
浅析狭义和广义相对论：一种拉格朗日表述

A Brief Analysis of Special And General Relativity ——The Lagrangian Formalism

何亦白

2024 年 11 月 8 日

Abstract

我们希望在本文中介绍一种能够用较少的数学知识来快速把握广义相对论的核心内容和观点的方法。我们首先通过对洛伦兹群和闵可夫斯基时空（下称闵氏时空）性质的研究重写闵氏时空中的作用量；然后通过定性分析得出爱因斯坦-希尔伯特作用量并由其导出爱因斯坦场方程；最后探讨广义相对论的替代理论：标量张量理论并推导其中最著名的 Brans-Dicke 理论的场方程。

Keywords 狭义相对论 · 广义相对论 · 最小作用量原理 · 爱因斯坦-希尔伯特作用量 · 爱因斯坦场方程 · 标量张量理论 · Brans-Dicke 理论

1 引言

在麦克斯韦发现麦克斯韦方程且经过无数的实验证实后，人们却发现其不满足长久以来所预设的物理理论在伽利略变换下协变的定理。亨德里克·洛伦兹首先发现了洛伦兹变换：在洛伦兹变换下，麦克斯韦方程组保持协变。庞加莱深化了洛伦兹变换，找到了洛伦兹变换的更一般的形式；最终爱因斯坦建立了相对论时空观，在其中所有的物理量在洛伦兹变换下都应是协变的而非在伽利略变换下协变。之后，爱因斯坦从等效原理出发，将引力融入狭义相对论的理论框架中，发展出了广义相对论。这一突破性的理论后由无数的观测证实，包括且不限于水星进动问题、星光偏折等。广义相对论涉及许多高深的数学理论，学习时较为困难；本文通过最小作用量原理导出了狭义相对论的基本内容和爱因斯坦场方程，某种程度上规避了复杂的数学内容并保留了对广义相对论添加修正项的可能性。最后本文简短的推导并研究了标量张量理论。

2 狭义相对论

2.1 洛伦兹群

狭义相对论的背景时空被称作闵可夫斯基时空，其几何结构由庞加莱群决定；庞加莱群是闵氏时空中的等距同构群。换句话说，庞加莱群是这么一种集合，其中的元素是保持闵氏时空中时空间隔不变的坐标变换矩阵。通过将闵氏时空中的向量略作修改，我们发现一个维持时空间隔的坐标变换关系可以写作：

$$\begin{pmatrix} \mathbf{x}' \\ 1 \end{pmatrix} = \begin{pmatrix} \Lambda & \Delta \mathbf{x} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \quad (1)$$

其中 Λ 代表了一个 $O(3, 1)$ 的元素，代表了闵氏时空中坐标架的旋转； $\Delta \mathbf{x}$ 是任意一个向量，表征了一个时空平移变换。 $O(3, 1)$ 正是我们常说的洛伦兹群。我们现在仅取 $O(3, 1)$ 中行列式为 1 的连通区域，并要求 $\Lambda^0_0 \geq 1$ ；满足这些要求的元素仍然组成一个群 $SO(3, 1)^\uparrow$ ，被称作固有正时洛伦兹群。我们一般所谓的洛伦兹群都是指这种固有正时洛伦兹群。洛伦兹群（或者说庞加莱群）保持时空间隔不变，即应该有关系

$$(\mathrm{d}s)^2 = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = \eta'_{\rho\sigma} \Lambda^\rho_\mu \mathrm{d}x^\mu \Lambda^\sigma_\nu \mathrm{d}x^\nu = \eta'_{\rho\sigma} \mathrm{d}x'^\rho \mathrm{d}x'^\sigma = (\mathrm{d}s')^2 \quad (2)$$

也就是说，时空间隔 $\mathrm{d}s = \sqrt{-\eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu}$ 是一个洛伦兹标量。

2.2 相对论下的作用量

拉格朗日力学中定义的作用量应当是一个洛伦兹标量，因为我们不希望坐标变化下粒子的运动方程发生改变。然而由于一般的作用量被定义为 $S = \int L \mathrm{d}t$ ，我们会发现拉氏量不是一个洛伦兹标量。这引导我们在闵氏时空中重写作用量。注意到闵氏时空中最自然、最简单的洛伦兹标量便是时空间隔；我们就可以取作用量为

$$S = -mc \int \mathrm{d}s. \quad (3)$$

我们可以通过参数化将 $\mathrm{d}s$ 写作如下形式：

$$\mathrm{d}s = \mathrm{d}\lambda \sqrt{-\eta_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}} \quad (4)$$

事实上，我们可以用任意参数来参数化时空间隔。定义固有时 $\mathrm{d}\tau^2 = -1/c^2 \mathrm{d}s^2$ ，那么我们就可以用固有时来参数化世界线，即将 (4) 中的 $\mathrm{d}\lambda$ 更改为 $\mathrm{d}\tau$ 即可。如此作用量就变成了

$$S = -mc \int \mathrm{d}\tau \sqrt{-\eta_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}}$$

为了验证这一作用量形式的正确性，我们在 (3) 中对 x^μ 变分：

$$\begin{aligned} \delta S &= -mc \int \delta \sqrt{-\mathrm{d}x^\mu \mathrm{d}x_\mu} \\ &= mc \int \frac{\mathrm{d}x_\mu \delta \mathrm{d}x^\mu}{\mathrm{d}s} \\ &= mc \frac{\mathrm{d}x_\mu}{\mathrm{d}s} \delta x^\mu \Big|_a^b - mc \int \delta x^\mu \frac{\mathrm{d}^2 x_\mu}{\mathrm{d}s^2} \mathrm{d}s = 0 \end{aligned} \quad (5)$$

这就说明自由粒子在闵氏时空中的运动方程必须满足 $\mathrm{d}^2 x^\mu / \mathrm{d}\tau^2 = 0$ 。这一微分方程的通解是时空中的直线方程，也就是说自由粒子在空间中静止或匀速直线运动。

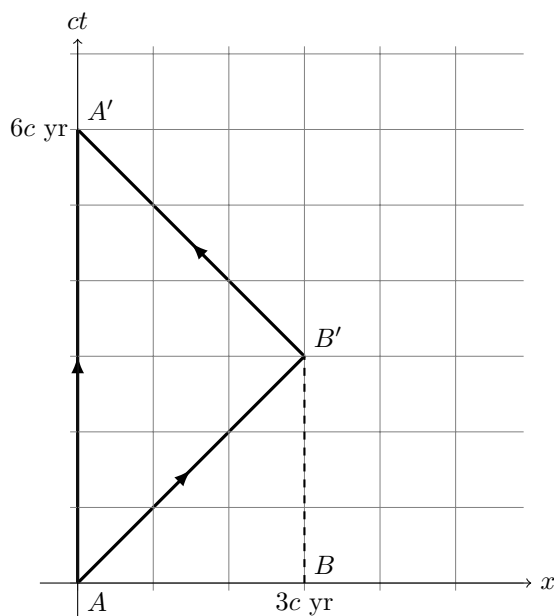
我们现在试图导出我们常见的洛伦兹因子 $\gamma \equiv 1/\sqrt{1-v^2/c^2}$ 。依旧考虑 (3) 中定义的作用量。我们将求和展开，得到

$$\begin{aligned} S &= -mc \int \mathrm{d}s \\ &= -mc \int \sqrt{-[c^2(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2]} \\ &= -mc^2 \int \mathrm{d}x^0 \sqrt{1 - \frac{1}{c^2} \left[\left(\frac{\mathrm{d}x^1}{\mathrm{d}x^0} \right)^2 + \left(\frac{\mathrm{d}x^2}{\mathrm{d}x^0} \right)^2 + \left(\frac{\mathrm{d}x^3}{\mathrm{d}x^0} \right)^2 \right]} \\ &= -mc^2 \int \mathrm{d}x^0 \sqrt{1 - \frac{v^2}{c^2}} \end{aligned} \quad (6)$$

