微分几何入门

Le Yang
yangle0125@qq.com

目录

第一章　拓扑空间简介

1.1　集论初步

确切地指定了的若干事物的全体叫做一个集合，简称集。集中的每一事物叫一个元素或点。若 x 是集 X 的元素，则说 “x 属于 X”，并记作 $x \in X$。符号 $\notin$ 则代表 “不属于”。有两种表示集合的方法，一种是一一列出其元素，元素间用逗号隔开，全体元素用花括号括起来，如

$$X = \{1, 4, 5.6\}$$

表示由实数 1,4 以及 5.6 构成的集。另一种表示方法是指出集中元素的共性，如

$$X = \{x \mid x \text{ 为实数}\}$$

表示 X 是全体实数的集合（这一特定集的通用记号为 $\mathbb{R}$）, 而

$$X = \{x \in \mathbb{R} \mid x > 9\}$$

则表示全体大于 9 的实数的集合。

不含元素的集叫空集，记作 $\varnothing$。

定义 1.1. *若集 A 的每一元素都属于集 X，就说 A 是 X 的*子集*，也说 A 含于 X 或 X 含 A，记作 $A \subset X$ 或 $X \supset A$。规定 $\varnothing$ 是任一集合的子集。A 称为 X 的*真子集*，若 $A \subset X$ 且 $A \neq X$。集 X 和 Y 称为*相等*的（记作 $X = Y$），若 $X \subset Y$ 且 $Y \subset X$*

注 1.1. *子集定义的更确切表述本应是 “集 A 叫集 X 的子集，当且仅当 A 的每一元素都属于 X”。但为方便起见，凡在定义中的 “若” 或 “当” 都是 “当且仅当” 之意。*

本书用 $:=$ 代表 “定义为”，用 $\equiv$ 代表 “恒等” 或 “记作”，例如 $C \equiv A - B$ 的含义是 “把 $A - B$ 记作 C”。采用这两个符号无非是为增加明确性，都换成等号也无妨。

定义 1.2. *集合 A，B 的并集，交集，差集和补集定义为：*

并集 $A \cup B := \{x \mid x \in A \text{ 或 } x \in B\}$

交集 $A \cap B := \{x \mid x \in A, x \in B\}$ *（条件 “$x \in A, x \in B$” 是 “$x \in A$且$x \in B$” 的简写，下同。）*

差集 $A - B := \{x \mid x \in A, x \notin B\}$（数学书常把差集记作 $A\backslash B$或$A \sim B$，本书一律记作 $A - B$。）

若 A 是 X 的子集，则 A 的**补集**$-A$ 定义为 $-A := X - A$

定理 1.1. 以上集运算符合如下规律：

交换律 $A \cup B = B \cup A$, $A \cap B = B \cap A$

结合律 $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$

分配律 $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$, $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

De Morgan **律** $A - (B \cup C) = (A - B) \cap (A - C)$, $A - (B \cap C) = (A - B) \cup (A - C)$

证明. 交换律和结合律根据定义显然。

$A \subset A \cup C, B \subset B \cup C$，所以若 $x \in A, x \in B$ 就有 $x \in A \cup C, x \in B \cup C$，即 $A \cap B \subset (A \cup C) \cap (B \cup C)$。同理，$C \subset (A \cup C) \cap (B \cup C)$。所以有

$$(A \cap B) \cup C \subset (A \cup C) \cap (B \cup C)$$

反过来，若 $x \in (A \cup C) \cap (B \cup C)$，也即 $x \in (A \cup C), x \in (B \cup C)$，而这可以有 4 种组合 $x \in A, x \in B$或$x \in A, x \in C$或$x \in C, x \in B$或$x \in C, x \in C$ 即 $x \in (A \cap B) \cup (A \cap C) \cup (B \cap C) \cup C$，注意到 $(B \cap C) \cup C = C$，就有 $x \in (A \cap B) \cup C$。这就说明

$$(A \cup C) \cap (B \cup C) \subset (A \cap B) \cup C$$

综上 $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$。

类似地，$A \cap C \subset A, B \cap C \subset B$，可以证明

$$(A \cap C) \cup (B \cap C) \subset (A \cup B) \cap C$$

反过来，若 $x \in (A \cup B) \cap C$, 也即 $x \in (A \cup B), x \in C$, 这可以有两种情况，$x \in A, x \in C$ 或 $x \in B, x \in C$，因此

$$(A \cup B) \cap C \subset (A \cap C) \cup (B \cap C)$$

综上 $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

对于 De Morgan 律,$x \in A - (B \cup C)$ 也即 $x \in A, x \notin (B \cup C)$,即 $x \in A, x \notin B, x \in A, x \notin C$。因此，$x \in (A - B) \cap (A - C)$，反之也成立，这说明

$$A - (B \cup C) = (A - B) \cap (A - C)$$

另一方面，$A - B \subset A, A - C \subset A$，因此 $(A - B) \cup (A - C) \subset A$。又 $x \in A - B$ 说明 $x \notin B$，因此 $x \in (A - B) \cup (A - C)$ 一方面说明 $x \in A$，另一方面说明 $x \notin B$ 或 $x \notin C$ 即 $x \notin B \cap C$。因此，$x \in A - (B \cap C)$，反之也成立，这说明

$$(A - B) \cup (A - C) = A - (B \cap C)$$

□

定义 1.3. 非空集合 X,Y 的卡氏积 $X\times Y$ 定义为

$$X\times Y\coloneqq\{(x,y)\mid x\in X, y\in Y\}$$

就是说，$X\times Y$ 是这样一个集合，它的每一个元素是由 X 的一个元素 x 和 Y 的一个元素 y 组成的一个有序对 (x,y)。多个（有限个[1]）集合的卡氏积可以类似定义，例如

$$X\times Y\times Z\coloneqq\{(x,y,z)\mid x\in X, y\in Y, z\in Z\}$$

而且规定卡氏积满足结合律，即 $(X\times Y)\times Z=X\times(Y\times Z)$

例 1.1. $\mathbb{R}^2\coloneqq\mathbb{R}\times\mathbb{R}$, $\mathbb{R}^n\coloneqq\mathbb{R}\times\cdots\times\mathbb{R}$（共 n 个 $\mathbb{R}$）. 既然 $\mathbb{R}^2$ 的元素是由两个实数构成的有序对，这两个实数就称为该元素的**自然坐标**。类似地，$\mathbb{R}^n$ 的每一个元素有 n 个自然坐标。可见 $\mathbb{R}^n$ 天生就是有坐标的，但其他集合则未必。利用自然坐标可以给 $\mathbb{R}^n$ 的任意两个元素定义距离的概念。

定义 1.4. $\mathbb{R}^n$ 的任意两个元素 $x=(x^1,\cdots,x^n)$, $y=(y^1,\cdots,y^n)$ 之间的**距离**$|y-x|$ 定义为

$$|y-x|\coloneqq\sqrt{\sum_{i=1}^{n}(y^i-x^i)^2}$$

定义 1.5. 设 X,Y 为非空集合。一个从 X 到 Y 的**映射**（记作 $f\colon X\to Y$）是一个法则，它给 X 的每一个元素指定 Y 的唯一的对应元素。若 $y\in Y$ 是 $x\in X$ 的对应元素，就写 $y=f(x)$，并称 y 为 x 在映射 f 下的**像**，称 x 为 y 的**原像**（或**逆像**）。X 称为映射 f 的**定义域**，X 的全体元素在映射 f 下的像的集合（记作 $f[X]$）称为映射 $f\colon X\to Y$ 的**值域**。映射 $f\colon X\to Y$ 和 $f'\colon X\to Y$ 称为**相等**的，若 $f(x)=f'(x)\ \forall x\in X$。

注 1.2. 通常也把 $y=f(x)$ 写成 $f\colon x\mapsto y$。请注意 $\mapsto$ 和 $\to$ 的区别：$f\colon X\to Y$ 中的 $\to$ 表示 f 是从 X 到 Y（集合到集合）的映射；而 $f\colon x\mapsto y$ 中的 $\mapsto$ 则表示 $x\in X$ 在映射 f 下的像是 y（元素到元素）。

注 1.3. 设 $A\subset X$，则 A 的元素在 f 下的像组成的子集记作 $f[A]$，即

$$f[A]\equiv\{y\in Y\mid \exists x\in A\text{使得}y=f(x)\}\subset Y$$

例 1.2. 普通微积分中的单值函数 $y=f(x)$ 就是一个由 $\mathbb{R}$（或其子集）到 $\mathbb{R}$ 的映射。

注 1.4. 从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的映射给出一个二元函数。同理，从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的映射给出 m 个 n 元函数。

定义 1.6. 映射 $f\colon X\to Y$ 叫**一一**的，若任一 $y\in Y$ 有不多于一个逆像（可以没有）。$f\colon X\to Y$ 叫**到上**的，若任一 $y\in Y$ 都有逆像（可以多于一个）。[2]

[1] 无限多个集合的卡氏积也可定义，但已超出本书范围。

[2] 不少数学书把本书的一一和到上映射分别叫**单射**和**满射**，把既是单射又是满射的映射叫**一一映射**（又称**双射**）。于是它们的一一映射强于本书的一一映射。

注 1.5. ① f 为到上映射的充要条件是值域 $f[X]=Y$。② 若 f 为一一映射，则存在逆映射 $f^{-1}\colon f[X]\to X$。然而，不论 $f\colon X\to Y$ 是否有逆，都可以定义任一子集 $B\subset Y$ 在 f 下的"逆像"$f^{-1}[B]$ 为

$$f^{-1}[B]\coloneqq\{x\in X\mid f(x)\in B\}\subset X$$

注意，这里的"逆像"是 X 的子集而不是 X 的元素。例如，如果 X 有（且仅有）两个元素 x 和 x' 在 f 作用（即映射）下的像都是 $y\in Y$，虽然逆映射 $f^{-1}\colon Y\to X$ 不存在，但把 y 看作 Y 的独点子集（即 $\{y\}$）时 $f^{-1}[\{y\}]$（简记作 $f^{-1}[y]$）仍有意义，含义为 $f^{-1}[y]=\{x,x'\}\subset X$。

定义 1.7. $f\colon X\to Y$ 称为**常值映射**，若 $f(x)=f(x')\ \forall x,x'\in X$

定义 1.8. 设 X,Y,Z 为集，$f\colon X\to Y$ 和 $g\colon Y\to Z$ 为映射，则 f 和 g 的**复合映射**$g\circ f$ 是从 X 到 Z 的映射，定义为 $(g\circ f)(x)\coloneqq g(f(x))\in Z\ \forall x\in X$

注 1.6. 若 $X=Y=Z=\mathbb{R}$，则复合映射 $g\circ f$ 就是熟知的一元复合函数。

若 X 和 Y 是一般的集合，对 X 与 Y 之间的映射就只能提出"一一"和"到上"这两个要求；但若 X 和 Y 还指定了某种结构，则往往可对 $f\colon X\to Y$ 提出更多要求，例如可要求 $f\colon\mathbb{R}\to\mathbb{R}$ 是连续的甚至光滑的。一元函数 $f\colon\mathbb{R}\to\mathbb{R}$ 的连续性在微积分中早有定义（"$\varepsilon-\delta$ 定义"），重述如下：① 称 f 在点 x 连续，若 $\forall\varepsilon>0\exists\delta>0$ 使得当 $|x'-x|<\delta$ 时有 $|f(x')-f(x)|<\varepsilon$；② 称 f 在 $\mathbb{R}$ 上连续，若它在 $\mathbb{R}$ 的任一点连续。这一定义依赖于 $\mathbb{R}$ 中任二元素的距离概念（对 $\mathbb{R}$ 而言距离就是坐标之差），似乎无法推广到没有距离定义的两个集合之间的映射。然而细想发现，$\varepsilon-\delta$ 定义可用开区间概念（而无需距离概念）重新表述如下：设 $X=Y=\mathbb{R}$，映射 $f\colon X\to Y$ 叫做连续的，若 Y 中任一开区间的"逆像"都是 X 的开区间之并（或是空集）。以上讨论从一个侧面说明"开区间之并"这一概念的用处：可以定义映射 $f\colon\mathbb{R}\to\mathbb{R}$ 的连续性。其实这一概念还有很多用处，因此往往有必要推广到除 $\mathbb{R}$ 外的集合 X。为方便起见，把 $\mathbb{R}$ 的任一可以表为开区间之并的子集（连同空集 $\varnothing$）称为开子集。为把开子集概念推广到任意集合 X，应先找出 $\mathbb{R}$ 的开子集的本质的、抽象的（因而可以推广的）性质。它们是：

(a) $\mathbb{R}$ 和空集 $\varnothing$ 都是开子集

(b) 有限个开子集之交仍是开子集

(c) 任意个开子集之并仍是开子集

把这三个性质推广，就可给任意集合 X 定义开子集概念。定义了开子集的集合叫拓扑空间。由开子集的概念出发又可定义许多概念并证明许多定理，从而发展为一门完整丰富的学科分支 — 点集拓扑学。下面将对拓扑空间的最基本内容做一介绍。

1.2 拓扑空间

如前所述，$\mathbb{R}$ 的子集分为开子集和非开子集两大类（任一子集要么是开的，要么是非开的。不要把非开子集称为闭子集。根据后面要讲的闭子集的定义，子集可以不开不闭，也可以既开又闭。）开子集具有上述三个性质。对任意非空集合 X 也可用适当方式指定其中的某些子集是开的，其他为非开的。为使这种指定有用，我们约定任何指定方式都要满足三个要求：

(a) X 本身和空集 $\varnothing$ 为开子集

(b) 有限个开子集之交为开子集

(c) 任意个（可以有限个也可以无限个）开子集之并为开子集

对同一集合，满足这三个要求的指定方式常常是很多的。例如，设 X 为任意集合，可以指定 X 及 $\varnothing$ 为开子集，其他子集都为非开。这当然满足上述三要求，其特点是开子集最少，只有两个。然而也可采用另一种极端的指定，即指定 X 的任意子集都是开子集。不难看出这种指定也满足上述三要求。上述两种指定虽然未必有太多用处，但它们至少能说明满足上述三要求的指定方式不止一种。我们说，每种满足上述三要求的指定给集合 X 赋予了一种附加结构，称为拓扑结构。对定义了拓扑结构的集合可以指着它的任一子集问："这是开子集吗？" 答案非 "是" 即 "否"，泾渭分明。反之，对没有定义拓扑结构的集合，这样的问题毫无意义。定义了拓扑结构的集合 X 的全体开子集也组成一个集合，称为 X 的一个拓扑，记作 $\mathscr{T}$（是 topology 的首字母的花体大写）。用 $\mathscr{P}$ 代表由 X 的全体子集组成的集合，则 X 的任一开子集 O 和任一非开子集 V 都是 $\mathscr{P}$ 的元素。X 的全体开子集组成 $\mathscr{P}$ 的一个子集 $\mathscr{T}$（注意，它不是 X 的子集），它就是 X 的拓扑。请注意符号 $\subset$ 同 $\in$ 的区别：$O \subset X$ 只表明 O 是 X 的子集，而 $O \in \mathscr{T}$ 则表明 O 是 X 的开子集。以上铺垫有助于理解如下用数学语言表述的定义。

定义 1.9. *非空集合 X 的一个***拓扑** $\mathscr{T}$ *是 X 的若干子集的集合，满足：*

(*a*) $X, \varnothing \in \mathscr{T}$

(*b*) 若 $O_i \in \mathscr{T}, i = 1, 2, \cdots, n$, 则 $\bigcap\limits_{i=1}^{n} O_i \in \mathscr{T}$

(*c*) 若 $O_\alpha \in \mathscr{T}\ \forall\alpha$, 则 $\bigcup\limits_{\alpha} O_\alpha \in \mathscr{T}$

定义 1.10. *定义了拓扑 $\mathscr{T}$ 的集合 X 称为***拓扑空间**。*拓扑空间 X 的子集 O 称为***开子集**（*简称***开集**）*, 若 $O \in \mathscr{T}$。*

对同一集合 X 可以定义不同的拓扑 $\mathscr{T}$（满足定义的 $\mathscr{T}$ 可以很多）。设 $\mathscr{T}_1$ 和 $\mathscr{T}_2$ 都是 X 的拓扑，则 X 的子集 A 可能满足 $A \in \mathscr{T}_1, A \notin \mathscr{T}_2$，即 A 对 $\mathscr{T}_1$ 而言（用 $\mathscr{T}_1$ 衡量）是开

集而对 $\mathscr{T}_2$ 而言不是开集。可见 $\mathscr{T}_1$ 和 $\mathscr{T}_2$ 把 X 定义为两个不同的拓扑空间。为明确所选拓扑起见，可用 $(X, \mathscr{T})$ 代表拓扑空间。于是 $(X, \mathscr{T}_1)$ 和 $(X, \mathscr{T}_2)$ 代表不同的拓扑空间，虽然它们的“底集”都是 X。在明确选定拓扑后也可只用 X 代表拓扑空间。

对给定的具体集合 X，应该选哪个拓扑使之成为一个拓扑空间？这取决于 X 自身的性质以及我们关心哪些方面的问题。例如，对集合 $\mathbb{R}^n$，在通常关心的大多数问题中都选所谓的通常拓扑为拓扑。

例 1.3. 设 X 为任意非空集合，令 $\mathscr{T}$ 为 X 的全部子集的集合，则它满足拓扑的定义，故构成 X 的一个拓扑，叫**离散拓扑**。

例 1.4. 设 X 为任意非空集合，令 $\mathscr{T} = \{X, \varnothing\}$，则它满足拓扑的定义，故构成 X 的一个拓扑，叫**凝聚拓扑**。凝聚拓扑是元素最少的拓扑，而离散拓扑是元素最多的拓扑。

例 1.5. (1) 设 $X = \mathbb{R}$，则 $\mathscr{T}_u \coloneqq \{空集或\mathbb{R}中能表为开区间之并的子集\}$ 称为 $\mathbb{R}$ 的**通常拓扑**。

(2) 设 $X = \mathbb{R}^n$，则 $\mathscr{T}_u \coloneqq \{空集或\mathbb{R}^n中能表为开球之并的子集\}$ 称为 $\mathbb{R}^n$ 的**通常拓扑**，其中，**开球**的定义为 $B(x_0, r) \coloneqq \{x \in \mathbb{R}^n \mid |x - x_0| < r\}$，$x_0$ 称为球心，$r > 0$ 称为半径。$\mathbb{R}^2$ 中的开球亦称**开圆盘**，$\mathbb{R}$ 中的开球就是开区间。

根据上述定义，$\mathbb{R}$ 中任一开区间用 $\mathscr{T}_u$ 衡量都是开集。然而，原则上也可以选其他拓扑使 $\mathbb{R}$ 成为不同于 $(\mathbb{R}, \mathscr{T}_u)$ 的拓扑空间。例如，若以凝聚拓扑衡量，则除 $\mathbb{R}$ 及 $\varnothing$ 之外都不是开集；反之，若以离散拓扑衡量，则 $\mathbb{R}$ 中任一子集（包括闭区间和半闭区间）都是开集。今后在把 $\mathbb{R}$ 看做拓扑空间时，如无声明就是指 $(\mathbb{R}, \mathscr{T}_u)$

例 1.6. 设 $(X_1, \mathscr{T}_1), (X_2, \mathscr{T}_2)$ 为拓扑空间，$X = X_1 \times X_2$，定义 X 的拓扑为

$$\mathscr{T} \coloneqq \{O \in X \mid O可表示为形如O_1 \times O_2的集合之并, O_1 \in \mathscr{T}_1, O_2 \in \mathscr{T}_2\}$$

则 $\mathscr{T}$ 称为 X 的**乘积拓扑**。

例 1.7. 设 $(X, \mathscr{T})$ 为拓扑空间，A 为 X 的任一非空子集。把 A 看作集合，当然也可指定拓扑（记作 $\mathscr{S}$，是 S 的花体）使 A 成为拓扑空间，记作 $(A, \mathscr{S})$。由于 A 是 X 的子集，我们希望 $\mathscr{S}$ 与 $\mathscr{T}$ 有尽量密切的联系。如果 $A \in \mathscr{T}$，问题很简单，只须定义 $\mathscr{S} \coloneqq \{V \subset A \mid V \subset \mathscr{T}\}$。然而，如果 $A \notin \mathscr{T}$，按上述定义就有 $A \notin \mathscr{S}$，违背定义的条件。因此 $\mathscr{S}$ 的上述定义不合法。一个巧妙的定义是

$$\mathscr{S} \coloneqq \{V \subset A \mid \exists O \in \mathscr{T} 使 V = A \cap O\}$$

由上式可以证明即使 $A \notin \mathscr{T}$ 也有 $A \in \mathscr{S}$，而且 $\mathscr{S}$ 满足定义的其他条件。这样定义的 $\mathscr{S}$ 叫做 A 的、由 $\mathscr{T}$ 导出的**诱导拓扑**。以后在把 $(X, \mathscr{T})$ 的子集 A 看作拓扑空间时，如无声明都指 $(A, \mathscr{S})$，其中 $\mathscr{S}$ 是由 $\mathscr{T}$ 诱导的拓扑。$(A, \mathscr{S})$ 称为 $(X, \mathscr{T})$ 的**拓扑子空间**。

下面的例子有助于加深对诱导拓扑的理解。$\mathbb{R}^2$ 中以 x_0 为心的单位圆周 S^1 定义为 $S^1 \coloneqq \{x \in \mathbb{R}^2 \mid |x - x_0| = 1\}$。设 $A \subset \mathbb{R}^2$ 是 S^1，由于它不能表为 $\mathbb{R}^2$ 中的开球之并（一条线窄到装不下任何开圆盘），A 用 $\mathbb{R}^2$ 的 $\mathscr{T}_u$ 衡量不是开的。用前述定义给 A 定义诱导拓扑 $\mathscr{S}$，则不但 A 用 $\mathscr{S}$ 衡量是开的，而且，设 V 是 A 中的任意一段（不含首末两点），则虽然 V 用 $\mathscr{T}_u$ 衡量不是开集，用 $\mathscr{S}$ 衡量却是开的，因为存在开圆 $O \in \mathscr{T}_u$ 使 $V = A \cap O$。

利用开集概念可对拓扑空间之间的映射定义连续性。下面给出两个等价的连续定义

定义 1.11.（*a*） 设 $(X, \mathscr{T})$ 和 $(Y, \mathscr{S})$ 为拓扑空间。映射 $f\colon X \to Y$ 称为**连续的**，若 $f^{-1}[O] \in \mathscr{T}\ \forall O \in \mathscr{S}$。

（*b*） 设 $(X, \mathscr{T})$ 和 $(Y, \mathscr{S})$ 为拓扑空间。映射 $f\colon X \to Y$ 称为**在点 $x \in X$ 处连续**，若对所有满足 $f(x) \in G'$ 的 $G' \in \mathscr{S}$，存在 $G \in \mathscr{T}$ 使 $x \in G$ 且 $f[G] \subset G'$。$f\colon X \to Y$ 称为连续，若它在所有点 $x \in X$ 上连续。

注 1.7. 不难看出，若 $X = Y = \mathbb{R}$，$\mathscr{T} = \mathscr{S} = \mathscr{T}_u$，上述定义就回到 $\varepsilon - \delta$ 定义。

定义 1.12. 拓扑空间 $(X, \mathscr{T})$ 和 $(Y, \mathscr{S})$ 称为**互相同胚**，若存在映射 $f\colon X \to Y$，满足（*a*）f 是一一到上的；（*b*）f 及 f^{-1} 都连续。[3] 这样的 f 称为从 $(X, \mathscr{T})$ 到 $(Y, \mathscr{S})$ 的**同胚映射**，简称**同胚**。

普通函数 $y = f(x)$ 的连续性和可微性用 C^r 表示，其中 r 为非负整数，C^0 代表连续，C^r 代表 r 阶导数存在并连续，C^∞ 代表任意阶导数存在并连续（称为**光滑**）。虽然用开集概念可以巧妙地把 C^0 性推广到拓扑空间之间的映射，但 $r > 0$ 的 C^r 性则不能。事实上，对拓扑空间之间的映射的最高要求已体现在同胚的定义中。同胚映射 $f\colon X \to Y$ 不仅在 X 和 Y 的点之间建立了一一对应的关系，而且还在 X 的开子集和 Y 的开子集之间建立了一一对应的关系，因而一切由拓扑决定的性质都可“全息”地被 f“携带”到 Y 中。因此，从纯拓扑学角度看，两个互相同胚的拓扑空间就“像得不能再像”，可以视作相等。

例 1.8. 任一开区间 $(a, b) \subset \mathbb{R}$ 与 $\mathbb{R}$ 同胚。

例 1.9. 圆周 $S^1 \subset \mathbb{R}^2$ 配以诱导拓扑（由 $\mathbb{R}^2$ 的 $\mathscr{T}_u$ 诱导）可看做拓扑空间。它与 $\mathbb{R}$ 不同胚。但挖去一点的圆周与 $\mathbb{R}$ 同胚。

例 1.10. 考虑欧氏平面上的一个圆和一个椭圆（均指圆周）。从拓扑学的角度看，$(\mathbb{R}^2, \mathscr{T}_u)$ 是拓扑空间，圆 S^1 和椭圆 E 是 $\mathbb{R}^2$ 的两个子集：$S^1, E \in \mathbb{R}^2$。可用 $\mathscr{T}_u$ 给 S^1 和 E 分别定义诱导拓扑使成两个拓扑空间 $(S^1, \mathscr{S}_{S^1})$ 及 $(E, \mathscr{S}_E)$。可以证明存在同胚映射 $f\colon (S^1, \mathscr{S}_{S^1}) \to (E, \mathscr{S}_E)$，所以从纯拓扑眼光看两者完全一样。

定义 1.13. $N \subset X$ 称为 $x \in X$ 的一个**邻域**，若存在 $O \in \mathscr{T}$ 使 $x \in O \subset N$。自身是邻域的开集称为**开邻域**。

[3]的确存在其逆不连续的一一到上的连续映射（用离散和凝聚拓扑）。

注 1.8. 设 $X=\mathbb{R}$，$N=[a,b]$，则 N 按定义是 x 的邻域，当且仅当 $a<x<b$。请特别注意“擦边”情况：若 $x=a$，则 N 并非 x 的邻域，因为 $\mathbb{R}$ 不存在开集 O 使 $x\in O\subset N$。直观地说，要使 $[a,b]$ 是 x 的邻域，x 在 $[a,b]$ 中应有“左邻右舍”。而 $x=a$ 的任何“左邻”都不属于 $[a,b]$，故 $[a,b]$ 不应是 $x=a$ 的邻域。另请注意如下微妙的例子：在拓扑空间 $[0,\infty)\subset\mathbb{R}$ 中，$[0,1)$ 是 0 的开邻域，$[0,1]$ 是 0 的邻域[4]。

定义 1.14. $N\subset X$ 称为 $A\subset X$ 的**一个邻域**，若存在 $O\in\mathscr{T}$ 使 $A\subset O\subset N$。

定理 1.2. $A\subset X$ 是开集，当且仅当 A 是 x 的邻域 $\forall x\in A$。

证明.（A） 设 A 为开，则 $\forall x\in A,\exists A\in\mathscr{T}$ 使 $x\in A\subset A$，故由前述定义知 A 是 x 的邻域。

（B） 设 A 是 x 的邻域 $\forall x\in A$，令 $O=\bigcup\limits_{x\in A}O_x$（$O_x\in\mathscr{T}$ 且满足 $x\in O_x\subset A$），则 $O=A$。又根据定义知 $O\in\mathscr{T}$，故 $A\in\mathscr{T}$，即 A 为开集。

□

定义 1.15. $C\subset X$ 叫**闭集**，若 $-C\in\mathscr{T}$。

定理 1.3. 闭集有以下性质：

(*a*) 任意个闭集的交集是闭集

(*b*) 有限个闭集的并集是闭集

(*c*) X 及 $\varnothing$ 是闭集

证明. 由定义和 De Morgan 律易证。 □

可见任何拓扑空间 $(X,\mathscr{T})$ 都有两个既开又闭的子集，即 X 和 $\varnothing$

定义 1.16. 拓扑空间 $(X,\mathscr{T})$ 称为**连通**的，若它除 X 和 $\varnothing$ 外没有既开又闭的子集。

例 1.11. 设 A 和 B 是 $\mathbb{R}$ 的开区间，$A\cap B=\varnothing$，以 $\mathscr{T}$ 代表由 $\mathbb{R}$ 的通常拓扑在子集 $X\equiv A\cup B$ 上的诱导拓扑，则拓扑空间 $(X,\mathscr{T})$ 的既开又闭的子集除 X 和 $\varnothing$ 外还有 A 和 B（A 和 B 在诱导拓扑下自然是开的，又因为 A 和 B 互为补集，故又都是闭的），所以 $(X,\mathscr{T})$ 是不连通的。这同 A 和 B 在直观上互不连通相吻合[5]。

设 $(X,\mathscr{T})$ 为拓扑空间，$A\subset X$。分别定义 A 的闭包、内部和边界如下：

定义 1.17. A 的**闭包**$\bar{A}$ 是所有含 A 的闭集的交集，即

$$\bar{A}\coloneqq\bigcap_{\alpha}C_\alpha,\ A\subset C_\alpha,\ -C_\alpha\in\mathscr{T}$$

[4]将 $[0,\infty)$ 记作 A。视其为拓扑空间即为其配备由 $\mathscr{T}_u$ 定义的诱导拓扑 $\mathscr{T}_A=\{V\subset A\mid\exists O\in\mathscr{T}_u 使 V=A\cap O\}$。现在只要取 $O=(-1,1)$，则 $V=A\cap O=[0,1)$ 是 A 中的开集，且满足 $x\in V\subset[0,1]$。

[5]与直观感觉更一致的是称为“弧连通”的概念。它与连通概念有微妙差别。

定义 1.18. A 的**内部** $i(A)$ 是所有含于 A 的开集的并集，即

$$i(A) := \bigcup_{\alpha} O_\alpha,\ O_\alpha \subset A,\ O_\alpha \in \mathscr{T}$$

定义 1.19. A 的**边界** $\dot{A} := \bar{A} - i(A)$，$x \in \dot{A}$ 称为**边界点**。$\dot{A}$ 也记作 ∂A

定理 1.4. $\bar{A}$，$i(A)$ 及 $\dot{A}$ 有以下性质：

(*a*) ①$\bar{A}$ 为闭集，②$A \subset \bar{A}$，③$A = \bar{A}$ 当且仅当 A 为闭集。

(*b*) ①$i(A)$ 为开集，②$i(A) \subset A$，③$i(A) = A$ 当且仅当 A 为开集。

(*c*) $\dot{A}$ 为闭集

证明. (a)、(b) 易证。(c) 的证明如下：$X - \dot{A} = X - [\bar{A} - i(A)] = (X - \bar{A}) \cup i(A)$。因为 $\bar{A}$ 为闭，故 $X - \bar{A}$ 为开，加之 $i(A)$ 为开，故 $X - \dot{A}$ 为开，因而 $\dot{A}$ 为闭。 □

定义 1.20. X 的开子集的集合 $\{O_\alpha\}$ 叫 $A \subset X$ 的一个**开覆盖**，若 $A \subset \bigcup_{\alpha} O_\alpha$。也可以说 $\{O_\alpha\}$ 覆盖 A。

1.3 紧致性

定义 1.21. 设 $\{O_\alpha\}$ 是 $A \subset X$ 的开覆盖。若 $\{O_\alpha\}$ 的有限个元素构成的子集 $\{O_{\alpha_1}, \cdots, O_{\alpha_n}\}$ 也覆盖 A，就说 $\{O_\alpha\}$ 有**有限子覆盖**。

定义 1.22. $A \subset X$ 叫**紧致的**，若它的任一开覆盖都有有限子覆盖。

例 1.12. 设 $x \in X$，则独点子集 $A \equiv \{x\}$ 必紧致。

证明. 设 $\{O_\alpha\}$ 是 $A \subset X$ 的任一开覆盖，则 $\{O_\alpha\}$ 中至少存在一个元素（记作 O_{α_1}）满足 $x \in O_{\alpha_1}$。于是 $\{O_{\alpha_1}\}$（作为 $\{O_\alpha\}$ 的子集）是 $A \equiv \{x\}$ 的开覆盖，故 $\{O_\alpha\}$ 有有限子覆盖。 □

例 1.13. $A \equiv (0, 1] \subset \mathbb{R}$ 不是紧致的。

证明. 以 $\mathbb{N}$ 代表自然数集，则 $\{(1/n, 2) \mid n \in \mathbb{N}\}$ 是 A 的开覆盖，它没有有限子覆盖。 □

类似地，$\mathbb{R}$ 中任一开区间或半开区间都非紧致。

例 1.14. $\mathbb{R}$ 不是紧致的。

定理 1.5. $\mathbb{R}$ 的任一闭区间都紧致。

注 1.9. 不要以为闭集一定紧致（就连 $\mathbb{R}$ 中也有非紧致闭集）。紧性与闭性有密切联系，但不等价，其关系体现在以下两定理中。

为讲述以下定理，先补充以下定义。

定义 1.23. 拓扑空间 $(X,\mathscr{T})$ 叫 T_2 **空间**或**豪斯多夫空间**，若

$$\forall x,y\in X,\ x\neq y,\ \exists O_1,O_2\in\mathscr{T}\text{使}x\in O_1,\ y\in O_2\text{且}O_1\cap O_2=\varnothing$$

注 1.10. 常见的拓扑空间（如 $\mathbb{R}^n$）都是 T_2 空间。凝聚拓扑空间是非 T_2 空间的一例。

定理 1.6. 若 $(X,\mathscr{T})$ 为 T_2 空间，$A\subset X$ 为紧致，则 A 为闭集。

定理 1.7. 若 $(X,\mathscr{T})$ 为紧致且 $A\subset X$ 为闭集，则 A 为紧致。

定义 1.24. $A\subset\mathbb{R}^n$ 叫**有界的**，若存在开球 $B\subset\mathbb{R}^n$ 使 $A\subset B$。

定理 1.8. $A\subset\mathbb{R}$ 为紧致，当且仅当 A 为有界闭集。

定理 1.9. 设 $A\subset X$ 紧致，$f\colon X\to Y$ 连续，则 $f[A]\subset Y$ 紧致。

由以上定理可得推论：同胚映射保持子集的紧致性。

定义 1.25. 在同胚映射下保持不变的性质称为**拓扑性质**或**拓扑不变性**。

例 1.15. 紧致性、连通性和 T_2 性都是拓扑性质。有界性不是拓扑性质，例如开区间 (a,b) 同胚于 $\mathbb{R}$，但前者有界而后者无界。由此还可看出长度也不是拓扑性质。

数学分析中有个熟知定理：闭区间上的连续函数必在该区间上取得其最大值和最小值。下述定理是这一定理的推广。

定理 1.10. 设 X 紧致，$f\colon X\to\mathbb{R}$ 连续，则 $f[X]\subset\mathbb{R}$ 有界并取得其最大值和最小值。

定理 1.11. 设 $(X_1,\mathscr{T}_1),(X_2,\mathscr{T}_2)$ 紧致，则 $(X_1\times X_2,\mathscr{T})$ 紧致（$\mathscr{T}$ 为 $\mathscr{T}_1$ 和 $\mathscr{T}_2$ 的乘积拓扑）。

定理 1.12. $A\subset\mathbb{R}^n$ 紧致，当且仅当它是有界闭集。

简单应用举例　考虑 $(\mathbb{R}^2,\mathscr{T}_u)$。设 S^1 是 $\mathbb{R}^2$ 中的任一圆周，易见它是有界闭集，于是它为紧致。因为连续映射保紧致性，而 $\mathbb{R}$ 及其任一开区间都不紧致，可见 S^1 不可能与 $\mathbb{R}$ 或其任一开区间同胚。类似地，$\mathbb{R}$ 中任一闭区间都不可能与 $\mathbb{R}$ 或其任一开区间同胚。

定义 1.26. 映射 $S\colon\mathbb{N}\to X$ 叫 X 中的**序列**。

注 1.11. 通常把序列记作 $\{x_n\}$，其中 $x_n\equiv S(n)\in X$，$n\in\mathbb{N}$。$\{x_n\}$ 其实就是 X 中编了次序的一串点。

定义 1.27. $x\in X$ 叫序列 $\{x_n\}$ 的**极限**，若对 x 的任一开邻域 O 存在 $N\in\mathbb{N}$ 使 $x_n\in O\ \forall n>N$。若 x 是 $\{x_n\}$ 的极限，就说 $\{x_n\}$ **收敛**于 x。

定义 1.28. $x \in X$ 叫序列 $\{x_n\}$ 的**聚点**，若 x 的任一开邻域都含 $\{x_n\}$ 的无限多点。

注 1.12. x 为 $\{x_n\}$ 的极限 $\Rightarrow x$ 为 $\{x_n\}$ 的聚点，但反之不一定。

下述定理中有一条件涉及"第二可数"概念。元素个数有限的集成为**有限集**，否则称为**无限集**。对有限集总可将其元素编号以便一个一个地数，所以有限集一定是可数集。但无限集也不一定不可数，例如 $\mathbb{N}$ 就是可数的无限集。有限集比无限集简单，可数无限集比不可数无限集简单。拓扑空间 $(X, \mathscr{T})$ 称为**第二可数的**，若 $\mathscr{T}$ 存在可数子集 $\{O_1, \cdots, O_k\} \subset \mathscr{T}$（或 $\{O_1, \cdots\} \subset \mathscr{T}$）使得任一 $O \in \mathscr{T}$ 可被表示为 $\{O_1, \cdots, O_k\}$（或 $\{O_1, \cdots\}$）的元素之并。例如，$(\mathbb{R}^n, \mathscr{T}_u)$ 是第二可数的，因为 $\mathscr{T}_u$ 有这样的子集（它的每个元素 O_i 是一个开球，球心的每个自然坐标都是有理数，半径也是有理数），使得任一 $O \in \mathscr{T}_u$ 可被表为该子集的元素之并。

定理 1.13. 若 $A \subset X$ 紧致，则 A 中任一序列都有在 A 内的聚点。反之，若 X 为第二可数且 $A \subset X$ 中任一序列都有在 A 内的聚点，则 A 紧致。

第二章 流形和张量场

2.1 微分流形

定义 2.1. 拓扑空间 $(M,\mathscr{T})$ 称为**n 维微分流形**，简称 n 维流形，若 M 有开覆盖 O_α，即 $M=\bigcup\limits_\alpha O_\alpha$，满足

（a） 对每一 O_α，存在同胚 $\psi_\alpha\colon O_\alpha\to V_\alpha$（$V_\alpha$ 是 $\mathbb{R}^n$ 用通常拓扑衡量的开子集）；

（b） 若 $O_\alpha\cap O_\beta\neq\varnothing$，则复合映射 $\psi_\beta\circ\psi_\alpha^{-1}$ 是 C^∞（光滑）的。[1]

注 2.1. ①$\psi_\beta\circ\psi_\alpha^{-1}$ 是从 $\psi_\alpha[O_\alpha\cap O_\beta]\subset\mathbb{R}^n$ 到 $\psi_\beta[O_\alpha\cap O_\beta]\subset\mathbb{R}^n$ 的映射。因 $\mathbb{R}^n$ 的每点都有 n 个自然坐标，故 $\psi_\beta\circ\psi_\alpha^{-1}$ 提供了 n 个 n 元函数。所谓 $\psi_\beta\circ\psi_\alpha^{-1}$ 是 C^∞ 的，就是指这每一个 n 元函数都是 C^∞ 的（n 元函数的 C^∞ 性在微积分中早有定义）。[2] ②设 $p\in O_\alpha$，则 $\psi_\alpha(p)\in\mathbb{R}^n$，故 $\psi_\alpha(p)$ 有 n 个自然坐标。很自然地把这 n 个数称为 p 点在映射 ψ_α 下获得的**坐标**。M 作为拓扑空间，其元素本来一般没有坐标，但作为流形，M 中位于 O_α 内的元素（点）就可以通过映射 ψ_α 获得坐标。若 $O_\alpha\cap O_\beta\neq\varnothing$，则 $O_\alpha\cap O_\beta$ 内的点既可以通过 ψ_α 又可以通过 ψ_β 获得坐标，这两组坐标一般不同。我们说 (O_α,ψ_α) 构成一个（局域）**坐标系**，其**坐标域**为 O_α；(O_β,ψ_β) 构成另一坐标系，其坐标域为 O_β。于是 $O_\alpha\cap O_\beta$ 内的点至少有两组坐标，分别记作 $\{x^\mu\}$ 和 $\{x'^\nu\}$（$\mu,\nu=1,\cdots,n$）。由映射 $\psi_\beta\circ\psi_\alpha^{-1}$ 提供的、体现两组坐标之间关系的 n 个 n 元函数

$$x'^1=\phi^1(x^1,\cdots,x^n),\cdots,x'^n=\phi^n(x^1,\cdots,x^n)$$

就称为一个**坐标变换**。定义的条件（b）保证坐标变换中的函数关系 $x'^\mu=\phi^\mu(x^1,\cdots,x^n)$ 都是 C^∞ 的。为方便起见也常称 $\{x^\mu\}$ 为坐标系，虽然从 $\{x^\mu\}$ 中看不出坐标域的范围。物理学家也常把 $x'^\mu=\phi^\mu(x^1,\cdots,x^n)$ 记作 $x'^\mu=x'^\mu(x^1,\cdots,x^n)$。

定义 2.2. 坐标系 (O_α,ψ_α) 在数学上又叫**图**，满足定义条件（a）、（b）的全体图的集合 $\{(O_\alpha,\psi_\alpha)\}$ 叫**图册**。条件（b）又称**相容性条件**，因此说一个图册中的任意两个图都是相容的。

[1] 此定义是光滑流形的一般定义。本书（以及通常在物理中）涉及的流形 M 还满足以下附加条件：作为拓扑空间，M 是 Hausdorff 的和第二可数的。今后谈到流形都指这种流形。

[2] 上述定义中的流形是光滑流形的简称，把条件（b）的 C^∞ 改为 C^r（r 为含零自然数）就成为 C^r 流形的定义。

例 2.1. 设 $M = (\mathbb{R}^2, \mathscr{T}_u)$。选 $O_1 = \mathbb{R}^2, \psi_1 =$ 恒等映射，则 $\{(O_1, \psi_1)\}$ 便是只含一个图的图册，故 $\mathbb{R}^2$ 是 2 维流形，而且是能用一个坐标域覆盖的流形，称为**平凡流形**。根据这个图册，$\mathbb{R}^2$ 中每点的坐标就是它作为 $\mathbb{R}^2$ 的元素天生就有的自然坐标。$\mathbb{R}^2$ 的点当然也可用其他坐标（如极坐标）描述。其实这无非是选择与图 (O_1, ψ_1) 相容的另一个图 (O_2, ψ_2)，其中 ψ_2 把 $p \in O_2$ 映为 $\psi_2(p) \in \mathbb{R}^2$，再把 $\psi_2(p)$ 的自然坐标称为 p 点的新坐标而已。但应注意坐标域 O_2 未必能包括 $\mathbb{R}^2$ 的全体点（例如极坐标）。

同理可知 $\mathbb{R}^n$ 是 n 维平凡流形。

例 2.2. 设 $M = (S^1, \mathscr{S})$，其中 $S^1 \coloneqq \{x \in \mathbb{R}^2 \mid |x - o| = 1\}$ 是以原点 o 为心的单位圆周，$\mathscr{S}$ 是 $\mathbb{R}^2$ 的 $\mathscr{T}_u$ 在 S^1 上诱导的拓扑。则可证明 S^1 是 1 维流形。

例 2.3. 设 $M = (S^2, \mathscr{S})$，其中 $S^2 \coloneqq \{x \in \mathbb{R}^3 \mid |x - o| = 1\}$ 是以原点 o 为心的单位球面，$\mathscr{S}$ 是 $\mathbb{R}^3$ 的 $\mathscr{T}_u$ 在 S^2 上诱导的拓扑。则可证明 S^2 是 2 维流形。

设图册 $\{(O_\alpha, \psi_\alpha)\}$ 把拓扑空间 M 定义为一个流形，则此图册中的任意两个图自然是相容的。但也可用另一图册 $\{(O'_\beta, \psi'_\beta)\}$ 把同一 M 定义为流形，这时有两种可能：①这两个图册不相容，即存在 O_α 和 O'_β 使 $O_\alpha \cap O'_\beta \neq \varnothing$，且在 $O_\alpha \cap O'_\beta$ 上 ψ_α 与 ψ'_β 不满足定义条件（b），这时就说这两个图册把 M 定义为两个不同的微分流形，并说这两个图册代表两种不同的**微分结构**；②这两个图册是相容的，这时就说它们把 M 定义为同一个微分流形（只有一种微分结构）。为方便起见，不妨把 $\{(O_\alpha, \psi_\alpha); (O'_\beta, \psi'_\beta)\}$ 看成一个图册。更进一步，索性把所有与 (O_α, ψ_α) 相容的图都放到一起造出一个最大的图册。今后说到 M 是一个流形时，总是默认已选定某一最大的图册作为微分结构。这使我们可以进行任意坐标变换。

微分流形与拓扑空间的重要区别是前者除有拓扑结构外还有微分结构，因此两个流形之间的映射不但可谈及是否连续，还可谈及是否可微，乃至是否 C^∞。设 M 和 M' 是两个流形，维数依次为 n 和 n'，$\{(O_\alpha, \psi_\alpha)\}$ 和 $\{(O'_\beta, \psi'_\beta)\}$ 依次为两者的图册，$f\colon M \to M'$ 是一个映射。$\forall p \in M$，任取坐标系 (O_α, ψ_α) 使 $p \in O_\alpha$ 以及坐标系 (O'_β, ψ'_β) 使 $f(p) \in O'_\beta$，则 $\psi'_\beta \circ f \circ \psi_\alpha^{-1}$ 是从 $V_\alpha \equiv \psi_\alpha[O_\alpha]$ 到 $\mathbb{R}^{n'}$ 的映射，因此相应于 n' 个 n 元函数，它们的 C^r 性可用以定义 $f\colon M \to M'$ 的 C^r 性。

定义 2.3. $f\colon M \to M'$ 称为**C^r 类映射**，如果 $\forall p \in M$，映射 $\psi'_\beta \circ f \circ \psi_\alpha^{-1}$ 对应的 n' 个 n 元函数是 C^r 类的。

注 2.2. 由于同一图册中各图相容，上述定义与坐标系 $\{(O_\alpha, \psi_\alpha)\}$ 及 $\{(O'_\beta, \psi'_\beta)\}$ 的选择无关。

定义 2.4. 微分流形 M 和 M' 称为**互相微分同胚**，若存在 $f\colon M \to M'$，满足（*a*）f 是一一到上的；（*b*）f 及 f^{-1} 是 C^∞ 的。这样的 f 称为从 M 到 M' 的**微分同胚映射**，简称**微分同胚**。

注 2.3. ① 微分同胚是对流形间的映射可以提出的最高要求（若流形上还有附加结构则另当别论），互相微分同胚的流形可视作相等。② 只有维数相等的流形才可能微分同胚。③ 若把流形定义中的 O_α 和 V_α 看做流形，则 ψ_α 也是微分同胚。

映射 $f\colon M \to M'$ 的一个重要而简单的特例是 $M' = \mathbb{R}$ 的情况。这时 M 的每点对应着一个实数，于是有如下定义：

定义 2.5. $f\colon M \to \mathbb{R}$ 称为M **上的函数**或M **上的标量场**。若 f 为 C^∞ 的，则称为M **上的光滑函数**。M 上全体光滑函数的的集合记作 $\mathscr{F}_M$，在不会混淆时简记为 $\mathscr{F}$。今后在提到函数而不加声明时都是指光滑函数。

例 2.4. $\mathbb{R}^3$ 中位于 q 点的点电荷的电势是流形 $M \equiv \mathbb{R}^3 - \{q\}$ 上的光滑函数。

例 2.5. 坐标系 (O, ψ) 的第 μ 坐标 x^μ 是定义在 O 上的光滑函数。

函数 $f\colon M \to \mathbb{R}$ 与坐标系 (O, ψ) 结合可得一个 n 元函数 $F(x^1, \cdots, x^n)$，因为 n 个坐标决定 O 中的唯一一点 p，而由 $f\colon M \to \mathbb{R}$ 可得唯一的实数 $f(p)$。然而 f 与另一坐标系 (O', ψ') 结合将给出另一 n 元函数 $F'(x'^1, \cdots, x'^n)$，函数关系 F 和 F' 不同，因为 $F = f \circ \psi^{-1}$ 而 $F' = f \circ \psi'^{-1}$。可见与函数 $f\colon M \to \mathbb{R}$ 相应的多元函数（指函数关系）是坐标系依赖的。应注意区分函数（标量场）f 和它与坐标系结合而得的多元函数 F。

设 M, N 为流形，则它们必为拓扑空间，故 $M \times N$ 也是拓扑空间。不难利用 M, N 的流形结构把 $M \times N$ 进一步定义为流形。设 M, N 的维数分别为 m, n，则 $M \times N$ 的维数是 $m + n$，即 $\dim(M \times N) = \dim M + \dim N$。

2.2 切矢和切矢场

2.2.1 切矢量

定义 2.6. 实数域上的一个**矢量空间**是一个集合 V 配以两个映射，即 $V \times V \to V$（叫**加法**）及 $\mathbb{R} \times V \to V$（叫**数乘**），满足如下条件：

（*a*） $v_1 + v_2 = v_2 + v_1,\ \forall v_1, v_2 \in V$;

（*b*） $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3),\ \forall v_1, v_2, v_3 \in V$;

（*c*） $\exists \underline{0} \in V$, 使 $\underline{0} + v = v,\ \forall v \in V$;

（*d*） $\alpha_1(\alpha_2 v) = (\alpha_1 \alpha_2) v,\ \forall v \in V,\ \alpha_1, \alpha_2 \in \mathbb{R}$;

（*e*） $(\alpha_1 + \alpha_2) v = \alpha_1 v + \alpha_2 v,\ \forall v \in V,\ \alpha_1, \alpha_2 \in \mathbb{R}$;

（*f*） $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2,\ \forall v_1, v_2 \in V,\ \alpha \in \mathbb{R}$;

（g） $1\cdot v = v,\ \forall v \in V$。

注 2.4. 由此 7 条可以推出：（1）$0\cdot v = \underline{0}$，（2）$\forall v \in V,\ \exists u \in V$, 使$v + u = \underline{0}$。约定把 u 记作 $-v$。

今后也常把 V 的零元简写作 0，即符号 0 既代表 $0 \in \mathbb{R}$ 又代表 $\underline{0} \in V$。

定义 2.7. 映射 $v\colon \mathscr{F}_M \to \mathbb{R}$ 称为**点 $p \in M$ 的一个矢量**，若 $\forall f, g \in \mathscr{F}_M$，$\alpha, \beta \in \mathbb{R}$，有

（a）（线性性）$v(\alpha f + \beta g) = \alpha v(f) + \beta v(g)$；

（b）（莱布尼茨律）$v(fg) = f|_p v(g) + g|_p v(f)$，其中 $f|_p$ 代表函数 f 在 p 点的值，亦可记作 $f(p)$。

注 2.5. 因 f 和 g 是 M 上的函数，故 fg 也是 M 上的函数，它在 M 的任一点 p 的值定义为 $f(p)$ 与 $g(p)$ 之积。

根据定义，要定义 p 点的一个矢量只须指定一个从 $\mathscr{F}_M$ 到 $\mathbb{R}$ 的、满足条件（a）、（b）的映射，就是说，指定一个对应规律（法则），根据这一规律，每一 $f \in \mathscr{F}_M$ 对应于一个确定的实数。因为这种映射很多，所以 p 点有很多（无限多）矢量。例如，设 (O, ψ) 是坐标系，其坐标为 x^μ，则 M 上任一光滑函数 $f \in \mathscr{F}_M$ 与 (O, ψ) 结合得 n 元函数 $F(x^1, \cdots, x^n)$，借此可给 O 中任一点 p 定义 n 个矢量，记作 X_μ（其中 $\mu = 1, \cdots, n$），它（们）作用于任一 $f \in \mathscr{F}_M$ 的结果 $X_\mu(f)$ 定义为如下实数：

$$X_\mu(f) \coloneqq \left.\frac{\partial F(x^1, \cdots, x^n)}{\partial x^\mu}\right|_p, \ \forall f \in \mathscr{F}_M$$

其中 $\partial F(x^1, \cdots, x^n)/\partial x^\mu|_p$ 是 $\partial F/\partial x^\mu|_{(x^1(p), \cdots, x^n(p))}$ 的简写。今后也把 $\partial F(x^1, \cdots, x^n)/\partial x^\mu$ 简写为 $\partial f(x^1, \cdots, x^n)/\partial x^\mu$ 或 $\partial f(x)/\partial x^\mu$ 甚至 $\partial f/\partial x^\mu$。应认出 $\partial f/\partial x^\mu$ 中的 f 代表 n 元函数 $F(x^1, \cdots, x^n)$ 而非标量场 f。于是上式可简写为

$$X_\mu(f) \coloneqq \left.\frac{\partial f(x)}{\partial x^\mu}\right|_p, \ \forall f \in \mathscr{F}_M$$

定理 2.1. 以 V_p 代表 M 中点 p 所有矢量的集合，则 V_p 是 n 维矢量空间（n 是 M 的维数），即 $\dim V_p = \dim M \equiv n$。

定义 2.8. 坐标域内任一点 p 的 $\{X_1, \cdots, X_n\}$ 称为 V_p 的一个**坐标基底**，每个 X_μ 称为一个**坐标基矢**，$v \in V_p$ 用 $\{X_\mu\}$ 线性表出的系数 v^μ 称为 v 的**坐标分量**。

定理 2.2. 设 $\{x^\mu\}$ 和 $\{x'^\nu\}$ 为两个坐标系，其坐标域的交集非空，p 为交集中的一点，$v \in V_p$，$\{v^\mu\}$ 和 $\{v'^\nu\}$ 是 v 在这两个系的坐标分量，则

$$v'^\nu = \left.\frac{\partial x'^\nu}{\partial x^\mu}\right|_p v^\mu$$

其中 x'^ν 是两系间坐标变换函数 $x'^\nu(x^\mu)$ 的简写。

证明. 根据坐标分量的定义，$v = v^\mu X_\mu = v'^\nu X'_\nu$。因此可以通过 X_μ 和 X'_ν 之间的关系来找到 v^μ 和 v'^ν 之间的关系。为此，设 q 是两坐标域交集内的任一点，则标量场 f 在 q 点的值 $f|_q$ 满足 $f|_q = f(x(q)) = f'(x'(q))$，简记为 $f(x) = f'(x')$。另一方面，每一 x'^ν 又是 n 个 x^μ 的函数（坐标变换关系），简记为 $x'^\nu = x'^\nu(x)$，故 $f(x) = f'(x'(x))$。于是

$$X_\mu(f) = \left.\frac{\partial f(x)}{\partial x^\mu}\right|_p = \left.\frac{\partial f'(x'(x))}{\partial x^\mu}\right|_p = \left(\frac{\partial f'(x')}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\mu}\right)_p = \left.\frac{\partial x'^\nu}{\partial x^\mu}\right|_p X'_\nu(f),\ \forall f \in \mathscr{F}_M$$

上式表明映射 X_μ 和 $\partial x'^\nu/\partial x^\mu|_p X'_\nu$ 相等，即

$$X_\mu = \left.\frac{\partial x'^\nu}{\partial x^\mu}\right|_p X'_\nu$$

所以 $v = v^\mu X_\mu = v'^\nu X'_\nu$ 可表为

$$v^\mu \left.\frac{\partial x'^\nu}{\partial x^\mu}\right|_p X'_\nu = v'^\nu X'_\nu$$

因 $\{X'_\nu\}$ 中的 n 个基矢彼此线性独立，故得证。 □

上述定理称为**矢量**（的分量）**变换式**，许多书籍采用此式作为矢量的定义。

定义 2.9. 设 I 为 $\mathbb{R}$ 的一个区间，则 C^r 类映射 $C: I \to M$ 称为 M 上的一条 C^r 类的**曲线**。今后如无声明，“曲线”均指光滑（C^∞ 类）曲线。对任一 $t \in I$，有唯一的点 $C(t) \in M$ 与之对应。t 称为曲线的**参数**。

注 2.6. 此处的曲线与直观的曲线概念有密切联系，但也有差别。直观的曲线往往是指上述映射 $C: I \to M$ 的像，即 M 的子集 $C[I]$，并且不提及参数。上述定义的曲线则是指映射本身，是“带参数的曲线”。[3] 设映射 $C: I \to M$ 和 $C': I' \to M$ 的像重合，则直观上往往认为它们是同一曲线，但只要 C 和 C′ 是不同映射，根据定义，它们就是不同曲线。不过在大多数情况下可以说 C 和 C′ 是“同一曲线”的两种参数化，准确地说，曲线 $C': I' \to M$ 称为曲线 $C: I \to M$ 的**重参数化**，若存在到上映射 $\alpha: I \to I'$，满足（*a*）$C = C' \circ \alpha$；（*b*）由 α 诱导的函数 $t' = \alpha(t)$ 有处处非零的导数。解释：由 $C = C' \circ \alpha$ 得

$$C(t) = C'(\alpha(t)) = C'(t'),\ \forall t \in I$$

映射 α 的到上性保证 $C'[I'] = C[I]$，即两曲线映射有相同的像。[4]

注 2.7. ①曲线 C 的像也常记作 $C(t)$（而不是 $C[I]$），以表明曲线的参数为 t。应注意，若 t 为某一具体值（“死的”），则 $C(t)$ 只代表曲线像中的一点；只当把 t 理解为“可跑遍 I”时（“活的”），$C(t)$ 才代表曲线的像。往往也把曲线的像简称为曲线。②设 (O, ψ) 是坐标系，$C[I] \in O$，

[3] 但也存在这样的曲线映射 $C: I \to M$，其像竟然铺满整个 M，很不像直观上的曲线。

[4] α 满足条件（b）保证 α 有一一性，加上到上性便知 C 也是 C′ 的重参数化。

则 $\psi \circ C$ 是从 $I \subset \mathbb{R}$ 到 $\mathbb{R}^n$ 的映射，相当于 n 个一元函数 $x^\mu = x^\mu(t), \mu = 1, \cdots, n$。这 n 个等式称为曲线的**参数方程**或**参数表达式**或**参数式**。一个简单的例子是 $\mathbb{R}^n$ 中以原点为心的单位圆周，其在自然坐标系中的参数式为 $x^1 = \cos t, x^2 = \sin t$。

定义 2.10. 设 (O, ψ) 为坐标系，x^μ 为坐标，则 O 的子集

$$\{p \in O \mid x^2(p) = 常数, \cdots, x^n(p) = 常数\}$$

可以看成以 x^1 为参数的一条曲线（的像）（改变 $x^2, \cdots, x^n$ 的常数值则得另一曲线），叫做x^1 **坐标线**。x^μ **坐标线**可仿此定义。

例 2.6. 在 2 维欧氏空间中，笛卡儿系 $\{x, y\}$ 的 x 及 y 坐标线是互相正交的两组平行直线，极坐标系 $\{r, \varphi\}$ 的 φ 坐标线是以原点为心的无数同心圆，r 坐标线是从原点出发的无数半直线。

直观的想法认为曲线上一点有无限多个彼此平行的切矢。但若把曲线定义为映射（“带参数的曲线”），则一条曲线的一点只有一个切矢。定义如下：

定义 2.11. 设 $C(t)$ 是流形 M 上的 C^1 曲线，则线上 $C(t_0)$ 点切于 $C(t)$ 的切矢 T 是 $C(t_0)$ 点的矢量，它对 $f \in \mathscr{F}_M$ 的作用定义为：

$$T(f) \coloneqq \left.\frac{d(f \circ C)}{dt}\right|_{t_0}, \ \forall f \in \mathscr{F}_M$$

注 2.8. ①$f\colon M \to \mathbb{R}$ 是 M 上的函数（标量场），不是什么一元函数，但与曲线 $C\colon I \to M$ 的结合 $f \circ C$ 便是以 t 为自变数的一元函数（也可记作 $f(C(t))$）。在不会混淆的情况下，$d(f \circ C)/dt$ 也可简写成 df/dt。②$C(t_0)$ 点切于 $C(t)$ 的切矢 T 也常记作 $\partial/\partial t|_{C(t_0)}$，于是上式也可写成

$$\left.\frac{\partial}{\partial t}\right|_{C(t_0)}(f) \coloneqq \left.\frac{d(f \circ C)}{dt}\right|_{t_0} \coloneqq \left.\frac{df(C(t))}{dt}\right|_{t_0}, \ \forall f \in \mathscr{F}_M$$

例 2.7. x^μ 坐标线是以 x^μ 为参数的曲线，又 p 点的坐标基矢 X_μ 就是过 p 的 x^μ 坐标线的切矢，故也常记作 $\partial/\partial x^\mu|_p$，于是坐标基矢 X_μ 对 f 的作用又可表为

$$\left.\frac{\partial}{\partial x^\mu}\right|_p (f) \coloneqq \left.\frac{\partial f(x)}{\partial x^\mu}\right|_p, \ \forall f \in \mathscr{F}_M$$

可见符号 $\partial f/\partial x^\mu$ 既可理解为 $\partial F(x^1, \cdots, x^n)/\partial x^\mu$，又可理解为坐标线的切矢 $\partial/\partial x^\mu$ 对标量场 f 的作用。

定理 2.3. 设曲线 $C(t)$ 在某坐标系中的参数式为 $x^\mu = x^\mu(t)$，则线上任一点的切矢 $\partial/\partial t$ 在该坐标基底的展开式为

$$\frac{\partial}{\partial t} = \frac{dx^\mu(t)}{dt}\frac{\partial}{\partial x^\mu}$$

就是说，曲线 $C(t)$ 的切矢 $\partial/\partial t$ 的坐标分量是 $C(t)$ 在该系的参数式 $x^\mu(t)$ 对 t 的导数。

证明. 注意到 $f(C(t_0)) = f(x^\mu(t_0)) \equiv f(p)$, $\forall f \in \mathscr{F}_M$，因此

$$\left.\frac{\partial}{\partial t}\right|_{C(t_0)}(f) = \left.\frac{df(C(t))}{dt}\right|_{t_0} = \left.\frac{df(x^\mu(t))}{dt}\right|_{t_0} = \left.\frac{\partial f}{\partial x^\mu}\frac{dx^\mu(t)}{dt}\right|_{t_0}, \ \forall f \in \mathscr{F}_M$$

对比等式的两边，可以发现 $\partial/\partial t$ 和 $\frac{dx^\mu(t)}{dt}\partial/\partial x^\mu$ 在同一点 $p = C(t_0)$ 上对任一 f 的作用是相等的，故定理得证。 □

定义 2.12. 非零矢量 $u, v \in V_p$ 称为**互相平行**的，若存在 $\alpha \in \mathbb{R}$ 使 $v = \alpha u$。

由定义可知曲线的切矢依赖于曲线的参数化，一条曲线 $C(t)$ 的一点 $C(t_0)$ 只有一个切于 $C(t)$ 的切矢。直观上之所以认为曲线上一点有无数（互相平行的）切矢，是因为把曲线理解为映射的像而不是映射本身（把无数个有相同像的曲线映射“简并化”为一条曲线）。下面的定理表明，若两条曲线 C 和 C' 的像相同，则它们在任一像点的切矢互相平行。

定理 2.4. 设曲线 $C': I' \to M$ 是 $C: I \to M$ 的重参数化，则两者在任一像点的切矢 $\partial/\partial t$ 和 $\partial/\partial t'$ 有如下关系：

$$\frac{\partial}{\partial t} = \frac{dt'(t)}{dt}\frac{\partial}{\partial t'}$$

其中 $t'(t)$ 是由映射 $\alpha: I \to I'$ 诱导而得的一元函数，即 $\alpha(t)$。

根据定义，$\forall p \in M$，若指定任一曲线 $C(t)$ 使 $p = C(t_0)$，则 V_p 中必有一元素可被视为该曲线在 $C(t_0)$ 点的切矢。现在问：指定 V_p 中任一元素 v，可否找到过 p 的曲线，其在 p 点的切矢是 v。答案是肯定的：这种曲线不但存在，而且很多。例如，任选坐标系 $\{x^\mu\}$ 使 p 含于其坐标域内，则以 $x^\mu(t) = x^\mu|_p + v^\mu(t)$ 为参数式的曲线便是所需曲线，其中 v^μ 是 v 在该系的坐标分量。

综上所述，V_p 中任一元素可视为过 p 的某曲线的切矢，因此 p 点的矢量亦称**切矢量**，V_p 则称为 p 点的**切空间**。

2.2.2 流形上的矢量场

定义 2.13. 设 A 为 M 的子集。若给 A 中每点指定一个矢量，就得到一个定义在 A 上的**矢量场**。

例 2.8. 非自相交曲线 $C(t)$ 上每点的切矢构成 $C(t)$（看作 M 的子集）上的一个矢量场。

设 v 是 M 上的矢量场，f 是 M 上的函数，则 v 在 M 的任一点 p 的值 $v|_p$ 将按定义把 f 映射为一个实数 $v|_p(f)$，它在 p 点跑遍 M 时构成 M 上的一个函数 $v(f)$。因此，矢量场 v 可视为把函数 f 变为函数 $v(f)$ 的映射。

定义 2.14. M 上的矢量场 v 称为**C^∞ 类（光滑）**的，若 v 作用于 C^∞ 类函数的结果仍为 C^∞ 类函数，即 $v(f) \in \mathscr{F}_M$, $\forall f \in \mathscr{F}_M$。$v$ 称为**C^r 类**的，若 v 作用于 C^∞ 类函数得 C^r 类函数。

今后如无声明，“矢量场”均指光滑（C^∞）矢量场。

例 2.9. (*1*) 坐标基矢 $\{X_\mu \equiv \partial/\partial x^\mu\}$ 构成坐标域上的 n 个光滑矢量场，叫**坐标基矢场**。

(*2*) $\mathbb{R}^3$ 中位于 q 的点电荷的静电场强 $\vec{E}$ 是流形 $M \equiv \mathbb{R}^3 - q$ 上的光滑矢量场。

定理 2.5. M 上的矢量场 v 是 C^∞（或 C^r）类的充要条件是它在任一坐标基底的分量 v^μ 为 C^∞（或 C^r）类函数。

设 v 为 M 上的光滑矢量场，则 $v(f) \in \mathscr{F}_M$，$\forall f \in \mathscr{F}_M$。若 u 为 M 上另一光滑矢量场，则 $u(v(f)) \in \mathscr{F}_M$。但函数 $u(v(f))$ 未必等于 $v(u(f))$，于是有如下定义

定义 2.15. 两个光滑矢量场 u 和 v 的**对易子**是一个光滑矢量场 $[u,v]$，定义为

$$[u,v](f) := u(v(f)) - v(u(f)),\ \forall f \in \mathscr{F}_M$$

注 2.9. 上式是对易子 $[u,v]$（作为矢量场）的定义式，它在每点 $p \in M$ 的值 $[u,v]|_p$（作为 p 点的矢量，即从 $\mathscr{F}_M$ 到 $\mathbb{R}$ 的映射）的定义应理解为

$$[u,v]|_p(f) := u|_p(v(f)) - v|_p(u(f)),\ \forall f \in \mathscr{F}_M$$

要确信上式定义的 $[u,v]|_p$ 是 p 点的矢量，还应证明它满足矢量定义的两个条件。

定理 2.6. 设 $\{x^\mu\}$ 为任一坐标系，则 $[\partial/\partial x^\mu, \partial/\partial x^\nu] = 0,\ \mu,\nu = 1,\cdots,n$。

上述定理表明任一坐标系的任意两个基矢场都互相对易。[5]

定义 2.16. 曲线 $C(t)$ 叫矢量场 v 的**积分曲线**，若其上每点的切矢等于该点的 v 值。

定理 2.7. 设 v 是 M 上的光滑矢量场，则 M 的任一点 p 必有 v 的唯一的积分曲线 $C(t)$ 经过（满足 $C(0) = p$）。

下面要用到群论的初步知识，故补充如下定义：

定义 2.17. 一个**群**是一个集合 G 配以满足以下条件的映射 $G \times G \to G$（叫**群乘法**，元素 g_1 和 g_2 的乘积记作 $g_1 g_2$）：

