

东华理工大学理学院物理系

# 第 3 章 对称性

v. 1.0.0

陈童

tongchen@ecut.edu.cn

July 31, 2021



## 什么是对称性？

### 对称性与守恒定律

对称性就是作用量的不变性

诺特定理

动量守恒、角动量守恒、能量守恒

### 对称性决定作用量

相对论不变性

与电磁场的耦合

# 什么是对称性



设想你在一艘太空飞船里，离所有的星星都很远，船窗密闭，你看不到外面，也无法接收外界的信息，并且假设推进器已经不向外喷射物质了，总之，假设你和飞船成了一个真正的封闭系统。不过，只要局限于飞船之内，你可以做任何事情。请问，你是否能判断飞船在太空中的坐标，飞船的朝向，以及飞船是否在飞行？



答案当然是不能，你的这种不能就是物理定律的对称性。简言之，物理定律的对称性，就是对某些物理性质的无法辨测性。伽利略最早领悟了对称性在物理学中的重要性，伽利略的这种领悟后来被爱因斯坦发扬光大，最终成为现代物理学的基石。那么，为什么无法辨测呢？你为什么无法辨测飞船的坐标和方向呢？原因在于，即使把飞船平移一段距离并旋转一个角度，飞船之中的你也无法觉察到任何的不同。即是说，这种无法辨测性起因于你可以对系统进行一些操作，但操作前后你归纳出来的物理定律完全一样，有完全一样的数学形式。所以，物理定律的对称性其实就是物理定律的形式在操作前后的不变性。

# 对称性与守恒定律

对称性就是作用量的不变性



所谓的某系统有一个对称性，我们是指（为了公式简洁起见，我们省略广义坐标上面的分量指标）：首先，系统按照某条物理路径  $q(t)$  在位形空间中演化，这条物理路径满足基本方程  $\delta S[q(t)] = 0$ 。其次，假设在对称操作的作用之下，原来的  $q(t)$  演化路径变换成了新的路径  $\tilde{q}(t) = F(q(t))$ 。所谓的对称操作，我们是指这条新的路径满足同样形式的基本方程  $\delta S[\tilde{q}(t)] = 0$ 。即，对于对称操作作用前后的两条路径  $q(t)$  和  $\tilde{q}(t) = F(q(t))$ ，有

$$\delta S[q(t)] = 0 \Leftrightarrow \delta S[\tilde{q}(t)] = 0. \quad (1)$$

很显然，满足这一要求的最直接方式是

$$S[\tilde{q}(t)] = S[q(t)]. \quad (2)$$

即是说，在对称操作的作用下，作用量泛函保持不变，**对称性就是作用量的不变性。**

# 对称性与守恒定律

## 诺特定理



对称性可以分成两类，像镜像对称这样的只包含有限个不同操作的对称性，称为离散对称性。像空间旋转这样的对称性称为连续对称性，因为它有无穷多不同的对称操作（这里即是旋转操作），这些对称操作依赖于某些连续的参数（在这里就是旋转角度）。本章主要考察连续对称性。原因之一在于，连续对称性出人意料地和所谓的守恒定律密切相关，可以证明，每一种连续对称性都对应一条守恒定律，这就是著名的诺特定理。

诺特定理的证明并不复杂，下面我们给出一个绝妙的证明。下面的证明假设了对称操作不改变时间  $t$ 。不过，即使将一般情形包括进来证明的关键也还是类似的。为了让读者看清楚证明的关键，这里先处理简化情形，然后，我们将在下一节中用对时间平移对称性的讨论说清楚如何推广到更一般情形。



假设系统有某种连续对称性 (即是说, 系统的作用量在这样的连续操作之下保持不变), 记为  $G(\theta)$ ,  $\theta$  就是这种对称性所依赖的连续参数,  $\theta = 0$  表示对称操作还来不及加上, 即表示不进行任何操作, 也称恒等操作。下面考虑参数  $\theta = \epsilon$  的无穷小对称操作,  $\epsilon$  为无穷小量。假设在此对称操作之下, 时间  $t$  保持不变, 但  $q$  变为  $\tilde{q}$

$$q(t) \rightarrow \tilde{q}(t) = q(t) + \epsilon F(q(t)), \quad (3)$$

式中  $F(q(t))$  为某个表达式, 它的形式取决于对称操作的具体定义。则根据对称性的定义, 作用量泛函将在此对称操作下保持不变, 即

$$\delta S = S[\tilde{q}(t)] - S[q(t)] = \int dt L(\tilde{q}, \dot{\tilde{q}}) - \int dt L(q, \dot{q}) = 0. \quad (4)$$



下面是关键的证明技巧。前面的  $\epsilon$  是一个无穷小常数，现在，设想我们将这个无穷小常数变成一个关于  $t$  的任意函数  $\epsilon(t)$  (函数值为无穷小量)，并要求在路径的起末两端 (分别对应时间  $t_i$  和  $t_f$ )，这个函数取值为零，即

$$\epsilon(t_i) = \epsilon(t_f) = 0. \quad (5)$$

我们考虑修改的变换

$$\begin{aligned} q(t) &\rightarrow \tilde{q}(t) = q(t) + \epsilon(t)F(q(t)), \\ \text{即 } \delta q &= \tilde{q}(t) - q(t) = \epsilon(t)F(q(t)). \end{aligned} \quad (6)$$

