目录

第二章　狭义相对性原理以及能量-动量张量

陈童

本章主要是帮助不太熟悉狭义相对论和与之相关的经典场论的读者快速熟悉这两方面的内容。对这两方面很熟悉的读者可以跳过本章的大部分内容。但建议不要跳过最后引入能量动量张量的那一节，因为它对于本书后续章节的讲解有很关键的作用，即使你已经了解了相关内容，也请务必浏览一下最后这一节。

2.1　相对性原理

狭义相对论的基本原理是相对性原理，有时候也称之为狭义相对性原理，它说的是：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。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，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**$c=1$。

2.1.1 间隔不变性与闵可夫斯基几何

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。

很显然，狭义相对论中的间隔平方完全类似于平坦的欧几里得几何中的线元ds^2，只不过现在处理的是时空而不是单纯的空间，而且在dt^2前面多了一个负号。基于这个观察，闵可夫斯基建议将狭义相对论看成是一种平坦时空的几何，这种几何类似于描述平坦空间的欧几里得几何，通常称作闵可夫斯基几何。根据这个观点，惯性参考系之间的坐标变换就类似于欧几里得几何中的坐标系旋转，(只不过现在是时空坐标的“旋转”), 所以保持“两点距离”的平方，即间隔平方(现在也同样称作闵氏几何的线元)。

人们通常约定$t = ct = x^0 (c = 1), x = x^1, y = x^2, z = x^3$，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，我们就可以将闵氏几何的线元重写为

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

式中$\eta_{\mu\nu}$为，$-\eta_{00} = \eta_{11} = \eta_{22} = \eta_{33} = 1$, 其它指标分量都等于零。类似于欧几里得空间的度规张量$\delta_{\mu\nu}$, 我们称$\eta_{\mu\nu}$为四维闵可夫斯基时空的度规张量。值得注意的是，$\eta_{\mu\nu}$关于它的两个指标是对称的，即满足$\eta_{\mu\nu} = \eta_{\nu\mu}$。

也可以将$\eta_{\mu\nu}$看成是一个4×4矩阵的分量形式(记这个矩阵为η)

$$\eta_{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1), \tag{2.9}$$

进而引入这个矩阵的逆矩阵η^{-1}，其分量形式记为$\eta^{\mu\nu}$，

$$\eta_{\alpha\beta}\eta^{\beta\gamma} = \delta_\alpha^\gamma, \tag{2.10}$$

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

$$(\Delta s)^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu = -(\Delta t)^2 + (\Delta \mathbf{x})^2. \tag{2.11}$$

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

$$s^2 = -t^2 + \mathbf{x}^2. \tag{2.12}$$

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。

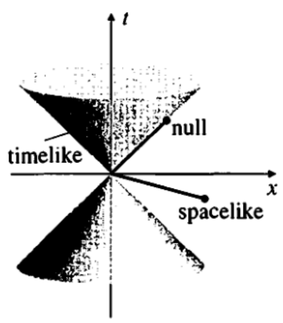

图 2.1: 时空图上原点处的光锥。

对于$s^2 < 0$情形，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。通常称上半光锥为未来光锥，称下半光锥为过去光锥。

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

$$\eta_{\mu\nu} dx^\mu dx^\nu = -d\tau^2. \tag{2.13}$$

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。

利用固有时，我们可以定义粒子的四维速度u^μ

$$u^\mu = \frac{dx^\mu}{d\tau}. \tag{2.14}$$

根据(2.13)式，显然有

$$\eta_{\mu\nu} u^\mu u^\nu = -1. \tag{2.15}$$

我们称u^μ为一个类时矢量。即是说，假如把u^μ的起点画在光锥原点的话，u^μ将完全躺在光锥之内。这就说明，一切从光锥原点经过的粒子，其在时空中的运动轨迹都将在这一点的光锥之内。当然，光子例外，光子的运动轨迹躺在光锥面上，因为光子的固有时总是零！

假设粒子的速度为$\mathbf{v}$，即$d\mathbf{x} = \mathbf{v} dt$，则根据(2.13)式即有

$$d\tau^2 = dt^2 - d\mathbf{x}^2 = dt^2(1 - \mathbf{v}^2) \Rightarrow dt = \frac{d\tau}{\sqrt{1 - \mathbf{v}^2}}. \tag{2.16}$$

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。

光锥与世界线

在狭义相对论中，由于类空相间事件的先后顺序是相对的，因此两个不同地的事件是否同时发生也是相对的，依赖于参考系，这就是所谓的同时的相对性。在牛顿力学中，我们可以绝对地将时空用等时的空间超曲面分割，因为同时的相对性，在狭义相对论中却无法做到这一点。但这并不说明狭义相对论的时空完全没有结构。相反，在任何一个时空点，我们都可以定义一个光锥，如图(2.2)所示，它就是经过这点的光线在时空中的运动轨迹之集合。由于在狭义相对论中，一切粒子的运动速度都不能超过光速，所以经过任何时空点的粒子，其在四维时空中的运动方向都只能在这点的光锥之内，最多是在光锥面上(对于光子)，而决定不能在光锥之外。

我们把粒子在时空中的运动轨迹称作粒子的世界线，记为$x^\mu(\sigma)$, 很多时候人们会选择粒子的固有时τ作为世界线的参数，即取$\sigma = \tau$。根据上一段所说可知，粒子的世界线只能从其上各点的光锥内部穿过，如图(2.2)所示。

前面我们说过运动的钟会变慢，同时我们也说过运动是相对的。那么，假设有两只钟，一只为Bob所持有，而Bob静止不动，另一只为Alice持有，它跟着Alice沿着闭合路径运动一圈再回到起点与Bob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。

