目录

第二章　相对性原理与场论

陈童

上一章说过，我们主要考察满足狭义相对论的经典场论。本章就从相对论的基本原理出发，考察狭义相对论将给经典场论施加什么限制。

在本章中，对于一个带指标的量，将使用上、下两种指标，并且两者通常是有区别的。我们引入一个所谓的**求和约定**，即当一个表达式中某个上指标和某个下指标符号相同时，就默认对这个指标的所有可能性进行求和。

2.1　相对性原理

2.1.1　信息传播的最大速度

伽利略最早论述了相对性原理，它说的是，在所有惯性系中力学规律都是一样的。力学规律当然也要包括物体之间相互作用的规律，但是在伽利略和牛顿时代建立的理论中，两个物体之间的相互作用是瞬时进行的，比如牛顿的万有引力定律就是这样，这就是所谓的超距相互作用。正如牛顿已经注意到的，这肯定是有问题的，一个物体要和另一个物体相互作用，也就是要对另一个物体施加影响，这种影响不可能即时，而必定有一个时间的延迟，但是伽利略的力学相对性原理没有考虑到这一点。

爱因斯坦将相对性原理进行了推广，爱因斯坦提出，在所有惯性系中，一切物理规律——包括相互作用的传播规律——都是相同的。特别的，爱因

斯坦提出，相互作用的传播速度不是无穷大，而是有限的。这个相互作用传播速度有限的相对性原理就是爱因斯坦的相对性原理，今天谈相对性原理时都默认指的是这条原理。

的确，按照今天对经典物理的理解，一个物体要对另一个物体施加作用，就要向它发出一个信号，而受作用的物体只是对这个信号进行响应，在场论中，这个传播相互作用的信号就是场的波动。总之，物体间的相互作用需要信息的传递，但是信息传递的速度不可能像超距作用说的那样是无穷大，而必定是有限的。不妨记信息传播的最大速度为c(当然，正如后面的章节中将会证明的，c就是真空中的光速)，按照相对性原理，c必定不依赖于所选的惯性系，而是一个不变的常数，**因此我们当然可以合适地选择时间的单位，使得$c=1$。**

同时的相对性

信息传播的最大速度是常数的一个立即推论即是，两个不同地点的事件是否同时发生是依赖于参考系的，在一个惯性系中同时发生的两件事，在另一个惯性系中往往不同时，这就是同时的相对性。同时的相对性意味着，在狭义相对论中，时间不是绝对的，不同的惯性系将有不同的时间！

为了理解同时的相对性，让我们考虑两个惯性系，分别称为S系和S'系，假设S'沿着S的x轴正方向匀速运动。设想在S系的x轴上两个不同点A和B(假设$x_A<x_B$) 同时发生了两个事件，并且在事件发生的同时，A、B两地分别向着AB的中点以最大传播速度c发出一个信号，很显然，在S系位于AB中点的观察者看来，这两束信号将同时到达。并且正是根据这两者的同时到达，这个观察者才可以推断A、B两点的事件是同时发生的。现在假设在S'系的AB中点也有一个观察者，同样的道理，这个观察者对A、B两地的事件是否同时发生的判断也是依据他是否同时接收到事件发生时发出的那两个信号。

那么S'系的这个观察者是否会同时接收到那两个信号呢？根据c的不变性，答案将是否定的，实际上这个观察者将先接收到B地发出的信号。这是因为，根据最大信息传播速度的不变性，A、B两地发出的信号在S'的观察者看来依然以同样的速度c传播，但是，现在这两个信号将要到达的目的地–也就是AB在S'系中的中点–在朝着B地发出的信号运动(同时也在背着A地发出的信号运动)，因此S'系位于中点的这个观察者无疑会先接收到B地发出的信号。因此，在S'系的这个观察者就会进而得出结论，即A、

B两地的这两个事件不是同时发生的，而是B点的事件先发生，A点的事件后发生，因此同时是相对的。请千万要记住，在整个分析过程中，你心中绝对不能预设A、B两地的信号在S'系中也是同时发出的。

2.1.2　间隔不变性

依然考察S系和S'系这两个惯性系，假定同一事件在这两个参考系中的时空坐标分别是(t,x,y,z)和$(t',x',y'.z')$。则，由于时空的均匀性，这两个参考系之间的坐标变换一定是一个**线性变换**。

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

$$-dt^2+dx^2+dy^2+dz^2=0. \tag{2.1}$$

根据c的不变性，同样的两个事件在S'系看来也得满足

$$-dt'^2+dx'^2+dy'^2+dz'^2=0. \tag{2.2}$$

换言之，两个邻近事件在两不同参考系中的时空坐标必得满足一个约束关系，即当(2.1)成立时必有(2.2)成立，反之亦然。又由于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

$$-dt'^2+dx'^2+dy'^2+dz'^2=D(\mathbf{v})(-dt^2+dx^2+dy^2+dz^2), \tag{2.3}$$

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，就有

$$-dt^2+dx^2+dy^2+dz^2=D(-\mathbf{v})(-dt'^2+dx'^2+dy'^2+dz'^2). \tag{2.4}$$

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，$D(-\mathbf{v})D(\mathbf{v})=(D(v))^2=1$，即$D(v)=\pm1$。又由于$D(v)$是$v$的连续函数，而且$D(0)=1$(对应$S$和$S'$为同一个参考系的情形)，因此必有$D(v)=1$。因此，对于任意两个邻近事件我们必有

$$-dt'^2+dx'^2+dy'^2+dz'^2=-dt^2+dx^2+dy^2+dz^2. \tag{2.5}$$

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

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 = -dt^2 + d\mathbf{x}^2. \tag{2.6}$$

用这个记号，方程(2.5)就可以简记成

$$ds'^2 = ds^2, \tag{2.7}$$

称为两个事件的间隔不变性。

人们通常约定$t = ct = x^0, x = x^1, y = x^2, z = x^3$，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

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \tag{2.8}$$

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$\eta_{\mu\nu} = \eta_{\nu\mu}$。

我们可以将$\eta_{\mu\nu}$看成是一个4×4矩阵的分量形式(记这个矩阵为η)，进而引入这个矩阵的逆矩阵η^{-1}，其分量形式记为$\eta^{\mu\nu}$，

$$\eta_{\alpha\beta}\eta^{\beta\gamma} = \delta_\alpha^\gamma, \tag{2.9}$$

式中δ_α^γ为4×4单位矩阵的分量形式。很容易看出，$\eta^{\mu\nu}$也为，$-\eta^{00} = \eta^{11} = \eta^{22} = \eta^{33} = 1$, 其它指标分量都等于零。

