对于任意 $g \in G$, 存在相应的逆元 $g^{-1} \in G$ 使得

$$g \circ g^{-1} = e = g^{-1} \circ g$$

- 结合律: 对于任意 $g_1, g_2, g_3 \in G, g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$

总而言之: 某对象在一些变换下保持不变, 所有的这些变换组成⁴的集合叫做对称群。对于 Minkowski 时空, Minkowski 度规⁷在变换下保持不变, 相应的对称群称为 Poincaré 群。

要注意群的定义完全与变换的对象是啥没有关系。因此我们可以脱离特定对象而只研究对称变换本身, 群的定义将变换从对象中‘提取’出来了。这是非常有用的, 许多不同事物具有同样或同类的对称性。群论让我们不用管变换的对象⁵, 只研究变换 (例如旋转) 的普遍性质。

⁷ 复习一下, Minkowski 度规就是在 Minkowski 空间中用来计算距离和长度的工具, 见第2章。

3.2 Rotations in two Dimensions 二维旋转

我们从一个简单而重要的例子开始学习群论。考虑那些让二维平面中的向量长度不变的对称变换。符合条件的有旋转和反射⁸。它们也是圆的对称变换, 一个单位元旋转或反射之后还是单位圆。可见同一个群 (对应一种变换) 可以作用于不同类对象: 圆 (几何图形), 或者向量⁶。本节考虑向量, 可以用旋转矩阵表示向量的旋转或反射⁹, 将起点位于原点的向量绕原点逆时针旋转 θ 角度的旋转矩阵为⁷

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (3.3)$$

关于 x 轴与 y 轴的反射变换用矩阵表示为:

$$P_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad P_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.4)$$

将矩阵乘法作为群乘法 \circ , 可以验证对于所有的 θ , $R(\theta), P_x, P_y$ 与矩阵乘法符合群公理, 因而构成一个群, 亦即旋转与反射变换构成一个群。

⁴我在想这里该用‘构成’还是‘组成’, 后来我觉得无所谓, 因为这不是初中化学...(分子构成物质, 元素组成物质...) — 译者 (SI)

⁵译者注: 圆还是正方形

⁶可以简单考虑为一个点, 其相对于原点就是该向量。---译者注

⁷译者注: 原文此式符号有误。

⁸ 当然还有平移。平移的数学描述与前两者有些不同, 我们之后再讨论它。

⁹ 旋转矩阵的推导见附录A.2。

可以用更抽象的方式表达‘所有让二维向量长度不变的变换’，向量长度就是向量与自身点乘的平方根 ($|a| = \sqrt{a \cdot a}$)。向量长度在变换 $a \rightarrow a'$ 下不变意味着⁸：

$$a' \cdot a' \stackrel{!}{=} a \cdot a \quad (3.5)$$

将线性变换对应的矩阵记作 \mathcal{O} ，变换即为 $a \rightarrow a' = \mathcal{O}a$ 。

$$a \cdot a = a^T a \rightarrow a'^T a' = (\mathcal{O}a)^T \mathcal{O}a = a^T \mathcal{O}^T \mathcal{O}a \stackrel{!}{=} a^T a = a \cdot a \quad (3.6)$$

由此可见，使向量长度不变的变换必须满足⁹：

$$\mathcal{O}^T \mathcal{O} = I \quad (3.7)$$

其中 I 表示单位矩阵。¹⁰。前文的旋转和反射矩阵都满足此条件¹¹ 2×2 矩阵中所有满足(3.7)式的矩阵构成了 $\mathcal{O}(2)$ 群，即所有 2×2 正交矩阵构成的群¹²。我们可以找出这个群中描述旋转变换的那一部分（构成一个子群）。根据(3.7)式：

$$\begin{aligned} \det(\mathcal{O}^T \mathcal{O}) &= \det(\mathcal{O}^T) \det(\mathcal{O}) \stackrel{!}{=} \det(I) = 1 \\ (\det(\mathcal{O}))^2 &\stackrel{!}{=} 1 \rightarrow \det(\mathcal{O}) \stackrel{!}{=} \pm 1 \end{aligned} \quad (3.8)$$

$\det \mathcal{O} = 1$ 的矩阵对应旋转变换¹³。而条件

$$\mathcal{O}^T \mathcal{O} = I \quad (3.9)$$

$$\det \mathcal{O} = 1 \quad (3.10)$$

定义了 $\mathcal{SO}(2)$ 群，这里的 \mathcal{S} 表示‘特殊’ (special)， \mathcal{O} 表示‘正交’ (orthogonal)。

$\mathcal{SO}(2)$ 包含的旋转变换保持了坐标系的取向，即一个右手坐标系¹⁴经旋转变换后还是右手系，而反射变换会改变它的取向。用线性代数的概念来说，我们规定 $\mathcal{SO}(2)$ 中的矩阵行列式必须为 +1。

⁸等号上面的! 只是起强调、提示作用 — 译者 (SI)

⁹其中是由于 a 任意从而能够得到 $\mathcal{O}^T \mathcal{O} = I$ 。

¹⁰

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

¹¹ 例如(3.3)式的矩阵， $R_\theta^T R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & 0 \\ 0 & \sin^2 \theta + \cos^2 \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

¹² 严谨地说，任意 2×2 正交矩阵可以表示为(3.3)、(3.4)式，或者它们的乘积的形式。

¹³ 见(3.3)与(3.4)式。反射矩阵的行列式等于 -1。

¹⁴ 右手坐标系与左手坐标系见附录A.5。

3.2.1 Rotations with Unit Complex Numbers 用单位复数表示旋转变换

还可以用单位复数来表示二维的旋转变换：绕原点旋转 θ 角可以表示为乘以单位复数 z (单位复数意为 $z = a+ib, |z|^2 = z^*z = 1$)¹⁵。

所有单位复数构成一个群，称为¹⁶ $\mathcal{U}(1)$ ，群乘法即为复数乘法，不难验证它符合群公理。为了看出它与之前引入的 $\mathcal{O}(2), \mathcal{SO}(2)$ 群¹⁰的关系，我们将 $\mathcal{U}(1)$ 群的条件 — 复数的模（长度）为 1 表示为¹⁷， $\forall \mathcal{U} \in \mathcal{U}(1)$ ：

$$\mathcal{U}^*\mathcal{U} = 1 \quad (3.11)$$

单位复数另一种写法是¹⁸：

$$R_\theta = e^{i\theta} = \cos \theta + i \sin \theta \quad (3.12)$$

R_θ 的模平方为：

$$R_\theta^* R_\theta = e^{-i\theta} e^{i\theta} = (\cos \theta - i \sin \theta)(\cos \theta + i \sin \theta) = 1 \quad (3.13)$$

例：将向量 $(3, 5)$ 对应的复数 $z = 3 + 5i$ 旋转 90° ：

$$z \rightarrow z' = e^{i90^\circ} z = \left(\underbrace{\cos 90^\circ}_{=0} + i \underbrace{\sin 90^\circ}_{=1} \right) (3+5i) = i(3+5i) = 3i-5 \quad (3.14)$$

图¹⁸绘制了旋转前后的向量（复数），图中可见复数乘以 e^{i90° 旋转了 90° 。需要注意 e^{i90° 作用于（乘以）向量对应的复数而非向量本身。描述二维旋转只需要一个参数：旋转角 ϕ 。而复数有两个自由度（实部和虚部），因此我们加上单位复数的限制： $|z| = \text{实部}^2 + \text{虚部}^2 = 1$ ，刚好只剩一个自由度。

描述二维旋转的两种方式 — 单位复数与 2×2 矩阵（矩阵元都是实数）通过如下方式相联系。定义：

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (3.15)$$

¹⁰译者注：原文有误，现已更正

¹⁵ 上标 * 表示复数的共轭复数：

$z = a + ib \rightarrow z^* = a - ib$

¹⁶ \mathcal{U} 表示么正性，它的含义为 $\mathcal{U}^*\mathcal{U} = 1$ 。

¹⁷ 定义中包含复数乘法的群的普遍性质见3.10节的附录。

¹⁸ 附录B.4.2推导了传说中的欧拉公式。对任意复数 $z = a+ib$, a 称为 z 的实部，记为 $\text{Re}(z) = a$ ； b 称为虚部，记为 $\text{Im}(z) = b$ 。欧拉公式中 R_θ 的实部为 $\cos \theta$ ，虚部为 $\sin \theta$ 。

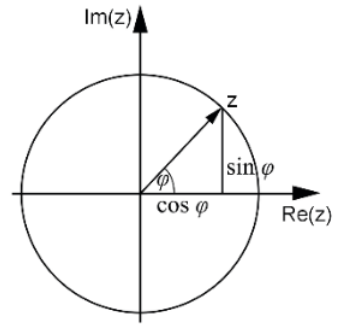


图 3.5: 单位复数在复平面的单位圆上

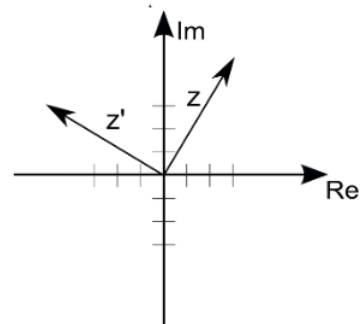


图 3.6: 复数通过乘以单位复数进行旋转

不难验证这样定义的 $\mathbf{1}, \mathbf{i}$ 仍然满足:

$$\mathbf{1}^2 = \mathbf{1}, \quad \mathbf{i}^2 = -\mathbf{1}, \quad \mathbf{1}\mathbf{i} = \mathbf{i}\mathbf{1} = \mathbf{i} \quad (3.16)$$

这样单位复数对应的 2×2 实数矩阵为¹¹

$$R_\theta = \cos \theta + i \sin \theta = \cos \theta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (3.17)$$

可见, 定义了复单位 $i \rightarrow$ 实矩阵 的对应关系后, 单位复数就回到了熟悉的旋转矩阵。还有一点要注意: 旋转矩阵作用于向量 (列矩阵), 而我们将 i 与 2×2 的实矩阵相联系, 因此被旋转的向量 (与一个复数对应) 也将变成一个 2×2 矩阵。

例: 任意向量 (a, b) 对应的复数对应的 2×2 矩阵为:

$$z = a + ib = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad (3.18)$$

将旋转矩阵作用于 z :

$$\begin{aligned} z' &= \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = R_\theta z = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \\ &= \begin{pmatrix} a \cos \theta - b \sin \theta & -b \cos \theta - a \sin \theta \\ a \sin \theta + b \cos \theta & -b \sin \theta + a \cos \theta \end{pmatrix} \end{aligned} \quad (3.19)$$

由上式可得

$$a' = a \cos \theta - b \sin \theta \quad (3.20)$$

$$b' = a \sin \theta + b \cos \theta \quad (3.21)$$

这与旋转矩阵作用于向量 (列矩阵形式) 相同:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \cos \theta - b \sin \theta \\ a \sin \theta + b \cos \theta \end{pmatrix} = \begin{pmatrix} a' \\ b' \end{pmatrix} \quad (3.22)$$

我们看到这两种表示方法是一回事儿, 用数学术语来说, $\mathcal{SO}(2)$ 与

¹¹译者注: 原文此式有误, 译文已参照勘误表修改。

$\mathcal{U}(1)$ 间有一个同构映射¹⁹。这一点太重要了，之后的篇幅会不断出现这种思想。

下面研究三维旋转，尝试找出它的两种描述方法²⁰。

3.3 Rotations in three Dimensions 三维旋转

三维向量旋转变换当然可以用 3×3 旋转矩阵¹²：

$$\begin{aligned} R_x &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} & R_y &= \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \\ R_z &= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (3.23)$$

类比 $\mathcal{SO}(2)$ 群，上面三个矩阵构成了 $\mathcal{SO}(3)$ 群的一组基²¹。这意味着 $\mathcal{SO}(3)$ 中的任意群元（矩阵）都可以写为 R_x, R_y, R_z 的线性组合，且系数唯一。将向量

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

绕 z 轴旋转²²就是用旋转矩阵乘以向量：

$$R_z(\theta)\vec{v} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix} \quad (3.24)$$

¹³ 为找出描述三维旋转的第二法，我们尝试把复数的概念扩展到三维空间。首先试着把 2 维的复数¹⁴扩展为 3 维复数，但前人探索发现没有 3 维的，但存在 4 维‘复数’，称为四元数 (quaternions)。下面就会看到四元数正是描述三维旋转的第二法，并且它还能深刻揭示宇宙 — 四维时空的规律。描述三维旋转需要三个参数（绕 x, y, z 轴的旋转角），而四元数有四个参数，这与二维旋转的情况

¹⁹ 如果映射 $\Pi : G \rightarrow G'$ 是一一映射，并且满足： $\forall g_1, g_2 \in G, \Pi(g_1) \circ \Pi(g_2) = \Pi(g_1 \circ g_2)$ ，则 Π 就是同构映射，并且称群 G, G' 是同构的。

²⁰ 其中蕴含着十分硬却失挺 (interesting) 的内容... 与二维旋转类似，我们将找出三维旋转的第二法，这个第二法能够揭示大自然更深刻的内容。

²¹ 基的概念见附录 A.1。

²² 三维向量旋转的一般情况的推导见附录 A.2。

¹² 原文公式有误，译文已按勘误表改正。

¹³ 译者注：原文符号有误，现已订正。

¹⁴ 实部虚部两个自由度

²³ 单位复数描述二维旋转。

类似²³。四元数加上单位长度的限制 — 单位四元数（三个自由参数）能够描述三维旋转。

3.3.1 Quaternions 四元数

我们类比复数来构造四元数，复数只有一个虚单位 i ，而四元数里定义三个虚单位，记作 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ ，它们仍然满足

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1 \quad (3.25)$$

这样任意一个四元数 q 表示为

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \quad (3.26)$$

其中 $a, b, c, d \in \mathbb{R}$, $\mathbf{1}$ 就是正常的实数 1。¹⁵

要想定义四元数间的乘法，必须定义虚单位间的乘法规则，比如 $\mathbf{ij} = ?$ 。为此再定义

$$\mathbf{ijk} = -1 \quad (3.27)$$

这样虚单位间的乘法就没问题了，例如，为导出 $\mathbf{ij} = \mathbf{k}$ ，在(3.27)式等号两侧右乘 \mathbf{k} ：

$$\mathbf{ij} \underbrace{\mathbf{kk}}_{=-1} = -\mathbf{k} \rightarrow \mathbf{ij} = \mathbf{k} \quad (3.28)$$

²⁴ 符号 \dagger 称为 ‘dagger’，表示转置加取复共轭，即 $a^\dagger = (a^*)^T$ 。实数向量间的点乘定义为 $a \cdot b = a^T b$ ，向量用列矩阵表示，其转置为行矩阵，于是点乘的定义符合矩阵乘法规则。四元数对应的列矩阵含复数，为使四元数与它自身相乘结果为实数（实数结果可表示 ‘长度’），在转置之外还加上取复共轭。

单位四元数 $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ 满足²⁴：

$$\begin{aligned} q^\dagger q &\stackrel{!}{=} 1 \\ \rightarrow (a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k})(a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) &= a^2 + b^2 + c^2 + d^2 \stackrel{!}{=} 1 \end{aligned} \quad (3.29)$$

就像单位复数构成一个群（群乘法为复数乘法）那样，单位四元数也构成一个群（群乘法为四元数乘法）。

类似于我们将二维复数与 2×2 实数矩阵建立联系的做法，四元数也可如此 — 将四个基 $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ 与合适的矩阵一一对应。其中

¹⁵ $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ 用黑体表示是为了强调它们是四元数的基。---译者 (SI)

一种方式（还有别的方式）是将它们与 2×2 复数矩阵对应¹⁶：

$$\begin{aligned} \mathbf{1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \mathbf{i} &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \mathbf{j} &= \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, & \mathbf{k} &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \end{aligned} \quad (3.30)$$

不难验证上式满足四元数乘法，比如(3.25)、(3.27)式。这样任意一个四元数都可用矩阵表示为：

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = \begin{pmatrix} a + di & b + ci \\ -b + ci & a - di \end{pmatrix} \quad (3.31)$$

由上式可见

$$\det(q) = a^2 + b^2 + c^2 + d^2 \quad (3.32)$$

上式与(3.29)式比较可以发现单位四元数对应的矩阵的行列式为1。因此单位四元数对应的 2×2 复数矩阵 \mathcal{U} 满足：

$$\mathcal{U}^\dagger \mathcal{U} = 1, \quad \text{且 } \det(\mathcal{U}) = 1 \quad (3.33)$$

注意(3.30)式中的矩阵是线性独立的²⁵，它们构成群 $SU(2)$ 的基。与之前相同， \mathcal{S} 表示特殊 (special)，含义为 $\det(\mathcal{U}) = 1$ ； \mathcal{U} 表示幺正 (unitary)²⁶，即 $\mathcal{U}^\dagger \mathcal{U} = 1$ 。任意单位四元数都唯一对应 $SU(2)$ 群中的一个群元。

$SU(2)$ 群如何与三维旋转联系？和之前不同， $SU(2)$ 群与 $SO(3)$ ²⁷ 之间的对应关系并不像二维旋转中 $U(1)$ 与 $SO(2)$ 那样简单。

2 维虚数 $z = a + ib$ 与 2 维向量很容易对应。单位复数保证了矢量长度在旋转变换下不变²⁸： $(Rz)^* R z = z^* R^* R z = z^* z$ 。四元数有 4 个参数，它与三维向量的对应关系是不明显。我们尝试将三维向量 $\vec{v} = (x, y, z)$ 定义为如下四元数¹⁷：

$$v \equiv x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad (3.34)$$

¹⁶译者注：原文此式及后续(3.31)、(3.38)、(3.39)、(3.43)和(3.44)均有误，已据勘误表订正。

¹⁷注意下式中的 \mathbf{ijk} 是四元数的‘基’而非三维直角坐标系的基矢。——译者

²⁵ 线性独立的概念见附录A.1。

²⁶ 进一步介绍见本章附录 — 3.10节。

²⁷ $SO(3)$ 群就是三维旋转矩阵构成的群。

²⁸ 注意这里的 R 表示一个单位复数，即 $R^* R = 1$ ， R 与复数 z 相乘就将 z 进行旋转变换。

利用前述四元数的矩阵表示可得：

$$\det(v) = x^2 + y^2 + z^2 \quad (3.35)$$

于是保持向量 (x, y, z) 长度不变的变换就对应于保持矩阵行列式不变的矩阵变换。而单位四元数对应矩阵都具有单位行列式，它与任意矩阵相乘不改变该矩阵的行列式¹⁸。进展很顺利，但现在有个微妙的状况：继续猜想下去，我们会认为用单位四元数 u 旋转向量（对应的四元数） v 直接用 u 乘以 v 就可以了（按照四元数乘法规则）。但细想起来这却不行，因为根据(3.34)式，我们把任意三维向量对应的四元数定义在 $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ 的范围内， u 与 v 的四元数乘积会超出这个范围，这样向量旋转后可能不是个向量！这当然不行。

事实上，旋转变换表示为：

$$v' = q^{-1}vq \quad (3.36)$$

其中 v, v' 分别为旋转前后的向量， q 为表示旋转的单位四元数。

这样我们终于实现了描述三维旋转的第二法 — 单位四元数。

例：定义 u 是 $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ 中的某一单位向量，则任意单位四元数 t 可表示为：

$$t = \cos \theta + u \sin \theta \quad (3.37)$$

利用(3.34)式，任意三维向量 \vec{v} 可表示为：

$$\vec{v} = (v_x, v_y, v_z)^T = v_x \mathbf{i} + v_y \mathbf{j} + v_z \mathbf{k} \stackrel{(3.31) \text{式}}{=} \begin{pmatrix} iv_z & v_x + iv_y \\ -v_x + iv_y & -iv_z \end{pmatrix} \quad (3.38)$$

简明起见我们举个特例，把向量 $\vec{v} = (1, 0, 0)^T$ 绕 z 轴旋转。

$$\vec{v} = (1, 0, 0)^T \rightarrow v = 1\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (3.39)$$

¹⁸利用 $\det(AB) = \det(A)\det(B)$ ，— 译者

$$R_z(\theta) = \cos \theta \mathbf{1} + \sin \theta \mathbf{k} = \begin{pmatrix} \cos \theta + i \sin \theta & 0 \\ 0 & \cos \theta - i \sin \theta \end{pmatrix} \quad (3.40)$$

上式也可用欧拉公式写出²⁹：

²⁹ 欧拉公式的推导见附录B.4.2

$$e^{ix} = \cos x + i \sin x \Rightarrow R_z(\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \quad (3.41)$$

四元数旋转矩阵的逆矩阵为：

$$R_z^{-1}(\theta) = \begin{pmatrix} \cos \theta - i \sin \theta & 0 \\ 0 & \cos \theta + i \sin \theta \end{pmatrix} = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \quad (3.42)$$

根据(3.36)式旋转向量 v ：

$$\begin{aligned} v' &= R_z^{-1}(\theta) v R_z(\theta) = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \\ &= \begin{pmatrix} 0 & e^{-i2\theta} \\ -e^{i2\theta} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \cos(2\theta) - i \sin(2\theta) \\ -\cos(2\theta) - i \sin(2\theta) & 0 \end{pmatrix} \end{aligned} \quad (3.43)$$

另一方面，向量 v' 用四元数表示为：

$$v' = \begin{pmatrix} iv'_z & v'_x + iv'_y \\ -v'_x + iv'_y & -iv'_z \end{pmatrix} \quad (3.44)$$

上式与(3.43)式比较可得：

$$v'_x = \cos(2\theta), \quad v'_y = -\sin(2\theta), \quad v'_z = 0 \quad (3.45)$$

所以

$$\rightarrow \vec{v'} = (\cos(2\theta), -\sin(2\theta), 0)^T \quad (3.46)$$

由上式可见从 \vec{v} 到 $\vec{v'}$ 确实进行了旋转³⁰，但是要注意，上式表示我们没有把 \vec{v} 旋转 θ 角度，而是旋转了 2θ ！因此我们定义 $\phi \equiv 2\theta$ ，

³⁰ 见(3.24)式，那里我们用三维旋转矩阵旋转向量。

等变换（恒等变换：什么都不干）的群元。相对的，离散群的群元个数是可数的，不存在无限接近于恒等元的群元。例如把正方形绕中心转 0.000001° ，这与恒等变换（转 0° ）非常接近，但它不是正方形的恒等变换。而把圆绕圆心转 0.000001° 就是对称的。圆的对称群（对称变换组成的群）是连续的，因为变换参数（旋转角）可以取任意（连续）的值。用数学符号来表示：把恒等元记作 I ，无限接近恒等元的群元 g 可表示为：

$$g(\epsilon) = I + \epsilon X \quad (3.48)$$

其中 ϵ 表示小量（数学上总是用 ϵ 表示小量）， X 称为生成元，稍后就讨论它。这样一个轻微的小变换作用在对象上几乎啥也不变，有时称 $g(\epsilon)$ 为无穷小变换。把无穷小变换重复许多次就能得到一个有限大小的变换。以旋转为例：许多朝同一方向小旋转等效于一次有限大的旋转。用数学语言来说，我们可以将无穷小变换重复许多次：

$$h(\theta) = \underbrace{(I + \epsilon X)(I + \epsilon X) \dots (I + \epsilon X)}_{k \uparrow (I + \epsilon X)} = (I + \epsilon X)^k \quad (3.49)$$

其中 k 表示无穷小变换重复的次数。如果 θ 表示有限大的旋转角，比如 50° 什么的，然后 N 表示超大的数，则无限接近于恒等变换的旋转变换可表示为：

$$g\left(\frac{\theta}{N}\right) = I + \frac{\theta}{N} X \quad (3.50)$$

要使上式表示的变换尽可能小，则让 N 尽可能大，令 $N \rightarrow \infty$ 。为了从这样一个无穷小变换得到一个有限变换，需要把无穷小变换重复无限多次，即：

$$h(\theta) = \lim_{N \rightarrow \infty} \left(I + \frac{\theta}{N} X \right)^N \quad (3.51)$$

微积分告诉我们这个极限就是指数函数³⁴：

$$h(\theta) = \lim_{N \rightarrow \infty} \left(I + \frac{\theta}{N} X \right)^N = e^{\theta X} \quad (3.52)$$

³⁴(3.51)式经常用作指数函数的定义。指数的级数表示与(3.51)式等价性的证明见附录B.4.1。这在几乎所有的数学分析课本中都能找到。

上式有 X 生成了有限变换 $h(\theta)$ 的感觉, 因此 X 成为生成元。生成元的严格定义之后详述, 下面先从另一个角度看这个问题。

考虑某个用矩阵表示的连续变换群, 对任意群元, 在恒等元 I 处做 Taylor 展开³⁵。

³⁵ 如果你从未听说过 Taylor 展开或 Taylor 级数, 作者推荐你看一看附录B.3

$$h(\theta) = I + \frac{1}{1!} \frac{dh}{d\theta} \Big|_{\theta=0} \theta + \frac{1}{2!} \frac{d^2h}{d\theta^2} \Big|_{\theta=0} \theta^2 + \cdots = \sum_n \frac{1}{n!} \frac{d^n h}{d\theta^n} \Big|_{\theta=0} \theta^n \quad (3.53)$$

利用指数函数的级数展开可将上式的级数表示为更紧凑的形式³⁶: ²⁰

³⁶ 附录B.4.1中有推导

$$h(\theta) = \exp \left(\frac{dh}{d\theta} \Big|_{\theta=0} \theta \right) \equiv \sum_n \frac{1}{n!} \frac{d^n h}{d\theta^n} \Big|_{\theta=0} \theta^n \quad (3.54)$$

这与之前的描述有联系。比较上式与(3.52)式可得:

$$X = \frac{dh}{d\theta} \Big|_{\theta=0} \quad (3.55)$$

蕴含在上述推导中的思想是: 通过研究群中重要的无穷小元素 — 生成元²¹ (上面的 X) 就能获得群的许多重要信息。

矩阵 Lie 群 (记作 G) 的 Lie 代数就是下面的集合:

$$\left\{ X \mid \text{若 } e^X \in G \right\}$$

即如果 X 满足 e^X 是 G 的群元, 则 X 就是 G 的 Lie 代数 (它是一个集合!) 中的元素。这个简单定义只对矩阵 Lie 群管用。之后我们引入 Lie 代数的一般定义。

³⁷ 群 G 的 Lie 代数通常用哥特体活字 (Fraktur) 字母 \mathfrak{g} 表示。

上面的矩阵 Lie 群的 Lie 代数的定义用数学语言表述为³⁷:

$n \times n$ 矩阵 Lie 群 G 的 Lie 代数 \mathfrak{g} 是满足如下条件的 $n \times n$ 矩阵 X 的集合:

$$e^{tX} \in G, t \in \mathbb{R}$$

根据群的定义, 群无非是一些变换的集合。群乘法 \circ 告诉我们群元之间怎样结合。矩阵 Lie 群的群乘法就是矩阵乘法, 我们可能 naive 地认为 Lie 代数的元素之间采用相同的方式结合 (\circ), 但这

²⁰原文下式有误, 译文已按勘误表修改。

²¹注意生成元 X 并非群元, $I + \epsilon X$ 才是。

并不正确！诚然，（矩阵 Lie 群的）Lie 代数的元素都是矩阵³⁸，但两个 Lie 代数的矩阵乘法的结果往往不再是 Lie 代数的元素。Lie 代数元素间有另外的结合规则，当然它与原来群的群乘法直接有关。

Lie 群乘法与 Lie 代数结合法则之间的关系由著名的 Baker-Campbell-Hausdorff 公式（以下简称 BCH 公式）给出³⁹：

$$e^X \circ e^Y = e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots} \quad (3.56)$$

其中⁴⁰ $X, Y \in \mathfrak{g}$ ，即 X, Y 是群 G 的生成元。 e^X, e^Y 是 G 的群元，把它们分别记作 g, h ，这样上式写成：

$$\underbrace{g}_{\in G} \circ \underbrace{h}_{\in G} = e^X \circ e^Y = \underbrace{e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\dots}}_{\in G} \quad (3.57)$$

上式等号右侧是群的一个群元，可见两个群元 (g, h) 相乘可表示为一些 Lie 代数元素的和（再取指数）²²。上两式出现的新运算 $[\cdot]$ 称为 **Lie 括号**，对于矩阵 Lie 群， $[X, Y] = XY - YX$ ， $[X, Y]$ 称为 X, Y 的对易子。注意 XY 和 YX 一般不是 Lie 代数的元素，但它们的差一定是⁴¹！

由 BCH 公式可知 Lie 代数元素间的乘法规则是 Lie 括号 $[\cdot]$ ，而非最初认为的矩阵乘法。就像群在群乘法 \circ 运算（矩阵群的 \circ 就是矩阵乘法）下封闭那样，我们称 Lie 代数在 Lie 括号运算下封闭。集合在某运算下封闭的含义是集合的任意两个元素进行该运算所得结果仍在此集合内⁴²。

在学习一个 Lie 代数的例子之后，我们将讨论 Lie 代数的现代定义。现代定义是从群的生成元在 Lie 括号运算下的行为来定义的。利用更广泛的现代定义就能看出哪些不同的群具有**相同的** Lie 代数，只从上面的 Lie 代数定义出发看不出这一点来。因此 Lie 代数的新定义能让我们更深刻地认识一种变换的基本特征。同一 Lie 代数可对应许多 Lie 群，Lie 理论的一条重要定理告诉我们这其中有一个**特别的** Lie 群。引入 Lie 群的现代定义之后上面所说的就具体起来了。

下面我们就举一个从给定的群导出相应 Lie 代数的例子。

³⁸ Lie 群理论的著名定理 — Ado 定理 (Ado's Theorem) 告诉我们任意 Lie 代数与矩阵 Lie 代数同构。

³⁹ 我们不讲这个公式的证明，在大多数关于 Lie 理论的教材都能找到，比如 William Fulton and Joe Harris. Representation Theory: A First Course. Springer, 1st corrected edition, 8 1999. ISBN 9780387974958

⁴⁰ 重复一下，群 G 的 Lie 代数通常用哥特体字母 \mathfrak{g} 表示。

⁴¹ John Stillwell.

Naive Lie Theory. Springer, 1st edition, August 2008a. ISBN 978-0387782140 里面有漂亮的证明。

⁴² 用数学语言表示封闭性： $\forall g, h \in G, g \circ h \in G$ 。对于 $X, Y \in \mathfrak{g}, [X, Y] \in \mathfrak{g}, X \circ Y \notin \mathfrak{g}$

²² 正如后面所说，这句话的正确性是由于(3.57)中的诸如 $[X, Y], [X, [X, Y]]$ 都是 Lie 代数元素。---译者注

3.4.1 The Generators and Lie Algebra of $\mathcal{SO}(3)$ $\mathcal{SO}(3)$ 群的生成元与 Lie 代数

$\mathcal{SO}(3)$ 群的定义为 ((3.10)式):

$$\mathcal{O}^T \mathcal{O} \stackrel{!}{=} I, \quad \det(\mathcal{O}) \stackrel{!}{=} 1 \quad (3.58)$$

任意群元 \mathcal{O} 用相应的生成元 J 表示为:

$$\mathcal{O} = e^{\phi J} \quad (3.59)$$

上式带入群的第一个定义条件得:

$$\mathcal{O}^T \mathcal{O} = e^{\phi J^T} e^{\phi J} \stackrel{!}{=} 1 \quad \rightarrow \quad J^T + J \stackrel{!}{=} 0 \quad (3.60)$$

⁴³ $\text{tr}(A)$ 表示矩阵 A 的迹, 也就是 A 的主对角线上所有元素的和。例如

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

的迹为 $\text{tr}(A) = A_{11} + A_{22}$ 。

⁴⁴ 线性无关的概念见附录A.1

带入第二个定义条件, 再利用等式⁴³ $\det(e^A) = e^{\text{tr}(A)}$ 得:

$$\begin{aligned} \det(\mathcal{O}) \stackrel{!}{=} 1 &\quad \rightarrow \quad \det(e^{\phi J}) = e^{\phi \text{tr}(J)} \stackrel{!}{=} 1 \\ &\quad \rightarrow \quad \text{tr}(J) \stackrel{!}{=} 0 \end{aligned} \quad (3.61)$$

满足(3.60)、(3.61)式的三个线性无关的⁴⁴矩阵为:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.62)$$

⁴⁵Levi-Civita 符号的含义见附录B.5.5

这三个矩阵构成 $\mathcal{SO}(3)$ 群的生成元的基。即群的任意生成元 J 可惟一表示为这三者的线性组合: $J = aJ_1 + bJ_2 + cJ_3$, 其中 a, b, c 表示实常数。 J_1, J_2, J_3 可用 Levi-Civita 符号更紧凑地表示⁴⁵:

$$(J_i)_{jk} = -\epsilon_{ijk}, \quad i, j, k = 1, 2, 3. \quad (3.63)$$

其中 j, k 表示生成元 J_i 的分量。例:

$$(J_1)_{jk} = -\epsilon_{1jk} \longleftrightarrow \begin{pmatrix} (J_1)_{11} & (J_1)_{12} & (J_1)_{13} \\ (J_1)_{21} & (J_1)_{22} & (J_1)_{23} \\ (J_1)_{31} & (J_1)_{32} & (J_1)_{33} \end{pmatrix}$$

$$= - \begin{pmatrix} \epsilon_{111} & \epsilon_{112} & \epsilon_{113} \\ \epsilon_{121} & \epsilon_{122} & \epsilon_{123} \\ \epsilon_{131} & \epsilon_{132} & \epsilon_{133} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (3.64)$$

这些生成元可以生成有限大小的变换矩阵。以 J_1 为例，我们可以只关注非零部分 — 右下角的 2×2 矩阵，把它记作 j_1 ⁴⁶。

⁴⁶ 其实子矩阵 j_1 也可用二维 Levi-Civita 符号表示, $(j_1)_{ij} = \epsilon_{ij}$ 。见附录 B.5.5。 j_1 是二维旋转群 $SO(2)$ 的生成元。

$$J_1 = \begin{pmatrix} 0 & & \\ & \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\equiv j_1} & \\ & & \end{pmatrix} \quad (3.65)$$

计算可得²³

$$j_1^2 = -\mathbf{1} \quad (3.66)$$

接着：

$$j_1^3 = \underbrace{j_1^2}_{=-1} j_1 = -j_1, \quad j_1^4 = +\mathbf{1}, \quad j_1^5 = +j_1 \quad (3.67)$$

一般情况为：

$$j_1^{2n} = (-1)^n I, \quad j_1^{2n+1} = (-1)^n j_1 \quad (3.68)$$

利用上式可计算 j_1 的指数函数的级数展开⁴⁷：

⁴⁷ 推导见附录 B.4.1。相关技巧在 B.4.2. 中进一步解释。正弦和余弦函数的级数展开也在 B.4.1 中。

$$\begin{aligned} R_1 &= e^{\phi j_1} = \sum_{n=0}^{\infty} \frac{\phi^n j_1^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \underbrace{j_1^{2n}}_{(-1)^n I} + \sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \underbrace{j_1^{2n+1}}_{(-1)^n j_1} \\ &= \underbrace{\left(\sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} (-1)^n \right)}_{=\cos \phi} I + \underbrace{\left(\sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} (-1)^n \right)}_{=\sin \phi} j_1 \\ &= \cos \phi \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin \phi \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

²³ 以下符号 $\mathbf{1}$ 、 I 均表示单位矩阵。 — 译者

$$= \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \quad (3.69)$$

再利用 $e^0 = 1$ 计算左上角的 0 对应的元素，得到完整的有限变换矩阵：

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{pmatrix} \quad (3.70)$$

我们已经得出 $\mathcal{SO}(3)$ 群的生成元显式的矩阵形式 ((3.62) 式)，这样就能直接计算生成元之间的 Lie 括号了⁴⁸，不难得到⁴⁹：

$$[J_i, J_j] = \epsilon_{ijk} J_k \quad (3.71)$$

其中 ϵ_{ijk} 是 Levi-Civita 符号。

出于物理层面的考虑，在 $\mathcal{SO}(3)$ 生成元 $e^{\phi J}$ 的幂附加一个虚单位 i ，即 $e^{i\phi J}$ ，生成元的基变成：

$$J_1 = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_2 = i \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_3 = i \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.72)$$

于是 Lie 代数⁵⁰为：

$$[J_i, J_j] = i\epsilon_{ijk} J_k \quad (3.73)$$

将生成元加 i 是为了让它们成为厄米生成元⁵¹，厄米生成元的含义是 $J^\dagger \equiv (J^*)^T \stackrel{!}{=} J$ 。加 i 是从物理意义考虑的：厄米矩阵的本征值都是实数，这在量子力学里面极其重要，因为生成元的本征值是实验的可观测量，第 8.3 节对此详细讨论。如果不加 i ，生成元就是反厄米的了： $J^\dagger = (J^*)^T = -J$ ，反厄米矩阵的本征值是虚数。

还有另一种导出生成元的方法，利用 (3.55) 式 — $X = \frac{dh}{d\theta}|_{\theta=0}$ ，以及三维旋转矩阵的形式 ((3.23)、(3.70) 式) 可得²⁴：

²⁴原文下式有误，译文已按勘误表修改。

⁴⁸ 前面说过，Lie 代数元素间的‘乘法’是 Lie 括号。直接计算生成元的基之间的 Lie 括号，就可得到所有生成元之间的 Lie 括号（任意生成元都是基的线性组合）。之后会看到生成元的基在 Lie 括号运算下的行为极其重要。Lie 代数的所有重要信息都蕴含在生成元的基的 Lie 括号中。

⁴⁹ 例如： $[J_1, J_2] = J_1 J_2 - J_2 J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \underbrace{\epsilon_{12k}}_{=0, \text{除了 } k=3} J_k = \epsilon_{123} J_3 = J_3.$

⁵⁰ 重复一下，我们把生成元的基的 Lie 括号运算结果等同为 Lie 代数，因为 Lie 代数的所有重要信息都蕴含在那里。

⁵¹ 例如 $J_1^* = -i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ ，而 $J_1^\dagger = (J_1^*)^T = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = J_1$ 。

$$\begin{aligned}
J_1 &= \left. \frac{dR_1}{d\theta} \right|_{\theta=0} = \left. \frac{d}{d\theta} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \right|_{\theta=0} \\
&= \left. \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin \theta & -\cos \theta \\ 0 & \cos \theta & -\sin \theta \end{pmatrix} \right|_{\theta=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (3.74)
\end{aligned}$$

这正是(3.26)式导出的第一个生成元。第一种方法更加普遍，因为像这样已知有限变换矩阵具体形式的好事不太常见。对于之后的 Lorentz 群，我们先根据群的定义导出其生成元的基，然后才得出 Lorentz 变换的显式形式。当然如果已知变换矩阵形式那么利用(3.55)式就能最简便地导出生成元。

下面开始填作者之前挖的坑吧：我们看看 Lie 代数的现代定义到底是啥。

3.4.2 The Abstract Definition of a Lie Algebra Lie 代数的抽象定义

前面介绍了简化版的 Lie 代数定义。群 G 的 Lie 代数是满足如下条件的 X 的集合： $e^X \in G$ 。下面我们会看到²⁵，群的重要部分——群乘法（规定群元之间怎样结合），可以从 Lie 代数元素的 Lie 括号运算中导出。

就像定义 Lie 群那样，我们把 Lie 代数的关键特征表示为公理，这就形成了 Lie 代数的抽象定义：

Lie 代数是一个向量空间 \mathfrak{g} 配备一个二元运算 $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ ，且二元运算 $[\cdot, \cdot]$ 满足如下公理：

- 双线性： $\forall a, b \in \mathbb{R}$ 以及 $\forall X, Y, Z \in \mathfrak{g}$, $[aX + bY, Z] = a[X, Z] + b[Y, Z]$, $[Z, aX + bY] = a[Z, X] + b[Z, Y]$ 。
- 反交换律： $\forall X, Y \in \mathfrak{g}$, $[X, Y] = -[Y, X]$ 。
- Jacobi 恒等式： $\forall X, Y, Z \in \mathfrak{g}$, $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$ 。

之前我们把矩阵的对易子作为 Lie 代数的二元运算 $[\cdot, \cdot]$ ，不难验证对易子满足上面的要求。其实有许多和对易子丝毫不同的二元运算照样满足这些公理，比如经典力学中著名的 Poisson 括号²⁶。

注意 Lie 代数的抽象定义完全不依赖群，这十分重要哟²⁷。

²⁵作者似乎很喜欢挖这种坑然后慢慢填。——译者

²⁶Poisson 括号可见朗道，力学，高等教育出版社，2007，§42，泊松括号。ISBN: 9787040208498。——译者

²⁷但是恐怕仅仅在这个地方读者很难感受到这句话的意义。没事，译者也是很

在下一小节我们推导 $SU(2)$ 群的生成元的基（再通过线性组合就能得到所有生成元），并且我们会看到它与 $SO(3)$ 的生成元满足同样的 Lie 括号关系（(3.73)式）。这就是 $SU(2)$ 与 $SO(3)$ 群有相同 Lie 代数的含义。这个超级重要的结论将揭示 $SU(2)$ 与 $SO(3)$ 间许多不为人知的关系。

3.4.3 The Generators and Lie Algebra of $SU(2)$ $SU(2)$ 的生成元与 Lie 代数

⁵² 复习一下，这意味着两个 $SU(2)$ 中的两个群元对应 $SO(3)$ 群中的同一元素。

⁵³ 具有单位行列式正是“特殊”(S) 的含义。

前面我们历经磨难终于找到描述三维旋转的第二法—— $SU(2)$ 群，并且发现 $SU(2)$ 是 $SO(3)$ 的双覆盖⁵²。

$SU(2)$ 群的群元是具有单位行列式⁵³的幺正的 2×2 矩阵，即：

$$U^\dagger U = U U^\dagger = \mathbf{1} \quad (3.75)$$

$$\det(U) = 1 \quad (3.76)$$

⁵⁴ 前面说过，为了让生成元是厄米矩阵，我们在 e 指数上加了 i ，这样可以确保量子力学预言的实验值为实数。

从群元满足的条件可导出其生成元满足的条件。设生成元为 J_1, J_2, \dots ²⁸，将群元用生成元表示，带入上两式⁵⁴：

$$U^\dagger U = (e^{iJ_i})^\dagger e^{iJ_i} \stackrel{!}{=} \mathbf{1} \quad (3.77)$$

$$\det(U) = \det(e^{iJ_i}) \stackrel{!}{=} 1 \quad (3.78)$$

结合(3.77)式、BCH 公式（(3.56)式）和 $[J_i, J_i] = 0$ ²⁹：

$$\begin{aligned} (e^{iJ_i})^\dagger e^{iJ_i} &= e^{-iJ_i^\dagger} e^{iJ_i} \stackrel{!}{=} \mathbf{1} \\ \rightarrow e^{-iJ_i^\dagger + iJ_i + \frac{1}{2}[J_i^\dagger, J_i] + \dots} &\stackrel{!}{=} \mathbf{1} \\ \underbrace{\rightarrow}_{e^0=1} J_i^\dagger &\stackrel{!}{=} J_i \end{aligned} \quad (3.79)$$

满足 $J_i^\dagger = J_i$ 的矩阵称为厄米矩阵，由上可见 $SU(2)$ 的生成元必须是厄米的。

久之后才领会到的，但是请一定把这句话记住。——译者

²⁸在下面的式子中请注意区分虚数单位 i 和表示数字指标的下标 i ，我们对作者采用这样的符号很抱歉。——译者

²⁹这由 Lie 括号的反对称性导出： $[J_i, J_i] = -[J_i, J_i] \rightarrow [J_i, J_i] = 0$ ，——译者

利用恒等式 $\det(e^A) = e^{\text{tr}(A)}$ ³⁰ 与 (3.78) 式可得：

$$\det(e^{iJ_i}) = e^{i\text{tr}(J_i)} = 1 \xrightarrow[e^0=1]{} \text{tr}(J_i) \stackrel{!}{=} 0 \quad (3.80)$$

综上， $SU(2)$ 的生成元一定是无迹的厄米矩阵。 2×2 无迹厄米矩阵的一组基为⁵⁵：

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.81)$$

也就是说任意 2×2 无迹厄米矩阵都可唯一表示为上面三个矩阵的线性组合。这三个矩阵称为 **Pauli 矩阵**。

把 Pauli 矩阵扔进 Lie 括号中进行计算可得：

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \quad (3.82)$$

其中 ϵ_{ijk} 是 Levi-Civita 符号。上式右侧的 2 很多余，因此通常将 $SU(2)$ 的生成元的基定义为 $J_i \equiv \frac{1}{2}\sigma_i$ ，这样 Lie 括号计算结果为：

$$[J_i, J_j] = i\epsilon_{ijk}J_k \quad (3.83)$$

注意这与 $SO(3)$ 的生成元形式 ((3.73) 式) 相同！因此我们说 $SU(2)$ 与 $SO(3)$ 有相同的 Lie 代数，因为 Lie 代数是通 Lie 括号定义的！利用 Lie 代数的抽象定义可以用不同的方式描述 $SU(2)$ 群对应的变换，即， $SU(2)$ 群对应的变换可以不用 2×2 矩阵表示。为了做到这一点，我们还需要知道 Lie 群的抽象定义。因为一直以来就是用 2×2 矩阵定义 $SU(2)$ 的，(SU ‘2’ 嘛！) 换一种表示方法 (例如用 3×3 矩阵表示) 则不知所云。Lie 群的抽象定义能够让我们发现同一变换的不同表示之间的关系。抽象定义将 Lie 群与一种几何结构——流形相等同，利用这种抽象结构来定义一个群。这一想法初看十分怪异，但在学习两个例子之后就会发现它是超级有意义的。

⁵⁵ 2×2 复数矩阵有 4 个复数元素，因此有 8 个自由度。无迹、厄米性的条件限制了只剩 3 个自由度。

³⁰ 可将 A 化成上三角矩阵来证明。— 译者

3.4.4 The Abstract Definition of a Lie Group Lie 群的抽象定义

我们讨论过纯良的 $U(1)$ 群，它由全体单位复数构成。定义为 $z^*z = 1$ ，设 $z = a + ib$ ，则：

$$z^*z = (a + ib)^*(a + ib) = (a - ib)(a + ib) = a^2 + b^2 = 1 \quad (3.84)$$

这也是单位圆的定义条件⁵⁶。可见单位复数的集合正是复平面上的单位圆。此外，我们知道 $U(1)$ 与 $SO(2)$ 群间存在一一映射⁵⁷。因此这两个群可以与几何对象——单位圆——视为等同。群的抽象定义并非将这个群用 $SO(2)$ ，或是不同的 $U(1)$ 体现，因为它们都是由特定维数的对象定义的。可以将该群直接利用单位圆定义。即二维旋转变换（对应的 Lie 群）等同于单位圆，然后我们可以用 $SO(2)$ 群表示这个变换（ 2×2 矩阵），也可用 $U(1)$ 群来表示（单位复数）。

下面讨论升级版—— $SO(3)$ 与 $SU(2)$ 。前面讲过在 $SU(2)$ 与单位四元数之间有一一映射。单位四元数就是满足下列条件的一般四元数 $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ ：

$$a^2 + b^2 + c^2 + d^2 \stackrel{!}{=} 1 \quad (3.85)$$

这与三维超球面 S^3 的定义⁵⁸相同！上式给出了一个 $SU(2)$ 与 S^3 之间的映射，该映射是同构映射（是单射也是满射），因此我们可以把 $SU(2)$ 与三维超球面 S^3 等同。

结合上面的例子，Lie 群的现代定义就不难理解了⁵⁹：

Lie 群是一个群，也是一个微分流形⁶⁰。当然，这个流形要满足如下条件：

群乘法 \circ 诱导出的从流形到流形自身的映射必须是可微的³¹。这称为相容性条件，它保证了群定义与流形定义的兼容。例如，群 G 的任意群元 a 诱导出了从 G 到 G 的映射：给定 $\forall b \in G$ ， $c \equiv a \circ b \in G$ ，映射 $a : G \rightarrow G$ ³² 必须是可微的。

流形中的任意点都有相应的坐标，可以用坐标的语言表示上述内容： $c = ab$ 对应的坐标必须是 b 的坐标的可微函数。

从 Lie 代数的抽象定义可以看出 $SO(3)$ 与 $SU(2)$ 有相同的

⁵⁶ 单位圆记作 S^1 ，它是二维空间中到原点的距离为 1 的所有点构成的集合。用数学语言就是 $S^1 = \{(x_1, x_2) | x_1^2 + x_2^2 = 1\}$

⁵⁷ 严谨的说法是：同构映射。数学上把两个对象‘相同’称为它们‘同构’。如果两个对象之间存在一个同构映射，那么它们就是同构的。

⁵⁸ 单位圆 S^1 的定义是 $\{(x_1, x_2) | x_1^2 + x_2^2 = 1\}$ 。二维球面 S^2 的定义是 $\{(x_1, x_2, x_3) | x_1^2 + x_2^2 + x_3^2 = 1\}$ ，它是三维球体的表面。三维超球面 S^3 的定义是 $\{(x_1, x_2, x_3, x_4) | x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}$ 。 S 的上标表示维数。在二维空间中加上单位长度的条件之后得到了一维对象 S^1 ，相同地，在四维空间加上单位长度条件 $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$ 后得到三维对象 S^3 ，它是四维超球体的“表面”，注意区分 S^3 与三维球体。

⁵⁹ 下面定义里的数学细节不必在意，本书用不着那些内容。但必须记住现代定义的灵魂：Lie 群 = 流形。

⁶⁰ 流形是满足某些条件的点集，例如球面就是一个流形，它局部看起来就像平坦 Euclidean 空间 \mathbb{R}^n 。或者， n 维流形就是满足如下条件的点的集合：任意一点可以对应 n 个独立坐标。本章末的 3.11 节有流形的详细说明。

³¹ 把函数看做一个映射，则映射的可微性就是函数的可微性。流形之间映射的可微性的一般定义可以见 微分几何入门与广义相对论，梁灿彬，上册，科学出版社，第二版，2006，ISBN 7-03-016469-1 第二章定义 3。——译者

³² 或者记作 $a : b \in G \mapsto c \in G$ 。

Lie 代数 ((3.83)式), 现在有了 Lie 群的抽象定义, 我们可以填在 3.4 节结尾处开的一个坑了:

“...Lie 理论的一条重要定理告诉我们这其中有一个特别的 Lie 群。”

为啥有一个特别的群? 因为这个特别的群是单连通的。单连通性是流形的性质, 它的含义是流形上的任意闭合曲线可以平滑地收缩为一点⁶¹。只有通过现代定义将群与流形联系起来才能找出特别的群。

把上面的重点重复一遍⁶²:

任一 Lie 代数仅对应一个单连通 Lie 群。

给定 Lie 代数, 它可对应许多 Lie 群, 而其中唯一的单连通 Lie 群可以视为其他群的“母体”, 因为总存在从这个单连通群到其他群的映射, 反之则不一定。何妨把这个单连通群称为母群呢? 数学家们给它起了一个不太文艺的名字: 覆盖群。并且称该群覆盖了同一 Lie 代数对应的其他群。我们已经遭遇过这种情况: $SU(2)$ 群是 $SO(3)$ 群的双覆盖, 即两个 $SU(2)$ 元素对应一个 $SO(3)$ 的元素。

$SU(2)$ 就是三维超球面, 它正是一个单连通流形。因此 Lie 代数 (3.83) 式对应的覆盖群就是 $SU(2)$ 。从 $SU(2)$ 出发可以构造出其他所有对应的群。

$SO(3)$ 是什么流形? 已知 $SU(2)$ 到 $SO(3)$ 的映射将 $SU(2)$ 的两个点映射为 $SO(3)$ 中的一个。因此 $SO(3)$ 是单位超球面的一半。

因此从流形的角度看, $SU(2)$ 群比 $SO(3)$ 更完整, $SO(3)$ 只是其中的一部分。

本书的指导思想之一就是用最基本的群描述大自然最基本的对称性。具体到三维旋转的情况, 我们采用 $SU(2)$ 而非 $SO(3)$ 。之后其他的对称性我们都这么处理。

世界似乎就是这么设计的!³³ 比如必须用 Poincaré 群的双覆盖表示来描述基本粒子, 而常用却不完整的 4-向量表示就不行。要从最深刻的层面描述自然, 必须使用覆盖群, 而非从覆盖群导出的其他群。

给定一个 Lie 代数, 我们可以通过 Lie 代数的表示来导出它的覆盖群的表示⁶³, 然后把 Lie 代数元素 (生成元) 的矩阵表示扔

⁶¹ 本书不讨论这些抽象内容。感兴趣的读者可以阅读本章末尾的推荐教材。在这里我们知道一个 Lie 代数总是有一个特别的群就够了。

⁶² 这个命题的证明可见 Michael Spivak.

A Comprehensive Introduction to Differential Geometry, Vol.1, 3rd Edition. Publish or Perish, 3rd edition, 1 1999. ISBN 9780914098706

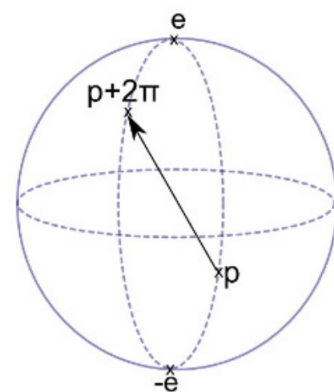


图 3.7: 三维超球面 S^3 的一个二维切面 (S^3 自身是四维超球的“表面”, 愚蠢的人类没法直接把它画出来), 图中可见球面的上半部分对应 $SO(3)$, 因为从 $SU(2)$ 到 $SO(3)$ 的映射将前者中的 p 点与 $p + 2\pi$ 映射到后者中的同一点。

⁶³ “表示”的具体含义下一节就讲。

³³ 先不管设计者有没有、到底是谁吧

进 e 指数里面就能得到任意群元的矩阵表示。

这就是 Lie 理论的威力所在。人类通过纯数学就能揭示自然的某些深刻规律。狭义相对论的标准对称群隐藏了一些真相⁶⁴，因为它（尽管常用）不是这一对称性对应的最基本的覆盖群。Poincaré 群⁶⁵的覆盖群才是最基本的，我们用它描述自然。

总结一下⁶⁶：

- $S^1 \cong U(1) \xleftrightarrow{\text{一一对应}} \mathcal{SO}(2)$
 - $S^3 \cong SU(2) \xleftrightarrow{\text{二对一}} \mathcal{SO}(3) \cong S^3$ 的一半
- $\Rightarrow SU(2)$ 是 Lie 代数 $[J_i, J_j] = i\epsilon_{ijk} J_k$ ((3.83)式) 的覆盖群，因为 S^3 是单连通流形。

下一节介绍 Lie 理论的重要分支 — 表示论。表示论使我们能够从任意 Lie 群导出相应基本群。

3.5 Representation Theory 表示论

群论可以描述变换本身而不必联系现实中的特定对象，这是群论的重要特征。

理论上经常从抽象观点来考虑一个群。这意味着用流形结构与群乘法定义群。例如 $SU(2)$ 就是 S^3 (超球面)，群元就是流形中的点，群乘法规定两个群元 a, b 怎样结合，且满足群公理。搞物理的更关心一个群到底能做什么，即群作用。

同一个群能作用于许多不同种类的东西⁶⁷，这一重要事实使我们定义了表示的概念：群 G 的一个表示就是从 G 到某个向量空间 V 上全体线性变换组成的集合⁶⁸的映射，将表示记作 R , 映射

$$R : \underbrace{g}_{\in G} \mapsto R(g) \quad (3.86)$$

其中 $R(g)$ 表示 V 上的某个线性变换。映射 R 必须满足如下条件⁶⁹：

- $R(e) = I$, 恒等元对应的线性变换是恒等变换。
- $R(g^{-1}) = (R(g))^{-1}$, 逆元对应的变换是逆变换。
- $R(g) \circ R(h) = R(g \circ h)$, 群元 g, h 结合对应的变换等于他们分别对应变换的结合。

群的表示⁷⁰将抽象的群元（群流形上的点）与向量空间中的线性变换相联系。尽管按定义来讲，表示是个映射，但大多数情况把

⁶⁴ 学过量子力学的人可能听说过：标准对称群不能告诉我们自旋！

⁶⁵ 简明起见，我们省略了“双覆盖”或“覆盖群”的说法，我们利用 Poincaré 群的一个表示导出相应的 Lie 代数，然后利用此 Lie 代数导出相应覆盖群的表示，也就是 Poincaré 双覆盖群的表示。

⁶⁶ S^2 未出现，没有与 S^2 对应的 Lie 群，因为没有三维复数。我们从二维复数 i 直接跳到了四维的四元数 i, j, k 。

⁶⁷ 这话的含义稍后会更加清楚。

⁶⁸ 本书中 V 上的线性变换其实就是矩阵，即表示 R 把群元映射为矩阵，矩阵通过矩阵乘法作用于向量空间中的向量。

⁶⁹ 满足这些条件的映射称为同态映射，同构映射当然是同态的，并且比同态性还多了一一映射的特征。

⁷⁰ 表示的定义可以进行推广，不一定非得是向量空间（别的空间也行）上的线性变换（非线性也可能），拓展后的映射叫 **realization**。物理主要考虑向量空间中的线性变换（例如量子力学的 Hilbert 空间、狭义相对论的 Minkowski 空间），因此对物理而言表示比更广泛的 realization 更重要。

表示看做许多矩阵的集合就行了。例如，向量空间⁷¹ \mathbb{R}^3 中的旋转矩阵组成的集合就是 $\mathcal{SO}(3)$ 群的一个表示，旋转矩阵就是 \mathbb{R}^3 上的线性变换。注意 $\mathcal{SO}(3)$ 群对应的变换也可以作用到其他向量空间。

表示论让我们能系统研究同一个群对不同向量空间的作用，这让群论变得精彩多了。

$SU(2)$ 是在物理里最常用的群之一。它可以作用到许多向量空间上，比如作用于一维复数向量空间 C^1 ，稍后会看到结果是很平凡的。再比如作用于 C^2 ，下面我们会详细探究这种作用， C^2 的元素是二维复数向量，因此 $SU(2)$ 的群元以 2×2 复数矩阵的形式出现，这是 $SU(2)$ 最常用的表示，之前一直采用此表示。 $SU(2)$ 还可作用到 C^3 ，前人已经对此有完备的研究， $SU(2)$ 作用于 3 维复向量的时候以 3×3 矩阵的形式出现。剧透： $SU(2)$ 的 C^3 版本的生成元是⁷²：³⁴

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (3.87)$$

把生成元扔进 e 指数里就可得到 $SU(2)$ 的群元在这个表示之下对应的矩阵的形式。

上面的路可以一直走到 C^n 。因此作者纳闷人们为啥叫把这个群叫 $SU(2)$ 而非 S^3 。⁷³但 $SU(2)$ 通常定义为满足如下条件的 2×2 (是 2 哟!) 复数矩阵集合 (即(3.33)式)

$$U^\dagger U = \mathbf{1}, \quad \det(U) = 1 \quad (3.88)$$

下面会把 $SU(2)$ 的群元写成 3×3 矩阵的形式，因此要牢记当谈论一般情况 $SU(2)$ 或其他群群元的时候，我们指的是抽象的群流形中的点，而非只是 2×2 矩阵或者别的。

本章前面都是从群的某个表示出发定义一个群的。比如用 2×2 矩阵定义 $SU(2)$ ，此方法的优点是群的含义十分具体。在入门之后，把群视为抽象流形的观点⁷⁴更有效，因为这样可以更方便地找出其他有用的群表示。

⁷¹ \mathbb{R}^3 表示三维 Euclidean 空间，它的元素就是三维向量，附录 A.1 里面有很多例子。

⁷² 下一节 (3.6.5) 就推倒 (无误) 它，这里只是为了说明就算是 $SU(2)$ 也有不得不守护的东西，也可以作用于三维向量。

⁷³ 在作者撰写本书手稿时就把这个群叫 S^3 ，考虑到这个非标准名字会对初学者造成困扰所以改了，因为主流教科书都用 $SU(2)$ 。

⁷⁴ 例如把 $SU(2)$ 视为 S^3

³⁴原文下式 J_3 的形式有误，译文已按勘误表修改。

下面定义几个之后要用的概念，它们与表示的“等级”有关（不是所有表示都是基本的）。

首先是**相似变换**。给定任意矩阵 R 与可逆⁷⁵矩阵 S 。则通过可逆矩阵进行相似变换得到：

$$R \rightarrow R' = S^{-1}RS \quad (3.89)$$

相似变换与表示的关系在于，给定群 G ，群的表示 $R(G)$ 以及任意可逆矩阵 S ，则 $S^{-1}RS$ 也是群表示。这可以从群的定义直接看出来： $\forall g_1, g_2 \in G$ ，表示 R 把群元映射为向量空间 V 上的线性变换 $R(g_1), R(g_2)$ 。群的定义要求：

$$R(g_1)R(g_2) = R(g_1g_2) \quad (3.90)$$

现在把表示进行相似变换：

$$S^{-1}R(g_1)\underbrace{SS^{-1}}_{=1}R(g_2)S = S^{-1}R(g_1)R(g_2)S = S^{-1}R(g_1g_2)S \quad (3.91)$$

上式可得 $S^{-1}RS$ 也是一个表示。如果我们有一个表示，那么就能利用任意可逆矩阵变出好多其他表示，新表示的形式可以更漂亮（比如让它是对角矩阵）。

第二个概念：**不变子空间**。给定群 G 在向量空间 V 上的表示 R ，如果 $V' \subset V$ 满足：

$$\forall v \in V', g \in G, \text{ 都有 } R(g)v \in V'$$

则 V' 称为 V 的不变子空间⁷⁶。

任意群元作用于 V' 中的任意向量所得向量仍然在 V' 内，这就可以在 V' 上定义群 G 的表示 R' 了：

$$R'(g)v = R(g)v, \quad \forall v \in V', g \in G \quad (3.92)$$

这种情况下，表示 R 就有一种不够基本的感觉，它更像是多个小基本部件所构成的。我们称原来的表示 R 不是原来的向量空间 V 的基本表示， R' 称为子表示。

定义了上述概念之后就能引入超级重要的不可约表示。群 G

⁷⁵ 矩阵 S 可逆的含义是存在矩阵 T 使得 $ST = TS = \mathbf{1}$ ， S 的逆矩阵一般记作 S^{-1} 。

⁷⁶ 自然 $v \in V$ ，子空间 $V' \subset V$ ，换言之 V' 的任意元素都是 V 的元素。

在向量空间 V 上的不可约表示 R 意味着除了 0 空间和 V 自身之外再无不变子空间。不可约表示是基本的，它们不能再分解为子表示。利用所有的不可约表示可导出其他所有表示。

还可以从另一个角度看不可约表示：不可约表示不能通过相似变换变成分块对角矩阵形式。而可约表示能被相似变换变成分块对角形式。表示的可约性很关键，比如描述基本粒子要用不可约表示⁷⁷，基本粒子在变换下的行为由相应对称群的不可约表示描述。

一个群有许多表示⁷⁸，怎么知道哪个是描述自然的基本表示？Casimir 元可以给我们答案。Casimir 元（记作 C ）从 Lie 代数的元素 — 生成元导出，它的定义是，对任意生成元 X 都有

$$[C, X] = 0 \quad (3.93)$$

著名的 Schur 引理⁷⁹告诉我们，给定一个不可约群表示 $R: \mathfrak{g} \rightarrow GL(V)$ ³⁵，如果某个线性变换 $T: V \rightarrow V$ （即 $T \in GL(V)$ ）与所有 $R(g), \forall g \in G$ 对易³⁶，则 T 必为恒等变换的常数倍。因此 Casimir 元是群表示的‘常数’，这些‘常数’的值自然地成群表示‘编号’⁸⁰。之后我们就从 Casimir 元对应编号最低的不可约表示开始。

群表示的定义与向量空间 V 有关。一个表示就是将抽象群到 V 的线性变换的映射。线性代数告诉我们以任意矩阵的本征向量为基可以张成一个向量空间。任意 Lie 群的生成元至少有一个可以通过相似变成对角矩阵⁸¹，然后这个对角生成元的本征向量就可作为向量空间的基。

下一节导出 $SU(2)$ 群的不可约表示。研究 $SU(2)$ 是因为 Lorentz 群的 Lie 代数是两份 $SU(2)$ 的 Lie 代数。Lorentz 群是 Poincaré 群的一部分，本章由浅入深对他们逐一讨论。

3.6 $SU(2)$ 特殊么正群 $SU(2)$

我们在 3.4.3 小节中用具体的矩阵形式进行计算得到了 $SU(2)$ 生成元的 Lie 括号⁸²。通过这一点便能得到 $SU(2)$ 群的进一步表示。回到开头我们对 $SU(2)$ 群的表示，即 2×2 的么正矩阵，其行列式的值为 1，但这只是一个特例罢了。在进一步处理这个问题之前，我们可以想一想大概能得到什么样的表示。

³⁵ $GL(V)$ 表示向量空间 V 上所有线性变换组成的集合。

³⁶ 即 $\forall g \in G, [R(g), T] = 0$ 。

⁷⁷ 要不用啥呢？（译者注：原文为 ‘What else?’）

⁷⁸ 比如我们前面讨论了二维旋转变换的两种方法 — 单位复数、 2×2 矩阵。它们都是群流形 S^1 的表示。

⁷⁹ Schur 引理是群论的基本结论之一，它在任何一本群论书都有。

⁸⁰ 具体内容见下一节的例子。

⁸¹ 对角生成元的集合称为 Cartan 子代数，相应的集合元素称为 Cartan 生成元。它们在量子场论中大放异彩，因为 Cartan 生成元的本征值可以给基本粒子“编号”。例如量子色动力学使用 $SU(3)$ 群，它有两个 Cartan 生成元。因此色动力学描述的粒子都有两个“编号”。通常不用数字表示编号，而是用“色”（红、蓝、绿…）来标注。再比如描述弱相互作用的 $SU(2)$ 群只有一个 Cartan 生成元，因此每个粒子只有一个标签 — Cartan 生成元的本征值。

⁸² 想当年我们就是用这东西抽象地定义了群的 Lie 代数。具体结果见 (3.83) 式

3.6.1 The Finite-dimensional Irreducible Representations of $SU(2)$

$SU(2)$ 群的有限维不可约表示

为了知道 $SU(2)$ 群有限维不可约表示的一些东西，我们用在3.4.3小节中出现的算符的线性组合来定义一个新的算符⁸³：

$$J_+ = \frac{1}{\sqrt{2}}(J_1 + iJ_2) \quad (3.94)$$

$$J_- = \frac{1}{\sqrt{2}}(J_1 - iJ_2) \quad (3.95)$$

用(3.83)式可以验证新定义的算符满足如下对易关系：

$$[J_3, J_{\pm}] = \pm J_{\pm} \quad (3.96)$$

$$[J_+, J_-] = J_3. \quad (3.97)$$

⁸³ 我们总是能让一个生成元对角化。遵从惯例，我们选 J_3 为对角阵，由此也生出了矢量空间的基矢。此外，人们一般都是按书上的这种方法引入 J_{\pm} 的。

⁸⁴ 即 $J_3 v = bv$ ，详见附录C.4。

若 v 为 J_3 本征矢，本征值为 b ⁸⁴，用新的算符作用在 v 上会得到很有意思的结果：

$$\begin{aligned} J_3(J_{\pm}v) &= J_3(J_{\pm}v) + \underbrace{J_{\pm}J_3v - J_3J_{\pm}v}_{=0} \\ &= \underbrace{J_{\pm}J_3v}_{=J_{\pm}bv} + \underbrace{J_3J_{\pm}v - J_{\pm}J_3v}_{=[J_3, J_{\pm}]v} \\ &\stackrel{(3.96)\text{式}}{=} (b \pm 1)J_{\pm}v \end{aligned} \quad (3.98)$$

可见 $J_{\pm}v$ 也是 J_3 的本征矢，本征值为 $(b \pm 1)$ ，记该本征矢为 w ，则：

$$J_3w = (b \pm 1)w \quad \text{以及} \quad w = J_{\pm}v \quad (3.99)$$

算符 J_- 和 J_+ 通常称为降算符和升算符，也称**阶梯算符 (ladder operators)**。我们可以通过重复利用阶梯算符 J_{\pm} 来得到关于 J_3 的本征矢，但这个过程不能无休止地进行，因为不同本征值的本征矢之间是线性独立的，而我们处理的是有限维的表示。这意味着相应的矢量空间是有限维的，我们只能得到有限个独立的矢量。

我们推断一定有一本征矢对应着最大的本征值 v_{\max} 。若把 J_+ 作用到 v 上 N 次以后，得到本征值最大的本征矢 v_{\max}

$$v_{\max} = J_+^N v \quad (3.100)$$

我们有

$$J_+ v_{max} = 0 \quad (3.101)$$

因为 v_{max} 已经是有最大本征值的本征矢了。称最大本征值为 $j := b + N$ 。相应地有对应着最小本征值的本征矢 v_{min} ，满足

$$J_- v_{min} = 0 \quad (3.102)$$

把 J_- 作用到 v_{max} 上 M 次以后，得到本征值最小的本征矢 v_{min}

$$v_{min} = J_-^M v_{max} \quad (3.103)$$

因此， v_{min} 对应的本征值为 $j - M$ 。为进一步讨论，我们需要知道 J_{\pm} 作用到本征矢上的具体形式。上述的计算说明， $J_- v_k$ 应等于 v_{k-1} 乘上一常数：

$$J_- v_k = \alpha_k v_{k-1} \quad (3.104)$$

如果非要深究 J_- 作用到 v_{max} 上的细节的话，那我只能告诉你⁸⁵ α 满足

$$\alpha_{j-k} = \frac{1}{\sqrt{2}} \sqrt{(2j-k)(k+1)} \quad (3.105)$$

注意到当 $k = 2j$ 时 α 为 0，因此，经过 $2j$ 步之后，我们从梯子的顶端到达了底端。 v_{min} 的本征值即是 $j - 2j = -j$ 。总的来说，我们得到了 $2j + 1$ 个本征态的本征值

$$\{-j, -j+1, \dots, j-1, j\} \quad (3.106)$$

此式仅当 j 为整数或半整数⁸⁶时成立。因为我们有 $2j + 1$ 个线性独立的本征矢，故矢量空间应为 $2j + 1$ 维⁸⁷。这些 J_3 的本征矢构成了完备的矢量空间，因为 J_1 和 J_2 能够被 J_+ 和 J_- 线性表示，就算是一个任意的线性组合 $\sum_i a_i v_i$ 也能变换成另一种线性组合 $\sum_i b_i v_i$ ，其中 a_i, b_i 是常数。因此， J_3 的本征矢构成了 V 的一个非平庸不变子空间，而我们也正在寻找能构成完备矢量空间 V 的不可约表示。

我们可以利用上面定义的 $2j + 1$ 维的矢量空间及 J_3 的本征

⁸⁵ 见 Matthew Robinson. Symmetry and the Standard Model. 第 90 页, Springer, 1st edition, August 2011. ISBN 978-1-4419-8267-4.