当然，现在因为  $\epsilon(t)$  不再是对称变换参数 (对称变换参数不依赖于  $t$ )，所以这个变换并不是一个对称性，因此不能保持作用量不变，但是如果  $\epsilon(t)$  变回常数函数，那作用量将是不变的。





这就意味着，在 (6) 的变换下，作用量的改变量必然具有如下形式

$$\delta S = \int_{t_i}^{t_f} dt (L(\tilde{q}, \dot{\tilde{q}}) - L(q, \dot{q})) = \int_{t_i}^{t_f} dt Q(q, \dot{q}) \dot{\epsilon}(t). \quad (7)$$

式中  $Q(q, \dot{q})$  为某个表达式。我们简单解释一下为什么是这样，原因有三：第一，作用量是某个函数（拉格朗日量）对时间的积分，所以作用量的改变量必然具有时间积分的形式。第二，拉格朗日量中只含有  $q(t)$  的一阶导数，因此作用量的改变量中最多只含有  $\epsilon(t)$  的一阶导数。第三，将作用量的改变量计算到一阶无穷小项，那这个改变量必然不能含  $\epsilon(t)$ ，因为这样的项在  $\epsilon(t)$  变回常数函数时不能变成零。所以这个改变量的一阶小量只能含  $\dot{\epsilon}$ （因此当  $\epsilon(t)$  变回常数函数时自动得到零）。综上，我们必有表达式 (7)。

# 对称性与守恒定律

诺特定理



上面讨论中的路径  $q(t)$  是任意路径，不必是真实演化路径。现在，假设我们取  $q(t)$  为真实的物理演化路径，那最小作用量原理将告诉我们，对于任何两端固定的无穷小改变 (变分)，将有  $\delta S = 0$ 。(6) 正是这样的一个两端固定的无穷小变分，因此这就意味着，对于真实的物理路径  $q(t)$ ，(7) 式得出的作用量改变量必定恒为零，即

$$\delta S = \int_{t_i}^{t_f} dt Q(q, \dot{q}) \epsilon(t) = 0. \quad (8)$$

将上式分部积分，就可以得到

$$\delta S = - \int_{t_i}^{t_f} dt \dot{Q}(q, \dot{q}) \epsilon(t) = 0. \quad (9)$$

但是， $\epsilon(t)$  的函数形式是任意的，因此这就说明，对于物理路径  $q(t)$ ，我们有

$$\frac{dQ}{dt} = 0. \quad (10)$$

即  $Q(q, \dot{q})$  为一个守恒量！可见，相应于每一个连续的对称性，都必定存在一个守恒量，这，就是诺特定理！

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



动量守恒、角动量守恒、还有能量守恒，这是最为人们熟知的守恒定律，也是最为基本的物理学定律之一。这些定律为什么如此基本呢？按照诺特定理的精神，原因应该在于导致这些守恒定律的对称性很基本，那么它们分别是什么对称性呢？结果表明，动量守恒是所谓的空间平移不变性的结果，角动量守恒是空间旋转不变性的结果，它们也就是本章一开始所谈的宇宙飞船位置坐标和空间方向的无法辨测性。而能量守恒则是时间平移不变性的结果，所谓的时间平移不变性也就是时间的零点是人们任意规定的，无法通过物理定律辨测时间的零点，因为明天做实验将和今天做实验得到同样的规律。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



我们先来讨论空间平移对称性和动量守恒。考虑一个多粒子系统，记其拉格朗日量为  $L(\mathbf{x}_i, \dot{\mathbf{x}}_i)$  ( $L(\mathbf{x}_1, \dots, \mathbf{x}_N, \dot{\mathbf{x}}_1, \dots, \dot{\mathbf{x}}_N)$  的简记)，因此系统的作用量为  $S = \int dt L(\mathbf{x}_i, \dot{\mathbf{x}}_i)$ 。假设系统具有空间平移不变性，即在如下坐标平移下，作用量保持不变

$$\mathbf{x}_i \rightarrow \tilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{a}. \quad (11)$$

很显然，这里作用量保持不变其实就相当于拉格朗日量在坐标平移下保持不变，即有  $L(\mathbf{x}_i + \mathbf{a}, \dot{\mathbf{x}}_i) = L(\mathbf{x}_i, \dot{\mathbf{x}}_i)$ 。考察无穷小平移，即取平移量  $\mathbf{a} = \epsilon$  为无穷小量，从而有

$$0 = \delta L = L(\mathbf{x}_i + \epsilon, \dot{\mathbf{x}}_i) - L(\mathbf{x}_i, \dot{\mathbf{x}}_i) = \epsilon \cdot \sum_j \frac{\partial L}{\partial \mathbf{x}_j}. \quad (12)$$