假设我们考察的不是两个邻近时空点，而是两个有限间隔的时空点x_1^μ和x_2^μ, 记$\Delta x^\mu = x_2^\mu - x_1^\mu$，记这两个事件的时空间隔为$\Delta s$，则有

$$(\Delta s)^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu = -(\Delta t)^2 + (\Delta \mathbf{x})^2. \tag{2.10}$$

同样，无论在哪个参考系中计算，间隔Δs都是不变的。1. 如果$(\Delta s)^2 < 0$, 我们就称x_1^μ和x_2^μ这两个事件**类时相间**(timelike separated)。由于这时候$(\frac{\Delta \mathbf{x}}{\Delta t})^2 < 1$，所以我们总可以用信号将这两个事件联系起来，所以这两个事件就存在因果关系，而且正如马上就会看到的，在不同的惯性系中，事件的因果关系将会保持不变，先发生的事件在任何惯性系中都会先发生。2. 如果$(\Delta s)^2 = 0$, 我们就称这两个事件**类光相间**(lightlike separated)。这

时候可以用光信号将两个事件联系起来。3. 如果$(\Delta s)^2 > 0$, 我们就称这两个事件**类空相间**(spacelike separated)。这时候两事件没有任何因果关系，不可能用任何信号将它们联系起来，同时它们的先后顺序也是相对的，在不同参考系中对哪个事件先发生会有不同的看法。

为了将上述两事件间的关系看得更清楚，我们取其中一个事件为$x_1^\mu = 0$, 即位于时空图的坐标原点，另一个事件$x_2^\mu = x^\mu$，记事件x^μ与原点事件的间隔为s，则

$$s^2 = -t^2 + \mathbf{x}^2. \tag{2.11}$$

对于$s^2 = 0$的类光相间情形，方程$-t^2 + \mathbf{x}^2 = 0$给出的是时空图上以原点为顶点的圆锥面，称之为光锥(lightcone), 如图(2.1)所示。 对于$s^2 < 0$情形，

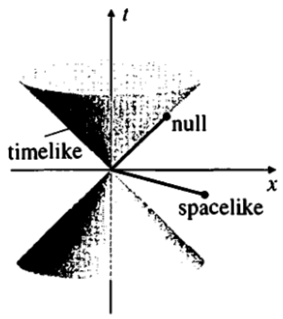

图 2.1: 时空图上原点处的光锥。

这时候方程$t^2 - \mathbf{x}^2 = -s^2 > 0$给出的是具有两支的双曲面，一支位于上半光锥所包围的内部区域，一支位于下半光锥所包围的内部区域，由于不同惯性系中$t^2 - \mathbf{x}^2 = -s^2$保持不变，因此在不同惯性系中事件2的坐标$x^\mu$只能在这个双曲面上变动。不过由于双曲面被光锥分隔成了两支，所以在不同惯性系中上半支的点只能在上半支上变动，即恒有$t > 0$，而下半支的点则恒有$t < 0$, 即是说，处于原点未来的事件在任何惯性系中都保持在未来，而处于原点过去的事件在任何惯性系中都在过去，即，不同参考系不会改变事件间的因果关系。要让双曲面上半支的点变到下半支只能是经过一个$t \to -t$的时间反演。最后，对于$s^2 > 0$情形, 方程$\mathbf{x}^2 - t^2 = s^2 > 0$给出的是一个连通的双曲面，原则上，这个双曲面上的任何点可以在一个合适的参考系中变到双曲面上的任何其它点，特别的，$t > 0$的点可以变到$t < 0$，反之亦然。即是说，与原点类空相间的事件是先于原点发生还是后发生并没有绝对的意义！

固有时

假如原来有一个参考系，有一个粒子从参考系的(t,x,y,z)点运动到$(t+dt,x+dx,y+dy,z+dz)$点，这两点(它们当然类时相间)间的间隔当然满足$ds^2=\eta_{\mu\nu}dx^\mu dx^\nu$。现在，假设有一个钟固定在这个粒子上，记此钟走过的时间为$d\tau$，并且我们依托这个钟建立一个固定在粒子上的参考系，那么在这个参考系中，粒子的空间位移当然是零，从而从这个固定在粒子的参考系看来，间隔ds应该满足，$ds^2=-d\tau^2$，即

$$\eta_{\mu\nu}dx^\mu dx^\nu=-d\tau^2. \tag{2.12}$$

上式中的τ就称之为粒子走过的固有时，而$x^0=t$则称之为坐标时。很显然，固有时就是固连在粒子上的钟所走过的时间。

假设粒子的速度为$\mathbf{v}$，即$d\mathbf{x}=\mathbf{v}dt$，则根据(2.12)式即有

$$d\tau^2=dt^2-d\mathbf{x}^2=dt^2(1-\mathbf{v}^2)\Rightarrow dt=\frac{d\tau}{\sqrt{1-\mathbf{v}^2}}. \tag{2.13}$$

即是说，坐标时总是比固有时长的。测固有时的钟当然是相对于原参考系运动的钟，所以这个结果常常也被人们说成是，运动的钟会变慢，因为对同一个参考系中的过程，它测出来的时间更短。

但是，这并非运动的钟本身有什么问题，而是从原参考系的静止观察者来看，运动的一切事物都变慢了，运动的人的生命过程也变慢了。不过，从运动的人自己来看，他自己的一切都是正常的，在他看来，反而是原参考系中的观察者在运动(运动是相对的)，反而是这观察者的生命过程变慢了。

那么假设有两只钟，一只是静止的，另一只沿着闭合路径运动一圈再回到起点与静止的钟比较，那到底哪只钟慢了呢? 回答是，运动的钟绝对地慢了。但是运动不是相对的吗? 从运动钟来看，不是静止的钟在运动吗? 但是，从这以运动的钟为参考系的后一种观点导不出静止的钟变慢的结论，因为这时候这个参考系不是一个惯性系，这是由于这个运动的钟是沿着闭合路径运动一圈而不是作匀速直线运动。

2.1.3 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

$$x'^\mu=\Lambda^\mu{}_\nu x^\nu, \tag{2.14}$$

式中$\Lambda^\mu_{\ \nu}$ 构成变换矩阵Λ的分量形式。根据间隔不变性，我们有

$$\eta_{\mu\nu}dx'^\mu dx'^\nu = \eta_{\mu\nu}\Lambda^\mu_{\ \alpha}\Lambda^\nu_{\ \beta}dx^\alpha dx^\beta = \eta_{\alpha\beta}dx^\alpha dx^\beta, \tag{2.15}$$

从而即有

$$\eta_{\mu\nu}\Lambda^\mu_{\ \alpha}\Lambda^\nu_{\ \beta} = \eta_{\alpha\beta}. \tag{2.16}$$

