第一章　微分同胚（曲线坐标系）

1.1　微分同胚

1.1.1　双射

设 f 是集合 A 到 B 的映照．如果 A 中不同的元素有不同的像，则称 f 为**单射**（也叫“一对一”）；如果 B 中每个元素都是 A 中元素的像，则称 f 为**满射**；如果 f 既是单射又是满射，则称 f 为**双射**（也叫“一一对应”）．三种情况的示意见图 1.1．

Images/Three_Mappings.PNG

图 1.1: 单射、满射与双射

设开集 $\mathfrak{D}_X, \mathfrak{D}_x \subset \mathbb{R}^m$，它们之间存在双射，即一一对应关系：

$$\boldsymbol{X}(\boldsymbol{x}) : \mathfrak{D}_x \ni \boldsymbol{x} = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix} \mapsto \boldsymbol{X}(\boldsymbol{x}) = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix}(\boldsymbol{x}) \in \mathfrak{D}_X. \tag{1.1}$$

由于该映照实现了 $\mathfrak{D}_x$ 到 $\mathfrak{D}_X$ 之间的双射，因此它存在逆映照：

$$\boldsymbol{x}(\boldsymbol{X}) : \mathfrak{D}_X \ni \boldsymbol{X} = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix} \mapsto \boldsymbol{x}(\boldsymbol{X}) = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix}(\boldsymbol{X}) \in \mathfrak{D}_x. \tag{1.2}$$

我们把 $\mathfrak{D}_X$ 称为**物理域**，它是实际物理事件发生的区域；$\mathfrak{D}_x$ 则称为**参数域**．由于物理域通常较为复杂，因此我们常把参数域取为规整的形状，以便之后的处理．

设物理量 $f(\boldsymbol{X})$ 定义在物理域 $\mathfrak{D}_{\boldsymbol{X}} \subset \mathbb{R}^m$ 上[1]，则 f 就定义了一个**场**：

$$f\,:\,\mathfrak{D}_{\boldsymbol{X}} \ni \boldsymbol{X} \mapsto f(\boldsymbol{X}). \tag{1.3}$$

所谓的“场”，就是自变量用位置刻画的映照．它可以是**标量场**，如温度、压强、密度等，此时 $f(\boldsymbol{X}) \in \mathbb{R}$；也可以是**向量场**，如速度、加速度、力等，此时 $f(\boldsymbol{X}) \in \mathbb{R}^m$；对于更深入的物理、力学研究，往往还需引入**张量场**，此时 $f(\boldsymbol{X}) \in \mathscr{T}^r(\mathbb{R}^m)$．

$\boldsymbol{X}$ 存在于物理域 $\mathfrak{D}_{\boldsymbol{X}}$ 中，我们称它为**物理坐标**．由于上文已经定义了 $\mathfrak{D}_{\boldsymbol{x}}$ 到 $\mathfrak{D}_{\boldsymbol{X}}$ 之间的双射（不是 f！），因此 $\mathfrak{D}_{\boldsymbol{x}}$ 中就有唯一的 $\boldsymbol{x}$ 与 $\boldsymbol{X}$ 相对应，它称为**参数坐标**（也叫**曲线坐标**）．又因为物理域 $\mathfrak{D}_{\boldsymbol{X}}$ 上已经定义了场 $f(\boldsymbol{X})$，参数域中必然唯一存在场 $\tilde{f}(\boldsymbol{x})$ 与之对应：

$$\tilde{f}\,:\,\mathfrak{D}_{\boldsymbol{x}} \ni \boldsymbol{x} \mapsto \tilde{f}(\boldsymbol{x}) = f \circ \boldsymbol{X}(\boldsymbol{x}) = f\big(\boldsymbol{X}(\boldsymbol{x})\big). \tag{1.4}$$

$\boldsymbol{x}$ 与 $\boldsymbol{X}$ 是完全等价的，因而 $\tilde{f}$ 与 f 也是完全等价的，所以同样有

$$f(\boldsymbol{X}) = \tilde{f}\big(\boldsymbol{x}(\boldsymbol{X})\big). \tag{1.5}$$

物理域中的场要满足守恒定律，如质量守恒、动量守恒、能量守恒等．从数学上看，这些守恒定律就是 $f(\boldsymbol{X})$ 需要满足的一系列偏微分方程．将场变换到参数域后，它仍要满足这些方程．但我们已经设法将参数域取得较为规整，故在其上进行数值求解就会相当方便．

1.1.2　参数域方程

上文已经提到，物理域中的场 $f(\boldsymbol{X})$ 需满足守恒定律，这等价于一系列偏微分方程（PDE）．在物理学和力学中，用到的 PDE 通常是二阶的，它们可以写成

$$\forall\, \boldsymbol{X} \in \mathfrak{D}_{\boldsymbol{X}},\quad \sum_{\alpha=1}^{m} A_\alpha(\boldsymbol{X}) \frac{\partial f}{\partial X^\alpha}(\boldsymbol{X}) + \sum_{\alpha=1}^{m}\sum_{\beta=1}^{m} B_{\alpha\beta}(\boldsymbol{X}) \frac{\partial^2 f}{\partial X^\beta \partial X^\alpha}(\boldsymbol{X}) = 0 \tag{1.6}$$

