数学编程与制图 python 与 PostScript

龙冰、钟尔杰、杨益民等

2021-10-17

前言

我们编写本书，暂定书名《数学编程与制图 - python 与 PostScript》...

记得上世纪八十年代初我们在四川大学数学系学习，1958 年设立的计算机专业还是数学系的一部分，分出去成立计算机系并壮大成计算机（及软件）学院是以后的事了。计算机程序设计是数学系的必修课，但由于各种条件的限制，课程开设成了摆设，猜想大多数同学认为是浪费时间。我们在学习的语言是 ALGOL 60，编程只能写在作业本上。学校一台巨大的当时高级的计算机放在一大房间里精心维护着，工作人员穿着白大掛，地板打着蜡。可是我我们平时没有机会接触。只是到了期末考试，我们才有机会进机房。开卷考试题是解二次方程 $ax^2 + bx + c = 0$。由专门操作人员在纸带上打孔，我们也不懂打孔，只听机器卡嚓卡嚓响，我们就都稀里糊涂地通过了。由于计算机技术的巨大进步，任何一台个人电脑的都会拥有比那台宝贝大不是知多少倍的计算能力。在个人电脑编程也是件再容易不过的事儿了，普通中学数学教育足以

龙冰、钟尔杰、杨益民
2021 年 10 月

目录

第一部分　开源软件安装

流变学（rheology）在牛津语言网站上的释义是：研究物质的变形和流动——特别是液体的非牛顿流动和固体的塑性流动——的科学（The branch of science that deals with the deformation and flow of matter, esp. the non-Newtonian flow of liquids and the plastic flow of solids）。

流变学是一门力学，属于经典力学范畴。力学包括静力学（statics）、运动学（kinematics）和动力学 (dynamics）三种问题。静力学集中考虑物体处于力的平衡之下的问题。运动学定量描述物体的位置和状态随时间的变化关系。动力学研究和分析物体运动状态改变的原因。[1]p. 1

在经典力学的运动学中，为了描述物体的位置，我们需要先选择参考系（frame of reference），建立坐标系（coordinates），使用位置向量来表示质点的位移。我们假设时间是连续的，把质点的位移写成时间的函数，在可导的前提下，其速度和加速度分别是位移的一阶和二阶时间导数[1] 附录 A,p. 422（如图I.0.1所示）：

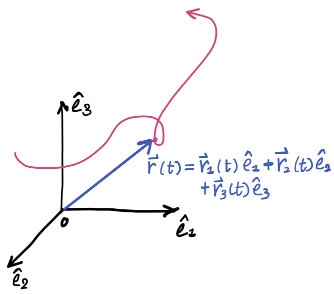

图 I.0.1: 质点运动学

$$
\begin{aligned}
&\mathbf{r} = \mathbf{r}\left(t\right) = r_1\widehat{\mathbf{e}}_1 + r_2\widehat{\mathbf{e}}_2 + r_3\widehat{\mathbf{e}}_3,\\
&\mathbf{v} = \frac{d\mathbf{r}}{dt}, \mathbf{a} = \frac{d^2\mathbf{r}}{dt^2}
\end{aligned}
$$

经典力学的动力学遵循牛顿运动定律，引入了“力”的概念作为物体运动状态变化的原因（第一定律），并建立了定量关系（第二定律），以及一个守恒律（第三定律）。牛顿运动定律在且仅在惯性系下成立。作为质点的动力学，在选定惯性系下，

$$
\mathbf{F} = m\frac{d\mathbf{v}}{dt}
$$

其中 m 是质点的质量，是惯性（改变物定运动状态的代价）大小的量度。上式可以推广到多个质点的体系，

对于一个可以发生形变和流动的物体，上述的质点运动学和动力学在描述形变和流动问题中是不够用的。

以下是两种适用的力学：

- 连续介质力学（continuum mechanics）：把宏观物体视为可无限微分的连续体。从热力学基本定律和牛顿运动定律（惯性系）出发，建立物体运动与形变的本构关系。
- 统计力学（statistical mechanics）：考虑组成物质的微观粒子的运动规律，基于统计力学的假设推出宏观物体的本构关系。高分子物理中，从链统计出发建立的粘流模型和橡胶弹性模型就是统计力学在高分子物理中的应用。

描述宏观物体形变规律的语言是连续介质力学。统计力学只是进一步给出微观原因，它关于宏观规律的预测仍旧要使用连续介质力学的语言来表述。

在连续介质力学中，一个物体的独特力学性质，是由它的本构关系/方程（constitutive relation/equation）来描述的。在一般物理学中，本构关系是指描述物质对外场激励的响应的定量关系：外场 $\mathcal{S}(\mathbf{r},t) \to$ 材料特性函数 $\mathcal{G}(\Delta\mathbf{x},\Delta t) \to$ 材料响应 $\mathcal{R}(\mathbf{x},t)$。其中，$\mathbf{r}$ 是空间任一位置，$\mathbf{x}$ 是处于材料内部的某位置。$\mathcal{G}(\Delta\mathbf{x},\Delta t)$ 是材料本身的性质，称为物料函数（material function）。我们留意到，$\mathcal{S}$ 和 $\mathcal{R}$ 都是场函数。物料函数是对材料观测的时间和空间间隔的函数，在数学形式上也是场函数[2]§ 9.7,p.192。以下我们通过两个例题来理解本构关系的概念。这两个例题都是以前的课程中介绍过的，请注意参见例中标注的引用。

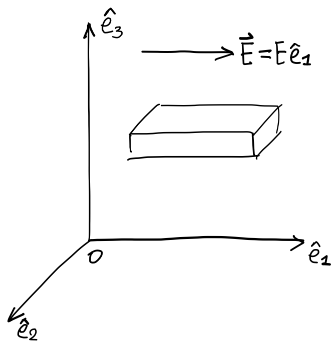

图 I.0.2: 例I.0.1

例 I.0.1. 此例中，外场是稳恒电场 $\mathbf{E}(\mathbf{r},t)=\mathbf{E}=E\widehat{\mathbf{e}}_1$，材料的物性函数是电极化率 $\chi(\Delta\mathbf{x},\Delta t)$（如图*I.0.2*所示）。对于各向同性均匀材料 $\chi = const$。材料的响应是电极化强度密度 $\mathbf{P}(\mathbf{x},t)$，表示材料内 $\mathbf{x}$ 处单位体积的电极化强度，即 $\mathbf{P}(\mathbf{x},t)\,dV$ 表示 $\mathbf{x}$ 处的微体积 dV 的电极化强度。实验表明，当电场不太大时*[3]§18.2.3,p. 57*，

$$\mathbf{P}(\mathbf{x},t) = \chi\varepsilon_0\mathbf{E}(\mathbf{r}=\mathbf{x},t)$$

其中 ε_0 是真空电容率。故本例中

$$\mathbf{P}(\mathbf{x},t) = \chi\varepsilon_0 E\widehat{\mathbf{e}}_1$$

这是该材料的电极化本构方程。

例 I.0.2. 此例中外场是一个梯度场，例如温度场 $T(\mathbf{r},t) = kr_1\widehat{\mathbf{e}}_1, \nabla T(\mathbf{r},t) = k\widehat{\mathbf{e}}_1$。材料的物料函数是传输物料函数，例如热导率 $\kappa = const$（假设各向同性均匀材料）。材料的响应是流（*flux*），即单位时间经过 $\mathbf{x}$ 处面积元 $\mathbf{n}dS$ 的物理量。例如热流：$\mathbf{q}(\mathbf{x},t)$，$\mathbf{q}(\mathbf{x},t)\cdot\mathbf{n}dS$ 是单位时间通过面积元 $\mathbf{n}dS$ 的热量。$\mathbf{q}$ 的方向与等温面垂直并指向温度减小的方向。实验表明，当温度梯度不太大时，热流满足傅立叶热传导定律*[4]§1.1.3.1,p. 9*：

$$\mathbf{q}(\mathbf{x},t) = -\kappa\nabla T(\mathbf{r}=\mathbf{x},t) = -k\kappa\widehat{\mathbf{e}}_1$$

这是该材料的热传导本构方程。

流变学中的本构关系主要是应力与应变（或应变速率）的关系。

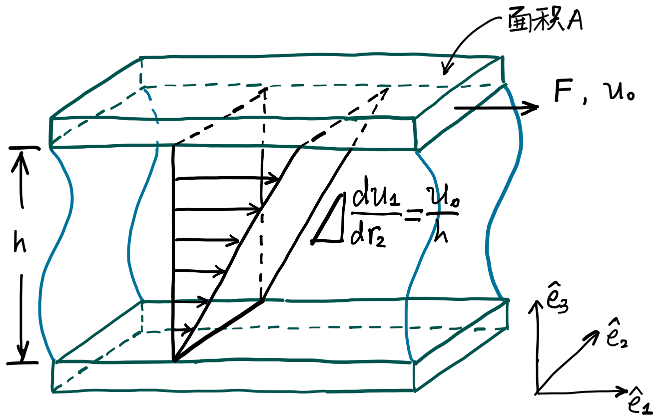

图 I.0.3: 例I.0.3

例 I.0.3 (简单剪切下的牛顿流体本构关系). 如图*I.0.3*所示的间距为 h、面积足够大的两平行板间充满粘度为 μ 牛顿流体。按照下板固定建立参考系，建立如图所示的坐标系。上板以速率 u_0 往平行于下板的方向运动。假设流体与平板接触的界面处相对流速为零；液体的流速 $\mathbf{u}(\mathbf{r},t)=u_0r_2\widehat{\mathbf{e}}_1/h$。若上板面积为 A、所受的拖曳力大小为 F（方向与上板运动速度相同）。F 与 u_0 的关系是

$$\tau=\frac{F}{A}=\mu\frac{u_0}{h}$$

这是一个本构关系：u_0 是外场，F 是材料响应，μ 是物料函数。

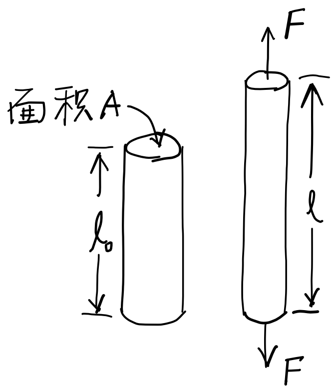

图 I.0.4: 例I.0.4

例 I.0.4 (轴向拉伸下的虎克固体本构关系). 如图*I.0.4*所示的虎克固体直圆杆，弹性模量为 E，初始截面积为 A，初始长度为 l_0。在杆的轴向受到大小为 F 的力时，杆的长度为 l，则 F 与

l 的关系是

$$\sigma = \frac{F}{A} = E\frac{l-l_0}{l_0}$$

这是一个本构关系：F 是外场，l 是材料响应，E 是物料函数。

上面两例中的本构方程形式只适用于特定形变方式。要完整描述一般形变过程的运动学和动力学，需要使用向量和张量以及它们的代数和微积分数学基础。

实际物体的运动和形变是在一定的初条件（initial condition）和边界条件（boundary condition）下，遵循守恒律（balance law）和本构关系的共同限定下才被具体确定的。在采用质点来简化的问题中，这些要素往往是隐含的条件。下面以一个例子说明，在解决一个具体流动问题的过程中，初条件、边界条件、守恒律和本构关系的角色。

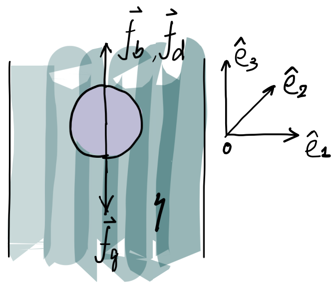

图 I.0.5: 落球粘度计原理图

例 I.0.5 (落球粘度计). *半径为 a 的球浸没在粘度为 η 的牛顿流体中。以液体的容器为参考系，建立如图I.0.5所示的坐标系。球的运动满足牛顿第二定律，即：$\mathbf{F} = md\mathbf{v}/dt$（动量守恒），其中 $m = \text{const}$ 是球的质量（质量守恒）。此时，我们并不能确定球的具体运动。*