更一般地，人们也把任何满足上式的时空坐标变换$\Lambda^\mu_{\ \nu}$称作洛伦兹变换，哪怕它不是源于不同惯性系之间的变换。

我们当然可以将(2.16)式写成矩阵形式，即

$$\Lambda^T\eta\Lambda = \eta. \tag{2.17}$$

很容易验证，如果洛伦兹变换Λ_1满足上面式子，Λ_2也满足上面式子，则$\Lambda_1\Lambda_2$必定也满足上面式子，从而也是洛伦兹变换。即是说，所有洛伦兹变换的集合在矩阵乘法下封闭。另外，对(2.17)式两边求行列式，并注意到$\det(\eta) = -1$, 从而即可得

$$[\det(\Lambda)]^2 = 1 \Rightarrow \det(\Lambda) = \pm 1. \tag{2.18}$$

由此可知矩阵Λ必定存在逆矩阵Λ^{-1}，并且很明显Λ^{-1}也是洛伦兹变换，即任何洛伦兹变换都有逆变换。满足乘法封闭性，并且存在逆元素的元素集合就是数学上所谓的群，所以，所有洛伦兹变换的集合构成一个群，称作洛伦兹群，常常记作$O(1,3)$。很明显，所有$\det(\Lambda) = 1$的洛伦兹变换也构成一个群，它是$O(1,3)$的子群，通常记作$SO(1,3)$，实际上，人们在谈到洛伦兹群的时候更多都是指的这个$SO(1,3)$群。

进一步，在(2.16)式中取$\alpha = \beta = 0$, 即可得

$$(\Lambda^0_{\ 0})^2 = 1 + \sum_{i=1,2,3}(\Lambda^i_{\ 0})^2 \geq 1 \ \Rightarrow \ \Lambda^0_{\ 0} \geq 1 \text{ or } \Lambda^0_{\ 0} \leq -1. \tag{2.19}$$

从而根据$\det(\Lambda)$的正负以及$\Lambda^0_{\ 0}$的正负，我们可以将洛伦兹变换的集合分成四个子集。其中所谓的正洛伦兹变换要求满足下面条件

$$\det(\Lambda) = 1, \quad \Lambda^0_{\ 0} \geq 1. \tag{2.20}$$

容易验证，正洛伦兹变换的集合也构成一个群，称作正洛伦兹群，它是洛伦兹群的子群，通常记为$SO^+(1,3)$。

一般来说，洛伦兹变换依赖于某些连续参数，比如两个参考系的相对速度以及坐标轴的相对角度等等。最简单的洛伦兹变换就是$\Lambda^\mu{}_\nu = \delta^\mu_\nu$的恒等变换，它描写两个始终完全重合的参考系。但是，我们可以设想连续地改变洛伦兹变换的参数，使得这两个参考系变得不再重合，从而使得恒等变换变成非平凡的洛伦兹变换。而正洛伦兹变换的集合正是能够由恒等变换连续地变化过来的所有洛伦兹变换。这是因为，恒等变换显然满足$\det(\Lambda) = 1$, $\Lambda^0{}_0 \geq 1$, 而连续的变化不可能使得$\det(\Lambda)$从$+1$突变到-1, 也不可能使得$\Lambda^0{}_0$ 从≥ 1突变到≤ -1，所以能和恒等变换连续过渡的洛伦兹变换一定是正洛伦兹变换。

另外，按照洛伦兹变换的一般定义，$t' = -t$, $\mathbf{x}' = \mathbf{x}$的时间反演变换显然也是一种洛伦兹变换，它的变换矩阵Λ显然满足，$\Lambda^0{}_0 = -1$, $\det(\Lambda) = -1$。因此，在任何正洛伦兹变换的基础上再进行一个时间反演变换，得到的就是$\det(\Lambda) = -1$, $\Lambda^0{}_0 \leq -1$的洛伦兹变换。类似的，空间反演变换$t' = t, \mathbf{x}' = -\mathbf{x}$也是洛伦兹变换，它的变换矩阵满足$\Lambda^0{}_0 = 1$, $\det(\Lambda) = -1$，因此，在正洛伦兹变换的基础上再进行一个空间反演变换，得到的就是$\det(\Lambda) = -1$, $\Lambda^0{}_0 \geq 1$的洛伦兹变换。而在正洛伦兹变换的基础上再同时进行时间和空间的反演，得到的就是$\det(\Lambda) = 1$, $\Lambda^0{}_0 \leq -1$的洛伦兹变换。

正洛伦兹变换的一个例子就是三维空间旋转，即如下洛伦兹变换

$$\Lambda^0{}_0 = 1, \quad \Lambda^0{}_i = \Lambda^i{}_0 = 0, \quad \Lambda^i{}_j = R_{ij}, \tag{2.21}$$

式中$i, j = 1, 2, 3$，而3×3矩阵R是特殊正交矩阵，即满足$R^T R = 1, \det(R) = 1$。不过，通常我们谈到洛伦兹变换的时候不是指这种旋转变换，虽然它的确很重要。

一些重要的洛伦兹变换

下面我们来构造一个最为通常而重要的洛伦兹变换。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(2.6)中的y, z坐标，进而将间隔公式简化为

$$ds^2 = -dt^2 + dx^2 = dx^+ dx^-, \tag{2.22}$$

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$x'^+=x'+t', x'^-=x'-t'$必然满足$dx'^+dx'^-=dx^+dx^-$。显然这就意味着必定有某个参数ω使得$x'^+=e^{-\omega}x^+$, $x'^-=e^{\omega}x^-$, 即有

$$\begin{aligned} x' &= \cosh(\omega)x-\sinh(\omega)t, \qquad (2.23)\\ t' &= -\sinh(\omega)x+\cosh(\omega)t. \qquad (2.24)\end{aligned}$$

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

$$\begin{aligned} t' &= \frac{t-vx}{\sqrt{1-v^2}},\\ x' &= \frac{x-vt}{\sqrt{1-v^2}}, \qquad (2.25)\end{aligned}$$

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。而且，反过来，S系也在以$-v$的速度相对S'系运动，所以必定有

$$\begin{aligned} t &= \frac{t'+vx'}{\sqrt{1-v^2}},\\ x &= \frac{x'+vt'}{\sqrt{1-v^2}}. \qquad (2.26)\end{aligned}$$

另外，也容易得到$e^{-\omega}=\cosh\omega-\sinh\omega=\left(\frac{1-v}{1+v}\right)^{\frac{1}{2}}$, 从而即有

$$x'+t'=\left(\frac{1-v}{1+v}\right)^{\frac{1}{2}}(x+t), \quad x'-t'=\left(\frac{1+v}{1-v}\right)^{\frac{1}{2}}(x-t). \qquad (2.27)$$

