

《场论》笔记

张大鹏

2016 年 7 月 11 日

目 录

第一章 相对性原理	1	3.6 Christoffel 符号与度规张量的关系	28
1.1 相互作用的传播速度	1	3.7 引力场中粒子的运动	28
1.2 间隔	2	3.8 恒定引力场	28
1.2.1 事件的间隔	2	3.9 旋转	28
1.2.2 类时间隔、类空间隔和类光间隔	4	第四章 引力场方程	29
1.3 固有时	6	4.1 曲率张量	29
1.4 Lorentz 变换	7	4.2 曲率张量的特性	29
1.5 速度的变换	9	4.3 引力场的作用量	29
1.6 四维矢量和四维张量	10	4.4 能量动量张量	29
1.6.1 四维矢量	11	4.5 Einstein 方程	29
1.6.2 四维张量	13	4.6 引力场的能动张量	29
1.6.3 四维张量分析	16	4.7 同步参考系	29
1.7 四维速度	18	4.8 Einstein 方程的标架表示	29
第二章 相对论力学	20	第五章 引力物体的场	30
2.1 Hamilton 原理	20	5.1 Newton 定律	30
2.2 能量与动量	21	5.2 中心对称的引力场	30
2.3 分布函数的变换	24	5.3 中心对称的引力场中的运动	30
2.4 粒子的衰变	25	5.4 球形物体的引力坍缩	30
2.5 粒子的弹性碰撞	26	5.5 类尘埃球的坍缩	30
2.6 角动量	27	5.6 非球形转动物体的引力坍缩	30
第三章 引力场中的粒子	28	5.7 物體遠距離處的引力場	30
3.1 非相对论力学中的引力场	28	5.8 二級近似下物體系統的運動方程	30
3.2 相对论力学中的引力场	28	插图目录	31
3.3 曲线坐标	28	参考文献	32
3.4 距离与时间间隔	28		
3.5 协变微分	28		

第一章 相对性原理

1.1 相互作用的传播速度

为了描述发生的事件，必须有一个所谓的**参考系**。与经典力学中的参考系不同的是，现在的参考系不仅需要有一个坐标系，还需要一个固定在其中的钟。此处坐标系用来刻画其中物体的位置，而钟用来指示时间。

经典力学中已经提到过所谓 **Galileo 相对性原理**：力学定理在所有惯性系中形式相同。一切惯性系等价，无法通过力学实验区分哪个惯性系更基本。换句话说，表示力学定律的方程对于由一个惯性系到另一个惯性系的时间和坐标各种变换来说是不变的。也就是说，描述力学定律的方程，用不同的惯性系的坐标与时间写出来，将具有相同的形式。而且，有理由认为：所有的自然定律在所有惯性系中都是相同的，即所谓的**相对性原理**。

粒子间的相互作用在经典力学中是瞬时传播的，然而，实验表明，瞬时的相互作用在自然界中是不存在的。实际上，如果相互作用的物体中的一个发生任何变动，仅仅在过了一段时间之后才能影响到其他物体。用这两个物体之间的距离除以这段时间，就得到**相互作用的传播速度**。

这个速度，严格地说，应该称为相互作用的最大传播速度。这个速度仅仅决定于某一物体的变动开始在第二个物体上表现出来所需要的时间。而相互作用的最大传播速度的存在，也就暗示着，自然界中物体运动的速度一般不可能大于这个速度。假如真的有这种运动存在，我们可以利用这种运动实现一个相互作用，使得此相互作用的传播速度大于前面的最大传播速度。从一个粒子向另一个粒子传播的相互作用往往叫做“信号”，它由第一个粒子发出，将第一个粒子所经历的变化“通知”第二个粒子。因此相互作用的传播速度也可以称为**信号速度**。根据相对性原理可以推断相互作用的传播速度在所有惯性系中都是相同的，即相互作用的传播速度是一个普适常数。

后面将会证明，这个恒定速度就是光在真空中的速度。**光速**常用 c 来表示，其值等于

$$c = 2.998 \times 10^8 \text{ m/s}$$

这个速度很大，因此经典力学在大多数情况下都足够精确。

把相对性原理和相互作用传播速度的有线性结合起来，就得到 **Einstein 相对性原理**，它不同于 Galileo 相对性原理，Galileo 的相对性原理基于无限大的相互作用传播速度。

以 Einstein 相对性原理（简称为相对性原理）为基础的力学，称为**相对论力学**。在运动物体的速度远小于光速的极限情形下，可略去相互作用传播速度有限对于运动的影响，此时相对论力学就变成**经典力学**。在相对论力学的公式中，取 $c \rightarrow \infty$ 的极限，就可由相对论力学在形式上过渡到经典力学。

很容易说明，绝对时间的概念是与 Einstein 相对性原理完全冲突的。在以绝对时间为基础的经典力学中，速度合成的通用法则是有效的，即复合运动的速度等于这个运动的各个速度的矢量和。将此法则用于相互作用的传播即可得出，互相作用在不同的惯性系中的

传播速度必然是不同的，这就与相对性原理冲突了。但是，实验则完全证实了相对性原理，1881 年 Michelson-Morley 实验表明，光速与其传播方向并无关系；然而，按照经典力学，光速在与地球运动相同的方向上，应该比在与地球运动方向相反的方向上小。

因此，相对性原理直接导出时间不是绝对的，即在不同的参考系中，时间的流逝也是不同的。所以，“两个不同的事件之间有一定的时间间隔”这样的陈述，仅在指定了参考系的情况下才有意义。特别地，在某一个参考系中同时发生的事件，对另一个参考系来说并不一定是同时的。

设有两个惯性系 K 和 K' ，其坐标轴分别为 xyz 和 $x'y'z'$ ，而 K' 则相对于 K 沿 x 和 x' 轴向右运动（如图 1.1 所示）。

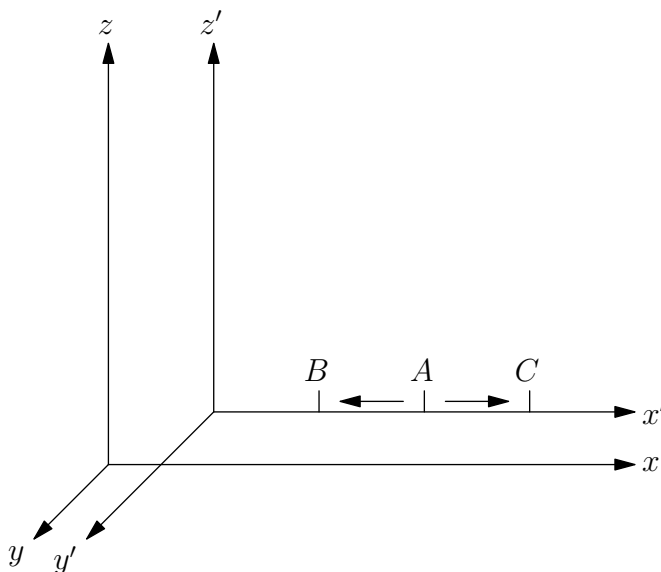


图 1.1: 同时的相对性

设信号从 x' 轴上某一点 A 向两个相反的方向发出，既然信号在 K' 系中的传播速度在所有的惯性系中都一样，在两个方向上都等于 c ，因此 K' 系中的观察者将观察到信号同时到达 B 点和 C 点。

然而在 K 系的观察者看来，同样的两个事件（即信号到达 B 点和 C 点）并不是同时的。按照相对性原理，信号相对于 K 系的速度也是 c ，而且因为 B 点在 K 系中向着信号发出的位置移动，而 C 点则背离信号发出的位置，所以在 K 系中，信号到达 B 点要比到达 C 点早。

1.2 间隔

1.2.1 事件的间隔

一个事件是由其发生的地点及其发生的时间来描述的，因此在某一实物粒子上所发生的事件可由粒子的三个坐标以及事件发生的时间来决定。

为表述便利起见，建立一个假象的四维空间，在此四维空间的四个轴中，三个用来刻画位置坐标，一个用来标示时间。在这个空间内，事件可用点来表示，这个点称为世界点。在

这个假象的四维空间内，每个粒子都对应于一条线，称为**世界线**。这条在线的各点决定了粒子在所有时刻的坐标。

现在考虑用数学形式来表示光速不变原理。为此，我们考虑两个彼此以恒定速度作相对运动的参考系 K 和 K' 。其中两个坐标系取为 x 轴和 x' 轴重合，而 y 轴和 z 轴则分别与 y' 轴和 z' 轴平行，并以 t 和 t' 分别表示在 K 和 K' 参考系内的时间。

设第一个事件是：在 K 系内的 t_1 时刻从具有坐标 x_1, y_1, z_1 （在同一参考系中）的点发出一个光速传播的信号。再设第二个事件是：信号在 t_2 时刻到达点 x_2, y_2, z_2 。由于信号传播的速度是 c ，因此可得两个事件坐标关系

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = 0 \quad (1.1)$$

而同样的两个事件，也可以在 K' 系中观察。设第一个事件在 K' 内的坐标为 x'_1, y'_1, z'_1, t'_1 ，而第二个事件的坐标则为 x'_2, y'_2, z'_2, t'_2 。根据光速不变原理，在 K' 系中信号传播的速度也是 c ，因此可有

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = 0 \quad (1.2)$$

假如 x_1, y_1, z_1, t_1 和 x_2, y_2, z_2, t_2 是任意两个事件的坐标，则

$$s_{12} = \sqrt{c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2} \quad (1.3)$$

称为这两个事件的**间隔**。因此，由光速不变原理可得，如果两个事件的间隔在某一个坐标系中为零，那么它在所有其他坐标系中均为零。

如果两个事件彼此无限接近，那么其间隔 ds 将满足

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (1.4)$$

从数学形式上来看，式(1.3)和(1.4)的形式可以设想为四维空间内两点之间的距离（该空间的四个轴分别为 x, y, z 和 ct ）。但是与普通几何的法则之间有一个根本区别，即间隔的平方和中沿不同轴的坐标差平方是以相异而非相同的运算符号求和的。^①

前面已经证明，如果在某一惯性系中有 $ds = 0$ ，则在任意其他惯性系中同样有 $ds' = 0$ 。此外， ds 和 ds' 为同阶的两个无穷小量。由此可知， ds^2 与 ds'^2 必须成比例，即

$$ds^2 = a ds'^2 \quad (1.5)$$

而且其中系数 a 仅与两个惯性系的相对速度大小有关。系数 a 不可能与坐标或时间有关系，否则就违背了空间和时间的均匀性。系数 a 也不可能与惯性系的相对速度方向有关，否则就违背了空间的各向同性。