(*a*) $(g_1 g_2) g_3 = g_1(g_2 g_3)$，$\forall g_1, g_2, g_3 \in G$；

(*b*) 存在**恒等元** $e \in G$ 使 $eg = ge = g$，$\forall g \in G$；

(*c*) $\forall g \in G$，存在**逆元** $g^{-1} \in G$ 使 $g^{-1} g = g g^{-1} = e$。

[5] 反之，设 $X_1, \cdots, X_n$ 是点 $p \in M$ 的某邻域 N 上的 n 个处处线性独立的 C^∞ 矢量场，且

$$[X_\mu, X_\nu] = 0,\ \mu,\nu = 1,\cdots,n$$

则必存在坐标系 $\{x^\mu\}$，其坐标域 $O \subset N$ 含 p，且在 O 上有 $X_\mu = \partial/\partial x^\mu$，$\mu = 1,\cdots,n$。

对称性在物理学中有重要意义，群论是研究对称性的有力工具。如果某一对象在某一变换下不变，就说它有该变换下的对称性。考虑带电面上的一个动点，它沿 x（或 y）轴平移。由于动点的电荷面密度 σ 在平移时不变，我们说 σ 具有沿 x（或 y）轴的平移对称性。更明确地说，σ 沿 x 轴的平移对称性是指函数 $\sigma(x,y,z)$ 满足

$$\sigma(x,y,z)=\sigma(x+a,y,z),\ \forall a\in\mathbb{R},$$

其中 $x\mapsto x+a, y\mapsto y, z\mapsto z$ 所代表的平移变换称为沿 x 轴的一个**平移**。设 G 是沿 x 轴的所有平移的集合，则 G 中的元素由实数 a 表征，记作 $\phi_a\in G$。把 $p\equiv(x,y,z)$ 和 $q\equiv(x+a,y,z)$ 看做 $\mathbb{R}^3$ 的点，则上述变换式相当于映射 $\phi_a\colon\mathbb{R}^3\to\mathbb{R}^3$（满足 $\phi_a(p)=q$），而且是微分同胚映射。再者，对 G 用下式定义群乘法

$$\phi_a\phi_b\coloneqq\phi_{a+b},\ \forall\phi_a,\phi_b\in G$$

则 G 构成群（ϕ_0 是恒等元，ϕ_{-a} 是 ϕ_a 的逆元）。这个群的无限多个元素可用实数 a 表征，因此称 a 为**参数**，称 G 为**单参数群**。又因每一群元素 $\phi_a\in G$ 都是 $\mathbb{R}^3$ 上的一个微分同胚，故又称 G 为 $\mathbb{R}^3$ 上的**单参微分同胚群**。

设 M 是流形，则 $\mathbb{R}\times M$ 是比 M 高一维的流形。设有映射 $\phi\colon\mathbb{R}\times M\to M$，则它能把一个实数 $t\in\mathbb{R}$ 和一个点 $p\in M$ 变为一个点 $\phi(t,p)\in M$。若给定 t，记 $\phi_t\equiv\phi(t,\bullet)$，则有映射 $\phi_t\colon M\to M$。同理，若给定 p，记 $\phi_p\equiv\phi(\bullet,p)$，则有映射 $\phi_p\colon\mathbb{R}\to M$。

定义 2.18. C^∞ 映射 $\phi\colon\mathbb{R}\times M\to M$ 称为 M 上的一个**单参微分同胚群**，若

（*a*） $\phi_t\colon M\to M$ 是微分同胚 $\forall t\in\mathbb{R}$；

（*b*） $\phi_t\circ\phi_s=\phi_{t+s}$，$\forall t,s\in\mathbb{R}$。

注 2.10. 集合 $\{\phi_t\mid t\in\mathbb{R}\}$ 是以复合映射为乘法的群，各群元 ϕ_t 是从 M 到 M 的微分同胚映射，ϕ_0 为恒等元（由定义知 $\phi_t\circ\phi_0=\phi_t$，故 ϕ_0 是恒等映射）。所谓 $\phi\colon\mathbb{R}\times M\to M$ 是 M 上的一个单参微分同胚群，其实是指集合 $\{\phi_t\mid t\in\mathbb{R}\}$ 是一个单参微分同胚群。

设 $\phi\colon\mathbb{R}\times M\to M$ 是单参微分同胚群，则 $\forall p\in M$，$\phi_p\colon\mathbb{R}\to M$ 是过 p 点的一条光滑曲线（满足 $\phi_p(0)=p$），叫做这个单参微分同胚群过 p 点的**轨道**。把这条曲线在点 $\phi_p(0)$ 的切矢记作 $v|_p$，便得 M 上的一个光滑矢量场 v。可见 M 上的一个单参微分同胚群给出 M 上的一个光滑矢量场。再看逆命题是否成立。设 v 是 M 上的光滑矢量场，看来 $\forall t\in\mathbb{R}$ 可以借用其积分曲线定义从 M 到 M 的微分同胚映射 ϕ_t（$\forall p\in M$，定义 $\phi_t(p)$ 为这样一个点，它位于过 p 的积分曲线上，其参数值与 p 的参数值之差为 t）。于是似乎可以得到一个单参微分同胚群。然而可能出现如下问题：某条积分曲线当参数取某些值时像点不存在（人为挖去 M 的某一区域就可造出这种情况），因此只能说 M 上的一个光滑矢量场给出一个**单参微分同胚局部群**。

2.3 对偶矢量场

定义 2.19. 设 V 是 $\mathbb{R}$ 上的有限维矢量空间。线性映射 $\omega\colon V \to \mathbb{R}$ 称为 V 上的**对偶矢量**。V 上全体对偶矢量的集合称为 V 的**对偶空间**，记作 V^*。[6]

注 2.11. 由于 V 上有加法和数乘，对映射 ω 的线性要求有确切含义，即

$$\omega(\alpha v + \beta u) = \alpha\omega(v) + \beta\omega(u),\ \forall v, u \in V,\ \alpha, \beta \in \mathbb{R}$$

例 2.10. 设 V 为全体 2×1 实矩阵的集合，则它在矩阵加法和数乘规则下构成 2 维矢量空间。以 ω 代表任一 1×2 实矩阵 (c, d)，其对 V 的任一元素 $v = \begin{pmatrix} a \\ b \end{pmatrix}$ 的作用可用矩阵乘法定义：$\omega(v) \coloneqq (c, d)\begin{pmatrix} a \\ b \end{pmatrix} = (ac + bd)$，结果是一个 1×1 矩阵，可认同为一个实数 $ac + bd$。这样定义的映射 $\omega\colon V \to \mathbb{R}$ 显然是线性的，可见任一 1×2 实矩阵都是 V 上的对偶矢量。推广可知：若把列矩阵（$n \times 1$ 矩阵）看做矢量，则行矩阵（$1 \times n$ 矩阵）就是对偶矢量。

定理 2.8. V^* 是矢量空间，且 $\dim V^* = \dim V$。

证明. 易证 V^* 是矢量空间，只需定义合适的加法、数乘和零元即可。

设 $\{e_\mu\}$ 是 V 的一组基矢，今定义 $e^{\mu *}$ 为 $e^{\mu *}(e_\nu) \coloneqq \delta^\mu{}_\nu$，$\mu, \nu = 1, \cdots, n$。下面证明 $\{e^{\mu *}\}$ 是 V^* 的一组基矢。

易证 $e^{1*}, \cdots, e^{n*}$ 彼此线性独立。

$\forall \omega \in V^*$，令 $\omega_\mu \equiv \omega(e_\mu)$，$\mu = 1, \cdots, n$，则易证 $\omega = \omega_\mu e^{\mu *}$。

因此 V^* 中任一元素可用 $\{e^{\mu *}\}$ 线性表出，故 $\{e^{\mu *}\}$ 是 V^* 的一个基底，叫做 $\{e_\mu\}$ 的**对偶基底**。于是可知 $\dim V^* = \dim V$。 □

两个矢量空间叫**同构**的，若两者之间存在一一到上的线性映射（这种映射称为**同构映射**）。两矢量空间同构的充要条件是维数相同。

由于 $\dim V^* = \dim V$，V^* 与 V 当然同构。同构映射不难找到。例如，设 $\{e_\mu\}$ 是 V 的一个基底，$\{e^{\mu *}\}$ 是其对偶基底，则由 $e_\mu \mapsto e^{\mu *}$ 定义的线性映射就是一个同构映射。但 $\{e_\mu\}$ 的选择十分任意，而基底改变后按上述方式定义的同构映射随之而变，故 V^* 与 V 之间不存在一个特殊的（与众不同的）同构映射，除非在 V 上另加结构。

V^* 既然是矢量空间，自然也有对偶空间，记作 V^{**}。有别于 V 与 V^* 的关系，V 与 V^{**} 间存在一个自然的、与众不同的同构映射，定义如下：$\forall v \in V$，欲给它自然地定义一个像 $v^{**} \in V^{**}$。因为 V^{**} 是 V^* 的对偶空间，v^{**} 应该是从 V^* 到 $\mathbb{R}$ 的线性映射。对它下定义无非是给出一个规律，按照这一规律，每一 $\omega \in V^*$ 对应于唯一的实数 $v^{**}(\omega)$。因为 v^{**}

[6]今后谈及 V 上的对偶矢量（及张量）时如无声明一律默认 V 是实数域上的有限维矢量空间。

有待定义为 v 的像，所以 $v^{**}(\omega)$ 与 v 和 ω 都应有关，而由 v 和 ω 构造的最简单的实数就是 $\omega(v)$，自然把 v^{**} 定义为

$$v^{**}(\omega) \coloneqq \omega(v),\ \forall \omega \in V^*$$

这一映射 $V \to V^{**}$ 是同构映射。这一自然同构关系表明 V 和 V^{**} 可视为同一空间（把每一 $v \in V$ 与其像 $v^{**} \in V^{**}$ 认同）。所以，真正有用的是 V 和 V^*，再取对偶（不论多少次）也得不到更多有用的空间。

定理 2.9. *若矢量空间 V 中有一基底变换 $e'_\mu = A^\nu{}_\mu e_\nu$（$A^\nu{}_\mu$ 无非是新基矢 e'_μ 用原基底展开的第 ν 分量），以 $A^\nu{}_\mu$ 为元素排成的（非退化）方阵记作 A，则相应的对偶基底变换为*

$$e'^{\mu *} = (\tilde{A}^{-1})_\nu{}^\mu e^{\nu *}$$

其中 $\tilde{A}$ 是 A 的转置矩阵，$\tilde{A}^{-1}$ 是 $\tilde{A}$ 之逆。

证明. 只须证明等式两边作用于 e'_α 所得结果相同，证明如下

$$\begin{aligned}
(\tilde{A}^{-1})_\nu{}^\mu e^{\nu *}(e'_\alpha) &= (\tilde{A}^{-1})_\nu{}^\mu e^{\nu *}(A^\beta{}_\alpha e_\beta) \\
&= (\tilde{A}^{-1})_\nu{}^\mu A^\beta{}_\alpha e^{\nu *}(e_\beta) \\
&= (A^{-1})^\mu{}_\nu A^\beta{}_\alpha \delta^\nu{}_\beta \\
&= (A^{-1})^\mu{}_\nu A^\nu{}_\alpha \\
&= \delta^\mu{}_\alpha \\
&= e'^{\mu *}(e'_\alpha)
\end{aligned}$$

□

以上属代数范畴，下面回到流形 M。因 $p \in M$ 有矢量空间 V_p，故也有 $V_p{}^*$。若在 M（或 $A \in M$）上每点指定一个对偶矢量，就得到 M（或 A）上的一个**对偶矢量场**。M 上的对偶矢量场 ω 叫做**光滑**的，若对所有光滑矢量场 v 有 $\omega(v) \in \mathscr{F}_M$。

设 $f \in \mathscr{F}_M$，我们来说明 f 自然诱导出 M 上的一个对偶矢量场，记作 df。要定义 df 只须说明它在 M 的任一点 p 的值 $df|_p \in V_p{}^*$ 的定义，而要定义 $df|_p$ 只须给出它对 p 点任一矢量 $v \in V$ 作用所得的实数，这个实数应与 f 和 v 都有关，而由 f 和 v 能构造的最自然（最简单）的实数便是 $v(f)$，因此定义 $df|_p$ 为

$$df|_p(v) \coloneqq v(f),\ \forall v \in V_p$$

由此易证

$$d(fg)|_p = f|_p(dg)|_p + g|_p(df)|_p$$

这正是微分算符 d 所满足的莱布尼茨律。

设 (O,ψ) 是一坐标系，则第 μ 个坐标 x^μ 可以看作 O 上的函数，于是 dx^μ（看作特殊的 df）是定义在 O 上的对偶矢量场。设 $p\in O$，$\partial/\partial x^\nu$ 是 V_p 的第 ν 个坐标基矢，则在 p 点有

$$dx^\mu\left(\frac{\partial}{\partial x^\nu}\right)=\frac{\partial}{\partial x^\nu}(x^\mu)=\delta^\mu{}_\nu$$

由此可见 $\{dx^\mu\}$ 正是与坐标基底 $\{\partial/\partial x^\nu\}$ 对应的**对偶坐标基底**。上式对 O 的任一点成立，因此，同 $\partial/\partial x^\nu$ 是 O 上的第 ν 个坐标基矢场类似，dx^μ 是 O 上的第 μ 个对偶坐标基矢场，$\{dx^\mu\}$ 则是 O 上的一个对偶坐标基底场。O 上任一对偶矢量场 ω 可借 $\{dx^\mu\}$ 展开：

$$\omega=\omega_\mu dx^\mu$$

其中 ω_μ 称为 ω 在该系的坐标分量，其表达式为：

$$\omega_\mu=\omega(\partial/\partial x^\mu)$$

定理 2.10. 设 (O,ψ) 是一坐标系，f 是 O 上的光滑函数，$f(x)$ 是 $f\circ\psi^{-1}$ 对应的 n 元函数 $f(x^1,\cdots,x^n)$ 的简写，则 df 可用对偶坐标基底 $\{dx^\mu\}$ 展开如下：

$$df=\frac{\partial f(x)}{\partial x^\mu}dx^\mu,\ \forall f\in\mathscr{F}_O$$

证明. 只须验证两边作用于任一坐标基矢 $\partial/\partial x^\nu$ 得相同结果，甚易。 □

定理 2.11. 设坐标系 $\{x^\mu\}$ 和 $\{x'^\nu\}$ 的坐标域有交，则交域中任一点 p 的对偶矢量 ω 在两坐标系中的分量 ω_μ 和 ω'_ν 的变换关系为

$$\omega'_\nu=\left.\frac{\partial x^\mu}{\partial x'^\nu}\right|_p\omega_\mu$$

2.4　张量场

定义 2.20. 矢量空间 V 上的一个**(k,l) 型张量**是一个多重线性映射

$$T\colon\underbrace{V^*\times\cdots\times V^*}_{k\text{ 个}}\times\underbrace{V\times\cdots\times V}_{l\text{ 个}}\to\mathbb{R}$$

注 2.12. T 可比喻为一部机器，有 k 个 “上槽” 和 l 个 “下槽”，只要在上、下槽分别输入 k 个对偶矢量和 l 个矢量，便生产出一个实数，且此实数对每个输入量都线性依赖（多重线性映射的含义）。

例 2.11.（*1*） V 上的对偶矢量是 V 上的 $(0,1)$ 型张量。

（*2*） V 的元素 v 可看作 V 上的 $(1,0)$ 型张量（因 v 可被认同为 v^{**}，而 v^{**} 是从 V^* 到 $\mathbb{R}$ 的线性映射）。

今后用 $\mathscr{T}_V(k,l)$ 表示 V 上全体 (k,l) 型张量的集合，于是 $V=\mathscr{T}_V(1,0)$，$V^*=\mathscr{T}_V(0,1)$。

设 $T\in\mathscr{T}_V(1,1)$，则 $T\colon V^*\times V\to\mathbb{R}$。但 T 也可看成另一映射。因为 $\forall\omega\in V^*, v\in V$ 有 $T(\omega;v)\in\mathbb{R}$，所以 $T(\omega;\bullet)$ 是一部只有下槽的机器，能把一个矢量线性地变为实数，这表明 $T(\omega;\bullet)$ 是 V 上的对偶矢量，即 $T(\omega;\bullet)\in V^*$。给定 T 后，再给一个 $\omega\in V^*$ 便能造出 $T(\omega;\bullet)\in V^*$，故 T 也可看作把对偶矢量 ω 变为对偶矢量 $T(\omega;\bullet)$ 的映射（而且是线性映射），即 $T\colon V^*\xrightarrow[\text{线性地}]{}V^*$。类似地还可以把 T 看成 $T\colon V\xrightarrow[\text{线性地}]{}V$。对同一 $T\in\mathscr{T}_V(1,1)$ 的这 3 种看法是等价的。为便于陈述，我们称这种把同一张量看成不同映射的做法为“张量面面观”。能够用“面面观”想问题是用映射定义张量的重要好处之一。今后将常用到。

定义 2.21. V 上的 (k,l) 和 (k',l') 型张量 T 和 T' 的**张量积**$T\otimes T'$ 是一个 $(k+k',l+l')$ 型张量，定义如下：

$$\begin{aligned}&T\otimes T'(\omega^1,\cdots,\omega^k,\omega^{k+1},\cdots,\omega^{k+k'};v_1,\cdots,v_l,v_{l+1},\cdots,v_{l+l'})\\ \coloneqq&T(\omega^1,\cdots,\omega^k;v_1,\cdots,v_l)T'(\omega^{k+1},\cdots,\omega^{k+k'};v_{l+1},\cdots,v_{l+l'})\end{aligned}$$

欧氏空间矢量场论中的并矢 $\vec{v}\vec{u}$ 其实就是矢量 $\vec{v}$ 和 $\vec{u}$ 的张量积，只不过略去 $\otimes$ 号。[7]

张量积是否满足交换律？设 $\omega\in V^*, v\in V\equiv V^{**}$，则 $v\otimes\omega\in\mathscr{T}_V(1,1), \omega\otimes v\in\mathscr{T}_V(1,1)$。由定义知 $\forall\mu\in V^*, u\in V$ 有 $v\otimes\omega(\mu;u)=v(\mu)\omega(u)=\omega(u)v(\mu)=\omega\otimes v(\mu;u)$（其中 $v(\mu)$ 应理解为 $v^{**}(\mu)$），故 $v\otimes\omega=\omega\otimes v$。但两个矢量（或两个对偶矢量）的张量积在交换顺序后一般成为另一张量，即 $v\otimes u\neq u\otimes v, \omega\otimes\mu\neq\mu\otimes\omega$。例如，欧氏空间的并矢就不满足交换律。

定理 2.12. $\mathscr{T}_V(k,l)$ 是矢量空间，$\dim\mathscr{T}_V(k,l)=n^{k+l}$。

证明.（A） 用自然的方法定义加法、数乘和零元使 $\mathscr{T}_V(k,l)$ 成为矢量空间。

（B） 证明其基矢共 n^{k+l} 个。以 $n=2,k=2,l=1$ 为例（不难推广至一般情况）。设 $\{e_1,e_2\}$ 为 V 的一个基底，$\{e^{1*},e^{2*}\}$ 为其对偶基底。只须证明以下 8 个元素构成 $\mathscr{T}_V(2,1)$ 的一个基底：

$$\begin{gathered}e_1\otimes e_1\otimes e^{1*},\ e_1\otimes e_1\otimes e^{2*},\ e_1\otimes e_2\otimes e^{1*},\ e_1\otimes e_2\otimes e^{2*},\\ e_2\otimes e_1\otimes e^{1*},\ e_2\otimes e_1\otimes e^{2*},\ e_2\otimes e_2\otimes e^{1*},\ e_2\otimes e_2\otimes e^{2*}.\end{gathered}$$

先证它们线性独立，再证任意 $T\in\mathscr{T}_V(2,1)$ 可表为：

$$T=T^{\mu\nu}{}_{\sigma}e_\mu\otimes e_\nu\otimes e^{\sigma*},$$

其中

$$T^{\mu\nu}{}_{\sigma}=T(e^{\mu*},e^{\nu*};e_\sigma).$$

□

[7]类似地，量子力学的 $|\psi\rangle|\phi\rangle$ 也是 $|\psi\rangle$ 和 $|\phi\rangle$ 的张量积，只不过略去 $\otimes$ 号。但量子力学中 $|\psi\rangle$ 所在的矢量空间是复数域上的无限维矢量空间，比现在讨论的实数域上的有限维矢量空间复杂。

注 2.13. $T^{\mu\nu}{}_{\sigma}$ 是张量 T 在基底 $\{e_\mu \otimes e_\nu \otimes e^{\sigma *}\}$ 的分量，简称为 T 在基底 $\{e_\mu\}$ 的分量。

下面介绍张量的另一重要运算，即缩并。如前所述，$(1,1)$ 型张量 T 可看作从 V 到 V 的线性映射，其实它就是线性代数所讲的线性变换。T 在任一基底 $e_\mu \otimes e^{\nu *}$ 的分量排成的矩阵（$T^{\mu}{}_{\nu}$）显然与基底有关，不难证明同一 T 在任意两个基底的分量对应的两个矩阵（$T^{\mu}{}_{\nu}$）和（$T'^{\mu}{}_{\nu}$）互为相似矩阵，证明如下：

$$\begin{aligned}T'^{\mu}{}_{\nu} &= T(e'^{\mu *}; e'_\nu) = T((\tilde{A}^{-1})_\rho{}^{\mu} e^{\rho *}; A^{\sigma}{}_{\nu} e_\sigma) = (\tilde{A}^{-1})_\rho{}^{\mu} A^{\sigma}{}_{\nu} T(e^{\rho *}; e_\sigma) \\ &= (\tilde{A}^{-1})_\rho{}^{\mu} A^{\sigma}{}_{\nu} T^{\rho}{}_{\sigma} = (A^{-1})^{\mu}{}_{\rho} T^{\rho}{}_{\sigma} A^{\sigma}{}_{\nu} = (A^{-1}TA)^{\mu}{}_{\nu}\end{aligned}$$

于是有矩阵等式 $T' = A^{-1}TA$（其中 T'，A，T 都代表矩阵。T 有时代表张量，有时代表矩阵，读者应能根据上下文识别。）。可见 T' 与 T 互为相似矩阵。以 $T'^{\mu}{}_{\mu}$（是 $\sum_{\mu=1}^{n} T'^{\mu}{}_{\mu}$ 的简写）和 $T^{\rho}{}_{\rho}$ 分别代表矩阵 T' 和 T 的迹，则由上式易得：

$$T'^{\mu}{}_{\mu} = (A^{-1})^{\mu}{}_{\rho} T^{\rho}{}_{\sigma} A^{\sigma}{}_{\mu} = A^{\sigma}{}_{\mu} (A^{-1})^{\mu}{}_{\rho} T^{\rho}{}_{\sigma} = \delta^{\sigma}{}_{\rho} T^{\rho}{}_{\sigma} = T^{\rho}{}_{\rho}$$

这就证明了同一 $(1,1)$ 型张量在不同基底的矩阵有相同的迹。在关心张量时，应该抓住其与基底无关的性质，$(1,1)$ 型张量 T 的迹 $T^{\mu}{}_{\mu}$ 就是这样一种性质，通常把它称为 T 的**缩并**或**收缩**，暂记作 CT，即

$$CT \coloneqq T^{\mu}{}_{\mu} = T(e^{\mu *}; e_\mu)$$

再讨论 $(2,1)$ 型张量 T 的缩并。T 可记作 $T(\bullet, \bullet; \bullet)$，它有两个上槽和一个下槽，故有两种可能的缩并：①第一上槽与下槽的缩并 $C_1^1 T \coloneqq T(e^{\mu *}, \bullet; e_\mu)$；②第二上槽与下槽的缩并 $C_1^2 T \coloneqq T(\bullet, e^{\mu *}; e_\mu)$。若改用另一基底 e'_ρ 定义这两种缩并，分别记作 $(C_1^1 T)'$ 和 $(C_1^2 T)'$，则易证 $C_1^1 T = (C_1^1 T)'$，$C_1^2 T = (C_1^2 T)'$。由“张量面面观”可知 $C_1^1 T$ 和 $C_1^2 T$ 都是 $(1,0)$ 型张量，它们在任一基底的分量可用 T 在该基底的分量表为 $(C_1^1 T)^\nu = T(e^{\mu *}, e^{\nu *}; e_\mu) = T^{\mu\nu}{}_{\mu}$，$(C_1^2 T)^\nu = T^{\nu\mu}{}_{\mu}$（已略去求和号）。不难推广上述讨论而得出 (k,l) 型张量的缩并定义如下：

定义 2.22. $T \in \mathscr{T}_V(k,l)$ 的第 i 上标（$i \leqslant k$）与第 j 下标（$j \leqslant l$）的**缩并**定义为：

$$C_j^i T \coloneqq T(\underbrace{\bullet, \cdots, e^{\mu *}}_{\substack{\uparrow \\ \text{第 } i \text{ 上槽}}}, \bullet, \cdots; \underbrace{\bullet, \cdots, e_\mu}_{\substack{\uparrow \\ \text{第 } j \text{ 下槽}}}, \bullet, \cdots) \in \mathscr{T}_V(k-1, l-1) \text{（要对 } \mu \text{ 求和）}$$

注 2.14. ①$C_j^i T$ 与基底选择无关。②由上式易见 (k,l) 型张量的每一缩并都是一个 $(k-1, l-1)$ 型张量。③联合使用张量积和缩并运算可从原有张量得到各种类型的新张量。例如，设 $v \in V$，$\omega \in V^*$，则 $v \otimes \omega$ 是 $(1,1)$ 型张量，而 $C(v \otimes \omega)$ 则是 $(0,0)$ 型张量（标量）。

后面经常遇到先求张量积再做缩并的运算，其结果可看做张量对矢量（或对偶矢量）的作用。作为例子，请考虑如下 3 个等式。

（a） $C(v \otimes \omega) = \omega_\mu v^\mu = \omega(v) = v(\omega)$，$\forall v \in V, \omega \in V^*$。

（b） $C_2^1(T \otimes v) = T(\bullet, v)$, $\forall v \in V, T \in \mathscr{T}_V(0,2)$。

（c） $C_2^2(T \otimes \omega) = T(\bullet, \omega; \bullet)$, $\forall \omega \in V^*, T \in \mathscr{T}_V(2,1)$。

我们只给出（b）的证明。待证等式左边的 $T \otimes v$ 是 $(1,2)$ 型张量，是一部有 1 个上槽、2 个下槽的机器，可表为 $T \otimes v(\bullet; \bullet, \bullet)$，故

$$C_2^1(T \otimes v) = T \otimes v(e^{\mu *}; \bullet, e_\mu)$$

所以只须证明下式

$$T \otimes v(e^{\mu *}; \bullet, e_\mu) = T(\bullet, v)$$

而此式是对偶矢量的等式，只须证明两边作用于任一 $(u \in V)$ 给出相同实数。

$$T \otimes v(e^{\mu *}; u, e_\mu) = T(u, e_\mu) v(e^{\mu *}) = T(u, e_\mu) v^\mu = T(u, v)$$

除以上三式外还有许多类似等式。这些等式是如下规律的表现："T 对 ω（或 v）的作用就是先求 T 与 ω（或 v）的张量积再缩并"，或者粗略地说，"作用就是先积后并"。对两个张量先求张量积再缩并的操作常又简称为对它们做缩并，因此上述粗略提法还可简化为"作用就是缩并"。

下面回到流形 M。M 中任一点 p 的切空间 V_p 的全体 (k,l) 型张量的集合自然记作 $\mathscr{T}_{V_p}(k,l)$。设 $\{e_\mu\}$ 和 $\{e^{\mu *}\}$ 是 V_p 的任一基底及对偶基底，则 T 同样可写成展开式。若选坐标系 $\{x^\mu\}$ 使坐标域含 p，则可用坐标基矢 $\partial/\partial x^\mu$ 和对偶坐标基矢 dx^μ 充当 e_μ 和 $e^{\mu *}$，即

$$T = T^{\mu\nu}{}_\sigma \frac{\partial}{\partial x^\mu} \otimes \frac{\partial}{\partial x^\nu} \otimes dx^\sigma$$

其中坐标分量 $T^{\mu\nu}{}_\sigma$ 可表为

$$T^{\mu\nu}{}_\sigma = T(dx^\mu, dx^\nu; \partial/\partial x^\sigma)$$

若在流形 M（或 $A \subset M$）上每点指定一个 (k,l) 型张量，就得到 M（或 A）上的一个 (k,l) 型**张量场**。M 上张量场 T 称为**光滑的**，若对所有光滑对偶矢量场 $\omega^1, \cdots, \omega^k$ 及光滑矢量场 $v_1, \cdots, v_l$ 有 $T(\omega^1, \cdots, \omega^k; v_1, \cdots, v_l) \in \mathscr{F}_M$。今后如无声明，"张量场"均指光滑（$C^\infty$）张量场。

定理 2.13. (k,l) 型张量在两个坐标系中的分量的变换关系为（简称**张量变换律**）

$$T'^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} = \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \cdots \frac{\partial x'^{\mu_k}}{\partial x^{\rho_k}} \frac{\partial x^{\sigma_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\sigma_l}}{\partial x'^{\nu_l}} T^{\rho_1 \cdots \rho_k}{}_{\sigma_1 \cdots \sigma_l}$$

注 2.15. 许多教科书采用上式作为张量定义。

2.5 度规张量场

定义 2.23. 矢量空间 V 上的一个**度规**g 是 V 上的一个对称、非退化的 $(0,2)$ 型张量。对称是指 $g(v,u)=g(u,v)\ \forall v,u\in V$，非退化是指 $g(v,u)=0\ \forall u\in V\Rightarrow v=0\in V$。

注 2.16. 这一抽象的非退化性定义与矩阵的非退化性（行列式非零）有密切联系。可以证明，若 g 非退化，则它在 V 的任一基底 $\{e_\mu\}$ 的分量 $g_{\mu\nu}\equiv g(e_\mu,e_\nu)$ 排成的矩阵也非退化。反之，若 V 有基底使 g 的分量矩阵非退化，则 g 非退化。

度规很像内积。但上述度规 g 与一般内积的区别在于 $g(v,v)$ 可以为负，且 $g(v,v)=0$ 不意味着 $v=0$。今后也常把 $g(v,u)$ 称为 v 和 u 在度规 g 下的内积。矢量空间 V 一旦定义了度规 g，其元素的长度及元素间的正交性就可以定义如下：

定义 2.24. $v\in V$ 的**长度**或**大小**定义为 $|v|\coloneqq\sqrt{|g(v,v)|}$。矢量 $v,u\in V$ 叫**相互正交**的，若 $g(v,u)=0$。V 的基底 $\{e_\mu\}$ 叫**正交归一**的，若任二基矢正交且每一基矢 e_μ 满足 $g(e_\mu,e_\mu)=\pm1$（不对 μ 求和）。

注 2.17. 上述定义表明度规 g 在正交归一基底的分量满足

$$g_{\mu\nu}=\begin{cases}0, & \mu\neq\nu\\ \pm1, & \mu=\nu\end{cases}$$

因此，度规在正交归一基底的分量排成的矩阵是对角矩阵，且对角元为 $+1$ 或 -1。

定理 2.14. 任何带度规的矢量空间都有正交归一基底。度规写成对角矩阵时对角元中 $+1$ 和 -1 的个数与所选正交归一基底无关。

定义 2.25. 用正交归一基底写成对角矩阵后，对角元全为 $+1$ 的度规叫**正定**的或**黎曼**的，对角元全为 -1 的度规叫**负定**的，其他度规叫**不定**的，只有一个对角元为 -1 的不定度规叫**洛伦兹**的。对角元之和叫度规的**号差**。相对论中用得最多的是洛伦兹度规和正定度规。

定义 2.26. 带洛伦兹度规 g 的矢量空间 V 的元素可分为三类：①满足 $g(v,v)>0$ 的 v 称为**类空矢量**；②满足 $g(v,v)<0$ 的 v 称为**类时矢量**；③满足 $g(v,v)=0$ 的 v 称为**类光矢量**。

注 2.18. 若度规是洛伦兹的，则 $g(v,v)=0$ 未必导致 $v=\underline{0}$。非零的 4 维类光矢量在相对论中有重要地位，例如便于描写电磁波及引力波在 4 维时空中的传播。

度规 g 是 $(0,2)$ 型张量，即由 $V\times V$ 到 $\mathbb{R}$ 的双重线性映射，所以 $\forall v,u\in V$ 有 $g(v,u)\in\mathbb{R}$，因而 $g(v,\bullet)\in V^*$。给定 g 后，再给一个 $v\in V$ 便可造出 $g(v,\bullet)\in V^*$，故 g 可看作由 V 到 V^* 的线性映射，即 $g\colon V\xrightarrow[\text{线性地}]{}V^*$，这是一个同构映射。因此，在 V 选定度规后就有了一

个自然的、与众不同的从 V 到 V^* 的同构映射，我们很自然地用这一映射把 V 与 V^* 认同。小结：无论有无度规，V 都与 V^{**} 自然认同；如果有度规，则 V 与 V^* 也自然认同。

下面回到流形 M 上来。

定义 2.27. M 上的对称的、处处非退化的 $(0,2)$ 型张量场称为**度规张量场**。

注 2.19. 这里我们只关心号差处处一样的度规场。

度规场的一大用处就是定义曲线长度。先讨论 2 维欧氏空间。设曲线 $C(t)$ 在自然坐标系 $\{x,y\}$ 的参数式为 $x=x(t), y=y(t)$，则曲线元段线长的平方 dl^2（是 $(dl)^2$ 的简写）为

$dl^2 = dx^2 + dy^2 = [(dx/dt)^2 + (dy/dt)^2]dt^2 = [(T^1)^2 + (T^2)^2]dt^2 = |T|^2dt^2$，

其中 T 是 $C(t)$ 的切矢。由上式得

$$dl = |T|dt$$

于是 $C(t)$ 的线长为

$$l = \int |T|dt$$

上式可推广至带有正定度规场 g 的任意流形 M 上。设 $C(t)$ 是 M 上任一 C^1 曲线，T 是其切矢，即 $T \equiv \partial/\partial t$，则 $|T| = \sqrt{g(T,T)}$，故 $C(t)$ 的线长自然定义为

$$l \coloneqq \int \sqrt{g(T,T)}dt$$

对有洛伦兹度规场 g 的流形 M，在定义线长前应注意曲线的类型。若 C^1 曲线 $C(t)$ 各点的切矢都类空，则 $C(t)$ 叫**类空曲线**。类似地可定义**类时曲线**和**类光曲线**。类空和类光曲线的线长仍由上式定义（因此类光曲线的线长恒为零）。注意到类时曲线有 $g(T,T) < 0$，其元线长应定义为 $dl \coloneqq \sqrt{-g(T,T)}dt$。于是有如下定义：

定义 2.28. 设流形 M 上有洛伦兹度规场 g，则 M 上的类空、类光及类时曲线 $C(t)$ 的线长定义为

$$l \coloneqq \int \sqrt{|g(T,T)|}dt,\ \text{其中} T \equiv \partial/\partial t$$

对于从类时转向类空（或相反）的曲线（“不伦不类” 的曲线），线长没有定义。下面对线长的讨论虽是就洛伦兹度规而言的，但对正定度规也适用（把所有曲线看作类空曲线）。

不难证明曲线的线长与其参数化无关，就是说，曲线重参数化（保持映射的像不变而适当改变参数）不改变线长。此外，由于线长的定义不涉及坐标系，线长当然与坐标系无关。但是，如果曲线位于坐标系 $\{x^\mu\}$ 的坐标域内，线长也可借助于坐标系计算。因为

$g(T,T) = g(T^\mu\partial/\partial x^\mu, T^\nu\partial/\partial x^\nu) = T^\mu T^\nu g(\partial/\partial x^\mu, \partial/\partial x^\nu) = (dx^\mu/dt)(dx^\nu/dt)g_{\mu\nu}$

（最后一步用到 “曲线切矢的坐标分量等于曲线在该系的参数式对参数的导数”，即 $T^\mu = dx^\mu/dt$。）所以元线长

$$dl = \sqrt{|g_{\mu\nu}dx^\mu dx^\nu|}$$

引入记号 $ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu$，则线长

$$l=\int\sqrt{ds^2}\,(\text{对类空曲线})$$

$$l=\int\sqrt{-ds^2}\,(\text{对类时曲线})$$

记号 ds^2 在微分几何中经常出现，通常称为**线元**。对类空曲线，ds^2 等于元段长 dl 的平方 dl^2; 对类时曲线，ds^2 等于 $-dl^2$，因而不是任何实数的平方。实际上，ds^2 只是一个记号，对类时曲线它根本不是任何实数的平方，但因 $ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu$ 右边含有度规 g 在所涉及的坐标系的全部分量 $g_{\mu\nu}$，从线元表达式可以直接“读出”度规的全体坐标分量。例如，设 2 维流形上的度规 g 在某坐标系 $\{t,x\}$ 的线元表达式为

$$ds^2=-xdt^2+dx^2+4dtdx$$

便可读出 g 在该系的分量为 $g_{tt}=-x, g_{xx}=1, g_{tx}=g_{xt}=2$。可见给定线元（表达式）相当于给定度规场。

设 $C\colon I\to M$ 是类空或类时曲线，则线上任一点 $C(t)$ 的切矢 T 的长度 $|T|$ 是 t 的函数，可记作 $|T|(t)$。任意指定线上一点 $C(t_0)$ 作为测量线长的起点，则介于 $C(t_0)$ 点和 $C(t)$ 点的曲线段的线长 $l(t)=\int_{t_0}^t|T|(t')dt'$ 是 t 的函数。l 也可充当该线的参数，称为**线长参数**。由 $dl\equiv\sqrt{|g(T,T)|}dt$ 可知，以线长为参数的曲线切矢满足 $|g(T,T)|=1$，即有单位长。

定义 2.29. *设流形 M 上给定度规场 g，则 (M,g) 叫***广义黎曼空间***（若 g 为正定，叫***黎曼空间***；若 g 为洛伦兹，叫***伪黎曼空间***，物理上叫***时空***。*[8]*）。*

下面介绍广义黎曼空间的两个简单而重要的例子，即欧氏空间和闵氏空间。

定义 2.30. *设 $\{x^\mu\}$ 是 $\mathbb{R}^n$ 的自然坐标，在 $\mathbb{R}^n$ 上定义度规张量场 δ 为*

$$\delta\coloneqq\delta_{\mu\nu}dx^\mu\otimes dx^\nu$$

则 $(\mathbb{R}^n,\delta)$ 称为n **维欧氏空间**，*δ 称为***欧氏度规**。

上式表明 δ 在自然坐标系的对偶坐标基底 $\{dx^\mu\otimes dx^\nu\}$ 的分量为 $\delta_{\mu\nu}\equiv\begin{cases}0, & \mu\neq\nu,\\ +1, & \mu=\nu,\end{cases}$ 因此，欧氏度规在自然坐标系的线元表达式应为 $ds^2=\delta_{\mu\nu}dx^\mu dx^\nu$。若 $n=2$，便有 $ds^2=(dx^1)^2+(dx^2)^2$。这正是熟知的 2 维欧氏空间的线元表达式。由定义可知自然坐标基底用欧氏度规衡量是正交归一的，因为由

$\delta(\partial/\partial x^\alpha,\partial/\partial x^\beta)=\delta_{\mu\nu}dx^\mu\otimes dx^\nu(\partial/\partial x^\alpha,\partial/\partial x^\beta)=\delta_{\mu\nu}dx^\mu(\partial/\partial x^\alpha)dx^\nu(\partial/\partial x^\beta)$

易见

$$\delta(\partial/\partial x^\alpha,\partial/\partial x^\beta)=\delta_{\alpha\beta}$$

[8] 准确地说，(M,g) 称为时空，若 M 为连通流形，g 为有足够可微程度的洛伦兹度规场。

但满足上式的坐标系未必是自然坐标系。例如，对 2 维欧氏空间，由自然坐标系按下式定义的坐标系

$$x' = x + a,\ y' = y + b\text{（}a, b\text{ 为常数）}$$

的基底 $\{\partial/\partial x', \partial/\partial y'\}$ 也满足上式（因而也正交归一）。进一步，不难验证由以下三式分别定义的 $\{x', y'\}$ 的坐标基底 $\{\partial/\partial x', \partial/\partial y'\}$ 也满足上式：

$$x' = x\cos\alpha + y\sin\alpha,\ y' = -x\sin\alpha + y\cos\alpha\text{（}\alpha\text{ 为常数）},$$

$$x' = -x,\ y' = y,$$

$$x' = x,\ y' = -y.$$

定义 2.31. n 维欧氏空间中满足 $\delta(\partial/\partial x^\alpha, \partial/\partial x^\beta) = \delta_{\alpha\beta}$ 的坐标系叫**笛卡尔坐标系**或**直角坐标系**。换句话说，一个坐标系叫笛卡尔系，若其坐标基底用欧氏度规 δ 衡量为正交归一。

注 2.20. ①因 $\delta \coloneqq \delta_{\mu\nu} dx^\mu \otimes dx^\nu$ 与上述定义等价，也可说满足此式的坐标系是笛卡尔系。②自然坐标系当然是笛卡尔系。③2 维欧氏空间中任意两个笛卡尔系之间的关系只能取上述四式中的一种形式（或它们的复合）。前二种分别称为**平移**和**转动**，后二种的每一种称为**反射**。④要分清符号 δ 和 $\delta_{\mu\nu}$。δ 代表欧氏度规，是张量场；而 $\delta_{\mu\nu}$ 则是 δ 在笛卡尔系的分量。还要注意 δ 在非笛卡尔系的分量不是 $\delta_{\mu\nu}$。

极坐标系 $\{r, \varphi\}$ 是 2 维欧氏空间中非笛卡尔系的一例。物理书中使用极坐标系时，相应的基底常用 $\{\hat{e}_r, \hat{e}_\varphi\}$（顶上加^表示单位矢），它是正交归一的，但却不是极坐标系的坐标基底 $\{\partial/\partial r, \partial/\partial\varphi\}$，关键在于 $\partial/\partial\varphi$ 不归一，因 $\delta(\partial/\partial\varphi, \partial/\partial\varphi) = r^2 \neq 1$[9]。实际上，$\hat{e}_\varphi$ 是对 $\partial/\partial\varphi$ 归一化的产物，即 $\hat{e}_\varphi \coloneqq r^{-1}\partial/\partial\varphi$。可见物理书中常用的 $\{\hat{e}_r, \hat{e}_\varphi\}$ 不是极坐标系的坐标基底而是与极坐标系相应的正交归一基底。

欧氏空间是最简单的黎曼空间。下面介绍最简单的伪黎曼空间—闵氏空间。4 维洛伦兹度规在对角化后的对角元为 $(-1, 1, 1, 1)$，为了突出这个唯一的 -1，我们把它所在的行、列记为 0 行 0 列，三个 $+1$ 所在行、列分别记为 $1, 2, 3$ 行和 $1, 2, 3$ 列。这种对角矩阵的元素记作 $\eta_{\mu\nu}$（以区别于 $\delta_{\mu\nu}$），即 $\eta_{00} \equiv -1, \eta_{11} \equiv \eta_{22} \equiv \eta_{33} \equiv 1$。推广到 n 维则有

$$\eta_{\mu\nu} \equiv \begin{cases} 0,\ \mu \neq \nu, \\ -1,\ \mu = \nu = 0, \\ +1,\ \mu = \nu = 1, \cdots, n-1. \end{cases}$$

下面给出闵氏空间的定义。

定义 2.32. 设 $\{x^\mu\}$ 是 $\mathbb{R}^n$ 的自然坐标，在 $\mathbb{R}^n$ 上定义度规张量场 η 为

$$\eta \coloneqq \eta_{\mu\nu} dx^\mu \otimes dx^\nu$$

[9] 按照矢量的分量变换律，有 $\partial/\partial\varphi = (dx^\mu/d\varphi)(\partial/\partial x^\mu)$。再根据极坐标与直角坐标的关系 $x^1 = r\cos\varphi$，$x^2 = r\sin\varphi$，带入即可求出。

则 $(\mathbb{R}^n, \eta)$ 称为**n 维闵氏空间**（物理上称为**n 维闵氏时空**），η 称为**闵氏度规**。

由定义可知闵氏度规在自然坐标系的线元表达式为 $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$。以 $n = 4$ 为例，有 $ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2$。这正是熟知的 4 维闵氏时空的线元（狭义相对论中的元间隔）表达式。不难证明

$$\eta(\partial/\partial x^\alpha, \partial/\partial x^\beta) = \eta_{\alpha\beta}$$

可见自然坐标基底 $\{\partial/\partial x^\mu\}$ 用闵氏度规衡量也是正交归一的（第 0 坐标基矢归 -1，其他归 $+1$）。但满足上式的却不一定是自然坐标系。例如，以 2 维闵氏空间为例，设 $\{t, x\}$ 是自然坐标，则

$$t' = t + a,\ x' = x + b \text{（}a, b \text{ 为常数）}$$

的坐标基底 $\{\partial/\partial t', \partial/\partial x'\}$ 也满足上式。不难验证，由以下三式分别定义的 $\{t', x'\}$ 系的坐标基底 $\{\partial/\partial t', \partial/\partial x'\}$ 也满足上式：

$$t' = t\,\mathrm{ch}\,\lambda + x\,\mathrm{sh}\,\lambda,\ x' = t\,\mathrm{sh}\,\lambda + x\,\mathrm{ch}\,\lambda \text{（}\lambda \text{ 为常数），}$$
$$t' = -t,\ x' = x,$$
$$t' = t,\ x' = -x.$$

定义 2.33. n 维闵氏空间中满足 $\eta(\partial/\partial x^\alpha, \partial/\partial x^\beta) = \eta_{\alpha\beta}$ 的坐标系叫**洛伦兹坐标系**或**伪笛卡尔坐标系**，也有文献称之为笛卡尔坐标系。

注 2.21. ①闵氏空间的自然坐标当然是洛伦兹坐标。②2 维闵氏空间中的任意两个洛伦兹坐标系之间的关系只能取上述四式中的一种形式（或它们的复合）。第一种称为**平移**，第二种称为**伪转动**，后两种的每一种称为**反射**。③闵氏度规张量 η 在非洛伦兹坐标基底的分量不等于 $\eta_{\mu\nu}$

2.6 抽象指标记号

表示张量的常用方法有两种。第一种是用不带指标的字母（如 T）代表张量，这有两个缺点：① 看不出张量类型；② 不易表明哪一上槽与哪一下槽做缩并（前面所用的记号 C^i_jT 是暂时的，在运算中有诸多不便。）第二种表示法是用分量（如 $T^{\mu\nu}{}_\sigma$）代表张量，用分量服从的等式代表张量服从的等式。分量等式是数量等式，因此在采用这种表示法的文献中所有等式都是数量等式。这种表示法可以克服第一种表示法的两个困难，但自身却有一严重缺点：有时由于选用某一（或某类）特殊基底而得到较为简单的分量等式，它只对特殊基底成立，因而不代表张量等式。我们希望知道哪些等式能够代表张量等式而哪些不能，然而在分量表示法中难以区分。为了克服这一缺点（同时保留分量表示法的所有优点），Penrose 首创“抽象指标记号”，要点如下：

1. (k,l) 型张量用带有 k 个上标和 l 个下标的字母表示，上下标为小写拉丁字母，只表示张量类型，故称**抽象指标**。例如，v^a 代表矢量，上标 a 与 $\vec{v}$ 中的 $\vec{}$ 作用一样（故不能谈及 $a=1$ 或 $a=2$ 的问题），ω_a 代表对偶矢量，$T^{ab}{}_c$ 代表 $(2,1)$ 型张量，等等。v^a 和 v^b 代表相同的矢量（即矢量 $\vec{v}$），但写等式时要注意"指标平衡"，例如可以写 $\alpha u^a + v^a = w^a$ 或 $\alpha u^b + v^b = w^b$ 而不可以写 $\alpha u^a + v^b = w^a$。

2. 重复上下抽象指标表示对这两个指标求缩并。例如

$$T^a{}_a = T(e^{\mu *}; e_\mu) = T^\mu{}_\mu, T^{ab}{}_a = T(e^{\mu *}, \bullet; e_\mu), T^{ab}{}_b = T(\bullet, e^{\mu *}; e_\mu)$$

3. 张量积记号省略。例如，设 $T \in \mathscr{T}_V(2,1), S \in \mathscr{T}_V(1,1)$，则 $T \otimes S$ 写成 $T^{ab}{}_c S^d{}_e$。在不用指标的张量表示法中，一般来说 $\omega \otimes \mu \neq \mu \otimes \omega$，因为当作用于对象 (v,u) 时，ω 作用于 v 还是 u 的问题由字母顺序决定（$\omega \otimes \mu$ 的第一字母 ω 作用于 (v,u) 的第一字母 v）。在抽象指标记号中，由于重复上下指标代表缩并，$\omega \otimes \mu(v,u)$ 既可写成 $\omega_a \mu_b v^a u^b$ 又可写成 $\mu_b \omega_a v^a u^b$（都代表 $\omega(v)\mu(u)$）。既然在这种写法中 $\omega_a \mu_b$ 和 $\mu_b \omega_a$ 的作用对象都是 $v^a u^b$，便有 $\omega_a \mu_b = \mu_b \omega_a$。就是说，代表张量的字母带着自己的抽象指标可以交换。张量积顺序的不可交换性体现为 $\omega_a \mu_b \neq \omega_b \mu_a$。

4. 涉及张量分量时，相应指标用小写希腊字母 μ，ν，α，β 等（正如前面一直用的），这种指标称为**具体指标**，可以问及 $\mu = 1$ 还是 $\mu = 2$ 的问题。张量在基矢上的展开式 $T = T^{\mu\nu}{}_\sigma e_\mu \otimes e_\nu \otimes e^{\sigma *}$ 现在写成

$$T^{ab}{}_c = T^{\mu\nu}{}_\sigma (e_\mu)^a (e_\nu)^b (e^\sigma)_c$$

（$(e^\sigma)_c$ 的抽象下标 c 已表明它是对偶基矢，无须写为 $(e^{\sigma *})_c$。）而 $T^{\mu\nu}{}_\sigma = T(e^{\mu *}, e^{\nu *}; e_\sigma)$ 现在写成

$$T^{\mu\nu}{}_\sigma = T^{ab}{}_c (e^\mu)_a (e^\nu)_b (e_\sigma)^c$$

注意，上两式的指标（无论抽象的还是具体的）都是"平衡"的。设 $T \in \mathscr{T}_V(0,2)$，则 T 应记作 T_{ab}。令 e_μ 为某基底的第 μ 基矢，则可知 $T(\bullet, e_\mu) = C_2^1(T \otimes e_\mu)$，而 $T \otimes e_\mu$ 用抽象指标应记为 $T_{ab}(e_\mu)^c$，故 $T(\bullet, e_\mu)$ 应记作 $T_{ab}(e_\mu)^b$，也可简记作 $T_{a\mu}$，即

$$T(\bullet, e_\mu) \equiv T_{ab}(e_\mu)^b \equiv T_{a\mu}$$

这是既有抽象指标又有具体指标的张量的表达方式，不妨认为 $T_{a1}, \cdots, T_{an}$ 代表 n 个对偶矢量，其中 $T_{a\mu}$ 代表"第 μ 个对偶矢量"。

5. 由"张量面面观"可知，V 上的 $(1,1)$ 型张量 $T^a{}_b$ 既可看作从 V 到 V 的线性映射又可看作从 V^* 到 V^* 的线性映射。就是说，$T^a{}_b$ 作用于矢量 $v^b \in V$ 仍为矢量，记作 $u^a \equiv T^a{}_b v^b \in V$；$T^a{}_b$ 作用于对偶矢量 $\omega_a \in V^*$ 仍为对偶矢量，记作 $\mu_b \equiv T^a{}_b \omega_a \in V^*$。其实，由抽象指标也可一望而知 $T^a{}_b v^b$ 和 $T^a{}_b \omega_a$ 分别是矢量和对偶矢量，可见抽象指标

记号是“张量面面观”的一种简单而直观的体现。以 $\delta^a{}_b$ 代表从 V 到 V 的恒等映射，即 $\delta^a{}_b v^b := v^a\ \forall v^b \in V$，则易见它也是从 V^* 到 V^* 的恒等映射，即 $\delta^a{}_b \omega_a = \omega_b\ \forall \omega_a \in V^*$。进一步不难证明 $\delta^a{}_b$ 与任一张量缩并的结果是把该张量的上标 b 换为 a（或把下标 a 换为 b），例如 $\delta^a{}_b T_{ac} = T_{bc}, \delta^a{}_b T^{cb}{}_e = T^{ca}{}_e$。设 $\{(e_\mu)^a\}$ 是 V 的基底，$\{(e^\mu)_a\}$ 是其对偶基底，则

$$(e^\mu)_a (e_\mu)^b = \delta^b{}_a$$

这是 $(1,1)$ 型张量等式，证明时只须验证两边作用于任一矢量 v^a 得相同结果。设 $\{(e_\mu)^a\}$ 是 V 的基底，$\{(e^\mu)_a\}$ 是其对偶基底，则 $\delta^a{}_b$ 在此基底的分量 $\delta^\mu{}_\nu \equiv \delta^a{}_b (e^\mu)_a (e_\nu)^b$ 满足 $\delta^\mu{}_\nu = \begin{cases} +1, & (\mu = \nu) \\ 0, & (\mu \neq \nu) \end{cases}$。证明很简单，以 $\delta^1{}_1$ 为例，$\delta^1{}_1 = \delta^a{}_b (e^1)_a (e_1)^b = (e^1)_a (e_1)^a = 1$。请注意，即使是洛伦兹号差的情况下也有 $\delta^0{}_0 = +1$。

6. 因度规 $g \in \mathscr{T}_V(0,2)$，故应记为 g_{ab}。设 $v \in V$，则 $g(\bullet, v) \in V^*$。把 g 看作 T，便得 $g(\bullet, v) = C^1_2(g \otimes v) = C^1_2(g_{ab} v^c) = g_{ab} v^b$，故 $g(\bullet, v)$ 应记作 $g_{ab} v^b$。又因有度规 g 时 V 与 V^* 在同构映射 $g\colon V \to V^*$ 下自然认同，而 $g_{ab} v^b \equiv g(\bullet, v)$ 正是 v^a 在这一映射下的像，故应与 v^a 认同，索性就把 $g_{ab} v^b$ 记作 v_a（可看作 v_a 的定义）。就是说，虽然在数学上 v^a 与 v_a 是两种不同性质的量（矢量和对偶矢量），但在应用上两者代表的是同一事物（故都用 v 表示）。于是常写

$$v_a = g_{ab} v^b$$

又由于 $g\colon V \to V^*$ 是同构映射，其逆映射 g^{-1} 自然存在。不难论证 g^{-1} 是 $(2,0)$ 型张量，本应记作 $(g^{-1})^{ab}$，但通常都简记为 g^{ab}（有上指标就不会与 g_{ab} 混淆）。根据类似推理，任一 $\omega_b \in V^*$ 在 g^{ab} 映射下的像为 $g^{ab}\omega_b$，索性记作 ω^a，以表示与 ω_a 代表同一事物，于是

$$\omega^a = g^{ab} \omega_b$$

上述两式表明可用 g_{ab} 及 g^{ab} 对上、下指标分别做“下降”和“上升”处理。这种升降指标操作适用于任何张量中的任何抽象指标。例如，$(1,1)$ 型张量 T 在抽象指标记号中可表为 $T^a{}_b$，所谓用度规对它降指标，其实是用 g 和 T 通过张量积及缩并运算求得一个 $(0,2)$ 型张量 $g(\bullet, e_\mu) \otimes T(e^{\mu *}; \bullet)$，在抽象指标记号中就把它记作 T_{ab}，即 $T_{ab} \equiv g_{ac} T^c{}_b$。

依次使用上述两式得

$$\omega^a = g^{ab}\omega_b = g^{ab}(g_{bc}\omega^c),\ \forall \omega^a \in V$$

故

$$g^{ab} g_{bc} = \delta^a{}_c$$

其实这是 g^{ab} 作为 g_{ab} 的逆映射的必然结果。

设 $\{(e_\mu)^a\}$ 是 V 的任一基底，$\{(e^\mu)_a\}$ 是其对偶基底，以 $g_{\mu\nu}$ 和 $g^{\mu\nu}$ 分别代表 g_{ab} 和 g^{ab} 在这一基底的分量，则

$$g^{\mu\nu}g_{\nu\sigma} = g^{ab}(e^\mu)_a(e^\nu)_b g_{cd}(e_\nu)^c(e_\sigma)^d = g^{ab}(e^\mu)_a g_{bd}(e_\sigma)^d = (e^\mu)_a(e_\sigma)^a = \delta^\mu{}_\sigma$$

其中第二、三步分别用到 $(e^\mu)_a(e_\mu)^b = \delta^b{}_a$ 和 $g^{ab}g_{bc} = \delta^a{}_c$。上式表明度规 g_{ab} 在任一基底的分量 $g_{\mu\nu}$ 的矩阵有逆（逆矩阵就是度规 g_{ab} 之逆 g^{ab} 在同一基底的分量 $g^{\mu\nu}$ 的矩阵），因而非退化。可见 g_{ab} 的非退化性保证它在任一基底的矩阵 $(g_{\mu\nu})$ 的非退化性。反之，设存在基底 $\{(e_\mu)^a\}$ 及其对偶基底 $\{(e^\mu)_a\}$ 使 $(g_{\mu\nu})$ 为非退化，则 $(g_{\mu\nu})$ 有逆矩阵 $(g^{\mu\nu})$，令 $g^{ab} \equiv g^{\mu\nu}(e_\mu)^a(e_\nu)^b$，则由 $g^{\mu\nu}g_{\nu\sigma} = \delta^\mu{}_\sigma$ 易证 $g^{ab}g_{bc} = \delta^a{}_c$，可见 $g_{ab}: V \to V^*$ 有逆映射 g^{ab}，因而非退化（$g_{ab}: V \to V^*$ 有逆表明它是一一映射，而如果 g_{ab} 退化，则 V 中除零元外还有 $v^a \neq 0$，其像也为 $\underline{0} \in V^*$，与一一性矛盾。）。

不难看出，用度规及其逆的分量 $g_{\mu\nu}$ 及 $g^{\mu\nu}$ 可对张量分量的上、下具体指标做下降和上升处理。例如，可把 $g_{\mu\nu}v^\nu$ 写成 v_μ，因为

$$g_{\mu\nu}v^\nu = g_{ab}(e_\mu)^a(e_\nu)^b v^\nu = g_{ab}(e_\mu)^a v^b = v_a(e_\mu)^a = v_\mu$$

作为抽象指标记号的例子，此处介绍 4 维闵氏度规 η_{ab} 的抽象指标表达式。闵氏度规的定义式在抽象指标记号中应表为

$$\eta_{ab} \coloneqq \eta_{\mu\nu}(\mathrm{d}x^\mu)_a(\mathrm{d}x^\nu)_b$$

其中 $\{(\mathrm{d}x^\mu)_a\}$ 是洛伦兹坐标系的对偶基底。以 $\{t, x, y, z\}$ 代表 $\{x^0, x^1, x^2, x^3\}$，则因为非零的 $\eta_{\mu\nu}$ 只有 $\eta_{00} = -1, \eta_{11} = \eta_{22} = \eta_{33} = 1$，上式可表为

$$\eta_{ab} = -(\mathrm{d}t)_a(\mathrm{d}t)_b + (\mathrm{d}x)_a(\mathrm{d}x)_b + (\mathrm{d}y)_a(\mathrm{d}y)_b + (\mathrm{d}z)_a(\mathrm{d}z)_b$$

与线元表达式 $\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$ 相应。如果改用球坐标系 $\{t, r, \theta, \varphi\}$，则由

$$x = r\sin\theta\cos\varphi, y = r\sin\theta\sin\varphi, z = r\cos\theta$$

不难从前式导出

$$\eta_{ab} = -(\mathrm{d}t)_a(\mathrm{d}t)_b + (\mathrm{d}r)_a(\mathrm{d}r)_b + r^2(\mathrm{d}\theta)_a(\mathrm{d}\theta)_b + r^2\sin^2\theta(\mathrm{d}\varphi)_a(\mathrm{d}\varphi)_b$$

与线元表达式 $\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2)$ 相应。

在许多不用抽象指标记号的文献中，4 维时空和 3 维黎曼空间的分量指标分别用希腊字母 $\mu, \nu, \cdots$（都从 0 取到 3）和拉丁字母 $i, j, k, \cdots$（都从 1 取到 3）。拉丁字母在本书中本应代表抽象指标，然而为区分 4 维和 3 维的分量指标，我们允许一个例外，即凡涉及 3 维黎曼空间时用拉丁字母中从 i 起的若干字母 $i, j, k, \cdots$ 充当具体指标（都从 1 取到 3），其他拉丁字母（如 a, b, c 等）仍作为抽象指标。例如 3 维矢量 $\vec{v}$ 可表为 $v^a = v^i(\partial/\partial x^i)^a$（$i$ 从 1 到 3 取和）。

在抽象指标记号中，坐标基矢记作 $(\partial/\partial x^\mu)^a$，对偶坐标基矢记作 $(dx^\mu)_a$。用度规 g_{ab} 和 g^{ab} 对前者和后者分别降、升指标，得对偶矢量 $g_{ab}(\partial/\partial x^\mu)^b$ 和矢量 $g^{ab}(dx^\mu)_b$。以 ω_a 简记 $g_{ab}(\partial/\partial x^\mu)^b$ 并用对偶坐标基矢展开为 $\omega_a = \omega_\nu (dx^\nu)_a$，两边作用于 $(\partial/\partial x^\sigma)^a$ 后得 $g_{\sigma\mu} = \omega_\sigma$[10]，故

$$g_{ab}(\partial/\partial x^\mu)^b = g_{\mu\nu}(dx^\nu)_a$$

可见 $g_{ab}(\partial/\partial x^\mu)^b$ 一般不等于 $(dx^\mu)_a$。类似可得

$$g^{ab}(dx^\mu)_b = g^{\mu\nu}(\partial/\partial x^\nu)^a$$

当 $g_{ab} = \delta_{ab}$（欧氏度规）且 $\{x^\mu\}$ 为笛卡尔系时上二式简化为

$$\delta_{ab}(\partial/\partial x^\mu)^b = (dx^\mu)_a,\ \delta^{ab}(dx^\mu)_b = (\partial/\partial x^\mu)^a$$

当 $g_{ab} = \eta_{ab}$（以 4 维闵氏度规为例）且 $\{x^\mu\}$ 为洛伦兹系时则有

$$\begin{aligned}\eta_{ab}(\partial/\partial x^0)^b = -(dx^0)_a,\ \eta_{ab}(\partial/\partial x^i)^b = (dx^i)_a;\\ \eta^{ab}(dx^0)_b = -(\partial/\partial x^0)^a,\ \eta^{ab}(dx^i)_b = (\partial/\partial x^i)^a.\end{aligned}$$

其中的 $i = 1, 2, 3$，此时的 i 不是抽象指标。

张量的上指标和下指标在文献中又常分别称为**逆变指标**和**协变指标**。相应地，矢量 v^a 和对偶矢量 ω_a 也分别称为**逆变矢量**和**协变矢量**。

张量的对称性可方便地用抽象指标表述如下：

定义 2.34. *$T \in \mathscr{T}_V(0,2)$ 称为**对称的**，若 $T(u,v) = T(v,u)$，$\forall u, v \in V$。*

由于 $T(u,v) = T_{ab}u^a v^b, T(v,u) = T_{ab}v^a u^b = T_{ba}u^a v^b$，故在抽象指标记号中 T 为对称的充要条件是 $T_{ab} = T_{ba}$。$(0,2)$ 型张量在抽象指标记号中本来既可记作 T_{ab} 又可记作 T_{ba}，两者代表同一张量。然而只当 T 为对称张量时才允许写为 $T_{ab} = T_{ba}$，可见用抽象指标写等式时比用它单独表示一个张量时要更为小心。同理，$(1,1)$ 型张量既可表为 $T^a{}_b$，也可表为 $T_b{}^a$，用度规降指标后分别为 $g_{ca}T^a{}_b = T_{cb}$ 和 $g_{ca}T_b{}^a = T_{bc}$。虽然两者代表同一张量，但只有降指标后为对称张量的 $(1,1)$ 型张量才允许写为 $T^a{}_b = T_b{}^a$。在不用度规升降指标时，(k,l) 型张量的上、下指标各排各的序，两个上下指标之间没有顺序问题。因此，如果愿意，$(1,1)$ 型张量可写为 T^a_b，$(2,1)$ 型张量可写为 T^{ab}_c 等。然而这种写法在用度规升降指标时出现不确定性。由于经常要升降指标，本书从一开始就把上、下两排指标错开，例如写成 $T^{ab}{}_c$。

以上讨论表明抽象指标记号在形式上与具体指标记号极为相似，这正是抽象指标记号的一大好处：它既可表示张量等式，又保留了具体指标记号的许多优点。

[10]左边 $= g_{ab}(\partial/\partial x^\mu)^b(\partial/\partial x^\sigma)^a = g_{a\mu}(\partial/\partial x^\sigma)^a = g_{\sigma\mu}$，右边 $= \omega_\nu(dx^\nu)_a(\partial/\partial x^\sigma)^a = \omega_\nu\delta^\nu{}_\sigma = \omega_\sigma$

定义 2.35. $(0,2)$ 型张量 T_{ab} 的**对称部分**（记作 $T_{(ab)}$）和**反称部分**（记作 $T_{[ab]}$）分别定义为

$$T_{(ab)} \coloneqq \frac{1}{2}(T_{ab}+T_{ba}),\ T_{[ab]} \coloneqq \frac{1}{2}(T_{ab}-T_{ba})$$

一般地，$(0,l)$ 型张量 $T_{a_1\cdots a_l}$ 的对称和反称部分定义为

$$T_{(a_1\cdots a_l)} \coloneqq \frac{1}{l!}\sum_{\pi} T_{a_{\pi(1)}\cdots a_{\pi(l)}},$$
$$T_{[a_1\cdots a_l]} \coloneqq \frac{1}{l!}\sum_{\pi} \delta_\pi T_{a_{\pi(1)}\cdots a_{\pi(l)}},$$

其中 π 代表 $(1,\cdots,l)$ 的一种排列，$\pi(1)$ 是指 π 所代表的那种排列中的第 1 个数字，$\sum\limits_{\pi}$ 代表对各种排列取和，$\delta_\pi \equiv \pm 1$（偶排列取 $+$，奇排列取 $-$。）。例如

$$T_{(a_1a_2a_3)} \coloneqq \frac{1}{6}(T_{a_1a_2a_3}+T_{a_3a_1a_2}+T_{a_2a_3a_1}+T_{a_1a_3a_2}+T_{a_3a_2a_1}+T_{a_2a_1a_3}),$$
$$T_{[a_1a_2a_3]} \coloneqq \frac{1}{6}(T_{a_1a_2a_3}+T_{a_3a_1a_2}+T_{a_2a_3a_1}-T_{a_1a_3a_2}-T_{a_3a_2a_1}-T_{a_2a_1a_3}).$$

定义 2.36. $T\in\mathscr{T}_V(0,l)$ 称为**全对称的**，若 $T_{a_1\cdots a_l}=T_{(a_1\cdots a_l)}$；T 称为**全反称的**，若 $T_{a_1\cdots a_l}=T_{[a_1\cdots a_l]}$。

以上三个定义也适用于 $(k,0)$ 型张量。例如，T 叫全对称的，若 $T^{a_1\cdots a_l}=T^{(a_1\cdots a_l)}$。

注 2.22. 任一 $(0,2)$ 型张量可表为其对称和反称部分之和，即 $T_{ab}=T_{(ab)}+T_{[ab]}$，但对 $l>2$ 的 $(0,l)$ 型张量不成立。例如，$T_{abc}\neq T_{(abc)}+T_{[abc]}$，然而 $T_{abc}=T_{(abc)}\Rightarrow T_{[abc]}=0$（见下文定理）。