如果先将S系变换到以速度v_1沿着x轴运动的S'系，从而即有$x'+t'=\left(\frac{1-v_1}{1+v_1}\right)^{\frac{1}{2}}(x+t)$，接着再变换到以速度$v_2$相对$S'$系沿$x$运动的$S''$系，即有$x''+t''=\left(\frac{1-v_2}{1+v_2}\right)^{\frac{1}{2}}(x'+t')$。则$S''$系相对$S$系的变换将是，$x''+t''=\left(\frac{1-v}{1+v}\right)^{\frac{1}{2}}(x+t)=\left(\frac{1-v_2}{1+v_2}\right)^{\frac{1}{2}}\left(\frac{1-v_1}{1+v_1}\right)^{\frac{1}{2}}(x+t)$，式中$v$为$S''$系相对于$S$系的速度。从而即有，

$$\left(\frac{1-v}{1+v}\right)^{\frac{1}{2}}=\left(\frac{1-v_1}{1+v_1}\right)^{\frac{1}{2}}\cdot\left(\frac{1-v_2}{1+v_2}\right)^{\frac{1}{2}}. \qquad (2.28)$$

很明显，只要$v_1<1, v_2<1$，则上面等式右边就是两个正实数的乘积，因此结果也必定是正实数，从而必有$v<1$，即是说，我们不可能通过多次洛伦兹变换使得相对运动速度大于光速。另外，从(2.28)式也容易得到

$$\frac{1-v}{1+v}=\frac{1-v_1-v_2+v_1v_2}{1+v_1+v_2+v_1v_2}. \qquad (2.29)$$

也即有

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}, \tag{2.30}$$

这就是相对论的速度合成公式。

下面，我们对(2.25)式给出的洛伦兹变换进行一个推广。依然考虑两个参考系S与S'之间的变换，同样假设初始时这两个参考系的坐标轴完全重合。不过，现在S'系相对于S系以任意方向的速度$\mathbf{v}$运动。以$v = |\mathbf{v}|$表示速度的大小，以$v_i, i = 1, 2, 3$表示速度$\mathbf{v}$的各分量。为了利用(2.25)式，我们将坐标矢量$\mathbf{x}$正交分解成平行于速度$\mathbf{v}$的$\mathbf{x}_\parallel$和垂直于速度$\mathbf{v}$的$\mathbf{x}_\perp$两部分，它们满足

$$\mathbf{x}_\parallel = \frac{\mathbf{v}(\mathbf{v}\cdot\mathbf{x})}{v^2}, \quad \mathbf{x}_\perp = \mathbf{x} - \mathbf{x}_\parallel. \tag{2.31}$$

则根据(2.25)式，容易写出下面洛伦兹变换关系

$$\begin{aligned} t' &= \frac{t - \mathbf{v}\cdot\mathbf{x}}{\sqrt{1-v^2}}, \\ \mathbf{x}'_\parallel &= \frac{\mathbf{x}_\parallel - \mathbf{v}t}{\sqrt{1-v^2}}, \\ \mathbf{x}'_\perp &= \mathbf{x}_\perp. \end{aligned} \tag{2.32}$$

由此即可以得到，

$$\begin{aligned} \mathbf{x}' &= \mathbf{x}'_\perp + \mathbf{x}'_\parallel = (\mathbf{x} - \mathbf{x}_\parallel) + \frac{\mathbf{x}_\parallel - \mathbf{v}t}{\sqrt{1-v^2}} \\ &= \mathbf{x} + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\frac{\mathbf{v}(\mathbf{v}\cdot\mathbf{x})}{v^2} - \frac{\mathbf{v}}{\sqrt{1-v^2}}t. \end{aligned} \tag{2.33}$$

将以上结果用洛伦兹变换矩阵$\Lambda^\mu_{\ \nu}$写出来即是，

$$\begin{aligned} &\Lambda^0_{\ 0} = \frac{1}{\sqrt{1-v^2}} \\ &\Lambda^0_{\ i} = -\frac{v_i}{\sqrt{1-v^2}} = \Lambda^i_{\ 0} \\ &\Lambda^i_{\ j} = \delta_{ij} + \left(\frac{1}{\sqrt{1-v^2}} - 1\right)\frac{v_i v_j}{v^2}. \end{aligned} \tag{2.34}$$

很显然，当速度$\mathbf{v} \to 0$时，这种洛伦兹变换就退化为恒等变换，所以它可以与恒等变换连续过渡，从而必定是正洛伦兹变换。实际上，任何正洛伦兹变换都可以表示成以上洛伦兹变换和空间旋转变换R的乘积。

利用上面的洛伦兹变换，我们可以计算一下两个事件在不同参考系中的时间差，即有

$$\Delta t' = \frac{\Delta t - v|\Delta \mathbf{x}_{\parallel}|}{\sqrt{1-v^2}} = \Delta t\Big(\frac{1 - v\frac{|\Delta \mathbf{x}_{\parallel}|}{\Delta t}}{\sqrt{1-v^2}}\Big). \tag{2.35}$$

如果两事件类时相间，则有$|\Delta \mathbf{x}_{\parallel}| \le |\Delta \mathbf{x}| < |\Delta t|$, 即$\frac{|\Delta \mathbf{x}_{\parallel}|}{|\Delta t|} < 1$, 从而

$$1 - v\frac{|\Delta \mathbf{x}_{\parallel}|}{\Delta t} \ge 1 - v\frac{|\Delta \mathbf{x}_{\parallel}|}{|\Delta t|} \ge 1 - \frac{|\Delta \mathbf{x}_{\parallel}|}{|\Delta t|} > 0. \tag{2.36}$$

从而，对于这种类时相间的情况，(2.35)式子左边$\Delta t'$的正负符号与右边Δt的正负符号相同，即是说，这时候洛伦兹变换不会改变事件的先后顺序，也就是会保持事件之间的因果关系！这个结论当然是我们已经知道了的。并且，从上面的讨论也可以看出，如果两事件类空相间，那它们之间的先后顺序就是相对的，在不同的参考系看来可能有不同的结论。

2.2　四维时空的矢量和张量

2.2.1　四维矢量和张量

上一节说过，在不同的参考系中，时空坐标按照下面的洛伦兹变换而变换，

$$dx'^{\mu} = \Lambda^{\mu}_{\ \nu} dx^{\nu}, \tag{2.37}$$

式中变换矩阵Λ满足

$$\Lambda^T \eta \Lambda = \eta \ \Leftrightarrow \ \Lambda^{-1} = \eta^{-1}\Lambda^T \eta. \tag{2.38}$$