⁸⁶ 不信的话你就用分数试试看！

⁸⁷ 见此书 189 页: Nadir Jeevanjee. An Introduction to Tensors and Group Theory for Physicists. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

⁸⁸ 例如此书 190 页: Nadir Jeevanjee. An Introduction to Tensors and Group Theory for Physicists. Birkhaeuser, 1st edition, August 2011. ISBN 978- 0817647148

⁸⁹ 回想一下, Casimir 算符指的是那些与群中每个生成元都对易的群元, 即 $[C, X] = 0$ 。

矢 v_k 来给出 $SU(2)$ 的一个表示。进一步说, 我们能证明 $SU(2)$ 的每一种不可约表示都等价于这一种⁸⁸。

3.6.2 The Casimir Operator of $SU(2)$ $SU(2)$ 的 Casimir 算符

如第3.5节所述, 我们可以用 Casimir 算符来标记不同的群表示。 $SU(2)$ 恰好有一个 Casimir 算符⁸⁹:

$$J^2 := (J_1)^2 + (J_2)^2 + (J_3)^2 \quad (3.107)$$

满足定义的条件:

$$[J^2, J_i] = 0 \quad (3.108)$$

我们可以用(3.95)式和(3.94)式定义的 J_{\pm} 来重新表示 J^2 :

$$\begin{aligned} J^2 &= J_+ J_- + J_- J_+ + (J_3)^2 \\ &= \frac{1}{2}(J_1 + iJ_2)(J_1 - iJ_2) + \frac{1}{2}(J_1 - iJ_2)(J_1 + iJ_2) + (J_3)^2 \\ &= \frac{1}{2}((J_1)^2 - iJ_1 J_2 + iJ_2 J_1 + (J_2)^2) + \frac{1}{2}((J_1)^2 + iJ_1 J_2 - iJ_2 J_1 + (J_2)^2) \\ &\quad + (J_3)^2 \\ &= (J_1)^2 + (J_2)^2 + (J_3)^2 \quad \checkmark \end{aligned} \quad (3.109)$$

⁹⁰ 这些东西只是一些归一化常数, 如果我们将 J_{\pm} 作用到归一态上, 得到的结果一般都不是归一化的。然而, 学物理的都喜欢用归一态, 详细的原因以后再谈。这些常数的推导有些无聊, 大体思路是先假设 c 是个归一化常数, 然后从 $J_{\pm} v_k = c v_{k \pm 1}$ 开始。计算过程能在大多数的量子力学教材中找到 (讲角动量和角动量算符那块儿)。如果你不知道也没关系, 因为这节讲的对接下来要讲的东西不是太重要。

带入⁹⁰

$$J_+ v_k = \frac{1}{\sqrt{2}} \sqrt{(j+k+1)(j-k)} v_{k+1} \quad (3.110)$$

以及

$$J_- v_k = \frac{1}{\sqrt{2}} \sqrt{(j+k)(j-k+1)} v_{k-1} \quad (3.111)$$

我们就能计算对于每个表示都成立的系数³⁷:

$$\begin{aligned} J^2 v_k &= (J_+ J_- + J_- J_+ + (J_3)^2) v_k \\ &= J_+ J_- v_k + J_- J_+ v_k + (J_3)^2 v_k \\ &= J_+ \frac{1}{\sqrt{2}} \sqrt{(j+k)(j-k+1)} v_{k-1} + J_- \frac{1}{\sqrt{2}} \sqrt{(j+k+1)(j-k)} v_{k+1} + k^2 v_k \\ &= \frac{1}{\sqrt{2}} \sqrt{(j+k)(j-k+1)} J_+ v_{k-1} + \frac{1}{\sqrt{2}} \sqrt{(j+k+1)(j-k)} J_- v_{k+1} + k^2 v_k \end{aligned}$$

³⁷原文下式有误, 译文已按勘误表修改。

$$\begin{aligned}
&= \frac{1}{\sqrt{2}} \sqrt{(j+k)(j-k+1)} \frac{1}{\sqrt{2}} \sqrt{(j+(k-1)+1)(j-(k-1))} v_k \\
&\quad + \frac{1}{\sqrt{2}} \sqrt{(j+k+1)(j-k)} \frac{1}{\sqrt{2}} \sqrt{(j+(k+1))(j-(k+1)+1)} v_k + k^2 v_k \\
&= \frac{1}{2} (j+k)(j-k+1) v_k + \frac{1}{2} (j-k)(j+k+1) v_k + k^2 v_k \\
&= (j^2 + j) v_k = j(j+1) v_k
\end{aligned} \tag{3.112}$$

现在会看这种表示下的几个特例，当然，从最简单的低维数情况开始。

3.6.3 The Representation of $SU(2)$ in one Dimension $SU(2)$ 的一维表示

j 的最小可能值为 0，这样群的表示是一 $2j+1 = 2 \cdot 0 + 1 = 1$ 维矢量空间。可以看出这个表示是平庸的，因为只有 1×1 矩阵满足 $SU(2)$ 李代数的对易关系 $[J_l, J_m] = i\varepsilon_{lmn} J_n = 0$ 。而生成元 0 的指数映射得到的是变换 $U = e^0 = 1$ ，即群的恒等元。

3.6.4 The Representation of $SU(2)$ in two Dimension $SU(2)$ 的二维表示

接下来我们讨论 $j = \frac{1}{2}$ 的情况，这个表示是 $2j+1 = 2 \cdot \frac{1}{2} + 1 = 2$ 维的。群的生成元 J_3 的本征值为 $\frac{1}{2}$ 和 $\frac{1}{2} - 1 = -\frac{1}{2}$ ，正如从(3.106)式中所见，因此 J_3 可表示为

$$J_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.113}$$

此处我们选择 J_3 为对角化的生成元⁹¹。对应其本征值 $+\frac{1}{2}$ 和 $-\frac{1}{2}$ 的本征矢为：

$$v_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{以及} \quad v_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.114}$$

可以用在这一基矢下的阶梯算符来得到 $SU(2)$ 另外两个生成元的矩阵形式

$$J_1 = \frac{1}{\sqrt{2}} (J_- + J_+) \tag{3.115}$$

$$J_2 = \frac{i}{\sqrt{2}} (J_- - J_+) \tag{3.116}$$

⁹¹ 由于对易关系， $SU(2)$ 群仅有一个对角化的生成元。还有就是记住我们可以用相似变换把一个不对角的生成元变成对角的。

这一点可以直接从(3.95)式和(3.94)式关于 J_{\pm} 的定义的得到。回忆一下, 表示空间中的基矢是由 J_3 的本征矢给出的, 而我们也基矢给出用 J_1, J_2 。换句话说, J_1, J_2 是由它们作用在 J_3 的本征矢上的效果来定义的。例如

$$J_1 v_{\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- + J_+)v_{\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- v_{\frac{1}{2}} + \underbrace{J_+ v_{\frac{1}{2}}}_{=0}) = \frac{1}{\sqrt{2}}J_- v_{\frac{1}{2}} = \frac{1}{2}v_{-\frac{1}{2}} \quad (3.117)$$

式中用到 $J_+ v_{\frac{1}{2}} = 0$, 因为 $\frac{1}{2}$ 是最大的本征值。最后结果的系数 $\frac{1}{2}$ 是从(3.105)式得到的。同理可得

$$J_1 v_{-\frac{1}{2}} = \frac{1}{\sqrt{2}}(J_- + J_+)v_{-\frac{1}{2}} = \frac{1}{2}v_{-\frac{1}{2}} \quad (3.118)$$

而我们之前得到的基矢是 $v_{\frac{1}{2}} = (1, 0)^T$ 和 $v_{-\frac{1}{2}} = (0, 1)^T$, 因此 J_1 的矩阵为:

$$J_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (3.119)$$

⁹² 推导出的(3.117)式为 $J_1 v_{\frac{1}{2}} = \frac{1}{2}v_{-\frac{1}{2}}$ 。带入这里具体的矩阵形式可得 $J_1 v_{\frac{1}{2}} =$

$$\frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}v_{-\frac{1}{2}} \quad \checkmark。$$

⁹³ 再次说明, $SU(2)$ 最开始的定义是全体行列式值为 1 的 2×2 么正矩阵的集合, 而在这儿我们采用抽象的定义, 即其对应的流形 S^3 , 接下来还要讨论这个群的更高维表示, 比如 3×3 矩阵的表示。其实叫成别的名字可能更合适 (比如说对应的流形 S^3), 但按照传统, 它就叫 $SU(2)$ 。

⁹⁴ 仍然是从对角化的生成元 J_3 出发。学过的人都知道 J_3 的本征值为 $(1, 0, -1)$, 因此立刻就能得到其矩阵形式。对于 J_1, J_2 而言, 仍然是从其作用结果出发, 即把他们写成 J_{\pm} 的线性组合然后作用到 J_3 的本征矢上。

可以验证上式满足我们所说的 J_1 作用到基矢上的效果⁹²。同理可得

$$J_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (3.120)$$

正如我们在这一章开头寻找 $SU(2)$ 群 Lie 代数时所发现的一样 ((3.81)式), 这些生成元 $J_i = \frac{1}{2}\sigma_i$ 正是 Pauli 矩阵。可以看出, 现在使用表示的是二维表示。当然除此之外还有其他的表示形式, 例如我们下节会看到的三维空间中的表示⁹³。

3.6.5 The Representation of $SU(2)$ in two Dimension $SU(2)$ 的三维表示

重复与二维情况相似的操作⁹⁴ 可以得到³⁸:

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} J_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (3.121)$$

这就是 $SU(2)$ 的三维表示。如果你有兴趣的话，还可以推出三维表示下 $SU(2)$ 群元的形式 (把生成元扔进 e 指数里)。介于现在所学已足够理解最重要的 Lorentz 群的表示，我们将不会推导更高维的 $SU(2)$ 表示。

3.7 The Lorentz Group $\mathcal{O}(1, 3)$ Lorentz 群 $\mathcal{O}(1, 3)$

抽象总是从具体的现实中产生的 ... 你总得从某件事物出发，最后才能把所有现实的痕迹抹除掉。

— Pablo Picasso⁹⁵

本节从 Lorentz 群一种常用的表示出发导出相应的 Lie 代数。推导过程与之前 $SU(2)$ 的情形相同 — 用 $SU(2)$ 的 2×2 矩阵表示导出了 Lie 代数。我们会发现 Lorentz 群的 Lie 代数可以用两份 $SU(2)$ Lie 代数表示，这会导出 Lorentz 群的更多表示。著名的向量表示 — 即 Lorentz 群群元为 4×4 矩阵，作用于 4 维向量 — 仅仅是众多表示中的一种。新表示能描述向量表示不能描述的物理系统。这就是 Lie 理论的威力所在，它使我们找出对称性隐藏的抽象结构，从而可以描述自然最基本的层面。

首先导出 Lorentz 群及其子群的特性。Lorentz 群定义为 Minkowski 空间中所有保证内积不变的变换组成的集合⁹⁶：

$$x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (3.122)$$

其中 $\eta_{\mu\nu}$ 表示 Minkowski 度规：

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (3.123)$$

⁹⁵ 引自 Robert S. Root-Bernstein and Michele M. Root-Bernstein. Sparks of Genius. Mariner Books, 1st edition, 8 2001. ISBN 9780618127450

⁹⁶ 下式的推导在第 2 章。重复一下，Lorentz 群的定义借鉴了 Euclidean 空间中旋转、反射变换保证 Euclidean 空间中内积不变的特性。

³⁸ 原文下式 J_3 的形式有误，译文已按勘误表修改。

这也就是我们把 Lorentz 群叫做 $\mathcal{O}(1,3)$ 的原因，类似地，群 $\mathcal{O}(4)$ 就是保持内积 $(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$ 的群。这个定义的条件对变换（矩阵）有什么限制？Lorentz 群的群元 — Lorentz 变换通常记作 Λ ，考虑变换 $x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu$ ，变换前后的内积为：

$$x^\mu \eta_{\mu\nu} x^\nu \rightarrow x'^\sigma \eta_{\sigma\rho} x'^\rho = (x^\mu \Lambda^\sigma_\mu) \eta_{\sigma\rho} (\Lambda^\rho_\nu x^\nu) \stackrel{!}{=} x^\mu \eta_{\mu\nu} x^\nu \quad (3.124)$$

由 x^μ 的任意性可得：

$$\Lambda^\sigma_\mu \eta_{\sigma\rho} \Lambda^\rho_\nu \stackrel{!}{=} \eta_{\mu\nu} \quad (3.125)$$

⁹⁷ 复习：两个向量的内积用矩阵乘法表示是，左侧的向量转置成行矩阵。因此下式左侧有 Λ^T 。

上式写成矩阵形式⁹⁷：

$$\Lambda^T \eta \Lambda \stackrel{!}{=} \eta \quad (3.126)$$

上式就是 Lorentz 变换 Λ 定义的数学形式！对上式取行列式，利用 $\det(AB) = \det(A) \det(B)$ 可得：

$$\det(\Lambda) \underbrace{\det(\eta)}_{=-1} \det(\Lambda) = \underbrace{\det(\eta)}_{=-1} \rightarrow \det(\Lambda)^2 \stackrel{!}{=} 1 \quad (3.127)$$

$$\rightarrow \det(\Lambda) \stackrel{!}{=} \pm 1 \quad (3.128)$$

⁹⁸ 稍后就会看到这有什么用。

取(3.125)式的 $\mu = \nu = 0$ 分量⁹⁸：

$$\begin{aligned} \Lambda^\sigma_0 \eta_{\sigma\rho} \Lambda^\rho_0 &\stackrel{!}{=} \underbrace{\eta_{00}}_{=1} \\ \rightarrow \Lambda^\sigma_0 \eta_{\sigma\rho} \Lambda^\rho_0 &= (\Lambda^0_0)^2 - \sum_i (\Lambda^i_0)^2 \stackrel{!}{=} 1 \end{aligned} \quad (3.129)$$

由此可得：

$$\Lambda^0_0 \stackrel{!}{=} \pm \sqrt{1 + \sum_i (\Lambda^i_0)^2} \quad (3.130)$$

⁹⁹ 这意味着右手坐标系在变换后仍为右手系，左手坐标系在变换后仍为左手系。左手、右手坐标系的定义见附录A.5

根据(3.128)与(3.130)式的正负可以把 Lorentz 群分成 4 分支。保证坐标系取向⁹⁹不变的两个分支满足 $\det(\Lambda) = +1$ 。如果想保证

时间轴的方向不变, 则需要条件 $\Lambda_0^0 \geq 0$, 因为

$$x^0 = t \rightarrow x'^0 = t' = \Lambda_\nu^0 x^\nu = \Lambda_0^0 t + \Lambda_1^0 x^1 + \Lambda_2^0 x^2 + \Lambda_3^0 x^3 \quad (3.131)$$

如果 $\Lambda_0^0 \geq 0$, 则 t' 与 t 同号。

满足 $\det(\Lambda) \geq 0$ 与 $\Lambda_0^0 \geq 0$ 的分支称为 $\mathcal{SO}(1, 3)^\uparrow$ 。我们主要研究这个子群, 它的名字很漂亮 — 正规¹⁰⁰或正时¹⁰¹Lorentz 群。Lorentz 群的这四分支是分离的, 因为从一个分支出发, 仅用该分支的 Lorentz 变换无法得到其他任何分支的元素。但引入两个新变换之后就可以从 $\mathcal{SO}(1, 3)^\uparrow$ 出发得到其他分支¹⁰²:

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (3.132)$$

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.133)$$

Λ_P 称为宇称算符, 相应的宇称变换其实就是镜面反射。 Λ_T 称为时间反演算符。

于是完整的 Lorentz 群 $\mathcal{O}(1, 3)$ 可以表示为如下集合:

$$\mathcal{O}(1, 3) = \{\mathcal{SO}(1, 3)^\uparrow, \Lambda_P \mathcal{SO}(1, 3)^\uparrow, \Lambda_T \mathcal{SO}(1, 3)^\uparrow, \Lambda_P \Lambda_T \mathcal{SO}(1, 3)^\uparrow\} \quad (3.134)$$

这样, 研究 Lorentz 群的表示就可简化为研究 $\mathcal{SO}(1, 3)^\uparrow$ 的表示, 因为只要再找出 Λ_P 、 Λ_T 的表示, Lorentz 群的其他分支就都有啦。

3.7.1 One Representation of the Lorentz Group Lorentz 群的 1 表示

下面利用 Lorentz 群的定义 ((3.125) 式) 构造 Lorentz 变换的显式矩阵表示。首先考虑矩阵的形式。对 Lorentz 群, 它的群元作用于 4 向量¹⁰³, 因此群元的形式为 4×4 实数矩阵。矩阵元必须

¹⁰⁰ ‘正规’ (proper) 意思是 $\mathcal{SO}(1, 3)^\uparrow$ 子群的群元都是保证坐标系取向和宇称不变的变换。

¹⁰¹ ‘正时’ (orthochronous) 是指子群的群元保持时间轴的方向不变。

¹⁰² 在现在的群表示下这两个算符的矩阵长这样。后面换用不同的表示, 它们会大不一样。

¹⁰³ 通常狭义相对论涉及的向量空间是四维实 Minkowski 空间 $R^{(1,3)}$ 。Lorentz 群就是定义在这个空间的 (Lorentz 变换就是保证 4×4 度规不变的变换), 因此首先研究群在 $R^{(1,3)}$ 上的表示。类似地, $SU(2)$ 群起初是用 2×2 复数矩阵表示的, 从那里出发可以得到 $SU(2)$ 好多其他的表示。

¹⁰⁴ 自行验证 $\Lambda^T \eta \Lambda = \eta$ 对 Λ 的限制条件是 10 个。留做习题

¹⁰⁵ 空间部分是指分量 $\mu = 1, 2, 3$ 对应的部分。空间部分通常记作 η_{ij} , 人们习惯用拉丁字母 (例如 i, j) 表示从 1 到 3, 而希腊字母 (例如 μ, ν) 表示从 0 到 3。

¹⁰⁶ $\eta_{11} = \eta_{22} = \eta_{33} = -1, \eta_{ij} = 0, \text{若 } i \neq j$

是实数, 因为 Minkowski 空间 $R^{(1,3)}$ 是实数空间。一般的 4×4 实数矩阵有 16 个参数, 而 Lorentz 群的定义规定了 10 个条件¹⁰⁴, 这样还剩 6 个自由参数。也就是说描述最一般的 Lorentz 变换需要 6 个参数。因此只要集齐 6 个线性无关的生成元, 就可以召唤 Lorentz 群的 Lie 代数。这 6 个生成元构成 Lie 代数的一组基, 即任意生成元可唯一地表示为它们的线性组合。然后就能计算这些基在 Lie 括号下的运算结果, 从而导出 Lie 代数的抽象定义。

注意三维 Euclidean 空间中的旋转矩阵 (它们只变换空间坐标而不管时间) 满足定义(3.125)式。这是因为 Minkowski 度规的空间部分¹⁰⁵与 3×3 单位矩阵成比例¹⁰⁶。因此对于只涉及空间坐标的变换, 由(3.125)可得:

$$\begin{aligned} -R^T I_{3 \times 3} R &= -R^T R \stackrel{!}{=} -I_{3 \times 3} \\ \rightarrow R^T I_{3 \times 3} R &= R^T R \stackrel{!}{=} I_{3 \times 3} \end{aligned}$$

上式正是 $\mathcal{O}(3)$ 群的定义式。再加上条件

$$\det(\Lambda) \stackrel{!}{=} 1$$

就是 $\mathcal{SO}(3)$ 群的定义。只涉及空间坐标的 Lorentz 变换的形式为:

$$\Lambda_{\text{rot}} = \begin{pmatrix} 1 & \\ & R_{3 \times 3} \end{pmatrix}$$

其中旋转矩阵 $R_{3 \times 3}$ 的形式见(3.23)式, 相关推导见 3.4.1 节。相应的生成元的推导与 3.4.1 节相似, 结果为:

$$J_i = \begin{pmatrix} 0 & \\ & J_i^{3\text{dim}} \end{pmatrix}, \quad i = 1, 2, 3. \quad (3.135)$$

比如, 由(3.65)式可得:

$$J_1 = \begin{pmatrix} 0 & \\ & J_1^{3\text{dim}} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (3.136)$$

下面包含时空坐标的变换。按 Lie 理论的套路, 考虑无穷小变

换¹⁰⁷:

$$\Lambda_\rho^\mu \approx \delta_\rho^\mu + \epsilon K_\rho^\mu \quad (3.137)$$

将上式带入 Λ 的定义条件(3.125)式:

$$\begin{aligned} \Lambda_\rho^\mu \eta_{\mu\nu} \Lambda_\sigma^\nu &\stackrel{!}{=} \eta_{\rho\sigma} \\ \rightarrow (\delta_\rho^\mu + \epsilon K_\rho^\mu) \eta_{\mu\nu} (\delta_\sigma^\nu + \epsilon K_\sigma^\nu) &\stackrel{!}{=} \eta_{\rho\sigma} \\ \rightarrow \eta_{\rho\sigma} + \epsilon K_\rho^\mu \eta_{\mu\sigma} + \epsilon K_\sigma^\nu \eta_{\rho\nu} + \underbrace{\epsilon^2 K_\rho^\mu \eta_{\mu\nu} K_\sigma^\nu}_{\text{略去高阶小量}\epsilon^2\text{项}} &= \eta_{\rho\sigma} \\ \rightarrow K_\rho^\mu \eta_{\mu\sigma} + K_\sigma^\nu \eta_{\rho\nu} &= 0 \end{aligned} \quad (3.138)$$

写成矩阵形式¹⁰⁸:

$$K^T \eta = -\eta K \quad (3.139)$$

上面导出了时空变换的生成元 K 需满足的条件。这些生成元生成的变换称为**推动**。推动的意思是变换到相对于初始坐标系作匀速直线运动的另一坐标系。比如从一个相对于观者静止的坐标系变换到相对于观者匀速直线运动的坐标系。下面考虑2.1节中的例子: 沿 x 轴的推动变换。由于 $y' = y, z' = z$, 故生成元的形式为:

$$K_x = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\equiv k_x} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (3.140)$$

只需要考虑非零的 2×2 矩阵部分就行了, (3.139)式简化为:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

不难解得¹⁰⁹:

$$k_x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

¹⁰⁷ Kronecker δ 符号定义为

$$\delta_\rho^\mu = \begin{cases} 1, & \mu = \rho \\ 0, & \mu \neq \rho \end{cases}$$

将 Kronecker δ 符号写成矩阵形式就是恒等矩阵 I .

¹⁰⁸ 第一个下标表示行指标, 第二个是列指标。不过指标为一上一下的时候会混淆, 这时我们把上下标错开表示, 比如 $K_\rho^\mu \equiv K_\rho^\mu \rightarrow (K^T)^\mu_\rho = K_\rho^\mu$ 。矩阵乘法的定义是用行乘以列, 因此 $K_\sigma^\nu \eta_{\rho\nu} = \eta_{\rho\nu} K_\sigma^\nu$ 的含义为, 矩阵 η 的第 ρ 行乘以 K 的第 σ 列, 用矩阵表示为 ηK 。此外, $K_\rho^\mu \eta_{\mu\sigma} = K_\rho^\mu \eta_{\mu\sigma} = (K^T)_\rho^\mu \eta_{\mu\sigma}$, 为将上式表示为矩阵乘法而引入矩阵 K 的转置, 只有这样才能表示为行乘以列的形式: K^T 的第 ρ 行乘以 η 的第 σ 列, 上式用矩阵表示为 $K^T \eta$ 。用指标记号可以自由移项, 比如 K_ρ^μ 只是矩阵 K 得一个元素, 即一个数。

$$\begin{aligned} &^{109} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \\ &\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \text{ 以及} \\ & - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \\ & - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

因此沿 x 轴的推动变换的生成元为：

$$K_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (3.141)$$

同理可以求出沿 y 、 z 轴推动变换的生成元：

$$K_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (3.142)$$

¹¹⁰ 生成元 K_x, K_y, K_z 都是厄米的，即 $K_i^\dagger = K_i$ 。因此 e 指数里不加虚数单位 i ，因为加了之后就变成反厄米的了。

根据 Lie 理论，把生成元扔进 e 指数里就能得到有限大小的变换¹¹⁰：

$$\Lambda_x(\phi) = e^{\phi K_x}$$

¹¹¹ 不难算出 $k_x^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ，然后 $k_x^4 = \mathbf{1}$ 等等等等。对于奇次幂有 $k_x^3 = k_x, k_x^5 = k_x, \dots$

还以 K_x 为例，简明起见，只考虑其中 exciting 的部分 — 2×2 矩阵 k_x （见(3.140)式）。将它的指数函数用级数展开式计算，并利用¹¹¹ $k_x^2 = \mathbf{1}$ ：

$$\begin{aligned} \Lambda_x(\phi) &= e^{\phi k_x} = \sum_{n=0}^{\infty} \frac{\phi^n k_x^n}{n!} = \sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \underbrace{k_x^{2n}}_{=1} + \sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \underbrace{k_x^{2n+1}}_{=k_x} \\ &= \left(\sum_{n=0}^{\infty} \frac{\phi^{2n}}{(2n)!} \right) I + \left(\sum_{n=0}^{\infty} \frac{\phi^{2n+1}}{(2n+1)!} \right) k_x \\ &= \cosh(\phi) I + \sinh(\phi) k_x \\ &= \begin{pmatrix} \cosh(\phi) & 0 \\ 0 & \cosh(\phi) \end{pmatrix} + \begin{pmatrix} 0 & \sinh(\phi) \\ \sinh(\phi) & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\phi) & \sinh(\phi) \\ \sinh(\phi) & \cosh(\phi) \end{pmatrix} \end{aligned} \quad (3.143)$$

计算过程与第3.4.1节的内容相似。这里的求和式里没出现 $(-1)^n$ ，因此结果不是 $\sin(\phi)$ 和 $\cos(\phi)$ ，而是双曲正弦 $\sinh(\phi)$ 和双曲余

弦 $\cosh(\phi)$ 。沿 x 轴推动变换的完整 4×4 矩阵为：

$$\Lambda_x = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) & 0 & 0 \\ \sinh(\phi) & \cosh(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.144)$$

同理可导出沿 y, z 轴的推动变换矩阵：

$$\Lambda_y = \begin{pmatrix} \cosh(\phi) & 0 & \sinh(\phi) & 0 \\ 0 & 1 & 0 & 0 \\ \sinh(\phi) & 0 & \cosh(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (3.145)$$

$$\Lambda_z = \begin{pmatrix} \cosh(\phi) & 0 & 0 & \sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(\phi) & 0 & 0 & \cosh(\phi) \end{pmatrix} \quad (3.146)$$

任意的推动变换都可以分解为上述三个变换矩阵的乘积。

3.7.2 Generators of the Other Components of the Lorentz Group Lorentz 群其他部分的生成元

为得到 Lorentz 群其他部分¹¹²生成元的形式，只需将宇称算符 Λ_P 和时间反演算符 Λ_T 作用于刚刚导出的生成元 J_i, K_i 上，用矩阵指标记法写为¹¹³：

$$(\Lambda_P)_{\alpha'}^{\alpha} (\Lambda_P)_{\beta'}^{\beta} (J_i)^{\alpha' \beta'} \underbrace{\hat{=} \Lambda_P J_i (\Lambda_P)^T}_{\text{变成矩阵形式}} = J_i \hat{=} (J_i)^{\alpha \beta} \quad (3.147)$$

$$(\Lambda_P)_{\alpha'}^{\alpha} (\Lambda_P)_{\beta'}^{\beta} (K_i)^{\alpha' \beta'} \underbrace{\hat{=} \Lambda_P K_i (\Lambda_P)^T}_{\text{变成矩阵形式}} = -K_i \hat{=} - (K_i)^{\alpha \beta} \quad (3.148)$$

¹¹² 复习：Lorentz 群可以分成 $\mathcal{O}(1,3) = \{\mathcal{SO}(1,3)^\uparrow, \Lambda_P \mathcal{SO}(1,3)^\uparrow, \Lambda_T \mathcal{SO}(1,3)^\uparrow, \Lambda_P \Lambda_T \mathcal{SO}(1,3)^\uparrow\}$ ，上一节导出了 $\mathcal{SO}(1,3)^\uparrow$ 的生成元。

¹¹³ 对 J, K 变换需要两个 Λ_P ，因为 J, K 都是二维矩阵，一维一个 Λ_P 。这就是算符矩阵在坐标变换下的一般形式，喵。

上两式可以用上节导出的 J, K 显式矩阵形式暴力计算, 例如³⁹:

$$J_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow J'_x = \Lambda_P J_x (\Lambda_P)^T = J_x \quad (3.149)$$

因为

$$\begin{aligned} \Lambda_P J_x (\Lambda_P)^T &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^T \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{aligned} \quad (3.150)$$

同理,

$$K_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow K'_x = \Lambda_P K_x (\Lambda_P)^T = -K_x \quad (3.151)$$

因为

$$\begin{aligned} \Lambda_P K_x (\Lambda_P)^T &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^T \\ &= - \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (3.152)$$

³⁹原文(3.149), (3.150)式有误, 译文已按勘误表修改。

综上，在宇称变换下：

$$\underbrace{J_i \longrightarrow J_i}_{\Lambda_P}, \quad \underbrace{K_i \longrightarrow -K_i}_{\Lambda_P} \quad (3.153)$$

它们之后很有用，因为宇称算符在其他表示之下的形式不会像现在的向量表示这样显然。时间反演算符 Λ_T 的推导过程相似，由于 Λ_T 只变换第一个分量（时间分量），因此只对 K_i 有影响：

$$\underbrace{(\Lambda_T)_{\alpha'}^{\alpha} (\Lambda_T)_{\beta'}^{\beta} (J_i)^{\alpha' \beta'} \triangleq \Lambda_T J_i (\Lambda_T)^T}_{\text{变成矩阵形式}} = J_i \triangleq (J_i)^{\alpha \beta} \quad (3.154)$$

$$\underbrace{(\Lambda_T)_{\alpha'}^{\alpha} (\Lambda_T)_{\beta'}^{\beta} (K_i)^{\alpha' \beta'} \triangleq \Lambda_T K_i (\Lambda_T)^T}_{\text{变成矩阵形式}} = -K_i \triangleq -(K_i)^{\alpha \beta} \quad (3.155)$$

或者简写为：

$$\underbrace{J_i \longrightarrow J_i}_{\Lambda_T}, \quad \underbrace{K_i \longrightarrow -K_i}_{\Lambda_T} \quad (3.156)$$

3.7.3 The Lie Algebra of the Proper Orthochronous Lorentz Group

$\mathcal{O}(1,3)^{\uparrow}$ 群的 Lie 代数

利用上节导出的 $\mathcal{SO}(1,3)^{\uparrow}$ 生成元的矩阵形式¹¹⁴可以直接计算相应的 Lie 代数¹¹⁵：

$$[J_i, J_j] = i\epsilon_{ijk} J_k \quad (3.157)$$

$$[J_i, K_j] = i\epsilon_{ijk} K_k \quad (3.158)$$

$$[K_i, K_j] = -i\epsilon_{ijk} J_k \quad (3.159)$$

其中 J_i 是旋转变换的生成元， K_i 是推动的生成元。一般形式的 Lorentz 变换可以表示为：

$$\Lambda = e^{iJ\theta + iK\Phi} \quad (3.160)$$

(3.158)式表明两种生成元 (J_i 与 K_i) 之间不对易。旋转生成元在对易运算下封闭¹¹⁶，而推动生成元则不是这样¹¹⁷。从 J_i, K_i 出发

¹¹⁴ 推动变换的生成元见(3.141)式，旋转变换的生成元见(3.62)式。

¹¹⁵ Levi-Civita 符号 ϵ_{ijk} 的定义见附录B.5.5

¹¹⁶ 在对易运算下封闭意味着对易子 $[J_i, J_j] = J_i J_j - J_j J_i$ 还是一个旋转变换的生成元。(3.157)体现了这一点。

¹¹⁷ (3.159)表明两个推动生成元 K_i, K_j 的对易子并非推动生成元，而是个转动生成元。

可以定义在对易运算下封闭并且相互对易的新生成元⁴⁰:

$$N_i^\pm \equiv \frac{1}{2}(J_i \pm iK_i) \quad (3.161)$$

不难算出对易关系为:

$$[N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+ \quad (3.162)$$

$$[N_i^-, N_j^-] = i\epsilon_{ijk}N_k^- \quad (3.163)$$

$$[N_i^+, N_j^-] = 0 \quad (3.164)$$

这正是 $SU(2)$ Lie 代数的形式! 我们发现了 $\mathcal{SO}(1,3)^\dagger$ 群的 Lie 代数包含两份 $SU(2)$ 的 Lie 代数。

我们搞了个大新闻, 因为前面已经构造了 $SU(2)$ 的所有不可约表示。然而就像 $\mathcal{SO}(3)$ 那样, Lorentz 群不是单连通群¹¹⁸, Lie 理论告诉我们, 对于非单连通群, 不存在 Lie 代数的不可约表示和群的表示之间的一一映射¹¹⁹。通过推导 Lorentz 群 Lie 代数的不可约表示, 可以导出 Lorentz 群的覆盖群的不可约表示, 只要把相应的生成元扔进 e 指数里就行了。这样导出的一部分表示是 Lorentz 群的表示, 还有一些不是。这些额外的表示很有用处, 它们可以用来描述某些基本粒子。

简明起见, 我们把之后要推导的 Lorentz 群 Lie 代数的表示, 或者 Lorentz 群双覆盖的表示, 全部简称为 Lorentz 群的表示。

$SU(2)$ 的每一不可约表示都可以用 $SU(2)$ 的 Casimir 算符对应的标量 j 来标记。因此 Lorentz 群覆盖群¹²⁰的不可约表示可以用两个数 j_1, j_2 (它们都是整数或半整数) 来标记。下面依次研究 Lorentz 群的 (j_1, j_2) 表示, 并利用前面导出的 $SU(2)$ 群的 $j_1, j_2 = 0, \frac{1}{2}, 1, \dots$ 表示的形式。

Lorentz 群的 Lie 代数可以用 $M_{\mu\nu}$ 表示为更紧凑的形式, $M_{\mu\nu}$ 的定义为:

$$J_i = \frac{1}{2}\epsilon_{ijk}M_{jk} \quad (3.165)$$

$$K_i = M_{0i} \quad (3.166)$$

¹¹⁸ 本书不证明这个命题, 证明过程很复杂并且与物理关系不大。

¹¹⁹ 这话理解起来有点难。前面提过, 一个 Lie 代数总对应一个特殊的群, 这个群的特殊之处在于它是单连通的。导出一个 Lie 代数的不可约表示之后, 再将 Lie 代数的元素 (即生成元) 扔进 e 指数里就能得到单连通群 (覆盖群) 的表示。只有对于单连通群, 前面说的一一映射才存在。

¹²⁰ Lorentz 群的覆盖群记作 $SL(2, C)$, 它是具有单位行列式 (行列式为 1) 的 2×2 复数矩阵的集合。 $SL(2, C) \rightarrow \mathcal{SO}(1,3)$ 之间的关系与之前讨论的 $SU(2) \rightarrow \mathcal{SO}(3)$ 之间的关系类似。

⁴⁰ 注意区分之后式子中的虚数单位 i 与指标 i , 我们对作者采用这样的符号很抱歉。

这样 Lie 代数可表示为:

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}) \quad (3.167)$$

下面从 Lorentz 群的 Lie 代数构造不可约表示, さ、いくぞ!

3.7.4 The (0, 0) Representation (0, 0) 表示

$SU(2)$ 的最低阶表示是平凡的, 因为 Lie 代数对应是 1 维的向量空间, 因此生成元必须是 1×1 矩阵, 满足对易关系的只有平凡的 0:

$$N_i^+ = N_i^- = 0 \rightarrow e^{N_i^+} = e^{N_i^-} = e^0 = 1 \quad (3.168)$$

因此 Lorentz 群的 (0, 0) 表示作用于在 Lorentz 变换下不变的对象。(0, 0) 表示又称为 Lorentz 标量表示。

3.7.5 The $(\frac{1}{2}, 0)$ Representation $(\frac{1}{2}, 0)$ 表示

$(\frac{1}{2}, 0)$ 表示就是对 $SU(2)$ 的一份 Lie 代数 N_i^+ 使用 2 维表示¹²¹, 另一份 Lie 代数 N_i^- 使用 1 维表示, 即 $N_i^- = 0$ 。由 N_i^- 的定义式(3.161)可得:

$$N_i^- = \frac{1}{2}(J_i - iK_i) = 0 \quad (3.169)$$

$$\rightarrow J_i = iK_i \quad (3.170)$$

此外, 根据3.6.4节中导出的 $SU(2)$ 的 2 维表示的形式:

$$N_i^+ = \frac{\sigma_i}{2} \quad (3.171)$$

其中 σ_i 表示 Pauli 矩阵, 其定义见(3.81)式。

另一方面:

$$N_i^+ \stackrel{(3.161)\text{式}}{=} \frac{1}{2}(J_i + iK_i) \stackrel{(3.170)\text{式}}{=} \frac{1}{2}(iK_i + iK_i) = iK_i \quad (3.172)$$

比较(3.171)与(3.172)式可得:

$$iK_i = \frac{\sigma_i}{2} \rightarrow K_i = \frac{\sigma_i}{2i} = \frac{i\sigma_i}{2i^2} = -\frac{i}{2}\sigma_i \quad (3.173)$$

$$(3.170)\text{式} \rightarrow J_i = iK_i = -\frac{i^2}{2}\sigma_i = \frac{1}{2}\sigma_i \quad (3.174)$$

¹²¹ 复习: 表示对应的向量空间的维数是 $2j + 1$, 因此 $j = 1/2$ 表示对应 $2 \times \frac{1}{2} + 1 = 2$ 维向量空间。

于是在这一表示下, Lorentz 旋转变换为:⁴¹

$$R_{\vec{\theta}} = e^{i\vec{\theta}\vec{J}} = e^{i\vec{\theta}\frac{\vec{\sigma}}{2}} \quad (3.175)$$

Lorentz 推动变换为⁴²:

$$B_{\vec{\phi}} = e^{i\vec{\phi}\vec{K}} = e^{\vec{\phi}\frac{\vec{\sigma}}{2}} \quad (3.176)$$

将生成元扔进 e 指数函数里, 再将 e 指数函数用级数展开, 就可以算出 Lorentz 群元的表示。例如, 绕 x 轴的旋转变换为:

$$R_x(\theta) = e^{i\theta J_1} = e^{i\theta\frac{1}{2}\sigma_1} = 1 + \frac{i}{2}\theta\sigma_1 + \frac{1}{2!}\left(\frac{i}{2}\theta\sigma_1\right)^2 + \cdots + \frac{1}{n!}\left(\frac{i}{2}\theta\sigma_1\right)^n + \cdots \quad (3.177)$$

¹²² 具体过程与 3.4.1 中的推导十分相似。

将 σ_1 的显式形式 ((3.81) 式) 代入, 并利用 $\sigma_1^2 = \mathbf{1}$ 可得¹²²:

$$\begin{aligned} R_x(\theta) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{i}{2}\theta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\left(\frac{\theta}{2}\right)^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \cdots \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & i \sin \frac{\theta}{2} \\ i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} \end{aligned} \quad (3.178)$$

关于其他轴的旋转及推动变换可类似地导出。注意 Lorentz 变换在这儿用 2×2 复矩阵表示, 它们当然不是作用在 Minkowski 空间的 (Minkowski 空间的元素是 4 维向量)。Lorentz 变换在 $(\frac{1}{2}, 0)$ 表示下作用的 2 分量¹²³对象称为左手旋量¹²⁴:

$$\chi_L = \begin{pmatrix} (\chi_L)_1 \\ (\chi_L)_2 \end{pmatrix} \quad (3.179)$$

旋量在这里指的是 2 分量对象。左手旋量的另一种定义是在 Lorentz 变换下根据 Lorentz 群的 $(\frac{1}{2}, 0)$ 表示变换的对象。注意旋量并不只是 2 分量的向量, 旋量有很多新的性质。比如 e 指数中的 $\frac{1}{2}$ 因子, 这个因子使得旋量¹²⁵在旋转 2π 之后不再和原来相同, 而多了一个负号。这个性质十分奇特, 因为日常遇到的对象在转 2π 之后总是和原来一样。

¹²³ 之后会看到这些二分量对象与自旋向上和自旋向下的状态有关。

¹²⁴ 这个名字在引入右手旋量之后的含义会更清楚。我们会看到宇称变换将左手旋量变换为右手旋量, 反之亦然。旋量通常分为左手和右手的, 它们的起源与螺旋性有关 — 螺旋性与手性不完全相同。重复一遍宇称算符的效果: 将左手坐标系变换为右手坐标系, 反之亦然。左手、右手旋量的名字就是从那儿来的。

¹²⁵ 旋量的更多性质可见,

J. J. Sakurai.

Modern Quantum Mechanics.

Addison Wesley, 1st edition, 9

1993. ISBN 9780201539295 中

的第 3.2 章

⁴¹ 译者注: 式子采用了一个缩写, 见边注 3.129。

⁴² 原文为 B_{θ} , 有误, 已据勘误表修改为 B_{ϕ} — 译者

“旋量可以说是 Lorentz 变换最基本的数学对象。”

— A. M. Steane¹²⁶

¹²⁶ Andrew M. Steane.
<https://arxiv.org/abs/1312.3824> An introduction to spinors. ArXiv e-prints, December 2013

3.7.6 The $(0, \frac{1}{2})$ Representation $(0, \frac{1}{2})$ 表示

$(0, \frac{1}{2})$ 表示的构造过程与之前的 $(\frac{1}{2}, 0)$ 相似。这里 N_i^+ 采用 1 维表示, 即 $N_i^+ = 0$, N_i^- 采用 2 维表示, 即 $N_i^- = \frac{1}{2}\sigma_i$ 。容易认为 $(0, \frac{1}{2})$ 表示会与 $(\frac{1}{2}, 0)$ 表示极为相似, 但并非如此! 根据 (3.161) 式可得:

$$N_i^+ = \frac{1}{2}(J_i + iK_i) = 0 \quad (3.180)$$

$$\rightarrow J_i = -iK_i \quad (3.181)$$

注意上面的负号。 N_i^- 采用 $SU(2)$ 的二维表示 (在 3.6.4 节推导):

$$N_i^- = \frac{1}{2}\sigma_i = \frac{1}{2}(J_i - iK_i) \underset{(3.181)\text{式}}{=} \frac{1}{2}(-iK_i - iK_i) = -iK_i \quad (3.182)$$

可以得到 $(0, \frac{1}{2})$ 表示的生成元为:

$$-iK_i = \frac{1}{2}\sigma_i \rightarrow K_i = -\frac{1}{2i}\sigma_i = \frac{i}{2}\sigma_i \quad (3.183)$$

再结合 (3.181) 式:

$$J_i = -iK_i = \frac{1}{2}\sigma_i \quad (3.184)$$

于是, 在这一表示之下, Lorentz 旋转变换为:

$$R_\theta = e^{i\vec{\theta}\vec{J}} = e^{i\vec{\theta}\frac{\vec{\sigma}}{2}} \quad (3.185)$$

Lorentz 推动变换为⁴³:

$$B_\phi = e^{i\vec{\phi}\vec{K}} = e^{-\vec{\phi}\frac{\vec{\sigma}}{2}} \quad (3.186)$$

可见 $(0, \frac{1}{2})$ 表示的旋转变换与 $(\frac{1}{2}, 0)$ 一样, 但推动变换的 e 指数幂差一个负号。我们推测这两个表示作用的对象相似但不完全相

⁴³原文为 B_θ , 有误, 已据勘误表修改为 B_ϕ — 译者

同。Lorentz 群的 $(0, \frac{1}{2})$ 表示作用的对象称为右手旋量：

$$\chi_R = \begin{pmatrix} (\chi_R)^1 \\ (\chi_R)^2 \end{pmatrix} \quad (3.187)$$

左手/右手旋量统称为 Weyl 旋量。

3.7.7 Van der Waerden Notation Van der Waerden 符号

本节引入便于处理旋量的标记法。上面导出了按不同方式变换的两种旋量，如何将它们区分？它们确实不同，但也并非毫无关系。实际上左手旋量（按 $(\frac{1}{2}, 0)$ 表示变换）与右手旋量（按 $(0, \frac{1}{2})$ 表示变换）之间存在联系。引入带点与不带点的指标来表示不同旋量——这一记法以发明者命名，称为 Van der Waerden 符号。这能清楚地表示何种对象按何种方式变换。Van der Waerden 符号的优点在完整的体系建立起来之后会更清楚。

左手旋量 χ_L 用不带点的下指标表示：

$$\chi_L = \chi_a \quad (3.188)$$

右手旋量用带点的上指标：

$$\chi_R = \chi^{\dot{a}} \quad (3.189)$$

¹²⁷ 注意下式就是二维 Levi-Civita 符号，其定义见附录 B.5.5。

下面引入“旋量度规”的概念。旋量度规使得右手旋量可以变换为左手旋量，反之亦然。旋量度规定义为¹²⁷：

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (3.190)$$

不难验证它满足上述性质。再定义¹²⁸：

$$\chi_L^C \equiv \epsilon \chi_L^* \quad (3.191)$$

其中 $*$ 表示取复共轭。下面将考虑 χ_L^C 怎样在 Lorentz 变换下变换，并会发现它正如同右手旋量一样变换。右手旋量由它在 Lorentz 变换下的变换方式所定义，因此认为 χ_L^C 正是一个右手旋量。下面让

¹²⁸ 这里对 χ_L^C 做一简介。上标 C 表示电荷共轭性，3.7.10 节具体说明它的含义。我们看到 ϵ 的作用结果是把对象的“标签”翻转——将左手旋量变换为右手旋量。之后会发现 ϵ 可以翻转所有“标签”，比如，电荷的正负。

推动变换作用到 χ_L^C ，利用⁴⁴：

$$(-\epsilon)(\epsilon) = 1 \quad (3.192)$$

$$(\epsilon)\sigma_i^*(-\epsilon) = -\sigma_i \quad (3.193)$$

上式不难用 Pauli 矩阵 σ_i 的显式形式直接验证。变换过程为¹²⁹：

$$\begin{aligned} \chi_L^C \rightarrow (\chi')_L^C &= \epsilon(\chi')_L^* \\ &= \epsilon(e^{\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L)^* \\ &= \epsilon(e^{\frac{\vec{\phi}}{2}\vec{\sigma}}(-\epsilon)(\epsilon)\chi_L)^* \\ &\quad \quad \quad =1, \text{见(3.192)式} \\ &= \underbrace{\epsilon(e^{\frac{\vec{\phi}}{2}\vec{\sigma}*}(-\epsilon)(\epsilon)\chi_L^*)}_{\text{(3.193)式} \rightarrow e^{-\frac{\vec{\phi}}{2}\vec{\sigma}}} \\ &= e^{-\frac{\vec{\phi}}{2}\vec{\sigma}} \underbrace{\epsilon\chi_L^*}_{=\chi_L^C} \\ &= e^{-\frac{\vec{\phi}}{2}\vec{\sigma}}\chi_L^C \end{aligned} \quad (3.194)$$

由此可见 χ_L^C 按右手旋量的形式变换¹³⁰。上式第 5 行利用了 $e^{\frac{\vec{\phi}}{2}\vec{\sigma}}$ 的 Taylor 级数展开，并对每一项利用了 (3.193) 式。不难验证 $\epsilon(\chi_L^*)$ 在旋转变换下仍与原来的变换方式相同，理应如此，因此 χ_L 与 χ_R 在旋转变换形式相同：

$$\chi_L^C \rightarrow (\chi')_L^C = \epsilon(\chi')_L^* = \epsilon(e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\chi)_L^* = e^{\frac{i\vec{\theta}}{2}\vec{\sigma}}\epsilon(\chi_L)^* \quad (3.195)$$

还可以验证以下结果： ϵ 在任意 Lorentz 变换下不变；将右手旋量变换为左手旋量需要用 $(-\epsilon)$ 。

仿照狭义相对论中的张量记号，定义“旋量度规”可以升降旋量的指标：

$$\underbrace{\epsilon\chi_L}_{\text{用指标记号表示}} = \epsilon^{ac}\chi_c = \chi^a \quad (3.196)$$

其中重复的上下标表示求和（即 Einstein 求和约定）。此外，从上面可知如果想把 χ_R 变换为 χ_L 还需要取复共轭：

$$\chi_R = \epsilon\chi_L^* \quad (3.197)$$

¹²⁹下面

$\vec{\phi}\vec{\sigma} \equiv \sum_i \phi_i \sigma_i$ 求和约定
， $\vec{\sigma}$ 并非表示“向量”，而只是为了方便而采用的缩写。

¹³⁰ 右手旋量在推动变换下的形式见(3.186)式： $B_\phi = e^{i\vec{\phi}\vec{K}} = e^{-\vec{\phi}\frac{\vec{\sigma}}{2}}$ 。注意它与左手旋量的变换 ((3.176)式)的不同： $B_\phi = e^{i\vec{\phi}\vec{K}} = e^{\vec{\phi}\frac{\vec{\sigma}}{2}}$

⁴⁴其中要注意 $(-\epsilon) = \epsilon^T$ ，所以才这么写。--译者注

用 Van der Waerden 符号, ϵ 作用在左手旋量的复共轭上得到的右手旋量用带点的上指标表示:

$$\chi_R = \epsilon \chi_L^* = \chi^{\dot{a}} \quad (3.198)$$

再与指标记号相比较, 可以发现 χ_L 的复共轭可以用带点的下指标表示:

$$\chi_L^* = \chi_a^* = \chi_{\dot{a}} \quad (3.199)$$

同理, 右手旋量 χ_R 的复共轭可用不带点的上指标表示:

$$\chi_R^* = (\chi^{\dot{a}})^* = \chi^a \quad (3.200)$$

下面考虑 $\chi_{\dot{a}}$ 与 χ^a 在 Lorentz 变换下如何改变, 这非常重要, $\chi_{\dot{a}}$ 与 χ^a 可以构造在 Lorentz 变换下不变的某种旋量的组合形式¹³¹。

从左手旋量在 Lorentz 变换下的行为出发:

$$\chi_L = \chi_a \xrightarrow{\text{Lorentz 变换}} \chi'_a = \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^b \chi_b \quad (3.201)$$

考虑旋量 $\chi_{\dot{a}}$ 的变换:

$$\begin{aligned} \chi_L^* = \chi_a^* = \chi_{\dot{a}} &\xrightarrow{\text{Lorentz 变换}} \chi'_{\dot{a}} = (\chi'_a)^* = \left(\left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^b \right)^* \chi_b^* \\ &= \left(e^{-i\vec{\theta} \frac{\vec{\sigma}^*}{2} + \vec{\phi} \frac{\vec{\sigma}^*}{2}} \right)_{\dot{a}}^{\dot{b}} \chi_{\dot{b}} \end{aligned} \quad (3.202)$$

类似地, 利用右手旋量的变换形式:

$$\chi_R \xrightarrow{\text{Lorentz 变换}} \chi'_R = \chi'^{\dot{a}} = \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} - \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_{\dot{b}}^{\dot{a}} \chi^{\dot{b}} \quad (3.203)$$

可以导出旋量 χ^a 的变换:

$$\begin{aligned} \chi_R^* = (\chi^{\dot{a}})^* = \chi^a &\xrightarrow{\text{Lorentz 变换}} \chi'^a = (\chi'^{\dot{a}})^* = \left(\left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} - \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_{\dot{b}}^{\dot{a}} \right)^* (\chi^{\dot{b}})^* \\ &= \left(e^{-i\vec{\theta} \frac{\vec{\sigma}^*}{2} - \vec{\phi} \frac{\vec{\sigma}^*}{2}} \right)_b^a \chi^b \end{aligned} \quad (3.204)$$

¹³¹ 这些旋量构造出的不变量极为重要, 因为我们要导出在所有惯性系下都成立的物理定律。这些不变量的重要性之后会不断体现。

为了组合出神秘的 Lorentz 不变量，还需要一件事：考虑两个向量标量积的定义： $\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b}$ 。仿照这个我们在旋量的标量积定义里也引入转置操作。注意上面新旋量变换式的 e 指数中出现了 Pauli 矩阵的复共轭 σ_i^* （例如 $e^{-i\vec{\theta} \frac{\vec{\sigma}^*}{2}}$ ），再加上转置操作就变成厄米共轭的形式： $\sigma_i^\dagger = (\sigma_i^*)^T$ ，其中前面已经解释过符号 \dagger （称为 dagger）的含义。所有 Pauli 矩阵的厄米共轭仍是 Pauli 矩阵自身，即：

$$\sigma_i^\dagger = (\sigma_i^*)^T = \sigma_i \quad (3.205)$$

这从 Pauli 矩阵的形式 (3.81) 式容易导出。

比较(3.201)与(3.204)式，再结合(3.205)式，可以看出旋量 χ_a 的转置与旋量 χ^a 分别对应的变换形式正好抵消，这意味着 $(\chi^a)^T \chi_a$ 是 Lorentz 变换下的不变量（即在 Lorentz 变换下不变），因为¹³²：

$$\begin{aligned} (\chi^a)^T \chi_a &\xrightarrow{\text{Lorentz 变换}} \\ (\chi'^a)^T \chi'_a &= \left(\left(e^{-i\vec{\theta} \frac{\vec{\sigma}^*}{2} - \vec{\phi} \frac{\vec{\sigma}^*}{2}} \right)_b^a \chi^b \right)^T \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^c \chi_c \\ &= (\chi^b)^T \left(\exp \left(-i\vec{\theta} \frac{(\vec{\sigma}^*)^T}{2} - \vec{\phi} \frac{(\vec{\sigma}^*)^T}{2} \right) \right)_b^a \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^c \chi_c \\ &\stackrel{(3.205) \text{ 式}}{=} (\chi^b)^T \underbrace{\left(e^{-i\vec{\theta} \frac{\vec{\sigma}}{2} - \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_b^a \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^c}_{=\delta_b^c} \chi_c \\ &= (\chi^c)^T \chi_c \end{aligned} \quad (3.206)$$

类似地，将 $\chi_{\dot{a}}$ 与 $\chi^{\dot{a}}$ 适当组合，利用(3.202)、(3.203)式可以得出另一个不变量 $(\chi_{\dot{a}})^T \chi^{\dot{a}}$ 。但是形如 $(\chi^{\dot{a}})^T \chi_a \doteq \chi_R^T \chi_L$ 的项不是 Lorentz 变换下的不变量，因为：

$$\chi_R^T \chi_L = (\chi^{\dot{a}})^T \chi_a \rightarrow (\chi'^{\dot{a}})^T \chi'_a = \chi^{\dot{b}} \underbrace{\left(e^{i\vec{\theta} \frac{\vec{\sigma}^T}{2} - \vec{\phi} \frac{\vec{\sigma}^T}{2}} \right)_{\dot{b}}^{\dot{a}} \left(e^{i\vec{\theta} \frac{\vec{\sigma}}{2} + \vec{\phi} \frac{\vec{\sigma}}{2}} \right)_a^c}_{\neq \delta_{\dot{b}}^c} \chi_c \quad (3.207)$$

可见这样的左手旋量与右手旋量结合的项并非 Lorentz 不变量。综上，Lorentz 不变量必须是一对同类上下标组合成的量¹³³。或者，不变量必须是右手旋量的厄米共轭左乘左手旋量： $\chi_R^\dagger \chi_L = (\chi_R^*)^T \chi_L \doteq (\chi^a)^T \chi_a$ ，或者左手旋量的厄米共轭左乘右手旋量： $\chi_L^\dagger \chi_R = (\chi_L^*)^T \chi_R \doteq (\chi_{\dot{a}})^T \chi^{\dot{a}}$ 。Lorentz 不变量是之后推导的自然界的普适定

¹³² 附录B.5.5中提到， δ_b^c 称为 Kronecker 符号，它是单位矩阵的指标形式，即 $\delta_b^c = \begin{cases} 1, & b = c \\ 0, & b \neq c \end{cases}$

¹³³ 即上下标都带点 \dot{a} 或都不带点 \dot{a}

律的关键。

在上述基础上, ϵ^{ab} 的名字“旋量度规”就变得合情合理了, 因为旋量间的内积 (在 Lorentz 变换下不变, 见(3.206)式) 可以写为:

$$\chi_a^T \chi^a \underset{(3.196) \text{ 式}}{=} \chi_a^T \epsilon^{ab} \chi_b \quad (3.208)$$

与第二章 Minkowski 空间的内积 (它是 Lorentz 变换的不变量) 的定义式(2.31)相对比:

$$x_\mu y^\mu = x_\mu \eta^{\mu\nu} y_\nu \quad (3.209)$$

其中 $\eta^{\mu\nu}$ 是 Minkowski 度规。可见旋量度规与四维向量的 Minkowski 度规地位相同¹³⁴。

在定义了上述符号之后, 不难写出旋量度规的下标形式:

$$\epsilon_{ab} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (3.210)$$

因为只有用 $(-\epsilon)$ 才能把 χ_R 转换为 χ_L 。^{135, 45} 一般的 Lorentz 变换 (转动 + 推动) 通常用 Λ 表示, 当 Λ 的指标带点, 说明它作用于右手旋量, 并由此推断相应的具体形式:

$$\chi_R \rightarrow \chi'_R = \chi^{\dot{a}} = \Lambda_{\dot{b}}^{\dot{a}} \chi^{\dot{b}} = (e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}})^{\dot{a}}_{\dot{b}} \chi^{\dot{b}} \quad (3.211)$$

同理, 左手旋量情形:

$$\chi_L \rightarrow \chi'_L = \chi'_a = \Lambda_a^b \chi_b = (e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}})^b_a \chi_b \quad (3.212)$$

这样 $(\frac{1}{2}, 0)$ 与 $(0, \frac{1}{2})$ 表示下的 Lorentz 变换记作:

$$\Lambda_{(\frac{1}{2}, 0)} = (e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}})^b_a \triangleq \Lambda_a^b \quad (3.213)$$

$$\Lambda_{(0, \frac{1}{2})} = (e^{i\vec{\theta}\frac{\vec{\sigma}}{2} - \vec{\phi}\frac{\vec{\sigma}}{2}})^b_a \triangleq \Lambda_{\dot{b}}^{\dot{a}} \quad (3.214)$$

⁴⁵根据(3.198), $\chi_R = \epsilon \chi_L^*$, 取共轭得到 $\chi_R^* = \epsilon \chi_L$, 再同时左乘 $\epsilon^{-1} = -\epsilon$ 得到

$$(-\epsilon) \chi_R^* = \epsilon^{-1} \epsilon \chi_L = \chi_L$$

¹³⁴ 为啥上式 (四维向量的内积的定义式) 当中没有转置操作? 其实上式既可以看做向量方程也可视为分量方程, 符号 x_μ 既代表向量本身又代表向量的分量, 不幸的是这有时会造成混乱, 比如现在。将上式看做分量方程, 则无需转置, 这对旋量内积也一样。然而转置运算不可或缺, 为了避免遗忘, 我们在旋量内积的定义式中加上上标 T 表示强调。对于常用的三维向量有更好的符号避免上述混乱: 向量自身用 \vec{a} , 向量分量用 a_i 表示。

¹³⁵ 读者可自行检验, 但这对之后的内容关系不大。

Van der Waerden 符号非常便捷，它有效地体现了 χ_L 与 χ_R 之间的联系。左右旋量之间可以互相变换，因此必然需要一套能将它们联系起来的符号体系。

之后考虑下一个不可约表示，有一位故人正在那里久候。

3.7.8 The $(\frac{1}{2}, \frac{1}{2})$ Representation $(\frac{1}{2}, \frac{1}{2})$ 表示

在 $(\frac{1}{2}, \frac{1}{2})$ 表示之下，两份 $SU(2)$ Lie 代数¹³⁶ N_i^+, N_i^- 都采用 2 维表示。首先考虑此表示下的 Lorentz 变换作用的对象是什么。两份 Lie 代数之间无相互作用，因为 N_i^+, N_i^- 是对易的，即 $[N_i^+, N_i^-] = 0$ 。因此作用的对象（记作 v ）在两份 Lie 代数下各自独立变换。 v 有两个指标 v_a^b ，每个指标在独立的 $SU(2)$ Lie 代数下变换。 v_a^b 即采用了上节介绍的 Van der Waerden 符号。

v_a^b 的每个指标有两个取值 ($\frac{1}{2}$ 和 $-\frac{1}{2}$)，⁴⁶ 因为每个表示都是 2 维的， v 共有 4 个分量。因此 v 可表示为 2×2 矩阵，但 v 也可以表示为四维向量，稍后会看到¹³⁷。

首先考虑复数矩阵，一个一般的 2×2 复数矩阵有 4 个复数元素，因此有 8 个自由参数。而据前所述我们只需要 4 个自由参数。任意复矩阵 M 可表示为厄米矩阵 $H (H^\dagger = H)$ 与反厄米矩阵 $A (A^\dagger = -A)$ 之和： $M = H + A$ 。 2×2 厄米矩阵和反厄米矩阵都有 4 个自由参数。此外，下面会证明厄米矩阵经 $(\frac{1}{2}, \frac{1}{2})$ 表示下的 Lorentz 变换后仍为厄米矩阵，反厄米矩阵也一样。这意味着厄米矩阵与反厄米矩阵分别组成的集合都是不变子空间（见 3.5 节），亦即，一般的 2×2 复矩阵上的表示是可约的。综上所述，下面研究 2×2 厄米矩阵（组成的向量空间）上的不可约表示。

2×2 厄米矩阵的一组基¹³⁸是 Pauli 矩阵¹³⁹ $\sigma_i, i = 1, 2, 3$ 加上单位矩阵。

首先考虑 v_{ab} ，而非 v_a^b ，这样可以利用(3.81)式的 Pauli 矩阵的显式形式。注意 v_a^b, v_{ab} 之间可用旋量度规 ϵ^{bc} 相互变换，因此容易从 v_a^b 回到 v_{ab} 。

记 $\sigma^0 = I_{2 \times 2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ，有⁴⁷

$$v_{ab} = v_\nu \sigma_{ab}^\nu = v_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + v_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + v_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

⁴⁶似乎指标应该为 1, 2 ? — 译者

⁴⁷原文下式等号右侧的 v 应为下指标，译文已改正。

¹³⁶ 数学上记作 $(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$ 。

¹³⁷ 在之前的二维旋转中遇到过类似情况，二维旋转操作可以用单位复数作用到向量对应的复数上来表示，在定义了复数与实数矩阵间的映射后，旋转也可以用实矩阵作用于列向量来表示。

¹³⁸ 它的含义是任意 2×2 厄米矩阵可以唯一地表示为基的线性组合： $a_0 \mathbf{1} + a_i \sigma_i$ 。

¹³⁹ 定义见(3.81)式。

(3.215)

¹⁴⁰ 这仅是基的选取问题，这里借助 Pauli 矩阵来选取具有我们之前导出的关于向量变换性质的基底。

上面解释过，我们也可以使用¹⁴⁰ $v_a^{\dot{b}} = v_\mu \sigma_{ac}^\mu \epsilon^{\dot{b}c}$ ，这样的话选用的基就是 $(\tilde{\sigma}_a^{\dot{b}})^\mu = \sigma_{ac}^\mu \epsilon^{\dot{b}c}$ 。⁴⁸一般的厄米矩阵表示为：

$$v_{ab} = \begin{pmatrix} v_0 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & v_0 - v_3 \end{pmatrix} \quad (3.216)$$

上节讲过，不同类的指标按不同方式变换，具体来说：带点的下标与不带点的下标变换方式不同。

下面研究 v_{ab} 在 Lorentz 变换下如何表现，利用不带点的下标的变换式(3.202)：

$$\begin{aligned} v \xrightarrow{\text{Lorentz 变换}} v' = v'_{ab} &= \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c v_{cd} \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}^*}{2} + \vec{\phi}\frac{\vec{\sigma}^*}{2}} \right)_b^{\dot{d}} \right)^T \\ &= \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c v_{cd} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2} + \vec{\phi}\frac{\vec{\sigma}^\dagger}{2}} \right)_b^{\dot{d}} \\ &\stackrel{\sigma^\dagger = \sigma}{=} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c v_{cd} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_b^{\dot{d}} \end{aligned} \quad (3.217)$$

¹⁴¹ 反厄米矩阵经 Lorentz 变换仍为反厄米矩阵的证明过程同理。只要把下面最后一步 $v_{cd}^\dagger = v_{cd}$ 改成 $v_{cd}^\dagger = -v_{cd}$ 。

前面说过，厄米矩阵经 Lorentz 变换后仍为厄米的，下面来验证¹⁴¹：

$$\begin{aligned} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c v_{cd} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_b^{\dot{d}} &\rightarrow \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c v_{cd} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_b^{\dot{d}} \right)^\dagger \\ &\stackrel{(ABC)^\dagger = C^\dagger B^\dagger A^\dagger}{=} \left(\left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_b^{\dot{d}} \right)^\dagger v_{cd}^\dagger \left(\left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c \right)^\dagger \\ &= \left(e^{i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2} + \vec{\phi}\frac{\vec{\sigma}^\dagger}{2}} \right)_b^{\dot{d}} v_{cd}^\dagger \left(e^{-i\vec{\theta}\frac{\vec{\sigma}^\dagger}{2} + \vec{\phi}\frac{\vec{\sigma}^\dagger}{2}} \right)_a^c \\ &\stackrel{\text{若 } v_{cd}^\dagger = v_{cd}}{=} \left(e^{i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_b^{\dot{d}} v_{cd} \left(e^{-i\vec{\theta}\frac{\vec{\sigma}}{2} + \vec{\phi}\frac{\vec{\sigma}}{2}} \right)_a^c \quad \checkmark \end{aligned} \quad (3.218)$$

¹⁴² 一般过程的推导可见 Matthew Robinson. *Symmetry and the Standard Model*. Springer, 1st edition, August 2011. ISBN 978-1-4419-8267-4, 128 页。

¹⁴³ 沿 z 轴意味着 $\vec{\phi} = (0, 0, \phi)^T$ 。沿 z 轴的推动变换形式最简单，因为只有 σ_3 是对角矩阵，其他方向的推动变换的 e 指数计算较复杂。

一般形式变换的计算过程冗长无聊，这里只推导一个特例¹⁴²，考虑将 v 沿 z 轴做推动变换¹⁴³：

⁴⁸原文该式有误，译文已按勘误表修改。

$$\begin{aligned}
v_{ab} \xrightarrow{\text{沿} z \text{轴推动}} v'^{ab} &= \left(e^{\phi \frac{\sigma_3}{2}}\right)_a^c \left(e^{\phi \frac{\sigma_3}{2}}\right)_b^d v_{cd} \\
&= \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix} \begin{pmatrix} v_0 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & v_0 - v_3 \end{pmatrix} \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix} \\
&= \begin{pmatrix} e^{\phi}(v_0 + v_3) & v_1 - iv_2 \\ v_1 + iv_2 & e^{-\phi}(v_0 - v_3) \end{pmatrix} \quad (3.219)
\end{aligned}$$

其中利用了 σ_3 是对角矩阵¹⁴⁴以及对角矩阵 e 指数的计算公式： $\begin{pmatrix} e^{A_{11}} & 0 \\ 0 & e^{A_{22}} \end{pmatrix}$ 。将变换后的 v' 用一般形式写出，并与(3.219)式进行比较：

$$^{144} \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$v'_{ab} = \begin{pmatrix} v'_0 + v'_3 & v'_1 - iv'_2 \\ v'_1 + iv'_2 & v'_0 - v'_3 \end{pmatrix} = \begin{pmatrix} e^{\phi}(v_0 + v_3) & v_1 - iv_2 \\ v_1 + iv_2 & e^{-\phi}(v_0 - v_3) \end{pmatrix}$$

比较可得对象 v 的分量在沿 z 轴的推动变换下的形式为¹⁴⁵：

$$\begin{aligned}
\rightarrow v'_0 + v'_3 &= e^{\phi}(v_0 + v_3) = (\cosh \phi + \sinh \phi)(v_0 + v_3) \\
\rightarrow v'_0 - v'_3 &= e^{-\phi}(v_0 - v_3) = (\cosh \phi - \sinh \phi)(v_0 - v_3)
\end{aligned}$$

¹⁴⁵ 下式利用了双曲函数的定义，即由 $\cosh \phi \equiv \frac{1}{2}(e^{\phi} + e^{-\phi})$, $\sinh \phi \equiv \frac{1}{2}(e^{\phi} - e^{-\phi})$ 得出 $e^{-\phi} = (\cosh \phi - \sinh \phi)$, $e^{\phi} = (\cosh \phi + \sinh \phi)$ 。当然不这么做也可以，不过用双曲函数使得形式更加紧凑。

解得：

$$\begin{aligned}
\rightarrow v'_0 &= \cosh(\phi)v_0 + \sinh(\phi)v_3 \\
\rightarrow v'_3 &= \sinh(\phi)v_0 + \cosh(\phi)v_3 \quad (3.220)
\end{aligned}$$

上面正是四维向量的 Lorentz 变换式¹⁴⁶：

¹⁴⁶ 见(3.146)式。

$$\begin{aligned}
\begin{pmatrix} v'_0 \\ v'_1 \\ v'_2 \\ v'_3 \end{pmatrix} &= \begin{pmatrix} \cosh \phi & 0 & 0 & \sinh \phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh \phi & 0 & 0 & \cosh \phi \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix} \\
&= \begin{pmatrix} \cosh(\phi)v_0 + \sinh(\phi)v_3 \\ v_1 \\ v_2 \\ \sinh(\phi)v_0 + \cosh(\phi)v_3 \end{pmatrix} \quad (3.221)
\end{aligned}$$

上述结果对任意 Lorentz 变换都成立，其他情形（如转动）由读者自行计算验证。由此可见 $(\frac{1}{2}, \frac{1}{2})$ 表示正是 4 向量表示。用向量形

式可以简化计算，因为用变换矩阵乘以向量比三个矩阵相乘要简单。人们熟悉的 4 向量其实与更基本的旋量有关，一个 4 向量是二阶旋量，即具有两个分量，根据 Lorentz 群的 $(\frac{1}{2}, \frac{1}{2})$ 表示变换的量。此外，4 向量并非描述最基本物理体系的量，它们不够基本，有一些系统它们无法描述。

¹⁴⁷ 不妨把二阶张量简单看成一个二阶矩阵 $M_{\mu\nu}$

有人说“旋量是向量的平方根”，这话与“向量是二阶张量¹⁴⁷的平方根”相似，一个二阶张量有两个向量指标，一个向量有两个旋量指标。因此 Lorentz 变换作用的最基本的对象其实是旋量。

前面讲过 Lorentz 群有四部分。不同部分之间通过宇称算符与时间反演算符联系¹⁴⁸，因此为了描述所有保证光速不变的变换，还需要考虑宇称和时间反演算符在各个表示下的形式。本书只考虑宇称变换，因为实验表明自然在宇称变换下并不总是对称的（之后的章节讨论），而且下一节有关宇称变换表示的推导也可应用于时间反演变换。

¹⁴⁸ 见(3.134)式。

3.7.9 Spinors and Parity 旋量和宇称

上面没有说明为啥 $(\frac{1}{2}, 0)$ 表示变换的对象叫左手旋量，而 $(0, \frac{1}{2})$ 表示变换的对象叫右手旋量。在简要介绍宇称变换之后这一问题就可以解释了。

前面已经导出了 Lorentz 群的生成元在宇称变换下的表现，见(3.153)式，便利起见重写一遍：

$$J_i \xrightarrow{P} J_i, \quad K_i \xrightarrow{P} -K_i \quad (3.222)$$

生成元 N_i^\pm 的定义式(3.161)式，在此也重写一遍：

$$N_i^\pm = \frac{1}{2}(J_i \pm iK_i) \quad (3.223)$$

由此可见在宇称变换之下 $N^+ \leftrightarrow N^-$ 。因此 $(0, \frac{1}{2})$ 表示下的 Lorentz 变换经宇称变换之后变成 $(\frac{1}{2}, 0)$ 表示下的变换，反之亦然。这就是左手、右手旋量名字的来历¹⁴⁹。就像右手坐标系在宇称变换下变为左手系那样，Lorentz 群的这两个表示也在宇称变换下来回变。

旋转变换在两个表示之下相同，而推动变换差一个负号，上述宇称变换的效果可总结为：

$$(\Lambda_{\vec{K}})_{(\frac{1}{2}, 0)} = e^{\vec{\phi} \vec{K}} \xrightarrow{P} e^{-\vec{\phi} \vec{K}} = (\Lambda_{\vec{K}})_{(0, \frac{1}{2})} \quad (3.224)$$

¹⁴⁹ 常见的名字是 left- / right-handed 旋量，但这个名字会引起混乱，因为 left-handed、right-handed 与螺旋性 (helicity) 的概念有关，螺旋性与手征性 (chirality) 是不同的东西。尽管如此这个名字还是反映了在宇称变换下某个“左”的东西变成了“右”的。

$$(\Lambda_{\vec{K}})_{(0, \frac{1}{2})} = e^{-\vec{\phi}\vec{K}} \underbrace{\rightarrow}_{\text{P}} e^{\vec{\phi}\vec{K}} = (\Lambda_{\vec{K}})_{(\frac{1}{2}, 0)} \quad (3.225)$$

可见，描述一个在宇称变换下不变的物理体系同时需要左手旋量和右手旋量。最简单的做法是把它们上下放置在一个对象中，该对象称为 **Dirac 旋量**：

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_a \\ \xi^{\dot{a}} \end{pmatrix} \quad (3.226)$$

前面提过左手、右手旋量统称 **Weyl 旋量**，我们称一个 Dirac 旋量 Ψ 包含两个 Weyl 旋量 χ_L 和 ξ_R 。注意 Ψ 是一般形式，即 χ_L 与 ξ_R 之间并未钦点任何关系。而具有如下形式的 Dirac 旋量

$$\Psi_M = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \quad (3.227)$$

就是一个特例，称为 Majorana 旋量。Dirac 旋量或 Majorana 旋量不是四维向量，它们的变换方式截然不同。Dirac 旋量按 Lorentz 群的 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示¹⁵⁰变换，把左右旋量分别对应的变换矩阵写成一个分块对角的大矩阵就是新的变换矩阵：

$$\Psi \xrightarrow{\text{Lorentz 变换}} \Psi' = \Lambda_{(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})} \Psi = \begin{pmatrix} \Lambda_{(\frac{1}{2}, 0)} & 0 \\ 0 & \Lambda_{(0, \frac{1}{2})} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (3.228)$$

例如， Ψ 的推动变换为：

$$\Psi \xrightarrow{\text{推动变换}} \Psi' = \begin{pmatrix} e^{\vec{\phi}\vec{\sigma}_2} & 0 \\ 0 & e^{-\vec{\phi}\vec{\sigma}_2} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (3.229)$$

我们研究 Dirac 旋量在宇称变换下的行为，是因为可以从它出发检验某个理论在宇称变换下是否不变。我们知道宇称变换下 $N^+ \leftrightarrow N^-$ ，因此 Dirac 旋量经宇称变换后不再是 Dirac 旋量（即那个根据 Lorentz 群的 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示变换的对象），再考虑到：

$$\left(0, \frac{1}{2}\right) \underbrace{\leftrightarrow}_{\text{P}} \left(\frac{1}{2}, 0\right) \quad (3.230)$$

¹⁵⁰ $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示是一个可约表示，这从它的变换矩阵是分块对角矩阵可以看出。而四维向量按 $(\frac{1}{2}, \frac{1}{2}) = (\frac{1}{2}, 0) \otimes (0, \frac{1}{2})$ 表示变换。

可以断定, 因为 Dirac 旋量按 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示变换, 故变换后的对象按照 $(0, \frac{1}{2}) \oplus (\frac{1}{2}, 0)$ 表示变换。

$$\Psi^P \xrightarrow{\text{Lorentz 变换}} (\Psi^P)' = \Lambda_{(0, \frac{1}{2}) \oplus (\frac{1}{2}, 0)} \Psi^P = \begin{pmatrix} \Lambda_{(0, \frac{1}{2})} & 0 \\ 0 & \Lambda_{(\frac{1}{2}, 0)} \end{pmatrix} \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix} \quad (3.231)$$

因此

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \xrightarrow{\text{宇称变换}} \Psi^P = \begin{pmatrix} \xi_R \\ \chi_L \end{pmatrix} \quad (3.232)$$

宇称变换将 Dirac 旋量 (包含 ξ_R, χ_L) 变换为另一个同样具有 ξ_R, χ_L 的对象, 只是上下位置变了。宇称变换的效果并非 $\xi_L \rightarrow \xi_R$, 具有这一效果的变换之后再讲述。

3.7.10 Spinors and Charge Conjugation 旋量与电荷共轭

3.7.7 节遇到了能使左手、右手旋量来回变换的变换 (好绕), $\chi_L \rightarrow \chi_L^C = \epsilon \chi_L^* = \chi_R$, $\xi_R \rightarrow \xi_R^C = (-\epsilon) \xi_R^* = \xi_L$ 。此变换不属于 Lorentz 群, 下面从另一角度来理解这一变换。

到目前为止这个变换仅仅是升降旋量指标的技巧, 现在考虑 Dirac 旋量在该变换下的表现, 最简单的想可以写成:

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \rightarrow \tilde{\Psi} = \begin{pmatrix} \chi_L^C \\ \xi_R^C \end{pmatrix} = \begin{pmatrix} \chi_R \\ \xi_L \end{pmatrix} \quad (3.233)$$

¹⁵¹ 与之前的宇称变换不同, 这里可以选择, 我们更喜欢新对象与原来的对象种类相同。 $\tilde{\Psi}$ 可以视为经宇称变换和电荷共轭变换后的 Dirac 旋量。

不幸的是新生儿 $\tilde{\Psi}$ 的变换方式与 Dirac 旋量不同¹⁵¹, Dirac 旋量在推动变换下的表现是:

$$\Psi \rightarrow \Psi' = \begin{pmatrix} e^{\frac{\vec{\theta} \cdot \vec{\sigma}}{2}} & 0 \\ 0 & e^{-\frac{\vec{\theta} \cdot \vec{\sigma}}{2}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (3.234)$$

$\tilde{\Psi}$ 按下式变换:

$$\tilde{\Psi} \rightarrow \tilde{\Psi}' = \begin{pmatrix} e^{-\frac{\vec{\theta} \cdot \vec{\sigma}}{2}} & 0 \\ 0 & e^{\frac{\vec{\theta} \cdot \vec{\sigma}}{2}} \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (3.235)$$

它是一种新对象, 因为它按照 Lorentz 群的不同表示进行变换。通

过这些认识，我们可以令：

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \rightarrow \Psi^C = \begin{pmatrix} \xi_R^C \\ \chi_L^C \end{pmatrix} = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix} \quad (3.236)$$

这样的变换就能满足 Dirac 旋量的 Lorentz 变换了。这个新变换通常称为**荷共轭 (charge conjugation)** 变换，这个名字有点误导性。由上式可知该变换将左手旋量变为右手旋量，也就是把描述基本粒子的一个标签翻转一下¹⁵²。之后会看到该算符不止能翻转这一个标签，而是对基本粒子的所有标签都有效。其中一个标签是电荷，因此该变换称为电荷共轭，在推导如何翻转电荷之前，先要学习电荷到底是啥。有一点是很重要的：不止电荷能翻，别的标签都可以。

我们可以接着推导 Lorentz 群的更高维表示，但到现在为止所有物理中常用的有限维不可约表示都已导出。物理上还要用到无限维表示，用它来变换物理场。

¹⁵² 选读：注意 Weyl 旋量实际上是个二分对象，稍后我们会定义称为‘自旋’的可观测量，它由 $\frac{1}{2}\sigma_3$ 描述。矩阵 ϵ 可以将自旋算符 $\frac{1}{2}\sigma_3$ 本征值 $+\frac{1}{2}$ 的本征向量变换为本征值 $-\frac{1}{2}$ 对应的本征向量。这称为自旋反转，即一个自旋 $\frac{1}{2}$ 的对象变换为自旋 $-\frac{1}{2}$ 的对象。

3.7.11 Infinite-Dimensional Representations 无限维表示

上面几节讨论了 Lorentz 群的有限维表示及其分类，这些有限维表示下的群元作用于 1-、2-或 4-分量的对象。物理中经常处理在时空中动力学演化的对象，下面考虑它们如何变换。前面处理的 Lorentz 变换可以写成如下形式：

$$\Phi_a \rightarrow \Phi'_a = M_{ab}(\Lambda)\Phi_b \quad (3.237)$$

其中 $M_{ab}(\Lambda)$ 是 Lorentz 变换 Λ 在某个有限维表示之下的矩阵形式。例如 $M_{ab}(\Lambda)$ 是个作用于 Weyl 旋量的 2×2 矩阵，作用的结果无非是新对象的分量为旧对象分量的线性组合。如果 Φ 分布在四维时空，即¹⁵³ $\Phi = \Phi(x)$ ，则 Lorentz 变换（除了变换 Φ 之外）还将变换时空坐标，一般情况下：

$$x^\mu \rightarrow \Lambda^\mu_\nu x^\nu \quad (3.238)$$

其中 Λ^μ_ν 是向量表示（即 $(\frac{1}{2}, \frac{1}{2})$ 表示）下的 Lorentz 变换。场的 Lorentz 变换为¹⁵⁴：

$$\Phi_a(x) \rightarrow M_{ab}(\Lambda)\Phi_b(\Lambda x) \quad (3.239)$$

¹⁵³ x 是时空坐标 t, x, y, z 的缩写。

¹⁵⁴ 大部分文献采用对称操作的 Wigner 惯例形式： $\Phi_a(x) \rightarrow M_{ab}(\Lambda)\Phi_b(\Lambda^{-1}x)$ ，但我们没有什么理由在这里这么写。

¹⁵⁵ Φ 的每个分量都是坐标 x 的函数。相应的算符作用在 $\Phi_a(x)$ 上, $\Phi_a(x)$ 是时空坐标的函数, 而在这种语境下的函数空间是无限维的。(原文为: ‘The corresponding operators act on $\Phi_a(x)$, i.e. functions of the coordinates and the space of functions is in this context infinite-dimensional.’) 函数空间之所以是无限维是由于基函数的个数是无穷多的。怎么把任意函数用无穷个基函数展开会在附录 D.1 Fourier 展开中讲述。

¹⁵⁶ 下式的符号 ∂^ν 是偏导数符号 $\frac{\partial}{\partial x_\nu}$ 的简写。

¹⁵⁷ $M^{\mu\nu}$ 在(3.165)式定义。分量 $\omega_{\mu\nu}$ 与通常的转动角度 θ_i 的关系为 $\theta_i = \frac{1}{2}\epsilon_{ijk}\omega_{jk}$, 与推动参数 ϕ_i 的关系为 $\phi_i = \omega_{0i}$ 。

上式的变换包含两部分。一部分变换作用在 Φ_a 上, 对应于有限维表示; 而第二部分变换作用于坐标。第二部分作用于一个无限维的¹⁵⁵向量空间, 因此需要用无限维表示。Lorentz 群的无限维表示由如下微分算符¹⁵⁶给出:

$$M_{\mu\nu}^{\text{inf}} = i(x^\mu \partial^\nu - x^\nu \partial^\mu) \quad (3.240)$$

读者可自行验证 $M_{\mu\nu}^{\text{inf}}$ 满足 Lorentz Lie 代数(3.167)式, 并且可以变换坐标。

这样坐标变换的形式为^{157, 49}:

$$\Phi(\Lambda x) = e^{-i\frac{\omega_{\mu\nu}}{2}M_{\mu\nu}^{\text{inf}}}\Phi(x) \quad (3.241)$$

其中 e 指数函数定义为它的级数展开, 前面已经讲过。完整的 Lorentz 变换是有限维表示 $M_{\mu\nu}^{\text{fin}}$ 与无限维表示 $M_{\mu\nu}^{\text{inf}}$ 变换的组合:

$$\Phi_a(x) \xrightarrow{\text{Lorentz变换}} \left(e^{-i\frac{\omega_{\mu\nu}}{2}M_{\mu\nu}^{\text{fin}}} \right)_a^b e^{-i\frac{\omega_{\mu\nu}}{2}M_{\mu\nu}^{\text{inf}}} \Phi_b(x) \quad (3.242)$$

矩阵 $M_{\mu\nu}^{\text{fin}}$ 是有限维的, 它可以并入无限维矩阵中:

$$\Phi_a(x) \xrightarrow{\text{Lorentz变换}} \left(e^{-i\frac{\omega_{\mu\nu}}{2}M_{\mu\nu}} \right)_a^b \Phi_b(x) \quad (3.243)$$

其中 $M_{\mu\nu} = M_{\mu\nu}^{\text{fin}} + M_{\mu\nu}^{\text{inf}}$ 。上式对应的 Lorentz 群表示称为场表示。

现在可以讨论平移变换, 即时空中某点的对象变换到另一点的变换。平移变换的结果并非各分量的组合, 因此不需要有限维表示描述平移, 而很容易地找到相应的无限维表示。这不再是 Lorentz 群的部分, 但是自然定律应该在时空各点都相同。Lorentz 群(推动 + 转动)加上平移变换称为 Poincaré 群, 下一节具体讲述。这里首先引入平移的无限维表示。简明起见, 考虑一维情形, 一个函数沿 x 轴的无穷小平移变换为:

$$\Phi(x) \rightarrow \Phi(x + \epsilon) = \Phi(x) + \underbrace{\partial_x \Phi(x)}_{\text{沿 } x \text{ 轴的 ‘变化率’}} \epsilon$$

⁴⁹ 其中, 下面的 $\omega^{\mu\nu}$ 是 Lorentz 变换的参数。

上式就是 Taylor 级数展开的前两项。按照惯例，相应的生成元额外乘以 $-i$ ：

$$P_i \equiv -i\partial_i \quad (3.244)$$

这样，任意的有限大小的平移变换为：

$$\Phi(x) \rightarrow \Phi(x+a) = e^{-ia^i P_i} \Phi(x) = e^{a^i \partial_i} \Phi(x)$$

其中 a^i 表示沿各个方向的平移数量。把 e 指数函数展开为 Taylor 级数¹⁵⁸就能看出上式就是 $\Phi(x+a)$ 的 Taylor 展开¹⁵⁹。平移到另一个时间点使用 $P_0 = i\partial_0$ ，平移到另一空间点使用 $P_i = -i\partial_i$ 。

¹⁵⁸ 这在附录 B.4.1 中推导。

¹⁵⁹ 这在附录 B.3 中推导。

3.8 The Poincaré group Poincaré 群

现在考虑自然界完整的时空对称群：Poincaré 群。Lorentz 群包含转动和推动变换，而保证光速不变的还有时空中的平移变换，因为在不同时空点测量的光速是相同的。等价地说，光速与所用坐标系原点的位置无关。将这一对称性加入 Lorentz 群之后就得到 Poincaré 群¹⁶⁰：

$$\begin{aligned} \text{Poincaré 群} &= \text{Lorentz 群} + \text{平移} \\ &= \text{旋转} + \text{推动} + \text{平移} \end{aligned} \quad (3.245)$$

¹⁶⁰ Poincaré 群并非 Lorentz 群与平移变换的直和 (direct sum)，而是半直和 (semi-direct sum)，但现在可以忽略此数学细节。

