

中国科学技术大学

电动力学讲义

陶鑫

March 17, 2023

前言

2023 版说明：重写了前四章，使之与上课内容更加对应，并补充了一些上课时略过的内容。

2022 版说明：将原英文讲义非常快速地翻成了中文，加了一部分原讲义里没有，但讲课时觉得有用的内容。因为翻译的过于快速，错误很多。

2018 版说明：一个主要变化是重新加入了从分析力学的角度推导 Maxwell 方程的章节，因为这一章节非常有助于从场论的角度理解电动力学，同时也强调了场作为另一种描述物理现象的方式与粒子的同等重要地位。与 2013 年的讲义不同的是，这次的场的分析力学部分为选修章节，不会计入最终成绩，也不会参加正常的考试部分。

2015 版说明：这本讲义主要是我自己讲课时参考所用，其内容主要参考了 Landau 和 Lifshitz 的《The Classic Theory of Fields》，David J. Griffiths 的《Introduction to Electrodynamics (4th Edition)》。其余部分还参考了 Jackson 的《Classical Electrodynamics》。在这三本书中，Landau 和 Jackson 的书并不太适合一般本科生阅读。这个讲义里对 Landau 的书作了大幅简化，而 Jackson 的书则只是参考了一小部分数学证明和一些个别有意思的章节。正常情况下，我推荐的阅读书是 Griffiths 的电动力学导论。

这个讲义第一版的使用是在 2014 年春季学期，当时采用的单位制是 Gauss 单位制，因为我觉得 Gauss 单位制更适合理论电动力学。但在实行过程中，不少同学向我建议换回国际单位制，以求更好地和之前的课程衔接。因此从 2015 年春季学期开始，我对这个讲义作了比较大幅度的改写，一是增加更多的基础数学知识，二是去除场的分析力学部分，三是将 Gauss 单位制换为 SI 单位制。在这里要特别感谢金泽宇，王贤瞳，和吴昊楠三位同学的协助，他们在改单位制的过程中付出了大量劳动，并且修正了 2014 年讲义中的很多错误。2014 年春季学期班上还有好几位其他同学在 14 年教学过程中也陆续地指出讲义中的错误，但遗憾当时并没有记录名单，在此一并感谢。

陶鑫

二零二一年七月

目录

1	简要数学基础	1
1.1	向量分析	1
1.2	向量运算技巧	4
1.3	张量简介	8
1.4	狄拉克 δ 函数	12
2	狭义相对论	15
2.1	简要历史背景	15
2.2	狭义相对论基础	16
2.3	狭义相对论的几何表达	27
2.4	张量表达与物理规律的不变性	32
3	外场中的带电粒子运动	35
3.1	经典粒子的最小作用量原理	35
3.2	自由粒子的相对论动力学	36
3.3	自由粒子运动方程的协变形式	38
3.4	给定外场中电荷的运动方程	39
3.5	电磁场势函数	43
3.6	电磁场张量	47
4	电磁场方程	49
4.1	拉格朗日方程：从离散到连续	49
4.2	场的最小作用量原理与拉格朗日方程	50
4.3	电磁场的拉格朗日量与场方程	55
4.4	连续性方程	59
4.5	场的能量与动量	61
4.6	诺特定理：对称性与守恒量 *	64
5	静电场	70
5.1	静电场的方程	70
5.2	静电场的能量	71
5.3	静电场的多极展开	72
5.4	外场中的电荷系统	76
5.5	介质的极化过程	77
5.6	一般的边界条件	79
5.7	边值问题与唯一性定理	79
5.8	静像电荷法	81
5.9	分离变量法	83
6	静磁学	88
6.1	静磁场	88
6.2	磁场的多极展开	89
6.3	外场中小电流体系的受力	91

6.4 介质的磁化	93
7 时变电磁场	95
7.1 真空中的电磁场方程	95
7.2 单色平面电磁波	96
7.3 波动的极化	97
7.4 线性介质中的电磁波	99
7.5 介质中的电磁波	100
7.6 折射率对频率的依赖关系模型	103
7.7 导体中的电磁波	105
7.8 导体表面电磁波的反射	107
7.9 时变电磁场的边值问题	108
8 运动粒子的场	112
8.1 一般情况下的解：推迟势	112
8.2 点电荷的推迟势：李纳-维谢尔势 (Lienard-Wiechert potentials)	114
9 电磁波的辐射	121
9.1 点电荷的辐射	121
9.2 辐射阻尼	125
9.3 电荷体系在远场近似下的辐射	127
9.4 电偶极辐射	129
9.5 磁偶极与电四极辐射	131
9.6 自由电子对电磁波的散射	133
9.7 束缚电子对电磁波的散射	135

第一章 简要数学基础

本章里我们介绍全书的主要数学基础，包括两个方面。第一是向量与张量的定义和相关运算。场作为时间与空间的函数，场的各种行为以及其与粒子间的相互作用常常与场的梯度，散度或旋度有关。因此，为了更好地理解电磁场理论，我们需要熟悉向量和张量的定义与运算技巧，以达到不依赖于公式表即可运算的地步。同时本课程会着重从坐标变换的角度去定义向量与张量，这一点对于理解物理规律的相对论不变性至为重要。第二是狄拉克 δ 函数。物理学中的很多理论是建立在点粒子模型之上的。而在处理与点粒子相关的物理量，特别是各种“密度”类的物理量时，不可避免地会运用到 δ 函数，电磁场理论也不例外。我们的介绍更多的是关注这些数学工具在物理学中的使用，但在数学严格性上可能会有所欠缺。

1.1 向量分析

1.1.1 向量定义

在之前的物理课里，向量通常被简单定义为有大小和方向的量，比如位移，力，速度等。这种定义在牛顿力学里直观易懂，也很容易图象化。但是在更高维度或非欧式空间里，这种定义往往不是特别方便直观。在这一节里，我们引入另一种向量的数学定义，针对欧式空间中的三维向量，如果任意三个数的组合，在坐标系旋转的时候，其转换的方式与位移矢量一致，则该三个数构成一个三维向量。

从以上定义里可以看到，位移矢量可以被认为所有矢量的“原型矢量”。这种向量的定义对于理解后续相对论四维时空下的（四维）向量特别重要，也有助于理解物理量在不同惯性参考系之间的变换规律。例如将光速记为 c ，则四维时空的基本位移矢量是 (ct, \mathbf{x}) ，而能量 ε ，与动量 \mathbf{p} 在一起，也会构成四维动量向量， $(\varepsilon/c, \mathbf{p})$ 。从向量定义可以知道， ε 与 \mathbf{p} 在不同惯性参考系之间的变换方式会分别类似于 t 和 \mathbf{x} 。

1.1.2 爱因斯坦求和约定

在讨论向量的定义与运算之前，我们先介绍使用爱因斯坦求和约定去表达向量以常见向量运算。使用该约定会大大方便向量分析，少写很多求和符号。据说爱因斯坦在做运算的时候，发现可以省略求和符号 \sum ，而不影响结果。比如

$$\mathbf{A} \cdot \mathbf{B} = \sum_{i=1}^3 A_i B_i = A_i B_i, \quad (1.1)$$

因此，只要在上式中将重复的下标 i 理解为对其求和，则式子的含义不变。这里，重复的指标 i 被称为“哑指标” (dummy index)。除了哑指标外，向量运算里还会用到不重复的下标，称为“自由指标” (free index)。例如， \mathbf{A} 的第 i 个分量需要表达为 A_i 。在实际使用中，遵守以下三条规则，就可以方便正确地使用爱因斯坦求和约定来进行运算：

1. 默认对哑指标求和，除非特别说明；
2. 一个运算式里，不同项中应该包含相同的自由指标；
3. 在任一项里，同一个指标最多出现两次。

利用求和约定，我们可以得到 $\mathbf{A} = A_i \mathbf{e}_i$ ，因为 $A_i \mathbf{e}_i = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2 + A_3 \mathbf{e}_3$ 。同样的，如果三维位移矢量 $\mathbf{x} = x_i \mathbf{e}_i$ ，则基于求和约定

$$\frac{\partial x_i}{\partial x_i} = 3. \quad (1.2)$$

初学者或注意力不集中时容易得出错误的结果 $\partial x_i / \partial x_i = 1$ ，因为这里默认对 i 从 1 到 3 求和。

例题 1.1 从向量 $\mathbf{A} = A_i \mathbf{e}_i$ 与 $\mathbf{B} = B_i \mathbf{e}_i$ 出发，得到两个向量的点乘表达式。

解 利用 \mathbf{A} 与 \mathbf{B} 的定义, 我们有

$$\mathbf{A} \cdot \mathbf{B} = A_i \mathbf{e}_i \cdot B_j \mathbf{e}_j = A_i B_j \mathbf{e}_i \cdot \mathbf{e}_j = A_i B_j \delta_{ij} = A_i B_i, \quad (1.3)$$

在上式中的 δ 符号定义为

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \quad (1.4)$$

注意, 这里为了避免 i 指标出现超过两次, 我们为 \mathbf{A} 与 \mathbf{B} 向量选用了不同的哑指标。否则, $\mathbf{A} \cdot \mathbf{B} = A_i \mathbf{e}_i \cdot B_i \mathbf{e}_i$, 这是因为在错误的式子里, i 指标出现了 4 次。

注意到哑指标是可以随意更换的, 即 $\mathbf{A} \cdot \mathbf{B} = A_j B_j = A_k B_k$, 只要不违反在求和约定中, 指标最多出现两次的规定即可。

向量间的主要运算有两个, 一个是点乘, 一个是叉乘。用爱因斯坦求和约定表达这两种运算是向量运算的基础。点乘的表达比较简单, 已在上例中给出。但叉乘要稍微复杂一点, 因为从分量来看, 在直角坐标系中

$$\mathbf{A} \times \mathbf{B} = (A_2 B_3 - A_3 B_2) \mathbf{e}_1 + (A_3 B_1 - A_1 B_3) \mathbf{e}_2 + (A_1 B_2 - A_2 B_1) \mathbf{e}_3.$$

上式并非直接以求和的形式表达的。因为要用爱因斯坦求和约定来表达叉乘, 我们需要引入新的符号, Levi-Civita 符号 ϵ_{ijk} , 定义为:

$$\epsilon_{ijk} = \begin{cases} +1 & i, j, k \text{ 为 } 1, 2, 3 \text{ 的偶数次置换;} \\ -1 & i, j, k \text{ 为 } 1, 2, 3 \text{ 的奇数次置换;} \\ 0 & \text{其它情况: 两个或三个指标相同.} \end{cases} \quad (1.5)$$

利用该符号, 我们可以很容易得出,

$$\mathbf{A} \times \mathbf{B} = \epsilon_{ijk} \mathbf{e}_i A_j B_k = \epsilon_{ijk} A_j B_k \mathbf{e}_i, \quad (1.6)$$

或者两个向量叉乘的第 i 个分量为

$$(\mathbf{A} \times \mathbf{B})_i = \epsilon_{ijk} A_j B_k. \quad (1.7)$$

Levi-Civita 符号本身的一些性质在向量运算中也非常重要。比如从定义中看 $\epsilon_{ijk} = \epsilon_{jki}$, $\epsilon_{ijk} = -\epsilon_{jik}$, 以及 $\epsilon_{iik} = 0$ 。另外, 一个关于 ϵ_{ijk} 的重要等式是

$$\epsilon_{ijk} \epsilon_{imn} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}. \quad (1.8)$$

这个式子并不难记, 只需要注意右边第一项中的 δ 符号下标对应着左边 Levi-Civita 符号的同位下标组合, 而第二项则是 Levi-Civita 符号的异位下标组合。这种式子常常会出现在三个向量的叉乘式中。下面举一个例子。

例题 1.2 用求和约定写出表达式 $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$ 的分量形式。

解

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \epsilon_{ijk} \mathbf{e}_i (\mathbf{A} \times \mathbf{B})_j C_k. \quad (1.9)$$

注意回避已有哑指标, 利用 $(\mathbf{A} \times \mathbf{B})_j = \epsilon_{jmn} A_m B_n$, 我们有

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \epsilon_{ijk} \mathbf{e}_i \epsilon_{jmn} A_m B_n C_k = \mathbf{e}_i \epsilon_{jki} \epsilon_{jmn} A_m B_n C_k \quad (1.10)$$

在最后一步, 我们运用了 $\epsilon_{ijk} = \epsilon_{jki}$ 。利用

$$\epsilon_{jki} \epsilon_{jmn} = \delta_{km} \delta_{in} - \delta_{kn} \delta_{im}. \quad (1.11)$$

因此

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{e}_i (\delta_{km} \delta_{in} - \delta_{kn} \delta_{im}) A_m B_n C_k = A_m C_m B_i \mathbf{e}_i - B_k C_k A_i \mathbf{e}_i. \quad (1.12)$$

仔细看最后一步的分量表达式对应的向量式为 $(\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{B} \cdot \mathbf{C}) \mathbf{A}$ 。

更多关于上述式子的运用, 我们会在之后的章节里碰到。

1.1.3 ∇ 算符

对初学者而言，向量张量运算之所以复杂，很多都是由 ∇ 算符引起的。在这一小节里，我们用求和约定来表达 ∇ 算符，以方便进行相关的向量运算。 ∇ 算符的表达形式可以从标量的梯度来得到， ∇T 。在直角坐标系里，

$$\nabla T = \mathbf{e}_1 \frac{\partial T}{\partial x_1} + \mathbf{e}_2 \frac{\partial T}{\partial x_2} + \mathbf{e}_3 \frac{\partial T}{\partial x_3} = \mathbf{e}_i \frac{\partial T}{\partial x_i}. \quad (1.13)$$

所以有

$$\nabla = \mathbf{e}_i \frac{\partial}{\partial x_i}. \quad (1.14)$$

这个等式清楚地表明 ∇ 算符同时具有微分算符和向量的性质，只要我们注意到这一点，相关的运算便不会太难。同时，上述表达式会是本课程中进行与 ∇ 算符相关的向量运算的基础。

利用 (1.14)，我们可以快速得到直角坐标系中一些常见的与 ∇ 相关物理量/算符的表达式。例如， $\nabla^2 \equiv \nabla \cdot \nabla$ 称为拉普拉斯算符 (Laplacian)，其作用在标量 f 上的表达式为

$$\nabla^2 f = \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot \mathbf{e}_j \frac{\partial}{\partial x_j} f = \delta_{ij} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f = \frac{\partial^2 f}{\partial x_i^2}. \quad (1.15)$$

又如一个向量的散度为

$$\nabla \cdot \mathbf{A} = \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot A_j \mathbf{e}_j = \frac{\partial A_i}{\partial x_i}. \quad (1.16)$$

类似的，一个向量 \mathbf{A} 的旋度为

$$\nabla \times \mathbf{A} = \epsilon_{ijk} \mathbf{e}_i \nabla_j A_k = \epsilon_{ijk} \mathbf{e}_i \frac{\partial}{\partial x_j} A_k = \epsilon_{ijk} \frac{\partial A_k}{\partial x_j} \mathbf{e}_i. \quad (1.17)$$

所以， \mathbf{A} 向量的旋度的第 i 个分量为

$$(\nabla \times \mathbf{A})_i = \epsilon_{ijk} \frac{\partial A_k}{\partial x_j}. \quad (1.18)$$

需要注意的是，上述关于 ∇ 的表达式是针对直角坐标系的。在一般的曲线坐标系下，相应的式子会稍加复杂，在本课程里不予讨论。这里，我们仅列出特殊而常见的柱坐标系与球坐标系下的相应表达式，以供需要之时查阅。

在球坐标下 $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi)$,

标量梯度：

$$\nabla f = \frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \mathbf{e}_\phi,$$

向量散度：

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta A_\theta) + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi},$$

向量旋度：

$$\begin{aligned} (\nabla \times \mathbf{A})_r &= \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta A_\phi) - \frac{1}{r \sin \theta} \frac{\partial A_\theta}{\partial \phi} \\ (\nabla \times \mathbf{A})_\theta &= \frac{1}{r \sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{1}{r} \frac{\partial}{\partial r} (r A_\phi), \\ (\nabla \times \mathbf{A})_\phi &= \frac{1}{r} \frac{\partial}{\partial r} (r A_\theta) - \frac{1}{r} \frac{\partial A_r}{\partial \theta}. \end{aligned}$$

拉普拉斯算符作用在标量上:

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}.$$

在柱坐标系 $(\mathbf{e}_r, \mathbf{e}_\phi, \mathbf{e}_z)$ 下,

标量梯度:

$$\nabla f = \frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \phi} \mathbf{e}_\phi + \frac{\partial f}{\partial z} \mathbf{e}_z,$$

向量散度:

$$\nabla \cdot \mathbf{A} = \frac{1}{r} \frac{\partial}{\partial r} (r A_r) + \frac{1}{r} \frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z},$$

向量旋度:

$$\begin{aligned} (\nabla \times \mathbf{A})_r &= \frac{1}{r} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z} \\ (\nabla \times \mathbf{A})_\phi &= \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}, \\ (\nabla \times \mathbf{A})_z &= \frac{1}{r} \frac{\partial}{\partial r} (r A_\phi) - \frac{1}{r} \frac{\partial A_r}{\partial \phi}. \end{aligned}$$

拉普拉斯算符作用在标量上

$$\nabla^2 f = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}.$$

1.2 向量运算技巧

在这一章里,我们将介绍向量运算的技巧。掌握了这些技巧,就可以在向量运算时不查询公式表了,也对后续场论的学习至关重要。

常见的向量运算一般有

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} \times \mathbf{C} &= \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} = \mathbf{A} \times \mathbf{B} \cdot \mathbf{C} \\ \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= (\mathbf{C} \times \mathbf{B}) \times \mathbf{A} = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C} \\ \nabla \times \nabla f &= 0 \\ \nabla \cdot \nabla \times \mathbf{A} &= 0 \\ \nabla(fg) &= f\nabla g + g\nabla f \\ \nabla \cdot (f\mathbf{A}) &= f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f \\ \nabla \times (f\mathbf{A}) &= f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A} \\ \nabla \cdot (\mathbf{A} \times \mathbf{B}) &= \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B} \\ \nabla(\mathbf{A} \cdot \mathbf{B}) &= \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} \\ \nabla \times (\nabla \times \mathbf{A}) &= \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \end{aligned}$$

接下来,我们就按难度递进的顺序介绍如何通过一些比较巧的方法记住或掌握上述向量等式。

1.2.1 基本复合向量运算

上述向量运算表中的头两行等式不包含 ∇ 算符，属于最基本的复合向量运算。对于这两行等式，我们介绍记住它们的记巧。

首先, 对于 $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$ ，我们可以交换点乘和叉乘的顺序。

$$\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{A} \times \mathbf{B} \cdot \mathbf{C}.$$

或者我们可以循环顺移向量的位置，比如

$$\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{B} \cdot \mathbf{C} \times \mathbf{A}.$$

这两种方式都可以在不同情况下方便计算。

其次, 对于三个向量的叉乘, $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$, 我们可以使用”先中间后外边”法¹。这里的中间和外边, 指的是括号里的两个向量在三个向量中的相对位置。比如,

$$\begin{aligned} \mathbf{A} \times (\underbrace{\mathbf{B}}_{\text{中间}} \times \underbrace{\mathbf{C}}_{\text{外边}}) &= \underbrace{\mathbf{B}}_{\text{中间}} (\underbrace{\mathbf{A} \cdot \mathbf{C}}_{\text{另两个向量点乘}}) - \underbrace{\mathbf{C}}_{\text{外边}} (\underbrace{\mathbf{B} \cdot \mathbf{A}}_{\text{另两个向量点乘}}). \\ (\underbrace{\mathbf{B}}_{\text{外边}} \times \underbrace{\mathbf{C}}_{\text{中间}}) \times \mathbf{A} &= \underbrace{\mathbf{C}}_{\text{中间}} (\underbrace{\mathbf{B} \cdot \mathbf{A}}_{\text{另两个向量点乘}}) - \underbrace{\mathbf{B}}_{\text{外边}} (\underbrace{\mathbf{A} \cdot \mathbf{C}}_{\text{另两个向量点乘}}). \end{aligned}$$

利用这个“先中间后外边”法处理三个向量的叉乘会非常方便。同时, 从上面的式子中, 我们也可以看到, 与普通数字的乘法不同, 向量的叉乘不满足结合率, 即 $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$ 。

例题 1.3 假如有两个向量 \mathbf{A} 与 \mathbf{B} , 证明 \mathbf{A} 向量垂直于 \mathbf{B} 向量的分量可以表达为 $\mathbf{A}_\perp = (\hat{\mathbf{B}} \times \mathbf{A}) \times \hat{\mathbf{B}}$ 。这里 $\hat{\mathbf{B}} \equiv \mathbf{B}/B$, 即 \mathbf{B} 向量方向上的单位矢量。

证明 利用 “先中间后外边” 法则, 我们有

$$(\hat{\mathbf{B}} \times \mathbf{A}) \times \hat{\mathbf{B}} = \mathbf{A}(\hat{\mathbf{B}} \cdot \hat{\mathbf{B}}) - \hat{\mathbf{B}}(\mathbf{A} \cdot \hat{\mathbf{B}}) = \mathbf{A} - A_\parallel \hat{\mathbf{B}} = \mathbf{A}_\perp. \quad (1.19)$$

在进行复杂一点的向量运算时, 我们常常会用到上述两条基本运算规律。

1.2.2 含有 ∇ 算符的向量运算

对于含有 ∇ 算符的运算, 我们需要时刻注意该算符本身的微分-向量双重属性。因为该算符的向量属性, 很容易得到

$$\begin{aligned} \nabla \times \nabla f &= 0, \\ \nabla \cdot \nabla \times \mathbf{A} &= 0. \end{aligned}$$

即一个标量的梯度其旋度为 0, 而一个向量的旋度其散度为 0。同时, 对 ∇ 算符而言, 微分原则适用, 即 $(ab)' = a'b + ab'$ 。所以有

$$\nabla(fg) = g\nabla f + f\nabla g.$$

同样的道理, 以下两个运算等式并不难理解。

$$\begin{aligned} \nabla \cdot (f\mathbf{A}) &= f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f, \\ \nabla \times (f\mathbf{A}) &= f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A}. \end{aligned}$$

¹Fundamentals of Plasma Physics by Paul M. Bellan

这里只需要注意到对标量 f 而言, 被 ∇ 算符作用的方式, 只有 ∇f 一种。同时, 等式两边 ∇ 与向量 \mathbf{A} 的运算方式不变, 即叉乘还是叉乘, 点乘还是点乘。

对于一些更复杂的式子, 比如 $\nabla \times (\mathbf{A} \times \mathbf{B})$, 我们需要一些技巧才能同时兼顾 ∇ 算符的两种属性。例如, 如果我们直接把 ∇ 当成普通向量, 按照上面的“先中间后外边”法则, 很容易得出

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A}). \quad (1.20)$$

或者类似的式子。但这个结论显然是错误的。原因就在于上式没有充分考虑到 ∇ 算符的微分属性。接下来我们介绍两个方法来有效地处理复杂的向量运算。

1.2.3 下标法

这个方法主要包含三步, 其使用的主要技巧是利用下标标记, 将 ∇ 算符的微分向量属性分开。为方便表示, 我们称该方法为“下标法”。让我们还是以 $\nabla \times (\mathbf{A} \times \mathbf{B})$ 为例子来讲解该方法。

第一步: 将 ∇ 当作一个算符, 因为其微分属性, 必然即作用于 \mathbf{A} 也作用于 \mathbf{B} 上。所以我们要将 ∇ 加上下标, 以表明它所作用的向量, 即

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = \nabla_{\mathbf{A}} \times (\mathbf{A} \times \mathbf{B}) + \nabla_{\mathbf{B}} \times (\mathbf{A} \times \mathbf{B}).$$

这一步的主要目的, 在于分离了 ∇ 算符的微分属性。完成这一步后, $\nabla_{\mathbf{A}}$ 与 $\nabla_{\mathbf{B}}$ 就可以当作普通向量来进行运算。

第二步: 因为 $\nabla_{\mathbf{A}}$ 只作用在 \mathbf{A} 上, 而 $\nabla_{\mathbf{B}}$ 只作用在 \mathbf{B} 上, 所以接下来只需将它们按照普通向量的规律, 重新调整式子, 使得 $\nabla_{\mathbf{A}}$ 只作用在 \mathbf{A} 上, 而 $\nabla_{\mathbf{B}}$ 只作用在 \mathbf{B} 上。

$$\begin{aligned} \nabla_{\mathbf{A}} \times (\mathbf{A} \times \mathbf{B}) &= (\nabla_{\mathbf{A}} \cdot \mathbf{B})\mathbf{A} - \mathbf{B}(\nabla_{\mathbf{A}} \cdot \mathbf{A}) \quad \text{中间-外边法} \\ &= (\mathbf{B} \cdot \nabla_{\mathbf{A}})\mathbf{A} - \mathbf{B}(\nabla_{\mathbf{A}} \cdot \mathbf{A}). \quad \text{点乘向量交换次序} \end{aligned}$$

类似的, 我们有

$$\begin{aligned} \nabla_{\mathbf{B}} \times (\mathbf{A} \times \mathbf{B}) &= (\nabla_{\mathbf{B}} \cdot \mathbf{B})\mathbf{A} - (\nabla_{\mathbf{B}} \cdot \mathbf{A})\mathbf{B} \\ &= (\nabla_{\mathbf{B}} \cdot \mathbf{B})\mathbf{A} - (\mathbf{A} \cdot \nabla_{\mathbf{B}})\mathbf{B}. \end{aligned}$$

第三步: 完成上两步后, 我们可以将相应的下标标记去掉, 所以最终的结果是

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - \mathbf{B}(\nabla \cdot \mathbf{A}) + (\nabla \cdot \mathbf{B})\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B}.$$

例题 1.4 用下标法证明 $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$.

证明 第一步, 我们给 ∇ 算符加上 \mathbf{A}, \mathbf{B} 下标, 即

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \nabla_{\mathbf{A}} \cdot (\mathbf{A} \times \mathbf{B}) + \nabla_{\mathbf{B}} \cdot (\mathbf{A} \times \mathbf{B}).$$

第二步, 将加了下标后的 ∇ 算符, 当普通向量进行运算, 使得 ∇ 算符作用到相对应的量上。

$$\begin{aligned} \nabla_{\mathbf{A}} \cdot (\mathbf{A} \times \mathbf{B}) &= \nabla_{\mathbf{A}} \times \mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \nabla_{\mathbf{A}} \times \mathbf{A}, \\ \nabla_{\mathbf{B}} \cdot (\mathbf{A} \times \mathbf{B}) &= -\nabla_{\mathbf{B}} \times \mathbf{B} \cdot \mathbf{A} = -\mathbf{A} \cdot \nabla_{\mathbf{B}} \times \mathbf{B}. \end{aligned}$$

第三步, 去掉下标, 得到相应等式。

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$$

在熟悉这个方法后, 只要记得清楚, 也可以不加下标。

1.2.4 分量法

分量法利用了向量式是不依赖于具体坐标系的原则。同样的向量式不管在直角坐标系还是球坐标系中都是成立的。利用这一点，我们可以将一个向量式表达为在直角坐标系下的分量形式，然后对该分量式进行必要的运算与分析，最后再将分量式重新表达为向量式。对初学者来说，最后一步是往往是比较难的。使用这个方法时，我们一般都会用到爱因斯坦求和约定。现在我们来看几个例子。

例题 1.5 假如 $\mathbf{x} = x_i \mathbf{e}_i$ 为位移矢量，证明 $\nabla \cdot \mathbf{x} = 3$ 以及 $\nabla \times \mathbf{x} = 0$ 。

证明 以第一式为例，第一步：直角坐标系下的分量形式

$$\nabla \cdot \mathbf{x} = \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot x_j \mathbf{e}_j.$$

第二步，进行相关分析

$$\mathbf{e}_i \frac{\partial}{\partial x_i} \cdot x_j \mathbf{e}_j = \frac{\partial x_j}{\partial x_i} \delta_{ij} = \frac{\partial x_i}{\partial x_i} = 3.$$

因为结果已经是个常数，相当于第三步已经完成。

类似的道理，

$$\nabla \times \mathbf{x} = \mathbf{e}_i \epsilon_{ijk} \frac{\partial}{\partial x_j} x_k = \mathbf{e}_i \epsilon_{ijk} \delta_{jk} = 0. \quad (1.21)$$

另外几个与位移矢量 \mathbf{x} 相关的常用向量等式有：

$$\begin{aligned} \nabla \mathbf{x} &= \mathbf{I}, \\ \nabla r &= \mathbf{x}/r = \mathbf{n}, \\ \nabla \times [f(r)\mathbf{x}] &= 0. \end{aligned}$$

这里 $r = |\mathbf{x}|$, \mathbf{n} 为 \mathbf{x} 方向上的单位矢量，而 \mathbf{I} 为下一节会介绍的单位张量。这几个式子使用分量法可以很容易证明。

例题 1.6 利用分量法，证明

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot \nabla \times \mathbf{B}.$$

证明 首先我们写出上式在直角坐标系中的形式并进行相关运算，

$$\begin{aligned} \nabla \cdot (\mathbf{A} \times \mathbf{B}) &= \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot (\epsilon_{jkl} A_k B_l) \mathbf{e}_j = \frac{\partial}{\partial x_i} (\epsilon_{ikl} A_k B_l) \\ &= \epsilon_{ikl} \left(\frac{\partial A_k}{\partial x_i} B_l + A_k \frac{\partial B_l}{\partial x_i} \right) = \epsilon_{ikl} \frac{\partial A_k}{\partial x_i} B_l + \epsilon_{ikl} A_k \frac{\partial B_l}{\partial x_i} \\ &= \epsilon_{lik} \frac{\partial A_k}{\partial x_i} B_l - A_k \epsilon_{kil} \frac{\partial B_l}{\partial x_i}. \end{aligned}$$

接下来，我们得到上述分量形式的向量式。注意到

$$\epsilon_{lik} \frac{\partial A_k}{\partial x_i} = (\nabla \times \mathbf{A})_l,$$

因此，

$$\epsilon_{lik} \frac{\partial A_k}{\partial x_i} B_l = (\nabla \times \mathbf{A}) \cdot \mathbf{B}. \quad (1.22)$$

类似的道理，

$$A_k \epsilon_{kil} \frac{\partial B_l}{\partial x_i} = \mathbf{A} \cdot (\nabla \times \mathbf{B}).$$

综上, 可得

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$$

例题 1.7 利用分量法证明

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla) \mathbf{A} - \mathbf{B}(\nabla \cdot \mathbf{A}) + (\nabla \cdot \mathbf{B}) \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B}.$$

证明 首先, 我们写出左边式子在直角坐标系中的分量形式,

$$\begin{aligned} \nabla \times (\mathbf{A} \times \mathbf{B}) &= \epsilon_{ijk} \frac{\partial}{\partial x_j} (\mathbf{A} \times \mathbf{B})_k \mathbf{e}_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} \epsilon_{klm} A_l B_m \mathbf{e}_i \\ &= \epsilon_{ijk} \epsilon_{klm} \left(\frac{\partial A_l}{\partial x_j} B_m + \frac{\partial B_m}{\partial x_j} A_l \right) \mathbf{e}_i. \end{aligned}$$

重新排列一下两个 Levi-Civita 符号的下标顺序, 好利用等式 (1.8), 即

$$\begin{aligned} \nabla \times (\mathbf{A} \times \mathbf{B}) &= \epsilon_{kij} \epsilon_{klm} \left(\frac{\partial A_l}{\partial x_j} B_m + \frac{\partial B_m}{\partial x_j} A_l \right) \mathbf{e}_i \\ &= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \left(\frac{\partial A_l}{\partial x_j} B_m + \frac{\partial B_m}{\partial x_j} A_l \right) \mathbf{e}_i. \end{aligned}$$

利用 δ_{ij} 符号的性质, 我们很容易得到,

$$\begin{aligned} \nabla \times (\mathbf{A} \times \mathbf{B}) &= \delta_{il} \delta_{jm} \left(\frac{\partial A_l}{\partial x_j} B_m + \frac{\partial B_m}{\partial x_j} A_l \right) \mathbf{e}_i - \delta_{im} \delta_{jl} \left(\frac{\partial A_l}{\partial x_j} B_m + \frac{\partial B_m}{\partial x_j} A_l \right) \mathbf{e}_i \\ &= \left(B_j \frac{\partial A_i}{\partial x_j} + A_i \frac{\partial B_j}{\partial x_j} - B_i \frac{\partial A_j}{\partial x_j} - A_j \frac{\partial B_i}{\partial x_j} \right) \mathbf{e}_i. \end{aligned}$$

最后我们从分量形式回复到向量表式, 注意到 $B_i \mathbf{e}_i \rightarrow \mathbf{B}$ 以及

$$B_j \frac{\partial}{\partial x_j} \rightarrow \mathbf{B} \cdot \nabla, \quad \frac{\partial B_j}{\partial x_j} \rightarrow \nabla \cdot \mathbf{B}, \quad \dots,$$

所以我们有

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla) \mathbf{A} + (\nabla \cdot \mathbf{B}) \mathbf{A} - (\nabla \cdot \mathbf{A}) \mathbf{B} - (\mathbf{A} \cdot \nabla) \mathbf{B}.$$

对于本课程的学习, 需要通过练习去掌握这两种向量运算的技巧。特别的, 掌握分量法对于后续四维向量的相关运算也有帮助。

1.3 张量简介

1.3.1 张量的引入

在定义张量之前, 我们先举个简单的例子来说明, 为什么张量不是一个特别复杂的概念, 并且引入张量在物理学中是一件自然而然的事情。我们要用到的物理量是通量 (flux) 的概念, 比如电荷通量, 能量通量。一个物理量 X 的通量, 指的是单位时间通过单位面积的 X 的量是多少, 方向沿与流速方向一致。如果单位体积的流体中 X 的量是 ρ_X (X 的密度), 而流速是 \mathbf{v} , 则单位时间内将要通过单位面积的流体位于以单位面积为底, 以 $\mathbf{v} \times$ 单位时间为高的柱体内。因此, 不难理解 X 的通量 (\mathbf{P}_X) 应该具有如下形式:

$$\mathbf{P}_X = \rho_X \times \text{单位面积} \times \mathbf{v} \times \text{单位时间} = \rho_X \mathbf{v}. \quad (1.23)$$

这样我们可以写出比如电荷通量, 能量通量等。假设电荷密度是 ρ , 能量密度是 ϵ , 则电荷通量是 $\rho \mathbf{v}$, 即电流密度 \mathbf{j} , 而能量通量则是 $\epsilon \mathbf{v}$ 。这些都是向量, 方向是流速方向。对于一个给定的面, \mathbf{A} , 通过该面的电荷通量是 $\rho \mathbf{v} \cdot \mathbf{A}$,

即当该面的法向与 \mathbf{v} 一致时，单位时间内通过的电荷量最大为 $\rho v A$ ，而当该面的法向与 \mathbf{v} 垂直是，无电荷通过。类似的，单位时间内经过该面的能量则是 $\varepsilon \mathbf{v} \cdot \mathbf{A}$ 。

流体除了携带能量外，显然还有动量。假定粒子数密度为 N ，而 \mathbf{p} 是平均的单个粒子动量，则单位体积的流体动量为 $\mathbf{g} = N\mathbf{p}$ 。而随着流体的流动，显然也有动量通量的概念。那对于动量通量，我们该如何表达？首先，我们可以先看动量的某个分量的通量，比如 g_x, x 方向的动量。其通量显然具有 $g_x v$ 的形式。类似的，我们会有 $g_y v, g_z v$ 这样的各分量通量。然而如果只能按分量写动量通量的话，会有三个分量，更重要的是所写出来的分量通量会依赖于具体选取的坐标系（例如 $g_x v$ ）。那怎么写出单个不依赖于具体坐标系的表达式来表达动量通量呢？答案是很明显的，动量通量应该具有 $\mathbf{g} \mathbf{v}$ 或者 $\mathbf{v} \mathbf{g}$ 的形式。这种形式的物理量就是一个典型的张量。

广义而言，向量，标量都是张量。向量是一阶张量，而标量是零阶张量。很容易发现，张量的阶数等于其分量自由指标的个数。如二阶张量有两个自由指标， $g_i v_j$ 。向量（一阶张量）则只有一个自由指标， A_i 。而标量（零阶张量）则没有自由指标， ϕ 。在不会引起歧义的情况下，我们将二阶张量简称为张量。

1.3.2 张量的定义

上述 $\mathbf{v} \mathbf{g}$ 形式的张量称为并矢 (dyad)。记法就是两个向量（又称矢量）并在一起，中间没有其它符号，比如， \mathbf{AB} 就是一个由 \mathbf{A} 和 \mathbf{B} 构成的并矢。它有两个方向，一个来自 \mathbf{A} ，一个来自 \mathbf{B} 。用爱因斯坦求和约定，

$$\mathbf{AB} = A_i \mathbf{e}_i B_j \mathbf{e}_j = A_i B_j \mathbf{e}_i \mathbf{e}_j, \quad (1.24)$$

其中 $\mathbf{e}_i \mathbf{e}_j$ 方向上的分量即是 $A_i B_j$ 。在这门课里，你应该习惯用方程 (1.24) 的方式去表达张量以及进行相关运算。

基本的张量运算可以通过理解向量与并矢的点乘得到。一个向量 \mathbf{X} 点乘并矢 \mathbf{AB} 的结果是

$$\mathbf{X} \cdot \mathbf{AB} \equiv (\mathbf{X} \cdot \mathbf{A}) \mathbf{B},$$

或者并矢 \mathbf{AB} 点乘向量 \mathbf{Y} 的结果是

$$\mathbf{AB} \cdot \mathbf{Y} \equiv \mathbf{A}(\mathbf{B} \cdot \mathbf{Y}).$$

即将点乘符号两边的向量先点乘，得到的标量结果再乘剩下的向量。如果用分量形式表达，则

$$\mathbf{X} \cdot \mathbf{AB} = X_i \mathbf{e}_i \cdot A_j B_k \mathbf{e}_j \mathbf{e}_k = X_i A_j B_k (\mathbf{e}_i \cdot \mathbf{e}_j) \mathbf{e}_k = X_i A_i B_k \mathbf{e}_k.$$

或者

$$\mathbf{AB} \cdot \mathbf{Y} = A_i B_j \mathbf{e}_i \mathbf{e}_j \cdot Y_k \mathbf{e}_k = A_i B_j Y_k \mathbf{e}_i (\mathbf{e}_j \cdot \mathbf{e}_k) = A_i B_k Y_k \mathbf{e}_i.$$

从以上的运算过程就可以看出一般情况下 $\mathbf{AB} \neq \mathbf{BA}$ ，因为 $\mathbf{AB} \cdot \mathbf{C}$ 是一个平行于 \mathbf{A} 的向量，而 $\mathbf{BA} \cdot \mathbf{C}$ 是一个平行于 \mathbf{B} 的向量。在这门课里张量的最主要形式就是并矢，例如后续会讲到的麦克斯韦应力张量 (Maxwell stress tensor)。类似的，一个向量也可以叉乘一个并矢，结果仍然是一个并矢。例如

$$\mathbf{X} \times \mathbf{AB} = (\mathbf{X} \times \mathbf{A}) \mathbf{B}, \quad (1.25)$$

$$\mathbf{AB} \times \mathbf{X} = \mathbf{A}(\mathbf{B} \times \mathbf{X}). \quad (1.26)$$

因此可见，涉及到并矢的计算通常会比较简单。此外，并矢在不同坐标系的变换方式与两个向量的积相同。例如，如果在从 K' 到 K 系，向量 \mathbf{A}, \mathbf{B} 分别变成 \mathbf{A}', \mathbf{B}' ，则并矢 \mathbf{AB} 会变成 $\mathbf{A}' \mathbf{B}'$ 。

并非所有的张量都是并矢的形式。一般的二阶张量是

$$\mathbf{F} = F_{ij} \mathbf{e}_i \mathbf{e}_j.$$

虽然 \mathbf{F} 不是并矢的形式，但可以写成多个并矢的和。例如，

$$\mathbf{F} = \mathbf{ab} + \mathbf{cd} + \mathbf{ef} + \cdots$$

在三维情况下，最少需要三个并矢的和去表达一个一般的二阶张量。从一般张量与并矢的关系可以看出，其不同坐标系的变换方式与两个向量的积相同。有一些特殊的张量具有对称或反对称性。如果 $T_{ij} = T_{ji}$ ，则该张量为对称张量；而如果 $T_{ij} = -T_{ji}$ ，则该张量为反对称张量。对于反对称张量，我们显然有 $T_{ii} = 0$ ，这里不对 i 求和。对于一个一般的二阶张量 F_{ij} ，我们可以将其拆成一个对称部分，与反对称部分。利用

$$F_{ij} = \underbrace{\frac{1}{2}(F_{ij} + F_{ji})}_{S_{ij}} + \underbrace{\frac{1}{2}(F_{ij} - F_{ji})}_{A_{ij}}$$

显然这里 S_{ij} 为 \mathbf{F} 的对称部分，而 A_{ij} 为 \mathbf{F} 的反对称部分。

例题 1.8 假如从两个坐标系 K 与 K' 系， K' 系由 K 以 z 轴旋转得到。显然， K 系中的向量 \mathbf{A} 与 K' 系中的向量的坐标变换关系具有的形式为

$$A'_i = M_{ij}A_j$$

求张量 \mathbf{T}' 与 \mathbf{T} 间的变换式。

解 利用张量的定义， \mathbf{T}' 与 \mathbf{T} 之间的变换式类似于两个向量的积，或者将其形式上看成一个并矢， \mathbf{AB} 。所以有

$$T'_{ij} \sim A'_i B'_j = M_{ik} A_k M_{jl} B_l \sim M_{ik} M_{jl} T_{kl}.$$

例如，假定 M_{ij} 具有如下形式

$$M_{ij} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

即

$$A'_1 = aA_1 + bA_2,$$

$$A'_2 = cA_1 + dA_2,$$

$$A'_3 = A_3.$$

而

$$\begin{aligned} T'_{12} \sim A'_1 B'_2 &= (aA_1 + bA_2)(cB_1 + dB_2) = acA_1B_1 + adA_1B_2 + bcA_2B_1 + bdA_2B_2 \\ &\sim acT_{11} + adT_{12} + bcT_{21} + bdT_{22}. \end{aligned}$$

或者直接利用

$$T'_{12} = M_{1k} M_{2l} T_{kl} = acT_{11} + adT_{12} + bcT_{21} + bdT_{22}.$$

类似的方式可以得到 \mathbf{T} 张量的其它分量的变换式。

如果 T_{ij} 为对称或反对称张量，则显然上式还可以进一步简化。例如，如果 T_{ij} 为对称张量，则

$$T'_{12} = acT_{11} + (ad + bc)T_{12} + bdT_{22}$$

而如果 T_{ij} 为反对称张量，则

$$T'_{12} = acT_{11} + (ad - bc)T_{12} + bdT_{22}.$$

一般的张量与向量的点乘运算也可以利用并矢与向量的点乘得到，其结果是一个向量，例如，

$$\mathbf{F} \cdot \mathbf{A} = F_{ij} \mathbf{e}_i \mathbf{e}_j \cdot A_k \mathbf{e}_k = F_{ij} A_k \delta_{jk} = F_{ij} A_j \mathbf{e}_i.$$

所以一次点乘会将相关项的张量阶数和降 2。例如在上式中，左边张量的阶数和为 $2 + 1 = 3$ ，而右边的向量阶数为 1。又比如，一个张量与两个向量的点乘结果是一标量 ($2 + 1 + 1 - 2 - 2 = 0$)，即

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{B} = A_i \mathbf{e}_i \cdot F_{jk} \mathbf{e}_j \mathbf{e}_k \cdot B_l \mathbf{e}_l = A_i F_{jk} B_l \delta_{ij} \delta_{kl} = A_i F_{ik} B_k.$$

这种两次点乘也可以用双点乘 (:) 表达：

$$\mathbf{A} \cdot \mathbf{F} \cdot \mathbf{B} \equiv \mathbf{BA} : \mathbf{F} \equiv \mathbf{F} : \mathbf{BA}.$$

所以一个张量与一个并矢的双点乘的结果为该张量先与其中一个向量点乘，所得结果（向量）再与剩下的向量点乘，最终的结果为一标量。

一个特殊的张量是单位张量 \mathbf{I} ，该张量的特点是

$$\mathbf{I} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I} = \mathbf{A}. \quad (1.27)$$

其分量形式为 $\mathbf{I} = \delta_{ij} \mathbf{e}_i \mathbf{e}_j$ 。对于该张量，很容易证明

$$\mathbf{I} : \mathbf{AB} = \mathbf{A} \cdot \mathbf{B},$$

$$\nabla \cdot (\varphi \mathbf{I}) = \nabla \varphi,$$

$$\mathbf{I} : \nabla \mathbf{B} = \nabla \cdot \mathbf{B}.$$

例题 1.9 证明 $\mathbf{I} : \nabla \mathbf{B} = \nabla \cdot \mathbf{B}$.

证明 利用分量形式，我们有

$$\begin{aligned} \mathbf{I} : \nabla \mathbf{B} &= \delta_{ij} \mathbf{e}_i \mathbf{e}_j : \mathbf{e}_k \frac{\partial}{\partial x_k} B_m \mathbf{e}_m = \delta_{ij} (\mathbf{e}_i \mathbf{e}_j \cdot \mathbf{e}_k) \cdot \mathbf{e}_m \frac{\partial}{\partial x_k} B_m \\ &= \delta_{ij} \delta_{im} \delta_{jk} \frac{\partial}{\partial x_k} B_m = \frac{\partial}{\partial x_k} B_k = \nabla \cdot \mathbf{B}. \end{aligned}$$

例题 1.10 证明

$$\nabla \cdot (\mathbf{fg} \times \mathbf{r}) = [\nabla \cdot (\mathbf{fg})] \times \mathbf{r} + \mathbf{g} \times \mathbf{f}.$$

证明 我们用分量法来进行证明如下：

$$\begin{aligned} \nabla \cdot (\mathbf{fg} \times \mathbf{r}) &= \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot f_k (\mathbf{g} \times \mathbf{r})_l \mathbf{e}_k \mathbf{e}_l = \frac{\partial}{\partial x_i} [f_i (\mathbf{g} \times \mathbf{r})_l] \mathbf{e}_l = \epsilon_{lkm} \frac{\partial}{\partial x_i} (f_i g_k r_m) \mathbf{e}_l \\ &= \epsilon_{lkm} \frac{\partial}{\partial x_i} (f_i g_k) r_m \mathbf{e}_l + \epsilon_{lkm} f_i g_k \frac{\partial}{\partial x_i} (r_m) \mathbf{e}_l \\ &= \epsilon_{lkm} [\nabla \cdot (\mathbf{fg})]_k r_m \mathbf{e}_l + \epsilon_{lkm} g_k (\mathbf{f} \cdot \nabla \mathbf{r})_m \mathbf{e}_l \\ &= [\nabla \cdot (\mathbf{fg})] \times \mathbf{r} + \mathbf{g} \times (\mathbf{f} \cdot \mathbf{I}) = [\nabla \cdot (\mathbf{fg})] \times \mathbf{r} + \mathbf{g} \times \mathbf{f} \end{aligned}$$

例题 1.11 证明

$$\nabla \cdot (\mathbf{fgh}) = (\nabla \cdot \mathbf{f}) \mathbf{gh} + (\mathbf{f} \cdot \nabla \mathbf{g}) \mathbf{h} + \mathbf{g} (\mathbf{f} \cdot \nabla \mathbf{h}).$$

证明 对于一些涉及到并矢的张量运算，我们也可以用下标法来做，依然按之前的三步来进行。

1. 首先，我们加上相应下标

$$\nabla \cdot (\mathbf{fgh}) = \nabla_f \cdot (\mathbf{fgh}) + \nabla_g \cdot (\mathbf{fgh}) + \nabla_h \cdot (\mathbf{fgh})$$

2. 利用向量与张量相关运算规则，将带下标的算符移到相应向量前。

$$\begin{aligned}\nabla_f \cdot (fgh) &= (\nabla_f \cdot f)gh \\ \nabla_g \cdot (fgh) &= (\nabla_g \cdot f)gh = (f \cdot \nabla_g)gh = (f \cdot \nabla_g g)h \\ \nabla_h \cdot (fgh) &= (\nabla_h \cdot f)gh = (f \cdot \nabla_h)gh = g(f \cdot \nabla_h h)\end{aligned}$$

3. 去掉下标，合并所有项。

$$\nabla \cdot (fgh) = (\nabla \cdot f)gh + (f \cdot \nabla g)h + g(f \cdot \nabla h)$$

1.4 狄拉克 δ 函数

物理学里的基本模型常常要用到点粒子，本课程也不例外。但是点粒子常常会涉及到一些无穷大的物理量，而这些物理量本身对一定大小的粒子也为有限值。例如，对于一质量为 m 的球形粒子，如果该球半径为 r ，则粒子的平均密度为有限值 $\rho = m/(4\pi r^3/3)$ 。可是对于位于空间坐标 \mathbf{x}_0 处，质量为 m 的点粒子，其密度在 $\mathbf{x} = \mathbf{x}_0$ 时为无穷大，在其它 $\mathbf{x} \neq \mathbf{x}_0$ 时为零。因此，我们有必要引入一类函数来表达点粒子相关的物理量。这类函数便是狄拉克 δ 函数²。

1.4.1 一维狄拉克 δ 函数

一维狄拉克 δ 函数的定义是

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0, \\ \infty, & \text{if } x = 0, \end{cases} \quad (1.28)$$

并且

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (1.29)$$

从上面的定义也可以看出，只要 $0 \in [a, b]$ ，我们便有

$$\int_a^b \delta(x) dx = 1. \quad (1.30)$$

如果 $f(x)$ 是一个一般的连续函数，则从 $\delta(x)$ 的定义中，我们很容易得出

$$f(x)\delta(x) = f(0)\delta(x). \quad (1.31)$$

所以，

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx = f(0) \int_{-\infty}^{\infty} \delta(x) dx = f(0). \quad (1.32)$$

如果无穷大的点不是在 $x = 0$ ，而是在 $x = a$ ，则对狄拉克 δ 函数进行平移，得到 $\delta(x - a)$ ，其定义类似为

$$\delta(x - a) = \begin{cases} 0, & \text{if } x \neq a, \\ \infty, & \text{if } x = a, \end{cases} \quad (1.33)$$

²狄拉克 δ 函数在物理学其它领域如强相互作用或量子力学里都有重要而广泛的应用。特别的，狄拉克是在量子力学里引入的这个函数，尽管该函数的定义可以追述到更早的工作中。

并且

$$\int_{-\infty}^{\infty} \delta(x-a) dx = 1. \quad (1.34)$$

显然我们也有

$$\int_{-\infty}^{\infty} f(x) \delta(x-a) dx = f(a). \quad (1.35)$$

严格地说, $\delta(x)$ 并不是一个普通函数, 因为在 $x=0$ 时, $\delta(x) = \infty$ 。但是我们也可以将其看作一些函数积分的极限情况。常见的例如

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2), \text{ 或 } \delta_n(x) = \frac{\sin(nx)}{\pi x}. \quad (1.36)$$

可以证明, 虽然对这两个函数而言, 当 $n \rightarrow \infty$ 时, δ_n 的极限并不存在, 但是它们都满足一个特点, 即

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(x) f(x) dx = \int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0). \quad (1.37)$$

这两个 $\delta_n(x)$ 的形式也表明狄拉克 δ 函数为偶函数, 即 $\delta(x) = \delta(-x)$ 。

狄拉克 δ 函数的这些定义显示, 对该函数相关的表达式而言, 我们通常都应该把它们放在积分符号下理解。例如, 两个含有 δ 函数的表达式, $D_1(x)$ 和 $D_2(x)$ 相等, 只有当对所有的常规函数 $f(x)$, 表达式

$$\int_{-\infty}^{\infty} f(x) D_1(x) dx = \int_{-\infty}^{\infty} f(x) D_2(x) dx,$$

都成立。例如, 一个关于狄拉克 δ 函数的重要性质是, 如果 $k \neq 0$,

$$\delta(kx) = \frac{1}{|k|} \delta(x). \quad (1.38)$$

而要证明此式, 我们仅需证明, 对任意的常规函数 $f(x)$,

$$\int_{-\infty}^{\infty} f(x) \delta(kx) dx = \frac{1}{|k|} \int_{-\infty}^{\infty} \delta(x) dx \quad (1.39)$$

这一式显然可以通过简单的积分变量替换得到。更广义一点, 如果狄拉克 δ 函数的变量是另一常规连续可导函数 $g(x)$, 则我们有

$$\delta[g(x)] = \sum_i \frac{\delta(x-x_i)}{|g'(x_i)|},$$

这里 x_i 是 $g(x) = 0$ 的根, 并且 $g'(x) \neq 0$ 。类似的道理, 也可以定义狄拉克 δ 函数的导数, $\delta'(x-a)$, 即

$$\int_{-\infty}^{\infty} \delta'(x-a) f(x) dx = - \int_{-\infty}^{\infty} f'(x) \delta(x-a) dx = -f'(a). \quad (1.40)$$

这里在第二步我们利用了分部积分。

1.4.2 三维狄拉克 δ 函数

三维狄拉克函数的定义为

$$\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z),$$

这里 $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$ 是位移矢量。从定义里可以看出,

$$\delta(\mathbf{r}) = \begin{cases} 0, & \text{if } \mathbf{r} \neq 0, \\ \infty, & \text{if } \mathbf{r} = 0, \end{cases}$$

且

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\mathbf{r}) dx dy dz = 1.$$

显而易见,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{r}) \delta(\mathbf{r} - \mathbf{a}) dx dy dz = f(\mathbf{a}).$$

例题 1.12 现在有一个电量为 q 的点电荷, 其位置为 $\mathbf{x}_q(t)$ 。请问其电荷密度如何表达?

解 注意到对任意大小的体积, ΔV , 点电荷密度满足

$$\int_{\Delta V} \rho dV' = \begin{cases} q, & \text{如果电荷位于 } \Delta V \text{ 内,} \\ 0, & \text{其它情况。} \end{cases}$$

对于 $\delta(\mathbf{x})$ 的定义, 很容易得出

$$\rho(\mathbf{x}) = q \delta(\mathbf{x} - \mathbf{x}_q(t)).$$

类似的道理, 如果我们有 N 个点粒子, 其中第 i 个粒子的电荷为 q_i , 且位于 \mathbf{x}_i 处, 则该粒子体系的密度分布可以写为

$$\rho(\mathbf{x}) = \sum_{i=1}^N q_i \delta(\mathbf{x} - \mathbf{x}_i).$$

注: 这个表达式在后续的章节里会经常用到。

第二章 狭义相对论

2.1 简要历史背景

2.1.1 电磁波的传播需要什么样的介质?

在 19 世纪末, 依据法拉弟与麦克斯韦等人的工作, 建立了电磁场理论, 从而电现象与磁现象被统一在优美的麦克斯韦方程组里。并且, 光作为该方程组的一种解, 被理解为一种电磁波动, 并得到赫兹等人实验上的证实。然而作为一种新的理论, 其中仍然存在两个明显的问题。一, 如果光是一种电磁波动的话, 它需要在什么样的介质中传播? 这个问题在今天看起来答案是清楚的, 然而当时已知的其它所有波动都是在介质中传播的, 例如声波。所以在面对当时新出现的事物(电磁波动)时, 思考其传播所需要的介质是一个非常自然的问题, 这也是科学探索中的一种常见过程。最终人们猜测可能存在一种叫以太的介质来承担电磁波传播, 而且这种以太的唯一作用似乎就是作为电磁波的传播介质。二, 麦克斯韦方程在伽利略 (Galileo) 变换下的参考系依赖性。在当时的已知力学相关物理过程都可以用牛顿定律来表达, 而不同惯性参考系里的力学规律在伽利略变换下具有形式上的不变性。假设我们有两个参考系, K' 和 K , 其中 K' 以速度 \mathbf{V} 相对 K 运动。则空间坐标在伽利略变换下的规律为

$$\mathbf{x} = \mathbf{x}' + \mathbf{V}t', \quad (2.1)$$

$$t = t'. \quad (2.2)$$

这里 \mathbf{x}, t 为 K 系中的物理量, 而相应的带撇的量为 K' 系中的物理量。对于力而言, $\mathbf{F} = m\mathbf{a}$ 在不同的惯性参考系间不改变, 即 $\mathbf{F} = \mathbf{F}'$ 。这样如果一个质点的运动规律在 K 系中表达为

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}, \quad (2.3)$$

$$\frac{d\mathbf{v}}{dt} = \mathbf{a}. \quad (2.4)$$

则在 K' 系中, 依伽利略变换给出的运动规律为

$$\frac{d\mathbf{x}'}{dt'} = \mathbf{v}' = \mathbf{v} - \mathbf{V} \equiv \mathbf{v}', \quad (2.5)$$

$$\frac{d\mathbf{v}'}{dt'} = \mathbf{a}' = \mathbf{a}. \quad (2.6)$$

这样不论是在 K 系还是 K' 系中。如果我们测量各自质量的受力得到加速度, 从而得到速度与位移, 则该质点在不同惯性参考系间满足的物理规律都具有相同的形式。换一种说法是, 我们在一惯性参考系内, 没有办法用力学规律去定出我们到底处在哪一个惯性系, 即所有的惯性参考系都是等价的。需要注意的是, 上式同时也定义了不同参考系间的速度变换, 即 $\mathbf{v} = \mathbf{v}' + \mathbf{V}$ 。另外, 时间独立于不同的惯性参考系。因此在牛顿力学与伽利略变换的理论框架下, 时间是绝对的, 通常没有必要去区分 t 与 t' 。

然而用于描述真空中电磁波传播的方程具有波动方程的基本形式, 即

$$\left(\frac{1}{\mu_0 \epsilon_0} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \psi = 0. \quad (2.7)$$

这里 $\nabla^2 = \nabla \cdot \nabla$, 且 ψ 为代表电磁场的任一分量。很容易证明波动方程的形式在伽利略变换下, 在不同的惯性参考系间的形式不同。这一点在力学问题里的波动上也是一样的。例如, 在相对介质静止的参考系中, 声波的基本方程为

$$\left(\frac{1}{c_s^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \delta p = 0. \quad (2.8)$$

这里 c_s 为声音传播的速度, 而 δp 为压强的扰动量。如果换到另一惯性参考系中, 声波的方程不再具有上述形式。然而这种形式上的不一样, 对声波而言是符合其规律的¹, 它并没有改变更为根本的牛顿力学规律。

¹例如, 声波的基本色散关系是在相对于介质静止的参考系里才成立的。

2.1.2 麦克逊-莫雷的实验与洛伦兹的解释

回到电磁现象上,当时多位实验学家试图找到以太存在的实验证据,从而更完美地理解电磁规律。其中最为著名的莫过于麦克逊-莫雷 (Michelson-Morley) 实验,用于测量地球相对于静止以太运动的实验。然而实验结果是否定的—当相对于假设的以太具有不同的运动方向时,该实验并没有探测到光会具有不同的速度。所以以太的存在与否受到了实验上的挑战。于此同时, FitzGerald 与 Lorentz 为了在以太的理论框架下,解释该实验结果,提出了长度收缩的概念,即物体的长度在沿运动方向上会发生收缩。后续经过洛伦兹与庞加莱²等人一系列的发展,提出了现在我们所熟知的洛伦兹变换。该变换可以解释一系列相关的实验结果,且可以证明该变换可以让麦克斯韦方程的形式在不同参考系间保持不变。但需要注意的是,到这一步,该变换是用来解释以太探测不到的实验结果的,并没有否定以太的存在。

麦克逊-莫雷的实验结果,以及作为描述其基本规律的麦克斯韦方程在不同参考系间在伽利略变换下具有形式上的改变,这意味着什么?一种可能是,我们可以利用电磁规律来判断我们所处的惯性参考系,因为毕竟电磁现象不是力学。第二种可能是,麦克斯韦方程不是描述电磁现象最为基本的方程,存在着更为根本的电磁规律,在伽利略变换下,具有形式上的不变性。第三种可能是,物理学并不需要一个实验测量不到的以太去担当电磁波传波的介质,而是需要重新去审视之前的物理理论,包括牛顿力学规律与伽利略变换。

2.1.3 《论动体的电动力学》

在 1905 年,爱因斯坦发表了四篇极为重要的论文³,其中一篇为《论动体的电动力学》⁴。在这篇文章里,爱因斯坦思考了在不同惯性参考系里同一过程不等价。一个与这篇论文等价的例子如下。我们考虑一个导线以及其周围一个运动的电荷。在导线静止的参考系里,电荷受到一个磁场力 $q\mathbf{v} \times \mathbf{B}$,这个磁场是由导线的电流产生的。因此这一现象为磁场对电荷的作用。然而,如果我们选择运动的电荷为参考系,则该电荷不受任何磁场力,因为在该参考系里,该电荷是静止的。但可以证明,其依然受力,而且受到的是电场力。同样的一个物理过程,其解释为电现象还是磁现象依赖于所选择的参考系。基于这一问题,爱因斯坦提出了两条原理,即相对性原理与光速不变性原理。仅从这两条原理出发,通过构建两个惯性参考系里的光信号时钟,爱因斯坦推导出了在不同惯性参考系间空间坐标与时间的变换关系,即洛伦兹变换。同时从洛伦兹变换出发,爱因斯坦得到了两个不同的惯性参考系里麦克斯韦方程的形式,并进一步依据相对性原理,通过要求物理规律在不同惯性参考系里不改变,得到了电场与磁场在不同参考系中的关系。虽然部分内容(如相对性原理,光速不变性,洛伦兹变换等)与前人结果一致,但这一工作的重要不同点是其仅从相对性原理与光速不变性出发,并抛弃了实验中测量不到的以太的概念,所得结果极大地改变了人们对于时间与空间的理解。

2.2 狭义相对论基础

2.2.1 相对性原理与光速不变性

在牛顿力学里,伽利略相对性原理指的是所有的力学规律在所有的惯性参考系里都有相同的形式。这里惯性参考系被定义为时间均匀,空间均匀且各向同性的参考系。在这样的参考系里,一个不受任何外力的粒子,因为时空均匀,其运动自然是匀速的(牛顿第一定律),且物理规律具有最简洁的形式。本课程里的参考系,除非特别指明,否则都是惯性参考系。

²Henri Poincaré, 法国数学家与物理学家。在相对论的创立过程中作出了重要贡献。在他 1898 年的文章 (The Measure of Time), 他提出为了让物理规律具有最简洁的形式, 光速应当是不变的。在他 1900 年的文章中, 讨论过相对运动, 并在 1904 年的文章里提出过“相对性原理”——没有任何物理实验可以区分两个惯性参考系; 在他 1905 年的文章 (On the Dynamics of the Electron) 中, 提出引力的作用并不是瞬时的, 而是以光速传播。也意识到 $x^2 + y^2 + z^2 - c^2t^2$ 为一个不同惯性参考系间的不变量, 以及洛伦兹变换对应着四维时空的旋转——第四维度为 ict 。

³其它三篇分别为: 布朗运动, 光电效应, 以及 $E = mc^2$ 。物理学界也把这一年称为“爱因斯坦的奇迹年”。

⁴原论文为: Einstein, A. (1905), Zur Elektrodynamik bewegter Körper. Ann. Phys., 322: 891-921. 英文翻译的版本标题一般为: “On the Electrodynamics of Moving Bodies”。

狭义相对论将伽利略不变性原理从力学扩大到所有的自然规律，以下两条作为基本假定 (postulates)：

1. 相对性原理：所有的物理规律在所有的惯性参考系里都是等价的。
2. 光速不变性：真空中的光速是一个常数，与光源是否运动无关，其数值为 $c = 2.998 \times 10^8 \text{ m/s}$ 。

我们没有办法有任何物理实验去定出我们所在的惯性参考系。而光速不变性则显然与伽利略变换以及时间的绝对性会有矛盾。因为时间的绝对性，伽利略速度变换为

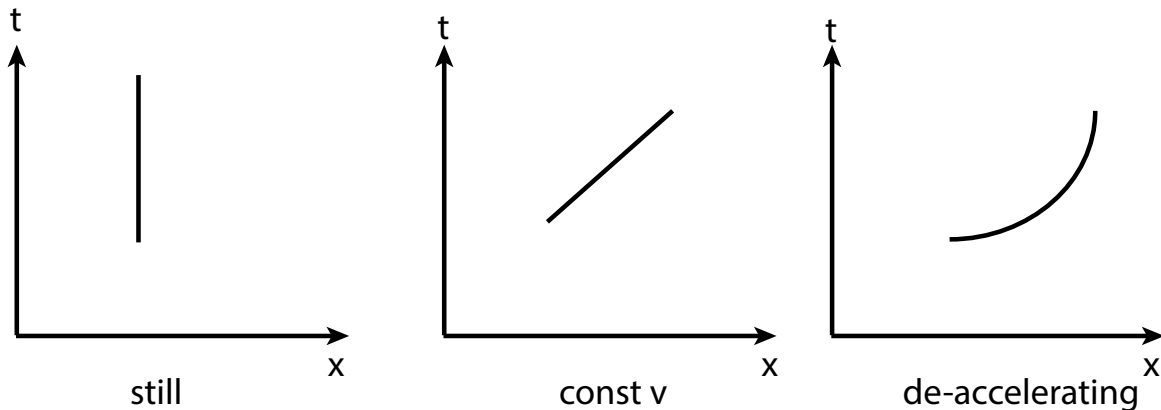
$$\mathbf{v} = \mathbf{v}' + \mathbf{V}. \quad (2.9)$$