的形式．我们的目标是把该物理域方程转化为参数域方程，即关于 $\tilde{f}(\boldsymbol{x})$ 的 PDE．多元微积分中已经提供了解决方案：**链式求导法则**．

考虑到

$$f(\boldsymbol{X}) = \tilde{f}\big(\boldsymbol{x}(\boldsymbol{X})\big) = \tilde{f}\big(x^1(\boldsymbol{X}), \cdots, x^m(\boldsymbol{X})\big), \tag{1.7}$$

于是有

$$\frac{\partial f}{\partial X^\alpha}(\boldsymbol{X}) = \sum_{s=1}^{m} \frac{\partial \tilde{f}}{\partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \frac{\partial x^s}{\partial X^\alpha}(\boldsymbol{X}). \tag{1.8}$$

这里用到的链式法则，由复合映照可微性定理驱动，它要求 $\tilde{f}$ 关于 $\boldsymbol{x}$ 可微，同时 $\boldsymbol{x}$ 关于 $\boldsymbol{X}$ 可微．

对于更高阶的项，往往需要更强的条件．一般地，我们要求

$$\begin{cases} \boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p\big(\mathfrak{D}_{\boldsymbol{x}}; \mathbb{R}^m\big); & \text{(1.9-a)} \\ \boldsymbol{x}(\boldsymbol{X}) \in \mathscr{C}^p\big(\mathfrak{D}_{\boldsymbol{X}}; \mathbb{R}^m\big). & \text{(1.9-b)} \end{cases}$$

这里的 $\mathscr{C}^p$ 指直至 p 阶偏导数（存在且）连续的映照全体；$p = 1$ 时，它就等价于可微．至于 p 的具体取值，则由 PDE 的阶数所决定．

① 实际的物理事件当然只会发生在三维 Euclid 空间中（只就“空间”而言），但在数学上也可以推广到 m 维．

通常情况下，已知条件所给定的往往都是 $\mathfrak{D}_{\boldsymbol{x}}$ 到 $\mathfrak{D}_{\boldsymbol{X}}$ 的映照

$$\boldsymbol{X}(\boldsymbol{x}):\ \mathfrak{D}_{\boldsymbol{x}} \ni \boldsymbol{x} = \begin{bmatrix} x^1 \\ \vdots \\ x^m \end{bmatrix} \mapsto \boldsymbol{X}(\boldsymbol{x}) = \begin{bmatrix} X^1 \\ \vdots \\ X^m \end{bmatrix}(\boldsymbol{x}) \in \mathfrak{D}_{\boldsymbol{X}}, \tag{1.10}$$

用它不好直接得到式 (1.8) 中的 $\partial x^s / \partial X^\alpha$ 项，但获得它的“倒数” $\partial X^\alpha / \partial x^s$ 却很容易，只需利用 **Jacobi 矩阵**：

$$\mathrm{D}\boldsymbol{X}(\boldsymbol{x}) \triangleq \begin{bmatrix} \dfrac{\partial X^1}{\partial x^1} & \cdots & \dfrac{\partial X^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial X^m}{\partial x^1} & \cdots & \dfrac{\partial X^m}{\partial x^m} \end{bmatrix}(\boldsymbol{x}) \in \mathbb{R}^{m\times m}, \tag{1.11}$$

它是一个方阵.

有了 Jacobi 矩阵，施加一些手法就可以得到所需要的 $\partial x^s / \partial X^\alpha$ 项．考虑到

$$\forall\, \boldsymbol{X} \in \mathfrak{D}_{\boldsymbol{X}}, \quad \boldsymbol{X}\big(\boldsymbol{x}(\boldsymbol{X})\big) = \boldsymbol{X}, \tag{1.12}$$

并且其中的 $\boldsymbol{X}(\boldsymbol{x})$ 和 $\boldsymbol{x}(\boldsymbol{X})$ 均可微，可以得到

$$\mathrm{D}\boldsymbol{X}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \mathrm{D}\boldsymbol{x}(\boldsymbol{X}) = \boldsymbol{I}_m, \tag{1.13}$$

其中的 $\boldsymbol{I}_m$ 是单位阵．因此

$$\mathrm{D}\boldsymbol{x}(\boldsymbol{X}) \triangleq \begin{bmatrix} \dfrac{\partial x^1}{\partial X^1} & \cdots & \dfrac{\partial x^1}{\partial X^m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x^m}{\partial X^1} & \cdots & \dfrac{\partial x^m}{\partial X^m} \end{bmatrix}(\boldsymbol{X}) = (\mathrm{D}\boldsymbol{X})^{-1}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial X^1}{\partial x^1} & \cdots & \dfrac{\partial X^1}{\partial x^m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial X^m}{\partial x^1} & \cdots & \dfrac{\partial X^m}{\partial x^m} \end{bmatrix}^{-1}(\boldsymbol{x}). \tag{1.14}$$