所以有 $\int -mc^2 d\tau = \int -mc^2 \sqrt{1 - v^2/c^2} dx^0$ 。这样我们能得到 $dt/d\tau = 1/\sqrt{1 - v^2/c^2} = \gamma$ 。考虑固有时的物理意义：固有时是相对于坐标架静止的观者所测的时间，那么从这个式子中自然的就可以看出钟慢效应。

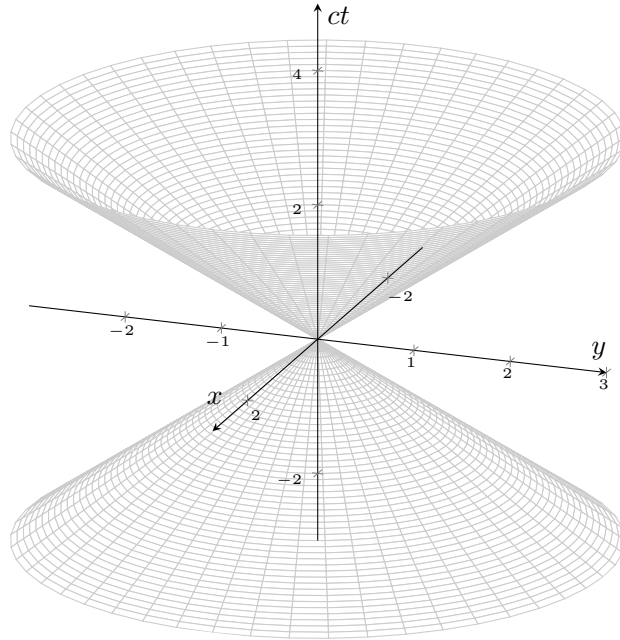
2.3 直观理解

上一节中对固有时的引入让我们可以解释双子悖论。考虑 A 处一对双胞胎，为了方便起见我们称他们为哥哥和弟弟（但是他们年龄、外貌均一致）。哥哥在 A 处乘坐宇宙飞船以速度 v 前往空间中 B 处（在时空中即 B' 处），再以相同的速度返回 A ，如下图所示。



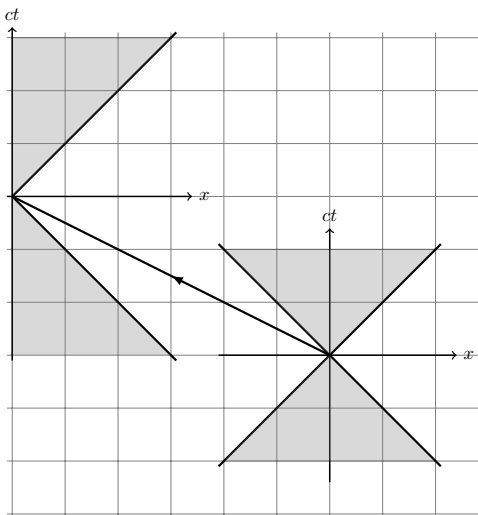
哥哥步下宇宙飞船时，他应当看到一个比自己年迈的弟弟。年龄，或者说年迈程度实际上只与随个人运动的坐标系下的时间有关，因此年迈程度就能反映固有时的长短。弟弟在时空中走的路径是 $A \rightarrow A'$ ，那么他的固有时按照定义就应该是 6 年，即可以说他老了六岁。然而在上面给出的极端例子里，我们算得的哥哥的固有时应当是 $\tau = 2\sqrt{-(3c \text{ yr})^2 + (3c \text{ yr})^2}/c = 0!$ 这就意味着，在哥哥看来，他在“一瞬间”就完成了六光年的路程，然而弟弟老去了六岁。由于固有时是洛伦兹变换下的不变标量，那么双子悖论中所谓的“因参考系的改变而产生的哥哥和弟弟之间年龄差异的矛盾”就不存在了。

从上面的例子可以看出，时空图（即上面 $ct - x$ 图）在研究狭义相对论甚至广义相对论中都是很有用的。而且我们还能发现，在时空图中， $ct = \alpha x (\alpha \leq 1)$ 的部分对应的粒子运动速度会“超光速”；这并不是一个“真实”的路径，这一部分对应的时空间隔是正的。我们可以绘制 $ct - xOy$ 时空图中所有的光速粒子的世界线，得到下图



我们称这样一个两个锥体组成的形状叫光锥。光锥将时空区分成了三块：光锥之外的部分，光锥之上的部分，光锥之外的部分。简单的计算可知，光锥之外的部分时空间隔是正的，即 $\delta_{ij}dx^i dx^j > (cdt)^2$ ，因此我们称在这个区域的世界线是类空的。真实的粒子速度永远无法超过光速，因此其测地线永远不会是类空的。光锥之上的部分 $\delta_{ij}dx^i dx^j = (cdt)^2$ ；这就是光的世界线，因此我们称光锥上的世界线是类光的，对应的粒子以光速运动。光锥内的部分 $\delta_{ij}dx^i dx^j < (cdt)^2$ ，因此我们称在这里的世界线是类时的：绝大多数粒子的世界线都是类时的。

光锥有什么意义？考虑一个任意粒子的世界线。那么我们可以任在其任意一点处取一个光锥； ct_+ 区域的类时和类光部分表征了粒子所有可能的世界线。如果一个事件在光锥之外，那么由于这个事件无法通过一个类光或是类时的粒子将信息传递给这个粒子（如下图），因此无法在这两个事件之间建立因果联系。因此我们也说光锥给闵氏时空赋予了因果结构。



2.4 测地线方程

在上一节中我们对闵氏时空中的作用量作变分得到了闵氏时空中的自由粒子的运动轨迹。我们现在不免要问：在更一般的时空中，自由粒子的运动方程又将如何？

我们设一般时空中的度规为 $g_{\mu\nu}$ ，那么线元 ds^2 仍可以被定义为 $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ 。仿照上一节的做法依旧对其参数化：

$$ds = d\lambda \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \quad (7)$$

仍考虑作用量 $S_g = \int ds$ 。变分得

$$\begin{aligned} \delta S_g &= \int d\lambda \delta \mathcal{L} \\ &= \int d\lambda \left[\frac{\partial \mathcal{L}}{\partial (x^\rho)} \delta x^\rho + \frac{\partial \mathcal{L}}{\partial \left(\frac{dx^\rho}{d\lambda}\right)} \delta \left(\frac{dx^\rho}{d\lambda}\right) \right] \\ &= \int d\lambda \left[\frac{\partial \mathcal{L}}{\partial (x^\rho)} \delta x^\rho - \frac{d}{d\lambda} \frac{\partial \mathcal{L}}{\partial \left(\frac{dx^\rho}{d\lambda}\right)} \delta x^\rho \right] + \frac{\partial \mathcal{L}}{\partial \left(\frac{dx^\rho}{d\lambda}\right)} \delta x^\rho \Big|_a^b \\ &= \int d\lambda \delta x^\rho \left[-\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\rho} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \frac{1}{\sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}} + \frac{1}{2} \frac{d}{d\lambda} \left(g_{\mu\rho} \frac{dx^\mu}{d\lambda} + g_{\rho\nu} \frac{dx^\nu}{d\lambda} \right) \frac{1}{\sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}} \right] \\ &= 0 \end{aligned} \quad (8)$$

由此得

$$\begin{aligned} &\frac{\partial g_{\mu\nu}}{\partial x^\rho} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} - 2 \frac{d}{d\lambda} \left(g_{\mu\rho} \frac{dx^\mu}{d\lambda} \right) \\ &= \frac{\partial g_{\mu\nu}}{\partial x^\rho} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} - 2 \left(\frac{\partial g_{\mu\rho}}{\partial x^\sigma} \frac{dx^\sigma}{d\lambda} \frac{dx^\mu}{d\lambda} + g_{\mu\rho} \frac{d^2 x^\mu}{d\lambda^2} \right) \\ &= \frac{\partial g_{\mu\nu}}{\partial x^\rho} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} - 2 \frac{\partial g_{\mu\rho}}{\partial x^\nu} \frac{dx^\nu}{d\lambda} \frac{dx^\mu}{d\lambda} - 2 g_{\mu\rho} \frac{d^2 x^\mu}{d\lambda^2} \\ &= -2 \left[g_{\mu\rho} \frac{d^2 x^\mu}{d\lambda^2} + \frac{1}{2} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \left(\frac{\partial g_{\mu\rho}}{\partial x^\nu} + \frac{\partial g_{\nu\rho}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right) \right] \\ &= 0 \end{aligned} \quad (9)$$

