相对论

目录

1 前言&记号约定

您可以在下面链接中下载此文档的最新版本：
`https://ghe0000.github.io/Cloud/notes`

- **粗体表示在高中范围内介绍过的内容，可能出现在地方卷题目之中**
- 蓝色字体表示推导的重点和关键概念，但高中不要求掌握

在高中阶段，有

- **狭义相对论适用于高速物体（一般认为大于光速10%）运动、在惯性系中**
- **广义相对论适用于强引力场、非惯性系中**

下文中$\frac{df(t)}{dt}$代表$f(t)$的导数，即$f'(t) = \frac{df(t)}{dt}$。对于dt，可认为是很小的Δt

2 狭义相对论的基本假设

一般而言，我们更愿意相信更简洁优美的解释。举个例子，你独自走在大街上，突然你的帽子被什么掀掉了，你愿意用一阵风还是火龙的尾巴来解释？（此梗出自“奥卡姆剃刀”原理）显然人们更愿意相信前者。物理学中也是如此，人们更愿意相信简洁，优美，对称的理论。人们相信，物理定律在不同惯性参考系下应保持不变。没有人可以通过实验判断自己是静止还是在某个惯性系中，没有人可以测出自己相对某个绝对静止惯性系的绝对速度。任何惯性系都是平权的，没有哪个惯性系是特殊的。这也就是相对性原理。

1887年，迈克尔逊（Michelson）和莫雷（Morley）用迈克尔逊干涉仪测量光速，结果发现，测量出的真空中的光速恒为c，也就是299792458m/s。这是难以置信的。这就好比如，你在路上看见一只小鸟以5m/s从你身旁飞过，于是你骑车以6m/s的速度去追它，可它仍旧离你越来越远，仍旧相对你5m/s。这件事可以用细思即恐来形容。但很明显，自然说了，光速测不变。

于是我们可以得出狭义相对论的两条基本假设：

- **一切惯性系平权，物理定律在不同惯性参考系下应保持不变，没有人可以测出自己的绝对速度（相对性原理）**
- **光速在不同惯性参考系下测出来均为c（光速测不变）**

3 时空间隔

3.1 距离

在高中物理一开始，我们就学习了位移的概念，而距离就是位移的模长。距离是两点之间的一种关系，而描述点最常用的方法就是他们的坐标，在直角坐标系下，我们可以很轻易地写出两个点的坐标为

$$A(x_A, y_A, z_A)$$
$$B(x_B, y_B, z_B)$$

于是我们便可以算出AB两点之间的距离为

$$|AB| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}$$

如果我们有另外一个直角坐标系，在这个坐标系中，A，B两点的坐标为

$$A(x'_A, y'_A, z'_A)$$
$$B(x'_B, y'_B, z'_B)$$

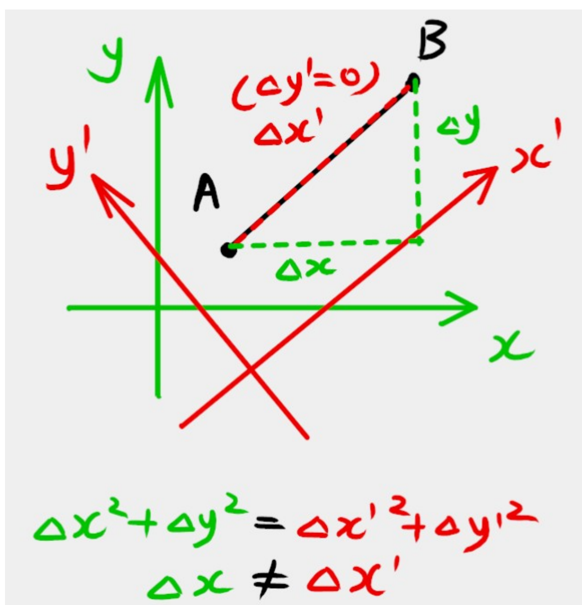

图 1: 距离

而距离的神奇之处，就在于

$$\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2} = \sqrt{(x'_A - x'_B)^2 + (y'_A - y'_B)^2 + (z'_A - z'_B)^2}$$