考虑三个参考系 K, K_1, K_2 ，令 V_1, V_2 为 K_1, K_2 相对于 K 的速度，此时有

$$ds^2 = a(V_1) ds_1^2, \quad ds^2 = a(V_2) ds_2^2$$

以及

$$ds_1^2 = a(V_{12}) ds_2^2$$

式中 V_{12} 是 K_2 相对于 K_1 速度的大小。比较这三组关系可得

$$\frac{a(V_2)}{a(V_1)} = a(V_{12}) \quad (1.6)$$

^① 二次式(1.4)所描述的四维几何，是 Minkowski 为相对论而引入的。这种几何称为**伪 Euclid 几何**。有些文献中将四维空间的四个坐标轴记作 x, y, z 和 ict ，这样间隔将在形式上表示为平方和的形式。

而 V_{12} 不仅依赖于矢量 \mathbf{V}_1 和 \mathbf{V}_2 的大小,还依赖于它们之间的夹角。^①但这个夹角在式(1.6)的左端并未出现,因此显然式(1.6)只有当函数 $a(V)$ 为常数时才成立,而且该常数必然为 1。因此,

$$ds^2 = ds'^2 \quad (1.7)$$

再从无限小间隔相等可得有限间隔相等,即 $s = s'$ 。

因此,我们得到一个很重要的结论:任意两个事件的间隔在所有惯性系中都是相等的,即当由一个惯性系变换到另一个惯性系时,这两个事件的间隔是不变的。这个不变性就是光速不变的数学表示。

1.2.2 类时间隔、类空间隔和类光间隔

再次假设 x_1, y_1, z_1, t_1 和 x_2, y_2, z_2, t_2 是在某一个惯性系 K 内的两个事件的坐标,是否存在一个惯性系 K' ,使得这两个事件在同一位置发生?在 K 系中,两个事件之间的间隔为

$$s^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 = c^2 t_{12}^2 - l_{12}^2$$

在 K' 系中,两个事件之间的间隔则为

$$s'^2 = c^2 t_{12}'^2 - l_{12}'^2$$

根据间隔的不变性,可得

$$c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 - l_{12}'^2$$

在 K' 系中两个事件在同一位置发生,即要求 $l_{12}'^2 = 0$,此时即有

$$s^2 = c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 > 0$$

由此可得,如果 $s^2 > 0$,即如果两个事件的间隔是实数的话,则存在一个惯性系,使得两个事件在同一位置发生。实数间隔称为**类时间隔**^②。

因此,若两个事件的间隔是类时的,那么就有这样一个参考系存在,在其中两个事件发生于同一位置。若任何两个事件在同一物体上发生,那么它们之间的间隔必然是类时的。

依然假设 x_1, y_1, z_1, t_1 和 x_2, y_2, z_2, t_2 是在某一个惯性系 K 内的两个事件的坐标,是否存在一个惯性系 K' ,使得这两个事件在同一时刻发生?于之前的讨论相同,同样有

$$c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 - l_{12}'^2$$

在 K' 系中两个事件在同一时刻发生,即要求 $t_{12}'^2 = 0$,此时即有

$$s^2 = c^2 t_{12}^2 - l_{12}^2 = -l_{12}'^2 < 0$$

因此,仅当两个事件的间隔是虚数的情况下,才存在一个惯性系使得两个事件同时发生。虚数间隔称为**类空间隔**。

因此,若两个事件的间隔是类空的,那么就有这样一个参考系存在,在其中两个事件同时发生。

恒为零的间隔称为**类光间隔**。

^① 虽然在相对论下,速度的合成公式于经典力学下不同,但两个速度的合成速度依然是依赖于两个速度的大小和它们的夹角。

^② 在一些其他文献上,间隔的符号与此处相反,因此间隔的类时性和类空性判定也是相反的。

由于间隔的不变性，将间隔分为类时间隔、类空间隔和类光间隔具有绝对的意义，即间隔的类空性、类时性或类旋光性与参考系无关。

取某一个事件 O 作为时间及空间坐目标原点，现在我们来研究所有其他事件对于本事件 O 的关系。首先只考虑一维空间的情况，如图1.2所示。一个当 $t = 0$ 时经过 $x = 0$ 点的粒子的匀速直线运动可以用一条直线来表示，这条直线过 O 点，对 t 轴的斜率等于粒子的速度。因为物体最大的速度为 c ，所以这条直线与 t 轴成角也有一个最大值。图1.2中有两条直线，代表两个信号经过事件 O （即当 $t = 0$ 时经过 $x = 0$ ）以光速向相反的两个方向传播。所有代表粒子运动的直线只能在 aOc 和 dOb 两个区域内。显然，在直线 ab 及 cd 上有 $x = \pm ct$ 。世界点位于区域 aOc 中的那些事件满足 $c^2t^2 - x^2 > 0$ 。即，这个区域内的任何事件与事件 O 之间的间隔都是类时的。在这个区域内 $t > 0$ ，即其中所有的事件都发生在事件 O 之“后”。但是两个事件若被类时间隔所分开，无论在哪一个参考系中都不可能同时发生。因此也不可能找到一个参考系使得 aOc 区域中的任何事件会在事件 O 之“前”发生，即在 $t < 0$ 时发生。因此， aOc 区域内的所有事件对 O 来说在任何参考系中都是未来的事件。所以，这个区域对事件 O 来说可称为**绝对未来**。

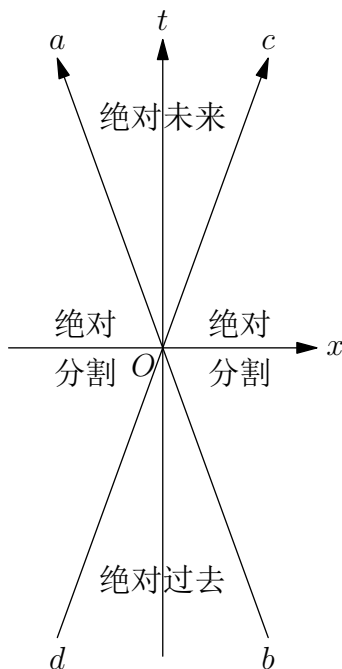


图 1.2: 光锥与因果性

完全类似地，所有在区域 bOd 内的事件对 O 来说都是**绝对过去**，即本区域内的事件，无论在何参考系中都在事件 O 之前发生。

最后，再研究 dOa 及 cOb 两个区域。本区域内的任何事件与事件 O 的间隔都是类空的。这些事件在任何参考系中都发生在空间的不同位置。因此，这些区域对于 O 来说都可称为**绝对分割**。但是，关于这些事件的“同时”、“较早”和“较晚”等概念都是相对的。对这些区域中的任何事件来说，在一些参考系中，此事件在事件 O 之后发生；在另一些参考系中，此事件在事件 O 之前发生；也存在一个参考系，在其中此事件与事件 O 同时发生。

如果考虑三个空间坐标轴，那么图1.2中两条相交的直线，在四维空间坐标系 x, y, z, t 中将称为一个“圆锥体” $x^2 + y^2 + z^2 - c^2t^2 = 0$ ，圆锥体的轴与 t 轴重合，称为**光锥**。光锥内的区域即为事件 O 的绝对未来和绝对过去。

两个事件仅仅在其间隔是类时间隔的情况下，彼此才能有因果关系；这可以由相互作用的传播速度不能大于光速这一事实直接推出来。也只有对这些事件来说，“较早”和“较晚”才具有绝对意义，而这一点又是使因果概念具有意义的必要条件。

1.3 固有时

假设我们在某一惯性参考系（不妨称之为静止系）中观察一只钟，这只钟相对于我们可作任意形式的运动。在各个不同的时刻，该钟的运动可以认为是匀速的。因此，在每一时刻，我们可以引入一个固连于运动钟上的坐标系，这个坐标系也是一个惯性系。

在无限小的时间间隔 dt 内，运动的钟前进的距离是 $\sqrt{dx^2 + dy^2 + dz^2}$ 。而在与运动钟固连的坐标系内，钟是静止的，即 $dx' = dy' = dz' = 0$ ，由间隔的不变性，可得

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt'^2$$

由此可得在与钟固连坐标系中的时间为

$$d\tau = dt \sqrt{1 - \frac{dx^2 + dy^2 + dz^2}{c^2 dt^2}} \quad (1.8)$$

再由

$$\frac{dx^2 + dy^2 + dz^2}{dt^2} = v^2$$

其中 v 为运动钟在静止系中的速度，所以有

$$d\tau = \frac{ds}{c} = dt \sqrt{1 - \frac{v^2}{c^2}} \quad (1.9)$$

将上式积分，我们可以得到，当静止的钟所行走的时间为 $t_2 - t_1$ 时，运动的钟所指示的时间为

$$\tau_2 - \tau_1 = \int_{t_1}^{t_2} dt \sqrt{1 - \frac{v^2}{c^2}} \quad (1.10)$$

随着某一给定物体一同运动的钟所指示的时间，称为该物体的**固有时**。由式(1.9)和(1.10)可知，一个运动物体的固有时永远比在静止系内相对应的时间间隔小。换句话说，运动的钟比静止的钟走得慢些。

假设有一只钟相对于某一惯性系 K 作匀速直线运动。同这只钟固连着的参考系 K' 也是惯性系。从 K 系的观察者来看， K' 系的钟走得慢。反过来说，从 K' 系的观察者来看， K 系的钟走得慢。为了是我们相信这是不矛盾的，可以注意下面的事实。为了确定 K' 系内的钟比 K 系内的钟慢，我们必须按下述方法来做。假设在某一时刻， K' 内的那只钟经过 K 内的一只钟旁边，而在这一时刻，两只钟所指示的时间恰好一样。为了比较 K 及 K' 内钟的快慢，我们必须再次将 K' 内同一只动钟的读数与 K 内的钟的读数作比较。但是新的时刻， K' 内的钟将从 K 内的另一些钟旁边经过，现在我们就将运动的钟与那些钟比较。这时我们发现， K' 内的钟比藉以比较的 K 内的钟走得慢。由此可见，为了比较两个参考系内的钟的快慢，我们需要在一个参考系内有几只钟，而在另一个参考系内只有一只钟。因此这种过程对这两个参考系来说，并不是对称的。与另一参考系内不同的一些钟比较的那一只钟总是走得慢的钟。

假设我们有两只钟，其中之一描绘一闭合路径，又回到出发点（即静止的钟所在之点），显然，与静止的钟相比，运动的钟慢了。相反，若设想运动的钟静止（即以运动的钟为参考

系), 就不能做上面的推论了。因为那只钟既然描绘了一条闭合曲线, 它的运动就不是匀速直线运动, 因而与之相连的参考系就不再是惯性系。因为自然定律只有在惯性系内才是一样的, 与静止的钟相连的系统(惯性系)及与运动的钟相连的系统(非惯性系)具有不同的特性, 因而导致静止的钟应当变慢这个结论的论证就不对了。