在等式两侧分别作用一度规张量的逆：

$$\frac{d^2 x^\sigma}{d\lambda^2} + \frac{1}{2} g^{\rho\sigma} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \left(\frac{\partial g_{\mu\rho}}{\partial x^\nu} + \frac{\partial g_{\nu\rho}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right) = 0 \quad (10)$$

定义克里斯托费尔符号（下简称克氏符）为

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\frac{\partial g_{\mu\rho}}{\partial x^\nu} + \frac{\partial g_{\nu\rho}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right) \quad (11)$$

那么 (10) 就可以简记为

$$\frac{d^2 x^\rho}{d\lambda^2} + \Gamma^\rho_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0 \quad (12)$$

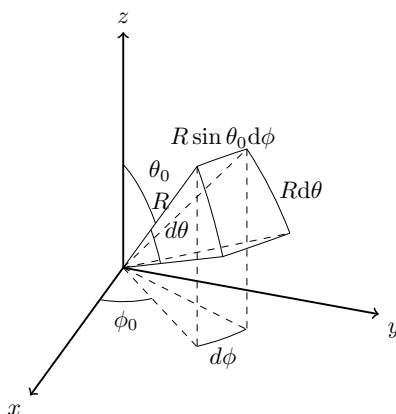
这即是一般时空下的自由粒子运动方程。如果取闵氏时空，那么克氏符为 0，此时这个方程退化为 (5) 中所示的闵氏时空中的自由粒子运动方程。

2.5 球面上的测地线

测地线 (Geodesic) 一词来源于希腊语，意为对地球的测量。同时也发展出一门学科，名叫测地学 (Geodesy)。现在我们虽然是在更一般的对象上研究测地线，然而通过对地球表面 (亦即 S^2) 的研究势必能加深我们对测地线的理解。

通俗来讲，测地线是弯曲空间中两点之间的最短路径。例如在空间中，测地线就是所有的直线；如果对所有的直线参数化那么就自然能得到我们在 (5) 中得到的微分方程 $d^2x^i/dt^2 = 0$ 。这也正是我们在 (12) 中令 $\Gamma = 0$ 会得到的结论。所以我们似乎能发现， $\Gamma^\rho_{\mu\nu}$ 包含了空间弯曲的信息，只需要知道一个空间的所有 $\Gamma^\rho_{\mu\nu}$ 就能得到其中的测地线方程。因此我们尝试证明：地球上的经纬线都是球面上的测地线。

要找到球面上所有的克氏符，那么我们就必须先找到球面上的度规。方便起见我们将两个极点从球面上扣掉。这样取球坐标系 $\{\theta, \phi\}$ 我们就可以通过下图得到球面上的线元 $ds^2 = R^2 (d\theta^2 + \sin^2 \theta d\phi^2)$ 。



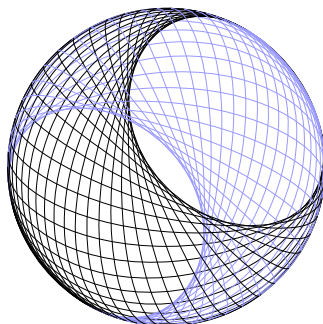
因此球面上的度规就是

$$g_{\mu\nu} = \begin{pmatrix} R^2 & \\ & R^2 \sin^2 \theta \end{pmatrix}$$

这样通过计算所有非零的克氏符只有 $\Gamma^\theta_{\phi\phi} = -\sin \theta \cos \theta$ 和 $\Gamma^\phi_{\theta\phi} = \cot \theta$ ，因此我们能得到两个测地线方程

$$\begin{aligned} \frac{d^2\theta}{dt^2} - \sin \theta \cos \theta \left(\frac{d\phi}{dt} \right)^2 &= 0 \\ \frac{d^2\phi}{dt^2} + \cot \theta \frac{d\theta}{dt} \frac{d\phi}{dt} &= 0 \end{aligned}$$

这个方程组解起来有些复杂，我们在此仅给出其通解为球面上的大圆线，如下图：



如果把地球放在球坐标系下考察，那么我们就写出经纬线的方程是 $\theta = \text{const}, \phi = \text{const}$ 。代回上面的测地线方程可以立刻发现所有的经纬线都是测地线，而这也说明满足测地线方程的曲线是空间中两点间的最短路径的结论。

3 广义相对论

3.1 爱因斯坦-希尔伯特作用量

在场论中，作用量经常会用另一种方式来给出：

$$S = \int d^n x \mathcal{L} \quad (13)$$

这里的拉氏量是一个“密度”，而作用量是拉氏量在整个空间中的一个积分。然而如果考虑一个坐标变换 $x \rightarrow x'$ ，那么就会出现

$$S = \int d^4 x \mathcal{L} = \int d^4 x' \frac{\partial x}{\partial x'} \mathcal{L}' \neq \int d^4 x' \mathcal{L}' \quad (14)$$

这就启发我们要在四维时空中找一个不变体元，满足关系 $a d^4 x = a' d^4 x'$ ，也就是说要有 $a' = (\partial x / \partial x') a$ 。立即意识到度规的变换就满足这样的规则：

$$g'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma} \quad (15)$$

那么我们就可以用度规来构造这么一个不变体元。由于拉氏量是一个标量，那么不变体元也需要是一个标量，我们必须对两侧的度规取行列式，并定义 $\det(g_{\mu\nu}) = g$ ，我们就可以得到

$$\begin{aligned} g' &= \left(\frac{\partial x}{\partial x'} \right)^2 g \\ \sqrt{-g'} &= \frac{\partial x}{\partial x'} \sqrt{-g} \end{aligned} \quad (16)$$

这即是我们的不变体元。这样作用量就可以写作

$$S = \int d^4 x \sqrt{-g} \mathcal{L} \quad (17)$$

作用量应该怎么取？考虑牛顿力学中引力场的泊松方程 $\nabla^2 \phi = 4\pi G\rho$ ：我们有理由相信，拉氏量也应该是某个东西的二次导数构造出的标量。里奇曲率标量正好满足这一要求： $R = g^{\mu\nu} R_{\mu\nu}$ ，其中 $R_{\mu\nu}$ 是里奇曲率张量，是黎曼张量缩并后的结果。这样我们就能取拉氏量为一正比于曲率标量的值： $\mathcal{L} \propto R$ 。因此我们可以将作用量取作

$$S = \frac{1}{2\kappa} \int d^4 x \sqrt{-g} R \quad (18)$$

我们暂且将常数记成 $1/2\kappa$ 的形式；在后面将决定其具体数值。

3.2 真空爱因斯坦场方程推导

前一节中我们得出了 EH 作用量

$$S = \frac{1}{2\kappa} \int d^4 x R \sqrt{-g} \quad (19)$$

变分：

$$\begin{aligned} \delta S &= \frac{1}{2\kappa} \int d^4 x \delta(R \sqrt{-g}) \\ &= \frac{1}{2\kappa} \int d^4 x \delta(g^{\mu\nu} R_{\mu\nu} \sqrt{-g}) \\ &= \frac{1}{2\kappa} \int d^4 x [\delta(g^{\mu\nu} \sqrt{-g}) R_{\mu\nu} + g^{\mu\nu} \sqrt{-g} \delta(R_{\mu\nu})]. \end{aligned} \quad (20)$$

我们先来考察一下变分后的第二项。里奇曲率张量被定义作 $R_{\mu\rho} = R_{\mu\nu\rho}{}^\nu$ ，所以对里奇曲率张量的变分可化作对黎曼张量缩并后的变分。考虑黎曼张量的定义：

$$R_{\mu\nu\rho}{}^\sigma \omega_\sigma = 2\nabla_{[\mu} \nabla_{\nu]} \omega_\rho \quad (21)$$

