

电动力学笔记

Buwai. Lee@School of Physics, NJU

2015 秋季学期

Table of Formulas in 3D Vector Analysis

On grad:

$$\nabla(fg) = f\nabla g + g\nabla f$$

$$\nabla(A \cdot B) = (A \cdot \nabla)B + (B \cdot \nabla)A + A \times (\nabla \times B) + B \times (\nabla \times A)$$

$$\nabla \left(\frac{1}{|r - r'|} \right) = - \frac{r - r'}{|r - r'|^3}$$

On div:

$$\nabla \cdot (fA) = (\nabla f) \cdot A + f\nabla \cdot A$$

$$\nabla \cdot (A \times B) = B \cdot (\nabla \times A) - A \cdot (\nabla \times B)$$

$$\nabla \cdot \left(\frac{r - r'}{|r - r'|^3} \right) = 4\pi\delta^3(r - r')$$

On rot:

$$\nabla \times (fA) = f\nabla \times A + \nabla f \times A$$

$$\nabla \times (A \times B) = (B \cdot \nabla)A - (A \cdot \nabla)B + (\nabla \cdot B)A - (\nabla \cdot A)B$$

$$\nabla \times (\nabla \times A) = \nabla(\nabla \cdot A) - \nabla \cdot (\nabla A) = \nabla(\nabla \cdot A) - \nabla^2 A$$

$$\hat{n} \cdot (\nabla \times \hat{n}) = 0 \text{ } (\hat{n} \text{ is the normal vector of a surface})$$

On ∇^2 :

$$\nabla^2(fg) = g\nabla^2 f + 2\nabla f \cdot \nabla g + f\nabla^2 g$$

$$\nabla \cdot (f\nabla g) = f\nabla^2 g + \nabla f \cdot \nabla g$$

$$\nabla \cdot (f\nabla g - g\nabla f) = f\nabla^2 g - g\nabla^2 f$$

$$\nabla \cdot (\nabla^2 f) = \nabla^2(\nabla \cdot f)$$

$$\nabla^2 \left(\frac{1}{|r - r'|} \right) = -4\pi\delta^3(r - r')$$

On Integral:

$$\int_S d\omega = \int_{\partial S} \omega \text{ (Stokes Theorem)}$$

$$\int_V dV(\nabla * f) = \int_{\partial V} d\sigma * f \text{ (When } * = \cdot, \text{ it is the famous Gauss Formula.)}$$

$$\int_V dV(A \cdot \nabla)B = \int_{\partial V} A \cdot d\sigma B \text{ (if } \nabla \cdot A = 0)$$

$$\int_S d\sigma \cdot (\nabla \times A) = \int_{\partial S} dl \cdot A \text{ (the traditional Stokes Formula)}$$

$$\int_S (d\sigma \times \nabla) \times A = \int_{\partial S} dl \times A$$

$$\int_S d\sigma \times \nabla f = \int_{\partial S} dl f$$

目录

| | | |
|----------|---------------------------|-----------|
| 1 | 狭义相对论与 Maxwell 方程组 | 1 |
| 1.1 | 狭义相对论基本框架 | 1 |
| 1.2 | 相对论力学 | 3 |
| 1.3 | 电磁场中的电荷 | 6 |
| 1.4 | 电磁场张量 | 7 |
| 1.5 | 场的变换 | 9 |
| 1.6 | Maxwell 方程组 | 10 |
| 1.7 | 能量动量张量 | 14 |
| 1.8 | Noether 定理 | 18 |
| 1.9 | 微分形式语言描述的 Maxwell 方程组 | 20 |
| 2 | Hamilton 形式化 | 25 |
| 2.1 | 正则方程和 Poisson 括号 | 25 |
| 2.2 | Hamilton 形式化的困难：约束系统 | 28 |
| 2.3 | Dirac 括号 | 30 |
| 2.4 | 带电粒子的运动 | 33 |
| 3 | 电磁波 | 35 |
| 3.1 | 波动方程 | 35 |
| 3.2 | 单色波 | 37 |
| 3.3 | 场的本征振动 | 39 |
| 3.4 | 部分偏振波 | 40 |
| 3.5 | 光线 | 40 |
| 4 | 恒定电磁场 | 41 |
| 4.1 | Coulomb 定律 | 41 |
| 4.2 | 匀速运动的电荷产生的场 | 43 |
| 4.3 | Biot-Savart 定律 | 45 |
| 4.4 | 多极矩 | 46 |
| 4.5 | 磁矩 | 47 |
| 4.6 | 外场中的电荷体系 | 48 |

| | | |
|----------|--------------------|-----------|
| 5 | 辐射 | 49 |
| 5.1 | 推迟势 | 49 |
| 5.2 | Lienard-Wiechert 势 | 51 |
| 5.3 | 电荷体系在远处所产生的场 | 52 |
| 6 | 联络与曲率 | 53 |
| 6.1 | 测地线 | 54 |
| 6.2 | 矢量丛及其联络 | 55 |
| 6.3 | 曲率 | 60 |
| 6.4 | Levi-Civita 联络 | 61 |
| 6.5 | Hodge 星算子 | 64 |
| 6.6 | 主丛及其联络 | 67 |
| 6.7 | 联络形式 | 68 |
| 6.8 | 规范场 | 71 |

符号约定

本文所指相对论皆狭义相对论。先指出指标在狭义相对论的运用是自然的。

物理上，我们并不能直接观测到数学上所谓的和（惯性）参考系选取无关的矢量，我们测量到的都是分量，然后判别矢量与否需要通过看这些分量是否在参考系变换下符合相应规则。所以狭义相对论里面记矢量一般为 x^μ ，即其 x 的分量形式，他在不同参考系下的值是不同的。

在 Einstein 记号上，我们约定，如果使用希腊字母，如 μ ，则滚动范围是 0 到 3，如果使用英文字母，如 i ，则滚动范围是 1 到 3。

约定在使用坐标或者矩阵表示的时候，第四个分量才是第零分量，譬如 $(0, 0, 0, 1)$ 是指第零分量为 1。此外，度规 $\eta_{\mu\nu}$ 和 $\eta^{\mu\nu}$ 我们都选取 $\text{diag}(-1, -1, -1, 1)$ ，这俩还用作提升或者下降指标，譬如

$$\eta_{\mu\nu} a^\nu = a_\mu, \quad \eta^{\mu\nu} A_\mu{}^\sigma = A^{\nu\sigma}.$$

直观地来看，提升或降低一下非第零指标都需要变号，提升或降低一个第零指标数值不变。有时候，我们将四维物理量 x^μ 写作 (\mathbf{x}, y) ，这表示， x^μ 的空间部分为 \mathbf{x} ，时间部分为 y 。也有时候我们用数学上不写出分量的形式，避免指标过烦，如 $\eta_{\mu\nu} x^\mu y^\nu = x_\mu y^\mu$ 就写作 $\mathbf{x} \cdot \mathbf{y}$ ，而

$$\mathbf{x} \cdot \mathbf{y} = -x^i y_i = x^1 y^1 + x^2 y^2 + x^3 y^3,$$

其余比如 $|p| = \sqrt{p \cdot p} = \sqrt{p_\mu p^\mu}$ 等也是清楚的。此外，对任意三维矢量 \mathbf{v} 我们用 $\hat{\mathbf{v}}$ 来表示 \mathbf{v} 方向的单位矢量。

需要特别留意的，对一个矢量 u^μ 求偏导数的算符应看作下标，同理，对一个矢量 u_μ 求偏导数的算符应看作上标，特别是对坐标的导数我们记

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\nabla, \frac{1}{c} \frac{\partial}{\partial t} \right), \quad \partial^\mu = \frac{\partial}{\partial x_\mu} = \left(-\nabla, \frac{1}{c} \frac{\partial}{\partial t} \right).$$

对于逆序数符号，我们约定 $\epsilon^{123} = 1$ 和 $\epsilon_{123} = 1$ ，两个并不是升降指标的关系。当然对于四个指标的， $\epsilon^{0123} = 1$ 。

最后的最后，本笔记最主要的参考书是 Landau 的经典场论，甚至说是缩写了一遍 Landau 都不为过。其他的选材，全凭兴趣。

Chapter 1

狭义相对论与 Maxwell 方程组

1.1 狭义相对论基本框架

狭义相对论的背景是 Minkowski 时空，首先他是一个 $\mathcal{M} = \mathbb{R}^3 \times \mathbb{R}$ ，其次，他上面有内积 $\langle \star, \star \rangle : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ ，满足

$$\langle x, y \rangle = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3,$$

采用 Einstein 记号，我们通常将其写作 $x_\mu y^\mu = \eta_{\mu\nu} x^\mu y^\nu$ 。

在众多的参考系中，有一类参考系，在这类参考系中观察一个自由粒子，他一定做匀速直线运动。这样的参考系称为惯性参考系。如果我们改用不同的惯性参考系来观测同一个物理量，他在我们看到的值上面诱导一个变换，我们称之为 Lorentz 变换，即

$$T : \mathcal{M} \rightarrow \mathcal{M}, \quad T : u \mapsto T(u).$$

所谓矢量，就是在参考系变换下，值如上式变换的量。

考虑到惯性参考系的性质，对于自由粒子，需要速度恒定¹ $x = (vt, ct)$ ，在变换后 $T(x) = x'$ 应该也有 $x' = (v't', ct')$ ，两边对 t 求导就有

$$T'(x)(v, c) = (v', c) \frac{dt'}{dt}, \quad (1.1)$$

这限制了变换 T 的形式。

一个粒子运动的整个时空信息表现为一条 \mathcal{M} 上的曲线，我们称之为一条世界线，世界线上的点称为粒子的世界点，或者其四维坐标，物理上看来就是四维矢量 $x^\mu = (\mathbf{x}, ct)$ ，其中 \mathbf{x} 就是熟悉的三维坐标， c 是光速。每一个世界点对应一个事件，那么两个世界点之间的距离 $|x - y|$ 就称为两个时间的间隔，这个量的物理意义是很基本的，如果从 x 到 y 是一个光速传播的粒子的两个世界点，由于光速不变原理，在任意一个参考系中都成立

$$(x - y) \cdot (x - y) = c^2(t_x - t_y)^2 - (\mathbf{x} - \mathbf{y})^2 = (t_x - t_y)^2(c^2 - c^2) = 0,$$

接着考察世界线的无穷小线元

$$ds = \sqrt{dx^\mu dx_\mu} = \sqrt{c^2 dt^2 - d\mathbf{x}^2} = c dt \sqrt{1 - \frac{v^2}{c^2}},$$

¹这就是说，惯性参考系之间的变换在几何上就是将直线变成直线。

若在一个参考系中 $ds = 0$ ，则在其他参考系中 $ds' = 0$ ，此外， ds 和 ds' 必须是同阶的无穷小，因此两者成正比，写作 $ds^2 = a ds'^2$ 。按照均匀性的假设，平移不变导致 a 只和两个参考系的相对坐标有关，而空间各向同性假设表明 a 只能相对坐标大小有关，由于 a 是关于无穷小间隔的系数，所以我们可以等价地说他只和相对速率有关。

现在考虑三个惯性系 K, K_1, K_2 ，分别有

$$ds^2 = a(v_1)ds_1^2, \quad ds^2 = a(v_2)ds_2^2,$$

但是 K_1 相对 K_2 我们也可以写出

$$ds_1^2 = a(v_{12})ds_2^2,$$

其中 v_{12} 是 K_1 相对 K_2 的速率，是两个速度矢量 $\mathbf{v}_1, \mathbf{v}_2$ 的函数，则

$$a(v_{12}) = \frac{a(v_2)}{a(v_1)},$$

左边依赖于 $\mathbf{v}_1, \mathbf{v}_2$ 之间的夹角，而右边并不，此时 a 只能是一个常函数，所以 $a = a/a = 1$ 。综合上面讨论的，我们得到了 $ds^2 = ds'^2$ ，即 ds^2 在参考系变换下是一个不变量。那么对于有限间隔，从无穷小间隔积分就可以发现他是一个不变量。

由于参考系诱导了一个 Lorentz 变换，那么这个应该也要保持有限间隔 $s = |x - y|$ 是一个不变量。假如现在有一个 Lorentz 变换 T ，则必有

$$|T(x) - T(y)|^2 = |x - y|^2, \quad (1.2)$$

则应该对任意的 x 和 y 都应该成立。现在令 $y \rightarrow x$ ，则

$$c^2 dt'^2 \left(1 - \frac{\mathbf{v}'^2}{c^2}\right) = ds'^2 = ds^2 = c^2 dt^2 \left(1 - \frac{\mathbf{v}^2}{c^2}\right),$$

则²

$$\frac{dt'}{dt} = \sqrt{\frac{c^2 - \mathbf{v}^2}{c^2 - \mathbf{v}'^2}},$$

现在，将其代入(1.1)，则

$$T'(x)(\mathbf{v}, c) = (\mathbf{v}', c) \sqrt{\frac{c^2 - \mathbf{v}^2}{c^2 - \mathbf{v}'^2}}$$

对任意的 x 都应该成立，此时只有 $T'(x)$ 是一个常矩阵才能满足，因此他 T 是一个仿射变换。仿射变换由一个平移和一个线性变换构成，不妨将其记作 $T(\Lambda, a)$ ，则

$$T(\Lambda, a)u = \Lambda u + a,$$

其中 $(\Lambda u)^\mu = \Lambda^\mu_\nu u^\nu$ 。

将任意的仿射变换代入(1.2)中，我们 $|\Lambda(x - y)|^2 = |x - y|^2$ ，由于 x 和 y 任意，所以也可以写作 $(\Lambda u) \cdot (\Lambda u) = u \cdot u$ ，写成指标形式

$$\eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = \eta_{\rho\sigma}.$$

因此不含平移的 Lorentz 变换构成一种正交群，我们称为 Lorentz 群。可以直接验证 $(\det \Lambda)^2 = 1$ ，这只要对上面求行列式就可以了。此外可以直接计算得 $(\Lambda^{-1})^\rho_\nu = \Lambda_\nu^\rho$ 。

²开根不妨取正，物理上就是说两个时间方向相同。

上面确定了不含平移的 Lorentz 变换需要满足的形式, 反之, 对于任意两个矢量的内积 $x \cdot y$, 我们可以检验他是不变量。所以在不含平移的 Lorentz 变换下不变的量称为标量, 也有时候我们会把在所有 Lorentz 变换下不变的量称为标量。值得注意的是, 体积元 $d^4x = dx^0 dx^1 dx^2 dx^3$ 也是一个不变量, 因为

$$dx'^0 dx'^1 dx'^2 dx'^3 = |\det(\Lambda)| dx^0 dx^1 dx^2 dx^3 = dx^0 dx^1 dx^2 dx^3,$$

其中 $|\det(\Lambda)| = 1$ 上面已经求过了。

注意 Λ 作为坐标变换和参考系变换的意义是不同的, 这两者不是对应关系, 坐标变换是作用在矢量的分量上面的, 或者说反变矢量上面的, 而参考系变换是作用在矢量的基上面的, 或者说协变矢量上面的。所以当参考系变换为 Λ^{-1} 的时候, 坐标变换为 Λ . 或者反过来, 当参考系变换为 Λ 的时候, 坐标变换为 Λ^{-1} , 在这个意义上, 矢量 p^μ 在参考系经过两次变换 $\Lambda_1 \Lambda_2$ 后就是

$$(\Lambda_1^{-1})^\mu{}_\rho (\Lambda_2^{-1})^\rho{}_\nu p^\nu = ((\Lambda_2 \Lambda_1)^{-1})^\mu{}_\nu p^\nu,$$

这就是所谓的反变矢量。指标在下面的就是协变矢量, 对应的是基, 所以一般所说的矢量就是反变矢量。

一个尤其重要的 Lorentz 变换如下: 假设我们在 O 系看到了一个物体沿着 x^3 方向以速度 v 匀速运动, 则 O 系看到的四维矢量 u^μ 和固定在物体上的 \bar{O} 系中看到的矢量 \bar{u}^μ 有如下关系:

$$\begin{pmatrix} u^1 \\ u^2 \\ u^3 \\ u^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma & \beta/\gamma \\ 0 & 0 & \beta/\gamma & 1/\gamma \end{pmatrix} \begin{pmatrix} \bar{u}^1 \\ \bar{u}^2 \\ \bar{u}^3 \\ \bar{u}^0 \end{pmatrix},$$

其中 $\beta = v/c$, $\gamma = \sqrt{1 - \beta^2}$. 显然, 该矩阵的行列式为 1. 更有趣的是, $v \rightarrow -v$ 即完成了逆变换, 物理上来看, O 系看 O' 系的速度是 v , 则 O' 系中看 O 系的速度为 $-v$.

下面我们要转向动力学, 首先, 如果按照通常方法定义速度 $v^\mu = dx^\mu/dt$, 他不是四维矢量, 因为 dt 不是标量, 取而代之的, 我们选取标量 ds 来定义 $u^\mu = dx^\mu/ds$, 这是一个四维矢量, 称为四维速度。此时

$$u^\mu = \frac{dx^\mu}{ds} = \frac{1}{c} \left(1 - \frac{\mathbf{v}^2}{c^2} \right)^{-1/2} \frac{dx^\mu}{dt},$$

所以

$$u^\mu = \left(\frac{\mathbf{v}}{c} \left(1 - \frac{\mathbf{v}^2}{c^2} \right)^{-1/2}, \left(1 - \frac{\mathbf{v}^2}{c^2} \right)^{-1/2} \right)$$

因此,

$$u \cdot u = c^2 (u^0)^2 - \mathbf{u}^2 = \left(1 - \frac{\mathbf{v}^2}{c^2} \right)^{-1} - \left(1 - \frac{\mathbf{v}^2}{c^2} \right)^{-1} \frac{\mathbf{v}^2}{c^2} = 1. \quad (1.3)$$

可以看到, 四维速度其实四个分量不是相互独立的。几何上, 这代表着 u^μ 是与世界线相切的一个四维单位矢量。此外, 对他求导就可以发现

$$u_\mu \frac{du^\mu}{ds} = 0.$$

1.2 相对论力学

作用量原理依然可以用来抽象相对论力学。对于一个自由粒子, 他的作用量需要当他匀速运动 (即世界线是直线) 的时候取极值, 而世界线的线长正是如此的一个泛函, 所以不妨猜测自由粒子的世界线具有如下形式

$$S = -\alpha \int_a^b ds = - \int_{t_a}^{t_b} dt \alpha c \sqrt{1 - \frac{\mathbf{v}^2}{c^2}},$$

所以 Lagrangian 形式³为

$$L = -\alpha c \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} dt,$$

当 \mathbf{v}^2 远比 c^2 小的时候, 对其展开有

$$L \approx -\alpha c dt + \frac{\alpha \mathbf{v}^2}{2c} dt.$$

除去恰当形式项 $-\alpha c dt$, 第二项应该就是经典力学的动能项 $m\mathbf{v}^2 dt/2$, 所以 $\alpha = mc$, 因此自由粒子的 Lagrangian 形式为

$$L = -mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} dt,$$

作用量为

$$S = -mc \int_a^b ds.$$

先固定端点对 S 求变分, 并按照作用量原理, 易得

$$0 = \delta S = -mc \int_a^b \frac{dx_\mu}{ds} d(\delta x^\mu) = mc \int_a^b \delta x^\mu \frac{du_\mu}{ds} ds,$$

于是, 自由粒子运动有 $du_\mu/ds = 0$, 即四维速度恒定。

反之, 为了确定动力学量, 假设粒子运动在真实轨道上, 而放松一个端点, 此时

$$\delta S = -mc \int^s \frac{dx_\mu}{ds} d(\delta x^\mu) = -mc u_\mu \delta x^\mu$$

所以 $\partial_\mu S = -mc u_\mu$, 在经典力学中已经知道, 作用量关于空间坐标的偏导数就是动量, 于是

$$p^i = \eta^{i\mu} \partial_\mu S = mc u^i,$$

作用量关于时间的偏导数的相反数就是能量,

$$E = -c \partial_0 S = mc^2 u_0 = mc^2 u^0,$$

综合起来即

$$p^\mu = (p^1, p^2, p^3, E/c) = (\mathbf{p}, E/c) = mc u^\mu = -\partial^\mu S,$$

这就是四维动量。他正比于四维速度, 正好类似于空间动量正比于空间速度。利用(1.3), 可以得到相对论下的动量-能量关系:

$$p^\mu p_\mu = m^2 c^2 u^\mu u_\mu = m^2 c^2,$$

具体写成三维形式即

$$\frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2.$$

利用动量能量关系就可以得到

$$\partial^\mu S \partial_\mu S = m^2 c^2, \quad (1.4)$$

这就是相对论下的 Hamilton-Jacobi 方程。

回到熟悉的三维形式, 我们这里求一下对应着 \mathbf{v} 的广义动量

$$\frac{\partial L_t}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}} = m\mathbf{c}\mathbf{u} = \mathbf{p},$$

³强调他是一个 1-形式, 主要是因为这是和选取坐标无关的, 将来有可能换成其他的变量, 比如 x^μ 和 u^μ . 如果需要我们常说的 Lagrangian, 使用 L_t 表示, 此时 $L = L_t dt$.

可以看到此处的广义动量即三维动量，因此刚刚对于四维动量的定义是比较合适的。

现在动力学位量只剩下角动量了，假设坐标有一个无穷小变换

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + (\delta\theta)^\mu{}_\nu$$

利用 Λ 的正交性（保内积），容易计算得 $(\delta\theta)_{\mu\nu} = -(\delta\theta)_{\nu\mu}$ 。

考虑作用量的改变

$$\delta S|_a^b = -p_\mu \delta x^\mu|_a^b = -p_\mu (\delta\theta)^\mu{}_\nu x^\nu|_a^b = -(\delta\theta)_{\mu\nu} p^\mu x^\nu|_a^b,$$

由于 $\omega_{\mu\nu}$ 反对称，于是求和中 $p^\mu x^\nu$ 的对称部分求和为 0，那么

$$\delta S|_a^b = -\frac{1}{2}(\delta\theta)_{\mu\nu}(p^\mu x^\nu - x^\mu p^\nu)|_a^b,$$

因此 $\delta S|_a^b = 0$ 就是说

$$(p^\mu x^\nu - x^\mu p^\nu)_a = (p^\mu x^\nu - x^\mu p^\nu)_b$$

是一个守恒量，定义四维角动量张量为

$$M^{\mu\nu} = x^\mu p^\nu - p^\mu x^\nu.$$

他是一个反对称张量。对应到三维角动量 $\mathbf{M} = \mathbf{r} \times \mathbf{p}$ 有

$$M^{23} = M_x, \quad M^{13} = -M_y, \quad M^{12} = M_z.$$

现在，如果希望用 $L = -mc ds$ 用来求运动方程和广义动量，可以发现 $L_s = -mc$ 是一个常数，运动方程和广义动量似乎都是 0，这是怎么回事呢？这是因为四维速度的四个分量不是独立的，它满足 $u^\mu u_\mu = 1$ 。所以，为了看起来独立，我们需要在 L_s 加上如下乘子，构造新的 Lagrangian

$$L_s^* = L_s + \lambda(x, u, s)(u^\mu u_\mu - 1),$$

为使下面的讨论对一般的 Lagrangian 都成立，设

$$L_s = L_{0,s} + L_{1,s} = -mc + L_{1,s},$$

此时的 Lagrange 方程写作

$$\frac{\partial L_s^*}{\partial x^\mu} - \frac{d}{ds} \frac{\partial L_s^*}{\partial u^\mu} = \frac{\partial L_s}{\partial x^\mu} - \frac{d}{ds} \frac{\partial L_s}{\partial u^\mu} + \frac{\partial}{\partial x^\mu} [\lambda(u^\mu u_\mu - 1)] - \frac{d}{ds} \frac{\partial}{\partial u^\mu} [\lambda(u^\mu u_\mu - 1)] = 0. \quad (1.5)$$

代入约束和 L 的具体形式

$$\frac{\partial L_{1,s}}{\partial x^\mu} - \frac{d}{ds} \frac{\partial L_{1,s}}{\partial u^\mu} = \frac{d}{ds} (2\lambda u_\mu) = 2u_\mu \frac{d\lambda}{ds} + 2\lambda \frac{du_\mu}{ds}$$

左边不妨记作 l_μ ，对于自由粒子， $l_\mu = 0$ ，那么显然 $2\lambda u_\mu$ 为常数，乘一个 u^μ ，就可以发现 λ 是常数，因此 u_μ 是常数，这个结论前面我们已经知道了。

如果 l_μ 任意，右边是两个正交矢量 u_μ 和 du^μ/ds 的叠加，提取 $d\lambda/ds$ 的系数，那么

$$l_\mu u^\mu = 2u_\mu u^\mu \frac{d\lambda}{ds} + 2\lambda \frac{du_\mu}{ds} u^\mu = 2 \frac{d\lambda}{ds},$$

如果我们有 $l_\mu u^\mu = 0$ ，这个时候， $d\lambda/ds = 0$ ，这意味着 λ 是一个常数。不妨选取 $\lambda = -mc/2$ ，则

$$L = -mc ds - \frac{mc}{2} (u^\mu u_\mu - 1) ds + L_1 = -\frac{mc}{2} ds - \frac{mc}{2} u^\mu u_\mu ds + L_1,$$

或者等价地去掉一个恰当形式项，把

$$L = -\frac{mc}{2}u^\mu u_\mu ds + L_1 \quad (1.6)$$

看作 u^μ 自由时候的 Lagrangian 形式。广义动量即

$$-P_\mu = \frac{\partial L_s}{\partial u^\mu} = -\frac{mc}{2} \frac{\partial(u^\nu u_\nu)}{\partial u^\mu} + \frac{\partial L_{1,s}}{\partial u^\mu} = -p_\mu + \frac{\partial L_{1,s}}{\partial u^\mu},$$

在自由粒子的情况下，广义动量即 $-p_\mu$ ，符合上面的定义，尤其是空间部分，可以看到就是 \mathbf{p} ，因此 $\lambda = -mc/2$ 的选取是合理的。

同时，写运动方程也没什么好顾虑的，直接对 L 写出

$$0 = \partial_\mu L_{1,s} + \frac{dP_\mu}{ds} = \partial_\mu L_{1,s} + \frac{dp_\mu}{ds} - \frac{d}{ds} \frac{\partial L_{1,s}}{\partial u_\mu} = l_\mu + \frac{dp_\mu}{ds},$$

即

$$l_\mu = -\frac{dp_\mu}{ds}, \quad (1.7)$$

或者用上面的

$$l_\mu = 2u_\mu \frac{d\lambda}{ds} + 2\lambda \frac{du_\mu}{ds} = -\frac{dp_\mu}{ds}.$$

1.3 电磁场中的电荷

相对论中不存在刚体的概念，所以在相对论中不能认为粒子有有限的尺寸而只能认为一个粒子是一个几何点。

既然引入了电荷，那么在作用量中又要引入相关于电荷的一项，一般来说，添加的作用量的形式⁴为

$$S_{mf} = -e \int_a^b A_\mu dx^\mu,$$

这是物质 (matter) 和场 (field) 作用的作用量。其中 A_μ 是场，目前只认为坐标的函数，而 e 是电荷量，不一定指电子电荷量。值得注意的是， A_μ 不是唯一确定的，比如，我们做如下变换 $A_\mu \rightarrow A_\mu + \partial_\mu f$ ，其中 f 是一个光滑实函数，那么作用量积分就变成

$$S_{mf} \rightarrow S_{mf} - e \int_a^b \partial_\mu f dx^\mu = S_{mf} - ef|_a^b,$$

可是作用量多了一个常数 $-ef|_a^b$ 对运动方程是不会有半点影响的。所以，只有在 $A_\mu \rightarrow A_\mu + \partial_\mu f$ 下不变的量，才有实际的物理意义，特别是，所有方程在 $A_\mu \rightarrow A_\mu + \partial_\mu f$ 应该不变，这种不变性称之为规范不变性。

回到作用量，结合自由粒子的作用量和粒子受场作用的作用量，总作用量写作

$$S = -mc \int_a^b ds - e \int_a^b A_\mu dx^\mu,$$

⁴尽管只在形式上，作用量中似乎还可以加上一个

$$\int ds B(x, u),$$

其中 B 是标量函数。取文章中这种形式的原因，一方面物理上被认为是实验结果，另一方面是数学上的，后面将看到，如果没有 B 这项就可以满足 $l_\mu u^\mu = 0$ ，此时 Lagrangian 可以形式地看作一个自由粒子 Lagrangian 和一个势的和，否则，Lagrangian 中的乘子，将看作变成自由粒子 Lagrangian 和势的一个耦合。方程变复杂，美感上也说不过去。

如果记 $A^\mu = (\mathbf{A}, \varphi/c)$, 那么

$$S = \int_{t_a}^{t_b} dt \left(-mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} + e\mathbf{A} \cdot \mathbf{v} - e\varphi \right),$$

此时 Lagrangian 写作

$$L_t = -mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} + e\mathbf{A} \cdot \mathbf{v} - e\varphi,$$

对应着 \mathbf{v} 的广义动量为

$$\mathbf{P} = \frac{\partial L_t}{\partial \mathbf{v}} = \mathbf{p} + e\mathbf{A},$$

应用 Lagrange 方程, 就有运动方程如下:

$$\dot{\mathbf{P}} = \frac{d}{dt} \frac{\partial L_t}{\partial \mathbf{v}} = \frac{\partial L_t}{\partial \mathbf{r}} = \nabla L_t = e\nabla(\mathbf{A} \cdot \mathbf{v} - \varphi) = e\nabla(\mathbf{A} \cdot \mathbf{v}) - e\nabla\varphi.$$

我们在幼儿园⁵已经熟知以下矢量公式

$$\nabla(\mathbf{A} \cdot \mathbf{v}) = (\mathbf{A} \cdot \nabla)\mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{A} + \mathbf{A} \times (\nabla \times \mathbf{v}) + \mathbf{v} \times (\nabla \times \mathbf{A}),$$

由之可以计算得

$$\frac{d}{dt}(\mathbf{p} + e\mathbf{A}) = e(\mathbf{v} \cdot \nabla)\mathbf{A} + e\mathbf{v} \times (\nabla \times \mathbf{A}) - e\nabla\varphi,$$

但是 \mathbf{A} 对时间的全微分可以分成本身随时间变化的一部分和位置矢量随时间变化的一部分:

$$\dot{\mathbf{A}} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A},$$

所以结合上面一大堆公式, 最后可以得到

$$\frac{d\mathbf{p}}{dt} = -e \frac{\partial \mathbf{A}}{\partial t} + e\mathbf{v} \times (\nabla \times \mathbf{A}) - e\nabla\varphi,$$

这就是著名的 Lorentz 力公式。

定义

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\varphi, \quad \mathbf{B} = \nabla \times \mathbf{A},$$

前者称为电场强度, 后者称为磁感应强度。在此定义下, Lorentz 力公式可以改成我们熟悉的形式

$$\dot{\mathbf{p}} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (1.8)$$

1.4 电磁场张量

这节就是用指标的语言重新写一遍上一节的内容, 指标的优点显然是将任意惯性参考系之间的转变都变得相当清晰。而且, 我们将会看到, 在这种情况下, 自由粒子部分的 Lagrangian 依然可以写作我们希望的那种形式(1.6)。先写出作用量

$$S = S_m + S_{mf} = \int_a^b L(x^\mu, u^\mu) = \int_a^b (L_m(x^\mu, u^\mu) + L_{mf}(x^\mu, u^\mu))$$

其中

$$L_{mf}(x^\mu, u^\mu) = -eA_\mu dx^\mu = -eA_\mu u^\mu ds$$

⁵Weinberg 的著名梗。

所以

$$l_\mu = \frac{\partial L_{\text{mf},s}}{\partial x^\mu} - \frac{d}{ds} \frac{\partial L_{\text{mf},s}}{\partial u^\mu} = -e \left(u^\nu \partial_\mu A_\nu - \frac{dA_\mu}{ds} \right) = -e u^\nu (\partial_\mu A_\nu - \partial_\nu A_\mu), \quad (1.9)$$

我们可以检验 $l_\mu u^\mu = 0$ ，这个只要直接计算 $u^\nu u^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu)$ 为零即可，而因为这是一个对称张量和反对称张量求和。当 $l_\mu u^\mu = 0$ ，在相对论力学里面讨论过，运动方程写作(1.7)，这种情况下即

$$e u^\nu (\partial_\mu A_\nu - \partial_\nu A_\mu) = \frac{dp_\mu}{ds}.$$

引入符号

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

其中 $F_{\mu\nu}$ 就是这节的主角，电磁场张量。这时候运动方程写作

$$\frac{dp_\mu}{ds} = e F_{\mu\nu} u^\nu.$$

这就是四维形式的电荷运动方程，注意到 $p^\mu = m c u^\mu$ ，电荷运动方程可以改写成下面两种形式

$$m c \frac{dp_\mu}{ds} = e F_{\mu\nu} p^\nu, \quad m c \frac{du_\mu}{ds} = e F_{\mu\nu} u^\nu.$$

将 $A^\mu = (\mathbf{A}, \varphi/c)$ 或者 $A_\mu = (-\mathbf{A}, \varphi/c)$ 代入 $F_{\mu\nu}$ 的表达式中，可以写出矩阵

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -B_z & B_y & -E_x/c \\ B_z & 0 & -B_x & -E_y/c \\ -B_y & B_x & 0 & -E_z/c \\ E_x/c & E_y/c & E_z/c & 0 \end{pmatrix},$$

或者简记为 $F_{\mu\nu} = (\mathbf{B}, \mathbf{E}/c)$ ，容易证明 $F^{\mu\nu} = (\mathbf{B}, -\mathbf{E}/c)$ 。

观测两个运动方程，可以发现，我们通常所说的 Lorentz 力方程(1.8)只有三个分量，而四维形式的电荷运动方程实际上有四个分量，所以似乎就多出来一个方程：

$$m c \frac{dp_0}{ds} = e F_{0\nu} p^\nu = \frac{e}{c} \mathbf{E} \cdot \mathbf{p},$$

把他换到三维形式即

$$\frac{dE}{dt} = e \mathbf{E} \cdot \mathbf{v}, \quad (1.10)$$

这就是粒子动能的变化，其中没有出现 \mathbf{B} 说明磁场并不改变其中带电粒子的动能。既然(1.10)阐述了内能随时间的变化，这个多出来的方程似乎很重要。

实则不然，这个方程并不独立于其他三个方程，由于 $u^\mu du_\mu/ds = 0$ ，所以将四维形式的电荷运动方程和 u^μ 做内积就有

$$0 = m c u^\mu \frac{du_\mu}{ds} = e F_{\mu\nu} u^\mu u^\nu,$$

右边为零是一个反对称张量和一个对称张量求和的自然结果。

这里来求四维广义动量，

$$P_\mu = -\partial_\mu L_s = p_\mu - \frac{\partial L_{\text{mf},s}}{\partial u^\mu} = p_\mu + e A_\mu,$$

也就是 $P^\mu = p^\mu + e A^\mu = (\mathbf{p} + e \mathbf{A}, (E + e\varphi)/c)$ 。

1.5 场的变换

这节聚焦在场在 Lorentz 变换下的表现，尤其是在那个 O 系观测沿 $z = x^3$ 轴以速度 v 运动的 O' 的变换，我们称为 z -变换，首先矢势是一个矢量，则变换规则为

$$A^\mu \rightarrow \Lambda^\mu{}_\nu A^\nu,$$

在 z -变换下，

$$\begin{pmatrix} A^1 \\ A^2 \\ A^3 \\ A^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma & \beta/\gamma \\ 0 & 0 & \beta/\gamma & 1/\gamma \end{pmatrix} \begin{pmatrix} A'^1 \\ A'^2 \\ A'^3 \\ A'^0 \end{pmatrix},$$

其中 $\beta = v/c$, $\gamma = \sqrt{1 - \beta^2}$.