其实，这两只钟分别测量的是Bob和Alice的固有时，而Bob与Alice在时空中的世界线如图(2.3)所示。对于有限长世界线，固有时T是$d\tau$沿着世

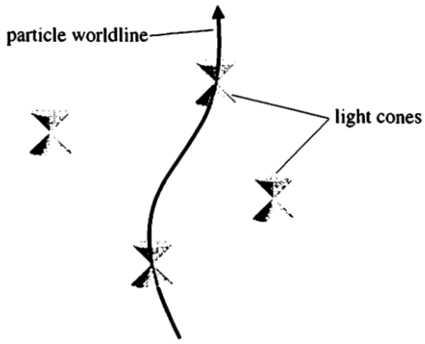

图 2.2: 时空图上的光锥和世界线。

界线的积分

$$T = \int d\tau = \int \sqrt{dt^2 - d\mathbf{x}^2}. \tag{2.17}$$

注意这个积分中的时间微元贡献和空间微元贡献是相减的关系，Bob世界线的空间完全固定，而Alice的世界线在空间上扫过，由于空间的贡献是要减去的，所以很显然Bob的固有时比较长，而Alice的固有时比较短，所以Alice的钟慢！

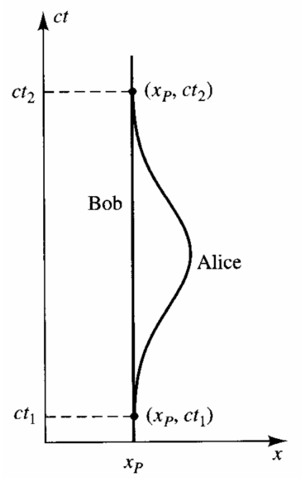

图 2.3: Bob与Alice在时空中的世界线。

这就是闵可夫斯基几何与欧几里得几何不同的地方，在欧几里得几何

中，两点之间直线总是最短的。但是在闵可夫斯基几何中，对于运动粒子的世界线，两点之间直线的固有时反而是最长的，而曲线由于要减去更多的空间贡献，固有时反而会比较短。

2.1.2　庞加莱群

狭义相对论中，两个惯性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

$$dx'^{\mu} = \Lambda^{\mu}{}_{\nu} dx^{\nu}, \tag{2.18}$$

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

$$\eta_{\mu\nu} dx'^{\mu} dx'^{\nu} = \eta_{\mu\nu} \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} dx^{\alpha} dx^{\beta} = \eta_{\alpha\beta} dx^{\alpha} dx^{\beta}, \tag{2.19}$$

从而即有

$$\eta_{\mu\nu} \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} = \eta_{\alpha\beta}. \tag{2.20}$$

我们当然可以将(2.20)式写成矩阵形式，即

$$\Lambda^T \eta \Lambda = \eta. \tag{2.21}$$

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(2.21)式两边求行列式，并注意到$\det(\eta) = -1$, 从而即可得

$$[\det(\Lambda)]^2 = 1 \Rightarrow \det(\Lambda) = \pm 1. \tag{2.22}$$

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(2.20)式中取$\alpha = \beta = 0$, 即可得

$$(\Lambda^{0}{}_{0})^2 = 1 + \sum_{i=1,2,3} (\Lambda^{i}{}_{0})^2 \geq 1 \;\Rightarrow\; \Lambda^{0}{}_{0} \geq 1 \text{ or } \Lambda^{0}{}_{0} \leq -1. \tag{2.23}$$

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

$$\det(\Lambda) = 1, \quad \Lambda^0_{\ 0} \geq 1. \tag{2.24}$$

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$SO^+(1,3)$。

除了惯性参考系之间的洛伦兹变换之外，很显然，时空坐标的平移

$$x^\mu \to x'^\mu = x^\mu + a^\mu, \tag{2.25}$$

这也同样保持事件间的间隔不变。洛伦兹变换再加上时空平移就构成一个比洛伦兹群更大的群，称作庞加莱群。由于四维时空闵可夫斯基几何的ds^2在庞加莱群变换下保持不变，所以它是闵可夫斯基时空的对称群。

2.1.3　四维闵可夫斯基时空的矢量和张量

上一小节说过，在不同的惯性参考系中，时空坐标按照下面的洛伦兹变换而变换，

$$dx'^\mu = \Lambda^\mu_{\ \nu} dx^\nu, \tag{2.26}$$

式中变换矩阵Λ满足

$$\Lambda^T \eta \Lambda = \eta \ \Leftrightarrow \ \Lambda^{-1} = \eta^{-1} \Lambda^T \eta. \tag{2.27}$$

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

$$A'^\mu = \Lambda^\mu_{\ \nu} A^\nu, \tag{2.28}$$

则我们称A^μ为一个四维闵可夫斯基时空的矢量，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

$$A^2 = \eta_{\mu\nu} A^\mu A^\nu = -(A^0)^2 + \mathbf{A}^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2. \tag{2.29}$$

很明显，A^2在洛伦兹变换下是不变的。假如$A^2 < 0$，我们就称A^μ为一个类时矢量，假如$A^2 = 0$，我们就称之为一个类光矢量或者零性矢量，假如$A^2 > 0$，就称之为一个类空矢量。

我们也可以定义下指标的四分量量A_μ为，

$$A_\mu = \eta_{\mu\nu} A^\nu, \tag{2.30}$$

写得更清楚一点就是

$$A_\mu = (A_0, A_1, A_2, A_3) = (-A^0, A^1, A^2, A^3) = (A_0, \mathbf{A}). \tag{2.31}$$

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

$$A'_\mu A'^\mu = A_\mu A^\mu. \tag{2.32}$$

注意到A^μ在洛伦兹变换下按照(2.28)式变换，因此上式就意味着A_μ必然按照下式变换

$$A'_\mu = (\Lambda^{-1})^\nu{}_\mu A_\nu. \tag{2.33}$$

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

$$A^\mu = \eta^{\mu\nu} A_\nu. \tag{2.34}$$

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

$$\partial'_\mu = (\Lambda^{-1})^\nu{}_\mu \partial_\nu. \tag{2.35}$$

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，有时候也称作矢量A和B的缩并。两个四矢量的内积(缩并)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

$$B'^{\mu\nu} = \Lambda^{\mu}_{\ \alpha}\Lambda^{\nu}_{\ \beta}B^{\alpha\beta}, \tag{2.36}$$

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，

$$B'_{\mu\nu} = (\Lambda^{-1})^{\alpha}_{\ \mu}(\Lambda^{-1})^{\beta}_{\ \nu}B_{\alpha\beta}. \tag{2.37}$$

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，

$$B'^{\mu}_{\nu} = \Lambda^{\mu}_{\ \alpha}(\Lambda^{-1})^{\beta}_{\ \nu}B^{\alpha}_{\beta}. \tag{2.38}$$

