Chapter 2

狭义相对论

现在的确否认了绝对空间及其上的运动，那绝对时间呢？前文提到要在电磁场上找灵感，我们先来看看当时对电磁场的认识发生了怎样的革命.

2.1 相对性原理

19 世纪，光作为"光介质"或发光以太的扰动的波动理论被广泛接受，该理论在 James Maxwell 的著作中达到了最发达的形式. 根据 Thomas Young(1804) 和 Augustin-Jean Fresnel(1816) 的工作，人们认为光在称为发光以太的弹性介质中以横波的形式传播. 然而，光学现象和电动力学现象是有区别的，因此有必要为所有现象创建特定的以太模型. 试图统一这些模型或创建对它们的完整机械描述的尝试没有成功，但经过诸如 Faraday、Kelvin 男爵等许多科学家的大量工作，Maxwell 于 1864 年提出了一个准确的电磁理论，推导出一组电、磁和电感方程. 真空无电荷的假设下的 Maxwell 方程组为

$$\begin{cases} \boldsymbol{\nabla}\cdot\boldsymbol{E}=0, \quad \boldsymbol{\nabla}\times\boldsymbol{E}=-\partial_t\boldsymbol{B}, \\ \boldsymbol{\nabla}\cdot\boldsymbol{B}=0, \quad \boldsymbol{\nabla}\times\boldsymbol{B}=\mu_0\epsilon_0\partial_t\boldsymbol{E}, \end{cases} \tag{2.1.1}$$

这里使用了当时的国际单位制 (SI)，其中 μ_0 称为真空磁导率 *(vacuum permeability)*，ϵ_0 称为真空介电常数 *(vacuum permittivity)*，数值为

$$\mu_0 \approx 1.3\times 10^{-6}\ \mathrm{N/A^2}, \quad \epsilon_0 \approx 8.9\times 10^{-12}\ \mathrm{F/m}.$$

为从中推出真空电磁场的波动方程，注意

$$\mu_0\epsilon_0\partial_{tt}\boldsymbol{E}=\boldsymbol{\nabla}\times(-\boldsymbol{\nabla}\times\boldsymbol{E})=\nabla^2\boldsymbol{E}-\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{E})=\nabla^2\boldsymbol{E}; \tag{2.1.2}$$

同理

$$\mu_0\epsilon_0\partial_{tt}\boldsymbol{B} = \nabla^2\boldsymbol{B}, \tag{2.1.3}$$

于是发现，真空电磁波以某种常速率 c 传播，数值上是

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} \approx 3 \times 10^8 \text{ m/s}, \tag{2.1.4}$$

这与光速的实验测定高度吻合，这预言了“光是一种电磁波”的结论，并在 1887 年由 Hertz 实验证实．随后，以 Maxwell 方程组为核心的经典电磁理论的正确性被大量实验所证实. 但是这里的光速是相对于谁的? 注意到我们并没有事先假设任何特定

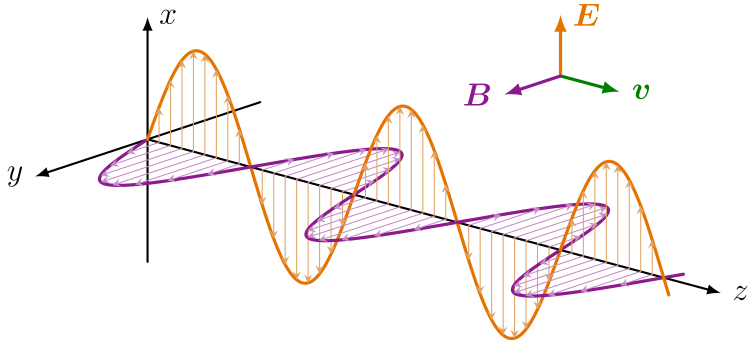


图 2.1: 电磁波

参考系，就已经得到了一个“绝对”的数值．这不符合当时人们的直觉．我们知道，Newton 的经典物理学建立在绝对时间和空间的基础上，坚持 Galileo 的相对性原理，将其精确地重述为机械系[1]．这可表述为：就力学定律而言，所有处于惯性运动中的参考系都享有同等的特权，不能将任何特定的惯性系归因于偏好的运动状态．不可通过观察粒子及其他物体的动力学行为识别一个参考系是静止的还是在做匀速运动. 因此在经典力学的世界里力学定律在所有的惯性系中都成立：不存在静止的绝对标准；只有相对运动是可观察的．Galileo 在其《关于两大世界体系的对话》中通过辩论说明这一原理，“*如同在甲板下的船舱中观察到的那样，在平静海面上船的匀速行驶不影响鱼、蝴蝶和其他运动物体的行为*”．相对性理论将这一原理作为基础，正如关于空间和时间本性的表述等同于关于 Newton 运动方程特性的表述．但若给予其这种普遍意义，则必须适用于全部物理学，而不只是适用于 Newton 力学. 最初这似乎不成问题，很难想像它在这种基本水平上成立，但不适用于更复杂的物理相互作用．尽管如此，当尝试将其扩展到电磁学时，由于 *Galileo 协变性*与 Maxwell 方程组相抵触，深度的问题还是显露出来.

[1] 《原理》运动定律之推论 V.

对含有低速运动电荷及低速变化的电场、磁场的系统，一切显得直截了当. 它们看来像是被不同匀速运动参考系之间变换的不变性定律所支配. 可以想像，*Galileo* 船的现代方案也可以携带若干磁体、电池和半导体以及其他电性元件. 主人公“Salviati”对相对性的论证似乎是强有力的.

但当考虑迅变场，尤其是考虑光的传播时，问题就出现了. 后来的 Einstein 在他的原始论文《论动体的电动力学》[2]中指出，“*Maxwell* 电动力学⋯⋯，当应用到运动物体时导致一些不对称性，而该不对称性似乎并非现象所固有的”. 如前文所述，主要的困难是 Maxwell 方程组给光和其他电磁波一个确定的速度，在真空中它以相同的速度沿各方向传播，与源的运动无关，这一事实与 Galileo 变换不相容，光在一个参考系中以速率 c 传播，在向着光源以速率 u 运动的参考系中必定具有速率 $c+u$. 因此，对于光以任意速率传播必定是可能的，光在一个光源静止的参考系中以速率 c 传播，在运动参考系中应有某种其他速率，所以 Galileo 变换意味着光速依赖于源的运动.

为解决这一矛盾，Maxwell 放弃相对性原理，首先提出光实际上是在同一个以太介质中的波动 (电磁辐射)，认为方程组只对一个绝对参考系 (以太) 成立，它是电磁现象的起因. 然而，Maxwell 的理论在运动物体的光学方面并不令人满意，虽然他能够提出一个完整的数学模型，但他无法提供对以太的连贯机械描述. 据这一假说，由方程组计算得到的真空光速是相对于以太的速度，在相对于“以太”运动的参考系中，光速具有不同的数值. 根据 Maxwell 的理论，所有光学和电学现象都通过该介质传播，这表明应该可以通过实验确定相对于以太的运动. 然而，任何已知的探测以太运动的实验都失败了. Fizeau 实验和 Michelson-Morley 实验表明光速与参考系的运动无关. 该实验结果否定了以太假说，表明相对性原理的正确性.

这导致 Hendrik Lorentz 从 1892 年开始发展了一种基于不动的发光以太（虽然没有推测其物质构成）、物理长度收缩和“当地时间”的理论，其中 Maxwell 方程在所有惯性系中保持其形式. Lorentz 的假说解决了上述矛盾，但他不能对 Lorentz 变换的物理本质做出合理的解释. 在 Lorentz 的以太理论之后，Henri Poincaré 猜测 Lorentz 变换和时空性质有关，他早先提出了“相对性原理”作为物理学的“一般”规律（包括电动力学和万有引力），在 1905 年利用这一原理修正了 Lorentz 的初步变换公式，也即后来的 boost 变换. 然而，只有 Einstein 的深刻洞察力，才把这种不变性提升到基本原理的高度.

Einstein 从 16 岁就开始思考的光与以太的问题，在此之前他已经明白电磁理论与经典力学的 Galileo 相对性不相容，他对其印象深刻：单个物体的速度没有意义、

[2] [?]，这是狭义相对论（下称狭相）确立的重要标志. 第一部分主要解决运动学上的矛盾，他没完全使用初等数学而独立地得到了 Lorentz 的结果；而在第二部分中，他用半边篇幅系统地描述了 Maxwell 方程组在不同惯性参考系之间的变换，但由于并非使用矢量微积分语言，略显繁琐. 论文发表于 1905 年 6 月 30 日. 这一年他共发表 6 篇论文，在 5 个不同领域中取得 4 个重大历史性成就，故亦史称“奇迹年”.

只有物体之间的相对速度才有意义. 请注意, 从某种意义上说, 这是 Newton 证明 "真正的运动" 的企图一个微小但重要的失败. 对 Einstein 来说, 这是一个暗示, 表明 Newton (以及用经典力学看待电动力学) 的概念有问题.

尽管 Newton 理论在经验上取得了成功, 但其关于绝对空间的想法仍令人深感不安. 正如 Leibniz、Mach 和其他许多人所强调的, 空间是一种超感官的实体, 它 "作用" 于物体, 但物体不能 "作用" 于它. Einstein 确信, 这种绝对空间的想法是错误的. 不可能有绝对的空间, 不可能有 "真正的运动". 只有相对运动, 因此相对加速度必须具有物理意义而绝对加速度不应加入物理方程. 通过狭义相对论, Einstein 成功地将伽利略有关于速度的相对论从 Maxwell 理论的挑战中解放出来. 然后, 他确信他可以为整个 Aristote-Descartes 的相对论进行辩护. 用 Einstein 的话说, "不仅是在惯性系中, 运动定律在所有参考系中都应该是一样的". 事物的运动是相对于彼此而言的, 而不是相对于一个绝对的空间而言的; 不可能有任何绝对运动的物理效应[3].

Einstein 意识到 Galileo 变换实际上是 Newton 经典时空观的体现, 如果承认 "真空光速独立于参考系" 这一实验事实为基本原理, 可以建立起一种新的相对时空观. 设 S', S 初态 $t = t' = 0$ 时, 空间轴重合, 但 S' 相对于 S 以沿共同 x, x' 轴正方向的 v 运动, 稍后证明可导出如下 *boost* 变换 *(transformation)*[4]:

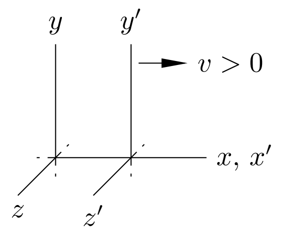


图 2.2: boost 变换示意图

$$\begin{cases} y' = y, \quad x' = \dfrac{x - vt}{\sqrt{1 - v^2/c^2}}, \\ z' = z, \quad t' = \dfrac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}. \end{cases} \tag{2.1.5}$$

进而地就可以明确规定: 描述物理定律的基本方程都应当在 Lorentz 变换中保持形式不变, 其结果是时空平移、空间旋转、boost 变换以及三者的任意复合. 而经典力学的 Galileo 变换

$$\begin{cases} y' = y, \quad x' = x - vt, \\ z' = z, \quad t' = t \end{cases} \tag{2.1.6}$$

其实就可以看作 boost 在 $c \to \infty$ 时的极限. 如此便调和了经典力学和电动力学. 这一理论完满解决如前所述的矛盾.

承认光速的不变将导致一些有趣的现象. Einstein 首先提出了一个关键性的概念: "同时" 具有相对性. "那列火车 7 点钟到达这里", 这实际上表述的是 "火车到达" 和 "表的短针指向 7" 是同时的事件. 但我们在这一参考系中观察到的同时事件,

[3]根据许多当代物理学家的说法, 这是对 "哲学" 思想的过度重视, 它不应该在物理学中发挥作用. 但 Einstein 在物理学上的成就远比这些物理学家获得的成就更为重要.

[4]其含义取时间轴上的推动之意. 一个较好的翻译是 "伪转动" (pseudo rotation), 我们将在本节末解释这一点.

在另一个参考系中不一定也“同时”. 在另外一个参考系中, 火车有可能不是在 7 点钟到达的, 而是“在 6 点或者 8 点到达的”.

由于同时的相对性, 当我们无法与某物体保持相对静止时, 比如你想要测量一量高速行驶的列车某一车厢的“长度”, 你必然会使用到计时器、标尺等工具, 我们就会注意到以下现象: 物体运动时, 它的一切 (物理、化学变化等) 从参照系的时间来看都会变慢; 重新审视长度的真正含义, 并会发现, 如上述的列车在某一参考系的运动测量值会短于静止测量值; 速度将不再是满足线性叠加的矢量, 因为我们所熟知的“位移矢量对时间求导”的“时间”将不再绝对; 光速不变避免了在电动力学中对发光以太的任何参考的需要, 并提出所谓的*相对论限制 (restriction): 任意惯性参考系下, 均不容许有质粒子以快于或等于光速的速度运动*.

