目录

第三章　对称性

陈童

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，飞船之内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。

关于物理定律的对称性的这一定义与几何对称性是一致的。比如，当我们说圆有很大的对称性时，我们指的是，圆的几何图形在绕圆心旋转任意角的操作下保持不变。当我们说人体外观左右对称时(并不严格)，我们指的是将人体沿着中线左右对调，人体外观保持不变。

3.1　对称性与守恒定律

3.1.1　对称性就是作用量的不变性

前面的章节中我们讲了经典物理学的基本原理，最小作用量原理，上面又定义了何为物理定律的对称性，下面我们将这两者结合起来。结合对称性与最小作用量原理的关键在于，根据最小作用量原理，基本力学定律由作用量的变分等于零给出。当然，这个作用量可以是相空间作用量，也可以是坐标空间作用量，不过，相对来说用坐标空间的作用量来分析对称性更加方便，因此本章将集中于此。而在相空间中分析对称性需要的准备知识就稍微多一些，我们将推迟到后面的章节中进行。

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$\widetilde{q}(t) = F(q(t))$。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$\delta S[\widetilde{q}(t)] = 0$。即，对于对称操作作用前后的两条路径$q(t)$和$\widetilde{q}(t) = F(q(t))$, 有

$$\delta S[q(t)] = 0 \Leftrightarrow \delta S[\widetilde{q}(t)] = 0. \tag{3.1}$$

很显然，满足这一要求的最直接方式是

$$S[\widetilde{q}(t)] = S[q(t)]. \tag{3.2}$$

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**。

3.1.2　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。我不知道这个证明应该归功于谁，我自己最早是从温伯格的量子场论书中学的，但我也在好几个大师的书或者课程中或多或少见过这个证明的关键想法。

值得说明的是，下面的证明也有一定的局限性，即它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，只是如果想一般性地写出来那就要复杂一些。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$G(\theta)$，θ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, ϵ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$\widetilde{q}$

$$q(t) \to \widetilde{q}(t) = q(t) + \epsilon F(q(t)), \tag{3.3}$$

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

$$\delta S = S[\widetilde{q}(t)] - S[q(t)] = \int dt L(\widetilde{q}, \dot{\widetilde{q}}) - \int dt L(q, \dot{q}) = 0. \tag{3.4}$$

下面是关键的证明技巧。前面的ϵ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

$$\epsilon(t_i) = \epsilon(t_f) = 0. \tag{3.5}$$

我们考虑修改的变换

$$\begin{aligned} q(t) \to \widetilde{q}(t) &= q(t) + \epsilon(t) F(q(t)), \\ \text{即} \quad \delta q &= \widetilde{q}(t) - q(t) = \epsilon(t) F(q(t)). \end{aligned} \tag{3.6}$$

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(3.6)的变换下，作用量的改变量必然具有如下形式

$$\delta S = \int_{t_i}^{t_f} dt \big(L(\widetilde{q}, \dot{\widetilde{q}}) - L(q, \dot{q}) \big) = \int_{t_i}^{t_f} dt Q(q, \dot{q}) \dot{\epsilon}(t). \tag{3.7}$$

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(3.7)。

以上都是数学技巧，下面是整个证明的物理部分。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$\delta S = 0$。(3.6)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**$q(t)$，(3.7)式得出的作用量改变量必定恒为零，即

$$\delta S = \int_{t_i}^{t_f} dt Q(q, \dot{q}) \dot{\epsilon}(t) = 0. \tag{3.8}$$

将上式分部积分，就可以得到

$$\delta S = -\int_{t_i}^{t_f} dt \dot{Q}(q, \dot{q}) \epsilon(t) = 0. \tag{3.9}$$

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

$$\frac{dQ}{dt} = 0. \tag{3.10}$$

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！

值得强调的是，以上诺特定理的证明过程也给我们指出了一条如何找到守恒量$Q(q, \dot{q})$的道路，我们下面的物理讨论都基于这一道路。

3.1.3　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。

空间平移对称性与动量守恒

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$L(\mathbf{x}_i, \dot{\mathbf{x}}_i)$($L(\mathbf{x}_1, ..., \mathbf{x}_N, \dot{\mathbf{x}}_1, ..., \dot{\mathbf{x}}_N)$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

$$\mathbf{x}_i \to \widetilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{a}. \tag{3.11}$$

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

$$0 = \delta L = L(\mathbf{x}_i + \epsilon, \dot{\mathbf{x}}_i) - L(\mathbf{x}_i, \dot{\mathbf{x}}_i) = \epsilon \cdot \sum_j \frac{\partial L}{\partial \mathbf{x}_j}. \tag{3.12}$$

下面使用关键技巧，即将无穷小参数ϵ变成$\epsilon(t)$，即考察

$$\mathbf{x}_i \to \widetilde{\mathbf{x}}_i = \mathbf{x}_i + \epsilon(t). \Leftrightarrow \delta \mathbf{x}_i = \epsilon(t). \tag{3.13}$$

