

数学物理笔记

现代理论基础结构的介绍

[美] Sadri Hassani 著
罗仁卿 编

0.1 序言

给刚刚踏上追寻智慧和真理之路的初学者。愿你们不忘初心，永远秉持着一颗热衷于刨根问底的内心。

宏伟华丽的建筑总是需要有个深入地底稳固无比的地基，我们对智慧和学问的追寻亦是如此。如果我们一味的追求处在高出的上层建筑，满足于现代理论的华丽结论，和有趣的实验与运用，那么我们总会意识到自己大多的理解都是无法深入且似懂非懂的，甚至于还会有一些难以察觉的谬误。最为严重的是，我们将必然慢慢地看不见脚下的路，在黑暗中迷茫失去了方向和理想。不过对于物理学来说，我并不是极端地声张一开始就一定要去深究其根基和原理，因为以往的物理学大厦本就是从上层探索发现建立而成的，并且根基需要太多的耐心去钻研而又缺乏显而易见的趣味和成果。当然，如果你本就热衷于探究苦涩艰难的抽象原理的话，那么从根基出发坚持不懈说不定将会开出一番新天地，毕竟谁也不知道将来的物理学将会如何的修建。

Hassani 的数学物理，在我打开这本书的一开始，从它的序言和目录结构读起，就已经让我不由的期待万分和激动不已，因为这就是我一直以来所梦寐以求的大而全的教材，它能够让我完整的建立和巩固我的物理学在数学上的框架和根基。不但能够扫除我长久依赖的一些似是而非的困惑，或许还能给予我一些可遇不可求的灵感。不过，这本书像是一本没有什么废话的百科全书，使得其中的内容太密集根本就难以读下去。因此，我便从最为简单和基础的集合开始研读，对每一个知识点我都一点一点细致的剖析，思考，和举例。很多时候，我还会在原来的基础上进行一些拓展和深入，或者补充说明和更加详细的解释。自然而又慢慢地也就有了这本数学物理笔记，不单单是为了督促和勉励自己，得以坚持下来深入地思考和研读完整本书。此外，另一个目的是为了和大家一起分享我对物理数学的热爱和一些微不足道的见解，希望能够给大家追寻梦想带来一些的帮助，最好我希望还能遇到一些志同道合的爱好者一起加入到对物理数学的探究和讨论中。我想大概许多和我一样热爱物理数学但是天赋平平的朋友们面对这样一本枯燥乏味庞大全

面的书籍都会感到望而生畏吧。所以，我在记笔记的同时不仅写的能够让我自己看懂和学习，我也尽可能的写的通俗易懂，给出一些例子和插图，将一些难点和令人疑惑的地方做好充分的解释试图让所有的爱好者都能够从中更加容易的理解和掌握更多深刻的思想。

不过由于作者本身也未曾尽管专业的训练，水平十分的有限，不免会出现一些谬误或者解释不清楚的地方，恳请发现纰漏的朋友能够之处遗漏所在，我的邮箱是 1084172661@qq.com。或者如果可以我恳请所有正在使用或者阅读这本笔记的朋友能够加入 QQ 群：153815103。这样我们就能一起学习，讨论，和思考，我也十分的欢迎各种提议和批评。



群名称: Hassani 数学物理讨论组
群 号: 153815103

此外，我也由衷的期望着能够有越来越多的人开始思考和研究数学物理建筑结构这样一门纯粹艺术。毕竟，这个世界实在有着太多美丽，巧妙，神奇，和崇高之处了，以至于我们的理论作为其倒影都是那样的优雅。甚至于有时，数学的理论实在过于的完美和纯粹，以至于我差点忘记了它不过人为一厢情愿的“虚假的完美”创造。因此，无论在这笔记中我是多么的偏向于探讨数学理论，而将物理学搁置一旁，都请牢记我们身处于眼前的世界中，所探究物理学中数学基础最终目的，归根结底是为了更好的领悟大自然运行规律的玄妙与其鬼斧神工的创造。

最后，我想分享一段一直以来对我的人生有着深远影响的一段话与君共勉，出自《小逻辑》，[德] 黑格尔著，贺麟译：“不去认识真理，只去认识那表面的有时间性的偶然的東西，只去认识虚浮的东西，这种虚浮习气在哲学里已经广泛地造成，在我们的时代里更为流行，甚至还加以大吹大擂。但我要特别呼吁青年的精神，因为青春是生命中最美好的一段时间，尚没有受到迫切需要的狭隘目的系统的束缚，而且还有从事于无关自己利益的科学工作的自由。同样青年人也还没有受过虚妄性的否定精神，和一种仅只是批判劳作的无内容的哲学的沾染。一个有健全心情的青年还有勇气去追求真理。”

目录

0.1	序言	2
第一章	预备 (Preliminaries)	6
1.1	集合 (sets)	6
1.1.1	等价关系 (Equivalence Relations)	12
1.2	映射 (Maps)	16
1.2.1	序列 (Sequence)	18
1.2.2	基数 (Cardinality)	20
1.2.3	可逆映射, 商映射	21
1.3	度量空间 (Metric Space)	25
1.3.1	基础拓扑 (Topology)	30
1.4	数学归纳法 (Mathematical Induction)	31
1.5	习题	32
第二章	向量和线性映射	35
2.1	向量空间 (Vector Spaces)	35
2.1.1	向量子空间 (Subspaces)	38
2.1.2	商空间 (Quotient (or Factor) Space)	40
2.1.3	直和 (Direct Sums)	42

2.1.4	张量积 (Tensor Product)	46
2.2	内积 (Inner Product)	47
2.2.1	正交性 (Orthogonality)	50
2.2.2	格拉姆-施密特正交化 (The Gram-Schmidt Process)	51
2.2.3	施瓦茨不等式 (Schwarz Inequality)	53
2.2.4	向量的模长	54
2.3	线性映射 (Linear Maps)	55
2.3.1	线性映射的核 (kernel)	60
2.3.2	线性同构 (Linear Isomorphism)	62
2.4	复结构 (Complex Structures)	65
2.5	线性泛函 (Linear Functionals)	68
2.6	多重线性映射 (Multilinear Maps)	72
2.6.1	线性算子的行列式	79
2.6.2	经典伴随 (Classical Adjoint)	80
2.7	习题	82
第三章	代数 (Algebra)	83
3.1	从向量空间到代数	83
3.1.1	广义性质 (General Properties)	84
3.1.2	同态 (Homomorphisms)	92
3.2	理想 (Ideals)	96

第一章 预备 (Preliminaries)

1.1 集合 (sets)

集合论是近代数学的基石,但是这也并不意味着集合是一个十分抽象和深奥的概念。计算机科学中集合本质上是一种数据结构,我认为这是一个十分优雅和直观的理解。关于集合的最为基础部分的讨论在这里将会比较的紧凑,因为这个部分理应对大家来说是复习,当然如果你已经生疏或者初学,那么你还是需要多看几遍牢记一些基本概念和符号的。

定义 1.1.1 集合 (*set*) 是将事物或者对象聚集在一起的方式,满足:

1. **无序性**: 集合的每个元素之间没有次序之分同时也没有权重差异。
2. **互异性**: 集合的每个元素都是唯一的,任意两个元素都是不同的。
3. **确定性**: 元素要么属于集合要么不属于,不允许模棱两可的存在。

我们将集合 A 中的每一个成员 a 称之为该集合的元素 (*elements*)。

比如学校里所有的学生就构成了一个集合。学校里面的任意一个学生也就是该集合的元素。有时,为了表示几何涵义也将 a 称之为一个点 (*point*)。

定义 1.1.2 a 是集合 A 中的一个元素写做, $a \in A$ 。 a 不是集合 A 中的一个元素写做 $a \notin A$ 。

定义 1.1.3 集合通常是由可枚举的元素在一对花括号里面来表示。如果一个集合只有一个元素，那么我们称其为**单元素集合 (Singleton)**。

例 1.1.1 $\{3, 1, 4, 2\}$ 是由四个数组成的集合； $\{a\}$ 是单元数集合。

此外， $\{\dots, -4, -2, 0, 2, 4, \dots\}$ 这样一个集合或许你能够一眼就看出这是所有偶数的集合。注意我们之前已经明确了集合是无序的，这里为了让大家能够明白我所想要表达的特点所以是按照顺序书写的。不过这样的符号表述有着明显的缺陷，那就是无法严格和清晰的表达该集合中元素的特点。

定义 1.1.4 很多情况下我们会使用 $\{n|P\}$ 来标记和定义一个集合。这样的符号我们可以将其读作“对于该集合中的每一个元素 n 都符合该数学真命题性条件 P 。”或者“集合中的所有元素都是在条件为真的情况下成立的”。

例 1.1.2 运用我们所学过的符号我们就能够将之前的偶数集合表示写成一个更加严谨的形式， $Even \equiv \{n|n = 2a \wedge a \in \mathbb{Z}\}$ 。不过，这样的写法虽然规范和严格，但是有些的累赘。因此，大多数情况下我们会将其简化为：

$$Even \equiv \{2n|n \in \mathbb{Z}\} \quad (1.1)$$

定义 1.1.5 如果集合 A 中的所有元素都同样是集合 B 中的元素，那么我们称集合 A 是集合 B 的**子集 (Subset)**。符号上我们将其记作 $A \subseteq B$ 或者 $B \supseteq A$ 。上述定义写做符号形式为， $(\forall a \in A \implies a \in B) \iff A \subseteq B$ 。如果 A 不是 B 的子集我们就写做， $C \not\subseteq B$ 。如果 A 是 B 的子集并且 $A \neq B$ ，那么 A 是 B 的**真子集 (proper subset)**。

如果一个事物它不是其自身，则违背了基本逻辑，因此它不可能存在。

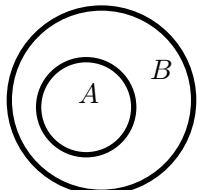
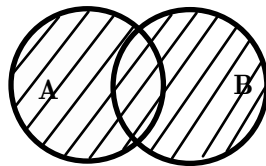
图 1.1: A 是 B 的子集

图 1.2: 并集

定义 1.1.6 空集 (empty set) 指的就是一个不存在任何元素的集合, 定义为: $\{a | a \neq a\}$ 。类似的, 我们还可以定义**零**为: 若事物 A 等价于非 A , 也就是 $A \Leftrightarrow \neg A$, 则 A 为零 (不存在)。

推论 1.1.1 空集是任意集合的子集。

这里我和你或许都会有些疑惑, 之前我们对子集的定义似乎并不支持这一推论。那么, 我们可以试着换个更加完善的说法:

定理 1.1.1 A 是 B 的**子集**当且仅当 A 中没有任何不属于 B 的元素。

定义 1.1.7 有限集合 A 中的所有子集构成的集合我们将其称之为**幂集 (Power Set)**, 记作 2^A 。

命题 1.1.1 集合 A 的子集数是 2^N , 其中 N 是 A 集合中元素的个数。

扩展 1.1.1 为什么集合 A 的子集数目是 2^N ? 每当我看到这样的问题, 我总是会回想起抛掷三枚硬币会带来 8 种组合的例子, 我们同样可以把这个思想转化到处理集合 $\{a, b, c\}$ 的所有子集。各硬币在每个事件中有正反两个状态, 正如各元素在每个集合中有存在或不存在两个状态。比如空集就是当所有不存在, 对应三个硬币全是反面。子集 $\{b, c\}$ 就是

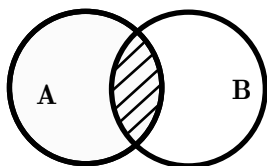


图 1.3: 交集

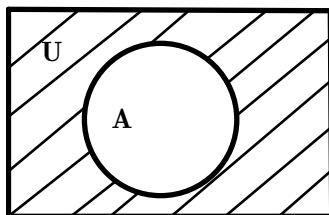


图 1.4: 补集

当 a 不存在但 b, c 存在, 对应三个硬币分别为反正正。这样我们就不难看出集合中任意元素对于任意该集合的子集都有两种状态, 那么显然总共的子集构成方式就有 2^N 个。

两个习题作为回顾中学知识的练习: 1. 如果集合有 N 个元素, 试着推导一个函数 Ω 来输出由 n 个元素构成子集的所有方式的数目 $\Omega(N, n)$ 。2. 由公式推导出 $\sum_{n=0}^N \Omega(N, n) = 2^N$ 。集合作为储存和表达数据的结构, 和数字一样自然有着一些特定的数学运算和操作。

定义 1.1.8 如果一个集合是由分别来自集合 A 或者集合 B 中的所有元素所构成的, 那么我们称其为 A 和 B 的**并集 (Union)**, 记作:

$$A \cup B \equiv \{c | c \in A \vee c \in B\}. \quad (1.2)$$

例 1.1.3 并集的例子: $\{1, 4, 7\} \cup \{e, d, c, b\} = \{e, d, 1, 4, 7, c, b\}$ 。

定义 1.1.9 如果一个集合中的任意元素属于集合 A 并且也属于集合 B , 那么我们称其为 A 和 B 的**交集 (Intersection)**, 记作:

$$A \cap B \equiv \{c | c \in A \wedge c \in B\}. \quad (1.3)$$

例 1.1.4 交集的例子: $\{e, 4, 7\} \cap \{e, d, c, b, 4\} = \{e, 4\}$

定义 1.1.10 对于三个及以上的集合的并集, 也就指的是 $A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n$, 我们将其记作 $\bigcup_{i=1}^n A_i$ 。同样的对于交集是 $\bigcap_{i=1}^n A_i$ 。

定义 1.1.11 如果一个集合中的任意元素都不属于集合 A , 那么我们称其集合 A 的**补集 (complement)**为将其记作 \bar{A} 或者:

$$\sim A \equiv \{c | c \notin A\} \quad (1.4)$$

如果我们单单就说一个集合是 A 的补集, 那么很显然它将会是一个有无穷多个在 A 集合之外的元素构成的集合, 并且这样的集合是无法给我们任何明确的意义。这就类似于我若告诉你一个事物它不是猫, 你还是不能知道我说的究竟是什么。因此, 给定集合 A 以外一个范围是很且重要的。

定义 1.1.12 我们将集合 D 作为关于 A 的补集的范围, 也就是将集合 A 在 D 上的补集定义为: $D \sim A \equiv \{c | c \in D \text{ and } c \notin A\}$ 。

例 1.1.5 你或许已经发现了这样的定义会带来两种不同的案例:

第一种: A 并非 D 的子集 $D \sim A \equiv \{e, 4, 7, a\} \sim \{e, d, c, b, 4\} = \{a, 7\}$ 。

