幼儿园版广义相对论

Kindergarten General Relativity

黄志琦著

First printing, December 2021

Contents

序

谁还不是个宝宝呢?

1. 标架和张量

1.1 狭义相对论

即使忽略引力，我们所处的“平直”四维时空并不是一个欧几里德 (Euclid) 空间，因为“时间”这个维度和空间维度有明显的不同。狭义相对论的给出的“平直时空”坐标系 (t,x,y,z)（其中 t 为时间，x,y,z 为三维直角坐标）对应的度规（邻近点之间的距离平方和坐标差的关系）为

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \tag{1.1}$$

通常采用 $c=1$ 的自然单位制（可以理解为我们把 ct 简单记作了 t）。这样

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \tag{1.2}$$

这和我们在小班学过的欧几里德空间中的勾股定理有所不同：$\mathrm{d}t^2$ 前面的符号是负的。

如果惯性参考系 K'（坐标 (t',x',y',z')）相对于惯性参考系 K（坐标 (t,x,y,z)）以速度 $(\beta,0,0)$ 运动，且两个坐标系的坐标轴均对齐，在 $t=t'=0$ 时刻两坐标系原点重合。那么有如下的洛伦兹变换公式：

$$t' = \frac{t-\beta x}{\sqrt{1-\beta^2}}, \tag{1.3}$$

$$x' = \frac{x-\beta t}{\sqrt{1-\beta^2}}, \tag{1.4}$$

$$y' = y, \tag{1.5}$$

$$z' = z. \tag{1.6}$$

其反变换为

$$
\begin{aligned}
t &= \frac{t'+\beta x'}{\sqrt{1-\beta^2}}, && (1.7)\\
x &= \frac{x'+\beta t'}{\sqrt{1-\beta^2}}, && (1.8)\\
y &= y', && (1.9)\\
z &= z'. && (1.10)
\end{aligned}
$$

例题 1: 假设某种机器人的使用寿命为 $\tau = 100$ 年。让机器人乘坐宇宙飞船去探索 $L = 1000$ 光年外的行星，飞船的速率 v 至少需要多大?

解答：在地球参考系，机器人到达行星事件坐标为 $t = L/v$，$x = L$。在机器人参考系

$$t' = \frac{t - vx}{\sqrt{1-v^2}} = \frac{L}{v}\sqrt{1-v^2}$$

令 $t' = 100$ 年，即可算出 v 的下限为 $\frac{10}{\sqrt{101}}$。

例题 2: 孪生兄弟张三和李四同时出生，在他们 20 岁的时候，张三乘坐宇宙飞船以大小为 $v_0 = \frac{3}{5}$ 的初始速度从地球出发，李四则留在地球上生活。张三的宇宙飞船在地球参考系中做匀加速直线运动（加速度方向指向地球）。在李四 80 岁时，张三回到了地球，这时张三多少岁?

解答：容易看出在地球参考系中，张三的速度为 $v(t) = \frac{3}{5} - at$，这里的 $a = \frac{1}{60\,\mathrm{yr}}\frac{6}{5} = \frac{1}{50\,\mathrm{yr}}$。张三在每段短暂的时间间隔内都可以看成一个近似的惯性系 K', $\mathrm{d}x' = 0$, 于是根据四维间隔不变性 $-(\mathrm{d}t')^2 = -\mathrm{d}t^2 + \mathrm{d}x^2$, 以及 $\mathrm{d}x = v(t)\mathrm{d}t$, 有 $\mathrm{d}t' = \sqrt{1-\left(\frac{3}{5}-at\right)^2}\mathrm{d}t$。积分得到

$$t' = \int_0^{6/(5a)} \sqrt{1-\left(\frac{3}{5}-at\right)^2}\,\mathrm{d}t = \frac{1}{a}\int_{-3/5}^{3/5}\sqrt{1-u^2}\,\mathrm{d}u = \frac{1}{a}\left(\frac{12}{25}+\arcsin\frac{3}{5}\right) = \left(24+50\arcsin\frac{3}{5}\right)\,\mathrm{yr}.$$

于是张三是 $44 + 50\arcsin\frac{3}{5} \approx 76.175$ 岁。

要注意区分狭义相对论默认的“上帝视角”叙事方式（用坐标描述事件）和现实观测的差别：实际的观测者并不能随意地获得事件的坐标，而需要通过实测手段。请仔细体会下面这道例题中的“动钟变慢”效应和其他教科书上的“上帝视角”描述有何不同。

例题 3: 一个速度为 $v=\frac{5}{13}$ 的航天旅行者和他的地球上的朋友在出发时对好了钟的时刻为 $t'=0$ (旅行者的时钟) 和 $t=0$ (地球上的钟)。地球上的朋友同时观察两个钟，直接观察 t，用望远镜观察 t'。当看到 t' 读数为 $1\,\mathrm{h}$ 时，看到 t 的读数为多少 h?

解答：旅行者时钟读数为 $1\,\mathrm{h}$ 的事件 A 在旅行者参考系的坐标为：$t'_A=1\,\mathrm{h}$，$x'_A=0$。在地球坐标系

$$t_A=\frac{t'_A+vx'_A}{\sqrt{1-v^2}}=\frac{13}{12}\mathrm{h}.$$

$$x_A=\frac{x'_A+vt'_A}{\sqrt{1-v^2}}=\frac{5}{12}\mathrm{h}.$$

这里的 $t_A>t'_A$ 就是通常所说的狭义相对论“动钟变慢”效应，然而对于题中所述的实际观测，我们需要考虑这个事件以光信号的形式传播到地球，被地球上的朋友观测到，还需要 $x_A=\frac{5}{12}\mathrm{h}$，因此 $t=t_A+x_A=\frac{3}{2}\mathrm{h}$。

1.2 标架、度规和逆标架

在比较形式的讨论中，我们经常用度规矩阵 $\eta_{\mu\nu}=\mathrm{diag}(-1,1,1,1)$（即 $\eta_{00}=-1,\eta_{11}=\eta_{22}=\eta_{33}=1$,其余元素均为零)来描述狭义相对论的闵可夫斯基(Minkowski)时空。时空的任意点用四维坐标 x^0,x^1,x^2,x^3 来描述（注意这里的上标只是个指标，并不代表次幂)，指标 0 代表时间维度，指标 1,2,3 代表空间维度。邻近点的时空距离可以写作：

$$\mathrm{d}s^2=\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu. \tag{1.11}$$

在上面的写法中，我们使用了爱因斯坦求和规则：默认对一上一下重复的希腊字母指标对空间所有维度指标（在狭义相对论里，就是从 0 到 3）求和。$\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ 实际上代表 $\sum_{\mu=0}^{3}\sum_{\nu=0}^{3}\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$。**爱因斯坦求和规则的写法将贯穿本书**。

狭义相对论的惯性系可以用标架（一组基矢）$\vec{e}_0,\vec{e}_1,\vec{e}_2,\vec{e}_3$ 来描述，它们之间的内积满足：$\vec{e}_\mu\cdot\vec{e}_\nu=\eta_{\mu\nu}$（注意这意味着时间基矢满足 $\vec{e}_0\cdot\vec{e}_0=-1$)。由于惯性系的“平直性”，在任何点的四维矢量 $\vec{A}$ 都可以用这个固定的标架线性组合出来，写成 $A^\mu\vec{e}_\mu$ （这里的 A^μ 的上标 μ 也只是个指标，不代表次幂；注意这里对 μ 指标进行了默认求和）的形式。但是在我们即将讨论的更一般理论（广义相对论）中，一般不存在这样的“普适”标架，每个时空点的标架可以不同。因此每个矢量都是定义在一个时空点上的，它可以分解为这个时空点的标架的线性组合。

在非欧氏空间，我们要区分矢量“按标架分解的系数”和“在标架上的投影”——两者一般来说是不同的，前者指把矢量 $\vec{A}$ 写成 $A^\mu\vec{e}_\mu$ 的形式，后者指把矢量 $\vec{A}$

和各个基矢取内积 $A_\mu \equiv \vec{A}\cdot\vec{e}_\mu$。我们把“在标架上的投影”$A_\mu$ 称为“协变分量”，而把“按标架分解的系数”A^μ 称为“逆变分量”。这个“逆”字涉及到我们实际上如何计算 A^μ：我们将构造一组“逆标架”$\vec{e}^0,\vec{e}^1,\vec{e}^2,\vec{e}^3$，使得**矢量在逆标架上的投影就等于它按标架分解的系数**，即 $A^\mu = \vec{A}\cdot\vec{e}^\mu$。对狭义相对论时空而言，很容易发现只要取 $\vec{e}^0=-\vec{e}_0, \vec{e}^1=\vec{e}_1, \vec{e}^2=\vec{e}_2, \vec{e}^3=\vec{e}_3$ 就能满足要求。不过，我们要稍微深入讨论下对任意标架该怎样获得逆标架——这将是讨论广义相对论的关键步骤。

假设（未必是平直的）m 维空间的每个点都有一组标架 $\vec{n}_1,\vec{n}_2,\ldots,\vec{n}_m$，满足 $\vec{n}_\mu\cdot\vec{n}_\nu = g_{\mu\nu}$(这里我们要求内积满足交换律, 即度规矩阵 $g_{\mu\nu}$ 为对称矩阵)。如果存在一组完备的“逆标架”$\vec{N}^1,\vec{N}^2,\ldots,\vec{N}^m$ 使得 $\vec{n}_\mu\cdot\vec{N}^\nu=\delta_\mu^\nu$, 这里克罗内克 (Kronecker) 符号 δ_μ^ν 即当且仅当 $\mu=\nu$ 时取值为 1，否则取值为零，我们就可以证明 $A_\mu\vec{N}^\mu$ 和 $A^\mu\vec{n}_\mu$ 都和 $\vec{A}$ 相等（这里 $A_\mu\equiv\vec{A}\cdot\vec{n}_\mu$，$A^\mu\equiv\vec{A}\cdot\vec{N}^\mu$）。事实上，对任意指标 ν，有

$$(A_\mu\vec{N}^\mu)\cdot\vec{n}_\nu = A_\mu\delta_\nu^\mu = A_\nu = \vec{A}\cdot\vec{n}_\nu,$$

根据 $\vec{n}_1,\vec{n}_2,\ldots,\vec{n}_m$ 的完备性即可知道 $A_\mu\vec{N}^\mu=\vec{A}$。

$$(A^\mu\vec{n}_\mu)\cdot\vec{N}^\nu = A^\mu\delta_\mu^\nu = A^\nu = \vec{A}\cdot\vec{N}^\nu.$$

根据 $\vec{N}^1,\vec{N}^2,\ldots,\vec{N}^m$ 的完备性即可知道 $A^\mu\vec{n}_\mu=\vec{A}$。

那么问题就归结为寻找满足 $\vec{n}_\mu\cdot\vec{N}^\nu=\delta_\mu^\nu$ 的 $\vec{N}^1,\vec{N}^2,\ldots,\vec{N}^m$。我们只需要把 $g_{\mu\nu}$ 看成一个 $m\times m$ 的矩阵并求逆，把逆矩阵记作 $M^{\mu\nu}$（这里我们把“逆”矩阵的指标放到了上边，是为了更方便地使用爱因斯坦规则）。那么容易验证 $\vec{N}^\nu\equiv M^{\nu\alpha}\vec{n}_\alpha$ 满足要求。实际上对任意指标 μ，有

$$(M^{\nu\alpha}\vec{n}_\alpha)\cdot\vec{n}_\mu = M^{\nu\alpha}g_{\alpha\mu} = \delta_\mu^\nu.$$

我们后面会做一件现在看上去还比较奇怪的事情：把 $g_{\mu\nu}$ 的逆矩阵 M 也换成用 g 来表示，通过它的指标在上边，我们可以判断 $g^{\mu\nu}$ 代表的是逆矩阵的元素。我们把 $g^{\mu\nu}$ 叫作度规的“逆变形式”——它和度规的“协变形式”$g_{\mu\nu}$ 都是物理量“度规”的不同数学表述形式。

总结一下：我们通过计算逆标架的流程如下：

(1) 计算标架的内积获得度规矩阵 $g_{\mu\nu}=\vec{n}_\mu\cdot\vec{n}_\nu$。

(2) 把度规矩阵求逆获得度规的逆变形式 $g^{\mu\nu}$。

(3) 用度规的逆变形式对标架进行线性组合获得“逆标架”：$\vec{N}^\mu=g^{\mu\nu}\vec{n}_\nu$。

标架和逆标架的地位是完全对称的, 容易验证标架也可以用逆标架表示出来：$\vec{n}_\mu = g_{\mu\nu}\vec{N}^\nu$。矢量按照标架分解的系数是矢量在逆标架上的投影；矢量按照逆标架分解的系数是矢量在标架上的投影。

1.3 张量

k 阶张量 A 是这样的“机器”：它依次“吃”进 k 个矢量 $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k$，“吐”出一个实数 $A(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k)$。这个吐出的实数线性地依赖于每一个“吃”进去的矢量。也就是说对任意指标 $1\le i\le k$ 以及实数 a,b，都有

$$A(\ldots,a\vec{v}_i+b\vec{u}_i,\ldots)=aA(\ldots,\vec{v}_i,\ldots)+bA(\ldots,\vec{u}_i,\ldots). \tag{1.12}$$

从一般的时空理论考虑，每个张量必须“居住”在时空点上——它的“食谱”也仅限于这个时空点的矢量。

如果在时空每个点都放置一个 k 阶张量，我们就称之为 k 阶张量场。

标量不需要“吃”进矢量，直接就“吐”出一个实数，所以标量就是零阶张量。

矢量 $\vec{A}$ “吃”进任意矢量 $\vec{B}$ 后，可以吐出一个实数 $\vec{A}\cdot\vec{B}$，所以矢量就是一阶张量。

度规张量 g 是这样定义的：它吃进任意矢量 $\vec{A},\vec{B}$，吐出实数 $g(\vec{A},\vec{B})=\vec{A}\cdot\vec{B}$。显然度规张量是二阶张量。

在一般的 m 维空间中，如果 k 阶张量 A 所居住的点建立了标架 $\vec{n}_1,\vec{n}_2,\ldots,\vec{n}_m$（并用上一小节所述方法求出了逆标架），就可以用“协变分量”，“逆变分量”或者“混合分量”来表示 A。协变分量 $A_{\mu_1\mu_2\ldots\mu_k}$ 是指把 $\vec{n}_{\mu_1},\vec{n}_{\mu_2},\ldots,\vec{n}_{\mu_k}$ 喂给 A 之后获得的实数，也就是说 $A_{\mu_1\mu_2\ldots\mu_k}$ 是 $A\left(\vec{n}_{\mu_1},\vec{n}_{\mu_2},\ldots,\vec{n}_{\mu_k}\right)$ 的简写。逆变分量 $A^{\mu_1\mu_2\ldots\mu_k}$ 则是 $A\left(\vec{N}^{\mu_1},\vec{N}^{\mu_2},\ldots,\vec{N}^{\mu_k}\right)$ 的简写。混合分量则是把标架基矢和逆标架基矢混合着喂给张量获得的实数。例如：二阶张量的混合分量 $A_\mu{}^\nu=A\left(\vec{n}_\mu,\vec{N}^\nu\right)$；四阶张量的混合分量 $A_\alpha{}^{\beta\mu}{}_\nu=A\left(\vec{n}_\alpha,\vec{N}^\beta,\vec{N}^\mu,\vec{n}_\nu\right)$。

利用标架和逆标架之间的转换关系：$\vec{N}^\mu=g^{\mu\nu}\vec{n}_\nu$ 和 $\vec{n}_\mu=g_{\mu\nu}\vec{N}^\nu$，容易把张量的分量的指标进行“升降”。例如，

$$A^\mu{}_\nu=A(\vec{N}^\mu,\vec{n}_\nu)=A(g^{\mu\alpha}\vec{n}_\alpha,\vec{n}_\nu)=g^{\mu\alpha}A_{\alpha\nu}.$$

$$A_{\mu\nu\rho}=A(\vec{n}_\mu,\vec{n}_\nu,\vec{n}_\rho)=A(g_{\mu\alpha}\vec{N}^\alpha,g_{\nu\beta}\vec{N}^\beta,g_{\rho\gamma}\vec{N}^\gamma)=g_{\mu\alpha}g_{\nu\beta}g_{\rho\gamma}A^{\alpha\beta\gamma}.$$

熟练之后，就无须写中间过程。

最后，我们来研究一下度规张量的协变、逆变和混合分量。

按照定义，$g_{\mu\nu}=g(\vec{n}_\mu,\vec{n}_\nu)=\vec{n}_\mu\cdot\vec{n}_\nu$，这和一开始引入度规时的表达式一致。

把协变形式的度规矩阵 $g_{\mu\nu}$ 的逆矩阵记作 M，则 $g^{\mu\nu}=\vec{N}^\mu\cdot\vec{N}^\nu=(M^{\mu\alpha}\vec{n}_\alpha)\cdot(M^{\nu\beta}\vec{n}_\beta)=M^{\mu\alpha}M^{\nu\beta}g_{\beta\alpha}=M^{\mu\alpha}\delta_\alpha^\nu=M^{\mu\nu}$。这验证了我们之前直接把 M 写成 g 的合法性。

最后，度规张量的混合形式 $g_\mu{}^\nu=g(\vec{n}_\mu,\vec{N}^\nu)=\vec{n}_\mu\cdot\vec{N}^\nu=\delta_\mu^\nu$，以及 $g^\mu{}_\nu=g(\vec{N}^\mu,\vec{n}_\nu)=\vec{N}^\mu\cdot\vec{n}_\nu=\delta_\nu^\mu$。就是说，度规张量的混合形式等价于克罗内克符号。

1.4 构造新张量的方法

1.4.1 张量的指标重排序

设 A 为 k 阶张量，对 $1,2,\ldots,k$ 的任意一个排列 $P=(i_1,i_2,\ldots,i_k)$，都可以定义一个张量

$$AP(\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k)\equiv A\left(\vec{v}_{i_1},\vec{v}_{i_2},\ldots,\vec{v}_{i_k}\right).$$

这相当于把 A 的输入端重新排了一下次序。

定义偶排列（$1,2,\ldots,k$ 经过偶数次置换得到的排列）P 的取值为 $\sigma(P)=1$，奇排列 P 的取值为 $\sigma(P)=-1$，可以构造完全反对称张量（指交换任意两个输入矢量，都会使得输出的实数反号）：

$$[A]=\sum_P \sigma(P)AP. \tag{1.13}$$

例如，二阶张量 A 经过反对称化之后为：

$$[A](\vec{v}_1,\vec{v}_2)=A(\vec{v}_1,\vec{v}_2)-A(\vec{v}_2,\vec{v}_1) \tag{1.14}$$