最后, 相对性原理除了引出新的时空观念外, 还导致了质量和能量的深刻重组. 同年 9 月 27 日, Einstein 专门发表了另一篇论文《物体的惯性同它所含的能量有关吗?》[?] 以讨论力学中的能量问题. 他企图讨论 Newton 第二定律推广至这个新时空观时发现了如下说法: 一个物体的能量在自洽意义上可以定义为

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}}. \tag{2.1.7}$$

其中 m 为物体的质量. 上式被后人称为*质能方程*. Einstein 为保持在形式上与 Newton 力学类似, 捏造了一个“动质量”概念 $m_k = m/\sqrt{1 - v^2/c^2}$, 此外又称静止系中测得的“静能”为 $E_0 = mc^2$. 现在人们一般不再使用这种“余孽”定义, 而选择将其扔进历史的纸篓桶里. 注意, 在相对论中, 这代表着质量和能量的“等同”而非“转化”. 而后会解释这样一种单位制, 其中 $c = 1$, 则直接有 $E_0 = m$, 此即说任何静止物体当然都含有“能量”, 毕竟其有内蕴的“质量”. 可见 Einstein 对质量及能量做了统一的诠释.

2.2 参考系变换

从形式上看, boost 变换是一个线性变换 Λ, 连接着原惯性系 S 所测量的 x, y, z, t 和新惯性系 S' 所测量的 x', y', z', t'. 整个变换式表为

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \Lambda \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}, \tag{2.2.1}$$

其中给时间乘上这么一个常数 c 以使量纲一致, 这样 Λ 的分量就是无量纲的. 为了学习以后的张量知识, 我们现在可以开始介绍一点用以运算的符号了. 我们给如上

的列矢量的分量进行编号:

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z, \tag{2.2.2}$$

当然，这里的编号不是乘幂，而只是指标 *(index)*，并且是上指标 *(upper/raised indices)*，因为在给行矢量、矩阵分量编号的时候还可以用下指标 *(lower indices)*，为此而作上下之区分. 不必担心，在很多情况下，我们是不会与乘幂相混淆的；实在冲突时，只需文字标注即可. 现在用 $\Lambda^\mu{}_\nu$ 来表示 Λ 的矩阵元，其中位于左边的 μ 称为行指标、右边的 ν 称为列指标，那么根据矩阵乘法，我们很容易发现每个 μ 分量都满足等式

$$x'^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu x^\nu,$$

可见，对一个列矢量左乘一个矩阵 Λ ，相当于与其矩阵元 $\Lambda^\mu{}_\nu$ 下方的列指标 ν “相靠”后，在上方留一个行指标 μ. Einstein 认为，既然线性代数的求和运算都已事先蕴含在矩阵乘法的定义之中，那何必不直接“甩掉求和号”，而只关注分量本身呢? 毕竟这样的话

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.2.3}$$

似乎更像是原来的矩阵方程. 而矩阵的乘法体现在哪儿? 体现在这两个实数的相乘，以及指标 ν 的“相靠”和“一上一下”之中. 像这样，*如果在一个式子中，同一个指标既作为上标出现，又作为下标出现，就意味着要对该指标求和*. Einstein 这个懒偷得非常巧妙但又十分深邃，所以后人给其取名为*求和约定 (summation convention)*. 关于“一上一下”这件事，或许有些书会选择统一放置在下方成为 $\Lambda_{\mu\nu}$ 而只强调指标“总共重复两次”的条件，更有甚者会选择将矩阵元直接写为 Λ^μ_ν 而不区分其左右位置. 这样在讨论更深刻的问题时都会出现缺陷：前者容易混淆矩阵的转置和逆，后者说不清楚“谁与谁相靠”，并且在“升降指标”这个操作时出现顺序混淆的问题[5]. 因此建议只在部分情况不会混淆时，再采取这种简化.

当然，我们还未说清 boost 变换的来历. Einstein 欲从相对性原理和光速不变导出变换式，但这两句话的描述是模糊的，并不能推出变换一定线性. 实际上，我们还需要在此二“公理”外另加约束：时空均匀性、空间的各向同性. 相对性原理断言，存在一类特殊的参考系（称为惯性系），其中物理定律一致；在惯性系中做匀速运动的惯性系仍为惯性系. 这说明，我们要在惯性系类中区分各惯性系，需要依靠是否有相对匀速运动，没有相对速度的便视为同一个惯性系. 那么，这可能与坐标或者时间有关吗？无关，否则不同空间点或者不同时刻就不再等价，这与时空均匀性相悖. 会与相对速度方向有关吗？无关，否则与空间各向同性相悖. 可见，以某一惯性系为基

[5]我们以后会知道，指标的上下代表着矢量、对偶矢量或者逆变、协变的区分，指标的左右代表着缩并作用、基底张量积的顺序.

准后，不同惯性系的特征由相对速率 $|v|$ 描述. 相对性原理说，这是可以互换角色的，即二者都应认为对方相对自己以 $|v|$ 运动. 可见，任意匀速运动应当在变换下仍保持匀速，而线性变换就是在数学上保持任意直线的变换，说明变换一定线性.

而光速不变原理说，所有惯性系所测光速一致. 在某系，光的速度表为

$$\frac{\sqrt{\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2}}{\mathrm{d}t}=c, \tag{2.2.4}$$

此即是说

$$\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2-c^2\mathrm{d}t^2=0 \tag{2.2.5}$$

应当在任意惯性系成立. 该等式便可表为

$$\begin{bmatrix} c\,\mathrm{d}t & \mathrm{d}x & \mathrm{d}y & \mathrm{d}z \end{bmatrix}\begin{bmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}\begin{bmatrix} c\,\mathrm{d}t \\ \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix}=0,$$

因此按分量表示为 $\eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu=0$，这里 $\eta_{\mu\nu}$ 表示对角矩阵 $\mathrm{diag}(-1,1,1,1)$ 的矩阵元. 可见这里一个因式就出现了两个指标 μ,ν. 不过，当多对指标重复时，由乘法的交换、结合及分配律知，多组求和顺序是无所谓的，即 $\sum_{\mu,\nu}=\sum_\mu\sum_\nu=\sum_\nu\sum_\mu$. 这再一次体现了求和约定对其分量关注的精神，一般而言不会混淆.

令*线元平方 (square of line element)* 为

$$\mathrm{d}s^2=\eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu. \tag{2.2.6}$$

由于在 S' 下同样有 $\mathrm{d}s'^2=\eta_{\mu\nu}\mathrm{d}x'^\mu\mathrm{d}x'^\nu=0$. 在更广泛的情形下，考虑到 $\mathrm{d}s,\mathrm{d}s'$ 乃同阶的无穷小，而对光而言二者同为零，故二者应成比例：

$$\mathrm{d}s'^2=k\,\mathrm{d}s^2, \tag{2.2.7}$$

其中系数 $k\neq0$ 只与两惯性系的 $|v|$ 有关. 互换角色，在物理上同样应有 $\mathrm{d}s^2=k\,\mathrm{d}s'^2$，因此 $k=\pm1$，而当 $|v|=0$ 时两惯性系等同，故应取 $k=1$，即

$$\mathrm{d}s'^2=\mathrm{d}s^2. \tag{2.2.8}$$

这个性质就称作*间隔不变性 (interval invariance)*. 坐标变换总存在全微分关系 $\mathrm{d}x^\mu=(\partial x^\mu/\partial x'^\sigma)\mathrm{d}x'^\sigma$（指标可任选），因此对任意 x'^σ，上式可写为

$$\eta_{\sigma\lambda}\mathrm{d}x'^\sigma\mathrm{d}x'^\lambda=\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu\Leftrightarrow\eta_{\sigma\lambda}=\eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x^\nu}{\partial x'^\lambda}, \tag{2.2.9}$$

分母下方是上指标，而这相当于分子下指标，因此这样写是指标平衡的. 在这个特殊例子里，坐标由线性变换 $x'^\nu={\Lambda^\nu}_\lambda x^\lambda$ 联系，因此 ${\Lambda^\mu}_\sigma=\partial x^\mu/\partial x'^\sigma$，故

$${\Lambda^\mu}_\sigma\eta_{\mu\nu}{\Lambda^\nu}_\lambda=\eta_{\sigma\lambda}, \tag{2.2.10}$$

关于矩阵的转置，比如对于 $\Lambda^{\mu}{}_{\sigma}$，我们像这样左右交换，即对调行列指标为 $\left(\Lambda^{\top}\right)_{\sigma}{}^{\mu}$，有时简写为 $\Lambda_{\sigma}{}^{\mu}$，这样 $\Lambda_{\sigma}{}^{\mu}\eta_{\mu\nu}\Lambda^{\nu}{}_{\lambda}$ 是指标相靠的，相当于按此顺序乘出另一矩阵. 这样的 Λ 就称作 *Lorentz* 变换. 在 boost 变换的简单情形下，相当于只研究 t, x，这样 Λ 在实质上就成为了一个二阶矩阵（对另两个空间坐标而言是恒等变换 I）. 于是方程 (2.2.10) 成为了

$$\begin{bmatrix} (\Lambda^0{}_1)^2 - (\Lambda^0{}_0)^2 & \Lambda^0{}_1\Lambda^1{}_1 - \Lambda^0{}_0\Lambda^1{}_0 \\ \Lambda^0{}_1\Lambda^1{}_1 - \Lambda^0{}_0\Lambda^1{}_0 & (\Lambda^1{}_1)^2 - (\Lambda^1{}_0)^2 \end{bmatrix} = \begin{bmatrix} -1 & \\ & 1 \end{bmatrix}, \tag{2.2.11}$$

这里括号外的 “2” 当然就表示平方了. 为得到合乎情理的解，可做如下限制：排除时空反演. 排除时间反演意味着 $\Lambda^0{}_0 > 0$，排除空间反演意味着 $\Lambda^1{}_1 > 0$. 如此对角元为相等正数，再由 $(\eta_{\mu\nu})$ 的零元，则 Λ 成为对称矩阵，就能保证行列式 $|\Lambda| = 1 > 0$. 根据物理意义（或者说那些额外的 “公理”），一个 v 就决定了一个 Λ. 在 boost 变换的情形中，有

$$\begin{bmatrix} ct' \\ 0 \end{bmatrix} = \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 \\ \Lambda^0{}_1 & \Lambda^1{}_1 \end{bmatrix} \begin{bmatrix} ct \\ vt \end{bmatrix},$$

观察第二行有 $\Lambda^0{}_1 = -v\Lambda^1{}_1/c$，记 *Lorentz* 因子 *(factor)* 为

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}, \tag{2.2.12}$$

这样再代回 (2.2.11) 解得

$$\Lambda^0{}_0 = \Lambda^1{}_1 = \gamma, \quad \Lambda^0{}_1 = \Lambda^1{}_0 = -\gamma v/c. \tag{2.2.13}$$

这样便得到了 boost 变换. 我们可求出沿任意 $\boldsymbol{v}$ 方向的 boost 变换. 空间坐标构成矢量 $\boldsymbol{x}$，将其按平行、垂直 $\boldsymbol{v}$ 方向分解，即

$$\boldsymbol{x} = \frac{(\boldsymbol{x} \cdot \boldsymbol{v})\boldsymbol{v}}{v^2} + \left(\boldsymbol{x} - \frac{(\boldsymbol{x} \cdot \boldsymbol{v})\boldsymbol{v}}{v^2}\right),$$

垂直分量保持不变，而平行分量与时间混合起来 boost. 时间上有

$$ct' = \gamma(ct - \boldsymbol{v} \cdot \boldsymbol{x}/c),$$

空间上有

$$\begin{aligned} \boldsymbol{x}' &= \gamma\left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}}{v} - vt\right)\frac{\boldsymbol{v}}{v} + \left(\boldsymbol{x} - \frac{(\boldsymbol{x} \cdot \boldsymbol{v})\boldsymbol{v}}{v^2}\right) \\ &= \boldsymbol{x} + \frac{\gamma - 1}{v^2}(\boldsymbol{v} \cdot \boldsymbol{x})\boldsymbol{v} - \gamma\boldsymbol{v}t. \end{aligned}$$

用 δ^i_j, δ_{ij} 表示恒等变换 I 的矩阵元. 不区分 δ^i_j 指标左右是考虑到 I 是对称矩阵. 有必要约定一下：一般像 i,j 这种拉丁字母默认取遍 $1,2,3$，而希腊字母 α,μ 等默认取遍 $0,1,2,3$. 这样，上式的分量表达为

$$\begin{aligned} ct' &= \gamma ct - \frac{\gamma v_i}{c} x^i, \\ x'^i &= -\frac{\gamma v^i}{c} ct + \left(\delta^i_j + v^i v_j \frac{\gamma - 1}{v^2} \right) x^j. \end{aligned}$$

其中 $v_i = \delta_{ij} v^j$. 综上，Λ 表为

$$\Lambda^0{}_0 = \gamma, \quad \Lambda^i{}_0 = -\gamma v^i / c, \quad \Lambda^0{}_i = -\gamma v_i / c, \quad \Lambda^i{}_j = \delta^i_j + v^i v_j \frac{\gamma - 1}{v^2}. \tag{2.2.14}$$

这便是任意方向 boost 变换.

2.3 时空度规

绝对的“深刻事物”在不同参考系下测得相对的“不同侧面”，这个思想引导狭相成为了一门几何学：Einstein 对狭相的理解是“代数”的，而 Minkowski 于 1908 年给 Einstein 的代数理论提供了深藏其下的“几何”解释.