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



下面使用关键技巧，即将无穷小参数  $\epsilon$  变成  $\epsilon(t)$ ，即考察

$$\mathbf{x}_j \rightarrow \tilde{\mathbf{x}}_j = \mathbf{x}_j + \epsilon(t). \Leftrightarrow \delta \mathbf{x}_j = \epsilon(t). \quad (13)$$

从而容易计算出作用量的改变量，有

$$\begin{aligned} \delta S &= \int dt \delta L = \int dt \left[ \sum_j \frac{\partial L}{\partial \mathbf{x}_j} \cdot \delta \mathbf{x}_j + \sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \delta \dot{\mathbf{x}}_j \right] \\ &= \int dt \left[ \sum_j \frac{\partial L}{\partial \mathbf{x}_j} \cdot \epsilon + \sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \cdot \dot{\epsilon} \right] \end{aligned} \quad (14)$$

利用平移不变性的 (12) 式，即有

$$\delta S = \int dt \left( \sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \right) \cdot \dot{\epsilon}. \quad (15)$$

根据上一节关于诺特定理的证明容易知道，上式说明  $(\sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j})$  是守恒量。根据上一章关于勒让德变换的讨论， $\frac{\partial L}{\partial \dot{\mathbf{x}}_j}$  就是第  $j$  个粒子的动量  $\mathbf{p}_j$ ，即  $\mathbf{p}_j = \frac{\partial L}{\partial \dot{\mathbf{x}}_j}$ 。所以，守恒的正是系统的总动量  $\mathbf{p}$ 。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



13

下面讨论旋转不变性和角动量守恒。这一次我们考察单粒子情形 (推广到多粒子情形是直接了当的), 不过, 我们假设这个粒子在  $d$  维空间中运动 (即不限于通常的 3 维空间)。这里依然采用直角坐标, 从而粒子的位置矢量  $\mathbf{x} = (x^1, x^2, \dots, x^d)$ , 我们将各分量记为  $x^\alpha, \alpha = 1, 2, \dots, d$ 。这里采用了上指标来标识各直角坐标分量, 但其实对于欧氏空间的直角坐标, 用上指标还是下指标并没有区别, 即有  $x_\alpha = x^\alpha$ 。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



14

首先需要研究一下如何处理  $d$  维空间中的无穷小空间旋转。记无穷小旋转操作之下，坐标的改变量为  $\delta x^\alpha$ ，很显然， $\delta x^\alpha$  应该与原来的矢量  $\{x^\alpha\}$  成线性关系，我们设

$$\delta x_\alpha = \epsilon_{\alpha\beta} x^\beta. \quad (16)$$

式中的  $\epsilon_{\alpha\beta}$  为刻画旋转的无穷小量。很明显，空间旋转不会改变一个矢量的长度，从而  $\delta(x^2) = 0$ ，即有

$$0 = \delta(x^2)/2 = \mathbf{x} \cdot \delta \mathbf{x} = x^\alpha \delta x_\alpha = \epsilon_{\alpha\beta} x^\alpha x^\beta. \quad (17)$$

注意到  $x^\alpha x^\beta$  关于指标  $\alpha, \beta$  对称，所以满足上式的唯一可能性是， $\epsilon_{\alpha\beta}$  关于两个指标反对称<sup>1</sup>，即

$$\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha}. \quad (18)$$

所以， $d$  维空间中独立的无穷小旋转有  $d(d-1)/2$  个，对于  $d=3$  的三维空间，独立的无穷小旋转数目刚好是 3！

<sup>1</sup>因为反对称和对称乘在一起才等于零，证明如下：设  $A_{\alpha\beta}$  关于两指标反对称， $B^{\alpha\beta}$  关于两指标对称，即  $A_{\alpha\beta} = -A_{\beta\alpha}$ ， $B^{\alpha\beta} = B^{\beta\alpha}$ 。则有  $A_{\alpha\beta} B^{\alpha\beta} = -A_{\beta\alpha} B^{\alpha\beta} = -A_{\beta\alpha} B^{\beta\alpha} = -A_{\alpha\beta} B^{\alpha\beta}$ ，从而  $A_{\alpha\beta} B^{\alpha\beta} = 0$ 。以上推导的最后一步是重命名了一下指标。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



因此，在无穷小旋转操作之下，我们有

$$\delta x_\alpha = \epsilon_{\alpha\beta} x^\beta, \quad \delta \dot{x}_\alpha = \epsilon_{\alpha\beta} \dot{x}^\beta. \quad (19)$$

如果系统具有旋转不变性，从而在上面的无穷小旋转之下拉格朗日量保持不变，即有  $\delta L(\mathbf{x}, \dot{\mathbf{x}}) = 0$ 。