定理 2.15.（*a*） 设 $T_{a_1\cdots a_l}=T_{(a_1\cdots a_l)}$，则

$$T_{a_1\cdots a_l}=T_{a_{\pi(1)}\cdots a_{\pi(l)}}(\pi \text{ 代表任一种排列})$$

即 $T_{(a_1\cdots a_l)}$ 展开式（共 $l!$ 项）中每一项都等于 $T_{a_1\cdots a_l}$，例如

$$T_{abc}=T_{(abc)}\Rightarrow T_{abc}=T_{acb}=T_{cab}=T_{cba}=T_{bca}=T_{bac}$$

（*b*） 设 $T_{a_1\cdots a_l}=T_{[a_1\cdots a_l]}$，则

$$T_{a_1\cdots a_l}=\delta_\pi T_{a_{\pi(1)}\cdots a_{\pi(l)}}$$

即 $T_{[a_1\cdots a_l]}$ 展开式中的偶排列项等于 $T_{a_1\cdots a_l}$，奇排列项等于 $-T_{a_1\cdots a_l}$，例如

$$T_{abc}=T_{[abc]}\Rightarrow T_{abc}=-T_{acb}=T_{cab}=-T_{cba}=T_{bca}=-T_{bac}$$

对 $(k,0)$ 型（上指标）全对称和全反称张量也有类似结论。

今后会经常遇到对带有圆、方括号的指标的运算，下面的定理对许多运算将带来很大方便。

定理 2.16.（a） 缩并时括号有“传染性”，即

$$T_{[a_1\cdots a_l]}S^{a_1\cdots a_l} = T_{[a_1\cdots a_l]}S^{[a_1\cdots a_l]} = T_{a_1\cdots a_l}S^{[a_1\cdots a_l]}$$

对圆括号亦然。

（b） 括号内的同种子括号可随意增删，例如

$$T_{[[ab]c]} = T_{[abc]}，\text{其中 } T_{[[ab]c]} \equiv \frac{1}{2}(T_{[abc]} - T_{[bac]})$$

（c） 括号内加异种子括号得零，例如

$$T_{[(ab)c]} = 0,\ T_{(a[bcd])} = 0$$

（d） 异种子括号缩并得零，例如

$$T^{(abc)}S_{[abc]} = 0$$

（e）

$$T_{a_1\cdots a_l} = T_{(a_1\cdots a_l)} \Rightarrow T_{[a_1\cdots a_l]} = 0$$
$$T_{a_1\cdots a_l} = T_{[a_1\cdots a_l]} \Rightarrow T_{(a_1\cdots a_l)} = 0$$

对 $(k, 0)$ 型（上指标）全对称和全反称张量也有类似结论。

第三章 黎曼（内禀）曲率张量

3.1 导数算符

欧氏空间有熟知的导数算符 $\vec{\nabla}$，它作用于函数（标量场）f 得矢量场 $\vec{\nabla}f$（梯度），作用于矢量场 $\vec{v}$（再求缩并）得标量场 $\vec{\nabla}\cdot\vec{v}$（散度）等。由于存在欧氏度规 δ_{ab}，欧氏空间的矢量 v^a 与对偶矢量 $v_a = \delta_{ab}v^b$ 自然认同。现在要把 $\vec{\nabla}$ 推广到任意流形，其上可以没有度规，所以要分清矢量和对偶矢量。研究发现在推广时 $\vec{\nabla}$ 更像对偶矢量，故应记作 ∇_a。其实 ∇ 本身是算符，既非矢量也非对偶矢量，所谓把 ∇ 看作对偶矢量是指它作用于函数 f 的结果 $\nabla_a f$ 是对偶矢量。推而广之，∇ 作用于任一 (k,l) 型张量场的结果是 $(k,l+1)$ 型张量场。于是有如下定义：

定义 3.1. 以 $\mathscr{F}_M(k,l)$ 代表流形 M 上全体 C^∞ 的 (k,l) 型张量场的集合（函数 f 可看作 $(0,0)$ 型张量场（标量场），故 $\mathscr{F}_M(0,0)\equiv\mathscr{F}_M$）。映射 $\nabla\colon \mathscr{F}_M(k,l)\to\mathscr{F}_M(k,l+1)$ 称为 M 上的（无挠）**导数算符**[1]，若它满足如下条件：

（*a*） 具有线性性：

$$\nabla_a(\alpha T^{b_1\cdots b_k}{}_{c_1\cdots c_l} + \beta S^{b_1\cdots b_k}{}_{c_1\cdots c_l}) = \alpha\nabla_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l} + \beta\nabla_a S^{b_1\cdots b_k}{}_{c_1\cdots c_l}$$
$$\forall T^{b_1\cdots b_k}{}_{c_1\cdots c_l}, S^{b_1\cdots b_k}{}_{c_1\cdots c_l}\in\mathscr{F}_M(k,l),\ \alpha,\beta\in\mathbb{R}$$

（*b*） 满足莱布尼茨律：

$$\nabla_a(T^{b_1\cdots b_k}{}_{c_1\cdots c_l}S^{d_1\cdots d_{k'}}{}_{e_1\cdots e_{l'}}) = T^{b_1\cdots b_k}{}_{c_1\cdots c_l}\nabla_a S^{d_1\cdots d_{k'}}{}_{e_1\cdots e_{l'}} + S^{d_1\cdots d_{k'}}{}_{e_1\cdots e_{l'}}\nabla_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l}$$
$$\forall T^{b_1\cdots b_k}{}_{c_1\cdots c_l}\in\mathscr{F}_M(k,l),\ S^{d_1\cdots d_{k'}}{}_{e_1\cdots e_{l'}}\in\mathscr{F}_M(k',l')$$

（*c*） 与缩并可交换顺序

（*d*） $v(f) = v^a\nabla_a f,\ \forall f\in\mathscr{F}_M, v\in\mathscr{F}_M(1,0)$

（*e*） 具有无挠性：$\nabla_a\nabla_b f = \nabla_b\nabla_a f,\ \forall f\in\mathscr{F}_M$

[1] $\mathscr{F}(k,l)$ 可放宽为全体 C^1 类 (k,l) 型张量场的集合。就是说，∇_a 可作用于任一 C^1 类张量场。

注 3.1.（*1*） 条件（*c*）又可表为 $\nabla \circ C = C \circ \nabla$，其中 C 代表缩并。今后将常写

$$\nabla_a(v^b\omega_b) = v^b\nabla_a\omega_b + \omega_b\nabla_a v^b$$

一类的式子，这就要用到条件（*c*），因为上式的导出过程为

$$\begin{aligned}\nabla_a(v^b\omega_b) &= \nabla_a[C(v^b\omega_c)] = C_2^1[\nabla_a(v^b\omega_c)] \\ &= C_2^1(v^b\nabla_a\omega_c) + C_2^1[(\nabla_a v^b)\omega_c] = v^b\nabla_a\omega_b + \omega_b\nabla_a v^b\end{aligned}$$

（*2*） 条件（*d*）左边的函数 $v(f)$ 不宜记作 $v^a(f)$，因 $v^a(f)$ 易被误以为矢量场。这是应写而不写抽象指标的少数例子之一。对条件（*d*）可用欧氏空间的 $\vec{\nabla}$ 为例理解。设 v^a 为欧氏空间中任一矢量场，其在笛卡尔系坐标基底的展开式为

$$v^a = v^1(\partial/\partial x)^a + v^2(\partial/\partial y)^a + v^3(\partial/\partial z)^a$$

则它对函数 f 的作用可表为

$$v(f) = v^1(\partial f/\partial x)^a + v^2(\partial f/\partial y)^a + v^3(\partial f/\partial z)^a = \vec{v}\cdot\vec{\nabla} f = v^a\nabla_a f$$

可见条件（*d*）是这一性质对任意流形的推广。

（*3*） 设 ∇_a 是任一导数算符，则由条件（*d*）易证

$$\nabla_a f = (\mathrm{d}f)_a,\ \forall f \in \mathscr{F}_M$$

其中，$(\mathrm{d}f)_a$ 是函数 f 生成的对偶矢量场 df 的抽象指标表示。

（*4*） 条件（*e*）实质上是下式的抽象指标表述：

$$(\nabla\nabla f)(u,v) = (\nabla\nabla f)(v,u),\ \forall u,v \in \mathscr{F}_M(1,0)$$

亦即 $\nabla\nabla f$ 是个对称的 $(0,2)$ 型张量。

（*5*） 满足条件（*a*）（*d*）而不满足条件（*e*）的导数算符叫**有挠导数算符**。广义相对论中只用无挠导数算符。本书的 ∇_a 在不加声明时一律代表无挠导数算符。

任何流形必定存在满足上述定义的导数算符。事实上，导数算符不但存在，而且很多。下面讨论多到什么程度。由 $\nabla_a f = (\mathrm{d}f)_a$ 可知任意两个导数算符 ∇_a 和 $\tilde{\nabla}_a$ 作用于同一函数的结果相同，即

$$\nabla_a f = \tilde{\nabla}_a f = (\mathrm{d}f)_a,\ \forall f \in \mathscr{F}_M$$

可见 ∇_a 与 $\tilde{\nabla}_a$ 的不同只能体现在对非 $(0,0)$ 型张量场的作用上。先讨论 $(0,1)$ 型张量场（对偶矢量场）。设在点 $p \in M$ 给定一个对偶矢量 $\mu_b \in V_p^*$，考虑 M 上的任意两个对偶矢量场 $\omega_b, \omega'_b \in \mathscr{F}_M(0,1)$，满足 $\omega'_b|_p = \omega_b|_p = \mu_b$（$\omega_b$ 和 ω'_b 称为 μ_b 在 M 上的两个延拓）。

设 ∇_a 为导数算符，则 $\nabla_a\omega'_b|_p$ 与 $\nabla_a\omega_b|_p$ 一般并不相同。这类似于以下事实：两个一元函数 $f(x)$ 和 $f'(x)$ 在 x_0 点取值相同（$f'(x_0) = f(x_0)$）并不保证 $(df'/dx)|_{x_0} = (df/dx)|_{x_0}$。然而下面要证明，对 M 上任意两个导数算符 ∇_a 和 $\tilde{\nabla}_a$，只要 $\omega'_b|_p = \omega_b|_p$ 就有

$$[(\tilde{\nabla}_a - \nabla_a)\omega'_b]_p = [(\tilde{\nabla}_a - \nabla_a)\omega_b]_p$$

其中 $(\tilde{\nabla}_a - \nabla_a)\omega_b$ 是 $\tilde{\nabla}_a\omega_b - \nabla_a\omega_b$ 的简写。

定理 3.1. 设 $p \in M$，$\omega_b, \omega'_b \in \mathscr{F}(0,1)$ 满足 $\omega'_b|_p = \omega_b|_p$，则

$$[(\tilde{\nabla}_a - \nabla_a)\omega'_b]_p = [(\tilde{\nabla}_a - \nabla_a)\omega_b]_p$$

证明. 只须证明

$$[\nabla_a(\omega'_b - \omega_b)]_p = [\tilde{\nabla}_a(\omega'_b - \omega_b)]_p$$

设 $\Omega_b \equiv \omega'_b - \omega_b$，选坐标系 $\{x^\mu\}$ 使其坐标域含 p，则 $\omega'_b|_p = \omega_b|_p$ 导致 $\Omega_\mu(p) = 0$，其中 Ω_μ 是 Ω_b 的坐标分量。于是对 p 点有

$$\begin{aligned}[\nabla_a(\omega'_b - \omega_b)]_p &= [\nabla_a\Omega_b]_p = \{\nabla_a[\Omega_\mu(dx^\mu)_b]\}|_p \\ &= \Omega_\mu(p)[\nabla_a(dx^\mu)_b]_p + [(dx^\mu)_b\nabla_a\Omega_\mu]_p = [(dx^\mu)_b\nabla_a\Omega_\mu]_p\end{aligned}$$

同理有 $[\tilde{\nabla}_a(\omega'_b - \omega_b)]_p = [(dx^\mu)_b\tilde{\nabla}_a\Omega_\mu]_p$。又 $[\nabla_a\Omega_\mu]_p = [\tilde{\nabla}_a\Omega_\mu]_p$，得证。 □

虽然导数 $[\nabla_a\omega_b]_p$ 和 $[\tilde{\nabla}_a\omega_b]_p$ 依赖于 ω_b 在 p 点的一个邻域内的值，然而上述定理表明 $[(\tilde{\nabla}_a - \nabla_a)\omega_b]_p$ 只依赖于 ω_b 在 p 点的值，这说明 $(\tilde{\nabla}_a - \nabla_a)$ 是把 p 点的对偶矢量 $\omega_b|_p$ 变为 p 点的 $(0,2)$ 型张量（$[(\tilde{\nabla}_a - \nabla_a)\omega_b]_p$）的线性映射（给定 p 点的任一对偶矢量 μ_b，任选对偶矢量场 ω_b 使它在 p 点的值 $\omega_b|_p = \mu_b$，则 $[(\tilde{\nabla}_a - \nabla_a)\omega_b]_p$ 便是 μ_b 在该映射下的像。）。所以 $(\tilde{\nabla}_a - \nabla_a)$ 在 p 点对应于一个 $(1,2)$ 型张量 $C^c{}_{ab}$，满足

$$[(\tilde{\nabla}_a - \nabla_a)\omega_b]_p = C^c{}_{ab}\omega_c|_p$$

因为 p 点可任选，所以 M 上的两个导数算符 ∇_a 和 $\tilde{\nabla}_a$ 在对 ω_b 的作用上的差别体现为 M 上的一个 $(1,2)$ 型张量场 $C^c{}_{ab}$，即

定理 3.2. $\nabla_a\omega_b = \tilde{\nabla}_a\omega_b - C^c{}_{ab}\omega_c$，$\forall\omega_b \in \mathscr{F}(0,1)$

∇_a 的无挠性导致张量场 $C^c{}_{ab}$ 的如下对称性：

定理 3.3. $C^c{}_{ab} = C^c{}_{ba}$

证明. 令 $\omega_b = \nabla_b f = \tilde{\nabla}_b f$，其中 $f \in \mathscr{F}_M$，则上式给出 $\nabla_a\nabla_b f = \tilde{\nabla}_a\tilde{\nabla}_b f - C^c{}_{ab}\nabla_c f$。交换指标 a，b 得 $\nabla_b\nabla_a f = \tilde{\nabla}_b\tilde{\nabla}_a f - C^c{}_{ba}\nabla_c f$。两式相减，注意到无挠性条件，便有 $C^c{}_{ab}\nabla_c f = C^c{}_{ba}\nabla_c f$。令 $T^c{}_{ab} \equiv C^c{}_{ab} - C^c{}_{ba}$，则对所有的 $f \in \mathscr{F}_M$ 有 $T^c{}_{ab}\nabla_c f = 0$，于是 $T^c{}_{ab}$ 在任一坐标基底的分量 $T^\sigma{}_{\mu\nu} = T^c{}_{ab}(dx^\sigma)_c(\partial/\partial x^\mu)_a(\partial/\partial x^\nu)_b = 0$（其中第二步是因为 $T^c{}_{ab}(dx^\sigma)_c = T^c{}_{ab}\nabla_c x^\sigma = 0$（把 x^σ 看作 f）），因而 $T^c{}_{ab} = 0$。 □

定理 3.4. $\nabla_a v^b = \tilde{\nabla}_a v^b + C^b{}_{ac} v^c$, $\forall v^b \in \mathscr{F}_M(1,0)$

证明. 设 ω_b 为 M 上任一对偶矢量场，则

$$\nabla_a(\omega_b v^b) = \omega_b \nabla_a v^b + v^b \nabla_a \omega_b = \omega_b \nabla_a v^b + v^b(\tilde{\nabla}_a \omega_b - C^c{}_{ab}\omega_c)$$

另一方面,$\tilde{\nabla}_a(\omega_b v^b) = \omega_b \tilde{\nabla}_a v^b + v^b \tilde{\nabla}_a \omega_b$。而 $\omega_b v^b$ 为标量场,故 $\nabla_a(\omega_b v^b) = \tilde{\nabla}_a(\omega_b v^b)$，故以上两式右边相等，因而得

$$\omega_b \nabla_a v^b = \omega_b \tilde{\nabla}_a v^b + C^c{}_{ab} v^b \omega_c = \omega_b \tilde{\nabla}_a v^b + C^b{}_{ac} v^c \omega_b,\ \forall \omega_b \in \mathscr{F}_M(0,1)$$

于是有定理结论。 □

用类似方法可以证明 ∇_a 与 $\tilde{\nabla}_a$ 作用于任一 (k,l) 型张量场 $T^{a_1\cdots a_k}{}_{b_1\cdots b_l}$ 所得结果之差 $\nabla_a T^{a_1\cdots a_k}{}_{b_1\cdots b_l} - \tilde{\nabla}_a T^{a_1\cdots a_k}{}_{b_1\cdots b_l}$ 可表为 $k+l$ 项，每项都含 $C^c{}_{ab}$，与 T 的某一上指标缩并的 k 项前面为 + 号，与 T 的某一下指标缩并的 l 项前面为 − 号，例如

$$\nabla_a T^b{}_c = \tilde{\nabla}_a T^b{}_c + C^b{}_{ad} T^d{}_c - C^d{}_{ac} T^b{}_d$$

一般形式见下面的定理：

定理 3.5.

$$\nabla_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l} = \tilde{\nabla}_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l} + \sum_i C^{b_i}{}_{ad} T^{b_1\cdots d\cdots b_k}{}_{c_1\cdots c_l} - \sum_j C^d{}_{ac_j} T^{b_1\cdots b_k}{}_{c_1\cdots d\cdots c_l},\ \forall T \in \mathscr{F}_M(k,l)$$

上述定理表明任意两个导数算符的差别仅体现在一个张量场 $C^c{}_{ab}$ 上。反之也不难验证，任给一个导数算符 $\tilde{\nabla}_a$ 和一个下标对称的光滑张量场 $C^c{}_{ab}$，由上式定义的 ∇_a 必满足导数算符的全部条件，因而也是一个导数算符。可见流形上只要有一个导数算符就会有许多导数算符。选定导数算符 ∇_a 后的流形 M 可记作 (M,∇_a)，它比 M 本身有更多结构（∇_a 提供附加结构），例如可谈及矢量沿曲线的平移及 (M,∇_a) 的曲率。

设 $\{x^\mu\}$ 是 M 的一个坐标系,其坐标基底和对偶坐标基底分别为 $\{(\partial/\partial x^\mu)^a\}$ 和 $\{(dx^\mu)_a\}$。在坐标域 O 上定义映射 $\partial_a\colon \mathscr{F}_O(k,l) \to \mathscr{F}_O(k,l+1)$ 如下（仅以 $T^b{}_c \in \mathscr{F}_O(1,1)$ 为例写出）：

$$\partial_a T^b{}_c \coloneqq (dx^\mu)_a (\partial/\partial x^\nu)^b (dx^\sigma)_c \partial_\mu T^\nu{}_\sigma$$

其中 $T^\nu{}_\sigma$ 是 $T^b{}_c$ 在该坐标系的分量，∂_μ 是对坐标 x^μ 求偏导数的符号 $\partial/\partial x^\mu$ 的简写。不难验证 ∂_a 满足导数算符定义的 5 个条件，可见 ∂_a 是 O 上的一个导数算符。这是一个从定义起就依赖于坐标系的导数算符，而且只在该坐标系的坐标域上有定义，称为该坐标系的**普通导数算符**。上式表明 $\partial_\mu T^\nu{}_\sigma$ 是张量场 $\partial_a T^b{}_c$ 在该坐标系的分量，所以 ∂_a 的定义亦可表为：张量场 $T^{b_1\cdots b_k}{}_{c_1\cdots c_l}$ 的普通导数 $\partial_a T^{b_1\cdots b_k}{}_{c_1\cdots c_l}$ 的坐标分量等于该张量场的坐标分量对坐标的偏导数 $\partial(T^{\nu_1\cdots\nu_k}{}_{\sigma_1\cdots\sigma_l})/\partial x^\mu$。由此易见：

（1） 任一坐标系的 ∂_a 作用于该系的任一坐标基矢和任一对偶坐标基矢结果为零[2]，即

$$\partial_a(\partial/\partial x^\nu)^b = 0,\ \partial_a(\mathrm{d}x^\nu)_b = 0$$

（2） ∂_a 满足比导数算符的无挠性条件强得多的条件，即

$$\partial_a\partial_b T^{\cdots}{}_{\cdots} = \partial_b\partial_a T^{\cdots}{}_{\cdots},\ \text{或}\partial_{[a}\partial_{b]}T^{\cdots}{}_{\cdots} = 0$$

其中 $T^{\cdots}{}_{\cdots}$ 是任意型张量场。[3]

∂_a 虽可看作 ∇_a 的特例，但其定义依赖于坐标系。我们把与坐标系（或其他人为因素）无关的那些 ∇_a 称为**协变导数算符**，∂_a 不在此列。

定义 3.2. 设 ∂_a 是 (M,∇_a) 上任给的坐标系的普通导数算符，则体现 ∇_a 与 ∂_a 的差别的张量场 $C^c{}_{ab}$（把 ∂_a 看作 $\tilde{\nabla}_a$）称为 ∇_a 在该坐标系的**克氏符**（*Christoffel symbol*），记作 $\Gamma^c{}_{ab}$。

注 3.2. 一般书强调克氏符不是张量，本书却说它是张量。这没有实质性矛盾，只是克氏符的定义有微妙的不同。不用抽象指标的书把克氏符定义为与坐标系有关的一堆数，在坐标变换下不服从张量变换律，故不构成张量。我们一开始就用映射语言把克氏符 $\Gamma^c{}_{ab}$ 定义为张量，但因它与 ∂_a 相应，而 ∂_a 依赖于坐标系，故克氏符是依赖于坐标系的张量（坐标系改变时张量本身要变）。设 ∇_a 是 M 上指定的导数算符，$\{x^\mu\}$ 和 $\{x'^\mu\}$ 是 M 上的两个坐标系，坐标域之交为 U，∇_a 在两系中的克氏符分别为 $\Gamma^c{}_{ab}$ 和 $\bar{\Gamma}^c{}_{ab}$，它们（在 U 中）既可用 $\{x^\mu\}$ 系也可用 $\{x'^\mu\}$ 系求分量，设 $\Gamma^c{}_{ab}$ 在 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 系的分量为 $\{\Gamma^\sigma{}_{\mu\nu}\}$ 和 $\{\Gamma'^\sigma{}_{\mu\nu}\}$（这两堆数当然满足张量变换律），$\bar{\Gamma}^c{}_{ab}$ 在 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 系的分量为 $\{\bar{\Gamma}^\sigma{}_{\mu\nu}\}$ 和 $\{\bar{\Gamma}'^\sigma{}_{\mu\nu}\}$（也满足张量变换律），但 $\{\Gamma^\sigma{}_{\mu\nu}\}$ 和 $\{\bar{\Gamma}'^\sigma{}_{\mu\nu}\}$ 不满足张量变换律。而一般书恰恰是把 $\{\Gamma^\sigma{}_{\mu\nu}\}$ 和 $\{\bar{\Gamma}'^\sigma{}_{\mu\nu}\}$ 分别定义为 $\{x^\mu\}$ 系和 $\{x'^\mu\}$ 系的克氏符，自然不构成张量。一般书强调“克氏符不是张量”是对的，本书则应强调“克氏符是坐标系依赖的张量”。只要采用抽象指标并按照上述思路讨论（包括使用“张量面面观”的优雅论辩），自然要承认反映两个导数算符 ∇_a 与 $\tilde{\nabla}_a$ 之差的 $C^c{}_{ab}$ 是张量。在 M 指定了一个导数算符 ∇_a 的前提下，选定坐标系就有导数算符 ∂_a，把 ∂_a 看作 $\tilde{\nabla}_a$，则反映 ∇_a 与 ∂_a 之差的 $C^c{}_{ab}$（此时记作 $\Gamma^c{}_{ab}$）当然是张量。然而与此同时应该强调 $\Gamma^c{}_{ab}$ 是坐标系依赖的张量（有多少个坐标系就有多少个不同的 ∂_a，因而有多少个不同的 $\Gamma^c{}_{ab}$。）。这种强调的实质就是在强调一般书所强调的“克氏符不是张量”。两种强调只是同一问题的两种提法。重要的不在于提法而在于充分注意问题的实质，即切记两组分量 $\{\Gamma^\sigma{}_{\mu\nu}\}$ 和 $\{\bar{\Gamma}'^\sigma{}_{\mu\nu}\}$ 之间不遵守张量变换律。

类似地，设 v^b 是矢量场，则 $\partial_a v^b$ 也是坐标系依赖的张量场，把 $\partial_a v^b$ 在 ∂_a 所在坐标系展开：

$$\partial_a v^b = (\mathrm{d}x^\mu)_a(\partial/\partial x^\nu)^b v^\nu{}_{,\mu}$$

[2]例如，第 ν 坐标基矢 $(\partial/\partial x^\nu)$ 的第 σ 坐标分量为 $(\partial/\partial x^\nu)^\sigma \equiv (\partial/\partial x^\nu)^b(\mathrm{d}x^\sigma)_b = \delta^\sigma{}_\nu$，故其偏导数为零。

[3]任意型张量场的坐标分量都是标量场（即 n 元函数），故求偏导数的结果与求导顺序无关。

其中 $v^\nu{}_{,\mu} \equiv \partial_\mu v^\nu \equiv \partial v^\nu/\partial x^\mu$（逗号代表求偏导数），一般书强调 $v^\nu{}_{,\mu}$ 不构成张量，我们说 $\partial_a v^b$ 是坐标系依赖的张量，也是同一问题的两种提法。说得更具体些，设 ∂_a 和 ∂'_a 分别是坐标系 $\{x^\mu\}$ 和 $\{x'_\mu\}$ 的普通导数算符，则一般有 $\partial_a v^b \neq \partial'_a v^b$（所以说 $\partial_a v^b$ 是坐标系依赖的张量）。把 $\partial_a v^b$ 和 $\partial'_a v^b$ 在各自坐标基底展开：

$$\partial_a v^b = (dx^\mu)_a(\partial/\partial x^\nu)^b v^\nu{}_{,\mu} \; \partial'_a v^b = (dx'^\mu)_a(\partial/\partial x'^\nu)^b v'^\nu{}_{,\mu}$$

其中 $v'^\nu{}_{,\mu} \equiv \partial v'^\nu/\partial x'^\mu$，则 $\partial_a v^b \neq \partial'_a v^b$ 导致 $v^\nu{}_{,\mu}$ 与 $v'^\nu{}_{,\mu}$ 之间一般不满足张量分量变换律。所以一般书说 $v^\nu{}_{,\mu}$ 不是张量。至于 $\nabla_a v^b$，则是与坐标系无关的张量，它在坐标系中的分量通常记为 $v^\nu{}_{;\mu}$，即 $\nabla_a v^b = v^\nu{}_{;\mu}(dx^\mu)_a(\partial/\partial x^\nu)^b$。由于 $\nabla_a v^b$ 与坐标系无关，$v^\nu{}_{;\mu}$ 满足张量变换律，故一般书说它是张量（其实是张量的分量），并称之为 v^ν 的**协变导数**（其实是协变导数 $\nabla_a v^b$ 的坐标分量）。类似地，$\nabla_a \omega_b$ 的坐标分量 $\omega_{\nu;\mu}$ 也被称为 ω_ν 的协变导数。

定理 3.6. $v^\nu{}_{;\mu} = v^\nu{}_{,\mu} + \Gamma^\nu{}_{\mu\sigma} v^\sigma$，$\omega_{\nu;\mu} = \omega_{\nu,\mu} - \Gamma^\sigma{}_{\mu\nu}\omega_\sigma$

其中 v^ν 及 ω_ν 为任意矢量场和对偶矢量场在任一坐标基底的分量，$\Gamma^\nu{}_{\mu\sigma}$ 是该系的克氏符 $\Gamma^b{}_{ac}$ 在该基底的分量（一般书的说法是 “$\Gamma^\nu{}_{\mu\sigma}$ 是该系的克氏符”，本书后面为简单起见也常用这一说法。）。

定理 3.7. 导数算符定义中与缩并可交换顺序的条件等价于

$$\nabla_a \delta^b{}_c = 0$$

其中 $\delta^b{}_c$ 看作 $(1,1)$ 型张量场，其在每点 $p \in M$ 的定义为 $\delta^b{}_c v^c = v^b$，$\forall v^c \in V_p$。

流形 M 上矢量场对易子 $[u,v]^a$ 的定义无需 M 有附加结构，但该式的不方便之处在于它不能脱离被作用对象（标量场 f）。现在，在有了导数算符的概念之后，就可借助于随便一个导数算符 ∇_a 写出矢量场对易子 $[u,v]^a$ 的显表达式，见如下定理：

定理 3.8. $[u,v]^a = u^b\nabla_b v^a - v^b\nabla_b u^a$

其中 ∇_b 是任一无挠导数算符。

证明. 对所有 $f \in \mathscr{F}$ 有

$$\begin{aligned}
[u,v](f) &= u(v(f)) - v(u(f)) = u^b\nabla_b(v^a\nabla_a f) - v^b\nabla_b(u^a\nabla_a f) \\
&= u^b(\nabla_b v^a)\nabla_a f + u^b v^a \nabla_b\nabla_a f - v^b(\nabla_b u^a)\nabla_a f - v^b u^a \nabla_b \nabla_a f \\
&= (u^b\nabla_b v^a - v^b\nabla_b u^a)\nabla_a f
\end{aligned}$$

而 $[u,v](f) = [u,v]^a\nabla_a f$，证毕。 □

注 3.3. 取任一坐标系 $\{x^\mu\}$ 的普通导数算符 ∂_b 作为上式的 ∇_b，便有

$$[u,v]^\mu = (dx^\mu)_a[u,v]^a = u^\nu\partial_\nu v^\mu - v^\nu\partial_\nu u^\mu$$

3.2 矢量场沿曲线的导数和平移

3.2.1 矢量场沿曲线的平移

在流形 M 上选定一个导数算符 ∇_a 后，就有矢量沿曲线平移的概念。

定义 3.3. 设 v^a 是沿曲线 $C(t)$ 的矢量场。v^a 称为沿 $C(t)$ **平移**的，若 $T^b\nabla_b v^a = 0$，其中 $T^a \equiv (\partial/\partial t)^a$ 是曲线的切矢。

正如 $T^a\nabla_a f = T(f)$ 可解释为函数 f 沿 T^a（即沿 $C(t)$）的导数那样，$T^b\nabla_a v^b$ 可解释为矢量场 v^a 沿 T^b 的导数。于是上述定义也可解释为：v^a 沿 $C(t)$ 平移的充要条件是它沿 T^b 的导数为零。

定理 3.9. 设曲线 $C(t)$ 位于坐标系 $\{x^\mu\}$ 的坐标域内，曲线的参数式为 $x^\mu(t)$。令 $T^a \equiv (\partial/\partial t)^a$，则沿 $C(t)$ 的矢量场 v^a 满足

$$T^b\nabla_b v^a = (\partial/\partial x^\mu)^a(dv^\mu/dt + \Gamma^\mu{}_{\nu\sigma}T^\nu v^\sigma)$$

证明. 以 ∂_a 代表坐标系 $\{x^\mu\}$ 的普通导数算符，则

$$\begin{aligned}T^b\nabla_b v^a &= T^b(\partial_b v^a + \Gamma^a{}_{bc}v^c) = T^b[(dx^\nu)_b(\partial/\partial x^\mu)^a\partial_\nu v^\mu + \Gamma^a{}_{bc}v^c] \\ &= T^\nu(\partial/\partial x^\mu)^a(\partial v^\mu/\partial x^\nu) + \Gamma^a{}_{bc}T^b v^c = (\partial/\partial x^\mu)^a[T^\nu(\partial v^\mu/\partial x^\nu) + \Gamma^\mu{}_{\nu\sigma}T^\nu v^\sigma]\end{aligned}$$

其中 T^ν 是曲线切矢 T^b 的坐标分量。又 $T^\nu = dx^\nu(t)/dt$，故

$$T^\nu(\partial v^\mu/\partial x^\nu) = [dx^\nu(t)/dt][\partial v^\mu(t(x))/\partial x^\nu] = dv^\nu(t)/dt$$

证毕。 □

定理 3.10. 曲线上一点 $C(t_0)$ 及该点的一个矢量决定唯一的沿曲线平移的矢量场。

证明. 若存在把整条曲线含于其坐标域内的坐标系，则由前述定理可知平移定义 $T^b\nabla_b v^a = 0$ 等价于

$$dv^\nu/dt + \Gamma^\mu{}_{\nu\sigma}T^\nu v^\sigma = 0,\ \mu = 1, \cdots, n$$

这是关于 n 个待求函数 $v^\mu(t)$ 的 n 个一阶常微分方程（注意 $\Gamma^\mu{}_{\nu\sigma}$ 及 T^ν 均为 t 的已知函数），而给定 $C(t_0)$ 点的一个矢量就是给定初始条件 $v^\nu(t_0)$，故有唯一解 $v^\nu(t)$。使用“接力法”不难把上述证明推广至曲线不能被一个坐标域覆盖的情况。 □

设 $p, q \in M$，则 V_p 和 V_q 是两个矢量空间，两者的元素无法比较。但若有一曲线 $C(t)$ 联接 p, q，就可用下法定义一个由 V_p 到 V_q 的映射：对所有的 $v^a \in V_p$，由上述定理知在 $C(t)$ 上有唯一的平移矢量场（其在 p 点的值为 v^a），它在 q 点的值就定义为 v^a 的像。注意，这是一个曲线依赖的映射，另选一条联接 p, q 的曲线，v^a 的像可能不同。然而，无论

如何，∇_a 的存在毕竟使原来毫无联系的 V_p 与 V_q 发生了某种联系（虽然曲线依赖），因此也把 ∇_a 叫做**联络**。

为什么说 ∇_a 是 3 维欧氏空间中熟知的 $\vec{\nabla}$ 在一般流形上的某种推广？为什么把 $T^b\nabla_b v^a$ 解释为 v^a 沿 T^b 的导数？为什么把满足 $T^b\nabla_b v^a = 0$ 的 v^a 称为沿曲线平移的矢量场？下面我们就来回答这些问题。

3.2.2 与度规相适配的导数算符

从本章开始至今未涉及度规，只假定 M 上选了一个联络 ∇_a。如果 M 上还指定了度规 g_{ab}，矢量之间就可以谈及内积。为使平移概念与欧氏空间中熟知的平移一致，应补充以下要求：设 u^a，v^a 为沿 $C(t)$ 平移的矢量场，则 $u^a v_a$（$\equiv g_{ab}u^a v^b$）在 $C(t)$ 上是常数（两个矢量平移时“内积”不变）。设 T^a 为曲线 $C(t)$ 的切矢，则这一要求等价于

$$0 = T^c\nabla_c(g_{ab}u^a v^b) = g_{ab}u^a T^c\nabla_c v^b + g_{ab}v^b T^c\nabla_c u^a + u^a v^b T^c\nabla_c g_{ab} = u^a v^b T^c\nabla_c g_{ab}$$

上式对所有曲线以及沿它平移的任意两个矢量场 u^a，v^a 成立的充要条件为

$$\nabla_c g_{ab} = 0$$

没有度规时，∇_c 的选择非常任意。指定度规后，选 ∇_c 时就宜满足附加要求 $\nabla_c g_{ab} = 0$。下面证明这一要求决定了唯一的 ∇_a。

定理 3.11. *流形 M 上选定度规场 g_{ab} 后，存在唯一的 ∇_a 使 $\nabla_a g_{bc} = 0$。*

证明. 设 $\tilde{\nabla}_a$ 为任一导数算符，欲求适当的 $C^c{}_{ab}$ 使它与 $\tilde{\nabla}_a$ 决定的 ∇_a 满足 $\nabla_a g_{bc} = 0$。已知

$$\nabla_a g_{bc} = \tilde{\nabla}_a g_{bc} - C^d{}_{ab}g_{dc} - C^d{}_{ac}g_{bd} = \tilde{\nabla}_a g_{bc} - C_{cab} - C_{bac}$$

故由 $\nabla_a g_{bc} = 0$ 得

$$C_{cab} + C_{bac} = \tilde{\nabla}_a g_{bc}$$

同理有

$$C_{cba} + C_{abc} = \tilde{\nabla}_b g_{ac}$$
$$C_{bca} + C_{acb} = \tilde{\nabla}_c g_{ab}$$

前两式相加减第三式并利用 $C_{cab} = C_{cba}$，得

$$2C_{cab} = \tilde{\nabla}_a g_{bc} + \tilde{\nabla}_b g_{ac} - \tilde{\nabla}_c g_{ab}$$

或

$$C^c{}_{ab} = \frac{1}{2}g^{cd}(\tilde{\nabla}_a g_{bd} + \tilde{\nabla}_b g_{ad} - \tilde{\nabla}_d g_{ab})$$

这 $C^c{}_{ab}$ 与 $\tilde{\nabla}_a$ 结合而得的 ∇_a 便是方程 $\nabla_a g_{bc} = 0$ 的解。这必定是唯一解，因若 ∇'_a 也满足 $\nabla'_a g_{bc} = 0$，把 ∇'_a 作为上式中的 $\tilde{\nabla}_a$ 便知反映 ∇_a 与 ∇'_a 差别的 $C^c{}_{ab}$ 为零。 □

满足 $\nabla_a g_{bc} = 0$ 的 ∇_a 称为**与 g_{ab} 适配（或相容）的导数算符**。今后如无声明，在有 g_{ab} 时谈到 ∇_a 都是指与 g_{ab} 适配的导数算符。可以证明，$\nabla_a g_{bc} = 0$ 保证 $\nabla_a g^{bc} = 0$（反之亦然）。这为计算带来很大方便。

例 3.1. 欧氏空间存在无数满足定义的导数算符，但与欧式度规 δ_{ab} 相适配的导数算符只有一个，这就是笛卡尔坐标系 $\{x^\mu\}$ 的普通导数算符 ∂_a（所有笛卡尔系的 ∂_a 都相同），因为由 δ_{ab} 的定义式得 $\partial_c \delta_{ab} = (dx^\sigma)_c (dx^\mu)_a (dx^\nu)_b \partial_\sigma \delta_{\mu\nu} = 0$。对 3 维欧氏空间，笛卡尔坐标系的 ∂_a 就是普通矢量场论中熟知的 $\vec{\nabla}$。

设 ∇_a 与 g_{bc} 相适配，取 $\tilde{\nabla}_a$ 为任一坐标系的 ∂_a，则前述定理中的 $C^c{}_{ab}$ 便是 ∇_a 在该坐标系的克氏符 $\Gamma^c{}_{ab}$，由该定理不难推得 $\Gamma^c{}_{ab}$ 在该系的分量 $\Gamma^\sigma{}_{\mu\nu}$ 的如下表达式:

$$\Gamma^\sigma{}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}(g_{\rho\mu,\nu} + g_{\nu\rho,\mu} - g_{\mu\nu,\rho})$$

推导如下：把 ∂_a 看作前述定理中的 $\tilde{\nabla}_a$，则式中的 $C^c{}_{ab}$ 就是 $\Gamma^c{}_{ab}$，故

$$\Gamma^c{}_{ab} = \frac{1}{2} g^{cd}(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab})$$

$$\begin{aligned}
\Gamma^\sigma{}_{\mu\nu} &= \Gamma^c{}_{ab}(dx^\sigma)_c(\partial/\partial x^\mu)^a(\partial/\partial x^\nu)^b \\
&= \frac{1}{2}(dx^\sigma)_c(\partial/\partial x^\mu)^a(\partial/\partial x^\nu)^b g^{cd}(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \\
&= \frac{1}{2}(\partial/\partial x^\mu)^a(\partial/\partial x^\nu)^b g^{\sigma\rho}(\partial_a g_{b\rho} + \partial_b g_{a\rho} - \partial_\rho g_{ab}) \\
&= \frac{1}{2} g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) \\
&= \frac{1}{2} g^{\sigma\rho}(g_{\nu\rho,\mu} + g_{\mu\rho,\nu} - g_{\mu\nu,\rho})
\end{aligned}$$

利用对称性 $g_{ab} = g_{ba}$ 不难看出这就是要推的结果。此式与 $v^\nu{}_{;\mu} = v^\nu{}_{,\mu} + \Gamma^\nu{}_{\mu\sigma} v^\sigma$，$\omega_{\nu;\mu} = \omega_{\nu,\mu} - \Gamma^\sigma{}_{\mu\nu}\omega_\sigma$ 结合可方便地求得矢量场 v^a 和对偶矢量场 ω_a 的协变导数 $\nabla_a v^b$ 和 $\nabla_a \omega_b$ 的坐标分量 $v^\nu{}_{;\mu}$ 和 $\omega_{\nu;\mu}$。

注 3.4. $\Gamma^\sigma{}_{\mu\nu}$ 既依赖于 M 上选定的 ∇_a 又依赖于坐标系。若 M 有度规 g_{ab}，则不加声明时 ∇_a 就是指同 g_{ab} 适配的导数算符，而“某系的克氏符”就是指这个 ∇_a 在该系的克氏符。例如，谈到 3 维欧氏空间中某坐标系的克氏符时，指的就是同欧氏度规相适配的 ∇_a（即笛卡尔系的 ∂_a）在该系的克氏符。一个笛卡尔系的 ∂_a 在任一笛卡尔系的克氏符显然为零。

设 $\vec{T}$ 是 3 维欧氏空间的矢量场，则 $\vec{T}\cdot\vec{\nabla} f$ 是梯度 $\vec{\nabla} f$ 在 $\vec{T}$ 方向的分量，即 f 沿 $\vec{T}$ 方向的导数。另一方面，由导数算符的条件（d）有 $T^a\partial_a f = T(f)$，而右边正是 f 沿 T^a 的导数。可见 $T^a\partial_a f = \vec{T}\cdot\vec{\nabla} f$。进一步的问题是：$T^b\partial_b v^a$ 代表什么？答案自然是 v^a 沿 T^a 的导数。详见下节。

3.2.3 矢量场沿曲线的导数与沿曲线的平移的关系

先讨论最简单的特例，即欧氏空间。欧氏空间有一类特殊坐标系（笛卡尔系），用它可定义矢量的绝对（非曲线依赖）平移。

定义 3.4. 欧氏空间中 p 点的矢量 $\vec{\tilde{v}}$ 称为 q 点的矢量 $\vec{v}$ **平移**至 p 点的结果，若两者在同一笛卡尔系的分量相等（注：对某一笛卡尔系为平移则对任意笛卡尔系也为平移）。

定义 3.5. 欧氏空间中曲线 $C(t)$ 上的矢量场 $\vec{v}$ 沿线的导数 $d\vec{v}/dt$ 定义为

$$\left.\frac{d\vec{v}}{dt}\right|_p \coloneqq \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\vec{\tilde{v}}|_p - \vec{v}|_p)\ \forall p\in C(t)$$

其中 $\vec{\tilde{v}}|_p$ 是把 $\vec{v}|_q$（q 为 p 在线上的邻点）平移至 p 点的结果，$\Delta t \equiv t(q)-t(p)$。

现在证明 $d\vec{v}/dt$ 在抽象指标记号中就是 $T^b\partial_b v^a$（其中 T^b 是 $C(t)$ 的切矢，∂_b 是笛卡尔系的普通导数算符。），为此只须证明两者在笛卡尔系 $\{x^i\}$ 的分量相等：

$$T^b\partial_b v^a \text{的第 } i \text{ 分量} = T^b\partial_b v^a (dx^i)_a = T^b\partial_b v^i = T(v^i) = dv^i/dt$$

另一方面，由上述定义可知

$$\left.\frac{d\vec{v}}{dt}\right|_p \text{的第 } i \text{ 分量} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\tilde{v}^i|_p - v^i|_p) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\tilde{v}^i|_q - v^i|_p) = \left.\frac{dv^i}{dt}\right|_p$$

（最后一步无非是函数 $v^i(t)$ 的导数的定义）对比两式可见

$$d\vec{v}/dt = T^b\nabla_b v^a$$

推广到带任意 ∇_a 的任意流形 M 的任意曲线 $C(t)$，便自然地把 $T^b\nabla_b v^a$ 称为 v^a 沿 T^b（或沿 $C(t)$）的导数。有时也用记号 Dv^a/dt 代表这一导数，即

$$Dv^a/dt \equiv T^b\nabla_b v^a$$

然而上述绝对平移的定义不能推广到带有任意联络 ∇_a 的任意流形 (M,∇_a)。后面我们将介绍流形的内禀曲率概念（并将看到欧氏和闵氏空间的内禀曲率为零），还将指出只有内禀曲率为零的空间才有绝对的（即与曲线无关的）平移概念。不过，鉴于欧氏空间中 $d\vec{v}/dt=0$ 与 $\vec{v}$ 沿 $C(t)$ 平移等价（见上述定义），可以自然地把 $T^b\nabla_b v^a=0$ 作为 (M,∇_a) 中曲线 $C(t)$ 上的矢量场 v^a 的平移定义，这就解释了本节一开始给出的平移定义提出的动机（但应特别注意这样定义的平移一般是曲线依赖的）。在这种讲法中，我们是先定义 v^a 沿曲线的导数 $T^b\nabla_b v^a$ 再用它定义 v^a 沿曲线的平移（与欧氏空间的顺序相反）。由于 v^a 沿曲线的导数 $T^b\nabla_b v^a$ 毕竟比较抽象，在有了曲线依赖的平移概念后，借用平移这一术语反过来对 $T^b\nabla_b v^a$ 作一解释是有益的。这一解释的实质就是上述定义所表达的含义，见如下定理。

定理 3.12. 设 v^a 是 (M,∇_a) 的曲线 $C(t)$ 上的矢量场，T^b 是 $C(t)$ 的切矢，p,q 是 $C(t)$ 上的邻点，则

$$T^b\nabla_b v^a|_p = \lim_{\Delta t\to 0}\frac{1}{\Delta t}(\tilde{v}^a|_p - v^a|_p)$$

其中 $\Delta t \equiv t(q)-t(p)$，$\tilde{v}^a|_p$ 是 $v^a|_q$ 沿 $C(t)$ 平移至 p 点的结果。

3.3 测地线

定义 3.6. (M,∇_a) 上的曲线 $\gamma(t)$ 叫**测地线**，若其切矢 T^a 满足 $T^b\nabla_b T^a = 0$。

注 3.5. ① 可见测地线的充要条件是其切矢沿曲线平移。② $T^b\nabla_b T^a = 0$ 称为**测地线方程**。③ 设流形 M 上有度规场 g_{ab}，则 (M,g_{ab}) 的测地线是指 (M,∇_a) 的测地线，其中 ∇_a 与 g_{ab} 适配。

设测地线 $\gamma(t)$ 位于某坐标系的坐标域内，则以 T^a 代替上节第一个定理中的 v^a 便得

$$\frac{dT^\mu}{dt} + \Gamma^\mu{}_{\nu\sigma}T^\nu T^\sigma = 0,\ \mu = 1,\cdots,n$$

设 $x^\nu = x^\nu(t)$ 是测地线 $\gamma(t)$ 的参数式，则 $T^\mu = dx^\mu/dt$，故上式可改写为

$$\frac{d^2x^\mu}{dt^2} + \Gamma^\mu{}_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt} = 0,\ \mu = 1,\cdots,n$$

这就是测地线方程的坐标分量表达式。

例 3.2. 欧（闵）氏度规在笛卡尔（洛伦兹）系的克氏符为零，测地线方程的通解为 $x^\mu(t) = a^\mu t + b^\mu$（其中 a^μ, b^μ 是常数）。如果把欧（闵）氏空间中在笛卡尔（洛伦兹）系的参数式为 $x^\mu(t) = a^\mu t + b^\mu$ 的曲线称为直线（段），欧（闵）氏空间的测地线便同义于直线（段）。可见测地线可看作欧氏空间直线概念向广义黎曼空间的推广。

例 3.3. 设 S^2 是 3 维欧氏空间的 2 维球面，以球心为原点建球坐标系，则 3 维欧氏线元为 $ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$。若元线段躺在 S^2 上，则 $r = R$（球半径）导致 $dr = 0$，故球面上的 “诱导线元”（称为标准球面线元）为 $d\hat{s}^2 = R^2(d\theta^2 + \sin^2\theta d\varphi^2)$，就是说，3 维欧氏度规 δ_{ab} 在球面 S^2 上诱导出 2 维度规 g_{ab}，其在坐标基底 $\{(\partial/\partial\theta)^a, (\partial/\partial\varphi)^a\}$ 的分量为 $g_{\theta\theta} = R^2, g_{\varphi\varphi} = R^2\sin^2\theta, g_{\theta\varphi} = g_{\varphi\theta} = 0$。从测地线方程的坐标分量表达式出发可以证明，以这个度规衡量，球面上的曲线为测地线当且仅当它是大圆弧（并配以适当的参数化）。

定理 3.13. 设 $\gamma(t)$ 为测地线，则其重参数化 $\gamma'(t')[=\gamma(t)]$ 的切矢 T'^a 满足

$$T'^b\nabla_b T'^a = \alpha T'^a（\alpha \text{ 为 } \gamma(t) \text{ 上的某个函数}）$$

证明.

$$T^a = \left(\frac{\partial}{\partial t}\right)^a = \left(\frac{\partial}{\partial t'}\right)^a\frac{dt'}{dt} = \frac{dt'}{dt}T'^a$$

$$0 = T^b\nabla_b T^a = \frac{dt'}{dt}T'^b\nabla_b\left(\frac{dt'}{dt}T'^a\right) = \left(\frac{dt'}{dt}\right)^2 T'^b\nabla_b T'^a + T'^a\frac{dt'}{dt}T'^b\nabla_b\left(\frac{dt'}{dt}\right)$$

右边第二项 $= T'^a\frac{dt'}{dt}\frac{d}{dt'}\left(\frac{dt'}{dt}\right) = T'^a\frac{d^2t'}{dt^2}$，故 $T'^b\nabla_b T'^a = -\left(\frac{dt}{dt'}\right)^2\frac{d^2t'}{dt^2}T'^a$。令 $\alpha \equiv -\left(\frac{dt}{dt'}\right)^2\frac{d^2t'}{dt^2}$，得证。 □

定理 3.14. 设曲线 $\gamma(t)$ 的切矢 T^a 满足 $T^b\nabla_b T^a = \alpha T^a$（$\alpha$ 为 $\gamma(t)$ 上的函数），则存在 $t' = t'(t)$ 使得 $\gamma'(t')[= \gamma(t)]$ 为测地线。

定义 3.7. 能使曲线成为测地线的参数叫该曲线的**仿射参数**。

注 3.6. 有时也把满足 $T^b\nabla_b T^a = \alpha T^a$ 的曲线叫测地线。不过，为避免混淆，最好称之为“非仿射参数化的测地线”。

定理 3.15. 若 t 是某测地线的仿射参数，则该线的任一参数 t' 是仿射参数的充要条件为 $t' = at + b$（其中 a, b 为常数且 $a \neq 0$）。

定理 3.16. 流形 M 的一点 p 及 p 点的一个矢量 v^a 决定唯一的测地线 $\gamma(t)$，满足

（1） $\gamma(0) = p$；

（2） $\gamma(t)$ 在 p 点的切矢等于 v^a。

证明. 任取一坐标系 $\{x^\mu\}$ 使坐标域含 p，则测地线方程的坐标分量表达式可看作关于 n 个待求函数 $x^\mu(t)$ 的 n 个二阶常微分方程，则给定 $p \in M$ 及 $v^a \in V_p$ 就是给定初始条件 $x^\mu(0) = x^\mu|_p$ 及 $(dx^\mu/dt)|_0 = v^a$，故有唯一解。 □

定理 3.17. 测地线的线长参数必为仿射参数。

众所周知，欧氏空间中两点之间直线（段）最短。现在讨论这一结论在多大程度上适用于带洛伦兹度规的流形（时空）。

定理 3.18. 设 g_{ab} 是流形 M 上的洛伦兹度规场，$p, q \in M$，则 p, q 间的光滑类空（类时）曲线为测地线当且仅当其线长取极值。

注 3.7. ① 本定理也适用于 g_{ab} 为正定度规的情况（这时曲线的定语“类空（类时）”略去）。② 线长取极值的含义如下：设 C 是 p, q 间的类空（类时）曲线，则可对它做微小修改而得到 p, q 间的与 C“无限邻近”的许多类空（类时）曲线。上述定理断定 C 为测地线的充要条件是其长度在所有这些类空（类时）曲线的长度中取极值。一元函数 $f(x)$ 取极值的条件是一阶导数为零，然而上述定理涉及的长度 l（看作“因变量”）相应的“自变量”不是实数而是曲线，关心的是从一条曲线变到另一曲线时 l 的变化，因此 l 不是一元函数而是泛函。根据变分理论，l 取极值的充要条件是 l 的变分 δl 为零。

一元函数 $f(x)$ 的极值可为极小（其充分条件为 $f'(x)=0, f''(x)>0$），也可为极大（其充分条件为 $f'(x)=0, f''(x)<0$），还可为既非极小又非极大（其必要条件为 $f'(x)=0, f''(x)=0$）。与此类似，线长的极值也有上述三种可能，讨论如下。

先讨论 g_{ab} 为正定度规的情况。给定 p, q 之间的任一曲线，总可略加修改而得长度更大的曲线，故 p, q 间的曲线长度无极大可言。设 C 是 p, q 间长度极小的一条曲线，由前述定理知它必为测地线。然而 p, q 间的测地线却未必长度极小，因为极值可以既非极小又非极大。可以证明，测地线长度取极小值的充要条件是线上不存在共轭点对。欧氏空间当然没有共轭点对，因此两点之间直线（段）最短。

再讨论 g_{ab} 为洛伦兹度规的情况。先看闵氏时空这一最简单特例。前已说过闵氏时空中直线与测地线同义。设 p, q 间有类时测地线 γ 相连，它是否为 p, q 间的最短线？否。由于类光曲线长度为零，任一类时曲线 C 都不是最短线，因为总可对它略加修改而成为足够接近类光的类时曲线 C'，其长度小于 C。事实上，类时测地线 γ 非但不是最短线，而且是 p, q 间的最长线。以 2 维闵氏时空为例证明（容易推广至任意维闵氏时空）。由于 γ 的参数式 $x^\mu(t)$ 为线性函数，可通过洛伦兹坐标的平移和伪转动选择洛伦兹系 $\{x^0, x^1\}$ 使其 x^0 坐标线与 γ 重合。设 C 是 p, q 间的任一类时非测地线，用许多等 x^0 线把 γ 分成许多元段，由闵氏线元表达式可知 γ 和 C 的元段的线长依次为

$$dl_\gamma = \sqrt{-ds^2} = \sqrt{-[-(dx^0)^2 + 0]} = dx^0$$
$$dl_C = \sqrt{-[-(dx^0)^2 + (dx^1)^2]} < dx^0 = dl_\gamma$$

上述结果也适用于其他任一对元段，可见 $l_\gamma > l_C$，即闵氏时空的类时测地线是最长类时线。换句话说，闵氏时空中两点之间（类时）直线（段）最长。又因为最长线必为测地线，所以对闵氏时空而言，两点间的类时线为最长线的充要条件是它为测地线。再讨论一般时空，设 C 是 p, q 间长度极大的类时线，则由定理可知它是测地线。然而反过来却未必，因为定理只保证 p, q 间的类时测地线长取极值，不保证它是极大（当然，由于类光曲线长度为零，它肯定也不是极小。）。可以证明，任意时空中类时测地线长为极大的充要条件是线上不存在共轭点对。小结：对任意时空中有类时联系的两点：① 两点间的最长线是类时测地线；② 两点间的类时测地线未必是最长线（对闵氏时空一定是）；③ 两点间没有最短类时线。

3.4 黎曼曲率张量

3.4.1 黎曼曲率的定义和性质

导数算符 ∇_a 的无挠性保证 $(\nabla_a\nabla_b - \nabla_b\nabla_a)f = 0$，即 $\nabla_a\nabla_b f$ 是个对称的 $(0,2)$ 型张量。把算符 $(\nabla_a\nabla_b - \nabla_b\nabla_a)$ 称为导数算符 ∇_a 的**对易子**，则 ∇_a 的无挠性体现为其对易子对函数的作用结果为零。然而无挠导数算符的对易子对其他型号的张量场的作用结果却未必为零，黎曼曲率张量正是这种非对易性的表现。

定理 3.19. 设 $f \in \mathscr{F}, \omega_a \in \mathscr{F}(0,1)$，则

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)(f\omega_c) = f(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c$$

证明. 把 $\nabla_a\nabla_b(f\omega_c)$ 和 $\nabla_b\nabla_a(f\omega_c)$ 分别展为 4 项，两者相减，注意到无挠性条件，便可得证。 □

定理 3.20. 设 $\omega_c, \omega'_c \in \mathscr{F}(0,1)$ 且 $\omega'_c|_p = \omega_c|_p$，则

$$[(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega'_c]|_p = [(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c]|_p$$

证明. 令 $\Omega_c \equiv \omega'_c - \omega_c$，则要证的结果等价于 $[(\nabla_a\nabla_b - \nabla_b\nabla_a)\Omega_c]|_p = 0$，将其中的 Ω_c 展开为坐标分量表达式，利用前一定理的结果即可得证。 □

上述定理表明 $(\nabla_a\nabla_b - \nabla_b\nabla_a)$ 是把 p 点的对偶矢量 $\omega_c|_p$ 变为 p 点的 $(0,3)$ 型张量 $[(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c]|_p$ 的线性映射，做法是：把 $\omega_c|_p$ 任意延拓而得一个定义于 p 点某邻域的对偶矢量场 ω_c，求出 $(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c$，再取其在 p 点的值便得映射的像。该定理保证这个像与延拓方式无关。于是 $(\nabla_a\nabla_b - \nabla_b\nabla_a)$ 对应于 p 点的一个 $(1,3)$ 型张量，叫**黎曼曲率张量**，记作 $R_{abc}{}^d$。又因 p 点任意，故 $R_{abc}{}^d$ 是张量场。于是有

定义 3.8. 导数算符 ∇_a 的**黎曼曲率张量场**$R_{abc}{}^d$ 由下式定义

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c = R_{abc}{}^d\omega_d,\ \forall\omega_c \in \mathscr{F}(0,1)$$

黎曼张量场反映导数算符的非对易性，是描述 (M, ∇_a) 内禀性质的张量场。只要选定导数算符就可谈及其黎曼张量。当然，对广义黎曼空间 (M, g_{ab}) 也可谈及黎曼张量，亦称 g_{ab} 的黎曼张量，是指与 g_{ab} 适配的那个导数算符 ∇_a 的黎曼张量场。黎曼张量场为零的度规称为**平直度规**。下面证明欧氏和闵氏度规都是平直度规。

定理 3.21. 欧氏空间 $(\mathbb{R}^n, \delta_{ab})$ 和闵氏空间 $(\mathbb{R}^n, \eta_{ab})$ 的黎曼曲率张量场为零。

证明. 欧（闵）氏空间任一笛卡尔（洛伦兹）系的普通导数算符 ∂_a 是与 δ_{bc} 适配的那个特定的导数算符。而

$$(\partial_a\partial_b - \partial_b\partial_a)\omega_c = (dx^\mu)_a(dx^\nu)_b(dx^\sigma)_c(\partial_\mu\partial_\nu\omega_\sigma - \partial_\nu\partial_\mu\omega_\sigma) = 0,\ \forall\omega_c$$

故 ∂_a 的 $R_{abc}{}^d$ 为零。 □

因此欧氏空间和闵氏空间都称为**平直空间**。事实上，闵氏空间在许多方面类似于欧氏空间，故又称为**伪欧空间**。

黎曼曲率张量反映导数算符作用于对偶矢量场的非对易性。由此可推知导数算符作用于任意型张量场 $T^{c_1\cdots c_k}{}_{d_1\cdots d_l}$ 的非对易性，即推出用 $R_{abc}{}^d$ 表述 $(\nabla_a\nabla_b - \nabla_b\nabla_a)T^{c_1\cdots c_k}{}_{d_1\cdots d_l}$ 的公式。我们有如下定理：

定理 3.22. $(\nabla_a\nabla_b-\nabla_b\nabla_a)v^c=-R_{abd}{}^c v^d, \forall v^c\in\mathscr{F}(1,0)$

证明. 对所有的 $\omega_c\in\mathscr{F}(0,1)$，有 $v^c\omega_c\in\mathscr{F}$，故由无挠性条件得

$$\begin{aligned}0&=(\nabla_a\nabla_b-\nabla_b\nabla_a)(v^c\omega_c)=\nabla_a(v^c\nabla_b\omega_c+\omega_c\nabla_b v^c)-\nabla_b(v^c\nabla_a\omega_c+\omega_c\nabla_a v^c)\\&=v^c\nabla_a\nabla_b\omega_c+\omega_c\nabla_a\nabla_b v^c-v^c\nabla_b\nabla_a\omega_c-\omega_c\nabla_b\nabla_a v^c\end{aligned}$$

从而 $\omega_c(\nabla_a\nabla_b-\nabla_b\nabla_a)v^c=-v^c(\nabla_a\nabla_b-\nabla_b\nabla_a)\omega_c=-v^cR_{abc}{}^d\omega_d=-\omega_cR_{abd}{}^cv^d$ 于是定理得证。 □

定理 3.23. 对于所有的 $T^{c_1\cdots c_k}{}_{d_1\cdots d_l}\in\mathscr{F}(k,l)$ 有

$$(\nabla_a\nabla_b-\nabla_b\nabla_a)T^{c_1\cdots c_k}{}_{d_1\cdots d_l}=-\sum_{i=1}^{k}R_{abe}{}^{c_i}T^{c_1\cdots e\cdots c_k}{}_{d_1\cdots d_l}+\sum_{j=1}^{l}R_{abd_j}{}^{e}T^{c_1\cdots c_k}{}_{d_1\cdots e\cdots d_l}$$

定理 3.24. 黎曼曲率张量有以下性质

(*1*) $R_{abc}{}^d=-R_{bac}{}^d$

(*2*) $R_{[abc]}{}^d=0$（循环恒等式）

(*3*) $\nabla_{[a}R_{bc]d}{}^e=0$（***Bianchi*** 恒等式）

若 M 上有度规场 g_{ab} 且 $\nabla_a g_{bc}=0$，则可定义 $R_{abcd}\equiv g_{de}R_{abc}{}^e$，且 R_{abcd} 还满足

(*4*) $R_{abcd}=-R_{abdc}$

(*5*) $R_{abcd}=R_{cdab}$

证明.（1） 由定义显见。

（2） 因 $R_{[abc]}{}^d\omega_d=\nabla_{[a}\nabla_b\omega_{c]}-\nabla_{[b}\nabla_a\omega_{c]}=2\nabla_{[a}\nabla_b\omega_{c]}$，故只须证

$$\nabla_{[a}\nabla_b\omega_{c]}=0,\ \forall\omega_c\in\mathscr{F}(0,1)$$

而（取 $\tilde\nabla_a$ 为 ∂_a）

$$\begin{aligned}\nabla_a(\nabla_b\omega_c)&=\partial_a(\nabla_b\omega_c)-\Gamma^d{}_{ab}\nabla_d\omega_c-\Gamma^d{}_{ac}\nabla_b\omega_d\\&=\partial_a(\partial_b\omega_c-\Gamma^e{}_{bc}\omega_e)-\Gamma^d{}_{ab}\nabla_d\omega_c-\Gamma^d{}_{ac}\nabla_b\omega_d\\&=(\partial_a\partial_b\omega_c-\partial_a\Gamma^e{}_{bc}\omega_e-\Gamma^e{}_{bc}\partial_a\omega_e)-\Gamma^d{}_{ab}\nabla_d\omega_c-\Gamma^d{}_{ac}\nabla_b\omega_d\end{aligned}$$

故 $\nabla_{[a}\nabla_b\omega_{c]}=\partial_{[a}\partial_b\omega_{c]}-\partial_{[a}\Gamma^e{}_{bc]}\omega_e-\Gamma^e{}_{[bc}\partial_{a]}\omega_e-\Gamma^d{}_{[ab}\nabla_{|d|}\omega_{c]}-\Gamma^d{}_{[ac}\nabla_{b]}\omega_d$ 下标 $[ab|d|c]$ 中的 $|d|$ 表明 d 不参与反称化。注意到 $\nabla_a\nabla_b\omega_c=\nabla_b\nabla_a\omega_c$ 和 $\Gamma^e{}_{bc}=\Gamma^e{}_{cb}$，可知上式右边每项都为零。

（3） 只须证 $\nabla_{[a}R_{bc]d}{}^e\omega_e = 0,\ \forall\omega_e\in\mathscr{F}(0,1)$，而

$$\begin{aligned}
\nabla_a R_{bcd}{}^e\omega_e &= \nabla_a(R_{bcd}{}^e\omega_e) - R_{bcd}{}^e\nabla_a\omega_e\\
&= \nabla_a(\nabla_b\nabla_c\omega_d - \nabla_c\nabla_b\omega_d) - R_{bcd}{}^e\nabla_a\omega_e\\
&= \nabla_a\nabla_b\nabla_c\omega_d - \nabla_a\nabla_c\nabla_b\omega_d - R_{bcd}{}^e\nabla_a\omega_e\\
&= \nabla_a\nabla_b\nabla_c\omega_d - \nabla_b\nabla_a\nabla_c\omega_d - R_{bcd}{}^e\nabla_a\omega_e\\
&= (\nabla_a\nabla_b - \nabla_b\nabla_a)\nabla_c\omega_d - R_{bcd}{}^e\nabla_a\omega_e\\
&= R_{abd}{}^e\nabla_c\omega_e + R_{abc}{}^e\nabla_e\omega_d - R_{bcd}{}^e\nabla_a\omega_e
\end{aligned}$$

现在取反称化，并注意到循环恒等式，有

$$\begin{aligned}
\nabla_{[a}R_{bc]d}{}^e\omega_e &= R_{[ab|d|}{}^e\nabla_{c]}\omega_e + R_{[abc]}{}^e\nabla_e\omega_d - R_{[bc|d|}{}^e\nabla_{a]}\omega_e\\
&= R_{[ab|d|}{}^e\nabla_{c]}\omega_e - R_{[ab|d|}{}^e\nabla_{c]}\omega_e\\
&= 0
\end{aligned}$$

（4） 由 $\nabla_a g_{cd} = 0$ 得

$$0 = (\nabla_a\nabla_b - \nabla_b\nabla_a)g_{cd} = R_{abc}{}^e g_{ed} + R_{abd}{}^e g_{ce} = R_{abcd} + R_{abdc}$$

（5） 把循环恒等式展开，即 $R_{abc}{}^d + R_{cab}{}^d + R_{bca}{}^d - R_{acb}{}^d - R_{cba}{}^d - R_{bac}{}^d = 0$，而带负号的每一项都等于带正号的对应项，所以 $R_{abc}{}^d + R_{cab}{}^d + R_{bca}{}^d = 0$，因此 $R_{abc}{}^d = R_{cad}{}^b + R_{cba}{}^d$。为了让等式中出现 $R_{cda}{}^b$，我们分别交换 a,d 和 b,d，得到 $R_{dbc}{}^a = R_{cda}{}^b + R_{cbd}{}^a$ 和 $R_{adc}{}^b = R_{cab}{}^d + R_{cda}{}^b$。将此三式相加，得 $R_{abc}{}^d + R_{dbc}{}^a + R_{adc}{}^b = 2R_{cda}{}^b$。把等号左边的 c 都调到最后，即 $R_{bad}{}^c + R_{bda}{}^c + R_{dab}{}^c = R_{bad}{}^c - R_{abd}{}^c = 2R_{bad}{}^c = 2R_{cda}{}^b$，而这就等价于我们要证的。