三阶张量 A 经过反对称化之后为

$$[A](\vec{v}_1,\vec{v}_2,\vec{v}_3)=A(\vec{v}_1,\vec{v}_2,\vec{v}_3)+A(\vec{v}_2,\vec{v}_3,\vec{v}_1)+A(\vec{v}_3,\vec{v}_1,\vec{v}_2)-A(\vec{v}_1,\vec{v}_3,\vec{v}_2)-A(\vec{v}_2,\vec{v}_1,\vec{v}_3)-A(\vec{v}_3,\vec{v}_2,\vec{v}_1). \tag{1.15}$$

k 阶完全反对称张量也叫做 k-形式。

1.4.2 张量积

我们可以把一个 p 阶张量 A 和一个 q 阶张量 B “组合”成一个 $p+q$ 阶张量：当你依次输入 $p+q$ 个矢量时，我把前 p 个依次喂给 A，后 q 个依次喂给 B，并把它们吐出来的实数相乘，就得到一个输出实数。这样的“组装机”我们称之为张量 A 和张量 B 的“张量积”[1]，记作 $A\otimes B$。

如果建立了标架 $\vec{n}_1,\vec{n}_2,\ldots,\vec{n}_m$。我们可以用标架或者逆标架构造张量积，$k$ 阶张量 A 可以写成

$$A=A^{\mu_1\mu_2\ldots\mu_k}\vec{n}_{\mu_1}\otimes\vec{n}_{\mu_2}\otimes\ldots\otimes\vec{n}_{\mu_k}. \tag{1.16}$$

或

$$A=A_{\mu_1\mu_2\ldots\mu_k}\vec{N}^{\mu_1}\otimes\vec{N}^{\mu_2}\otimes\ldots\otimes\vec{N}^{\mu_k} \tag{1.17}$$

[1]有的场合也叫“直积”，但直积这个名字容易和线性空间的直积概念混淆。

所以高阶张量虽然不一定直接能写成一堆矢量的张量积，但它一定能写成这样的张量积的线性组合！所以我们可以把矢量当成构建张量世界的基本砖块。

等式 (1.16) 的证明如下：对任意一组空间维度指标 $\alpha_1,\alpha_2,\ldots,\alpha_k$，有

$$A^{\mu_1\mu_2\ldots\mu_k}\left(\vec{n}_{\mu_1}\otimes\vec{n}_{\mu_2}\otimes\ldots\otimes\vec{n}_{\mu_k}\right)\left(\vec{N}^{\alpha_1},\vec{N}^{\alpha_2},\ldots,\vec{N}^{\alpha_k}\right)=A^{\mu_1\mu_2\ldots\mu_k}\delta^{\alpha_1}_{\mu_1}\delta^{\alpha_2}_{\mu_2}\ldots\delta^{\alpha_k}_{\mu_k}=A^{\alpha_1\alpha_2\ldots\alpha_k}.$$

也就是说 (1.16) 两边作用于 $\vec{N}^{\alpha_1},\vec{N}^{\alpha_2},\ldots,\vec{N}^{\alpha_k}$ 的结果相同。由逆标架的完备性即完成证明。

1.4.3 张量指标的缩并

我们来考虑怎样降低张量的阶数。设有 k 阶张量 A，它有 k 个“输入端”。在建立了标架之后，我们可以给一个输入端投喂标架 $\vec{n}_\mu$，另一个输入端投喂逆标架 $\vec{N}^\mu$，并对 μ 按爱因斯坦求和规则进行求和。问题是：这样得到的是一个（和标架的选择无关的）张量吗？

假设有另一组标架 $\vec{n}'_1,\vec{n}'_2,\ldots,\vec{n}'_m$ 以及逆标架 $\vec{N}'^1,\vec{N}'^2,\ldots,\vec{N}'^m$。设标架之间的变换关系为 $\vec{n}'_\mu=T_\mu{}^\alpha\vec{n}_\alpha$，$\vec{N}'^\nu=S_\beta{}^\nu\vec{N}^\beta$，这里 T,S 均为 $m\times m$ 的常数矩阵。那么 $\delta^\nu_\mu=\vec{n}'_\mu\cdot\vec{N}'^\nu=T_\mu{}^\alpha S_\beta{}^\nu\delta^\beta_\alpha=T_\mu{}^\alpha S_\alpha{}^\nu$。于是 T 和 S 互为逆矩阵，也满足 $S_\mu{}^\alpha T_\alpha{}^\nu=\delta^\nu_\mu$。那么

$$\begin{aligned}
A\left(\ldots,\vec{N}'^\mu,\ldots,\vec{n}'_\mu,\ldots\right) &= A\left(\ldots,S_\beta{}^\mu\vec{N}^\beta,\ldots,T_\mu{}^\alpha\vec{n}_\alpha,\ldots\right)\\
&= S_\beta{}^\mu T_\mu{}^\alpha A\left(\ldots,\vec{N}^\beta,\ldots,\vec{n}_\alpha,\ldots\right)\\
&= \delta^\alpha_\beta A\left(\ldots,\vec{N}^\beta,\ldots,\vec{n}_\alpha,\ldots\right)\\
&= A\left(\ldots,\vec{N}^\alpha,\ldots,\vec{n}_\alpha,\ldots\right).
\end{aligned}\tag{1.18}$$

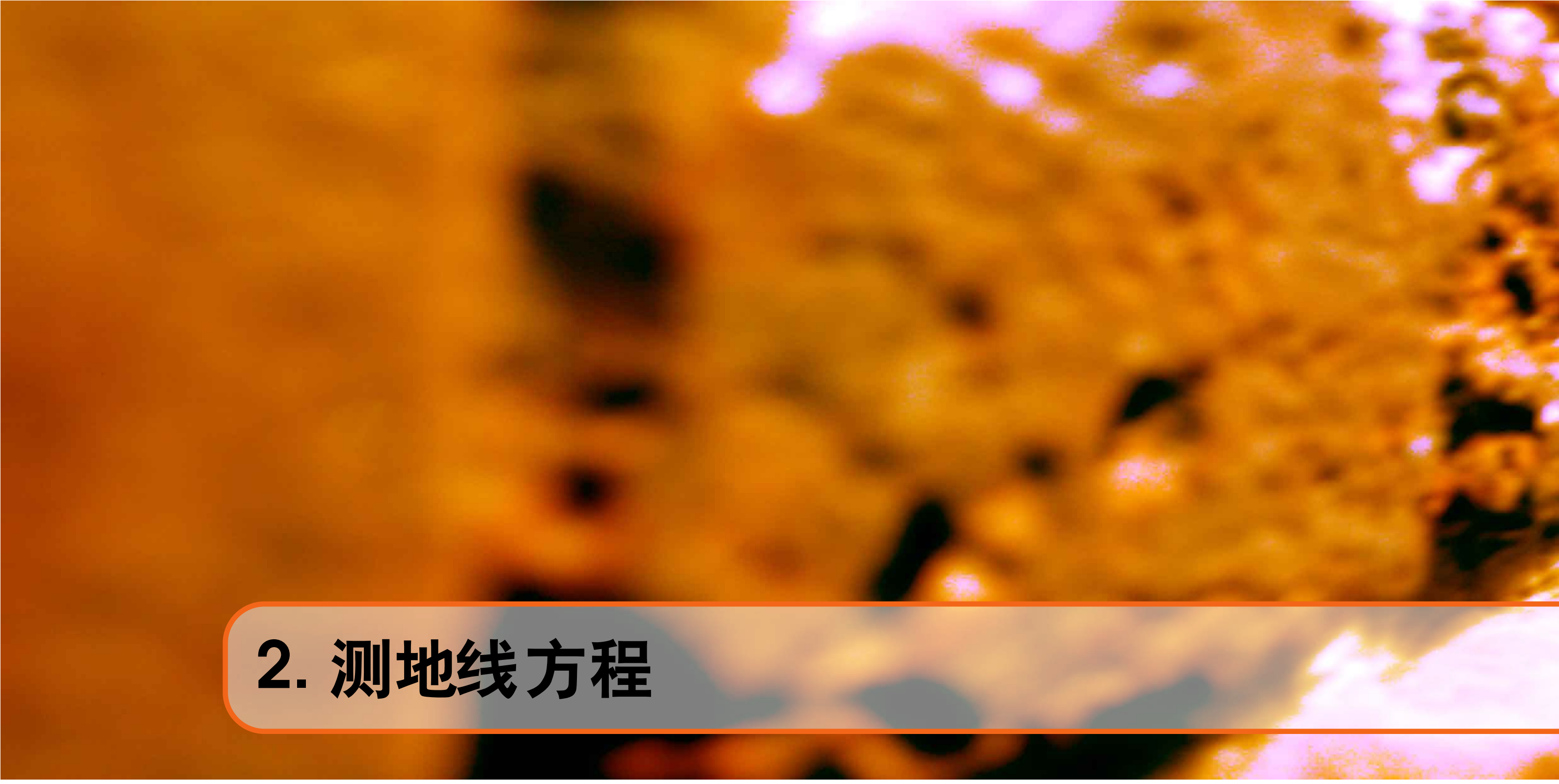
另外，要注意缩并只和选定的两个输入端有关，和哪个指标输入端喂逆标架，哪个输入端喂标架没有关系。这是因为缩并可以写成下面的对称形式：

$$A\left(\ldots,\vec{N}^\mu,\ldots,\vec{n}_\mu,\ldots\right)=g^{\nu\mu}A\left(\ldots,\vec{n}_\nu,\ldots,\vec{n}_\mu,\ldots\right).\tag{1.19}$$

在建立了标架，张量用分量来表示的情况下，我们可以直接把张量的缩并用重复指标写出来。例如二阶张量 A 的缩并 $A^\mu{}_\mu$ 为一个标量。三阶张量 B 的前两个指标缩并后的 $B^\mu{}_{\mu\nu}$ 为矢量。

1.5 课后练习

习题 1: 证明两个坐标分别为 (t_1,x_1,y_1,z_1) 和 (t_2,x_2,y_2,z_2) 的事件之间的距离平方 $\Delta s^2\equiv-(t_1-t_2)^2+(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2$ 在洛伦兹变换下是不变量。



2. 测地线方程

2.1 坐标系和自然标架

在这一章里，我们将用坐标系来描述弯曲的空间。我们假想一个嵌入在 n 维平直空间（由固定的标架 $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ 以及固定度规 $h_{ij} = \vec{e}_i \cdot \vec{e}_j$ 描述）中的 m 维弯曲空间。弯曲空间可以嵌入在高维的平直空间（例如球面可以嵌入三维欧几里德空间）其实是个完全没有必要的假设（高斯首先发现了这一点），做这个假设只是为了照顾我们（或者其实只有我）可怜的想象力。

在 n 维平直空间（后面我们称之为超空间）中，每个点可以用坐标 $(x^1, x^2, \ldots, x^n)$ 描述，对应的矢量 $\vec{x} = x^i \vec{e}_i$（拉丁字母自动从 1 到 n 求和）。在我们的想象中，每个点对应的 $\vec{x}$ 是一个物理存在，和固定标架 $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ 的选择无关。

在超空间中，两个邻近点之间的“距离平方”由矢量差的自内积给出：

$$\mathrm{d}s^2 = h_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{2.1}$$

和之前的讨论一样，我们把度规 h_{ij} 看成一个 $n \times n$ 矩阵并计算它的逆矩阵，把逆矩阵元素记为 h^{ij}。超空间中的逆标架 $\vec{e}^i = h^{ij}\vec{e}_j$。超空间中的每个点 $\vec{x}$ 也可以用它的协变分量 $x_i \equiv \vec{x} \cdot \vec{e}_i = x^j \vec{e}_j \cdot \vec{e}_i = h_{ij}x^j$ 来表示。

当然，超空间只是为了描述问题方便而引入的假想空间。现在我们回到“真实”的 m 维弯曲空间，为了更加形象我后面将称之为“曲面”。假设在曲面上建立了光滑的坐标系 $(u^1, u^2, \ldots, u^m)$。我们可以结合超空间的概念和坐标系来建立标架：

$$\vec{n}_\mu = \frac{\partial \vec{x}}{\partial u^\mu}. \tag{2.2}$$

就是说 $\vec{n}_\mu$ 是沿着 u^μ 坐标线的切矢量。标架 (2.2) 称为坐标系 $(u^1, u^2, \ldots, u^m)$ 的自然标架。

从超空间的角度看，自然标架张成的矢量空间代表了曲面的“切平面”——该点的张量都只“吃”切平面上的矢量。我们可以用上一章的方法构造出自然标架的逆标架 $\vec{N}^1, \vec{N}^2, \ldots, \vec{N}^n$。逆标架只是标架的重新线性组合，所以张成的空间仍然是切平面。

曲面上的度规 $g_{\mu\nu} = \vec{n}_\mu \cdot \vec{n}_\nu$ 给出了曲面上邻近点的距离平方：

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}u^\mu \mathrm{d}u^\nu. \tag{2.3}$$

2.2 张量分量的坐标系变换

如果曲面上有另外一个坐标系 $(\tilde{u}^1, \tilde{u}^2, \ldots, \tilde{u}^n)$，它的自然标架为

$$\vec{\tilde{n}}_\mu = \frac{\partial \vec{x}}{\partial \tilde{u}^\mu} = \frac{\partial \vec{x}}{\partial u^\nu}\frac{\partial u^\nu}{\partial \tilde{u}^\mu} = \frac{\partial u^\nu}{\partial \tilde{u}^\mu}\vec{n}_\nu. \tag{2.4}$$

利用 Jacobi 矩阵 $\frac{\partial u^\mu}{\partial \tilde{u}^\nu}$ 和 $\frac{\partial \tilde{u}^\mu}{\partial u^\nu}$ 互为逆矩阵，可以证明逆标架

$$\vec{\tilde{N}}^\mu = \frac{\partial \tilde{u}^\mu}{\partial u^\nu}\vec{N}^\nu. \tag{2.5}$$

事实上，按上式给出的 $\vec{\tilde{N}}^\mu$，

$$\vec{\tilde{N}}^\mu \cdot \vec{\tilde{n}}_\nu = \frac{\partial \tilde{u}^\mu}{\partial u^\alpha}\frac{\partial u^\beta}{\partial \tilde{u}^\nu}\vec{N}^\alpha \cdot \vec{n}_\beta = \frac{\partial \tilde{u}^\mu}{\partial u^\alpha}\frac{\partial u^\beta}{\partial \tilde{u}^\nu}\delta^\alpha_\beta = \frac{\partial \tilde{u}^\mu}{\partial u^\alpha}\frac{\partial u^\alpha}{\partial \tilde{u}^\nu} = \delta^\mu_\nu. \tag{2.6}$$

满足逆标架的定义。

利用(2.4) 和 (2.5) 可以把张量的任意形式分量在不同坐标系之间转换。例如

$$\tilde{A}^\mu{}_{\alpha\beta} = \frac{\partial \tilde{u}^\mu}{\partial u^{\mu'}}\frac{\partial u^{\alpha'}}{\partial \tilde{u}^\alpha}\frac{\partial u^{\beta'}}{\partial \tilde{u}^\beta}A^{\mu'}{}_{\alpha'\beta'} \tag{2.7}$$

请仔细体会上述混合分量的变换规则——它可以推广到任意阶张量的任意混合分量形式。

上述坐标变换规则的更重要应用是它的逆命题也成立：分量在坐标变换下满足上述变换规则的量是张量。事实上，很多教材把坐标变换规则作为张量的定义。

2.3 张量场的梯度

我们对一个固定点处的张量的研究到此告一段落。从现在开始，我们要考虑曲面上的张量场的梯度。

2.3.1 联络

要定义张量场的梯度，必须把邻近点的张量作一个比较。在超空间里，由于标架是固定的，不同点张量的比较非常容易，只要喂给它们相同的矢量，然后看它们输出的实数的差异就行了。但是，曲面上的张量“吃”的矢量只能在切平面里，而一般来说不同点的切平面（从超空间的视角看）不一样，根本就没法喂给它们相同的矢量！

因为“梯度”只比较邻近的点的张量，我们先来研究邻近点的切平面的变化，我们把自然标架的变化率分解为其自身的线性组合：

$$\frac{\partial \vec{n}_\mu}{\partial u^\nu} = \Gamma^\lambda_{\mu\nu}\vec{n}_\lambda. \tag{2.8}$$

上式中的“联络”(connection) $\Gamma^\lambda_{\mu\nu}$ 作为“按标架分解的系数”，等价于被分解量在逆标架上的投影：

$$\Gamma^\lambda_{\mu\nu} \equiv \vec{N}^\lambda \cdot \frac{\partial \vec{n}_\mu}{\partial u^\nu}. \tag{2.9}$$

因为 $\vec{N}^\lambda \cdot \vec{n}_\mu = \delta^\lambda_\mu$ 不依赖于坐标，所以上式也可以写成

$$\Gamma^\lambda_{\mu\nu} = -\vec{n}_\mu \cdot \frac{\partial \vec{N}^\lambda}{\partial u^\nu}. \tag{2.10}$$

上式又等价于

$$\frac{\partial \vec{N}^\lambda}{\partial u^\nu} = -\Gamma^\lambda_{\mu\nu}\vec{N}^\mu. \tag{2.11}$$

也就是说：“逆标架的联络”和“标架的联络”相差一个符号。

有趣的是，我们可以直接对度规求导计算出联络。利用

$$\Gamma^{\lambda}_{\mu\nu} = \vec{N}^{\lambda} \cdot \frac{\partial \vec{n}_{\mu}}{\partial u^{\nu}} = g^{\lambda\alpha} \vec{n}_{\alpha} \cdot \frac{\partial \vec{n}_{\mu}}{\partial u^{\nu}} = g^{\lambda\alpha} \frac{\partial \vec{x}}{\partial u^{\alpha}} \cdot \frac{\partial^2 \vec{x}}{\partial u^{\nu} \partial u^{\mu}}. \tag{2.12}$$