定义 1.1.13 若 A 并非 D 的子集, 则其为 A 在 D 中的**相对补集** (A, D 的**差集 (difference)**) 记作 $D \sim A$ 或 $D - A$ 。

第二种如图 1.4 所示, A 是 D 的子集: 比如, $\mathbb{Z} \sim \mathbb{Z}^- = \mathbb{N}$:

定义 1.1.14 如果 A 是 D 的子集, 则 $D \sim A$ 称之为 A 在 D 中的**绝对补集**。 D 是**全集 (Universal Set)** 往后都将记作 U 。

命题 1.1.2 任意集合都可以是全集。

不过依据不同特定场合，全集会有着更加清晰明确的使用。比如，在实分析中实数集就是全集，复分析中复数集就是全集。

定理 1.1.2 交集，并集，补集的定理：

交换律 (Commutative property):

$$A \cup B = B \cup A, \quad A \cap B = B \cap A. \quad (1.5)$$

结合律 (Associative property):

$$(A \cup B) \cup C = A \cup (B \cup C), \quad (1.6)$$

$$(A \cap B) \cap C = A \cap (B \cap C). \quad (1.7)$$

分配律 (Distributive property):

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C), \quad (1.8)$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C). \quad (1.9)$$

德摩根定律 (De Morgan's Law):

$$A \sim (B \cup C) = (A \sim B) \cap (A \sim C), \quad (1.10)$$

$$A \sim (B \cap C) = (A \sim B) \cup (A \sim C). \quad (1.11)$$

集合中元素的顺序是不重要的。然而，在很多情况下元素的顺序是不得不被考虑的。例如，二维平面的坐标 (x, y) 代表了二维平面上的任意一点，其中 $(1, 2) \neq (2, 1)$ 。这就引出了下面的定义：

定义 1.1.15 有序对 (Ordered Pairs)的概念, 我们将其记作 (a, b) 。如果 a, b 分别属于集合 A, B 中的元素, 那么一个由 a, b 的有序对组成的集合就是 A, B 的笛卡儿积, 记作: $A \times B = \{(a, b) | a \in A \wedge b \in B\}$ 。对于三个及以上的集合的笛卡儿积, 我们将其**广义化定义**记作: $A_1 \times A_2 \cdots \times A_n = \{(a_1, \cdots, a_n) | a_i \in A_i\}$ 。若集合 $A_1 = A_2 \cdots = A_n$, 我们可以将 $A_1 \times A_2 \cdots \times A_n$ 简写为 A^n 。

比如一副扑克牌的花色和大小的笛卡儿积就构成了 52 张牌。

例 1.1.6 关于广义化的例子就是当集合 $A = \mathbb{R}$, A^n 就是 n 维笛卡尔坐标中全部有序对 (x_1, \cdots, x_n) 的 n **元组 (n-tuples)** 实数集合 \mathbb{R}^n 。

1.1.1 等价关系 (Equivalence Relations)

在许多情况下, 集合中的元素之间往往都有着某些特定的联系, 因而它们都能够自然地组合或划分成该集合的多个子集。为了描述两个及以上或同一集合上的元素之间的关系, 我们将会定义一种叫做关系的数学结构。

定义 1.1.16 若 A, B 为两个非空集合, 从 A 到 B 的**关系 (Relations)**指的就是对 $A \times B$ 中的任意一个有序对 (a, b) 中元素 a 与 b 的一个**条件比较判断 (comparison test)**, 如果 a 与 b 满足条件并通过比较判断, 那么 a 与 b **有关系**, 记作 $a \triangleright b$ (作为一个存放着所有符合条件的有序对的一个 $A \times B$ 的子集)。否则 a 与 b 没有关系。任意 A^n 的子集都可以被当作集合 A 上的 n **元关系**。

例 1.1.7 设条件为 $(a > b) \wedge (a + b > 4)$, $\{1, 3, 7\} \times \{1, 2, 4\} = \{(1, 2), (3, 2), (7, 1), (3, 1), (1, 4), (7, 4), (3, 4), (1, 1), (7, 2)\}$ 其中 $(7, 1)$ 符合条件, $(7 > 1) \wedge (7 + 1 > 4)$ 。同理有, $\{1, 3, 7\} \triangleright \{1, 2, 4\} = \{(7, 4), (3, 2), (7, 2), (7, 1)\}$

定义 1.1.17 如果 \triangleright 是集合 A 上的一个二元关系且满足以下三个性质,

- **自反性 (Reflexively):** $a \triangleright a, \quad \forall a \in A$
- **对称性 (Symmetry):** $a \triangleright b \implies b \triangleright a, \quad \forall a, b \in A$
- **传递性 (Transitivity):** $a \triangleright b \wedge b \triangleright c \implies a \triangleright c, \quad \forall a, b, c \in A$

那么我们就认为它是一个在集合 A 上的**等价关系**。

等价关系是一类特殊的二元关系, 它是拥有一个等价条件判断的二元关系, 因此它也是 $A \times A$ 的子集。另外关于定义中所使用的符号 $a \triangleright b$ 我们称之为“ a (在某种条件或者设定下) 等价于 b ”。等价关系是一个十分重要的数学构造, 因为它是广义上的对两个事物按照其属性带来的等价性的规范。比如, 我们最为熟悉的实数之间的等价也是符合广义等价关系的三个性质的。此外, 三角形的全等和相似也是等价关系。

定义 1.1.18 如果 \triangleright 是集合 A 上的等价关系并且 $a \in A$, 对于任意的 $b \in A$ 都有 $b \triangleright a$, 也就是集合中所有元素都等价于 a , 那么我们称在该条件下构成的集合为元素 a 的**等价类 (Equivalence Classes)**, 记作

$$[a] = \{b \in A | b \triangleright a\} \quad (1.12)$$

写代码的时候有时我们需要把一个大的循环拆分成由多个步骤组成的小周期来处理。毕竟我们常常会希望我们 n 个线程同时处理整个大循环, 因此我们自然的就会将大循环按照整数索引的余数 n 划分为 n 个小任务让线程分别处理。这样就自然产生了等价类。

例 1.1.8 设集合 $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$, 该集合上的关系条件为 $a \equiv b \pmod{3}$, 这样的条件使其成为等价关系。注意 $a b$ 是该集合自身的笛卡儿积带来的有序对 (a, b) 。该集合上的等价关系就是:

$A \triangleright A = \{(0, 0), (2, 8), (1, 4), (0, 6), (0, 3), (2, 2), (2, 5), (1, 1), (1, 7)\}$ 。其中有三个等价类分别为 $[0] = \{0, 3, 6\}, [1] = \{1, 4, 7\}, [2] = \{2, 5, 8\}$ 。

定理 1.1.3 如果 \triangleright 是在集合 A 上的一个等价关系并且有 $a, b \in A$, 那么要么有 $[a] \cap [b] = \emptyset$ 或者是 $[a] = [b]$ 。

这个定理其实挺显然的, 毕竟逻辑上一个事物不能同时等价于两个及以上不相同的其他事物, 它本质上是因为等价关系的传递性。如果你还不能理解那就再回过头来看看上面的例子。由该定理可得知, 如果 $b \in [a]$, 那么元素 b 绝对不会出现在不等于 $[a]$ 的其他任何等价类中。因此, 一个等价类中的任意元素都是该等价类所独属的, 并且我们将等价类的任何元素作为该类的**代表 (representative)**。由于等价关系的对称性, 我们将其记作 \bowtie 。

等价关系的概念是自然的存在于物理学中的, 因为决定同一物理性质的数学对象应该被认为是等价的。比如, 单位长度的**量子态函数**与幅角不同的单位长度复数的乘积, 可以组合在一起, 因为它们都代表相同的物理状态。此外, 我们知道**磁场 (Magnetic Field)**是**螺线矢量场 (solenoid vector field)**因为其**散度 (divergence)**为零, $\nabla \cdot \mathbf{B} = 0$ 。磁场的**向量势 (vector potential)** \mathbf{A} 的定义为, $\mathbf{B} = \nabla \times \mathbf{A}$ 。那么, 我们就有 $\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0$ 。这就说明该向量势在任意**组成方向 (components)**上都有连续二阶偏导数。这里我们也就很自然的可以联想到, 如果让向量势加上一个有连续二阶偏导数的标量函数 f , 那么它并不会改变原来的磁场。

所以, 螺线向量场也就是磁场所对应的向量势不是唯一的, 这就表明不同的向量势可以计算出同一个磁场。也就是说, 设 $\mathbf{A} = \mathbf{A}' + \nabla f$, 因为 $\nabla \times \nabla f = 0$, 所以 $\mathbf{B} = \nabla \times \mathbf{A}' + \nabla \times \nabla f = \mathbf{B}'$ 。那么, 我们就很自然的联想到了之前讨论过的等价关系和等价类。其中, $\mathbf{A} = \mathbf{A}' + \nabla f$ 就是关于 $\mathbf{A} \bowtie \mathbf{A}'$ 的一个等价条件判断。它们之间的等价关系是建立在彼此都可以计算出同一磁场这一条件或者设定下成立的, 也就是说等价类 $[\mathbf{A}]$ 是一个包含了所有能够计算出同一磁场的不同向量势。

定义 1.1.19 如果集合 A 的子集所构成的一个集合 $\{A_1, A_2, \dots, A_n\}$ 满足以下三个条件, **交集为空 (disjoint)**: $A_i \cap A_j = \emptyset, \forall i \neq j$, **不含空集**: $A_i \neq \emptyset, \forall i$, **并集为 A** : $\bigcup_{i=1}^n A_i = A$, 我们就将其称之为集合 A 的**划分 (partition)** P 。集合 P 中的每个子集我们称之为**类 (class)**。

划分的定义是十分显然和易懂的, 划分集合就像划分蛋糕一样。你应该已经看出来了划分和等价类有着直接显然的联系, 那就是对于一个集合 A 上的任意等价关系都可以确定一种划分方式, 其所有等价类构成的集合是集合 A 上的一个划分, 每一个等价类都是集合 A 的类。反之亦然, 给定一个集合上的划分, 我们也可以归纳出该集合上的等价关系, 因为集合中两个元素等价仅当其存在于同一个划分的区块中。

定义 1.1.20 集合 A 的所有等价类所组成的集合 $\{[a] | a \in A\}$ 称为 A 在等价关系上的**商集 (Quotient Set)**, 记作 A/\sim 。

例 1.1.9 回顾例 1.3.1, 现在我们考虑整数集 \mathbb{Z} 上的**同余关系 (Congruence Relation)** $a \triangleright b \iff k|a-b$ 。这里我只不过是换了个表达 (因为这样的表达能够更容易的推导出同余关系是等价关系), 意思是 $a-b$ 可以被 k 整除, 也可以记作 $a \equiv b \pmod{k}$ 。当我们从模 3 推广到了模 4 之后, 我们可以很轻易的看出整数集 \mathbb{Z} 在同余关系上的**商集** 为, $\mathbb{Z}/\sim = \{[0], [1], \dots, [k-1]\}$ 。其中 $[0] = \{\dots, -3k, -2k, -k, 0, k, 2k, 3k, \dots\}$, $[k-1] = \{\dots - 2k-1, -k-1, -1, k-1, 2k-1, 3k-1, \dots\}$ 。

定义 1.1.21 设集合 \mathbb{R}^n 为是实数集 \mathbb{R} 的**广义笛卡儿积**。那么, \mathbb{R}^n 就是由 n **元组有序对** 构成的 n 维空间。若点 $X = (x_1, x_2, \dots, x_n), Y = (y_1, y_2, \dots, y_n)$ 是在原点出发的同一条射线, 则它们等价。也就是如果 $x_i = ky_i, \forall i$ 那么 $X \propto Y$ 为**等价关系**。这里面的一个**等价类**就是一条射

线，一个射影空间 (Projective Space) \mathcal{P}^{n-1} 的定义就是 \mathbb{R}^n 在该等价关系上的商集， \mathbb{R}^n / \sim 。

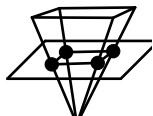


图 1.5: 射影平面

这也类似于绘画中的透视技巧，透视需要先选择透视点 (可以在纸外)，然后每对本应当平行的线条相交于透视点。这样一张画纸就相当于被三维空间中的一个点上的所有射线所穿过。这些射线所构成的集合就是一个 2 维的射影空间。因此， n 维度的射影空间应当对应 \mathbb{R}^{n+1} 。

1.2 映射 (Maps)

之前我们讨论了集合间元素的关系和分类，接下来我们将要深入的探讨在集合之间的联系。我们所熟悉的函数也可以称作映射就是这样的联系。所谓映射就是一个输入某个事物然后经过“加工”后输出“成品”的“工厂”。我们当然不会满足于这样类比，所以接下来我们将对其给以明确的定义。

定义 1.2.1 如果非空集合 X 中的任意一个元素 x 能够按照特定的法则 f 唯一地对应或者联系到集合 Y 上的一个元素 y ，那么我们就称其为**从集合 X 到集合 Y 的一个映射 f (Map)**，记作 $f: X \rightarrow Y$ 或 $X \xrightarrow{f} Y$ 。对于集合内的元素，我们将元素 y 记作 $f(x)$ ，来表示元素 x 经过映射 f 对应到了的元素 y ，也就是 $x \xrightarrow{f} y$ 。数学上**函数 (Function)**一般指的是一个陪域为实数集 \mathbb{R} 或复数集 \mathbb{C} 的映射。

定义 1.2.2 恒等映射 (Identity map)或者恒等函数是一个特殊的映射 $id_A(a): A \rightarrow A$ ，我们将其定义为， $id_A(a) = a, \forall a \in A$ ，简单来说，它就是一个输入和输出是完全相等的映射。

如图 1.6 所示, 我们将输出的 $f(x)$ 称作输入的 x 在映射 f 之下的**像 (image)**。因此, 根据之前映射的定义, $x \in X$ 有且只有一个像。如果映射 $f: X \rightarrow Y$ 和映射 $g: X \rightarrow Y$ 的**像**是一样的, $f(x) = g(x), \forall x \in X$, 那么这两个映射是**相等的**。所有像所构成的集合我们称作 f 的**值域 (range)**, 记作 $f(X)$ 。此外, 我们将集合 X 称作**定义域 (Domain)**, 将集合 Y 称为**陪域 (Co-domain)**或者**目标域 (target space)**。