类比于四分量的dx^{μ}，假设一个任意的四分量量$A^{\mu} = (A^0, \mathbf{A})$在参考系的变换下与$dx^{\mu}$的变换规则相同，即满足

$$A'^{\mu} = \Lambda^{\mu}_{\ \nu} A^{\nu}, \tag{2.39}$$

则我们称A^{μ}为一个四维时空的矢量，简称四矢量，当然严格来讲A^{μ}是四矢量的分量形式。与时空间隔类似，我们可以定义四矢量的平方A^2为

$$A^2 = \eta_{\mu\nu} A^{\mu} A^{\nu} = -(A^0)^2 + \mathbf{A}^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2. \tag{2.40}$$

很明显，A^2在洛伦兹变换下是不变的。

我们也可以定义下指标的四分量量A_μ为，

$$A_\mu = \eta_{\mu\nu} A^\nu, \tag{2.41}$$

写得更清楚一点就是

$$A_\mu = (A_0, A_1, A_2, A_3) = (-A^0, A^1, A^2, A^3) = (A_0, \mathbf{A}). \tag{2.42}$$

则A^2就可以写成$A^2 = A_\mu A^\mu$，而A^2在洛伦兹变换下的不变性则意味着

$$A'_\mu A'^\mu = A_\mu A^\mu. \tag{2.43}$$

注意到A^μ在洛伦兹变换下按照(2.39)式变换，因此上式就意味着A_μ必然按照下式变换

$$A'_\mu = (\Lambda^{-1})^\nu{}_\mu A_\nu. \tag{2.44}$$

在洛伦兹变换下按照这样变换的量同样叫做四矢量。不过为了区分上指标的四矢量和下指标的四矢量，有时候人们称A^μ为四矢量的逆变分量，而称A_μ为四矢量的协变分量。利用$\eta_{\mu\nu}$我们可以把上指标降下来，进而将逆变分量转化为协变分量，反过来，我们也可以利用$\eta^{\mu\nu}$将下指标升上去，即

$$A^\mu = \eta^{\mu\nu} A_\nu. \tag{2.45}$$

假设记$\partial_\mu = \frac{\partial}{\partial x^\mu}$，则不难明白全微分$d = dx^\mu \partial_\mu$是不依赖于坐标系的，由此即可以看出，偏导运算$\partial_\mu$在洛伦兹变换下和协变四矢量的变换规则相同，即按下式变换

$$\partial'_\mu = (\Lambda^{-1})^\nu{}_\mu \partial_\nu. \tag{2.46}$$

归纳一下即是，在洛伦兹变换下，四矢量的逆变分量和dx^μ的变换规则相同，而协变分量则和∂_μ的变换规则相同。

很显然，任意一个逆变四矢量A^μ和任意一个协变四矢量B_μ都可以构成一个在洛伦兹变换下保持不变的量，这个量即是$A^\mu B_\mu$, 有时候也记作$A \cdot B$，称作两个四矢量A和B的内积。两个四矢量的内积是洛伦兹不变的，称作一个四维标量，四维标量即是在洛伦兹变换下保持不变的量。

四维矢量是只有一个指标的量，我们当然可以进一步考察多个指标的量，比如$B^{\mu\nu}$, 如果这个量的每一个指标在洛伦兹变换下都按逆变矢量那样变，即是说

$$B'^{\mu\nu} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}B^{\alpha\beta}, \tag{2.47}$$

我们就称$B^{\mu\nu}$为一个2阶逆变张量，或者记作$(2,0)$张量，$(2,0)$代表它有2个上指标0个下指标。类似的，我们也可以考察$(0,2)$张量，它即是两个下指标，且在洛伦兹变换下按照下式变换的量，

$$B'_{\mu\nu} = (\Lambda^{-1})^{\alpha}{}_{\mu}(\Lambda^{-1})^{\beta}{}_{\nu}B_{\alpha\beta}. \tag{2.48}$$

进一步，也可以考察混合型张量，比如$(1,1)$张量，它即是一个上指标一个下指标，且在洛伦兹变换下按照下式变换的量，

$$B'^{\mu}_{\nu} = \Lambda^{\mu}{}_{\alpha}(\Lambda^{-1})^{\beta}{}_{\nu}B^{\alpha}_{\beta}. \tag{2.49}$$

类似的概念可以很容易推广到有p个上指标q个下指标的(p,q)张量。特别的，$(0,0)$张量就是四维标量，$(1,0)$张量就是四维逆变矢量，而$(0,1)$张量则是四维协变矢量。

当然，完全类似于四矢量情形，我们同样可以用$\eta_{\mu\nu}$来将张量的上指标降下来，也可以用$\eta^{\mu\nu}$来将张量的下指标升上去。而且，对于一个(p,q)张量，我们可以让它的某个上指标和某个下指标相同，从而默认对这个指标求和，结果就是一个$(p-1,q-1)$张量，这就叫做张量的缩并。比如说，对于$(1,1)$张量B^{μ}_{ν}，我们可以考察$B^{\mu}_{\mu} = B^{0}_{0} + B^{1}_{1} + B^{2}_{2} + B^{3}_{3}$，注意它的上指标和下指标已经求和掉了，从而人们很容易验证它是洛伦兹不变的，即是一个$(0,0)$张量，或者说是一个四维标量。

另外，比如说对于$(2,0)$张量$B^{\mu\nu}$，我们可以进一步要求它的两个指标对称，即满足$B^{\mu\nu} = B^{\nu\mu}$，这就叫二阶对称张量。而如果我们要求两个指标反对称，即满足$B^{\mu\nu} = -B^{\nu\mu}$，那就叫二阶反对称张量。对于反对称张量$B^{\mu\nu}$，我们有$B^{00} = B^{11} = B^{22} = B^{33} = 0$, 这是因为比如说$B^{00} = -B^{00}$，从而必有$B^{00} = 0$。对于$(0,p)$张量$C_{\mu_1\mu_2\ldots\mu_p}$，如果它的任意两个指标均反对称，我们就称之为$p$阶反对称张量。但是在四维时空中，必定有$p \leq 4$。这是因为，在四维时空中，任何指标都只能取$0,1,2,3$，从而对于$p > 4$的情形，$C_{\mu_1\mu_2\ldots\mu_p}$的任意$p$个下指标中必有两个取相同值，考虑到反对称这就意味着$C_{\mu_1\mu_2\ldots\mu_p} = 0$, 即高于4阶的反对称张量必定为零。进一步，由于$p$阶反