以及

$$\frac{\partial g_{\mu\alpha}}{\partial u^{\nu}} = \frac{\partial^2 \vec{x}}{\partial u^{\mu} \partial u^{\nu}} \cdot \frac{\partial \vec{x}}{\partial u^{\alpha}} + \frac{\partial \vec{x}}{\partial u^{\mu}} \cdot \frac{\partial^2 \vec{x}}{\partial u^{\alpha} \partial u^{\nu}}, \tag{2.13}$$

$$\frac{\partial g_{\nu\alpha}}{\partial u^{\mu}} = \frac{\partial^2 \vec{x}}{\partial u^{\nu} \partial u^{\mu}} \cdot \frac{\partial \vec{x}}{\partial u^{\alpha}} + \frac{\partial \vec{x}}{\partial u^{\nu}} \cdot \frac{\partial^2 \vec{x}}{\partial u^{\alpha} \partial u^{\mu}}, \tag{2.14}$$

$$\frac{\partial g_{\mu\nu}}{\partial u^{\alpha}} = \frac{\partial^2 \vec{x}}{\partial u^{\mu} \partial u^{\alpha}} \cdot \frac{\partial \vec{x}}{\partial u^{\nu}} + \frac{\partial \vec{x}}{\partial u^{\mu}} \cdot \frac{\partial^2 \vec{x}}{\partial u^{\alpha} \partial u^{\nu}}, \tag{2.15}$$

就可以得到

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\alpha} \left(\frac{\partial g_{\mu\alpha}}{\partial u^{\nu}} + \frac{\partial g_{\nu\alpha}}{\partial u^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial u^{\alpha}} \right). \tag{2.16}$$

上面这个由度规给出的联络叫作克利斯朵夫符号 (Cristoffel Symbols)。嗯，别问还有别的什么联络。幼儿园的宝宝们不宜知道得太多！

要注意的是：联络并不是张量。在坐标变化 $u \to \tilde{u}$ 下，

$$\tilde{\vec{N}}^{\lambda} = \frac{\partial \tilde{u}^{\lambda}}{\partial u^{\lambda'}} \vec{N}^{\lambda'}. \tag{2.17}$$

$$\frac{\partial \tilde{\vec{n}}_{\mu}}{\partial \tilde{u}^{\nu}} = \frac{\partial}{\partial \tilde{u}^{\nu}} \left(\frac{\partial u^{\mu'}}{\partial \tilde{u}^{\mu}} \vec{n}_{\mu'} \right) = \frac{\partial u^{\mu'}}{\partial \tilde{u}^{\mu}} \frac{\partial \vec{n}_{\mu'}}{\partial \tilde{u}^{\nu}} + \frac{\partial^2 u^{\mu'}}{\partial \tilde{u}^{\mu} \partial \tilde{u}^{\nu}} \vec{n}_{\mu'} = \frac{\partial u^{\mu'}}{\partial \tilde{u}^{\mu}} \frac{\partial u^{\nu'}}{\partial \tilde{u}^{\nu}} \frac{\partial \vec{n}_{\mu'}}{\partial u^{\nu'}} + \frac{\partial^2 u^{\mu'}}{\partial \tilde{u}^{\mu} \partial \tilde{u}^{\nu}} \vec{n}_{\mu'}. \tag{2.18}$$

取上面两式的点乘，得到

联络的坐标变换公式：

$$\tilde{\Gamma}^{\lambda}_{\mu\nu} = \frac{\partial \tilde{u}^{\lambda}}{\partial u^{\lambda'}} \frac{\partial u^{\mu'}}{\partial \tilde{u}^{\mu}} \frac{\partial u^{\nu'}}{\partial \tilde{u}^{\nu}} \Gamma^{\lambda'}_{\mu'\nu'} + \frac{\partial \tilde{u}^{\lambda}}{\partial u^{\alpha}} \frac{\partial^2 u^{\alpha}}{\partial \tilde{u}^{\mu} \partial \tilde{u}^{\nu}}. \tag{2.19}$$

右边第一项按照张量的变换规则给出，第二项的存在说明联络不是张量。

2.3.2 局域惯性系

坐标系选取的任意性不代表我们可以在所有点随意地画坐标线（取自然标架），因为那样就会没法拼成一个连续光滑的坐标网格。不过，我们总是可以从一个固定点开始画坐标线。也就是说，我们可以在一个选定的点任意地取自然标架以及自然标架的变化率，然后其他点的标架和标架变化率需要满足一定的光滑性限制。这类似于微分方程的初条件可以随意选取。

那么，对任意选定的一个点 P，能否让自然标架的局域变化率为零呢？从超空间的视角看，答案一般是否定的：从超空间的角度看，不同点的切平面不同，当从一个点开始移动时，自然标架也必须随之变化。不过，我们总是可以在该点让自然标架"最经济"地变化，即变化率是垂直于切平面的（一边移动一边不停地把标架"压"回到切平面上）。也就是说：所有的 $\frac{\partial\vec{n}_\mu}{\partial u^\nu}$ 都和 $\vec{N}^\lambda$ 正交，即在该点的联络的所有的分量都是零。我们称这样建立的坐标系为 P 点的**局域惯性系**。从曲面上的蚂蚁的视角看，在 P 点的局域惯性系里，标架的变化率在 P 点就消失了（因为标架的"真正"变化是垂直于曲面的，蚂蚁看不到）。

如果某个点上的某个张量在一个坐标系里的所有分量都是零，那么它在任何坐标系里都是零；反之，一个非零张量，无论怎样作坐标变换，也不可能把它所有分量变为零。不过，联络并不是张量：它的坐标变换规则里除了张量项之外，还多了一项坐标的二次偏导和一次偏导的乘积。正是这个多出来的项可以和张量项抵消，使得在一个选定点的联络完全消失。要注意的是，不同点的局域惯性系一般而言不同，我们一般只能选取坐标系使得一个点的联络消失，而不是所有点的联络同时消失——除非空间本身就是平直的。

2.3.3 矢量的"平行"移动

在超空间中，一个矢量 $\vec{A}$ 可以平行地从一个点移动到另一个点，从而建立不同点的矢量的等同关系。在曲面上，矢量 $\vec{A}$ 移动时会脱离切平面，我们在移动它的过程中必须不断地把它"压回"到切平面上。最经济的办法当然是沿垂直于切平面的方向压，也就是 $\vec{A}$ 的变化率和标架、逆标架都正交：

$$\frac{\partial\vec{A}}{\partial u^\nu}\cdot\vec{n}_\mu = 0. \tag{2.20}$$

把 $\vec{A} = A_\lambda\vec{N}^\lambda$ 代入上式，得到

$$\frac{\partial A_\lambda}{\partial u^\nu}\left(\vec{N}^\lambda\cdot\vec{n}_\mu\right) + A_\lambda\vec{n}_\mu\cdot\frac{\partial\vec{N}^\lambda}{\partial u^\nu} = 0. \tag{2.21}$$

利用 (2.10)，就有

$$\frac{\partial A_\mu}{\partial u^\nu} - A_\lambda\Gamma^\lambda_{\mu\nu} = 0. \tag{2.22}$$

我们可以从另一个角度来理解 (2.22)：因为矢量的分量 $A_\mu = \vec{A}\cdot\vec{n}_\mu$ 是矢量在标架上的投影，在矢量“不变”的情况下，标架的变化是导致矢量的分量变化的唯一因素。于是可以写出

$$\frac{\partial A_\mu}{\partial u^\nu} = \vec{A}\cdot\frac{\partial \vec{n}_\mu}{\partial u^\nu} = A_\lambda \vec{N}^\lambda\cdot\frac{\partial \vec{n}_\mu}{\partial u^\nu}. \tag{2.23}$$

再利用 (2.9)，就得到和 (2.22) 等价的

$$\frac{\partial A_\mu}{\partial u^\nu} = \Gamma^\lambda_{\mu\nu}A_\lambda. \tag{2.24}$$

把 (2.24) 沿着矢量的移动路径积分，由于“沿途”矢量 $\vec{A}$ 的改变量都不可见（和切平面正交），曲面上的蚂蚁可以认为矢量“没有变化”。这样，我们就建立了一种把矢量从一个点“平行”地移到另一个点的方案。在移动距离的一阶近似下，(2.24) 实际上给出了平行移动的“唯一”结果。但要注意的是，对有限大小的移动距离，矢量平行移动的结果一般和移动路径有关。我们在下一章会对此进行深入的讨论。

在坐标系中，设 $\vec{N}^\mu(P)$ 沿着 $\vec{n}_\nu$ 平行移动 $\mathrm{d}u^\nu$ 后得到（一阶近似意义下）$\vec{B}$，那么在移动终点 Q 处 $\vec{B}$ 的分量为：

$$B_\alpha = \delta^\mu_\alpha + \Gamma^\beta_{\alpha\nu}\delta^\mu_\beta \mathrm{d}u^\nu = \delta^\mu_\alpha + \Gamma^\mu_{\alpha\nu}\mathrm{d}u^\nu. \tag{2.25}$$

也就是在移动终点 Q，

$$\vec{B} = B_\alpha \vec{N}^\alpha(Q) = \vec{N}^\mu(Q) + \Gamma^\mu_{\alpha\nu}\mathrm{d}u^\nu \vec{N}^\alpha. \tag{2.26}$$

（由于右边第二项已经是一阶小量，我们不强调位置）。

设有 k 阶张量 A，在移动前的 P 点，A 吃进 $\vec{N}^{\mu_1}(P),\vec{N}^{\mu_2}(P),\ldots,\vec{N}^{\mu_k}(P)$，吐出 $A^{\mu_1\mu_2\cdots\mu_k}(P)$。然后沿着 $\vec{n}_\nu$ 移动 $\mathrm{d}u^\nu$ 到达 Q 点，在移动终点，让张量吃进“平行”移动后的 $\vec{N}^{\mu_1}(Q)+\Gamma^{\mu_1}_{\alpha\nu}\mathrm{d}u^\nu\vec{N}^\alpha,\vec{N}^{\mu_2}(Q)+\Gamma^{\mu_2}_{\alpha\nu}\mathrm{d}u^\nu\vec{N}^\alpha,\ldots,\vec{N}^{\mu_k}(Q)+\Gamma^{\mu_k}_{\alpha\nu}\mathrm{d}u^\nu\vec{N}^\alpha$，吐出（一阶近似意义下）$A^{\mu_1\mu_2\cdots\mu_k}(Q)+\mathrm{d}u^\nu\left(\Gamma^{\mu_1}_{\alpha\nu}A^{\alpha\mu_2\mu_3\cdots\mu_k}+\Gamma^{\mu_2}_{\alpha\nu}A^{\mu_1\alpha\mu_3\cdots\mu_k}+\ldots+\Gamma^{\mu_k}_{\alpha\nu}A^{\mu_1\mu_2\cdots\mu_{k-1}\alpha}\right) = A^{\mu_1\mu_2\cdots\mu_k}(P)+\mathrm{d}u^\nu\left(\frac{\partial A^{\mu_1\mu_2\cdots\mu_k}}{\partial u^\nu}+\Gamma^{\mu_1}_{\alpha\nu}A^{\alpha\mu_2\mu_3\cdots\mu_k}+\Gamma^{\mu_2}_{\alpha\nu}A^{\mu_1\alpha\mu_3\cdots\mu_k}+\ldots+\Gamma^{\mu_k}_{\alpha\nu}A^{\mu_1\mu_2\cdots\mu_{k-1}\alpha}\right)$。于是张量的梯度的分量可以定义为：

$$(\nabla A)^{\mu_1\mu_2\cdots\mu_k}{}_\nu \equiv \frac{\partial A^{\mu_1\mu_2\cdots\mu_k}}{\partial u^\nu}+\Gamma^{\mu_1}_{\alpha\nu}A^{\alpha\mu_2\mu_3\cdots\mu_k}+\Gamma^{\mu_2}_{\alpha\nu}A^{\mu_1\alpha\mu_3\cdots\mu_k}+\ldots+\Gamma^{\mu_k}_{\alpha\nu}A^{\mu_1\mu_2\cdots\mu_{k-1}\alpha}. \tag{2.27}$$

2.3.4 张量的协变导数

在其他文献中，张量的梯度也叫做张量的“协变导数”或者“协变微商”。我们会交替使用这些名称，但以文献中最常见的名称“协变导数”为主。

在上一小节中，我们建立了邻近点地矢量的“等价关系”，并借此推导出了矢量场的逆变形式的梯度的表达式 (2.27)。下面我们采用更加简明的一种方式来推导

张量的梯度的任意形式分量表达式：在定义联络时，我们已经默认忽略了标架在不可见维度上的变化，这相当于把标架的基矢平行移动到邻近点和那里的标架基矢做比较。所以我们可以通过把标架的变化率写成联络的办法来取代平行移动。

考虑曲面上的一个 k 阶张量场 A，它的协变导数 ∇A 会使 A 的分量随着坐标变化；同时，标架本身的变化也会导致 A 的分量随着坐标变化。我们只要把标架本身的变化的贡献给“抠掉”，就能得到 ∇A。

对标量场 φ 而言，它不需要在标架上进行投影，因此啥也不需要抠——协变导数就是普通偏导数。

$$(\nabla\varphi)_\mu = \frac{\partial\varphi}{\partial u^\mu}. \tag{2.28}$$

对矢量场 $\vec{A}$，我们要从其协变分量的变化率里抠掉标架变化率的贡献。在下面的推导中，我们把 $A(\ldots)$ 看成一个吃进一个矢量，吐出一个实数的操作。

$$\begin{aligned}
(\nabla A)_{\mu\nu} &= \frac{\partial A_\mu}{\partial u^\nu} - A\left(\frac{\partial \vec{n}_\mu}{\partial u^\nu}\right) \\
&= \frac{\partial A_\mu}{\partial u^\nu} - A\left(\Gamma^\lambda_{\mu\nu}\vec{n}_\lambda\right) \\
&= \frac{\partial A_\mu}{\partial u^\nu} - \Gamma^\lambda_{\mu\nu}A\left(\vec{n}_\lambda\right) \\
&= \frac{\partial A_\mu}{\partial u^\nu} - \Gamma^\lambda_{\mu\nu}A_\lambda.
\end{aligned} \tag{2.29}$$

对逆变分量，我们要抠掉逆标架变化率的贡献，注意逆标架的变化率由负的联络给出（见 (2.11)）

$$\begin{aligned}
(\nabla A)^\mu{}_\nu &= \frac{\partial A^\mu}{\partial u^\nu} - A\left(\frac{\partial \vec{N}^\mu}{\partial u^\nu}\right) \\
&= \frac{\partial A^\mu}{\partial u^\nu} - A\left(-\Gamma^\mu_{\lambda\nu}\vec{N}^\lambda\right) \\
&= \frac{\partial A^\mu}{\partial u^\nu} + \Gamma^\mu_{\lambda\nu}A\left(\vec{N}^\lambda\right) \\
&= \frac{\partial A^\mu}{\partial u^\nu} + \Gamma^\mu_{\lambda\nu}A^\lambda.
\end{aligned} \tag{2.30}$$

类似地，对二阶张量 A，有

$$(\nabla A)_{\mu\nu\rho} = \frac{\partial A_{\mu\nu}}{\partial u^\rho} - A\left(\frac{\partial \vec{n}_\mu}{\partial u^\rho}, \vec{n}_\nu\right) - A\left(\vec{n}_\mu, \frac{\partial \vec{n}_\nu}{\partial u^\rho}\right) = \frac{\partial A_{\mu\nu}}{\partial u^\rho} - A_{\alpha\nu}\Gamma^\alpha_{\mu\rho} - A_{\mu\alpha}\Gamma^\alpha_{\nu\rho}, \tag{2.31}$$

$$(\nabla A)^{\mu\nu}{}_\rho = \frac{\partial A^{\mu\nu}}{\partial u^\rho} - A\left(\frac{\partial \vec{N}^\mu}{\partial u^\rho}, \vec{N}^\nu\right) - A\left(\vec{N}^\mu, \frac{\partial \vec{N}^\nu}{\partial u^\rho}\right) = \frac{\partial A^{\mu\nu}}{\partial u^\rho} + A^{\alpha\nu}\Gamma^\mu_{\alpha\rho} + A^{\mu\alpha}\Gamma^\nu_{\alpha\rho}, \tag{2.32}$$

和

$$(\nabla A)^{\mu}{}_{\nu\rho} = \frac{\partial A^{\mu}{}_{\nu}}{\partial u^{\rho}} - A\left(\frac{\partial \vec{N}^{\mu}}{\partial u^{\rho}}, \vec{n}_{\nu}\right) - A\left(\vec{N}^{\mu}, \frac{\partial \vec{n}_{\nu}}{\partial u^{\rho}}\right) = \frac{\partial A^{\mu}{}_{\nu}}{\partial u^{\rho}} + A^{\alpha}{}_{\nu}\Gamma^{\mu}_{\alpha\rho} - A^{\mu}{}_{\alpha}\Gamma^{\alpha}_{\nu\rho}, \tag{2.33}$$

显然，对更高阶的张量的协变导数也可以如法炮制。我们的口诀是：上指标加，下指标减。

协变导数具有如下重要的性质：

1 **度规的协变导数都是零**。这是因为度规的定义不依赖于坐标：$g(\vec{v}_1,\vec{v}_2)=\vec{v}_1\cdot\vec{v}_2$。

2 **协变导数是一个线性操作，因此和缩并可交换，并能保留求协变导数前的张量的指标对称性**。

3 **和普通导数类似，张量积的协变导数可以按照莱布尼兹法则展开，但要注意求导指标要添加在协变导数的作用对象上**。例如，如果 φ 是标量场，$\vec{A}$ 是矢量场，那么 $\left[\nabla(\varphi\vec{A})\right]_{\mu\nu} = A_{\mu}(\nabla\varphi)_{\nu} + \varphi(\nabla A)_{\mu\nu}$。又如 A 和 B 都是二阶张量，那么 $(\nabla(A\otimes B))_{\alpha\beta\mu\nu\rho} = A_{\alpha\beta}(\nabla B)_{\mu\nu\rho} + (\nabla A)_{\alpha\beta\rho}B_{\mu\nu}$。

4 **和偏导数不同，两次协变导数一般不能交换次序**。例如对矢量场 $\vec{A}$，$\left[\nabla\left(\nabla\vec{A}\right)\right]_{\lambda\mu\nu}$ 一般和 $\left[\nabla\left(\nabla\vec{A}\right)\right]_{\lambda\nu\mu}$ 不同。