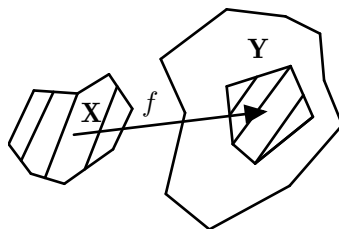


图 1.6: 映射 f 将 X 对应到 Y 中的子集, Y 中的阴影是 f 的值域 $f(X)$ 。

关于函数和映射我们还需要明确定义中的概念。1. 定义域中**所有的元素**都必须映射到陪域上的某个元素。2. 定义域中的任意一个元素都必须有**唯一**的映射到陪域上。也就是说任意一个 x **必须有对应并且只能对应一个 $f(x)$** , 不过一个 $f(x)$ **不必须要有对应且可以对应多个 x** 。

定义 1.2.3 映射 $f: A \rightarrow B$ 的 Γ_f **图像 (Graph)** 是 $A \times B$ 的子集, 定义为 $\Gamma_f = \{(a, f(a)) | a \in A\} \subset A \times B$ 。

定义 1.2.4 确实有些让人迷惑, 不过我们思考后应该就能够很快的发现它不过就是定义了我们初中就学过的函数图像。我们先从最为熟悉的实数函数来考虑, 然后很容易地就可以注意到 $(a, f(a))$ 说的不过就是函数线条上的任意一个点, 那么这些点所构成的集合 Γ_f 自然就是函数图像了。那么, 我们为什么要强调他是 $A \times B$ 的子集呢? 对于实数函数来说, $A \times B$ 无非就是整个笛卡尔坐标系的 xy 平面。我们的函数图像要在 xy 平面上或者范围里自然也是必须且重要的。

例 1.2.1 集合 C 存放了三个不同的硬币 $\{a, b, c\}$, 假设一个函数 f 会将其对应到有正反两种状态构成的集合 $S = \{0, 1\}$ 上, 那么就会有 $2^3 = 8$ 种可能的映射和其图像。它们分别是:

$$\begin{aligned}\Gamma_f 1 &= \{(a, 0), (b, 0), (c, 0)\}, \Gamma_f 2 = \{(a, 1), (b, 0), (c, 0)\}, \\ \Gamma_f 3 &= \{(a, 0), (b, 1), (c, 0)\}, \Gamma_f 4 = \{(a, 0), (b, 0), (c, 1)\}, \\ \Gamma_f 5 &= \{(a, 1), (b, 1), (c, 0)\}, \Gamma_f 6 = \{(a, 1), (b, 0), (c, 1)\}, \\ \Gamma_f 7 &= \{(a, 0), (b, 1), (c, 1)\}, \Gamma_f 8 = \{(a, 1), (b, 1), (c, 1)\}\end{aligned}$$

将大小为 N 的集合映射到大小为 n 的集合上会带来 n^N 种映射。

1.2.1 序列 (Sequence)

接下来的内容是一个十分重要的综合性例子, 我们将会探讨如何运用集合和映射的思想, 有艺术性地创造和定义一个新的结构。

定义 1.2.4 如果 $\mathbb{Z}_0^+ = \{0, 1, 2, \dots, a\}$ 并且 Y 是任意集合, 那么任何映射 $l: \mathbb{Z}_0^+ \rightarrow Y$ 都叫做 Y 上的**序列**。我们将序列定义域中的任意一个元素 $n \in \mathbb{Z}_0^+$ 叫做**索引**, 序列 l 的一个**像**, $l(n) = u_n \in Y$ 称作序列的**项**, 这样序列 l 就记作 $[u_n]$ 。

既然提到了序列, 这里我们不妨就来给它个定义。**序列 (Sequence)**通俗的讲就是将一堆事物按照顺序且可以重复的放在一起。用数学的话来说, 一堆事物就指的是**集合**, 把一堆事物按照规则放在一起指的就是**映射**, **序列**就是把一个集合有规则映射到另一个集合上从而体现**顺序和可重复性**。虽然说集合中的元素有无序性, 但是集合本身可以是建立在一个有序的框架条件上的, 有限大小为 a 的**非负整数集** $\mathbb{Z}_0^+ = \{0, 1, 2, \dots, a\}$ 就是一个天然就有顺序的集合。我们的直觉可能会是要把一个集合当作输入映射到正整数集上**排号**输出。然而这是**错误**的, 只要稍加思考我们就会发现如果要满足**待排序集合**上的事物具有**可重复性**, 那么根据映射性质它就不能作为定义域输入, 因为定义域上的元素只能有一个对应。因此, **正确**的做法是将正整数集

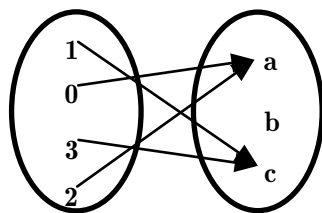


图 1.7: 序列 $[a, c, a, c]$

作为定义域映射到一个集合上, 由此达到为该集合排序的目的。具体的映射作为序列的例子如图 1.7 所示, 序列 $f: \mathbb{Z}_0^{+a} \rightarrow a, b, c$ 将 0 对应到 a 以此类推, 排列在一起构成序列 $[a, c, a, c]$ 。

定理 1.2.1 如果有映射 $f: X \rightarrow Y$ 和映射 $g: Y \rightarrow Z$, 那么 f 与 g 的**复合 (composition)** $h: X \rightarrow Z$ 也是一个映射, 记作 $h = g \circ f$ 。此外, 复合映射 h 的像为 $h(x) = g(f(x))$ 。

例 1.2.2 映射 $f: \{1, 2\} \rightarrow \{3, 4\}$ 和 $g: \{3, 4\} \rightarrow \{5, 6\}$ 的复合函数为 $h: \{1, 2\} \rightarrow \{5, 6\}$ 。如果 $\Gamma_f = \{(1, 3), (2, 4)\}$, $\Gamma_g = \{(3, 5), (4, 6)\}$, 那么 $\Gamma_h = \{(1, 5), (2, 6)\}$ 。

定义 1.2.5 如果有映射 $f: X \rightarrow Y$ 并且 A 是集合 X 的子集, 那么 $f(A) = \{f(x) | x \in A\}$ 就是 A 的**像**。反过来, 如果 $B \subseteq Y$, 那么 $f^{-1}(B) = \{x \in X | f(x) \in B\}$ 就是 B 的**原像 (Preimage)**。如果 B 由单个元素 b 组成的, 则 $f^{-1}(B) = \{x \in X | f(x) = b\}$ 是由 X 中的所有映射到 B 的元素组成的一个**纤维 (Fiber)**。

由于定义域中的多个元素可以映射到一个陪域中的元素上, 同时不一定所有陪域上的元素都要有对应。因此, 一个纤维不一定是单元素集合而是可以由**零个及以上**的元素构成的集合。根据定义, 定义域子集可以**映射**到一个陪域子集 (**像**) 上, 然而如果推广一下我们之前所讨论的内容, 就会发现陪域子集对应到定义域子集 (**原像**) 上的联系**不一定**是满足映射条件的。因而若是想要了解在怎样的条件下陪域也可以**映射**到定义域上的话, 我们就需要先明确定义和讨论映射的方式和规则。

定义 1.2.6 如果 $x_1, x_2 \in X$ 并且每当 $f(x_1) = f(x_2)$ 都有 $x_1 = x_2$, 那么就意味着任意 $f(x) \in Y$ 最多只有一个原像 $f^{-1}(f(x)) \in X$, 我们将这样的映射 f 称作**单射 (injective or one-to-one)**, 记作 $1-1$ 。如果映

射 f 满足 $f(X) = Y$, 那么 $f(x) \in Y$ 至少有一个原像 $f^{-1}(f(x)) \in X$, 我们将这样的映射 f 称作**满射** (*onto or surjective*)。如果 f 既是单射又是满射, 那么任意的像 $f(x) \in Y$ 一定**有且只有一个**原像, 我们称其为**双射** (*bijective*)。

在我看来数学上的定义大致分为**规则性**定义和**结构性**定义。一般来说, 我都会在能力范围内尽可能的长篇大论地深入探讨**结构性**定义, 就像之前探讨该如何定义序列那样, 因为其中往往可能会有着很多**深刻的**思想和图景。对于**规则性**定义来说, 定义本身并不会对我们的理解起到很大的帮助, 更多的时候我们对其的理解是建立在对具体例子思考和**归纳**后得出的。

1.2.2 基数 (Cardinality)

当你尝试思考了不同实例后, 或者说你天赋异禀一开始就从定义显然的知道了下面的规律。单射的定义域元素数量一定**小于等于**陪域, 满射的定义域元素数量一定**大于等于**陪域, 双射的定义域元素数量一定**等于**陪域。如果定义域和陪域元素数量**相等**, 那么无论是**单射**还是**满射**其效果彼此相同且都和**双射**一样。看来集合的**元素数量**还是个挺重要的玩意儿, 我们不妨给它起个名字-**基数** (cardinal)。集合 $\mathbb{N}_n = \{1, 2, \dots, n\}$ 是**有限的** (finite)并且基数为 n , 任意集合 A 是有限的如果 A 和 \mathbb{N} 是等价的 $A \bowtie \mathbb{N}_n$ 。也就是集合 A 和 \mathbb{N} 之间存在**双射**, 这里集合 A 的基数也是 n 。如果 A 不是有限的, 那么它就是**无限的** (infinite)。一个集合 (所有自然数的集合) 可以与其子集 (例如所有偶数的集合) 双射, 这是任意无限集共有的特性。如果 $A \bowtie \mathbb{N}$, 那么 A 是**可数的** (countably infinite)。如果集合 A 既不是有限的, 也不是可数的, 那么它是**不可数的** (uncountable)。

在**闭区间** $[0, 1]$ 中间挖去一个**开区间** $(\frac{1}{3}, \frac{2}{3})$, 也就像是把一根线段平均分成三份, 拿走中间的一份, 剩下的就是 $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, 继续这样的操作有 $[0, \frac{1}{9}] \cap [\frac{2}{9}, \frac{1}{9}] \cap [\frac{2}{9}, \frac{7}{9}] \cap [\frac{8}{9}, 1]$, 让其无限地一直挖下去, 剩下的部分就是**康托尔集** (Cantor set), 康托尔集的基数和原先的 $[0, 1]$ 是一样的。

1.2.3 可逆映射, 商映射

定理 1.2.2 结合之前我们讨论过的**复合函数**, 如果 $f \circ g$ 是单射, 则 g 一定是单射而 f **不必**是单射。如果 $f \circ g$ 为满射, 则 g **不必**是满射而 f **一定**是满射。如果 $f \circ g$ 是双射, 则 g **不必**是满射而 f **不必**是单射。

回到我们之前所讨论的原像上, 如果是映射是**单射**的话, 那么**陪域**上可能存在没有与**定义域**对应的元素, 因而像对应到原像**不满足**映射的条件。如果是**满射**, 那么尽管**陪域**上所有元素都能够有对应, 然而其不一定是**唯一**的对应, 因此还是**不满足**映射性质。如果映射 f 为**双射**, 那么其**陪域**和**定义域**的元素数量根据定义是**相等**的并且对于任意一个 $y \in Y$ 都有且只有一个 $x \in X$ 使得 $f(x) = y$ 。这样一来就满足了映射的性质, 我们就有 $f^{-1}: Y \rightarrow X$ 满足 $f^{-1}(y) = x$ 。

定义 1.2.7 映射 f 是**可逆的** (*invertible*) 并且 f^{-1} 是 f 的**反函数** (*inverse function*), 如果 $f(x) = y \iff f^{-1}(y) = x, \forall y \in Y, x \in X$ 。或者说, f 可逆当且仅当有 $f^{-1}(f(x)) = x$ 并且 $f(f^{-1}(y)) = y, \forall y \in Y, x \in X$ 也就是 $f^{-1} \circ f = id_Y \wedge f \circ f^{-1} = id_X$ 。

例 1.2.3 $f(x) = y = e^x$ 是可逆的其反函数为 $f^{-1}(y) = x = \ln y$, 因为 $f^{-1}(f(x)) = \ln(e^x) = x$ 并且 $f(f^{-1}(y)) = e^{\ln y} = y$ 。

例 1.2.4 $f(x) = y = 2x + 3$ 是可逆的其反函数为 $f^{-1}(y) = x = \frac{y-3}{2}$ 或者 $f^{-1}(x) = \frac{x-3}{2}$ 。设映射 $f: X \rightarrow Y$ 图像的子集为: $\{(3, 0), (5, 1), (9, 3), (7, 2)\}$, 那么其逆映射 $f^{-1}: Y \rightarrow X$ 对应的图像子集为 $\{(2, 7), (1, 5), (0, 3), (3, 9)\}$ 。

我们不难发现映射图像和逆映射图像集合 Γ_f 和 $\Gamma_{f^{-1}}$ 上的每个有序对是有一**对一关系** (双射) 的且有**对称性**, 设 $a \triangleright b \equiv (a, b)$, 那么就有 $(x, y) \implies (y, x)$ 。有序对的对称只是一个比较直观的展示, 真正的对称是

发生在映射本身上的 $f: X \rightarrow Y \implies f^{-1}: Y \rightarrow X$ 。那么, 我们可以定义一个求逆的映射 $p: \Gamma_f \rightarrow \Gamma_{f^{-1}}$ 。我们规定的求逆映射正是利用了映射与逆映射图像之间的对称性, 从而交换有序对中 x, y 的位置得出的。在有序对内位置被对调的时候, 满足条件 $x = y$ 的有序对是**前后不会发生改变的**也就是图像集合 $\{(x, y) | x = y, x \in X, y \in Y\}$ 。因而 $f(x)$ 和 $f^{-1}(x)$ 是关于 $f(x) = x$ 镜像对称的。

例 1.2.5 利用对称性**展示**复合函数 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$:

复合映射 $g \circ f = X \xrightarrow{f} Y \xrightarrow{g} Z$ 的**逆**为 $f^{-1} \circ g^{-1} = Z \xrightarrow{g^{-1}} Y \xrightarrow{f^{-1}} X$ 。

定理 1.2.3 给定映射 $f: X \rightarrow Y$, 集合 X 上有**等价关系** $x_1 \bowtie x_2 \implies f(x_1) = f(x_2)$ 。**等价类**是 X 的子集, 其所有元素映射到 Y 中的同一元素。若是 X, Y 有**相同基数**且 f 为单射或者满射, 则有**等价关系** $X \bowtie Y$ 。

证明可考虑等价关系性质: 自反, 对称, 传递。关于 $x_1 \bowtie x_2$ 的**等价类**也就是说, 一个**满射的陪域**中任意元素的**原像**都可以被认为是一个等价类 $[x] = f^{-1}(f(x))$, 因而我们可以得到以该**等价类**组成的原定义域的**商集**。