对称张量场p个指标必须全不相同，所以在四维时空中，它的独立分量个数就是$\frac{4!}{p!(4-p)!}$。

最后，四维张量的概念很容易推广到场，如果一个量既是一个四维张量，同时还是一个场，那就叫做张量场，比如一个$(0,2)$型二阶张量场可以写成$B_{\mu\nu}(x)$，式中x表示时空点。

2.2.2　四维时空中的微分形式

在《理论力学新讲》的第一章，我们介绍过微分形式和外微分的知识，实际上，微分形式与反对称张量场有密切的联系，这一小节就让我们从这个联系开始。具体来说，对于一个p-形式C，我们可以写出它的分量形式，

$$C = \frac{1}{p!} C_{\mu_1\mu_2...\mu_p} dx^{\mu_1} \wedge dx^{\mu_2} \wedge .... \wedge dx^{\mu_p}. \tag{2.50}$$

式中$C_{\mu_1\mu_2...\mu_p}$为这个p形式的分量，它的指标是两两反对称的。事实上，$C_{\mu_1\mu_2...\mu_p}$必定是一个p阶反对称张量场。这是因为，p-形式本身不依赖于坐标系，因此洛伦兹变换到x'^μ坐标后必有

$$C'_{\mu_1\mu_2...\mu_p}(x') dx'^{\mu_1} \wedge .... \wedge dx'^{\mu_p} = C_{\nu_1\nu_2...\nu_p}(x) dx^{\nu_1} \wedge .... \wedge dx^{\nu_p}. \tag{2.51}$$

由此易知$C_{\mu_1\mu_2...\mu_p}$在洛伦兹变换下必定按照$(0,p)$张量的变换规则变换。从而p-形式的分量和p阶反对称张量场一一对应。

在四维时空中，由于p阶反对称张量场必须满足$p \leq 4$，因此在四维时空中最多考虑4-形式，更高阶的微分形式自动为零。

有一个特殊的4-形式值得专门讲一下，那就是所谓的四维时空的体积形式Ω，

$$\Omega = dt \wedge dx \wedge dy \wedge dz. \tag{2.52}$$

但这个四形式严格来说算不上一个微分形式，而是所谓的赝形式，原因在于它在坐标变换以后会相差一个雅可比行列式，因此洛伦兹变换以后

$$\Omega' = dt' \wedge dx' \wedge dy' \wedge dz' = \det(\Lambda)\Omega. \tag{2.53}$$

因此对于$\det(\Lambda) = 1$的洛伦兹变换，它和一个真正的微分形式一样不依赖于坐标系。但是，对于$\det(\Lambda) = -1$的洛伦兹变换，比如空间反演变换或时间反演变换，那它就要多出一个负号，而不是真正不依赖于坐标系。像这

样的在空间反演和时间反演之下多出一个负号的“微分形式”就叫做赝形式。但是，对于物理学研究来说，由于主要考虑的是正洛伦兹变换，这时候自动有$\det(\Lambda)=1$, 因此我们常常将赝形式和真正的微分形式同等对待。

赝形式的分量也不是真正的反对称张量，而是所谓的赝张量，即它在空间反演或时间反演之下相比于真正张量的变换规则会多出一个负号。但是，和平等对待赝形式一样，在物理学中，我们也常常将赝张量和张量平等对待。因此，以后除非某些地方需要特别做出区分，否则我们就将赝形式同样称为微分形式，也将赝张量同样称为张量。

为了写出体积形式Ω所对应的反对称张量，我们引入如下定义

$$\epsilon_{\mu_1\mu_2\mu_3\mu_4}=\begin{cases}1, & (\mu_1\mu_2\mu_3\mu_4)\text{是}(0123)\text{的偶置换}\\ -1, & (\mu_1\mu_2\mu_3\mu_4)\text{是}(0123)\text{的奇置换}\\ 0, & \text{其它情况}\end{cases}. \tag{2.54}$$

容易验证有，

$$\Omega=\frac{1}{4!}\epsilon_{\mu_1\mu_2\mu_3\mu_4}dx^{\mu_1}\wedge dx^{\mu_2}\wedge dx^{\mu_3}\wedge dx^{\mu_4}. \tag{2.55}$$

可见$\epsilon_{\mu_1\mu_2\mu_3\mu_4}$正是体积形式所对应的四阶反对称张量(赝张量)，不过，由于$\epsilon_{\mu_1\mu_2\mu_3\mu_4}$的定义不依赖于坐标系，所以这个反对称张量实际上在洛伦兹变换下变换的结果是**保持原值不变**。

我们注意到p形式的独立分量个数与$4-p$形式的独立分量个数相同，均为$\frac{4!}{p!(4-p)!}$, 这使得人们想到也许可以建立两者间的一一对映。的确，这样的对映是存在的，它叫做霍奇对偶(Hodge duality)，通常用$*$号来标记。具体来说，我们定义$*$号为一个线性映射，它在微分形式上的作用如下

$$*(dx^{\mu_1}\wedge...\wedge dx^{\mu_p})=\frac{1}{(4-p)!}\epsilon^{\mu_1...\mu_p}{}_{\nu_{p+1}...\nu_4}dx^{\nu_{p+1}}\wedge...\wedge dx^{\nu_4}. \tag{2.56}$$

特别的，$*1$将映射到体积形式Ω

$$*1=\frac{1}{4!}\epsilon_{\mu_1\mu_2\mu_3\mu_4}dx^{\mu_1}\wedge dx^{\mu_2}\wedge dx^{\mu_3}\wedge dx^{\mu_4}=\Omega. \tag{2.57}$$

而对于任意的p形式C,

$$C=\frac{1}{p!}C_{\mu_1\mu_2...\mu_p}dx^{\mu_1}\wedge dx^{\mu_2}\wedge....\wedge dx^{\mu_p}, \tag{2.58}$$

根据线性性，我们有

$$
\begin{aligned}
*C &= \frac{1}{p!} C_{\mu_1 \dots \mu_p} * (dx^{\mu_1} \wedge \dots . \wedge dx^{\mu_p}) \\
&= \frac{1}{(4-p)!p!} C_{\mu_1 \dots \mu_p} \epsilon^{\mu_1 \dots \mu_p}{}_{\nu_{p+1} \dots \nu_4} dx^{\nu_{p+1}} \wedge \dots \wedge dx^{\nu_4}, \qquad (2.59)
\end{aligned}
$$

结果显然为$(4-p)$形式。

特别的，2形式的霍奇对偶依然为2形式！对于2形式，我们也常常可以等价地认为霍奇对偶是作用在它的分量上，这是通过定义$*C_{\mu_1\mu_2} = (*C)_{\mu_1\mu_2}$。则根据(2.59)式，易有

$$
*C_{\mu_1\mu_2} = \frac{1}{2} C_{\nu_3\nu_4} \epsilon^{\nu_3\nu_4}{}_{\mu_1\mu_2}. \qquad (2.60)
$$