从而容易计算出作用量的改变量，有

$$\begin{aligned}
\delta S &= \int dt \delta L = \int dt \big[\sum_j \frac{\partial L}{\partial \mathbf{x}_j} \cdot \delta \mathbf{x}_j + \sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \delta \dot{\mathbf{x}}_j \big] \\
&= \int dt \big[\sum_j \frac{\partial L}{\partial \mathbf{x}_j} \cdot \epsilon + \sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \cdot \dot{\epsilon} \big]
\end{aligned} \tag{3.14}$$

利用平移不变性的(3.12)式，即有

$$\delta S = \int dt \big(\sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \big) \cdot \dot{\epsilon}. \tag{3.15}$$

根据上一节关于诺特定理的证明容易知道，上式说明$\big(\sum_j \frac{\partial L}{\partial \dot{\mathbf{x}}_j} \big)$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$\mathbf{p}$

$$\mathbf{p} = \sum_j \mathbf{p}_j. \tag{3.16}$$

所以空间平移不变性会导致系统的总动量守恒！

空间旋转不变性与角动量守恒

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$\mathbf{x}=(x^1,x^2,...,x^d)$, 我们将各分量记为$x^\alpha,\alpha=1,2,...,d$。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$x_\alpha=x^\alpha$。

首先需要研究一下如何处理d维空间中的空间旋转，核心是如何处理无穷小旋转(因为诺特定理的关键在于无穷小旋转)。记无穷小旋转操作之下，坐标的改变量为δx^α，很显然，δx^α应该与原来的矢量$\{x^\alpha\}$成线性关系(因为一个矢量在无穷小旋转之下改变多少显然是正比于原来的矢量的)，我们设

$$\delta x_\alpha=\epsilon_{\alpha\beta}x^\beta. \tag{3.17}$$

这里使用了求和约定，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$\delta(\mathbf{x}^2)=0$，即有

$$0=\delta(\mathbf{x}^2)/2=\mathbf{x}\cdot\delta\mathbf{x}=x^\alpha\delta x_\alpha=\epsilon_{\alpha\beta}x^\alpha x^\beta. \tag{3.18}$$

注意到$x^\alpha x^\beta$关于指标α,β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[1]，即

$$\epsilon_{\alpha\beta}=-\epsilon_{\beta\alpha}. \tag{3.19}$$

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!

因此，在无穷小旋转操作之下，我们有

$$\delta x_\alpha=\epsilon_{\alpha\beta}x^\beta,\quad \delta\dot{x}_\alpha=\epsilon_{\alpha\beta}\dot{x}^\beta. \tag{3.20}$$

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)

$$\delta\dot{x}^\alpha=\epsilon^{\alpha\beta}\dot{x}_\beta+\dot{\epsilon}^{\alpha\beta}x_\beta. \tag{3.21}$$

[1]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。

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

$$
\begin{aligned}
\delta L &= \frac{\partial L}{\partial \dot{x}^{\beta}}\dot{\epsilon}^{\beta\alpha}x_{\alpha} = p_{\beta}x_{\alpha}\dot{\epsilon}^{\beta\alpha} \\
&= -\frac{1}{2}\big(x_{\alpha}p_{\beta} - x_{\beta}p_{\alpha}\big)\dot{\epsilon}^{\alpha\beta}.
\end{aligned}
\tag{3.22}
$$

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

$$
\delta S = -\frac{1}{2}\int dt\big(x_{\alpha}p_{\beta} - x_{\beta}p_{\alpha}\big)\dot{\epsilon}^{\alpha\beta} \tag{3.23}
$$

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$J_{\alpha\beta}$,

$$
J_{\alpha\beta} = x_{\alpha}p_{\beta} - x_{\beta}p_{\alpha}. \tag{3.24}
$$

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

$$
J_{12} = xp_y - yp_x, \quad J_{23} = yp_z - zp_y, \quad J_{31} = zp_x - xp_z. \tag{3.25}
$$

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ϵ^{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

$$
J^i = \frac{1}{2}\epsilon^{ijk}J_{jk}. \tag{3.26}
$$

不过，很显然，高维空间的角动量只能记为$J_{\alpha\beta}$。

时间平移对称性与能量守恒

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$t \to \widetilde{t} = t + a$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$q(t) \to \widetilde{q}(t) = q(\widetilde{t})$，从而

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

式中最后一个等号只是将$\widetilde{t}$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。对于开放系统，通常拉格朗日量是要显含t的，即要写成$L(q, \dot{q}, t)$，这时候当然就没有时间平移不变性了。