你可能会说，这不是废话吗？物体的长度是物体本身的属性，怎么会因为坐标系选取的不同而改变？但这时，我们注意到，在这两个坐标系中，$|AB|$在x轴上的投影在两个坐标系上并不相同，换句话说，如果我们发现物体的长度会随着坐标系的变化而变化，那便是因为我们没有考虑完所有的维度。

3.2 坐标

如果我们要描述一个事件，我们不但需要知道这个事件发生的位置x、y、z，还需要知道这个事件发生的时刻t，于是我们可以把这个事件在时空中的位置记为(t, x, y, z)，简记为$(t, \vec{r})$，其中$\vec{r} = (x, y, z)$。

但为什么在刚才我们只考虑了两个事件在x、y、z三个方向的差值，忽略了t的维度，便可以得到一个不随坐标系变化而变化的距离呢？这是因为在牛顿时空观中时间流逝的速度恒定不变，也就是说，两个事件在t方向上的差值不会随着坐标系变化而变化，也因此，我们才能在牛顿时空观下得到不随坐标系变化而变化的距离。

3.3 矛盾

迈克尔逊和莫雷的实验告诉了我们，真空中的光速恒为c。这本身与牛顿时空观相违背。假设一艘飞船相对地面运动速度为v，飞船向前进方向发射了一个光子。在飞船看来，光子的速度为c，光子运动了t秒，相对飞船运动的距离为ct。但在地面看来，光子的速度仍为c，光子相对飞船的速度为$c - v$，光子相对飞船的运动距离为$(c - v)t$，不等于ct。

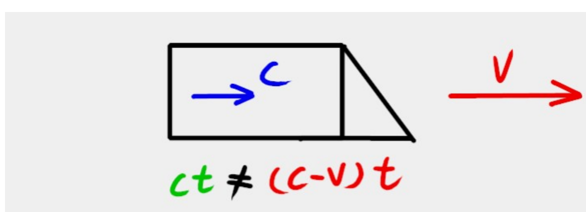

图 2: 经典时空观与光速不变的矛盾（绿：飞船、红：地面）

距离随着惯性系的改变而发生了改变！这也就意味着原本忽略的时间维度其实不能忽略，原本认为不变的时间流逝速度其实发生了改变。我们需要改写距离的表达式以考虑时间维度上

的间隔，以获得一个在惯性系变换下真正不变的距离。

3.4 时空间隔

由于长度和时间的单位不相同，所以我们并不能直接把时间维度上的间隔与空间维度上的距离直接利用勾股定理相加，他们之间有一个转换系数a，即$s^2 = r^2 + (a\Delta t)^2$，其中$r^2 = \Delta x + \Delta y + \Delta z$为空间上的距离。

于是，在飞船上，光子运动t秒的时空“距离”为$s^2 = (c\Delta t)^2 + (a\Delta t)^2$，在地面看来，时空“距离”为$s'^2 = c'\Delta t'^2 + (a\Delta t')^2$

由于我们要让这个“距离”在不同惯性系下保持不变，于是有$s^2 = s'^2$，因此有

$$(c\Delta t)^2 + (a\Delta t)^2 = c'\Delta t'^2 + (a\Delta t')^2$$

由于光速在不同参考系下不变，即$c = c'$，又由于时间的流逝速度不同，我们有$\Delta t \neq \Delta t'$，为使得上式恒成立，就必须有$a^2 = -c^2$，即$a = ic$。也就是说，时间和空间的比例关系是ic。光速c的特殊意义不在于他它是光的速度，而在于它是时间和空间之间的比例系数。换句话说，我们每次测量光速时，我们是在测量时间和空间之间的转换关系!