对于张量的变化，

$$F^{\mu\nu} \rightarrow \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma},$$

为了方便计算，我们将其改成

$$F^\mu{}_\nu \rightarrow \Lambda^\mu{}_\rho F^\rho{}_\sigma \Lambda^\sigma{}_\nu = \Lambda^\mu{}_\rho F^\rho{}_\sigma (\Lambda^{-1})^\sigma{}_\nu,$$

这就是三个矩阵复合，如果是 z -变换，则

$$\begin{aligned} & \Lambda^\mu{}_\rho F^\rho{}_\sigma (\Lambda^{-1})^\sigma{}_\nu \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma & \beta/\gamma \\ 0 & 0 & \beta/\gamma & 1/\gamma \end{pmatrix} \begin{pmatrix} 0 & B_z & -B_y & E_x/c \\ -B_z & 0 & B_x & E_y/c \\ B_y & -B_x & 0 & E_z/c \\ E_x/c & E_y/c & E_z/c & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma & -\beta/\gamma \\ 0 & 0 & -\beta/\gamma & 1/\gamma \end{pmatrix} \\ &= \begin{pmatrix} 0 & B_z & -B_y/\gamma - E_x\beta/c\gamma & B_y\beta/\gamma + E_x/c\gamma \\ -B_z & 0 & B_x/\gamma - E_y\beta/c\gamma & -B_x\beta/\gamma + E_y/c\gamma \\ B_y/\gamma + E_x\beta/c\gamma & -B_x/\gamma + E_y\beta/c\gamma & 0 & E_z/c \\ \beta B_y/\gamma + E_x/c\gamma & -B_x\beta/\gamma + E_y/c\gamma & E_z/c & 0 \end{pmatrix}, \end{aligned}$$

所以

$$F_{\mu\nu} \rightarrow \begin{pmatrix} 0 & -B_z & B_y/\gamma + E_x\beta/c\gamma & -B_y\beta/\gamma - E_x/c\gamma \\ B_z & 0 & -B_x/\gamma + E_y\beta/c\gamma & B_x\beta/\gamma - E_y/c\gamma \\ -B_y/\gamma - E_x\beta/c\gamma & B_x/\gamma - E_y\beta/c\gamma & 0 & -E_z/c \\ \beta B_y/\gamma + E_x/c\gamma & -B_x\beta/\gamma + E_y/c\gamma & E_z/c & 0 \end{pmatrix},$$

那么从 $(\mathbf{B}', \mathbf{E}')$ 到 (\mathbf{B}, \mathbf{E}) 的变换公式也已经给出了：

$$\mathbf{B} = \left(\frac{B'_x - E'_y\beta/c}{\sqrt{1 - \beta^2}}, \frac{B'_y + E'_x\beta/c}{\sqrt{1 - \beta^2}}, B'_z \right),$$

以及

$$\mathbf{E} = \left(\frac{E'_x + \beta c B'_y}{\sqrt{1 - \beta^2}}, \frac{E'_y - \beta c B'_x}{\sqrt{1 - \beta^2}}, E'_z \right).$$

可以看到,即使是足够简单的 z -变换,我们也计算得如此辛苦,所以通常我们会去寻找标量,即在任何参考系下都相同的量。显然,从四维指标形式来看,

$$F_{\mu\nu}F^{\mu\nu}, \quad \epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$$

将是两个标量,对应到三维,

$$\mathbf{E}^2 - c^2\mathbf{B}^2, \quad \mathbf{E} \cdot \mathbf{B}$$

是两个标量。

有了这两个不变量,对于匀强电磁场,可以做这样的断言:如果在某个参考系, $\mathbf{E} \cdot \mathbf{B} = 0$, 那么在任意的惯性参考系,我们都将发现 \mathbf{E} 和 \mathbf{B} 垂直,特别地,可以找到一个参考系,在其中只有纯电场或者纯磁场,由于 $\mathbf{E}^2 - c^2\mathbf{B}^2$ 是不变量,所以当 $\mathbf{E}^2 - c^2\mathbf{B}^2 > 0$ 的时候,我们能发现纯电场,反之,能发现纯磁场。

1.6 Maxwell 方程组

回到定义,

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\varphi, \quad \mathbf{B} = \nabla \times \mathbf{A},$$

对 \mathbf{E} 求旋度,对 \mathbf{B} 求散度就得到了第一对 Maxwell 方程:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0.$$

这两个方程写成四维矢量形式即

$$\partial_{[\mu}F_{\nu\rho]} = \partial_{\mu}F_{\nu\rho} - \partial_{\nu}F_{\mu\rho} + \partial_{\rho}F_{\mu\nu} = 0.$$

狭义相对论的几个假设的改变是很基本的,在经典力学中,所谓的场的作用是瞬时的,可是到了相对论下自然是不对的,粒子激发了场,场再作用粒子,而场的传播速度是不可能超过光速的。因此,场不再只是一个数学模型,而可能具有物理实体的意义。下面我们就来谈论场的作用量,最后写出物质、场以及物质和场作用的作用量之和,并给出完整的场方程,即 Maxwell 方程组。

我们希望得到 S_f 是场本身的作用量,即场在没有电荷时候的作用量。由于没有电荷,这一项并不对运动在场内的粒子有影响,反之,若是想要寻找确定场的方程,这一项显然是必须的。为了求出作用量的形式,我们参考实验事实,场是可以叠加的,所以场方程必然是线性方程,所以 S_f 这项肯定是场的二次式,但是势又不能包含在 S_f 内,因为势没有被确定下来。因此 S_f 应当是 $F_{\mu\nu}$ 的函数的积分。作用量是标量,所以可能是一个标量的积分。综合上面所有的条件,我们有理由相信,作用量 S_f 只能正比于 $F_{\mu\nu}F^{\mu\nu}$ 的积分。所以,

$$S_f = -a \int d^4x F_{\mu\nu}F^{\mu\nu},$$

系数 a 决定单位,以后我们再说,那么总的作用量就可以写作

$$S = S_m + S_f + S_{mf} = -mc \int ds - \int e dx^\mu A_\mu - a \int d^4x F_{\mu\nu}F^{\mu\nu},$$

这个作用量在规范变化下,对运动方程没有影响。更一般地,我们经常考察多粒子系统,这种时候

$$S = S_m + S_f + S_{mf} = - \sum mc \int ds - \sum e \int dx^\mu A_\mu - a \int d^4x F_{\mu\nu}F^{\mu\nu}.$$

为了数学上面的方便，我们不把电荷当点看，而经常是看作连续分布的，这种时候需要谈及所谓的电荷密度 ρ ，电荷密度 ρ 关于体积的积分即这块体积内的电荷量，

$$e = \int \rho dV.$$

对于离散的点电荷，密度取成 δ 函数形，即

$$\rho = \sum_a e_a \delta^3(\mathbf{r} - \mathbf{r}_a).$$

电荷是一个标量，但是 ρ 却不一定是标量，为了电荷是一个标量，我们只需要 ρdV 是个标量就可以了。在 ρdV 两边乘以 dx^μ ，那么

$$\rho dV dx^\mu = \rho dV dx^0 \frac{dx^\mu}{dx^0} = \rho \frac{dx^\mu}{dx^0} d^4x,$$

左边是一个四维矢量，右边也是，而 d^4x 是标量，因此

$$j^\mu = c\rho \frac{dx^\mu}{dx^0}$$

是一个四维矢量，称为四维电流密度矢量。其空间部分有 $\mathbf{j} = \rho \mathbf{v}$ ，而时间部分为 $c\rho$ ，因此 $j^\mu = (\mathbf{j}, c\rho) = (\rho \mathbf{v}, c\rho)$ 。

如果对电荷密度全空间积分，那么就得到了全空间内的电荷量

$$\int_V \rho dV = \frac{1}{c} \int j^0 dS_0 = \frac{1}{c} \int j^\mu dS_\mu,$$

其中积分在整个与 x^0 轴垂直的四维空间的超平面上，其中 dS_μ 是四维空间中的三维面元，类似于三维空间中的二维面元， $dV = dS_0$ 是其中的一个元素，方向与 x^0 轴相同，因此在被积分曲面上， $j^0 dS_0 = j^\mu dS_\mu$ ，故而最后一个等式成立。

有了电流密度矢量，可以改写 S_{mf} 如下

$$S_{mf} = - \int dV \rho \int dx^\mu A_\mu = - \int d^4x \rho A_\mu \frac{dx^\mu}{dx^0} = - \frac{1}{c} \int d^4x A_\mu j^\mu.$$

最后，作用量为

$$S = S_m + S_f + S_{mf} = - \sum mc \int ds - \frac{1}{c} \int d^4x A_\mu j^\mu - a \int d^4x F_{\mu\nu} F^{\mu\nu}.$$

考察一个体积内电荷对时间的变化，即考察 $\partial_t \int \rho dV$ 。另一方面，单位时间内电荷的变化等于单位时间内离开这个体积而走到外面去的电荷量，或者反过来进来的电荷量，在单位时间内经过包围该体积的曲面的面元 $d\mathbf{f}$ 的电荷等于 $\rho \mathbf{v} \cdot d\boldsymbol{\sigma}$ ，此处的 \mathbf{v} 是该处面元所在的空间点的速度。 $d\boldsymbol{\sigma}$ 方向是外法向，假若电荷离开，则 $\rho \mathbf{v} \cdot d\boldsymbol{\sigma} > 0$ ，假若进来，则为负。因此，单位时间离开体积的总电荷就是 $\rho \mathbf{v} \cdot d\boldsymbol{\sigma}$ 对曲面的积分，所以

$$\partial_t \int \rho dV = - \int_{\partial V} \rho \mathbf{v} \cdot d\boldsymbol{\sigma} = - \int_{\partial V} \mathbf{j} \cdot d\boldsymbol{\sigma} = - \int_V dV \nabla \cdot \mathbf{j},$$

所以

$$\partial_t \rho + \nabla \cdot \mathbf{j} = \partial_0(c\rho) + \nabla \cdot \mathbf{j} = \partial_\mu j^\mu = 0,$$

这就是电荷连续性方程。我们已经知道，沿着某一个超平面，总电荷可以表达成

$$\frac{1}{c} \int j^\mu dS_\mu,$$

那么对任意的, 包含整个三维空间的超曲面, 他和这个超平面构成一个闭合曲面 S , 沿着这个闭合曲面积分, 我们按照电荷连续性方程和 Stokes 公式

$$\frac{1}{c} \int_{\partial\Omega} j^\mu dS_\mu = \frac{1}{c} \int_{\Omega} \partial_\mu j^\mu d^4x = 0,$$

所以 j^μ/c 在两个超曲面上积分都是全空间的总电荷。

现在考察电荷守恒和规范不变性的联系, 如果做变换 $A_\mu \rightarrow A_\mu + \partial_\mu f$, 那么 S_{mf} 满足

$$S_{\text{mf}} \rightarrow S_{\text{mf}} - \frac{1}{c} \int d^4x \partial_\mu f j^\mu = S_{\text{mf}} + \frac{1}{c} \int d^4x f \partial_\mu j^\mu,$$

按照电荷连续性方程 (电荷守恒), 后面一个积分项为 0, 保证了运动方程的不变。所以电荷守恒可以理解成规范变换的需求, 或者在逻辑上反过来, 真是因为有电荷守恒, 所以电磁场和物质的作用量项才可以有 $\int d^4x A_\mu j^\mu$ 的形式。

现在推导第二对 Maxwell 方程, 首先, 我们认为电荷的运动是已知的, 所以并不需要变分粒子的坐标, 转而要变分势 (势就类似于场的坐标)。这正好比在求粒子的运动方程的时候, 我们只变分粒子的坐标, 而认为场是已知的。

写出作用量

$$S = S_f + S_{\text{mf}} = - \int d^4x \left(\frac{1}{c} A_\mu j^\mu + a F_{\mu\nu} F^{\mu\nu} \right),$$

其中 $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ 。Larangian 密度为

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{c} A_\mu j^\mu - a F_{\mu\nu} F^{\mu\nu}.$$

为以后的方便, 我们这里先考虑一般的 \mathcal{L} 下的场方程。对 $S = \int d^4x \mathcal{L}$ 固定边界求变分有,

$$\begin{aligned} 0 = \delta S &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial A_\mu} \delta A_\mu + \frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} \delta \partial_\nu A_\mu \right) \\ &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial A_\mu} \delta A_\mu + \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} \delta A_\mu \right) - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} \delta A_\mu \right) \\ &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} \right) \delta A_\mu, \end{aligned}$$

因此, Larangian 密度应该满足方程

$$\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} = 0. \quad (1.11)$$

代入电磁场的 Larangian 密度, 可以计算有

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu A_\mu} \\ &= -\frac{1}{c} j^\mu - \partial_\nu \left(-2a F^{\rho\sigma} \frac{\partial F_{\rho\sigma}}{\partial \partial_\nu A_\mu} \right) \\ &= -\frac{1}{c} j^\mu + 2a \partial_\nu [F^{\rho\sigma} (\delta_\rho^\nu \delta_\sigma^\mu - \delta_\sigma^\nu \delta_\rho^\mu)] \\ &= -\frac{1}{c} j^\mu + 2a \partial_\nu [F^{\nu\mu} - F^{\mu\nu}] \\ &= -\frac{1}{c} j^\mu + 4a \partial_\nu F^{\nu\mu}, \end{aligned}$$

所以我们要求的场方程即

$$\partial_\nu F^{\nu\mu} = \frac{1}{4ac} j^\mu.$$

令 $\mu = 1$,

$$\frac{\partial F^{11}}{\partial x} + \frac{\partial F^{12}}{\partial y} + \frac{\partial F^{13}}{\partial z} + \frac{1}{c} \frac{\partial F^{10}}{\partial t} = -\frac{1}{4ac} j^1,$$

即

$$-\frac{\partial B_z}{\partial y} + \frac{\partial B_y}{\partial z} + \frac{1}{c^2} \frac{\partial E_x}{\partial t} = -\frac{1}{4ac} j^1,$$

连同 $\mu = 2, 3$ 可以给出方程

$$\nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \frac{1}{4ac} \mathbf{j}.$$

剩下的只有 $\mu = 0$ 了, 即

$$\frac{\partial F^{01}}{\partial x} + \frac{\partial F^{02}}{\partial y} + \frac{\partial F^{03}}{\partial z} + \frac{1}{c} \frac{\partial F^{00}}{\partial t} = -\frac{1}{4ac} j^0,$$

也就是

$$-\frac{1}{c} \frac{\partial E_x}{\partial x} - \frac{1}{c} \frac{\partial E_y}{\partial y} - \frac{1}{c} \frac{\partial E_z}{\partial z} = -\frac{1}{4ac} c\rho,$$

或者

$$\nabla \cdot \mathbf{E} = \frac{c}{4a} \rho.$$

前面说过 a 有关于单位制, 对于 SI 制, $a = c\epsilon_0/4$, 其中 ϵ_0 是真空电容率或真空介电常数, 此时

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \frac{1}{c^2 \epsilon_0} \mathbf{j},$$

通常我们还会定义 μ_0 使得 $\mu_0 \epsilon_0 = 1/c^2$, 则

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{j},$$

这就是我们通常在 SI 制下看到的第二组 Maxwell 方程。

以后我们并不采用 SI 制, 而采用 Gauss 制, 此时 $a = c/16\pi$, 那么的第二组 Maxwell 方程写作

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi \mathbf{j} \right).$$

单位转换并不复杂, 只要注意量纲即可, 比如 $e\mathbf{E}$ 就是 Energy/Length² 的量纲, 假设基本力学量的量纲不变, 只要 \mathbf{E} 改变量纲, 就会引起 e 反着变。知道了 \mathbf{E} , 那么只要知道 $c\mathbf{E}$ 和 \mathbf{B} 同量纲, 那么 \mathbf{B} 也清楚了。

现在把已知的两组方程在 Gauss 制下写出来:

$$\begin{cases} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{B} &= \frac{1}{c^2} \left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi \mathbf{j} \right), \end{cases} \quad (1.12)$$

这就是大名鼎鼎的 Maxwell 方程组。

Landau 在他的第二卷采用了另一个物理量 $\mathbf{H} = c\mathbf{B}$ 来表示磁场, 他有着和 \mathbf{E} 一样的量纲, 此时 Maxwell 方程组写作

$$\begin{cases} \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, \\ \nabla \cdot \mathbf{H} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{H} &= \frac{1}{c} \left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi \mathbf{j} \right). \end{cases}$$

四维形式的方程则写作

$$\begin{cases} \partial_{[\mu} F_{\nu\rho]} = 0, \\ \partial_\nu F^{\nu\mu} = 4\pi j^\mu / c^2. \end{cases} \quad (1.13)$$

1.7 能量动量张量

场作为物理实体，我们已知了其本身的 Lagrangian 密度，也应该讨论其能量和动量。

对于场而言，势才是“坐标”，那么类比经典力学，此时 x^μ 应该是做为“时间”用的才对，所以类比于能量关于时间的守恒，我们可以考虑 \mathcal{L} 关于坐标的守恒。为此，考察其导数

$$\partial_\mu \mathcal{L} = \frac{\partial \mathcal{L}}{\partial A^\nu} \partial_\mu A^\nu + \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\nu} \partial_\mu \partial_\rho A^\nu,$$

利用场方程(1.11)，

$$\partial_\mu \mathcal{L} = \partial_\rho \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\nu} \partial_\mu A^\nu + \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\nu} \partial_\mu \partial_\rho A^\nu = \partial_\rho \left(\frac{\partial \mathcal{L}}{\partial \partial_\rho A^\nu} \partial_\mu A^\nu \right),$$

或者改写成

$$\partial_\rho \left(\frac{\partial \mathcal{L}}{\partial \partial_\rho A^\nu} \partial_\mu A^\nu - \delta_\mu^\rho \mathcal{L} \right) = 0,$$

定义

$$T^\nu_\mu = \frac{\partial \mathcal{L}}{\partial \partial_\nu A^\rho} \partial_\mu A^\rho - \delta_\mu^\nu \mathcal{L},$$

或者

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \partial_\mu A^\rho} \partial^\nu A^\rho - \eta^{\mu\nu} \mathcal{L}, \quad (1.14)$$

称之为能量动量张量（简称能动张量）。它满足守恒方程 $\partial_\mu T^{\mu\nu} = 0$ 。注意 $T^{\mu\nu}$ 不一定是唯一的，他始终可以加上一个 $f^{\mu\nu}$ ，其满足 $\partial_\mu f^{\mu\nu} = 0$ ，则保持 $\partial_\mu T^{\mu\nu} = 0$ 不变。

将 $\partial_\mu T^{\mu\nu} = 0$ 对某个四维体积分（包含整个三维空间）有

$$0 = \int_V d^4x \partial_\mu T^{\mu\nu} = \int_{\partial V} dS_\mu T^{\mu\nu},$$

这就是说

$$P^\nu = a \int_{\partial V} dS_\mu T^{\mu\nu}$$

对任意包含整个三维空间的超曲面 S 都相等，其中 a 是一个常数。换句话说， P^μ 是守恒量。而且，从定义中可以直接看到，尽管 $T^{\mu\nu}$ 不是唯一确定的，但 P^μ 是唯一确定。

P^μ 将被理解成这个体系的四维动量矢量，特别地， P^0 应该等于这个体系的能量除以 c ，为此，我们将积分选取在 x^0 为常数的超平面上，就有

$$P^0 = a \int dS_\nu T^{\nu\mu} = a \int dV T^{00},$$

而 T^{00} 可以被计算出来为

$$T^{00} = \dot{A}^\mu \frac{\partial \mathcal{L}}{\partial \dot{A}^\mu} - \mathcal{L},$$

A 既然被认为是场的“坐标”，那么 \dot{A} 就自然被理解成场的“速度”，所以 T^{00} 不是别的，就是场的“能量密度”，为了考察他和真实能量之间的关系，我们只要看 \mathcal{L} 的量纲就好了，因此 $\mathcal{L} d^4x = L_t dt$ ，而 L_t

的量纲就是能量，所以 $c\mathcal{L}dV$ 的量纲才是真实能量，而 $c\mathcal{L}$ 是能量密度的量纲。因此，这里的能量密度是 cT^{00} 。综合上面的量纲分析，我们最后得到 a 应该取 1，那么

$$P^\mu = \int_S dS_\nu T^{\nu\mu}.$$

因为能量动量张量不是唯一确定的，我们可以找一个条件来确定他，这就是体系的角动量守恒。类比

$$M^{\mu\nu} = \sum (x^\mu p^\nu - p^\mu x^\nu) = C,$$

我们写出体系的角动量守恒方程，主要注意把动量换掉，并把离散体系换成连续体系，即

$$C = M^{\mu\nu} = \int_S (x^\mu dP^\nu - x^\nu dP^\mu) = \int_S dS_\rho (x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu}).$$

为了守恒，只有被积式子的散度为 0，即

$$0 = \partial_\rho (x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu}) = T^{\mu\nu} - T^{\nu\mu},$$

这样确定的 $T^{\mu\nu}$ 是一个对称张量。虽然一般来说， $T^{\mu\nu}$ 并不对称，但是我们可以找到一个项 $f^{\mu\nu}$ ，使得 $T^{\mu\nu} + f^{\mu\nu}$ 对称，并且变换后依然满足方程 $\partial_\nu T^{\mu\nu} = 0$ 。

假设我们确定了 $T^{\mu\nu}$ 使得他是一个对称张量。接着我们来确定他的分量，前面已经算过 $cT^{00} = w$ 是能量密度，依然在 x^0 为常数的超平面上积分，则

$$P^\mu = \int dV T^{0\mu},$$

左边是动量，那么右边的被积分的自然就是动量密度的三个分量 (T^{01}, T^{02}, T^{03}) 。

为了明白 $T^{\mu\nu}$ 其他分量的意义，将守恒方程 $\partial_\mu T^{\mu\nu} = 0$ 分成空间和时间部分：

$$\frac{1}{c^2} \partial_t w + \partial_i T^{0i} = 0, \quad \frac{1}{c} \partial_t T^{0i} + \partial_j T^{ji} = 0.$$

第一个方程在空间的一个体积 V 内积分

$$\partial_t \int_V dV w = -c^2 \int_V dV \partial_i T^{i0} = -c^2 \int_{\partial V} d\sigma_i T^{i0},$$

其中 $d\sigma_i$ 是三维空间标准的面元。上式左边是体积内能量的改变量，右边是穿过界面体积的能量，类比电荷守恒方程，其中电流矢量的类比物即矢量 $\mathbf{S} = (c^2 T^{10}, c^2 T^{20}, c^2 T^{30})$ ，我们称之为能流密度，是单位时间内穿过单位面积的能量。由于 $T^{\mu\nu}$ 对称，所以能流密度等于动量密度乘以 c^2 。

对第二个方程，我们同样在空间的一个体积 V 内积分

$$\partial_t \int_V dV T^{0i} = -c \int_{\partial V} d\sigma_j T^{ji},$$

因为 T^{0i} 是动量密度，类比能流密度，右边定义的张量 T^{ji} 应该被称为动量流密度，将之记作 $-\sigma^{ji}$ ，称为应力张量。综上，可以写出 $T^{\mu\nu}$ 矩阵

$$(T^{\mu\nu}) = \frac{1}{c} \begin{pmatrix} -\sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} & S_x/c \\ -\sigma_{yx} & -\sigma_{yy} & -\sigma_{yz} & S_y/c \\ -\sigma_{zx} & -\sigma_{zy} & -\sigma_{zz} & S_z/c \\ S_x/c & S_y/c & S_z/c & w \end{pmatrix}.$$

迄今为止，我们未用到 \mathcal{L} 的具体形式，所以对于任意的 \mathcal{L} 其实都成立，只不过一般情况下变量不是 A^μ 而是 q^a 而已，其中 q 按 a 要求和。具体到电磁场，可以计算

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \partial_\mu A^\rho} \partial^\nu A^\rho - \eta^{\mu\nu} \mathcal{L},$$

其中

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{c}{16\pi} F_{\mu\nu} F^{\mu\nu}.$$

而因为是场本身的 \mathcal{L} ，所以此时 Maxwell 方程组(1.13)中的源 $j^\mu = 0$ ，Maxwell 方程组写作

$$\begin{cases} \partial_{[\mu} F_{\nu\rho]} &= 0, \\ \partial_\nu F^{\nu\mu} &= 0. \end{cases}$$

现在开始计算：

$$T^{\mu\nu} = -\frac{c}{8\pi} F^{\mu'\nu'} \frac{\partial F_{\mu'\nu'}}{\partial \partial_\mu A_\rho} \partial^\nu A_\rho - \eta^{\mu\nu} \mathcal{L} = -\frac{c}{4\pi} F^{\mu\rho} \partial^\nu A_\rho + \frac{c}{16\pi} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma}.$$

这个能动张量并不对称，我们要将之调整为对称的保证角动量守恒，为此做变换

$$\begin{aligned} T^{\mu\nu} &\rightarrow T^{\mu\nu} + \frac{c}{4\pi} F^{\mu\rho} \partial_\rho A^\nu = -\frac{c}{4\pi} F^{\mu\rho} \partial^\nu A_\rho + \frac{c}{4\pi} F^{\mu\rho} \partial_\rho A^\nu + \frac{c}{16\pi} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \\ &= -\frac{c}{4\pi} F^{\mu\rho} (\partial^\nu A_\rho - \partial_\rho A^\nu) + \frac{c}{16\pi} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \\ &= -\frac{c}{4\pi} \eta^{\nu\sigma} F^{\mu\rho} F_{\sigma\rho} + \frac{c}{16\pi} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma}, \end{aligned}$$

可以检验他是对称的，剩下的只要检验添上的一项满足守恒方程即可：

$$\partial_\nu (F^{\nu\rho} \partial_\rho A^\mu) = F^{\nu\rho} \partial_\nu \partial_\rho A^\mu + \partial_\rho A^\mu \partial_\nu F^{\nu\rho} = 0,$$

第一项为零是 $F^{\nu\rho}$ 的反对称性，第二项为零是 $F^{\nu\rho}$ 满足的 Maxwell 方程。检验完毕。

现在，变换后的能动张量

$$T^{\mu\nu} = -\frac{c}{4\pi} \eta^{\mu\sigma} F^{\nu\rho} F_{\sigma\rho} + \frac{c}{16\pi} \eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma}$$

是对称的了。现在，我们来求他的分量，首先

$$F_{\rho\sigma} F^{\rho\sigma} = 2\mathbf{B}^2 - \frac{2}{c^2} \mathbf{E}^2,$$

而 $C^{\mu\nu} = -\eta^{\nu\sigma} F^{\mu\rho} F_{\sigma\rho}$ 分成三部分为：

$$C^{ij} = -\frac{E_i E_j}{c^2} - B_i B_j + \delta_{ij} \mathbf{B}^2, \quad C^{0i} = \frac{1}{c} (\mathbf{E} \times \mathbf{B})_i, \quad C^{00} = \frac{\mathbf{E}^2}{c^2}.$$

综上

$$T^{ij} = \frac{c}{4\pi} C^{ij} - \frac{c}{8\pi} \delta_{ij} \left(\mathbf{B}^2 - \frac{1}{c^2} \mathbf{E}^2 \right) = \frac{c}{8\pi} \left(-2\frac{E_i E_j}{c^2} - 2B_i B_j + \delta_{ij} \mathbf{B}^2 + \frac{\delta_{ij}}{c^2} \mathbf{E}^2 \right),$$