Poincaré 群的生成元是 Lorentz 群的生成元 J_i, K_i 加上平移变换的在 Minkowski 空间的生成元 P_μ 。

用 J_i, K_i, P_μ 表示的 Lie 代数为¹⁶¹：

$$[J_i, J_j] = i\epsilon_{ijk} J_k \quad (3.246)$$

$$[J_i, K_j] = i\epsilon_{ijk} K_k \quad (3.247)$$

$$[K_i, K_j] = -i\epsilon_{ijk} J_k \quad (3.248)$$

$$[J_i, P_j] = i\epsilon_{ijk} P_k \quad (3.249)$$

$$[J_i, P_0] = 0 \quad (3.250)$$

$$[K_i, P_j] = i\delta_{ij} P_0 \quad (3.251)$$

$$[K_i, P_0] = -iP_i \quad (3.252)$$

¹⁶¹ 这些式子不怎么好看... 这里列出只是为了完整。

上面的形式比较凌乱，定义 $M_{\mu\nu}$ 来表述会整洁的多， $M_{\mu\nu}$ 的定义

为:

$$J_i = \frac{1}{2} \epsilon_{ijk} M_{jk} \quad (3.253)$$

$$K_i = M_{0i} \quad (3.254)$$

用 $M_{\mu\nu}$ 表示的 Poincaré Lie 代数为:

$$[P_\mu, P_\nu] = 0 \quad (3.255)$$

$$[M_{\mu\nu}, P_\rho] = i(\eta_{\mu\rho} P_\nu - \eta_{\nu\rho} P_\mu) \quad (3.256)$$

由定义式可得

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\mu\rho} M_{\nu\sigma} - \eta_{\mu\sigma} M_{\nu\rho} - \eta_{\nu\rho} M_{\mu\sigma} + \eta_{\nu\sigma} M_{\mu\rho}) \quad (3.257)$$

¹⁶² 复习: Casimir 算符是从生成元构造的、与所有生成元都对易的算符。

Poincaré 群是相当复杂的, 有必要将它的表示用 Casimir 算符固有的标量值加以标记。Poincaré 群有两个 Casimir 算符¹⁶², 第一个是:

$$P_\mu P^\mu =: m^2 \quad (3.258)$$

¹⁶³ 别担心, 后面会详谈。

这个标量值记作 m^2 , 它的含义你懂的。稍后我们就看到它就是粒子质量¹⁶³。

¹⁶⁴ $\epsilon^{\mu\nu\rho\sigma}$ 是四维 Levi-Civita 符号, 在附录 B.5.5 中定义。

第二个 Casimir 算符为 $W_\mu W^\mu$, 其中¹⁶⁴:

$$W^\mu = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} P_\nu M_{\rho\sigma} \quad (3.259)$$

W^μ 称为 **Pauli-Lubanski 四维向量**。经过超级繁琐的计算之后可以证明。相应的标量值记作 $j \equiv j_1 + j_2$, 它通常称为自旋。现在“自旋”只是个名字, 稍后会看到它有什么含义。对于 Lorentz 群, 它的两个¹⁶⁵ $SU(2)$ Lie 代数每个都对应一个 j_i 。

例如, $(j_1, j_2) = (0, 0)$ 表示称为自旋 0 表示¹⁶⁶, $(j_1, j_2) = (\frac{1}{2}, 0)$ 与 $(j_1, j_2) = (0, \frac{1}{2})$ 都称为自旋 $\frac{1}{2}$ 表示¹⁶⁷, 然后 $(j_1, j_2) = (\frac{1}{2}, \frac{1}{2})$ 表示称为自旋 1 表示¹⁶⁸。

综上, Poincaré 群的表示用两个标量来标记: m, j 。其中 m 可以取任意值, j 只能取整数或半整数。

¹⁶⁵ 复习: Lorentz 群的 Lie 代数可以视为两份 $SU(2)$ Lie 代数 $SU(2)$ 群的表示可以用一个数 j 标记, 于是 Lorentz 群双覆盖的表示可以用两个数 (j_1, j_2) 标记。

¹⁶⁶ $j_1 + j_2 = 0 + 0 = 0$ 。

¹⁶⁷ $j_1 + j_2 = \frac{1}{2} + 0 = 0 + \frac{1}{2} = \frac{1}{2}$ 。

¹⁶⁸ $j_1 + j_2 = \frac{1}{2} + \frac{1}{2} = 1$ 。

3.9 Elementary Particles 基本粒子

Poincaré 群不可约表示的标记就是物理学对基本粒子的标记¹⁶⁹：质量 m 与自旋（这里就是 j ）。对给定质量 m 与自旋（例如 $j = \frac{1}{2}$ ）的基本粒子的描述，需要按 Poincaré 群的 m , 自旋 $\frac{1}{2}$ 表示变换。

粒子的更多特征量（例如电荷）将从内禀对称性导出，这些标签用来定义一种基本粒子。例如，电子的定义为：

- 质量： $9.109 \times 10^{-31}\text{kg}$
- 自旋： $\frac{1}{2}$
- 电子电荷： $1.602 \times 10^{-19}\text{C}$
- 弱荷，称为弱同位旋： $-\frac{1}{2}$
- 强荷，称为色荷：0

这些标签决定了一种基本粒子在实验中的观测值。本章导出的各种表示使我们能在数学上描述它们。基本粒子可分为如下几种¹⁷⁰：

- 自旋 0，由标量（记作 Φ ）描述，标量按 Poincaré 群的 $(0, 0)$ 表示（称为自旋 0 表示或标量表示）变换。例如 Higgs 粒子由一个标量场描述。
- 自旋 $\frac{1}{2}$ ，由旋量（记作 Ψ ）描述。旋量按 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示（称为自旋 $\frac{1}{2}$ 表示或旋量表示）变换。例如电子和夸克由旋量描述。
- 自旋 1，由向量（记作 A ）描述。向量按照 $(\frac{1}{2}, \frac{1}{2})$ 表示（称为自旋 1 表示或向量表示）变换。例如光子由向量描述。

上面的结论是物理学重要、深刻而优美的创见，因此再重复一遍：

Poincaré 群的不可约表示是描述所有基本粒子的数学工具。描述标量粒子（例如 Higgs Boson）的数学对象称为标量，它根据自旋 0 表示变换。描述自旋 $\frac{1}{2}$ 粒子（例如电子、中微子、夸克等等）的数学对象称为旋量，它根据自旋 $\frac{1}{2}$ 表示变换。描述光子或其它自旋 1 的粒子所用的对象称为向量，它根据自旋 1 表示变换。

关于自旋这个充满暗示的名字的详细解释在讨论完实验观测之后的 8.5.5 节讲述。想要弄清楚自旋的含义，我们必须要知道如何通过实验观测研究对象是否在“自旋”。现在把自旋仅仅当做一个名词就好。

¹⁶⁹ 其他说法：基本粒子是 Poincaré 群的不可约表示。

¹⁷⁰ 第一章导论介绍不能推导的内容说过，没有理由只有三个表示对应基本粒子。我们可以继续考虑更高维的表示，但没发现对应的粒子，比如自旋 $\frac{3}{2}$ 的粒子。然而这些表示可以描述复合对象。此外，许多物理学家认为传递引力作用的基本粒子——引力子的自旋为 2，因此必须用相应的高维表示来描述它。

Further Reading Tips 阅读建议

¹⁷¹ John Stillwell. *Naive Lie Theory*. Springer, 1st edition, 8 2008b. ISBN 9780387782140

¹⁷² Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhaeuser, 1st edition, August 2011. ISBN 978-0817647148

- John Stillwell - 你们啊 Naive Lie Theory¹⁷¹ 是一本非常容易读的、介绍 Lie 理论的数学教材。
- N. Jeevanjee - An Introduction to Tensors and Group Theory for Physicists¹⁷² 是一本非常出色的导论教材，重点关注群论在物理学中的应用。

3.10 Appendix: Rotations in a Complex Vector Space 附录：复向量空间中的旋转

在复向量空间也可定义保证内积不变的变换。我们希望一个向量与其自身的内积为实数，这样才能让它等于向量的模的平方，模长为复数没有物理意义。因此复向量内积的定义（除了转置以外还）附加了取复共轭的操作¹⁷³：

$$a \cdot a = (a^T)^* a = a^\dagger a \quad (3.260)$$

符号 \dagger 称为 dagger，表示厄米共轭操作，即取复共轭 + 转置。保证内积不变的变换 \mathcal{U} 必须满足 $\mathcal{U}^\dagger \mathcal{U} = 1$ ：

$$(\mathcal{U}a)^\dagger \cdot (\mathcal{U}a) = a^\dagger \mathcal{U}^\dagger \mathcal{U} a = a^\dagger a \quad (3.261)$$

这种形式的变换记作 $\mathcal{U}(n)$ ，其中 n 表示复向量空间的维数，“ \mathcal{U} ”表示么正性。复习一下， $SU(n)$ 群称为特殊 (special) 么正群，因为它的群元还满足 $\det(\mathcal{U}) = 1$ 。

3.11 Appendix: Manifolds 附录：流形

流形 M 是满足如下条件的点的集合：对于 M 的任意开邻域，存在将开邻域映射到 \mathbb{R}^n 开子集的一一映射。简单地说，该定义意味着流形 M 局部地与 \mathbb{R}^n 类似。 M 的开邻域到 \mathbb{R}^n 的映射将 M 中的点 P 对应到一个 n 元数组 $(x_1(P), \dots, x_n(P))$ ，其中 $x_1(P), \dots, x_n(P)$ 称为 P 点的坐标。因此 n 维流形可以看做这样的点集：任意点的邻域都可对应 n 个独立坐标。

流形的典型例子是球面，三维球体的表面称为 2-球面 (two-sphere)，记作 S^2 ，定义为 \mathbb{R}^3 当中满足 $x^2 + y^2 + z^2 = r^2$ 的点的集合，其中 r 为球半径。注意三维球体的表面是二维的，因为它的

¹⁷³ $z = a + ib, z^* = a - ib$ ，因此 $z^* z = (a + ib)(a - ib) = a^2 + b^2 \in \mathbb{R}$ 。

定义有三个坐标与一个约束条件，这样减少了一个自由度。这也是数学上称之为 2-球面的原因。为了表明球面是个流形，我们需要找出从球面到 \mathbb{R}^2 的映射，这个映射可以由通常的球坐标给出。

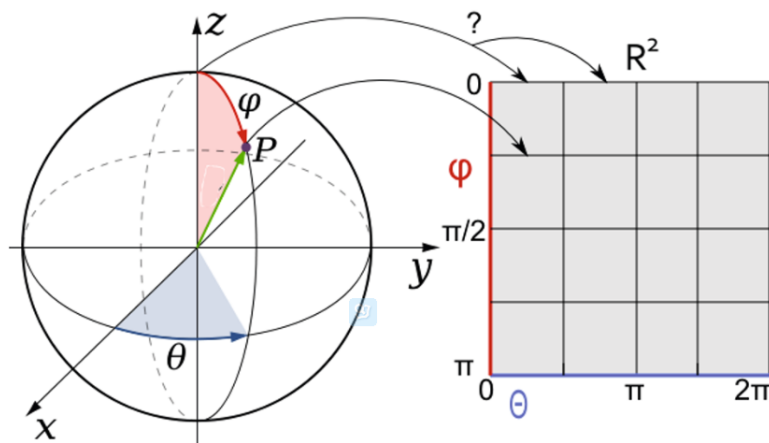


图 3.8: 从球面的一个邻域到 \mathbb{R}^2 映射的图示

球面上几乎所有的点都可明确地对应于坐标 (φ, θ) ，几乎所有... 但是极点 $\varphi = 0$ 映射到哪儿？不可能有一对一的点，因为极点对应的是一整条线，在上图已经画出。因此球坐标映射不能覆盖整个球面，还需要一个映射将极点及其邻域映射到 \mathbb{R}^2 。相似的问题还出现在 $\theta = 0$ 对应的半圆处， $\theta = 0$ 与 $\theta = 2\pi$ 映射到 \mathbb{R}^2 的同一个点，因此也不是一一映射。这说明对流形而言，一般不存在一个覆盖流形所有点的全局坐标系，而只有局域坐标系，即仅覆盖某一邻域的坐标系。不过这没关系，因为流形的特征是局部像 \mathbb{R}^n 。

上面说到，球坐标系映射只对开邻域 $0 < \varphi < \pi, 0 < \theta < 2\pi$ 有效，还需要另外的映射来覆盖整个球面。可以使用第二个球坐标系——坐标轴的指向不同，使得有问题的极点不再对应 $\varphi = 0$ ，在第二个映射补充下，球面上所有的点都可映射到 \mathbb{R}^2 ，于是 2-球面确实是个流形。

流形的平凡例子是 \mathbb{R}^n ，根据定义容易看出。

4. The Framework 理论框架

这一章的基本思路是，我们要通过使某些东西取极小值，得到正确的关于自然的方程。某些东西是什么？至少我们知道一件事：它不应该在 Lorentz 变换下改变，否则我们会在不同的参考系下得到不同的自然规律。在数学意义上，它意味着我们寻找的这个东西是个标量，依照 Lorentz 群的 $(0,0)$ 表示作变换。再加上到自然对简单规律的偏好，导出关于自然的方程绰绰有余。

从这个想法出发，我们将会引入拉格朗日形式 (Lagrangian formalism)。通过使理论的中心对象取极小值，可以得到用以描述问题中的物理系统的运动方程。极小化过程的结果被称作欧拉-拉格朗日方程 (Euler-Lagrange equations)。

通过拉格朗日形式，可以得到物理中最重要的定理之一：Noether 定理 (Noether's Theorem)。这个定理揭示了对称性和守恒量¹之间的深刻联系。我们将在下一章中利用它来理解，理论是如何来描述实验测量量的。

4.1 Lagrangian Formalism 拉格朗日形式

拉格朗日形式是在基础物理中被广泛运用的一个强有力的框架²。由于理论的基本对象——拉格朗日量 (Lagrangian) 是一个标量³，它相对简单。如果你希望从对称性的观点考虑问题，这种形式将会是非常有用的。要求拉格朗日量的积分，作用量 (action)，

¹ 守恒量指的是不随时间变化的物理量。例如一个给定体系的能量或动量。数学上意味着 $dQ/dt = 0 \rightarrow Q = \text{常数}$ 。

² 物理中当然有其他框架，例如以哈密顿量 (Hamiltonian) 为中心对象的哈密顿形式 (Hamiltonian formalism)。哈密顿量的问题在于它不是洛伦兹不变的，因为它所代表的能量，仅仅是协变能动矢量 (covariant energy-impulse vector) 的一个分量

³ 标量指依照洛伦兹群的 $(0,0)$ 表示作变换的对象。这意味着它不在洛伦兹变换下改变

在某些对称变换下不变，即保证了体系的动力学遵从该对称性。

4.1.1 Fermat 原理

自然界中不论发生什么过程，由此造成作用量的改变总尽可能的小。

⁴Recherche des lois du mouvement (1746)

- Pierre de Maupertius ⁴

拉格朗日形式的思想源于 Fermat 原理：光在两空间点间传播总依耗时最短的路径 $q(t)$ 而行。数学上来讲，如果我们定义给定路径 $q(t)$ 的作用量为

$$S_{\text{light}}[\mathbf{q}(t)] = \int dt$$

⁵ 此处的作用量仅仅是沿给定路径对时间的积分，但一般而言作用量会更加复杂，我们待会儿就能见到

⁶ 一般而言，我们希望找到极值 (extremums)，即极小值和极大值。下一节中的方法足以找到这两者。无论如何，我们将继续谈论极小值

而我们的任务便是找到一条特定的路径 $q(t)$ 使作用量取极小值⁵为了得到一个给定函数的极小值⁶，我们可以求得其导函数并令其为零；而为了找到泛函 $S[\mathbf{q}(t)]$ ¹——函数 $\mathbf{q}(t)$ 的函数 S ——的极小值，就得要一个新的数学工具：变分法。

4.1.2 变分法：基本思想

在思考如何发展一套能够找到泛函极值的新理论之前，我们需要倒回去想想什么给出一个数学上的极小点。变分法给出的答案是，极小点由极小点邻域的性质决定。例如，让我们尝试寻找一个寻常函数 $f(x) = 3x^2 + x$ 的极小点 x_{\min} 。我们从一个特定点 $x = a$ 出发，仔细考察其邻域。数学上它意味着 $a + \epsilon$ ，其中 ϵ 代表无穷小量（可正可负）。我们将 a 的变分代入函数 $f(x)$ ：

$$f(a + \epsilon) = 3(a + \epsilon)^2 + (a + \epsilon) = 3(a^2 + 2a\epsilon + \epsilon^2) + a + \epsilon。$$

如果 a 是极小点， ϵ 的一阶变分必须为零，否则我们可以取 ϵ 为负 $\epsilon < 0$ ，这样 $f(a + \epsilon)$ 就会比 $f(a)$ 更小²。因此，我们将线性依赖于 ϵ 的项取出并令其为零。

$$3 \cdot 2a\epsilon + \epsilon \stackrel{!}{=} 0 \rightarrow 6a + 1 \stackrel{!}{=} 0。$$

¹译者注：这里的 $q(t)$ 原文一会儿用粗体，一会不用，不知道想表达什么，此处与原文保持一致。

²译注：此处讨论有误。可改作“否则若 ϵ 的一阶变分为正，我们可以...，反之亦然。”

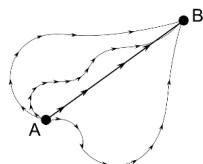


图 4.1: 对于给定初末端点的路径的扰动

由此我们找到极小点

$$x_{\min} = a = -\frac{1}{6},$$

它自然和我们求导 $f(x) = 3x^2 + x \rightarrow f'(x) = 6x + 1$ 并令其为零的办法得到的结果一致。对于寻常函数而言，这只是一个用来干同一件事不同方法而已³，但是变分法却能找到泛函的极值点。我们马上就能看到，应当如何处理一个一般的作用量泛函。

拉格朗日形式的中心思想在于，对于有质量的物体也存在一个与光的 Fermat 原理相类似的原理。当然，它不可能直接遵从费马原理，但是我们可以从一个更一般的形式出发

$$S[q(t)] = \int \mathcal{L} dt$$

其中 \mathcal{L} 一般是一个非常数的参量，称为拉格朗日量。对于光而言，这个参量是个常数。一般的，拉格朗日量依赖于物体的坐标和速度 $\mathcal{L} = \mathcal{L}(q(t), \frac{\partial}{\partial t} q(t))$ 。下一节中我们将仔细讨论这件事⁷。在仔细讨论如何对这样一个泛函使用变分法之前，我们需要先讲讲两个小问题。

4.2 Restrictions 限制

正如1.1节所提到的，现有理论中有些限制条件无法从第一性原理中得到。我们所能知道是，如果想得到一个有意义的理论，我们必须加上这些限制。

一个重要的限制是，我们仅允许拉格朗日量中出现尽可能低阶的非平凡⁴导数。这里平凡指的是对于系统的动力学，即运动方程，没有影响。某些理论会包含一阶导数，另外一些带二阶导数。一个给定理论所含的最低阶导数由该条件决定：拉格朗日量在 Lorentz 变换下不变⁸，否则我们将在不同的参考系下导出不同的运动方程⁵。对于某些理论，我们无法写出一个仅含一阶导数项的不变量，那么此时二阶导数便成了最低阶导数项。

我们并不清楚如何处理包含高阶导数的理论，它们有一些深刻的、系统性的困难⁹。另外，拉格朗日量中的高阶导数会使得运

³译注：对于受微元法荼毒的物竞生而言，求导才是用来干这件事的不同方法。

⁴译注：原文为 ‘non-trivial’。Trivial 这个词常包含简单、弱智、无意义之义。

⁵译注：此处讨论疑有误，作者此前并没有说明积分中的时间是坐标时还是固有时

⁷ 我们的任务是找到对于给定拉格朗日量和初始条件有着最小作用量的路径 $q(t)$ 。在此之前，我们得先找到正确的拉格朗日量，用以描述问题中的物理系统。这是我们在上一章中所讨论的对称性所能发挥作用的地方。通过要求拉格朗日量在洛伦兹群的所有变换下不变，我们就能找到正确的拉格朗日量

⁸ 实际上，作用量才应该是 Lorentz 不变的。但如果拉格朗日量满足这个条件，那作用量自然也是。

⁹ 这些困难称为 Ostrogradski 不稳定性，即包含高阶导数的系统的能量没有下界，以至于系统中的所有态总会自发衰变到能量更低的态上去。这类系统中找不到稳定的状态。

动方程中也包含高阶导数，这样就必须知道更多的初始条件才能决定物体的运动。

有些人会宣称对理论所包含导数阶数的限制源于我们希望得到一个局域¹⁰理论，但这个要求仅会排除掉那些无穷阶导数。一个非局域相互作用有如下形式¹¹

$$\Phi(x-h)\Phi(x) \quad (4.1)$$

即，时空中距离为 h 的两个点的场相互作用。利用 Taylor 展开我们有⁶

$$\Phi(x-h) = \sum_{k=0}^{\infty} \left(\left(\frac{\partial}{\partial x} \right)^k \Phi(x) \Big|_x \right) \frac{(-h)^k}{k!} \quad (4.2)$$

即包含无穷阶导数会导致一个非局域的理论。

另外一个限制是，要得到一个自由（无相互作用）场/粒子的理论，我们只能写到二次项。这意味着我们只会考虑¹²

$$\Phi^0, \Phi^1, \Phi^2$$

这样的项。例如，形如 $\Phi^2 \partial_\mu \Phi$ 的项是 Φ 的三阶项，因而不会被包含在我们自由理论的拉格朗日量中。

4.3 Particle Theories vs. Field Theories 粒子理论与场论

我们目前有两套用以描述自然的理论框架。一套是粒子理论，用依赖于时间的粒子位置描述物理系统，即 $\vec{q} = \vec{q}(t)$ 。由于我们不是非得在使用笛卡尔坐标系¹³，所以用的是字母 q 而不是 x 。对于这样的理论，拉格朗日量依赖于坐标 \vec{q} ，速度 $\partial_t \vec{q}$ 和时间 t ：⁷

$$\mathcal{L} = \mathcal{L}(\vec{q}, \partial_t \vec{q}, t) \quad (4.3)$$

一个著名的拉格朗日量是 $\mathcal{L} = \frac{1}{2} m \dot{\vec{q}}^2$ ，它能导出经典力学中的 Newton 运动方程。在稍后将进行相当详细的讨论。

另一套是场论：使用场而不是独立粒子的坐标来描述自然¹⁴。在这套理论中，时空构成了场 $\Phi(\vec{x}, t)$ 表演的舞台。利用前面提到

¹⁰ 局域性是狭义相对论的基本假定，2.4节已作阐述

¹¹ 我们将会讨论粒子的拉格朗日理论：寻求粒子的路径；也会讨论场的拉格朗日理论：寻求场函数 $\Phi(x)$ 。这将是下一节的主题。

¹² 从另外一个角度来看，这再次表示我们只引入最低阶的非平凡项。我们随后就会看到，含 Φ^0 和 Φ^1 的项是平凡的，因此我们这次用的是 Φ 的最低阶非平凡项

¹³ 例如，我们可以用球坐标系

¹⁴ 量子场论有一个特别优美的性质是关于如何引进粒子的。我们将在第6章中看到场有能力产生和湮灭粒子

⁶译注：此式有误，已更正

⁷译注：严格说来，依赖于坐标函数 $\vec{q}(t)$ ，速度函数 $\partial_t \vec{q}(t)$ 和时间 t

过的限制，我们得到¹⁵

$$\mathcal{L} = \mathcal{L}(\Phi(\vec{x}, t), \partial_\mu \Phi(\vec{x}, t), \vec{x}) \quad (4.4)$$

这里一个著名的拉格朗日量密度 $\mathcal{L} = \frac{1}{2} (\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2)$ ，它可以导出 Klein-Gordon 方程。

场论的一大优点是它将时空等价的对待。在粒子理论中我们使用空间坐标 $\vec{q}(t)$ 作为时间的函数来描述粒子。尤其是，拉格朗日量中没有类似于 $\partial_{\vec{q}} t$ 的项（如果出现了这类项，我们该怎样理解它呢？）当我们讨论粒子的坐标时，意义是清楚的，但对时间做类似的陈述时，却很难有明确的意义⁸。

讨论了这么多，我们终于可以回到这章开头所说的极小化问题上来了。我们希望找到某些泛函

$$S[q(t)] = \int \mathcal{L} dt$$

的极小值，以得到正确的运动方程。

对于粒子而言，方程的解是使得泛函取极小值的正确路径；而对于场而言，解是正确的场函数。

此刻请不用担心 \mathcal{L} 的具体形式，下面数章将详细讨论如何对问题中的系统导出正确的拉格朗日量 \mathcal{L} 。现在，我们将使用之前所介绍过的变分法，对于一个一般的 \mathcal{L} 导出泛函 $S[q(t)]$ 的极小值。极小化过程将给出系统的运动方程。

4.4 Euler-Lagrange equation 欧拉-拉格朗日方程

我们从粒子理论出发，它能让我们搞清楚，在给定初末端点后，粒子是如何在两点间运动的。数学上，即寻找函数 $q(t)$ 使得作用量⁹

$$S = \int_{t_1}^{t_2} \mathcal{L} \left(q(t), \frac{dq(t)}{dt}, t \right)$$

取极值（极大值或极小值）。

⁸译注：利用不那么古老的相对论性语言，粒子理论下时空坐标是可以在形式上被同等对待的：用以求导的东西只能是曲线的参数，它可以取作粒子世界线的固有时、某一坐标系的坐标时、或者其他什么东西（只要性质合适，也可以是空间坐标）；而时空坐标则被同样的求导。

⁹译注：此处 \mathcal{L} 原文字体有误，已更正

¹⁵ 这里我故意使用了不同的符号 \mathcal{L} ，因为在场论中大多数时候我们都在和拉格朗日量密度 \mathcal{L} 打交道。两者之间满足关系式 $\mathcal{L} = \int d^3x \mathcal{L}$

使用记号

$$\dot{q}(t) = \frac{dq(t)}{dt}.$$

与之前的例子类似，取 $q(t) = a(t)$ 并让对这个函数进行一个扰动

$$a(t) + \epsilon(t)$$

其中 ϵ 依旧是一个无穷小量。对于一个粒子而言，我们必须同时改变其速度 $\dot{a}(t) + \dot{\epsilon}(t)$ ，其中 $\dot{\epsilon}(t) = \frac{d\epsilon(t)}{dt}$ 。

在边界上，扰动后的路径应与原路径相同：

$$0 = \epsilon(t_1) = \epsilon(t_2) \quad (4.5)$$

这是因为我们寻找的是使作用量积分取极值的给定初末端点的路径。

这个扰动使得泛函变为

$$S = \int_{t_1}^{t_2} \mathcal{L}(q + \epsilon, \dot{q} + \dot{\epsilon}, t) dt.$$

¹⁶ 我们使用的是这个公式的多变量形式，参见附录B.3：

$$\mathcal{L}(q + \epsilon, \dot{q} + \dot{\epsilon}, t) = \mathcal{L}(q, \dot{q}, t) + \epsilon \frac{\partial \mathcal{L}}{\partial q} + \dot{\epsilon} \frac{\partial \mathcal{L}}{\partial \dot{q}} + \dots$$

¹⁷ 分部积分是莱布尼兹律的直接推论，详见附录B.2

与之前的例子类似，极小值要求线性依赖于扰动 ϵ 的项为零。由于我们处理的是一般的 \mathcal{L} ，故将其展开成 Taylor 级数¹⁶，并令一阶项为零

$$\int_{t_1}^{t_2} dt \left[\epsilon(t) \frac{\partial \mathcal{L}}{\partial q} + \left(\frac{d}{dt} \epsilon(t) \right) \frac{\partial \mathcal{L}}{\partial \dot{q}} \right] \stackrel{!}{=} 0. \quad (4.6)$$

对后一项做分部积分¹⁷得到

$$\begin{aligned} & \int_{t_1}^{t_2} dt \left(\frac{d}{dt} \epsilon(t) \right) \frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} = \\ & \epsilon(t) \frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \epsilon(t) \frac{d}{dt} \left(\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} \right) \end{aligned}$$

利用(4.5)式有

$$\epsilon(t) \frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} \Big|_{t_1}^{t_2} = 0$$

因此, 我们可以将(4.6)改写为

$$\int_{t_1}^{t_2} dt \epsilon(t) \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} \right) \right] \stackrel{!}{=} 0$$

易见上式对于任意扰动 $\epsilon(t)$ 为零仅当方括号 $[\]$ 内的表达式为零。
故有¹⁸

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}} \right) = 0 \quad (4.7)$$

这就是著名的欧拉-拉格朗日方程 (**Euler-Lagrange equation**)

我们可以用类似的办法处理场论。首先, 注意到场论将时空同等对待。因此引入拉格朗日量密度:

$$\mathcal{L} = \int d^3\mathbf{x} \mathcal{L}(\Phi^i, \partial_\mu \Phi^i) \quad (4.8)$$

并用拉格朗日量密度来表示作用量

$$S = \int dt \mathcal{L} = \int d^4\mathbf{x} \mathcal{L}(\Phi^i, \partial_\mu \Phi^i) \quad (4.9)$$

依照上面讲过的步骤, 我们可以得到场的运动方程组^{10 19}:

$$\frac{\partial \mathcal{L}}{\partial \Phi^i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^i)} \right) = 0 \quad (4.10)$$

下一节中将在拉格朗日形式下导出近代物理学中重要的一个定理。从中可以看到对称性与守恒量之间的深刻联系。守恒量是描述自然的合适参量²⁰, 这个定理将教会我们如何和它打交道。

4.5 Noether's Theorem Noether 定理

Noether 定理表示, 拉格朗日量的每一个对称性都直接对应一个守恒量。换言之, 物理学家用于描述自然(守恒量)的记号和对称性直接相关。这确实是科学史上最美丽的创见之一。

4.5.1 粒子理论的 Noether 定理

先看看能对粒子理论中的守恒量做点什么。我们限于讨论连续对称性, 这样才能使用无穷小变换。如同上一章所言, 重复进行无穷小变换可以得到一个有限变换。拉格朗日量在无穷小变换²¹ $q \rightarrow$

¹⁸ 可能你会对于两类不同的导数符号感到困惑。 $\frac{d}{dt}$ 被称作对 t 的全导数, 而 $\frac{\partial}{\partial t}$ 被称为偏导数。全导数给出总改变量, 函数 f 的变化率, 即偏导数, 乘以变量自身的改变, 之和。例如, 三维空间中的函数 $f(x(t), y(t), z(t))$ 的总变化率为 $\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial t}$ 。即变化率乘以自身的改变量。相对的, 偏导数仅给出总改变量的一部分。对于一个不显式的依赖于 t 的函数, 其偏导数为零。例如, 对于 $f(x(t), y(t)) = x^2 y + y^3$, 我们有 $\frac{\partial f}{\partial t} = 0$, 但 $\frac{\partial f}{\partial x} = 2xy \neq 0$, $\frac{\partial f}{\partial y} = x^2 + 3y^2 \neq 0$ 。因此 $\frac{df}{dt} = 2xy \frac{\partial x}{\partial t} + (x^2 + 3y^2) \frac{\partial y}{\partial t}$ 。作为对比, 对于另一个函数 $g(x(t), y(t), t) = x^2 t + y$ 我们有 $\frac{\partial g}{\partial t} = x^2$

¹⁹ 在这里用“组”字是因为对于每一个场分量 Φ^1, Φ^2, \dots 我们都可以得到一个方程

²⁰ 我们将其视做锚定这个复杂的世界之物。不论万事万物如何改变, 守恒量依旧是守恒量。

²¹ 用符号 δ 来表述小扰动, 希望不会和用来表述偏导数的 ∂ 相混淆。

¹⁰ 不对 i 求和

$q' = q + \delta q$ 下不变的性质用数学表述如下¹¹

$$\begin{aligned}\delta\mathcal{L} &= \mathcal{L}(q, \frac{dq}{dt}, t) - \mathcal{L}(q + \delta q, \frac{d(q + \delta q)}{dt}, t) \\ &= \mathcal{L}(q, \frac{dq}{dt}, t) - \mathcal{L}(q + \delta q, \frac{dq}{dt} + \frac{d\delta q}{dt}, t) \stackrel{!}{=} 0\end{aligned}\quad (4.11)$$

要求拉格朗日量是变换不变的，这个限制太强了。为了保证动力学性质一样，要求不变的应该是作用量而不是拉格朗日量。当然，如果拉格朗日量不变，作用量自然不变：

$$\delta S = \int dt \mathcal{L}(q, \frac{dq}{dt}, t) - \int dt \mathcal{L}(q + \delta q, \frac{dq}{dt}, t) = \int dt \delta\mathcal{L} \stackrel{\text{若 } \delta\mathcal{L}=0}{=} 0 \quad (4.12)$$

在什么情况下，拉格朗日量改变而作用量不变呢？答案是，给拉格朗日量加上一个任意函数 G 对时间的全导数

$$\mathcal{L} \rightarrow \mathcal{L} + \frac{dG}{dt}$$

这时作用量不变²²

$$\delta S \rightarrow \delta S' = \delta S + \int_{t_1}^{t_2} dt \frac{d}{dt} \delta G = \delta S + \underbrace{\frac{\partial G}{\partial q} \delta q}_{=0 \text{ 因为 } \delta q(t_1)=\delta q(t_2)=0} \Big|_{t_1}^{t_2}.$$

最后一步中用到了扰动 δq 在初末时刻¹² t_1, t_2 为零。由此可知，拉格朗日量的扰动 $\delta\mathcal{L}$ 不必非要为零，而仅需满足一个更弱的条件

$$\delta\mathcal{L} \stackrel{!}{=} \frac{dG}{dt}. \quad (4.13)$$

这意味着拉格朗日量可以在不改变作用量和运动方程的情况下，加减某个函数的全微分 $\frac{dG}{dt}$ 。将(4.11)式的右端由 0 改作 $\frac{dG}{dt}$ ，我们有

$$\delta\mathcal{L} = \mathcal{L}(q, \frac{dq}{dt}, t) - \mathcal{L}(q + \delta q, \frac{dq}{dt} + \frac{d(\delta q)}{dt}, t) \stackrel{!}{=} \frac{dG}{dt} \quad (4.14)$$

¹¹译者注：该书中 $\delta S, \delta\mathcal{L}$ 与平常的约定都差了一个负号。

¹²译注：右侧删去了原文的一对括号，原文为 (t_1, t_2)

²² 用到了 $\delta G = \frac{\partial G}{\partial q} \delta q$ 。这是因为 $G = G(q)$ 而我们改变了 q ，故 G 的扰动等于改变率 $\frac{\partial G}{\partial q}$ 乘上 q 的扰动

将第二项按 Taylor 级数展开到 δq 的一阶项, 并使用记号 $\frac{dq}{dt} = \dot{q}$

$$\begin{aligned} \rightarrow \delta \mathcal{L} &= \mathcal{L} - \mathcal{L} - \frac{\partial \mathcal{L}}{\partial q} \delta q - \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} \stackrel{!}{=} \frac{dG}{dt} \\ \rightarrow \delta \mathcal{L} &= -\frac{\partial \mathcal{L}}{\partial q} \delta q - \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} \stackrel{!}{=} \frac{dG}{dt} \end{aligned} \quad (4.15)$$

利用欧拉-拉格朗日方程²³重写(4.15)式, 有

$$\rightarrow \delta \mathcal{L} = -\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q - \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} \stackrel{!}{=} \frac{dG}{dt}$$

利用莱布尼兹律²⁴, 我们有

$$\begin{aligned} \rightarrow \delta \mathcal{L} &= -\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q \right) \stackrel{!}{=} \frac{dG}{dt} \\ \rightarrow \frac{d}{dt} \underbrace{\left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q + G \right)}_{\equiv J} &\stackrel{!}{=} 0 \end{aligned} \quad (4.16)$$

由此得到了一个不随时间变化的物理量 J :

$$J = \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q + G, \quad (4.17)$$

这是因为

$$\frac{dJ}{dt} = 0 \rightarrow J = \text{常数}。$$

为了看清楚这一点, 先借用一下后面章节的结论: 常质量自由粒子的 Newton 第二定律是²⁵

$$m\ddot{\vec{q}} = 0. \quad (4.18)$$

相应的拉格朗日量是²⁶

$$\mathcal{L} = \frac{1}{2} m \dot{\vec{q}}^2 \quad (4.19)$$

你可以将其代入欧拉-拉格朗日方程 ((4.7)式) 中去验证它。

让我们从这个拉格朗日量出发, 计算不同对称性所分别对应的守恒量。

由于 $\mathcal{L} = \frac{1}{2} m \dot{\vec{q}}^2$ 不依赖于 \vec{q} , 故此拉格朗日量在空间平移变换 $\vec{q} \rightarrow \vec{q} + \vec{a}$ 不变 ($\delta \mathcal{L} = 0$)。这个对应的守恒量满足条件式(4.13),

²³(4.7)式: $\frac{\partial \mathcal{L}}{\partial q} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right)$

²⁴ 如果你对莱布尼兹律仍有疑问, 请参考附录B.1

²⁵ 如果 q 代表某个物体的位置, 则 $\frac{dq}{dt} = \dot{q}$ 是这个物体的速度, $\frac{d}{dt} \frac{dq}{dt} = \frac{d^2 q}{dt^2} = \ddot{q}$ 是加速度。

²⁶ 我们现在处理的是高于一维的问题, 所以使用矢量 \vec{q} 和 \vec{a} 而不是 q 和 a 。

即 $G = 0$ 的情形¹³

$$J_{\text{trans}} = \frac{\partial \mathcal{L}}{\partial \dot{\vec{q}}} \vec{a} = m \dot{\vec{q}} \vec{a} = \vec{p} \vec{a}, \quad (4.20)$$

其中 $\vec{p} = m \dot{\vec{q}}$ 是经典力学中的**动量 (momentum)**。由于等式 $\frac{dJ}{dt}$ 对于任意 \vec{a} 都成立，所以说：拉格朗日量的空间平移不变性导致了**动量守恒**。

现在考虑旋转变换，这得需要大于一的维度，因为一维上的旋转是没有意义的。下面将不会使用类似 \vec{q} 的矢量符号，而是更方便的指标记号。我们能将所有方程简单的用 $\vec{q} \rightarrow q_i$ 改写。考虑无穷小旋转²⁷ $q_i \rightarrow q_i + \epsilon_{ijk} q_j a_k$ ，由此 $\delta q_i = \epsilon_{ijk} q_j a_k$ 。仍然是因为由于 $\mathcal{L} = \frac{1}{2} m \dot{\vec{q}}^2$ 不依赖于 \vec{q} ，我们的拉格朗日量在这样的变换下也不改变，其对应的守恒量是²⁸

$$\begin{aligned} J_{\text{rot}} &= \frac{\partial \mathcal{L}(q_i, \dot{q}_i, t)}{\partial \dot{q}_i} \delta q_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \epsilon_{ijk} q_j a_k \\ &= m \dot{q}_i \epsilon_{ijk} q_j a_k = m p_i \epsilon_{ijk} q_j a_k \end{aligned} \quad (4.21)$$

$$\rightarrow J_{\text{rot}} = (\vec{p} \times \vec{q}) \cdot \vec{a} \equiv \vec{L} \cdot \vec{a}$$

在最后一步中我们将表达式用矢量记号重新写出，其中 \times 称作叉乘，而 \vec{L} 是经典力学中**角动量 (angular momentum)**。即：旋转不变性导致了角动量守恒。

接下来让我们看看时间平移不变性²⁹。一个无穷小的时间平移 $t \rightarrow t' = t + \epsilon$ 会导致

$$\begin{aligned} \delta \mathcal{L} &= \mathcal{L} \left(q(t), \frac{dq(t)}{dt}, t \right) - \mathcal{L} \left(q(t + \epsilon), \frac{dq(t + \epsilon)}{dt}, t + \epsilon \right) \\ &= -\frac{\partial \mathcal{L}}{\partial q} \frac{\partial q}{\partial t} \epsilon - \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial t} \epsilon - \underbrace{\frac{\partial \mathcal{L}}{\partial t} \epsilon}_{\text{等式左端正好为全导数}} - \frac{d\mathcal{L}}{dt} \epsilon, \end{aligned} \quad (4.22)$$

这表明一般而言 $\delta \mathcal{L} \neq 0$ ，但有 $G = -\mathcal{L}$ 。

将其代入(4.16)式，有¹⁴

$$\frac{d}{dt} \underbrace{\left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} - \mathcal{L} \right)}_{\equiv \mathcal{H}} = 0 \quad (4.23)$$

守恒量 \mathcal{H} 称为**哈密顿量 (Hamiltonian)**，代表了系统的总能量。

²⁷ 这里我们用 Levi-Civita 符号来写旋转生成元。参见(3.63)式下的文字。

²⁸ 下式由(4.16)式导出，这里对应的依旧是 $G = 0$ 的情况

²⁹ 这意味着物理不会管我们是在昨天、今天还是五十年后做的实验，给定相同的初始条件，物理定律不会变化。

¹³ 译注：使用了求和约定，见第3章边注129

¹⁴ 译注：原文有字体混乱，已更正

对于这个例子，我们有

$$\begin{aligned}\mathcal{H} &= \frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} - \mathcal{L} = \underbrace{\left(\frac{\partial}{\partial \dot{q}} \frac{1}{2} m \dot{q}^2 \right)}_{=m\dot{q}} \dot{q} - \frac{1}{2} m \dot{q}^2 \\ &= m \dot{q}^2 - \frac{1}{2} m \dot{q}^2 = \frac{1}{2} m \dot{q}^2,\end{aligned}\quad (4.24)$$

正好是系统的动能。由于这里没有势场或外力，故动能就是总能量。能导出在外势场中运动的粒子的 Newton 第二定律 $m\ddot{q} = -\frac{\partial V}{\partial q}$ ¹⁵ 的拉格朗日量是

$$\mathcal{L} = \frac{1}{2} m \dot{q}^2 - V(q)。$$

其哈密顿量为

$$\mathcal{H} = \frac{1}{2} m \dot{q}^2 + V$$

的确是正确的总能量（等于动能加势能）。

由**推动不变性 (boost invariance)** 导致的守恒量略有些奇怪，相应的计算放在了这一章的附录4.6中。其结果为

$$\tilde{J}_{\text{boost}} = \underbrace{pt - \frac{1}{2} mvt}_{\equiv \tilde{p}t} - mq = \tilde{p}t - mq \quad (4.25)$$

我们可以看到这个量依赖于时间零点，可以通过恰当的选取时间零点使其为零。由于这个量是守恒量，这个守恒律告诉我们零始终都是零。¹⁶

作为粒子理论的总结，我们得到了如下关系：

- 空间平移不变性 \Rightarrow 动量守恒
- 推动不变性³⁰ $\Rightarrow \tilde{p}t - mq$ 守恒
- 旋转不变性 \Rightarrow 角动量守恒
- 时间平移不变性 \Rightarrow 能量守恒

Noether 定理向我们展示了，为什么这些记号³¹会以这个或那个形式出现在形形色色的物理理论中。只要我们依旧有通常的时空对称性，动量、能量和角动量就是守恒量。在场论中，有两类对称

³⁰ 推动的另一个叫法是动量空间的平移，因为变换 $q \rightarrow q + vt$ 使动量 $m\dot{q} \rightarrow m(\dot{q} + v)$

³¹ 除了推动不变性导致的守恒量

¹⁵ 原文有误，已更正

¹⁶ 译注：这个守恒率没有说的那么平庸，事实上，它等价于质心运动定理；
又注：该结果以及附录4.6的推导有误，参见译注 23

性：一方面，拉格朗日量会在时空变换（如旋转）下不变；另一方面，我们也会有在场自身的变换下的不变性，称为内禀对称性 (internal symmetries)。

4.5.2 场论中的 Noether 定理——时空对称性

对于场，我们得区分在时空变换下两类不同的改变。当观测者 S 看到场 $\Psi(x)$ 时，观测者 S' 看到的则是场 $\Psi'(x')$ 。这是同一个场，只是从不同的视角来看，会有不同的分量值。这两类描述之间由 Lorentz 群中的一个合适的变换相联系。我们现在将会使用在 3.7.11 节中介绍过的场表示。（无穷维）微分算符表示 x 变换到 x' 。这意味着利用这个表示我们能在不同的时空点或者在一个旋转参考系中计算场分量。Lorentz 群的有限维变换使 Ψ 映射到 Ψ' ，即，混合场分量³²。

对于一个依赖于时空的场的具体变换，需要同时考虑两个部分。我们将分别处理它们，首先来看 x 到 x' 的这一项。对于旋转来说，单由这一个部分推出的守恒量并不会真的守恒，这是因为我们忽略了变换的第二项。仅当将变换的两项 $x \rightarrow x'$ 和 $\Psi \rightarrow \Psi'$ 分别导致的守恒量加起来之后得到的，才是真正守恒的。

为了使上面这段论证更有说服力，我们考虑一个一般的拉格朗日量密度 $\mathcal{L}(\Phi(x_\mu), \partial_\mu \Phi(x_\mu), x_\mu)$ ¹⁷。对称性意味着

$$\mathcal{L}(\Phi(x_\mu), \partial_\mu \Phi(x_\mu), x_\mu) = \mathcal{L}(\Phi'(x'_\mu), \partial_\mu \Phi'(x'_\mu), x'_\mu)。 \quad (4.26)$$

一般的，当函数自身发生改变且其计算的点也发生改变时，函数的函数的总改变量由下式给出³³

$$\delta f(g(x), h(x), \dots) = \frac{\partial f}{\partial g} \delta g + \frac{\partial f}{\partial h} \delta h + \dots + \frac{\partial f}{\partial x}。 \quad (4.27)$$

运用到拉格朗日量上有

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \Phi} \delta \Phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \delta (\partial_\mu \Phi) + \frac{\partial \mathcal{L}}{\partial x_\mu} \delta x_\mu, \quad (4.28)$$

用欧拉-拉格朗日方程³⁴将其改写为

³² 请回忆这个例子，Weyl 旋量有两个分量而矢量场有四个分量。如果我们从一个不同的视角来考察矢量场 $A_\mu = \begin{pmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{pmatrix}$ ，

即，从旋转系看，它会形如 $A'_\mu = \begin{pmatrix} A'_0 \\ A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \begin{pmatrix} A_0 \\ -A_2 \\ A_1 \\ A_3 \end{pmatrix}$ 。 A'_μ 和 A_μ

描述了从两个相差了绕 z 轴旋转 90° 的坐标系分别观察到的同一个场。

³³ 如果你对此不熟悉：这常被称为全导数。总改变量等于变化率，即偏导数，乘以变量自身的改变，之和。例如，三维空间中的函数 $f(x, y, z)$ 的总该变量为 $\frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial z} \delta z$ ，即变化率乘以自身改变的距离。我们考察的是无穷小变化，故我们可以略去 Taylor 展开的高阶项仅考虑第一项。

³⁴ (4.10) 式：

$\frac{\partial \mathcal{L}}{\partial \Phi} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right)$

¹⁷ 译注：原文多出一左括号

$$\begin{aligned}\delta\mathcal{L} &= \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \right) \delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \underbrace{\delta(\partial_\mu\Phi)}_{=\partial_\mu\delta\Phi} + \frac{\partial\mathcal{L}}{\partial x_\mu} \delta x_\mu \\ &\stackrel{\text{莱布尼兹律}}{=} \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \delta\Phi \right) + \frac{\partial\mathcal{L}}{\partial x_\mu} \delta x_\mu\end{aligned}\quad (4.29)$$

其中 $\delta\Phi$ 有两项

$$\delta\Phi = \epsilon_{\mu\nu} S^{\mu\nu} \Phi(x) - \frac{\partial\Phi(x)}{\partial x_\mu} \delta x_\mu, \quad (4.30)$$

其中 $\epsilon_{\mu\nu}$ 为旋转参数, $S_{\mu\nu}$ 为相应的有限维群表示中的旋转算符, 额外的负号仅为了方便。 $S_{\mu\nu}$ 是由旋转生成元 $S_i = \frac{1}{2}\epsilon_{ijk}S_{jk}$ 和推动生成元 $K_i = S_{i0}$ 构成, 类似于(3.165)式中 $M_{\mu\nu}$ 的定义。 $S_{\mu\nu}$ 的这个定义使得我们同时进行旋转和推动。

第一项仅在旋转和推动时较重要, 因为平移不会导致场分量的混合。推动所对应的守恒量依旧不是特别有意思, 就像在粒子理论中得到的那样。所以实际上这一项只和旋转相关。

让我们从最简单的场变换出发: 时空平移, 即

$$x_\mu \rightarrow x'_\mu = x_\mu + \delta x_\mu = x_\mu + a_\mu \quad (4.31)$$

将 $\epsilon_{\mu\nu} = 0$ (在平移下场分量不会混合) 代入方程(4.30)

$$\delta\Phi = -\frac{\partial\Phi(x)}{\partial x_\mu} \delta x_\mu$$

再由(4.29)式, 考察不变性 ($\delta\mathcal{L} = 0$) 所导致的守恒量

$$-\partial_\nu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\Phi)} \frac{\partial\Phi(x)}{\partial x_\mu} \delta x_\mu \right) + \frac{\partial\mathcal{L}}{\partial x_\mu} \delta x_\mu = 0 \quad (4.32)$$

$$\rightarrow -\partial_\nu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\Phi)} \frac{\partial\Phi(x)}{\partial x^\mu} - \delta_\mu^\nu \mathcal{L} \right) \delta x^\mu = 0 \quad (4.33)$$

由(4.31)式我们有 $\delta x^\mu = a^\mu$, 代入(4.33)式中可得¹⁸

$$-\partial_\nu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu\Phi)} \frac{\partial\Phi(x)}{\partial x^\mu} - \delta_\mu^\nu \mathcal{L} \right) a^\mu = 0, \quad (4.34)$$

我们定义能动张量

$$T_\mu^\nu := \frac{\partial\mathcal{L}}{\partial(\partial_\nu\Phi)} \frac{\partial\Phi(x)}{\partial x^\mu} - \delta_\mu^\nu \mathcal{L}. \quad (4.35)$$

¹⁸译者注: 原文上下数式的指标不平衡, 现已修正。

由于 a_μ 是任意的, (4.34)式告诉我们 T_μ^ν 满足连续性方程

$$\partial_\nu T_\mu^\nu = 0 \quad (4.36)$$

$$\rightarrow \partial_\nu T_\mu^\nu = \partial_0 T_\mu^0 - \partial_i T_\mu^i = 0 \quad (4.37)$$

对于所有 μ 成立。这直接告诉我们有守恒量的存在, 例如取 $\mu = 0$ 我们有³⁵

$$\begin{aligned} \partial_0 T_0^0 - \partial_i T_0^i &= 0 \rightarrow \partial_0 T_0^0 = \partial_i T_0^i \\ \partial_t T_0^0 &= -\nabla \cdot \vec{T} \xrightarrow{\text{在某个无限的空间域 } V \text{ 上积分}} \int_V d^3\mathbf{x} \partial_t T_0^0 = - \int_V d^3\mathbf{x} \nabla \cdot \vec{T} \end{aligned}$$

$$\partial_t \int_V d^3\mathbf{x} T_0^0 = - \int_V d^3\mathbf{x} \nabla \cdot \vec{T} \xrightarrow{\text{散度定理: } \delta V \text{ 代表 } V \text{ 的表面}} - \int_{\delta V} d^2\mathbf{x} \vec{T} \cdot \hat{n} \underset{\text{场在无穷远处为零}}{=} 0 \quad (4.38)$$

$$\rightarrow \partial_t \int_V d^3\mathbf{x} T_0^0 = 0 \quad (4.39)$$

最后一步中是这样处理的: 考虑一个无限大的空间域, 如一个有着无限大半径 r 的球体, 我们需要在其表面计算积分, 即在 $r = \infty$ 处计算场值。在2.3节中, 我们发现了物理中任何东西都有一个速度上界。由此无穷远处的场不可能对有限远 x 处的物理有任何影响, 故我们说场在无穷远处为零¹⁹。

最终的结论是: 时空平移不变性导致四个守恒量³⁶, (4.39)式告诉我们它们是

$$E = \int d^3\mathbf{x} T_0^0 \quad (4.40)$$

$$P_i = \int d^3\mathbf{x} T_i^0 \quad (4.41)$$

其中 $i = 1, 2, 3$ 。它们分别被称为系统的总能量——因时间平移 $x_0 \rightarrow x_0 + a_0$ 不变性而守恒; 场的总动量——因空间平移 $x_i \rightarrow x_i + a_i$ 不变性而守恒。

³⁵ 使用了 $\partial_0 = \partial_t$

和 $\partial^i T_i^0 = \nabla \cdot \vec{T}$,

以及著名的散度定理 $\int_V d^3\mathbf{x} \nabla A = \int_{\delta V} d^2\mathbf{x} A$, 即能让体积分化作面积分。对于散度定理, 在那儿叫 Gauss 定理, 有一个非常有启发性的证明, 可以在 Richard P. Feynman, Robert B. Leighton, and Matthew Sands.

The Feynman Lectures on Physics: Volume 2. Addison-Wesley,

1st edition, 2 1977.

ISBN 9780201021172 的第三章找到, 可在此免费在线查看:

http://www.feynmanlecture.s.caltech.edu/II_03.html。

(译注: 有中译本: 费恩曼, 莱顿, 桑兹著; 李洪芳, 王子辅, 钟万衡译。费恩曼物理学讲义, 新千年版, 第二卷, 上海: 上海科学技术出版社, 4 2013, ISBN 9787547816370)

³⁶ 因为对于任意 a_μ 都有 $\partial_0 T_\mu^0 a_\mu = \partial_0 T_0^0 a_0 - \partial_i T_i^0 a^i = 0$ 成立, 所以对于每一个分量我们都能得到一个连续性方程。

¹⁹译注: 这段讨论怪怪的, 正确性令人起疑。

4.5.3 旋转和推动

接下来,我们会仔细研究旋转和推动下的不变性。先从(4.30)式的第二项出发,随后讨论第一项的含义。我们将从第一项和第二项分别得到一个量,两项之和是守恒的。标量场没有可以混合的分量,故对于它而言第一项始终为零,即 $S_{\mu\nu} = 0$; 也可以用3.7.4节的结论解释:我们用标量作为 Lorentz 群一维表示的生成元。故在第一部分中导出的将会是对应于标量场真实守恒量。

变换的第二项由下式给出

$$x_\mu \rightarrow x'_\mu = x_\mu + \delta x_\mu = x_\mu + M_\mu^\sigma x_\sigma$$

其中 M_μ^σ 是转动和推动的在无穷维表示下的生成元,即在(3.240)式中所定义微分算符。

将 $\delta x_\mu = M_\mu^\sigma x_\sigma$ 代入(4.33)中得

$$\begin{aligned} \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu \Phi)} \frac{\partial \Phi(x)}{\partial x^\mu} - \delta_\mu^\nu \mathcal{L} \right) M^{\mu\sigma} x_\sigma &= 0 \\ \underbrace{\rightarrow}_{(4.35)\text{式}} \partial_\nu T_\mu^\nu M^{\mu\sigma} x_\sigma &= 0 \end{aligned} \quad (4.42)$$

我们能将它改写成更常见的形式

$$\begin{aligned} &= \frac{1}{2} (\partial^\nu T_\nu^\mu M_{\mu\sigma} x^\sigma + \partial^\nu T_\nu^\mu \underbrace{M_{\mu\sigma} x^\sigma}_{=-M_{\sigma\mu}}) = \frac{1}{2} (\partial^\nu T_\nu^\mu M_{\mu\sigma} x^\sigma + \partial^\nu T_\nu^\mu - M_{\sigma\mu} x^\sigma) \\ &\stackrel{\text{替换哑指标}}{=} \frac{1}{2} (\partial^\nu T_\nu^\mu M_{\mu\sigma} x^\sigma + \partial^\nu T_\nu^\sigma - M_{\mu\sigma} x^\mu) = \frac{1}{2} (\partial^\nu T_\nu^\mu x^\sigma + \partial^\nu T_\nu^\sigma x^\mu) M_{\mu\sigma} \\ &= \frac{1}{2} \partial_\nu (T^{\mu\nu} x^\sigma + \partial^\nu T^{\sigma\nu} x^\mu) M_{\mu\sigma} = 0 \\ &\rightarrow \text{引入 } (J^\nu)^{\sigma\mu} \equiv T^{\mu\nu} x^\sigma - \partial^\nu T^{\sigma\nu} x^\mu \rightarrow \partial_\nu (J^\nu)^{\sigma\mu} = 0. \end{aligned} \quad (4.43)$$

在最后一步引入的 J^ν 被称为 **Noether 流 (Noether current)**。注意到,由于在(3.165)式定义 M 时引入的反对称性 $M_{\mu\sigma} = -M_{\sigma\mu}$, 有 $M_{\mu\mu} = 0$ 。由此我们找到了六个不同的连续性方程 $\partial_\nu (J^\nu) = 0$, 分别对应于每一个非零的 $M_{\mu\sigma}$ ²⁰。使用与(4.39)式一样的方式, 我们可以论证如下的量不随时间改变

$$Q^{\mu\sigma} = \int d^3\mathbf{x} (T^{\mu 0} x^\sigma - T^{\sigma 0} x^\mu) \quad (4.44)$$

²⁰译注: 排除了零元的同时, 也将始终互为相反数的一对元视作一个

³⁷ 回忆一下(3.165)式中提到的转动生成元 J_i 和 $M_{\mu\nu}$ 之间的关系: $J_i = \frac{1}{2}\epsilon_{ijk}M_{jk}$, 这里我们使用拉丁字母 $i, j \in \{1, 2, 3\}$ 而不是希腊字母 $\mu, \nu \in \{0, 1, 2, 3\}$, 将其限制在空间分量上。

³⁸ 下标 orbit (轨道) 的含义马上就会清楚, 因为它实际上只是守恒量的一部分, 我们马上就会去考察第二项。

³⁹ 请回忆推动生成元 K_i 和 $M_{\mu\nu}$ 之间的关系: $K_i = M_{0i}$

由此我们得到了对应于转动³⁷不变性的守恒量

$$L_{\text{orbit}}^i = \frac{1}{2}\epsilon_{ijk}Q_{jk} = \frac{1}{2}\epsilon_{ijk} \int d^3\mathbf{x} (T^{\mu 0}x^\sigma - \partial^\nu T^{\sigma 0}x^\mu), \quad (4.45)$$

其被称为场的轨道角动量³⁸。

同样的, 我们也有所对应的守恒量³⁹

$$Q^{0i} = \int d^3\mathbf{x} (T^{i0}x^0 - T^{00}x^i) \quad (4.46)$$

将会在附录4.7节中详细讨论。

4.5.4 自旋

接下来我们希望捡回原先在导出守恒量时所丢掉的(4.30)中的第一项。现在考察

$$\delta\Phi = \epsilon_{\mu\nu}S^{\mu\nu}\Phi(x), \quad (4.47)$$

⁴⁰ 场分量的混合是由有限维表示来进行的。例如转动生成元的二维表示 $J_i = \frac{1}{2}\sigma_i$ 混合了 Weyl 旋量的分量。

其中 $S^{\mu\nu}$ 是问题中的变换⁴⁰的恰当有限维表示表示。这使(4.29)式多了一项

$$\partial_\rho \left(\frac{\partial \mathcal{L}}{\partial(\partial_\rho \Phi)} \epsilon_{\mu\nu}S^{\mu\nu}\Phi(x) \right) \quad (4.48)$$

同时也使(4.45)式多了一项。完整的守恒量是

$$L^i = \frac{1}{2}\epsilon_{ijk}Q^{jk} = \frac{1}{2}\epsilon_{ijk} \int d^3\mathbf{x} \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi)} S^{jk}\Phi(x) + (T^{k0}x^j - T^{j0}x^k) \right) \quad (4.49)$$

据此写出

$$L^i = L_{\text{spin}}^i + L_{\text{orbit}}^i. \quad (4.50)$$

第一项是个新东西, 但和我们之前考虑的轨道角动量有类似之处, 因为这是当我们讨论同一个不变性时出现而加在一起的。标准的观点是这个守恒量是某种内禀⁴¹角动量。

回忆我们曾用某种也叫做自旋 (spin) 的东西来给 Poincaré 群的表示作标签。现在这个记号再一次出现了。我们将在下一章, 并在8.5.5节中相当仔细地, 学习如何能测量这类新的角动量, 自旋,

⁴¹ 我们随后将会看到场产生湮灭粒子。一个自旋 $\frac{1}{2}$ 场产生自旋 $\frac{1}{2}$ 粒子, 这是基本粒子的固有属性。所以说我们用“内禀”这个词。轨道角动量是用来描述两个或更多粒子如何相互旋转的一个量。

以及认识到它和我们用以给 Poincaré 群的表示作标签的东西，确实是相一致的。

4.5.5 场论中的 Noether 定理——内禀对称性

我们现在来仔细研究内禀对称性⁴²。在某些场自身的无穷小变换

$$\Phi_i \rightarrow \Phi'_i = \Phi_i + \delta\Phi_i \quad (4.51)$$

下拉格朗日量的不变性 ($\delta\mathcal{L} = 0$)，数学上表达为²¹

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}(\Phi_i, \partial_\mu \Phi_i) - \mathcal{L}(\Phi_i + \delta\Phi_i, \partial_\mu(\Phi_i + \delta\Phi_i)) \\ &= -\frac{\partial\mathcal{L}}{\partial\Phi_i}\delta\Phi_i - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\partial_\mu\delta\Phi_i \stackrel{!}{=} 0, \end{aligned} \quad (4.52)$$

即因变换是无穷小的，故 Taylor 展开到 $\delta\Phi_i$ 第一项。依场的欧拉-拉格朗日方程(4.10)可知

$$\frac{\partial\mathcal{L}}{\partial\Phi} = \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \right)。$$

代入(4.42)式，我们得到

$$\delta\mathcal{L} = -\partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)} \right) \delta\Phi_i - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)} \partial_\mu\delta\Phi_i = 0$$

再用上莱布尼兹律，得到

$$\partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)} \delta\Phi_i \right) = 0 \quad (4.53)$$

这意味着，如果拉格朗日量在变换 $\Phi_i \rightarrow \Phi'_i = \Phi_i + \delta\Phi_i$ 下不变，那我们又得到一个被称为了 Noether 流的量

$$J^\mu \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} \delta\Phi, \quad (4.54)$$

其满足连续性方程

$$(4.53)\text{式} \rightarrow \partial_\mu J^\mu = 0. \quad (4.55)$$

更进一步，用导出(4.39)式同样的办法，能得到一个对时间的守恒

⁴² 毋庸置疑，对于相互作用场论，内禀对称性是相当重要的。除此之外，我们将从最简单的内禀对称性引出的守恒量出发开始学习量子场论

²¹ 译注：原文上下标混乱，已修改。

量

$$\partial_t \int d^3\mathbf{x} J^0 = 0. \quad (4.56)$$

⁴³ 注意此时我们并没有假定场是实的还是复的。常数 ϵ 是任取的，若我们想让场平移量是实的，我们当然可以令它是虚数。这里用 i 是为了使生成元是厄米的（现在来看还没什么意义），这样其本征值就是实的。现在不要太困扰，把这个额外的 i 当做一个约定就好，只要通过定义 $\epsilon' := i\epsilon$ ，就能把它吸收到常数里去。

我们再来看一个重要的例子：在场自身的平移下不变性⁴³，即在(4.51)式中的 $\delta\Phi = -i\epsilon$ ：

$$\Phi \rightarrow \Phi' = \Phi - i\epsilon. \quad (4.57)$$

或者对于不止一个分量的场

$$\Phi_i \rightarrow \Phi'_i = \Phi_i - i\epsilon_i. \quad (4.58)$$

这个变换的生成元是 $-i\frac{\partial}{\partial\Phi}$ ，因为

$$\Phi' = e^{-i\epsilon\frac{\partial}{\partial\Phi}}\Phi = (1 - i\epsilon\frac{\partial}{\partial\Phi} + \dots)\Phi \approx \Phi - i\epsilon \quad (4.59)$$

对应的守恒量被称为共轭 (conjugate) 动量 Π 。方便起见，引入共轭动量密度⁴⁴ $\pi \equiv J_0$ ，通过(4.53)、(4.56)式，并利用守恒量中不需要包含 $\delta\Phi = -i\epsilon$ 这件事。我们得到⁴⁵

$$\begin{aligned} \partial_t \Pi = 0 \rightarrow \partial_t \Pi &= \partial_t \int d^3\mathbf{x} J_0 = \partial_t \int d^3\mathbf{x} \pi = 0 \\ &\text{用到了 } J_0 = \pi = \frac{\partial\mathcal{L}}{\partial(\partial_0\Phi)} \end{aligned} \quad (4.60)$$

注意这和源于空间平移不变性 $\Phi(x) \rightarrow \Phi(x') = \Phi(x + \epsilon)$ 物理的场动量密度不同，那东西长这样

$$p^i = T^{0i} = \frac{\partial\mathcal{L}}{\partial(\partial_0\Phi)} \frac{\partial\Phi}{\partial x_i},$$

即之前在(4.41)式中导出的东西。下面数章中我们将会看到更多不同的内禀对称性，它们在发展相互作用理论时是无价的。现在来总结一下我们学过的场论中的对称性和守恒量：

- 空间平移不变性 \Rightarrow 物理动量

$$P^i = \int d^3\mathbf{x} T^{0i} = \int d^3\mathbf{x} \frac{\partial\mathcal{L}}{\partial(\partial_0\Phi)} \frac{\partial\Phi}{\partial x_i} \text{ 守恒}$$

- 推动不变性⁴⁶ \Rightarrow 能量中心的速度是常数
- 转动不变性 \Rightarrow 总角动量

$$L^i = \frac{1}{2}\epsilon_{ijk} \int d^3\mathbf{x} \left(\frac{\partial\mathcal{L}}{\partial(\partial_0\Phi)} S^{jk}\Phi(x) + (T^{k0}x^j - T^{j0}x^k) \right)$$

⁴⁴ 这个相当概括的量是量子场论最重要的记号，我们将在第6章中看到这一点

⁴⁵ (4.54)式中的 $\delta\Phi$ 成为一个常数 ϵ ，因为这个量对于任何 ϵ 都守恒，所以我们在定义守恒量时不需要它。

⁴⁶ 推动又被称为在动量的空间平移

$= L_{\text{spin}}^i + L_{\text{orbit}}^i$, 包含自旋和轨道角动量, 守恒

- 时间平移不变性 \Rightarrow 能量

$$E = \int d^3\mathbf{x} T^{00} = \int d^3\mathbf{x} \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi)} \frac{\partial \Phi}{\partial x_i} - \mathcal{L} \right) \text{ 守恒}$$

- 场自身的平移不变性 \Rightarrow 共轭动量

$$\Pi = \int d^3\mathbf{x} \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi)} = \int d^3\mathbf{x} \pi \text{ 守恒}$$

Further Reading Tips 阅读建议

- **Cornelius Lanczos - The Variational Principles of Mechanics**⁴⁷ 是一本关于拉格朗日体系在经典力学中的应用的绝妙的书。
- **Herbert Goldstein - Classical Mechanics**⁴⁸ 是经典力学的标准教材, 其中有许多对拉格朗日体系非常好的解释。

⁴⁷Cornelius Lanczos. The Variational Principles of Mechanics. Dover Publications, 4th edition, 3 1986. ISBN 9780486650678

⁴⁸Herbert Goldstein, Charles P. Poole Jr., and John L. Safko. Classical Mechanics. Addison-Wesley, 3rd edition, 6 2001. ISBN 9780201657029

4.6 Appendix: Conserved Quantity from Boost Invariance for Particle Theories 附录：粒子理论中的推动不变性与其守恒量

此附录中我们想找到对于示例拉格朗日量 $\mathcal{L} = \frac{1}{2}m\dot{q}^2$ 其来自于推动不变性的守恒量。一个推动变换即 $q \rightarrow q' = q + \delta q = q + vt$, 其中 v 代表一个常值速率。该拉格朗日量在推动下并不是不变的, 但正如之前就看到过的, 这并不是一个问题:

为了使运动方程在变换下不变, 拉格朗日量至多允许变化一个任意函数的全导数。这个条件在推动变换 $q \rightarrow q' = q + vt \Rightarrow \dot{q}' = \dot{q} + v$
 $v=\text{常数!}$

下是显然满足的, 即我们的拉格朗日量

$$\mathcal{L} = \frac{1}{2}m\dot{q}^2$$

的改变量为

$$\begin{aligned} \delta \mathcal{L} &= \mathcal{L}(q, \dot{q}, t) - \mathcal{L}(q', \dot{q}', t) = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\dot{q}'^2 \\ &= \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m(\dot{q} + v)^2 = -m\dot{q}v - \frac{1}{2}mv^2, \end{aligned} \quad (4.61)$$

等于下面这个函数的全导数⁴⁹

$$G \equiv -m\dot{q}v - \frac{1}{2}mv^2t.$$

⁴⁹ $\frac{dG}{dt} = \frac{d}{dt}(-m\dot{q}v - \frac{1}{2}mv^2t) = -m\dot{q}v - \frac{1}{2}mv^2$, 因为 v 和 m 都是常数。

⁵⁰ 看到这一点的另一个方法是直接看运动方程: $m\ddot{q} = 0$ 。我们有变换 $q \rightarrow q' = q + \delta q = q + vt$, 其中 v 是常数, 且有 $\dot{q} \rightarrow \dot{q}' = \dot{q} + v$ 以及 $\ddot{q} \rightarrow \ddot{q}' = \ddot{q}$ 。

⁵¹ 拉格朗日量改变了, 即 $\delta\mathcal{L} \neq 0$, 但是作用量值最多改变一个全导数, 故我们得到了相同的运动方程。

我们断言运动方程⁵⁰不变⁵¹, 故我们找到了一个守恒量。

对于推动, 守恒量是 ((4.16)式)²²

$$J_{\text{boost}} = \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}}}_{\equiv p} \underbrace{\delta q}_{=vt} + G = pvt - mvq - \frac{1}{2}mv^2t,$$

由于 (4.16) 式右端为零, 故我们能除去一个 v 的公因子²³。由此有

$$\tilde{J}_{\text{boost}} = \underbrace{pt - \frac{1}{2}mvt}_{\equiv \tilde{p}t} - mq = \tilde{p}t - mq \quad (4.62)$$

这是一个奇怪的守恒量, 因为它的值依赖于时间零点的选取。我们能选择一个合适的时间零点使得这个量为零。由于它是守恒量, 故这个零始终为零。

4.7 Appendix: Conserved Quantity from Boost Invariance for Field Theories 附录：场论中的推动不变性与其守恒量

由场论中的推动不变性, 我们有如下守恒量

$$Q^{0i} = \int d^3\mathbf{x} (T^{i0}x^0 - T^{00}x^i), \quad (4.63)$$

即在(4.46)式中所导出的, 利用了 $K_i = M_{0i}$ 。

我们怎样诠释这个守恒量呢? Q^{0i} 是守恒的, 意味着

$$\begin{aligned} 0 &= \frac{\partial Q^{0i}}{\partial t} = \underbrace{\int d^3\mathbf{x} \frac{\partial T^{i0}}{\partial t} x^0}_{=x^0 \frac{\partial}{\partial t}} + \int d^3\mathbf{x} T^{i0} \underbrace{\frac{\partial x^0}{\partial t}}_{=1 \text{ 因为 } x^0=t} - \frac{\partial}{\partial t} \int d^3\mathbf{x} T^{00} x^i \\ &= t \frac{\partial P^i}{\partial t} + P^i - \frac{\partial}{\partial t} \int d^3\mathbf{x} T^{00} x^i \end{aligned}$$

由(4.41)式²⁴可知 P^i 是守恒的 $\frac{\partial P^i}{\partial t} = 0 \rightarrow P^i = \text{常量}$, 故我们断言

²²译注: 原文中下式有符号错误, 已改正。

²³译注: 由于之前关于守恒量的诸多推导都使用了变换是无穷小变换这一条件, 故这里 v 必须取无穷小量, 即略去 J_{boost} 中 v 的二阶项, 最后得到的才是守恒量。否则再利用其它条件我们会得到 $t \equiv 0$ 的荒谬结论。

²⁴译注: 原文公式引用有误, 已改正。

$$\frac{\partial}{\partial t} \int d^3x T^{00} x^i = P^i = \text{常数}. \quad (4.64)$$

故这个守恒律告诉我们⁵²能量中心以一个恒定的速度运动。或者从另一个视角来看：动量等于总能量乘以质心速度。

不论如何，(4.46)式中导出的守恒量还是显得很奇怪。它能够通过恰当选取起始时刻而取作零。我们能选择一个合适的时间零点使得这个量为零。由于它是守恒量，故这个零始终是零。²⁵

⁵² 不要太在意这一项的意义，因为它并不像其他守恒量那样有用。

²⁵ 译注：很是不能理解作者的意思。毕竟我们可以选取合适的能量零点使能量为零、合适的惯性系使动量为零、以及合适的匀速转动系使角动量为零。其他的守恒量并没有更优越的地方。

The Equations of Nature

自然界的方程

5	测量	119
5.1	量子力学中的算符	
5.2	量子场论的算符	
6	自由场理论	123
6.1	Lorentz 协变性与不变量	
6.2	Klein-Gordon 方程	
6.3	Dirac 方程	
6.4	Proca 方程	
7	相互作用理论	133
7.1	$U(1)$ 相互作用	
7.2	$SU(2)$ 相互作用	
7.3	质量项、 $U(1)$ 和 $SU(2)$ 的统一	
7.4	宇称破坏	
7.5	轻子质量项	
7.6	夸克质量项	
7.7	同位旋	
7.8	$SU(3)$ 相互作用	
7.9	Bosons 和 Fermions 间的游戏	

“如果你正在从给方程添加美感的方面考虑，你就是在成功的路上。”

– Paul A. M. Dirac

in *The evolution of the physicist's picture of nature*.

"If one is working from the point of view of getting beauty into one's equation,...one is on a sure line of progress." Scientific American, vol.208 Issue 5,pp.45-53.