但为什么时间是虚数维呢？时间与距离的关系为什么会出现i？实际上，虽然时间不应该被看作完全独立的一维，时间与空间具有共同的特性，但时间仍旧是比较特殊的一维。这和我们的日常经验是相同的。日常生活中我们可以把球从空间中的一点抛向空间中任意的另外一点，却无法把球从未来抛向过去。如果时空之间的关系没有虚数单位i，时间和空间就没有什么区别了，过去和现在的概念也不复存在。比例是实数的世界，将比现在更难想象。

回到时空“距离”，根据上面的推导，我们可以得到时空“距离”的表达式，即

$$s^2 = r^2 - (c\Delta t)^2$$

我们一般将其称为时空间隔 [1]。我们可以看到，相对性原理体现在不同惯性系中时空间隔不变，光速不变则为时空间隔的一个特例，时空间隔很好地体现了两个基本假设。基于此，我们便可以推导出狭义相对论所有的结论。

---

[1]在一些书中时空间隔的表达式取了相反数，即

$$s^2 = (c\Delta t)^2 - r^2$$

这是为了保持一般情况下“间隔”大于零，取相反数并不影响间隔的性质。用广义相对论的术语来说，号差+2与−2等效。在此不取相反数只是由于习惯和统一性。

4 效应

4.1 同时的相对性

真空中光速不变，这一个有悖于常识的定理，而这也带来了一个“崭新”的观念：**“同时”是相对的**。

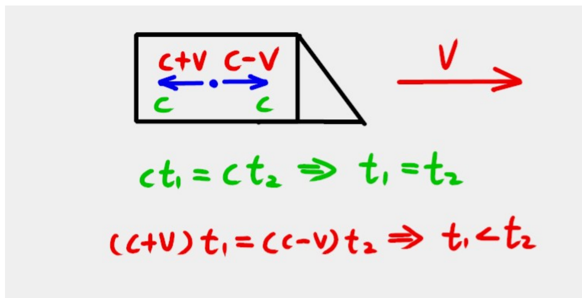

图 3: 同时的相对性（绿：飞船、红：地面）

我们假设飞船中间同时向飞船两边发射一个光子，在飞船上的人看来，两个光子以相同的速度c飞过了相同的距离，两个光子同时到达飞船两头。但在地面上观察，向飞船尾部发射的光子与飞船的相对速度为$c+v$，而向飞船头部发射的光子与飞船的相对速度为$c-v$，两个光子以不同的速度飞过了相同的距离，两个光子到达的时间不同，**原本在飞船上认为同时的事件在地面上变得不同时，而是一前一后发生的**。这与我们的日常经验严重冲突，可以说，理解狭义相对论的关键就在于理解“同时”的相对性。

4.2 动钟变慢

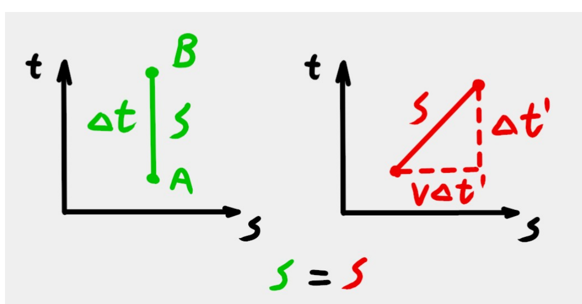

图 4: 动钟变慢时空图（绿：飞船、红：地面）

现在我们假设飞船上有一个相对飞船静止的光源，它每隔一个固定的时间就闪烁一下。在飞船看来，光源是静止的，两次闪烁之间的时间间隔为Δt，位移$r=0$，于是飞船上测得的两次闪烁之间的时空间隔为

$$s^2=-c^2\Delta t^2$$

在地面看来，光源是运动的，两次闪烁之间的时间间隔为$\Delta t'$，位移$r=v\Delta t'$，其中v是飞船相对于地面的速度。于是飞船上测得的两次闪烁之间的时空间隔为

$$s^2=v^2\Delta t'^2-c^2\Delta t'^2$$

上面两式联立，解得

$$\Delta t'=\gamma\Delta t$$

其中$\gamma=\frac{1}{\sqrt{1-\beta^2}}$，$\beta=\frac{v}{c}$，并且$\beta\leq 1$，$\gamma\geq 1$。