设 $t = 0$ 时刻球速 $\mathbf{v} = \mathbf{0}$（初条件），液体的密度为 ρ_{w}、球的密度为 ρ_{s}，则球受到的重力是 $\mathbf{f}_{\text{g}} = -4\pi\rho_{\text{s}} g a^3 \widehat{\mathbf{e}}_3/3$，浮力 $\mathbf{f}_{\text{b}} = 4\left(\rho_{\text{w}} - \rho_{\text{s}}\right)\pi g a^3 \widehat{\mathbf{e}}_3$，液体对球的拖曳力 $\mathbf{f}_{\text{d}} = -k\pi\eta a\mathbf{v}$（本构关系）。其中，当球面处的相对流速总为零时，$k = 6$（边界条件）。由牛顿第二定律，

$$\frac{4/3}{\pi}\rho_{\text{s}} g a^3 \frac{d\mathbf{v}}{dt} = \frac{4}{3}\pi g a^3\left(\rho_{\text{w}} - \rho_{\text{s}}\right) - 6\pi\eta a\mathbf{v}$$

解方程得：

$$\begin{aligned}
&\mathbf{v}\left(t\right) = v_1\left(t\right)\widehat{\mathbf{e}}_1 + v_2\left(t\right)\widehat{\mathbf{e}}_2 + v_3\left(t\right)\widehat{\mathbf{e}}_3 \\
&v_1\left(t\right) = v_2\left(t\right) \equiv 0 \\
&v_3\left(t\right) = \frac{2\left(\rho_{\text{w}} - \rho_{\text{s}}\right) g a^3}{9\eta}\left(1 - e^{\frac{-9\eta t}{2\rho_{\text{s}} a^2 g}}\right)
\end{aligned}$$

当 $t \to \infty$ 时，

$$v_3(t) \equiv v_\infty = \frac{2(\rho_\mathrm{w} - \rho_\mathrm{s})ga^2}{9\eta}$$

$$\eta = \frac{2(\rho_\mathrm{w} - \rho_\mathrm{s})ga^2}{9v_\infty}$$

上例中的考察对象（球）是刚体，在运动过程中形状不变。在流变学中，我们要描述物体在形变过程中如何保持质量、动量和能量守恒，这需要借助场函数的一系列积分定理，相关的知识将会在数学准备部分介绍。

第二部分 python

II.1 待定

集合是近代数学的基本语言。连续介质力学中的大量概念都依赖集合来定义。而集合本身却是难以借助更基础的概念进行定义的“最原始概念”。因此在下述集合的定义只能假定每一位读者都能一致地理解。

定义 II.1.1 (集合)**.** 集合（*set*）是具有某种特性的事物的整体。构成集合的事物或对象称为元素（*element*）或成员（*member*）。集合还必须满足:

- 无序性：一个集合中，每个元素的地位是相同的，元素之间是无序的
- 互异性：一个集合中，任何两个元素都不相同，即每个元素只出现一次
- 确定性：给定一个集合及一个事物，该事物要么属于要么不属于该集合，不允许模棱两可。

我们已经学习过，集合的表示方式有列举法和描述法。

例 II.1.1.

$$
\begin{aligned}
&A = \{2, 4, 6, 8\} \\
&B = \{n|n\text{是偶数且}1 < n < 9\} \\
&1 \notin A, 2 \in B
\end{aligned}
$$

以下定义和定理是关于集合之间的包含与被包含关系的概念，读者应该已经熟悉。

定义 II.1.2 (子集、包含、集合的相等)**.** 如果只要 $a \notin B$，则必有 $a \in A$，则称 B 是 A 的子集，或称 A 包含 B，记为 $B \subseteq A$，$A \supseteq B$。如果 $A \subseteq B$ 且 $B \subseteq A$ 则称 $A = B$。

定理 II.1.1. *1.* $S \subseteq S, \forall S$，

2. 若 $A \subseteq B$ 且 $B \subseteq C$ 则 $A \subseteq C$。

定义 II.1.3 (真子集、真包含)**.** 若 $B \subseteq A$ 且 $B \neq A$，则称 B 是 A 的真子集，或称 A 真包含 B，记为 $B \subset A$，$A \supset B$。

定理 II.1.2. *1.* $\forall S$，$S \subset S$ 均不成立。

2. 若 $A \subset B$ 则 $B \subset A$ 不成立。

3. 若 $A \subset B$ 且 $B \subset C$，则 $A \subset C$。

定义 II.1.4 (空集)**.** 空集 $\emptyset = \{a|a \neq a\}$

可验证如下定理:

定理 II.1.3. 1. $\emptyset \subseteq A, \forall A$。
2. $\emptyset \subset A, \forall A \neq \emptyset$。

集合与集合之间可以通过并集和交集的“运算”得出新的集合。

定义 II.1.5 (并集)**.** 若 $a \in A$ 或 $a \in B, \forall a, A, B$，则 $a \in A \cup B$，称 A 与 B 的并集。

可验证:

定理 II.1.4. 并集关系具有以下性质:
1. 结合律: $A \cup (B \cup C) = (A \cup B) \cup C$。
2. 交换律: $A \cup B = B \cup A$。
3. $\emptyset \cup A = A, \forall A$。

定义 II.1.6 (交集)**.** 若 $a \in A$ 且 $a \in B, \forall a, A, B$，则 $a \in A \cap B$，称 A 与 B 的交集。

定义 II.1.7 (笛卡尔积)**.** 两个集合 A 和 B 的笛卡尔积是所有有序对 $(a, b), a \in A, b \in B$ 的集合，即 $A \times B = \{(a, b)\,|a \in A且b \in B\}$

我们可以推广这一定义至任意个集合的笛卡尔积。集合 $A_1, A_2, \cdots, A_n$ 的笛卡尔积是 $A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \cdots, a_n)\,|a_i \in A_i\}$，一个有序 n 元组的集合。如果 $A_1 = A_2 = \cdots = A_n = A$，则记 $A_1 \times A_2 \times \cdots \times A_n = A^n$。例如，$\mathbb{R}$ 是实数集，$\mathbb{R}^n$ 是 n 元有序实数组的集合，在解析几何中常用以表示 n 维欧几里德空间中一个点的坐标。

一个集合的元素可以与另一个集合的元素建立对应关系。我们主要关心的是满足某种规定的对应关系，称为映射，定义如下。

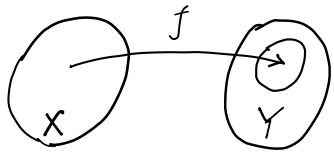

图 II.1.1: 映射

定义 II.1.8 (映射)**.** (如图II.1.1所示)从集合 X 到集合 Y 的映射(*mapping*)f，记为 $f : X \to Y$，是 X 的所有元素与 Y 的部分或所有元素之间的对应关系，且每个 X 的元素只对应 Y 的一个元素。如果 $y \in Y$ 是 $x \in X$ 通过映射 f 的对应，则可写成 $y = f(x)$，并称 $f(x)$ 是 x 的

像（*image*）。按上述定义，x 只有一个像。X 称为该映射 f 的定义域（*domain*），记为 $\mathrm{dom}f$，Y 称为该映射 f 的陪域（*codomain*）或到达域（*target domain*）。x 的像 $f(x)$ 是 Y 的子集，称为该映射 f 的值域（*range*），记为 $\mathrm{ran}f$。

当映射的陪域是实数集 $\mathbb{R}^n$ 或复数集 $\mathbb{C}^n$ 时，我们常常也称其为函数（function）。

定义 II.1.9 (映射的相等). 若映射 $f: X \to Y$ 和 $g: X \to Y$ 满足 $f(x) = g(x), \forall x \in X$，则这两个映射相等。

定义 II.1.10 (恒等映射). 恒等映射（*identity mapping*）$\mathrm{id}_A: A \to A$ 是由集合 A 到其自身的映射：$\mathrm{id}_A(a) = a, \forall a \in A$。

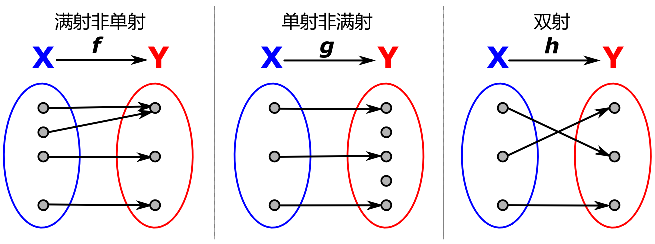

图 II.1.2: 满射、单射和双射

定义 II.1.11 (单射、双射、满射). （如图II.1.2所示）对于映射 $f: X \to Y$，若 $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$，则映射 f 是单射（*injective mapping*）。若 $f(X) = Y$，则称 f 是满射（*surjective mapping*）。既是单射又是满射的映射叫双射（*bijective mapping*）。

读者可自行画图表示非满射非单射的情况。

定义 II.1.12 (复合映射). 令 $f: X \to Y$ 和 $g: Y \to Z$ 是两个映射。则 f 和 g 的复合映射（*composite mapping*），记为 $g \circ f$，是从 X 到 Z 的映射：

$$g \circ f(x) = g(f(x)), \forall x \in X$$

定理 II.1.5. 对于映射 $f: X \to Y$ 和 $g: Y \to Z$，

- 如果 f 和 g 都是满射，则 $g \circ f$ 是满射
- 如果 $g \circ f$ 是满射，则 g 是满射

证明. 留作练习。 □

定理 II.1.6. 对于映射 $f: X \to Y$ 和 $g: Y \to Z$，

- 如果 f 和 g 都是单射，则 $g \circ f$ 是单射
- 如果 $g \circ f$ 是满射，则 f 是单射

证明. 留作练习。 □

定义 II.1.13 (逆映射)**.** 对于映射 $f : X \to Y$，如果存在一个映射 $g : Y \to X$，使得复合映射 $g \circ f$ 是恒等映射 id_X，则称映射 f 是可逆的（*invertible*），映射 g 是 f 的逆映射（*inverse mapping*）。

关于逆映射，有一条重要的定理，使得在今后的数学陈述和推理中，我们可以默认——

定理 II.1.7. 双射必存在唯一逆映射。双射的逆映射也是双射。

证明. 为了证明这一定理，我们首先证明一个引理：任一单射非满射均存在逆映射。

设 $f : X \to Y$ 是一个单射非满射，即 $\exists y \notin \mathrm{f}, y \in Y$。由集合的相关定义此处必有 $\{y|y \in \mathrm{ran}f\} \cup \{y|y \notin \mathrm{ran}f, y \in Y\} = Y$。

现定义 $g : Y \to X$，为使 g 为一个映射，它必须对 $y \in \mathrm{ran}f$ 和 $y \notin \mathrm{ran}f, y \in Y$ 均有定义。现将其定义为：

$$g(y) = \begin{cases} x|_{f(x)=y}, & y \in \mathrm{ran}f \\ \text{任一}x \in X, & y \notin \mathrm{ran}f, y \in Y \end{cases}$$

