

第五章 狭义相对论

狭义相对论是在二十世纪初由庞加莱 (Jules Henri Poincaré, 1854 - 1912) 和爱因斯坦 (Albert Einstein, 1879 - 1955) 各自独立建立的。在历史的长河中，人类很早就开始通过哲学思辨去探讨和理解时间和空间的本质，但是这些思考并没有带来任何实质性的进展。狭义相对论不但彻底改变了人类对时空的看法，而且让人类认识到物理是更有效的理解时间和空间的方式。

现在流行的观点是，对于狭义相对论的建立，爱因斯坦做了主要贡献，而庞加莱只是发展了一些相关的数学形式。非常遗憾，这个流行的观点和历史事实并不相符。庞加莱完全理解光速不变，同时相对性等相对论效应；而且作为他那个时代最伟大的数学家之一，庞加莱发展了四维时空的数学理论，指出洛伦兹变换实质上是一个非欧线性空间的旋转群。这个非欧四维时空现在经常被称为明可夫斯基空间，其实应该被称为庞加莱空间，至少也应该被称为庞加莱-明可夫斯基空间。庞加莱对狭义相对论的贡献被忽略是历史的不幸。

狭义相对论是关于时空结构的基本理论，和物理体系的具体形式无关。任何一个物理体系，无论它是由什么粒子构成的，它在时空中运动和变化都应该符合狭义相对论。如果有一个物理理论违反狭义相对论，大概有两个可能：(1) 粒子运动速度远小于光速，相对论效应可以忽略；(2) 这个理论是错的。

狭义相对论的核心是两个惯性参照系之间的洛伦兹变换。常见的洛伦兹变换推导方法来自爱因斯坦，它的基本出发点是相对性原理和光速不变假设。由于狭义相对论是关于时空的基本理论，这个推导方法很容易造成一个错误印象：光是具有和时间、空间一样地位的基本存在。根据现代物理，光只是传递电磁相互作用的场或粒子，它和传递强相互作用的胶子，弱相互作用的 W

和 Z 玻色子具有相同的地位. 在这里我们将放弃光速不变假设, 只利用相对性原理推导洛伦兹变换.

5.1 洛伦兹变换

如图 5.1 所示, 我们考虑两个惯性参照系 K 和 K' , 它们之间的相对速度 v_k 沿 x 轴. 在第三章我们讨论过不同惯性参照系之间的变换. 我们发现利用惯性定律和速度相加定律, 这两个惯性参照系 K 和 K' 之间有伽利略变换

$$x' = x - v_k t, \quad y' = y, \quad z' = z, \quad t' = t, \quad (5.1)$$

其中 (t, x, y, z) 是参照系 K 的时间和空间坐标而 (t', x', y', z') 是参照系 K' 的时间和空间坐标. 伽利略变换确实和我们的日常经验吻合. 但是狭义相对论告诉我们, 速度相加定律和伽利略变换只是近似成立, 它们和我们的日常经验吻合是因为日常生活中出现的速度 v 远远小于光速. 当物体的运动速度接近光速时, 速度相加定律不再成立, 伽利略变换也随之不成立.

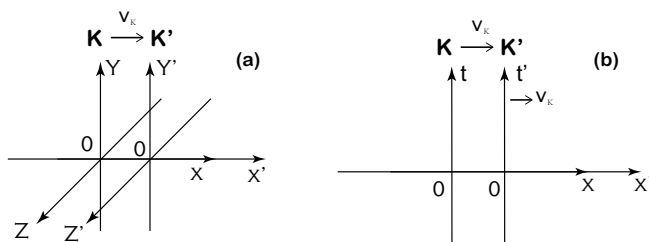


图 5.1: 两个惯性参照系 K 和 K' . (a) 参照系的空间部分; (b) 忽略 y 和 z 轴的时空坐标. 相对速度 v_k 沿 x 轴.

两个不同的惯性参照系之间的正确变换是洛伦兹变换. 这个变换最早是洛伦兹 (Hendrik Antoon Lorentz, 1853 - 1928) 在研究麦克斯韦方程组的时候发现的. 但是洛伦兹并没有认识到这个变换后面深刻而丰富的物理内涵, 他甚至都没有抛弃以太的概念, 认为其中的一个惯性参照系固定在以太上, 只有这个参照系中观测到的时间才是真正的物理时间. 当庞加莱看到洛伦兹的结果, 他立刻认识到了它的重要性. 他以洛伦兹变换为起点发展了完整的狭义

相对论。他的理论不但比同阶段爱因斯坦的理论更完整，而且也得到了光速不变、同时相对性、长度收缩、时间延缓等大家熟知的相对论效应。爱因斯坦独立地发展了狭义相对论，他的出发点和庞加莱不同。庞加莱是直接承认洛伦兹变换的正确性，以此为起点发展他的理论。爱因斯坦则是先假设了两个原理，然后利用这两个原理推导出了洛伦兹变换。爱因斯坦设定的原理是：

相对性原理 在所有惯性参照系中，物理定律具有相同的数学表达形式。

光速不变原理 在所有惯性参照系中，真空中的光速相同。

光速不变原理简洁明了，非常好理解；相对而言，相对性原理则有些晦涩：正确理解“物理定律具有相同的数学表达形式”对于初学者不是一件轻松的事情。幸好，为了推导洛伦兹变换，只需要这样理解相对性原理：一个受力为零的物体在所有惯性参照系中都受力为零，做匀速直线运动或静止。这其实就是第三章讨论过的惯性定律或者牛顿第一定律。

我们的出发点既不同于庞加莱的也和爱因斯坦的不一样。早在 1910 年俄罗斯物理学家伊戈那拓斯基(Vladimir Ignatowski, 1875 -1942) 就发现，只用相对性原理就可以推导出两个惯性参照系之间的洛伦兹变换。伊戈那拓斯基的推导解决一个重要的令人困惑的问题，为什么光在物理中占据这么特殊的位置，出现在基本时空结构的定义中？现代物理告诉我们，自然界有四种基本相互作用，每种相互作用都是由一种或多种粒子传递的，光作为传递电磁相互作用的粒子只是其中一员。虽然大家对光更为熟悉和了解，但是从基础理论的角度看，光子并不比传递其他基本相互作用的粒子，胶子、W 和 Z 玻色子，以及可能存在的引力子，具有更特殊的地位。而且在推导洛伦兹变换时，我们只考虑匀速直线运动，不涉及任何相互作用。另一方面，洛伦兹变换是自然界存在的一个非常基本的变换，适用于所有的基本粒子，比如胶子、电子。人们不得不问为什么一个基于光速或麦克斯韦方程推导出来的变换具有这么高的普适性？我们似乎只有一种选择：自然界存在一个普适速度，光只是碰巧具有这个速度。伊戈那拓斯基的推导表明这个观点是对的：相对性原理要求自然界存在一个普适速度，在我们生活的宇宙中，光和任何静止质量为零的粒子恰好具有这个速度。最后，从这个推导出发也更容易理解为什么

洛伦兹虽然最早发现了洛伦兹变换但是却没有发展出狭义相对论. 洛伦兹变换中的光速一定让洛伦兹误以为这个变换只适用于电磁现象, 不是普适的. 基于这些考虑, 我们在这里按照伊戈那拓斯基的方法推导洛伦兹变换. 在下面的讨论中, 我们忽略和相对速度 v_k 垂直的 y 轴和 z 轴, 因为这两个空间方向不受相对运动的影响. 为此我们建立一个只有 x 轴和时间轴的坐标系, 如图5.1(b) 所示.