也就是说，**在地面看来，飞船上的一秒对应地面上的γ秒，飞船上的时钟变慢了，也就是动钟变慢。**

如果事件始终相对自己静止，也就是说r恒为0，那么时空间隔为$s^2=-c^2\Delta t^2$，也就是说，所有惯性系都可以得到相对事件静止的惯性系测得的时间间隔，我们称其为固有时，记为τ，于是有$s^2=-c^2\tau^2$。

4.3 动尺收缩

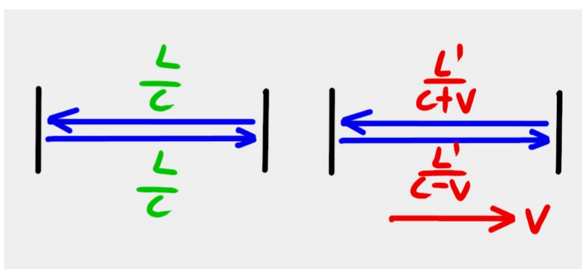

图 5: 动尺收缩（绿：飞船、红：地面）

现在我们假设飞船上有两面镜子，镜子的法向量平行于飞船前进方向，在飞船看来，两面镜子之间的距离为L，在镜子中有一个光子在这两面镜子之间往返，往返时间为Δt，于是有

$$\Delta t = \frac{2L}{c}$$

在地面看来，两面镜子之间的距离为L'，光运动方向与飞船前进方向同向时，$\Delta t_1' = \frac{L}{c-v}$，光运动方向与飞船前进方向反向时，$\Delta t_2' = \frac{L}{c+v}$，于是在地面看来，光往返一次的时间为

$$\Delta t' = \Delta t_1' + \Delta t_2' = \frac{2L'}{c^2 - v^2}$$

由前文的动钟变慢，有

$$\gamma = \frac{\Delta t'}{\Delta t} = \frac{\gamma^2 L'}{L}$$

故有

$$L' = \frac{L}{\gamma} = \sqrt{1-\beta^2}L$$

也就是说，在地面看来，飞船上看长1米的杆在地面看来长$\frac{1}{\gamma}$米，即**地面测出运动的飞船上的长度比原来要小，也就是动尺收缩。**[2]

4.4 双生子佯谬

根据前文的讨论，在地面看来，飞船上的钟比地面上的钟慢，但在飞船看来，飞船是不动的，地面（地球）是运动的，也就是说，在飞船看来，地面上的钟比飞船上的慢！那这时，我们便要问，究竟是哪个钟慢了？

举个例子，一艘飞船从地球出发，飞往火星，在火星调头，再飞回地球，这时候我们要问，是飞船上经过的时间长还是地面经过的时间长？

这个问题常常描述得很复杂，甚至用到了广义相对论，但实际上，我们可以通过时空间隔来简单讨论。我们可以画出两个参考系在时空中的"轨迹"，我们记地面参考系的“轨迹”为A，飞船参考系的轨迹为B。