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

$$
t \to \widetilde{t}(t) = t + \epsilon(t), \tag{3.27}
$$

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

$$
\begin{aligned}
S[q(\widetilde{t})] &= \int dt L(q(\widetilde{t}), \frac{d}{dt} q(\widetilde{t})) = \int d\widetilde{t} (\frac{dt}{d\widetilde{t}}) L\Big(q(\widetilde{t}), (\frac{d\widetilde{t}}{dt}) \dot{q}(\widetilde{t})\Big) \\
&= \int d\widetilde{t} (\frac{1}{1 + \dot{\epsilon}}) L\Big(q(\widetilde{t}), (1 + \dot{\epsilon}) \dot{q}(\widetilde{t})\Big) \\
&= \int d\widetilde{t} L(q(\widetilde{t}), \dot{q}(\widetilde{t}))) - \int d\widetilde{t} (\dot{\epsilon}) L\Big(q(\widetilde{t}), \dot{q}(\widetilde{t})\Big) + \int d\widetilde{t} \frac{\partial L}{\partial\big(\dot{q}(\widetilde{t})\big)} \dot{q}(\widetilde{t}) \big(\dot{\epsilon}\big) \\
&= \int dt L(q(t), \dot{q}(t))) + \int dt \Big[\frac{\partial L}{\partial\big(\dot{q}(t)\big)} \dot{q}(t) - L(q(t), \dot{q}(t))\Big] \dot{\epsilon},
\end{aligned} \tag{3.28}
$$

式中最后一行是将变量$\widetilde{t}$重命名成了t。从上面计算可以得到，变换前后作用量的改变量为

$$
\delta S = S[q(\widetilde{t})] - S[q(t)] = \int dt \Big[\frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q})\Big] \dot{\epsilon}. \tag{3.29}
$$

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

$$
E(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}). \tag{3.30}
$$

上面的推导过程忽略了广义坐标的指标，如果恢复指标，那守恒量$E(q,\dot{q})$的表达式就将是

$$E(q,\dot{q})=\frac{\partial L}{\partial \dot{q}^a}\dot{q}^a-L(q,\dot{q}) \tag{3.31}$$

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！

上面这个将时间平移对称性和能量守恒定律联系起来的推导过程正好完善了我们上一小节对诺特定理的证明。因为这里涉及到的对称性正好是需要将时间t变换为$\widetilde{t}$的那种对称性，而上面的推导正好可以说明在这种情形下如何处理关于诺特定理的证明。

以上关于能量守恒定律的证明看起来有些不必要的复杂。但是，这一证明的好处是，它最具有诺特定理的精神，这同样也是我们关于动量守恒以及角动量守恒的讨论过程的优点，正因为如此，以上关于这些守恒定律的讨论过程完全可以平行地照搬到场论中。而且，由于现在的物理系统比场要简单一些，所以这里的讨论可以写得更加清晰，从而更能凸显诺特定理相关推理和推导的本质，我想，读者日后学习量子场论时回头再来看这里的推导一定更能领悟这一推导的巧妙。

举例

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$S[y(x)]=\int dxL(y,y')$, 式中

$$L(y,y')=n(y)\sqrt{1+(y')^2}. \tag{3.32}$$

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$x\to\widetilde{x}=x+a$的变换下不变。因此这个例子也有一个“守恒能量”，为

$$E(y,y')=\frac{\partial L}{\partial y'}y'-L(y,y')=-\frac{n(y)}{\sqrt{1+(y')^2}}. \tag{3.33}$$

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。

3.1.4 宇宙的总能量是零

既然已经讲到费马原理的例子，不妨让我们看一下最一般的光程泛函$S[\mathbf{x}(s)]$有什么对称性，又会导致什么守恒定律。这里,

$$S[\mathbf{x}(s)] = \int n(\mathbf{x})|\frac{d\mathbf{x}}{ds}|ds. \tag{3.34}$$

很明显，这里有参数s的平移不变性。即在$s \to \widetilde{s} = s + a$的变换下，$S[\mathbf{x}(s)]$保持不变。因此按照前面关于能量守恒的处理可以知道，必定存在一个守恒的“能量”$E(\mathbf{x}(s), \dot{\mathbf{x}}(s))$。

下面要做的就是找到$E(\mathbf{x}(s), \dot{\mathbf{x}}(s))$的表达式。为此，考察如下无穷小变换

$$s \to \widetilde{s}(s) = s + \epsilon(s), \tag{3.35}$$

这里$\widetilde{s}$是原参数s的函数。在此变换之下，$S[\mathbf{x}(s)]$将变换为

$$\begin{aligned} S[\mathbf{x}(s)] \to S[\mathbf{x}(\widetilde{s})] &= \int n\big(\mathbf{x}(\widetilde{s})\big)|\frac{d}{ds}\mathbf{x}(\widetilde{s})|ds \\ &= \int n\big(\mathbf{x}(\widetilde{s})\big)|\frac{d}{d\widetilde{s}}\mathbf{x}(\widetilde{s})|d\widetilde{s} \\ &= \int n\big(\mathbf{x}(s)\big)|\frac{d}{ds}\mathbf{x}(s)|ds = S[\mathbf{x}(s)]. \end{aligned} \tag{3.36}$$