定义 1.2.8 等价类是 X 的子集, 其所有元素映射到 Y 中的同一个点。也就是说, 有个与之对应的映射 $\tilde{f}: X / \bowtie \rightarrow Y$, 或者 $\tilde{f}([x]) = f(x)$, 我们将其称之为**商映射 (Quotient Map or Factor Map)**。

比如, 设函数为 $f(x) = x^2$ 就有 $\tilde{f}(\{\pm\sqrt{x} | x \in X\}) = f(x)$ 。

定理 1.2.4 映射 $\tilde{f}: X / \bowtie \rightarrow f(X)$ 是**双射**。

证明: 商映射是单射, 因为 $\tilde{f}([x_1]) = \tilde{f}([x_2]) \implies f(x_1) = f(x_2)$, 所以 x_1 和 x_2 属于同一等价类 $[x_1] = [x_2]$ 。映射 f 为**满射**当且仅当 $f(X) = Y$, 并且 f 满射意味着每一个 $y \in Y$ 都有与之对应的**等价类**。如果 f **不是满射**, 则任意的 $y \in f(X)$ 都有对应的**等价类**使得 \tilde{f} 既是**满射**也是**单射**。♣

接下来我们将会讨论几个较为复杂的综合性例子作为映射部分的总结和巩固。第一个例子需要我们对**区间 (Intervals)**有最为基本的认识。

定义 1.2.9 设 $a, b \in \mathbb{R}$, **开 (Open) 区间**: $]a, b[= \{x \in \mathbb{R} | a < x < b\}$ 。

闭 (Close) 区间: $[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$ 。另外, $[a, b[= \{x \in \mathbb{R} | a \leq x < b\}$, $]a, b] = \{x \in \mathbb{R} | a < x \leq b\}$ 。

映射 $\sin : \mathbb{R} \rightarrow [-1, 1] \subset \mathbb{R}$ 和 $\cos : \mathbb{R} \rightarrow [-1, 1] \subset \mathbb{R}$ 有关于陪域元素 0 的原像为: $\sin^{-1}(0) = \{n\pi\}_{n=-\infty}^{\infty}$, $\cos^{-1}(0) = \{\frac{\pi}{2} + n\pi\}_{n=-\infty}^{\infty}$ 。类似的, **闭区间** $[0, \frac{1}{2}] \subset \mathbb{R}$ 的原像, 包含所有 x 轴上的粗线的**并集**。如图 1.8 所示, x 轴上的所展示的粗线为函数 \sin 或 \cos 的**原像** $\sin^{-1}[0, \frac{1}{2}]$ 或者 $\cos^{-1}[0, \frac{1}{2}]$ 的子集。函数图像上的粗线为 \sin 或者 \cos 的**像**。

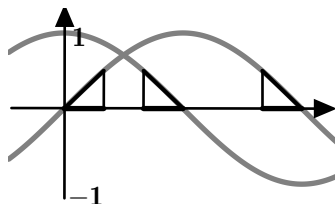


图 1.8: 三角函数

定义 1.2.10 一个从 X 到 Y 的 n 元映射 (n -ary map) 为映射 $f : X^n \rightarrow Y$ 有 $f(x_1, \dots, x_n) = y$, 这就是我们所熟悉的多元函数。

射影映射 (projection map) 为 $\text{Proj}_i : X^n \rightarrow X_i$ 给定 $\mathbf{x} = (x_1, \dots, x_n) \in X^n$, $\text{Proj}_i(\mathbf{x}) = x_i$ 。

定义 1.2.11 给定任意集合 X 上有等价关系 \bowtie , 那么根据定义自然 (*natural or canonical map*) 就有一个称作**正则射影 (canonical projection)**的映射 $p : X \rightarrow X/\bowtie$, 将一个集合 X 映射到该集合所以等价类组成的商集 X/\bowtie 上。对于任意的 $x \in X$ 也就是 $p(x) = [x]$ 。

因为不同的 x 可以映射到同一等价类上 $p(x_1) = p(x_2) \implies x_1 \bowtie x_2$, 所以**射影**显然是满射非单射的。如果该等价关系为恒等映射 $\bowtie = \text{id}_X$, 那么该映射就变为双射, 记作 $X \cong X/\text{id}_X$ 。

例 1.2.6 单射和满射依赖于定义域和值域。映射 $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$ 和 $g: \mathbb{R} \rightarrow]-1, 1[, g(x) = \tanh x$ 都是**双射**。接下来, 我们就映射 $f(x) = x^2$ 来考虑在不同定义域和值域下的情况。

1. **满射非单射** $f: \mathbb{R} \rightarrow \mathbb{R}^+, \pm x$ 都二对一映射到全体正实数上。
2. **单射非满射** $f: \mathbb{R}^+ \rightarrow \mathbb{R}, +x$ 一对一映射到实数的正数部分 x^2 上。
3. **双射** $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+, +x$ 都一对一映射到全体正实数上。
4. **非单射非满射** $f: \mathbb{R} \rightarrow \mathbb{R}, \pm x$ 二对一映射到实数的正数部分 x^2 上。

例 1.2.7 设 $M^{n \times n}$ 表示 $n \times n$ 实矩阵的集合。定义一个求**行列式** (*determinant*) 的函数 $\det: M^{n \times n} \rightarrow \mathbb{R}$ 。这个映射是满射而不是单射, 因为不同矩阵可以有同一个行列式且行列式的取值范围为整个实数集。

例 1.2.8 映射 $f: \mathbb{C} \rightarrow \mathbb{R}, f(z) = |z|$ 既不是单射也不是满射。原像 $f^{-1}(1)$ 为复平面上的**单位圆** (*unit circle*)。

不同复数的模长可以相同因此不单射。复数的模一定是零和正实数的并集 $f(\mathbb{C}) = \{0\} \cup \mathbb{R}^+$ 故不满射。该映射可以引申出一个 \mathbb{C} 上的**等价关系**: 如果 z_1, z_2 的模长相等在同一个圆上, 那么就有 $z_1 \bowtie z_2$ 。这样我们就有商集 \mathbb{C}/\bowtie 为复平面上圆心为原点的所有圆的集合。 $\tilde{f}: \mathbb{C}/\bowtie \rightarrow \{0\} \cup \mathbb{R}^+$ 是将每个圆对应到其半径的一个双射。正如我们之前在 n 元映射的定义中所见, 映射的定义域可以是集合的笛卡尔积。和二元关系一样 n 元映射中的**二元映射** $f: X \times X \rightarrow Y$ 也是尤为的重要的。

例 1.2.9 点乘就是将任意两个向量对应到实数集的二元映射 $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, 设向量 $\mathbf{a}, \mathbf{b} = (a_1, \dots, a_n), (b_1, \dots, b_n)$, 给定 $f(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$ 。

定义 1.2.12 对于二元映射 $f: X \times X \rightarrow Y$, 如果有 $Y = X$, 我们称这样的映射为**二元运算** (*binary operation*)。从集合 X 中任意选取两

个元素映射到一个仍然属于 X 的元素上的运算是**封闭的** (*closed*)。

例 1.2.10 实数上的四则运算的加、减、乘、除都是二元运算。**叉乘**就是将任意两个向量对应到新向量的二元运算 $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ，设向量 $\mathbf{a}, \mathbf{b} = (a_1, \dots, a_n), (b_1, \dots, b_n)$ ，给定 $f(\mathbf{a}, \mathbf{b}) = \mathbf{a} \times \mathbf{b}$ 。

1.3 度量空间 (Metric Space)

尽管集合论是现代数学的根基，但它本身还是过于的抽象和形式化。因此，我们将会基于集合介绍更多的数学结构让它变得更为实用。数学上通常会以两种方式来实现这样的构造，对应于两类数学的主要分支—**代数** (algebra)和**分析** (analysis)。我们可以通过在一个集合上引入二元运算把它变成一个**代数结构**。例如，**向量加法**是封闭的二元运算。一个**群** (group)是由一种集合以及一个二元运算所组成的代数结构。使用集合的概念对分析学进行抽象时，还能引导出**开集与闭集**之类的**拓扑**性质，其中**连续性的**概念起着关键作用。我们不会很的深入这两个领域，但我们还是会按照循序直观地介绍一些我们所需要的最为基本的概念，那我们先从介绍**度量空间**开始。

定义 1.3.1 集合 X 是一个**度量空间**，若任意两点 x, y 经过二元映射 $d: X \times X \rightarrow \mathbb{R}$ 能够对应到 x, y 之间的距离 $d(x, y)$ 上，且满足：

- 1. $d(x, y) \geq 0 \forall x, y$ **非负性**：两点间距离永不为负
- 2. $d(x, y) = 0 \iff x = y$ **唯一性**：两点相等距离为零
- 3. $d(x, y) = d(y, x)$ **对称性**：两点互换距离不变
- 4. $d(x, y) \leq d(x, z) + d(z, y)$ **三角不等式**

看到这样的定义很多人或许都会先入为主的带入笛卡尔坐标系或者欧几里得几何上的点。这或许是抽象的程度和深度不足的体现，因此我们在阅读定义时，更多的是要思考和分析那些从对象中**抽象**出来的**性质**。我们需要将目光暂时从对象上移开，也就是说，我们需要时刻的清楚集合 X **不需要**

任何的结构, 任意随便的什么集合都可以是集合 X 。因此, 度量空间的关键是距离函数对于任意两个元素满足的那四个性质。

定理 1.3.1 如果在距离函数 d 下, X 是一个度量空间, 那么任意一个 X 的非空子集 S 也是函数 d 下的一个度量空间。

很显然, 因为任意 X 中的两个元素在距离函数 d 下都满足四个性质的话, 那么其部分元素构成的子集一样满足。

例 1.3.1 各种度量空间的例子:

1. 我们所最为熟知也是很重要的一个度量空间就是**欧氏空间 (Euclidean space)** \mathbb{R}^n 。设非负实数距离函数为二元映射 $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{0\}$, 设向量 $\mathbf{a}, \mathbf{b} = (a_1, \dots, a_n), (b_1, \dots, b_n)$, $d(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\| = \sqrt{(a_1 - b_1)^2, \dots, (a_n - b_n)^2}$ 。此外, **复平面**也可以被看做一个二维的欧氏空间, 尽管复数不能比较大小, 复数之间是有距离的。因而我们可以将实数上的分析学思想搬到复数上。

2. **有理数集**为实数集的子集, 那么易证明其广义笛卡儿积也满足 $\mathbb{Q}^n \subset \mathbb{R}^n$, 因此在同一距离函数下, \mathbb{Q}^n 也是度量空间。

3. 设 X 为包含球面上所有点的集合。我们可以定义两个不同的距离函数。 $d_1(P, Q)$ 为连接球面上点 P, Q 的**弦长 (chord length)**, $d_2(P, Q)$ 为穿过点 P, Q 最大的一个圆上以 P 为起点 Q 为终点的**弧长**。对于 d_2 如果三个点有两个是**极点** N, S , 那么三角不等式就变为等式 $d(N, S) = d(N, P) + d(P, S)$, 因为极点的弦穿过球心, 因而穿过极点的任意一个圆都为球面最大的圆, 并且能够覆盖整个球面的所有点, 使得三个点总会在同一弧上。

4. 让 $C^0[a, b]$ 为闭区间 $[a, b]$ 上的连续实值函数 $[a, b] \rightarrow \mathbb{R}$ 的集合, $f, g \in C^0(a, b)$ 。我们可以定义距离函数 $d(f, g) = \int_a^b |f(x) - g(x)| dx$ 。这个距离描述的就是函数 f, g 的图像在区间 $[a, b]$ 上所包围起来的面积。

5. 让 $C_B(a, b)$ 为闭区间 $[a, b]$ 上的**有界 (bounded)**连续实值函数

$[a, b] \rightarrow \mathbb{R}$ 的集合, $f, g \in C_B(a, b)$ 。距离函数 $d(f, g) = \max_{a \leq x \leq b} \{|f(x) - g(x)|\}$ 。这个距离描述的就是函数 f, g 的图像在区间 $[a, b]$ 上在同一 x 上最长的一条垂线距离。

根据度量空间的定义, 两点间距离为正实数也是欧几里得度量空间的性质。不过, 相对论的闵可夫斯基度量空间 (Minkowskian metric space) 中距离的平方可以是负实数。由此看见, 若要判定闵可夫斯基空间为度量空间的话, 用闵可夫斯基度量 (Minkowski metric) $s^2 = -t^2 + x^2 + y^2 + z^2$ 是行不通的, 也就是说闵可夫斯基度量不是该空间上的距离函数。

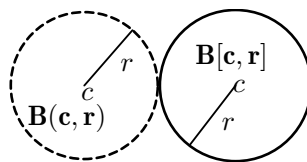


图 1.9: 开球, 闭球

根据度量的性质, $B(c, 0) = \emptyset$; $B[c, 0] = \{c\}$ 。开球和闭球的定义会让我们不由自主的将其带入到欧氏空间, 因此值得强调开球和闭球可以在任意度量空间。如图 1.9 所示, 我们也可以更为直观地将开球集合比作一个用虚线围成的圆盘, 将闭球集合比作用实线围成的圆盘。

定义 1.3.2 若 X 为度量空间有 $c, x \in X, r \geq 0$, 那么一个以 c 为球心 r 为半径的**开球 (Open ball)**为 $B(c, r) = \{x | d(c, x) < r\}$, 也就是由所有到球心距离小于半径的点构成的集合。一个以 c 为球心 r 为半径的**闭球 (Closed ball)**为 $B[c, r] = \{x | d(c, x) \leq r\}$, 也就是由所有到球心距离小于等于半径的点构成的集合。

定义 1.3.3 一个度量空间 X 上的**序列**是从自然数集到度量空间上的映射 $x: \mathbb{N} \rightarrow X$ 。该序列的像属于度量空间, $x(n) = x_n \in X$ 。我们用 $\{x_n\}_{n=1}^{\infty}$ 来表示对序列所有项的枚举。

我们在之前已经强调了序列的有序性。距离函数的规定使得在度量空间中点与点之间自然有了“**紧密度 (closeness)**”的概念。因而, 我们可以将序列引入到度量空间上去, 为其增加元素间紧密度的新特性。

定义 1.3.4 给定某个 $x \in X$, 若对于任意正实数 $\epsilon > 0$, 都存在自然数 N , 使得每当 $n > N$ 时, 都有一个围绕点 x 的开球 $B(x_n, \epsilon)$, 也就是 $d(x_n, x) < \epsilon$, 则序列 $\{x_n\}_{n=1}^{\infty}$ **收敛 (converges)** 于 x 并记作 $\lim_{n \rightarrow \infty} d(x_n, x) = 0$ 或者 $x_n \rightarrow x$ 。反之我们则称其为**发散 (divergence)**。**柯西序列 (cauchy sequence)**满足 $\lim_{n, m \rightarrow \infty} d(x_n, x_m) = 0$