前面三条性质可以用协变导数的计算法则严格证明。最后一条会引出一个重要的概念——黎曼张量——它是下一章的主题。

2.3.5 沿着参数化曲线的协变导数

假设有一条曲线 $u^{\mu}(t)$（这里 t 是任意的参数，不一定是“长度参量”），张量场可以沿着曲线对参量 t 求协变导数——即当沿着曲线移动很小距离时，张量变化量和参量 t 的变化量之比。张量 A 沿着曲线的协变导数可以通过把 ∇A 求导指标（按我们的约定，为最后一个指标）和 $\frac{\mathrm{d}u^{\mu}}{\mathrm{d}t}$ 缩并获得。我们以矢量 A 为例来说明，沿曲线的协变导数的协变分量为

$$\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)_{\mu} = \left(\frac{\partial A_{\mu}}{\partial x^{\nu}} - \Gamma^{\lambda}_{\mu\nu}A_{\lambda}\right)\frac{\mathrm{d}x^{\nu}}{\mathrm{d}t} = \frac{\mathrm{d}A_{\mu}}{\mathrm{d}t} - \Gamma^{\lambda}_{\mu\nu}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}t}A_{\lambda}, \tag{2.34}$$

逆变分量为

$$\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)^{\mu} = \left(\frac{\partial A^{\mu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\lambda\nu}A^{\lambda}\right)\frac{\mathrm{d}x^{\nu}}{\mathrm{d}t} = \frac{\mathrm{d}A^{\mu}}{\mathrm{d}t} + \Gamma^{\mu}_{\lambda\nu}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}t}A^{\lambda}. \tag{2.35}$$

如果曲线是物理的（不依赖于坐标系选取的），那么 k 阶张量场沿着曲线的协变导数仍然是 k 阶张量。但是它仅在曲线上有定义，所以严格来说不能说是张量场。

2.4 测地线方程

在曲面上参数化的曲线 $u^\mu(\lambda)$ 的长度可以写成

$$\int \frac{\mathrm{d}s}{\mathrm{d}\lambda}\,\mathrm{d}\lambda = \int \sqrt{g_{\rho\beta}\frac{\mathrm{d}u^\rho}{\mathrm{d}\lambda}\frac{\mathrm{d}u^\beta}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda. \tag{2.36}$$

取极值对应的欧拉-拉格朗日方程是

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{g_{\rho\beta}\frac{\mathrm{d}u^\beta}{\mathrm{d}\lambda}}{\frac{\mathrm{d}s}{\mathrm{d}\lambda}} - \frac{\frac{\partial g_{\alpha\beta}}{\partial u^\rho}\frac{\mathrm{d}u^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}u^\beta}{\mathrm{d}\lambda}}{2\frac{\mathrm{d}s}{\mathrm{d}\lambda}} = 0. \tag{2.37}$$

如果用“长度参量” s 来替代任意的参量 λ，上面的方程等价于

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(g_{\rho\beta}\frac{\mathrm{d}u^\beta}{\mathrm{d}s}\right) - \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial u^\rho}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = 0 \tag{2.38}$$

也就是

$$\frac{\partial g_{\rho\beta}}{\partial u^\alpha}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} + g_{\rho\beta}\frac{\mathrm{d}^2u^\beta}{\mathrm{d}s^2} - \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial u^\rho}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = 0. \tag{2.39}$$

即

$$g_{\rho\beta}\frac{\mathrm{d}^2u^\beta}{\mathrm{d}s^2} + \frac{\partial g_{\rho\beta}}{\partial u^\alpha}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} - \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial u^\rho}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = 0. \tag{2.40}$$

两边乘以 $g^{\mu\rho}$ 得到

$$\frac{\mathrm{d}^2u^\mu}{\mathrm{d}s^2} + g^{\mu\rho}\frac{\partial g_{\rho\beta}}{\partial u^\alpha}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} - \frac{1}{2}g^{\mu\rho}\frac{\partial g_{\alpha\beta}}{\partial u^\rho}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = 0. \tag{2.41}$$

注意到第二项里的指标 α 和 β 是可以交换的，并利用 Cristoffel 符号，就立刻得到

测地线方程:

$$\frac{\mathrm{d}^2u^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = 0. \tag{2.42}$$

我们换一种更加直观的方法来理解测地线方程。曲线的单位切矢量可以写成 $\frac{\mathrm{d}\vec{x}}{\mathrm{d}s}$。用逆标架可以投影得到单位切矢量的分量

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}s} = \vec{N}^\mu \cdot \frac{\mathrm{d}\vec{x}}{\mathrm{d}s}. \tag{2.43}$$

两边对 s 求导，得到

$$\frac{\mathrm{d}^2u^\mu}{\mathrm{d}s^2} = \frac{\mathrm{d}\vec{N}^\mu}{\mathrm{d}s}\cdot\frac{\mathrm{d}\vec{x}}{\mathrm{d}s} + \vec{N}^\mu \cdot \frac{\mathrm{d}^2\vec{x}}{\mathrm{d}s^2}. \tag{2.44}$$

因为要求切矢量始终在曲面上，单位切矢量“不得不变化”：$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}s^2}$ 一般而言非零。但是，为了让路径最短，单位切矢量的“最经济”的变化方式当然是垂直于曲面变化。也就是说，对测地线，有 $\vec{N}^\mu \cdot \frac{\mathrm{d}^2\vec{x}}{\mathrm{d}s^2} = 0$。于是对于测地线，有

$$\frac{\mathrm{d}^2 u^\mu}{\mathrm{d}s^2} = \frac{\mathrm{d}\vec{N}^\mu}{\mathrm{d}s} \cdot \frac{\mathrm{d}\vec{x}}{\mathrm{d}s}. \tag{2.45}$$

注意到 $\frac{\mathrm{d}\vec{x}}{\mathrm{d}s} = \frac{\mathrm{d}u^\beta}{\mathrm{d}s}\vec{n}_\beta$，以及 $\frac{\mathrm{d}\vec{N}^\mu}{\mathrm{d}s} = \frac{\partial \vec{N}^\mu}{\partial u^\alpha}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}$，就得到

$$\frac{\mathrm{d}^2 u^\mu}{\mathrm{d}s^2} = \vec{n}_\beta \cdot \frac{\partial \vec{N}^\mu}{\partial u^\alpha}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} = -\Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s}. \tag{2.46}$$

在最后一步我们用了(2.10)。

上面的“最经济”的表述可以用沿着曲线的协变导数来更准确地描述：**测地线的单位切向量沿着测地线本身的协变导数为零**。如果考虑一般的参数 t，直接用沿曲线的协变导数公式(2.35)可以写出：

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}\vec{x}}{\mathrm{d}s}\right)^\mu = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}u^\mu}{\mathrm{d}s}\right) + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}t} = 0. \tag{2.47}$$

这里的 $\frac{\mathrm{d}u^\mu}{\mathrm{d}s}$ 是单位切矢量 $\frac{\mathrm{d}\vec{x}}{\mathrm{d}s}$ 的分量。在 (2.47) 中取 $t = s$ 显然又回到了 (2.42)。

关于测地线方程，我们来划几条重点：

1. 测地线方程(2.42)里的 s 可以换成 s 的任意线性变换 $as+b$ $(a \neq 0)$，但不能换成任意参量（因为那样会对二阶导那项产生额外的贡献）。
2. 测地线和平行移动有很直接的联系：测地线就是不断把轨迹的单位切矢量做平行移动形成的。
3. 过一个给定点的测地线一般有无穷多条——你可以随意选定方向开始移动。但一旦选定初始方向后，后续就是不断把轨迹的单位切矢量做平行移动，测地线就唯一了。
4. 连接两个点的测地线使路径长度取“稳定值”（路径长度变分为零），但不一定让路径长度取到最小值或最大值。连接两点的测地线也不一定只有有限条（例如球面的南北极之间的大圆弧有无穷多条）。

2.5 广义相对论中的粒子测地线方程

在广义相对论中，经常用 (x^0, x^1, x^2, x^3) 来描述时空点或者质点的坐标，$\mathrm{d}\tau = \sqrt{g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu}$（在平直时空或者狭义相对论中 $\mathrm{d}\tau = \sqrt{\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2}$）描述有静止质量的粒子的“原时”或“固有时”（就是粒子“感觉”自身经过的时间，或者说假想粒子携带的钟走过的时间）。对有质量且不受外力（在纯几何观点的广义相对论里，“引力”被认为并不存在）的点粒子，就可以用测地线方程(2.42)来描述

它的运动:

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}+\Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau}=0. \tag{2.48}$$

或者用粒子的四维动量 $p^\mu=m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ 来描述，就是

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau}+\Gamma^\mu_{\alpha\beta}\frac{p^\alpha p^\beta}{m}=0. \tag{2.49}$$

有时，我们会用“时间坐标” x^0 来书写上述方程，利用上面的测地线方程以及 $p^0=m\frac{\mathrm{d}x^0}{\mathrm{d}\tau}$，就有

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}x^0}+\Gamma^\mu_{\alpha\beta}\frac{p^\alpha p^\beta}{p^0}=0. \tag{2.50}$$

对静止质量为零的粒子，例如光子，它的原时为零（嗯，光子无法“感受”到时间的流动！）但是四维动量非零，我们就只能使用方程(2.50)。

最后，我们把联络的表达式 (2.16) 代入 (2.50)，得到

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}x^0}+\frac{1}{2}g^{\mu\nu}\left(\frac{\partial g_{\alpha\nu}}{\partial x^\beta}+\frac{\partial g_{\beta\nu}}{\partial x^\alpha}-\frac{\partial g_{\alpha\beta}}{\partial x^\nu}\right)\frac{p^\alpha p^\beta}{p^0}=0. \tag{2.51}$$

方程两边同时乘以 $g_{\lambda\mu}$，得到

$$g_{\lambda\mu}\frac{\mathrm{d}p^\mu}{\mathrm{d}x^0}+\frac{1}{2}\left(\frac{\partial g_{\alpha\lambda}}{\partial x^\beta}+\frac{\partial g_{\beta\lambda}}{\partial x^\alpha}-\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\right)\frac{p^\alpha p^\beta}{p^0}=0. \tag{2.52}$$

注意到括号里前两项的贡献是一样的，因为

$$\frac{\partial g_{\alpha\lambda}}{\partial x^\beta}\frac{p^\beta p^\alpha}{p^0}=\frac{\partial g_{\alpha\lambda}}{\partial x^\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}x^0}p^\alpha=\frac{\mathrm{d}g_{\alpha\lambda}}{\mathrm{d}x^0}p^\alpha=\frac{\mathrm{d}g_{\mu\lambda}}{\mathrm{d}x^0}p^\mu \tag{2.53}$$

$$\frac{\partial g_{\beta\lambda}}{\partial x^\alpha}\frac{p^\alpha p^\beta}{p^0}=\frac{\partial g_{\beta\lambda}}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}x^0}p^\beta=\frac{\mathrm{d}g_{\beta\lambda}}{\mathrm{d}x^0}p^\beta=\frac{\mathrm{d}g_{\mu\lambda}}{\mathrm{d}x^0}p^\mu \tag{2.54}$$

把上面两式代入(2.52)，并把括号里最后一项的贡献移到方程右边，就得到一个非常有趣的结果:

$$\frac{\mathrm{d}(g_{\mu\lambda}p^\mu)}{\mathrm{d}x^0}=\frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\frac{p^\alpha p^\beta}{p^0}. \tag{2.55}$$

也就是

协变动量的测地线方程:

$$\frac{\mathrm{d}p_\lambda}{\mathrm{d}x^0}=\frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\frac{p^\alpha p^\beta}{p^0}. \tag{2.56}$$

这个方程的意义在于它指出了这样一件事情: 如果度规不依赖于某个坐标，那么对应指标的动量协变分量是个守恒量！我们后面在实战中会频繁用到这个结论。

2.6 课后练习

习题 2: 用张量的坐标变换规则证明张量的缩并仍然是张量。

3. 黎曼张量

空间的内禀（指曲面上的蚂蚁就能感受到的，而非超空间视角的）弯曲以“度量的扭曲”为特征。我们已经知道，度量由度规张量描述，而度规张量的本质是取内积的运算。当两个矢量沿同一条路径平行移动时，它们的内积保持不变（请思考为什么），所以直接对度规张量计算协变导数行不通（度规的协变导数恒为零）。不过，我们可以保持一个矢量不动，另一个矢量“移动一圈”回来，这时两矢量之间的内积就会发生变化。“移动一圈”的路径至少涉及两个两个方向。这样，描述空间内禀弯曲的张量至少要涉及两个矢量以及两个不同的移动方向，也就是说，它至少得是四阶张量。这一章的主题就是研究这样一个四阶张量，它叫曲率张量或者黎曼张量，一般用字母 R 来表示。

内禀弯曲的空间还有很多其他的特征，例如：平行线会相交（三角形内角和不等于 π）；协变导数不可交换次序等等；这些性质都可以用矢量平行移动来描述，因此也可以被黎曼张量描述。

要注意的是，和上一章讨论协变导数时只需保留一阶小量的情况不同，在讨论曲率时需要保留二阶小量。

3.1 矢量“移动一圈”

我们在上一章找到了把矢量在曲面上“平行”移动的方法。不过这种“平行”从超空间的角度来看并非真正平行，而是忽略了垂直于曲面的“不可观测分量”的变化。

当曲面可以“展开”成一个平直空间（例如圆柱面或圆锥面可以展成平面），

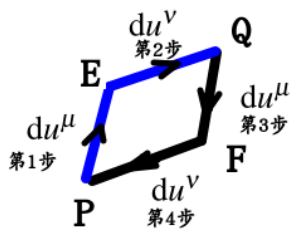


Figure 3.1: 矢量移动一圈的路径

曲面上的“平行”移动可以简化为在展开的空间中的严格平行的移动（把矢量“压向”曲面的操作和把曲面“展开”是等价的）。因此，在平直空间或者可以展开成平直空间的曲面上，“平行”移动是一个只和起点、终点相关的操作，与移动路径无关。在这种情况下，我们可以在任意两点间建立矢量的等价关系，而不只是在邻近点取一阶近似意义下的等价关系。

像球面这样不可展开为平直空间的“真正弯曲”的曲面，“平行”移动的结果一般和路径相关。为了更加具体地说明这个问题，如图3.1所示，我们来选取一个非常简单的小平行四边形：第一步沿着 $\vec{n}_\mu$ 方向移动 $\mathrm{d}u^\mu$，第二步沿着 $\vec{n}_\nu$ 方向移动 $\mathrm{d}u^\nu$，第三步沿着 $\vec{n}_\mu$ 方向移动 $-\mathrm{d}u^\mu$，第四步沿着 $\vec{n}_\nu$ 方向移动 $-\mathrm{d}u^\nu$。注意在一阶近似下，第一步和第三步中 $\vec{A}$ 的变化的贡献互相抵消，第二步和第四步中 $\vec{A}$ 的变化互相抵消，所以最后 $\vec{A}$ 在一阶近似意义下不变。但是，如果考虑二阶修正，情况就不同了：第一步和第三步导致 A_α 的净变化为

$$\Delta A_\alpha = -\frac{\partial\left(A_\beta\Gamma^\beta_{\alpha\mu}\right)}{\partial u^\nu}\mathrm{d}u^\nu\mathrm{d}u^\mu, \tag{3.1}$$

这里右边的 A_β 要视为在平行移动过程中的 A 的分量，满足 $\frac{\partial A_\beta}{\partial u^\nu} = \Gamma^\lambda_{\beta\nu}A_\lambda$。

第二步和第四步导致 A_α 的净变化为

$$\Delta A_\alpha = \frac{\partial\left(A_\beta\Gamma^\beta_{\alpha\nu}\right)}{\partial u^\mu}\mathrm{d}u^\mu\mathrm{d}u^\nu. \tag{3.2}$$

等式右边的 A_β 同样视为在平行移动过程中的 A 的分量，满足 $\frac{\partial A_\beta}{\partial u^\mu} = \Gamma^\lambda_{\beta\mu}A_\lambda$。

于是最后净变化

$$\begin{aligned}\Delta A_\alpha &= \left[\frac{\partial\left(A_\beta\Gamma^\beta_{\alpha\nu}\right)}{\partial u^\mu} - \frac{\partial\left(A_\beta\Gamma^\beta_{\alpha\mu}\right)}{\partial u^\nu}\right]\mathrm{d}u^\mu\mathrm{d}u^\nu \\ &= \left[A_\lambda\Gamma^\lambda_{\mu\beta}\Gamma^\beta_{\alpha\nu} + A_\beta\frac{\partial\Gamma^\beta_{\alpha\nu}}{\partial u^\mu} - A_\lambda\Gamma^\lambda_{\nu\beta}\Gamma^\beta_{\alpha\mu} - A_\beta\frac{\partial\Gamma^\beta_{\alpha\mu}}{\partial u^\nu}\right]\mathrm{d}u^\mu\mathrm{d}u^\nu \\ &= A_\lambda\left[\Gamma^\lambda_{\mu\beta}\Gamma^\beta_{\alpha\nu} + \frac{\partial\Gamma^\lambda_{\alpha\nu}}{\partial u^\mu} - \Gamma^\lambda_{\nu\beta}\Gamma^\beta_{\alpha\mu} - \frac{\partial\Gamma^\lambda_{\alpha\mu}}{\partial u^\nu}\right]\mathrm{d}u^\mu\mathrm{d}u^\nu \end{aligned} \tag{3.3}$$

定义黎曼张量的分量为

$$R^{\lambda}{}_{\alpha\mu\nu}=\Gamma^{\lambda}_{\mu\beta}\Gamma^{\beta}_{\alpha\nu}+\frac{\partial\Gamma^{\lambda}_{\alpha\nu}}{\partial u^{\mu}}-\Gamma^{\lambda}_{\nu\beta}\Gamma^{\beta}_{\alpha\mu}-\frac{\partial\Gamma^{\lambda}_{\alpha\mu}}{\partial u^{\nu}}. \tag{3.4}$$

就有

$$\Delta A_{\alpha}=R^{\lambda}{}_{\alpha\mu\nu}A_{\lambda}\,\mathrm{d}u^{\mu}\mathrm{d}u^{\nu}. \tag{3.5}$$

在 P 点黎曼张量吃掉四个在 P 点的四个矢量 $\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4$，就进行如下操作，把 $\vec{v}_1$ 沿着 $\vec{v}_3$ “平行”移动 ε（这里的 $\varepsilon\to 0^{+}$）；然后沿着 $\vec{v}_4$ “平行”移动 ε；然后沿着 $\vec{v}_3$ “平行”移动 $-\varepsilon$；然后沿着 $\vec{v}_4$ “平行”移动 $-\varepsilon$ 回到 P 点，矢量 $\vec{v}_1$ 的变化量除以 ε^2 后再和 $\vec{v}_2$ 取点积作为输出的实数。