Publication Date:
05/1963

5. Measuring Nature 测量

现在已经发现对称性与守恒量之间的关系了，接下来就可以利用这种联系。用更专业的话说，Noether 定理建立了对称变换的生成元与相应守恒量之间的联系。本章将利用这种联系。

守恒量常常被物理学家们用来描述物理系统，因为无论这个系统经历了怎样复杂的变化，守恒量都是一样的。比如，物理学家们为了描述一个实验会使用动量，能量或角动量。Noether 定理提示了一个至关重要的思想：描述自然的物理量等同于其对应的生成元：

$$\text{物理量} \Rightarrow \text{对应对称性的生成元} \quad (5.1)$$

我们会看到这种选择会自然地导出量子理论。下面通过考虑粒子理论来实现。

5.1 Operators of Quantum Mechanics 量子力学中的算符

惯例用一个尖帽子 \hat{O} 来表示一个算符

拉格朗日量在空间平移变换的生成元的作用下不变可导出动

量守恒。因此，由此可得¹

$$\text{动量 } \hat{p}_i \rightarrow \text{空间平移操作的生成元: } -i\partial_i$$

类似的，时间平移生成元带来的不变性给出能量守恒，也就是

$$\text{能量 } \hat{E} \rightarrow \text{时间平移操作的生成元 } -i\partial_0$$

¹ 或者说，我们可以观察守恒量与对应的推动下的不变性。注意到我们在4.6节中得到的非相对论性 Galilei 推动下的守恒量是从非相对论性的拉格朗日量得出来的。同样，我们可以对相对论性拉格朗日量做一样的事情，得到守恒量 $t\vec{p} - \vec{x}E$ 。相对论性能量由 $E = \sqrt{m^2 + p^2}$ 给出。在相对论性极限 $c \rightarrow \infty$ 下，我们有 $E \approx m$ ，并且，Lorentz 推动下的守恒量回到我们的 Galilei 推动的结果。因此，粒子理论的守恒量为 $M_i = (tp_i - x_i E)$ 。推动的生成元（见(3.240)，其中 $K_i = M_{0i}$ ）为 $K_i = i(x_0 \partial_i - x_i \partial_0)$ 。比较两者，其中 $x_0 = t$ ，给出 $M_i = (tp_i - x_i E) \leftrightarrow K_i = i(t\partial_i - x_i \partial_0)$ 。因此，利用我们之前的选择，很直观的我们有，坐标： $\hat{x}_i \rightarrow x_i$

² 如果你完全不了解量子力学，为什么这些简单的式子会如此重要对你来说看起来比较奇怪。你也听说过 Heisenberg 不确定性原理。在第8.3节，我们会更进一步的研究量子力学的形式与结构，然后就可以看到这个方程表明不能够以任意精度同时测量一个粒子的动量和坐标。我们的物理量的解释是算符的测量，而这个方程告诉我们先测量动量再测量位置与先测量位置再测量动量是完全不一样的。

³ 如果你以前没有过这个观念，注意我们可以把每一个矢量方程都写成只包含数的形式。比如，牛顿第二定律 $\vec{F} = m\vec{\ddot{x}}$ ，可以对于任意矢量 \vec{C} 写出 $\vec{F}\vec{C} = m\vec{\ddot{x}}\vec{C}$ 。不仅如此，如果对于任意的 \vec{C} 都成立，那么一直把这个 \vec{C} 写着也就没什么意义了。

没有关于“位置守恒”的对称性，所以位置并没有伴随生成元，我们有¹

$$\text{位置 } \hat{x}_i \rightarrow x_i$$

上面利用算符给出了描述自然的理论中使用的物理量。接下来我们自然要问：这些算符作用在什么上面？它们是如何与实验中的可观测量联系起来的呢？下一章仔细讨论这件事情，在这里只需要知道物理量，也就是算符，是作用在某物上的。不妨就认为它作用于一个抽象的对象上，称之为 Ψ ，后面会探究这个“某物”是什么鬼。

现在可以导出一些非常重要的²量子力学的方程了。就像上面已经解释过的那样，我们假定算符作用在抽象的 Ψ 上。于是有

$$\begin{aligned} [\hat{p}_i, \hat{x}_j]\Psi &= (\hat{p}_i \hat{x}_j - \hat{x}_j \hat{p}_i)\Psi = i(-\partial_i \hat{x}_j + \hat{x}_j \partial_i)\Psi \\ &\stackrel{\text{莱布尼兹律}}{=} -i(\partial_i \hat{x}_j)\Psi - i\hat{x}_j(\partial_i \Psi) + i\hat{x}_j(\partial_i \Psi) \stackrel{\partial_i \hat{x}_j = \frac{\partial x_j}{\partial x_i}}{=} -i\delta_{ij}\Psi \end{aligned} \quad (5.2)$$

这个方程对任意的 Ψ 都成立，因为我们没有对 Ψ 做任何假设，因此我们可以不考虑它³而直接的写出

$$[\hat{p}_i, \hat{x}_j] = -i\delta_{ij} \quad (5.3)$$

5.1.1 Spin and Angular Momentum 自旋与角动量

上一章的第4.5.4节中讲过，对标量理论而言，旋转不变性带来的守恒量有两个部分，第二部分对应轨道角动量，称之为无限维的

¹译者注：这里再次提醒一下，本书中虚数单位与指标经常共用字母 i ，相信聪明的读者有能力分辨。

轨道角动量生成元表示。

轨道角动量 $\hat{L}_i \rightarrow$ 转动生成元（无限维表示） $i\frac{1}{2}\epsilon_{ijk}(x^j\partial^k - x^k\partial^j)$

类似的，第一部分称作自旋，对应有限维度的生成元⁴的表示。

自旋 $\hat{S}_i \rightarrow$ 转动生成元（有限维表示） S_i

如前面(4.30)式解释过的那样， $S_{\mu\nu}$ 和转动算符生成元 S_i 的关系为 $S_i = \frac{1}{2}\epsilon_{ijk}S_{jk}$ 。

比如，当我们考虑自旋 $\frac{1}{2}$ 场的时候，我们需要用到我们在3.7.5节中得到的二维表示：

$$\hat{S}_i = i\frac{\sigma_i}{2} \quad (5.4)$$

其中 σ_i 是 Pauli 矩阵。在学习如何利用本章导出的这些算符之后，我们在8.5.5节深入讨论这一有趣而奇怪类型的角动量。时刻注意，只有这两部分角动量的和是守恒的。

⁴ 这是从场的分量的线性组合下的不变性带来的守恒量，因此是有限维的表示。

5.2 The Operators of Quantum Field Theory 量子场论的算符

场论的最重要的客体当然就是作为时空位置的函数⁵的场了 $\Phi = \Phi(x)$ 。我们接下来希望描述时空上的一个点发生的相互作用，从而我们需要考虑动力学变量 $\pi = \pi(x)$ 的密度，而不是其总量 $\Pi = \int d^3x \pi(x) \neq \Pi(x)$ 。

我们上一章发现了场自身的平移不变性 $\Phi \rightarrow \Phi - i\epsilon$ 产生了一种新的守恒量，称为共轭动量 Π 。就像我们上一章做的那样，我们用对应的生成元(4.60)式来表示共轭动量密度：

$$\text{共轭动量密度} \pi(x) \rightarrow \text{关于场自身平移的生成元：} -i\frac{\partial}{\partial\Phi(x)}$$

后面会看到量子场论与这里处理的稍稍不同，而现在的处理已经足够了。

就像上一节讨论的那样，我们需要给算符一些作用的对象。这里同样使用抽象的 Ψ ，后面的章节会具体说明。同样可以得到一个针对量子场论⁶十分重要的方程。

⁵ 这里， $x = x_0, x_1, x_2, x_3$ ，包含时间 $x_0 = t$ 。

⁶ 最后一步的时候我们类似 $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$ 给出 delta 分布 $\frac{\partial f(x_i)}{\partial f(x_j)} = \delta(x_i - x_j)$ ，它实际上可以严格的说明的。更具体的内容请看附录 D.2。

$$\begin{aligned}
[\Phi(x), \pi(y)]\Psi &= \left[\Phi(x), -i \frac{\partial}{\partial \Phi(y)} \right] \Psi \\
&\stackrel{\text{莱布尼兹律}}{=} \cancel{-i\Phi(x) \frac{\Psi}{\partial \Phi(y)}} + \cancel{i\Phi(x) \frac{\Psi}{\partial \Phi(y)}} + i \left(\frac{\partial \Phi(x)}{\partial \Phi(y)} \right) \Psi = i\delta(x-y)\Psi
\end{aligned} \tag{5.5}$$

同样的，由于这个式子对任意的 Ψ 都成立，可以写成：

$$[\Phi(x), \pi(y)] = i\delta(x-y) \tag{5.6}$$

类似的，对于场有多于一个分量的情况，我们有

$$[\Phi_i(x), \pi_j(y)] = i\delta(x-y)\delta_{ij} \tag{5.7}$$

后面我们会看到，几乎所有的量子场论的内容都要从这个简单的方程出发。

6. Free Theory 自由场理论

本章建立自由体系的物理理论，也就是说无相互作用的情况下的对称性给出的场¹。我们会

- 利用 Lorentz 群的 $(0,0)$ 表示给出 Klein-Gordon 方程
- 利用 Lorentz 群的 $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$ 表示给出 Dirac 方程
- 利用 Lorentz 群的矢量 $(\frac{1}{2},\frac{1}{2})$ 表示给出 Proca 方程，在无质量极限下退化到著名的 Maxwell 方程组

¹ 尽管这里针对的是场，后面会看到这里得到的方程同样可以用来处理粒子

6.1 Lorentz Covariance and Invariance Lorentz 协变性与不变量

本章会导出粒子物理的标准模型（这是我们拥有的最好的物理理论）中的运动方程。我们期待这些方程在所有的惯性系中都一样，因为如果不这样的话，每一个参考系都有各自的方程，而狭义相对论指出并没有一个特别的参考系，所以各惯性系中方程形式相同。这称为 **Lorentz 协变性**。一个 Lorentz 协变的对象是说它按照某一种给定的 Lorentz 群的表示来进行变换。比如说，矢量 A_μ ，通过 $(\frac{1}{2},\frac{1}{2})$ 表示变换，因此是 Lorentz 协变的；这其实就是说，Lorentz 变换： $A_\mu \rightarrow A'_\mu$ ， A_μ, A'_μ 实际是同一个对象（而非完全不同的）。另一方面，比如说， $A_1 + A_3$ 就不是 Lorentz 协变的，因为它不按照 Lorentz 群的某种表示来变换。但这并不是说我们不知道它是如何变换的，因为它的变换规律从 A_μ 的变换规律里

面可以简单地得出，但是在不同的惯性系中它看起来完全不一样。经推动变换到另一个惯性系之后它看起来或许像 $A_2 + A_4$ 。只包含 Lorentz 协变的对象的方程称为 Lorentz 协变方程，比如

$$A_\mu + 7B_\mu + C_\nu A^\nu D_\mu = 0$$

就是一个 Lorentz 协变方程，因为在另一个参考系中，它看起来就是

$$A_\mu + 7B_\mu + \underbrace{C_\nu A^\nu}_{\text{= 一个 Lorentz 标量, 根据 (0,0) 表示来进行变换}} D_\mu = A'_\mu + 7B'_\mu + C_\nu A^\nu D'_\mu = 0$$

它是一个 Lorentz 标量，根据 (0,0) 表示来进行变换

它在不同惯性系看起来是一样的。只有一部分是这样的方程一般不是 Lorentz 协变的，它在另一个惯性系里面会不一样。

为了导出 Lorentz 协变的方程，我们要求作用量 S 是 Lorentz 不变的。这是说 S 只能包含在参考系变换下保持不变的成分。换句话说，作用量里面只可以包含 Lorentz 变换下不变的成分。从作用量 S 中我们可以得到运动方程²。如果 S 依赖于参考系，那么运动方程就不是 Lorentz 协变的了。

就像前一章提到的，可以用更严格的约束条件 — 要求拉格朗日量是 Lorentz 不变的，这样作用量也就肯定是不变的了。

6.2 Klein-Gordon Equation Klein-Gordon 方程

下面从最简单的情况开始：标量，它按照 Lorentz 群的 (0,0) 表示变换。我们要找出对应的拉格朗日量来确定标量的运动方程。符合要求的拉格朗日量的一般形式为³：

$$\mathcal{L} = A\Phi^0 + B\Phi + C\Phi^2 + D\partial_\mu\Phi + E\partial_\mu\Phi\partial^\mu\Phi + F\Phi\partial_\mu\Phi \quad (6.1)$$

注意我们考虑的是拉格朗日量密度 \mathcal{L} 而不是 L 本身，物理理论从作用量

$$S = \int d\mathbf{x} \mathcal{L} \quad (6.2)$$

² 回忆一下，我们使作用量取极小值进而得到的欧拉-拉格朗日方程就是系统的运动方程。

³ 在 4.2 节中有所讨论：我们只考虑 0, 1 和 2 阶的 Φ 。只考虑最低可能的导数项的原因在后面就会说明。

得出, 其中 $d\mathbf{x}$ 可以理解为对时间和空间的积分。因此, 诸如 $\Phi \partial_\mu \partial^\mu \Phi$ 的项实际是没必要的, 因为它等效于 $\partial_\mu \Phi \partial^\mu \Phi$, 从分部积分就可以直接看出⁴。

不仅如此, Lorentz 不变性要求拉格朗日量必须是标量。因此, 所有奇数阶的 ∂_μ , 比如 $\partial_\mu \Phi$ 是禁止的。值得一提的是, 如果常数, 比如 a, c , 有一个 Lorentz 指标的话, 这说明这些常数是一个 4-矢量, 标定了时空的方向从而破坏了空间的各向同性。常数项实际上可忽略, 也就是 $A = 0$, 因为物理理论从欧拉-拉格朗日方程中导出, 因此一个常数项对运动方程没有影响⁵。除此之外, 我们可以忽略掉 Φ 的线性项, 也就是说 $B = 0$, 因为从欧拉-拉格朗日方程中可以看出这一项只会给运动方程增加一个常数⁶。那么, 剩下的项就是:

$$\mathcal{L} = C\Phi^2 + E\partial_\mu \Phi \partial^\mu \Phi \quad (6.3)$$

本章开始就指明了要建立一个自由场的理论, 这意味着只有一个 Φ 而不会有形如 $\Phi_1 \Phi_2$ 这种相互作用形式, 这样的项在下一章 — 描述相互作用的理论中考虑。

还有一件事情需要注意: 最后留下了两个常数 C 和 E , 利用变分法可以把它俩最终收为一个常数, 因为拉格朗日量整体差一个常数因子并无影响⁷。不仅如此, 通常我们都在拉格朗日量前增加一个 $\frac{1}{2}$, 然后称另外的常数⁸为 $-m^2$ 。因此, 我们最终得到了

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \quad (6.4)$$

现在利用变分法, 也就是说把拉格朗日量带入到欧拉-拉格朗日方程(4.10)式, 我们就得到了运动方程

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right) &= 0 \\ \rightarrow \frac{\partial}{\partial \Phi} \left(\frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \right) - \partial_\mu \left(\frac{\partial}{\partial (\partial_\mu \Phi)} \left(\frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \right) \right) &= 0 \\ \rightarrow (\partial_\mu \partial^\mu + m^2) \Phi &= 0 \end{aligned} \quad (6.5)$$

这就是著名的 **Klein-Gordon 方程**, 是描述自旋 0 的自由场/粒子的方程。

⁴ 分部积分产生的边界项为零, 因为在远处场很小。值得注意的是, 这是由于我们物理里面的速度是有上限的 (第2.3节)。因此, 很远的场对于有限远的 x 位置的场不会有影响。

⁵ 考虑(4.10)式:

$$\frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right) = 0$$

因此 $\mathcal{L} \rightarrow \mathcal{L} + A$ 中的常数项 A 并不改变任何东西: $\frac{\partial(\mathcal{L}+A)}{\partial \Phi} - \partial_\mu \left(\frac{\partial(\mathcal{L}+A)}{\partial (\partial_\mu \Phi)} \right) = \frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right)$ 。

⁶ $\mathcal{L} \rightarrow \mathcal{L} + B\Phi$ 使得 $\frac{\partial(\mathcal{L}+B\Phi)}{\partial \Phi} - \partial_\mu \left(\frac{\partial(\mathcal{L}+B\Phi)}{\partial (\partial_\mu \Phi)} \right) = \frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right) + B = 0$, 只产生运动方程中的可加常数。

⁷ $\mathcal{L} \rightarrow C\mathcal{L}$ 得到

$$\begin{aligned} \frac{\partial(C\mathcal{L})}{\partial \Phi} - \partial_\mu \left(\frac{\partial(C\mathcal{L})}{\partial (\partial_\mu \Phi)} \right) &= 0 \\ \downarrow \\ \frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right) &= 0 \end{aligned}$$

⁸ 后面我们会明白的看到为什么这样称呼这个常数: 它正是拉格朗日量中的粒子质量。

6.2.1 Complex Klein-Gordon Field 复 Klein-Gordon 场

对于自旋 0 的场可以不用标量场的复共轭构造 Lorentz 不变的拉格朗日量。然而对于自旋 $\frac{1}{2}$ 场 Ψ 来说就不一样了，这会带来非常有趣的结果⁹。而且，我们甚至也可以像很多教科书一样研究等价的 Lorentz 不变的拉格朗日量

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi$$

这实际上与考虑两个等质量标量场的相互作用是一样的：

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi_1 \partial^\mu \phi_1 - \frac{1}{2} m^2 \phi_1^2 + \frac{1}{2} \partial_\mu \phi_2 \partial^\mu \phi_2 - \frac{1}{2} m^2 \phi_2^2$$

因为我们有

$$\phi \equiv \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$$

Lorentz 对称性决定了拉格朗日量的形式。基本的标量 (自旋 0) 粒子非常罕见，事实上，只有 Higgs 玻色子是实验上验证了的，但是这个拉格朗日量可以描述复合系统，比如介子。因此，下面不再研究这种拉格朗日量了，而且大多数教科书中也仅仅是用它来作为习题。

6.3 Dirac Equation Dirac 方程

Dirac 的另一个故事是他第一次遇到 Feynman 的时候，在漫长的寂静之后他说：“我有一个方程。你也有吗？”

—Anthony Zee¹⁰

我们这节想要研究自由的自旋 $\frac{1}{2}$ 场/粒子的运动方程。我们会使用 Lorentz 群的 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示，因为要描述宇称变换下的对称性的理论必须同时有 $(\frac{1}{2}, 0)$ 和 $(0, \frac{1}{2})$ 表示¹¹。而在这种对称群表示下做变换的对象称为 Dirac 旋量，包括右手旋量和左手旋量。

3.7.9 节讲过，Dirac 旋量定义为

$$\Psi \equiv \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_a \\ \xi^{\dot{a}} \end{pmatrix} \quad (6.6)$$

现在需要用 Dirac 旋量构造某个 Lorentz 不变量，从而放到拉

⁹ 剧透：每个自旋 $\frac{1}{2}$ 的粒子都有反粒子。用复数场实际上就是同时考虑两个场，接下来会说到。因此，我们强制的令 Lorentz 不变性同时在两个 (紧密联系的) 场上作要求，这通常就解释为粒子和反粒子场。

¹⁰ Anthony Zee.
Quantum Field Theory in a Nutshell. Princeton University Press, 1st edition, 3 2003.
ISBN 9780691010199

¹¹ 我们在第 3.7.9 节中说明了，宇称变换会把 $(\frac{1}{2}, 0)$ 表示变成 $(0, \frac{1}{2})$ 表示

格朗日量里面。第一步是从左手旋量和右手旋量出发构造不变量。

利用在第3.7.7节介绍的 Van-der-Waerden 记号。两种可能的形式分别是¹²:

$$I_1 := \chi_a^T \xi^{\dot{a}} = (\chi_a^*)^T \xi^{\dot{a}} = (\chi_a)^\dagger \xi^{\dot{a}} = (\chi_L)^\dagger \xi_R \quad (6.7)$$

和

$$I_2 := (\xi^a)^T \chi_a = ((\xi^a)^*)^T \chi_a = (\xi_R)^\dagger \chi_L \quad (6.8)$$

3.7.7节讲过, Lorentz 不变量总是要把带点的下标和带点的上标, 或者不带点的下标和不带点的上标结合。这里再次看到右手旋量和左手旋量需要成对出现。

更进一步的, 我们可以构造两种组合, 其涉及到一阶导数且满足 Lorentz 协变, 我们一会儿就会看到。但是首先需要理解旋量的导数意味着什么。3.7.8节讲过如何从旋量构造 4-矢量

$$v_{ab} = v_\nu \sigma_{ab}^\nu$$

其中 v_ν 像 4-矢量一样变换。因此, 旋量表述下的导数算符就是

$$\partial_{ab} = \partial_\nu \sigma_{ab}^\nu \quad (6.9)$$

惯例上会定义 $\bar{\sigma}^0 = I_{2 \times 2}$, $\bar{\sigma}^i = -\sigma^i$, 因此我们可以构建 Lorentz 不变的项

$$I_3 := (\chi_{\dot{a}})^T \partial_\mu (\sigma^\mu)^{\dot{a}b} \chi_b = (\chi_L)^\dagger \partial_\mu \bar{\sigma}^\mu \chi_L \quad (6.10)$$

和

$$I_4 := (\xi^a)^T \partial_\mu (\sigma^\mu)_{ab} \xi^{\dot{b}} = (\xi_R)^\dagger \partial_\mu \sigma^\mu \xi_R \quad (6.11)$$

除了 $(\sigma^\mu)_{ab}$ 之外, 我们还需要 $(\sigma^\mu)^{\dot{a}b}$ 。¹第一个指标必须是不带点的, 第二个指标必须是带点的, 这样才可以和其它旋量指标合

¹² 这里实际上有另外两种可能, 称为 Majorana 质量项。我们已经知道如何使用自旋“度规” ϵ 来把自旋指标升或者降。我们可以写一个 Lorentz 不变的项, 诸如 $\epsilon(\xi_L)^\dagger \xi_L$, 因为 $\epsilon(\xi_L)^\dagger = \xi^a$ 的形式, 有一个上标的右手旋量; 但是这样写的话就会缺少自由度。 χ_R 和 ξ_L 都有两个分量, 因此我们在诸如 $(\chi_R)^\dagger \xi_L$ 的项里面有四个自由度, 而在 $\epsilon(\xi_L)^\dagger \xi_L = \xi^a \xi_a$ 中, 像右手旋量一样变换的东西并不依赖于 ξ_L , 我们从而只有两个自由度。关于 Majorana 旋量实际上还有很多可以说的, 它到现在还在进行 (实验上的) 研究来确定中微子到底属于哪一种。再多说一句, Majorana 旋量是“实”的 Dirac 旋量。这里我打引号的原因是通常实的就是说 $\Psi^* \stackrel{!}{=} \Psi$ 。对于旋量来说, 这个条件并不 Lorentz 协变 (因为 Lorentz 变换的表示是复的)。如果我们强制要求我们的标准要求 $(\Psi^* \stackrel{!}{=} \Psi)$ 的话, 对于换一个参考系这事情一般来说就不对了。然而, 我们还是给一个对 Dirac 旋量的 Lorentz 协变的“实”条件: $\begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix} \Psi^* \stackrel{!}{=} \Psi$, 二者通常诠释为电荷共轭。这种诠释在第7.1.5中会进行解释。因此, Majorana 旋量实际上描述的是一种粒子, 它的电荷共轭粒子, 通常叫反粒子, 就是它自己。Majorana 粒子是自己的反粒子, 而 Majorana 旋量就是 Dirac 旋量再加一个附加条件: $\Psi_M \equiv \begin{pmatrix} \xi_L \\ \epsilon \xi_L^* \end{pmatrix}$ 或 $\Psi_M \equiv \begin{pmatrix} -\epsilon \chi_R^* \\ \chi_R \end{pmatrix}$

¹原文为 $(\sigma^\mu)^{\dot{a}b}$, 译文已按勘误表改正。

适的组合。我们通过作用旋量度规两次得到 $(\sigma^\mu)^{ab}$:

$$\begin{aligned}
 (\sigma^\mu)^{ab} &= ((\sigma^\mu)^T)^{ba} = \epsilon^{\dot{b}\dot{c}}((\sigma^\mu)^T)_{\dot{c}d}(\epsilon^{ad})^T \\
 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\sigma^\mu)^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^T \\
 &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\sigma^\mu)^T \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \bar{\sigma}_\mu
 \end{aligned} \tag{6.12}$$

比如, 对 σ_3 来说,

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{=\sigma_3^T} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{=\bar{\sigma}_3 = -\sigma_3}$$

千万别因为 ∂_μ 只作用在一个旋量上就困惑, 因为我们要在拉格朗日量中使用这些不变量, 它们总要积分¹³, 因此我们可以进行分部积分得到其它的可能的形式。因此这里的选择是不加约束的¹⁴。

如果我们定义矩阵¹⁵

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix} \rightarrow \gamma_\mu = \begin{pmatrix} 0 & \bar{\sigma}_\mu \\ \sigma_\mu & 0 \end{pmatrix} \tag{6.13}$$

我们可以用 Dirac 旋量的形式来写我们刚刚发现的不变量

$$\Psi^\dagger \gamma_0 \Psi \text{ 和 } \Psi^\dagger \gamma_0 \gamma^\mu \partial_\mu \Psi, \tag{6.14}$$

因为

$$\Psi^\dagger \gamma_0 \Psi = \begin{pmatrix} (\chi_L)^\dagger & (\xi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \underbrace{(\chi_L)^\dagger \bar{\sigma}_0 \xi_R}_{=I_1} + \underbrace{(\xi_R)^\dagger \sigma_0 \chi_L}_{=I_2}$$

这与我们之前找到的两个不变量完全一致¹⁶, 并且

$$\begin{aligned}
 \Psi^\dagger \gamma_0 \gamma^\mu \partial_\mu \Psi &= \begin{pmatrix} (\xi_L)^\dagger & (\chi_R)^\dagger \end{pmatrix} \begin{pmatrix} 0 & \bar{\sigma}_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^\mu \partial_\mu \\ \sigma^\mu \bar{\partial}_\mu & 0 \end{pmatrix} \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix} \\
 &= \underbrace{(\xi_L)^\dagger \bar{\sigma}_0 \sigma^\mu \partial_\mu \chi_R}_{=I_3} + \underbrace{(\chi_R)^\dagger \sigma_0 \sigma^\mu \partial_\mu \xi_L}_{=I_4}
 \end{aligned}$$

¹³ 我们得到运动方程是从作用量来的, 作用量就是拉格朗日量的积分。

¹⁴ 我们得到对应的运动方程的时候会看的更清楚。我们不论在哪里放 ∂_μ 我们都会得到相同的运动方程, 当然我们也可以把两中可能的组合都放到拉格朗日量里面去。这就会使得式子更长, 但是并没有什么新的东西。

¹⁵ 我们得到有下标的矩阵通过使用矢量指标的升降 $\gamma_\mu = \eta_{\mu\nu} \gamma^\nu = \eta_{\mu\nu} \begin{pmatrix} 0 & \sigma^\nu \\ \bar{\sigma}^\nu & 0 \end{pmatrix} = \begin{pmatrix} 0 & \eta_{\mu\nu} \sigma^\nu \\ \eta_{\mu\nu} \bar{\sigma}^\nu & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix}$, 因为 $\eta_{\mu\nu} \sigma^\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \sigma^0 \\ \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix} = \begin{pmatrix} \sigma^0 \\ -\sigma^1 \\ -\sigma^2 \\ -\sigma^3 \end{pmatrix} = \bar{\sigma}_\mu$

¹⁶ σ_0 只是单位矩阵而已, 因此 $\bar{\sigma}_0 = \sigma_0$ 。

就想说过的那样给出了另外两个不变量¹⁷。通常引入记号²

$$\bar{\Psi} := (\Psi)^\dagger \gamma_0 \quad (6.15)$$

现在可以写出利用 Dirac 旋量构造的 Lorentz 不变的拉格朗日量的完整形式了，它符合 4.2 节的所有限制³：

$$\mathcal{L} = A\Psi^\dagger \gamma_0 \Psi + B\Psi^\dagger \gamma_0 \gamma^\mu \partial_\mu \Psi = A\bar{\Psi}\Psi + B\bar{\Psi}\gamma^\mu \partial_\mu \Psi$$

带入常数 ($A = -m, B = i$)，我们得到了完整的 **Dirac** 拉格朗日量

$$\mathcal{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma^\mu \partial_\mu \Psi = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \quad (6.16)$$

注意 $\mathcal{L}_{\text{Dirac}}$ 里面实际上有两个不同的场，因为 Ψ 是复的¹⁸。不然就没法 Lorentz 不变了。更准确的说，我们有

$$\Psi = \Psi_1 + i\Psi_2$$

其中 Ψ_1, Ψ_2 是实数场。通常我们使用两个不同的复数场 Ψ 和 $\bar{\Psi}$ 比两个实数场要更方便。把这个拉格朗日量代入到欧拉-拉格朗日方程（方便起见，这里再写一遍），

$$\frac{\partial \mathcal{L}}{\partial \Psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Psi)} \right) = 0 \quad (6.17)$$

我们有

$$-m\bar{\Psi} - i\partial_\mu \bar{\Psi} \gamma^\mu = 0 \rightarrow (i\partial^\mu \bar{\Psi} \gamma_\mu + m\bar{\Psi}) = 0$$

同样将场 $\bar{\Psi}$ 带入欧拉-拉格朗日方程

$$\frac{\partial \mathcal{L}}{\partial \bar{\Psi}} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \bar{\Psi})} \right) = 0$$

¹⁷ 注意 $\sigma^\mu \partial_\mu = \partial_\mu \sigma^\mu$ ，因为 σ^μ 仅仅是常数矩阵而已。

¹⁸ 因此左手旋量和右手旋量也都是复的。

² 译者注：在一些其它教科书里，一般也会写成 $\bar{\Psi}$ 。

³ 译者注：这句话太绕口了。直接看英文会好点：Now we have everything we need to construct a Lorentz-invariant Lagrangian that is in agreement with the restrictions discussed in Sec. 4.2 using Dirac spinors.

我们有

$$(i\gamma_\mu \partial^\mu - m)\Psi = 0 \quad (6.18)$$

这就是著名的 **Dirac 方程**。注意我们对拉格朗日量进行分部积分也会得到完全一样的东西

$$-m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu \partial^\mu \Psi \underbrace{=}_{\text{在作用量积分中分部积分}} -m\bar{\Psi}\Psi - (i\partial^\mu \bar{\Psi})\gamma_\mu \Psi$$

然后利用欧拉-拉格朗日方程，

$$\rightarrow -m\Psi + i\partial_\mu \gamma^\mu \Psi = 0$$

¹⁹ 你完全可以使用更长的拉格朗日量来包含更多的可能的项，但事实上结果没有区别。

因此拉格朗日量的形式是没有约束的，除了它的反对称性的要求¹⁹。

6.4 Proca Equation Proca 方程

现在想找那些根据 Lorentz 群的 $(\frac{1}{2}, \frac{1}{2})$ 表示变换的对象的运动方程。前面已经看到这个表示是矢量表示，因此这件事就比较简单了。考虑一个任意的矢量场 A_μ ，然后构造所有符合第4.2节约束的 Lorentz 不变的项。我们必须结合下标和上标，因为我们在第2.4节定义的 Minkowski 空间的标量积是这样的，而标量才是拉格朗日量里面需要的。可能的不变量有²⁰

$$\begin{aligned} I_1 &= \partial^\mu A^\nu \partial_\mu A_\nu, & I_2 &= \partial^\mu A^\nu \partial_\nu A_\mu \\ I_3 &= A^\mu A_\mu, & I_4 &= \partial^\mu A_\mu \end{aligned}$$

而拉格朗日量就写为

$$\mathcal{L}_{\text{Proca}} = C_1 \partial^\mu A^\nu \partial_\mu A_\nu + C_2 \partial^\mu A^\nu \partial_\nu A_\mu + C_3 A^\mu A_\mu + C_4 \partial^\mu A_\mu \quad (6.19)$$

我们可以忽略最后一项 $\partial^\mu A_\mu$ ，因为根据欧拉-拉格朗日方程，它只在运动方程里面提供常数项 C_4 ，并不产生什么影响。因此，一阶的 ∂_μ 项是平凡的。

利用欧拉-拉格朗日方程分别计算每个场分量的运动方程，

$$\frac{\partial \mathcal{L}}{\partial A_\rho} = \partial_\sigma \left(\frac{\partial \mathcal{L}}{\partial (\partial_\sigma A_\rho)} \right)$$

²⁰ 由于可以分部积分，我们再一次扔掉了 $\partial_\mu \partial^\mu A^\nu A_\nu$ 这项，而仅留着 $\partial^\mu A^\nu \partial_\mu A_\nu$ 。

需要非常仔细的处理这些指标。我们先看一下欧拉-拉格朗日方程的右边，并计算与 C_1 有关的项：

$$\begin{aligned}
 & \partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma A_\rho)} (C_1 \partial^\mu A^\nu \partial_\mu A_\nu) \right) \\
 & \stackrel{\text{product rule}}{=} C_1 \partial_\sigma \left((\partial_\mu A_\nu) \frac{\partial(\partial^\mu A^\nu)}{\partial(\partial_\sigma A_\rho)} + (\partial^\mu A^\nu) \frac{\partial(\partial_\mu A_\nu)}{\partial(\partial_\sigma A_\rho)} \right) \\
 & \stackrel{\text{用度规降指标}}{=} C_a \partial_\sigma \left((\partial_\mu A_\nu) g^{\mu\kappa} g^{\nu\lambda} \frac{\partial(\partial_\kappa A_\lambda)}{\partial(\partial_\sigma A_\rho)} + (\partial^\mu A^\nu) \frac{\partial(\partial_\mu A_\nu)}{\partial(\partial_\sigma A_\rho)} \right) \\
 & = C_a \partial_\sigma \left((\partial_\mu A_\nu) g^{\mu\kappa} g^{\nu\lambda} \delta_\kappa^\sigma \delta_\lambda^\rho + (\partial^\mu A^\nu) \delta_\mu^\sigma \delta_\nu^\rho \right) \\
 & = C_1 \partial_\sigma (\partial^\sigma A^\rho + \partial^\sigma A^\rho) = 2C_1 \partial_\sigma \partial^\sigma A^\rho
 \end{aligned}$$

类似的，我们可以计算出

$$\partial_\sigma \left(\frac{\partial}{\partial(\partial_\sigma A^\rho)} (C_2 \partial^\mu A^\nu \partial_\nu A_\mu) \right) = 2C_2 \partial^\rho (\partial_\sigma A^\sigma)$$

因此，(6.19)的拉格朗日量得到的运动方程为

$$2C_3 A^\rho = 2C_1 \partial_\sigma \partial^\sigma A^\rho + 2C_2 \partial^\rho (\partial_\sigma A^\sigma)$$

带入惯例的常数得到

$$\rightarrow m^2 A^\rho = \frac{1}{2} \partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) \quad (6.20)$$

这被称为 **Proca 方程组**²¹，这是有质量自旋为 1 的粒子的波动方程。对于无质量 ($m = 0$) 自旋为 1 的粒子，即光子，方程为

$$\rightarrow 0 = \frac{1}{2} \partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) \quad (6.21)$$

这就是没有电流情况下的非齐次 **Maxwell 方程组**。通常我们定义电磁张量

$$F^{\sigma\rho} := \partial^\sigma A^\rho - \partial^\rho A^\sigma \quad (6.22)$$

而非齐次 Maxwell 方程组则是

$$\partial_\rho F^{\sigma\rho} = 0 \quad (6.23)$$

²¹ 称为方程组是因为对每一个分量 ρ 都有一个方程

注意我们可以对无质量的自旋 1 场重新写拉格朗日量

$$\mathcal{L} = \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu)$$

为

$$\begin{aligned}\mathcal{L}_{\text{Maxwell}} &= \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &= \frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) \\ &= \underbrace{\frac{1}{4}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu A_\mu + \partial^\mu A^\nu \partial_\mu A_\nu)}_{\text{对哑指标重新命名}} \\ &= \frac{1}{4}(2\partial^\mu A^\nu \partial_\mu A_\nu - 2\partial^\mu A^\nu \partial_\nu A_\mu) \\ &= \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) \quad \checkmark\end{aligned}\tag{6.24}$$

$\mathcal{L}_{\text{Maxwell}} = \frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ 是我们惯用的拉格朗日量的形式。等价的，有质量的自旋 1 的场的拉格朗日量可以写成

$$\mathcal{L}_{\text{Proca}} = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + m^2 A_\mu A^\mu = \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) + m^2 A_\mu A^\mu\tag{6.25}$$

本章导出了描述自由场/粒子们的运动方程。我们想知道这些方程能预言什么实验现象，但在此之前需要更多的方程才行，因为实验永远是涉及相互作用的。比如只能通过其他粒子，比如光子，来测量电子。因此，下一章将推导描述不同的场/粒子之间相互作用的拉格朗日量。

7. Interaction Theory 相互作用理论

Summary 总结

在这一章中我们将导出不同的场之间的相互作用。例如描述电子是如何和光子作用的¹。

内禀对称性，或者这里叫做规范² 对称性，能指引我们得到拉格朗日量的正确形式。我们将从局域³ $\mathcal{U}(1)$ 对称性出发来得到拉格朗日量

$$\mathcal{L} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_{\mu}\partial^{\mu}\Psi + A_{\mu}\bar{\Psi}\gamma^{\mu}\Psi + \partial^{\mu}A^{\nu}\partial_{\mu}A_{\nu} - \partial^{\mu}A^{\nu}\partial_{\nu}A_{\mu},$$

即量子电动力学 (quantum electrodynamics) 的拉格朗日量。这个拉格朗日量描述了有电荷有质量的场和无质量自旋为 1 的场 (光子场) 之间的相互作用。只有在排除了形如 $mA_{\mu}A^{\mu}$ 的“质量项”后 (这和用 A_{μ} 来描述的光子是无质量的实验事实相符)，拉格朗日量才是局域 $\mathcal{U}(1)$ 不变的¹。利用 Noether 定理，我们能从 $\mathcal{U}(1)$ 对称性中导出一个新的守恒量，它一般称为电荷 (electric charge)。

接下来是 $SU(2)$ 对称性。引入一个二分量的二重态 (doublet)

$$\bar{\Psi} := \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix},$$

¹译注：应该为“只有这样才是且仅是”。

¹ 从另外一个视角看：电子 (自旋 $\frac{1}{2}$ 且有质量) 场如何和光子 (自旋为 1) 场作用。

² 马上就会解释这个奇怪的名字

³ 这意味着不仅仅是一个 $e^{i\alpha}$ ，而是在每一个时空点上作用一个不同的因子，数学上写作 $e^{i\alpha(x)}$ 。或者这样说：变换参数 $\alpha = \alpha(x)$ 现在是 x 的函数，在不同的时空点上有不同的值。

这样一个二重态场中包含了两个自旋 $\frac{1}{2}$ 场，例如，电子和电子中微子场在 $SU(2)$ 的“旋转”变换下相互转换。

我们可以使用二重态记号写下局域 $SU(2)$ 不变的拉格朗日量⁴

$$\mathcal{L} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + \bar{\Psi}\gamma_\mu\sigma_j W_j^\mu\Psi - \frac{1}{2}(W_{\mu\nu})_i(W^{\mu\nu})_i,$$

其中包含了三个自旋为 1 的场 W_j^μ 。由于 $SU(2)$ 群的生成元有三个基 $J_i = \sigma_i/2$ ，所以我们需要三个场来保证拉格朗日量是局域 $SU(2)$ 不变的。我们将看到局域 $SU(2)$ 对称性仅当形如 $m\bar{\Psi}\Psi$ 、 $mW_\mu W^\mu$ 的质量项（其中质量 m 是任意给定矩阵）不存在时才有可能实现，这是因为 Ψ 现在是二分量对象了。所以此时不仅自旋为 1 的场 W_j^μ 得是无质量的了，自旋 $\frac{1}{2}$ 场也一样。另外一种可能是两个自旋 $\frac{1}{2}$ 等质量场，但是它被实验否决了：电子质量远大于电子中微子质量。除此之外，我们从实验知道三个自旋为 1 的场 W_j^μ 不是无质量的。这常解释成 $SU(2)$ 对称性被破缺了。

这是随后被引入的 **Higgs 机制 (Higgs formalism)** 的想法来源。这个机制能使我们得到一个包含质量项的 $SU(2)$ 不变的拉格朗日量。它通过引入一个与零自旋场，即 Higgs 场，的相互作用来达成目标。相同的机制使我们能给自旋 $\frac{1}{2}$ 场加上任意的质量项以与实验相符。最后的相互作用拉格朗日量描述了一种新的相互作用：**弱相互作用 (weak interaction)**，由三个⁵有质量自旋为 1 的场： W^+ 、 W^- 和 Z ，作为传播媒介。利用 Noether 定理，我们能从 $SU(2)$ 对称性得到一个新的守恒量：**同位旋 (isospin)**，它相当于于电磁相互作用中的电荷，是弱相互作用中的荷。

最后，我们将考虑内禀 $SU(3)$ 对称，它将会将我们能描述一类新的相互作用，**强相互作用 (strong interaction)**，的拉格朗日量。为此我们将引入一个三重态

$$Q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix},$$

在 $SU(3)$ 下变换，包含三个自旋 $\frac{1}{2}$ 场。这三个自旋 $\frac{1}{2}$ 场被解释为有不同**颜色 (color)** 的**夸克 (quarks)**，这是电磁作用中的电荷、或者弱作用中的同位旋在强相互作用中的对应物。一样，质量项是

⁴ $W_j^{\mu\nu}$ 这玩意会通过三个 W_j^μ 定义，就像从 $U(1)$ 规范场 A^μ 定义 $F_{\mu\nu}$ 一样。

⁵ 准确的讲：我们从 $U(1) \otimes SU(2)$ 对称破缺到 $U(1)$ 。这个过程产生了三个有质量矢量 bosons: W^+ 、 W^- 、 Z ，和一个无质量矢量 bosons: γ 光子。所以说为什么常讲电磁学（来自 $U(1)$ 对称性）和弱相互作用理论（来自 $SU(2)$ ）是统一的。一开始的 $U(1)$ 对称性和最后留下来的 $U(1)$ 对称性不是一个东西。光子和 Z -Bosons 可以看做是两个矢量 Bosons 的线性组合，常记做 B 和 W^3 ，用以使拉格朗日量具有局域 $U(1)$ (B -bosons) 和 $SU(2)$ (W^1 、 W^2 、 W^3 -bosons) 不变性。

被禁止的，但这次和实验结论一致：8 个相应的 bosons⁶，称为胶子，是无质量的⁷。根据实验我们知道只有夸克（自旋 $\frac{1}{2}$ ）和胶子（自旋为 1）携带颜色。最后，拉格朗日量的结果是

$$\mathcal{L} = -\frac{1}{4}F_{\alpha\beta}^A F_A^{\alpha\beta} + \bar{Q}(iD_\mu\gamma^\mu - m)Q,$$

咱仅仅引用一下装个逼，因为推导太繁琐了，并且和我们之前做过的完全类似。

总结一下：

$U(1)$	\rightarrow	1 个规范场	\rightarrow	无质量光子	\rightarrow	电荷
$SU(2)$	\rightarrow	3 个规范场	\rightarrow	有质量 W- 和 Z-bosons（需要 Higgs）	\rightarrow	同位旋
$SU(3)$	\rightarrow	8 个规范场	\rightarrow	无质量胶子	\rightarrow	色荷

⁶ 数字 8 源于 $SU(3)$ 的生成元有 8 个基。

⁷ 作为补充，我们从实验知道 $SU(3)$ 的三重态中的场有相同的质量。这是个好消息，因为局域 $SU(3)$ 对称性禁止形如 $m\bar{Q}Q$ 这样带有任意质量矩阵 m 的项存在，但允许 $\bar{Q}\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}Q$ 这样一项，这代表三重态里的项有相同的质量。由此局域 $SU(3)$ 不变性不会在拉格朗日量的质量项上造成新的障碍， $SU(3)$ 对称也不没有被破缺。

7.1 $U(1)$ Interaction $U(1)$ 相互作用

为了得到拉格朗日量中正确的相互作用项，我们得使用内禀对称性，或者称作规范对称性 (gauge symmetries)。规范对称这个叫法有一些历史上的原因，和我们现在要讨论的东西之间关系不是太大。Weyl 曾尝试将电磁学⁸“作为时空对称性的结果，特别是在局域尺度变换下的对称性”。将其称作规范对称性是有原因的，因为它意味着，例如，我们能任意改变用于定义长度标准的铂金尺（用来规范²实验中测长的仪器），而不改变物理。这个尝试没有成功，但稍晚些时候，Weyl 找到了能正确导出电磁学的对称性，并仍用这个词来命名。

⁸Frank Wilczek.

Riemann-einstein structure from volume and gauge symmetry. Phys. Rev. Lett., 80:4851–4854, Jun 1998. doi: 10.1103/PhysRevLett.80.4851

7.1.1 自旋 $\frac{1}{2}$ 自由场的内禀对称性

再来看一眼用来导出自由自旋 $\frac{1}{2}$ 理论的拉格朗日量((6.16)式)

$$\mathcal{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi = \bar{\Psi}(i\gamma_\mu\partial^\mu - m)\Psi \quad (7.1)$$

它是在 Lorentz 对称性的要求下导出的，但如果我们仔细观察它，能发现这个拉格朗日量中蕴含着另一个对称性。拉格朗日量在场

²译注：没有规矩，不成方圆。线长是某个线性空间中的某种范数，将改变线长的定义的变换，称作规范变换，应当是一个合理的称呼。关于 gauge 这个词的命名以及为何翻译成“规范”，可以参见曹则贤. Norm and Gauge[J]. 物理, 2013, 42(11): 815-819.

Ψ 的如下变换下不变

$$\begin{aligned}\Psi &\rightarrow \Psi' = e^{ia}\Psi \\ \Rightarrow \bar{\Psi} &\rightarrow \bar{\Psi}' = \Psi'^{\dagger}\gamma_0 = (e^{ia}\Psi)^{\dagger}\gamma_0 = \bar{\Psi}e^{-ia},\end{aligned}\quad (7.2)$$

⁹ 别忘了 $\bar{\Psi} = \Psi^{\dagger}\gamma_0$

其中负号来源于复共轭⁹，而 a 是一个任给实数。为了看清这一点，我们将变换后的场的拉格朗日量明确写出

$$\begin{aligned}\mathcal{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_{\mu}\partial^{\mu}\Psi' \\ &= -m(\bar{\Psi}e^{-ia})(e^{ia}\Psi) + i(\bar{\Psi}e^{-ia})\gamma_{\mu}\partial^{\mu}(e^{ia}\Psi) \\ &= -m\bar{\Psi}\Psi \underbrace{e^{-ia}e^{ia}}_{=1} + i\bar{\Psi}\gamma_{\mu}\partial^{\mu}\Psi \underbrace{e^{-ia}e^{ia}}_{=1} \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_{\mu}\partial^{\mu}\Psi = \mathcal{L}_{\text{Dirac}},\end{aligned}\quad (7.3)$$

¹⁰ 技术上讲：一个复数和所有矩阵（例如 γ_{μ} ）对易。

其中我们利用了 e^{ia} 仅是一个复数这一点来将它自由前后挪动¹⁰。回忆我们在第3章中曾讲过的，任何单位复数（能写作 e^{ia} 的形式）构成了一个 $\mathcal{U}(1)$ 群，由此可以将我们刚才发现的事实用数学语言表达出来：拉格朗日量是 $\mathcal{U}(1)$ 不变的。由于它丝毫不涉及时空的变换而只改变场自身，所以说这是一个内禀对称性。猛地一看，这玩意还挺萌的，但好像没什么用。不要走开，稍后的节目会告诉你它有多精彩！

现在咱再来钻深一点。我们刚展示了，给我们的场乘一个任意单位复数什么都不会改变。这个对称变换 $\Psi \rightarrow \Psi' = e^{ia}\Psi$ 叫做**全局 (global)** 变换，因为我们在每一个点 x 给场 $\Psi = \Psi(x)$ 乘上了一个相同的因子 e^{ia} 。

但是，为什么一个时空点的相位因子会和另一个时空点的相关呢？一个点的选取不应该瞬间改变整个宇宙的情况。这是挺奇怪的，因为正如2.4节中所述，狭义相对论告诉我们没有信息能比光传播得更快。而全局对称性选择会瞬间应用到在整个宇宙中的每一个点上³。

让我们来检查一下，如果给每一个时空点变换一个不同的因子 $a = a(x)$ ，拉格朗日量是不是还是不变的。这叫做**局域 (local)** 变换。

³ 译注：引入局域规范的另一种讲法是，我们不能保证外星人里物理学家和我们选取同一套规范（这里的 a 是可以随意选取的，但是必须要选取一个才能得到完整的理论。 a 的选取被称为规范的选取），但是物理定律不应该有区别。

作变换

$$\begin{aligned}\Psi &\rightarrow \Psi' = e^{ia(x)}\Psi \\ \Rightarrow \bar{\Psi} &\rightarrow e^{-ia(x)}\bar{\Psi},\end{aligned}\quad (7.4)$$

现在 $a = a(x)$ 依赖于坐标了, 我们得到变换后的拉格朗日量¹¹

$$\begin{aligned}\mathcal{L}'_{\text{Dirac}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= -m(\bar{\Psi}e^{-ia(x)})\underbrace{(e^{ia(x)}\Psi)}_{=1} + i(\bar{\Psi}e^{-ia(x)})\gamma_\mu\partial^\mu(e^{ia(x)}\Psi) \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi \underbrace{e^{-ia}e^{ia}}_{=1 \text{ 莱布尼兹律}} + i(e^{-ia(x)}\bar{\Psi})\gamma_\mu\Psi(\partial^\mu e^{ia(x)}) \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i^2(\partial^\mu a(x))\bar{\Psi}\gamma_\mu\Psi \neq \mathcal{L}_{\text{Dirac}}\end{aligned}\quad (7.5)$$

由于莱布尼兹律导致了一个额外的项, 所以说我们的拉格朗日量在局域 $U(1)$ 对称变换下不是不变的。但按照上面所讨论的, 拉格朗日量应该是局域不变的, 在这里却不是。这里有一些事情可做, 但我们得先研究另一种对称性。

¹¹ 可能会好奇包含全部可能导数项的拉格朗日量 (由简洁性的要求被忽略) 是不是局域 $U(1)$ 不变的: $\mathcal{L} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i(\partial^\mu\Psi)\gamma_\mu\Psi$ 。你可以检查一下, 这个拉格朗日量自然是局域 $U(1)$ 不变的, 但注意第二项和第三项之和为零: $i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i(\partial^\mu\Psi)\gamma_\mu\Psi = \underbrace{i(\partial^\mu\Psi)\gamma_\mu\Psi}_{\text{分部积分}} - i(\partial^\mu\Psi)\gamma_\mu\Psi = 0$ 。包含了全部可能导数的正确拉格朗日量的两导数项间必须有一个相对负号: $\mathcal{L}_{\text{Dirac}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - i(\partial^\mu\Psi)\gamma_\mu\Psi$, 故它并不是局域 $U(1)$ 不变的。

7.1.2 自旋为 1 自由场的内禀对称性

接下来, 来看看自旋为一的粒子¹²

$$\mathcal{L}_{\text{Proca}} = \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu + m^2 A_\mu A^\mu. \quad (7.6)$$

这儿也能找到一个全局内禀对称性。如果我们作变换

$$A_\mu \rightarrow A'_\mu = A_\mu + a_\mu \quad (7.7)$$

其中 a_μ 为任意常数, 拉格朗日量变为

$$\begin{aligned}\mathcal{L}'_{\text{Proca}} &= \partial^\mu A'^\nu \partial_\mu A'_\nu - \partial^\mu A'^\nu \partial_\nu A'_\mu + m^2 A'^\mu A'_\mu \\ &= \partial^\mu (A^\nu + a^\nu) \partial_\mu (A_\nu + a_\nu) - \partial^\mu (A^\nu + a^\nu) \partial_\nu (A_\mu + a_\mu) + m^2 (A_\mu + a_\mu) (A^\mu + a^\mu) \\ &= \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu + m^2 (A_\mu + a_\mu) (A^\mu + a^\mu)\end{aligned}\quad (7.8)$$

¹² 见(6.25)式, 注意我们为了简明起见而丢掉的 $\frac{1}{2}$ 因子

由此我们作出结论，对于无质量场（ $m = 0$ ），这个变换是个全局对称的变换。

那局域对称性又会怎么样呢？作变换

$$A_\mu \rightarrow A'_\mu = A_\mu + a_\mu(x) \quad (7.9)$$

无质量的拉格朗日量变成

$$\begin{aligned} \mathcal{L}'_{\text{Maxwell}} &= \partial^\mu A'^\nu \partial_\mu A'_\nu - \partial^\mu A'^\nu \partial_\nu A'_\mu \\ &= \partial^\mu (A^\nu + a^\nu(x)) \partial_\mu (A_\nu + a_\nu(x)) - \partial^\mu (A^\nu + a^\nu(x)) \partial_\nu (A_\mu + a_\mu(x)) \\ &= \partial^\mu A^\nu \partial_\mu A_\nu + \partial^\mu a^\nu(x) \partial_\mu A_\nu + \partial^\mu A^\nu \partial_\mu a_\nu(x) + \partial^\mu a^\nu(x) \partial_\mu a_\nu(x) \\ &\quad - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu a_\mu(x) - \partial^\mu a^\nu(x) \partial_\nu A_\mu - \partial^\mu a^\nu(x) \partial_\nu a_\mu(x) \end{aligned} \quad (7.10)$$

即没有内禀对称性。

然而，如果作变换 $A_\mu \rightarrow A'_\mu + \partial_\mu a(x)$ ，我们能弄出一个局域内禀对称性来。这意味着我们加的是一个任意函数的导数，而不是一个任意函数。这会导致⁴

$$\begin{aligned} \mathcal{L}'_{\text{Maxwell}} &= \partial^\mu A'^\nu \partial_\mu A'_\nu - \partial^\mu A'^\nu \partial_\nu A'_\mu \\ &= \partial^\mu (A^\nu + \partial^\nu a(x)) \partial_\mu (A_\nu + \partial_\nu a(x)) \\ &\quad - \partial^\mu (A^\nu + \partial^\nu a(x)) \partial_\nu (A_\mu + \partial_\mu a(x)) \\ &= \partial^\mu A^\nu \partial_\mu A_\nu + \partial^\mu (\partial^\nu a(x)) \partial_\mu A_\nu + \partial^\mu A^\nu \partial_\mu \partial_\nu a(x) \\ &\quad + \partial^\mu \partial^\nu a(x) \partial_\mu \partial_\nu a(x) \\ &\quad - \partial^\mu A^\nu \partial_\nu A_\mu - \partial^\mu A^\nu \partial_\nu (\partial_\mu a(x)) - \partial^\mu (\partial^\nu a(x)) \partial_\nu A_\mu \\ &\quad - \partial^\mu (\partial^\nu a(x)) \partial_\nu (\partial_\mu a(x)) \\ &\quad \underbrace{=}_{\partial_\nu \partial_\mu = \partial_\mu \partial_\nu \text{ 以及 重命名哑指标}} \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu = \mathcal{L}_{\text{Maxwell}} \end{aligned} \quad (7.11)$$

所以说这里的确有一个局域内禀对称变换。这看上去也像是一个技术细节。好的，我们找到了一些局域的、内禀的对称性，然后呢？

7.1.3 把奇怪的东西堆一起

总结一下咱现在都干了啥：

- 我们发现自旋 $\frac{1}{2}$ 自由场的拉格朗日量有一个全局的内禀对

⁴译注：原文缺少一个 (x)

称性 $\Psi \rightarrow \Psi' = e^{ia}\Psi$ 。或者这样说：自旋 $\frac{1}{2}$ 自由场的拉格朗日量在全局 $\mathcal{U}(1)$ 变换下不变。

- 这个对称性不是局域的（尽管它应当是），因为在 $a = a(x)$ 的情况下，我们在拉格朗日量中得到如下形式的额外项（(7.5)式）

$$-(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi. \quad (7.12)$$

换言之：拉格朗日量不是局域 $\mathcal{U}(1)$ 不变的。

- 上一节中我们发现了无质量自旋为一的场的一个局域内禀对称性

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu a(x), \quad (7.13)$$

它仅当我们加上的是一个任意函数的导数 $\partial_\mu a(x)$ ，而不是一个任意函数 $a_\mu(x)$ 时，才是一个局域对称性。

这两坨奇怪的东西看上去真该加在一起：拉格朗日量中的一个额外项 $A_\mu\bar{\Psi}\gamma^\mu\Psi$ 会变换作

$$A_\mu\bar{\Psi}\gamma^\mu\Psi \rightarrow (A_\mu + \partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi = A_\mu\bar{\Psi}\gamma^\mu\Psi + \partial_\mu a(x)\bar{\Psi}\gamma^\mu\Psi \quad (7.14)$$

比较 (7.12) 式中的第二项⁵。拉格朗日量中这个将 Ψ 、 $\bar{\Psi}$ 和 A_μ 耦合在一起的新项，正好能干掉那个不自量力阻挡自旋 $\frac{1}{2}$ 自由场拥有局域 $\mathcal{U}(1)$ 不变性的家伙。换言之：通过添加这个新项使得拉格朗日量局域 $\mathcal{U}(1)$ 不变。

再来仔细研究一下这件事。首先习惯上都会在局域 $\mathcal{U}(1)$ 变换的指数上添一个常数因子 g ： $e^{iga(x)}$ 。这样额外项变成

$$-(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi \rightarrow -g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi. \quad (7.15)$$

正如我们下面将看到的，这个额外的因子 g 代表一个任意的耦合常数¹³ g 。将这个新项

$$gA_\mu\bar{\Psi}\gamma^\mu\Psi$$

加到自旋 $\frac{1}{2}$ 自由场的拉格朗日量上，其中 γ^μ 用来保证这一项

¹³ 耦合常数总能告诉我们相互作用的强度是多少。这里讨论的是电磁相互作用， g 代表着它的强度。

⁵译注：此处疑引用有误

的 Lorentz 不变性, 否则带有一个自由指标 μ 的 A_μ 和拉格朗日量密度这个标量无法匹配, 另外也插入了耦合常数 (coupling constant)¹⁴ g 。这样拉格朗日量

$$\mathcal{L}_{\text{Dirac+ 额外项}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi。$$

¹⁴ 我们能看到这里 g 决定了 Ψ 、 $\bar{\Psi}$ 和 A_μ 耦合在一起的强度。

¹⁵ Ψ 、 $\bar{\Psi}$ 和 A_μ 的联合变换被称作 $\mathcal{U}(1)$ 规范变换。

依据 Ψ 、 $\bar{\Psi}$ 和 A_μ 局域变换的规则来变换这个拉格朗日量, 得到¹⁵

$$\begin{aligned}\mathcal{L}'_{\text{Dirac+ 额外项}} &= -m\bar{\Psi}'\Psi' + i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' + gA'_\mu\bar{\Psi}'\gamma^\mu\Psi' \\ &\stackrel{\text{见(7.5)式}}{=} -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi + gA'_\mu\bar{\Psi}'\gamma^\mu\Psi' \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi \\ &\quad + g(A_\mu + \partial_\mu a(x))(e^{-iga(x)}\bar{\Psi})\gamma^\mu(e^{iga(x)}\Psi) \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \underline{g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi} \\ &\quad + gA_\mu\bar{\Psi}\gamma^\mu\Psi + \underline{g(\partial_\mu a(x))\bar{\Psi}\gamma^\mu\Psi} \\ &= -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi = \mathcal{L}_{\text{Dirac+ 额外项}}。 \tag{7.16}\end{aligned}$$

由此, 在加上一个额外项后, 我们得到了, 正如所期待的, 一个局域 $\mathcal{U}(1)$ 不变的拉格朗日量。为了描述一个包含了包含有质量自旋 $\frac{1}{2}$ 场和无质量自旋为一的场的系统, 我们必须将无质量自旋为一的自由场的拉格朗日量也加到拉格朗日量中来。这样就得到了完整的拉格朗日量

$$\mathcal{L}_{\text{Dirac+ 额外项 + Maxwell}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi + \partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu。 \tag{7.17}$$

为了方便起见, 引入记号

$$D_\mu \equiv i\partial_\mu + gA_\mu, \tag{7.18}$$

称作协变导数 (covariant derivative)⁶。这样拉格朗日量写作

$$\mathcal{L}_{\text{Dirac+ 额外项 + Maxwell}} = -m\bar{\Psi}\Psi + \bar{\Psi}\gamma_\mu \underbrace{(i\partial^\mu + gA^\mu)}_{\equiv D^\mu} \Psi + \partial^\mu A^\nu\partial_\mu A_\nu$$

⁶译注: 常见的记号似乎与其差了一个虚数单位

$$= -\partial^\mu A^\nu \partial_\nu A_\mu - m \bar{\Psi} \Psi + \bar{\Psi} \gamma_\mu D^\mu \Psi + \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu \quad (7.19)$$

这即为电磁场的量子场论（常称为量子电动力学）的正确拉格朗日量。我们通过观察自旋 $\frac{1}{2}$ 自由场和自旋为一的自由场的内禀对称性，简单的得到了它。

我们将要回答的下一个问题是：这个拉格朗日量会导致什么样的运动方程？

7.1.4 非齐次麦克斯韦方程和最小耦合

谜底揭晓：这个拉格朗日量给出有电流情况下的非齐次 Maxwell 方程组。

推导是很直白的：将拉格朗日量⁷

$$\mathcal{L}_{\text{Dirac} + \text{额外项} + \text{Maxwell}} = -m \bar{\Psi} \Psi + i \bar{\Psi} \gamma_\mu \partial^\mu \Psi + g A_\mu \bar{\Psi} \gamma^\mu \Psi + \partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu$$

代入每一个场的拉格朗日方程中

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \Psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Psi)} \right) &= 0 \\ \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \bar{\Psi})} \right) &= 0 \\ \frac{\partial \mathcal{L}}{\partial A_\rho} - \partial_\sigma \left(\frac{\partial \mathcal{L}}{\partial (\partial_\sigma A_\rho)} \right) &= 0. \end{aligned}$$

得到

$$\bar{\Psi} (i \gamma_\mu \partial^\mu + m) + g A_\mu \bar{\Psi} \gamma^\mu = 0 \quad (7.20)$$

$$(i \gamma_\mu \partial^\mu - m) \Psi - g A_\mu \gamma^\mu \Psi = 0 \quad (7.21)$$

$$\partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) + g \bar{\Psi} \gamma^\mu \Psi = 0 \quad (7.22)$$

头两个方程描述了自旋 $\frac{1}{2}$ 粒子/场在外加电磁场下的行为。在许多书里这些方程是利用一个叫做**最小耦合 (minimal coupling)** 的办法得到的，即当外场存在时，导数 ∂_μ 应当换成协变导数⁸

$$\partial_\mu \rightarrow D_\mu = i \partial_\mu + A_\mu \quad (7.23)$$

⁷译注：原文下式标记有误，已更正。

⁸译注：此处虚单位 i 的位置似有不妥，同后文

来导出正确的方程。“最小”这个词的意思是，这里只用了一个规范场 A_μ 的四个分量⁹。

现在我们有用来描述 Dirac 旋量在外加电磁场中的行为的方程 ((7.21)式)，所以说可以来兑现在3.7.10节中许下的承诺了。在那儿我们宣称电荷共轭变换（充满暗示的名字）改变其描述对象的电荷。换言之，如果 Ψ 描述带电荷 $+e$ 的某些东西，而电荷共轭旋量 Ψ^C 描述带电荷 $-e$ 的某些东西。电荷决定了自旋 $\frac{1}{2}$ 粒子/场和外加自旋为一的场的耦合强度，我们将考虑 Ψ^C 的运动方程是什么。之后我们将谈谈第三个方程，(7.22)式。

7.1.5 再一次电荷共轭变换

为了得到对应的方程，我们需要找到 Dirac 旋量的电荷共轭变换算符确切的形式。我们在3.7.10节中得到了变换 ((3.236)式)

$$\Psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \rightarrow \Psi^C = \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}. \quad (7.24)$$

这个变换可以简单地用一个 γ_μ 矩阵来描述。利用(6.13)式中 γ_2 的定义，我们有

$$\Psi^C = i\gamma_2 \Psi^* = i \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \chi_L^* \\ \xi_R^* \end{pmatrix}, \quad (7.25)$$

利用 $i\sigma_2 = \epsilon$ 正好就是旋量度规，改写作

$$= \begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix} \begin{pmatrix} \chi_L^* \\ \xi_R^* \end{pmatrix} = \begin{pmatrix} \epsilon \xi_R^* \\ -\epsilon \chi_L^* \end{pmatrix}. \quad (7.26)$$

等价于

$$= \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}, \quad (7.27)$$

正如3.7.7节，特别是(3.194)式中展示的一样。故我们从(7.21)式开始：

$$(i\gamma_\mu \partial^\mu - m)\Psi - gA_\mu \gamma^\mu \Psi = (\gamma_\mu (i\partial^\mu - gA^\mu) - m)\Psi = 0, \quad (7.28)$$

⁹译注：否则，我们可以使用张量场 $\bar{\Psi}\sigma_{\mu\nu}F^{\mu\nu}\Psi$

并取其复共轭，作为得到 Ψ^C 的方程的第一步：

$$\rightarrow (\gamma_\mu^*(-i\partial^\mu - gA^\mu) - m) \Psi^* = 0. \quad (7.29)$$

然后左乘一个 γ_2 ，并在 Ψ 前插入一个 $1 = \gamma_2^{-1}\gamma_2$ ：

$$\rightarrow \gamma_2 (\gamma_\mu^*(-i\partial^\mu - gA^\mu) - m) \underbrace{\gamma_2^{-1}\gamma_2}_{=1} \Psi^* = 0. \quad (7.30)$$

$$\rightarrow \underbrace{(\gamma_2\gamma_\mu^*\gamma_2^{-1})}_{=-\gamma_\mu} (-i\partial^\mu - gA^\mu) - m\gamma_2\gamma_2^{-1} \gamma_2 \Psi^* = 0. \quad (7.31)$$

$$\underbrace{\rightarrow}_{\text{方程乘上 } i} (-\gamma_\mu(-i\partial^\mu - gA^\mu) - m) \underbrace{i\gamma_2\Psi^*}_{=\Psi^C \text{ 见(7.25)式}} = 0. \quad (7.32)$$

$$\rightarrow (\gamma_\mu(i\partial^\mu + gA^\mu) - m) \Psi^C = 0, \quad (7.33)$$

这和 Ψ 的运动方程只在耦合常数上差了个负号 $g \rightarrow -g$ 。这与其名字，电荷共轭，相符¹⁶。

接下来，我们转向上一节中的第三个式子，(7.22)式，称为有电流情况下的非齐次麦克斯韦方程 (inhomogeneous Maxwell equation in the presence of an electric current)。为了讲清楚这个名字，我们又需要 Noether 定理了。

¹⁶ 注意 $\Psi^C \neq \bar{\Psi}$ 。 $\Psi^C = i\gamma_2\Psi^*$ 而 $\bar{\Psi} = \Psi^\dagger\gamma_0 = (\Psi^*)^T\gamma_0$ 。电荷共轭变换让我们能用反粒子来解释一些东西，稍后便会详细讨论。

7.1.6 内禀 $U(1)$ 对称性的 Noether 定理

在4.5.5节中我们得知 Noether 定理将所有内禀对称性和一个守恒量联系了起来。和我们刚才发现的 $U(1)$ 对称性相联系的守恒量是什么？Noether 定理告诉我们，从如下形式的变换

$$\Psi \rightarrow \Psi' = \Psi + \delta\Psi$$

能导出 Noether 流

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} \delta\Psi$$

其满足连续性方程

$$\partial_\mu J^\mu = 0. \quad (7.34)$$

一个全局¹⁷ $U(1)$ 变换是

¹⁷ 记住，自旋 $\frac{1}{2}$ 自由场的拉格朗日量仅仅是全局 $U(1)$ 不变的。而上一节最终得到的拉格朗日量是全局 $U(1)$ 不变的。全局对称性是局域对称性在 $a = \text{常数}$ 时的特殊情况。所以说，如果我们有一个局域 $U(1)$ 不变的拉格朗日量，它自动是全局 $U(1)$ 不变的。这里考虑的全局 $U(1)$ 对称性将导出一个对自由场和相互作用场都守恒的量

$$\Psi \rightarrow \Psi' = e^{iga} = (1 + iga + \dots)\Psi$$

如通常所做的那样，我们仅将指数函数展开到第一项，因为 $\mathcal{U}(1)$ 是一个 Lie 群，而任意变换都可以由无穷小的来生成。一个无穷小的变换写作

$$\Psi \rightarrow \Psi' = \Psi + iga\Psi$$

故我们有 $\delta\Psi = iga\Psi$ 。就如4.5.5节中所做的一样，导出 Noether 流

$$\begin{aligned} J^\mu &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} \delta\Psi \\ &= \frac{\partial(-m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi)}{\partial(\partial_\mu \Psi)} iga\Psi \\ &= -\bar{\Psi}\gamma^\mu ga\Psi = -ga\bar{\Psi}\gamma^\mu\Psi. \end{aligned} \quad (7.35)$$

¹⁸ 保留常数 g ，它的值由实验决定，而不是任意选取的。

由于连续性方程对于任意 a 都成立，故略去¹⁸任意的常数因子 a 。故，我们定义

$$J^\mu = -g\bar{\Psi}\gamma^\mu\Psi \quad (7.36)$$

这常被称作**四电流 (electric four-current)**。零分量正是**电荷密度**，其空间积分对时间守恒¹⁹

¹⁹ 类似的推导参见(4.39)式。

$$Q = \int d^3\mathbf{x} \underbrace{\rho}_{\text{电荷密度}} = \int d^3\mathbf{x} J^0 = -g \int d^3\mathbf{x} \bar{\Psi}\gamma^0\Psi \quad (7.37)$$

在量子框架下， Ψ 和概率幅相关，有要求 $\int d^3\mathbf{x} \bar{\Psi}\gamma^0\Psi = 1$ ，即总概率必须为 100% = 1。所以说守恒量实际就是耦合强度 g ，在电磁作用意义下正比于电荷。所以，全局 $\mathcal{U}(1)$ 对称性导致了电荷守恒。

注意(7.22)式，我们可以借助(7.36)式将其改写成

$$\begin{aligned} \partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma) + \underbrace{g\bar{\Psi}\gamma^\rho\Psi}_{=-J^\rho} &= 0 \\ \rightarrow \partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma) &= J^\rho. \end{aligned} \quad (7.38)$$

借助(6.22)式中定义的电磁张量，改写作

$$\partial_\sigma F^{\rho\sigma} = J^\rho. \quad (7.39)$$

这即是在有电流情况下的非齐次 Maxwell 方程组。这些方程组²⁰，和非齐次 Maxwell 方程组（可以由 $F^{\rho\sigma}$ 的定义直接得到²¹）一起，是经典电动力学的基础。

接下来我们将快速的过一遍有质量自旋为一的场和零自旋场的相互作用。

²⁰ “组”，因为对于每一个分量都有一个方程

²¹ 在11章中将会见到

7.1.7 零自旋有质量场的相互作用

注意我们导出的零自旋场的拉格朗日量

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2)$$

不是 $U(1)$ 不变的，做变换 $\Phi \rightarrow \Phi' = e^{ia}\Phi$ 看一下就知道了。

但是，复标量理论

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi) \quad (7.40)$$

具有 $U(1)$ 对称性，因为有 $\Phi \rightarrow \Phi' = e^{ia}\Phi$ 及 $\Phi^* \rightarrow (\Phi^*)' = e^{-ia}\Phi^*$ 。所以我们会像7.1节中对自旋 $\frac{1}{2}$ 理论所做的那样，导出一个相互作用理论。推导完全是类似的²²，结果是

$$\mathcal{L} = \frac{1}{2}(((\partial_\mu - iqA_\mu)\Phi^*)((\partial^\mu + iqA^\mu)\Phi) - m^2 \Phi^* \Phi) \quad (7.41)$$

代入拉格朗日方程

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi)} \right) &= 0 \\ \frac{\partial \mathcal{L}}{\partial \Phi^*} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi^*)} \right) &= 0, \end{aligned}$$

最终得到相应的运动方程¹⁰

$$(\partial_\mu + iqA_\mu)(\partial^\mu - iqA^\mu)\Phi^* - m^2 \Phi^* = 0 \quad (7.42)$$

$$(\partial_\mu - iqA_\mu)(\partial^\mu + iqA^\mu)\Phi - m^2 \Phi = 0, \quad (7.43)$$

²² 正确的拉格朗日量可以仅通过做替换 $\partial_\mu \rightarrow D_\mu = i\partial_\mu + A_\mu$ 来得到

¹⁰译注：下式符号有误，已更正

描述了每一个零自旋场和自旋为一的无质量场相耦合。

7.1.8 自旋为 1 有质量场的相互作用

自旋为 1 有质量场和自旋为 1 无质量场的相互作用也可以简单的由对称性得到。自旋为 1 无质量场的拉格朗日量为

$$\mathcal{L}_{\text{Maxwell}} = \frac{1}{4} F^{\mu\nu} F_{\mu\nu}.$$

为了区分有质量场和无质量场，我们将有质量场命名为 B^μ ，并定义

$$G^{\mu\nu} := \partial^\mu B^\nu - \partial^\nu B^\mu$$

自旋为 1 的有质量场的拉格朗日量写作 ((6.19)式)

$$\mathcal{L}_{\text{Proca}} = \frac{1}{2} G^{\mu\nu} G_{\mu\nu} + m^2 B^\mu B_\mu$$

Lorentz 对称性表明拉格朗日量中的相互作用项一定具有形式

$$\mathcal{L}_{\text{Proca-相互作用}} = C G_{\mu\nu} F^{\mu\nu}$$

其中耦合常数 C 需要由实验来定。感兴趣的读者可以借助欧拉-拉格朗日方程自行推导相应的运动方程。

7.2 $SU(2)$ Interactions $SU(2)$ 相互作用

$\mathcal{U}(1)$ 对称性的成功激励着我们提出这个问题： $\mathcal{U}(1)$ 对称性是我们的拉格朗日量唯一的内禀对称性吗？

事实上，我们能够在两个无质量自旋 $\frac{1}{2}$ 粒子上找到另一个内禀对称性。将 6.3 节中得到的拉格朗日量作两份加在一起，就是两个无质量自旋 $\frac{1}{2}$ 粒子的拉格朗日量了。最后的结果即 (6.16) 式：

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma_\mu \partial^\mu - m)\psi.$$

这里我们将质量项略去，即 $m = 0$ ，否则马上就能看到拉格朗日量不是变换不变的。再之后将看到我们是怎么在不破坏对称性的情况下把质量项捡回来的。“作两份加在一起”指的是

$$\mathcal{L}_{D_1+D_2} = i\bar{\psi}_1\gamma_\mu\partial^\mu\psi_1 + i\bar{\psi}_2\gamma_\mu\partial^\mu\psi_2 \quad (7.44)$$

通过新定义二重态 (doublet) Ψ

$$\begin{aligned}\Psi &:= \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \\ \rightarrow \bar{\Psi} &:= (\bar{\psi}_1 \quad \bar{\psi}_2)\end{aligned}$$

我们能将前式改写成

$$\mathcal{L}_{D_1+D_2} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi. \quad (7.45)$$

这个拉格朗日量在全局 $SU(2)$ 变换

$$\Psi \rightarrow \Psi' = e^{ia_i\frac{\sigma_i}{2}}\Psi \quad (7.46)$$

$$\Rightarrow \bar{\Psi} \rightarrow \bar{\Psi}' = \bar{\Psi}e^{-ia_i\frac{\sigma_i}{2}}, \quad (7.47)$$

下不变。其中重复指标 i 代表求和, a_i 为任意实数而 $\frac{\sigma_i}{2}$ 是 $SU(2)$ 群的常用生成元, σ_i 是泡利矩阵。

为了看明白不变性, 我们将变换后的拉格朗日量写出^{23 11}

$$\begin{aligned}\mathcal{L}'_{D_1+D_2} &= i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\ &= i\bar{\Psi}e^{-ia_i\frac{\sigma_i}{2}}\gamma_\mu\partial^\mu e^{ia_i\frac{\sigma_i}{2}}\Psi \\ &= i\bar{\Psi}\gamma_\mu\partial^\mu\Psi = \mathcal{L}_{D_1+D_2}, \quad \checkmark\end{aligned} \quad (7.48)$$

最后一行源于我们的变换 $e^{ia_i\frac{\sigma_i}{2}}$ 作用在我们的新定义的二分量对象 Ψ 上, 而 γ_μ 作用在二重态里面的对象上, 即 Dirac 旋量¹²。我们可以用指标来将其写清楚¹³

$$\left[\left(e^{-ia_i\frac{\sigma_i}{2}}\right)_{ab}\delta_{\alpha\beta}\right]\left[\delta_{bc}\gamma_\mu^{\beta\delta}\right]\left[\left(e^{ia_i\frac{\sigma_i}{2}}\right)_{cd}\delta_{\delta\epsilon}\right] = [\delta_{ad}\gamma_\mu^{\alpha\epsilon}]$$

这个对称性也该作为局域对称性。 $SU(2)$ 变换将二重态的两个分量混合起来。稍后我们将会将这两个场称作电子和电子中微子场。对称性告诉我们, 把哪个叫电子哪个叫电子中微子其实无所谓, 因为用了 $SU(2)$ 变换, 想怎么换就怎么换。如果这仅仅是一个全局变换, 那么当我们确定一个选择²⁴, 即哪个叫电子哪个叫电

²³ 这里我们丢掉了质量项。形如 $-m_1\bar{\Psi}\Psi$ 和 $-m_2\bar{\Psi}\Psi$ 的项可以借助 Ψ 的二分量定义和质量矩阵

$$m := \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$$

写作

$$\mathcal{L}_{D_1+D_2} = -\bar{\Psi}m\Psi。$$

不幸的是, 这一项并不是 $SU(2)$ 不变的, 即 $\mathcal{L}'_{D_1+D_2} = -\bar{\Psi}'m\Psi' = i\bar{\Psi}e^{-ia_i\frac{\sigma_i}{2}}\underbrace{me^{ia_i\frac{\sigma_i}{2}}}_{\neq m}\Psi$ 。当质量

相等 $m_1 = m_2$ 时, 它自然是不变的, 但是等会能看到我们可以在不违背对称性的情况下包含任意大小的质量。实验告诉我们二重态的两个场并没有产生相同质量的粒子, 即 $m_1 \neq m_2$ 。稍后会有更加详细的讨论。

²⁴ 这叫规范选取 (choosing a gauge)。

¹¹ 译注: 原文边注公式缺一', 已更正

¹² 译注: 用线性代数的话来讲, 两个变换分别作用在两个子空间上, 而这两个子空间的直积才是二重态所处的线性空间。

¹³ 译注: 指标真尼玛多。

子中微子后，这个选择将瞬间适用到整个宇宙上去。故我们得研究局域变换。我们再一次发现它不是的，但正如局域 $\mathcal{U}(1)$ 对称性的情形一样，我们将尽最大努力让拉格朗日量具有局域 $SU(2)$ 不变性。

问题依旧出在求导上，这会导致额外项：

$$\begin{aligned}
 \mathcal{L}'_{D_1+D_2} &= i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' \\
 &= i\bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\partial^\mu e^{ia_i(x)\frac{\sigma_i}{2}}\Psi \\
 &\stackrel{\text{莱布尼兹律}}{=} i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi \neq \mathcal{L}_{D_1+D_2}.
 \end{aligned} \tag{7.49}$$

有着7.1节中处理 $\mathcal{U}(1)$ 的经验，我们很清楚接下来该做什么。我们将一步一步的导出一个局域 $SU(2)$ 不变的拉格朗日量，这会花一些时间。为了避免读者陷入迷惑，在这里先列出接下来要进行的步骤：

- 最后的结论是，加上用以描述两个自旋 $\frac{1}{2}$ 场 $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ 和三个有质量自旋为 1 的场 W_i^μ 相互作用的额外项 $i\bar{\Psi}\gamma_\mu W_i^\mu\sigma_i\Psi$ 后，拉格朗日量成为局域 $SU(2)$ 不变的。
- 注意，不变性仅当按照下式变换

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu a_i(x) + \epsilon_{ijk}a_j(x)(W_\mu)_k$$

而不是

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu a_i(x),$$

即不同于 $\mathcal{U}(1)$ 对称性时的情况。

- 仅当自由场拉格朗日量为如下形式¹⁴

$$\mathcal{L} = \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})^i$$

其中

$$(W_{\mu\nu})_i = \partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i + \epsilon_{ijk}(W_\mu)_j(W_\nu)_k$$

¹⁴译注：原文字体错误，已更正。

而不是

$$(W_{\mu\nu})_i = \partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i。$$

时, 对称性才存在。换言之, 为了得到一个局域 $SU(2)$ 不变的拉格朗日量, 我们必须重新定义三个 W_i^μ 场的拉格朗日量, 使得其在变换

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu a_i(x) + \epsilon_{ijk} a_j(x) (W_\mu)_k$$

下不变。

现在我们来具体执行这些步骤。从(7.11)式中我们得知自旋为 1 的场有内禀对称性

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu a(x)$$

而这里我们将会使用三个自旋为 1 的场 W_1^μ, W_2^μ 和 W_3^μ , 这是因为 $SU(2)$ 有三个生成元 $\frac{\sigma_i}{2}$, 由指数上的求和 $a_i(x) \frac{\sigma_i}{2}$ 产生的每一个额外项都需要一个场。这导致了局域内禀对称性

$$W_i^\mu \rightarrow W_i'^\mu = W_i^\mu + \partial^\mu a_i(x)$$

故我们尝试在拉格朗日量里添加项

$$\bar{\Psi} \gamma_\mu \frac{\sigma_j}{2} W_j^\mu \Psi。$$

不幸的是, 这样并不能得到局域 $SU(2)$ 不变的拉格朗日量¹⁵

$$\begin{aligned} \mathcal{L}'_{D_1+D_2+\text{额外项}} &= i\bar{\Psi}' \gamma_\mu \partial^\mu \Psi' + \bar{\Psi}' \gamma_\mu \frac{\sigma_j}{2} W_j'^\mu \Psi' \\ &= i\bar{\Psi} e^{-ia_i(x) \frac{\sigma_i}{2}} \gamma_\mu \partial^\mu e^{ia_i(x) \frac{\sigma_i}{2}} \Psi \\ &\quad + i\bar{\Psi} e^{-ia_i(x) \frac{\sigma_i}{2}} \gamma_\mu \frac{\sigma_j}{2} \left(W_j^\mu + \partial^\mu a_j(x) \right) e^{ia_i(x) \frac{\sigma_i}{2}} \Psi \\ &= i\bar{\Psi} \gamma_\mu \partial^\mu \Psi - \bar{\Psi} \gamma_\mu \left(\partial^\mu a_i(x) \frac{\sigma_i}{2} \right) \Psi \\ &\quad + \underbrace{\bar{\Psi} e^{-ia_i(x) \frac{\sigma_i}{2}} \gamma_\mu \frac{\sigma_j}{2} W_j^\mu e^{ia_i(x) \frac{\sigma_i}{2}} \Psi}_{\neq \gamma_\mu \frac{\sigma_j}{2} W_j^\mu \text{ 因为 } [\frac{\sigma_i}{2}, \frac{\sigma_j}{2}] = \epsilon_{ijk} \frac{\sigma_k}{2} \neq 0} \end{aligned}$$

¹⁵ 译注: 原文下式上下标混乱, 以及最后两行漏 $1/2$ 因子, 以及倒数第二个等号应该为不等号, 已更正。

$$\begin{aligned}
& + i\bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\frac{\sigma_j}{2}(\partial^\mu a_j(x))e^{ia_i(x)\frac{\sigma_i}{2}}\Psi \\
& \neq i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_i}{2}W_i^\mu\Psi \\
& \quad + \bar{\Psi}\gamma_\mu\left(\partial^\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi \\
& = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_i}{2}W_i^\mu\Psi = \mathcal{L}_{D_1+D_2+\text{额外项}} \quad \checkmark
\end{aligned} \tag{7.50}$$