用代数的方法总可以求出

$$\varphi^s_\alpha := \frac{\partial x^s}{\partial X^\alpha}, \tag{1.15}$$

它是通过求逆运算确定的函数，即位于矩阵 $\mathrm{D}\boldsymbol{x}$ 第 s 行第 α 列的元素．这样就有

$$\frac{\partial f}{\partial X^\alpha}(\boldsymbol{X}) = \sum_{s=1}^{m} \frac{\partial \tilde{f}}{\partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \varphi^s_\alpha\big(\boldsymbol{x}(\boldsymbol{X})\big). \tag{1.16}$$

接下来处理二阶偏导数．由上式，

$$\begin{aligned} \frac{\partial^2 f}{\partial X^\beta \partial X^\alpha}(\boldsymbol{X}) = \sum_{s=1}^{m} \Bigg[& \left(\sum_{k=1}^{m} \frac{\partial^2 \tilde{f}}{\partial x^k \partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \frac{\partial x^s}{\partial X^\beta}(\boldsymbol{X}) \right) \cdot \varphi^s_\alpha\big(\boldsymbol{x}(\boldsymbol{X})\big) \\ & + \frac{\partial \tilde{f}}{\partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \left(\sum_{k=1}^{m} \frac{\partial \varphi^s_\alpha}{\partial x^k}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \frac{\partial x^k}{\partial X^\beta}(\boldsymbol{X}) \right) \Bigg] \end{aligned}$$

继续利用式 (1.15)，有

$$= \sum_{s=1}^{m} \Bigg[\left(\sum_{k=1}^{m} \frac{\partial^2 \tilde{f}}{\partial x^k \partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big) \cdot \varphi^s_\beta\big(\boldsymbol{x}(\boldsymbol{X})\big) \right) \cdot \varphi^s_\alpha\big(\boldsymbol{x}(\boldsymbol{X})\big)$$

$$+\frac{\partial \tilde{f}}{\partial x^s}\big(\boldsymbol{x}(\boldsymbol{X})\big)\cdot\left(\sum_{k=1}^{m}\frac{\partial \varphi_\alpha^s}{\partial x^k}\big(\boldsymbol{x}(\boldsymbol{X})\big)\cdot\varphi_\beta^k\big(\boldsymbol{x}(\boldsymbol{X})\big)\right)\Bigg]. \tag{1.17}$$

这样，就把一阶和二阶偏导数项全部用关于 $\boldsymbol{x}$ 的函数[①]表达了出来．换句话说，我们已经把物理域中 f 关于 $\boldsymbol{X}$ 的 PDE，转化成了参数域中 $\tilde{f}$ 关于 $\boldsymbol{x}$ 的 PDE．这就是上文要实现的目标．

1.1.3 微分同胚的定义

上文已经指出了 $\mathfrak{D}_{\boldsymbol{x}}$ 到 $\mathfrak{D}_{\boldsymbol{X}}$ 的映照 $\boldsymbol{X}(\boldsymbol{x})$ 所需满足的一些条件．这里再次罗列如下:

1. $\mathfrak{D}_{\boldsymbol{X}}, \mathfrak{D}_{\boldsymbol{x}} \subset \mathbb{R}^m$ 均为**开集**[②]；
2. 存在 $\mathfrak{D}_{\boldsymbol{x}}$ 同 $\mathfrak{D}_{\boldsymbol{X}}$ 之间的**双射** $\boldsymbol{X}(\boldsymbol{x})$，即存在**一一对应**关系；
3. $\boldsymbol{X}(\boldsymbol{x})$ 和它的逆映照 $\boldsymbol{x}(\boldsymbol{X})$ 满足一定的**正则性**要求．

对第 3 点要稍作说明．

如果满足这三点，则称 $\boldsymbol{X}(\boldsymbol{x})$ 为 $\mathfrak{D}_{\boldsymbol{x}}$ 与 $\mathfrak{D}_{\boldsymbol{X}}$ 之间的 **$\mathscr{C}^p$-微分同胚**，记为 $\boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p\big(\mathfrak{D}_{\boldsymbol{x}}; \mathfrak{D}_{\boldsymbol{X}}\big)$．把物理域中的一个部分对应到参数域上的一个部分，需要的仅仅是双射这一条件；而要使得物理域中所满足的 PDE 能够转换到参数域上，就需要“过去”和“回来”都满足 p 阶偏导数连续的条件（即正则性要求）.