应力张量有

$$\sigma_{ij} = cT^{ij} = \frac{c^2}{4\pi} \left(B_i B_j + \frac{E_i E_j}{c^2} \right) - \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right) \delta_{ij}.$$

能流密度 $\mathbf{S} = (c^2 T^{01}, c^2 T^{02}, c^2 T^{03})$,

$$(\mathbf{S})_i = c^2 T^{0i} = \frac{c^3}{4\pi} C^{0i} = \frac{c^2}{4\pi} (\mathbf{E} \times \mathbf{B})_i,$$

即

$$\mathbf{S} = \frac{c^2}{4\pi} \mathbf{E} \times \mathbf{B}.$$

能量密度

$$w = cT^{00} = \frac{c^2}{4\pi} C^{00} + \frac{c^2}{8\pi} \left(\mathbf{B}^2 - \frac{1}{c^2} \mathbf{E}^2 \right) = \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right).$$

经过上述计算, 我们还可以得到电磁场能动张量的另一个性质 $T_\mu{}^\mu = 0$, 计算如下

$$\begin{aligned} T_\mu{}^\mu &= T_0{}^0 + T_i{}^i = T^{00} - \sum_i T^{ii} \\ &= \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right) + \frac{c^2}{8\pi} \sum_i \left(2 \frac{E_i^2}{c^2} + 2B_i^2 - \mathbf{B}^2 - \frac{1}{c^2} \mathbf{E}^2 \right) \\ &= \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right) - \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right) = 0. \end{aligned}$$

我们上面对场的能量密度、能流密度有了定义, 这是建立在场和粒子之间没有相互作用上的。现在, 我们引入粒子, 考察能量守恒定律。为此, 我们对电磁场本身能量求时间的导数

$$\partial_t \int_V w dV = \frac{c^2}{4\pi} \int_V dV \left(\mathbf{B} \cdot \partial_t \mathbf{B} + \frac{1}{c^2} \mathbf{E} \cdot \partial_t \mathbf{E} \right),$$

这里使用含源的 Maxwell 方程组(1.12)替换掉场对时间的偏导数,

$$\partial_t \int_V w dV = \frac{c^2}{4\pi} \int_V dV \left(-\mathbf{B} \cdot (\nabla \times \mathbf{E}) + \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \frac{4\pi}{c^2} \mathbf{E} \cdot \mathbf{j} \right),$$

使用我们幼儿园就熟知的矢量分析公式

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})$$

以及能流密度 $\mathbf{S} = c^2 \mathbf{E} \times \mathbf{B} / 4\pi$, 则

$$\partial_t \int_V w dV = - \int_V dV \nabla \cdot \mathbf{S} - \int_V dV \mathbf{E} \cdot \mathbf{j} = - \int_{\partial V} d\boldsymbol{\sigma} \cdot \mathbf{S} - \int_V dV \mathbf{E} \cdot \mathbf{j},$$

如果源满足 $\mathbf{j} = 0$, 则又回到我们前面见过的形式, 现在多了最后一项, 肯定和粒子有关。单独考察这一项, 化连续为离散, 改写为

$$- \sum e \mathbf{E} \cdot \mathbf{v},$$

回忆(1.10), 我们就会发现, 这项就是所有带电粒子构成的系统的能量(动能)的导数之和 $-\sum dE/dt$ 。所以

$$\partial_t \int_V w dV = - \int_{\partial V} d\boldsymbol{\sigma} \cdot \mathbf{S} - \frac{d}{dt} \sum E,$$

也即

$$\partial_t \int_V w dV + \frac{d}{dt} \sum E = - \int_{\partial V} d\boldsymbol{\sigma} \cdot \mathbf{S}. \quad (1.15)$$

选定一个空间封闭曲面, 假设在我们关心的时刻, 带电粒子本身不穿过这个封闭曲面, 此时(1.15)可以写作

$$\partial_t \left(\int_V w dV + \sum E \right) = - \int_{\partial V} d\boldsymbol{\sigma} \cdot \mathbf{S},$$

左边括号里面的项是整个系统所具有的总能量, 即粒子和场所具有的总能量。如果右边积分为 0, 比如 V 取做整个三维空间, 再比如能流密度为 0, 那么左边为 0 就是在说在这个封闭曲面内的系统能量恒定。如

果带电粒子本身在我们关心的那个时刻穿过这个封闭曲面，则此时右边应该添加一项粒子输运的项，粒子和场总系统的能量不仅由场带来带去，也由由粒子本身带走或者带来了。

对于含源的能量动量张量的考察，本质上只要对含源的 \mathcal{L} 重复一次上面的分析即可，同样，上面关于能量守恒的考察也自然地会出现在守恒方程中。我们不再重复，却有理由相信，系统的能量是粒子的动能和电磁场的能量之和，系统的动量也是粒子的动量和电磁场的动量之和。在电磁场中的粒子虽不一定动量、能量守恒，但是只要和电磁场一起考虑，则动量、能量守恒依然成立。

1.8 Noether 定理

经典场论是对一般的 \mathcal{L} 处理作用量

$$S = \int d^4x \mathcal{L}(\phi^a, \partial_\mu \phi^a)$$

的，其中 a 是场分量的编号，不一定只有三个或者四个。关于这种作用量，我们前面已经求过他的 Lagrange 方程：

$$\frac{\partial \mathcal{L}}{\partial \phi^a} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi^a} = 0. \quad (1.16)$$

此外，我们对一般的 \mathcal{L} ，定义了能量动量张量：

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^a} \partial^\nu \phi^a - \eta^{\mu\nu} \mathcal{L}, \quad (1.17)$$

它满足守恒方程：

$$\partial_\mu T^{\mu\nu} = 0, \quad (1.18)$$

同时也用能量动量张量定义了场的四维动量和角动量张量：

$$P^\mu = \int dS_\nu T^{\nu\mu}, \quad M^{\mu\nu} = \int dS_\rho (x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu}). \quad (1.19)$$

现在我们讨论所谓的 Noether 定理，他建立了对称性和守恒律的联系。为方便起见，下面我们设 ϕ^a 中的指标只能取一个值，也就是只有单变量 ϕ 。

假设对场 ϕ ，我们有无穷小变换

$$\phi(x) \rightarrow \phi(x) + \alpha(x) \Delta \phi(x), \quad (1.20)$$

那么他就会诱导 \mathcal{L} 产生变换

$$\mathcal{L} \rightarrow \mathcal{L} + \Delta \mathcal{L} = \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi} \alpha(x) \Delta \phi(x) + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \partial_\nu [\alpha(x) \Delta \phi(x)].$$

如果

$$\int d^4x \Delta \mathcal{L} = 0,$$

那么这个时候作用量是不变的

假设 $\alpha(x) = \alpha$ 是一个常数，此时

$$\Delta \mathcal{L} = \alpha \left(\frac{\partial \mathcal{L}}{\partial \phi} \Delta \phi(x) + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \partial_\nu [\Delta \phi(x)] \right) = \alpha \Delta \phi(x) \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \right) + \alpha \partial_\nu \left(\Delta \phi(x) \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \right),$$

如果我们的场满足 Lagrange 方程，则第一项自然为 0，第二项作为散度项，对全空间的积分可以化作无穷远处曲面上的积分，如果场在无穷远处消失了，则

$$\int d^4x \Delta \mathcal{L} = 0$$

此时是自然满足的。因此，若是我们的场满足 Lagrange 方程，那么(1.20)对应的 $\alpha(x)$ 是一个常数的变换保持作用量不变。这样的变换称为全局变换，而上面的话可以重新表述为：如果场满足 Lagrange 方程，全局变换保持作用量不变。全局变换诱导的对称性通常被称为全局对称性。与全局对称性联系的是有限参数决定的 Lie 群，反之，局部对称性依赖于时间和空间的函数。

所以我们关注的是更一般的情况，即就算场不满足 Lagrange 方程，场的变换却依然保持作用量不变，这样的变换称之为对称变换。此时，场的变化诱导的作用量变化的一般形式为

$$\Delta S = \int d^4x J^\mu \partial_\mu \alpha(x) = - \int d^4x \alpha(x) \partial_\mu J^\mu, \quad (1.21)$$

其中分部积分后散度项的消失是场在足够远处消失的要求，这里的 J 被称为守恒荷。如果为了保证变换是一种对称变换，即 $\Delta S = 0$ ，就必须

$$\partial_\mu J^\mu = 0,$$

这个方程称为守恒方程，他诱导了一个守恒量

$$F = \int_V J^0 dV.$$

为了看出他是守恒的，直接将对时间求导

$$\dot{F} = c \int_V \partial_0 J^0 dV = -c \int_V \partial_i J^i dV = -c \int_V J^i d\sigma_i$$

J 在无穷远消失的话，则 F 守恒。因此，Noether 定理告诉我们，对称性诱导了守恒律。

很多时候，对称变换不只保持作用量不变，而且当(1.20)中的 α 为常数时还保持 Lagrangian 密度 \mathcal{L} 不变，即

$$\frac{\partial \mathcal{L}}{\partial \phi} \Delta \phi(x) + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \partial_\nu [\Delta \phi(x)] = 0,$$

那么

$$\Delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi} \alpha(x) \Delta \phi(x) + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \partial_\nu [\alpha(x) \Delta \phi(x)] = \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \Delta \phi(x) \partial_\nu \alpha(x),$$

或者

$$\Delta S = \int d^4x \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \Delta \phi(x) \partial_\nu \alpha(x),$$

对比一般的表达式，我们就有

$$J^\mu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \Delta \phi(x).$$

作为例子，考察电磁场的平移变换

$$A^\mu(x) \rightarrow A^\mu(x + \epsilon(x)) = A^\mu + \epsilon^\nu(x) \partial_\nu A^\mu(x),$$

所以

$$(\Delta A^\mu)_\nu = \partial_\nu A^\mu(x),$$

然后考察作用量的变化⁶

$$\begin{aligned}
 \Delta S &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial A^\mu} \epsilon^\nu(x) \partial_\nu A^\mu(x) + \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\mu} \partial_\rho [\epsilon^\nu(x) \partial_\nu A^\mu(x)] \right) \\
 &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial A^\mu} \partial_\nu A^\mu(x) + \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\mu} \partial_\rho \partial_\nu A^\mu(x) \right) \epsilon^\nu(x) + \int d^4x \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\mu} \partial_\rho \epsilon^\nu(x) \partial_\nu A^\mu(x) \\
 &= \int d^4x \left(\epsilon^\nu(x) \partial_\nu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \partial_\rho A^\mu} \partial_\nu A^\mu(x) \partial_\rho \epsilon^\nu(x) \right) \\
 &= \int d^4x \left(\delta_\nu^\mu \epsilon^\nu(x) \partial_\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \partial_\mu A^\rho} \partial_\nu A^\rho(x) \partial_\mu \epsilon^\nu(x) \right) \\
 &= \int d^4x \left(-\delta_\nu^\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \partial_\mu A^\rho} \partial_\nu A^\rho(x) \right) \partial_\mu \epsilon^\nu(x),
 \end{aligned}$$

所以守恒荷为

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial \partial_\mu A^\rho} \partial_\nu A^\rho(x) - \delta_\nu^\mu \mathcal{L}$$

满足守恒方程

$$\partial_\mu T^\mu_\nu = 0.$$

这个其实就是能动张量。

相对论场论, 很重要的变换就是 Lorentz 变换, 对应着无穷小变换 $\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu$, 其中 $\omega_{\mu\nu} = -\omega_{\nu\mu}$, 如果作用量有 Lorentz 不变性, 则自然有着对应的守恒荷 $M^{\rho\mu\nu}$, 当然对应着角动量, 通过对这个守恒荷的计算和分析, 我们就有希望建立一个对称的能量动量张量, 这里不再计算, 具体可以查看 Weinberg 的 QFT V1 中的 section 7.4。

1.9 微分形式语言描述的 Maxwell 方程组

这节来稍稍讲讲微分形式语言描述的 Maxwell 方程组, 假定读者有足够的微分几何知识。现在开始, 我们讨论的背景定为 n 维定向伪 Riemann 流形上。此外, 在这节里面不妨令 $c = 1$, 则此时, Gauss 制下的 Maxwell 方程写作:

$$\begin{cases} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t} + 4\pi\mathbf{j}, \end{cases} \quad (1.22)$$

和

$$\begin{cases} \partial_{[\mu} F_{\nu\rho]} &= 0, \\ \partial_\nu F^{\nu\mu} &= -4\pi j^\mu, \end{cases} \quad (1.23)$$

此时

$$F^{\mu\nu} F_{\mu\nu} = 2(\mathbf{B}^2 - \mathbf{E}^2).$$

Definition 1. 在 p 阶形式空间 $\Omega^p(M)$ 和 $\Omega^{n-p}(M)$ 定义一个线性算符 $\star: \Omega^p(M) \rightarrow \Omega^{n-p}(M)$ 使得, 对任意的 $\omega, \mu \in \Omega^p(M)$ 成立

$$\omega \wedge (\star\mu) = \langle \omega, \mu \rangle \text{vol},$$

⁶这是一个很好的算例, 从始至终, 没有使用 Lagrange 方程, 就是说, 即使 \mathcal{L} 不满足 Lagrange 方程, \mathcal{L} 本身也不是不变的, 作用量的改变量却还是有(1.21)的形式。

其中 $\langle *, * \rangle$ 是两个 p 形式之间的内积⁷, 而 vol 是 M 上的体积形式, 在 $\mathbb{R}^{s+(n-s)}$ 即 $\text{vol} = dx^1 \wedge \cdots \wedge dx^n$ 。这样定义的 \star 被称为 *Hodge 星算子*。

Proposition 1. 在 $\mathbb{R}^{s+(n-s)}$ 上, 在 p -形式上成立恒等式

$$\star\star = (-1)^{s+p(n-p)}.$$

证明. 直接计算

$$dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge (\star dx^{i_1} \wedge \cdots \wedge dx^{i_p}) = g_{i_1 i_1} \cdots g_{i_p i_p} \text{sign}(\sigma) dx^{i_1} \wedge \cdots \wedge dx^{i_n},$$

其中 $\sigma(1, \dots, n) = (i_1, \dots, i_n)$. 所以

$$\star dx^{i_1} \wedge \cdots \wedge dx^{i_p} = g_{i_1 i_1} \cdots g_{i_p i_p} \text{sign}(\sigma) dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n},$$

同样

$$\begin{aligned} dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n} \wedge (\star dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n}) \\ = g_{i_{p+1} i_{p+1}} \cdots g_{i_n i_n} \text{sign}(\sigma) dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n} \\ = (-1)^{p(n-p)} g_{i_{p+1} i_{p+1}} \cdots g_{i_n i_n} \text{sign}(\sigma) dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}. \end{aligned}$$

因此

$$\star dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n} = (-1)^{p(n-p)} g_{i_{p+1} i_{p+1}} \cdots g_{i_n i_n} \text{sign}(\sigma) dx^{i_1} \wedge \cdots \wedge dx^{i_p},$$

最后

$$\star\star = (-1)^{p(n-p)} g_{i_1 i_1} \cdots g_{i_n i_n} \text{sign}(\sigma)^2 = (-1)^{s+p(n-p)}.$$

对于我们熟悉的 3 + 1 度规, 此时 $\star\star = (-1)^{3+p(4-p)} = (-1)^{p^2-1}$. □

现在我们扯回 Maxwell 方程, 在三维空间上, 我们定义⁸电场强度 \mathbf{E} 是一个 1-形式 $\mathbf{E} = E_i dx^i$, 磁感应强度 \mathbf{B} 是一个 2-形式

$$\mathbf{B} = \frac{1}{2} \sum_k B_k \epsilon_{kij} dx^i \wedge dx^j = B_1 dx^2 \wedge dx^3 - B_2 dx^1 \wedge dx^3 + B_3 dx^1 \wedge dx^2.$$

假如在 \mathbb{R}^{3+1} 上定义 2-形式

$$\mathbf{F} = -\mathbf{B} - \mathbf{E} \wedge dx^0,$$

我们可以将其写成分量形式, 即 $\mathbf{F} = F_{\mu\nu} dx^\mu \wedge dx^\nu / 2$, 一般来说 $F_{\mu\nu}$ 并不唯一确定, 但是由于 $dx^\mu \wedge dx^\nu$ 是反对称的, 因此我们规定 $F_{\mu\nu}$ 也是反对称的, 对比上面的式子 $\mathbf{F} = -\mathbf{B} - \mathbf{E} \wedge dx^0$, 可以得到矩阵 $(F_{\mu\nu})$ 为

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -B_3 & B_2 & -E_1 \\ B_3 & 0 & -B_1 & -E_2 \\ -B_2 & B_1 & 0 & -E_3 \\ E_1 & E_2 & E_3 & 0 \end{pmatrix},$$

⁷假设 e^i 和 f^i 都是 1-形式, 对于 p -形式 $e^1 \wedge \cdots \wedge e^p$ 和 $f^1 \wedge \cdots \wedge f^p$, 他们的内积定义为

$$\langle e^1 \wedge \cdots \wedge e^p, f^1 \wedge \cdots \wedge f^p \rangle = \det(g^{-1}(e^i, f^j)),$$

其中 g^{-1} 是余切空间中由切空间中的度规 g 诱导的度规, 他的矩阵 g^{ij} 是切空间度规矩阵 g_{ij} 的逆, 即 $g^{ij} g_{jk} = \delta_k^i$. 因为内积对函数线性, 所以内积就对任意两个 p -形式都有了定义。

⁸下面定义的形式虽然是粗体, 但不是斜体, 粗斜体还是交给我们熟悉的三维矢量。

这就是我们熟悉的电磁场张量。

可以计算

$$\mathbf{F} \wedge \star \mathbf{F} = \langle \mathbf{B} + \mathbf{E} \wedge dx^0, \mathbf{B} + \mathbf{E} \wedge dx^0 \rangle \text{vol} = (\mathbf{B}^2 - \mathbf{E}^2) \text{vol} = \frac{1}{2} F^{\mu\nu} F_{\mu\nu} \text{vol}.$$

因此 $\mathbf{F} \wedge \star \mathbf{F}$ 正比于场的 Lagrangian 密度, 而以前也说明, 正比的系数只决定我们所选取的单位制, 就理论而言, 并不重要。将来若是要改写 Lagrangian 密度到一个坐标无关且适合任意流形的形式, 则 $\mathbf{F} \wedge \star \mathbf{F}$ 是要比 $F^{\mu\nu} F_{\mu\nu}/2$ 合适一些。

作为 2-形式, \mathbf{F} 被 Hodge 星算子作用后的 $\star \mathbf{F}$ 也是 2-形式, 即

$$\star \mathbf{F} = -\star \mathbf{B} - \star(\mathbf{E} \wedge dx^0),$$

利用我们的度规, 则

$$\star dx^i \wedge dx^j = \sum_k \epsilon^{ijk} dx^k \wedge dx^0, \quad \star dx^i \wedge dx^0 = -\frac{1}{2} \sum_{jk} \epsilon^{ijk} dx^j \wedge dx^k.$$

于是我们可以写出矩阵

$$(\star F_{\mu\nu}) = \begin{pmatrix} 0 & E_3 & -E_2 & -B_1 \\ -E_3 & 0 & E_1 & -B_2 \\ E_2 & -E_1 & 0 & -B_3 \\ B_1 & B_2 & B_3 & 0 \end{pmatrix},$$

即做了如下变换

$$\mathbf{E} \rightarrow \mathbf{B}, \quad \mathbf{B} \rightarrow -\mathbf{E}.$$

现在我们计算 \mathbf{F} 的外微分

$$\begin{aligned} d\mathbf{F} &= -d\mathbf{B} - (d\mathbf{E}) \wedge dx^0 \\ &= -\nabla \cdot \mathbf{B} dx^1 \wedge dx^2 \wedge dx^3 - \partial_0 \mathbf{B} \wedge dx^0 - \partial_i E_j dx^i \wedge dx^j \wedge dx^0 \\ &= -\nabla \cdot \mathbf{B} dx^1 \wedge dx^2 \wedge dx^3 - \left(\partial_i E_j + \frac{1}{2} \sum_k \partial_0 B_k \epsilon_{kij} \right) dx^i \wedge dx^j \wedge dx^0, \end{aligned} \tag{1.24}$$

磁感应强度的散度自然为 0, 第二项括号里面的项由 Maxwell 方程 $\nabla \times \mathbf{E} + \partial_0 \mathbf{B} = 0$ 也可以知道为 0, 所以由第一组 Maxwell 方程可知 $d\mathbf{F} = 0$, 反之, 如果 $d\mathbf{F} = 0$ 成立, 则第一组 Maxwell 方程成立。所以第一组 Maxwell 方程等价于说 \mathbf{F} 是一个闭形式。

剩下的自然是第二组 Maxwell 方程, 第二组含源, 所以我们再引入一个电流矢量

$$\mathbf{j} = \rho \partial_0 + j^i \partial_i,$$

利用度规转化为电流 1-形式

$$\mathbf{j} = \rho dx^0 - \sum_i j^i dx^i,$$

我们下面证明, 第二组 Maxwell 方程写作

$$\star d \star \mathbf{F} = -4\pi \mathbf{j}.$$

首先计算 $d \star \mathbf{F}$, 利用变换 $\mathbf{E} \rightarrow \mathbf{B}, \mathbf{B} \rightarrow -\mathbf{E}$ 和(1.24)可以直接得到

$$d \star \mathbf{F} = \nabla \cdot \mathbf{E} dx^1 \wedge dx^2 \wedge dx^3 - \left(\partial_i B_j - \sum_k \partial_0 E_k \epsilon_{kij} \right) dx^i \wedge dx^j \wedge dx^0,$$

所以

$$\star d \star \mathbf{F} = -\nabla \cdot \mathbf{E} dx^0 + \sum_l \left(\partial_i B_j - \frac{1}{2} \sum_k \partial_0 E_k \epsilon_{kij} \right) \epsilon^{lij} dx^l,$$

因此, $\star d \star \mathbf{F} = -4\pi \mathbf{j}$ 等价于

$$\nabla \cdot \mathbf{E} = 4\pi \rho, \quad \nabla \times \mathbf{B} = \partial_0 \mathbf{E} + 4\pi \mathbf{j}.$$

这就是第二组 Maxwell 方程。

如果我们的空间是可缩的, 此时第一组 Maxwell 方程的闭形式则等价于恰当形式, 即存在 1-形式 \mathbf{A} 使得 $d\mathbf{A} = \mathbf{F}$, 此时结合第二组方程则写作

$$-4\pi \mathbf{j} = \star d \star \mathbf{F} = \star d \star d\mathbf{A},$$

这就是势 \mathbf{A} 所需要满足的方程。

将 $\star d \star \mathbf{F} = -4\pi \mathbf{j}$ 改写成 $d \star \mathbf{F} = \pm 4\pi \star \mathbf{j}$, 两边取外微分有

$$0 = \pm 4\pi d \star \mathbf{j},$$

则 $d \star \mathbf{j} = 0$ 其实就是我们的电荷守恒方程。

现在假设是无源的, 则两个方程写作

$$d\mathbf{F} = 0, \quad d \star \mathbf{F} = 0.$$

所以 $\mathbf{F} \rightarrow \star \mathbf{F}$ 保持方程不变。我们现在就讨论这种变换, 由于 \star 是线性算子, 故可以考察其本征值。因为在 \mathbb{R}^{3+1} 下, 对于 2-形式有 $\star \star = (-1)^3 = -1$, 所以, 我们大概可以相信 \star 的本征值为 $\pm i$, 对应的本征矢量为 \mathbf{F}_\pm , 其余的矢量可以写作

$$\mathbf{F} = \mathbf{F}_+ + \mathbf{F}_-,$$

其中

$$\star \mathbf{F}_\pm = \pm i \mathbf{F}_\pm.$$

对于 \mathbf{F}_\pm 而言, $d\mathbf{F} = 0$ 就可以推出

$$0 = \pm i d\mathbf{F}_\pm = d \star \mathbf{F}_\pm.$$

这节的最后, 我们改写电磁场的作用量为

$$S = - \int_M \mathbf{A} \wedge \star \mathbf{j} - \frac{1}{8\pi} \int_M \mathbf{F} \wedge \star \mathbf{F}.$$

Chapter 2

Hamilton 形式化

前面一章主要讨论了经典场论的 Lagrange 形式化，正如在 Noether 定理那里看到的一样，在 Lagrange 形式化下，对称性是容易分析的。这章我们来讨论 Hamilton 形式化，当我们想要量子化一个经典场，Hamilton 形式化就会更加实用，一方面因为 Hamilton 算子在量子力学里面是自然的，另一方面，Poisson 括号和对易子存在方便的对应。这章假设 $c = 1$ 。

2.1 正则方程和 Poisson 括号

从经典力学中已经得知，对于 L_t ，运动方程即 Lagrange 方程写作

$$\frac{d}{dt} \frac{\partial L_t}{\partial \dot{q}^a} - \frac{\partial L_t}{\partial q^a} = 0.$$

广义坐标 q^a 对应的广义动量定义为

$$p_a = \frac{\partial L_t}{\partial \dot{q}^a}.$$

而 Hamiltonian 定义为

$$H(p_a, q^a) = p_a \dot{q}^a - L_t,$$

其中所有的量都应该写成 p_a 和 q^a 的函数。此时运动方程，即 Hamilton 方程写作

$$\dot{p}_a = -\frac{\partial H}{\partial q^a}, \quad \dot{q}^a = \frac{\partial H}{\partial p_a}.$$

这些都是大家熟悉的。

Poisson 括号 $[\star, \star]_P$ 定义如下¹

$$[F, G]_P = \frac{\partial F}{\partial q^a} \frac{\partial G}{\partial p_a} - \frac{\partial F}{\partial p_a} \frac{\partial G}{\partial q^a},$$

所以 $[p_a, q^b]_P = -\delta_a^b$ 以及

$$\dot{f}(t, p, q) = \frac{\partial f}{\partial t} - [H, f]_P,$$

此时运动方程写作

$$\dot{p}_a = [p_a, H]_P, \quad \dot{q}^a = [q^a, H]_P.$$

¹量子化时， $\hbar[\star, \star]_P \rightarrow [\star, \star]/i$ ，后者是算符之间的对易子。

用一个例子，我们来演示如何将这套用在场论里。假设有 N 个质点，他们处于一条直线上且相距为 a ，之间有劲度系数相同的轻弹簧相连，静止时总长为 L 。对于这样一个系统，选取每个质点离开原点的位移量 x_n 作为广义坐标，此时的 Lagrangian 写作

$$L_t = \sum_{i=1}^N \frac{m}{2} \dot{x}_i^2 - \sum_{i=1}^{N-1} \frac{m\omega^2}{2} (x_{i+1} - x_i - a)^2,$$

设 $x_i = (i-1)a + \phi_i$ ，这就是说 ϕ_i 是第 i 个质点偏离其初始位置的位移，此时 Lagrangian 写作

$$L_t = \sum_{i=1}^N \frac{m}{2} \dot{\phi}_i^2 - \sum_{i=1}^{N-1} \frac{m\omega^2}{2} (\phi_{i+1} - \phi_i)^2.$$

现在假设 N 很大，做变换

$$\phi_n \rightarrow \sqrt{a}\phi(x)|_{x=na}, \quad \phi_{n+1} - \phi_n \rightarrow a^{3/2}\partial_x\phi(x)|_{x=na}, \quad \sum_i \rightarrow \frac{1}{a} \int_0^L dx,$$

则系统的 Lagrangian 可以近似写作

$$L_t = \int_0^L dx \left(\frac{m}{2} \dot{\phi}^2 - \frac{m\omega^2 a^2}{2} (\partial_x \phi)^2 \right),$$

即

$$\mathcal{L} = \frac{m}{2} \dot{\phi}^2 - \frac{m\omega^2 a^2}{2} (\partial_x \phi)^2. \quad (2.1)$$

计算原来系统的广义动量

$$\pi_n = \frac{\partial L_t}{\partial \dot{\phi}_n} = m\dot{\phi}_n, \quad (2.2)$$

所以 Hamilton 写作

$$H = \sum_{i=1}^N \frac{\pi_i^2}{2m} + \sum_{i=1}^{N-1} \frac{m\omega^2}{2} (\phi_{i+1} - \phi_i)^2,$$

一样得，在大 N 下作变换可以得到近似

$$H = \int_0^L dx \left(\frac{\pi^2}{2m} + \frac{m\omega^2 a^2}{2} (\partial_x \phi)^2 \right). \quad (2.3)$$

从(2.2)来看，在大 N 下，有 $\pi(x) = m\dot{\phi}(x)$ ，广义坐标们变成了场 $\phi(x)$ 而广义动量则变成了另一个场 $\pi(x)$ ，从(2.1)中容易读出

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \quad (2.4)$$

以及从(2.3)中读出

$$H = \int_0^L dx \pi \dot{\phi} - L_t.$$

可见，到了场论，为了完成 Hamilton 形式化，就必须利用(2.4)去寻找 ϕ 的对偶场 π ，这就如同在经典力学里面去寻找广义动量一样。找到对偶场之后，将原本的场关于时间的导数反解出来，就写出了 Hamiltonian.