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，这同样也叫做张量的缩并。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。$B_{\mu\nu}(x)$在洛伦兹变换下按照下式变

$$B_{\mu\nu}(x) \to B'_{\mu\nu}(x') = (\Lambda^{-1})^{\alpha}{}_{\mu}(\Lambda^{-1})^{\beta}{}_{\nu}B_{\alpha\beta}(x). \tag{2.39}$$

其它四维张量场的变换规则可以类似地推广。特别的，对于标量场$\Phi(x)$，我们有

$$\Phi(x) \to \Phi'(x') = \Phi(x). \tag{2.40}$$

2.2　狭义相对性原理与经典物理

2.2.1　狭义相对性原理与经典场论

现在我们可以将狭义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。我们知道，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$S=\int d^4x\mathcal{L}$，注意到由于$\det(\Lambda)=1$，因此体积元d^4x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！

标量场

进一步，假设我们考虑的是一个标量场论，场变量记为ϕ，则拉氏密度实际上是ϕ和$\partial_\mu\phi$的函数(这里$\partial_\mu=\frac{\partial}{\partial x^\mu}$)，记为$\mathcal{L}=\mathcal{L}(\phi,\partial_\mu\phi)$。根据最小作用量原理，我们可以进行如下推导

$$\begin{aligned}
\delta S[\phi(x)] &= \int d^4x\delta\mathcal{L}(\phi,\partial_\mu\phi)\\
&= \int d^4x\Big[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi)\Big] = \int d^4x\Big[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi)\Big]\\
&= \int d^4x\Big[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Big)\Big]\delta\phi + \int d^4x\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big).
\end{aligned} \tag{2.41}$$

很显然，最后一行的第二项是全微分项，因此积分结果完全取决于边界项，为

$$\int d^4x\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big)=\int dS_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\Big)|_\infty. \tag{2.42}$$

式中dS_μ表示四维时空无穷远三维边界的体积元。总之，结果仅仅在四维时空的无穷远边界上有贡献。但是，最小作用量原理要求在时空的边界上场位形是固定的，从而$\delta\phi=0$，因此这一项的最终结果其实等于零。

从而即有

$$\delta S[\phi(x)]=\int d^4x\Big[\frac{\partial\mathcal{L}}{\partial\phi}-\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Big)\Big]\delta\phi, \tag{2.43}$$

进一步应用最小作用量原理$\delta S[\phi(x)]=0$，即可得到场方程，

$$\frac{\delta S}{\delta\phi(x)}=\frac{\partial\mathcal{L}}{\partial\phi}-\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\Big)=0. \tag{2.44}$$

而$\partial_\mu\phi$能构造出来的最简单洛伦兹标量就是

$$\partial_\mu\phi\partial^\mu\phi=\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi=-(\partial_t\phi)^2+(\nabla\phi)^2. \tag{2.45}$$

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

$$-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi. \tag{2.46}$$

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

$$\mathcal{L}=-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\mathcal{U}(\phi). \tag{2.47}$$

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，

$$\mathcal{L}=-\frac{1}{2}g(\phi)\partial_\mu\phi\partial^\mu\phi-\mathcal{U}(\phi), \tag{2.48}$$

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

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi + f(\phi)(\partial_\mu\partial_\nu\phi)(\partial^\mu\partial^\nu\phi). \tag{2.49}$$

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，

$$\mathcal{L} = -g_{ab}(\phi)\partial_\mu\phi^a\partial^\mu\phi^b. \tag{2.50}$$

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

$$\mathcal{L} = -\partial_\mu\overline{\phi}\partial^\mu\phi - \mathcal{U}(\overline{\phi}\phi). \tag{2.51}$$

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，

$$\phi \to e^{i\theta}\phi, \quad \overline{\phi} \to e^{-i\theta}\overline{\phi}. \tag{2.52}$$

式中θ为一个任意常数。

电磁场

以上只考虑了四维闵可夫斯基时空中的标量场论。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。比方说，对于电磁场四矢量$A^\mu = (\phi, \mathbf{A})$，我们还要加上规范对称性，即要求物理可观测量和作用量在如下规范变换下保持不变

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \varepsilon(x). \tag{2.53}$$

式中$\varepsilon(x)$为任意函数。这时候系统的作用量S_g可以写成，

$$S_g = -\int d^4x \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{2.54}$$

式中$F_{\mu\nu}$称作规范场强，它是一个二阶反对称张量场，定义为

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.55}$$

很容易验证，$F_{\mu\nu}$在规范变换下保持不变。

将上面这个作用量对A_μ变分，可得

$$\begin{aligned} \delta S_g &= -\int d^4x \frac{1}{2} F^{\mu\nu} \delta F_{\mu\nu} = -\int d^4x F^{\mu\nu} \partial_\mu \delta A_\nu \\ &= \int d^4x \partial_\mu F^{\mu\nu} \delta A_\nu. \end{aligned} \tag{2.56}$$

为了得出上式最后一行的结果，我们需要分部积分，并丢弃边界项(因为假设无穷远边界上场的变分等于零)。

人们也可以将电磁场与电流四矢量$J^\mu = (\rho, \mathbf{J})$相耦合，即给电磁场作用量加上一项$S_I$, 它满足

$$\delta S_I = \int J^\mu \delta A_\mu. \tag{2.57}$$

相应的，根据最小作用量原理$0 = \delta S = \delta S_g + \delta S_I$，即可得规范场$A_\mu$满足的运动微分方程，

$$-\partial_\mu F^{\mu\nu} = J^\nu. \tag{2.58}$$

另外，根据$F_{\mu\nu}$的定义(2.55)，很容易验证它还必然满足如下比安奇恒等式

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0. \tag{2.59}$$

(2.58)式和(2.59)式一起就构成麦克斯韦方程组的四维协变形式，更多的讨论可以参看我的《经典场论新讲》。

2.2.2　狭义相对性原理与经典力学

如果我们要考察的不是一个场论系统，而是一个粒子，那狭义相对性原理告诉我们，相对论粒子的作用量也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(由于$E=mc^2$,而$c=1$)，从而我们知道，相对论粒子的作用量必定可以写成

$$S[x(\sigma)]=-m\int d\tau=-m\int d\sigma\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}. \tag{2.60}$$