□

注 3.8. 设 $\dim M = n$，则 $R_{abc}{}^d$ 的分量 $R_{\mu\nu\sigma}{}^\rho$ 共有 n^4 个。但由于满足前述定理中的代数等式，独立分量数仅为

$$N = n^2(n^2-1)/12$$

选定度规后，每个 $(0,2)$ 型张量 T_{ab} 对应于一个 $(1,1)$ 型张量 $T^a{}_b \equiv g^{ac}T_{cb}$，即矢量空间上的一个线性变换，其在任意基底的分量组成一个矩阵，不同基底对应的矩阵互为相似矩阵，故有相同的迹，其值为 $T^a{}_a = g^{ac}T_{ac}$，称为张量 $T^a{}_b$ 的迹，也称为 T_{ab} 的迹。类似地，给定 $(0,4)$ 型张量 R_{abcd}，原则上可通过缩并得到以下六个“迹”（每个“迹”是一个 $(0,2)$ 型张量）：$g^{ab}R_{abcd}$，$g^{ac}R_{abcd}$，$g^{ad}R_{abcd}$，$g^{bc}R_{abcd}$，$g^{bd}R_{abcd}$ 及 $g^{cd}R_{abcd}$。然而，由于黎曼张量 $R_{abc}{}^d$ 降指标后所得 R_{abcd} 的性质及 g^{ac} 的对称性，易见上述六个缩并中的第一、六个为零；第二、五个相等（理由：$g^{ac}R_{abcd} = g^{ac}R_{badc}$，与 $g^{bd}R_{abcd}$ 实质一样，

只因要照顾指标平衡而不写 $g^{ac}R_{abcd}=g^{bd}R_{abcd}$）；第三、四个相等并等于第二、五个之负，故六个缩并中只有一个独立，例如可取 $g^{bd}R_{abcd}=R_{abc}{}^{b}$，记作 R_{ac}，称为**里奇张量**。应该强调的是为定义里奇张量无需借用度规，因为 $R_{ac}\equiv R_{abc}{}^{b}$ 天生就有明确意义。R_{ac} 还可借度规求迹，即 $g^{ac}R_{ac}$，记作 R，称为**标量曲率**。易证 $R_{ac}=R_{ca}$。此外还应掌握 $R_{abc}{}^{d}$ 的无迹部分，叫外尔张量，定义如下：

定义 3.9. 对维数 $n\geqslant 3$ 的广义黎曼空间，**外尔张量**C_{abcd} 由下式定义：

$$C_{abcd}\coloneqq R_{abcd}-\frac{2}{n-2}(g_{a[c}R_{d]b}-g_{b[c}R_{d]a})+\frac{2}{(n-1)(n-2)}Rg_{a[c}g_{d]b}$$

定理 3.25. 外尔张量有以下性质：

(*1*) $C_{abcd}=-C_{bacd}=-C_{abdc}=C_{cdab}$，$C_{[abc]d}=0$

(*2*) C_{abcd} 的各种迹都为零，例如 $g^{ac}C_{abcd}=0$

注 3.9. 外尔张量的定义说明 R_{abcd} 是其无迹部分 C_{abcd} 与有迹部分

$$\frac{2}{n-2}(g_{a[c}R_{d]b}-g_{b[c}R_{d]a})-\frac{2}{(n-1)(n-2)}Rg_{a[c}g_{d]b}$$

之和。

定义 3.10. 广义黎曼空间的**爱因斯坦张量**G_{ab} 由下式定义

$$G_{ab}\coloneqq R_{ab}-\frac{1}{2}Rg_{ab}$$

定理 3.26. $\nabla^{a}G_{ab}=0$（其中 $\nabla^{a}G_{ab}\equiv g^{ac}\nabla_{c}G_{ab}$）。

证明. 由 Bianchi 恒等式有 $0=\nabla_{a}R_{bcd}{}^{e}+\nabla_{c}R_{abd}{}^{e}+\nabla_{b}R_{cad}{}^{e}$。指标 a 同 e 缩并得 $0=\nabla_{a}R_{bcd}{}^{a}+\nabla_{c}R_{abd}{}^{a}+\nabla_{b}R_{cad}{}^{a}=\nabla_{a}R_{bcd}{}^{a}-\nabla_{c}R_{bd}+\nabla_{b}R_{cd}$。以 g^{bd} 作用得

$$\begin{aligned}0&=g^{bd}\nabla_{a}R_{bcd}{}^{a}-g^{bd}\nabla_{c}R_{bd}+g^{bd}\nabla_{b}R_{cd}\\&=\nabla_{a}R_{c}{}^{a}-\nabla_{c}R+\nabla_{b}R_{c}{}^{b}=2\nabla_{a}R_{c}{}^{a}-\nabla_{c}R\end{aligned}$$

故 $\nabla^{a}G_{ab}=\nabla^{a}R_{ab}-\frac{1}{2}R\nabla^{a}g_{ab}=\nabla_{a}R_{b}{}^{a}-\frac{1}{2}R\nabla_{b}=0$，其中第二步用到 $R_{ab}=R_{ba}$ 以及 $\nabla^{a}G_{ab}$ 的定义，第三步用到上式。 □

爱因斯坦张量 G_{ab} 及其满足的上述定理对建立广义相对论的爱因斯坦方程有重要作用。

3.4.2 由度规计算黎曼曲率

设 M 上给定度规 g_{ab}，由 $\nabla_{a}g_{bc}=0$ 便决定唯一的联络 ∇_{a}，因而有黎曼张量 $R_{abc}{}^{d}$。常见的问题是已知 g_{ab} 欲求 $R_{abc}{}^{d}$。所谓计算某张量，就是求出它在某基底的分量。基底分

为坐标基底和非坐标基底两大类，本小节只讲用坐标基底求曲率的方法，用非坐标基底的方法将在后面的章节介绍。

任选坐标系后，度规分量 $g_{\mu\nu}$ 便是已知量，满足 $\nabla_a g_{bc} = 0$ 的联络 ∇_a 在此坐标系下的体现就是它在该系的克氏符

$$\Gamma^{\sigma}{}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}(-g_{\mu\nu,\rho} + g_{\rho\mu,\nu} + g_{\nu\rho,\mu})$$

$\Gamma^{\sigma}{}_{\mu\nu}$ 有三个具体指标，故 $\Gamma^{\sigma}{}_{\mu\nu}$ 含 n^3 个数。对称性 $\Gamma^{\sigma}{}_{\mu\nu} = \Gamma^{\sigma}{}_{\nu\mu}$ 使这 n^3 个数中只有 $n^2(n+1)/2$ 个独立（当 $n = 4$ 时有 40 个数独立，其分别为 $\Gamma^0{}_{00}, \Gamma^0{}_{01}, \Gamma^0{}_{02}, \Gamma^0{}_{03}, \Gamma^0{}_{11}, \Gamma^0{}_{12}, \Gamma^0{}_{13}, \Gamma^0{}_{22}, \Gamma^0{}_{23}, \Gamma^0{}_{33}; \Gamma^1{}_{0}$
计算的第一步就是从已知的 $g_{\mu\nu}$ 求出全部非零的 $\Gamma^{\sigma}{}_{\mu\nu}$。

由黎曼张量定义有 $R_{abc}{}^d \omega_d = 2\nabla_{[a}\nabla_{b]}\omega_c$，其中 $\nabla_a \nabla_b \omega_c$ 可借前面证明中的等式表为 6 项（原式共 5 项，第 5 项的 $\nabla_b \omega_d$ 又可展为两项，即 $\partial_b \omega_d - \Gamma^e{}_{bd}\omega_e$），对每项的指标 a, b 反称化，注意到 $\partial_{[a}\partial_{b]}\omega_c = 0$，$\Gamma^d{}_{[ab]} = \Gamma^d{}_{[(ab)]} = 0$，便得

$$\begin{aligned} R_{abc}{}^d \omega_d &= 2(\partial_{[a}\partial_{b]}\omega_c - \partial_{[a}\Gamma^e{}_{b]c}\omega_e - \Gamma^e{}_{c[b}\partial_{a]}\omega_e - \Gamma^d{}_{[ab]}\nabla_d \omega_c - \Gamma^d{}_{c[a}\nabla_{b]}\omega_d) \\ &= 2(-\partial_{[a}\Gamma^e{}_{b]c}\omega_e - \Gamma^e{}_{c[b}\partial_{a]}\omega_e - \Gamma^d{}_{c[a}\partial_{b]}\omega_d + \Gamma^d{}_{c[a}\Gamma^e{}_{b]d}\omega_e) \\ &= -2\partial_{[a}\Gamma^d{}_{b]c}\omega_d + 2\Gamma^e{}_{c[a}\Gamma^d{}_{b]e}\omega_d, \ \forall \omega_d \in \mathscr{F}(0,1) \end{aligned}$$

故 $R_{abc}{}^d = -2\partial_{[a}\Gamma^d{}_{b]c} + 2\Gamma^e{}_{c[a}\Gamma^d{}_{b]e}$，其坐标分量为

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

其中 $\Gamma^{\rho}{}_{\mu\sigma,\nu} \equiv \partial\Gamma^{\rho}{}_{\mu\sigma}/\partial x^{\nu}$。由上式又可得里奇张量的坐标分量表达式

$$R_{\mu\sigma} = R_{\mu\nu\sigma}{}^{\nu} = \Gamma^{\nu}{}_{\mu\sigma,\nu} - \Gamma^{\nu}{}_{\nu\sigma,\mu} + \Gamma^{\lambda}{}_{\sigma\mu}\Gamma^{\nu}{}_{\nu\lambda} - \Gamma^{\lambda}{}_{\sigma\nu}\Gamma^{\nu}{}_{\mu\lambda}$$

3.5 内禀曲率和外曲率

根据直觉，平面是平直的，曲面是弯曲的。说得准确些，这些“平面”和“曲面”都是指镶嵌在 3 维欧氏空间中的 2 维面（后者例如球面和柱面）。现在问：对给定的 n 维流形，可否也仿照这一思路谈及它是否弯曲？只要它能被镶嵌进 $n+1$ 维流形，就可这样讨论。把流形镶进高一维流形所定义的曲率叫“外曲率”，有准确定义。3 维欧氏空间中的球面和圆柱面由这一定义求得的外曲率都非零，同直观感觉吻合。然而本章介绍的黎曼张量却是内禀曲率，它反映流形 M 在指定联络 ∇_a 后的“内禀弯曲性”，无须镶嵌进高一维的流形去判断，与外曲率并不相同（一般而言，(M, g_{ab}) 中凡是只由 g_{ab}（而不必嵌入高一维流形）决定的性质都称为 (M, g_{ab}) 的**内禀性质**。）。“内禀弯曲性”的“弯曲”一词反映的只是以下三个等价性质，具有这些性质的广义黎曼空间叫弯曲空间。

（1） 导数算符的非对易性，即 $(\nabla_a\nabla_b-\nabla_b\nabla_a)\omega_c=R_{abc}{}^d\omega_d$，$\forall\omega_d\in\mathscr{F}(0,1)$，其中非零张量场 $R_{abc}{}^d$ 被用作内禀（黎曼）曲率的定义。

（2） 矢量平移的曲线依赖性

(M,∇_a) 中两点 p,q 的切空间 V_p 和 V_q 之间存在一个曲线依赖的平移映射：对 p,q 之间的一段曲线，p 点的任一矢量 v^a 决定线上的一个平移矢量场 $\tilde{v}^a$（满足 $\tilde{v}^a|_p=v^a$），它在 q 点的值 $\tilde{v}^a|_q$ 就定义为 v^a 的像，或说 $\tilde{v}^a|_q$ 是 v^a 沿线平移至 q 的结果。对欧氏、闵氏空间以及所有平直空间，这一平移与曲线无关，因此在谈到"把 p 点的矢量平移到 q 点"时不必说明沿哪条曲线，这种简单性称为平移的绝对性，是人们十分熟悉的。然而弯曲空间就不如此简单。可以证明，内禀曲率 $R_{abc}{}^d$ 非零的充要条件是存在这样的闭合曲线，线上某点的一个矢量沿线平移一周后不复原（所得矢量与原矢量不等），因此平移同曲线有关（只存在曲线依赖的平移概念）。

（3） 存在初始平行后来不平行的测地线

平直空间的曲率张量场 $R_{abc}{}^d$ 为零，因此不具有以上三个性质中的任一个。具体地说，① 与平直度规适配的导数算符 ∂_a（即笛卡尔或洛伦兹系的普通导数算符）不存在非对易性；② 矢量平移同曲线无关，因此可谈及矢量的"绝对平移"；③ 平行直线永不相交。

内禀曲率与外曲率是不同的概念。例如，3 维欧氏空间中的 2 维圆柱面的外曲率非零而内禀曲率为零。圆柱面可看作平面上介于两条直线 l_1 和 l_2 之间的部分在两直线认同（粘合）后的结果。由于 p 点的 $R_{abc}{}^d$ 的计算只涉及 p 点的一个邻域，p 点的 $R_{abc}{}^d$ 不会因 l_1 和 l_2 的认同而变得非零。

第四章
李导数、Killing 场和超曲面

4.1 流形间的映射

设 M、N 为流形（维数可不同），$\phi\colon M \to N$ 为光滑映射。以 $\mathscr{F}_M$ 和 $\mathscr{F}_N$ 分别代表 M 和 N 上光滑函数的集合，$\mathscr{F}_M(k,l)$ 和 $\mathscr{F}_N(k,l)$ 分别代表 M 和 N 上光滑 (k,l) 型张量场的集合。ϕ 自然诱导出一系列映射如下。

定义 4.1. **拉回映射**$\phi^*\colon \mathscr{F}_N \to \mathscr{F}_M$ 定义为

$$(\phi^* f)|_p \coloneqq f|_{\phi(p)},\ \forall f \in \mathscr{F}_N, p \in M$$

即 $\phi^* f = f \circ \phi$。

由定义不难证明：

（1） $\phi^*\colon \mathscr{F}_N \to \mathscr{F}_M$ 是线性映射，即

$$\phi^*(\alpha f + \beta g) = \alpha\phi^*(f) + \beta\phi^*(g)\ \forall f, g \in \mathscr{F}_N,\ \alpha, \beta \in \mathbb{R}$$

（2） $\phi^*(fg) = \phi^*(f)\phi^*(g),\ \forall f, g \in \mathscr{F}_N$

定义 4.2. 对 M 中任一点 p 可定义**推前映射**$\phi_*\colon V_p \to V_{\phi(p)}$ 如下：$\forall v^a \in V_p$，定义其像 $\phi_* v^a \in V_{\phi(p)}$ 为

$$(\phi_* v)(f) \coloneqq v(\phi * f),\ \forall f \in \mathscr{F}_N$$

还应证明这样定义的 $\phi_* v^a$ 满足矢量定义的两个要求，从而确是 $\phi(p)$ 点的矢量。许多文献也把 ϕ_* 称为 ϕ 的**切映射**。

定理 4.1. $\phi_*\colon V_p \to V_{\phi(p)}$ 是线性映射，即

$$\phi_*(\alpha u^a + \beta v^a) = \alpha\phi_* u^a + \beta\phi_* v^a\ \forall u^a, v^a \in V_p,\ \alpha, \beta \in \mathbb{R}$$

定理 4.2. 设 $C(t)$ 是 M 中的曲线，T^a 为曲线在 $C(t_0)$ 点的切矢，则 $\phi_* T^a \in V_{\phi(C(t_0))}$ 是曲线 $\phi(C(t))$ 在 $\phi(C(t_0))$ 点的切矢（曲线的切矢的像是曲线像的切矢）。

证明. 根据推前映射、曲线切矢以及拉回映射的定义

$$\phi_* T^a(f) = T^a(\phi^* f) = \left.\frac{d(\phi^* f(C))}{dt}\right|_{t_0} = \left.\frac{df(\phi(C))}{dt}\right|_{t_0}$$

而这就是曲线 $\phi(C(t))$ 在 $\phi(C(t_0))$ 点的切矢。 □

定义 4.3. 拉回映射可按如下方式延拓至 $\phi^*: \mathscr{F}_N(0,l) \to \mathscr{F}_M(0,l)$：$\forall T \in \mathscr{F}_N(0,l)$，定义 $\phi^*(T) \in \mathscr{F}_M(0,l)$ 为

$$(\phi^* T)_{a_1 \cdots a_l}|_p (v_1)^{a_1} \cdots (v_l)^{a_l} \coloneqq T_{a_1 \cdots a_l}|_{\phi(p)} (\phi_* v_1)^{a_1} \cdots (\phi_* v_l)^{a_l},\ \forall p \in M, v_1, \cdots, v_l \in V_p$$

定义 4.4. $\forall p \in M$，推前映射 ϕ_* 可按如下方式延拓至 $\phi_*: \mathscr{T}_{V_p}(k,0) \to \mathscr{T}_{V_{\phi(p)}}(k,0)$（即 ϕ_* 是把 p 点的 $(k,0)$ 型张量变为 $\phi(p)$ 点的同型张量的映射）：$\forall T \in \mathscr{T}_{V_p}(k,0)$，其像 $\phi_*(T) \in \mathscr{T}_{V_{\phi(p)}}(k,0)$ 由下式定义：

$$(\phi_* T)^{a_1 \cdots a_l} (\omega^1)_{a_1} \cdots (\omega^k)_{a_k} \coloneqq T^{a_1 \cdots a_l} (\phi^* \omega^1)_{a_1} \cdots (\phi^* \omega^k)_{a_k},\ \forall \omega^1, \cdots, \omega^k \in V^*_{\phi(p)}$$

其中 $(\phi^* \omega)_a$ 定义为 $(\phi^* \omega)_a v^a \coloneqq \omega_a (\phi_* v)^a$，$\forall v^a \in V_p$。

注 4.1. 拉回映射 ϕ^* 能把 N 上的 $(0,l)$ 型张量场变为 M 上的同型张量场，是场变为场的映射；推前映射 ϕ_* 只能把 M 中一点 p 的 $(k,0)$ 型张量变为其像点 $\phi(p)$ 的同型张量。可否将 ϕ_* 延拓为把 M 上的 $(k,0)$ 型张量场变为 N 上的同型张量场的映射？在一般情况下不能。以矢量场为例。关键在于，给定 M 上一个矢量场 v 后，要定义 N 上的像矢场 $\phi_* v$ 就要对 N 的任一点 q 定义一个矢量，而这势必涉及 q 点的逆像 $\phi^{-1}(q)$。如果 ϕ 不是到上映射，则 $\phi^{-1}(q)$ 可能不存在，从而无法用 $\phi^{-1}(q)$ 点的 v 作为右边的 v；如果 ϕ 不是一一映射，则逆像 $\phi^{-1}(q)$ 可能多于一点，从而无法确定该用哪一逆像点的 v 作为右边的 v。这暗示，如果 ϕ 只是光滑映射，则 ϕ_* 未必能把场推前为场。然而，如果 $\phi: M \to N$ 是微分同胚映射，则上述困难不复存在，推前映射 ϕ_* 可看作把 M 上的 $(k,0)$ 型张量场变为 N 上同型张量场的映射，即 $\phi_*: \mathscr{F}_M(k,0) \to \mathscr{F}_N(k,0)$。再者，由于 ϕ^{-1} 存在而且光滑，其拉回映射 ϕ^{-1*} 把 $\mathscr{F}_M(0,l)$ 映到 $\mathscr{F}_N(0,l)$，这可看作 ϕ 的推前映射 ϕ_*，于是 ϕ_* 又可进一步推广为 $\phi_*: \mathscr{F}_M(k,l) \to \mathscr{F}_N(k,l)$。例如，设 $T^a{}_b \in \mathscr{F}_M(1,1)$，则 $(\phi_* T)^a{}_b \in \mathscr{F}_N(1,1)$ 定义为

$$(\phi_* T)^a{}_b|_q \omega_a v^b \coloneqq T^a{}_b|_{\phi^{-1}(q)} (\phi^* \omega)_a (\phi^* v)^b,\ \forall q \in N, \omega_a \in V^*_q, v^b \in V_q$$

其中，$(\phi^* v)^b$ 应理解为 $(\phi_*^{-1} v)^b$。同理，拉回映射也可推广为 $\phi^*: \mathscr{F}_N(k,l) \to \mathscr{F}_M(k,l)$。推广后的 ϕ_* 和 ϕ^* 仍为线性映射，而且互逆。

设 $\phi\colon M \to N$ 是微分同胚，$p \in M$，$\{x^\mu\}$ 和 $\{y^\mu\}$ 分别是 M 和 N 的局域坐标系，坐标域 O_1 和 O_2 满足 $p \in O_1, \phi(p) \in O_2$。于是 $p \in \phi^{-1}[O_2]$。ϕ 为微分同胚保证 M 和 N 的维数相等，故 $\{x^\mu\}$ 和 $\{y^\mu\}$ 的 μ 都是从 1 到 n。微分同胚本是点的变换，但也可等价地看作坐标变换，因为可用 $\phi\colon M \to N$ 在 $\phi^{-1}[O_2]$ 上定义一组新坐标 $\{x'^\mu\}$ 如下：$\forall q \in \phi^{-1}[O_2]$，定义 $x'^\mu(q) \coloneqq y^\mu(\phi(q))$。可见微分同胚映射 ϕ 在 p 的邻域 $O_1 \cap \phi^{-1}[O_2]$ 上自动诱导出一个坐标变换 $x^\mu \mapsto x'^\mu$。由前述定理不难证明 $\forall q \in O_1 \cap \phi^{-1}[O_2]$ 有

$$\phi_*[(\partial/\partial x'^\mu)^a|_q] = (\partial/\partial y^\mu)^a|_{\phi(q)}$$

由此又可证明

$$\phi_*[(dx'^\mu)^a|_q] = (dy^\mu)^a|_{\phi(q)}$$

于是对微分同胚映射 $\phi\colon M \to N$ 就存在两种观点：① **主动观点**, 它如实地认为 ϕ 是点的变换（把 p 变为 $\phi(p)$）以及由此导致的张量变换（把 p 点的张量 T 变为 $\phi(p)$ 点的张量 $\phi_* T$）；② **被动观点**，它认为点 p 及其上的所有张量 T 都没变，$\phi\colon M \to N$ 的后果是坐标系有了变换（从 $\{x^\mu\}$ 变为 $\{x'^\mu\}$）。这两种观点虽然似乎相去甚远，但在实用上是等价的。下面的定理可以看作等价性的某种表现。

定理 4.3. $(\phi_* T)^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}|_{\phi(p)} = T'^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}|_p$, $\forall T \in \mathscr{F}_M(k,l)$

式中左边是新点 $\phi(p)$ 的新张量 $\phi_* T$ 在老坐标系 $\{y^\mu\}$ 的分量，右边是老点 p 的老张量 T 在新坐标系 $\{x'^\mu\}$ 的分量。

注 4.2. *上式是实数等式，左边是由主动观点（认为点和张量变了而坐标系没变）所得的数，右边是由被动观点（认为点和张量没变但坐标系变了）所得的数。两边相等就表明两种观点在实用上等价。*

例 4.1. *设 $v^a \in V_p$，令 $u^a \equiv \phi_* v^a \in V_{\phi(p)}$，则由上述定理不难证明*

$$u^\mu = v^\nu (\partial x'^\mu/\partial x^\nu)|_p$$

4.2 李导数

根据前面的知识我们知道 M 上的一个光滑矢量场 v^a 给出一个单参微分同胚群 ϕ。[1]设 $T^{\cdots}{}_{\cdots}$ 是 M 上的光滑张量场，则 $\phi_t^* T^{\cdots}{}_{\cdots}$ 也是同型光滑张量场，其中 ϕ_t 是单参微分同胚群 ϕ 的一个群元。这两个张量场在点 $p \in M$ 的值之差 $\phi_t^* T^{\cdots}{}_{\cdots}|_p - T^{\cdots}{}_{\cdots}|_p$ 是 p 点的张量，它与 t 之商 $(\phi_t^* T^{\cdots}{}_{\cdots}|_p - T^{\cdots}{}_{\cdots}|_p)/t$ 在 t 趋于零时的极限可看作张量场 $T^{\cdots}{}_{\cdots}$ 在 p 点的某种导数，于是有以下定义：

定义 4.5. $\mathscr{L}_v T^{a_1\cdots a_k}{}_{b_1\cdots b_l} \coloneqq \lim\limits_{t\to 0} \frac{1}{t}(\phi_t^* T^{a_1\cdots a_k}{}_{b_1\cdots b_l} - T^{a_1\cdots a_k}{}_{b_1\cdots b_l})$

*称为张量场 $T^{a_1\cdots a_k}{}_{b_1\cdots b_l}$ 沿矢量场 v^a 的***李导数***（$\mathscr{L}_v$ 中的 v 不写为 v^a，以免误解。）。*

[1]若 v^a 不完备，则只能给出单参微分同胚局部群。本节只涉及局部性质，无须明确区分局部和整体。

注 4.3. 因 ϕ_t^* 为线性映射，故李导数是由 $\mathscr{F}_M(k,l)$ 到 $\mathscr{F}_M(k,l)$ 的线性映射。还可证明 $\mathscr{L}_v$ 同缩并可交换顺序。

定理 4.4. $\mathscr{L}_v f = v(f),\ \forall f \in \mathscr{F}$

证明. $\forall p \in M$，设 $C(t)$ 是 ϕ 过 p 点的轨道，$p = C(0)$，则 $\phi_t(p) = C(t)$ 且 $v^a|_p \equiv (\partial/\partial t)^a|_p$ 是 $C(t)$ 在 p 点的切矢，故

$$\begin{aligned}\mathscr{L}_v f|_p &= \lim_{t\to 0}\frac{1}{t}(\phi_t^* f - f)|_p = \lim_{t\to 0}\frac{1}{t}[f(\phi_t(p)) - f(p)]\\ &= \lim_{t\to 0}\frac{1}{t}[f(C(t)) - f(C(0))] = \frac{d}{dt}(f\circ C)|_{t=0} = v(f)|_p\end{aligned}$$

□

下面以 $n = 2$ 为例介绍一种对计算李导数很有用的坐标系。设 $\{x^1, x^2\}$ 为坐标系，则 x^1 坐标线和 x^2 坐标线组成坐标“网格”，欲知坐标域中某点的坐标，只须看它位于网格的哪两条坐标线的交点。求李导数时总要给定矢量场 v^a，可以选定它的积分曲线为 x^1 坐标线（t 充当 x^1），再相当任意地选定另一组与这组曲线横截（即交点上两线切矢不平行）的曲线作为 x^2 坐标线，这样得到的坐标系称为矢量场 v^a 的**适配坐标系**。换句话说，矢量场 v^a 就是其适配坐标系的第一坐标基矢场，即 $v^a = (\partial/\partial x^1)^a$。以上讨论可推广至任意维流形。

定理 4.5. 张量场 $T^{a_1\cdots a_k}{}_{b_1\cdots b_l}$ 沿 v^a 的李导数在 v^a 的适配坐标系的分量

$$(\mathscr{L}_v T)^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = \frac{\partial T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}}{\partial x^1}$$

注 4.4. 上式左边在坐标变换时满足张量变换律而右边则否，故不能改写为张量等式。

证明. 仅以 $n = 2, k = l = 1$ 为例（容易推广至一般情况）。因 $\phi_t^* = (\phi_t^{-1})_* = \phi_{-t*}$，李导数定义式在任一坐标系（现在取适配坐标系）的分量式为

$$(\mathscr{L}_v T)^\mu{}_\nu|_p = \lim_{t\to 0}\frac{1}{t}[(\phi_{-t*}T)^\mu{}_\nu|_p - T^\mu{}_\nu|_p],\ \forall p \in M$$

令 $q \equiv \phi_t(p)$，因上式只涉及 p 点附近的情况，总可以认为 p, q 点都在同一适配坐标域内。对 ϕ_{-t} 而言，q 为老点，p 为新点，故由坐标变换“主动观点”和“被动观点”的等价性得

$$(\phi_{-t*}T)^\mu{}_\nu|_p = T'^\mu{}_\nu|_q = \left[\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}T^\rho{}_\sigma\right]_q$$

其中第二步用到张量的分量变换律。式中 x^σ 为适配（老）坐标，x'^μ 是由 ϕ_{-t} 诱导的新坐标。上式右边涉及新老坐标间的偏导数在 q 点的值，要计算就必须找出 q 点的一个小邻域 N 内的坐标变换。$\forall \bar{q} \in N$，记 $\bar{p} \equiv \phi_{-t}(\bar{q})$。由适配坐标的定义知 $x^1(\bar{q}) = x^1(\bar{p}) + t, x^2(\bar{q}) = x^2(\bar{p})$，而按定义，$\phi_{-t}$ 在 $\bar{q}$ 诱导的新坐标则为 $x'^1(\bar{q}) \equiv x^1(\bar{p}), x'^2(\bar{q}) \equiv x^2(\bar{p})$，故 $x'^1(\bar{q}) = x^1(\bar{q}) - t, x'^2(\bar{q}) = x^2(\bar{q})$。因为 $\bar{q}$ 为 N 内任一点，故对 N 有 $x'^1 = x^1 - t, x'^2 = x^2$，求导得 $(\partial x'^\mu/\partial x^\rho)|_q = \delta^\mu{}_\rho, (\partial x^\sigma/\partial x'^\nu)|_q = \delta^\sigma{}_\nu$，于是上式成为 $(\phi_{-t*}T)^\mu{}_\nu|_p = T^\mu{}_\nu|_q$，带入前式便得 $(\mathscr{L}_v T)^\mu{}_\nu|_p = \partial T^\mu{}_\nu/\partial x^1|_p$。 □

由上述定理可知 $\mathscr{L}_v$ 满足莱布尼茨律。

定理 4.6. $\mathscr{L}_v u^a = [v,u]^a$, $\forall u^a, v^a \in \mathscr{F}(1,0)$

或者，借助于对易子的表达式，有

$$\mathscr{L}_v u^a = v^b \nabla_b u^a - u^b \nabla_b v^a$$

其中 ∇_a 为任一无挠导数算符。

证明. 待证命题是矢量等式，只须证明它在某一坐标系的分量等式 $(\mathscr{L}_v u)^\mu = [v,u]^\mu$ 成立。最方便的当然是适配坐标系。设 v^a 的适配坐标系 $\{x^\mu\}$ 的普通导数算符是 ∂_a，则

$$[v,u]^\mu = (dx^\mu)_a [v,u]^a = (dx^\mu)_a (v^b \partial_b u^a - u^b \partial_b v^a) = v^b \partial_b u^\mu = v(u^\mu) = \partial u^\mu / \partial x^1 = (\mathscr{L}_v u)^\mu$$

其中第三步是因为 $v^a = (\partial/\partial x^1)^a$，而普通导数算符作用于任一坐标基矢为零，第四步用到导数算符的定义条件，最后一步用到前面的定理。 □

定理 4.7. $\mathscr{L}_v \omega_a = v^b \nabla_b \omega_a + \omega_b \nabla_a v^b$, $\forall \omega^a \in \mathscr{F}(0,1), v^a \in \mathscr{F}(1,0)$

其中 ∇_a 为任一无挠导数算符。

证明. 一方面 $\mathscr{L}_v(\omega_a u^a) = v^b \nabla_b (\omega_a u^a) = v^b \nabla_b \omega_a u^a + v^b \omega_a \nabla_b u^a$。另一方面 $\mathscr{L}_v(\omega_a u^a) = \mathscr{L}_v \omega_a u^a + \omega_a \mathscr{L}_v u^a = \mathscr{L}_v \omega_a u^a + \omega_a v^b \nabla_b u^a - \omega_a u^b \nabla_b v^a$。因此,$\mathscr{L}_v \omega_a u^a = v^b \nabla_b \omega_a u^a + \omega_a u^b \nabla_b v^a$。该式对任意的 u^a 都成立，因此命题得证。 □

定理 4.8.

$$\mathscr{L}_v T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} = v^c \nabla_c T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} - \sum_{i=1}^{k} T^{a_1 \cdots c \cdots a_k}{}_{b_1 \cdots b_l} \nabla_c v^{a_i} + \sum_{j=1}^{l} T^{a_1 \cdots a_k}{}_{b_1 \cdots c \cdots b_l} \nabla_{b_j} v^c$$

$\forall T \in \mathscr{F}(k,l), v \in \mathscr{F}(1,0)$，$\nabla_a$ 为任一导数算符。

4.3 Killing 矢量场

本章至此未涉及度规及与之适配的导数算符，李导数的定义不要求流形 M 有附加结构。但若 M 上选定了度规场 g_{ab}，则对微分同胚映射 $\phi\colon M \to M$ 还可提出更高的要求，即 $\phi^* g_{ab} = g_{ab}$。于是有如下定义：

定义 4.6. 微分同胚 $\phi\colon M \to M$ 称为**等度规映射**，简称**等度规**，若 $\phi^* g_{ab} = g_{ab}$。

注 4.5. ① 等度规映射是特殊的微分同胚映射,其特殊性在于“保度规”性,即 $\phi^* g_{ab} = g_{ab}$。注意这是张量场的等式，其含义是每点 p 的两个张量 $g_{ab}|_p$ 和 $\phi^* g_{ab}|_p$ 相等。② 由 ϕ^{-1*} 。$\phi^* = (\phi \circ \phi^{-1})^* =$ 恒等映射 易见 $\phi\colon M \to M$ 为等度规映射当且仅当 $\phi^{-1}\colon M \to M$ 为等度规映射。

流形 M 上众多矢量场中有一类特殊矢量场，即光滑矢量场。每一光滑矢量场给出一个单参微分同胚群。如果 M 上指定了度规场 g_{ab}，则众多光滑矢量场中还可挑出特殊的一个子类，其中每个矢量场给出的单参微分同胚群是单参等度规群，即每个群元 $\phi_t: M \to M$ 是 M 上的一个等度规映射。于是有以下定义：

定义 4.7. (M, g_{ab}) 上的矢量场 ξ^a 称为 ***Killing* 矢量场**，若它给出的单参微分同胚（局部）群是单参等度规（局部）群。等价地（还可以证明），ξ^a 称为 *Killing* 矢量场，若 $\mathscr{L}_\xi g_{ab} = 0$。

定理 4.9. ξ^a 为 *Killing* 矢量场的充要条件是 ξ^a 满足如下的 ***Killing* 方程**：

$$\nabla_a \xi_b + \nabla_b \xi_a = 0,\ 或 \nabla_{(a}\xi_{b)} = 0,\ 或 \nabla_a \xi_b = \nabla_{[a}\xi_{b]}$$

（其中 ∇_a 满足 $\nabla_a g_{bc} = 0$）

证明. $0 = \mathscr{L}_\xi g_{ab} = \xi^c \nabla_c g_{ab} + g_{cb}\nabla_a \xi^c + g_{ac}\nabla_b \xi^c = \nabla_a \xi_b + \nabla_b \xi_a$ □

定理 4.10. 若存在坐标系 $\{x^\mu\}$ 使 g_{ab} 的全部分量满足 $\partial g_{\mu\nu}/\partial x^1 = 0$，则 $(\partial/\partial x^1)^a$ 是坐标域上的 *Killing* 矢量场。

证明. $\{x^\mu\}$ 是 $(\partial/\partial x^1)^a$ 的适配坐标系。因此，$(\mathscr{L}_{\partial/\partial x^1} g)_{\mu\nu} = \partial g_{\mu\nu}/\partial x^1 = 0$，故 $\mathscr{L}_{\partial/\partial x^1} g_{ab} = 0$，即 $(\partial/\partial x^1)^a$ 为 Killing 矢量场。 □

定理 4.11. 设 ξ^a 为 *Killing* 矢量场，T^a 为测地线的切矢，则 $T^a\nabla_a(T^b\xi_b) = 0$，即 $T^b\xi_b$ 在测地线上为常数。

证明.

$$T^a\nabla_a(T^b\xi_b) = T^a\nabla_a T^b \xi_b + T^a T^b \nabla_a \xi_b = T^{(a}T^{b)}\nabla_a\xi_b = T^a T^b \nabla_{(a}\xi_{b)} = 0$$

其中用到测地线方程和 Killing 方程。 □

设 ξ^a, η^a 是 Killing 矢量场，α, β 是常实数，则由 Killing 方程的线性性知 $\alpha\xi^a + \beta\eta^a$ 也是 Killing 矢量场。不难看出 M 上所有 Killing 矢量场的集合是个矢量空间。还可证明对易子 $[\xi, \eta]^a$ 也是 Killing 矢量场。

定理 4.12. (M, g_{ab}) 上最多有 $n(n+1)/2$ 个独立的 *Killing* 矢量场（$n \equiv \dim M$），即 M 上所有 *Killing* 矢量场的集合（作为矢量空间）的维数小于等于 $n(n+1)/2$。

注 4.6. ① 等度规映射可看作一种“保度规”的对称变换，所以一个 *Killing* 矢量场代表 (M, g_{ab}) 的一个对称性，具有 $n(n+1)/2$ 个独立 *Killing* 矢量场的广义黎曼空间 (M, g_{ab}) 称为最高对称性空间。② 寻找 (M, g_{ab}) 的全体 *Killing* 矢量场的一般方法是求 *Killing* 方程的通解。然而对某些简单的 (M, g_{ab}) 还存在容易得多的方法。下面仅举数例。

例 4.2. 找出下列广义黎曼空间的全体独立的 *Killing* 矢量场。

（*1*） 2 维欧氏空间 $(\mathbb{R}^2, \delta_{ab})$。设 $\{x, y\}$ 为笛卡尔坐标系，则 $ds^2 = dx^2 + dy^2$，即欧氏度规 δ_{ab} 在此系中的全部分量为常数，故由前述定理知 $(\partial/\partial x)^a$ 和 $(\partial/\partial y)^a$ 为 *Killing* 矢量场。我们相信欧氏空间有最高对称性，$n = 2$ 时应有 3 个独立的 *Killing* 矢量场。果然，若改用极坐标系，便有 $ds^2 = dr^2 + r^2 d\varphi^2$，可见欧氏度规 δ_{ab} 在此系中的全部分量与 φ 无关，所以 $(\partial/\partial\varphi)^a$ 为 *Killing* 矢量场，它在笛卡尔系的坐标基底的展开式为 $(\partial/\partial\varphi)^a = -y(\partial/\partial x)^a + x(\partial/\partial y)^a$。展开系数与坐标有关，由此不难证明 $(\partial/\partial\varphi)^a$ 独立于前两个 *Killing* 场。$(\partial/\partial x)^a$ 和 $(\partial/\partial y)^a$ 的 *Killing* 性反映 2 维欧氏度规沿 x 和 y 轴的平移不变性，$(\partial/\partial\varphi)^a$ 的 *Killing* 性表明它有旋转不变性。

（*2*） 3 维欧氏空间 $(\mathbb{R}^3, \delta_{ab})$。因为 $n = 3$，故有 6 个独立的 *Killing* 矢量场，即 $(\partial/\partial x)^a, (\partial/\partial y)^a, (\partial/\partial z)^a, -y(\partial/\partial x)^a + x(\partial/\partial y)^a, -z(\partial/\partial y)^a + y(\partial/\partial z)^a$ 和 $-x(\partial/\partial z)^a + z(\partial/\partial x)^a$。前 3 个反映 3 维欧氏度规沿 x, y, z 轴的平移不变性；后 3 个反映它绕 z, x, y 轴的旋转不变性。

（*3*） 2 维闵氏空间 $(\mathbb{R}^2, \eta_{ab})$。在洛伦兹坐标系 $\{t, x\}$ 中有 $ds^2 = -dt^2 + dx^2$，故知 $(\partial/\partial t)^a$ 和 $(\partial/\partial x)^a$ 为 *Killing* 场。为求第三个，用下式定义新坐标 ψ, η：

$$x = \psi \operatorname{ch}\eta,\ t = \psi \operatorname{sh}\eta,\ 0 < \psi < \infty,\ -\infty < \eta < \infty$$

闵氏线元可用新坐标表为 $ds^2 = d\psi^2 - \psi^2 d\eta^2$。上式表明 η_{ab} 在新坐标系的全体分量与坐标 η 无关，故 $(\partial/\partial\eta)^a$ 也是 *Killing* 矢量场（其积分曲线是双曲线），它在洛伦兹坐标基底的展开式为

$$(\partial/\partial\eta)^a = t(\partial/\partial x)^a + x(\partial/\partial t)^a$$

由展开系数与坐标有关可知 $(\partial/\partial\eta)^a$ 与前两个 *Killing* 场独立。不难验证 $(\partial/\partial\eta)^a$ 是 $\mathbb{R}^2$ 上的 *Killing* 矢量场，叫做**伪转动**（*boost*）*Killing* 矢量场，表明闵氏度规具有伪转动下的不变性，对应于洛伦兹变换。

（*4*） 4 维闵氏空间 $(\mathbb{R}^4, \eta_{ab})$。因 $n = 4$，故独立的 *Killing* 场共 10 个，分三组：

（*a*） 4 个平移 $(\partial/\partial t)^a, (\partial/\partial x)^a, (\partial/\partial y)^a, (\partial/\partial z)^a$;

（*b*） 3 个空间转动 $-y(\partial/\partial x)^a + x(\partial/\partial y)^a, -z(\partial/\partial y)^a + y(\partial/\partial z)^a, -x(\partial/\partial z)^a + z(\partial/\partial x)^a$;

（*c*） 3 个伪转动 $t(\partial/\partial x)^a + x(\partial/\partial t)^a, t(\partial/\partial y)^a + y(\partial/\partial t)^a, t(\partial/\partial z)^a + z(\partial/\partial t)^a$

组（*a*）反映闵氏度规沿 t, x, y, z 轴的平移不变性，组（*b*）反映它绕 z, y, x 轴的空间旋转不变性，组（*c*）反映它在 tx, ty, tz 面内的伪转动下的不变性。

定理 4.13. 设 $\{x, t\}$ 是 2 维闵氏空间 $(\mathbb{R}^2, \eta_{ab})$ 的洛伦兹坐标系，$\phi_\lambda\colon \mathbb{R}^2 \to \mathbb{R}^2$ 是伪转动 *Killing* 场 $\xi^a \equiv t(\partial/\partial x)^a + x(\partial/\partial t)^a$ 对应的单参等度规群的一个群元（即以参数 $\lambda \in \mathbb{R}$ 刻画的那个等度规映射），则由 ϕ_λ 诱导的坐标变换 $\{x, t\} \mapsto \{x', t'\}$ 是洛伦兹变换。

注 4.7. 本定理表明伪转动和洛伦兹变换是同一变换的两种（主动与被动）提法。类似地，欧氏空间的转动 *Killing* 场 $-y(\partial/\partial x)^a + x(\partial/\partial y)^a$ 与坐标变换

$$x' = x\cos\alpha - y\sin\alpha,\ y' = x\sin\alpha + y\cos\alpha$$

也是同一变换的两种提法。

证明. 矢量场 $\xi^a \equiv (\partial/\partial\eta)^a$ 的积分曲线的参数方程为 $dx^\mu(\eta)/d\eta = \xi^\mu$（$\mu = 0, 1$）。注意到 $\xi^a \equiv t(\partial/\partial x)^a + x(\partial/\partial t)^a$，便得

$$\frac{dx(\eta)}{d\eta} = t(\eta),\ \frac{dt(\eta)}{d\eta} = x(\eta)$$

$\forall p \in \mathbb{R}^2$，设 $C(\eta)$ 是满足 $p = C(0)$ 的积分曲线，即 $x(0) = x_p, t(0) = t_p$，则不难证明上述方程的特解（即该线的参数式）为

$$x(\eta) = x_p \operatorname{ch}\eta + t_p \operatorname{sh}\eta,\ t(\eta) = x_p \operatorname{sh}\eta + t_p \operatorname{ch}\eta$$

设 $q \equiv \phi_\lambda(p)$，则 q 就是 $C(\eta)$ 上参数值 $\eta = \lambda$ 的点，即 $q = C(\lambda)$，故由 ϕ_λ 诱导的新坐标 t' 和 x' 满足

$$x'_p \equiv x_q = x_p \operatorname{ch}\lambda + t_p \operatorname{sh}\lambda,\ t'_p \equiv t_q = x_p \operatorname{sh}\lambda + t_p \operatorname{ch}\lambda$$

因 p 点任意，故可去掉下标 p 而写成

$$x' = x\operatorname{ch}\lambda + t\operatorname{sh}\lambda = \operatorname{ch}\lambda(x + t\operatorname{th}\lambda),\ t' = t\operatorname{ch}\lambda + x\operatorname{sh}\lambda = \operatorname{ch}\lambda(t + x\operatorname{th}\lambda)$$

令 $v \equiv \operatorname{th}\lambda, \gamma \equiv (1 - v^2)^{-1/2} = \operatorname{ch}\lambda$，则

$$x' = \gamma(x + vt),\ t' = \gamma(t + vx)$$

这便是熟知的洛伦兹变换（注意，我们用几何单位制，其中光速 $c = 1$。）。 □

由 $ds^2 = -dt^2 + dx^2$ 和上述定理易得 $ds^2 = -dt'^2 + dx'^2$，可见伪转动对应的等度规映射诱导的坐标变换把洛伦兹系 $\{t, x\}$ 变为洛伦兹系 $\{t', x'\}$。此结果可推广为如下定理：

定理 4.14. 设 $\{x^\mu\}$ 是 $(\mathbb{R}^n, \eta_{ab})$ 的洛伦兹坐标系，则 $\{x'^\mu\}$ 也是洛伦兹坐标系的充要条件是它由 $\{x^\mu\}$ 通过等度规映射 $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ 诱导而得。

证明. 把 η_{ab} 记作 g_{ab}，其在 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 系的分量分别记作 $g_{\mu\nu}$ 和 $g'_{\mu\nu}$。

（A） 设 $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ 是等度规映射（即 $\phi^* g_{ab} = g_{ab}$），$\{x'^\mu\}$ 是由洛伦兹系 $\{x^\mu\}$ 通过 ϕ 诱导而得的坐标系，则 $\forall p \in \mathbb{R}^n$ 有 $g'_{\mu\nu}|_p = (\phi_* g)_{\mu\nu}|_{\phi(p)} = (\phi^{-1*} g)_{\mu\nu}|_{\phi(p)} = g_{\mu\nu}|_{\phi(p)} = \eta_{\mu\nu}$。其中第一步是坐标变换的主被动观点的等价性，最后一步用到 $\{x^\mu\}$ 的洛伦兹性。上式说明 p 点的 g_{ab} 在 $\{x'^\mu\}$ 系的分量为 $\eta_{\mu\nu}$，故 $\{x'^\mu\}$ 为洛伦兹系。

（B） 设 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 都是洛伦兹系，$\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ 是与坐标变换 $\{x^\mu\} \mapsto \{x'^\mu\}$ 对应的微分同胚映射，则 $\forall p \in \mathbb{R}^n$ 有 $(\phi^{-1*}g)_{\mu\nu}|_p = (\phi_* g)_{\mu\nu}|_p = g'_{\mu\nu}|_{\phi^{-1}(p)} = \eta_{\mu\nu} = g_{\mu\nu}|_p$，其中，最后两步用到 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 的洛伦兹性。上式表明 $\phi^{-1*}g_{ab} = g_{ab}$，故 ϕ^{-1}（因而 ϕ）是等度规映射。

□

注 4.8. 本定理也适用于欧氏空间，只须把洛伦兹系改为笛卡尔系。可以说等度规映射保持坐标系的洛伦兹（笛卡尔）性。

4.4 超曲面

定义 4.8. 设 M,S 为流形，$\dim S \leqslant \dim M \equiv n$。映射 $\phi\colon S \to M$ 称为**嵌入**, 若 ϕ 是一一和 C^∞ 的, 而且 $\forall p \in S$，推前映射 $\phi_*\colon V_p \to V_{\phi(p)}$ 非退化（$V_{\phi(p)}$ 是指 $\phi(p)$ 作为 M 的一点的切空间）, 即 $\phi_* v^a = 0 \Rightarrow v^a = 0$.

注 4.9. 嵌入的上述条件使 S 的拓扑和流形结构可自然地被带到 $\phi[S]$ 上去, 从而使 $\phi\colon S \to \phi[S]$ 成为微分同胚。

定义 4.9. 嵌入 $\phi\colon S \to M$ 称为 M 的一个嵌入子流形，简称子流形。也常把映射的像 $\phi[S]$ 称为**嵌入子流形**。若 $\dim S = n - 1$, 则 $\phi[S] \subset M$ 称为 M 的一张**超曲面**。

例 4.3. 设 U 是 M 的开子集, 把 M 的流形结构限制在 U 上，U 便成为与 M 同维的流形。把 U 看作定义中的 S, 令 $\phi\colon U \to M$ 为恒等映射, 则 $U \equiv \phi[U]$ 便是 M 的一个嵌入子流形（同维嵌入）。

例 4.4. 设 S 是 $\mathbb{R}^3$（看作 M）中的单位球面 S^2, 则恒等映射 $\phi\colon S^2 \to \mathbb{R}^3$ 给出 $\mathbb{R}^3$ 的一个嵌入子流形。注意到 S^2 比 $\mathbb{R}^3$ 低一维, 可知 S^2 是 $\mathbb{R}^3$ 的一个超曲面。

设 $\phi[S]$ 是 M 的超曲面,$q \in \phi[S] \subset M$。作为 M 的一点，q 有 n 维切空间 V_q。若 $w^a \in V_q$ 是过 q 且躺在 $\phi[S]$ 上的某曲线的切矢（“躺在” 是指曲线每点都在 $\phi[S]$ 上）, 则说 w^a 切于 $\phi[S]$。V_q 中全体切于 $\phi[S]$ 的元素构成的子集记作 W_q。超曲面的定义保证 W_q 是 V_q 的 $n-1$ 维子空间。谈到超曲面时自然想到它的法矢。设 $\phi[S]$ 是超曲面,$q \in \phi[S]$, 则过 q 点的法矢 n^a 应定义为与 q 点所有切于 $\phi[S]$ 的矢量正交的矢量。然而正交性只有在指定度规后才有意义。当 M 没有度规时，不能定义法矢 n^a，但可定义 “法余矢”n_a。**余矢**是对偶矢量的别名。由于对偶矢量作用于矢量给出实数（无需度规）, 可定义法余矢如下:

定义 4.10. 设 $\phi[S]$ 是超曲面，$q \in \phi[S]$。非零对偶矢量 $n_a \in V_q^*$ 称为 $\phi[S]$ 在 q 点的**法余矢**, 若 $n_a w^a = 0\ \forall w^a \in W_q$。

定理 4.15. 超曲面 $\phi[S]$ 上任一点 q 必有法余矢 n_a。法余矢不唯一, 但 q 点的任意两个法余矢之间只能差一实数因子。

证明. 设 $\{(e_2)^a, \cdots, (e_n)^a\}$ 为 W_q 任一基底, 因 $\dim V_q = n$, V_q 必有与 $\{(e_2)^a, \cdots, (e_n)^a\}$ 线性无关的元素, 任取其一并记作 $(e_1)^a$, 则 $\{(e_\mu)^a \mid \mu = 1, \cdots, n\}$ 为 V_q 的基底，其对偶基底记作 $\{(e^\mu)_a\}$。令 $n_a = (e^1)_a$，则 $n_a(e_\tau)^a = \delta^1{}_\tau = 0$ $(\tau = 2, \cdots, n)$, 故 $n_a w^a = 0\ \forall w^a \in W_q$, 可见 n_a 为法余矢。若存在 m_a 满足 $m_a(e_\tau)^a = 0$ $(\tau = 2, \cdots, n)$，则其在对偶基底 $\{(e^\mu)_a\}$ 的分量 $m_\tau = m_a(e_\tau)^a = 0$ $(\tau = 2, \cdots, n)$, 因而 $m_a = m_1(e^1)_a = m_1 n_a$, 即 m_a 与 n_a 只差一因子 m_1。 □

注 4.10. 非超曲面嵌入子流形（如 3 维流形中的曲线）的法余矢没有这样好的唯一性。

定理 4.16. 设 $\phi[S]$ 是由 $f =$ 常数 给出的超曲面，则面上的 $\nabla_a f$ 是超曲面的法余矢。

证明. 只须对任一 $q \in \phi[S]$ 证明 $w^a \nabla_a f = 0\ \forall w^a \in W_q$。因 w^a 总切于过 q 并躺在 $\phi[S]$ 上的某曲线 $C(t)$, 故 $w^a \nabla_a f = \frac{\partial}{\partial t}(f) = 0\ \forall w^a \in W_q$, 最后一步是因 f 在 $C(t)$ 上为常数。 □

若 M 上有度规 g_{ab}，则 $n^a \equiv g^{ab} n_b \in V_q$ 与 $\phi[S]$ 的所有矢量正交（因 $g_{ab} n^a w^b = n_b w^b = 0\ \forall w^a \in W_q$）, 故 n^a 叫超曲面 $\phi[S]$ 在 q 点的**法矢**。若 g_{ab} 为正定度规（例如 $\mathbb{R}^2$ 嵌入 3 维欧氏空间）,n^a 自然不属于 W_q, 即 $n^a \in V_q - W_q$; 然而, 若 g_{ab} 为洛伦兹度规，n^a 却有可能属于 W_q。以下就 g_{ab} 为洛伦兹度规的情况进行讨论。

定理 4.17. $n^a \in W_q$ 的充要条件为 $n^a n_a = 0$。

证明.（A） 设 $n^a \in W_q$, 则 n^a 可看作 $n_a w^a = 0$ 中的 w^a，故 $n_a n^a = 0$。

（B） 由前面定理的证明知道对任一法余矢 n_a 存在基底 $\{(e_\mu)^a\}$ 使 $(e_2)^a, \cdots, (e_n)^a \in W_q$ 且 $n_a = (e^1)_a$, 故 n^a 在该基底的第一分量 $n^1 = n^a(e^1)_a = n^a n_a$。因此 $n_a n^a = 0 \Rightarrow n^1 = 0 \Rightarrow n^a = \sum_{\tau=2}^{n} n^\tau (e_\tau)^a \in W_q$。

□

例 4.5. 设 $S = \mathbb{R}, M = \mathbb{R}^2, M$ 上度规 $g_{ab} = \eta_{ab}, \phi\colon \mathbb{R} \to \mathbb{R}^2$ 为嵌入, 则 $\phi[\mathbb{R}]$ 是 2 维闵氏时空中的超曲面。设 t, x 为洛伦兹坐标，讨论以下三种有代表性的情况:

（1） $\phi[\mathbb{R}]$ 与 x 轴平行。$\forall q \in \phi[\mathbb{R}]$，令 $(e_2)^a = (\partial/\partial x)^a$，选

$$(e_1)^a = \alpha(\partial/\partial t)^a + \beta(\partial/\partial x)^a, (\alpha, \beta \text{ 可为任意实数，但 } \alpha \neq 0。)$$

则不难验证 $(e^1)_a = \alpha^{-1}(dt)_a$。根据前面定理的证明过程可知 $(e^1)_a$ 为法余矢 n_a，相应的法矢为 $n^a = \alpha^{-1} g^{ab}(dt)_b = -\alpha^{-1}(\partial/\partial t)^a$，满足 $n^a \notin W_q$ 且 $n_a n^a < 0$（即 n^a 为类时）。

（2） $\phi[\mathbb{R}]$ 与 t 轴平行。$\forall q \in \phi[\mathbb{R}]$，令 $(e_2)^a = (\partial/\partial t)^a$，选

$$(e_1)^a = \alpha(\partial/\partial t)^a + \beta(\partial/\partial x)^a, (\alpha, \beta \text{ 可为任意实数，但 } \beta \neq 0。)$$

则 $(e^1)_a = \beta^{-1}(dx)_a$。取 $(e^1)_a$ 为法余矢 n_a，相应的法矢为 $n^a = \beta^{-1}(\partial/\partial x)^a$，满足 $n^a \notin W_q$ 且 $n_a n^a > 0$（即 n^a 为类空）。

(3) $\phi[\mathbb{R}]$ 与 x 轴夹 45∘ 角（按欧氏）。$\forall q \in \phi[\mathbb{R}]$，令 $(e_2)^a = (\partial/\partial t)^a + (\partial/\partial x)^a$，选

$$(e_1)^a = \alpha(\partial/\partial t)^a + \beta(\partial/\partial x)^a, \alpha \neq \beta$$

则 $(e^1)_a = (\alpha-\beta)^{-1}[(dt)_a - (dx)_a]$。取 $(e^1)_a$ 为法余矢 n_a，相应的法矢为 $n^a = (\alpha-\beta)^{-1} g^{ab}[(dt)_b - (dx)_b] = -(\alpha-\beta)^{-1}[(\partial/\partial t)^a + (\partial/\partial x)^a] = -(\alpha-\beta)^{-1}(e_2)^a$，满足 $n^a \in W_q$ 且 $n_a n^a = 0$（即 n^a 为类光）。这种情况下，超曲面的法矢既与面上所有切矢垂直（法矢定义），本身又是切矢之一！

定义 4.11. 超曲面叫**类空**的，若其法矢处处类时（$n^a n_a < 0$）；超曲面叫**类时**的，若其法矢处处类空（$n^a n_a > 0$）；超曲面叫**类光**的，若其法矢处处类光（$n^a n_a = 0$）。

若 $n^a n_a \neq 0$，今后谈法矢时都指归一化法矢，即 $n^a n_a = \pm 1$。

定义 4.12. 设 $\phi[S]$ 是流形 M 中的嵌入子流形（不一定是超曲面），$q \in \phi[S], W_q$ 是 q 点切于 $\phi[S]$ 的切空间。W_q 的度规 h_{ab} 叫做由 V_q 的度规 g_{ab} 生出的诱导度规，若

$$h_{ab} w_1^a w_2^b = g_{ab} w_1^a w_2^b, \ \forall w_1^a, w_2^b \in W_q$$

诱导度规 h_{ab} 实质上是把 V_q 上度规 g_{ab} 的作用对象限制于 W_q 的结果。当 $\phi[S]$ 为类时或类空超曲面时，诱导度规 h_{ab} 可用其归一化法矢（$n^a n_a = 1$）方便地表示为

$$h_{ab} \equiv g_{ab} \mp n_a n_b, (n^a n_a = +1 \text{ 时取 } -，n^a n_a = -1 \text{ 时取 } +)$$

因为 $\forall w_1^a, w_2^b \in W_q$ 有 $h_{ab} w_1^a w_2^b = g_{ab} w_1^a w_2^b \mp n_a w_1^a n_b w_2^b = g_{ab} w_1^a w_2^b$，即满足定义。

设 $\phi[S]$ 为类时或类空超曲面，$q \in \phi[S]$，h_{ab} 满足前式。令

$$h^a{}_b \equiv g^{ac} h_{cb} = \delta^a{}_b \mp n^a n_b$$

则 $\forall v^a \in V_q$ 有 $h^a{}_b v^b = v^a \mp n^a(n_b v^b)$，或

$$v^a = h^a{}_b v^b \pm n^a(n_b v^b)$$

上式代表矢量 v^a 的一种分解, 其中 $\pm n^a(n_b v^b)$ 与法矢 n^a 平行，称为法向分量, $h^a{}_b v^b$ 与法矢 n^a 垂直（因为 $n_a(h^a{}_b v^b) = 0$）, 称为切向分量（切于 $\phi[S]$ 的分量）。$h^a{}_b$ 称为从 V_q 到 W_q 的**投影映射**。

定理 4.18. 类光超曲面上的诱导"度规"是退化的（因而没有诱导度规）。

证明. 以 h_{ab} 代表诱导"度规"。超曲面的类光性导致 $n^a \in W_q$, 故 W_q 有非零元素 n^a 使 $h_{ab} n^a w^b = g_{ab} n^a w^b = 0$, $\forall w^a \in W_q$。可见 h_{ab} 是 W_q 上的退化张量。 □

例 4.6. 设 t, x, y, z 是 4 维闵氏空间 $(\mathbb{R}^4, \eta_{ab})$ 的洛伦兹坐标, 则 η_{ab} 可用对偶坐标基矢表为

$$\eta_{ab} = \eta_{\mu\nu}(dx^\mu)_a(dx^\nu)_b = -(dt)_a(dt)_b + (dx)_a(dx)_b + (dy)_a(dy)_b + (dz)_a(dz)_b$$

方程 $t - (x^2 + y^2 + z^2)^{1/2} = 0$ 定义了一个类光超曲面 $\mathscr{S}$, 它是以原点为锥顶的圆锥面。$\forall q \in \mathscr{S} \subset \mathbb{R}^4$，有 4 维切空间 V_q 和 3 维切空间（切于 $\mathscr{S}$）$W_q \subset V_q$。令

$$n^a|_q \equiv (\partial/\partial t)^a|_q + (\partial/\partial z)^a|_q(\text{以下略去下标 } q)$$

则 n^a 是 q 点的类光法矢, 故 $n^a \in W_q$, 因而 $\{(\partial/\partial x)^a, (\partial/\partial y)^a, n^a\}$ 是 W_q 的基底。现在计算 η_{ab} 在 W_q 上的诱导"度规"h_{ab} 在此基底的分量 $h_{\mu\nu}$。

$$h_{11} = h_{ab}(\partial/\partial x)^a(\partial/\partial x)^b = \eta_{ab}(\partial/\partial x)^a(\partial/\partial x)^b = 1$$

类似地，有 $h_{22} = h_{ab}(\partial/\partial y)^a(\partial/\partial x)^b = 1$。而 $h_{\mu\nu}$ 的第三个对角元（记作 h_{nn}）则为

$$h_{nn} = h_{ab}n^a n^b = \eta_{ab}[(\partial/\partial t)^a + (\partial/\partial z)^a][(\partial/\partial t)^b + (\partial/\partial z)^b] = 1 - 1 = 0$$

而且容易验证 $h_{\mu\nu}$ 的所有非对角元为零，故

$$(h_{\mu\nu}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

因而 h_{ab} 退化（也说其"号差"为 $(+,+,0)$）。可见 η_{ab} 在类光超曲面 $\mathscr{S}$ 上无诱导度规。然而, 令 S 为 $\mathscr{S}$ 与任一等 t 面（$t > 0$）之交（是 2 维球面），以 $\hat{W}_q \subset W_q$ 代表 W_q 中所有切于 S 的元素组成的子空间, 则 η_{ab} 在 $\hat{W}_q$ 却有诱导度规, 记作 $\hat{h}_{ab}$，而且不难验证

$$\hat{h}_{ab} = (dx)_a(dx)_b + (dy)_a(dy)_b$$

读者不难对 $(\mathbb{R}^4, \eta_{ab})$ 中由 t–$z = 0$ 定义的类光超曲面做类似讨论。

第五章　微分形式及其积分

5.1　微分形式

先介绍 n 维矢量空间 V 上的“形式”，再讨论 n 维流形 M 上的“微分形式场”。

定义 5.1. $\omega_{a_1\cdots a_l} \in \mathscr{T}_V(0,l)$ 叫 V 上的**l 次形式**（简称**l 形式**），若

$$\omega_{a_1\cdots a_l} = \omega_{[a_1\cdots a_l]}$$

为书写方便，有时略去下标而把 l 形式 $\omega_{a_1\cdots a_l}$ 写为 $\boldsymbol{\omega}$。

定理 5.1.（a）　$\omega_{a_1\cdots a_l} = \omega_{[a_1\cdots a_l]} \Rightarrow$ 对任意基底有 $\omega_{\mu_1\cdots\mu_l} = \omega_{[\mu_1\cdots\mu_l]}$

（b）　存在基底使 $\omega_{\mu_1\cdots\mu_l} = \omega_{[\mu_1\cdots\mu_l]} \Rightarrow \omega_{a_1\cdots a_l} = \omega_{[a_1\cdots a_l]}$

定理 5.2.（a）　$\omega_{a_1\cdots a_l} = \delta_\pi \omega_{a_{\pi(1)}\cdots a_{\pi(l)}}$

（b）　对任意基底 $\omega_{\mu_1\cdots\mu_l} = \delta_\pi \omega_{\mu_{\pi(1)}\cdots\mu_{\pi(l)}}$

其中 π 代表 $(1,\cdots,l)$ 的一种排列，$\pi(1)$ 是指 π 所代表的那种排列中的第 1 个数字，$\delta_\pi \equiv \pm 1$（偶排列取 $+$，奇排列取 $-$）。

由上可知，在 l 形式的分量 $\omega_{\mu_1\cdots\mu_l}$ 中，凡有重复具体指标者必为零，例如

$$\omega_{112} = \omega_{133} = \omega_{212} = 0$$

V 上全体 l 形式的集合记作 $\Lambda(l)$。1 形式其实就是 V 上的对偶矢量，故 $\Lambda(1) = V^*$。约定把任一实数称为 V 上的 **0 形式**，则 $\Lambda(0) = \mathbb{R}$。l 形式既然是 $(0,l)$ 型张量，自然有 $\Lambda(l) \subset \mathscr{T}_V(0,l)$。不但如此，还容易证明 $\Lambda(l)$ 是 $\mathscr{T}_V(0,l)$ 的线性子空间，其维数的计算可从关于 $\mathscr{T}_V(k,l)$ 维数的计算得到启发：为求 $\mathscr{T}_V(k,l)$ 的维数可先找一个基底，而为此则要先定义张量积。然而两个微分形式（作为两个张量）的张量积并非全反称，故不再是微分形式。但可对全体指标施行全反称操作使之成为微分形式，于是有如下定义

定义 5.2. 设 $\boldsymbol{\omega}$ 和 $\boldsymbol{\mu}$ 分别为 l 形式和 m 形式，则其**楔形积**（简称**楔积**）是按下式定义的 $l+m$ 形式：

$$(\omega \wedge \mu)_{a_1\cdots a_l b_1\cdots b_m} \coloneqq \frac{(l+m)!}{l!m!} \omega_{[a_1\cdots a_l} \mu_{b_1\cdots b_m]}$$

或者说，楔积是满足上式的映射 $\wedge\colon \Lambda(l)\times\Lambda(m)\to\Lambda(l+m)$

楔积 $(\omega\wedge\mu)_{a_1\cdots a_l b_1\cdots b_m}$ 亦可记作 $\omega_{a_1\cdots a_l}\wedge\mu_{b_1\cdots b_m}$，也常简记为 $\boldsymbol{\omega}\wedge\boldsymbol{\mu}$。

由定义可知楔积满足结合律和分配律，即 $(\boldsymbol{\omega}\wedge\boldsymbol{\mu})\wedge\boldsymbol{\nu}=\boldsymbol{\omega}\wedge(\boldsymbol{\mu}\wedge\boldsymbol{\nu})$（因而 $\boldsymbol{\omega}\wedge\boldsymbol{\mu}\wedge\boldsymbol{\nu}$ 有明确意义）和 $\boldsymbol{\omega}\wedge(\boldsymbol{\mu}+\boldsymbol{\nu})=\boldsymbol{\omega}\wedge\boldsymbol{\mu}+\boldsymbol{\omega}\wedge\boldsymbol{\nu}$。但楔积一般不服从交换律，例如对 1 形式 $\boldsymbol{\omega}$ 和 $\boldsymbol{\mu}$ 有

$$\boldsymbol{\omega}\wedge\boldsymbol{\mu}\equiv\omega_a\wedge\mu_b\equiv(\omega\wedge\mu)_{ab}=2\omega_{[a}\mu_{b]}=\omega_a\mu_b-\omega_b\mu_a$$

$$\boldsymbol{\mu}\wedge\boldsymbol{\omega}\equiv(\mu\wedge\omega)_{ab}=2\mu_{[a}\omega_{b]}=\mu_a\omega_b-\mu_b\omega_a$$

可见对两个 1 形式的楔积有 $\boldsymbol{\omega}\wedge\boldsymbol{\mu}=-\boldsymbol{\mu}\wedge\boldsymbol{\omega}$。推广至一般情况，设 $\boldsymbol{\omega}$ 和 $\boldsymbol{\mu}$ 分别是 l 和 m 形式，则

$$\boldsymbol{\omega}\wedge\boldsymbol{\mu}=(-1)^{lm}\boldsymbol{\mu}\wedge\boldsymbol{\omega}$$

定理 5.3. 设 $\dim V=n$，则 $\dim\Lambda(l)=C_n^l=\frac{n!}{l!(n-l)!}$，若 $l\leqslant n$；$\Lambda(l)=\{0\}$（只有零元），若 $l>m$。

下面回到流形 M 上来。若对 M（或 $A\subset M$）的任一点 p 指定 V_p 上的一个 l 形式，就得到 M（或 A）上的一个 l 形式场（“场”字常略去）。1 形式场和 0 形式场分别是对偶矢量场和标量场。M 上的光滑的 l 形式场称为 l **次微分形式场**，也简称作 l **形式场**或 l **形式**。