定理： 对于p形式C,

$$
**C = (-)(-)^{p(4-p)} C. \qquad (2.61)
$$

即，对于$p = 1,3$形式，有$**C = C$; 对于$p = 0,2,4$形式，有$**C = -C$; 特别的，对于$p = 2$形式，有$**C = -C$。

证明如下。重复应用$*$映射的作用，有

$$
\begin{aligned}
**C &= \frac{1}{(4-p)!p!} C_{\mu_1 \dots \mu_p} \epsilon^{\mu_1 \dots \mu_p}{}_{\nu_{p+1} \dots \nu_4} * (dx^{\nu_{p+1}} \wedge \dots \wedge dx^{\nu_4}) \\
&= \frac{1}{(4-p)!p!p!} C_{\mu_1 \dots \mu_p} \epsilon^{\mu_1 \dots \mu_p}{}_{\nu_{p+1} \dots \nu_4} \epsilon^{\nu_{p+1} \dots \nu_4}{}_{\rho_1 \dots \rho_p} dx^{\rho_1} \wedge \dots \wedge dx^{\rho_p} \\
&= \frac{(-)^{p(4-p)}}{(4-p)!p!p!} C_{\mu_1 \dots \mu_p} \epsilon^{\mu_1 \dots \mu_p}{}_{\nu_{p+1} \dots \nu_4} \epsilon_{\rho_1 \dots \rho_p}{}^{\nu_{p+1} \dots \nu_4} dx^{\rho_1} \wedge \dots \wedge dx^{\rho_p} \\
&= \frac{(-)^{p(4-p)}}{(4-p)!p!p!} C_{\mu_1 \dots \mu_p} \epsilon^{\mu_1 \dots \mu_p \nu_{p+1} \dots \nu_4} \epsilon_{\rho_1 \dots \rho_p \nu_{p+1} \dots \nu_4} dx^{\rho_1} \wedge \dots \wedge dx^{\rho_p} \\
&= (-)(-)^{p(4-p)} \frac{1}{p!} C_{\mu_1 \dots \mu_p} dx^{\mu_1} \wedge \dots \wedge dx^{\mu_p} = (-)(-)^{p(4-p)} C. \qquad (2.62)
\end{aligned}
$$

式中我们利用了恒等式

$$
\begin{aligned}
& \frac{1}{p!(4-p)!} \epsilon^{\mu_1 \dots \mu_p \nu_{p+1} \dots \nu_4} \epsilon_{\rho_1 \dots \rho_p \nu_{p+1} \dots \nu_4} dx^{\rho_1} \wedge \dots \wedge dx^{\rho_p} \\
= & \ (-) dx^{\mu_1} \wedge \dots \wedge dx^{\mu_p}, \qquad (2.63)
\end{aligned}
$$

这里的负号来自于$\epsilon^{0123} = -\epsilon_{0123} \Rightarrow \epsilon^{\mu_1 \dots \mu_4} = (-)\epsilon_{\mu_1 \dots \mu_4}$。

霍奇对偶运算可以和外微分运算结合起来，为了本书后面章节的应用，让我们考察一个重要的例子。假设我们考察一个2形式F,

$$F=\frac{1}{2}F_{\mu_1\mu_2}dx^{\mu_1}\wedge dx^{\mu_2}, \tag{2.64}$$

让我们来计算一下$*d*F$的结果是什么。

$$\begin{aligned}
*d*F &= \frac{1}{2}\frac{1}{2}*d\Big(F_{\mu_1\mu_2}\epsilon^{\mu_1\mu_2}{}_{\nu_3\nu_4}dx^{\nu_3}\wedge dx^{\nu_4}\Big)\\
&= \frac{1}{2}\frac{1}{2}\partial_\rho F_{\mu_1\mu_2}\epsilon^{\mu_1\mu_2}{}_{\nu_3\nu_4}*(dx^\rho\wedge dx^{\nu_3}\wedge dx^{\nu_4})\\
&= \frac{1}{2}\frac{1}{2}\partial_\rho F_{\mu_1\mu_2}\epsilon^{\mu_1\mu_2}{}_{\nu_3\nu_4}\epsilon^{\rho\nu_3\nu_4}{}_\sigma dx^\sigma\\
&= \frac{1}{2}\frac{1}{2}\partial^\rho F_{\mu_1\mu_2}\epsilon^{\mu_1\mu_2\nu_3\nu_4}\epsilon_{\rho\sigma\nu_3\nu_4}dx^\sigma\\
&= -\partial^\rho F_{\rho\sigma}dx^\sigma.
\end{aligned} \tag{2.65}$$

式中我们再次应用了前面用过的恒等式

$$\frac{1}{2}\frac{1}{2}\epsilon^{\mu_1\mu_2\nu_3\nu_4}\epsilon_{\rho\sigma\nu_3\nu_4}F_{\mu_1\mu_2}=-F_{\rho\sigma}. \tag{2.66}$$

不妨小结一下推导的结果，即是

$$*d*F=\big(-\partial^\mu F_{\mu\nu}\big)dx^\nu, \tag{2.67}$$

是一个1形式。

完全类似的推导可以得到，对于1形式$A=A_\mu dx^\mu$有

$$*d*A=-\partial^\mu A_\mu, \tag{2.68}$$

结果为0形式。而对于3形式$C=\frac{1}{3!}C_{\mu_1\mu_2\mu_3}dx^{\mu_1}\wedge dx^{\mu_2}\wedge dx^{\mu_3}$,

$$*d*C=\frac{1}{2!}\big(-\partial^\mu C_{\mu\rho\sigma}\big)dx^\rho\wedge dx^\sigma. \tag{2.69}$$

结果为2形式。

实际上，无需计算即可知，对于任何p形式C, $*d*C$必定为一个$p-1$形式，原因其实很简单，我们留给读者自己思索。即是说，$*d*$运算的效果刚好与外微分运算相反，外微分运算会将微分形式升高1次，而$*d*$运算则会将微分形式降低1次。

另外，也可以把霍奇对偶与外微分最漂亮的结论，即广义斯托克斯定理，结合起来。为此我们先回顾一下广义斯托克斯定理，它说的是，对于时空中任何一个$p+1$维超曲面D，其p维边界我们记为∂D，有

$$\int_{\partial D} C_p = \int_D dC_p, \tag{2.70}$$

式中C_p表示一个任意p形式。

下面取D为四维时空中的一个区域，∂D为它的三维边界，则对于任意1形式$A = A_\mu dx^\mu$我们有

$$\int_{\partial D} *A = \int_D d*A = -\int_D **d*A = \int_D \partial^\mu A_\mu * 1 = \int_D (\partial^\mu A_\mu)\Omega. \tag{2.71}$$