则有如下几条结论：

1. $g(y)$ 是映射。因为它对每一 $y \in Y$ 均有定义且一个 $y \in Y$ 只对应一个 $x \in X$。
2. g 是满射。因为，仅 $y \in \mathrm{ran}f$ 情况的定义式就已决定了 $\mathrm{ran}g = X$。
3. g 是非单射。因为 g 是满射，再考虑 $y \notin \mathrm{ran}f, y \in Y$ 情况的定义式，就可知 $\exists x \in X$ 满足 $x = g(y) = g(y')$，其中 $y \neq y', y \in \mathrm{ran}f, y' \notin \mathrm{ran}f, y' \in Y$。
4. g 是 f 的逆映射。因为，对于任一 $x \in X$ 均有 $g \circ f(x) \equiv g(f(x)) = x$，即 $g \circ f = \mathrm{id}_X$。
5. 一般地，g 是不唯一的。因为 $y \notin \mathrm{f}, y \in Y$ 的情况可定义 $g(x)$ 等于任一 $x \in X$，故只要集合 X 不是只有一个元素，那么 g 都不唯一。

该引理证毕。

现在正式证定理II.1.7。从上面定义的这个 g 继续，如果 g 是双射，则 g 不仅是满射，还是单射。由刚刚证完的引理，可用类似方法给 g 找一个逆映射 $f' : X \to Y$。而且，由于 $\mathrm{ran}g \equiv X$，我们无需像定义 g 那样为 f' 分出 $x \notin \mathrm{ran}g, x \in X$ 的情况，因为不存在这种情况。故

$$f'(x) = y|_{g(y)=x}$$

是 g 的逆映射，且 f' 是满射。而且，把 g 的定义代入上式有 $f'(x) = y|_{g(y)=x} = y|_{x|_{f(x)=y}} = f(x)$，即 f' 不是别的映射而恰为 $f(x)$。即 g 的逆映射是唯一的。因 f' 是满射故 f 是满射，而 f 本身就是单射，故 f 是双射。 $\square$

第三部分 PostScript

III.1 Some graphs

III.1.1 一些数学公的几何解释

$(a + b)^2 = a^2 + 2ab + b^2$

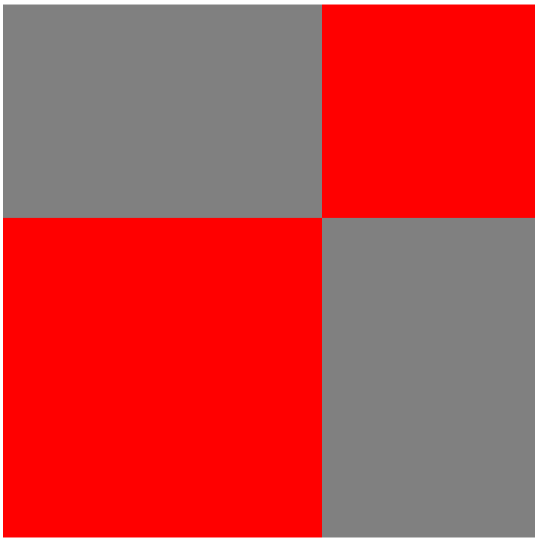

图 III.1.1: $(a + b)^2 = a^2 + 2ab + b^2$

例 III.1.1 (“我们”vs “球上的人”). 在一个相对我们作匀速（低速）圆周运动的球上站着的人(他看不到我们，但能与我们通讯交流)，通过抓稳该球保持与球相对静止。他以球作参照物，研究球上的物体的运动，会得出：保持物体的静止需要一个方向恒定的力，因为球上的物体都需要绑在球上才能保持相对于球静止，他通过弹簧秤可以测量出这一力的大小，进一步发现保持物体静止的力与该物体的质量成正比。如果把某物体松绑，使其不受任何力，物体会开始作一个曲线运动。我们则能不通过测量力就看到球和球上的物体的圆周运动，在我们眼中它们的运动状态是一直在改变中的，而不是静止的；给物体松绑后，物体以松绑时的速度作为初速度作匀速直线运动（如图III.1.1所示）。故我们的观察符合“物体在没有外力作用的情况静止或作匀速直线运动”的论断。如果球的匀速圆周运动的角速度发生变化，我们仍然能不通过力的测量而看到这一点，但球上的人只会发现，保持球上物体静止所需的力的大小在莫名其妙地变化。如果他在球上设置一个单摆，他将发现单摆的摆动方向在随时间变化，而我们认为这个单摆的摆动方向没有变化。总而言之，我们认为牛顿力学定律符合实际，但球上的人将不会

同意。但他和我们通过通讯至少能确认双方所观察的现象是同一个物理现象，按照物理客观性原则理应受同一套自然定律统治。摆在他和我们面前的任务只是如何构造一套理论统一地描述好这一自然定律。

在经典力学中，一个物理事件的发生，同时具有“发生的位置”和“发生的时刻”这两个属性。一名观察者使用时钟测量两个物理事件发生的时间间隔，使用直尺来测量两个物理事件发生位置之间的距离。其中距离的测量总是针对同时刻发生的两个事件。

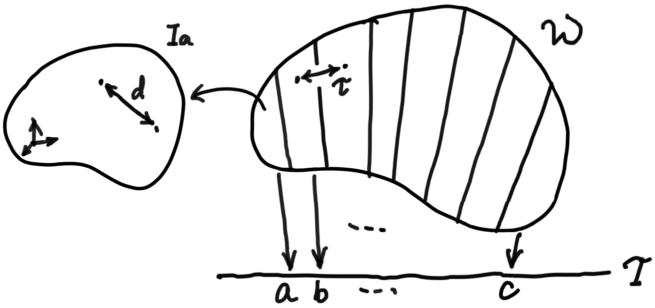

图 III.1.2: 事件世界与同时等价类示意图

引入事件世界 $\mathcal{W}$。

定义 III.1.1 (事件世界、时间). $\mathcal{W}$ 是一个非空集，其元素 $w \in \mathcal{W}$ 称为一个事件。我们给予 $\mathcal{W}$ 一个函数 $\tau : \mathcal{W}^2 \to \mathbb{R}$，满足：

- $\forall a, b \in \mathcal{W} : \tau(a, b) = -\tau(b, a)$
- $\forall a, b, c \in \mathcal{W} : \tau(a, b) + \tau(b, c) = \tau(a, c)$
- $\forall a \in \mathcal{W}, t \in \mathbb{R}, \exists b \in \mathcal{W} : \tau(a, b) = t$

我们称 τ 为（两事件的）时间间隔或时长。我们还能称

- 事件 a 早于事件 b 当且仅当 $\tau(a, b) > 0$
- 事件 a 晚于事件 b 当且仅当 $\tau(a, b) < 0$
- 事件 a 与 b 同时当且仅当 $\tau(a, b) = 0$

上面的第三条给出了两事件的同时性。易验同时性是一个等价关系，即 $S = \{(a, b) | \tau(a, b) = 0\}$ 满足自反性、传递性、对称性。这一等价关系将事件世界 $\mathcal{W}$ 划分为等价类，即：

$$\mathcal{W} = \bigcup_{a \in \mathcal{T}} I_a, I_a = \{x \in \mathcal{W} | \tau(a, x) = 0\}$$

其中等价类 I_a 称为时刻 a 对应的同时等价类，它是所有与事件 a 同时的事件的子集，而 a 在此变成了时刻的标记，集合 $\mathcal{T}$ 称为时间，是所有（不同）时刻的集合（如图*III.1.2*所示）。

我们还给予事件世界 $\mathcal{W}$ 以一个度量 $d : I_a \times I_a \to \mathbb{R}, a \in \mathcal{T}$，称两事件间的“距离”，它还要满足：

- d 是 I_a 上的一个欧几里德度量，故 I_a 是一个欧几里德空间，带有一个平移空间 $\mathcal{V}_a$。
- $\forall a \in \mathcal{T} : \dim \mathcal{V}_a = 3$

注意到，这一度量只能作用在同时发生的两个事件上的（如图*III.1.2*所示）。

结合度量 τ 的定义，$(\mathcal{T}, |\tau|)$ 形成一个度量空间。在此度量空间上的等距变换拥有之前介绍过的一切性质。值得注意的是，由 τ 的定义，可知 $\mathcal{T}$ 的平移空间是一维的、完备的。我们可直接使用实数集 $\mathbb{R}$ 作为其平移空间。选定某时刻 $t_0 \in \mathcal{T}$ 作为原点，$\mathbb{R}$ 的正或负作为时间流逝的方向以及单位时长（相当于在作为向量空间的 $\mathbb{R}$ 中选择了一个规范正交基），则由等距变换的表示定理（定理??）对任一 $\mathcal{T}$ 上的等距变换 $i : \mathcal{T} \to \mathcal{T}$ 有 $i\,(t) = i\,(t_0) + Q\,(t - t_0)$。其中 $Q = \pm 1$ 时间正方向的选择。由于我们很少遇到两个观察者记录时间的方式是反号的情况，因此常常只讨论（默认）$Q = 1$，即

$$i\,(t) = i\,(t_0) + (t - t_0)$$

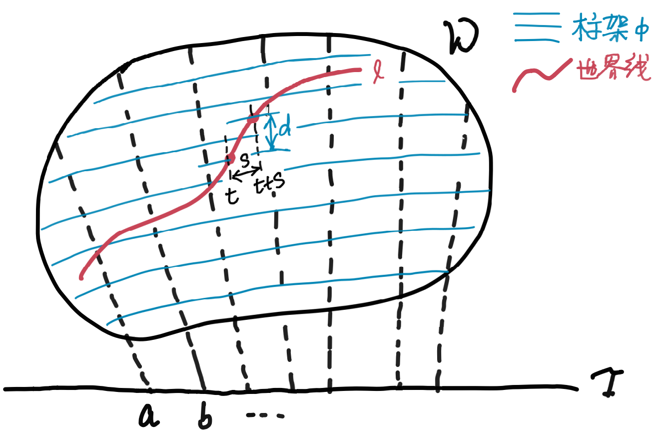

图 III.1.3: 世界线及事件世界的标架示意图

III.1.2 世界线与标架

引入世界线的概念。

定义 III.1.2 (世界线). 世界线是由时间 $\mathcal{T}$ 到事件世界 $\mathcal{W}$ 的单射 $l : \mathcal{T} \to \mathcal{W}$。

一条世界线 $l\,(t)\,, t \in \mathcal{T}$ 的任意两个值都是属于不同时刻的事件（如图III.1.3所示）。一条世界线是某质点运动轨迹的客观存在（即不依赖人的观察与否与观察方式的独立概念）。

依赖世界线的概念，我们可以引入“事件世界的标架”的概念。

定义 III.1.3 (事件世界的标架). 事件世界的标架 ϕ 是一组“平行”的世界线 $\{l, \cdots\}$（如图*III.1.3*所示），其元素 $l \in \phi$ 称为标架线。“平行”的意思是对任意两时刻 $t_1, t_2 \in \mathcal{T}$ 两条世界线 $l_1, l_2 \in \phi$

有 $d(l_1(t_1), l_2(t_!)) = d(l_1(t_2), l_2(t_2))$。一个标架的所有标架线还要覆盖整个事件世界的所有事件，即

$$\bigcup_{l_i \in \phi} \mathrm{ran} l_i = \mathcal{W}$$

使得没有一个事件是不属于任何标架线的。映射 $l_\phi : \mathcal{W} \to \phi, l_\phi(a) = \{l | l \in \phi, a \in \mathrm{ran} l\} \, \forall a \in \mathcal{W}$ 是为任一事件找到其所属的标架线的映射。由标架线“平行”的性质易验 $\{l | l \in \phi, a \in \mathrm{ran} l\}$ 有且只有一个元素，即任一事件 $a \in \mathcal{W}$ 对应且只对应一条标架线，且有 $d(l_\phi(a)(t), l_\phi(b)(t)) =$ 常数$\forall a, b \in \mathcal{W}, t \in \mathcal{T}$，故可由此定义“标架线之间的距离”:

$$d(l_\phi(a), l_\phi(b)) = d(l_\phi(a)(t), l_\phi(b)(t)), t \in \mathcal{T}$$

映射 l_ϕ 还具有性质 $l_\phi(l_\phi(a)(t)) = l_\phi(a) \, \forall t \in \mathcal{T}, a \in \mathcal{W}$。