式中σ为粒子世界线的参数，第一个式子的负号是为了使得作用量有极小值(因为固有时有极大值，直线的固有时最长)。

值得指出的是，上述作用量(2.60)显然具有重参数不变性，即在$\sigma\to\widetilde{\sigma}(\sigma)$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世界线参数σ。最常见的选择有两种，第一种是，取$\sigma=\tau$，即取固有时本身为路径的参数。第二种选择是取$\sigma=x^0=t$，即取通常的时间坐标为参数，这时候注意到

$$d\tau^2=dt^2-d\mathbf{x}^2=dt^2(1-\mathbf{v}^2). \tag{2.61}$$

式中$\mathbf{v}=\frac{d\mathbf{x}}{dt}$是粒子的速度。从而就可以将作用量(2.60)写成

$$S[\mathbf{x}(t)]=-m\int dt\sqrt{1-\mathbf{v}^2}. \tag{2.62}$$

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(2.62)近似成

$$S=-m\int dt+\int dt\frac{1}{2}m\mathbf{v}^2+... \tag{2.63}$$

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!

为了导出相对论粒子在闵可夫斯基时空中的运动方程，我们记

$$L=\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}, \tag{2.64}$$

因此当取固有时本身为世界线参数时，即$\sigma=\tau$时，我们有$L=1$，即

$$-\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}=1. \tag{2.65}$$

现在，我们可以把相对论粒子的作用量写成$S[x(\sigma)]=-m\int d\sigma L$, 为了导出粒子的运动方程，我们需要计算变分$\delta S[x(\sigma)]=-m\int d\sigma\delta L$。利用

$$\begin{aligned}L\delta L=\frac{1}{2}\delta(L^2)&=-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\delta(\frac{dx^\nu}{d\sigma})\\&=-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{d(\delta x^\nu)}{d\sigma}.\end{aligned} \tag{2.66}$$

可以得到

$$\begin{aligned}\delta S[x(\sigma)]&=-m\int d\sigma\delta L=m\int d\sigma\frac{1}{L}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{d(\delta x^\nu)}{d\sigma}\\&=m\int d\sigma\frac{d}{d\sigma}(\frac{1}{L}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\delta x^\nu)-m\int d\sigma\frac{d}{d\sigma}(\frac{1}{L}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma})\delta x^\nu\\&=-m\int d\sigma\frac{d}{d\sigma}(\frac{1}{L}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma})\delta x^\nu.\end{aligned} \tag{2.67}$$

式中最后一个等于号是利用了路径两端是固定的(即在两端$\delta x^\nu=0$)，从而全微分项的积分结果为零。进而根据最小作用量原理$\delta S[x(\sigma)]=0$，就可以得到运动微分方程

$$\frac{\delta S}{\delta x^\nu(\sigma)}=-m\frac{d}{d\sigma}(\frac{1}{L}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma})=0. \tag{2.68}$$

我们可以通过取$\sigma=\tau$来简化这个运动方程，这时候$L=1$，从而即有

$$\frac{d^2x^\mu}{d\tau^2}=0. \tag{2.69}$$

(2.69)式就是自由的相对论性粒子在时空中的运动微分方程，这个方程的解显然是

$$x^\mu(\tau)=u^\mu\tau+a^\mu, \tag{2.70}$$

式中u^μ和a^μ均是常矢量，u^μ当然就是粒子的四维速度，而且(2.65)式告诉我们

$$\eta_{\mu\nu}u^\mu u^\nu=-1. \tag{2.71}$$

也即是说，相对论性自由粒子在闵可夫斯基时空中是作匀速直线运动！也可以定义粒子的四维动量p^μ，为

$$p^\mu = mu^\mu = m\frac{dx^\mu}{d\tau}. \tag{2.72}$$

如果允许引入一个辅助性的力学变量$e(\sigma)$，那我们还可以将作用量(2.60)写成一个更加顺眼的形式，

$$S[e(\sigma), x(\sigma)] = \frac{1}{2}\int d\sigma\Big(e^{-1}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma} - em^2\Big). \tag{2.73}$$

为了证明这个作用量与前面的(2.60)相等价，人们只要先对辅助力学变量$e(\sigma)$使用最小作用量原理即可，即先利用下式求出$e(\sigma)$，再代入上面作用量中消去$e(\sigma)$

$$\frac{\delta S}{\delta e(\sigma)} = 0. \tag{2.74}$$

特别的，(2.73)允许我们取$m=0$的极限，从而就得到零质量粒子(比如光子)的经典作用量，为

$$S[e(\sigma), x(\sigma)] = \frac{1}{2}\int d\sigma\Big(e^{-1}\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}\Big). \tag{2.75}$$

这是引入辅助变量$e(\sigma)$之后的一个意外好处，即允许我们用作用量原理统一描述有质量粒子和零质量的粒子。

上面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？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。

仔细思考以后我们可能会发现下面这一项

$$+q\int_a^b A_\mu dx^\mu. \tag{2.76}$$

它就是电磁势沿着带电粒子世界线的积分, 式中我们假设这条世界线起于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(2.53)的作用下，它会变为

$$q\int_a^b A_\mu dx^\mu \to q\int_a^b A_\mu dx^\mu + q\int_a^b d\varepsilon \tag{2.77}$$

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

$$q\int_a^b d\varepsilon = q\big(\varepsilon(b)-\varepsilon(a)\big). \tag{2.78}$$

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(2.76)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，(2.76)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

$$S[x(\sigma)] = -m\int d\sigma\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}} + q\int d\sigma A_\mu\frac{dx^\mu}{d\sigma}. \tag{2.79}$$

如果将参数σ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

$$S[\mathbf{x}(t)] = -m\int dt + \int dt\Big[\frac{1}{2}m\mathbf{v}^2 - q\phi + q\mathbf{A}\cdot\mathbf{v}\Big] + ... \tag{2.80}$$