在第5.1章我们推导了两个惯性参照系 \mathcal{K} 和 \mathcal{K}' 之间的变换关系. 在利用伽利略速度相加定律以前, 我们得到了如下结果

$$t' = g_1(v_k)x + f_1(v_k)t, \quad (5.2)$$

$$x' = f_1(v_k)(x - v_k t). \quad (5.3)$$

引入第三个惯性参照系 \mathcal{K}'' , 它相对于 \mathcal{K}' 以速度 u_k 沿 x 轴运动. 它们之间的变换是

$$t'' = g_1(u_k)x' + f_1(u_k)t', \quad (5.4)$$

$$x'' = f_1(u_k)(x' - u_k t'). \quad (5.5)$$

结合上面两个变换可以得到 \mathcal{K} 和 \mathcal{K}'' 之间的变换

$$t'' = [f_1(v_k)g_1(u_k) + f_1(u_k)g_1(v_k)]x + f_1(v_k)[f_1(u_k) - v_k g_1(u_k)]t, \quad (5.6)$$

$$x'' = f_1(u_k)[f_1(v_k) - u_k g_1(v_k)]x - (v_k + u_k)f_1(v_k)f_1(u_k)t. \quad (5.7)$$

另外一方面, 根据相对性原理, 参照系 \mathcal{K}'' 相对于参照系 \mathcal{K} 肯定也是在做匀速直线运动, 假设这个速度是 w_k . 那么这两个参照系之间应该有如下的变换关系

$$t'' = g_1(w_k)x + f_1(w_k)t, \quad (5.8)$$

$$x'' = f_1(w_k)(x - w_k t). \quad (5.9)$$

变换 (5.35, 5.38) 和变换 (5.39, 5.40) 应该是同一个变换. 通过比较各项对应

的系数，我们得到

$$\begin{aligned} & f_1(u_k)f_1(v_k) - v_k f_1(v_k)g_1(u_k) \\ &= f_1(v_k)f_1(u_k) - u_k g_1(v_k)f_1(u_k) = f_1(w_k), \end{aligned} \quad (5.10)$$

$$(v_k + u_k)f_1(v_k)f_1(u_k) = w_k f_1(w_k), \quad (5.11)$$

$$f_1(v_k)g_1(u_k) + f_1(u_k)g_1(v_k) = g_1(w_k). \quad (5.12)$$

从上面第一式我们可以得到

$$\frac{v_k f_1(v_k)}{g_1(v_k)} = \frac{u_k f_1(u_k)}{g_1(u_k)} = -C. \quad (5.13)$$

上式第一项只是 v_k 的函数而第二项只是 u_k 的函数，因此 C 必须是一个和速度 v_k, u_k 无关的常数。利用这个关系，我们一方面可以验证公式 (5.11) 和公式 (5.12) 是等价的，另一方面可以进一步简化惯性参照系 \mathcal{K} 和 \mathcal{K}' 之间的变换

$$t' = f_1(v_k)\left(t - \frac{v_k}{C}x\right), \quad (5.14)$$

$$x' = f_1(v_k)(x - v_k t). \quad (5.15)$$

现在只剩函数 $f_1(v_k)$ 和常数 C 未定。上面这个变换的逆变换是

$$t = \frac{1}{f_1(v_k)(1 - v_k^2/C)}\left(t' + \frac{v_k}{C}x'\right), \quad (5.16)$$

$$x = \frac{1}{f_1(v_k)(1 - v_k^2/C)}(x' + v_k t'). \quad (5.17)$$

从另外一个角度，参照系 \mathcal{K} 是以速度 $-v_k$ 相对 \mathcal{K}' 运动，所以按照公式 (5.14), (5.15) 我们应该有

$$t = f_1(v_k)\left(t' + \frac{v_k}{C}x'\right), \quad (5.18)$$

$$x = f_1(v_k)(x' + v_k t'). \quad (5.19)$$

上面我们已经用了在第 4 章得到的结论 $f_1(v_k) = f_1(-v_k)$ 。上面两个变换应该是同一个变换，因为这就要求

$$f_1(v_k) = \pm \frac{1}{\sqrt{1 - v_k^2/C}}. \quad (5.20)$$

这里只能取 + 号, 因为物理上要求 $f_1(0) = 1$. 所以惯性参照系 \mathcal{K} 和 \mathcal{K}' 之间的变换关系是

$$t' = \frac{1}{\sqrt{1 - v_k^2/C}} \left(t - \frac{v_k}{C} x \right), \quad (5.21)$$

$$x' = \frac{1}{\sqrt{1 - v_k^2/C}} (x - v_k t). \quad (5.22)$$

现在还有一个参数 C 需要确定. 我们把上面这个变换称为泛洛伦兹变换因为当 $C = c^2$ (这里 c 是光速) 时它就是洛伦兹变换.

我们继续上面的推导. 结合公式 (5.10) 和 (5.13), 我们有

$$f_1(v_k) f_1(u_k) \left(1 + \frac{u_k v_k}{C} \right) = f_1(w_k). \quad (5.23)$$

再利用公式 (5.11), 我们可以得到速度 w_k

$$w_k = \frac{u_k + v_k}{1 + u_k v_k / C}. \quad (5.24)$$

这个公式实质上是一个速度相加规则: 如果乙相对于甲的速度是 v_k , 丙相对于乙的速度是 u_k , 那么丙相对于甲的速度就是上面的 w_k . 显然, 这个速度相加规则只有在 $C \rightarrow \infty$ 时才和伽利略的速度相加定律一样.

对于常数 C , 纯粹从理论上, 我们不但无法确定它的大小, 甚至无法确定它的正负. 这个常数只能通过实验来确定. 为此我们重新考虑在第 5.1 章中的一个思想实验. 如图 5.2 所示, 在这个实验里有三个箱子, 箱子中各有一个观察者, 小方、小李和小邓. 在小李看来, 小方的箱子向左以速度 v 运动, 小邓的箱子向右以速度 v 运动. 和第 5.1 章中的思想实验相比, 这里多了一个相对于小邓静止的装置, 它能发射一种粒子“戊”. 小方、小李和小邓分别在自己的惯性参照系里测量了粒子“戊”的速度, 发现都是 η , 即粒子“戊”的速度在任何惯性参照系里都是 η . 根据公式 (5.24), 小李和小邓的测量结果应该满足如下关系

$$\eta = \frac{\eta + v}{1 + \eta v / C}. \quad (5.25)$$

由此我们得到 $C = \eta^2$. 这是小方、小李和小邓生活的宇宙中常数 C 的值. 在我们的宇宙中也存在一种粒子, 实验发现它的速度在所有惯性参照系都是一样的. 这个粒子就是光子, 所以在我们的宇宙中常数 $C = c^2$. 我们甚至都不用直接测量光速. 大量的实验已经证实麦克斯韦方程组是对的, 而麦克斯韦方程组必须在洛伦兹变换下不变, 这就要求 $C = c^2$. 日常生活中的速度都远小于光速 c , 这时 c 可以被当着无穷大, 洛伦兹变换退化成伽利略变换, 公式 (5.24) 中的速度相加规则退化为伽利略速度相加定律.

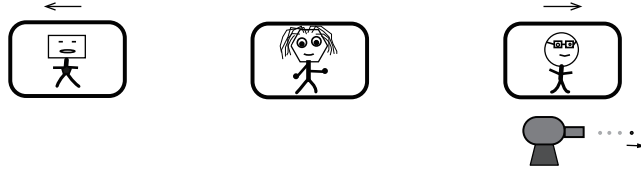


图 5.2: 小方相对小李向左以速度 v 运动, 小邓相对小李向右以速度 v 运动. 有一个装置, 相对于小邓静止, 在水平方向发射了粒子“戊”. 小方、小李和小邓测量了“戊”的速度, 发现是一样的.

假设小方、小李和小邓厌倦了他们宇宙中的物理, 并且发现了一个方法, 成功地将 $C = \eta^2$ 改成了 $C = -\zeta^2$, 并把这个新的宇宙命名为桃花源. 小方、小李和小邓很快发现桃花源宇宙里所有物质的速度都会依赖惯性参照系的选择, 而且还有更奇怪的事情发生. 假设小方、小李和小邓之间的相对速度是 $v = 2\zeta/3$, 在小邓的参照系里粒子“戊”的速度也是 $u = 2\zeta/3$. 那么根据公式 (5.24), 小李认为“戊”的速度是

$$u' = \frac{\frac{2\zeta}{3} + \frac{2\zeta}{3}}{1 - \frac{4\zeta^2}{9\zeta^2}} = \frac{12}{5}\zeta. \quad (5.26)$$

小方认为“戊”的速度是

$$u'' = \frac{\frac{2\zeta}{3} + \frac{12\zeta}{5}}{1 - \frac{8\zeta^2}{5\zeta^2}} = -\frac{46}{9}\zeta. \quad (5.27)$$

也就是说, 小方认为粒子“戊”在向左运动, 而小李和小邓都认为粒子“戊”在向右运动. 桃花源宇宙确实是一个不同的世界, 在那里电磁波的行为也一

定非常古怪. 这里不再深入讨论.