有了微分同胚，物理域中的位置就可用参数域中的位置等价地进行刻画．因此我们也把微分同胚称为**曲线坐标系**．

1.2 向量值映照的可微性

1.2.1 可微性的定义

设 $\boldsymbol{x}_0$ 是参数域 $\mathfrak{D}_{\boldsymbol{x}}$ 中的一个内点．在映照 $\boldsymbol{X}(\boldsymbol{x})$ 的作用下，它对应到物理域 $\mathfrak{D}_{\boldsymbol{X}}$ 中的点 $\boldsymbol{X}\big(\boldsymbol{x}_0\big)$．参数域是一个开集．根据开集的定义，必然存在一个实数 $\lambda > 0$，使得以 $\boldsymbol{x}_0$ 为球心、λ 为半径的球能够完全落在定义域 $\mathfrak{D}_{\boldsymbol{x}}$ 内，即

$$\mathfrak{B}_\lambda\big(\boldsymbol{x}_0\big) \subset \mathfrak{D}_{\boldsymbol{x}}, \tag{1.18}$$

其中的 $\mathfrak{B}_\lambda\big(\boldsymbol{x}_0\big)$ 表示 $\boldsymbol{x}_0$ 的 λ 邻域．

如果 $\exists\ \mathrm{D}\boldsymbol{X}\big(\boldsymbol{x}_0\big) \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^m)$[③]，满足

$$\forall\, \boldsymbol{x}_0 + \boldsymbol{h} \in \mathfrak{B}_\lambda\big(\boldsymbol{x}_0\big), \quad \boldsymbol{X}\big(\boldsymbol{x}_0 + \boldsymbol{h}\big) - \boldsymbol{X}\big(\boldsymbol{x}_0\big) = \mathrm{D}\boldsymbol{X}\big(\boldsymbol{x}_0\big)(\boldsymbol{h}) + o\left(\|\boldsymbol{h}\|_{\mathbb{R}^m}\right) \in \mathbb{R}^m, \tag{1.19}$$

则称向量值映照 $\boldsymbol{X}(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 点**可微**．其中，$\mathscr{L}(\mathbb{R}^m, \mathbb{R}^m)$ 表示从 $\mathbb{R}^m$ 到 $\mathbb{R}^m$ 的**线性变换**全体．

根据这个定义，所谓可微性，指由自变量变化所引起的因变量变化，可以用一个线性变换近似，而误差为一阶无穷小量．自变量可见到因变量空间最简单的映照形式就是线性映照（线性变换），因而具有可微性的向量值映照具有至关重要的作用．

① 当然它仍然是 $\boldsymbol{X}$ 的隐函数：$\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{X})$．

② 用形象化的语言来说，如果在区域中的任意一点都可以吹出一个球，并能使球上的每个点都落在区域内，那么这个区域就是**开集**．这是复合映照可微性定理的一个要求．

③ 正如之前已经定义的，$\mathrm{D}\boldsymbol{X}$ 已经用来表示 Jacobi 矩阵．这里还是请先暂时将它视为一种记号，其具体形式将在下一小节给出．

1.2.2 Jacobi 矩阵

下面我们研究 $\mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right) \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^m)$ 的表达形式．由于 $\boldsymbol{h} \in \mathbb{R}^m$，所以

$$\boldsymbol{h} = \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix} = h^1 \boldsymbol{e}_1 + \cdots + h^i \boldsymbol{e}_i + \cdots + h^m \boldsymbol{e}_m. \tag{1.20}$$

另一方面，$\mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right) \in \mathscr{L}(\mathbb{R}^m, \mathbb{R}^m)$ 具有线性性：

$$\forall\, \alpha,\, \beta \in \mathbb{R} \text{ 和 } \tilde{\boldsymbol{h}},\, \hat{\boldsymbol{h}} \in \mathbb{R}^m, \quad \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\alpha \tilde{\boldsymbol{h}} + \beta \hat{\boldsymbol{h}}\right) = \alpha\, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\tilde{h}\right) + \beta\, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\hat{h}\right). \tag{1.21}$$

这样就有

$$\begin{aligned} \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)(\boldsymbol{h}) &= \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(h^1 \boldsymbol{e}_1 + \cdots + h^i \boldsymbol{e}_i + \cdots + h^m \boldsymbol{e}_m\right) \\ &= h^1\, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_1\right) + \cdots + h^i\, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_i\right) + \cdots + h^m\, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_m\right) \end{aligned} \tag{1.22}$$

注意到 $h^i \in \mathbb{R}$ 以及 $\mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_i\right) \in \mathbb{R}^m$，因而该式可以用矩阵形式表述：

$$= \left[\mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_1\right), \cdots, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_m\right)\right] \begin{bmatrix} h^1 \\ \vdots \\ h^m \end{bmatrix}. \tag{1.23}$$

最后一步要用到分块矩阵的思想：左侧的矩阵为 1“行”m 列，每一“行”是一个 m 维列向量；右侧的矩阵（向量）则为 m 行 1 列．两者相乘，得到 1“行”1 列的矩阵（当然实际为 m 行），即之前的 (1.22) 式．在线性代数中，$m \times m$ 的矩阵 $\begin{bmatrix} \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_1\right) & \cdots & \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_m\right) \end{bmatrix}$ 通常称为**变换矩阵**（也叫**过渡矩阵**）.