设 (O,ψ) 为一坐标系，则 O 上的 l 形式场可方便地用对偶坐标基底场 $\{(dx^\mu)_a\}$ 逐点线性表出。

$$\omega_{a_1\cdots a_l}=\sum_C\omega_{\mu_1\cdots\mu_l}(dx^{\mu_1})_{a_1}\wedge\cdots\wedge(dx^{\mu_l})_{a_l}$$

其中

$$\omega_{\mu_1\cdots\mu_l}=\omega_{a_1\cdots a_l}(\partial/\partial x^{\mu_1})^{a_1}\cdots(\partial/\partial x^{\mu_l})^{a_l}$$

是 O 上的函数。一个重要的特例是 $l=n$ 的情况。因为 $C_n^l=C_n^n=1$，上式右边的求和只有一项，即

$$\omega_{a_1\cdots a_n}=\omega_{1\cdots n}(dx^1)_{a_1}\wedge\cdots\wedge(dx^n)_{a_n}$$

简写为

$$\boldsymbol{\omega}=\omega_{1\cdots n}dx^1\wedge\cdots\wedge dx^n$$

上式也可以这样理解：M 中任一点 p 的所有 n 形式的集合是 1 维矢量空间，只有一个基矢，可取为 $dx^1\wedge\cdots\wedge dx^n|_p$，上式就是 $\boldsymbol{\omega}|_p$ 用这一基矢的展开式。注意，展开系数 $\omega_{1\cdots n}$ 对不同点可以不同，因而是坐标域上的函数，也可表为坐标的 n 元函数 $\omega_{1\cdots n}(x^1,\cdots,x^n)$。

我们以 $\Lambda_M(l)$ 代表 M 上全体 l 形式场的集合。

定义 5.3. 流形 M 上的**外微分算符**是一个映射 $d\colon\Lambda_M(l)\to\Lambda_M(l+1)$，定义为

$$(d\omega)_{ba_1\cdots a_l}:=(l+1)\nabla_{[b}\omega_{a_1\cdots a_l]}$$

其中 ∇_b 为任一导数算符（因由 $C^c{}_{ab} = C^c{}_{ba}$ 可证对任意 ∇_a 和 $\tilde{\nabla}_a$ 有 $\tilde{\nabla}_{[b}\omega_{\dots]} = \nabla_{[b}\omega_{\dots]}$）。可见在定义外微分之前无须在 M 上指定导数算符（及任何其他附加结构，如度规）。

例 5.1. 第二章曾定义过 $(df)_a$，第三章又知 $(df)_a = \nabla_a f$，可见 $(df)_a$ 就是 $f \in \Lambda_M(0)$ 的外微分，这正是当时用符号 df 的原因。

把 l 形式场 ω 写成对偶坐标基矢的展开式的一个好处是便于计算 $d\omega$，请看如下定理

定理 5.4. 设 $\omega_{a_1\cdots a_l} = \sum\limits_C \omega_{\mu_1\cdots\mu_l}(dx^{\mu_1})_{a_1} \wedge \cdots \wedge (dx^{\mu_l})_{a_l}$，则

$$(d\omega)_{ba_1\cdots a_l} = \sum_C (d\omega_{\mu_1\cdots\mu_l})_b \wedge (dx^{\mu_1})_{a_1} \wedge \cdots \wedge (dx^{\mu_l})_{a_l}$$

证明. 选该系的普通导数算符 ∂_b 作为 ∇_b，结合 $(d\omega)_{ba_1\cdots a_l}$ 的定义即可得证。 □

定理 5.5. $d \circ d = 0$

证明. 选任一坐标系的导数算符 ∂_b 作为 ∇_b，便有

$$[d(d\omega)]_{cba_1\cdots a_l} = (l+2)(l+1)\partial_{[c}\partial_{[b}\omega_{a_1\cdots a_l]]} = (l+2)(l+1)\partial_{[[c}\partial_{b]}\omega_{a_1\cdots a_l]} = 0$$

其中第二步是因为同种括号内的子括号可以随意增删，最后一步是因为 $\partial_{[a}\partial_{b]}T^{\cdots}{}_{\cdots} = 0$。 □

定义 5.4. 设 ω 为 M 上的 l 形式场。ω 叫**闭的**，若 $d\omega = 0$；ω 叫**恰当的**，若存在 $l-1$ 形式场 μ 使 $\omega = d\mu$。

因此，前述定理亦可表述为：若 ω 是恰当的，则 ω 是闭的。然而，要使逆命题成立则还须对流形 M 提出一定要求（略）。平凡流形 $\mathbb{R}^n$ 满足这一要求，而流形一定局域平凡，所以对任意流形 M 而言，闭的 l 形式场至少是局域恰当的。就是说，设 ω 是流形 M 上的闭的 l 形式场，则 M 的任一点 p 必有邻域 N，在 N 上存在 $l-1$ 形式场 μ 使 $\omega = d\mu$。

定理 5.6. 当 $M = \mathbb{R}^2$ 时，上述定理及其逆定理给出普通微积分的下述命题：给定函数 $X(x,y)$ 及 $Y(x,y)$，存在函数 $f(x,y)$ 使 $df = Xdx + Ydy$ 的充要条件是 $\partial X/\partial y = \partial Y/\partial x$。

证明. 由定义知 1 形式场 $Xdx + Ydy$ 的外微分为

$$\begin{aligned} d(Xdx + Ydy) &= dX \wedge dx + dY \wedge dy = (\frac{\partial X}{\partial x}dx + \frac{\partial X}{\partial y}dy) \wedge dx + (\frac{\partial Y}{\partial x}dx + \frac{\partial Y}{\partial y}dy) \wedge dy \\ &= \frac{\partial X}{\partial y}dy \wedge dx + \frac{\partial Y}{\partial x}dx \wedge dy = (\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y})dx \wedge dy \end{aligned}$$

（A） 若存在函数 f 使 1 形式等式 $df = Xdx + Ydy$ 成立，则

$$(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y})dx \wedge dy = ddf = 0$$

故 $\partial X/\partial y = \partial Y/\partial x$。

（B） 若 $\partial X/\partial y = \partial Y/\partial x$，则 $d(Xdx + Ydy) = 0$，即 1 形式场 $Xdx + Ydy$ 为闭，于是 $Xdx + Ydy$ 为恰当，即存在函数 f 使 $df = Xdx + Ydy$。

□

5.2 流形上的积分

先以 3 维欧氏空间 $(\mathbb{R}^3, \delta^a{}_b)$ 为例。设 $\vec{v}$ 为矢量场，L 为光滑曲线，S 为光滑曲面。在指定 L 的方向和 S 的法向之前，积分 $\int_L \vec{v}\cdot d\vec{l}$ 和 $\iint\limits_S \vec{v}\cdot d\vec{S}$ 都只唯一确定到差一个负号的程度。要完全确定这两个积分，就要指定 L 的方向和 S 的法向。推而广之，计算任意流形上的积分之前应指定该流形的“定向”。然而，并非所有流形都是可定向的。

定义 5.5. n 维流形称为**可定向的**，若其上存在 C^0 且处处非零的 n 形式场 ε。

例 5.2. $\mathbb{R}^3$ 是可定向流形，因为其上存在 C^∞ 的 3 形式场 $\varepsilon \equiv dx \wedge dy \wedge dz$，其中 x, y, z 为自然坐标。

例 5.3. 莫比乌斯带是不可定向流形。

定义 5.6. 若在 n 维可定向流形 M 上选定一个 C^0 且处处非零的 n 形式场 ε，就说 M 是**定向的**（“已经定向”之意）。设 ε_1 和 ε_2 是两个 C^0 且处处非零的 n 形式场，若存在处处为正的函数 h 使 $\varepsilon_1 = h\varepsilon_2$，就说 ε_1 和 ε_2 给出 M 的**同一个定向**。

注 5.1. 从给出 M 的定向这个角度看，满足 $\varepsilon_1 = h\varepsilon_2$ $(h > 0)$ 的 ε_1 和 ε_2 是等价的。由于 n 维流形 M 上每点的全体 n 形式的集合是 1 维矢量空间，任意两个 n 形式场 ε_1 和 ε_2 必有关系 $\varepsilon_1 = h\varepsilon_2$ 其中 h 是 M 上的函数。若 ε_1 和 ε_2 处处非零，则 h 处处非零；若 ε_1 和 ε_2 为 C^0 的，则 h 为 C^0 的。对连通流形[1]来说（我们只讨论连通流形），一个处处非零的连续函数只能处处为正或处处为负。可见连通流形只能有两种定向。

定义 5.7. M 上选好以 ε 为代表的定向后，开域 $O \subset M$ 上的基底场 $\{(e_\mu)^a\}$ 叫做以 ε 衡量为**右手的**，若 O 上存在处处为正的函数 h 使 $\varepsilon = h(e^1)_{a_1} \wedge \cdots \wedge (e^n)_{a_n}$，其中 $\{(e^\mu)_a\}$ 是 $\{(e_\mu)^a\}$ 的对偶基（否则称为**左手的**）。一个坐标系叫**右（左）手系**，若其坐标基底是右（左）手的。

下面介绍 n 维定向流形 M 上的 n 形式场 ω 的积分。ω 可用对偶坐标基矢的楔形积 $dx^1 \wedge \cdots \wedge dx^n$ 展开为

$$\omega = \omega_{1\cdots n}(x^1, \cdots, x^n)dx^1 \wedge \cdots \wedge dx^n$$

[1]拓扑空间 $(X, \mathscr{T})$ 称为**连通的**，若它只有两个既开又闭的子集；称为**弧连通的**，若 X 的任意两点可被一条在 X 中的连续曲线连接。流形称为连通的（或弧连通的），若其底拓扑空间是连通的（或弧连通的）。对拓扑空间，弧连通必定连通，但连通不一定弧连通（存在“擦边性”反例）。对流形，弧连通与连通等价。

可见每一 n 形式场 ω 在坐标域上给出一个 n 元函数 $\omega_{1\cdots n}(x^1,\cdots,x^n)$，我们就把这个 n 元函数的普通 n 重积分称为 n 形式场 ω 的积分，准确定义如下：

定义 5.8. 设 (O,ψ) 是 n 维定向流形 M 上的右手坐标系，ω 是开子集 $G\subset O$ 上的连续 n 形式场，则 ω 在 G 上的**积分**定义为

$$\int_G \omega \coloneqq \int_{\psi[G]} \omega_{1\cdots n}(x^1,\cdots,x^n)dx^1\cdots dx^n$$

上式右边是 n 元函数 $\omega_{1\cdots n}(x^1,\cdots,x^n)$ 在 $\mathbb{R}^n$ 的开子集 $\psi[G]$ 上的普通积分[2]，早已有定义。

注 5.2.（*1*） 为说明上述定义的合理性，还应证明 ω 在 G 上的积分与所选右手坐标系无关。仅以 $n=2$ 为例证明如下。

设 (O,ψ) 和 (O',ψ') 为右手坐标系，满足 $G\subset O\cap O'$，两坐标分别记作 x^1,x^2 和 x'^1,x'^2，则

$$\omega=\omega_{12}dx^1\wedge dx^2=\omega'_{12}dx'^1\wedge dx'^2$$

令 $\int_G\omega\equiv\int_{\psi[G]}\omega_{12}dx^1dx^2$，$(\int_G\omega)'\equiv\int_{\psi'[G]}\omega'_{12}dx'^1dx'^2$，欲证

$$(\int_G\omega)'=\int_G\omega$$

由张量变换律知 $\omega'_{12}=\frac{\partial x^1}{\partial x'^1}\frac{\partial x^2}{\partial x'^2}\omega_{12}+\frac{\partial x^2}{\partial x'^1}\frac{\partial x^1}{\partial x'^2}\omega_{21}=\omega_{12}\det(\frac{\partial x^\mu}{\partial x'^\nu})$，其中

$$\det(\frac{\partial x^\mu}{\partial x'^\nu})\equiv\begin{vmatrix}\frac{\partial x^1}{\partial x'^1} & \frac{\partial x^1}{\partial x'^2}\\ \frac{\partial x^2}{\partial x'^1} & \frac{\partial x^2}{\partial x'^2}\end{vmatrix}$$

[2]指 Riemann 或 Lebesgue 积分

是这个变换的雅克比行列式。根据二重积分的换元公式[3],

$$\int_{\psi[G]} \omega_{12} dx^1 dx^2 = \int_{\psi'[G]} \omega_{12} \det(\partial x^\mu / \partial x'^\nu) dx'^1 dx'^2 = \int_{\psi'[G]} \omega'_{12} dx'^1 dx'^2$$

故得证。

然而，如果 $\{x^\mu\}$ 和 $\{x'^\mu\}$ 分别是右、左手系，则 $\det(\partial x^\mu/\partial x'^\nu) < 0$，等号右边的 $\det(\partial x^\mu/\partial x'^\nu)$ 应改为 $|\det(\partial x^\mu/\partial x'^\nu)| = -\det(\partial x^\mu/\partial x'^\nu)$, 故上式变为

$$\int_{\psi[G]} \omega_{12} dx^1 dx^2 = -\int_{\psi'[G]} \omega_{12} \det(\partial x^\mu / \partial x'^\nu) dx'^1 dx'^2 = -\int_{\psi'[G]} \omega'_{12} dx'^1 dx'^2$$

因此，为了定义出同一积分，当 $\{x^\mu\}$ 是左手系时应把 $\int_G \omega$ 定义为

$$\int_G \omega := -\int_{\psi[G]} \omega_{1\cdots n}(x^1, \cdots, x^n) dx^1 \cdots dx^n$$

(2) 一个坐标系是右手系还是左手系取决于流形所选的定向，故 $\int_G \omega$ 是依赖于由 ε 所给出的定向的，定向改变后积分变号。

(3) 以上之定义了 ω 在坐标域内开子集 G 上的积分。ω 在全流形 M 上的积分 $\int_M \omega$ 可由局部积分“缝合”而成，其定义涉及“单位分解”, 从略。

设 S, M 是流形，维数分别是 l 和 $n(> l)$，$\phi\colon S \to M$ 是嵌入。因为 $\phi[S]$ 是子流形，当然可谈及其上的 l 形式场 μ 的积分。然而 “$\phi[S]$ 嵌入在 M 内” 的事实导致 “$\phi[S]$ 上的 l 形式场” 具有两种可能的含义。正如 “$\phi[S]$ 上的切矢场” 有切于和不切于 $\phi[S]$ 之分那样，“$\phi[S]$ 上的 l 形式场” 也可分为 “切于” 和不 “切于”$\phi[S]$ 两种。准确地说，$\phi[S]$ 上的 l 形式场 μ 称为 “切于”$\phi[S]$ 的，如果 $\forall q \in \phi[S]$，$\mu|_q$ 是 W_q（而非 V_q）上的 l 形式（能把 W_q 的任意 l 个元素变为一个实数的线性映射）。“$\phi[S]$ 上的 l 形式场” 既可能是 “切于”$\phi[S]$ 的，也可能不是 “切于”$\phi[S]$ 的。因为谈及 l 形式场在 $\phi[S]$ 上的积分时是把 $\phi[S]$ 作为独立流形

[3] 设 $f(x, y)$ 在 xOy 平面上的闭区域 D 上连续，变换

$$T\colon x = x(u, v),\ y = y(u, v)$$

将 uOv 平面上的闭区域 D' 变为 xOy 平面上的 D, 且满足

(1) $x(u, v)$, $y(u, v)$ 在 D' 上具有一阶连续偏导数

(2) 在 D' 上雅克比行列式

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)} \neq 0$$

(3) 变换 $T\colon D' \to D$ 是一一到上的

则有

$$\iint_D f(x, y) dx dy = \iint_{D'} f[x(u, v), y(u, v)] |J(u, v)| du dv$$

看待的（不顾及它“外面”的情况），所以只有“切于”$\phi[S]$ 的 l 形式场 μ 的积分才有意义。不过，既然 $\phi[S]$ 上的、不“切于”$\phi[S]$ 的 l 形式场 μ 是能把每点 $q \in \phi[S]$ 的 V_q（而不只是 W_q）的任意 l 个元素变为一个实数的线性映射，而 W_q 无非是 V_q 的子空间，只要把 μ 的作用范围限制在 W_q 便得到一个“切于”$\phi[S]$ 的 l 形式场，我们把它记作 $\tilde{\mu}$, 并称之为 μ 的限制。准确说来有如下定义

定义 5.9. 设 $\mu_{a_1\cdots a_l}$ 是 l 维子流形 $\phi[S] \subset M$ 上的 l 形式场。$\phi[S]$（看作脱离 M 而独立存在的流形）上的 l 形式场 $\tilde{\mu}_{a_1\cdots a_l}$ 称为 $\mu_{a_1\cdots a_l}$ 在 $\phi[S]$ 上的**限制**，若

$$\tilde{\mu}_{a_1\cdots a_l}|_q(\omega_1)^{a_1}\cdots(\omega_l)^{a_l} = \mu_{a_1\cdots a_l}|_q(\omega_1)^{a_1}\cdots(\omega_l)^{a_l},\ \forall q \in \phi[S],\ (\omega_1)^a,\cdots(\omega_l)^a \in W_q$$

今后凡谈及 l 形式场 μ 在 l 维子流形 $\phi[S]$ 上的积分时，一律理解为 μ 的限制 $\tilde{\mu}$ 的积分，即总把 $\int_{\phi[S]}\mu$ 理解为 $\int_{\phi[S]}\tilde{\mu}$。

5.3 Stokes 定理

3 维欧氏空间的 Stokes 定理

$$\iint\limits_S (\vec{\nabla}\times\vec{A})\cdot d\vec{S} = \oint\limits_L \vec{A}\cdot d\vec{l}$$

和 Gauss 定理

$$\iiint\limits_V (\vec{\nabla}\cdot\vec{A})dV = \oiint\limits_S \vec{A}\cdot\vec{n}dS$$

的共性是反映区域上的积分和它的边界上的积分的联系。在介绍一般的 Stokes 定理前，先引入“带边流形”的概念。n 维带边流形的最简单例子是

$$\mathbb{R}^{n-} \coloneqq \{(x^1,\cdots,x^n)\in\mathbb{R}^n \mid x^1 \leqslant 0\}$$

其中 $x^1,\cdots,x^n$ 是自然坐标，由 $x^1 = 0$ 的所有点组成的子集叫 $\mathbb{R}^{n-}$ 的边界，它本身是个 $n-1$ 维流形（其实就是 $\mathbb{R}^{n-1}$）。推广至一般情况，n **维带边流形** N 与 n 维流形定义相仿，只是把该定义中的 $\mathbb{R}^n$ 改为 $\mathbb{R}^{n-}$，即 N 的开覆盖 O_α 中的每一元素 O_α 都应同胚于 $\mathbb{R}^{n-}$ 的一个子集，N 中全体被映射到 $x^1 = 0$ 处的点组成 N 的**边界**，记作 ∂N。请注意 ∂N 是 $n-1$ 维流形；$i(N) \equiv N - \partial N$ 是 n 维流形。例如，$\mathbb{R}^3$ 中的实心球体 B 是 3 维带边流形，其边界（2 维球面 S^2）是 2 维流形，$i(B)$ 则是 3 维流形。

定理 5.7. 设 n 维定向流形 M 的紧致子集 N 是个 n 维带边流形，ω 是 M 上的 $n-1$ 形式场（可微性至少为 C^1），则

$$\int_{i(N)} d\boldsymbol{\omega} = \int_{\partial N}\boldsymbol{\omega}$$

注 5.3. 把 M 的定向 ε 限制在 N 上便得到 N 的定向，仍记作 ε, 它在 N 的边界 ∂N（M 中的超曲面）上自然诱导出一个定向，记作 $\bar{\varepsilon}$, 是 $\bar{\varepsilon}_{a_1\cdots a_{n-1}}$ 的简写。仅以 $\mathbb{R}^{2-}$ 为例介绍，这时 $M=\mathbb{R}^2$, $N=\mathbb{R}^{2-}$, $\partial N=\{(x^1,x^2)\mid x^1=0\}$。设 $\mathbb{R}^2$（因而 $\mathbb{R}^{2-}$）的定向为 $\varepsilon_{ab}=(dx^1)_a\wedge(dx^2)_b$，则 $\{x^1,x^2\}$ 以 ε_{ab} 衡量为右手系。因 $x^1|_{\partial N}=0$，把 x^1 开除后所得的 $\{x^2\}$ 便是 ∂N 的一个坐标系。我们这样定义 ∂N 的诱导定向 $\bar{\varepsilon}_a$，使坐标系 $\{x^2\}$ 以 $\bar{\varepsilon}_a$ 衡量为右手系。选 $\bar{\varepsilon}_a=(dx^2)_a$ 便可满足这一要求。诱导定向的这个基本要求可以推广到任意带边流形 N。上式左边是 n 形式场 $d\omega$ 在 n 维流形 $i(N)$（以 ε 为定向）上的积分，右边是 $n-1$ 形式场 ω 在 $n-1$ 维流形 ∂N（以 $\bar{\varepsilon}$ 为定向）上的积分。

例 5.4. 设 $\vec{A}$ 是 2 维欧氏空间的矢量场，L 是 $\mathbb{R}^2$ 中的光滑闭合曲线，S 是由 L 包围的开子集，x^1,x^2 为笛卡尔坐标，则熟知的 2 维欧氏空间 *Stokes* 定理（又称 *Green* 定理）为

$$\iint\limits_S(\partial A_2/\partial x^1-\partial A_1/\partial x^2)dx^1dx^2=\oint\limits_L A_l dl$$

现在说明上式是前述定理的特例。令 $M=\mathbb{R}^2$，则 $S\cup L$ 可认为是 N，其中 S 和 L 分别对应 $i(N)$ 和 ∂N。用欧氏度规 δ_{ab} 把 A^a 变为 1 形式场（即对偶矢量）$A_a\equiv\delta_{ab}A^b$，则 A_a 对应 $\boldsymbol{\omega}$。把 A_a 用笛卡尔系对偶坐标基矢展开：$\boldsymbol{\omega}=A_a=A_\mu(dx^\mu)_a$，则

$$d\boldsymbol{\omega}=dA_\mu\wedge dx^\mu=\frac{\partial A_\mu}{\partial x^\nu}dx^\nu\wedge dx^\mu=\frac{\partial A_1}{\partial x^2}dx^2\wedge dx^1+\frac{\partial A_2}{\partial x^1}dx^1\wedge dx^2=(\frac{\partial A_2}{\partial x^1}-\frac{\partial A_1}{\partial x^2})dx^1\wedge dx^2$$

其中第二个等号是 df 用对偶坐标基矢的展开式，可见上式左边可以表示为 $\int_{i(N)}d\boldsymbol{\omega}$。另一方面，$\int_{\partial N}\boldsymbol{\omega}=\int_{\partial N}\tilde{\boldsymbol{\omega}}$。选线长 l 为 L 的局部坐标，把 $\tilde{\boldsymbol{\omega}}$ 用坐标基矢展开为 $\tilde{\omega}_a=\tilde{\omega}_1(l)(dl)_a$，两边与 $(\partial/\partial l)^a$ 缩并得

$$\tilde{\omega}_1(l)=\tilde{\omega}_a(\partial/\partial l)^a=\omega_a(\partial/\partial l)^a=A_a(\partial/\partial l)^a=A_l$$

故 $\tilde{\boldsymbol{\omega}}=A_l dl$，于是

$$\oint\limits_L A_l dl=\int_{\partial N}\boldsymbol{\omega}$$

可见，上式是 *Stokes* 定理的特例。

以上介绍了微分形式在流形上的积分及有关定理。为了讲解函数在流形上的积分，先介绍体元的概念。

5.4 体元

定义 5.10. n 维可定向流形 M 上的任一个 C^0 而且处处非零的 n 形式场 ε 称为一个**体元**。

注 5.4. 体元与定向的区别在于：若 ε_1 和 ε_2 是两个 C^0 且处处非零的 n 形式场，而且有处处为正的函数 h 使 $\varepsilon_1 = h\varepsilon_2$，则 ε_1 和 ε_2 代表同一个定向，但只要 $\varepsilon_1 \neq \varepsilon_2$ 它们就是两个不同体元。对可定向连通流形，定向只有两个，而体元却有无限多个。谈及定向流形上的积分和体元时不要求流形上选定度规场，这时体元的选择十分任意（只有一个要求，就是体元与定向相容，即代表体元的 ε 与代表定向的 ε 之间的乘子为正。），没有一个与众不同的体元。然而，如果流形上给定了度规场 g_{ab}，便存在选择特定体元的自然方法。

先考虑带度规 g_{ab} 的 2 维定向流形。设 $\varepsilon_{a_1a_2}$ 为任一体元，则 $\varepsilon^{a_1a_2} \equiv g^{a_1b_1}g^{a_2b_2}\varepsilon_{b_1b_2}$ 有意义，且 $\varepsilon^{a_1a_2}\varepsilon_{a_1a_2}$ 是标量场，可借任一基底计算。我们用正交归一基底。若 g_{ab} 为正定度规，则

$$\varepsilon^{a_1a_2}\varepsilon_{a_1a_2} = \delta^{\mu_1\nu_1}\delta^{\mu_2\nu_2}\varepsilon_{\nu_1\nu_2}\varepsilon_{\mu_1\mu_2} = \delta^{11}\delta^{22}\varepsilon_{12}\varepsilon_{12} + \delta^{22}\delta^{11}\varepsilon_{21}\varepsilon_{21} = 2(\varepsilon_{12})^2$$

若 g_{ab} 为洛伦兹度规，则

$$\varepsilon^{a_1a_2}\varepsilon_{a_1a_2} = \eta^{11}\eta^{22}\varepsilon_{12}\varepsilon_{12} + \eta^{22}\eta^{11}\varepsilon_{21}\varepsilon_{21} = -2(\varepsilon_{12})^2$$

推广至带任意度规 g_{ab} 的 n 维流形有

$$\varepsilon^{a_1\cdots a_n}\varepsilon_{a_1\cdots a_n} = (-1)^s n!(\varepsilon_{1\cdots n})^2$$

其中 $\varepsilon_{1\cdots n}$ 是 $\varepsilon_{a_1\cdots a_n}$ 在正交归一基底的分量，s 是 g_{ab} 在正交归一基底的分量中 -1 的个数，例如正定度规有 $s = 0$，洛伦兹度规有 $s = 1$。所谓借用度规选择一个特定的体元，是指规定体元 $\varepsilon_{a_1\cdots a_n}$ 在正交归一基 $\{(e^\mu)_a\}$ 的分量满足如下的简单性要求：

$$\varepsilon_{1\cdots n} = \pm 1$$

即

$$\varepsilon_{a_1\cdots a_n} = \pm(e^1)_{a1} \wedge \cdots \wedge (e^n)_{a_n}（对正交归一基）$$

这相当于要求

$$\varepsilon^{a_1\cdots a_n}\varepsilon_{a_1\cdots a_n} = (-1)^s n!$$

满足上式的 $\varepsilon_{a_1\cdots a_n}$ 称为**与度规 g_{ab} 相适配（相容）的体元**。上式只把体元确定到差一个负号的程度，加上“体元与定向相容”的要求才确定唯一的体元。于是上式右边的 $+$ 和 $-$ 号分别对应于右手和左手正交归一基。

小结　涉及积分时，我们只关心可定向流形 M。首先选好一个定向使 M 成为定向流形。任一基底的右（左）手性由所选定向规定。没有度规场 g_{ab}（或其他可资利用的几何结构）时，体元相当任意，但要求与定向相容。指定 g_{ab} 后，体元 $\varepsilon_{a_1\cdots a_n}$ 由 g_{ab} 以及“体元与定向相容”的要求唯一确定，简称**适配体元**。今后如无特别声明，在有度规时提到体元都指这个唯一的适配体元。

在 3 维欧氏空间 $(\mathbb{R}^3, \delta_{ab})$ 中任取一个符合直观含义的右手笛卡尔系 $\{x, y, z\}$ 并指定 3 形式场 $\varepsilon = dx \wedge dy \wedge dz$ 为定向，则 $\{x, y, z\}$ 按定义就是以 ε 衡量的右手系。把 $\varepsilon = dx \wedge dy \wedge dz$ 与上式对比可知 ε 是适配体元。设 G 是 $\mathbb{R}^3$ 的开子集且普通积分 $\iiint\limits_G dxdydz$ 存在，则此积分自然代表 G 的体积（按普通微积分学的体积定义）。另一方面，3 形式场 ε 在 $G \subset \mathbb{R}^3$ 上的积分 $\int_G \varepsilon$ 正是 $\iiint\limits_G dxdydz$，可见 $\int_G \varepsilon$ 就是 G 的体积。推广至任意带正定度规 g_{ab} 的定向流形 N，设 ε 为适配体元，若 $\int_N \varepsilon$ 存在，就称它为 N 的（用 g_{ab} 衡量的）**体积**（对 1, 2 维流形又分别叫**长度**和**面积**）。这可看作把 ε 称为体元的由来。

定理 5.8. 设 ε 为适配体元，$\{(e_\mu)^a\}$ 及 $\{(e^\mu)_a\}$ 为基底及其对偶基底，g 为 g_{ab} 在此基底的分量组成的行列式，$|g|$ 为 g 的绝对值，则（式中 +、− 号分别适用于右手和左手基底）

$$\varepsilon_{a_1\cdots a_n} = \pm\sqrt{|g|}(e^1)_{a_1} \wedge \cdots \wedge (e^n)_{a_n}$$

注 5.5. 对正交归一基底有 $|g| = 1$。

定理 5.9. 设 ∇_a 和 ε 分别是与度规相适配的导数算符和体元，则

$$\nabla_b \varepsilon_{a_1\cdots a_n} = 0$$

证明. 由 $\nabla_b g_{ac} = 0$ 可得 $\varepsilon^{a_1\cdots a_n}\nabla_b \varepsilon_{a_1\cdots a_n} = 0$[4]，于是对任一矢量场 v^b 有（两边做缩并）

$$\varepsilon^{a_1\cdots a_n} v^b \nabla_b \varepsilon_{a_1\cdots a_n} = 0$$

因 M 中任一点的 n 形式的集合是 1 维矢量空间，故该点的任意两个 n 形式只能差到一个乘子 h（对不同点 h 可不同），因此 $v^b\nabla_b \varepsilon_{a_1\cdots a_n} = h\varepsilon_{a_1\cdots a_n}$。代入上式便给出 $h = 0$，所以 $v^b\nabla_b \varepsilon_{a_1\cdots a_n} = 0$。因 v^b 为任意矢量场，故 $\nabla_b \varepsilon_{a_1\cdots a_n} = 0$。 □

下面要证明关于体元的两个十分有用的等式，为此先要介绍下述引理。

定理 5.10. $\delta^{[a_1}{}_{a_1} \cdots \delta^{a_j}{}_{a_j} \delta^{a_{j+1}}{}_{b_{j+1}} \cdots \delta^{a_n]}{}_{b_n} = \dfrac{(n-j)!j!}{n!} \delta^{[a_{j+1}}{}_{b_{j+1}} \cdots \delta^{a_n]}{}_{b_n}$

根据上述引理，可以证明

定理 5.11. (*a*) $\varepsilon^{a_1\cdots a_n}\varepsilon_{b_1\cdots b_n} = (-1)^s n! \delta^{[a_1}{}_{b_1} \cdots \delta^{a_n]}{}_{b_n}$

(*b*) $\varepsilon^{a_1\cdots a_j a_{j+1}\cdots a_n}\varepsilon_{a_1\cdots a_j b_{j+1}\cdots b_n} = (-1)^s (n-j)! j! \delta^{[a_{j+1}}{}_{b_{j+1}} \cdots \delta^{a_n]}{}_{b_n}$

证明. $\varepsilon^{a_1\cdots a_n}\varepsilon_{b_1\cdots b_n} = \varepsilon^{[a_1\cdots a_n]}\varepsilon_{[b_1\cdots b_n]}$ 表明 $\varepsilon^{a_1\cdots a_n}\varepsilon_{b_1\cdots b_n}$ 对全部上标和全部下标都为反称。不难证明这种 (n, n) 型张量的集合是 1 维矢量空间，而 $\delta^{[a_1}{}_{b_1} \cdots \delta^{a_n]}{}_{b_n}$ 属于这类张量

[4] 因为 $\varepsilon^{a_1\cdots a_n}\varepsilon_{a_1\cdots a_n} = (-1)^s n!$，两边取导数得 $\nabla_b(\varepsilon^{a_1\cdots a_n}\varepsilon_{a_1\cdots a_n}) = \nabla_b \varepsilon^{a_1\cdots a_n}\varepsilon_{a_1\cdots a_n} + \varepsilon^{a_1\cdots a_n}\nabla_b \varepsilon_{a_1\cdots a_n} = 0$。又 $\nabla_b g_{ac} = 0$ 可导出 $\nabla_b g^{ac} = 0$，故前式第一项为 0。

（因为不难证明 $\delta^{[a_1}{}_{b_1}\cdots\delta^{a_n]}{}_{b_n}=\delta^{[a_1}{}_{[b_1}\cdots\delta^{a_n]}{}_{b_n]}$），故任何这类张量与 $\delta^{[a_1}{}_{b_1}\cdots\delta^{a_n]}{}_{b_n}$ 只差一个乘子，从而 $\varepsilon^{a_1\cdots a_n}\varepsilon_{b_1\cdots b_n}=K\delta^{[a_1}{}_{b_1}\cdots\delta^{a_n]}{}_{b_n}$。与 $\varepsilon_{a_1\cdots a_n}\varepsilon^{b_1\cdots b_n}$ 缩并，左边结果为 $(-1)^s n!(-1)^s n!$，右边结果为 $K\varepsilon_{b_1\cdots b_n}\varepsilon^{b_1\cdots b_n}=K(-1)^s n!$，于是 $K=(-1)^s n!$，故得式（a）。两边分别对前 j 个上、下指标缩并得

$$\begin{aligned}\varepsilon^{a_1\cdots a_j a_{j+1}\cdots a_n}\varepsilon_{a_1\cdots a_j b_{j+1}\cdots b_n}&=(-1)^s n!\delta^{[a_1}{}_{a_1}\cdots\delta^{a_j}{}_{a_j}\delta^{a_{j+1}}{}_{b_{j+1}}\cdots\delta^{a_n]}{}_{b_n}\\&=(-1)^s(n-j)!j!\delta^{[a_{j+1}}{}_{b_{j+1}}\cdots\delta^{a_n]}{}_{b_n}\end{aligned}$$

（其中最末一步用到上述引理）可见式（b）成立。 □

5.5 函数在流形上的积分，Gauss 定理

定义 5.11. 设 ε 为流形 M 上的任一体元，f 为 M 上的 C^0 函数，则**f 在 M 上的积分**（记作 $\int_M f$）定义为 n 形式场 $f\varepsilon$ 在 M 上的积分，即

$$\int_M f\coloneqq\int_M f\varepsilon$$

由定义知函数的积分与体元的选择有关。只要流形上给定度规，我们约定总是用适配体元来定义函数的积分。这样，在带度规的定向流形上，函数给定后其积分也就确定。以 3 维欧氏空间 $(\mathbb{R}^3,\delta_{ab})$ 为例。设 $\{x,y,z\}$ 为右手笛卡尔系，则 $\varepsilon=dx\wedge dy\wedge dz$ 是适配体元，于是 $(\mathbb{R}^3,\delta_{ab})$ 上的函数 $f\colon\mathbb{R}^3\to\mathbb{R}$ 的积分按定义为 $\int_{\mathbb{R}^3}f=\int_{\mathbb{R}^3}f\varepsilon$，而右边无非是 3 形式场 $\omega=f\varepsilon$ 的积分。设 f 与笛卡尔系 $\{x,y,z\}$ 结合而得的 3 元函数为 $F(x,y,z)$，则

$$\omega=F(x,y,z)dx\wedge dy\wedge dz$$

故

$$\int f=\int f\varepsilon=\int\omega=\iiint F(x,y,z)dxdydz$$

如果愿意，也可用（右手）球坐标系 $\{r,\theta,\varphi\}$ 计算。由线元式 $ds^2=dr^2+r^2(d\theta^2+\sin^2\theta d\varphi^2)$ 可知 $g=r^4\sin^2\theta$，故 $\varepsilon=r^2\sin\theta dr\wedge d\theta\wedge d\varphi$，于是 $\omega\equiv f\varepsilon=\hat{F}(r,\theta,\varphi)r^2\sin\theta dr\wedge d\theta\wedge d\varphi$（其中 $\hat{F}(r,\theta,\varphi)$ 是由 f 与 $\{r,\theta,\varphi\}$ 结合而得的 3 元函数），故

$$\int f=\int f\varepsilon=\int\omega=\iiint\hat{F}(r,\theta,\varphi)r^2\sin\theta drd\theta d\varphi$$

下面介绍一般形式的 Gauss 定理。读者熟悉的 Gauss 定理是

$$\iiint\limits_V(\vec{\nabla}\cdot\vec{A})dV=\oint\!\!\!\oint\limits_S\vec{A}\cdot\vec{n}dS$$

上式左右两边可分别形象地说成是“函数 $\vec{\nabla}\cdot\vec{A}$ 与体元 dV 的乘积的积分”和“函数 $\vec{A}\cdot\vec{n}$ 与面元（2 维体元）dS 的乘积的积分”。下面分两步证明由 Stokes 定理可以导出一个公式，它把上式作为特例包括在内。第一步是导出如下定理，其左边可看作上式左边的推广。

定理 5.12. 设 M 是 n 维定向流形，N 是 M 中的 n 维紧致带边嵌入子流形，g_{ab} 是 M 上的度规，ε 和 ∇_a 分别是适配体元和适配导数算符，v^a 是 M 上的 C^1 矢量场，则

$$\int_{i(N)}(\nabla_b v^b)\varepsilon=\int_{\partial N}v^b\varepsilon_{ba_1\cdots a_{n-1}}$$

注 5.6. 上式左边可看作上上式左边的推广。

证明. $n-1$ 形式场 $\omega\equiv v^b\varepsilon_{ba_1\cdots a_{n-1}}$ 的外微分 $d\omega=n\nabla_{[c}(v^b\varepsilon_{|b|a_1\cdots a_{n-1}]})$ 是 n 形式场，其中 ∇_c 可为任意导数算符。N 中任一点的 n 形式的集合是 1 维矢量空间，故该点的两个 n 形式 $d\omega$ 与 ε 只差一个因子，即

$$n\nabla_{[c}(v^b\varepsilon_{|b|a_1\cdots a_{n-1}]})=h\varepsilon_{ca_1\cdots a_{n-1}}$$

其中 h 是 N 上的函数，可求之如下：上式两边与 $\varepsilon^{ca_1\cdots a_{n-1}}$ 缩并，右边得 $(-1)^s hn!$，左边得

$$\begin{aligned}
&n\varepsilon^{ca_1\cdots a_{n-1}}\nabla_{[c}(v^b\varepsilon_{|b|a_1\cdots a_{n-1}]})=n\varepsilon^{[ca_1\cdots a_{n-1}]}\nabla_c(v^b\varepsilon_{ba_1\cdots a_{n-1}})\\
=&n\varepsilon^{ca_1\cdots a_{n-1}}\varepsilon_{ba_1\cdots a_{n-1}}\nabla_c v^b=n(-1)^s(n-1)!\delta^c{}_b\nabla_c v^b=(-1)^s n!\nabla_b v^b
\end{aligned}$$

（注意加入 δ 时是把 ε 的一个上标 c 换成了 b）故 $h=\nabla_b v^b$，$d\omega=\varepsilon\nabla_b v^b$。根据 Stokes 定理，上式得证。 □

下面进一步把上式右边改写为同上上式右边类似的形式。由于后者涉及边界 S 上的体元 dS，我们先从 ∂N 的体元谈起。此处只讨论 ∂N 不是类光超曲面的情况，这时可谈及 ∂N 的归一化法矢 n^a，满足 $n^a n_a=\pm1$。N 上的度规 g_{ab} 在 ∂N 上的诱导度规为 $h_{ab}=g_{ab}\mp n_a n_b$。把 ∂N 看作带度规 h_{ab} 的 $n-1$ 维流形，其体元（记作 $\hat\varepsilon_{a_1\cdots a_{n-1}}$）应满足两个条件：① 与 ∂N 的诱导定向（记作 $\bar\varepsilon_{a_1\cdots a_{n-1}}$）相容；② 与度规 h_{ab} 相适配，即

$$\hat\varepsilon^{a_1\cdots a_{n-1}}\hat\varepsilon_{a_1\cdots a_{n-1}}=(-1)^{\hat s}(n-1)!$$

其中 $\hat\varepsilon^{a_1\cdots a_{n-1}}$ 是用 h_{ab} 对 $\hat\varepsilon_{a_1\cdots a_{n-1}}$ 升指标的结果，$\hat s$ 是 h_{ab} 的对角元中负数的个数。∂N 上满足这两个条件的体元 $\hat\varepsilon_{a_1\cdots a_{n-1}}$ 称为**诱导体元**。设 n^b 是 ∂N 的外向单位法矢（以 $i(N)$ 作为内部，“外向”有明确意义），则诱导体元 $\hat\varepsilon_{a_1\cdots a_{n-1}}$ 与 N 上体元 $\varepsilon_{a_1\cdots a_{n-1}}$ 有如下关系：

$$\hat\varepsilon_{a_1\cdots a_{n-1}}=n^b\varepsilon_{ba_1\cdots a_{n-1}}$$

下面就是把上上式作为特例包括在内的一般 Gauss 定理。

定理 5.13. 设 M 是 n 维定向流形，N 是 M 中的 n 维紧致带边嵌入子流形，g_{ab} 是 M 上的度规，ε 和 ∇_a 分别是适配体元和适配导数算符，$\hat\varepsilon$ 是 ∂N 上的诱导体元，∂N 的外向法矢 n^a 满足 $n^a n_a=\pm1$，v^a 是 M 上的 C^1 矢量场，则

$$\int_{i(N)}(\nabla_a v^a)\varepsilon=\pm\int_{\partial N}v^a n_a\hat\varepsilon\ (n^a n_a=+1\text{ 时取 }+,\ n^a n_a=-1\text{ 时取 }-)$$

证明. 由上一定理知只须证明 $\int_{\partial} v^b \varepsilon_{ba_1\cdots a_{n-1}} = \pm\int_{\partial N} v^a n_a \hat{\varepsilon}$。令 $\omega = v^b\varepsilon_{ba_1\cdots a_{n-1}}$, 注意到 $\int_{\phi[S]}\omega \equiv \int_{\phi[S]}\tilde{\omega}$, 只须证明

$$\tilde{\omega}_{a_1\cdots a_{n-1}} = \pm v^b n_b \hat{\varepsilon}_{a_1\cdots a_{n-1}},\ \forall q\in\partial N$$

其中 n^a 为 ∂N 的单位外向法矢。上式两边都是 W_q 上的 $n-1$ 形式，故必有 K 使

$$\tilde{\omega}_{a_1\cdots a_{n-1}} = K v^b n_b \hat{\varepsilon}_{a_1\cdots a_{n-1}}$$

于是只须证明 $K=\pm1$。设 $\{(e_0)^a = n^a, (e_1)^a, \cdots, (e_{n-1})^a\}$ 是 V_q 的一个右手正交归一基底，用 $(e_1)^{a_1}\cdots(e_{n-1})^{a_{n-1}}$ 缩并上式，右边给出

$$K v^b n_b \hat{\varepsilon}_{12\cdots(n-1)} = \pm K v^b (e^0)_b \hat{\varepsilon}_{12\cdots(n-1)} = \pm K v^0$$

（其中 $\hat{\varepsilon}_{12\cdots(n-1)}$ 是 $\hat{\varepsilon}_{a_1\cdots a_{n-1}}$ 在用基底 $\{(e_1)^a,\cdots,(e_{n-1})^a\}$ 表示时的分量，第一个等号是因为 $n_b = \pm(e^0)_b$[5]，第二个等号是因为右手基底保证 $\hat{\varepsilon}_{12\cdots(n-1)} = 1$。）另一方面，左边缩并结果为

$$\begin{aligned}&\tilde{\omega}_{a_1\cdots a_{n-1}}(e_1)^{a_1}\cdots(e_{n-1})^{a_{n-1}} = \omega_{a_1\cdots a_{n-1}}(e_1)^{a_1}\cdots(e_{n-1})^{a_{n-1}}\\ =&v^b\varepsilon_{ba_1\cdots a_{n-1}}(e_1)^{a_1}\cdots(e_{n-1})^{a_{n-1}} = v^\mu\varepsilon_{\mu12\cdots(n-1)} = v^0\varepsilon_{012\cdots(n-1)} = v^0\end{aligned}$$

对比两式得 $K=\pm1$。 □

注 5.7. 上述定理的条件是 n^a 为 ∂N 的外向单位法矢。把规定改为“当 $n^a n_a = +1$ 时，n^a 朝外向，$n^a n_a = -1$ 时 n^a 朝内向”，则上述定理右边的 $\pm$ 号消失，*Gauss* 定理改取如下形式

$$\int_{i(N)}(\nabla_a v^a)\varepsilon = \int_{\partial N} v^a n_a \hat{\varepsilon}$$

若 ∂N 为类光超曲面，即 $n^a n_a = 0$，则上式仍然成立，但 $\hat{\varepsilon}$ 要按下式定义：

$$\frac{1}{n}\varepsilon_{a_1\cdots a_n} = n_{[a_1}\hat{\varepsilon}_{a_2\cdots a_n]}$$

5.6 对偶微分形式

以 $\Lambda_p(l)$ 代表 $p\in M$ 的全部 l 形式的集合，则有

$$\dim\Lambda_p(l) = C_n^l = C_n^{n-l} = \dim\Lambda_p(n-l)$$

若 M 为带度规 g_{ab} 的定向流形，ε 为适配体元，则可用 ε 及 g_{ab} 在 $\Lambda_M(l)$ 和 $\Lambda_M(n-l)$ 之间定义一个同构映射如下：

[5] $n_b = n_0(e^0)_b$，其中 $n_0 \equiv n_b(e_0)^b = n_b n^b = \pm1$。

定义 5.12. $\forall\omega\in\Lambda_M(l)$，定义 ω 的**对偶微分形式**$^*\omega\in\Lambda_M(n-l)$ 为

$$^*\omega_{a_1\cdots a_{n-l}}\coloneqq\frac{1}{n!}\omega^{b_1\cdots b_l}\varepsilon_{b_1\cdots b_l a_1\cdots a_{n-l}}$$

其中

$$\omega^{b_1\cdots b_l}=g^{b_1c_1}\cdots g^{b_lc_l}\omega_{c_1\cdots c_l}$$

注 5.8. 以上定义的 * 称为 ***Hodge star***。不难看出：① $^*:\Lambda_M(l)\to\Lambda_M(n-l)$ 是同构映射；② $f\in\mathscr{F}_M$ 作为 0 形式场，其对偶形式场按定义为：

$$^*f_{a_1\cdots a_n}=\frac{1}{0!}f\varepsilon_{a_1\cdots a_n}=f\varepsilon_{a_1\cdots a_n}$$

即 $^*f$ 等于度规适配的体元 ε 的 f 倍，因此可以说函数 f 的积分定义为其对偶形式场的积分。对上式再取 * 得

$$^*(^*f)={}^*(f\varepsilon)=\frac{1}{n!}f\varepsilon^{b_1\cdots b_n}\varepsilon_{b_1\cdots b_n}=(-1)^sf$$

这一结果可推广为如下定理：

定理 5.14. $^{**}\omega=(-1)^{s+l(n-l)}\omega$

现在用微分几何观点重新观察早已熟悉的 3 维欧氏空间 $(\mathbb{R}^3,\delta_{ab})$ 上的矢量代数和矢量场论（其中 M 就是 $\mathbb{R}^3$）。

（1） 为什么过去从未听说过 1、2 和 3 形式场？首先，利用欧式度规 δ_{ab} 可把对偶矢量场 ω_a 变为矢量场 $\omega^a=\delta^{ab}\omega_b$，从而消除了使用 1 形式场的必要性。今后在只涉及 $(\mathbb{R}^3,\delta_{ab})$ 时将不再认真区分上下标。其次，由于 $n=3$，$\Lambda_M(2)$ 与 $\Lambda_M(1)$ 维数相同并可用同构映射 $^*:\Lambda_M(2)\to\Lambda_M(1)$ 把 $\omega\in\Lambda_M(2)$ 与 $^*\omega\in\Lambda_M(1)$ 认同，从而消除了使用 2 形式场的必要性。类似地，$\Lambda_M(3)$ 与 $\Lambda_M(0)$ 维数相同并可用同构映射 $^*:\Lambda_M(3)\to\Lambda_M(0)$ 把 $\omega\in\Lambda_M(3)$ 与 $^*\omega\in\Lambda_M(0)$ 认同，而后者（0 形式场）就是 $\mathbb{R}^3$ 上的函数。可见 3 维欧氏空间的微分形式场都可由函数和矢量场代替。

（2） 现在讨论矢量代数的点乘和叉乘运算。把矢量 $\vec{A}$ 和 $\vec{B}$ 分别记作 A^a 和 B^b。$\vec{A}$ 和 $\vec{B}$ 的点乘积 $\vec{A}\cdot\vec{B}$ 自然就是 A_aB^a，但对叉乘积 $\vec{A}\times\vec{B}$ 如何理解？令

$$\omega_{ab}\equiv A_a\wedge B_b=2A_{[a}B_{b]}$$

则

$$^*\omega_c=\frac{1}{2}\omega^{ab}\varepsilon_{abc}=\varepsilon_{abc}A^{[a}B^{b]}=\varepsilon_{abc}A^aB^b$$

其中，ε_{abc} 是欧氏度规的适配体元。设 $\{x,y,z\}$ 为右手笛卡尔坐标系，则其坐标基底正交归一，于是 ε_{abc} 在此系的非零分量 ε_{ijk} 为

$$\varepsilon_{123}=\varepsilon_{312}=\varepsilon_{231}=-\varepsilon_{132}=-\varepsilon_{213}=-\varepsilon_{321}=1$$

可见 ε_{ijk} 就是熟知的 Levi–Civita 记号[6]。于是 $^*\omega_c$ 在该笛卡尔系的第 k 分量为

$$^*\omega_k = \varepsilon_{ijk}A^iB^j,\ k = 1, 2, 3$$

上式右边正是 $\vec{A} \times \vec{B}$ 按叉乘定义的第 k 分量 $(\vec{A} \times \vec{B})_k$，故 $\vec{A} \times \vec{B}$ 可看作 $^*\boldsymbol{\omega}$（准确地说是对偶矢量 $^*\boldsymbol{\omega}$ 对应的矢量）。而 $\boldsymbol{\omega} = A \wedge B$，可见对 $\vec{A}$ 和 $\vec{B}$ 求叉积就是先求楔积 $A \wedge B$ 再求其对偶形式，可直观地表达为 $\times = ^* \circ \wedge$。

（3） 再从微分几何看 3 维欧氏空间的矢量场论。如前所述，矢量场论的 $\vec{\nabla}$ 就是与欧氏度规 δ_{ab} 适配的导数算符 ∂_a，涉及 $\vec{\nabla}$ 的公式原则上都可用 ∂_a 表出。例如

（a） $\vec{\nabla} f = \partial_a f$

（b） $\vec{\nabla} \cdot \vec{A} = \partial_a A^a$

（c） $\vec{\nabla} \times \vec{A} = \varepsilon^{abc}\partial_a A_b$ （推导参考前述内容）

（d） $\vec{\nabla} \cdot (\vec{A}\vec{B}) = \partial_a(A^aB^b)$

（e） $\vec{\nabla}\vec{A} = \partial^a A^b$

（f） $\nabla^2 f = \partial_a\partial^a f$

（g） $\nabla^2\vec{A} = \partial_a\partial^a A^b$

借用 ∂_a 及抽象指标还可使一些常用公式的推证简化而且理由清晰。仅举二例。

例 5.5. 用 ∂_a 证明

$$\vec{\nabla} \cdot (\vec{A} \times \vec{B}) = \vec{B} \cdot (\vec{\nabla} \times \vec{A}) - \vec{A} \cdot (\vec{\nabla} \times \vec{B})$$

证明. $\vec{\nabla} \cdot (\vec{A} \times \vec{B}) = \partial_c(\varepsilon^{cab}A_aB_b) = \varepsilon^{cab}(A_a\partial_c B_b + B_b\partial_c A_a)$

而 $\vec{B} \cdot (\vec{\nabla} \times \vec{A}) = B_b(\vec{\nabla} \times \vec{A})^b = B_b(\varepsilon^{bac}\partial_c A_a) = \varepsilon^{cab}B_b\partial_c A_a$

$-\vec{A} \cdot (\vec{\nabla} \times \vec{B}) = -A_a(\vec{\nabla} \times \vec{B})^a = -A_a(\varepsilon^{acb}\partial_c B_b) = \varepsilon^{cab}A_a\partial_c B_b$

代入便得证。 □

例 5.6. 用 ∂_a 证明

$$\vec{\nabla}(\vec{A} \cdot \vec{B}) = (\vec{A} \cdot \vec{\nabla})\vec{B} + (\vec{B} \cdot \vec{\nabla})\vec{A} + \vec{A} \times (\vec{\nabla} \times \vec{B}) + \vec{B} \times (\vec{\nabla} \times \vec{A})$$

---

[6] $\varepsilon_{a_1 a_2 a_3 \dots a_n} \coloneqq \begin{cases} +1 & \text{如果}(a_1, a_2, a_3, \dots, a_n)\text{是}(1, 2, 3, \dots, n)\text{的一个偶排列} \\ -1 & \text{如果}(a_1, a_2, a_3, \dots, a_n)\text{是}(1, 2, 3, \dots, n)\text{的一个奇排列} \\ 0 & \text{其他情况} \end{cases}$

证明. 右边第一项 $= A_a\partial^a B^b$

右边第二项 $= B_a\partial^a A^b$

$$\begin{aligned}\text{右边第三项} &= \vec{A}\times(\varepsilon^{cde}\partial_d B_e) = \varepsilon^{bac}A_a(\varepsilon_{cde}\partial^d B^e)\\ &= 2\delta^b_{[d}\delta^a_{e]}A_a\partial^d B^e = (\delta^b_d\delta^a_e - \delta^b_e\delta^a_d)A_a\partial^d B^e = A_a\partial^b B^a - A_a\partial^a B^b\end{aligned}$$

同理，右边第四项 $= B_a\partial^b A^a - B_a\partial^a A^b$

故，右边 $= A_a\partial^b B^a + B_a\partial^b A^a = \partial^b(A_a B^a) = \vec{\nabla}(\vec{A}\cdot\vec{B})$ □

（4） 3 维欧氏空间中的梯度、旋度和散度可用外微分简单表述如下：

定理 5.15. 设 f 和 $\vec{A}$ 是 3 维欧氏空间的函数和矢量场，则

$$\mathrm{grad} f = \mathrm{d}f,\ \mathrm{curl}\vec{A} = {}^*\mathrm{d}A,\ \mathrm{div}\vec{A} = {}^*\mathrm{d}({}^*A)$$

证明.

$$\mathrm{d}A = 2\partial_{[b}A_{a]}$$

$$^*\mathrm{d}A = \partial^{[b}A^{a]}\varepsilon_{bac} = \vec{\nabla}\times\vec{A}$$

$$^*A = A^a\varepsilon_{abc}$$

$$\mathrm{d}^*A = 3\partial_{[d}A^a\varepsilon_{|a|bc]}$$

$$^*\mathrm{d}^*A = \frac{3}{3!}\partial^{[d}A_a\varepsilon^{|a|bc]}\varepsilon_{bcd} = \frac{1}{2}\partial^d A_a\varepsilon^{bca}\varepsilon_{bcd} = \frac{1}{2}\partial^d A_a 2!\delta^a_d = \partial_a A^a$$

□

$\mathbb{R}^3$ 是平凡流形保证 $\mathbb{R}^3$ 上的闭形式场必恰当，同上述定理结合便很容易证明 3 维欧氏空间场论中并不易证的下列熟知命题：

（1） 无旋矢量场必可表为梯度，即

$$\mathrm{curl}\vec{E} = 0 \Rightarrow \exists\text{标量场}\phi\ \text{使}\vec{E} = \mathrm{grad}\phi$$

（2） 无散矢量场必可表为旋度，即

$$\mathrm{div}\vec{B} = 0 \Rightarrow \exists\text{矢量场}\vec{A}\ \text{使}\vec{B} = \mathrm{curl}\vec{A}$$

证明.（1） $\mathrm{curl}\vec{E} = 0 \Rightarrow {}^*\mathrm{d}E = 0 \Rightarrow^* ({}^*\mathrm{d}E) =^* 0 = 0$ 而根据 $^{**}\boldsymbol{\omega} = (-1)^{s+l(n-l)}\boldsymbol{\omega}$，可知 $^*({}^*\mathrm{d}E) = \mathrm{d}E$

（2） $\mathrm{div}\vec{B} = 0 \Rightarrow {}^*\mathrm{d}({}^*B) = 0 \Rightarrow \mathrm{d}({}^*B) = 0 \Rightarrow^* B = \mathrm{d}A \Rightarrow B = {}^*\mathrm{d}A$

□

第六章　量子力学数学基础简介

量子力学数学基础的主要内容是泛函分析，特别是希尔伯特空间及其线性算符的理论。读过本书前两章的读者对集合、映射、拓扑、矢量空间及其对偶空间、张量、度规等概念比较熟悉，这为学习泛函分析提供了一定的方便。学过上述概念的物理系学生往往跃跃欲试地把量子力学的内积、左右矢、线性算符以及波函数用正交归一基底展开等一系列问题同上述概念联系起来思考，力图求得更为深入和清晰的理解。他们希望有一份简明读物作参考。本章主要为满足这种需要而产生。讲解中尽量与本书前两章以及量子力学的有关内容（特别是 Dirac 的左右矢记号）相联系或对比。为了尽量节省读者的时间，我们只介绍最为必需的概念和定理，而且略去其中少数概念的定义和某些定理的证明。想深入学习泛函分析的读者则应阅读有关教程或专著。

6.1　Hilbert 空间初步

6.1.1　Hilbert 空间及其对偶空间

第二章中定义的矢量空间称为实矢量空间。把定义中的 $\mathbb{R}$ 改为全体复数的集合 $\mathbb{C}$ 便得到复矢量空间。定义了内积的复矢量空间 V 称为（复）内积空间。内积是一个从 $V\times V$ 到 $\mathbb{C}$ 的映射。以 i 代表这一映射，即 $i\colon V\times V\to\mathbb{C}$，则 i 可看作有两个槽的机器，在两槽中分别输入 $f,g\in V$，便得一个复数。因此，i 也可形象地记作 $i(\bullet,\bullet)$，括号中的两个圆点代表两个槽。为简单起见，索性把 $i(\bullet,\bullet)$ 简写为 $(\bullet,\bullet)$，确切定义如下：

定义 6.1. 复矢量空间 V 称为内积空间，若存在内积映射 $i\colon V\times V\to\mathbb{C}$，对任意 $f,g,h\in V$ 和任意 $c\in\mathbb{C}$ 满足

（a）　$(f,g+h)=(f,g)+(f,h)$；

（b）　$(f,cg)=c(f,g)$；

（c）　$(f,g)=\overline{(g,f)}$（其中 $\overline{(g,f)}$ 代表复数 (g,f) 的共轭复数）；

（d）　$(f,f)\geqslant 0$，且 $(f,f)=0\Leftrightarrow f=0$。[1]

[1] 只含零元的空间也可看作内积空间，但本章在不加声明时只讨论维数大于零的内积空间。

注 6.1. 上述定义的前两条件表明内积映射 $(\bullet,\bullet)$ 对第二槽是线性的。第三条件(配以前两条件)则表明 $(f+g,h)=(f,h)+(g,h)$ 及 $(cf,g)=\bar{c}(f,g)$。由于具有这种性质,我们说映射 $(\bullet,\bullet)$ 对第一槽是反线性(或共轭线性)的。最后一个条件表明 $(f,g)=0\ \forall g\in V \Rightarrow f=0$,因可取 $g=f$。

注 6.2. 上述定义对实矢量空间也适用,只须把 $\mathbb{C}$ 改为 $\mathbb{R}$。实空间的内积就是第二章定义的正定度规,但是,由于非正定度规不满足上述定义的最后一个条件,所以度规未必是内积。

例 6.1.

$$C[a,b]\equiv\{[a,b]\subset\mathbb{R}\text{上连续复值函数}f(x)\}$$

是无限维复矢量空间。定义内积

$$(f,g):=\int_b^a\overline{f(x)}g(x)dx,\ \forall f,g\in C[a,b]$$

则 $C[a,b]$ 是内积空间。

利用内积可自然定义内积空间中任意两点的距离,进而定义空间的拓扑。

定义 6.2. 内积空间 V 中任意两元素 f 和 g 的距离定义为

$$d(f,g):=\sqrt{(f-g,f-g)}$$

易见 $d(f,g)=d(g,f)$。设 $f\in V, r>0$,则以 f 为心、以 r 为半径的开球定义为

$$B(f,r):=\{g\in V\mid d(g,f)<r\}$$

用开球可给 V 定义拓扑(类似于第一章中的"通常拓扑"):

$$\mathscr{T}:=\{\text{空集或 }V\text{ 中能表为开球之并的子集}\}$$

可见内积空间 V 可自然地被定义为一个拓扑空间。上式定义的拓扑 $\mathscr{T}$ 叫做 V 的自然拓扑。今后如无特别声明,凡涉及 V 的拓扑时一律指这一拓扑。

第二章讲过(有限维)实矢量空间 V 的对偶空间 V^*,它是由 V 到 $\mathbb{R}$ 的全体线性映射的集合。对(有限或无限维)复矢量空间 V,对偶矢量可定义为由 V 到 $\mathbb{C}$ 的线性映射。然而内积空间除了是复矢量空间外还是拓扑空间,因此对映射 $\eta\colon V\to\mathbb{C}$ 还可问及是否连续的问题。(这时还涉及 $\mathbb{C}$ 的拓扑,其定义也很自然:设 $z_1,z_2\in\mathbb{C}$ 且 $z_1=a_1+ib_1, z_2=a_2+ib_2$(其中 $a_1,a_2,b_1,b_2\in\mathbb{R}$),则 z_1,z_2 的距离可定义为 $d(z_1,z_2):=[(a_1-a_2)^2+(b_1-b_2)^2]^{1/2}$,用此距离便可定义开球,从而定义 $\mathbb{C}$ 的拓扑。)在泛函分析中,每个连续的线性映射 $\eta\colon V\to\mathbb{C}$ 称为 V 上的一个连续线性泛函。我们关心 V 上全体连续线性泛函的集合(它有许多好的性质),并称它为 V 的对偶空间。[2]

[2]若 V 为有限维,则 V 上的线性泛函必定连续。因此只当 V 为无限维时连续性才是对线性泛函有实质意义的要求。

定义 6.3. 内积空间 V 的对偶空间（又称共轭空间）定义为

$$V^* \coloneqq \{\eta\colon V \to \mathbb{C} \mid \eta\text{为连续的线性映射}\}$$

V^* 也可看作复矢量空间，为此只须用如下的自然方法定义加法和数乘：

加法 $(\eta_1 + \eta_2)(f) \coloneqq \eta_1(f) + \eta_2(f)$, $\forall \eta_1\eta_2 \in V^*, f \in V$;

数乘 $(c\eta)(f) \coloneqq c \cdot \eta(f)$, $\forall \eta \in V^*, c \in \mathbb{C}$

（由此可知零元是 V^* 中这样的元素，它作用于任意 $f \in V$ 都得零。）

读者至此自然会联想到量子力学的右矢和左矢空间，并猜想右矢空间是内积空间 V，而左矢空间则是 V^*。然而事情比此略为复杂。要保证 Dirac 的左右矢记号得心应手，左右逢源，左、右矢空间应该 “完全对等”。这似乎不难：根据第二章，有限维实矢量空间 V 上的度规自然诱导一个从 V 到 V^* 的一一到上的线性映射。然而，由于复空间的内积与实空间的度规的少许不同，复空间 V 上的内积自然诱导的从 V 到 V^* 的映射不是线性而是反线性的。不过这不构成什么问题，真正构成问题的是量子力学中用到的内积空间多数是无限维的，而这导致上述映射未必到上，即 V 与 V^*“不一样大”。先看如下命题：

定理 6.1. 内积映射 $(\bullet,\bullet)$ 自然诱导出一个一一的、反线性的映射 $\nu\colon V \to V^*$。

证明. 设 $f \in V$，则 $\eta_f \equiv (f,\bullet)$ 是从 V 到 $\mathbb{C}$ 的映射。由内积定义可知它是线性的。还可证明（略）它是连续的，因此 $\eta_f \in V^*$。具体说，η_f 是 V^* 的这样一个元素，它作用于 $g \in V$ 的结果为 $\eta_f(g) \coloneqq (f, g)$。可见 $(\bullet,\bullet)$ 自然诱导出一个映射 $\nu\colon V \to V^*$，定义为 $\nu(f) \coloneqq \eta_f$。根据映射 ν 的定义，其反线性性是显然的，其一一性证明如下：设 $\exists f_1, f_2 \in V$ 使得 $\nu(f_1)(g) - \nu(f_2)(g) = 0\ \forall g \in V$，则有 $(f_1 - f_2, g) = 0\ \forall g \in V$，因此就有 $f_1 = f_2$。 □

若 V 是有限维空间，则 V^* 与 V 有相同的维数（请参考第二章中对实矢量空间的证明），因而上述反线性映射 $\nu\colon V \to V^*$ 一定到上。若 V 是无限维的，则 V^* 也是无限维的，这时 $\nu\colon V \to V^*$ 不一定到上。直观地说，V^* 有可能 “比 V 大”，即 $\nu[V] \subset V^*$ 但 $\nu[V] \neq V^*$。然而，为保证 Dirac 的左右矢运算给出正确结果，我们需要 V^* 同 V“一样大”，即 $\nu[V] = V^*$。不久将看到，其实 V^* 比 V 最多 “只多一层皮”，只要给 V 适当地 “补上一层皮”，$\nu\colon V \to V^*$ 就可到上，V 与 V^* 就 “一样大”。为了用准确语言表述这些直观思考，我们先介绍下列数学概念。

定义 6.4. 设 V 为内积空间，$f \in V$，$\{f_n\}$ 是 V 中的一个序列（见第一章最后一节中的定义）。我们说 $\{f_n\}$ 收敛于 f（记作 $f_n \to f$），若 $\lim\limits_{n\to\infty} d(f, f_n) = 0$。$f$ 称为序列 $\{f_n\}$ 的极限，[3]记作 $f = \lim\limits_{n\to\infty} f_n$。

[3]内积空间可自然看作拓扑空间，而拓扑空间中的点序列的极限已有定义（见第一章最后一节）。不难证明该定义同本页定义等价。

定义 6.5. V 中的序列 $\{f_n\}$ 称为柯西序列，若 $\forall\varepsilon > 0, \exists N > 0$ 使当 $n, m \geqslant N$ 时有 $d(f_n, f_m) < \varepsilon$。

可以证明，收敛（于任一 $f \in V$）的序列一定是柯西序列，但反之不然。

定义 6.6. 内积空间 V 称为完备的，若其中任一柯西序列收敛。

前述例子中的 $C[a, b]$ 是不完备的内积空间。因为存在这样的柯西序列，其收敛于一个不连续函数（因此它在 $C[a, b]$ 中不收敛）。柯西序列 $\{f_n(x)\}$ 在 $C[a, b]$ 内没有极限表明 $C[a, b]$ 不完备。为使之完备化，可把 $[a, b]$ 上虽不连续却平方可积的复值函数（例如上述柯西序列的极限）包含进去，扩大后的空间记作 $L^2[a, b]$，即

$$L^2[a, b] \equiv \{f\colon [a, b] \to \mathbb{C} \mid \int_a^b |f(x)|^2 dx < \infty\}^4$$