让我们回到我们的宇宙, 在这里 $C = c^2$, 因此最后的变换是

$$t' = \frac{1}{\sqrt{1 - v_k^2/c^2}} \left(t - \frac{v_k}{c^2} x \right), \quad (5.28)$$

$$x' = \frac{1}{\sqrt{1 - v_k^2/c^2}} (x - v_k t). \quad (5.29)$$

这就是洛伦兹变换. 当相对速度 v_k 远远小于 c 时, 洛伦兹变换会退化成伽利略变换. 利用公式 (5.28, 5.29), 我们可以推导速度在不同惯性参照系之间的变换关系. 具体推导如下,

$$v'_x = \frac{dx'}{dt'} = \frac{v_x - v_k}{1 - v_k v_x / c^2}, \quad (5.30)$$

$$v'_y = \frac{dy'}{dt'} = \frac{dy}{dt'} = \frac{\sqrt{1 - v_k^2/c^2}}{1 - v_k v_x / c^2} v_y, \quad (5.31)$$

$$v'_z = \frac{dz'}{dt'} = \frac{dz}{dt'} = \frac{\sqrt{1 - v_k^2/c^2}}{1 - v_k v_x / c^2} v_z. \quad (5.32)$$

类似地还可以推导出加速度之间的变换关系, 这个关系相当复杂但不是特别重要, 有兴趣的读者可以尝试推导.

5.2 相对论效应

狭义相对论里有很多令人惊叹的超越日常经验的物理效应, 这既包括刚刚讨论过的光速不变也包括同时的相对性, 时间延缓和长度收缩. 时间延缓也被称为时间膨胀. 这些效应彻底改变了人们对时间和空间的看法. 我们将先从洛伦兹变换出发来讨论这些效应, 然后利用一个思想实验重新讨论时间延缓和长度收缩, 这个思想实验的装置和迈克尔逊-莫雷 (Michelson-Morley) 实验类似. 下面的讨论中, 我们依然忽略和相对速度垂直的两个空间维度.

考虑两个时钟, 甲和乙, 它们都相对于小李静止. 小李设法让这两个钟同步. 这不难做到: 测量这两个钟之间的距离 L_0 , 确认时钟甲的计时 t_0 , 向时钟

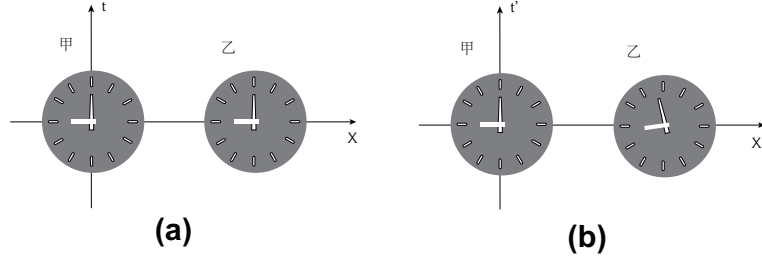


图 5.3: 两个时钟, 甲和乙. (a) 在小李的参照系里, 它们静止并且相距 L_0 , 并且同步计时. (b) 但是在小邓的参照系里 (相对小李向右以速度 v 运动), 它们并不同步.

乙发射光信号, 将时钟乙的时间设置为 $t_0 + L_0/c$. 自此以后, 只要两个时钟计时准确, 它们的指针将永远指向相同的时刻. 在将两个时钟同步以后, 通过适当选择时间和空间的原点, 在小李的惯性参考系 \mathcal{K}_2 里, 它们的世界线分别是 $(t, 0, 0, 0)$ 和 $(t, L_0, 0, 0)$. 根据洛伦兹变换, 在小邓的参考系 \mathcal{K}_3 里 (\mathcal{K}_3 和 \mathcal{K}_2 的时间和空间原点重合), 时钟甲的世界线是

$$t'_1 = \gamma t, \quad x'_1 = -\gamma vt, \quad (5.33)$$

而时钟乙的世界线是

$$t'_2 = \gamma(t - \frac{vL_0}{c^2}), \quad (5.34)$$

$$x'_2 = \gamma(L_0 - vt). \quad (5.35)$$

这里

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}}, \quad \beta = v/c. \quad (5.36)$$

注意: $t'_1 \neq t'_2$, 也就是说在小邓的眼里, 两个时钟并没有同步. 由此可见, 时钟之间的同步是相对的, 依赖惯性参照系的选择.

从上面的公式中还能很快得到一个有趣的结果, 时间延缓. 如果在参照系 \mathcal{K}_2 中, 时钟甲经过了时间 T_0 , 那么根据公式 (5.33) 的第一式, 我们发现在参照系 \mathcal{K}_3 中时钟甲走过的时间是

$$T = \frac{T_0}{\sqrt{1 - v^2/c^2}}. \quad (5.37)$$

注意这个公式中的 T_0 和 T 是同一个时钟，时钟甲，在不同惯性参考系中的两个时间，不是同一个惯性参考系中不同时钟的时间。从公式 (5.34) 可以看出，这个结论对时钟乙同样成立。因此，对于任意时钟，它走过的时间会依赖惯性参考系的选择：它相对于惯性参考系的速度越快，它走得越慢。这就是著名的时间延缓效应。

现在来看一下长度收缩效应。这要求测量时钟甲和时钟乙之间的距离。这个任务对小李比较简单，因为两个时钟相对于她都是静止的。她可以先确认时钟甲确实是在原点，在稍后的任意时刻 t 再确认时钟乙的位置 L_0 。由于时钟甲是静止的，而且已经和时钟乙同步，因此当小李在时刻 t 确认时钟乙的位置时，她知道在这个时刻时钟甲依然在原点，因此它们之间的距离是 L_0 。

对于小邓来说，测量两个时钟之间距离的任务就不是这么容易了。小邓需要确认在同一个时刻时钟甲的位置 X'_1 和时钟乙的位置 X'_2 ，然后利用它们的差得到二者之间的距离 $L = X'_2 - X'_1$ 。但是在小邓的参照系 \mathcal{K}_3 中，时钟甲和时钟乙现在不是同步的。因此小邓需要先将时钟甲和时钟乙同步。当同步完成以后，时钟甲在 \mathcal{K}_3 中的世界线是 $(t', -vt', 0, 0)$ ，时钟乙是 $(t', L - vt', 0, 0)$ 。 L 是参照系 \mathcal{K}_3 中两个时钟之间的距离。根据洛伦兹变换，在小李的参考系 \mathcal{K}_2 里时钟甲的世界线是

$$t_1 = \gamma(t' - \frac{v^2}{c^2}t') , \quad x_1 = \gamma(vt' - vt') = 0 , \quad (5.38)$$

而时钟乙的世界线是

$$t_2 = \gamma(t' + \frac{vL}{c^2}) , \quad (5.39)$$

$$x_2 = \gamma(L + vt') . \quad (5.40)$$

在参考系 \mathcal{K}_2 中，两个时钟都是静止的，因此它们的坐标一定要满足 $x_2 - x_1 = L_0$ 。利用这个关系，再结合公式 (5.38) 的第二式和公式 (5.40)，我们得到

$$L = \frac{x_2 - x_1}{\gamma} = \sqrt{1 - v^2/c^2} L_0 . \quad (5.41)$$

这就是小邓的测量结果，他发现两个时钟之间的距离相对于静止距离 L_0 变短了。这是长度收缩效应。另外，利用公式 (5.38) 的第一式和公式 (5.39)，

我们有

$$t_2 - t_1 = \frac{vL_0}{c^2}. \quad (5.42)$$

这表明虽然时钟甲和时钟乙在小邓的参照系 \mathcal{K}_3 中是同步的, 但是在小李的参照系里它们是不同步的. 这个公式也为小邓提供了一种在参照系 \mathcal{K}_3 中同步时钟甲和时钟乙的方式. 实验操作在静止参照系中总是相对简单和方便. 小邓可以让小李帮忙, 在静止参照系 \mathcal{K}_2 中有意把时钟乙根据上面这个公式往前拨一点. 这样在小邓参照系里, 这两个时钟就是同步的.

在上面的讨论和推导中, 有一个很明显的事实, 如果 $\beta = v/c \ll 1$, 那么所有的相对论效应, 同步的相对性、时间延缓和长度收缩, 都会很小, 以致可以忽略不计. 这正是我们在日常生活中看不到相对论效应的原因. 下面我们换一种更物理的方式推导时间延缓和长度收缩效应, 在这个推导里我们将尽量利用光速不变这个实验事实, 而不是直接利用洛伦兹变换.