我们刚才计算的其实是 $R(\vec{N}^{\lambda},\vec{n}_{\alpha},\vec{n}_{\mu},\vec{n}_{\nu})$，所以确实是黎曼张量的分量 $R^{\lambda}{}_{\alpha\mu\nu}$。

按照黎曼张量的定义，知道 $R(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4)$ 关于 $\vec{v}_3,\vec{v}_4$ 是反对称的，即 $R(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4)=-R(\vec{v}_1,\vec{v}_2,\vec{v}_4,\vec{v}_3)$。另外，如果同时把 $\vec{v}_1$ 也沿着闭合曲线“平行”移动，由于 $\vec{v}_1\cdot\vec{v}_2$ 是一个标量，在移动过程中并不会改变（或者按照“最经济”的规则，直接计算 $\mathrm{d}(\vec{v}_1\cdot\vec{v}_2)=\vec{v}_2\cdot\mathrm{d}\vec{v}_1+\vec{v}_1\cdot\mathrm{d}\vec{v}_2=0+0=0$）。也就是说：$(\vec{v}_1+\Delta\vec{v}_1)\cdot(\vec{v}_2+\Delta\vec{v}_2)=\vec{v}_1\cdot\vec{v}_2$。因为 $\Delta\vec{v}_1$ 和 $\Delta\vec{v}_2$ 都是 $O(\varepsilon^2)$，所以 $\Delta\vec{v}_1\cdot\Delta\vec{v}_2$ 可以忽略掉。那就表示 $\vec{v}_1\cdot\Delta\vec{v}_2=-\vec{v}_2\cdot\Delta\vec{v}_1$。也就是说黎曼张量对前两个吃掉的矢量也是反对称的：$R(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4)=-R(\vec{v}_2,\vec{v}_1,\vec{v}_3,\vec{v}_4)$。

由于黎曼张量的坐标分量仅仅是把标架喂给它的输出，就有

黎曼张量的反对称性质：

$$R_{\alpha\beta\mu\nu}=-R_{\beta\alpha\mu\nu} \tag{3.6}$$

$$R_{\alpha\beta\mu\nu}=-R_{\alpha\beta\nu\mu}. \tag{3.7}$$

3.2 比安基（Bianchi）代数恒等式

在局域惯性系里面，

$$R^{\lambda}{}_{\alpha\mu\nu}+R^{\lambda}{}_{\mu\nu\alpha}+R^{\lambda}{}_{\nu\alpha\mu}=\frac{\partial\Gamma^{\lambda}_{\alpha\nu}}{\partial u^{\mu}}-\frac{\partial\Gamma^{\lambda}_{\alpha\mu}}{\partial u^{\nu}}+\frac{\partial\Gamma^{\lambda}_{\mu\alpha}}{\partial u^{\nu}}-\frac{\partial\Gamma^{\lambda}_{\mu\nu}}{\partial u^{\alpha}}+\frac{\partial\Gamma^{\lambda}_{\nu\mu}}{\partial u^{\alpha}}-\frac{\partial\Gamma^{\lambda}_{\nu\alpha}}{\partial u^{\mu}}=0. \tag{3.8}$$

因为张量恒等式和坐标系选择无关，所以这个等式在任意坐标系成立，它被称之为：

比安基代数恒等式，或第一比安基恒等式

$$R_{\lambda\alpha\mu\nu}+R_{\lambda\mu\nu\alpha}+R_{\lambda\nu\alpha\mu}=0. \tag{3.9}$$

（注意我们把指标 λ 降了下来，因右边是零，这不影响等式的成立）。

轮换指标，容易写出

$$R_{\alpha\lambda\mu\nu}+R_{\alpha\mu\nu\lambda}+R_{\alpha\nu\lambda\mu}=0. \tag{3.10}$$

$$R_{\mu\alpha\lambda\nu}+R_{\mu\lambda\nu\alpha}+R_{\mu\nu\alpha\lambda}=0. \tag{3.11}$$

$$R_{\nu\alpha\lambda\mu}+R_{\nu\lambda\mu\alpha}+R_{\nu\mu\alpha\lambda}=0. \tag{3.12}$$

把上述四个等式全部相加，并利用黎曼张量的反对称性质，就得到

$$2R_{\lambda\nu\alpha\mu}+2R_{\mu\alpha\lambda\nu}=0. \tag{3.13}$$

也就是

黎曼张量的块对称性：

$$R_{\alpha\beta\mu\nu}=R_{\mu\nu\alpha\beta}. \tag{3.14}$$

注意，它是由第一比安基恒等式和黎曼张量的两条反对称性质推导出来，并不是一个独立的对称性。

容易证明：黎曼张量的对称性对“吃”进去的矢量也是成立的。例如，

$$\begin{aligned} R(\vec{v}_3,\vec{v}_4,\vec{v}_1,\vec{v}_2) &= (\vec{v}_3\cdot\vec{N}^\alpha)(\vec{v}_4\cdot\vec{N}^\beta)(\vec{v}_1\cdot\vec{N}^\mu)(\vec{v}_2\cdot\vec{N}^\nu)R_{\alpha\beta\mu\nu} \\ &= (\vec{v}_3\cdot\vec{N}^\alpha)(\vec{v}_4\cdot\vec{N}^\beta)(\vec{v}_1\cdot\vec{N}^\mu)(\vec{v}_2\cdot\vec{N}^\nu)R_{\mu\nu\alpha\beta} \\ &= R(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4) \end{aligned} \tag{3.15}$$

3.2.1 黎曼张量的独立分量个数

在 m 维空间里，虽然黎曼张量有 m^4 个分量，但是它们很多恒为零（平凡分量），还有很多相同或者只相差一个符号。那么一共有多少个非平凡的独立分量呢？

考虑非平凡的分量 $R_{\mu\nu\alpha\beta}$，根据反对称性和块对称性，不妨可以设 $\mu<\nu$，$\alpha<\beta$。这样 (μ,ν) 有 $N=\frac{m(m-1)}{2}$ 种选择。同样 (α,β) 也有 N 种选择。总计 N^2 种选

Table 3.1: 黎曼张量非恒为零的独立分量的个数

空间维数	1	2	3	4	5	6
独立 $R_{\mu\nu\alpha\beta}$ 个数	0	1	6	20	50	105

择。但是里面 $N(N-1)$ 种 $(\mu,\nu)\neq(\alpha,\beta)$ 的情况其实被重复计算了（因为根据块对称性，交换 (μ,ν) 和 (α,β) 不产生新的独立分量）。这样有 $N+\frac{N(N-1)}{2}=\frac{N(N+1)}{2}=\frac{m(m-1)(m^2-m+2)}{8}$ 种分量。

问题在于，这些分量也不完全是独立的。我们还要扣去 $\frac{m(m-1)(m-2)(m-3)}{24}$ 个四个指标都不重复的比安基恒等式代表的自由度（指标重复的情况或者是平凡的，或者已经在块对称性里完全使用了；四个指标不重复的情况只能推出三个块对称性，所以还留有一个自由度）。这样最后一共有

$$\frac{m(m-1)(m^2-m+2)}{8}-\frac{m(m-1)(m-2)(m-3)}{24}=\frac{m^2(m^2-1)}{12}$$

个独立分量。由此我们制作了表 3.1。可以看到，1 维空间没有非零分量（没有内禀弯曲），2 维空间只有一个独立分量（曲面在每点的内禀弯曲可以由单个参数表述）。

3.3 比安基（Bianchi）微分恒等式

在 P 点的局域惯性系里，P 点的所有联络为零（但是联络的偏导数一般不为零），协变导数和普通偏导数相等。因此

$$(\nabla R)^{\alpha}{}_{\beta\mu\nu\rho}=\frac{\partial^2\Gamma^{\alpha}_{\beta\nu}}{\partial u^{\mu}\partial u^{\rho}}-\frac{\partial^2\Gamma^{\alpha}_{\beta\mu}}{\partial u^{\nu}\partial u^{\rho}}. \tag{3.16}$$

轮换指标可以得到：

$$(\nabla R)^{\alpha}{}_{\beta\nu\rho\mu}=\frac{\partial^2\Gamma^{\alpha}_{\beta\rho}}{\partial u^{\nu}\partial u^{\mu}}-\frac{\partial^2\Gamma^{\alpha}_{\beta\nu}}{\partial u^{\rho}\partial u^{\mu}}. \tag{3.17}$$

$$(\nabla R)^{\alpha}{}_{\beta\rho\mu\nu}=\frac{\partial^2\Gamma^{\alpha}_{\beta\mu}}{\partial u^{\rho}\partial u^{\nu}}-\frac{\partial^2\Gamma^{\alpha}_{\beta\rho}}{\partial u^{\mu}\partial u^{\nu}}. \tag{3.18}$$

把上面三个等式相加，并降下指标 α，就得到

比安基微分恒等式，也叫做第二比安基恒等式：

$$(\nabla R)_{\alpha\beta\nu\rho\mu}+(\nabla R)_{\alpha\beta\nu\rho\mu}+(\nabla R)_{\alpha\beta\rho\mu\nu}=0. \tag{3.19}$$

虽然比安基微分恒等式是在局域惯性系推导的，但是张量恒等式和坐标系选择无关，所以它在任意坐标系都成立。

由于 ∇ 是线性操作，∇R 对前四个指标（或前四个“吃”进去的矢量）保留了 R 的所有对称性。写出黎曼张量的梯度的显示表达，也可以更加清晰地确认这一点。例如，对比

$$(\nabla R)_{\alpha\beta\mu\nu\rho} = \frac{\partial R_{\alpha\beta\mu\nu}}{\partial u^\rho} - R_{\sigma\beta\mu\nu}\Gamma^\sigma_{\alpha\rho} - R_{\alpha\sigma\mu\nu}\Gamma^\sigma_{\beta\rho} - R_{\alpha\beta\sigma\nu}\Gamma^\sigma_{\mu\rho} - R_{\alpha\beta\mu\sigma}\Gamma^\sigma_{\nu\rho}, \tag{3.20}$$

和

$$(\nabla R)_{\beta\alpha\mu\nu\rho} = \frac{\partial R_{\beta\alpha\mu\nu}}{\partial u^\rho} - R_{\sigma\alpha\mu\nu}\Gamma^\sigma_{\beta\rho} - R_{\beta\sigma\mu\nu}\Gamma^\sigma_{\alpha\rho} - R_{\beta\alpha\sigma\nu}\Gamma^\sigma_{\mu\rho} - R_{\beta\alpha\mu\sigma}\Gamma^\sigma_{\nu\rho}, \tag{3.21}$$

就可以确认 $(\nabla R)_{\alpha\beta\mu\nu\rho} = -(\nabla R)_{\beta\alpha\mu\nu\rho}$。

结合比安基代数恒等式，(3.20) 和

$$(\nabla R)_{\alpha\mu\nu\beta\rho} = \frac{\partial R_{\alpha\mu\nu\beta}}{\partial u^\rho} - R_{\sigma\mu\nu\beta}\Gamma^\sigma_{\alpha\rho} - R_{\alpha\sigma\nu\beta}\Gamma^\sigma_{\mu\rho} - R_{\alpha\mu\sigma\beta}\Gamma^\sigma_{\nu\rho} - R_{\alpha\mu\nu\sigma}\Gamma^\sigma_{\beta\rho}, \tag{3.22}$$

以及

$$(\nabla R)_{\alpha\nu\beta\mu\rho} = \frac{\partial R_{\alpha\nu\beta\mu}}{\partial u^\rho} - R_{\sigma\nu\beta\mu}\Gamma^\sigma_{\alpha\rho} - R_{\alpha\sigma\beta\mu}\Gamma^\sigma_{\nu\rho} - R_{\alpha\nu\sigma\mu}\Gamma^\sigma_{\beta\rho} - R_{\alpha\nu\beta\sigma}\Gamma^\sigma_{\mu\rho}, \tag{3.23}$$

就可以得到 $(\nabla R)_{\alpha\beta\mu\nu\rho} + (\nabla R)_{\alpha\mu\nu\beta\rho} + (\nabla R)_{\alpha\nu\beta\mu\rho} = 0.$

3.4 黎曼张量的缩并（里奇张量和黎曼曲率）

定义黎曼张量的非平凡缩并：

$$\mathscr{R}_{\mu\nu} \equiv R^\rho{}_{\mu\rho\nu}. \tag{3.24}$$

为里奇张量（Ricci tensor）。里奇张量再次缩并，就得到黎曼曲率

$$\mathsf{R} = \mathscr{R}^\mu{}_\mu = R^{\rho\mu}{}_{\rho\mu}. \tag{3.25}$$

黎曼曲率 R 也叫里奇标量（Ricci scalar）

最后，我们定义爱因斯坦张量

$$G = \mathscr{R} - \frac{1}{2}\mathsf{R}g. \tag{3.26}$$

它的协变导数满足下面的重要性质：

$$\boxed{(\nabla G)^\mu{}_{\nu\mu} = 0.} \tag{3.27}$$

证明如下：注意到

$$G^{\mu}{}_{\nu}=\mathscr{R}^{\mu}{}_{\nu}-\frac{1}{2}\mathrm{R}\delta^{\mu}{}_{\nu}. \tag{3.28}$$

以及 $\nabla\delta^{\mu}{}_{\nu}=0$，就有

$$(\nabla G)^{\mu}{}_{\nu\mu}=(\nabla\mathscr{R})^{\mu}{}_{\nu\mu}-\frac{1}{2}\delta^{\mu}{}_{\nu}(\nabla\mathrm{R})_{\mu}=(\nabla\mathscr{R})^{\mu}{}_{\nu\mu}-\frac{1}{2}(\nabla\mathrm{R})_{\nu}. \tag{3.29}$$

把上式右边写成完整的 ∇R 的形式，并利用它的对称性：

$$\begin{aligned}
(\nabla G)^{\mu}{}_{\nu\mu} &= (\nabla R)^{\alpha\mu}{}_{\alpha\nu\mu}-\frac{1}{2}(\nabla R)^{\alpha\mu}{}_{\alpha\mu\nu}.\\
&= (\nabla R)^{\alpha\mu}{}_{\alpha\nu\mu}+\frac{1}{2}\left[(\nabla R)^{\alpha\mu}{}_{\mu\nu\alpha}+(\nabla R)^{\alpha\mu}{}_{\nu\alpha\mu}\right].\\
&= (\nabla R)^{\alpha\mu}{}_{\alpha\nu\mu}+\frac{1}{2}\left[-(\nabla R)^{\mu\alpha}{}_{\mu\nu\alpha}-(\nabla R)^{\alpha\mu}{}_{\alpha\nu\mu}\right].\\
&= \frac{1}{2}(\nabla R)^{\alpha\mu}{}_{\alpha\nu\mu}-\frac{1}{2}(\nabla R)^{\mu\alpha}{}_{\mu\nu\alpha}.\\
&= 0.
\end{aligned} \tag{3.30}$$

在最后一步，我们在第二项交换了求和指标 μ 和 α。

3.5 矢量场的二次协变导数的不可交换性

考虑矢量场 A，它取二次协变导数后成为三阶张量 $(\nabla(\nabla A))$。和普通偏微商不同，协变导数一般是不可交换次序的，也就是说 $\nabla(\nabla A)$ 对两个求导指标一般不是对称的。我们先来硬核计算一下：

$$\begin{aligned}
(\nabla(\nabla A))_{\alpha\mu\nu} &= \frac{\partial(\nabla A)_{\alpha\mu}}{\partial u^{\nu}}-\Gamma^{\lambda}_{\alpha\nu}(\nabla A)_{\lambda\mu}-\Gamma^{\lambda}_{\mu\nu}(\nabla A)_{\alpha\lambda}\\
&= \frac{\partial^2 A_{\alpha}}{\partial u^{\mu}\partial u^{\nu}}-\frac{\partial(\Gamma^{\lambda}_{\alpha\mu}A_{\lambda})}{\partial u^{\nu}}-\Gamma^{\lambda}_{\alpha\nu}\left(\frac{\partial A_{\lambda}}{\partial u^{\mu}}-\Gamma^{\alpha}_{\lambda\mu}A_{\alpha}\right)-\Gamma^{\lambda}_{\mu\nu}(\nabla A)_{\alpha\lambda}\\
&= \frac{\partial^2 A_{\alpha}}{\partial u^{\mu}\partial u^{\nu}}-\left(\Gamma^{\lambda}_{\alpha\mu}\frac{\partial A_{\lambda}}{\partial u^{\nu}}+\Gamma^{\lambda}_{\alpha\nu}\frac{\partial A_{\lambda}}{\partial u^{\mu}}\right)-\Gamma^{\lambda}_{\mu\nu}(\nabla A)_{\alpha\lambda}-\frac{\partial\Gamma^{\beta}_{\alpha\mu}}{\partial u^{\nu}}A_{\beta}+\Gamma^{\lambda}_{\alpha\nu}\Gamma^{\beta}_{\lambda\mu}A_{\beta}
\end{aligned} \tag{3.31}$$

上述结果中只有最后两项关于 μ,ν 不对称，所以就有

$$(\nabla(\nabla A))_{\alpha\mu\nu}-(\nabla(\nabla A))_{\alpha\nu\mu}=A_{\beta}\left(\Gamma^{\lambda}_{\alpha\nu}\Gamma^{\beta}_{\lambda\mu}-\Gamma^{\lambda}_{\alpha\mu}\Gamma^{\beta}_{\lambda\nu}-\frac{\partial\Gamma^{\beta}_{\alpha\mu}}{\partial u^{\nu}}+\frac{\partial\Gamma^{\beta}_{\alpha\nu}}{\partial u^{\mu}}\right). \tag{3.32}$$

利用 (3.4)，就有

$$(\nabla(\nabla A))_{\alpha\mu\nu}-(\nabla(\nabla A))_{\alpha\nu\mu}=R^{\beta}{}_{\alpha\mu\nu}A_{\beta}. \tag{3.33}$$