将作用量(2.79)对规范势A_μ变分(为了避免混淆，下面将带电粒子在时空中的位置坐标改记为$x_e^\mu(\sigma)$)，即可得到

$$\begin{aligned}\delta S &= q\int d\sigma\frac{dx_e^\mu}{d\sigma}\delta A_\mu\\ &= \int d^4x\Big[q\int d\sigma\frac{dx_e^\mu}{d\sigma}\delta^4\big(x-x_e(\sigma)\big)\Big]\delta A_\mu.\end{aligned} \tag{2.81}$$

将这个式子与标准的电流四矢量J^μ与规范势的耦合$\delta S = \int d^4x J^\mu\delta A_\mu$进行比较，即可得到此带电粒子系统的电流四矢量，为

$$J^\mu(x) = q\int d\sigma\frac{dx_e^\mu}{d\sigma}\delta^4\big(x-x_e(\sigma)\big). \tag{2.82}$$

为了看清楚(2.82)式的物理意义，我们取世界线参数$\sigma = x_e^0$, 并做出(2.82)式中的积分，则有

$$\begin{aligned}\rho &= q\delta^3(\mathbf{x}-\mathbf{x}_e(t))\\ \mathbf{J} &= q\mathbf{v}_e\delta^3(\mathbf{x}-\mathbf{x}_e(t)).\end{aligned} \tag{2.83}$$

很显然，结果正符合我们对电荷密度以及电流密度表达式的预期。

2.3 时空对称性与能量动量张量

这一节我们将引入物质系统的一个特征性物理量，即所谓的能量-动量张量。根据著名的诺特定理，任何一个连续的对称性都必然有一个相应的守恒流四矢量，能量-动量张量就是与时空平移对称性相对应的守恒流，由于时空平移本身是一个四维矢量，所以它对应的守恒流就是一个二阶张量。

2.3.1 场论系统的能量-动量张量

标量场

为了简单起见，我们首先讲述标量场系统的能量-动量张量。

假设有一个标量场$\phi(x)$，其拉格朗日密度可以写成$\mathcal{L}(\phi,\partial_\mu\phi)$，特别的，这个拉格朗日密度不显含时空坐标$x$。现在，假定将整个场论系统进行一个时空平移，使得$x$点的场平移到$x'$点，

$$x^\mu \to x'^\mu = x^\mu + a^\mu, \tag{2.84}$$

其中a^μ为某常数四矢量。记平移之后的场为$\phi'(x)$，很显然ϕ'在x'点的场值来自于平移之前ϕ在x点的场值，即

$$\phi'(x') = \phi(x). \tag{2.85}$$

很容易验证作用量$S[\phi] = \int d^4x\mathcal{L}(\phi,\partial_\mu\phi)$在此平移之下保持不变，具体验证过程如下

$$\begin{aligned}
S[\phi'] &= \int d^4x\mathcal{L}(\phi'(x),\partial_\mu\phi'(x)) \\
&= \int d^4x'\mathcal{L}(\phi'(x'),\partial'_\mu\phi'(x')) = \int d^4x'\mathcal{L}(\phi(x),\partial'_\mu\phi(x)) \\
&= \int d^4x\mathcal{L}(\phi(x),\partial_\mu\phi(x)) = S[\phi]
\end{aligned} \tag{2.86}$$

式中第二个等号只是将x变量重写成了x'，倒数第二个等号是用了$d^4x' = d^4x$以及$\partial'_\mu = \partial_\mu$。

为了通过诺特定理引入相应的守恒流，下面考察无穷小时空平移，即将a^μ取成无穷小量ϵ^μ。进一步，我们使用一个**关键技巧**，即将ϵ^μ 变成依赖于时空坐标x的无穷小量$\epsilon^\mu(x)$, 即考察如下无穷小时空平移

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x). \tag{2.87}$$

当然我们依然有

$$\phi'(x') = \phi(x). \tag{2.88}$$

但是，在这种依赖于时空点x的局域平移之下，作用量当然无法保持不变，因为这种局域平移根本不是系统的对称性。

很显然，在一阶近似上有

$$\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \delta^{\mu}_{\nu} + \partial_{\nu}\epsilon^{\mu}, \quad \frac{\partial x^{\nu}}{\partial x'^{\mu}} = \delta^{\nu}_{\mu} - \partial_{\mu}\epsilon^{\nu}. \tag{2.89}$$

另外，记坐标变换的雅可比行列式为$|\frac{\partial x'}{\partial x}|$, 利用矩阵恒等式$\det(A) = e^{\mathrm{Tr}(\ln A)}$, 易得在一阶近似上有

$$|\frac{\partial x'}{\partial x}| = 1 + \partial_{\mu}\epsilon^{\mu}(x). \tag{2.90}$$

下面我们来计算在局域平移(2.87)之前和之后，系统作用量的改变量。平移以后的作用量$S[\phi'(x)]$为

$$\begin{aligned}
S[\phi'(x)] &= \int d^4x \mathcal{L}(\phi'(x), \partial_{\nu}\phi'(x)) \\
&= \int d^4x' \mathcal{L}(\phi'(x'), \partial'_{\mu}\phi'(x')) = \int d^4x' \mathcal{L}\big(\phi(x), \partial'_{\mu}\phi(x)\big) \\
&= \int d^4x |\frac{\partial x'}{\partial x}| \mathcal{L}\big(\phi(x), \frac{\partial x^{\nu}}{\partial x'^{\mu}}\partial_{\nu}\phi(x)\big) \\
&= \int d^4x (1 + \partial_{\mu}\epsilon^{\mu}(x)) \mathcal{L}\big(\phi(x), \partial_{\mu}\phi(x) - \partial_{\mu}\epsilon^{\nu}\partial_{\nu}\phi(x)\big) \\
&= \int d^4x \Big[-\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\phi)}\partial_{\nu}\phi(x) + \delta^{\mu}_{\nu}\mathcal{L}\big(\phi(x), \partial_{\rho}\phi(x)\big)\Big]\partial_{\mu}\epsilon^{\nu} \\
&\quad + \int d^4x \mathcal{L}\big(\phi(x), \partial_{\nu}\phi(x)\big).
\end{aligned} \tag{2.91}$$