接下来要搞清楚变换矩阵的具体形式．取

$$\boldsymbol{h} = \left[0, \cdots, \lambda, \cdots, 0\right]^{\mathsf{T}} = \lambda\, \boldsymbol{e}_i \in \mathbb{R}^m, \tag{1.24}$$

即除了 $\boldsymbol{h}$ 的第 i 各元素为 λ 外，其余元素均为 0（$\lambda \neq 0$）．因而有 $\|\boldsymbol{h}\|_{\mathbb{R}^m} = \lambda$．代入可微性的定义 (1.19) 式，可得

$$\begin{aligned} &\boldsymbol{X}\left(\boldsymbol{x}_0 + \boldsymbol{h}\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right) = \boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right) \\ &= \left[\mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_1\right), \cdots, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_i\right), \cdots, \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_m\right)\right]\left[0, \cdots, \lambda, \cdots, 0\right]^{\mathsf{T}} + o(\lambda) \\ &= \lambda \cdot \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_i\right) + o(\lambda). \end{aligned} \tag{1.25}$$

由于 λ 是非零实数，故可以在等式两边同时除以 λ 并取极限：

$$\lim_{\lambda \to 0} \frac{\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right)}{\lambda} = \mathrm{D}\boldsymbol{X}\left(\boldsymbol{x}_0\right)\left(\boldsymbol{e}_i\right), \tag{1.26}$$

这里的 $o(\lambda)$ 根据其定义自然趋于 0．该式左侧极限中的分子部分，是自变量 $\boldsymbol{x}$ 第 i 个分量的变化所引起因变量的变化；而分母，则是自变量第 i 个分量的变化大小．我们引入下面的记号：

$$\frac{\partial \boldsymbol{X}}{\partial x^i}\left(\boldsymbol{x}_0\right) \coloneqq \lim_{\lambda \to 0} \frac{\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right)}{\lambda} \in \mathbb{R}^m, \tag{1.27}$$

它表示因变量 $\boldsymbol{X} \in \mathbb{R}^m$ 作为一个整体，相对于自变量 $\boldsymbol{x} \in \mathbb{R}^m$ 第 i 个分量 $x^i \in \mathbb{R}$ 的“变化率”，即 $\boldsymbol{X}$ 关于 x^i（在 $\boldsymbol{x}_0$ 处）的**偏导数**．由于我们没有定义向量的除法，因此自变量作为整体所引起因变量的变化，是没有意义的．利用偏导数的定义，可有

$$\begin{aligned}&\left[\mathsf{D}\boldsymbol{X}(\boldsymbol{x}_0)(\boldsymbol{e}_1),\ \cdots,\ \mathsf{D}\boldsymbol{X}(\boldsymbol{x}_0)(\boldsymbol{e}_i),\ \cdots,\ \mathsf{D}\boldsymbol{X}(\boldsymbol{x}_0)(\boldsymbol{e}_m)\right]\\ =&\left[\frac{\partial \boldsymbol{X}}{\partial x^1}(\boldsymbol{x}_0),\ \cdots,\ \frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x}_0),\ \cdots,\ \frac{\partial \boldsymbol{X}}{\partial x^m}(\boldsymbol{x}_0)\right] \in \mathbb{R}^{m\times m}.\end{aligned} \tag{1.28}$$

下面给出 $\partial \boldsymbol{X} / \partial x^i\ (\boldsymbol{x}_0)$ 的计算式．根据定义，有

$$\begin{aligned}\frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x}_0) &\coloneqq \lim_{\lambda\to 0}\frac{\boldsymbol{X}(\boldsymbol{x}_0+\lambda\boldsymbol{e}_i)-\boldsymbol{X}(\boldsymbol{x}_0)}{\lambda} \in \mathbb{R}^m\\ &= \lim_{\lambda\to 0}\frac{1}{\lambda}\cdot\left(\begin{bmatrix}X^1\\ \vdots\\ X^m\end{bmatrix}(\boldsymbol{x}_0+\lambda\boldsymbol{e}_i)-\begin{bmatrix}X^1\\ \vdots\\ X^m\end{bmatrix}(\boldsymbol{x}_0)\right)\\ &= \lim_{\lambda\to 0}\begin{bmatrix}\dfrac{X^1(\boldsymbol{x}_0+\lambda\boldsymbol{e}_i)-X^1(\boldsymbol{x}_0)}{\lambda}\\ \vdots\\ \dfrac{X^m(\boldsymbol{x}_0+\lambda\boldsymbol{e}_i)-X^m(\boldsymbol{x}_0)}{\lambda}\end{bmatrix}.\end{aligned} \tag{1.29}$$

向量极限存在的充要条件是各分量极限均存在，即存在

$$\frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}_0) \coloneqq \lim_{\lambda\to 0}\frac{X^\alpha(\boldsymbol{x}_0+\lambda\boldsymbol{e}_i)-X^\alpha(\boldsymbol{x}_0)}{\lambda} \in \mathbb{R}, \tag{1.30}$$

其中的 $\alpha = 1, \cdots, m$．这其实就是我们熟知的多元函数偏导数的定义．用它来表示向量值映照的偏导数，可有

$$\frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x}_0) = \begin{bmatrix}\dfrac{\partial X^1}{\partial x^i}(\boldsymbol{x}_0)\\ \vdots\\ \dfrac{\partial X^m}{\partial x^i}(\boldsymbol{x}_0)\end{bmatrix} = \sum_{\alpha=1}^m \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}_0)\,\boldsymbol{e}_\alpha. \tag{1.31}$$