如前文所述的一切“奇特”现象，皆是“光速”与相关运动定律在任意惯性系不变所带来的，究其直觉，就是“同时的相对性”. 这使我们构建了一个颠覆性的先进观念：那就是要区分“绝对”和“相对”！虽然确实可以说在物体相对静止系下所测数据为原有的“实际值”，而在运动系下所测为“观测值”. 然而，讨论物理量必经由实验测量，而从测量角度而言，没有何种的测量结果能鹤立鸡群，独特地作为“原有的”、“实际的”，尽管确实在不同参考系中测量“同一对象的物理量”，却得到“不同结果”. 这意味着，我们有必要认为或许测量值本身是可以相对变化的，而不变的应该就是那个“深刻事物”. 这些测量只是该事物的不同侧面. 根据间隔不变性，考虑定义这样一种物理量，它是物体在其静止系所测的内在时间，称固有时 *(proper time)* τ，则根据定义取空间分量为零，有

$$\mathrm{d}s^2 = -c^2 \mathrm{d}\tau^2, \tag{2.3.1}$$

那么就有可能通过给空间再“增添”一个时间维度，使我们所讨论的物理背景为一个 4 维的时空连续统 *(space-time continuum)*，简称时空 *(spacetime)*，而找到那些“深刻事物”. Minkowski 于 1908 年在德国科隆的自然哲学家大会 (Naturforscher Versammlung) 上以“空间与时间”为题发表了一篇演讲，其开篇是：“我打算提给你们的空间和时间的概念，是从实验物理学的土壤上生长出来的，它的力量正在于此. 这种观念是革命性的. 由于有了这个观念，空间本身和时间本身，注定了会逐渐

消逝，成为仅仅是影子，只有二者的联合才会保持为独立的现实性.”该演说又发表为了论文 [**?**]. 足以见得时空概念的革命性.

或许负数开方总是会让读者有些不适，但不必担心，我们不需要讨论 $\mathrm{d}s$ 的虚实. 为了消除这一点，部分书籍会选择令 $(\eta_{\mu\nu}) = \mathrm{diag}(1,-1,-1,-1)$，从而使得 $\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 > 0$. 如果定义 $\eta_{\mu\nu}$ 的 Lorentz *号差 (signature)* 为其矩阵的迹，则这种选择就是 ${\eta^\mu}_\mu = -2$，而本书一般选择 $+2$. 更有甚者选择将时间分量排在最后作为第 4 分量，这样就有 $\mathrm{diag}(1,1,1,-1)$，但这并不影响号差. 无论怎样，这些选择纯粹只是偏好，不会带来任何实质区别[6].

我们知道，Newton 时空观认为空间平直线性，其上运动可由三维速度描述，因此时间是割裂于空间的自变量. 但正如前文讨论，狭相时空观已经如此地暗示了时间和空间的深刻联系，那么 Minkowski 对 Einstein 理论的理解是非常自然的. 可见，间隔不变性保证了均匀流逝的固有时的存在性，并使之为绝对事物，就像质量一样，而真正成为可以表征物体自身固有属性的*内禀 (intrinsic) 物理量*.

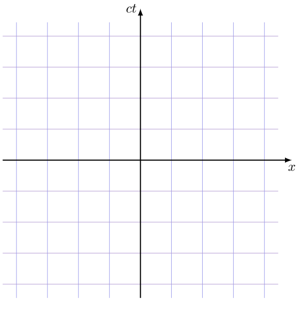


图 2.3: 时空图

对初学者来说，关键总是在于究竟如何想象时空. 还记得 x-t 图吗？为便于理解，我们就以这样一张 2 维白纸 $\mathbf{R}^2$ 举例，其上只有两个任意选取的自然坐标. 当选定某种原点及其坐标系时，坐标数字才具有位置的意义. 图上的点都被赋予了两个坐标，因此也称作事件 *(event)*，毕竟为了叙述一个生活意义上的事件，其实和写记叙文没什么区别：那便是地点和时间[7]. 经典力学中，两个事件 $(t,x),(t,x')$ 相当于自然语言的“在 x 处及另一不同 x' 处于相同时刻 t 发生的事”. Newton 习惯像这样把“处”和“时刻”分割开来看，但在狭相中很容易找到一个 Lorentz 变换使得这两个事件不再具有相同的 t. 我故意没有提及“人物”：不同的人当然可以在同一个时空点干不同的“事”，但为了区分这种“事”，只需要说清楚对应的是谁即可. 实际上，如果我们定义*世界线 (worldline)* 就是物体所经历的一系列事件构成的曲线，那在一部小说中，当然也就可以区分谁干了什么事，因为不同的事在不同的“故事线”上. 因此，我们就不再专门设置这样一个鸡肋的概念，而把时空的点就称作事件，而把物体等同于一条世界线. 像这样示意时空背景及其上曲线的图就是*时空图 (spacetime diagram)*. 我们选择时间轴正方向朝上，而空间轴朝右，在习惯上与经典力学有微小区别. 考虑到我们所处的是三维空间，暂且不严谨地说，像这样的点 (ct,x,y,z) 所构

[6] 可能会带来哲学区别的是这样的“Wick 技巧”：如果设 $x^0 = \mathrm{i}ct$，则 $\eta_{\mu\nu} = \delta_{\mu\nu}$，我们后续会知道，这意味着区别于闵氏几何的欧氏勾股定理. 那么 boost 变换就变成了 x-$\mathrm{i}ct$ 图的“旋转”. 一般除讨论量子力学复结构的哲学外，不建议使用这种做法.

[7] 至于起因、经过和结果，对应着的是动力学以及因果结构的内容，要在后续章节中才会讲到.

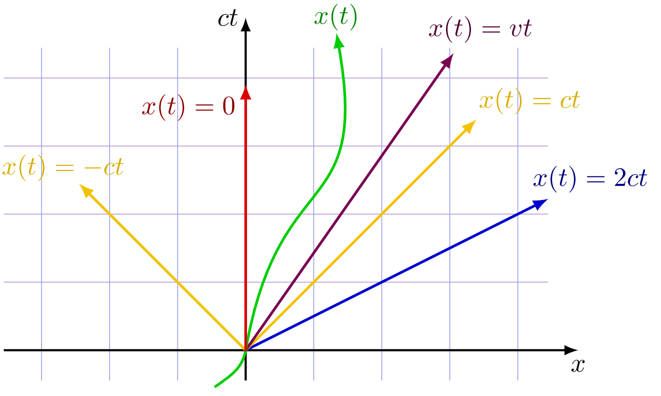


图 2.4: 世界线

成的集合就叫时空. 这里乘上 c 的理由如前文所述. 等等，那我们在推导 boost 时，矢量 $(ct,x,y,z)^\top$ 不正对应着时空坐标吗？的确，我们那时已隐约窥见了时空概念. 并且还可以告诉读者一个好消息：矢量正好属于我们所寻觅的“深刻事物”，其不同侧面正是分量. 间隔不变性意味着如下某种速度矢量

$$\left(\frac{\mathrm{d}x^0}{\mathrm{d}\tau},\frac{\mathrm{d}x^1}{\mathrm{d}\tau},\frac{\mathrm{d}x^2}{\mathrm{d}\tau},\frac{\mathrm{d}x^3}{\mathrm{d}\tau}\right)^\top$$

的“模长”平方

$$-\left(\frac{\mathrm{d}x^0}{\mathrm{d}\tau}\right)^2+\left(\frac{\mathrm{d}x^1}{\mathrm{d}\tau}\right)^2+\left(\frac{\mathrm{d}x^2}{\mathrm{d}\tau}\right)^2+\left(\frac{\mathrm{d}x^3}{\mathrm{d}\tau}\right)^2=-c^2 \tag{2.3.2}$$

是恒等的. 也就是说，我们最好认为至少这个矢量是个“几何绝对”的东西[8]，而相对体现在其分量的投影上. 我们知道三维空间里速率是

$$v=\frac{\sqrt{\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2}}{\mathrm{d}t}, \tag{2.3.3}$$

上下同除 $\mathrm{d}\tau$ 得

$$v=\frac{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\tau}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\tau}\right)^2}}{\mathrm{d}t/\mathrm{d}\tau}=\frac{c\sqrt{(\mathrm{d}t/\mathrm{d}\tau)^2-1}}{\mathrm{d}t/\mathrm{d}\tau},$$

因此实际上

$$\frac{\mathrm{d}t}{\mathrm{d}\tau}=\frac{1}{\sqrt{1-v^2/c^2}}=\gamma. \tag{2.3.4}$$

可见，相对速率 $|v|$ 的不同所带来的，不外乎就是 t,τ 间相对流逝“快慢”的变化.

[8] 对此，Einstein 曾表示相比于“相对论”这个名字，他更喜欢“绝对论”.

我们实际上还未细究一件事：究竟什么是模长呢？虽然坐标可以定位一个事件，但它并没给出几何信息，比如长度、夹角什么的. Minkowski 类比了一下几何学知识，指出狭相实际上给出了几何信息：正是这个“模长”. 换句话说，间隔不变性已包含了狭相的所有物理学. 是的，我们回忆一下，间隔不变性就是指线元

$$\mathrm{d}s^2 = \mathrm{d}x^\mu\,\eta_{\mu\nu}\mathrm{d}x^\nu$$

在不同惯性系中是绝对的. 而我们之前的所有物理结论确实可从该“公理”推出. 上式可看作二次型，类比一下空间的勾股定理，它是

$$\mathrm{d}\ell^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \mathrm{d}x^i\delta_{ij}\mathrm{d}x^j,$$

可见，要想定义距离，我们有必要重新审视一下几何学. 为了测量距离，当然需要一个“度量衡”标准，或者说一把“规尺”. 在几何学中，定义距离的是其相应的*度规 (metric)* 概念，这个命名十分自然. 假设球面上有两点 P, Q，它们之间的“距离”可以怎样确定？如果我们定义这个球面是所有使得 $x^2+y^2+z^2=1$ 的点 (x,y,z) 之集合，那么 P,Q 就是 $\mathbf{R}^3$ 中的点. 所以我们可以用勾股定理确定它们的距离，例如 $(1,0,0)$ 和 $(0,1,0)$ 的距离就是 $\sqrt{2}$. 然而我们真的想去计算直线段 PQ 的长度吗？这个直线段并不完全躺在球面上，所以用直线段来定义出来的长度似乎“什么也不是”. 在球面几何学里，我们知道，可以定义 P,Q 之间的距离为连接这两点，而完全贴于球面上的“最短路径”之长. 但无论如何，在球面的局部上总是有勾股定理的. 因此我们可以尝试用空间的勾股定理去*诱导 (induce)* 出在球面上的“勾股定理”.

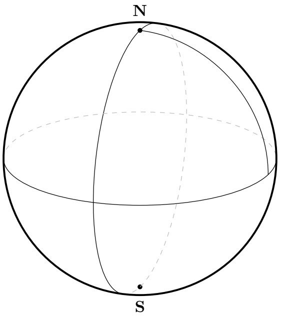


图 2.5: 球面

现在用上微积分的语言. 考虑一个 $\mathbf{R}^3$ 上的路径长度. 路径可以看成是一个动点 $(x(t), y(t), y(t))$ 之轨迹，假设 t 从 0 开始而于 1 停止. 如果 Δt 很小，则 $x(t+\Delta t)$ 可以近似为 $x(t)+\dot{x}(t)\Delta t = x+\Delta x$；对于 y, z 同理. 所以我们知道附近两点的距离就相当于“最短的直线长” $\Delta t\sqrt{\dot{x}^2(t)+\dot{y}^2(t)+\dot{z}^2(t)}$，或者说 $\sqrt{\Delta x^2+\Delta y^2+\Delta z^2}$. 要计算光滑路径的长度，就在此路径上取很多很多点，而相邻二者非常接近. 将这些线元加起来，可以给出一个很好的近似. 取得越多，逼近越好. 当 $\Delta t \to 0$ 时，所有沿路径的距离之和会得到一个积分，即

$$\ell = \int_0^1 \sqrt{\dot{x}^2(t)+\dot{y}^2(t)+\dot{z}^2(t)}\,\mathrm{d}t.$$

如果我们换成微分记号 $\frac{\mathrm{d}}{\mathrm{d}t}$, 则可以改写成

$$\int \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2} = \int \mathrm{d}\ell.$$

这样我们发现，想要诱导球面上的线长 ℓ，只需要讨论空间线元平方 $\mathrm{d}\ell^2$ 在球面上的表达式即可，有时在数学中叫 *Riemann* 度量，但是就叫它度规也未尝不可. 为了找到球面度规，我们只需要如下的球坐标变换即可:

$$\begin{cases} x = r\sin\theta\cos\phi, \\ y = r\sin\theta\sin\phi, \\ z = r\cos\theta. \end{cases} \tag{2.3.5}$$

代入到 $\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$ 之中，可以计算得到

$$\mathrm{d}\ell^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\phi^2, \tag{2.3.6}$$

设球面在球坐标下是 $r = R$，则

$$\mathrm{d}\ell^2 = R^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2),$$

有时把括号写作 $\mathrm{d}\Omega^2$，这里 Ω 正是立体角. 至此，给定球面上的曲线，即经纬度参数式或关系式, 那么就可利用上述线元计算球面距离. 比如，假定曲线是整个赤道 ε，此即 $\phi = 2\pi t$ 而 $\theta = \pi/2$，则线元化简为 $\mathrm{d}\ell = R\,\mathrm{d}\phi$，积分得到 $\int_\varepsilon \mathrm{d}\ell = 2\pi R$.