[2]你可能会问，能不能用时空间隔来推导动尺收缩？可以，但由于要先推导出Lorentz变换，相对麻烦，故不在此使用

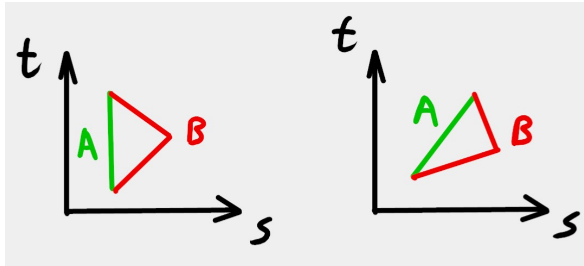

图 6: 双生子佯谬时空图（红：飞船、绿：地面、左：相对地面静止惯性系、右：相对地面运动惯性系）

我们可以看出来，无论在哪个惯性系，轨迹A的空间的路程都小于轨迹B，带入时空间隔的表达式，我们可以得到，无论在哪个惯性系，都测出轨迹B经过的时间小于轨迹A。

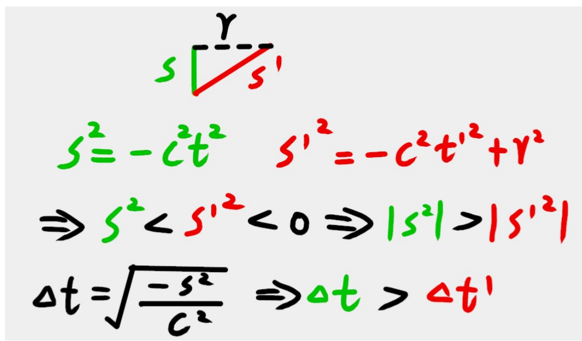

图 7: 轨迹A、B的微元（红：飞船、绿：地面）

也就是说，无论在哪个惯性系，都可以得到飞船经过的时间小于地面经过的时间。[3] 这在实验中得到了精密的验证。

5 相对论动力学

5.1 四维位移

根据前文的时空间隔，我们可以类比出四维下的位移，我们用下标μ来说明这个矢量是四维的，于是我们有

[3]你可能会问，如果观察者与飞船一起运动，相对飞船静止，那如何得到上述结论？对于这个问题，由于观察者不是惯性系（有加减速），故解决此问题要用到微分几何。时空间隔不变只能得出在任意惯性系中的结论

$$\vec{X}_\mu = (ict, x, y, z) = (ict, \vec{r})$$

其中$\vec{r} = (x, y, z)$为空间位移。可以注意到，四维位移的模长就是时空间隔，不随着惯性系改变而改变。之后我们得到的其他力学量都应该有这种性质。

我们现在把动力学中最基本的位移都改变了，因此我们需要改写其他力学量到四维形式。

5.2　四维速度与光速

速度是位移关于时间的导数，由于我们要让四维量的模长为惯性系变换下的不变量，所以时间要选用任何惯性系都能得到相同差值的时间，也就是固有时，记为τ，因此有$\vec{U}_\mu = \frac{d\vec{X}_\mu}{d\tau}$。对于矢量关于时间的导数，我们可以对矢量的各个分量分别对时间求导[4]。记观察飞船运动的我们测得的时间为t，由动钟变慢，有$\frac{dt}{d\tau} = \gamma$。于是我们有

$$\vec{U}_\mu = (ic\gamma, \gamma\vec{v})$$

我们可以求出$\vec{U}_\mu$的模方，即$\vec{U}_\mu^2 = -c^2$。也就是说，在四维时空观，世界万物的速度都是光速。并且这个结论与动钟变慢可以完美结合。我们记$\vec{v_s}$为空间维度速度，记$\vec{v_t}$为时间维度速度，记$s_t = v_t\Delta t = c\Delta\tau$，于是有

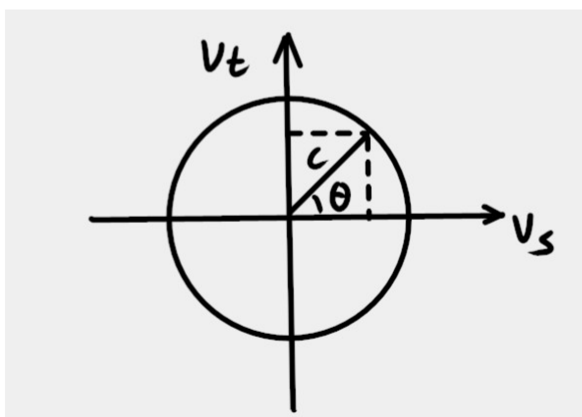

图 8: 四维速度

$$\vec{v_t} + \vec{v_s} = \vec{c}$$

[4]这里在此处没有问题，但如果基矢随着时间变化，那么我们必须求出基矢关于时间的导数

$$\beta = sin\theta = \frac{v_s}{c}$$
$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{cos\theta}$$
$$v_t = ccos\theta$$
$$\Delta t = \frac{c}{v_t}\Delta\tau = \gamma\Delta\tau$$