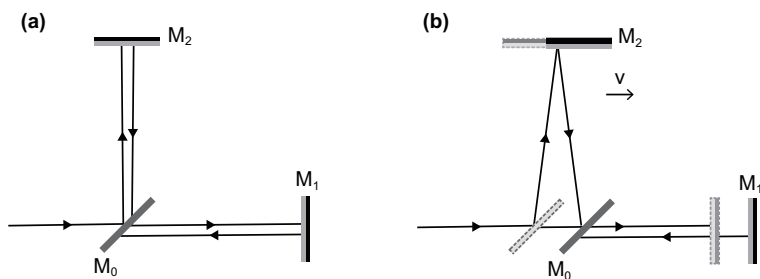


图 5.4: (a) 实验室惯性参照系; (b) 另一个惯性参照系, 在这个参照系中, 实验室以速度 v 向右运动. M_0 是一个半透镜, 光线经过后, 一半继续平行往前, 一半被反射后垂直向上; M_1 和 M_2 都是全反射镜. 在实验室坐标系中, M_0 和 M_1 之间的距离与 M_0 和 M_2 之间的距离一样, 都是 L_0 . 为了清楚显示, 反射回来的光线被人为平移了.

小邓正在他的箱子里做一个实验, 这个实验的装置和迈克尔逊-莫雷实验类似 (见图 5.4). 实验中, 光入射到半透镜 M_0 上, 被分为两束, 一束继续向前, 一束被反射后向上. 它们分别被反射镜 M_1 和 M_2 反射, 重新汇聚于镜子 M_0 . 在小邓的参照系 \mathcal{K}_3 中, 实验装置都是静止的. 小邓可以很容易测得

M_0 和 M_1 之间的距离, 以及 M_0 和 M_2 间的距离, 它们都是 L_0 . 同时测得光从 M_0 到 M_1 或 M_2 , 最后回到 M_0 的时间 $T_0 = 2L_0/c$.

在小李的参照系 \mathcal{K}_2 中, 情况有些不同. 如图5.4(b) 所示, 对于 M_0 到 M_2 之间的反射, 光走过的实际路径不再直上直下, 而是一条折线. 假设整个来回反射过程的时间是 T , 由于光速不变, 我们应该有

$$(cT/2)^2 = (vT/2)^2 + L_0^2. \quad (5.43)$$

由此可得

$$T = \frac{2L_0}{\sqrt{c^2 - v^2}} = \frac{T_0}{\sqrt{1 - (v/c)^2}}. \quad (5.44)$$

对于小李来说, 光线在 M_0 到 M_2 间来回的时间变长了. 这就是时间延缓.

现在考虑光在 M_0 到 M_1 之间的反射. 在小邓的参照系 \mathcal{K}_3 中, M_0 到 M_1 之间的距离显然是 L_0 ; 但是在小李眼里这个距离却不是 L_0 , 我们来看看为什么. 假设在小李的参照系 \mathcal{K}_2 中, M_0 和 M_1 之间的距离是 L . 由于镜子 M_1 在以速度 v 向前运动和光速不变原理, 光从 M_0 到达 M_1 的时间是

$$t_1 = \frac{L}{c - v}. \quad (5.45)$$

由于镜子 M_0 在以速度 v 向前运动和光速不变原理, 光从 M_1 回到 M_0 的时间是

$$t_2 = \frac{L}{c + v}. \quad (5.46)$$

在这个实验中光被分光镜 M_0 分为两束, 然后分别经历了两个不同的物理过程: 一束光在垂直方向运动; 另一束光在水平方向运动. 但是这两束光同时从 M_0 出发 (事件 1), 最后又同时回到 M_0 (事件 2). 这个基本物理事实—事件 1 和事件 2 发生了—不会随惯性参照系改变. 因此在小李的参照系 \mathcal{K}_3 中, 两束光依然同时从 M_0 出发; 同时回到 M_0 . 这样, 如果我们分别利用这两束光经历的不同过程去计算起点和终点之间的时间间隔, 它们必须相同. 于是我们有

$$t_1 + t_2 = \frac{L}{c + v} + \frac{L}{c - v} = T = \frac{T_0}{\sqrt{1 - v^2/c^2}}. \quad (5.47)$$

利用 $T_0 = L_0/c$, 我们最后得到

$$L = \sqrt{1 - v^2/c^2} L_0. \quad (5.48)$$

在小李看来, M_0 和 M_1 之间的距离变短了. 这就是长度收缩效应.

注意在上面的推导中, 我们也应用了相对性原理 (或洛伦兹变换), 只是没有明确提. 比如, 反射镜 M_1 在小邓的参照系里是静止的, 所以它在小李的参照系里做匀速直线运动. 这里隐含了相对性原理的应用.

考虑一个具体的例子. 宇航员小李驾驶飞船以速度 v 从地球到达冥王星, 地球和冥王星之间的距离是 d_0 . 宇航员小李携带了一个时钟记录自己的旅行时间, 地面控制人员小方也有一个时钟记录地面的时间. 为了方便我们不妨把小李的时钟叫时钟甲, 小方的时钟叫时钟乙. 这里有两个自然的惯性参照系: 地面参照系 \mathcal{K} 和飞船参照系 \mathcal{K}' . 设定它们的空间和时间原点在飞船出发的一刻是重合的. 对于地面参照系 \mathcal{K} , 飞船于时刻 $t_1 = 0$ 从 $x_1 = 0$ 出发, 在时刻 $t_2 = d_0/v$ 到达 $x_2 = d_0$. 洛伦兹变换告诉我们这两个事件在飞船参照系 \mathcal{K}' 中的坐标分别是, $t'_1 = 0$, $x'_1 = 0$ 和 $t'_2 = d_0/(v\gamma)$, $x'_2 = 0$. 现在考虑地面小方的时钟乙. 在地面参照系 \mathcal{K} 中, 飞船出发时它的时空坐标是 $\tilde{t}_1 = 0$, $\tilde{x}_1 = 0$, 飞船到达冥王星时, 它的时空坐标是 $\tilde{t}_2 = d_0/v$, $\tilde{x}_2 = 0$. 类似地利用洛伦兹变换, 我们有这两个事件在 \mathcal{K}' 中的坐标, $\tilde{t}'_1 = 0$, $\tilde{x}'_1 = 0$ 和 $\tilde{t}'_2 = d_0\gamma/v$, $\tilde{x}'_2 = -\gamma d_0$. 这里 $\gamma = 1/\sqrt{1 - v^2/c^2}$.

从上面的计算可以得出这样的结论. 地面的小方根据自己的时钟乙认为飞船飞行的时长是 d_0/v ; 宇航员小李根据飞船上的时钟甲, 认为飞行时长是 $d_0\sqrt{1 - v^2/c^2}/v$. 这是合理的, 因为小李认为路程只有 $d_0\sqrt{1 - v^2/c^2}$. 同时, 小李认为在整个旅程中, 地面的时钟乙记录的时长是 $d_0/(v\sqrt{1 - v^2/c^2})$.

这个例子虽然简单, 依然经常引起理解上的混乱. 为了避免理解上可能的混淆和矛盾, 我们必须牢记, 这个问题里有两位观察者, 小李和小方, 以及他们各自的时钟甲和乙. 在整个飞船从地球出发抵达冥王星的过程中

- 小李认为时钟甲记录的时长是 $\frac{d_0}{v}\sqrt{1 - v^2/c^2}$;

- 小方认为时钟甲记录的时长是 $\frac{d_0}{v}$;
- 小李认为时钟乙记录的时长是 $\frac{d_0}{v\sqrt{1-v^2/c^2}}$;
- 小方认为时钟乙记录的时长是 $\frac{d_0}{v}$.

无论是对于时钟甲还是时钟乙, 上面的结论和时间迟缓公式 (5.37) 都是一致的.

5.3 庞加莱-明可夫斯基空间

普通的三维空间, 它有两个基本特征: (1) 其中的点 \vec{r} 由三个坐标 (x, y, z) 确定; (2) 不同两个点, \vec{r} 和 \vec{r}' , 之间的距离由勾股定理给出, 即

$$|\vec{r} - \vec{r}'| = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}. \quad (5.49)$$

这个距离在旋转变换下不变. 在前面讨论四维时空时, 我们只讨论了其中的点, 也就是事件, 它由四个坐标 (t, x, y, z) 确定和描述. 在下面的讨论中我们会看到在事件之间也可以定义“距离”, 这个“距离”在洛伦兹变换下不变.