由上面的推导易知，变换前后作用量的改变量为

$$\begin{aligned}
\delta S &= S[\phi'(x)] - S[\phi(x)] \\
&= -\int d^4x \Big[\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\phi)}\partial_{\nu}\phi(x) - \delta^{\mu}_{\nu}\mathcal{L}\Big]\partial_{\mu}\epsilon^{\nu} \\
&= \int d^4x T^{\mu}{}_{\nu}\partial_{\mu}\epsilon^{\nu} = \int d^4x T_{\mu\nu}\partial^{\mu}\epsilon^{\nu}.
\end{aligned} \tag{2.92}$$

式中

$$T^{\mu\nu} = -\Big[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi(x) - \eta^{\mu\nu}\mathcal{L}\Big]. \tag{2.93}$$

注意(2.92)式，当ϵ^ν是一个不依赖于时空点的整体无穷小平移时，即有$\delta S=0$，这当然是因为这种整体时空平移是系统的对称性。正因为如此，不仅对于标量场系统可以找到这样一个$T^{\mu\nu}$，实际上对闵可夫斯基时空的任何物质系统都能找到一个类似的$T^{\mu\nu}$，因为任何物质系统的作用量在局域平移(2.87) 前后的改变量都必定能写成如下形式

$$\delta S = \int d^4x T_{\mu\nu}\partial^\mu\epsilon^\nu. \tag{2.94}$$

而这又是因为，当将ϵ^ν取成常矢量(即考察无穷小整体时空平移时)时，由于时空平移对称性，必有$\delta S=0$，所以在局域化平移情形，δS只能依赖于ϵ^ν的偏导$\partial^\mu\epsilon^\nu$，从而δS必定能写成(2.94)式的形式！

下面进一步假定$\epsilon^\nu(x)$在时空的无穷远边界上都趋于零。另外，上面的(2.94)式对于任意场位形都成立，下面我们考察真实的满足场运动微分方程的场位形，那这时候由于这些场位形满足最小作用量原理，当然就有$\delta S=0$对于任意的在时空无穷远处趋于零的场变分都成立，那当然也对局域时空平移(2.87)引起的场改变成立。即是说，对于真实场位形，必有

$$0 = \delta S = \int d^4x T_{\mu\nu}\partial^\mu\epsilon^\nu. \tag{2.95}$$

将上式分部积分，即有

$$0 = \int d^4x T_{\mu\nu}\partial^\mu\epsilon^\nu = -\int d^4x \partial^\mu T_{\mu\nu}\epsilon^\nu(x). \tag{2.96}$$

由于$\epsilon^\nu(x)$为任意无穷小函数，从而即有

$$\partial^\mu T_{\mu\nu} = 0. \tag{2.97}$$

即是说，$T_{\mu\nu}$就是与时空平移对称性相应的守恒流！特别的，如果取$\epsilon^\mu=\epsilon\delta^{\mu 0}$, 即假设考察的是时间平移，那相应的守恒流$T^{\mu 0}$当然就是能量流四矢量，换言之，$T^{00}$必定为能量密度，$T^{i0}(i=1,2,3)$为能量流密度。而如果取$\epsilon^\mu=\epsilon\delta^{\mu j}$为空间平移, 那相应的守恒流$T^{\mu j}$当然就是动量流四矢量，即是说$T^{0j}$必为动量密度，而$T^{ij}$为动量流密度。将所有这些合起来，我们就

称$T_{\mu\nu}$为能量动量张量！上面的讨论也说明，**闵可夫斯基时空的任何物质系统都存在这么一个守恒的能量动量张量。**

比如，假设我们考虑$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \mathcal{U}(\phi)$的场论模型，则容易算得，

$$\begin{aligned} T^{\mu\nu} &= \partial^\mu\phi\partial^\nu\phi + \eta^{\mu\nu}\mathcal{L} \\ &= \partial^\mu\phi\partial^\nu\phi - \frac{1}{2}\eta^{\mu\nu}\partial_\rho\phi\partial^\rho\phi - \eta^{\mu\nu}\mathcal{U}(\phi). \end{aligned} \tag{2.98}$$

很明显，这个$T^{\mu\nu}$的两个指标是对称的。实际上，可以证明，对于任何洛伦兹不变的标量场论，其$T^{\mu\nu}$都必定是对称张量。但是当我们的考察范围超出标量场论时，其按照上面类似的办法求出来的$T^{\mu\nu}$就不一定为对称张量了，比方说对于矢量场，它的$T^{\mu\nu}$就不是对称张量。不过，$T^{\mu\nu}$的流守恒方程$\partial_\mu T^{\mu\nu} = 0$告诉我们，$T^{\mu\nu}$的定义不是唯一的，实际上人们很容易看出，对于任何$X^{\rho\mu\nu}$，只要$X^{\rho\mu\nu} = -X^{\mu\rho\nu}$，则$T^{\mu\nu} + \partial_\rho X^{\rho\mu\nu}$同样满足流守恒方程，因此可以定义为新的能动量张量。从而只要我们合适地选取$X^{\rho\mu\nu}$, 我们总可以让重新定义以后的能动量张量为一个对称张量。实际上，马上我们就会给出一个找到这样一个对称能量动量张量的巧妙办法。

洛伦兹对称性

我们考察的经典场论都是具有洛伦兹不变性的场论，即拉格朗日密度为洛伦兹标量的理论，具体来说即是拉氏密度在如下坐标变换下保持不变的理论，

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{2.99}$$

$\Lambda^\mu{}_\nu$就是所谓的洛伦兹变换。那么，洛伦兹变换对应的守恒流是什么呢?

为此我们需要考察无穷小洛伦兹变换，即取

$$\Lambda^\mu{}_\nu = \delta^\mu_\nu + \epsilon^\mu{}_\nu, \tag{2.100}$$

式中$\epsilon^\mu{}_\nu$为无穷小量。很显然，无穷小洛伦兹变换由于可以和恒等变换连续过渡，从而必定是正洛伦兹变换，即满足[1]

$$\det(\Lambda) = 1 \Rightarrow \epsilon^\mu{}_\mu = 0. \tag{2.101}$$

进一步，利用洛伦兹变换的定义$\eta_{\mu\nu}\Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta = \eta_{\alpha\beta}$, 易得

$$\eta_{\mu\nu}(\delta^\mu_\alpha + \epsilon^\mu{}_\alpha)(\delta^\nu_\beta + \epsilon^\nu{}_\beta) = \eta_{\alpha\beta} \Rightarrow \epsilon_{\alpha\beta} + \epsilon_{\beta\alpha} = 0. \tag{2.102}$$

[1]利用矩阵恒等式$\det(A) = e^{\mathrm{Tr}(\ln A)}$.