早在古希腊芝诺的乌龟就已经对趋近作为不断追赶动态过程这一思想有了大致的雏形, 在数学上我们为了追求严格和精准的描述, 巧妙地运用了不断选取 $n > N$ 让 ϵ 趋近于无穷小。

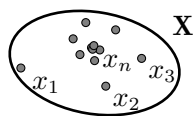


图 1.10: 柯西序列

我们有时可能无法直接用该定义检验序列的敛散性, 因为这需要给定我们一个 x 作为**极限点 (Limit point)**也是**聚集点 (cluster point)**。不过如图 1.12 所示, 我们还是可以用接下来的方法考察序列中点与点间的距离不断**缩小**, 随着 n 的取值不断**增大**。

定义 1.3.5 一个**完备度量空间 (complete metric space)**是所有柯西序列收敛的度量空间。

注意: 柯西序列不一定是收敛的。♠ 例如, 设度量空间为有理数集 \mathbb{Q} 有度量 $d(x, y) = |x - y|$, 考虑序列 $\{x_n\}_{n=1}^{\infty}$ 有 $x_n = \sum_{k=1}^n (-1)^{k+1}/k$, 这是一个柯西序列有 $|x_m - x_n| \rightarrow 0$ (证明留做习题)。然而, $\lim_{n \rightarrow \infty} x_n = \ln 2$ 不是**有理数**, 不属于该度量空间。

由此可见, 有理数空间**不是**一个完备度量空间。不过, 如果我们将所有柯西序列的 (无理数) **极限点**都加进 \mathbb{Q} 中, 那么这个空间就变成**完备**的了, 当然也变成了 \mathbb{R} 。因而, **所有的不完备空间都可以被扩充为完备空间**。

定义 1.3.6 映射 $f: X \rightarrow Y$ 在点 $x_0 \in X$ 上**连续**, 如果序列 $\{x_n\}$ 收敛到 x_0 , 那么序列 $\{f(x_n)\}$ 也会收敛 $f(x_0)$ 。

连续性的概念是分析学中极为重要的, 你或许曾经见过 $\epsilon - \delta$ 语言版本的连续性定义, 不过本质上其思想和我们之前定义收敛的都是一样的, 不断选取 $n > N$ 让 ϵ 趋近于无穷小都有 $d(x_n, x_0) < \epsilon$ 的同时设置一个 $\delta > 0$ 跟着一起趋近无穷小都有 $d(f(x_n), f(x_0)) < \delta$ 。

定义 1.3.7 度量空间 X 的子集是**有界的 (bounded)**, 如果它是被包含在任意一个有限大小半径 r 的开球 $B(c, r)$ 之中。如果该子集不是有界的, 我们称其为**无界的 (unbounded)**。

所有度量空间里的收敛序列或者柯西序列都是**有界的**。我们用收敛序列的证明作为例子, **证明:** 如果 $\{x_n\}_{n=1}^{\infty}$ 收敛, 那么

$$\exists N_1 \in \mathbb{N}, \forall n > N_1, d(x_n, x) < 1, \quad (1.13)$$

$$r = \max\{d(x_1, x), \dots, d(x_{N_1}, x), 1\} \quad (1.14)$$

$$\forall n \geq 1, d(x_n, x) < r. \quad (\text{Q.E.D}) \quad (1.15)$$

定义 1.3.8 度量空间 (X, d) 的子集 U 是**开集 (open sets)**如果对于子集 U 中任意的点 x , 存在大于零的实数 ϵ 使得对于 X 中任意的点 y 只要满足 $d(x, y) < \epsilon$ 那么 y 也在 U 中。也就是, U 是开集如果 $\forall x \in U, \exists \epsilon > 0 \in \mathbb{R}$ 使得 $\forall y \in X, d(x, y) < \epsilon \implies y \in U$ 。

用开球来定义开集就是, 如果 U 中任意的点 x 为球心, $\epsilon > 0$ 为半径的开球 $B(x, \epsilon)$ 都被包含在 U 内, 则 U 是开集。我们同样可以用开集来定义**极限点** x_0 : 如果任意包含 x_0 的**开集**至少包含一个不是 x_0 的点 x_n , 根据定义就有 $d(x_n, x_0) < \epsilon$ 使得 x_0 为极限点。**连续性**, 设 $f(x_0) \in V, x_0 \in U$ 并且 U, V 都是开集, 那么若是 $f(U) \subset V$ 则有 f 在点 x_0 上连续。

1.3.1 基础拓扑 (Topology)

定义 1.3.9 拓扑 (Topology) $\tau = \{A_i\}$ 是非空集合 A 的子集的集合:

- 1. \emptyset, A 都属于拓扑 τ 。
- 2. τ 的任意子集的**并集**也在 τ 中, $\cup A_i \in \tau$ 。
- 3. τ 的**有限**多个子集的**交集**也在 τ 中, $\cap A_i \in \tau$ 。

有序对 (A, τ) 称作一个**拓扑空间 (topological space)**。

开集 U 是 τ 的成员, **闭集** $\sim U$ 是开集 U 的**补集**。因而空集和 A 也是开集 (闭集), 任意开集 (闭集) 的**并集**以及有限开集 (闭集) 的**交集**也还是开集 (闭集)。因此, A, \emptyset 既是开集也是闭集。点 x 是拓扑空间 A 的子集 X 的元素, 如果有开集 U 使得 $U \subset X, x \in U$, 则其为 x 是**邻域** X 的**内点**, X 的**内部 (interior)** X° 是其中所有内点的集合, 所以 X° 也是开集, x 的邻域 X 是 A 的子集, 包含一个开集 A_i 并且 $x \in A_i$ 。

定义 1.3.10 映射 $f: X \rightarrow Y$ 的 X, Y 为拓扑空间, 若 $\forall x \in X$ 都有 $f(x) \in Y$ 的邻域 V 使得 $f^{-1}(V)$ 为 x 的领域, 则 f 为**连续映射**。

若双射 $f: X \rightarrow Y$ 及其逆 f^{-1} 都为连续, 则 f 为**同胚映射 (homeomorphism)**, X 和 Y **同胚 (homeomorphic)** 记作 $X \cong Y$ 。

集合 U 被所有闭集包含, 那么 U 的**闭包 (Closure)** \bar{U} 是所有闭集的**交集**。闭包 \bar{U} 也是由 U 所有内点和极限点所组成的一个集合。类似的, **内部** U° 是包含 U 的所有开集的**并集**。**边界 (Boundary)** ∂U 是将闭包的内部刨开 (差集), 也就是 $\bar{U} \sim U^\circ$ 。集合 U 的**覆盖** P 的成员是集合 Q_i 满足, $U \subseteq \cup Q_i$ 。 S 的**子覆盖 (subcover)** V 是 S 的覆盖 P 的子集并且覆盖 S 。**开覆盖 (Open cover)** C 的成员是开集 A_i 并且 U 是所有成员的并集的子集, $U \subset \cup A_i$ 。 U 是**紧致的 (compact)** 如果所有的该集合的开覆盖都有**有限**的子覆盖。也就是, 尽管有一个无限开集构成的集合覆盖整个集合, 我们仍旧能够选取一个有限的开集覆盖整个集合。

1.4 数学归纳法 (Mathematical Induction)

定义 1.4.1 数学归纳法: 假设每个自然数 n 都有一个命题 S_n , 那么对于任意的 n 都有 S_n 成立, 只要满足: 1. 命题 S_1 成立 (真); 2. 给定任意自然数 m , 如果 S_m 成立, 则 S_{m+1} 也成立。

例 1.4.1 用数学归纳法证明二项式定理 (Binomial theorem):

$$\begin{aligned}(a+b)^m &= \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k = \sum_{k=0}^m \frac{m!}{k!(m-k)!} a^{m-k} b^k \quad (1.16) \\ &= a^m + ma^{m-1}b + \frac{m(m-1)}{2!} a^{m-2}b^2 + \cdots + mab^{m-1} + b^m\end{aligned}$$

当 S_1 时, $(a+b)^1 = a^1 + b^1$ 成立。现在, 我们需要证明在 S_{m+1} 时同样成立。也就是要**证明**, $(a+b)^{m+1} = \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m+1-k} b^k$ 。

$$(a+b)^{m+1} = (a+b)^m(a+b) = \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k (a+b) \quad (1.17)$$

$$= \sum_{k=0}^m \binom{m}{k} a^{m-k+1} b^k + \sum_{k=0}^m \binom{m}{k} a^{m-k} b^{k+1} \quad (1.18)$$

$$= a^{m+1} + \sum_{k=0}^m \binom{m}{k} + \binom{m}{k-1} a^{m-k+1} b^k + b^{m+1}$$

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1} \quad (\text{Pascal's rule}) \quad (1.19)$$

$$\begin{aligned}&= a^{m+1} + \sum_{k=0}^m \binom{m+1}{k} a^{m-k+1} b^k + b^{m+1} \\ &= \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m+1-k} b^k \quad (Q.E.D) \quad (1.20)\end{aligned}$$

1.5 习题

习题 1.1: 详细证明一个基数为 n 的集合的所有子集的数量为 2^n 。

习题 1.2: 设 A, B, C 为一个全集 U 的子集, 证明:

$$A \subset B \wedge B \subset C \implies A \subset C. \quad (1.21)$$

$$A \subset B \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B. \quad (1.22)$$

$$A \subset B \wedge B \subset C \implies (A \cup B) \subset C. \quad (1.23)$$

$$A \cup B = (A \sim B) \cup (A \cap B) \cup (B \sim A) \quad (1.24)$$

习题 1.3: 假设对于每一个 $n \in \mathbb{N}$, 都有:

$$I_n = \{x : |x - 1| < n \wedge |x + 1| > 1/n\}. \quad (1.25)$$

现在试着找到 $\cup_n I_n$ 和 $\cap_n I_n$ 。

习题 1.4: 证明下述关于等价类和代表的命题。

命题 1.5.1 如果 $a' \in [a]$, 那么有 $[a'] = [a]$ 成立。

习题 1.5: 在平面, 三维空间, 和四维空间中的向量集合上定义一个“乘法”的二元运算上的向量。在每种情况下, 将一个乘积的分量用其两个向量的分量表示。

习题 1.6: 详细证明

命题 1.5.2 当 f 和 g 都是双射时有: $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ 。

习题 1.7: 考虑序列 $\{x_n\}_{n=1}^{\infty}$ 有 $x_n = \sum_{k=1}^n (-1)^{k+1}/k$, 证明这是一个柯西序列有 $|x_m - x_n| \rightarrow 0$ 。在不失一般性的前提下, 假设 $n > m$ 和 $n-m$ 是偶数 (奇数 $n-m$ 的情况也可以用类似的方法处理)。

(a) 证明:

$$x_n - x_m = (-1)^m \sum_{j=1}^{n-m} \frac{(-1)^j}{j+m}. \quad (1.26)$$

(b) 分离求和的偶数和奇数部分并证明:

$$x_n - x_m = (-1)^m \left\{ \sum_{k=1}^{(n-m)/2} \frac{1}{2k+m} - \sum_{k=1}^{(n-m)/2} \frac{1}{2k-1+m} \right\}. \quad (1.27)$$

(c) 合并两个部分并证明:

$$x_n - x_m = -(-1)^m \sum_{k=1}^{(n-m)/2} \frac{1}{(2k+m)(2k+m-1)}. \quad (1.28)$$

(d) 证明:

$$|x_n - x_m| \leq \frac{1}{(1+m)^2} + \sum_{k=2}^{(n-m)/2} \frac{1}{(2k+m-1)^2}. \quad (1.29)$$

(e) $\int_1^s f(x)dx \leq \sum_{k=2}^s f(k)$ 对任何连续函数 f 都成立, 带入 (d) 小问:

$$|x_n - x_m| \leq \frac{1}{(1+m)^2} + \int_1^{(n-m)/2} \frac{1}{(2k+m-1)^2}. \quad (1.30)$$

$$\lim_{n,m \rightarrow \infty} |x_n - x_m| = \lim_{n,m \rightarrow \infty} \frac{1}{(1+m)^2} + \frac{1}{n} - \frac{1}{2} \left(\frac{1}{n-1} - \frac{1}{m+1} \right) = 0. \quad (1.31)$$

习题 1.8: 试着找到从自然数对应到整数的一个双射 $f: \mathbb{N} \rightarrow \mathbb{Z}$.

习题 1.9: 取任意两个开区间 (a,b) 和 (c,d) , 证明无论区间的范围大小是多少, 第一个区间中的点与第二个区间中的点一样多. **提示:** 找出两个区

间点之间的 (线性) 代数关系。

习题 1.10: 用数学归纳法推出**莱布尼兹乘积法则** (Leibniz rule):

$$\frac{d^n}{dx^n}(f \cdot g) = \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \frac{d^{n-k} g}{dx^{n-k}} \quad (1.32)$$

习题 1.11: 用数学归纳法推出下列两个求和式子:

$$\sum_{k=0}^n r^k = \frac{r^{n+1} - 1}{r - 1}, \quad \sum_{k=0}^n k = \frac{n(n+1)}{2} \quad (1.33)$$

扩展 1.5.1 定义 1.5.1 归纳定义 (inductive definitions) 是用数学归纳法的思想去定义一些包含**整数**的递归。

例 1.5.1 用归纳法定义幂 (powers):

1. 从命题 S_0 开始: $a^0 = 1$ 。
2. 给出 S_m 的规则: $a^m = a^{m-1}a, \forall m > 0$ 。

阶乘 (factorial):

1. 从命题 S_0 开始: $0! = 1$ 。
2. 给出 S_m 的规则: $m! = (m-1)! \cdot m, \forall m > 0$ 。

习题 1.12: **二叉树 (binary tree)** 是每个**节点 (node)** 最多只有两个**分支 (branch)** 的**树状结构 (tree structure)**。二叉树的分支具有左右次序。试着用归纳思想递归地定义正整数为节点的**二叉树** (**提示:** 从空树-只有一个节点没有分支的树开始)

第二章 向量和线性映射

N -维笛卡尔向量 (Cartesian vectors) 有 N 个分量 (components)。笛卡尔向量有两个限制：1. 它们的分量一定都是实数。2. 维度是有限的。在物理学上有时我们需要打破实数或者有限的限制。因而，这里我们将试着抛开维度和实数的一些固有观点。

2.1 向量空间 (Vector Spaces)

狄拉克符号 (Dirac notation) 使用左矢 (bra), $\langle v|$ 和右矢 (ket), $|v\rangle$ 来表示向量。这在复向量空间上十分的实用，左矢是右矢的厄米共轭 (Hermitian adjoint) 将在第四章讨论。

定义 2.1.1 一个复数域 \mathbb{C} 之上的向量空间 \mathcal{V} 是由向量 (vectors) $|a\rangle, |b\rangle, |x\rangle$, 等构成的集合，有以下性质：

1. 任意两对向量之和属于向量空间, $|a\rangle + |b\rangle \in \mathcal{V}, \forall |a\rangle, |b\rangle \in \mathcal{V}$, 有:
 - (a) 交换律 (commutativity): $|a\rangle + |b\rangle = |b\rangle + |a\rangle$.
 - (b) 结合律 (associativity): $|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle$.
 - (c) 存在唯一零向量 $|0\rangle \in \mathcal{V}$ 使得 $|a\rangle + |0\rangle = |0\rangle, \forall |a\rangle \in \mathcal{V}$.
 - (d) 对于任意向量 $|a\rangle \in \mathcal{V}$ 都唯一存在另一个向量 $|-a\rangle \in \mathcal{V}$ 使其和为零向量, $|a\rangle + |-a\rangle = |0\rangle$.