作用在同一个对偶矢量场上相等，因此展开可得到：

$$\begin{aligned}
 R_{\mu\nu\rho}{}^\sigma &= -2\partial_{[\mu} \Gamma^\sigma{}_{\nu]\rho} + 2\Gamma^\lambda{}_{\rho[\mu} \Gamma^\sigma{}_{\nu]\lambda} \\
 &= -(\partial_\mu \Gamma^\sigma{}_{\nu\rho} - \partial_\nu \Gamma^\sigma{}_{\mu\rho}) + (\Gamma^\lambda{}_{\rho\mu} \Gamma^\sigma{}_{\nu\lambda} - \Gamma^\lambda{}_{\rho\nu} \Gamma^\sigma{}_{\mu\lambda}) \\
 &= \partial_\nu \Gamma^\sigma{}_{\mu\rho} - \partial_\mu \Gamma^\sigma{}_{\nu\rho} + \Gamma^\lambda{}_{\rho\mu} \Gamma^\sigma{}_{\nu\lambda} - \Gamma^\lambda{}_{\rho\nu} \Gamma^\sigma{}_{\mu\lambda}
 \end{aligned} \tag{22}$$

等式两侧取变分：

$$\begin{aligned}
 \delta R_{\mu\nu\rho}{}^\sigma &= \delta(\partial_\nu) \Gamma^\sigma{}_{\mu\rho} + \partial_\nu \delta(\Gamma^\sigma{}_{\mu\rho}) - [\delta(\partial_\mu) \Gamma^\sigma{}_{\nu\rho} + \partial_\mu \delta(\Gamma^\sigma{}_{\nu\rho})] \\
 &\quad + [\delta(\Gamma^\lambda{}_{\rho\mu}) \Gamma^\sigma{}_{\nu\lambda} + \Gamma^\lambda{}_{\rho\mu} \delta(\Gamma^\sigma{}_{\nu\lambda})] - [\delta(\Gamma^\lambda{}_{\rho\nu}) \Gamma^\sigma{}_{\mu\lambda} + \Gamma^\lambda{}_{\rho\nu} \delta(\Gamma^\sigma{}_{\mu\lambda})] \\
 &= \partial_\nu \delta \Gamma^\sigma{}_{\mu\rho} - \partial_\mu \delta \Gamma^\sigma{}_{\nu\rho} + \Gamma^\sigma{}_{\nu\lambda} \delta \Gamma^\lambda{}_{\rho\mu} + \Gamma^\lambda{}_{\rho\mu} \delta \Gamma^\sigma{}_{\nu\lambda} - \Gamma^\sigma{}_{\mu\lambda} \delta \Gamma^\lambda{}_{\rho\nu} - \Gamma^\lambda{}_{\rho\nu} \delta \Gamma^\sigma{}_{\mu\lambda}
 \end{aligned} \tag{23}$$

式中只出现了克氏符的变分项。虽然克氏符本身不是张量，但是克氏符的变分却是一个张量，因此我们可以将导数算符作用在克氏符的变分项上：

$$\nabla_\nu \delta \Gamma^\sigma{}_{\mu\rho} = \partial_\nu \delta \Gamma^\sigma{}_{\mu\rho} + \Gamma^\sigma{}_{\lambda\nu} \delta \Gamma^\lambda{}_{\mu\rho} - \Gamma^\lambda{}_{\mu\nu} \delta \Gamma^\sigma{}_{\lambda\rho} - \Gamma^\lambda{}_{\rho\nu} \delta \Gamma^\sigma{}_{\mu\lambda} \tag{24}$$

同理交换一下角标 $\mu \leftrightarrow \nu$ ，我们能得到：

$$\nabla_\mu \delta \Gamma^\sigma{}_{\nu\rho} = \partial_\mu \delta \Gamma^\sigma{}_{\nu\rho} + \Gamma^\sigma{}_{\lambda\mu} \delta \Gamma^\lambda{}_{\nu\rho} - \Gamma^\lambda{}_{\nu\mu} \delta \Gamma^\sigma{}_{\lambda\rho} - \Gamma^\lambda{}_{\rho\mu} \delta \Gamma^\sigma{}_{\nu\lambda} \tag{25}$$

对比一下 (24)，(25) 和 (23) 中的项，可以发现有以下关系：

$$\delta R_{\mu\nu\rho}{}^\sigma = \nabla_\nu \delta \Gamma^\sigma{}_{\mu\rho} - \nabla_\mu \delta \Gamma^\sigma{}_{\nu\rho} \tag{26}$$

又因为里奇曲率张量实际上是黎曼张量第二、四指标缩并后得到的结果，缩并 ν, σ 再交换 $\rho \leftrightarrow \nu$ 能得到

$$\begin{aligned}
 \delta R_{\mu\rho} &= \nabla_\nu \delta \Gamma^\nu{}_{\mu\rho} - \nabla_\mu \delta \Gamma^\nu{}_{\nu\rho} \\
 \Rightarrow \delta R_{\mu\nu} &= \nabla_\rho \delta \Gamma^\rho{}_{\mu\nu} - \nabla_\mu \delta \Gamma^\rho{}_{\rho\nu}
 \end{aligned} \tag{27}$$

这即我们想得到的对里奇曲率张量变分的结果。

我们现在单独计算 (20) 中第二项的积分。将 (27) 代入 (20) 中的第二项可以得到：

$$\begin{aligned}
 \frac{1}{2\kappa} \int d^4x g^{\mu\nu} \sqrt{-g} \delta R_{\mu\nu} &= \frac{1}{2\kappa} \int d^4x g^{\mu\nu} \sqrt{-g} (\nabla_\rho \delta \Gamma^\rho{}_{\mu\nu} - \nabla_\mu \delta \Gamma^\rho{}_{\rho\nu}) \\
 &= \frac{1}{2\kappa} \int d^4x \sqrt{-g} (\nabla_\rho g^{\mu\nu} \delta \Gamma^\rho{}_{\mu\nu} - \nabla_\mu g^{\mu\nu} \delta \Gamma^\rho{}_{\rho\nu}) \\
 &= \frac{1}{2\kappa} \int d^4x \sqrt{-g} (\nabla_\rho g^{\mu\nu} \delta \Gamma^\rho{}_{\mu\nu} - \nabla_\rho g^{\rho\nu} \delta \Gamma^\mu{}_{\mu\nu})
 \end{aligned} \tag{28}$$

注意到括号中的项的指标缩并完后实际上是一个矢量，因此我们可以写作如下形式：

$$\frac{1}{2\kappa} \int d^4x \sqrt{-g} \nabla_\rho V^\rho \tag{29}$$

这正是高斯定理的形式。高斯定理告诉我们，上式的积分等于边界面的通量和。我们取无穷远为高斯面，那么穿过无穷远的通量和应当是 0，上式的积分就自然等于 0，我们只需要考虑 (20) 中积分的第一项。

现在我们来算第一项。将变分展开可得

$$\frac{1}{2\kappa} \int d^4x R_{\mu\nu} \delta(\sqrt{-g} g^{\mu\nu}) = \frac{1}{2\kappa} \int d^4x R_{\mu\nu} [\delta(\sqrt{-g}) g^{\mu\nu} + \sqrt{-g} \delta g^{\mu\nu}] \tag{30}$$

有关系 $\delta\sqrt{-g} = -1/2\sqrt{-g}\delta g$ 。 δg 可以通过以下推导得出 [2]：

$$\begin{aligned}
 \ln(\det M) &= \text{tr}(\ln(M)) \\
 \frac{1}{\det M} \delta(\det M) &= \text{tr}(M^{-1} \delta M)
 \end{aligned} \tag{31}$$