即$\epsilon^{\mu\nu}$是一个二阶反对称张量。

注意到无穷小洛伦兹变换为$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu{}_\nu x^\nu$。为了利用诺特定理，我们使用关键技巧，即将无穷小参数$\epsilon^\mu{}_\nu$变成依赖于时空坐标的$\epsilon^\mu{}_\nu(x)$，进而考察如下无穷小局域时空坐标变换，

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x), \quad \text{这里} \quad \epsilon^\mu(x) = \epsilon^\mu{}_\nu(x) x^\nu. \tag{2.103}$$

则完全类似于前面对(2.94)式的论证，可知在此时空变换之下作用量的改变量必定可以写成

$$\begin{aligned} \delta S &= \int d^4x T^\mu{}_\nu \partial_\mu \epsilon^\nu(x) \\ &= \int d^4x T^\mu{}_\nu \big(\epsilon^\nu{}_\mu + x^\rho \partial_\mu \epsilon^\nu{}_\rho\big). \end{aligned} \tag{2.104}$$

(注意，对于比方说矢量场，它的各指标分量也要按照无穷小局域洛伦兹变换$\Lambda^\mu{}_\nu(x) = \delta^\mu_\nu + \epsilon^\mu{}_\nu(x)$的规则变，这是不同于前面时空平移情形的。在时空平移情形，矢量场的分量指标并不变换。因此上式中的$T_{\mu\nu}$一般不同于前面按照时空平移直接求出来的$T_{\mu\nu}$。) 注意到$\epsilon_{\mu\nu}$关于指标反对称，从而有

$$\begin{aligned} \delta S &= \frac{1}{2}\int d^4x \Big[(T^{\mu\nu} - T^{\nu\mu})\epsilon_{\nu\mu} + (x^\rho T^{\mu\nu} - x^\nu T^{\mu\rho})\partial_\mu \epsilon_{\nu\rho}\Big] \\ &= -\frac{1}{2}\int d^4x \Big[(T^{\mu\nu} - T^{\nu\mu})\epsilon_{\mu\nu}\Big] - \frac{1}{2}\int d^4x \Big[(x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu})\partial_\rho \epsilon_{\mu\nu}\Big]. \end{aligned} \tag{2.105}$$

如果我们将$\epsilon_{\mu\nu}$取回常数，则$\partial_\rho \epsilon_{\mu\nu} = 0$, 从而上式最后一行只剩下前面那项，但是洛伦兹不变性告诉我们，当$\epsilon_{\mu\nu}$为常数时作用量应该不变，即这时候必有$\delta S = 0$，由此可知

$$T^{\mu\nu} - T^{\nu\mu} = 0, \tag{2.106}$$

即这样求出来的能动量张量必定是对称张量！这就是我们给出的如何寻找对称能量动量张量的办法。

在(2.105)式中代入$T^{\mu\nu} = T^{\nu\mu}$，即有

$$\delta S = -\frac{1}{2}\int d^4x \Big[(x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu})\partial_\rho \epsilon_{\mu\nu}\Big]. \tag{2.107}$$

完全类似于前面导出能动量张量守恒的讨论，这个结果意味着$M^{\rho\mu\nu} = x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu}$为守恒流，满足守恒方程

$$\partial_\rho M^{\rho\mu\nu} = 0. \tag{2.108}$$

综合关于无穷小局域时空平移和无穷小局域洛伦兹变换的结果，可知，在如下无穷小局域时空变换之下

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x), \tag{2.109}$$

任何物质系统作用量的改变量必定可以写成

$$\delta S = \int d^4x T^{\mu\nu}\partial_\mu\epsilon_\nu(x) = \frac{1}{2}\int d^4x T^{\mu\nu}\big(\partial_\mu\epsilon_\nu + \partial_\nu\epsilon_\mu\big). \tag{2.110}$$

其中我们利用了$T^{\mu\nu}$是一个对称张量。这个式子是一个很**关键的式子**，在后面的章节中有大用。

电磁场

下面我们来考察一个自由电磁场的能量-动量张量。所谓的自由电磁场就是不与其它任何东西耦合，单独只由拉氏密度$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$描述的无源电磁场。当然，我们可以按照前面讨论无穷小局域洛伦兹变换中给出的办法直接找出对称的能量-动量张量。但更快的方法实际上是下面的办法。

我们已经知道标量场能动量张量的公式，为了得到电磁场的能动量张量，我们将$A_\mu(\mu = 0,1,2,3)$的每一个分量看成一个实标量场，从而容易给出电磁场能动量张量的公式

$$T^{\mu\nu} = -\Big[\frac{\partial\mathcal{L}}{\partial(\partial_\mu A^\rho)}\partial^\nu A^\rho - \eta^{\mu\nu}\mathcal{L}\Big]. \tag{2.111}$$

代入自由电磁场的拉氏密度可以得到，

$$T^{\mu\nu} = F^\mu{}_\rho\partial^\nu A^\rho - \frac{1}{4}\eta^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}. \tag{2.112}$$

很显然，这个能动量张量不是一个对称张量，而更为严重的问题是，它不是规范不变的！而作为物理可观测量，能动量张量必须规范不变！好在，正如前面所说的，我们可以通过给它加上一个合适的$\partial_\rho X^{\rho\mu\nu}$来重新定义能动量张量，使它成为一个对称张量。具体来说，我们可以给上述能动量张量加上$\partial^\rho(-F^\mu{}_\rho A^\nu) = -F^\mu{}_\rho\partial^\rho A^\nu$ (注意到对于自由电磁场，电流四矢量等于零，从而方程(2.58)变成$\partial^\rho F^\mu{}_\rho = 0$), 很显然，加上这一项修正以后，电磁场的能动量张量就变成了如下对称张量

$$T^{\mu\nu} = F^\mu{}_\rho F^{\nu\rho} - \frac{1}{4}\eta^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}. \tag{2.113}$$