上面这个例子挺有趣的，他用经典理论建立了一个简单的一维固体的模型，可以看作将场论引入凝聚态中的一个最简单的例子。继续他的求解只要写出场方程（是声波的波动方程）即可，这里略去。下面转而使用变分法给出场的 Hamilton 形式化的推导。

对于任何场 $Q^a(\mathbf{x}, t)$ 的泛函 $F[Q^a(t)]$ ，通过

$$\delta F[Q^a(t)] = \int d^3\mathbf{x} \frac{\delta F[Q^a(t)]}{\delta Q^a(\mathbf{x}, t)} \delta Q^a(\mathbf{x}, t)$$

可以定义变分算符

$$\frac{\delta F[Q^a(t)]}{\delta Q^a(\mathbf{x}, t)}, \quad (2.5)$$

对于固定的 \mathbf{x} ，那么 $Q^a(\mathbf{x}, t)$ 也是一种泛函²，所以很容易从定义中得到

$$\frac{\delta Q^a(\mathbf{x}', t)}{\delta Q^b(\mathbf{x}, t)} = \delta_b^a \delta^3(\mathbf{x}' - \mathbf{x}).$$

如果有不止一个可以独立变分的场（比如 Lagrangian 中的 \dot{Q}^a ），那么(2.5)中就有对其他可以独立变分的场的项。

下面开始正式推导。首先作用量是场的泛函

$$S[Q^a, \dot{Q}^a] = \int dt L_t[Q^a(t), \dot{Q}^a(t)],$$

做变分应该有

$$\delta S = \int dt \int d^3\mathbf{x} \left(\frac{\delta L_t}{\delta Q^a(\mathbf{x}, t)} \delta Q^a(\mathbf{x}, t) + \frac{\delta L_t}{\delta \dot{Q}^a(\mathbf{x}, t)} \delta \dot{Q}^a(\mathbf{x}, t) \right),$$

使用分部积分，然后去掉边界项，就得到了

$$\delta S = \int d^4x \left(\frac{\delta L_t}{\delta Q^a(\mathbf{x}, t)} - \frac{d}{dt} \frac{\delta L_t}{\delta \dot{Q}^a(\mathbf{x}, t)} \right) \delta Q^a(\mathbf{x}, t),$$

作用量原理告诉我们，场方程写作

$$\frac{\delta L_t}{\delta Q^a(\mathbf{x}, t)} - \frac{d}{dt} \frac{\delta L_t}{\delta \dot{Q}^a(\mathbf{x}, t)} = 0. \quad (2.6)$$

按照一般的假设， $x = (\mathbf{x}, t)$ 和

$$L_t = \int d^3\mathbf{x} \mathcal{L}[Q^a(x), \nabla Q^a(x), \dot{Q}^a(x)],$$

对他变分

$$\delta L_t = \int d^3\mathbf{x} \left(\frac{\partial \mathcal{L}}{\partial Q^a} \delta Q^a + \frac{\partial \mathcal{L}}{\partial \partial_i Q^a} \delta \partial_i Q^a + \frac{\partial \mathcal{L}}{\partial \dot{Q}^a} \delta \dot{Q}^a \right),$$

分部积分并去掉边界项

$$\delta L_t = \int d^3\mathbf{x} \left[\left(\frac{\partial \mathcal{L}}{\partial Q^a} - \partial_i \frac{\partial \mathcal{L}}{\partial \partial_i Q^a} \right) \delta Q^a + \frac{\partial \mathcal{L}}{\partial \dot{Q}^a} \delta \dot{Q}^a \right],$$

所以

$$\frac{\delta L_t}{\delta Q^a} = \frac{\partial \mathcal{L}}{\partial Q^a} - \partial_i \frac{\partial \mathcal{L}}{\partial \partial_i Q^a}, \quad \frac{\delta L_t}{\delta \dot{Q}^a} = \frac{\partial \mathcal{L}}{\partial \dot{Q}^a}. \quad (2.7)$$

观察(2.7)的第二项为

$$\frac{\delta L_t}{\delta \dot{Q}^a(x)} = \frac{\partial \mathcal{L}}{\partial \dot{Q}^a}(x) = P_a(x).$$

²或者写作

$$Q^a(\mathbf{x}, t) = \int d^3\mathbf{x}' Q^a(\mathbf{x}', t) \delta^3(\mathbf{x}' - \mathbf{x}),$$

所以对于固定的 \mathbf{x} ， $Q^a(\mathbf{x}, t)$ 是一个泛函。

至此,利用变分算符,我们重新改写了场方程的定义。从(2.7)和对偶场的定义可以看到,粒子的 Lagrange 描述到场的 Lagrange 描述,只需要将偏导数算符转为变分算符就可以了。如果利用(2.7),代入场方程(2.6)就得到了熟知的场方程

$$\frac{\partial \mathcal{L}}{\partial Q^a} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu Q^a} = 0.$$

利用变分算符,对于场论的 Hamilton 形式化就可以如下进行,首先寻找对偶场

$$\frac{\delta L_t}{\delta \dot{Q}^a(x)} = P_a(x),$$

然后使用 P_a 和 Q^a 反解出 \dot{Q}^a , 构造 Hamiltonian

$$H[Q^a(t), P_a(t)] = \int d^3\mathbf{x} P_a(x) \dot{Q}^a(x) - L_t[Q^a(t), \dot{Q}^a(t)],$$

此时正则方程写作

$$\dot{P}_a(x) = -\frac{\delta H}{\delta Q^a(x)}, \quad \dot{Q}^a(x) = \frac{\delta H}{\delta P_a(x)}.$$

我们还可以如下定义场论的 Poisson 符号

$$[F[Q^a(t), P_a(t)], G[Q^a(t), P_a(t)]]_P = \int d^3\mathbf{x} \left(\frac{\delta F}{\delta Q^a(\mathbf{x}, t)} \frac{\delta G}{\delta P_a(\mathbf{x}, t)} - \frac{\delta F}{\delta P_a(\mathbf{x}, t)} \frac{\delta G}{\delta Q^a(\mathbf{x}, t)} \right),$$

此时

$$[F, P_a(\mathbf{x}, t)] = \frac{\delta F}{\delta Q^a(\mathbf{x}, t)}, \quad [Q^a(\mathbf{x}, t), F] = \frac{\delta F}{\delta P_a(\mathbf{x}, t)}.$$

2.2 Hamilton 形式化的困难：约束系统

强匀强磁场中的电荷

前面已经知道,电磁场中的电荷的 Lagrangian 写作

$$L_t = -mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} + e\mathbf{A} \cdot \mathbf{v} - e\varphi,$$

如果磁场足够大,那么就可以无视掉动能项,只留下

$$L_t = e\mathbf{A} \cdot \mathbf{v} - e\varphi,$$

假设磁场是匀强,且方向为 \hat{z} , 设

$$\mathbf{A} = \frac{B_0}{2}(x\hat{y} - y\hat{x}) - e\varphi,$$

所以

$$L_t = \frac{eB_0}{2}(x\dot{y} - y\dot{x}) - e\varphi.$$

假设 φ 不显含 z 坐标,这样就可以顺便无视掉 \hat{z} 方向的匀速运动。我们考察这样一个二维系统。

首先 Lagrange 方程写作

$$B_0\dot{y} = \partial_x \varphi, \quad B_0\dot{x} = -\partial_y \varphi.$$

这就是正确的运动方程。

按照标准的 Legendre 变换,先计算正则动量

$$p_x = \partial_{\dot{x}} L_t = -\frac{eB_0}{2}y, \quad p_y = \partial_{\dot{y}} L_t = \frac{eB_0}{2}x.$$

然后

$$H = p_x \dot{x} + p_y \dot{y} - L_t = \frac{eB_0}{2}(x\dot{y} - \dot{x}y) - L_t = e\varphi(x, y).$$

不用继续做下去了，首先 \dot{x} 和 \dot{y} 不能用 p_x 和 p_y 反解出来，此外，即使写出了 Hamiltonian，因为他不显含正则动量，利用正则方程得到的运动方程是 $\dot{x} = 0$ 和 $\dot{y} = 0$ ，显然这和上面使用 Lagrange 方程确定的正则方程是不同的。

重新审视正则动量 $p_x = -eB_0 y/2$ 和 $p_y = eB_0 x/2$ ，他所有的变量不是正则坐标就是正则动量，所以这两个方程

$$f_1(x, y, p_x, p_y) = p_x + \frac{eB_0}{2}y = 0,$$

$$f_2(x, y, p_x, p_y) = p_y - \frac{eB_0}{2}x = 0,$$

在相空间确定了一个曲面，而 Hamiltonian 是在这个曲面上写成 $e\varphi(x, y)$ 。因此，在整个相空间上，Hamiltonian 具有形式

$$H = e\varphi + \eta_1 f_1 + \eta_2 f_2.$$

这样去写正则方程，给出

$$\dot{p}_x = -\partial_x H = -e\partial_x \varphi - f_1 \partial_x \eta_1 - f_2 \partial_x \eta_2 + \frac{eB_0}{2}\eta_2,$$

$$\dot{x} = \partial_{p_x} H = f_1 \partial_{p_x} \eta_1 + f_2 \partial_{p_x} \eta_2 + \eta_1,$$

代入约束 $f_1 = 0, f_2 = 0$ ，那么给出第一组正则方程为

$$\dot{p}_x = -e\partial_x \varphi + \frac{eB_0}{2}\eta_2, \quad \dot{x} = \eta_1,$$

同理可以给出第二组曲面上的正则方程

$$\dot{p}_y = -e\partial_y \varphi - \frac{eB_0}{2}\eta_1, \quad \dot{y} = \eta_2.$$

再利用约束 $p_x = -eB_0 y/2$ 和 $p_y = eB_0 x/2$ ，就得到了

$$\begin{aligned} -\frac{eB_0}{2}\dot{y} &= -e\partial_x \varphi + \frac{eB_0}{2}\dot{y}, \\ \frac{eB_0}{2}\dot{x} &= -e\partial_y \varphi - \frac{eB_0}{2}\dot{x}, \end{aligned}$$

这正是正确的运动方程。

电磁场

也许上面一个系统是因为我们无视掉了二次的动能项而导致的约束，那么电磁场的约束更加本质而且难以避免，即 $F^{00} = 0$ 。

写出电磁场的 Lagrangian 密度

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} - 1 A_\mu J^\mu,$$

求其对偶场

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} = \frac{\partial \mathcal{L}}{\partial \partial_0 A_\mu} = -\frac{1}{8\pi} F^{\rho\nu} \frac{\partial}{\partial \partial_0 A_\mu} F_{\rho\nu} = -\frac{1}{4\pi} F^{0\mu},$$

所以有自然的约束 $\Pi^\mu = 0$ 。这就意味着接下来如果用

$$H = \int d^4 \mathbf{x} (\Pi^\mu \dot{A}_\mu - \mathcal{L})$$

来写出 Hamiltonian 时，无法用 Π^μ 和 A_μ 解出全部的 \dot{A}_μ 。

2.3 Dirac 括号

假设一些数学上可能出现的问题这里都不会出现，再假设系统的 Hamiltonian 不含时，即这是一个能量守恒系统。对于经典力学而言，广义动量和广义坐标之间的约束就是确定了一个相空间之中的曲面，如果（至少在局部）能够选取新的广义动量和广义坐标，那么我们就又得到了一个无约束系统，他的运动方程直接由新坐标和动量的正则方程确定，这样、就解决了约束问题。这节的主要内容基于 Toshihide Maskawa 和 Hideo Nakajima 的论文 *Singular Lagrangian and the Dirac-Faddeev Method— Existence Theorem of Constraints in 'Standard Form'*。

为描述 Hamilton 力学，辛几何是方便的。辛几何默认存在了一个闭的非退化 2-形式

$$\Omega = - \sum_{i=1}^n dp^i \wedge dq^i, \quad (2.8)$$

通过这个 2-形式，可以定义一个从光滑函数到矢量场的映射 $f \mapsto X_f$ 通过

$$\Omega(X_f, Y) = df(Y) = Yf, \quad (2.9)$$

其中 Y 是任意矢量场，而 f 是一个光滑函数。

在局部，选取一个坐标卡，2-形式 Ω 写作

$$\Omega = \frac{1}{2} \Omega_{ij} dx^i \wedge dx^j,$$

其中 $\Omega_{ij} = [x^i, x^j]$ ，不妨再通过 $\Omega^{ij} \Omega_{jk} = -\delta_k^i$ 定义 Ω^{ij} ，这样 X_f 就写作

$$X_f = \Omega^{ij} \partial_j f \partial_i.$$

对于这样一个矢量场，可以通过

$$\left. \frac{d}{dt} \right|_{t=0} g_f^t(x) = X_f(x)$$

给出他的相流（单参同胚映射） g_f^t 。

通过 2-形式 Ω ，Possion 括号就可以写成坐标无关的形式

$$[f, g]_P = \Omega(X_f, X_g) = -X_f g. \quad (2.10)$$

使用 Possion 括号，则

$$\Omega^{ij} = [x^i, x^j]_P, \quad X_f = -[f, x^i] \partial_i,$$

且 Possion 括号和矢量场之间的 Lie 括号的关系为

$$[X_f, X_g] = -X_{[f, g]_P}. \quad (2.11)$$

靠着相流的语言，可以证明

$$[F, G]_P(x) = \left. \frac{d}{dt} \right|_{t=0} F(g_G^t(x)),$$

如果 $[F, G]_P = 0$ ，则称 F 和 G 相互对合。从相流方面来看， F 和 G 相互对合就是指 F 是 G 的相流的首次积分，这个很容易从上面式子的右侧看出来，同时，因为 $[G, F]_P = -[F, G]_P = 0$ ，所以 G 也是 F 的相流的首次积分，因此称为相互。

现在有 N 个约束 $\{f^i\}$ ，他决定了一个曲面 $M = M(f)$ ，我们只考虑矩阵 $[f^i, f^j]_P$ 在 M 上的限制 $[f^i, f^j]_P|_M$ 常秩的情况，设他的秩为 m 。适当对约束进行线性组合，不妨将 $[f^i, f^j]_P|_M$ 看作块对角矩阵，

除去右下角的 $m \times m$ 矩阵之外的矩阵元都是 0，他的行列式不为 0，但因为他是反对称的，所以 m 必须是偶数，记 $m = 2s$ 以及 $r = N - 2s$ ，此时我们称这个约束 M 是 (r, s) 型约束。其中 r 指的是第一类约束，对应于 $[f^i, \eta]_{\text{P}}|_M = 0$ 的情况，其中 $\eta = a_i f^i$ ， a_i 是任意函数。而 s 指的是第二类约束，即除了第一类约束的约束。

现在假设有 (r, s) 型约束 $\{\phi^i, \psi^\alpha\}$ ，其中 $\{\phi^i\}$ 是第一类约束，而 $\{\psi^\alpha\}$ 是第二类约束，定义

$$V = \{c_i \phi^i + d_\alpha \psi^\alpha : c_i, d_\alpha \text{ are arbitrary functions.}\},$$

那么

$$A = \{g : [g, V]_{\text{P}} \subseteq V\}$$

是那些满足 $[g, V]_{\text{P}}|_M = 0$ 的函数的集合，他构成一个代数，加法乘法显然，至于 Poisson 括号来自于 Jacobi 恒等式。可以看到，第一类约束属于 $V_0 = A \cap V$ 。他是 A 的子代数且满足 $AV_0 \subseteq V_0$ 和 $[A, V_0]_{\text{P}} \subseteq V_0$ ，所以 V_0 是 A 的理想，可以顺便搞一个商代数 $A^* = A/V_0$ 。值得一提的是，在 V_0 中还能有第二类约束的高阶项。

设第一类约束 ϕ^i 以及 $f \in A$ ，则 $[f, \phi^i]_{\text{P}}|_M = 0$ 。反过来，如果有一个函数 g 满足 $g|_M = f|_M$ 且对所有的第一类约束都有 $[g, \phi^i]_{\text{P}}|_M = 0$ ，可以计算有

$$[g, \psi^\alpha]_{\text{P}}|_M = [g, \psi^\mu]_{\text{P}}|_M \delta_\mu^\alpha,$$

定义矩阵 $D_{\mu\nu}$ 是矩阵 $[\psi^\mu, \psi^\nu]_{\text{P}}$ 的逆矩阵，即满足 $D_{\mu\nu}[\psi^\nu, \psi^\alpha]_{\text{P}} = \delta_\mu^\alpha$ ，那么

$$[g, \psi^\alpha]_{\text{P}}|_M = ([g, \psi^\mu]_{\text{P}} D_{\mu\nu} [\psi^\nu, \psi^\alpha]_{\text{P}})|_M,$$

构造

$$g^* = g - [g, \psi^\mu]_{\text{P}} D_{\mu\nu} \psi^\nu,$$

容易验证 $[g^*, \psi^\mu]_{\text{P}}|_M = 0$ ，即 $g^* \in A$ ，以及 $g^*|_M = g|_M = f|_M$ ，所以 g^* 和 f 最多只是差了一个 V_0 中的元素，而在 A^* 是唯一确定的。这样我们就看到，所有满足 $[g, \phi^i]_{\text{P}}|_M = 0$ 中的函数 g 都唯一确定了 A^* 中的元素 g^* 。

下面首先要消除所有的第一类约束。

Theorem 1. 如果 M 是 (r, s) 型约束，那么存在一套正则坐标使得

$$M = M(q^1, \dots, q^{r+s}; p^{r+1}, \dots, p^{r+s}).$$

即 M 由方程 $q^1 = 0, \dots, q^{r+s} = 0$ 和 $p^{r+1} = 0, \dots, p^{r+s} = 0$ 确定。

这个定理在原则上将一大类约束简化到了某些正则坐标为 0 的情况。对于原本的约束 $M = M(\phi^1, \dots, \phi^r; \psi^1, \dots)$ 应该有

$$\phi^i = \sum_{j=1}^r A_j^i q^j + \sum_{k=1}^{2s} B_k^i \xi^k,$$

其中 $\{\xi^k\} = \{q^{r+1}, \dots, q^{r+s}, p^{r+1}, \dots, p^{r+s}\}$ ，以及 $B_k^i \in A$ 且 $\det(A_j^i) \neq 0$ 。

如果函数 f (当然最重要的就是 H) 满足 $[f, \phi^i]_{\text{P}}|_M = 0$ ，那么

$$[f, \phi^i]_{\text{P}}|_M = \sum_{j=1}^r A_j^i|_M [f, q^j]_{\text{P}}|_M = - \sum_{j=1}^r A_j^i|_M \left. \frac{\partial f}{\partial p^j} \right|_M,$$

所以

$$\left. \frac{\partial f}{\partial p^j} \right|_M = 0.$$

对所有 $j \leq r$ 成立, 而 p^j 不是约束, 那么

$$\left. \frac{\partial f|_M}{\partial p^j} = \left. \frac{\partial f}{\partial p^j} \right|_M = 0,$$

这就是说 $f|_M$ 不显含 p^j , 所以我们可以对 $j \leq r$ 取 $p^j = 0$. 设对 $j \leq r$ 选定 $p^j = 0$ 构成的约束为 M' , 那么系统就应该在约束 $M_0 = M' \cap M = M_0(q^1, \dots, q^{r+s}; p^1, \dots, p^{r+s})$ 内, 这是一个 $(0, r+s)$ 型约束. 这样我们就通过添加几个约束消去了全部的第一类约束. 对任意的 g 都可以通过

$$g^* = g - [g, \psi^\mu]_{\text{P}} D_{\mu\nu} \psi^\nu$$

构造一个 A^* 中的元素 g^* . 他对 $i \leq r+s$ 成立

$$g^*|_{M_0} = g|_{M_0}, \quad \left. \frac{\partial g^*}{\partial q^i} \right|_{M_0} = \left. \frac{\partial g^*}{\partial p^i} \right|_{M_0} = 0. \quad (2.12)$$

所以 g^* 就可以看成形式上无约束系统的物理量.

更一般地, 如果系统已经选定了一套坐标 $\{\chi^1, \dots, \chi^r, \eta^{r+1}, \dots, \eta^{2n}\}$, 如果在 M 只有 $[\chi^i, \phi^j]_{\text{P}}$ 可能不为零, 那么按照上面的思路

$$[f, \phi^i]_{\text{P}}|_M = \sum_{j=1}^r \left. \frac{\partial f}{\partial \chi^j} \right|_M [\chi^j, \phi^i]_{\text{P}}|_M = 0,$$

如果附加条件 $\det([\chi^j, \phi^i]_{\text{P}})|_M \neq 0$, 那么自然就有

$$\left. \frac{\partial f|_M}{\partial \chi^j} = \left. \frac{\partial f}{\partial \chi^j} \right|_M = 0,$$

于是可以选定 $\chi^j = 0$. 此时 $(0, r+s)$ 型约束写作 $M_0(\psi^1, \dots, \psi^{2(r+s)})$. 这种约束最常见的是某个正则坐标为常数的情况, 对于这种约束, 直接将他的对偶坐标取作零即可.

运动方程的确定靠 Poisson 括号, 正如前面知道的, 在没有约束的时候, 运动方程写作 $\dot{p} = [p, H]_{\text{P}}$. 因为现在处理是约束系统, 对物理量而言, 需要改成形式上没有约束的物理量, 这样我们应该处理的是 $[f^*, g^*]_{\text{P}}|_{M_0}$, 直接计算易得

$$[f^*, g^*]_{\text{P}}|_{M_0} = ([f, g]_{\text{P}} - [f, \psi^\mu]_{\text{P}} D_{\mu\nu} [\psi^\nu, g]_{\text{P}})|_{M_0},$$

定义所谓的 Dirac 括号 $[*, *]_{\text{D}}$ 如下

$$[f, g]_{\text{D}} = [f, g]_{\text{P}} - [f, \psi^\mu]_{\text{P}} D_{\mu\nu} [\psi^\nu, g]_{\text{P}}, \quad (2.13)$$

他满足

$$[f, g]_{\text{D}}|_{M_0} = [f^*, g^*]_{\text{P}}|_{M_0}.$$

如果采用约束的标准形式 $M_0(q^1, \dots, q^{r+s}; p^1, \dots, p^{r+s})$, 此时 Dirac 括号 $[*, *]_{\text{D}}$ 写作

$$[f, g]_{\text{D}}|_{M_0} = \sum_{i=r+s+1}^n \left(\frac{\partial f|_{M_0}}{\partial q^i} \frac{\partial g|_{M_0}}{\partial p^i} - \frac{\partial f|_{M_0}}{\partial p^i} \frac{\partial g|_{M_0}}{\partial q^i} \right), \quad (2.14)$$

可以看到这就是对于后 $n - (r+s)$ 组正则变量的 Poisson 括号. 结合(2.12)和(2.14), 我们所做的就是对约束系统选取了一套正则坐标, 使得他形式上变成了没有最后几组正则坐标的无约束系统一样.

总结一下 Hamilton 形式化一个约束系统的步骤，首先写出所有约束，然后加进几个新的约束去掉所有的第一类约束，最后计算出 Dirac 括号就可以得到正确的运动方程。

回到强匀强磁场中的电荷的例子，两个约束分别为

$$\begin{aligned} f_1 &= p_x + \frac{eB_0}{2}y, \\ f_2 &= p_y - \frac{eB_0}{2}x, \end{aligned}$$

因为 $[f_1, f_2]_P = eB_0 \neq 0$ ，所以他们都是第二类约束，直接的计算就得到了

$$D = \frac{1}{eB_0} \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$

然后运动方程应该为

$$\begin{aligned} \dot{p}_x &= [p_x, H]_D \\ &= [p_x, H]_P - [p_x, f_1]_P D_{12} [f_2, H]_P - [p_x, f_2]_P D_{21} [f_1, H]_P \\ &= -\frac{e}{2} \partial_x \varphi, \end{aligned}$$

代入约束 $p_x = -eB_0 y/2$ 就得到一个运动方程

$$B_0 \dot{y} = \partial_x \varphi.$$

剩下一个运动方程同理由 $\dot{p}_y = [p_y, H]_D$ 给出。

尽管这节的全部内容都是在经典力学框架内的，但是通过导数到变分算符的转变，可以将其完全形式地移动到场论中去。对于约束系统场的量子化，Poisson 括号与对易子的转变这里应该改成 Dirac 括号与对易子的改变。

2.4 带电粒子的运动

这节开始

Chapter 3

电磁波

上一章我们主要讨论了狭义相对论下力学的可能性，重点在于运动方程、力学量等等，却几乎没有实际求解过运动方程或是分析具体的电磁场。因此，接着的这几章的目的，就是给出这样的应用。而作为开始，我们就谈谈 Maxwell 理论力量的体现，也即是 Maxwell 本人的伟大预言，电磁波。

3.1 波动方程

无源的 Maxwell 方程组写作：

$$\begin{cases} \partial_{[\mu} F_{\nu\rho]} = 0, \\ \partial_\nu F^{\nu\mu} = 0, \end{cases} \quad (3.1)$$

对第二对 Maxwell 方程中的 $F^{\mu\nu}$ 应用定义写成 A^μ 的形式，

$$\partial_\nu \partial^\mu A^\nu - \partial_\nu \partial^\nu A^\mu = 0,$$

因为 A 不是唯一确定的，我们定下一个规范 $\partial_\mu A^\mu = 0$ ，这个规范称为 Lorenz 规范¹，因此第二对 Maxwell 方程写作

$$\partial_\nu \partial^\nu A^\mu = 0,$$

或写作

$$\frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} - \nabla^2 \varphi = 0, \quad \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = 0.$$

这就是波动方程，波速为 c 。下面我们要去求解这个方程。第一个方程我们可以固定规范，使得 $\varphi = 0$ ，所以我们暂时不需要考虑他，而此时 Lorenz 规范写作 $\nabla \cdot \mathbf{A} = 0$ 。这种规范有时候被称为 Coulomb 规范。

首先考虑平面波的情况，即对于 A 的任意分量，他们空间坐标中只依赖于 $x^1 = x$ ，此时波动方程写作

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \frac{\partial^2 \mathbf{A}}{\partial x^2} = 0,$$

他的通解写作

$$\mathbf{A} = \mathbf{f}(ct - x) + \mathbf{g}(ct + x),$$

前者表示沿着 x 正方向运动的波，后者表示沿着 x 轴负方向运动的波。

¹这个 Lorenz 是 Ludvig Valentin Lorenz(1829–1891)，是个丹麦人，和他名字差不多的，也就是前面说的 Lorentz 变换的 Lorentz 是 Hendrik Antoon Lorentz(1853–1928)，是一个荷兰人。

现在只考虑向 x 正方向运动的波 $\mathbf{A} = \mathbf{f}(ct - x)$,

$$\begin{aligned}\mathbf{B} &= \nabla \times \mathbf{A} = -\partial_1 f^3(ct - x) \hat{\mathbf{y}} + \partial_1 f^2(ct - x) \hat{\mathbf{z}} = -\hat{\mathbf{x}} \times \frac{\partial \mathbf{f}}{\partial \tau}(ct - x) = \dot{\mathbf{A}} \times \hat{\mathbf{x}}, \\ \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} = -\partial_t \mathbf{f}(ct - x) = -c \frac{\partial \mathbf{f}}{\partial \tau}(ct - x) = -c \dot{\mathbf{A}},\end{aligned}$$

其中头上戴帽的矢量为该方向的单位矢量, 所以对于平面波, 有 $\mathbf{B} = \hat{\mathbf{x}} \times \mathbf{E}/c$, 更一般地, 对于任意的单方向平面波, 有

$$\mathbf{B} = \frac{1}{c} \hat{\mathbf{n}} \times \mathbf{E}, \quad (3.2)$$

其中 $\hat{\mathbf{n}}$ 是运动方向的单位矢量。

对于可以允许的平面波解, 一般而言 $\hat{\mathbf{n}} \cdot \mathbf{E} \neq 0$, 这是因为其中存在非电磁波项。为此, 考察 Lorenz 规范 $\nabla \cdot \mathbf{A} = 0$, 对单方向平面波即 $\partial_1 A^1 = 0$, 所以对该分量, 波动方程写作

$$\frac{\partial^2 A^1}{\partial t^2} = 0,$$

即 E_1 是一个和时间无关的电场, 他不是电磁波的一部分, 由叠加原理, 我们可以从全部的电场中去掉它, 只留下电磁波项, 此时 $\hat{\mathbf{n}} \cdot \mathbf{E} = 0$. 结合(3.2), 对于平面波就有

$$\mathbf{E} = c\mathbf{B} \times \hat{\mathbf{n}}. \quad (3.3)$$

考虑单方向平面波的能流密度

$$\mathbf{S} = \frac{c^2}{4\pi} \mathbf{E} \times \mathbf{B} = \frac{c}{4\pi} \mathbf{E} \times (\hat{\mathbf{n}} \times \mathbf{E}) = \frac{c}{4\pi} E^2 \hat{\mathbf{n}} - \frac{c}{4\pi} (\hat{\mathbf{n}} \cdot \mathbf{E}) \mathbf{E},$$

可以看到有依赖于 $\hat{\mathbf{n}} \cdot \mathbf{E}$ 的一项, 去掉非电磁波项, 此时 \mathbf{E} 与 $\hat{\mathbf{n}}$ 垂直, 即 $\hat{\mathbf{n}} \cdot \mathbf{E} = 0$, 所以

$$\mathbf{S} = \frac{c}{4\pi} E^2 \hat{\mathbf{n}}.$$

与此同时, 因为 \mathbf{E} 与 $\hat{\mathbf{n}}$ 垂直, 所以 $|\hat{\mathbf{n}} \times \mathbf{E}| = |\mathbf{E}|$, 所以

$$B^2 = \frac{1}{c^2} E^2,$$

因此, 能量密度写作

$$w = \frac{c^2}{8\pi} \left(B^2 + \frac{1}{c^2} E^2 \right) = \frac{c^2}{8\pi} \frac{2}{c^2} E^2 = \frac{E^2}{4\pi} = \frac{|\mathbf{S}|}{c},$$

或者

$$\mathbf{S} = cw\hat{\mathbf{n}}.$$

上面说明了两点, 第一, 对平面波而言, 磁场和电场具有的能量是相同的, 其次, 对平面波而言, 动量密度的大小成立

$$w - \frac{S^2}{c^2} = 0,$$

这和零质量粒子的动量能量关系一模一样, 即

$$\frac{E^2}{c^2} - \mathbf{p}^2 = 0,$$

这暗示了光的电磁波诠释和光子诠释的统一性。

本节的最后，演示使用三维 Mawell 方程推出电场和磁场的波动方程，首先无源方程写作

$$\begin{cases} \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} = 0, \\ \nabla \cdot \mathbf{E} = 0, \\ \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \end{cases} \quad (3.4)$$

首先

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right),$$

左边有

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E},$$

右边有

$$\nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right) = -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) = -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2},$$

所以

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$

磁场同理

$$\nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0.$$

3.2 单色波

单色波就是指频率恒定的波，即满足 $\partial_t^2 \mathbf{A} = -\omega^2 \mathbf{A}$ ，其中 ω 是常数。波动方程于是可以改为

$$\nabla^2 \mathbf{A} + \frac{\omega^2}{c^2} \mathbf{A} = 0,$$

这个方程一般人叫他 Helmholtz 方程。考虑单色平面单方向波，此时方程写作

$$\frac{\partial^2 A^1}{\partial x^2} + \frac{\omega^2}{c^2} A^1 = 0,$$

解为

$$A^1 = B \cos \left(-\frac{\omega}{c} (ct - x) \right)$$

或者在任意的方向 $\hat{\mathbf{n}}$ ，并将其改写成指数函数的实部

$$\mathbf{A} = \text{Re} \left[\mathbf{A}_0 \exp \left(-i \frac{\omega}{c} (ct - \hat{\mathbf{n}} \cdot \mathbf{r}) \right) \right],$$

以后，只要进行线性运算，我们都可以把 Re 省去，定义波矢 $\mathbf{k} = \hat{\mathbf{n}}\omega/c$ ，最后，我们将解写作

$$\mathbf{A} = \mathbf{A}_0 \exp (i(\mathbf{k} \cdot \mathbf{r} - \omega t)),$$

如果再定义四维波矢 $k^\mu = (\mathbf{k}, \omega/c)$ ，那么

$$k^\mu x_\mu = \omega t - \mathbf{k} \cdot \mathbf{r},$$

所以解还可以写作

$$\mathbf{A} = \mathbf{A}_0 \exp (-ik^\mu x_\mu).$$

对于四维波矢，可以注意到的是恒等式

$$k^\mu k_\mu = \frac{\omega^2}{c^2} - \mathbf{k}^2 = \frac{\omega^2}{c^2} - \frac{\omega^2}{c^2} = 0,$$

而能动张量写作

$$T^{\mu\nu} = \frac{wc^2}{\omega^2} k^\mu k^\nu.$$

接着求

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} = \nabla \times (\mathbf{A}_0 \exp(i(\mathbf{k} \cdot \mathbf{r} - \omega t))) = i\mathbf{k} \times \mathbf{A}, \\ \mathbf{E} &= -ik\mathbf{A}, \end{aligned}$$

这里，我们要更细致地讨论单色波场的方向，比如电场

$$\mathbf{E} = \text{Re}(\mathbf{E}_0 \exp(-ik^\mu x_\mu)),$$

一般来说， \mathbf{E}_0 是一个复矢量，而 E_0^2 是一个复数，设其幅角为 -2α ，定义一个复矢量 \mathbf{b} 使得 $\mathbf{b}^2 = |\mathbf{E}_0^2|$ 是 E_0^2 的模长，这时候

$$\mathbf{E} = \text{Re}(\mathbf{b} \exp(-ik^\mu x_\mu - i\alpha)),$$

将 \mathbf{b} 分成两个实矢量之和 $\mathbf{b} = \mathbf{b}_1 + i\mathbf{b}_2$ ，那么，因为他的平方是实的，在必然有 $\mathbf{b}_1 \cdot \mathbf{b}_2 = 0$ ，两者正交，于是，选 \mathbf{b}_1 的方向为 y 轴， \mathbf{b}_2 的方向为 z 轴，那么

$$\begin{aligned} E_y &= b_1 \cos(k^\mu x_\mu + \alpha) \\ E_z &= \pm b_2 \sin(k^\mu x_\mu + \alpha), \end{aligned}$$