取 $M = g_{\mu\nu}$ ，那么 $\det M = g$ ，上式就变为

$$\begin{aligned}\delta g &= g(g^{\mu\nu}\delta g_{\mu\nu}) = -g(g_{\mu\nu}\delta g^{\mu\nu}) \\ \delta(\sqrt{-g}) &= \frac{1}{2\sqrt{-g}}g(g_{\mu\nu}\delta g^{\mu\nu}) \\ &= -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}\end{aligned}\quad (32)$$

代入 (20) 自然得

$$\begin{aligned}\frac{1}{2\kappa}\int d^4x R_{\mu\nu}[\delta(\sqrt{-g})g^{\mu\nu} + \sqrt{-g}\delta g^{\mu\nu}] &= \frac{1}{2\kappa}\int d^4x R_{\mu\nu}\left[\left(-\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}\right)g^{\mu\nu} + \sqrt{-g}\delta g^{\mu\nu}\right] \\ &= \frac{1}{2\kappa}\int d^4x \sqrt{-g}\delta g^{\mu\nu}\left(-\frac{1}{2}g_{\mu\nu}R + R_{\mu\nu}\right) = 0\end{aligned}\quad (33)$$

其中 R 是里奇曲率张量和度规缩并得的曲率标量。上式即可得出真空爱因斯坦方程：

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0\quad (34)$$

3.3 考虑物质项的爱因斯坦场方程

考虑物质项后，在作用量中添加一项 $S_m = \int d^4x \sqrt{-g}\mathcal{L}_m$ ，其与真空下的作用量的变分应和为 0。单独考察其变分项：

$$\begin{aligned}\delta S_m &= \int d^4x \delta(\sqrt{-g}\mathcal{L}_m) \\ &= \int d^4x \delta g^{\mu\nu} \frac{\partial(\sqrt{-g}\mathcal{L}_m)}{\partial g^{\mu\nu}}\end{aligned}\quad (35)$$

由于这个积分是要与 (33) 中的积分相加的，我们就想要能够提出一个系数 $\sqrt{-g}$ 来。我们令

$$\begin{aligned}T_{\mu\nu} &= -\frac{2}{\sqrt{-g}} \frac{\partial(\sqrt{-g}\mathcal{L}_m)}{\partial g^{\mu\nu}} \\ \Rightarrow \frac{\partial(\sqrt{-g}\mathcal{L}_m)}{\partial g^{\mu\nu}} &= -\frac{1}{2}\sqrt{-g}T_{\mu\nu}\end{aligned}\quad (36)$$

代回 (35) 式再与 (33) 式相加：

$$\begin{aligned}-\frac{1}{2}\int d^4x \sqrt{-g}\delta g^{\mu\nu}T_{\mu\nu} + \frac{1}{2\kappa}\int d^4x \sqrt{-g}\delta g^{\mu\nu}\left(-\frac{1}{2}g_{\mu\nu}R + R_{\mu\nu}\right) &= 0 \\ \int d^4x \sqrt{-g}\delta g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \kappa T_{\mu\nu}\right) &= 0\end{aligned}\quad (37)$$

就可以得到有物质场的爱因斯坦场方程：

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}\quad (38)$$

3.4 确定系数

在接下来的讨论中我们将取 $G = c = 1$ 的几何单位制以简化运算。为了确定系数 κ ，我们在测地线方程中考虑一个低速稳定弱场近似 $g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}$ ，其中 $h_{\mu\nu}$ 为一度规小量（因此我们称 $g_{\mu\nu}$ 为弱场；其与闵氏度规之差仅为一个小量）。低速意味着三速度 $v \ll 1$ ，因此四速度 $u \approx (1, 0, 0, 0)$ ，稳定意味着 $\partial g_{\mu\nu}/\partial x^0 = 0$ 。因此有关系

$$\frac{dx^i}{d\tau} \ll \frac{dx^0}{d\tau} \approx 1\quad (39)$$

最后一个也正暗示着在低速下固有时与坐标系中的时间一致。将其代入测地线方程得

$$\frac{d^2x^\sigma}{d\tau^2} + \left(\frac{dx^0}{d\tau}\right)^2 \Gamma^\sigma_{00} = 0.\quad (40)$$

考虑克氏符的定义：

$$\Gamma^{\sigma}_{\mu\nu} = \frac{1}{2}g^{\rho\sigma} \left(\frac{\partial g_{\mu\rho}}{\partial x^{\nu}} + \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}} \right) \quad (41)$$

那么就有

$$\Gamma^{\sigma}_{00} = \frac{1}{2}g^{\sigma\rho} \left(\frac{\partial g_{0\rho}}{\partial x^0} + \frac{\partial g_{0\rho}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^{\rho}} \right) \quad (42)$$

考虑到弱场近似 $g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}$ ，那么 $\partial g_{0\rho}/\partial x^0 \approx \partial \eta_{0\rho}/\partial x^0 = 0$ ，就有

$$\begin{aligned} \Gamma^{\sigma}_{00} &= -\frac{1}{2}g^{\sigma\rho} \frac{\partial g_{00}}{\partial x^{\rho}} \\ &= -\frac{1}{2}(\eta^{\sigma\rho} + h^{\sigma\rho}) \partial_{\rho}(\eta_{00} + h_{00}) \\ &\approx -\frac{1}{2}\eta^{\sigma\rho} \partial_{\rho} h_{00} \end{aligned} \quad (43)$$

考虑 $\sigma = 0$ ，测地线方程就可以写成

$$\frac{d^2 x^0}{d\tau^2} - \frac{1}{2} \left(\frac{dx^0}{d\tau} \right)^2 \eta^{0\rho} \partial_{\rho} h_{00} = 0 \quad (44)$$

$\eta^{\mu\nu}$ 仅在 η^{00} 处有非零值；然而 $\partial_0 h_{00} = 0$ ，所以 $dx/d\tau = \text{const}$ 。因此我们直接考虑仅有空间项的测地线方程：

$$\begin{aligned} \frac{d^2 x^i}{dt^2} - \frac{1}{2} \eta^{ik} \partial_k h_{00} &= 0 \\ \frac{d^2 x^i}{dt^2} - \frac{1}{2} \partial^i h_{00} &= 0 \end{aligned} \quad (45)$$

我们对引力场做了低速稳定弱场近似，那么这个式子应该能够反映牛顿力学的原理。事实也正是这样；考虑牛顿力学框架下的引力势 Φ ，被写作

$$\frac{d^2 x^i}{dt^2} + \partial^i \Phi = 0 \quad (46)$$

因此我们应该取 $h_{00} = -2\Phi$ 来使我们的近似能够回到牛顿力学的框架中。

现在我们回到场方程中。我们不加说明的指出， T_{00} 为能量密度，因此泊松方程可以写作

$$\begin{aligned} \nabla^2 \Phi &= 4\pi\rho \\ -\frac{1}{2} \nabla^2 h_{00} &= 4\pi T_{00} \end{aligned} \quad (47)$$

再考虑稳定弱场近似下的克氏符：

$$\begin{aligned} \Gamma^{\sigma}_{\mu\nu} &= \frac{1}{2}g^{\rho\sigma} \left(\frac{\partial g_{\mu\rho}}{\partial x^{\nu}} + \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}} \right) \\ &= \frac{1}{2}\eta^{\rho\sigma} (\partial_{\nu} h_{\mu\rho} + \partial_{\mu} h_{\nu\rho} - \partial_{\rho} h_{\mu\nu}) \end{aligned} \quad (48)$$