向量值映照 $\boldsymbol{X}$ 关于 x^i 的偏导数，从代数的角度来看，是 Jacobi 矩阵的第 i 列；从几何的角度来看，则是物理域中 x^i 线的切向量；从计算的角度来看，又是（该映照）每个分量偏导数的组合．

现在我们重新回到 Jacobi 矩阵．情况已经十分明了：只需把之前获得的各列并起来，就可以得到完整的 Jacobi 矩阵．于是

$$\begin{aligned}\mathsf{D}\boldsymbol{X}(\boldsymbol{x}_0)(\boldsymbol{h}) &= \left[\frac{\partial \boldsymbol{X}}{\partial x^1},\ \cdots,\ \frac{\partial \boldsymbol{X}}{\partial x^m}\right](\boldsymbol{x}_0)(\boldsymbol{h})\\ &= \begin{bmatrix}\dfrac{\partial X^1}{\partial x^1} & \cdots & \dfrac{\partial X^1}{\partial x^m}\\ \vdots & \ddots & \vdots\\ \dfrac{\partial X^m}{\partial x^1} & \cdots & \dfrac{\partial X^m}{\partial x^m}\end{bmatrix}(\boldsymbol{x}_0)\cdot\begin{bmatrix}h^1\\ \vdots\\ h^m\end{bmatrix}.\end{aligned} \tag{1.32}$$

这与 1.1.2 小节中 (1.11) 式给出的定义是完全一致的.

1.2.3 偏导数的几何意义

这一小节中，我们要回过头来，澄清向量值映照偏导数的几何意义.

如图 1.2，$\boldsymbol{X}(\boldsymbol{x})$ 是定义域空间 $\mathfrak{D}_{\boldsymbol{x}} \subset \mathbb{R}^m$ 到值域空间 $\mathfrak{D}_{\boldsymbol{X}} \subset \mathbb{R}^m$ 的向量值映照. 在定义域空间 $\mathfrak{D}_{\boldsymbol{x}}$ 中，过点 $\boldsymbol{x}_0$ 作一条平行于 x^i 轴的直线，称为 x^i-线. x^i 轴定义了向量 $\boldsymbol{e}_i$，因而 x^i-线上的任意一点均可表示为 $\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i$，其中 $\lambda \in \mathbb{R}$.

Images/Vector-Value_Mapping.PNG

图 1.2: 向量值映照偏导数的几何意义

在 $\boldsymbol{X}(\boldsymbol{x})$ 的作用下，点 $\boldsymbol{x}_0$ 被映照到 $\boldsymbol{X}\left(\boldsymbol{x}_0\right)$，而 $\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i$ 则被映照到了 $\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i\right)$. 这样一来，$x^i$-线也就被映照到了值域空间 $\mathfrak{D}_{\boldsymbol{X}}$ 中，成为一条曲线.

根据前面的定义，当 $\lambda \to 0$ 时，

$$\frac{\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_i\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right)}{\lambda} \to \frac{\partial \boldsymbol{X}}{\partial x^i}\left(\boldsymbol{x}_0\right). \tag{1.33}$$

对应到图 1.2 中，就是 x^i-线（值域空间中）在 $\boldsymbol{X}\left(\boldsymbol{x}_0\right)$ 处的**切向量**.

完全类似，在定义域空间 $\mathfrak{D}_{\boldsymbol{x}}$ 中，过点 $\boldsymbol{x}_0$ 作出 x^j-线（自然是平行于 x^j 轴），其上的点可以表示为 $\boldsymbol{x}_0 + \lambda \boldsymbol{e}_j$. 映射到值域空间 $\mathfrak{D}_{\boldsymbol{X}}$ 上，则成为 $\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_j\right)$. 很显然，

$$\frac{\partial \boldsymbol{X}}{\partial x^j}\left(\boldsymbol{x}_0\right) = \frac{\boldsymbol{X}\left(\boldsymbol{x}_0 + \lambda \boldsymbol{e}_j\right) - \boldsymbol{X}\left(\boldsymbol{x}_0\right)}{\lambda} \tag{1.34}$$

就是 x^j-线在 $\boldsymbol{X}\left(\boldsymbol{x}_0\right)$ 处的切向量. 在定义域空间中，x^i-线作为直线共有 m 条，它们之间互相垂直. 作用到值域空间后，这样的 x^i-线尽管变为了曲线，但仍为 m 条. 相应的切向量，自然也有 m 个.

1.3 局部基

这里的讨论基于曲线坐标系（即微分同胚）$\boldsymbol{X}(\boldsymbol{x}) \in \mathscr{C}^p\left(\mathfrak{D}_{\boldsymbol{x}}; \mathfrak{D}_{\boldsymbol{X}}\right)$.