总结标架的特点：

- 由于标架线是世界线，故它贯穿了不同时刻；
- 任一事件必属于某标架线；
- 我们可以谈论两条标架线之间的距离；

可知，标架的重要意义是使得我们可以讨论不同时刻的任意两个事件 $a \in I_{t_1}, b \in I_{t_2}, t_1, t_2 \in \mathcal{T}$ 之间的距离，选定标架 ϕ，a 与 b 的距离就是 $d(l_\phi(a), l_\phi(b))$。

例如，既然一条世界线所代表的某质点的运动过程，那么我们直觉上会想要通过形如下式的导数来定义“速度”:

$$\lim_{s \to 0} \frac{l(t+s) - l(t)}{s}$$

但是，上式的分子中相减的两个事件是属于不同时刻的同时等价类的（$l(t+s) \in I_{t+s}, l(t) \in I_t, t+s, t \in \mathcal{T}$，如图III.1.3所示），因此这个减法的意义是没有定义的。通过事件世界的标架 ϕ，我们可以解决这个问题，从而定义速度 $\mathbf{v} : \mathcal{T} \to \mathcal{V}_\phi$：

$$\mathbf{v}(t_0) = \lim_{t \to t_0} \frac{l_\phi(l(t)) - l_\phi(l(t_0))}{t}$$

上式的速度定义利用了 (ϕ, d) 是一个欧几里德度量空间的事实，分子中相减的两个元素是标架 ϕ 中的两条世界线。它们的差对应 $\mathcal{V}_\phi$ 中的一个平移向量。故速度是一个向量且依赖标架 ϕ 的选择。注意到，在速度的定义过程中我们默认了一切需要的连续性。加速度也可以类似地被定义，且由于相同的原因加速度也依赖标架 ϕ 的选择。我们称标架 ϕ 的选择是主观的。依赖 ϕ 定义的概念都是不具有物理客观性的（除非增加某些限定条件）。

考虑任一时刻 $t \in \mathcal{T}$ 对应的同时等价类 I_t。选定某标架 ϕ 后，易证 I_t 中的任意两个不同的事件不可能属于同一条标架线，反之亦然。故映射 l_ϕ 限定在某同时等价类 I_t 上是双射。同为欧

几里德度量空间的 (I_t, d) 和 (ϕ, d)，它们对应的平移空间 $\mathcal{V}_t, \mathcal{V}_\phi$ 同构，故 $\mathrm{dim}\mathcal{V}_\phi = \mathrm{dim}\mathcal{V}_t = 3$。对于一系列不同时刻的同时等价类 $I_{t_1}, I_{t_2}, \cdots$，它们各自都对应一个平移空间 $\mathcal{V}_{t_1}, \mathcal{V}_{t_2}, \cdots$，且 $\mathrm{dim}\mathcal{V}_{t_1} = \mathrm{dim}\mathcal{V}_{t_2} = \cdots = \mathrm{dim}\mathcal{V}_\phi$，即标架的平移空间 $\mathcal{V}_\phi$ 与每个时刻的同时等价类的平移空间都同构。我们于是可以统一采用 $\mathcal{V}_\phi$ 作为每个时刻的同时类的平移空间，不同时刻下的事件的位置，我们都用统一的一套平移向量描述。

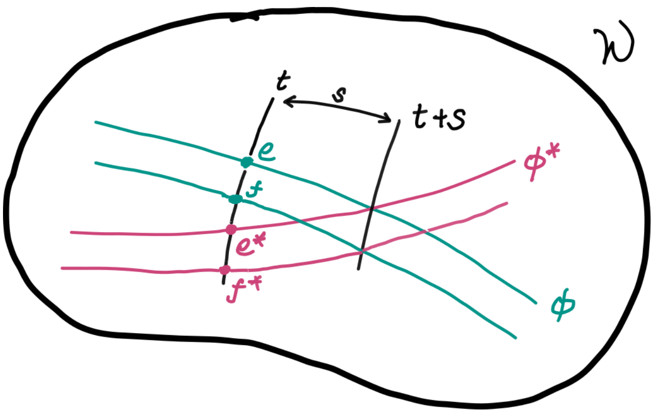

图 III.1.4: 事件世界的标架变换示意图

考虑事件世界 $\mathcal{W}$ 中的两组不同的标架 ϕ, ϕ^*。事件 $e \in I_t$ 的时刻是 $t \in \mathcal{T}$，其所属的 ϕ 中的标架线是 $l_\phi(e)$。考虑距时刻 t 间隔为 $s \in \mathbb{R}$ 的另一时刻 $t+s$，该时刻在标架线 $l_\phi(e)$ 上的事件是 $l_\phi(e)(t+s)$。该事件在标架 ϕ^* 中所属的标架线是 $l_{\phi^*}(l_\phi(e)(t+s))$，在该标架线上 t 时刻的事件是 $l_{\phi^*}(l_\phi(e)(t+s))(t) \in I_t$（如图III.1.4所示）。令

$$e^* = l_{\phi^*}(l_\phi(e)(t+s))(t)$$

则有 $l_{\phi^*}(e^*) = l_{\phi^*}(l_\phi(e)(t+s))$。设 $e, f \in I_t, e \neq f$，则有

$$d(e^*, f^*) = d(l_{\phi^*}(l_\phi(e)(t+s)), l_{\phi^*}(l_\phi(f)(t+s)))$$

由恒等关系 $l_{\phi^*}(l_\phi(e)(t))(t) = l_\phi(e)(t) \forall e \in \mathcal{W}, t \in \mathcal{T}$，上式 $\Leftrightarrow$

$$\begin{aligned} d(l_\phi(e)(t+s), l_\phi(f)(t+s)) &= d(l_\phi(e)(t), l_\phi(f)(t)) \\ &= d(e, f) \end{aligned}$$

因此，对任意 $e, f \in \mathcal{W}$，$d(l_\phi(e), l_\phi(f)) = d(l_{\phi^*}(e^*), l_{\phi^*}(f^*))$，即由 ϕ 到 ϕ^* 存在一个等距映射。由等距映射的表示定理，选定 ϕ 中的一标架线 l_0，则 $l*(s) = l_0^*(s) + \mathbf{Q}(s)(l - l_0), l \in \phi, l^* \in \phi^*$，其中 $l_0^* = l_{\phi^*}(l_\phi(l_0)(t+s))$。我们称这一变换为一个标架变换。一个标架变换不仅包括了上述的标架线之间的等距变换，还隐含包括了一个 $\mathcal{T}$ 的等距变换（即选定了参考时刻 t 并将任一时刻表示为 $t+s$）。

上述关于事件世界、世界线、标架的概念构建方式，代表着我们观测物理事件的实际方式。一位观察者总是把所观察到的空间理解为 3 维欧几里德空间。他用一把直尺测量同一时刻两物理事件的距离，用一个时钟测量两物理事件的时间间隔。在某时刻下，选定空间中一点作为参照点和过该点三条两两不共线的射线方向作为参考方向——即坐标系，则空间中任一点都能通过从参照点到该点的方向和距离（即平移向量）的得到唯一标示。选择某一时刻作为参考时刻，则任一时刻都可通过与参考时刻之间的时长得到标示。坐标系与这一参考时刻的选择，统称参考系。一个观察者在某时刻选择的坐标系会在其他时刻延用；这相当于选定了一个标架 ϕ 并实际统一采用 $\mathcal{V}_\phi$ 中的向量作来任一时刻下的平移向量。

两名观察者，观察同一事件，他们首先要相互对表，即在某时刻两人用各自的钟表同时开始计时，经过相同的时长后，两人得出双方钟表读数的差别。这是在时间 $\mathcal{T}$ 内的一个等距变换。对完表之后，两人选取同一时刻作为参考时刻，对同一物理事件进行观察。由于两人选取的坐标系可能是不同的，因此两人还需要在同一时刻确认对方所找的坐标系；然后，其中一人除了要观察所关注的物理事件在自己所选择的坐标系下随时间的变化之外，还要观察对方选择的坐标系在自己选择的坐标系下随时间的变化，才能知道任一时刻，同一物理事件在对方坐标系下的表示。每个时刻，两个坐标系的之间的关系都可能是不同的，故它们之间的转换是依赖时刻的等距变换。

III.1.3 标架的简化定义

一名观察者如果选定了参考时刻 t_0 和标架 ϕ，则任一时刻都能对应 $\mathbb{R}$ 中的一个实数。选定 $\mathcal{V}_\phi$ 中的一向量 $\widehat{\mathbf{o}}$ 和一组基 $\{\widehat{\mathbf{e}}_i\}$，则任一时刻发生的物理事件的位置都能对应到 $\mathbb{R}^3$ 中的一个坐标。为了实用性，我们可以把标架重新定义为一名观测者从 $\mathcal{W}$ 直接到 $\mathbb{R}^{3+1}$ 的对应过程，即 $\phi : \mathcal{W} \to \mathbb{R}^{3+1}$。这时，标架 ϕ 实际包括了一个观察者对 t_0、$\widehat{\mathbf{o}}$ 和 $\{\widehat{\mathbf{e}}_i\}$ 的主观选择。标架变换仍然包括时间和空间的两个等距变换，与原意义的标架变换是相同。

具体地，任一事件 $w \in \mathcal{W}$ 经标架 ϕ 映射为 $(\mathbf{x}, t), \mathbf{x} \in \mathbb{R}^3, t \in \mathbb{R}$，经标架 ϕ^* 映射为 $(\mathbf{x}^*, t^*), \mathbf{x}^* \in \mathbb{R}^3, t^* \in \mathbb{R}$。我们认为 ϕ, ϕ^* 是可逆的，故有 $(\mathbf{x}^*, t^*) = \phi^* \circ \phi^{-1}(\mathbf{x}, t)$。若 w 的时刻是 $a \in \mathcal{T}$，标架 ϕ 选择的时间原点是 t_0，标架 ϕ^* 选择的时间原点是 t_0^*，则时间的度量空间上的一个等距变换 $g : \mathcal{T} \to \mathcal{T}$ 可表示为

$$t^* = g(t) = t_0^* + \tau(t - t_0), t_0^* = g(t_0)$$

若标架 ϕ 和 ϕ^* 选择的空间原点分别是 $\mathbf{x}_0, \mathbf{x}_0^*$，则在时刻 a 由标架 ϕ 到标架 ϕ^* 的变换是一个等距变换 $i_a : \mathbb{R}^3 \to \mathbb{R}^3$，

$$\mathbf{x}^* = i_a(\mathbf{x}) = i_a(\mathbf{x}_0) + \mathbf{Q}_a(\mathbf{x} - \mathbf{x}_0), \mathbf{x}_0^* = i_a(\mathbf{x}_0)$$