然而，刚才只是应用这个公式作为开始，这个例子还没有显示出度规的真正好处. 在几何学中，度规的选择直接决定了该体系内的几何知识，因此上述所讨论的都称作欧氏几何 *(Euclidean geometry)*. 历史上为论证平面几何“第五公理”的独立性，曾为此构造了另一种度规:

$$\mathrm{d}\ell^2 = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{1 - x^2 - y^2},$$

为使距离为实数，定义域要求 $x^2 + y^2 < 1$，而这个区域是一块不含边界的单位圆盘. 类似地，路径在该度规下的长度为

$$\int_0^1 \sqrt{\frac{\dot{x}(t)^2 + \dot{y}(t)^2}{1 - x(t)^2 - y(t)^2}}\,\mathrm{d}t.$$

当接近圆盘边缘时，这个“距离”比欧氏距离会越来越大. 这个度规所对应的体系称作双曲几何 *(hyperbolic geometry)* 的圆盘模型.

无论怎样，上述例子具有这样一种共性：由于在局部上线元都是“直”的，因此距离的表达式必须是二次型:

$$\mathrm{d}\ell^2 = E(x,y)\mathrm{d}x^2 + 2F(x,y)\mathrm{d}x\,\mathrm{d}y + G(x,y)\mathrm{d}y^2,$$

使线元均为正只需使得 $EG > F^2$ 即可. 交叉项只是因为两个坐标的线元不一定“垂直”, 但总归还是个“勾股定理”. 此即说, 无论这个几何多么扭曲, 至少在局部上要还原至欧氏几何. 当然, 最好这三个函数也具有一些光滑性.

这个定义可以直截了当地推广至更高维的几何, 因为勾股定理的高维推广仍然是二次型. 对于 n 维而言, 它就是

$$\mathrm{d}\ell^2 = g_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu, \tag{2.3.7}$$

其中 $g_{\mu\nu}$ 可是各坐标 x^μ 的函数, 排成了一张 $n\times n$ 的对称表格 $(g_{\mu\nu})$. $g_{\mu\nu}$ 称为度规分量, 不混淆时可称度规. 把这个线元平方叫做度规亦可, 原因稍后阐述. 在欧氏几何下, 度规是正定的 *(positive definite)*, 即 $g_{\mu\nu} > 0$, 可作强调地称其 *Riemann* 的; 否则就叫伪 *Riemann* 的 *(pseudo Riemannian)*, 包括负定的 *(negative definite)*, 以及不定的 *(indefinite)* 即正负都有的.

度规不仅仅是定义了距离概念, 实际上也定义了“夹角”. 给背景空间取一个坐标系, 定义第 x^μ 坐标线 *(coordinate line)* 是其上除 x^μ 外的坐标皆为常数的曲线. 在欧氏空间的直角坐标系中, 任意矢量 $\boldsymbol{v}$ 可用与坐标线相切的单位基 $\{\boldsymbol{e}_\mu\}$ 展开为 $v^\mu\boldsymbol{e}_\mu$. 皆于此基与坐标系的相关性可称坐标基 *(coordinate basis)*. 直角坐标系的坐标基还是正交的, 即正交归一的: $\boldsymbol{e}_\mu\cdot\boldsymbol{e}_\mu = \delta_{\mu\nu}$. 可见第 x^ν 坐标基自身在坐标基下的分量为 $\mathrm{d}x^\mu/\mathrm{d}x^\nu = \delta^\mu_\nu$. 取另一矢量 $\boldsymbol{u} = u^\nu\boldsymbol{e}_\nu$, 则

$$\boldsymbol{v}\cdot\boldsymbol{u} = v^\mu u^\nu \boldsymbol{e}_\mu\cdot\boldsymbol{e}_\mu = \delta_{\mu\nu}v^\mu u^\nu.$$

可见矢量内积 *(inner product)* 与度规是有关系的. 在一般情形下, 我们就可以用度规 $g_{\mu\nu}$ 定义两矢量的内积, 进而给出了夹角信息. 因此我们说, 实际上定义度规就是在定义该几何体系中两矢量的内积.

对比间隔不变性和度规的定义, 可以迅速发现 $\eta_{\mu\nu}$ 可看作度规, 且是伪 Riemann 度规. 既然度规决定了该体系的所有几何知识, 而一方面间隔不变决定了狭相的所有物理知识, 我们就可以说, 狭相可以等价于一门几何学, 这门几何就按“主人”取名[9]为 Minkowski 几何, 简称闵氏几何, 与欧氏几何相对照. $\eta_{\mu\nu}$ 就叫做闵氏度规.

2.4 单位制转换

物理学是关乎实验测量的学科, 讨论物理量的单位是极为重要的. 在 SI 中, 我们认为时间、空间乃不同物理量, 因而自然不会谈及“1s 等于多少 m”这种看似奇怪的问题. 但在相对论中, 这个问题很有意义, 因为时空乃统一的概念. 设 SI 下 c 近似等于 3×10^8 m/s, 因此令 $c=1$ 这件事就相当于说

$$1\,\mathrm{s} = 3\times10^8\,\mathrm{m}.$$

[9]历史可见 [?].

这会与 SI 有任何冲突吗？不会, 因为在 SI 中不会讨论“1s 等于多少 m”这种问题, 因此我们说，这是 SI 中一个可以利用的自由度. 在以后学习的广义相对论中，还经常涉及引力常量，把它们的数值都取为 1，即

$$c = G = 1,$$

可简化大量公式的书写，这就是相对论所使用的几何单位制 *(system of geometrized units)*，简称 SG. 在 SI 中 G 近似等于 $6.67 \times 10^{-11}\,\mathrm{m}^3/(\mathrm{s}^2 \cdot \mathrm{kg})$，因此该操作相当于

$$1\,\mathrm{kg} = \frac{6.67 \times 10^{-11}}{(3 \times 10^8)^2}\,\mathrm{m},$$

可见这利用的是 SI 的另一自由度. SG 便统一了长度、质量和时间的单位. 不涉及引力的量子理论经常使用如下的自然单位制 *(system of natural units)*，简称 SN：

$$c = \hbar = 1,$$

这里约化 Plank 常量 $\hbar$ 在 SI 中近似等于 $6.58 \times 10^{-22}\,\mathrm{MeV} \cdot \mathrm{s}$. 然而在计算物理量的数值时几何单位就不再方便，因此我们要研究物理量、物理公式如何从几何制转换到其它单位制. 一般来讲，所研究物理量在当前单位制是带上单位的. 比如，SG 下粒子物理常用 eV 作为质量单位，因为质能方程在 SG 下统一了质量和能量单位. 欲将质量单位从 SG 还原为 SI，无非是在问 eV 等于多少 kg. 首先都转化到常用单位来，即

$$1\,\mathrm{eV} = 1.6 \times 10^{-19}\,\mathrm{C} \cdot \mathrm{V} = 1.6 \times 10^{-19}\,\mathrm{J} = 1.6 \times 10^{-19}\,\mathrm{kg} \cdot \frac{\mathrm{m}^2}{\mathrm{s}^2},$$

代入 SG 定义有

$$1\,\mathrm{eV} = \frac{1.6 \times 10^{-19}}{(3 \times 10^8)^2}\,\mathrm{kg}.$$

然而，若所研究物理量在当前单位制只是一个数字，则其通常只是在当前单位制下无单位，但在其它单位制中有单位，因此严格来说不能称其无量纲. 真正无量纲的量在任何单位制下都无单位，因此不受单位转换的影响，如精细结构常数

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137},$$

代入 SN 定义得 $e^2/4\pi\epsilon_0 = 1/137$，这说明兼容 SN 和精细结构常数的同时仍有自由度可利用. 令

$$c = \hbar = \epsilon_0 = 1,$$

因而 $e = \sqrt{4\pi/137}$. 由 $c = 1/\sqrt{\epsilon_0\mu_0}$ 得 $\mu_0 = 1$. 如上规定就称为 *Heaviside-Lorentz* 单位制，简称 SHL. 在 SHL 下不用再写出 Maxwell 方程中的常数. SG 和 SN 还可

共同构成 *Plank* 单位制，简称 SP，这在量子引力理论中用得多. 最后，对一个整体的物理公式而言，最便捷的做法不是如上这样分析单位，而是分析量纲，即根据量纲补充缺失的常数. 另外，有必要说明这样一种思想：有别于传统认知，物理公式应当视作“数”的公式而非“量”的公式. 具体细节放入附录 **??** 中. 简单来说，正如定义 $x^0 = ct$ 一样，亦可定义 SG 下的固有时为 $\tau' = c\tau$，物理方程中就再也不会出现 c，因为它全部被收进了时间里，类似于将 G 收进场源质量里，并不影响其它任何物理量的计算. 从重新定义的角度看，τ' 的单位同距离单位彻底一致，时间和距离“共享”同一单位，因而 $c = 1$ 是真正的“数”；从单位转换的角度看，这些数据并不需要收进任何一个物理量中，而是直接收进单位的相对关系里.

2.5 闵氏几何

2.5.1 闵氏时空

我们有必要先讨论 $\mathbf{R}^n$ 这种写法的由来. 在集合论定义了集合 *(set)* 和元素 *(element)* 此二概念后，可进一步定义如下四种常见运算[10]：

$$
\begin{aligned}
A \cup B &= \{x | x \in A \text{ 或 } x \in B\} && \text{并集}, \\
A \cap B &= \{x | x \in A \text{ 且 } x \in B\} && \text{交集}, \\
A \setminus B &= \{x | x \in A \text{ 且 } x \notin B\} && \text{余集、差集或补集}, \\
A \times B &= \{(x, y) | x \in A, y \in B\} && \text{卡氏积}.
\end{aligned}
$$

其中承认逻辑学的自然语言表达. 这里读者可能不熟悉的是第 4 种——*Descartes* 积，简称卡氏积，它不外乎是将两个集合的元素打包在一起，组成一个括号. 这种写法当然不是唯一的，实际上应当按照 x, y 各自的所有可能“相乘”，因为它们之间“独立”，就好像我们在排列组合学中看到的那样. 于是看到，在用分析学定义了实数集 $\mathbf{R}$ 后，就可以像这样复合出一个“2 维”集合：$\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$. 3 维欧氏空间就是在 $\mathbf{R}^3$ 上处处配以欧氏度规 δ_{ij}（构成度规场）. 可 δ_{ij} 具体来讲究竟是什么意思？类比矢量 $\boldsymbol{v}$ 分量需借助一组坐标基来确定，按理来说 δ_{ij} 是某个“东西” $\boldsymbol{\delta}$ 的分量，也需借助坐标系及其上坐标基，然而 $\mathbf{R}^3$ 上的坐标系无穷多，我们并未事先约定. 不同坐标基很可能给出不同分量. 比如，线元在直角坐标系下是

$$\mathrm{d}\ell^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \delta_{ij} \mathrm{d}x^i \mathrm{d}x^j,$$

而在球坐标系下却是

$$\mathrm{d}\ell^2 = \mathrm{d}r^2 + r^2 \, \mathrm{d}\theta^2 + r^2 \sin^2\theta \, \mathrm{d}\phi^2 = \delta'_{ij} \mathrm{d}x'^i \mathrm{d}x'^j,$$

[10]对于结构，诸如集合、群等，通常能以直接的方式定义“子结构”，例如子集、子群等. 故将不赘述这一点.

很明显有不同分量 δ_{ij},δ'_{ij}. 当然，坐标系变换不改变线元，我们可以说 $\boldsymbol{\delta}$ 是与坐标系无关的深刻事物，变化的仅是各分量，其间按链式法则有关系

$$\delta'_{ij}=\delta_{kl}\frac{\partial x^k}{\partial x'^i}\frac{\partial x^l}{\partial x'^j}, \tag{2.5.1}$$

因此严格来讲只有 $\boldsymbol{\delta}$ 才叫做度规. 实际上，如果在 $\mathbf{R}^3$ 的某一点找到了一组坐标基使 $\boldsymbol{\delta}$ 的分量为 δ_{ij}，以此坐标基延伸出的直线便可构成 $\mathbf{R}^3$ 上整体的坐标线，并确定出 $\mathbf{R}^3$ 上各点的坐标基. 如果场 $\boldsymbol{\delta}$ 在各处坐标基（即在此整体坐标系下）仍具备分量 δ_{ij}，我们便可确定此 $\boldsymbol{\delta}$ 的确为欧氏度规场.