这才是一个真正规范不变的量。这个能动量张量有一个重要的特征，即

$$T^\mu{}_\mu = 0. \tag{2.114}$$

2.3.2 多粒子系统的能量动量张量

考虑多个相对论性自由粒子所构成的系统，我们以$n = 1, 2, ..., N$来标记不同的粒子。根据本章前面的知识可以知道，这个系统的作用量可以写成

$$S[x(\sigma)] = -\sum_n m_n \int d\tau_n = -\sum_n m_n \int d\sigma_n \sqrt{-\eta_{\mu\nu}\frac{dx_n^\mu}{d\sigma_n}\frac{dx_n^\nu}{d\sigma_n}}. \tag{2.115}$$

式中x_n^μ为粒子n的时空坐标，m_n为它的质量，τ_n为它的固有时，σ_n为它的世界线参数。由于上式在$\sigma_n \to \tilde{\sigma}_n(\sigma_n)$的重参数化之下保持不变，所以$\sigma_n$的选择有很大的任意性，特别的，我们可以将$\sigma_n$选作固有时$\tau_n$。

不妨以固有时参数(即$\sigma_n = \tau_n$)下的粒子作用量来进一步讨论。作用量(2.115)显然具有$x_n^\mu \to {x'}_n^\mu = x_n^\mu + a^\mu$($a^\mu$为常矢量) 的时空坐标平移不变性。为了考察这一时空平移对称性所对应的能动量张量，我们考虑如下局域化的无穷小时空坐标变换

$$x_n^\mu \to {x'}_n^\mu = x_n^\mu + \epsilon^\mu(x_n) \Leftrightarrow \delta x_n^\mu = \epsilon^\mu(x_n). \tag{2.116}$$

(从而$\delta(dx_n^\mu) = d(\delta x_n^\mu) = \frac{\partial \epsilon^\mu}{\partial x_n^\nu}dx_n^\nu$) 注意到

$$\begin{aligned}
\delta S &= -\sum_n m_n \int \delta(d\tau_n) = -\sum_n m_n \int \frac{1}{2}\frac{\delta(d\tau_n^2)}{d\tau_n} \\
&= \sum_n m_n \int \frac{1}{2}\delta(\eta_{\mu\nu}dx_n^\mu dx_n^\nu)/d\tau_n \\
&= \sum_n m_n \int \frac{dx_n^\mu}{d\tau_n}\eta_{\mu\nu}\delta(dx_n^\nu) \\
&= \sum_n m_n \int \frac{dx_n^\mu}{d\tau_n}\eta_{\mu\nu}d(\delta x_n^\nu).
\end{aligned} \tag{2.117}$$

可得在上面的局域时空坐标变换下，作用量的改变量为

$$\begin{aligned}
\delta S &= \sum_n m_n \int \frac{dx_n^\mu}{d\tau_n}\eta_{\mu\nu}d(\delta x_n^\nu) \\
&= \sum_n m_n \int (\frac{dx_n^\mu}{d\tau_n})\frac{\partial \epsilon_\mu}{\partial x_n^\nu}dx_n^\nu \\
&= \sum_n m_n \int d\tau_n(\frac{dx_n^\mu}{d\tau_n}\frac{dx_n^\nu}{d\tau_n})\frac{\partial \epsilon_\mu}{\partial x_n^\nu}.
\end{aligned} \tag{2.118}$$

通过引入四维时空δ函数$\delta^4\big(x-x_n(\tau_n)\big)=\prod_{\mu=0}^{3}\delta\big(x^\mu-x_n^\mu(\tau_n)\big)$, 可以进一步将上面结果写成

$$\begin{aligned}\delta S &= \sum_n m_n \int d\tau_n \big(\frac{dx_n^\mu}{d\tau_n}\frac{dx_n^\nu}{d\tau_n}\big)\frac{\partial\epsilon_\mu}{\partial x_n^\nu} \\ &= \int d^4x \Big[\sum_n m_n \int d\tau_n \big(\frac{dx_n^\mu}{d\tau_n}\frac{dx_n^\nu}{d\tau_n}\big)\delta^4\big(x-x_n(\tau_n)\big)\Big]\partial_\nu\epsilon_\mu(x). \end{aligned} \tag{2.119}$$

完全类似于前面关于标量场能量-动量张量的讨论，可知，与时空平移对称性对应的能动量张量为

$$T^{\mu\nu}(x)=\sum_n m_n \int d\tau_n \big(\frac{dx_n^\mu}{d\tau_n}\frac{dx_n^\nu}{d\tau_n}\big)\delta^4\big(x-x_n(\tau_n)\big). \tag{2.120}$$

显然，这个能动量张量是一个对称张量。

为了看清楚上述能动量张量的意义，我们将它重写成，

$$T^{\mu\nu}=\sum_n m_n \int d\sigma_n \big(\frac{dx_n^\mu}{d\sigma_n}\frac{dx_n^\nu}{d\tau_n}\big)\delta^4\big(x-x_n(\sigma_n)\big). \tag{2.121}$$

然后取世界线参数$\sigma_n=x_n^0$, 则当我们做完对σ_n的积分后，即有能动量密度

$$T^{0\mu}=\sum_n m_n \big(\frac{dx_n^\mu}{d\tau_n}\big)\delta^3\big(\mathbf{x}-\mathbf{x}_n(t)\big)=\sum_n p_n^\mu\delta^3\big(\mathbf{x}-\mathbf{x}_n(t)\big). \tag{2.122}$$

类似的，在取$\sigma_n=x_n^0$并做完对σ_n的积分以后，也有流密度

$$\begin{aligned}T^{i\mu} &= \sum_n m_n \big(\frac{dx_n^\mu}{d\tau_n}\frac{dx_n^i}{dt}\big)\delta^3\big(\mathbf{x}-\mathbf{x}_n(t)\big) \\ &= \sum_n p_n^\mu v_n^i \delta^3\big(\mathbf{x}-\mathbf{x}_n(t)\big). \end{aligned} \tag{2.123}$$

综合(2.122)式和(2.123)式，可知粒子系统的能动量张量可以表达成

$$T^{\mu\nu}=\Big[\sum_n \frac{p_n^\mu p_n^\nu}{p_n^0}\delta^3\big(\mathbf{x}-\mathbf{x}_n(t)\big)\Big]. \tag{2.124}$$