如果你认为静止不动，其实是你的四维速度全部分到了时间维度上；你认为其他高速运动的物体流逝时间变慢了，其实只是他们把四维速度分了一部分到空间维度上；而光子，只不过是把所有的速度都分到了空间维度，它的一瞬即是永恒。

也就是说，无论是静止的，高速的，光速的，它们在四维时空中都以光速c运动。静止的你和以光速飞行的宇宙背景辐射在四维时空观中都以光速c来运动，只不过你在时间上以光速飞行，而光子在空间上以光速飞行。

在牛顿时空观中，万物的速度方向不同，大小也不同。而在四维时空观中，世界万物的速度都是c，只不过速度的方向指向不同而已。由此观之，狭义相对论其实比经典的牛顿力学更简洁，优美，对称。

5.3 四维动量与质量膨胀

根据动量的定义，记m_0为静止物体的质量，于是有$\vec{P}_\mu = m_0\vec{U}_\mu$，即

$$\vec{P}_\mu = (icm_0\gamma, \gamma m_0\vec{U}_\mu) = (icm, m\vec{U}_\mu)$$

其中$m = \gamma m_0$，**也就是说，运动的物体质量增加**，质量膨胀效应如白送一般便可以推出来。

5.4 四维力与质能方程

根据力的定义，有$\vec{K}_\mu = \frac{d\vec{P}_\mu}{d\tau}$，即

$$\vec{K}_\mu = (ic\gamma\frac{dm}{dt}, \gamma\vec{F})$$

根据功的定义，我们有

$$dW_\mu = \vec{K}_\mu \cdot d\vec{X}_\mu = -c^2\gamma dm + \gamma dw$$

由于四维功W_μ不随惯性系变化而变化，故$dW_\mu = 0$ [5] 。于是有

$$dw = c^2dm$$

[5] 这里由于$\vec{X}_\mu$不随着惯性系改变而改变，因此在改变惯性系时$d\vec{X}_\mu = 0$，因此$dW_\mu = 0$，与一般含义不同，需要注意区分

两边积分，将做功改为能量，于是我们便可以得到大名鼎鼎的**质能方程**$E = mc^2$。

这个方程将能量和质量紧紧联系在一起，一个物体的质量反应了这个物体的能量。可以说，质量几乎成了能量的同义语。

至此，狭义相对论的力学部分基本推导完了。可以说，利用时空间隔推导，是最简洁也最深刻的。由于实验观测得到光速c是定值，我们便得到原本不变的距离发生了变化，于是我们考虑缺失的时间维度，得到新的真正不变的距离，并由此构建不随惯性系变化而变化的位移，速度，动量，力，功的四维形式，建立起狭义相对论力学。

6 走向广义之路

6.1 惯性系的困惑

现在我们回忆一下狭义相对论的两个基本假设

- 一切惯性系平权，物理定律在不同惯性参考系下应保持不变，没有人可以测出自己的绝对速度（相对性原理）
- 光速在不同惯性参考系下测出来均为c（光速测不变）

其中第一条强调了一切惯性系平权，狭义相对论是在探讨不同惯性参考系的理论。

这里我们遇到了一个问题，如何定义惯性系？你可能会说，物体不受力时相对惯性系做静止或匀速直线运动。那么，如何定义不受力？我们只能说，“不受力”意味着物体在惯性系中保持静止或匀速直线运动。这样的定义形成了一个循环论证。显然是没有用的。