一只钟所指示的时间间隔, 等于沿着钟的世界线而取的积分 $\frac{1}{c} \int ds = \int d\tau$ 。假如钟是静止的, 则它的世界线是一条与 t 轴平行的直线; 假如钟在闭合路径上作非匀速运动而且又回到出发点, 那么它的世界线是一条曲线, 这条曲线经过静止钟的直的世界线的起点和终点。另一方面, 可以看到, 静止钟所指示的时间间隔永远比运动的钟所指示的大。因此, 我们得到一个结论, 在两个世界点之间所取的积分 $\int ds$, 只有在连接这两点之间的直线上能取得其最大值。^①

1.4 Lorentz 变换

在经典力学中, 已知一个惯性系 K 中事件的坐标 x, y, z, t , 获得相对 K 系以速度 V 匀速运动的 K' 系中的坐标 x', y', z', t' 是很简单的, 称为 Galileo 变换。即

$$\begin{cases} \mathbf{r} = \mathbf{r}' + \mathbf{V}t \\ t = t' \end{cases} \quad (1.11)$$

很容易证明, Galileo 变换不能满足相对论的要求, 即这个变换不能满足事件之间的间隔不变。

两个事件之间的间隔, 可以认为是四维 Minkowski 空间中两个世界点之间的距离。因此, 满足相对论要求的变换(即在四维 Minkowski 空间中的平移和旋转), 必须保持四维空间 x, y, z, ct 内的所有距离不变。对于平移, 这是显然的。因此只需考虑四维空间中的旋转即可。

四维空间内的一切转动可以分解为六个分别在六个平面 xy, yz, zx, tx, ty, tz 内的转动, 其中前三个转动仅仅变换空间坐标, 它们对应通常的空间转动, 只需将 K 系和 K' 系的坐标轴取得相互平行即可避免。

现在研究在 tx 平面内的转动, 这时 y 和 z 坐标是不变的。具体地说, 这个变换必须使差值 $c^2t^2 - x^2$, 即点 (ct, x) 到原点的“距离”的平方保持不变。因此, 新旧坐标关系最一般的表达式为

$$\begin{cases} x = x' \cosh \psi + ct' \sinh \psi \\ ct = x' \sinh \psi + ct' \cosh \psi \end{cases} \quad (1.12)$$

式中 ψ 为转动角。式(1.12)与坐标轴转动变换通常公式的不同之处在于, 此处的旋转公式中为双曲函数而非三角函数。这就是 Euclid 几何与伪 Euclid 几何之间的差别。假设 K' 系以速度 V 沿 x 轴相对 K 系作匀速运动, 则 x, t 与 x', t' 之间的变换关系式中转动角仅与相对速度 V 有关。

研究 K' 系的原点在 K 系中的运动。这时 $x' = 0$, 而式(1.12)可写成

$$x = ct' \sinh \psi, \quad ct = ct' \cosh \psi$$

^① 当然, 需假设 a 及 b 两点和连接它们的曲线满足该曲线上所有的线元 ds 都是类时的。这个性质与四维几何的伪 Euclid 特性有关。在 Euclid 空间中, 这个积分沿直线当然取得最小值, 此处定义的 Minkowski 空间的间隔与 Euclid 空间的间隔(距离)相比, 多了一个负号, 因此积分在直的世界在线取得最大值而非最小值。

由此可得

$$\frac{x}{ct} = \tanh \psi$$

而 $\frac{x}{t}$ 显然是 K' 相对 K 的速度 V , 因此有

$$\tanh \psi = \frac{V}{c}$$

由此可得

$$\begin{cases} x = \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}} \\ y = y' \\ z = z' \\ t = \frac{t' + \frac{V}{c^2}x'}{\sqrt{1 - \frac{V^2}{c^2}}} \end{cases} \quad (1.13)$$

这组变换称为 **Lorentz 变换**, 是今后讨论的基础。其逆变换只需将 V 替换为 $-V$ 即可得到^①。

由式(1.13)可见, 取 $c \rightarrow \infty$ 的经典力学极限, Lorentz 变换事实上就过渡到 Galileo 变换了。当 $V > c$ 时, 式(1.13)中的 x, t 变成虚数, 这与运动速度不可能大于光速的事实符合。此外我们也不可以用以光速运行的参考系, 因为在这种情形下, 式(1.13)的分母为零。

当 V 比光速小很多时, 我们可以用下面的近似公式代替式(1.13)

$$\begin{cases} x = x' + Vt' \\ y = y' \\ z = z' \\ t = t' + \frac{V}{c^2}x' \end{cases} \quad (1.14)$$

假设在 K 系中有一根平行于 x 轴的静止杆, 假定它在 K 系内测定的长度为 $\Delta x = x_2 - x_1$, 在 t' 时刻杆两端的坐标 x'_1, x'_2 与 x_1, x_2 的关系可由 Lorentz 变换(1.13)获得

$$x_1 = \frac{x'_1 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad x_2 = \frac{x'_2 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}$$

记杆在 K' 中的长度为 $\Delta x' = x'_2 - x'_1$, 则有

$$\Delta x = \frac{\Delta x'}{\sqrt{1 - \frac{V^2}{c^2}}}$$

杆的**固有长度**是它在与其相对静止的惯性系中的长度。如果用 $l_0 = \Delta x$ 表示这个固有长度, 用 l 表示它在任何其他参考系 K' 中的长度, 即有

$$l = l_0 \sqrt{1 - \frac{V^2}{c^2}} \quad (1.15)$$

因此, 杆在与其相对静止的惯性系中最长, 在相对于杆以速度 V 运动的惯性系中, 它的长度就要减少一个因子 $\sqrt{1 - \frac{V^2}{c^2}}$ 。相对论的这个结果称为 **Lorentz 收缩**。因为物体的横向尺

^① 当然直接求解方程组(1.13)也能得到相同的结果。

度都不因运动而改变, 所以它的体积 \mathcal{V} 也按照相似的公式收缩, 即

$$\mathcal{V} = \mathcal{V}_0 \sqrt{1 - \frac{V^2}{c^2}} \quad (1.16)$$

Lorentz 变换与 Galileo 变换另外一个不同点在于 Galileo 变换具有可对易性, 即连续两次 Galileo 变换 (具有不同的速度 \mathbf{V}_1 和 \mathbf{V}_2) 的结果与施行变换的顺序无关。然而连续两次 Lorentz 变换的结果一般依赖于它们的顺序。可类比于刚体的有限转动, 作为四维空间中的有限转动, 同样具有不可对易性, 除非 \mathbf{V}_1 和 \mathbf{V}_2 相互平行。

1.5 速度的变换

假设 K' 系相对于 K 系以速度 V 沿 x 轴运动, 则根据 Lorentz 变换(1.13)可有

$$\begin{cases} dx = \frac{dx' + Vdt'}{\sqrt{1 - \frac{V^2}{c^2}}} \\ dy = dy' \\ dz = dz' \\ dt = \frac{dt' + \frac{V}{c^2}dx'}{\sqrt{1 - \frac{V^2}{c^2}}} \end{cases}$$

考虑到

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}, \quad \mathbf{v}' = \frac{d\mathbf{r}'}{dt'}$$

则有

$$\begin{cases} v_x = \frac{dx}{dt} = \frac{v'_x + V}{1 + \frac{V}{c^2}v'_x} \\ v_y = \frac{dy}{dt} = \frac{v'_y \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c^2}v'_x} \\ v_z = \frac{dz}{dt} = \frac{v'_z \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c^2}v'_x} \end{cases} \quad (1.17)$$

这些公式就决定了速度的变换, 它们是相对论中的速度合成法则。在 $c \rightarrow \infty$ 的极限情形下, 它们就变为经典力学中的公式:

$$v_x = v'_x + V, \quad v_y = v'_y, \quad v_z = v'_z$$

在粒子沿 x 轴运动的特殊情况下, $v_x = v, v_y = v_z = 0$ 。那么, $v'_x = v', v'_y = v'_z = 0$, 并且有

$$v = \frac{v' + V}{1 + \frac{v'V}{c^2}} \quad (1.18)$$

很容易证明, 如果两个速度各小于或等于光速, 其合成速度根据这组公式也不会大于光速。

假如速度 V 比光速 c 小很多 (v 可以是任意的), 我们将近似地 (精确到 $\frac{V}{c}$ 的项) 得到

$$\begin{cases} v_x = v'_x + V \left(1 - \frac{v'^2_x}{c^2}\right) \\ v_y = v'_y - v'_x v'_y \frac{V}{c^2} \\ v_z = v'_z - v'_x v'_z \frac{V}{c^2} \end{cases} \quad (1.19)$$

这 3 个公式可以合写为一个矢量公式

$$\mathbf{v} = \mathbf{v}' + \mathbf{V} - \frac{1}{c^2} (\mathbf{V} \cdot \mathbf{v}') \mathbf{v}' \quad (1.20)$$

我们可以指出, 在相对论的速度合成公式(1.17)中, 相加的两个速度 \mathbf{v}' 和 \mathbf{V} 是以不对称的方式引入的 (如果它们都不沿 x 轴方向的话)。这个事实同 Lorentz 变换的非对易性有关, 将在下节提到。

让我们这样来选择坐标轴, 使粒子的速度在给定时刻是在 xy 平面内。这时粒子在 K 系内的速度分量是 $v_x = v \cos \theta$, $v_y = v \sin \theta$, 而在 K' 系内则为 $v'_x = v' \cos \theta'$, $v'_y = v' \sin \theta'$, 其中 v, v' 为速度在 K 和 K' 系中的大小, θ, θ' 为速度与 x 轴和 x' 轴的夹角。利用速度合成公式(1.17), 可得

$$\tan \theta = \frac{v' \sqrt{1 - \frac{V^2}{c^2}} \sin \theta'}{v' \cos \theta' + V} \quad (1.21)$$

这个公式决定了速度的方向从一个参考系变换到另一个参考系时的改变。

让我们来详尽地研究这个公式的另一个重要特例, 即光由一个参考系变换到另一个参考系时的偏差, 即所谓的光行差现象。在这种情形下 $v = v' = c$, 因而式(1.21)化为

$$\tan \theta = \frac{\sqrt{1 - \frac{V^2}{c^2}}}{\frac{V}{c} + \cos \theta'} \sin \theta' \quad (1.22)$$

再由式(1.17)可得

$$\sin \theta = \frac{\sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c} \cos \theta'} \sin \theta', \quad \cos \theta = \frac{\frac{V}{c} + \cos \theta'}{1 + \frac{V}{c} \cos \theta'} \quad (1.23)$$

如果 $V \ll c$, 由式(1.23)可得精确到数量级为 $\frac{V}{c}$ 项的公式如下:

$$\sin \theta - \sin \theta' = -\frac{V}{c} \sin \theta' \cos \theta'$$

若引入光行差角 $\Delta \theta = \theta' - \theta$, 我们就可得到同级的近似公式

$$\Delta \theta = \frac{V}{c} \sin \theta' \quad (1.24)$$

这就是著名的光行差的基本公式。

1.6 四维矢量和四维张量

方便起见, 本节在直角坐标系下进行讨论, 大部分结论都仅在直角坐标系中成立, 在曲线坐标系中, 本节中的很多形式将失去协变性。

1.6.1 四维矢量

一个事件的坐标 (ct, x, y, z) 可以看成是四维空间中一个四维径向矢量的分量。我们将它的分量记为 x^i ，这里指标 i 取值为 $0, 1, 2, 3$ ，而且