下面将  $\epsilon_{\alpha\beta}$  变成  $\epsilon_{\alpha\beta}(t)$ ，那现在  $\delta \dot{x}_\alpha$  就会多出一项，变成 (注意我们说过，上下指标其实没有区别)

$$\delta \dot{x}^\alpha = \epsilon^{\alpha\beta} \dot{x}_\beta + \dot{\epsilon}^{\alpha\beta} x_\beta. \quad (20)$$



# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



从而现在拉格朗日量就不再是不变了，其改变量会多出一个非零项，变成

$$\begin{aligned}\delta L &= \frac{\partial L}{\partial \dot{x}^\beta} \dot{\epsilon}^{\beta\alpha} x_\alpha = p_\beta x_\alpha \dot{\epsilon}^{\beta\alpha} \\ &= -\frac{1}{2} (x_\alpha p_\beta - x_\beta p_\alpha) \dot{\epsilon}^{\alpha\beta}.\end{aligned}\quad (21)$$

式中已经利用了动量的定义式  $p_\beta = \frac{\partial L}{\partial \dot{x}^\beta}$ 。当然，现在作用量也不是不变的，而是会改变为

$$\delta S = -\frac{1}{2} \int dt (x_\alpha p_\beta - x_\beta p_\alpha) \dot{\epsilon}^{\alpha\beta} \quad (22)$$

根据上一节诺特定理证明的有关讨论可以知道，上式意味着  $x_\alpha p_\beta - x_\beta p_\alpha$  是一个守恒量，习惯上记为  $J_{\alpha\beta}$ ，

$$J_{\alpha\beta} = x_\alpha p_\beta - x_\beta p_\alpha. \quad (23)$$

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



17

从而  $d$  维空间的旋转不变性意味着存在  $d(d-1)/2$  个守恒量，为  $J_{\alpha\beta}$ ，这些守恒量就是所谓的角动量。如果推广到多粒子情形，那守恒的就是总角动量。为了看清楚  $J_{\alpha\beta}$  的确是角动量，我们取  $d=3$ ，这时候有 3 个守恒量，分别为

$$J_{12} = xp_y - yp_x, \quad J_{23} = yp_z - zp_y, \quad J_{31} = zp_x - xp_z. \quad (24)$$

显然这正是通常的角动量，通常分别记作  $J^3, J^1, J^2$ 。利用上一章引入的列维-西维塔记号  $\epsilon^{ijk}$  (请不要和上文的各无穷小量混淆)，我们可以将 3 维空间角动量这两种写法之间的关系写为

$$J^i = \frac{1}{2} \epsilon^{ijk} J_{jk}. \quad (25)$$

不过，很显然，高维空间的角动量只能记为  $J_{\alpha\beta}$ 。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



18

我们知道，封闭系统的能量守恒，这条守恒定律源自于哪种对称性呢？结果表明能量守恒源自于封闭系统所具有的时间平移不变性，即你是今天观测这个系统还是明天观测这个系统得到的物理规律是一样的。从作用量和拉格朗日量来看的话，时间平移不变性来源于拉格朗日量不显含时间  $t$ ，从而在  $t \rightarrow \tilde{t} = t + a$  的时间平移下（相当于将整个系统往过去方向移动了  $a$ ，所以其时间坐标增加了  $a$ ），我们有  $q(t) \rightarrow \tilde{q}(t) = q(\tilde{t})$ ，从而

$$\begin{aligned} S[q(t)] &= \int_{t_i}^{t_f} dt L(q(t), \dot{q}(t)) \rightarrow \\ S[q(\tilde{t})] &= \int_{t_i - a}^{t_f - a} dt L(q(\tilde{t}), \frac{d}{dt} q(\tilde{t})) = \int_{t_i}^{t_f} d\tilde{t} L(q(\tilde{t}), \dot{q}(\tilde{t})) \\ &= \int_{t_i}^{t_f} dt L(q(t), \dot{q}(t)) = S[q(t)], \end{aligned}$$

式中最后一个等号只是将  $\tilde{t}$  重命名为了  $t$ 。可见，正因为  $L(q, \dot{q})$  不显含  $t$ ，所以作用量在时间平移下保持不变。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



下面考虑无穷小时间平移，并且和前面一样，将无穷小参数变为  $\epsilon(t)$ ，即考虑如下时间变换

$$t \rightarrow \tilde{t}(t) = t + \epsilon(t), \quad (26)$$

注意  $\tilde{t}$  为  $t$  的函数。现在，变换以后的作用量  $S[q(\tilde{t})]$  为

$$\begin{aligned} S[q(\tilde{t})] &= \int dt L(q(\tilde{t}), \frac{d}{dt}q(\tilde{t})) = \int d\tilde{t} \left( \frac{dt}{d\tilde{t}} \right) L(q(\tilde{t}), \left( \frac{d\tilde{t}}{dt} \right) \dot{q}(\tilde{t})) \\ &= \int d\tilde{t} \left( \frac{1}{1+\dot{\epsilon}} \right) L(q(\tilde{t}), (1+\dot{\epsilon})\dot{q}(\tilde{t})) \\ &= \int d\tilde{t} L(q(\tilde{t}), \dot{q}(\tilde{t})) - \int d\tilde{t} (\dot{\epsilon}) L(q(\tilde{t}), \dot{q}(\tilde{t})) + \int d\tilde{t} \frac{\partial L}{\partial (\dot{q}(\tilde{t}))} \dot{q}(\tilde{t}) (\dot{\epsilon}) \\ &= \int dt L(q(t), \dot{q}(t)) + \int dt \left[ \frac{\partial L}{\partial (\dot{q}(t))} \dot{q}(t) - L(q(t), \dot{q}(t)) \right] \dot{\epsilon}, \end{aligned} \quad (27)$$