这会破坏光速的不变性。除了麦克逊-莫雷实验外，后续也有一系列实验去验证光速不变性原理，在目前的测量精度范围内，并没有发现这条假定是错误的。

2.2.2 四维时空与时空变换的线性性

爱因斯坦在他的论文里通过构建一个以光为信号的时钟，以及不同惯性参考系间的等价性来得到满足以上两条假定的时空坐标变换关系。在本讲义里，我们通过另外一种办法来得到这种变换。首先我们引入几个基本概念。第一个概念是事件 (event)。一个事件由它发生的时间 (t) 和空间位置 (\mathbf{x}) 所定义。在相对论课程里，我们会经常讨论一个事件的时空坐标在不同惯性参考系下如何改变，并且事件还可以被用来定义许多其它的物理量。例如，假定我们有一个参考系 K ，里面有一个静止的时钟。则该时钟所测量的时间间隔可以利用两个事件来定义，(t_1, \mathbf{x}) 和 (t_2, \mathbf{x})。这里 t_1 是第一次读数的时间， t_2 是第二次读数的时间。因为该时钟静止，所以其两次测量时，空间坐标不变，均为 \mathbf{x} 。时钟所测量的时间间隔为 $t_2 - t_1$ 。又如，在 K 参考系，有一辆运动的车子，则可以通过两个事件来得到该车子的长度：分别为测量车头与车尾所在的时间与空间坐标，记为 (t_1, x_1), (t_2, x_2)。因为我们想测量长度的车子是运动的，所以需要在确保两个事件的时间坐标一致 $t_1 = t_2$ 的情况下，通过 $l = x_2 - x_1$ 来计算其长度。这一点是后续计算尺缩效应的基础。这两个例子中的结论看起来毫不意外，但理解事件的概念以及同时的相对性，是理解相对论里很多看似奇怪的结论的基本出发点。

第二个需要引入的概念是四维时空 (spacetime)，即一个以时间和三维空间构成的时空。这个四维时空的具体性质后续再详细讨论。通常为了四个维度有相同的量纲，我们用 ct 作为第四维度⁵。有了该四维时空，每一个事件的时空坐标便是该四维时空中的一个点，称为世界点 (world point)。而任何一个粒子在这样的四维时空里的轨迹都会是一条线，称为世界线 (world line)。例如，一个静止的粒子在这个四维空间里的世界线是一条平行于时间轴的直线。又或者，一个自由粒子，其世界线是一条斜线，而一个有加速度的粒子，其世界线是一条曲线。如图所示，



有了事件与四维时空的基本概念，我们可以得到第一个关于狭义相对论的重要结论：在两个不同的惯性参考系间的时空坐标变换一定是一个线性变换。这一结论从物理上不难理解。如果我们有二个以速度 V 相对运动的

⁵涉及到相对论时，一种简单的处理是让 $c = 1$ ，这样可以避免掉式子里出现的许多 c 。而为了回到 $c \neq 1$ 的式子，只需要按照相关的量纲补上 c 。例如，在 $c = 1$ 时，相对论因子 $\gamma = 1/\sqrt{1-v^2}$ 。要回到 $c \neq 1$ 的式子，我们只需要将 $v \rightarrow v/c$ ，使得其量纲与其它项一致即可，即 $\gamma = 1/\sqrt{1-(v/c)^2}$ 。

惯性参考系 K 与 K' 。现考虑任一 K 系中不受外力的质点，其在 K 系对应的四维时空 (t, x, y, z) 中的世界线一定是一条直线，对应在上图中，该直线的斜率为 $(v/c)^{-1}$ ，这里 v 为质点在 K 系中的速度。同时在 K' 系对应的四维时空 (t', x', y', z') 中的世界线也一定是一条直线，其斜率为 $(v'/c)^{-1}$ 。反过来该结论也同样成立，即在 K 系中的一条直线，一定也对应着 K' 系中的一条直线。因此，我们自然期待作为联系两个参考系时空坐标的变换应该是线性的。

2.2.3 时空间隔及其不变性

定义事件，我们可以进一步引入一个相对论中非常重要的概念：时空间隔 (spacetime interval)。如果我们有两个事件，其坐标分别是 (t_1, x_1, y_1, z_1) 与 (t_2, x_2, y_2, z_2) 。则我们可以定义这两个事件的时空间隔 Δs 为⁶。

$$\Delta s = (c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2)^{1/2} = (c^2 \Delta t^2 - \Delta \mathbf{x}^2)^{1/2}, \quad (2.10)$$

这里 $\Delta t = t_2 - t_1$, $\Delta x = x_2 - x_1$, 等等。

例题 2.1 假设在一个惯性参考系 K 中，有如下两个事件：

- 事件 1：在 t_1 时刻从位置 (x_1, y_1, z_1) 发出一束光。
- 事件 2：在 t_2 时刻从位置 (x_2, y_2, z_2) 接收到光。

请计算这两个事件的时空间隔。

解 因为光信号的速度为 c ，所以 $\Delta \mathbf{x}^2 = c^2 \Delta t^2$ 。这两个事件的时空间隔为

$$\Delta s^2 = c^2 \Delta t^2 - \Delta \mathbf{x}^2 = 0. \quad (2.11)$$

例题 2.2 我们有另一个惯性参考系 K' ，则上述两个事件在其中的时空间隔是多少？

解 因为时间不再是一个绝对不变的独立于参考系的存在，所以在 K' 中，两个事件的时空坐标为

- 事件 1： t'_1 与 \mathbf{x}'_1 ,
- 事件 2： t'_2 与 \mathbf{x}'_2 。

我们注意到相对论的第二条假设：光速不变性。所以在 K' 中，依然有 $\Delta \mathbf{x}'^2 = c^2 \Delta t'^2$ ，即在 K' 中，

$$\Delta s'^2 = c^2 \Delta t'^2 - \Delta \mathbf{x}'^2 = 0. \quad (2.12)$$

以上两个例子告诉我们，用光信号连接的两个事件，其时空间隔为 0，且在所有的惯性参考系里，都是如此。反过来也是一样，因为光速不变性，如果在一个惯性参考系 K 里， $\Delta s = 0$ ，则在任何其它惯性参考系 K' 里， $\Delta s' = 0$ 。一个自然而然的问题是，上面的例子的结论是否可以推广到 $\Delta s \neq 0$ 的情况？即 $\Delta s = \Delta s'$ 是否对任意的时空间隔都成立，而不仅是 $\Delta s = 0$ ？初看起来，答案似乎是不一定的，因为一般的事件不是用光信号联系的。

为了找到一般情况下， Δs 与 $\Delta s'$ 的关系，我们考虑任意两个事件在不同惯性参考系里的时空间隔⁷。分别为

$$\Delta s^2 = c^2 \Delta t^2 - \Delta \mathbf{x}^2 \text{ 与 } \Delta s'^2 = c^2 \Delta t'^2 - \Delta \mathbf{x}'^2. \quad (2.13)$$

注意到时空坐标变换是线性的，一般情形下，时空间隔 $\Delta s'^2$ 可以用 K 系的量表达为

$$\Delta s'^2 = \sum_{\alpha=0}^3 \sum_{\beta=0}^3 A_{\alpha\beta} \Delta x_{\alpha} \Delta x_{\beta}. \quad (2.14)$$

这里我们用希腊字母 α, β 表达从 0 到 3。其中 0 对应时间 t (即 $\Delta x_0 = c \Delta t$)，而用 i, j 等下标表达 1, 2, 3，对应着空间坐标。同时，因为我们还没有引入后文的四维矢量，为了区分，我们此处不用爱因斯坦求和法则表达，而直接写出求和符号。注意到方程右边的所有项都是以两个坐标差的关系，而因为是惯性系，改变坐标原点的选择不会改变其结果，因此这里的系数 $A_{\alpha\beta}$ 一定不依赖于具体的时空坐标，只能是两个参考系间相对运动的速度

⁶注意一下，不同的书籍或文章里时空间隔的定义可能会稍有差别，比如 $\Delta s^2 = \Delta \mathbf{x}^2 - c^2 t^2$

⁷这里的证明参考的是 Bernard Schultz 的 “A First Course in General Relativity”。

的函数, 即 $A_{\alpha\beta} = A_{\alpha\beta}(\mathbf{V})$ 。此外, $\Delta x^\alpha \Delta x^\beta$ 的项是对称的, 意味着我们可以选择 $A_{\alpha\beta}$ 为对称的, 这会简化我们的分析。

接下来, 我们利用当 $\Delta s = 0$ 时, $\Delta s' = 0$ 这个条件对任意的两个以光信号连接的事件都成立的条件来进一步的限制上式中相关系数的性质。首先, 我们有当 $\alpha \neq \beta$ 时 $A_{\alpha\beta} = 0$ 。为此, 我们选取两组以光信号连接的事件, 第一组的空间坐标差为 Δx_i , 而另一组则为 $-\Delta x_i$ 。这样我们有

$$0 = A_{00}\Delta t^2 + \sum_{\beta} A_{0\beta}\Delta t\Delta x_{\beta} + A_{ij}\Delta x_i\Delta x_j, \quad (2.15)$$

$$0 = A_{00}\Delta t^2 - \sum_{\beta} A_{0\beta}\Delta t\Delta x_{\beta} + A_{ij}\Delta x_i\Delta x_j. \quad (2.16)$$

将上面两个式子相减, 且注意到结论对任意的 Δx_{β} 都成立, 我们有 $A_{0\beta} = 0$ 。进一步的, 我们运用同样的方法, 分别依次改变 $\Delta x, \Delta y, \Delta z$ 的符号, 则会得到当 $i \neq j$ 时, 所有的 $A_{ij} = 0$ 。这样方程 (2.14) 可以写为

$$\Delta s'^2 = A_{00}c^2\Delta t^2 + A_{11}\Delta x^2 + A_{22}\Delta y^2 + A_{33}\Delta z^2. \quad (2.17)$$

其次, 当 $\Delta s^2 = 0$ 时, $c^2\Delta t^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$, 此时

$$\Delta s'^2 = (A_{00} + A_{11})\Delta x^2 + (A_{00} + A_{22})\Delta y^2 + (A_{00} + A_{33})\Delta z^2 = 0. \quad (2.18)$$

因为上一式对任意的用光信号连接的两个事件都要成立, 因此我们一定有 $A_{00} = -A_{11} = -A_{22} = -A_{33}$, 即

$$\Delta s'^2 = A(\mathbf{V})[c^2\Delta t^2 - (\Delta x^2 + \Delta y^2 + \Delta z^2)] = A(\mathbf{V})\Delta s^2. \quad (2.19)$$

这里 $A(\mathbf{V}) = A_{00}$ 。最后, 因为惯性参考系的空间各向同性, 而 Δs 本身和坐标系方向的选取也是没有关系的, 所以 $A(\mathbf{V})$ 不能和速度方向有关, 只能和速度大小有关, 因此最终我们有

$$\Delta s' = A(V)\Delta s. \quad (2.20)$$

类似的, 我们也有 $\Delta s = A(V)\Delta s'$ 。如果 $A^2 = 1$ 或者 $A = \pm 1$ 。再依据 $V = 0$ 时, 两参考系相同, 我们选择 $A = 1$ 。这样我们从光速不变性与惯性参考系的时空均匀性, 得出一个非常重要的结论: 两个事件的时空间隔在不同的惯性参考系间是不变的。作为两条假定的直接结论, 所有与狭义相对论相关的物理关系都需要满足时空间隔的不变性。

例题 2.3 从时空间隔的不变性, 证明如果一个粒子在一个惯性参考系里是以低于光速运动的, 则在所有其它的惯性参考系里其速度都小于光速。

证明 如果一个粒子在一个惯性参考系里以低于光速运动, 则我们选取其世界线的一个相邻两点, 并计算其时空间隔, $ds^2 = c^2 dt^2 - d\mathbf{x}^2 = (c^2 - |\mathbf{v}|^2)dt^2$ 。这里 $\mathbf{v} = d\mathbf{x}/dt$ 。因为 $v < c$, 则 $\Delta s^2 > 0$ 。因为时空间隔的不变性, 对于任意一个惯性参考系, 我们都有 $\Delta s'^2$, 因此在任意一个惯性参考系里, 该粒子都是以小于光速运动的。

2.2.4 洛伦兹变换

通过之前的讨论, 我们已经知道伽利略变换与光速不变性相矛盾, 所以显然它也与时空间隔不变性相矛盾。同时我们现在也知道时间与空间一样, 依赖于具体的参考系。假设我们有两个以某个速度相对运动的惯性参考系, K 与 K' , 那么一个事件的时空坐标 (t, \mathbf{x}) 可以通过什么样的线性变换得到另一个参考系里的坐标 (t', \mathbf{x}') , 能使得两事件的时空间隔保持不变⁸?

首先, 可以保持时空间隔不变的线性坐标变换的可能性有多种, 例如空间旋转变换, 平移变换等。但不是每一种变换都可以对应到两个相对运动的惯性参考系间的时空坐标变换⁹。我们先举一个可以保持时空间隔不变, 但只涉及到空间坐标变换的例子。

⁸本节的推导方法来自于 Landau 与 Lifshiz 所著的《理论物理从书第二卷: 经典场论》

⁹按习惯可以简单分为两类。一类是保持任一事件 (ct, \mathbf{x}) 与坐标原点 $(0, \mathbf{0})$ 之间的时空间隔 $(s^2 = c^2 t^2 - \mathbf{x}^2)$ 不变的坐标变换。所有的这一类变换构成了齐次洛伦兹群 (the homogeneous Lorentz group), 包括空间旋转与本课程里介绍的齐次洛伦兹变换。一类是保持任意两个事件之间的时空间隔 $(\Delta s^2 = c^2 \Delta t^2 - \Delta \mathbf{x}^2)$ 不变的坐标变换, 所有的这一类变换构成了庞加莱群 (the Poincaré group), 也称非齐次洛伦兹群 (the inhomogeneous Lorentz group), 包括平移变换以及齐次洛伦兹群里的坐标变换。显然满足后者的变换一定满足前者。

例题 2.4 如果我们仅改变两个空间坐标 $(x, y) \rightarrow (x', y')$, $t = t', z = z'$ 。请找到合适的坐标变换, 使得时空间隔

$$s^2 = c^2 t^2 - (x^2 + y^2 + z^2) \quad (2.21)$$

不改变。

解 依问题, 显然我们仅需要 $r^2 \equiv x^2 + y^2 = x'^2 + y'^2$ 不变, 即任意一点, 其到原点的距离不变。我们可以利用如下方式找到两组坐标之间的变换关系。因为 $\cos^2 \theta + \sin^2 \theta = 1$, 所以我们可以令

$$x = r \cos \theta, \quad (2.22)$$

$$y = r \sin \theta, \quad (2.23)$$

以及

$$x' = r \cos(\theta + \phi), \quad (2.24)$$

$$y' = r \sin(\theta + \phi), \quad (2.25)$$

这样, $x^2 + y^2 = r^2 = x'^2 + y'^2$ 恒成立, 且 (x, y) 与 (x', y') 不必相同。利用三角函数的和差化积公式,

$$x' = r \cos \theta \cos \phi - r \sin \theta \sin \phi, \quad (2.26)$$

$$y' = r \sin \theta \cos \phi + r \cos \theta \sin \phi. \quad (2.27)$$

再次利用 $x = r \cos \theta$ 与 $y = r \sin \theta$, 我们有

$$x' = x \cos \phi - y \sin \phi, \quad (2.28)$$

$$y' = y \cos \phi + x \sin \phi. \quad (2.29)$$

这样就得到两个坐标系间的坐标关系, 该变换关系可以保持相应的时空间隔不改变。显然, 这个坐标变换对应着普通的空间旋转, ϕ 角的含义显然是与空间转动角度有关。该变换无法描述两个以给定速度相对运动的惯性参考系之间的时空坐标变换。

其次, 如果我们同时改变时间与空间, 能否找到一个坐标变换, 使得时空间隔 $s^2 = c^2 t^2 - x^2$ 不改变? 为简单起见, 我们考虑只改变 t 与 x 坐标, 而保持 y, z 不变的坐标变换, 即 $(t, x) \rightarrow (t', x')$, 且 $y = y', z = z'$ 。显然此时, 我们仅需要考虑保持

$$c^2 t^2 - x^2 = c^2 t'^2 - x'^2 = s^2. \quad (2.30)$$

采用与上例的相同方式, 我们利用双曲函数 (\cosh, \sinh, \dots) 的特点, 即对任意的 θ , 都有

$$\cosh^2 \theta - \sinh^2 \theta = 1, \quad (2.31)$$

所以我们可以令

$$ct = s \cosh \theta,$$

$$x = s \sinh \theta.$$

且

$$ct' = s \cosh \theta',$$

$$x' = s \sinh \theta'.$$

与空间旋转类似，利用双曲函数对 t 和 x 参量化，让我们保证时空间隔的不变。接下来我们继续通过双曲函数的和差化积来得到相应坐标的联系。假设 $\theta = \theta' + \phi$ ，注意到 $ct' = s \cosh \theta'$ 和 $x' = s \sinh \theta'$ ，我们有

$$\begin{aligned} ct &= s \cosh(\theta' + \phi) = s \cosh \theta' \cosh \phi + s \sinh \theta' \sinh \phi \\ &= ct' \cosh \phi + x' \sinh \phi, \end{aligned}$$

$$\begin{aligned} x &= s \sinh(\theta' + \phi) = s \sinh \theta' \cosh \phi + s \cosh \theta' \sinh \phi \\ &= x' \cosh \phi + ct' \sinh \phi. \end{aligned}$$

重写方程为

$$x = x' \cosh \phi + ct' \sinh \phi, \quad (2.32)$$

$$ct = x' \sinh \phi + ct' \cosh \phi. \quad (2.33)$$

这些即是我们所需要的坐标变换方程，与例 2.4 中的空间坐标旋转类似， ϕ 在这里也是“旋转”角度，这种旋转被称为“双曲旋转”(hyperbolic rotation)。

进一步，以上的“双曲旋转”能否对应到两个以相对速度运动的两个惯性参考系 K 与 K' 间的时空坐标变换？为此，我们考虑一个特殊的点： $x' = 0, y' = 0, z' = 0$ 。该点对应的 (t, \mathbf{x}) 坐标显然为 $x = ct' \sinh \phi$ 与 $ct = ct' \cosh \phi$ ，且 $y = 0, z = 0$ 。因此， $x/ct = \sinh \phi / \cosh \phi$ 。如果我们将 (t, \mathbf{x}) 与 (t', \mathbf{x}') 分别理解为 K 系与 K' 系的时空坐标，则 $\mathbf{x}' = \mathbf{0}$ 点，代表了 K' 系的空间坐标原点。而相应的 x/ct 则代表了该原点在 K 系中沿 x 方向的运动速度（注意到 $y = y' = 0, z = z' = 0$ 不变）。因此，如果我们令

$$\sinh \phi / \cosh \phi = V/c = (e^\phi - e^{-\phi}) / (e^\phi + e^{-\phi}) \quad (2.34)$$

则解出

$$e^\phi = \sqrt{\frac{V+c}{V-c}}. \quad (2.35)$$

从而

$$\sinh \phi = \frac{e^\phi - e^{-\phi}}{2} = \frac{V/c}{\sqrt{1 - V^2/c^2}} \equiv \beta(V) \gamma(V) \quad (2.36)$$

$$\cosh \phi = \frac{e^\phi + e^{-\phi}}{2} = \frac{1}{\sqrt{1 - V^2/c^2}} \equiv \gamma(V) \quad (2.37)$$

这里 $\beta(V) \equiv V/c$ 且 $\gamma(V) = 1/\sqrt{1 - \beta^2}$ 。把 $\sinh \phi$ 和 $\cosh \phi$ 的表达式代入坐标变换方程里，同时考虑到 $y = y', z = z'$ ，我们有

$$t = \gamma(t' + \beta x'/c),$$

$$x = \gamma(x' + \beta ct'),$$

$$y = y',$$

$$z = z'.$$

这一变换也可以简单的用矩阵表达，

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix}.$$

因此包含 (t, x) 坐标的双曲旋转可以保持时空间隔不变, 且对应着相个沿 x 方向相对运动的两个惯性参考系间的时空坐标变换。注意到 $x' = 0, t' = 0$ 对应着 $x = 0, t = 0$, 即两个惯性参考系的时空坐标原点重合, 相应的变换称为齐次洛伦兹变换 (homogeneous Lorentz transformation)。通常在不会歧义的情况下, 也即称为洛伦兹变换¹⁰。类似的, $\mathbf{V} = V\mathbf{e}_y$ 的坐标变换是四维时空的 $y-t$ 平面的双曲旋转, 而 $\mathbf{V} = V\mathbf{e}_z$ 的坐标变换则是 $t-z$ 平面的双曲旋转。时空坐标的反变换既可以通过求解洛伦兹变换方程得到, 也可以利用 K 系相对 K' 系以速度 $-\mathbf{V}$ 运动得到。因此

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}.$$

显然, 如果 $c \rightarrow \infty$ 时 (非相对论情况), 则 $\gamma \rightarrow 1$, 这时洛伦兹变换近似为伽利略变换, $x = x' + Vt'$, 及 $t = t'$ 。但需要注意的是, 洛伦兹变换与伽利略变换有本质上的不同: 前者可以保持时空间隔不变, 符合狭义相对论的两条假定, 但后者不能。另一个不同点是, 双曲旋转与一般的旋转类似, 除非有着相同的转轴, 否则两次旋转的最终结果一般依赖于旋转的先后次序, 所以一般情况下洛伦兹变换不具有可对易性。但两次伽利略变换的结果并不依赖于变换的先后秩序, 具有对易性。

2.2.5 洛伦兹速度变换

从洛伦兹变换很容易得到相应的速度变换, 且可以直接证明洛伦兹变换符合光速不变性。假如一个粒子在 K 系里, 速度是 \mathbf{v} ; 而在 K' 系里是 \mathbf{v}' 。假如 K' 以速度 $\mathbf{V} = V\mathbf{e}_x$ 相对于 K 运动, 则从洛伦兹变换可以得到,

$$dt = \gamma(dt' + \beta dx'/c), \quad (2.38)$$

$$dx = \gamma(dx' + \beta c dt'), \quad (2.39)$$

$$dy = dy', \quad (2.40)$$

$$dz = dz'. \quad (2.41)$$

这样利用 $\mathbf{v} \equiv d\mathbf{x}/dt$ 以及 $\mathbf{v}' \equiv d\mathbf{x}'/dt'$, 我们得到

$$v_x = \frac{dx' + \beta c dt'}{dt' + \beta dx'/c} = \frac{v'_x + \beta c}{1 + \beta v'_x/c}, \quad (2.42)$$

$$v_y = \frac{v'_y/\gamma}{1 + \beta v'_x/c}, \quad (2.43)$$

$$v_z = \frac{v'_z/\gamma}{1 + \beta v'_x/c}. \quad (2.44)$$

该组速度变换显然可以保持光速不变, 例如 $v'_x = c, v'_y = 0, v'_z = 0$, 则可得到 $v_x = c, v_y = 0, v_z = 0$ 。另外, 在 $c \rightarrow \infty$ 时, 该组速度变换成为伽利略变换。

2.2.6 同时性的相对性与因果律

狭义相对论里, 时间失去了绝对性, 而是与空间坐标类似, 时间依赖于所在的参考系。两个事件发生的先后顺序也可能依赖于所在参考系。为了更好地解释这一现象, 我们先介绍时空间隔的分类。

两个事件间的时空间隔 Δs^2 定义为 $c^2 \Delta t^2 - \Delta \mathbf{x}^2$, 因此依据 $c^2 t^2$ 与 $\Delta \mathbf{x}^2$ 的相对大小关系, Δs^2 可正可负。据此, 一般将时空间隔分为类光 (lightlike) 间隔, 类空 (spacelike) 间隔和类时 (timelike) 间隔。类光间隔即用光信

¹⁰ 仅对应于一个惯性参考系相对另一惯性参考系沿某一方向运动的洛伦兹变换也被称为 Lorentz boost, 例如本课程中的变换。但 boost 一词在中文里我还没有找到比较好的对应翻译。如果时空变换还包括了坐标的平动 (translation), 则相应的变换称为庞加莱变换或非齐次洛伦兹变换。

号联系的两个事件间的间隔，其 $\Delta s^2 = 0$ ；而对类空间隔，则有 $\Delta s^2 < 0$ 或 $c^2\Delta t^2 < \Delta x^2$ ；最后，如果 $\Delta s^2 > 0$ 或 $c^2\Delta t^2 > \Delta x^2$ ，则该间隔为类时间隔。注意到，时空间隔是一个不依赖于具体惯性参考系的物理量，所以两个事件间的间隔类型也是不依赖于具体的惯性参考系的。

类空间隔的一个重要特点是，其对应的两个事件发生的同时性是不确定的，依赖于所选参考系。例如，在一个参考系里，两个事件不是同时发生的， $\Delta t > 0$ ，但是它们之间空间上相隔太远， $|\Delta x| > c\Delta t$ ，所以 $\Delta s^2 < 0$ ，为类空间隔。这样对任一参考系 K' 而言，我们也都有 $\Delta s'^2 < 0$ 。所以，我们可能会找到一个参考系， K' ，在该参考系里， $\Delta t' = 0$ 而不改变时空间隔，即

$$\Delta s'^2 = -\Delta x'^2 < 0. \quad (2.45)$$

具体的，利用洛伦兹变换¹¹，只要我们选择 K' 系相对于 K 系的速度 V 使其满足 $\Delta t = (\beta/c)\Delta x$ ，则 $\Delta t' = 0$ 。而因为 $c\Delta t < \Delta x$ ， $\beta < 1$ ，因此这样的参考系 K' 总是存在的。类似的，也可以找到一个参考系 K' ， $\Delta t' < 0$ ，即两事件的先后发生顺序在 K' 与 K 系中恰好相反。因此，类空间隔联系的两个事件，其发生的同时性与先后顺序依赖于所选择的参考系。很多与狭义相对论相关的佯谬都是因为所讨论的两个事件的时空间隔是类空型的。

但对类时间隔而言，情况则有所不同。同样我们假设， $\Delta t > 0$ ，但是此时 $\Delta s^2 > 0$ 。问题是我們有没有可能找到一个参考系，使得 $\Delta t' = 0$ ？答案显然是不可能的，否则此时 $\Delta s'^2 = -\Delta x'^2 < 0$ ，违背时空间隔的不变性¹²。进一步的，利用洛伦兹变换，

$$\Delta t' \propto c\Delta t - \beta\Delta x = c\Delta t[1 - \beta(\Delta x/c\Delta t)]. \quad (2.46)$$

因为 $\beta < 1$ ，且 $\Delta x < c\Delta t$ ，对类时间隔事件而言，在所有的惯性参考系里，两事件的时间差 $\Delta t'$ 与 Δt 具有相同的符号，或者说，其发生的先后顺序不依赖于所选择的惯性参考系。

由类时间隔事件先后顺序的绝对性，我们可以推出有因果关系的两个事件之间一定是类时间隔联系的以及作用 (interaction) 传递的最大速度为光速。因为如果事件 1 是事件 2 发生的原因，则不管在哪个参考系观察，因果律决定了一定是先有因（事件 1），后有果（事件 2）。而两事件的先后顺序不依赖于参考系，则其间隔一定是类时的。同时事件 1 的作用传递，导致事件 2 的发生，而作为类时间隔， $|\Delta x|/\Delta t < c$ ，因此，我们也可推出作用传递的最大速度为光速。

2.2.7 固有时

现在我们介绍相对论里的一个重要概念，固有时 (proper time)，指的是相对于一个运动物体或参考系静止的时钟所给出的时间 (τ)。在该参考系里，该固定时钟的两次测量事件之间的时空间隔为 $\Delta s = c\Delta\tau$ ，因为相应的 $\Delta x = 0$ 。所以从时空间隔的不变性，也可以看出固有时的不变性，因为 $\Delta\tau = \Delta s/c$ 。换个角度看，这个结论也是显然的，因为 K' 参考系的固有时，不管在哪个参考系看，都还是 K' 参考系的固有时，不会改变。

现在我们从时空间隔不变的角度，推导运动的时钟变缓这一结论。假设现在有一物体相对于 K 系以速度 v 运动，有一个时钟相对于该物体静止。在该物体静止的参考系 K' 里，该时钟两次测量的事件，都发生在同一位置，其时间间隔为该运动物体固有时的间隔 $d\tau = ds/c$ 。而在 K 系里，钟是速度 v 运动的，所以两次测量的事件，不是发生在同一位置，即 $dx \neq 0$ ，但其时空间隔，依然是 $ds = cd\tau$ 。因为 $dx/dt = v$ ，我们有

$$\sqrt{c^2dt^2 - dx^2} = cd\tau \Rightarrow d\tau = dt\sqrt{1 - v^2/c^2}. \quad (2.47)$$

利用相对论因子 $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$ ，我们得到 $d\tau = dt/\gamma$ 。而因为 $v < c$ ，我们有 $\gamma \geq 1$ 且 $d\tau \leq dt$ ，此处的等号仅当 $v = 0$ 时成立。或者我们可以对上式积分，得到

$$\tau_2 - \tau_1 = \int_{t_1}^{t_2} dt/\gamma < t_2 - t_1. \quad (2.48)$$

¹¹这里为了和上面的洛伦兹变换一致，我们假定 $V = Ve_x$ ， $\Delta x = \Delta xe_x$ 。

¹²容易证明，类时间隔不能保证事件发生的空间坐标是相同的。假设在 K 系里，有两个事件，其 $\Delta s^2 > 0$ ，但 $\Delta x \neq 0$ ，即不同位置发生的两个事件。问题是有可能找到一个参考系 K' 使得 $\Delta x' = 0$ ？答案是可能的，不违背时空间隔的不变性。因为 $\Delta s'^2 = c^2\Delta t'^2 > 0$ ，所以只要 $\Delta t' = \Delta s/c$ 就可以了。之后的固有时即属于这一例子。

即运动的时间变慢。该结论也可以直接从洛伦兹变换得到，即

$$\Delta t = \gamma(\Delta t' + \beta \Delta x'/c) \quad (2.49)$$

对 K' 内静止的时钟而言， $\Delta x' = 0$ ， $\Delta t' = \Delta \tau$ ，所以 $\Delta t = \gamma \Delta \tau$ ，与利用时空间隔所得结论相同。

有很多实验支持时间变缓的结论，例如， μ 子在静止参考系里寿命约为 $2\mu\text{s}$ 。如果它们以 $v \approx c$ 来运动，大概可以运动 600 m。可是尽管大气层顶距地面约为 100 km，我们在地面上却依然可以观测到从大气层顶下来的 μ 子。这是因为 μ 子高速运动，在实验室参考系看来，其可以存在的时间约为 $2\gamma\mu\text{s}$ ，其可运动的距离变大了 γ 倍。

时间变缓效应也常与人体衰老变缓在一起讨论。例如一个人以 $\gamma = 3$ 的速度相对于实验室参考系运动，其随身携带的钟表明时间已经过了 10 年，即 $\Delta \tau = 10$ 年。但是在实验室参考系里看来，其所经过的时间应为 $3\Delta \tau = 30$ 年。当然这个运动的人自己并不会觉得自己寿命变长了，多活了 20 年。对他而言，他只是过了 10 年而已。

另一方面，与同时性的相对性类似，时间变缓效应因为太过于颠覆性，常常被用来证反相对论，从而产生很多看似荒谬其实不然的结论，称为佯谬 (paradox)。以下举两个例子。

例题 2.5 如果 K' 相对于 K 以速度 \mathbf{V} 运动，按上文所述假定 K' 系固有时间隔为 $\Delta t'$ ，在 K 系看来 $\Delta t' = \Delta t/\gamma$ 。然后，如果从 K' 看， K 系是以速度 $-\mathbf{V}$ 运动的，所以用相对性原理，应该也有 $\Delta t = \Delta t'/\gamma$ 。这看起来显然是不通的。

解 要理解这个佯谬，我们需要注意上式中不同时间间隔的含义差义。为方便讨论，我们假定 $\mathbf{V} = V\mathbf{e}_x$ 。对 K' 系而言，我们定义其固定时钟两次测量的事件时空坐标为 (t'_1, x'_1) 以及 (t'_2, x'_1) 。相对应的 K 系里，两事件的时空坐标分别为 (t_1, x_1) , (t_2, x_2) ，以及 $\Delta t = t_2 - t_1 = \Delta t'/\gamma$ 。

首先，我们证明就算从 K' 系看， $\Delta t \neq \Delta t'/\gamma$ 。对于题中两个事件，其在 K' 中对应的时间为

$$t'_1 = \gamma(t_1 - \beta x_1/c), \quad (2.50)$$

$$t'_2 = \gamma(t_2 - \beta x_2/c). \quad (2.51)$$

注意到这时，因为 $x_1 \neq x_2$ ，我们不能得出 $\Delta t' = t'_2 - t'_1 = \gamma \Delta t$ ，而是应该为

$$t'_2 - t'_1 = \gamma[(t_2 - t_1) - \beta(x_2 - x_1)/c]. \quad (2.52)$$

再次利用 $x_2 - x_1 = V(t_2 - t_1)$ 得到

$$t'_2 - t'_1 = \gamma(t_2 - t_1)(1 - \beta^2) = (t_2 - t_1)/\gamma. \quad (2.53)$$

或者说 $\Delta t' = \Delta t/\gamma$ ，与原结论一致。

现在我们反过来考虑 K 系的固有时与 K' 系内时间的差异。我们使用固有时的定义，使用固定于某参考系的时钟来测量其固有时。在 K' 系看来，如果其中位于 x'_1 的时钟两次测量的时间差为 $\Delta t' = t'_2 - t'_1$ ，则其相对应的 $\Delta t = t_2 - t_1$ 是否可以代表 K 系的固有时间隔？答案是否定的，因为尽管在同一惯性参考系内，时间是均匀的，然而从 K' 系看 K 系，其中的时间与空间是紧密联系在一起的。当 K' 系内的时间为 t'_1 时， K 系内不同空间位置处的时间是不一样的，虽然位于 x_1 处的时钟对应时间为 t_1 ，即

$$t'_1 = \gamma(t_1 - \beta x_1/c) \quad (2.54)$$

但位于 x_2 处的时间却不是 t_2 ，记此时间为 t_0 ，从洛伦兹变换可得

$$t'_1 = \gamma(t_0 - \beta x_2/c) \quad (2.55)$$

从以上两式，可以得到

$$t_0 - \beta x_2/c = t_1 - \beta x_1/c \quad (2.56)$$

因此

$$t_0 = t_1 + \beta(x_2 - x_1)/c. \quad (2.57)$$

因此, 如果我们用 K 系中固定于 x_2 处的时钟所测量的时间作为其固有时, 从 K' 系看 K 系, 其固有时并非 $t_2 - t_1$, 而是 $t_2 - t_0$, 为

$$t_2 - t_0 = t_2 - t_1 - \beta(x_2 - x_1)/c. \quad (2.58)$$

注意到 K' 系相对 K 以 $V\mathbf{e}_x$ 来运动, 我们有 $x_2 - x_1 = V(t_2 - t_1)$, 因此 K 系固有时为

$$t_2 - t_0 = (t_2 - t_1)(1 - \beta^2) = \gamma(t'_2 - t'_1)(1 - \beta^2) = (t'_2 - t'_1)/\gamma = \Delta t'/\gamma. \quad (2.59)$$

符合狭义相对论的结论。这里我们运用了 $t_2 - t_1 = \gamma(t'_2 - t'_1)$ 。但与问题佯谬中的关键不同点在于此时 K 系的固有时间隔不是 $\Delta t = t_2 - t_1$ 而是 $t_2 - t_0$ 。

例题 2.6 双生子佯谬: 一对兄妹双胞胎, 妹妹坐高速飞船离去并返回, 问在他们再次相遇的时候, 哪一个更年轻? 在地球上的哥哥看来, 妹妹是高速运动的, 所以再见时, 妹妹更年轻。而在妹妹看来, 哥哥是运动的, 所以哥哥会更年轻。所以到底哪一位更年轻?

解 这个佯谬的关键在于妹妹因为要返回, 必然经历减速再加速的过程, 从而不再是一直处于惯性参考系里。所以惯性参考系的等价性不再适用。兄妹再见时, 是妹妹更年轻, 因为哥哥经历了 T 年, 而妹妹则是 $\int_0^T dt/\gamma < T$ 年。

这种双生子现象也被用来证明相对论的动钟变缓效应, 当然用的不是一对双胞胎, 而是用两个 μ 子。其中一个静止的, 另一个是运动的。当两个 μ 子再相遇时, 运动的 μ 子更年轻。

利用时空图, 我们也可以发现妹妹所对应的世界线的间隔长度 $\Delta s = c\Delta\tau$, 小于哥哥的世界线长度。如图所示。因此这个四维时空具有不同于一般欧氏空间里的性质。

2.2.8 固有长度与固有时

从洛伦兹变换里, 我们也可以发现平行于运动方向的长度会发生收缩 (length contraction), 类似于固有时, 我们定义在相对于被测量物体静止的参考系里, 得到的物体长度, 称为其固有长度 (proper length), 记为 l_0 。假设我们量的是一根棒子, 且其平行于 K' 参考系的 x' 轴, 则在 K' 参考系里, 其长度 $l_0 = |x'_1 - x'_2|$ 。这里 x'_1 与 x'_2 是棒子左右两端在 K' 系里的空间坐标。因为在 K' 参考系里, 棒子是静止的, 所以 x'_1 与 x'_2 坐标的获取是任意时刻都可以的。现在我们要在 K 系里测量该棒子的长度, $l = |x_1 - x_2|$ 。这里 x_1 与 x_2 是棒子左右两端在 K 系里的坐标。因为该棒子相对于 K 系是运动的, x_1 与 x_2 必须是同一个时间测量的坐标, 即 $t_1 = t_2$ 。利用洛伦兹变换 ($K \rightarrow K'$),

$$x'_1 = \gamma(x_1 - Vt_1), \quad (2.60)$$

$$x'_2 = \gamma(x_2 - Vt_2). \quad (2.61)$$

因为 $t_1 = t_2$, 我们有

$$x'_1 - x'_2 = \gamma(x_1 - x_2). \quad (2.62)$$

所以 $l_0 = \gamma l$ 或者 $l = l_0/\gamma$ 。因为 $l < l_0$, 这一效应称为长度收缩效应。

同样对于上一小节里 μ 子从大气层顶到地面的实验。如果我们选一个固定于 μ 子的参考系, 在此系里, μ 子的存活时间也只有 $2\mu\text{s}$ 。但是为什么其可以从大气层顶运动到地面? 这是因为在 μ 子看来, 大气层顶与地面间的距离不再是 100 km , 而是由于尺缩效应变成 $100/\gamma\text{ km}$ 。

例题 2.7 谷仓-梯子佯谬: 一个农夫有一个谷仓, 固有长度为 l_B , 他有一个梯子, 固有长度为 $l_L > l_B$ 。现在他想将梯子放在谷仓里。采取的方法如下: 让他女儿带着梯子以一定的速度 v 奔跑, 使得 $l_L/\gamma \leq l_B$ 。这里

$\gamma = 1/\sqrt{1-v^2/c^2}$ 。这样因为相对论尺缩效应，在谷仓参考系中，梯子可以放进谷仓，当梯子两端都在谷仓内时，两边的门会关闭。可是在女儿看来，不是梯子在动，而是谷仓在动，所以谷仓的长度从 l_B 变得更短了 (l_B/γ)，因此更不可能放下梯子。为简化问题，假定梯子的任何一端将要碰到谷仓仓门时，门会自动打开，不会阻碍梯子的运动。

解 两种描述其实都是对的，并不矛盾。而要理解该佯谬，我们需要注意的是对用类空间隔联系的两事件，其先后顺序是依赖于所选参考系的。

首先我们定义梯子前端到达谷仓右门的事件 1 在谷仓参考系里的时空坐标为 (t_1, x_1) ，而梯子后端到达谷仓左门的事件 2 的时空坐标为 (t_2, x_2) 。现在我们需要证明，这两个事件之间的间隔类型为类空，过程如下。显然 $x_2 - x_1 = l_B$ ，且要使梯子可以放进谷仓，我们需要 $t_1 \geq t_2$ ，即在谷仓系中，事件 2 先于事件 1 发生。从运动学角度，容易得到 $\Delta t = t_1 - t_2 = (l_B - l_L/\gamma)/v > 0$ ，因为 $l_B \geq l_L/\gamma$ 。这两个事件间的时空间隔为

$$\Delta s^2 = c^2 \Delta t^2 - l_B^2 = (l_B - l_L/\gamma)^2/\beta^2 - l_B^2 = (l_B^2/\gamma^2 + l_L^2/\gamma^2 - 2l_B l_L/\gamma)/\beta^2. \quad (2.63)$$

利用 $l_B < l_L$ ，以及 $l_B \geq l_L/\gamma$ ，我们一定有

$$\Delta s^2 < 2(l_L^2/\gamma^2 - l_B l_L/\gamma)/\beta^2 \leq 2(l_L^2/\gamma^2 - l_L^2/\gamma^2)/\beta^2 = 0. \quad (2.64)$$

因此两事件的时空间隔类型为类空型，其发生的先后顺序必然依赖于观察者所在参考系。在谷仓参考系中，事件 2 先于事件 1 发生，在一定时间内，梯子完全位于谷仓内。

现考虑梯子参考系 K' 中两事件的发生顺序。利用洛伦兹变换，两事件的时间坐标分别为

$$t'_1 = \gamma[t_1 - \beta x_1/c], \quad (2.65)$$

$$t'_2 = \gamma[t_2 - \beta x_2/c]. \quad (2.66)$$

然而

$$t'_1 - t'_2 = \gamma[t_1 - t_2 - \beta(x_1 - x_2)/c] = \gamma \left[\frac{l_B - l_L/\gamma}{v} - \frac{\beta l_B}{c} \right] = \frac{\gamma}{v} \left[\frac{l_B}{\gamma^2} - \frac{l_L}{\gamma} \right] \quad (2.67)$$

利用 $l_B \leq l_L/\gamma$ ，我们有

$$t'_1 - t'_2 \leq \frac{\gamma}{v} \left[\frac{l_L}{\gamma^2} - \frac{l_L}{\gamma} \right] = \frac{l_L}{\gamma v} [1 - \gamma] < 0. \quad (2.68)$$

因此，在梯子参考系里，事件 1 先于事件 2 发生。这一点与女儿所观察到的梯子长于谷仓的情况完全符合。因此上述佯谬其实是因为类空间隔事件的先后顺序依赖于所选参考系，其不同观察者看到的不同现象与狭义相对论完全符合，

最后本问题可以变得稍微更复杂一点：在谷仓系里，当梯子两端都在仓内时，两边的门关上后不再打开，这样就可以把梯子关在谷仓里了。这一点在梯子参考系中又如何理解呢？在梯子参考系中，事件 1 依然先发生，且会和谷仓的门发生碰撞。但是梯子的后端此时并不清楚前端发生了碰撞：作用的传递速度是有限的，最大为光速，因此会继续向前，直到后端也进入仓门。但同时，前端发生减速，此时的梯子参考系不再是惯性参考系，其上各点也不具有同样的速度，不能与谷仓参考系等价，因此也不存在佯谬。当梯子最终静止下来后，其长度会恢复原状，大于谷仓长度。因此，要么撑破谷仓门，要么梯子折断。注意到梯子上的不同点可以具有不同的速度，因此这一例子也表明，因为作用传递的速度有限性，刚体的概念在相对论下并不再适用。

最后，我们类似地引入固有体积 (proper volume) 的概念，即物体在一个相对其静止的参考系里测出的体积。因为仅仅平行的运动方向的长度缩小，而垂直于运动方向的尺度不变，所以

$$V = V_0/\gamma. \quad (2.69)$$

因为运动的时钟变缓 ($dt = \gamma d\tau$)，而运动的体积缩小 ($dV = dV_0/\gamma$)，所以如果对时间和空间构成的四维时空而言，其体积元 $dt dV$ 不随参考系发生变化，即 $dt dV = d\tau dV_0$ 。这一点在后面的课程里非常重要。

2.3 狭义相对论的几何表达

在爱因斯坦的相对论发表后，闵可斯基发现可以用优美的几何语言表达其规律。虽然爱因斯坦一开始对这种方式表示了怀疑，但是后来却显然意识到了其重要性，并运用在广义相对论里。用几何语言表达物理规律的一个重要原因是，物理规律会具有形式上的不变性，这样便可以自然而然地满足相对性原理的要求。例如，牛顿第二定律可以表达为 $\mathbf{F} = m\mathbf{a}$ ，即两个向量 \mathbf{F} 与 \mathbf{a} 间是一个简单的正比关系：向量的方向一致，大小相差 m 倍。这一关系显然不会因为选取的坐标系的改变而改变。

对于狭义相对论，闵可夫斯基注意到了时间和空间在其中的等价地位，并称：“时间和空间本身都将消失在阴影里，只有它们的联合体才能存在。¹³”。将时间与空间在一起构成的四维空间，称为闵氏空间。而符合狭义相对论要求的物理规律可以被表达成该空间里张量之间的关系。在这一节里，我们将介绍四维闵氏空间的基本性质，并利用该空间，介绍狭义相对论的张量表达。后续课程中，我们将依赖这种表达方式认识物理规律的协变性。

2.3.1 四维向量

闵氏空间里的基本向量，“位移”向量，是从原点到事件的时空坐标 (ct, x, y, z) 的向量。这是一个基本的四维向量 (four-vector 或者 4-vector)，也会是闵氏空间里其它四维向量的“原型”向量¹⁴，其分量可以写成

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z. \quad (2.70)$$

注意这里的各分量指标为上标，具体原因后面会有更仔细说明。该四维位移向量的模的平方是时空间隔，即

$$s^2 = c^2t^2 - x^2 - y^2 - z^2. \quad (2.71)$$

从四维位移向量模的定义可以看出，四维的闵氏空间与一般的欧几里德空间不同，否则向量的模的平方应为 $c^2t^2 + \mathbf{x}^2$ ，因此闵氏空间属于伪欧氏空间。

四维位移矢量在不同参考系间的变换方式满足洛伦兹变换，对应着四维时空包含了时间轴的双曲旋转。洛伦兹变换利用 x^μ ， $\mu = 0, 1, 2, 3$ 来表达具有更加优美与对称的形式，

$$x^0 = \gamma(x'^0 + \beta x'^1), \quad (2.72)$$

$$x^1 = \gamma(x'^1 + \beta x'^0), \quad (2.73)$$

$$x^2 = x'^2, \quad (2.74)$$

$$x^3 = x'^3. \quad (2.75)$$

与三维空间的向量类似，四维位移矢量与其在不同参考系里的变换方式定义了所有其它的四维向量。

定义 2.1 任何四个量，例如 A^0, A^1, A^2, A^3 ，在任意两个惯性参考系间变换的方式与四维位移矢量 x^μ 相同，即满足洛伦兹变换，则它们构成了一个四维向量。

根据该定义，对于两个惯性参考系 K' 与 K 中，该向量分量的联系式显然为

$$A^0 = \gamma(A'^0 + \beta A'^1), \quad (2.76)$$

$$A^1 = \gamma(A'^1 + \beta A'^0), \quad (2.77)$$

并且 $A^2 = A'^2$ ， $A^3 = A'^3$ 。这里 A^0 称为该向量的时间分量 (time component)，而 A^1, A^2, A^3 称为该向量的空间分量 (space components)。我们经常会将四维向量写成

$$A^\alpha = (A^0, \mathbf{A}), \quad (2.78)$$

¹³原文英文版是 “Space of itself, and time of itself will sink into mere shadows, and only a kind of union between them shall survive.”

¹⁴请参考第一章里向量的定义

这里 \mathbf{A} 是一个三维向量。在本课程里，我们用希腊字符 (α, β, \dots) 表示 $0, 1, 2, 3$ ，而且拉丁字符 (i, j, \dots) 表示 $1, 2, 3$ 。如果不加“四维”的说明，向量，标量，张量等一般指三维空间的量。

在实际运用中，我们一般不是将四个量组合在一起，通过看它们的变换是不是洛伦兹变换的方式来判断这四个量是否构成一个四维向量。更常见的情况是我们从已有的四维向量/标量/张量组合出新的四维向量，从而可以利用其定义，推断新的四维向量的各个分量在不同惯性参考系间必然满足洛伦兹变换。例如 x^α 是个四维向量，而固有时 τ 是一个不随参考系改变的量，因此 $dx^\alpha/d\tau$ 是一个四维向量，该向量在不同参考系间一定是用洛伦兹变换联系起来的。具体的例子见第 2.3.5 章。

现在我们再来看一下四维位移矢量的模平方用 $x^\alpha = (x^0, \mathbf{x})$ 的分量形式来写，

$$s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2, \quad (2.79)$$

类似的，任意四维向量 A^α 的模的平方为

$$(A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2. \quad (2.80)$$

初看起来，四维向量的模的平方的表达方式虽然符合物理，但数学上很难看，也没有办法用一个简单的求和方式或爱因斯坦求和约定来表达。这里我们引入表达四维向量分量的另一种方式¹⁵，协变分量 (covariant components), $x_\alpha = (x_0, x_1, x_2, x_3)$ 。相对应的， $x^\alpha = (x^0, x^1, x^2, x^3)$ 称为四维向量的逆变分量 (contravariant components)。在狭义相对论里，这两种形式之间的关系是

$$x_0 = x^0, x_1 = -x^1, x_2 = -x^2, x_3 = -x^3. \quad (2.81)$$

即时间分量一样，而空间分量差一个负号。注意这两种方式表达的是同一个四维向量，即四维位移矢量。记住逆变与协变分量与其指标的关系可以通过口诀“上逆下协”，或英文里逆变单词，**contravariant**，与上标的单词，**superscript**，都有“r”。需要注意的是，向量的逆变分量和协变分量并不是狭义相对论独有或独创的概念，相反，它们来源于一般的曲线坐标系中的向量张量分析。理解曲线坐标系里的向量张量分析，包括相应的逆变分量和协变分量，虽然不是学习相对论的必要前提，却大有帮助。

利用一个四维向量的两种分量形式，我们可以很方便的写出一个向量的模平方。例如，四维位移矢量的模平方为，

$$s^2 = \sum_{\alpha=0}^3 x^\alpha x_\alpha. \quad (2.82)$$

或任一四维向量， A^α ，其模平方为 $\sum_{\alpha=0}^3 A_\alpha A^\alpha$ 。对于四维向量而言，其运算也可以用爱因斯坦求和约定。只不过，对四维向量而言，多考虑两条规则。1，对哑指标的默认求和是从 0 到 3，同时哑指标的位置一定是一个上标一个下标。2，不同项里自由指标不仅要符号一样，上下位置也要一样。例如 $A_\alpha + B_\alpha$ 是可以的，但是 $A_\alpha + B^\alpha$ 是错误的式子。利用爱因斯坦求和约定，四维矢量 A^α 的模平方可以写为 $A^\alpha A_\alpha$ 。类似的，两个四维向量可以进行点乘，因为模的平方可以看成四维向量点乘自己。所以两个四维向量 A^α 与 B^α 的点乘定义为 $A^\alpha B_\alpha$ 或者 $A_\alpha B^\alpha$ 。这里可以进一步看到，同一个哑指标可以交换上下位置。注意，与欧氏空间不同，闵氏空间里四维向量的模可正可负也可 0，这一点可以类比于前文所讨论的不同时空间隔类型。

在一些教材或论文里，会引入复数时空。在此时空里，时间分量为虚数，即 ict ，从而四维矢量的模应该是 $x^2 + y^2 + z^2 + (ict)^2$ 。在复数时空里，洛伦兹变换相当于是保持向量 (ict, x, y, z) 的模不变，对应着该复数时空的普通旋转。而本课程里，则通过引入逆变和协变分量与实数时空 (real spacetime)。虽然选择实数空间会比用复数空间稍复杂一点，但这种表达能更清楚地表明时空的几何结构，且更容易拓展到非惯性参考系中的物理过程，在后续理论中（如广义相对论）里会很常用。

¹⁵更一般的情况下引入协变分量可见第章。

2.3.2 四维标量

四维位移矢量的模是时空间隔，其是一个不随参考系变化的量，同时也是一个四维标量 (four-scalar 或 4-scalar)。类似的，我们可以定义一般的四维标量。

定义 2.2 在洛伦兹变换下，不随参考系改变的物理量，称为四维标量。

依据前文据述，除时空间隔与固有时外，我们有四维空间的体积元， $d\Omega \equiv c dt dx dy dz = c dt dV$ ，也是一重要四维标量。

在很多物理学领域里，物理过程里的不变量往往是研究一个系统的物理性质非常重要的切入点。类似的，四维标量对于学习相对论和电动力学也很重要，因此我们需要知道利用四维张量构造四维标量的方法。例如，向量的模平方或两个向量的点乘 $x^\alpha x_\alpha$ 或 $A^\alpha B_\alpha$ 显然都是四维标量，它们并没有自由指标。随着我们遇到更多的四维向量，更多的四维标量也会被构造或发现。

2.3.3 四维张量

进一步的，我们引入（二阶）四维张量 (four-tensor 或者 4-tensor) 的概念，其定义与三维情况下的（二阶）张量类似。

定义 2.3 如果 16 个量 $F^{\alpha\beta}$ 放在一起，在洛伦兹变换情况下，与两个四维向量的积 ($A^\alpha B^\beta$) 的变换方式相同，则这 16 个量构成一个二阶四维张量。

F^{11} 与 $A^1 B^1$ 的变换方式相同，而后者为

$$A^1 B^1 = \gamma(A'^1 + \beta A'^0) \gamma(B'^1 + \beta B'^0) = \gamma^2 (A'^1 B'^1 + \beta A'^1 B'^0 + \beta A'^0 B'^1 + \beta^2 A'^0 B'^0). \quad (2.83)$$

因此，四维张量 $F^{\alpha\beta}$ 的 F^{11} 分量在不同参考系间的变换方式为

$$F^{11} = \gamma^2 (F'^{11} + \beta F'^{10} + \beta F'^{01} + \beta^2 F'^{00}). \quad (2.84)$$

类似的可以得到四维张量的所有其它分量在两个惯性参考系间的变换规律。对一个四维张量而言，指标的上下位置有多种组合形式，如 $F^{\alpha\beta}$ （逆变形式）， $F_{\alpha\beta}$ （协变形式）， F^α_β （混合形式）。在混合形式里，下指标里的点用以表达 β 指标在后， α 在前，因为混合形式还有一种，即 F_β^α 。

一个重要的四维张量是四维度规张量 $g^{\alpha\beta}$ ，也称为度规系数。对于闵氏空间，该系数为

$$g_{\alpha\beta} = g^{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha = \beta = 0 \\ -1 & \text{if } \alpha = \beta \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.85)$$

或者利用矩阵可表达为

$$g_{\alpha\beta} = g^{\alpha\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (2.86)$$

利用该度规张量，我们可以方便地在逆变和协变分量间进行转换¹⁶。对四维向量 A^α 而言，

$$A_\alpha = g_{\alpha\beta} A^\beta \text{ or } A^\alpha = g^{\alpha\beta} A_\beta. \quad (2.87)$$

¹⁶更一般情况下的这种转换可以用曲线坐标里度规系数的定义来理解。在曲线坐标系里，一个向量 \mathbf{A} 可以表达为 $A^i \mathbf{e}_i$ 或 $A_i \mathbf{e}^i$ ，这里 $A^i \equiv \mathbf{A} \cdot \mathbf{e}^i$ 与 $A_i \equiv \mathbf{A} \cdot \mathbf{e}_i$ 分别称为该向量的逆变与协变分量，而 \mathbf{e}_i 与 \mathbf{e}^i 则分别为协变基矢与逆变基矢，具有性质 $\mathbf{e}_i \cdot \mathbf{e}^j = \delta_i^j$ 。这里 δ_i^j 在 $i = j$ 时为 1，否则为 0。而度规系数的定义为 $g^{ij} \equiv \mathbf{e}^i \cdot \mathbf{e}^j$ 或 $g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j$ ，因此 $A^i = \mathbf{A} \cdot \mathbf{e}^i = A_j \mathbf{e}^j \cdot \mathbf{e}^i = g^{ij} A_j$ 。从该系数定义可以看出， $g^{ij} = g^{ji}$ ，且 $g_j^i \equiv \mathbf{e}^i \cdot \mathbf{e}_j = \delta_j^i$ 。

这样时间分量保持不变，而空间分量改变符号。例如，我们可以将时空间隔表达为 $ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta$ 。

对于一般的四维张量，我们也可利用度规系数在其不同形式间进行转换，而仅仅需要了解爱因斯坦求和约定便可写出其转换方式，例如 $F_{\mu\nu} = g_{\mu\alpha} F^{\alpha\beta} g_{\beta\nu}$, $F_\mu^\beta = g_{\mu\alpha} F^{\alpha\beta}$, \dots 。利用闵氏空间度规系数的定义，我们可以得出四维张量的简单转换规律，

1. 上下移动一次空间指标 (1, 2, 3)，改变一次正负号，
2. 上下移动时间指标 (0) 不改变正负号。

例如: $F_{00} = F^{00}$, $F_{01} = -F^{01}$, $F_{11} = F^{11}$, $F_0^0 = F^{00}$, $F_0^1 = F^{01}$, $F_1^0 = -F^{10}$, $F_1^1 = -F^{10}$, \dots 。同样的规定也适用于度规张量本身，所以其混合形式为 $g_\alpha^\beta = g_{\alpha\nu} g^{\nu\beta}$ 。很容易得到，

$$g_\beta^\alpha = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{otherwise} \end{cases} \quad (2.88)$$

这也是四维时空里的单位张量 δ_β^α 。这里因为该张量为对称张量，所以我们不用另外加点去区分 α, β 指标的先后顺序。

利用四维张量与四维向量，我们可以构造四维标量，保证最后的表达式里没有自由指标。例如利用四维张量 $F^{\alpha\beta}$ 双点乘自己，我们得到一四维标量

$$F_{\alpha\beta} F^{\beta\alpha} = \text{四维标量}. \quad (2.89)$$

或者利用四维向量和四维张量的点乘来构造四维标量，如 $A_\alpha B_\beta F^{\alpha\beta}$ 。

2.3.4 四维梯度与四维散度

四维张量也有相应的微分运算，为简单计，我们在这门课里只介绍基础的微分运算，包括四维梯度与四维散度。对一个四维标量 ϕ 而言，其四维梯度是一个四维向量，定义为

$$\frac{\partial \phi}{\partial x^\alpha} = \left(\frac{1}{c} \frac{\partial \phi}{\partial t}, \nabla \phi \right) \quad (2.90)$$

如何判断这是一个逆变还是协变分量？这里只需要注意到四维标量 ϕ 的微分可以写为

$$d\phi = dx^\alpha \frac{\partial \phi}{\partial x^\alpha}. \quad (2.91)$$

因为 dx^α 是一个逆变分量， $\partial\phi/\partial x^\alpha$ 必然是一个四维向量的协变分量，因此有时会将这种梯度记为

$$\frac{\partial \phi}{\partial x^\alpha} \equiv \partial_\alpha \phi.$$

类似的，我们也可以定义逆变形式的四维梯度，

$$\frac{\partial \phi}{\partial x_\alpha} \equiv \partial^\alpha \phi = \left(\frac{1}{c} \frac{\partial \phi}{\partial t}, -\nabla \phi \right).$$

从四维梯度的逆变与协变形式可以看出，不能仅依据空间分量的正负号来判断一个四维向量是逆变形式还是协变形式，而是应该从其构造过程出发，或者更一般的，从其在不同参考系间的变换规律来判断：逆变分量的变换规律与 x^α 相同，而协变分量的变换规律与 x_α 相同。

对一个四维向量， $A^\alpha = (A^0, \mathbf{A})$ ，也可以定义其四维散度

$$\frac{\partial A^\alpha}{\partial x^\alpha} = \partial_\alpha A^\alpha = \partial^\alpha A_\alpha = \frac{1}{c} \frac{\partial A^0}{\partial t} + \nabla \cdot \mathbf{A}. \quad (2.92)$$

显然，这个散度的结果是一个四维标量。例如，以后我们会学到，电荷密度与电流密度会组合成一个四维向量， $j^\alpha = (\rho c, \mathbf{j})$ 。该四维向量的散度为

$$\frac{\partial j^\alpha}{\partial x^\alpha} = \frac{1}{c} \frac{\partial(\rho c)}{\partial t} + \nabla \cdot \mathbf{j} = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j}. \quad (2.93)$$

利用电荷连续性方程，四维电流向量的散度为 0。因为一个四维标量的梯度也是一个四维向量，我们同样可以计算其散度

$$\partial_\alpha \partial^\alpha \phi = \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi \equiv \square \phi. \quad (2.94)$$

这里 $\square \equiv \partial_\alpha \partial^\alpha$ 称为达朗贝尔算符 (the d'Alembert operator)¹⁷，会是后续课程里常常出现的一个算符。从其构造过程可以看出，达朗贝尔算符可以看成四维空间里的拉普拉斯算符，即 $\nabla \cdot \nabla = \nabla^2$ 。

2.3.5 粒子的几个常见四维量

对于一个粒子而言，有一些基本的四维量。首先，其时空间隔 (s) 与固有时 (τ) 为四维标量。其次，其四维空间的位置为四维向量，即 $x^\mu = (ct, \mathbf{x})$ 。基于现有的四维张量，我们可以依据张量的规律来构造新的四维量。

第一个例子为基于四维位移矢量和固有时构建的四维速度，定义为 $u^\alpha = \mathrm{d}x^\alpha / \mathrm{d}\tau$ ，这里 τ 是粒子静止参考系里的固有时。在 K 中，粒子的速度为 v ，因此 $\mathrm{d}\tau = \mathrm{d}t / \gamma$ ，而 $\gamma \equiv 1 / \sqrt{1 - v^2 / c^2}$ 。所以我们有

$$u^0 = \frac{c \mathrm{d}t}{\mathrm{d}\tau} \gamma = \gamma c, \quad (2.95)$$

$$u^1 = \frac{\mathrm{d}x}{\mathrm{d}\tau} \gamma = \gamma v_x, \quad (2.96)$$

$$u^2 = \gamma v_y, \quad (2.97)$$

$$u^3 = \gamma v_z. \quad (2.98)$$

这里 v_x, v_y, v_z 为粒子在 K 中的运动速度，而四维速度也可以写成

$$u^\alpha = (\gamma c, \gamma v_x, \gamma v_y, \gamma v_z) = (\gamma c, \gamma \mathbf{v}) = (u^0, \mathbf{u}). \quad (2.99)$$

这里 $\mathbf{u} = \gamma \mathbf{v}$ ，为四维速度 u^α 的空间分量。如果利用 \mathbf{u} 向量，相对论因子可以表达为 $\gamma = \sqrt{1 + u^2 / c^2}$ 。因为四维速度是一个四维向量，其在两个不同的参考系 K' 与 K 里的变换方式为

$$u^0 = \gamma_V (u'^0 + \beta_V u'^1), \quad (2.100)$$

$$u^1 = \gamma_V (u'^1 + \beta_V u'^0), \quad (2.101)$$

$$u^2 = u'^2, \quad (2.102)$$

$$u^3 = u'^3. \quad (2.103)$$

这里为了区分，我们在 γ 与 β 的表达式里引入下标 V ，即 $\beta_V = V / c$ 而 $\gamma_V = 1 / \sqrt{1 - \beta_V^2}$ 。利用该变换，也可以求出速度 $\mathbf{v} = \mathrm{d}\mathbf{x} / \mathrm{d}t$ 在不同参考系里的变换关系，式 (2.42)-(2.44)。而四维速度的模平方为

$$u^\alpha u_\alpha = \frac{\mathrm{d}x^\alpha \mathrm{d}x_\alpha}{\mathrm{d}\tau \mathrm{d}\tau} = \frac{\mathrm{d}s^2}{\mathrm{d}s^2 / c^2} = c^2, \quad (2.104)$$

在上式里，我们运用了 $\mathrm{d}\tau = \mathrm{d}s / c$ 。或者也可以直接计算，得到

$$u^\alpha u_\alpha = \gamma^2 (c^2 - v^2) = c^2, \quad (2.105)$$

这个结果显然是一个四维标量。将四维速度乘上粒子质量 (mu^α) 还可以得到四维动量，具体的讨论放下一章节。

第二个例子为基于四维速度与固有时定义的物理量 $a^\mu = \mathrm{d}u^\mu / \mathrm{d}\tau$ 。该四维向量显然与三维的加速度类似，因此被称为四维加速度。四维加速度有一个有趣的性质：其与四维速度的点乘为 0，即四维加速度与四维速度是“垂直”的。因为

$$a^\mu u_\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} u_\mu = \frac{1}{2} \left(\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} u_\mu + \frac{\mathrm{d}u_\mu}{\mathrm{d}\tau} u^\mu \right). \quad (2.106)$$

¹⁷ 达朗贝尔算符也常被记为 \square^2 ，这种记法与拉普拉斯算符更有对称性，但本书为简单记为 \square 。此外，因为对应着波动方程，有时也会被称为波动算符。该算符以法国数学和物理学家 Jean le Rond d'Alembert 命名。