²⁵ 数学上将有非对易元素的群称为非阿贝尔的。作为对比, $\mathcal{U}(1)$ 是阿贝尔的, 此时一切都相对简单。

容易发现问题出在 $SU(2)$ 的生成元不对易上²⁵。我们来仔细研究一下这个困难。正如通常处理 Lie 群的那样, 将目光集中在无穷小变换上, 略去高阶项:

$$\begin{aligned}
e^{-ia_i(x)\frac{\sigma_i}{2}}\frac{\sigma_j}{2}e^{ia_i(x)\frac{\sigma_i}{2}} & \approx \left(1 - ia_i(x)\frac{\sigma_i}{2}\right)\frac{\sigma_j}{2}\left(1 + ia_i(x)\frac{\sigma_i}{2}\right) \\
& = \frac{\sigma_j}{2} - a_i(x)\frac{\sigma_i}{2}\frac{\sigma_j}{2} + a_i(x)\frac{\sigma_j}{2}\frac{\sigma_i}{2} + \mathcal{O}(a^2) \\
& = \frac{\sigma_j}{2} + a_i(x)\left[\frac{\sigma_j}{2}, \frac{\sigma_i}{2}\right] + \mathcal{O}(a^2) \\
& = \frac{\sigma_j}{2} - a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2} + \mathcal{O}(a^2) \neq \frac{\sigma_j}{2}
\end{aligned} \tag{7.51}$$

作为无穷小变换, $\mathcal{O}(n^2)$ 和更高阶的项都可以略去, 拉格朗日量为

$$\begin{aligned}
\mathcal{L}'_{D_1+D_2+\text{额外项}} & = i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' + \bar{\Psi}'\gamma_\mu\frac{\sigma_j}{2}W_j'^\mu\Psi' \\
& \stackrel{(7.50)\text{式}}{=} i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}\gamma_\mu\left(\partial_\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi \\
& \quad + \bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu e^{ia_i(x)\frac{\sigma_i}{2}}\Psi \\
& \quad + \bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\frac{\sigma_j}{2}(\partial_\mu a_j(x))e^{-ia_i(x)\frac{\sigma_i}{2}}\Psi \\
& \stackrel{(7.51)\text{式}}{\approx} i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}\gamma_\mu\left(\partial_\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi \\
& \quad + \bar{\Psi}\gamma_\mu\left(\frac{\sigma_j}{2} - a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)W_j^\mu\Psi \\
& \quad + \bar{\Psi}\gamma_\mu\left(\frac{\sigma_j}{2} - a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}\right)(\partial_\mu a_j(x))\Psi \\
& = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - \bar{\Psi}\gamma_\mu\left(\partial_\mu a_i(x)\frac{\sigma_i}{2}\right)\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu\Psi \\
& \quad - \bar{\Psi}\gamma_\mu a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}W_j^\mu\Psi + \bar{\Psi}\gamma_\mu(\partial^\mu a_j(x))\Psi \\
& \quad - \underbrace{\bar{\Psi}\gamma_\mu a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}(\partial^\mu a_j(x))\Psi}_{\mathcal{O}(a^2)}
\end{aligned}$$

$$= i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu\Psi - \bar{\Psi}\gamma_\mu a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}W_j^\mu\Psi \quad (7.52)$$

属于 $(W_\mu)_i$ 的能保证 $SU(2)$ 不变的内禀对称性只能是

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu a_i(x) + \epsilon_{ijk}a_j(x)(W_\mu)_k$$

因为它能够在拉格朗日量中添加一项，恰好抵消掉那个破坏对称性的家伙¹⁶

$$\begin{aligned} \mathcal{L}'_{D_1+D_2+\text{额外项}} &= i\bar{\Psi}'\gamma_\mu\partial^\mu\Psi' + \bar{\Psi}'\gamma_\mu\frac{\sigma_j}{2}W_j'^\mu\Psi' \\ &= i\bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\partial^\mu e^{ia_i(x)\frac{\sigma_i}{2}}\Psi \\ &\quad + i\bar{\Psi}e^{-ia_i(x)\frac{\sigma_i}{2}}\gamma_\mu\frac{\sigma_j}{2}\left(W_j^\mu + \partial^\mu a_j(x) + \epsilon_{jlm}a_l(x)(W^\mu)_m\right)e^{ia_i(x)\frac{\sigma_i}{2}}\Psi \\ &\stackrel{(7.52)\text{式}}{=} \underbrace{i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu\Psi - \bar{\Psi}\gamma_\mu a_i(x)\epsilon_{ijk}\frac{\sigma_k}{2}W_j^\mu\Psi}_{(7.52)\text{式}} + \bar{\Psi}\gamma_\mu\frac{\sigma_j}{2}\epsilon_{ilm}a_l(x)W_m^\mu\Psi \\ &= \mathcal{L}_{D_1+D_2+\text{额外项}} \quad \checkmark \end{aligned} \quad (7.53)$$

接下来得看看这个到底是不是三个自旋为 1 的自由粒子的拉格朗日量的内禀对称性。三个自旋为 1 的粒子的拉格朗日量是

$$\begin{aligned} \mathcal{L}_{3\times\text{Maxwell}} &= \frac{1}{4}(W^{\mu\nu})_1(W_{\mu\nu})_1 + \frac{1}{4}(W^{\mu\nu})_2(W_{\mu\nu})_2 + \frac{1}{4}(W^{\mu\nu})_3(W_{\mu\nu})_3 \\ &= \frac{1}{4}(W^{\mu\nu})_i(W_{\mu\nu})_i \end{aligned} \quad (7.54)$$

其中

$$(W_{\mu\nu})_i = \partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i.$$

我们希望拉格朗日量在变换

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu a_i(x) + \epsilon_{ijk}a_j(x)(W_\mu)_k \quad (7.55)$$

¹⁶译注：原文上下标混乱，且等号注释位置有误，已更正。

下不变。变换后的拉格朗日量是

$$\begin{aligned}
\mathcal{L}'_{3 \times \text{Maxwell}} &= \frac{1}{4} (W'^{\mu\nu})_i (W'_{\mu\nu})_i \\
&\stackrel{\text{见(6.25)式}}{=} \partial^\mu (W'^\nu)_i \partial_\mu (W'_\nu)_i - \partial^\mu (W'^\nu)_i \partial_\nu (W'_\mu)_i \\
&= \partial^\mu \left((W^\nu)_i + \partial^\nu a_i(x) + \epsilon_{ijk} a_j(x) (W^\nu)_k \right) \\
&\quad \partial_\mu \left((W_\nu)_i + \partial_\nu a_i(x) + \epsilon_{ijk} a_j(x) (W_\nu)_k \right) \\
&\quad - \partial^\mu \left((W^\nu)_i + \partial^\nu a_i(x) + \epsilon_{ijk} a_j(x) (W^\nu)_k \right) \\
&\quad \partial_\nu \left((W_\mu)_i + \partial_\mu a_i(x) + \epsilon_{ijk} a_j(x) (W_\mu)_k \right) \\
&\stackrel{\text{略去 } \mathcal{O}(a^2) \text{ 项}}{=} \partial^\mu (W^\mu)_i \partial_\mu (W_\nu)_i + \overline{\partial^\mu (W^\nu)_i \partial_\mu \partial_\nu a_i(x)} \\
&\quad + (\partial^\mu (W^\nu)_i) \partial_\mu \epsilon_{ijk} a_j(x) (W_\nu)_k \\
&\quad + \overline{\partial^\mu \partial^\nu a_i(x) (\partial_\mu (W_\nu)_i)} + \partial^\mu \epsilon_{ijk} a_j(x) (W^\nu)_k \partial_\mu (W_\nu)_i \\
&\quad - \partial^\mu (W^\nu)_i \partial_\nu (W_\mu)_i - \overline{(\partial^\mu (W^\nu)_i) \partial_\mu \partial_\nu a_i(x)} \\
&\quad - \partial^\mu (W^\nu)_i \partial_\nu \epsilon_{ijk} a_j(x) (W_\mu)_k \\
&\quad - \overline{\partial^\mu \partial^\nu a_i(x) (\partial_\nu (W^\mu)_i)} - \partial^\mu \epsilon_{ijk} a_j(x) (W^\nu)_k \partial_\nu (W_\mu)_i \\
&= \underbrace{\partial^\mu (W^\nu)_i \partial_\mu (W_\nu)_i - \partial^\mu (W^\nu)_i \partial_\nu (W_\mu)_i}_{\mathcal{L}_{3 \times \text{Maxwell}}} \\
&\quad + (\partial^\mu (W^\nu)_i) \partial_\mu \epsilon_{ijk} a_j(x) (W_\nu)_k \\
&\quad + \partial^\mu \epsilon_{ijk} a_j(x) (W^\nu)_k \partial_\mu (W_\nu)_i - \partial^\mu (W^\nu)_i \partial_\nu \epsilon_{ijk} a_j(x) (W_\mu)_k \\
&\quad - \partial^\mu \epsilon_{ijk} a_j(x) (W^\nu)_k \partial_\nu (W_\mu)_i \\
&\stackrel{\text{莱布尼兹律}}{=} \mathcal{L}_{3 \times \text{自旋1}} - (\partial^\mu (W^\nu)_i) \epsilon_{ijk} (\partial_\mu a_j(x)) (W_\nu)_k \\
&\quad - (\partial^\mu (W^\nu)_i) \epsilon_{ijk} a_j(x) (\partial_\mu (W_\nu)_k) \\
&\quad - \epsilon_{ijk} (\partial^\mu a_j(x)) (W^\nu)_k \partial_\mu (W_\nu)_i - \epsilon_{ijk} a_j(x) (\partial^\mu (W^\nu)_k) \partial_\mu (W_\nu)_i \\
&\quad + \partial^\mu (W_\nu)_i \epsilon_{ijk} (\partial_\nu a_j(x)) (W_\mu)_k + \partial^\nu (W^\nu)_i \epsilon_{ijk} a_j(x) (\partial_\nu (W_\mu)_k) \\
&\quad + \epsilon_{ijk} (\partial^\mu a_j(x)) (W^\nu)_k \partial_\nu (W_\mu)_i + \epsilon_{ijk} a_j(x) (\partial^\mu (W^\nu)_k) \partial_\nu (W_\mu)_i
\end{aligned} \tag{7.56}$$

显然没有内禀对称性。

尽管如此，我们可以通过在张量 $W_{\mu\nu}$ 上加一个额外项

$$-\epsilon_{ijk}(W_\mu)_j(W_\nu)_k$$

来保证内禀对称性。即

$$(W_{\mu\nu})_i = \partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i - \epsilon_{ijk}(W_\mu)_j(W_\nu)_k,$$

这样三个自旋 1 场便有了一个新的拉格朗日量¹⁷

$$\begin{aligned}\mathcal{L}_{3\times\text{Maxwell}} &= \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \\ &= \left(\partial_\mu(W_\nu)_i - \partial_\nu(W_\mu)_i - \epsilon_{ijk}(W_\mu)_j(W_\nu)_k \right) \\ &\quad \times \left(\partial^\mu(W^\nu)_i - \partial^\nu(W^\mu)_i - \epsilon_{ijk}(W^\mu)_j(W^\nu)_k \right) \\ &\quad (7.57)\end{aligned}$$

读者可自行完成冗长的计算以验证拉格朗日量的不变性。综上所述，我们找到了局域 $SU(2)$ 不变的拉格朗日量，写作

$$\mathcal{L}_{D_1+D_2+\text{相互作用}+3\times\text{Maxwell}} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + \bar{\Psi}\gamma_\mu\frac{\sigma_j}{2}W_j^\mu\Psi - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \quad (7.58)$$

7.3 Mass Terms and Unification of $\mathcal{U}(1)$ and $SU(2)$ 质量项、 $\mathcal{U}(1)$ 和 $SU(2)$ 的统一

在上一章中，为了防止 $SU(2)$ 对称性被破坏，我们不能往拉格朗日量中添加形如 $m_1\bar{\Psi}\Psi$ 或 $m_2(W_\mu)_i(W^\mu)_i$ 的质量项。而从实验中我们知道相应的粒子²⁶却有质量，这常被解释为 $SU(2)$ 对称性的破缺。这意味着对称性在高能标下存在，而在低能标下自发破缺。

之前我们得到了一个局域 $\mathcal{U}(1)$ 不变的拉格朗日量，在这里为方便起见我们将相应的自旋为 1 的场记做 B_μ ：

$$\mathcal{L}_{\text{局域 } \mathcal{U}(1) \text{ 不变}} = -m\bar{\psi}\psi + \bar{\psi}\gamma_\mu(i\partial^\mu + gB_\mu)\psi - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \quad (7.59)$$

²⁶ 例如，由 Ψ 描述的电子 e^- 和电子中微子 ν_e ，以及由 $(W_\mu)_i$ 描述的三个 bosons。

¹⁷译注：原文下式多一乘号，已更正

其中

$$B^{\mu\nu} := \partial^\mu B^\nu - \partial^\nu B^\mu$$

由于自旋为 1 的场 B^μ 能保证拉格朗日量的 $U(1)$ 不变性, 所以其常被称作 $U(1)$ 规范场。

²⁷ 之前为了简明起见略去的三个 W_i^μ 场的耦合常数, 现在记做 g' 。

局域 $SU(2)$ 不变的拉格朗日量是²⁷

$$\mathcal{L}_{\text{局域 } SU(2) \text{ 不变}} = \bar{\Psi} \gamma_\mu (i\partial^\mu + g' \sigma_j W_j^\mu) \Psi - \frac{1}{4} (W_{\mu\nu})_i (W^{\mu\nu})_i$$

其中

$$(W_{\mu\nu})_i = \partial_\mu (W_\nu)_i - \partial_\nu (W_\mu)_i + \epsilon_{ijk} (W_\mu)_j (W_\nu)_k$$

以及二重态

$$\Psi := \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

类似的, 由于这三个自旋为 1 的场 $(W_\nu)_i$ 能保证拉格朗日量的局域 $SU(2)$ 不变性, 故其常被称为 $SU(2)$ 规范场。

将它们加到一起我们能得到一个局域 $U(1)$ 且 $SU(2)$ 不变的拉格朗日量

$$\mathcal{L}_{SU(2) \text{ 且 } U(1)} = \bar{\Psi} \gamma_\mu (i\partial^\mu + g B^\mu + g' \sigma_j W_j^\mu) \Psi - \frac{1}{4} (W_{\mu\nu})_i (W^{\mu\nu})_i - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}. \quad (7.60)$$

现在, 该怎么在不破坏 $SU(2)$ 对称性的前提下加入质量项呢? 至今我们还没用过的唯一一个东西是零自旋场, 所以来看看这个。零自旋场具有全局 $U(1)$ 不变性的拉格朗日量由(7.40)式给出

$$\mathcal{L}_{\text{零自旋}} = \frac{1}{2} (\partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi). \quad (7.61)$$

我们还能在不破坏任何对称约束的情况下在这个拉格朗日量中添加 ϕ 的高阶项²⁸。故我们写出下式, 其中常数改作了惯用的记号

$$\mathcal{L}_{\text{零自旋} + \text{额外项}} = \partial_\mu \phi^\dagger \partial^\mu \phi + \rho^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2. \quad (7.62)$$

²⁸ 记住, 唯一真正禁止的是高阶导数项, 否则无法得到合理的理论。 ϕ 的高幂次项描述场 ϕ 的自作用, 在希望得到一个“自由”理论的时候被略去。

从(7.41)已经得知如何在使拉格朗日量局域 $\mathcal{U}(1)$ 不变的前提下，在零自旋场和 $\mathcal{U}(1)$ 规范场 B_μ 间添加耦合项：

$$\begin{aligned} \mathcal{L}_{\text{零自旋} + \text{额外项} + \text{自旋} 1 \text{ 耦合}} &= \left((\partial_\mu - i\frac{1}{2}gB_\mu)\phi^\dagger \right) \left((\partial^\mu + i\frac{1}{2}gB^\mu)\phi \right) \\ &\quad + \rho^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \end{aligned} \quad (7.63)$$

其具有对称性²⁹

$$B_\mu \rightarrow B'_\mu = B_\mu + \partial_\mu a(x) \quad (7.64)$$

$$\phi(x) \rightarrow \phi'(x) = e^{ia(x)}\phi(x). \quad (7.65)$$

²⁹ 见(7.13)式，并注意整体乘在拉格朗日量上的一个常数不会影响物理。

用我们在上一节中对自旋 $\frac{1}{2}$ 场导出局域 $SU(2)$ 不变拉格朗日量同样的办法，我们能对零自旋场的二重态³⁰写出一个局域 $SU(2)$ 不变的拉格朗日量

$$^{30}\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$\begin{aligned} \mathcal{L}_{SU(2) \text{ 且 } \mathcal{U}(1)} &= \left((\partial_\mu - ig'\sigma_i(W_\mu)_i - i\frac{1}{2}gB_\mu)\Phi^\dagger \right) \left((\partial^\mu + ig'\sigma_i(W^\mu)_i + i\frac{1}{2}gB^\mu)\Phi \right) \\ &\quad + \underbrace{\rho^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2}_{\equiv -V(\Phi)} \end{aligned} \quad (7.66)$$

其中 $\Phi := \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$ ，对称变换为 ((7.55)式)

$$(W_\mu)_i \rightarrow (W'_\mu)_i = (W_\mu)_i + \partial_\mu b_i(x) + \epsilon_{ijk}b_j(x)(W_\mu)_k \quad (7.67)$$

和

$$\Phi \rightarrow \Phi' = e^{ib_i(x)\sigma_i}\Phi, \quad \bar{\Phi} \rightarrow \bar{\Phi}' = \bar{\Phi}e^{-ib_i(x)\sigma_i} \quad (7.68)$$

我们从局域 $SU(2)$ 不变的拉格朗日量出发，研究它是如何为场 W_μ^i 和 B_μ^i 给出质量项的。手动添加质量项是不可行的，因为它会破坏对称性从而导致不可接受的理论³¹。

³¹ 这个理由很复杂，具体讨论不会在本书中给出。技术上讲：我们需要一个局域 $SU(2)$ 对称的拉格朗日量来得到一个可重整的理论。感兴趣的读者可以参阅本节末给出的参考书目。

我们在上面定义了

$$\begin{aligned} V(\Phi) &= -\rho^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \\ &= -\rho^2 \phi_1^\dagger \phi_1 + \lambda (\phi_1^\dagger \phi_1)^2 - \rho^2 \phi_2^\dagger \phi_2 + \lambda (\phi_2^\dagger \phi_2)^2 \quad (7.69) \\ &= V(\phi_1) + V(\phi_2) \end{aligned}$$

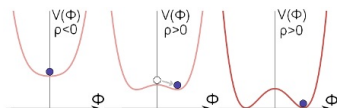


图 7.1: Higgs 势取不同参数 ρ 值的图像, ρ 被认为是由于宇宙暴涨所造成的宇宙冷却而改变。Figure adapted from “Spontaneous symmetry breaking” by FT2 (Wikimedia Commons) released under a CC BY-SA 3.0 licence: <http://creativecommons.org/licenses/by-sa/3.0/deed.en>. URL: [http://commons.wikimedia.org/wiki/File:Spontaneous_symmetry_breaking_\(explanatory_diagram\).png](http://commons.wikimedia.org/wiki/File:Spontaneous_symmetry_breaking_(explanatory_diagram).png), Accessed: 8.12.2014

这常称作 Higgs 势。其参数 ρ 取不同值而 $\lambda > 0$ 的图绘制于图 7.1

这个想法是这样的: 在相当高的温度下, 例如宇宙早期, 势能会像图左那样。极小值点, 这里成为真空值, 毫无疑问的在 $\phi = 0$ 处。参数 λ 和 ρ 随着温度的下降发生改变, 势能的形状也是一样。当温度低于某个临界值时, 如图右所示, $\phi = 0$ 处不再是极小点。此时将会有许多点都能取到极小值。

事实上, 这个势能有无穷多个极小值点。其可用通常的办法计算得到

$$V(\phi) = -\rho^2 |\phi|^2 + \lambda |\phi|^4 \quad (7.70)$$

$$\frac{\partial V(\phi)}{\partial \phi} = -2\rho^2 |\phi| + 4\lambda |\phi|^3 \quad (7.71)$$

$$\rightarrow |\phi|(-2\rho^2 + 4\lambda |\phi|^2) \stackrel{!}{=} 0 \quad (7.72)$$

$$\rightarrow |\phi|^2 \stackrel{!}{=} \frac{\rho^2}{2\lambda} \quad (7.73)$$

$$\rightarrow |\phi| \stackrel{!}{=} \sqrt{\frac{\rho^2}{2\lambda}} \quad (7.74)$$

$$\phi_{\min} = \sqrt{\frac{\rho^2}{2\lambda}} e^{i\varphi}. \quad (7.75)$$

这对于 φ 的任意取值都是极小值, 所以我们有无穷多个极小点。全体极小点躺在了半径为 $\sqrt{\frac{\rho^2}{2\lambda}}$ 的一个圆上。图 7.2 即为 Higgs 的三维图。就如同一粒弹珠从墨西哥草帽顶端自发而随机的滚落, 一个新的真空值有无穷多种选择。

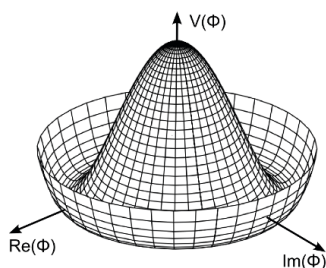


图 7.2: Higgs 的三维图。Figure adapted from “Mexican hat potential polar” by Rupert Millard (Wikimedia Commons) released under a public domain licence. URL: http://commons.wikimedia.org/wiki/File:Mexican_hat_potential_polar.svg, Accessed: 7.5.2014

从(7.69)中我们可知对于二重态, 两个分量都需要作出这种选择。则对于二重态其极小点为

$$\Phi_{\min} = \begin{pmatrix} \phi_{1 \min} \\ \phi_{2 \min} \end{pmatrix} \quad (7.76)$$

一个经济的选择³²是

³² 复习: 对称破缺意味着一个极小值有无穷多种选取方式

$$\Phi_{\min} = \begin{pmatrix} 0 \\ \sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix} \equiv \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \quad (7.77)$$

其中因子 $\frac{1}{2}$ 只是为了使计算简洁而引入，自然有 $v \equiv \sqrt{\frac{\rho^2}{\lambda}}$ 。下一步中我们会将场 Φ 在极小值附近展开以得到想要的物理。之后我们将会学到，由于得不到精确解，量子场论中的计算总是围绕着极小值点的一系列级数。为了得到有意义的结果，我们必须将场移动到新的极小点。故我们考虑

$$\Phi = \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} + \phi_{2r} + i\phi_{2c} \end{pmatrix}. \quad (7.78)$$

改写作³³

$$\Phi = e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}, \quad (7.79)$$

这是由于考虑将指数函数展开成级数，代入泡利矩阵 σ_i 矩阵的具体形式，在一阶意义上

$$\begin{aligned} e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} &\approx (1 + i\theta_i \frac{\sigma_i}{2}) \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= (1 + i\theta_1 \frac{\sigma_1}{2} + i\theta_2 \frac{\sigma_2}{2} + i\theta_3 \frac{\sigma_3}{2}) \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 1 + i\frac{1}{2}\theta_3 & \frac{1}{2}\theta_1 - i\frac{1}{2}\theta_2 \\ \frac{1}{2}\theta_1 + i\frac{1}{2}\theta_2 & 1 - i\frac{1}{2}\theta_3 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} (\frac{1}{2}\theta_1 - i\frac{1}{2}\theta_2) \frac{v+h}{\sqrt{2}} \\ (1 - i\frac{1}{2}\theta_3) \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ \text{重新定义} \rightarrow &\equiv \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} \phi_{2r} + i\phi_{2c} \end{pmatrix}. \end{aligned} \quad (7.80)$$

将零自旋复二重态写成这个形式，有助于利用局域 $\mathcal{SU}(2)$ （规范）对称性化简计算。为了得到物理结果，我们必须选取一个规范，而合适的选择会使生活更美好³⁴。

³³ 等会就能看到这个形式有多有用了。

³⁴ 不同的规范适用于不同的问题。而这里我们采用的叫做么正规范，这对理解一个理论的物理内容也是有益的。

一个一般的局域 $SU(2)$ 变换写作

$$\Phi \rightarrow \Phi' = e^{ib_i(x)\frac{\sigma_i}{2}} \Phi, \quad (7.81)$$

通过选取合适的 $b_i(x)$, 我们可以消去(7.79)式中的指数因子。在么正规范 (unitary gauge) 下, 复标量二重态为

$$\Phi_{un} = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}. \quad (7.82)$$

这件事也可以这样理解: 复标量二重态的四个分量和 $SU(2)$ 规范的自由度恰好相等³⁵。这样, 这三个场分量便是非物理的³⁶, 不可能被实验观测到。最后留下来的便是一个物理场 h , 称作 **Higgs 场**。

接下来, 我们希望看看拉格朗日量对称性破缺的具体含义。我们将(7.66)¹⁸式给出的拉格朗日量重新写在这里, 并调整一下记号

$$\begin{aligned} \mathcal{L} = & \left((\partial_\mu - ig'\frac{1}{2}\sigma_i(W_\mu)_i - i\frac{1}{2}gB_\mu)\Phi^\dagger \right) \left((\partial^\mu + ig'\sigma_i(W^\mu)_i + i\frac{1}{2}gB^\mu)\Phi \right) \\ & - V(\Phi) \end{aligned} \quad (7.83)$$

现在将场 Φ 用平移且取么正规范的场, 即(7.82)式代入。我们对新出现的包含真空值 v 的项尤其感兴趣。而其他项描述了 Higgs 场的自作用以及 Higgs 场与其他场的相互作用, 我们不会再对其做出进一步的讨论。我们将极小值点 $\Phi \rightarrow \Phi_{\min} = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ (即不计入 h), 代入得

$$\begin{aligned} & \left((\partial_\mu - ig'\frac{1}{2}\sigma_i(W_\mu)_i - i\frac{1}{2}gB_\mu)\Phi_{\min}^\dagger \right) \left((\partial^\mu + ig'\frac{1}{2}\sigma_i(W^\mu)_i + i\frac{1}{2}gB^\mu)\Phi_{\min} \right) \\ & = \left| (\partial^\mu + ig'\frac{1}{2}\sigma_i(W^\mu)_i + i\frac{1}{2}gB^\mu)\Phi_{\min} \right|^2 \\ & = \left| (\partial^\mu + ig'\frac{1}{2}\sigma_i(W^\mu)_i + i\frac{1}{2}gB^\mu)\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 \\ & = \frac{v^2}{8} \left| (g'\sigma_i(W^\mu)_i + gB^\mu) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|^2 \end{aligned}$$

¹⁸译注: 原文引用公式有误, 已更正

³⁵ 注意, 这仅在满足如下这些条件下时成立: 有一个局域 $SU(2)$ 理论, 场 $\theta = \theta(x)$ 理所应当的依赖于时空坐标。对于一个全局对称性, 这些分量怎么也不能被规范掉, 此时将其解释为无质量 boson, 称为 Goldstone bosons。
³⁶ 局域 $SU(2)$ 对称在实验上不存在任何可观测的效应。这仅仅是我们方程的对称性, 规范自由度从不在任何实验测量量中出现。否则我们将无法做出预言, 因为我们将有无穷多个相同可能性的预言(它们由 $SU(2)$ 变换相联系)。尽管这样, 这个对称性也不是没用的, 它能将我们导向拉格朗日量的正确形式。

现在将用 2×2 单位阵表出的 B_μ 和 Pauli 矩阵的具体形式³⁷代入得

$$\begin{aligned}
 &= \frac{v^2}{8} \left| \begin{pmatrix} g'W_3^\mu + gB^\mu & g'W_1^\mu - ig'W_2^\mu \\ g'W_1^\mu + ig'W_2^\mu & -g'W_3^\mu + gB^\mu \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|^2 \\
 &= \frac{v^2}{8} \left| \begin{pmatrix} g'W_1^\mu - ig'W_2^\mu \\ -g'W_3^\mu + gB^\mu \end{pmatrix} \right|^2 \quad (7.84) \\
 &= \frac{v^2}{8} \left((g')^2 ((W_1^\mu)^2 + (W_2^\mu)^2) + (g'W_3^\mu - gB^\mu)^2 \right)
 \end{aligned}$$

用已有的场定义两个新的自旋为 1 的场：

$$W_+^\mu \equiv \frac{1}{\sqrt{2}}(W_1^\mu - iW_2^\mu) \quad (7.85)$$

$$W_-^\mu \equiv \frac{1}{\sqrt{2}}(W_1^\mu + iW_2^\mu) \quad (7.86)$$

即 W_+^μ 是 W_-^μ 的复共轭。(7.84)式中的第一项可以写作

$$(W_1^\mu)^2 + (W_2^\mu)^2 = 2(W^+)_\mu (W^-)^\mu \quad (7.87)$$

把前面的常数添上，有

$$\left(\underbrace{\frac{g'v}{2}}_{\equiv m_W} \right) (W^+)_\mu (W^-)^\mu \quad (7.88)$$

这就像一个典型的“质量”项。

(7.84)式中的第二项可以写作矩阵形式¹⁹

$$(g'W_3^\mu - gB^\mu)^2 = \begin{pmatrix} (W_3)_\mu & B_\mu \end{pmatrix} \underbrace{\begin{pmatrix} g'^2 & -gg' \\ -gg' & g^2 \end{pmatrix}}_{\equiv G} \begin{pmatrix} (W_3)_\mu \\ B_\mu \end{pmatrix} \quad (7.89)$$

为了将其解释为质量项，我们需要将矩阵 G 对角化³⁸。线性代数说我们需要先得到矩阵 G 的本征值 λ_1, λ_2 以及归一³⁹的本征向

$${}^{37}\sigma_i W_i = \begin{pmatrix} W_3 & W_1 - iW_2 \\ W_1 + iW_2 & -W_3 \end{pmatrix}$$

¹⁹译注：原文下式有误，已更正。

³⁸ 马上就能看到一个对角矩阵会带给我们和其他质量项长得一模一样的项。这使我们能将其对应的场诠释成实验上能够观测到的物理场。我们当然也可以就用 W_3^μ 和 B^μ 干活，但是却难以找到其直接的物理意义。

³⁹ 长度为一，即 $v \cdot v = 1$ 。

量 v_1, v_2 。简单的计算得到

$$\begin{aligned}\lambda_1 = 0 & \rightarrow v_1 = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g \\ g' \end{pmatrix} \\ \lambda_2 = g^2 + g'^2 & \rightarrow v_2 = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g' \\ -g \end{pmatrix}\end{aligned}$$

用本征向量排成的矩阵 M 可以将 G 对角化, 即 $G_{\text{diag}} = M^{-1}GM$, 其中

$$M = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \quad (7.90)$$

以及

$$G_{\text{diag}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & g^2 + g'^2 \end{pmatrix} \quad (7.91)$$

由于我们用的是归一的本征向量, 故矩阵 M 是正交的 ($M^T = M^{-1}$):

$$\begin{aligned}M^T M &= \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \\ &= \frac{1}{g^2 + g'^2} \begin{pmatrix} g^2 + g'^2 & gg' - gg' \\ gg' - gg' & g^2 + g'^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned} \quad (7.92)$$

由此在(7.89)式中插入两个单位阵 $1 = M^T M$:

$$\begin{aligned}& \left((W_3)^\mu \quad B^\mu \right) \underbrace{MM^T}_{=1} G \underbrace{MM^T}_{=1} \begin{pmatrix} (W_3)_\mu \\ B_\mu \end{pmatrix} \\ &= \left((W_3)^\mu \quad B^\mu \right) M \underbrace{M^T G M}_{=G_{\text{diag}}} M^T \begin{pmatrix} (W_3)_\mu \\ B_\mu \end{pmatrix}.\end{aligned} \quad (7.93)$$

剩下的任务即是计算 $M^T \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix}$, 以得到两个新的, 有对应质量

项的场：

$$\begin{aligned} M^T \begin{pmatrix} W_3^\mu & B_\mu \end{pmatrix} &= \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & g' \\ g' & -g \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix} \\ &= \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} gW_3^\mu + g'B^\mu \\ gW_3^\mu - g'B^\mu \end{pmatrix} \equiv \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix} \end{aligned} \quad (7.94)$$

由此将第二项改做²⁰

$$\begin{aligned} \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} G_{\text{diag}} \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix} &= \begin{pmatrix} A^\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g^2 + g'^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z_\mu \end{pmatrix} \\ &= (g^2 + g'^2)(Z^\mu)^2 + 0 \cdot (A^\mu)^2. \end{aligned} \quad (7.95)$$

综上，我们从自旋为 1 的无质量场 W_i^μ 和 B^μ 的拉格朗日量出发

$$(\partial_\mu - ig' \frac{1}{2} \sigma_i (W_\mu)_i - ig B_\mu) \Phi^\dagger (\partial^\mu + ig' \frac{1}{2} \sigma_i (W^\mu)_i + ig B^\mu) \Phi. \quad (7.96)$$

通过自发对称性破缺，我们在拉格朗日量中得到了可以被解释成质量项的新项

$$\underbrace{\frac{1}{8} v^2 g'^2 (W^+)_\mu (W^-)^\mu}_{=\frac{1}{2} M_W^2} + \underbrace{\frac{1}{8} v^2 (g^2 + g'^2) Z_\mu^2}_{=\frac{1}{2} M_Z^2} + \underbrace{\frac{1}{8} v^2 \cdot 0}_{\text{光子质量为 0}} \cdot A_\mu^2 \quad (7.97)$$

我们能看到其中一个自旋为 1 的场 A_μ 在对称性自发破缺后仍然是无质量的。这是电磁的光子场，现有的所有实验都表明光子是无质量的⁴⁰。注意对应于 Z-bosons 的 Z_μ 和对应于光子的 A_μ 是场 B_μ 和 W_μ^3 的正交组合。它们竟有一个同样的起源！

同样的推演可以用来在不破坏局域 $SU(2)$ 对称性的情况下得到自旋 $\frac{1}{2}$ 场的质量项，但在具体讨论这个之前，我们需要先谈谈自然界中一个有趣的事实：宇称破坏。

⁴⁰ 注意，在这一节中，我略去了一些重要的概念：超荷和 Weinberg 角。Weinberg 角 θ_W 由 $\cos(\theta_W) = \frac{g}{g^2 + g'^2}$ 或 $\sin(\theta_W) = \frac{g'}{g^2 + g'^2}$ 来定义。通过它可以简化本节中的一些记号。超荷略复杂一些，感兴趣的读者可以参考量子场论的标准教材，部分推荐数目见第9章末。

²⁰ 译注：从而我们再次得到了(7.84)式中的第二项——作者你绕了这么大一个弯子累不累啊？

7.4 Parity Violation 宇称破坏

大自然不遵守宇称对称性，即宇称破坏，是科学史上最为重大的发现之一。通俗地说，宇称破坏使一部分物理实验表现出与其镜像实验所不同的物理现象。吴健雄在“吴实验”中最早发现了宇称破坏现象。这是一个漂亮的实验，但它的具体实现与本节主题无关，因此我们将只关注实验结果。

“吴实验”发现，弱相互作用的媒介子（ W^+ , W^- , Z 玻色子）仅与左手粒子耦合。换句话说，只有左手粒子才会通过弱作用力相互作用，这一物理机制与宇称破坏的关系将在本节结尾处加以解释。所有在弱相互作用中产生的粒子都是“左手”的。中微子只参与弱相互作用，因此右手中微子可能根本不存在⁴¹。另一方面，所有其它粒子都能通过另外三种相互作用产生，因此都可以具有“右手性”。

目前为止，我们用“左手性”和“右手性”，标注了对象在经历洛伦兹变换时所依据的群表示。这是一个抽象的概念，但我们却可以实际测量粒子的手性，因为抽象的“手性”概念直接关联于直观的“螺旋度”概念⁴²。

其实在多数情况下，微观粒子并没有确定的手性，也不处于确定的左手或右手态，而是由同时包含两种手性分量的狄拉克旋量所描述。宇称破坏并不是理论预言，相反，它的发现震惊了整个物理学界。即使到了今天，大自然这一奇特的性质依旧令人费解。然而，将这一发现纳入我们的理论框架却并不困难。我们只需利用某些数学工具，确保处理弱相互作用时只考虑左手旋量即可。

我们用 χ, ξ 表示二分量外尔旋量，而 ψ 代表四分量狄拉克旋量：

$$\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \quad (7.98)$$

Ψ 是狄拉克旋量的双重态：

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad (7.99)$$

⁴¹ 我们很快会看到，有质量的左手粒子总会在传播过程中获得右手分量。实验表明中微子有质量，因此也应当包含右手分量。然而这一右手分量不参与任何已知相互作用。

⁴² 此处我们不讨论这一概念，因其细节与本节叙述的主题无关。读者只需知道类似实验具有可行性。更多精彩的讨论可以参见以下文献：Alessandro Bettini. *Introduction to Elementary Particle Physics*. Cambridge University Press, 2nd edition, 4 2014. ISBN 9781107050402

上面提到的“某些工具”，是投影算符 P_L :

$$P_L \psi = P_L \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix} \equiv \psi_L \quad (7.100)$$

投影算符可以用如下矩阵构造⁴³:

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (7.101)$$

矩阵 γ_5 被称为**手性算符**，因为单一手性态 $\begin{pmatrix} \chi_L \\ 0 \end{pmatrix}$ 或 $\begin{pmatrix} 0 \\ \xi_R \end{pmatrix}$ ，正是 γ_5 本征值为 -1 和 $+1$ 的两个本征态。

于是，投影算符 P_L 可以写为⁴⁴

$$P_L = \frac{1 - \gamma_5}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (7.102)$$

类似的，我们定义

$$P_R = \frac{1 + \gamma_5}{2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (7.103)$$

现在为了限制只有左手粒子参与弱相互作用，我们必须在拉氏量中，描述各类场与 W_μ^\pm 与 Z_μ 相互作用的每一项里都插入 P_L 。这些项已在7.2节中导出，最终结果参看方程(7.58)，我们将其重写如下：

$$\mathcal{L} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i\bar{\Psi}\gamma_\mu\sigma_j W_j^\mu\Psi - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \quad (7.104)$$

其中需要处理的一项是 $\bar{\Psi}\gamma_\mu\sigma_j W_j^\mu\Psi$ ，我们将 P_L 插入：

$$\rightarrow \mathcal{L} = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + i\bar{\Psi}\gamma_\mu\sigma_j W_j^\mu P_L \Psi - \frac{1}{4}(W_{\mu\nu})_i(W^{\mu\nu})_i \quad (7.105)$$

此处 P_L 作用于双重态，得到

$$P_L \Psi = \begin{pmatrix} P_L & 0 \\ 0 & P_L \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} (\psi_1)_L \\ (\psi_2)_L \end{pmatrix} \quad (7.106)$$

⁴³ 请回忆方程(6.13)中 γ_μ 矩阵的定义，切记并无 γ_4 矩阵存在。由于另有规范以 γ_4 表示 γ_0 。为避免符号混淆，此处特将这一矩阵记为 γ_5 。

⁴⁴ 读者可能会问，为什么我们要用这样复杂的方式定义 P_L ，而不是直接从矩阵形式出发呢？这样做的目的，是为了保留在其它基下工作的可能性，那时 γ_μ 的矩阵形式可以非常不同。（读者可以在8.10节中找到更多信息）注意到拉氏量中狄拉克旋量总与 γ_μ 相伴出现。在这两者之间，我们总可以插入 $1 = \mathcal{U}^{-1}\mathcal{U}$ ，其中 \mathcal{U} 为任意可逆矩阵。例如 $\partial_\mu \bar{\Psi} \gamma_\mu \Psi = \partial_\mu \bar{\Psi} \underbrace{\mathcal{U}^{-1}}_{=1} \underbrace{\gamma_\mu \mathcal{U}^{-1}}_{=1} \Psi = \partial_\mu \bar{\Psi}' \underbrace{\mathcal{U}^{-1}}_{=\Psi'} \underbrace{\gamma_\mu \mathcal{U}^{-1}}_{=\gamma_\mu'} \underbrace{\Psi}_{=\Psi'} = \partial_\mu \bar{\Psi}' \gamma_\mu' \Psi'$ 。物理事实当然独立于这类基矢变换，我们却可以利用它简化计算。本书正文偏好使用的基矢，被称为 Weyl 基。在其它基矢下，狄拉克旋量的两个分量是 χ_L 与 ξ_R 的混合。尽管如此，通过 $P_L = \frac{1-\gamma_5}{2}$ 定义的投影算符，总能将其投影为左手分量。这是因为 $P_L^{\text{Weyl}} \Psi^{\text{Weyl}} = \Psi_L^{\text{Weyl}} \Rightarrow P_L' \Psi' = \frac{1-\gamma_5'}{2} \Psi' = \frac{1-\gamma_5}{2} \mathcal{U} \Psi^{\text{Weyl}} = \frac{1-\gamma_5}{2} \mathcal{U} \Psi^{\text{Weyl}} = \mathcal{U} \Psi_L^{\text{Weyl}} = \Psi_L'$ ✓

我们断言，只需在表达式中插入一个 P_L ，就可以将 $\bar{\Psi}$ 和 Ψ 同时投影到左手分量上。为了理解这一点，我们需要以下三个等式：

⁴⁵ 投影算符的定义中还包括 $P_L P_R = P_R P_L = 0$ ，此处可用 P_L, P_R 和 γ_5 的矩阵形式验证成立。

⁴⁶ 或可用定义 $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ 结合以下等式证明： $\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = \frac{1}{2}\eta_{\mu\nu}$ ，其中 $\eta_{\mu\nu}$ 为 Minkowski 度规。

- $(P_L)^2 = P_L$ ，这可以从算符的矩阵形式立即看出，同时也是投影算符必须满足的性质⁴⁵。两次投影应等价于一次投影。
- $\{\gamma_5, \gamma_\mu\} = \gamma_5\gamma_\mu + \gamma_\mu\gamma_5 = 0$ ，这可以通过简单粗暴的展开计算加以证明⁴⁶。
- $(P_L)^\dagger = P_L$ ，因为 γ_5 是实矩阵。而这一点可以从 γ_5 的矩阵形式直接看出 $\gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ 。

第二个等式直接推出 $\gamma_5\gamma_\mu = -\gamma_\mu\gamma_5$ ，即我们可以交换 γ_5 和 γ_μ 矩阵的位置，只需加入一个额外的负号。于是我们有：

$$\gamma_\mu P_L = \gamma_\mu \frac{1 - \gamma_5}{2} = \frac{1 + \gamma_5}{2} \gamma_\mu = P_R \gamma_\mu \quad (7.107)$$

现在我们可以将方程(7.105)中的相关项写为：

$$\begin{aligned} \bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi &= \bar{\Psi} \gamma_\mu \sigma_j W_j^\mu (P_L)^2 \Psi \\ &= \underbrace{\bar{\Psi}}_{\Psi^\dagger \gamma_0} \underbrace{\gamma_\mu P_L \sigma_j W_j^\mu}_{P_R \gamma_\mu} \underbrace{P_L \Psi}_{\Psi_L} \\ &= \Psi^\dagger \underbrace{\gamma_0 P_R \gamma_\mu \sigma_j W_j^\mu}_{P_L \gamma_0} \Psi_L \\ &= \underbrace{(P_L \Psi)^\dagger}_{\Psi_L^\dagger} \gamma_0 \gamma_\mu \sigma_j W_j^\mu \Psi_L \end{aligned}$$

利用 $P_L^\dagger = P_L$ 以及 $(AB)^\dagger = ((AB)^T)^* = (B^T A^T)^* = B^\dagger A^\dagger$

$$\begin{aligned} &= \underbrace{(P_L \Psi)^\dagger}_{\Psi_L^\dagger} \gamma_0 \gamma_\mu \sigma_j W_j^\mu \Psi_L \\ &= \bar{\Psi}_L \gamma_\mu \sigma_j W_j^\mu \Psi_L \quad \checkmark \end{aligned} \quad (7.108)$$

⁴⁷ 这就是说，若一项实验的结果全由拉氏量中的这一项所决定，那么当实验设置全部以镜像方式反转时，实验结果也会有所不同。

⁴⁸ 旋量宇称算符在3.7.9节中导出。用 γ_μ 矩阵，我们可以将宇称算符写作 $P = \gamma_0 = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

⁴⁹ 矢量宇称算符即为 $P_{\text{vector}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ ，如方程(3.132)所述。

至此，我们已经知道了，如何在数学上描述只有左手粒子参与弱相互作用的物理事实。然而，为什么这就意味着宇称破坏？为了理解这一点，我们需要对相互作用项进行宇称变换。因为在宇称变换下这一项不具有不变性，由这一项描述的物理系统，也将和它的“镜像”有所不同⁴⁷。这里我们用到旋量⁴⁸和矢量⁴⁹的宇称算符 P_{spinor} ， P_{vector} 。在宇称变换下：

$$\underbrace{\bar{\Psi}}_{\Psi^\dagger \gamma_0} \gamma_\mu \sigma_j (W_j^\mu) P_L \Psi \rightarrow (P_{\text{spinor}} \Psi)^\dagger \gamma_0 \gamma_\mu \sigma_j (P_{\text{vector}} W^\mu)_j P_L (P_{\text{spinor}} \Psi)$$

$$\begin{aligned}
&= (\Psi)^\dagger \gamma_0 \gamma_0 \gamma_\mu \sigma_j (P_{\text{vector}} W^\mu)_j P_L \gamma_0 \Psi \\
&= \underbrace{(\Psi)^\dagger \gamma_0 \gamma_0 \gamma_\mu \sigma_j (P_{\text{vector}} W^\mu)_j P_R \Psi}_{\text{利用 } \{\gamma_5, \gamma_0\}=0 \text{ 以及 } P_L = \frac{1-\gamma_5}{2}} \\
&\quad (7.109)
\end{aligned}$$

我们可以通过矩阵运算得到 $\gamma_0 \gamma_0 \gamma_0 = \gamma_0$ 及 $\gamma_0 \gamma_i \gamma_0 = -\gamma_i$ 。同时，我们通过 P_{vector} 的表达式看出 $P_{\text{vector}} W^0 = W^0$ 及 $P_{\text{vector}} W^i = -W^i$ 。此处两负号相互抵消，于是拉氏量中的这一项，在宇称变换下变为

$$(P_{\text{spinor}} \Psi)^\dagger \gamma_0 \gamma_\mu (P_{\text{vector}} W^\mu)_j P_L (P_{\text{spinor}} \Psi) = \bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_R \Psi \neq \bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi \quad (7.110)$$

因此这一项不具有不变性，也就违反了宇称对称性。

⁵⁰ 详见附录 A 中的解释。

宇称破坏还有一个重要的推论。还记得，我们会把那些在变换下相互转变的东西⁵⁰竖着放在一起，夹在两个括号之间。比如我们讨论四维矢量，是因为它的分量在空间转动和推动下可以相互转化。在这一节中，我们知道只有左手粒子参与弱相互作用，而描述这一现象的正确拉氏量写作 $\bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi = \bar{\Psi}_L \gamma_\mu \sigma_j W_j^\mu \Psi_L$ 。从物理上说，这一项表明组成左手旋量双重态的两个分量，即两个自旋 1/2 场 $(\psi_1)_L, (\psi_2)_L$ 可以通过弱作用相互转化。而右手场不参与弱相互作用，因而 $(\psi_1)_R, (\psi_2)_R$ 无法相互转化。可见，我们没有理由将两个右手分量写成双重态的形式。在数学上，这意味着右手场构成 $SU(2)$ 单态，也就是说，它们在对称操作下将根据 $SU(2)$ 群的一维表示变换。从 3.6.3 节的讨论中我们知道，这等于说右手场实际上在变换下不变。最后我们加以总结：

- 左手场构成 $SU(2)$ 双重态： $\Psi_L = \begin{pmatrix} (\psi_1)_L \\ (\psi_2)_L \end{pmatrix}$ ，因为它们都参与弱相互作用，可以互相转化。它们根据 $SU(2)$ 群的二维表示变换：

$$\Psi_L \rightarrow \Psi'_L = e^{i\vec{a} \cdot \frac{\vec{\sigma}}{2}} \Psi_L \quad (7.111)$$

- 右手场由 $SU(2)$ 单态描述： $(\psi_1)_R, (\psi_2)_R$ ，因为它们不参与弱相互作用，因而不能相互转化。它们根据 $SU(2)$ 的一维表

示变换:

$$\begin{aligned}(\psi_1)_R &\rightarrow (\psi_1)'_R = e^0(\psi_1)_R = (\psi_1)_R \\(\psi_2)_R &\rightarrow (\psi_2)'_R = e^0(\psi_2)_R = (\psi_2)_R\end{aligned}\quad (7.112)$$

下面我们将继续讨论如何在保持原有对称性的前提下，于拉氏量中加入自旋 $\frac{1}{2}$ 粒子的质量项。

7.5 Lepton Mass Terms 轻子质量项

在7.2节开头，我们发现无法既保证 $SU(2)$ 对称性，又在拉氏量中加入任意的质量项 $\bar{\Psi}m\Psi$ 。在这一节里，我们会看到宇称破坏让问题变得更加复杂。最后我们将再次引入希格斯机制解决这一难题。

在上一节中，我们讨论了耦合项 $\bar{\Psi}\gamma_\mu\sigma_j W_j^\mu P_L\Psi$ 的手性。然而质量项的手性又如何呢？回顾我们在方程(6.7)与(6.8)中导出的，不含导数的旋量不变量：

$$I_1 := (\chi_a)^\dagger \xi^a = (\chi_L)^\dagger \xi_R \quad \text{和} \quad I_2 := (\xi^a)^T \chi_a = (\xi_R)^\dagger \chi_L \quad (7.113)$$

⁵¹ 如上节所述，狄拉克旋量 ψ_L 和 ψ_R 由手性投影算符定义： $\psi_L = P_L\psi$ 和 $\psi_R = P_R\psi$ 。注意 $\bar{\psi} = \psi^\dagger\gamma_0$

我们可以用狄拉克旋量将其写为⁵¹：

$$\begin{aligned}\bar{\psi}\psi &= \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L \\&= \begin{pmatrix} \chi_L^\dagger & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \xi_R \end{pmatrix} \\&= \chi_L^\dagger \xi_R + \xi_R^\dagger \chi_L \quad \checkmark\end{aligned}\quad (7.114)$$

我们看到，洛伦兹不变的质量项包含左手与右手场的乘积。上节结尾处提到，左手场与右手场在 $SU(2)$ 变换下有不同的变换性质，这就产生了问题。左手场是双重态，而右手场是单态。双重态和单态的乘积不再是 $SU(2)$ 不变量。如：

$$\underbrace{\bar{\Psi}}_{\text{双重态}} \underbrace{\psi_R}_{\text{单态}} \rightarrow \bar{\Psi}'_L \psi'_R = \bar{\Psi}_L e^{-ib_i \frac{\sigma_i}{2}} \psi_R \neq \bar{\Psi}_L \psi_R \quad (7.115)$$

基于处理自旋 1 场质量项的经验，我们重规叠矩，不加入上述质量项，而是在拉氏量中加入与 0 自旋场耦合，且满足 $SU(2)$

不变的相互作用项。此后通过选择 0 自旋场的真空期望值，破缺这一对称性，进而制造出质量项。

耦合 0 自旋双重态与自旋 $\frac{1}{2}$ 场，且保证 $SU(2), \mathcal{U}(1)$ 与洛伦兹不变性的项，由下式给出：

$$\bar{\Psi}_L \Phi \psi_R \quad (7.116)$$

为验证其不变性，我们将 $SU(2)$ 变换作用于上式⁵²

$$\bar{\Psi}^L \Phi \psi^R \rightarrow \bar{\Psi}'^L \Phi \psi'^R = \bar{\Psi}^L e^{-ib_i(x)\sigma_i} \Phi e^{ib_i(x)\sigma_i} \psi^R = \bar{\Psi}^L \Phi \psi^R \quad \checkmark$$

同时作用 $\mathcal{U}(1)$ 变换

$$\bar{\Psi}^L \Phi \psi^R \rightarrow \bar{\Psi}'^L \Phi \psi'^R = \bar{\Psi}^L e^{-ia(x)} \Phi e^{ia(x)} \psi^R = \bar{\Psi}^L \Phi \psi^R \quad \checkmark$$

0 自旋场（洛伦兹标量）在洛伦兹变换下不变⁵³，于是这一项为洛伦兹不变量，因为它的其余因子与方程(7.114)相同。

这类相互作用项被称为 **Yukawa 耦合**。我们将这一项与其厄米共轭结合，再乘以耦合常数⁵⁴ $-\lambda_2$ ，加入到拉氏量之中。

$$\mathcal{L} = -\lambda_2 (\bar{\Psi}^L \Phi \psi_2^R + \bar{\psi}_2^R \Phi \Psi^L) \quad (7.117)$$

这一额外项不仅描述了 fermion 与 Higgs 场的耦合，也在 $SU(2)$ 对称性破缺后导致了 fermion 质量项的产生。我们在真空期望附近展开，选择方程(7.77)

$$\Phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}$$

将其代入拉氏量，得到

$$\begin{aligned} \mathcal{L} &= -\frac{\lambda_2}{\sqrt{2}} \left((\bar{\Psi}_1^L, \bar{\Psi}_2^L) \begin{pmatrix} 0 \\ v + h \end{pmatrix} \psi_2^R + \bar{\psi}_2^R (0, v + h) \begin{pmatrix} \Psi_1^L \\ \Psi_2^L \end{pmatrix} \right) \\ &= -\frac{\lambda_2(v + h)}{\sqrt{2}} (\bar{\Psi}_2^L \psi_2^R + \bar{\psi}_2^R \Psi_2^L) \end{aligned}$$

由方程(7.114)可知上式等价于

$$= -\frac{\lambda_2(v + h)}{\sqrt{2}} \bar{\psi}_2 \psi_2 \quad (7.118)$$

⁵² 注意到 $\Phi^L \rightarrow \Phi'^L = e^{ib_i(x)\sigma_i} \Phi^L$ ，及 $\sigma_i^\dagger = \sigma_i$ 。

⁵³ 0 自旋场根据定义，应随洛伦兹群的 $(0,0)$ 表示变换。在这一表示中，所有洛伦兹变换都是恒等变换。在3.7.4中我们导出了这一结论。

⁵⁴ 稍后，我们会讨论包含 ψ_1^R 和 $-\lambda_1$ 的相互作用项。届时，读者就能明白为什么我们将耦合常数记为 $-\lambda_2$ ，以及下式为何只包含 ψ_2^R 。

$$= \underbrace{-\frac{\lambda_2 v}{\sqrt{2}} \bar{\psi}_2 \psi_2}_{\text{Fermion 质量项}} - \underbrace{\frac{\lambda_f h}{\sqrt{2}} \bar{\psi}_2 \psi_2}_{\text{Fermion-Higgs 相互作用}} \quad (7.119)$$

可见, Higgs 机制确实导致了质量项的产生。再一次, 我们根据对称性约束为拉氏量添加 Higgs 耦合项, 并在对称破缺后产生出自旋 $\frac{1}{2}$ 场的质量项。值得注意的是, 只有双重态中的第二个场 ψ_2 获得了质量项。那么第一个场 ψ_1 怎么办呢?

为了获得 Ψ_1 的质量项, 我们考虑电荷共轭⁵⁵Higgs 场 $\tilde{\Psi} = \epsilon \Psi^*$ 的耦合项。由于

$$\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \rightarrow \tilde{\Phi} = \epsilon \Phi^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} \quad (7.120)$$

对电荷共轭 Higgs 场重复上面的推导步骤, 即可获得 Ψ_1 的质量项

$$\begin{aligned} \mathcal{L} &= -\lambda_f (\bar{\Psi}^L \tilde{\Phi} \psi_1^R + \bar{\psi}_1^R \tilde{\Phi} \Psi^L) \\ &= -\frac{\lambda_1}{\sqrt{2}} \left(\begin{pmatrix} \bar{\Psi}_1^L, \bar{\Psi}_2^L \end{pmatrix} \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} \Psi_1^R + \bar{\Psi}_1^R \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix} \begin{pmatrix} \Psi_1^L \\ \Psi_2^L \end{pmatrix} \right) \\ &= -\frac{\lambda_1(v+h)}{\sqrt{2}} (\bar{\Psi}_1^L \Psi_1^R + \bar{\Psi}_1^R \Psi_1^L) \end{aligned}$$

⁵⁶ ν 永远代表中微子。在此处, 我们给出双重态 ψ_1 和 ψ_2 中两分量的传统命名: 电子场 e 和电子中微子场 ν_e 。

为更好地理解这一抽象的双重态, 我们将其重写如下⁵⁶:

$$\Psi = \begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad (7.121)$$

其它轻子, 即 μ, ν_μ 与 τ, ν_τ 也有等价的写法。这一写法是由实验结果所确定的, 因为在弱相互作用下, 电子要么转化为动量可能不同的另一电子, 要么变为电子中微子 ν_e 与其它粒子的集合。在弱相互作用中, e 和 ν_e (同样包括 μ, ν_μ 与 τ, ν_τ) 总是相伴出现。这一事实可以用耦合项 $\bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi$ 的数学形式加以解释。如前一节所述, 这一项可以用 Pauli 矩阵 σ_i 的矩阵形式加以重写⁵⁷, 并给出双重态两分量间的耦合项。

⁵⁷ 这再次给出 $\sigma_i W_i^\mu =$

$$\begin{pmatrix} W_3^\mu & W_1^\mu - iW_2^\mu \\ W_1^\mu + iW_2^\mu & -W_3^\mu \end{pmatrix}.$$

这可以用下式进行简化:

$$W_\pm = \frac{1}{\sqrt{2}}(W_1 \pm iW_2) \Rightarrow$$

$$\sigma_i W_i^\mu = \begin{pmatrix} W_3^\mu & \sqrt{2}W_+^\mu \\ \sqrt{2}W_-^\mu & -W_3^\mu \end{pmatrix}$$

$$\bar{\Psi} \gamma_\mu \sigma_j W_j^\mu P_L \Psi = \begin{pmatrix} \bar{\nu}_e & \bar{e} \end{pmatrix} \gamma_\mu \begin{pmatrix} W_3^\mu & \sqrt{2}W_+^\mu \\ \sqrt{2}W_-^\mu & -W_3^\mu \end{pmatrix} P_L \begin{pmatrix} \nu_e \\ e \end{pmatrix}$$

$$\begin{aligned}
& \underbrace{\left((\bar{\nu}_e)_L \quad (\bar{e})_L \right)}_{\text{使用方程(7.105)}} \gamma_\mu \begin{pmatrix} W_3^\mu & \sqrt{2}W_+ \\ \sqrt{2}W_- & -W_3^\mu \end{pmatrix} P_L \begin{pmatrix} (\nu_e)_L \\ (e)_L \end{pmatrix} \\
&= (\bar{\nu}_e)_L \gamma_\mu W_3^\mu (\nu_e)_L + (\bar{\nu}_e)_L \gamma_\mu \sqrt{2}W_+ (e)_L \\
&\quad + (\bar{e})_L \gamma_\mu \sqrt{2}W_- (\nu_e)_L - (\bar{e})_L \gamma_\mu W_3^\mu (e)_L
\end{aligned} \tag{7.122}$$

如果我们想同时考虑包括 e, μ, τ 的所有轻子产生过程，需将相应的三个耦合项都写入拉氏量：

$$\bar{\Psi}_e \gamma_\mu \sigma_j W_j^\mu P_L \Psi_e + \bar{\Psi}_\mu \gamma_\mu \sigma_j W_j^\mu P_L \Psi_\mu + \bar{\Psi}_\tau \gamma_\mu \sigma_j W_j^\mu P_L \Psi_\tau \tag{7.123}$$

引入 $\Psi_l = \begin{pmatrix} \nu_l \\ l \end{pmatrix}$ ，其中 $l = e, \mu, \tau$ ，上式可写为更紧凑的形式：

$$\bar{\Psi}_L \gamma_\mu \sigma_j W_j^\mu P_L \Psi_L$$

定义 $l = \begin{pmatrix} l_L \\ l_R \end{pmatrix}$ ，质量项可写为：

$$\underbrace{-\frac{\lambda_l v}{\sqrt{2}}(\bar{l}l)}_{\text{Fermion 质量项}} - \underbrace{\frac{\lambda_f h}{\sqrt{2}}(\bar{l}l)}_{\text{Fermion-Higgs 相互作用}}$$

对中微子也有等价的形式。

这一拉氏量可以用来预测有关 Higgs 场的实验现象。轻子质量由下式给出：

$$m_l = \frac{\lambda_l v}{\sqrt{2}} \rightarrow \lambda_l = \frac{m_l \sqrt{2}}{v} \tag{7.124}$$

而轻子与 Higgs 场的耦合强度为：

$$c_l = \frac{\lambda_l h}{\sqrt{2}} \underbrace{=}_{\text{方程(7.124)}} \frac{m_l \sqrt{2} h}{\sqrt{2} v} = \frac{m_l h}{v} \tag{7.125}$$

从这一方程可以看出，Higgs 场与给定轻子的耦合强度，正比于这一轻子的质量。轻子质量越大，耦合强度越强。运用类似的推

导，我们可将这一结论推广到所有粒子。

参与弱相互作用的自旋 $\frac{1}{2}$ 粒子还包括**夸克**。同时，夸克还参与第三种相互作用：强相互作用。这一内容将在7.8节中讨论。首先我们将讨论夸克的质量项，可喜的是，这一主题可用类似引入轻子质量的方法加以处理。

7.6 Quark Mass Terms 夸克质量项

在上一节中我们知道， $SU(2)$ 双重态包含了两种可通过弱作用力相互转化的粒子。就夸克⁵⁸而言，这两种粒子是上夸克和下夸克。

$$q = \begin{pmatrix} u \\ d \end{pmatrix} \quad (7.126)$$

类似的，奇夸克、粲夸克，和顶夸克、底夸克也各自组成双重态。

与前文一样，考虑到只有左手粒子参与弱相互作用的实验事实，我们构造左手双重态和右手单态：

$$\underbrace{q_L}_{\text{双重态}} = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \rightarrow e^{ia_i \frac{\sigma_i}{2}} q_L \quad (7.127)$$

$$\begin{aligned} \underbrace{u_R}_{\text{单态}} &\rightarrow u_R \\ \underbrace{d_R}_{\text{单态}} &\rightarrow d_R \end{aligned} \quad (7.128)$$

右手粒子依旧不参与弱相互作用，不能转化为任何其它粒子，只构成 $SU(2)$ 单态（单分量的数学对象）。

这里，我们遇到了与处理轻子质量时相同的疑难：为了得到洛伦兹不变量，我们需要组合左、右手旋量。然而，这样的组合却不满足 $SU(2)$ 不变性，这让我们转而选择引入 Higgs 机制。换言之，我们放弃在拉氏量中直接加入质量项：

$$\bar{q}_L u_R + \bar{q}_L d_R + \bar{u}_R q_L + \bar{d}_R q_L \quad (7.129)$$

因为它不满足 $SU(2)$ 不变性。相反，我们考虑夸克与 0 自旋

⁵⁸ 没听说过夸克？请看1.3节。

场 Ψ 的耦合项:

$$\lambda_u \bar{q}_L \tilde{\Phi} u_R + \lambda_d \bar{q}_L \Phi d_R + \lambda_u \bar{u}_R \tilde{\Phi} q_L + \lambda_d \bar{d}_R \Phi q_L \quad (7.130)$$

这里我们引入耦合常数 λ_u, λ_d 和电荷共轭 Higgs 双重态⁵⁹, 后者为上夸克提供了质量项⁶⁰。

可见, 这里的全部数学构造, 都类比于处理轻子质量时所用的方法: 我们在最小值附近展开 Higgs 场⁶¹, 代入拉氏量, 制造出质量项和夸克——Higgs 耦合项。

7.7 Isospin 同位旋

现在讨论 $SU(2)$ 对称性所对应的守恒量。自由拉氏量只具有全局对称性, 而加入相互作用项, 则为其引入了局域对称性。注意到全局对称性是局域对称性的特例, 因此全局对称性也存在于所有局域对称的拉氏量之中, 其对应的守恒量也在两种情况下都守恒。由于 Noether 定理, $SU(2)$ 对称性给出了一个新的守恒量: 同位旋。同位旋与电荷类似, 后者对应全局 $U(1)$ 对称性。

将 Noether 定理应用于内部对称性(见 4.5.5 节, 参考方程(4.56))得到,

$$\partial_0 \int d^3x \underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi)} \delta \Psi}_{=Q} = 0 \quad (7.131)$$

拉氏量在下述形式的变换下不变:

$$\Psi \rightarrow e^{ia_i \frac{\sigma_i}{2}} \Psi = (1 + ia_i \frac{\sigma_i}{2} + \dots) \Psi \quad (7.132)$$

于是无穷小变分写为 $\delta \Psi = ia_i \frac{\sigma_i}{2} \Psi$, 其中 a_i 为任意系数。可见每个生成元都对应了一个守恒量, 因为取 a_i 中任意两个为 0、第三个不为 0 时, 拉氏量同样具有不变性。比如 $a_2 = a_3 = 0$ 而 $a_1 \neq 0$, 或 $a_1 = a_2 = 0$ 而 $a_3 \neq 0$ 。当然, 取 $a_1 \neq 0, a_2 \neq 0$ 且 $a_3 \neq 0$ 时我们还能得到一个守恒量, 但它其实是上述与生成元对应的三个守恒量之和。换言之, 一共有三个独立的守恒量, 它们与 $SU(2)$ 生成元一一对应。

全局不变的自由拉氏量写为(方程(7.45)):

$$\mathcal{L}_{D_1+D_2} = i \bar{\Psi} \gamma_\mu \partial^\mu \Psi$$

⁵⁹ 即 方程(7.120),
 $\tilde{\Phi} = \epsilon \Phi^* = \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix}$

⁶⁰ 我们将双重态定义为 $\begin{pmatrix} u \\ d \end{pmatrix}$ 。

其与 $\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}$ 的乘积正比于 d 。

⁶¹ $\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}$, 电荷共轭 Higgs 场有类似结论。

⁶² 见方程(7.131)，与往常一样，定义中不含任意常数 a_i 。

对应守恒量为 Q_i 。举例来说，对电子—中微子双重态，我们有⁶²

$$\begin{aligned} Q_i &= i\bar{\Psi}\gamma_0\frac{\sigma_i}{2}\Psi \\ &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}^\dagger \underbrace{\gamma_0\gamma_0}_{=1}\frac{\sigma_i}{2}\begin{pmatrix} \nu_e \\ e \end{pmatrix} \end{aligned} \quad (7.133)$$

注意 Pauli 矩阵中只有 σ_3 是对角矩阵。因此对双重态 (ν_e, e) 的两个分量来说，只有 $i = 3$ 的守恒量才有确定的取值。而分量 ν_e 与 e 不是另两个生成元 σ_1 和 σ_2 的本征态。当然，可以自由选取基矢以保证如 σ_2 处于对角形式。这样一来，我们可以重新定义 ν_e 和 e ，最终结果也并无不同。这里的重点在于，尽管存在三个与生成元一一对应的守恒量，但每一时刻却只能用其中一个来标记粒子或量子态。

对 $i = 3$ 我们有

$$\begin{aligned} Q_3 &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}^\dagger \frac{\sigma_3}{2} \begin{pmatrix} \nu_e \\ e \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \nu_e \\ e \end{pmatrix}^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix} \\ &= \frac{1}{2}\nu_e^\dagger\nu_e - \frac{1}{2}e^\dagger e \end{aligned} \quad (7.134)$$

于是可用 $Q_3(\nu_e) = \frac{1}{2}$ 及 $Q_3(e) = -\frac{1}{2}$ 分别标记两种粒子。反过来说，对 $i = 1$ ，我们有

$$\begin{aligned} Q_1 &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}^\dagger \frac{\sigma_1}{2} \begin{pmatrix} \nu_e \\ e \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \nu_e \\ e \end{pmatrix}^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix} \\ &= \frac{1}{2}\nu_e^\dagger e - \frac{1}{2}e^\dagger \nu_e \end{aligned} \quad (7.135)$$

矩阵 σ_1 不是对角阵，也就无法用来标记粒子。

7.7.1 Labelling States 标记量子态

在3.5节中,我们介绍了 Cartan 生成元的概念,它们是给定群的对角生成元。从上一节中我们知道,在为双重态中的粒子新增标记时,这些生成元十分有用。⁶³典型的 $SU(2)$ 双重态有以下形式

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad (7.136)$$

$SU(2)$ 只有一个 Cartan 生成元 $J_3 = \frac{1}{2}\sigma_3$, 它的本征值为 $+\frac{1}{2}$ 和 $-\frac{1}{2}$ 。(左手)中微子 $\begin{pmatrix} \nu_e \\ 0 \end{pmatrix}$ 是这一生成元的本征态,具有本征值 $+\frac{1}{2}$ 。这些本征值构成了粒子的新标记,亦即中微子和电子的同位旋。

按照同样的方法,我们可以用同位旋标记右手单重态。这些态随 $SU(2)$ 的一维表示变换。一维表示的生成元就是 0^{64} : $J_i = 0$ 。因此这些单态正是 Cartan 生成元 J_3 本征值为 0 的本征态。总之,右手单重态,如 e_R , 具有 0 同位旋。这与前文中右手场不参与弱相互作用的事实前后呼应。正如不带电的对象不参与电磁相互作用,不带同位旋的场也不参与弱相互作用。

最后,我们为规范场 $W_+^\mu, W_-^\mu, W_3^\mu$ 赋予同位旋,它们构成 $SU(2)$ 三重态:

$$W^\mu = \begin{pmatrix} W_+^\mu \\ W_-^\mu \\ W_3^\mu \end{pmatrix} \quad (7.137)$$

随 $SU(2)$ 的三维表示变换。此时, Cartan 生成元 J_3 有三个本征值: $+1, -1, 0^{65}$, 于是我们得出 $Q_3(W_+^\mu) = 1, Q_3(W_-^\mu) = -1, Q_3(W_3^\mu) = 0$ 。这便是 W_+ 与 W_- 玻色子的同位旋。

应注意,三重态 $\begin{pmatrix} W_1^\mu \\ W_2^\mu \\ W_3^\mu \end{pmatrix}$ 对应另一组基矢,其中 J_3 不是对角

阵。这是引入 W_\pm^μ 的另一个原因。

为了进一步阐明个中联系,我们回顾引入规范场 W_i^μ 的过程。在拉氏量中,这三个规范场与群的生成元相组合,以 $\sigma_i W_i^\mu$ 的形式出现。这可以看做对象 W^μ 以 σ_i 为基的线性展开: $W^\mu = \sigma_i W_i^\mu = \sigma_3 W_3^\mu + \sigma_2 W_2^\mu + \sigma_1 W_1^\mu$, 如同矢量展开为基矢量的线性

⁶³ 在后续章节中,我们将以同样的方法处理三重态,及 $SU(3)$ 不变性对应的守恒量。

⁶⁴ 3.7.4节解释了右手单重态不参与变换的原因。

⁶⁵ 这可由方程(3.121)中 J_3 的矩阵形式直接导出: $J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 。

⁶⁶ 李代数构成向量空间！

⁶⁷ 准确的说，我们要求映射是同态映射，应满足一些限制条件。

⁶⁸ 群本身不是向量空间。我们可以研究群如何作用于自身，但这不是群表示，而是群的具体化 (realization)。

$$^{69}\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$^{70}W^\mu = \begin{pmatrix} W_1^\mu \\ W_2^\mu \\ W_3^\mu \end{pmatrix}$$

组合： $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$ 。由于生成元是 $SU(2)$ Lie 代数中的元素⁶⁶， W^μ 也应是同一 Lie 代数中的元素。因此， W^μ 的变换规律将由 $SU(2)$ 在自身 Lie 代数上的表示所确定。也就是由 $SU(2)$ 的元素如何作用于自身 Lie 代数中的元素（生成元）所确定。这一观念看似奇特，实则自然而然。请读者回忆群表示的定义：群表示，是从群到某一线性空间中所有线性算符集合的映射，这些线性算符也构成线性空间⁶⁷。目前为止，我们只关注了“外在”的线性空间，如 Minkowski 空间。而与群相关的唯一⁶⁸内生线性空间却正是群的李代数！由此可见，研究群在这一线性空间上的表示并不奇怪。这一具有特别重要性的群表示，被称为群的伴随表示。

规范场（如 W_+, W_-, W_3 ）是规范群伴随表示中的元素。 $SU(2)$ Lie 代数是三维的，因为它由三个生成元生成，并具有三维的伴随表示。正如矢量的分量自上而下写在左右括号之间⁶⁹，我们也可以将 W^μ 的三个分量自上而下写在括号之间⁷⁰，并称之为三重态。这些伴随表示下的生成元，可以通过基矢变换与前面导出的三维生成元相联系。

在下一节中，我们将讨论“更高阶”的内部对称群 $SU(3)$ 。我们将看到，要求拉氏量满足局域 $SU(3)$ 不变性，可以导出描述强相互作用的拉氏量。

7.8 $SU(3)$ Interactions $SU(3)$ 相互作用

和上一节中在两个 fermion 场和 $SU(2)$ 中干的一样，我们可以在三个 fermion 场中找到一个局域 $SU(3)$ 不变的拉格朗日量。这个对称性不会破缺，对应于自旋为 1 的无质量胶子场。 $SU(3)$ 是由全体行列式为 1 的 3×3 么正矩阵构成的群，即

$$U^\dagger U = U U^\dagger = I \quad \det U = 1. \quad (7.138)$$

如通常处理 Lie 群时一样，我们能将其写成指数函数⁷¹

$$U = e^{iT_A \theta_A}. \quad (7.139)$$

同 $SU(2)$ 的情形一样⁷²，该群的定义式(7.138)要求其生成元是厄米且无迹的。

$$T_A^\dagger = T_A \quad (7.140)$$

⁷¹ 在2.4中已经提到了，大写罗马字母 A, B, \dots 总是要从 1 到 8 求和。

⁷² 见(3.80)式以及其后的文字，其基底由 2×2 Pauli 矩阵给出。

$$\text{Tr}(T_A) = 0 \quad (7.141)$$

一组这样的无迹厄米的基底,至少在一组表象下,由如下八个⁷³ 3×3 个矩阵给出,称为 Gell-Mann 矩阵:

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \quad (7.142)$$

该群的生成元与 Gell-Mann 矩阵通过 $T_A = \frac{1}{2}\lambda_A$ 相联系,就如同 Pauli 矩阵与 $SU(2)$ 群的生成元⁷⁴相关联一样。这个群的 Lie 代数由下式给出

$$[T^A, T^B] = if^{ABC}T^C, \quad (7.143)$$

这里仍约定诸如 A, B, C 的大写字母取遍 1 到 8 的全部值。 f^{ABC} 称为 $SU(3)$ 群的结构常数,对于 $SU(2)$ 群它就是 Levi-Civita 符号 ϵ_{ijk} 。这些系数可以直接暴力算出⁷⁵

$$f^{123} = 1 \quad (7.144)$$

$$f^{147} = -f^{156} = f^{246} = f^{257} = f^{345} = -f^{367} = \frac{1}{2} \quad (7.145)$$

$$f^{458} = f^{678} = \frac{\sqrt{3}}{2}, \quad (7.146)$$

剩下的都可以根据结构常数 f^{ABC} 对于指标是交换反对称的这一性质算出。例如

$$f^{ABC} = -f^{BAC} = -f^{CBA}. \quad (7.147)$$

还剩下一些不能通过重新给指标排序得到的项都是零。

⁷³ 一般的, $SU(N)$ 群有 $N^2 - 1$ 个生成元。生成元的数目常被称为群的秩 (rank)。

⁷⁴ $J_i = \frac{\sigma_i}{2}$, 见(3.82)式以及其后的解释。

⁷⁵ 这个结果不是很优美,但为了完整性我们还是将其列出

类似于我们在7.2节中对 $SU(2)$ 干过的事, 引入自旋 $\frac{1}{2}$ 场的三重态

$$Q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \quad (7.148)$$

同 $SU(2)$ 的情形一样, 我们给三重态中的对象赋予了新的指标, 在下一小节中会仔细讨论。

为了使拉格朗日量

$$\mathcal{L} = i\bar{Q}\partial_\mu\gamma^\mu Q - \bar{Q}mQ \quad (7.149)$$

局域 $SU(3)$ 不变要求添加自旋 $\frac{1}{2}$ 场和新的自旋 1 场的耦合项。推导类似于 $SU(2)$ 的情形, 但具体计算会更加麻烦, 所以我们仅给出最后的拉格朗日量⁷⁶

$$\mathcal{L} = -\frac{1}{4}F_{\alpha\beta}^A F_A^{\alpha\beta} + \bar{Q}(iD_\mu\gamma^\mu - m)Q, \quad (7.150)$$

自旋 1 胶子场 G_α^A 的场强张量 $F_{\alpha\beta}^A$ 定义为

$$F_{\alpha\beta}^A = \partial_\alpha G_\beta^A - \partial_\beta G_\alpha^A - gf^{ABC}G_\alpha^B G_\beta^C \quad (7.151)$$

其中 f^{ABC} 为前面在生成元的对易子中出现过的 $SU(3)$ 群的结构常数。此外, D_α 定义为

$$D_\alpha = \partial_\alpha + igT^C G_\alpha^C \quad (7.152)$$

其中 T_C 是此节一开始定义的 $SU(3)$ 群的生成元。可以看到, 除了生成元和对易关系有一点区别, 这里每一项都完全和 $SU(2)$ 的情形相类似。

7.8.1 Color 色

类似于7.7节中对 $SU(2)$ 的讨论, 我们可以通过 Noether 定理从 $SU(3)$ 对称性中得到新的守恒量。完全相同的思路, 这里我们有 8 个守恒量, 每个生成元对应一个。同样也只能使用对角生成元的守恒量来标记粒子。 $SU(3)$ 群有两个 Cartan⁷⁷生成元 $\frac{1}{2}\lambda_3$ 和 $\frac{1}{2}\lambda_8$ 。是故, 每个通过强作用力相互作用的粒子携带两个额外的指

⁷⁶ 记得对大写字母 (A, B, C, \dots) 从 1 到 8 求和。作者貌似已经重复三遍了

⁷⁷ 回忆, Cartan 生成元 = 对角生成元。

标, 对应于这两个 Cartan 生成元的本征值。

$$\frac{1}{2}\lambda_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ 的本征值}^{78} \text{ 是 } +\frac{1}{2}, -\frac{1}{2} \text{ 和 } 0。$$

$$\text{对于}^{21}\frac{1}{2}\lambda_8 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \text{ 其本征值}^{79} \text{ 为 } \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}} \text{ 和 }$$

$$\frac{-1}{\sqrt{3}}。$$

若将强相互作用 fermions 配成三重态 (在由 Cartan 生成元的本征向量构成的基下), 可以给它们指定如下带有任意旋量 ψ 的指标:

$$\left(+\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \text{ 给 } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \psi$$

人们常定义红色 (red): $= \left(+\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$ 。这说的是具有 $\begin{pmatrix} \psi \\ 0 \\ 0 \end{pmatrix}$ 形式

的东西叫做红的。

类似的

$$\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \text{ 给 } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \psi$$

并定义蓝色 (blue): $= \left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$, 同样的绿色 (green): $= \left(0, \frac{-1}{\sqrt{3}}\right)$ 。将这些东西称作颜色的想法来自于如下事实: 如果将三种色加在一起, 即

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \psi,$$

$$^{78} \text{ 本征向量自然是 } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\text{和 } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}。$$

$$^{79} \text{ 分别对应本征向量 } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ 和 } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}。$$

²¹译注: 原文少了一 $\frac{1}{2}$, 已更正。

我们便得到一个零荷的态（无色态），因为

$$\lambda_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \quad \text{以及} \quad \lambda_8 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0,$$

就像包含了所有颜色光的日光却是无色的一样。

完全类似于我们在 $SU(2)$ 上干的，将全体 $SU(3)$ 单态的色荷取做零，这样这些粒子间就不存在强相互作用了。细节上有点区别：它们是无色的。另外，也能通过完全类似于7.7节中给 W-Bosons 指定赝自旋的步骤，利用 $SU(3)$ 群的（8 维⁸⁰）伴随表示来给规范场 G_A^μ （胶子）来指定颜色。

⁸⁰ 由于我们有 8 个生成元，所以说 $SU(3)$ 群的伴随表示是 8 维的。

7.8.2 Quark Description 夸克描述

⁸¹ 由自旋 $\frac{1}{2}$ 场产生，在第6章中已经详细讨论

通过强力相互作用的自旋 $\frac{1}{2}$ 粒子⁸¹称为夸克。如果要对夸克进行讨论，我们得考虑一大堆事情：

- 夸克们是 $SU(3)$ 的三重态，用 Q 表示。一个三重态里有不同

颜色的同类夸克，例如上夸克： $u = \begin{pmatrix} u_r \\ u_g \\ u_b \end{pmatrix}$ 。为了保证 $SU(3)$

不变，三重态总是以 $\bar{Q}Q$ 的形式成对出现。我们也可以用指标号将其写作 $\bar{Q}Q = \bar{q}_c q_c$ ，其中指标 c 代表色 $c = r, g, b$ 。

- 另外，夸克之间还有弱相互作用，所以也是 $SU(2)$ 的二重态。二重态⁸²中的每一个元（= 每一个夸克）是一个三重态：

$q = \begin{pmatrix} u_c \\ d_c \end{pmatrix}$ 。这很是困扰人，所以说除非需要考虑强相互作用，色指标 c 通常省略。

- 这还不够，要记得每一个夸克都由一个 Dirac 旋量描述，这又是一个二分量的对象 $u_c = \begin{pmatrix} (\chi_u^L)_c \\ (\xi_u^R)_c \end{pmatrix}$ 。上分量描述右手夸克而下分量描述一个同样的左手夸克。

说了这些，再回到 $SU(3)$ 相互作用上来。令人开心的是，全部实验都表明 $SU(3)$ 的规范 boson，即胶子，是无质量的，故不需要给拉格朗日量添加一项规范 boson 的质量项。那么 $SU(3)$ 就不需要破缺。

除此之外，fermion 的质量项在 $SU(3)$ 对称性下也不是问题，

⁸² 每一个二重态里有两个不同的夸克，例如一个上夸克和一个下夸克，或者一个顶夸克和一个底夸克。

因为如下形式的项

$$\bar{Q}mQ \quad (7.153)$$

当三重态中的全部粒子都是等质量时是 $SU(3)$ 不变的。即 m 正比于单位阵⁸³。三重态中的东西描述了处于不同颜色的同样的夸克，应当有着相同的质量。例如，对于一个上夸克，其三重态为

$$U = \begin{pmatrix} u_r \\ u_b \\ u_g \end{pmatrix}, \quad (7.154)$$

其中 u_r 、 u_b 、 u_g 分别代表一个红色、蓝色和绿色的上夸克，都有着相同的质量。

其他的自旋 $\frac{1}{2}$ 粒子，像是电子和中微子，不携带色，故不能与胶子相耦合。由于其耦合常数比电磁相互作用 ($U(1)$) 和弱相互作用 ($SU(2)$) 大得多，从局域 $SU(3)$ 不变性导出的相互作用称为强相互作用 (strong interaction)。

$${}^{83}m = m \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ 而不是 } m = \begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}$$

7.9 The Interplay Between Fermions and Bosons Bosons 和 Fermions 间的游戏

在这一节我们将使用更加物理的语言，对本章得到的所有内容作一个总结。我们之后将学到自旋 $\frac{1}{2}$ 场能产生/湮灭自旋 $\frac{1}{2}$ 粒子，自旋 1 的场也能产生和湮灭自旋 1 的粒子。这一章中我们导出用于描述不同的场以及对应的粒子相互作用的拉格朗日量。

在1.3节中就已经提到，我们称自旋 $\frac{1}{2}$ 粒子为 fermion，自旋为 1 粒子为 boson。标准的说法是 fermion 构成了物质而 boson 作为物质间相互作用的媒介。现在我们能搞明白这句话说的是什么了。

这一章从第6章得到的自由场的拉格朗日量出发，发现了单个、两个、或三个自由的自旋 $\frac{1}{2}$ 场的拉格朗日量的内禀对称性；这些都只是全局对称性，由于狭义相对论的约束看上去相当不可信；而局域对称性则显得更加自然。

接着，我们发现能通过添加耦合项来保持拉格朗日量局域不变。这些耦合项描述了原有的自旋 $\frac{1}{2}$ 场和新的自旋为 1 的场的相互作用。由于一些历史上的原因，这里的内禀对称性被称为规范

对称性，相应的自旋为 1 的场也称为规范场。通过 Noether 定理，我们能从每一个内禀对称性得到一个新的守恒量。类似于从 $U(1)$ 对称性中得到的电荷，它们被诠释为荷。

- 为了得到一个局域 $U(1)$ 不变的拉格朗日量，我们需要一个规范场 A^μ 。最终的拉格朗日量正确的描述了电磁相互作用。 $U(1)$ 对称性告诉我们电荷是守恒的。
- 为了得到一个局域 $SU(2)$ 不变的拉格朗日量，我们需要三个规范场 $W_1^\mu, W_2^\mu, W_3^\mu$ 。最终的拉格朗日量正确的描述了弱相互作用。 $SU(2)$ 对称性告诉我们同位旋是守恒的。
- 为了得到一个局域 $SU(3)$ 不变的拉格朗日量，我们需要八个规范场 G_1^μ, G_2^μ, \dots 。最终的拉格朗日量正确的描述了强相互作用。 $SU(3)$ 对称性告诉我们色是守恒的。

不同种类的力由不同的 boson (自旋为 1 的粒子) 负责。电磁力的媒介是光子，由 $U(1)$ 规范场 A_μ 产生。弱力的媒介是 W^+ 、 W^- 和 Z boson，而强力是 8 个不同的胶子，分别由相应的 $SU(2)$ 和 $SU(3)$ 规范场产生。

此外，我们发现 $SU(2)$ 对称性禁止拉格朗日量中的质量项。而从实验中我们得知这是不对的。在不破坏任何对称性的前提先添加质量项的解决方案是 Higgs 机制。它通过添加描述已有的自旋为 1 场、自旋 $\frac{1}{2}$ 场与一个新的零自旋场 (Higgs 场) 的相互作用来实现。通过 $SU(2)$ 对称性的自发破缺，并将 Higgs 场在一个新的非对称极小点处展开，我们能得到拉格朗日量中所需要的质量项。



Applications 应用

8	量子力学	185
8.1	对应到粒子理论	
8.2	相对论性能量-动量关系	
8.3	量子力学的数学形式	
8.4	Schrödinger 方程	
8.5	从波动方程到粒子的运动	
8.6	Heisenberg 不确定性原理	
8.7	对几种诠释的评议	
8.8	附录：Dirac 旋量分量的诠释	
8.9	附录：解 Dirac 方程	
8.10	附录：Dirac 旋量的不同的基	
9	量子场论	219
9.1	对应到场理论	
9.2	自旋为 0 的自由场理论	
9.3	自由的自旋 $\frac{1}{2}$ 理论	
9.4	自旋 1 的自由场理论	
9.5	相互作用场论	
9.6	Klein-Gordon 方程的一般解	
10	经典力学	243
10.1	相对论力学	
10.2	非相对论力学的拉格朗日量	
11	电动力学	249
11.1	齐次 Maxwell 方程组	
11.2	Lorentz 力	
11.3	Coulomb 势	
12	引力	255
13	结束语	261

“中用的东西都不中看”

– Theophile Gautier
in *Mademoiselle de Maupin*.