2. 任意的复数 α 称作**标量 (scalar)**并且对于任意 $|a\rangle$ 对应的 $\alpha|a\rangle \in \mathcal{V}$, 满足: (a) $\alpha(\beta|a\rangle) = (\alpha\beta)|a\rangle$, (b) $1|a\rangle = |a\rangle$ 。

3. 涉及向量和标量的乘法是服从分配律的:

$$(a) \quad \alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle.$$

$$(b) \quad (\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle.$$

复数域上的向量空间叫做**复向量空间 (complex vector space)**, 实数域上的叫做**实向量空间 (real vector space)**, 它们之间向量空间定义的结构是完全一样的。向量空间也可以被称作**线性空间**, 线性所描述的是其性质。

实数和复数的数学结构都是**域 (field)**。一个域 \mathbb{F} 是一个集合, 在这个集合中可以对集合的元素 (**标量**) 进行**二元运算** (加减乘除), 其运算法则和实数是一样的。其中**加法恒等 (additive identity)**为 0, **乘法恒等 (multiplicative identity)**为 1。对于任意的 $\alpha \in \mathbb{F}$ 有**加法的逆 (additive inverse)**为 $-\alpha$, 并且有乘法的逆 α^{-1} 为 α 的乘法恒等。由此可见, **向量空间一定是建立在域之上的**, 域中标量乘向量仍在向量空间中。

定义 2.1.2 向量 $|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$, 是**线性不相关 (linearly independent)**如果对于 $\alpha_i \in \mathbb{C}$ 都有 $\sum_{i=1}^n \alpha_i |a_i\rangle = 0$ 成立仅当所有的 $\alpha_i = 0$, 反之则**线性相关**。 $\{|a_i\rangle\}_{i=1}^n$ 的**线性组合 (linear combination)**为 $\sum_{i=1}^n \alpha_i |a_i\rangle$ 。

线性组合乘积的**标量**一定是属于或者说对应在向量空间上的**域**上的。

例 2.1.1 向量空间的例子:

1. \mathbb{R}^n 是在实数 (作为标量) 的域 \mathbb{R} 上的一个向量空间; 向量空间 \mathbb{R} 上的任意多个向量之间都是线性相关的, 因为它们都在一条线上。

2. \mathbb{C} 可以是实数的域或者复数的域上的向量空间, 因为实数 (或复数) 乘以复数仍是复数。当在实数的域时, 该向量空间类似 \mathbb{R}^2 。

3. \mathbb{R} 不是在复数的域 \mathbb{C} 之上的向量空间, 因为其违背了定义中的第二条性质: 一个复数乘以一个实数不一定属于实数。

4. 还记的初中时期向量的合成与分解那个平行四边形吗? 其实我们在哪里就已经初步的认识向量空间了。平面 (或 n 维) 上的“箭头”的集合在平行四边形法则下, 构成实数域 \mathbb{R} 上的向量空间。

5. $\mathcal{P}^c[t]$ 是所有复数作为系数, t 作为变量组成的多项式的集合。设零向量为零多项式, 则 $\mathcal{P}^c[t]$ 是在复数域下满足多项式相加和复数相乘仍为多项式的向量空间。而实数作为系数的多项式 $\mathcal{P}^r[t]$ 是实数域上 (但不是在复数域上) 的向量空间。

6. \mathbb{C}^n 是包含所有复数 n 元组有序对的集合, 如 $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 让 γ 为一个复数, 那么我们定义:

$$|a\rangle + |b\rangle = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n), \quad (2.1)$$

$$\gamma|a\rangle = (\gamma\alpha_1, \gamma\alpha_2, \dots, \gamma\alpha_n), \quad (2.2)$$

$$|0\rangle = (0, 0, \dots, 0), \quad (2.3)$$

$$-|a\rangle = (-\alpha_1, -\alpha_2, \dots, -\alpha_n). \quad (2.4)$$

这样一来 \mathbb{C}^n 就是在复数域上的一个向量空间, 我们称其为 **n-维复数空间**。类似的, 我们所最为熟悉的 \mathbb{R}^n 是实数域上的向量空间称作 **n-维实数空间**。不过, 我们要适应和抛开维度和分量等熟悉的对向量狭义认知。

7. 所有 m 行 n 列的复数矩阵构成的集合 $M^{m \times n}$ 是一个复数域上的向量空间, 设零向量为所有位置为零的 $m \times n$ 的矩阵。

8. \mathbb{C}^∞ 为所有复数序列 $|a\rangle = \{\alpha_i\}_{i=1}^\infty, \sum_{i=1}^\infty |\alpha_i|^2 < \infty$ 的集合是复数域上的一个向量空间。所有在实数开区间 $]a, b[$ 的一元复值函数构成的集合是复数域上的一个向量空间。集合 $\mathcal{C}^n(a, b)$ 是所有在开区间 $]a, b[$ 上 n 阶导数有连续的一元实值函数, **光滑 (smooth) 函数** 为 $\mathcal{C}^\infty(a, b)$ 。该集合为实数域上的一个向量空间。可见, **线性性质和标量的域**是同样重要的。

2.1.1 向量子空间 (Subspaces)

向量空间 V 的子集 W 不一定是一个向量空间，因为其中的向量的线性组合不一定属于 W ，因而我们需要新的关于向量子空间的定义。

定义 2.1.3 向量空间 V 的向量子空间 W 是一个服从**封闭性** (W 中的任意向量的**线性组合**任旧属于 W) 的一个关于 V 的非空子集，也就是 $\forall |a\rangle, |b\rangle \in W \implies \alpha|a\rangle + \beta|b\rangle \in W, \forall \alpha, \beta \in \mathbb{C}$ 。

向量子空间是向量空间，两个向量子空间的交集也是一个向量子空间。因为交集中任意两个向量也同样属于这两个向量子空间，那么两个子空间也同时有这两个向量的线性组合，因而作为其交集也有，使得交集满足作为向量子空间的条件。

例 2.1.2 向量子空间的例子：

1. \mathbb{R} 是向量空间 \mathbb{C} 在实数域上的向量子空间，但 \mathbb{C} 若是在复数域上的则 \mathbb{R} 不是其子空间。一条穿过原点 (包含零向量) 的直线上的所有向量的集合是在平面或者空间在实数域上的向量的子空间。如果维度 $n < m$ 则 \mathbb{R}^n 是 \mathbb{R}^m 在 \mathbb{R} 上的子空间，同样的 \mathbb{C}^n 是 \mathbb{C}^m 在 \mathbb{C} 上的子空间。

2. $\mathcal{P}_n^c[t]$ 是小于等于 n 次幂的复数多项式，它是 $\mathcal{P}^c[t]$ 的子空间。若 $n < m$ 则 $\mathcal{P}_n^c[t]$ 是 $\mathcal{P}_m^c[t]$ 的子空间， $\mathcal{P}_n^r[t]$ 是 $\mathcal{P}_m^r[t]$ 的子空间。

3. 若 $r \leq m, s \leq n$ 则 $M^{r \times s}$ 是 $M^{m \times n}$ 的子空间。这里我们用最后 $m - r$ 行和 $n - s$ 列都为零的 $m \times n$ 矩阵来定义一个 $r \times s$ 矩阵。

4. x 轴上的一个向量与 y 轴上的一个向量的线性组合**不一定**在 x 或者 y 轴之上，因此**两个子空间的并集不是一个子空间**。

定理 2.1.1 如果 S 是向量空间 V 中的任意非空向量集，则 S 中所有向量线性组合的集合 W_S 是 V 的子空间。我们称 W_S 为 S 的**线性生成 (linear span)**记作 $\text{Span}\{S\}$ 。

m 个在 n 维实数空间 \mathbb{R}^n 的**线性无关**的向量, 其中 $m \leq n$, 可以生成一个 \mathbb{R}^m 的子空间。比如, 喷墨打印机的墨盒在 x 轴移动, 纸张在 y 轴移动就能打印一整个纸张平面。

定义 2.1.4 设 $B = \{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$ 为域 \mathbb{F} 上向量空间 \mathcal{V} 的有限子集, 如果 B 是**线性无关**的并且**生成**整个 \mathcal{V} (任意的 $|v\rangle \in \mathcal{V}$ 都能被 B 的**线性组合**所**唯一**合成), 则 B 为 \mathcal{V} 的**基 (basis)**, $|e_i\rangle$ 称作**基向量**。

换句话说, 若是我们能够将平面上的任意向量分解为线性无关的分量 (x 和 y), 那么我们也可以利用 x 和 y 分量 (单位向量) 合成出整个平面。

定义 2.1.5 设域 \mathbb{F} 上的 n 维向量空间 \mathcal{V} 的基 $B = \{|e_i\rangle\}_{i=1}^n$, 对于任意的 $|v\rangle \in \mathcal{V}$ 都唯一的存在一个 n **元组** $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{F}^n$, 使得 $|v\rangle = \sum_{i=1}^n \alpha_i |e_i\rangle$, 则该 n **元组** 被称为 $|v\rangle$ 在基 B 下的**分量 (components)**。

不难看出, 平面上 x, y 的单位向量的集合 $B \equiv \{(1, 0), (0, 1)\}$ 是描述向量空间的实用基本工具。 B 被称为线性空间 \mathbb{R}^2 的**基**其成员称作**基向量**。

定理 2.1.2 任意向量空间中的元素, 都可以唯一地表示成**基向量**的线性组合。如果基向量有限且有个数为 n 的基向量, 则称向量空间为**有限 n 维向量空间**, 反之则称为**无限维数向量空间**。

我们将域 \mathbb{F} 上向量空间 \mathcal{V} 的**基 B 的基数 (cardinality)**称为 \mathcal{V} 的**维数**, 用 $\dim \mathcal{V}$ 表示。为了强调它的标量域, 还可记作 $\dim_{\mathbb{F}} \mathcal{V}$ 。维数为 N 的向量空间 \mathcal{V} 有时用 \mathcal{V}_N 来表示。**注意:** 给定向量空间 \mathcal{V} 的一个**基 $B = \{|e_i\rangle\}_{i=1}^n$** , B 中任意 m 个基向量 ($m < n$) 的**生成空间**是 \mathcal{V} 的 m 维子空间。

例 2.1.3 基的例子:

1. 数字 1 (或者任意**非零实数**) 都是 \mathbb{R} 的一个**基**。数字 1 和 $i = \sqrt{-1}$ (或者任意实部和虚部都是**非零实数**的复数) 都是 \mathbb{C} 在域 \mathbb{R} 上一个的基。数字

1(或者任意**非零复数**)都是 \mathbb{C} 在域 \mathbb{C} 的一个**基**。

2. \mathbb{R}^n 在数域 \mathbb{R} 上的**标准基** (standard basis)为 $B = \{\hat{e}_j\}_{j=1}^n$, 其中 \hat{e}_j 是一个在位置 j 上为 1 其余位置为 0 的 n -元组 (**单位向量** (unit vector))。 \mathbb{C}^n 在数域 \mathbb{C} 上的**标准基** (standard basis)同样是 $B = \{\hat{e}_j\}_{j=1}^n$, 不过 \mathbb{C}^n 在数域 \mathbb{R} 上的**标准基** (standard basis)就变成了 $B = \{\hat{e}_j\}_{j=1}^n \cup \{\hat{e}_k\}_{k=1}^n$, 其中 \hat{e}_k 是一个在位置 k 上为 i 其余位置为 0 的 n -元组, 因此 \mathbb{C}^n 在 \mathbb{R} 上是 $2n$ 维度的空间; \mathbb{C}^n 在 \mathbb{C} 上是 n 维度空间。 $M^{m \times n}$ 的基为 $B = \{E_{ij}\}_{i,j=1}^{m,n}$, 其中 E_{ij} 在第 i 行第 j 列为 1, 其余位置为零的 $m \times n$ 矩阵, 因此我们可以用 $\sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} E_{ij}$ 来表示任意一个矩阵。

3. 向量间 $\mathcal{P}^c[t]$ 的基 $B = \{t^i\}_{i=0}^\infty$ 由单项式 $1, t, t^2, \dots$ 组成, 这是一个无限维度空间。 $\mathcal{P}_n^c[t]$ 的基为 $B = \{t^i\}_{i=0}^n$, 这是一个 $n+1$ 维度空间。任何一个 n 阶多项式都可以唯一地用一个 $n+1$ 元组来表示, 使得 $p[t] = \sum_{i=0}^n \alpha_i t^i$ 。根据**泰勒定理** (Taylor's theorem), 设 $a < 0 < b$, 则**光滑函数** $C^\infty(a, b)$ 的基为 $B = \{t^i\}_{i=0}^\infty$, 因为 $C^\infty(a, b)$ 中任意的函数都可以在零点 $x = 0$ 处被展开为一个**无穷幂级数** (infinite power series) $\sum_{i=0}^\infty \alpha_i t^i$ 。

2.1.2 商空间 (Quotient (or Factor) Space)

商空间就是将商集的概念转到线性空间上形成的。首先对于一个向量空间 \mathbb{R}^3 来说, 我们可以对其确定一个**等价关系**进而**划分出等价类**。