其中的 \pm 号说明是沿着 z 正负反向，此时分量满足

$$\frac{E_y^2}{b_1^2} + \frac{E_z^2}{b_2^2} = 1.$$

由此我们看到，在空间每一点，电场强度在垂直于波传播方向的一个平面内转动，端点绘出一个椭圆，这样的波被称为椭圆偏振波，如果 \pm 取正，则称其为右旋偏振，反之则左旋。

我们之所以谈论单色波，是因为我们希望对于任意的波，将其分解为不同频率的单色波的叠加。特别地，对于离散的谱，有周期的波，自然我们可以分解成单色波的 Fourier 级数：

$$f(t) = \sum_{n=-\infty}^{\infty} f_n e^{-in\omega_0 t},$$

其中 f 是任意表征波的量。他的分量由

$$f_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{in\omega_0 t} dt$$

确定，由于 $f(t)$ 是实数，所以系数有关系 $f_{-n} = f_n^*$ 。按照完全性关系，有

$$\bar{f}^2 = \sum_{n=-\infty}^{\infty} |f_n|^2.$$

另外的一类场可以展开为含连续分布的不同频率的 Fourier 积分，

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a(\omega) e^{i\omega t} d\omega,$$

他的系数有

$$a(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.$$

如果 $f(t)$ 是实数，则 $a(-\omega) = a(\omega)^*$ 。类似地，完全性关系给出

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |a(\omega)|^2 d\omega.$$

3.3 场的本征振动

我们来考虑空间一个有限体积内的电磁场，不妨假设是一个长方体，边长为 L_x, L_y, L_z ，此时我们可以把场展开成 Fourier 级数

$$\mathbf{A} = \sum_{\mathbf{k}} \mathbf{A}_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}),$$

其中

$$\mathbf{k} = \left(\frac{2\pi n_x}{L_x}, \frac{2\pi n_y}{L_y}, \frac{2\pi n_z}{L_z} \right)$$

而 n_x, n_y, n_z 都是整数。

因为 \mathbf{A} 是实的，所以 $\mathbf{A}_{-\mathbf{k}} = \mathbf{A}_{\mathbf{k}}^*$ ，从 $\nabla \cdot \mathbf{A} = 0$ 的规范，则

$$\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}} = 0,$$

这就是说 $\mathbf{A}_{\mathbf{k}}$ 和 \mathbf{k} 相互垂直。将分解代入波动方程，则可以得到分量的方程

$$\ddot{\mathbf{A}}_{\mathbf{k}} + c^2 \mathbf{k}^2 \mathbf{A}_{\mathbf{k}} = 0.$$

若边长足够大，则对于 \mathbf{k} 的小改变量，我们都可以考察他 \mathbf{n} 的改变量（即 \mathbf{k} 的小改变量内的可能值 \mathbf{n} 的数目），因此，

$$\Delta \mathbf{n} = \frac{1}{2\pi} \mathbf{L} \cdot \Delta \mathbf{k}$$

于是在 $\Delta \mathbf{k}$ 内的 \mathbf{k} 的可能值的总数为

$$\Delta n = \Delta n_1 \Delta n_2 \Delta n_3 = \frac{V}{(2\pi)^3} \Delta k_1 \Delta k_2 \Delta k_3,$$

将右边的 $\Delta k_1 \Delta k_2 \Delta k_3$ 变到球坐标

$$\Delta k_1 \Delta k_2 \Delta k_3 = \mathbf{k}^2 \Delta k \Delta \Omega,$$

其中 $\Delta k = |\Delta \mathbf{k}|$ ，而 $\Delta \Omega$ 为立体角元，对于任意方向，则以 4π 代之，那么在 Δk 内而方向任意的 $\Delta \mathbf{k}$ 的总数就是

$$\Delta n = \frac{V}{4\pi^2} \mathbf{k}^2 \Delta k.$$

来看看 Lagrangian,

$$L_t = -\frac{c^2}{16\pi} \int dV F_{\mu\nu} F^{\mu\nu} = \frac{c^2}{8\pi} \int dV \left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2 \right),$$

选取规范 $\varphi = 0, \nabla \cdot \mathbf{A} = 0$ ，则

$$L_t = \frac{1}{8\pi} \int dV ((\partial_t \mathbf{A})^2 - c^2 (\nabla \times \mathbf{A})^2),$$

带入 \mathbf{A} 的 Fourier 展开

$$L_t = \frac{1}{8\pi} \sum_{\mathbf{k}, \mathbf{k}'} \int dV e^{i(\mathbf{k} + \mathbf{k}') \cdot \mathbf{r}} \left(\dot{\mathbf{A}}_{\mathbf{k}} \cdot \dot{\mathbf{A}}_{\mathbf{k}'} + c^2 (\mathbf{k} \times \mathbf{A}_{\mathbf{k}}) \cdot (\mathbf{k}' \times \mathbf{A}_{\mathbf{k}'}) \right),$$

注意到 $\int dV \exp(i(\mathbf{k} + \mathbf{k}') \cdot \mathbf{r}) = V \delta_{\mathbf{k}, -\mathbf{k}'}$ ，所以

$$L_t = \frac{1}{8\pi} \sum_{\mathbf{k}} \left(\dot{\mathbf{A}}_{\mathbf{k}} \cdot \dot{\mathbf{A}}_{-\mathbf{k}} - c^2 (\mathbf{k} \times \mathbf{A}_{\mathbf{k}}) \cdot (\mathbf{k} \times \mathbf{A}_{-\mathbf{k}}) \right),$$

注意到 $\mathbf{A}_{-\mathbf{k}} = \mathbf{A}_{\mathbf{k}}^*$, 并打开三重积就可以得到

$$L_t = \frac{1}{8\pi} \sum_{\mathbf{k}} \left(\left| \dot{\mathbf{A}}_{\mathbf{k}} \right|^2 - c^2 (\mathbf{k}^2 |\mathbf{A}_{\mathbf{k}}|^2 - |\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}|^2) \right).$$

如若 $\mathbf{A}_{\mathbf{k}}$ 还是实的, 则 Lagrangian 可以写作

$$L_t = \frac{1}{8\pi} \sum_{\mathbf{k}} \left(\dot{A}_{\mathbf{k}}^2 - \lambda_{\mathbf{k}} A_{\mathbf{k}}^2 \right) = \frac{1}{8\pi} \sum_{\mathbf{k}} \sum_{i=1}^3 \left(\dot{A}_{\mathbf{k}i}^2 - \lambda_{\mathbf{k}} A_{\mathbf{k}i}^2 \right),$$

将 $A_{\mathbf{k}i}$ 看作广义坐标, 可以看到这正是谐振子的 Lagrangian, 此时 Lagrange 方程就给出了分量所需要满足的演化方程。

3.4 部分偏振波

每一个单色波都是偏振的, 实际上, 我们平时遇到的都是近似的单色波, 即频率范围很小的波, 假设 ω 是一个中间频率, 那么电场可以写作

$$\mathbf{E}_0(t) e^{-i\omega t},$$

这里的复振幅 $\mathbf{E}_0(t)$ 一个时间的缓变函数, 如果是单色波, 则就是常数了。 $\mathbf{E}_0(t)$ 决定波的偏振, 因此偏振也将随时间变化, 这样的波被称为部分偏振波。

3.5 光线

平面波那里已经提到过了, 电磁波和零质量粒子的力学关系式是类似的, 这里更是会强化这一点。

光的传播用光线处理, 可以看作波长趋向于 0 的情况, 或者说波数很大。如果我们把场的形式写作 $a \exp(i\psi)$, 这就是说相位 ψ 就很大。我们在局部用单色波代替 $a \exp(i\psi)$, 对足够小的时空间隔 (把一个端点取做零点), 把相位展开到一阶:

$$\psi = \psi_0 + \partial_i \psi x^i + \partial_t \psi t,$$

对比单色波的相位 $k_\mu x^\mu$, 于是得到 $\partial_\mu \psi = -k_\mu$, 对于电磁波的波数,

$$0 = k^\mu k_\mu = \partial_\mu \psi \partial^\mu \psi,$$

这就是程函方程。

这些材料就足够我们类比经典力学了, 作用量 S 和动量的关系 $\partial_\mu S = -p_\mu$ 以及动量之间的联系 $\partial_\mu S \partial^\mu S = p_\mu p^\mu = m^2 c^2$, 如果是零质量的粒子就有 $\partial_\mu S \partial^\mu S = 0$, 这就是说, 我们可以做出这样的类比 $\psi \sim S$ 以及 $k \sim p$ 。

回到程函方程, 对固定频率的波, 程函写作 $\psi = \omega t + \psi_0(r)$, 此时的程函方程为 $|\nabla \psi_0|^2 = \omega^2/c^2$. 因为光线在每一点都垂直于相应的波面, 所以光线的方向由 $\nabla \psi_0$ 确定, 换言之, 由 k 的空间分量完全确定。

对固定频率的波, 由于类比 $k \sim p$, 对应于能量守恒的体系, 则我们有 Maupertuis 原理,

$$\delta S = \delta \int d\mathbf{l} \cdot \mathbf{p} = 0,$$

于是对于几何光学, 我们有

$$\delta \psi = \delta \int d\mathbf{l} \cdot \mathbf{k} = \delta \int d\mathbf{l} \cdot n\omega/c = \frac{\omega}{c} \delta \int d\mathbf{l} = 0.$$

这就是 Fermat 原理。

Chapter 4

恒定电磁场

再写一遍 Maxwell 方程组

$$\begin{cases} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{B} &= \frac{1}{c^2} \left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi\mathbf{j} \right), \end{cases}$$

上一章处理的是 $\rho = 0, \mathbf{j} = 0$ 的情况，这章近似可以认为处理的是两个时间偏导数为 0 的情况，这就是所谓的恒定电磁场（不随时间变化的电磁场），方程写作

$$\begin{cases} \nabla \times \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{B} &= \frac{4\pi\mathbf{j}}{c^2}. \end{cases} \quad (4.1)$$

由于大家很熟悉 SI 制度，我们指出 Gauss 制往 SI 制度转化的一种方法：在 SI 制下观察式子的量纲，在缺少真空电容率 ϵ_0 量纲¹处乘以一个 $4\pi\epsilon_0$ 即可。

4.1 Coulomb 定律

这节暂时假设 $\mathbf{A} = 0$ ，也就是 $\mathbf{B} = 0$ ，此时只留下两个方程

$$\begin{cases} \nabla \times \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, \end{cases} \quad (4.2)$$

注意我们的定义， $\mathbf{E} = -\nabla\varphi$ ，则第一个方程自然满足，第二个方程写作

$$\Delta\varphi = \nabla^2\varphi = -4\pi\rho.$$

如果是无源的静电场， $\Delta\varphi = 0$ ，因此从这个方程可以看出，电场的势能在任何地方都不能有极大值或者极小值，否则 φ 各方向的一阶导数为 0，而他们的二阶导数有着相同的符号，于是 $\Delta\varphi \neq 0$ 。

¹就是 $\text{A}^2 \cdot \text{s}^4 \cdot \text{kg}^{-1} \cdot \text{m}^{-3}$ ，其中 A 是电流的单位 Ampere。

我们现在来解方程 $\Delta\varphi = -4\pi\rho$, 为此只要解出齐次方程的通解 $\Delta\varphi = 0$ 和一般方程的一个特解即可。而为了了解特解, 我们只要解出一个基本解就好, 即方程

$$\Delta\varphi = -4\pi e\delta^3(\mathbf{r}) \quad (4.3)$$

的特解就好, 其中 $\delta^3(\mathbf{x})$ 是三维 Dirac 函数, 而 e 是总电荷量。假如我们得到了一个基本解 $\psi(\mathbf{r})$, 此时

$$\begin{aligned} -4\pi\rho(x) &= -\int dV' 4\pi\rho(\mathbf{r}')\delta^3(\mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{e} \int dV' \rho(\mathbf{r}')\Delta'\psi(\mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{e} \int dV' \rho(\mathbf{r}')\Delta\psi(\mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{e} \Delta \int dV' \rho(\mathbf{r}')\psi(\mathbf{r}' - \mathbf{r}), \end{aligned}$$

其中 Δ' 是指求导对象为 \mathbf{r}' , 而 $\Delta'\psi(\mathbf{r}' - \mathbf{r}) = \Delta\psi(\mathbf{r}' - \mathbf{r})$ 是直接计算即可得到的。这样我们就找到了原方程 $\Delta\varphi = -4\pi\rho$ 的一个特解

$$\varphi(\mathbf{r}) = \frac{1}{e} \int dV' \rho(\mathbf{r}')\psi(\mathbf{r}' - \mathbf{r}).$$

因此我们来看基本解满足的方程(4.3), 这个方程的解就是熟知的

$$\varphi = \frac{e}{|\mathbf{r}|} = \frac{e}{r} = \frac{e}{\sqrt{x^2 + y^2 + z^2}}.$$

物理上这就是点电荷的电势, 电荷量为 e , 如果点电荷位于 \mathbf{r}_0 处, 则势写作

$$\varphi(\mathbf{r}) = \frac{e}{|\mathbf{r} - \mathbf{r}_0|}.$$

求个导数, 就得到了电场的表达式

$$\mathbf{E} = e \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|^3}.$$

对于由离散点电荷构成的系统, 他们的电势写作

$$\varphi(\mathbf{r}) = \sum_i \frac{e_i}{|\mathbf{r} - \mathbf{r}_i|}.$$

下面考察静电场的能量, 由前已知, 电磁场的能量密度为

$$w = \frac{c^2}{8\pi} \left(\mathbf{B}^2 + \frac{1}{c^2} \mathbf{E}^2 \right),$$

第一项为 0 因为我们考察的是纯静电场, 然后将 \mathbf{E} 改成 $-\nabla\varphi$, 此时

$$\begin{aligned} U &= -\frac{1}{8\pi} \int dV \mathbf{E} \cdot \nabla\varphi \\ &= -\frac{1}{8\pi} \int dV \nabla \cdot (\varphi\mathbf{E}) + \frac{1}{8\pi} \int dV \varphi \nabla \cdot \mathbf{E}, \\ &= \frac{1}{2} \int dV \varphi\rho. \end{aligned}$$

对于离散情况

$$U = \frac{1}{2} \sum_a e_a \varphi_a.$$

其中 φ_a 是指全部电荷所产生的电势在 e_a 所处位置的值。这样，电场的能量从纯电场的表示写成了势能和电荷表示。

若我们把这个公式应用到点电荷，此时能量写作 $e\varphi/2$ ，这个能量被理解为一种自能，这个电势 φ 是这个电荷自己产生的。可是，注意到点电荷的电势为 e/r ，当 $r \rightarrow 0$ 的时候，这个自能将要趋向于无穷，能量就是质量，所以也就有无穷大质量。你说这怎么可能，你一个电子怎么就有了无穷大质量呢？所以我们认为呐，这个是经典电动力学框架的漏洞。为了让这个漏洞不存在于这个理论框架种种，我们可以限制这个理论的最小尺度，这个时候自能大约是 mc^2 的量级，即

$$R \sim \frac{e^2}{mc^2}.$$

现在把自能从能量中去掉

$$U' = \frac{1}{2} \sum_a e_a \varphi'_a,$$

其中

$$\varphi'_a = \sum_{b \neq a} \frac{e_b}{|\mathbf{r}_a - \mathbf{r}_b|} = \sum_{b \neq a} \frac{e_b}{r_{ab}}.$$

所以

$$U' = \frac{1}{2} \sum_{b \neq a} \frac{e_a e_b}{r_{ab}} = \frac{1}{2} \sum_{b \neq a} U'_{ab},$$

其中 $U'_{ab} = e_a e_b / r_{ab}$ 是两个点电荷的相互作用能量。

4.2 匀速运动的电荷产生的场

因为我们知道四维势能满足 Lorentz 变换，所以只要我们在那个 O 系观测沿 $z = x^3$ 轴以速度 v 运动的 O' 内的静电场，按照一般的方法，任意的四维矢量 A^μ 按照如下变换：

$$\begin{pmatrix} A^1 \\ A^2 \\ A^3 \\ A^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\gamma & \beta/\gamma \\ 0 & 0 & \beta/\gamma & 1/\gamma \end{pmatrix} \begin{pmatrix} A'^1 \\ A'^2 \\ A'^3 \\ A'^0 \end{pmatrix},$$

其中 $\beta = v/c$, $\gamma = \sqrt{1 - \beta^2}$. 特别地，空间坐标有

$$x = x', \quad y = y', \quad z = \frac{z' + \beta ct'}{\sqrt{1 - \beta^2}} = \frac{z' + vt'}{\sqrt{1 - \beta^2}},$$

四维势有

$$A^1 = A^2 = 0, \quad A^3 = \frac{v\varphi'/c^2}{\sqrt{1 - \beta^2}}, \quad \varphi = \frac{\varphi'}{\sqrt{1 - \beta^2}}.$$

或者对任意方向的 \mathbf{v} ，此时

$$\mathbf{A} = \frac{\varphi' \mathbf{v}}{c^2 \sqrt{1 - \beta^2}}.$$

由空间坐标的变换

$$r^2 = x'^2 + y'^2 + \frac{(z' + vt')^2}{1 - \beta^2},$$

令 $v \rightarrow -v$ 就得到了

$$r'^2 = x^2 + y^2 + \frac{(z - vt)^2}{1 - \beta^2},$$

所以在 O' 系中点电荷静电场有

$$\varphi' = \frac{e}{r'},$$

变到 O 系中

$$\varphi = \frac{\varphi'}{\sqrt{1-\beta^2}} = \frac{e}{r' \sqrt{1-\beta^2}} = \frac{e}{r_*},$$

其中

$$r_* = r' \sqrt{1-\beta^2} = \sqrt{(x^2 + y^2)(1-\beta^2) + (z-vt)^2}.$$

当然我们前面也算过

$$\begin{aligned} \mathbf{E} &= \left(\frac{E'_x + \beta c B'_y}{\sqrt{1-\beta^2}}, \frac{E'_y - \beta c B'_x}{\sqrt{1-\beta^2}}, E'_z \right), \\ \mathbf{B} &= \left(\frac{B'_x - E'_y \beta / c}{\sqrt{1-\beta^2}}, \frac{B'_y + E'_x \beta / c}{\sqrt{1-\beta^2}}, B'_z \right), \end{aligned}$$

由于在 O' 系中 $\mathbf{B}' = 0$, 所以

$$\begin{aligned} \mathbf{E} &= \left(\frac{E'_x}{\sqrt{1-\beta^2}}, \frac{E'_y}{\sqrt{1-\beta^2}}, E'_z \right), \\ \mathbf{B} &= \left(\frac{-E'_y \beta / c}{\sqrt{1-\beta^2}}, \frac{E'_x \beta / c}{\sqrt{1-\beta^2}}, 0 \right), \end{aligned}$$

代入 $\mathbf{E}' = e\mathbf{r}'/r'^3$, 就得到了

$$\mathbf{E} = (1-\beta^2) \frac{e\mathbf{r}}{r_*^3},$$

其中 \mathbf{r} 其实是指离运动电荷 \mathbf{r} 处的坐标, 是以点电荷为中心来看的, 在 O 系中来看应该是 $\mathbf{r} - \mathbf{v}t$. 而容易验证

$$\mathbf{B} = \frac{1}{c^2} \mathbf{v} \times \mathbf{E},$$

因此, 我们下面更关注 \mathbf{E} .

如果对任意的 \mathbf{v} 来求 \mathbf{E} , 将 r_* 改写成坐标无关的形式更为方便

$$r_* = \sqrt{\mathbf{r}^2 - \beta^2(\mathbf{r}^2 - (\mathbf{r} \cdot \hat{\mathbf{v}})^2)} = |\mathbf{r}| \sqrt{1 - \beta^2(1 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{v}})^2)},$$

其中 $\beta^2 = \mathbf{v}^2/c^2$. 这时

$$\mathbf{E} = (1-\beta^2) \frac{e\mathbf{r}}{r_*^3}$$

就对任意的 \mathbf{v} 都成立了。

现在如果 \mathbf{r} 垂直于 \mathbf{v} , 则 $\hat{\mathbf{r}} \cdot \hat{\mathbf{v}} = 0$. 此时垂直于运动方向的电场分量为

$$E_{\perp} = \frac{1}{\sqrt{1-\beta^2}} \frac{e}{|\mathbf{r}|^2},$$

如果 \mathbf{r} 平行于 \mathbf{v} , 则 $(\hat{\mathbf{r}} \cdot \hat{\mathbf{v}})^2 = 1$, 此时平行于运动方向的电场分量为

$$E_{\parallel} = (1-\beta^2) \frac{e}{|\mathbf{r}|^2}.$$

可以看到, 当速度增加时, $1-\beta^2 \rightarrow 0$, 所以 E_{\perp} 增大, 而 E_{\parallel} 减小, 则点电荷电场从静止的一个正球到匀速运动时在运动方向被压扁了。

在速度不大的时候, 用 $\mathbf{E} = e\mathbf{r}/|\mathbf{r}|^3$ 代替真实电场, 则

$$\mathbf{B} = \frac{1}{c^2} \frac{e}{|\mathbf{r}|^3} \mathbf{v} \times \mathbf{r}.$$

4.3 Biot-Savart 定律

这节处理剩下两个方程，此时先假设 $\mathbf{E} = 0$ ，方程写作

$$\begin{cases} \nabla \cdot \mathbf{B} = 0, \\ \nabla \times \mathbf{B} = 4\pi \mathbf{j}/c^2. \end{cases} \quad (4.4)$$

$\mathbf{E} = 0$ 其实是个很愚蠢的假设，因为如果有电流密度，就意味着有运动的电荷，有运动电荷就不可能每时每刻每个地方都有 $\mathbf{E} = 0$ ，可是，后面我们将看到，在一类很广义的情况下，磁感应强度的平均值也满足这两个方程，则数学上，他们有相同的解。无论如何，即使是沉迷于解方程的恶趣味，讨论物理问题前解决这两个方程至少在数学上是有意义的，等到解完方程后，我们再回到物理上来。

设 $\mathbf{B} = \nabla \times \mathbf{A}$ ，则第一个方程是自然的，那么结合第二个方程，就得到了

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \Delta \mathbf{A} = \frac{4\pi \mathbf{j}}{c^2},$$

选取规范使得 $\nabla \cdot \mathbf{A} = 0$ ，此时

$$\Delta \mathbf{A} = -\frac{4\pi \mathbf{j}}{c^2}.$$

前面我们已经看到 $\Delta \varphi = -4\pi \rho$ 的解（将点电荷的基本解带到一般势能的解中）是

$$\varphi(\mathbf{r}) = \int dV' \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|},$$

因为三维矢势各个分量（直角坐标内）依然满足这个方程，所以就得到了

$$\mathbf{A}(\mathbf{r}) = \frac{1}{c^2} \int dV' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}.$$

对于离散电荷体系，我们有

$$\mathbf{A}(\mathbf{r}) = \frac{1}{c^2} \sum_a \frac{e_a \mathbf{v}_a}{|\mathbf{r} - \mathbf{r}_a|}.$$

下面我们求磁感应强度

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{1}{c^2} \int dV' \nabla \times \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} = \frac{1}{c^2} \int dV' \nabla \left(\frac{1}{|\mathbf{r}' - \mathbf{r}|} \right) \times \mathbf{j}(\mathbf{r}'),$$

计算出那个散度，最后就得到了

$$\mathbf{B} = -\frac{1}{c^2} \int dV' \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \times \mathbf{j}(\mathbf{r}') = \frac{1}{c^2} \int dV' \mathbf{j}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3},$$

这就是俗称的 Biot-Savart 定律。

前面指出过，因为存在运动电荷，所以不可能成立 $\mathbf{E} = 0$ 甚至 $\partial_t \mathbf{E} = 0$ ，所以磁场所满足的方程只能是

$$\begin{cases} \nabla \cdot \mathbf{B} = 0, \\ \nabla \times \mathbf{B} = \frac{1}{c^2} \left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi \mathbf{j} \right), \end{cases}$$

将两边求时间平均，则

$$\begin{cases} \nabla \cdot \bar{\mathbf{B}} = 0, \\ \nabla \times \bar{\mathbf{B}} = \frac{1}{c^2} \left(\frac{\partial \bar{\mathbf{E}}}{\partial t} + 4\pi \bar{\mathbf{j}} \right), \end{cases}$$

退而求次，现在我们只要让

$$\frac{\partial \bar{\mathbf{E}}}{\partial t} = 0,$$

就对 $\overline{\mathbf{B}}$ 得到类似的方程了。将时间平均写出来

$$\frac{\partial \overline{\mathbf{E}}}{\partial t} = \frac{1}{2T} \int_{-T}^T dt \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{2T} (\mathbf{E}(T) - \mathbf{E}(-T)),$$

如果场 \mathbf{E} 是有界的, 则随着 $T \rightarrow \infty$, 则自然得到了上面希望的 $\frac{\partial \overline{\mathbf{E}}}{\partial t} = 0$. 这是一类很广义的情况, 而且 $\overline{\mathbf{B}}$ 是一个物理上可测量的量。最后, 真正的 Biot-Savart 定律应该写作:

$$\overline{\mathbf{B}} = \frac{1}{c^2} \int dV' \overline{\mathbf{j}(\mathbf{r}')} \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}.$$

4.4 多极矩

实际过程中, 我们遇到的不是点电荷, 而是一个电荷体系, 比如带电物体, 当离他们足够远的时候, 可以将其看做一个点电荷。这节, 我们希望将这个问题做得更加精确, 得到他的场更精确的近似解。

首先, 对离散电荷体系, 将坐标零点设在电荷体系中, 势能写作

$$\varphi = \sum_a \frac{e_a}{|\mathbf{r} - \mathbf{r}_a|},$$

当 $|\mathbf{r}|$ 足够大的时候, 任意的光滑函数 $f(\mathbf{r} - \mathbf{r}_a)$ 有近似

$$f(\mathbf{r} - \mathbf{r}_a) = f(\mathbf{r}) - (\mathbf{r}_a \cdot \nabla) f(\mathbf{r}),$$

所以

$$\varphi = \frac{1}{|\mathbf{r}|} \sum_a e_a - \sum_a e_a \mathbf{r}_a \cdot \nabla \frac{1}{|\mathbf{r}|},$$

此时第一项是点电荷项, 等价于一个总电荷为整个体系电荷量的点电荷的势, 第二项是更精确的项, 设

$$\mathbf{d} = \sum_a e_a \mathbf{r}_a,$$

我们称 \mathbf{d} 为这个体系的偶极矩, \mathbf{d} 的值和选取的坐标有关, 设我们选了一个新的零点, 则所有的位置矢量都加上了一个常矢量 \mathbf{a} , 即

$$\mathbf{d}' = \sum_a e_a (\mathbf{r}_a + \mathbf{a}) = \mathbf{d} + \mathbf{a} \sum_a e_a,$$

在一种极其特殊的情况下, 即系统的总电荷为 0 的时候, 此时这个体系的偶极矩的值和我们选取的坐标无关。

此外, 如果有一个轴对称的电荷分布, 那么偶极矩为 0, 这点可以从

$$\mathbf{d} = \sum_a e_a \mathbf{r}_a = - \sum_a e_a \mathbf{r}_a$$

看出。

只考虑偶极矩项的势能

$$\varphi = \mathbf{d} \cdot \nabla \frac{1}{|\mathbf{r}|},$$

求其负梯度

$$\mathbf{E} = -\nabla \left(\mathbf{d} \cdot \nabla \frac{1}{|\mathbf{r}|} \right) = \frac{3(\hat{\mathbf{r}} \cdot \mathbf{d})\hat{\mathbf{r}} - \mathbf{d}}{|\mathbf{r}|^3},$$

这个场围绕 \mathbf{d} 的方向具有轴对称性。

现在, 我们试图利用 Taylor 公式来展开势能成更精确的形式

$$f(\mathbf{r} - \mathbf{r}_a) = \sum_{k=0}^n (-\mathbf{r}_a \cdot \nabla)^k f(\mathbf{r}),$$

这当然是可以的。但是, 在球谐函数理论中我们知道著名的公式

$$\frac{1}{|\mathbf{r} - \mathbf{r}_a|} = \sum_{l=0}^{\infty} \frac{|\mathbf{r}_a|^l}{|\mathbf{r}|^{l+1}} P_l(\hat{\mathbf{r}}_a \cdot \hat{\mathbf{r}}),$$

其中 $|\mathbf{r}| > |\mathbf{r}_a|$ 而 P_l 是 l 阶 Legendre 多项式, 全部累加起来

$$\varphi = \sum_a \frac{e_a}{|\mathbf{r} - \mathbf{r}_a|} = \sum_{l=0}^{\infty} \frac{1}{|\mathbf{r}|^{l+1}} \sum_a e_a |\mathbf{r}_a|^l P_l(\hat{\mathbf{r}}_a \cdot \hat{\mathbf{r}}),$$

如果是连续电荷体系

$$\varphi = \sum_{l=0}^{\infty} \frac{1}{|\mathbf{r}|^{l+1}} \int dV' \rho(\mathbf{x}') |\mathbf{r}'|^l P_l(\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}).$$

对于 l 级展开, 称之为 2^l 极矩。

4.5 磁矩

类似于考察电多极矩, 这节我们考察离开一个稳定运动的电荷体系很远处的平均磁场。

已知矢势有

$$\bar{\mathbf{A}}(\mathbf{r}) = \frac{1}{c^2} \sum_a \frac{e_a \mathbf{v}_a}{|\mathbf{r} - \mathbf{r}_a|},$$

然后展开到一阶

$$\bar{\mathbf{A}}(\mathbf{r}) = \frac{1}{c^2 |\mathbf{r}|} \sum_a e_a \bar{\mathbf{v}}_a - \frac{1}{c^2} \sum_a e_a \mathbf{v}_a \left(\overline{\mathbf{r}_a \cdot \nabla \frac{1}{|\mathbf{r}|}} \right),$$

因为 $\bar{\mathbf{v}}_a$ 是一个关于时间导数的量的平均, 而有限运动使得位置矢量有界, 因此 $\bar{\mathbf{v}}_a = 0$, 也就是第一项消失, 此时

$$\bar{\mathbf{A}}(\mathbf{r}) = -\frac{1}{c^2} \sum_a e_a \mathbf{v}_a \left(\overline{\mathbf{r}_a \cdot \nabla \frac{1}{|\mathbf{r}|}} \right) = \frac{1}{c^2 |\mathbf{r}|^3} \sum_a e_a \overline{\mathbf{v}_a (\mathbf{r}_a \cdot \mathbf{r})},$$

注意到 (特别留意, \mathbf{r} 现在固定, 是一个常量)

$$e_a \mathbf{v}_a (\mathbf{r}_a \cdot \mathbf{r}) = \frac{d}{dt} (e_a \mathbf{r}_a (\mathbf{r}_a \cdot \mathbf{r})) - e_a \mathbf{r}_a \frac{d}{dt} (\mathbf{r}_a \cdot \mathbf{r}),$$

在时间平均后

$$\overline{\mathbf{v}_a (\mathbf{r}_a \cdot \mathbf{r})} = -\overline{\mathbf{r}_a (\mathbf{v}_a \cdot \mathbf{r})},$$

为了后面的方便, 我们改写上式为

$$\overline{\mathbf{v}_a (\mathbf{r}_a \cdot \mathbf{r})} = \frac{1}{2} \left(\overline{\mathbf{v}_a (\mathbf{r}_a \cdot \mathbf{r})} - \overline{\mathbf{r}_a (\mathbf{v}_a \cdot \mathbf{r})} \right) = \frac{1}{2} \overline{(\mathbf{r}_a \times \mathbf{v}_a) \times \mathbf{r}},$$