式中最后一行是将变量  $\tilde{t}$  重命名成了  $t$ 。

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



从上面计算可以得到，变换前后作用量的改变量为

$$\delta S = S[q(\tilde{t})] - S[q(t)] = \int dt \left[ \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}) \right] \dot{\epsilon}. \quad (28)$$

完全类似于上一节证明诺特定理时的推理可以知道，上式意味着存在下面的守恒量

$$E(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}). \quad (29)$$

很明显，如果我们利用  $p_a = \frac{\partial L}{\partial \dot{q}^a}$  将这个守恒量中的变量  $\dot{q}^a$  用广义动量  $p_a$  来表示，那这个守恒量就正是哈密顿量  $H(q, p)$ 。这说明，与时间平移对称性对应的守恒量  $E(q, \dot{q})$  在物理上就代表能量。所以，时间平移对称性意味着能量守恒！

# 对称性与守恒定律

动量守恒、角动量守恒、能量守恒



不妨举一例子。上一章在讲费马原理时我们引入过这样一个光程泛函  $S[y(x)] = \int dx L(y, y')$ , 式中

$$L(y, y') = n(y) \sqrt{1 + (y')^2}. \quad (30)$$

显然, 从数学上来说, 这个例子的变量  $x$  就相当于我们上面的  $t$ , 函数  $y(x)$  就相当于上面的  $q(t)$ 。如此来看的话, 这个例子显然也有“时间”平移不变性, 即光程泛函在  $x \rightarrow \tilde{x} = x + a$  的变换下不变。因此这个例子也有一个“守恒能量”, 为

$$E(y, y') = \frac{\partial L}{\partial y'} y' - L(y, y') = -\frac{n(y)}{\sqrt{1 + (y')^2}}. \quad (31)$$

也即是说,  $E(y, y')$  是一个不依赖于  $x$  的积分常数, 称之为欧拉-拉格朗日方程的首积分。显然, 这正是我们上一章中找到了的那个首积分。



到现在为止，我们应该已经比较清楚作用量的重要性了，有了作用量的具体表达式，利用最小作用量原理就能得出运动微分方程。有了作用量里的拉格朗日量，经过勒让德变换就是得到描述系统能量的哈密顿量。更一般的，作用量的不变性意味着物理定律的对称性。总之，从作用量出发讨论物理问题非常简洁，也非常方便。问题是，如何写出一个物理系统的作用量呢？回答是，基本靠猜。

更何况，通常来说物理学家猜作用量可不是漫无目的的，物理学家们作猜测时有一些基本的原则，其中最重要的就是根据对称性。只要我们能预先猜测出一个物理系统的对称性，那这些对称性往往就能决定(至少是极大地限制)作用量的形式，这是因为，作用量必须得在对称变换下保持不变，作用量泛函不是一个任意的泛函，而是一种具有对称性的特殊泛函。

# 对称性决定作用量

相对论不变性



23

在非相对论物理中对称性对作用量有所限制，但这个限制还不够强，这就是为什么爱因斯坦之前的物理学很少谈对称性的原因。但是，相对论有很大的不同，相对论意味着一种很强的时空对称性。相对论不变性再结合一些其它的对称性，往往就能决定作用量。所以杨振宁先生曾经说：“对称性决定相互作用”，指的就是对称性往往能决定作用量。为了给读者示例对称性如何决定作用量，下面我们从相对论不变性开始谈起。



# 对称性决定作用量

相对论不变性



24

在非相对论物理中对称性对作用量有所限制，但这个限制还不够强，这就是为什么爱因斯坦之前的物理学很少谈对称性的原因。但是，相对论有很大的不同，相对论意味着一种很强的时空对称性。相对论不变性再结合一些其它的对称性，往往就能决定作用量。所以杨振宁先生曾经说：“对称性决定相互作用”，指的就是对称性往往能决定作用量。为了给读者示例对称性如何决定作用量，下面我们从相对论不变性开始谈起。

从现在开始我们合适地选取时间的单位，使得光速  $c = 1$ 。因此质量的量纲（记为  $[m]$ ）和能量量纲（记为  $[E]$ ）相同，因为  $[m] = [mc^2] = [E]$ 。同样，时间的量纲和空间长度的量纲也相同。