考虑到 $\mathbf{R}^n$ 是最简单的空间，而欧氏空间可记作 $(\mathbf{R}^3,\boldsymbol{\delta})$，我们有理由假定闵氏时空就是 $\mathbf{R}^4$ 配以闵氏度规场 $\boldsymbol{\eta}$ 并记作

$$\mathbf{R}^{3+1}=(\mathbf{R}^4,\boldsymbol{\eta}). \tag{2.5.2}$$

$\boldsymbol{\eta}$ 的定义可类比前文，相应坐标系称为惯性坐标系或 *Lorentz* 坐标系，称 t 为坐标时. $\boldsymbol{\eta}$ 的分量同理满足变换关系

$$\eta'_{\sigma\lambda}=\eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x^\nu}{\partial x'^\lambda}. \tag{2.5.3}$$

这与先前的 (2.2.9) 有微妙区别. 在一般的坐标变换下，分量 $\eta_{\mu\nu},\eta'_{\mu\nu}$ 可以不同，而 (2.2.9) 实际上规定的是一类特殊的变换（Lorentz 变换），使得若 $\boldsymbol{\eta}$ 在一个坐标系中表为 $\eta_{\mu\nu}$（因而是惯性坐标系），则满足 (2.2.9) 的新坐标系亦为惯性坐标系，可见 $\mathrm{d}s^2=\mathrm{d}s'^2$ 包含了相对性原理和光速不变. 然而从操作性上讲，定义坐标系须通过物理观测与记录，才能确定给事件赋予何种坐标. 干这件事的人或仪器在数学模型上视作一个质点，称为观者 *(observer)*. 为了观测，其手持一部走时准确的标准钟 *(standard clock)*，该钟读数正是其固有时. 除此之外，还需处处配以 *4-标架 (tetrad)*，也即其上任意点矢量的正交归一基 $\{\boldsymbol{e}_\mu\}$，满足 $\boldsymbol{e}_\mu\cdot\boldsymbol{e}_\nu=\eta_{\mu\nu}$. 在生活中，*3-标架 (frame)* 可以是由三根单位长的短直杆或刻度尺焊成的、两两正交的架子，每根直杆代表一个观测方向，它们在每一时刻的指向由该观者选定. 当然，读者亦可尝试用右手大拇指、食指及中指按右手螺旋比划出 3-标架. 在数学上，3-标架抽象为空间正交归一基 $\{\boldsymbol{e}_i\}$，“空间”指皆与观者世界线切矢 $\boldsymbol{e}_0$ 正交. 今后谈及 4-标架时默认为右手标架. 规定 $\boldsymbol{e}_0$ 所指方向为未来 *(future)*，反方向为过去 *(past)*.

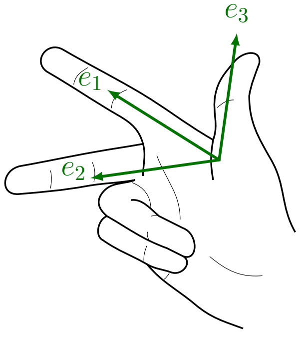
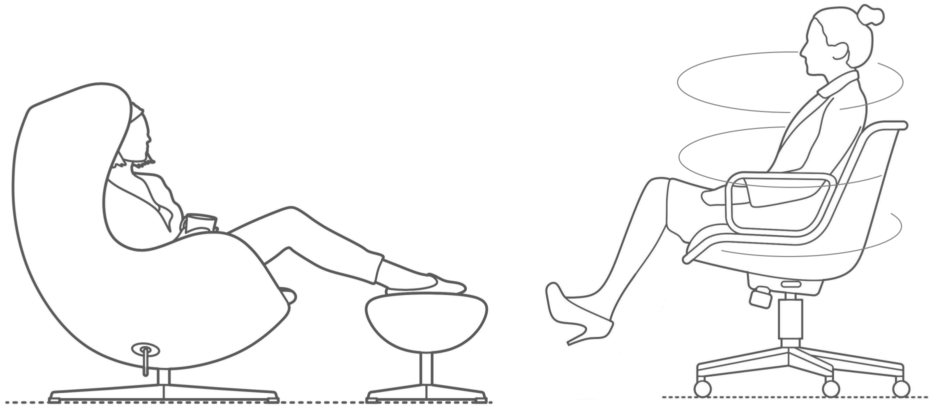


图 2.6: 空间右手 3-标架

惯性观者[11]就是世界线与某一惯性坐标系 t 坐标

[11]尽量和 Newton 力学意义的惯性观者相照应（第一定律），即在所引进的参考系中，自由粒子之轨迹或是一点，或是一条以不变速度运动的直线，这样在 $\mathbf{R}^{3+1}$ 中世界线就是直线.



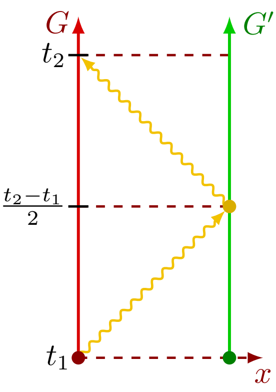
(a) Cynthia (b) Kaylor

图 2.7: Cynthia 可视为惯性观者而 Kaylor 不能

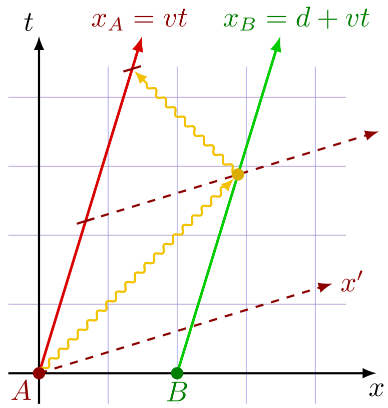
线重合的观者．这样惯性观者的 4-标架又称*惯性基矢*或*惯性标架*．我们还应希望惯性标架*无自转*，否则可能观测到赝力．如图 2.7，设 Cynthia、Kaylor 两人坐在地面的两把椅子上：Cynthia 坐底座固定的蛋壳椅；Kaylor 坐底座静置的办公椅且不停自转．二者世界线都走直线，但 Cynthia 可视为惯性观者而 Kaylor 不能．请注意，虽然观者概念本身已要求把此二人看成没有大小的点（于是由一世界线代表），但自转涉及的仅是线上各点每一空间基矢的方向沿线是否改变，故仍有明确意义（准确定义见 **??** 节）．然而在惯性观者所对应的惯性坐标系中，若令 4-标架 $\{\boldsymbol{e}_\mu\}$ 就是相应的坐标基，则直观看来就是无自转的．一般默认这一点．

严格来讲，一个观者只能对发生在自己世界线上的事件做*直接观测*．因而 4-标架场又称作观者所处处配备的“局部实验室”，所测物理量无非就是将相关的 4 维量投影到 3-标架上．欲对一定时空范围内的事件进行观测，要么需要处处设置观者，要么需要依靠信号传播[12]，我们暂时只谈前者．这样，一族观者世界线所组成的线汇 *(congruence)* 就构成一个参考系．可见参考系、坐标系应是不同概念．但我们往往选择让 t 的坐标线与参考系中某个观者的轨迹相重合，因此参考系和坐标系将紧密联系．一系列惯性观者就构成*惯性参考系*．在不必认真区分此二概念时，笼统称为*惯性系*未尝不可．一个有用的惯性系还应有如下要求．设某一观者位于坐标原点，光子可相对于该坐标系在任意方向上沿直线运动．注意，所谓实验事实上，“惯性系均承认相同的真空光速数值”，应理解为双程平均真空光速，而非单向真空光速．也即，应通过如下的 *Fizeau 程序*确定真空光速 c：时钟读数为 t_1 时，从原点发射一个光子，

[12]一般用光，因为光速不变是优良性质．可见光能给予视觉感受．



(a) 对钟



(b) 动系对钟

图 2.8: 对钟以确定同时面

然后由放置在 (x, y, z) 的镜子反射回来；时钟读数为 t_2 时，光子回到原点，测得

$$c = \frac{2\sqrt{x^2 + y^2 + z^2}}{t_2 - t_1}.$$

目前只有原点一只标准钟，因此只能观测原点处发生的事件. 为能够观察到一定范围内的情况，即建立一个参考系，就应该在每个空间点上都放置一只走时率 *(rate)* 全同的标准钟（此标准之含义），但起点 *(setting)* 却不一定一致. 然而我们在日常生活中总是会感觉到周围物体与自己“同时”，这相当于在“截取”时空时，“扔掉”时间，而只观察某时刻所有同时事件所构成的 3 维空间. 这在时空图中可以用一张 2 维平面 Σ_0 表示，因此称为同时面 *(simultaneity surface)*. 因此我们要求，在建立惯性参考系时，与所有观者世界线正交的同时面上应当有一致读数，这就是对钟 *(clock synchronization)*. Fizeau 程序提供了一种对钟方法：可通过从原点发射的球面电磁波（所有方向上都有光子传播）来同步各点时钟，即 (x, y, z) 点接收到这个电磁波时，该处时钟的读数就应调到

$$\frac{t_2 - t_1}{2} = \frac{\sqrt{x^2 + y^2 + z^2}}{c},$$

这样就对好了空间中所有的钟. 于是观者现在可以在他的参考系中，给发生在当地的事件指定 3 个空间坐标及 1 个时间坐标. 惯性系的时间坐标就称为惯性坐标时. 至今为止，单向光速从未被实验测量过，也不可能实验测量，除非找到光信号之外的对钟方式. 真空光速在各个方向上相同是假定性的原理. 观者起初在某种距离单位上测量时间 (如时针的轨迹)，距离均匀地走过其时间坐标 t 乘以 c. 于是，若定义时间坐标为 $x^0 = ct$，则时空的四个坐标将具有一致量纲. 进一步选取 SG 使得 $c = 1$，后

文将默认这一点. 对好钟后, 由正交性知, 惯性参考系确定出的坐标系才的确是惯性坐标系.

设矢量 $\boldsymbol{v}=v^\mu\boldsymbol{e}_\mu,\boldsymbol{w}=w^\nu\boldsymbol{e}_\nu$, 立即有 $\boldsymbol{v}\cdot\boldsymbol{w}=\eta_{\mu\nu}v^\mu w^\nu$. 若矢量 v 模方 $\eta_{\mu\nu}v^\mu v^\nu$ 为正、负或零, 那就分别称为类空 *(spacelike)*、类时 *(timelike)* 或类光 *(null)* 矢量. 所有类光矢量之集 C_N 称为光锥 *(light cone)*. $\mathbf{R}^{3+1}$ 中的事件可以视作 4 维位矢. 设两个不同事件位矢 $\boldsymbol{x}=x^\mu\boldsymbol{e}_\mu,\boldsymbol{x}_0=x_0^\nu\boldsymbol{e}_\nu$, 其差也是一个 4 维矢量, 即一个从 $\boldsymbol{x}_0$ 到 $\boldsymbol{x}$ 的矢量 $\boldsymbol{x}-\boldsymbol{x}_0$. 因此任意矢量可由 $\mathbf{R}^{3+1}$ 中事件点之差来定义. $x^\mu-x_0^\mu=\Delta x^\mu$ 定义为时空中的事件间隔, 因此严格来说线元是无穷小间隔. 考虑 2 维情况下做 boost 变换, 这样对任二不同事件的事件间隔, S 系所测与 S' 系所测之间的关系为

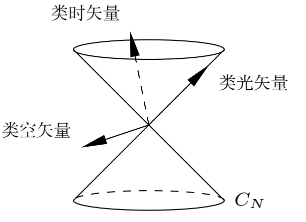


图 2.9: 3 维时空的光锥

$$\Delta t'=\gamma(\Delta t-v\Delta x),\quad \Delta x'=\gamma(\Delta x-v\Delta t). \tag{2.5.4}$$

如果 $\Delta t=0$, 则 $\Delta t'=-\gamma v\Delta x=-v\Delta x'\neq 0$. 这种现象正是同时的相对性. 现可用时空图简单说明之. 在平面中选取两条互相垂直的直线来代表 x 轴和 t 轴. 当然, 这两条线的欧氏垂直性没有任何物理上的必要, 但更直观一些. 而 x' 轴即 $t'=0$, 按照 boost 变换应是 $t=vx$. 同理, t' 轴就是 $t=(1/v)x$. 可见, 若以 S 系为基准, 则 S' 系的 t',x' 轴将关于类光方向对称. 这我们在对钟的时候已看到. 注意在推导 boost 变换时, (2.2.11) 式左矩阵对角元形似双曲线方程, 暗示我们可以定义