式中最后一行是将变量$\widetilde{s}$重命名为s。我们发现，即使在这种修改以后的无穷小变换下，$S[\mathbf{x}(s)]$仍然不变！事实上，上面的推导过程和$\widetilde{s}(s)$的具体形式没有多大关系，也即是说，其实$S[\mathbf{x}(s)]$在相当任意的$s \to \widetilde{s}(s)$这种重新参数化之下都会保持不变，$S[\mathbf{x}(s)]$具有关于s的重参数不变性！它意味着，在与“时间平移”有关的(3.35)变换下，$S[\mathbf{x}(s)]$的改变量将是零，我们写作

$$0 = \delta S = S[\mathbf{x}(\widetilde{s})] - S[\mathbf{x}(s)] = \int ds\Big(0\Big) \cdot \dot{\epsilon}(s). \tag{3.37}$$

因此，这也即是说，“守恒能量”$E(\mathbf{x}(s), \dot{\mathbf{x}}(s))$其实为零，即

$$E(\mathbf{x}(s), \dot{\mathbf{x}}(s)) = 0. \tag{3.38}$$

从这个例子我们学到的就是，如果一个物理系统的作用量不仅在时间平移下保持不变，而且实际上在$t \to \widetilde{t}(t)$的任意时间重参数化之下都保持不

变。那这个物理系统的总能量将不仅守恒，而且实际上这个总能量必定等于零！宇宙就是一个这样的系统，对宇宙作用量有贡献的包括引力场和物质场，是广义相对论的一个作用量，它具有所谓的微分同胚不变性，特别的，它具有$t \to \widetilde{t}(t)$的时间微分同胚不变性。因此，根据上面的推理，宇宙的总能量必定是零。这是因为，引力势能是负的，而物质场的能量是正的，正负抵消刚好得到零！

但是，上面的论证有一个漏洞，那就是宇宙毕竟是一个场论系统，不是一个点粒子系统，而场是一种空间分布。这使得当我们对场论系统进行类似上面那样的推导时，系统的渐近空间边界需要额外的特殊对待，要加上边界条件，它使得即使体内的时间有重参数不变性，这种重参数不变性也不能延展到渐近边界上去，这样一来渐近空间边界上就可能对总能量产生贡献。所以，如果宇宙在空间上有渐近边界，那它的总能量就不一定是零。不过，现在人们更倾向于认为宇宙在空间上并没有渐近边界，所以，总能量还是零，至少我们倾向于这么认为。

3.1.5 *伽利略不变性

在本章的一开始我们设想了一艘在太空中孤独地飞行的宇宙飞船，我们说了，飞船中的人不仅不能辨测空间坐标和方向，他也不能辨测自己是否在飞行。也即是说，可以给飞船一个任意常数速度$\mathbf{v}$，使得$\mathbf{x} \to \mathbf{x}+\mathbf{v}t$，而飞船中的人不会辨测到任何异常。

$$\mathbf{x} \to \mathbf{x}+\mathbf{v}t, \tag{3.39}$$

就是所谓的伽利略变换，它将一个惯性系变换到另外一个相对它以速度$\mathbf{v}$匀速飞行的惯性系。伽利略告诉我们，这种变换不会带来力学规律的变化，一切惯性系都是平权的。这种力学规律的伽利略变换不变性就是所谓的相对性原理。那么，伽利略不变性能给出什么守恒定律呢？

为了回答这个问题，我们首先构造一个具体的多粒子作用量，如下

$$S[\mathbf{x}(t)] = \int dt \Big[\frac{1}{2}\sum_i m_i \dot{\mathbf{x}}_i^2 - \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j)\Big]. \tag{3.40}$$

但是，这个作用量在$\mathbf{x}_i \to \widetilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{v}t$的伽利略变换下并非不变，而是会

变化为

$$
\begin{aligned}
S[\mathbf{x}(t)] \to S[\widetilde{\mathbf{x}}(t)] &= \int_{t_i}^{t_f} dt\Big[\frac{1}{2}\sum_i m_i(\dot{\mathbf{x}}_i+\mathbf{v})^2 - \sum_{i<j} V(\mathbf{x}_i-\mathbf{x}_j)\Big] \\
&= S[\mathbf{x}(t)] + \int_{t_i}^{t_f} dt\big[\sum_i m_i\dot{\mathbf{x}}_i\cdot\mathbf{v}\big] + \frac{1}{2}m\mathbf{v}^2(t_f-t_i) \\
&= S[\mathbf{x}(t)] + \sum_i m_i\big[\mathbf{x}_i(t_f)-\mathbf{x}_i(t_i)\big]\cdot\mathbf{v} + \frac{1}{2}m\mathbf{v}^2(t_f-t_i).
\end{aligned}
$$