我们将其带回黎曼张量的定义式并舍去高阶小量得

$$\begin{aligned} R_{\mu\nu\rho}{}^{\sigma} &= \partial_{\nu} \Gamma^{\sigma}_{\mu\rho} - \partial_{\mu} \Gamma^{\sigma}_{\nu\rho} \\ &= \frac{1}{2} \eta^{\lambda\sigma} [\partial_{\nu} (\partial_{\rho} h_{\mu\lambda} + \partial_{\mu} h_{\rho\lambda} - \partial_{\lambda} h_{\mu\rho}) - \partial_{\mu} (\partial_{\rho} h_{\nu\lambda} + \partial_{\nu} h_{\rho\lambda} - \partial_{\lambda} h_{\nu\rho})] \\ \Rightarrow R_{\mu\rho} &= R_{\mu\nu\rho}{}^{\nu} = \frac{1}{2} \eta^{\lambda\nu} [\partial_{\nu} (\partial_{\rho} h_{\mu\lambda} + \partial_{\mu} h_{\rho\lambda} - \partial_{\lambda} h_{\mu\rho}) - \partial_{\mu} (\partial_{\rho} h_{\nu\lambda} + \partial_{\nu} h_{\rho\lambda} - \partial_{\lambda} h_{\nu\rho})] \\ \Rightarrow R_{00} &= \frac{1}{2} \eta^{\lambda\nu} [\partial_{\nu} (\partial_0 h_{0\lambda} + \partial_0 h_{0\lambda} - \partial_{\lambda} h_{00}) - \partial_0 (\partial_0 h_{\nu\lambda} + \partial_{\nu} h_{0\lambda} - \partial_{\lambda} h_{\nu 0})] \\ &= -\frac{1}{2} \eta^{\lambda\nu} \partial_{\nu} \partial_{\lambda} h_{00} \\ &= -\frac{1}{2} \delta^{ij} \partial_i \partial_j h_{00} = -\frac{1}{2} \nabla^2 h_{00} = 4\pi T_{00}. \end{aligned} \quad (49)$$

现在我们需要做的就是找出 $R_{\mu\nu} - 1/2g_{\mu\nu}R$ 与 $T_{\mu\nu}$ 之间的关系。在场方程两侧同时乘以 $g^{\mu\nu}$ 能得到

$$R_{\mu\nu}g^{\mu\nu} - 1/2g_{\mu\nu}g^{\mu\nu}R = R - 2R = -R = \kappa T. \quad (50)$$

那么场方程就可以改写为

$$R_{\mu\nu} = \kappa \left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu} \right) \quad (51)$$

考虑低速条件，能动张量的迹 $T = g^{\mu\nu}T_{\mu\nu} = \text{tr}(T_{\mu\nu}) \approx -T_{00}$ ，再取 00 分量，我们能够得到

$$R_{00} = \frac{1}{2}\kappa T_{00} \quad (52)$$

与 (49) 中结论比对我们可以发现 $\kappa = 8\pi$ ，这是我们选取 $G = c = 1$ 的结果。现在我们通过量纲分析将 c 还原。

我们先来看左侧所谓爱因斯坦张量的量纲。由于黎曼张量作用于一个对偶矢量上等价于两个联络作用于一个对偶矢量上之差，而联络本身作为导数无量纲，因此黎曼张量无量纲。而对于里奇张量我们有 $R_{\mu\rho} = g^{\nu\sigma}R_{\mu\nu\rho\sigma} = g^{\nu\sigma}g_{\sigma\lambda}R_{\mu\nu\rho}{}^\lambda$ ，于是里奇张量也和黎曼张量一样无量纲。曲率标量 $R = R_{\mu\rho}g^{\mu\rho}$ 就应当具有 $g^{\mu\rho}$ 的量纲，即 L^{-2} ；因此等式左侧无量纲。

我们再来看右侧能动张量的量纲 [1]。考虑时空中的理想流体的能动张量：

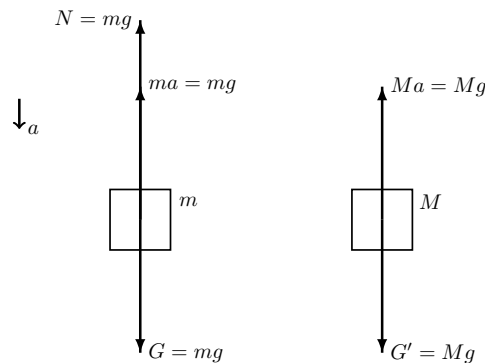
$$T_{\mu\nu} = \rho u_\mu u_\nu + pg_{\mu\nu} + pu_\mu u_\nu \quad (53)$$

由于相加项量纲一定一致，我们只需考虑 $pg_{\mu\nu}$ 的量纲。[$pg_{\mu\nu}$] = MLT^{-2} ，因此能动张量也应该具有量纲 MLT^{-2} 。但是等式左侧的爱因斯坦张量无量纲，所以 [κ] = $M^{-1}L^{-1}T^2$ 。由上述的讨论可以发现， κ 应当是正比于引力常数 G 的，而 [G] = $M^{-1}L^3T^{-2}$ ，因此 [κ/G] = $L^{-4}T^4 = [c^{-4}]$ ，因此我们可以确定 $\kappa = 8\pi Gc^{-4}$ ，这样场方程最终可以确定的写为

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} \quad (54)$$

3.5 爱因斯坦场方程的解

简单来讲，场方程描述了这么一件事：时空指导物质如何运动，物质指导时空如何弯曲。这一观点与牛顿力学有一个根本的区别。牛顿力学中，指导物质运动的永远都是一个力，这一论断由牛顿第二定律给出： $\mathbf{F} = d\mathbf{p}/dt$ 。即便是我们在引入场的概念后，我们研究场对粒子运动的影响也总是透过力这一个媒介；例如如果考虑一个引力场 $V = -GM/r$ ，那么我们需要有 $F/m = -dV/dr$ 然后才能得到粒子的运动方程。然而在广义相对论的框架下，引力不再是一个力，而是被认为是时空的一个内蕴性质；也就是说一切所谓“受到引力作用”的粒子的运动实际上都是时空中的测地运动。比如一个因引力而下落的苹果；在牛顿力学看来这个苹果在引力的作用下做加速运动；然而在广义相对论看来这个苹果做测地运动（走时空中的“最短”路径），其 4-加速度（在时空中的加速度）为 0，反而坐在椅子上的一个人有一个非 0 的 4-加速度。也可以说，广义相对论在局部上将一个引力场与一个局部的加速参考系等同起来。例如一个在地球表面静止的物体和一个自由落体的物体，如果在局部的一个向下 \mathbf{g} 的非惯性系中考虑，那么就可以对其受力分析：

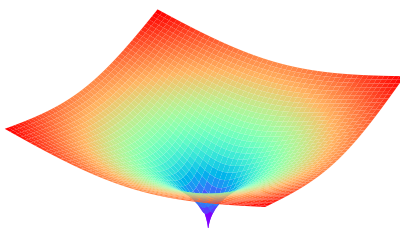


所以说在加速系中考虑的原静止的物块将有一个向上的 $a' = N/m = g$ 的加速度，而一个原自由落体的物体是“静止”的。这样一种等效的操作被称作等效原理。透过等效原理，我们可以完全抛弃引力作为一个力的存在而在空间中处处定义一个局部的加速参考系；这样也就等同于牛顿力学中引力场的概念。而在一般的时空中的测地线方程 (12) 里面我们看到了坐标的二次导数项，与加速度的定义相似。这启发我们用一种时空结构来在处处给出一个局部的加速参考系，这样我们就能得到“时空指导物质如何运动”的结论。

场方程本身的物理意义可能不是很直观，于是我们将在本节讨论他的一个最简单的解：史瓦西真空解。史瓦西真空解是球对称的真空爱因斯坦场方程的一个一般的解，描述了不带电、不自转的星体外部的时空弯曲方式，在球坐标系中写作如下形式：

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

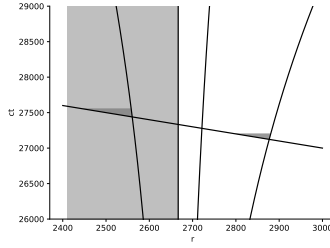
其中我们认为 M 代表了星体的质量 (事实上这个说法欠准确)。对于场方程的解，我们可以取时空中中的一个等时面、压缩掉两个空间维度再将其嵌入 \mathbb{R}^3 来研究其性质：