在最后一式的第二项里，我们交换了哑指标 μ 的上下位置，所以

$$a^\mu u_\mu = \frac{1}{2} \frac{d(u^\mu u_\mu)}{d\tau} = 0. \quad (2.107)$$

基于四维加速度，也可以构造四维力 $F^\mu = ma^\mu$ ，详见习题。

从以上两个四维向量的例子，我们可以发现，四维向量习惯上都是以其空间分量来命名的。

2.4 张量表达与物理规律的不变性

用张量这样的几何量表达物理规律的优点是其所写出的方程自然而然地具有协变性。在本课程里，如果一个物理规律可以用闵氏空间的四维张量来表达，则其自然而然的具有了洛伦兹协变性：在不同的惯性参考系里，具有同样的形式。我们首先对狭义相对论证明这一点。

2.4.1 狭义相对论中的物理规律不变性

对于狭义相对论，上述四维向量 x^μ 在不同参考系里的变换关系可以表达为

$$x^\mu = \Lambda^\mu_\nu x'^\nu \quad (2.108)$$

这里我们引入 Λ^μ_ν 来代表洛伦兹变换。因为 Λ^μ_ν 是对称的，即 $\Lambda^\mu_\nu = \Lambda^\nu_\mu$ ，所以我们不区分两个指标的前后次序。对于以相对速度为 $\mathbf{V} = V\mathbf{e}_x$ 运动的两个参考系 K 与 K' 而言， $\Lambda^0_0 = \gamma$ ， $\Lambda^0_1 = \beta\gamma$ ， $\Lambda^1_1 = \gamma$ ，而其它分量为 0。另一方面对于张量 $F^{\alpha\beta}$ 而言，其变换规律按定义显然为

$$F^{\mu\beta} = \Lambda^\alpha_\mu \Lambda^\beta_\nu F'^{\alpha\nu}. \quad (2.109)$$

而四维标量则在不同参考系里保持不变。

现在我们假定一个物理规律在 K 系中用张量表达为

$$F^{\alpha\beta} = T^{\alpha\beta}. \quad (2.110)$$

显然，在 K' 系中， $F \rightarrow F'$ ，而 $T \rightarrow T'$ 。那在 K' 系中，同样的物理规律，是否依然可以表达为 $F' = T'$ ？利用张量的变换规律，我们有

$$F'^{\alpha\beta} = \Lambda'^\alpha_\mu \Lambda'^\beta_\nu F^{\mu\nu}, \quad (2.111)$$

$$T'^{\alpha\beta} = \Lambda'^\alpha_\mu \Lambda'^\beta_\nu T^{\mu\nu}. \quad (2.112)$$

这里 Λ' 代表从 K 系到 K' 系的洛伦兹变换。显然， $F'^{\alpha\beta} = T'^{\alpha\beta}$ 中 K' 系中成立，物理规律具有不变性。

因此，直接从四维张量出发去探索新的基本物理规律，所得到的规律自然满足相对性原理。当然，如果一个物理规律不是用四维张量来表达，并不意味着其一定没有洛伦兹协变性。只是在这种情况下，要证明其具有洛伦兹协变性需要付出更多的努力。

2.4.2 更广义情况下的物理规律不变性

我们现在将上述用张量表达的物理规律不变性推广到更一般的情况，来进一步说明为什么张量在相对论部分应用如此广泛，顺便也通过张量分量的变换规律介绍逆变与协变张量更一般的定义¹⁸。如果我们有两个坐标系，其由 x^α 与 x'^α 所定义。则 dx^α 与 dx'^α 间显然可以由下式联系，

$$dx^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta} dx'^\beta. \quad (2.113)$$

¹⁸对于混合形式的张量分量，很容易从逆变与协变形式的分量变换规律的类似过程得到，不再赘述。

我们定义逆变形式的向量分量与上式所要求的变换规律相同，即

$$A^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta} A'^\beta. \quad (2.114)$$

对于标量，我们则要求其在不同参考系间不变，即 $\phi(x^\mu) = \phi'(x'^\mu)$ ，这里 x'^μ 为 x^μ 在 K' 系中的对应坐标。同样的，我们定义二阶张量的逆变形式为满足以下变换规律的物理量

$$F^{\alpha\beta} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} F'^{\mu\nu}. \quad (2.115)$$

可以看出，熟练地掌握爱因斯坦求和指标的规律，很容易写出上述变换式。依据以上定义，如果一个张量的所有分量在某个参考系中为 0，则在所有其它的参考系中也都为 0。因此，如果一个物理规律表达为

$$F^{\alpha\beta} = A^\alpha B^\beta, \quad (2.116)$$

则显然我们有 $T^{\alpha\beta} \equiv F^{\alpha\beta} - A^\alpha B^\beta = 0$ 对任意的 α, β 指标均成立。因此，在任意的其它参考系中，我们也一定有对应的关系 $T'^{\alpha\beta} \equiv F'^{\alpha\beta} - A'^\alpha B'^\beta = 0$ ，或者

$$F'^{\alpha\beta} = A'^\alpha B'^\beta. \quad (2.117)$$

即物理规律具有形式上的不变性。这一点也可直接利用张量的坐标变换规律得到，因为

$$F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} F^{\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} A^\mu B^\nu = A'^\alpha B'^\beta. \quad (2.118)$$

同样具有形式上的不变性。

注意到不是所有的物理量都具有上述的逆变分量变换规律。假如有一标量 ϕ ，其梯度的分量在两个参考系中的关系

$$\partial_\alpha \phi = \frac{\partial x'^\beta}{\partial x^\alpha} \partial'_\beta \phi \quad (2.119)$$

这里 $\partial'_\beta \equiv \partial/\partial x'^\beta$ 。因此，我们引入向量的协变形式，其变换规律与上式相同，即

$$A_\alpha = \frac{\partial x'^\beta}{\partial x^\alpha} A'_\beta. \quad (2.120)$$

类似的，对于协变形式的张量，我们有

$$F_{\alpha\beta} = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} F'_{\mu\nu}. \quad (2.121)$$

因此，与前文同样的证明过程，容易理解协变形式的张量表达的物理规律一样具有不变性。

以上关于张量分量的变换性质也有助于我们去判断一些量是否是张量，方法是看其在不同坐标系间的变换是否满足相应的变换规律。

例题 2.8 证明度规系数 $g_{\alpha\beta}$ 构成一个张量。

证明 要证明这一点，我们利用时空间隔在不同参考系间的不变性，即

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = g'_{\alpha\beta} dx'^\alpha dx'^\beta. \quad (2.122)$$

利用

$$dx'^\alpha = \frac{\partial x'^\alpha}{\partial x^\mu} dx^\mu, \quad (2.123)$$

$$dx'^\beta = \frac{\partial x'^\beta}{\partial x^\nu} dx^\nu. \quad (2.124)$$

我们有

$$g'_{\alpha\beta} dx'^\alpha dx'^\beta = g'_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} dx^\mu \frac{\partial x'^\beta}{\partial x^\nu} dx^\nu. \quad (2.125)$$

因此，一定有

$$g_{\mu\nu} = g'_{\alpha\beta} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}}. \quad (2.126)$$

符合张量协变分量在不同坐标系里的变换规律，因此度规系数 $g_{\alpha\beta}$ 构成一个二阶张量。

在本课程里，我们会从最小作用量原理出发，直接推导出用四维张量形式所表达的电磁场规律，这样所得到的规律自然满足爱因斯坦相对性原理的要求。同样的原则也可以被运用到其它物理过程上。

第三章 外场中的带电粒子运动

3.1 经典粒子的最小作用量原理

经典力学里的一个重要出发点是最低作用量原理。在本课程里，我们将最低作用量原理扩展到可以描述符合狭义相对论要求的粒子运动方程。而在下一章中，我们会进一步发现最低作用量原理也可以给出场所应当满足的方程。因此最低作用量原理成了现代物理学发展一个重要基石。

在经典力学里，一个粒子的作用量可以通过拉格朗日量 (Lagrangian) L 表达为

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt. \quad (3.1)$$

这里 q 是广义坐标，在 t_1 与 t_2 时刻，粒子位于 $q(t_1)$ 与 $q(t_2)$ ，即 $q(t_1) = q^{(1)}$ and $q(t_2) = q^{(2)}$ 。拉格朗日量 L 是广义坐标 (q)，广义速度 (\dot{q})，与时间的函数。在起点与终点之间，可以有无数条路径连接，沿每一条路径，我们都可以计算出 L 对时间的积分，从而得到该路径的作用量。最低作用量原理表明，沿着真实物理路径的作用量取极值，即

$$\delta S = 0. \quad (3.2)$$

通过该原理，我们便可以获得拉格朗日量 L 所满足的方程，从而得到真实路径的方程。

为了得到真实路径的方程，我们注意到如果 $q = q(t)$ 是真实路径，则对该路径 q 的微小扰动

$$q(t) + \delta q(t), \quad (3.3)$$

所引起的作用量 S 的变化为 0，即

$$\delta S = \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta \dot{q}, t) dt - \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = 0. \quad (3.4)$$

所限制的条件是首尾两点的广义坐标不变，即 $q(t_1) = q^{(1)}$ and $q(t_2) = q^{(2)}$ ，或者说

$$\delta q(t_1) = 0, \quad \delta q(t_2) = 0. \quad (3.5)$$

在 δS 里保留到一阶项，我们有

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt = 0. \quad (3.6)$$

注意到 $\delta \dot{q}$ 是由我们对 q 进行扰动所引起的 \dot{q} 的变化，所以

$$\delta \dot{q} = \frac{d}{dt} (q + \delta q) - \frac{d}{dt} (q) = \frac{d}{dt} \delta q, \quad (3.7)$$

即，变分算符 δ 与微分算符 d/dt 可以交换次序。所以

$$\delta S = \left[\frac{dL}{d\dot{q}} \delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt = 0. \quad (3.8)$$

因为 $\delta q(t_1) = 0$ 并且 $\delta q(t_2) = 0$ ，我们有

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt = 0. \quad (3.9)$$

因为 δS 必须对任意的 δq 为 0，所以我们有

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0, \quad (3.10)$$

即拉格朗日方程。有了该方程，如果我们知道一个系统的拉格朗日量 L ，就可以写出该系统的动力学方程。

以上的推导过程其实展现了通过最小作用量原理获取系统的演化方程的一个基本思路，即为了获得 q 所满足的方程，我们将 δS 写成，

$$\delta S = \int (\dots) \delta q \quad (3.11)$$

的形式。这一步还往往会用到分步积分与 q 在起始与末尾点扰动量为 0 的条件。因为 $\delta S = 0$ 对任意的扰动 δq 都成立，所以括号里的式子一定为 0，即演化方程。

作用量和拉格朗日量有一些重要的性质，即它们不是唯一的，因为最小作用量里的“大小”仅仅是一个相对的概念。所以如果我们在作用量里加一个常数，并不会影响沿真实和扰动路径计算出的作用量之间的相对大小。因此，如果我们有二个拉格朗日量： $L'(q, \dot{q}, t)$ 与 $L(q, \dot{q}, t)$ ，且它们之间仅相差一个函数 $f(q, t)$ 的全导数，即

$$L' = L + \frac{d}{dt} f(q, t), \quad (3.12)$$

则

$$S' = \int_{t_1}^{t_2} L' dt = \int_{t_1}^{t_2} L dt + \int_{t_1}^{t_2} \frac{df}{dt} dt = S + f(q^{(2)}, t) - f(q^{(1)}, t). \quad (3.13)$$

这种情况下， $\delta S'$ 和 δS 会产生完全相同的运动方程。

在经典力学里，最常见的一个例子是粒子在外场中的运动。例如一维情况

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \left(\frac{1}{2} m v^2 + \alpha \phi \right) dt \quad (3.14)$$

这里 α 为一个耦合系数，代表外场 ϕ 对粒子影响的强弱；同时，外场 ϕ 沿着粒子的路径取值。该作用量对应的欧拉-拉格朗日方程显然为

$$m \mathbf{a} = \alpha \nabla \phi. \quad (3.15)$$

该式中 $\mathbf{a} = d\mathbf{v}/dt$ 为加速度。从这里可见耦合系数 α 的作用：对同样质量的粒子， $|\alpha|$ 越小，则耦合越弱；反之亦然。

3.2 自由粒子的相对论动力学

以上的作用量 $S = \int L dt$ 并不是用四维量来表达的，因此其导出的方程（如 3.10）并不一定满足相对论所要求的协变性，即在不同参考系里都具有同样的形式。那我们该如何将以上最小作用量原理应用在相对论问题上，使得获得的物理规律方程自动满足相对性原理的要求？注意到，在相对论里，时间与空间处于同等地位，因此最小作用量原理要求沿实际路径的作用量取极值，应该理解为沿四维时空中粒子的世界线取极值，积分变量可以为 dx^μ 或者 $d\tau$ 这样的四维量。同时我们扰动的不再只是空间变量，而应该是粒子的世界线时空坐标 x^μ 。其次，物理规律的协变性要求我们用四维标量来表达作用量¹，这样得到的运动方程自然而然地具有协变性。因为选用了四维标量，在一个参考系里对应着真实路径的作用量，在其它任意参考系里也一定对应着真实路径，因为都对应对应着极值。

那么在相对论力学框架下，自由粒子的作用量 $S = \int_a^b dS$ 应该是什么样子？这里我们有两个基本的要求。一是四维标量原则：作用量需要满足在洛伦兹变换下是不变量，即其应该是四维标量²。二是对应性原则：即在非相对论情况下 ($c \rightarrow \infty$)，该作用量应该可以给出非相对论运动方程。利用这两点，我们可以去猜测一个相对论自由粒子的作用量。四维标量原则注定了作用量 dS 的可能形式并不太多。有 $u^\alpha du_\alpha$, $x^\alpha dx_\alpha$, $u^\alpha dx_\alpha$, $d\tau$, ds 。但

¹这一点在下一章场的方程里可以有更清楚的体现。

²更一般的，我们希望其也能在坐标平动变换时保持不变，即庞加莱不变性。

其中一些如 $u^\alpha \mathrm{d}u_\alpha \propto \mathrm{d}(u^\alpha u_\alpha) = \mathrm{d}c^2 = 0$, 显然不合适。其次对非相对论粒子, $\mathrm{d}S = (1/2)mv^2 \mathrm{d}t$, 这样我们猜测相应的相对论自由粒子作用量可能具有类似的形式, 如 $(1/2)mu^\alpha u_\alpha \mathrm{d}\tau$ 。这里我们仅仅将 $v^2 = \mathbf{v} \cdot \mathbf{v} \rightarrow u^\alpha u_\alpha$, $\mathrm{d}t \rightarrow \mathrm{d}\tau$ 。注意, 这只是一种猜测的技巧, 不是从非相对论过渡到相对论的规则。因为 $u^\alpha u_\alpha = c^2$, 我们猜测作用量应该有如下形式,

$$S = \int_a^b \frac{\alpha}{2} mc^2 \mathrm{d}\tau, \quad (3.16)$$

这里 α 是一个额外的常数系数, 可以作为一个自由参数用来进一步确保作用量满足第二个条件。

实验室参考系的时间 t 与粒子固有时 τ 间有 $\mathrm{d}t = \mathrm{d}\tau/\gamma$, 所以该作用量所对应的拉格朗日量为

$$S = \int L \mathrm{d}t \Rightarrow L = \frac{\alpha}{2} mc^2 / \gamma. \quad (3.17)$$

在非相对论情况下 $c \rightarrow \infty$,

$$L = \frac{\alpha}{2} mc^2 \sqrt{1 - \frac{v^2}{c^2}} \rightarrow \frac{\alpha}{2} mc^2 - \frac{\alpha mv^2}{4}. \quad (3.18)$$

注意到 $(\alpha/2)mc^2$ 是个常数, 所以只要 $\alpha = -2$, 则上述拉格朗日量在非相对论情况下为 $-mc^2 + (1/2)mv^2$, 其所给出的运动方程与非相对论力学完全一致。这样自由粒子的相对论作用量应为

$$S = -mc^2 \int_a^b \mathrm{d}\tau. \quad (3.19)$$

相应的拉格朗日量是

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}, \quad (3.20)$$

或者 $L = -mc^2/\gamma$ 。这里 $\gamma \equiv 1/\sqrt{1 - v^2/c^2}$ 即为相对论因子。

3.2.1 相对论能量和动量

从相对论自由粒子的拉格朗日量出发, 我们可以定义一个粒子的相对论动量

$$\mathbf{p} \equiv \frac{\partial L}{\partial \mathbf{v}}. \quad (3.21)$$

利用 $L = -mc^2 \sqrt{1 - v^2/c^2}$, 所以

$$\mathbf{p} = -mc^2 \frac{1}{2} \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \left(-\frac{2\mathbf{v}}{c^2}\right) = \gamma m \mathbf{v},$$

或者

$$\mathbf{p} = \frac{m \mathbf{v}}{\sqrt{1 - v^2/c^2}} = \gamma m \mathbf{v}.$$

如果 $c \rightarrow \infty$, $\mathbf{p} = m \mathbf{v}$, 即非相对论动量。

一个粒子的能量同样可以用以上拉格朗日量 L 定义为

$$\mathcal{E} \equiv \mathbf{v} \cdot \frac{\partial L}{\partial \mathbf{v}} - L = \mathbf{p} \cdot \mathbf{v} - L. \quad (3.22)$$

利用 $\mathbf{p} = \gamma m \mathbf{v}$, 我们有

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = \gamma mc^2. \quad (3.23)$$

如果 $v \rightarrow 0$, $\mathcal{E} \rightarrow mc^2$, 即一个粒子的静止能量。相应的, $\mathcal{E}_k \equiv (\gamma - 1)mc^2$ 称为粒子的动能。如果 $c \rightarrow \infty$, 则 $\mathcal{E} = mc^2 + \frac{1}{2}mv^2$ 。有的教科书或者文章会将上式中的 m 称为“静止质量”, 但我们则直接称其为质量。不额外定义其它质量, 如 γm 。注意到一个系统的时间对称性是与能量守恒紧密联系在一起的, 因此能量守恒定律总是成立的。但从上述能量定义, 我们很容易理解, 在相对论力学框架下, 质量守恒一般是不成立的, 典型的例子如核聚变与核裂变。

粒子的能量与动量是有着特殊联系的。注意到四维速度的定义 $u^\alpha = (\gamma c, \gamma \mathbf{v})$ 。如果我们用质量 m 去乘四维向量 u^α , 则可以得到一个新的四维向量, 四维动量向量,

$$p^\alpha = mu^\alpha = (\gamma mc, \gamma m\mathbf{v}) = \left(\frac{\mathcal{E}}{c}, \mathbf{p} \right). \quad (3.24)$$

这里 $\mathbf{p} = (p_x, p_y, p_z)$ 。能量与动量构成四维向量 p^α 的时间与空间分量。因为 p^α 是一四维向量, 其在不同惯性参考系里的变换与 x^α 相同, 满足洛伦兹变换, 即

$$\mathcal{E} = \frac{\mathcal{E}' + Vp'_x}{\sqrt{1 - V^2/c^2}}, \quad (3.25)$$

$$p_x = \frac{p'_x + (V/c^2)\mathcal{E}'}{\sqrt{1 - V^2/c^2}}, \quad (3.26)$$

$$p_y = p'_y, \quad (3.27)$$

$$p_z = p'_z. \quad (3.28)$$

这样我们就知道了粒子的能量与动量在不同参考系里是如何变换的。对一光子而言, 其能量为 $\mathcal{E} = \hbar\omega$, 动量为 $\hbar\mathbf{k}$, 从以上变换规律便可以得到电磁波频率在不同参考系间的变换关系, 即光的相对论多普勒效应³。

四维动量向量的模可以得到能量与动量大小的关系。因为 $p^\alpha = mu^\alpha$, 所以

$$p_\alpha p^\alpha = m^2 c^2. \quad (3.29)$$

代入 $p_\alpha = (\mathcal{E}/c, -\mathbf{p})$ 以及 $p^\alpha = (\mathcal{E}/c, \mathbf{p})$, 我们得到

$$\frac{\mathcal{E}^2}{c^2} - p^2 = m^2 c^2, \quad (3.30)$$

或者

$$\mathcal{E}^2 = p^2 c^2 + m^2 c^4 \quad (3.31)$$

此即能量-动量方程。而对光子而言, $m = 0$, 我们有 $p = \frac{\mathcal{E}}{c}$ 。这个式子也近似适用于 $\mathcal{E} \gg \mathcal{E}_0$ 的情况, 此时 $p \approx \mathcal{E}/c$ 。

例题 3.1 电子的静止能量约为 0.511 MeV。请估计动能为 50 MeV 的电子其相对论因子与动量大小。

解 利用动能定义, 显然我们有 $\gamma = \mathcal{E}_k/\mathcal{E}_0 + 1 \approx 101$ 。而其动量大小约为 $p \approx \mathcal{E}/c \approx \mathcal{E}_k/c = 50 \text{ MeV}/c$ 。

3.3 自由粒子运动方程的协变形式

利用作用量 S 和最小作用量原理, 我们可以得到满足相对论原理所要求的协变方程, 具体过程经典力学里的过程相似。首先固定粒子轨迹两端的时空坐标, 并扰动粒子在四维空间里的路径 $x^\alpha \rightarrow x^\alpha + \delta x^\alpha$ 。依最小作用量原理, 沿真实路径的作用量最小即 $\delta S = 0$ 。注意到

$$S = -mc^2 \int_a^b d\tau = -mc \int_a^b \sqrt{dx_\alpha dx^\alpha}. \quad (3.32)$$

扰动 x^α 为 $x^\alpha + \delta x^\alpha$, 相应的作用量的扰动量为

$$\delta S = -mc^2 \delta \int_a^b d\tau = -mc \delta \int_a^b \sqrt{dx_\alpha dx^\alpha}. \quad (3.33)$$

³频率与动量也构成四维波矢向量 $k^\mu \equiv (\omega/c, \mathbf{k})$, 这一点可以从波动相位 $\phi = \mathbf{k} \cdot \mathbf{x} - \omega t = -k^\mu x_\mu$ 为四维标量得到。具体见电磁波一章。

为了得到沿真实路径的运动方程，我们只需要将上式写成

$$\delta S = \int_a^b (\dots) \delta x^\alpha \quad (3.34)$$

的形式。利用最小作用量原理， $\delta S = 0$ ，便可得到括号里的式子即真实路径所满足的方程。

首先，我们计算 $\delta\sqrt{dx_\alpha dx^\alpha}$ 。

$$\delta\sqrt{dx_\alpha dx^\alpha} = \frac{1}{2ds} \delta(dx_\alpha dx^\alpha) = \frac{1}{2ds} (\delta dx_\alpha dx^\alpha + dx_\alpha \delta dx^\alpha) \quad (3.35)$$

注意我们可以交换上下哑指标，

$$\delta dx_\alpha dx^\alpha = \delta dx^\alpha dx_\alpha. \quad (3.36)$$

所以， $\delta\sqrt{dx_\alpha dx^\alpha} = dx_\alpha \delta dx^\alpha / ds$ 。同时 $ds = c d\tau$ ， $dx_\alpha / ds = u_\alpha / c$ ，并且

$$\delta\sqrt{dx_\alpha dx^\alpha} = u_\alpha \delta dx^\alpha / c = u_\alpha d\delta x^\alpha / c. \quad (3.37)$$

利用 $\delta\sqrt{dx_\alpha dx^\alpha} = u_\alpha d\delta x^\alpha / c$ ，我们有

$$\delta S = -m \int_a^b u_\alpha d\delta x^\alpha = -m u_\alpha \delta x^\alpha \Big|_a^b + m \int_a^b \delta x^\alpha \frac{du_\alpha}{d\tau} d\tau. \quad (3.38)$$

因为在扰动过程里，首尾的位置， $x^\alpha(a)$ 与 $x^\alpha(b)$ ，是固定不变的，所以

$$\delta S = 0 \Rightarrow m \frac{du_\alpha}{d\tau} = \frac{dp_\alpha}{d\tau} = 0 \quad (3.39)$$

即四维速度 u_α 或四维动量 p_α 沿自由粒子的世界线是个常数。

3.4 给定外场中电荷的运动方程

与经典力学类似，在外场中运动的电荷，其作用量可以写成如下形式：

$$S = S_p + S_{pf} \quad (3.40)$$

这里 $S_p = -mc^2 \int d\tau$ 是粒子的作用量，而 S_{pf} 则表达粒子和场的耦合或相互作用的作用量。现在要做的是找到 S_{pf} 的形式，使其符合两个条件：一， S_{pf} 应为一四维标量，二，相应运动方程应该与实验结果（洛伦兹运动方程）吻合。在本节里，我们先举一个例子，介绍从最小作用量原理得到相对论粒子在标量场中的运动方程。

例题 3.2 一个简单的粒子与一个标量场 ϕ 的耦合作用量可以表达为 $S_{pf} = \int \alpha c^2 \phi(x^\mu) d\tau$ ，这里 α 为粒子与该标量场的耦合系数，而 c^2 只是为了让最后的结果看起来简单一些。这个作用量是沿粒子在四维时空的世界线取值的， $\phi(x^\mu)$ ，且显然是一个四维标量。请利用最小作用量原理得到粒子在该场中的运动方程。

解 系统的作用量为

$$S = \int -mc^2 d\tau + \int \alpha c^2 \phi(x^\mu) d\tau \quad (3.41)$$

利用最小作用量原理，我们可以得到相应的协变性运动方程。显然

$$\delta S = \delta S_p + \delta S_{pf} = 0 \quad (3.42)$$

之前我们已经得到过 $\delta S_p = \int (dp_\mu / d\tau) \delta x^\mu d\tau$ 。因此，我们仅需得到 δS_{pf} 即可。从其定义，可知

$$\delta S_{pf} = \alpha c^2 \int (\delta \phi d\tau + \phi \delta d\tau). \quad (3.43)$$

利用前一节的结论, $\delta d\tau = (u_\mu/c^2)d\delta x^\mu$, 我们有

$$\delta S_{pf} = \int \alpha c^2 \partial_\mu \phi \delta x^\mu d\tau + \int \alpha \phi u_\mu d\delta x^\mu. \quad (3.44)$$

显然, 右边的第一项已经符合 $\int (\dots) \delta x^\mu d\tau$ 的形式。对于第二项, 我们利用 $\delta x^\mu(a) = \delta x^\mu(b) = 0$, 及分步积分可得

$$\int \alpha \phi u_\mu d\delta x^\mu = \alpha \phi u_\mu \delta x^\mu \Big|_a^b - \int \delta x^\mu d(\alpha \phi u_\mu) = - \int \delta x^\mu d(\alpha \phi u_\mu) = - \int \delta x^\mu \frac{d(\alpha \phi u_\mu)}{d\tau} d\tau. \quad (3.45)$$

因此, 我们有

$$\delta S = \int \left[\frac{dp_\mu}{d\tau} + \alpha c^2 \partial_\mu \phi - \frac{d(\alpha \phi u_\mu)}{d\tau} \right] \delta x^\mu d\tau = 0. \quad (3.46)$$

所以四维形式的粒子运动方程为

$$\frac{dp_\mu}{d\tau} - \frac{d(\alpha \phi u_\mu)}{d\tau} = \frac{d}{d\tau} [(m - \alpha \phi) u_\mu] = -\alpha c^2 \partial_\mu \phi. \quad (3.47)$$

注意到 $\partial_\mu \phi$ 为标量场 ϕ 的四维梯度, 以上方程与方程 3.15 高度相似, 只是加速度被替换成了 $d[(1 - \alpha \phi/m) u_\mu]/d\tau$ 。从某种角度看, 粒子与该标量场的部分作用相当于是其质量从 m 变为 $\bar{m} \equiv m - \alpha \phi$ 。令 $\mu = 0$, 利用 $d\tau = dt/\gamma$, 我们得到该方程的时间分量为

$$\frac{d\bar{\mathcal{E}}}{dt} = -\frac{\alpha c^2}{\gamma} \frac{\partial \phi}{\partial t}, \quad (3.48)$$

这里 $\bar{\mathcal{E}} \equiv \gamma \bar{m} c^2$ 。类似的, 相应的空间分量 ($\mu = 1, 2, 3$) 为

$$\frac{d\bar{\mathbf{p}}}{dt} = \frac{\alpha c^2}{\gamma} \nabla \phi. \quad (3.49)$$

这里 $\bar{\mathbf{p}} \equiv \gamma \bar{m} \mathbf{v}$

如果从上述作用量得到拉格朗日量, 利用欧拉-拉格朗日方程, 则可以得到关于 $\bar{\mathbf{p}}$ 的方程。但与能量相关的方程需要另外得到。

3.4.1 带电粒子与外场耦合作用量

上例中的标量场 ϕ 所得到的运动方程显然不能用来描述带电粒子在电磁场中的运动, 与实验结果 (洛伦兹运动方程) 并不一致。所以接下来我们尝试第二个例子: 粒子在一个用四维向量描述的场中的运动方程。

假定该向量场为 $A^\mu = (A^0, \mathbf{A})$, 粒子与场的耦合系数为 α 。电荷的电量因为该场为向量场, 所以沿粒子世界线积分的耦合作用量不可能是 $\int A^\mu d\tau$ 的形式, 因为这不是一个四维标量。合适的作用量应该是 $A_\mu dx^\mu$ 的形式。显然, 两种沿粒子路径的积分 $d\tau$ (或 ds) 与 dx^μ 的差别类似于三维情况下 dl (长度微分) 与 $d\mathbf{x}$ (位移微分) 之间的区别。这样代表耦合的作用量为

$$S_{pf} = \alpha \int_a^b A_\mu dx^\mu, \quad (3.50)$$

在该积分中, 场 A_μ 是沿粒子的世界线取值的。而粒子在该场中运用的总作用量为

$$S = \int_a^b (-mc^2 d\tau + \alpha A_\mu dx^\mu). \quad (3.51)$$

现在我们利用该作用量, 得到粒子的运动方程。

注意到 $A_\mu dx^\mu = cA_0 dt - \mathbf{A} \cdot d\mathbf{x} \equiv \phi dt - \mathbf{A} \cdot d\mathbf{x}$ 。这里我们定义 $\phi = cA_0$, 因此 $A^\mu = (\phi/c, \mathbf{A})$ 。利用 $d\mathbf{x}/dt = \mathbf{v}$, 可得粒子的拉格朗日量为

$$L = -mc^2/\gamma + \alpha \phi - \alpha \mathbf{A} \cdot \mathbf{v}. \quad (3.52)$$

显然我们有

$$\frac{\partial L}{\partial \mathbf{x}} = \nabla L = \alpha \nabla \phi - \alpha \nabla (\mathbf{A} \cdot \mathbf{v}). \quad (3.53)$$

注意到

$$\nabla (\mathbf{A} \cdot \mathbf{v}) = \underbrace{(\mathbf{A} \cdot \nabla) \mathbf{v}}_{=0} + (\mathbf{v} \cdot \nabla) \mathbf{A} + \mathbf{v} \times (\nabla \times \mathbf{A}) + \underbrace{\mathbf{A} \times (\nabla \times \mathbf{v})}_{=0}. \quad (3.54)$$

所以

$$\nabla L = \alpha \nabla \phi - \alpha (\mathbf{v} \cdot \nabla) \mathbf{A} - \alpha \mathbf{v} \times (\nabla \times \mathbf{A}). \quad (3.55)$$

另一方面,

$$\frac{\partial L}{\partial \mathbf{v}} = \mathbf{p} - \alpha \mathbf{A}. \quad (3.56)$$

因此, 从欧拉-拉格朗日方程, 我们得到粒子的运动方程为

$$\frac{d}{dt} (\mathbf{p} - \alpha \mathbf{A}) = -\alpha (\mathbf{v} \cdot \nabla) \mathbf{A} - \alpha \mathbf{v} \times (\nabla \times \mathbf{A}) + \alpha \nabla \phi \quad (3.57)$$

然而,

$$-\alpha \frac{d}{dt} \mathbf{A} = -\alpha \left[\frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{A} \right]. \quad (3.58)$$

所以,

$$\frac{d\mathbf{p}}{dt} = \alpha \frac{\partial \mathbf{A}}{\partial t} + \alpha \nabla \phi - \alpha \mathbf{v} \times (\nabla \times \mathbf{A}). \quad (3.59)$$

该方程显然与带电粒子在电磁场中的运动方程吻合。定义 $\alpha = -q$, 这里 q 为电荷电量, 而负号的出现仅仅是因为习惯上定义电子带负电荷, 而质子带正电荷。进一步的, 定义电场 \mathbf{E} 为

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi. \quad (3.60)$$

而磁场为

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (3.61)$$

利用 \mathbf{E} 和 \mathbf{B} , 粒子的运动方程可以写为

$$\frac{d\mathbf{p}}{dt} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}, \quad (3.62)$$

与实验以及带电粒子在电磁场中的运动方程完全一致, 右边的电磁场力称为洛伦兹力 (Lorentz force)。

小结一下, 带电子粒子在外加电磁场中的作用量为

$$S = \int_a^b (-mc^2 d\tau - qA_\mu dx^\mu). \quad (3.63)$$

这里的四维向量 $A^\mu = (\phi/c, \mathbf{A})$, 其时间分量 ϕ 称为电磁场的标势, 而空间分量 \mathbf{A} 称为电磁场的矢势。相应的拉格朗日量为

$$L = -mc^2/\gamma + q\mathbf{A} \cdot \mathbf{v} - q\phi. \quad (3.64)$$

作用量与拉格朗日量都由两部分构成, 对应着自由粒子以及粒子与场的耦合。

注意到在这本章的计算里, 我们均假设外场给定, 推导粒子的运动方程。这里一个隐含的假设是, 粒子对场的影响可以忽略不计, 这种计算也被称为测试粒子 (test particle) 计算。

3.4.2 协变形式运动方程

利用最小作用量原理，我们可以得到带电粒子在外场中运动的协变形式。对作用量 $S = S_p + S_{pf}$ 沿真实路径扰动 $x^\mu \rightarrow x^\mu + \delta x^\mu$ ，同时要求首尾点 $\delta x^\mu(a) = \delta x^\mu(b) = 0$ ，则最小作用量原理表明

$$\delta S = \delta S_p + \delta S_{pf} = 0. \quad (3.65)$$

与例 3.2 类似过程，从自由粒子运动方程中，我们已经有 $\delta S_p = \int (\mathbf{dp}_\mu / \mathbf{d}\tau) \delta x^\mu \mathbf{d}\tau$ 。因此，接下来我们仅需将 δS_{pf} 写为 $\int (\cdots) \delta x^\mu \mathbf{d}\tau$ 的形式即可。

注意到

$$\delta S_{pf} = \int_a^b (-q A_\mu \mathbf{d}x^\mu - q \delta A_\mu \mathbf{d}x^\mu). \quad (3.66)$$

右边第二项显然可以通过 $\delta A_\mu = (\partial A_\mu / \partial x^\nu) \delta x^\nu$ 写为

$$\int_a^b -q \delta A_\mu \mathbf{d}x^\mu = -q \int \frac{\partial A_\mu}{\partial x^\nu} \delta x^\nu \mathbf{d}x^\mu = -q \int \frac{\partial A_\mu}{\partial x^\nu} u^\nu \delta x^\mu \mathbf{d}\tau. \quad (3.67)$$

这里 $u^\mu = \mathbf{d}x^\mu / \mathbf{d}\tau$ 为四维速度，且上式形式与所要求的基本一致。但是注意到 δS_p 一式中， δx 的指标选用了 μ ，因此我们利用哑指标更换不影响结果这一性质，将上式中的 μ 与 ν 指标交换，所以

$$\int_a^b -q \delta A_\mu \mathbf{d}x^\mu = -q \int \frac{\partial A_\nu}{\partial x^\mu} u^\nu \delta x^\mu \mathbf{d}\tau. \quad (3.68)$$

而对 δS_{pf} 式右边第一项，我们显然需要使用分布积分并结合首尾点不扰动的特点将其改写为

$$\int_a^b -q A_\mu \mathbf{d}x^\mu = -q A_\mu \delta x^\mu \Big|_a^b + q \int_a^b \delta x^\mu \mathbf{d}A_\mu = q \int_a^b \delta x^\mu \frac{\partial A_\mu}{\partial x^\nu} \mathbf{d}x^\nu = q \int_a^b \frac{\partial A_\mu}{\partial x^\nu} u^\nu \delta x^\mu \mathbf{d}\tau. \quad (3.69)$$

与所求的形式完全一致。

由 δS_p 与 δS_{pf} 的表达式，我们可以得到

$$\delta S = \int_a^b \left[\frac{\mathbf{d}p_\mu}{\mathbf{d}\tau} - q \left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} \right) u^\nu \right] \delta x^\mu \mathbf{d}\tau \quad (3.70)$$

定义四维电磁场张量 $F^{\mu\nu}$ 为

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (3.71)$$

则

$$\delta S = \int_a^b \left[\frac{\mathbf{d}p_\mu}{\mathbf{d}\tau} - q F_{\mu\nu} u^\nu \right] \delta x^\mu \mathbf{d}\tau. \quad (3.72)$$

从最小作用量原理 $\delta S = 0$ ，我们得到相应的运动方程为

$$\frac{\mathbf{d}p_\mu}{\mathbf{d}\tau} = q F_{\mu\nu} u^\nu. \quad (3.73)$$

该方程即是带电粒子在给定外场中运动方程的协变形式，其左边为四维力，而右边为四维电磁场张量与粒子四维速度的点乘。

为了得到上述协变形式方程的时间与空间分量，我们首先得到四维电磁场张量 $F_{\mu\nu}$ 的各分量形式。从其定义可知， $F_{\mu\nu}$ 是反对称张量，即 $F_{\mu\nu} = -F_{\nu\mu}$ ，且我们有 $F_{\alpha\alpha} = 0$ ，此处不对 α 求和⁴。至于其它分量，从其定义且注意到 $A_\alpha = (\phi/c, -\mathbf{A})$ ，我们有

$$F_{10} = \frac{\partial A_0}{\partial x^1} - \frac{\partial A_1}{\partial x^0} = \frac{1}{c} \frac{\partial \phi}{\partial x} + \frac{1}{c} \frac{\partial A_x}{\partial t} = -E_x/c. \quad (3.74)$$

⁴实际上，在使用逆变分量与协变分量时，对指标求和的正确表达式应为 F_α^α 或 F^α_α ，对应着该张量的迹 (trace)，显然，对于一个反对称张量而言， $F_\alpha^\alpha = 0$ 。

依次类推，在国际单位制中，

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (3.75)$$

从 $F_{\mu\nu}$ ，我们也可以得到该张量的逆变形式

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (3.76)$$

显然的，我们有 $F_{0i} = (\mathbf{E})_i/c$ ，而 $F_{ij} = -\epsilon_{ijk}(\mathbf{B})_k$ 。注意，此处我们对电磁场 \mathbf{E} 与 \mathbf{B} 所采用的指标 i, j, k 对应的上下位置仅是一种表达上的方便，并不意味着电场 \mathbf{E} 或 \mathbf{B} 的逆变与协变分量，因为它们不构成四维向量。相反的，应将其简单理解为直角坐标系中相应物理量的 x, y, z 分量，且改变指标上下位置并不改变其正负号。这里我们选用 $(\mathbf{E})_i$ 或 $(\mathbf{E})^i$ 这样的表达方式，以区别于四维向量的空间分量，如 A_i 或 A^i 。利用 $F_{\mu\nu}$ ，我们可以得到协变形式运动方程 (3.73) 的各分量。对时间分量 ($\mu = 0$)，

$$\frac{dp_0}{d\tau} = qF_{0\mu}u^\mu = qF_{0i}u^i = \frac{q}{c}(\mathbf{E})_i\gamma v^i = \frac{q}{c}\gamma\mathbf{E} \cdot \mathbf{v}. \quad (3.77)$$

在右边第二步里，我们使用了 $F_{00} = 0$ 。注意到 $p_0 = \gamma mc = \mathcal{E}/c$ 以及 $d\tau = dt/\gamma$ ，我们有

$$\frac{d\mathcal{E}}{dt} = q\mathbf{E} \cdot \mathbf{v}. \quad (3.78)$$

因此协变方程的时间分量对应于粒子能量的变化来源于电磁场所做的功 $\mathbf{E} \cdot \mathbf{v}$ 。类似的，对于其空间分量，我们有

$$\frac{dp_i}{d\tau} = qF_{i\mu}u^\mu = qF_{i0}u^0 + qF_{ij}u^j = -q(\mathbf{E})_i\gamma - q\epsilon_{ijk}(\mathbf{B})_k\gamma v^j \quad (3.79)$$

$$= \gamma[-q(\mathbf{E})_i + q\epsilon_{ikj}(\mathbf{B})_k v^j] = q\gamma(-\mathbf{E} + \mathbf{B} \times \mathbf{v})_i. \quad (3.80)$$

注意到 $p_\mu = (\mathcal{E}/c, -\mathbf{p})$ ，因此 $p_i = -(\mathbf{p})_i$ 。这样协变形式运动方程的空间分量为

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (3.81)$$

对应着洛伦兹运动方程 (3.62)。

3.5 电磁场势函数

3.5.1 势函数的变换规律

在上一节中我们引入了电磁场标势 ϕ 与矢势 \mathbf{A} ，且它们构成四维向量 $A^\mu = (\phi/c, \mathbf{A})$ 。因此，我们很容易知道这两个物理量在不同参考系间的变换规律，即

$$\phi = \gamma(V)(\phi' + c\beta A'_x), \quad (3.82)$$

$$A_x = \gamma(V)[A'_x + (\beta/c)\phi'], \quad (3.83)$$

$$A_y = A'_y, \quad (3.84)$$

$$A_z = A'_z. \quad (3.85)$$

当然, 这里假设 K' 以速度 $\mathbf{V} = V\mathbf{e}_x$ 相对于 K 运动。需要注意的是, 这里 \mathbf{A} 与 ϕ 是 (t, \mathbf{x}) 的函数, 而 ϕ' 与 \mathbf{A}' 是 (t', \mathbf{x}') 的函数。所以如果我们希望通过 \mathbf{A} 与 ϕ 得到电磁场 \mathbf{E} 与 \mathbf{B} , 则还需要注意时空坐标的变换关系。

例题 3.3 利用标势与矢势的变换规律, 得到匀速运动带电粒子的电磁场。

解 我们通过如下两步来分析该问题,

1. 在电荷静止的参考系 (K') 里, 仅仅有电场 \mathbf{E}' , 而磁场 $\mathbf{B}' = 0$ 。
2. 利用电磁场的变换规律得到实验室参考系 K 系中的电磁场。

在 K' 系里, 假设该静止的点电荷 q 位于坐标原点 $x' = y' = z' = 0$ 。则其势函数为

$$\phi' = \frac{q}{4\pi\epsilon_0 r'} \quad \text{and} \quad \mathbf{A}' = 0. \quad (3.86)$$

我们可以通过洛伦兹变换, 从 ϕ' 与 \mathbf{A}' 得到 K 系中的 ϕ 和 \mathbf{A} ,

$$\phi/c = \gamma(\phi'/c + \beta A'_x) = \frac{\phi'(\mathbf{r}')/c}{\sqrt{1 - v^2/c^2}}, \quad (3.87)$$

$$A_x = \gamma(A'_x + \beta\phi'/c) = \frac{(v/c^2)\phi'(\mathbf{r}')}{\sqrt{1 - v^2/c^2}}, \quad (3.88)$$

$$A_y = A'_y = 0, \quad (3.89)$$

$$A_z = A'_z = 0, \quad (3.90)$$

因为 $A^\alpha = (\phi/c, \mathbf{A})$ 是一个四维向量。注意到方程右边的量都是 t' 与 \mathbf{r}' 的函数。而为了从 ϕ 和 \mathbf{A} 得到 K 系中的电磁场, 我们需要对 t 和 \mathbf{r} 求导。所以我们还需要再进行一步, 将 ϕ 和 \mathbf{A} 写成 t 和 \mathbf{r} 的函数。我们知道 $r' = \sqrt{x'^2 + y'^2 + z'^2}$ 且

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \quad (3.91)$$

$$y' = y, \quad (3.92)$$

$$z' = z. \quad (3.93)$$

所以

$$r' = \sqrt{\frac{(x - vt)^2}{1 - v^2/c^2} + y^2 + z^2}. \quad (3.94)$$

利用方程 (3.94), 将 $r' = r'(x, y, z)$ 代入 ϕ 和 \mathbf{A} , 方程 (3.87)-(3.90) 得出

$$\phi = \frac{q}{4\pi\epsilon_0 r' \sqrt{1 - v^2/c^2}} \quad (3.95)$$

$$= \frac{q}{4\pi\epsilon_0 \sqrt{(x - vt)^2 + (1 - v^2/c^2)(y^2 + z^2)}} \quad (3.96)$$

$$\equiv \frac{q}{4\pi\epsilon_0 r^*}, \quad (3.97)$$

$$A_x = \frac{v}{c^2} \phi, \quad (3.98)$$

$$A_y = A_z = 0. \quad (3.99)$$

从 ϕ 和 \mathbf{A} , 我们现在可以直接得到 \mathbf{E} 与 \mathbf{B} 。

注意到 $r^* = \sqrt{(x - vt)^2 + (1 - v^2/c^2)(y^2 + z^2)}$,

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} = -\nabla\phi - \frac{v}{c^2} \frac{\partial\phi}{\partial t} \mathbf{e}_x. \quad (3.100)$$

这里 $\phi = q/4\pi\epsilon_0 r^*$, 所以我们有

$$\frac{\partial\phi}{\partial x} = -\frac{q}{4\pi\epsilon_0 r^{*2}} \frac{\partial r^*}{\partial x} = -\frac{q}{4\pi\epsilon_0 r^{*3}} (x - vt) \quad (3.101)$$

$$\frac{\partial\phi}{\partial y} = -\frac{q}{4\pi\epsilon_0 r^{*2}} \frac{\partial r^*}{\partial y} = -\frac{q}{4\pi\epsilon_0 r^{*3}} \left(1 - \frac{v^2}{c^2}\right) y \quad (3.102)$$

$$\frac{\partial\phi}{\partial z} = -\frac{q}{4\pi\epsilon_0 r^{*2}} \frac{\partial r^*}{\partial z} = -\frac{q}{4\pi\epsilon_0 r^{*3}} \left(1 - \frac{v^2}{c^2}\right) z \quad (3.103)$$

$$\frac{\partial\phi}{\partial t} = -\frac{q}{4\pi\epsilon_0 r^{*2}} \frac{\partial r^*}{\partial t} = -\frac{q}{4\pi\epsilon_0 r^{*3}} (x - vt)(-v) \quad (3.104)$$

将方程 (3.101)-(3.104) 代入 (3.100) 得到

$$\mathbf{E} = -\nabla\phi - \frac{v}{c^2} \frac{\partial\phi}{\partial t} \mathbf{e}_x \quad (3.105)$$

$$\begin{aligned} &= \frac{q}{4\pi\epsilon_0 r^{*3}} (x - vt) \mathbf{e}_x + \frac{q}{4\pi\epsilon_0 r^{*3}} \left(1 - \frac{v^2}{c^2}\right) y \mathbf{e}_y \\ &\quad + \frac{q}{4\pi\epsilon_0 r^{*3}} \left(1 - \frac{v^2}{c^2}\right) z \mathbf{e}_z - \frac{v^2}{c^2} \frac{q}{4\pi\epsilon_0 r^{*3}} (x - vt) \mathbf{e}_x \end{aligned} \quad (3.106)$$

$$= \frac{q}{4\pi\epsilon_0 r^{*3}} \left(1 - \frac{v^2}{c^2}\right) [(x - vt) \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z] \quad (3.107)$$

$$= \left(1 - \frac{v^2}{c^2}\right) \frac{q}{4\pi\epsilon_0 r^{*3}} \mathbf{R}, \quad (3.108)$$

这里 $\mathbf{R} \equiv (x - vt) \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z$ 是从电荷当前时刻 (t) 的位置指向场点 (x, y, z) 的向量, 因为粒子轨迹是 $\mathbf{x} = vt \mathbf{e}_x$.

类似的过程很容易得到 $\mathbf{B} = (\mathbf{v}/c^2) \times \mathbf{E}$. 接下来我们仅仅讨论电场. 引入角度 θ , 其定义为向量 \mathbf{R} 和运动方向 (\mathbf{e}_x) 间的夹角. 这样

$$R_{\parallel} = x - vt, \quad (3.109)$$

$$R_{\perp} = y^2 + z^2 = R^2 \sin^2 \theta, \quad (3.110)$$

$$r^{*2} = R_{\parallel}^2 + R_{\perp}^2 = R^2 \left(1 - \frac{v^2}{c^2} \sin^2 \theta\right). \quad (3.111)$$

利用方程 (3.109)-(3.111), 我们将方程 (3.108) 中的 \mathbf{E} 表达成 θ 和 \mathbf{R} 的函数, 即

$$\mathbf{E} = \left(1 - \frac{v^2}{c^2}\right) \frac{q \mathbf{R}}{4\pi\epsilon_0 R^3 [1 - (v^2/c^2) \sin^2 \theta]^{3/2}} \quad (3.112)$$

$$= \frac{q \mathbf{R}}{4\pi\epsilon_0 R^3} \frac{1 - \beta^2}{(1 - \beta^2 \sin^2 \theta)^{3/2}}. \quad (3.113)$$

对一个给定距离 R , 电场 \mathbf{E} 的值在 $\theta = \pi/2$ (垂直方向) 达到最大,

$$E_{\perp} = E_{\theta=\pi/2} = \frac{q}{4\pi\epsilon_0 R^2} (1 - \beta^2)^{-1/2} = \frac{q}{4\pi\epsilon_0 R^2} \gamma, \quad (3.114)$$

且在平行方向 ($\theta = 0$ or π) 最小,

$$E_{\parallel} = E_{\theta=0} = E_{\theta=\pi} = \frac{q}{4\pi\epsilon_0 R^2} (1 - \beta^2) = \frac{1}{\gamma^2} \frac{q}{4\pi\epsilon_0 R^2}. \quad (3.115)$$

所以随着粒子速度的增加, 对距离电荷相同的点而言, E_{\parallel} 减少, 且 E_{\perp} 增加. 换句话说, 电场在平行于运动方向上收缩了. 如果 $v \rightarrow c$, 则 $E_{\parallel} \rightarrow 0$ 且 $E_{\perp} \rightarrow \infty$, 因为

$$\mathbf{E} = \frac{q \mathbf{R}}{4\pi\epsilon_0 R^3} \frac{1 - \beta^2}{(1 - \beta^2 \sin^2 \theta)^{3/2}}. \quad (3.116)$$

随着 $\beta \rightarrow 1$, 分母上也会在 $\Delta\theta$ 的小角度上 $1 - \beta^2 \sin^2 \theta \rightarrow 0$, 即对一给定距离而言, \mathbf{E} 仅仅在一很小角度非常大, 其它角度近似为 0.

3.5.2 规范不变性

标势与矢势的一个重要性质是它们不是唯一的。我们知道，描述电磁场对带电粒子的作用时，即可以用势函数 A^α （如在作用量 S_{pf} 中），或者像在运动方程里，仅仅使用 \mathbf{E} 与 \mathbf{B} ，

$$\frac{d\mathbf{p}}{dt} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}. \quad (3.117)$$

但这两类物理量其实不太一样。比如，从实验室里测量的角度来看，我们可以通过粒子的受力，完全地定出 \mathbf{E} 和 \mathbf{B} 。然而，我们没有办法用同样的方法定出 \mathbf{A} 和 ϕ ，因为势函数是通过时空导数进入粒子运动方程的。利用 \mathbf{E} 与 \mathbf{B} 的定义，如果我们对 A^α 作如下的替换

$$\mathbf{A}' = \mathbf{A} + \nabla f \quad (3.118)$$

$$\phi' = \phi - \frac{\partial f}{\partial t} \quad (3.119)$$

完全不会改变 \mathbf{E} 和 \mathbf{B} ，从而不会改变粒子的受力和运动方程。或者说运动方程在势函数的这种变换下具有不变性，这种特殊的不变性称为规范不变性（gauge invariance）或规范对称性。规范不变性是电磁场与现代场论的重要核心之一，且与时间对称性或平移对称性对应着系统的能量与动量守恒类似，规范对称性也对应着某类物理量的守恒，具体会在下一章讨论。

势函数 A^α 的规范不变性，在作用量里也可以体现。因为作用量不是唯一的。例如假如有一标量 $f = f(x^\alpha)$ ，则耦合作用量

$$S_{pf} = -q \int_a^b A_\alpha dx^\alpha \quad (3.120)$$

与

$$S'_{pf} = -q \int_a^b A'_\alpha dx^\alpha = -q \int_a^b A_\alpha dx^\alpha + q \int_a^b df = S_{pf} + q \int_a^b df \quad (3.121)$$

对应着同样的运动过程，因数 $\delta S_{pf} = 0$ 与 $\delta S'_{pf} = 0$ 所得到的运动方程是一样的。然而，两者所用到的势函数不同，且

$$A'_\alpha = A_\alpha - \partial_\alpha f. \quad (3.122)$$

该式的时空分量则给出方程 (3.118)-(3.119) 中的规范变换。

规范不变性允许我们选定一个额外条件去进一步限制势函数 ϕ 和 \mathbf{A} ，而选定一个条件对应着选定一个规范。实际工作中，常见的两种规范条件是库仑规范 (Coulomb gauge) 和洛伦茨规范 (Lorenz gauge)⁵。库仑规范是指

$$\nabla \cdot \mathbf{A} = 0. \quad (3.123)$$

该规范的优点是可以很容易的将电场分为两部分：无源场与无旋场，或者称为横场 (transverse field, \mathbf{E}_T) 和纵场 (longitudinal field, \mathbf{E}_L)⁶，其中

$$\mathbf{E}_T = -\frac{\partial \mathbf{A}}{\partial t}, \quad (3.124)$$

$$\mathbf{E}_L = -\nabla \phi \quad (3.125)$$

⁵这里的 Lorenz 是来自丹麦的物理学家和数学家 Ludvig Lorenz，而非提出洛伦兹变换的 Hendrik Lorentz。为了区分两人，现行做法通常是将 Lorenz 译为洛伦茨，而将 Lorentz 译为洛伦兹。

⁶这里横场与纵场的名称可以从波动问题中理解：在平面波分析中，常常使用复数表达场的量。通过傅立叶变换，可将 ∇ 替换为 $i\mathbf{k}$ ，这里 \mathbf{k} 为波矢。所以 \mathbf{E}_T 满足 $\mathbf{k} \cdot \mathbf{E}_T = 0$ ，即 $\mathbf{k} \perp \mathbf{E}_T$ ，这样波动传播的方向 \mathbf{k} 与振动物理量 (\mathbf{E}_T) 的方向垂直，对应着横波。而对 \mathbf{E}_L ，则有 $\mathbf{k} \parallel \mathbf{E}_L$ ，即波动传播方向与振动物理量 (\mathbf{E}_L) 方向相平行，对应纵波。

这样, $\nabla \cdot \mathbf{E}_T = 0$ 且 $\nabla \times \mathbf{E}_L = 0$ 。要理解库伦规范总是可以选择的, 现假定原本的矢势 \mathbf{A} 满足 $\nabla \cdot \mathbf{A} = g$, 则我们需要做的是选择一个规范变换, $\mathbf{A}' = \mathbf{A} + \nabla f$, 使得 $\nabla \cdot \mathbf{A}' = 0$, 因此等价于要求我们找到了一个 f , 使其满足 $0 = \nabla \cdot \nabla f + g$ 或者说 $\nabla^2 f = -g$, 即 f 满足非齐次的拉普拉斯方程 (泊松方程) 就可以使得新的矢势函数满足库伦规范, 当然同时标势 $\phi' = \phi - \partial f / \partial t$ 。

另一个重要的规范, 洛伦茨规范, 在理论分析中常常特别有用。该规范的限制条件是

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0. \quad (3.126)$$

这个规范的特别之处在于其对应着四维形式

$$\frac{\partial A^\alpha}{\partial x^\alpha} = 0. \quad (3.127)$$

即四维势函数的散度为 0。注意上述方程是一个四维标量方程, 即在所有惯性参考系中都是不变的。所以如果 \mathbf{A} 和 ϕ 在一个惯性参考系中满足洛伦茨规范, 则它们在所有其它的惯性参考系中都满足该规范。同样的, 洛伦茨规范总是可以选择的。假定原本的势函数 A^α 满足 $\partial_\alpha A^\alpha = g$, 则需要作的是选择一个规范变换, $A'^\alpha = A^\alpha - \partial^\alpha f$, 使得 $\partial_\alpha A'^\alpha = 0$, 因此等价于要求我们找到了一个 f , 使其满足

$$0 = g - \partial_\alpha \partial^\alpha f \Rightarrow \partial_\alpha \partial^\alpha f = g. \quad (3.128)$$

即

$$\frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} - \nabla^2 f = g. \quad (3.129)$$

这是一个非齐次的波动方程。以上过程也让我们可以看到洛伦茨规范另一个有趣的性质: 它不能唯一地确定势函数, 因为我们可以进一步作一个规范变换 $A''^\alpha = A'^\alpha - \partial^\alpha h$ 而不改变电磁场 \mathbf{E} 与 \mathbf{B} , 只要 h 满足 $\partial_\alpha \partial^\alpha h = 0$ 即可。

3.6 电磁场张量

在本章中引入的另一个重要概念是电磁场张量, 其各分量为电场与磁场。从这一张量我们可以看到, 电场与磁场本质上为同一物理量 $F^{\mu\nu}$ 的不同分量, 在不同的参考系里体现出不同的性质而已。理解了这一点, 也就可以在概念上理解爱因斯坦在其论文《论动体的电动力学》的摘要里所提出的问题。在下一章中我们会进一步看到该电磁场张量在电磁场理论中扮演关键角色。

3.6.1 电磁场的变换规律

本节里, 我们利用电磁场张量来得到电磁场在不同参考系间的变换规律。注意到 $F^{\mu\nu}$ 是一个反对称四维张量, 对电场 \mathbf{E} ,

$$E_x = E'_x, \quad E_y = \gamma(E'_y + c\beta B'_z), \quad E_z = \gamma(E'_z - c\beta B'_y). \quad (3.130)$$

而对磁场 \mathbf{B} ,

$$B_x = B'_x, \quad B_y = \gamma[B'_y - (\beta/c)E'_z], \quad B_z = \gamma[B'_z + (\beta/c)E'_y]. \quad (3.131)$$

如果两个参考系间的相对速度方向是任意, 则利用相应的洛伦兹变换, 可以得到

$$\mathbf{E} = \gamma(\mathbf{E}' + c\mathbf{B}' \times \boldsymbol{\beta}) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{E}'), \quad (3.132)$$

$$\mathbf{B} = \gamma[\mathbf{B}' - (\mathbf{E}'/c) \times \boldsymbol{\beta}] - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta}(\boldsymbol{\beta} \cdot \mathbf{B}'). \quad (3.133)$$

这种变换的非相对论形式常常见到。在 $V/c \ll 1$ 时，上述变换式为

$$\mathbf{E} = \mathbf{E}' + c\mathbf{B}' \times \boldsymbol{\beta}, \quad (3.134)$$

$$\mathbf{B} = \mathbf{B}' - (\mathbf{E}'/c) \times \boldsymbol{\beta}. \quad (3.135)$$

这里 $\boldsymbol{\beta} = \mathbf{V}/c, \gamma = 1/\sqrt{1-\beta^2}$. 从以上变换式可以看出，在一个参考系里的电场可以转换成另一个参考系里的磁场，反之亦然。进一步地体现了电场与磁场是紧密联系在一起的。

3.6.2 两个重要的洛伦兹不变量

除了上述电磁场变换式之外，还可以通过电磁场张量 $F^{\alpha\beta}$ 引入两个与电磁场相关的重要洛伦兹不变量。利用这两个不变量，有时可以帮助我们快速地确定变换后电磁场的可能性质。第一个不变量是通过四维张量 $F_{\alpha\beta}$ 双点乘自己得到

$$F_{\alpha\beta}F^{\alpha\beta} = 2[B^2 - (E/c)^2]. \quad (3.136)$$

因此 $B^2 - (E/c)^2$ 为一洛伦兹不变量，限定了电磁场大小的相对关系在不同参考系间保持不变。为了得到第二个重要的洛伦兹不变量，我们首先引入 $e^{\alpha\beta\mu\nu}$ ，四阶完全反对称单位张量。张量 $e^{\alpha\beta\mu\nu}$ 有如下性质，

1. $e^{0123} = 1$.
2. 如果调换任意一对指标，则 $e^{\alpha\beta\mu\nu}$ 改变符号。
3. 如果任意两个指标相同，则 $e^{\alpha\beta\mu\nu} = 0$ 。

利用 $e^{\alpha\beta\mu\nu}$ ，我们首先构建一个新的四维张量， $\mathcal{F}^{\alpha\beta}$ ，其定义为

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2}e^{\alpha\beta\mu\nu}F_{\mu\nu} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix} \quad (3.137)$$

这一张量称为电磁场张量的对偶张量。显然，该张量的分量相当于是将 $F^{\alpha\beta}$ 中的 \mathbf{E}/c 替换成 \mathbf{B} ，且将 \mathbf{B} 替换成 $-\mathbf{E}/c$ 。将 $\mathcal{F}^{\alpha\beta}$ 与电磁场张量双点乘可得

$$\mathcal{F}^{\mu\nu}F_{\mu\nu} = -\frac{2}{c}\mathbf{E} \cdot \mathbf{B}, \quad (3.138)$$

因此 $\mathbf{E} \cdot \mathbf{B}$ 也为一洛伦兹不变量，该不变量在 $\mathbf{E} \cdot \mathbf{B} = 0$ 时特别有用。利用上述两个洛伦兹不变量，我们有如下结论：

1. 如果在 K 系中， $E > cB$ ，则在所有其它参考系中 $E > cB$ ，反过来也一样。
2. 如果在 K 系中， $\mathbf{E} \perp \mathbf{B}$ ，则我们可以找到一个 K' 系，使得
 - 如果 $B^2 - (E/c)^2 > 0$ ，则 $\mathbf{E} = 0$ 。
 - 如果 $B^2 - (E/c)^2 < 0$ ，则 $\mathbf{B} = 0$ 。
3. 如果中 K 系中， $\mathbf{E} \perp \mathbf{B}$ ，则在其它任意 K' 系中，如果 $E' \neq 0$ 且 $B' \neq 0$ ，都有 $\mathbf{E}' \perp \mathbf{B}'$ 。

利用这两个洛伦兹不变量，我们很容易理解匀速运动的带电粒子，其电场与磁场一定是垂直的， $\mathbf{E} \cdot \mathbf{B} = 0$ ，且 $B < E/c$ ，因为在粒子静止的参考系里 $\mathbf{B}' = 0$ 。这一点与例 3.3 中的结论 $\mathbf{B} = (\mathbf{v}/c^2) \times \mathbf{E}$ 相符。

第四章 电磁场方程

4.1 拉格朗日方程：从离散到连续

考虑一维的沿 x 方向的均匀密度与弹性的棒子。如果有一扰动，则该扰动会沿棒子以波动的形式沿弹性棒传播。相应的物理量，如棒子上某一点的位移量 η 或密度的扰动量 $\delta\rho$ ，均满足一维的波动方程

$$\frac{\partial^2 \eta}{\partial t^2} - v^2 \frac{\partial^2 \eta}{\partial x^2} = 0. \quad (4.1)$$

通常这一方程在力学里是通过分析一个小微元的运动方程得到的。现在我们来从最小作用量原理出发，得到上述方程。

作为对连续的弹性棒的近似，我们先考虑一由 N 个一维的弹簧振子构成的离散系统，每个振子的平衡态长度为 $\Delta x = L/N$ 。弹簧本身无质量，但是振子的质量为 m ，且 $m = \rho \Delta x$ ，这里 ρ 为棒子静止时的密度。现取第 i 个振子偏离其初始位置的距离为 η_i ，则其偏离平衡位置的量为 $\eta_{i+1} - \eta_i$ ，因此具有的势能为 $V_i = k(\eta_{i+1} - \eta_i)^2/2$ ，而动能为 $T_i = m\dot{\eta}_i^2/2$ ，因此对第 i 个振子而言，其拉格朗日量为

$$L_i = \frac{m}{2} \dot{\eta}_i^2 - \frac{k}{2} (\eta_{i+1} - \eta_i)^2. \quad (4.2)$$

这样整个系统的拉格朗日量为

$$L = \sum_{i=1}^N \left[\frac{m}{2} \dot{\eta}_i^2 - \frac{k}{2} (\eta_{i+1} - \eta_i)^2 \right]. \quad (4.3)$$

这是一个自由度为 N 的离散体系拉格朗日量，为了方便下一步，我们将该拉格朗日量写为

$$L = \frac{1}{2} \sum_{i=1}^N \Delta x \left[\frac{m}{\Delta x} \dot{\eta}_i^2 - k \left(\frac{\eta_{i+1} - \eta_i}{\Delta x} \right)^2 \right]. \quad (4.4)$$

现在令 $\Delta x \rightarrow 0$ ，这一过程也等价于 $N \rightarrow \infty$ ，即自由度从有限值趋向无穷大，则显然有

$$\frac{\eta_{i+1} - \eta_i}{\Delta x} \rightarrow \frac{\partial \eta}{\partial x} \quad (4.5)$$

且对每一个小弹簧振子的求和变为对 x 的积分，即 $\sum_i \Delta x \rightarrow \int dx$ ，而相应的拉格朗日量为

$$L = \frac{1}{2} \int dx \left[\rho \left(\frac{\partial \eta}{\partial t} \right)^2 - \mu \left(\frac{\partial \eta}{\partial x} \right)^2 \right] = \int \mathcal{L} dx \quad (4.6)$$

注意这里我们用 $\partial \eta / \partial t$ 替换了 $\dot{\eta}$ ，且 $\mu \equiv \lim_{\Delta x \rightarrow 0} k \Delta x$ 。后面通过对比相应的方程，我们会知道 μ 的含义为弹性棒的杨氏模量。而

$$\mathcal{L} = \frac{1}{2} \left[\rho \left(\frac{\partial \eta}{\partial t} \right)^2 - \mu \left(\frac{\partial \eta}{\partial x} \right)^2 \right] \quad (4.7)$$

则是系统的拉格朗日量密度。

从上述例子及其分析过程中我们可以看到，对上述离散系统，拉氏量是 η_i 与 $\dot{\eta}_i$ 的函数，对应着 N 个自由度与关于时间 t 的常微分方程。而对连续体系，拉氏量是 $\partial \eta / \partial t$ 与 $\partial \eta / \partial x$ 的函数， x 作为独立变量代替了 i 的位置，而相应的方程则为一个关于 $\eta(t, x)$ 的偏微分方程。在更一般的一维情况下，拉格朗日密度则为 $\mathcal{L} = \mathcal{L}(\eta, \partial_t \eta, \partial_x \eta, t, x)$ 。当然 \mathcal{L} 也可能是 η 更高阶导数的函数，但本课里，我们只处理 \mathcal{L} 最多是 η 一阶时空导数的情况。注意在拉格朗日密度的定义里， $\eta, \partial_t \eta$ 与 $\partial_x \eta$ 是独立的自变量。

利用最小作用量原理，我们可以从上述作用量得到相应的拉格朗日方程。该系统的作用量为

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \mathcal{L}(\eta, \partial_t \eta, \partial_x \eta, t, x) \quad (4.8)$$

为了得到关于 η 的方程，我们对 η 作无穷小的扰动， $\eta \rightarrow \eta + \delta\eta$ ，且固定其在时间与空间边界上的值，即 $\delta\eta(x = x_1, x_2) = 0$ 且 $\delta\eta(t = t_1, t_2) = 0$ 。则依变分原理，我们有相应的作用量扰动 $\delta S = \int dt \int dx \delta\mathcal{L} = 0$ ，即

$$\delta S = \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \left[\frac{\partial \mathcal{L}}{\partial \eta} \delta\eta + \frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \delta(\partial_t \eta) + \frac{\partial \mathcal{L}}{\partial(\partial_x \eta)} \delta(\partial_x \eta) \right] = 0. \quad (4.9)$$

与获得离散体系运动方程类似，我们的目标是将上式写成 $\int dt \int dx (\cdots) \delta\eta = 0$ 的形式。注意到 $\delta(\partial_t \eta) = \partial_t(\delta\eta)$ ， $\delta(\partial_x \eta) = \partial_x(\delta\eta)$ 我们可以对上式右边的二、三项进行分布积分。例如，

$$\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \partial_t(\delta\eta) \right] = \int_{x_1}^{x_2} dx \int_{t_1}^{t_2} dt \left\{ \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \delta\eta \right] - \delta\eta \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \right] \right\} \quad (4.10)$$

$$= \int_{x_1}^{x_2} dx \left. \frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \delta\eta \right|_{t_1}^{t_2} - \int_{x_1}^{x_2} dx \int_{t_1}^{t_2} dt \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \right] \delta\eta = - \int_{x_1}^{x_2} dx \int_{t_1}^{t_2} dt \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \right] \delta\eta. \quad (4.11)$$

在上式最后一步，我们利用了最小作用量原理中，扰动量在边界上的变分为 0，即 $\delta\eta(t = t_1, t_2) = 0$ 。用类似的方式处理式 4.9 右边第三项，则

$$\delta S = \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \left\{ \frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \right] - \frac{\partial}{\partial x} \left[\frac{\partial \mathcal{L}}{\partial(\partial_x \eta)} \right] \right\} \delta\eta = 0. \quad (4.12)$$

因为对 η 的变分扰动是任意的，我们有该一维连续系统的拉格朗日方程为

$$\frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial}{\partial t} \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} \right] - \frac{\partial}{\partial x} \left[\frac{\partial \mathcal{L}}{\partial(\partial_x \eta)} \right] = 0. \quad (4.13)$$

对于上例中的一维简单弹性棒，利用其拉格朗日量密度，式 4.7，我们有

$$\frac{\partial \mathcal{L}}{\partial \eta} = 0, \quad \frac{\partial \mathcal{L}}{\partial(\partial_t \eta)} = \rho \frac{\partial \eta}{\partial t}, \quad \frac{\partial \mathcal{L}}{\partial(\partial_x \eta)} = -\mu \frac{\partial \eta}{\partial x}. \quad (4.14)$$

代入拉格朗日方程，式 4.13，我们得到相应的运动方程为

$$\frac{\partial^2 \eta}{\partial t^2} - v^2 \frac{\partial^2 \eta}{\partial x^2} = 0, \quad (4.15)$$

这是一个典型的波动方程，其中 $v = \sqrt{\mu/\rho}$ 为波动的传播速度。从这里也可以看出 $\mu = \lim_{\Delta x \rightarrow 0} k \Delta x$ 为该弹性棒的弹性模量。

4.2 场的最小作用量原理与拉格朗日方程

4.2.1 一维标量场方程

以上利用最小作用量原理得到一个力学连续系统方程的过程同样可以运用在描述场的行为上。通常一个场随时间与空间变化的物理量，即 $\phi = \phi(t, \mathbf{x})$ ，其对应的拉格朗日量密度一般为 $\mathcal{L} = \mathcal{L}(\phi, \partial_t \phi, \partial_{\mathbf{x}} \phi, t, \mathbf{x})$ 。利用 $x^\mu = (ct, \mathbf{x})$ ，拉格朗日量密度也可记为 $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi, x^\mu)$ ，这不仅仅是记法的方便，而是更有助于直接得到符合相对论协变规律的场的方程。通常情况下，如果是无可相互作用粒子存在时的自由场 (free field)，则拉格朗日量密度为 $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$ 。

作为一个简单的例子，我们先介绍一个一维的标量场，并通过最小作用量原理得到该场的拉格朗日方程。假定一个标量场， $\phi(t)$ ，与空间无关，仅与时间有关。则其作用量可以通过拉格朗日量写成

$$S = \int \mathcal{L}(\phi, \dot{\phi}) dt. \quad (4.16)$$

这里 $\dot{\phi} = d\phi/dt$. 现在我们需要通过最小作用量原理来得到一个场的方程, 也就是 ϕ 是满足何种方程, 则我们可以沿着真实的场来扰动 $\phi \rightarrow \phi + \delta\phi$. 最小作用量原理会表明, $\delta S = 0$. 依照之前分析力学中推导欧拉-拉格朗日方程的过程, 我们可有

$$\delta S = \int \left[\mathcal{L} \left(\phi + \delta\phi, \frac{d}{dt}(\phi + \delta\phi) \right) - \mathcal{L}(\phi, \dot{\phi}) \right] dt = 0, \quad (4.17)$$

简单化简后可得

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \delta\dot{\phi} \right] dt = 0. \quad (4.18)$$

对括号里的第二项用分步积分法, 并注意到边界处 (无穷远的过去和将来), ϕ 是固定不扰动的, 我们得到

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \phi} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) \right] \delta\phi dt = 0. \quad (4.19)$$

从而得到该一维标量场的拉格朗日方程为

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (4.20)$$

这个方程与粒子的拉格朗日方程完全一致. 其实这个也不难理解, 方程 (4.16) 完全和粒子的作用量类似, 即 $S = \int L(q, \dot{q}) dt$. 只不过是 q 换成了 ϕ .