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

该径向四维矢量“长度”的平方由下式给出

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

它在四维坐标系的任意转动下不变，特别是，它在 Lorentz 变换下不变。

为了表示方便起见，我们引入四维坐标的两种“类型”，用带上标和下标的符号 x^i 和 x_i 来标记它们，它们之间满足如下关系

$$x_0 = x^0, \quad x_1 = -x^1, \quad x_2 = -x^2, \quad x_3 = -x^3 \quad (1.25)$$

量 x^i 称为四维坐标的**协变分量**， x_i 称为四维坐标的**逆变分量**。而四维径向矢量的平方则记为

$$\sum_{i=0}^3 x^i x_i = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3$$

一般经常略去求和号，将这样的求和简单记作 $x^i x_i$ ，也就是约定遍历所有重复指标求和，而把求和号省去^①。每对指标中必须一个为上标，另一个为下标。这种遍历“傀”指标求和的约定非常方便，可大大简化公式的书写。我们将用拉丁字母 i, j, k, l, \dots 表示四维指标，其取值为 $0, 1, 2, 3$ 。

四维坐标系的任意坐标变换可以用

$$x^i = f^i(x'^0, x'^1, x'^2, x'^3) \quad (1.26)$$

来表示。在这个坐标变换下，协变坐标的微分为

$$dx^i = \frac{\partial x^i}{\partial x'^j} dx'^j \quad (1.27)$$

而一个标量函数 $\phi(x^0, x^1, x^2, x^3)$ 的梯度则满足关系

$$\frac{\partial \phi}{\partial x^i} = \frac{\partial x'^j}{\partial x^i} \frac{\partial \phi}{\partial x'^j} \quad (1.28)$$

更一般地，如果 4 个量 A^0, A^1, A^2, A^3 ，在四维坐标系的变换下像四维协变坐标的微分 dx^i 那样变换（即式(1.27)），即

$$A^i = \frac{\partial x^i}{\partial x'^j} A'^j \quad (1.29)$$

我们就将这 4 个量的集合称为**四维矢量** A^i ，而每个 A^i 则称为该四维矢量的**协变分量**。而如果 4 个量 A_0, A_1, A_2, A_3 ，在四维坐标系的变换下像标量函数的梯度 $\frac{\partial \phi}{\partial x^i}$ 那样变换（即式(1.28)），即

$$A_i = \frac{\partial x'^j}{\partial x^i} A'_j \quad (1.30)$$

则这 4 个量的集合称为四维矢量 A_i ，而每个 A_i 则称为该四维矢量的**逆变分量**。同一个四维矢量可以用协变分量来表示，也可以用逆变分量来表示。

^① 这个约定称为“Einstein 求和约定”。

特别地, 在 Lorentz 变换下, 有

$$\begin{cases} A^0 = \frac{A'^0 + \frac{V}{c} A'^1}{\sqrt{1 - \frac{V^2}{c^2}}} \\ A^1 = \frac{A'^1 + \frac{V}{c} A'^0}{\sqrt{1 - \frac{V^2}{c^2}}} \\ A^2 = A'^2 \\ A^3 = A'^3 \end{cases}, \quad \text{以及} \quad \begin{cases} A_0 = \frac{A'_0 - \frac{V}{c} A'_1}{\sqrt{1 - \frac{V^2}{c^2}}} \\ A_1 = \frac{A'_1 - \frac{V}{c} A'_0}{\sqrt{1 - \frac{V^2}{c^2}}} \\ A_2 = A'_2 \\ A_3 = A'_3 \end{cases} \quad (1.31)$$

与四维径向矢量的平方类似, 任一四维矢量的平方定义为:

$$A^i A_i = A^0 A_0 + A^1 A_1 + A^2 A_2 + A^3 A_3 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2$$

与四维矢量的平方类比, 我们可以构造两个不同四维矢量的标积:

$$A^i B_i = A_i B^i = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$$

显然, 这既可以写为 $A^i B_i$, 也可以写为 $A_i B^i$, 结果是相同的。一般来说, 一对傀指标中的上标和下标总是可以交换的。

积 $A^i B_i$ 是一个**四维标量**——它在四维坐标系的变换下是不变的。这点根据四维矢量协变分量的变换关系(1.29)和逆变分量的变换关系(1.30)可得

$$A^i B_i = \frac{\partial x^i}{\partial x'^j} A'^j \frac{\partial x'^k}{\partial x^i} B'_k = \frac{\partial x'^k}{\partial x^i} \frac{\partial x^i}{\partial x'^j} A'^j B'_k = \frac{\partial x'^k}{\partial x'^j} A'^j B'_k$$

其中显然有 $\frac{\partial x'^k}{\partial x'^j} = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases}$, 因此有 $A^i B_i = A'^j B'_j$, 由此便说明了四维矢量的平方与四维坐标系的变换无关。

分量 A^0 称为四维矢量的**时间分量**, A^1, A^2, A^3 称为四维矢量的**空间分量** (与四维径向矢量类比)。四维矢量的平方可以为正、负或零, 相应的四维矢量分别称为**类时矢量**、**类空矢量**和**类光矢量**^①。

在纯空间转动 (即不影响时间轴的变换) 下, 四维矢量 A^i 的三个空间分量构成一个三维矢量 \mathbf{A} 。该四维矢量的时间分量 (在这些变换下) 是一个三维标量。为了列举四维矢量的分量, 我们常将其写为

$$A^i = (A^0, \mathbf{A})$$

同一四维矢量的协变分量为 $A_i = (A^0, -\mathbf{A})$ 。该四维矢量的平方为

$$A^i A_i = (A^0)^2 - \mathbf{A}^2$$

因此, 对于四维径向矢量:

$$x^i = (ct, \mathbf{r}), \quad x_i = (ct, -\mathbf{r}), \quad x^i x_i = c^2 t^2 - \mathbf{r}^2$$

对于三维矢量, 在直角坐标系中没有必要区分逆变和协变分量。只要能够做到这一点而不至引起混淆, 我们将用希腊字母作为下标把这些分量记为 A_α ($\alpha = x, y, z$)。对于重复的希腊字母指针我们将假设其遍历 x, y, z 求和, 例如

$$\mathbf{A} \cdot \mathbf{B} = A_\alpha B_\alpha$$

^① 类光矢量也称为各向同性矢量。

1.6.2 四维张量

二阶四维张量是 16 个量 A^{ij} 的集合，它在坐标变换下像两个四维矢量分量的积那样变换，即

$$A^{ij} = \frac{\partial x^i}{\partial x'^k} \frac{\partial x^j}{\partial x'^l} A'^{kl} \quad (1.32)$$

类似地，可以定义更高阶的四维张量。一个二阶张量的分量可以写为三种形式：协变分量 A_{ik} 、逆变分量 A^{ik} 和混合分量 A^i_k 及 A_k^i 。不同类型分量之间的联系由一下通则决定：升或降一个空间指标（1, 2, 3）改变分量的正负号，而升或降时间指标（0）则不变号。因此有

$$\begin{aligned} A_{00} &= A^{00}, & A_{01} &= -A^{01}, & A_{11} &= A^{11}, \dots \\ A^0_0 &= A^{00}, & A^0_1 &= A^{01}, & A^0_{-1} &= -A^{01}, & A^1_1 &= -A^{11}, \end{aligned}$$

在纯空间变换下，9 个量 A^{11}, A^{12}, \dots 构成一个三维张量。三个分量 A^{01}, A^{02}, A^{03} 和三个分量 A^{10}, A^{20}, A^{30} 构成三维矢量，而分量 A^{00} 是一个三维标量。

如果 $A^{ik} = A^{ki}$ ，张量 A^{ik} 称为**对称的**；如果 $A^{ik} = -A^{ki}$ ，张量 A^{ik} 称为**反对称的**。在反对称张量中，所有对角分量（即分量 $A^{00}, A^{11}, A^{22}, A^{33}$ ）都是零。对于一个对称张量 A^{ik} ，混合分量 A^i_k 和 A_k^i 相等，在这样的情形下，我们把一个指标置于另一个上方，简单地记为 A^i_k 。

在每个张量方程中，等号两边所含的自由指标（区别于傀指标）必须字母相同且位置相同（即上或下）。张量方程中的自由指标可以上移或下移，但必须对方程中所有的项同时进行。让不同张量的协变和逆变分量相等是“非法的”，这样的方程即便碰巧在特定参考系中成立，在变换到另一个参考系时也会失效（即不具有“协变性”）。

通过对张量 A^{ik} 的分量求和可以形成一个标量

$$A^i_i = A^0_0 + A^1_1 + A^2_2 + A^3_3$$

这个和称为**张量的迹**，求得它的运算称为**缩并**。缩并任何一对指标会使张量的阶减去 2。例如二阶张量 $A^i B_k$ 缩并后 $A^i B_i$ 为四维标量，而四阶张量 A^{iklm} 的缩并 A^{ikli} 是一个二阶张量，等等。

单位四维张量 δ^i_k 满足如下条件：对于任意四维矢量 A^i ，有

$$\delta^i_k A^k = A^i \quad (1.33)$$

因此，这个张量的分量显然是

$$\delta^i_k = \begin{cases} 1, & \text{当 } i = k \\ 0, & \text{当 } i \neq k \end{cases} \quad (1.34)$$

它的迹是 $\delta^i_i = 4$ 。通过在 δ^i_k 中升一个指标或者降一个指标，我们就得到逆变张量 η^{ik} 或协变张量 η_{ik} ，称之为**Minkowski 空间度规张量**。张量 η^{ik} 和 η_{ik} 具有相同的分量，可以用矩阵表示为

$$(\eta^{ik}) = (\eta_{ik}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.35)$$

由此，显然有

$$\eta_{ik} A^k = A_i, \quad \eta^{ik} A_k = A^i \quad (1.36)$$

两个四维矢量的标积因而可以写成形式

$$A^i A_i = \eta_{ik} A^i A^k = \eta^{ik} A_i A_k \quad (1.37)$$

张量 δ_k^i 、 η_{ik} 和 η^{ik} 的特别之处在于，它们的分量在所有坐标系中都相同。四阶的全反对称单位张量 e^{iklm} 具有同样的性质。这个张量的分量在交换任意一对指标时变号，其非零分量为 ± 1 。从反对称性可知，有两个指标相同的所有分量均为零，所以仅有的非零分量是那些所有 4 个指标都不同者。我们令

$$e^{0123} = 1 \quad (1.38)$$

于是，所有其他的非零分量 e^{iklm} 等于 1 或 -1 ，依 i, k, l, m 这几个数能经偶数还是奇数次换位排成 0, 1, 2, 3 而定。由于

$$e_{0123} = \eta_{0i} \eta_{1j} \eta_{2k} \eta_{3l} e^{ijkl} = (-1)^3 e^{0123} = -1$$