这样就将矢势写作了

$$\bar{\mathbf{A}}(\mathbf{r}) = \frac{1}{2c^2 |\mathbf{r}|^3} \sum_a e_a \overline{(\mathbf{r}_a \times \mathbf{v}_a) \times \mathbf{r}}.$$

定义

$$\mathbf{m} = \frac{1}{2} \sum_a e_a \mathbf{r}_a \times \mathbf{v}_a,$$

称为磁矩，利用磁矩，可以将一级近似的矢势写作

$$\mathbf{A}(\mathbf{r}) = \frac{1}{c^2} \mathbf{m} \times \frac{\mathbf{r}}{|\mathbf{r}|^3} = \frac{1}{c^2} \nabla \frac{1}{|\mathbf{r}|} \times \mathbf{m}.$$

知道了矢势，就容易求得磁感应强度

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) = \frac{1}{c^2} \nabla \times \left(\nabla \frac{1}{|\mathbf{r}|} \times \mathbf{m} \right) = \frac{1}{c^2} (\mathbf{m} \cdot \nabla) \nabla \frac{1}{|\mathbf{r}|} + \frac{4\pi}{c^2} \delta^3(\mathbf{r}) \mathbf{m},$$

因为讨论的是足够远处的磁场，所以第二项可以去掉，第一项具体写作

$$\mathbf{B}(\mathbf{r}) = \frac{1}{c^2} (\mathbf{m} \cdot \nabla) \nabla \frac{1}{|\mathbf{r}|} = \frac{1}{c^2} \frac{3(\hat{\mathbf{r}} \cdot \mathbf{m})\hat{\mathbf{r}} - \mathbf{m}}{|\mathbf{r}|^3},$$

和电偶极子的公式几乎一模一样。

注意到，考察足够远处的磁场，他的第零阶近似始终为 0，而第一阶近似和电偶极子的电场一模一样。那么我们是否可以依据偶极子和磁矩的相似性做出如下猜想：磁场也存在类似于电荷的磁荷，但是他是一正一负成对出现的，这样才使得第零阶近似始终为 0。

这样的猜想是有理由的，对错先不说，甚至在实用上，认为存在磁荷的模型在求解一些物理问题时候也相当有用，当然这样做的根据是公式上的相似性，而不涉及本质上的相似性。

作为补充，这里求一下处于一个平面上的均匀电流环的磁矩，为此首先将磁矩改成连续体系的形式

$$\mathbf{m} = \frac{1}{2} \int dV' \mathbf{r}' \times \mathbf{j}(\mathbf{r}').$$

电流是一个标量，即在空间中考虑一个曲面 S ，定义通过他的电流强度（简称电流）为

$$I(S) = \int_S d\sigma \cdot \mathbf{j}(\mathbf{r}).$$

对于一个环面，在每一个 φ 确定的截面，都可以定义一个 I ，如果 $I(\varphi)$ 是一个常数，那么就称呼其为恒定电流的。在生活中，我们遇到的恒定电流系统，往往可能截面面积极小，比如电线，此时 $dV' \mathbf{j}(\mathbf{r}')$ 近似于 $I d\mathbf{l}'$ ，在截面面积趋向于 0 的时候

$$\mathbf{m} = \frac{I}{2} \int_{\partial S} \mathbf{r}' \times d\mathbf{l}'.$$

考虑 S 是平面上的均匀电流环的面积矢量，则

$$\mathbf{m} = -\frac{I}{2} \int_{\partial S} d\mathbf{l}' \times \mathbf{r}' = -\frac{I}{2} \int_S (d\boldsymbol{\sigma}' \times \nabla') \times \mathbf{r}' = -\frac{I}{2} \int_S d\boldsymbol{\sigma}' (\mathbf{k} \times \nabla') \times \mathbf{r}',$$

其中 \mathbf{k} 是面积矢量方向的单位矢量，不妨取做第三根轴，很容易计算那个积分为 $-2\mathbf{S}$ ，所以

$$\mathbf{m} = I\mathbf{S}.$$

4.6 外场中的电荷体系

假设给定的外场是 $\mathbf{E}(\mathbf{r})$ 和 $\mathbf{B}(\mathbf{r})$ ，那么对一个离散电荷体系来说，由 Lorentz 力公式，

$$\mathbf{F} = \sum_a e_a \mathbf{E}(\mathbf{r}_a) + \sum_a e_a \mathbf{v}_a \times \mathbf{B}(\mathbf{r}_a),$$

对于连续体系

$$\mathbf{F} = \int dV' \rho(\mathbf{r}') \mathbf{E}(\mathbf{r}') + \int dV' \mathbf{j}(\mathbf{r}') \times \mathbf{B}(\mathbf{r}').$$

Chapter 5

辐射

5.1 推迟势

含源电磁场的方程为

$$\partial_\nu \partial^\nu A^\mu - \partial^\mu \partial_\nu A^\nu = \partial_\nu F^{\nu\mu} = \frac{4\pi}{c^2} j^\mu,$$

在 Lorenz 规范 $\partial_\nu A^\nu = 0$ 下, 得到了方程

$$\frac{1}{c^2} \frac{\partial^2 A^\mu}{\partial t^2} - \nabla^2 A^\mu = \partial_\nu \partial^\nu A^\mu = \frac{4\pi}{c^2} j^\mu,$$

这是一个非齐次方程。对于非齐次方程, 他的通解是由非齐次方程本身的一个特解和他对应的齐次方程的通解叠加而成, 而他对应的齐次方程就是前面我们看到的电磁波方程。

为了求得特解, 我们采用基本解方法, 去解方程

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = \frac{4\pi}{c} \delta(t) \delta^3(\mathbf{r}),$$

或者

$$\frac{\partial^2 \psi}{\partial \tau^2} - \nabla^2 \psi = 4\pi \delta(\tau) \delta^3(\mathbf{r}),$$

其中 $\tau = ct$ 。两边对 τ 做 Fourier 变换, 设 τ 的共轭变量是 ω , 则

$$\omega^2 \hat{\psi} + \nabla^2 \hat{\psi} = -2\sqrt{2\pi} \delta^3(\mathbf{r}),$$

其中 $\hat{\psi}$ 是变换后的函数。

对于我们的真空无界情形, 必然是球对称的, 所以我们也去求球对称的特解。于是对 ∇^2 展开只剩下了径部,

$$\omega^2 \hat{\psi} + \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \hat{\psi}}{\partial r} \right) = -2\sqrt{2\pi} \delta^3(\mathbf{r}),$$

在 $r \neq 0$ 的地方, 这个方程就是普通的

$$r^2 \omega^2 \hat{\psi} + \frac{\partial}{\partial r} \left(r^2 \frac{\partial \hat{\psi}}{\partial r} \right) = 0,$$

令 $\hat{\psi} = \chi/r$, 则 $\partial^2 \chi / \partial r^2 + \omega^2 \chi = 0$, 该方程的解写作 $\chi = A \exp(i\omega r) + B \exp(-i\omega r)$, 所以

$$\hat{\psi} = A \frac{\exp(i\omega r)}{r} + B \frac{\exp(-i\omega r)}{r}.$$

现在将方程

$$\omega^2 \hat{\psi} + \nabla^2 \hat{\psi} = -2\sqrt{2\pi} \delta^3(\mathbf{r}),$$

两边在半径为 R 的球内积分, 则

$$\omega^2 \int_V dV \hat{\psi} + \int_V dV \nabla^2 \hat{\psi} = -2\sqrt{2\pi},$$

第一个积分写作

$$\int_V dV \hat{\psi} = 4\pi \int_0^R dr r (A \exp(i\omega r) + B \exp(-i\omega r)),$$

所以在 $R \rightarrow 0$ 的时候, 积分也趋向于 0, 所以在渐进过程中, 我们可以略去它。剩下的第二个积分使用 Stokes 定理

$$\int_{\partial V} d\boldsymbol{\sigma} \cdot \nabla \hat{\psi} = 4\pi R^2 \frac{\partial \hat{\psi}}{\partial r}(R),$$

在 $R \rightarrow 0$ 的时候, 代入已知的 $\hat{\psi}$ 形式, 可以计算得

$$\lim_{R \rightarrow 0} \int_{\partial V} d\boldsymbol{\sigma} \cdot \nabla \hat{\psi} = -4\pi (A + B) = -2\sqrt{2\pi},$$

可以解出 $A + B = 1/\sqrt{2\pi}$ 。因此可以看见, $\hat{\psi} = (A \exp(i\omega r) + B \exp(-i\omega r))/r$ 是在 $r = 0$ 处也成立的解, 所以他就是我们希望得到的球对称特解, 我们将其写作

$$\hat{\psi} = C \frac{\exp(i\omega r)}{\sqrt{2\pi r}} + (1 - C) \frac{\exp(-i\omega r)}{\sqrt{2\pi r}}.$$

其中 C 是待定常数。最后 Fourier 逆变换, 则基本解写作

$$\psi = C \frac{\delta(r + \tau)}{r} + (1 - C) \frac{\delta(r - \tau)}{r}.$$

回到原方程,

$$\begin{aligned} \frac{4\pi}{c^2} j^\mu(\mathbf{r}, t) &= \frac{4\pi}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \delta(\tau' - \tau) \delta^3(\mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \left(\frac{\partial^2 \psi}{\partial (\tau')^2} - (\nabla')^2 \psi \right) (\tau' - \tau, \mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \left(\frac{\partial^2 \psi}{\partial \tau^2} - \nabla^2 \psi \right) (\tau' - \tau, \mathbf{r}' - \mathbf{r}) \\ &= \left(\frac{\partial^2}{\partial \tau^2} - \nabla^2 \right) \frac{1}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \psi(\tau' - \tau, \mathbf{r}' - \mathbf{r}), \end{aligned}$$

所以

$$\begin{aligned} A^\mu &= \frac{1}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \psi(\tau' - \tau, \mathbf{r}' - \mathbf{r}) \\ &= \frac{1}{c^2} \int dV' d\tau' j^\mu(\mathbf{r}', \tau'/c) \left(C \frac{\delta(|\mathbf{r}' - \mathbf{r}| + \tau' - \tau)}{|\mathbf{r}' - \mathbf{r}|} + (1 - C) \frac{\delta(|\mathbf{r}' - \mathbf{r}| - \tau' + \tau)}{|\mathbf{r}' - \mathbf{r}|} \right) \\ &= \frac{1}{c^2} \int dV' \left(C \frac{j^\mu(\mathbf{r}', t - |\mathbf{r}' - \mathbf{r}|/c)}{|\mathbf{r}' - \mathbf{r}|} + (1 - C) \frac{j^\mu(\mathbf{r}', t + |\mathbf{r}' - \mathbf{r}|/c)}{|\mathbf{r}' - \mathbf{r}|} \right), \end{aligned}$$

记 $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ 以及对任意的函数 $f(t \pm R/c)$, 将其记作 $f_{t \pm R/c}$, 则

$$A^\mu(\mathbf{r}, t) = \frac{C}{c^2} \int dV' \frac{j_{t-R/c}^\mu(\mathbf{r}')}{R} + \frac{1-C}{c^2} \int dV' \frac{j_{t+R/c}^\mu(\mathbf{r}')}{R}.$$

剩下的，就是我们要确定常数 C 。

一般来说，电磁波的解写作

$$A^\mu(\mathbf{r}, t) = A_0^\mu + \frac{C}{c^2} \int dV' \frac{j_{t-R/c}^\mu(\mathbf{r}')}{R} + \frac{1-C}{c^2} \int dV' \frac{j_{t+R/c}^\mu(\mathbf{r}')}{R},$$

为了确定 A_0^μ 和 C ，我们只要知道一个初值就可以了。但实际上，我们并不处理这样的初始条件，作为代替，我们常常给出与电荷体系离得很远处在所有时刻的条件，就是说，辐射是从系统外入射到系统上的。与此对应，辐射与系统相互作用后建立的场和外场的差别只是系统发出的辐射，即 $A = A_0 + A_1$ ，其中 A_0 是外场而 A_1 是系统发出的辐射。这个辐射在远处看是波的形式，由系统向外（即 R 增加）传播。因此我们只能取 $t - R/c$ 的那个部分，此时 $C = 1$ ，否则 A^μ 在 t 时候的值将由在未来的分布 $j_{t+R/c}^\mu$ 确定，而这与随时间正向演化的传播过程显然是不符的，在这种情况下

$$A^\mu(\mathbf{r}, t) = A_0^\mu + \frac{1}{c^2} \int dV' \frac{j_{t-R/c}^\mu(\mathbf{r}')}{R} = A_0^\mu + A_{t-R/c}^\mu, \quad (5.1)$$

其中 A_0^μ 是外场，而 $A_{t-R/c}^\mu$ 被称为推迟势。这意味着对应一个电流/电荷分布，他需要时间 R/c 才能够对场有所影响，譬如，在真空中突然放入一个电荷，经过时间 t 后，电荷所产生的电场在 $R > ct$ 处依然为 0，这和经典情况的超距作用是不同的，他严格地遵守着相对论的基本原理，作用的传递速度是不可能超过光速。

5.2 Lienard-Wiechert 势

Lienard-Wiechert 势是一个运动电荷所产生的势能。对于这样一个运动电荷，如果假设我们能够观察到这个粒子，下面需要一个断言，即他的世界线和以我们作为观测者的下半光锥（由 $x_\mu x^\mu = 0$ 且 $x^0 < 0$ 的矢量构成的曲面）一定有唯一交点。

交点的存在性来自于在过去我们能够观察到这个粒子，因此他的世界线的一部分就在下半光锥内，而当 $t' = 0$ 的时候，他的世界点要么在 $\mathbf{r}(0) \neq 0$ 时处于在光锥外，要么在光锥上，所以世界线的连续性保证了他的世界线一定和下半光锥相交。

剩下的唯一性来自于粒子的速度不大于光速，如果有多次相交，那么找两个 $\mathbf{r}(t)^2 - c^2 t^2$ 的相邻零点 $t_0 < 0$ 和 $t_1 < 0$ ，那么肯定在其中的一点处， $\mathbf{r}(t)^2 - c^2 t^2 = 0$ 的导数值为负，不妨设为 t_0 ，此时

$$\mathbf{r}(t_0) \cdot \mathbf{v}(t_0) < c^2 t_0,$$

但是

$$|ct_0||\mathbf{v}(t_0)| = |\mathbf{r}(t_0)||\mathbf{v}(t_0)| \geq |\mathbf{r}(t_0) \cdot \mathbf{v}(t_0)| > c^2 |t_0|,$$

所以 $|\mathbf{v}(t_0)| > c$ ，这就造成了矛盾。

现在来正式推导 Lienard-Wiechert 势，如果电荷的轨迹为 $\mathbf{r}_0(t)$ ，那么他的电荷密度写作 $j^0(\mathbf{r}, t) = ec\delta^3(\mathbf{r} - \mathbf{r}_0(t))$ ，根据上节给出的基本解，推迟势写作

$$A^0 = \frac{1}{c^2} \int dV' d\tau' j^0(\mathbf{r}', \tau'/c) \frac{\delta(|\mathbf{r}' - \mathbf{r}| + \tau' - \tau)}{|\mathbf{r}' - \mathbf{r}|},$$

所以电势有

$$\varphi = e \int dV' dt' \delta^3(\mathbf{r}' - \mathbf{r}_0(t')) \frac{\delta(|\mathbf{r}' - \mathbf{r}|/c + t' - t)}{|\mathbf{r}' - \mathbf{r}|} = e \int dt' \frac{\delta(|\mathbf{r} - \mathbf{r}_0(t')|/c + t' - t)}{|\mathbf{r}_0(t') - \mathbf{r}|}.$$

注意到公式

$$\delta(f(t)) = \frac{\delta(t - t_0)}{|f'(t_0)|},$$

其中 t_0 是 $f(t)$ 的零点, 如果是多重零点, 则需要对诸零点求和。那么

$$\frac{\delta(|\mathbf{r} - \mathbf{r}_0(t')|/c + t' - t)}{|\mathbf{r}_0(t') - \mathbf{r}|} = \frac{\delta(t' - t_0)}{|\mathbf{r} - \mathbf{r}_0(t_0)| - (\mathbf{r} - \mathbf{r}_0(t_0)) \cdot \mathbf{v}_0(t_0)/c},$$

其中 t_0 是 $|\mathbf{r} - \mathbf{r}_0(t')|/c + t' - t = 0$ 的唯一零点¹, 记 $\mathbf{R} = \mathbf{r} - \mathbf{r}_0(t_0)$, 则

$$\varphi = \frac{e}{R - \mathbf{R} \cdot \mathbf{v}_0/c}.$$

利用 $\mathbf{j} = \mathbf{v}_0 \rho = e v_0 \delta^3(\mathbf{r} - \mathbf{r}_0(t))$, 可以类似地推出

$$\mathbf{A} = \frac{1}{c^2} \frac{e \mathbf{v}_0}{R - \mathbf{R} \cdot \mathbf{v}_0/c},$$

这些势能被称为 Lienard-Wiechert 势。

5.3 电荷体系在远处所产生的场

这节来考虑一个运动电荷体系在距离远大于该体系尺度的地方所产生的场。选择电荷体系内的任意一点为坐标原点, 用 \mathbf{r} 表示想要求场强位置的位矢。设 \mathbf{r}' 是电荷体系内电荷的位矢, 记 $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ 。

如果电荷体系离开很远, 则

$$R = |\mathbf{r} - \mathbf{r}'| = r - \mathbf{r}' \cdot \hat{\mathbf{r}},$$

将这个代入推迟势的公式(5.1), 则

$$A_{t-R/c}^\mu = \frac{1}{c^2} \int dV' \frac{j_{t-R/c}^\mu(\mathbf{r}')}{R} = \frac{1}{c^2} \int dV' \frac{j_{t-r/c-\mathbf{r}' \cdot \hat{\mathbf{r}}/c}^\mu(\mathbf{r}')}{r - \mathbf{r}' \cdot \hat{\mathbf{r}}},$$

一般来说, 分母可以直接写作 r , 因为 $\mathbf{r}' \cdot \hat{\mathbf{r}}$ 比起 r 来说小太多, 而在 $t - R/c$ 中却不可忽略: 是否可以略去这些项, 并不决定于 $\mathbf{r}' \cdot \hat{\mathbf{r}}/c$ 和 r/c 的相对大小, 而是 j^μ 在时间 $\mathbf{r}' \cdot \hat{\mathbf{r}}/c$ 内变化的多少。结合上面两点, 所以

$$A_{t-R/c}^\mu = \frac{1}{c^2 r} \int dV' j_{t-r/c-\mathbf{r}' \cdot \hat{\mathbf{r}}/c}^\mu(\mathbf{r}'). \quad (5.2)$$

在足够远处, 场在一个不大的空间区域内可以当做平面波。为此, 距离不仅必须比体系的尺度大很多, 而且还要比体系所辐射的电磁波波长大很多, 这样的区域被称为“波区”。

在平面波内, 电磁场之间的关系有(3.3), 即 $\mathbf{E} = c\mathbf{B} \times \hat{\mathbf{r}}$, 因为 $\mathbf{B} = \nabla \times \mathbf{A}$, 所以为了决定波区中的场, 只要求出 \mathbf{A} 就好了。使用(5.2), 并去掉 $1/r$ 的高阶项, 就有

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{1}{c} \dot{\mathbf{A}} \times \hat{\mathbf{r}}, \quad \mathbf{E} = (\dot{\mathbf{A}} \times \hat{\mathbf{r}}) \times \hat{\mathbf{r}}.$$

¹不妨设 $\mathbf{r} = 0$ 和 $t = 0$, 这在数学上即平移原点。这样方程变为 $|\mathbf{r}_0(t')| + ct' = 0$, 他的根一定是 $\mathbf{r}_0(t')^2 - (ct')^2 = (|\mathbf{r}_0(t')| + ct')(|\mathbf{r}_0(t')| - ct') = 0$ 在 $t' < 0$ 时的根。而 $\mathbf{r}_0(t')^2 - (ct')^2 = 0$ 如果有解 $t_0 < 0$, 那么就意味着 $(\mathbf{r}_0(t_0), ct_0)$ 在下半光锥上, 而也就是说 $(\mathbf{r}_0(t'), ct')$ 作为粒子的世界线就和下半光锥有交点。而这点的存在唯一性已经证明过了。

Chapter 6

联络与曲率

我们已经知道，在惯性参考系中，间隔写作 $ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$ ，如果采用保持惯性参考系的变换，即 Lorentz 变换，那么间隔不变。可是，如果我们变到了非惯性系，那么间隔一般来说不再能够写作 $ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$ ，比如有一个变换 A ，他不是仿射的，那么 $x_2 = A(x_1)$ 后（指标规则依然如同狭义相对论）

$$dx_2^\mu = \frac{\partial A^\mu}{\partial x_1^\nu}(x_1) dx_1^\nu$$

所以

$$(ds_2)^2 = \eta_{\mu\nu} dx_2^\mu dx_2^\nu = \eta_{\mu\nu} \frac{\partial A^\mu}{\partial x_1^\rho}(x_1) \frac{\partial A^\nu}{\partial x_1^\sigma}(x_1) dx_1^\rho dx_1^\sigma,$$

右边并不能写成 $\eta_{\mu\nu} dx_1^\mu dx_1^\nu$ ，而是 $g_{\mu\nu} dx_1^\mu dx_1^\nu$ ，其中

$$g_{\rho\sigma}(x_1) = \eta_{\mu\nu} \frac{\partial A^\mu}{\partial x_1^\rho}(x_1) \frac{\partial A^\nu}{\partial x_1^\sigma}(x_1)$$

不再是对角的，而且也不再是一个常矩阵。

Einstein 认为，非惯性参考系中产生的动力学效应可以等价为一个力场的作用（当然，这个力场和真的力场还是有些许不同的），从上面可以看到，这个力场应当完全由时空的几何结构（度规）决定。此外，Einstein 认为，真实存在的重力场也应该是时空的几何结构的改变，同样被度规决定。所以，时空的几何结构不再是时空的固有性质，不再是一切物理现象的背景，而本身参与到物理过程中，本身是一种物理对象。

这就是史上第一个非 Abel 场论，即广义相对论的一些基本想法。已经看到，前面所谈论的狭义相对论的基本背景是平直空间，比如 \mathbb{R}^4 或者 \mathbb{R}^{3+1} ，但现在我们必须扩展到流形上，为此伪 Riemann 几何是必须的，假设大家已经知道了一些流形上的微积分。

下面的讨论建立在伪 Riemann 流形（Semi-Riemannian Manifold）上，首先给出他的定义。

Definition 2. 假设 M 是一个可微流形，且在他每一点的切空间 $T_p M$ 都存在度规，即一个非退化¹的对称二重线性映射 $\langle \star, \star \rangle_p : T_p M \times T_p M \rightarrow \mathbb{R}$ 。

由于二重线性映射可以看作一个 2-形式 $g_p : T_p M \times T_p M \rightarrow \mathbb{R}$ ，因此度规可以看出一个 2-形式场 g 。局部来看

$$g_p(u, v) = u^i v^j g(e_i, e_j) = u^i v^j g_{ij},$$

¹即如果 $\langle a, a \rangle = 0$ ，则 $a = 0$ 。

最后的要求, (g_{ij}) 的负本征值的个数不变², 即如果在一点为 I 个, 其他点都是 I 个。这样的 (M, g) 就被称为一个伪 Riemann 流形。当 $I = 0$ 时候, 就变成 Riemann 流形。

因为 (g_{ij}) 对实对称矩阵, 所以可以对角化成实对角矩阵, 所以非退化条件等价于没有零本征值或者 $\det(g_{ij}) \neq 0$ 。

前面我们已经在 \mathbb{R}^4 上选取了度规 η 使得 \mathbb{R}^4 变成了一个伪 Riemann 流形 \mathbb{R}^{3+1} , 但是这样的说法还是有些模糊。之所以我们能说是在 \mathbb{R}^4 上选取了度规, 是因为 \mathbb{R}^4 在每一点的切空间 $T_p\mathbb{R}^4$ 都同构于 \mathbb{R}^4 , 并且, 在每一点定义的内积, 都可以看作全局定义的在 \mathbb{R}^4 上定义的度规。但是到了可微流形上, 就必须强调这些区别。

指标规则基本不变, 但以后不再用英文字母表示空间部分, 而是指任意需要累加的指标。

g^{il} 是 g_{il} 的逆, 更准确地说就是 $g^{il}g_{lm} = \delta_m^i$ 。

抬升和下降指标依然由度规 g 完成, 即 $x_i = g_{ij}x^j$ 和 $x^i = g^{ij}x_j$, 需要注意的是, 现在求导和提升或者下降指标不再对易, 因为 g 是可以改变的。

6.1 测地线

在伪 Riemann 流形上可以使用度规来定义可微曲线 $\gamma: [a, b] \rightarrow M$ 的长度

$$L(\gamma) = \int_a^b \sqrt{|\langle \gamma'(\tau), \gamma'(\tau) \rangle|} d\tau = \int_a^b \sqrt{|g(\gamma'(\tau), \gamma'(\tau))|} d\tau.$$

如果在物理上理解一条 γ 为一个粒子的世界线, 则还要加上 $\langle \gamma'(\tau), \gamma'(\tau) \rangle > 0$ 的假设, 此时 $L(\gamma)$ 就是世界线的线长, 其中参数 τ 不一定有什么物理意义。类比在 \mathbb{R}^{3+1} 里面解耦合的作用量的形式(1.6), 可以定义作用量积分为

$$S(\gamma) = -\frac{mc}{2} \int_a^b \langle \gamma'(\tau), \gamma'(\tau) \rangle d\tau,$$

而使得 $S(\gamma)$ 取极值的路径被称为测地线。

为了求出测地线, 将作用量积分在局部坐标下面写出来, 用 x 来表示从 M 到 \mathbb{R}^n 的局部平凡化, 就是说 $x(\tau) = x(\gamma(\tau))$ 是 \mathbb{R}^n 中的坐标, 此外用 \dot{f} 来表示 $df/d\tau$, 则

$$S(\gamma) = -\frac{mc}{2} \int_a^b g_{ij}(x(\tau)) \dot{x}^i(\tau) \dot{x}^j(\tau) d\tau,$$

他的 Lagrange 方程直接计算后, 发现可以写作

$$\ddot{x}^i(\tau) + \Gamma_{jk}^i(x(\tau)) \dot{x}^j(\tau) \dot{x}^k(\tau) = 0,$$

其中

$$\Gamma_{jk}^i = \frac{1}{2} g^{il} (\partial_k g_{jl} + \partial_j g_{kl} - \partial_l g_{jk})$$

被称为 Christoffel 记号, 下面我们讨论联络的时候还会再遇到。如果有了上述方程的一个解 $x(\tau)$, 下面将要指出的是 $\langle \dot{x}(\tau), \dot{x}(\tau) \rangle$ 是一个常数, 就像我们在狭义相对论里面说的 $u^\mu u_\mu = 1$ 一样。这点直接计算

$$\frac{d}{d\tau} \langle \dot{x}(\tau), \dot{x}(\tau) \rangle = \frac{d}{d\tau} (g_{ij}(x(\tau)) \dot{x}^i(\tau) \dot{x}^j(\tau)) = 0$$

就可以了, 其中要用到上面的 Lagrange 方程。这个结论可以理解成测地线的切矢量的模长 $\langle \gamma', \gamma' \rangle$ 是一个常数。

²根据惯性定理, 负本征值的个数和选取的基无关。

测地线局部的存在唯一性靠着二阶常微分方程的理论就可以得到, 从局部到整个空间, 我们可以一个一个坐标卡延拓过去, 但是否可行这里不做讨论。

在参数化过程中, 选取一个有意义的参数有利于我们观察问题, 这里, 如同狭义相对论里面一样, 可以采用线元来参数化我们的曲线, 这个时候

$$L(\gamma) = \int \sqrt{|\langle \gamma'(s), \gamma'(s) \rangle|} ds = \int_a^b \sqrt{|g(\gamma'(s), \gamma'(s))|} ds,$$

将 $u = \gamma'(s)$ 看成速度矢量, 可以类比定义动量 $p = mc u$. 又注意到 $\langle u, u \rangle$ 在选取 s 作为参数的时候满足 (不怎么严格, 但是用其他方法强行算一波还是有下面这个结论)

$$(ds)^2 \langle u, u \rangle = g_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} (ds)^2 = g_{ij} dx^i dx^j = (ds)^2,$$

所以 $\langle u, u \rangle = 1$, 因此

$$p^i p_i = g_{ij} p^i p^j = m^2 c^2.$$

反过来, 假如粒子的运动轨迹是测地线, 此时在线元参数化下 $\langle u, u \rangle = u^i u_i = 1$ 构成一个约束, 我们可以检查作用量依然可以写作

$$S = -mc \int ds$$

的形式, 检查的过程和相对论力学那节中的几乎一模一样。但是第一要注意在 $u^i u_i = g_{ij}(x) u^i u^j$ 不再不显含 x 了, 此外

$$0 = \frac{d}{ds} (g_{ij} u^i u^j) = u^j \frac{d}{ds} (g_{ij} u^i) + g_{ij} u^i (u^j)' = u^i (u_i)' + u_i (u^i)',$$

所以我们不能草率地说 u 和 du/ds 是相互正交的矢量。

只要注意上面两点, 运动方程写作

$$l_i = (2(u_i)') - \partial_i g_{mn} u^m u^n \lambda + 2u_i \lambda',$$

记 $2k_i = 2(u_i)' - \partial_i g_{mn} u^m u^n$, 只要存在一个和 k_i 正交但不和 u_i 正交的矢量 v^i , 这样

$$l_i v^i = 2v^i u_i \lambda',$$

只要满足 $l_i v^i = 0$ 就自然有 $\lambda' = 0$. 这是我们最感兴趣的情况, 因为这种情况下, 作用量的自由粒子部分可以看作

$$S = -\frac{mc}{2} \int_a^b g_{ij} u^i u^j ds.$$

而整个运动方程就写成

$$l_i = -mck_i = -\frac{mc}{2} (2(u_i)' - \partial_i g_{mn} u^m u^n).$$

6.2 矢量丛及其联络

联络的出现, 代数上实现了对矢量场的方向导数, 几何上实现了所谓的平行移动。讨论联络的基本背景是矢量场和主丛, 他们都是纤维丛的一个特例。

Definition 3. 纤维丛: 一个纤维丛 (E, π, M, F) 是指, E, M 和 F 都是可微流形, 投影 $\pi: E \rightarrow M$ 可微, 对于每一个 $x \in E$, 在 $\pi(x)$ 的附近可以找到一个邻域 U 使得存在一个微分同胚 φ 满足下列交换图:

其中 M 被称为底空间, 而 F 被称为纤维。值得注意的是, 因为有局部平凡化, 所以对每一点 $p \in M$ 都成立 $\pi^{-1}(p) \cong F$.

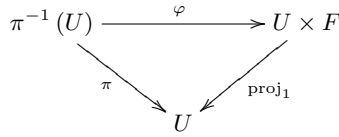


图 6.1: Locally Triviality

如果一个丛的纤维是矢量空间³，这个丛就被称为矢量丛，所以纤维是张量空间，就被称为张量丛。

Definition 4. 截面：一个纤维丛 (E, π, M, F) 的光滑截面 s 就是一个光滑函数 $s : M \rightarrow E$ 满足 $\pi \circ s = \text{Id}_M$. 所有光滑截面的集合我们记作 $\Gamma(E)$.