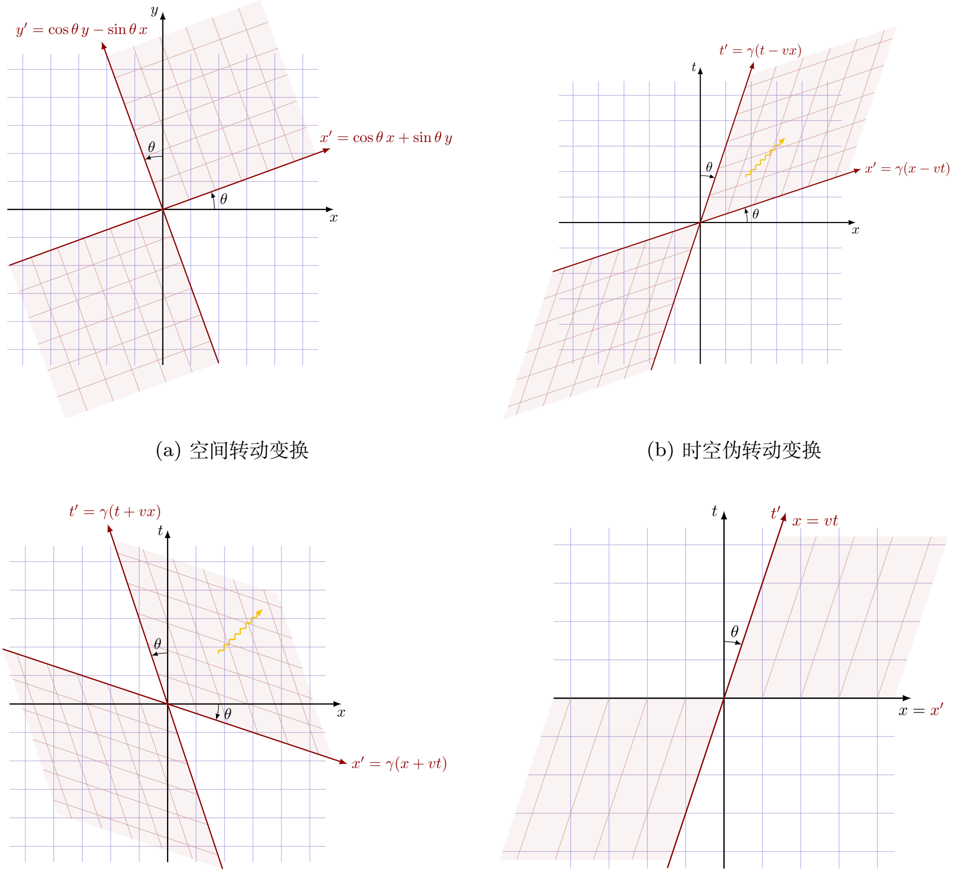
$$\cosh\theta=\frac{1}{\sqrt{1-v^2}},\quad \sinh\theta=\frac{v}{\sqrt{1-v^2}},$$

这相当于定义

$$v=\tanh\theta, \tag{2.5.5}$$

满足这种关系的 θ 称为快度. 用快度可形象地将轴的变化描述为一种 "旋转", 如图 2.10, 故 boost 变换又称为伪旋转变换. 在日常生活中, 光速相对于我们较大, 因此图上的光锥应当 "下压", 这时可见时间轴的变化并不明显, 我们退回到 Galilean 变换, 即大家都共用一个坐标时. 某事件的坐标值可作平行于另一轴的直线来投影得到. 例如, 某事件的 t 坐标可做平行于 x 轴的直线得到 (此即 S 系同时面). 而 t' 坐标可做平行于 x' 轴的直线而截得 (此即 S' 系同时面). 所以一般情况下, 在一个参考系同时的一对事件不会在另一参考系中同时. 注意, 这些同时面直线实际上表示 3 维空间, 故同时面交点实际是 2 维平面, 其上任意两个事件被两个观者都看做同时.

若 $\boldsymbol{x}-\boldsymbol{x}_0$ 类光, 则 $\Delta x^2=\Delta t^2$, 即事件 $\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 间的空间间隔等于光在事件间的时间间隔中传播的距离. 这对其他任意惯性基失一样成立, 因为 Lorentz 变换是保持度规不变的变换, 简称保度规 *(isometry)* 变换. $\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 可以 "由光子联系起来".



(a) 空间转动变换

(b) 时空伪转动变换

(c) 负方向伪转动

(d) Galilean 变换

图 2.10: boost 变换的几何直观

所有观者都会一致地看到两个事件中，哪一个是“发射”光子，哪一个是“接受”光子，因为可以证明，非零的类时或类光矢量在任意惯性基矢下，其 4 个坐标符号一致 (因为 ${\Lambda^0}_0 > 0$)．所以，$t - t_0$ 要么对任意惯性观者为正 (即 $\boldsymbol{x}_0$ 发生在 $\boldsymbol{x}$ 之前)，要么对任意惯性观者为负 (即 $\boldsymbol{x}_0$ 发生在 $\boldsymbol{x}$ 之后)．由于光子在惯性坐标系中沿直线运动，所以可以定义如下概念：若 $\boldsymbol{x} - \boldsymbol{x}_0$ 类光，则在 $\mathbf{R}^{3+1}$ 中过 $\boldsymbol{x}$ 和 $\boldsymbol{x}_0$ 的直线称为光子世界线，这条世界线被看作光子历史中经历的所有事件 (包括 $\boldsymbol{x}$ 和 $\boldsymbol{x}_0$) 之集.

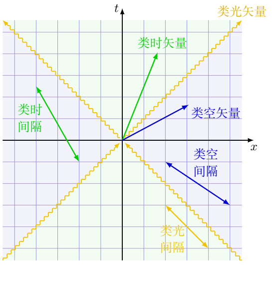
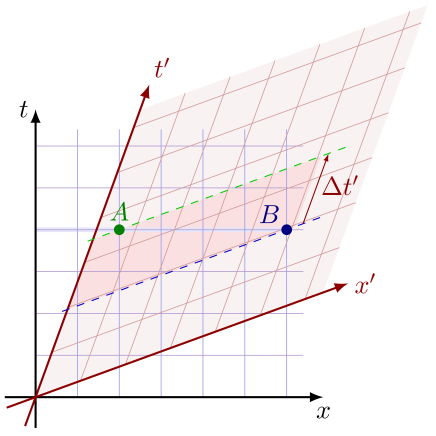
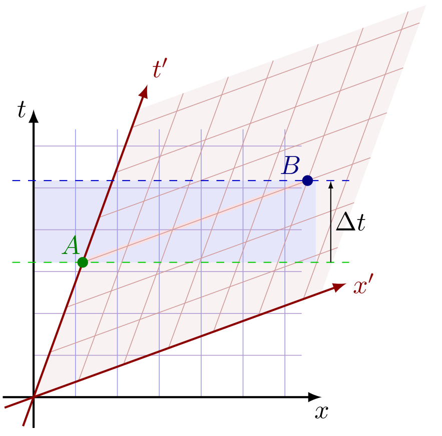


图 2.12: 2 维时空的光锥

若 $\boldsymbol{x} - \boldsymbol{x}_0$ 是类时矢量，则在任意惯性基矢下有 $\Delta x^2 < \Delta t^2$，此时必存在惯性基矢 $\{\boldsymbol{e}'_\mu\}$ 使得 $\Delta x' = 0$，即存在这样一个观者，在其看来，这两个事件一先一后地



(a) 在 S 中同时



(b) 在 S' 中同时

图 2.11: 同时的相对性

发生在空间中的同一位置. 设想该位置被某一质点所占据，如放在此处的观者时钟，我们将发现，事件 $\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 都由该质点经历，且 $\boldsymbol{x}-\boldsymbol{x}_0$ 的 "模长"，即对任意事件 $\boldsymbol{x}_0, \boldsymbol{x} \in \mathbf{R}^{3+1}$ 和任意惯性基，$\sqrt{-\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu}$ 就是由该质点携带的钟所记录的事件时间间隔. 对于其他任意惯性观者来说，这个质点看起来是自由而不受力的，因为其以不变速度沿直线运动. 这使我们可以有如下定义：若 $\boldsymbol{x}-\boldsymbol{x}_0$ 类时，则在 $\mathbf{R}^{3+1}$ 中过 $\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 的直线称为*自由质点世界线*，且

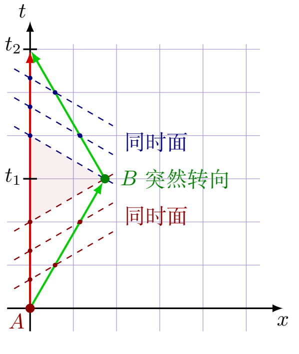


图 2.13: 双生子佯谬

$$\Delta\tau = \tau\left(\boldsymbol{x}-\boldsymbol{x}_0\right) = \sqrt{-\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu} \tag{2.5.6}$$

就是 $\boldsymbol{x}_0$ 和 $\boldsymbol{x}$ 间的固有时. 可见 $\tau\left(\boldsymbol{x}-\boldsymbol{x}_0\right)$ 就是 $\boldsymbol{x}-\boldsymbol{x}_0$ 的模长，衡量自身局部或内部演化的物理过程的时间. 这不同于欧氏几何的长度，其满足的不是欧氏三角不等式，而是如下 "反转" 形式：设 $\boldsymbol{x}, \boldsymbol{A}, \boldsymbol{B}$ 是 $\mathbf{R}^{3+1}$ 中的事件，且使 $\boldsymbol{x}-\boldsymbol{B}$ 和 $\boldsymbol{B}-\boldsymbol{A}$ 均为指向未来的类时矢量，则 $\boldsymbol{x}-\boldsymbol{A} = (\boldsymbol{x}-\boldsymbol{B}) + (\boldsymbol{B}-\boldsymbol{A})$ 亦是类时的，且

$$\tau\left(\boldsymbol{x}-\boldsymbol{A}\right) \geqslant \tau(\boldsymbol{x}-\boldsymbol{B}) + \tau\left(\boldsymbol{B}-\boldsymbol{A}\right), \tag{2.5.7}$$

当且仅当 $\boldsymbol{x}-\boldsymbol{B}$ 和 $\boldsymbol{B}-\boldsymbol{A}$ 共线时取等. 这将导致很有意思的事实：若两个质点都经历了 $\boldsymbol{x}$ 和 $\boldsymbol{A}$，则自由的那一个 (也即静止在某惯性坐标系中) 会需要更长的时间等候第二个事件发生，因为做变速运动的另一时钟整体走时较慢. 这一现象称为*双生子佯谬 (twin paradox)*，即双胞胎之一选择从地球（视为惯性系）启程，乘飞船在太空中遨游，凯旋时将比呆在地球上的那位更年轻.

对于任意类时世界线仍有双生子佯谬关系. $\boldsymbol{x}-\boldsymbol{x}_0$ 类时且指向未来，当且仅当存在一条类时世界线 α，其中 $\alpha(\tau_0)=x_0$ 和 $\alpha(\tau_1)=x$，且有 $T(\alpha)\leqslant\tau(\boldsymbol{x}-\boldsymbol{x}_0)$，这里当且仅当 α 是类时直线时取等. 这里 T 表示两点间一段世界线长或者说所经历的 *(elapsed along)* 固有时

$$T(\alpha)=\int_\alpha \mathrm{d}\tau=\int_\alpha \sqrt{-g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu}. \tag{2.5.8}$$

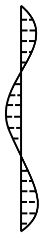


图 2.14: 示意图

这是因为，$\mathbf{R}^{3+1}$ 中任意两个事件间存在唯一的直线，这样就总可以这条直线建立惯性系，而连接两个事件的任意类时曲线都会比这条直线长（按闵氏度规衡量）. 可将惯性世界线分为若干微小片段来看到这一点. 对每一片段，竖直线上经历的固有时都是 $\mathrm{d}\tau'=\mathrm{d}t$，但做水平的同时线以截取曲线，则斜边所经历的固有时是 $\mathrm{d}\tau=\sqrt{\mathrm{d}t^2-\mathrm{d}x^2}$，故 $\mathrm{d}\tau<\mathrm{d}\tau'$.

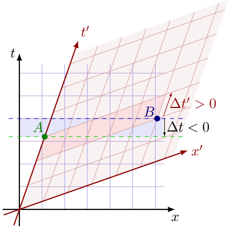


图 2.15: 类空间隔

考虑 $\boldsymbol{x}-\boldsymbol{x}_0$ 类空的情况，这时 $\Delta x^2>\Delta t^2$. 现在就不存在自由粒子或是光子能够经历这两个事件. 可以证明，对任意实数 T 都存在惯性基矢 $\{\boldsymbol{e}_\mu\}$ 使得 $\Delta t=T$. 如此，便很可能一些观者判定两个事件同时发生，另一些观者则断言 $\boldsymbol{x}_0$ 先于 $\boldsymbol{x}$，剩下的观者则反之. 这种类型的两个事件的先后次序没有意义，说明类空间隔无因果联系. 对于能够断定两个事件同时发生 ($\Delta t=0$) 的那些惯性观者而言，$\sqrt{\eta_{\mu\nu}\Delta x^\mu\Delta x^\nu}$ 这个量就是它们之间的空间距离，因而这个量称为固有空间距离 (无论 $\boldsymbol{x}-\boldsymbol{x}_0$ 是否类空).