4.2.2 与时间无关的标量场

在这一小节里, 我们假设标量场 $\phi = \phi(x, y, z)$ 与时间无关. 通过最小作用量原理, 我们可以得到相应的拉格朗日方程. 对该场而言, $\mathcal{L} = \mathcal{L}(\phi, \partial_x \phi, \partial_y \phi, \partial_z \phi)$ 且作用量

$$S = \int \mathcal{L}(\phi, \partial_x \phi, \partial_y \phi, \partial_z \phi) dx dy dz dt \quad (4.21)$$

为了得到场的方程, 我们扰动 $\phi \rightarrow \phi + \delta\phi$, 但同时保持边界上 $\delta\phi = 0$. 这样所有的空间导数 ($\partial_j \phi$) 也会发生扰动. 所以作用量的扰动为

$$\delta S = \frac{1}{c} \int \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial (\partial_x \phi)} \delta(\partial_x \phi) + \frac{\partial \mathcal{L}}{\partial (\partial_y \phi)} \delta(\partial_y \phi) + \frac{\partial \mathcal{L}}{\partial (\partial_z \phi)} \delta(\partial_z \phi) \right) d\Omega, \quad (4.22)$$

$$= \frac{1}{c} \int \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} \partial_i \delta\phi \right) d\Omega \quad (4.23)$$

为了获得场 ϕ 的方程, 我们将把作用量的扰动写成

$$\delta S = \int (\dots) \delta\phi \quad (4.24)$$

的形式. 这样依据最小作用量原理, 括号内的式子即是我们要的场的演化方程。

令

$$\mathbf{A} = \left(\frac{\partial \mathcal{L}}{\partial (\partial_x \phi)}, \frac{\partial \mathcal{L}}{\partial (\partial_y \phi)}, \frac{\partial \mathcal{L}}{\partial (\partial_z \phi)} \right), \quad (4.25)$$

则

$$\frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} \partial_i \delta\phi = \mathbf{A} \cdot \nabla \delta\phi. \quad (4.26)$$

这样,

$$\delta S = \frac{1}{c} \int \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \mathbf{A} \cdot \nabla \delta\phi \right) d\Omega. \quad (4.27)$$

利用

$$\oint \mathbf{f} \cdot d\mathbf{S} = \int \nabla \cdot \mathbf{f} dV. \quad (4.28)$$

选取 $\mathbf{f} = \delta\phi \mathbf{A}$, 则

$$\oint \delta\phi \mathbf{A} \cdot d\mathbf{S} = \int (\delta\phi \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla \delta\phi) dV, \quad (4.29)$$

此处我们用到了

$$\nabla \cdot (\delta\phi \mathbf{A}) = \delta\phi \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla \delta\phi. \quad (4.30)$$

所以,

$$\int \mathbf{A} \cdot \nabla \delta\phi dV = \oint \delta\phi \mathbf{A} \cdot d\mathbf{S} - \int \delta\phi \nabla \cdot \mathbf{A} dV. \quad (4.31)$$

注意到在边界上 $\delta\phi = 0$,

$$\int \mathbf{A} \cdot \nabla \delta\phi dV = - \int \delta\phi \nabla \cdot \mathbf{A} dV. \quad (4.32)$$

这样作用量的扰动量变成

$$\delta S = \frac{1}{c} \int \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi - \delta\phi \nabla \cdot \mathbf{A} \right) d\Omega, \quad (4.33)$$

$$= \frac{1}{c} \int \left(\frac{\partial \mathcal{L}}{\partial \phi} - \nabla \cdot \mathbf{A} \right) \delta\phi d\Omega. \quad (4.34)$$

到这一步, 我们已经将 δS 化成了想要的形式。所以从最小作用量原理, 场的演化方程应该是

$$\frac{\partial}{\partial x^i} \left(\frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad (4.35)$$

这里我们利用了

$$\nabla \cdot \mathbf{A} = \frac{\partial A^i}{\partial x^i} \quad \text{且} \quad A^i = \frac{\partial \mathcal{L}}{\partial (\partial_i \phi)}. \quad (4.36)$$

注意我们利用了逆变-协变形式表达了 $\nabla \cdot \mathbf{A}$, 为了和更一般的情况一致。

4.2.3 标量场的拉格朗日方程

对一般的标量场, $\phi = \phi(t, x, y, z) = \phi(x^\mu)$, 我们仅从方程 (4.20) 和 (4.35) 不难猜出, 其拉格朗日方程应该为¹

$$\frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right] - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (4.37)$$

这里是猜出来的。费曼曾稍为夸张地说过, 在你不知道一个问题的答案前, 你不应该去解它。从场的拉格朗日方程可以看出, 如果 \mathcal{L} 为一四维标量, 但上式所代表的场的方程自然而然具有协变性, 这一点与之前我们所强调的作用量为一四维标量一致。

现在我们从最小作用量原理得到上述拉格朗日方程, 但过程需要用要高维情况下的高斯定理。在三维空间里, 对于一个向量 \mathbf{A} , 我们有

$$\int \nabla \cdot \mathbf{A} dV = \int \mathbf{A} \cdot d\mathbf{S}. \quad (4.38)$$

¹在很多教科书里, $\partial_\mu \phi$ 也被记为 $\phi_{,\mu}$, 注意这里下标中的逗号“,”。如果是一个向量场 A^ν 或 A_ν 对 μ 的导数, 则分别可记成 $A^\nu_{,\mu}$ 或 $A_{\nu,\mu}$ 。这种记法的一个优点是场的拉格朗日方程看起来会更简洁些, 但本讲义里为简化起见, 后续不引入这种表达方式。

这里右边的面积分为沿着包裹体积 V 的曲面。如果我们将上式写成分量形式，则为

$$\int \partial_i A^i dV = \int A^i dS_i. \quad (4.39)$$

而对于四维时空下的向量，其同样存在类似表达的高斯定理，即

$$\int \partial_\alpha A^\alpha dV = \int A^\alpha dS_\alpha. \quad (4.40)$$

左边是 $n = 4$ 维空间里的一个散度体积分，右边是 $n - 1 = 3$ 维的包裹着该“体积”的“面”积分；当然，高维空间里的面比较难以形象的理解。这个定理的具体证明放在习题中。我们将要利用该定理，从最小作用量原理得到一般标量场的欧拉-拉格朗日方程。

对于标量场 $\phi = \phi(x^\mu)$ ，我们有作用量

$$S = \int \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) d(ct) dx dy dz = \int \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) d\Omega \quad (4.41)$$

如果我们将真实的场 $\phi(x^\mu)$ 扰动， $\phi \rightarrow \phi + \delta\phi$ ，且在边界面 $\partial\Omega$ 上 $\delta\phi = 0$ ，最小作用量原理表明， $\delta S = 0$ ，即

$$\delta S = \int \{ \mathcal{L}[\phi + \delta\phi, \partial_\mu(\phi + \delta\phi), x^\mu] - \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) \} d\Omega. \quad (4.42)$$

保留到一阶项，上式变为

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\mu(\delta\phi) \right] d\Omega. \quad (4.43)$$

在上式的右边第二项里，与之前一样，我们交换了变分 (δ) 与微分算符 (∂)。再一次的，为了从最小作用量原理中得到场的现在的目标是要将上式写成 $\delta S = \int (\dots) \delta\phi d\Omega$ 的形式。对于上式右边第二项我们有因此方程 4.43 右边第二项可以改写为

$$\int \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \frac{\partial}{\partial x^\mu} (\delta\phi) \right] d\Omega = \int \left\{ \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right] - \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right] \delta\phi \right\} d\Omega. \quad (4.44)$$

利用 4.40，上式右边第一项可以写为

$$\int \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right] d\Omega = \int \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi dS^\mu = 0. \quad (4.45)$$

最后一步，我们利用了边界“面”上， $\delta\phi = 0$ 。因此，

$$\delta S = \int \left\{ \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right] \right\} \delta\phi d\Omega = 0. \quad (4.46)$$

上式对任意的扰动 $\delta\phi$ 都成立，因此，场的拉格朗日方程为

$$\frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right] - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (4.47)$$

这里我们按照习惯交换了两项的顺序，没有本质上的改变。因为对时间坐标与空间坐标采取同等的处理，可以看出拉格朗日方程所描述的场论自然而然的具有符合相对论协变形式，只要 \mathcal{L} （或作用量 S ）是一个四维标量。

接下来我们举两个例子，从标量场的拉格朗日量得到场的方程。

例题 4.1 我们选的第一个拉格朗日量密度是 $\mathcal{L} = \frac{1}{2} \partial_\alpha \phi \partial^\alpha \phi$ ，这是一个简单但不无聊的场。如果是 $\mathcal{L} = \phi$ 这种形式，不仅简单，而且无聊。系数里的 $1/2$ 是为了在后面的例子里面可以消除掉。在本例里面，是没有用处的。

解 对于该拉氏量，显然有 $\partial \mathcal{L} / \partial \phi = 0$ 。所以拉格朗日方程为

$$\frac{\partial}{\partial x^\alpha} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi)} \right] = 0. \quad (4.48)$$

这里的一个数学问题是如何计算该式里的导数？我们需注意到两点。一，因为是对 x^α 求导，指标 α 已经用过。拉氏量里的哑指标不能再选用 α ，所以我们让 $\mathcal{L} \propto \partial_\nu \phi \partial^\nu \phi$ 。二，注意到 $\partial_\nu \phi = \delta_\nu^\alpha \partial_\alpha \phi$ 且 $\partial^\nu \phi = g^{\nu\alpha} \partial_\alpha \phi$ ，那么

$$\frac{1}{2} \frac{\partial}{\partial(\partial_\alpha \phi)} (\partial_\nu \phi \partial^\nu \phi) = \frac{1}{2} \left[\frac{\partial(\partial_\nu \phi)}{\partial(\partial_\alpha \phi)} \partial^\nu \phi + \frac{\partial(\partial^\nu \phi)}{\partial(\partial_\alpha \phi)} \partial_\nu \phi \right], \quad (4.49)$$

$$= \frac{1}{2} [\delta_\nu^\alpha \partial^\nu \phi + g^{\nu\alpha} \partial_\nu \phi], \quad (4.50)$$

$$= \frac{1}{2} [\partial^\alpha \phi + \partial^\alpha \phi] = \partial^\alpha \phi. \quad (4.51)$$

将该结果代入拉氏方程里，我们有

$$\frac{\partial}{\partial x^\alpha} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi)} \right] = \frac{\partial}{\partial x^\alpha} \left(\frac{\partial \phi}{\partial x_\alpha} \right) = 0. \quad (4.52)$$

或者说

$$\frac{\partial^2}{\partial x^\alpha \partial x_\alpha} \phi = 0. \quad (4.53)$$

这是场的方程的协变形式。其中

$$\frac{\partial^2}{\partial x^\alpha \partial x_\alpha} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2, \quad (4.54)$$

所以该场的方程也是

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = 0. \quad (4.55)$$

这是一个以光速 c 传播的波动方程。此式也可利用达朗贝尔算符

$$\square \equiv \frac{\partial^2}{\partial x^\alpha \partial x_\alpha} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \quad (4.56)$$

简单表达为 $\square \phi = 0$ 。

例题 4.2 现在考虑有一个粒子存在的情况下，上例中的场的方程为如何改变？

解 对于有粒子的情况，我们需要知道该粒子如何与场耦合。我们从分析力学里面学过，非相对论粒子在外场中的拉氏量为

$$L = \frac{1}{2} m v^2 - \alpha \phi(\mathbf{x}_p, t). \quad (4.57)$$

这里第二项表达的就是粒子与场的耦合项，即 $L_{pf} = -\alpha \phi$ ，且 ϕ 是沿着粒子轨迹计算的，所以是 \mathbf{x}_p 的函数。

那如何将这个耦合项运用在场在拉氏量里？我们需要注意到 L_{pf} 是拉格朗日量，而在场的拉格朗日方程里，需要用到的是拉格朗日量密度。所以我们用

$$L_{pf} = \int \mathcal{L}_{pf} dV \Rightarrow \mathcal{L}_{pf} = -\alpha \phi(\mathbf{x}, t) \delta(\mathbf{x} - \mathbf{x}_p). \quad (4.58)$$

得到 \mathcal{L}_{pf} 。所以在有粒子存在情况下，系统的拉格朗日量密度为

$$\mathcal{L} = \frac{1}{2} \partial_\alpha \phi \partial^\alpha \phi - \alpha \phi(\mathbf{x}, t) \delta(\mathbf{x} - \mathbf{x}_p) = \mathcal{L}_f + \mathcal{L}_{pf}. \quad (4.59)$$

将该拉格朗日量代入场的拉氏方程，注意到 \mathcal{L}_{pf} 与 $\partial_\mu \phi$ 无关。所以从上例中，我们有

$$\frac{\partial}{\partial x^\alpha} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi)} \right] = \frac{\partial}{\partial x^\alpha} \left(\frac{\partial \phi}{\partial x_\alpha} \right) \quad (4.60)$$

而 \mathcal{L}_f 与 ϕ 无关, 所以

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\alpha \delta(\mathbf{x} - \mathbf{x}_p). \quad (4.61)$$

这样有粒子情况下, 场的演化方程变成

$$\frac{\partial^2}{\partial x^\alpha \partial x_\alpha} \phi = -\alpha \delta(\mathbf{x} - \mathbf{x}_p). \quad (4.62)$$

或者

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = -\alpha \delta(\mathbf{x} - \mathbf{x}_p). \quad (4.63)$$

这是一个有源波动方程, 而粒子则扮演着源的角色。

4.2.4 向量场的拉格朗日方程

如果现在我们面对的场不是标量场, 而是一个向量场 $A^\nu = A^\nu(t, x, y, z)$, 其拉格朗日方程应该为何种形式? 首先, 该场的拉格朗日密度的形式应该为 $\mathcal{L} = \mathcal{L}(A^\nu, \partial_\mu A^\nu)$ 。这里 $\partial_\mu \equiv \partial/\partial x^\mu$ 。从分析力学里单个坐标的拉氏量与多坐标的拉氏量以及标量场的拉格朗日方程, 该向量场的拉格朗日方程应该只需要将标量场方程中的 $\partial_\mu \phi$ 替换成 $\partial_\mu A^\nu$ 即可, 所以有如下形式

$$\frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu A^\nu)} \right] - \frac{\partial \mathcal{L}}{\partial A^\nu} = 0. \quad (4.64)$$

该方程有 4 个分量, $\nu = 0, 1, 2, 3$ 。该场对应的作用量为 $S = \int \mathcal{L}(A^\nu, \partial_\mu A^\nu, x^\mu) d\Omega$ 。本课程里的电磁场的拉格朗日量与作用量便是一个关于 $\partial_\mu A^\nu$ 的四维标量函数。在下一节详细讨论。

4.3 电磁场的拉格朗日量与场方程

在之前的课程里, 我们已经知道电磁场是一个向量场。那它的拉氏量应该是什么? 这里的猜测也有一些限制条件。1, 该量是一个四维标量, 这一点对所有的作用量和拉格朗日量密度都是一个基本要求。在该要求下, 相应的场的方程自动满足相对性原理。2, 该场是一个规范场, 相应的量应该具有规范不变性, 所以像 $A^\mu A_\mu$ 这种形式的量不会满足规范不变性。而如果拉氏量是 \mathbf{E} 或 \mathbf{B} 的函数, 则规范不变性自动满足。3, 实验显示, 该场具有可叠加性。因此相应的场的方程应该是线性方程, 那么拉格朗日密度应该是 E^2 或 B^2 这种形式²。在这三个条件的限制下, 可以满足条件的拉格朗日量的形式不太多。实际中, 电磁场的拉格朗日量形式为

$$\mathcal{L} = -\frac{1}{4\mu_0 c} F^{\alpha\beta} F_{\alpha\beta}. \quad (4.65)$$

这里选用的系数 $1/4\mu_0 c$ 是因为我们选用的单位制。显而易见, 这个拉格朗日量满足以上所有条件。当然最终以实验检验, 该拉氏量所给出的方程满足实验的观测。

4.3.1 无粒子存在下的场方程

从该拉格朗日量, 以及方程 (4.64), 我们不难得到电磁场的方程。注意到该拉格朗日量其实类似于例 4.1 中的拉氏量, 在例 4.1 中, $\mathcal{L} \propto (\partial_\mu \phi)^2$, 而本例中 $\mathcal{L} \propto F^{\alpha\beta} F_{\alpha\beta} \propto (\partial_\alpha A^\beta)^2$, 其最终的场的方程也是类似的。首先, 电磁

²我们还可以要求在拉格朗日量密度中, A^α 对 x^μ 的导数的阶数最高只是一阶, 类似于粒子的拉格朗日量里只有速度 $(d\mathbf{x}/dt)$, 而没有加速度 $(d^2\mathbf{x}/dt^2)$ 。

场的拉格朗日方程与 A^β 无关，所以拉格朗日方程里的 $\partial L/\partial A^\beta = 0$ ，我们需要计算的仅仅是第一项里的导数。再次，我们计算 $\partial \mathcal{L}/\partial(\partial_\alpha A^\beta)$ ，这里

$$\mathcal{L} = -\frac{1}{4\mu_0 c} F^{\alpha\beta} F_{\alpha\beta} = -\frac{1}{4\mu_0 c} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)(\partial^\alpha A^\beta - \partial^\beta A^\alpha), \quad (4.66)$$

$$= -\frac{1}{4\mu_0 c} [(\partial_\alpha A_\beta)(\partial^\alpha A^\beta) - (\partial_\beta A_\alpha)(\partial^\alpha A^\beta) - (\partial_\alpha A_\beta)(\partial^\beta A^\alpha) + (\partial_\beta A_\alpha)(\partial^\beta A^\alpha)]. \quad (4.67)$$

这里我们首先注意到，方括号里的第一项和最后一项其实是一样的，因为哑指标是可以更改的。比如，最后一项里，我们将 $\alpha \rightarrow \beta$ 且同时将 $\beta \rightarrow \alpha$ ，得到 $(\partial_\alpha A_\beta)(\partial^\alpha A^\beta)$ ，与第一项相同。类似的道理，第二项与第三项其实是一样的。这样，拉氏量变成

$$\mathcal{L} = -\frac{1}{2\mu_0 c} [(\partial_\alpha A_\beta)(\partial^\alpha A^\beta) - (\partial_\beta A_\alpha)(\partial^\alpha A^\beta)]. \quad (4.68)$$

为了得到场的方程，我们需要将该拉格朗日量对 $\partial_\mu A^\nu$ 求导。与例 4.1 类似，这里只需要注意到，

$$\partial_\alpha A_\beta = \delta_\alpha^\mu (\partial_\mu A^\nu) g_{\nu\beta}, \quad (4.69)$$

$$\partial^\alpha A^\beta = g^{\alpha\mu} (\partial_\mu A^\nu) \delta_\nu^\beta. \quad (4.70)$$

不难得到，

$$\frac{\partial}{\partial(\partial_\mu A^\nu)} [(\partial_\alpha A_\beta)(\partial^\alpha A^\beta)] = \frac{\partial}{\partial(\partial_\mu A^\nu)} [(\partial_\alpha A_\beta)] (\partial^\alpha A^\beta) + \frac{\partial}{\partial(\partial_\mu A^\nu)} [(\partial^\alpha A^\beta)] (\partial_\alpha A_\beta) = 2\partial^\mu A_\nu. \quad (4.71)$$

类似的

$$\frac{\partial}{\partial(\partial_\mu A^\nu)} [(\partial_\beta A_\alpha)(\partial^\alpha A^\beta)] = 2\partial_\nu A^\mu. \quad (4.72)$$

所以我们可以得到

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\nu)} = \frac{1}{\mu_0 c} (\partial_\nu A^\mu - \partial^\mu A_\nu) = \frac{1}{\mu_0 c} F_\nu^\mu. \quad (4.73)$$

相应的场的方程为

$$\frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\nu)} \right] = 0 \Rightarrow \frac{1}{\mu_0 c} \frac{\partial F_\nu^\mu}{\partial x^\mu} = 0 \quad \text{或者} \quad \frac{\partial F_{\nu\mu}}{\partial x_\mu} = 0. \quad (4.74)$$

代入 $F_{\nu\mu} = \partial_\nu A_\mu - \partial_\mu A_\nu$ ，则场的方程为

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial A^\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x_\mu} \right) = 0. \quad (4.75)$$

这即是电磁场所满足的演化方程。

因为该场是规范场，所以我们可以选一个规范条件，进一步简化该方程。注意到上述方程的第一项可以写为

$$\frac{\partial}{\partial x^\mu} \frac{\partial A^\mu}{\partial x^\nu} = \frac{\partial}{\partial x^\nu} \frac{\partial A^\mu}{\partial x^\mu}, \quad (4.76)$$

而洛伦茨规范条件即是 A^μ 的四维散度为 0。所以如果我们选定洛伦茨规范，则电磁场方程的形式可以简化为

$$\frac{\partial^2}{\partial x^\mu \partial x_\mu} A_\nu = 0 \Rightarrow \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) A_\nu = 0. \quad (4.77)$$

这即是在洛伦茨规范条件下的电磁场方程，其分量形式为

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = 0, \quad \nu = 0, \quad (4.78)$$

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = 0, \quad \nu = 1, 2, 3. \quad (4.79)$$

显然，这描述了没有带电粒子存在情况下电磁场满足波动方程，其波速为 c 。

4.3.2 给定粒子分布下的场的方程

如果有带电粒子存在，场的方程该会是什么样子？这里我们显然需要用到存在带电粒子里，系统的拉格朗日量密度，其包含了场的拉格朗日量和表达场与粒子耦合的拉格朗日量。在前一章中，我们知道，表达带电粒子与电磁场耦合的作用量为

$$S_{pf} = - \int_a^b q A_\alpha dx^\alpha, \quad (4.80)$$

这里 A_α 是沿着粒子轨迹计算的，即 $A_\alpha = A_\alpha(t, \mathbf{x}_p)$ 。我们接下来需要做的是将该作用量写成 $S_{pf} = \int \mathcal{L}_{pf} d\Omega$ 的形式，从而得到相应的 \mathcal{L}_{pf} ，例 4.2 提供了一个示例。

$$S_{pf} = - \int_a^b q A_\alpha(t, \mathbf{x}_p) dx^\alpha = - \int_a^b q A_\alpha(t, \mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_p) \frac{dx^\alpha}{d(ct)} dV d(ct), \quad (4.81)$$

所以显而易见，如果存在一个带电粒子的话，

$$\mathcal{L}_{pf} = q A_\alpha(t, \mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_p) \frac{dx^\alpha}{d(ct)}. \quad (4.82)$$

而如果存在多个带电粒子的话，

$$\mathcal{L}_{pf} = \sum_i q_i A_\alpha(t, \mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_i) \frac{dx^\alpha}{d(ct)}. \quad (4.83)$$

这里下标 i 表示对第 i 个粒子求和。进一步的，我们利用电荷密度的定义，即

$$\rho = \sum_i q_i \delta(\mathbf{x} - \mathbf{x}_i), \quad (4.84)$$

则

$$\mathcal{L}_{pf} = \frac{1}{c} \rho A_\alpha(t, \mathbf{x}) \frac{dx^\alpha}{dt}. \quad (4.85)$$

从推导过程，该拉氏量显然是一个四维标量。而且从上下标的位置可以看出，我们可以将其理解为一个四维向量与 A_α 的点乘。我们定义四维电流向量，

$$j^\alpha = \rho \frac{dx^\alpha}{dt} = (\rho c, \mathbf{j}) \quad (4.86)$$

则

$$\mathcal{L}_{pf} = \frac{1}{c} j^\alpha A_\alpha(t, \mathbf{x}). \quad (4.87)$$

从四维电流的定义过程也可以看出，电荷密度并不是四维标量，因为 q 是标量，但是体积 V 不是。关于该四维向量，在后面的小节中再仔细讨论。

综合以上结果，有粒子存在情况下，系统的拉格朗日密度为

$$\mathcal{L} = -\frac{1}{4\mu_0 c} F_{\alpha\beta} F^{\alpha\beta} - \frac{1}{c} j^\alpha A_\alpha. \quad (4.88)$$

从该拉氏量密度，不难得到存在带电粒子时场的方程。注意，这里粒子的分布是给定的。将该拉格朗日量密度代入方程 (4.64)，注意到 \mathcal{L}_f 与 A_μ 无关，而 \mathcal{L}_{pf} 与 $\partial_\mu A^\nu$ 无关。所以

$$\frac{\partial \mathcal{L}}{\partial A^\nu} = \frac{\partial \mathcal{L}_{pf}}{\partial A^\nu} = -\frac{1}{c} j^\alpha g_{\alpha\nu} = -\frac{1}{c} j_\nu, \quad (4.89)$$

这里我们利用了 $A_\alpha = g_{\alpha\nu} A^\nu$ 。而

$$\frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu A^\nu)} \right] = \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}_f}{\partial (\partial_\mu A^\nu)} \right]. \quad (4.90)$$

所以利用方程 (4.73), 场的方程为

$$\frac{1}{\mu_0 c} \frac{\partial F_{\nu\mu}}{\partial x_\mu} = -\frac{1}{c} j_\nu \quad \text{或者} \quad \frac{\partial F_{\mu\nu}}{\partial x_\mu} = \mu_0 j_\nu. \quad (4.91)$$

如果我们再次选用洛伦茨规范, 则该方程可以写为

$$\frac{\partial^2}{\partial x^\mu \partial x_\mu} A_\nu = \mu_0 j_\nu. \quad (4.92)$$

其分量形式为

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = \frac{\rho}{\epsilon_0}, \quad \nu = 0 \quad (4.93)$$

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \mu_0 \mathbf{j}, \quad \nu = 1, 2, 3. \quad (4.94)$$

这里我们利用了 $c^2 = 1/\mu_0 \epsilon_0$ 。

4.3.3 用 \mathbf{E} 和 \mathbf{B} 表达的场方程

场的方程 (4.93) 是利用势函数 ϕ 和 \mathbf{A} 表达出来的。如果用 \mathbf{E} 和 \mathbf{B} 表达, 会是什么样的了? 这里我们需要用的 \mathbf{E} 和 \mathbf{B} 的定义, 以及洛伦茨规范条件。首先从定义出发,

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi, \quad (4.95)$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (4.96)$$

我们立刻有,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (4.97)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (4.98)$$

这两个方程仅仅是从定义得到的, 所以无论存不存在带电粒子, 它们都是成立的。其中第一个方程称为法拉第定理 (Faraday's law), 第二个方程并没有一个明确的名字, 可以看作磁场的高斯定理。

其次, 我们注意到洛伦茨规范条件, 即 $\partial_\alpha A^\alpha = 0$ 或者

$$\frac{1}{c^2} \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} = 0. \quad (4.99)$$

对 ϕ 的方程,

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \nabla \cdot \left(-\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \right) = \nabla \cdot \mathbf{E} = \rho/\epsilon_0. \quad (4.100)$$

这里我们对洛伦茨规范求了时间导, 即将 $\partial/\partial t$ 作用在方程 (4.99) 上, 得到

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\nabla \cdot \frac{\partial \mathbf{A}}{\partial t}. \quad (4.101)$$

该方程即为高斯定理。

另一方面, 对矢势 \mathbf{A} 而言,

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} + \frac{1}{c^2} \frac{\partial}{\partial t} \nabla \phi + \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}. \quad (4.102)$$

这里我们将 $\nabla \cdot$ 作用在洛伦茨规范 (4.99) 上, 得到

$$\frac{1}{c^2} \frac{\partial}{\partial t} \nabla \phi + \nabla(\nabla \cdot \mathbf{A}) = 0. \quad (4.103)$$

进一步的，利用 \mathbf{E} 和 \mathbf{B} 的定义，上式可以化成

$$-\frac{1}{c^2} \frac{\partial}{\partial t} \left(-\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \right) + \nabla \times (\nabla \times \mathbf{A}) = -\frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \nabla \times \mathbf{B} = \mu_0 \mathbf{j}. \quad (4.104)$$

整理可得

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (4.105)$$

此式即加了位移电流项的安培定理 (Ampere's law)。

总结一下，电磁场的方程可以表达为多种形式。形式一，协变形式

$$\frac{\partial F_{\mu\nu}}{\partial x_\mu} = \mu_0 j_\nu. \quad (4.106)$$

形式二，采用洛伦茨规范后，用势函数表达的形式为

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = \frac{\rho}{\epsilon_0}, \quad (4.107)$$

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \mu_0 \mathbf{j}. \quad (4.108)$$

或者形式三，用 \mathbf{E} 与 \mathbf{B} 所表达的形式为

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (4.109)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (4.110)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (4.111)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (4.112)$$

这些场的方程清楚的表明了场是如何随时空演化的，且又是如何受带电粒子影响的。在后续的课程里，我们将分各种特殊情况，一一求解该方程。从简单的静电场开始，到获得复杂的随时间变化又存在着源（带电粒子）的情况下这一套方程的解以及它们的性质。

麦克斯韦在电磁理论上的贡献，不仅仅在于他通过引入位移电流将电磁现象统一在一套麦克斯韦方程里。更重要的是他首先提出了场具有能量这一观点，从而不再只是一种数学描述上的方便，而是建立了场是物质的另一种形态这一重大物理学概念的突破，深深地影响了后世的物理学发展。与粒子一样，场具有能量，动量，和角动量。这些物理量也可以从场的作用量或拉格朗日量得到，这一点留在习题部分。在课程里，我们从这些物理量的守恒角度，得到场的能量密度和动量密度。在介绍具体的能量和动量守恒之前，我们先来看一下表述场这一类具有连续分布特性的物质时，守恒量的一般表达方式。

4.4 连续性方程

对于一个离散的粒子体系，如果我们讨论能量和动量守恒，通常的表达是在一过程前后， $\sum_i \mathcal{E}_i$ 和 $\sum_i \mathbf{p}_i$ 不变。这里 \sum_i 是对所有参与该过程的粒子求和。例如，两个粒子发生弹性碰撞，则碰撞前后，我们有简单的能量和动量守恒，

$$\mathcal{E}_1 + \mathcal{E}_2 = \mathcal{E}'_1 + \mathcal{E}'_2, \quad (4.113)$$

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}'_1 + \mathbf{p}'_2. \quad (4.114)$$

带'的量为碰撞后的。但在连续性物质里面的守恒律是如何表达呢？原则上如果我们将连续性介质当作无穷多粒子组成的体系，方程 (4.113)-(4.114) 显然都是正确的。但是另一方面，这也显然不太方便。所以在处理连续性介质的守恒律时，通常会采取另一种方式，这种方式产生的方程具有类似的形式，称为连续性方程³。

³这里所讨论的两种考虑方式差异与流体力学中的拉格朗日描述和欧拉描述的差异类似。

在这一小节里，我们讨论的物理量没有源 (source) 和汇 (sink)，即该物理量不能被创造也不能消失。电荷守恒为例，我们先讨论没有源和汇的连续性方程。在这一类方程里。如果我们考虑空间中一个区域 V 内的电荷量的变化，因为电荷守恒，我们很容易知道该区域内的电荷的单位时间的增加量，等于单位时间内通过该区域表面 (∂V) 流入该区域的电荷量的多少。而后者可以表达为电荷的通量占剩面积矢量，即 $-\mathbf{F}_\rho \cdot d\mathbf{S}$ ，然后对全表面积分。对电荷而言，其通量显然是 $\mathbf{F}_\rho = \rho\mathbf{v}$ 。这里 \mathbf{v} 是将电荷当成一个连续分布的介质后流动的速度。所以相应的方程是

$$\frac{\partial}{\partial t} \int \rho dV + \int \rho\mathbf{v} \cdot d\mathbf{S} = 0. \quad (4.115)$$

利用高斯定理，将面积分化成体积分，上式可以表达成

$$\frac{\partial}{\partial t} \int \rho dV + \int \nabla \cdot (\rho\mathbf{v}) dV = 0. \quad (4.116)$$

我们在固定的区域，所以可以交换 $\partial/\partial t$ 与积分符号，整理得

$$\int \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{v}) \right] dV = 0. \quad (4.117)$$

这即是连续性方程的积分形式。又因为电荷守恒对任一区域都成立，即守恒律的局地性⁴，所以方括号内的式子一定是 0，即

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{v}) = 0. \quad (4.118)$$

此即连续性方程的微分形式。具体对电荷而言，我们注意到电流密度 $\mathbf{j} = \rho\mathbf{v}$ 其实就是电荷通量，所以电荷的连续性方程为

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (4.119)$$

进一步的，该式可以利用四维电流向量 $j^\mu = (\rho/c, \mathbf{j})$ 表达为

$$\frac{\partial j^\alpha}{\partial x^\alpha} = 0, \quad (4.120)$$

或者说四维电流向量的四维散度等于 0，就是说任何参考系里，上式都是成立的。下一节中我们也可以利用场的拉格朗日量，从系统的对称性中，找到相应的守恒律。这些守恒律表达为一个四维向量的四维散度为 0。

相关的概念可以推广到其它守恒量上。比如流体力学里常常假设质量守恒（无创造和消失），所以流体力学（包括磁流体力学）里常见的一个方程便是，

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{v}) = 0. \quad (4.121)$$

下标 m 表示这是质量密度。类似的，如果我们假设粒子无创造或消失，则也有粒子数守恒。如果粒子数密度为 n ，则相应的连续性方程的形式为

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{v}) = 0. \quad (4.122)$$

所以总结一下，如果物理量 X 守恒，则其连续性方程可以表达为

$$\frac{\partial \rho_X}{\partial t} + \nabla \cdot \mathbf{F}_X = 0. \quad (4.123)$$

这里 ρ_X 为单位体积内 X 的量，而 \mathbf{F}_X 则为 X 的通量，一般形式为 $\rho_X \mathbf{v}_X$ 。这些连续性方程物理学里具有广泛的应用，理解并认识这种形式十分重要。数值计算方法里的有限元 (finite volume) 方法也是基于这一类连续性方程形式的，在数值流体动力学 (computational fluid dynamics) 中应用十分广泛。

⁴例如，如果电荷仅仅是全局守恒，则电荷守恒可以指的是某处消失 Q 的电荷，另一处会立刻增加 Q 的电荷。

连续性方程 (4.123) 的形式有时也很方便地让我们理解一些概念。以布朗运动为代表的扩散过程为例，描述密度 n 的随时空变化的方程通常具有的形式为

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial n}{\partial x} \right), \quad (4.124)$$

这种形式，即对时间一阶导，对空间（或者说是扩散发生的坐标）是二阶导。这里 D 是扩散系数。结合方程 (4.123) 很容易看出， $-D\partial n/\partial x$ 可以理解为扩散过程的通量，所以该通量与扩散系数和物理量的梯度均有关，方向和梯度方向相反，意味着虽然扩散过程微观上对应着随机运动，但其净结果是将该量从高往低运输。如果 $\partial n/\partial x = 0$ ，则扩散对应的净输运停止。

4.4.1 有源和汇的连续性方程

如果一个物理量有源或汇，那它的连续性方程应该是什么样子？比如，假设我们考虑的是所有种类的能量，则显然能量会满足方程 (4.123) 的形式。但如果我们只考虑场的能量呢？因为能量可以在不同形式之间进行转换，比如电磁场可以能过对带电粒子做功的方程将能量转移给粒子，那么对于电磁场而言，显然其能量是不一定守恒的，取决于有没有转移给带电粒子。从上一小节连续性方程的推导过程不难看出，这时候的连续性方程形式应该从方程 (4.123) 变为

$$\frac{\partial \rho_X}{\partial t} + \nabla \cdot \mathbf{F}_X = S, \quad (4.125)$$

这里 S 为单位时间单位体积内 X 产生或消失的量。如果 $S > 0$ ，则其为源；反之，则为汇。以电磁场的能量连续性方程，因为其可以通过对带电粒子做功而转移到粒子上，结合 S 的含义，其应该与电磁场对单位体积内带电粒子做功的功率有关。对单个带电粒子，场的做功功率为 $q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} = \mathbf{E} \cdot q\mathbf{v}$ 。如果我们假设单位体积内的电量为 ρ ，显然做功功率为 $\mathbf{E} \cdot \rho\mathbf{v} = \mathbf{E} \cdot \mathbf{j}$ 。同时该功率表达的是单位时间单位体积内带电粒子能量的增加率，所以它也是单位时间单位体积内场的能量的减少率，也即 $S = -\mathbf{E} \cdot \mathbf{j}$ 。因为场的能量的连续性方程具有的形式应该为

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{P} = -\mathbf{E} \cdot \mathbf{j}. \quad (4.126)$$

这里 \mathcal{E} 是场的能量密度，而 \mathbf{P} 是场的能量通量。

同样的道理，如果我们考虑电磁场在 x 方向的动量守恒，则也存在类似形式的方程，

$$\frac{\partial g_x}{\partial t} + \nabla \cdot \mathbf{F}_{g_x} = S_x \quad (4.127)$$

那这里的 S_x 又是什么？显然其与电磁场在单位时间内向单位体积内的带电粒子转移了多少动量相关，而带电粒子在单位时间内获得了多少动量，等于粒子所受的力，所以 S_x 与单位体积粒子在 x 方向的受力有关，与能量的道理类似， $S_x = -\rho(\mathbf{E} + \mathbf{v} \times \mathbf{B})_x$ 。这里我们讨论的是 x 方向的动量。如果对于整体的动量，在之前的例子里，我们说过其通量对应着一个张量，所以场的动量连续性方程可以写成

$$\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \mathbf{G} = -\rho\mathbf{E} + \mathbf{j} \times \mathbf{B}. \quad (4.128)$$

这里我们利用了 $\mathbf{j} = \rho\mathbf{v}$ 。在下一节中，我们将利用方程 (4.126) 与 (4.128) 获得场的能量与动量密度的表达式。

4.5 场的能量与动量

4.5.1 场的能量密度与能量通量

如何从方程 (4.126) 中得到场的能量密度表达式？已知的量是方程右边的项，所以整体的方向是将方程的右边改写成左边的形式，即一个标量的时间导数加上一个向量的散度，如果可以做到，则该标量可能对应着能量密度，而

向量则可能对应着能量的通量。其次,不管是场的能量密度还是能量通量,都应该只是场的函数,即 $\mathcal{E} = \mathcal{E}(\mathbf{E}, \mathbf{B})$, 且 $\mathbf{P} = \mathbf{P}(\mathbf{E}, \mathbf{B})$ 。最后,我们在之前的课程里知道, \mathbf{E} 和 \mathbf{B} 是同一个物理量 ($F^{\alpha\beta}$) 的不同分量而已,其地位是平等的。因此,在场的能量,动量等表达式里, \mathbf{E} 和 \mathbf{B} 应该也具有平等地位。考虑到这三点,我们利用麦克斯韦方程,改写 $-\mathbf{E} \cdot \mathbf{j}$ 。

首先,我们利用安倍定律将电流 \mathbf{j} 替换成场,这样所有的量就只与场相关了。由安倍定律得,

$$\mathbf{j} = \frac{1}{\mu_0} \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (4.129)$$

所以

$$\mathbf{j} \cdot \mathbf{E} = \frac{1}{\mu_0} \left(\nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot \mathbf{E} = \frac{1}{\mu_0} \nabla \times \mathbf{B} \cdot \mathbf{E} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \cdot \mathbf{E}. \quad (4.130)$$

这个方程隐约具有,但还没有完全达到,我们所要求的形式,因此,我们继续通过增加减少项来改变它。结合法拉第定律,

$$\mathbf{j} \cdot \mathbf{E} = \frac{1}{\mu_0} (\nabla \times \mathbf{B} \cdot \mathbf{E} - \nabla \times \mathbf{E} \cdot \mathbf{B}) + \frac{1}{\mu_0} \nabla \times \mathbf{E} \cdot \mathbf{B} - \frac{\epsilon_0}{2} \frac{\partial E^2}{\partial t}, \quad (4.131)$$

$$= -\frac{1}{\mu_0} \nabla \cdot (\mathbf{E} \times \mathbf{B}) - \left(\frac{1}{2\mu_0} \frac{\partial B^2}{\partial t} + \frac{\epsilon_0}{2} \frac{\partial E^2}{\partial t} \right). \quad (4.132)$$

将该表达式代入方程 (4.126), 有

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{P} = \frac{\partial}{\partial t} \left(\frac{\epsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2 \right) + \nabla \cdot \left(\frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right). \quad (4.133)$$

对比方程两边,我们有

$$\mathcal{E} = \frac{\epsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2, \quad (4.134)$$

$$\mathbf{P} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}. \quad (4.135)$$

此即电磁场的能量密度 \mathcal{E} 与能量通量 \mathbf{P} , 后者也称为波印廷矢量。

从推导过程上看,我们并没有完全的定义电磁场的能量密度的形式。因为,显然我们可以在 \mathcal{E} 上加上任意一个与时间无关的量,或者在 \mathbf{P} 上加上任意一个向量的旋度, $\nabla \times \mathbf{D}$, 都不会改变该连续性方程。那我们如何知道我们的表达式是否正确? 首先,能量密度与其通量的表达式并不唯一。在讨论能量守恒性质这一点上,我们更多的是关注能量的变化量,而不是其绝对值。其次,目前为止,所有的实验都与这两个表达式稳合的很好,或者说没有证否这两个表达式。因此,表达式 (4.134) 与 (4.135) 可以被认为是最简单的满足能量守恒性质的场的能量密度和能量通量。

4.5.2 场的动量密度和麦克斯韦应力张量

参照类似的过程,我们可以推导场的动量密度和动量张量。从方程 (4.128) 出发,利用高斯定理与安倍定律,我们有

$$\rho = \epsilon_0 \nabla \cdot \mathbf{E}, \quad (4.136)$$

$$\mathbf{j} = \frac{1}{\mu_0} \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (4.137)$$

代入方程 (4.128) 的右边项,得到

$$-\rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = -\epsilon_0 (\nabla \cdot \mathbf{E}) \mathbf{E} - \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B}, \quad (4.138)$$

$$= -\epsilon_0 (\nabla \cdot \mathbf{E}) \mathbf{E} - \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} + \epsilon_0 \frac{\partial}{\partial t} (\mathbf{E} \times \mathbf{B}) + \epsilon_0 \mathbf{E} \times (\nabla \times \mathbf{E}). \quad (4.139)$$

这里我们利用了

$$\frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} = \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) - \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) + \mathbf{E} \times (\nabla \times \mathbf{E}). \quad (4.140)$$

上式可以进一步整理成

$$-\rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \frac{\partial}{\partial t}(\epsilon_0 \mathbf{E} \times \mathbf{B}) - \epsilon_0(\nabla \cdot \mathbf{E})\mathbf{E} - \frac{1}{\mu_0}(\nabla \cdot \mathbf{B})\mathbf{B} - \frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B} + \epsilon_0 \mathbf{E} \times (\nabla \times \mathbf{E}). \quad (4.141)$$

可以看出，我们仅需要将右边的后四项整理成一个张量的散度形式即可。利用

$$\mathbf{E} \times (\nabla \times \mathbf{E}) = (\nabla \cdot \mathbf{E})\mathbf{E} - \nabla \cdot \left(\mathbf{E}\mathbf{E} - \frac{1}{2}E^2 \mathbf{I} \right), \quad (4.142)$$

我们有

$$-\mathbf{E}(\nabla \cdot \mathbf{E}) + \mathbf{E} \times (\nabla \times \mathbf{E}) = -\nabla \cdot \left(\mathbf{E}\mathbf{E} - \frac{1}{2}E^2 \mathbf{I} \right), \quad (4.143)$$

同理，

$$-\mathbf{B}(\nabla \cdot \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{B}) = -\nabla \cdot \left(\mathbf{B}\mathbf{B} - \frac{1}{2}B^2 \mathbf{I} \right). \quad (4.144)$$

所以如果我们定义张量

$$\mathbf{T} = \left[\epsilon_0 \mathbf{E}\mathbf{E} + \frac{1}{\mu_0} \mathbf{B}\mathbf{B} - \frac{1}{2} \left(\epsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right) \mathbf{I} \right] \quad (4.145)$$

则

$$-\rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \frac{\partial}{\partial t}(\epsilon_0 \mathbf{E} \times \mathbf{B}) - \nabla \cdot \mathbf{T}. \quad (4.146)$$

对比方程 (4.128) 可知

$$\mathbf{g} = \epsilon_0 \mathbf{E} \times \mathbf{B}, \quad (4.147)$$

$$\mathbf{G} = -\mathbf{T}. \quad (4.148)$$

这里 \mathbf{g} 是电磁场的动量密度， \mathbf{G} 是动量通量，而 \mathbf{T} 则被称为麦克斯韦应力张量 (Maxwell stress tensor)，是一个非常重要的张量。

4.5.3 场的能量和动量的一些性质

场的能量和动量具有一些有趣的性质。一，场的波印廷矢量（能量通量）和动量密度的形式非常相似，仅仅相差一个 c^2 ，即 $\mathbf{P} = (1/\mu_0\epsilon_0)\mathbf{g} = c^2\mathbf{g}$ 。这点对电磁场而言其实不难理解。如果我们将电磁场看成为同一方向运动的大量光子组成，每个光子的能量为 $\hbar\omega$ ，动量为 $\hbar\mathbf{k}$ ，且 $\omega/k = c$ 。那么场的能量密度可以表达为 $N\hbar\omega$ ，而动量密度为 $N\hbar\mathbf{k}$ 。这里 N 是光子数密度。那么能量通量可以表达为 $N\hbar\omega\mathbf{c}$ ，这里 \mathbf{c} 与 \mathbf{k} 同向。从而 $\mathbf{P} = N\hbar\omega\mathbf{c} = N\hbar kc\mathbf{c} = N\hbar k c^2 = c^2\mathbf{g}$ 。

那这个结论对一般的粒子（非光子）构成的体系是否依然成立？答案是肯定的。这里我们只需要注意到，对一般的粒子而言，其能量为 γmc^2 ，而动量为 $\gamma m\mathbf{v}$ 。这样该体系的能量通量为 $N\gamma mc^2\mathbf{v}$ ，而动量密度为 $N\gamma m\mathbf{v}$ ，两者间显然相差了一个 c^2 。

上述用来解释能量通量和动量密度的两个粒子模型显然极其粗糙，但其实这个结论与连续分布介质的四维能动张量有关，具体可以参考习题部分。

二是关于麦克斯韦应力张量的含义。应力张量的各分量是什么意思？我们注意到 T_{ij} 可以在形式上理解为 $p_i v_j$ ，这里 p_i 是 i 方向的动量，而 v_j 则是 j 速度。因此， T_{ij} 是向 j 方向，单位时间通过单位面积传输了多少 i

方向的动量，而后者可以理解为单位面积所受的 i 方向的力 ($F = \mathbf{dp}/dt$)。例如， T_{xx} 是向垂直于 x 方向的单位面积所感受到的 x 方向的力，即压强。而 T_{xy} 则是垂直于 y 方向的单位面积所感受到的 x 方向的力。注意，此时力的方向是平行于该单位面的或者说其方向与该面的法向垂直，这个力称为剪应力 (shear stress)。这样， \mathbf{T} 张量的对角分量表示的是压力，而非对角分量表示的是剪应力。

应力张量 (stress tensor) 的概念并非电磁场所独有，而是一个非常一般的量。例如，如果大气或流体里的压强为 p ，则一个大气元所受的力，可以表达为 $-\nabla p$ 。这个力也可以写成张量的形式，即 $-\nabla \cdot \mathbf{p}$ ，即该应力张量是一所有对角项相同的对角张量，没有剪应力项。更一般的流体里，压强所对应的力也是由张量表示的， $-\nabla \cdot \mathbf{P}$ ，这个张量称为压强张量 (pressure tensor)。

另一方面，如果我们有一个系统，其电磁场不随时间变化，则从动量的连续性方程有，

$$\rho \mathbf{E} + \mathbf{j} \times \mathbf{B} = \nabla \cdot \mathbf{T}. \quad (4.149)$$

因此，应力张量的散度为单位体积的电荷所受的力。这也是应力张量的物理含义之一。

4.6 诺特定理：对称性与守恒量 *

与离散体系相同，系统拉格朗日量的意义不仅仅在于其能给出运动方程或者场方程。更重要的，通过拉格朗日量所展示的对称性，我们可以相对容易地得到系统的重要守恒量。更进一步，现代物理在分析基本粒子间的相互作用时，更是利用已知的守恒量获得系统应该有的对称性，从而去限定其拉格朗日量所具有的可能形式。而体现对称性与守恒量的重要定理便是诺特定理 (Noether's theorem)⁵。在离散体系中，时间对称性与能量守恒，空间平移对称性与动量守恒均是诺特定理的体现。在本节里，我们首先介绍场的诺特定理。并以此时空平移对称性为例，应用诺特定理，得到场的能量与动量守恒。

4.6.1 诺特定理

对于一个连续系统如场，诺特定理的内容如下：对于系统的每一个连续对称性，都存在一个相应的守恒量。这里连续对称性的含义是，对一个无穷小的时空坐标变换， $x^\mu \rightarrow x'^\mu \equiv x^\mu + \delta x^\mu$ ，或场的变换 $\phi(x^\mu) \rightarrow \phi'(x'^\mu) = \phi(x^\mu) + \delta\phi$ ， $\delta\phi \equiv \epsilon\Phi(x^\mu)$ ，系统的拉格朗日量密度形式上保持不变，即

$$\mathcal{L}'[\phi'(x'^\mu), \partial'_\mu \phi', x'^\mu] = \mathcal{L}[\phi(x^\mu), \partial_\mu \phi, x^\mu]. \quad (4.150)$$

假定 $x^\alpha \rightarrow x'^\alpha = x^\alpha + \delta x^\alpha$ ， $\phi'(x') = \phi(x) + \delta\phi$ ，这里的 $\delta\phi$ 包含了场本身的变化以及时空坐标的改变所引起的场的变化。定义 $\bar{\delta}\phi = \phi'(x) - \phi(x)$ ，即在同一时空坐标处，场本身的变化。接下来我们有

$$\delta\mathcal{L} = \mathcal{L}[\phi'(x'), \partial_\mu \phi'(x'), x'] - \mathcal{L}[\phi(x), \partial_\mu \phi(x), x] \quad (4.151)$$

$$= \underbrace{\mathcal{L}[\phi'(x'), \partial_\mu \phi'(x'), x'] - \mathcal{L}[\phi'(x), \partial_\mu \phi'(x), x]}_{\text{第一项: } x \rightarrow x' \text{ 所引起的变化}} + \underbrace{\mathcal{L}[\phi'(x), \partial_\mu \phi'(x), x] - \mathcal{L}[\phi(x), \partial_\mu \phi(x), x]}_{\text{第二项: } \phi(x) \rightarrow \phi'(x) \text{ 所引起的变化}}. \quad (4.152)$$

对于第一项，我们有

$$\mathcal{L}[\phi'(x'), \partial_\mu \phi'(x'), x'] - \mathcal{L}[\phi'(x), \partial_\mu \phi'(x), x] = \delta x^\mu \partial_\mu \mathcal{L} = \partial_\mu (\mathcal{L} \delta x^\mu) \quad (4.153)$$

需要指出的是，这里的 $\partial_\mu \mathcal{L}$ 指的是固定其它的 $x^\nu (\nu \neq \mu)$ 分量，对 x^μ 作偏导。例如当 $\mu = 0$ 时， $\partial_0 \mathcal{L} = \partial_0 \mathcal{L}|_{x^1, x^2, x^3}$ 。有时容易将其错误的理解为固定拉格朗日量， $\mathcal{L}(\phi, \partial_\mu \phi, x^\mu)$ ，中与场相关的量 $(\phi, \partial_\mu \phi)$ 而对 x^μ 作偏导。

⁵ 艾米·诺特 (Amalie Emmy Noether): 犹太人，19 世纪末至 20 世纪初的德国女数学家，被包括爱因斯坦和 Hermann Weyl 在内的一些科学家称为数学史上最重要的女性。曾被 David Hilbert 邀请至哥廷根大学担任教职，但哲学系教授反对女性担任教职，只能以 Hilbert 之名教授了四年。后受纳粹迫害，前往美国。

而对于第二项，我们有

$$\mathcal{L}[\phi'(x), \partial_\mu \phi'(x), x] - \mathcal{L}[\phi(x), \partial_\mu \phi(x), x] = \mathcal{L}[\phi(x) + \bar{\delta}\phi, \partial_\mu \phi(x) + \partial_\mu \bar{\delta}\phi, x] - \mathcal{L}[\phi(x), \partial_\mu \phi(x), x] \quad (4.154)$$

$$= \frac{\partial \mathcal{L}}{\partial \phi} \bar{\delta}\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\mu \bar{\delta}\phi = \frac{\partial \mathcal{L}}{\partial \phi} \bar{\delta}\phi - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right] \bar{\delta}\phi + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \bar{\delta}\phi \right] \quad (4.155)$$

$$= \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \bar{\delta}\phi \right]. \quad (4.156)$$

在其中我们利用了场的拉格朗日方程。所以总的拉格朗日量的变化为

$$\delta \mathcal{L} = \partial_\mu \left[\delta x^\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \bar{\delta}\phi \right]. \quad (4.157)$$

现在我们将上式表达为 $\delta\phi$ ，而不是 $\bar{\delta}\phi$ 。利用 $\delta\phi$ 的定义，保留至一阶项，我们有

$$\delta\phi = \phi'(x') - \phi(x) = \phi'(x') - \phi'(x) + \phi'(x) - \phi(x) \approx \frac{\partial \phi}{\partial x^\nu} \delta x^\nu + \bar{\delta}\phi(x). \quad (4.158)$$

因此，

$$\bar{\delta}\phi = \delta\phi - \partial_\nu \phi \delta x^\nu. \quad (4.159)$$

将此式代入 $\delta \mathcal{L}$ ，我们有

$$\delta \mathcal{L} = \partial_\mu \left[\delta x^\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} (\delta\phi - \partial_\nu \phi \delta x^\nu) \right] = \partial_\mu \left\{ \left[\delta_\nu^\mu \mathcal{L} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi \right] \delta x^\nu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right\}. \quad (4.160)$$

这里我们利用了 $\delta x^\mu = \delta_\nu^\mu \delta x^\nu$ 。我们定义一个新的张量， $T^{\mu\nu}$ ，

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L} \quad (4.161)$$

则上述拉格朗日量的变化可以写为

$$\delta \mathcal{L} = -\partial_\mu \left[T^{\mu\nu} \delta x_\nu - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right] = -\partial_\mu \delta j^\mu. \quad (4.162)$$

因此，如果 $\delta \mathcal{L} = 0$ ，则我们有 $\partial_\mu \delta j^\mu = 0$ 。但为了避免在守恒量的定义(式 4.164)中引入任意的一个无穷小量，我们通常会让 δx_ν 与 $\delta\phi$ 正比于一个无穷小参数 ϵ ，例如

$$\delta x_\nu = \epsilon X_\nu, \delta\phi = \epsilon \Phi. \quad (4.163)$$

这样，我们便可以定义 j^μ 为

$$j^\mu = T^{\mu\nu} X_\nu - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Phi, \quad (4.164)$$

从而我们有

$$\delta \mathcal{L} = -\epsilon \partial_\mu j^\mu = 0 \Rightarrow \partial_\mu j^\mu = 0. \quad (4.165)$$

该向量称为“流”(current)，而该流的四维散度为 0。假定 $j^0 = \rho/c$ ，即 $j^\mu = (\rho/c, \mathbf{j})$ ，则这一方程的三维形式为

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \quad (4.166)$$

如前所述，这是一个典型的连续性方程，表明了该流所代表的物理量的局域守恒性，其相应的守恒量为该“流”所对应的“荷”(charge)。为了看清这一点，对上式进行全空间积分，则

$$\int_{-\infty}^{\infty} \frac{\partial \rho}{\partial t} dV + \int_{-\infty}^{\infty} \nabla \cdot \mathbf{j} dV = \int_{-\infty}^{\infty} \frac{\partial \rho}{\partial t} dV + \int_{r \rightarrow \infty} \mathbf{j} \cdot d\mathbf{S} = 0, \quad (4.167)$$

假定在无穷远处 ($r \rightarrow \infty$), $\mathbf{j} \rightarrow 0$, 我们有全空间的荷的总量为 Q , 定义为

$$Q = \int_{\infty}^{\infty} \rho dV, \quad (4.168)$$

其满足

$$\frac{dQ}{dt} = \int_{\infty}^{\infty} \frac{\partial \rho}{\partial t} dV = 0. \quad (4.169)$$

即总的 Q 守恒。上述连续性方程所表达的这种守恒性是局地守恒, 即不仅仅全空间中 Q 的总量守恒⁶, 而且对于一个任意的有限空间 V_0 , 我们该空间中荷的总量 Q_0 的增加一定等于通过该空间的表面 (∂V_0) 流入该体积的荷的量, 即

$$\frac{dQ_0}{dt} = \int \frac{\partial \rho}{\partial t} dV = - \int_{\partial V_0} \mathbf{j} \cdot d\mathbf{S}. \quad (4.170)$$

因此, \mathbf{j} 也是该荷的通量, 而 ρ 则可以理解为该荷的密度。最后, 诺特定理所导出的连续性方程是以一个四维向量的四度散度形式所表达的, 因此其结论在不同参考系里都是成立的。

4.6.2 时空平移对称性与能量动量守恒

例题 4.3 在经典力学里, 如果一个系统存在时间平移对称性, 则系统能量守恒。

解 当对时间作 $t \rightarrow t' = t + \epsilon$ 变换时, 不会引起场的变化, 即新的场 $\phi'(t') = \phi(t)$ 。这样我们可以有 $\delta x_\nu = \epsilon \delta_\nu^0$ (或 $X_\nu = \delta_\nu^0$), 而 $\delta \phi = 0$ 。所以相应的流 j^ν 定义为

$$j^\mu = T^{\mu\nu} \delta_\nu^0 = T^{\mu 0} = \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\nu \phi - \delta_\nu^\mu \mathcal{L} \right] \delta_\nu^0 = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_0 \phi - \delta_0^\mu \mathcal{L}. \quad (4.171)$$

因此,

$$j^0 = \frac{\partial \mathcal{L}}{\partial (\partial_t \phi)} \partial_t \phi - \mathcal{L}, \quad (4.172)$$

$$j^i = \frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} \partial_0 \phi. \quad (4.173)$$