所以，由 ϕ 到 ϕ^* 的标架变换包含一个时间的等距变换 g 和一个空间的等距变换 i_a，其中 a 是事件的时刻。

III.1.4 场函数的标架不变性

设 $(\mathbf{x}^*, t^*), (\mathbf{y}^*, t^*)$ 与 $(\mathbf{x}, t), (\mathbf{y}, t)$ 是同一事件在标架 ϕ, ϕ^* 下的坐标和时标，则有 $\mathbf{x}^* = \mathbf{x}_0^*(t) + \mathbf{Q}(t)[\mathbf{x} - \mathbf{x}_0], \mathbf{y}^* = \mathbf{x}_0^* + \mathbf{Q}(t)[\mathbf{y} - \mathbf{x}_0]$。两式相减得

$$\mathbf{y}^* - \mathbf{x}^* = \mathbf{Q}(t)(\mathbf{y} - \mathbf{x})$$

我们知道 $(\mathbf{x}^*, t^*)$ 与 $(\mathbf{x}, t)$ 对应是同一事件；同样地 $(\mathbf{y}^*, t^*)$ 与 $(\mathbf{y}, t)$ 也对应同一事件。故上式是同一向量在一个标架变换下的关系式。设 $\mathbf{A} \in \mathcal{L}(\mathbb{R}^3)$ 是 $\mathbb{R}^3$ 上的任一张量，$\mathbf{v} = \mathbf{y} - \mathbf{x}, \mathbf{v}^* = \mathbf{y}^* - \mathbf{x}^*$。则有

$$\mathbf{A}^*\mathbf{v}^* \equiv (\mathbf{A}\mathbf{v})^* = \mathbf{Q}(t)\mathbf{A}\mathbf{Q}^\intercal(t)\mathbf{v}^*$$

其中我们利用了 $\mathbf{Q}^{-1} = \mathbf{Q}^\intercal$。上式是同一张量在一个标架变换下的关系式。

设在标架 ϕ, ϕ^* 下，标量场函数 $h = h(\mathbf{x}, t), h^* = h^*(\mathbf{x}^*, t^*)$，向量场函数 $\mathbf{v} = \mathbf{v}(\mathbf{x}, t), \mathbf{v}^* = \mathbf{v}^*(\mathbf{x}^*, t^*)$，张量场函数 $\mathbf{A} = \mathbf{A}(\mathbf{x}, t), \mathbf{A}^* = \mathbf{A}^*(\mathbf{x}^*, t^*)$，则

- 当且仅当 $h^*(\mathbf{x}^*, t^*) = h(\mathbf{x}, t)$ 时称 h——
- 当且仅当 $\mathbf{v}^*(\mathbf{x}^*, t^*) = \mathbf{Q}(t)\mathbf{v}(\mathbf{x}, t)$ 时称 $\mathbf{v}$——
- 当且仅当 $\mathbf{A}^*(\mathbf{x}^*, t^*) = \mathbf{Q}(t)\mathbf{A}(\mathbf{x}, t)\mathbf{Q}^\intercal(t)$ 时称 $\mathbf{A}$——

具有标架不变性。

III.1.5 复合函数求导的链式法则

定理. 定理??：如果函数 $\mathbf{f}: \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ 在 $\mathbf{x}_0 \in D$ 处可微分；函数 $\mathbf{g}: \mathbb{R}^n \supset E \to \mathbb{R}^p$ 在 $\mathbf{f}(\mathbf{x}_0) \in E \cap D$ 处可微分，则复合函数 $\mathbf{g} \circ \mathbf{f}$ 在 $\mathbf{x}_0$ 处可微分，且其全导数

$$\left.\frac{d\mathbf{g} \circ \mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} = \left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)} \left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}$$

证明. 首先证明 $\mathbf{x}_0$ 处于复合函数 $\mathbf{g} \circ \mathbf{f}$ 的定义域内。由于 $\mathbf{f}(\mathbf{x}_0) \in \mathrm{dom}\mathbf{g}$ 且 $\mathbf{g}$ 在 $\mathbf{f}(\mathbf{x}_0)$ 处可微分，故总存在正实数 δ' 使得只要 $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \delta'$ 就有 $\mathbf{f}(\mathbf{x}) \in \mathrm{dom}\mathbf{g}$。又因为 $\mathbf{x}_0 \in \mathrm{dom}\mathbf{f}$ 且 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处可微分，故总存在正实数 δ 使得只要 $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ 则 $\mathbf{x} \in \mathrm{dom}\mathbf{f}$，同时还必存在 $\delta' > 0$ 使得这一 δ 选择下的 $\mathbf{x}$ 满足 $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \delta'$。所以任一满足 $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ 的 $\mathbf{x}$ 均在复合函数 $\mathbf{g} \circ \mathbf{f}$ 的定义域内。

按照全微分和全导数的定义，由于函数 $\mathbf{f}$ 和 $\mathbf{g}$ 分别在 $\mathbf{x}_0$ 和 $\mathbf{f}(\mathbf{x}_0)$ 可导，故存在函数 $\mathbf{z}_1$、$\mathbf{z}_2$ 满足 $\lim_{\mathbf{x}\to\mathbf{x}_0}\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)=\mathbf{0}$、$\lim_{\mathbf{f}(\mathbf{x})\to\mathbf{f}(\mathbf{x}_0)}\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))=\mathbf{0}$，且

$$\begin{gathered}
\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)-\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)=\|\mathbf{x}-\mathbf{x}_0\|\,\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\\
\mathbf{g}(\mathbf{f}(\mathbf{x}))-\mathbf{g}(\mathbf{f}(\mathbf{x}_0))-\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))=\|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)\|\,\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))
\end{gathered}$$

把 $\mathbf{g}(\mathbf{f}(\mathbf{x}))$ 记为 $\mathbf{g}\circ\mathbf{f}(\mathbf{x})$，并把上面的第一条式子代入第二条，得

$$\begin{aligned}
&\mathbf{g}\circ\mathbf{f}(\mathbf{x})-\mathbf{g}\circ\mathbf{f}(\mathbf{x}_0)-\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\left[\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)+\|\mathbf{x}-\mathbf{x}_0\|\,\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\right]\\
=&\|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)\|\,\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))\\
\Leftrightarrow&\mathbf{g}\circ\mathbf{f}(\mathbf{x})-\mathbf{g}\circ\mathbf{f}(\mathbf{x}_0)-\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)\\
=&\|\mathbf{x}-\mathbf{x}_0\|\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)+\|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)\|\,\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))
\end{aligned}$$

由三角不等式，又有*

$$\begin{aligned}
\|\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0)\|&=\left\|\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)+\|\mathbf{x}-\mathbf{x}_0\|\,\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\right\|\\
&\le\left\|\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)\right\|+\|\mathbf{x}-\mathbf{x}_0\|\,\|\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\|\\
&\le k\,\|\mathbf{x}-\mathbf{x}_0\|+\|\mathbf{x}-\mathbf{x}_0\|\,\|\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\|
\end{aligned}$$

故

$$\begin{aligned}
&\mathbf{g}\circ\mathbf{f}(\mathbf{x})-\mathbf{g}\circ\mathbf{f}(\mathbf{x}_0)-\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x}-\mathbf{x}_0)\\
\le&\|\mathbf{x}-\mathbf{x}_0\|\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)+\left(k\,\|\mathbf{x}-\mathbf{x}_0\|+\|\mathbf{x}-\mathbf{x}_0\|\,\|\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\|\right)\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))\\
\le&\|\mathbf{x}-\mathbf{x}_0\|\left\{\left\|\left.\frac{d\mathbf{g}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x}_0)}\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\right\|+\left(k+\|\mathbf{z}_1(\mathbf{x}-\mathbf{x}_0)\|\right)\mathbf{z}_2(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}_0))\right\}
\end{aligned}$$

由于函数 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处连续，即极限 $\lim_{\mathbf{x}\to\mathbf{x}_0}\mathbf{f}(\mathbf{x})=\mathbf{f}(\mathbf{x}_0)$，故上式最后的大括号在 $\mathbf{x}\to\mathbf{x}_0$ 时趋于 $\mathbf{0}$。按照全微分和全导数的定义，命题得证。 □

*此处用到定理??。

III.1.6 反函数定理和隐函数定理

引理 III.1.1. 线性变换 $\mathbf{L}:\mathcal{V}_n\to\mathcal{W}_m$ 是单射当且仅当存在正实数 $p>0$ 满足 $\|\mathbf{Lx}\|\geq p\|\mathbf{x}\|\,\forall\mathbf{x}\in\mathcal{V}_n$。

证明. 如果 $\mathbf{L}$ 不是单射，则存在 $\mathbf{x}_0\neq\mathbf{0}$ 满足 $\mathbf{Lx}_0=\mathbf{0}$，即有 $\|\mathbf{Lx}_0\|=0<m\|\mathbf{x}\|$。故其逆否命题成立。

如果 $\mathbf{L}$ 是单射，则其存在逆 $\mathbf{L}^{-1}$ 满足 $\mathbf{L}^{-1}\mathbf{L}=\mathbf{I}_{\mathcal{V}}$ 且 $\mathbf{L}^{-1}$ 也是线性变换。由定理??，$\exists k>0,\|\mathbf{L}^{-1}\mathbf{y}\|\leq k\|\mathbf{y}\|\,\forall\mathbf{y}\in\mathcal{W}_m$。令 $p=1/k$ 则有 $p\|\mathbf{x}\|=m\|\mathbf{L}^{-1}\mathbf{Lx}\|\leq pk\|\mathbf{Lx}\|=\|\mathbf{Lx}\|$ □

引理 III.1.2. 设 $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^m$ 是可微函数，且在 $\mathbf{x}_0$ 处连续可微。再假设 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处的导数 $\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}$ 是单射。则存在 $\delta>0$ 和 $M>0$ 使得只要 $\|\mathbf{x}-\mathbf{x}_0\|<\delta$ 就有

$$\left\|\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}\mathbf{y}\right\|\geq M\|\mathbf{y}\|\,\forall\mathbf{y}\in\mathbb{R}^n$$

证明. 记函数 $\mathbf{f}$ 的导函数为 $\mathbf{L}$。由引理III.1.1，存在 $p>0$ 使得 $\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|\geq p\|\mathbf{y}\|\,\forall\mathbf{y}\in\mathbb{R}^n$。同时，由于 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处连续可微，故对任一正实数——此处选择 $p/2>0$——存在 $\delta>0$ 使得只要 $\|\mathbf{x}-\mathbf{x}_0\|<\delta$ 就有 $\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_0)\|\leq p/2$。由线性变换的范的定义，有不等式

$$\|(\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_0))\mathbf{y}\|\leq\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_0)\|\,\|\mathbf{y}\|\leq\frac{p}{2}\|\mathbf{y}\|$$

由三角不等式又有

$$\begin{aligned}\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|&=\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}+\mathbf{L}(\mathbf{x})\mathbf{y}-\mathbf{L}(\mathbf{x})\mathbf{y}\|\\&\leq\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}-\mathbf{L}(\mathbf{x})\mathbf{y}\|+\|\mathbf{L}(\mathbf{x})\mathbf{y}\|\\\Leftrightarrow\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|-\|\mathbf{L}(\mathbf{x})\mathbf{y}-\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|\leq\|\mathbf{L}(\mathbf{x})\mathbf{y}\|\end{aligned}$$

上式不等号左边可代入刚刚确定的结论：$\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|\leq p\|\mathbf{y}\|,-\|\mathbf{L}(\mathbf{x})\mathbf{y}-\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|\geq-\frac{p}{2}$，得

$$\begin{aligned}\|\mathbf{L}(\mathbf{x})\mathbf{y}\|&\geq\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|-\|\mathbf{L}(\mathbf{x})\mathbf{y}-\mathbf{L}(\mathbf{x}_0)\mathbf{y}\|\\&\geq p\|\mathbf{y}\|-\frac{p}{2}\|\mathbf{y}\|=\frac{p}{2}\|\mathbf{y}\|\end{aligned}$$

故存在 $M=p/2>0$ 满足命题。 □

引理 III.1.3. 设 $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^m$ 是可微函数，且在 $\mathbf{x}_0$ 处连续可微。再假设 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处的导数 $\left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0}$ 是单射，则存在正实数 $\delta>0$ 和 $M>0$ 使得只要 $\|\mathbf{x}'-\mathbf{x}_0\|<\delta$ 就有 $\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|\leq M\|\mathbf{x}'-\mathbf{x}\|$。