根据公式 (5.28, 5.29), 我们有

$$c^2 t'^2 - x'^2 = \frac{1}{1 - v_k^2/c^2} [(ct - v_k x/c)^2 - (x - v_k t)^2] = c^2 t^2 - x^2. \quad (5.50)$$

由于 $y' = y$ 和 $z' = z$, 上面等式意味着

$$c^2 t'^2 - x'^2 - y'^2 - z'^2 = c^2 t^2 - x^2 - y^2 - z^2. \quad (5.51)$$

这个结果提示我们做如下定义

$$s = \sqrt{c^2 t^2 - x^2 - y^2 - z^2}. \quad (5.52)$$

由于没有物体的速度会超过光速, 上面根号下的量永远是正的. s 被看作两个事件 $(0, 0, 0, 0)$ 和 (t, x, y, z) 之间的“距离”, 被称为时空间隔或事件间隔. 上面的计算表明 s 在洛伦兹变换下不变. 以 s 定义距离的四维时空被称为

庞加莱-明可夫斯基空间. 如果我们定义 $\tilde{s} = \sqrt{c^2 t^2 + x^2 + y^2 + z^2}$, 那么这个四维时空是我们熟悉的欧几里得空间. 但是这个空间不描述任何真实的物理世界, 因为 \tilde{s} 在洛伦兹变换下会改变. 有趣的是, 在桃花源宇宙里, 由于 $C = -c^2$, $\tilde{s} = \sqrt{c^2 t^2 + x^2 + y^2 + z^2}$ 在洛伦兹变换下不变. 因此桃花源宇宙里的四维时空是欧几里得空间.

利用时空间隔的洛伦兹变换不变性, 我们也可以推导出时间延缓效应. 考虑一个时钟, 它一动不动地处于惯性参照系 \mathcal{K} 的原点. 在 \mathcal{K} 中, 它的世界线为 $(\tau, 0, 0, 0)$; 在另外一个惯性参照系 \mathcal{K}' 中, 它的世界线为 $(t', -v_k t', 0, 0)$. 根据公式 (5.51), 我们有

$$c^2 \tau^2 = c^2 t'^2 - v_k^2 t'^2. \quad (5.53)$$

经过简单的变换, 我们可以进一步得到

$$t' = \frac{\tau}{\sqrt{1 - v_k^2/c^2}}. \quad (5.54)$$

和前面得到的时间延缓公式 (5.37) 完全一致. 静止时钟经历的时间 τ 一般被称为固有时间(proper time).

为了方便, 我们把庞加莱-明可夫斯基空间中的一个点记为 $\mathbf{x} = (x_0, x_1, x_2, x_3)$. 其中 $x_0 = ct$; x_1, x_2, x_3 是空间坐标, 分别对应前面的 x, y, z . 我们把 \mathbf{x} 称为四维位置, 它的所有分量都具有相同的物理量纲 (长度). 我们同时把洛伦兹变换 (5.28, 5.29) 改写成矩阵形式

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \gamma_k & -\gamma_k \beta_k & 0 & 0 \\ -\gamma_k \beta_k & \gamma_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (5.55)$$

其中 $\beta_k = v_k/c$, $\gamma_k = \gamma(v_k) = 1/\sqrt{1 - \beta_k^2}$. 引入

$$\mathbb{L}_1 = \begin{pmatrix} \gamma_k & -\gamma_k\beta_k & 0 & 0 \\ -\gamma_k\beta_k & \gamma_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.56)$$

洛伦兹变换可以进一步简写成 $\mathbf{x}' = \mathbb{L}_1 \mathbf{x}$. \mathbb{L}_1 的逆矩阵是

$$\mathbb{L}_1^{-1} = \begin{pmatrix} \gamma_k & \gamma_k\beta_k & 0 & 0 \\ \gamma_k\beta_k & \gamma_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.57)$$

这个形式和我们的物理预期一致. \mathbb{L}_1^{-1} 对应从参照系 \mathcal{K}' 到参照系 \mathcal{K} 中的洛伦兹变换. 由于 \mathcal{K}' 相对于 \mathcal{K} 的速度是 $-v_k$, \mathbb{L}_1^{-1} 应该具有以上形式.

定义一个对角矩阵

$$\mathbf{g} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (5.58)$$

这个矩阵被称为度规矩阵因为庞加莱-明可夫斯基空间中向量的长度可以写成

$$\begin{aligned} x_0^2 - x_1^2 - x_2^2 - x_3^2 &= \mathbf{x}^T \mathbf{g} \mathbf{x} \\ &= \begin{pmatrix} x_0 & x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned} \quad (5.59)$$

通过直接计算我们可以证明 $\mathbb{L}_1^T \mathbf{g} \mathbb{L}_1 = \mathbf{g}$.

$$\mathbf{x}'^T \mathbf{g} \mathbf{x}' = \mathbf{x}^T \mathbb{L}_1^T \mathbf{g} \mathbb{L}_1 \mathbf{x} = \mathbf{x}^T \mathbf{g} \mathbf{x}. \quad (5.60)$$

这就是前面介绍的不变量 (5.51).

如果相对速度是沿任意方向 $\mathbf{v} = (v_x, v_y, v_z)$, 那么可以证明洛伦兹变换矩阵具有下面的形式

$$\mathbb{L} = \begin{pmatrix} \gamma_k & -\gamma_k \frac{v_x}{c} & -\gamma_k \frac{v_y}{c} & -\gamma_k \frac{v_z}{c} \\ -\gamma_k \frac{v_x}{c} & 1 + (\gamma_k - 1) \frac{v_x^2}{v^2} & (\gamma_k - 1) \frac{v_x v_y}{v^2} & (\gamma_k - 1) \frac{v_x v_z}{v^2} \\ -\gamma_k \frac{v_y}{c} & (\gamma_k - 1) \frac{v_x v_y}{v^2} & 1 + (\gamma_k - 1) \frac{v_y^2}{v^2} & (\gamma_k - 1) \frac{v_y v_z}{v^2} \\ -\gamma_k \frac{v_z}{c} & (\gamma_k - 1) \frac{v_x v_z}{v^2} & (\gamma_k - 1) \frac{v_y v_z}{v^2} & 1 + (\gamma_k - 1) \frac{v_z^2}{v^2} \end{pmatrix} \quad (5.61)$$

它也满足 $\mathbb{L}^T \mathbf{g} \mathbb{L} = \mathbf{g}$.

5.4 相对论中的速度、动量和能量

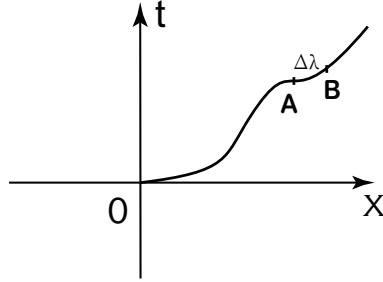


图 5.5: 庞加莱-明可夫斯基空间中的一条世界线. A 和 B 是曲线上两个相近的事件, $\Delta\lambda$ 是 AB 这段曲线的长度.

考虑一个粒子, 它的运动轨迹对应庞加莱-明可夫斯基空间中的一条世界线. 不失一般性, 假定运动的起点是坐标系的原点. 定义 $\lambda(t)$ 是从初始时刻到某时刻 t 粒子世界线的长度. 它和时间 t 显然有一对一的关系, 因此我们可以用它来代替时间 t , 把粒子的世界线表示为

$$\mathbf{x}(\lambda) = (x_0(\lambda), x_1(\lambda), x_2(\lambda), x_3(\lambda)). \quad (5.62)$$

如图 5.5 所示, 考察世界线上两个相近的事件 A 和 B, 它们之间世界线的长

度是 $\Delta\lambda$. 由此我们可以定义四个平均“速度”,

$$\bar{u}_0 = \frac{x_0^B - x_0^A}{\Delta\lambda} c, \quad \bar{u}_1 = \frac{x_1^B - x_1^A}{\Delta\lambda} c, \quad (5.63)$$

$$\bar{u}_2 = \frac{x_2^B - x_2^A}{\Delta\lambda} c, \quad \bar{u}_3 = \frac{x_3^B - x_3^A}{\Delta\lambda} c. \quad (5.64)$$