由于我们最后的结果是一个张量方程。张量方程两边遵循一样的坐标变换规则，所以只要在一个参考系里证明了它成立，它就在任何坐标系里成立。所以我们可以只在局域惯性系里证明它，这可以大大简化计算过程：在局域惯性系里，

$$(\nabla(\nabla A))_{\alpha\mu\nu} = \frac{\partial(\nabla A)_{\alpha\mu}}{\partial u^\nu} = \frac{\partial^2 A_\alpha}{\partial u^\mu \partial u^\nu} - \frac{\partial(\Gamma^\beta_{\alpha\mu} A_\beta)}{\partial u^\nu} = \frac{\partial^2 A_\alpha}{\partial u^\mu \partial u^\nu} - \frac{\partial \Gamma^\beta_{\alpha\mu}}{\partial u^\nu} A_\beta \quad (3.34)$$

交换 μ,ν 并做差，得到

$$(\nabla(\nabla A))_{\alpha\mu\nu} - (\nabla(\nabla A))_{\alpha\nu\mu} = -\frac{\partial \Gamma^\beta_{\alpha\mu}}{\partial u^\nu} + \frac{\partial \Gamma^\beta_{\alpha\nu}}{\partial u^\mu} = R^\beta{}_{\alpha\mu\nu} A_\beta. \quad (3.35)$$

我们看到，用局域惯性系推导张量方程非常便捷，不过要注意的是，局域惯性系里的扔掉所有联络的形式只对一点成立，不能直接对它进行求导。或者说，在局域惯性系里推导张量方程时如果涉及求导，那么求导前的表达式必须保留所有联络。

3.6 测地线偏离方程

我们对测地线方程 $u^\mu(s)$ 做一个扰动 $\delta u^\mu(s) = \zeta^\mu(s)$，我们将假设 ζ^μ 是非常小的量，并要求扰动后的方程描述曲线仍然是以 s 为长度参量的测地线，那么就有

$$\frac{\mathrm{d}^2\left(u^\mu + \zeta^\mu\right)}{\mathrm{d}s^2} + \left(\Gamma^\mu_{\alpha\beta} + \frac{\partial \Gamma^\mu_{\alpha\beta}}{\partial u^\nu}\zeta^\nu\right) \frac{\mathrm{d}(u^\alpha + \zeta^\alpha)}{\mathrm{d}s} \frac{\mathrm{d}(u^\beta + \zeta^\beta)}{\mathrm{d}s} = 0. \quad (3.36)$$

上式减去微扰前的测地线方程 (2.42)，并保留到 ζ 的一阶小量

$$\frac{\mathrm{d}^2\zeta^\mu}{\mathrm{d}s^2} + \frac{\partial \Gamma^\mu_{\alpha\beta}}{\partial u^\nu}\zeta^\nu \frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} + 2\Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}\zeta^\beta}{\mathrm{d}s} = 0. \quad (3.37)$$

在局域惯性系里，方程(3.37)简化为

$$\frac{\mathrm{d}^2\zeta^\mu}{\mathrm{d}s^2} = -\frac{\partial \Gamma^\mu_{\alpha\beta}}{\partial u^\nu}\zeta^\nu \frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\frac{\mathrm{d}u^\beta}{\mathrm{d}s} \quad (3.38)$$

在 ζ^μ 都非常小时，可以近似认为 ζ^μ 是一个矢量 $\vec{\zeta}$ 的逆变分量（思考为什么）。那么根据 (2.35)，矢量 $\vec{\zeta}$ 沿着（扰动前的）测地线的协变导数的逆变分量为：

$$\left(\frac{\mathrm{d}\zeta}{\mathrm{d}s}\right)^\mu = \frac{\mathrm{d}\zeta^\mu}{\mathrm{d}s} + \Gamma^\mu_{\alpha\nu}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\zeta^\nu. \quad (3.39)$$

如果再对 $\frac{\mathrm{d}\zeta}{\mathrm{d}s}$ 这个矢量沿着测地线计算协变导数，将会得到一个很复杂的表达式。不过，我们可以**在局域惯性系里**计算，

$$\left(\frac{\mathrm{d}^2\zeta}{\mathrm{d}s^2}\right)^\mu = \frac{\mathrm{d}^2\zeta^\mu}{\mathrm{d}s^2} + \frac{\partial \Gamma^\mu_{\alpha\nu}}{\partial u^\beta}\frac{\mathrm{d}u^\beta}{\mathrm{d}s}\frac{\mathrm{d}u^\alpha}{\mathrm{d}s}\zeta^\nu. \quad (3.40)$$

再利用局域惯性系里的(3.38)，就有

$$\left(\frac{\mathrm{d}^2\zeta}{\mathrm{d}s^2}\right)^{\mu}=\left(-\frac{\partial\Gamma^{\mu}_{\alpha\beta}}{\partial u^{\nu}}+\frac{\partial\Gamma^{\mu}_{\alpha\nu}}{\partial u^{\beta}}\right)\frac{\mathrm{d}u^{\beta}}{\mathrm{d}s}\frac{\mathrm{d}u^{\alpha}}{\mathrm{d}s}\zeta^{\nu}=R^{\mu}{}_{\alpha\beta\nu}\frac{\mathrm{d}u^{\beta}}{\mathrm{d}s}\frac{\mathrm{d}u^{\alpha}}{\mathrm{d}s}\zeta^{\nu}. \tag{3.41}$$

最后结果是一个张量方程，所以在任何坐标系成立。这个张量方程被称为

测地线偏离方程

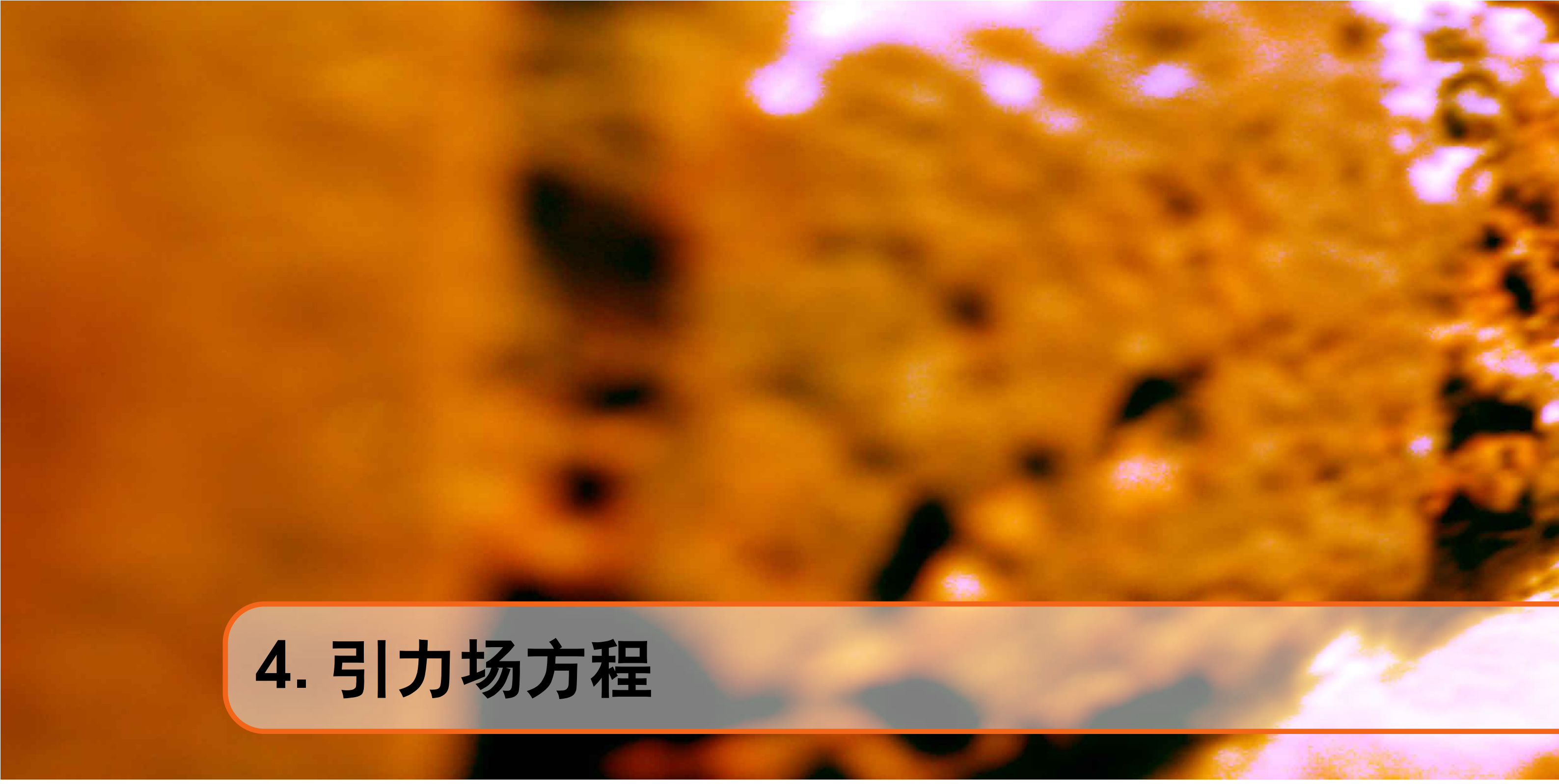
$$\left(\frac{\mathrm{d}^2\zeta}{\mathrm{d}s^2}\right)^{\mu}=R^{\mu}{}_{\alpha\beta\nu}\frac{\mathrm{d}u^{\beta}}{\mathrm{d}s}\frac{\mathrm{d}u^{\alpha}}{\mathrm{d}s}\zeta^{\nu}. \tag{3.42}$$

它是引力波测量实验的关键依据。请注意左边沿测地线的协变导数 $\left(\frac{\mathrm{d}^2\zeta}{\mathrm{d}s^2}\right)^{\mu}$ 里的括号不能去掉——因为在动态标架理论里面，求导和取指标是不可交换次序的（协变导数的公式里也可以看出这一点）。

从测地线偏离方程(3.42)可以看出，即使在 $s=0$ 时让沿测地线的协变导数 $\frac{\mathrm{d}\zeta}{\mathrm{d}s}$ 为零（两条测地线平行），当改变 s 往前移动时，黎曼曲率的存在会导致 $\frac{\mathrm{d}\zeta}{\mathrm{d}s}$ 逐渐偏离零（两条测地线变得不再平行）。

3.7 课后练习

习题 3: 计算在 $\mathfrak{R}^3$ 中的半径为 r 的球面的黎曼曲率 R。



4. 引力场方程

4.1 矩阵理论的回顾

度规的协变分量、逆变分量、混合分量都可以看成是一个实对称矩阵（混合分量实际上是单位矩阵）。为了方便我们后面做深入的研究，有必要回顾一下实对称矩阵的知识。在本小节的讨论中，我们把单位矩阵记作 I。

（如果你是线性代数大佬，请直接跳过本节）

4.1.1 共轭转置

在学习实数矩阵时，我们比较熟悉“转置”这个概念。一个矩阵 A 的转置记作 A^T。“转置”的概念推广到任意的复数矩阵就成为了“共轭转置”，就是把矩阵“放倒”的同时，顺手还把每个元素变为其共轭复数。复矩阵的共轭转置通常用 $A^\dagger$ 来表示。

例如

$$A=\begin{pmatrix}1+\mathrm{i} & 2+\mathrm{i} & \mathrm{i}\\ 3-\mathrm{i} & -4-2\mathrm{i} & 0\end{pmatrix} \tag{4.1}$$

取共轭转置后为

$$A^\dagger=\begin{pmatrix}1-\mathrm{i} & 3+\mathrm{i}\\ 2-\mathrm{i} & -4+2\mathrm{i}\\ -\mathrm{i} & 0\end{pmatrix} \tag{4.2}$$

容易证明，乘积的共轭转置等于共轭转置的倒序乘积

$$(AB)^\dagger = B^\dagger A^\dagger, \tag{4.3}$$

$$(ABC)^\dagger = C^\dagger B^\dagger A^\dagger, \tag{4.4}$$

$$\cdots \tag{4.5}$$

另外，共轭转置矩阵的行列式是原矩阵的行列式的共轭

$$\det\left(A^\dagger\right) = [\det(A)]^*, \tag{4.6}$$

这里 $\det(A)$ 表示 A 的行列式。

一个复数列向量 x 的长度平方就定义为 $x^\dagger x$。显然 $x^\dagger x$ 等于 x 的所有元素的模的平方之和，所以 $x^\dagger x \geq 0$，等号当且仅当 x 是零矢量时取到。

我们说 P 是正交矩阵，是指它是实数矩阵且满足 $P^T P = I$。推广到复数矩阵的情况，我们说 P 是酉矩阵（或幺正矩阵），是指它满足 $P^\dagger P = I$。酉矩阵对应于把矢量旋转的线性映射。

我们说 A 是实对称矩阵，是指它是实数矩阵且满足 $A = A^T$。推广到复数矩阵的情况，我们说 A 是厄米矩阵（或厄米特矩阵），是指它满足 $A = A^\dagger$。例如

$$A = \begin{pmatrix} 1 & 2+\mathrm{i} \\ 2-\mathrm{i} & -4 \end{pmatrix} \tag{4.7}$$

就是厄米矩阵。

设 A 是 $n \times n$ 厄米矩阵，通过求解方程

$$\det(\lambda I - A) = 0, \tag{4.8}$$

可以得到 n 个本征值（m 重根视为 m 个），这些本征值一定是实数。这是因为如果存在非零向量 x 使得 $Ax = \lambda x$，则 $\lambda x^\dagger x = x^\dagger A x = (Ax)^\dagger x = \lambda^* x^\dagger x$。

只出现 $m = 1$ 次的本征值对应的本征矢方向确定（但允许乘 -1），且和其他本征值的本征矢都正交。这是因为对两个本征值 $\lambda_1 \neq \lambda_2$（按前面讨论结果，它们都是实数），存在非零向量 x_1, x_2 使得 $Ax_1 = \lambda_1 x_1$，$Ax_2 = \lambda_2 x_2$，于是有 $\lambda_1 x_1^\dagger x_2 = (Ax_1)^\dagger x_2 = x_1^\dagger A x_2 = \lambda_2 x_1^\dagger x_2$，即 $x_1^\dagger x_2 = 0$。

重复 $m > 1$ 次的本征值的所有本征矢构成（和其他本征值的本征矢都正交的）m 维线性子空间。可以取一组正交基，人为地使这些本征矢也两两正交；这样的取法当然有无穷多种。

根据前述讨论，对厄米矩阵 A 总是可以取一组正交归一化的本征矢，令它们为列向量就得到一个酉矩阵 U。容易根据本征矢的定义直接验证 $U^\dagger A U$ 是一个以 A 的所有本征值为对角元的对角矩阵。

一般地，对一个方阵 S，如果存在一个酉矩阵 U 使得 $U^\dagger SU$ 是对角矩阵，则称 S 可以酉对角化。在矩阵论中有一个深奥的定理：一个矩阵 S 可以酉对角化的充分必要条件是它和自己的共轭矩阵对易 (也就是 $SS^\dagger = S^\dagger S$) ——这样的矩阵叫正规矩阵 (Normal Matrix)。显然，酉矩阵和厄米矩阵都是正规矩阵，所以都能酉对角化。

4.1.2 矩阵的迹

一个方阵 A 的所有对角元素之和称为 A 的迹（trace)，一般用 $\mathrm{Tr}(A)$ 来表示。

对一个 $n\times m$ 的矩阵 A 和一个 $m\times n$ 的矩阵 B，有

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}B_{ji} = \mathrm{Tr}(BA). \tag{4.9}$$

也就是说，两个矩阵乘积的迹和它们相乘的次序无关。

如果是多个矩阵相乘求迹，我们可以对它们进行“轮换”，例如

$$\mathrm{Tr}(ABCD) = \mathrm{Tr}(DABC) = \mathrm{Tr}(CDAB) = \mathrm{Tr}(BCDA). \tag{4.10}$$

对任意方阵 A 和同阶的满秩方阵 P 有

$$\mathrm{Tr}(P^{-1}AP) = \mathrm{Tr}(PP^{-1}A) = \mathrm{Tr}(A). \tag{4.11}$$

即相似变换保持迹不变。由此我们知道，可对角化的矩阵（例如，实对称矩阵）的迹等于它所有本征值之和。

4.1.3 对角矩阵的矩阵函数

考虑对角矩阵

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix} \tag{4.12}$$

显然对任意正整数 k，有

$$\Lambda^k = \begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \cdots & \\ & & & \lambda_n^k \end{pmatrix} \tag{4.13}$$

我们约定 $\Lambda^0 = 1$，这样上式对 $k=0$ 也成立。

显然，对任何多项式 P，有

$$P(\Lambda)=\begin{pmatrix} P(\lambda_1) & & & \\ & P(\lambda_2) & & \\ & & \dots & \\ & & & P(\lambda_n) \end{pmatrix} \tag{4.14}$$

我们还可以定义

$$e^{\Lambda}\equiv\sum_{k=0}^{\infty}\frac{1}{k!}\Lambda^k=\begin{pmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \dots & \\ & & & e^{\lambda_n} \end{pmatrix} \tag{4.15}$$

.

实际上，并不需要把讨论限制在可以展开为幂级数的函数。对任何一元实变量函数 f，我们都可以直接定义矩阵函数

$$f(\Lambda)\equiv\begin{pmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \dots & \\ & & & f(\lambda_n) \end{pmatrix} \tag{4.16}$$

.