证明. 记函数 $\mathbf{f}$ 的导函数为 $\mathbf{L}$。由引理III.1.1，存在 $m>0$ 使得 $\|\mathbf{L}(\mathbf{x}_0)\mathbf{y}\| \geq m\|\mathbf{y}\| \, \forall \mathbf{y} \in \mathbb{R}^n$。

由于 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处连续可微，故对任一正实数——此处选择 $M = m/(2\sqrt{n}) > 0$——存在 $\delta > 0$ 使得只要 $\|\mathbf{x}-\mathbf{x}_0\| < \delta$ 就有 $\|\mathbf{L}(\mathbf{x}) - \mathbf{L}(\mathbf{x}_0)\| \leq m/(2\sqrt{n})$。

按照命题叙述，设 $\mathbf{x}$ 和 $\mathbf{x}'$ 是 $\mathbb{R}^n$ 的任意两向量满足 $\|\mathbf{x}'-\mathbf{x}_0\| < \delta, \|\mathbf{x}-\mathbf{x}_0\| < \delta$，令 $\mathbf{z} = \mathbf{x}' - \mathbf{x}$，则对 $0 \leq t \leq 1$ 有

$$\begin{aligned}
\|\mathbf{x}+t\mathbf{z}-\mathbf{x}_0\| &= \|t\mathbf{x}' + (1-t)\mathbf{x} - \mathbf{x}_0\| \\
&= \|t(\mathbf{x}'-\mathbf{x}) + (1-t)(\mathbf{x}-\mathbf{x}_0)\| \\
&\leq t\|\mathbf{x}'-\mathbf{x}_0\| + (1-t)\|\mathbf{x}-\mathbf{x}_0\| < t\delta + (1-t)\delta = \delta
\end{aligned}$$

上述推导结论在几何上的意义是，只要点 $\mathbf{x}', \mathbf{x}$ 在由 $\|\mathbf{x}_0\| < \delta$ 的开集内部，则它们的连线上的点 $\mathbf{x}+t\mathbf{z}$ 都在此开集内部，或称“δ-球是凸的”。由于导函数连续是在整个 δ-球内都成立的，因此对由 $0 \leq t \leq 1$ 定义的所有点 $\mathbf{x}+t\mathbf{z}$ 均有 $\|\mathbf{L}(\mathbf{x}+t\mathbf{z}) - \mathbf{L}(\mathbf{x}_0)\| < m/(2\sqrt{n})$。又由线性变换的模的定义有 $\|(\mathbf{L}(\mathbf{x}+t\mathbf{z}) - \mathbf{L}(\mathbf{x}_0))\mathbf{y}\| \leq \|\mathbf{L}(\mathbf{x}+t\mathbf{z}) - \mathbf{L}(\mathbf{x}_0)|\,\|\mathbf{y}\| < M\|\mathbf{y}\| \, \forall \mathbf{y} \in \mathbb{R}^n$。

引入“取坐标函数”,$\pi_k(\mathbf{x}) = x_k, \mathbf{x} = (x_1, \cdots, x_n) \in \mathbf{R}^n, k = 1, \cdots, n$。易验证 $\frac{d\pi_k(\mathbf{x})}{d\mathbf{x}} \equiv \pi_k(\mathbf{x})$。若定义 $g_k(t) = \pi_k(\mathbf{f}(\mathbf{x}+t\mathbf{z})), 0 \leq t \leq 1$，则由链式法则可得如下关系

$$\frac{dg_k}{dt} = \pi_k(\mathbf{L}(\mathbf{x}+t\mathbf{z})\mathbf{z})$$

由微分中值定理，存在 $t_k \in [0,1]$ 使得 $g_k(1) - g_k(0) = \frac{dg_k}{dt_k}$。代入 g_k、$\frac{dg_k}{dt_k}$ 的表达式得 $\pi_k(\mathbf{f}(\mathbf{x}')) - \pi_k(\mathbf{f}(\mathbf{x})) = \pi_k(\mathbf{L}(\mathbf{x}+t_k\mathbf{z})\mathbf{z})$。注意到，函数 $\pi_k(\mathbf{x})$ 就是向量 $\mathbf{x}$ 在第 k 个基上的投影长度。由投影长度不大于向量长度（代数意义是使用柯西–施瓦茨不等式），有

$$\begin{aligned}
\|\mathbf{f}(\mathbf{x}') - \mathbf{f}(\mathbf{x})\| &\geq |\pi_k(\mathbf{f}(\mathbf{x}') - \mathbf{f}(\mathbf{x}))| \\
&= |\pi_k(\mathbf{f}(\mathbf{x}')) - \pi_k(\mathbf{f}(\mathbf{x}))| \\
&= |\pi_k(\mathbf{L}(\mathbf{x}+t_k\mathbf{z})\mathbf{z})|
\end{aligned}$$

另有以下三角不等式成立：

$$\|(\mathbf{L}(\mathbf{x}+t_k\mathbf{z}) - \mathbf{L}(\mathbf{x}_0))\mathbf{z}\| + \|\mathbf{L}(\mathbf{x}_0)\mathbf{z}\| \leq \|\mathbf{L}(\mathbf{x}+t_k\mathbf{z})\mathbf{z}\|$$

上式左右取投影也成立，即

$$|\pi_k((\mathbf{L}(\mathbf{x}+t_k\mathbf{z}) - \mathbf{L}(\mathbf{x}_0))\mathbf{z})| + |\pi_k(\mathbf{L}(\mathbf{x}_0)\mathbf{z})| \leq |\pi_k(\mathbf{L}(\mathbf{x}+t_k\mathbf{z})\mathbf{z})|$$

以上不等式联合有

$$\|\mathbf{f}(\mathbf{x}') - \mathbf{f}(\mathbf{x})\| \geq |\pi_k((\mathbf{L}(\mathbf{x}+t_k\mathbf{z}) - \mathbf{L}(\mathbf{x}_0))\mathbf{z})| + |\pi_k(\mathbf{L}(\mathbf{x}_0)\mathbf{z})|$$

由事实 $\|\mathbf{x}\| \le \sqrt{n}\max\{|x_i|\} \equiv \sqrt{n}\max\{\pi_i(\mathbf{x})\}$（之前在说明范的定义的等价性时证明过该事实）知，在 $k=1,\cdots,m$ 中至少有一个 k 满足

$$\sqrt{n}\left|\pi_k\left(\mathbf{L}\left(\mathbf{x}_0\right)\mathbf{z}\right)\right| \ge \left\|\mathbf{L}\left(\mathbf{x}_0\right)\mathbf{z}\right\|$$

再次利用投影不大于原长，有

$$\left|\pi_k\left(\left(\mathbf{L}\left(\mathbf{x}+t_k\mathbf{z}\right)-\mathbf{L}\left(\mathbf{x}_0\right)\right)\mathbf{z}\right)\right| \le \left\|\left(\mathbf{L}\left(\mathbf{x}+t_k\mathbf{z}\right)-\mathbf{L}\left(\mathbf{x}_0\right)\right)\mathbf{z}\right\|$$

再次联合这些不等式有

$$\begin{aligned}\left\|\mathbf{f}\left(\mathbf{x}'\right)-\mathbf{f}\left(\mathbf{x}\right)\right\| &\ge \frac{1}{\sqrt{n}}\left\|\mathbf{L}\left(\mathbf{x}_0\right)\mathbf{z}\right\| - \left\|\left(\mathbf{L}\left(\mathbf{x}+t_k\mathbf{z}\right)-\mathbf{L}\left(\mathbf{x}_0\right)\right)\mathbf{z}\right\| \\ &\ge 2M\|\mathbf{z}\| - M\|\mathbf{z}\| = M\left\|\mathbf{x}'-\mathbf{x}\right\|\end{aligned}$$

□

有了上面三个引理，我们可正式给出反函数定理的证明。

定理 III.1.1 (反函数定理). **设 $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微函数，记函数 $\mathbf{f}$ 的导函数为 $\mathbf{L}(\mathbf{x}) \equiv \left.\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}}$。若 $\mathbf{L}(\mathbf{x}_0)$ 是单射，则总存在 $\mathbf{x}_0$ 的一个邻域 N 使得 $\mathbf{f}$ 在 N 上有连续可导的逆函数 $\mathbf{f}^{-1}$；$\mathbf{f}$ 的像的集合 $\mathbf{f}(N)$ 也是开集；对 N 内任意一点 $\mathbf{x}$ 都有**

$$\left.\frac{d\mathbf{f}^{-1}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{f}(\mathbf{x})} = \mathbf{L}^{-1}(\mathbf{x})$$

其中 $\mathbf{y}=\mathbf{f}(\mathbf{x})$。

证明. 我们先列出引理III.1.2和III.1.3的结论。由于 $\mathbf{f}$ 在 $\mathbf{x}_0$ 处连续可微，且 $\mathbf{L}(\mathbf{x}_0)$ 是单射，故:

- 由引理III.1.2，对 $\mathbf{x}_0$ 的任一邻域 $N=\{\mathbf{x}|\,\|\mathbf{x}-\mathbf{x}_0\|<\delta\}$（$\delta>0$ 为任一正实数），都能找到正实数 $M(\mathbf{y})>0$ 满足 $\|\mathbf{L}(\mathbf{x})\mathbf{y}\| \ge M\|\mathbf{y}\|\,\forall\mathbf{y}\in\mathbb{R}^n$。进一步地，再由引理III.1.1可知导函数 $\mathbf{L}(\mathbf{x})$ 在 N 上的每个值都是单射线性变换。再由于 $\mathbf{L}(\mathbf{x})$ 在 N 上都是单射线性变换且其定义域和陪域维数相同，故 $\mathbf{L}(\mathbf{x})$ 在 N 上的每个值都是双射（同构）线性变换。
- 由引理III.1.3，对 N 内部任一 $\mathbf{x}'$，总能找到正实数 $M'(\mathbf{x}')>0$ 满足 $\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\| \ge M'\|\mathbf{x}'-\mathbf{x}\|$。

我们令 $M'=M$，这相当于联系了 $\mathbf{y}$ 和 $\mathbf{x}'$。

我们证明的任务包括:

I 函数 $\mathbf{f}$ 存在逆函数 $\mathbf{f}^{-1}$；

II 开集 N 经 $\mathbf{f}$ 的像集 $\mathbf{f}(N)$ 也是开集；

III $\forall \mathbf{x} \in N, \mathbf{L}^{-1}(\mathbf{x})$ 是 $\mathbf{f}^{-1}$ 的导数；

IV $\mathbf{f}^{-1}$ 连续可微。

I的证明：由引理III.1.3的结论，若 $\mathbf{x}' \neq \mathbf{x}$ 则 $\mathbf{f}(x) \neq \mathbf{f}(\mathbf{x}')$，即 $\mathbf{f}$ 是单射，故必存在逆 $\mathbf{f}^{-1}$。I证毕。

II的证明：首先我们确认一些比较直接的接论：

- 由于 N 是开集，故对任一 $\mathbf{x}_1 \in N$，总能找到足够小的 δ_1 使得 $B = \{\mathbf{x} | \|\mathbf{x} - \mathbf{x}_1\| \leq \delta_1\}$ 在 N 的内部。注意这里的 B 是一个闭集。
- 由于函数 $\mathbf{f}$ 存在逆函数 $\mathbf{f}^{-1}$，故 $\mathbf{x} \in N \Leftrightarrow N \ni \mathbf{x} = \mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))\mathbf{f}^{-1}(\mathbf{y}) \forall \mathbf{y} \in \mathbf{f}(N) = \{\mathbf{y} | \mathbf{y} = \mathbf{f}(\mathbf{x}), \mathbf{x} \in N\}$。即给定任一 $\mathbf{y}_1 \in \mathbf{f}(N)$ 有且只有一个 $\mathbf{x}_1 \in N$ 满足 $\mathbf{f}(\mathbf{x}_1) = \mathbf{y}_1$。