2.5.2 运动效应

狭相所有熟知的运动学效应均可从以下讨论定量地获得. 设 θ_1 和 θ_2 是两个快度，考虑 boost 变换 $\Lambda(\theta_1)$ 和 $\Lambda(\theta_2)$.根据双曲函数的性质，有 $\Lambda(\theta_1)\Lambda(\theta_2)=\Lambda(\theta_1+\theta_2)$. 代入 $v_i=\tanh\theta_i$，以及 $v=\tanh(\theta_1+\theta_2)$，那么根据 $\tanh\theta$ 求和公式有

$$v=\frac{v_1+v_2}{1+v_1v_2}. \tag{2.5.9}$$

其物理解释很简单. 设存在 3 个惯性观者，其空间轴的设置如前文所述. 若第二个观者相对于第一个的速度是 v_1，而第三个相对于第二个的速度是 v_2，那第三个相对于第一个的速度就不是经典

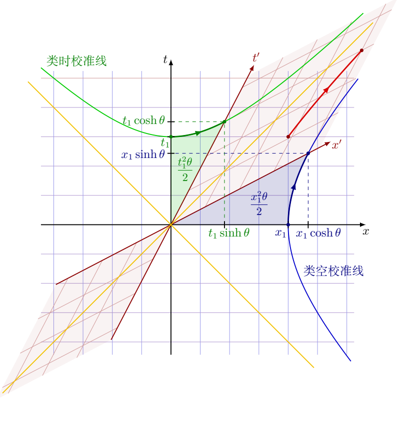


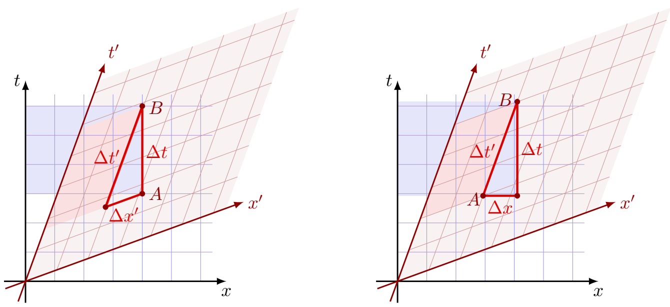
图 2.16: 校准线

力学的线性叠加所给出的 v_1+v_2，而是 $(v_1+v_2)/(1+v_1v_2)$．这就是相对论的速度叠加公式．

在时空图中比较线长的最佳方法是用双曲线校准．令双曲线 $x^2-t^2=x_1^2$，其与 x 轴交于 $(t,x)=(0,x_1)$．根据闵氏度规，双曲线上的点

$$t=x_1\sinh\theta,\quad x=x_1\cosh\theta$$

与原点连线是类空间隔且恒定，可用来比较类空间隔，故称类空校准线．从 $(0,x_1)$ 至 (t,x) 扫过角度 θ，易得扫过面积为 $x_1^2\theta/2$（类比欧式几何的圆）．同理推知类时校准线的结论．如此，考虑静止在 S 系中的原点，其世界线上的两个事件，由同一时



(a) 在 S 中静止 (b) 在 S' 中静止

图 2.17: 钟慢效应

钟分别测得两个读数．这种情况 $\Delta x=0$，所以

$$\Delta t'=\gamma\Delta t>\Delta t. \tag{2.5.10}$$

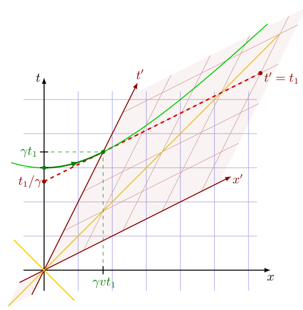
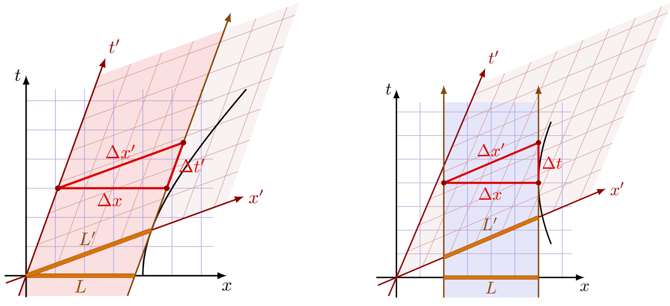


图 2.18: S' 系的钟根据同时面 $t'=t_1,t'=0$ 认为 S 的钟慢

这种效应完全对称，因为如果 $\Delta x'=0$，那么 boost 给出 $\Delta t=\gamma\Delta t'>\Delta t'$．每个观者都认为其他的相对运动的时钟走时慢．这种现象称为钟慢效应或时间膨胀．这在几何上用校准线便一目了然．例如，两个观者都看到 S 系原点上的时钟读数"0"，但直线 $t'=t_1$ 与该时钟世界线（t 轴）在 $(t,x)=(t_1,0)$ 下方的 $(t_1/\gamma,0)$ 相交．须强调，这种现象的物理意义是完全真实的．例如，在宇宙射线内发现的介子的寿命是如此之短，以至于即使以光速运动，它们穿透大气层所需时间也要比其通常寿命大几十倍．按理来说，其不可能到达地面，但其实还是可以到达，因为钟慢效应"使其保持年轻"，即在我们看来，其实际寿命比通常寿命长很多．

由于不同的惯性观者一般来说，对事件的同时性各抒己见，而测量一运动物体"长度"的唯一方法，就是要利用一把量尺"同时"测量物体两端；所以我们不应当

诧异，不同的惯性观者所测量的长度和时间的确会不同．考虑一把沿 S' 系 x' 轴静



(a) 在 S' 中静止　　(b) 在 S 中静止

图 2.19: 尺缩效应

置的测量尺子．在这个坐标系中它的“长度”是 $\Delta x'$．尺子两端的世界线是平行于 t' 轴的两条直线．若 S 观者在这些世界线上同时观测两个事件，那么他们的坐标将满足 $\Delta t = 0$，且由 boost 有 $\Delta x' = \gamma\Delta x$，则

$$\Delta x = \Delta x'/\gamma < \Delta x', \tag{2.5.11}$$

即量尺在其运动方向上缩短了 $1/\gamma$ 倍．这种现象称为*尺缩效应*或*长度收缩*．就像钟慢效应那样，尺缩效应完全对称且实在．沿 S 系 x 轴静置尺子的情况由上图道尽．

尺缩效应导致了一个有趣的*车库佯谬*[13]．设汽车与车库静长相等. 汽车朝着库匀速前进，司机认为“动库变短，不能放下”，司库认为“动车收缩，放下有余”．二人矛盾吗？为画出直观的时空图，设车库并无后墙 (其“后墙”类似于隧道洞口). 注意，画图时可借用校准曲线以保证车和库有相等静长．由图易见，以司库所在惯性系的同时面衡量，车短于库；以司机所在惯性系的同时面衡量，车长于库．两人同时面看法都对，关键是同时性的相对性导致结论的相对性．“到底放下还是放不下?”这种绝对式的问题没有意义，正如在尺缩问题中“到底哪把尺子较长”亦无意义．我们若真要“比较”，能让各位意见一致的做法只能是都放入

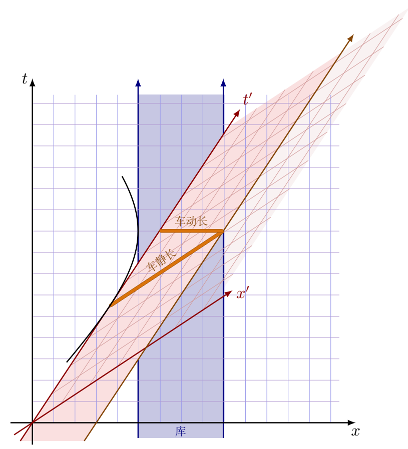


图 2.20: 汽车匀速进库

[13] 在英语世界里常称 ladder paradox，因为可借西式农耕文化里的梯子 (ladder) 和谷仓 (barn) 表述．一个经典的表述是用火车和隧道，这都大同小异.

静系中去. 当车库有坚硬后墙时，车头固然要撞墙，因而停止. 然而撞墙信息（以相互作用纵波）传至车身各部分乃至车尾皆需时间，只当车尾获悉后整个车身才能停住，因而汽车将被压缩到比司库期初所认为的长度还要短的程度. 这时，车的确在库中装下有余，谁看都装得下. 假设材料性质理想，使得信息传递为光速（在谁看来都是）且质点获悉信号后立即停止，有兴趣的读者不妨在图上补出撞车后的时空图.

2.5.3 动力学

设 $\mathbf{R}^{3+1}$ 中世界线 $\alpha(\tau)$ 在惯性系的表达式是 $x^\mu(\tau)$，则其速度矢量就是

$$\boldsymbol{U} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\boldsymbol{e}_\mu, \tag{2.5.12}$$

这称为 *4 维速度*，简称 *4-速*. 这样固有时参数使速度模方不变，且在 SG 下归一:

$$\eta_{\mu\nu}U^\mu U^\nu = -1, \tag{2.5.13}$$

说明 4-速为 α 的单位切矢. 当然，使其模长不变的参数当然有很多任意选择，统称为*仿射参数 (affine parameter)* λ. 容易发现，固有时参数就相当于线长参数（所经历的固有时为世界线线长），而固有时的任意线性函数仍属于仿射参数. 为使分量写法更类似于 $\boldsymbol{U}\cdot\boldsymbol{U}$，规定 $U_\mu = \eta_{\mu\nu}U^\nu$，这样便可写成 $U^\mu U_\mu = -1$，类似于前文的 $v_i = \delta_{ij}v^j$. 这种操作叫*降指标*，所得 U_μ 亦可看作矢量，但分量上不完全等同于原矢量，因为其时间分量带负号. 对 3 维矢量 v^i 无此问题，因为 $\eta_{ij} = \delta_{ij}$. 鉴于 U^μ,U_μ 的对应关系又称二者互为*对偶矢量 (dual vectors)*. 一般习惯将前者指为矢量而后者为对偶矢量. 同样我们可以*升指标*，那就是利用度规的逆 $\eta^{\mu\nu}$，即满足

$$\eta_{\mu\sigma}\eta^{\sigma\lambda} = \delta^\lambda_\mu. \tag{2.5.14}$$

升指标操作便可将对偶矢量变回原矢量.

考虑一个自由粒子，在 $\mathbf{R}^{3+1}$ 的任意惯性系中，其 4-速沿其世界线平移不变，即

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = U^\nu\partial_\nu U^\mu = 0, \tag{2.5.15}$$

其中 ∂_ν 表示 $\partial/\partial x^\nu$. 说明在 $\mathbf{R}^{3+1}$ 中自由粒子走直线. 结合 $U_\mu U^\mu = -1$ 可将其仿射参数唯一确定到线长参数. 设粒子相对于惯性系的 *3-速*为

$$\boldsymbol{v} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\boldsymbol{e}_i, \tag{2.5.16}$$

则易知

$$\boldsymbol{U} = \gamma\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\boldsymbol{e}_\mu = \gamma\left(\boldsymbol{v} + \boldsymbol{e}_0\right). \tag{2.5.17}$$

类比经典力学，*4-动量*在时空中可自然地定义为

$$\boldsymbol{P}=m\boldsymbol{U}, \tag{2.5.18}$$

这样甚至能统一能量（质量）和动量. 在任意惯性基矢下，有 $\boldsymbol{P}=\gamma m\,(\boldsymbol{v}+e_0)$，因此惯性系中 $\boldsymbol{P}$ 的空间部分是

$$\boldsymbol{p}=\gamma m\boldsymbol{v}, \tag{2.5.19}$$

这称为 *3-动量*. 记速率 $v=|\boldsymbol{v}|$，若将 $\gamma=1/\sqrt{1-v^2}$ 按二项式定理展开，则得到

$$\boldsymbol{p}=m\boldsymbol{v}+\frac{1}{2}mv^2\boldsymbol{v}+\cdots,$$

可见若 $v\ll 1$，那么 3-动量近似于 $m\boldsymbol{v}$. 若把 m 等同于经典力学中的惯性质量 (这是由相对于该粒子的速度很小的观者测量的)，则此即经典动量. 因此 3-动量给出了经典动量的相对论性修正. 再来看看 $\boldsymbol{P}$ 的时间分量，即

$$P^0=-\boldsymbol{P}\cdot\boldsymbol{e}_0=\gamma m=m+\frac{1}{2}mv^2+\cdots,$$

相应于 Newton 动能项 $(1/2)mv^2$ 的出现表明，P^0 是给定惯性观者对其测量的相对论性 (总) 能量，记作 E. 须知道，物理学中的 “能量” 概念是一个微妙的概念，把 P^0 称为能量没有带来任何物理内容. 这个名字是否合适只能由实验决定. 我们尤其应该问，上式出现的 m 这一项是否同 P^0 代表 “能量” 的观点相符. 观察 $v=0$ 的情况 (即粒子相对于观者静止)，则有质能方程 $E_0=m$，表明粒子即使静止也有 “静能”. 若这种静能真的是物理意义上的 “能量”，那它应该能释放出来并得到应用. 当然，事实上这也是可能的，并且已经被令人信服地证明了. 实际例子是家喻户晓的，比如核能. 最后，可证 4-动量模长恒定. 取其模长平方，一方面 $P^\mu P_\mu=mU^\mu mU_\mu=-m^2$，另一方面由 $\boldsymbol{P}=E\boldsymbol{e}_0+\boldsymbol{p}$ 和 $\boldsymbol{p}\cdot\boldsymbol{e}_0=0$ 可知 $P^\mu P_\mu=-E^2+p^2$，这里 $p=|\boldsymbol{p}|$. 故有如下的*质壳 (mass-shell) 关系*：

$$E^2=m^2+p^2. \tag{2.5.20}$$

惯性观者走直线，而更为任意的世界线便是做非惯性运动，这样仿照 3 维力学，可用如下的 *4 维加速度*，简称 *4-加速*刻画：

$$\boldsymbol{A}=\frac{\mathrm{d}U^\mu}{\mathrm{d}\tau}\boldsymbol{e}_\mu=U^\nu\partial_\nu U^\mu\boldsymbol{e}_\mu, \tag{2.5.21}$$