因为在每一点, $s(p) \in \pi^{-1}(p) \cong F$, 所以 $s(p)$ 可以看成是 F 中的元素。就是说, 一个光滑截面, 在流形 M 的每一点都赋予一个 $s(p) \in F$, 这就是我们熟知的场的概念的推广。

显然 $E = M \times F$ 是一个平凡的纤维丛, 就干脆称为平凡丛, 他的截面可以如图做出来:

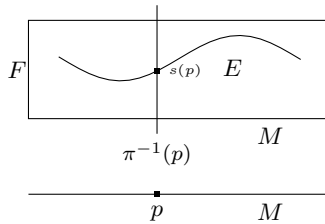


图 6.2: Trivial Bundle and its Section

一个矢量丛的重要例子就是切丛, 他是流形每一点切空间的不交并

$$TM = \coprod_{x \in M} T_x M = \bigcup_{x \in M} \{x\} \times T_x M = \bigcup_{x \in M} \{(x, y) \mid y \in T_x M\},$$

他有自然的投影 $\pi : TM \rightarrow M$ 使得 $\pi(x, v) = x$. 这个投影将一点的切空间 $T_p M$ 映到了点 p 上. 同样地, 余切丛、张量丛都可以这么定义。

矢量场是矢量丛的截面, 特别地, 度规是 2-形式场 (形式是余切矢量), 局部来看就是

$$g = g_{ij}(x) dx^i \otimes dx^j.$$

矢量丛上的联络的首要目的是为了对矢量场进行求导的。按照一般的思路, 将矢量场 s 参数化, 假设他在曲线 $\gamma(t)$ 上, $\gamma(0) = x_0$, 而 $\gamma'(0) = Y$, 则似乎他沿 Y 方向的导数应该定义成

$$\lim_{t \rightarrow 0} \frac{s(\gamma(t)) - s(x_0)}{t},$$

可是 $s(\gamma(t))$ 和 $s(x_0)$ 属于不同点的切空间, 不能相减。在 \mathbb{R}^n 很容易克服这个困难, 只要将两个切空间的矢量的端点平移到一起, 这样就可以相减了。但是到了流形上, 就会发现没那么简单, 最简单的问题, 什么是平移? 下面我们先抽象地定义联络, 然后再回来说这个问题。

³当然一个实数域上的 n 维矢量空间都同构于 \mathbb{R}^n , 看作 \mathbb{R}^n 不失一般性。

Definition 5. 令 E 是一个矢量丛, 一个 (线性) 联络或者说一个协变导数是指一个映射

$$D : \Gamma(E) \rightarrow \Gamma(E) \otimes \Gamma(T^*M)$$

满足下面几条规则: 记 $Ds(V) = D_V s$, 则 $D_V : \Gamma(E) \rightarrow \Gamma(E)$, 对于任意的 $V, W \in T_p M$, $r, s \in \Gamma(E)$ 和 $f, g \in C^\infty(M, \mathbb{R})$ 满足

$$\begin{aligned} D_{fV+gW}s &= fD_V s + gD_W s, \\ D_V(r+s) &= D_V r + D_V s, \\ D_V(fs) &= (Vf)s + fD_V s. \end{aligned}$$

$T\mathbb{R}^n$ 的截面可以写作 $s(x) = x^i(x)e_i$, 其中 e_i 是 \mathbb{R}^n 的标准基, 此时

$$D_V s = dx^i(V)e_i = (Vx^i)e_i,$$

是一个联络。

任何一个截面局部都可以写成 $s = s^i \mu_i$, 其中矢量场 μ_i 对应着矢量空间的一组基, 而对于切矢量来说, 局部可以写作 $X = X^i \partial_i$, 所以

$$D_X s = X^i D_{\partial_i} (s^j \mu_j) = X^i \partial_i s^j \mu_j + X^i s^j D_{\partial_i} \mu_j = X(s^j) \mu_j + X^i s^j D_{\partial_i} \mu_j,$$

由于 $D_{\partial_i} \mu_j$ 依然是一个截面, 所以他是 μ_i 的线性组合

$$D_{\partial_i} \mu_j = \Gamma_{ij}^k \mu_k,$$

其中 Γ_{ij}^k 被称为 Christoffel 符号, 以后会看到, 前面出现过的 Christoffel 符号只是这里的特例。此外, 以后在局部用 D_i 来记 D_{∂_i} 。综上, 联络在局部可以写作

$$D_X s = X(s^j) \mu_j + X^i s^j \Gamma_{ij}^k \mu_k.$$

有了联络的概念, 我们可以谈什么叫做平行移动, 假设现在我们有一条可微曲线 $c(t)$, 那么他的切向量是 $\dot{c}(c(t))$, 或者在局部写作 $\dot{c}^k(t) \partial_k$, 对于任意的 M 上的截面 s 可以计算沿着这个曲线切线的导数

$$\begin{aligned} D_{\dot{c}} s(c(t)) &= \dot{c}^k(t) \partial_k s^j \mu_j(c(t)) + (\dot{c}^i s^j \Gamma_{ij}^k)(c(t)) \\ &= \left(\dot{s}^j \mu_j + \dot{c}^i s^j \Gamma_{ij}^k \mu_k \right) (c(t)). \end{aligned}$$

Definition 6. 如果 $c(t)$ 是 M 上一条可微曲线, 则称满足 $D_{\dot{c}(t)} s(t) = 0$ 的解 $s(t)$ 是 $s(0)$ 在 E 中沿着曲线 c 的平行移动。

由于 $\dot{s}^j \mu_j + \dot{c}^i s^j \Gamma_{ij}^k \mu_k = 0$ 是系数 \dot{s}^j 的一阶常微分方程组, 所以解出系数后, 就解出了 s , 即给定一个初值 $s(0)$, 有唯一解 $s(t)$ 满足上述方程。

现在来检查联络是如何和平行移动联系起来的。对于 M 上的两个点, 找一条曲线 $c(t)$ 连接两点, 且在一点处的切矢量为 Y , 假设有一组联络的基 $\mu_i(t)$ 他们沿着 $c(t)$ 平行, 即 $D_{\dot{c}(t)} \mu_i(t) = 0$, 定义 $P_{c,t}(s(t))$ 将 $s(t)$ 平行移动回 $s(0)$, 局部来看就是

$$P_{c,t}(s^i(t) \mu_i(t)) = s^i(t) \mu_i(0).$$

此时按照联络的定义, 可以计算得到 $D_{\dot{c}(t)} s(t) = \dot{s}^i(t) \mu_i(t)$, 令 $t \rightarrow 0$, 则

$$D_Y s(0) = \lim_{t \rightarrow 0} \frac{s^i(t) \mu_i(0) - s^i(0) \mu_i(0)}{t} = \lim_{t \rightarrow 0} \frac{P_{c,t}(s(t)) - s(0)}{t},$$

这正是我们前面希望得到的矢量场沿一个方向的导数的定义。

现在解释“联络”的意思, 假如有一个矢量丛 E , 他的切丛为 TE , 在每一点 $p = (\pi(p), v)$, $T_p E$ 都有一个确定的子空间, 即每一点纤维的切空间 $T_v E_p$, 所谓的联络就是在每一点选定了 $T_p E$ 中 $T_v E_p$ 的补空间⁴ H_p , 即选定了 H_p 使得 $T_p E = T_v E_p \oplus H_p$. $T_v E_p$ 被称为垂直子空间, 或记做 V_p , 而 H_p 则被称为水平子空间, 这个命名的直观可以考察平凡丛 $E = M \times F$, $T_v E_p = T_v F$ 是切于 F 的, 而 F 可以看做和 M 垂直. 垂直子空间的不交并构成一个矢量丛, 我们称为垂直丛, 他是原本矢量丛的子丛, 同样我们有水平丛。

一旦有了一个联络, 任意 M 中的一条曲线 $c(t)$ 都可以将 $\dot{c}(0) = X$ 通过平行移动得到 E 中的一条曲线 $\psi(t)$. 对于同一个 X , 可以考察不同曲线 $c(t)$ 平行移动后得到的切矢量们 $\dot{\psi}(0)$, 他们构成 $T_{(c(0), X)} E = T_{s(0)} E$ 的一个子空间, 这就是 $H_{\psi(0)}$. 因为 $\psi(t)$ 是使用 D 平行移动而成, 他在纤维上的投影是 $P_{c,t}^{-1} s(0)$, 因此他的切矢量在垂直子空间上的投影 $p_v(\dot{\psi}(0))$ 是

$$p_v(\dot{\psi}(0)) = \lim_{t \rightarrow 0} \frac{P_{c,t}^{-1} s(0) - s(0)}{t} = - \lim_{t \rightarrow 0} P_{c,t}^{-1} D_X s(0) = 0,$$

因此 $\dot{\psi}(0) \in H_{\psi(0)}$.

上面这种看法可以看做联络的一种定义, 因为他是 Ehresmann 首先形式化定义的, 所以也被称为 Ehresmann 联络. 使用 Ehresmann 联络, 可以比较方便对联络的存在性等问题进行考察, 也方便搞清楚平行移动的直观等等, 详细而严谨的讨论可以参看 Jeffrey M. Lee 的 *Manifolds and Differential Geometry* 一书中的第 12 章。

重新观察

$$D_X s = X(s^j) \mu_j + s^j X^i \Gamma_{ij}^k \mu_k = ds^j(X) \mu_j + s^j X^i \Gamma_{ij}^k \mu_k,$$

将 $X^i \Gamma_{ij}^k \mu_k$ 写作 $A(X) \mu_j$, 那么就可以写出 $D = d + A$ 或者更细致一些

$$D(s) = D(s^i \mu_i) = ds^i \mu_i + s^i A \mu_i.$$

特别地, 对于基 μ_i , $D \mu_i = A \mu_i$. 现在来看看 A 到底是什么东西, 由于 $A(X) : \mu_j \mapsto X^i \Gamma_{ij}^k \mu_k = dx^i(X) \Gamma_{ij}^k \mu_k$, 或者将 $A(X)$ 写成矩阵

$$A_i^j(X) = \Gamma_{ki}^j dx^k(X)$$

因此 A 不是别的, 而是一个 1-形式值的矩阵, 写形式一点 $A \in \Gamma(\mathfrak{gl}(n, \mathbb{R}) \otimes T^*M|_U)$, 其中下标 U 指我们在局部考虑这个问题. 以后将 A 称为联络 D 的联络形式。

局部来看矢量丛 E , 设 U_α 是 M 的一个坐标图册, 那么丛的局部平凡化给出了 $\varphi_\alpha : \pi^{-1}(U_\alpha) : U_\alpha \rightarrow U_\alpha \times V$. 这样, 在非空的 $U_\alpha \cap U_\beta$ 上, 我们对于 E 的同一点 p 就有了按 φ_α 和 φ_β 不同的平凡化, 他们之间靠一个转移函数 $\varphi_{\beta\alpha} : U_\alpha \cap U_\beta \rightarrow \mathfrak{gl}(V)$ 相互联系, 即

$$\varphi_\beta \circ \varphi_\alpha(x, v) = (x, \varphi_{\beta\alpha}(v)),$$

这个式子也可以直接看做转移函数的定义。

对于同一处不同局部化的同一个截面 s , 我们使用转移函数将其联系起来 $s_\beta = \varphi_{\beta\alpha} s_\alpha$, 这样, 对于局部来看的联络 $D_\alpha = d + A_\alpha$ 和 $D_\beta = d + A_\beta$ 就有

$$\varphi_{\beta\alpha} D_\alpha s_\alpha = D_\beta s_\beta = D_\beta(\varphi_{\beta\alpha} s_\alpha),$$

⁴补空间不唯一, 所以需要选定, 比如在 \mathbb{R}^3 中, 一条直线的补空间可以是任意不和他平行的平面。

具体写出来

$$\begin{aligned}\varphi_{\beta\alpha}ds_\alpha + \varphi_{\beta\alpha}A_\alpha s_\alpha &= d(\varphi_{\beta\alpha}s_\alpha) + A_\beta(\varphi_{\beta\alpha}s_\alpha) \\ \Rightarrow \varphi_{\beta\alpha}A_\alpha s_\alpha &= d(\varphi_{\beta\alpha})s_\alpha + A_\beta(\varphi_{\beta\alpha}s_\alpha) \\ \Rightarrow A_\alpha &= \varphi_{\beta\alpha}^{-1}d\varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1}A_\beta\varphi_{\beta\alpha}.\end{aligned}$$

所以联络形式并不像一个张量一样变化，但是两个不同联络形式的差却是，所以两个联络的差是一个张量场。

一个矢量丛的对偶丛就是指纤维相互为对偶空间的丛，对偶丛上当然也会有联络，因为他也是一个矢量丛。对偶丛上的光滑截面取值为对偶矢量，因此，对偶丛上的光滑截面和矢量丛上的光滑截面在每点通过内积给出了一个值，这就是说 (μ, ν^*) 是 M 上的一个光滑实函数，因此我们可以对其求外微分，而两个丛的联络可以通过类似于 Leibniz 法则通过内积相互联系，即

$$d(\mu, \nu^*) = (D\mu, \nu^*) + (\mu, D^*\nu^*),$$

左边右边都是 1-形式，所以定义是合理的，这样，通过丛上的联络 D 就定义了对偶丛上的对偶联络 D^* 。令 μ_j 是局部的一组基，则 $(\mu_i, \mu^j) = \delta_i^j$ 且

$$0 = d(\mu_i, \mu^j) = (A^k_{\ i}\mu_k, \mu^j) + (\mu_i, A^{*j}_k\mu^k) = A^j_{\ i} + A^{*j}_i,$$

即 $A^* = -A^T$ ，其中上标 T 表示转置。如果用 Christoffel 符号表示对偶联络，则 $D^*_i\mu^j = \mu^k\Gamma_{ki}^j$ ，两个系数的关系是

$$\Gamma_{ij}^k = -\Gamma_{ji}^k$$

Definition 7. 矢量丛 E_1 和 E_2 的张量积 $E_1 \otimes E_2$ 指底空间 M 不变，而纤维变成原本纤维们的张量积。设在 E_1 和 E_2 上有联络 D_1 和 D_2 ，则在 E 上按如下方法定义了一个联络 D

$$D(s_1 \otimes s_2) = D_1s_1 \otimes s_2 + s_1 \otimes D_2s_2.$$

一个有趣的空间是 $\text{End}(E) = E \otimes E^*$ ，令 $\sigma = \sigma^i_j\mu_i \otimes \mu^j$ 是 $\text{End}(E)$ 的一个截面，则

$$\begin{aligned}D\sigma &= d\sigma^i_j\mu_i \otimes \mu^j + \sigma^i_jD\mu_i \otimes \mu^j + \sigma^i_j\mu_i \otimes D^*\mu^j \\ &= d\sigma + A^k_{\ i}\sigma^i_j\mu_k \otimes \mu^j + \sigma^i_j\mu_i \otimes A^{*j}_k\mu^k \\ &= d\sigma + A^k_{\ i}\sigma^i_j\mu_k \otimes \mu^j - \sigma^i_jA^j_{\ k}\mu_i \otimes \mu^k \\ &= d\sigma + [A, \sigma].\end{aligned}$$

类似的计算可以给出任意张量场 ω 的联络满足：

Proposition 2. 假如 ω 是一个 (r, s) 型张量场，则

$$\begin{aligned}X(\omega(\eta^i; X_j)) &= (D_X\omega)(\eta^i; X_j) \\ &\quad + \sum_i \omega(\eta^1, \dots, \eta^{i-1}, D_X\eta^i, \eta^{i+1}, \dots, \eta^r; X_j) \\ &\quad + \sum_j \omega(\eta^i; X_1, \dots, X_{j-1}, D_XX_j, X_{j+1}, \dots, X_s).\end{aligned}$$

这个等式可以理解成广义的 Leibniz 法则。

6.3 曲率

前面已经知道

$$D : \Gamma(E) \rightarrow \Gamma(E) \otimes \Gamma(T^*M) = \Gamma(E) \otimes \Omega^1(M),$$

其中 $\Omega^1(M)$ 是 1-形式场, $\Gamma(E) \otimes \Omega^1(M)$ 中的元素可以看做矢量值的 1-形式, 那么自然, 可以将 $\Gamma(E) \otimes \Omega^p(M)$ 看做矢量值的 p -形式, 不妨将其记做 $\Omega^p(E)$, 以及 $\Gamma(E) = \Omega^0(E)$, 则联络实际上是完成了矢量值的 0-形式到矢量值的 1-形式的转变:

$$D : \Omega^0(E) \rightarrow \Omega^1(E).$$

类似于 d 是从 p -形式到 $(p+1)$ -形式的转变, 我们希望拓展联络完成 $\Omega^p(E)$ 到 $\Omega^{p+1}(E)$ 的转变。

首先定义一个 s -形式 ω_1 和一个矢量值的 r -形式 $\sigma = \mu \otimes \omega_2 \in \Omega^r(E)$ 的楔积, 自然

$$\sigma \wedge \omega_1 = \mu \otimes (\omega_2 \wedge \omega_1)$$

就构造出了一个矢量值的 $(r+s)$ -形式。然后我们扩展联络如下: 对一个矢量值的 r -形式 $\sigma = \mu \otimes \omega$, 其中 ω 是一个 r -形式, 定义

$$D\sigma = D\mu \wedge \omega + \mu \otimes d\omega.$$

但是, 正如我们熟知的恒等式 $d^2 = 0$, 他其实在局部是等价于等式 $\partial_i \partial_j = \partial_j \partial_i$, 或者 $[\partial_i, \partial_j] = 0$, 直观上就是说, 沿着正交方向先后求导, 他们的结果是一样的。前面已经看到了, \mathbb{R}^n 上存在联络 $D = d$, 因此在 \mathbb{R}^n 上的这个联络依然是满足 $D^2 = 0$ 的。可以指出, $[\partial_i, \partial_j] = 0$ 是因为我们是在同一个切空间内求导的结果, 而这就忽略掉了流形本身的具体结构, 而联络并不会。

为了更清晰地看到这一点, 设 $s = s^i \mu_i$, 计算

$$\begin{aligned} D^2 s &= D(ds^i \mu_i + s^i A \mu_i) \\ &= D\mu_i \wedge ds^i + D(s^i A^j \mu_j) \\ &= A\mu_i \wedge ds^i + D(s^i \mu_j) \wedge A^j_i + s^i \mu_j dA^j_i \\ &= -ds^i \wedge A\mu_i + D(s^i \mu_j) \wedge A^j_i + (dA)s \\ &= s^i A\mu_j \wedge A^j_i + (dA)s \\ &= s^i \mu_k A^k_j \wedge A^j_i + (dA)s \\ &= (A \wedge A)s + (dA)s. \end{aligned}$$

所以作用在矢量值 0-形式上, $D^2 = A \wedge A + dA$, 这个量, 我们另外给个名字。

Definition 8. 一个联络 D 的曲率为

$$F := D^2 : \Omega^0(E) \rightarrow \Omega^2(E).$$

根据上面求的, 局部曲率算符有 $F = A \wedge A + dA$, 他一般来说不为 0, 如果曲率算符为 0, 则称这个联络是平坦的。

曲率算符可以看做 $\text{End}(E)$ 值的 2-形式, 因为 $F : \Omega^0(E) \rightarrow \Omega^2(E)$, 于是

$$F \in \Omega^2(E) \otimes (\Omega^0(E))^* = \Gamma(E) \otimes \Omega^2(M) \otimes \Gamma(E^*) = \Gamma(\text{End}(E)) \otimes \Omega^2(M),$$

这就是说 $F \in \Omega^2(\text{End}(E))$.

设 $A = A_i dx^i$, 其中 A_i 是 $n \times n$ 矩阵, 直接计算给出了

$$F = \frac{1}{2} (\partial_{[i} A_{j]} + [A_i, A_j]) dx^i \wedge dx^j.$$

观察这个式子是很有趣的, 前面在电磁学里面定义了电磁场张量 $F_{ij} = \partial_{[i} A_{j]}$, 其中 A_i 是四维势。如果四维势对应联络形式, 那么电磁场张量对应着曲率。曲率中多出来的 $[A_i, A_j]$ 在电磁学中是自然消失的。

直接的计算可以给出:

Theorem 2. *Bianchi* 等式:

$$DF = 0.$$

使用 $A_\alpha = \varphi_{\beta\alpha}^{-1} d\varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} A_\beta \varphi_{\beta\alpha}$, 我们考察 F 在坐标变换下的改变

$$\begin{aligned} F_\alpha &= A_\alpha \wedge A_\alpha + dA_\alpha \\ &= \varphi_{\beta\alpha}^{-1} A_\beta \wedge A_\beta \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} A_\beta \varphi_{\beta\alpha} \wedge \varphi_{\beta\alpha}^{-1} d\varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} d\varphi_{\beta\alpha} \wedge \varphi_{\beta\alpha}^{-1} A_\beta \varphi_{\beta\alpha} + dA_\alpha \\ &= \varphi_{\beta\alpha}^{-1} A_\beta \wedge A_\beta \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} A_\beta \wedge d\varphi_{\beta\alpha} - d\varphi_{\beta\alpha}^{-1} \varphi_{\beta\alpha} \wedge \varphi_{\beta\alpha}^{-1} A_\beta \varphi_{\beta\alpha} + dA_\alpha \\ &= \varphi_{\beta\alpha}^{-1} A_\beta \wedge A_\beta \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} A_\beta \wedge d\varphi_{\beta\alpha} - d\varphi_{\beta\alpha}^{-1} \wedge A_\beta \varphi_{\beta\alpha} + dA_\alpha \\ &= \varphi_{\beta\alpha}^{-1} A_\beta \wedge A_\beta \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} A_\beta \wedge d\varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} d(A_\beta \varphi_{\beta\alpha}) \\ &= \varphi_{\beta\alpha}^{-1} A_\beta \wedge A_\beta \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} (dA_\beta) \varphi_{\beta\alpha} \\ &= \varphi_{\beta\alpha}^{-1} F_\beta \varphi_{\beta\alpha}. \end{aligned}$$

在计算过程中, 使用了

$$0 = dI = d(\varphi_{\beta\alpha}^{-1} \varphi_{\beta\alpha}) = d\varphi_{\beta\alpha}^{-1} \varphi_{\beta\alpha} + \varphi_{\beta\alpha}^{-1} d\varphi_{\beta\alpha},$$

所以 F 变换得符合张量的变换方式, 即 F 是一个张量。可以将 Fs 写作 $R(\star, \star)s$, 其中 R 被称为曲率张量, 下面的定理实现了用联络表示曲率张量 R , 证明依然是直接的计算。

Theorem 3. 一个联络 D 的曲率张量 R 满足

$$R(X, Y)s = D_X D_Y s - D_Y D_X s - D_{[X, Y]} s.$$

因为他是张量, 所以

$$R(X, Y)(fs) = fR(X, Y)s,$$

此外对 X, Y 都函数线性从上面的定理来看是显然的, 还有 $R(X, Y) + R(Y, X) = 0$ 也是显然的。

前面提到了, 曲率涉及了流形本身 (包含联络) 的具体结构, 从曲率张量来看, 给定一个光滑曲线族 $f(s, t) : \mathbb{R} \times \mathbb{R} \rightarrow M$, 按照曲线 $f(s, 0)$ 将切矢量 v 从 $f(0, 0)$ 平行移动到 $f(1, 0)$, 然后按照曲线 $f(1, t)$ 平行移动到 $f(1, 1)$, 之后平行移动到 $f(0, 1)$, 最后平行移动回 $f(0, 0)$ 得到了新的切矢量 v' , 曲率就是度量了 v' 和 v 之间的差异。

6.4 Levi-Civita 联络

将上面两节的理论应用到伪 Riemann 流形, 应用到伪 Riemann 流形的切丛上。切丛上的联络不用 D 而写作 ∇ 。

由于伪 Riemann 流形的切丛上赋予了度规, 这就比前面单纯的矢量场上的联络理论要有更多内容, 这一点体现在联络和度规相容条件上:

$$d\langle \mu, \nu \rangle = \langle \nabla \mu, \nu \rangle + \langle \mu, \nabla \nu \rangle.$$

或者

$$X\langle\mu, \nu\rangle = \langle\nabla_X\mu, \nu\rangle + \langle\mu, \nabla_X\nu\rangle.$$

当然 Levi-Civita 联络要比这还要特殊一点。切丛上除了曲率，还能定义一个叫做挠率的张量

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

Levi-Civita 联络的挠率为 0，即 Levi-Civita 联络是无挠的。联络无挠在局部等价于 $\Gamma^k_{ij} = \Gamma^k_{ji}$ 。挠率的几何意义其实并不很清楚。

Theorem 4. 伪 Riemann 流形存在唯一的和度规相容的、无挠的联络，我们称之为 Levi-Civita 联络。

证明. 唯一性的证明，和度规相容的、无挠的联络 ∇ 一定满足：

$$\langle\nabla_X Y, Z\rangle = \frac{1}{2} (X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle + \langle[X, Y], Z\rangle - \langle[X, Z], Y\rangle - \langle[Y, Z], X\rangle).$$

利用和度规相容

$$X\langle Y, Z\rangle = \langle\nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle,$$

$$Y\langle X, Z\rangle = \langle\nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle,$$

$$Z\langle X, Y\rangle = \langle\nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle.$$

利用无挠性 $\nabla_X Y - \nabla_Y X = [X, Y]$ ，则

$$X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle = 2\langle\nabla_X Y, Z\rangle - \langle[X, Y], Z\rangle + \langle[X, Z], Y\rangle + \langle[Y, Z], X\rangle.$$

唯一性证明完毕。

存在性的证明，固定 X, Y ，令右边的为 $\omega(Z)$ ，则可以直接计算得 $\omega(fZ) = f\omega(Z)$ ，此外 $\omega(Z_1 + Z_2) = \omega(Z_1) + \omega(Z_2)$ 。那么从度规的非退化可以得知，存在唯一的 A 使得

$$\langle A, Z\rangle = \omega(Z),$$

这样子，令 $\nabla_X Y = A$ ，最后检验这是一个无挠、和度规相容的联络即可。 \square

伪 Riemann 流形的 Levi-Civita 联络的 Christoffel 记号为

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} (\partial_k g_{jl} + \partial_j g_{kl} - \partial_l g_{jk}),$$

注意到，这就是前面我们里面作用量求测地线方程时候出现的 Christoffel 记号。

Definition 9. 设有切丛上的联络 ∇ ，则一条可微曲线 $c(t)$ 称为自平行的或者称为测地线，如果其满足 $\nabla_{\dot{c}} \dot{c} = 0$ 。

利用

$$D_{\dot{c}} s(c(t)) = \left(\dot{s}^j \mu_j + \dot{c}^i s^j \Gamma^k_{ij} \mu_k \right) (c(t)),$$

直接给出测地线方程为

$$\nabla_{\dot{c}} \dot{c} = \ddot{c}^j \mu_j + \dot{c}^i \dot{c}^j \Gamma^k_{ij} \mu_k = 0,$$

再简单一些

$$\ddot{c}^k + \dot{c}^i \dot{c}^j \Gamma^k_{ij} = 0.$$

很久以前, 在谈论作用量原理的时候, 讲到要寻找和 k_i 正交的矢量 v^i , 其中

$$2k_i = 2(u_i)' - \partial_i g_{mn} u^m u^n,$$

现在我们指出 $k_i = (\nabla_u u)_i$, 所以寻找和 k_i 正交的矢量 v^i 就是指寻找一个矢量场 v 使得

$$\langle \nabla_u u, v \rangle = 0.$$

利用 Levi-Civita 联络的性质

$$u\langle u, u \rangle = 2\langle \nabla_u u, u \rangle.$$

因为 $\langle u, u \rangle = 1$, 所以 $v = u$ 就是一个自然的解, 这和狭义相对论里面讨论的一样。如果使用联络的记号, 那么在 $l_i u^i = 0$ 情况下粒子的运动方程写作

$$l_i = -mc(\nabla_u u)_i = -(\nabla_u p)_i.$$

本节的最后来谈谈活动标架法, 活动标架法利用的是局部基和对偶基来计算联络和曲率。从直接的计算开始, 可以得到:

Proposition 3. 设 D 是无挠联络, 则

$$d\omega(X_1, \dots, X_{r+1}) = \sum_{i=1}^{r+1} (-1)^{i-1} D_{X_i} \omega(X_1, \dots, \hat{X}_i, \dots, X_{r+1}),$$

其中 \hat{X}_i 指在括号中没有这项。

对 $r = 1$ 和 Levi-Civita 联络,

$$d\omega(X, Y) = \nabla_X^* \omega(Y) - \nabla_Y^* \omega(X),$$

对 μ^i 使用上式, 注意到 $\nabla^* \mu^i = A^* \mu^i = \mu^j A_j^{*i} = -\mu^j A_j^i$, 得到

$$\begin{aligned} d\mu^i(X, Y) &= \nabla_X^* \mu^i(Y) - \nabla_Y^* \mu^i(X) = -\mu^j(Y) A_j^i(X) + \mu^j(X) A_j^i(Y) \\ &= (\mu^j \otimes A_j^i - A_j^i \otimes \mu^j)(X, Y) \\ &= (\mu^j \wedge A_j^i)(X, Y), \end{aligned}$$

所以

$$d\mu^i = \mu^j \wedge A_j^i = -A_j^i \wedge \mu^j, \quad (6.1)$$

交换后有负号是因为这是两个 1-形式。

前面已经对 $s = s^i \mu_i$ 算过

$$Fs = s^i \mu_k A_j^k \wedge A_j^i + s^i \mu_j dA_j^i,$$

对 $s = \mu_i$ 有

$$F\mu_i = \mu_k A_j^k \wedge A_j^i + \mu_j dA_j^i,$$

所以

$$R(X, Y)\mu_i = F\mu_i(X, Y) = \mu_k A_j^k \wedge A_j^i(X, Y) + \mu_j dA_j^i(X, Y),$$

另一方面将 $R(X, Y)$ 写成分量的形式

$$R(X, Y)\mu_k = X^i Y^j R_{ijk}{}^l \mu_l = R_{ijk}{}^l \mu^i \otimes \mu^j(X, Y)\mu_l = \frac{1}{2} R_{ijk}{}^l \mu^i \wedge \mu^j(X, Y)\mu_l$$

其中 $R_{ijk}{}^l = \mu^l(R(\mu_i, \mu_j)\mu_k)$, 最后一个等号使用了 $R(X, Y)$ 关于 X, Y 是反对称的, 所以 $R_{ijk}{}^l$ 中的 i, j 也是反对称的, 于是

$$\frac{1}{2}R_{mni}{}^k\mu^m\wedge\mu^n=A^k{}_j\wedge A^j{}_i+dA^k{}_i. \quad (6.2)$$

式(6.1)和(6.2)统称 Cartan 结构方程。有时候会记

$$\Omega^k{}_i = -\frac{1}{2}R_{mni}{}^k\mu^m\wedge\mu^n,$$

称之为曲率形式, 那么结构方程写作

$$dA^k{}_i = -\Omega^k{}_i - A^k{}_j\wedge A^j{}_i.$$

从一般的曲率张量本身可以构造几个特殊的曲率, 比如

Definition 10. 设 Π 为切空间 T_pM 的二维子空间, 取他的一组基 X, Y , 定义 Π 的截面曲率为

$$K(\Pi) = \frac{\langle R(X, Y)X, Y \rangle}{\langle X, X \rangle \langle Y, Y \rangle - \langle X, Y \rangle^2}.$$

容易直接计算验证 $K(\Pi)$ 的定义和基的选取无关。

Definition 11. 取 $X, Y, Z \in T_pM$, 那么映射 $X \rightarrow R(Y, X)Z$ 就是一个 T_pM 上的自同态, 定义这个自同态的迹

$$\text{Ric}(Y, Z) = \text{tr}(X \rightarrow R(Y, X)Z)$$

为 Ricci 曲率张量。他直接出现在 Einstein 场方程中。

6.5 Hodge 星算子

假设流形是可定向的无边流形⁵, 此外下面谈到的形式都是有紧支集的。所以根据 Stokes 定理, $(n-1)$ -形式的外微分在整个流形上的积分为 0。

从体积形式开始, 在流形的局部, 比如考虑 (U, ϕ) , 上面的度规有 g_{ij} , 同样地, (U', ϕ') 和 g'_{ij} , 现在假设两个局部是相交的, 则体积形式作为几何量, 应该是不变的。设在 U 上的坐标为 x , U' 上的为 y , 所以

$$dy^1 \wedge \cdots \wedge dy^n = \det\left(\frac{\partial y}{\partial x}\right) dx^1 \wedge \cdots \wedge dx^n,$$