要证明 $\mathbf{f}(N)$ 是开集，就是要证明，对任一 $\mathbf{y}_1 \in \mathbf{f}(N)$，总能找到足够小的 $\widetilde{M} > 0$ 使得开集 $C = \left\{\mathbf{y} | \|\mathbf{y} - \mathbf{y}_1\| < \widetilde{M}\right\}$ 在 $\mathbf{f}(N)$ 的内部。

如何由已知条件来找到这个 $\widetilde{M}$ 呢？由于 N 是开集，我们通过 $\mathbf{x}_1 = \mathbf{f}^{-1}(\mathbf{y}_1) \in N$，可以找到使得闭集 $B = \{\mathbf{x} | \|\mathbf{x} - \mathbf{x}_1\| \leq \delta_1\}$ 在 N 的内部的一个正实数 δ_1。

如果 $\widetilde{M}$ 存在，则对任一 $\mathbf{y} \in C$，我们可以从 B 中找到一个 $\mathbf{x}'$ 使得 $\mathbf{y}' = \mathbf{f}(\mathbf{x}')$ 到 $\mathbf{y} \in C$ 的距离最短，并由引理III.1.3，总能找到足够小的正实数 M' 使得

$$\|\mathbf{f}(\mathbf{x}') - \mathbf{y}_1\| = \|\mathbf{f}(\mathbf{x}') - \mathbf{f}(\mathbf{x}_1)\| \geq M' \|\mathbf{x}' - \mathbf{x}_1\| = M'\delta_1$$

接下来我们将证明：

i 如果 $\widetilde{M} = M'\delta/2$，那么上述的 $\mathbf{x}'$ 在 B 的内部（即不在 B 的边界上）；

ii 这一 $\mathbf{y}'$ 就是 $\mathbf{y}$。

上面两条若得证，则给定任一 $\mathbf{y}_1 \in \mathbf{f}(N)$，总有正实数 $\widetilde{M}$（且具体地 $\widetilde{M} = M'\delta_1/2$）使得开集 $C = \left\{\mathbf{y} | \|\mathbf{y} - \mathbf{y}_1\| < \widetilde{M}\right\}$ 在 $\mathbf{f}(N)$ 的内部。II也就得证了。

i的证明：反证法。设 $\mathbf{x}'$ 在 B 的边界上，即 $\|\mathbf{x}' - \mathbf{x}_1\| = \delta_1$，则由引理III.1.3，总能找足够小的正实数 M' 使得

$$\|\mathbf{f}(\mathbf{x}') - \mathbf{y}_1\| = \|\mathbf{f}(\mathbf{x}') - \mathbf{f}(\mathbf{x}_1)\| \geq M' \|\mathbf{x}' - \mathbf{x}_1\| = M'\delta_1$$

那么，由三角不等式，对任一 $\mathbf{y} \in C$（即总有 $\|\mathbf{y}-\mathbf{y}_1\| < M'\delta_1/2$），

$$\begin{aligned}
\|\mathbf{f}(\mathbf{x}')-\mathbf{y}\| &\geq \|\mathbf{f}(\mathbf{x}')-\mathbf{y}_1\| - \|\mathbf{y}-\mathbf{y}_1\| \\
&> M'\delta_1 - \|\mathbf{y}-\mathbf{y}_1\| \\
&> M'\delta_1 - \frac{M'\delta_1}{2} \\
&= \frac{M'\delta_1}{2} \\
&> \|\mathbf{y}-\mathbf{y}_2\| \\
&= \|\mathbf{f}(\mathbf{x}_1)-\mathbf{y}\|
\end{aligned}$$

但这与“$\mathbf{y}'$ 到 $\mathbf{y}$ 的距离最短”相矛盾，故 $\mathbf{x}'$ 在 B 的内部。

ii的证明：设到 $\mathbf{y}$ 的距离平方函数

$$g(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})-\mathbf{y}\|^2 = (\mathbf{f}(\mathbf{x})-\mathbf{y}) \cdot (\mathbf{f}(\mathbf{x})-\mathbf{y})$$

则 $\mathbf{x}'$ 应使得该函数的一阶导数等于零，即 $\left.\frac{dg(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}'} = \mathbf{0}$（零变换）。由零变换性质和链式法则，对任一 $\mathbf{z} \in \mathbb{R}^n$，

$$0 = \left.\frac{dg(\mathbf{x})}{d\mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}'} \mathbf{z} = 2(\mathbf{f}(\mathbf{x}')-\mathbf{y}) \cdot (\mathbf{L}(\mathbf{x}')\mathbf{z})$$

由于 $\mathbf{L}(\mathbf{x})$ 在 N 上的每个值都是双射（同构）线性变换，故有且只有一个向量 $\mathbf{z} \in \mathbb{R}^n$ 满足 $\mathbf{L}(\mathbf{x}')\mathbf{z} = \mathbf{f}(\mathbf{x}')-\mathbf{y}$。故上式 $\Leftrightarrow$

$$0 = 2(\mathbf{f}(\mathbf{x}')-\mathbf{y}) \cdot (\mathbf{f}(\mathbf{x}')-\mathbf{y}) \Leftrightarrow \mathbf{f}(\mathbf{x}')-\mathbf{y} = \mathbf{0}$$

即，只要 $\mathbf{x}' \in N$ 是使 $\mathbf{y}' = \mathbf{f}(\mathbf{x}')$ 到任一 $\mathbf{y} \in \mathbf{f}(N)$ 的距离最短的点，则 $\mathbf{f}(\mathbf{x}') = \mathbf{y}' = \mathbf{y}$。ii证毕。II证毕。

III的证明：按照导数的定义，相当于要证明对任一 $\mathbf{x} \in N$，极限

$$\lim_{\mathbf{f}(\mathbf{x}')\to\mathbf{f}(\mathbf{x})} \frac{\mathbf{x}'-\mathbf{x}-\mathbf{L}^{-1}(\mathbf{x})(\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x}))}{\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|} = \mathbf{0}$$

令未求极限前的比增量为 $\mathbf{s}$，即

$$\mathbf{s} = \frac{\mathbf{x}'-\mathbf{x}-\mathbf{L}^{-1}(\mathbf{x})(\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x}))}{\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|}$$

由于 $\mathbf{L}(\mathbf{x})$ 在 $\mathbf{x} \in N$ 内都有定义，故极限

$$\lim_{\mathbf{x}'\to\mathbf{x}} \frac{\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})-\mathbf{L}(\mathbf{x})(\mathbf{x}'-\mathbf{x})}{\|\mathbf{x}'-\mathbf{x}\|} = \mathbf{0}$$

令

$$\mathbf{r} = \frac{\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})-\mathbf{L}(\mathbf{x})(\mathbf{x}'-\mathbf{x})}{\|\mathbf{x}'-\mathbf{x}\|}$$

则 $\lim_{\mathbf{x}'\to\mathbf{x}}\mathbf{r}=\mathbf{0}$。对 $\mathbf{x}'\in N,\mathbf{x}'\neq\mathbf{x}$，$\mathbf{s}$ 可由 $\mathbf{r}$ 表示为

$$\mathbf{s}=-\frac{\|\mathbf{x}'-\mathbf{x}\|}{\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|}\mathbf{L}^{-1}(\mathbf{x})\,\mathbf{r}$$

由引理III.1.3，存在足够小正实数 $\mathbf{M}$ 使得 $\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|\geq M\|\mathbf{x}'-\mathbf{x}\|$，故有

$$0\geq-\frac{\|\mathbf{x}'-\mathbf{x}\|}{\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|}\geq-\frac{1}{M}$$

即 $-\frac{\|\mathbf{x}'-\mathbf{x}\|}{\|\mathbf{f}(\mathbf{x}')-\mathbf{f}(\mathbf{x})\|}$ 是有界的。

又由定理??的推论，线性变换都是连续函数，故由复合函数的连续性，极限

$$\lim_{\mathbf{x}'\to\mathbf{x}}\mathbf{L}^{-1}(\mathbf{x})\,\mathbf{r}=\mathbf{L}^{-1}(\mathbf{x})\lim_{\mathbf{x}'\to\mathbf{x}}\mathbf{r}=\mathbf{0}$$

由于一个有界函数与一个有极限的函数的积的极限等于那个有极限的函数的极限*，故 $\lim_{\mathbf{x}'\to\mathbf{x}}\mathbf{s}=\mathbf{0}$。又由于当 $\mathbf{x}'\to\mathbf{x}$ 时 $\mathbf{f}(\mathbf{x}')\to\mathbf{f}(\mathbf{x})$，由 $\mathbf{s}$ 的形式有 $\lim_{\mathbf{x}'\to\mathbf{x}}\mathbf{s}=\mathbf{0}\Leftrightarrow\lim_{\mathbf{f}(\mathbf{x}')\to\mathbf{f}(\mathbf{x})}\mathbf{s}=\mathbf{0}$。III证毕。

IV的证明：要证 $\mathbf{f}^{-1}$ 的导函数连续，即对任一 $\mathbf{x}_1\in N,\mathbf{y}_1=\mathbf{f}(\mathbf{x}_1)$ 有

$$\lim_{\mathbf{y}\to\mathbf{y}_1}\left.\frac{d\mathbf{f}^{-1}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}}=\left.\frac{d\mathbf{f}^{-1}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}_1}$$

由III的证明我们已经有

$$\left.\frac{d\mathbf{f}^{-1}(\mathbf{y})}{d\mathbf{y}}\right|_{\mathbf{y}=\mathbf{y}}=\mathbf{L}^{-1}(\mathbf{x}),\mathbf{y}=\mathbf{f}(\mathbf{x}),\forall\mathbf{x}\in N$$

故只需证

$$\lim_{\mathbf{x}\to\mathbf{x}_1}\mathbf{L}^{-1}(\mathbf{x})=\mathbf{L}^{-1}(\mathbf{x}_1)$$

由引理III.1.3，总存在足够小正实数 M 满足 $\|\mathbf{L}(\mathbf{x})\,\mathbf{y}\|\geq M\|\mathbf{y}\|\,\forall\mathbf{y}\in\mathbb{R}^n$，故令 $\mathbf{z}=\mathbf{L}(x)\,\mathbf{y}$，则 $\|\mathbf{z}\|\geq M\|\mathbf{L}^{-1}(\mathbf{x})\,\mathbf{z}\|$。

由于 $\mathbf{f}$ 是连续可微函数，设 $\mathbf{x}_1\in N$，对任一 $\epsilon'>0$，总存在 $\delta>0$，使得只要 $\|\mathbf{x}-\mathbf{x}_1\|<\delta$ 就有 $\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\|<\epsilon'$。具体的，设由 δ 定义的 $\mathbf{x}_1$ 的邻域 $N_1=\{\mathbf{x}|\,\|\mathbf{x}-\mathbf{x}_1\|\}<\delta$ 在 N 的内部，则对任一 $\mathbf{x}\in N_1$，以下不等式成立

$$\begin{aligned}\left\|\left(\mathbf{L}^{-1}(\mathbf{x})-\mathbf{L}^{-1}(\mathbf{x}_1)\right)\mathbf{z}\right\|&=\left\|\mathbf{L}^{-1}(\mathbf{x})\left(\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\right)\mathbf{L}^{-1}(\mathbf{x}_1)\,\mathbf{z}\right\|\\&\leq\frac{1}{M}\left\|\left(\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\right)\mathbf{L}^{-1}(\mathbf{x}_1)\,\mathbf{z}\right\|\\&\leq\frac{1}{M}\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\|\left\|\mathbf{L}^{-1}(\mathbf{x}_1)\,\mathbf{z}\right\|\\&\leq\frac{1}{M^2}\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\|\,\|\mathbf{z}\|\end{aligned}$$