"There is nothing truly
beautiful but that
which can never be
of any use whatsoever;
everything useful
is ugly."
Mademoiselle de
Maupin. Wildside
Press, 11 2007. ISBN
9781434495556

8. Quantum Mechanics 量子力学

Summary 总结

本章讲述量子力学，基于第5章讲过的对应关系¹，由此首先可导出相对论性能量-动量关系。

在建立量子力学的基本框架后，我们就对 Klein-Gordon 方程（最简单的标量型粒子的运动方程）取非相对论极限，得到著名的 **Schrödinger 方程**，方程的解被诠释为概率幅，随后我们用**波动力学**的方法分析两个简单例子。

之后呢，我们引入了 **Dirac 符号**，这对于理解量子力学的结构十分有帮助。系统的初始态被一个抽象的态矢量 $|i\rangle$ 标记，称之为右矢。测量该初始态得到某一特定末状态的概率幅可以用左矢（记作 $\langle f|$ ，表示末态）乘以右矢 $|i\rangle$ 计算。左矢与右矢的乘积 $\langle f|i\rangle$ 是一个复数，即为状态 $i \rightarrow f$ 的概率幅，发生这个过程的概率就是 $|\langle f|i\rangle|^2$ 。之后讨论**投影算符**。我们会看到它是如何与**完备性关系**一起用来把任意状态用任意一个算符的本征态来进行展开的。之前常用的波动力学方法可以视为将态矢在坐标基底展开的特例。在 Dirac 符号下，Schrödinger 方程用于计算态的时间演化。为了阐明这一联系，我们会用 Dirac 符号来重新讨论一个前面用波动力学求解过的例子。

¹物理量 \rightarrow 对称性生成元，见(5.1)式。

8.1 Particle Theory Identifications 对应到粒子理论

¹Klein-Gordon, Dirac, Proca, Maxwell 方程

² 见(3.240), (3.244)和第5章。

前面导出的方程¹可以用在粒子理论和场理论中。本章考虑它们应用于粒子理论的情形。因此动力学变量就是粒子（们）的坐标、能量和动量。第5章讲过，这些物理量可视为相应对称性的生成元：²

- 动量 $\hat{p}_i = -i\partial_i$
- 坐标 $\hat{x}_i = x_i$
- 能量 $\hat{E} = i\partial_0$
- 角动量 $\hat{L}_i = i\frac{1}{2}\epsilon_{ijk}(x^j\partial^k - x^k\partial^j)$

在讨论这些算符是如何应用于量子力学之前，我们先用它们来将现代物理最重要的方程之一导出。

8.2 Relativistic Energy-Momentum Relation 相对论性能量-动量关系

第6.2节推导了自旋 0 自由场的运动方程——Klein-Gordon 方程：

$$(\partial_\mu\partial^\mu + m^2)\Phi = 0,$$

利用上一节的那些对应关系³,

$$\begin{aligned} (\partial_\mu\partial^\mu + m^2)\Phi &= (\partial_0\partial_0 - \partial_i\partial_i + m^2)\Phi \\ &= \left(\left(\frac{1}{i}E\right)\left(\frac{1}{i}E\right) - \left(-\frac{1}{i}p_i\right)\left(-\frac{1}{i}p_i\right) + m^2\right)\Phi \\ &= (-E^2 + \vec{p}^2 + m^2)\Phi = 0 \end{aligned} \quad (8.1)$$

$$\rightarrow E^2 = \vec{p}^2 + m^2 \quad \text{或者用 4-矢量表示: } p_\mu p^\mu = m^2 \quad (8.2)$$

这就是著名的狭义相对论能量-动量关系 (energy-momentum relation)。上式应用于静止粒子 ($\vec{p} = 0$) 即得到 Einstein 著名的质能方程：

$$E^2 = m^2 \rightarrow E = mc^2$$

其中为了明确起见我们把 c^2 写了回来。现在就理解了 Poincaré 群的第一个 Casimir 算符 $p_\mu p^\mu$ 的标量值 (见(3.258)式) 为何要用 m^2 这样的符号来标记了。 $p_\mu p^\mu$ 就代表着粒子质量的平方, 而通过实验, 比如说测量粒子的能量和动量, 就可测量质量 $m = \sqrt{E^2 - \vec{p}^2}$ 。同理, 我们也可以理解为什么6.2节要把拉格朗日量里面的系数常数记作 m^2 。

8.3 The Quantum Formalism 量子力学的数学形式

前面说过, 物理量用算符表示, 这些算符作用于什么对象? 首先需要注意每个算符都有一组本征函数, 就像矩阵的本征向量一样。矩阵是有限维的, 因此矩阵有有限多个本征向量; 而算符通常作用在无限维的矢量空间中, 而得到无穷多个本征函数。比如, 动量算符的本征函数满足的本征方程为

$$\underbrace{-i\partial_i}_{\text{算符}} \Psi = \underbrace{p_i}_{\text{本征值}} \underbrace{\Psi}_{\text{本征函数}} \quad (8.3)$$

其中, p_i 是一个数。一个显然的解为

$$\underbrace{C}_{=\text{常数}} e^{ip_i x_i} \rightarrow -i\partial_i C e^{ip_i x_i} = p_i C e^{ip_i x_i} \quad \checkmark \quad (8.4)$$

但要注意, 对于任意的 p_i , 这都是一个解。因此我们发现有无无穷多个动量算符 $\hat{p}_i = -i\partial_i$ 的本征函数。能量本征函数也能用同样的办法来得到

$$i\partial_0 \Phi = E\Phi \quad (8.5)$$

角动量的本征函数也是一样⁴。类比于矩阵的本征矢量, 这些本征函数可以看做基⁵, 而这就意味着我们可以把任意的函数 Ψ 按照这些基进行展开。例如, 在动量本征函数⁶进行展开 (简明起见, 考虑一维情形)

$$\Psi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp \Psi_p e^{-ipx} \quad (8.6)$$

⁴ 角动量本征函数的求解更加复杂, 因为角动量算符比其它算符本身就更复杂。我们找不到一组三个角动量分量共同的本征态, 因为 $[\hat{L}_i, \hat{L}_j] \neq 0$ 。这在后面就会详细讨论, 最终选取的本征函数是角动量第三个方向 \hat{L}_3 (和角动量算符的平方 \hat{L}^2 , 它与其它的角动量分量都对易 $[\hat{L}^2, \hat{L}_j] = 0$) 的完备归一基, 即著名的球谐函数基。

⁵ 在矩阵中, 本征矢量是矩阵对应的矢量空间的基

⁶ 注意, 这其实就是我们在附录D.1中介绍的 Fourier 变换, 因子 $\frac{1}{\sqrt{2\pi}}$ 是出于习惯约定。

其中, Ψ_p 是展开系数, 类似于矢量 $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$ 中的 v_1, v_2, v_3 。对于一些系统来说, 我们的边界条件使得最后得到的基是离散而不连续的。我们可以, 比如说, 用能量本征态 Φ_{E_c} 进行展开:

$$\Psi = \sum_n c_n \Phi_{E_n} \quad (8.7)$$

注意, 一般来说, 一个算符的本征函数并不是另一个算符的本征函数。只有当两个算符对易, 即 $[A, B] = AB - BA = 0$ 时, 才可以找到一组同时是两个算符本征函数的基。下面证明这个命题。设 $[C, D] \neq 0 \rightarrow CD \neq DC$, 而对于 C 的某个本征函数 Ψ , 有 $C\Psi = c\Psi$, 其中 c 是它的本征值。如果 Ψ 同时也是 D 的本征函数 $D\Psi = d\Psi$, 则有

$$\begin{aligned} CD\Psi &= C d\Psi = \underbrace{dC\Psi}_{\text{因为 } d \text{ 只是一个数}} = dc\Psi \\ DC\Psi &= D c\Psi = cD\Psi = cd\Psi = \underbrace{cd\Psi}_{\text{因为数之间互相对易}} = dc\Psi \end{aligned}$$

$$\rightarrow DC = CD \quad \text{这与 } [C, D] \neq 0 \text{ 矛盾, 因而不存在共同的本征函数。} \quad (8.8)$$

⁷ 这里使用 Ψ 是出于量子力学的惯例, 尽管前面只用它表示描述自旋 0 粒子的旋量。

一般来说, 物理量算符作用在描述物理系统的态 (state)⁷ Ψ 上, 通过求解运动方程得到 Ψ 。

一般的解按照某组基展开都会超过一项。举个例子, 考虑两个能量本征态² $\Psi = c_1\Phi_{E_1} + c_2\Phi_{E_2}$ ⁸, 能量算符作用于它可得

$$\hat{E}\Psi = \hat{E}(c_1\Phi_{E_1} + c_2\Phi_{E_2}) = c_1E_1\Phi_{E_1} + c_2E_2\Phi_{E_2} \neq E(c_1\Phi_{E_1} + c_2\Phi_{E_2}) \quad (8.9)$$

由此可见不同能量 (的能量本征态) 叠加 (superposition) 形成的态一般不是能量算符的本征态, 因为本征态需要满足定义——存在某个数 E 使得 $\hat{E}\Psi = E\Psi$ 。不过, 处于 Ψ 状态系统的能量是多少呢? 两个能量本征态的叠加态又意味着什么呢? 它的物理意义是什么?

⁸ 这意味着所有的展开系数 $\Psi = \sum_n c_n \Phi_{E_n}$ 是零。

²即能量算符 (哈密顿算符) \hat{H} 的本征态, 能量本征值为 E 的本征态记作 Φ_E 。

首先我们能从拉格朗日量的 $\mathcal{U}(1)$ 对称性中得到一点提示，它表明从 Ψ 的运动方程得到的解没有直接的物理意义⁹。

另一方面，注意到运动方程所有的解都是 \vec{x}, t 的函数¹⁰，即 $\Psi = \Psi(\vec{x}, t)$ 。

标准诠释是，波函数 $\Psi(\vec{x}, t)$ 的模方 $|\Psi(\vec{x}, t)|^2$ 表示粒子位置的概率密度。注意 $\mathcal{U}(1)$ 变换对于 $|\Psi|^2 = \Psi^\dagger \Psi \rightarrow (\Psi')^\dagger (\Psi') = \Psi^\dagger e^{-i\alpha} e^{i\alpha} \Psi = \Psi^\dagger \Psi$ 这样的量是没有影响的。换句话说（以一维情形为例）， $\Psi(x, t)$ 是在 $[x, x + dx]$ 区间内测量到粒子的概率密度³。因此，如果对整个空间积分的话，必定有

$$\int dx \Psi^*(x, t) \Psi(x, t) \stackrel{!}{=} 1 \quad (8.10)$$

这称为波函数的归一化，因为在全空间中找到一个粒子的概率必定为 100% = 1。

如果要预言其他物理量，则必须把波函数用相应的本征函数基展开。比如说用能量本征基展开： $\Psi = c_1 \Phi_{E_1} + c_2 \Phi_{E_2} + \dots$ 。标准的量子力学诠释是，测量 Ψ 状态的体系的能量，所得结果只能为能量本征值，测量结果为本征值 E_1 的概率为 Ψ 与能量本征函数 Φ_{E_1} 的“交叠”（的绝对值平方）：

$$P(E_1) = \left| \int dx \Phi_{E_1}^*(x, t) \Psi(x, t) \right|^2$$

在上面的例子中，它就是

$$\begin{aligned} P(E_1) &= \left| \int dx \Phi_{E_1}^*(x, t) \Psi(x, t) \right|^2 = \left| \int dx \Phi_{E_1}^*(x, t) (c_1 \Phi_{E_1} + c_2 \Phi_{E_2} + \dots) \right|^2 \\ &= \left| \underbrace{c_1 \int dx \Phi_{E_1}^*(x, t) \Phi_{E_1}}_{=1 \text{ 如之前所讲}} + \underbrace{\int dx \Phi_{E_1}^*(x, t) (c_2 \Phi_{E_2} + \dots)}_{=0 \text{ 因为本征态是正交的}} \right|^2 \\ &= |c_1|^2 \end{aligned} \quad (8.11)$$

类似的，可以将 $\Psi(\vec{x}, t)$ 用动量本征函数展开：

$$\Psi(\vec{x}, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d^3p \Psi(\vec{p}, t) e^{-i\vec{p} \cdot \vec{x}},$$

³即在区间 $[x, x + dx]$ 测量到粒子的概率为 $|\Psi(x, t)|^2 dx$ 。

⁹ 如果假定 Ψ 直接描述粒子，那么它的 $\mathcal{U}(1)$ 变换 $\Psi' = e^{i\alpha} \Psi$ （也是运动方程的解）描述的是什么呢？

¹⁰ 下一节会对这点详细说明。

则 $\Psi(\vec{p}, t)$ 是系统动量处于 $[p, p + dp]$ 区间内的概率幅。

这种解释可以用来对系统做几率性预测，比如用下一节介绍的统计期望值的办法。之后我们会导出非相对论性量子力学的运动方程，并看两个例子。

8.3.1 Expectation Value 期望值

统计学中，期望值是类比于加权平均定义的。比如，如果掷骰子十次得到的结果为 2, 4, 1, 3, 3, 6, 3, 1, 4, 5，那么它的平均值就是

$$\langle x \rangle = (2 + 4 + 1 + 3 + 3 + 6 + 3 + 1 + 4 + 5) \cdot \frac{1}{10} = 3.2$$

而另一种计算这个的方法是先将相同结果放在一起，通过权重给出平均值

$$\langle x \rangle = \frac{2}{10} \cdot 1 + \frac{1}{10} \cdot 2 + \frac{3}{10} \cdot 3 + \frac{2}{10} \cdot 4 + \frac{1}{10} \cdot 5 + \frac{1}{10} \cdot 6 = 3.2$$

一般的，我们可以写

$$\langle x \rangle = \sum_i \rho_i x_i \quad (8.12)$$

其中 ρ_i 表示几率。对一个连续分布的情况同样有

$$\langle x \rangle = \int dx \rho(x) x \quad (8.13)$$

类比上面的形式，在量子力学中物理量 \hat{O} 期望值的定义为

$$\langle \hat{O} \rangle = \int d^3x \Psi^* \hat{O} \Psi \quad (8.14)$$

一般的，要把 Ψ 按照 \hat{O} 的本征函数展开，比如说动量本征函数。然后，把算符 \hat{O} 作用在这些态上，得到对应的本征值，这样就得到了加权和。

举个例子，处于 Ψ 态的粒子的坐标期望值为

$$\langle \hat{x} \rangle = \int d^3x \Psi^* \hat{x} \Psi = \int d^3x \Psi^* x \Psi = \int d^3x x \underbrace{\Psi^* \Psi}_{\text{在位置 } x \text{ 的几率密度}} \quad (8.15)$$

为了计算方便，下面取 Klein-Gordon 方程的非相对论极限。

8.4 Schrödinger Equation Schrödinger 方程

Klein-Gordon 方程有平面波解

$$\Phi = e^{\pm i p_\mu x^\mu} \equiv e^{\pm i p \cdot x}$$

其中, $p_\mu = (E, \vec{p})^T$, 是粒子四动量 (满足守恒律):

$$\begin{aligned} 0 &= (\partial_\mu \partial^\mu + m^2) \Phi \\ &= (\partial_\mu \partial^\mu + m^2) e^{\pm i p_\mu x^\mu} \\ &= (i^2 p_\mu p^\mu + m^2) e^{\pm i p_\mu x^\mu} = 0 \\ &= (-m^2 + m^2) e^{\pm i p_\mu x^\mu} = 0 \quad \checkmark \end{aligned} \tag{8.16}$$

我们可以把解写成稍不同的形式

$$\Phi = e^{\pm i p_\mu x^\mu} = \Phi = e^{i(-Et + \vec{x} \cdot \vec{p})}$$

Φ 的时间部分是 e^{-iEt} , 即 $\Phi \propto e^{-iEt}$ 。从(8.2)式可知

$$E = \sqrt{\vec{p}^2 + m^2} = \sqrt{m^2 \left(\frac{\vec{p}^2}{m^2} + 1 \right)} = m \sqrt{\frac{\vec{p}^2}{m^2} + 1}.$$

在非相对论极限 $|\vec{p}| \ll m$ 下, 即研究对象的移动速度比光速要慢很多, 从而其动量比质量也要小很多的情形下, 可以对能量进行 Taylor 展开:

$$\begin{aligned} E &= m \left(1 + \frac{1}{2} \frac{\vec{p}^2}{m^2} + \dots \right) \\ &\rightarrow E \approx \underbrace{m}_{\text{静质量}} + \underbrace{\frac{\vec{p}^2}{2m}}_{\text{动能}}. \end{aligned}$$

于是

$$\Phi = e^{i(-Et + \vec{x} \cdot \vec{p})} \approx e^{-imt} \underbrace{e^{i\vec{p} \cdot \vec{x} - i(p^2/2m)t}}_{\equiv \phi(\vec{x}, t)} = e^{-imt} \phi(\vec{x}, t) \tag{8.17}$$

由 $|\vec{p}| \ll m$ 可知静质能远大于粒子动能, 因此 $\phi(\vec{x}, t)$ 的振动

频率远低于 e^{-imt} . 把 $e^{-imt}\phi(\vec{x}, t)$ 作为试探解⁴带入 Klein-Gordon 方程可得⁵

$$(\partial_\mu \partial^\mu + m^2)e^{-imt}\phi(\vec{x}, t) = (\partial_0 \partial^0 - \partial_i \partial^i + m^2)e^{-imt}\phi(\vec{x}, t) = 0.$$

根据莱布尼兹律, $\partial_t e^{-imt}(\dots) = e^{-imt}(-im + \partial_t)(\dots)$, 即简单的把微分算符作用出来, 我们得到

$$e^{-imt}((-im + \partial_t)^2 + \partial_i \partial^i + m^2)\phi(\vec{x}, t) = 0,$$

两边可以除掉 e^{-imt} , 因为它肯定不是零。因此,

$$\rightarrow ((-im + \partial_t)^2 + \nabla^2 + m^2)\phi(\vec{x}, t) = 0. \quad (8.18)$$

$$7 \rightarrow (-m^2 - 2im\partial_t + (\partial_t)^2 + \nabla^2 + m^2)\phi(\vec{x}, t) = 0.$$

把第三项

$$\begin{aligned} (\partial_t)^2 \phi(\vec{x}, t) &= (\partial_t)^2 \exp[i\vec{p} \cdot \vec{x} - i(\vec{p}^2/2m)t] \\ &= \left(\frac{\vec{p}^2}{2m}\right)^2 \exp[i\vec{p} \cdot \vec{x} - i(\vec{p}^2/2m)t] \propto \frac{p^4}{m^2}. \end{aligned} \quad (8.19)$$

与第二项对比

$$\begin{aligned} im\partial_t \phi(\vec{x}, t) &= im\partial_t \exp[i\vec{p} \cdot \vec{x} - i(\vec{p}^2/2m)t] \\ &= m \left(\frac{\vec{p}^2}{2m}\right) \exp[i\vec{p} \cdot \vec{x} - i(\vec{p}^2/2m)t] \propto p^2. \end{aligned} \quad (8.20)$$

这表明, 在 $|\vec{p}| \ll m$ 的极限下, 第三项可忽略, 因此⁶

$$\begin{aligned} (-2im\partial_t + \nabla^2)\phi(\vec{x}, t) &= 0 \\ \underbrace{\rightarrow}_{\text{除以}(-2m)} \left(i\partial_t - \frac{1}{2m}\nabla^2 \right) \phi(\vec{x}, t) &= 0 \end{aligned}$$

⁴译者注: ansatz 是个很奇怪的词, 我不知道怎么翻译。。。比如 Bethe-Ansatz, 叫 Bethe 近似嘛也不对, 叫啥都不对。

⁵原文下式有误, 译文已按勘误表修改。

⁶原文下式有误, 译文已按勘误表修改

$$\rightarrow \left(i\partial_t + \frac{\nabla^2}{2m} \right) \phi(\vec{x}, t) = 0, \quad (8.21)$$

这就是著名的 **Schrödinger 方程**。利用本章开头物理量与算符的对应关系，Schrödinger 方程写为⁷

$$\begin{aligned} &\rightarrow \left(E - \underbrace{\frac{\vec{p}^2}{2m}}_{=\text{动能}} \right) \phi(\vec{x}, t) = 0 \\ &\rightarrow E = \frac{\vec{p}^2}{2m}. \end{aligned} \quad (8.22)$$

此即为非相对论性能量—动量关系。

从该式出发很容易考虑粒子位于外势场 V 中的情况，因为外势场中粒子的能量就是动能加势能：

$$\rightarrow E = \frac{\vec{p}^2}{2m} + V.$$

典型的例子有谐振子势 $V = -kx^2$ 。Schrödinger 方程通常用所有能量算符之和（例如动能⁸— $\frac{\nabla^2}{2m}$ 加势能 \hat{V} ）——哈密顿算符 \hat{H} 表示，即：

$$\begin{aligned} i\partial_t \phi(\vec{x}, t) &= -\underbrace{\frac{\nabla^2}{2m}}_{\equiv \hat{H}} \phi(\vec{x}, t) \\ &\rightarrow i\partial_t \phi(\vec{x}, t) = \hat{H} \phi(\vec{x}, t). \end{aligned} \quad (8.23)$$

Dirac 方程的非相对论极限的计算过程类似，所得结果称为 Pauli 方程。

8.4.1 Schrödinger Equation with External Field 有外场的 Schrödinger 方程

此外，我们可以对有相互作用的 Klein-Gordon 方程(7.43)式，即描述有质量的 0 自旋场与无质量的自旋 1 场（光子场）相互作用的方程，做同样的非相对论性近似。所得结果为

$$\left(i\partial_t + \frac{1}{2m} \left(\nabla - ia\vec{A} \right)^2 + q\Phi \right) \phi(\vec{x}, t) = 0 \quad (8.24)$$

⁷原文下式有误，译文已按勘误表修改

⁸译注：原文缺负号，已更正。

8.5 From Wave Equations to Particle Motion 从波动方程到粒子的运动

本节讨论量子力学标准诠释的两个典型例子。

8.5.1 Example: Free Particle 例：自由粒子

自由粒子（无外部势场）Schrödinger 方程(8.21)式的一个特解（称为平面波解）为：

$$\Psi = e^{-i(Et - \vec{p} \cdot \vec{x})}, \quad (8.25)$$

验证：

$$\begin{aligned} i\partial_t e^{-i(Et - \vec{p} \cdot \vec{x})} &= -\frac{\nabla^2}{2m} e^{-i(Et - \vec{p} \cdot \vec{x})} \\ \rightarrow E e^{-i(Et - \vec{p} \cdot \vec{x})} &= \frac{\vec{p}^2}{2m} e^{-i(Et - \vec{p} \cdot \vec{x})}. \end{aligned} \quad (8.26)$$

其中 E 是自由粒子总能量的值， $E = \frac{\vec{p}^2}{2m}$ ((8.22)式)。各特解的线性组合为更一般的解：

$$\Psi = A e^{-i(Et - \vec{p} \cdot \vec{x})} + B e^{-i(E't + \vec{p}' \cdot \vec{x})} + \dots$$

波函数 Ψ 被诠释为概率幅，因此描述真实粒子的波函数必定是归一化的，因为在粒子在全空间的概率为 $100\% = 1$ 。上面的波函数不能归一化，它散布在全空间。因此采用平面波的合适的线性组合（称为波包 (wave-packet)）来描述一个真实的自由粒子：

$$\Psi_{WP}(\vec{x}, t) = \int d^3p A(\vec{p}) e^{i(\vec{p} \cdot \vec{x} - Et)}, \quad (8.27)$$

其中系数 $A(\vec{p})$ 为复数，由归一化条件求出。典型的波包是高斯波包，它的 $A(\vec{p})$ 服从高斯分布：

$$\Psi_{GWP} = \int d^3p A(\vec{p}) e^{i(\vec{p} \cdot \vec{x} - Et)} = \int d^3p \psi_0 e^{i(\vec{p} - \vec{p}_0)^2 / 4\sigma^2} e^{i(\vec{p} \cdot \vec{x} - Et)}.$$

其中 \vec{p} 为常数向量。

典型的高斯波包如图8.1。在实际计算中经常采用特殊技巧避免求解复杂波包，而是处理更简单的波函数。

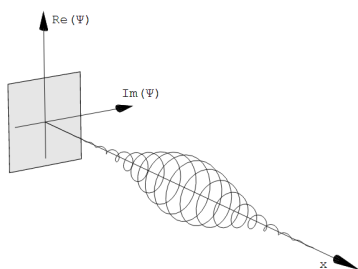


图 8.1: 自由高斯波包图示。Figure by Inductiveload (Wikimedia Commons) released under a public domain license. URL: https://upload.wikimedia.org/wikipedia/commons/4/49/Travelling_Particle_Wavepacket.svg

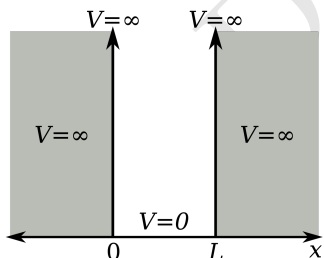


图 8.2: 一维无限深势阱（硬盒）。Figure by Benjamin D. Esham (Wikimedia Commons) released under a public domain licence. URL: https://upload.wikimedia.org/wikipedia/commons/2/27/Infinitesimal_potential_well.svg, Accessed: 4.5.2014

8.5.2 Example: Particle in a Box 例：盒子中的粒子

下面讨论量子力学的标准例子⁹：一个粒子被限制在一维的、势垒无限高的硬盒内，即，盒子内的势能为零，盒外的势能为无穷大，如图8.2。

无限深势阱数学上用一段函数表示：

$$V = \begin{cases} 0, & 0 < x < L, \\ \infty, & \text{其它}. \end{cases} \quad (8.28)$$

因此需要分段求解 Schrödinger 方程

$$i\partial_t \Psi(\vec{x}, t) = -\frac{1}{2m} \partial_x^2 \Psi(x, t) + V(x) \Psi(x, t).$$

- 势阱内 $V = 0, 0 < x < L$ ，方程的解即为自由粒子波函数。
- 势阱外 $V = \infty$ ，具有物理意义的解仅有 $\Psi(x, t) = 0$ ，即由于势垒无穷大，粒子不可能在势阱外出现。

将自由粒子波函数的一般形式写为¹¹：

$$\begin{aligned} \Psi(x, t) &= Ae^{-i(Et - \vec{p} \cdot \vec{x})} + Be^{-i(Et + \vec{p} \cdot \vec{x})} \\ &= (C \sin(\vec{p} \cdot \vec{x}) + D \cos(\vec{p} \cdot \vec{x})) e^{-iEt}, \end{aligned}$$

利用(8.22)式表示的非相对论能量-动量关系，上式写为：

$$\begin{aligned} E &= \frac{\vec{p}^2}{2m} \rightarrow |\vec{p}| = \sqrt{2mE} \\ \Psi(x, t) &= (C \sin(\sqrt{2mE}x) + D \cos(\sqrt{2mE}x)) e^{-iEt} \end{aligned} \quad (8.29)$$

利用波函数的连续性¹²可得边界条件为 $\Psi(0) = \Psi(L) \stackrel{!}{=} 0$ 。带入上式，由于 $\cos 0 = 1$ 故而 $D \stackrel{!}{=} 0$ 。此外还有

$$\sqrt{2mE} \stackrel{!}{=} \frac{n\pi}{L}, \quad n = 1, 2, 3, \dots \quad (8.30)$$

于是满足边界条件的波函数为¹³：

$$\Phi_n(x, t) = C \sin\left(\frac{n\pi}{L}x\right) e^{-iE_n t}. \quad (8.31)$$

¹¹ 下式利用了 $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$ 以及 $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ ，这两个式子可以从 $\cos x, \sin x, e^{ix}$ 的级数展开式导出，详见附录B.4.1。

¹² 如果波函数在某处不连续，则粒子动量 $\hat{p}_x \Psi = -i\partial_x \Psi$ 为无穷大（跃变点处的导数无穷大），不符合基本法。

¹³ 注意下面给波函数添加了编号 n ，因为由上式可见每个 n 都对应一个解。

⁹ 几乎任何一本量子力学教材都会有吧……

验证:

$$\rightarrow \Phi_n(L, t) = C \sin\left(\frac{n\pi}{L}L\right) e^{-iE_n t} = C \sin(n\pi) e^{-iE_n t} = 0 \quad \checkmark$$

$$\rightarrow \Phi_n(0, t) = C \sin\left(\frac{n\pi}{L}0\right) e^{-iE_n t} = C \sin(0) e^{-iE_n t} = 0 \quad \checkmark$$

不难求出归一化常数 $C = \sqrt{\frac{2}{L}}$, 注意在盒内 ($0 < x < L$) 找到粒子的概率为 $100\% = 1$, 而盒外的概率为零:

$$\begin{aligned} P &= \int_0^L dx \Phi_n^*(x, t) \Phi_n(x, t) \stackrel{!}{=} 1 \\ P &= \int_0^L dx C^2 \sin\left(\frac{n\pi}{L}x\right) e^{+iE_n t} \sin\left(\frac{n\pi}{L}x\right) e^{-iE_n t} \\ &= C^2 \int_0^L dx \sin^2\left(\frac{n\pi}{L}x\right) = C^2 \left[\frac{x}{2} - \frac{\sin\left(\frac{2n\pi}{L}x\right)}{4\frac{n\pi}{L}} \right]_0^L \\ &= C^2 \left(\frac{L}{2} - \frac{\sin\left(\frac{2n\pi}{L}L\right)}{4\frac{n\pi}{L}} \right) = C^2 \frac{L}{2} \stackrel{!}{=} 1 \\ \rightarrow C^2 &\stackrel{!}{=} \frac{2}{L} \quad \checkmark \end{aligned}$$

解(8.30)式得到盒中粒子的能级为

$$E_n \stackrel{!}{=} \frac{n^2 \pi^2}{L^2 2m}. \quad (8.32)$$

粒子可能具有的能量是量子化的, 即能量只能是某个常量 (这里是 $\frac{\pi^2}{L^2 2m}$) 的整数倍, 这就是量子力学名字的来历。

每个 n 都对应一个解, 解的线性组合

$$\Phi(x, t) = A\Phi_1(x, t) + B\Phi_2(x, t) + \dots$$

仍然是 Schrödinger 方程的解。各线性系数 A, B, \dots 要满足归一化条件¹⁴。

下面考虑能量的问题¹⁵, 例如, 测量一个 (归一化的) 状态为

$$\Psi(x, t) = \sqrt{\frac{3}{5}}\Phi_2(x, t) + \sqrt{\frac{2}{5}}\Phi_3(x, t).$$

的粒子, 测得能量 $E = E_2 = \frac{2^2 \pi^2}{L^2 2m}$ 的概率是多少?

量子力学经典诠释给出的答案是: 概率 $P(E = E_2)$ 等于粒子

¹⁴ 大于 $1 = 100\%$ 的概率没有意义。

¹⁵ 前面讨论的是粒子位置的概率分布。

状态 Ψ 与 E_2 对应的本征态 Φ_2 “重叠”绝对值的平方：

$$P(E = \frac{2^2\pi^2}{L^2 2m}) = \left| \int dx \Phi_2^*(x, t) \Psi(x, t) \right|^2.$$

Φ_2 与 Ψ 的重叠可以视为它们的标量积¹⁶：

$$(\Phi_2, \Psi) = \int dx \Phi_2^* \Psi = \underbrace{c}_{\text{复数}}.$$

上述积分可以利用本征状态的正交归一性大大简化计算

$$\int dx \Phi_n^*(x, t) \Phi_{n'}(x, t) = \delta_{nn'}$$

例如¹⁷：

$$\begin{aligned} \int_0^L dx \Phi_2^* \Phi_3(x, t) &= \int_0^L dx C \sin\left(\frac{2\pi}{L}x\right) e^{+iEt} C \sin\left(\frac{3\pi}{L}x\right) e^{-iEt} \\ &= C^2 \int_0^L dx \sin\left(\frac{2\pi}{L}x\right) \sin\left(\frac{3\pi}{L}x\right) = 0. \end{aligned}$$

因此测量得到能量 $E = \frac{2^2\pi^2}{L^2 2m}$ 的概率为

$$\begin{aligned} P(E = \frac{2^2\pi^2}{L^2 2m}) &= \left| \int dx \Phi_2^*(x, t) \Psi(x, t) \right|^2 \\ &= \left| \int dx \Phi_2^*(x, t) \left(\sqrt{\frac{3}{5}} \Phi_2(x, t) + \sqrt{\frac{2}{5}} \Phi_3(x, t) \right) \right|^2 \\ &= \left| \int dx \left(\underbrace{\sqrt{\frac{3}{5}} \Phi_2^*(x, t) \Phi_2(x, t)}_{\text{积分后等于1}} + \underbrace{\sqrt{\frac{2}{5}} \Phi_2^*(x, t) \Phi_3(x, t)}_{\text{积分后等于0}} \right) \right|^2 \\ &= \left(\sqrt{\frac{3}{5}} \right)^2. \end{aligned} \quad (8.33)$$

(8.31)式的波函数解称为能量算符 $i\partial_t$ 的本征态，或者哈密顿算符¹⁸ $\hat{H} \equiv -\frac{\partial_x^2}{2m}$ 的本征态，因为

$$\hat{H}\Phi_n = E_n\Phi_n. \quad (8.34)$$

能量算符作用于能量本征态得到的结果为该本征态的常数倍，这个常数即为该状态的能量。注意能量算符（或哈密顿算符）作用于一般状态造成的变化不止变成常数倍这么简单。例如，考虑如下能

¹⁶ 实际上， Ψ, Φ_n 是 Hilbert 空间的向量， (Ψ, Φ_n) 正是它们的标量积。

¹⁷ 不难用分部积分或者 Wolframalpha 检验计算结果。

¹⁸ 这可以从 Schrödinger 方程直接导出：

$$i\partial_t \Phi = -\frac{\partial_x^2}{2m} \Phi \equiv \hat{H}\Phi.$$

量本征态的线性组合：

$$\Psi = \sqrt{\frac{3}{5}}\Phi_2 + \sqrt{\frac{2}{5}}\Phi_3.$$

哈密顿算符作用其上得到：

$$\hat{H}\Psi = \hat{H}\left(\sqrt{\frac{3}{5}}\Phi_2 + \sqrt{\frac{2}{5}}\Phi_3\right) \underset{(8.34)\text{式}}{=} \sqrt{\frac{3}{5}}E_2\Phi_2 + \sqrt{\frac{2}{5}}E_3\Phi_3.$$

作用后的结果不能再写成 Ψ 的常数倍了，因为 $E_2 \neq E_3$ 。因此 Ψ 不是能量算符的本征态。然而 Ψ 可以表示为各 Φ_n 的线性组合，因为能量本征态 $\{\Phi_n\}$ 构成一组完备基。

下一节介绍由 Dirac 发明的十分便捷的 Dirac 符号，它有助于我们理解量子力学的框架结构。

¹⁹ 它的含义稍后会进一步说明。

8.5.3 Dirac Notation Dirac 符号

Dirac 符号中，系统状态由右矢 (**ket**)¹⁹表示，记作

$$|\Psi\rangle, \quad (8.35)$$

例如，一个硬盒中的状态为 Φ_n 的粒子的能量本征方程用 Dirac 符号写为

$$\hat{H}|\Phi_n\rangle = E_n|\Phi_n\rangle.$$

每个右矢可以定义相应的左矢 (**bra**)，记作 $\langle\Psi|$ ，定义为

$$\langle\Psi|^\dagger = |\Psi\rangle, \quad (8.36)$$

其中符号 \dagger 表示厄米共轭，即转置 + 复共轭。左矢是作用于右矢的数学对象，右矢之间的内积定义为

$$(|\Phi\rangle, |\Psi\rangle) \equiv \langle\Phi||\Psi\rangle.$$

左矢作用与右矢产生的结果为一个复数，记作

$$\langle\Phi||\Psi\rangle = c. \quad (8.37)$$

复数 c 即为 $|\Psi\rangle$ 状态的粒子观测位于 $|\Phi\rangle$ 状态的概率幅，即相应的概率为 $|\langle\Phi|\Psi\rangle|^2$ 。例如测量一个 $|\Psi\rangle$ 状态的粒子位于 $[x, x+dx]$ 的概率为

$$\langle x|\Psi\rangle \equiv \Psi(x).$$

$\Psi(x)$ 即为前面使用的波函数。再比如，观测粒子的动量位于 $[p, p+dp]$ 的概率幅是多少？答案用 Dirac 符号表示为

$$\langle p|\Psi\rangle \equiv \Psi(p).$$

我们一直采用的是概率性诠释，因而物理状态必须是归一化的²⁰，例如

$$\int dx |\Psi(x, t)|^2 = \int dx \Psi^\dagger(x, t) \Psi(x, t) \stackrel{!}{=} 1 \quad (8.38)$$

$$\int dp |\Phi(p, t)|^2 = \int dp \Phi^\dagger(p, t) \Phi(p, t) \stackrel{!}{=} 1. \quad (8.39)$$

上两式用 Dirac 符号表示为

$$\int dx |\langle x|\Psi\rangle|^2 = \int dx (\langle x|\Psi\rangle)^\dagger \langle x|\Psi\rangle = \int dx \langle \Psi|x\rangle \langle x|\Psi\rangle \stackrel{!}{=} 1 \quad (8.40)$$

$$\int dp |\langle p|\Phi\rangle|^2 = \int dp (\langle p|\Phi\rangle)^\dagger \langle p|\Phi\rangle = \int dp \langle \Phi|p\rangle \langle p|\Phi\rangle \stackrel{!}{=} 1. \quad (8.41)$$

上面出现了一种新算符： $|p\rangle\langle p|$ 与 $|x\rangle\langle x|$ ，这种形式的算符称为**投影算符 (projection operators)**²¹。它们是算符，因为它们作用于右矢，得到新的右矢。例如

$$|x\rangle \underbrace{\langle x|\Psi\rangle}_{\text{等于某个复数, 记作}\Psi(x)} = |x\rangle \Psi(x),$$

等于某个复数，记作 $\Psi(x)$

等号右侧是复数乘以右矢，因而作用结果是一个右矢，而算符就是作用于右矢得到右矢的东西。由(8.40)式可以引入一个新算符：

$$\int dx \langle \Psi|x\rangle \langle x|\Psi\rangle = \langle \Psi | \underbrace{\left(\int dx |x\rangle \langle x| \right)}_{\equiv \hat{I}} | \Psi \rangle = \langle \Psi | \hat{I} | \Psi \rangle \stackrel{!}{=} 1. \quad (8.42)$$

²⁰ 粒子在全空间中出现的概率为 $100\% = 1$ ，或等价地说，粒子的动量取任意值的概率为 $100\% = 1$ 。换句话说，所有状态的概率之和等于 1。

²¹ 就像前面介绍过的左手旋量、右手旋量的投影算符那样，这些投影算符也满足条件 $P^2 = P$ 。

由此可得

$$\hat{I}|\Psi\rangle \stackrel{!}{=} |\Psi\rangle. \quad (8.43)$$

任何表示物理状态的 $|\Psi\rangle$ 都满足 $\langle\Psi|\Psi\rangle = 1$, 因为观测一个处于 $|\Psi\rangle$ 状态的系统, 所得结果为 $|\Psi\rangle$ 的概率为 $100\% = 1$. 例如, 在 x_0 处放置一个粒子, 即在 x_0 处观测到该粒子的概率为 1, 则 $\langle x_0|x_0\rangle = 1$. \hat{I} 由此称为单位算符, 它的地位就像数字 1 在数字乘法中那样。关系式

$$\int dx |x\rangle\langle x| = \hat{I}, \quad (8.44)$$

或者当基矢为离散的情况下

$$\sum_i |i\rangle\langle i| = \hat{I} \quad (8.45)$$

称为完备性关系 (completeness relations)。一般的右矢 $|a\rangle$ 在基 $\{|i\rangle\}$ 下的分量为

$$|i\rangle^\dagger|a\rangle \equiv \langle i|a\rangle \equiv a_i,$$

分量 a_i 是复数。利用完备性关系, 即 $\sum_i |i\rangle\langle i| = \hat{I}$ 可得

$$|a\rangle = \sum_i |i\rangle\langle i|a\rangle = \sum_i |i\rangle a_i.$$

²² 这与三维矢量用基向量展开 (例如 $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3$) 相似, 详见附录A.1。

这可以看做是 $|a\rangle$ 在基 $\{|i\rangle\}$ 下的展开式²², a_i 可以视为展开系数。对于连续基, 展开式写为

$$|\Psi\rangle = \int dx |x\rangle \underbrace{\langle x|\Psi\rangle}_{\equiv \Psi(x), \text{复数}} = \int dx |x\rangle \Psi(x).$$

8.3.1节介绍的期望值用 Dirac 符号表示为

$$\langle\hat{O}\rangle = \langle\Psi|\hat{O}|\Psi\rangle.$$

下面重新回顾一下硬盒中的粒子, 用 Dirac 符号表述。

8.5.4 Example: Particle in a Box, Again 例：盒子中的粒子，续

考虑前面求解过的问题：观测位于 $|\Psi\rangle$ 状态粒子的能量，所得结果为 $E = E_2 = \frac{n^2\pi^2}{2mL^2}$ 的概率是多少？这个问题用 Dirac 符号可以简明地表示，概率为

$$\begin{aligned} P(E_2) &= \left\langle E = \frac{n^2\pi^2}{2mL^2} \middle| \Psi \right\rangle = \left\langle E = \frac{n^2\pi^2}{2mL^2} \middle| \underbrace{\left(\int dx |x\rangle \langle x| \right)}_{=\hat{I}} \middle| \Psi \right\rangle \\ &= \int dx \left\langle E = \frac{n^2\pi^2}{2mL^2} \middle| x \right\rangle \langle x | \Psi \rangle = \int dx \Phi_2^*(x) \Psi(x) \end{aligned}$$

答案与 8.5.2 节用波动力学标准方法的结果相同。

8.5.5 Spin 自旋

本节讨论 5.1.1 节所导出的新的角动量——自旋。关于自旋目前我们有两条线索。首先，我们用自旋标记了 Lorentz 群的表示。例如描述自旋 $\frac{1}{2}$ 的基本粒子需要用按照 Lorentz 群 $\frac{1}{2}$ 表示进行变换的对象。其次，自旋是旋转对称性对应守恒量的一部分²³。5.1.1 节导出的自旋算符作用于态矢给出粒子的自旋。对于标量表示（所描述的粒子），自旋算符 $\hat{S} = 0$ ，即标量粒子任意状态的自旋为 0。

再考虑 $(\frac{1}{2}, 0)$ 表示，用二维矩阵表示的旋转生成元为（推导见 3.7.5 节）：

$$\hat{S}_i = \frac{\sigma_i}{2} \quad (8.46)$$

其中 σ_i 为 Pauli 矩阵。描述粒子自旋的状态记作 Ψ ，下面将自旋算符作用与 Ψ 。例如算符 \hat{S}_3 作用的结果是粒子自旋在第 3 方向，通常称为 z 方向的分量。同样 \hat{S}_2 表示自旋 y 分量， \hat{S}_1 表示 x 分量。

算符 \hat{S}_3 的显式形式为

$$\hat{S}_3 = \frac{\sigma_3}{2} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \quad (8.47)$$

相应的本征态为

$$v_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad v_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (8.48)$$

²³ 见 4.5.4 节。

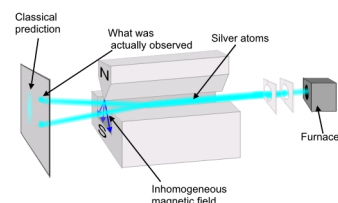


图 8.3: Stern-Gerlach 实验图示。实验在最初使用银原子，它的自旋仅由原子轨道最外层电子的自旋决定，因此实验结果与电子相同。一束粒子流通过不均匀磁场，对于经典类型的轨道角动量，粒子在磁场偏离形成连续分布。然而实验只观测到两种偏离，即粒子流经磁场后只分为两束，一束为自旋 $\frac{1}{2}$ ，另一束为 $-\frac{1}{2}$ 。图片作者为 Theresa Knott (Wikimedia Commons) distributed under a CC BY-SA 3.0 license: <http://creativecommons.org/licenses/by-sa/3.0/deed.en> 图片地址: https://upload.wikimedia.org/wikipedia/commons/e/ee/Stern-Gerlach_experiment_svg.svg

²⁴ 回顾：旋量是坐标变换下按照 Lorentz 群的 $(\frac{1}{2}, 0)$ 表示或 $(0, \frac{1}{2})$ 表示或 $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ 表示变换的量。

相应的本征值分别为 $\frac{1}{2}, -\frac{1}{2}$ 。这意味着旋量²⁴描述的粒子的自旋为 $\frac{1}{2}$ ，方向与坐标轴平行或反向。自旋 $\frac{1}{2}$ 表示的名字正是来源于此。量子理论诠释为自旋只能取 $\frac{1}{2}$ 或 $-\frac{1}{2}$ 。[4.5.4](#)节讲过，自旋与轨道角动量有些相似，它们都源自旋转不变性。从上面的讨论可见自旋的观测结果有着奇特的性质，该性质最著名的实验验证是 Stern-Gerlach 实验（见图[8.3](#)）。

自旋的这一性质在任意方向都成立。测量自旋的 x, y 或 z 分量的结果都只可能有 $\frac{1}{2}$ 或 $-\frac{1}{2}$ 。

下面看一个用量子理论描述自旋测量的例子。前面（[\(5.4\)](#)式）说过，自旋算符（以 z 分量为例）的显式形式为：

$$\hat{S}_z = \frac{1}{2}\sigma_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \quad (8.49)$$

本征态为 $|\frac{1}{2}\rangle_z \hat{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 以及 $|\frac{-1}{2}\rangle_z \hat{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ，其中下标 z 表示与 \hat{S}_z 有关。一般的旋量并非本征态自旋，而是它们的线性叠加，即 $|X\rangle = a|\frac{1}{2}\rangle_z + b|\frac{-1}{2}\rangle_z$ 。系数决定特定状态。如果观测一束粒子流的 z 分量，并“过滤”掉自旋 $-\frac{1}{2}$ 的部分¹⁰，则筛选后的粒子自旋状态的系数 $b=0, a=1$ ，在筛选之前，状态为 $a=b=\frac{1}{\sqrt{2}}$ ，这意味着两种自旋的概率²⁵相等均为 $\frac{1}{2}$ 。先后测量粒子自旋在不同方向分量（例如，先测量 S_z 再测量 S_x ）的结果十分有趣。即便把所有自旋 $-\frac{1}{2}$ 粒子都剔除出去，其余的仍有 $S_x = -\frac{1}{2}$ 的粒子。

自旋算符 \hat{S}_z 作用于态矢意味着测量自旋的 z 分量。对于一般状态 $|X\rangle$ ，测量结果为 $\frac{1}{2}$ 或 $-\frac{1}{2}$ 的概率由系数 a, b 决定。如果测量结果为 $-\frac{1}{2}$ 的概率已知，则相应的概率幅为

$${}_z\langle -\frac{1}{2} | X \rangle = a \underbrace{{}_z\langle -\frac{1}{2} | \frac{1}{2} \rangle_z}_{=0} + b \underbrace{{}_z\langle -\frac{1}{2} | -\frac{1}{2} \rangle_z}_{=1} = b. \quad (8.50)$$

因此测量自旋 z 分量所得结果为 $-\frac{1}{2}$ 的概率为 $P_{z=-\frac{1}{2}} = |b|^2$ 。测量其他方向的分量，例如 x 轴，即用自旋算符 \hat{S}_x 作用于态矢， S_x 的矩阵形式为（[\(5.4\)](#)式）：

$$S_x = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \quad (8.51)$$

¹⁰这可以用前面提到的 Stern-Gerlach 实验装置完成。

²⁵ 系数 a, b 与概率幅直接相关，下面就会看到。

它的本征值仍为 $\pm\frac{1}{2}$, 本征矢分别为 $|\frac{1}{2}\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 以及 $|\frac{1}{2}\rangle_x \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$,

要计算 $S_x = -\frac{1}{2}$ 的概率, 需要将 $|\frac{1}{2}\rangle_z$ 与 $|\frac{1}{2}\rangle_x$ 用 $|\frac{1}{2}\rangle_x, |-\frac{1}{2}\rangle_x$ 表示:

$$\underbrace{|\frac{1}{2}\rangle_z}_{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \frac{1}{\sqrt{2}} \left(\underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} + \underbrace{|-\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}} \right) \quad (8.52)$$

$$\underbrace{|-\frac{1}{2}\rangle_z}_{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} = \frac{1}{\sqrt{2}} \left(\underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} - \underbrace{|-\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}} \right) \quad (8.53)$$

因此

$$|X\rangle = a|\frac{1}{2}\rangle_z + b|-\frac{1}{2}\rangle_z = a\frac{1}{\sqrt{2}} \left(|\frac{1}{2}\rangle_x + |-\frac{1}{2}\rangle_x \right) + b\frac{1}{\sqrt{2}} \left(|\frac{1}{2}\rangle_x - |-\frac{1}{2}\rangle_x \right). \quad (8.54)$$

测量自旋 x 分量所得结果为 $-\frac{1}{2}$ 的概率幅为

$$\begin{aligned} {}_x\langle -\frac{1}{2} || X \rangle &= {}_x\langle -\frac{1}{2} | \left(a\frac{1}{\sqrt{2}} \left(|\frac{1}{2}\rangle_x + |-\frac{1}{2}\rangle_x \right) + b\frac{1}{\sqrt{2}} \left(|\frac{1}{2}\rangle_x - |-\frac{1}{2}\rangle_x \right) \right) \\ &= \frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}} \end{aligned} \quad (8.55)$$

$$\text{即 } P_{S_x=-\frac{1}{2}} = \left| \frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}} \right|^2.$$

现在考虑开头提出的问题。将所有 $S_z = -\frac{1}{2}$ 的粒子剔除后, 粒子的状态为

$$|X\rangle_{z\text{方向筛选后}} = |\frac{1}{2}\rangle_z. \quad (8.56)$$

即 $a=1, b=0$, 因此 $S_x = -\frac{1}{2}$ 的概率为 $P_{S_x=-\frac{1}{2}} = \left| \frac{1}{\sqrt{2}} - \frac{0}{\sqrt{2}} \right|^2 = \frac{1}{2}$. 现在再筛选掉 $S_x = -\frac{1}{2}$ 的粒子, 然后测量 S_z , 所得结果十分神奇。筛选掉所有 $S_x = -\frac{1}{2}$ 后粒子的状态为

$$|X\rangle_{x\text{方向筛选后}} = |\frac{1}{2}\rangle_x. \quad (8.57)$$

要计算测量 S_z 所得结果为 $-\frac{1}{2}$ 的概率需要将状态 $|\frac{1}{2}\rangle_x$ 用 $|\frac{1}{2}\rangle_z, |-\frac{1}{2}\rangle_z$ 展开:

$$|X\rangle_{x\text{方向筛选后}} = \underbrace{|\frac{1}{2}\rangle_x}_{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \frac{1}{\sqrt{2}}\left(\underbrace{|\frac{1}{2}\rangle_z}_{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} + \underbrace{|-\frac{1}{2}\rangle_z}_{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}\right) \quad (8.58)$$

于是 $P_{z=-\frac{1}{2}} = |\langle-\frac{1}{2}|z|X\rangle|^2 = \frac{1}{2}$. 将上述观测结果进行总结:

- 初始时粒子(流)的状态任意, 观测自旋的 z 分量, 筛选掉所有 $S_z = -\frac{1}{2}$ 的粒子, 筛选后粒子的状态为

$$|X\rangle_{z\text{方向筛选后}} = |\frac{1}{2}\rangle_z. \quad (8.59)$$

- 测量筛选后粒子的自旋 z 分量的结果自然是全部为 $+\frac{1}{2}$:

$${}_z\langle\frac{1}{2}||X\rangle_{z\text{方向筛选后}} = 1. \quad (8.60)$$

$${}_z\langle-\frac{1}{2}||X\rangle_{z\text{方向筛选后}} = 0 \quad (8.61)$$

- 再测量粒子的自旋 x 分量发现其中一半为 $S_x = \frac{1}{2}$, 另一半为 $-\frac{1}{2}$. 剔除其中所有 $S_x = -\frac{1}{2}$ 的粒子后粒子的状态为

$$|X\rangle_{x\text{方向筛选后}} = |\frac{1}{2}\rangle_x.$$

- 现在再次测量粒子的自旋 z 分量, 我们得到令人吃惊的测量结果: 一半粒子为 $S_z = \frac{1}{2}$, 另一半为 $S_z = -\frac{1}{2}$. 尽管在第一步已经筛选掉所有 $S_z = -\frac{1}{2}$ 的粒子!

关于这部分内容的精彩讨论(包括实际的观测仪器)可见 Feynman 物理学讲义第三卷²⁶.

²⁶ Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Volume 3*. Addison Wesley, 1st edition, 1971. ISBN 9780201021189

8.6 Heisenberg's Uncertainty Principle Heisenberg 不确定性原理

现在可以讨论量子力学中最奇妙的特征之一了。从上节可知, 对粒子自旋 x 分量的测量会破坏已知的自旋 z 分量的信息, 这一情况在量子力学的很多测量中都存在。我们可以从²⁷ $\hat{S}_x\hat{S}_z \neq \hat{S}_z\hat{S}_x$ 出发来考察这一点, 此式说明先对自旋 z 分量后对 x 分量进行测量和先对 x 分量后对 z 分量进行测量的结果是不同的。这一点并不令人惊奇, 因为对 z 方向自旋进行测量后, 系统应处于 \hat{S}_z 的一

²⁷ 我们把自旋算子和对应的有限维旋转生成元认同, 其满足对易关系 $[J_i, J_j] = J_i J_j - J_j J_i = i\epsilon_{ijk} J_k \neq 0 \rightarrow J_i J_j \neq J_j J_i$. 举个例子来说, 描述自旋 $1/2$ 的粒子应采用对应的二维表示 $J_i = \frac{\sigma_i}{2}$.

个本征态上，而对 x 方向自旋进行测量后，系统应处于 \hat{S}_x 的本征态上， \hat{S}_z 和 \hat{S}_x 的本征态均不同，最后的结果也就不同。

上述现象可以从另一视角描述：**自旋的 x 分量与 z 分量不能同时确定！**每次测量 z 方向的自旋时，都会让 x 方向的自旋信息再次变为未知，反之亦然。这一结论对于 z 方向/ x 方向和 y 方向的自旋也是一样的。

你也许觉得自旋是比较奇怪的物理量，但在粒子位置和动量的测量上也发现了同样的结果。回顾一下式(5.3)，为了方便重写一遍：

$$[\hat{p}_i, \hat{x}_j] = \hat{p}_i \hat{x}_j - \hat{x}_j \hat{p}_i = i\delta_{ij} \quad (8.62)$$

由上可见，对粒子位置的 x 分量进行测量，会导致之前对动量 x 分量的测量结果变得不确定。换句话说，我们不能同时足够精确地确定粒子在同一方向上的位置与动量。注意，只有沿相同方向的位置和动量算符的对易子才不为零²⁸，所以测量动量的 y 分量不会干扰对位置的 x 分量的测量。

每次测量动量会使位置信息变得不确定，反之亦然，这就是著名的 **Heisenberg 不确定原理**。同样的现象也发生在角动量的不同分量之间，因为它们的对易子不为零。总而言之，我们可以计算任意两个物理量的对易性，如果它们不对易，它们就不能同时足够精确地被确定。

这其实不太奇怪。量子力学使用相应对称性的生成元作为测量算符，比如说，对动量的测量等价于平移生成元的作用²⁹，这一生成元将我们的系统平移了一小点，所以粒子的位置也被改变了。真正令人惊奇的是大自然就是这么运作的，许多年来有很多实验验证 Heisenberg 不确定原理，它们都证明了这个原理的正确性。

²⁸Kronecker delta 函数 δ_{ij} 在 $i \neq j$ 时为 0， $i = j$ 时为 1，其定义在附录 B.5.5。

²⁹ 复习：系统的平移不变性导出动量守恒定律。

8.7 Comments on Interpretations 对几种诠释的评议

本章采用的都是量子力学的标准诠释与通用符号，然而还有其他同样有效的体系。例如，就结果而言，Feynman 路径积分体系³⁰与本章描述的波动力学完全等价。但二者的计算过程相当不同。在路径积分体系中计算粒子从点 a 到 b 的概率，需要将 a 到 b 的所有可能的路径叠加。尽管听起来很奇怪，但是可以证明它的计算结果和波动力学相同。Freeman Dyson 曾经讲过关于路径积分的一个故事³¹：

³⁰ 这方面内容可参考 Richard P. Feynman and Albert R. Hibbs. *Quantum Mechanics and Path Integrals*: Emended Edition. Dover Publications, emended edition, 7 2010. ISBN 9780486477220

³¹ Harry Woolf, editor. *Some Strangeness in the Proportion*. Addison-Wesley, 1st edition, 2 1981. ISBN 9780201099249

Dick Feynman 向我讲述他量子力学“历史态叠加”的新理论。“电子它想干啥就干啥”，他说：“它朝着任意方向以任何速度往过去或未来行进，不管它怎么跑，你把这些的振幅全加起来，就能得到波函数。”我对他说：“你疯啦。”但实际上他没疯。

量子力学基本方程的另一种基本诠释（比起路径积分更加偏离主流）是 Bohm 诠释。它的出发点是在 Schrödinger 方程中加入拟设项 Re^{St} 。方程的实部与虚部构成两个方程，其中一个可以类比为经典力学中的 Hamilton-Jacobi 方程（有附加项，附加项可以解释为与量子效应有关的额外势能），进一步的计算就与经典力学非常相似。对额外势求梯度得到新的力，可以带入 Newton 的经典方程 $F = ma$ 中。因此在 Bohm 诠释中仍然存在经典的粒子轨道。据我所知，Bohm 诠释的计算结果与标准非相对论量子力学的计算结果相同。然而由于额外势的非定域性使得它不受欢迎。

阅读建议

- **Richard P. Feynman - The Feynman Lectures on Physics, Vol. 3**³²¹¹ 是量子力学入门的出色教材。量子力学的大部分概念都在此书中有更清楚的解释。
- **David J. Griffiths - Introduction to Quantum Mechanics**³³¹² 是非常易读的、具有启发性的教材。
- **J. J. Sakurai - Modern Quantum Mechanics**³⁴¹³ 是一本精彩著作，它提供了量子力学的独特视角。
- **Paul A. M. Dirac - Lectures on Quantum Mechanics**³⁵¹⁴ 是量子力学的原始文献之一。尽管古老，但时至今日它的评价依然很高，因为阅读量子力学创始人的思想是非常具有启示意义。
- Dirac 旋量的分量意义的详细诠释可参考附录 8.8。

³²Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Volume 3*. Addison Wesley, 1st edition, 1 1971. ISBN 9780201021189

³³David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, 2nd edition, 4 2004. ISBN 9780131118928

³⁴J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley, 1st edition, 9 frn[o]-993. ISBN 9780201539295

³⁵Paul A. M. Dirac and Physics. *Lectures on Quantum Mechanics*. Dover Publications, 1st edition, 3 2001. ISBN 9780486417134

¹¹译注：此书新版可在此免费在线查看：http://www.feynmanlectures.caltech.edu/III_toc.html。

另有新版中译本：费恩曼，莱顿，桑兹著；李洪芳，王子辅，钟万衡译。费恩曼物理学讲义，新千年版，第三卷，上海：上海科学技术出版社，4 2013，ISBN 9787547816387。

¹²译注：国内有由机械工业出版社出版的影印本和中译本，影印本 ISBN 为 9787111182948，贾瑜等的中译本 ISBN 为 9787111278771。

¹³译注：本书第二版在国内有由世界图书出版公司出版的影印本，ISBN 为 9787510035067。

¹⁴译注：国内有科学出版社的影印本。

8.8 Appendix: Interpretation of the Dirac Spinor Components 附录：Dirac 旋量分量的诠释

在接下来的附录中， \mathbf{u}, \mathbf{v} 表示 Dirac 旋量里面的二分量对象， u, v 表示四分量对象。例如， u_1, u_2 描述两个不同的四分量对象。一个一般的四分量对象 u 相应的二分量对象记作 $\mathbf{u}_1, \mathbf{u}_2$ ，即

$$u = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}.$$

目前我们对 Dirac 旋量中的两个 Weyl 旋量的认识还比较模糊。它们究竟代表什么？应该被诠释成什么？此外，Dirac 旋量里的 Weyl 旋量也具有两个分量，这该怎么解释？下面就来讨论。

简言之：

- Dirac 旋量 $\psi \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}$ 中的两个 Weyl 旋量 χ_L, ξ_R 描述“不同的粒子”。然而惯例上仍然称它们是相同的粒子，只是手征性不同，例如对电子而言：

- χ_L 描述左手电子 (left-chiral electron)。
- ξ_R 描述右手电子 (right-chiral electron)。

实际上它们是十分不同的粒子或场³⁶，因为它们不能由宇称变换或电荷共轭相联系。因此采用不同的符号 χ, ξ 表示它们。当然它们存在某些关系，不然为啥写在同一个旋量里呢，这一点稍后再详述。

- Weyl 旋量的两个分量描述粒子不同的自旋构型³⁷。
 - 正比于 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 的 Weyl 旋量描述自旋向上 (spin-up) 的粒子。
 - 正比于 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 的 Weyl 旋量描述自旋向下 (spin-down) 的粒子。
 - 其他 Weyl 旋量仅仅是自旋向上/向下状态的叠加。

下面进行详细推导：

弱相互作用和宇称破缺使得我们明白了左手手性和右手手性是截然不同的粒子这一重要观点。左手手性的粒子携带弱荷（同位旋），因此通过弱力来相互作用。右手手性的粒子则不携带，因此不相互作用，我们在7.7.1节中有过解释。

我们自然界中的每一个粒子都有它对应的反粒子，而其反粒子的描述就是通过对粒子的荷进行共轭操作。荷的共轭操作会翻

³⁶ 它们由不同的量子数标记，因此实验性质相当不同！

³⁷ 自旋构型在8.5.5节讨论。复习一下，自旋与普通角动量的测量方法相似，只是例如自旋 $\frac{1}{2}$ 粒子自旋所有方向分量的测量结果只能是 $\frac{1}{2}$ 或 $-\frac{1}{2}$ ，它们通常称作自旋向上与自旋向下。

转所有的粒子的标签，当然包括同位旋。我们可以看看，对应于电子，其对应什么样的粒子。我们有

- 一个左手手性的电子 χ_L ，同位旋 $-\frac{1}{2}$ ，带电 $-e$ ，处于双重态。
- 一个右手手性的反-左手手性电子 $(\chi_L)^c = \chi_R$ ，将有同位旋 $\frac{1}{2}$ ，电荷 $+e$ ，也同样是处于双重态。这种特性不会因为荷共轭就消失。这可能看起来有点迷茫，但是到现在为止我们只讨论粒子¹⁵间弱力带来的耦合。右手反粒子也会被弱相互作用耦合。
- 右手电子 ξ_R ，同位旋 0，电荷 $-e$ 。
- 左手反-右手手性电子 $(\xi_R)^c = \xi_L$ ，同位旋 0，电荷 $+e$ 。

因此，当讨论电子的时候，我们实际上有四种不同的需要考虑的“东西”。因为他们都是不同的粒子，我们实际上需要给他们不同的名字！一般我们只会说两种：电子和正电子，我们马上会看到这意味着什么。

我们接下来的讨论仅限于问题中的粒子静止的参考系。在其他的参考系中，这些讨论是类似的，但是会有琐碎的细节。

Dirac 旋量和其荷共轭与这四种粒子直接相关。物理的电子和物理的正电子一般被认为是 Dirac 方程的解。Dirac 方程是运动方程，也就是决定粒子的动力学的方程，其解为 Dirac 旋量的时间演化。我们要求出来我们的粒子的时间演化，这就需要我们同时考虑这两种粒子。

附录8.9节中有严格的推导，这里仅给出结论。Dirac 方程有四个线性无关的解，其中两个有形式

$$\psi_i = \begin{pmatrix} u_i \\ u_i \end{pmatrix} \quad (8.63)$$

比如³⁸ $u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-imt}$ 和 $u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-imt}$ ，而另外两个有形式

$$\bar{\psi}_i = \begin{pmatrix} -v_i \\ v_i \end{pmatrix} \quad (8.64)$$

比如 $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{+imt}$ 和 $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{+imt}$ 。

³⁸ 这实际上是一种基的选取。任何选取只需要线性无关就可以。

¹⁵ 而不是反粒子，切记切记

Dirac 方程告诉我们两个粒子的自旋由 Dirac 旋量里面的上下 Weyl 旋量来描述。不仅如此，他们的时间依赖必须一样。这些解就是我们通常认为的不同自旋的**物理电子** (**physical electron**) 和**物理正电子** (**physical positron**)³⁹。

- ψ_1 是自旋向上的电子
- ψ_2 是自旋向下的电子
- $\bar{\psi}_1$ 是自旋向上的电子
- $\bar{\psi}_2$ 是自旋向下的电子

我们可以看到一个物理电子有一个左手（上面的两分量）和右手（下面的两分量）成分。一个物理电子的左右手成分的自旋、时间依赖必须是一样的。就如上面所讨论的，左手和右手成分的行为是非常不一样的，因为他们有不同的弱荷！而为了描述他们的动力学，物理电子必须同时有这两个成分。

注意这些解并不意味着我们用来描述物理电子用的是严格的上和下分量。只需要他们的自旋和时间依赖必须一样。上部分仍然是双重态，而下部分则不是。上部分按照 $SU(2)$ 变化而下部分不是。用前面的符号来看的话，我们有

$$\text{物理电子} = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} \propto \begin{pmatrix} u \\ u \end{pmatrix} \quad (8.65)$$

这并不说明 $\chi_L = \xi_R$ 。Weyl 旋量 χ_L 是有同位旋的双重态，而由 ξ_R 描述的粒子则没有同位旋。不仅如此，我们也知道了 Dirac 旋量里面的上和下 Weyl 旋量在 Lorentz 推动下是不一样变化的。因此，我们使用不同的记号是非常有必要的。换句话说，左手电子 χ_L 带一个 $SU(2)$ 指标，因为 χ_L 按照 $SU(2)$ 双重态的一部分来变换。 ξ_R 则没有指标，按照 $SU(2)$ 单重态来变化呢。

一样的，我们有

$$\text{物理正电子} = \begin{pmatrix} -\xi_L \\ \chi_R \end{pmatrix} \propto \begin{pmatrix} -v \\ v \end{pmatrix} \quad (8.66)$$

我们可以通过荷共轭看到⁴⁰

³⁹ 注意这些解彼此之间的荷共轭。利用荷共轭算符来仔细计算： $\psi_1^c = i\gamma_2\psi_1^*$ ，因此 $(\psi_1)^c = \bar{\psi}_2$ ，而 $(\psi_2)^c = \bar{\psi}_1$ 。

⁴⁰ 见第7.1.5节，并利用(6.13)定义的矩阵的精确形式： $\gamma_2 = \begin{pmatrix} 0 & \bar{\sigma}_2 \\ \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$ 。

$$\begin{aligned}
(\text{物理电子})^c &= i\gamma_2 \begin{pmatrix} \xi_L \\ \chi_R \end{pmatrix}^* = \begin{pmatrix} -\chi_L \\ \xi_R \end{pmatrix} \propto i\gamma_2 \begin{pmatrix} u \\ u \end{pmatrix}^* = \begin{pmatrix} -u^c \\ u^c \end{pmatrix} \\
&= \text{物理正电子}
\end{aligned} \tag{8.67}$$

这告诉我们自然界中观测到的物理电子实际上大部分时候都是两种粒子的叠加：携带同位旋的左手电子和无同位旋的右手电子！一样的我们知道，物理正电子也是由携带同位旋的反-左手电子和无同位旋的反-右手粒子的混合。

Dirac 方程的解则告诉我们粒子如何进行时间演化。比如说我们考虑在一个弱相互作用下出现的一个自旋向上的电子，然后它如何进行时间演化呢？首先他是一个纯的左手电子，可以写为

$$e_L^\uparrow = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{8.68}$$

而它并不是 Dirac 方程的解，因此，为了得到它的时间演化，我们可以用 Dirac 方程的解展开：

$$e_L^\uparrow = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right) = \Psi_1(t=0) - \bar{\Psi}_1(t=0) \tag{8.69}$$

我们知道 Ψ_1 和 $\bar{\Psi}_1$ 如何时间演化

$$\Psi_1(t) - \bar{\Psi}_1(t) \Rightarrow \frac{1}{2} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{-imt} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imt} \right) \tag{8.70}$$

对于 $t=0$ ，这就是左手态，但是演化了一段时间，比如 $t = \frac{\pi}{2m}$ 之

后，我们有

$$\rightarrow \frac{1}{2} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underbrace{e^{-i\frac{\pi}{2}}}_{=-i} - \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underbrace{e^{i\frac{\pi}{2}}}_{=i} \right) = -i \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = -ie_R^\uparrow \quad (8.71)$$

这描述了一个自旋向上的右手电子！这里我们了解到，随着时间演化，左手粒子会变成右手粒子，反之亦然。为了描述电子的演化，我们必须有 e_L 和 e_R ，这也就是我们把它们在一起写成 Dirac 旋量的原因。对正电子亦然。

记得两种不同的粒子 e_L, e_R 有不同的弱荷，即同位旋，而时间演化可以使二者互相转换。大多数时候它们都是两种的混合物，而不是确切的处于某个本征态。同位旋和手性因此在时间演化中就不守恒。

我们现在可以看到 Dirac 旋量的符号是必须的，因为我们有一个封闭的将本节开始提到的四种粒子中任意两种联系起来的动力学行为。手性和同位旋在传播过程中也就不守恒了。可以发现，同一个电子，随着时间演化有时处于左手态，有时也会处于右手态。

8.9 Appendix: Solving the Dirac Equation 附录：解 Dirac 方程

和上节一样，我们把 Dirac 自旋量中的二分量量记做 u 和 v ，把四分量量记做 u 和 v 。也就是说例如 u_1, u_2 的记号代表的是四分量量，在不特别区分这些量的时候我们只写 u 。同时 u_1 和 u_2 是这个四分量量中的两个二分量量： $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ 。

本附录中我们将在手性基底上解静止参考系下的 Dirac 方程。其他任意参考系的解能够通过本节解做伪转动变换得到。除了我们在上一节的讨论，我们在第9章讨论量子场论的时候也会用到这些解。