如果对比离散体系, 不难看出这里的 j^0 定义与哈密顿量相似, 代表着哈密顿密度, 或能量密度。对应的守恒量 $Q = \int (j^0/c) dV$ 是系统的能量。而 $\mathbf{j} = (j^1, j^2, j^3)$ 则对应着能流。

以上过程显然可以对任一个空间任标重复, 即如果系统存在着空间平移对称性, 即 $x_i \rightarrow x'_i = x_i + \epsilon$, 则对应着相应的动量守恒。其中对应着 x_i 的流为 $T^{\mu\nu} \delta_\nu^i = T^{\mu i}$, 而守恒式的表达为 $\partial_\mu T^{\mu i} = 0$ 。这里 T^{0i} 是对应于空间平移对称性时的守恒量, 与离散体系类比, 此守恒量为场的动量密度。

上述时间平移与空间平移可以合在一起, 构成时空平移对称性, 相应的守恒量为一个张量, 称为能量-动量张量 (stress-energy tensor)。为了得到这一点, 我们令 $x_\nu \rightarrow x'_\nu = x_\nu + \epsilon_\nu$, 而时空平移时, 不论是向量场还是标量场都不会改变, 即 $\phi'(x'^\mu) = \phi(x^\mu)$ 或 $\delta \phi = 0$ 。因此, 我们可以从 $\delta \mathcal{L} = -\epsilon_\nu \partial_\mu T^{\mu\nu} = 0$ 对任意的无穷小平移 ϵ_ν 都成立, 从而得到

$$\partial_\mu T^{\mu\nu} = 0. \quad (4.174)$$

显然上式的 $\nu = 0$ 分量对应着能量守恒, 而 $\nu = i$ 分量对应着动量守恒, 因此, 张量 $T^{\mu\nu}$ 被称为能量-动量张量。如果我们记 $\mathcal{E} \equiv T^{00}/c$, $\mathbf{P} \equiv (T^{10}, T^{20}, T^{30})$, 则能量守恒表达为

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathbf{P} = 0. \quad (4.175)$$

⁶如果仅是总量守恒, 则可以出现的情况是, 在宇宙中某处减少了 ΔQ 的同时, 在另一处增加了 ΔQ , 在两个地点之间不存在联系。

同时记动量密度为 \mathbf{g} ，其分量为 $g^i = T^{0i}/c$ ，且定义动量张量 \mathbf{G} ，其分量 $G^{ij} = T^{ij}$ ，则 $\nu = i$ 时的动量守恒式可表达为

$$\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \mathbf{G} = 0. \quad (4.176)$$

因此能量-动量张量 $T^{\mu\nu}$ 的各项应为

$$T^{\mu\nu} = \begin{pmatrix} c\mathcal{E} & \mathbf{P} \\ c\mathbf{g} & \mathbf{G} \end{pmatrix} \quad (4.177)$$

同时，从上式可以看出 $T^{\mu\nu}$ 的定义不是唯一的。如果我们有一三阶张量， $K^{\lambda\mu\nu}$ ，且其关于 μ 与 λ 指标反对称，即 $K^{\lambda\mu\nu} = -K^{\mu\lambda\nu}$ ，则我们可以构建一个新的二阶张量， $\hat{T}^{\mu\nu} = T^{\mu\nu} + \partial_\lambda K^{\lambda\mu\nu}$ 。因为 K 张量的反对称性，容易证明 $\partial_\mu \partial_\lambda K^{\lambda\mu\nu} = 0$ 。据此，新的张量 \hat{T} 同样满足其四维散度为零，即 $\partial_\mu \hat{T}^{\mu\nu} = 0$ 。利用这一方式，我们可以将不对称的能动张量构造为对称张量。这样，从 \hat{T} 我们也可以看出能流 ($P^i = \hat{T}^{i0}$) 与动量密度 ($g^i = \hat{T}^{0i}/c$) 之间差 $1/c$ 因子，与我们在前一章里所得结论相同。

从粒子体系的拉格朗日量很容易看出，对场而言，如果 $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$ ，即其不显示的依赖于时空坐标时，则系统具有时空平移对称性，此时场的能量动量守恒性得到满足。这一点也可以直接从诺特定理得到。当时空作一无穷小平移时， $x'_\mu = x_\mu + \epsilon_\mu$ ，则相应的拉格朗日量变为

$$\delta \mathcal{L} = \mathcal{L}(\phi'(x'), \partial'_\mu \phi(x'), x') - \mathcal{L}(\phi(x), \partial_\mu \phi(x), x) \quad (4.178)$$

因为在时空平移时， $\phi'(x') = \phi(x)$ ，且 $\partial'_\mu \phi' = \partial_\mu \phi$ ，因此我们有

$$\delta \mathcal{L} = \mathcal{L}(\phi(x), \partial_\mu \phi(x), x') - \mathcal{L}(\phi(x), \partial_\mu \phi(x), x) = \epsilon_\mu \left. \frac{\partial \mathcal{L}}{\partial x^\mu} \right|_{\phi, \partial_\mu \phi} \quad (4.179)$$

这里最后一项中对 x^μ 的偏导是同时固定了 ϕ 与 $\partial_\mu \phi$ 。因此，当 $\epsilon_\mu \neq 0$ 时，要使 $\delta \mathcal{L} = 0$ ，我们一定要求

$$\left. \frac{\partial \mathcal{L}}{\partial x^\mu} \right|_{\phi, \partial_\mu \phi} = 0 \Rightarrow \mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi). \quad (4.180)$$

即拉格朗日量密度不显式的依赖于时空坐标。因此，之前的例子里，自由场的拉格朗日量都具有这一形式。反过来，如果 $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi)$ ，则利用场的拉格朗日方程，也可以直接证明 $\partial_\mu T^{\mu\nu} = 0$ 。这一问题放在习题中。

4.6.3 电磁场的能量-动量张量

在本节的最后，我们将上述能量-张量定义运用到电磁场中，得到电磁场的能量-动量张量。利用 $\mathcal{L} = -(1/4\mu_0)F^{\alpha\beta}F_{\alpha\beta}$ ，我们有

$$\Pi^{\mu\lambda} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\lambda)} = -\frac{1}{4\mu_0} \frac{\partial}{\partial(\partial_\mu A_\lambda)} (F^{\alpha\beta} F_{\alpha\beta}). \quad (4.181)$$

从之前的计算中，我们有

$$\frac{\partial}{\partial(\partial_\mu A_\lambda)} F^{\alpha\beta} = \frac{\partial}{\partial(\partial_\mu A_\lambda)} (\partial^\alpha A^\beta - \partial^\beta A^\alpha) = g^{\alpha\mu} g^{\beta\lambda} - g^{\beta\mu} g^{\alpha\lambda}. \quad (4.182)$$

类似的，

$$\frac{\partial}{\partial(\partial_\mu A_\lambda)} F_{\alpha\beta} = \delta_\alpha^\mu \delta_\beta^\lambda - \delta_\beta^\mu \delta_\alpha^\lambda. \quad (4.183)$$

因此，

$$\Pi^{\mu\lambda} = -\frac{1}{4\mu_0} \left\{ [g^{\alpha\mu} g^{\beta\lambda} - g^{\beta\mu} g^{\alpha\lambda}] F_{\alpha\beta} - [\delta_\alpha^\mu \delta_\beta^\lambda - \delta_\beta^\mu \delta_\alpha^\lambda] F^{\alpha\beta} \right\} = -\frac{1}{\mu_0} F^{\mu\lambda}. \quad (4.184)$$

在最后一步，我们利用了 $F^{\mu\lambda} = -F^{\lambda\mu}$. 因此，相应的能动张量，暂记为 \tilde{T} ,

$$\tilde{T}^{\mu\nu} = \Pi^{\mu\lambda} \partial^\nu A_\lambda - \eta^{\mu\nu} \mathcal{L} = -\frac{1}{4\mu_0} [4F^{\mu\lambda} \partial^\nu A_\lambda - \eta^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta}]. \quad (4.185)$$

从右式第一项可以看出，这一能动张量并不对称，即 $\tilde{T}^{\mu\nu} \neq \tilde{T}^{\nu\mu}$ ，而且其各项的定义与前一节中所得电磁场能量动量相关物理量并不一致。

现在我们利用一三阶张量 $K^{\lambda\mu\nu}$ 将上式表达为一对称的能量动量电磁场张量，而所选的张量为

$$K^{\lambda\mu\nu} = -\frac{1}{\mu_0} F^{\lambda\mu} A^\nu, \quad (4.186)$$

显然

$$\partial_\lambda K^{\lambda\mu\nu} = -\frac{1}{\mu_0} F^{\lambda\mu} \partial_\lambda A^\nu = \frac{1}{\mu_0} F^{\mu\lambda} \partial_\lambda A^\nu. \quad (4.187)$$

这里我们利用了真空中场的方程 $\partial_\lambda F^{\lambda\mu} = 0$. 因此，新构建的张量

$$T^{\mu\nu} = \tilde{T}^{\mu\nu} + \partial_\lambda K^{\lambda\mu\nu} = \frac{1}{\mu_0} \left(F^{\mu\lambda} F_{\lambda}^{\nu} + \frac{1}{4} \eta^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta} \right). \quad (4.188)$$

现在我们来查看这一张量的各个分量。首先我们有 $F^{\alpha\beta} F_{\alpha\beta} = 2(B^2 - E^2/c^2) = 2(B^2 - \mu_0 \epsilon_0 E^2)$, 因此，

$$\frac{1}{4\mu_0} \eta^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta} = \eta^{\mu\nu} \left(\frac{1}{2\mu_0} B^2 - \frac{1}{2} \epsilon_0 E^2 \right). \quad (4.189)$$

另一方面我们有 $F^{\mu\lambda} F_{\lambda}^{\nu} = F^{\mu\lambda} F_{\lambda\rho} \eta^{\rho\nu}$. 这样，

$$\frac{1}{\mu_0} F^{0\lambda} F_{\lambda\rho} \eta^{\rho 0} = \frac{1}{\mu_0} F^{0i} F_{i0} \eta^{00} = \frac{1}{\mu_0} (E^i/c)(E_i/c) = \frac{1}{\mu_0 c^2} E^2 = \epsilon_0 E^2. \quad (4.190)$$

因此，

$$T^{00} = \frac{1}{2} \left(\epsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right), \quad (4.191)$$

即电磁场的能量密度。而

$$\frac{1}{\mu_0} F^{0\lambda} F_{\lambda\rho} \eta^{\rho i} = -\frac{1}{\mu_0} F^{0\lambda} F_{\lambda i} = -\frac{1}{\mu_0} F^{0j} F_{ji} = -\frac{1}{\mu_0} (E^j/c) \epsilon_{jik} B^k = \frac{1}{c} \frac{1}{\mu_0} (\mathbf{E} \times \mathbf{B})_i \quad (4.192)$$

因为 $\eta^{0i} = 0$, 所以我们有

$$T^{0i} = \frac{1}{c} \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}. \quad (4.193)$$

即电磁场的能流 (或坡印廷矢量) 除以光速，或者是电磁场的动量密度。这样 $\partial_\mu T^{0\mu}$ 对应的方程即为

$$\frac{\partial}{\partial t} \left[\frac{1}{2} \left(\epsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right) \right] + \nabla \cdot \left(\frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) = 0. \quad (4.194)$$

即能量守恒方程的连续性形式。

最后，对于 $\mu = i \neq 0, \nu = j \neq 0$, 我们有

$$\frac{1}{\mu_0} F^{i\lambda} F_{\lambda\rho} \eta^{\rho j} = \frac{1}{\mu_0} F^{i0} F_{0\rho} \eta^{\rho j} + \frac{1}{\mu_0} F^{ik} F_{kl} \eta^{lj} \quad (4.195)$$

其中第一项为

$$\frac{1}{\mu_0} F^{i0} F_{0\rho} \eta^{\rho j} = \frac{1}{\mu_0} \frac{E^i}{c} \frac{E_\rho}{c} \eta^{\rho j} = \epsilon_0 E^i E_\rho \eta^{\rho j} = -\epsilon_0 E^i E^j. \quad (4.196)$$

这里我们利用了 $\eta^{jj} = -1$ (此处不对 j 求和)。而第二项为

$$\frac{1}{\mu_0} F^{ik} F_{kl} \eta^{lj} = \frac{1}{\mu_0} \epsilon^{ikm} B_m \epsilon_{kln} B^n \eta^{lj} = -\frac{1}{\mu_0} [\delta_l^i \delta_n^m - \delta_n^i \delta_l^m] B_m B^n \eta^{lj} = -\frac{1}{\mu_0} B^2 \eta^{ij} - \frac{1}{\mu_0} B^i B^j. \quad (4.197)$$

将上述两项与 $(1/4\mu_0)\eta^{ij}F^{\alpha\beta}F_{\alpha\beta}$ 结合在一起，我们有

$$T^{ij} = -\epsilon_0 E^i E^j - \frac{1}{\mu_0} B^i B^j - \frac{1}{2\mu_0} B^2 \eta^{ij} - \frac{1}{2} \epsilon_0 E^2 \eta^{ij}. \quad (4.198)$$

如果有三维的向量与张量表示，则有

$$T^{ij} = -\left[\epsilon \mathbf{E} \mathbf{E} + \frac{1}{\mu_0} \mathbf{B} \mathbf{B} - \left(\frac{1}{2\mu_0} B^2 + \frac{1}{2} \epsilon_0 E^2 \right) \mathbf{I} \right]. \quad (4.199)$$

这里我们同样利用了当 $i = j$ 时， $\eta^{ij} = -1$ 。注意到 $T^{i0} = T^{0i}$ ，因此，方程 $\partial_\mu T^{i\mu} = 0$ 可以表示为

$$\frac{\partial}{\partial t} (\epsilon_0 \mathbf{E} \times \mathbf{B}) + \nabla \cdot \mathbf{G} = 0. \quad (4.200)$$

这里 $G_{ij} = -T_{ij}$ ，即为电磁场的动量通量张量。

第五章 静电场

5.1 静电场的方程

在这一章里，我们考虑不随时间变化的场。当不随时间变化时，麦克斯韦方程为

$$\nabla \times \mathbf{E} = 0, \quad (5.1)$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (5.2)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}, \quad (5.3)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (5.4)$$

从这组方程可以明显的看出，当场不随时间变化时，电场和磁场是解耦的。这一点从它们的定义也可以看出，当 $\partial/\partial t = 0$ 时，

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} = -\nabla\phi,$$

$$\mathbf{B} = \nabla \times \mathbf{A}.$$

意味着 \mathbf{E} 完全由 ϕ 决定，而 \mathbf{B} 完全由 \mathbf{A} 决定。所以我们可以分开考虑静电和静磁场。

在本章，我们考虑静电场。对静电场而言， $\mathbf{E} = -\nabla\phi$ 。显然，一般情况下用 ϕ 来研究静电场问题会大大简化分析，因为只需处理一个标量。一个例外是如果所考虑的问题具有某种对称性，则可以考虑利用静电场的高斯定理去直接获得电场 \mathbf{E} 。在静电场情况下，标势 ϕ 也即静电势，其所满足的方程为

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0}. \quad (5.5)$$

如果所考虑的区域无源， $\rho = 0$ ，则

$$\nabla^2 \phi = 0. \quad (5.6)$$

这一方程称为拉普拉斯方程 (Laplace equation)。相应的， ∇^2 称为拉普拉斯算符。注意到在不随时间变换的场的情况下，矢势 \mathbf{A} 所满足的方程为 $\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j}$ ，与标势 ϕ 的方程高度相似。更进一步的，在无源情况下，磁场 \mathbf{B} 的旋度为 0，因此，也可以用一标势函数 ψ 代表磁场即， $\mathbf{B} = -\nabla\psi$ 。因此，在本章中针对静电场问题所发展起来的许多计算技巧也可以运用在静磁场问题上。

为了得到这个方程的通解，我们利用场的叠加性。从点电荷的静电场出发。假设点电荷位于原点，这个问题显然具有对称性，我们可以利用高斯定理获得相应的解。在此情况下， $\mathbf{E} = E(r)\mathbf{e}_r$ ，这里 r 是场点到点电荷的距离，且 $\mathbf{e}_r = \mathbf{r}/r$ 。从高斯定理，

$$\oint \mathbf{E} \cdot d\mathbf{S} = \int \nabla \cdot \mathbf{E} dV = \int \frac{\rho}{\epsilon_0} dV = \frac{q}{\epsilon_0}, \quad (5.7)$$

这里我们利用了 $\rho = q\delta(\mathbf{r})$ 。将积分区域选为以点电荷为中心的球，则

$$\oint \mathbf{E} \cdot d\mathbf{S} = 4\pi r^2 E = \frac{q}{\epsilon_0}. \quad (5.8)$$

所以，一个点电荷的静电场为

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \mathbf{e}_r.$$

而相应的势函数显然为

$$\phi = \frac{1}{4\pi\epsilon_0} \frac{q}{r},$$

因为

$$-\nabla\phi = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \nabla r = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \mathbf{e}_r = \mathbf{E}. \quad (5.9)$$

利用上式, 我们也可以得到一个关于 $\delta(\mathbf{r})$ 的有用等式。将 $\phi = q/4\pi\epsilon_0 r$ 代入 $\nabla^2\phi = -\rho/\epsilon_0$, 对一个位于原点的点电荷而言, $\rho = q\delta(\mathbf{r})$, 因此

$$\nabla^2 \left(\frac{q}{4\pi\epsilon_0 r} \right) = -\frac{q\delta(\mathbf{r})}{\epsilon_0} \quad (5.10)$$

或者

$$\nabla^2 \left(\frac{1}{r} \right) = -4\pi\delta(\mathbf{r}) \quad (5.11)$$

对于一般的电荷分布, $\rho = \rho(\mathbf{r})$, 其电场可以利用叠加原理得到。

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \sum \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^2} \mathbf{e}_{\mathbf{r}-\mathbf{r}'} \Delta V' = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{R^2} \mathbf{e}_R dV'. \quad (5.12)$$

在最后一步里, 我们定义 $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}'$, 为从源点 (\mathbf{r}') 指向场点的向量, 而且 $R = |\mathbf{R}|$ 。类似的, 对一般的电荷分布而言,

$$\phi = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{R} dV'. \quad (5.13)$$

此即方程 (5.5) 的解。

5.2 静电场的能量

对静电场而言, 场的总能量为

$$U = \frac{\epsilon_0}{2} \int E^2 dV, \quad (5.14)$$

这里对全空间区域积分。利用 $\mathbf{E} = -\nabla\phi$, 我们可将方程 (5.14) 改写为

$$U = -\frac{\epsilon_0}{2} \int \mathbf{E} \cdot \nabla\phi dV = -\frac{\epsilon_0}{2} \int \nabla \cdot (\mathbf{E}\phi) dV + \frac{\epsilon_0}{2} \int \phi \nabla \cdot \mathbf{E} dV. \quad (5.15)$$

再利用高斯定理,

$$\frac{\epsilon_0}{2} \int \nabla \cdot (\mathbf{E}\phi) dV = \frac{\epsilon_0}{2} \oint \phi \mathbf{E} \cdot d\mathbf{S} = 0, \quad (5.16)$$

因为 $\phi \mathbf{E} \sim R^{-3}$, $S \sim R^2$, 且积分面是以 R 为半径的球面, 所以该积分随距离变化的关系为 R^{-1} 。随着 $R \rightarrow \infty$, 积分趋向 0。因此, 静电场的总能量为

$$U = \frac{\epsilon_0}{2} \int \phi \nabla \cdot \mathbf{E} dV = \frac{1}{2} \int \rho \phi dV, \quad (5.17)$$

这里我们利用了 $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ 。对一个电荷系统而言, $\rho = \sum q_i \delta(\mathbf{x} - \mathbf{x}_a)$,

$$U = \frac{1}{2} \int \rho \phi dV = \frac{1}{2} \sum_a q_a \phi_a, \quad (5.18)$$

这里 $\phi_a \equiv \phi(\mathbf{x}_a)$ 所有电荷在 \mathbf{x}_a 处的电势。

现在我们考虑单个点电荷, $U = q\phi/2$ 。但是, $\phi = q/r$, 在此情况下, $r = 0$ 导致 $U = \infty$! 如果 U 是无穷大, 则相应的质量 $m = U/c^2$ 必然也是无穷大。这种不符合实际的结果显示电磁场理论是有局限性的。这种能量无穷大的问题被称为自能 (self-energy) 问题, 是目前物理学里仍然没有很好解决的问题。

自能问题来自于我们假定电荷是点粒子。如果我们将电子看作半径为 R_0 的球，则 $U = e^2/8\pi\epsilon_0 R_0$ ，不再是无穷大。该能量应该于 mc^2 同一量级，所以我们定义

$$R_0 \sim \frac{e^2}{4\pi\epsilon_0 mc^2}, \quad (5.19)$$

为经典电子半径。经典电磁场理论在 $r < R_0$ 是有问题的。

现在我们回到多电荷体系，

$$\phi_a = \frac{1}{4\pi\epsilon_0} \sum_b \frac{q_b}{r_{ab}}, \quad (5.20)$$

这里 $r_{ab} = |\mathbf{x}'_a - \mathbf{x}'_b|$ 是两个电荷间的距离。这里 ϕ 包含两部分：无穷大的自能部分与电荷间的相互作用能。从现在开始，我们讨论场的能量时，不再考虑自能，而只讲相互作用能。从而

$$U^* = \frac{1}{2} \sum_a q_a \phi_a^*, \quad (5.21)$$

这里

$$\phi_a^* = \frac{1}{4\pi\epsilon_0} \sum_{b \neq a} \frac{q_b}{r_{ab}}, \quad (5.22)$$

为其它所有电荷在电荷 a 处的势，即去除了无穷大的自势 (self-potential)。相互作用的静电能为

$$U^* = \frac{1}{2} \sum_a q_a \sum_{b \neq a} \frac{q_b}{4\pi\epsilon_0 r_{ab}} = \frac{1}{8\pi\epsilon_0} \sum_{a \neq b} \frac{q_a q_b}{r_{ab}}. \quad (5.23)$$

例如，两个电荷构成的体系，其相互作用能量 U^* 为

$$U^* = \frac{1}{4\pi\epsilon_0} \left(\underbrace{\frac{1}{2} \frac{q_1 q_2}{r_{12}}}_{i=1, j=2} + \underbrace{\frac{1}{2} \frac{q_2 q_1}{r_{21}}}_{i=2, j=1} \right) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}}. \quad (5.24)$$

5.3 静电场的多极展开

虽然上述通解形式简单，但运用在实际问题上（如解释介质在电场中的极化现象），我们常常不会将每一个点电荷在空间某处 \mathbf{x} 的电场求出，然后对所有点电荷的相应电场求和。例如，一团水分子在某处的电场为何？直观反应是电场应该是 0，因为水分子整体都是电中性的。这里的思考过程实际上是先求出整个水分子的整体电荷为 0，然后由此得到空间某处的静电场为 0。可是如果再进一步，每个水分子的正负电荷中心并不是完全重合的，所以其在远处的电场应该不为 0。这两个回答，看似矛盾，其实不然。因为前者对应的是远处电场的 0 阶解，而后者则对应着高阶解，或者说是由电荷分布中更精细的结构导致的场。另一个不会用点电荷的场叠加求解实际问题的原因是计算量过大，细节在后面介绍。

在这一节里，我们介绍如何获得静电场的近似解。通常获得的是近似解的途径是通过找到系统里的一小量，然后利用该小量进行泰勒展开¹。我们先建立坐标系原点在电荷体系内，因此系统的空间尺度为 x' ，而场点到源点的距离的尺度为 x ，我们利用泰勒展开的主要项获得近似解的前提条件就是场点到源点的尺度远大于源本身的空间尺度，即 $x' \ll x$ 。此时我们的系统也被称为双空间尺度系统，其小量可以理解为 $\epsilon \equiv x'/x$ 。因为存在该小量，静电场的低阶项占主导，可以看成其近似解。对于静电场问题的泰勒展开近似解，其过程等价于函数的多极展开。因此对于静电场与磁场，习惯上把这种展开过程称为多极展开 (multipole expansion)，关于这一点，参见习题。

¹ 泰勒展开或多极展开本身是不需要小量的。但是物理学里，这种展开很有用的原因通常是几个低阶项为精确解提供了很好的近似，这里小量的存在就很重要了。

首先，对场点 \mathbf{x} ，其电势为

$$\phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \sum_a \frac{q_a}{|\mathbf{x} - \mathbf{x}'_a|}. \quad (5.25)$$

利用 $|\mathbf{x}'| \ll |\mathbf{x}|$ ，可以将一般的函数 $f(\mathbf{x} - \mathbf{x}')$ 作泰勒展开

$$f(\mathbf{x} - \mathbf{x}') = f(\mathbf{x}) - \mathbf{x}' \cdot \nabla f(\mathbf{x}) + \frac{1}{2} x'_i x'_j \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}) + \cdots \quad (5.26)$$

在我们的问题里， $f(\mathbf{x} - \mathbf{x}') = \phi(\mathbf{x} - \mathbf{x}')$ ，且

$$\phi = \frac{1}{4\pi\epsilon_0} \sum_a \frac{q}{|\mathbf{x} - \mathbf{x}'|} = \phi^{(0)} + \phi^{(1)} + \phi^{(2)} + \cdots, \quad (5.27)$$

这里 $\phi^{(i)}$ 代表第 i 阶项。我们可以比较不同项的量级，如 $\phi^{(1)}$ 和 $\phi^{(0)}$ ，

$$\frac{\phi^{(1)}}{\phi^{(0)}} \sim \frac{\mathbf{x}' \cdot \nabla \phi}{\phi} \sim \frac{|\mathbf{x}'| |\phi| / |\mathbf{x}|}{\phi} \sim \frac{|\mathbf{x}'|}{|\mathbf{x}|} \ll 1. \quad (5.28)$$

显然 $\phi^{(1)}/\phi^{(0)} \sim \mathcal{O}(\epsilon)$ 。类似的，你可以证明 $\phi^{(2)}/\phi^{(0)} \sim \mathcal{O}(\epsilon^2)$ 。需要注意的一点是，当我们说某一项是 $\mathcal{O}(\epsilon^n)$ 时，指的并非这一项的绝对大小，而是说，当 $\epsilon \rightarrow 0$ 时，这一项趋向 0 的速度与 ϵ^n 一致。所以 10^{-16} 是 0 阶项，而 1000ϵ 是一阶项。

将方程 (5.26) 应用到 (5.27)，我们有

$$\phi^{(0)} = \phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \sum_a \frac{q}{r}, \quad \text{with } r \equiv |\mathbf{x}|, \quad (5.29)$$

$$\phi^{(1)} = -\frac{1}{4\pi\epsilon_0} \sum_a q x'_a \cdot \nabla \frac{1}{r}, \quad (5.30)$$

$$\phi^{(2)} = \frac{1}{8\pi\epsilon_0} \sum_a q x'_i x'_j \frac{\partial^2}{\partial x_i \partial x_j} \left(\frac{1}{r} \right). \quad (5.31)$$

如果我们仔细观察这些项，有两点值得注意。首先，每一阶的近似项都具有 $f(\mathbf{x}')g(\mathbf{x})$ 的形式。其中 f 只于 \mathbf{x}' 有关，而 g 只于 \mathbf{x} 有关。这一点与一般解的表达式 (5.27) 不同。正是因为这一点，利用近似解在计算机求解时往往会大大节约计算时间。如果我们的电荷系统有 N 个例子，而我们需要计算 M 个场点的电场。利用一般解 (5.27)，则需要的计算次数为 $\mathcal{O}(NM)$ ，因为对每一个场点，都需要完成计算 N 次。而如果例用近似解，则需要的计算量为 $\mathcal{O}(N + M)$ 。其中 N 次计算是获得 $f(\mathbf{x}')$ ，对每个场点，只需要计算 $g(\mathbf{x})$ 就可以。因此，对 N 和 M 非常大的情况，近似解会大大节约计算时间。例如 $N = 10^{10}$ 个粒子，而 $M = 10^5$ 个场点，一次计算约耗时 10^{-10} 秒，则通解需要 $10^{15} \times 10^{-10} = 10^5$ 秒或者 27 小时，而近似解则只需要 $10^{10} + 10^5 \approx 10^{10}$ 次计算，耗时约为 1 秒。两者相差 10^5 倍。还有一些实际的问题是现有的计算机能力不可能解决的。因此，利用多极展开来计算场是粒子类模拟的一种重要手段。其次，我们发现， $f(\mathbf{x}')$ 的形式很有规律：

$$0 \text{ 阶} \cdots \cdots \sum_a q = \int \rho(\mathbf{x}') dV', \quad (5.32)$$

$$1 \text{ 阶} \cdots \cdots \sum_a q x'_a = \int \rho(\mathbf{x}') \mathbf{x}' dV', \quad (5.33)$$

$$2 \text{ 阶} \cdots \cdots \sum_{a,b} q x'_a x'_b = \int \rho(\mathbf{x}') \mathbf{x}' \mathbf{x}' dV'. \quad (5.34)$$

一般的，如果 $f(x)$ 是关于某个函数，则 $\int f(x) x^n dx$ 称为该函数的 n 阶矩 (moment)，而且每一阶矩都有其含义，且表达了该函数的形状的部分性质。例如，设 $f(\mathbf{x}) = \rho_m(\mathbf{x})$ 是质量密度分布函数，则 $\int \rho_m(\mathbf{x}) dV$ 表示的是所有质量汇在一起的总质量； $\int \rho_m(\mathbf{x}) \mathbf{x} dV$ 则表示了重心的位置； $\int \rho_m(\mathbf{x}) \mathbf{x} \mathbf{x} dV$ 则与转动惯量张量相关。又比如，假设 $f(x)$ 是一个概率分布函数，即相关过程发生在 x 与 $x + dx$ 间的概率为 $f(x) dx$ ，则其 1 阶矩， $\int f(x) x dx$ 表示的是该过程的期望值；二阶矩 $\int f(x) x^2 dx$ 表达的是方差 (variance)²。矩的概念在后续的课程里也会经常遇到。

²更一般的 n 阶矩的定义形式为 $f(x)(x - \mu)^n dx$ ，这里 μ 是一常数。如果是方差，则 $n = 2$ 且 μ 是 x 的平均值。

5.3.1 电单极矩与电偶极矩

回到电场的展开式。显然，0 阶项， $\phi^{(0)}$ ，为电荷分布函数的 0 阶矩，一般称为电单极矩 (monopole)，

$$\phi^{(0)} = \phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \frac{\sum_i q_i}{r}. \quad (5.35)$$

即忽略电荷分布，将所有电荷聚在一起的总电荷 $Q = \sum_a q$ 所对应的势。显然，如果 $Q \neq 0$ ，则 $\phi^{(0)}$ 占主导，为主要近似解。该主要近似解也与我们的直观感受一致，如果我们离一堆电荷很远，则所在点的电场为将所有电荷聚在一起所对应的电场。具体到水分子的例子，我们会直观的认为相应的电场为 0，应该是因为大脑自动取了 0 阶近似。

如果 $Q = 0$ ，则 $\phi^{(0)} = 0$ 。如果我们需要得到一个更好的近似解，则需要更高阶项。一阶近似解， $\phi^{(1)}$ ，与电荷分布函数的 1 阶矩有关。引入电偶极矩 (electrostatic dipole moment) \mathbf{p} ，定义为

$$\mathbf{p} = \sum_a q\mathbf{x}' = \int \rho(\mathbf{x}')\mathbf{x}'dV. \quad (5.36)$$

则我们可将电偶极势 $\phi^{(1)}$ 表达为

$$\phi^{(1)} = -\frac{1}{4\pi\epsilon_0} \sum_a q\mathbf{x}' \cdot \nabla \frac{1}{r} = -\frac{\mathbf{p}}{4\pi\epsilon_0} \cdot \nabla \frac{1}{r} = \frac{\mathbf{p} \cdot \mathbf{x}}{4\pi\epsilon_0 r^3}. \quad (5.37)$$

对应的电场可以从 $\mathbf{E} = -\nabla\phi$ 得到。需要注意的是，这里 $\nabla \equiv \partial/\partial\mathbf{x}$ 是对场点 \mathbf{x} 求偏导，因此源点坐标 \mathbf{x}' 是固定不变的，这样电偶极矩 $\mathbf{p}(\mathbf{x}')$ 在求偏导时可以当作一个常量。因此，

$$\mathbf{E} = -\nabla \frac{\mathbf{p} \cdot \mathbf{x}}{4\pi\epsilon_0 r^3} = -\frac{1}{4\pi\epsilon_0 r^3} \nabla(\mathbf{p} \cdot \mathbf{x}) - \frac{1}{4\pi\epsilon_0} (\mathbf{p} \cdot \mathbf{x}) \nabla \frac{1}{r^3} \quad (5.38)$$

$$= -\frac{1}{4\pi\epsilon_0 r^3} \mathbf{p} \cdot \nabla \mathbf{x} - \frac{\mathbf{p} \cdot \mathbf{x}}{4\pi\epsilon_0} \frac{3}{r^4} \nabla r \quad (5.39)$$

$$= -\frac{\mathbf{p}}{4\pi\epsilon_0 r^3} + \frac{3\mathbf{p} \cdot \mathbf{x}}{4\pi\epsilon_0 r^4} \frac{\mathbf{x}}{r} \quad (5.40)$$

如果令 $\mathbf{n} = \mathbf{x}/r$ ，则

$$\mathbf{E}^{(1)} = \frac{3(\mathbf{n} \cdot \mathbf{p})\mathbf{n} - \mathbf{p}}{4\pi\epsilon_0 r^3}. \quad (5.41)$$

注意 $\mathbf{E}^{(1)} \sim r^{-3}$ ，因此随距离衰减的比 $\mathbf{E}^{(0)} \sim r^{-2}$ 更快。

电偶极场的分量可以从式 (5.41) 直接获得。假设 $\mathbf{p} = p\mathbf{e}_z$ ， $\theta = \langle \mathbf{n}, \mathbf{e}_z \rangle$ ， $\mathbf{e}_\perp \perp \mathbf{e}_z$ ，则

$$E_z = \frac{3(\mathbf{n} \cdot \mathbf{p})\mathbf{n} - \mathbf{p}}{4\pi\epsilon_0 r^3} \cdot \mathbf{e}_z = p \frac{3\cos^2\theta - 1}{4\pi\epsilon_0 r^3}, \quad (5.42)$$

$$E_\perp = \frac{3(\mathbf{n} \cdot \mathbf{p})\mathbf{n} - \mathbf{p}}{4\pi\epsilon_0 r^3} \cdot \mathbf{e}_\perp = p \frac{3\cos\theta\sin\theta}{4\pi\epsilon_0 r^3}. \quad (5.43)$$

或者使用球坐标系，

$$E_r = \mathbf{E} \cdot \mathbf{n} = \frac{3(\mathbf{n} \cdot \mathbf{p})(\mathbf{n} \cdot \mathbf{n}) - \mathbf{p} \cdot \mathbf{n}}{4\pi\epsilon_0 r^3} = p \frac{2\cos\theta}{4\pi\epsilon_0 r^3} = p \frac{\cos\theta}{2\pi\epsilon_0 r^3}, \quad (5.44)$$

$$E_\theta = \mathbf{E} \cdot \mathbf{e}_\theta = \frac{3(\mathbf{n} \cdot \mathbf{p})(\mathbf{n} \cdot \mathbf{e}_\theta) - \mathbf{p} \cdot \mathbf{e}_\theta}{4\pi\epsilon_0 r^3} = p \frac{\sin\theta}{4\pi\epsilon_0 r^3}. \quad (5.45)$$

两者可以合并为

$$\mathbf{E} = \frac{p}{4\pi\epsilon_0 r^3} (2\cos\theta\mathbf{e}_r + \sin\theta\mathbf{e}_\theta). \quad (5.46)$$

后面我们会发现，磁偶极场也具有一样的形式，仅仅是将 p/ϵ_0 替换成 $\mu_0 m$ 即可，其中 m 是磁偶极矩。

关于电偶极矩，有两个性质值得一提。一，如果总电荷 $\sum_a q = 0$ ，则电偶极矩 \mathbf{p} 与坐标原点的选取无关。例如，我们选另一坐标系， $\mathbf{x}'' = \mathbf{x}' + \mathbf{a}$ ，则

$$\mathbf{p}' = \sum q \mathbf{x}'' = \sum q \mathbf{x}' + \sum q \mathbf{a} = \sum q \mathbf{x}' = \mathbf{p}. \quad (5.47)$$

而如果 $Q \equiv \sum_a q \neq 0$ ，则 $\mathbf{p}' = \mathbf{p} + Q\mathbf{a}$ 。二，对很多物理问题，有一个简单的模型对理解物理过程很有帮助。一个常用的电偶极模型是一对正负电荷（电偶极子） $|q|(\mathbf{x}'_+)$ 与 $-|q|(\mathbf{x}'_-)$ ，其电偶极矩为

$$\mathbf{p} = |q|\mathbf{x}'_+ - |q|\mathbf{x}'_- = |q|\mathbf{l}, \quad \text{with } \mathbf{l} = \mathbf{x}'_+ - \mathbf{x}'_-. \quad (5.48)$$

5.3.2 电四极矩

接下来我们继续分析二阶势， $\phi^{(2)}$ 。如果 Q 与 \mathbf{p} 都为 0，则其为主导项。将方程 (5.31) 重写为

$$\phi^{(2)} = \frac{1}{8\pi\epsilon_0} \sum_a q x'_i x'_j \frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{r}. \quad (5.49)$$

暂时我们可以定义系统的电四阶矩 (quadrupole moment) 为

$$\mathbf{D} = \sum_a 3q \mathbf{x}' \mathbf{x}', \quad \text{or} \quad D_{ij} = \sum_a 3q x'_i x'_j, \quad (5.50)$$

则

$$\phi^{(2)} = \frac{D_{ij}}{24\pi\epsilon_0} \frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{r}. \quad (5.51)$$

称为电四阶势。注意到 \mathbf{D} 有 6 个分量，

$$D_{11}, D_{22}, D_{33}, D_{12} = D_{21}, D_{13} = D_{31}, D_{23} = D_{32}. \quad (5.52)$$

但其实只有 5 个是独立的。对电荷系统外的场点 \mathbf{x} 而言，

$$\nabla^2 \frac{1}{r} = 0 \quad \text{一般情况下该式等于 } -4\pi\delta(\mathbf{r}). \quad (5.53)$$

我们可以将上式重写为

$$\nabla^2 \frac{1}{r} = \frac{\partial^2}{\partial x_i \partial x_i} \frac{1}{r} = \delta_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{r} = 0. \quad (5.54)$$

利用方程 (5.54)，电四极势 $\phi^{(2)}$ 可以写成

$$\phi^{(2)} = \frac{1}{24\pi\epsilon_0} (D_{ij} - r'^2 \delta_{ij}) \frac{\partial^2}{\partial x_i \partial x_j} \frac{1}{r}. \quad (5.55)$$

这里 $r' = |\mathbf{x}'|$ 。因此我们可以正式的定义电四极矩为

$$D_{ij} = \sum_a q (3x'_i x'_j - r'^2 \delta_{ij}). \quad (5.56)$$

该张量的迹 (trace) $D_{ij} = D_{11} + D_{22} + D_{33} = 0$ ，因此该张量仅有 5 个独立分量。电四极矩的这种定义方式与物体的转动惯量张量异曲同工。

5.4 外场中的电荷系统

在这一节里我们考虑一个在外场中的电荷系统在该场中的受力。这里的场是外部施加的，电荷系统自己产生的场相对外场是忽略的，因此这里也是测试粒子问题。我们先计算电荷体系在外场中的势能， $\phi(\mathbf{x})$,

$$U = \sum_a q_a \phi(\mathbf{x}'_a) = \sum_a q \phi(\mathbf{x}'). \text{ 省略“}a\text{”下标} \quad (5.57)$$

同样的将坐标系原点选在电荷体系内部，电荷 a 的坐标是 \mathbf{x}'_a 。假设 ϕ 在电荷体系的空间尺度上变化缓慢，我们可以利用电荷体系的空间尺度与外场变化的空间尺度之比作为泰勒展开的小量。对于场 $\phi(\mathbf{x})$ ，其空间变换尺度可以表示成

$$\frac{\phi}{\nabla \phi} \sim \frac{\phi}{\phi/L} \sim L, \quad (5.58)$$

所以我们选取的展开小量为

$$\epsilon \sim \frac{x'}{L}. \quad (5.59)$$

这样 $\phi(\mathbf{x}')$ 可以展开为

$$\phi(\mathbf{x}') = \phi(0) + \mathbf{x}' \cdot \nabla \phi + \frac{1}{2} x'_i x'_j \frac{\partial^2 \phi}{\partial x_i \partial x_j} + \dots \quad (5.60)$$

如果我们比较右边的第一项和第二项，

$$\frac{\mathbf{x}' \cdot \nabla \phi}{\phi(0)} \sim \frac{|\mathbf{x}'| \phi/L}{\phi} \sim \frac{|\mathbf{x}'|}{L} \sim \mathcal{O}(\epsilon). \quad (5.61)$$

即第二项比第一项高一阶。利用 ϕ 的展开式，类似于多极展开过程，电荷体系在外场中的势能为

$$U = U^{(0)} + U^{(1)} + U^{(2)} + \dots, \quad (5.62)$$

这里

$$U^{(0)} = \sum_a q \phi(0), \quad (5.63)$$

$$U^{(1)} = \sum_a q \mathbf{x}' \cdot (\nabla \phi)_0 = \mathbf{p} \cdot \nabla \phi. \quad (5.64)$$

我们该展开的每一阶显然都对应着电荷体系电荷分布的各阶矩。

对 0 阶而言， $U^{(0)} = \sum_a q \phi(0)$ ，所以是电单极对应的势能，即将所有电荷放在一起的总电荷 $Q = \sum_a q$ 放在外场中的势能。

第一阶对应电偶极矩项， $U^{(1)} = \mathbf{p} \cdot \nabla \phi$ ，或者用 \mathbf{E} 表达，则是 $U^{(1)} = -\mathbf{p} \cdot \mathbf{E}(0)$ 。利用该势能，我们可以写出电荷体系所感受到的力

$$\mathbf{f} = -\nabla(U^{(0)} + U^{(1)} + \dots) \approx \sum_a q \mathbf{E}(0) + \nabla(\mathbf{p} \cdot \mathbf{E}) \quad (5.65)$$

$$= Q \mathbf{E}(0) + (\mathbf{p} \cdot \nabla) \mathbf{E}. \quad (5.66)$$

如果 $Q = 0$ ，则

$$\mathbf{f} = (\mathbf{p} \cdot \nabla) \mathbf{E}. \quad (5.67)$$

在推导上式过程中，我们利用了³

$$\nabla(\mathbf{p} \cdot \mathbf{E}) = (\mathbf{p} \cdot \nabla) \mathbf{E} + \mathbf{p} \times (\nabla \times \mathbf{E}) = (\mathbf{p} \cdot \nabla) \mathbf{E}, \quad (5.68)$$

因为对静电场而言， $\nabla \times \mathbf{E} = 0$ 。

³注意 $\nabla \cdot \mathbf{p} = 0$ 且 $\nabla \times \mathbf{p} = 0$ ，因为 \mathbf{p} 是用电荷体系坐标 \mathbf{x}' 的函数，而从势求场的过程是对场点求偏导，即 $\nabla \equiv \partial/\partial \mathbf{x}$ 。

5.5 介质的极化过程

一般我们讨论的介质有两种：导体（conductors）和绝缘体（dielectrics）。在导体里，有可以自由移动的电荷，一般是电子。另一方面，绝缘介质里则没有自由电子，所有的电子都是被束缚在原子里的。尽管如此，这些被束缚住的电子还是可以在外场作用下发生位置的一点点改变。这样总的叠加效应可能会非常重要，会在介质内产生新的电场。在本小节里，我们利用之前的知识来理解介质的电极化过程。

5.5.1 非极化原子

假设我们有一个放在外场 \mathbf{E} 中的中性原子。因为原子整体呈电中性，所以很自然的其电单极项为 0。但是在外加电场作用下，正负电荷中心分离，从而获得了电偶极矩⁴。在外场 \mathbf{E} 不是太强的情况下，正负电荷会在新的位置上达到平衡，此时原子被极化（polarized），形成小的电偶极子。显然该电偶极子对应的电偶极矩方向与外场相同 \mathbf{E} ，因为正电荷沿着 \mathbf{E} ，而负电荷反向 \mathbf{E} 移动。可以简单证明此时 \mathbf{p} 大约正比于外场 \mathbf{E} ，或者

$$\mathbf{p} = \alpha \mathbf{E} \quad (5.69)$$

这里 α 被称为原子极化率（atomic polarizability），其数值依赖于具体的原子。

这里我们利用一个简单的原子模型来得到 α 的表达式。假设该原子为一正电荷的原子核（ q ）和一个球状电子云构成，电子云半径为 a ，总电量为 $-q$ 。没有外场时，电子云的中心与原子核中心重合，此时电偶极矩也为 0。但是如果有外场 \mathbf{E} ，原子核在沿电场方向上移动到新的位置，假设为距电子云中心 d 外。显然在平衡状态下，原子核受力平衡，即其所受外加电场力等于电子云对其吸引力。我们假定电子云仍然近似为半径为 a 的球形，则电子云在原子核位置处的电场为

$$E_e = \frac{qd}{4\pi\epsilon_0 a^3}. \quad (5.70)$$

在平衡时， $E_e = E$ ，或者

$$\frac{qd}{4\pi\epsilon_0 a^3} = E, \quad \text{or} \quad qd = 4\pi\epsilon_0 a^3 E. \quad (5.71)$$

注意到 qd 即是新形成的电偶极子对应的电偶极矩，所以对简单原子而言，

$$\mathbf{p} = 4\pi\epsilon_0 a^3 \mathbf{E}, \quad (5.72)$$

或者说原子极化率为 $\alpha = 4\pi\epsilon_0 a^3$ 。

对复杂原子而言，其极化形成的电偶极矩不一定与外加电场在同一方向。例如，对 CO_2 分子而言， $\mathbf{p} = \alpha_{\perp} \mathbf{E}_{\perp} + \alpha_{\parallel} \mathbf{E}_{\parallel}$ 。更一般的情况下，电极化率是用张量来表达的， $\mathbf{p} = \mathbf{A} \cdot \mathbf{E}$ ，这里 \mathbf{A} 称为极化张量（polarizability tensor）。

5.5.2 极化分子

如果分子原本的正负电荷是分离的，则其就算没有外加电场，都具有有的电偶极矩 \mathbf{p} ，例如水分子。但是因为热运动，这些分子的电偶极矩是杂乱无章的，整体上的电偶极矩， $\sum \mathbf{p}$ ，可以忽略不计。但是如果施加了外加电场，每个电偶极矩都会受到一个力矩， $\mathbf{N} = \mathbf{p} \times \mathbf{E}$ 。在该力矩作用下，各分子都会发生偏转，使得电偶极矩 \mathbf{p} 尽量与 \mathbf{E} 相同。虽然因为热运动， \mathbf{p} 不可能会完全平行于 \mathbf{E} ，但整体上还是会在 \mathbf{E} 的方向上形成可观的整体电偶极矩，从而形成可观的极化电场。

一般而言，上述两种极化过程是同时发生的，只不过重要性在不对情况下，有相对差别。

⁴如果电场过强，正负电荷还会分开，原子被电离

5.5.3 极化介质的电场

在之前的两个小节里，我们讨论了介质极化的两种可能过程。这一节里，我们来看一看，在介质因为极化具有了宏观上的偶极矩后，其产生的电场是什么样子。这里要利用到第 5.3 小节的知识。

首先，我们定义一下宏观的偶极矩。与连续分布介质的物理量一般用密度表示一下，我们定义单位体积内的总电偶极矩为 \mathbf{P} ，即极化强度 (polarization)。在定义了该物理量后，对于 dV 体积内的介质，其总的电偶极矩为 $d\mathbf{p} = \mathbf{P}dV$ ，从而产生的场为

$$d\varphi = \frac{1}{4\pi\epsilon_0} \frac{d\mathbf{p} \cdot \mathbf{R}}{R^3}, \quad (5.73)$$

这里 $\mathbf{R} = \mathbf{x} - \mathbf{x}'$ ，我们假定考虑的体积元位于 \mathbf{x}' 处。这样极化介质产生的总的场为

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \frac{\mathbf{P} \cdot \mathbf{R}}{R^3} dV' \quad (5.74)$$

如果对该电势求导数，便可以得到极化对应的电场。

另一方面，介质因为极化也会形成物理上的电荷聚集，从而有非零的电荷密度，这种电荷称为束缚电荷 (bounded charges)，与自由电荷对应。这样的束缚电荷有两种，一种是面束缚电荷密度， $\sigma = -\mathbf{P} \cdot \mathbf{n}$ ，另一种为体束缚面电荷密度， $\rho = -\nabla \cdot \mathbf{P}$ 。所以也可以从这些束缚电荷密度出发，直接求得相应电势和电场。我们现在可以证明两种描述是等价的。首先利用，

$$\nabla' \left(\frac{1}{R} \right) = \frac{\mathbf{R}}{R^3}. \quad (5.75)$$

注意，这里 $\nabla' \equiv \partial/\partial\mathbf{x}'$ ，是对源点坐标求偏导。这样极化势可以写成

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \mathbf{P} \cdot \nabla' \left(\frac{1}{R} \right) dV'. \quad (5.76)$$

利用分布积分，

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \nabla' \cdot \left(\frac{\mathbf{P}}{R} \right) dV' - \frac{1}{4\pi\epsilon_0} \int \frac{\nabla' \cdot \mathbf{P}}{R} dV', \quad (5.77)$$

$$= \frac{1}{4\pi\epsilon_0} \int \frac{\mathbf{P}}{R} \cdot d\mathbf{s}' - \frac{1}{4\pi\epsilon_0} \int \frac{\nabla' \cdot \mathbf{P}}{R} dV'. \quad (5.78)$$

在第二步里，运用了高斯定理将向量散度的体积分化成了向量的面积分。对比给定电荷 dq 的电势，

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \frac{dq}{R}. \quad (5.79)$$

我们显然证明了介质极化形成的电场等价于束缚面电荷 $\sigma = \mathbf{P} \cdot \mathbf{n}$ 和体电荷 $\rho = -\nabla \cdot \mathbf{P}$ 所形成的电场。

5.5.4 介质内的电场

刚刚我们证明了介质极化所形成的电场，利用了电场的多极展开里，电偶极矩所对应的电场。但我們需要注意一点的是，在我们进行多极展开时，有一个基本假设 $x' \ll x$ ，即电荷体系的空间尺度远远小于场点到电荷的距离。所以上式在介质外是没有问题的，这一条件可以被满足，因为电荷体系的空间尺度很小。但是在介质内，这一条件却不一定处处成立。想象一下，如果我们选的介质内的场点附近有一极化分子，且 $x' \sim x$ ，则似乎电场的多极展开的结论不能用了，且如果我们足够靠近相应分子里的电荷，则电场理论上可以非常大，而稍远点则会下降的非常快。这样电场随距离变化的图，其在介质内是忽大忽小的。然后实际中，我们并不太关心这种微观尺度的抖动非常厉害的电场，而更是关心平均后的电场。同时我们可以证明，如果我们将介质的电场理解为一个小区间内的平均电场，则上述结论依然成立的。这个平均用的体积元，一般是宏观小微观大。宏观小是其尺度相对于相应宏观物理量的变化尺度要小，而微观大则是相对于原子分子尺度，使得其包含足够多的电偶极子⁵。

⁵H. A. Lorentz 在 1902 年发表了题为 “The fundamental equations for electromagnetic phenomena in ponderable bodies deduced from the theory of electrons” 的文章。他在其中详细讨论了如何从微观麦克斯韦方程得到宏观麦克斯韦方程。

正因此，我们在前面几章基于点电荷所建立的电磁场方程也称为微观麦克斯韦方程 (microscopic Maxwell equations)，而介质相关的电磁场方程里的场是平均后的场，相应的场方程也称为宏观麦克斯韦方程 (macroscopic Maxwell equations)。

5.6 一般的边界条件

从高斯定理出发，如果我们有二个区域，I 和 II，则有

$$\oint \mathbf{D} \cdot d\mathbf{S} = \int \rho dV = Q = \sigma A, \quad (5.80)$$

这里 σ 是边界面上的电荷密度，而 A 是边界的面积。在边界上取一厚度 ε ，底面 A 横越两区域的小体积元，则方程 (5.80) 给出

$$(D_{\perp}^I - D_{\perp}^{II})A = \sigma A, \quad (5.81)$$

这里我们让 $\varepsilon \rightarrow 0$ 。所以有

$$\epsilon_I E_{\perp}^I - \epsilon_{II} E_{\perp}^{II} = \sigma \text{ or } (\epsilon_I \mathbf{E}^I - \epsilon_{II} \mathbf{E}^{II}) \cdot \mathbf{n} = \sigma, \quad (5.82)$$

这里 \mathbf{n} 是从区域 II 指向 I 的单位向量。

电场 \mathbf{E} 的切向分量是连续的，这点可以从如下的式子得到

$$\int \nabla \times \mathbf{E} \cdot d\mathbf{S} = \oint \mathbf{E} \cdot d\mathbf{l} = 0. \quad (5.83)$$

如果我们垂直于边界建立一个积分面，两边的高度为 ε ，宽为 l 。让 $\varepsilon \rightarrow 0$ ，则从方程 (5.83) 可以得到

$$E_{\parallel}^I - E_{\parallel}^{II} = 0. \quad (5.84)$$

我们可以将上述两个方程 (5.84) 与 (5.82) 合并在一起为

$$\epsilon_I \mathbf{E}^I - \epsilon_{II} \mathbf{E}^{II} = \sigma \mathbf{n}. \quad (5.85)$$

类似的考虑，电势在横越两个区域时是连续的，因为

$$\varphi^I - \varphi^{II} = - \int_a^b \mathbf{E} \cdot d\mathbf{l}, \quad (5.86)$$

如果 $a \rightarrow b$ ，右边为 0；所以 $\varphi^I = \varphi^{II}$ 。然后，电势的梯度不是连续的，因为

$$-\mathbf{n} \cdot \nabla \varphi = - \frac{\partial \varphi}{\partial n} = \mathbf{E} \cdot \mathbf{n}, \quad (5.87)$$

或者

$$\epsilon_I \frac{\partial \varphi^I}{\partial n} - \epsilon_{II} \frac{\partial \varphi^{II}}{\partial n} = -\sigma. \quad (5.88)$$

5.7 边值问题与唯一性定理

在静电学里一大常见问题是在我们感兴趣的区域内部没有电荷，这也是这一章里我们主要考虑的问题。在此情形下，电势 φ 满足

$$\nabla^2 \varphi = 0. \quad (5.89)$$

这个方程称为拉普拉斯方程, ∇^2 也称为拉普拉斯算符。拉氏方程的解有个特点就是其最大值和最小值不会在区域内部, 只能出现在边界上。要理解这一点, 我们可以将上述方程看成一个扩散问题的稳态解, 即

$$\frac{\partial \varphi}{\partial t} = \nabla^2 \varphi. \quad (5.90)$$

对于这样一个扩散过程, 显然在系统达到稳态时 ($\partial \varphi / \partial t \rightarrow 0$), 区域内部不可能有最大值或最小值, 否则系统还没有到达稳态, 会继续演化。

现在我们所要求解的问题显然是一个边值条件问题 (boundary value problem), 因此如果要定出这个问题的解, 我们还需要相应的边界条件。问题是什么样的边界条件是恰好够我们定出内部唯一的电场? 以一个简单的一维问题为例

$$\frac{\partial^2 \varphi}{\partial x^2} = 0, \quad x \in [a, b]. \quad (5.91)$$

这个方程的通解显然是 $\varphi = Ax + B$. 为了得到一个定出 φ , 我们需要 A 和 B ⁶。这样如下的几种 φ 的边界条件都是合适的:

- 给定 $\varphi(a)$ 与 $\varphi(b)$;
- 给定 $\partial \varphi(a) / \partial x$ and $\varphi(b)$;
- 给定 $\partial \varphi(b) / \partial x$ and $\varphi(a)$.

而另一方面, 如果我们知道的是 $\partial \varphi(b) / \partial x$ 与 $\partial \varphi(a) / \partial x$, 我们不能同时定出 A 与 B , 而且如果这两个边界条件处理不当, 还可能给出不自洽的结果。或者如果我们仅仅知道 $\varphi(a)$, 显然也不能同时决定 A 和 B 。

一般的三维问题会比这个更复杂。我们先证明如下的第一种解的唯一性定理: “如果 φ 在边界上处处给定, 则 $\nabla^2 \varphi = 0$ 解是唯一确定的。”在静电学里有好几种唯一性定理, 其证明过程大同小异。接下来我们通过证明第一唯一性定理来展示这一证明过程。

证明 假设我们有两个解 φ_1 和 φ_2 , 那么 $\nabla^2 \varphi_1 = 0$ 且 $\nabla^2 \varphi_2 = 0$ 。所以现在要证明第一唯一性定理, 我们仅需证明 $\varphi_1 = \varphi_2$ 即可。

现定义 $\varphi_3 \equiv \varphi_1 - \varphi_2$, 则我们需要证明在区域内部, φ_3 处处为 0。显然, φ_3 也满足拉普拉斯方程 $\nabla^2 \varphi_3 = 0$ 。因为在边界上 $\varphi_1 = \varphi$ 且 $\varphi_2 = \varphi$, 所以在边界上 $\varphi_3 = 0$ 。简而言之, φ_3 所满足的方程是

$$\nabla^2 \varphi_3 = 0, \text{ 且 } \varphi_3 = 0 \text{ 在 } \partial V \text{ 上}. \quad (5.92)$$

利用拉普拉斯方程解的性质, 即其解在区域内部不可能有最大值或最小值, 这些极值只能出现在边界上。那么因为 φ_3 在边界上处处为 0, 意味着在区域 V 内部, φ_3 的最大值和最小值均为 0, 所以 φ_3 处处为 0, 即

$$\varphi_1 = \varphi_2. \quad (5.93)$$

这样第一唯一性定理的证明完成。

以上的唯一性定理有个有意思的推论。可以发现, 如果 φ 满足的方程如下

$$\nabla^2 \varphi = -\frac{\rho}{\epsilon}, \quad (5.94)$$

即区域内的介质与电荷密度 ρ 给定, 且 φ 在边界上给定, 则方程 (5.94) 中 φ 的解是唯一确定的。显然引入给定电荷密度, 不会改变上述证明过程。

第一唯一性定理有时可以让我们靠“猜”或者构造一个等价问题来求解静电场的边值条件问题。对于适用上述定理的问题, 无论用什么方法, 只要我们可以找一个 φ , 在内部满足静电方程, 且边界上取值于给定条件相同, 则所找到的 φ 一定是原问题的唯一解。镜象电荷法即是这一定理的直接应用。

⁶但是如果我们的仅仅是希望定出电场 $E = d\varphi/dx$, 我们并不需要知道 B 是什么。

5.7.1 导体与第二唯一性原理

有时候在我们处理的静电问题中会出现导体或者我们并不知道边界上的电势，而仅仅是只到相关导体上的总电量，例如，导体 a 上有 Q_a 的电量，导体 b 上有 Q_b 的电量，但显然这些电荷也不会均匀分布在导体表面。第二唯一性原理便是处理这种情况。假定在导体之前的区域，电荷密度 ρ 的分布是给定的，则第二唯一性原理表述如下：

“区域 V 被导体包围或者无穷远且包含给定电荷分布 ρ ，如果区域内每个导体上的总电量是给定的，则区域内的电场 E 是唯一确定的。”

证明 该定理的证明与之前类似，但更复杂一点。假设我们有两个电场 E_1, E_2 均在区域内满足相应的静电方程，即

$$\nabla \cdot E_1 = \frac{\rho}{\epsilon} \text{ 且 } \nabla \cdot E_2 = \frac{\rho}{\epsilon}, \quad (5.95)$$

显然它们包围着各导体的面上也满足如下的积分形式方程

$$\oint_{i^{\text{th}} \text{ conductor}} E_1 \cdot ds = \frac{Q_i}{\epsilon} \text{ and } \oint_{i^{\text{th}} \text{ conductor}} E_2 \cdot ds = \frac{Q_i}{\epsilon}. \quad (5.96)$$

对于外边界有，

$$\oint_{\text{outer boundary}} E_1 \cdot ds = \frac{Q_{\text{tot}}}{\epsilon} \text{ and } \oint_{\text{outer boundary}} E_2 \cdot ds = \frac{Q_{\text{tot}}}{\epsilon}. \quad (5.97)$$

与之前类似，我们定义 $E_3 = E_1 - E_2$ ，且证明 E_3 在待求解区域处处为 0。

显然 E_3 在导体间区域满足 $\nabla \cdot E_3 = 0$ ，且沿包围各导体面的积分为

$$\oint E_3 \cdot ds = 0. \quad (5.98)$$

我们知道导体都是等势体，所以对每一个导体而言 φ_1, φ_2 与 φ_3 都是常数。这里我们利用

$$\nabla \cdot (\varphi_3 E_3) = \varphi_3 \nabla \cdot E_3 + E_3 \cdot \nabla \varphi_3 = -E_3^2, \quad (5.99)$$

在上式中，我们用了 $\nabla \cdot E_3 = 0$ 且 $\nabla \varphi_3 = -E_3$ 。沿整体区域积分，我们有

$$\int \nabla \cdot (\varphi_3 E_3) dV = \oint_S \varphi_3 E_3 \cdot ds = \varphi_3 \oint_S E_3 \cdot ds = 0, \quad (5.100)$$

这里我们利用了沿每个包裹着导体的积分面 φ_3 是一个常数。利用方程 (5.99)，式 (5.100) 也可以被写为

$$\int \nabla \cdot (\varphi_3 E_3) dV = - \int E_3^2 dV \leq 0. \quad (5.101)$$

将式 (5.101) 与 (5.100) 相结合，我们发现

$$\int E_3^2 dV = 0, \quad (5.102)$$

或者 $E_3 = 0$ 。因此 $E_1 = E_2$ 。

5.8 静像电荷法

5.8.1 经典静像电荷问题

镜像法是一个利用了第一唯一性原理来求解静电问题的一个巧妙方法，虽然不适用于一般的情况，但当它适用时，可以很快地让我们找到问题的解。一个经典的镜像问题如下：

一个点电荷 q 放在离一无限大接地平板距离为 d 的位置，求平板之上半空间的电势。为简化分析，我们建立一坐标系，平板位于 $z = 0$ 平面，而电荷在 $(0, 0, d)$ 处。我们需要求解的是 $z > 0$ 处的电势。

数学上，这个问题我们要求解电势 φ 所满足的泊松方程，且边界条件如下

- $\varphi = 0$ 当 $z = 0$,
- $\varphi = 0$ 当 $z \rightarrow \infty$.

所以第一唯一性定理适用, 该问题存在一唯一解。如果我们可以猜出一个 φ 满足相应的方程和边界条件, 则该猜出来的 φ 一定是原问题的唯一解。

当然, 我们没办法瞎猜。具体的, 我们构造另一个问题, 这个问题里的 $\tilde{\varphi}$ 是由 q 与另一个位于 $(0, 0, -d)$ 的电荷 $-q$ 产生的。显然, 空间中任一点的总电势为

$$\tilde{\varphi} = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{x^2 + y^2 + (z-d)^2}} - \frac{q}{\sqrt{x^2 + y^2 + (z+d)^2}} \right). \quad (5.103)$$

且该电势满足

- $\tilde{\varphi} = 0$ 当 $z = 0$.
- $\tilde{\varphi} = 0$ 当 $z \rightarrow \infty$.

从第一唯一性原理, $\varphi = \tilde{\varphi}$ 一定是原问题的解。在这个方法里 $-q$ 称为镜像电荷, 原因显而易见。从镜像法所依据的原理可以看出, 在使用镜像法的时候, 一定不能在求解区域引入新的电荷, 因为这会改变原电势所满足的方程, 不再适用第一唯一性原理的条件。

5.8.2 表面感应电荷

我们可以很容易的算出导体平板表面的感应电荷。导体内的电势为 0, 所以

$$\sigma = -\epsilon_0 \frac{\partial \varphi}{\partial n} = -\epsilon_0 \frac{\partial \varphi}{\partial z}. \quad (5.104)$$

在 $z = 0$, 并利用式 (5.103),

$$\sigma = -\frac{qd}{2\pi} \frac{1}{(x^2 + y^2 + d^2)^{3/2}}. \quad (5.105)$$

所以通过对 x, y 积分, 总的感应电荷为:

$$Q = \int \sigma dx dy = \int \sigma r dr d\phi = -q. \quad (5.106)$$

我们也可以计算 q 所受的力, 在原问题与镜像问题里肯定是一样的, 因为 q 电荷处的场是一样的。或者也可以计算总的场的能量, 这时原问题与镜像问题答案不同, 因为原问题里, 场仅局限于上半空间。

5.8.3 其它镜像问题

假设一点电荷 q 被放在离一半径为 R 的接地导体球中心距离为 a 的位置 ($a > R$)。请问球外的电势如何?