# 对称性决定作用量

相对论不变性



25

如大家知道的，爱因斯坦基于两条基本的原理建立起了整个狭义相对论。这两条基本原理分别是：1. 相对性原理。即物理定律在所有的惯性系中有相同的形式。这当然是伽利略相对性原理的简单推广。2. 光速不变原理。即在任何给定的惯性系中，光速  $c$  总是一样的，并与光源的运动无关。

让我们考虑两个惯性系，分别称为  $S$  系和  $S'$  系，假定同一个事件在这两个参考系中的时空坐标分别是  $(t, x, y, z)$  和  $(t', x', y', z')$ 。那么根据相对性原理以及时空的均匀性，我们可以知道，这两个参考系之间的坐标变换一定是一个线性变换。

# 对称性决定作用量

相对论不变性



26

现在，设想在  $S$  系中，从  $(t, x, y, z)$  点发出一束光到达  $(t + dt, x + dx, y + dy, z + dz)$  点，由于  $c = 1$ ，显然我们有

$$-dt^2 + dx^2 + dy^2 + dz^2 = 0. \quad (32)$$

根据光速不变原理，同样的两个事件在  $S'$  系看来也得满足

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = 0. \quad (33)$$

换言之，两个邻近的事件在两个参考系中的时空坐标必得满足一个约束关系，即当 (32) 成立时必有 (33) 成立，反之亦然。又由于相对性原理所告诉我们的，两参考系之间的坐标变换是线性变换，因此，对任意的两个邻近事件，我们必有

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = D(\mathbf{v})(-dt^2 + dx^2 + dy^2 + dz^2), \quad (34)$$

式中  $\mathbf{v}$  是  $S'$  系相对于  $S$  系的速度。同样的，由于  $S$  与  $S'$  地位平等，如果从  $S'$  变换到  $S$ ，我们就有

$$-dt^2 + dx^2 + dy^2 + dz^2 = D(-\mathbf{v})(-dt'^2 + dx'^2 + dy'^2 + dz'^2). \quad (35)$$

换言之，我们必有  $D(-\mathbf{v})D(\mathbf{v}) = 1$ 。

# 对称性决定作用量

相对论不变性



又由于空间的各向同性可知， $D$  对相对速度  $\mathbf{v}$  的依赖只能是依赖于其大小  $v$ ，而必定和其方向无关，因此我们必定有

$D(-\mathbf{v}) = D(\mathbf{v}) = D(v)$ 。因此， $(D(v))^2 = 1$ ，即  $D(v) = \pm 1$ 。又由于  $D(v)$  是  $v$  的连续函数，而且  $D(0) = 1$  (对应  $S$  和  $S'$  变成同一个参考系)，因此我们必有  $D(v) = 1$ 。因此，对于任意两个邻近事件我们必有

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = -dt^2 + dx^2 + dy^2 + dz^2. \quad (36)$$

通常将  $-dt^2 + dx^2 + dy^2 + dz^2$  称为两个邻近事件的时空间隔平方，简称间隔，并记为  $-d\tau^2$ ，即

$$-d\tau^2 = -dt^2 + dx^2 + dy^2 + dz^2 = -dt^2 + dx^2. \quad (37)$$

用这个记号，方程 (36) 就可以简记成

$$d\tau'^2 = d\tau^2, \quad (38)$$

称为两个事件的间隔不变性。

# 对称性决定作用量

相对论不变性



28

假如有一个粒子从  $(t, x, y, z)$  运动到  $(t + dt, x + dx, y + dy, z + dz)$ , 那么其间隔  $d\tau^2$  就等于一个固定在粒子上的钟在这个过程中所走过的时间的平方 (因为在粒子本身的参考系中这个钟的空间坐标改变量为零), 所以这时候  $d\tau$  就是这个钟走过的时间, 称为粒子的固有时。

从这里的讨论可以看到, 间隔的定义和是否有粒子运动无关, 而固有时的定义却依赖于一个运动的粒子。因此, 这实际上是两个不同的概念, 人们常常用不同的记号表示它们。

# 对称性决定作用量

相对论不变性



29

人们通常约定  $t = x^0, x = x^1, y = x^2, z = x^3$  这样就把四个时空坐标统一地记成了  $x^\mu, \mu = 0, 1, 2, 3$ . 利用这个记号, 我们就可以将间隔的计算公式重写为

$$-d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \quad (39)$$

式中我们默认了求和约定(后文也都默认求和约定), 而  $\eta_{\mu\nu}$  为,  $-\eta_{00} = \eta_{11} = \eta_{22} = \eta_{33} = 1$ , 其它指标分量都等于零。 $\eta_{\mu\nu}$  称为四维闵可夫斯基时空的度规张量, 所谓的闵可夫斯基时空指的就是狭义相对论中的平直时空。值得注意的是,  $\eta_{\mu\nu}$  关于它的两个指标是对称的, 即满足  $\eta_{\mu\nu} = \eta_{\nu\mu}$ 。