Dirac 方程为：

$$(i\partial_\mu \gamma^\mu - m)\psi = 0 \quad (8.72)$$

我们考虑其平面波解，设 $\Psi = ue^{-ipx}$ ，其中 u 是一四分量量，因

为矩阵 γ_μ 是 4×4 的。在静止参考系中, 三维动量 $\vec{p} = 0$, 指数项因此变为 $-ipx = -i(p_0x_0 - \vec{p}\vec{x}) = -ip_0x_0$ 。现在使用我们本章初推导得到的相对论能量动量关系 $E = \sqrt{|\vec{p}|^2 + m^2}$, 已知 $p_0 = E$ 和 $x_0 = t$, 我们得到 $-ipx = -iEt = -i\sqrt{|\vec{p}|^2 + m^2}t = -imt$ 。把此式带入到 Dirac 方程中:

$$\begin{aligned}
 (i\partial_\mu\gamma^\mu - m)ue^{-imt} &= 0 \\
 \rightarrow (i(\partial_0\gamma^0 - \partial_i\gamma_i) - m)ue^{-imt} &= 0 \\
 \rightarrow (i(-im)\gamma^0 - m)ue^{-imt} &= 0 \\
 \rightarrow (m\gamma^0 - m)u &= 0 \\
 \rightarrow \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)u &= 0 \\
 \rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} &= 0 \\
 \rightarrow \begin{pmatrix} -u_1 + u_2 \\ u_1 - u_2 \end{pmatrix} &= 0
 \end{aligned} \tag{8.73}$$

注意这里矩阵中的 1 实际上代表的是 2×2 的单位矩阵, 所以 u_1 和 u_2 是二分量子。我们看出我们假设的平面波解在 $u_1 = u_2$ 的条件下符合方程, 由此我们找到了 Dirac 方程两个线性独立的解:

$$\Psi_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{-imt}, \quad \Psi_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} e^{-imt} \tag{8.74}$$

我们可以通过设 $\tilde{\Psi} = ve^{ipx}$ 来得到另外两个解, 在静止参考系下这一解形式化为 $\tilde{\Psi} = ve^{imt}$, 带入方程得

$$\begin{aligned}
 (i\partial_\mu\gamma^\mu - m)ve^{imt} &= 0 \\
 \rightarrow (-m\gamma^0 - m)v &= 0 \\
 \rightarrow \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= 0 \\
 \rightarrow \begin{pmatrix} -v_1 - v_2 \\ -v_1 - v_2 \end{pmatrix} &= 0
 \end{aligned} \tag{8.75}$$

我们由此得到时间依赖为 e^{imt} 的两个解，其满足 $-v_1 = v_2$ 的条件。此时两个线性独立的解为：

$$\tilde{\Psi}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} e^{imt}, \quad \tilde{\Psi}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} e^{imt} \quad (8.76)$$

8.10 Appendix: Dirac Spinors in Different Bases 附录：Dirac 旋量的不同的基

在拉格朗日量中，Dirac 旋量 ψ 总是和矩阵 γ_μ 组合出现。这个特性可以通过转换到一个不同的基来简化计算。这种简化之所以行得通，是因为对于任意的可逆矩阵 N ，我们都可以通过在 ψ 和 γ_μ 之间插入形如 $1 = N^{-1}N$ 的项来重新定义两者。

$$\partial_\mu \bar{\psi} \gamma_\mu \psi = \partial_\mu \bar{\psi} \underbrace{N^{-1}N}_{=1} \gamma_\mu \underbrace{N^{-1}N}_{=1} \psi = \partial_\mu \underbrace{\bar{\psi} N^{-1}}_{\equiv \bar{\psi}'} \underbrace{N \gamma_\mu N^{-1}}_{\equiv \gamma'_\mu} \underbrace{N \psi}_{\equiv \psi'} \quad (8.77)$$

目前为止我们在本书中使用的基被称为手性基或者 Weyl 基。而习惯上 Dirac 方程的求解中用的是另一种基，被称为质量基或者 Dirac 基。在我们目前为止使用的手性或 Weyl 基中，Dirac 拉格朗日量

$$\mathcal{L}_D = i\chi_L^\dagger \sigma^\mu \partial_\mu \chi_L + i\xi_R^\dagger \bar{\sigma}^\mu \partial_\mu \xi_R - m\chi_L^\dagger \xi_R - m\xi_R^\dagger \chi_L \quad (8.78)$$

具有非对角的质量项——也就是混合了不同的态的质量项。我们可以利用选择基的自由度来找到一组使得质量项对角的基，并将之称为质量基。

这意味着我们想要形如 $\psi^\dagger m \psi$ 的质量项，其中 $m = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$ ，

于是我们就得到了质量项的形式是¹⁶

$$\bar{\psi}' M' \psi' = (\psi')^\dagger \gamma'_0 M' \psi' = \begin{pmatrix} u' \\ v' \end{pmatrix}^\dagger \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} = (u')^\dagger m_1 u' + (v')^\dagger m_2 v', \quad (8.79)$$

而我们正在处理的是

$$\bar{\psi} M \psi = \psi^\dagger \gamma_0 M \psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}^\dagger \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix} = m \chi_L^\dagger \xi_R + m \xi_R^\dagger \chi_L. \quad (8.80)$$

我们目前为止一直使用的后一个基很容易体现出手性，然而质量基下的 Dirac 旋量会更容易和物理中的传播粒子建立联系。

为了得到第二个形式和第一个形式间的关系，我们需要把矩阵 $M = \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 对角化。这可以通过矩阵 $N = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ 来实现，也就是说

$$\underbrace{N^{-1} \begin{pmatrix} -m & 0 \\ 0 & m \end{pmatrix} N}_{\equiv M'} = M \quad (8.81)$$

$$\rightarrow m \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (8.82)$$

于是我们就可以相应地重新定义 Dirac 旋量

$$\begin{aligned} \bar{\psi} M \psi &= \bar{\psi} \underbrace{N N^{-1}}_{=-1} M \underbrace{N N^{-1}}_{=-1} \psi \\ &= \underbrace{\bar{\psi} N}_{\equiv \bar{\psi}'} \underbrace{N^{-1} M N}_{\equiv M'} \underbrace{N^{-1} \psi}_{\equiv \psi'} \\ &= \bar{\psi}' M' \psi' \end{aligned} \quad (8.83)$$

¹⁶此处原书中不带', 应为印刷错误, 但是勘误表中没有。带 ' 应为正确形式, 故改正并说明。

观察一下在这个基下手性投影算符 $P_L = \frac{1-\gamma_5}{2}$ 的形式是很有帮助的。因此我们需要得出相应的 γ_5 矩阵

$$\begin{aligned}\bar{\gamma}_5 &= N^{-1}\gamma_5 N = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\end{aligned}\tag{8.84}$$

与此对应的本征向量是 $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ 和 $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 。这表示手性本征态现在由具有上下分量的 Dirac 旋量描述, 例如, 左手态在这个基下的形式就是 $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ 相比之下, 在手性基下 γ_5 是对角的, 左手本征态由只有上分量的 Dirac 旋量 $\psi_L = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix}$ 给出, 右手本征态则由只有下分量的 Dirac 旋量 $\psi_R = \begin{pmatrix} 0 \\ \xi_R \end{pmatrix}$ 给出。

在这个基下手性投影算符写成

$$P_L = \frac{1-\gamma_5}{2} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\tag{8.85}$$

8.10.1 Solutions of the Dirac Equation in the Mass Basis 质量基下 Dirac 方程的解

我们可以求解质量基下的 Dirac 方程

$$(i\gamma_\mu \partial^\mu - m)\Psi = 0\tag{8.86}$$

通过假设 $\psi = ue^{-ipx}$, 可以导出

$$(\gamma_\mu p^\mu - m)ue^{-ipx} = 0$$

$$\rightarrow (\gamma_\mu p^\mu - m)u = 0.$$

等价地，我们可以假设 $\psi = ve^{-ipx}$ 导出

$$(-\gamma_\mu p^\mu - m)ve^{ipx} = 0$$

$$\rightarrow (-\gamma_\mu p^\mu - m)v = 0.$$

类似在手性基中的解，我们这里也在静止坐标系中求解，也就是 $\vec{p} = 0$ 。我们做这样的选择之所以可行，是因为物理规律在所有参照系中都是相同的，所以我们可以选择最符合我们需求的一个。在这个参照系中，由于 $p_i = 0$ ，所以我们有

$$\rightarrow (\gamma_0 p^0 - m)u = 0$$

$$\rightarrow (-\gamma_0 p^0 - m)v = 0.$$

此外还有 $p_0 = E$ ，并且我们可以使用我们在本章开头（式(8.2)）推导的相对论能-动量关系。在静止参照系中我们有 $E = \sqrt{(p_i)^2 + m^2} = m$ 。现在我们可以利用上面给出的矩阵 N 和 $\gamma'_0 = N^{-1}\gamma_0 N$ 计算出 γ_0 在质量基或 Dirac 基中的精确形式。要记得在 m 中我们还包含着一个单位矩阵，因此

$$\rightarrow \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} m - m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) u = 0$$

$$\rightarrow \left(- \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} m - m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) v = 0.$$

$$\rightarrow \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} u = 0$$

$$\rightarrow \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} v = 0.$$

要记得每个 Dirac 旋量都包含两个二分量的变量，我们可以推出 u 的下面的二分量变量和 v 的上面的二分量变量必定是 0：

$$\rightarrow \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2u_2 \end{pmatrix} = 0 \rightarrow u_2 = 0$$

$$\rightarrow \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -2v_1 \\ 0 \end{pmatrix} = 0 \rightarrow v_1 = 0.$$

我们可以看到在这个基下物理的传播粒子（也就是 Dirac 方程的解）可以用只含有上分量的旋量描述，而等价地对于反粒子则只含有下分量。从而我们可以再次得到四个线性独立的解

$$\Psi'_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \quad \Psi'_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} e^{-imt} \quad (8.87)$$

和

$$\Psi'_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} e^{imt} \quad \Psi'_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{imt} \quad (8.88)$$

在这个参照系中这个基下的一般解，可以写成线性组合的形式

$$\psi = ue^{-ipx} + ve^{ipx} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix} e^{-ipx} + \begin{pmatrix} 0 \\ v_1 \end{pmatrix} e^{ipx} \quad (8.89)$$

并且我们可以通过将这个解做一个 Lorentz 推动变换得到任意一个参照系中的解。此外，最一般的解是所有动量和自旋态⁴¹的解的叠加

⁴¹ 注意 Weyl 旋量的两个分量代表不同的自旋态。

$$\Psi = \sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{E_p}} \left(c_r(p) u_r(p) e^{-ipx} + d_r^\dagger(p) v_r(p) e^{ipx} \right) \quad (8.90)$$

9. Quantum Field Theory 量子场论

总结 剧透

译注：本节的内容是介绍本章将要做些什么也就是剧透，第一遍读不懂属于正常情况。

本章介绍量子场论的基本框架。由第5章导出的公式

$$[\Phi(x), \pi(y)] = i\delta(x - y)$$

可见场本身就是算符。自旋为 0, 1/2, 1 的场的运动方程的解是用场的 **Fourier 展开**¹写出来的。从上式的对易关系可见 Fourier 系数也是算符。之后，我们会看到这些**算符**（当然还包括导出它们的场）是如何**产生和湮灭**粒子的，用对应场的拉格朗日量可以导出代表能量的哈密顿算符。

之后建立**相互作用理论**，它是量子场论的核心。我们会看到在相互作用理论中，哈密顿量是由自由场的哈密顿量和**相互作用的哈密顿量**线性组合而成的。这一思路之后可应用于**相互作用绘景**中，在这个绘景下场的时间演化是由自由哈密顿量主导的，而态的时间演化是由相互作用哈密顿量主导的。在这个绘景下，能够推导出散射过程的几率幅，用 Dirac 记号可以表示成

¹Fourier 变换背后的思想在附录D.1中进行了解释。

$$\langle f | \hat{S} | i \rangle,$$

这里 \hat{S} 表示描述散射过程的算符, $|i\rangle$ 是初态, $\langle f|$ 是末态。我们会发现算子 \hat{S} 可以用相互作用哈密顿量 H_I 写成

$$\hat{S}(t, t_i) = e^{-i \int_{t_i}^t dt' H_I}.$$

这个式子解不出来, 因此采用级数展开式来计算这个指数。对于大多数实验而言, 前面几项就足以精确的描述了。

级数展开中的每一项可以解释为描述不同种类的散射过程。相互作用哈密顿量包含了场的线性组合, 就如上面提到的, 是用来产生或者湮灭场的。对于最低的非平凡阶次, 我们得到 8 项, 且第一项描述了一个从 $e^-e^+ \rightarrow \gamma$ 的散射过程。这意味着我们从一个由电子和正电子组成的初态 $|e^-e^+\rangle$ 出发, 这个初态被自旋 1/2 的场算符湮灭, 并且之后一个光子 $\langle \gamma|$ 由光子场产生。其他项作用在初态 $|e^-e^+\rangle$ 时的结果是 0。

这个级数的下一阶由许多项组成, 我们只考虑其中的一项, 同样地, 从初态 $|e^-e^+\rangle$ 开始, 其中的一项描述了过程 $e^-e^+ \rightarrow \gamma \rightarrow e^-e^+$, 这里初末的正负电子对一般具有不同的动量。

同样地, 所有这些项可以理解成所有的相互作用哈密顿量。一种简化这类计算的图形化方法是著名的 Feynman 图。在这个图中每一条线, 每一个顶点代表了上面计算的东西的一个因子。

9.1 Field Theory Identifications 对应到场理论

本节探究从对称性限制得到的拉格朗日量是怎样应用到场论的框架中的。要得到一个描述自然界的场论, 第一步是将拉格朗日量和第5章内容, 尤其是(5.5)式¹结合起来, 为了方便, 将(5.5)式重写一遍:

$$[\Phi(x), \pi(y)] = i\delta(x - y) \quad (9.1)$$

其中共轭动量密度 $\pi(y)$ 为

$$\pi(y) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi(y))} \quad (9.2)$$

¹译注: 似乎应该为(5.6)式。

9.2 Free Spin 0 Field Theory 自旋为 0 的自由场理论

“一个产生的艺术首先是一个毁灭的艺术。”

——Pablo Picasso²

从最简单的例子——自由的 0 自旋场——出发,它由在 Lorentz 变换下不变的标量来描述(在 3.7.4 节推导)。我们已经在 6.2 节中推出了相对应的拉格朗日量²

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \quad (9.3)$$

它的运动方程叫做 Klein-Gordon 方程

$$(\partial_\mu \partial^\mu + m^2)\Phi = 0 \quad (9.4)$$

可直接计算出共轭动量

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi(x))} = \frac{\partial}{\partial(\partial_0 \Phi(x))} \frac{1}{2}(\partial_\mu \Phi(x) \partial^\mu \Phi(x) - m^2 \Phi^2(x)) = \partial_0 \Phi(x)$$

Klein-Gordon 方程的最一般的解可以写成 Fourier 展开的形式³

$$\Phi(x) = \int dk^3 \frac{1}{(2\pi)^3 2\omega_k} (a(k)e^{-ikx} + b(k)e^{ikx}) \quad (9.5)$$

其中³ $\omega_k^2 \equiv \vec{k}^2 + m^2$, 如果限制在实标量场的情况上式可写为

$$\Phi(x) = \int dk^3 \frac{1}{(2\pi)^3 2\omega_k} (a(k)e^{-ikx} + a^\dagger(k)e^{ikx}) \quad (9.6)$$

因为 $c + c^\dagger = \underbrace{\text{Re}(c) + i \cdot \text{Im}(c)}_c + \underbrace{\text{Re}(c) - i \cdot \text{Im}(c)}_{c^\dagger} = 2\text{Re}(c)$.

现在考虑方程(9.1), 即不为零的对易子 $[\Phi(x), \pi(y)] \neq 0$ 有什么含义。这意味着 $\Phi(x)$ 和 $\pi(y)$ 不是普通的函数而必须是算符, 因为一般的函数是对易的: $(3+x)(7xy) = (7xy)(3+x)$ 。由方程(9.6)可以得到结论: Fourier 系数 $a(k)$ 和 $a(k)^\dagger$ 是算符, 因为 $e^{\pm ikx}$ 只是复数, 而复数是对易的。

从方程(9.1)可以计算出⁴

$$[a(k), a^\dagger(k')] = (2\pi)^3 \delta^3(\vec{k} - \vec{k}') \quad (9.7)$$

² 引自 Rollo May, *The Courage to Create*. W.W.Norton and Company, reprint edition, 3 1994. ISBN 9780393311068

³ 如果想知道关于积分测度的详细的计算过程, 以及为什么写出这种形式的解的验证请看本章最后 9.6 的附录。

⁴ 例如可以参考 Lewis H. Ryder 书的 4.1 节, *Quantum Field Theory*. Cambridge University Press, 2nd edition, 6 1996. ISBN 9780521478114

² 原文下式有误, 译文已按勘误表修改。

³ 原文此式有误, 译文已按勘误表修改。

$$[a(k), a(k')] = 0 \quad (9.8)$$

$$[a^\dagger(k), a^\dagger(k')] = 0 \quad (9.9)$$

既然已经知道场就是算符，接下来自然要问：它作用在什么上面？在粒子理论中，动力学变量与算符相等同，算符作用的数学对象（波函数，抽象的 Dirac 右矢等等）描述粒子。在场论中，目前没有什么东西来描述粒子。在这一点上，粒子是怎么出现在场论中是完全不清楚的。然而，通过考虑场的展开系数 $a(k)$ 和 $a^\dagger(k)$ 作用于什么抽象的数学对象上，就可以知道场作用在什么对象上的了。为了建立直观认识，我们从熟悉的能量入手。

一个标量场的能量 E 由时间平移不变性导出的方程(4.40)给出：

$$\begin{aligned} E &= \int d^3x T^{00} \\ &= \int d^3x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi)} \underbrace{\frac{\partial \Phi}{\partial x_0}}_{\partial_0 \Phi} - \mathcal{L} \right) \\ &= \int d^3x \left((\partial_0 \Phi)^2 - \frac{1}{2} (\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2) \right) \\ &\stackrel{\partial_\mu \partial^\mu = \partial_0 \partial_0 - \partial_i \partial_i}{=} \frac{1}{2} \int d^3x \left((\partial_0 \Phi)^2 + (\partial_i \Phi)^2 + m^2 \Phi^2 \right) \end{aligned} \quad (9.10)$$

将方程(9.6)代入方程(9.10)，并利用对易关系 (9.7) ~ (9.9) 式可得

$$\begin{aligned} E &= \frac{1}{2} \int dk^3 \frac{1}{(2\pi)^3} \omega_k \left(a^\dagger(k) a(k) + a(k) a^\dagger(k) \right) \\ &\stackrel{(9.7) \text{ 式}}{=} \int dk^3 \frac{1}{(2\pi)^3} \omega_k \left(a^\dagger(k) a(k) + \frac{1}{2} (2\pi)^3 \delta^3(0) \right) \end{aligned} \quad (9.11)$$

这时我们的理论发散了，积分中的第二项是无穷大的。我们可以在这里停下并且说这个理论无效。然而，一些勇敢的人进行了深入的探索，忽略无穷大项并且发现了一个精确描述自然的理论。从这里继续的标准的做法就是忽略第二项，这种做法的可行性尚未得到解释。不过这里的关键在于发散项出现在每个系统的能量中，并且我们只能测量能量的差值。因此这个无穷大的常数项不会对测量造成影响。

按照习惯，写作算符的能量叫做哈密顿量 \hat{H} ，我们可以计算

\hat{H} 和 Fourier 系数 $a(k)$ 以及 $a^\dagger(k)$ 的对易子⁵。我们得到了

⁵ 之后会看到它有什么用。

$$\begin{aligned}
 [\hat{H}, a^\dagger(k')] &= \int dk^3 \frac{1}{(2\pi)^3} \omega_k [a^\dagger(k) a(k), a^\dagger(k')] \\
 &\stackrel{[a^\dagger(k), a^\dagger(k')]=0}{=} \int dk^3 \frac{1}{(2\pi)^3} \omega_k a^\dagger(k) [a(k), a^\dagger(k')] \\
 &= \int dk^3 \omega_k a^\dagger(k) \delta^3(k - k') \\
 &\stackrel{\text{见附录 D.2}}{=} \omega_{k'} a^\dagger(k')
 \end{aligned} \tag{9.12}$$

同样的

$$[\hat{H}, a(k')] = -\omega_{k'} a(k') \tag{9.13}$$

量子理论中的算符作用于描述系统状态的数学对象上（见8.3节）。在这个情况下，如果用能量算符，即哈密顿算符 \hat{H} 作用在抽象的对象 $|?\rangle$ 上来描述物理系统，就能得到系统的能量：

$$\hat{H}|?\rangle = E|?\rangle \tag{9.14}$$

现在回到开始的那个问题：场⁶是如何作用于物理系统的？先来考虑⁷第一个 Fourier 系数（它是个算符）对于系统能量的影响：

⁶ 记住：场 = 算符

⁷ 这里我们采取了非常聪明的处理方式，它是 Dirac 在处理量子力学中谐振子的问题时首先被引入的

$$\begin{aligned}
 \hat{H}(a(k')|?\rangle) &= \left(a(k')\hat{H} + \underbrace{\hat{H}a(k') - a(k')\hat{H}}_{[\hat{H}, a(k')]} \right) |?\rangle \\
 &= a(k') \underbrace{\hat{H}|?\rangle}_{=E|?\rangle} + [\hat{H}, a(k')]|?\rangle \\
 &= (a(k')E + [\hat{H}, a(k')])|?\rangle \\
 &\stackrel{\text{方程(9.13)}}{=} (a(k')E - \omega_{k'}a(k'))|?\rangle \\
 &= (E - \omega_{k'})a(k')|?\rangle
 \end{aligned} \tag{9.15}$$

同样地，对于第二个 Fourier 系数有

$$\hat{H}a^\dagger(k')|?\rangle = (E + \omega_{k'})a^\dagger(k')|?\rangle \tag{9.16}$$

它们该如何解释？我们看到 $a(k')|?\rangle$ 可以理解成一个具有能量 $E -$

ω_k 的新系统。更确切的，我们定义

$$|?_2\rangle \equiv a(k')|?\rangle$$

于是

$$\hat{H}|?_2\rangle \underbrace{=}_{(9.15)\text{式}} (E - \omega_{k'})|?_2\rangle$$

它暗示了场的作用效果。想象一个真空态系统 $|0\rangle$ ，依定义有 $H|0\rangle = 0|0\rangle$ 。如果用 $a^\dagger(k')$ 作用于真空态 $|0\rangle$ ，由上可知它将真空的体系变到了一个具有能量 $\omega_{k'}$ 的体系

$$\hat{H}a^\dagger(k')|0\rangle \underbrace{=}_{(9.16)\text{式}} \omega_{k'}a^\dagger(k')|0\rangle \quad (9.17)$$

可见 $a^\dagger(k')$ 在一个全空的体系中产生了一些具有能量为 $\omega_{k'}$ 的东西，我们把它叫做一个动量为 k' 的粒子！如果再用 $a^\dagger(k')$ 作用在这个系统上一次，就得到了具有相同动量的第二个粒子。如果用 $a^\dagger(k'')$ 作用，就可以得到具有动量 k'' 的粒子等等。因此， a^\dagger 称为产生算符。与 $a^\dagger(k')$ 的理解过程类似， $a(k')$ 破坏或者湮灭一个具有能量为 $\omega_{k'}$ 的粒子，所以叫做湮灭算符或消灭算符。为了使其更加精确，我们引入关于粒子态的记号

$$a^\dagger(k)|0\rangle \equiv |1_k\rangle \quad (9.18)$$

$$a^\dagger(k)|1_k\rangle \equiv |2_k\rangle \quad (9.19)$$

$$a^\dagger(k')|2_k\rangle \equiv |2_k, 1_{k'}\rangle \quad (9.20)$$

再来看一下能量

$$E = \int dk^3 \frac{1}{(2\pi)^3} \omega_k a^\dagger(k) a(k).$$

如果这个算符作用在像 $|2_{k_1}, k_2\rangle$ 这样的态上会发生什么呢？结果应该是

$$E = 2\omega_{k_1} + \omega_{k_2}$$

它是两个能量为 ω_{k_1} 和一个能量为 ω_{k_2} 的粒子的能量之和。因此，

出现在这里的算符

$$N(k) \equiv a^\dagger(k)a(k) \quad (9.21)$$

是粒子数算符，记作 $N(k)$ ，它可以从一个态中提取出动量为 k 的粒子的数目

$$N(k)|n_k, n'_k, \dots\rangle = n_k|n_k, n'_k, \dots\rangle. \quad (9.22)$$

这时，能量算符可以写成

$$E = \int dk^3 \frac{1}{(2\pi)^3} \omega_k N(k)$$

此外，注意到有些物理系统的动量谱是离散而非连续的⁸，对于这些系统，所有的积分变成了求和，比如能量的形式为

$$E = \sum_k \omega_k N(k).$$

对易关系变成了

$$[a(k), a^\dagger(k')] = \delta_{k,k'} \quad (9.23)$$

注意量子场论就像量子力学一样，是一个做概率性预言的理论。因此物理态需要归一化 $\langle k, k', \dots | k, k', \dots \rangle = 1$ ，因为超过 $100\% = 1$ 的概率没有意义。如果用一个像 $a^\dagger(k)$ 这样的算符作用在右矢上，所得新的右矢的模不一定是 1，因此可以写为

$$a^\dagger(k)|n_k\rangle = C|n_k + 1\rangle \quad (9.24)$$

其中 n_k 是动量为 k 的粒子数， C 是某个常数。上式取厄米共轭得到⁹：

$$\begin{aligned} (a^\dagger(k)|n_k\rangle)^\dagger &= (C|n_k + 1\rangle)^\dagger \\ \rightarrow \langle n_k|a(k) &= \langle n_k + 1|C^\dagger \end{aligned} \quad (9.25)$$

⁸ 例如硬盒中的粒子，它的动量谱是离散的。假设问题中的系统是局限在有限的体积 V 中是量子场论中一个常用的技巧，在计算的最后我们再取 $V \rightarrow \infty$ 这个极限。

⁹ $|n_k\rangle^\dagger = \langle n_k|$ ，当然也有 $(a^\dagger)^\dagger = a$ 。

由此可得

$$\underbrace{\langle n_k | a(k)}_{\text{方程(9.25)}} \underbrace{a^\dagger(k) | n_k \rangle}_{\text{方程(9.24)}} = \langle n_k + 1 | \underbrace{C^\dagger C}_{\text{一个数而不是算符}} | n_k + 1 \rangle = C^\dagger C \underbrace{\langle n_k + 1 | n_k + 1 \rangle}_{=1} \quad (9.26)$$

或者用离散的对易关系 (方程(9.23)) 写为

$$\begin{aligned} \langle n_k | a(k) a^\dagger(k) | n_k \rangle &= \langle n_k | (\underbrace{a^\dagger(k) a(k)}_{=N(k), \text{见(9.21)式}} + \underbrace{\delta_{k,k}}_{=1}) | n_k \rangle \\ &\stackrel{\text{方程(9.22)}}{=} \underbrace{\langle n_k | (n_k + 1) | n_k \rangle}_{\text{一个数而不是算符}} = (n_k + 1) \underbrace{\langle n_k | n_k \rangle}_{=1} \end{aligned} \quad (9.27)$$

结合方程(9.26)和(9.27)得到

$$C^\dagger C = n_k + 1 \rightarrow C = \sqrt{n_k + 1} \quad (9.28)$$

从而有

$$a^\dagger(k) | n_k \rangle = \sqrt{n_k + 1} | n_k + 1 \rangle \quad (9.29)$$

同样的步骤可得⁴

$$a(k) | n_k \rangle = \sqrt{n_k} | n_k - 1 \rangle \quad (9.30)$$

这时出现了两个问题。首先，如果在全空的系统中湮灭一个粒子会发生什么？第二：能量守恒和电荷守恒怎么办？如何能够从真空产生一些东西而不违背守恒律？首先，守恒律从来没有被违背，但是这一点只有在进一步深入这个理论之后才会变的清晰起来。Richard Feymann 也有同样的问题¹⁰，看一看或许对我们有帮助

¹⁰Feynman 的 Nobel 演讲 (December 11, 1965)

我记得当某个人教我关于产生和湮灭算符的时候，当说到这个算符产生了一个电子，我说：“你怎么能够产生一个电子？它不符合电荷守恒。”就这样，我从思想上的抵触阻碍了学习一个非常实际的计算方法。

⁴原文此式有误，译文已按勘误表修改。

第二, 我们当然永远不能湮灭那些一开始就不存在的东西, 如果湮灭算符作用到一个全空的体系 $|0\rangle$, 从方程(9.30)得到

$$a(k)|0_k\rangle = \sqrt{0}|0_k - 1_k\rangle = 0 \quad (9.31)$$

或者, 同样地

$$a(k')|1_k\rangle = \sqrt{0}|1_k, 0_{k'} - 1_{k'}\rangle = 0 \quad (9.32)$$

由此可见, 如果湮灭算符 $a(k')$ 作用到一个不包含动量 k' 的粒子的右矢 (例如 $|k\rangle$) 上, 结果依然是 0. 产生算符和湮灭算符出现在场的 Fourier 展开中, 这个 Fourier 展开包含了对于所有可能动量的积分 (或求和), 因此当这些场作用到一个像 $|k\rangle$ 的右矢上时, 只有一个湮灭算符会导致不为 0 的结果。这一点在我们试图用量子场论来描述相互作用的时候是非常重要的。

在开始研究相互作用之前, 我们先来简单的看看自旋 $1/2$ 和自旋 1 的自由场。

9.3 Free Spin $\frac{1}{2}$ Theory 自旋 $\frac{1}{2}$ 的自由场理论

在6.3节已经推导了自旋 $\frac{1}{2}$ 自由场的运动方程——Dirac 方程:

$$(i\gamma_\mu\partial^\mu - m)\Psi = 0$$

Dirac 方程的一般解可以写成如下形式¹¹

$$\begin{aligned} \Psi &= \sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{\omega_p}} \left(c_r(p) u_r(p) e^{-ipx} + d_r(p) v_r(p) e^{+ipx} \right) \\ &= \Psi^+ + \Psi^- \end{aligned} \quad (9.33)$$

与前面不同, 这里不局限于实数场, 因为6.3节表明 Lorentz 不变的拉格朗日量需要复的自旋 $\frac{1}{2}$ 场。按照惯例将通解写成

$$\Psi = \sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{\omega_p}} \left(c_r(p) u_r(p) e^{-ipx} + d_r^\dagger(p) v_r(p) e^{+ipx} \right) \quad (9.34)$$

¹¹Dirac 方程的求解过程见(8.9)节, 这与上节 Klein-Gordon 方程的过程相似。

因为这样, $d_r^\dagger(p)$ 可以看作是产生了一个反粒子。如果在解中把它命名为 $d_r(p)$ 会导致混淆, 因为 $c_r^\dagger(p)$ 产生而 $c_r(p)$ 湮灭粒子。将 Fourier 系数命名为 $d_r^\dagger(p)$ 而非 $d_r(p)$, 使得反粒子也有相似的解释: $d_r^\dagger(p)$ 产生而 $d_r(p)$ 湮灭。

6.3节还导出了伴随 Dirac 方程:

$$(i\partial^\mu \bar{\Psi} \gamma_\mu + m \bar{\Psi}) = 0$$

其解为

$$\begin{aligned} \Psi &= \sum_r \sqrt{\frac{m}{(2\pi)^3}} \int \frac{d^3p}{\sqrt{\omega_p}} \left(c_r^\dagger(p) \bar{u}_r(p) e^{+ipx} + d_r(p) \bar{v}_r(p) e^{-ipx} \right) \\ &= \Psi^+ + \Psi^- \end{aligned} \quad (9.35)$$

¹² 8.10节导出了静止参考系 ($p_i = 0$) 下的旋量基底, 任意参考系下的旋量基底可以通过把静止参考系下的旋量基底做一个推动变换得到。

其中 u_1, u_2, v_1, v_2 是任意参考系下的“旋量基底”。¹²

$$u_1 = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_3}{E+m} \\ \frac{p_1+ip_2}{E+m} \end{pmatrix}, \quad u_2 = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 0 \\ 1 \\ \frac{p_1-ip_2}{E+m} \\ \frac{-p_3}{E+m} \end{pmatrix} \quad (9.36)$$

$$v_1 = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_1-ip_2}{E+m} \\ \frac{-p_3}{E+m} \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} \frac{p_3}{E+m} \\ \frac{p_1+ip_2}{E+m} \\ 1 \\ 0 \end{pmatrix}. \quad (9.37)$$

¹³ 这微小的差异及其重要! 如果自旋 $\frac{1}{2}$ 场和标量场完全相同, 那么宇宙中没有任何东西是稳定的。

¹⁴ $[\Phi(x), \pi(y)] = \Phi(x)\pi(y) - \pi(y)\Phi(x) = i\delta(x-y)$.

自旋 $\frac{1}{2}$ 自由场理论其余的推导过程和标量理论相似, 但是有一点小的不同。¹³可以看出标量场中用到的对易关系并不适用于自旋 $\frac{1}{2}$ 场。在运动方程 (即 Dirac 方程) 的通解中系数 c^\dagger 产生粒子, 而 d^\dagger 产生反粒子。利用对易关系¹⁴计算自旋 $\frac{1}{2}$ 场的哈密顿量, 会得到如下形式:

$$H \sim \int c^\dagger c - d^\dagger d$$

这一形式表明反粒子的能量是负的。这是一个严重的问题, 因为任

意一个态都可以无休止的衰变到更低能量的状态，宇宙中的任何东西都不能是稳定的。和标量情况不同，这里的哈密顿量不是有下界的。但如果用反对易关系来替代：

$$\{\Phi(x), \pi(y)\} = \Phi(x)\pi(y) + \pi(y)\Phi(x) = i\delta(x-y)$$

则哈密顿量含有反粒子产生和湮灭算符的那一项变成了正的，问题也就解决了。

这个看上去很笨拙的技巧，却产生了非常有趣的结果。例如¹⁵

$$\{c^\dagger(k), c^\dagger(k')\} = 0$$

由此可得

$$\begin{aligned} \{c^\dagger(k), c^\dagger(k)\} &= c^\dagger(k)c^\dagger(k) + c^\dagger(k)c^\dagger(k) = 2c^\dagger(k)c^\dagger(k) = 0 \\ \Rightarrow c^\dagger(k)c^\dagger(k) &= 0 \end{aligned} \quad (9.38)$$

两个相同的产生算符作用得到的结果总是 0！¹⁶。这意味着不能产生两个相同的自旋 $\frac{1}{2}$ 粒子，这就是著名的 **Pauli 不相容原理**。

其他的 Fourier 系数反对易关系可以通过场和共轭动量的反对易关系得到：

$$\begin{aligned} \{c_r(p), c_s^\dagger(p')\} &= \delta_{rs}\delta(p-p') \\ \{d_r(p), d_s^\dagger(p')\} &= \delta_{rs}\delta(p-p') \end{aligned} \quad (9.39)$$

其他可能的组合都等于 0。因此，这些系数与上一节自旋 0 的场的系数具有相同的性质。它们产生和湮灭粒子，不同的是相同的算符作用两次的结果是 0。自旋 $\frac{1}{2}$ 自由场的哈密顿量可以和自旋 0 的自由标量场用相似的方式得到

$$H_{\text{free}}^{\frac{1}{2}} = \int d^3x (-i\bar{\Psi}\gamma_i\partial^i\Psi - m\bar{\Psi}\Psi) \quad (9.40)$$

或者用 Fourier 系数来表示

$$\mathcal{H}_{\text{free}}^{\frac{1}{2}} = \sum_r \int d^3p \omega_p \left(c_r^\dagger c_r(p) + d_r^\dagger(p) d_r(p) + \text{const} \right) \quad (9.41)$$

这里常数项导致了一项无穷大的项，同样地，我们忽略它。

¹⁵ 下式类似于标量场的(9.9)式，但是现在用反对易子替换了对易子 $[\cdot] \rightarrow \{\cdot\}$

¹⁶ 对于自旋为 0 的场，对应的方程不是很令人吃惊，因为这里我们用了反对易子 $[c^\dagger(k), c^\dagger(k)] = c^\dagger(k)c^\dagger(k) - c^\dagger(k)c^\dagger(k) = 0$

9.4 Free Spin 1 Theory 自旋 1 的自由场理论

6.4节导出了自旋 1 自由场的运动方程——Proca 方程：

$$m^2 A^\rho = \frac{1}{2} \partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) \quad (9.42)$$

它的解和自旋 0 场的解是类似的，形式为

$$A_\mu = \int \frac{d^3k}{\sqrt{(2\pi)^3 2\omega_k}} \left(\epsilon_{r,\mu}(k) a_r(k) e^{-ikx} + \epsilon_{r,\mu}(k) a_r^\dagger(k) e^{ikx} \right) \quad (9.43)$$

$$A_\mu = A_\mu^+ + A_\mu^- \quad (9.44)$$

这里 $\epsilon_{r,\mu}(k)$ 是基矢量，叫做偏振矢量。对于自旋 1 的场，可以使用对易子而非反对易子，因此系数 a_r , a_r^\dagger 和自旋 0 的场的系数具有相同的性质。

9.5 Interacting Field Theory 相互作用场论

下一步考虑不同自旋的场之间的相互作用。相应拉格朗日量已经在先前章节中，通过 Lorentz 对称性与规范对称性导出。举例来说，有质量自旋 $\frac{1}{2}$ 场与无质量自旋 1 场之间的相互作用，可用以下拉格朗日量描述（方程(7.17)）

$$\mathcal{L}_{\text{Dirac+Extra-Term}} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_\mu\partial^\mu\Psi + gA_\mu\bar{\Psi}\gamma^\mu\Psi \quad (9.45)$$

¹⁷ 很快会用到这一哈密顿量！

由此可导出哈密顿量¹⁷

$$\begin{aligned} H &= \int d^3x T^{00} \\ &= \int d^3x \left(\underbrace{\frac{\partial \mathcal{L}}{\partial(\partial_0\Psi)}}_{i\bar{\Psi}\gamma_0} \partial_0\Psi - \mathcal{L} \right) \\ &= \int d^3x \left(i\bar{\Psi}\gamma_0\partial_0\Psi + m\bar{\Psi}\Psi - i\bar{\Psi} \underbrace{\gamma_\mu\partial^\mu}_{\gamma_0\partial_0 - \gamma_i\partial_i} \Psi - gA_\mu\bar{\Psi}\gamma^\mu\Psi \right) \\ &= \underbrace{\int d^3x (m\bar{\Psi}\Psi + i\bar{\Psi}\gamma_i\partial^i\Psi)}_{=H_{\text{free}}^{\frac{1}{2}}} - \underbrace{\int d^3x (gA_\mu\bar{\Psi}\gamma^\mu\Psi)}_{\equiv -H_I} \\ &= H_{\text{free}}^{\frac{1}{2}} + H_I \end{aligned} \quad (9.46)$$

9.5.1 Scatter Amplitudes 散射振幅

计算动量为 p_1, p_2, \dots, p_n 的 n 个粒子¹⁸转变为动量为 $q_1, q_2, \dots, q_{n'}$ 的数量、动量都可能不同的 n' 个粒子的概率，是量子场论的基本目标之一。概率幅用 Dirac 符号写为：

$$\langle q_1, q_2, \dots, q_{n'} | \hat{S} | p_1, p_2, \dots, p_n \rangle \quad (9.47)$$

其中 \hat{S} 为描述散射过程的算符。下一节将导出这一算符的具体形式。在上式描述的过程中，初始时刻 t_i 时有数个粒子存在于空间点 x_i ，而在相互作用后的 t_f 时刻，（可能不同的）新粒子则存在于新位置 x_f 。

于是问题来了：量子场论如何描述量子态（记作 $|\psi\rangle$ ）的演化？

9.5.2 Time Evolution of States 态的时间演化

为了解答这个问题，注意到能量算符一方面由时间平移算符 $i\partial_0$ 给出，另一方面也是系统的哈密顿量。举例来说，自旋 0 场有哈密顿量

$$H_{\text{free}}^0 = \frac{1}{2} \int d^3x \left((\partial_0 \Phi)^2 + (\partial_i \Phi)^2 + m^2 \Phi^2 \right) \quad (9.48)$$

而自旋 $\frac{1}{2}$ 场有

$$H_{\text{free}}^{\frac{1}{2}} = \int d^3x (m \bar{\Psi} \Psi + i \bar{\Psi} \gamma_i \partial^i \Psi) \quad (9.49)$$

考虑到“能量算符”在场论中的两种表示，可以写出

$$i\partial_0 |\psi(t)\rangle = H |\psi(t)\rangle \quad (9.50)$$

这也是量子场论中态的演化方程。从这一方程出发，可以定义相应的时间演化算符 $\mathcal{U}(t)$ ，其效果是将某一时刻的量子态变换到另一时刻。为了简便，我们可取初始时刻 $t = 0$ ，于是有¹⁹

$$|\psi(t)\rangle = \mathcal{U}(t) |\psi(0)\rangle \quad (9.51)$$

代入方程(9.50)有

$$i\partial_0 \mathcal{U}(t) |\psi(0)\rangle = H \mathcal{U}(t) |\psi(0)\rangle \quad (9.52)$$

¹⁸ 这便是物理学家在对撞实验中所制备的。

¹⁹ 一般来说 $|\psi(t)\rangle = \mathcal{U}(t = t') |\psi(t')\rangle$

²⁰ 我们在5.1节的边注中讨论过这一思路。

这一方程对任意 $|\psi(0)\rangle$ 成立，因此²⁰

$$i\partial_0\mathcal{U}(t) = H\mathcal{U}(t) \quad (9.53)$$

可以找到通解

$$\mathcal{U}(t) = e^{-i\int_0^t dx_0 H} \quad (9.54)$$

代入即可验证：

$$\begin{aligned} i\partial_0\mathcal{U}(t) &= H\mathcal{U}(t) \rightarrow i\partial_0 e^{-i\int_0^t dx_0 H} = H e^{-i\int_0^t dx_0 H} \\ &\rightarrow i(-iH)e^{-i\int_0^t dx_0 H} = H e^{-i\int_0^t dx_0 H} \\ &\rightarrow H e^{-i\int_0^t dx_0 H} = H e^{-i\int_0^t dx_0 H} \quad \checkmark \end{aligned} \quad (9.55)$$

实验无法测量右矢 $|\psi(t)\rangle$ ，而只能测左矢与右矢的乘积，一般来说，也就是下述形式的对象：

$$\langle f(t)|\hat{O}|i(t)\rangle, \quad (9.56)$$

其中 \hat{O} 是某个算符， $|i(t)\rangle$ 是初态， $\langle f(t)|$ 是末态。此处量子态在 \mathcal{U} 的作用下演化，而算符则不含时。这一稍有些任意的选择被称为 **Schrödinger 绘景**，下面将讨论其它不同的绘景。

²¹ 记得 $\langle f| = |f\rangle^\dagger$

将上式用时间演化算符表出，得到²¹

$$\langle f(0)|\mathcal{U}^\dagger(t)\hat{O}\mathcal{U}(t)|i(0)\rangle \quad (9.57)$$

这里有一个重要的思想：我们可以改变观点，认为算符 \hat{O} 按照 $\mathcal{U}^\dagger(t)\hat{O}\mathcal{U}(t)$ 的形式随时间演化，而左右矢都不含时。这种新观点被称为 **Heisenberg 绘景**。

下面讲一个在相互作用理论中非常有用的传统技巧。哈密顿量总可以写成自由哈密顿量和相互作用哈密顿量的和（方程(9.46)）

$$H = H_{\text{free}} + H_I \quad (9.58)$$

这里的技巧便是混用上述两种观点：让量子态随 H_I 演化，而让算符随 H_{free} 演化²²。这个技巧非常有用，因为在这一观点下自由场论中的所有结果都可以加以利用。这种观点称为**相互作用绘景**。

²² 注意在量子场论中场即是算符

在相互作用绘景中，量子态是

$$|i(t)\rangle_I \equiv \mathcal{U}_{\text{free}}^\dagger |i(t)\rangle_S \quad (9.59)$$

其中 $\mathcal{U}_{\text{free}} = e^{-i \int_0^t dx_0 H_{\text{free}}}$ ，指标 S 表示 **Schrödinger** 绘景，也就是量子态随总哈密顿量演化，且算符不含时的标准观点。

将方程(9.59)与(9.50)合并可得²³相互作用绘景下的演化方程。

$$\begin{aligned} i\partial_0 |i(t)\rangle_S &= H |i(t)\rangle_S \\ \rightarrow i\partial_0 \mathcal{U}_{\text{free}} |i(t)\rangle_I &= H \mathcal{U}_{\text{free}} |i(t)\rangle_I \\ \rightarrow i\partial_0 e^{-i \int_0^t dx_0 H_{\text{free}}} |i(t)\rangle_I &= (H_{\text{free}} + H_I) e^{-i \int_0^t dx_0 H_{\text{free}}} |i(t)\rangle_I \\ \underbrace{\rightarrow H_{\text{free}} e^{-i \int_0^t dx_0 H_{\text{free}}} |i(t)\rangle_I}_{\text{莱布尼兹律}} + i e^{-i \int_0^t dx_0 H_{\text{free}}} \partial_0 |i(t)\rangle_I & \\ &= (\cancel{H_{\text{free}}} + H_I) e^{-i \int_0^t dx_0 H_{\text{free}}} |i(t)\rangle_I \\ \rightarrow i e^{-i \int_0^t dx_0 H_{\text{free}}} \partial_0 |i(t)\rangle_I &= H_I e^{-i \int_0^t dx_0 H_{\text{free}}} |i(t)\rangle_I \\ \rightarrow i\partial_0 |i(t)\rangle_I &= \underbrace{e^{i \int_0^t dx_0 H_{\text{free}}} H_I e^{-i \int_0^t dx_0 H_{\text{free}}}}_{= H_I^{\text{int}}, \text{相互作用绘景下的相互作用哈密顿量}} |i(t)\rangle_I \\ \rightarrow i\partial_0 |i(t)\rangle_I &= H_I^{\text{int}} |i(t)\rangle_I \end{aligned} \quad (9.60)$$

总之，此时量子态的时间演化实际上由相互作用哈密顿量 H_I^{int} 决定。这一简短的方程在后续讨论中价值连城。

现在回到最初的问题：怎么计算散射过程

$$\langle f(t_f) | \hat{S}(t_f, t_i) | i(t_i) \rangle \quad ? \quad (9.61)$$

算符 $\hat{S}(t_f, t_i)$ 将 t_i 时刻的初态 $|i(t_i)\rangle$ 变换为 t_f 时刻的末态 $|\Psi(t_f)\rangle$ 。一般来说，末态不是一个确定的粒子态，而是多种可能结果的叠加。反过来说，若非如此则粒子物理学未免太过无聊：相互作用会永远导致同样的结果。好在现实中的情况更加多变，任意线性叠加都可能存在。这一点在详细讨论算符 \hat{S} 后就能明了。

具体来说， \hat{S} 将一个初态变换为一系列末态的线性叠加。为简洁起见，下面用 t 来表示 t_f ，有²⁴

$$\hat{S}(t_f, t_i) |i(t_i)\rangle = |\Psi(t)\rangle = \sum_f \underbrace{S_{fi}}_{\text{复数}} |f(t)\rangle \quad (9.62)$$

²³ 这一形式的方程在 Schrödinger 绘景下成立。因此需由 $|i(t)\rangle_I \equiv \mathcal{U}_{\text{free}}^\dagger |i(t)\rangle_S$ 求解 $|i(t)\rangle_S$ 。在方程两边乘 $\mathcal{U}_{\text{free}}$ ，注意 $\mathcal{U}_{\text{free}} \mathcal{U}_{\text{free}}^\dagger = 1$ 。由此可得 $|i(t)\rangle_S = \mathcal{U}_{\text{free}} |i(t)\rangle_I$ 。这一结果随后可代入方程(9.50)。

²⁴ 这可以看成量子态 $|\Psi(t)\rangle$ 对粒子态 $|f(t)\rangle$ 的级数展开。

左乘 $\langle f'(t)|$, 只有一项非 0

$$\langle f(t)| \sum_f S_{fi} |f(t)\rangle = \sum_f S_{fi} \underbrace{\langle f'(t)| |f(t)\rangle}_{\delta_{ff'}} = \sum_f S_{fi} \delta_{ff'} = S_{f'i} \quad (9.63)$$

于是这一过程发生的概率为 $|S_{f'i}|^2$ 。

现在详述散射算符 \hat{S} 。出于这一目的, 请回顾上面导出的时间演化方程²⁵

$$\rightarrow i\partial_t |\Psi(t)\rangle_I = H_I^{\text{int}} |\Psi(t)\rangle_I \quad (9.64)$$

利用方程(9.62), 上式可用初态和算符 \hat{S} 重写为

$$\begin{aligned} \rightarrow i\partial_t S(\hat{t}, t_i) |i(t_i)\rangle_I &= H_I \hat{S} |i(t_i)\rangle_I \\ \underbrace{\rightarrow i(\partial_t S(\hat{t}, t_i)) |i(t_i)\rangle_I}_{\text{莱布尼兹律}} + iS(\hat{t}, t_i) \underbrace{\partial_t |i(t_i)\rangle_I}_{=0, \text{因为}|i(t_i)\rangle_I \text{不含} t} &= H_I \hat{S} |i(t_i)\rangle_I \end{aligned}$$

这对任意初态适用, 于是

$$i\partial_t \hat{S}(t, t_i) = H_I \hat{S} \quad (9.65)$$

有通解²⁶

$$\hat{S}(t, t_i) = e^{-i \int_{t_i}^t dt' H_I} \quad (9.66)$$

乍一看问题很简单。已知 H_I 求上述积分, 就得到了算符 $\hat{S}(t, t_i)$, 再将其作用到初态上就得到了不同过程的散射振幅。可惜这一积分却无法求解。为在数学上简化问题的形式, 初始时刻可取为 $t_i = -\infty$, 末尾时刻可取为 $t = \infty$ 。这样就避免了散射振幅的时间依赖性。比如说, 在粒子散射实验中, 特定过程发生的概率在“碰撞”发生后 10^{-24} 秒时, 与“碰撞”后 $2 \cdot 10^{-24}$ 秒时就不尽相同, 因为相互作用还在持续! 将时间尺度取得很大可以帮助避免这类问题, 因为这样一来, 测量结果就是相互作用完全结束后的数据。

9.5.3 Dyson Series Dyson 级数

由于散射振幅的指数积分无法解析求解, 我们把它展开成 Taylor 级数²⁷

²⁵ 即方程(9.60)。为避免符号混乱我们省略“int”, 即相互作用绘景的脚标。

²⁶ 此处我们忽略了一个重要因素: 编时性。这将在下一节讨论。

²⁷ 这在附录 B.4.1 中导出: $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$

$$\begin{aligned}
\hat{S}(-\infty, \infty) &= e^{-i \int_{-\infty}^{\infty} dt' H_I(t')} \\
&= 1 - i \int_{-\infty}^{\infty} dt_1 H_I(t_1) \\
&\quad - \frac{1}{2!} \left(\int_{-\infty}^{\infty} dt_1 H_I(t_1) \right) \left(\int_{-\infty}^{\infty} dt_2 H_I(t_2) \right) + \dots
\end{aligned} \tag{9.67}$$

这被称为 **Dyson 级数**。注意第三项， $H_I(t_1)$ 与 $H_I(t_2)$ 不是数值而是作用于右矢的算符。所以应保证时间靠前的算符先作用于右矢，之后再由时间靠后的算符作用。即算符应按照时间顺序依次作用于右矢。反过来说，先将 $H_I(t = 5s)$ ，再将 $H_I(t = 2s)$ 作用于量子态是没有意义的。在上述级数中，若 $t_1 < t_2$

$$\left(\int_{-\infty}^{\infty} dt_2 H_I(t_2) \right) \left(\int_{-\infty}^{\infty} dt_1 H_I(t_1) \right)$$

而当 $t_2 < t_1$ 时则有

$$\left(\int_{-\infty}^{\infty} dt_1 H_I(t_1) \right) \left(\int_{-\infty}^{\infty} dt_2 H_I(t_2) \right)$$

为此可定义抽象化的编时算符 T ：

$$T\{A(x)B(y)\} := \begin{cases} A(t_1)B(t_2) & \text{若 } t_1 > t_2 \\ B(t_2)A(t_1) & \text{若 } t_1 < t_2 \end{cases} \tag{9.68}$$

于是级数展开变成

$$\begin{aligned}
\hat{S}(\infty, -\infty) &= \underbrace{1}_{S^{(0)}} - i \underbrace{\int_{-\infty}^{\infty} dt_1 H_I(t_1)}_{S^{(1)}} \\
&= \underbrace{-\frac{1}{2!} T \left\{ \left(\int_{-\infty}^{\infty} dt_1 H_I(t_1) \right) \left(\int_{-\infty}^{\infty} dt_2 H_I(t_2) \right) \right\}}_{S^{(2)}} + \dots
\end{aligned} \tag{9.69}$$

或者写为求和形式⁵

$$\hat{S}(\infty, -\infty) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \cdots \int_{-\infty}^{\infty} dt_n T H_I(t_1) H_I(t_2) \cdots H_I(t_n)$$

⁵原文下式有误，译文已修改。

$$= \sum_{n=0}^{\infty} S^{(n)} \quad (9.70)$$

²⁸ 见方程(9.46)，其中耦合常数记作 e 。

这一级数可用于实际计算，因为 H_I 包含一个数值因子²⁸：对应相互作用的耦合常数，即 $H_I \propto g$ 。如在电磁相互作用中，这一耦合常数小于一。于是级数展开中第二项 $S^{(2)} \propto (H_I)^2 \propto g^2$ ，贡献小于第一项 $S^{(1)} \propto g$ 。更高阶项贡献更小。为了描述考察的系统，计算级数展开的前几项常常就足够了。更高阶项导致的修正往往超出了实验测量的范围。

然而沿这一思路深挖，需要进行繁冗的代数运算。第一步是建立 Wick 定理，将算符的编时乘积表示为算符对易子的组合。接下来需要计算这些对易子，得到著名的 Feynman 传播子。尽管如此，我们也确实希望继续讨论²⁹，所以我们略去证明、直接使用这些结果。感兴趣的读者可以参考量子场论的标准文献。³⁰

²⁹ 因为有趣的部分才刚要开始啊！

³⁰ 本章最后一节列出了推荐的参考书目。

9.5.4 Evaluating the Series 计算级数

回到本节开始的例子：有质量自旋 $\frac{1}{2}$ 场与无质量自旋 0 场的相互作用。相互作用哈密顿量为（方程(9.46)）

$$H_I = - \int d^3x g A_\mu \bar{\Psi} \gamma^\mu \Psi$$

如上所述，我们单独考察级数中的每一项。

级数的第一项是平凡项，直接等于恒等算符

$$S^{(0)} = I \quad (9.71)$$

第二项更令人激动

$$S^{(1)} = -i \int_{-\infty}^{\infty} dt H_I = ig \int_{-\infty}^{\infty} d^4x A_\mu \bar{\Psi} \gamma^\mu \Psi \quad (9.72)$$

这可以用方程(9.33)和(9.44)重写为

$$S^{(1)} = ig \int_{-\infty}^{\infty} d^4x (A_\mu^+ + A_\mu^-) (\bar{\Psi}^+ + \bar{\Psi}^-) \gamma^\mu (\Psi^+ + \Psi^-) \quad (9.73)$$

可以看出，第二项实际含有 8 项。考虑将其中一项 $S_1^{(1)}$ 作用于由

一个电子和一个正电子组成的量子态 $|e^+(p_1), e^-(p_2)\rangle$ 上

$$ig \int_{-\infty}^{\infty} d^4x A_{\mu}^+ \bar{\Psi}^+ \gamma^{\mu} \Psi^+ |e^+(p_1), e^-(p_2)\rangle$$

Ψ^+ 包括所有动量上³¹粒子湮灭³²算符与常数和 e^{-ipx_1} 因子的乘积。

$$\Psi^+ \propto \int d^3p c_r(p) e^{-ipx}$$

除 $c_r(p) = c_r(p_2)$ 一项外，上式中所有算符都摧毁这一右矢，即作用后得数为 0，因这些项所湮灭的粒子动量态并不存在于右矢之中。于是算符作用的结果是

$$\Psi^+ |e^+(p_1), e^-(p_2)\rangle \propto e^{-ip_2x} |e^+, 0\rangle$$

同样的，作用 $\bar{\Psi}^+$ 于右矢得到

$$\bar{\Psi}^+ e^{-ip_2x} |e^+(p_1), 0\rangle \propto e^{-ip_2x} e^{-ip_1x} |0, 0\rangle$$

剩下纯粹的真空态与数个常数的乘积。最后 A_{μ}^+ 作用于右矢，产生所有动量的光子。

定性来说，这一项对产生一个动量 k' 光子态 $|\gamma\rangle$ 的散射振幅有贡献

$$\begin{aligned} \langle f | S_1^{(1)} | i \rangle &= \langle \gamma_{k'} | S_1^{(1)} | e^+(p_1), e^-(p_2) \rangle \\ &= \int_{-\infty}^{\infty} d^4x \langle \gamma_{k'} | \sum_k \text{constant}(k) | \gamma_k \rangle e^{-ix(p_1+p_2-k)} \\ &= \int_{-\infty}^{\infty} d^4x \sum_k \text{constant}(k) \underbrace{\langle \gamma_{k'} | | \gamma_k \rangle}_{\delta_{kk'}} e^{-ix(p_1+p_2-k)} \\ &= \int_{-\infty}^{\infty} d^4x \text{constant}(k') e^{-ix(p_1+p_2-k)} \end{aligned} \quad (9.74)$$

对 x 积分导致 delta 函数 $\delta(p_1 + p_2 - k)$ ，它代表了动量守恒条件³³。实验中无法测量、制备确定动量的粒子，真实动量总处于一个范围内。因此在计算的最后，应将散射振幅对相应的动量范围积分。

注意到其它七项对 $\hat{S}^{(1)}$ 贡献为 0，因为他们湮灭初态中不存在的粒子，如一个光子。若选取其它初态，如一个光子和一个正电

³¹ 我们对所有可能动量积分！

³² 记住带有 \dagger 的算符是产生算符，没有的就是湮灭算符。 Ψ^+ 由方程(9.33)定义

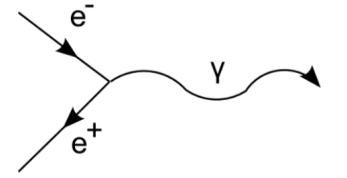


图 9.1: 过程 $e^+e^- \rightarrow \gamma$ 的 Feynman 图

³³ e^- 的 4-动量 ($p_0 =$ 能量 E , 空间部分为通常的动量 p_i) 与 e^+ 的 4-动量之和，必须等于光子 γ 4-动量: $p_1 + p_2 - k \stackrel{!}{=} 0$ 。否则由附录 D.2 中的解释，有 $\delta(p_1 + p_2 - k) = 0$ 。

子 $|\gamma, e^+\rangle$, 那么第一项贡献将为 0, 其余项则非 0。

下面我们定性考察 \hat{S} 级数展开的第三项。

$$\begin{aligned} S^{(2)} &= -\frac{1}{2!} T \left\{ \left(\int_{-\infty}^{\infty} d^4 x_1 \mathcal{H}_I(x_1) \right) \left(\int_{-\infty}^{\infty} d^4 x_2 \mathcal{H}_I(x_2) \right) \right\} \\ &= -\frac{1}{2!} T \left\{ \left(\int_{-\infty}^{\infty} d^4 x_1 g A_\mu(x_1) \bar{\Psi}(x_1) \gamma^\mu \Psi(x_1) \right) \right. \\ &\quad \left. \left(\int_{-\infty}^{\infty} d^4 x_2 g A_\mu(x_2) \bar{\Psi}(x_2) \gamma^\mu \Psi(x_2) \right) \right\} \end{aligned} \quad (9.75)$$

其中编时乘积可用 Wick 定理重写为对易子的正规序乘积之和, 记为 $N\{\}$ 。正规序意味着将所有产生算符放到左侧, 湮灭算符则放到右侧。比如, $N\{aa^\dagger a^\dagger a\} = a^\dagger a^\dagger aa$ 。举例来说, 求和中的一项是

$$-\frac{1}{2!} g^2 \int \int d^4 x_1 d^4 x_2 N \left\{ \bar{\Psi}(x_1) \gamma^\mu \Psi(x_1) [A_\mu(x_1), A_\mu(x_2)] \bar{\Psi}(x_2) \gamma^\mu \Psi(x_2) \right\}$$

其中对易子可由 A_μ 的显式解求出, 被称为光子传播子³⁴ $[A_\mu(x_1), A_\mu(x_2)] \equiv iD_\mu(x_1 - x_2)$ 。展开这一项后会得到许多项, 因为每一个 $\Psi, \bar{\Psi}$ 等算符都是两项之和, 现在只考虑其中一项。如果再次选择包含电子和正电子的初态, 可以定性得到其中一项³⁵的表达式, 如下所示:

$$-\frac{1}{2!} g^2 \int \int d^4 x_1 d^4 x_2 \bar{\Psi}^-(x_1) \Psi^-(x_1) D_\mu(x_1 - x_2) \bar{\Psi}^+(x_2) \Psi^+(x_2) |e^+, e^-\rangle \quad (9.76)$$

这可以从物理角度加以理解:

- 初态中的两个粒子在 x_2 处被 $\bar{\Psi}^+(x_2) \Psi^+(x_2)$ 湮灭。
- 之后传播子在 x_2 处产生一个“虚”光子, 虚光子传播到 x_1 湮灭。
- 最后 $\bar{\Psi}^-(x_1) \Psi^-(x_1)$ 再次于 x_1 处产生一个电子和一个正电子。

通过这一项可以计算出反应 $e^+e^- \rightarrow e^+e^-$ 的概率幅, 其中入射和出射粒子的动量当然可以完全不同, 但总动量保持守恒。所有其它项也同样可以诠释为有质量³⁶自旋 $\frac{1}{2}$ 场和无质量³⁷自旋 0 场。

计算中得到的这类概率幅可以直接在实验中加以检验, 因为概率幅直接联系于实验可测量量: 散射截面。

本章描述的技术可以用于推导量子场论中的许多重要结果。之前导出的其它相互作用项也可被加入相互作用哈密顿量, 而后对应的概率幅也可类似求出。不行的是, 针对 $SU(3)$ 规范对称性导

³⁴ 传播子是量子场论中最难推导的对象之一。比如考虑标量传播子, 推导的出发点是 $i\Delta = \langle 0|T\{\Phi(x)\Phi^\dagger(y)\}|0\rangle$, 即从真空中创造一个 y 处的粒子, 再于 x 处令其湮灭。这可以重写为对易子 $i\Delta = \langle 0|[\Phi^{\dagger+}(y), \Phi^-(x)]|0\rangle$ 。通过繁杂的数学计算可以得到 $i\Delta = \frac{-i}{2\pi^3} \int \frac{d^3 k}{\omega_k} e^{-ik(x-y)}$, 这也是实际计算中使用的表达式。这些粗略描述有多处不完全正确, 但它表达了这些数学构造的基本思想。

³⁵ 我们又选择了电子和正电子碰撞时非零的一项

³⁶ 此处是电子和正电子。其它可能性也包括夸克和另两种轻子 μ 和 τ 。

³⁷ 此处为光子

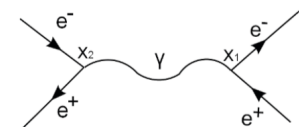


图 9.2: 过程 $e^+e^- \rightarrow \gamma \rightarrow e^+e^-$ 的 Feynman 图

致的相互作用项, 这一思路并不总是有效, 因为强相互作用的耦合常数过大, 以致无法用级数展开加以计算。若耦合常数大于 1, 高阶项大于低阶项, 便不能用级数的前几项近似级数和。这一物理学分支被称为量子色动力学, 在这门学科中需要用到其他计算方案, 超出了本书讨论的范围。

Further Reading Tips 进一步的阅读建议

- **Robert D. Klauber - 简明量子场论 (Student Friendly Quantum Field Theory)**³⁸在我看来是最好的量子场论导论教材。所有章节都极富教学技巧, 因为作者曾花费许多时间考虑读者学习量子场论时的难点。
- **Francis Halzen, Alan D. Martin - 夸克与轻子: 现代粒子物理学导论 (Quarks and Leptons: An Introductory Course in Modern Particle Physics)**³⁹是一本聚焦于应用量子场论计算方案的极佳之作。
- **Anthony Zee - 果壳中的量子场论 (Quantum Field Theory in a Nutshell)**⁴⁰包含许多独具匠心的论述。不过有些章节太短, 初学者很难理解。对用其它书籍学过量子场论的读者, 这本书十分值得一读。
- **Franz Mandl, Graham Shaw - 量子场论 (Quantum Field Theory)**⁴¹是学习弱、强相互作用理论的绝佳起点。
- **Michele Maggiore - 量子场论现代导引 (A Modern Introduction to Quantum Field Theory)**⁴²是一本极好的导论书籍, 着重讨论了场论中的群论概念。
- **Ian J. R. Aitchison and Anthony J. G. Hey - 粒子物理学中的规范场论 (Gauge Theories in Particle Physics)**⁴³非常清晰地诠释了规范场论的思想。

³⁸Robert D. Klauber. Student Friendly Quantum Field Theory. Sandtrove Press, 2nd edition, 12 2013. ISBN 9780984513956

³⁹Francis Halzen and Alan D. Martin. Quarks and Leptons: An Introductory Course in Modern Particle Physics. Wiley, 1st edition, 1 1984. ISBN 9780471887416

⁴⁰Anthony Zee. Quantum Field Theory in a Nutshell. Princeton University Press, 1st edition, 3 2003. ISBN 9780691010199

⁴¹Franz Mandl and Graham Shaw. Quantum Field Theory. Wiley, 2nd edition, 5 2010. ISBN 9780471496847

⁴²Michele Maggiore. A Modern Introduction to Quantum Field Theory. Oxford University Press, 1st edition, 2 2005. ISBN 9780198520740

⁴³Ian J.R. Aitchison and Anthony J.G. Hey. Gauge Theories in Particle Physics. CRC Press, 4th edition, 12 2012. ISBN 9781466513174

9.6 Appendix: Most General Solution of the Klein-Gordon Equation Klein-Gordon 方程的一般解

不难得到 Klein-Gordon 方程的一个解: 平面波即满足方程。

$$\Phi(x) = ae^{i(px-Et)} = e^{-ip_\mu x^\mu}$$

由狭义相对论能量-动量关系(8.2)式可得：

$$E^2 = \vec{p}^2 + m^2 \rightarrow p_\mu p^\mu = m^2$$

于是有

$$\begin{aligned} (\partial_\mu \partial^\mu + m^2)\Phi &= (\partial_\mu \partial^\mu + m^2)e^{-ip_\mu x^\mu} \\ &= (-p_\mu p^\mu + m^2)e^{-ip_\mu x^\mu} \\ &= (-m^2 + m^2)e^{-ip_\mu x^\mu} \\ &= 0 \quad \checkmark \end{aligned} \quad (9.77)$$

由于推导中对场求导两次，指数中的符号不会影响最终结果。因而，

$$\Phi^\dagger(x) = a^\dagger e^{-i(px-Et)} = a^\dagger e^{ip_\mu x^\mu}$$

也是方程的解。将以上两解线性叠加，可以得到更多解。一般解由所有可能解得叠加得出，亦可写为解的 Fourier 展开⁴⁴

$$\Phi(x) = \int \frac{d^4 p}{(2\pi)^4} \left(a(p) e^{-i(p_\mu x^\mu)} + a^\dagger(p) e^{i(p_\mu x^\mu)} \right)$$

注意到 $a = a(p)$ ，因为不同 p 值可以对应不同的乘积因子，求和中每一项都是各自独立的解。

在文献中，波数 $k_i = \frac{p_i}{\hbar}$ 和频率 $k_0 = \omega = \frac{E}{\hbar}$ 常被用以代替能量和动量。由于我们使用单位制 $\hbar = 1$ ，只需将上面公式中的变量重新命名，即可得到标准教科书中的表达式。同时，缩写 $kx \equiv k_\mu x^\mu$ 也很常用。于是可以写出

$$\Phi(x) = \int \frac{d^4 k}{(2\pi)^4} \left(a(k) e^{-i(kx)} + a^\dagger(k) e^{i(kx)} \right)$$

注意到并非所有 Klein-Gordon 方程的解都适用于描述自然，因为需满足“质壳”条件 ("mass-shell" condition)

$$p_\mu p^\mu = k_\mu k^\mu = m^2 \rightarrow k_0^2 - k_i^2 \stackrel{!}{=} m^2 \rightarrow k^2 \stackrel{!}{=} m^2 \quad (9.78)$$

只有满足质壳条件的解才与相对论能量-动量关系相协调(方程(8.2))。通过插入 delta 分布排除非物理(不在壳里面)解后⁴⁵，这一条件

⁴⁴ 这便是 2π 因子的来源。同时，这也是解构成正交集的条件： $\int dk e^{ikx} e^{-ikx'} = \int dk e^{ik(x-x')} = 2\pi \delta(x-x')$ 。于是 2π 因子即为归一化常数。

⁴⁵ 解释见附录 D.2

可被纳入上面的方程中。

$$\Phi(x)_{\text{physical}} = \int \frac{dk^4}{(2\pi)^4} 2\pi \delta(k^2 - m^2) \left(a(k) e^{-i(kx)} + a^\dagger(k) e^{i(kx)} \right)$$

除此之外，只有正能解有物理意义，原因前面已经提到过，否则能量没有下限、系统无法处于稳定态。利用 Heavside 函数 $\theta(k_0)$ ($k_0 < 0$ 时为 0, $k_0 \geq 0$ 时为 1) 可将这一约束纳入方程。方程中的积分变为：

$$\Phi(x)_{\text{physical}} = \int \frac{1}{(2\pi)^3} \underbrace{dk^4 \delta(k^2 - m^2) \theta(k_0)}_{\text{测度}} \left(a(k) e^{-i(kx)} + a^\dagger(k) e^{i(kx)} \right)$$

等式中的测度可以重写为

$$\begin{aligned} dk^4 \delta(k^2 - m^2) \theta(k_0) &= dk^4 \delta(k_0^2 - \underbrace{\vec{k}^2 - m^2}_{\text{定义为 } -\omega_k^2}) \theta(k_0) \\ &= dk^4 \delta(k_0^2 - \omega_k^2) \theta(k_0) \\ &= dk^4 \delta((k_0 - \omega_k)(k_0 + \omega_k)) \theta(k_0) \\ &\underbrace{=} dk^4 \frac{1}{2k_0} \left(\delta(k_0 - \omega_k) + \delta(k_0 + \omega_k) \right) \theta(k_0) \\ \text{使用 } \delta(f(x)) &= \sum_i \frac{\delta(x - a_i)}{\left| \frac{df}{dx}(a_i) \right|}, \text{ 其中 } a_i \text{ 为函数 } f \text{ 的根, 即 } f(a_i) = 0 \\ &\underbrace{=} dk^4 \frac{1}{2k_0} \delta(k_0 - \omega_k) \\ \text{因为 } \delta(k_0 + \omega_k) &\text{ 的参数在 } k_0 \geq 0 \text{ 时永不为 } 0 \\ &= dk^3 dk_0 \frac{1}{2k_0} \delta(k_0 - \omega_k) \\ &\underbrace{=} dk^3 \frac{1}{2\omega_k} \quad (9.79) \\ &\text{对 } k_0 \text{ 积分} \end{aligned}$$

最终得到满足物理条件的 Klein-Gordon 方程一般解

$$\Phi(x) = \int dk^3 \frac{1}{(2\pi)^3 2\omega_k} \left(a(k) e^{-i(kx)} + a^\dagger(k) e^{i(kx)} \right) \quad (9.80)$$

10. Classical Mechanics 经典力学

本章探究量子力学与经典力学之间的联系，我们会看到动量算符期望值对时间的导数正是 Newton 第二定律（经典力学的基础之一）的形式。

力学量算符 \hat{O} 的期望值为 ((8.14)式)：

$$\langle \hat{O} \rangle = \int d^3x \Psi^* \hat{O} \Psi$$

势场 V 中单粒子的 Schrödinger 方程 ((8.23)式)：

$$\begin{aligned} (i\frac{d}{dt} + \frac{\nabla^2}{2m})\Psi - V\Psi &= 0 \\ \rightarrow i\frac{d}{dt}\Psi &= \underbrace{\left(-\frac{\nabla^2}{2m} + V\right)}_{=:H} \Psi \\ &\rightarrow \frac{d}{dt}\Psi = \frac{1}{i}H\Psi \\ \underbrace{\frac{d}{dt}\Psi^*}_{H^\dagger=H} &= -\frac{1}{i}\underbrace{\Psi^*}_{\Psi^\dagger=\Psi^*}H \end{aligned}$$

期望值对时间的导数为：

$$\frac{d}{dt}\langle \hat{O} \rangle = \int d^3x \left(\left(\frac{d}{dt}\Psi^* \right) \hat{O}\Psi + \Psi^* \left(\frac{d}{dt}\hat{O} \right) \Psi + \Psi^* \hat{O} \left(\frac{d}{dt}\Psi \right) \right)$$

对大部分算符都有 $\frac{d}{dt}\hat{O} = 0$, 例如动量 $\hat{O} = \hat{p} = -i\vec{\nabla} \neq \hat{O}(t)$. 我们还利用 Schrödinger 方程来重写上式中波函数及其共轭对时间的导数项:

$$\begin{aligned}
 \frac{d}{dt}\langle\hat{O}\rangle &= \int d^3x \left(\left(-\frac{1}{i}\Psi^*H \right) \hat{O}\Psi + \Psi^*\hat{O}\left(\frac{1}{i}H\Psi \right) \right) \\
 &= \frac{1}{i} \int d^3x ((-\Psi^*H)\hat{O}\Psi + \Psi^*\hat{O}(H\Psi)) \\
 &= \frac{1}{i} \int d^3x \Psi^*[\hat{O}, H]\Psi \\
 &= \frac{1}{i} \langle[\hat{O}, H]\rangle \quad (10.1)
 \end{aligned}$$

上式是著名的 **Ehrenfest 定理**. 比如, 考虑 $\hat{O} = \hat{p}$, $H = \frac{p^2}{2m} + V$:

$$\begin{aligned}
 \frac{d}{dt}\langle\hat{p}\rangle &= \frac{1}{i} \langle[\hat{p}, H]\rangle \\
 &= \frac{1}{i} \left\langle \left[\hat{p}, \frac{\hat{p}^2}{2m} + V \right] \right\rangle \\
 &= \frac{1}{i} \left\langle \underbrace{\left[\hat{p}, \frac{\hat{p}^2}{2m} \right]}_{=0} + [\hat{p}, V] \right\rangle \\
 &= \frac{1}{i} \langle[\hat{p}, V]\rangle \\
 &= \frac{1}{i} \int d^3x \Psi^*[\hat{p}, V]\Psi \\
 &= \frac{1}{i} \int d^3x \Psi^* \hat{p} V \Psi - \frac{1}{i} \int d^3x \Psi^* V \hat{p} \Psi \\
 &= \frac{1}{i} \int d^3x \Psi^* (-i\nabla) V \Psi - \frac{1}{i} \int d^3x \Psi^* V (-i\nabla) \Psi \\
 &\stackrel{\text{乘积法则}}{=} - \int d^3x \Psi^* (\nabla V) \Psi - \int d^3x \Psi^* V \nabla \Psi + \int d^3x \Psi^* V \nabla \Psi \\
 &= - \int d^3x \Psi^* (\nabla V) \Psi \\
 &= \langle -\nabla V \rangle = \langle F \rangle \quad (10.2)
 \end{aligned}$$

上式意味着动量期望对时间的导数等于负的势能梯度 (也就是力) 的期望值, 这正是 **Newton 第二定律**¹. 它可以计算宏观物体的运动轨迹. 历史上这个关于力的定律是从实验中总结的唯象定律, 作用于一个物体上的力在方程的右端线性相加. 定义唯象的物理量 **动量** $p_{\text{mak}} = mv$, Newton 第二定律可写为 $\frac{d}{dt}p_{\text{mak}} = \frac{d}{dt}mv$, 当物体的质量不变时等号右侧即为 $m\frac{d}{dt}v$. 速度就是坐标对时间的导数², 综上可得:

¹ 在4.5节由 Noether 定理解释守恒量的时候直接用过这个式子 (没推导), 这里正式把它推倒 (无误) 啦。

² 换句话说, 速度 $v = \frac{d}{dt}x(t) = \dot{x}(t)$ 是物体坐标的时间变化率, 同理, 加速度 $a = \frac{d}{dt}\frac{d}{dt}x(t) = \ddot{x}(t)$ 是速度的时间变化率。

$$m \frac{d^2}{dt^2} x = F_1 + F_2 + \dots \quad (10.3)$$

解这个微分方程就得到物体的运动轨迹 $x = x(t)$ 。下一章有一个关于经典力学这方面的例子。

10.1 Relativistic Mechanics 相对论力学

拉格朗日形式提供了有关经典力学相当不同的一种视角。描述一个粒子的运动需要一个方程。本书总是假设正确的方程是使某物取极小值而导出的。第4章中已经讲过它必须是 Lorentz 变换下的不变量，否则不同惯性系的运动规律会不同。

在狭义相对论中有一个现成的 Lorentz 不变量：时空间隔（在2.1节推导）

$$(ds)^2 = (cd\tau)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \quad (10.4)$$

其中 τ 是固有时（见2.2节）。 $(ds)^2$ 的平方根自然是不变量，由此可以构造一个用来取极小值，并且形式上最简单的量：

$$S = \int C d\tau, \text{ 其中 } C \text{ 是常数, } d\tau = \frac{1}{c} \sqrt{(cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2} \quad (10.5)$$

正确的常数应当是 $C = -mc^2$ ，即我们要使

$$S = -mc^2 \int d\tau \quad (10.6)$$

取极小值。

简明起见，只考虑一维情形：

$$\begin{aligned} d\tau &= \frac{1}{c} \sqrt{(cdt)^2 - (dx)^2} = \frac{1}{c} \sqrt{(cdt)^2 \left(1 - \frac{(dx)^2}{c^2(dt)^2}\right)} \\ &= \frac{1}{c} (cdt) \sqrt{1 - \frac{1}{c^2} \left(\frac{dx}{dt}\right)^2} \stackrel{\frac{dx}{dt} = \dot{x}, \text{ 即为粒子速度.}}{=} dt \sqrt{1 - \frac{\dot{x}^2}{c^2}}. \end{aligned} \quad (10.7)$$

带入(10.6)式可得：

$$S = \int \underbrace{-mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}}}_{\equiv \mathcal{L}} dt \quad (10.8)$$

像以前那样, 通过将 \mathcal{L} 带入欧拉-拉格朗日方程(4.7)式就可得到使 S 取极小值的方程:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) &= 0 \\ \rightarrow \underbrace{\frac{\partial}{\partial x} \left(-mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} \right)}_{=0} - \frac{d}{dt} \left(\frac{\partial}{\partial \dot{x}} \left(-mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} \right) \right) &= 0 \\ \rightarrow c^2 \frac{d}{dt} \left(\frac{-m \frac{\dot{x}}{c^2}}{\sqrt{1 - \frac{\dot{x}^2}{c^2}}} \right) &= 0 \rightarrow \frac{d}{dt} \left(\frac{m \dot{x}}{\sqrt{1 - \frac{\dot{x}^2}{c^2}}} \right) = 0 \end{aligned} \quad (10.9)$$

这正是相对论下的自由粒子的运动方程。若粒子在外势场 $V(x)$ 中运动, 则在拉格朗日量中附上势能项就可以了³:

$$\mathcal{L} = -mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} - V(x) \quad (10.10)$$

这个拉格朗日量带入欧拉-拉格朗日方程得到:

$$\frac{d}{dt} \left(\frac{m \dot{x}}{\sqrt{1 - \frac{\dot{x}^2}{c^2}}} \right) = -\frac{dV}{dx} \equiv F \quad (10.11)$$

注意在非相对论极限 $\dot{x} \ll c$ 之下 $\sqrt{1 - \frac{\dot{x}^2}{c^2}} \approx 1$, 于是上式化为 Newton 第二定律形式 ((10.3)) 式。

10.2 The Lagrangian of Non-Relativistic Mechanics 非相对论力学的拉格朗日量

上一节导出的拉格朗日量在非相对论极限下的行为非常值得讨论, 非相对论极限指粒子速度远小于光速的情形: $\dot{x} \ll c$. 由此可利用 Taylor 公式⁴:

$$-mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} = -mc^2 \left(1 - \frac{1}{2} \frac{\dot{x}^2}{c^2} + \dots \right). \quad (10.12)$$

在 $\dot{x} \ll c$ 的极限下可以忽略 \dot{x}/c 的高阶小量, 于是拉格朗日量化为:

$$\mathcal{L} = -mc^2 + \frac{1}{2} m \dot{x}^2 - V(x) \quad (10.13)$$

³ 实际上添加的是 $-V(x)$ 而非 $V(x)$, 这是为了保证待会儿根据 Noether 定理计算体系的能量时, 势能项为 $+V(x)$.