式中$m = \sum_i m_i$为系统的总质量。我们发现，作用量在伽利略变换下多出来了两项。但是，多出来的这两项仅仅影响路径的起末两个端点，而在应用最小作用量原理时这两个端点却是固定的，因此多出来的这两项不会对变分产生任何影响。也即是说，虽然作用量并非伽利略不变，但是$\delta S = 0$的运动微分方程却是伽利略不变的。所以，这里虽然需要略微推广一下我们关于对称性的定义，但伽利略变换的确是我们所构造的这个力学系统的对称性。

下面我们考察无穷小伽利略变换，即设变换速度$\mathbf{v} = \epsilon$。和前面一样，我们令ϵ为依赖于时间t的任意无穷小函数，记为$\epsilon(t)$。即是说，我们考虑如下变换

$$
\mathbf{x}_i \to \widetilde{\mathbf{x}}_i = \mathbf{x}_i + \epsilon(t)t. \tag{3.41}
$$

在这个变换的作用下，作用量$S[\mathbf{x}(t)]$变换为

$$
\begin{aligned}
S[\mathbf{x}(t)] &\to S[\widetilde{\mathbf{x}}(t)] = \int_{t_i}^{t_f} dt\Big[\frac{1}{2}\sum_i m_i(\dot{\mathbf{x}}_i+\epsilon+\dot{\epsilon}t)^2 - \sum_{i<j} V(\mathbf{x}_i-\mathbf{x}_j)\Big] \\
&= S[\mathbf{x}(t)] + \int_{t_i}^{t_f} dt\big[\sum_i m_i\dot{\mathbf{x}}_i\big]\cdot\epsilon(t) + \int_{t_i}^{t_f} dt\big[\sum_i m_i\dot{\mathbf{x}}_i\big]t\cdot\dot{\epsilon}
\end{aligned}
$$

对第二行中间那项分部积分，就可以得到

$$
\begin{aligned}
\delta S &= S[\widetilde{\mathbf{x}}(t)] - S[\mathbf{x}(t)] \\
&= -\int_{t_i}^{t_f} dt\big[\sum_i m_i\mathbf{x}_i\big]\cdot\dot{\epsilon} + \int_{t_i}^{t_f} dt\big[\sum_i m_i\dot{\mathbf{x}}_i\big]t\cdot\dot{\epsilon} \\
&= \int_{t_i}^{t_f} dt\big[\big(\sum_i m_i\dot{\mathbf{x}}_i\big)t - \sum_i m_i\mathbf{x}_i\big]\cdot\dot{\epsilon}.
\end{aligned} \tag{3.42}
$$

上面这个式子意味着，$\left(\sum_i m_i\dot{\mathbf{x}}_i\right)t-\sum_i m_i\mathbf{x}_i$是守恒量，不妨记为$\mathbf{K}$。很显然$\sum_i m_i\dot{\mathbf{x}}_i$就是系统的总动量$\mathbf{p}$, $\mathbf{p}=\sum_i m_i\dot{\mathbf{x}}_i$。由于伽利略不变的系统一定具有空间平移不变性，所以这个总动量$\mathbf{p}$本身就是守恒量。另外，我们可以引入质心坐标$\mathbf{x}_c$，从而有$\sum_i m_i\mathbf{x}_i=m\mathbf{x}_c$。因此守恒量$\mathbf{K}$其实就是

$$\mathbf{K}=\mathbf{p}t-m\mathbf{x}_c. \tag{3.43}$$

不过，$\mathbf{K}$守恒其实是显然的，因为$\frac{d\mathbf{K}}{dt}=\mathbf{p}-m\frac{d\mathbf{x}_c}{dt}=0$显然成立，我们上面推导的关键不过是指出这个守恒量和伽利略对称性密切相关。

3.2　对称性决定作用量

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。其实，牛顿运动定律也是猜的，麦克斯威方程组也是猜的，从根本上来说，要得到一个物理系统的运动微分方程，我们都是靠猜。而面对一个全新的物理系统，猜它的作用量可比猜运动微分方程容易多了！我想，上一章我们关于洛伦兹力的讨论足以说明这一点。

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。

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。

3.2.1　相对论不变性

本小节讨论相对论不变性，这里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。

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。

现在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，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。这是因为，麦克斯韦方程作为一组特定的物理学定律必须在这两个参考系中取相同的形式，而麦克斯韦方程是一组关于时间和空间坐标的一阶偏微分方程，它如果要保持形式不变的话，除非两组时空坐标之间的变换是线性变换才可能。

另外，根据光速不变原理我们可以知道，在从S系到S'系的变换中，t'不仅依赖于原来的时间坐标t，而且还一定同时依赖于原来的空间坐标$\mathbf{x}$(与牛顿的绝对时间观念是完全不同的)。这是因为，根据光速不变原理，在S系中不同地点同时发生的两个事件在S'看来是不同时的，这就是所谓的同时的相对性。

间隔不变性

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

$$-dt^2 + dx^2 + dy^2 + dz^2 = 0. \tag{3.44}$$

根据光速不变原理，同样的两个事件在S'系看来也得满足

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = 0. \tag{3.45}$$

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(3.44)成立时必有(3.45)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