比如, 我们设在同一平面上的两个向量有等价关系 (重力势能相等), 那么我们就能够划分出平面 \mathbb{R}^2 。根据我们之前所讨论过划分的类**交集为空**的性质, 这些平面必须要**平行**于彼此。因此, 代表这些平面的法向量 \hat{n} 一定是在同一条直线上的 \mathbb{R} (重力势能仅和一维的高度相关)。而这条直线作为所有等价类的集合就是商集, 同时它也是一个实数域上的一个线性空间, 因为其满足平面的法向量的加法和标量乘法。我们也不难看出, 其中必须唯一存在一个穿过原点的平面, 对于这样的商集我们就称其为**商空间**。

我们还可以设在同一条直线上的两个向量有等价关系 (同一磁场线上),

划分出 \mathbb{R}^3 的子空间 \mathbb{R} 。我们可以简化的考虑这些直线都垂直于 xy 平面 (均匀磁场), 那么我们的**商空间**就是 xy 平面 \mathbb{R}^2 。之前我们用一维的向量表示二维的平面, 这里我们又反过来用二维的向量来代表一个一维直线“穿过” xy 平面的位置。我们当然不能够满足这样直观和浅显的具体解释, 以及对其性质和维度上联系的一些模糊的领悟, 因而接下来我们将会从更具有普适性的角度利用基本性质去抽象的定义和构造, 形而上的讨论其本质和证明。

定义 2.1.6 设 W 为域 \mathbb{F} 上的向量空间 V 中的一个**子空间**, 如果 $|a\rangle \in V \wedge |b\rangle \in V$, 那么称 $|a\rangle$ 和 $|b\rangle$ 有**关系**。如果 $|a\rangle - |b\rangle \in W$ 则 $|a\rangle$ 和 $|b\rangle$ 有**等价关系** $|a\rangle \sim |b\rangle$ 。 $|a\rangle$ 的**等价类**为 $[a] = \{|a\rangle + w_i | w_i \in W\}$; **商集**为 $V/W = \{[a] | |a\rangle \in V\}$ 。

下面我们通过定义向量的**加法和标量乘法**将商集转化为**商空间** V/W 。 $\alpha[a] + \beta[b] = [\alpha a + \beta b]$ 其中 $[\alpha a + \beta b]$ 是 $\alpha|a\rangle + \beta|b\rangle$ 的等价类。这个等式与等价类**代表**的选择是**无关的**。也就是说, 如果 $[a'] = [a], [b'] = [b]$, 那么 $[\alpha a' + \beta b'] = [\alpha a + \beta b]$ 也成立。因而我们必须有以下式子成立,

$$(\alpha|a'\rangle + \beta|b'\rangle) - (\alpha|a\rangle + \beta|b\rangle) \in W \quad (2.5)$$

既然, $|a'\rangle \in [a]$, 那么就有 $|w_1\rangle \in W$, 使得 $|a'\rangle = |a\rangle + |w_1\rangle$, 同样的有 $|b'\rangle = |b\rangle + |w_2\rangle$ 。因此有, $(\alpha|a'\rangle + \beta|b'\rangle) - (\alpha|a\rangle + \beta|b\rangle) = \alpha|w_1\rangle + \beta|w_2\rangle$ 因为 W 是一个子空间, 所以有 $\alpha|w_1\rangle + \beta|w_2\rangle \in W$, 使得**式 2.6** 成立, 因而我们的对其进行的加法和乘法的定义是合理的。

定理 2.1.3 定义了加法和乘法后, 我们再设其加法恒元为 $[a] + |0\rangle = [a] \in V/W$, 有时我们也将 $[a]$ 记作 $|a\rangle + W$, 那么**商空间** $V/W = \{|a\rangle + W | |a\rangle \in V\}$ 就是一个**线性空间**。

定理 2.1.4 设 W 的基为 $\{|a_i\rangle\}_{i=1}^n$, 对其的扩充 $\{|a_i\rangle, |b_j\rangle\}_{i=1, j=n+1}^{n, m}$ 为 V 的基, 那么 V/W 的基就是 $\{[b_j]\}_{j=n+1}^m$ 。

证明： 设 $[a] \in V/W, |a\rangle \in V$ ，那么就有，

$$[a] \equiv |a\rangle + W = \sum_{i=1}^n \alpha_i |a_i\rangle (\in W) + \sum_{j=n+1}^m \beta_j |b_j\rangle + W \quad (2.6)$$

$$= \sum_{j=n+1}^m \beta_j |b_j\rangle + W = \left[\sum_{j=n+1}^m \beta_j |b_j\rangle \right] = \sum_{j=n+1}^m \beta_j [b_j] \quad (2.7)$$

因此，任意属于 V/W 的向量 $[v]$ 都能够表示为 $\{[b_j]\}_{j=n+1}^m$ 的线性组合， $[v] = \sum_{j=n+1}^m \beta_j ([b_j] + W)$ 。♣

定理 2.1.5 上述论证引出了，如果 $\{[b_j]\}_{j=n+1}^m$ 是生成 V/W 的一个基，并且 V 是有限维的，那么 $\dim(V/W) = m - n = \dim V - \dim W$ 。

接下来，我们就只需用证明 $\{[b_j]\}_{j=n+1}^m$ 是**线性无关**的了。

证明： 假设 $\sum_{j=n+1}^m \beta_j ([b_j] + W) = [0]$ ，那么就有 $\sum_{j=n+1}^m \beta_j |b_j\rangle + W = |0\rangle + W = W$ ，意味着 $\sum_{j=n+1}^m \beta_j |b_j\rangle \in W$ ，使得

$$\sum_{j=n+1}^m \beta_j |b_j\rangle = \sum_{i=1}^n \alpha_i |a_i\rangle, \quad \sum_{j=n+1}^m \beta_j |b_j\rangle - \sum_{i=1}^n \alpha_i |a_i\rangle = 0 \quad (2.8)$$

回顾设定： $\{|a_i\rangle, |b_j\rangle\}_{i=1, j=n+1}^{n, m}$ 为 V 的基，所以它是线性无关的并且有 $\alpha_1 = \cdots = \alpha_n = \beta_{n+1} = \cdots = \beta_m = 0$ 使得推论式子 2.10 成立，证明结束。

2.1.3 直和 (Direct Sums)

将一个向量空间划分成为**不相交**的子空间在一些情况下是可行且方便的。当然，我们也可以反过来将两个子空间拼接成另一个子空间。接下来的例子十分重要，特别是如果你还没能够完全理解定义 2.1.6 中我们为什么会那样设置等价关系？等价类是怎么得出的？

例 2.1.4 直和与定义 2.1.6 的例子：

我们可以用等价关系在 \mathbb{R}^3 建立一个向心力场的结构。首先，我们等价关系设置为：如果描述向心力的向量 $|a\rangle, |b\rangle \in \mathbb{R}^3$ 投影到了同一个平面上，则它们对应的角动量相等。然后，让不同平面互相平行且唯一对应一个角动量。设子空间 L 为穿过原点的直线 \mathbb{R} 且角动量 $|l\rangle$ 属于 L ，设另一个**穿过原点** (交集为零向量) 的且垂直于 L 的平面 \mathbb{R}^2 为子空间 S ，有位矢 $|r\rangle \in S$ 。不难看出，对应任意的 $|a\rangle \in \mathbb{R}^3$ 都有与之对应的 $|l\rangle$ 使得 $|l\rangle - |a\rangle \in S$ ，因为 L 过原点垂直于 S ，这里的 $|l\rangle - |a\rangle$ 不过就是 $|a\rangle$ 投影在 S 平面上的向量。

因此，我们设置**等价关系**为：如果 $|l\rangle - |a\rangle \in S$ 那么就有 $|l\rangle \sim |a\rangle$ ，这样我们就有等价类 $[l] = \{|l\rangle + |s_i\rangle \mid |s_i\rangle \in S\}$ 。也就是，**等价类** $[l] = S + |l\rangle$ 是将原有平面 S 沿着“轴线” L 相互平行地平移了 $|l\rangle$ 而形成的一个新的平面 S_l 该平面唯一对应一个角动量 $|l\rangle$ 。等价类 $[l]$ 中 (投影在同平面) 的任意的向量 $|a\rangle \in \mathbb{R}^3$ 的角动量相等。**商集**作为这些等价类的集合就是我们一开始设定的直线 L 上所有向量构成的一个线性空间，因此我们也称其为商空间。

实际上，这里我们建立的 \mathbb{R}^3 向心力场是描述一个粒子在平面 \mathbb{R}^2 受向心力而运动的。我们设置了一个角动量垂直于平面的角动量 \mathbb{R} 使得对运动平面 \mathbb{R}^2 的描述拓展到了 \mathbb{R}^3 上。任意 $|r\rangle \in S, |l\rangle \in L$ 都能够唯一的对应到 $|a\rangle \in \mathbb{R}^3$ ，也就是 $|r\rangle + |l\rangle = |a\rangle, \forall (|r\rangle \in S, |l\rangle \in L, |a\rangle \in \mathbb{R}^3)$ 。比如，设 L 为 z 轴， S 为 xy 平面，那么就有 $(x, y, z) = (0, 0, z) + (x, y, 0)$ 。那么我们就将其记作 $\mathbb{R}^3 = L + S = \mathbb{R}^2 + \mathbb{R}$ 。这意味着，向量 $|a\rangle \in \mathbb{R}^3$ 同时包含着位置矢量和角动量的信息，或者说向心力场向量的分量为位置矢量和角动量。

我们可以将这个例子推广到任意的线性空间上，如下定义所示：

定义 2.1.7 设 U 和 W 为线性空间 V 的**子空间**，我们将 V 中所有能够表示为子空间 U, W 中两个向量 $|u\rangle, |w\rangle$ 之**和**的向量所组成的集合记作 $U + W = \{|u\rangle + |w\rangle \mid \forall (|u\rangle + |w\rangle) \in V\}$ 。实际上，我们还不难看出 $U + W$ 也是 V 的子空间。

不过,我们发现有时 $U+W$ 中的向量并不能唯一的表示为两个子空间中向量之和,有可能会有 $|a\rangle = |u_1\rangle + |w_3\rangle = |u_2\rangle + |w_1\rangle$ 。比如,设 U 为 xy 平面 \mathbb{R}^2 , W 为 yz 平面 \mathbb{R}^2 ,它们和 $U+W$ 都是 \mathbb{R}^3 的子空间,事实上 $U+W = \mathbb{R}^3$ 。因为,对于任意的 $(x, y, z) \in \mathbb{R}^3$ 都有 $(x, y, z) = (x, \frac{1}{2}y, 0) \in U + (0, \frac{1}{2}y, z) \in W$ 。显然这不是唯一的,我们还能写做 $(x, y, z) = (x, \frac{1}{3}y, 0) + (0, \frac{2}{3}y, z)$ 等等。

定义 2.1.8 U, W 为 V 的子空间,如果 $V = U+W$ 并且 $U \cap W = \{|0\rangle\}$, 那么我们称 V 为 U, W 的直和记作 $V = U \oplus W$ 。

设 U 为过原点的一个 xy 平面, W 为过原点的一条 z 轴上的直线,则 U, W 的直和为 $\mathbb{R}^3 = U \oplus W$ 。我们发现对于任意的 $(x, y, z) \in \mathbb{R}^3$ 都能够唯一的等于 U, W 中两向量的和, $(x, y, z) = (x, y, 0) + (0, 0, z)$ 。

定理 2.1.6 $V = U \oplus W$ 中的任意一个向量 $|v\rangle$ 都能够唯一的表示为 U, W 中各取一个向量之间的和 $|u\rangle + |w\rangle$ 。

证明: 假设 V 中存在一个 $|v\rangle$ 不能够唯一的被表示,而是有

$$|v\rangle = |u\rangle + |w\rangle = |u'\rangle + |w'\rangle \Leftrightarrow |u\rangle - |u'\rangle = |w\rangle - |w'\rangle. \quad (2.9)$$

W 中的向量 $|w\rangle - |w'\rangle$ 和 U 中的向量 $|u\rangle - |u'\rangle$ 相等,意味着 U 和 W 的交集有 $|u\rangle - |u'\rangle$ 也是 $|w\rangle - |w'\rangle$ 。不过根据定义, $U \cap W = \{|0\rangle\}$ 使得:

$$|u\rangle - |u'\rangle = |w\rangle - |w'\rangle = |0\rangle \implies |u\rangle = |u'\rangle; |w\rangle = |w'\rangle. \quad (2.10)$$

我们证明了一个属于 U 的向量和一个属于 W 的向量有且只有一种方式相加唯一地得到一个属于 V 的向量。由此可见,定义中交集为零向量是引出唯一性的根源。♣ 其实,反过来先来定义唯一性然后推出交集为零也是可以的。如果 $U \cap W = \{|a\rangle\}$, 有标量乘法 $\alpha|a\rangle$ 也属于 U, W , 使得存在不唯一的 $|a\rangle = \frac{1}{3}|a\rangle + \frac{2}{3}|a\rangle = \frac{1}{4}|a\rangle + \frac{3}{4}|a\rangle$, 因此这违背了唯一性的定义, U, W 共同的向量只能是零向量。该结论能被推广至两个以上子空间的直和。

定义 2.1.9 设 $\{U_i\}_{i=1}^r$ 为 V 的子空间, 有 $V = \sum_{i=1}^r U_i$ and $U_i \cap U_j = \{0\}$, $\forall i, j = 1, \dots, r$. 则称 V 为 $\{U_i\}_{i=1}^r$ 的直和并记作 $\bigoplus_{i=1}^r U_i$.

设 U_1 为 x 轴上直线 \mathbb{R} , U_2 为 y 轴上直线 \mathbb{R} , U_3 为 z 轴上直线 \mathbb{R} . 显然有 $U_1 + U_2 + U_3 = \mathbb{R}^3$, 对于任意向量 $(x, y, z) \in \mathbb{R}^3$ 都唯一地有, $(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z)$, 使得 $\mathbb{R}^3 = U_1 \oplus U_2 \oplus U_3$.