以及 e^{iklm} 的非零分量数为 $4! = 24$ ，可得

$$e^{iklm} e_{iklm} = 24(-1)^3 (e^{iklm})^2 = -24 \quad (1.39)$$

对于坐标系的转动而言， e^{iklm} 诸量的特性与张量分量的特性相同，但是如果我们改变 1 个或 3 个坐标的正负号，分量 e^{iklm} 并不改变，因为根据定义，它们在所有坐标系中都相同，而张量的分量在这种情况下是应当变号的。所以，严格地说， e^{iklm} 并不是张量，而是一个赝张量。任意阶的赝张量，特别是赝标量，在所有的坐标变换下都具有张量的性质，只有那些不能归结为转动的变换，即反射（不能归结为转动的坐标正负号改变）是例外。

乘积 $e^{iklm} e_{prst}$ 构成一个 8 阶四维张量，它是一个真正的张量，通过缩并其一对或多对指标可以得到 6 阶、4 阶和 2 阶张量。所有这些张量在所有坐标系中具有相同的形式。所以，它们的分量必须表示为单位张量 δ_k^i （其分量在所有坐标系中都相同的唯一真张量）分量乘积的组合。

记 σ 为一个 4 阶置换，其作用在 0, 1, 2, 3 中的某数将得到其重排之后所对应的数。由于 e^{iklm} 的值只有在其指标互不相同才非零，因此有

$$i, k, l, m = \sigma(0), \sigma(1), \sigma(2), \sigma(3)$$

用 $\text{sgn } \sigma$ 表示置换 σ 的符号，如果置换后的元素经过偶数次交换可以恢复原本顺序，则该置换符号值为 1，否则为 -1 。根据全反对称单位张量的定义，即有 $e^{\sigma(0)\sigma(1)\sigma(2)\sigma(3)} = \text{sgn } \sigma$ ，由此即有

$$e^{iklm} e_{prst} = e^{\sigma(0)\sigma(1)\sigma(2)\sigma(3)} e_{\gamma(0)\gamma(1)\gamma(2)\gamma(3)} = -\text{sgn } \sigma \text{sgn } \gamma$$

考虑到指标排列所必须具有的对称性，可有

$$e^{iklm} e_{prst} = e^{\sigma(0)\sigma(1)\sigma(2)\sigma(3)} e_{\gamma(0)\gamma(1)\gamma(2)\gamma(3)} = -\text{sgn } \sigma \text{sgn } \gamma = -\text{sgn } \sigma \text{sgn } \gamma \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$= -\text{sgn } \sigma \text{sgn } \gamma \begin{vmatrix} \delta_0^0 & \delta_1^0 & \delta_2^0 & \delta_3^0 \\ \delta_0^1 & \delta_1^1 & \delta_2^1 & \delta_3^1 \\ \delta_0^2 & \delta_1^2 & \delta_2^2 & \delta_3^2 \\ \delta_0^3 & \delta_1^3 & \delta_2^3 & \delta_3^3 \end{vmatrix} = -\begin{vmatrix} \delta_{\gamma(0)}^{\sigma(0)} & \delta_{\gamma(1)}^{\sigma(0)} & \delta_{\gamma(2)}^{\sigma(0)} & \delta_{\gamma(3)}^{\sigma(0)} \\ \delta_{\gamma(0)}^{\sigma(1)} & \delta_{\gamma(1)}^{\sigma(1)} & \delta_{\gamma(2)}^{\sigma(1)} & \delta_{\gamma(3)}^{\sigma(1)} \\ \delta_{\gamma(0)}^{\sigma(2)} & \delta_{\gamma(1)}^{\sigma(2)} & \delta_{\gamma(2)}^{\sigma(2)} & \delta_{\gamma(3)}^{\sigma(2)} \\ \delta_{\gamma(0)}^{\sigma(3)} & \delta_{\gamma(1)}^{\sigma(3)} & \delta_{\gamma(2)}^{\sigma(3)} & \delta_{\gamma(3)}^{\sigma(3)} \end{vmatrix}$$

$$= - \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i & \delta_t^i \\ \delta_p^k & \delta_r^k & \delta_s^k & \delta_t^k \\ \delta_p^l & \delta_r^l & \delta_s^l & \delta_t^l \\ \delta_p^m & \delta_r^m & \delta_s^m & \delta_t^m \end{vmatrix}$$

由此，可有

$$e^{iklm}e_{prst} = \delta_t^i \begin{vmatrix} \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \\ \delta_p^m & \delta_r^m & \delta_s^m \end{vmatrix} - \delta_t^k \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^l & \delta_r^l & \delta_s^l \\ \delta_p^m & \delta_r^m & \delta_s^m \end{vmatrix} + \delta_t^l \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^m & \delta_r^m & \delta_s^m \end{vmatrix} - \delta_t^m \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix}$$

缩并其中的 m 和 t 可得

$$\begin{aligned} e^{iklm}e_{prsm} &= \begin{vmatrix} \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \\ \delta_p^i & \delta_r^i & \delta_s^i \end{vmatrix} - \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix} + \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix} - 4 \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix} \\ &= - \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix} \end{aligned}$$

由此，可有

$$e^{iklm}e_{prsm} = -\delta_s^i \begin{vmatrix} \delta_p^k & \delta_r^k \\ \delta_p^l & \delta_r^l \end{vmatrix} + \delta_s^k \begin{vmatrix} \delta_p^i & \delta_r^i \\ \delta_p^l & \delta_r^l \end{vmatrix} - \delta_s^l \begin{vmatrix} \delta_p^i & \delta_r^i \\ \delta_p^k & \delta_r^k \end{vmatrix}$$

缩并其中的 l 和 s 可得

$$e^{iklm}e_{prlm} = - \begin{vmatrix} \delta_p^k & \delta_r^k \\ \delta_p^i & \delta_r^i \end{vmatrix} + \begin{vmatrix} \delta_p^i & \delta_r^i \\ \delta_p^k & \delta_r^k \end{vmatrix} - 4 \begin{vmatrix} \delta_p^i & \delta_r^i \\ \delta_p^k & \delta_r^k \end{vmatrix} = -2 \begin{vmatrix} \delta_p^i & \delta_r^i \\ \delta_p^k & \delta_r^k \end{vmatrix} = -2(\delta_p^i \delta_r^k - \delta_r^i \delta_p^k)$$

再缩并其中的 k 和 r 可得

$$e^{iklm}e_{pklm} = -2(4\delta_p^i - \delta_p^i) = -6\delta_p^i$$

综上，可有如下关系式

$$e^{iklm}e_{prst} = - \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i & \delta_t^i \\ \delta_p^k & \delta_r^k & \delta_s^k & \delta_t^k \\ \delta_p^l & \delta_r^l & \delta_s^l & \delta_t^l \\ \delta_p^m & \delta_r^m & \delta_s^m & \delta_t^m \end{vmatrix} \quad (1.40)$$

$$e^{iklm}e_{prsm} = - \begin{vmatrix} \delta_p^i & \delta_r^i & \delta_s^i \\ \delta_p^k & \delta_r^k & \delta_s^k \\ \delta_p^l & \delta_r^l & \delta_s^l \end{vmatrix} \quad (1.41)$$

$$e^{iklm}e_{prlm} = -2(\delta_p^i \delta_r^k - \delta_r^i \delta_p^k) \quad (1.42)$$

$$e^{iklm}e_{pklm} = -6\delta_p^i \quad (1.43)$$

如果 A^{ik} 是一个反对称张量，则张量 A^{ik} 和赝张量 $A^{*ik} = \frac{1}{2}e^{iklm}A_{lm}$ 称为彼此对偶。类似地， $e^{iklm}A_m$ 是一个与矢量 A^i 对偶的三阶反对称赝张量。对偶张量的乘积 $A^{ik}A_{ik}^*$ 显然是一个赝标量。

联系这里的讨论, 三阶全反对称单位赭张量 $e_{\alpha\beta\gamma}$ 是这样一些量的集合, 它们在任何一对指标换位时变号。因此在 $e_{\alpha\beta\gamma}$ 的分量中, 只有那些具有 3 个不同指标者才非零。我们令 e_{123} , 于是其他分量等于 1 或 -1 则根据 α, β, γ 这个序列能经偶数还是奇数次换位排成 x, y, z 的顺序而定 (即等于该置换的符号)。乘积 $e_{\alpha\beta\gamma}e_{\lambda\mu\nu}$ 构成一个 6 阶真三维张量, 因而可以表示为单位三维张量 $\delta_{\alpha\beta}$ 分量乘积的组合。对于三阶全反对称单位赭张量也有类似的如下关系式:

$$e_{\alpha\beta\gamma}e_{\lambda\mu\nu} = \begin{vmatrix} \delta_{\alpha\lambda} & \delta_{\alpha\mu} & \delta_{\alpha\nu} \\ \delta_{\beta\lambda} & \delta_{\beta\mu} & \delta_{\beta\nu} \\ \delta_{\gamma\lambda} & \delta_{\gamma\mu} & \delta_{\gamma\nu} \end{vmatrix} \quad (1.44)$$

$$e_{\alpha\beta\gamma}e_{\lambda\mu\gamma} = \delta_{\alpha\lambda}\delta_{\beta\mu} - \delta_{\alpha\mu}\delta_{\beta\lambda} \quad (1.45)$$

$$e_{\alpha\beta\gamma}e_{\lambda\beta\gamma} = 2\delta_{\alpha\lambda} \quad (1.46)$$

$$e_{\alpha\beta\gamma}e_{\alpha\beta\gamma} = 6 \quad (1.47)$$

在坐标系的反射 (即所有坐标变号) 下, 一个普通矢量也变号。这样的矢量称为**极矢量**。一个矢量若能写成两个极矢量的矢积, 则其分量在反演下不变号, 这样的矢量称为**轴矢量**。一个极矢量和一个轴矢量的标积并不是一个真标量, 而是一个赭标量, 它在坐标反演下变号。轴矢量是赭矢量, 对偶于某反对称张量。因此, 如果 $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, 那么

$$C_\alpha = \frac{1}{2}e_{\alpha\beta\gamma}C_{\beta\gamma}, \quad \text{其中 } C_{\beta\gamma} = A_\beta B_\gamma - A_\gamma B_\beta$$