这里的 c 是为了让 \bar{u}_α 具有速度的量纲. 当 B 趋近于 A 时, 这段运动近似为一个匀速直线运动, 它的速度是粒子在 A 点的瞬时速度 \mathbf{v} . 想象一个参照系, 相对于现在的参照系以速度 \mathbf{v} 运动. 在这个参照系里, 粒子是近似静止的, 事件 A 和 B 之间世界线的长度是 $c\Delta\tau$, 这里的 τ 就是公式 (5.54) 中的固有时间. $\Delta\lambda$ 是事件 A 和 B 之间的事件间隔, 在洛伦兹变换下不变. 因此我们有 $\Delta\lambda \approx c\Delta\tau$. 这样当 B 趋近于 A 时, 公式 (5.63, 5.64) 就成为

$$u_0 = \frac{dx_0}{cd\tau} c = c \frac{dt}{d\tau}, \quad u_1 = \frac{dx_1}{cd\tau} c = \frac{dx}{dt} \frac{dt}{d\tau} = v_x \frac{dt}{d\tau}, \quad (5.65)$$

$$u_2 = \frac{dx_2}{cd\tau} c = \frac{dy}{dt} \frac{dt}{d\tau} = v_y \frac{dt}{d\tau}, \quad u_3 = \frac{dx_3}{cd\tau} c = \frac{dz}{dt} \frac{dt}{d\tau} = v_z \frac{dt}{d\tau}. \quad (5.66)$$

根据公式 (5.54) (注意这个公式里的 t' 是这里的 t), 我们有 $dt/d\tau = \gamma(v)$, 其中 $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$. 最后我们得到

$$u_0 = \gamma(v)c, \quad u_1 = \gamma(v)v_x, \quad u_2 = \gamma(v)v_y, \quad u_3 = \gamma(v)v_z. \quad (5.67)$$

这个速度 $\mathbf{u} = (u_0, u_1, u_2, u_3)$ 通常被称为四维速度, 它有一个非常好的性质, 满足洛伦兹变换.

事实上不难看出 \mathbf{u} 满足洛伦兹变换. 根据上面的讨论, 我们有

$$\mathbf{u} = c \frac{d\mathbf{x}}{d\tau}. \quad (5.68)$$

显然, $d\mathbf{x}$ 在不同惯性参照系之间的变换是洛伦兹变换. 同时 $d\tau$ 是洛伦兹变换不变量, c 是一个常数, 因此 $\mathbf{u} = (u_0, u_1, u_2, u_3)$ 满足洛伦兹变换, 即 $\mathbf{u}' = \mathbb{L}_1 \mathbf{u}$. 展开后的具体矩阵形式如下

$$\begin{pmatrix} u'_0 \\ u'_1 \\ u'_2 \\ u'_3 \end{pmatrix} = \begin{pmatrix} \gamma_k & -\gamma_k \beta_k & 0 & 0 \\ -\gamma_k \beta_k & \gamma_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix}. \quad (5.69)$$

可以验证这个速度的变换关系和公式 (5.30, 5.31, 5.32) 是一致的. 在验证时, 必须注意一点, 上式中有三个不同的 γ ,

$$\gamma' = \frac{1}{\sqrt{1 - v'^2/c^2}}, \quad \gamma_k = \frac{1}{\sqrt{1 - v_k^2/c^2}}, \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \quad (5.70)$$

其中 $v'^2 = v_x'^2 + v_y'^2 + v_z'^2$, $v^2 = v_x^2 + v_y^2 + v_z^2$. v 和 v' 分别是粒子在惯性参照系 \mathcal{K} 和 \mathcal{K}' 中的速度, v_k 则是参照系 \mathcal{K}' 相对于 \mathcal{K} 的速度. 利用公式 (5.30, 5.31, 5.32), 我们可以得到

$$\begin{aligned} c^2 - v'^2 &= c^2 - \frac{(v_x - v_k)^2 + (1 - v_k^2/c^2)v_y^2 + (1 - v_k^2/c^2)v_z^2}{(1 - v_kv_x/c^2)^2} \\ &= \frac{c^2(1 - v_kv_x/c^2)^2 - (v_x - v_k)^2 - (1 - v_k^2/c^2)(v^2 - v_x^2)}{(1 - v_kv_x/c^2)^2} \\ &= \frac{(c^2 - v^2)(1 - v_k^2/c^2)}{(1 - v_kv_x/c^2)^2}. \end{aligned} \quad (5.71)$$

这意味着

$$\frac{\gamma}{\gamma'} = \frac{1}{(1 - v_kv_x/c^2)\gamma_k}. \quad (5.72)$$

利用这个公式, 我们可以很容易从公式 (5.30, 5.31, 5.32) 推导出 $\mathbf{u}' = \mathbb{L}_1 \mathbf{u}$, 即公式 (5.69).

虽然公式 (5.30, 5.31, 5.32) 中的速度变换和公式 (5.69) 在数学上是等价的, 但公式 (5.69) 的优势是显而易见的. 首先, 公式 (5.69) 更简洁; 更重要的是它告诉我们, 除了 \mathbf{x} 是一个满足洛伦兹变换的四维向量外, \mathbf{u} 也是. 既然 \mathbf{u} 是四维向量, 那么它的长度

$$u_0^2 - u_1^2 - u_2^2 - u_3^2 = \mathbf{u}^T \mathbf{g} \mathbf{u} \quad (5.73)$$

是一个洛伦兹变换下的不变量. 如果 \mathbf{u} 描述一个粒子, 它的速度 $v = \sqrt{v_x^2 + v_y^2 + v_z^2} < c$. 我们总是可以选择一个惯性参照系, 粒子在这个参照系的速度是零, 也就是它的四维速度是 $(c, 0, 0, 0)$. 在这个参照系里

$$u_0^2 - u_1^2 - u_2^2 - u_3^2 = c^2. \quad (5.74)$$

由于这是洛伦兹变换下的不变量，这个等式在所有惯性参照系中成立。

事实上，还有很多由其他物理量构成的四维向量。四维动量 \mathbf{p} 就是其中的一个，它的定义为 $\mathbf{p} = \tilde{m}\mathbf{u}$ ，即 $p_\alpha = \tilde{m}u_\alpha$ 。这里 \tilde{m} 是一个具有质量量纲的常数。我们引入 $E = \tilde{m}u_0c$ 将 p_0 改写成 $p_0 = E/c$ 。注意 E 具有能量的量纲，我们马上会看到它确实是一个自由粒子（也就是不受任何外力的粒子）的能量。考虑一个质量为 m_0 的做匀速直线运动的粒子，其速度 v 远远小于光速 c ，即 $v \ll c$ 。对于这个粒子，我们有

$$\begin{aligned} E = \tilde{m}u_0c &= \tilde{m}\gamma c^2 = \frac{\tilde{m}c^2}{\sqrt{1-v^2/c^2}} \\ &\approx \tilde{m}c^2\left(1 + \frac{v^2}{2c^2} + \frac{3v^4}{8c^4} + \cdots\right) \end{aligned} \quad (5.75)$$

我们让 $\tilde{m} = m_0$ ，上面第二行右边的第二项就是我们熟悉的粒子的动能。所以在相对论里，一个运动速度为 v 质量为 m_0 自由粒子的能量是

$$E = \gamma m_0 c^2 = \frac{m_0 c^2}{\sqrt{1-v^2/c^2}}. \quad (5.76)$$

令 $m = \gamma m_0$ ，上式简化为 $E = mc^2$ 。这就是著名的质能公式。如果粒子静止，那么它的能量是 $E = m_0 c^2$ 。 m_0 因此被称为静止质量；相应地， m 被称为运动质量。对于四维动量 \mathbf{p} 的其它三个分量，我们有

$$p_1 = \gamma m_0 v_x = m v_x = p_x, \quad (5.77)$$

$$p_2 = \gamma m_0 v_y = m v_y = p_y, \quad (5.78)$$

$$p_3 = \gamma m_0 v_z = m v_z = p_z. \quad (5.79)$$

它们就是我们熟悉的动量，唯一的差别是质量 m 会随运动速度变化。

基于以上分析，我们重写四维动量 \mathbf{p} 的各个分量 $\mathbf{p} = (E/c, p_x, p_y, p_z)$ 。在这里，能量和动量具有同等的地位，是同一个向量的四个分量。这和四维位置 \mathbf{x} 形成呼应，在 \mathbf{x} 中时间和空间位置具有同等的地位，是同一个向量的四个分量。由于 \mathbf{p} 和 \mathbf{u} 只差一个常数系数 m_0 ，四维动量 \mathbf{p} 显然满足洛伦兹

变换. 选择一个惯性参照系它对应的不变量是

$$E^2 - p_x^2 c^2 - p_y^2 c^2 - p_z^2 c^2 = m_0^2 c^4. \quad (5.80)$$

上式左边是一个质量为 m_0 的做匀速直线运动的粒子在能量-动量四维空间中“长度”的一般表达式. 在相对这个粒子静止的参照系里, 由于粒子速度为零, 这个“长度”只有一项, 即上式右边. 将上式左边的后三项移到右边, 我们有

$$E^2 = m_0^2 c^4 + p_x^2 c^2 + p_y^2 c^2 + p_z^2 c^2. \quad (5.81)$$

从这里可以清楚地看到, 一个粒子的能量有两部分: (1) 和静止质量相关的一部分, 这部分能量和运动完全无关; (2) 和运动相关的一部分. 令 $p^2 = p_x^2 + p_y^2 + p_z^2$, 上式简化为

$$E^2 = m_0^2 c^4 + p^2 c^2. \quad (5.82)$$

任何一个自由粒子的能量和动量都满足这个关系. 如果一个粒子的静止质量为零, 即 $m_0 = 0$, 我们有 $E = pc$. 光的静止质量是零, 所以光具有这个简洁的能量和动量关系.