# 对称性决定作用量

相对论不变性



$dx^\mu$  这样的四维时空的上指标量就称为四矢量，一个一般的四矢量可以记为  $A^\mu$ 。但是，我们也可以利用  $\eta_{\mu\nu}$  将  $A^\mu$  的指标降下来，即约定

$$A_\mu = \eta_{\mu\nu} A^\nu. \quad (40)$$

反过来，人们还会引入一个  $\eta^{\mu\nu}$ ，它同样满足  $-\eta^{00} = \eta^{11} = \eta^{22} = \eta^{33} = 1$ ，其它指标分量都等于零。利用  $\eta^{\mu\nu}$  我们也可以反过来将  $A_\mu$  的指标升上去，即约定

$$A^\mu = \eta^{\mu\nu} A_\nu. \quad (41)$$

人们也可以定义两个四矢量  $A^\mu, B^\mu$  之间的内积，记为  $A \cdot B$ ，其定义为

$$A \cdot B = A^\mu B_\mu = A_\mu B^\mu = \eta_{\mu\nu} A^\mu B^\nu. \quad (42)$$

两个四矢量的内积是参考系变换不变的，称为洛伦兹标量。

# 对称性决定作用量

相对论不变性



31

相对论不变性告诉我们，相对论粒子的作用量得是洛伦兹变换（即不同参考系的坐标变换）不变的！唯一的和粒子坐标有关的这种不变量就是粒子的固有时  $d\tau$ 。所以，粒子的作用量必定正比于  $d\tau$  沿着粒子运动路径的积分。固有时具有时间量纲，而作用量的量纲为能量量纲乘以时间量纲，刚好粒子质量  $m$  是能量量纲，从而我们知道，相对论粒子的作用量必定可以写成

$$S[x(s)] = -m \int d\tau = -m \int ds \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}}. \quad (43)$$

式中  $s$  为粒子运动路径的参数，第一个式子的负号有多个理解角度，这里我们想指出的是，有了这个负号，这个作用量的非相对论极限才会对，请大家在待会儿我们讨论非相对论极限时留心。



# 对称性决定作用量

相对论不变性



32

上述作用量 (43) 显然具有重参数不变性, 即在  $s \rightarrow \tilde{s}(s)$  的参数变换下保持不变。从而我们可以任意选择参数  $s$ 。常见的选择有两种, 第一种是, 取  $s = \tau$ , 从而作用量 (43) 可以写成

$$S[x(\tau)] = -m \int d\tau = -m \int d\tau \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}}. \quad (44)$$

第二种选择是取  $s = x^0 = t$ , 即取通常的时间坐标为参数, 这时候注意到

$$d\tau^2 = dt^2 - dx^2 = dt^2(1 - v^2). \quad (45)$$

式中  $v = \frac{dx}{dt}$  是粒子的速度。从而就可以将作用量 (43) 写成

$$S[x(t)] = -m \int dt \sqrt{1 - v^2}. \quad (46)$$

假设粒子的运动速度远低于光速, 即  $v \ll 1$ , 那这时候就可以利用关于  $v^2$  的泰勒展开将相对论的作用量 (46) 近似成

$$S = -m \int dt + \int dt \frac{1}{2} m v^2 + \dots \quad (47)$$

# 对称性决定作用量

相对论不变性



33

另外，上述相对论粒子作用量显然具有  $x^\mu \rightarrow \tilde{x}^\mu = x^\mu + a^\mu$  ( $a^\mu$  为常矢量) 的时空坐标平移不变性。不妨以固有时参数的 (44) 来进一步讨论。为了考察这一时空平移对称性所对应的守恒量，我们考虑如下无穷小变换

$$x^\mu \rightarrow \tilde{x}^\mu = x^\mu + \epsilon^\mu(\tau) \Leftrightarrow \delta x^\mu = \epsilon^\mu(\tau). \quad (48)$$

在此变换下，作用量的改变量为

$$\begin{aligned} \delta S &= -m \int \delta(d\tau) = -m \int \frac{1}{2} \frac{\delta(d\tau^2)}{d\tau} \\ &= m \int \frac{1}{2} \delta(\eta_{\mu\nu} dx^\mu dx^\nu) / d\tau \\ &= m \int \frac{dx^\mu}{d\tau} \eta_{\mu\nu} \delta(dx^\nu) \\ &= m \int d\tau \left( \frac{dx^\mu}{d\tau} \right) \dot{\epsilon}_\mu. \end{aligned} \quad (49)$$

这意味着与时空坐标平移对应的守恒量是  $m \frac{dx^\mu}{d\tau}$ ，记作  $p^\mu$ 。



利用固有时  $\tau$  和坐标时  $t$  之间的关系式 (45) 容易得出

$$p^0 = \frac{m}{\sqrt{1-v^2}}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-v^2}}. \quad (50)$$