现在回到四维张量上来。反对称张量 A^{ik} 的空间分量 ($i, k = 1, 2, 3$) 对于纯空间变换构成一个三维反对称张量。根据上面的论述, 其分量可以用一个三维轴矢量的分量来表示。对于同样的变换, 分量 A^{01}, A^{02}, A^{03} 构成一个三维极矢量。因此一个反对称四维张量可以写成矩阵

$$(A^{ik}) = \begin{pmatrix} 0 & p_x & p_y & p_z \\ -p_x & 0 & -a_z & a_y \\ -p_y & a_z & 0 & -a_x \\ -p_z & -a_y & a_x & 0 \end{pmatrix} \quad (1.48)$$

这里对于空间变换, \mathbf{p} 和 \mathbf{a} 分别为极矢量和轴矢量。在列出反对称四维张量的分量时, 可以将它们写成形式

$$A^{ik} = (\mathbf{p}, \mathbf{a})$$

其协变分量则为

$$A_{ik} = (-\mathbf{p}, \mathbf{a})$$

1.6.3 四维张量分析

前面已经提到, 标量 ϕ 的四维梯度是四维矢量 $\frac{\partial\phi}{\partial x^i}$, 它可以展开写为

$$\frac{\partial\phi}{\partial x^i} = \left(\frac{1}{c} \frac{\partial\phi}{\partial t}, \nabla\phi \right) \quad (1.49)$$

而坐标的微分 dx^i 与四维梯度的缩并即为该标量的微分

$$d\phi = \frac{\partial\phi}{\partial x^i} dx^i$$

由此可以看出，它具有两个四维矢量标积的形式，也是一个标量。

一般说来，对于坐标 x^i 微分的算符 $\frac{\partial}{\partial x^i}$ 应当看成是该算符四维矢量的协变分量。如果我们对于“协变坐标” x_i 进行微分，则导数

$$\frac{\partial \phi}{\partial x_i} = \eta^{ik} \frac{\partial \phi}{\partial x^k} = \left(\frac{1}{c} \frac{\partial \phi}{\partial t}, -\nabla \phi \right)$$

构成一个四维矢量的逆变分量。对于坐目标偏导数可以采用缩写

$$\partial_i = \frac{\partial}{\partial x^i}, \quad \partial^i = \frac{\partial}{\partial x_i}$$

以这种形式书写微分算符时，由它们构成的量协变或逆变性质是一目了然的。

在三维空间中，可以沿体积、曲面或曲线进行积分。在四维空间中积分有四种类型：

1. 沿四维空间中一条曲线的积分。积分元就是线元，即四维矢量 dx^i 。

2. 沿四维空间中的一个（二维）曲面的积分。在三维空间中，无穷小有向面积元素在坐标平面 $x_\alpha x_\beta$ 上的投影为微分 2- 形式 $dx_\alpha \wedge dx_\beta$ 。类似地，在四维空间中，无限小面元由二阶反对称张量 $df^{ij} = dx^i \wedge dx^j$ 给定，其分量为该面元在坐标平面上的投影。在三维空间中，我们不用张量 $df_{\alpha\beta}$ 而用与之对偶的矢量 df_γ 来表示该面元 $df_\gamma = \frac{1}{2} e_{\alpha\beta\gamma} df_{\alpha\beta}$ 。在几何上，这是一个与面元垂直的矢量，其绝对值等于面元的面积。在四维空间中无法构造这样的矢量，但可以构造与张量 df^{ij} 对偶的张量 df^{*ij} ，满足

$$df^{*ij} = \frac{1}{2} e^{ijkl} df_{kl} \quad (1.50)$$

它们之间满足关系 $df^{*ij} df_{ij} = 0$ ，这是由于

$$\begin{aligned} df^{*ij} df_{ij} &= \frac{1}{2} e^{ijkl} df_{kl} df_{ij} = -\frac{1}{2} e^{ljk i} df_{ij} df_{kl} = -\frac{1}{2} e^{jkl i} df_{ij} df_{kl} = \frac{1}{2} e^{jikl} df_{ij} df_{kl} \\ &= -\frac{1}{2} e^{ijkl} df_{ij} df_{kl} = -df^{*ij} df_{ij} \end{aligned}$$

由此即有 $df^{*ij} df_{ij} = 0$ 。在几何上，它描述了一个正交于面元 df_{ij} 而且与之大小相等的面元。

3. 沿一个超曲面，即沿一个三维流形的积分。在三维空间中，无穷小有向体积元素可以表示为微分 3- 形式 $dx \wedge dy \wedge dz$ 。类似地，在四维空间中，一个无穷小三维超曲面的有向面积元素可以表示为 $dS^{ijk} = dx^i \wedge dx^j \wedge dx^k$ 。这构成一个三阶张量，对所有 3 个指标都是反对称的，在此情形下，使用与张量 dS^{ijk} 对偶的四维矢量 dS^{*i} 更方便：

$$dS^{*i} = -\frac{1}{6} e^{ijkl} dS_{jkl}, \quad dS_{jkl} = e_{njkl} dS^{*n} \quad (1.51)$$

这里

$$dS^{*0} = dS^{123}, \quad dS^{*1} = dS^{023}, \quad dS^{*2} = dS^{013}, \quad dS^{*3} = dS^{012}$$

在几何上， dS^{*i} 是一个四维矢量，数值上等于超曲面元的“面积”，并与该面元垂直。特别是， $dS^{*0} = dx \wedge dy \wedge dz$ ，这就是三维体积元 dV ，即超曲面在超平面 $x^0 = \text{const}$ 上的投影。

4. 沿一个四维体积的积分，积分元是标量

$$d\Omega = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 = c dt \wedge dV \quad (1.52)$$

类似三维矢量分析中的 Gauss 定理和 Stokes 定理，利用微分几何中的广义 Stokes 公式

$$\oint_{\partial\Omega} \omega = \int_{\Omega} d\omega \quad (1.53)$$

可以得到四维积分的变换，其中的算符 d 表示外微分，设有外形式为

$$\omega = f dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}$$

其运算规则为

$$d\omega = \partial_j f dx^j \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}$$

沿一闭合超曲面的积分可以变换到沿其包含的四维体积的积分，即

$$\begin{aligned} \oint A^i dS_i^* &= \oint A^i \left(-\frac{1}{6} e_{ijkl} dS^{jkl} \right) = -\frac{1}{6} \oint A^i e_{ijkl} dx^j \wedge dx^k \wedge dx^l \\ &= -\frac{1}{6} \int d(A^i e_{ijkl} dx^j \wedge dx^k \wedge dx^l) = -\frac{1}{6} \int \partial_n A^i e_{ijkl} dx^n \wedge dx^j \wedge dx^k \wedge dx^l \end{aligned}$$

根据微分形式的反对称性和四阶全反对称张量 e^{ijkl} 的定义可有

$$dx^n \wedge dx^j \wedge dx^k \wedge dx^l = e^{njkl} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 = e^{njkl} d\Omega$$

所以有

$$\oint A^i dS_i^* = -\frac{1}{6} \int \partial_n A^i e_{ijkl} e^{njkl} d\Omega = -\frac{1}{6} \int \partial_n A^i (-6\delta_i^n) d\Omega = \int \partial_i A^i d\Omega$$

更一般地，任意阶张量沿一闭合超曲面的积分都可以变换到沿其包含的四维体积的积分，只需在积分式中用算符

$$dS_i^* \rightarrow d\Omega \frac{\partial}{\partial x^i} = d\Omega \partial_i \quad (1.54)$$

代替积分元 dS_i^* 即可。这个公式是三维空间中 Gauss 定理的推广。

沿一个二维曲面的积分可以变换为“包含”它的超曲面的积分，用算符

$$df_{ik}^* \rightarrow dS_i \frac{\partial}{\partial x^k} - dS_k \frac{\partial}{\partial x^i} = dS_i \partial_k - dS_k \partial_i \quad (1.55)$$

代替积分元 df_{ik}^* 。例如，对于反对称张量 A^{ik} 的积分，可有

$$\frac{1}{2} \oint A^{ik} df_{ik}^* = \frac{1}{2} \int (dS_i \partial_k A^{ik} - dS_k \partial_i A^{ik}) = \int dS_i \partial_k A^{ik} \quad (1.56)$$

沿一条四维闭合曲线的积分可以通过代换

$$dx^i \rightarrow df^{ki} \frac{\partial}{\partial x^k} = df^{ki} \partial_k \quad (1.57)$$

变换为“包含”它的曲面的积分。因此，对于一个矢量的积分，可有

$$\oint A_i dx^i = \int df^{ki} \partial_k A_i = \frac{1}{2} \int df^{ki} (\partial_i A_k - \partial_k A_i) \quad (1.58)$$

这是 Stokes 定理的推广。

1.7 四维速度

由普通的三维速度矢量，我们可以构造一个四维矢量。一个粒子的**四维速度**（四速度）是矢量

$$u^i = \frac{dx^i}{d\tau} \quad (1.59)$$

为了求出它的分量，根据式(1.9)可有

$$d\tau = dt \sqrt{1 - \frac{v^2}{c^2}} \quad (1.60)$$

其中 v 为粒子的普通三维速度。因此

$$u^1 = \frac{dx^1}{d\tau} = \frac{dx}{dt \sqrt{1 - \frac{v^2}{c^2}}} = \frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}}$$

由此可得

$$u^i = \left(\frac{c}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \quad (1.61)$$

注意到，四维速度是一个具有速度的量纲。

四维速度的分量之间并不彼此独立，由于 $dx^i dx_i = ds^2 = c^2 d\tau^2$ ，因此有

$$u^i u_i = c^2 \quad (1.62)$$

因此，四维速度的几何意义是粒子世界线的一个四维切矢量，其大小为 c 。

与四维速度的定义类似，二阶导数

$$w^i = \frac{d^2 x^i}{d\tau^2} = \frac{du^i}{d\tau}$$

可以称为**四维加速度**。微分式(1.62)可得

$$u_i w^i = 0 \quad (1.63)$$

即四维速度矢量与四维加速度矢量是相互正交的。

第二章 相对论力学

2.1 Hamilton 原理

现在我们从 Hamilton 原理开始, 研究实物粒子的运动规律。Hamilton 原理指出: 对于每一个力学体系, 有一个叫做作用量的积分 S 存在, 这个积分对于实际运动有最小值^①, 即它的变分 δS 为零。

一个自由实物粒子的作用量积分必然与参考系的选择无关, 即, 它必须对于 Lorentz 变换保持不变。因此, 它必须为一个标量函数。另外显然其被积函数必须是一个微分 1- 形式。对于一个自由粒子, 所能构造出的唯一的这种标量, 仅仅是间隔 ds , 或者固有时 $d\tau$, 或者它们乘以一个常数 αds 。这样一来, 对于一个自由粒子, 作用量的积分必须取下面的形式:

$$S = -\alpha \int_{\tau_1}^{\tau_2} ds$$

其中 \int_a^b 表示沿着粒子在两个特定事件间的世界线的积分, 这两个事件就是粒子在 t_1 时刻到达初位置和在 t_2 时刻到达末位置, 也就是说 \int_a^b 是沿着两个世界点之间的世界线的积分; 而 α 为表征该粒子的一个常数。在第 1.3 节中, 积分 $\int_a^b ds$ 沿着一条直的世界线的值最大; 沿着一条弯曲的世界线, 可以使得积分值为任意小^②