这个问题也可以用镜像法解决。所构造的镜像问题如下, 我们将一电量为 $q' = -(R/a)q$ 的电荷放置在球中心和 q 之间。则显然, 如果该电荷到中心的距离为 b 时

$$b = R^2/a. \quad (5.107)$$

这两电荷在球表面产生的电势为 0。在任一点, 其电势为

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{R_q} + \frac{q'}{R'_q} \right), \quad (5.108)$$

这里 R_q 是到 q 的距离, 而 R'_q 是到 q' 的距离。而 $\varphi(\mathbf{r})$ $r > R$ 即是原问题的解。

再一次的我们可以看到, 当镜像法适用的时候, 该方法非常简单。但一般情况下, 找到合适的镜像电荷非常困难或者不可能。这时我们需要更一般的方法来求解静电边值条件问题。

5.9 分离变量法

分离变量法的数学依据和过程在数理方法课程上有详细解释。这里我们仅展示利用该方法求解静电问题的两个例子，一个为直角坐标系，一个为球坐标系。

5.9.1 直角坐标系

例题 5.1 一个无穷长的矩形接地金属管边长为 a 和 b 。但在, $x = 0$, 的一端其电势维持在一个给定电势 $\varphi_0(y, z)$ 。求管内的电势。

解 我们要求解的方程为

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0, \quad (5.109)$$

而边界条件是

$$\varphi = 0 \text{ at } y = 0, y = a, z = 0, z = b; \quad (5.110)$$

$$\varphi = 0 \text{ at } x \rightarrow \infty; \quad (5.111)$$

$$\varphi = \varphi_0(y, z) \text{ 当 } x = 0. \quad (5.112)$$

利用分离变量法，我们让 $\varphi(x, y, z) = X(x)Y(y)Z(z)$ 。代入式 (5.109), 我们有

$$\frac{1}{X} \frac{\partial^2 X}{\partial x^2} + \frac{1}{Y} \frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = 0. \quad (5.113)$$

所以

$$\frac{1}{X} \frac{\partial^2 X}{\partial x^2} = C_1, \quad (5.114)$$

$$\frac{1}{Y} \frac{\partial^2 Y}{\partial y^2} = C_2, \quad (5.115)$$

$$\frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = C_3, \quad (5.116)$$

这里 $C_1 + C_2 + C_3 = 0$ 。

考虑到式 (5.110)-(5.112) 给出的边界条件，我们让 $C_2 = k^2$, $C_3 = l^2$, 则 $C_1 = -(k^2 + l^2)$ 。代入式 (5.114) - (5.116), 我们有

$$X = Ae^{\sqrt{k^2+l^2}x} + Be^{-\sqrt{k^2+l^2}x}, \quad (5.117)$$

$$Y = C \sin ky + D \cos ky, \quad (5.118)$$

$$Z = E \sin lz + F \cos lz. \quad (5.119)$$

从方程 (5.111), 显然 $A = 0$ 。在 $y = 0, Y = 0$, 则必然 $D = 0$ 。在 $y = a, Y = 0$, 我们有 $k = n\pi/a$ 。类似的 $F = 0$ 且 $l = m\pi/b$ 。所以最终 φ 的解为

$$\varphi = \sum_{n,m} C_{n,m} e^{-\pi \sqrt{(n/a)^2 + (m/b)^2} x} \sin(n\pi y/a) \sin(m\pi z/b). \quad (5.120)$$

剩下的常数 C 可以从边界条件: $\varphi(x=0) = \varphi_0(y, z)$ 得到, 即

$$C_{n,m} = \frac{4}{ab} \int \varphi_0(y, z) \sin\left(\frac{n\pi y}{a}\right) \sin\left(\frac{m\pi z}{b}\right) dy dz. \quad (5.121)$$

如果 $\varphi_0(y, z)$ 是一个常数 φ_0 , 则

$$C_{n,m} = \begin{cases} 0, & \text{if } n \text{ or } m \text{ is even,} \\ 16\varphi_0/\pi^2 mn, & \text{if } n \text{ and } m \text{ are odd.} \end{cases} \quad (5.122)$$

5.9.2 球坐标系

有的静电问题有天然的球对称性，比如涉及到球形物体。这种情况下，球坐标系成为一种自然而然的选择。

在球坐标系 (r, θ, ϕ) 中，拉普拉斯方程为

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \varphi}{\partial \phi^2} = 0. \quad (5.123)$$

在本课中，我们只处理对方位角对称的问题，即 $\varphi = \varphi(r, \theta)$ 与方位角 ϕ 无关。所以方程 (5.123) 变为

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial^2 \varphi}{\partial \theta^2} \right) = 0. \quad (5.124)$$

利用分离变量法，我们让 $\varphi = R(r)\Theta(\theta)$ ，则式 (5.124) 变成

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = 0. \quad (5.125)$$

第一第二项等于一个符号相反的常数。为方便计，在球坐标中，我们写成

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = l(l+1)R, \quad (5.126)$$

$$\frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = -l(l+1) \sin \theta \Theta. \quad (5.127)$$

数理方程的课程告诉我们上式的解为

$$R(r) = Ar^l + \frac{B}{r^{l+1}}, \quad (5.128)$$

$$\Theta(\theta) = P_l(\cos \theta). \quad (5.129)$$

这里 $P_l(\cos \theta)$ 为自变量为 $\cos \theta$ 的勒让得 (Legendre) 多项式。对于 l 比较小的几个勒让得多项式为

$$P_0(x) = 1, \quad (5.130)$$

$$P_1(x) = x, \quad (5.131)$$

$$P_2(x) = (3x^2 - 1)/2, \quad (5.132)$$

$$P_3(x) = (5x^3 - 3x)/2, \quad (5.133)$$

$$P_4(x) = (35x^4 - 30x^2 + 3)/8. \quad (5.134)$$

注意到 $P_l(x)$ 是关于 x 的 l 阶多项式；且当 l 为偶（奇）数时，只包含偶（奇）数次方项。勒让得多项式是正交的，即，

$$\begin{aligned} \int_{-1}^1 P_l(x) P_{l'}(x) dx &= \int_0^\pi P_l(\cos \theta) P_{l'}(\cos \theta) \sin \theta d\theta \\ &= \begin{cases} 0, & \text{if } l' \neq l, \\ \frac{2}{2l+1}, & \text{if } l' = l. \end{cases} \end{aligned} \quad (5.135)$$

显然在分离变量法中，我们需要用到这些性质。

最后球坐标系中拉普拉斯方程 (5.124) 在方位角对称下的通解为

$$\varphi(r, \theta) = \sum_{l=0}^{\infty} \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos \theta). \quad (5.136)$$

例题 5.2 一个半径为 R 的空心球，其表面电势为 $\varphi_0(\theta)$ 。求空心球内部的电势。

解 我们的一个边界条件是 $r = 0$ 处电势有限大, 所以通解中 $B_l = 0$, 即

$$\varphi(r, \theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta). \quad (5.137)$$

在 $r = R$,

$$\varphi(R, \theta) = \sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) = \varphi_0(\theta). \quad (5.138)$$

利用式 (5.135), 我们有

$$A_l = \frac{2l+1}{2R^l} \int_0^\pi \varphi_0(\theta) P_l(\cos \theta) \sin \theta d\theta. \quad (5.139)$$

代入通解, 即得球内电势。

例题 5.3 一个不带电半径为 R 的导体球放入一均匀外场 $\mathbf{E} = E_0 \mathbf{e}_z$ 中。求导体球外的电场。

解 让我们先确定边界条件是什么。首先, 导体球本身是一个等势体, 设其电势为 0。因为这里外部电场为一均匀场, 所以在无穷远处的电势对应的场为 \mathbf{E}_0 , 所以当 z 很大时, $\varphi = -E_0 z + C$ 。电场关于 $z = 0$ 的平面应该是对称的, 所以电势在 $z = 0$ 的平面应该是不变的, 这样 $C = 0$ 。最后, 边界条件总结如下

$$\varphi = 0, \quad \text{如果 } r = R, \quad (5.140)$$

$$\varphi = -E_0 r \cos \theta, \quad \text{如果 } r \gg R. \quad (5.141)$$

条件 (5.140) 导致

$$A_l R^l + \frac{B_l}{R^{l+1}} = 0, \quad (5.142)$$

或者

$$B_l = -A_l R^{2l+1}. \quad (5.143)$$

注意本问题在无穷远处的边界条件导致我们不能得到所有的 $A_l = 0$ 。方程 (5.136) 变为

$$\varphi(r, \theta) = \sum_{l=0}^{\infty} A_l \left(r^l - \frac{R^{2l+1}}{r^{l+1}} \right) P_l(\cos \theta). \quad (5.144)$$

对于 $r \gg R$, 括号里的第二项相对第一项可以忽略不计; 所以边界条件 (5.141) 给出

$$\varphi(r, \theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta) = -E_0 r \cos \theta. \quad (5.145)$$

利用 $P_1(x) = x$ 或 $P_1(\cos \theta) = \cos \theta$, 我们可以立刻得到

$$A_1 = -E_0, \quad (5.146)$$

且其它所有 A_l 都是 0。因此, 球外的电势 φ 为

$$\varphi(r, \theta) = -E_0 \left(r - \frac{R^3}{r^2} \right) \cos \theta. \quad (5.147)$$

显然, 第一项来源于外场, 而第二项来自导体球表面的感应电荷。这些感应电荷的电荷密度可以直接算出,

$$\sigma(\theta) = -\epsilon_0 \left. \frac{\partial \varphi}{\partial r} \right|_R = 3\epsilon_0 E_0 \cos \theta. \quad (5.148)$$

显然 σ 在北半球 ($z > 0$) 是正的, 而在南半球 ($z < 0$) 是负的。

例题 5.4 一半径为 R 的球壳上携带密度为 $\sigma_0(\theta)$ 的面电荷。求球内外的电势。

解 这个问题与之前的不同之处在于有两个求解区域。在这种情况下，我们可以分别求出各自区域的解。这些解一般含有待定常数。然后我们再通过连接两区域的连接条件定出这些常数。在本问题中，连接边界位于 $r = R$ 的球面。

对 $r \leq R$,

$$\varphi = \sum_l A_l r^l P_l(\cos \theta). \quad (5.149)$$

而对 $r \geq R$,

$$\varphi = \sum_l \frac{B_l}{r^{l+1}} P_l(\cos \theta). \quad (5.150)$$

电势在 $r = R$ 处连续，因此

$$A_l R^l = \frac{B_l}{R^{l+1}}, \quad (5.151)$$

或者

$$B_l = A_l R^{2l+1}. \quad (5.152)$$

利用式 (5.88),

$$\epsilon_0 \left. \frac{\partial \varphi_{\text{out}}}{\partial r} \right|_{r=R} - \epsilon_0 \left. \frac{\partial \varphi_{\text{in}}}{\partial r} \right|_{r=R} = -\sigma_0(\theta). \quad (5.153)$$

或者

$$\epsilon_0 \sum_l (2l+1) A_l R^{l-1} P_l(\cos \theta) = \sigma_0(\theta). \quad (5.154)$$

利用式 (5.135),

$$A_l = \frac{1}{2\epsilon_0 R^{l-1}} \int_0^\pi \sigma_0(\theta) P_l(\cos \theta) \sin \theta d\theta. \quad (5.155)$$

这样我们就得到了球内外的电势。

5.9.3 含有线性介质的边界条件问题

对线性介质而言，边界条件式 5.80 为

$$(\epsilon^{\text{I}} \mathbf{E}^{\text{I}} - \epsilon^{\text{II}} \mathbf{E}^{\text{II}}) \cdot \mathbf{n} = \sigma_f, \quad (5.156)$$

这里 \mathbf{n} 是从区域 II 指向区域 I 的法向矢量，或者用电势

$$\epsilon^{\text{I}} \frac{\partial \varphi^{\text{I}}}{\partial n} - \epsilon^{\text{II}} \frac{\partial \varphi^{\text{II}}}{\partial n} = -\sigma_f. \quad (5.157)$$

电势本身依然是连续的，即

$$\varphi^{\text{II}} = \varphi^{\text{I}}. \quad (5.158)$$

在均匀各向同性的线性介质内部，我们很容易得到一个很有用的关系，即束缚电荷密度 ρ_b 正比于自由电荷密度 ρ_f ,

$$\rho_b = -\nabla \cdot \mathbf{P} = -\nabla \cdot \mathbf{D} \left(\frac{\epsilon - \epsilon_0}{\epsilon} \right) = - \left(\frac{\epsilon - \epsilon_0}{\epsilon} \right) \rho_f. \quad (5.159)$$

这里用了 $\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E} = (1 - \epsilon_0/\epsilon) \mathbf{D}$ 。

例题 5.5 一均匀线性介质材料做成的球被放入均匀外场 \mathbf{E}_0 中。求球内的电场。

解 因为球是均匀线性极化的，所以球内 $\rho_b = \rho_f = 0$ 。因此，球内外电势都满足

$$\nabla^2 \varphi = 0. \quad (5.160)$$

而边界条件为

$$\varphi_{\text{in}} = \varphi_{\text{out}}, \quad \varphi \text{ 连续} \quad (5.161)$$

$$\epsilon \frac{\partial \varphi_{\text{in}}}{\partial r} = \epsilon_0 \frac{\partial \varphi_{\text{out}}}{\partial r}, \quad \sigma_f = 0 \text{ 球表面}, \quad (5.162)$$

$$\varphi_{\text{out}} = -E_0 r \cos \theta, \quad \text{对 } r \gg R \text{ 而言}. \quad (5.163)$$

综合边界条件，在球内电势为

$$\varphi = \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta); \quad (5.164)$$

而在球外，

$$\varphi = -E_0 r \cos \theta + \sum_{l=0}^{\infty} \frac{B_l}{r^{l+1}} P_l(\cos \theta). \quad (5.165)$$

边界条件式 (5.161) 要求

$$\sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) = -E_0 R \cos \theta + \sum_{l=0}^{\infty} \frac{B_l}{R^{l+1}} P_l(\cos \theta), \quad (5.166)$$

所以

$$A_1 R = -E_0 R + \frac{B_1}{R^2}, \quad (5.167)$$

$$A_l R^l = \frac{B_l}{R^{l+1}}, \quad \text{对 } l \neq 1. \quad (5.168)$$

条件 (5.162) 给出

$$\epsilon \sum_{l=0}^{\infty} l A_l R^{l-1} P_l(\cos \theta) = -\epsilon_0 \mathbf{E} \cos \theta - \epsilon_0 \sum_{l=0}^{\infty} \frac{(l+1) B_l}{R^{l+2}} P_l(\cos \theta), \quad (5.169)$$

因此

$$\epsilon l A_l R^{l-1} = -\frac{\epsilon_0 (l+1) B_l}{R^{l+2}}, \quad \text{for } l \neq 1, \quad (5.170)$$

$$\epsilon A_1 = -\epsilon_0 E_0 - \frac{2\epsilon_0 B_1}{R^3}. \quad (5.171)$$

方程 (5.168) 和 (5.170) 给出

$$A_l = B_l = 0, \quad \text{for } l \neq 1, \quad (5.172)$$

而方程 (5.171) 与 (5.167) 可以解出

$$A_1 = -\frac{3\epsilon_0}{\epsilon + 2\epsilon_0} E_0, \quad (5.173)$$

$$B_1 = \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} R^3 E_0. \quad (5.174)$$

因此

$$\varphi_{\text{in}} = -\frac{3\epsilon_0}{\epsilon + 2\epsilon_0} E_0 r \cos \theta = -\frac{3\epsilon_0}{\epsilon + 2\epsilon_0} E_0 z, \quad (5.175)$$

且球内电场为

$$\mathbf{E} = \frac{3\epsilon_0}{\epsilon + 2\epsilon_0} \mathbf{E}_0. \quad (5.176)$$

可见介质球会部分屏蔽电场，与导体类似，只是后者会完全屏蔽电场。

第六章 静磁学

6.1 静磁场

我们首先推导静磁场的场方程。从麦克斯韦方程出发，在静场条件下，

$$\nabla \cdot \mathbf{B} = 0, \quad (6.1)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}. \quad (6.2)$$

利用矢势 \mathbf{A} , 磁场 \mathbf{B} 可以表示为

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (6.3)$$

将 (6.3) 代入式 (6.2), 我们得到 \mathbf{A} 满足

$$\nabla \times (\nabla \times \mathbf{A}) = \mu_0 \mathbf{j} \quad (6.4)$$

注意到 $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, 并选择库伦规范

$$\nabla \cdot \mathbf{A} = 0, \quad (6.5)$$

这样 \mathbf{A} 的方程变为

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j}. \quad (6.6)$$

该方程与泊松方程完全类似，

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0}. \quad (6.7)$$

所以我们可以直接利用静电学里的结论来直接得到方程 (6.6) 里 \mathbf{A} 的解。对点电荷而言，

$$\phi = \frac{1}{4\pi\epsilon_0} \frac{q}{|\mathbf{x} - \mathbf{x}'|}, \quad (6.8)$$

泊松方程 (6.7) 的一般解为

$$\phi = \frac{1}{4\pi\epsilon_0} \int \frac{\rho}{|\mathbf{x} - \mathbf{x}'|} dV', \quad \text{with } dV' = d^3x' \quad (6.9)$$

这里 \mathbf{x} 是场点的坐标而 \mathbf{x}' 是电荷的坐标。因此为了得到方程中 (6.6) \mathbf{A} 的解, 只需 $\rho/\epsilon_0 \rightarrow \mu_0 \mathbf{j}$, 所以

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}}{|\mathbf{x} - \mathbf{x}'|} dV'. \quad (6.10)$$

这便是静磁场中矢势 \mathbf{A} 在库伦规范条件下的解。

静磁场 \mathbf{B} 可以从 \mathbf{A} 利用 $\mathbf{B} = \nabla \times \mathbf{A}$ 得到，

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\mu_0}{4\pi} \int \nabla \times \frac{\mathbf{j}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} dV'. \quad (6.11)$$

注意在这晨我们将 ∇ 算符与积分交换了一下，因为 ∇ 算符是对 \mathbf{x} 求导，不是 \mathbf{x}' 。计算出导数，我们有

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int \nabla \times \left(\frac{\mathbf{j}}{R} \right) dV' \quad (6.12)$$

$$= \frac{\mu_0}{4\pi} \int \nabla \frac{1}{R} \times \mathbf{j} dV'. \quad (6.13)$$

或者说，由稳恒电流 \mathbf{j} 产生的磁场为

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j} \times \mathbf{R}}{R^3} dV'. \quad (6.14)$$

这便是毕奥-萨伐尔定律 (the Biot-Savart law)。如果是一个局地环电流，我们可以替换 $\mathbf{j}dV' \rightarrow I d\mathbf{l}$ ，这里 I 是电流。这样

$$\mathbf{A} = \frac{\mu_0}{4\pi} \oint \frac{I d\mathbf{l}}{R}, \quad (6.15)$$

$$\mathbf{B} = \frac{\mu_0}{4\pi} \oint \frac{I d\mathbf{l} \times \mathbf{R}}{R^3}. \quad (6.16)$$

6.2 磁场的多极展开

在这一节里，我们考虑一个局地电流体系在远处 (\mathbf{x}) 所产生的磁场。所使用的方法与静电场的多极展开类似。为方便讨论，将坐标原点建立在电流系统内部，这样 $r \gg r'$,

$$f(\mathbf{x} - \mathbf{x}') = f(\mathbf{x}) - \mathbf{x}' \cdot \nabla f(\mathbf{x}) + \dots \quad (6.17)$$

因此，利用 $r \equiv |\mathbf{x}|$,

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} \approx \frac{1}{r} - \mathbf{x}' \cdot \nabla \frac{1}{r} + \dots, \quad i = 1, 2, 3. \quad (6.18)$$

所有的展开原理上与静电部分类似。但是，因为矢势 \mathbf{A} 是一个矢量，相应的展开过程会稍微复杂一点。我们也仅仅考虑到磁偶极项为止。

在场点 \mathbf{x} 处的矢势 \mathbf{A} 的形式为

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} dV' \quad (6.19)$$

其多极展开中的 0 阶和 1 阶项 \mathbf{A} 分别为

$$\mathbf{A}^{(0)} = \frac{\mu_0}{4\pi r} \int \mathbf{j} dV' \quad (6.20)$$

$$\mathbf{A}^{(1)} = -\frac{\mu_0}{4\pi} \int \mathbf{j}(\mathbf{x}' \cdot \nabla \frac{1}{r}) dV' \quad (6.21)$$

$$= \frac{\mu_0}{4\pi} \frac{\mathbf{x}}{r^3} \cdot \int \mathbf{x}' \mathbf{j} dV'. \quad (6.22)$$

首先我们考虑 $\mathbf{A}^{(0)}$ ，磁单极项。从直觉上看， $\mathbf{A}^{(0)} = 0$ ，因为目前的理论和实验不支持存在类似于电单极的磁单极子。我们也可以直接证明这一点。利用

$$\int (f \mathbf{j} \cdot \nabla' g + g \mathbf{j} \cdot \nabla' f) dV' = 0. \quad (6.23)$$

这里 $\nabla' = \partial/\partial \mathbf{x}'$ 。为简化计，我们称该等式为 $f - g$ 等式。这个等式的证明留在作业里，仅需注意到 \mathbf{j} 是一个局地电流系，所以在边界上有 $\mathbf{j} \cdot d\mathbf{S} = 0$ 。让 $f = 1, g = \mathbf{x}' \cdot \mathbf{e}_i = x'_i$ ，这样 $\nabla' g = \mathbf{e}_i$ 。利用 $f - g$ 等式 (6.23)，

$$A_i^{(0)} = \int \mathbf{j} \cdot \mathbf{e}_i dV' = 0, \quad i = 1, 2, 3. \quad (6.24)$$

所以

$$\mathbf{A}^{(0)} = 0. \quad (6.25)$$

接下来我们计算 1 阶项，我们需要计算

$$\mathbf{x} \cdot \int \mathbf{x}' \mathbf{j} dV'. \quad (6.26)$$

让 $f = x'_i$ 且 $g = x'_l$ ，这样

$$\int (x'_i \mathbf{j} \cdot \mathbf{e}_l + x'_l \mathbf{j} \cdot \mathbf{e}_i) dV' = 0, \quad (6.27)$$

或者

$$\int (x'_i j_l + x'_l j_i) dV' = 0. \Rightarrow \int x'_l j_i dV' = - \int x'_i j_l dV' \quad (6.28)$$

这样 $A^{(1)}$ 相关项可以被写成

$$\mathbf{x} \cdot \int \mathbf{x}' j_i dV' = x_l \int x'_l j_i dV' \quad (6.29)$$

$$= x_l \int \left(\frac{x'_l j_i - x'_i j_l}{2} \right) dV' \quad (6.30)$$

$$= \frac{1}{2} x_l \int \epsilon_{kli} (\mathbf{x}' \times \mathbf{j})_k dV' \quad (6.31)$$

$$= -\frac{1}{2} \epsilon_{ilk} x_l \int (\mathbf{x}' \times \mathbf{j})_k dV' \quad (6.32)$$

$$= -\frac{1}{2} \left[\mathbf{x} \times \int (\mathbf{x}' \times \mathbf{j}) dV' \right]_i. \quad (6.33)$$

定义磁偶极矩为

$$\boxed{\mathbf{m} = \frac{1}{2} \int \mathbf{x}' \times \mathbf{j} dV'}, \quad (6.34)$$

则

$$\mathbf{A}^{(1)} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \mathbf{x}}{r^3}. \quad (6.35)$$

因为 $\mathbf{A}^{(0)}$ 总是为 0，一般情况下磁偶极项 $\mathbf{A}^{(1)}$ 占主导地位。对一小电流环， $\mathbf{j} dV' \rightarrow I d\mathbf{l}'$ ，

$$\boxed{\mathbf{m} = \frac{1}{2} \oint \mathbf{x}' \times I d\mathbf{l}'} \quad (6.36)$$

定义矢量面积 $\mathbf{S} \equiv \frac{1}{2} \oint \mathbf{x}' \times d\mathbf{l}'$ 。对一平面电流环， $\mathbf{m} = I\mathbf{S}$ ，这里 \mathbf{S} 的大小是环的面积，其方向由电流方向结合右手法则决定。

偶极磁场 \mathbf{B} 可以直接从 \mathbf{A} 中得到。

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\mu_0}{4\pi} \nabla \times \left(\frac{\mathbf{m} \times \mathbf{x}}{r^3} \right) \quad (6.37)$$

或者

$$\boxed{\mathbf{B} = \frac{\mu_0}{4\pi} \frac{3(\mathbf{n} \cdot \mathbf{m})\mathbf{n} - \mathbf{m}}{r^3}, \quad \text{with } \mathbf{n} \equiv \frac{\mathbf{x}}{r}} \quad (6.38)$$

该场的表达式类似于电偶极场，仅仅是 \mathbf{p} 由 \mathbf{m} 替换。

6.3 外场中小电流体系的受力

我们考虑一个小电流体系在给定外场 \mathbf{B} 中的受力。与之前类似，这里的外场是由其它的源所产生的。我们来计算小电流体系所受到的力与力矩。电流的受力为

$$\mathbf{F} = \int \mathbf{j} \times \mathbf{B}(\mathbf{x}') dV'. \quad (6.39)$$

对磁场进行多极展开, $\mathbf{B} = B_i \mathbf{e}_i$,

$$B_i(\mathbf{x}') = B_i(0) + \mathbf{x}' \cdot \nabla B_i(0) + \dots, \quad (6.40)$$

这里坐标系的源点在电流体系内部。

将 $\mathbf{B}(\mathbf{x}')$ 的展开式代入 \mathbf{F} , 则

$$F_i = \int [(\mathbf{j} \times \mathbf{B}(0))_i + (\mathbf{j} \times (\mathbf{x}' \cdot \nabla) \mathbf{B})_i + \dots] dV' \quad (6.41)$$

$$= F_i^{(0)} + F_i^{(1)} + \dots. \quad (6.42)$$

首先计算 $\mathbf{F}^{(0)}$, 利用之前的结论,

$$\mathbf{F}^{(0)} = \left(\int \mathbf{j} dV' \right) \times \mathbf{B}(0) = 0, \quad (6.43)$$

偶极项所感受的力为

$$\mathbf{F}^{(1)} = \int \mathbf{j} \times (\mathbf{x}' \cdot \nabla) \mathbf{B}(0) dV'. \quad (6.44)$$

其第 i 个分量为

$$F_i^{(1)} = \epsilon_{iln} \int j_l \mathbf{x}' \cdot \nabla B_n(0) dV' = \epsilon_{iln} \int j_l x'_k \nabla_k B_n(0) dV'. \quad (6.45)$$

之前推导 \mathbf{A} 的多极展开时, 我们利用了

$$A_l^{(1)} = \frac{\mu_0}{4\pi} \frac{x_k}{r^3} \int x'_k j_l dV' \Rightarrow A_l^{(1)} = \frac{\mu_0}{4\pi} \frac{(\mathbf{m} \times \mathbf{x})_l}{r^3} \quad (6.46)$$

我们可以得到

$$\int x'_k j_l x_k dV' = (\mathbf{m} \times \mathbf{x})_l \quad (6.47)$$

替换 $x_k \rightarrow \nabla_k B_n$, 我们有

$$\int x'_k j_l \nabla_k B_n(0) dV' = (\mathbf{m} \times \nabla)_l B_n(0), \quad (6.48)$$

因此

$$F_i^{(1)} = \epsilon_{iln} (\mathbf{m} \times \nabla)_l B_n(0), \quad (6.49)$$

或者

$$\mathbf{F}^{(1)} = (\mathbf{m} \times \nabla) \times \mathbf{B} \quad (6.50)$$

利用先中间后外边法则, $\nabla = \nabla_{\mathbf{B}}$, 且 ∇ 是对 \mathbf{x} 求导, 不是 \mathbf{x}' , 我们可以得到

$$(\mathbf{m} \times \nabla) \times \mathbf{B} = \nabla(\mathbf{m} \cdot \mathbf{B}) - \mathbf{m}(\nabla \cdot \mathbf{B}),$$

因为 $\nabla \cdot \mathbf{B} = 0$ ，所以有

$$\boxed{\mathbf{F}^{(1)} = \nabla(\mathbf{m} \cdot \mathbf{B})}. \quad (6.51)$$

该式可以被写成另外的形式，利用

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A}$$

且 $\nabla \equiv \partial/\partial \mathbf{x}$ ，我们有

$$\nabla(\mathbf{m} \cdot \mathbf{B}) = \mathbf{m} \times (\nabla \times \mathbf{B}) + (\mathbf{m} \cdot \nabla)\mathbf{B}; \quad (6.52)$$

在产生外场的电流体系区域之外， $\nabla \times \mathbf{B} = 0$ 。因此

$$\boxed{\mathbf{F}^{(1)} = (\mathbf{m} \cdot \nabla)\mathbf{B}}. \quad (6.53)$$

这是磁偶极子在外场中的受力，形式也与电偶极子的受力类似 ($\mathbf{p} \rightarrow \mathbf{m}$)。

小电流体系在外场中的势能为

$$U = -\mathbf{m} \cdot \mathbf{B}, \quad (6.54)$$

这一点可以从力与势能的关系得到，

$$\mathbf{F} = -\nabla U = \nabla(\mathbf{m} \cdot \mathbf{B}). \quad (6.55)$$

与电偶极子的势能相似， $U = \mathbf{p} \cdot \mathbf{E}$ 。

小电流体系所受的力矩为

$$\mathbf{K} = \int \mathbf{x}' \times \mathbf{f} dV' = \int \mathbf{x}' \times (\mathbf{j} \times \mathbf{B}) dV'. \quad (6.56)$$

这里最低阶项为

$$\mathbf{K} \approx \mathbf{K}^{(0)} = \int \mathbf{x}' \times [\mathbf{j} \times \mathbf{B}(0)] dV' \quad (6.57)$$

$$= \int [(\mathbf{x}' \cdot \mathbf{B})\mathbf{j} - \mathbf{B}(\mathbf{x}' \cdot \mathbf{j})] dV'. \quad (6.58)$$

右边第一项为

$$\int (\mathbf{x}' \cdot \mathbf{B})\mathbf{j} dV' = \int [\mathbf{B} \cdot \mathbf{x}']\mathbf{j} dV' = \mathbf{m} \times \mathbf{B}(0). \quad (6.59)$$

第二项可以从 $f \cdot g$ 等式得到为 0 ($f = g = r'$)。这样最低阶非零力矩为

$$\mathbf{K} = \mathbf{m} \times \mathbf{B}(0). \quad (6.60)$$

从 \mathbf{K} 和 \mathbf{m} 的表达式，我们可以得到一个有趣的现象。一个体系的角动量 \mathbf{L} 与力矩的关系为

$$\frac{d\mathbf{L}}{dt} = \mathbf{K} = \mathbf{m} \times \mathbf{B}_0, \quad (6.61)$$

这里 $\mathbf{B}_0 = \mathbf{B}(0)$ ，且 $\mathbf{L} = \sum \mathbf{x}' \times \mathbf{p} = \sum \mathbf{x}' \times m\mathbf{v}$ 。这里是对所有粒子求和。磁偶极矩为

$$\mathbf{m} = \frac{1}{2} \int \mathbf{x}' \times \mathbf{j} dV' = \frac{1}{2} \int \mathbf{x}' \times \rho \mathbf{v} dV' = \frac{1}{2} \sum \mathbf{x}' \times q\mathbf{v}. \quad (6.62)$$

我们看到，如果所有粒子的荷质比 q/m 相同，则

$$\mathbf{m} = \frac{1}{2} \sum (\mathbf{x}' \times m\mathbf{v}) \frac{q}{m} = \frac{q}{2m} \sum \mathbf{x}' \times m\mathbf{v} = \frac{q}{2m} \mathbf{L}. \quad (6.63)$$

这样角动量为

$$\boxed{\frac{d\mathbf{L}}{dt} = \frac{q}{2m} \mathbf{L} \times \mathbf{B}_0 = \mathbf{L} \times \boldsymbol{\Omega}_L} \quad (6.64)$$

这里 $\boldsymbol{\Omega}_L = (q/2m)\mathbf{B}_0$ 是拉莫进动 (Larmor precession) 的频率。

6.4 介质的磁化

6.4.1 抗磁, 顺磁, 与铁磁

微观上讲, 磁场都是电荷的运动所形式的。电子绕原子核运动, 且会自旋。我们可以将这些运动对应的小的电流环看成磁偶极子。一般情况下, 这些磁偶极子是随机排列的, 它们产生的磁场互相抵消, 宏观上不显示磁性。但是如果将这些物质放入一个外场中, 各小磁偶极子在外场作用下重新排列, 形成静的整体磁偶极矩。这样介质便被磁极化 (magnetically polarized), 或简称为磁化。

与电极化不同, 介质的磁化通常更为复杂。有几种不同的磁化。如果磁化后的磁矩平行与外场 \mathbf{B} , 则为顺磁 (paramagnets); 反之则为抗磁 (diamagnets)。对部分物质而言, 磁化的结果依赖于磁化的历史, 这种称为铁磁 (ferromagnets)。在本节中, 因铁磁性过于复杂, 我们仅仅讨论顺磁和抗磁形成的微观过程。注意的是, 严格来讲, 经典物理是无法解释磁化的, 需要的是量子力学相关过程。

6.4.2 顺磁

介质的顺磁性电子自旋形成的磁偶极子。当施加外场 \mathbf{B} 后, 每个磁偶极子感受到一个力矩, 该力矩会尽量将偶极子与外场平行从而降低势能。这样形成的整体磁矩便平行于外加磁场 \mathbf{B} , 从而介质显出顺磁性。这种机制初听之下容易让人觉得每种介质都该是顺磁的, 然后量子力学里的泡利不相容原理通常会导致相反自旋的电子成对出现, 而力矩作用在这样的一对电子上的净效果基本为 0。顺磁性通常出现在那些原子中有未配对电子的物质上。

6.4.3 抗磁

电子不仅自旋, 还会绕原子核转动。假设该运动为圆周, 其半径为 R , 则形成的电流为

$$I = -\frac{e}{T} = -\frac{ev}{2\pi R}. \quad (6.65)$$

将 z 方向定义为角速度方向, 则对应的磁偶极矩为

$$\mathbf{m} = I\mathbf{S} = -\frac{evR}{2}\mathbf{e}_z \quad (6.66)$$

当我们施加外场 \mathbf{B} 后, 该轨道磁偶极矩会感受到一个力矩 $\mathbf{m} \times \mathbf{B}$ 。但是改变轨道的方向远难过改变自旋的方向, 所以轨道磁偶极矩对顺磁的贡献比较小。尽管如此, 在外场作用下, 电子会加速或减速, 取决于 \mathbf{B} 的方向。为证明这一点, 在施加外场 \mathbf{B} 之前, 我们有

$$\frac{1}{4\pi\epsilon_0} \frac{e^2}{R^2} = m_e \frac{v^2}{R}. \quad (6.67)$$

当我们施加外场 \mathbf{B} 后, 有一个额外的向心力 $-e(\mathbf{v}/c) \times \mathbf{B}$, 且电子的新速度为 \bar{v} 。假设 $\mathbf{B} = B\mathbf{e}_z$, 新的运动方程为

$$\frac{1}{4\pi\epsilon_0} \frac{e^2}{R^2} + e\bar{v}B = m_e \frac{\bar{v}^2}{R}. \quad (6.68)$$

定义 $\Delta v = \bar{v} - v$ 且仅保留到 Δv 的一阶项, 式 (6.68) 与 (6.67) 给出

$$e\bar{v}B = m_e \frac{\bar{v}}{R} 2\Delta v, \quad (6.69)$$

或者

$$\Delta v = \frac{eRB}{2m_e}. \quad (6.70)$$

所以 $\mathbf{B} = B\mathbf{e}_z$ 施加后，电子速度增加。与此对应式 (6.66) 表明 \mathbf{m} 也会变化，

$$\Delta\mathbf{m} = -\frac{eR}{2}\Delta v\mathbf{e}_z = -\frac{e^2R^2}{4m_e}\mathbf{B}. \quad (6.71)$$

注意 $\Delta\mathbf{m}$ 与外场 \mathbf{B} 相反。因此当施加一个外场 \mathbf{B} 后，每一个绕核运动的电子都获得了一个额外的磁偶极矩，方向都与外场相反。这样整体上介质便获得了一个净的与外场相反的磁偶极矩，显示出抗磁性。这种抗磁性一般远比顺磁性弱，所以通常只在原子中电子完全配对的情况下才显现出来。

第七章 时变电磁场

在本章里，我们忽略源项（ ρ 和 \vec{j} ），但考虑随时间变化的场，即真空中传播的电磁波。同时也会介绍简单线性介质里的时变电磁场，以与真空中的电磁场直接对比，解释其中相似与不同之处。

7.1 真空中的电磁场方程

在真空中，电荷密度与电流密度均为 0，此时麦克斯韦方程简化为

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (7.1)$$

$$\nabla \cdot \vec{E} = 0, \quad (7.2)$$

$$\nabla \times \vec{B} = \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}, \quad (7.3)$$

$$\nabla \cdot \vec{B} = 0. \quad (7.4)$$

这一组方程可以化为关于 \vec{E} 和 \vec{B} 的波动方程。将 $\partial/\partial t$ 作用在 $\nabla \times \vec{E}$ 的方程上，

$$\nabla \times \frac{\partial \vec{E}}{\partial t} = \frac{\partial}{\partial t} \nabla \times \vec{E} = -\frac{\partial^2 \vec{B}}{\partial t^2}. \quad (7.5)$$

利用安培定律，我们可将该方程改写为

$$\nabla \times \frac{\partial \vec{E}}{\partial t} = \nabla \times \left(c^2 \nabla \times \vec{B} \right). \quad (7.6)$$

或

$$\nabla \times \left(\nabla \times \vec{B} \right) = -\frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2}. \quad (7.7)$$

利用向量运算公式

$$\nabla \times (\nabla \times \vec{B}) = \nabla(\nabla \cdot \vec{B}) - \nabla^2 \vec{B} = -\nabla^2 \vec{B}, \quad (7.8)$$

上一方程变成

$$-\nabla^2 \vec{B} = -\frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2}, \quad (7.9)$$

或者

$$\boxed{\frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0.} \quad (7.10)$$

类似的，我们可以得到电场 \vec{E} 满足

$$\boxed{\frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0.} \quad (7.11)$$

从电磁场的协变方程出发，我们可以直接得到电磁势 ϕ 和 \vec{A} 的波动方程：

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 0, \quad (7.12)$$

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = 0. \quad (7.13)$$

这一组方程也被称为达朗贝尔 (d'Alembert) 方程，且

$$\boxed{\square \equiv \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2,} \quad (7.14)$$

也被称为达朗贝尔算符。该算符也可以被理解为四维时空下的拉普拉斯算符。

7.2 单色平面电磁波

我们考虑一个形式如

$$f(x, t) = f_0 e^{i(kx - \omega t)}, \quad (7.15)$$

的波动。该波动只与一个空间坐标有关，所以是平面波；其只有一个频率，所以该波动是单色的¹。因此，该波动称为单色平面波。如果从数学求解波动方程的角度，单色平面波也会自然地出现在对时间和空间作傅立叶变换后。

我们注意到在上述表达式里，我们用了一个复数变量来表达一个波动物理量，这里是利用了 $e^{i\theta} = \cos\theta = i \sin\theta$ 。因为复数的一些性质，在波动分析和许多其它物理问题里这会非常方便。例如对上述平面波动，它的时空导数可以很容易写出

$$\frac{\partial f}{\partial t} = -i\omega f; \quad \frac{\partial f}{\partial x} = ikf. \quad (7.16)$$

对于更一般点的平面电磁波，

$$f = f_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}, \quad (7.17)$$

也有

$$\nabla f = i\vec{k}f. \quad (7.18)$$

如果用实数来表示波动，例如 $f = f_0 \cos(kx - \omega t)$ ，上述导数则为

$$\frac{\partial f}{\partial t} = \omega f_0 \sin(kx - \omega t); \quad \frac{\partial f}{\partial x} = -k f_0 \sin(kx - \omega t). \quad (7.19)$$

这在很多的分析里非常的不方便。

如果我们用一个复变量 f 去表示一个物理量，则该实际物理量可以取为该复变量的实部，即 $\text{Re}(f)$ 。那用复变量的前提条件是什么？物理量 f 所满足的方程应当是一个线性方程。假设 \mathcal{L} 是一个线性算符，则

$$\mathcal{L}\{f\} = \mathcal{L}\{\text{Re}(f) + i \text{Im}(f)\} = \mathcal{L}\{\text{Re}(f)\} + i \mathcal{L}\{\text{Im}(f)\} = 0, \quad (7.20)$$

意味着 $\mathcal{L}\{\text{Re}(f)\} = 0$ 且 $\mathcal{L}\{\text{Im}(f)\} = 0$ ，所以其实部自动满足相关的方程。麦克斯韦方程是线性的，所以我们可以用复数去表示 \vec{E} 与 \vec{B} 。

例如对于平面电磁波动量 f (标量) 和 \vec{A} (矢量)，我们有

$$\nabla \times \vec{A} = i\vec{k} \times \vec{A}, \quad \frac{\partial \vec{A}}{\partial t} = -i\omega \vec{A}, \quad \nabla f = i\vec{k}f \quad \text{and} \quad \frac{\partial f}{\partial t} = -i\omega f. \quad (7.21)$$

这样，相应的电场和磁场为

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} = -i\vec{k}\phi + i\omega \vec{A}, \quad (7.22)$$

$$\vec{B} = i\vec{k} \times \vec{A} = \vec{k} \times \vec{E}/\omega. \quad (7.23)$$

从这里我们可以看到 $\vec{E} \perp \vec{B}$ ，且 $\vec{B} \perp \vec{k}$ 。第二个式子也告诉我们电场和磁场是同相位振荡的。要获取物理的“ \vec{E} ”和“ \vec{B} ”，仅仅只需取它们的实部即可： $\text{Re}(\vec{E})$ 与 $\text{Re}(\vec{B})$ 。

那么场的能量该如何用复电场 \vec{E} 来表达？为了区分，我们用 $\hat{\vec{E}}$ 表达复电场，而用 \vec{E} 表达物理的电场，这样

$$\vec{E} = \text{Re}\hat{\vec{E}}. \quad (7.24)$$

¹这个名词显然由光的颜色和频率的关系而来

电场的能量密度是 $w = \epsilon_0 \vec{E} \cdot \vec{E} / 2$ 。但是

$$w = \text{Re} \left\{ \frac{\epsilon_0}{2} \hat{\vec{E}} \cdot \hat{\vec{E}} \right\}, \quad (7.25)$$

是否正确? 我们需要仔细的分析一下。假设 $\hat{\vec{E}} = \vec{E}_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}$, 且简化计, \vec{E}_0 is 是实数。²

$$\hat{\vec{E}} \cdot \hat{\vec{E}} = \vec{E}_0^2 e^{2i\varphi} \quad \text{with } \varphi \equiv \vec{k} \cdot \vec{x} - \omega t, \quad (7.26)$$

则

$$\text{Re} \left\{ \frac{\epsilon_0}{2} \hat{\vec{E}} \cdot \hat{\vec{E}} \right\} = \frac{\epsilon_0}{2} E_0^2 \cos 2\varphi. \quad (7.27)$$

然而, 场的能量密度用物理电场 $\vec{E} = \vec{E}_0 \cos \varphi$ 表达, 则是

$$w = \frac{\epsilon_0}{2} \vec{E}_0^2 \cos^2 \varphi \neq \text{Re} \left\{ \frac{\epsilon_0}{2} \hat{\vec{E}} \cdot \hat{\vec{E}} \right\}. \quad (7.28)$$

正确地用复数表达场的能量密度的应该是

$$w = \frac{\epsilon_0}{2} \vec{E} \cdot \vec{E} = \frac{\epsilon_0}{2} \frac{\hat{\vec{E}} + \hat{\vec{E}}^*}{2} \cdot \frac{\hat{\vec{E}} + \hat{\vec{E}}^*}{2}. \quad (7.29)$$

这个式子看起来啰嗦, 但后续会利用波动的特性将其简化。

7.3 波动的极化

波动的另一个重要性质是其极化特性。考虑一个平面单色电磁波, 其波矢 $\vec{k} = k\vec{e}_x$, 则

$$\vec{E} = \vec{E}_0 e^{i(kx - \omega t)}, \quad (7.30)$$

因为 $\vec{k} = k\vec{e}_x$, 所以 $\vec{E} = \vec{E}_y + \vec{E}_z$, 或者

$$\vec{E} = (\vec{E}_{0y} + \vec{E}_{0z}) e^{i(kx - \omega t)} = (\hat{E}_{0y} \vec{e}_y + \hat{E}_{0z} \vec{e}_z) e^{i(kx - \omega t)}. \quad (7.31)$$

注意这里振幅 \hat{E}_{0y} 与 \hat{E}_{0z} 一般都是复常数。我们讲波动极化的时候, 讲的是空间某个固定点观察的时候, 波动的电场矢量末端所画出的轨迹是什么样子的。从上式可以看出, 该矢量是两个互相垂直的振动所合成出来的。根据两个垂直方向振动的相位和振幅的相对关系, 总电场矢量的极化可以分为线性极化, 圆极化, 和最一般的椭圆极化。

我们从最简单的线性极化 (linear polarization) 开始。如果 \hat{E}_{0y} 与 \hat{E}_{0z} 有相同的相位 α , 即

$$\hat{E}_{0y} = E_{0y} e^{i\alpha} \quad \text{且} \quad \hat{E}_{0z} = E_{0z} e^{i\alpha}. \quad (7.32)$$

则

$$\vec{E} = (E_{0y} \vec{e}_y + E_{0z} \vec{e}_z) e^{i(kx - \omega t + \alpha)}, \quad (7.33)$$

我们可以看到电场矢量沿着一个固定的方向 \hat{n} 在振荡。该方向 \hat{n} 与 \vec{e}_y 的夹角为

$$\langle \hat{n}, \vec{e}_y \rangle = \tan^{-1} \left(\frac{E_{0z}}{E_{0y}} \right), \quad (7.34)$$

这个波动即是线性极化的波动。

²一般情况下, \vec{E}_0 是一个复常数, 有自己的常数复相位。

第二种特殊极化为圆极化 (circular polarization)。这时两个垂直的振动具有相同的振幅, $E_{0y} = E_{0z} = E_0$, 且其相位之间差 $\pi/2$, 即

$$\hat{E}_{0y} = E_0 e^{i\alpha} \quad \text{and} \quad \hat{E}_{0z} = E_0 e^{i(\alpha \pm \pi/2)}, \quad (7.35)$$

这样

$$\vec{E} = E_0 (\vec{e}_y \pm i\vec{e}_z) e^{i(kx - \omega t + \alpha)}. \quad (7.36)$$

实际的物理电场为

$$E_y = E_0 \cos(kx - \omega t + \alpha), \quad (7.37)$$

$$E_z = \mp E_0 \sin(kx - \omega t + \alpha). \quad (7.38)$$

在一个固定点 $z = z_0$, 电场矢量以频率 ω , 振幅 E_0 画出一个圆。这个波动称为圆极化波动。注意到 $\hat{E}_{0z}/\hat{E}_{0y} = i$ 或者 $-i$, 电场的转动方向 (direction of rotation, DOR) 在这两种情况下是不同的。当一个观察者是面向波来的方向时 (或者说面向波源), 则

1. $\hat{E}_{0z}/\hat{E}_{0y} = i$, DOR 是逆时针的
2. $\hat{E}_{0z}/\hat{E}_{0y} = -i$, DOR 是顺时针的。

$$\vec{E} = E_0 (\vec{e}_y + i\vec{e}_z) e^{i(kx - \omega t + \alpha)}$$

圆极化的波按旋转方向还被称为左旋或右旋, 即波动的矢量是按照左手四指弯曲的方向还是右手四指弯曲的方向, 但是这也显然取决于大拇指的指向。在不同的物理领域里, 大拇指的指向定义是不一样的。

方式一: 从源的角度来定义

将大拇指指离源, 与 \vec{k} 相同, 利用四指弯曲的方向定义左右旋, 则

1. $\hat{E}_{0z}/\hat{E}_{0y} = i$, 右旋极化
2. $\hat{E}_{0z}/\hat{E}_{0y} = -i$, 左旋极化

这种定义方式在工程类, 量子物理和射电天文里用得比较多。

方式二: 从接收者的角度

将大拇指指向源, 与 \vec{k} 方向相反, 利用四指弯曲的方向定义左右旋, 则

1. $\hat{E}_{0z}/\hat{E}_{0y} = i$, 左旋极化
2. $\hat{E}_{0z}/\hat{E}_{0y} = -i$, 右旋极化

这种定义方式一般在光学里用得比较多。

方式三: 利用背景磁场的方向

不管 \vec{k} 的方向如何, 将大拇指指向 \vec{B}_0 方向, 利用四指弯曲的方向定义左右旋, 则

1. $\hat{E}_{0z}/\hat{E}_{0y} = i$, 左旋或右旋极化
2. $\hat{E}_{0z}/\hat{E}_{0y} = -i$, 左旋或右旋极化

这种定义方式在等离子体物理里使用。因为右旋波与电子回旋运动方向一致, 而左旋波与离子回旋运动方向一致。

一般来说 \hat{E}_{0y} 和 \hat{E}_{0z} 有不同于 $\pi/2$ 的相位和振幅, 此时波动是椭圆极化 (elliptical polarization) 的。我们接下来证明这一点。利用

$$\vec{E} = \vec{E}_0 e^{i(kx - \omega t)}, \quad (7.39)$$

我们令 $\vec{E}_0 = \vec{b} e^{i\alpha}$ 。注意这里 \vec{b} 是一个选定的复常数, 其实部和虚部为

$$\vec{b} = \vec{b}_1 + i\vec{b}_2, \quad (7.40)$$

这里 \vec{b}_1 与 \vec{b}_2 均为实向量。我们选择一个 \vec{b} 使得 $\vec{b} \cdot \vec{b} = \vec{b}_1^2 + \vec{b}_2^2 + 2i\vec{b}_1 \cdot \vec{b}_2$ 为实数, 这样 $\vec{b}_1 \cdot \vec{b}_2$ 一定是 0, 即 $\vec{b}_1 \perp \vec{b}_2$ 。为简化讨论, 我们令 $\vec{b}_1 = b_1 \vec{e}_1$ 且 $b_1 > 0$ 。

如果定义另一个单位矢量 \vec{e}_2 定义为 $\vec{e}_x \times \vec{e}_1 = \vec{e}_2$, 则

$$\vec{b}_2 = \pm b_2 \vec{e}_2, \quad b_2 > 0. \quad (7.41)$$

现在我们可以利用 \vec{e}_1 和 \vec{e}_2 来表达 \vec{E} 的分量:

$$\vec{E} = \vec{b} e^{i(kx - \omega t + \alpha)} = (b_1 \vec{e}_1 \pm ib_2 \vec{e}_2) e^{i(kx - \omega t + \alpha)}. \quad (7.42)$$

$$\hat{E}_1 = b_1 e^{i(kx - \omega t + \alpha)} \vec{e}_1, \quad (7.43)$$

$$\hat{E}_2 = \pm b_2 e^{i(kx - \omega t + \alpha + \pi/2)} \vec{e}_2. \quad (7.44)$$

因此, 在 \vec{e}_1 和 \vec{e}_2 方向上,

$$E_1 = b_1 \cos(kx - \omega t + \alpha), \quad (7.45)$$

$$E_2 = \mp b_2 \sin(kx - \omega t + \alpha). \quad (7.46)$$

显而易见, E_1 与 E_2 满足

$$\frac{E_1^2}{b_1^2} + \frac{E_2^2}{b_2^2} = 1. \quad (7.47)$$

在空间中的任一个固定点, 向量 \vec{E} 的末端轨迹为由上式所定义的一个椭圆, 其长短轴为 $\max(b_1, b_2)$ 和 $\min(b_1, b_2)$ 。这个波即是椭圆极化的。作为椭圆极化的特殊情况, 如果 $b_1 = b_2$, 则波动为圆极化; 而如果其中一个为 0, 则 $E_1 = 0$ 或 $E_2 = 0$, 此时波动是线偏振的。

7.4 线性介质中的电磁波

接下来我们考虑没有自由电荷和电流的线性介质中的电磁波。在这种情况下, 麦克斯韦方程为

$$\nabla \cdot \vec{E} = 0, \quad (7.48)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (7.49)$$

$$\nabla \cdot \vec{B} = 0, \quad (7.50)$$

$$\nabla \times \vec{B} = \mu\epsilon \frac{\partial \vec{E}}{\partial t}. \quad (7.51)$$

而边界条件与之前类似:

$$\epsilon_1 E_1^\perp = \epsilon_2 E_2^\perp, \quad (7.52)$$

$$\vec{E}_1^\parallel = \vec{E}_2^\parallel, \quad (7.53)$$

$$B_1^\perp = B_2^\perp, \quad (7.54)$$

$$\frac{\vec{B}_1^\parallel}{\mu_1} = \frac{\vec{B}_2^\parallel}{\mu_2}. \quad (7.55)$$

显然 \vec{E} 与 \vec{B} 满足如下的波动方程

$$\frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = 0, \quad (7.56)$$

这里 $v = 1/\sqrt{\mu\epsilon}$ 为光在介质中的相速度。该方程的平面波解与之前类似为

$$u(x, t) = f(x - vt) + g(x + vt). \quad (7.57)$$

或者利用单色平面波，

$$u = u_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}, \quad (7.58)$$

这里

$$k = \frac{\omega}{v} = \omega \sqrt{\mu\epsilon}, \quad (7.59)$$

是波数。我们也定义折射率为

$$n = \frac{c}{v} = \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}} = \sqrt{\epsilon_r\mu_r}. \quad (7.60)$$

对于大部分介质而言, $\mu_r \approx 1$, 因此, $n \approx \sqrt{\epsilon_r}$ 。因此, 如果我们知道了介质的电学性质, 便得到了光在该介质中的折射率。

利用类似于 (7.58) 的单色平面波解, 我们可以从麦克斯韦方程中得到

$$\vec{k} \cdot \vec{E} = 0 \text{ and } \vec{k} \cdot \vec{B} = 0, \quad (7.61)$$

因此, 介质中的电磁波也是横波。同时, 法拉弟方程给出

$$i\vec{k} \times \vec{E} = -(-i\omega)\vec{B} \quad (7.62)$$

或

$$\vec{n} \times \frac{\vec{E}}{c} = \vec{B}, \quad (7.63)$$

这里 $\vec{n} = n\hat{k} = \vec{k}c/w$ 。方程 (7.63) 显示介质中的电磁波满足 $B = nE/c$ 。

然而, 尽管上述介质中的麦克斯韦方程与真空中的类似, 但是其中的电磁波性质却可能比上述简单图象更为复杂。原因在于介质中的电荷可以在电磁波的电磁场作用下运动, 从而与电磁波发生能量交换。如果是介质吸收了电磁波的能量, 则会造成波动的能量衰减。所以在更为一般的情况下, 上述方程里的波矢 k 或频率 ω 不一定为实数。为简化讨论, 我们设 ω 为实数, 而 \vec{k} 为复数, 这里可以很方便的表达波在传播过程中的能量变化。例如, $\vec{k} = \vec{k}_r + i\vec{k}_i$, 则此时单色平面波的解为

$$u = u_0 e^{i(\vec{k}_r \cdot \vec{x} - \omega t)} e^{-\vec{k}_i \cdot \vec{x}}. \quad (7.64)$$

这意味着波动在沿 $+x$ 传波时可以衰减, 如果 $k_i > 0$, 或增长, 如果 $k_i < 0$ 。

7.5 介质中的电磁波

7.5.1 介绍

在这一节里, 我们将利用电动力学知识里电磁场在边界处的行为来解释电磁波(光)在介质分界处的折射(reflection)与反射(refraction), 主要利用的是方程 (7.52)-(7.55)。我们将解释以下光的现象:

- 入射, 折射与反射波的波矢位于一个平面内 (也称为入射平面)
- 入射角等于反射角

- 斯奈尔定理 (Snell's law), 即

$$\frac{\sin \theta_i}{\sin \theta_r} = \frac{n'}{n}, \quad (7.65)$$

这里 θ_i 与 θ_r 分别是入射与折射角, 而 n 与 n' 则是光在入射与折射介质里的折射率。

我们还将分析以下过程

- 反射与折射光的强度随入射角的变化,
- 光的相位变化和极化性质。

为方便讨论, 我们将建立如下坐标系: z 轴垂直于介质分界面, 且分界面位于 $z = 0$ 处。对 $z > 0$ 而言, 线性介质的磁常数与电常数为 μ' 与 ϵ' ; 对 $z < 0$ 的介质, 相应常数为 μ 与 ϵ 。显然, 相应介质的折射率为 $n' = \sqrt{\mu'\epsilon'/\mu_0\epsilon_0}$ ($z > 0$) 与 $n = \sqrt{\mu\epsilon/\mu_0\epsilon_0}$ ($z < 0$)。一个波矢为 \vec{k} , 频率为 ω 的平面波 ω 从 $z < 0$ 入射, 在 $z = 0$ 处发生反射, 折射光进入 $z > 0$ 。折射与反射波的波矢分别为 \vec{k}' 与 \vec{k}'' 。我们定义 θ_i 为入射角 (i-incident), θ_r 为反射角, 而 θ_t 为折射角。

在这样一个系统里有三支波, 而显然电磁场在边界处所要满足的边界条件决定了这三支波之间是有联系的:

- 入射波

$$\vec{E} = \vec{E}_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}, \quad (7.66)$$

$$\vec{B} = \vec{n} \times \frac{\vec{E}}{c} = \sqrt{\mu\epsilon} \frac{\vec{k}}{k} \times \vec{E}. \quad (7.67)$$

- 折射波

$$\vec{E}' = \vec{E}'_0 e^{i(\vec{k}' \cdot \vec{x} - \omega t)}, \quad (7.68)$$

$$\vec{B}' = \vec{n}' \times \frac{\vec{E}'}{c} = \sqrt{\mu'\epsilon'} \frac{\vec{k}'}{k'} \times \vec{E}'. \quad (7.69)$$

- 反射波

$$\vec{E}'' = \vec{E}''_0 e^{i(\vec{k}'' \cdot \vec{x} - \omega t)}, \quad (7.70)$$

$$\vec{B}'' = \vec{n}'' \times \frac{\vec{E}''}{c} = \sqrt{\mu\epsilon} \frac{\vec{k}''}{k''} \times \vec{E}''. \quad (7.71)$$

这里我们利用了 $n'' = n = \sqrt{\mu\epsilon/\mu_0\epsilon_0}$ 与 $k'' = k = \omega\sqrt{\mu\epsilon/\mu_0\epsilon_0}/c = \omega\sqrt{\mu\epsilon}$ 。显然, $k' = \omega\sqrt{\mu'\epsilon'/\mu_0\epsilon_0}/c = \omega\sqrt{\mu'\epsilon'}$ 。

在 $z < 0$ 区域, 电场与磁场分别为 $\vec{E} + \vec{E}''$ 和 $\vec{B} + \vec{B}''$, 而在 $z > 0$ 区域, 电场与磁场为 \vec{E}' 和 \vec{B}' 。这样在 $z = 0$ 处, 边界条件 (7.52)-(7.55) 一定具有如下的形式

$$(\dots)e^{i(\vec{k} \cdot \vec{x} - \omega t)} + (\dots)e^{i(\vec{k}'' \cdot \vec{x} - \omega t)} = (\dots)e^{i(\vec{k}' \cdot \vec{x} - \omega t)}. \quad (7.72)$$

这个等式在边界上 ($z = 0$) 任意一点恒成立。因此, 我们一定有, 所有的相位因子一定是相等的, 即

$$\vec{k} \cdot \vec{x} = \vec{k}'' \cdot \vec{x} = \vec{k}' \cdot \vec{x}, \text{ at } z = 0, \quad (7.73)$$

或者说

$$k_x x + k_y y = k''_x x + k''_y y = k'_x x + k'_y y. \quad (7.74)$$

这些方程对所有点 (x and y) 都成立, 有且仅有

$$k_x = k''_x = k'_x, \quad (7.75)$$

与

$$k_y = k''_y = k'_y. \quad (7.76)$$

接下来我们设置 x 轴, 使得 \vec{k} 位于 xz 平面内, 即 $k_y = 0$ 。方程 (7.76) 得出 $k_y'' = k_y' = 0$; 即 \vec{k}'' 与 \vec{k}' 都是位于 xz 平面内。所以我们证明了第一个结论:

入射, 反射与折射波的波矢位于同一个平面内, 该平面被称为入射平面, 且该平面也包括了分界面的法向矢量 (此处为 \vec{e}_z)。

第二, 从方程 (7.75) 得到

$$k \sin \theta_i = k'' \sin \theta_{r'} = k' \sin \theta_r. \quad (7.77)$$

因为 $k = k''$, 我们立刻得到 $\theta_i = \theta_{r'}$, 即

入射角等于反射角。

而对于折射角 θ_r 而言, 我们有

$$\frac{\sin \theta_i}{\sin \theta_r} = \frac{k'}{k} = \frac{n'}{n}. \quad (7.78)$$

这即是折射定律, 或斯奈尔定律。

为了得到三支波的相位与强度关系, 我们需要运用具体的边界条件将这三支波联系起来。为简化讨论, 我们假设入射波是线极化的。一般情况下的波可以分解为两支线性极化波。有两种情况可以讨论, 一是入射波电场 \vec{E} 是位于入射平面内的。二是入射波电场是垂直于入射平面的。

我们先考虑 \vec{E} 垂直于入射平面的情况。让所有 \vec{E} 位于 y 方向, 即电场振幅 $\vec{E}_0 = E_0 \vec{e}_y$ 等。所有磁场 \vec{B} 则为 $\vec{B} = \vec{n} \times \vec{E}/c$ 。利用边界条件 (7.53) 与 (7.55), 我们有

$$E_0 + E_0'' - E_0' = 0, \quad (7.79)$$

$$\sqrt{\frac{\epsilon}{\mu}}(E_0 - E_0'') \cos \theta_i - \sqrt{\frac{\epsilon'}{\mu'}} E_0' \cos \theta_r = 0. \quad (7.80)$$

另外两个边界条件所给出的信息是一样的。利用入射波电场振幅, 反射与折射波的振幅为

$$\frac{E_0'}{E_0} = \frac{2n \cos \theta_i}{n \cos \theta_i + (\mu/\mu') \sqrt{n'^2 - n^2 \sin^2 \theta_i}}, \quad (7.81)$$

$$\frac{E_0''}{E_0} = \frac{n \cos \theta_i - (\mu/\mu') \sqrt{n'^2 - n^2 \sin^2 \theta_i}}{n \cos \theta_i + (\mu/\mu') \sqrt{n'^2 - n^2 \sin^2 \theta_i}}, \quad (7.82)$$

这里我们利用了斯奈尔定律。这里的振幅关系是对 \vec{E} 垂直于入射平面的情况。

现在如果 \vec{E} 是平行于入射平面的。我们让所有的磁场位于 $-\vec{e}_y$ 方向。边界条件 (7.53) 与 (7.55) 给出

$$\cos \theta_i (E_0 - E_0'') - \cos \theta_r E_0' = 0, \quad (7.83)$$

$$\sqrt{\frac{\epsilon}{\mu}}(E_0 + E_0'') - \sqrt{\frac{\epsilon'}{\mu'}} E_0' = 0. \quad (7.84)$$

这些联合起来可以得到

$$\frac{E_0'}{E_0} = \frac{2nn' \cos \theta_i}{(\mu/\mu') n'^2 \cos \theta_i + n \sqrt{n'^2 - n^2 \sin^2 \theta_i}}, \quad (7.85)$$

$$\frac{E_0''}{E_0} = \frac{(\mu/\mu') n'^2 \cos \theta_i - n \sqrt{n'^2 - n^2 \sin^2 \theta_i}}{(\mu/\mu') n'^2 \cos \theta_i + n \sqrt{n'^2 - n^2 \sin^2 \theta_i}}. \quad (7.86)$$

对于垂直入射情况 $\theta_i = 0$ ，两种结论给出相同的结果

$$\frac{E'_0}{E_0} = \frac{2}{\sqrt{\mu\epsilon'/\mu'\epsilon + 1}}, \quad (7.87)$$

$$\frac{E''_0}{E_0} = \frac{\sqrt{\mu\epsilon'/\mu'\epsilon} - 1}{\sqrt{\mu\epsilon'/\mu'\epsilon} + 1}. \quad (7.88)$$

如果我们进一步假设 $\mu' = \mu$ ，则

$$\frac{E'_0}{E_0} = \frac{2n}{n' + n}, \quad (7.89)$$

$$\frac{E''_0}{E_0} = \frac{n' - n}{n' + n}. \quad (7.90)$$

7.5.2 反射光的极化

从上述的振幅关系我们可以得到一个关于反射波有意思且有用的结论。如果 \vec{E} 是位于入射平面内的，则从方程 (7.86)，我们可以看到存在一个特殊的入射角， θ_B ，使得反射波完全消失，即 $E''_0 = 0$ 。依旧为简化分析假设 $\mu' = \mu$ ，这个特殊的角度所满足

$$n'^2 \cos \theta_B - n \sqrt{n'^2 - n^2 \sin^2 \theta_B} = 0, \quad (7.91)$$

从中可以解出

$$\theta_B = \tan^{-1} \left(\frac{n'}{n} \right). \quad (7.92)$$

这个角度称为布鲁斯特角 (Brewster's angle)。注意到如果 \vec{E} 是垂直于入射平面时，反射波是不会消失的。现在我们假设入射波是非极化的，在平行和垂直于入射平面的两个方向上有相同的电场。当这样的波以布鲁斯特角入射时，反射波的电场会完全垂直于入射平面，因为另一个方向上的反射电场消失了。所以反射波此时是完全线性极化的。而一般情况下，在其它角度，反射波是部分偏振的。这样的一个物理过程被广泛用于摄影的圆偏振滤镜或偏振太阳镜中用来部分或完全地消除地面或水面的反光。