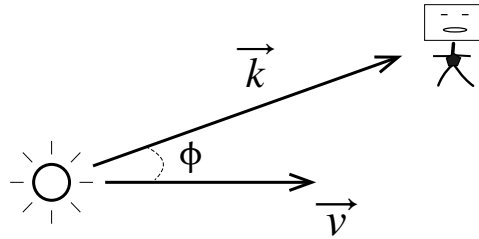


图 5.6: 光的多普勒效应. 光源相对于观察者小方以速度 \vec{v} 运动, \vec{k} 是传向小方的光的波矢.

根据麦克斯韦的电磁理论, 光是电磁波, 具有频率 ω 和三个波矢 k_x, k_y, k_z . 利用麦克斯韦方程, 可以严格证明在和光速 c 恰当组合后, 这四个量也可以组成一个四维矢量, 满足洛伦兹变换. 有意思的是利用量子力学可以更

直接地得到这个结论. 根据量子力学, 光是由一个一个光子构成的. 对于一个频率为 ω 波长为 λ 的光子, 它的能量是 $E = \hbar\omega$, 动量是 $\vec{p} = \hbar\vec{k}$, 其中 $k = |\vec{k}| = 2\pi/\lambda$ 是光的波矢. 利用光的能量和动量关系 $E = pc$, 我们有 $\omega = 2\pi c/\lambda$, 这正是大家熟知的光的频率和波长关系. 由于 E/c 和 \vec{p} 组成一个四维矢量, 而 \hbar 是一个自然常数和参考系无关, 这意味着光的频率 ω 和波矢 \vec{k} 正好构成一个四维矢量 $\mathbf{k} = (\omega, k_x c, k_y c, k_z c)$. 下面利用这个事实讨论光的多普勒效应.

考虑一个光源 S , 它发出的光被远处的一个接收器 R 接收. 如果光源 S 静止, 它发出的光的角频率是 ω_0 波长是 λ_0 , 这也是接收器接收到的频率和波长. 但是如果光源 S 在运动, 那么接收器接收到的光的频率会发生变化, 这就是光的多普勒效应. 不失一般性, 我们假设光源 S 的速度 \vec{v} 沿 x 轴方向. 由于光源在动, 所以光源参照系里的四维波矢 \mathbf{k}' 和接收器参照系里的四维波矢 \mathbf{k} 之间满足如下洛伦兹变换,

$$\omega' = \gamma(\omega - \frac{v}{c} k_x c) = \gamma(\omega - vk \cos \phi). \quad (5.83)$$

这里 ϕ 是波矢 \vec{k} 和波源速度 \vec{v} 之间的夹角 (参见图5.6). 注意光源参照系里的频率 ω' 就是光源静止时光的频率 ω_0 , 同时利用 $k = \omega/c$, 我们有

$$\omega = \frac{\sqrt{1 - v^2/c^2}}{1 - v \cos \phi/c} \omega_0. \quad (5.84)$$

这就是光频率的多普勒公式.

5.5 相对论力学

在上一节, 我们通过对 \mathbf{x} 微分得到了四维动量 \mathbf{p} . 继续对 \mathbf{p} 微分, 我们会得到另外一个四维物理量, 它具有力的量纲. 这就是四维力, 它的具体形式是

$$\mathbf{F} = \frac{d\mathbf{p}}{d\tau} = \gamma \frac{d\mathbf{p}}{dt}. \quad (5.85)$$

上面利用了 $dt/d\tau = \gamma$. 注意到前面四维速度的三个空间分量和普通的三维速度有这个关系, $u_{1,2,3} = \gamma v_{x,y,z}$. 类似地, 我们应该有

$$F_{1,2,3} = \gamma f_{x,y,z}, \quad (5.86)$$

这里 $\vec{f} = (f_x, f_y, f_z)$ 是普通的三维力. 这样从公式 (5.85) 我们就得到了符合相对论的牛顿第二定律

$$\vec{f} = \frac{d\vec{p}}{dt} = \frac{d}{dt} \left(\frac{m_0 \vec{v}}{\sqrt{1 - v^2/c^2}} \right). \quad (5.87)$$

现在考察 F_0 的物理意义. 根据公式 (5.85), 我们有

$$F_0 = \frac{\gamma}{c} \frac{dE}{dt}, \quad (5.88)$$

其中的 E 是粒子的静止能量加动能, 满足 $E^2 = m_0^2 c^4 + p^2 c^2$. 对这个等式两边求导,

$$E \frac{dE}{dt} = \vec{p} \cdot \frac{d\vec{p}}{dt} c^2 = m \vec{v} \cdot \vec{f} c^2 = E \vec{v} \cdot \vec{f}. \quad (5.89)$$

进一步整理,

$$\frac{dE}{dt} = \vec{f} \cdot \vec{v}. \quad (5.90)$$

由于粒子的静止能量不会改变, 这个式子有非常明确的物理含义: 粒子动能的改变等于外力做的功. 结合上面的结果, 我们有

$$F_0 = \gamma \vec{v} \cdot \vec{f} / c. \quad (5.91)$$

这表明 F_0 正比于外力 \vec{f} 在单位时间里做的功, 即功率.

这里讨论的相对论力学在研究高速带电粒子在电磁场中的运动时非常有用, 在日常生活中的应用非常有限. 主要原因是日常生活中物体的速度远小于光速, 同时也忽视传递力所需要的时间.

5.6 相对论碰撞

高能加速器现在是人类探测微观基本粒子最有力的实验手段. 加速器中的粒子速度都接近光速, 它们之间的碰撞必须在相对论的框架下进行处理. 在高能加速器中经常发生两种碰撞, 一束高能粒子轰击一个静止的靶或者两束高能粒子对撞. 我们将主要讨论这两种碰撞.

考虑两个粒子，它们的静止质量分别为 m_{01} 和 m_{02} . 相对于加速器，粒子 1 的速度是 \vec{v}_1 ，粒子 2 静止. 所以这个系统的总能量是

$$E = E_1 + E_2 = m_1 c^2 + m_{02} c^2, \quad (5.92)$$

总动量是

$$\vec{p} = \vec{p}_1 + \vec{p}_2 = m_1 \vec{v}_1. \quad (5.93)$$

在相对论碰撞中，由于粒子的质量会随速度显著变化，质心和质心参照系不再是非常有用的概念和工具. 取而代之的是动量中心参照系(center of momentum)，在这个惯性参照系中系统的总动量为零，即

$$\vec{p}' = m_1' \vec{v}_1' + m_2' \vec{v}_2' = 0. \quad (5.94)$$

在动量中心参照系中系统的总能量是

$$E' = m_1' c^2 + m_2' c^2. \quad (5.95)$$

不失一般性，我们假设粒子 1 的速度 \vec{v}_1 是沿 x 方向. 这样 \vec{v}_1' 和 \vec{v}_2' 也是沿 x 方向. 由于粒子 2 在加速器参照系中是静止，所以如果动量中心相对加速器的速度是 \vec{v} ，那么 $\vec{v} = -\vec{v}_2'$. 另外，根据相对论速度相加规则，粒子 1 在两个参照系中的速度有如下关系

$$v_1' = \frac{v_1 - v}{1 - v_1 v / c^2}. \quad (5.96)$$

利用这个关系，并结合上面的公式 (5.94)，我们可以确定动量中心参照系相对于加速器的速度 v . 在加速器参照系，这个系统四维动量 \mathbf{p} 的长度是

$$\mathbf{p}^T \mathbf{g} \mathbf{p} = \frac{E^2}{c^2} - \vec{p} \cdot \vec{p} = m_{01}^2 c^2 + m_{02}^2 c^2 + 2m_1 m_{02} c^2. \quad (5.97)$$

在动量中心参照系，这个系统四维动量 \mathbf{p}' 的长度是

$$\mathbf{p}'^T \mathbf{g} \mathbf{p}' = \frac{E'^2}{c^2} = m_1'^2 c^2 + m_2'^2 c^2 + 2m_1' m_2' c^2. \quad (5.98)$$

这个长度在洛伦兹变换下不变，上面两式相等

$$\frac{E'^2}{c^4} = m_1'^2 + m_2'^2 + 2m_1' m_2' = m_{01}^2 + m_{02}^2 + 2m_1 m_{02}. \quad (5.99)$$

这个等式将在下面被用来计算基本粒子碰撞反应的阈值.