如果 f 是一个可导函数，矩阵函数 $f(\Lambda)$ 的微小变化量就可以写成

$$\begin{aligned}
\delta f(\Lambda) &= \begin{pmatrix} \delta f(\lambda_1) & & & \\ & \delta f(\lambda_2) & & \\ & & \dots & \\ & & & \delta f(\lambda_n) \end{pmatrix} \\
&= \begin{pmatrix} f'(\lambda_1)\delta\lambda_1 & & & \\ & f'(\lambda_2)\delta\lambda_2 & & \\ & & \dots & \\ & & & f'(\lambda_n)\delta\lambda_n \end{pmatrix} \\
&= \begin{pmatrix} f'(\lambda_1) & & & \\ & f'(\lambda_2) & & \\ & & \dots & \\ & & & f'(\lambda_n) \end{pmatrix}\begin{pmatrix} \delta\lambda_1 & & & \\ & \delta\lambda_2 & & \\ & & \dots & \\ & & & \delta\lambda_n \end{pmatrix} \\
&= f'(\Lambda)\delta\Lambda
\end{aligned} \tag{4.17}$$

由于两个对角矩阵相乘是可以交换次序的，所以我们也可以这样写：$\delta f(\Lambda)=\delta\Lambda f'(\Lambda)$。

为了节省纸张（万一这本瞎写的书将来出版了呢！），我们下面把以 $\lambda_1,\lambda_2,\dots,\lambda_n$ 为对角元的对角矩阵简记为 $\mathrm{diag}(\lambda_1,\lambda_2,\dots,\lambda_n)$

4.1.4 实对称矩阵的矩阵函数

现在我们考虑任意的一个实对称矩阵 A，它可以正交对角化为

$$P^T A P = \Lambda, \tag{4.18}$$

这里 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ 是由本征值构成的对角矩阵。P 是正交矩阵（满足 $P^T P = I$）。

等式 (4.18) 两边右乘 P^T，左乘 P，就得到

$$A = P\Lambda P^T, \tag{4.19}$$

对任意一元实变量函数 f，我们定义

$$f(A) \equiv P f(\Lambda) P^T, \tag{4.20}$$

注意我们通常用 A^{-1} 来表示 A 的逆矩阵。但是对于 $f(x) = x^{-1}$，我们是否能把 $f(A)$ 写成 A^{-1}？这会引起符号的混乱吗？

其实并不会！因为这两个涵义对应的是同一个矩阵。证明如下，设 $f(x) = x^{-1}$，则

$$\Lambda f(\Lambda) = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}) = I \tag{4.21}$$

于是

$$A f(A) = P\Lambda P^T P f(\Lambda) P^T = P\Lambda f(\Lambda) P^T = P P^T = I. \tag{4.22}$$

so far so good! 但是当我们试图把(4.17)推广到实对称矩阵时，却遭遇了巨大的困难：由于 A 的微小变化 δA 和 A 未必对易，一般来说 $\delta f(A)$ 既不能写成 $f'(A)\delta A$，也不能写成 $\delta A f'(A)$，而是要写成更加“对称”的形式。

我们先从具体的几个例子入手。

对 $f(A) = A^n$ （n 为正整数），可以直接计算得出

$$\delta(A^n) = (A + \delta A)^n - A^n = A^{n-1}\delta A + A^{n-2}\delta A A + \ldots + A\delta A A^{n-2} + \delta A A^{n-1}. \tag{4.23}$$

对于 $f(A) = A^{-1}$，由于 $A f(A) = I$ 是不变量，所以 $\delta A f(A) + A\delta f(A) = 0$，左乘 A^{-1} 就可以得到

$$\delta\left(A^{-1}\right) = -A^{-1}\delta A A^{-1}. \tag{4.24}$$

容易由此得到对任意正整数 n，

$$\delta\left(A^{-n}\right) = -A^{-1}\delta A A^{-n} - A^{-2}\delta A A^{-n+1} - \ldots - A^{-n+1}\delta A A^{-2} - A^{-n}\delta A A^{-1}. \tag{4.25}$$

但是，如果是 $f(x)=e^x$ 这样有无穷多个幂次的，或者像 $f(x)=\ln|x|$ 这样都无法展成统一的幂级数的函数，该怎么办呢?

对于一般的可导函数 $f(x)$，我们很难获得 $\delta f(A)$ 和 δA 之间的明确关系。不过，我们可以退而求其次，计算 $\delta f(A)$ 的迹。由于迹的运算允许矩阵相乘交换次序，就可以得到

$$\mathrm{Tr}\left[\delta f(A)\right]=\mathrm{Tr}\left[f'(A)\delta A\right]=\mathrm{Tr}\left[\delta A f'(A)\right]. \tag{4.26}$$

这个结论的证明过程如下：

由于 $P^TP=I$ 是不变量，所以

$$\delta P^T P+P^T\delta P=0. \tag{4.27}$$

此外，有

$$\delta A=\delta(P\Lambda P^T)=P\delta\Lambda P^T+\delta P\Lambda P^T+P\Lambda\delta P^T. \tag{4.28}$$

和

$$\begin{aligned}\delta f(A) &= \delta(Pf(\Lambda)P^T)\\ &= P\left[\delta f(\Lambda)\right]P^T+\delta P f(\Lambda)P^T+Pf(\Lambda)\delta P^T\\ &= P\delta\Lambda f'(\Lambda)P^T+\delta P f(\Lambda)P^T+Pf(\Lambda)\delta P^T.\end{aligned} \tag{4.29}$$

以及按照定义

$$f'(A)=Pf'(\Lambda)P^T. \tag{4.30}$$

于是有

$$\begin{aligned}\mathrm{Tr}\left[\delta f(A)\right] &= \mathrm{Tr}\left[P\delta\Lambda f'(\Lambda)P^T\right]+\mathrm{Tr}\left[\delta P f(\Lambda)P^T\right]+\mathrm{Tr}\left[Pf(\Lambda)\delta P^T\right]\\ &= \mathrm{Tr}\left[\delta\Lambda f'(\Lambda)\right]+\mathrm{Tr}\left[\left(P^T\delta P+\delta P^T P\right)f(\Lambda)\right]\\ &= \mathrm{Tr}\left[\delta\Lambda f'(\Lambda)\right]\end{aligned} \tag{4.31}$$

和

$$\begin{aligned}\mathrm{Tr}\left[f'(A)\delta A\right] &= \mathrm{Tr}\left[Pf'(\Lambda)P^T\left(P\delta\Lambda P^T+\delta P\Lambda P^T+P\Lambda\delta P^T\right)\right]\\ &= \mathrm{Tr}\left[Pf'(\Lambda)P^TP\delta\Lambda P^T\right]+\mathrm{Tr}\left[Pf'(\Lambda)P^T\delta P\Lambda P^T\right]+\mathrm{Tr}\left[Pf'(\Lambda)P^TP\Lambda\delta P^T\right]\\ &= \mathrm{Tr}\left[Pf'(\Lambda)\delta\Lambda P^T\right]+\mathrm{Tr}\left[\Lambda P^TPf'(\Lambda)P^T\delta P\right]+\mathrm{Tr}\left[Pf'(\Lambda)\Lambda\delta P^T\right]\\ &= \mathrm{Tr}\left[P^TPf'(\Lambda)\delta\Lambda\right]+\mathrm{Tr}\left[\Lambda f'(\Lambda)P^T\delta P\right]+\mathrm{Tr}\left[f'(\Lambda)\Lambda\delta P^T P\right]\\ &= \mathrm{Tr}\left[f'(\Lambda)\delta\Lambda\right]+\mathrm{Tr}\left[\Lambda f'(\Lambda)\left(P^T\delta P+\delta P^T P\right)\right]\\ &= \mathrm{Tr}\left[f'(\Lambda)\delta\Lambda\right]\end{aligned} \tag{4.32}$$

结合 (4.31) 和 (4.32) 即得证 (4.26).

4.1.5 行列式的绝对值

由于正交矩阵的行列式为 1，等式(4.19)意味着

$$\det(A)=\det(\Lambda)=\lambda_1\lambda_2\ldots\lambda_n. \tag{4.33}$$

如果 A 是满秩矩阵（意味着 $\det(A)\neq 0$），就可以 (4.33) 两边取绝对值后再取对数，

$$\ln|\det(A)|=\ln|\lambda_1|+\ln|\lambda_2|+\ldots+\ln|\lambda_n|=\mathrm{Tr}\left(\ln|\Lambda|\right). \tag{4.34}$$

需要注意的是，在很多文献（尤其是数学书）里，$|\Lambda|$ 表示 Λ 的行列式，我们这里不采用这个习惯。在上面的写法里，$\ln|\Lambda|$ 表示对 Λ 取 $f(x)=\ln|x|$ 对应的矩阵函数。

因为 $\mathrm{Tr}[f(A)]=\mathrm{Tr}[Pf(\Lambda)P^T]=\mathrm{Tr}[P^TPf(\Lambda)]=\mathrm{Tr}[f(\Lambda)]$。令 $f(x)=\ln|x|$，即有 $\mathrm{Tr}\left(\ln|A|\right)=\mathrm{Tr}\left(\ln|\Lambda|\right)$。再结合(4.34)，就有

$$\ln|\det(A)|=\mathrm{Tr}\left(\ln|A|\right). \tag{4.35}$$

注意到 $f(x)=\ln|x|$ 这个函数的导函数为 $\frac{1}{x}$（注意分母是 x 而不是 $|x|$，因为 $\ln|x|$ 在 $x<0$ 时是单调下降的），就得到

$$\delta\ln|\det(A)|=\mathrm{Tr}\left(A^{-1}\delta A\right). \tag{4.36}$$

或者其等价的形式

$$\delta|\det(A)|=|\det(A)|\mathrm{Tr}\left(A^{-1}\delta A\right). \tag{4.37}$$

如果取 A 为度规的协变形式 $g_{\mu\nu}$，那么根据 (4.37)就有

$$\delta|\det(g)|=|\det(g)|g^{\mu\nu}\delta g_{\mu\nu}. \tag{4.38}$$

这里我们用 $\det(g)$ 来表示协变形式 $g_{\mu\nu}$ 的行列式。之所以可以把下标省略掉，是因为在广义相对论里一般都不讨论逆变 $g^{\mu\nu}$ 的行列式。

根据(4.24)就有

$$\delta g^{\mu\nu}=-g^{\mu\alpha}g^{\nu\beta}\delta g_{\alpha\beta}. \tag{4.39}$$

我们绕了一个大圈，其实主要就是为了获得(4.38) 和 (4.39) 这两个结果。

4.2 物理体积元

在坐标变换 $u\to\tilde{u}$ 下，按张量的变换规则

$$\tilde{g}_{\mu\nu}=\frac{\partial u^{\alpha}}{\partial\tilde{u}^{\mu}}g_{\alpha\beta}\frac{\partial u^{\beta}}{\partial\tilde{u}^{\nu}}. \tag{4.40}$$

把上式右边看成三个矩阵相乘，按照乘积的行列式等于行列式的乘积，知道

$$\det(\tilde{g})=\left[\det\left(\frac{\partial u}{\partial\tilde{u}}\right)\right]^2\det(g). \tag{4.41}$$

注意一个有趣的地方：坐标变换不改变 $\det(g)$ 的正负号。

注意到坐标体积元的变换规则为

$$\left|\det\left(\frac{\partial u}{\partial\tilde{u}}\right)\right||\mathrm{d}\tilde{u}^1\mathrm{d}\tilde{u}^2\ldots\mathrm{d}\tilde{u}^m|=|\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m| \tag{4.42}$$

我们可以取 $\sqrt{|\det(g)|}|\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m|$ 作为空间的“物理体积元”——它在坐标变换下是不变量（也就是标量）！为了书写方便，我们采用 $\mathrm{d}V_{\mathrm{phys}}$ 这个符号来标记物理体积元，用 $\mathrm{d}V_{\mathrm{coor}}$ 这个符号来标记空间体积元。也就是说 $\mathrm{d}V_{\mathrm{coor}}\equiv|\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m|$，$\mathrm{d}V_{\mathrm{phys}}\equiv\sqrt{|\det(g)|}\mathrm{d}V_{\mathrm{coor}}=\sqrt{|\det(g)|}|\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m|$。

体积元当然是用来进行体积分的，我们在进行体积分时，默认都沿着所有坐标增大的方向进行积分，以此来保证 $\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m=|\mathrm{d}u^1\mathrm{d}u^2\ldots\mathrm{d}u^m|$。当然，你可能会产生一些困惑——是不是允许“体积元”带符号会更加自然和灵活？答案是肯定的，但是这样会涉及很多可能会毁掉幼儿园快乐时光的新概念（例如，微分形式），我暂时还不打算这么干。

4.2.1 弯曲空间的高斯定理

有了物理体积的概念，我们可以把静电学的高斯定理推广到弯曲空间。

用联络的表达式(2.16)可以直接得到下面的恒等式：

$$\Gamma^{\lambda}_{\mu\lambda}=\frac{1}{2}g^{\alpha\beta}\frac{\partial g_{\alpha\beta}}{\partial u^{\mu}} \tag{4.43}$$

再利用(4.38)，有

$$\frac{\partial\sqrt{|\det(g)|}}{\partial u^{\mu}}=\frac{1}{2\sqrt{|\det(g)|}}\frac{\partial|\det(g)|}{\partial u^{\mu}}=\frac{\sqrt{|\det(g)|}}{2}g^{\alpha\beta}\frac{\partial g_{\alpha\beta}}{\partial u^{\mu}}=\sqrt{|\det(g)|}\Gamma^{\lambda}_{\mu\lambda}. \tag{4.44}$$

由此可以证明任意一个矢量场 A 的“物理散度” $(\nabla A)^{\mu}{}_{\mu}$ 和物理体积元的乘积等于 $\sqrt{|\det(g)|}A$ 的“坐标散度”和坐标体积元的乘积，也就是

$$(\nabla A)^{\mu}{}_{\mu}\mathrm{d}V_{\mathrm{phys}}=\frac{\partial\left(\sqrt{|\det(g)|}A^{\mu}\right)}{\partial u^{\mu}}\mathrm{d}V_{\mathrm{coor}}. \tag{4.45}$$

实际上，上式两边都代表 A 在小体积元表面上的法向通量（净流出量）。只不过左边是在物理空间进行表述的，右边是在数学坐标空间进行表述的。由于表面净流出量是可加量，所以对一块空间区域 Ω 而言，

$$\int_{\Omega}(\nabla A)^{\mu}{}_{\mu}\mathrm{d}V_{\mathrm{phys}}=\int_{\Omega}\frac{\partial\left(\sqrt{|\det(g)|}A^{\mu}\right)}{\partial u^{\mu}}\mathrm{d}V_{\mathrm{coor}}=\int_{\partial\Omega}A^{\mu}\mathrm{d}S_{\mu}. \tag{4.46}$$

这里 $\partial\Omega$ 是区域 Ω 的边界，$\mathrm{d}S$ 是区域表面上的法向面积元。

4.3 在度规扰动下张量的变化

4.3.1 关于 $\delta g_{\mu\nu}$ 是不是张量的理解

为了更轻松地理解这一章地内容，我们可以采用一种非主流的观点：度规协变形式的的变分可以看成是对超空间度量的法则进行了改动，导致 $g_{\mu\nu}=\vec{n}_\mu\cdot\vec{n}_\nu$ 发生了改变，变化量为 $\delta g_{\mu\nu}$。那么在这种观点下，$\delta g_{\mu\nu}$ 是不是张量呢？从我们对张量的幼儿园版定义（吃进矢量吐出数的机器）来看，它还不能完全算张量。微扰前的空间和微扰后的空间其实是两个**共享坐标系的**不同空间。我们可以用原空间的度规对原空间的张量进行指标升降，也可以用新空间的度规（协变形式为 $g_{\mu\nu}+\delta g_{\mu\nu}$，逆变形式为 $g^{\mu\nu}+\delta g^{\mu\nu}=g^{\mu\nu}-g^{\mu\alpha}g^{\nu\beta}\delta g_{\alpha\beta}$）对新空间的张量进行指标升降。事实上，容易验证 (4.39) 和下面的“新坐标系指标升降规则”是等价的：

$$g^{\mu\nu}+\delta g^{\mu\nu}=(g^{\mu\alpha}+\delta g^{\mu\alpha})(g^{\nu\beta}+\delta g^{\nu\beta})(g_{\alpha\beta}+\delta g_{\alpha\beta}). \tag{4.47}$$

但是，δg 作为两个不同空间的张量之差，却不是任何一个空间里的张量——它不能用任何一空间的度规进行指标升降——这就是为什么 (4.39) 中 δg 不符合指标升降规则的原因。

不过，“共享坐标系”这个特殊设定使得微扰前和微扰后的空间里的张量服从一样的坐标变换规则，那么两个空间的张量差 δg 也服从同样的坐标变换规则。从这个观点看，δg 又像是张量。为了更好地区分，我们可以给这种“满足坐标变换规则，但是不满足指标升降规则”的 δg 起个名字，比如叫作**张量变分**。

但是等等，为什么要自己随便起个名字？文献中没有标准的名字吗？大概率是没有的（至少我没看到过）。我们选择的超空间叙事并不是文献中的标准视角。文献中一般的看法是：只要满足坐标变换规则的就是张量，指标升降会把一个张量变成另一个（即不承认不同形式的分量是同一个机器吐的数）。按照这种观点，$\delta g_{\mu\nu}$ 就是货真价实的张量；$\delta g^{\mu\nu}$ 是 $g_{\mu\nu}$ 的逆矩阵的变分，和把 $\delta g_{\mu\nu}$ 指标升上来的张量是两个不同的张量（相差一个负号）。

无论采用何种观点，为了不在混乱的符号中迷失，我们可以**定义一个既满足坐标变换规则也满足指标升降规则的货真价实张量 $\delta\hat{g}$，并约定 $\delta\hat{g}_{\mu\nu}=\delta g_{\mu\nu}$（或者**

等价的 $\delta\hat{g}^{\mu\nu}=-\delta g^{\mu\nu}$)。要注意的是，$\delta\hat{g}^{\mu}{}_{\nu}\equiv g^{\mu\lambda}\delta\hat{g}_{\lambda\nu}$ 和 $\delta g^{\mu}{}_{\nu}$ 之间并无关系，前者是通过把 $\delta g_{\mu\nu}$ 升指标获得的（一般非零），后者是扰动前后两个空间的 $g^{\mu}{}_{\nu}=\delta^{\mu}_{\nu}$ 之差（恒为零）。

4.3.2 其他张量在度规扰动下的变化

虽然会被冠以非主流之名，但为了保持“幼儿园版”的风格，我就继续以我们的叙事角度出发讨论问题。这会导致更多“离经叛道”的理解方式的出现——但不影响任何计算结果。

从超空间的视角看，扰动后的标架没有改变（虽然它们之间的内积被重新定义了），所以所有其他没有被扰动的（即和空间几何无关的）张量的协变分量不变。但是，由于 $g_{\mu\nu}$ 改变了，其逆矩阵——也就是用来组装逆标架的系数都发生了改变，这导致扰动后的空间里的逆标架和扰动前空间里的逆标架不同。既然逆标架都发生了改变，那么所有和空间几何无关的张量的逆变分量也都发生了改变——这实际上正是空间几何和物质作用量对话的方式之一（另一种方式是通过影响物理体积元的大小）。例如，设 A 是矢量场，在扰动 $\delta g_{\mu\nu}$ 下，其协变分量不变，逆变分量的变化量 $\delta A^{\mu}=\delta(g^{\mu\nu}A_{\nu})=\delta g^{\mu\nu}A_{\nu}$。