我们已经知道，$\boldsymbol{X}(\boldsymbol{x})$ 的 Jacobi 矩阵可以表示为

$$\mathsf{D}\boldsymbol{X}(\boldsymbol{x}) = \left[\frac{\partial \boldsymbol{X}}{\partial x^1}, \cdots, \frac{\partial \boldsymbol{X}}{\partial x^i}, \cdots, \frac{\partial \boldsymbol{X}}{\partial x^m}\right](\boldsymbol{x}) \in \mathbb{R}^{m\times m}, \tag{1.35}$$

式中的

$$\frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x}) = \lim_{\lambda\to 0} \frac{\boldsymbol{X}\left(\boldsymbol{x} + \lambda\, \boldsymbol{e}_i\right) - \boldsymbol{X}(\boldsymbol{x})}{\lambda}. \tag{1.36}$$

在参数域 $\mathfrak{D}_{\boldsymbol{x}}$ 中作出 x^i-线．映照到物理域后，它变成一条曲线，我们仍称之为 x^i-线．1.2.3 小节已经说明，(1.36) 式表示物理域中 x^i-线的切向量．在张量分析中，我们通常把它记作 $\boldsymbol{g}_i(\boldsymbol{x})$．

由于微分同胚要求是双射，因而 Jacobi 矩阵

$$\mathsf{D}\boldsymbol{X}(\boldsymbol{x}) = \left[\boldsymbol{g}_1, \cdots, \boldsymbol{g}_i, \cdots, \boldsymbol{g}_m\right](\boldsymbol{x}) \in \mathbb{R}^{m\times m} \tag{1.37}$$

必须是非奇异的．这等价于

$$\left\{\boldsymbol{g}_i(\boldsymbol{x}) = \frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x})\right\}_{i=1}^{m} \subset \mathbb{R}^m \tag{1.38}$$

线性无关．由此，它们可以构成 $\mathbb{R}^m$ 上的一组基．

用任意的 $\boldsymbol{x} \in \mathfrak{D}_{\boldsymbol{x}}$ 均可构建一组基．但选取不同的 $\boldsymbol{x}$，将会使所得基的取向有所不同．因而这种基称为**局部协变基**．和之前一样，我们用“协变”表示指标在下方．

有了局部协变基 $\{\boldsymbol{g}_i(\boldsymbol{x})\}_{i=1}^m$，根据 交叉引用第一章（已注释掉） 小节中的讨论，必然唯一存在与之对应的**局部逆变基** $\{\boldsymbol{g}^i(\boldsymbol{x})\}_{i=1}^m$，满足

$$\left[\boldsymbol{g}^1(\boldsymbol{x}), \cdots, \boldsymbol{g}^m(\boldsymbol{x})\right]^{\mathsf{T}} \left[\boldsymbol{g}_1(\boldsymbol{x}), \cdots, \boldsymbol{g}_m(\boldsymbol{x})\right] = \begin{bmatrix} (\boldsymbol{g}^1)^{\mathsf{T}} \\ \vdots \\ (\boldsymbol{g}^m)^{\mathsf{T}} \end{bmatrix}(\boldsymbol{x}) \cdot \mathsf{D}\boldsymbol{X}(\boldsymbol{x}) = \boldsymbol{I}_m. \tag{1.39}$$

下面我们来寻找逆变基 $\{\boldsymbol{g}^i(\boldsymbol{x})\}_{i=1}^m$ 的具体表示．考虑到①

$$\boldsymbol{X}\left(\boldsymbol{x}(\boldsymbol{X})\right) = \boldsymbol{X} \in \mathbb{R}^m, \tag{1.40}$$

并利用复合映照可微性定理，可知

$$\mathsf{D}\boldsymbol{X}\left(\boldsymbol{x}(\boldsymbol{X})\right) \cdot \mathsf{D}\boldsymbol{x}(\boldsymbol{X}) = \boldsymbol{I}_m, \tag{1.41}$$

即有

$$\mathsf{D}\boldsymbol{x}(\boldsymbol{X}) = (\mathsf{D}\boldsymbol{X})^{-1}\left(\boldsymbol{x}(\boldsymbol{X})\right). \tag{1.42}$$

于是

$$\begin{bmatrix} (\boldsymbol{g}^1)^{\mathsf{T}} \\ \vdots \\ (\boldsymbol{g}^m)^{\mathsf{T}} \end{bmatrix}(\boldsymbol{x}) = (\mathsf{D}\boldsymbol{X})^{-1}(\boldsymbol{x}) = \mathsf{D}\boldsymbol{x}(\boldsymbol{X}) = \begin{bmatrix} \dfrac{\partial x^1}{\partial X^1} & \cdots & \dfrac{\partial x^1}{\partial X^m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x^m}{\partial X^1} & \cdots & \dfrac{\partial x^m}{\partial X^m} \end{bmatrix}(\boldsymbol{X}). \tag{1.43}$$

① 这里的几步推导在 1.1.2 小节中也有所涉及．

这样我们就得到了局部逆变基的具体表示（注意转置）：

$$\boldsymbol{g}^i(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial x^i}{\partial X^1} \\ \vdots \\ \dfrac{\partial x^i}{\partial X^m} \end{bmatrix} (\boldsymbol{X}) = \sum_{\alpha=1}^{m} \frac{\partial x^i}{\partial X^\alpha}(\boldsymbol{X})\, \boldsymbol{e}_\alpha . \tag{1.44}$$