⁴ Taylor 公式的简介见附录 B.3.

前面说过，拉格朗日量中 $-mc^2$ 这样的常数对运动方程没有影响，因此非相对论力学的拉格朗日量为：

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x) \quad (10.14)$$

在外势场为零，即 $V(x) = 0$ 的情形下， $\mathcal{L} = \frac{1}{2}m\dot{x}^2$ 正是4.5.1节中由 Noether 定理导出守恒量所用的拉格朗日量。把(10.14)式的拉格朗日量带入欧拉-拉格朗日方程(4.7)式可得：

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) &= 0 \\ \rightarrow \frac{\partial}{\partial x} \left(\frac{1}{2}m\dot{x}^2 \right) - \frac{d}{dt} \left(\frac{\partial}{\partial \dot{x}} \left(\frac{1}{2}m\dot{x}^2 - V(x) \right) \right) &= 0 \\ \rightarrow -\frac{\partial}{\partial x} V(x) - \frac{d}{dt} (m\dot{x}) &= 0 \\ \rightarrow \frac{d}{dt} (m\dot{x}) &= -\frac{\partial}{\partial x} V(x) \quad (10.15) \end{aligned}$$

这正是本章开头导出的 Newton 第二定律(10.3)式。

11. Electrodynamics 电动力学

第7章已经导出了经典电动力学最重要的方程 — 非齐次 Maxwell 方程（组）(7.22)式：

$$\partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) = J_\rho \quad (11.1)$$

利用电磁张量¹可将上式写为更紧凑的形式：

$$\partial_\sigma F^{\sigma\rho} = J^\rho \quad (11.2)$$

7.1.6节导出了 J_ρ 为 Noether 流，即 $\partial_\rho J^\rho = 0$ 。这一守恒流即为宏观理论中的四维电流。电磁张量 $F^{\sigma\rho}$ 是反称张量 ($F^{\sigma\rho} = -F^{\rho\sigma}$)，这可以从定义式 $F^{\sigma\rho} \equiv \partial^\sigma A^\rho - \partial^\rho A^\sigma$ 直接看出，因此 $F^{\sigma\rho}$ 只有 6 个独立分量，其中三个是

$$F^{i0} = \partial^i A^0 - \partial^0 A^i \quad (11.3)$$

另外三个是

$$F^{ij} = \partial^i A^j - \partial^j A^i = (\delta_\ell^i \delta_m^j - \delta_m^i \delta_\ell^j) \partial^\ell A^m \underbrace{=}_{\delta_\ell^i \delta_m^j - \delta_m^i \delta_\ell^j = \epsilon^{ijk} \epsilon^{klm}, \text{证明作为习题}} \epsilon^{ijk} \epsilon^{klm} \partial_\ell A^m \quad (11.4)$$

¹ 电磁张量

$F^{\sigma\rho} \equiv (\partial^\sigma A^\rho - \partial^\rho A^\sigma)$.

这些独立分量通常记为：

$$\partial^i A^0 - \partial^0 A^i \equiv E^i \quad (11.5)$$

$$\epsilon^{ijk} \partial^j A^k \equiv -B^i \quad (11.6)$$

因此

$$F^{i0} = E^i \quad (11.7)$$

$$F^{ij} = \epsilon^{ijk} \epsilon^{klm} \partial^l A^m = -\epsilon^{ijk} B^k \quad (11.8)$$

²(11.2)式称为方程组是因为它代表了 $\rho = 0, 1, 2, 3$ 一共四个方程。

将非齐次 Maxwell 方程组²重新写为：

$$\partial_\sigma F^{\rho\sigma} = \partial_0 F^{\rho 0} - \partial_k F^{\rho k} = J^\rho \quad (11.9)$$

³ $\epsilon^{jk\ell} \partial_k B^\ell = (\nabla \times \vec{B})^j$ ，后者就是向量分析里面的旋度运算 (nabla 算子叉乘向量)。

对于三个空间分量 ($\rho \rightarrow i$) 有³：

$$\begin{aligned} \partial_0 F^{i0} - \partial_k F^{ik} &= \partial_0 E^i + \epsilon^{ik\ell} \partial_k B^\ell = J_i \\ \rightarrow \partial_t \vec{E} + \nabla \times \vec{B} &= \vec{J} \end{aligned} \quad (11.10)$$

对于时间分量 ($\rho = 0$ 分量)：

$$\begin{aligned} \partial_0 \underbrace{F^{00}}_{=0, \text{见 } F^{\rho\sigma} \text{ 定义式}} - \partial_k \underbrace{F^{0k}}_{\substack{F^{\mu\nu} = -F^{\nu\mu} \\ \text{(11.7) 式}}} &= \partial_k E^k = J^0 \\ \rightarrow \nabla \cdot \vec{E} &= J_0 \end{aligned} \quad (11.11)$$

这是非齐次 Maxwell 方程组在工程等领域中的常用形式。

⁴ 四维 Levi-Civita 符号 $\epsilon^{\mu\nu\rho\sigma}$ 的定义见附录 B.5.5。

11.1 The Homogeneous Maxwell Equations 齐次 Maxwell 方程组

定义对偶 (dual) 电磁张量 $\tilde{F}^{\mu\nu}$ 为四维 Levi-Civita 符号⁴与电磁张量的缩并：

$$\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \quad (11.12)$$

从 $F^{\mu\nu}$ 的定义可以导出, $F^{\mu\nu}$ 与任意全反称张量 (比如 $\epsilon^{\mu\nu\rho\sigma}$) 缩并后的导数为零, 即:

$$\partial_\mu \tilde{F}^{\mu\nu} = \partial_\mu \epsilon^{\mu\nu\rho\sigma} (\partial_\sigma A_\rho - \partial_\rho A_\sigma) = 0 \quad (11.13)$$

这一结论实际是如下命题的特殊情形: 两个对称指标与两个反对称指标缩并¹所得结果为零⁵。以(11.13)式括号内的第一项为例:

⁵ 详见附录B.5.4。

$$\begin{aligned} \epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho &= \frac{1}{2} (\epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho + \epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho) \\ &\stackrel{\text{重命名哑指标}}{=} \frac{1}{2} (\epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho + \epsilon^{\sigma\nu\rho\mu} \partial_\sigma \partial_\mu A_\rho) \\ &\stackrel{\epsilon^{\mu\nu\rho\sigma} = -\epsilon^{\sigma\nu\rho\mu}}{=} \frac{1}{2} (\epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho - \epsilon^{\mu\nu\rho\sigma} \partial_\mu \partial_\sigma A_\rho) = 0 \quad \checkmark \end{aligned} \quad (11.14)$$

同理可导出括号中的第二项为零。方程组⁶

⁶ 下式代表 $\nu = 0, 1, 2, 3$ 四个方程组成的方程组。

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (11.15)$$

称为齐次 Maxwell 方程组 (homogeneous Maxwell equations)。从推导过程可见它源自电磁张量 $F^{\mu\nu}$ 的定义。下面将它写成用电场 \vec{E} 与磁场 \vec{B} 表示的形式, 考虑 $\nu = 0$ 对应的方程:

$$\begin{aligned} 0 &= \partial_\mu \tilde{F}^{\mu 0} \\ &= \partial_\mu \epsilon^{\mu 0 \rho \sigma} F^{\rho \sigma} \\ &= \partial_0 \underbrace{\epsilon^{00 \rho \sigma}}_{=0} F^{\rho \sigma} - \partial_i \epsilon^{i 0 \rho \sigma} F^{\rho \sigma} \\ &\stackrel{\epsilon^{i 0 \rho \sigma}=0, \text{ 对于 } \rho=0 \text{ 或 } \sigma=0}{=} -\partial_i \epsilon^{i 0 j k} F^{j k} \\ &= \partial_i \epsilon^{0 i j k} F^{j k} \\ &\stackrel{(11.8) \text{ 式}}{=} -\partial_i \underbrace{\epsilon^{0 i j k} \epsilon^{\ell j k}}_{=2\delta^{i\ell}} B^\ell \\ &= -2\partial_i \delta_{i\ell} B^\ell = -2\partial_i B^i \\ &\Rightarrow \partial_i B^i = 0, \text{ 用向量记号表示: } \nabla \cdot \vec{B} = 0 \end{aligned} \quad (11.16)$$

¹ 例如, 张量 $A^{\alpha\beta\mu\nu} = A^{\alpha\beta\nu\mu}$, μ, ν 是一对对称指标; 张量 $B_{\rho\sigma\eta\delta} = -B_{\sigma\rho\eta\delta}$, ρ, σ 是一对反对称指标。一对对称指标与反对称指标的缩并结果必为零 (即所得张量分量均为 0): $C_{\eta\delta}^{\alpha\beta} \equiv A^{\alpha\beta\gamma\varphi} B_{\gamma\varphi\eta\delta} = 0$ 。

考虑空间分量: $\nu = i$ 对应的方程可以导出:

$$\nabla \times \vec{E} + \partial_t \vec{B} = 0 \quad (11.17)$$

这是齐次 Maxwell 方程组在应用领域中的常见形式。

11.2 The Lorentz Force Lorentz 力

利用上一章导出的量子力学与经典力学的联系 (Ehrenfest 定理) 可以推导出著名的 Lorentz 力公式。考虑在电磁场中运动的非相对论性且无自旋的粒子, 其 Schrödinger 方程包含电磁场的耦合项 ((8.24)式)⁷

$$i\partial_t \Psi = \underbrace{\left(\frac{1}{2m} (\vec{p} - q\vec{A})^2 + q\Phi \right)}_{\equiv H} \Psi \quad (11.18)$$

定义该系统的动量⁸为 $\vec{\Pi} = \vec{p} - q\vec{A}$, 则系统的哈密顿量为:

$$H = \frac{1}{2m} |\vec{\Pi}|^2 + q\Phi$$

有了哈密顿量, 再重复与上一章相同的过程, 就能得到类似(10.1)式的结果。因为这里涉及到依赖于时间的算符, 因而推导中包含对时间的偏导数项:

$$\begin{aligned} \frac{d}{dt} \langle \hat{O} \rangle &= \frac{1}{i} \langle [O, H] \rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle \\ \rightarrow \frac{d}{dt} \langle \vec{\Pi} \rangle &= \frac{1}{i} \langle [\vec{\Pi}, H] \rangle + \left\langle \frac{\partial \vec{\Pi}}{\partial t} \right\rangle, \end{aligned}$$

$\partial \vec{\Pi} / \partial t \neq 0$, 因为 \vec{A} 显含时间。将 H 的具体形式带入得:

$$\begin{aligned} \rightarrow \frac{d}{dt} \langle \vec{\Pi} \rangle &= \frac{1}{i} \langle [\vec{\Pi}, \frac{1}{2m} |\vec{\Pi}|^2 + q\Phi] \rangle + \left\langle \frac{\partial \vec{\Pi}}{\partial t} \right\rangle \\ \rightarrow \frac{d}{dt} \langle \vec{\Pi} \rangle &= \frac{1}{i} \langle [\vec{\Pi}, \frac{1}{2m} |\vec{\Pi}|^2] \rangle + \underbrace{\frac{1}{i} \langle [\vec{\Pi}, q\Phi] \rangle}_{=\langle q\nabla\Phi \rangle} + \left\langle \frac{\partial \vec{\Pi}}{\partial t} \right\rangle. \end{aligned}$$

最后一步中 $[\vec{\Pi}, q\Phi]$ 的推导可类比于上一章中 $[\hat{p}, V]$ 的推导 (因为 $[\vec{A}, \Phi] = 0$)。接下来计算 $[\vec{\Pi}^2, \vec{\Pi}]$, 它不等于零, 因为各分量 Π_i

⁷ 该式是从与外电磁场耦合的 Klein-Gordon 方程导出, 亦即, 从与自旋 1 无质量场 (电磁场) 耦合的自旋 0 的粒子的拉格朗日量导出的。因此推导的出发点还是用来导出拉格朗日量的 Lorentz 对称性与规范对称性。

⁸ 这个定义来自 Noether 定理: 系统的平移不变性对应守恒量-动量, $\frac{\partial L}{\partial \dot{x}} = \Pi$.

之间不对易：

$$[\Pi_i, \Pi_j] = -\frac{q}{i} \left(\frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \right) = -\frac{q}{i} \epsilon_{ijk} \underbrace{B_k}_{=\epsilon_{k\ell m} \frac{\partial}{\partial x_\ell} A_m} \quad (11.19)$$

上面利用了磁场定义 $\vec{B} = \nabla \times \vec{A}$ 的指标形式。粒子的速度⁹定义为 $\vec{v} \equiv \vec{\Pi}/m$ ，由此可得：

⁹ (Newton 极限下) 动量除以质量即为速度： $p = mv$ 。

$$\begin{aligned} \frac{d}{dt} \langle \vec{\Pi} \rangle &= \langle q \nabla \Phi \rangle - \langle q(\vec{v} \times \vec{B}) \rangle + \underbrace{\left\langle \frac{\partial \vec{\Pi}}{\partial t} \right\rangle}_{=-q \frac{\partial \vec{A}}{\partial t}} \\ \frac{d}{dt} \langle \vec{\Pi} \rangle &= -q \langle (\vec{v} \times \vec{B}) \rangle + q \underbrace{\left\langle \nabla \Phi - \frac{\partial \vec{A}}{\partial t} \right\rangle}_{=\langle E \rangle, \text{见(11.5)式}} \end{aligned}$$

最终得到

$$\frac{d}{dt} \langle \vec{\Pi} \rangle \equiv F_{\text{Lorentz}} = -q \langle (\vec{v} \times \vec{B}) + \langle E \rangle \rangle. \quad (11.20)$$

上式即为带电粒子在外电磁场中的运动方程，其解即为粒子轨迹。

11.3 Coulomb Potential Coulomb 势

第7章讲过拉格朗日量具有内禀对称性，可以利用这一自由度简化计算，也就是说将场进行内禀变换以使计算大大简化。只要进行规范变换（使拉格朗日量不变），则变换前后的场蕴含的物理规律是相同的。通常选择的规范是关于光子场 A^μ 的，规定¹⁰ $\partial_\mu A^\mu = 0$ ，这称为 Lorentz 规范。利用这一规范条件将非齐次 Maxwell 方程组简化为：

¹⁰ 可以证明这一规范条件的存在性（略），相关细节可参见电动力学教材。

$$\partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma) = \partial_\sigma \partial^\sigma A^\rho = J^\rho \quad (11.21)$$

考虑静电情况，静止的电荷关于坐标系原点对称分布。求解外部区域（无源，即 $J_\rho = 0$ ）的电磁场，外部区域的 Maxwell 方程组为²：

$$\partial_\sigma \partial^\sigma A^\rho = \partial_0 \partial^0 A^\rho - \partial_i \partial^i A^\rho = 0. \quad (11.22)$$

²原文下式等号右侧上下标有误，译文已更正。

考虑到系统是静态 ($\partial_0 A^\rho = 0$)、对称的, 因此将方程用球坐标表示¹¹。这样可以忽略除了包含 ∂_r 之外的所有项¹², 由此可得:

$$\rightarrow \frac{\partial^2}{\partial r^2}(rA^\mu) = 0. \quad (11.23)$$

上式具有通解:

$$A^\mu = \epsilon^\mu \frac{C}{r} + \epsilon^\mu D \quad (11.24)$$

其中 ϵ^μ 表示某个常数 4-向量, C, D 表示常数。场 A^μ 在无穷远处为零, 故 $D = 0$, A^μ 的 0 分量即为著名的 Coulomb 势:

$$A^0 = \Phi = \frac{C}{r} = \frac{Ze}{r}, \quad (11.25)$$

其中 Z 为整数, e 为元电荷 (电子电荷)。由于电子电荷是量子化的, 故常数 C 写为 Ze 的形式。标准模型不能对电荷量子化做出满意的解释。

Further Reading Tips 阅读建议

- **Richard P. Feynman - The Feynman Lectures on Physics Volume 2**¹³ 是不错的电动力学入门教材。
- **David J. Griffiths - Introduction to Electrodynamics**¹⁴ 是更详细的电动力学优秀入门教材。

¹¹ 除了用直角坐标 x, y, z 表示物体的位置外, 还可使用球坐标 r, θ, ϕ 表示。
 $\partial_i \partial^i A^\rho = \frac{\partial^2}{\partial r^2}(rA^\rho) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial A^\rho}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 A^\rho}{\partial \phi^2}$. 当系统具有球对称性时选择球坐标十分方便, 因为物理量只与 r (到原点的距离) 有关。

¹² 对于球对称性的场, $\partial_\theta A = \partial_\phi A = 0$ 。

¹³ Richard P. Feynman, Robert B. Leighton, and Matthew Sands. The Feynman Lectures on Physics: Volume 2. Addison-Wesley, 1st edition, 2 1977. ISBN 9780201021172

¹⁴ David J. Griffiths. Introduction to Electrodynamics. Addison-Wesley, 4th edition, 10 2012. ISBN 9780321856562

³ 译注: 此书新版可在此免费在线查看: http://www.feynmanlectures.ca/1tech.edu/II_toc.html。

另有新版中译本: 费恩曼, 莱顿, 桑兹著; 李洪芳, 王子辅, 钟万衡译. 费恩曼物理学讲义, 新千年版, 第三卷, 上海: 上海科学技术出版社, 4 2013, ISBN 9787547816370。

⁴ 译注: 国内有由机械工业出版社出版的影印本和中译本, 影印本 ISBN 为 9787506272896, 贾瑜等的中译本 ISBN 为 9787111444046。

12. Gravity 引力

非常不幸，目前最靠谱的引力理论与本书其他部分（对称性）的三观不合。这是现代物理学的最大问题之一，本章尝试使读者对这一疑难有所认识。

现代的引力理论是 **Einstein** 的广义相对论 (**general relativity**)。广相的基本思想是引力为时空弯曲的结果。质量/能量改变时空的曲率，变化了的曲率进而影响质量/能量的运动。能量与曲率的这一互动由著名的 Einstein 场方程描述：

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \quad (12.1)$$

等号左侧是描述曲率的 Einstein 张量 $G_{\mu\nu}$ ，右侧是能量-动量张量¹ $T_{\mu\nu}$ ， G 是引力常数。

在现代观点——引力 = 时空曲率的指导思想下，Einstein 场方程的形式比较容易导出²。怎么导出？首先，能量-动量守恒是最重要的物理定律之一，这可以从时空的均匀性直接推导（见第四章），在均匀时空中物理规律在时空的平移变换下不变。因此，将时空的均匀性作为基本假设的话，能量与动量自然就是守恒的，这一守恒定律的数学形式为 (4.36式)：

$$\partial^\mu T_{\mu\nu} = 0. \quad (12.2)$$

¹ 能量-动量张量是与平移对称性直接相关的量，见4.36式。

² 100年前，Einstein 经过大概6年导出了场方程的正确形式。100年后，在现代的较高观点之下，场方程的导出更加直接明快。

³ 目前为止用到的度规包括

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

和 Euclidean 度规

$$\delta^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

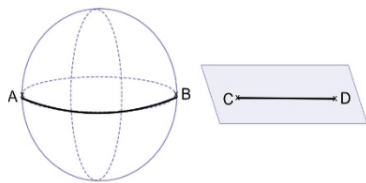


图 12.1: 弯曲空间与平直空间中的两点间距

⁴ 即 $\partial^\mu G_{\mu\nu} = 0$.

⁵ 二阶张量是指它有两个指标 $\mu\nu$, 这与散度为零一样也是限制之一, 因为 $T_{\mu\nu}$ 是二阶张量。

下面介绍描述关于曲率的数学工具, 它是广义相对论繁琐计算的根源。我们已经见过关于曲率的重要的量——度规, 它用于计算两点间距离³。由图12.1可见, 弯曲空间中的两点间距与平直时空的两点不同, 因而度规在曲率的数学描述中起到十分重要的作用。

现在就能“导出”Einstein 场方程了, 因为能放在能动张量对面的量只有一个: Einstein 张量 $G_{\mu\nu}$ 。这是因为 Einstein 张量是度规 $g_{\mu\nu}$ 及其一阶、二阶导数的函数中唯一¹ 的无散度⁴函数。不管其多复杂, 放到等号一边描述曲率的量只能是它, 因为

$$T_{\mu\nu} = CG_{\mu\nu}, \quad \text{由 } \partial^\mu T_{\mu\nu} = 0 \rightarrow \partial^\mu G_{\mu\nu} = 0. \quad (12.3)$$

Einstein 张量是二阶张量⁵, 正好符合要求。

Einstein 张量通过 Ricci 张量 $R_{\mu\nu}$ 、Ricci 标量 R (即 Ricci 张量的迹: $R \equiv R^\nu_\nu$) 和度规张量 $g_{\mu\nu}$ 定义:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}. \quad (12.4)$$

其中 Ricci 张量由 Christoffel 符号 $\Gamma^\mu_{\nu\rho}$ 定义:

$$R_{\alpha\beta} = \partial_\rho \Gamma^\rho_{\beta\alpha} - \partial_\beta \Gamma^\rho_{\rho\alpha} + \Gamma^\rho_{\rho\lambda} \Gamma^\lambda_{\beta\alpha} - \Gamma^\rho_{\beta\lambda} \Gamma^\lambda_{\rho\alpha} \quad (12.5)$$

其中 Christoffel 符号通过度规张量定义:

$$\Gamma^d_{ab} = \frac{1}{2}g^{cd} \left(\frac{\partial g_{ca}}{\partial x^b} + \frac{\partial g_{cb}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^c} \right) = \frac{1}{2}g^{cd} (\partial_b g_{ca} + \partial_a g_{cb} - \partial_c g_{ab}). \quad (12.6)$$

上面的式子复杂得吓人, 这稍微揭示了广义相对论计算的繁复程度。

现在考虑物体在弯曲时空中的行为。设自由物体经过弯曲时空中的 A, B 两点, 物体在 A, B 之间的轨迹是怎样的? 容易猜想物体沿弯曲时空中两点间的最短路径 (称为短程线) 运动, 确实如此 (证明略)。这样, 给定质能分布 $T_{\mu\nu}$ 之后就可根据 Einstein 场方程计算度规与 Christoffel 符号, 从而得到自由质点的轨迹——

¹译注: 至多差一常数与度规张量的乘积, 即著名的宇宙学常数 Λ

短程线方程 (geodesic equation):

$$\frac{d^2 x^\lambda}{dt^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} = 0. \quad (12.7)$$

测地线是流形上的两点间局域最短⁶的连线。

有趣的是, Einstein 将 Christoffel 符号视作引力场⁷:

若 $\Gamma_{\nu\rho}^\mu$ 消失, 则 (自由) 质点沿直线匀速运动, 因此 $\Gamma_{\nu\rho}^\mu$ 描述了关于匀速直线的偏离, 它们是引力场的分量。

– Albert Einstein⁸

从 12.7 式就能理解 Einstein 所说的, 令 $\Gamma_{\nu\rho}^\mu = 0$, 测地线方程退化为:

$$\frac{d^2 x^\lambda}{dt^2} = 0. \quad (12.8)$$

方程的解是一条直线。

弯曲时空的另一有趣事实是微分符号的变化。平直时空中的导数利用下式定义:

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}. \quad (12.9)$$

上式用到了函数 f 在两个不同点的值。弯曲时空的情形不像平直时空那样简单。如图 12.2, 如果想比较球面上不同位置的两个向量的差值, 怎样才能得到它们真正的差而除去弯曲空间造成的影响? 微分几何告诉我们可以进行平移, 将一个向量平移到另一个向量的位置, 从而可以⁹比较它们。

弯曲空间中的导数称为协变导数 (covariant derivative)¹⁰:

$$D_b v^a \equiv \partial_b v^a + \Gamma_{bc}^a v^c. \quad (12.10)$$

因此, 弯曲时空中的方程需要将普通导数替换为协变导数:

$$\partial_b \rightarrow D_b = \partial_b + \Gamma_{bc}^a. \quad (12.11)$$

这很眼熟。见 7.18 式, 我们在先前的章节中讲过, 自旋 0 或自旋 $\frac{1}{2}$ 场拉格朗日量的局域 $\mathcal{U}(1)$ 不变性要求与自旋 1 的场有特定的耦

⁶ 这一说法是过于简单的, 正确的概念需要微分几何的相关知识, 超出本书范围。

⁷ 一般将度规视为引力场。

⁸ Albert Einstein. The foundation of the general theory of relativity. 1916

⁹ 实际上, 微分几何中的坐标系都是局域的。3.11 节讲过, 流形的主要特征是局域平直 (像 Euclidean 空间)。因此, 流形上对象的坐标只在一个坐标域内有效, 同样只能在同一坐标域内比较不同对象的坐标差值。

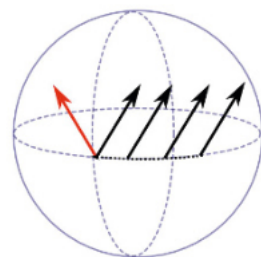


图 12.2: 为了比较红色箭头与黑色箭头, 我们将黑色箭头平移到红色箭头的位置。

¹⁰ 这里的 Christoffel 符号 Γ_{bc}^a 通常称为联络系数, 因为它们联系了进行比较的两点。

合。这一耦合可以用下式描述：

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ieA_\mu. \quad (12.12)$$

注意这不只是个数学技巧，它导出了正确的电磁学理论。弱相互作用的情形也类似：

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + ig \underbrace{W_\mu}_{=W_\mu^i \sigma^i}. \quad (12.13)$$

强相互作用也是：

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig' \underbrace{G_\mu}_{=T^C G_\mu^C}. \quad (12.14)$$

尽管上面各式的形式相似，但引力无法像其他相互作用那样量子化。其他相互作用都是用量子理论描述的，都给出概率性的实验预言。而广义相对论是经典理论，因为广相中的粒子沿确定的轨迹运动，没有概率性因素。

更糟糕的是目前没有任何实验能检测引力与其他力之间的相互作用。基本粒子间的引力太小，无法测量。因此忽视了引力，只考虑强、弱与电磁相互作用的标准模型就可以与实验符合得相当好。大质量物体的广义相对论效应实验上可以测量，但量子效应却十分微弱，因为大质量物体包含大量基本粒子，所有量子效应被平均掉了。第10章导出了平均值的运动方程，它正是经典物理的形式，毫无量子效应。

引力的量子化的困难可以列出长长的列表，不过 Einstein 简明地论述了引力与其他相互作用的差异：

... 根据广义相对论，引力与其他相互作用（特别是电磁力）相比地位特殊，因为描述引力场的十个方程²同时确定了时空的度量性质。

¹¹ Albert Einstein and Francis A. Davis. *The Principle of Relativity*. Dover Publications, reprint edition, 6 1952. ISBN 9780486600819

— Albert Einstein¹¹

将普通导数替换为协变导数就能描述弯曲时空中的量子粒子，但这并非引力的动力学理论。Einstein 场方程等号右侧（能动张

²或许指 Einstein 场方程，代表十个方程。——译者

量)容易量子化,选定合适的生成元就行了。困难的是等号左边的 Einstein 张量, 它用 Christoffel 符号写为:

$$G_{\alpha\beta} = (\delta_{\alpha}^{\gamma}\delta_{\beta}^{\zeta} - \frac{1}{2}g_{\alpha\beta}g^{\gamma\zeta})(\partial_{\epsilon}\Gamma_{\gamma\zeta}^{\epsilon} - \partial_{\zeta}\Gamma_{\gamma\epsilon}^{\epsilon} + \Gamma_{\epsilon\sigma}^{\epsilon}\Gamma_{\gamma\zeta}^{\sigma} - \Gamma_{\zeta\sigma}^{\epsilon}\Gamma_{\epsilon\gamma}^{\sigma}). \quad (12.15)$$

由此猜想 Einstein 场方程或许是 $\Gamma_{\zeta\sigma}^{\epsilon}$ 的场方程¹², 并且 $\partial_b \rightarrow D_b = \partial_b + \Gamma_{bc}^a$ 描述引力场 $\Gamma_{\zeta\sigma}^{\epsilon}$ 与其他场的耦合。

无论将度规张量(两个指标)还是 Christoffel 符号(三个指标)视为引力场, 描述它们都需要 Poincaré 群的 (1,1) 以及更高维的表示, 或者说自旋 2, 3, ... 表示。目前大多数物理学家认为引力场量子化后的引力子 (graviton) 是自旋为 2 的 boson。

目前为止, 没有靠谱的量子引力理论¹³能比下一节列出教科书上的标准引力理论给出进一步的信息。

¹² 就像 Maxwell 方程组与电磁场的关系那样。

¹³ 许多尝试过程产生了无穷多个无穷大项, 这对概率性预言来说十分要命。

Further Reading Tips 阅读建议

引力的标准理论——Einstein 的广义相对论的更多信息可以参考以下文献:

- **Ta-Pei Cheng - Relativity, Gravitation and Cosmology**¹⁴是一本内容丰富、起点低的广义相对论入门教材, 其中有大量启发性的论述。是快速入门的首选。
- **A. Zee - Einstein Gravity in a Nutshell**¹⁵是广义相对论的最佳教材。此书从基本概念开始, 避免不必要而复杂的数学概念, 精彩地解释了广义相对论的起源与应用。
- **Charles W. Misner, Kip S. Thorne, John Archibald Wheeler - Gravitation**¹⁶是一本关于引力的鸿篇巨著, 其中阐明了有大量其他教材未深入讨论的问题。

关于尝试引力量子化的更多信息可见

- **John C. Baez, Javier P. Muniain - Gauge Fields, Knots, and Gravity**¹⁷以面向物理学家的方式介绍了引力量子化的数学工具。

¹⁴Ta-Pei Cheng. *Relativity, Gravitation and Cosmology: A Basic Introduction*. Oxford University Press, 2nd edition, 1. 2010. ISBN 9780199573646

¹⁵Anthony Zee. *Einstein Gravity in a Nutshell*. Princeton University Press, 1st edition, 5. 2013. ISBN 9780691145587

¹⁶ Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. W.H. Freeman, 1st edition, 9. 1973. ISBN 9780716703440

¹⁷ John C. Baez and Javier P. Muniain. *Gauge Fields, Knots, and Gravity*. World Scientific Pub Co Inc, 1st edition, 9. 1994. ISBN 9789810220341

13. 结束语

在我看来，目前人类距离描述万物的万有理论还十分遥远。即使是本书主体部分所描述的优美理论也还存在着许多有待阐明的地方¹。此外，如上一章所述，我们在量子引力理论方面仍一筹莫展。

而且现在存在着表明现有理论仍有待改进的实验证据（主要是宇宙学与天体粒子物理学的暗物质与暗能量疑难）。

我个人觉得现有疑难的背后仍蕴藏着大量内容，甚至或许是物理学的崭新框架。无论如何，未来的进展都非常有趣，衷心祝愿你会沿着这条路走向更远。

¹比如本书多得要命的 typo，赶快出修订版吧！——译者

Appendix 附录

A	矢量代数	265
A.1	基矢	
A.2	坐标系变换	
A.3	矩阵乘法	
A.4	标量	
A.5	左手/右手坐标系	
B	微积分	273
B.1	Product Rule	
B.2	分部积分	
B.3	Taylor 级数	
B.4	级数	
B.5	指标记号	
C	线性代数	285
C.1	基本变换	
C.2	矩阵指数函数	
C.3	行列式	
C.4	本征值与本征向量	
C.5	对角化	
D	其他数学概念	289
D.1	Fourier 变换	
D.2	δ 分布函数	

A. Vector calculus 矢量代数

物理中常用三个数 $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ 来描述一个东西的位置。我们将其称作矢量 \vec{v} ，在它头上放了一个箭头。这三个数称作矢量沿三个坐标轴的分量。第一个数说了这个矢量沿 x 方向走了多远，第二个是沿 y 方向以及第三个是沿 z 方向。例如， $\vec{w} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$ 是一个指向 y 轴方向的矢量。

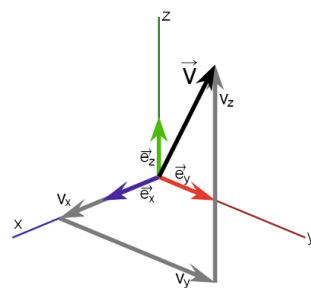
矢量间可以相加

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad \vec{w} = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \quad \rightarrow \quad \vec{v} + \vec{w} = \begin{pmatrix} v_x + w_x \\ v_y + w_y \\ v_z + w_z \end{pmatrix} \quad (\text{A.1})$$

也可以相乘

$$\vec{v} \cdot \vec{w} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \cdot \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} = v_x w_x + v_y w_y + v_z w_z. \quad (\text{A.2})$$

乘法的结果是一个数 (= 标量) 而不是矢量，其有特定的名



称：标量积 (scalar product)。计算矢量与自身的标量积可以得到其长度：

$$\text{长度}(\vec{v}) = \sqrt{\vec{v} \cdot \vec{v}}. \quad (\text{A.3})$$

注意并不是随便三个量都可以写在一起，用括号包起来，作为一个矢量。例如，把一个房间里的温度 T ，压强 P 和湿度 H 放在一起：

$$\begin{pmatrix} T \\ P \\ H \end{pmatrix}. \quad (\text{A.4})$$

没人阻止我们把他们放在一起，但这结果毫无意义，而且显然不是一个矢量，因为不存在能将它们混在一起的线性关系。而如果我们仅仅是换一个视角¹去考察一个矢量，它的坐标分量将会互相转换¹。故将这些坐标分量写在一起用括号包起来是有用的。另外，如果分量会因视角的改变而发生混合的话，将其写成位置矢量的形式是有用的。

从现在开始，我们将完全按照位置矢量 \vec{v} 的变换方式的量称为一个矢量。这句话的意思是，如果在某种变换下我们有 $\vec{v} \rightarrow \vec{v}' = M\vec{v}$ ，则任何依 $\vec{w} \rightarrow \vec{w}' = M\vec{w}$ 作变换的量都是矢量。动量和加速度是典型的例子。

物理中我们将经常碰到这种思路。当我们把一些量写在一起放在一对括号中时，它们未必是矢量，但一定是在某些线性算符下相互变换的量。线性算符常通过与矩阵的乘法来表示。

¹ 这一点等会儿会有更清楚的说明。

² 一组矢量 $\{\vec{a}, \vec{b}, \vec{c}\}$ 被称为是线性独立的，若方程 $c_1\vec{a} + c_2\vec{b} + c_3\vec{c} = 0$ 成立当且仅当 $c_1 = c_2 = c_3 = 0$ 。这说的其实就是其中任何一个矢量不能表达为其他矢量的线性组合：如果有 $c_1\vec{a} + c_2\vec{b} + c_3\vec{c} = 0$ 对非零的系数成立，则有 $c_1\vec{a} + c_2\vec{b} = -c_3\vec{c}$ 。

A.1 Basis Vectors 基矢

我们可以通过引入基矢来更一般的讨论沿坐标轴的分量这件事。基矢是一组线性独立²且长度为一²的矢量。三维里我们需要三个基矢，从而能将任何矢量都用基矢量的组合来表达。一个简单的

¹译注：本意，即视线的角度。

²译注：一般资料上并不要求最后一条；满足这一条的称为单位基矢。

选择是：

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (\text{A.5})$$

这样任意三维矢量 \vec{v} 都能用基矢表出

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (\text{A.6})$$

v_1, v_2, v_3 三个数叫做 \vec{v} 的分量。注意分量是依赖于基矢的。

前面提到的矢量³ \vec{w} 由此可以写成 $\vec{w} = 0\vec{e}_1 + 4\vec{e}_2 + 0\vec{e}_3$ 。对基矢另一个同样好的选法是

$${}^3\vec{w} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$$

$$\tilde{\vec{e}}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \tilde{\vec{e}}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \tilde{\vec{e}}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (\text{A.7})$$

在这组基下矢量 \vec{w} 看起来会有点不同：

$$\vec{w} = 2\sqrt{2}\tilde{\vec{e}}_1 - 2\sqrt{2}\tilde{\vec{e}}_2 + 0\tilde{\vec{e}}_3 = 2\sqrt{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 2\sqrt{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}. \quad (\text{A.8})$$

从而将 \vec{w} 用在新基矢下的分量写出

$$\tilde{\vec{w}} = \begin{pmatrix} 2\sqrt{2} \\ -2\sqrt{2} \\ 0 \end{pmatrix}.$$

这并不是一个不同的矢量，而只是一个不同的描述！准确地讲， $\tilde{\vec{w}}$ 是原坐标系中的矢量 \vec{w} 在一个相对有旋转的坐标系下描述。³

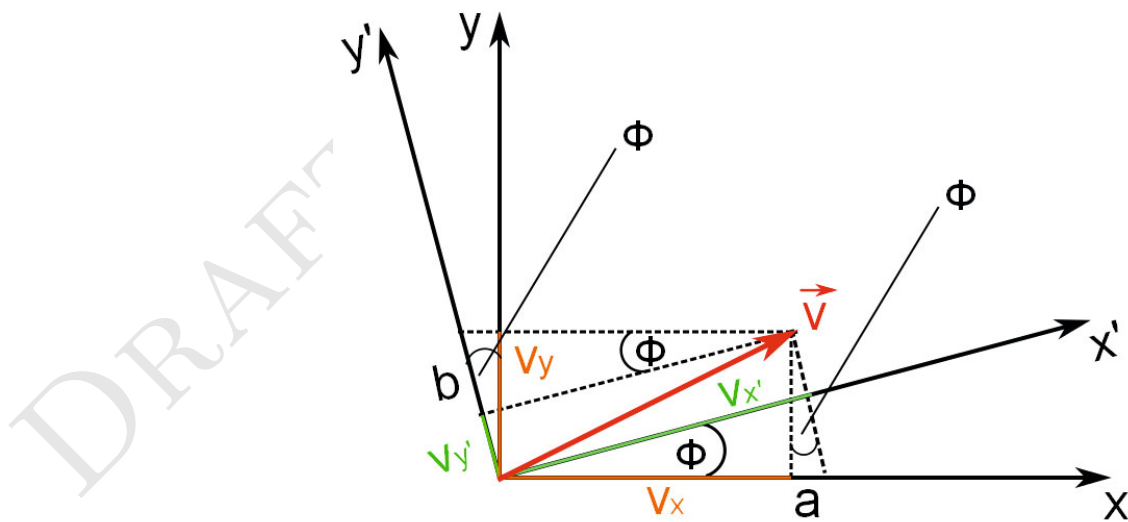
³译注：在整个附录A里作者就没说过几句准确的话，这句话也说得挺糙的。

A.2 Change of Coordinate Systems 坐标系变换

通过矩阵，我们可以更精确地描述不同坐标系间的联系。两个不同的坐标系可以代表实验中两个持不同视角的观察者，或者仅仅是一个想使用一套新的基矢的观察者。这些描述之间的关系是什么？为了避免复杂的讨论，让我们假定这两个坐标系的原点和 z 轴都是一样的。这样的话就只是 x 和 y 轴有区别了。进一步假定实验中某个重要的量用矢量 \vec{v} 描述。

如果第一个观察者看到的是矢量 $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$ ，我们能通过三角函数 $\sin(\phi)$ ， $\cos(\phi)$ 和 $\tan(\phi) = \frac{\sin(\phi)}{\cos(\phi)}$ 来计算出第二个观察者看到的 $\vec{v} = \begin{pmatrix} v_{x'} \\ v_{y'} \\ v_{z'} \end{pmatrix}$ ，如图A.1所示：

图 A.1: 矢量在两个不同的坐标系下分量的示意图。细节见正文。



计算 v_x 和 $v_{x'}$ 的关系，由

$$\cos(\phi) = \frac{v_{x'}}{v_x + a} \rightarrow v_{x'} = (v_x + a) \cos(\phi)$$

以及

$$\tan(\phi) = \frac{a}{v_y} \rightarrow a = v_y \tan(\phi)。$$

得到

$$\begin{aligned} v_{x'} &= (v_x + v_y \tan(\phi)) \cos(\phi) = \left(v_x + v_y \frac{\sin(\phi)}{\cos(\phi)} \right) \cos(\phi) \\ &= v_x \cos(\phi) + v_y \sin(\phi)。 \end{aligned}$$

类似的，利用

$$\cos(\phi) = \frac{v_y}{v_{y'} + b} \rightarrow v_{y'} = v_y \frac{1}{\cos(\phi)} - b$$

和

$$\tan(\phi) = \frac{v_y}{v_{y'} + b} \rightarrow b = v_{x'} \tan(\phi)$$

再借助 $\sin^2(\phi) + \cos^2(\phi) = 1$ 导出

$$\begin{aligned} v_{y'} &= v_y \frac{1}{\cos(\phi)} - v_{x'} \tan(\phi) = v_y \frac{1}{\cos(\phi)} - (v_x \cos(\phi) + v_y \sin(\phi)) \frac{\sin(\phi)}{\cos(\phi)} \\ &= v_y \frac{\sin^2(\phi) + \cos^2(\phi)}{\cos(\phi)} - v_x \sin(\phi) - v_y \frac{\sin^2(\phi)}{\cos(\phi)} = v_y \cos(\phi) - v_x \sin(\phi) \end{aligned}$$

最终有 $v_{y'} = -v_x \sin(\phi) + v_y \cos(\phi)$ 。

我们能将其用一个旋转矩阵来表达：

$$\begin{aligned} \begin{pmatrix} v_{x'} \\ v_{y'} \\ v_{z'} \end{pmatrix} &= R_z(\phi) \vec{v} = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \\ &= \begin{pmatrix} \cos(\phi)v_x + \sin(\phi)v_y \\ -\sin(\phi)v_x + \cos(\phi)v_y \\ v_z \end{pmatrix}。 \end{aligned} \quad (\text{A.9})$$

将矩阵的每一行与未旋转的矢量相乘，便得到了旋转后的矢量。如之前所提到的，沿 z 轴的分量 v_3 对于不同的观察者而言是相同的。矩阵 $R_z(\phi)$ 描述了一个绕 z 轴转动了角度 ϕ 的旋转。

A.3 Matrix Multiplication 矩阵乘法

像这样通过矩阵来作计算是一个极大的简化。矩阵乘法的规则就是行乘列。如果将一个矢量看做一个一列三行的矩阵（一个 3×1 矩阵），则标量积也可以看做是矩阵乘法的一种特殊情况。即

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = \begin{pmatrix} v_x & v_y & v_z \end{pmatrix} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} = v_x w_x + v_y w_y + v_z w_z, \quad (\text{A.10})$$

其中 T 代表转置，即行变成列，列变成行。那么 \vec{v}^T 便是一个 1 行 3 列的矩阵。这样写的话标量积也就成了行乘列的矩阵乘法。

类似的，两个矩阵的乘法也就是左边矩阵的每一行乘右边矩阵的每一列，如图 A.2。下面是一个具体的例子

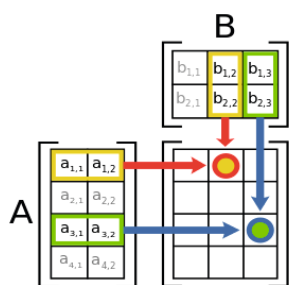


图 A.2: 矩阵乘法示意图。时时刻刻记住行乘列。第一个指标代表所在的行数，第二个代表列数。在例子中，乘积矩阵红色的元素 $c_{1,2} = a_{1,1}b_{1,2} + a_{1,2}b_{2,2}$ ，蓝色的元素 $c_{3,3} = a_{3,1}b_{1,3} + a_{3,2}b_{2,3}$ 。一般的， $c_{i,j} = a_{i,k}b_{k,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \dots$
Figure by Olivier Perrin (Bilou Wikimedia Commons) released under a CC BY-SA 3.0 licence: <http://creativecommons.org/licenses/by-sa/3.0/deed.en>. URL: http://commons.wikimedia.org/wiki/File:Matrix_multiplication_diagram_2.svg, Accessed: 28.1.2015

$$M = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} 0 & 1 \\ 4 & 8 \end{pmatrix}$$

$$MN = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 4 & 8 \end{pmatrix} = \begin{pmatrix} 2 \cdot 0 + 3 \cdot 4 & 2 \cdot 1 + 3 \cdot 8 \\ 1 \cdot 0 + 0 \cdot 4 & 1 \cdot 1 + 0 \cdot 8 \end{pmatrix} = \begin{pmatrix} 12 & 26 \\ 0 & 1 \end{pmatrix} \quad (\text{A.11})$$

时刻把行乘列记在脑子里⁴。注意两个矩阵的乘法是不对易的，即一般来说 $MN \neq NM$ 。

A.4 Scalars 标量

另一个值得注意的事情是，两个矢量标量积的结果对于所有观察者而言都是相同的。这看起来可以作为标量的定义：标量是对于全部观察者都相同的量。这不只是简单的在说每一个数都是标量，因为矢量的每一个分量都是一个数，但是它们对于不同的观察者而言却不一样。作为对比，两个矢量的标量积对于全体观察者却是相同的。这源于，矢量的长度与其与自身的标量积直接相关这一事实。改变观察的视角或位置并不会改变任何东西的长度。矢量的长度被称作旋转下的不变量，即无论我们怎样旋转系统其都保持原样。

⁴译注：说起来，什么是行什么是列也是一个需要记牢的东西，尤其是与台湾友人交流时。

A.5 Right-handed and Left-handed Coordinate Systems 左手/右手坐标系

当我们谈论两个观察者时，我们默认了它们对坐标系的定义都是相同的。而事实上，我们有两种可能的选择，它们这是依旧能通过矩阵乘法相联系，却不能靠转动了。两个观察者可能是一个选择了所谓的右手坐标系，而另一个选择了左手坐标系。

我们无法将一个左手系旋转成一个右手系。事实上，这两类坐标系通过镜面反射相关联。这就是说，右手系和左手系由如下形式的变换相联系⁵

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \rightarrow \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}, \quad (\text{A.12})$$

就是说把全部空间坐标都反一个号。人们也习惯于将这类变换称作**宇称变换 (parity transformation)**。我们可以用如下方式来表示一个宇称变换

$$\vec{v} \rightarrow \vec{v}' = P\vec{v} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}. \quad (\text{A.13})$$

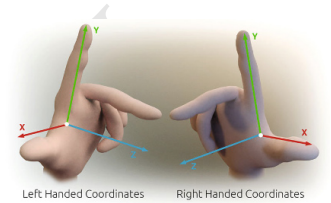


图 A.3: 右手坐标系和左手坐标系。Figure by Primalshell (Wikimedia Commons) released under a CC-BY-SA-3.0 licence: <http://creativecommons.org/licenses/by-sa/3.0/deed.en>. URL: http://commons.wikimedia.org/wiki/File:3D_Cartesian_Coordinate_Handedness.jpg, Accessed: 1.12.2014

⁵译注：原文这里给出的其实是一个相对原点的反射，而不是镜面反射。当然，这两种反射之间只差一个转动。

B. Calulus 微积分

B.1 Product Rule 莱布尼兹律

莱布尼兹律¹

$$\frac{d}{dx}(f(x)g(x)) = \left(\frac{df(x)}{dx}\right)g(x) + f(x)\left(\frac{dg(x)}{dx}\right) \equiv f'g + fg' \quad (\text{B.1})$$

可直接由导数的定义得到

$$\begin{aligned}\frac{d}{dx}[f(x)g(x)] &= \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{[f(x+h)g(x+h) - f(x+h)g(x)] + [f(x+h)g(x) - f(x)g(x)]}{h} \\ &= \lim_{h \rightarrow 0} f(x+h) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h} \\ &= f(x)g'(x) + g(x)f'(x)\end{aligned}$$

B.2 Integration by Parts 分部积分

据说，Peter Lax 被授予美国国家科学奖章时，其他获奖人（可能不是数学家）问他凭借什么拿到的这个奖。

Lax 回答：“我做了个分部积分。”

- Willie Wrong¹

¹ 是他在 www.math.stackexchange.com 上说的。

¹ 译注：直译为“乘积规则”，鉴于这种叫法在中文文本中较罕见，故译作“莱布尼兹律”。

² 见(B.1)式, 并利用微积分基本定理 $\int_a^b h'(x) = h(b) - h(a)$ 。

从莱布尼兹律可以直接得到一个重要的积分手段。将莱布尼兹律一式积分²

$$\underbrace{\int_a^b dx \frac{d}{dx} (f(x)g(x))}_{= f(x)g(x)|_a^b} = \int_a^b dx \left(\frac{df(x)}{dx} \right) g(x) + \int_a^b dx f(x) \left(\frac{dg(x)}{dx} \right) \quad (\text{B.2})$$

移项, 得到

$$\int_a^b dx \left(\frac{df(x)}{dx} \right) g(x) = f(x)g(x)|_a^b - \int_a^b dx f(x) \left(\frac{dg(x)}{dx} \right). \quad (\text{B.3})$$

³ 2.4节中我们说到了没什么能比光速还快, 故无穷远处的场一定不会对有限 x 处的物理产生影响。

当讨论场的物理时, 这个手段很是有用: 在全空间作积分时, 即 $a = -\infty, b = \infty$, 由于一切场量在无穷远处都必须为零³, 边界项为零 $f(x)g(x)|_a^b = 0$ 。

B.3 The Taylor Series Taylor 级数

Taylor 级数使我们能将任何无穷阶可微函数写成幂级数形式

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2}(x-a)^2 f''(a) + \dots \quad (\text{B.4})$$

- 一方面我们能用它来求某些函数的级数形式。这可以用来, 例如, 说明 $e^{ix} = \cos(x) + i\sin(x)$ 。
- 另一方面, 也能用它来得到函数在某点附近的近似值。当出于某些原因我们可以略去高阶项, 不用考虑无穷多项时, 这公式的用处会很大。比如说我们要算函数 $f(x)$ 在 y 的邻域中一点 $y + \Delta y$ 的值, 可以写出⁴

$$f(y + \Delta y) = f(y) + (y + \Delta y - y)f'(y) + \dots = f(y) + \Delta y f'(y) + \dots \quad (\text{B.5})$$

这就是说, 通过在 y 处函数的性质, 我们能得到其在 $y + \Delta y$ 处的近似值。对于无穷小邻域² $\Delta y = \epsilon \rightarrow 0$ 的极端情况, 函数的改变量

⁴ 不要被这些变量的名字搞混了。上面的公式里我们是从 a 出发计算函数在 x 处的值。而这里我们是希望利用 y 处的信息来得到 f 在 $y + \Delta y$ 处的性质。准确的关系是: $x = y + \Delta y$ 和 $a = y$ 。

²译注: 原文有误, 已更正。

可以仅用 Taylor 级数的一（线性）项来表示

$$\Delta f = f(y+\epsilon) - f(y) = f(y) + \epsilon f'(y) + \cdots - f(y) = \epsilon f'(y) + \underbrace{\cdots}_{\approx 0 \text{ 由于 } \epsilon^2 \approx 0} \quad (\text{B.6})$$

这个式子是最有用的数学工具之一，它也能通过微积分基本定理和分部积分推出来。基本定理说

$$\int_a^x dt f'(t) = f(x) - f(a) \rightarrow f(x) = f(a) + \int_a^x dt f'(t). \quad (\text{B.7})$$

可以将第二项用分部积分写开：在 $f'(t)$ 前面补一个 1 成 $1f'(t)$ ，利用(B.3)式取 $g'(t) = 1$ 的情况。分部积分能改写积分⁵ $\int_a^b v'u = vu|_a^b - \int_a^b vu'$ ，积掉一项而对另一项微分。这里我们积 $g'(t) = 1$ 而微分 $f'(t)$ ，即式子里 $g' = v'$ 及 $f' = u$ 。求得

$$f(x) = f(a) + \int_a^x dt f'(t) = f(a) + g(t)f(t)|_a^x - \int_a^b dt g(t)f''(t). \quad (\text{B.8})$$

现在的问题是 $g(t)$ 是啥。我们目前只知道 $g'(t) = 1$ ，但却有无穷多个函数的导数都是这个：对于任意常数 c ，函数 $g = t + c$ 都满足 $g'(t) = 1$ 。我们选取⁶ $g = t - x$ ，即将积分上界的负值 $-x$ 作为我们的常数 c 。从而上式的第二项就可以写作

$$g(t)f(t)|_a^x = (t-x)f(t)|_a^x = \underbrace{(x-x)}_{=0} f(x) - (a-x)f(a) = (x-a)f(a) \quad (\text{B.9})$$

从而整个式子现在写作

$$f(x) = f(a) + (x-a)f'(a) + \int_a^x dt (x-t)f''(t). \quad (\text{B.10})$$

对最后一项再次使用分部积分，现在是⁷ $v' = (x-t)$ 和 $u = f''(t)$ ：

$$\rightarrow \int_a^x dt \underbrace{(x-t)}_{=v'} \underbrace{f''(t)}_{=u} = \underbrace{-\frac{1}{2}(x-t)^2}_{=v} \underbrace{f''(t)}_{=u} \Big|_a^x - \int_a^x dt \underbrace{\left(-\frac{1}{2}(x-t)^2\right)}_{=v} \underbrace{f'''(t)}_{=u'}$$

⁵ 译注：原文有误，已更正。

⁶ 方程对任意 c 都成立，我们自然可以随便选一个，但那样就推不出我们想要的公式了。选取恰当的 c 是为了得到一个有用的关于 $f(x)$ 的公式。否则 $f(x)$ 将会同时出现在等式两侧。

⁷ 注意 $v' = (x-t)$ 对 t 的积分会出一个符号： $v = -(x-t)^2 + d$ 积分常数 d 取作零。

边界项同样简单的是零

$$-\frac{1}{2}(x-t)^2 f''(t) \Big|_a^x = -\frac{1}{2} \underbrace{(x-x)^2}_{=0} f''(x) + \frac{1}{2}(x-a)^2 f''(a)。$$

得到公式

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2}(x-a)^2 f''(a) + \int_a^x dt \frac{1}{2}(x-t)^2 f'''(t)。 \quad (\text{B.11})$$

可以继续将最后一项分部积分，但是我们应该能直接肉眼观察到规律了。Taylor 级数可以利用求和号 Σ 写成更紧凑的形式：

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \\ &= \frac{f^{(0)}(a)(x-a)}{0!} + \frac{f^{(1)}(a)(x-a)^1}{1!} + \frac{f^{(2)}(a)(x-a)^2}{2!} \quad (\text{B.12}) \\ &= \frac{f^{(3)}(a)(x-a)^3}{3!} + \dots, \end{aligned}$$

其中 $f^{(n)}$ 代表 f 的 n 阶导数，例如 $f^{(2)} = f''$ ， $n!$ 代表 n 的阶乘，即 $n! = 1 \cdot 2 \cdot 3 \dots n$ 。比如对 $n = 5$ 有 $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ ， $2! = 2 \cdot 1 = 2$ ，以及定义 $0! = 1$ 。级数是下一节的主题。

B.4 Series 级数

上一节中我们碰巧发现了含有无穷求和的一个重要等式。在这一节里将介绍关于求和的一些小技巧 and 几个重要级数。

B.4.1 Important Series 几个重要的级数

从上一节中我们学到了任意无穷可微的函数都可以写成级数的形式³。让我们从最重要的例子开始：指数函数 e^x 。它的泰勒展开式可以直接写出：

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad (\text{B.13})$$

³译注：一个典型的反例是 $f(x) = \exp(-1/x^2)$ ，它在原点处（补充定义后）任意阶导数都为零。

用到了指数函数的⁴定义，导函数等于自身： $(e^x)' = e^x$ 。将 $a = 0$ 带入 Taylor 级数，利用 $e^0 = 1$ ，便得到指数函数的 Taylor 级数 ((B.12)式)：

$$e^x = \sum_{n=0}^{\infty} \frac{e^0(x-0)^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (\text{B.14})$$

这个级数也可以作为 e^x 的定义式。

另外两个无穷可微的重要函数是 $\sin(x)$ 和 $\cos(x)$ 。利用 $(\sin(x))' = \cos(x)$ ， $(\cos(x))' = -\sin(x)$ ， $\cos(0) = 1$ 以及 $\sin(0) = 0$ ，可以推得它们的泰勒级数。

$$\sin(x) = \sum_{n=0}^{\infty} \frac{\sin^{(n)}(0)(x-0)^n}{n!}。$$

由于 $\sin(0) = 0$ ，故全部偶数 n 项都为零，由此可以将求和分开：注意到

$$\sum_{n=0}^{\infty} n = \sum_{n=0}^{\infty} (2n+1) + \sum_{n=0}^{\infty} (2n) \quad (\text{B.15})$$

$$1 + 2 + 3 + 4 + 5 + 6 \dots = 1 + 3 + 5 \dots + 2 + 4 + 6 + \dots$$

将 $\sin(x)$ 的求和分裂开来

$$\begin{aligned} \sin(x) &= \sum_{n=0}^{\infty} \frac{\sin^{(2n+1)}(0)(x-0)^{2n+1}}{(2n+1)!} + \underbrace{\sum_{n=0}^{\infty} \frac{\sin^{(2n)}(0)(x-0)^{2n}}{(2n)!}}_{=0} \\ &= \sum_{n=0}^{\infty} \frac{\sin^{(2n+1)}(0)(x-0)^{2n+1}}{(2n+1)!}。 \end{aligned} \quad (\text{B.16})$$

$\sin(x)$ 的全部偶数阶导数， $\sin^{(2n)}$ 也是 $\sin(x)$ (除去可能的负号)，由于 $\sin(0) = 0$ 所以说第二项为零。 $\sin(x)$ 全部奇数阶导数

⁴译注：某一种

是 $\cos(x)$ ，除去可能的负号。我们有

$$\begin{aligned}\sin(x)^{(1)} &= \cos(x) \\ \sin(x)^{(2)} &= \cos'(x) = -\sin(x) \\ \sin(x)^{(3)} &= -\sin'(x) = -\cos(x) \\ \sin(x)^{(4)} &= -\cos'(x) = \sin(x) \\ \sin(x)^{(5)} &= \sin'(x) = \cos(x)\end{aligned}\tag{B.17}$$

规律是 $\sin^{(2n+1)}(x) = (-1)^n \cos(x)$ ，你可以将任意整数 n 带入验证⁸。由此将 (B.16) 式改写作

$$\begin{aligned}\sin(x)^{(1)} &= \sin(x)^{(2 \cdot 0 + 1)} \\ &= (-1)^0 \cos(x) \\ &= \cos(x), \sin(x)^{(3)} \\ &= \sin(x)^{(2 \cdot 1 + 1)} \\ &= (-1)^1 \cos(x) \\ &= -\cos(x)\end{aligned}$$

$$\begin{aligned}\sin(x) &= \sum_{n=0}^{\infty} \frac{\sin^{(2n+1)}(0)(x-0)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \cos(0)(x-0)^{2n+1}}{(2n+1)!} \\ &= \underbrace{\cos(0)=1}_{\cos(0)=1} \sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n+1}}{(2n+1)!}\end{aligned}\tag{B.18}$$

这就是 $\sin(x)$ 的 Taylor 级数，也可以作为 $\sin(x)$ 的定义式。类似的，我们可以得到

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n}}{(2n)!},\tag{B.19}$$

其奇数阶导数都正比于 $\sin(0) = 0$ 。

B.4.2 Splitting Sums 分裂求和

上一节中我们提到了一个有用的计算技巧。那儿我们用了下面这个例子

$$\sum_{n=0}^{\infty} n = \sum_{n=0}^{\infty} (2n+1) + \sum_{n=0}^{\infty} (2n)\tag{B.20}$$

$$1 + 2 + 3 + 4 + 5 + 6 \dots = 1 + 3 + 5 \dots + 2 + 4 + 6 + \dots$$

来表明我们可以将任何求和分解成偶数和奇数项。 $2n$ 总是偶数，而 $2n+1$ 总是奇数。在上一节中我们就看到了这很有用，但现在可以再看看另一个例子。如果将带有复变量 ix 的指数函数的级数

分裂开来会怎样？

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(ix)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(ix)^{2n+1}}{(2n+1)!} \end{aligned} \quad (\text{B.21})$$

用 $(ix)^{2n} = i^{2n} x^{2n}$ 和 $i^{2n} = (i^2)^n = (-1)^n$ 来写开。另外也需要 $i^{2n+1} = i \cdot i^{2n} = i(-1)^n$ 。由此有

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n+1}}{(2n+1)!} \\ &= \underbrace{\sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n}}{(2n)!}}_{=\cos(x) \text{ 见(B.19)式}} + i \underbrace{\sum_{n=0}^{\infty} \frac{(-1)^n (x)^{2n+1}}{(2n+1)!}}_{=\sin(x) \text{ 见(B.18)式}} \\ &= \cos(x) + i \sin(x) \end{aligned} \quad (\text{B.22})$$

B.4.3 Einstein's Sum Convention Einstein 求和约定

物理中经常有求和，老是写大大的求和号 Σ 是很烦人的一件事。所以说某个聪明人引入了一个约定，叫 Einstein 求和约定。根据这个约定每当一个指标在某些项中重复出现两次，如同上面的 n ，那么就意味着一个隐藏的求和号。即

$$a_n a_n \equiv \sum_n a_n b_n. \quad (\text{B.23})$$

其他形式的例子

$$a_n b_n c_m \equiv \sum_n a_n b_n c_m. \quad (\text{B.24})$$

$$a_m b_n c_m \equiv \sum_m a_m b_n c_m. \quad (\text{B.25})$$

但是

$$a_n b_m \neq \sum_n a_n b_m, \quad (\text{B.26})$$

因为一般而言 $n \neq m$ 。没有同伴的指标叫做自由指标 (free index)，而有同伴要求和的叫做哑指标 (dummy index)。具体原因会在下一节中解释。

在 $a_n b_n c_m \equiv \sum_n a_n b_n c_m$ 一例中， n 是哑指标，而 m 是自由指标。对于 $a_m b_n c_m \equiv \sum_m a_m b_n c_m$ ，指标 m 是哑的而 n 是自由的。

B.5 Index Notation 指标记号

B.5.1 Dummy Indices 哑指标

⁹ 当然我们不能将一个指标改成另一类指标。例如，我们可以将改 $i \rightarrow j$ 但不能改 $i \rightarrow \mu$ ，因为 μ 这样的希腊字母指标总是从 0 加到 3，而 i 这样的罗马数字指标是从 1 加到 3。

一定要注意到的一点是，有伴的指标事实上什么用处都没有。将 n 重命名为 k ，什么事都不改变⁹，即 n 已经被缩并掉了

$$a_n b_n c_m = a_k b_k c_m \equiv \sum_n a_n b_n c_m \equiv \sum_k a_k b_k c_m. \quad (\text{B.27})$$

另一方面，自由指标却不能自由的被重命名。将 m 改写作 q 会造成区别，因为等式另一端也必须要有相同的自由指标。这就是说当我们单独考察类似于 $a_n b_n c_m$ 这样的项时，总需要将其他带有自由指标 m 的项拉进来，指标要改名就得一起改名。我们来看一个例子

$$F_i = \epsilon_{ijk} a_j b_k. \quad (\text{B.28})$$

这里除了一个新东西 ϵ_{ijk} ，带着多个指标，我们暂且不去管它，待会儿会仔细讨论。这里我们能将 $\epsilon_{ijk} a_j b_k$ 的指标 j, k 随意重命名，因为它们被缩并掉了。例如 $j \rightarrow u, k \rightarrow z$ ，这样得到 $\epsilon_{iuz} a_u b_z$ 和之前的完全是一样的。另一头， i 就不是一个哑指标了，我们不能将其重命名 $i \rightarrow m$ ： $\epsilon_{muz} a_u b_z$ ，因为这样的话上式就会变成

$$F_i = \epsilon_{muz} a_u b_z. \quad (\text{B.29})$$

再看之前那句话这就显得有点迂腐了，因为很明显的将 F_i 里的 i 改掉来得到一个有意义的等式。但一般来说我们都会单独的讨论一个对象，那就得清楚我们不能去改变哪些东西。

B.5.2 Objects with more than One Index 带多个指标的对象

现在可以谈谈那些有多个指标的对象。带两个指标的自然就是矩阵了。第一个指标告诉我们变量是在哪一行而第二个指标代

表哪一行。例如

$$M_{ij} \equiv \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}. \quad (\text{B.30})$$

对于 M_{12} 就是在第一行第二列。

利用指标可以简单的写出矩阵乘法。两个矩阵的乘积是

$$MN = (MN)_{ij} = M_{ik}N_{kj}. \quad (\text{B.31})$$

在左边是矩阵乘积 (MN) 的 i 行 j 列的元素，可以由 M 的第 i 行和 N 的第 j 列相乘得到。指标 k 出现了两次，代表一个隐含的求和。有时我们会给一个对象两个或更多的指标（张量）。在这本书中，两个一般就够用了，唯一的例外将在下面几节中讨论到。

B.5.3 Symmetric and Antisymmetric Indices 对称/反对称指标

一个矩阵被称为是对称的，若 $M_{ij} = M_{ji}$ 。对于二维情形来说就是 $M_{12} = M_{21}$ ，一个对称矩阵的例子如下

$$\begin{pmatrix} 9 & 3 \\ 3 & 17 \end{pmatrix} \quad (\text{B.32})$$

一个矩阵被称为是全反对称的 (totally antisymmetric)，若 $M_{ij} = -M_{ji}$ 对于全体 i, j 成立。一个示例是

$$\begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix} \quad (\text{B.33})$$

注意对角元都是零，这是因为 $M_{11} = -M_{11}$ 仅当 $M_{11} = 0$ 时成立，对于 M_{22} 也一样。

B.5.4 Antisymmetric \times Symmetric Sums 对称和反对称求和

一个重要的事实：当我们将带有完全相同的指标的对称和反对称的对象乘起来求和时，结果为零：

$$\sum_{ij} a_{ij}b_{ij} = 0 \quad (\text{B.34})$$

若有 $a_{ij} = -a_{ji}$ 和 $b_{ij} = b_{ji}$ 对于全部 i, j 成立。这个结论很容易看出来：改写

$$\sum_{ij} a_{ij} b_{ij} = \frac{1}{2} \left(\sum_{ij} a_{ij} b_{ij} + \sum_{ij} a_{ij} b_{ij} \right) \quad (\text{B.35})$$

正如之前解释过的，我们能自由的将哑指标重命名 $i \rightarrow j$ 及 $j \rightarrow i$ ，仅对第二项作用

$$\rightarrow \sum_{ij} a_{ij} b_{ij} = \frac{1}{2} \left(\sum_{ij} a_{ij} b_{ij} + \sum_{ij} a_{ji} b_{ji} \right) \quad (\text{B.36})$$

然后利用 b_{ij} 的对称性和 a_{ij} 的反对称性，来翻转第二项中的指标，得到¹⁰

$$\begin{aligned} \rightarrow \sum_{ij} a_{ij} b_{ij} &= \frac{1}{2} \left(\sum_{ij} a_{ij} b_{ij} + \sum_{ij} \underbrace{a_{ji}}_{=-a_{ij}} \underbrace{b_{ji}}_{=b_{ij}} \right) \\ &= \frac{1}{2} \left(\sum_{ij} a_{ij} b_{ij} - \sum_{ij} a_{ij} b_{ij} \right) = 0 \end{aligned} \quad (\text{B.37})$$

¹⁰ 如果你觉得这是一个无聊的把戏，算几个真实的例子来验证其正确性。

B.5.5 Two Important Symbols 两个重要的符号

第一个是单位矩阵。二维情形下为

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (\text{B.38})$$

单位矩阵用指标记号来标记时称 Kronecker 符号 δ_{ij} ，对任意维度都可由下式定义

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (\text{B.39})$$

Kronecker 符号是对称的，有 $\delta_{ij} = \delta_{ji}$ 。

第二个是 Levi-Civita 符号 ϵ_{ijk} ，在二维下由下式定义

$$\epsilon_{ij} = \begin{cases} 1 & (i, j) = (1, 2) \\ 0 & i = j \\ -1 & (i, j) = (2, 1) \end{cases} \quad (\text{B.40})$$

矩阵形式写作

$$\epsilon_{ij} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (\text{B.41})$$

三维下的 Levi-Civita 符号是

$$\epsilon_{ijk} = \begin{cases} 1 & (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2) \\ 0 & i = j \text{ 或 } j = k \text{ 或 } k = i \\ -1 & (i, j, k) = (1, 3, 2), (3, 2, 1), (2, 1, 3) \end{cases} \quad (\text{B.42})$$

对于四维

$$\epsilon_{ijkl} = \begin{cases} 1 & \text{若 } (i, j, k, l) \text{ 是 } (1, 2, 3, 4) \text{ 的偶置换} \\ -1 & \text{若 } (i, j, k, l) \text{ 是 } (1, 2, 3, 4) \text{ 的奇置换} \\ 0 & \text{其他} \end{cases} \quad (\text{B.43})$$

例如 $(1, 2, 4, 3)$ 和 $(2, 1, 4, 3)$ 就分别是 $(1, 2, 3, 4)$ 的奇置换（换了一个）和偶置换（换了两个）。

Levi-Civita 符号是全反对称的：根据定义，当置换两个指标时，总是多一个负号：

$$\epsilon_{ijk} = -\epsilon_{jik}, \quad \epsilon_{ijk} = -\epsilon_{ikj} \text{ 等。或者二维下 } \epsilon_{ij} = \epsilon_{ji}。$$

C. Linear Algebra 线性代数

利用矩阵等线性代数工具可以大幅简化计算，下面从基本变换开始对线性代数做一回顾。

C.1 Basic Transformations 基本变换

矩阵的复共轭定义为：

$$M_{ij}^* = \begin{pmatrix} M_{11}^* & M_{12}^* \\ M_{21}^* & M_{22}^* \end{pmatrix}, \quad (\text{C.1})$$

即对每个矩阵元取复共轭¹。

矩阵的转置定义为 $M_{ij}^T = M_{ji}$ ，用矩阵形式表示为：

$$M_{ij} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \rightarrow M_{ij}^T = \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix}, \quad (\text{C.2})$$

即交换矩阵的行与列。两个矩阵的乘积的转置 $(MN)^T \neq M^T N^T$ ，而是 $(MN)^T = N^T M^T$ ，即乘积的转置需要调换乘积顺序，这用指标记号不难证明：

$$\begin{aligned} MN &\equiv (MN)_{ij} = M_{ik} N_{kj} \\ (MN)^T &\equiv ((MN)_{ij})^T = (MN)_{ji} = (M_{ik} N_{kj})^T \end{aligned}$$

¹ 复数 $z = a + ib$ (a 为 z 的实部, b 为虚部, a, b 均为实数) 的复共轭为 $z^* = a - ib$ 。

$$(M_{ik}N_{kj})^T = M_{ik}^T N_{kj}^T = M_{ki}N_{jk} = N_{jk}M_{ki} \equiv N^T M^T, \quad (\text{C.3})$$

其中最后一步利用了矩阵乘法的指标记号，左侧矩阵的行乘以右侧矩阵的列。我们将 $M_{ki}N_{jk}$ 写为 $N_{jk}M_{ki}$ 的形式，后者即为左侧矩阵的行乘以右侧矩阵的列的形式，这样就能用矩阵乘法记作 $N^T M^T$ 。

注意在指标记号下乘积中的各项顺序可自由颠倒，例如 $M_{ki}N_{jk} = N_{jk}M_{ki}$ ，因为各 M_{ki} 与 N_{jk} 只代表数字。

C.2 Matrix Exponential Function 矩阵指数函数

前面已经导出了指数函数的级数形式，由此可定义矩阵的指数函数 e^M ，其中 M 为任意方阵：

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}. \quad (\text{C.4})$$

必须注意，一般 $e^M e^N \neq e^{M+N}$ 。 $e^M e^N = e^{M+N}$ 当且仅当 $MN = NM$ 。

C.3 Determinants 行列式

² 若矩阵 M 可逆，则它存在逆矩阵，记作 M^{-1} ，满足 $M^{-1}M = MM^{-1} = \mathbf{1}$ (单位矩阵)。

矩阵行列式的概念不太直观，但是相当有用。例如，若矩阵的行列式非零，则该矩阵可逆²。这个命题的证明有点复杂，读者可在任何一本线性代数教材中查阅。

3×3 矩阵的行列式 $\det(A)$ 可以用 Levi-Civita 符号定义：

$$\det(A) = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \epsilon_{ijk} A_{1i} A_{2j} A_{3k}, \quad (\text{C.5})$$

$n \times n$ 矩阵的行列式的定义类似：

$$\det(A) = \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n \epsilon_{i_1 i_2 \dots i_n} A_{1i_1} A_{2i_2} \cdots A_{ni_n}. \quad (\text{C.6})$$

例： 2×2 矩阵的行列式

$$\det(A) = \det \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} = (3 \cdot 2) - (5 \cdot 1) = 1$$

一般 3×3 矩阵的行列式：

$$\det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1). \quad (\text{C.7})$$

C.4 Eigenvalues and Eigenvectors 本征值与本征向量

本征值与本征向量是物理学中大用特用的线性代数概念。矩阵 M 的特征值 λ 与特征向量 \vec{v} 满足：

$$M\vec{v} = \lambda\vec{v}. \quad (\text{C.8})$$

上式两侧作用的是同一向量 \vec{v} ，这意味着 \vec{v} 经矩阵 M 变换后只改变原来的常数 λ 倍。每一本征向量对应一个本征值。计算本征值与本征向量的过程有点繁复，可以参考任何一本线性代数教材。

本征向量重要性的一个例子是旋转变换，我们用旋转矩阵描述旋转操作，用旋转矩阵的本征向量定义各旋转轴。

C.5 Diagonalization 对角化

利用本征值与本征向量可以使矩阵对角化，这能简化计算，并且有助于解释方程的物理意义。可以证明任意可对角化的矩阵 M 能写为如下形式³：

$$M = N^{-1}DN, \quad (\text{C.9})$$

其中矩阵 N 由 M 的本征向量组成，矩阵 D 是以 M 的本征值为对角元的对角矩阵：

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = N^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} N = (\vec{v}_1, \vec{v}_2)^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} (\vec{v}_1, \vec{v}_2). \quad (\text{C.10})$$

³ 一般情况，变换 $M' = N^{-1}MN$ 意味着基的变换， M' 是 M 在新坐标系中的形式。因此矩阵对角化的含义就是可以找到一个坐标系使得矩阵 M 在该坐标系中的形式简单（对角矩阵）。

D. Additional Mathematical Notions

其他数学概念

D.1 Fourier 变换

Fourier 变换的思想类似于把一个任意向量按照一组基底矢量 $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ 做展开¹，在三维欧氏空间下最常见的基底选择为

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (\text{D.1})$$

任意三维矢量 \vec{v} 可以按这组基底展开为

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3 = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (\text{D.2})$$

傅里叶变换的概念就是对函数做类似的展开²，对于周期函数来说展开的基底是 $\sin(kx)$ 和 $\cos(kx)$ ，也就是说任意周期函数都可以

¹关于后者的详细介绍参见附录 A.1

²从更抽象的层面来说，函数是广义上的矢量，也就是说它们是一个矢量空间中的元素。对于不同种类的函数它们所组成的矢量空间也是不同的。这种广义矢量空间的定义和通常欧氏矢量空间的类似（后者就是我们常见的带小箭头 $\vec{}$ 的位置矢量构成的空间），比如说，你可以像加两个向量一样把两个函数加起来，并可以对函数定义标量积。

写成如下形式³

$$f(x) = \sum_{k=0}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \quad (\text{D.3})$$

其中 a_k 和 b_k 为常系数。

任意（无需周期条件）函数可以按基底 e^{ikx} 和 e^{-ikx} 展开，但是这时候求和号被积分号代替⁴：

$$f(x) = \int_0^{\infty} dk (a_k e^{ikx} + b_k e^{-ikx}) \quad (\text{D.4})$$

我们也可以把上式写成

$$f(x) = \int_{-\infty}^{\infty} dk f_k e^{-ikx} \quad (\text{D.5})$$

展开系数 f_k 可以写成 $\tilde{f}(k)$ ，其常被称为 $f(x)$ 的傅里叶变换。

D.2 δ 分布函数

从某种角度上来说， δ 分布函数在积分中起的作用和 Kronecker δ 符号⁵在求和中起的作用类似。举个例子，考虑

$$\sum_{i=1}^3 a_i b_j = a_1 b_j + a_2 b_j + a_3 b_j \quad (\text{D.6})$$

现在我们想把求和的第二项提取出来，我们可以利用 Kronecker δ 符号 δ_{2i} ，这样得到

$$\sum_{i=1}^3 \delta_{2i} a_i b_j = \delta_{21} a_1 b_j + \delta_{22} a_2 b_j + \delta_{23} a_3 b_j = a_2 b_j \quad (\text{D.7})$$

写成更一般的式子即为

$$\int_{i=3}^3 \delta_{ik} a_i b_j = a_k b_j \quad (\text{D.8})$$

³在某些点上未必如此，具体取决于函数和边界条件。更详细的介绍可以参见任意一本数学物理方法教材（鉴于大多数读者都是物理系而非数学系的学生）——译者

⁴复习一下，积分实际上就是求和的极限过程， \sum_k 中分立变量 k 在 $\int dk$ 中变成了连续变量。

⁵Kronecker δ 符号定义见附录 B.5.5

δ 分布函数 $\delta(x - y)$ 由下式定义：

$$\int dx f(x) \delta(x - y) = f(y) \quad (\text{D.9})$$

和 Kronecher δ 符号非常类似， δ 分布函数从积分中提取出了一项⁶，我们还可以通过相似性从式

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \quad (\text{D.10})$$

推出

$$\frac{\partial f(x_i)}{\partial f(x_j)} = \delta(x_i - x_j) \quad (\text{D.11})$$

通过相似性来导出此式当然不能算是一个证明，如果你愿意你也可以通过严格的数学推导得到这个式子。关于 δ 分布函数还有很多内容可以讲，但就此书的目的而言你只需要知道它是干什么的就足够了。事实上，这也是 Dirac 最早引入这一函数的方式。

⁶即 $x = y$ 的一项，例如， $\int dx f(x) \delta(x - 2) = f(2)$.