7.6 折射率对频率的依赖关系模型

之前提到电磁波在介质中传播的折射率是可以由介质的电学性质计算出来的。在这一章里，我们利用 $n \approx \sqrt{\epsilon_r}$ ，从电动力学角度来建立一个关于折射率的模型。这个模型可以解释常见的折射率对频率的依赖关系，或者说光的色散 (dispersion)。我们从建立一个简单的关于 ϵ 的模型出发，来得到折射率 n 。

与之前静电学里的关于极化过程的模型类似，我们现在需要考虑的是在一个外加电磁场作用下，电子的运动位移，并从相应的电偶极矩 \vec{p} 与外加电场之间的关系得到 ϵ 。介质内部的束缚电子在外加电磁场作用下的运动方程可以由模型近似：

$$\ddot{\vec{x}} + \gamma \dot{\vec{x}} + \omega_0^2 \vec{x} = -\frac{e}{m} \vec{E}(\vec{x}, t) = -\frac{e}{m} \vec{E}_0 e^{-i\omega t}. \quad (7.93)$$

这里 ω_0 电子被束缚在势阱里的自然振动频率， ω 为外加电磁波的频率。为了防止在共振处出现的无穷大振幅的解，我们引入阻尼，由 γ 表示，注意 $\gamma \ll (\omega_0, \omega)$ 。这里我们忽略了磁场所施加的力，这是束缚电子通常可以认为是非相对论的 ($v \ll c$)，此时 $vB/E \sim v/c \ll 1$ 。整个模型是一个在外加电场力作用下带阻尼的谐振子模型。这个方程的解可以简单表达为。

$$\vec{x} = -\frac{e}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \vec{E}. \quad (7.94)$$

相应的电偶极矩为

$$\vec{p} = -e\vec{x} = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \vec{E}. \quad (7.95)$$

现在我们假设每单位体积里有 N 个分子，而每个分子有 Z 个电子。为了建立一个更有趣的模型，我们假设这些电子的本征频率不是单一的，即每个分子里有 f_j 个电子，其本征频率为 ω_j ，阻尼系数为 γ_j 。这样极化矢量 \vec{P} 为

$$\vec{P} = \frac{Ne^2}{m} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\gamma_j} \vec{E} = \chi_e \epsilon_0 \vec{E}. \quad (7.96)$$

这里 χ_e 是电极化率。因此相对介电常数 ϵ 为

$$\epsilon_r = 1 + \chi_e = 1 + \frac{Ne^2}{m\epsilon_0} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\gamma_j}. \quad (7.97)$$

注意这里 ϵ_r 是一个复数，因此 $n = \sqrt{\epsilon_r}$ 与 $k = n\omega/c$ 都是复数，这显然代表了电磁波的能量变化。如果 $k = k_r + ik_i$ 且 $k_i > 0$ ，则波沿 $+x$ 是衰减的。因为波的能量是与 $Re(\vec{E})^2$ 成正比的，所以我们会定义

$$k = \beta + i\frac{\alpha}{2}, \quad (7.98)$$

这样波的强度会以 $e^{-\alpha x}$ 变化。

对大部分气体而言，方程 (7.97) 右边的第二项很小。利用 $\sqrt{1+x} \approx 1 + x/2$ ，我们可以近似得到

$$n \approx \sqrt{\epsilon_r} \approx 1 + \frac{Ne^2}{2m\epsilon_0} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\gamma_j} \quad (7.99)$$

这样我们将折射率 n 分成实数和虚数部分， $n = n_r + in_i$ ，或者

$$n_r = 1 + \frac{Ne^2}{2m\epsilon_0} \sum_j \frac{f_j(\omega_j^2 - \omega^2)}{(\omega_j^2 - \omega^2)^2 + (\omega\gamma_j)^2}, \quad (7.100)$$

$$n_i = \frac{Ne^2\omega}{2m\epsilon_0} \sum_j \frac{f_j\gamma_j}{(\omega_j^2 - \omega^2)^2 + (\omega\gamma_j)^2} \quad (7.101)$$

这里 n_r 便是光学里学过的普通折射率，而 n_i 则代表着波的衰减或增长。代入 $k = n\omega/c$ 并利用方程 (7.101)，我们有

$$\alpha = \frac{Ne^2\omega^2}{m\epsilon_0 c} \sum_j \frac{f_j\gamma_j}{(\omega_j^2 - \omega^2)^2 + (\omega\gamma_j)^2} \quad (7.102)$$

这样折射率便由方程 (7.101) 给出，同时衰减系数是方程 (7.102)。

我们可以从方程 (7.101) 看到，一般情况下，在远离共振时，折射率随频率增加而增加（例如牛顿的色散实验）。然而，在共振频率附近，折射率会随频率增加而快速减小。这一现象是后来发现的，所以这种色散也被称为反常色散（anomalous dispersion）。同时从方程 (7.102) 出发，我们看到在共振点附近， α 变得非常大：很多电磁波的能量被介质共振吸收了。

如果我们远离所有共振频率，我们可以得到一个更简化的折射率公式。从上述方程出发，折射率可以近似为

$$n_r = 1 + \frac{Ne^2}{2m\epsilon_0} \sum_j \frac{f_j}{\omega_j^2 - \omega^2}. \quad (7.103)$$

对透明物质而言，例如大气，最接近可见光的共振频率在极紫外区间，所以 $\omega < \omega_j$ 。在这种情况下

$$\frac{1}{\omega_j^2 - \omega^2} \approx \frac{1}{\omega_j^2} \left(1 + \frac{\omega^2}{\omega_j^2} \right), \quad (7.104)$$

因此,

$$n_r = 1 + \left(\frac{Ne^2}{2m\epsilon_0} \sum_j \frac{f_j}{\omega_j^2} \right) + \omega^2 \left(\frac{Ne^2}{2m\epsilon_0} \sum_j \frac{f_j}{\omega_j^4} \right). \quad (7.105)$$

用真空中的光的波长 ($\lambda = 2\pi c/\omega$) 表达,

$$n = 1 + A(1 + \frac{B}{\lambda^2}). \quad (7.106)$$

这便是光学里的柯西公式 (Cauchy's equation); 系数 A 称为折射系数, 而 B 称为色散系数。柯西公式在可见光范围内对大部分气体适用。

在我们结束本节之前, 我们还可以讨论两个 ϵ 的简单极限。

情况一: 低频极限 $\omega \rightarrow 0$ 。如果所有的 $\omega_j \neq 0$, 则我们可以得到静电学里的分子极化率。

情况二: 高频极限。如果对所有的 j , $\omega \gg \omega_j$, 则

$$\epsilon_r = 1 + \frac{Ne^2}{m\epsilon_0} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\gamma_j} \approx 1 - \frac{NZe^2}{m\epsilon_0} \frac{1}{\omega^2} \equiv 1 - \frac{\omega_p^2}{\omega^2}. \quad (7.107)$$

这里

$$\omega_p^2 = \frac{NZe^2}{m\epsilon_0}, \quad (7.108)$$

是非常重要的等离子体频率 (plasma frequency)。利用 $n^2 = \epsilon_r$, 相应的色散关系为

$$\omega^2 = \omega_p^2 + c^2 k^2. \quad (7.109)$$

这个色散关系在等离子体科学里也被用来测量电子的密度, 因为当 $\omega \rightarrow \omega_p$ 时, $k \rightarrow 0$, 意味着波不能在这个等离子体里传播。

7.7 导体中的电磁波

在导体中, 因为存在可以自由移动的电子, 从上面的模型所导出的介电常数一般有实部和虚部。通常习惯性地 将实部称为 ϵ , 而虚部则对应着电导率 σ 。利用欧姆定律, $\vec{j}_f = \sigma \vec{E}$, 我们可以得到导体中的麦克斯韦方程为

$$\nabla \cdot \vec{E} = \frac{\rho_f}{\epsilon}, \quad (7.110)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (7.111)$$

$$\nabla \cdot \vec{B} = 0, \quad (7.112)$$

$$\nabla \times \vec{B} = \mu\epsilon \frac{\partial \vec{E}}{\partial t} + \mu\sigma \vec{E}. \quad (7.113)$$

利用连续性方程和高斯定理 (方程 (7.110)), 我们得到

$$\frac{\partial \rho_f}{\partial t} = -\nabla \cdot \vec{j}_f = -\sigma (\nabla \cdot \vec{E}) = -\frac{\sigma}{\epsilon} \rho_f. \quad (7.114)$$

这个方程有如下形式的解

$$\rho_f = \rho_{f0} e^{-t/\tau}, \quad (7.115)$$

这里时间尺度 $\tau = \epsilon/\sigma$ 。这个方程表明在导体中以任何形式聚集的自由电荷都会在 τ 这么长的时间尺度内消失。对于一个理想导体, $\sigma = \infty$, 所以 $\tau = 0$ 。一般情况下, 我们可以将 τ 与相关时间尺度比较, 从而来判断一个导体

有多接近理想导体。对于频率为 ω 的时变电磁波而言，相关时间尺度为 $1/\omega$ 。所以，我们定义如果 $\tau \ll 1/\omega$ ，则该导体为良导体；而如果 $\tau \gg 1/\omega$ ，则为差的导体。这两个方程可以被写为

$$\begin{cases} \frac{\sigma}{\epsilon\omega} \gg 1, \\ \frac{\sigma}{\epsilon\omega} \ll 1, \end{cases} \quad (7.116)$$

现在起我们仅考虑良导体。相较之下， τ 是非常非常小的，一般为 10^{-14} 秒量级。这样对于大部分物理现象 ($T \gg \tau$)，我们不用考虑上述麦克斯韦方程里的自由电荷，而可以将其设为 0。这样麦克斯韦方程为

$$\nabla \cdot \vec{E} = 0, \quad (7.117)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (7.118)$$

$$\nabla \cdot \vec{B} = 0, \quad (7.119)$$

$$\nabla \times \vec{B} = \mu\epsilon \frac{\partial \vec{E}}{\partial t} + \mu\sigma \vec{E}. \quad (7.120)$$

该方程与之前介质中的麦克斯韦方程类似，除了在安培定理中多了一个从自由电流而来的 $\mu\sigma \vec{E}$ 一项。

类似于之前，我们可以得到 \vec{E} 与 \vec{B} 的方程：

$$\nabla^2 \vec{E} = \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} + \mu\sigma \frac{\partial \vec{E}}{\partial t}, \quad (7.121)$$

$$\nabla^2 \vec{B} = \mu\epsilon \frac{\partial^2 \vec{B}}{\partial t^2} + \mu\sigma \frac{\partial \vec{B}}{\partial t}. \quad (7.122)$$

加入阻尼谐振子的讨论。

假设 $\vec{E} = \vec{E}_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)}$ ， \vec{B} 也类似，则我们会得到一个复数的 k ，即

$$k^2 = \mu\epsilon\omega^2 + i\mu\sigma\omega. \quad (7.123)$$

这就是导体中的电磁波的色散关系。

让 $k = \beta + i\alpha/2$ ，我们有

$$\beta = \sqrt{\frac{\mu\epsilon}{2}} \omega \left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} + 1 \right]^{1/2}, \quad (7.124)$$

$$\frac{\alpha}{2} = \sqrt{\frac{\mu\epsilon}{2}} \omega \left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} - 1 \right]^{1/2}. \quad (7.125)$$

对于一个良导体而言， $\sigma/\epsilon\omega \gg 1$ ， $\beta \approx \alpha/2$ ，则

$$k = (1 + i)\sqrt{\mu\sigma\omega/2} \quad (7.126)$$

之前讲过， k 的虚数部分表示的是波振幅的增加或衰减，此时波的振幅有 $e^{-k_i x}$ 的形式，其空间尺度显然为 $1/k_i$ 。对导体中的电磁波而言 $k_i > 0$ ，所以电磁波在导体里是衰减的。其衰减的空间尺度，即振幅减少 $1/e \approx 1/3$ 所需传播距离，称为趋肤深度 (skin depth) δ ，即

$$\delta = \frac{2}{\alpha} \approx \sqrt{\frac{2}{\mu\sigma\omega}}. \quad (7.127)$$

对于像铜这样的良导体，对 60 Hz 的电磁波， $\delta \approx 0.85$ ，而对于 100 Mhz 的波， $\delta \approx 0.71 \times 10^{-3}$ cm。

我们可以通过法拉弟定委来得到导体中电磁波的磁场 \vec{B} 和电场 \vec{E} 的关系。对良导体，

$$\vec{B} = \vec{n} \times \frac{\vec{E}}{c} = \frac{c\vec{k}}{\omega} \times \frac{\vec{E}}{c} = \frac{1}{\omega} \sqrt{\mu\sigma\omega/2} (1 + i) \vec{e}_k \times \vec{E}, \quad (7.128)$$

即

$$\vec{B} = \sqrt{\frac{\mu\sigma}{\omega}} e^{i\pi/4} \vec{e}_k \times \vec{E}. \quad (7.129)$$

因此, 导体中的电磁波磁场与电场间有一个相位差。对良导体而言, \vec{B} 在相位上落后 \vec{E} $\pi/4$ 。另一方面, 线性介质中电磁场的能量密度为

$$w = \frac{1}{2}(\vec{E} \cdot \vec{D} + \vec{B} \cdot \vec{H}) = \frac{1}{2}(\epsilon E^2 + \frac{1}{\mu} B^2). \quad (7.130)$$

所以在良导体中, 电磁波的磁场能量密度和电场能量密度的比值为

$$\frac{w_B}{w_E} = \frac{B^2/\mu}{\epsilon E^2} = \frac{\sigma}{\epsilon\omega} \gg 1. \quad (7.131)$$

因为与真空中不同, 导体中的电磁波能量大部分是在磁场中。

7.8 导体表面电磁波的反射

我们考虑电磁波从真空垂直入射一个良导体。相应边界条件是

$$\vec{E}_1^{\parallel} = \vec{E}_2^{\parallel}, \quad (7.132)$$

$$\vec{H}_1^{\parallel} - \vec{H}_2^{\parallel} = \vec{k}_f \times \vec{n}. \quad (7.133)$$

注意, 对一个良导体而言, σ 是有限的, 而 $\vec{j}_f = \sigma \vec{E}$ 意味着, 我们不同有一个像 δ 函数一样的面电流密度 ($\sigma \rightarrow 0$), 除非 \vec{E} 无穷大³。因此良导体的边界条件为

$$\vec{E}_1^{\parallel} = \vec{E}_2^{\parallel}, \quad (7.134)$$

$$\vec{H}_1^{\parallel} = \vec{H}_2^{\parallel}. \quad (7.135)$$

利用之前章节的符号与推导过程, 我们得到

$$E_0 + E_0'' = E_0', \quad (7.136)$$

$$H - H'' = H'. \quad (7.137)$$

假设 $\mu_r \approx \mu_r' \approx 1$ 并利用

$$H' = \sqrt{\frac{\sigma}{2\mu_0\omega}} (1 + i) E_0', \quad (7.138)$$

$$H = E/\mu_0 c, \quad (7.139)$$

$$H'' = E''/\mu_0 c. \quad (7.140)$$

我们将方程 (7.137) 写为

$$E_0 - E_0'' = c \sqrt{\frac{\mu_0\sigma}{2\omega}} (1 + i) E_0'. \quad (7.141)$$

求解方程 (7.141) 与 (7.136), 我们得到

$$\frac{E_0''}{E_0} = -\frac{1 + i - \sqrt{2\omega/\mu_0\sigma}/c}{1 + i + \sqrt{2\omega/\mu_0\sigma}/c}. \quad (7.142)$$

³在理想导体里, 面电流是可以存在的。

相应的垂直入射时的反射系数为

$$R = \left| \frac{E_0''}{E_0} \right|^2 = \frac{\left(1 - \sqrt{2\omega/\mu_0\sigma/c}\right)^2 + 1}{\left(1 + \sqrt{2\omega/\mu_0\sigma/c}\right)^2 + 1} \approx 1 - \frac{2}{c} \sqrt{\frac{2\omega}{\mu_0\sigma}} = 1 - 2\sqrt{\frac{2\omega\epsilon_0}{\sigma}}. \quad (7.143)$$

对良导体而言, $\omega\epsilon_0/\sigma \ll 1$, 因此 $R \rightarrow 1$ 。这也是为什么良导体可以做成镜子的原因。在中国古代, 有以铜为镜, 而现在有以玻璃背后镀银作镜子。

7.9 时变电磁场的边值问题

目前为止, 我们仅考虑了无限均匀介质中的时变电磁场。在本节, 我们来看一看当电磁波的传播受到边界条件限制时, 其性质会发生什么样的变化。所考虑的例子是中空金属管中的电磁波。如果这个金属管两端也是封闭的, 则其称为一个谐振腔 (cavity); 否则, 则称为一个波导管 (waveguide)。注意, 这里的波导管是一个很广义的词, 其应用不局限于电磁波情况—任何可以引导波传播的物质都可以被称为波导管。为了简化讨论, 我们假设波导的边界是由理想导体组成的, 且其截面是尺坟固定的矩形。我们重点考虑这些边界条件会如何改变波的性质。我们会发现, 一般而言, 波导管中的电磁波不再是横波。

我们假设波导管中由均匀非耗散的线性介质组成, 其电与磁常数分别为 ϵ 和 μ 。如果我们假设波的频率为 ω , 这样介质中的麦克斯韦方程是

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} = i\omega \vec{B}, \quad (7.144)$$

$$\nabla \cdot \vec{E} = 0, \quad (7.145)$$

$$\nabla \cdot \vec{B} = 0, \quad (7.146)$$

$$\nabla \times \vec{B} = \mu\epsilon \frac{\partial \vec{E}}{\partial t} = -i\mu\epsilon\omega \vec{E}. \quad (7.147)$$

这些方程与之前章节中的方程类似, 我们很容易得到 \vec{E} 和 \vec{B} 的方程为

$$(\nabla^2 + \mu\epsilon\omega^2) \begin{pmatrix} \vec{E} \\ \vec{B} \end{pmatrix} = 0, \quad (7.148)$$

其边界条件是

$$\vec{E}^{\parallel} = 0, \quad (7.149)$$

$$B^{\perp} = 0. \quad (7.150)$$

注意这里我们没有假设电磁场对空间具有平面波的依赖关系, $e^{i\vec{k}\cdot\vec{x}}$, 因为这取决于我们的边界条件。

7.9.1 波导

假设沿波传播的方向为 z 。在此情况下, 波沿 z 方向传播; 因此, 我们可以假设对 z 的依赖关系为 e^{ikz} , 即

$$\vec{E}(x, y, z, t) = \vec{E}(x, y)e^{i(kz - \omega t)}, \quad (7.151)$$

$$\vec{B}(x, y, z, t) = \vec{B}(x, y)e^{i(kz - \omega t)}. \quad (7.152)$$

则场方程 (7.148) 变成二维情况,

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + (\mu\epsilon\omega^2 - k^2) \right] \begin{pmatrix} \vec{E} \\ \vec{B} \end{pmatrix} = 0. \quad (7.153)$$

这样我们可以通过求解此方程得到电磁场的三个分量对 (x, y) 的依赖关系, 例如分离变量法。但在此之前, 我们可以利用麦克斯韦方程中的法拉弟与安培定律发现场的三个分量并不是独立的, 例如, 如果我们知道 E_z 与 B_z 的话, E_x 和 E_y 是完全可以确定的。

从法拉弟与安培定律得到

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = i\omega B_z, \quad (7.154)$$

$$\frac{\partial E_z}{\partial y} - ikE_y = i\omega B_x, \quad (7.155)$$

$$ikE_x - \frac{\partial E_z}{\partial x} = i\omega B_y, \quad (7.156)$$

$$\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} = -i\mu\epsilon\omega E_z, \quad (7.157)$$

$$\frac{\partial B_z}{\partial y} - ikB_y = -i\mu\epsilon\omega E_x, \quad (7.158)$$

$$ikB_x - \frac{\partial B_z}{\partial x} = -i\mu\epsilon\omega E_y. \quad (7.159)$$

方程 (7.155), (7.156), (7.158) 与 (7.159) 可以结合得到

$$E_x = \frac{i}{\mu\epsilon\omega^2 - k^2} \left(k \frac{\partial E_z}{\partial x} + \omega \frac{\partial B_z}{\partial y} \right), \quad (7.160)$$

$$E_y = \frac{i}{\mu\epsilon\omega^2 - k^2} \left(k \frac{\partial E_z}{\partial y} - \omega \frac{\partial B_z}{\partial x} \right), \quad (7.161)$$

$$B_x = \frac{i}{\mu\epsilon\omega^2 - k^2} \left(k \frac{\partial B_z}{\partial x} - \mu\epsilon\omega \frac{\partial E_z}{\partial y} \right), \quad (7.162)$$

$$B_y = \frac{i}{\mu\epsilon\omega^2 - k^2} \left(k \frac{\partial B_z}{\partial y} + \mu\epsilon\omega \frac{\partial E_z}{\partial x} \right). \quad (7.163)$$

注意, 这里的条件是 $\mu\epsilon\omega^2 - k^2 \neq 0$ 。从这些方程中我们可以看到, 如果我们知道了 E_z 与 B_z , 则电磁场的 x, y 分量完全确定。而 E_z 与 B_z 的方程容易从式 (7.153) 中得到,

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + (\mu\epsilon\omega^2 - k^2) \right] \begin{pmatrix} E_z \\ B_z \end{pmatrix} = 0. \quad (7.164)$$

取决于 E_z 与 B_z , 波导管中可以存在三种基本的模式。如果 $E_z = 0$, 我们称这些波为 TE (“transverse electric”) 模式的波动; 如果 $B_z = 0$, 我们称其为 TM (“transverse magnetic”) 模式的波。如果 $E_z = 0$ 且 $B_z = 0$, 则我们称其为 TEM (“transverse electromagnetic”) 模式的波。

注意 TEM 模式的波动不能出现在中空的波导中。如果 $E_z = 0$ 且 $B_z = 0$, 我们不能再利用方程 (7.160)-(7.163); 因为它们不是互相独立的, 因而不确定的。这时我们需要回到麦克斯韦方程。利用高斯定理

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0. \quad (7.165)$$

如果 $B_z = 0$, 则法拉弟定理给出

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = 0. \quad (7.166)$$

场 $\vec{E} = E_x \vec{e}_x + E_y \vec{e}_y$ 为无源且无旋。所以我们可以利用一个标势来表达 \vec{E} , 且该标势满足拉普拉斯方程。但是 \vec{E} 的边界条件为 $\vec{E}^\parallel = 0$, 或者说导体边界意味着等势面, 而拉氏方程的解在求解区域内部不会有极大或极小值, 而以标势的解必然也为常数, 也即意味着 $\vec{E} = 0$, 因为 TEM 模不能在中空的波导中出现。但是如果波导的中心还有一个导体, 例如同轴传输线, 则标势在内外导体上不一定相同, 此时波导中可以存在非零的电磁场, 允许 TEM 模式的传播。

7.9.2 矩形波导中的 TE 模式波动

假设我们有一个高为 a 宽为 b 的矩形波导，现在来看看其中传播的 TE 模有何性质。从之前的讨论，我们仅需求解

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + (\mu\epsilon\omega^2 - k^2) \right] B_z = 0. \quad (7.167)$$

现在我们需要确定其边界条件。因为 $B^\perp = 0$ ，所以从方程 (7.162) 与 (7.163)，我们得到边界条件为

$$\frac{\partial B_z}{\partial x} = 0, \text{ at } x = 0, a, \quad (7.168)$$

$$\frac{\partial B_z}{\partial y} = 0, \text{ at } y = 0, b. \quad (7.169)$$

利用分离变量法, $B_z(x, y) = X(x)Y(y)$, 则

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \mu\epsilon\omega^2 - k^2 = 0. \quad (7.170)$$

因此，我们有

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k_x^2, \quad (7.171)$$

$$\frac{1}{Y} \frac{d^2 Y}{dy^2} = -k_y^2, \quad (7.172)$$

$$k_x^2 + k_y^2 + k^2 = \mu\epsilon\omega^2. \quad (7.173)$$

而 X 和 Y 的一般解为

$$X(x) = A \sin(k_x x) + B \cos(k_x x), \quad (7.174)$$

$$Y(y) = C \sin(k_y y) + D \cos(k_y y). \quad (7.175)$$

利用边界条件

$$X(x) = B \cos(k_x x), \text{ with } k_x = m\pi/a, \quad (7.176)$$

$$Y(y) = D \cos(k_y y), \text{ with } k_y = n\pi/b. \quad (7.177)$$

这里 $m = 0, 1, 2, 3, \dots$, 而 $n = 0, 1, 2, 3, \dots$ 。因此，

$$B_z = B_0 \cos(m\pi x/a) \cos(n\pi y/b). \quad (7.178)$$

这个解对应的波的模式称为 TE_{mn} 模式。习惯上 $m \geq n$ ，注意 m 和 n 不能同时为 0。 TE_{mn} 模式的波数为

$$k = \sqrt{\mu\epsilon\omega^2 - \pi^2[(m/a)^2 + (n/b)^2]}. \quad (7.179)$$

如果

$$\omega < \frac{1}{\sqrt{\mu\epsilon}} \pi \sqrt{(m/a)^2 + (n/b)^2} \equiv \omega_{mn}, \quad (7.180)$$

波数变成虚数，意味着波不能在导管中传播很远。因此 ω_{mn} 称为 TE_{mn} 的截止频率 (cutoff frequency)。对一个给定尺寸的波导管而言，其最低截止频率为

$$\boxed{\omega_{10} = \frac{\pi}{\sqrt{\mu\epsilon a}}}. \quad (7.181)$$

低于这个频率的波不能在该波导管中传播。

利用截止频率，我们可以将波数 k 写为

$$k = \sqrt{\mu\epsilon} \sqrt{\omega^2 - \omega_{mn}^2}. \quad (7.182)$$

波的相速度为

$$v_{ph} = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon} \sqrt{1 - (\omega_{mn}/\omega)^2}}. \quad (7.183)$$

而群速度为

$$v_g = \frac{\partial\omega}{\partial k} = \frac{1}{\sqrt{\mu\epsilon}} \sqrt{1 - (\omega_{mn}/\omega)^2}. \quad (7.184)$$

如果 $\mu = \epsilon = 1$ ，则 $v_{ph} > c$ 且 $v_g < c$ 。

第八章 运动粒子的场

在这一章里，我们将麦克斯韦方程的所有元素（时变，源）都予以保留，找出此时方程的解。这种解对应着给定的源 (ρ, \vec{j}) 所产生的电磁场应是什么样子。

8.1 一般情况下的解：推迟势

我们从麦克斯韦方程的四维形式出发，

$$\frac{\partial F^{\alpha\beta}}{\partial x^\beta} = \mu_0 j^\alpha. \quad (8.1)$$

利用势函数 $A^\alpha = (\phi/c, \vec{A})$ 并选洛伦茨规范，这些方程可以写为

$$\frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = \mu_0 \vec{j}, \quad (8.2)$$

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \frac{\rho}{\epsilon_0}. \quad (8.3)$$

这里算符

$$\square \equiv \partial_\mu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \quad (8.4)$$

称为达朗贝尔算符。

如果场是静态的，则

$$\nabla^2 \vec{A} = -\mu_0 \vec{j}, \quad (8.5)$$

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0}. \quad (8.6)$$

另一面，如果无源，即 $\rho = 0$ 且 $\vec{j} = 0$ ，则

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = 0, \quad (8.7)$$

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0. \quad (8.8)$$

这里对应着无源（真空）情况下场的方程，也是齐次波动方程。在之前的章节里我们都处理这两种情况下场的解。

一般情况下的麦克斯韦方程即非齐次方程，其一般解的形式为非齐次方程的特解加上齐次方程的通解。比如，我们用 ϕ 方程举例，其一般的解为

$$\phi = \phi_1 + \phi_0, \quad (8.9)$$

这里 ϕ_1 是个特解，满足 $\square \phi = \rho/\epsilon_0$ 。而 ϕ_0 则是齐次方程 (8.8) 的通解。

为了找到特解，我们利用叠加原理，考虑位于 \vec{x}' 处电量为 $dq(t)$ 的电荷所产生的势。该电荷的电荷密度为 $\rho = dq(t)\delta(\vec{R})$ ，这里 $\vec{R} = \vec{x} - \vec{x}'$ 。相应的 ϕ 方程 (8.3) 可以写为

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \frac{dq(t)}{\epsilon_0} \delta(\vec{R}). \quad (8.10)$$

在空间其它点 $\vec{R} \neq 0$ ，方程 (8.10) 为

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 0. \quad (8.11)$$

我们所用的方法是先求解方程 (8.11), 得到 $\vec{R} \neq 0$ 处的 ϕ , 然后将这个解与 ϕ 在 $\vec{R} = 0$ 连接起来。

显然对一个点电荷 $dq(t)$, ϕ 是球对称的, 即, $\phi = \phi(R)$ 。利用球坐标系 $\phi = \phi(R)$, 我们将方程 (8.11) 写为

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial \phi}{\partial R} \right) = 0. \quad (8.12)$$

令 $\phi(R, t) = \chi(R, t)/R$, 方程 (8.12) 变成

$$\frac{1}{c^2} \frac{\partial^2 \chi}{\partial t^2} - \frac{\partial^2 \chi}{\partial R^2} = 0. \quad (8.13)$$

这与平面波方程形式相同, 其一般解为

$$\chi(R, t) = f(t - R/c) + g(t + R/c). \quad (8.14)$$

因为我们仅需要一个特解, 我们可以选择 f 或者 g 。注意 f 代表沿 $+\vec{R}$ 方向传播的波, 而 g 代表沿 $-\vec{R}$ 方向传播的波。因为源位于 $R = 0$ 处, 我们从物理上选择 $\chi = f(t - R/c)$, 则

$$\phi = \frac{\chi(t - R/c)}{R}. \quad (8.15)$$

注意这里 χ 的形式是未知的, 我们需要利用边界条件, 即 $\vec{R} = 0$ 处的 ϕ , 来得到 χ 的函数形式。

当 $R \rightarrow 0$, $\phi \rightarrow \infty$, 因此 ϕ 在空间上的变化占主导, 即 $\partial/\partial x \gg \partial/\partial t$ 。此时, 我们可以忽略时间导数, 因此 ϕ 在 $\vec{R} = 0$ 附近的方程成为

$$\nabla^2 \phi = -\frac{dq(t)}{\epsilon_0} \delta(\vec{R}), \quad \vec{R} = 0 \text{ 附近}. \quad (8.16)$$

注意到这即是一个位于 \vec{x}' 处的点电荷 dq 所对应的静电势的方程。利用库伦定理, 我们得到

$$\phi = \frac{dq(t)}{4\pi\epsilon_0 R} \quad \vec{R} = 0 \text{ 附近}. \quad (8.17)$$

方程 (8.15) 所得到的 ϕ 在 $\vec{R} \rightarrow 0$ 时一定与此 ϕ 形式相同, 因此

$$\phi = \frac{dq(t - R/c)}{4\pi\epsilon_0 R} \quad \text{for all } \vec{R}. \quad (8.18)$$

对于任一形式分布的电荷系统, ϕ 方程的解显然是

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 R} \rho\left(\vec{x}', t - \frac{R}{c}\right) dV' + \phi_0. \quad (8.19)$$

这里 \vec{x} 是场点向量, $dV' = dx' dy' dz'$ 。类似的, \vec{A} 的解为

$$\vec{A}(\vec{x}, t) = \int \frac{\mu_0}{4\pi R} \vec{j}\left(\vec{x}', t - \frac{R}{c}\right) dV' + \vec{A}_0. \quad (8.20)$$

现在我们来看看上述解的物理含义。势函数 ϕ 与 \vec{A} 的特解

$$\phi_1(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 R} \rho\left(\vec{x}', t - \frac{R}{c}\right) dV', \quad (8.21)$$

$$\vec{A}_1(\vec{x}, t) = \int \frac{\mu_0}{4\pi R} \vec{j}\left(\vec{x}', t - \frac{R}{c}\right) dV', \quad (8.22)$$

被称为推迟势 (retarded potentials), 因为依赖关系 $t - R/c$ 。这里我们定义 $t_r \equiv t - R/c$ 为推迟时刻。即在场点 \vec{x} 处, t 时刻的场依赖于源 ρ 和 \vec{j} 在推迟时刻 $t_r = t - R/c$ 时的状态。这显然与场的传播是有限速度的相关。注意到源的不同点, 其推迟时刻是不一样的。例如, 假定源上两点 \vec{x}'_1 与 \vec{x}'_2 , 到场点 \vec{x} 的距离为 R_1 与 R_2 。则势函数 $\phi(\vec{x}, t)$ 是由 $\rho(\vec{x}'_1, t - R_1/c)$ 与 $\rho(\vec{x}'_2, t - R_2/c)$ 决定的。这里两点的推迟时刻分别为 $t - R_1/c$ 和 $t - R_2/c$, 显然一般情况下 $R_1 \neq R_2$, 所以两个推迟时刻不同。

齐次方程的一般解一般由边界或初始条件决定。而上述特解代表的是源电荷和电流所产生的向外传播的场。在剩下的课程里，我们仅考虑此特解所对应的场，来理解时变的源会向外辐射出什么样的电磁场。

如果我们对比静态场情况下的势函数的解，

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 R} \rho(\vec{x}') dV', \quad (8.23)$$

$$\vec{A}(\vec{x}, t) = \int \frac{\mu_0}{4\pi R} \vec{j}(\vec{x}') dV'. \quad (8.24)$$

与推迟势，很容易得到它们之间的区别仅仅是在源里加入了对推迟时间 $t - R/c$ 的依赖。然而这个小技巧并不适用于一般的 \vec{E} 与 \vec{B} 。

如果要得到一般情况下 \vec{E} 与 \vec{B} 的解，我们可以从推迟势出发，利用

$$\vec{E} = -\nabla\phi - \frac{\partial\vec{A}}{\partial t}, \quad (8.25)$$

$$\vec{B} = \nabla \times \vec{A}. \quad (8.26)$$

这里的空间导数是对场点 \vec{x} 。注意到推迟势的形式，

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 R} \rho\left(\vec{x}', t - \frac{R}{c}\right) dV', \quad (8.27)$$

$$\vec{A}(\vec{x}, t) = \int \frac{\mu_0}{4\pi R} \vec{j}\left(\vec{x}', t - \frac{R}{c}\right) dV'. \quad (8.28)$$

它们对 \vec{x} 的依赖在两个地方，即 $R = |\vec{x} - \vec{x}'|$ 和 $t_r = t - R/c$ 。作为一个练习，很容易证明

$$\begin{aligned} \vec{E}(\vec{x}, t) = & \frac{1}{4\pi\epsilon_0} \int \left[\left(\frac{\rho(\vec{x}', t_r)}{|\vec{x} - \vec{x}'|^3} + \frac{1}{|\vec{x} - \vec{x}'|^2 c} \frac{\partial \rho(\vec{x}', t_r)}{\partial t} \right) (\vec{x} - \vec{x}') \right. \\ & \left. - \frac{1}{|\vec{x} - \vec{x}'|^2 c} \frac{\partial \vec{j}(\vec{x}', t_r)}{\partial t} \right] d^3\vec{x}' \end{aligned} \quad (8.29)$$

$$\vec{B}(\vec{x}, t) = \frac{\mu_0}{4\pi} \int \left[\frac{\vec{j}(\vec{x}', t_r)}{|\vec{x} - \vec{x}'|^3} + \frac{1}{|\vec{x} - \vec{x}'|^2 c} \frac{\partial \vec{j}(\vec{x}', t_r)}{\partial t} \right] \times (\vec{x} - \vec{x}') d^3\vec{x}' \quad (8.30)$$

这些方程称为 Jefimenko 方程。然后它们的应用不如推迟势广泛。一般我们都是选得到推迟势，然后从推迟势里得到电场与磁场。

8.2 点电荷的推迟势：李纳-维谢尔势 (Lienard-Wiechert potentials)

我们现在考虑一个任意形式运动的点电荷 q 的推迟势。假如电荷的轨迹是由 $\vec{x}_p(t)$ 与 $\vec{v}(t)$ 给定。从推迟势出发， $\phi(\vec{x}, t)$ 与 $\vec{A}(\vec{x}, t)$ 是由 $\vec{x}_p(t_r)$ 和 $\vec{v}(t_r)$ 来决定的。这里我们介绍两种不同的方法，来得到点电荷的推迟势，也称为李纳-维谢尔势。

方法一：一个很自然的想法是利用上述一般的推迟势结合点电荷的电荷与电流密度表达示，来得到点电荷的推迟势。对这种方法的掌握有助于理解为什么李纳-维谢尔势具有那样的形式。我们以标势 ϕ 为例说明。

对上述点电荷，其电荷密度可表达为 $\rho(\vec{x}, t) = q\delta[\vec{x} - \vec{x}_p(t)]$ 。将其代入推迟势的表达示，我们有

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 |\vec{x} - \vec{x}'|} q\delta[\vec{x}' - \vec{x}_p(t_r)] dV', \quad (8.31)$$

这里的积分函数包含一个 δ 函数，但是这里的积分结果并不是 $\phi = q/4\pi\epsilon_0 |\vec{x} - \vec{x}_p|$ 。这是因为推迟时刻 $t_r = t - |\vec{x} - \vec{x}'| = t_r(\vec{x}')$ 。所以这里涉及到 δ 函数的积分具有如下形式，

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 R} q\delta[\vec{g}(\vec{x}')] dV'. \quad (8.32)$$

这是一个关于三维复合 δ 函数的积分，其积分稍显复杂。但一个简便方法如下：

$$\phi(\vec{x}, t) = \int \frac{1}{4\pi\epsilon_0 |\vec{x} - \vec{x}'|} q \delta[\vec{x}' - \vec{x}_p(t_r)] dV' = \frac{q}{4\pi\epsilon_0} \iint \frac{\delta[\vec{x}' - \vec{x}_p(\tau)]}{|\vec{x} - \vec{x}'|} \delta[\tau - t_r(\vec{x}')] d\tau dV'. \quad (8.33)$$

这里通过引入一个关于独立变量 τ 的积分，使得关于 \vec{x}' 的积分变成普通的三维 δ 函数积分，不再是上述的三维复合 δ 函数形式，因而非常简单。先对 \vec{x}' 积分，得

$$\phi(\vec{x}, t) = \frac{q}{4\pi\epsilon_0} \int d\tau \left\{ \int \frac{\delta[\vec{x}' - \vec{x}_p(\tau)]}{|\vec{x} - \vec{x}'|} \delta[\tau - t_r(\vec{x}')] dV' \right\} = \frac{q}{4\pi\epsilon_0} \int d\tau \frac{\delta(\tau - t_r)}{|\vec{x} - \vec{x}_p(\tau)|}. \quad (8.34)$$

注意，最后一步式子里的推迟时刻因为积分的关系变为 $t_r = t - |\vec{x} - \vec{x}_p(\tau)|/c$ 。虽然这仍然是一个复合 δ 函数，但是因为是一维的，比较容易处理。接下来我们利用

$$\delta[\tau - t_r(\tau)] = \delta[g(\tau)] = \frac{\delta(\tau - t_{r0})}{|g'(t_{r0})|}. \quad (8.35)$$

这里 t_{r0} 是 $g(\tau) = 0$ 的根，即 $t_{r0} = t - |\vec{x} - \vec{x}_p(t_{r0})|/c$ 。注意 $g(\tau) = 0$ 最多只有一个根，因为带电粒子的运动速度小于光速。而

$$g'(t_{r0}) = \left. \frac{\partial g(\tau)}{\partial \tau} \right|_{t_{r0}} = 1 + \frac{1}{c} \left. \frac{\partial R}{\partial \tau} \right|_{t_{r0}} = 1 - \frac{\vec{R}}{R} \cdot \frac{\vec{v}(t_{r0})}{c} = 1 - \vec{e}_R \cdot \vec{\beta}(t_{r0}). \quad (8.36)$$

这里 $\vec{e}_R \equiv \vec{R}/R$, $\vec{\beta}(t_{r0}) = \vec{v}(t_{r0})/c$ ，且我们用了 $R^2 = \vec{R} \cdot \vec{R}$ 及

$$\frac{\partial R}{\partial \tau} = \frac{\vec{R}}{R} \cdot \left(-\frac{\partial \vec{x}_p}{\partial \tau} \right). \quad (8.37)$$

利用上式完成对 τ 的积分，我们有

$$\phi(\vec{x}, t) = \frac{q}{4\pi\epsilon_0} \int d\tau \frac{\delta(\tau - t_{r0})}{R[1 - \vec{e}_R \cdot \vec{\beta}(\tau)]} = \frac{q}{4\pi\epsilon_0 R} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}(t_{r0})}. \quad (8.38)$$

类似的过程我们有

$$\vec{A}(\vec{x}, t) = \frac{\mu_0 q \vec{v}}{4\pi R} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}(t_{r0})} = \frac{\vec{v}}{c^2} \phi(\vec{x}, t). \quad (8.39)$$

方法二：因为推迟势的表达势仅依赖于粒子的速度和位置，所以我们可以通过建立瞬时惯性参考系的方法结合洛伦兹变换得到运动电荷的推迟势，就像匀速运动的电荷的场时所用方法一样。唯一的区别是这里我们建立的 \tilde{K} 系以速度 $\vec{v}(t_r)$ 运动，而该速度不是常数，因此， \tilde{K} 是一个瞬时惯性参考系。

首先，我们得到粒子静止参考系 \tilde{K} 里的势函数。因此 $\tilde{\phi}$ 与 $\tilde{\vec{A}}$ 是

$$\tilde{\phi} = \frac{q}{4\pi\epsilon_0 \tilde{R}}, \text{ and } \tilde{\vec{A}} = 0. \quad (8.40)$$

这里 $\tilde{R} = c(\tilde{t} - \tilde{t}_r)$ 是 \vec{x} 和 \vec{x}_p 在 \tilde{K} 中的距离。

现在我们从 \tilde{K} 通过洛伦兹变换得到 K 系里的 ϕ 和 \vec{A} 。因为 \tilde{K} 以速度 $\vec{v}(t_r)$ 相对于 K 运动，

$$\phi = \frac{\tilde{\phi}}{\sqrt{1 - v(t_r)^2/c^2}} = \frac{1}{\sqrt{1 - v(t_r)^2/c^2}} \frac{q}{4\pi\epsilon_0 \tilde{R}}, \quad (8.41)$$

$$\vec{A} = \frac{(\vec{v}(t_r)/c^2) \tilde{\phi}}{\sqrt{1 - v(t_r)^2/c^2}} = \frac{\vec{v}(t_r)}{c^2} \phi. \quad (8.42)$$

这里我们有在 ?? 一章里遇到的同样的问题： ϕ 和 \vec{A} 是用 \tilde{R} 表示的，而不是 K 系中的空间坐标 R ，因此我们还需要用 R 表达 \tilde{R} 。

利用洛伦兹变换，我们有

$$\tilde{R} = c(\tilde{t} - \tilde{t}_r) = \frac{c}{\sqrt{1 - v(t_r)^2/c^2}} \left[t - t_r - \frac{\vec{v}(t_r)}{c^2} \cdot (\vec{x} - \vec{x}_p) \right], \quad (8.43)$$

$$= \frac{1}{\sqrt{1 - v^2/c^2}} \left[c(t - t_r) - \frac{\vec{v}}{c} \cdot (\vec{x} - \vec{x}_p) \right], \quad (8.44)$$

$$= \gamma R(1 - \vec{e}_R \cdot \vec{\beta}). \quad (8.45)$$

注意，所有右边的物理量 ($\vec{v}, \vec{\beta}, \vec{R}, \gamma$) 都是在推迟时刻 t_r 时的值，即， $\gamma = 1/\sqrt{1 - v(t_r)^2/c^2}$, $\vec{\beta} = \vec{v}(t_r)/c$, $\vec{R} = \vec{x} - \vec{x}_p(t_r)$, 且 $R = c(t - t_r)$ 。

将方程 (8.45) 代入方程 (8.41) 与 (8.42)，得到点电荷的推迟势

$$\phi(\vec{x}, t) = \gamma \frac{q}{4\pi\epsilon_0 R(1 - \vec{e}_R \cdot \vec{\beta})} \frac{1}{\gamma} = \frac{q}{4\pi\epsilon_0 R(1 - \vec{e}_R \cdot \vec{\beta})}, \quad (8.46)$$

$$\vec{A}(\vec{x}, t) = \frac{1}{c} \frac{q\vec{\beta}}{4\pi\epsilon_0 R(1 - \vec{\beta} \cdot \vec{e}_R)}. \quad (8.47)$$

同样右边所有的量都是在推迟时刻 t_r 取值，即

$$\vec{\beta} = \vec{v}/c = \vec{v}(t_r)/c, \quad \text{and} \quad \vec{R} = \vec{x} - \vec{x}_p(t_r). \quad (8.48)$$

方程 (8.46) 与 (8.47) 即是点电荷的推迟势，也叫 Liénard-Wiechert 势。

8.2.1 点电荷的电磁场

从推迟势里，点电荷的 \vec{E} 和 \vec{B} 为

$$\vec{E} = -\nabla\phi - \frac{\partial\vec{A}}{\partial t}, \quad (8.49)$$

$$\vec{B} = \nabla \times \vec{A}. \quad (8.50)$$

这里的空间和时间偏导稍有复杂，因为这里的势函数对空间和时间的依赖关系即有显示的，也有隐式的。

我们来分析 ϕ 和 \vec{A} 如何依赖于 t 和 \vec{x} 。以 ϕ 为例，显然我们有 $\phi = \phi[\vec{x}, t_r(\vec{x}, t)]$ 。这里 ϕ 显示的依赖于 \vec{x} 因为 $\vec{R} = \vec{x} - \vec{x}_p(t_r)$ 。除此之外， ϕ 还通过推迟时刻 t_r 隐式的依赖于 \vec{x} 。对于时间 t , ϕ 只通过 t_r 隐式的依赖于 t 。矢量 \vec{A} 也有同样的时间空间依赖关系。

注意到获得 \vec{E} 和 \vec{B} 的式中， $\nabla = \nabla|_t$ 且 $\partial/\partial t = \partial/\partial t|_{\vec{x}}$ ，这里我们显示的写出求偏导时所固定的物理量。在厘清相应势函数的时间空间依赖关系后，我们有

$$\frac{\partial}{\partial t} = \frac{\partial t_r}{\partial t} \bigg|_{\vec{x}} \frac{\partial}{\partial t_r} \bigg|_{\vec{x}}, \quad (8.51)$$

$$\nabla = \nabla|_{t_r} + \nabla t_r|_t \cdot \frac{\partial}{\partial t_r}. \quad (8.52)$$

因此为了从 ϕ 与 \vec{A} 得到 \vec{E} 和 \vec{B} ，我们需要计算 $\partial t_r/\partial t$ 与 ∇t_r 。

首先我们计算 $\partial t_r/\partial t$ 。利用 $R = c(t - t_r)$,

$$\frac{\partial R}{\partial t} = c \left(1 - \frac{\partial t_r}{\partial t} \right) = \frac{\partial R}{\partial t_r} \frac{\partial t_r}{\partial t}. \quad (8.53)$$

这里的目的是从这个方程求解 $\partial t_r/\partial t$ 。利用式 (8.37)，我们有

$$\frac{\partial R}{\partial t_r} = -\vec{e}_R \cdot \vec{v}(t_r), \quad (8.54)$$

从而得到

$$1 - \frac{\partial t_r}{\partial t} = -\vec{e} \cdot \vec{\beta} \frac{\partial t_r}{\partial t}, \quad (8.55)$$

从中可得

$$\frac{\partial t_r}{\partial t} = \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}}. \quad (8.56)$$

因此

$$\frac{\partial}{\partial t} = \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} \cdot \frac{\partial}{\partial t_r}. \quad (8.57)$$

现在我们求 ∇t_r , 即 t_r 对 \vec{x} 的偏导, 这里是固定时间 t 的。再次利用 $R = c(t - t_r)$, 有

$$\nabla t_r = -\nabla R(\vec{x}, t_r)/c = -\frac{1}{c} \left(\nabla|_{t_r} R + \nabla t_r \cdot \frac{\partial R}{\partial t_r} \right) \quad (8.58)$$

现在我们仅需计算 $\nabla|_{t_r} R$ 。利用 $R = \sqrt{(\vec{x} - \vec{x}_p(t_r))^2}$, 可得

$$\nabla|_{t_r} R = \frac{\vec{R}}{R} \cdot \nabla \vec{x} = \vec{e}_R. \quad (8.59)$$

这里我们利用了 $\nabla \vec{x} = \mathbf{I}$ 及 $\vec{R} \cdot \mathbf{I} = \vec{R}$ (\mathbf{I} 为单位张量)。结合式 (8.37), 我们有

$$\nabla t_r = -\frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e} \cdot \vec{\beta}}. \quad (8.60)$$

因此,

$$\nabla = \nabla|_{t_r} - \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial}{\partial t_r}. \quad (8.61)$$

有了这些准备, 我们现在可以计算电场和磁场。

$$\vec{E} = -\nabla|_{t_r} \phi + \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \phi}{\partial t_r} - \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \vec{A}}{\partial t_r}, \quad (8.62)$$

$$\vec{B} = \left[\nabla|_{t_r} - \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial}{\partial t_r} \right] \times \vec{A}. \quad (8.63)$$

注意到 $\vec{A} = (\vec{v}/c^2)\phi$, 我们可以先将 \vec{E} 和 \vec{B} 化简为只用 ϕ 来表达。我们需要使用

$$\frac{\partial \vec{A}}{\partial t_r} = \frac{1}{c^2} \frac{\partial(\vec{v}\phi)}{\partial t_r} = \frac{\vec{\beta}}{c} \frac{\partial \phi}{\partial t_r} + \frac{\vec{a}}{c^2} \phi. \quad (8.64)$$

这里 $\vec{a} \equiv \partial \vec{v} / \partial t_r$ 为粒子在推迟时刻的加速度, 以及

$$\nabla|_{t_r} \times \vec{A} = \frac{1}{c^2} \nabla|_{t_r} \phi \times \vec{v}. \quad (8.65)$$

这里我们使用了 $\nabla|_{t_r} \times \vec{v} = 0$ 。因此,

$$\vec{E} = -\nabla|_{t_r} \phi + \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \phi}{\partial t_r} - \frac{1}{c} \frac{\vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \phi}{\partial t_r} - \frac{1}{c^2} \frac{\vec{a}\phi}{1 - \vec{e}_R \cdot \vec{\beta}} \quad (8.66)$$

$$= -\nabla|_{t_r} \phi + \frac{1}{c} \frac{\vec{e}_R - \vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \phi}{\partial t_r} - \frac{1}{c^2} \frac{\vec{a}\phi}{1 - \vec{e}_R \cdot \vec{\beta}}. \quad (8.67)$$

而磁场 \vec{B} 为

$$\vec{B} = \nabla|_{t_r} \times \vec{A} - \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \times \frac{\partial \vec{A}}{\partial t_r} \quad (8.68)$$

$$= \frac{1}{c^2} \nabla|_{t_r} \phi \times \vec{v} - \frac{1}{c} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} \times \left(\frac{\vec{\beta}}{c} \frac{\partial \phi}{\partial t_r} + \frac{\vec{a}}{c^2} \phi \right) \quad (8.69)$$

$$= \frac{1}{c^2} \nabla|_{t_r} \phi \times \vec{v} - \frac{1}{c^2} \frac{\vec{e}_R \times \vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{\partial \phi}{\partial t_r} - \frac{1}{c^3} \frac{\vec{e}_R \times \vec{a}}{1 - \vec{e}_R \cdot \vec{\beta}} \phi \quad (8.70)$$

注意到磁场的表达式与电场的表达式高度相似。

接下来我们求具体的电场表达式，注意到

$$\nabla|_{t_r} \phi = \frac{q}{4\pi\epsilon_0} \nabla|_{t_r} \left(\frac{1}{R} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} \right) = \frac{q}{4\pi\epsilon_0} \left[\left(\nabla|_{t_r} \frac{1}{R} \right) \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} + \frac{1}{R} \nabla|_{t_r} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} \right] \quad (8.71)$$

$$= \frac{q}{4\pi\epsilon_0} \left[-\frac{1}{R^2} \frac{\vec{e}_R}{1 - \vec{e}_R \cdot \vec{\beta}} + \frac{1}{R} \frac{\nabla|_{t_r} (\vec{e}_R \cdot \vec{\beta})}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right] \quad (8.72)$$

这里我们需要使用

$$\nabla|_{t_r} (\vec{e}_R \cdot \vec{\beta}) = \left[\frac{1}{R} \nabla|_{t_r} \vec{R} - \frac{1}{R^2} (\nabla|_{t_r} R) \vec{R} \right] \cdot \vec{\beta} \quad (8.73)$$

$$= \left[\frac{1}{R} - \frac{\vec{e}_R}{R} \vec{e}_R \right] \cdot \vec{\beta} = \frac{\vec{\beta} - \vec{e}_R (\vec{e}_R \cdot \vec{\beta})}{R}. \quad (8.74)$$

代入上式，得

$$\nabla|_{t_r} \phi = \frac{q}{4\pi\epsilon_0 R^2} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} \left[-\vec{e}_R + \frac{\vec{\beta} - \vec{e}_R (\vec{e}_R \cdot \vec{\beta})}{1 - \vec{e}_R \cdot \vec{\beta}} \right] = \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{\beta} - \vec{e}_R}{(1 - \vec{e}_R \cdot \vec{\beta})^2}. \quad (8.75)$$

从这个表达式里，有一个有趣的结论，即

$$\nabla|_{t_r} \phi \times \vec{\beta} = \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{\beta} - \vec{e}_R}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \times \vec{\beta} = \frac{q}{4\pi\epsilon_0 R^2} \frac{-\vec{e}_R \times \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} = \frac{q}{4\pi\epsilon_0 R^2} \frac{-\vec{e}_R \times (\vec{\beta} - \vec{e}_R)}{(1 - \vec{e}_R \cdot \vec{\beta})^2} = -\vec{e}_R \times \nabla|_{t_r} \phi. \quad (8.76)$$

因此，对比 \vec{B} 与 \vec{E} 的表达式，我们立刻得到

$$\vec{B} = \vec{e}_R \times \vec{E}/c. \quad (8.77)$$

这里我们在 \vec{B} 的式子里，还运用了 $-\vec{e}_R \times \vec{\beta} = \vec{e}_R \times (\vec{e}_R - \vec{\beta})$ 。

下面我们来求 $\partial\phi/\partial t_r$ ，

$$\frac{\partial \phi}{\partial t_r} = \frac{q}{4\pi\epsilon_0} \left[\frac{\partial}{\partial t_r} \left(\frac{1}{R} \right) \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} + \frac{1}{R} \frac{\partial (\vec{e}_R \cdot \vec{\beta}) / \partial t_r}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right] \quad (8.78)$$

$$= \frac{q}{4\pi\epsilon_0} \left[\frac{1}{R^2} \frac{\vec{e}_R \cdot \vec{v}}{1 - \vec{e}_R \cdot \vec{\beta}} + \frac{1}{R} \frac{1}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \left(\frac{\partial \vec{e}_R}{\partial t_r} \cdot \vec{\beta} + \vec{e}_R \cdot \frac{\partial \vec{\beta}}{\partial t_r} \right) \right] \quad (8.79)$$

$$= \frac{q}{4\pi\epsilon_0 R} \left[\frac{\vec{e}_R \cdot \vec{v}}{R} \frac{1}{1 - \vec{e}_R \cdot \vec{\beta}} + \frac{1 - \vec{v} \cdot \vec{\beta} + (\vec{e}_R \cdot \vec{v})(\vec{e}_R \cdot \vec{\beta})}{R (1 - \vec{e}_R \cdot \vec{\beta})^2} + \frac{1}{c} \frac{\vec{e}_R \cdot \vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right]. \quad (8.80)$$

这里我们使用了

$$\frac{\partial \vec{e}_R}{\partial t_r} = \frac{1}{R} \frac{\partial \vec{R}}{\partial t_r} - \frac{\vec{R}}{R^2} \frac{\partial R}{\partial t_r} = -\frac{\vec{v}}{R} + \frac{\vec{e}_R \cdot \vec{v}}{R} \vec{e}_R \quad (8.81)$$

上述 $\partial\phi/\partial t_r$ 式可以继续化简为

$$\frac{\partial\phi}{\partial t_r} = \frac{q}{4\pi\epsilon_0 R} \left[\frac{\vec{v}}{R} \cdot \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} + \frac{1}{c} \frac{\vec{e}_R \cdot \vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right] \quad (8.82)$$

$$= \frac{q\vec{v}}{4\pi\epsilon_0 R^2} \cdot \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} + \frac{1}{c} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R \cdot \vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \quad (8.83)$$

现在将 $\nabla|_{t_r}\phi$, $\partial\phi/\partial t_r$ 与 ϕ 的表达式, 代入 \vec{E} 。

$$\vec{E} = -\frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{\beta} - \vec{e}_R}{(1 - \vec{e}_R \cdot \vec{\beta})^2} + \frac{1}{c} \frac{\vec{e}_R - \vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \left[\frac{q}{4\pi\epsilon_0 R^2} \vec{v} \cdot \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right] \quad (8.84)$$

$$+ \frac{1}{c} \frac{\vec{e}_R - \vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \left[\frac{1}{c} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R \cdot \vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \right] - \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2}. \quad (8.85)$$

注意到结果表达式里的相关项可以分成两类：一类仅与速度 \vec{v} 有关，一类还与加速度 \vec{a} 有关。我们有

$$\vec{E} = \vec{E}_{\vec{v}} + \vec{E}_{\vec{a}}. \quad (8.86)$$

这里与速度有关的电场（速度场或广义库伦场）

$$\vec{E}_{\vec{v}} = \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} + \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{1 - \vec{e}_R \cdot \vec{\beta}} \frac{(\vec{e}_R - \vec{\beta}) \cdot \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \quad (8.87)$$

$$= \frac{q}{4\pi\epsilon_0 R^2} \frac{1}{(1 - \vec{e}_R \cdot \vec{\beta})^3} \left[(\vec{e}_R - \vec{\beta})(1 - \vec{e}_R \cdot \vec{\beta}) + (\vec{e}_R - \vec{\beta})(\vec{e}_R \cdot \vec{\beta} - \beta^2) \right] \quad (8.88)$$

$$= \frac{q}{4\pi\epsilon_0 R^2} \frac{(\vec{e}_R - \vec{\beta})(1 - \beta^2)}{(1 - \vec{e}_R \cdot \vec{\beta})^3} = \frac{1}{\gamma^2} \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^3}. \quad (8.89)$$

而与加速度有关的部分为（加速度场或辐射场）

$$\vec{E}_{\vec{a}} = \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^3} (\vec{e}_R \cdot \vec{a}) - \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{a}}{(1 - \vec{e}_R \cdot \vec{\beta})^2} \quad (8.90)$$

$$= \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{1}{(1 - \vec{e}_R \cdot \vec{\beta})^3} \left[(\vec{e}_R - \vec{\beta})(\vec{e}_R \cdot \vec{a}) - \vec{a}(1 - \vec{e}_R \cdot \vec{\beta}) \right] \quad (8.91)$$

$$= \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{1}{(1 - \vec{e}_R \cdot \vec{\beta})^3} \left[(\vec{a} \times \vec{e}_R) \times \vec{e}_R + (\vec{\beta} \times \vec{a}) \times \vec{e}_R \right] \quad (8.92)$$

$$= \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R \times [(\vec{e}_R - \vec{\beta}) \times \vec{a}]}{(1 - \vec{e}_R \cdot \vec{\beta})^3}. \quad (8.93)$$

这里我们使用了

$$\vec{e}_R(\vec{e}_R \cdot \vec{a}) - \vec{a} = (\vec{a} \times \vec{e}_R) \times \vec{e}_R, \quad (8.94)$$

$$\vec{a}(\vec{e}_R \cdot \vec{\beta}) - \vec{\beta}(\vec{e}_R \cdot \vec{a}) = (\vec{\beta} \times \vec{a}) \times \vec{e}_R. \quad (8.95)$$

综合上述计算结果，我们可得任意形式运动的带电粒子，其电场与磁场具有如下的形式

$$\vec{E} = \underbrace{\frac{1}{\gamma^2} \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^3}}_{\vec{E}_{\vec{v}}: \text{速度场}} + \underbrace{\frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R \times [(\vec{e}_R - \vec{\beta}) \times \vec{a}]}{(1 - \vec{e}_R \cdot \vec{\beta})^3}}_{\vec{E}_{\vec{a}}: \text{辐射场}}, \quad (8.96)$$

$$\vec{B} = \vec{e}_R \times \vec{E}/c. \quad (8.97)$$

例题 8.1 在电磁场一章，我们讨论过匀速直线运动的带电粒子的场，其方向为从粒子当前时刻的位置指向场点。然后，我们从推迟势出发的电场，与粒子相关的物理量都是在推迟时刻的值，似乎与当前时刻无关。请解释。

解 对于匀速运动粒子，其电场仅含速度场，即

$$\vec{E} = \frac{1}{\gamma^2} \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^3}. \quad (8.98)$$

此时方程右边所有的一切粒子相关物理量都是在推迟时刻定义的， $t_r = t - R(t_r)/c$ 。然而因为粒子是匀速运动，所以 $\vec{v}(t_r) = \vec{v}$ 为常数，且

$$\vec{R}(t) = \vec{R}(t_r) + \vec{v}(t - t_r) = \vec{R}(t_r) - \vec{v}R(t_r)/c = \vec{R}(t_r) - \vec{\beta}R(t_r) = [\vec{e}_R(t_r) - \vec{\beta}]R(t_r). \quad (8.99)$$

因此电场表达式中的 $\vec{e}_R(t_r) - \vec{\beta} = \vec{R}(t)/R(t_r)$ ，其方向为从粒子当前时刻 (t) 的位置指向场点。

例题 8.2 比较非相对论粒子的辐射场与速度场的大小关系。

解 对于非相对论粒子，其速度场和辐射场可分别简化为

$$\vec{E}_{\vec{v}} = \frac{1}{\gamma^2} \frac{q}{4\pi\epsilon_0 R^2} \frac{\vec{e}_R - \vec{\beta}}{(1 - \vec{e}_R \cdot \vec{\beta})^3} \approx \frac{q}{4\pi\epsilon_0 R^2} \vec{e}_R, \quad (8.100)$$

$$\vec{E}_{\vec{a}} = \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \frac{\vec{e}_R \times [(\vec{e}_R - \vec{\beta}) \times \vec{a}]}{(1 - \vec{e}_R \cdot \vec{\beta})^3} \approx \frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \vec{e}_R \times (\vec{e}_R \times \vec{a}) = -\frac{1}{c^2} \frac{q}{4\pi\epsilon_0 R} \vec{a}_{\perp}. \quad (8.101)$$

这里运用了 $\vec{a}_{\perp} = (\vec{e}_R \times \vec{a}) \times \vec{e}_R$ 。

这两类场之类的量级关系为 $E_{\vec{v}}/E_{\vec{a}} \sim c^2/Ra$ 。这里 c^2/a 为一特征空间尺度。因此对于 $R < c^2/a$ 时，速度场占主导；超过此距离时，辐射场占主导。且我们总有 $R \rightarrow \infty$ 时， $E_{\vec{v}} \ll E_{\vec{a}}$ 。

从上述辐射场的表达式，我们也可以得到非相对论粒子，其辐射电场的大小在给定距离 R 时，正比于 \vec{a} 在垂直于 \vec{e}_R 方向上的投影， \vec{a}_{\perp} 。如果我们定义 $\theta \equiv \langle \vec{e}_R, \vec{a} \rangle$ ，则

$$|\vec{E}_{\vec{a}}| \propto |\vec{a}_{\perp}| = a \sin \theta. \quad (8.102)$$

因此，如果粒子的加速度是沿着 \vec{e}_R 方向时，辐射电场为 0。

第九章 电磁波的辐射

这里辐射的意思是指的是通过电磁场的形式将能量传至无穷远处。因此在这一章里我们主要计算运动电荷所辐射的能量与功率。

9.1 点电荷的辐射

我们从比较简单的点电荷开始，计算一个点电荷在单位时间内因为运动向外辐射的功率。显然，我们需要用到点电荷的电场和磁场：

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0\gamma^2} \frac{\mathbf{R} - R\boldsymbol{\beta}}{S^3} + \frac{q}{4\pi\epsilon_0 c} \frac{\mathbf{R} \times [(\mathbf{R} - R\boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{S^3}, \quad (9.1)$$

$$\mathbf{B} = \mathbf{e}_R \times \mathbf{E}/c. \quad (9.2)$$

我们已经知道上式 (9.1) 右边的第一项为速度场，该场正比于 R^{-2} 。因此，它不能将能量传至远处，即不能辐射。第二项正比于 R^{-1} ，可以将能量传至无穷远。因此，我们在讨论电荷的辐射的时候，仅仅讨论第二项，即辐射场的贡献。利用 $S = R - \boldsymbol{\beta} \cdot \mathbf{R}$,

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 c} \frac{\mathbf{R} \times [(\mathbf{R} - R\boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(R - \boldsymbol{\beta} \cdot \mathbf{R})^3} \quad (9.3)$$

$$= \frac{q}{4\pi\epsilon_0 c} \frac{\mathbf{e}_R \times [(\mathbf{e}_R - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{e}_R)^3 R}, \quad (9.4)$$

$$\mathbf{B} = \mathbf{e}_R \times \mathbf{E}/c. \quad (9.5)$$

注意到 $\mathbf{E} \cdot \mathbf{e}_R = 0$, 坡印廷矢量为

$$\mathbf{S}(\mathbf{r}, t) = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} = \frac{1}{\mu_0 c} E^2 \mathbf{e}_R, \quad (9.6)$$

$$= \frac{1}{\mu_0 c} \left(\frac{q}{4\pi\epsilon_0 c} \right)^2 \frac{|\mathbf{e}_R \times [(\mathbf{e}_R - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]|^2}{(1 - \boldsymbol{\beta} \cdot \mathbf{e}_R)^6 R^2} \mathbf{e}_R \quad (9.7)$$

坡印廷矢量的含义为单位时间内能过法向为 \mathbf{S} 方向的单位面元的能量。因此，从 $t_1(t'_1)$ 到 $t_2(t'_2)$ 通过方向为 \mathbf{e}_R 的单位面元所辐射出的能量可以表达为

$$\Delta E = \mathbf{S} \cdot \mathbf{e}_R \Delta t. \quad (9.8)$$

而从大小为 $R^2 d\Omega$ 大小的面元所辐射的总能量为

$$\Delta \mathcal{E} = \mathbf{S} \cdot \mathbf{e}_R \Delta t R^2 d\Omega. \quad (9.9)$$

这里 $d\Omega \equiv \sin \theta d\theta d\phi$ 为立体角。注意到，该能量是粒子在 t'_1 至 t'_2 之间辐射出的，因此粒子的通过该面元的辐射功率为

$$\Delta P \equiv \frac{\Delta \mathcal{E}}{\Delta t'} = \mathbf{S} \cdot \mathbf{e}_R \frac{\Delta t}{\Delta t'} R^2 d\Omega. \quad (9.10)$$

这里 $\Delta t' = t'_2 - t'_1$ 。将差分换成微分，则上式中的辐射功率可以写成

$$dP = \mathbf{S} \cdot \mathbf{e}_R \frac{\partial t}{\partial t'} R^2 d\Omega. \quad (9.11)$$

这里的偏导是固定空间坐标而进行的。习惯上，我们会用通过 \mathbf{e}_R 方向单位立体角粒子所辐射出的功率，即

$$\frac{dP}{d\Omega} = \mathcal{S} \cdot \mathbf{e}_R R^2 \frac{\partial t}{\partial t'}. \quad (9.12)$$

利用

$$\frac{\partial t}{\partial t'} = 1 - \mathbf{e}_R \cdot \boldsymbol{\beta}, \quad (9.13)$$

则

$$\frac{dP}{d\Omega} = \mathcal{S} \cdot \mathbf{e}_R R^2 (1 - \mathbf{e}_R \cdot \boldsymbol{\beta}). \quad (9.14)$$

将式 (9.7) 代入方程 (9.14)，则

$$\frac{dP}{d\Omega} = \frac{q^2 \mu_0 c}{16\pi^2} \frac{\left| \mathbf{e}_R \times [(\mathbf{e}_R - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}] \right|^2}{(1 - \boldsymbol{\beta} \cdot \mathbf{e}_R)^5}. \quad (9.15)$$

这即是任意形式运动电荷的辐射功率的一般表达式。

9.1.1 非相对论粒子的辐射功率

首先，我们计算非相对论粒子的辐射功率，即 $\beta \ll 1$ ，则 $1 - \boldsymbol{\beta} \cdot \mathbf{e}_R \approx 1$ 。此时，方程 (9.15) 变成

$$\frac{dP}{d\Omega} = \frac{q^2 \mu_0 c}{16\pi^2} \left| \mathbf{e}_R \times (\mathbf{e}_R \times \dot{\boldsymbol{\beta}}) \right|^2 \quad (9.16)$$

$$= \frac{q^2 \mu_0 c}{16\pi^2} \left| (\mathbf{e}_R \cdot \dot{\boldsymbol{\beta}}) \mathbf{e}_R - \dot{\boldsymbol{\beta}} \right|^2 \quad (9.17)$$

$$= \frac{q^2 \mu_0 c}{16\pi^2} \left| \dot{\boldsymbol{\beta}}_{\perp} \right|^2 \quad (9.18)$$

$$= \frac{q^2 \mu_0}{16\pi^2 c} \dot{v}^2 \sin^2 \Theta, \quad (9.19)$$

这里 $\Theta \equiv \langle \dot{\mathbf{v}}, \mathbf{e}_R \rangle = \langle \mathbf{a}, \mathbf{e}_R \rangle$ 。进一步注意到单个点电荷的电偶极矩为 $\mathbf{p} = \sum q \mathbf{r}' = q \mathbf{r}'$ ，则 $\dot{\mathbf{p}} = q \mathbf{v}$ 且 $\ddot{\mathbf{p}} = q \mathbf{a}$ 。因此式 (9.19) 可以利用 \mathbf{p} 写成

$$\frac{dP}{d\Omega} = \frac{\mu_0 |\ddot{\mathbf{p}}|^2}{16\pi^2 c} \sin^2 \Theta. \quad (9.20)$$

非相对论粒子的总辐射功率为

$$P = \int dP = \int_0^{2\pi} d\phi \int_0^{\pi} d\Theta \frac{\mu_0 |\ddot{\mathbf{p}}|^2}{16\pi^2 c} \sin^3 \Theta = \frac{\mu_0 |\ddot{\mathbf{p}}|^2}{6\pi c} = \frac{\mu_0 q^2 a^2}{6\pi c}. \quad (9.21)$$

这个公式称为拉莫公式 (Larmor formula)。从该式可以看出，非相对论点电荷的辐射功率正比于其加速度的平方。

9.1.2 相对论粒子的辐射

对于 $\beta \sim 1$ 的相对论粒子，我们无法忽略方程 (9.15) 中的 β 或者 $\boldsymbol{\beta}$ ，此时计算会更加复杂。这里我们首先讨论两种特殊情况： $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$ 与 $\boldsymbol{\beta} \perp \dot{\boldsymbol{\beta}}$ ，而将一般情况下的辐射功率计算留在作业里。

首先我们计算 $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$ 时的辐射功率。如果 $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$ ，则方程 (9.15) 变成

$$\frac{dP}{d\Omega} = \frac{q^2 \mu_0 c}{16\pi^2} \frac{\left| \mathbf{e}_R \times [\mathbf{e}_R \times \dot{\boldsymbol{\beta}}] \right|^2}{(1 - \boldsymbol{\beta} \cdot \mathbf{e}_R)^5}. \quad (9.22)$$

定义 $\theta = \langle \boldsymbol{\beta}, \mathbf{e}_R \rangle$. 当 $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$, $\langle \dot{\boldsymbol{\beta}}, \mathbf{e}_R \rangle = \theta$.

$$\frac{dP}{d\Omega} = \frac{q^2 \mu_0 c}{16\pi^2} \frac{\dot{\boldsymbol{\beta}}^2 \sin^2 \theta}{(1 - \beta \cos \theta)^5}, \quad (9.23)$$

此时辐射功率的角分布很有意思. 如果 $\theta = 0$, 则 $dP/d\Omega = 0$. 然后, 对 $\beta \rightarrow 1$ 的超相对论粒子,

$$1 - \beta \cos \theta \sim 1 - \cos \theta \rightarrow 0, \quad \text{as } \theta \rightarrow 0. \quad (9.24)$$

因此 $dP/d\Omega$ 在 $\theta \sim 0$ 处 (但不是在 $\theta = 0$) 非常大. 注意到辐射功率的角分布与加速度的符号无关, 即无论该粒子是加速还是减速. 当一个高速粒子作减速运动时 (例如撞到另一个粒子上), 会向外辐射出能量. 此时这种辐射被称为轫致辐射, 刹车辐射或制动辐射 (bremsstrahlung).