定义标量场 $f(\boldsymbol{x})$ 的**梯度**为

$$\nabla f(\boldsymbol{x}) \triangleq \sum_{\alpha=1}^{m} \frac{\partial f}{\partial x^\alpha}(\boldsymbol{x})\, \boldsymbol{e}_\alpha , \tag{1.45}$$

则局部逆变基又可以表示成

$$\boldsymbol{g}^i(\boldsymbol{x}) = \nabla x^i(\boldsymbol{X}). \tag{1.46}$$

此处的梯度实际上就是我们熟知的三维情况在 m 维下的推广.

在 1.2.3 小节中已经指出，局部协变基的几何意义是 x^i-线的切向量. 现在，我们来讨论局部逆变基的几何意义.

Images/Local_Basis.PNG

图 1.3: 局部逆变基的几何意义

如图 1.3 所示，在参数空间域中，过点 $\boldsymbol{x}$ 作垂直于 x^i 轴的平面，记为 x^i-面. 在 x^i-面上，自然有 $x^i = \text{const}$. 映照到物理空间后，x^i-面变为一个曲面，其上仍有 $x^i(\boldsymbol{X}) = \text{const.}$，即它是一个等值面. 等值面的梯度方向显然与该曲面的法向相同. 因此，局部逆变基 $\boldsymbol{g}^i(\boldsymbol{x})$ 的几何意义就是 x^i-面的**法向量**.

现在来验证一下**对偶关系**.

$$\left(\boldsymbol{g}_i(\boldsymbol{x}), \boldsymbol{g}^j(\boldsymbol{x})\right)_{\mathbb{R}^m} = \left(\frac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x}), \nabla x^j(\boldsymbol{X})\right)_{\mathbb{R}^m}$$

$$= \left(\sum_{\alpha=1}^{m} \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x})\, \boldsymbol{e}_\alpha,\ \sum_{\beta=1}^{m} \frac{\partial x^j}{\partial X^\beta}(\boldsymbol{X})\, \boldsymbol{e}_\beta \right)_{\mathbb{R}^m}$$

利用内积的线性性，有

$$\begin{aligned}
&= \sum_{\alpha=1}^{m} \sum_{\beta=1}^{m} \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}) \frac{\partial x^j}{\partial X^\beta}(\boldsymbol{X}) \cdot \left(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta\right)_{\mathbb{R}^m} \\
&= \sum_{\alpha=1}^{m} \sum_{\beta=1}^{m} \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}) \frac{\partial x^j}{\partial X^\beta}(\boldsymbol{X}) \cdot \delta_{\alpha\beta}
\end{aligned}$$

合并掉指标 β，可得

$$\begin{aligned}
&= \sum_{\alpha=1}^{m} \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}) \frac{\partial x^j}{\partial X^\alpha}(\boldsymbol{X}) \\
&= \sum_{\alpha=1}^{m} \frac{\partial x^j}{\partial X^\alpha}(\boldsymbol{X}) \frac{\partial X^\alpha}{\partial x^i}(\boldsymbol{x}). \qquad (1.47)
\end{aligned}$$

最后一步求和号中的第一项位于 Jacobi 矩阵 $\mathrm{D}\boldsymbol{x}(\boldsymbol{X})$ 的第 j 行第 α 列，而第二项位于 $\mathrm{D}\boldsymbol{X}(\boldsymbol{x})$ 的第 α 行第 i 列，因此关于 α 的求和结果便是乘积矩阵的第 j 行第 i 列．根据式 (1.41)，这两个 Jacobi 矩阵的乘积为单位阵，所以有

$$\left(\boldsymbol{g}_i(\boldsymbol{x}), \boldsymbol{g}^j(\boldsymbol{x})\right)_{\mathbb{R}^m} = \delta_i^j. \qquad (1.48)$$

总结一下我们得到的结果．对于体积形态的连续介质，存在着

$$\begin{cases}
\text{局部协变基：} & \left\{\boldsymbol{g}_i(\boldsymbol{x}) \triangleq \dfrac{\partial \boldsymbol{X}}{\partial x^i}(\boldsymbol{x})\right\}_{i=1}^{m}, \\
\text{局部逆变基：} & \left\{\boldsymbol{g}^i(\boldsymbol{x}) \triangleq \nabla x^i(\boldsymbol{X})\right\}_{i=1}^{m},
\end{cases}$$

它们满足对偶关系

$$\left(\boldsymbol{g}_i(\boldsymbol{x}), \boldsymbol{g}^j(\boldsymbol{x})\right)_{\mathbb{R}^m} = \delta_i^j. \qquad (1.50)$$

这样，在研究连续介质中的一个点时，我们就有三种基可以使用：局部协变基、局部逆变基，当然还有典则基 $\left\{\boldsymbol{e}_i\right\}_{i=1}^{m}$.

1.4　标架运动方程