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = D(\mathbf{v})(-dt^2 + dx^2 + dy^2 + dz^2), \qquad (3.46)$$

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

$$-dt^2 + dx^2 + dy^2 + dz^2 = D(-\mathbf{v})(-dt'^2 + dx'^2 + dy'^2 + dz'^2). \qquad (3.47)$$

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

$$-dt'^2 + dx'^2 + dy'^2 + dz'^2 = -dt^2 + dx^2 + dy^2 + dz^2. \qquad (3.48)$$

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

$$-d\tau^2 = -dt^2 + dx^2 + dy^2 + dz^2 = -dt^2 + d\mathbf{x}^2. \qquad (3.49)$$

用这个记号，方程(3.48)就可以简记成

$$d\tau'^2 = d\tau^2, \qquad (3.50)$$

称为两个事件的间隔不变性。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[2]。

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

$$-d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \qquad (3.51)$$

[2]从这里的讨论读者可以看到，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，这是两个不同的概念，人们常常用不同的记号表示它们。

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$\eta_{\mu\nu} = \eta_{\nu\mu}$。

dx^μ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^μ。但是，我们也可以利用$\eta_{\mu\nu}$将A^μ的指标降下来，即约定

$$A_\mu = \eta_{\mu\nu} A^\nu. \tag{3.52}$$

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

$$A^\mu = \eta^{\mu\nu} A_\nu. \tag{3.53}$$

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

$$A \cdot B = A^\mu B_\mu = A_\mu B^\mu = \eta_{\mu\nu} A^\mu B^\nu. \tag{3.54}$$

两个四矢量的内积是参考系变换不变的，称为洛伦兹标量。

洛伦兹变换

狭义相对论中，两个参考系之间的时空坐标变换称作洛伦兹变换，根据前面所说，它是一个保持间隔不变性的线性变换，通常写成

$$x'^\mu = \Lambda^\mu_\nu x^\nu, \tag{3.55}$$

式中Λ^μ_ν构成一个常数矩阵。

下面我们来构造一个具体的洛伦兹变换。为此，让我们考虑一种特殊情况，假如S'系沿着S系的x轴正方向以匀速v运动。假设我们这么选取S'的坐标轴，以使得初始时两个坐标系的坐标轴完全重合。这样一来，由于两坐标系的相对运动只发生在x方向上，与y, z方向无关，所以显然有$y' = y, z' = z$。因此，我们只需要考虑事件的t坐标和x坐标在参考系变换下如何变换，为此不妨暂时忽略间隔公式(3.49)中的y, z坐标，进而将间隔公式简化为

$$-d\tau^2 = -dt^2 + dx^2 = dx^+ dx^-, \tag{3.56}$$

式中$x^+ = x + t, x^- = x - t$。间隔不变性告诉我们，参考系变换以后的$x^{+\prime} = x' + t', x^{-\prime} = x' - t'$必然满足$dx^{+\prime} dx^{-\prime} = dx^+ dx^-$。显然这就意味着必定有某个参数$\omega$使得$x^{+\prime} = e^{-\omega} x^+$, $x^{-\prime} = e^{\omega} x^-$, 即有

$$x' = \cosh(\omega)x - \sinh(\omega)t, \tag{3.57}$$

$$t' = -\sinh(\omega)x + \cosh(\omega)t. \tag{3.58}$$

考虑S'系的坐标原点$x' = 0$，它在S中的运动速度为v, 因此我们有$\tanh(\omega) = v$，利用双曲函数的相关公式，容易得到$\cosh(\omega) = 1/\sqrt{1-v^2}, \sinh(\omega) = v/\sqrt{1-v^2}$, 也即是说，$x, t$坐标的变换公式为

$$t' = \frac{t - vx}{\sqrt{1-v^2}}, \tag{3.59}$$

$$x' = \frac{x - vt}{\sqrt{1-v^2}}, \tag{3.60}$$

再加上前面的$y' = y, z' = z$，这四个式子就是最常用的一组洛伦兹变换。

3.2.2 相对论粒子的作用量

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

$$S[x(s)] = -m\int d\tau = -m\int ds\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}}. \tag{3.61}$$

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，请读者在待会儿我们讨论非相对论极限时留心。

值得指出的是，上述作用量(3.61)显然具有重参数不变性，即在$s \to \tilde{s}(s)$的参数变换下保持不变。这意味着我们可以在一定意义上任意选择参数s。最常见的选择有两种，第一种是，取$s = \tau$，即取固有时本身为路径的参数，从而作用量(3.61)可以写成

$$S[x(\tau)] = -m\int d\tau = -m\int d\tau\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}}. \tag{3.62}$$