另一方面，我们知道，时间坐标平移对应能量守恒，空间坐标平移对应动量守恒。因此  $p^0$  就是相对论粒子的能量， $\mathbf{p}$  就是相对论粒子的动量。这两者一起构成了一个四矢量  $p^\mu$ ，称作四动量。很容易看出，四动量满足如下关系

$$-p^\mu p_\mu = (p^0)^2 - \mathbf{p}^2 = m^2. \quad (51)$$



上一小节所讨论的只不过是一个自由的相对论粒子。而实际的相对论粒子会和诸如引力场或者电磁场这样的基本力场发生相互作用，如何写出一个包含了相互作用的作用量呢？这一小节我们以电荷为  $q$  的带电粒子与电磁场的相互作用为例来说明这一点。要点依然是通过考虑对称性，不过，这时候我们要考虑的是电磁场的规范对称性。



在电动力学中，我们会引入所谓的矢量势  $\mathbf{A}$ ，它和电势  $\phi$  一起构成了一个四矢量  $A^\mu$ ，它的定义是  $A^0 = \phi, A^i = A_i (i = 1, 2, 3)$ 。但是  $A^\mu$  本身不是物理的，真正物理上可测量的是电场强度  $\mathbf{E}$  和磁场强度  $\mathbf{B}$ ，这两者可以统一成一个两指标的所谓二阶反对称张量  $F_{\mu\nu}$ ，具体来说即是  $F_{i0} = E_i, F_{ij} = \epsilon_{ijk} B^k (i, j = 1, 2, 3)$ ，其中  $F_{ij}$  我们已经在上一章碰到过了。 $F_{\mu\nu}$  与  $A_\mu$  的关系是

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (52)$$

式中  $\partial_\mu = \frac{\partial}{\partial x^\mu}$ 。

# 对称性决定作用量

与电磁场的耦合



从  $F_{\mu\nu}$  的关系式 (52) 可以看出，它在如下变换下保持不变

$$A_\mu \rightarrow A_\mu + \partial_\mu \varepsilon. \quad (53)$$

式中  $\varepsilon(x)$  为时空坐标  $x^\mu$  的任意函数。这个变换就称之为电磁场的规范变换，从这个变换可以知道， $A_\mu$  本身不是物理的，因为不同的相差一个规范变换的  $A_\mu$  描述了同样的电磁场强度。

我们也可以用微分形式的语言，将四矢量  $A_\mu$  写成四维时空上的电磁场 1-形式  $A$ ,

$$A = A_\mu dx^\mu. \quad (54)$$

用这套微分形式的语言，规范变换其实就是如下变换

$$A \rightarrow A + d\varepsilon. \quad (55)$$

带电粒子和电磁场相互作用，它的作用量也得在这个规范变换下保持不变，这就是所谓的规范对称性。

# 对称性决定作用量

与电磁场的耦合



为了满足规范对称性，我们当然可以只用  $F_{\mu\nu}$  来和带电粒子相耦合，因为它是规范不变的。但我们同时还要求作用量具有洛伦兹不变性，也就是说作用量得是一个洛伦兹标量，因此不能有没被内积掉的指标存在。 $F_{\mu\nu}$  关于两个下标反对称，因此最简单的洛伦兹标量  $F_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}$  实际上等于零。用  $F_{\mu\nu}$  构造出来的最简单非零洛伦兹标量是  $F_{\mu\nu} F^{\mu\nu}$ ，很显然它是电磁场场强的二次方项，如果电磁场不是非常强，那它就是一个二阶小量，那这样的项实际上贡献就非常小，因此可以从带电粒子的作用量中忽略！

# 对称性决定作用量

与电磁场的耦合



仔细思考以后我们可能会发现下面这一项

$$+q \int_a^b A. \quad (56)$$

这一项规范不变吗？很明显不是，因为在规范变换 (55) 的作用下，它会变为

$$q \int_a^b A \rightarrow q \int_a^b A + q \int_a^b d\varepsilon \quad (57)$$

但是，我们发现多出来的部分是一个全微分，它完全取决于路径的两个端点，即

$$q \int_a^b d\varepsilon = q(\varepsilon(b) - \varepsilon(a)). \quad (58)$$

而我们早就知道，在使用最小作用量原理时，路径的两个端点是固定不变的，因此规范变换下多出来的这一项对变分完全没有贡献。即是说，虽然给作用量加上的这一项 (56) 不是规范不变的，但是，由它导出来的运动微分方程却是规范不变的！





因此，我们可以写出相对论性带电粒子完整的作用量，为

$$\begin{aligned} S[x(s)] &= -m \int d\tau + q \int A \\ &= -m \int ds \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}} + q \int ds A_\mu \frac{dx^\mu}{ds}. \end{aligned} \quad (59)$$

如果将参数  $s$  取成坐标时  $t$ ，并考虑  $v \ll 1$  的非相对论极限，那上式就可以近似成

$$S[\mathbf{x}(t)] = -m \int dt + \int dt \left[ \frac{1}{2} m \mathbf{v}^2 - q\phi + q\mathbf{A} \cdot \mathbf{v} \right] + \dots \quad (60)$$

显然，除了前面那个无关紧要的常数项外，这个作用量正是我们上一章纯粹碰运气猜出来的带电粒子的作用量。但是现在，在一定的意义上我们是通过对称性的考虑推导出了它！



**Thank You!**