这是代入了太阳的数值后的代表太阳周围时空弯曲模式的图，曲面是压缩了一个时间维度和一个空间维度的史瓦西时空。在这样的一个嵌入图中，我们可以将太阳想象成一个放在中央的球，时空在球的作用下向“下”弯曲；然而嵌入到 \mathbb{R}^3 中的曲面的曲率是史瓦西时空的内蕴曲率 (这代表着，他的曲率并不需要在更高维的空间中去得到)，因此通过嵌入图我们可以更直观的看出一个时空的内蕴曲率，从而看出其代表的“引力”和产生这一种内蕴曲率的原理。

下面我们介绍一些广义相对论应用的实例。

黑洞 黑洞是一个极其特殊的星体。在人们观测到 M87* 之前，黑洞的存在仅仅是广义相对论通过如下推导得出的一个特殊星体：如果在史瓦西时空中取 $r = 2M$ ，那么我们就发现在此处的线元没有意义，但是可以通过取一个恰当的坐标系使得其有意义。尽管如此，这一半径的特殊取值仍然是非常重要的。通过上面的量纲分析我们可以知道 $r = 2GM/c^2$ ，这一半径被称作史瓦西半径。对于恒星等星体而言，如果其所有质量都被堆积在了这个半径的球体之内，那么它将坍缩成一个黑洞。通过牛顿力学中一个滑稽的等式同样也可以推出星体的史瓦西半径： $1/2mc^2 = GMm/r \implies r = 2GM/c^2$ 。这是一个神奇的巧合，但也反映了事件视界的意义：事件视界内部的粒子需要“超光速”以逃离黑洞，但当然这是无法完成的事情。也就是说，虽然一个人可以自由的从外部跨过一个黑洞的事件视界，他却永远无法跨出来了。通过取恰当的坐标系再计算史瓦西时空中的测地线，我们可以画出类似下图的 $ct - r$ 图：



其中的直线和对数曲线均为类光世界线，两个世界线的交点处的阴影代表了焦点处的未来光锥。左侧阴影部分是史瓦西半径内部，即事件视界内部。如此我们就能看出，如果一个粒子跨入了事件视界，那么其全部的未来光锥都在事件视界内部；任何类时或是类光的世界线都无法逃逸事件视界。这也正是“粒子需要超光速以逃离黑洞”的原理。这图中还可以看出，由于越靠近事件视界的外沿的世界线越陡，其到达远处某 r_0 处的时间就越长，最终趋近无穷大。所以如果事件视界外一个观者看着一个人掉入事件视界，那个人掉进去的那一瞬间的影像将永远定格在事件视界表面。（由于这个过程往往还伴随着严重的红移，我们假设这个观者能接收全部的频率的信息）

水星进动 牛顿力学下的行星轨道应该是一个封闭的椭圆形，恒星（太阳）在其中一个焦点处。如果考虑外围行星对其的引力作用，那么行星的轨道在近日点处会发生进动现象，也就是说近日点会在一个以太阳为圆心的圆上运动。水星离太阳最近、与太阳的质量差异最显著、观测也较方便，因此研究水星的进动问题常常是最简单的。1859年法国天文学家于尔班·勒威耶所测量的水星的进动速率比牛顿所给出的进动速率要快一些；当时的物理学家无法对这一误差做出合理的解释。爱因斯坦发展出广义相对论后，通过广义相对论的方法计算出了正确的水星进动速率，详细过程可参考 Biesel, O. (2008)[4]。

4 标量张量理论

前一节中，我们假设了爱因斯坦-希尔伯特作用量的拉氏量正比于里奇标量。但是这其实是一种取巧的做法；实际上，如果我们假设拉氏量是一个标量场与里奇标量的耦合量，那么它仍然可以是一个爱因斯坦-希尔伯特作用量的拉氏量，并且给出与常见的场方程不一样的场方程。其中一个比较著名的理论是 Brans-Dicke 理论（下简称 BD 理论）。BD 理论中的场方程由如下作用量给出：

$$S = \frac{1}{16\pi} \int d^4x \sqrt{-g} \left[\phi R - \frac{\omega}{\phi} \nabla_\mu \phi \nabla^\mu \phi - V(\phi) \right] + \int d^4x \sqrt{-g} \mathcal{L}_m \quad (55)$$

我们现在试图通过变分法来得出 BD 理论下的场方程。

发现这是一个有两个宗量的泛函，因此我们先对 ϕ 变分：

$$\begin{aligned} \delta S &= \frac{1}{16\pi} \int d^4x \sqrt{-g} \left(R - \frac{\omega}{\phi^2} \nabla_\mu \phi \nabla^\mu \phi + \frac{2\omega}{\phi} \square \phi - \frac{dV}{d\phi} \right) \delta \phi = 0 \\ \Rightarrow R - \frac{\omega}{\phi^2} \nabla_\mu \phi \nabla^\mu \phi + \frac{2\omega}{\phi} \square \phi - \frac{dV}{d\phi} &= 0 \end{aligned} \quad (56)$$

其中 $\square = g^{\mu\nu} \nabla_\mu \nabla_\nu$ 。

再对 $g^{\mu\nu}$ 变分。我们先重新写一下作用量：

$$S = \int d^4x \left\{ \frac{1}{16\pi} \left[\sqrt{-g} \phi R - \frac{\omega}{\phi} \sqrt{-g} \nabla_\mu \phi \nabla^\mu \phi - \sqrt{-g} V(\phi) \right] + \sqrt{-g} \mathcal{L}_m \right\} \quad (57)$$

4.1 对耦合项的变分

第一部分 $1/16\pi \sqrt{-g} \phi R$ 的变分比较复杂：

$$\delta(\sqrt{-g} \phi R) = \delta(\phi) \sqrt{-g} R + \phi \delta(\sqrt{-g} R) \quad (58)$$

先考虑第二部分的变分。注意到这实际上和广义相对论框架下的变分推导一致，因此第二部分的变分可写作

$$\phi \delta(\sqrt{-g}R) = \phi \sqrt{-g} \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^{\mu\nu} \equiv \phi \sqrt{-g} G_{\mu\nu} \delta g^{\mu\nu} \quad (59)$$

其中我们定义 $R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ 为一个新的张量 $G_{\mu\nu}$ ，称作爱因斯坦张量。

下面我们来看第一项。使用分部积分法，我们可以得到

$$\delta(\phi) \sqrt{-g} R = \delta(\phi) \sqrt{-g} R_{\mu\nu} g^{\mu\nu} = \sqrt{-g} g^{\mu\nu} [\delta(R_{\mu\nu} \phi) - \phi \delta R_{\mu\nu}] = -\sqrt{-g} g^{\mu\nu} \phi \delta R_{\mu\nu} \quad (60)$$

再次利用一下上一节中的结论，在测地坐标下 $\delta R_{\mu\nu} = \nabla_\rho \delta \Gamma^\rho_{\mu\nu} - \nabla_\mu \delta \Gamma^\rho_{\rho\nu}$ ，我们会有

$$\begin{aligned} g^{\mu\nu} \delta R_{\mu\nu} &= g^{\mu\nu} (\nabla_\rho \delta \Gamma^\rho_{\mu\nu} - \nabla_\mu \delta \Gamma^\rho_{\rho\nu}) \\ &= \nabla_\rho g^{\mu\nu} \delta \Gamma^\rho_{\mu\nu} - \nabla^\nu \delta \Gamma^\rho_{\rho\nu} \\ &= \nabla_\sigma (g^{\mu\nu} \delta \Gamma^\sigma_{\mu\nu} - g^{\nu\sigma} \delta \Gamma^\rho_{\rho\nu}) \end{aligned} \quad (61)$$

我们再来考察克氏符的变分项。我们可以取测地坐标使得 $\partial_\rho g_{\mu\nu} = 0$ 。将克氏符的变分项展开能得到：

$$\delta \Gamma^\sigma_{\mu\nu} = \frac{1}{2} \left\{ \delta g^{\sigma\lambda} [\partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}] + g^{\sigma\lambda} [\delta(\partial_\mu g_{\nu\lambda}) + \delta(\partial_\nu g_{\mu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})] \right\} \quad (62)$$