其中 $\partial y / \partial x$ 是坐标变换的 Jacobian, 这里选取的坐标变换是保向的, 即 Jacobian 的行列式大于零。

体积形式必然同时正比于 $dy^1 \wedge \cdots \wedge dy^n$ 和 $dx^1 \wedge \cdots \wedge dx^n$, 当然这是不够的, 他还依赖于度规的选取。我们考虑一下度规的变化

$$g'_{ij} = g(\partial'_i, \partial'_j) = g\left(\frac{\partial x^k}{\partial y^i} \partial_k, \frac{\partial x^l}{\partial y^j} \partial_l\right) = \frac{\partial x^k}{\partial y^i} \frac{\partial x^l}{\partial y^j} g_{kl},$$

为了消掉上面那个行列式, 我们考虑一下上式的行列式

$$\det(g'_{ij}) = \det\left(\frac{\partial x}{\partial y}\right)^2 \det(g_{ij}) = \det\left(\frac{\partial y}{\partial x}\right)^{-2} \det(g_{ij})$$

⁵Hodge 用一套整体分析的方法来研究了 de Rham 上同调群, 而他的星算子就是这套方法中很重要的一个部分。稍稍具体一点, Hodge 通过在 Riemann 流形上引入度量, 利用度量在每一个同调类中选出代表元, 而每个代表元都是椭圆算子的核, 然后就可以使用椭圆算子核的性质。具体的展开这里不可能表。

这里出现了 2，所以还要开方一下，则有

$$\sqrt{|\det(g'_{ij})|} = \det\left(\frac{\partial y}{\partial x}\right)^{-1} \sqrt{|\det(g_{ij})|}$$

综上

$$\sqrt{|\det(g'_{ij})|} dy^1 \wedge \cdots \wedge dy^n = \sqrt{|\det(g_{ij})|} dx^1 \wedge \cdots \wedge dx^n,$$

这是一个几何量，因此他就是体积形式。

Hodge 星算子的定义和第一章里面的一样

$$\omega \wedge (\star \mu) = \langle \omega, \mu \rangle \text{vol},$$

通过耐心细致但不复杂的计算可以得到，

Proposition 4. 在伪 Riemann 流形 M 上，如果他的度规的负本征值个数为 s ，那么在 p -形式上成立恒等式

$$\star \star = (-1)^{s+p(n-p)}.$$

证明很类似在 $\mathbb{R}^{s+(n-s)}$ 上的证明。

有了 Hodge 星算子，可以在任意两个 p -形式之间定义一个新的内积

$$(\omega, \mu) = \int \omega \wedge (\star \mu),$$

很容易看出他是双线性的。

现在假如有一个 $(p-1)$ 形式 ω ，那么对任意的 p -形式 μ ， $(d\omega, \mu)$ 是可以定义的，这样，类似于定义转置算符，我们可以定义 d 的对偶算子 d^* 如下

$$(\omega, d^* \mu) = (d\omega, \mu).$$

所以 d^* 完成的是从 p 形式到 $(p-1)$ -形式的转变，只是是否存在还尚且未知。

对 $\omega \wedge \star \mu$ 求外微分

$$\begin{aligned} d(\omega \wedge \star \mu) &= d\omega \wedge \star \mu + (-1)^{p-1} \omega \wedge d(\star \mu) \\ &= d\omega \wedge \star \mu + (-1)^{p-1} \omega \wedge d(\star \mu) \\ &= d\omega \wedge \star \mu + (-1)^{p-1} (-1)^{s+(p-1)(n-p+1)} \omega \wedge \star \star d(\star \mu) \\ &= d\omega \wedge \star \mu - (-1)^{s+(p+1)n+1} \omega \wedge \star (\star d \star \mu), \end{aligned}$$

两边积分后，左边由 Stokes 公式直接为 0，所以

$$0 = (d\omega, \mu) - (-1)^{s+(p+1)n+1} (\omega, \star d \star \mu),$$

最后就求得了

$$d^* = (-1)^{s+(p+1)n+1} \star d \star.$$

所以前面写过的含源部分的 Maxwell 方程组 $\star d \star \mathbf{F} = -4\pi \mathbf{j}$ 就等价于

$$d^* \mathbf{F} = -4\pi \mathbf{j}.$$

Definition 12. *Hodge-Laplace* 算子定义如下:

$$\Delta = d^*d + dd^*.$$

他将 p -形式变成 p -形式。

他的一些性质直接计算即可, 比如 $\Delta\star = \star\Delta$. 他是 Laplace 算子在流形上的推广, 假如我们的度规是对角化的, 记 $g_i = g_{ii}$, 则对光滑实函数 f 有,

$$\begin{aligned}\Delta f &= d^*df + dd^*f \\ &= \sum_i (-1)^{s+1} \star d \star \left(\frac{\partial f}{\partial x^i} dx^i \right) \\ &= \sum_i (-1)^{s+1} \star d \left((-1)^{i-1} \frac{\sqrt{|\det(g)|}}{g_i} \frac{\partial f}{\partial x^i} dx^1 \wedge \cdots \wedge \hat{dx}^i \wedge \cdots \wedge dx^n \right) \\ &= \sum_i (-1)^{s+1} \frac{\partial}{\partial x^i} \left(\frac{\sqrt{|\det(g)|}}{g_i} \frac{\partial f}{\partial x^i} \right) \frac{\star \text{vol}}{\sqrt{|\det(g)|}} \\ &= \sum_i \frac{(-1)^{s+1}}{\sqrt{|\det(g)|}} \frac{\partial}{\partial x^i} \left(\frac{\sqrt{|\det(g)|}}{g_i} \frac{\partial f}{\partial x^i} \right),\end{aligned}$$

其中 $dd^*f = 0$ 是因为 d^*f 是一个 n -形式。

对于 \mathbb{R}^n , 有 $s = 0, g_i = 1$, 则

$$\Delta f = - \sum_i \frac{\partial}{\partial x^i} \left(\frac{\partial f}{\partial x^i} \right),$$

仅仅和我们熟知的 Laplace 算子差一个负号而已。

对于 $s = 3$ 的 \mathbb{R}^{3+1} 有

$$\Delta f = \frac{\partial}{\partial x^0} \left(\frac{\partial f}{\partial x^0} \right) - \sum_{i=1}^3 \frac{\partial}{\partial x^i} \left(\frac{\partial f}{\partial x^i} \right),$$

就是我们熟知的 D'Alembert 算子。

如果我们现在有一个矢量值形式, 即 $w \otimes \omega$, 其中 $w \in W$ 是一个矢量空间中的元素, 而 ω 是一个 p -形式。如果 W 有一个非退化双线性函数 $B: W \otimes W \rightarrow \mathbb{R}$, 那么对于两个 W 值 p -形式,

$$v = \omega \otimes X, \quad w = \mu \otimes Y,$$

其中 $X, Y \in \mathfrak{g}$, 定义他们的内积为

$$\langle v, w \rangle = B(X, Y) \langle \omega, \mu \rangle,$$

其中 $\langle \omega, \mu \rangle$ 则是 p -形式之间的内积, 即

$$\langle e^1 \wedge \cdots \wedge e^p, f^1 \wedge \cdots \wedge f^p \rangle = \det(g^{-1}(e^i, f^j)).$$

有了上面这个内积的定义, 我们可以拓展星算子定义到 W 值形式依然如

$$v \wedge (\star w) = \langle v, w \rangle \text{vol}.$$

特别地, 如果 $W = \mathfrak{g}$ 是一个 Lie 代数, B 是这个 Lie 代数上面的非退化的双线性度量, 那么我们就得到了 \mathfrak{g} 形式上的星算子定义。如果 Lie 群是半单的, 那么 Lie 代数上面必然存在一个双不变度量, 即 Killing 形式 $(A, B)_K = \text{tr}(\text{ad}(A) \circ \text{ad}(B))$ 。Killing 形式本质上就是 Lie 群在双不变度量下的 Ricci 曲率, Killing 形式本身是双不变的, 而且在半单假设下, 他是非退化的。

6.6 主丛及其联络

主丛是另一种纤维丛，他在规范理论里面处于基本的地位。

Definition 13. 设有纤维丛 (P, π, M, G) ，他被称为一个主 G -丛如果

(1) G 是一个 Lie 群。

(2) G 自由右作用于 P 上：即存在运算 $P \times G \rightarrow P$ ，并且对任意的 p 成立 $p \cdot g = p$ 当且仅当 g 是 G 的单位元。 M 微分同胚于 G 在 P 上面轨道的集合 P/G 。

(3) 局部平凡化的时候，如果 $p \in P$ 局部写作 $(\pi(p), g)$ ，那么同一个轨道上的 $p \cdot g'$ 的局部平凡化写作 $(\pi(p), gg')$ 。其中 gg' 之间的乘法就是群乘法。

一个主丛自然地和一个矢量丛联系起来，通过群表示的方式。找一个矢量空间 V ，现在群 G 在 V 有表示，是左作用的。那么我们可以在 $P \times V$ 上通过

$$(p, v) \cdot g = (p \cdot g, \rho(g^{-1})v)$$

定义一个自由的右作用，其中 ρ 是 G 在 V 上的群表示。这样 $P \times V/G \rightarrow P/G$ 就是一个矢量丛，他的纤维同构于 V 。这可以很简单检验，考虑 $p \cdot G \in P/G$ ，那么他在 $P \times V/G$ 中的原象就是 $(p \cdot G, V)$ 。

如果考虑矢量丛 $P \times V/G \rightarrow M \cong P/G$ 上两个不同平凡化之间的转移函数，就会发现转移函数还是 G 左作用于 V 。转而考察主丛 $P \rightarrow M$ ，他的转移函数就是 G 的元素，由于满足

$$g\beta\alpha g\alpha\beta = e, \quad g\beta\gamma g\alpha\beta = g\alpha\gamma, \quad g\gamma\sigma g\beta\gamma g\alpha\beta = g\alpha\sigma.$$

所以转移函数的取值构成一个群，他是 G 的子群 H ，这样，就称呼 H 是主丛 $P \rightarrow M$ 的结构群。

要在主丛上定义联络，Ehresmann 联络就很方便。

Definition 14. 主丛上的联络是指在主丛的切丛上光滑地指派一个子丛 H ，他在每一点都是垂直子空间的补空间，并且对于 R_g 诱导的切空间的映射满足 $H_{p \cdot g} = (R_g)_* H_p$ 。

现在，我们来看一些主丛的例子，首先是球面上的主丛。在 \mathbb{R}^n 中，一组标准正交基 $\{e_i\}$ 称为一个标准正交标架，因为两组正交标架之间差一个正交变换 $O(n)$ ，定义映射

$$\begin{aligned} \pi: O(n) &\rightarrow S^{n-1} \\ \{e_i\} &\mapsto e_n \end{aligned}$$

当 e_n 固定时， $\{e_1, \dots, e_{n-1}\}$ 是 e_n 的正交补空间中的标准正交标架，这些标架看到全体可以等同于 $O(n-1)$ ，可验证 π 给出了 S^{n-1} 上的纤维丛，该纤维丛也是主丛，结构群是 $O(n-1)$ ，在 $O(n)$ 上的右作用为

$$\begin{aligned} O(n) \times O(n-1) &\rightarrow O(n) \\ (A, B) &\mapsto A \cdot \text{diag}(B, 1), \end{aligned}$$

其中 $\text{diag}(B, 1)$ 指的是对角为 B 和 1 的矩阵，乘法即矩阵乘法。同样地，如果考虑定向标准正交基，那么 $SO(n)$ 就是 S^{n-1} 上的主丛，纤维是 $SO(n-1)$ 。

再比如 Riemann 流形上的标架丛。设 M 是一个 Riemann 流形， $T_p M$ 中的一组标准正交基称为标架，令 $F(M)$ 是流形各处标架的不交并， $F(M)$ 上可以自然定义微分结构使其变成一个可微流形，映射 $\pi: F(M) \rightarrow M$ 显示了 $F(M)$ 其实是一个主丛，纤维是 $O(n)$ 。

Definition 15. 设 Lie 群 G 光滑右作用于流形 M ，在 G 的 Lie 代数 \mathfrak{g} 中任取矢量 A ，他都可以有一个 M 的单参变换群 $R_{\exp(tA)}$ ，这个变换群的无穷小生成元是一个 M 上的矢量场，称为由 A 诱导的基本矢量场，记做 A^* 。或者将 $(A^*)_p$ 看做曲线 $\sigma(t) = R_{\exp(tA)}p = p \cdot \exp(tA)$ 在 $\sigma(0) = p$ 处的切矢量 $\dot{\sigma}(0)$ 。

当 G 为自由作用时, 只要 $A \neq 0$, 则 A^* 处处非零. 对于主丛 P , G 自由右作用于他, 所以每一个矢量 $A \in \mathfrak{g}$ 都诱导了 P 上的基本矢量场 A^* , 映射 $A \rightarrow (A^*)_u$ 是 $\mathfrak{g} \rightarrow T_u G = V_u P$ 的线性同构, 所以 A^* 作为 TP 的切矢量场, 他只有垂直分量.

设主丛 P 上有联络 H_p , 对于每一个 $p \in P$, 可以定义线性映射 $\omega_p : T_p P \rightarrow \mathfrak{g}$ 通过对任意的 $X \in T_p P$ 指定 $\omega_p(X) = A \in \mathfrak{g}$ 为满足 $(A^*)_p = p_v(X)$ 的唯一矢量 A , 这样我们就得到了 P 上 \mathfrak{g} 值 1-形式 ω , 称为给定联络的联络形式.

很容易根据定义得到, $\forall A \in \mathfrak{g}$ 可以得到 $\omega(A^*) = A$, 由于 $A^* \in V_p P$, 或者 $p_v(A^*) = A^*$, 那么自然 $\omega(A^*) = A$.

ω 作为形式, 自然可以拉回, 尤其是右作用 R_g 的拉回 $(R_g)^*$.

Proposition 5. 联络形式满足: $(R_g)^* \omega = \text{Ad}(g^{-1})\omega$.

证明. 使用拉回的定义, 只要对任意的矢量场 X 考察 $((R_g)^* \omega)(X)$ 就可以知道 $(R_g)^* \omega$ 了. 由于线性性, 我们可以分成垂直方向和水平方向考虑, 对于竖直方向, 因为每一点都有 $\mathfrak{g} \cong V_p P$, 所以不妨将竖直矢量场看做一个基本矢量场 $X = A^*$

$$((R_g)^* \omega)(A^*) = \omega((R_g)_* A^*).$$

现在来求 $(R_g)_* A^*$,

$$\left. \frac{d}{dt} \right|_{t=0} (R_g R_{\exp(tA)})(p) = \left. \frac{d}{dt} \right|_{t=0} (R_g R_{\exp(tA)} R_g^{-1})(p \cdot g) = \left. \frac{d}{dt} \right|_{t=0} (R_{g^{-1} \exp(tA)g})(p \cdot g),$$

因为 $g^{-1} \exp(tA)g = \exp(t \text{Ad}(g^{-1})A)$, 上面的等式就是说 $\text{Ad}(g^{-1})A$ 生成了基本矢量场 $(R_g)_* A^*$.

综上

$$((R_g)^* \omega)(A^*) = \omega((R_g)_* A^*) = \text{Ad}(g^{-1})A = \text{Ad}(g^{-1})(\omega(A^*)),$$

这就是说在竖直方向 $(R_g)^* \omega = \text{Ad}(g^{-1})\omega$.

对于水平方向, 由于 $p_v(X) = 0$ 所以 $\omega(X) = 0$, 此外 $(R_g)_*$ 由于将水平空间映到水平空间, 所以 $(R_g)_* X$ 依然是水平矢量场, 于是 $(R_g)^* \omega = \text{Ad}(g^{-1})\omega$, 因为两边都等于 0. \square

反之, 如果给定了 P 上一个 \mathfrak{g} 值 1-形式 ω , 满足 $\omega(A^*) = A$ 和 $(R_g)^* \omega = \text{Ad}(g^{-1})\omega$, 那么可以定义一个联络通过

$$H_p = \ker(\omega_p) = \{X \in T_p P : \omega_p(X) = 0\}.$$

6.7 联络形式

这里在局部计算主丛的联络形式, 并将其和矢量丛里面的联络形式联系起来.

首先在主丛 P 有局部平凡化 (U_α, ψ_α) , 记微分同胚 $\psi_\alpha(p) = (\pi(p), g_\alpha(p))$. 那么在两个相交的平凡化上, 转移函数满足

$$g_\beta(p) = g_{\beta\alpha} g_\alpha(p),$$

并且主丛的定义要求平凡化要在同一个轨道里, 所以

$$\psi_\alpha(p \cdot g) = (\pi(p), g_\alpha(p)g).$$

在每一个 α , 定义 U_α 上的局部截面 σ_α 为

$$\sigma_\alpha(x) = \psi_\alpha^{-1}(x, e),$$

转移到其他平凡化里面

$$\sigma_\beta(x) = \psi_\beta^{-1}(x, e) = \psi_\alpha^{-1}(x, g_{\alpha\beta}(x)) = \psi_\alpha^{-1}(x, e) \cdot g_{\alpha\beta}(x) = \sigma_\alpha(x) \cdot g_{\alpha\beta}(x).$$

对每一个 α , 通过联络形式 ω 可以定义一个 \mathfrak{g} 值 1-形式 ω_α , 那么通过截面 $\sigma_\alpha : U_\alpha \rightarrow P$ 诱导的 $\sigma_\alpha^* : T^*P \rightarrow T^*U_\alpha$, 定义

$$\omega_\alpha = \sigma_\alpha^*(\omega).$$

关于他在转移函数下的表现, 考察 $x \in U_\alpha \cap U_\beta$, 设 $X \in T_x M$, 由截面在转移函数下的表现 $\sigma_\beta(x) = R_{g_{\alpha\beta}(x)}\sigma_\alpha(x)$, 求个导数 (定义 $L_p(g) = p \cdot g$)

$$(\sigma_\beta(x))_* X = (R_{g_{\alpha\beta}(x)})_*(\sigma_\alpha(x))_* X + (L_{\sigma_\alpha(x)})_*(g_{\alpha\beta})_* X,$$

上面式子的第二项是因为 $g_{\alpha\beta}(x)$ 也是和 x 有关的。于是

$$\begin{aligned} \omega_\beta(x)X &= ((\sigma_\beta)^*\omega(x))X \\ &= \omega(\sigma_\beta(x))(\sigma_\beta(x))_* X \\ &= \omega(\sigma_\beta(x))((R_{g_{\alpha\beta}(x)})_* \sigma_\alpha(x))_* X \\ &= \omega(\sigma_\beta(x)) \left((R_{g_{\alpha\beta}(x)})_*(\sigma_\alpha(x))_* X + (L_{\sigma_\alpha(x)})_*(g_{\alpha\beta})_* X \right) \\ &= \text{Ad}(g_{\alpha\beta}^{-1}) (\omega(\sigma_\alpha(x))(\sigma_\alpha(x))_* X) + \omega(\sigma_\beta(x)) ((L_{\sigma_\alpha(x)})_*(g_{\alpha\beta})_* X), \\ &= \text{Ad}(g_{\alpha\beta}^{-1}) (\omega_\alpha(x)X) + L_{\sigma_\alpha(x)}^*(\omega)(g_{\alpha\beta}(x)) ((g_{\alpha\beta})_* X) \\ &= \text{Ad}(g_{\alpha\beta}^{-1}) (\omega_\alpha(x)X) + g_{\alpha\beta}^* \circ L_{\sigma_\alpha(x)}^*(\omega)(x) (X) \end{aligned}$$

定义 $\theta = L_{\sigma_\alpha(x)}^*(\omega)$, 为了证实这个定义是合理的, 我们需要检验这和选取的平凡化无关。选一个 $Y \in T_g G$ 的元素, 再重用一下符号 L , 将其看做群的左乘, 那么 $L_{g^{-1}}$ 的导数 $(L_{g^{-1}})^* : T_g G \rightarrow T_e G = \mathfrak{g}$ 且 $(L_g^{-1})_* Y = Y_0$. 现在

$$\begin{aligned} L_{\sigma_\alpha(x)}^*(\omega)(g)(Y) &= L_{\sigma_\alpha(x)}^*(\omega)(g)((L_g)_* Y_0) \\ &= (\omega)(\sigma_\alpha(x) \cdot g) (L_{\sigma_\alpha(x)}_* L_g^* Y_0) \\ &= (\omega)(\sigma_\alpha(x) \cdot g) ((L_{\sigma_\alpha(x) \cdot g})_* Y_0), \end{aligned}$$

注意, $(L_{\sigma_\alpha(x) \cdot g})_* Y_0$ 其实是 $(\sigma_\alpha(x) \cdot g) \cdot \exp(tY_0)$ 的切矢量, 所以他是 Y_0 诱导的基本矢量场的矢量 $(Y_0)_{\sigma_\alpha(x) \cdot g}^*$, 对于基本矢量场, 按照联络形式的性质 $\omega_{\sigma_\alpha(x) \cdot g}((Y_0)_{\sigma_\alpha(x) \cdot g}^*) = Y_0$, 综上

$$\theta(Y) = L_{\sigma_\alpha(x)}^*(\omega)(g)(Y) = \omega_{\sigma_\alpha(x) \cdot g}(Y_0^*) = Y_0 = (L_{g^{-1}})_* Y,$$

所以 θ 的定义是合理的。在 Lie 群理论里面, 这个 θ 称为典则左不变 \mathfrak{g} 值 1-形式, 或者 Maurer-Cartan 形式, 称为左不变的是因为对任意的群元 g 都有 $L_g^* \theta = \theta$, 从定义看这是显然的。

最后 ω_α 在转移函数下的表现就是

$$\omega_\beta = \text{Ad}(g_{\alpha\beta}^{-1})\omega_\alpha + g_{\alpha\beta}^* \theta,$$

反之, 给定了所有的 ω_β 也就给出了联络形式, 当然也给出了联络。

前面说过, 一个主丛可以自然地通过表示和一个矢量丛 (称为伴随丛) 联系在一起。如果已知了表示 $\rho : G \rightarrow \mathfrak{gl}(V)$, 我们就可以几乎把所有的 $g \in G$ 换成 $\rho(g) \in \mathfrak{gl}(V)$ 而没什么问题, 比如现在的矢量丛上的转移函数即 $\varphi_{\alpha\beta} = \rho(g_{\alpha\beta})$.

给定主丛上的联络, 即联络形式 ω , 可以布局定义 ω_α , 由于 $\rho: G \rightarrow \mathfrak{gl}(V)$ 是 Lie 群同态, 当然就诱导了 Lie 代数同态 $\rho_*: \mathfrak{g} \rightarrow \mathfrak{gl}(V)$, 那么 $A_\alpha = \rho_*\omega_\alpha$ 就是 $\mathfrak{gl}(V)$ 值 1-形式。

因为一般线性群有直接继承自 \mathbb{R}^{n^2} 的微分结构, 所以

$$(L_g)_*v = \lim_{t \rightarrow 0} \frac{1}{t} (L_g(a + tv) - L_g(a)) = \lim_{t \rightarrow 0} \frac{1}{t} (L_g(tv)) = L_g(v) = gv.$$

其中 g 和 v 都是矩阵, 矩阵乘矩阵还是矩阵, 所以 Lie 代数也是矩阵的形式。设 $dg = (dx_{ij})$, 那么 v 就可以写成 $dg(v)$, 因为 $dx_{ij}(v) = v_{ij}$, 则

$$\theta = g^{-1}dg.$$

所以一般线性群上的 Maurer-Cartan 形式即 $g^{-1}dg$, 将所有的 \mathfrak{g} 值形式 ω 换成 $\rho_*\omega$ 就好, 那么按照这个原则

$$\begin{aligned} A_\beta &= \rho(g_{\alpha\beta})^{-1} A_\alpha \rho(g_{\alpha\beta}) + \rho(g_{\alpha\beta})^* (\rho(g)^{-1} d\rho(g)) \\ &= \rho(g_{\alpha\beta})^{-1} A_\alpha \rho(g_{\alpha\beta}) + \rho(g_{\alpha\beta})^{-1} d\rho(g_{\alpha\beta}) \\ &= \varphi_{\alpha\beta}^{-1} A_\alpha \varphi_{\alpha\beta} + \varphi_{\alpha\beta}^{-1} d\varphi_{\alpha\beta}, \end{aligned}$$

这是前面我们已经见到过的变换方式。利用 A_α 可以构造矢量丛上的联络, 事实上, 固定 V 的一组基 e_i , 那么可以在每一个 U_α 上定义一个伴随丛的局部截面

$$\mu_{\alpha,i}(x) = \varphi_\alpha^{-1}(x, e_i)$$

记 $\mu_\alpha = (\mu_{\alpha,1}, \dots, \mu_{\alpha,n})$ 是一个行矢量, 那么矩阵乘法给出

$$\mu_\alpha \cdot \rho(g_{\alpha\beta}) = \mu_\alpha \cdot \varphi_{\alpha\beta} = \mu_\beta.$$

将 A_α 写成矩阵形式 $A_{\alpha,j}^i$, 我们可以定义些局部联络

$$D_\alpha \mu_{\alpha,i} = \mu_{\alpha,j} \otimes A_{\alpha,i}^j,$$

或者写成矩阵形式

$$D_\alpha \mu_\alpha = \mu_\alpha \otimes A_\alpha.$$

如果 $s = s^i \mu_{\alpha,i}$ 是 U_α 中的截面, 则定义局部联络算子

$$D_\alpha s = \mu_{\alpha,i} \otimes ds^i + s^i D_\alpha \mu_{\alpha,i}.$$

容易通过 A_α 和 μ_α 在转移函数下的行为证明在 $U_\alpha \cap U_\beta \neq \emptyset$ 上, $D_\alpha = D_\beta$ 。这样, 我们就给出了一个矢量丛上的联络 D 。

下面讨论主丛上的曲率。设 η_1, η_2 都是某个流形上的 \mathfrak{g} 值微分形式, 取 \mathfrak{g} 的一组基 e_i , 所以

$$\eta_1 = e_i \otimes \eta_1^i, \quad \eta_2 = e_i \otimes \eta_2^i,$$

定义, 两个 \mathfrak{g} 值微分形式的对易为

$$[\eta_1, \eta_2] = [e_i, e_j] \otimes \eta_1^i \wedge \eta_2^j,$$

容易验证这个定义和基的选取无关。根据定义, 直接计算可以得到

$$[\eta_2, \eta_1] = (-1)^{\deg \eta_1 \cdot \deg \eta_2 + 1} [\eta_1, \eta_2], \quad d[\eta_1, \eta_2] = [d\eta_1, \eta_2] + (-1)^{\deg \eta_1} [\eta_1, d\eta_2],$$

此外对于两个 1-形式，直接的计算也有

$$[\eta_1, \eta_2](X, Y) = [\eta_1(X), \eta_2(Y)] - [\eta_1(Y), \eta_2(X)].$$

注意，根据上面的两个计算有

$$[\eta_1, \eta_1](X, Y) = [\eta_1(X), \eta_1(Y)] - [\eta_1(Y), \eta_1(X)] = 2[\eta_1(X), \eta_1(Y)],$$

此时一般是 $[\eta_1, \eta_1] \neq 0$ 。类比矢量丛 $F = dA + A \wedge A$ ，我们定义主丛上的曲率形式为

$$F = d\omega + \frac{1}{2}[\omega, \omega].$$

曲率这样定义的合理性以及下面这个定理，不再详细叙述， $1/2$ 来自于 $[\eta_1, \eta_1]$ 的计算，其余的可以自行检验两个表达式在矢量丛上的统一性。

Theorem 5. *Bianchi 等式*: $dF = [F, \omega]$.

6.8 规范场

定义一个规范场论需要一个 Lie 群 G ，一个主 G -丛 P ，底流形为 M ，此外 Lie 代数 \mathfrak{g} 上面有一个双不变度量，即 $\langle \text{Ad}_g X, \text{Ad}_g Y \rangle = \langle X, Y \rangle$ 对任意的 $g \in G$ 和 $X, Y \in \mathfrak{g}$ 都成立。

P 上的一个联络被称为一个规范场，正如上面看到的，给出规范场只要给出其规范形式 A 即可，所以经常称呼 A 为规范场，但实际上， A 不是坐标无关的，他依赖于局部化的选取，所以一个规范场是确定到一个联络上面的。在不同的局部化选取，或者说在不同规范下，规范场 A 的变换为

$$A_\beta = \text{Ad}(g_{\alpha\beta}^{-1})A_\alpha + g_{\alpha\beta}^* \theta,$$

如果 G 是一个矩阵群，那么变换写作

$$A_\beta = g_{\alpha\beta}^{-1} \cdot A_\alpha \cdot g_{\alpha\beta} + g_{\alpha\beta}^{-1} \cdot dg_{\alpha\beta}.$$

规范场 A 的曲率 F_A 由下式定义：

$$F_A = dA + \frac{1}{2}[A, A].$$

定一个规范，矩阵群下设 $A = A_i dx^i$ ，那么

$$F = \frac{1}{2}(\partial_{[i} A_{j]} + [A_i, A_j]) dx^i \wedge dx^j.$$

现在来看经典电动力学确实是一种规范场论，取 M 为 \mathbb{R}^{3+1} ， $G = U(1)$ ， $P = M \times G$ 是一个平凡主 G -丛。由于 $\mathfrak{u}(1) = \mathbb{R}$ ，所以所谓的 $\mathfrak{u}(1)$ 值 1-形式 $A = A_i dx^i$ 其实就是普通的 1-形式，其中 A_i 是标量函数。记规范场为 $A = (\mathbf{A}, \varphi)$ ，转移函数 $\psi: M \rightarrow U(1)$ 的作用就变成了

$$A' = A + \psi^{-1} d\psi = A + d \ln(\psi),$$

由于定义了同一个联络，所以也就是同一个场。此外，曲率 F_A 就是

$$F = \frac{1}{2} \partial_{[i} A_{j]} dx^i \wedge dx^j,$$

他的矩阵当然就是熟悉的定义 $F_{ij} = \partial_{[i} A_{j]}$ ，即电磁场张量。因为这里曲率是反对称的，他的 $1+2+3=6$ 个独立分量确定了 \mathbf{E} 和 \mathbf{B} 的全部空间分量。Bianchi 等式 $dF_A = [F_A, A]$ 或在矢量丛下写作 $DF_A = 0$ 就告诉了我们第一组 Maxwell 方程。

第二组 Maxwell 方程的写出需要作用量，类比电磁场，一个纯经典规范场的作用量写作

$$S_{YM}(A) = -\frac{1}{2} \int_M F_A \wedge (\star F_A),$$

他被称为 Yang-Mills 作用量。一般在定义星算子的时候，会选择 Killing 形式作为非退化双线性型（加上半单假设）。经典电动力学的作用量中的 Killing 形式即两个实数相乘。如果含有源，则作用量写作

$$S(A) = - \int_M A \wedge \star j - \frac{1}{2} \int_M F_A \wedge \star F_A.$$

如果考虑 Lie 群是矩阵群，那么可以在矢量丛下对其变分就得到了第二组 Maxwell 方程

$$\star D \star F_A = -j.$$