*这个基本定理可由极限的 $\delta-\epsilon$ 语言证明，很多地方有，此略。

由线性变换的范的定义（最大下界），上述不等式 $\Leftrightarrow$

$$\left\|\mathbf{L}^{-1}(\mathbf{x})-\mathbf{L}^{-1}(\mathbf{x}_1)\right\| \leq \frac{1}{M^2}\left\|\mathbf{L}(\mathbf{x})-\mathbf{L}(\mathbf{x}_1)\right\| \leq \frac{\epsilon'}{M^2}$$

令 $\epsilon=\frac{\epsilon'}{M^2}$，我们就有对于任一 $\mathbf{x}_1 \in N$ 和任一 $\epsilon>0$，总有 $\delta>0$ 使得只要 $\|\mathbf{x}-\mathbf{x}_1\|<\delta$ 就有 $\left\|\mathbf{L}^{-1}(\mathbf{x})-\mathbf{L}^{-1}(\mathbf{x}_1)\right\|<\epsilon$。具体地，这个 δ 总存在是由于 M 总存在。这相当于说 $\lim_{\mathbf{x}\to\mathbf{x}_1}\mathbf{L}^{-1}(\mathbf{x})=\mathbf{L}^{-1}(\mathbf{x}_1)$，IV证毕。 □

III.1.7 等距变换的表示定理

引理 III.1.4. 设 $(\mathcal{E},d)$ 是一个欧几里得空间，$\mathcal{V}$ 是其平移空间，$i:\mathcal{E}\to\mathcal{E}$ 是 $\mathcal{E}$ 上的一个等距变换。记 $\mathbf{u}'=i(X+\mathbf{u})-i(X), \mathbf{v}'=i(X+\mathbf{v})-i(X), \mathbf{u},\mathbf{u}',\mathbf{v},\mathbf{v}'\in\mathcal{V}, X\in\mathcal{E}$，则 $(\mathbf{u}'|\mathbf{v}')=(\mathbf{u}|\mathbf{v})$ 对任意 $\mathbf{u},\mathbf{u}',\mathbf{v},\mathbf{v}'\in\mathcal{V}$ 均成立。

证明. 利用欧几里得空间的定义和等距变换的定义，有

$$(\mathbf{u}'|\mathbf{u}')=d^2(i(X),i(X+\mathbf{u}))=d^2(X,X+\mathbf{u})=(\mathbf{u}|\mathbf{u})$$

同理有 $(\mathbf{v}'|\mathbf{v}')=(\mathbf{v}|\mathbf{v})$。由内积的性质，$(\mathbf{u}-\mathbf{v}|\mathbf{u}-\mathbf{v})=(\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v})-2(\mathbf{u}|\mathbf{v})$，则

$$\begin{aligned}
2(\mathbf{u}'|\mathbf{v}') &= (\mathbf{u}'|\mathbf{u}')+(\mathbf{v}'|\mathbf{v}')-(\mathbf{u}'-\mathbf{v}'|\mathbf{u}'-\mathbf{v}') \\
&= (\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v}) \\
&\quad -(i(X+\mathbf{u})-i(X)-i(X+\mathbf{v})+i(X)\,|i(X+\mathbf{u})-i(X)-i(X+\mathbf{v})+i(X)) \\
&= (\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v}) \\
&\quad -d^2(i(X+\mathbf{v}),i(X+\mathbf{v})+(i(X+\mathbf{u})-i(X+\mathbf{v}))) \\
&= (\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v})-d^2(X+\mathbf{v},X+\mathbf{u}) \\
&= (\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v})-(X+\mathbf{u}-(X+\mathbf{V})\,|X+\mathbf{u}-(X+\mathbf{v})) \\
&= (\mathbf{u}|\mathbf{u})+(\mathbf{v}|\mathbf{v})-(\mathbf{u}-\mathbf{v}|\mathbf{u}-\mathbf{v}) \\
&= 2(\mathbf{u}|\mathbf{v})
\end{aligned}$$

□

引理 III.1.5. 设 $(\mathcal{E},d)$ 是一个欧几里得空间，$\mathcal{V}$ 是其平移空间，$i:\mathcal{E}\to\mathcal{E}$ 是 $\mathcal{E}$ 上的一个等距变换。定义关于一点 $X\in\mathcal{E}$ 的映射 $\mathbf{Q}_X:\mathcal{V}\to\mathcal{V}, \mathbf{Q}_X\mathbf{u}=i(X+\mathbf{u})-i(X)\,\forall\mathbf{u}\in\mathcal{V}$，则

1. $\mathbf{Q}_X$ 与 X 的选择无关（故可略去下标 $\mathbf{X}$）；

2. $\mathbf{Q}$ 是正交算符;
3. 对给定的 i, Q 唯一存在。

证明. 由命题中 $\mathbf{Q}_X$ 的定义有

$$\mathbf{Q}_X (Y - X) = i(Y) - i(X) \forall X, Y \in \mathcal{E}$$

由引理III.1.4有

$$\begin{aligned}(\mathbf{Q}_X\mathbf{u}|\mathbf{Q}_X\mathbf{v}) &= (i(X+\mathbf{u}) - i(X)|i(X+\mathbf{u}) - i(X)) \\ &= (\mathbf{u}|\mathbf{v}), \forall \mathbf{u}, \mathbf{v} \in \mathcal{V}\end{aligned}$$

如果 $\mathbf{Q}_X$ 是线性算符，则上述结论就证明了 $\mathbf{Q}_X$ 是正交算符。以下进一步证明映射 $\mathbf{Q}_X$ 是一个线性算符。由

$$\begin{aligned}&(\mathbf{Q}_X(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2) - \mathbf{Q}_X(\alpha_1\mathbf{u}_1) - \mathbf{Q}_X(\alpha_2\mathbf{u}_2)|\mathbf{Q}_X\mathbf{v}) \\ =& (i(X+\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2) - i(X) - [i(X+\alpha_1\mathbf{u}_1) - i(X)] - [i(X+\alpha_2\mathbf{u}_2) - i(X)] \\ &|i(X+\mathbf{v}) - i(X)) \\ =& (i(X+\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2) - i(X)|i(X+\mathbf{v}) - i(X)) \\ &- (i(X+\alpha_1\mathbf{u}_1) - i(X)|i(X+\mathbf{v}) - i(X)) \\ &- (i(X+\alpha_2\mathbf{u}_2) - i(X)|i(X+\mathbf{v}) - i(X)) \quad \text{(用到引理III.1.4)} \\ =& (\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2|\mathbf{v}) - (\alpha_1\mathbf{u}_1|\mathbf{v}) - (\alpha_2\mathbf{u}_2|\mathbf{v}) \\ =& 0\end{aligned}$$

得

$$\mathbf{Q}_X(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2) - \mathbf{Q}_X(\alpha_1\mathbf{u}_1) - \mathbf{Q}_X(\alpha_2\mathbf{u}_2) = \mathbf{0}$$

即 $\mathbf{Q}_X$ 是线性算符。结合上一结论 $\mathbf{Q}_X$ 就是正交算符。第 2 个命题得证。

要证明 $\mathbf{Q}_X$ 不依赖 X 的选择，设有另一 $X' \in \mathcal{E}$，则对任意 $\mathbf{u} \in \mathcal{V}$ 有

$$\begin{aligned}\mathbf{Q}_{X'}\mathbf{u} &= i(X'+\mathbf{u}) - i(X') \\ &= i(X + [(X'-X)\mathbf{u}]) - i(X + (X'-X)) \\ &= \mathbf{Q}_X((X'-X)+\mathbf{u}) - \mathbf{Q}_X(X'-X) \\ &= \mathbf{Q}_X[(X'-X)+\mathbf{u}-(X'-X)] \\ &= \mathbf{Q}_X\mathbf{u}\end{aligned}$$

即得证。

至此，对于给定的 i，$\mathbf{Q}$ 的存在性已证明。$\mathbf{Q}$ 的唯一性也是显然的。设另有线性算符 $\mathbf{P}:\mathcal{V}\to\mathcal{V},\mathbf{P}\mathbf{u}=i\left(X+\mathbf{u}\right)-i\left(X\right)\forall\mathbf{u}\in\mathcal{V}$，则易验

$$\left(\mathbf{P}-\mathbf{Q}\right)\mathbf{u}=\mathbf{P}\mathbf{u}-\mathbf{Q}\mathbf{u}=\mathbf{0}$$

□

定理 (等距变换的表示定理). 定理??: 设 $(\mathcal{E},d)$ 是一个欧几里德空间，$\mathcal{V}$ 是其平移空间，选定任一点 $X_0\in\mathcal{E}$，则 $\mathcal{E}$ 上的任一等距变换 $i:\mathcal{E}\to\mathcal{E},i\in\mathcal{I}$ 都可表示为

$$i\left(X\right)=i\left(X_0\right)+\mathbf{Q}_i\left(X-X_0\right)$$

其中 $\mathbf{Q}_i$ 是一个正交算符，关于 i 唯一存在。

证明. 由引理III.1.5，任一 $\mathcal{E}$ 上的等距变换 i 都唯一对应一个 $\mathcal{V}$ 上的正交算符 $\mathbf{Q}$ 满足 $\mathbf{Q}\mathbf{u}=i\left(X+\mathbf{u}\right)-i\left(X\right)\forall X\in\mathcal{E},\mathbf{u}\in\mathcal{V}$。对任一 $X\in\mathcal{E}$，可令 $\mathbf{u}=X-X_0$，则对任一 i 存在唯一 $\mathbf{Q}_i$ 满足 $\mathbf{Q}_i\mathbf{u}=\mathbf{Q}_i\left(X-X_0\right)=i\left(X\right)-i\left(X_0\right)\Rightarrow i\left(X\right)=i\left(X_0\right)+\mathbf{Q}_i\left(X-X_0\right)$。 □

参考文献

[1] 钟尔杰. 大学物理（上册）[M]. 华南理工大学出版社, 2009.

[2] 钟尔杰. 数学实验方法[M]. 电子科技大学出版社, 2009.

[3] 邓文基. 大学物理（下册）[M]. 华南理工大学出版社, 2009.

[4] 钟理, 伍钦, 马四朋. 化工原理（上册）[M]. 化学工业出版社, 2008.

[5] 华罗庚. 数论导引[M]. 科学出版社, 1957.

[6] HOFFMAN K, KUNZE R. Linear Algebra[M]. 2nd ed. Prentice-Hall, Inc., 1971.

[7] HASSANI S. Mathematical physics : a modern introduction to its foundations[M]. New York: Springer, 1999.

[8] NOLL W. Lectures on the Foundations of Continuum Mechanics and Thermodynamics[G] //The Foundations of Mechanics and Thermodynamics: Selected Papers. Berlin, Heidelberg: Springer Berlin Heidelberg, 1974: 293-324.

[9] 钟尔杰. 数值分析讲义[M]. 电子科技大学出版社, 2009.

[10] WILLIAMSON R E, CROWELL R H, TROTTER H F. Calculus of Vector Functions[M]. 3rd ed. Prentice Hall, Inc., 1972.

[11] REISS J, SPRENGER J. Scientific Objectivity[G]//ZALTA E N. The Stanford Encyclopedia of Philosophy. Winter 2020. Metaphysics Research Lab, Stanford University, 2020.