测地坐标下第一项自然全部为 0。注意到对导数的变分可以写为变分的导数，因此克氏符的变分可以写作

$$\begin{aligned} \delta \Gamma^\sigma_{\mu\nu} &= \frac{1}{2} g^{\sigma\lambda} (\partial_\mu \delta g_{\nu\lambda} + \partial_\nu \delta g_{\mu\lambda} - \partial_\lambda \delta g_{\mu\nu}) = \frac{1}{2} (g^{\sigma\lambda} \partial_\mu \delta g_{\nu\lambda} + g^{\sigma\lambda} \partial_\nu \delta g_{\mu\lambda} - \partial^\sigma \delta g_{\mu\nu}) \\ \delta \Gamma^\rho_{\rho\nu} &= \frac{1}{2} g^{\rho\gamma} (\partial_\rho \delta g_{\nu\gamma} + \partial_\nu \delta g_{\rho\gamma} - \partial_\gamma \delta g_{\rho\nu}) = \frac{1}{2} (\partial^\gamma \delta g_{\nu\gamma} + g^{\rho\gamma} \partial_\nu \delta g_{\rho\gamma} - \partial^\rho \delta g_{\rho\nu}) \end{aligned} \quad (63)$$

代回 (61)：

$$\begin{aligned} g^{\mu\nu} \delta R_{\mu\nu} &= \nabla_\sigma \left[g^{\mu\nu} \left(\frac{1}{2} g^{\sigma\lambda} \partial_\mu \delta g_{\nu\lambda} + \frac{1}{2} g^{\sigma\lambda} \partial_\nu \delta g_{\mu\lambda} - \frac{1}{2} \partial^\sigma \delta g_{\mu\nu} \right) - \right. \\ &\quad \left. g^{\nu\sigma} \left(\frac{1}{2} \partial^\gamma \delta g_{\nu\gamma} + \frac{1}{2} g^{\rho\gamma} \partial_\nu \delta g_{\rho\gamma} - \frac{1}{2} \partial^\rho \delta g_{\rho\nu} \right) \right] \\ &= \nabla_\sigma \left(\frac{1}{2} g^{\sigma\lambda} \partial^\nu \delta g_{\nu\lambda} + \frac{1}{2} g^{\sigma\lambda} \partial^\mu \delta g_{\mu\lambda} - \frac{1}{2} g^{\mu\nu} \partial^\sigma \delta g_{\mu\nu} \right. \\ &\quad \left. - \frac{1}{2} g^{\nu\sigma} \partial^\gamma \delta g_{\nu\gamma} - \frac{1}{2} g^{\rho\gamma} \partial^\sigma \delta g_{\rho\gamma} + \frac{1}{2} g^{\nu\sigma} \partial^\rho \delta g_{\rho\nu} \right) \end{aligned} \quad (64)$$

观察指标求和的形式，并利用测地坐标下 $\nabla_\mu = \partial_\mu$ 可以得到

$$\begin{aligned} g^{\mu\nu} \phi \delta R_{\mu\nu} &= \phi \nabla_\sigma (g^{\sigma\nu} \partial^\mu - g^{\mu\nu} \partial^\sigma) \delta g_{\mu\nu} \\ &= \phi (-\nabla_\nu \nabla_\mu + g_{\mu\nu} g^{\rho\sigma} \nabla_\sigma \nabla_\rho) \delta g^{\mu\nu} \\ &= (\nabla_\mu \nabla_\nu \phi - g_{\mu\nu} \square \phi) \delta g^{\mu\nu} \end{aligned} \quad (65)$$

因此对耦合项的变分是

$$\delta(\sqrt{-g} \phi R) = [\phi \sqrt{-g} G_{\mu\nu} - \sqrt{-g} (\nabla_\mu \nabla_\nu \phi - g_{\mu\nu} \square \phi)] \delta g^{\mu\nu} \quad (66)$$

4.2 对剩余项的变分

现在我们来考察第二项，即 $-\omega/\phi \sqrt{-g} \nabla_\mu \phi \nabla^\mu \phi$ 项的变分。这一项的变分比较简单：

$$\begin{aligned} \delta \left(-\frac{\omega}{\phi} \sqrt{-g} \nabla_\mu \phi \nabla^\mu \phi \right) &= \delta \left(-\frac{\omega}{\phi} \sqrt{-g} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi \right) \\ &= -\frac{\omega}{\phi} (\delta \sqrt{-g} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi + \sqrt{-g} \delta g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi) \\ &= -\frac{\omega}{\phi} \left(-\frac{1}{2} \sqrt{-g} g_{\mu\nu} \nabla_\rho \phi \nabla^\rho \phi + \sqrt{-g} \nabla_\mu \phi \nabla_\nu \phi \right) \delta g^{\mu\nu} \end{aligned} \quad (67)$$

第三项变分结果如下：

$$-\delta\sqrt{-g}V(\phi) = \frac{1}{2}\sqrt{-g}g_{\mu\nu}V(\phi)\delta g^{\mu\nu} \quad (68)$$

对第四项的变分我们在前一章也有遇到：

$$\delta(\sqrt{-g}\mathcal{L}_m) = -\frac{1}{2}\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu} \quad (69)$$

4.3 对 Brans-Dicke 理论的分析

因此整个作用量对 $g^{\mu\nu}$ 的变分是

$$\left\{ \left[\phi\sqrt{-g}G_{\mu\nu} - \sqrt{-g}(\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\square\phi) \right] - \frac{\omega}{\phi} \left(-\frac{1}{2}\sqrt{-g}g_{\mu\nu}\nabla_\rho\phi\nabla^\rho\phi + \sqrt{-g}\nabla_\mu\phi\nabla_\nu\phi \right) + \frac{1}{2}\sqrt{-g}g_{\mu\nu}V(\phi) - 8\pi\sqrt{-g}T_{\mu\nu} \right\} \delta g^{\mu\nu} = 0 \quad (70)$$

第二个场方程可以被写作

$$G_{\mu\nu} = \frac{8\pi}{\phi}T_{\mu\nu} + \frac{\omega}{\phi^2} \left(\nabla_\mu\phi\nabla_\nu\phi - \frac{1}{2}g_{\mu\nu}\nabla^\rho\phi\nabla_\rho\phi \right) + \frac{1}{\phi}(\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\square\phi) - \frac{V}{2\phi}g_{\mu\nu}. \quad (71)$$

接下来我们来分析一下 BD 理论中标量场的意义。取场方程的迹 [3] 可以得到一个 R 的表达式：

$$R = -\frac{8\pi}{\phi}T + \frac{\omega}{\phi^2}\nabla^\mu\phi\nabla_\mu\phi + \frac{3}{\phi}\square\phi + \frac{2V}{\phi} \quad (72)$$

将其代回 (56) 可以得到：

$$\square\phi = \frac{1}{2\omega + 3} \left(8\pi T + \phi \frac{dV}{d\phi} - 2V \right). \quad (73)$$

这是一个标量场的微分方程。这个式子告诉我们，标量场是由物质项 T 给出的，然而在作用量中并没有与 \mathcal{L}_m 耦合，而是通过度规对物质运动产生影响。同时对比一下 (71) 与 (54) 我们可以发现，在 BD 理论中 $G = 1/\phi$ ，从一个不变的常量成为了一个随标量场改变而改变的量，因而变成了一个时空坐标的函数。

5 总结

本文中我们利用对物理模型的作用量变分的方法导出了狭义相对论中自由粒子的运动轨迹、一般时空中的测地线方程以及广义相对论和 Brans-Dicke 理论中的场方程，强调了作用量在物理理论中的重要地位，用拉氏量将狭义相对论、广义相对论以及其他理论联系起来，并给出了修正引力的理论导出场方程的一种方法。

参考文献

- [1] 梁灿彬, & 周彬. (2009). 微分几何入门与广义相对论. 科学出版社.
- [2] Carroll, S. M. (2019). Spacetime and geometry. Cambridge University Press.
- [3] Faraoni, V., & Faraoni, V. (2004). Scalar-Tensor Gravity (pp. 1-53). Springer Netherlands.
- [4] Biesel, O. (2008). The precession of mercury's perihelion. Leiden University.