仍用前面例子中的内积定义，则可证 $L^2[a, b]$ 是完备的内积空间。

内积空间 $C[a, b]$ 的这一完备化程序很有启发性。事实上，任何不完备的内积空间 V 都可以完备化，为此只须把它略加扩大—把所有柯西序列“应有的极限点”都补进 V 中。可以证明，对任何不完备内积空间 V，总可找到完备的内积空间 $\tilde{V}$，使得 $V \subset \tilde{V}$ 而且 $\bar{V} = \tilde{V}$，其中 $\bar{V}$ 是把 $\tilde{V}$ 看作拓扑空间（用其内积定义拓扑）时子集 V 的闭包。直观地可以说 $\tilde{V}$ 比 V“最多只多一层皮”。

定义 6.7. 完备的内积空间叫希尔伯特空间，记作 $\mathscr{H}$。

注 6.3. 有限维的内积空间一定完备，因此一定是 *Hilbert* 空间。然而量子力学中用到的 *Hilbert* 空间多数是无限维的。无限维是许多问题变得复杂的根源。

同 $L^2[a, b]$ 相仿，

$$L^2(\mathbb{R}^n) \equiv \{f\colon \mathbb{R}^n \to \mathbb{C} \mid \int |f|^2 d^n x < \infty\}\text{（内积定义仿照前述例子）}$$

也是 Hilbert 空间，其中 $L^2(\mathbb{R}^3)$ 是量子力学常用的波函数空间。

由于具有完备性，Hilbert 空间有许多很好的性质，其中对我们特别有用的就是 $\mathscr{H}$ 与其对偶空间 $\mathscr{H}^*$“一样大”，即 $\nu[\mathscr{H}] = \mathscr{H}^*$，见如下命题：

定理 6.2. 设 $\mathscr{H}$ 是 *Hilbert* 空间，$\mathscr{H}^*$ 是其对偶空间，则 $\forall\eta \in \mathscr{H}^*$，有唯一的 $f_\eta \in \mathscr{H}$ 使 $\eta(g) = (f_\eta, g)\ \forall g \in \mathscr{H}$。

证明. 见任一泛函分析教程中关于 Riesz 表现定理的证明。 □

[4]式中的积分是指 Lebesgue 积分。严格地说，函数 $f(x)$ 还应具有“可测性”。可以这样说：物理上遇到的函数都满足这一要求。此外，两个函数如果只在测度为零的集上不同就应被视为 $L^2[a, b]$ 的同一元素。本脚注同样适用于 $L^2(\mathbb{R}^n)$。

注 6.4. 上述定理表明，对 *Hilbert* 空间，前述命题中的 $\nu\colon \mathscr{H} \to \mathscr{H}^*$ 是到上的，即 ν 是一一到上的反线性映射。这个命题的重要性在于保证 $\mathscr{H}$ 同 $\mathscr{H}^*$"一样大"，请注意不完备内积空间没有这样好的结论。可见物理学不但需要内积空间，而且需要完备的内积空间—*Hilbert* 空间。

利用映射 $\nu\colon \mathscr{H} \to \mathscr{H}^*$ 还可把 $\mathscr{H}^*$ 定义为 Hilbert 空间：$\forall \eta, \xi \in \mathscr{H}^*$，由上述定理可知有唯一的 $f_\eta, f_\xi \in \mathscr{H}$ 使 $\eta = \nu(f_\eta), \xi = \nu(f_\xi)$。定义 η 和 ξ 的内积为

$$(\eta, \xi) \coloneqq (f_\xi, f_\eta)$$

则不难验证 (η, ξ) 满足内积定义，故 $\mathscr{H}^*$ 是内积空间。还可证明 $\mathscr{H}^*$ 是完备的，因而也是 Hilbert 空间。可见 $\mathscr{H}^*$ 与 $\mathscr{H}$ 实在非常"像"：它们之间不但存在一一到上的反线性映射，而且这一映射还在上式的意义上保内积。$\mathscr{H}$ 和 $\mathscr{H}^*$ 的这种相像性使我们可以把它们分别用作量子力学中的右矢和左矢空间（详见第四小节）。

6.1.2 Hilbert 空间的正交归一基

N 维矢量空间 V 的一个基底无非是由 N 个元素组成的、满足如下两个要求的一个序列 $\{e_1, \cdots, e_N\}$：① $\{e_1, \cdots, e_N\}$ 线性独立；② V 的任一元素 f 可由 $\{e_1, \cdots, e_N\}$ 线性表出。我们想把基底概念推广至无限维的 Hilbert 空间 $\mathscr{H}$。

定义 6.8. $\mathscr{H}$ 的有限子集 $\{f_1, \cdots, f_N\}$ 称为线性独立的，若

$$\sum_{n=1}^{N} c_n f_n = 0 \Rightarrow c_n = 0, n = 1, \cdots, N$$

$\mathscr{H}$ 的任一子集 $\{f_\alpha\}$ 称为线性独立的，若 $\{f_\alpha\}$ 的任一非空有限子集线性独立。

如果 $\mathscr{H}$ 中存在满足以下两条件的无限序列 $\{e_n\}$：① $\{e_n\}$ 线性独立；② $\mathscr{H}$ 的任一元素 f 可由 $\{e_n\}$ 线性表出：

$$f = \sum_{n=1}^{\infty} c_n e_n,\ c_n \in \mathbb{C}$$

就说 $\{e_n\}$ 构成 $\mathscr{H}$ 的一个基底。（上式中的 $f = \sum_{n=1}^{\infty} c_n e_n$ 是 $f = \lim_{N\to\infty} \sum_{n=1}^{N} c_n e_n$ 的简写。请注意这里的极限涉及 $\mathscr{H}$ 的拓扑，对未定义拓扑的矢量空间无意义。）

下面再讨论 $\mathscr{H}$ 的正交归一基。

定义 6.9. *Hilbert* 空间 $\mathscr{H}$ 中的序列 $\{f_n\}$ 叫正交归一序列，若

$$(f_m, f_n) = \delta_{mn}$$

不难证明 $\mathscr{H}$ 中的任一正交归一序列都线性独立。所以，若 $\mathscr{H}$ 的维数有限，则当 $\{f_n\}$ 的元素个数等于 $\mathscr{H}$ 的维数时，$\{f_n\}$ 自然构成 $\mathscr{H}$ 的一个基底，而且是正交归一基。当 $\mathscr{H}$ 是无限维时，要成为正交归一基，$\{f_n\}$ 的元素必须无限多个。然而，并非由无限多个元素构成的正交归一序列 $\{f_n\}$ 都是正交归一基，这里有一个是否已把元素“选够”的问题（例如，设 $\{f_n\}$ 是基底，则只取 n 为偶数的子集也含无限多个元素，但却不是基底），只有满足下面定义的完备性条件的 $\{f_n\}$ 才能成为正交归一基。

定义 6.10. $\mathscr{H}$ 中的正交归一序列 $\{f_n\}$ 叫完备的，若 $\mathscr{H}$ 中除零元外不存在与每个 f_n 都正交的元素（即不能通过给 $\{f_n\}$ 添加新元素而得到“更大”的正交归一序列）。

$\{f_n\}$ 的完备性保证 $\mathscr{H}$ 的任一元素 f 都可用 $\{f_n\}$ 线性表出，因此 $\mathscr{H}$ 的任一完备的正交归一序列（如果存在）都是 $\mathscr{H}$ 的正交归一基。改用 $\{e_n\}$ 代表完备的正交归一序列，则任一 $f \in \mathscr{H}$ 都可用 $\{e_n\}$ 线性表出。由前述两式得

$$c_n = (e_n, f)$$

故

$$f = \sum_{n=1}^{\infty} (e_n, f) e_n$$

6.1.3 Hilbert 空间上的线性算符

定义 6.11. 映射 $A\colon \mathscr{H} \to \mathscr{H}$ 称为 $\mathscr{H}$ 上的算符，数学书一般译作算子。A 作用于 $f \in \mathscr{H}$ 的结果记作 Af。A 称为线性算符，若

$$A(c_1 f_1 + c_2 f_2) = c_1 A f_1 + c_2 A f_2,\ \forall f_1, f_2 \in \mathscr{H}, c_1, c_2 \in \mathbb{C}$$

定义 6.12. 算符 $A\colon \mathscr{H} \to \mathscr{H}$ 和 $B\colon \mathscr{H} \to \mathscr{H}$ 称为相等的，若 $Af = Bf\ \forall f \in \mathscr{H}$。

今后如无特别声明，行文中的算符均指线性算符。$\mathscr{H}$ 上全体线性算符的集合 $\mathscr{L}(\mathscr{H})$ 也是个复矢量空间[5]，只要用如下的自然方式定义加法和数乘：

加法 $(A_1 + A_2)f \coloneqq A_1 f + A_2 f,\ \forall A_1 A_2 \in \mathscr{L}(\mathscr{H}), f \in \mathscr{H}$；

数乘 $(cA)f \coloneqq c(Af),\ \forall A \in \mathscr{L}(\mathscr{H}), c \in \mathbb{C}$

（$\mathscr{L}(\mathscr{H})$ 的零元（也叫零算符）是这样的算符，它作用于任意 $f \in \mathscr{H}$ 都得 $\mathscr{H}$ 的零元。）

算符分为有界算符和无界算符两大类（定义见下节）。本节只讨论有界算符。

定义 6.13. $\mathscr{H}$ 上的一个线性算符 $A\colon \mathscr{H} \to \mathscr{H}$ 自然诱导出 $\mathscr{H}^*$ 上的一个线性算符 $A^*\colon \mathscr{H}^* \to \mathscr{H}^*$，定义为

$$(A^*\eta)(f) \coloneqq \eta(Af),\ \forall f \in \mathscr{H}, \eta \in \mathscr{H}^*$$

[5]注意和 $\mathscr{H}^*$ 区分，$\mathscr{H}^*$ 中的元素皆是从 $\mathscr{H}$ 到 $\mathbb{C}$ 的连续线性映射。

易见 $A^*\eta$（作为从 $\mathscr{H}$ 到 $\mathbb{C}$ 的映射）是线性的，还可证明它是连续的，因此 $A^*\eta \in \mathscr{H}^*$，从而保证 A^* 是从 $\mathscr{H}^*$ 到 $\mathscr{H}^*$ 的映射。这样定义的 A^* 称为算符 A 的对偶算符。

注 6.5. ① A 和 A^* 是两个不同 *Hilbert* 空间上的算符，其中 $A\colon \mathscr{H} \to \mathscr{H}$ 而 $A^*\colon \mathscr{H}^* \to \mathscr{H}^*$。② 不难证明：

（a） A^* 的确是线性算符

（b） A^* 与 A 的对应关系是线性的，即

$$(A_1 + A_2)^* = A_1^* + A_2^*,\ (cA)^* = cA^*$$

$\mathscr{H}$ 上的任一线性算符 A 的对偶算符 A^* 又可自然诱导出 $\mathscr{H}$ 上的一个线性算符 $A^\dagger\colon \mathscr{H} \to \mathscr{H}$。设 $\nu\colon \mathscr{H} \to \mathscr{H}^*$ 是前述一一、到上、反线性映射，则可定义 $A^\dagger$ 为如下的复合映射：

$$A^\dagger \coloneqq \nu^{-1} \circ A^* \circ \nu$$

ν 的反线性性导致 ν^{-1} 的反线性性，加上 A^* 的线性性，便知 $A^\dagger$ 是线性的。

定义 6.14. 如上定义的 $A^\dagger\colon \mathscr{H} \to \mathscr{H}$ 叫 A 的伴随算符。

注 6.6. A 和 $A^\dagger$ 都是 $\mathscr{H}$ 上的算符，但 A^* 是 $\mathscr{H}^*$ 上的算符。

定理 6.3. 设 $A^\dagger$ 是 A 的伴随算符，则

$$(f, Ag) = (A^\dagger f, g),\ \forall f, g \in \mathscr{H}$$

反之，若 $B\colon \mathscr{H} \to \mathscr{H}$ 满足

$$(f, Ag) = (Bf, g),\ \forall f, g \in \mathscr{H}$$

则 $B = A^\dagger$。

证明. $(f, Ag) = \eta_f(Ag) = (A^*\eta_f)(f) \equiv \eta_h(g) = (h, g) = (A^\dagger f, g)$, 其中第一步用到 η_f 的定义，第二步用到 A^* 的定义，第三步无非是把 $A^*\eta_f$ 记作 η_h，最后一步用到 $A^\dagger$ 的定义。反之，

$$0 = (Bf, g) - (A^\dagger f, g) = (Bf - A^\dagger f, g),\ \forall f, g \in \mathscr{H}$$

故 $0 = Bf - A^\dagger f = (B - A^\dagger)f\ \forall f \in \mathscr{H}$，因而 $B = A^\dagger$。 □

注 6.7. 可见上式可用作 $A^\dagger$ 的等价定义。

定理 6.4. $A^\dagger$ 与 A 的对应关系是反线性的，即

$$(A_1 + A_2)^\dagger = A_1^\dagger + A_2^\dagger$$
$$(cA)^\dagger = \bar{c}A^\dagger,\ \forall c \in \mathbb{C}$$

证明. $((A_1 + A_2)^\dagger)f, g) = (f, (A_1 + A_2)g) = (f, A_1 g) + (f, A_2 g) = (A_1^\dagger f, g) + (A_2^\dagger f, g) = ((A_1^\dagger + A_2^\dagger)f, g)$

$((cA)^\dagger f, g) = (f, cAg) = c(f, Ag) = c(A^\dagger f, g) = (\bar{c}A^\dagger f, g)$ □

定理 6.5. 设 A 为 $\mathscr{H}$ 上的有界算符，则 $A^{\dagger\dagger} = A$

证明. $(A^{\dagger\dagger}f, g) = (f, A^\dagger g) = (\overline{A^\dagger g, f}) = (\overline{g, Af}) = (Af, g)$ □

定义 6.15. （有界）线性算符 $A\colon \mathscr{H} \to \mathscr{H}$ 称为自伴的或厄米的，若 $A = A^\dagger$，即

$$(f, Ag) = (Af, g),\ \forall f, g \in \mathscr{H}$$

注 6.8. “厄米算符就是自伴算符”的说法只对有界算符成立。对无界算符，自伴性强于厄米性，详见下节。

6.1.4 Dirac 的左右矢记号

在 Dirac 的记号中，每一 $f \in \mathscr{H}$ 记作 $|f\rangle$，称为右矢；每一 $\eta \in \mathscr{H}^*$ 记作 $\langle\eta|$，称为左矢。$\langle\eta|$ 作用于 $|f\rangle$ 所得复数记作 $\langle\eta|f\rangle$，即 $\langle\eta|f\rangle \equiv \eta(f)$。物理学家常把 $\langle\eta|f\rangle$ 称为 $\langle\eta|$ 与 $|f\rangle$ 的内积，在泛函分析中 $\langle\eta|f\rangle$ 则是 g_η 与 f 的内积，其中 $g_\eta \equiv \nu^{-1}(\eta) \in \mathscr{H}$。记号 $\langle\eta|f\rangle$ 中的 $\langle\rangle$ 可看作一个尖括号，“括号”在英文中是“bracket”，去掉字母 c 并拆为两半，自然把左半 $\langle|$ 称为 bra，而右半 $|\rangle$ 称为 ket。$|f\rangle \in \mathscr{H}$ 在 $\nu\colon \mathscr{H} \to \mathscr{H}^*$ 映射下的像 $\eta_f \in \mathscr{H}^*$ 本应记作 $\langle\eta_f|$，但可简记为 $\langle f|$，这不会与 $|f\rangle$ 混淆，却可形象地表明 $\langle f|$ 就是 $|f\rangle$ 在 ν 映射下的对应物。通常也把这种对应关系记作 $\langle f| \leftrightarrow |f\rangle$。同样，$\langle\eta| \in \mathscr{H}^*$ 的逆像 $\nu^{-1}(\eta)$ 可简记为 $|\eta\rangle$，即 $\langle\eta| \leftrightarrow |\eta\rangle$，原来的 $(f, g) = \eta_f(g)$ 则可表为 $(f, g) = \langle f|g\rangle$。实际上，使用 Dirac 记号后无需再以拉丁字母 $f, g, \cdots$ 和希腊字母 $\eta, \xi, \cdots$ 分别代表 $\mathscr{H}$ 和 $\mathscr{H}^*$ 的元素。设 $|\psi\rangle \in \mathscr{H}, c \in \mathbb{C}$，则 $c\,|\psi\rangle \in \mathscr{H}$，则可记作 $|c\psi\rangle$，即 $c\,|\psi\rangle \equiv |c\psi\rangle$。由内积空间的定义有

$$\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle},$$
$$\langle\psi|c\phi\rangle = c\,\langle\psi|\phi\rangle,$$
$$\langle c\psi|\phi\rangle = \bar{c}\,\langle\psi|\phi\rangle$$

其中 $\langle c\psi|$ 代表 $|c\psi\rangle$ 在映射 ν 下的像，即 $\langle c\psi| \leftrightarrow |c\psi\rangle$。注意到 ν 的反线性性，得

$$\langle c\psi| = \nu(c\psi) = \bar{c}\nu(\psi) = \bar{c}\,\langle\psi|$$

故映射 $\nu\colon \mathscr{H} \to \mathscr{H}^*$ 的反线性性体现为

$$\langle c\psi| = \bar{c}\,\langle\psi|$$

或

$$c\,|\psi\rangle \leftrightarrow \bar{c}\,\langle\psi|$$

其中 $\bar{c}\langle\psi|$ 是复数 $\bar{c}$ 乘左矢 $\langle\psi|$ 所得的左矢，也可记为 $\langle\psi|\,\bar{c}$

算符 A 作用于右矢 $|\psi\rangle \in \mathscr{H}$ 所得的右矢记作 $|A\psi\rangle$，即 $A|\psi\rangle \equiv |A\psi\rangle$。把用 A^* 作用于左矢 $\langle\eta| \in \mathscr{H}^*$ 的结果记作 $A^*\langle\eta|$，则 A^* 的定义式可表为

$$(A^*\langle\eta|)\,|f\rangle = \langle\eta|\,(A\,|f\rangle),\ \forall\,|f\rangle \in \mathscr{H}, \langle\eta| \in \mathscr{H}^*$$

现在说明上式右边的圆括号可以去掉。先回到不用 Dirac 记号的式子。因为 $A\colon \mathscr{H} \to \mathscr{H}$ 而 $\eta\colon \mathscr{H} \to \mathbb{C}$，所以 $\eta \circ A\colon \mathscr{H} \to \mathbb{C}$。$\eta$ 是 $\mathscr{H}$ 上的连续线性泛函保证 $\eta \circ A$ 也是，故 $\eta \circ A \in \mathscr{H}^*$。进一步把 $\eta \circ A$ 简记作 ηA，则有

$$(A^*\eta)(f) = (\eta \circ A)(f) = (\eta A)(f)$$

因而 $A^*\eta = \eta A$（$\in \mathscr{H}^*$），用 Dirac 记号则为 $A^*\langle\eta| = |\eta\rangle A$, 于是前式左边等于 $(\langle\eta|\,A)\,|f\rangle$，前式就成为

$$(\langle\eta|\,A)\,|f\rangle = \langle\eta|\,(A\,|f\rangle)$$

上式表明圆括号没有必要，$\langle\eta|\,A\,|f\rangle$ 有明确含义，它既可理解为 $(\langle\eta|\,A)\,|f\rangle$，也可理解为 $\langle\eta|\,(A\,|f\rangle)$。有人把 $A^\dagger$ 记作 A^*，则他们的 $\langle\eta|\,A^*$ 是我们的 $\langle\eta|\,A^\dagger = A^{\dagger *}\langle\eta|$。

定理 6.6. $A\,|\psi\rangle \leftrightarrow \langle\psi|\,A^\dagger$

证明. 因 $A\,|\psi\rangle \equiv |A\psi\rangle \leftrightarrow \langle A\psi|$，故只须证 $\langle A\psi|\phi\rangle = (\langle\psi|\,A^\dagger)\,|\phi\rangle\ \forall\,|\phi\rangle \in \mathscr{H}$。由上式得

$$(\langle\psi|\,A^\dagger)\,|\phi\rangle = \langle\psi|\,(A^\dagger\,|\phi\rangle) = \langle\psi|A^\dagger\phi\rangle = \langle A\psi|\phi\rangle$$

其中最后一步用到 $A^{\dagger\dagger} = A$。 □

由上可知，任一算符 A 的本征方程 $A\,|\psi\rangle = c\,|\psi\rangle$ 的左矢形式为

$$\langle\psi|\,A^\dagger = \langle\psi|\,\bar{c}$$

设 $\{|e_n\rangle\}$ 是 $\mathscr{H}$ 的正交归一基，则可定义 $\mathscr{H}$ 上的线性算符 $\sum_n |e_n\rangle\langle e_n|$ 为

$$(\sum_n |e_n\rangle\langle e_n|)\,|\psi\rangle \coloneqq \sum_n |e_n\rangle\langle e_n|\psi\rangle \in \mathscr{H},\ \forall\,|\psi\rangle \in \mathscr{H}$$

注意到 $\sum_n |e_n\rangle\langle e_n|\psi\rangle = |\psi\rangle$ 就有

$$\sum_n |e_n\rangle\langle e_n| = I$$

其中 I 代表单位算符（恒等映射），其定义为 $I\,|\psi\rangle \coloneqq |\psi\rangle\ \forall\,|\psi\rangle \in \mathscr{H}$。上式便是量子力学中常用的完备性关系。

6.1.5 态矢和射线

量子系统每一时刻的态由 Hilbert 空间 $\mathscr{H}$ 中的一个矢量（右矢 $|\psi\rangle$）表示，因此右矢叫做态矢。然而态矢与态的对应关系不是一一的。Dirac 说过（大意）: 设由 $|\psi\rangle$ 代表的态与自己叠加，结果将对应于态矢 $c_1|\psi\rangle + c_2|\psi\rangle = (c_1+c_2)|\psi\rangle$，其中 c_1 和 c_2 是任意复数。我们应该假定，除了 $c_1+c_2=0$ 的情况外，结果态 $(c_1+c_2)|\psi\rangle$ 与原始态 $|\psi\rangle$ 相同，即态矢 $(c_1+c_2)|\psi\rangle$ 和 $|\psi\rangle$ 应代表相同的态。（“自己与自己叠加不会得出新态”，请注意这与经典物理非常不同。）就是说，右矢 $|\psi\rangle$ 和 $c|\psi\rangle$（c 为任意非零复数）代表同一状态。于是，若对 $\mathscr{H}$ 的任意非零元素 $|\psi\rangle$ 定义 $\mathscr{H}$ 的子集 $r_\psi \coloneqq \{c|\psi\rangle \mid c\in\mathbb{C}, c\neq 0\}$，并称 r_ψ 为过 $|\psi\rangle$ 的一条射线，则一条射线对应于量子系统的一个态。以 $\mathscr{H}$ 中的所有射线为元素的集合 $\mathscr{R}$ 叫射线空间。

6.2 无界算符及其自伴性

前面一节讨论的是 “$\mathscr{H}$ 上的”“有界” 线性算符。下面先解释这两个定语的含义。

“$\mathscr{H}$ 上的算符（operator on $\mathscr{H}$）” 是指从 $\mathscr{H}$ 到 $\mathscr{H}$ 的映射，其定义域是整个 $\mathscr{H}$。然而量子力学中大多数算符 A 的定义域（记作 D_A）都只能是 $\mathscr{H}$ 的一个真子集，即 A 只能是从 $D_A\subset\mathscr{H}$（$D_A\neq\mathscr{H}$）到 $\mathscr{H}$ 的映射。这称为 $\mathscr{H}$ 中的算符（operator in $\mathscr{H}$）。[6]以 1 维空间的波函数空间 $\mathscr{H}=L^2(\mathbb{R})$ 为例。（如前所述，此乃无限维 Hilbert 空间。）定义位置算符 $X\colon D_X\to L^2(\mathbb{R})$ 为

$$(X\psi)(x) \coloneqq x\psi(x),\ \forall\psi\in D_X, x\in\mathbb{R}$$

上式的含义是：X 作用于 D_X 的任一元素 ψ 的结果是这样一个波函数，其在点 $x\in\mathbb{R}$ 的值等于 ψ 在点 x 的值 $\psi(x)$ 乘以 x。算符 X 的定义域 $D_X\neq L^2(\mathbb{R})$，因为 $\exists\psi\in L^2(\mathbb{R})$ 使 $X\psi\notin L^2(\mathbb{R})$。例如，函数

$$\psi(x)=\begin{cases}1/x, & x\geqslant 1\\ 0, & x<1\end{cases}$$

是平方可积的，但函数

$$(X\psi)(x)=x\psi(x)=\begin{cases}1, & x\geqslant 1\\ 0, & x<1\end{cases}$$

却非平方可积，说明 $\psi\in L^2(\mathbb{R})$ 而 $\psi\notin D_X$。因而位置算符 X 只是 $L^2(\mathbb{R})$ 中（而非 $L^2(\mathbb{R})$ 上）的线性算符。第二个例子是动量算符 $P\colon D_P\to L^2(\mathbb{R})$，其定义为

$$(P\psi)(x)\coloneqq -i\hbar\frac{d\psi(x)}{dx},\ \forall\psi\in D_P$$

[6]这种用 on 和 in 形容算符定义域的约定只在部分文献中采用。在不采用这种约定的文献中，“operator on $\mathscr{H}$” 不表明该算符的定义域是 $\mathscr{H}$。

D_P 是 $L^2(\mathbb{R})$ 的这样的子集，其中每个元素 $\psi(x)$ 几乎处处可微，[7]而且

$$\frac{d\psi(x)}{dx} \in L^2(\mathbb{R})$$

并非 $L^2(\mathbb{R})$ 的每个元素都满足这一条件，所以 $D_P \neq L^2(\mathbb{R})$，即动量算符 P 也只是 $L^2(\mathbb{R})$ 中（而非 $L^2(\mathbb{R})$ 上）的线性算符。

下面介绍有界算符的定义。

定义 6.16. 线性算符 $A\colon D_A(\subset \mathscr{H}) \to \mathscr{H}$ 称为有界算符，若 $\exists M > 0$ 使

$$\sqrt{(Af, Af)} \leqslant M\sqrt{(f, f)},\ \forall f \in D_A$$

否则称为无界算符。

定理 6.7. 设 A 是 $\mathscr{H}$ 上的有界线性算符，则 $A^\dagger$ 也是 $\mathscr{H}$ 上的有界线性算符。

对有限维 Hilbert 空间，所有线性算符都是有界的。反之，量子力学用到的无限维 Hilbert 空间中许多重要线性算符（例如位置算符 X，动量算符 P 和哈氏算符 H）都是无界算符。X 的无界性可证明如下：考虑波函数序列 $\{\psi_1(x), \psi_2(x), \cdots, \psi_n(x), \cdots\}$，其中 $\psi_n(x)$ 定义为

$$\psi_n(x) = \begin{cases} 1, & x \in [n, n+1) \\ 0, & x \notin [n, n+1) \end{cases}$$

则易见 $\psi_n \in D_X$。有

$$(\psi_n, \psi_n) = \int_{-\infty}^{+\infty} \overline{\psi_n(x)}\psi_n(x)dx = \int_n^{n+1} dx = 1$$

另一方面，

$$(X\psi_n, X\psi_n) = \int_{-\infty}^{+\infty} x^2\overline{\psi_n(x)}\psi_n(x)dx = \int_n^{n+1} x^2 dx > n^2$$

因为 n 可为任意大的自然数，所以不存在 $M > 0$ 使

$$\sqrt{(X\psi_n, X\psi_n))} \leqslant M\sqrt{(\psi_n, \psi_n)},\ \forall \psi_n(x)$$

可见 X 无界。

可以证明，有界算符的定义域总可延拓至整个 $\mathscr{H}$，因此讨论有界算符时可只关心 $\mathscr{H}$ 上的算符。然而对无界算符不能如此简化，所以涉及无界算符时要格外注意定义域问题。

设 A 是 $\mathscr{H}$ 中的无界算符，定义域 $D_A \neq \mathscr{H}$。我们遇到的第一个问题是如何定义其伴随算符 $A^\dagger$。回忆上节对 $\mathscr{H}$ 上的算符 A 的对偶算符 A^* 和伴随算符 $A^\dagger$ 的定义。先用下式定义 A^*：

$$(A^*\eta)(f) \coloneqq \eta(Af),\ \forall f \in \mathscr{H}$$

[7]关于 D_P 的进一步限制可见后文。

再用 A^* 定义 $A^\dagger$。可以证明这样定义的 A^*（和 $A^\dagger$）是 $\mathscr{H}^*$ 上（和 $\mathscr{H}$ 上）的有界算符。然而，A^* 的上述定义对 $\mathscr{H}$ 中的算符 A 不适用。由于 $D_A \neq \mathscr{H}$，Af 对于不在 D_A 中的 f 无意义，因此上式中的 “$\forall f \in \mathscr{H}$” 应改为 “$\forall f \in D_A$”。然而这样改后的式子就不成其为 A^* 的定义，因为要定义 $\mathscr{H}^*$ 的元素 $A^*\eta$ 必须定义它对 $\mathscr{H}$ 的每一元素的作用。可见当 $D_A \neq \mathscr{H}$ 时 $A^\dagger$ 需要另给定义。注意到其等价定义，可考虑用下法直接定义 $A^\dagger$。对给定的 $f \in \mathscr{H}$，若存在唯一的 $h_f \in \mathscr{H}$ 使

$$(f, Ag) = (h_f, g),\ \forall g \in D_A$$

就把 $A^\dagger f$ 定义为 h_f，即

$$A^\dagger f \coloneqq h_f$$

$A^\dagger$ 的定义域自然为

$$D_{A^\dagger} = \{f \in \mathscr{H} \mid \exists \text{唯一} h_f \in \mathscr{H} \text{使} (f, Ag) = (h_f, g)\ \forall g \in D_A\}$$

用此法定义 $A^\dagger$ 的可能性取决于满足式 $(f, Ag) = (h_f, g),\ \forall g \in D_A$ 的 h_f 的唯一性，其充要条件将由下述定理给出。为讲此定理先介绍两个术语：拓扑空间 X 的子集 $U \subset X$ 称为稠密子集，若 $\bar{U} = X$（$\bar{U}$ 代表 U 的闭包）；Hilbert 空间 $\mathscr{H}$ 中的算符 A 称为稠定的，若 $\overline{D_A} = \mathscr{H}$。

定理 6.8. 设 A 是 $\mathscr{H}$ 中的线性算符，$f \in \mathscr{H}$，则满足

$$(f, Ag) = (h, g),\ \forall g \in D_A$$

的 $h \in \mathscr{H}$ 是唯一的当且仅当 A 是稠定的。

由此可知，当且仅当 $\mathscr{H}$ 中的算符 A 为稠定时，其伴随算符 $A^\dagger$ 可由前式定义，其定义域亦如前所述。我们有

$$(f, Ag) = (A^\dagger f, g),\ \forall g \in D_A, f \in D_{A^\dagger}$$

不难证明这样定义的 $A^\dagger$ 是线性的，而且若 A 是 $\mathscr{H}$ 上的有界算符，则这样定义的 $A^\dagger$ 与上节定义的 $A^\dagger$ 相同（这时 $A^\dagger$ 也是 $\mathscr{H}$ 上的有界算符）。考虑到上述定理，今后凡谈到算符 A 的伴随算符 $A^\dagger$ 时都默认 A 是稠定的。可以证明位置和动量算符都是稠定算符。应该指出，D_A 稠密并不保证 $D_{A^\dagger}$ 稠密。在最坏的情况下，$D_{A^\dagger}$ 可以 “小” 到只含零元的程度，即 $D_{A^\dagger} = \{0\}$。幸好这种情况很少出现。定义域不稠密的无界算符在量子力学中没有什么用处。

算符的厄米性对量子力学的重要性是众所周知的。上节已对 $\mathscr{H}$ 上的算符 A 的厄米性下了如下定义：

$$(f, Ag) = (Af, g),\ \forall f, g \in \mathscr{H}$$

$\mathscr{H}$ 中的算符 $A\colon D_A \to \mathscr{H}$ 的厄米性仍可用上式定义，只须把式中的 $\forall f, g \in \mathscr{H}$ 改为 $\forall f, g \in D_A$。然而，泛函分析有个 Hellinger–Toeplitz 定理，它断言 Hilbert 空间 $\mathscr{H}$ 上满足

上式的算符 A 必然有界。这就导致如下严酷的结论：无界的厄米算符的定义域不可能是全 $\mathscr{H}$。因此涉及无界厄米算符时必须格外注意定义域问题。下面将看到，无界算符的厄米性和自伴性的关键区别就在于定义域。

设 A, B 是 $\mathscr{H}$ 上的算符，则其和、积及相等的定义十分简单：

和 $(A+B)f \coloneqq Af + Bf,\ \forall f \in \mathscr{H}$，

积 $(AB)f \coloneqq A(Bf),\ \forall f \in \mathscr{H}$,

相等 $A = B \Leftrightarrow Af = Bf,\ \forall f \in \mathscr{H}$。

然而对 $\mathscr{H}$ 中的算符就要麻烦一些。例如，$A+B$ 的定义域只能是 $D_A \cap D_B$，而 AB 的定义域则还涉及 B 的值域，因为只当 $Bf \in D_A$ 时 ABf 才有意义。A 与 B 的相等性和相互包含性则由如下定义规定：

定义 6.17. 设 A, B 是 $\mathscr{H}$ 中的线性算符，其定义域分别为 D_A 和 D_B，则

(*a*) $A = B$ 若 $D_A = D_B$ 而且 $Af = Bf,\ \forall f \in D_A = D_B$；

(*b*) $A \subset B$ 若 $D_A \subset D_B$ 而且 $Af = Bf,\ \forall f \in D_A$（把 B 称为 A 的延拓或扩张）。

定义 6.18. $\mathscr{H}$ 中的任意（有界或无界）稠定线性算符 A 称为厄米的，若

$$(f, Ag) = (Af, g),\ \forall f, g \in D_A$$

定理 6.9. $\mathscr{H}$ 中的稠定算符 A 为厄米算符的充要条件是 $A \subset A^\dagger$。[8]

证明.（A） 设 A 为厄米，则 $\forall f \in D_A$，有 $(f, Ag) = (h_f, g)\ \forall g \in D_A$。对比 $D_{A^\dagger}$ 的定义发现 $f \in D_{A^\dagger}$（其中的 h_f 的唯一性由 A 的稠定性保证，h_f 就是 Af）。所以，$f \in D_A \Rightarrow f \in D_{A^\dagger}$，因而 $A \subset A^\dagger$。

（B） 设 $A \subset A^\dagger$，则 $D_A \subset D_{A^\dagger}$ 且 $\forall f \in D_A$ 有 $A^\dagger f = Af$，故

$$(f, Ag) = (Af, g),\ \forall f, g \in D_A$$

可见 A 是厄米的。

□

注 6.9. 讨论表明（此处只介绍结论），若厄米算符 A 有界，可把 A 唯一地延拓为 $\mathscr{H}$ 上的有界算符，仍记作 A（唯一性由 D_A 的稠密性保证）。同样，$A^\dagger$ 也可被唯一地延拓为 $\mathscr{H}$ 上的有界算符，而且 $(f, Ag) = (Af, g) = (A^\dagger f, g)\ \forall f, g \in \mathscr{H}$。可见上述命题中的 $A \subset A^\dagger$ 对有界厄米算符 A 可表为 $A = A^\dagger$。然而确实存在 $D_A \neq D_{A^\dagger}$（因而 $A \neq A^\dagger$）的无界厄米算符 A，对这种算符一般只能写 $A \subset A^\dagger$。这说明对无界算符 A 来说，$A = A^\dagger$ 是比 $A \subset A^\dagger$ 更高的要求。满足 $A = A^\dagger$ 的算符非常重要，值得赋予专名：

[8] A 为厄米不保证 $A^\dagger$ 为厄米。事实上，对厄米算符 A 有 $A \subset A^{\dagger\dagger} \subset A^\dagger$（却未必有 $A^\dagger \subset A^{\dagger\dagger}$）。

定义 6.19. $\mathscr{H}$ 中的任意（有界或无界）稠定算符 A 称为自伴的，若 $A = A^\dagger$。

由此可知，无界算符的自伴性强于厄米性。偏偏量子力学中多数算符是无界算符，因此区分厄米性和自伴性就成为重要问题。厄米性和自伴性的关键差别在于定义域，所以对量子力学的算符应该特别注意定义域及其有关问题（例如定义域的延拓或压缩以及算符的性质在定义域延拓或压缩时的可能改变）。讨论量子力学问题时出现的若干混淆正是不注意算符的定义域问题所致。

定理 6.10. $L^2(\mathbb{R})$ 中的位置算符 $X\colon D_X \to L^2(\mathbb{R})$ 是自伴算符。

证明. 前已述及 X 是稠定的，由下式易见 X 是厄米的：

$$(f, Xg) = \int_{-\infty}^{+\infty} \overline{f(x)} x g(x) dx = \int_{-\infty}^{+\infty} \overline{xf(x)} g(x) dx = (Xf, g)$$

因此有 $X \subset X^\dagger$，余下只须证明 $X^\dagger \subset X$。设 $f \in D_{X^\dagger}$，则根据伴随算符的定义有 $(f, Xg) = (X^\dagger f, g)\ \forall g \in D_X$，即

$$\int_{-\infty}^{+\infty} \overline{f(x)} x g(x) dx = \int_{-\infty}^{+\infty} \overline{(X^\dagger f)(x)} g(x) dx$$

故

$$\int_{-\infty}^{+\infty} [\overline{f(x)} x - \overline{(X^\dagger f)(x)}] g(x) dx = 0,\ \forall g \in D_X, f \in D_{X^\dagger}$$

因为在有界区间 (a, b) 外为零的任一函数都属于 D_X，可取

$$g(x) = \begin{cases} xf(x) - (X^\dagger f)(x), & x \in (a, b) \\ 0, & x \notin (a, b) \end{cases}$$

从而得 $\int_b^a |xf(x) - (X^\dagger f)(x)|^2 dx = 0$，故在 (a, b) 上[9]有

$$xf(x) = (X^\dagger f)(x)$$

因 (a, b) 任意，故 $\forall x \in \mathbb{R}$ 有 $xf(x) = (X^\dagger f)(x)$。而 $f \in D_{X^\dagger}$ 保证 $X^\dagger f \in L^2(\mathbb{R})$，所以上式可改写为 $(Xf)(x) = (X^\dagger f)(x)$。可见 $f \in D_{X^\dagger}$ 导致 $f \in D_X$。加之 $Xf = X^\dagger f\ \forall f \in D_{X^\dagger}$，便有 $X^\dagger \subset X$。 □

定理 6.11. $L^2(\mathbb{R})$ 中的动量算符 $P\colon D_P \to L^2(\mathbb{R})$ 是自伴算符。

注 6.10. P 的定义式中的 $-i$ 对保证厄米性[10]有关键作用。

[9]更准确的提法是 “在 (a, b) 上几乎处处”，即测度为零的集可以例外。

[10]证明厄米性需要分部积分。为此对 P 的定义域 D_P 要作进一步限制：D_P 只含 $L^2(\mathbb{R})$ 中这样的函数，它们在每一有界区间 $[a, b]$ 上绝对连续且导数属于 $L^2(\mathbb{R})$。可证 D_P 是稠密的。

知道 X 和 P 是自伴算符后，自然希望哈氏算符 H 也是自伴算符。虽然动能和势能算符的自伴性的证明相对简单，作为它们之和的哈氏算符的自伴性的证明却要困难得多（这也许有点出人意料）。幸好，泛函分析中存在各种定理（如 Kato 定理），在许多场合下可用以证明哈氏算符的自伴性。在一般性讨论中则假定哈氏算符是自伴算符。

无论从历史发展还是当今教学的角度看，如果一味追求数学严格性，量子力学也许寸步难行。物理学家通常采用的实际可行的做法是默认有限维的结论也适用于无限维（把有限项之和改为无限项之和或积分），借助于这种默认并配以 Dirac 巧妙发明的 δ 函数，往往可以简单快捷地得出正确结果。不妨称这种做法为物理做法。然而学生在学习过程中可能出现某些困扰，甚至达到百思不解的程度。例如，教科书上都说动量算符 P 和位置算符 X 的本征矢（本征函数）构成 Hilbert 空间的正交归一基底，P 的本征矢是平面波 $\exp(i\hbar^{-1}\vec{p}\cdot\vec{r})$，X 的本征矢是 $\delta^3(\vec{r}-\vec{r}_0)$。然而前者因为非平方可积而不属于 Hilbert 空间（指 $L^2(\mathbb{R}^3)$），后者则根本不是通常意义下的函数，更谈不上属于 $L^2(\mathbb{R}^3)$。连空间的元素都不是，怎么竟成为该空间的基矢？这正如把不是委员的人说成是常委那样不可思议。由此还会引申出许多难以回答的问题，例如：谐振子的波函数 $|\psi\rangle$ 既可用能量本征矢表为分立的取和式 $|\psi\rangle=\sum_{n=1}^{\infty}\psi_n|n\rangle$，也可用位置（或动量）本征矢表为连续的取和式 $|\psi\rangle=\int_{-\infty}^{+\infty}|x\rangle\langle x|\psi\rangle\,dx$，这个 Hilbert 空间的维数如何考虑？（虽然都是 ∞，但可数的 ∞ 与不可数的 ∞ 并不相同。）其实，泛函分析对所有这些问题都能给出数学上严密的回答（因而可充当“后台”）。下面略做一点粗浅介绍。

众所周知，有限维 Hilbert 空间 $\mathscr{H}$ 的厄米（因而自伴）算符 A 存在有限个实本征值，全部线性独立的本征矢组成 $\mathscr{H}$ 的一组正交归一基底，任一矢量 $f\in\mathscr{H}$ 可用此基底展开，任一算符可用此基底表为 $N\times N$ 矩阵（$N\equiv\dim\mathscr{H}$）。特别是，A 自己在此基底下的矩阵是对角矩阵，且对角元为实本征值。可惜这些既简单又实用的结论对无限维 Hilbert 空间并不成立。幸好无限维 Hilbert 空间中的自伴算符（不论有界无界）具有一系列好的性质，足以保证默认有限维的结论适用于无限维的许多物理做法给出正确结果。例如，自伴算符的确可以实现“对角化”，不过不是表现为有限维空间中那样简单的形式，因此最好称为广义对角化。在介绍广义对角化之前，有必要说明无限维 Hilbert 空间在涉及自伴算符的问题上与有限维 Hilbert 空间的诸多区别。首先，无限维 Hilbert 空间中无界的厄米算符未必是自伴算符，而非自伴的厄米算符未必能实现广义对角化。其次，即使是自伴算符，也未必存在本征值和本征矢，更不用说本征矢构成基底。以最常用的 $L^2(\mathbb{R}^3)$ 中的位置、动量和哈氏算符为例。位置算符 X 和动量算符 P 虽然都是自伴算符，却根本没有本征值和本征矢（本征函数），因为根本不存在 $\lambda\in\mathbb{C}$ 和 $\psi\in L^2(\mathbb{R}^3)$ 满足本征方程 $X\psi=\lambda\psi$ 和 $P\psi=\lambda\psi$ （如前所述，$\exp(i\hbar^{-1}\vec{p}\cdot\vec{r})$ 和 $\delta^3(\vec{r}-\vec{r}_0)$ 都不属于 $L^2(\mathbb{R}^3)$）。哈氏算符 H 的情况则有所不同：某些系统（如谐振子）的 H 有完备的本征矢（分立的能量本征态），而某些系统（如自由质点）则没有：其 H 的本征方程的解是平面波，因而不是 $L^2(\mathbb{R}^3)$ 的元素（矢量），更谈不上是本征矢。[11] 第三，即使自伴算符存在本征值和本征矢，其本征矢也未必足以构成基底，只用本

[11] 并非不能用平面波展开波函数，事实上，对任意波包作平面波展开不但允许（无非是把函数写成傅里叶积分）而且很重要，只是不应称之为用基底展开。根据泛函分析，Hilbert 空间必有基底，因而任一元素可用基底展开，但展开式一定是可数项之

征值也未必足以解决算符的广义对角化问题。解决问题的根本途径是依靠本征值概念的推广—谱—及其有关定理。然而谱定理不适用于非自伴的厄米算符，这就是严格规定可观察量必须是自伴算符（而不只是厄米算符）的理由之一。（然而通常的物理讲法却把可观察量定义为有完备本征矢的厄米算符。现在可看出这种讲法的不严格性。）顺便一提，作为自伴算符重要性的另一例子，我们指出只有自伴算符 A 才可用指数方式生成一个决定量子系统动力学的单参酉群 $U(t) = e^{itA}$，它在量子力学中的重要作用是人所共知的。谱理论是泛函分析中的一个精彩而重要的部分，已超出本章范围。

和（$\psi = \sum_{n=1}^{\infty} c_n e_n$）而绝决不是积分。（就连不可分的 Hilbert 空间（它只有不可数基底）也如此，更何况动量算符涉及的 Hilbert 空间 $L^2(\mathbb{R}^3)$ 是可分空间。）

第七章　李群和李代数

7.1　群论初步

定义 7.1. 集合 G 配以满足以下条件的映射 $G\times G\to G$（叫群乘法）称为群：

（a）　$(g_1g_2)g_3=g_1(g_2g_3)\ \forall g_1,g_2,g_3\in G$

（b）　存在恒等元 $e\in G$，使 $eg=ge=g$，$\forall g\in G$

（c）　$\forall g\in G$，存在逆元 $g^{-1}\in G$，使 $g^{-1}g=gg^{-1}=e$

注 7.1. 恒等元是唯一的[1]，任一群元的逆元也是唯一的[2]。

定义 7.2. 乘法满足交换律的群（即 $gh=hg\ \forall g,h\in G$）称为阿贝尔群。只含有限个元素的群叫有限群，否则叫无限群。群 G 的子集 H 称为 G 的子群，若 H 用 G 的乘法为乘法也构成群。

定义 7.3. 设 G 和 G' 是群。映射 $\mu\colon G\to G'$ 叫同态，若

$$\mu(g_1g_2)=\mu(g_1)\mu(g_2),\forall g_1,g_2\in G$$

定理 7.1. 同态映射 $\mu\colon G\to G'$ 有以下性质：

（a）　若 e,e' 各为 G,G' 的恒等元，则 $\mu(e)=e'$[3]

（b）　$\mu(g^{-1})=\mu(g)^{-1},\forall g\in G$[4]

（c）　$\mu[G]$ 是 G' 的子群[5]，当 G 是阿贝尔群时 $\mu[G]$ 是 G' 的阿贝尔子群[6]。

[1]考虑有两个恒等元 e_1,e_2，有 $e_1e_2=e_1=e_2$

[2]考虑任一群元 g 有两个逆元 g_1^{-1},g_2^{-1}，有 $g_1^{-1}=(g_2^{-1}g)g_1^{-1}=g_2^{-1}(gg_1^{-1})=g_2^{-1}$

[3]$\mu(g)=\mu(ge)=\mu(g)\mu(e)$，$\forall\mu(g)\in G'$，考虑到恒等元的唯一性，则必有 $\mu(e)=e'$。

[4]$\mu(g^{-1})=\mu(g^{-1})e'=\mu(g^{-1})\mu(g)\mu(g)^{-1}=\mu(e)\mu(g)^{-1}=\mu(g)^{-1}$

[5]显然 $\mu[G]$ 是 G' 的子集（否则 μ 就不构成 $G\to G'$ 的映射）。按 μ 的定义，容易验证 $\mu[G]$ 上的乘法是结合的，单位元是 $\mu(e)$，$\mu(g)$ 逆元为 $\mu(g^{-1})$

[6]若 G 上的乘法可交换，则 μ 的定义保证 $\mu[G]$ 上的乘法也可交换。注意这并不说明 G' 上的乘法可交换。

定义 7.4. 一一到上的同态映射称为同构。当有可能与矢量空间之间的同构混淆时又明确地把群之间的同构称为群同构。同构 $\mu\colon G \to G$ 称为群 G 上的自同构。

例 7.1. $\forall g \in G$，可构造一个称为伴随同构的自同构映射，又称内自同构，记作 $I_g\colon G \to G$，定义[7]为

$$I_g(h) \coloneqq ghg^{-1}, \forall h \in G$$

注 7.2. 今后常把两个同构的群视作一样，并用等号表示。

定义 7.5. 群 G 和 G'（看作两个集合）的卡氏积 $G \times G'$ 按下列乘法

$$(g_1, g_1')(g_2, g_2') \coloneqq (g_1 g_2, g_1' g_2'),\ \forall g_1, g_2 \in G,\ g_1', g_2' \in G'$$

构成的群称为 G 和 G' 的直积群。

例 7.2. 以加法为群乘法，则 $\mathbb{R}$ 是群。$\mathbb{R}^2 \equiv \mathbb{R} \times \mathbb{R}$ 配以由上式定义的乘法就构成直积群，而且此乘法正好是 $\mathbb{R}^2$ 上（自然定义）的加法。

定义 7.6. 设 H 是群 G 的子群，$g \in G$，则 $gH \equiv \{gh \mid h \in H\}$ 称为 H 的含 g 的左陪集。类似地可定义右陪集。

注 7.3. 若子群 H 的两个左陪集有交，则两者必相等[8]。

定义 7.7. 群 G 的子群 H 称为正规子群或不变子群，若

$$ghg^{-1} \in H,\ \forall g \in G, h \in H$$

定义 7.8. 设 G 是群，则 $A(G) \equiv \{\mu\colon G \to G \mid \mu\text{为自同构映射}\}$ 以映射的复合为群乘法构成群，称为群 G 的自同构群。“以映射的复合为乘法”是指 $\forall \mu, \nu \in A(G)$，群乘积 $\mu\nu \in A(G)$ 定义为 $(\mu\nu)(g) \equiv \mu(\nu(g))\ \forall g \in G$。

定理 7.2. 以 $A_I(G)$ 代表 G 上全体内自同构映射的集合，即

$$A_I(G) \equiv \{I_g\colon G \to G \mid g \in G\} \subset A(G)$$

则 $A_I(G)$ 是群 $A(G)$ 的一个正规子群。

证明. $A_I(G)$ 显然是 $A(G)$ 的子集，并且容易证明，以映射复合为群乘法，$A_I(G)$ 也构成群，所以 $A_I(G)$ 是 $A(G)$ 的子群。其恒等元为 I_e，其相应于 I_g 的逆元为 $I_{g^{-1}}$。为了验证它还是正规子群，考虑 $(\mu I_g \mu^{-1})(h) = \mu(g(\mu^{-1}(h))g^{-1}) = \mu(g)\mu(\mu^{-1}(h))\mu(g^{-1}) = I_{\mu(g)}(h), \forall \mu \in A(G), I_g \in A_I(G), h \in G$。所以有 $\mu I_g \mu^{-1} = I_{\mu(g)} \in A_I(G), \forall \mu \in A(G), I_g \in A_I(G)$。 □

[7] 当然应该验证这一定义符合同构映射的要求。首先 $I_g(h_1 h_2) = I_g(h_1) I_g(h_2)$ 说明这是同态映射。其次，$I_g(h_1) = I_g(h2) \Rightarrow h_1 = h_2, \forall h_1, h_2 \in G$ 说明该映射是一一的。最后，$I_g(g^{-1}hg) = h, \forall h \in G$ 说明该映射是到上的。

[8] 设 $p = g_1 h_1 = g_2 h_2$ 是 $g_1 H \cap g_2 H$ 中的一个元素。$\forall x = g_1 h_x \in g_1 H$，有 $x = g_1 h_1 h_1^{-1} h_x = g_2 h_2 h_1^{-1} h_x \in g_2 H$，反之亦然。

定义 7.9. 设 H 和 K 是群，且存在同态映射 $\mu\colon K \to A(H)$。$\forall k \in K$，把 $\mu(k) \in A(H)$ 简记作 μ_k，则 $G \equiv H \times K$ 配以下式定义的群乘法

$$(h, k)(h', k') := (h\mu_k(h'), kk'),\ \forall h, h' \in H, k, k' \in K$$

所构成的群称为 H 和 K 的半直积群[9]，记作 $G \equiv H \otimes_S K$。

7.2 李群

定义 7.10. 若 G 既是 n 维（实）流形又是群，其群乘映射 $G \times G \to G$（请注意 $G \times G$ 也是流形）和求逆元映射 $G \to G$ 都是 C^∞ 的，则 G 叫 n 维（实）李群[10]。

例 7.3. 以加法为群乘法，则 $\mathbb{R}$ 是 1 维李群。

例 7.4. $\mathbb{R}$ 和 $\mathbb{R}$ 的直积群 $\mathbb{R}^2$ 是 2 维李群。推而广之，$\mathbb{R}^n$ 是 n 维李群。

例 7.5. 设 $\phi\colon \mathbb{R} \times M \to M$ 是流形 M 上的任一单参微分同胚群，则 $\{\phi_t \mid t \in \mathbb{R}\}$ 是 1 维李群[11]，同构于 $\mathbb{R}$。

例 7.6. 易证广义黎曼空间 (M, g_{ab}) 上的两个等度规映射的复合也是等度规映射，因此 (M, g_{ab}) 上全体等度规映射的集合以复合映射为乘法构成群，称为 (M, g_{ab}) 的等度规群。还可验证等度规群是李群。闵氏时空的等度规群是 10 维李群，施瓦西时空的等度规群是 4 维李群。一般地，n 维广义黎曼空间 (M, g_{ab}) 的等度规群的维数 $m \leqslant n(n+1)/2$。但 (M, g_{ab}) 上全体微分同胚的集合则“大”到不能构成有限群。事实上，它是一个无限维群。

注 7.4. 本章只限于讨论有限维李群，虽然许多结论对无限维李群也适用。

以下如无特别声明，G 一律代表李群。李群的双重身份（即是群又是流形）使得用几何语言研究李群成为可能。群乘映射和求逆元映射的光滑性则使李群具有一系列好性质。

定义 7.11. 李群 G 和 G' 之间的 C^∞ 同态映射 $\mu\colon G \to G'$ 称为李群同态。李群同态 μ 称为李群同构，若 μ 为微分同胚。

---

[9] 真是一个巧妙的定义。

结合律：$[(h, k)(h', k')](h'', k'') = (h\mu_k(h'), kk')(h'', k'') = (h\mu_k(h')\mu_{kk'}(h''), kk'k'')$ 而 $(h, k)[(h', k')(h'', k'')] = (h, k)(h'\mu_{k'}(h''), k'k'') = (h\mu_k(h'\mu_{k'}(h'')), kk'k'') = (h\mu_k(h')\mu_k\mu_{k'}(h''), kk'k'') = (h\mu_k(h')\mu_{kk'}(h''), kk'k'')$ 最后两个等号是因为 $A(H)$ 中的元素皆为同态映射，而且 μ 本身也是同态映射（因此 $\mu(ab) = \mu(a)\mu(b)$）。

单位元：注意到同态映射保单位元，而 $A(H)$ 中的单位元就是恒等映射，故 $(e_h, e_k)(h, k) = (e_h\mu_{e_k}(h), e_k k) = (e_h h, e_k k)$ 而 $(h, k)(e_h, e_k) = (h\mu_k(e_h), ke_k) = (he_h, ke_k)$

逆元：$(\mu_{k-1}(h^{-1}), k^{-1})(h, k) = (\mu_{k-1}(h^{-1})\mu_{k-1}(h), k^{-1}k) = (e_h, e_k)$ 而 $(h, k)(\mu_{k-1}(h^{-1}), k^{-1}) = (h\mu_k(\mu_{k-1}(h^{-1})), kk^{-1}) = (e_h, e_k)$

[10] 约定把有离散拓扑的可数群称为零维李群。因为有限群的默认拓扑是离散拓扑，所以有限群都可看作零维李群。

[11] $\phi(t, p) = p\ \forall p \in M, t \in \mathbb{R}$ 的情况可视为例外。这是与 M 上的 0 矢量场对应的那个特殊的单参微分同胚群，是只含恒等元的独点群，可看作零维李群。

定义 7.12. 李群 G 的子集 H 称为 G 的李子群，若 H 既是 G 的子流形又是 G 的子群。

定义 7.13. $\forall g \in G$，映射 $L_g : h \mapsto gh\ \forall h \in G$ 叫做由 g 生成的左平移。

注 7.5. ① 由李群定义中关于群乘映射和求逆元映射的 C^∞ 性可知左平移 $L_g : G \to G$ 是微分同胚映射。② 易见 $L_{gh} = L_g \circ L_h$。

以下的讨论经常涉及 G 的一点的矢量和 G 的一个子集上的矢量场，并要对两者作明确区分。我们将用 $A, B, \cdots$ 代表一点的矢量，用 $\bar{A}, \bar{B}, \cdots$ 代表矢量场，用 $\bar{A}_g$ 代表矢量场 $\bar{A}$ 在点 $g \in G$ 的值。为简化表达式，本章中所有矢量（除少数情况外）都不加抽象指标。

定义 7.14. G 上的矢量场 $\bar{A}$ 叫左不变的，若

$$L_{g*}\bar{A} = \bar{A},\ \forall g \in G$$

其中 L_{g*} 是由左平移映射 $L_g : G \to G$ 诱导的推前映射。

注 7.6. ① 左不变矢量场必为 C^∞ 矢量场；② 不难看出左不变矢量场的定义式等价于

$$(L_{g*}\bar{A})_{gh} = \bar{A}_{gh}$$

推前映射一般只能把一点的张量映射为张量，但当 $\phi : M \to N$ 是微分同胚时，ϕ_* 可把 M 上张量场 v 映为 N 上张量场 $\phi_* v$，定义为 $(\phi_* v)|_{\phi(p)} = \phi_*(v|_p)\ \forall p \in M$。用于现在的情况，就是 $(L_{g*}\bar{A})_{gh} = L_{g*}(\bar{A}_h)$。于是上式又等价于

$$\bar{A}_{gh} = L_{g*}(\bar{A}_h)$$

上式可作为左不变矢量场 $\bar{A}$ 的等价定义。

不难看出左不变矢量场之和以及左不变矢量场乘以常数仍为左不变矢量场，故 $\mathscr{L} \equiv \{\bar{A} \mid \bar{A}\text{是 } G \text{ 上的左不变矢量场}\}$ 是矢量空间。

定理 7.3. G 上全体左不变矢量场的集合 $\mathscr{L}$ 与 G 的恒等元 e 的切空间 V_e（作为两个矢量空间）同构。

证明. $\forall A \in V_e$，用下式定义 G 上的矢量场 $\bar{A}$:

$$\bar{A}_g \coloneqq L_{g*}A,\ \forall g \in G$$

由此得 $\bar{A}_e = A$。把上式的 g 改为 gh 得

$$\bar{A}_{gh} = L_{gh*}A = (L_g \circ L_h)_* A = (L_{g*} \circ L_{h*})A = L_{g*}(L_{h*}A) = L_{g*}\bar{A}_h$$

因此[12] $\bar{A}$ 是左不变矢量场。可见上式定义了一个映射 $\eta\colon V_e \to \mathscr{L}$（把 A 映为 $\bar{A}$）。L_{g*} 的线性性保证 η 的线性性，由 $\bar{A}_e = A$ 易见[13] η 是一一映射。于是欲证 η 为同构只须证 η 为到上映射。$\forall \bar{A} \in \mathscr{L}$，有 $\bar{A}_e \in V_e$。注意到 $\eta(\bar{A}_e)|_g = L_{g*}\bar{A}_e = \bar{A}_{ge} = \bar{A}_g,\ \forall g \in G$。这说明 $\eta(\bar{A}_e) = \bar{A}\ \forall \bar{A} \in \mathscr{L}$。即 η 为到上映射[14]。 □

7.3 李代数

在矢量空间 $\mathscr{V}$ 上定义某种称为“乘法”的映射就得到一个代数。一种重要的乘法叫李括号，记作 $[,]\colon \mathscr{V} \times \mathscr{V} \to \mathscr{V}$，它是满足以下两条件的双线性映射：

（a） $[A, B] = -[B, A],\ \forall A, B \in \mathscr{V}$

（b） $[A, [B, C]] + [C, [A, B]] + [B, [A, C]] = 0,\ \forall A, B, C \in \mathscr{V}$

第二个条件称为雅克比恒等式。

定义 7.15. 定义了李括号的矢量空间称为李代数。任意两个元素的李括号都为零的李代数称为阿贝尔李代数。

本章只讨论有限维实矢量空间上的李代数（实李代数），虽然许多结论对无限维（实或复）李代数也适用。

例 7.7. 把 $\mathbb{R}^3$ 看作 3 维矢量空间，用下式定义李括号

$$[\vec{v}, \vec{u}] \coloneqq \vec{v} \times \vec{u}, \forall \vec{v}, \vec{u} \in \mathbb{R}^3$$

则 $\mathbb{R}^3$ 成为李代数。

例 7.8. $\mathscr{M} \equiv \{m \times m\text{矩阵}\}$ 显然为 m^2 维矢量空间。用矩阵对易子定义李括号，即

$$[A, B] \coloneqq AB - BA, \forall A, B \in \mathscr{M}, \text{(其中 } AB \text{ 是 } A, B \text{ 的矩阵积)}$$

则 $\mathscr{M}$ 是李代数。

定理 7.4. G 上全体左不变矢量场的集合 $\mathscr{L}$ 是李代数。

证明. 以矢量场对易子为李括号。因为 $\forall \bar{A}, \bar{B} \in \mathscr{L}$ 有

$$L_{g*}[\bar{A}, \bar{B}] = [L_{g*}\bar{A}, L_{g*}\bar{B}] = [\bar{A}, \bar{B}]$$

---

[12] 关于复合映射的推前等于推前映射的复合，证明如下：$(L_g \circ L_h)_* A(f) \coloneqq A((L_g \circ L_h)^* f)\ \forall f \in \mathscr{F}$，而 $(L_g \circ L_h)^* f|_p \coloneqq f|_{(L_g \circ L_h)(p)} = f|_{L_g(L_h(p))} = L_g^* f|_{L_h(p)} = L_h^*(L_g^* f)|_p = (L_h^* \circ L_g^*) f|_p\ \forall p$。因此 $(L_g \circ L_h)_* A(f) = A((L_h^* \circ L_g^*) f)\ \forall f \in \mathscr{F}$。另一方面，$[(L_{g*} \circ L_{h*})A](f) = [L_{g*}(L_{h*}A)](f) = (L_{h*}A)(L_g^* f) = A(L_h^* L_g^* f) = A((L_h^* \circ L_g^*) f)\ \forall f \in \mathscr{F}$

[13] 假设有 $A \neq B$，则 $\eta(A) = \bar{A}, \eta(B) = \bar{B}$，因其在 e 点不等（$\bar{A}_e = A \neq B = \bar{B}_e$），故 $\bar{A} \neq \bar{B}$。

[14] 就是说，任给一个 $\bar{A} \in \mathscr{L}$，都可以找到其原像 $\bar{A}_e \in V_e$。

所以 $[\bar{A},\bar{B}]\in\mathscr{L}\ \forall\bar{A},\bar{B}\in\mathscr{L}$，可见对易子的确是从 $\mathscr{L}\times\mathscr{L}$ 到 $\mathscr{L}$ 的映射。对易子当然是双线性的和反称的。而且可以验证对易子也满足雅克比恒等式。 □

定义 7.16. 设 $\mathscr{V}$ 和 $\mathscr{W}$ 是李代数。线性映射 $\beta\colon\mathscr{V}\to\mathscr{W}$ 称为李代数同态，若它保李括号，即 $\beta([A,B])=[\beta(A),\beta(B)]\ \forall A,B\in\mathscr{V}$。李代数同态 $\beta\colon\mathscr{V}\to\mathscr{W}$ 称为李代数同构，若 β 是一一到上映射。

注 7.7. 今后常把两个同构的李代数视作相同，并用等号表示。

对李群 G 的恒等元 e 的切空间 V_e 用下式定义李括号：

$$[A,B]\coloneqq[\bar{A},\bar{B}]_e,\ \forall A,B\in V_e$$

（其中 $\bar{A},\bar{B}$ 分别是 A,B 对应的左不变矢量场。）则 V_e 成为李代数[15]，称为李群 G 的李代数，记作 $\mathscr{G}$。易证 $\mathscr{G}$ 与 $\mathscr{L}$ 有李代数同构关系，η 可充当同构映射。反之，给定一个李代数，是否也可找到一个李群，它的李代数是所给的李代数？答案是：这样的李群一定存在，并且唯一到只差整体拓扑结构的程度。（例如，以流形 S^1 上的角坐标之和作为群乘法，则 S^1 是 1 维李群，它与 1 维李群 $\mathbb{R}$ 不同—有不同的整体拓扑—但却有相同的李代数。第 6 节末还将给出有相同李代数的不同李群的另外两个重要例子。）准确地说，给定一个李代数，总可找到唯一的单连通李群（其流形为单连通流形[16]的李群），它以所给李代数为李代数。这是李群理论中的一个重要定理。李群和李代数的这一密切联系使李群的讨论大为简化，因为李代数比李群简单得多。

定理 7.5. 设 $\mathscr{G}$ 和 $\hat{\mathscr{G}}$ 分别是李群 G 和 $\hat{G}$ 的李代数，$\rho\colon G\to\hat{G}$ 是同态映射，则 ρ 在点 $e\in G$ 诱导的推前映射 $\rho_*\colon\mathscr{G}\to\hat{\mathscr{G}}$ 是李代数同态。

定义 7.17. 李代数 $\mathscr{G}$ 的子空间 $\mathscr{H}$ 称为 $\mathscr{G}$ 的李子代数，若

$$[A,B]\in\mathscr{H},\ \forall A,B\in\mathscr{H}$$

其中 $[A,B]$ 是把 A,B 看作 $\mathscr{G}$ 的元素时的李括号，现在也称为子代数 $\mathscr{H}$ 的李括号。

定理 7.6. 设 H 是李群 G 的李子群，则 H 的李代数 $\mathscr{H}$ 是 $\mathscr{G}$ 的李子代数。

定义 7.18. 李代数 $\mathscr{G}$ 的子代数 $\mathscr{H}$ 称为 $\mathscr{G}$ 的理想，若

$$[A,\mu]\in\mathscr{H},\ \forall A\in\mathscr{G},\mu\in\mathscr{H}$$

注 7.8. 理想在李代数理论中的角色相当于正规子群在群论中的角色。

定理 7.7. 设 $\mathscr{H}\subset\mathscr{G}$ 是理想，$\mathscr{G}/\mathscr{H}$ 代表以等价类为元素的集合（$A,B\in\mathscr{G}$ 叫等价的，若 $A-B\in\mathscr{H}$），则 $\mathscr{G}/\mathscr{H}$ 是李代数，称为商李代数。

[15] 这巧妙地利用了 V_e 线性同构于 $\mathscr{L}$ 这一事实。

[16] 任一闭曲线可通过连续变形缩为一点的连通流形称为单连通流形。

定义 7.19. 李代数 $\mathscr{G}$ 称为单李代数，若它不是阿贝尔代数而且除 $\mathscr{G}$ 及 $\{0\}$ 外不含理想。$\mathscr{G}$ 称为半单李代数，若它不含非零的阿贝尔理想。相应地，李群 G 称为单李群，若它不是阿贝尔群而且除 G 外不含正规子群。G 称为半单李群，若它不含阿贝尔正规子群。

7.4 单参子群和指数映射

定义 7.20. C^∞ 曲线 $\gamma\colon \mathbb{R} \to G$ 叫李群 G 的单参子群，若

$$\gamma(s+t) = \gamma(s)\gamma(t),\ \forall s, t \in \mathbb{R}$$

其中 $\gamma(s)\gamma(t)$ 代表群元 $\gamma(s)$ 和 $\gamma(t)$ 的群乘积。

注 7.9. 上式表明单参子群 $\gamma\colon \mathbb{R} \to G$ 是从李群 $\mathbb{R}$ 到 G 的李群同态映射。

注 7.10. 按上述定义，单参子群是满足上式的一条 C^∞ 曲线，但不妨把映射的像集 $\{\gamma(t) \mid t \in \mathbb{R}\} \subset G$ 看作单参子群。上式保证：① 子集的任意两个元素按 G 的乘法的积仍在子集内；② 子集含 G 的恒等元 e（由上式得 $\gamma(t) = \gamma(0)\gamma(t)$，以 $\gamma(t)$ 的逆元右乘得 $e = \gamma(0)$）③ 子集的任一元素 $\gamma(t)$（作为 G 的元素）的逆元（就是 $\gamma(-t)$）在子集内。所以子集 $\{\gamma(t) \mid t \in \mathbb{R}\}$ 构成子群[17]（而且是阿贝尔子群）。