在加速器中, 只要给两个粒子足够的动能, 这两个粒子会利用这些动能在碰撞后产生更多的粒子. 比如在历史上物理学家曾经在加速器上利用下面这个碰撞产生了反质子 \bar{p} ,

$$p + p \rightarrow p + p + p + \bar{p} \quad (5.100)$$

如果碰撞前两个质子相对加速器都是静止的, 这个反应显然不会发生. 只有通过加速其中的一个质子或同时加速两个质子, 给它们足够动能才能保证这个反应发生. 在设计和运行加速器时, 物理学家需要知道这个动能的最小值, 也就是碰撞反应阈值. 动量中心参照系为计算这个碰撞动能阈值提供了方便.

假设碰撞后产生了 n 个粒子, 它们的静止质量分别是 $m_{01}, m_{02}, \dots, m_{0n}$. 由于动量守恒, 在动量中心参照系里, 这些粒子的总动量是零. 如果碰撞前两个粒子的动能正好是阈值, 那么产生的这些粒子在动量中心参照系里应该正好是静止的. 动量中心系中碰撞前系统的能量是 E' , 因此根据能量守恒我们有

$$E' = \left(\sum_{r=1}^n m_{0r} \right) c^2. \quad (5.101)$$

先考虑前面讨论的情况, 碰撞前粒子 1 具有速度 \vec{v}_1 , 粒子 2 静止. 这种情况下, 碰撞前的总动能是

$$T = m_1 c^2 - m_{01} c^2. \quad (5.102)$$

利用公式 (5.99), 我们有

$$E'^2 = m_{01}^2 c^4 + m_{02}^2 c^4 + 2(T + m_{01} c^2) m_{02} c^2 = (m_{01} + m_{02})^2 c^4 + 2T m_{02} c^2. \quad (5.103)$$

当 T 正好是阈值 T_1 时, 这个公式中的 E' 和公式 (5.101) 中的 E' 相等, 因此

$$\frac{T_1}{m_{01} c^2} = \frac{\left(\sum_{r=1}^n m_{0r} \right)^2 - (m_{01} + m_{02})^2}{2m_{01} m_{02}}. \quad (5.104)$$

对于上面的质子碰撞 (5.100), 我们有

$$T_1 = 6m_p c^2. \quad (5.105)$$

这里 m_p 是质子的静止质量. 也就是说为了产生质子和反质子对, 至少需要把其中一个质子加速到它的动能是它静止质量的 6 倍.

现在考虑两个粒子对撞的情况, 进一步假设两个粒子有相同的静止质量 m_0 而且它们的速度大小相等且方向相反. 这种情况下, 加速器参考系和动量中心参考系是同一个参考系, 系统能量

$$E = E' = 2T + 2m_0 c^2. \quad (5.106)$$

由此得到对撞的动能阈值

$$\tilde{T}_1 = \left(\sum_{r=1}^n m_{0r} \right) c^2 / 2 - m_0 c^2. \quad (5.107)$$

因此利用质子和质子对撞去产生质子-反质子对, 动能的阈值是

$$\tilde{T}_1 = m_p c^2. \quad (5.108)$$

比前面用运动质子轰击静止质子小了接近一个数量级. 因此现在的高能加速器都是对撞机. 物理上这不难理解. 在加速器参照系, 系统的总动量不会改变, 如果一开始不是零就永远不是零, 这部分动能不能被转化为其它能量, 比如产生新的粒子.

习题

题 5.1 — 如图5.7所示, 考虑排成一列的物体 m_1, m_2, \dots, m_n , 它们之间没有任何相互作用. m_{j+1} 相对于 m_j 的速度是 u ($0 < u < c$), u 为正表示方向向右. 假设在实验室参照系中, m_1 的速度是 v_1 ($0 < v_1 < c$). 根据相对论速度相加规则 (5.24), 在实验室参照系中 m_2 的速度是

$$v_2 = \frac{u + v_1}{1 + uv_1/C},$$

物体 m_{j+1} 的速度是

$$v_{j+1} = \frac{u + v_j}{1 + uv_j/C}.$$

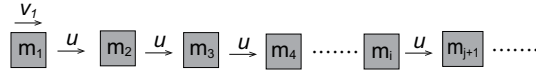


图 5.7: 质量相同的一系列物体, 相邻物体之间的相对速度是 u .

- (1) 如果 $C = c^2$, 证明 $\lim_{n \rightarrow \infty} v_n = c$.
- (2) 如果 $C = -c^2$, 用解析方法证明或数值计算演示极限 $\lim_{n \rightarrow \infty} v_n$ 不存在.

题 5.2 — 在惯性参照系 \mathcal{K} 中有一个静止的木棒, 它的长度是 L_0 , 和 x 轴正向的夹角是 $\theta < \pi/2$. 另有一个惯性参照系 \mathcal{K}' 相对于 \mathcal{K} 沿 x 轴以速度 v 运动. 求惯性参照系 \mathcal{K}' 中木棒的长度以及和 x 轴的夹角.

题 5.3 — 在一次观测中, 物理学家发现一个 μ 介子从产生到衰变穿越了 5 公里大气层, 它的速率是 $v = 0.9c$ (c 是光速).

- (1) 物理学家测量的 μ 介子的寿命是多长? 在相对 μ 介子静止的惯性参照系里, μ 介子的寿命是多长?
- (2) 在相对 μ 介子静止的惯性参照系里, 该介子穿过了多厚的大气层.

题 5.4 — 一个粒子在惯性参照系 \mathcal{K} 中的速度是 $\vec{v} = (v_x, v_y, v_z)$, 在另外一个惯性参照系 \mathcal{K}' 中的速度是 $\vec{v}' = (v'_x, v'_y, v'_z)$. 参照系 \mathcal{K}' 相对于 \mathcal{K} 沿 x 轴运动, 速度是 v_k . 证明速度变换关系 (5.30, 5.31, 5.32) 和四维速度 $\mathbf{u} = (u_0, u_1, u_2, u_3)$ 在参照系 \mathcal{K} 和 \mathcal{K}' 之间的洛伦兹变换是一致的.

题 5.5 — 在图5.4所示的实验中, 只考虑光束在 M_0 和 M_2 之间的运动. (a) 在实验室参照系, 光直上直下来回反射; (b) 在另一个参照系中, 光的运动线路则是一条折线. 利用速度变换公式 (5.30, 5.31, 5.32) 证明, 当光沿 (b) 中的折线运动时, 它的速度确实仍然是 c .

题 5.6 — 质子的静止质量是 $m_p = 938\text{MeV}$. 在宇宙射线中检测到一个能量为 10^{10}GeV 的质子. 假设具有这个能量的质子穿过一个直径为 10^5 光年的星系. 在相对质子静止的参照系中测量, 这需要多少时间? (由于 $E = mc^2$, 在高能物理里, 能量和质量经常被当作一个物理量. 因此虽然质子的静止质量是 $m_p = 938\text{MeV}/c^2$, 但人们经常省略 c^2 .)

题 5.6 — (20 分) 一个伽马光子轰击相对实验室静止的质子产生质子-反质子对,

$$\gamma + p \longrightarrow p + p + \bar{p}.$$

问光子必须具有的最小能量 E 是多少? (以质子静止能量 $m_p c^2$ 表达.) 和第5.6节中产生质子-反质子对的两个阈值进行比较.