当 $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$ 时, 辐射总功率为

$$P_{\parallel} = \int \frac{q^2 \mu_0 c}{16\pi^2} \frac{\dot{\boldsymbol{\beta}}^2 \sin^2 \theta}{(1 - \beta \cos \theta)^5} d\Omega \quad (9.25)$$

$$= \frac{q^2 \mu_0 c \dot{\boldsymbol{\beta}}^2}{16\pi^2} \int_0^{2\pi} d\phi \int_0^{\pi} \frac{\sin^2 \theta}{(1 - \beta \cos \theta)^5} d\theta \quad (9.26)$$

$$= \frac{q^2 \mu_0 c \dot{\boldsymbol{\beta}}^2}{16\pi^2} 2\pi \frac{4}{3(1 - \beta^2)^3} \quad (9.27)$$

$$= \frac{q^2 \mu_0 c \dot{\boldsymbol{\beta}}^2}{6\pi(1 - \beta^2)^3} \quad (9.28)$$

$$= \frac{q^2 \mu_0}{6\pi c} \gamma^6 \dot{\boldsymbol{v}}^2 = \frac{q^2 \mu_0}{6\pi c} \gamma^6 a^2. \quad (9.29)$$

我们也可以利用粒子的受力来表示其辐射总功率, 这在有些问题里会非常方便. 对 $\mathbf{v} \parallel \dot{\mathbf{v}}$, 我们从动量 \mathbf{p} 的时间变化率里得到,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\gamma m \mathbf{v}) = \dot{\gamma} m \mathbf{v} + \gamma m \dot{\mathbf{v}}. \quad (9.30)$$

注意到 $\beta^2 = (\gamma^2 - 1)/\gamma^2$, 且对 $\mathbf{v} \parallel \dot{\mathbf{v}}$ 而言有,

$$\dot{\gamma} m \mathbf{v} = \gamma^3 \beta^2 m \dot{\mathbf{v}} = (\gamma^3 - \gamma) m \dot{\mathbf{v}}. \quad (9.31)$$

将方程 (9.31) 代入式 (9.30) 我们有

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \gamma^3 m \dot{\mathbf{v}}. \quad (9.32)$$

因此利用力 \mathbf{F} , $\boldsymbol{\beta} \parallel \dot{\boldsymbol{\beta}}$ 时的总辐射功率为

$$P_{\parallel} = \frac{q^2 \mu_0}{6\pi c m^2} \left(\frac{d\mathbf{p}}{dt} \right)^2. \quad (9.33)$$

因此, 在给定外力下, 此时的辐射功率与粒子的能量即 γ 无关.

第二种特殊情况是 $\boldsymbol{\beta} \perp \dot{\boldsymbol{\beta}}$. 这种情况比上一种稍微复杂些. 我们首先建立一个坐标系, $\mathbf{v} = v \mathbf{e}_z$, 且 $\dot{\mathbf{v}} = \dot{v} \mathbf{e}_x$, 则

$$\dot{\boldsymbol{\beta}} \cdot \mathbf{e}_R = \dot{\beta} \sin \theta \cos \phi, \quad (9.34)$$

$$\boldsymbol{\beta} \cdot \mathbf{e}_R = \beta \cos \theta. \quad (9.35)$$

这里 (R, θ, ϕ) 构成一个球坐标系. 利用这两个方程

$$\begin{aligned} & \mathbf{e}_R \times [(\mathbf{e}_R - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}] \\ &= (\mathbf{e}_R \cdot \dot{\boldsymbol{\beta}})(\mathbf{e}_R - \boldsymbol{\beta}) - [(\mathbf{e}_R - \boldsymbol{\beta}) \cdot \mathbf{e}_R] \dot{\boldsymbol{\beta}} \end{aligned} \quad (9.36)$$

$$= \dot{\beta} \sin \theta \cos \phi (\mathbf{e}_R - \boldsymbol{\beta}) - (1 - \beta \cos \theta) \dot{\boldsymbol{\beta}}. \quad (9.37)$$

因此,

$$\begin{aligned} & |e_R \times [(e_R - \beta) \times \dot{\beta}]|^2 \\ &= (\dot{\beta} \sin \theta \cos \phi)^2 (e_R - \beta) \cdot (e_R - \beta) + (1 - \beta \cos \theta)^2 \dot{\beta} \cdot \dot{\beta} \\ &\quad - 2\dot{\beta} \sin \theta \cos \phi (1 - \beta \cos \theta) (e_R - \beta) \cdot \dot{\beta} \end{aligned} \quad (9.38)$$

$$= \dot{\beta}^2 [(1 - \beta \cos \theta)^2 + (\sin \theta \cos \phi)^2 (\beta^2 - 1)]. \quad (9.39)$$

将 (9.39) 代入 (9.15) 得

$$\boxed{\frac{dP}{d\Omega} = \frac{q^2 \mu_0 c}{16\pi^2} \frac{\dot{\beta}^2 [(1 - \beta \cos \theta)^2 + (\sin \theta \cos \phi)^2 (\beta^2 - 1)]}{(1 - \beta \cos \theta)^5}}. \quad (9.40)$$

对式 (9.40) 积分得到总辐射功率

$$P_{\perp} = \frac{q^2 \mu_0 c \dot{\beta}^2}{16\pi^2} 2\pi \frac{4}{3(1 - \beta^2)^2} = \frac{q^2 \mu_0}{6\pi c} \dot{v}^2 \gamma^4. \quad (9.41)$$

因为当 $\mathbf{v} \perp \dot{\mathbf{v}}$ 时, 力 $d\mathbf{p}/dt$ 为

$$\frac{d\mathbf{p}}{dt} = \gamma m \dot{\mathbf{v}} \Rightarrow \left(\frac{d\mathbf{p}}{dt} \right)^2 = \gamma^2 m^2 \dot{v}^2, \quad (9.42)$$

这里我们利用了 γ 是一个常数。总辐射功率用力来表达, 即为

$$\boxed{P_{\perp} = \frac{q^2 \mu_0}{6\pi c} \frac{\gamma^2}{m^2} \left(\frac{d\mathbf{p}}{dt} \right)^2}. \quad (9.43)$$

因此对一个给定的外力, 辐射功率正比于 γ^2 。能量越高辐射越大。

这里关于 P_{\parallel} 和 P_{\perp} 的结论对设计加速器很重要。实际应用中有两种加速器: 线性加速器 ($\mathbf{v} \parallel \dot{\mathbf{v}}$) 和回旋加速器 ($\mathbf{v} \perp \dot{\mathbf{v}}$)。对一个给定外力, $P_{\perp} = P_{\parallel} \gamma^2$ 。因此, 随着能量增加, 回旋加速器里的能量损失要比线性加速器里严重得多。所以原则上, 利用线性加速器更容易将粒子加速到很高的能量。另一方面, 回旋加速器的一个重要应用是获取辐射如 X 光, 从而利用这些辐射来进行其它的实验, 如 X 光衍射实验来确定物质的性质。

利用上述两种特殊情况下的辐射, 我们可以计算任意形式运动的带电粒子的辐射总功率。注意到 \mathbf{E} 和 \mathbf{B} 满足线性叠加原理, 因此 $\mathbf{E} = \mathbf{E}(\dot{\beta}_{\parallel}) + \mathbf{E}(\dot{\beta}_{\perp})$, 这里 \parallel 与 \perp 都是相对于 \mathbf{v} 定义的。电场为

$$\mathbf{E} \propto e_R \times [(e_R - \beta) \times \dot{\beta}] \quad (9.44)$$

$$= e_R \times [(e_R - \beta) \times (\dot{\beta}_{\parallel} + \dot{\beta}_{\perp})] \quad (9.45)$$

$$= e_R \times [(e_R - \beta) \times \dot{\beta}_{\parallel} + (e_R - \beta) \times \dot{\beta}_{\perp}]. \quad (9.46)$$

因此, 如果我们写成 $\mathbf{E} = \mathbf{E}_{\parallel} + \mathbf{E}_{\perp}$, 则

$$E^2 = E_{\parallel}^2 + E_{\perp}^2 + 2\mathbf{E}_{\parallel} \cdot \mathbf{E}_{\perp}. \quad (9.47)$$

所以方程 (9.15) 变成

$$\frac{dP}{d\Omega} = \left(\frac{dP}{d\Omega} \right)_{\parallel} + \left(\frac{dP}{d\Omega} \right)_{\perp} + \left(\frac{dP}{d\Omega} \right)_{\text{cross}}. \quad (9.48)$$

从式 (9.46) 得到,

$$\left(\frac{dP}{d\Omega} \right)_{\text{cross}} \propto 2\{e_R \times [(e_R - \beta) \times \dot{\beta}_{\parallel}]\} \cdot \{e_R \times [(e_R - \beta) \times \dot{\beta}_{\perp}]\}. \quad (9.49)$$

总辐射功率为

$$P = \int \left[\left(\frac{dP}{d\Omega} \right)_{\parallel} + \left(\frac{dP}{d\Omega} \right)_{\perp} + \left(\frac{dP}{d\Omega} \right)_{\text{cross}} \right] d\Omega \quad (9.50)$$

$$= P_{\parallel} + P_{\perp} + \int \left(\frac{dP}{d\Omega} \right)_{\text{cross}} d\Omega. \quad (9.51)$$

作为练习，你可以证明

$$\int \left(\frac{dP}{d\Omega} \right)_{\text{cross}} d\Omega = 0. \quad (9.52)$$

因此利用方程 (9.29) 与 (9.41),

$$P = P_{\parallel} + P_{\perp} = \frac{q^2 \mu_0}{6\pi c} \gamma^6 a_{\parallel}^2 + \frac{q^2 \mu_0}{6\pi c} \gamma^4 a_{\perp}^2 = \frac{q^2 \mu_0}{6\pi c} (\gamma^6 a_{\parallel}^2 + \gamma^4 a_{\perp}^2). \quad (9.53)$$

这即是任意运动的相对论粒子所辐射出的总功率，也称为李纳德公式 (Lienard formula)。显然，在 $\gamma \rightarrow 1$ 时，李纳德公式回复到拉莫公式。

9.2 辐射阻尼

在这一节里，我们讨论辐射对于粒子的效应。因为辐射会将能量以场的形式损失掉，所以粒子会减少能量。这一过程被称为辐射阻尼 (radiation damping)。显然，对于一个给定的外力，在其它条件相同的情况下，一个带电的粒子会比中性粒子要加速的慢。等效来讲，相当于辐射过程在带电粒子上施加了一个额外的力， \mathbf{F}_{rad} 。这个力称为辐射阻尼力。本节的目的是计算这个辐射阻尼力的大小。显然，同自能一样，这种粒子自己给自己施加的力 (self-force) 在现代物理学里也没有被很完整地理解。

为了简化讨论，我们仅考虑非相对论粒子，其单位时间内的辐射能量由拉莫公式给出。从能量守恒角度出发，也许你会得到粒子在单位时间内的能量损失为该阻尼力所做的功，即

$$\mathbf{F}_{\text{rad}} \cdot \mathbf{v} = -P = -\frac{\mu_0 q^2 a^2}{6\pi c}. \quad (9.54)$$

但这个公式实际上是不对的。但我们计算辐射功率的时候，我们忽略了速度场的能量，因为这个能量不能被传至无穷远。然而，当粒子加速或减速时，速度场也会发生变化，此时在速度场里的总能量也会增加或减少。因此，严格的能量守恒需要将速度场的能量变化考虑在内。即，

$$\mathbf{F}_{\text{rad}} \cdot \mathbf{v} = -\frac{\mu_0 q^2 a^2}{6\pi c} - \frac{d\varepsilon_v}{dt}. \quad (9.55)$$

这里 ε_v 指的是速度场的能量。但我们可以考虑一个简单的情境，即在一个时间间隔内，粒子的初始和最终状态是相同的，这样存储在速度场里的能量是完全相同的。这样对这一段间隔而言，粒子唯一的能量变化在于通过辐射场的形式所损失的能量。这样尽管方程 (9.54) 在每一个时刻是不正确的，但是如果我们对这一个时间段作一个积分的话，结果是对的：

$$\int_{t_1}^{t_2} \mathbf{F}_{\text{rad}} \cdot \mathbf{v} dt = -\frac{\mu_0 q^2}{6\pi c} \int_{t_1}^{t_2} a^2 dt. \quad (9.56)$$

我们如果对右边作分步积分，则有

$$\int_{t_1}^{t_2} a^2 dt = \int_{t_1}^{t_2} \frac{d\mathbf{v}}{dt} \cdot \frac{d\mathbf{v}}{dt} dt = \mathbf{v} \cdot \frac{d\mathbf{v}}{dt} \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d^2\mathbf{v}}{dt^2} \cdot \mathbf{v} dt. \quad (9.57)$$

因为系统的状态在 t_1 和 t_2 时刻是相同的，

$$\int_{t_1}^{t_2} a^2 dt = - \int_{t_1}^{t_2} \frac{d^2\mathbf{v}}{dt^2} \cdot \mathbf{v} dt = - \int_{t_1}^{t_2} \dot{\mathbf{a}} \cdot \mathbf{v} dt. \quad (9.58)$$

将该式代入方程 (9.57) 有

$$\int_{t_1}^{t_2} \left(\mathbf{F}_{\text{rad}} - \frac{\mu_0 q^2}{6\pi c} \dot{\mathbf{a}} \right) \cdot \mathbf{v} dt = 0. \quad (9.59)$$

显然方程 (9.59) 会被满足, 如果

$$\mathbf{F}_{\text{rad}} = \frac{\mu_0 q^2}{6\pi c} \dot{\mathbf{a}}. \quad (9.60)$$

这种辐射阻尼力的表达式称为辐射阻尼力的 **Abraham-Lorentz** 公式。尽管如此, 我们应该看到, 这个力的形式是猜出来的, 且我们可以在方程 (9.59) 加上任一个垂直于 \mathbf{v} 的项而不改变结果。显然方程 (9.60) 是一种时间平均的辐射阻尼力, 且是目前已知的最简单的表达式。在本书里, 我们将用该式来表达带电粒子所受到的辐射阻尼力。

尽管如此, 如前所述, 辐射阻尼力这种自我作用力会导致一些奇怪的结论。为展示这一点, 我们首先写出粒子在考虑了辐射阻尼力后的运动方程:

$$m\mathbf{a} = \frac{\mu_0 q^2}{6\pi c} \dot{\mathbf{a}} + \mathbf{F}_0, \quad (9.61)$$

这里 \mathbf{F}_0 代表其它的任何力。为简化起见, 我们仅考虑一个一维问题。方程 (9.61) 也可以被写为

$$a = \tau \dot{a} + F_0/m, \quad (9.62)$$

这里

$$\tau = \frac{\mu_0 q^2}{6\pi mc} = \frac{2r_q}{3c}, \quad (9.63)$$

且 $r_q = q^2/4\pi\epsilon_0 mc^2$ 是带电粒子的经典半径。对一个电子而言, $\tau \sim 6 \times 10^{-24} \text{ s}$ 。注意到当我们考虑了辐射阻尼力后, 加速度是一个时间的连续性函数, 就如同速度和空间坐标一样。为证明这一点, 我们首先将方程 (9.62) 从 $t - \epsilon$ 到 $t + \epsilon$ 积分,

$$\int_{t-\epsilon}^{t+\epsilon} a dt = \int_{t-\epsilon}^{t+\epsilon} \tau \dot{a} dt + \int_{t-\epsilon}^{t+\epsilon} F_0/m dt, \quad (9.64)$$

这与

$$v(t + \epsilon) - v(t - \epsilon) = \tau[a(t + \epsilon) - a(t - \epsilon)] + \int_{t-\epsilon}^{t+\epsilon} F_0/m dt \quad (9.65)$$

相同。注意到 v 是时间 t 的连续性函数, 这样如果令 $\epsilon \rightarrow 0$,

$$0 = \tau[a(t + \epsilon) - a(t - \epsilon)] + 0, \text{ as } \epsilon \rightarrow 0. \quad (9.66)$$

因此, 随着 $\epsilon \rightarrow 0$, 我们有 $a(t + \epsilon) - a(t - \epsilon) \rightarrow 0$, 或者说 a 是 t 的连续性函数。

现在我们考虑一个带电粒子在 $t = 0$ 时从区域 I 运动至区域 II。在区域 I 内, 粒子受到一个恒定的外力 F_0 , 而在区域 II 内, 没有其它外力。方程 (9.62) 在区域 I 的解为

$$a^{\text{I}} = a_0^{\text{I}} e^{t/\tau} + F_0/m, \quad (9.67)$$

在区域 II 的解为

$$a^{\text{II}} = a_0^{\text{II}} e^{t/\tau}. \quad (9.68)$$

让 a^{II} 和 a^{I} 在 $t = 0$ 时相等, 我们有

$$a_0^{\text{I}} + F_0/m = a_0^{\text{II}}. \quad (9.69)$$

除非 $a_0^{\text{I}} = -F_0/m$, 否则 $a_0^{\text{II}} = 0$ 。只要 $a_0^{\text{II}} \neq 0$, 在区域 II 的粒子在无任何外力的情况下会无限加速或减速! 这种情况在辐射阻尼力占主导情况下也存在, 因为此时的运动方程为

$$a = \tau \dot{a}, \quad (9.70)$$

其解为 $a = a_0 e^{t/\tau}$ 。在之后的内容里, 我们都将假设辐射阻尼力远小于其它外力。

9.3 电荷体系在远场近似下的辐射

如之前一样，为简化讨论，我们将坐标系的原点选在电荷体系的内部。定义电荷位置 \mathbf{x}' ，场点坐标 \mathbf{x} ，且

$$\mathbf{R} = \mathbf{x} - \mathbf{x}'. \quad (9.71)$$

我们通常都会考虑波动形式的辐射。相对应有三个空间尺度， $|\mathbf{x}|$ ， $|\mathbf{x}'|$ ，和 λ ，典型的波长。为简化分析，我们作几个假设。首先，假设 $|\mathbf{x}'| \ll (|\mathbf{x}|, \lambda)$ ；即系统的空间尺度要远小于波长和场点到源点的距离。在此基础上，我们可以通过 λ 与 $|\mathbf{x}|$ 的大小关系定义如下三个辐射区域

1. 近场区 $|\mathbf{x}| \ll \lambda$.
2. 中间场区: $|\mathbf{x}| \sim \lambda$.
3. 远场区（辐射场区）: $|\mathbf{x}| \gg \lambda$.

在本课里，我们仅考虑远场区的电磁场。

我们先介绍一下 $|\mathbf{x}'| \ll \lambda$ 的物理含义。假设对应于 λ 的典型波动周期为 T ，或者说 $\lambda \sim Tc$ ，这里 c 是真空中光速。现在考虑一个粒子在 T 时间内的运动，其距离的量级为 $|\mathbf{x}'|$ ，即系统的空间尺度，这样粒子的特征速度可以表示为 $v \sim |\mathbf{x}'|/T$ 。所以 $|\mathbf{x}'| \ll \lambda$ 意味着

$$|\mathbf{x}'| \ll \lambda \sim Tc \quad (9.72)$$

或者

$$\frac{|\mathbf{x}'|}{T} \sim v \ll c; \quad (9.73)$$

即 $|\mathbf{x}'| \ll \lambda$ 意味着粒子是非相对论的。

现在我们来分析一个电荷系统在远场处在电磁场与辐射。因为这里我们考虑的是电荷体系的辐射，我们需要用一般的推迟势来计算 \mathbf{E} 和 \mathbf{B} 。我们可以用傅立叶变换来分析场，因此我们可以仅仅考虑一个频率为 ω 的成分，即假定源 ρ 和 \mathbf{j} 的形式分别为 $\rho(\mathbf{x}')e^{-i\omega t}$ 和 $\mathbf{j} = \mathbf{j}(\mathbf{x}')e^{-i\omega t}$ 。一般情况下的 $\rho(t)$ 和 $\mathbf{j}(t)$ 可以通过傅立叶变换写成一系统频率的叠加。这样，对应的推迟势为

$$\phi(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho_0(\mathbf{x}')e^{-i\omega(t-R/c)}}{R} dV', \quad (9.74)$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}_0(\mathbf{x}')e^{-i\omega(t-R/c)}}{R} dV'. \quad (9.75)$$

对真空中的电磁波而言，我们有 $k = \omega/c$ ，上式可以重写为

$$\phi(\mathbf{x}, t) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho_0(\mathbf{x}')e^{i(kR-\omega t)}}{R} dV', \quad (9.76)$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}_0(\mathbf{x}')e^{i(kR-\omega t)}}{R} dV'. \quad (9.77)$$

因为势函数里的时间依赖关系为 $e^{-i\omega t}$ ，所以

$$\frac{\partial}{\partial t} = -i\omega, \quad (9.78)$$

这样在源区之外，我们有

$$\nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = -i\frac{\omega}{c^2} \mathbf{E} = -ik\mathbf{E}/c, \quad (9.79)$$

或者

$$\mathbf{E} = i\frac{c}{k} \nabla \times \mathbf{B}, \text{ and } \mathbf{B} = \nabla \times \mathbf{A}. \quad (9.80)$$

因此, 对一个电荷系统的辐射而言, 一种较简便的计算方法是先利用矢势 \mathbf{A} 计算 \mathbf{B} , 然后利用式 (9.80) 去计算电场 \mathbf{E} 。注意到势函数的时间部分和空间部分可以很容易的分离开, 我们有

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}(\mathbf{x})e^{-i\omega t}, \quad (9.81)$$

这里

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int \frac{\dot{\mathbf{j}}_0(\mathbf{x}')e^{ikR}}{R} dV'. \quad (9.82)$$

在以下的计算中, 我们仅需 $\mathbf{A}(\mathbf{x})$, 因为对时间的依赖关系是固定的, 且容易计算。

现在我们回到方程 (9.77) 并利用远场近似, 即 $|\mathbf{x}| \gg \lambda \gg |\mathbf{x}'|$, 来简化 $\mathbf{A}(\mathbf{x})$ 的表达式. 与之前静电和静磁不同的是, 这里有两个小量

$$\epsilon \equiv |\mathbf{x}'|/|\mathbf{x}| \ll 1, \quad (9.83)$$

和

$$\eta \equiv |\mathbf{x}'|/\lambda \sim v/c \ll 1. \quad (9.84)$$

如果我们比较 η 与 ϵ , 容易发现

$$\eta/\epsilon = |\mathbf{x}|/\lambda \gg 1. \quad (9.85)$$

在之后的泰勒展开中, 我们最多保留分析到 $\mathcal{O}(\eta)$ 项, 这样所以 $\mathcal{O}(\epsilon)$ 的项都不会讨论, 因为是高阶项。

首先, 我们通过 R 进行泰勒展开,

$$R = \sqrt{(\mathbf{x} - \mathbf{x}')^2} = (\mathbf{x}^2 + \mathbf{x}'^2 - 2\mathbf{x} \cdot \mathbf{x}')^{1/2} \approx r - \mathbf{x}' \cdot \mathbf{n} + \dots, \quad (9.86)$$

这里 $r = |\mathbf{x}|$, $\mathbf{n} \equiv \mathbf{x}/r$. 将方程 (9.86) 代入式 (9.82), 我们有

$$\mathbf{A}(\mathbf{x}) \approx \frac{\mu_0}{4\pi r} \int \dot{\mathbf{j}}_0(\mathbf{x}')e^{ikr - ik\mathbf{x}' \cdot \mathbf{n}} dV' = \frac{\mu_0 e^{ikr}}{4\pi r} \int \dot{\mathbf{j}}_0(\mathbf{x}')e^{-ik\mathbf{x}' \cdot \mathbf{n}} dV'. \quad (9.87)$$

注意到这里我们忽略了分母里的 $\mathbf{x}' \cdot \mathbf{n}$ 而保留了相位里的 $ik\mathbf{x}' \cdot \mathbf{n}$. 原因如下, 首先

$$\frac{1}{R} = \frac{1}{r - \mathbf{x}' \cdot \mathbf{n}} \approx \frac{1}{r} \left(1 + \frac{\mathbf{x}' \cdot \mathbf{n}}{r} \right) \sim \frac{1}{r} (1 + r'/r) = \frac{1}{r} (1 + \epsilon), \quad (9.88)$$

另一方面

$$e^{-ik\mathbf{x}' \cdot \mathbf{n}} \approx 1 - ik\mathbf{x}' \cdot \mathbf{n} \sim 1 - ir'/\lambda = 1 - i\eta \quad (9.89)$$

因为 $\eta/\epsilon \gg 1$ 而我们最多保留到 $\mathcal{O}(\eta)$ 项, 而不是 ϵ , 因此我们一般不能忽略 $e^{-ik\mathbf{x}' \cdot \mathbf{n}}$.

我们也可以从推迟势的角度简单理解上述结论. 利用方程 (9.86),

$$\rho(\mathbf{x}', t') \approx \rho(\mathbf{x}', t - r/c + \mathbf{x}' \cdot \mathbf{n}/c) \quad (9.90)$$

$$\approx \rho(\mathbf{x}', t - r/c) + \left. \frac{\partial \rho}{\partial t'} \right|_{t=r/c} \frac{\mathbf{x}' \cdot \mathbf{n}}{c}. \quad (9.91)$$

因此, 能否忽略第二项取决于 ρ 在 $\mathbf{x}' \cdot \mathbf{n}/c$ 这么长的时间尺度内变化了多少, 而不是其大小本身。

本章剩下的分析均利用了少量 $\eta = |\mathbf{x}'|/\lambda$, 过程类似于静电与静磁的多极展开。

9.4 电偶极辐射

仅保留到 $\eta = |\mathbf{x}'|/\lambda$ 的最低阶项, 方程 (9.87) 变成

$$\mathbf{A}(\mathbf{x}) \approx \frac{\mu_0 e^{ikr}}{4\pi r} \int \mathbf{j}_0(\mathbf{x}') dV'. \quad (9.92)$$

因为 $\mathbf{j}_0(\mathbf{x}') = \rho_0(\mathbf{x}')\mathbf{v}$, 积分成为

$$\int \mathbf{j}_0(\mathbf{x}') dV' = \int \rho_0(\mathbf{x}') \mathbf{v} dV' = \sum_a q_a \mathbf{v}_a = \dot{\mathbf{p}}. \quad (9.93)$$

这里我们利用了

$$\rho(\mathbf{x}') = \sum_a q_a \delta(\mathbf{x}' - \mathbf{x}'_a), \quad (9.94)$$

和电偶极矩的定义

$$\mathbf{p} = \sum_a q_a \mathbf{x}'_a \Rightarrow \dot{\mathbf{p}} = \sum_a q_a \dot{\mathbf{x}}'_a = \sum_a q_a \mathbf{v}_a. \quad (9.95)$$

因此, 式 (9.92) 变成

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0 e^{ikr}}{4\pi r} \dot{\mathbf{p}}. \quad (9.96)$$

这即是电偶极辐射在远场处对应的矢势。

从式 (9.96), 我们容易得到磁场 \mathbf{B} 为

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\mu_0 e^{ikr}}{4\pi} \left(\frac{ik}{r} - \frac{1}{r^2} \right) \nabla r \times \dot{\mathbf{p}} \quad (9.97)$$

$$= \frac{ik\mu_0 e^{ikr}}{4\pi r} \left(1 - \frac{1}{ikr} \right) \mathbf{n} \times \dot{\mathbf{p}}. \quad (9.98)$$

利用 $1/kr \sim \lambda/r \sim \epsilon/\eta \ll 1$,

$$\mathbf{B} = \frac{ik\mu_0 e^{ikr}}{4\pi r} \mathbf{n} \times \dot{\mathbf{p}} = \frac{i\omega\mu_0 e^{ikr}}{4\pi cr} \mathbf{n} \times \dot{\mathbf{p}}. \quad (9.99)$$

因为时间依赖关系为 $e^{-i\omega t}$, $i\omega\dot{\mathbf{p}} = -\ddot{\mathbf{p}}$, 且

$$\mathbf{B} = \frac{\mu_0 e^{ikr}}{4\pi cr} \ddot{\mathbf{p}} \times \mathbf{n}. \quad (9.100)$$

从上述分析, 我们可以看到, 当计算 ∇ , 且我们仅需要 λ/r 的最低阶项时, 我们可以简单的将 ∇ 替换为 $ik\mathbf{n}$.

电场可以通过式 (9.80) 得到,

$$\mathbf{E} = i\frac{c}{k} \nabla \times \mathbf{B} = i\frac{c}{k} ik\mathbf{n} \times \mathbf{B} = c\mathbf{B} \times \mathbf{n} = \frac{\mu_0 e^{ikr}}{4\pi r} (\ddot{\mathbf{p}} \times \mathbf{n}) \times \mathbf{n}. \quad (9.101)$$

如果我们定义 $\theta \equiv \langle \mathbf{p}, \mathbf{n} \rangle$, 注意到 $\ddot{\mathbf{p}} = -\omega^2 \mathbf{p}$, 因此 θ 也是 \mathbf{p} 和 \mathbf{n} 之间的夹角。电场 \mathbf{E} (式 (9.101)) 和磁场 \mathbf{B} (式 (9.100)) 分别为

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0 e^{ikr}}{4\pi cr} \ddot{\mathbf{p}} \sin\theta \mathbf{e}_\phi, \quad (9.102)$$

$$\mathbf{E}(\mathbf{x}) = \frac{\mu_0 e^{ikr}}{4\pi r} \ddot{\mathbf{p}} \sin\theta \mathbf{e}_\theta. \quad (9.103)$$

这里 (r, θ, ϕ) 构成球坐标系。这便是电偶极辐射在远场处的电场和磁场。

为了计算电偶极辐射的功率，我们从玻印廷矢量出发

$$\mathbf{S} = \frac{1}{\mu_0} \text{Re}(\mathbf{E}(\mathbf{x}, t)) \times \text{Re}(\mathbf{B}(\mathbf{x}, t)) = \frac{1}{\mu_0} \left(\frac{\mathbf{E} + \mathbf{E}^*}{2} \right) \times \left(\frac{\mathbf{B} + \mathbf{B}^*}{2} \right) \quad (9.104)$$

$$= \frac{1}{4\mu_0} (\mathbf{E} \times \mathbf{B} + \mathbf{E} \times \mathbf{B}^* + \mathbf{E}^* \times \mathbf{B} + \mathbf{E}^* \times \mathbf{B}^*). \quad (9.105)$$

利用方程 (9.102) 与 (9.103),

$$\mathbf{E} \times \mathbf{B} = \mathbf{E}(\mathbf{x}) \times \mathbf{B}(\mathbf{x}) e^{-2i\omega t}, \quad (9.106)$$

$$\mathbf{E} \times \mathbf{B}^* = \mathbf{E}(\mathbf{x}) \times \mathbf{B}(\mathbf{x}). \quad (9.107)$$

因为玻印廷矢量有正比于 $e^{-2i\omega t}$ ，即快速变化的项。在此情况下，通常我们更关心短时间平均后的玻印廷矢量在长时间尺度上如何变化

$$\langle \mathbf{S} \rangle = \frac{1}{T} \int_t^{t+T} \mathbf{S}(\tau) d\tau. \quad (9.108)$$

注意到

$$\langle e^{im\omega t} \rangle = \frac{1}{T} \int_t^{t+T} e^{im\omega\tau} d\tau = 0 \text{ for } m \neq 0. \quad (9.109)$$

因此时间平均后的玻印廷矢量为

$$\langle \mathbf{S} \rangle = \frac{1}{4\mu_0} (\mathbf{E} \times \mathbf{B}^* + \mathbf{E}^* \times \mathbf{B}) = \frac{1}{2\mu_0} \text{Re}(\mathbf{E} \times \mathbf{B}^*). \quad (9.110)$$

对电偶极辐射, $\mathbf{E} = c\mathbf{B} \times \mathbf{n}$, 这样

$$\mathbf{E} \times \mathbf{B}^* = (c\mathbf{B} \times \mathbf{n}) \times \mathbf{B}^* = c\mathbf{n}(\mathbf{B} \cdot \mathbf{B}^*) = c|\mathbf{B}|^2 \mathbf{n}, \quad (9.111)$$

因为从式 (9.100) 可以得到 $\mathbf{B} \perp \mathbf{n}$ 。将式 (9.111) 代入方程 (9.110) 得到

$$\langle \mathbf{S} \rangle = \frac{c}{2\mu_0} |\mathbf{B}|^2 \mathbf{n}. \quad (9.112)$$

利用式 (9.100), 辐射功率的角分布为

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \langle \mathbf{S} \rangle \cdot \mathbf{n} r^2 = \frac{\mu_0}{32\pi^2 c} |\ddot{\mathbf{p}} \times \mathbf{n}|^2, \quad (9.113)$$

或者

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{\mu_0}{32\pi^2 c} |\ddot{\mathbf{p}}|^2 \sin^2 \theta = \frac{\mu_0 \omega^4}{32\pi^2 c} |\mathbf{p}|^2 \sin^2 \theta, \quad (9.114)$$

这里我们利用了 $|\ddot{\mathbf{p}}| = \omega^2 |\mathbf{p}|$ ，且在我们所做假设下，粒子是非相对论的。总的辐射功率为

$$\langle P \rangle = \int \left\langle \frac{dP}{d\Omega} \right\rangle d\Omega = \int \left\langle \frac{dP}{d\Omega} \right\rangle \sin \theta d\theta d\varphi \quad (9.115)$$

$$= \frac{\mu_0}{32\pi^2 c} |\ddot{\mathbf{p}}|^2 \int_0^{2\pi} d\varphi \int_0^\pi \sin^3 \theta d\theta \quad (9.116)$$

$$= \frac{\mu_0 |\ddot{\mathbf{p}}|^2}{12\pi c} = \frac{\mu_0 \omega^4 |\mathbf{p}_0|^2}{12\pi c}, \quad (9.117)$$

这里我们利用了 $\mathbf{p} = \mathbf{p}_0 e^{-i\omega t}$ 。

作为电偶极辐射的一个应用，我们考虑一个线性天线的辐射。这个天线的长度 $d \ll \lambda$ 。将天线的中心定义为坐标系的源点，并假设天线中的电流为

$$\mathbf{I} = I\mathbf{e}_z = I(z)e^{-i\omega t}\mathbf{e}_z = I_0\left(1 - \frac{2|z|}{d}\right)e^{-i\omega t}\mathbf{e}_z. \quad (9.118)$$

从连续性方程,

$$\nabla \cdot \mathbf{j} - i\omega\rho = 0 \quad (9.119)$$

或者 $\rho = \nabla \cdot \mathbf{j}/i\omega$, 所以线电荷密度为

$$\rho' = \rho \Delta A = \frac{\nabla \cdot \mathbf{j} \Delta A}{i\omega} = \frac{\nabla \cdot (I \mathbf{e}_z)}{i\omega} = \frac{1}{i\omega} \frac{dI}{dz} = i \frac{2I_0}{\omega d} \text{sgn} z. \quad (9.120)$$

相应的电偶极矩为

$$\mathbf{p} = \sum_a q_a \mathbf{x}'_a = \sum q(z) z \mathbf{e}_z = \int_{-d/2}^{d/2} \rho' z dz \mathbf{e}_z. \quad (9.121)$$

将式 (9.120) 代入 (9.121) 得到

$$\mathbf{p} = 2 \int_0^{d/2} i \frac{2I_0}{\omega d} z dz \mathbf{e}_z = i \frac{I_0 d}{2\omega} \mathbf{e}_z. \quad (9.122)$$

将式 (9.122) 代入 (9.114) 与 (9.117) 得到

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{\mu_0 \omega^4}{32\pi^2 c} \left(\frac{I_0 d}{2\omega} \right)^2 \sin^2 \theta = \frac{\mu_0 c I_0^2}{128\pi^2} (kd)^2 \sin^2 \theta, \quad (9.123)$$

和

$$\langle P \rangle = \frac{\mu_0 c}{48\pi} I_0^2 (kd)^2. \quad (9.124)$$

总的辐射功率正比于 kd 或者说 d/λ 。所以对于一给定波长, 为了增加辐射功率, 我们可以增加 d 。但是上述公式的一个前提是 $d/\lambda \ll 1$, 或 $kd \ll 1$, 因此不能无限增大 d 且还利用方程 (9.124) 去计算辐射功率。

9.5 磁偶极与电四极辐射

通常情况下上述电偶极辐射占主导。然而, 当电偶极辐射为 0 时, 或者我们需要准备到 $\mathcal{O}(\eta)$ 的辐射时, 我们需要展开 $e^{-ik\mathbf{x}' \cdot \mathbf{n}}$ 且保留 $\mathcal{O}(\eta)$ 的相关项。

泰勒展开势函数为

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0 e^{ikr}}{4\pi r} \int \mathbf{j}(\mathbf{x}') e^{-ik\mathbf{x}' \cdot \mathbf{n}} dV' \quad (9.125)$$

$$\approx \frac{\mu_0 e^{ikr}}{4\pi r} \int \mathbf{j}(\mathbf{x}') (1 - ik\mathbf{x}' \cdot \mathbf{n}) dV'. \quad (9.126)$$

在小量 r'/λ 的一阶项为

$$\mathbf{A}(\mathbf{x}) = -\frac{ik\mu_0 e^{ikr}}{4\pi r} \int \mathbf{j} \mathbf{x}' \cdot \mathbf{n} dV' = -\frac{ik\mu_0 e^{ikr}}{4\pi r} \int \mathbf{n} \cdot \mathbf{x}' \mathbf{j} dV'. \quad (9.127)$$

任何一个张量都可以写成一个对称张量和反对称张量的和。将此运用到 $\mathbf{x}' \mathbf{j}$ 上

$$\mathbf{x}' \mathbf{j} = \frac{1}{2}(\mathbf{x}' \mathbf{j} + \mathbf{j} \mathbf{x}') + \frac{1}{2}(\mathbf{x}' \mathbf{j} - \mathbf{j} \mathbf{x}'). \quad (9.128)$$

将 (9.128) 从左边点乘 \mathbf{n}/c , 我们有

$$\mathbf{n} \cdot \mathbf{x}' \mathbf{j} = \frac{1}{2}[(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} + (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] + \frac{1}{2}[(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} - (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] \quad (9.129)$$

$$= \frac{1}{2}[(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} + (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] + \frac{1}{2}(\mathbf{x}' \times \mathbf{j}) \times \mathbf{n}. \quad (9.130)$$

将方程 (9.128) 代入式 (9.127)。我们先考虑反对称项。该项的积分为

$$\frac{1}{2} \int (\mathbf{x}' \times \mathbf{j}) \times \mathbf{n} dV' = -\mathbf{n} \times \frac{1}{2} \int (\mathbf{x}' \times \mathbf{j}) dV' = -\mathbf{n} \times \mathbf{m}, \quad (9.131)$$

这里

$$\mathbf{m} = \frac{1}{2} \int (\mathbf{x}' \times \mathbf{j}) dV' \quad (9.132)$$

是之前学过的磁偶极矩。

对称项的积分为

$$\frac{1}{2} \int [(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} + (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] dV', \quad (9.133)$$

注意到 $\mathbf{j} = \sum q_a \delta(\mathbf{x}' - \mathbf{x}'_a) \mathbf{v}_a$, 积分变成

$$\frac{1}{2} \int [(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} + (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] dV' \quad (9.134)$$

$$= \frac{1}{2} \sum q_a [(\mathbf{n} \cdot \mathbf{x}'_a) \mathbf{v}_a + (\mathbf{n} \cdot \mathbf{v}_a) \mathbf{x}'_a]. \quad (9.135)$$

利用 $\mathbf{v}_a = d\mathbf{x}'_a/dt$, 我们有

$$(\mathbf{n} \cdot \mathbf{x}'_a) \mathbf{v}_a + (\mathbf{n} \cdot \mathbf{v}_a) \mathbf{x}'_a \quad (9.136)$$

$$= (\mathbf{n} \cdot \mathbf{x}'_a) \frac{d\mathbf{x}'_a}{dt} + (\mathbf{n} \cdot \frac{d\mathbf{x}'_a}{dt}) \mathbf{x}'_a \quad (9.137)$$

$$= \mathbf{n} \cdot \frac{d}{dt} (\mathbf{x}'_a \mathbf{x}'_a). \quad (9.138)$$

将方程 (9.138) 代入式 (9.135), 我们有

$$\frac{1}{2} \int [(\mathbf{n} \cdot \mathbf{x}') \mathbf{j} + (\mathbf{n} \cdot \mathbf{j}) \mathbf{x}'] dV' = \frac{1}{2} \mathbf{n} \cdot \frac{d}{dt} \sum q \mathbf{x}' \mathbf{x}'. \quad (9.139)$$

因为 $\mathbf{B} = \nabla \times \mathbf{A} = ik\mathbf{n} \times \mathbf{A}$, 我们可以在 \mathbf{A} 里加入任意一个方向为 \mathbf{n} 的向量。所以我们在方程 (9.138) 加入

$$-\frac{1}{6} \mathbf{n} \cdot \frac{d}{dt} \left(\sum q r'^2 \mathbf{l} \right) \quad (9.140)$$

结果为

$$\frac{\mathbf{n}}{2} \cdot \frac{d}{dt} \sum q \mathbf{x}' \mathbf{x}' - \frac{\mathbf{n}}{6} \cdot \frac{d}{dt} \left(\sum q r'^2 \mathbf{l} \right) \quad (9.141)$$

$$= \frac{\mathbf{n}}{6} \cdot \frac{d}{dt} \sum q (3\mathbf{x}' \mathbf{x}' - r'^2 \mathbf{l}) \quad (9.142)$$

$$= \mathbf{n} \cdot \frac{1}{6} \dot{\mathbf{D}} \equiv \frac{1}{6} \dot{\mathbf{D}}, \quad (9.143)$$

这里的张量 \mathbf{D} 即是之前学过的电四极张量, 且向量 $\mathbf{D} \equiv \mathbf{n} \cdot \mathbf{D}$ 。

将对称与反对称部分合并, 我们有总的矢势为

$$\mathbf{A} = -\frac{ik\mu_0 e^{ikr}}{4\pi r} \left(\frac{1}{6} \dot{\mathbf{D}} - \mathbf{n} \times \mathbf{m} \right). \quad (9.144)$$

从上述过程可以看出磁偶极辐射与电四极辐射是同一量级的。辐射磁场可以从 \mathbf{A} 中得到, 为

$$\mathbf{B} = \nabla \times \mathbf{A} = ik\mathbf{n} \times \frac{-ik\mu_0 e^{ikr}}{4\pi r} \left(\frac{1}{6} \dot{\mathbf{D}} - \mathbf{n} \times \mathbf{m} \right) \quad (9.145)$$

$$= \frac{k^2 \mu_0 e^{ikr}}{4\pi r} \mathbf{n} \times \left(\frac{1}{6} \dot{\mathbf{D}} - \mathbf{n} \times \mathbf{m} \right). \quad (9.146)$$

我们也可以分离磁偶极辐射与电四极辐射的磁场。磁偶极部分的磁场为

$$\begin{aligned}\mathbf{B}_m &= \frac{k^2 \mu_0 e^{ikr} \mathbf{n}}{4\pi r} \times (-\mathbf{n} \times \mathbf{m}) \\ &= (\mathbf{n} \times \mathbf{m}) \times \mathbf{n} \frac{k^2 \mu_0 e^{ikr}}{4\pi r} \\ &= (\ddot{\mathbf{m}} \times \mathbf{n}) \times \mathbf{n} \frac{\mu_0 e^{ikr}}{4\pi c^2 r}.\end{aligned}\quad (9.147)$$

电四极部分的磁场为

$$\mathbf{B}_D = \mathbf{n} \times \dot{\mathbf{D}} \frac{k^2 \mu_0 e^{ikr}}{24\pi r} = -i\mathbf{n} \times \mathbf{D} \frac{k^3 c \mu_0 e^{ikr}}{24\pi r} = \ddot{\mathbf{D}} \times \mathbf{n} \frac{\mu_0 e^{ikr}}{24\pi c^2 r}.\quad (9.148)$$

电场可以从 $\mathbf{E} = c\mathbf{B} \times \mathbf{n}$ 得到，相应的

$$\mathbf{E}_m = c\mathbf{B}_m \times \mathbf{n} = (\mathbf{n} \times \ddot{\mathbf{m}}) \frac{\mu_0 e^{ikr}}{4\pi cr},\quad (9.149)$$

$$\mathbf{E}_D = c\mathbf{B}_D \times \mathbf{n} = (\ddot{\mathbf{D}} \times \mathbf{n}) \times \mathbf{n} \frac{\mu_0 e^{ikr}}{24\pi cr}.\quad (9.150)$$

利用式 (9.147) 与 (9.148), 我们可以得到辐射功率的角分布

$$\left\langle \frac{dP}{d\Omega} \right\rangle_m = \left\langle \frac{c}{\mu_0} |\mathbf{B}_m|^2 r^2 \right\rangle = \left\langle \frac{c}{\mu_0} \left| \frac{\mu_0 (\ddot{\mathbf{m}} \times \mathbf{n}) \times \mathbf{n}}{4\pi c^2} \right|^2 \right\rangle = \frac{\mu_0 |\ddot{\mathbf{m}}|^2}{32\pi^2 c^3} \sin^2 \theta.\quad (9.151)$$

$$\left\langle \frac{dP}{d\Omega} \right\rangle_D = \left\langle \frac{c}{\mu_0} |\mathbf{B}_D|^2 r^2 \right\rangle = \left\langle \frac{c}{\mu_0} \left| \frac{\mu_0 \ddot{\mathbf{D}} \times \mathbf{n}}{24\pi c^2} \right|^2 \right\rangle = \frac{\mu_0}{1152\pi^2 c^3} |\ddot{\mathbf{D}} \times \mathbf{n}|^2.\quad (9.152)$$

总的磁偶极辐射功率为

$$\langle P \rangle_m = \int \left\langle \frac{dP}{d\Omega} \right\rangle d\Omega = \iint \frac{\mu_0 |\ddot{\mathbf{m}}|^2}{16\pi^2 c^3} \sin^3 \theta d\theta d\varphi = \frac{\mu_0 |\ddot{\mathbf{m}}|^2}{12\pi c^3}.\quad (9.153)$$

此式与方程 (9.117) 电偶极辐射的总功率表达式高度相似。通过量级分析，你可以知道为什么磁偶极辐射要比电偶极辐射高一阶。

因为 $\mathbf{D} = \mathbf{D} \cdot \mathbf{n}$, 电四极辐射的角功率分布通常会非常复杂。这里我们举一简单例子来计算其辐射功率。假定有三个电荷 $(q, q, -2q)$ 分别位于 $\mathbf{x}'(q) = l\mathbf{e}_z$, $\mathbf{x}'(q) = -l\mathbf{e}_z$, 和 $\mathbf{x}'(-2q) = 0\mathbf{e}_z$, 则

$$\mathbf{D} = 6ql^2 \mathbf{e}_z \mathbf{e}_z,\quad (9.154)$$

所以 $\mathbf{D} = \mathbf{D} \cdot \mathbf{n} = 6ql^2 \mathbf{e}_z \cos \theta$, 这样

$$|\mathbf{D} \times \mathbf{n}| = |6ql^2 \cos \theta \mathbf{e}_z \times \mathbf{n}| = 6ql^2 \cos \theta \sin \theta,\quad (9.155)$$

且

$$|\ddot{\mathbf{D}} \times \mathbf{n}|^2 = 36q^2 l^4 \omega^6 \cos^2 \theta \sin^2 \theta.\quad (9.156)$$

总的辐射功率为

$$\langle P \rangle_D = \frac{\mu_0}{1152\pi^2 c^3} \int_0^{2\pi} d\varphi \int_0^\pi 36q^2 l^4 \omega^6 \cos^2 \theta \sin^3 \theta d\theta = \frac{\mu_0 q^2 l^4 \omega^6}{60\pi c^3}.$$

9.6 自由电子对电磁波的散射

现在我们考虑一个与电荷辐射紧密相关的问题：电荷对电磁波的散射。一个例子是白天虽然不是正对着太阳看，我们也能看到明亮的天空。这是因为空气中的分子，原子等造成了太阳光的散射，而这些散射的光进入了我们

的眼。在经典电动力学框架里，这一过程可以理解如下：当电磁波入射到电荷上，电荷感受到洛伦兹力被加速从而导致电荷辐射出新的电磁波。

能常情况下我们讨论散射的时候，都会使用一个物理量称为散射截面。该截面分两种：微分散射截面和总散射截面。其中微分散射截面 (**differential cross-section**) 定义为某个给定方向上单位时间内散射系统所发出的能量， $\langle dP \rangle$ ，与入射电磁波的能量通量 ($\langle S \rangle$) 的比值。该比值显然有面积的量纲。这里 $\langle \dots \rangle$ 表示时间平均。因此

$$d\sigma = \frac{\langle dP \rangle}{\langle S \rangle}. \quad (9.157)$$

我们可以对 $d\sigma$ 积分得到总散射截面 σ (total scattering cross-section).

在本小节里，我们考虑一个非相对论自由电子对电磁波的散射。该电子原本是静止的。对非相对论粒子而言，其所受磁场力与电场力的比值为

$$\frac{|\mathbf{v} \times \mathbf{B}|}{|\mathbf{E}|} \sim \frac{vB}{E} \sim \frac{v}{c} \ll 1. \quad (9.158)$$

因此，我们可以忽略磁场力而仅考虑电场力。假设入射波是一个线偏振的单色电磁波，即

$$\mathbf{E} = \mathbf{E}_0 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha), \quad (9.159)$$

这样粒子的运动方程为

$$\ddot{\mathbf{x}} = \frac{q}{m} \mathbf{E} = \frac{q}{m} \mathbf{E}_0 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha). \quad (9.160)$$

因为我们假设非相对论，电子的运动振幅 $l \sim vT \ll cT = \lambda$ ，这里 T 是波动周期， λ 是波长。我们可以忽略相位里的空间因子变化，即 $\mathbf{k} \cdot \mathbf{x} = \mathbf{k} \cdot \mathbf{x}_0$ 。为简便，我们让 $\mathbf{x}_0 = 0$ 。方程 (9.160) 变成

$$\ddot{\mathbf{x}} = \frac{q}{m} \mathbf{E}_0 \cos(-\omega t + \alpha). \quad (9.161)$$

用复数表达式，

$$\ddot{\mathbf{x}} = \frac{q}{m} \hat{\mathbf{E}}_0 e^{-i\omega t}, \text{ with } \hat{\mathbf{E}}_0 = \mathbf{E}_0 e^{i\alpha}. \quad (9.162)$$

因为电子有非零加速度，它会辐射出电磁波，其最低阶辐射为电偶极辐射。电偶极矩为 $\mathbf{p} = q\mathbf{x}$ 且

$$\ddot{\mathbf{p}} = q\ddot{\mathbf{x}} = \frac{q^2}{m} \hat{\mathbf{E}}_0 e^{-i\omega t}. \quad (9.163)$$

所以平均的辐射功率可以用式 (9.114) 得到，

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{\mu_0}{32\pi^2 c} |\ddot{\mathbf{p}}|^2 \sin^2 \theta = \frac{\mu_0}{32\pi^2 c} \frac{q^4}{m^2} E_0^2 \sin^2 \theta, \quad (9.164)$$

或者

$$\langle dP \rangle = \frac{\mu_0}{32\pi^2 c} \frac{q^4}{m^2} E_0^2 \sin^2 \theta d\Omega. \quad (9.165)$$

而入射波的能量流为

$$S = \frac{1}{\mu_0} |\mathbf{E} \times \mathbf{B}| = \frac{1}{\mu_0 c} E^2, \quad (9.166)$$

因此

$$\langle S \rangle = \frac{1}{\mu_0 c} \frac{E_0^2}{2} = \frac{1}{2\mu_0 c} E_0^2. \quad (9.167)$$

将式 (9.167) 和 (9.165) 代入方程 (9.157), 我们可以得到散射截面为

$$d\sigma = \frac{\mu_0^2 q^4}{16\pi^2 m^2} \sin^2 \theta d\Omega = \left(\frac{q^2}{4\pi\epsilon_0 mc^2} \right)^2 \sin^2 \theta d\Omega. \quad (9.168)$$

显然, 自由电子的散射截面 $d\sigma$ 与入射波的频率无关。

总散射截面可以通过积分得到, 即 $\sigma = \int d\sigma$, 或者

$$\sigma = \left(\frac{q^2}{4\pi\epsilon_0 mc^2} \right)^2 \int_0^\pi \sin^3 \theta d\theta \int_0^{2\pi} d\varphi = \frac{8\pi}{3} \left(\frac{q^2}{4\pi\epsilon_0 mc^2} \right)^2. \quad (9.169)$$

利用经典电子半径 $r_e = q^2/4\pi\epsilon_0 mc^2$, 则

$$\sigma = \frac{8\pi}{3} r_e^2. \quad (9.170)$$

非相对论电子对电磁波的散射被称为 Thomson 散射, 散射截面 (9.169) 或 (9.170) 被称为 Thomson 公式。

9.7 束缚电子对电磁波的散射

现在我们计算一个束缚电子对电磁波的散射, 例如原子中的电子。如之前, 我们用一个本征频率为 ω_0 谐振子模型来描述电子的运动。相应的电子运动方程为

$$\ddot{\mathbf{x}} + \omega_0^2 \mathbf{x} = (\mathbf{F}_{\text{damping}} + \mathbf{F}_{\text{drive}})/m, \quad (9.171)$$

显然如果 $\omega_0 \rightarrow 0$, 则回复到自由电子的状态。这里 $\mathbf{F}_{\text{damping}}$ 是阻尼项, 为了不在共振时出现无穷大的解。本节里, 我们仅考虑辐射阻尼, 因此

$$\mathbf{F}_{\text{damping}} = m\tau \dot{\mathbf{x}} = m\tau \ddot{\mathbf{x}} = -m\tau\omega^2 \dot{\mathbf{x}} = -m\gamma \dot{\mathbf{x}}, \quad (9.172)$$

这里 $\gamma = \tau\omega^2$ 。另一方面, $\mathbf{F}_{\text{drive}}$ 是入射波施加的力,

$$\mathbf{F}_{\text{drive}} = q\mathbf{E}e^{-i\omega t}. \quad (9.173)$$

将 $\mathbf{F}_{\text{damping}}$ 和 $\mathbf{F}_{\text{drive}}$ 代入方程 (9.171), 我们得到

$$\ddot{\mathbf{x}} + \omega_0^2 \mathbf{x} + \gamma \dot{\mathbf{x}} = \frac{q}{m} \mathbf{E}e^{-i\omega t}. \quad (9.174)$$

注意, 我们假设 $\gamma \ll (\omega_0, \omega)$ 。这样在非共振时, 我们可以忽略阻尼项。

从上式, \mathbf{x} 有一个形如 $\mathbf{x}_0 e^{-i\omega t}$ 的解为

$$\mathbf{x} = \frac{(q/m)\mathbf{E}_0 e^{-i\omega t}}{\omega_0^2 - \omega^2 - i\gamma\omega} = \frac{(q/m)\mathbf{E}_0 e^{-i(\omega t - \delta)}}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}, \quad (9.175)$$

这里 $-\delta$ 是 $\omega_0^2 - \omega^2 - i\gamma\omega$ 的相位, 且

$$\tan \delta = \frac{\gamma\omega}{\omega_0^2 - \omega^2}. \quad (9.176)$$

从方程 (9.175), 电偶极矩 \mathbf{p} 的二阶时间导数为

$$\ddot{\mathbf{p}} = q\ddot{\mathbf{x}} = -\frac{(q^2/m)\omega^2 \mathbf{E}_0 e^{-i(\omega t - \delta)}}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}. \quad (9.177)$$

剩下的散射截面的计算与前一节相同。电偶极的辐射角功率分布为

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{\mu_0}{32\pi^2 c} |\ddot{\mathbf{p}}|^2 \sin^2 \theta \quad (9.178)$$

$$= \frac{\mu_0}{32\pi^2 c} \left(\frac{q^2}{m} \right)^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} E_0^2 \sin^2 \theta. \quad (9.179)$$

总辐射功率为 $\int \langle \frac{dP}{d\Omega} \rangle d\Omega$, 或者

$$\langle P \rangle = \frac{\mu_0 q^4}{12\pi m^2 c} \frac{\omega^4 E_0^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \quad (9.180)$$

因此总的散射截面为

$$\sigma = \frac{\langle P \rangle}{\langle S \rangle} = \frac{\langle P \rangle}{(1/2\mu_0 c) E_0^2} \quad (9.181)$$

$$= \frac{\mu_0^2 q^4}{6\pi m^2} \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}, \quad (9.182)$$

这里我们代入了方程 (9.167) 中 $\langle S \rangle$ 的表达式。利用经典电子半径 $r_e = q^2/4\pi\epsilon_0 mc^2$, 散射截面为

$$\sigma = \frac{8\pi}{3} r_e^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \quad (9.183)$$

我们可以考虑方程 (9.183) 的几种极限情况。如果 $\omega \ll \omega_0$, 则

$$\sigma \approx \frac{8\pi}{3} r_e^2 \frac{\omega^4}{\omega_0^4}. \quad (9.184)$$

这种散射称为瑞利 (Rayleigh) 散射。蓝色的天空, 经色的晚霞和月全食时红色的月亮都是由此造成的, 因为对蓝光的散射比红光强很多。另一方面, 如果 $\omega \gg \omega_0$, 则

$$\sigma \approx \frac{8\pi}{3} r_e^2, \quad (9.185)$$

即回复到自由电子的 Thomson 散射。此时电子相对于外加电磁波来讲, 其本征运动如此之慢, 可以近似为一个自由电子。第三种特殊情况是, 当 $\omega \sim \omega_0$, 此时发生共振的散射,

$$\sigma = \frac{8\pi}{3} r_e^2 \left(\frac{\omega}{\gamma} \right)^2, \quad (9.186)$$

因为共振, 此时的散射截面远大于 Thomson 散射截面。