作用量可以表示为对时间的积分 $S = \int_{t_1}^{t_2} L dt$, 其中 L 即为这个力学体系的 Lagrange 函数。利用式(1.9), 可得

$$S = - \int_{t_1}^{t_2} \alpha c \sqrt{1 - \frac{v^2}{c^2}} dt$$

其中 v 为实物粒子的速度, 即实物粒子的 Lagrange 函数为

$$L = -\alpha c \sqrt{1 - \frac{v^2}{c^2}} \quad (2.1)$$

上面已经提到, α 是表征该粒子的一个量。在经典力学中, 这个量就是该粒子的质量 m 。当我们取 $c \rightarrow +\infty$ 的极限时, L 的表达式应该过渡到它的经典表达式 $L = \frac{1}{2}mv^2$ 。将 L 按 $\frac{v}{c}$ 展开至 $\frac{v^2}{c^2}$ 项可得

$$L = -\alpha c \sqrt{1 - \frac{v^2}{c^2}} = -\alpha c + \frac{\alpha v^2}{2c}$$

Lagrange 函数中的常数项对运动方程没有影响, 可以略去。略去常数项 $-\alpha c$ 之后, 与经典力学中自由粒子的 Lagrange 函数 $L = \frac{1}{2}mv^2$ 比较, 可得 $\alpha = mc$ 。

^① 严格的讲, 应该是驻值。

^② 因为这些积分的值都是负值。

所以, 自由实物粒子的作用量是

$$S = -mc \int_a^b ds \quad (2.2)$$

而 Lagrange 函数是

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} \quad (2.3)$$

2.2 能量与动量

类似于经典 Hamilton 力学中的做法, 将矢量 $\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}}$ ^① 称为该粒子的动量。利用式(2.3)可得

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.4)$$

对于很小的速度 ($v \ll c$), 或者 $c \rightarrow +\infty$ 的极限情形下, 式(2.4)就变为经典的动量公式 $\mathbf{p} = m\mathbf{v}$ 。而当 $v = c$ 时, 实物粒子的动量就变为无穷大。

动量对时间的导数就是作用于粒子的力。将速度矢量记作 $\mathbf{v} = v\mathbf{e}$, 其中 \mathbf{e} 是与 \mathbf{v} 通向的方向矢量, 则有

$$\frac{d\mathbf{p}}{dt} = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{d\mathbf{e}}{dt} + \frac{m\mathbf{e}}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{3}{2}}} \frac{dv}{dt} \quad (2.5)$$

当粒子的速度大小不变仅有方向改变时, 即力与速度方向垂直时, 有

$$\frac{d\mathbf{p}}{dt} = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{d\mathbf{e}}{dt}$$

当粒子的速度方向不变仅有大小改变时, 即力与速度方向平行时, 有

$$\frac{d\mathbf{p}}{dt} = \frac{m\mathbf{e}}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{3}{2}}} \frac{dv}{dt}$$

即力与加速度之比已经不再是常数, Newton 第二定律失效。

粒子的能量 \mathcal{E} ^②可以定义为

$$\mathcal{E} = \mathbf{p} \cdot \mathbf{v} - L \quad (2.6)$$

将式(2.4)和式(2.3)代入可得

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.7)$$

相对论力学与经典力学的一个重要差别即为, 粒子的能量在 $v = 0$ 时并不为零, 而是有一个有限值

$$\mathcal{E}_0 = mc^2 \quad (2.8)$$

① $\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}}$ 表示 $p^i = \frac{\partial L}{\partial v^i}$ 。

② 此处用 \mathcal{E} 而不用 E 来表示能量是为了避免与后面涉及的电场强度 \mathbf{E} 相混淆。

这个量称为粒子的**静能**。当速度远小于光速时，将式(2.7)展开为 $\frac{v^2}{c^2}$ 的幂级数，可以得到

$$\mathcal{E} \approx mc^2 + \frac{1}{2}mv^2$$

即，扣除静能后，能量表达式(2.7)退化为经典结果。

虽然我们考虑的是一个自由粒子的情形，但是这些公式同样可以适用于许多粒子构成的复合物体。此时 m 即为该物体的总质量， \mathbf{v} 是指它整体的运动速度。在相对论力学中，任何封闭系统的能量都是一个完全确定的量，它总是正的并与该系统的质量直接相关。而在经典力学中，封闭系统的能量可以有一个任意常数的差别，而且既可以为正值，也可以为负值。

一个静止系统的能量，除了其组成粒子的静能外，还包括它们的动能和它们之间的相互作用能。因此， mc^2 并不等于 $\sum m_\alpha c^2$ （其中 m_α 为各个粒子单独存在时的质量），也就是说 m 并不等于 $\sum m_\alpha$ 。在相对论力学中，质量守恒定律不再成立，复合物体的质量不再等于其各部分质量之和，而只有包含了粒子静能在内的能量守恒定律是成立的。

利用式(2.4)和式(2.7)消去粒子的速度，可得自由实物粒子能量与动量之间的关系

$$\frac{\mathcal{E}^2}{c^2} = p^2 + m^2 c^2 \quad (2.9)$$

进而可获得粒子的 Hamilton 函数 \mathcal{H} ^①：

$$\mathcal{H} = c\sqrt{p^2 + m^2 c^2} \quad (2.10)$$

对于低速情况 $\frac{p}{mc} \ll 1$ ，近似地有

$$\mathcal{H} \approx mc^2 + \frac{p^2}{2m}$$

扣除静能后，Hamilton 函数也退化为经典结果。

利用式(2.4)和式(2.7)我们可以得到一个自有例子的能量、动量和速度之间的关系为

$$\mathbf{p} = \frac{\mathcal{E}\mathbf{v}}{c^2} \quad (2.11)$$

当 $v = c$ 时，粒子的动量和能量都变为无穷大，因此，一个粒子如果其质量不为零，就不可能以光速运动。在相对论力学中，可以存在质量为零同时以光速运动的粒子（例如光子和中微子）。根据式(2.11)，我们有

$$p = \frac{\mathcal{E}}{c} \quad (2.12)$$

同样的公式对于非零质量的粒子在**极端相对论**情况下**近似成立**，此时粒子的能量将远远大于其静能 mc^2 。

下面来推导上面所得到所有关系式的四维形式。根据最小作用量原理

$$\delta S = -mc\delta \int_a^b ds$$

由于 $ds = \sqrt{dx^i dx_i}$ ，故有

$$\begin{aligned} \delta S &= -mc \int_a^b \frac{dx_i \delta dx^i}{ds} = -m \int_a^b u_i \delta dx^i = -m u_i \delta x^i \Big|_a^b + m \int_a^b \delta x^i \frac{du_i}{ds} ds \\ &= -m u_i \delta x^i \Big|_a^b + \frac{m}{c} \int_a^b \delta x^i \frac{du_i}{d\tau} d\tau \end{aligned}$$

① 此处用 \mathcal{H} 而不用 H 来表示 Hamilton 函数是为了避免与后面涉及的磁场强度 \mathbf{H} 相混淆。

由 $\delta S = 0$ 和固定边界条件可得自由实物粒子的运动方程即为 $\frac{du_i}{d\tau} = 0$ 。

但现在我们需要将作用量的变分表示为坐标的函数，此时需将 a 点当做固定的，令粒子按照实际的轨道运动，考察 δS 与 $(\delta x^i)_b$ 之间的关系。根据 b 点所取的任意性，将 $(\delta x^i)_b$ 记作 δx^i ，则有

$$\delta S = -mu_i \delta x^i \quad (2.13)$$

由此，称四维矢量

$$p_i = -\frac{\partial S}{\partial x^i} = mu_i \quad (2.14)$$

为四维动量矢量。在经典力学中，导数 $\frac{\partial S}{\partial x}$, $\frac{\partial S}{\partial y}$, $\frac{\partial S}{\partial z}$ 是粒子动量矢量 \mathbf{p} 的三个分量，而导数 $-\frac{\partial S}{\partial t}$ 是粒子的能量 \mathcal{E} 。因此，即有四维动量的协变分量为

$$p_i = \left(\frac{\mathcal{E}}{c}, -\mathbf{p} \right) \quad (2.15)$$

而逆变分量则为

$$p^i = \left(\frac{E}{c}, \mathbf{p} \right) \quad (2.16)$$

因此，在相对论力学中，能量与动量是同一个四维矢量的分量。故可直接由此获得能量与动量由一个惯性系变换到另一个惯性系时的变换公式，即

$$\begin{cases} p_x = \frac{p'_x + \frac{V}{c^2} \mathcal{E}'}{\sqrt{1 - \frac{V^2}{c^2}}} \\ p_y = p'_y \\ p_z = p'_z \\ \mathcal{E} = \frac{\mathcal{E}' + V p'_x}{\sqrt{1 - \frac{V^2}{c^2}}} \end{cases} \quad (2.17)$$

式中 p_x, p_y, p_z 是三维矢量 \mathbf{p} 的分量（即四维矢量 p^i 的逆变分量 p^1, p^2 和 p^3 ）。

四维动量满足

$$p^i p_i = m^2 u^i u_i = m^2 c^2 \quad (2.18)$$

将式(2.15)和式(2.16)代入，即可得到关系式(2.9)。

而类比于力的通常定义，四维力矢量可以定义为

$$f^i = \frac{dp^i}{d\tau} = m \frac{du^i}{d\tau} \quad (2.19)$$

其满足恒等式

$$f^i u_i = 0 \quad (2.20)$$

这个四维矢量可以用通常的三维力矢量 $\mathbf{F} = \frac{d\mathbf{p}}{dt}$ 表示为

$$f^i = \left(\frac{\mathbf{f} \cdot \mathbf{v}}{c \sqrt{1 - \frac{v^2}{c^2}}}, \frac{\mathbf{f}}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \quad (2.21)$$

其时间分量与外力对该粒子的功率相关。

将式(2.14)代入式(2.18)就得到相对论的 Hamilton-Jacobi 方程

$$\frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x_i} = \eta^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = m^2 c^2 \quad (2.22)$$

或者将其展开为

$$\frac{1}{c^2} \left(\frac{\partial S}{\partial t} \right)^2 - \left(\frac{\partial S}{\partial x} \right)^2 - \left(\frac{\partial S}{\partial y} \right)^2 - \left(\frac{\partial S}{\partial z} \right)^2 = m^2 c^2 \quad (2.23)$$