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

$$d\tau^2 = dt^2 - d\mathbf{x}^2 = dt^2(1-\mathbf{v}^2). \tag{3.63}$$

式中$\mathbf{v} = \frac{d\mathbf{x}}{dt}$是粒子的速度。从而就可以将作用量(3.61)写成

$$S[\mathbf{x}(t)] = -m\int dt\sqrt{1-\mathbf{v}^2}. \tag{3.64}$$

假设粒子的运动速度远低于光速，即$\mathbf{v} \ll 1$，那这时候就可以利用关于$\mathbf{v}^2$的泰勒展开将相对论的作用量(3.64)近似成

$$S = -m\int dt + \int dt\frac{1}{2}m\mathbf{v}^2 + ... \tag{3.65}$$

省略号表示$\mathbf{v}^2$的高阶项。很显然，除了相差一个对变分没有影响的常数项$-m\int dt$之外，这个近似作用量正是非相对论的自由粒子的作用量!

另外，上述相对论粒子作用量显然具有$x^\mu \to \widetilde{x}^\mu = x^\mu + a^\mu$($a^\mu$为常矢量)的时空坐标平移不变性。不妨以固有时参数的(3.62)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

$$x^\mu \to \widetilde{x}^\mu = x^\mu + \epsilon^\mu(\tau) \Leftrightarrow \delta x^\mu = \epsilon^\mu(\tau). \tag{3.66}$$

在此变换下，作用量的改变量为

$$\begin{aligned}
\delta S &= -m\int \delta(d\tau) = -m\int \frac{1}{2}\frac{\delta(d\tau^2)}{d\tau} \\
&= m\int \frac{1}{2}\delta(\eta_{\mu\nu}dx^\mu dx^\nu)/d\tau \\
&= m\int \frac{dx^\mu}{d\tau}\eta_{\mu\nu}\delta(dx^\nu) \\
&= m\int d\tau\big(\frac{dx^\mu}{d\tau}\big)\dot{\epsilon}_\mu.
\end{aligned} \tag{3.67}$$

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^μ,

$$p^\mu = m\frac{dx^\mu}{d\tau}. \tag{3.68}$$

利用固有时τ和坐标时t之间的关系式(3.63)容易得出

$$p^0 = \frac{m}{\sqrt{1-\mathbf{v}^2}}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2}}. \tag{3.69}$$

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^μ，称作四动量。很容易看出，四动量满足如下关系

$$-p^\mu p_\mu = (p^0)^2 - \mathbf{p}^2 = m^2. \tag{3.70}$$

如果允许引入一个辅助性的力学变量$e(s)$，那我们还可以将作用量(3.61)写成一个更加顺眼的形式，

$$S[e(s), x(s)] = \frac{1}{2}\int ds\Big(e^{-1}\eta_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} - em^2\Big). \tag{3.71}$$

为了证明这个作用量与前面的(3.61)相等价，人们只要首先利用下式消去辅助力学变量$e(s)$即可

$$\frac{\delta S}{\delta e(s)} = 0. \tag{3.72}$$

这个过程就和我们上一章从相空间最小作用量原理推导坐标空间最小作用量原理时所做的类似。

3.2.3 与电磁场的耦合

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。

正如上一章涉及过的，在电动力学中，我们会引入所谓的矢量势$\mathbf{A}$，它和电势ϕ一起构成了一个四矢量A^μ，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

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{3.73}$$

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

从$F_{\mu\nu}$的关系式(3.73)可以看出，它在如下变换下保持不变

$$A_\mu \to A_\mu + \partial_\mu \varepsilon. \tag{3.74}$$

式中$\varepsilon(x)$为时空坐标x^μ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_μ本身不是物理的，因为不同的相差一个规范变换的A_μ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_μ写成四维时空上的电磁场1-形式A,

$$A = A_\mu dx^\mu. \tag{3.75}$$

用这套微分形式的语言，规范变换其实就是如下变换

$$A \to A + d\varepsilon. \tag{3.76}$$

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。

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，如何电磁场不是非常强，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！

看起来好像没什么不可忽略的既规范不变又洛伦兹不变的项可以加到相对论粒子的作用量中去！这当然有问题，因为带电粒子显然和电磁场有不可忽略的相互作用。仔细思考以后我们可能会发现下面这一项

$$+q\int_a^b A. \tag{3.77}$$

它就是电磁场1-形式沿着带电粒子时空运动路径的积分, 式中我们假设这条路径起于a点终止于b点。这一项显然是洛伦兹不变的，那它规范不变吗?很明显不是，因为在规范变换(3.76)的作用下，它会变为

$$q\int_a^b A \to q\int_a^b A + q\int_a^b d\varepsilon \tag{3.78}$$

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

$$q\int_a^b d\varepsilon = q\big(\varepsilon(b)-\varepsilon(a)\big). \tag{3.79}$$

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(3.77)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！所以，(3.77)这一项实际上符合要求！