本节的一个重点是证明如下结论："单参子群是左不变矢量场过 e 的不可延积分曲线，反之亦然。" 单参子群按定义是从 $\mathbb{R}$ 到 G 的映射，定义域为全 $\mathbb{R}$。因此，要接受上述结论，至少要证明左不变矢量场过 e 的积分曲线的参数可取遍全 $\mathbb{R}$。下面的定理对此给出保证（而且更强）。

定理 7.8. 任一左不变矢量场 $\bar{A}$ 都是完备矢量场，就是说，它的每一不可延积分曲线的参数都可取遍全 $\mathbb{R}$。

证明. 本书惯用 $\partial/\partial t$ 代表曲线 $\gamma(t)$ 的切矢，由于本章涉及多条曲线的切矢，为避免混淆，我们改用 $\frac{d}{dt}\gamma(t)$ 代表 $\gamma(t)$ 的切矢。设 $\mu(t)$ 是 $\bar{A}$ 的、满足 $\mu(0) = e$ 的积分曲线，不失一般性，设其定义域为开区间 $(-\varepsilon, \varepsilon)$，即 $\mu\colon (-\varepsilon, \varepsilon) \to G$。在线上取点 $h \equiv \mu(\varepsilon/2)$，令 $\nu(t) \equiv h\mu(t - \varepsilon/2)$，则有曲线映射 $\nu\colon (-\varepsilon/2, 3\varepsilon/2) \to G$，且其切矢

$$\left.\frac{d}{dt}\right|_t \nu(t) = \left.\frac{d}{dt}\right|_t L_h\mu(t-\varepsilon/2) = L_{h*}\left.\frac{d}{dt}\right|_t \mu(t-\varepsilon/2)$$

其中第二步用到"曲线像的切矢等于曲线切矢的像"。令 $t'(t) \equiv t - \varepsilon/2$，则

$$\left.\frac{d}{dt}\right|_t \mu(t-\varepsilon/2) = \left.\frac{d}{dt}\right|_t \mu(t(t)) = \left[\left.\frac{d}{dt'}\right|_{t(t')} \mu(t')\right]\frac{dt'}{dt} = \left.\frac{d}{dt'}\right|_{t'(t)} \mu(t') = \bar{A}_{\mu(t-\varepsilon/2)}$$

[17]在 $\gamma\colon \mathbb{R} \to G$ 是多对一映射的情况下（例如后面的独点线及第 5 节的 $SO(2)$ 群），$\exists t_1, t_2 \in \mathbb{R}$ 使 $\gamma(t_1) = \gamma(t_2) \in \{\gamma(t) \mid t \in \mathbb{R}\}$。子群首先是子集，所以 $\gamma(t_1)$ 和 $\gamma(t_2)$ 理应看作子群 $\{\gamma(t) \mid t \in \mathbb{R}\}$ 的同一群元。但若称 $\{\gamma(t) \mid t \in \mathbb{R}\}$ 为单参子群，则不同参数值 t_1 和 t_2 又应给出不同群元，就是说，一旦在"子群"前冠以"单参"就应有 $\gamma(t_1) \neq \gamma(t_2)$。可见"单参子群"一词有不妥之处，至少不应把子群 $\{\gamma(t) \mid t \in \mathbb{R}\}$ 称为单参子群。

最后一步是因为 $\mu(t)$ 是 $\bar{A}$ 的积分曲线。代入上式便得

$$\left.\frac{d}{dt}\right|_t \nu(t) = L_{h*}\bar{A}_{\mu(t-\varepsilon/2)} = \bar{A}_{h\mu(t-\varepsilon/2)} = \bar{A}_{\nu(t)}$$

这就表明，ν 也是 $\bar{A}$ 的积分曲线，而 $\nu(\varepsilon/2) = h\mu(0) = h = \mu(\varepsilon/2)$ 则说明 ν 与 μ 有交，因此由积分曲线的（局域）唯一性可知在两积分曲线 ν 和 μ 的定义域的交集 $(-\varepsilon/2, \epsilon)$ 上有 $\nu = \mu$。可见 μ 的定义域已被延拓至 $(-\varepsilon, 3\varepsilon/2)$。重复以上操作便得 $\bar{A}$ 的一条定义域为全 $\mathbb{R}$ 的积分曲线 $\gamma: \mathbb{R} \to G$。作为证明的第二步，$\forall g \in G$，令

$$\beta(t) \equiv g\gamma(t)$$

则仿照上述推导有

$$\left.\frac{d}{dt}\right|_t \beta(t) = \bar{A}_{g\gamma(t)} = \bar{A}_{\beta(t)}$$

表明 $\beta: \mathbb{R} \to G$ 是 $\bar{A}$ 过 g 的积分曲线（满足 $d\beta/dt|_{t=0} = \bar{A}_g$）。可见 $\bar{A}$ 的任一不可延积分曲线的参数都可取遍全 $\mathbb{R}$。 □

定理 7.9. *设 $\gamma: \mathbb{R} \to G$ 是左不变矢量场 $\bar{A}$ 的、满足 $\gamma(0) = e$ 的积分曲线，则 γ 是 G 的一个单参子群。*

证明. 只须证明 γ 满足单参子群的定义。前面已经证明 $\beta(t) \equiv g\gamma(t)$ 是 $\bar{A}$ 的、满足 $\beta(0) = g$ 的积分曲线。取 γ 线上一点 $\gamma(s)$ 作为 g，便知

$$\beta(t) \equiv \gamma(s)\gamma(t)$$

是满足 $\beta(0) = \gamma(s)$ 的积分曲线。另一方面，由 $\gamma_1(t) \equiv \gamma(s+t)$ 定义的曲线 $\gamma_1: \mathbb{R} \to G$ 在 $\gamma_1(t)$ 点的切矢为

$$\left.\frac{d}{dt}\right|_t \gamma_1(t) = \left.\frac{d}{dt}\right|_t \gamma(s+t) = \left[\left.\frac{d}{d(s+t)}\right|_{s+t} \gamma(s+t)\right] \frac{d(s+t)}{dt} = \bar{A}_{\gamma(s+t)} = \bar{A}_{\gamma_1(t)}$$

上式表明 γ_1 也是 $\bar{A}$ 的积分曲线，而且也满足 $\gamma_1(0) = \gamma(s)$。由积分曲线的唯一性可知 $\gamma_1(t) = \beta(t)$，于是有 $\gamma(s+t) = \gamma(s)\gamma(t)$。 □

定理 7.10. *设单参子群 $\gamma: \mathbb{R} \to G$ 在恒等元 e 的切矢为 A，则 $\gamma(t)$ 是 A 对应的左不变矢量场 $\bar{A}$ 的积分曲线。*

证明. 只须证明对 γ 的任一点 $\gamma(s)$ 有 $\bar{A}_{\gamma(s)} = d/dt|_{t=s}\gamma(t)$，而这由下式显见：

$$\begin{aligned}\bar{A}_{\gamma(s)} &= L_{\gamma(s)*}A = L_{\gamma(s)*}\left.\frac{d}{dt}\right|_{t=0}\gamma(t) = \left.\frac{d}{dt}\right|_{t=0} L_{\gamma(s)}\gamma(t) \\ &= \left.\frac{d}{dt}\right|_{t=0}\gamma(s)\gamma(t) = \left.\frac{d}{dt}\right|_{t=0}\gamma(s+t) = \left.\frac{d}{dt'}\right|_{t'=s}\gamma(t') = \left.\frac{d}{dt}\right|_{t=s}\gamma(t)\end{aligned}$$

□

由上述两定理可知左不变矢量场与单参子群之间有一一对应关系。注意到 V_e 与左不变矢量场的集合 $\mathscr{L}$ 一一对应，便知李群 G 的李代数 $\mathscr{G}$（$=V_e$）的每一元素 A 生成 G 的一个单参子群 $\gamma(t)$，所以 $\mathscr{G}$ 的每一元素称为一个（无限小）生成元。物理文献常又只把 $\mathscr{G}$ 的一个基底的元素称为生成元。

注 7.11. 与 V_e 的零元（$A=0$）对应的单参子群[18]就是只含 e 的子群，即满足 $\gamma[\mathbb{R}]=e$ 的独点线 $\gamma\colon\mathbb{R}\to G$。

定义 7.21. 李群 G 上的指数映射 $\exp\colon V_e\to G$ 定义为

$$\exp(A)\coloneqq\gamma(1),\ \forall A\in\mathscr{G}$$

其中 $\gamma\colon\mathbb{R}\to G$ 是与 A 对应的那个单参子群。

定理 7.11.

$$\exp(sA)=\gamma(s),\ \forall s\in\mathbb{R},A\in V_e$$

其中 $\gamma(s)$ 是由 A 决定的单参子群。

证明. 令 $A'\equiv sA\in V_e$。以 $\bar{A},\bar{A}'$ 分别代表 A 和 A' 对应的左不变矢量场。$\eta\colon V_e\to\mathscr{L}$ 是同构映射导致 $\bar{A}'=s\bar{A}$。以 $\gamma(t)$ 代表 $\bar{A}$ 对应的单参子群，用 $\gamma'(t)\equiv\gamma(st)$ 定义曲线 $\gamma'\colon\mathbb{R}\to G$，则

$$\left.\frac{d}{dt}\right|_t\gamma'(t)=\left.\frac{d}{dt}\right|_t\gamma(st)=\left[\left.\frac{d}{d(st)}\right|_{st}\gamma(st)\right]\frac{d(st)}{dt}=s\bar{A}_{\gamma(st)}=\bar{A}'_{\gamma'(t)}$$

说明 $\gamma'(t)$ 是 $\bar{A}'$ 的积分曲线。注意到 $\gamma'(0)=\gamma(0)=e$，可知 $\gamma'(t)$ 是 $\bar{A}'$ 对应的单参子群。于是 $\exp(sA)=\exp(A')=\gamma'(1)=\gamma(s)$。 □

注 7.12. 设 $\gamma(t)$ 是由 $A\in V_e$ 决定的单参子群，则上述定理表明

$$\gamma(t)=\exp(tA)$$

今后也常用 $\exp(tA)$ 代表由 A 决定的单参子群。

定理 7.12. 设 $\phi\colon\mathbb{R}\times G\to G$ 是由 $A\in V_e$ 对应的左不变矢量场 $\bar{A}$ 产生的单参微分同胚群，则

$$\phi_t(g)=g\exp(tA),\ \forall g\in G,t\in\mathbb{R}$$

证明. 设 $\gamma(t)$ 是由 A 决定的单参子群，前述定理表明 $\gamma(t)=\exp(tA)$。由本节一开始的定理之证明又知 $\beta(t)\equiv g\gamma(t)$ 是 $\bar{A}$ 过 g 点的积分曲线，且 $\beta(0)=g$。由单参微分同胚群的构造[19]便知

$$\phi_t(g)=\beta(t)=g\gamma(t)=g\exp(tA)$$

□

[18] 其对应的左不变矢量场为零矢量场。

[19] 定义 $\phi_t(p)$ 为这样一个点，它位于过 p 的积分曲线上，其参数值与 p 的参数值之差为 t。
换言之，左不变矢量场 $\bar{A}$ 对应的单参子群就是其给出的单参微分同胚群过 e 点的轨道。

7.5 常用李群及其李代数

7.5.1 $GL(m)$ 群（一般线性群）

设 V 是 m（$< \infty$）维实矢量空间，以 $GL(m)$ 代表由 V 到 V 的全体可逆线性映射的集合，以映射的复合为乘法，则不难证明 $GL(m)$ 是群，称为 m 阶（实）一般线性群。因为 V 到 V 的线性映射就是 V 上的 $(1,1)$ 型张量，所以 $GL(m)$ 的任一群元 $T \in \mathscr{T}_V(1,1)$。取定 V 的任一基底（及其对偶基底）后，T 就有 m^2 个分量，自然对应于一个 $m \times m$ 实矩阵。映射 T 的可逆性保证其矩阵（仍记作 T）是可逆矩阵，即 $\det T \neq 0$。不难看出全体 $m \times m$ 可逆矩阵在矩阵乘法下构成群（以单位矩阵 I 为恒等元），而且与 $GL(m)$ 同构，于是

$$GL(m) = \{m \times m\text{实矩阵}T \mid \det T \neq 0\}$$

因此也常把 $GL(m)$ 看作实矩阵群（但上式等号代表的同构关系依赖于 V 的基底的选取）。另一方面，集合 $\mathbb{R}^{m^2}$ 的每点因为由 m^2 个实数构成而可排成一个 $m \times m$ 的实矩阵，故 $GL(m)$ 可看作 $\mathbb{R}^{m^2}$ 的子集。对矩阵求行列式的操作可看作连续映射 $\det\colon \mathbb{R}^{m^2} \to \mathbb{R}$，满足

$$GL(m) = \det{}^{-1}(-\infty, 0) \cup \det{}^{-1}(0, \infty) \subset \mathbb{R}^{m^2}$$

$(-\infty, 0)$ 和 $(0, \infty)$ 显然是 $\mathbb{R}$ 的开子集,故映射 $\det$ 的连续性保证 $\det^{-1}(-\infty, 0)$ 和 $\det^{-1}(0, \infty)$ 是 $\mathbb{R}^{m^2}$ 的开子集，这又导致 ① $GL(m)$ 是 $\mathbb{R}^{m^2}$ 的开子集; ② $\det^{-1}(-\infty, 0)$ 和 $\det^{-1}(0, \infty)$ 是 $GL(m)$ 的开子集（用诱导拓扑衡量）。而 $\det^{-1}(-\infty, 0)$ 和 $\det^{-1}(0, \infty)$ 又因为互为补集而都是 $GL(m)$ 的闭子集，可见 $GL(m)$ 含有除自身和 $\varnothing$ 外的既开又闭的子集，因而是非连通流形。还可证明，$GL(m)$ 是有两个连通分支的非连通流形。最简单的例子是

$$GL(1) = (-\infty, 0) \cup (0, \infty)$$

当说到 $GL(m)$ 的群元是从 V 到 V 的可逆线性映射时，我们是在用纯几何语言。当说到 $GL(m)$ 的群元是 $m \times m$ 可逆矩阵时，我们是在用坐标语言—给 V 选定基底（及其对偶基底）后，从 V 到 V 的任一可逆线性映射 T 就有了 m^2 个分量，选作坐标，便得流形 $GL(m)$ 上的一个坐标系。以 $\mathfrak{gl}(m)$ 代表 $GL(m)$ 的李代数，设 $A \in \mathfrak{gl}(m)$，则它是恒等元 $I \in GL(m)$ 的矢量，在上述坐标系的基底中自然有 m^2 个分量，因而也对应于一个 $m \times m$ 的实矩阵。反之，任一 $m \times m$ 实矩阵的 m^2 个矩阵元配以 I 点的坐标基矢便给出 I 点的一个矢量。可见，虽然只有可逆实矩阵 $m \times m$ 才是 $GL(m)$ 的元素，任意 $m \times m$ 实矩阵都是 $\mathfrak{gl}(m)$ 的元素。事实上，全体 $m \times m$ 实矩阵构成的矢量空间与 I 点的切空间 V_I 有同构关系，于是有如下定理

定理 7.13. $\mathfrak{gl}(m) = \{m \times m\text{实矩阵}\}$（等号代表两个矢量空间同构，同构关系依赖于 V 的基底的选取）

对任一 $m \times m$ 矩阵 A 引入符号

$$\mathrm{Exp}(A) \equiv I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

其中 $A^2 \equiv AA$（矩阵相乘），$A^3 \equiv AAA$。可以证明：① 上式右边收敛，所以左边有意义，是一个 $m \times m$ 矩阵；② 若矩阵 A, B 对易，即 $AB = BA$，则

$$\mathrm{Exp}(A + B) = (\mathrm{Exp}(A))(\mathrm{Exp}(B))$$

下面的定理表明，对 $\mathfrak{gl}(m)$ 的元素 A（看作矩阵），Exp 就是指数映射的 exp。

定理 7.14. $\mathrm{Exp}(A) = \exp(A),\ \forall A \in \mathfrak{gl}(m)$

证明. $\forall s, t \in \mathbb{R}$，由上式得

$$\mathrm{Exp}[(s + t)A] = \mathrm{Exp}(sA)\,\mathrm{Exp}(tA),\ \forall A \in \mathfrak{gl}(m)$$

取 $s = 1, t = -1$，则上式给出 $\mathrm{Exp}(A)\,\mathrm{Exp}(-A) = \mathrm{Exp}(0) = I$。上式表明矩阵 $\mathrm{Exp}(A)$ 有逆（就是 $\mathrm{Exp}(-A)$），故 $\mathrm{Exp}(A) \in GL(m)$。式中的 A 可为 $\mathfrak{gl}(m)$ 的任一元素，所以

$$\mathrm{Exp}(tA) \in GL(m),\ \forall A \in \mathfrak{gl}(m), t \in \mathbb{R}$$

令 $\gamma(t) \equiv \mathrm{Exp}(tA)$，则有 $\gamma(s + t) = \gamma(s)\gamma(t)$，且根据 Exp 的定义有

$$\left.\frac{d}{dt}\right|_{t=0} \mathrm{Exp}(tA) = A$$

可见 $\gamma(t) \equiv \mathrm{Exp}(tA)$ 是由 A 唯一决定的单参子群 $\exp(tA)$，从而 $\mathrm{Exp}(tA) = \exp(tA)$ □

本节的第一个定理表明 $\mathfrak{gl}(m)$ 和 $\{m \times m\text{实矩阵}\}$ 作为矢量空间是同构的。自然要问他们作为李代数是否也同构。答案是肯定的。以 V_e 作为 $\mathfrak{gl}(m)$，其任意二元素 $A, B \in V_e$ 的李括号定义为 $[A, B] = [\bar{A}, \bar{B}]_e$（见第三节）。另一方面，$A, B \in V_e$ 所对应的 $m \times m$ 矩阵（仍记作 A, B）的李括号定义为矩阵对易子 $AB - BA$。因此，欲证 $\mathfrak{gl}(m)$ 和 $\{m \times m\text{实矩阵}\}$ 是同构李代数只须证明 $[\bar{A}, \bar{B}]_e = AB - BA$。而这正是如下定理的结论。

定理 7.15. 设 G 是 $GL(m)$ 的李子群，则其李代数元 $A, B \in \mathfrak{g} \subset \mathfrak{gl}(m)$ 的李括号 $[A, B]$ 对应于 A, B 所对应的矩阵（仍记作 A, B）的对易子 $AB - BA$，即

$$[A, B] = AB - BA$$

7.5.2 O(m) 群（正交群）

如前所说，$GL(m)$ 群是 m 维矢量空间 V 上所有可逆 $(1, 1)$ 型张量 T 的集合，除可逆性外没有其他要求，所以称为一般线性群。对 T 再提出某些适当要求则可得到 $GL(m)$ 群的

某些子群。$O(m)$ 群就是重要一例，其要求是 T 保度规。以下把 $O(m)$ 群的元素专记作 Z。设 (V, g_{ab}) 是带正定度规的 m 维矢量空间。线性映射 $Z: V \to V$ 称为保度规的，若

$$g_{ab}(Z^a{}_c v^c)(Z^b{}_d u^d) = g_{cd} v^c u^d, \ \forall v^c, u^d \in V$$

取 $u = v$，则上式给出

$$g_{ab}(Z^a{}_c v^c)(Z^b{}_d v^d) = g_{cd} v^c v^d, \ \forall v^c \in V$$

可见保度规的 Z 对矢量的作用必保长度。反之，利用 $g_{ab} = g_{(ab)}$ 可证任何满足上式的 Z 必保度规。所以保长性等价于保度性。保度规性又等价于

$$Z^a{}_c Z^b{}_d g_{ab} = g_{cd}$$

令

$$O(m) \equiv \{Z^a{}_b \in \mathscr{T}_V(1,1) \mid Z^a{}_c Z^b{}_d g_{ab} = g_{cd}\}$$

把 $Z^a{}_b$ 看作从 V 到 V 的映射，以复合映射为群乘法，则 $O(m)$ 是群，而且是 $GL(m)$ 的李子群。用 V 的任一正交归一基底可把上式改写为分量形式[20]：

$$\delta_{\sigma\rho} = Z^\mu{}_\sigma Z^\nu{}_\rho \delta_{\mu\nu} = (Z^T)_\sigma{}^\mu \delta_{\mu\nu} Z^\nu{}_\rho$$

其中 Z^T 代表矩阵 Z 的转置矩阵。把上式写成矩阵等式即为

$$I = Z^T I Z = Z^T Z$$

表明 $Z^T = Z^{-1}$，即 Z 是正交矩阵。可见 $O(m)$ 同构于 $m \times m$ 正交矩阵群（由全体 $m \times m$ 正交矩阵以矩阵乘法为群乘法构成的群）。因为对任一矩阵 T 有 $\det T^T = \det T$，由上式得

$$1 = (\det Z^T)(\det Z) = (\det Z)^2$$

故

$$\det Z = \pm 1$$

上式表明群元的行列式不是 1 就是 -1，这就注定了 $O(m)$ 群的流形是非连通的。

以上抽象定义的 $O(m)$ 可用具体对象来体现。先考虑 $O(1)$ 群。1 维欧氏空间 $(\mathbb{R}, \delta_{ab})$ 可看作带正定度规的 1 维矢量空间，因而可以充当用以定义 $O(1)$ 群的那个 (V, g_{ab})。$O(1)$ 群的每个群元 Z 就是把空间的任一点（即起自原点的任一矢量）$\vec{v}$ 变为长度相等的矢量 $\vec{v}' \equiv Z(\vec{v})$ 的线性映射。由于长度相等，只有 $\vec{v}' = \vec{v}$ 和 $\vec{v}' = -\vec{v}$ 两种可能。可见 $O(1)$ 只有两个群元，其中一个是恒等元，另一个称为反射，记作 r，于是 $O(1) = \{e, r\}$。再讨论 $O(2)$ 群。2 维欧氏空间 $(\mathbb{R}^2, \delta_{ab})$ 可看作带正定度规的 2 维矢量空间。$O(2)$ 的每个群元 Z 就是

[20] $g_{\sigma\rho}(e^\sigma)_c(e^\rho)_d = Z^\mu{}_\sigma Z^\nu{}_\rho g_{\mu\nu}(e^\sigma)_c(e^\rho)_d$ 因 $(e^\sigma)_c(e^\rho)_d$ 正交归一并注意到我们约定 g_{ab} 为正定度规，就有 $g_{\sigma\rho}(e^\sigma)_c(e^\rho)_d = \delta_{\sigma\rho}$

把空间中起自原点的任一矢量 $\vec{v}$ 变为长度相等的矢量 $\vec{v}' \equiv Z(\vec{v})$ 的线性映射。首先想到的自然是令 $\vec{v}$ 转某一角度 α 的转动，这种映射记作 $Z(\alpha)$，可用矩阵表示为

$$Z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

这是行列式为 $+1$ 的正交矩阵。上式的特例是 $Z(0) = \mathrm{diag}(1,1)$，即恒等元 I。然而，不难验证下面的 $Z'(\alpha)$ 也保长度：

$$Z'(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$$

这是行列式为 -1 的正交矩阵，其特例是 $Z'(0) = \mathrm{diag}(1,-1)$，它作用于 $\vec{v}$ 的结果 $\vec{v}'$ 与 $\vec{v}$ 关于 x 轴对称，故 $Z'(0)$ 代表以 x 轴为对称轴的反射，记作 r_y（只改变 y 分量），即 $\vec{v}' = r_y(\vec{v})$。不要误以为 r_y 也可看作转动，因为它作用于另一矢量 $\vec{u}$ 得 $r_y(\vec{u})$，其与 $\vec{u}$ 的夹角不同于 $r_y(\vec{v})$ 与 $\vec{v}$ 的夹角。同理，$Z'(\pi) = \mathrm{diag}(-1,1)$ 代表以 y 轴为对称轴的反射 r_x。此外，$Z(\pi) = \mathrm{diag}(-1,-1)$ 则代表以原点为对称点的反演，记作 i_{xy}，易证 $i_{xy} = r_x r_y$。

Z 的 4 个矩阵元由于条件 $Z^T Z = I$ 而受到 3 个方程的限制[21]，只有 1 个独立，所以 $O(2)$ 是 1 维李群。由 $Z^T Z = I$ 还可证明 $O(2)$ 的群元被前述两式所穷尽，其中第一式代表的子集构成 $O(2)$ 的李子群，称为 2 维空间的转动群，记作 $SO(2)$，字母 S 代表 “特殊”，是指每个群元 $Z \in SO(2)$ 满足 $\det Z = +1$。这也适用于其他常见李群，例如 $GL(m)$ 中满足 $\det T = +1$ 的子集也构成子群，称为 $SL(m)$ 群（特殊线性群）。$O(2)$ 群中由前述第二式表示的子集不含恒等元 e，因此不是子群。两个子集的 α 都可解释为转角（$Z'(\alpha)$ 中的 α 代表（对 x 轴）反射后再转的角度），取值范围是 $0 \leqslant \alpha < 2\pi$。所以 $O(2)$ 群的流形可看作两个互不连通的 S^1（圆周）之并，是一个非连通流形，每个圆周是一个连通分支。

推而广之，因为 $O(m)$ 的群元 Z 满足 $\det Z = \pm 1$，其流形总是由两个连通分支组成的非连通流形，其中含 e 的分支记作 $SO(m)$，即

$$SO(m) = \{Z \in O(m) \mid \det Z = +1\}$$

这是 $O(m)$ 的李子群，称为特殊正交群，最常用的是 3 维空间的转动群 $SO(3)$，它的每一群元可看作把 3 维欧氏空间中起自原点的任一矢量 $\vec{v}$ 绕过原点的某轴转某角的映射（而且除恒等元外的每个群元的转轴是唯一的，这一结论称为欧拉定理）。这使我们可用起自原点的一个箭头代表 $SO(3)$ 的一个群元：箭头所在直线代表转轴，箭头长度（规定从 0 到 π）代表转角，方向相反的箭头代表沿相反方向的转动。由于沿某方向转 π 角相当于沿反方向转 π 角，同一直线段的两端代表同一群元，应当认同。于是，李群 $SO(3)$ 的流形是 3 维实心球体，球面上位于每一直径两端的一对点要认同（简称对径认同）。这一流形记作 $\mathbb{RP}^3$。（3 维实射影空间 $\mathbb{RP}^3$ 本有其他定义方式，但可以证明 $SO(3)$ 与 $\mathbb{RP}^3$ 微分同胚。）

[21] $z_{11}^2 + z_{21}^2 = 1, z_{12}^2 + z_{22}^2 = 1, z_{11}z_{12} + z_{21}z_{22} = 0$

下面讨论李群 $O(m)$ 和 $SO(m)$ 的李代数 $\mathfrak{o}(m)$ 和 $\mathfrak{so}(m)$。这两个李代数必定一样，因为设 $A \in \mathfrak{o}(m)$，则 A 决定 $O(m)$ 中的一个单参子群 $\gamma(t)$，注意到 ① $\det I = 1$; ② $O(m)$ 中任一群元的行列式只能为 ± 1，则由单参子群的连续性可知 $\gamma(t)$ 上任一点 Z 都有 $\det Z = 1$，可见对任一 t 有 $\gamma(t) \in SO(m)$，从而 A，作为 $\gamma(t)$ 在 I 点的切矢，必属于 $\mathfrak{so}(m)$

定理 7.16. $\mathfrak{o}(m) = \{m \times m\text{实矩阵}A \mid A^T = -A\text{（即 }A\text{ 为反称阵）}\}$

证明.（A） $\forall A \in \mathfrak{o}(m)$，设 $Z(t)$ 是李群 $O(m)$ 中的曲线，且 $Z(0) = I, d/dt|_{t=0} Z(t) = A$，则 $Z(t)$（对每一 t 值）是一个满足

$$Z^T(t)Z(t) = I$$

的矩阵[22]。对上式求导并在 $t = 0$ 取值得

$$0 = \left[Z^T(t)\frac{d}{dt}Z(t)\right]_{t=0} + \left[\left(\frac{d}{dt}Z^T(t)\right)Z(t)\right]_{t=0} = Z^T(0)A + \left[\frac{d}{dt}Z^T(t)\right]_{t=0} Z(0)$$

$Z(0) = I$ 等价于 $Z^T(0) = I$，$A = d/dt|_{t=0}Z(t)$ 等价于[23]$A^T = d/dt|_{t=0}Z^T(t)$，因而上式给出 $0 = A + A^T$，可见 A 为反称矩阵。

（B） 设 A 为任一 $m \times m$ 反称实矩阵，注意到 $(A^2)^T = (A^T)^2, (A^3)^T = (A^T)^3, \cdots$，因此有 $(\mathrm{Exp}\, A)^T = \mathrm{Exp}(A^T)$[24]所以

$$(\mathrm{Exp}\, A)^T(\mathrm{Exp}\, A) = (\mathrm{Exp}\, A^T)(\mathrm{Exp}\, A) = \mathrm{Exp}(-A)(\mathrm{Exp}\, A) = I$$

上式表明 $\mathrm{Exp}\, A$ 是正交矩阵，因而 $\mathrm{Exp}\, A \in O(m)$。由此又知[25]$\mathrm{Exp}(tA) \in O(m)$（对每一 t 值）。又根据 Exp 的定义知 $d/dt|_{t=0}\,\mathrm{Exp}(tA) = A$，可见 A 是李群 $O(m)$ 中过 I 的曲线 $\mathrm{Exp}(tA)$ 在 I 点的切矢，所以 $A \in \mathfrak{o}(m)$

□

利用上述定理可方便地决定李群 $O(m)$ 及 $SO(m)$ 的维数。因为 $m \times m$ 反称矩阵的 m 个对角元全为零，其余 $m^2 - m$ 个元素中只有半数独立，所以决定一个 $m \times m$ 的反称实矩阵需要 $(m^2 - m)/2 = m(m-1)/2$ 个实数。于是上述定理导致

$$\dim O(m) = \dim \mathfrak{o}(m) = \frac{1}{2}m(m-1)$$

具体说，

$$\dim O(1) = 0\text{（零维李群）}, \dim O(2) = 1, \dim O(3) = 3, \dim O(4) = 6, \cdots,$$

[22]原因为 $Z(t) \in SO(m)$，参考前述分析

[23]$\lim\limits_{t\to 0}\dfrac{Z^T(t) - Z^T(0)}{t} = \lim\limits_{t\to 0}\dfrac{Z^T(t) - I^T}{t^T} = \lim\limits_{t\to 0}\left(\dfrac{Z(t) - I}{t}\right)^T = A^T$

[24]直观想来这很显然，但因涉及无限级数，求转置与求极限操作的可交换问题并不简单。以下遇到类似问题时不再指出。

[25]把前述论证中的 A 替换为 $A' \equiv tA$ 即可。

7.5.3 O(1,3) 群（洛伦兹群）

定义 $O(m)$ 群时曾约定 (V, g_{ab}) 中的 g_{ab} 为正定度规，现在放宽为任意度规。设 g_{ab} 在正交归一基底下的矩阵有 m' 个对角元为 -1，m'' 个对角元为 $+1$，则 V 上所有保度规的线性映射的集合在以复合映射为群乘法下构成群，记作 $O(m', m'')$，即

$$O(m', m'') \coloneqq \{\Lambda^a{}_b \in \mathscr{T}_V(1,1) \mid \Lambda^a{}_c \Lambda^b{}_d g_{ab} = g_{cd}\}$$

同 $O(m)$ 类似，$O(m', m'')$ 也是 $GL(m)$（其中 $m = m' + m''$）的李子群，$O(m)$ 可看作 $O(m', m'')$ 在 $m' = 0, m'' = m$ 时的特例。

抽象定义的 $O(m', m'')$ 群也可与 $O(m)$ 类似地用具体对象体现。我们只讨论 $m' = 1$ 的情况，并且最关心 $O(1,3)$ 群（4 维闵氏时空的洛伦兹群），但先从最简单的 $O(1,1)$ 群讲起。2 维闵氏时空 $(\mathbb{R}^2, \eta_{ab})$ 可看作用来定义 $O(1,1)$ 群的那个带洛伦兹度规的 (V, g_{ab})。用 $(\mathbb{R}^2, \eta_{ab})$ 的任一正交归一基可把保度规条件 $\Lambda^a{}_c \Lambda^b{}_d g_{ab} = g_{cd}$ 改写为分量形式：

$$\eta_{\sigma\rho} = \Lambda^\mu{}_\sigma \Lambda^\nu{}_\rho \eta_{\mu\nu} = (\Lambda^T)_\sigma{}^\mu \eta_{\mu\nu} \Lambda^\nu{}_\rho$$

写成矩阵等式即为

$$\eta = \Lambda^T \eta \Lambda, \text{其中} \eta \equiv \mathrm{diag}(-1, 1)$$

上式表明 $\det \Lambda = \pm 1$，因此，与 $O(m)$ 群类似，$O(1,1)$ 群也是非连通的。注意到洛伦兹变换 $t' = \gamma(t - vx), x' = \gamma(x - vt)$ 保闵氏线元，自然猜想它是 $O(1,1)$ 的群元。把变换的参数 v 改写为 $\mathrm{th}\,\lambda$，其中 $\lambda \in (-\infty, \infty)$，则 $\gamma \equiv (1 - v^2)^{-1/2} = \mathrm{ch}\,\lambda$，故洛伦兹变换可改写为

$$t' = t\,\mathrm{ch}\,\lambda - x\,\mathrm{sh}\,\lambda,\ x' = -t\,\mathrm{sh}\,\lambda + x\,\mathrm{ch}\,\lambda$$

不难验证这一变换的矩阵

$$\Lambda(\lambda) = \begin{bmatrix} \mathrm{ch}\,\lambda & -\mathrm{sh}\,\lambda \\ -\mathrm{sh}\,\lambda & \mathrm{ch}\,\lambda \end{bmatrix}$$

满足上式，可见 $\forall \lambda \in (-\infty, \infty)$，$\Lambda(\lambda)$ 是 $O(1,1)$ 的群元。对上式的 Λ 有 $\det \Lambda = +1$，然而前面已说明 $\det \Lambda = \pm 1$，看来还应有其他连通分支。其实，$O(1,1)$ 的非连通性除了来自 Λ 所受的限制 $\det \Lambda = \pm 1$ 之外还来自 $\Lambda^0{}_0$ 所受到的限制。设 $\{(e_\mu)^a\}$ 是 V 的正交归一基底，用 $\Lambda^a{}_c \Lambda^b{}_d g_{ab} = g_{cd}$ 作用于 $(e_0)^c (e_0)^d$ 得[26]$-(\Lambda^0{}_0)^2 + (\Lambda^1{}_0)^2 = -1$，所以 $(\Lambda^0{}_0)^2 \geqslant 1$。$\det \Lambda$ 与 $\Lambda^0{}_0$ 的允许值的配合使李群 $O(1,1)$ 由如下 4 个互不连通的连通分支组成：

(I) 子集 $O_+^\uparrow(1,1)$，其元素 $\Lambda(\lambda) = \begin{bmatrix} \mathrm{ch}\,\lambda & -\mathrm{sh}\,\lambda \\ -\mathrm{sh}\,\lambda & \mathrm{ch}\,\lambda \end{bmatrix}$ 满足 $\det \Lambda = +1, \Lambda^0{}_0 \geqslant +1$；

(II) 子集 $O_-^\uparrow(1,1)$，其元素 $\Lambda(\lambda) = \begin{bmatrix} \mathrm{ch}\,\lambda & -\mathrm{sh}\,\lambda \\ \mathrm{sh}\,\lambda & -\mathrm{ch}\,\lambda \end{bmatrix}$ 满足 $\det \Lambda = -1, \Lambda^0{}_0 \geqslant +1$；

[26]左边 $= \Lambda^a{}_0 \Lambda^b{}_0 g_{ab} = -(\Lambda^0{}_0)^2 + (\Lambda^1{}_0)^2$，右边 $= -1$。

(III) 子集 $O^{\downarrow}_{-}(1,1)$，其元素 $\Lambda(\lambda)=\begin{bmatrix}-\operatorname{ch}\lambda & \operatorname{sh}\lambda\\ -\operatorname{sh}\lambda & \operatorname{ch}\lambda\end{bmatrix}$ 满足 $\det\Lambda=-1, \Lambda^0{}_0\leqslant -1$；

(IV) 子集 $O^{\downarrow}_{+}(1,1)$，其元素 $\Lambda(\lambda)=\begin{bmatrix}-\operatorname{ch}\lambda & \operatorname{sh}\lambda\\ \operatorname{sh}\lambda & -\operatorname{ch}\lambda\end{bmatrix}$ 满足 $\det\Lambda=+1, \Lambda^0{}_0\leqslant -1$；

$(\mathbb{R}^2, \eta_{ab})$ 可看作带洛伦兹度规的 2 维矢量空间，$\Lambda^a{}_b\in O(1,1)$ 就是把空间中起自原点的任一矢量 v^a 变为 $\Lambda^a{}_b v^b$ 的保度规线性映射。略去抽象指标，用正交归一基底把 v 表为列矢量 $v=\begin{bmatrix}v^0\\ v^1\end{bmatrix}$，我们来看 4 个连通分支中 $\lambda=0$ 的元素 $\Lambda(0)$ 对它的作用。对 $O^{\uparrow}_{+}(1,1)$ 而言，$\Lambda(0)$ 就是恒等元，它把 v 映射为 v。对 $O^{\uparrow}_{-}(1,1)$，有 $\Lambda(0)=\operatorname{diag}(1,-1)$，作用于 v 所得矩阵为 $v'=\begin{bmatrix}v^0\\ -v^1\end{bmatrix}$，称为空间反射，记作 r_x。对 $O^{\downarrow}_{-}(1,1)$，有 $\Lambda(0)=\operatorname{diag}(-1,1)$，作用于 v 所得矩阵为 $v'=\begin{bmatrix}-v^0\\ v^1\end{bmatrix}$，称为时间反射，记作 r_t。对 $O^{\downarrow}_{+}(1,1)$，有 $\Lambda(0)=\operatorname{diag}(-1,-1)$，作用于 v 所得矩阵为 $v'=\begin{bmatrix}-v^0\\ -v^1\end{bmatrix}$，称为时空反演，记作 i_{tx}。

因为每个连通分支的元素由一个参数 λ 决定，而且 $\lambda\in(-\infty,\infty)$，所以每一连通分支都可看作流形 $\mathbb{R}$，整个李群 $O(1,1)$ 是 1 维非连通流形。4 个连通分支中只有 $O^{\uparrow}_{+}(1,1)$ 是 $O(1,1)$ 的子群（只有它包含恒等元），因而是李子群。（因为有这样的一般结论：李群的含恒等元的连通分支是它的李子群。）$O(1,1)$ 与 $O(2)$ 有一重要区别：$O(2)$ 的流形紧致而 $O(1,1)$ 非紧致。我们称 $O(2)$ 为紧致李群而 $O(1,1)$ 为非紧致李群。

在以上基础上就不难介绍 $O(m',m'')$ 中对物理最有用的特例，即洛伦兹群 $O(1,3)$，简记为 L，其群元 $\Lambda\in L$（作为 4×4 矩阵）的充要条件与前述一样，只是改为 4×4 矩阵等式，即

$$\eta=\Lambda^T\eta\Lambda, \text{其中}\eta\equiv\operatorname{diag}(-1,1,1,1)$$

L 是由 4 个连通分支组成的 6 维非连通流形，这 4 个连通分支是

$$L^{\uparrow}_{+}=\{\Lambda\in L\mid\det\Lambda=+1,\Lambda^0{}_0\geqslant +1\},\ L^{\uparrow}_{-}=\{\Lambda\in L\mid\det\Lambda=-1,\Lambda^0{}_0\geqslant +1\},$$
$$L^{\downarrow}_{-}=\{\Lambda\in L\mid\det\Lambda=-1,\Lambda^0{}_0\leqslant -1\},\ L^{\downarrow}_{+}=\{\Lambda\in L\mid\det\Lambda=+1,\Lambda^0{}_0\leqslant -1\}.$$

每个连通分支含一个最简单的元素，依次记作 I, r_s, r_t, i_{ts}，其中

$$I\equiv\operatorname{diag}(1,1,1,1)\in L^{\uparrow}_{+}\text{是 } L \text{ 的恒等元}$$
$$r_s\equiv\operatorname{diag}(1,-1,-1,-1)\in L^{\uparrow}_{-}\text{是 } L \text{ 的空间反射元}$$
$$r_t\equiv\operatorname{diag}(-1,1,1,1)\in L^{\downarrow}_{-}\text{是 } L \text{ 的时间反射元}$$
$$i_{ts}\equiv r_t r_s=-I\in L^{\downarrow}_{+}\text{是 } L \text{ 的时空反演元}$$

上述 4 个连通分支中只有 $L_+^\uparrow$ 是子群（只有它包含恒等元），而且是李子群，称为固有洛伦兹群[27]，是 6 维连通流形，流形结构为 $\mathbb{R}^3 \times SO(3)$。其他 3 个连通分支都是子群 $L_+^\uparrow$ 的左陪集：

$$L_-^\uparrow = r_s L_+^\uparrow,\ L_-^\downarrow = r_t L_+^\uparrow,\ L_+^\downarrow = i_{ts} L_+^\uparrow$$

同 O(3) 群类似，洛伦兹群 O(1,3) 和它的子群 $L_+^\uparrow$ 有相同的李代数，记作 $\mathfrak{o}(1,3)$。

定理 7.17. $\mathfrak{o}(1,3) = \{4\times 4\text{实矩阵}A \mid A^T = -\eta A\eta\}$

证明.（A） $\forall A \in \mathfrak{o}(1,3)$，考虑 O(1,3) 中满足以下两条件的曲线 $\Lambda(t)$：

（1） $\Lambda(0) = I$，

（2） $d/dt|_{t=0}\Lambda(t) = A$

注意到 $\Lambda(t) \in O(1,3)$ 则有 $\Lambda^T(t)\eta\Lambda(t) = \eta\ \forall t \in \mathbb{R}$。对 t 求导并在 $t = 0$ 取值得

$$0 = \left[\frac{d}{dt}\Lambda^T(t)\right]_{t=0}\eta\Lambda(0) + \Lambda^T(0)\eta\left[\frac{d}{dt}\Lambda(t)\right]_{t=0} = A^T\eta I + I\eta A = A^T\eta + \eta A$$

以 η^{-1} 右乘上式，注意到 $\eta^{-1} = \eta$，得 $A^T = -\eta A\eta$。（η 与 η^{-1} 作为矩阵相等，但写矩阵元时最好分清上下标，即 η 的矩阵元为 $\eta_{\mu\nu}$ 而 η^{-1} 的矩阵元为 $\eta^{\mu\nu}$。）

（B） 设 A 是满足 $A^T = -\eta A\eta$ 的 4×4 实矩阵。仿照讨论 $O(m)$ 时的证明，只须证明 $\mathrm{Exp}(A) \in O(1,3)$。因为对任意矩阵 M, N 有

$$\begin{aligned}
&N^{-1}(\mathrm{Exp}\,M)N = N^{-1}(I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \cdots)N \\
=&I + N^{-1}MN + \frac{1}{2!}(N^{-1}MN)(N^{-1}MN) + \frac{1}{3!}(N^{-1}MN)(N^{-1}MN)(N^{-1}MN) + \cdots \\
=&\mathrm{Exp}(N^{-1}MN)
\end{aligned}$$

取 $N = \eta, M = A$ 便得 $\eta(\mathrm{Exp}\,A)\eta = \mathrm{Exp}(\eta A\eta)$，因而 $\eta(\mathrm{Exp}\,A) = \mathrm{Exp}(\eta A\eta)\eta$。于是 $(\mathrm{Exp}\,A)^T\eta(\mathrm{Exp}\,A) = (\mathrm{Exp}\,A^T)\eta(\mathrm{Exp}\,A) = (\mathrm{Exp}\,A^T)\,\mathrm{Exp}(\eta A\eta)\eta = \mathrm{Exp}(A^T + \eta A\eta)\eta = \eta$，（其中用到 $A^T = -\eta A\eta$。）可见 $\mathrm{Exp}(A) \in O(1,3)$。

□

要利用计算 $O(m)$ 维数的方法来计算 O(1,3) 群的维数，可以借用如下事实：$\forall A \in \mathfrak{o}(1,3)$，令 $B \equiv \eta A$，则由 $A^T = -\eta A\eta$ 易见 $B^T = -B$，即 $B \in \mathfrak{o}(4)$。这表明存在从 $\mathfrak{o}(1,3)$ 到 $\mathfrak{o}(4)$ 的映射，而且是（矢量空间之间的）同构映射，因而 $\dim\mathfrak{o}(1,3) = \dim\mathfrak{o}(4)$。这一

[27] ${\Lambda^0}_0 \geqslant +1$ 的洛伦兹变换不改变被作用矢量 v^a 的时间分量 v^0 的正负性，故称保时向的（orthochronous），反之 ${\Lambda^0}_0 \leqslant -1$ 的洛伦兹变换称为反时向的（antichronous）。另一方面，“固有”（proper）一词则常用以形容 $\det\Lambda = +1$。因此 $L_+^\uparrow$ 的全称应是固有、保时向洛伦兹群，而固有洛伦兹群一词则应留给 L 的非连通子群 $L_+ = \{\Lambda \in L \mid \det\Lambda = +1\}$。也有文献称 $L_+^\uparrow$ 为限制（restricted）洛伦兹群或正常洛伦兹群。考虑到简单性和习惯性，本书仍仿照多数文献称 $L_+^\uparrow$ 为固有洛伦兹群。

结论只依赖于 η 的对称性和非退化性。对 $O(m', m'')$ 群，令 η 代表这样的对角矩阵，其前 m' 个对角元为 -1，后 m'' 个对角元为 $+1$，便知也有类似结论，即

$$\dim O(m', m'') = \dim O(m' + m'')$$

$\Lambda \in L$ 满足 $\eta = \Lambda^T \eta \Lambda$，其中 $\eta \equiv \mathrm{diag}(-1, 1, 1, 1)$。而

$$\eta = \Lambda^T \eta \Lambda \Leftrightarrow \eta \Lambda^{-1} = \Lambda^T \eta \Leftrightarrow \Lambda^{-1} = \eta^{-1} \Lambda^T \eta$$

再用 $\eta^{-1} = \eta$ 便得 $\Lambda^{-1} = \eta \Lambda^T \eta$。借此可方便地求得 $\Lambda \in L$ 的逆矩阵 Λ^{-1}：先写出 Λ 的转置矩阵 Λ^T，再对 Λ^T 的各元素依下列法则改变符号：$\begin{bmatrix} + - - - \\ - + + + \\ - + + + \\ - + + + \end{bmatrix}$，结果即为 Λ^{-1}。

7.5.4 $U(m)$ 群（酉群）

$GL(m)$ 群是用 m 维实矢量空间 V 定义的。把 V 改为 m 维复矢量空间所得的一般线性群记作 $GL(m, \mathbb{C})$，是 $2m^2$ 维（实）李群。与 $GL(m, \mathbb{R})$（即 $GL(m)$）不同，$GL(m, \mathbb{C})$ 是连通李群。我们只介绍这个群的一个重要子群—酉群 $U(m)$。正如 $GL(m)$ 群的子群 $O(m)$ 要求保度规那样，酉群 $U(m)$ 是对 $GL(m, \mathbb{C})$ 要求保内积的结果。复矢量空间的内积与实矢量空间的内积类似而又不同，最重要的区别在于，实矢量空间的内积对两个矢量的作用都为线性，而复矢量空间的内积[28]只对第二个矢量为线性，对第一个矢量则为反线性。定义了内积的复矢量空间称为内积空间。设 V 是有限维内积空间，线性映射 $A\colon V \to V$ 称为 V 上的线性算符，简称算符，它自然诱导出 V 上的另一线性算符 $A^\dagger$，称为 A 的伴随算符，满足

$$(A^\dagger f, g) = (f, Ag), \ \forall f, g \in V$$

其中 Ag 代表 A 作用于 g 所得矢量，即 $A(g)$ 的简写。可以证明每一 A 按上式决定唯一的 $A^\dagger$，上式可作为 $A^\dagger$ 的定义。

定义 7.22. 内积空间 V 上的算符 U 称为酉算符（或幺正算符），若其作用保内积，即

$$(Uf, Ug) = (f, g), \ \forall f, g \in V$$

定理 7.18. 算符 U 为酉算符的充要条件为

$$U^\dagger U = \delta$$

其中 δ 代表恒等算符，即从 V 到 V 的恒等映射。

[28]如下条件可作为其定义：① $(f, g+h) = (f, g) + (f, h)$；② $(f, cg) = c(f, g)$；③ $(f, g) = \overline{(g, f)}$；④ $(f, f) \geqslant 0$，且$(f, f) = 0 \Leftrightarrow f = 0$；

证明. 若 $U^\dagger U=\delta$，则根据伴随算符的定义有 $(Uf,Ug)=(U^\dagger Uf,g)=(f,g)$, $\forall f,g\in V$，故 U 为酉算符。反之，若 U 为酉算符，则 $\forall f,g\in V$ 有

$$0=(Uf,Ug)-(f,g)=(U^\dagger Uf,g)-(f,g)=((U^\dagger U-\delta)f,g)$$

取 $g=(U^\dagger U-\delta)f$，则 $0=(g,g)$，故 $g=0$，即 $(U^\dagger U-\delta)f=0\ \forall f\in V$，从而 $U^\dagger U=\delta$。 □

选定 V 的一个正交归一基底 $\{e_i\}$ 后，V 上任一算符 A 可用矩阵表示，其矩阵元定义为[29]

$$A_{ij}:=(e_i,Ae_j)$$

由上式得

$$A_{ij}=(A^\dagger e_i,e_j)=\overline{(e_j,A^\dagger e_i)}=\overline{A^\dagger_{ji}}$$

所以，若以 A 和 $A^\dagger$ 分别代表算符 A 和 $A^\dagger$ 在同一基底的矩阵，以 $\overline{A}^T$ 代表 A 的转置矩阵的矩阵元都取复数共轭所得的矩阵，则

$$A^\dagger=\overline{A}^T$$

定理 7.19. 算符 U 为酉算符的充要条件是其在正交归一基底下的矩阵 U 满足矩阵等式

$$U^{-1}=U^\dagger,\text{即}U^{-1}=\overline{U}^T$$

证明. 前述酉算符满足的算符等式在正交归一基底下对应于矩阵等式 $U^\dagger U=I$ 故 $U^{-1}=U^\dagger$。 □

定义 7.23. 满足上式的复矩阵 U 称为酉矩阵（或幺正矩阵）。

注 7.13. 如果 V 为实矢量空间，则 $U^{-1}=U^\dagger$ 成为 $U^{-1}=U^T$，即 U 为正交矩阵。所以酉矩阵可看作正交矩阵在复矩阵的推广。

定理 7.20. 设 U 为酉矩阵，则

$$\det U=e^{i\phi}(\text{其中 }\phi\in\mathbb{R})\text{，即 }|\det U|=1$$

证明. 对前述矩阵等式取行列式得 $1=(\det\overline{U}^T)(\det U)=(\overline{\det U})(\det U)$，得证。 □

定义 7.24. 令 $U(m)\equiv\{m$ 维内积空间 V 上的酉算符$\}=\{m\times m$酉矩阵$\}$，以复合映射（或矩阵乘法）为群乘法，则不难验证 $U(m)$ 是李群，称为酉群（或幺正群）。

[29]可以证明，这样定义的矩阵满足如下两个条件：① A 是零算符当且仅当其矩阵 A 为零矩阵；② 算符 AB 的矩阵为算符 A 的矩阵与算符 B 的矩阵的乘积 AB。

例 7.9. （酉群 $U(1)$）选 1 维复矢量空间 $\mathbb{C}$ 作为 V 并按下式定义内积使成为内积空间：

$$(f, g) := \bar{f}g,\ \forall f, g \in \mathbb{C}\text{（容易验证满足内积定义各条件）}$$

复数 $U \in \mathbb{C}$ 可看作 $\mathbb{C}$ 上的线性算符，它作用于任一 $f \in \mathbb{C}$ 的结果定义为复数乘积 Uf。易证 U 为酉算符当且仅当 $|U| = 1$，所以 $\mathbb{C}$ 上全体酉算符的集合是复平面上的单位圆。可见李群 $U(1)$ 的流形是一个圆周，是个紧致的连通流形（连通性的证明见下文定理）。

定义 7.25. 复方阵 A 称为厄米的，若 $A^\dagger = A$；A 称为反厄米的，若 $A^\dagger = -A$。

注 7.14. ① 对实方阵，厄米就是对称；反厄米就是反称。② 易证：A 为厄米 $\Leftrightarrow$ iA 为反厄米。

与 $GL(m)$ 群的李代数 $\mathfrak{gl}(m)$ 同构于全体 $m \times m$ 实矩阵构成的李代数类似，$GL(m, \mathbb{C})$ 群的李代数 $\mathfrak{gl}(m, \mathbb{C})$ 同构于全体 $m \times m$ 复矩阵构成的李代数，即

$$\mathfrak{gl}(m, \mathbb{C}) = \{m \times m\text{复矩阵}\}\text{（等号代表李代数同构）}$$

定理 7.21. 酉群 $U(m)$ 的李代数

$$\mathfrak{u}(m) = \{A \in \mathfrak{gl}(m, \mathbb{C}) \mid A^\dagger = -A\} = \{m \times m\text{反厄米矩阵}\}$$

证明. 与正交群对应的定理证明类似，只须把 $Z^T(t)$ 和 A^T 分别改为 $U^\dagger(t)$ 和 $A^\dagger$。 □

定理 7.22.

$$\dim U(m) = \dim \mathfrak{u}(m) = m^2$$

证明. $A \in \mathfrak{u}(m)$ 是有 m^2 个复元素的矩阵，由 $2m^2$ 个实数决定。但 A 的反厄米性 $A_{ij} = -\bar{A}_{ji}$ 导致 m^2 个实方程，理由是：① 反厄米性要求对角元为虚数，故每个对角元的反厄米条件相当于一个实方程（实部 = 0），因而所有对角元提供 m 个实方程；② 反厄米性使每对非对角元提供 2 个实方程，而非对角元共有 $(m^2 - m)/2$ 对，故全部非对角元共提供 $2 \times (m^2 - m)/2 = m^2 - m$ 个方程。可见 $2m^2$ 个实数要受 m^2 个实方程限制，只有 m^2 个独立。 □

例 7.10. （李代数 $\mathfrak{u}(1)$）由第一个例子得知李群 $U(1)$ 可表为 $U(1) = \{e^{-i\theta} \mid \theta \in \mathbb{R}\}$，其中 $e^{-i\theta} = \mathrm{Exp}(-i\theta)$ 可看作以 θ 为参数的单参子群（等于 $U(1)$ 自身），相应的李代数元为

$$A = \left.\frac{d}{d\theta}\right|_{\theta=0} e^{-i\theta} = -i \in \mathfrak{u}(1)\text{（不难验证 }A\text{ 的确满足 }A^\dagger = -A\text{）}$$

以 A 为基矢生成的 1 维（实）矢量空间就是 $\mathfrak{u}(1)$，故 $\mathfrak{u}(1) = \{-i\alpha \mid \alpha \in \mathbb{R}\}$

酉矩阵满足的 $\det U = e^{i\phi}$ 和正交矩阵满足的 $\det Z = \pm 1$ 很不一样。这同如下事实密切相关：

定理 7.23. 与 $O(m)$ 群相反，酉群 $U(m)$ 是连通流形。

证明. 只须证明 $\forall U \in U(m)$，存在连续曲线 $\gamma(t) \subset U(m)$ 使 $\gamma(0) = I, \gamma(1) = U$。由线性代数可知任意酉矩阵可用酉变换化为对角形，且对角元的模皆为 1。就是说，$\forall U \in U(m), \exists W \in U(m)$ 使

$$WUW^{-1} = D$$

式中 D 为对角矩阵，对角元为 $e^{i\varphi_1}, \cdots, e^{i\varphi_m}$（其中 $\varphi_1, \cdots, \varphi_m \in \mathbb{R}$）。设 Φ 是以 $\varphi_1, \cdots, \varphi_m$ 为对角元的实对角矩阵，则可验证 $D = \mathrm{Exp}(i\Phi)$。由上式得

$$U = W^{-1}DW = W^{-1}(\mathrm{Exp}(i\Phi))W = \mathrm{Exp}(iW^{-1}\Phi W)$$

其中最末一步可参考前述证明中的类似处理。令 $A \equiv iW^{-1}\Phi W$，则

$$A^{\dagger} = \overline{A^{T}} = iW^{T}\overline{\Phi^{T}(W^{-1})}^{T} = -iW^{-1}\Phi W = -A$$

表明 A 为反厄米矩阵，故 $A \in \mathfrak{u}(m)$。于是上式相当于 $U = \mathrm{Exp}\,A$，可见 $U(m)$ 中存在连续曲线 $\gamma(t) \equiv \mathrm{Exp}(tA)$ 使 $\gamma(0) = I, \gamma(1) = U$。 □

注 7.15. 以上证明中的连续曲线 $\gamma(t)$ 其实是单参子群，因此我们证明了比定理的结论更强的结论：酉群的任一群元都属于某一单参子群。这一结论也适用于 $SO(m)$ 群。然而并非任一李群的含 e 的连通分支的任一群元都属于某一单参子群。例如 $\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \in GL(2)$ 就不是。存在这样的定理：紧致的连通李群的指数映射是到上映射，这等价于任一群元都属于某一单参子群。

$U(m)$ 的子集 $SU(m) \equiv \{U \in U(m) \mid \det U = 1\}$ 是 $U(m)$ 的李子群，称为特殊酉群。同 $U(m)$ 一样，$SU(m)$ 也是紧致的连通李群。

定理 7.24. 设 A 为任意 $m \times m$ 矩阵，则

$$\det(\mathrm{Exp}\,A) = e^{\mathrm{tr}\,A}$$

证明. 设 $f(t) \equiv \det(\mathrm{Exp}\,tA)$，则

$$\frac{d}{dt}f(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} f(t+s) = \left.\frac{\partial}{\partial s}\right|_{s=0} [\det(\mathrm{Exp}(t+s)A)] = \left.\frac{\partial}{\partial s}\right|_{s=0} [\det(\mathrm{Exp}\,tA)\det(\mathrm{Exp}\,sA)]$$
$$= [\det(\mathrm{Exp}\,tA)] \left.\frac{\partial}{\partial s}\right|_{s=0} [\det(\mathrm{Exp}\,sA)] = [\det(\mathrm{Exp}\,tA)]\,\mathrm{tr}\,A = f(t)\mathrm{tr}A$$

其中第五步将在下行之后补证。由上式得 $d\ln f(t)/dt = \mathrm{tr}A$，加之 $f(0) = \det I = 1$，有 $f(t) = e^{(\mathrm{tr}\,A)t}$。令 $t = 1$ 便得上式。 □

最后补证上式的第五步，即 $d/ds|_{s=0}[\det(\operatorname{Exp} sA)] = \operatorname{tr} A$。设 X 为任意 $m \times m$ 矩阵，$X^i{}_j$ 是其矩阵元，$\varepsilon_{i_1\cdots i_m}$ 代表 m 维 Levi–Civita 记号，则

$$\det X = \varepsilon_{i_1\cdots i_m} X^{i_1}{}_1 \cdots X^{i_m}{}_m$$

取 $X = \operatorname{Exp}(sA) = I + sA + s^2A^2/2! + \cdots$，有 $X^i{}_j = \delta^i{}_j + sA^i{}_j + s^2A^i{}_kA^k{}_j/2! + \cdots$，故

$$X^i{}_j|_{s=0} = \delta^i{}_j, \quad \left.\frac{d}{ds}\right|_{s=0} X^i{}_j = A^i{}_j$$

对前式求导并利用上式便可证得 $d/ds|_{s=0}[\det(\operatorname{Exp} sA)] = \operatorname{tr} A$：

$$\begin{aligned}&\left.\frac{d}{ds}\right|_{s=0} \det X = \varepsilon_{i_1\cdots i_m} \left.\left(\frac{dX^{i_1}{}_1}{ds} X^{i_2}{}_2 \cdots X^{i_m}{}_m + \cdots + X^{i_1}{}_1 X^{i_2}{}_2 \cdots \frac{dX^{i_m}{}_m}{ds}\right)\right|_{s=0}\\ &= \varepsilon_{i_1\cdots i_m}(A^{i_1}{}_1\delta^{i_2}{}_2\cdots\delta^{i_m}{}_m + \cdots + \delta^{i_1}{}_1\delta^{i_2}{}_2\cdots A^{i_m}{}_m) = (A^1{}_1 + \cdots + A^m{}_m) = \operatorname{tr} A\end{aligned}$$

定理 7.25. 特殊酉群 $SU(m)$ 的李代数

$$\mathfrak{su}(m) = \{A \in \mathfrak{u}(m) \mid \operatorname{tr} A = 0\} = \{m \times m\text{无迹反厄米矩阵}\}$$

证明.（A） 设 $A \in \mathfrak{su}(m)$，则 $\forall t \in \mathbb{R}$ 有 $\operatorname{Exp}(tA) \in SU(m)$，故 $\det(\operatorname{Exp} tA) = 1$，根据上述定理得 $e^{\operatorname{tr}(tA)} = 1$。因而 $\operatorname{tr}(tA) = i2k\pi$（其中 $k = 0$ 或整数）。但 $\operatorname{tr}(tA) = t(\operatorname{tr} A)$，逼出 $k = 0, \operatorname{tr} A = 0$，即 A 为无迹矩阵。另一方面，由 $A \in \mathfrak{su}(m)$ 易见 $A \in \mathfrak{u}(m)$，故 A 为反厄米矩阵。

（B） 设 A 是无迹反厄米矩阵，则 $\forall t \in \mathbb{R}$，tA 也是无迹反厄米矩阵。反厄米性导致 $tA \in \mathfrak{u}(m)$，从而 $\operatorname{Exp}(tA) \in U(m)$。$tA$ 的无迹性配以前述定理导致 $\det(\operatorname{Exp} tA) = 1$，由 $\operatorname{Exp}(tA) \in U(m)$ 便知 $\operatorname{Exp}(tA) \in SU(m)$。把 $\operatorname{Exp}(tA)$ 写成级数式，求导便得曲线 $\operatorname{Exp}(tA) \subset SU(m)$ 在 I 点的切矢 $d/dt|_{t=0} \operatorname{Exp}(tA) = A$。可见 $A \in \mathfrak{su}(m)$。

□

上述定理表明酉群 $U(m)$ 与其子群 $SU(m)$ 有不同的李代数。

定理 7.26.

$$\dim SU(m) = \dim \mathfrak{su}(m) = m^2 - 1$$

证明. $\mathfrak{su}(m)$ 是 $\mathfrak{u}(m)$ 的子代数：$\mathfrak{u}(m)$ 的元素 A 只当满足 $\operatorname{tr} A = 0$ 才是 $\mathfrak{su}(m)$ 的元素。A 的反厄米性使对角元为虚数，所以 $\operatorname{tr} A = 0$ 只提供一个实方程，由此便知 $\dim \mathfrak{su}(m) = m^2 - 1$。 □

7.5.5 $E(m)$ 群（欧氏群）

m 维欧氏空间的等度规群称为欧氏群，记作 $E(m)$。欧式空间有最高对称性，所以 $\dim E(m) = m(m+1)/2$。以 $E(2)$ 为例。2 维欧氏空间有 3 个独立的 Killing 矢量场，这使我们想到其上的等度规映射既有平移又有转动。先看平移。以 $\vec{v}, \vec{a}$ 代表欧氏空间的任二点，则映射 $\vec{v} \mapsto \vec{v} + \vec{a}$ 称为平移，记作 $T_{\vec{a}}$，即

$$T_{\vec{a}}\vec{v} \coloneqq \vec{v} + \vec{a}$$

取笛卡尔系使 x 轴沿 $\vec{a}$ 向。因为 Killing 场 $\partial/\partial x$ 对应的单参等度规群过任一点的轨道都是平行于 x 轴的直线，而 $T_{\vec{a}}$ 对 $\vec{v}$ 的作用是让点 $\vec{v}$ 沿平行于 x 轴的直线移过距离 a，所以 $T_{\vec{a}}$ 是该群的一元，因而是等度规映射。因为平移的复合还是平移（$T_{\vec{a}} \circ T_{\vec{b}} = T_{\vec{a}+\vec{b}} = T_{\vec{b}} \circ T_{\vec{a}}$），所以 $T(2) \equiv \{T_{\vec{a}} \mid \vec{a} \in \mathbb{R}^2\}$ 以复合映射为乘法构成群（而且是阿贝尔群）。又因 $\vec{a}$ 相当于 2 个实数，所以 $\dim T(2) = 2$。再看转动。因为 Killing 场 $\partial/\partial\varphi \equiv -y\partial/\partial x + x\partial/\partial y$ 对应的单参等度规群过任一点的轨道是以原点为心的圆周，不难看出绕原点转 α 角的映射 $Z(\alpha)$ 是该群的一元，所以 $Z(\alpha)$ 也是等度规映射。集合 $\{Z(\alpha) \mid 0 \leqslant \alpha < 2\pi\}$ 正是李群 $SO(2)$。设 E 是 $E(2)$ 的任一元素，它把欧氏空间的原点 $\vec{o} \equiv (0,0)$ 映为点 $\vec{a}$，即 $E\vec{o} = \vec{a}$，则 $T_{-\vec{a}}E\vec{o} = \vec{o}$，故 $T_{-\vec{a}}E \in E(2)$ 是保持原点不动的等度规映射。可以证明，保原点的等度规映射要么是转动，要么是反演，故 $T_{-\vec{a}}E \in O(2)$。记 $Z \equiv T_{-\vec{a}}E$，则 $E = T_{\vec{a}}Z$。可见 $\forall E \in E(2)$，存在唯一的 $T_{\vec{a}} \in T(2), Z \in O(2)$ 使 $E = T_{\vec{a}}Z$，说明 E 可表为有序对 $(T_{\vec{a}}, Z)$，因而作为集合或拓扑空间有 $E(2) = T(2) \times O(2)$。因为 $E(2)$ 的群乘法定义为映射的复合，所以 $(T_{\vec{a}}, Z)$ 对 $\vec{v}$ 的作用表现为

$$(T_{\vec{a}}, Z)\vec{v} = T_{\vec{a}}Z\vec{v} = Z\vec{v} + \vec{a}$$

设 $T_{\vec{a}_1}, T_{\vec{a}_2} \in T(2), Z_1, Z_2 \in O(2)$，注意到

$$(T_{\vec{a}_1}, Z_1)(T_{\vec{a}_2}, Z_2)\vec{v} = (T_{\vec{a}_1}, Z_1)(Z_2\vec{v} + \vec{a}_2) = Z_1(Z_2\vec{v} + \vec{a}_2) + \vec{a}_1 = (T_{\vec{a}_1 + Z_1\vec{a}_2}, Z_1Z_2)\vec{v}$$