式中第2个等号利用了对于4形式C_4有$**C_4 = -C_4$(而$d*A$正是一个4形式)，另外，式中第3个等号是代入了(2.68)式，最后一个等号是利用了$*1 = \Omega$。不妨将最终的结果写清楚一点，即

$$\int_{\partial D} *A = \int_D (\partial^\mu A_\mu)\Omega. \tag{2.72}$$

这正是四维时空中的高斯定理。

2.3　相对性原理与经典场论

现在我们可以将相对性原理重新表述为，任何物理规律都应该在洛伦兹变换下保持不变。因此本书要考察的经典场论当然也要满足这种洛伦兹不变性，进一步，从第一章我们已经知道，经典场论的规律(也就是场方程) 可以由最小作用量原理导出，因此这就意味着经典场论的作用量泛函必须在洛伦兹变换下保持不变！这就意味着经典场论的作用量泛函必须是洛伦兹标量。

另外，对于局域场论，作用量S总可以写成拉格朗日密度$\mathcal{L}$的积分，即$S = \int d^4x\mathcal{L}$，注意到体积元d^4x等价于体积形式Ω，因此显然是洛伦兹不变的，因此S要是洛伦兹标量当且仅当拉格朗日密度$\mathcal{L}$为洛伦兹标量！

进一步，假设我们考虑的是一个标量场论，场变量记为ϕ，则从第一章可以知道$\mathcal{L} = \mathcal{L}(\phi, \dot{\phi}, \nabla\phi)$。很显然，场的时间导数$\dot{\phi}$可以和空间导数$\nabla\phi$结合成一个四维矢量$\partial_\mu\phi = (\dot{\phi}, \nabla\phi)$。所以拉氏密度实际上是$\phi$和$\partial_\mu\phi$的函数，记为$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$。从而最小作用量原理导出来的场方程就是，

$$\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Big) = \frac{\partial\mathcal{L}}{\partial\phi}. \tag{2.73}$$

而$\partial_\mu\phi$能构造出来的最简单洛伦兹标量就是

$$\partial_\mu\phi\partial^\mu\phi = \eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = -(\partial_t\phi)^2 + (\nabla\phi)^2. \tag{2.74}$$

要求动能项为正，并进一步通过将一个合适的常数吸收进场ϕ的定义之中，我们总能将$\partial_\mu\phi$对拉氏密度最简单的贡献写作

$$-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi. \tag{2.75}$$

另外，很显然，ϕ的任意函数$-\mathcal{U}(\phi)$都是洛伦兹标量，因此可以加到拉氏密度中去，进而就得到如下最简单的洛伦兹不变的拉氏密度

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \mathcal{U}(\phi) \\ &= \frac{1}{2}[(\partial_t\phi)^2 - (\nabla\phi)^2] - \mathcal{U}(\phi).\end{aligned} \tag{2.76}$$

很显然，这给出的正是前面第一章中所考察的经典场论例子。所以，第一章的例子不是完全任意写出来的，它实际上是洛伦兹不变性限制下的最简单例子！

当然，洛伦兹不变性并不能完全决定拉氏密度，比如，读者很容易发现下面的拉氏密度同样洛伦兹不变，

$$\mathcal{L} = -\frac{1}{2}g(\phi)\partial_\mu\phi\partial^\mu\phi - \mathcal{U}(\phi), \tag{2.77}$$

式中$g(\phi)$为ϕ的任意函数。这也是一种很常见的标量场模型，虽然人们对它的研究可能比上面那个更简单的模型略少。在这个模型中取$g(\phi)$为常数g_0，然后再将$\sqrt{g_0}$吸收到ϕ场的定义中去，就回到了上面那个更简单的模型。

读者可能会想为什么只用ϕ和$\partial_\mu\phi$构造拉氏密度呢？为什么不考虑二阶导数$(\partial_\mu\partial_\nu\phi)$，甚至更高阶导数呢？的确，考虑二阶导数也能轻易构造出洛伦兹不变的拉氏密度，比如

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi + f(\phi)(\partial_\mu\partial_\nu\phi)(\partial^\mu\partial^\nu\phi). \tag{2.78}$$

但实际上，人们几乎不会研究这种场论模型，原因有两个：第一，这种场论模型用最小作用量原理导出的场方程是四阶微分方程，而我们通常要求物理系统的运动微分方程为二阶微分方程。第二，可以证明，这样含高阶导数的模型导出来的哈密顿量(也就是能量)没有下界，即没有最低能量，从而物理上是不允许的，这就是所谓的Ostrogradsky 不稳定性。

前面的标量场模型很容易推广，比如说，我们可以同时考察n个标量场，记为$\phi^a, a=1,2,...,n$，这时候很容易构造出如下拉氏密度，

$$\mathcal{L}=-g_{ab}(\phi)\partial_\mu\phi^a\partial^\mu\phi^b. \tag{2.79}$$

式中$g_{ab}(\phi)$是ϕ^a的函数，实际上人们通常让它是场空间的黎曼度规。这样的场论模型就是所谓的非线性sigma模型。之所以没有在非线性sigma模型的拉氏密度中加上$-\mathcal{U}(\phi)$这样的项，是因为我们还要求了场空间的微分同胚不变性，$\mathcal{U}(\phi)$这样的项会破坏这种不变性。

前面考察的标量场ϕ都是实数值的，我们当然也可以考察复数值的标量场，不过由于作用量和拉氏密度必须是实数值的，所以这时候需要同时考虑ϕ以及它的复共轭场$\overline{\phi}$。很显然，这时候最简单的拉氏密度可以取下面的形式

$$\mathcal{L}=-\partial_\mu\overline{\phi}\partial^\mu\phi-\mathcal{U}(\overline{\phi}\phi). \tag{2.80}$$

很容易看出，除了洛伦兹不变性之外，这个拉氏密度还在下面变换下保持不变，

$$\phi\to e^{i\theta}\phi, \quad \overline{\phi}\to e^{-i\theta}\overline{\phi}. \tag{2.81}$$

式中θ为一个任意常数。

以上只考虑了四维时空中的标量场论。四维矢量场甚至高阶张量场当然也能构造相应的拉氏密度，进而得到相应的经典场论。但这时候为了得到真正有用的经典场论，往往需要在洛伦兹不变性之外进一步对系统加上更多的限制，我们还是留到后面的章节中讨论吧。