将方程(2.23)过渡到经典力学极限需要首先作变换

$$S = S' - mc^2 t \quad (2.24)$$

将 S' 作为新的作用量。这是由于相对论力学中一个粒子的能量包括了一个经典力学中没有的 mc^2 项。将式(2.24)代入(2.23)中, 可得

$$\frac{1}{2mc^2} \left(\frac{\partial S'}{\partial t} \right)^2 - \frac{\partial S'}{\partial t} - \frac{1}{2m} \left[\left(\frac{\partial S'}{\partial x} \right)^2 + \left(\frac{\partial S'}{\partial y} \right)^2 + \left(\frac{\partial S'}{\partial z} \right)^2 \right] = 0$$

在 $c \rightarrow +\infty$ 的极限情况下, 这个方程就退化为经典的 Hamilton-Jacobi 方程。

2.3 分布函数的变换

在许多物理问题中都需要处理粒子动量的分布函数 $f(\mathbf{p})$, 乘积 $f(\mathbf{p})d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 表示动量分量在给定间隔 $d\mathbf{p}_x, d\mathbf{p}_y, d\mathbf{p}_z$ 内的粒子数, 即在动量空间内体积元 $d^3\mathbf{p} = d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 内的粒子数。于是当我们从一个参考系变换到另一个参考系中时, 需要同时获得分布函数 $f(\mathbf{p})$ 的变换。

因为 $f(\mathbf{p})d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 表示粒子数, 这显然是一个不变量, 所以只需要确定“体积元” $d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 在 Lorentz 变换下的性质。引入一个四维坐标系, 其坐标为四维动量矢量, 则 $d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 可以看作是由方程 $p^i p_i = m^2 c^2$ 定义的超曲面元的零分量。这个超曲面元是沿该超曲面法线指向的四维矢量。超曲面 $p^i p_i = m^2 c^2$ 的法线方向与四维矢量 p_i 一致。由此可知, 比值

$$\frac{d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z}{\mathcal{E}} \quad (2.25)$$

是一个不变量, 因为它是两个平行四维矢量的对应分量的比值。由于 $f(\mathbf{p})d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z$ 也是不变量, 将其写作

$$f(\mathbf{p}) \mathcal{E} \frac{d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z}{\mathcal{E}}$$

由此可得乘积 $f(\mathbf{p})\mathcal{E}$ 是不变量。因此可以得到 K' 系中的分布函数与 K 系中分布函数的关系为

$$f'(\mathbf{p}')\mathcal{E}' = f(\mathbf{p})\mathcal{E}$$

由此即有

$$f'(\mathbf{p}') = \frac{f(\mathbf{p})\mathcal{E}}{\mathcal{E}'} \quad (2.26)$$

其中 \mathbf{p} 和 \mathcal{E} 需要利用变换公式(2.17)表示。

在气体动理论中, 分布函数则表示为 $f(\mathbf{r}, \mathbf{p})$, 乘积 $f(\mathbf{r}, \mathbf{p})d\mathbf{p}_x d\mathbf{p}_y d\mathbf{p}_z dV$ 表示体积元 dV 中动量在间隔 $d\mathbf{p}_x, d\mathbf{p}_y, d\mathbf{p}_z$ 内的粒子数。函数 $f(\mathbf{r}, \mathbf{p})$ 称为相空间中的分布函数, 微分形式 $d\tau = d^3\mathbf{p}dV$ 是这个空间的体积元。

方便起见, 再引入一个具有给定动量的粒子在其中处于静止的参考系 K_0 , 粒子在该参考系中占有的体积元即为其固有体积 dV_0 。由于 K_0 相对于该粒子是静止的, 所以 K 系和 K' 系相对于 K_0 系的速度与这些粒子在 K 和 K' 系中的速度相同, 因此有

$$dV = dV_0 \sqrt{1 - \frac{v^2}{c^2}}, \quad dV' = dV_0 \sqrt{1 - \frac{v'^2}{c^2}}$$

由此即有

$$\frac{dV}{dV'} = \frac{\sqrt{1 - \frac{v^2}{c^2}}}{\sqrt{1 - \frac{v'^2}{c^2}}} = \frac{\mathcal{E}'}{\mathcal{E}}$$

又由于 $\frac{d^3p}{\mathcal{E}} = \frac{d^3p'}{\mathcal{E}'}$, 即有

$$d\tau = d\tau'$$

即相空间的体积元是不变量^①。因为粒子数 $f(\mathbf{r}, \mathbf{p})d\tau$ 根据定义也是不变量, 因而可得: 相空间的分布函数是不变量, 即

$$f'(\mathbf{r}', \mathbf{p}') = f(\mathbf{r}, \mathbf{p}) \quad (2.27)$$

式中的 \mathbf{r}', \mathbf{p}' 由 Lorentz 变换公式与 \mathbf{r}, \mathbf{p} 相联系。

2.4 粒子的衰变

现在来考虑一个质量为 M 的粒子自发衰变为质量分别为 m_1 和 m_2 两部分的情形。将能量守恒定律应用于原粒子处于静止的参考系中, 可得

$$Mc^2 = \mathcal{E}_{10} + \mathcal{E}_{20} \quad (2.28)$$

式中 \mathcal{E}_{10} 和 \mathcal{E}_{20} 是出射粒子的能量。因为它们的动量都非零, 故 $\mathcal{E}_{10} > m_1 c^2, \mathcal{E}_{20} > m_2 c^2$, 因此仅当 $M > m_1 + m_2$ 时, 式(2.28)才能满足, 即一个粒子可以自发衰变为质量之和小于该粒子的两部分。另一方面, 如果 $M < m_1 + m_2$, 则粒子对于该类型的衰变是稳定的, 不会自发衰变。如果需要在这种情况下人为引起衰变, 必须从外界对其注入至少等于其“束缚能” $(m_1 + m_2 - M)c^2$ 的能量。

衰变过程中, 动量也是守恒的。因为粒子的初始动量是零, 故出射粒子的动量之和也是零, 即

$$\mathbf{p}_{10} + \mathbf{p}_{20} = \mathbf{0} \quad (2.29)$$

因而有 $p_{10}^2 = p_{20}^2$, 即有

$$\mathcal{E}_{10}^2 - m_1^2 c^4 = \mathcal{E}_{20}^2 - m_2^2 c^4 \quad (2.30)$$

式(2.28)和式(2.30)唯一决定了出射粒子的能量:

$$\begin{cases} \mathcal{E}_{10} = \frac{M^2 + m_1^2 - m_2^2}{2M} c^2 \\ \mathcal{E}_{20} = \frac{M^2 - m_1^2 + m_2^2}{2M} c^2 \end{cases} \quad (2.31)$$

^① 此处相当于是将经典统计力学中的 Liouville 定理推广至了相对论情形。

在一定意义上, 这个问题的逆问题就是计算两个碰撞粒子在其总动量为零的参考系^①中的总能量 Mc^2 。这个量的计算给出了伴随碰撞粒子状态改变或新粒子“产生”的各种非弹性碰撞过程可能存在的判据。仅当“反应产物”的质量之和不超过 M 时, 这类过程才能够发生。

假设在初始参考系(即实验室系)中, 一个质量为 m_1 能量为 \mathcal{E}_1 的粒子同另一个质量为 m_2 的静止粒子相撞, 这两个粒子的总能量是

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 = \mathcal{E}_1 + m_2 c^2$$

它们的总动量则为

$$\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_1$$

将两个粒子看成一个复合系统, 由式(2.11)可得其整体的运动速度为

$$\mathbf{V} = \frac{\mathbf{p}c^2}{\mathcal{E}} = \frac{\mathbf{p}_1 c^2}{\mathcal{E} + m_2 c^2} \quad (2.32)$$

这个速度即为 C 系相对于实验室系 (L 系) 的运动速度。

当然在测定 M 时, 没有必要从一个参考系变换到另一个参考系^②。直接利用实物粒子的能量动量关系(2.9)即有

$$M^2 c^4 = \mathcal{E}^2 - p^2 c^2 = (\mathcal{E}_1 + m_2 c^2)^2 - (\mathcal{E}_1^2 - m_1^2 c^4)$$

由此可有

$$M^2 = m_1^2 + m_2^2 + \frac{2m_2 \mathcal{E}_1}{c^2} \quad (2.33)$$

2.5 粒子的弹性碰撞

现在我们从相对论力学的观点来考虑粒子的弹性碰撞。将两个碰撞粒子的质量、动量和能量分别记作 $m_1, \mathbf{p}_1, \mathcal{E}_1$ 和 $m_2, \mathbf{p}_2, \mathcal{E}_2$, 用带撇号的相应量表示碰撞后的量。

碰撞过程中的动量守恒定律可以一并写为四维动量守恒方程

$$p_1^i + p_2^i = p_1'^i + p_2'^i \quad (2.34)$$

从这个四维矢量方程可以构造出有助于进一步计算的不变关系式。将式(2.34)改写为形式

$$p_1^i + p_2^i - p_1'^i = p_2'^i \quad (2.35)$$

将其两边平方, 即有

$$(p_1^i + p_2^i - p_1'^i)(p_{1i} + p_{2i} - p_{1i}') = p_2'^i p_{2i}'$$

注意到其中 $p_1^i p_{1i} = p_1'^i p_{1i}' = m_1^2 c^4$, $p_2^i p_{2i} = p_2'^i p_{2i}' = m_2^2 c^4$, 可得

$$m_1^2 c^4 + p_{1i} p_2^i - p_{1i} p_1'^i - p_{2i} p_1'^i = 0 \quad (2.36)$$

类似的将式 $p_1^i + p_2^i - p_2'^i = p_1'^i$ 两边平方, 可得

$$m_2^2 c^4 + p_{1i} p_2^i - p_{2i} p_2'^i - p_{1i} p_2'^i = 0 \quad (2.37)$$

^① 这个参考系简称为动量中心系或“C 系”。

^② 因为质量 M 是四维标量。

2.6 角动量

第三章 引力场中的粒子

3.1 非相对论力学中的引力场

3.2 相对论力学中的引力场

3.3 曲线坐标

3.4 距离与时间间隔

3.5 协变微分

3.6 Christoffel 符号与度规张量的关系

3.7 引力场中粒子的运动

3.8 恒定引力场

3.9 旋转

第四章 引力场方程

4.1 曲率张量

4.2 曲率张量的特性

4.3 引力场的作用量

4.4 能量动量张量

4.5 Einstein 方程

4.6 引力场的能动张量

4.7 同步参考系

4.8 Einstein 方程的标架表示

第五章 引力物体的场

5.1 Newton 定律

5.2 中心对称的引力场

5.3 中心对称的引力场中的运动

5.4 球形物体的引力坍缩

5.5 类尘埃球的坍缩

5.6 非球形转动物体的引力坍缩

5.7 物體遠距離處的引力場

5.8 二級近似下物體系統的運動方程

第五章 插图目录

第五章 参考文献

- [1] 郎道, 栗弗席兹. 力学 (第五版). 北京: 高等教育出版社, 2013.
- [2] 郎道, 栗弗席兹. 场论 (第八版). 北京: 高等教育出版社, 2012.