因此，我们可以写出相对论性带电粒子完整的作用量，为

$$\begin{aligned} S[x(s)] &= -m\int d\tau + q\int A \\ &= -m\int ds\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}} + q\int dsA_\mu\frac{dx^\mu}{ds}. \end{aligned} \tag{3.80}$$

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

$$S[\mathbf{x}(t)] = -m\int dt + \int dt\Big[\frac{1}{2}m\mathbf{v}^2 - q\phi + q\mathbf{A}\cdot\mathbf{v}\Big] + ... \tag{3.81}$$

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！

3.2.4　对称性决定一切？

相对论不变性加上规范不变性不仅可以决定带电粒子的作用量，它也可以决定电磁场本身的作用量，从而也就决定了电磁场的麦克斯威方程。电磁场的规范对称性后来被杨振宁和米尔斯所推广，变成所谓的非阿贝尔规范对称性。最后人们发现，某种非阿贝尔规范对称性决定了强相互作用，而另一种非阿贝尔规范对称性决定了弱相互作用。不仅如此，爱因斯坦的另外一个伟大发现就是，万有引力定律也遵循某种奇妙的“对称性”，这种“对称性”既决定了处于引力场中的粒子的作用量，也决定了引力场本身的作用量，从而也就决定了引力场本身的爱因斯坦方程。

到此为止，我们可以说，自然界中的四种基本相互作用，电磁相互作用、弱相互作用、强相互作用、以及万有引力相互作用，它们的作用量都

是由对称性所决定的。对称性是否决定了一切呢？这个问题我们无法回答，不过我们可以回顾一下，决定了万有引力规律的是何种对称性，并谈谈今天人们对自然界的对称性又有什么新的猜测。

万有引力和惯性力之间的对称性

爱因斯坦曾经向别人讲述过他所谓的“这辈子最幸福的想法”，他说(大意)：有一天他坐在专利局里，思考如何推广他的狭义相对论，使之能自洽地包含万有引力[3]。突然，他想到一个人在电梯里随着电梯自由下落的情形。爱因斯坦仿佛可以切身地体会到这个人感受到的失重，他意识到，在这个自由下落的人看来，地球重力不存在了，因为它被加速运动的惯性力抵消了！爱因斯坦抓住了这个念头的关键点，即是说，惯性力可以抵消万有引力，即是说，惯性力和万有引力其实是不可区分的！

这就是万有引力所遵循的奇妙对称性。根据这种对称性，对于一个在外太空加速运动的人而言，他会感受到一个加速运动的惯性力，但是，假设他不能向外看，不知道飞船在加速，他完全可以认为自己是在一个引力场中。所以伽利略发现，只要不向外看，飞船的速度就是不可辨测的，而爱因斯坦进一步发现，飞船的加速度也是不可辨测的，因为它等效于一个引力场！

当然，这里有一个微妙的关键，即惯性力和万有引力之间的这种等价性只对局限在一个局部小区域(比如一部电梯)之内的观察者才成立。即是说，万有引力的这种奇妙对称性是一种局域对称性(和电磁场的规范对称性有点类似)。爱因斯坦称他发现的这种万有引力的局域对称性为等效原理，意即，局域而言，惯性力和万有引力等效。

将等效原理所描述的这种对称性和狭义相对论不变性结合起来，就能决定万有引力场的作用量，也就是所谓的希尔伯特-爱因斯坦作用量。对这个作用量应用最小作用量原理，就能导出引力场的爱因斯坦方程。而这一整个理论就是著名的广义相对论。

---

[3]那时候人们只知道电磁力和万有引力，强相互作用力和弱相互作用力还没发现。电磁力天然与狭义相对论自洽，但是牛顿的万有引力理论是一种超距作用力，即是说牛顿认为万有引力的传播不需要时间，这和狭义相对论所说的相互作用的传播速度不能超过光速矛盾了！

玻色子与费米子的对称性？

我们已经看到了对称性的神奇力量。人们当然想继续延伸这种力量的作用范围，使之无远弗届。在现代物理中，对称性观念的最重要推广也许要算超对称。读者也许知道，自然界中的基本粒子可以分成两大类，即费米子和玻色子，像电子这样的粒子就是费米子，而像光子这样的粒子就是玻色子。费米子和玻色子性质截然不同，两者在物理理论中的作用也不一样。然而，超对称说，也许费米子与玻色子之间也存在着一种对称性，也许每一个玻色子都有一个费米子伴侣，反之也一样，也许玻色子和费米子在理论中的作用是平等的。

当然，超对称依然是一个猜测，目前还没有被实验所验证。但是，物理学家们的确构造出了很多具有超对称性的理论模型，这些模型中有些能解决现代物理中一些重大的难题。不仅如此，超对称的理论模型由于具有更大的对称性，因此理论的性质也会更好，所以常常成为人们求解物理问题以及发展物理想法的玩具模型，当前数学物理研究的一个重要方向就是研究各种超对称场论模型。我们这里无法用简单的几句话描绘出超对称的美妙，因为这需要在量子场论中才能看清楚。