另一个想法是，定义一无所有的空间中静止或做匀速直线运动的参考系为惯性系，但是，如果空间一无所有，我们根本无法标记和区分各种运动，任何参考系都建立不起来。

可见，惯性系的定义成为狭义相对论的一个困难。狭义相对论的整个理论都建立在惯性系的基础之上，但我们却无法定义惯性系。于是整个理论好像建立在沙滩之上。

6.2 万有引力的遗难

狭义相对论的另一个困难与引力有关。狭义相对论把运动学写成了四维的不随惯性系变化而变化的形式。然而所有把万有引力定律改写成四维形式的尝试都失败了，该定律无法与狭义相对论兼容。当时人们只知道电磁相互作用和万有引力相互作用，其中一种就与相对论不兼容，这显然不令人满意。

6.3 等效原理和广义相对性原理

由于无法定义惯性系，爱因斯坦想干脆取消掉惯性系在相对论上的特殊地位，认为光速不变和相对性原理在任何参考系上成立，于是狭义相对性原理推广成广义相对性原理，光速测不变也推广到所有参考系：

- **一切参考系平权，物理定律在任何参考系下保持不变，没人可以测出自己的绝对加速度。**
- **光速在不同一切参考系下测出来均为c。**

在解决上面两个问题的过程中，爱因斯坦抓住了“等效原理”这把“金钥匙”。而爱因斯坦关于升降机的实验，最清楚地体现了他的等效原理的思想。设想一个观察者处在一个封闭的升降机内，得不到升降机的任何外部信息。当他看到一切物体都以加速度a在下落，下落加速度a与与物体大小与物质组成无关，那么他无法断定自己属于下列情况的哪一种：

- 升降机静止在一个重力加速度为a的星球表面
- 升降在无引力场的太空中以加速度a运动

当观测者感到一切物体都处于失重状态时，那么他同样无法断定自己属于下列情况的那一种：

- 升降机在一个重力加速度为a的星球表面做自由落体运动
- 升降在无引力场的太空中静止或者匀速直线运动

造成上述现象的原因在于，力学中引力场与惯性加速场在小范围（局域）等效，即弱等效原理。

爱因斯坦则对上述原理进行了推广：**所有物理效应在引力场与惯性加速场的小范围（局域）中等效，即强等效原理**[6]。

6.4 引力的解释

在多年的思考之后，爱因斯坦写下了著名的爱因斯坦场方程：

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}$$

其中$R_{\mu\nu}$为Ricci曲率张量，R为Ricci曲率标量，$g_{\mu\nu}$为度规，它们用于描述时空的弯曲。$T_{\mu\nu}$为能动张量，描述物质的能量和运动。如同牛顿第二定律$F=ma$将力学和运动学向结合一样，这

[6]高中只强调引力与加速等效，不强调小范围

个式子将物质运动和时空弯曲紧密结合。相对论专家惠勒形象地称其为："物质决定时空如何弯曲，时空决定物质如何运动。"