并把有序对 $(T_{\vec{a}}, Z)$ 简记作 $(\vec{a}, Z)$，则上式表明 $E(2)$ 的群元 $(T_{\vec{a}_1}, Z_1)$ 与 $(T_{\vec{a}_2}, Z_2)$ 间的乘法为

$$(\vec{a}_1, Z_1)(\vec{a}_2, Z_2) = (\vec{a}_1 + Z_1\vec{a}_2, Z_1Z_2)$$

同半直积群的定义对比可知 $E(2)$ 是 $T(2)$ 和 $O(2)$ 的半直积群，其中 Z_1 同态对应于 $T(2)$ 的自同构群 $A(T(2))$ 的元素 ψ_{Z_1}，它对 $T(2)$ 的作用为 $\psi_{Z_1}\vec{a}_2 = Z_1\vec{a}_2\ \forall \vec{a}_2 \in T(2)$（请注意 $Z_1\vec{a}_2$ 是 $T_{Z_1\vec{a}_2}$ 的简写。）于是

$$E(2) = T(2) \otimes_S O(2)$$

以上讨论的思路也适用于 $E(3)$，因而也有

$$E(3) = T(3) \otimes_S O(3)$$

表 7.1: 常用矩阵李群一览表

符号	李群名称	连通性	矩阵	维数	其李代数的矩阵
$GL(m)$	一般线性群（实）	不连通	$m \times m$ 可逆实矩阵	m^2	$m \times m$ 任意实矩阵
$GL(m, \mathbb{C})$	一般线性群（复）	连通	$m \times m$ 可逆复矩阵	$2m^2$	$m \times m$ 任意复矩阵
$SL(m)$	特殊线性群（实）	连通	行列式为 1 的 $m \times m$ 可逆实矩阵	$m^2 - 1$	$m \times m$ 无迹实矩阵
$SL(m, \mathbb{C})$	特殊线性群（复）	连通	行列式为 1 的 $m \times m$ 可逆复矩阵	$2m^2 - 2$	$m \times m$ 无迹复矩阵
$O(m)$	正交群	不连通	正交实矩阵	$\frac{m(m-1)}{2}$	$m \times m$ 反称实矩阵
$SO(m)$	转动群（特殊正交群）	连通	行列式为 1 的正交实矩阵	$\frac{m(m-1)}{2}$	$m \times m$ 反称实矩阵
$O(1,3)$	洛伦兹群	不连通	4×4 实矩阵 Λ 满足 $\eta = \Lambda^T \eta \Lambda$	6	4×4 实矩阵 A 满足 $A^T = -\eta A \eta$
$L_+^\uparrow$	固有洛伦兹群	连通	4×4 实矩阵 Λ 满足 $\eta = \Lambda^T \eta \Lambda$ $\det \Lambda = 1, \Lambda^0{}_0 \geqslant 1$	6	4×4 实矩阵 A 满足 $A^T = -\eta A \eta$
$U(m)$	酉群	连通	$m \times m$ 酉矩阵	m^2	$m \times m$ 反厄米复矩阵
$SU(m)$	特殊酉群	连通	行列式为 1 的 $m \times m$ 酉矩阵	$m^2 - 1$	$m \times m$ 反厄米无迹矩阵

7.5.6 Poincaré 群（庞加莱群）

4 维闵氏时空的等度规群称为 Poincaré 群，记作 P。闵氏时空的最高对称性导致 Poincaré 群是 10 维李群。沿着对 $E(2)$ 群的讨论思路，可知 Poincaré 群是 4 维平移群 $T(4)$ 与 6 维洛伦兹群 $L \equiv SO(1,3)$ 的半直积群。即

$$P = T(4) \otimes_S L$$

具体说，以 a 代表闵氏时空起自原点的矢量，则 $T_a \in T(4)$。设 $T_{a_1}, T_{a_2} \in T(4)$，$\Lambda_1, \Lambda_2 \in L$，则有序对 $(T_{a_1}, \Lambda_1), (T_{a_2}, \Lambda_2) \in T(4) \times L$。把有序对简记作 $(a_1, \Lambda_1), (a_2, \Lambda_2)$，按复合映射得

$$(a_1, \Lambda_1)(a_2, \Lambda_2) = (a_1 + \Lambda_1 a_2, \Lambda_1 \Lambda_2)$$

故 P 是 $T(4)$ 与 L 的半直积群。把上式的 L 改为 $L_+^\uparrow$ 得到的 P 的子群称为固有 Poincaré 群，记作 P_P。

7.6 李代数的结构常数

设 $\mathscr{V}$ 是李代数，则李括号 $[,]: \mathscr{V} \times \mathscr{V} \to \mathscr{V}$ 是双线性映射。由“张量面面观”可知它可以看作 $\mathscr{V}$ 上的一个 $(1,2)$ 型张量。把这一张量记作 $C^c{}_{ab}$（a, b, c 代表矢量空间 $\mathscr{V}$ 的抽象指标），便有

$$[v,u]^c = C^c{}_{ab}v^a u^b, \ \forall v^a, u^b \in \mathscr{V}$$

由上式定义的张量 $C^c{}_{ab}$ 称为李代数 $\mathscr{V}$ 的结构（常）张量。阿贝尔李代数的结构常张量显然为零。可以证明，连通李群为阿贝尔群当且仅当其李代数为阿贝尔李代数。

定理 7.27. 李代数的结构张量 $C^c{}_{ab}$ 有以下性质：

(*a*) $C^c{}_{ab} = -C^c{}_{ba}$

(*b*) $C^c{}_{a[b}C^a{}_{de]} = 0$

证明. 由 $[v,u]^a = -[u,v]^a$ 易见第一条成立。利用李括号满足的雅克比恒等式可证明第二条也成立。 □

设 $\{(e_\mu)^a\}$ 是 $\mathscr{V}$ 的任一基底，则

$$[e_\mu, e_\nu]^c = C^c{}_{ab}(e_\mu)^a(e_\nu)^b = C^\sigma{}_{\mu\nu}(e_\sigma)^c$$

上式的 $C^\sigma{}_{\mu\nu}$ 称为李代数 $\mathscr{V}$ 的结构常数。与结构张量 $C^c{}_{ab}$ 不同，结构常数 $C^\sigma{}_{\mu\nu}$ 除依赖李代数本身外还依赖于对基底的人为选择。基底变换时结构常数按张量分量变换律变换。李氏的一个基本定理断言：给定一组满足以下两条件的常数 $C^\sigma{}_{\mu\nu}$：① $C^\sigma{}_{\mu\nu} = -C^\sigma{}_{\nu\mu}$；② $C^\sigma{}_{\mu[\nu}C^\mu{}_{\rho\tau]} = 0$，必存在李群，其李代数以 $C^\sigma{}_{\mu\nu}$ 为结构常数。

例 7.11. 李群 $\mathbb{R}^2$ 的李代数的结构张量。

把 $\mathbb{R}^2$ 的每点看作从 $(0,0)$ 点出发的一个矢量，以矢量加法为群乘法，则 $\mathbb{R}^2$ 是以 $(0,0)$ 点为恒等元 e 的 2 维李群。易见这是阿贝尔群。可以证明：① 过 e 的每一个直线是一个单参子群；② $\mathbb{R}^2$ 上任一左不变矢量场 $\bar{A}$ 满足 $\partial_a \bar{A}^b = 0$（其中 ∂_a 是自然坐标系的普通导数算符）。等价地，$\bar{A}$ 的积分曲线族是平行直线族[30]。由②可知任意两个左不变矢量场 $\bar{A}$，$\bar{B}$ 的对易子为零[31]：$[\bar{A}, \bar{B}] = 0$，故结构张量 $C^a{}_{bc} = 0$。这是当然的，因为 $\mathbb{R}^2$ 是连通的阿贝尔李群，其李代数自然是阿贝尔代数。类似地可定义李群 $\mathbb{R}^n$，它当然也是阿贝尔群，其李代数也是阿贝尔李代数。

例 7.12. 李代数 $\mathfrak{so}(3)$ 的结构张量。

[30] 考虑任一左不变矢量场过 e 的不可延积分曲线，在这种情况下，它就是 $\mathbb{R}^2$ 的单参子群，即过 e 的一条直线。因此，$\bar{A}_{\gamma(s)} = \frac{d}{dt}\big|_{t=s}\gamma(t)$。设 $\gamma(t) = (ta, tb)$（即其斜率为 b/a），则 $\bar{A}_{\gamma(s)} = (a, b)$。因此，$\partial_a \bar{A}^b$ 的坐标分量 $\partial \bar{A}^\nu / \partial x^\mu = 0$。

[31] $[u,v]^\mu = u^\nu \partial_\nu v^\mu - v^\nu \partial_\nu u^\mu = 0$

前已证明 $\dim \mathfrak{so}(3) = 3$。为求得 $\mathfrak{so}(3)$ 的 3 个基矢，可借用 SO(3) 群的 3 个典型的单参子群，它们分别由绕 x 轴、y 轴和 z 轴的转动组成，以转角 α 为参数，则三个单参子群的矩阵依次为

$$Z_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, Z_y(\alpha) = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}, Z_z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

（容易验证它们满足单参子群的条件，如 $Z_z(\alpha)Z_z(\beta) = Z_z(\alpha+\beta)$）每一子群在恒等元 I 的切矢给出 $\mathfrak{so}(3)$ 的一个元素（该单参子群的生成元），分别是

$$A_1 = \left.\frac{d}{d\alpha}\right|_{\alpha=0} Z_x(\alpha) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$A_2 = \left.\frac{d}{d\alpha}\right|_{\alpha=0} Z_y(\alpha) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

和

$$A_3 = \left.\frac{d}{d\alpha}\right|_{\alpha=0} Z_z(\alpha) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A_1, A_2, A_3 显然线性独立，因而构成 $\mathfrak{so}(3)$ 的一个基底。注意到矩阵李群的李代数的李括号等于矩阵对易子，就不难验证这 3 个基矢中任意两个的李括号为

$$[A_1, A_2] = A_3, [A_2, A_3] = A_1, [A_3, A_1] = A_2$$

即 $[A_i, A_j] = \varepsilon^k{}_{ij} A_k$，$i, j, k = 1, 2, 3$ 其中 $\varepsilon^k{}_{ij}$ 是 *Levi–Civita* 记号。可见 $\mathfrak{so}(3)$ 在基底 $\{A_1, A_2, A_3\}$ 下的结构常数是 $\varepsilon^k{}_{ij}$。

例 7.13. 洛伦兹李代数 $\mathfrak{o}(1,3)$ 的结构张量。

仿照上例的做法，先写出固有洛伦兹群 $L_+^\uparrow$ 的 6 个典型的单参子群的矩阵：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\alpha & -\sin\alpha \\ 0 & 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & 0 & \sin\alpha \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\alpha & 0 & \cos\alpha \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} \operatorname{ch}\lambda & -\operatorname{sh}\lambda & 0 & 0 \\ -\operatorname{sh}\lambda & \operatorname{ch}\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \operatorname{ch}\lambda & 0 & -\operatorname{sh}\lambda & 0 \\ 0 & 1 & 0 & 0 \\ -\operatorname{sh}\lambda & 0 & \operatorname{ch}\lambda & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \operatorname{ch}\lambda & 0 & 0 & -\operatorname{sh}\lambda \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\operatorname{sh}\lambda & 0 & 0 & \operatorname{ch}\lambda \end{bmatrix}$$

其中前 3 个代表 4 维闵氏时空中的空间转动，后 3 个代表伪转动。把它们的生成元分别记作 r_1, r_2, r_3 和 b_1, b_2, b_3，则由求导可得

$$r_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, r_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, r_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$b_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, b_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, b_3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

易见 $\mathfrak{o}(1,3)$ 的这 6 个元素线性独立，因而构成一个基底。不难验证它们的李括号满足下式（它的某些重要后果将在第 9 节讨论）：

$$[r_i, r_j] = \varepsilon^k{}_{ij} r_k,\ [b_i, r_j] = \varepsilon^k{}_{ij} b_k,\ [b_i, b_j] = -\varepsilon^k{}_{ij} r_k,\ i, j, k = 1, 2, 3$$

李代数 $\mathfrak{o}(1,3)$ 在基底 $\{r_1, r_2, r_3, b_1, b_2, b_3\}$ 下的结构常数可从上式读出。引入符号 $l_{\mu\nu}$（$\mu, \nu = 0, 1, 2, 3$ 且 $l_{\mu\nu}$ 反称），其含义为

$$l_{01} \equiv b_1, l_{02} \equiv b_2, l_{03} \equiv b_3, l_{12} \equiv r_3, l_{23} \equiv r_1, l_{31} \equiv r_2$$

则不难验证前式可统一表为下式：

$$[l_{\mu\nu}, l_{\rho\sigma}] = \eta_{\mu\rho} l_{\nu\sigma} + \eta_{\nu\sigma} l_{\mu\rho} - \eta_{\mu\sigma} l_{\nu\rho} - \eta_{\nu\rho} l_{\mu\sigma}$$

每一个 $l_{\mu\nu}$ 可看作一个反称实矩阵，矩阵元可表为

$$(l_{\mu\nu})^\alpha{}_\beta = -\delta^\alpha{}_\mu \eta_{\beta\nu} + \delta^\alpha{}_\nu \eta_{\beta\mu}$$

例 7.14. 特殊酉群 $SU(2)$ 的李代数 $\mathfrak{su}(2)$ 的结构张量。

前面已知 $\dim \mathfrak{su}(2) = 3$，且 $\mathfrak{su}(2)$ 的元素是 2×2 无迹反厄米矩阵。大家熟知的 3 个泡利矩阵 $\tau_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tau_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \tau_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 是无迹厄米矩阵，乘以 i 就是无迹反厄米矩阵，而且互相线性独立，可充当 $\mathfrak{su}(2)$ 的基底。为了得到简单的结构常数，我们不用 i 而用 $-i/2$ 乘以泡利矩阵，从而得到 $\mathfrak{su}(2)$ 的如下 3 个基矢：

$$E_i = -i\tau_i/2,\ i = 1, 2, 3$$

不难验证

$$[E_1, E_2] = E_3,\ [E_2, E_3] = E_1,\ [E_3, E_1] = E_2$$

即

$$[E_i, E_j] = \varepsilon^k{}_{ij} E_k,\ i, j, k = 1, 2, 3$$

这表明 $\mathfrak{su}(2)$ 在基底 $\{E_1, E_2, E_3\}$ 下的结构常数与 $\mathfrak{so}(3)$ 在基底 $\{A_1, A_2, A_3\}$ 下的结构常数相同。（如果以 i 而不是 $-i/2$ 乘泡利矩阵，则结构常数将不满足上式。）令线性映射 $\psi\colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ 满足 $\psi(E_i) = A_i$（$i = 1, 2, 3$），就可看出 ψ 保李括号，所以是 $\mathfrak{su}(2)$ 与 $\mathfrak{so}(3)$ 之间的李代数同构。应该注意，虽然 $\mathfrak{su}(2)$ 的元素是复矩阵，但 $\mathfrak{su}(2)$ 仍是实的 3 维矢量空间：只有以实数为组合系数才能确保所构成的线性组合仍在 $\mathfrak{su}(2)$ 之内。

$\mathfrak{su}(2)$ 与 $\mathfrak{so}(3)$ 同构并不意味着 $SU(2)$ 与 $SO(3)$ 同构。事实上，不存在 $SU(2)$ 与 $SO(3)$ 之间的同构映射，但存在 2 对 1 的李群同态映射 $\pi\colon SU(2) \to SO(3)$，即 $SU(2)$ 的两个不同群元对应于 $SO(3)$ 的一个群元。前面说过，一个李群决定唯一的李代数，但一个李代数并不决定唯一的李群（除非待定的是单连通李群）。$SU(2)$ 与 $SO(3)$ 的关系是这一重要结论的一个很好的例子：

$$\mathfrak{su}(2) = \mathfrak{so}(3) \nRightarrow SU(2) = SO(3)$$

$SU(2)$ 与 $SO(3)$ 都是连通流形，但可以证明 $SU(2)$ 是单连通流形而 $SO(3)$ 不是。正是这一非单连通性使 $SU(2) \neq SO(3)$。与此类似，群 $SL(2, \mathbb{C})$ 与固有洛伦兹群 $L_+^\uparrow$ 有相同李代数，但两群只同态不同构。存在 2 对 1 的李群同态映射 $\pi\colon SL(2, \mathbb{C}) \to L_+^\uparrow$。

7.7 李变换群和 Killing 矢量场

把流形 M 上的单参微分同胚群 $\phi\colon \mathbb{R} \times M \to M$ 中的 1 维阿贝尔李群 $\mathbb{R}$ 推广为任意李群 G 便有如下定义

定义 7.26. 设 G 是李群，M 是流形。C^∞ 映射 $\sigma\colon G \times M \to M$ 称为 M 上的一个李变换群，若

（*a*） $\forall g \in G, \sigma_g\colon M \to M$ 是微分同胚；

（*b*） $\sigma_{gh} = \sigma_g \circ \sigma_h,\ \forall g, h \in G$

容易验证 $\{\sigma_g \mid g \in G\}$ 以映射的复合为群乘法构成群，恒等元就是从 M 到 M 的恒等映射（等于 σ_e，其中 e 为 G 的恒等元），且 $(\sigma_g)^{-1} = \sigma_{g^{-1}}\ \forall g \in G$。有些文献把 $\{\sigma_g \mid g \in G\}$ 称为李变换群（我们有时也用这一称谓），有些则把 G 称为李变换群。$\forall p \in M$，有 $\sigma_p\colon G \to M$。σ_p 和 σ_g 都是由 σ 诱导出的映射，三者关系为

$$\sigma_p(g) = \sigma(p, g) = \sigma_g(p),\ \forall g \in G, p \in M$$

李群按定义是一个高度抽象的概念，而李变换群却比较具体：它的群元 σ_g 是流形 M 上的点变换（微分同胚）。例如，抽象定义的李群 $SO(3)$ 若借用一个 2 维球面 S^2 来想象就

变得具体：每一 $g \in SO(3)$ 对应于 S^2 上这样一个点变换，它让 S^2 的每点绕过球心的同一轴转同一角。事实上，上述定义的条件（b）保证存在从李群 $G = \{g\}$ 到李变换群 $\{\sigma_g\}$ 的同态映射，所以后者的确可看作前者的具体化，于是有以下定义：

定义 7.27. 从李群 $G = \{g\}$ 到李变换群 $\{\sigma_g : M \to M \mid g \in G\}$ 的上述同态映射 $G \to \{\sigma_g\}$ 称为 G 的一个实现，M 称为实现空间。若此同态为同构，则称为忠实实现。

注 7.16. 因为每一 $g \in G$ 对应于一个左平移映射 $L_g : G \to G$（看作 $\sigma_g : M \to M$，其中 $M = G$），所以映射 $G \to \{L_g \mid g \in G\}$ 是 G 的一个实现（实现空间就是 G 本身），而且是忠实实现。

定义 7.28. 李群 G 在流形 M 上的一个实现 $G \to \{\sigma_g\}$ 称为 G 的一个表示，若 M 是矢量空间且 $\sigma_g : M \to M$（$\forall g \in G$）是线性变换。这时称 M 为表示空间。往往也把映射 $G \to \{\sigma_g\}$ 的像 $\{\sigma_g\}$ 称为 G 的表示（但应注意不同映射有相同像的情况）。若忠实实现是表示，则称为忠实表示。

注 7.17. ① 可见 G 的每一表示 $\{\sigma_g\}$ 可看作一个矩阵群。② 因为 S^2 不是矢量空间，把 $SO(3)$ 的群元看作 S^2 上的转动这种对应关系只是实现而非表示。但若选 $M = \mathbb{R}^3$，则 $SO(3)$ 的每一群元对应于 $\mathbb{R}^3$ 上的这样一个微分同胚，它让每点绕过原点的同一轴转同一角。这种对应关系就是一个表示，而且是忠实表示。

设 $\gamma : \mathbb{R} \to G$ 是 G 的、与 $A \in V_e$ 对应的单参子群，则 $\{\gamma(t) \mid t \in \mathbb{R}\}$ 是 G 中的曲线。把群 $\{\sigma_g \mid g \in G\}$ 中 g 的取值范围限制在曲线 $\gamma(t)$ 上便得子集 $\{\sigma_{\gamma(t)} \mid t \in \mathbb{R}\}$，其中每个 $\sigma_{\gamma(t)} : M \to M$ 都是微分同胚，而且 $\sigma_{\gamma(t)} \circ \sigma_{\gamma(s)} = \sigma_{\gamma(t+s)}$，可见 G 的每一单参子群决定 M 上的一个单参微分同胚群 $\{\sigma_{\gamma(t)} \mid t \in \mathbb{R}\}$。我们想找出与之对应的 C^∞ 矢量场（记作 $\bar{\xi}$）。因为 $\bar{\xi}$ 的积分曲线与单参微分同胚群 $\{\sigma_{\gamma(t)} \mid t \in \mathbb{R}\}$ 的轨道重合，所以 $\forall p \in M$，$\bar{\xi}$ 等于过 p 点的轨道在 p 点的切矢。群 $\{\sigma_{\gamma(t)} \mid t \in \mathbb{R}\}$ 过 p 的轨道是点集 $\{\sigma_{\gamma(t)}(p) \mid t \in \mathbb{R}\}$，而 $\sigma_{\gamma(t)}(p) = \sigma(\gamma(t), p) = \sigma_p(\gamma(t))$，于是

$$\bar{\xi}_p = \left.\frac{d}{dt}\right|_{t=0} \sigma_p(\gamma(t)) = \sigma_{p*} \left.\frac{d}{dt}\right|_{t=0} \gamma(t) = \sigma_{p*} A$$

可见，给定李变换群 $\sigma : G \times M \to M$ 后，G 的李代数 V_e 的每一元素 A 对应于 M 上的一个 C^∞ 矢量场 $\bar{\xi}$，故有映射 $\chi : V_e \to \{\bar{\xi}\}$。为理解 $\{\bar{\xi}\}$ 的意义，先看一类特殊情况：M 是带度规 g_{ab} 的流形，G 是 (M, g_{ab}) 的等度规群，李变换群 $\sigma : G \times M \to M$ 的每一 $\sigma_g : M \to M$ 是等度规映射。这时 G 的每一单参子群 $\gamma(t)$ 产生的单参微分同胚群 $\{\sigma_{\gamma(t)} \mid t \in \mathbb{R}\}$ 升格为单参等度规群，其轨道的切矢 $\bar{\xi}$ 就是 (M, g_{ab}) 上的 Killing 矢量场，即 $\{\bar{\xi}\} \subset \mathscr{K}$（$\mathscr{K}$ 代表 (M, g_{ab}) 上全体 Killing 场构成的矢量空间），故 χ 可看作映射 $V_e \to \mathscr{K}$（可证其为线性映射）。以矢量场对易子为 $\mathscr{K}$ 的李括号，则 $\mathscr{K}$ 成为李代数，且可证明 $\chi : V_e \to \mathscr{K}$ 对李括号保到只差一个负号的程度：

$$\chi([A, B]) = -[\chi(A), \chi(B)], \ \forall A, B \in V_e$$

用下式定义新映射 $\psi\colon V_e \to \mathscr{K}$：

$$\psi(A) = -\chi(A),\ \forall A \in V_e$$

则易见 ψ 保李括号，即

$$\psi([A,B]) = [\psi(A), \psi(B)],\ \forall A, B \in V_e$$

故 ψ 是李代数同态。如果每个 Killing 场都是完备矢量场，则每个都产生单参等度规群，故等度规群 G（因而 V_e）与 $\mathscr{K}$ 维数相同，$\psi\colon V_e \to \mathscr{K}$ 是李代数同构。反之，不完备 Killing 场的单参等度规群 $\{\phi\colon M \to M\}$ 是只含恒等元的 0 维（而不是 1 维）李群，因为无论参数 t 取何值（只要非零），总有 $p \in M$ 使 p 在 ϕ_t 下无像。于是 $\dim G$ 减少而 $\dim \mathscr{K}$ 不变，导致 $\dim G = \dim V_e < \dim \mathscr{K}$，意味着 V_e 与 $\mathscr{K}$ 只同态不同构。下面是简单特例：把 $(\mathbb{R}^2, \delta_{ab})$ 的一半（连边）挖去，则除了平行于切痕的那些平移 Killing 场之外的 Killing 场（含旋转 Killing 场）都不再完备，故 $\dim G = \dim V_e = 1$（而 $\dim \mathscr{K}$ 仍为 3）。于是有

定理 7.28. (M, g_{ab}) 的等度规群的李代数 V_e 同构于其上全体 *Killing* 场的集合 $\mathscr{K}$ 的李子代数；当每一 *Killing* 场都完备时 $V_e = \mathscr{K}$（李代数同构）

利用 Killing 场的对易子可方便地求得等度规群的李代数的结构常数，我们只举例说明 V_e 同构于 $\mathscr{K}$ 时的计算过程（略去抽象指标）。

例 7.15. 3 维欧氏空间 $(\mathbb{R}^3, \delta_{ab})$ 中的 2 维球面 (S^2, h_{ab})（其中 h_{ab} 是 δ_{ab} 的诱导度规）的等度规群是 $SO(3)$。(S^2, h_{ab}) 上的全体 *Killing* 矢量场的集合 $\mathscr{K}$ 是 3 维李代数，不难验证

$$\begin{aligned}
\xi_1 &\equiv \sin\varphi(\partial/\partial\theta) + \cot\theta\cos\varphi(\partial/\partial\varphi) \\
\xi_2 &\equiv -\cos\varphi(\partial/\partial\theta) + \cot\theta\sin\varphi(\partial/\partial\varphi) \\
\xi_3 &\equiv -\partial/\partial\varphi
\end{aligned}$$

满足 *Killing* 方程，可选作 $\mathscr{K}$ 的一组基矢。对易子的计算表明

$$[\xi_1, \xi_2] = \xi_3,\ [\xi_2, \xi_3] = \xi_1,\ [\xi_3, \xi_1] = \xi_2$$

这可看作李代数 $\mathfrak{so}(3)$ 的结构常数表达式。以 A_1, A_2, A_3 代表 $\mathfrak{so}(3)$ 的一组基矢，则由 $\psi(A_i) \equiv \xi_i$（$i = 1, 2, 3$）定义的映射 $\psi\colon \mathfrak{so}(3) \to \mathscr{K}$ 正是前述的 ψ，由上述定理可知它是李代数同构。

例 7.16. 作为 2 维欧氏空间的等度规群，欧氏群 $E(2)$ 的李代数 $\mathfrak{e}(2)$ 同构于 $(\mathbb{R}^2, \delta_{ab})$ 上全体 *Killing* 矢量场的集合 $\mathscr{K}$。设 $\{x, y\}$ 是笛卡尔系，则

$$\xi_{t_1} \equiv -\frac{\partial}{\partial x},\ \xi_{t_2} \equiv -\frac{\partial}{\partial y},\ \xi_r \equiv y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$$

是独立的 *Killing* 矢量场，可充当 $\mathscr{K}$ 的一组基矢。简单的对易子的计算表明

$$[\xi_{t_1},\xi_{t_2}]=0,\ [\xi_r,\xi_{t_1}]=\xi_{t_2},\ [\xi_r,\xi_{t_2}]=-\xi_{t_1}$$

这可看作李代数 $\mathfrak{e}(2)$ 的结构常数表达式。

例 7.17. 作为 4 维闵氏时空的等度规群，*Poincaré* 群 P 的李代数 $\mathfrak{p}$ 同构于 $(\mathbb{R}^4,\eta_{ab})$ 上全体 *Killing* 场的集合 $\mathscr{K}$。设 $\{t,x,y,z\}$ 是任一洛伦兹系，则

$$\xi_{t_0}\equiv-\frac{\partial}{\partial t},\ \xi_{t_1}\equiv-\frac{\partial}{\partial x},\ \xi_{t_2}\equiv-\frac{\partial}{\partial y},\ \xi_{t_3}\equiv-\frac{\partial}{\partial z}$$

$$\xi_{r_1}\equiv z\frac{\partial}{\partial y}-y\frac{\partial}{\partial z},\ \xi_{r_2}\equiv x\frac{\partial}{\partial z}-z\frac{\partial}{\partial x},\ \xi_{r_3}\equiv y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y}$$

$$\xi_{b_1}\equiv t\frac{\partial}{\partial x}-x\frac{\partial}{\partial t},\ \xi_{b_2}\equiv t\frac{\partial}{\partial y}-y\frac{\partial}{\partial t},\ \xi_{b_3}\equiv t\frac{\partial}{\partial z}-z\frac{\partial}{\partial t}$$

是独立的 *Killing* 场，可充当 $\mathscr{K}$ 的一组基矢。简单的计算表明全部非零对易子为

$$[\xi_{r_i},\xi_{r_j}]=\varepsilon^k{}_{ij}\xi_{r_k},\ [\xi_{r_i},\xi_{b_j}]=\varepsilon^k{}_{ij}\xi_{b_k},\ [\xi_{b_i},\xi_{b_j}]=-\varepsilon^k{}_{ij}\xi_{r_k},\ [\xi_{b_i},\xi_{t_0}]=-\xi_{t_i},$$
$$[\xi_{r_i},\xi_{t_j}]=\varepsilon^k{}_{ij}\xi_{t_k},\ [\xi_{b_i},\xi_{t_i}]=-\xi_{t_0}\text{（不对 } i \text{ 求和）},\ i,j,k=1,2,3$$

上式也可浓缩地表为（其中 $l_{\mu\nu}$ 的含义参考前文）

$$[\xi_{l_{\mu\nu}},\xi_{l_{\rho\sigma}}]=\eta_{\mu\rho}\xi_{l_{\nu\sigma}}+\eta_{\nu\sigma}\xi_{l_{\mu\rho}}-\eta_{\mu\sigma}\xi_{l_{\nu\rho}}-\eta_{\nu\rho}\xi_{l_{\mu\sigma}},$$
$$[\xi_{l_{\mu\nu}},\xi_{t_\sigma}]=\eta_{\mu\sigma}\xi_{t_\nu}-\eta_{\nu\sigma}\xi_{t_\mu},\ \mu,\nu,\sigma=0,1,2,3$$

7.8 伴随表示和 Killing 型

设 V 是 m（$<\infty$）维实矢量空间，令

$$\mathscr{L}(V)\equiv\{\text{线性变换}\psi\colon V\to V\}\equiv\{V\text{上}(1,1)\text{型张量}\psi^a{}_b\}$$

则 $\mathscr{L}(V)$ 是 m^2 维矢量空间，任选 V 的基底后又对应于 $\mathfrak{gl}(m,\mathbb{R})$。在 $\mathscr{L}(V)$ 上可自然定义乘法：$\forall\psi,\varphi\in\mathscr{L}(V)$，其乘积 $\psi\varphi$ 定义为复合映射 $\psi\circ\varphi\in\mathscr{L}(V)$。由此又可自然定义李括号映射

$$[,]\colon\mathscr{L}(V)\times\mathscr{L}(V)\to\mathscr{L}(V)$$

为

$$[\psi,\varphi]\coloneqq\psi\varphi-\varphi\psi,\ \forall\psi,\varphi\in\mathscr{L}(V)$$

从而使 $\mathscr{L}(V)$ 成为 m^2 维李代数。

定义 7.29. 李代数同态映射 $\beta\colon\mathscr{G}\to\mathscr{L}(V)$ 称为李代数 $\mathscr{G}$ 的表示。

同一李群（或李代数）可有无数表示，本节的重点是介绍李群和李代数的伴随表示。根据第一节，用群 G 的每一元素 $g \in G$ 可构造一个称为伴随同构的自同构映射 $I_g\colon G \to G$。对李群 G，这还是个微分同胚，因此是从 G 到 G 的李群同构。按定义，$I_g(h) \equiv ghg^{-1}\ \forall h \in G$，所以 $I_g(e) = e$，它在 e 点所诱导的推前映射（切映射）I_{g*}（是 I_{g*e} 的简写）是从 V_e 到 V_e 的映射，简记为 Ad_g，即 $\mathrm{Ad}_g \equiv I_{g*e}$。因 V_e 就是 G 的李代数 $\mathscr{G}$，故 $\mathrm{Ad}_g\colon \mathscr{G} \to \mathscr{G}$ 是 $\mathscr{G}$ 上的线性变换。虽然 I_g 作用于 e 得 e，但作用于过 e 的曲线一般得另一曲线（两者在 e 点有不同切矢），故 Ad_g 作用于 $A \in \mathscr{G}$ 未必得 A。

定理 7.29. 设 $\mathscr{G}$ 是李群 G 的李代数，则 $\forall g \in G, A \in \mathscr{G}$ 有

$$\exp(t\mathrm{Ad}_g A) = g(\exp tA)g^{-1}$$

证明. 令 $\gamma(t) \equiv \exp(tA), \gamma'(t) \equiv g(\exp tA)g^{-1}$，则由单参子群的定义式 $\gamma(t+s) = \gamma(t)\gamma(s)$ 易证 $\gamma'(t+s) = \gamma'(t)\gamma'(s)$，故 $\gamma'(t)$ 也是单参子群。注意到上式左边是由 $\mathrm{Ad}_g A$ 生成的单参子群，为证此式只须证明两边在 e 点的切矢相同，证明如下：

$$\begin{aligned}\text{右边的切矢} &= \left.\frac{d}{dt}\right|_{t=0}[g(\exp tA)g^{-1}] = \left.\frac{d}{dt}\right|_{t=0}[I_g(\exp tA)] \\ &= I_{g*}[\left.\frac{d}{dt}\right|_{t=0}\exp(tA)] = \mathrm{Ad}_g A = \text{左边的切矢}\end{aligned}$$

□

定理 7.30. 设 H 是李群 G 的正规子群，$\mathscr{H}$ 是 H 的李代数，则

$$\mathrm{Ad}_g B \in \mathscr{H},\ \forall B \in \mathscr{H}, g \in G$$

证明. 由 $B \in \mathscr{H}$ 可知 $\exp(tB)$ 是 H 的单参子群，由正规子群的定义又知 $g\exp(tB)g^{-1} \subset H$，于是上述定理表明 $\exp(t\mathrm{Ad}_g B)$ 是 H 的单参子群，因而

$$\mathrm{Ad}_g B \in \mathscr{H}$$

□

定理 7.31. 设 $\mathscr{G}$ 是李群 G 的李代数，则 $\forall A, B \in \mathscr{G}$ 有

$$[A, B] = \left.\frac{d}{dt}\right|_{t=0}(\mathrm{Ad}_{\exp(tA)}B)$$

注 7.18. $\mathrm{Ad}_{\exp(tA)}B$ 在 t 固定时是 e 点的矢量，在 t 变动时是 t 的、矢量取值的函数，等式右边代表此函数的导函数在 $t = 0$ 的值（仍是 e 点的矢量），即

$$[A, B] = \left.\frac{d}{dt}\right|_{t=0}(\mathrm{Ad}_{\exp(tA)}B) \equiv \lim_{t\to 0}\frac{1}{t}(\mathrm{Ad}_{\exp(tA)}B - \mathrm{Ad}_{\exp(0)}B)$$

证明. 以 $\bar{A}, \bar{B}$ 分别代表 $A, B \in \mathscr{G}$ 对应的左不变矢量场，$\phi\colon \mathbb{R}\times G \to G$ 代表 $\bar{A}$ 产生的单参微分同胚群，则 $\phi_t(e) = \exp(tA)$，即 $\bar{B}_{\phi_t(e)} = L_{\exp(tA)*}B$，因此

$$[A,B] = [\bar{A},\bar{B}]_e = (\mathscr{L}_{\bar{A}}\bar{B})_e = \lim_{t\to 0}\frac{1}{t}[(\phi_{-t*}\bar{B})_e - \bar{B}_e] = \lim_{t\to 0}\frac{1}{t}[\phi_{-t*}\bar{B}_{\phi_t(e)} - \phi_{-0*}\bar{B}_{\phi_0(e)}]$$
$$= \left.\frac{d}{dt}\right|_{t=0}[\phi_{-t*}\bar{B}_{\phi_t(e)}] = \left.\frac{d}{dt}\right|_{t=0}[\phi_{-t*}L_{\exp(tA)*}B] = \left.\frac{d}{dt}\right|_{t=0}[(\phi_{-t}\circ L_{\exp(tA)})_*B]$$

另一方面，由 $I_h(g) \equiv hgh^{-1}$ 又知 $\forall g\in G$ 有

$$I_{\exp(tA)}(g) = \exp(tA)g\exp(-tA) = \phi_{-t}[\exp(tA)g] = \phi_{-t}[L_{\exp(tA)}(g)] = (\phi_{-t}\circ L_{\exp(tA)})(g)$$

（其中第二步用到第四节末的定理。）所以 $I_{\exp(tA)} = \phi_{-t}\circ L_{\exp(tA)}$，代入上式便得

$$[A,B] = \left.\frac{d}{dt}\right|_{t=0}(I_{\exp(tA)*}B) = \left.\frac{d}{dt}\right|_{t=0}(\mathrm{Ad}_{\exp(tA)}B)$$

□

定理 7.32. 设 H 是连通李群 G 的连通李子群，$\mathscr{H}$ 和 $\mathscr{G}$ 分别是 H 和 G 的李代数，则 H 是 G 的正规子群当且仅当 $\mathscr{H}$ 是 $\mathscr{G}$ 的理想。

证明.（A） 因为 H 是 G 的正规子群，由前述定理知 $\mathrm{Ad}_{\exp(tA)}B\in\mathscr{H}$，结合上述定理就有 $[A,B]\in\mathscr{H}$，可见 $\mathscr{H}$ 是 $\mathscr{G}$ 的理想。

（B） 因为 $\mathscr{H}$ 是 $\mathscr{G}$ 的理想，故根据上述定理有 $\mathrm{Ad}_g B\in\mathscr{H}$。可以证明（略），只要 H 是连通李群的连通李子群，则 $\forall h\in H, \exists B_1, B_2, \cdots\in\mathscr{H}$ 使 $h = (\exp B_1)(\exp B_2)\cdots$（有限个指数之积）。所以

$$ghg^{-1} = g(\exp B_1)g^{-1}g(\exp B_2)g^{-1}\cdots = h_1h_2\cdots\in H$$

其中 $h_n \equiv g(\exp B_n)g^{-1} = \exp(\mathrm{Ad}_g B_n)\in H, n\in\mathbb{N}$。

□

注 7.19. 可见理想在李代数理论中的角色相当于正规子群在群论中的角色。

由推前映射的线性性可知 $I_{g*}\colon\mathscr{G}\to\mathscr{G}$ 是 $\mathscr{G}$ 上的线性映射，故 $\mathrm{Ad}_g\in\mathscr{L}(\mathscr{G})$。$I_g\colon G\to G$ 是微分同胚还保证 $I_{g*}\colon\mathscr{G}\to\mathscr{G}$ 是同构映射[32]，当然有逆，所以

$$\mathrm{Ad}_g\in\{\mathscr{G}\text{上可逆线性变换}\}\subset\mathscr{L}(\mathscr{G})$$

既然每一 $g\in G$ 对应于一个 Ad_g，便有从 G 到 $\{\mathscr{G}$上可逆线性变换$\}$ 的映射，记作 Ad，即

$$\mathrm{Ad}\colon G\to\{\mathscr{G}\text{上可逆线性变换}\}$$

[32]其实 $I_{g*}\colon\mathscr{G}\to\mathscr{G}$ 不止是矢量空间同构而且是李代数同构。

选定基底后 $\mathscr{G}$ 上的每个可逆线性变换对应于一个可逆矩阵，故 {$\mathscr{G}$上可逆线性变换} 对应于一个矩阵群（就是 $GL(m,\mathbb{R})$，其中 $m\equiv\dim(G)$）。下面证明 Ad 是同态映射，因而是李群 G 的表示。

定理 7.33. $\mathrm{Ad}\colon G\to\{\mathscr{G}$上可逆线性变换$\}$ 是同态映射。

证明. 不难证明 $\forall g,h\in G$ 有 $I_{gh}=I_g\circ I_h$，故

$$\mathrm{Ad}_{gh}=I_{gh*}=(I_g\circ I_h)_*=I_{g*}\circ I_{h*}=\mathrm{Ad}_g\circ\mathrm{Ad}_h$$

同态性于是得证。 □

定义 7.30. 同态映射 $\mathrm{Ad}\colon G\to\{\mathscr{G}$上可逆线性变换$\}$ 称为李群 G 的伴随表示。

至今已讲过三个与“伴随”有关的映射，即伴随同构 $I_g\colon G\to G$；伴随同构的切映射 $\mathrm{Ad}_g\colon\mathscr{G}\to\mathscr{G}$；伴随表示 $\mathrm{Ad}\colon G\to\{\mathscr{G}$上可逆线性变换$\}$。现在介绍第四个，即 $\mathrm{ad}_A\colon\mathscr{G}\to\mathscr{G}$（其中 A 不是群元而是代数元：$A\in\mathscr{G}$），定义为

$$\mathrm{ad}_A(B):=[A,B],\ \forall B\in\mathscr{G}$$

由李括号的线性性可知 $\mathrm{ad}_A\colon\mathscr{G}\to\mathscr{G}$ 有如下两个性质：

（a） $\forall B_1,B_2\in\mathscr{G},\beta_1,\beta_2\in\mathbb{R}$ 有 $\mathrm{ad}_A(\beta_1B_1+\beta_2B_2)=\beta_1\mathrm{ad}_A(B_1)+\beta_2\mathrm{ad}_A(B_2)$

（b） $\forall A_1,A_2\in\mathscr{G},\alpha_1,\alpha_2\in\mathbb{R}$ 有 $\mathrm{ad}_{\alpha_1A_1+\alpha_2A_2}=\alpha_1\mathrm{ad}_{A_1}+\alpha_2\mathrm{ad}_{A_2}$

性质（a）表明 ad_A 是 $\mathscr{G}$ 上的线性变换，即 $\mathrm{ad}_A\in\mathscr{L}(\mathscr{G})$，也可看作 $\mathscr{G}$ 上的 $(1,1)$ 型张量，所以上式在补上抽象指标后又可用 $\mathscr{G}$ 的结构张量 $C^c{}_{ab}$ 改写为

$$(\mathrm{ad}_A)^c{}_bB^b\equiv[A,B]^c=C^c{}_{ab}A^aB^b,\ \forall B^b\in\mathscr{G}$$

从而给出张量 $(\mathrm{ad}_A)^c{}_b$ 在甩掉作用对象 B^b 后的表达式

$$(\mathrm{ad}_A)^c{}_b=A^aC^c{}_{ab}$$

映射 ad_A（其中 $A\in\mathscr{G}$）虽不同于 Ad_g（其中 $g\in G$），但两者有如下密切关系：

$$\mathrm{Ad}_{\exp(A)}=\mathrm{Exp}(\mathrm{ad}_A)$$

其中 $\mathrm{Exp}(\mathrm{ad}_A)$ 的含义是

$$\mathrm{Exp}(\mathrm{ad}_A)\equiv\delta+\mathrm{ad}_A+\frac{1}{2!}(\mathrm{ad}_A)^2+\frac{1}{3!}(\mathrm{ad}_A)^3+\cdots$$

这可看作 $\mathscr{G}$ 上的 $(1,1)$ 型张量等式，δ 代表恒等映射（即 $\delta^a{}_b$），$(\mathrm{ad}_A)^2$ 代表 $(\mathrm{ad}_A)^a{}_c(\mathrm{ad}_A)^c{}_b$ 等等。因 ad_A 可看作抽象李代数 $\mathscr{L}(\mathscr{G})$ 的元素，$\mathrm{Exp}(\mathrm{ad}_A)$ 也可理解为 $\exp(\mathrm{ad}_A)$，即李代数 $\mathscr{L}(\mathscr{G})$ 相应的（单连通）李群上的指数映射。这一理解与用幂级数的理解等价。

既然每一 $A \in \mathscr{G}$ 对应于一个 $\mathrm{ad}_A \in \mathscr{L}(\mathscr{G})$，便有从 $\mathscr{G}$ 到 $\mathscr{L}(\mathscr{G})$ 的映射（与"伴随"有关的第五个映射），记作 $\mathrm{ad}\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$，$\mathrm{ad}_A$ 的性质（b）表明 ad 为线性映射。由李括号所满足的雅克比恒等式易证 ad_A 满足

$$\mathrm{ad}_{[A,B]} = \mathrm{ad}_A \mathrm{ad}_B - \mathrm{ad}_B \mathrm{ad}_A,\ \forall A, B \in \mathscr{G}$$

即 $\mathrm{ad}\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$ 保李括号，因而是同态映射。

定义 7.31. 同态映射 $\mathrm{ad}\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$ 称为李代数 $\mathscr{G}$ 的伴随表示。

注意到 $\{\mathscr{G}\text{上可逆线性变换}\} \subset \mathscr{L}(\mathscr{G})$，也可写 $\mathrm{Ad}\colon G \to \mathscr{L}(\mathscr{G})$。由 $\mathrm{Ad}_e\colon \mathscr{G} \to \mathscr{G}$ 是恒等映射可知 $\mathrm{Ad}(e)$ 是 $\mathscr{L}(\mathscr{G})$ 中的单位矩阵 I，故 $\mathrm{Ad}\colon G \to \mathscr{L}(\mathscr{G})$ 在点 e 诱导出推前映射 $\mathrm{Ad}_*\colon V_e \to V_I$。注意到 $V_e = \mathscr{G}$ 以及矢量空间 $\mathscr{L}(\mathscr{G})$ 可与其中任一点的切空间（现在取 V_I）认同，便有 $\mathrm{Ad}_*\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$，即映射 Ad_* 与 ad 有相同的定义域和值域。下面证明它们是相等的映射。

定理 7.34. 映射 $\mathrm{Ad}_*\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$ 与 $\mathrm{ad}\colon \mathscr{G} \to \mathscr{L}(\mathscr{G})$ 相等。

证明. 只须证 $\mathrm{Ad}_* A = \mathrm{ad}_A\ \forall A \in \mathscr{G}$，为此只须证 $(\mathrm{Ad}_* A)(B) = \mathrm{ad}_A(B)\ \forall A, B \in \mathscr{G}$。因

$$\mathrm{Ad}_* A = \mathrm{Ad}_*[\left.\frac{d}{dt}\right|_{t=0} \exp(tA)] = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\exp tA) = \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp tA}$$

故

$$\mathrm{Ad}_* A(B) = [\left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{\exp tA}](B) = \left.\frac{d}{dt}\right|_{t=0} (\mathrm{Ad}_{\exp tA} B) = [A, B] = \mathrm{ad}_A(B)$$

□

利用 $\mathscr{G}$ 的伴随表示可定义映射 $K\colon \mathscr{G} \times \mathscr{G} \to \mathbb{R}$ 如下：

$$K(A, B) \coloneqq \mathrm{tr}(\mathrm{ad}_A \mathrm{ad}_B),\ \forall A, B \in \mathscr{G}$$

其中 $\mathrm{ad}_A \mathrm{ad}_B$ 表示两个线性变换的相继作用，也可用抽象指标表为

$$(\mathrm{ad}_A)^a{}_c (\mathrm{ad}_B)^c{}_b = (\mathrm{ad}_A \mathrm{ad}_B)^a{}_b$$

而 $\mathrm{tr}(\mathrm{ad}_A \mathrm{ad}_B)$ 则代表 $(\mathrm{ad}_A \mathrm{ad}_B)^a{}_a$，即 $(1,1)$ 型张量 $(\mathrm{ad}_A \mathrm{ad}_B)^a{}_b$ 的迹。由上式可知 K 是把 $\mathscr{G}$ 上两个矢量 A, B 变为一个实数的对称双线性映射，因而是 $\mathscr{G}$ 上的 $(0,2)$ 型对称张量，在李代数理论中被称为 $\mathscr{G}$ 上的 Killing 型。由前述定理不难证明

$$K([A, B], C) = K(A, [B, C]),\ \forall A, B, C \in \mathscr{G}$$

提示：利用 tr 号内方阵顺序的可轮换性，例如 $\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$。

上式也可借抽象指标表为

$$K(A,B)=(\mathrm{ad}_A)^a{}_b(\mathrm{ad}_B)^b{}_a=C^a{}_{cb}C^b{}_{da}A^cB^d$$

另一方面，$K(A,B)$ 又可用抽象指标表为 $K_{cd}A^cB^d$，与上式结合便得

$$K_{cd}=C^a{}_{cb}C^b{}_{da}$$

K_{ab} 的对称性使它有可能充当矢量空间 $\mathscr{G}$ 上的度规，但为此还须 K_{ab} 非退化。可以证明，K_{ab} 非退化的充要条件是 $\mathscr{G}$ 为半单李代数。所以半单李代数 $\mathscr{G}$ 的 Killing 型 K_{ab} 可充当 $\mathscr{G}$ 的度规，称为嘉当度规。K_{ab} 的号差原则上因 $\mathscr{G}$ 而异，不过物理中遇到的李代数的 K_{ab} 多数是负定的，因而存在正交归一基底 $\{(E_\mu)^a\}$ 使 $K_{\mu\nu}=K_{ab}(E_\mu)^a(E_\nu)^b=-\delta_{\mu\nu}$。

嘉当度规 K_{ab} 的一大用处是给结构常数 $C^\sigma{}_{\mu\nu}$ 降指标，即可定义

$$C_{\rho\mu\nu}\equiv K_{\rho\sigma}C^\sigma{}_{\mu\nu}$$

定理 7.35. $C_{\rho\mu\nu}\equiv C_{[\rho\mu\nu]}$

证明. 下标本来就反称，故只须证 $C_{\rho\mu\nu}=-C_{\mu\rho\nu}$，即 $K_{\rho\sigma}C^\sigma{}_{\mu\nu}=-K_{\mu\sigma}C^\sigma{}_{\rho\nu}$。注意到 $K_{\rho\sigma}C^\sigma{}_{\mu\nu}=K(E_\rho,E_\sigma)C^\sigma{}_{\mu\nu}=K(E_\rho,C^\sigma{}_{\mu\nu}E_\sigma)=K(E_\rho,[E_\mu,E_\nu])$，只须证

$$K(E_\rho,[E_\mu,E_\nu])=-K(E_\mu,[E_\rho,E_\nu])$$

利用前述定理及 $[E_\rho,E_\mu]=-[E_\mu,E_\rho]$ 立即得证。 □

李代数 $\mathscr{G}$ 的 Killing 型是用伴随表示 ad 定义的，但可自然推广到任何表示 $\beta\colon\mathscr{G}\to\mathscr{L}(\mathscr{G})$。在 $\mathscr{G}$ 上定义对称双线性映射 $\tilde{K}\colon\mathscr{G}\times\mathscr{G}\to\mathbb{R}$ 如下：

$$\tilde{K}(A,B)\coloneqq\mathrm{tr}[\beta(A)\beta(B)],\ \forall A,B\in\mathscr{G}$$

当 β 为伴随表示 ad 时 $\tilde{K}$ 回到 K。与 K 类似，$\tilde{K}$ 也可看作 $\mathscr{G}$ 上的 $(0,2)$ 型对称张量，故也可记作 $\tilde{K}_{ab}$。当李代数 $\mathscr{G}$ 及其表示 β 满足适当条件时 $\tilde{K}$ 非退化，这时就可充当 $\mathscr{G}$ 上的度规，不妨称为广义嘉当度规。不难证明前述定理也适用于 $\tilde{K}$，即

$$\tilde{K}([A,B],C)=\tilde{K}(A,[B,C]),\ \forall A,B,C\in\mathscr{G}$$

第八章 纤维丛及其在规范场论的应用

本附录大量用到李群和李代数的知识，阅读前最好先读前一章。由于种种原因，本章（及前一章）的某些符号与本书他处的习惯不尽相同，例如，① 除少数地方外不用抽象指标，以减轻表达式的外观复杂度。② 曲线 $\eta(t)$ 的切矢用 $\dfrac{d\eta(t)}{dt}$（或 $d\eta(t)/dt$，或 $(d/dt)\eta(t)$，只在少数情况用 $\partial/\partial t$）代表。此外，本章经常涉及丛流形 P 和底流形 M，它们的点分别用 $p \in P$ 和 $x \in M$ 代表；$p \in P$ 的切空间在本书他处都记作 V_p，但在本章中 V_p 另有含义，代表 p 点切空间的竖直子空间。因此本章把 $p \in P$ 的切空间记作 T_pP。类似地，$x \in M$ 的切空间记作 T_xM。这不但符合大多数文献的习惯，而且也合理，因为 M 的切丛在文献中统一记作 TM, 点 $x \in M$ 的切空间既然对应于 x 上方的纤维，就对应于 TM 的一个子集，自然应记作 T_xM。

8.1 主纤维丛

8.1.1 主丛的定义和例子

定义 8.1. 李群 G 在流形 K 上的一个左作用是一个 C^∞ 映射 $L: G \times K \to K$ 满足：

（a） $L_g: K \to K$ 是微分同胚 $\forall g \in G$；

（b） $L_{gh} = L_g \circ L_h$，$\forall g, h \in G$

注 8.1. 与李变换群的定义比较可知李变换群 $G \times M \to M$ 就是 G 在 M 上的左作用。

定义 8.2. 李群 G 在流形 K 上的一个右作用是一个 C^∞ 映射 $R: K \times G \to K$ 满足：

（a） $R_g: K \to K$ 是微分同胚 $\forall g \in G$；

（b） $R_{gh} = R_h \circ R_g$，$\forall g, h \in G$

注 8.2. 仿照李变换群，也可以证明 L_e 及 R_e（其中 e 是 G 的恒等元）是恒等映射。

下面大量用到李群 G 在某流形（主丛流形）P 上的右作用，即 $R\colon P\times G\to P$。我们将把 $R_g(p)$（其中 $g\in G, p\in P$）简记作 pg。

定义 8.3. 子集 $\{pg \mid g\in G\}\subset P$ 称为右作用 $R\colon P\times G\to P$ 过点 $p\in P$ 的轨道。右作用 $R\colon P\times G\to P$ 称为自由的，若 $g\neq e\Rightarrow pg\neq p\ \forall p\in P$。左作用的自由性和轨道仿此定义。

现在介绍主纤维丛的定义。这一定义对初学者有一定难度，我们先给出定义，再以加注的方式详加解释。

定义 8.4. 主纤维丛由一个叫做丛流形的流形 P、一个叫做底流形的流形 M 和一个叫做结构群的李群 G 组成，满足以下要求：

（*a*） G 在 P 上有自由右作用 $R\colon P\times G\to P$;

（*b*） 存在 C^∞ 的、到上的投影映射 $\pi\colon P\to M$，满足

$$\pi^{-1}[\pi(p)]=\{pg \mid g\in G\},\ \forall p\in P$$

（*c*） 每一 $x\in M$ 有开邻域 $U\subset M$ 及微分同胚 $T_U\colon \pi^{-1}[U]\to U\times G$，且 T_U 取如下形式：

$$T_U(p)=(\pi(p),S_U(p)),\ \forall p\in\pi^{-1}[U]$$

其中映射 $S_U\colon\pi^{-1}[U]\to G$ 满足

$$S_U(pg)=S_U(p)g,\ \forall g\in G$$

注 8.3. 今后把主纤维丛简称主丛，并简记为 $P(M,G)$，甚至简记为 P。

注 8.4. 投影映射 $\pi\colon P\to M$ 一般不是一一映射，故逆映射 $\pi^{-1}\colon M\to P$ 一般不存在。但对子集 $U\subset M$ 而言，$\pi^{-1}[U]\equiv\{p\in P\mid \pi(p)\in U\}$ 有明确含义。同理，把 $x\in M$ 看作 M 的独点子集 $\{x\}\in M$，则 $\pi^{-1}[\{x\}]$ 有意义（简记作 $\pi^{-1}[x]$），称为点 $x\in M$ 上方的纤维。注意到前述定义，条件（*b*）实际上就是要求任一点 $p\in P$ 的投影 $\pi(p)\in M$ 上方的纤维 $\pi^{-1}[\pi(p)]$ 等于右作用 R 过点 p 的轨道。

注 8.5. 映射 $R\colon P\times G\to P$ 在给定 $p\in P$ 后诱导出映射 $R_p\colon G\to P$。既然 $\pi^{-1}[\pi(p)]$ 是 R 过 p 的轨道，映射 R_p 的值域就只能是 $\pi^{-1}[\pi(p)]\subset P$，故也可把 R_p 更明确地写成 $R_p\colon G\to\pi^{-1}[\pi(p)]$。令 $x\equiv\pi(p)$ 以便把 $R_p\colon G\to\pi^{-1}[\pi(p)]$ 简记为 $R_p\colon G\to\pi^{-1}[x]$。可以证明 $R_p\colon G\to P$ 是个嵌入映射，所以 G 的流形结构可被带到 $\pi^{-1}[x]$ 上，使 $\pi^{-1}[x]$ 成为 P 中的嵌入子流形，[1]而且 $R_p\colon G\to\pi^{-1}[x]$ 成为微分同胚。进一步自然要问：G 的群结构是否也可被带到 $\pi^{-1}[x]$ 上？答案是肯定的。首先，$R_p(e)=p$ 使我们想到可选 p 作为待定义李群 $\pi^{-1}[x]$ 的恒等元。其次，每一 $p'\in\pi^{-1}[x]$ 对应于 G 的一个元素 $g\equiv R_p^{-1}(p')$，故

$$p'=R_p(g)=R_g(p)=pg$$

[1]不难证明 $\pi^{-1}[x]$ 由此得到的微分（流形）结构与点 $p\in\pi^{-1}[x]$ 的选择无关。还可证明 $\pi^{-1}[x]$ 是正则嵌入子流形。

因而可借用 G 的群乘法给 $\pi^{-1}[x]$ 定义群乘法：

$$(pg)\cdot(ph) \coloneqq p(gh),\ \forall pg, ph \in \pi^{-1}[x]$$

于是每一纤维都可看作 G 的一个李群同构版本。但必须注意：由于 p 点在 $\pi^{-1}[x]$ 上可以任取（$\pi^{-1}[x]$ 中没有一点是天生与众不同的），在把 $\pi^{-1}[x]$ 定义为李群时不存在一种自然的定义方式（取任意 $p \in \pi^{-1}[x]$ 作为恒等元均可）。所以，与 G 不同，$\pi^{-1}[x]$ 不是一个自然的李群，或说它不存在天生的群结构。它与 G 之间的李群同构映射 R_p 是 p 点依赖的。

注 8.6. 条件（*c*）保证每一 $x \in M$ 都有开邻域 U，其逆像 $\pi^{-1}[U]$ 与乘积流形 $U \times G$ 微分同胚。在特殊情况下，这个 U 可能“大”到等于 M，这时 $\pi^{-1}[U] = P$，于是 P 与乘积流形 $M \times G$ 微分同胚，不妨写 $P = M \times G$。这种可表为乘积流形的主丛称为平凡主丛。一般主丛并不平凡，但条件（*c*）保证 $\pi^{-1}[U]$ 总与 $U \times G$ 微分同胚，所以说任何主丛都是局域平凡的。由于微分同胚映射 T_U 在此起关键作用，所以把 T_U 称为一个局域平凡化，简称局域平凡。

注 8.7. 条件（*c*）是对映射 T_U 的要求。$\forall p \in \pi^{-1}[U]$，映射的像 $T_U(p)$ 是 $U \times G$ 的一点，即 U 的一点与 G 的一点构成的有序对 $(\bullet, \bullet)$，由此不难理解前式右边写成 $(\pi(p), S_U(p))$ 的原因。括号中第一槽 $\pi(p)$ 表明 $T_U(p)$ 的第一要素等于 p 的投影 $\pi(p)$（$p \in \pi^{-1}[U]$ 保证 $\pi(p)$ 的确是 U 的元素），第二槽则要灵活得多，它只规定第二要素是 p 在映射 S_U 下的像，而对映射 S_U 的唯一要求是满足前述等式。

注 8.8. $\forall x \in U$，把 S_U 的定义域限制为 $\pi^{-1}[U]$ 的子集 $\pi^{-1}[x]$，便有微分同胚 $S_U : \pi^{-1}[x] \to G$（以保证 T_U 是微分同胚）。可见，选定一个局域平凡 T_U 就选定了 $\pi^{-1}[x]$ 与 G 之间的一个微分同胚映射，因而在 $\pi^{-1}[x]$ 上确定了一个“特殊点”$\check{p}_U$，满足 $S_U(\check{p}_U) = e \in G$。相应于这个 $\check{p}_U$ 又有微分同胚 $R_{\check{p}_U} : G \to \pi^{-1}[x]$，不难证明映射 $R_{\check{p}_U}$ 与 $S_U : \pi^{-1}[x] \to G$ 互逆，即

$$S_U \circ R_{\check{p}_U} : G \to G\text{是恒等映射}$$

注 8.9. 设 $P(M, G)$ 是主丛，则不难证明如下的有用公式：

$$R_g \circ R_p = R_{pg} \circ I_{g^{-1}},\ \forall p \in P, g \in G$$

其中，$I_{g^{-1}} : G \to G$ 是由 $g^{-1} \in G$ 构造的伴随同构。

例 8.1. 对任给的李群 G 和流形 M 总可按如下三步构造一个主丛：

（*1*） 选 $P \equiv M \times G$ 为丛流形，则 P 的任一点都可表为 $p = (x, g)$，其中 $x \in M, g \in G$。

（*2*） 定义自由右作用 $R : P \times G \to P$ 为

$$R_h(x, g) \coloneqq (x, gh),\ \forall x \in M,\ h, g \in G$$

（也可表为 $(x, g)h \coloneqq (x, gh)$）

(3) 定义投影映射为

$$\pi(x, g) := x, \ \forall x \in M, g \in G$$

以上三步足以保证 $P(M, G)$ 是个主丛，条件 (c) 自动满足，具体说，$\forall x \in M$ 都选 M 为条件中的开邻域 U，从而 $\pi^{-1}[U] = \pi^{-1}[M] = M \times G$，再把局域平凡 $T_U \colon \pi^{-1}[U] \to U \times G$ 定义为恒等映射便可。这样构造的主丛显然是平凡的。

设 $P(M, G)$ 是主丛，$T_U \colon \pi^{-1}[U] \to U \times G$ 和 $T_V \colon \pi^{-1}[V] \to V \times G$ 是两个局域平凡，且 $U \cap V \neq \varnothing$，则每一 $p \in \pi^{-1}[U \cap V]$ 在 G 中有两个像点，即 $g_U \equiv S_U(p)$ 和 $g_V \equiv S_V(p)$，故 $T_U(p) = (x, g_U), T_V(p) = (x, g_V)$。设 $\{U, V, \cdots\}$ 是 M 的一个开覆盖，以 Σ 代表 $U \times G, V \times G, \cdots$ 的非交并集（“非交”是指 $(x, g_U) \in U \times G$ 和 $(x, g_V) \in V \times G$ 即使在 $g_U = g_V$ 时也看作不同点），则 Σ 比 P 要“大”，因为每一 $p \in P$ 在 Σ 中对应不止一个像点。但只要把每点 $p \in P$ 在 Σ 的所有像点认同为一点，Σ 就可代表 P。由于 $g_U = S_U(p), g_V = S_V(p)$，且 g_U 和 g_V 都是群元，所以

$$g_U = g_U {g_V}^{-1} g_V = S_U(p) S_V(p)^{-1} S_V(p)$$

（其中 $S_V(p)^{-1}$ 代表群元 $S_V(p)$ 的逆元。）由此便引出如下定义。

定义 8.5. 设 $T_U \colon \pi^{-1}[U] \to U \times G$ 和 $T_V \colon \pi^{-1}[V] \to V \times G$ 是主丛 $P(M, G)$ 的两个局域平凡，$U \cap V \neq \varnothing$。映射 $g_{UV} \colon U \cap V \to G$ 称为由 T_U 到 T_V 的转换函数，若

$$g_{UV}(x) = S_U(p) S_V(p)^{-1}, \ \forall x \in U \cap V \text{（其中 } p \text{ 满足 } \pi(p) = x\text{）}$$

注 8.10. 条件 $\pi(p) = x$ 表明 $p \in \pi^{-1}[x]$。在 $\pi^{-1}[x]$ 上取另一点 p'（$\neq p$），它自然也满足 $\pi(p') = x$。如果用 p' 代替上式的 p 后竟给出不同的 $g_{UV}(x)$，则上式不能充当 g_{UV} 的合法定义。因此应验证由上式定义的 $g_{UV}(x)$ 的确与 p 的选择无关。设 p' 是 $\pi^{-1}[x]$ 的另一点，则因 $p, p' \in \pi^{-1}[x]$，必有 $g \in G$ 使 $p' = pg$，故

$$\begin{aligned} S_U(p') S_V(p')^{-1} &= S_U(pg) S_V(pg)^{-1} = S_U(p) g [S_V(p) g]^{-1} \\ &= S_U(p) g g^{-1} S_V(p)^{-1} = S_U(p) e S_V(p)^{-1} = S_U(p) S_V(p)^{-1} \end{aligned}$$

（其中第三步用到群元乘积之逆的计算法则，即 $(hg)^{-1} = g^{-1} h^{-1}$。）可见上式是 g_{UV} 的合法定义。

定理 8.1. g_{UV} 满足

(a) $g_{UU}(x) = e, \ \forall x \in U \cap V$;

(b) $g_{VU}(x) = g_{UV}(x)^{-1}, \ \forall x \in U \cap V$;

(c) $g_{UV}(x) g_{VW}(x) g_{WU}(x) = e, \ \forall x \in U \cap V \cap W$

刚才讲过，要使非并交集 Σ 等于 P，应把 Σ 中由每点 $p \in P$“裂变”出来的各点（分别属于 Σ 的各项 $U \times G, V \times G, \cdots$）认同。更准确地可以定义一个等价关系 $\sim$：设 $(x, g) \in U \times G, (x', g') \in V \times G$，当且仅当

$$x = x', \ g = g_{UV}(x)g'$$

时就说 (x, g) 等价于 (x', g')，记作 $(x, g) \sim (x', g')$。将 Σ 中所有等价的点认同，结果便是 P，记作 $P = \Sigma/\sim$。数学上的等价关系必须具备三个性质，以现在的例子陈述就是：

（a） 反身性，即 $(x, g) \sim (x, g)$；

（b） 对称性，即 $(x, g) \sim (x', g') \Leftrightarrow (x', g') \sim (x, g)$；

（c） 传递性，即 $(x, g) \sim (x', g'), (x', g') \sim (x'', g'') \Rightarrow (x, g) \sim (x'', g'')$

可以验证由上式定义的 $\sim$ 的确是等价关系。

定理 8.2. *设 g_{UV} 是从局域平凡 T_U 到 T_V 的转换函数，$x \in U \cap V$。以 $\breve{p}_U$ 和 $\breve{p}_V$ 分别代表由 T_U 和 T_V 在 $\pi^{-1}[x]$ 上确定的“特殊点”，即*

$$S_U(\breve{p}_U) = S_V(\breve{p}_V) = e \in G$$

则

$$\breve{p}_V = \breve{p}_U g_{UV}(x)$$

证明. $\breve{p}_U, \breve{p}_V \in \pi^{-1}[x]$ 保证存在 $g \in G$ 使

$$\breve{p}_V = \breve{p}_U g$$

则

$$g_{UV}(x) = S_U(\breve{p}_V)S_V(\breve{p}_V)^{-1} = S_U(\breve{p}_V)e = S_U(\breve{p}_U g) = S_U(\breve{p}_U)g = eg = g$$

代入上式便可得证。 □

定义 8.6. *设 $P(M, G)$ 是主丛，U 是 M 的开子集。C^∞ 映射 $\sigma\colon U \to P$ 称为一个局域截面，若*

$$\pi(\sigma(x)) = x, \ \forall x \in U$$

注 8.11. *上式保证每点 $x \in U$ 在 σ 映射下的像都在纤维 $\pi^{-1}[x]$ 内。*

定理 8.3. *局域截面与局域平凡之间存在自然的一一对应关系。*

证明. 给定 $T_U\colon \pi^{-1}[x] \to U \times G$ 后每一 $x \in U$ 的纤维有一个特殊点 $\breve{p}_U$，满足 $S_U(\breve{p}_U) = e \in G$。把 $\breve{p}_U$ 定义为 $\sigma(x)$ 便自然得到一个局域截面 $\sigma\colon U \to P$（光滑性是显然的）。反之，给定 $\sigma\colon U \to P$ 后，把每一 $x \in U$ 的像点 $\sigma(x)$ 作为待定义的 T_U 的特殊点便可定义 $T_U\colon \pi^{-1}[x] \to U \times G$。具体说，$\forall p \in \pi^{-1}[U]$ 令 $x \equiv \pi(p)$，则 p 与 $\sigma(x)$ 属于同一纤维 $\pi^{-1}[x]$，故存在唯一的 $g \in G$ 使 $p = \sigma(x)g$，把 $T_U(p)$ 定义为 (x, g) 便可。 □

8.2 主丛上的联络

8.3 与主丛相伴的纤维丛（伴丛）

8.4 物理场的整体规范不变性

8.5 物理场的局域规范不变性

8.6 截面的物理意义

8.7 规范势与联络

8.8 规范场强与曲率

8.9 矢丛上的联络和协变导数