那么和空间几何无关且没有逆变指标的标量，在度规扰动时就一定没有变化吗？未必！这涉及到标量是怎么定义的。如果是原生的标量（比如圆周率，或者一个底层的标量场），确实不会在度规扰动时发生变化。但是如果是用高阶张量缩并得到的标量，在度规扰动下就会发生改变——原因是我们定义缩并时需要同时投喂标架和逆标架，而逆标架被度规扰动给改变了。

例如，一个和空间几何无关的二阶张量场 T 的缩并 T^{μ}_{μ} 在度规扰动时的变化为

$$\delta(T^{\mu}_{\mu})=\delta(g^{\mu\nu}T_{\nu\mu})=T_{\nu\mu}\delta g^{\mu\nu}=-T_{\mu\nu}\delta\hat{g}^{\mu\nu}=-T^{\mu\nu}\delta\hat{g}_{\mu\nu}. \tag{4.48}$$

和空间几何无关的矢量场 A 的缩并 $A^{\mu}A_{\mu}$ 在度规扰动下的变化有三种不同的计算方法，第一种是把它看成是二阶张量场 $A\otimes A$ 的两个指标的缩并，直接套用上面的结果得到 $\delta(A_{\mu}A^{\mu})=-A^{\mu}A^{\nu}\delta\hat{g}_{\mu\nu}$。第二种是先计算逆变分量的变化

$$\delta A^{\mu}=\delta(g^{\mu\nu}A_{\nu})=A_{\nu}\delta g^{\mu\nu}=-A_{\nu}\delta\hat{g}^{\mu\nu}. \tag{4.49}$$

那么

$$\delta(A_{\mu}A^{\mu})=A_{\mu}\delta A^{\mu}=-A_{\mu}A_{\nu}\delta\hat{g}^{\mu\nu}=-A^{\mu}A^{\nu}\delta\hat{g}_{\mu\nu}. \tag{4.50}$$

第三种方法颇有点“故意多绕几个圈”的感觉，仅仅是为了演示各种算法的自洽

性，

$$\begin{aligned}
\delta(A_\mu A^\mu) &= \delta(A^\mu A^\nu g_{\mu\nu}) \\
&= (\delta A^\mu)A^\nu g_{\mu\nu} + A^\mu(\delta A^\nu)g_{\mu\nu} + A^\mu A^\nu \delta g_{\mu\nu} \\
&= (-\delta\hat{g}^{\mu\alpha}A_\alpha)A^\nu g_{\mu\nu} - A^\mu(-\delta\hat{g}^{\nu\alpha}A_\alpha)g_{\mu\nu} + A^\mu A^\nu \delta\hat{g}_{\mu\nu} \\
&= -A_\mu A_\alpha \delta\hat{g}^{\mu\alpha} - A_\nu A_\alpha \delta\hat{g}^{\nu\alpha} + A^\mu A^\nu \delta\hat{g}_{\mu\nu} \\
&= -2A_\mu A_\nu \delta\hat{g}^{\mu\alpha} + A^\mu A^\nu \delta\hat{g}_{\mu\nu} \\
&= -A^\mu A^\nu \delta\hat{g}_{\mu\nu}
\end{aligned} \tag{4.51}$$

最后，我们来看一个特殊的量——联络在度规扰动下的变化量 $\delta\Gamma^\lambda_{\ \mu\nu}$。注意到联络的坐标变换规则(2.19)里多出来的那项 $\frac{\partial\tilde{u}^\lambda}{\partial u^\alpha}\frac{\partial^2 u^\alpha}{\partial\tilde{u}^\mu\partial\tilde{u}^\nu}$ 仅依赖于新旧坐标之间的关系，而度规扰动前后的空间是共享坐标系的，所以在 $\delta\Gamma^\lambda_{\ \mu\nu}$ 的坐标变换规则里，多出来的项消掉了，使得 $\delta\Gamma^\lambda_{\mu\nu}$ 满足张量的坐标变换规则——也就是说联络本身虽然不是张量，但是它在度规扰动下的改变量 $\delta\Gamma^\lambda_{\ \mu\nu}$ 反而是一个张量。

事实上，利用 (2.16) 可以直接计算得到 $\delta\Gamma$ 的显式张量形式：

$$\delta\Gamma^\lambda_{\ \mu\nu} = \frac{1}{2}\left[(\nabla\delta\hat{g})_{\nu\ \mu}^{\ \lambda} + (\nabla\delta\hat{g})_{\mu\ \nu}^{\ \lambda} - (\nabla\delta\hat{g})_{\mu\nu}^{\quad\lambda}\right]. \tag{4.52}$$

或者把指标 λ 降下来，得到

$$\boxed{\delta\Gamma_{\lambda\mu\nu} = \frac{1}{2}\left[(\nabla\delta\hat{g})_{\nu\lambda\mu} + (\nabla\delta\hat{g})_{\mu\lambda\nu} - (\nabla\delta\hat{g})_{\mu\nu\lambda}\right].} \tag{4.53}$$

4.3.3 响应张量

反过来，如果知道一个标量 W 在任意坐标扰动下的变化为

$$\delta W = F^{\mu\nu}\delta\hat{g}_{\mu\nu}, \tag{4.54}$$

我们考虑在坐标变换 $u \to \tilde{u}$ 下 $F^{\mu\nu}$ 将如何改变。因为 δW 是坐标变换下的不变量（这一点和它是不是组装的没关系），而 $\delta\hat{g}$ 满足坐标变换规则

$$\delta\hat{g}_{\mu\nu} = \frac{\partial\tilde{u}^\alpha}{\partial u^\mu}\frac{\partial\tilde{u}^\beta}{\partial u^\nu}\tilde{\delta\hat{g}}_{\alpha\beta} \tag{4.55}$$

于是

$$\tilde{F}^{\alpha\beta}\tilde{\delta\hat{g}}_{\alpha\beta} = \delta W = F^{\mu\nu}\delta\hat{g}_{\mu\nu} = F^{\mu\nu}\frac{\partial\tilde{u}^\alpha}{\partial u^\mu}\frac{\partial\tilde{u}^\beta}{\partial u^\nu}\tilde{\delta\hat{g}}_{\alpha\beta} \tag{4.56}$$

由于度规扰动 $\tilde{\delta\hat{g}}_{\alpha\beta}$ 是任意的，所以

$$\tilde{F}^{\alpha\beta} = F^{\mu\nu}\frac{\partial\tilde{u}^\alpha}{\partial u^\mu}\frac{\partial\tilde{u}^\beta}{\partial u^\nu}. \tag{4.57}$$

这说明 F 的逆变形式满足张量的坐标变换规则。或者说，(4.54) 可以作为二阶张量 F 的定义式。

既然 F 是新定义的，我们无须担心它是否满足指标升降规则，而可以直接用张量指标升降规则定义 $F_{\mu\nu}$，这样

$$\delta W = F_{\alpha\beta}\delta\hat{g}^{\alpha\beta}. \tag{4.58}$$

也是成立的。

为了让语言更加形象，我们把在度规扰动下给出 (4.54)（或者等价的 (4.58)）的二阶张量 F 称为标量 W 的**响应张量**（嗯，这名字也是我随口编的）。根据(4.48)，我们立刻有了一个有趣的定理：

> 一个和空间几何无关的二阶张量 T 的迹 $(T^{\mu}_{\ \mu})$ 的响应张量就是 $-T$。

物理体积元作为标量，可以计算响应张量。根据 (4.38) 容易得到

$$\delta\sqrt{|\det(g)|} = \frac{\sqrt{|\det(g)|}}{2}g^{\mu\nu}\delta g_{\mu\nu}. \tag{4.59}$$

所以

$$\delta \mathrm{d}V_{\text{phys}} = \frac{1}{2}\mathrm{d}V_{\text{phys}}g^{\mu\nu}\delta g_{\mu\nu}. \tag{4.60}$$

也就是说，

> 物理体积元的响应张量等于它和度规的乘积的一半。

最后，我们来研究里奇标量的响应张量。

$$\delta \mathsf{R} = \delta\left(g^{\mu\nu}\mathscr{R}_{\mu\nu}\right) = \mathscr{R}_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta\mathscr{R}_{\mu\nu}. \tag{4.61}$$

在共享坐标系的前提下，偏微分算符和微扰算符 δ 可以交换次序，以及注意到 $\delta\Gamma$ 是张量

$$\begin{aligned}
\delta\mathscr{R}_{\mu\nu} &= \delta R^{\alpha}_{\ \mu\alpha\nu} \\
&= \delta\left(\frac{\partial\Gamma^{\alpha}_{\mu\nu}}{\partial u^{\alpha}} - \frac{\partial\Gamma^{\alpha}_{\mu\alpha}}{\partial u^{\nu}} + \Gamma^{\lambda}_{\mu\nu}\Gamma^{\alpha}_{\lambda\alpha} - \Gamma^{\lambda}_{\mu\alpha}\Gamma^{\alpha}_{\lambda\nu}\right) \\
&= \frac{\partial\delta\Gamma^{\alpha}_{\mu\nu}}{\partial u^{\alpha}} - \frac{\partial\delta\Gamma^{\alpha}_{\mu\alpha}}{\partial u^{\nu}} + \Gamma^{\alpha}_{\lambda\alpha}\delta\Gamma^{\lambda}_{\mu\nu} + \Gamma^{\lambda}_{\mu\nu}\delta\Gamma^{\alpha}_{\lambda\alpha} - \Gamma^{\alpha}_{\lambda\nu}\delta\Gamma^{\lambda}_{\mu\alpha} - \Gamma^{\lambda}_{\mu\alpha}\delta\Gamma^{\alpha}_{\lambda\nu} \\
&= \left(\frac{\partial\delta\Gamma^{\alpha}_{\mu\nu}}{\partial u^{\alpha}} + \Gamma^{\alpha}_{\lambda\alpha}\delta\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\lambda}_{\mu\alpha}\delta\Gamma^{\alpha}_{\lambda\nu} - \Gamma^{\lambda}_{\alpha\nu}\delta\Gamma^{\alpha}_{\mu\lambda}\right) \\
&\quad - \left(\frac{\partial\delta\Gamma^{\alpha}_{\mu\alpha}}{\partial u^{\nu}} + \Gamma^{\alpha}_{\lambda\nu}\delta\Gamma^{\lambda}_{\mu\alpha} - \Gamma^{\lambda}_{\mu\nu}\delta\Gamma^{\alpha}_{\lambda\alpha} - \Gamma^{\lambda}_{\alpha\nu}\delta\Gamma^{\alpha}_{\mu\lambda}\right) \\
&= (\nabla\delta\Gamma)^{\alpha}_{\ \mu\nu\alpha} - (\nabla\delta\Gamma)^{\alpha}_{\ \mu\alpha\nu}.
\end{aligned} \tag{4.62}$$

再注意到度规 g 的协变导数为零，所以直接可以把 $g^{\mu\nu}$ 挪到协变导数符号后面进行升指标

$$
\begin{aligned}
& g^{\mu\nu}\delta\mathscr{R}_{\mu\nu} \\
= & (\nabla\delta\Gamma)^{\alpha\nu}{}_{\nu\alpha} - (\nabla\delta\Gamma)^{\alpha\nu}{}_{\alpha\nu} \\
= & \frac{1}{2}\left[(\nabla\nabla\delta\hat{g})^{\nu\alpha}{}_{\nu\alpha} + (\nabla\nabla\delta\hat{g})_{\nu}{}^{\alpha\nu}{}_{\alpha} - (\nabla\nabla\delta\hat{g})^{\nu}{}_{\nu}{}^{\alpha}{}_{\alpha} - (\nabla\nabla\delta\hat{g})^{\nu\alpha}{}_{\alpha\nu} - (\nabla\nabla\delta\hat{g})_{\alpha}{}^{\alpha\nu}{}_{\nu} + (\nabla\nabla\delta\hat{g})^{\nu}{}_{\alpha}{}^{\alpha}{}_{\nu}\right] \\
= & (\nabla\nabla\delta\hat{g})^{\mu\nu}{}_{\mu\nu} - (\nabla\nabla\delta\hat{g})^{\mu}{}_{\mu}{}^{\nu}{}_{\nu}
\end{aligned}
\tag{4.63}
$$

最后的结果是

$$
\delta\mathsf{R} = \mathscr{R}_{\mu\nu}\delta g^{\mu\nu} + (\nabla F)^{\alpha}{}_{\alpha} = -\mathscr{R}^{\mu\nu}\delta\hat{g}_{\mu\nu} + \left[(\nabla\nabla\delta\hat{g})^{\mu\nu}{}_{\mu\nu} - (\nabla\nabla\delta\hat{g})^{\mu}{}_{\mu}{}^{\nu}{}_{\nu}\right]. \tag{4.64}
$$

虽然我们无法把后面括号内的两项写成一个张量乘以 $\delta\hat{g}_{\mu\nu}$ 的形式，但这两项具有散度的形式，在对物理体积元积分后会转化为边界积分。在做变分推导时，我们一般固定边界上的所有场，所以这两项没有积分贡献。所以，在作用量正比于 R 的情形我们通常认为 R 的“有效”响应张量是 $-\mathscr{R}$ 并直接忽略括号内的两项。

4.4 作用量和爱因斯坦方程

经典的广义相对论认为引力并不是一种真正的力（至少爱因斯坦是这样认为的，但也有一些大物理学家对此仍然持保留意见）：时空的弯曲影响着物质的运动，物质的分布和运动也影响着时空的弯曲。在暂时不考虑量子效应的情况下，我们希望用一个统一的、包含时空度规在内的作用量来描述物质和时空之间的这种互动。

在空间一小块区域内，有物质作用量 $\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}}$（这里 $\mathscr{L}_m$ 是物质作用量密度），我们还想添加描述这块区域有多弯曲的“几何作用量”。最最简单的，包含时空的内禀弯曲的信息的标量就是里奇标量 $\mathsf{R} \equiv R^{\alpha\beta}{}_{\alpha\beta}$。因此我们不妨假设几何作用量是 $\lambda \mathsf{R}\mathrm{d}V_{\mathrm{phys}}$，这里 λ 是个待实验测量的常量——它有两个作用：一是让作用量的量纲正确；二是允许我们调节这个理论中时空对物质的响应的强弱程度。

物质的“能量动量张量”定义为物质作用量 $\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}}$ 的响应张量的两倍。因此物质作用量的变化量可以用“能量动量张量密度” T 表示出来

$$
\delta(\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}}) = \frac{1}{2}\left(T^{\mu\nu}\mathrm{d}V_{\mathrm{phys}}\right)\delta\hat{g}_{\mu\nu}. \tag{4.65}
$$

如果我们在固定坐标区域内对度规进行扰动，并假设在其边界上所有场都没

有扰动（这样散度项的积分都可以忽略），那么作用量的变分等于

$$\begin{aligned}\delta S &= \delta\int\left(\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}} + \lambda \mathsf{R} \mathrm{d}V_{\mathrm{phys}}\right) \\ &= \int\left(\delta(\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}}) + \lambda \mathrm{d}V_{\mathrm{phys}}\delta\mathsf{R} + \lambda\mathsf{R}\delta\mathrm{d}V_{\mathrm{phys}}\right) \\ &= \int\left(\frac{1}{2}T^{\mu\nu} - \lambda\mathscr{R}^{\mu\nu} + \frac{1}{2}\lambda\mathsf{R}g^{\mu\nu}\right)\delta\hat{g}_{\mu\nu}\mathrm{d}V_{\mathrm{phys}} \\ &= \int\left(\frac{1}{2}T^{\mu\nu} - \lambda G^{\mu\nu}\right)\delta\hat{g}_{\mu\nu}\mathrm{d}V_{\mathrm{phys}} \end{aligned} \tag{4.66}$$

我们在推导过程中忽略了体积分为零的散度项。并在最后利用了爱因斯坦张量的定义 $G^{\mu\nu} = \mathscr{R}^{\mu\nu} - \frac{1}{2}\mathsf{R}g^{\mu\nu}$。

让作用量的变分为零，即得到爱因斯坦方程

$$G^{\mu\nu} = \frac{1}{2\lambda}T^{\mu\nu}. \tag{4.67}$$

我们后面会展示：通过在弱场低速近似下和牛顿引力理论对比，可确定 $\lambda = \frac{1}{16\pi G_N}$，这里 $G_N \approx 6.67\times 10^{-11}\mathrm{m}^3/\mathrm{s}^2/\mathrm{kg}$ 是牛顿引力常数。所以最后的

爱因斯坦方程:

$$G^{\mu\nu} = 8\pi G_N T^{\mu\nu}. \tag{4.68}$$

例题 4: 经典实标量场 φ 的作用量密度等于

$$\mathscr{L}_m = -\frac{1}{2}(\nabla\varphi)^\mu(\nabla\varphi)_\mu - V(\varphi). \tag{4.69}$$

计算它的能量动量张量密度。

解答：因为 $-\frac{1}{2}(\nabla\varphi)^\mu(\nabla\varphi)_\mu$ 是一个和空间几何无关的二阶张量 $-\frac{1}{2}(\nabla\varphi)^\mu(\nabla\varphi)^\nu$ 的迹，所以它的响应张量为自身乘以 -1，即 $\frac{1}{2}(\nabla\varphi)^\mu(\nabla\varphi)^\nu$；而 $V(\varphi)$ 是一个底层标量，响应张量为零。于是 $\mathscr{L}_m$ 的总响应张量为 $\frac{1}{2}(\nabla\varphi)^\mu(\nabla\varphi)^\nu$。

$$\delta(\mathscr{L}_m \mathrm{d}V_{\mathrm{phys}}) = \mathrm{d}V_{\mathrm{phys}}\delta\mathscr{L}_m + \mathscr{L}_m\delta\mathrm{d}V_{\mathrm{phys}} = \frac{1}{2}\left[(\nabla\varphi)^\mu(\nabla\varphi)^\nu + g^{\mu\nu}\mathscr{L}_m\right]\delta\hat{g}_{\mu\nu}\mathrm{d}V_{\mathrm{phys}}. \tag{4.70}$$

于是能量动量张量，也就是 $\mathscr{L}_m\mathrm{d}V_{\mathrm{phys}}$ 的响应张量的两倍，等于 $\left[(\nabla\varphi)^\mu(\nabla\varphi)^\nu + g^{\mu\nu}\mathscr{L}_m\right]\mathrm{d}V_{\mathrm{phys}}$。能量动量张量密度就是

$$T^{\mu\nu} = (\nabla\varphi)^\mu(\nabla\varphi)^\nu + g^{\mu\nu}\mathscr{L}_m. \tag{4.71}$$

5. 未完待续

Bibliography

References