因此我们说直线满足 $\boldsymbol{A}=\boldsymbol{0}$. 易知 $\boldsymbol{A},\boldsymbol{U}$ 正交，这是因为

$$U_\nu A^\nu=U_\nu U^\mu\partial_\mu U^\nu=\frac{1}{2}U^\mu\partial_\mu(U_\nu U^\nu)=0, \tag{2.5.22}$$

这说明 $\boldsymbol{U}$ 模长恒定而只改变方向，当然加速度与其垂直（伪转动），这与经典力学的直觉相符. 取惯性坐标系，*3-加速*可表为 $a^i=\mathrm{d}v^i/\mathrm{d}t=\mathrm{d}^2x^i/\mathrm{d}t^2$. 下面研究 4-加

速同 3-加速关系. 首先将 4-加速写为 $A^\mu = \gamma\,\mathrm{d}U^\mu/\mathrm{d}t$，其时间分量为

$$A^0 = \gamma\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \gamma^4 v\frac{\mathrm{d}v}{\mathrm{d}t} = \gamma^4(\boldsymbol{v}\cdot\boldsymbol{a}), \tag{2.5.23}$$

其中考虑到 $\boldsymbol{a}$ 在 $\boldsymbol{v}$ 上投影为 $\mathrm{d}v/\mathrm{d}t$；其空间分量为

$$A^i = \gamma\frac{\mathrm{d}(\gamma v^i)}{\mathrm{d}t} = \gamma^2\frac{\mathrm{d}v^i}{\mathrm{d}t} + v^i\gamma\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \gamma^2 a^i + \gamma^4(\boldsymbol{v}\cdot\boldsymbol{a})v^i. \tag{2.5.24}$$

这样 A 的模长满足

$$A^\mu A_\mu = -\gamma^8(\boldsymbol{v}\cdot\boldsymbol{a})^2 + \gamma^4 a^2 + 2\gamma^6(\boldsymbol{v}\cdot\boldsymbol{a})^2 + \gamma^8(\boldsymbol{v}\cdot\boldsymbol{a})^2 v^2 = \gamma^4 a^2 + \gamma^6(\boldsymbol{v}\cdot\boldsymbol{a})^2. \tag{2.5.25}$$

可见，$\boldsymbol{A} = \boldsymbol{0}$ 的质点相对于任意惯性系的 3-加速都为零.

同理，质点所受的 *4-力 (4-force)* 可定义为

$$\boldsymbol{F} = \frac{\mathrm{d}P^\mu}{\mathrm{d}\tau}\boldsymbol{e}_\mu = U^\nu\partial_\nu P^\mu\boldsymbol{e}_\mu, \tag{2.5.26}$$

容易验证 $F^0 = \mathrm{d}E/\mathrm{d}\tau$,$F^i = \gamma f^i$，其中 *3-力*满足 $f^i = \mathrm{d}p^i/\mathrm{d}t$. 若质量在运动中保持常数，则 $\boldsymbol{F} = m\boldsymbol{A}, \boldsymbol{f} = m\boldsymbol{a}$，且还有类似的功能关系

$$\boldsymbol{f}\cdot\boldsymbol{v} = \mathrm{d}E/\mathrm{d}t\,, \tag{2.5.27}$$

这便再次验证 $E = \gamma m$ 的合理性.

2.6 作用量

第 **??** 章提到过测度和泛函. 二者具有一定相似性，因为我们可用函数或方程来描述集合. 比如，注意 3 维区域 D 同 ∂D 一一对应，而 ∂D 可由曲面方程表述. 在微积分学中，曲面的普遍表示方法是参数式，即三个二元函数 $x^i = f^i(s,\tau)$. 注意到曲面具有无穷自由度，可见一个测度往往又相当于一个泛函. 比如，时空中一段世界线所经历的固有时

$$T = \int_\alpha \sqrt{-g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} = \int_{t_0}^{t_1}\sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}\,\mathrm{d}t \tag{2.6.1}$$

是 α 所代表集合的测度，同时也是关于其参数式 $x^\mu(t)$ 的长度泛函，其中 $\dot{x}^\mu = \mathrm{d}x^\mu/\mathrm{d}t$. 这个参数 t 可以是时空的坐标时、仿射参数亦或固有时等.

我们知道微积分学中，某种函数在某点取最值的必要条件是函数在该点的导数为零. 但反之并不一定保证最值，甚至也非极值而是*鞍点 (saddle point)*. 那么如何对泛函求导呢? 这就涉及*变分法 (variation method)*. 设 α 是时空上的一条类时世界线，$x^\mu(t)$ 是其在某坐标系的参数式，作为研究目标的泛函可表为

$$S(\alpha) = \int_{t_0}^{t_1} L(x^\mu,\dot{x}^\mu)\,\mathrm{d}t\,, \tag{2.6.2}$$

其中 L 称作拉氏量 *(Lagrangian)*. 物理学很少考虑含时情形，即默认 $\partial L/\partial t=0$，因此没有写作 $L(x^\mu,\dot{x}^\mu,t)$. 欲找到点 $x^\mu(t_0),x^\mu(t_1)$ 间使得泛函取极值的路径 α，因此变分时唯一的限制就是使可能的解都固定两端，即端点是给定而不随曲线变更. 根据微积分学，极值的含义就是取邻近自变量时因变量近似不变. 设 α' 是极值曲线 α 的一条邻近曲线，其参数式与原先的差是微小的 $\delta x^\mu=x'^\mu-x^\mu$，但要求边界条件 $\delta x^\nu(t_0)=\delta x^\nu(t_1)=0$. 这个 δ 记号就称为变分. 极值曲线的变分所引起的泛函变分满足 $\delta S=0$，即相对于路径变化是高阶无穷小. 注意

$$\delta S=\int_{t_0}^{t_1}L(x^\mu+\delta x^\mu,\dot{x}^\mu+\delta\dot{x}^\mu)\,\mathrm{d}t-\int_{t_0}^{t_1}L(x^\mu,\dot{x}^\mu)\,\mathrm{d}t=\int_{t_0}^{t_1}\delta L\,\mathrm{d}t\,,$$

因此端点不变使得变分就像挪进了积分内一样. 类似于微积分学的思路，

$$\begin{aligned}\delta L&=L(x^\mu+\delta x^\mu,\dot{x}^\mu+\delta\dot{x}^\mu)-L(x^\mu,\dot{x}^\mu+\delta\dot{x}^\mu)+L(x^\mu,\dot{x}^\mu+\delta\dot{x}^\mu)-L(x^\mu,\dot{x}^\mu)\\&=\frac{\partial L}{\partial x^\mu}\delta x^\mu+\frac{\partial L}{\partial\dot{x}^\mu}\delta\dot{x}^\mu,\end{aligned}$$

可见变分的代数运算与微分极为相似. 其中的 $\delta\dot{x}^\mu$ 不便于处理，注意

$$\delta\dot{x}^\mu=\frac{\mathrm{d}}{\mathrm{d}t}(x^\mu+\delta x^\mu)-\dot{x}^\mu=\frac{\mathrm{d}}{\mathrm{d}t}(\delta x^\mu),$$

因此可交换顺序，进而利用分部积分法转化，再用变分限制消去边界项得

$$\int\frac{\partial L}{\partial\dot{x}^\mu}\delta\dot{x}^\mu\,\mathrm{d}t=\left.\frac{\partial L}{\partial\dot{x}^\mu}\delta x^\mu\right|_{t_0}^{t_1}-\int\delta x^\mu\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{x}^\mu}\,\mathrm{d}t=-\int\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{x}^\mu}\delta x^\mu\,\mathrm{d}t\,,$$

代回去就可提出 δx^μ 为

$$\delta S=\int\left(\frac{\partial L}{\partial x^\mu}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{x}^\mu}\right)\delta x^\mu\,\mathrm{d}t=0,$$

该式对任意变分 δx^μ 成立，当且仅当括号恒为零，因此

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{x}^\mu}-\frac{\partial L}{\partial x^\mu}=0,\tag{2.6.3}$$

此式称为 *Euler-Lagrange 方程*，简称 E-L 方程，它是为“变分为零”程序化操作专设的充要条件.

现代常用*作用量 (action)* 描述各类物理学领域的理论. 对于任一力学系统，通常都存在一个作用量泛函 S，它是沿各种可能运动轨迹的积分[14]，而符合自然规律的 S 应在系统的实际运动轨迹（经典轨道）上满足 $\delta S=0$. 换句话说，力学体系的运动轨迹由此确定. 比如，在经典力学中，粒子体系的作用量常取

$$S=\int_{t_1}^{t_2}(T-V)\,\mathrm{d}t=\int_{t_1}^{t_2}\left(\frac{1}{2}mv^2-V\right)\mathrm{d}t\,,\tag{2.6.4}$$

[14]包括经典上允许或不允许的轨道，这里“经典”是指不涉及量子性质.

这样代入至 E-L 方程即得粒子的运动方程 $m\dot{\boldsymbol{v}}=-\boldsymbol{\nabla}V$. 下面考虑自由粒子. 狭相涉及参考系变换，而作为客观规律的的作用量应当与之选择无关，亦即，作用量在 Lorentz 变换下不变. 作用量本身是标量，因此只能找 $\mathrm{d}\tau, m, c=1$ 这种内禀标量来构造. 我们得先给世界线找参数，可见只能取 $S\propto\int\mathrm{d}\tau$. 在量纲上最终的 L 应与能量一致，因此必须出现 $S\propto m\int\mathrm{d}\tau$. 注意在 $v\ll 1$ 时

$$m\int_{\tau_1}^{\tau_2}\mathrm{d}\tau=m\int_{t_1}^{t_2}\sqrt{1-v^2}\,\mathrm{d}t\approx m\int_{t_1}^{t_2}\left(1-\frac{v^2}{2}\right)\mathrm{d}t\,,$$

其中常数不影响变分，可见若想还原至自由粒子的动能，只能取

$$L=-m\sqrt{1-v^2},\quad S=-m\int\mathrm{d}\tau. \tag{2.6.5}$$

这便是自由粒子拉氏量的相对论修正. 读者容易验证，在经典力学中，*正则动量*可由拉氏量按 $p_i=\partial L/\partial v^i$ 给出 (它是经典动量的对偶). 由 $v=\sqrt{\delta_{jk}v^jv^k}$ 知 $\partial v/\partial v^i=v_i/v$，故将正则动量沿用至此得

$$p_i=\frac{\partial L}{\partial v^i}=\gamma m v_i, \tag{2.6.6}$$

可见 3-动量定义的合理性. 同样，在经典力学中，能量就是由拉氏量按 *Legendre* 变换 $H=2T-L=pv-L$ 给出的*哈氏量 (Hamiltonian)*. 容易验证确实有 $H=T+V$. 易知 $p^2=m^2v^2/(1-v^2)$，则 $v=p/\sqrt{p^2+m^2},\gamma=\sqrt{p^2+m^2}/m$. 故将哈氏量沿用至此得

$$E=H=pv+\frac{m}{\gamma}=\sqrt{p^2+m^2}=\gamma m, \tag{2.6.7}$$

可见 E 定义的合理性. 实际上，注意

$$L=-m/\gamma=\frac{mU^\mu U_\mu}{\mathrm{d}t/\mathrm{d}\tau}=mU_\mu\dot{x}^\mu. \tag{2.6.8}$$

则直接得到

$$\frac{\partial L}{\partial\dot{x}^\mu}=mU_\mu=P_\mu, \tag{2.6.9}$$

这正是 4-动量的对偶，足以见得整套体系的自洽性. 由前文知，在 $\mathbf{R}^{3+1}$ 中，能使 $S=-m\int\mathrm{d}\tau$ 满足 $\delta S=0$ 的显然是直线. 考虑固定初始事件 $\boldsymbol{x}_0$，其与任意事件 $\boldsymbol{x}$ 之间有且只有一条直线. 记 $\boldsymbol{x}_0$ 对应固定参数 t_0，而末端 $\boldsymbol{x}$ 对应任意 $t>t_0$，则在直线轨迹上有

$$\delta S(x^\mu)=\delta\int_{t_0}^{t}L\,\mathrm{d}t'=\left.\frac{\partial L}{\partial\dot{x}^\mu}\delta x^\mu\right|_{t'=t}.$$

可见 $\partial S/\partial x^\mu=\partial L/\partial\dot{x}^\mu$，进而 $E=H=-\,\partial S/\partial t$,$p_i=\partial S/\partial x^i$，与经典力学结论一致，并由此推知 *Hamilton-Jacobi 方程*在狭相中表为

$$\eta^{\mu\nu}\frac{\partial S}{\partial x^\mu}\frac{\partial S}{\partial x^\nu}=P^\mu P_\mu=-m^2. \tag{2.6.10}$$