爱因斯坦的广义相对论给了我们一个与众不同的、新颖的关于引力的解释。在爱因斯坦看来，不存在什么超距作用的引力，而是时空本身的弯曲"拖拽"着物体下坠。

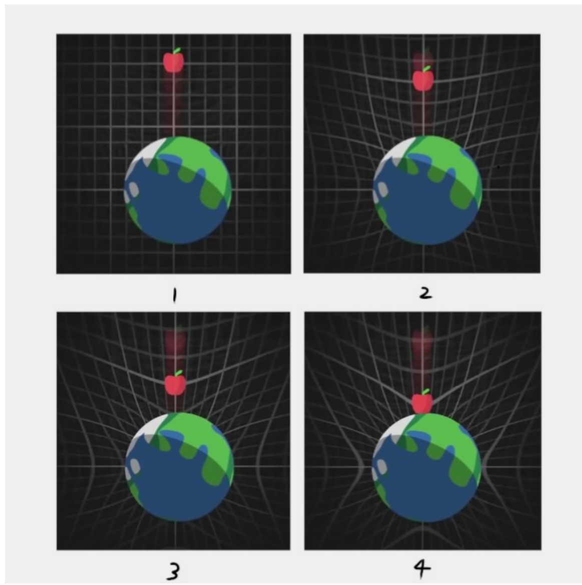

图 9: 广义相对论苹果下落

我们可以认为，时空由于地球的存在而一直在向地球中心"收缩"，地面上的观察者由于受到地面的压力而向背离地球中心的方向加速，由于时空"收缩"，观察者保持不动，于是感觉因为时空"收缩"在加速自由落体的自由落体受到了引力。这就好像在一个加速流动的河流之中，观察者的船受力而保持不动，看到飘浮在河上的物体在加速运动一样。用广义相对论的术语来说，就是测地线之间包含的体积由于曲率的存在随着时间而减小。

这里需要强调一点，时空并没有真正发生了"收缩"，是时间维度的曲率的存在使得弯曲空间中平行直线逐渐接近让我们感觉像是"收缩"。就比如在地球上，垂直于赤道的两条直线最终在极点相遇，在沿着直线往极点运动的观察者看来，平行直线逐渐靠近，空间像是在发生"收缩"，但实际上，空间并没有真正发生"收缩"，它只是我们的感觉而已。

并且，由于时间维度的曲率存在，我们还可以得到**引力场强的地方时间流速变慢**，也就是说，在塔底的钟比在塔顶的钟慢（忽略高度不同造成的速度差）

读者可能之前在其他地方了解过广义相对论，看见过如下经典的示意图。

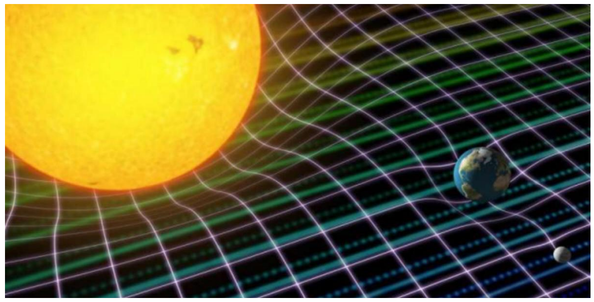

图 10: 广义相对论示意图

这张图片很好地体现了了引力是由于时空发生了弯曲产生的，但这张图仍旧有很多问题，如下

- 这张图给人以物质在时空之上的感觉，但实际上物质在时空之中
- 这张图在用“重力”解释“重力”，引入时空之外的某种“重力”是不可接受的
- 这张图给人二维的时空要在三维描述的感觉，但实际上，广义相对论并不需要除时间外额外的维度描述

因此，笔者认为，前文的解释相比该图具有优越性，测地线之间包含的体积由于曲率的存在随着时间而减小是可以通过计算得出的，该解释具有理论依据。

6.5 变化的“线元”

由于曲率的存在，狭义相对论的线元会发生改变。根据广义相对论，一个无自转的不带电的球体（地球可以近似为该球体，因为地球自转角速度很小），其周围空间的线元为：

$$ds^2 = -(1-\frac{2GM}{rc^2})c^2dt^2 + (1-\frac{2GM}{rc^2})^{-1}dr^2 + r^2(d\theta^2 + sin^2\theta d\phi^2)$$

由此，我们可以推导出光在太阳引力场的偏折、水星的进动、黑洞的视界、卫星与地面的时间修正，它们都得到了观测和实验的验证，都很好的验证了广义相对论。

对于卫星与地面的时间修正，由于在轨道上运动速度比地面快，根据狭义相对论时间流逝比地面慢；但引力场强比地面小，根据广义相对论的引力场减慢时钟效应，时间流逝比地面快。因此我们可以计算出，对于圆轨道，当轨道半径$r_0 \approx 9551km$，卫星的时间流逝速度与地面相同。当轨道半径低于r_0时，卫星的时间流逝速度比地球慢（对于中国空间站，每天约慢$25.4\mu s$，若不修正每天约能造成越$7.6km$的误差），当轨道半径高于r_0时（如在同步轨道上的北斗导航卫星），卫星的时间流逝速度比地面快。