定理 2.1.7 r 个不同子空间 U_i 中的向量集合 $\{|u_i\rangle\}_{i=1}^r$ 是线性无关的。

证明: 首先, 我们将 V 拆分成 $V = U_1 + W'$, 其中 $W' = U_2 \oplus \dots \oplus U_r$. 设 $\{|u_i\rangle\}_{i=1}^r$ 中所有向量都是非零向量且有 $u_i \in U_i$. 假设有等式,

$$\alpha_1|u_1\rangle + \alpha_2|u_2\rangle + \dots + \alpha_r|u_r\rangle = |0\rangle \Rightarrow \alpha_1|u_1\rangle = -\alpha|w'\rangle. \quad (2.11)$$

其中 $|w'\rangle \in W'$, 根据交集为零向量性质, $\alpha_1|u_1\rangle \in U_1$ 等于 $-\alpha|w'\rangle \in W'$ 仅当 $\alpha_1|u_1\rangle = |0\rangle$. 由于 $|u_1\rangle \neq |0\rangle$, 则有 $\alpha_1 = 0$. 继续设 $W' = U_2 \oplus W''$, $|w''\rangle \in W''$, 使得 $\alpha_2|u_2\rangle = -\beta|w''\rangle$, 同理得 $\alpha_2 = 0$, 那么我们一直递归下去, 那么式 2.11 成立当且仅当 $\alpha_1 = \dots = \alpha_r = 0$, 使得不同子空间 U_i 中的向量集合 $\{|u_i\rangle\}_{i=1}^r$ 是线性无关的。♣

定理 2.1.8 如果 $V = V_1 \oplus V_2$ 并且 $V = V_1 \oplus V_3$, 则有 $V_2 = V_3$ 成立。

定理 2.1.9 1. 设 U 是 V 的一个子空间, 则有 V 的子空间 W 使得 $V = U \oplus W$. 2. 直和的维度, 若 $V = U \oplus W$, 则有:

$$\dim V = \dim U + \dim W. \quad (2.12)$$

证明: 设 $\{|u_i\rangle\}_{i=1}^m$ 是 U 的一组基, 其基数为 m , 则存在一个子空间 $W = \text{Span}\{|w_j\rangle\}_{j=m+1}^n$, 使得这组基能扩展为 V 的一组基 $\{|u_i\rangle|w_j\rangle\}_{i=1, j=m+1}^{m, n}$, 其基数为 n , W 的基 $\{|w_j\rangle\}_{j=m+1}^n$ 基数为 $n - m$. ♣

定理 2.1.10 设 U 和 V 是 \mathbb{R} 或 \mathbb{C} 上的任意两个向量空间, 考虑笛卡尔积 $W \equiv U \times V$ 并且在 W 上定义零向量 $|0\rangle_W = (|0\rangle_U, |0\rangle_V)$ 以及,

$$\alpha(|u\rangle, |v\rangle) = (\alpha|u\rangle, \alpha|v\rangle) \quad (2.13)$$

$$(|u_1\rangle, |v_1\rangle) + (|u_2\rangle, |v_2\rangle) = (|u_1\rangle + |u_2\rangle, |v_1\rangle + |v_2\rangle). \quad (2.14)$$

为标量乘法和向量加法, 这样一来 $W = U \times V$ 就成为了一个**线性空间**。此外, 如果我们规定 U, V 中向量的形式为 $(|u\rangle, |0\rangle_V) \in U, (|0\rangle_U, |v\rangle) \in V$, 那么 U, V 为 W 的子空间且有 $W = U \oplus V$ 。

定理 2.1.11 设 $\{|u_i\rangle\}_{i=1}^M$ 是 U 的一组基, $\{|v_j\rangle\}_{j=1}^N$ 是 V 的一组基。设 $W = U \oplus V$, 定义向量 $|c_k\rangle_{k=1}^{M+N} \in W$ 为,

$$|c_i\rangle = (|u_i\rangle, |0\rangle_V), 1 \leq i \leq M \quad (2.15)$$

$$|c_i\rangle = (|0\rangle_U, |v_{i-M}\rangle), M+1 \leq i \leq M+N \quad (2.16)$$

那么这组向量的集合就是 W 的基, W 的维数是 $M+N$ 。

2.1.4 张量积 (Tensor Product)

定义 2.1.10 设 U 和 V 是向量空间并在它们的笛卡尔积上定义标量乘法和双线性条件 (*bilinearity conditions*):

$$\alpha(|u\rangle, |v\rangle) = (\alpha|u\rangle, |v\rangle) = (|u\rangle, \alpha|v\rangle) \quad (2.17)$$

$$(\alpha_1|u_1\rangle + \alpha_2|u_2\rangle, |v\rangle) = \alpha_1(|u_1\rangle, |v\rangle) + \alpha_2(|u_2\rangle, |v\rangle) \quad (2.18)$$

$$(|u\rangle, \beta_1|v_1\rangle + \beta_2|v_2\rangle) = \beta_1(|u\rangle, |v_1\rangle) + \beta_2(|u\rangle, |v_2\rangle). \quad (2.19)$$

该性质使得 $U \times V$ 转为一个称作**张量积**的线性空间记作 $U \otimes V$ ，张量积空间中的向量用 $|u\rangle \otimes |v\rangle$ 表示。

直和是用两个线性空间构造一个新的向量空间的一种方法。此外还有另一种方法为向量空间的张量积。若 $\{|u_i\rangle\}_{i=1}^M$ 和 $\{|v_j\rangle\}_{j=1}^N$ 分别是 U 和 V 的基，则式 2.16 和 2.17 分别有，

$$|u\rangle \otimes |v\rangle = \left(\sum_{i=1}^M \alpha_i |u_i\rangle\right) \otimes \left(\sum_{j=1}^N \beta_j |v_j\rangle\right) = \sum_{i=1}^M \sum_{j=1}^N \alpha_i \beta_j |u_i\rangle \otimes |v_j\rangle. \quad (2.20)$$

$$|0\rangle_U \otimes |v\rangle = (|u\rangle - |u\rangle) \otimes |v\rangle = |u\rangle \otimes |v\rangle - |u\rangle \otimes |v\rangle = |0\rangle_{U \otimes V} \quad (2.21)$$

例 2.1.5: 运用式 2.19 计算向量 $(x, y) \in \mathbb{R}^2$ 和 $(a, b, c) \in \mathbb{R}^3$ 的张量积, 设 \mathbb{R}^2 基 $\hat{e}_1 \equiv (1, 0), \hat{e}_2 \equiv (0, 1), \mathbb{R}^3$ 基 $\hat{E}_1 \equiv (1, 0, 0), \hat{E}_2 \equiv (0, 1, 0), \hat{E}_3 \equiv (0, 0, 1)$ 则有, $(x, y) \otimes (a, b, c) = xa(\hat{e}_1 \otimes \hat{E}_1) + ya(\hat{e}_2 \otimes \hat{E}_1) + xb(\hat{e}_1 \otimes \hat{E}_2) + yb(\hat{e}_2 \otimes \hat{E}_2) + xc(\hat{e}_1 \otimes \hat{E}_3) + yc(\hat{e}_2 \otimes \hat{E}_3)$. 我们还可以用一个 2×3 的矩阵来表示其结果, $M = \begin{pmatrix} xa & xb & xc \\ ya & yb & yc \end{pmatrix}$ 的每一个位置 m_{ij} 上的元素对应式 2.19 的 $\alpha_i \beta_j$ 。

定理 2.1.12 因此, $U \otimes V$ 的基为 $\{|u_i\rangle \otimes |v_j\rangle\}$ 并且 $\dim(U \otimes V) = \dim U \dim V$ 。此外有恒等式, $|0\rangle_U \otimes |v\rangle = |u\rangle \otimes |0\rangle_V = |0\rangle_{U \otimes V}$ 。

2.2 内积 (Inner Product)

定义所给出的向量空间过于一般化和无结构, 因而没有太多的物理意义。所以, 我们将向量空间上引入的一个实用的结构-**标量积** (scalar(dot) product)。例 1.2.1 的**点乘**也是标量积, 它将两个向量 \mathbf{a}, \mathbf{b} 与一个实数对应, 也就能够将其记作映射 $f: V \times V \rightarrow \mathbb{R}$ 有 $f(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$, 并且有对称性: $f(\mathbf{a}, \mathbf{b}) = f(\mathbf{b}, \mathbf{a})$, 还有线性: $f(\alpha \mathbf{a} + \beta \mathbf{b}, \mathbf{c}) = \alpha f(\mathbf{a}, \mathbf{c}) + \beta f(\mathbf{b}, \mathbf{c})$, 以及向量“长度”不为负: $|\mathbf{a}|^2 = f(\mathbf{a}, \mathbf{a}) = \mathbf{a} \cdot \mathbf{a} \geq 0$ 且 $f(\mathbf{a}, \mathbf{a}) = 0$ 当且仅当 $\mathbf{a} = \mathbf{0}$ 。

不过上述讨论都是在其标量为**实数域**上进行的, 下面我们先试着将标量积的性质推广到**复数域**上。然而, 这时就会产生一个矛盾: $f(i|a\rangle, i|a\rangle) = i^2 f(|a\rangle, |a\rangle) = -f(|a\rangle, |a\rangle)$ 违背了长度不为负的性质。在解决问题之前我们还需要一个前置知识,

定义 2.2.1 复数 $\alpha = a + bi = re^{i\theta}$ 的**共轭复数** (*complex conjugate*) 为 $\alpha^* = a - bi = re^{-i\theta}$ 。一元二次方程的复数根是共轭复数。

显然如果我们能够修改性质使得, $f(i|a\rangle, i|a\rangle) = i^* i f(|a\rangle, |a\rangle) = f(|a\rangle, |a\rangle)$ 问题就能够解决。我们发现问题的根源是出在**线性**性质, 那么我们让等式右侧首位标量对应应在等式左侧上的标量为其**共轭复数**就能够解决问题, 如下:

$$f(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha^* f(|a\rangle, |c\rangle) + \beta^* f(|b\rangle, |c\rangle). \quad (2.22)$$

注意: 第二个输入向量的变量, 也就是等式右侧函数 f 里面左侧的向量只是保持线性但是**不取共轭**。♣ 也就是,

$$f(|a\rangle, \beta|b\rangle, \gamma|c\rangle) = \beta f(|a\rangle, |b\rangle) + \gamma f(|a\rangle, |c\rangle) \quad (2.23)$$

(数学上通常约定正好和物理上**相反**: 内积首个向量是线性的而在第二个上是共轭线性的) 在内积这里狄拉克符号的实用性就得以体现了, 我们将 $f(|a\rangle, |b\rangle)$ 记作 $\langle a|b\rangle$ 。这里我们正是将函数内左侧的 $|a\rangle$ **取共轭**后表示为 $\langle a|$, 也就是 $\langle a| \equiv |a\rangle^*$, **这里尤为重要!**

例 2.2.1 复数域上的内积

设 $|a\rangle, |b\rangle \in \mathbb{C}^n$ 其中 $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $|b\rangle = (\beta_1, \beta_2, \dots, \beta_n)$, 那么 $\langle a| = (\alpha_1^*, \alpha_2^*, \dots, \alpha_n^*)$ 。我们定义 \mathbb{C}^n 上的内积为:

$$\langle a|b\rangle \equiv \alpha_1^* \beta_1 + \alpha_2^* \beta_2 + \dots + \alpha_n^* \beta_n = \sum_{i=1}^n \alpha_i^* \beta_i. \quad (2.24)$$

接下来我们验证这样的规定是符合内积的三个性质的，然而第一步我们似乎就遇到了问题 $f(|a\rangle, |b\rangle) \neq f(|b\rangle, |a\rangle)$, $\langle a|b\rangle \neq \langle b|a\rangle$, $\sum_{i=1}^n \alpha_i^* \beta_i \neq \sum_{i=1}^n \beta_i^* \alpha_i$ ，这是因为我们忘了在改变狄拉克符号同时需要取共轭，也就是说性质一实际上是 $\langle a|b\rangle = \langle b|a\rangle^*$ 。性质二就是式 2.21 和式 2.22，这里我们用狄拉克符号分别表示为：

$$(\alpha\langle a| + \beta\langle b|)|c\rangle = \alpha^*\langle a|c\rangle + \beta^*\langle b|c\rangle, \quad (2.25)$$

$$\langle a|(\beta|b\rangle + \gamma|c\rangle) = \beta\langle a|b\rangle + \gamma\langle a|c\rangle. \quad (2.26)$$

不过记法 $\langle a|(\beta|b\rangle + \gamma|c\rangle)$ 不方便和美观，因此我们将其记作 $\langle a|\beta b + \gamma c\rangle$ 。性质三就是 $\langle a|a\rangle = \alpha_1^* \alpha_1 + \cdots + \alpha_n^* \alpha_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2 \geq 0$ 。接下来我们就得以给出复数域上内积的定义了。

定义 2.2.2 在复数域上的向量空间 V 中，两个向量 $|a\rangle$ 和 $|b\rangle$ 的**内积**为一个复数 \mathbb{C} ，记作 $\langle a|b\rangle \in \mathbb{C}$ ，使得：

- 1. **共轭对称性** $\langle a|b\rangle = \langle b|a\rangle^*$
 - 2. **半双线性 (sesquilinear)** $\langle a|\beta b + \gamma c\rangle = \beta\langle a|b\rangle + \gamma\langle a|c\rangle$
 - 3. **正定性 (positive definite)** $\langle a|a\rangle \geq 0$, $\langle a|a\rangle = 0 \Leftrightarrow |a\rangle = |0\rangle$.
- 定义了内积结构的向量空间 V 叫做**内积空间 (inner product space)**。

正定实 (positive definite real)内积又称**欧几里得 (Euclidean)**内积。复数共轭导致复数域上的内积空间不是真正双线性空间而是半双线性，不过准确的说是**厄米共轭 (hermitian)**内积空间。因为我们将结合性质一和二构成了的特殊的半双线性形式称之为厄米共轭形式。

例 2.2.2 设 $f(x) g(x) \in \mathbb{C}(a, b)$ ，其中 $\mathbb{C}(a, b)$ 为所有 x 为变量的连续复数函数的线性空间。并且 $w(t)$ 称作**加权函数 (weight function)**是一个实值连续函数且在开区间 (a, b) 上任意点都为正数。定义内积为：

$$\langle x|y\rangle \equiv \int_a^b w(t)x^*(t)y(t)dt, \quad (2.27)$$

其中 a, b 为实数也可以是无穷大只要积分是存在的, 在无限维空间 $\mathbb{C}(a, b)$ 上取决于加权函数 $w(t)$ 可以定义不同的内积, 称作 $\mathbb{C}(a, b)$ 上的**标准内积** (*standard inner product*)。

2.2.1 正交性 (Orthogonality)

解析几何和微积分的向量通常用轴上的单位向量表示, 即具有单位长度和互相垂直的向量。这样的向量在抽象内积空间中也很重要。

定义 2.2.3 向量 $|a\rangle, |b\rangle \in V$ 是**正交** (*orthogonal*)的若 $\langle a|b\rangle = 0$, 是**标准** (*normal*)的若 $\langle a|b\rangle = 1$ 。一个**标准 (归一) 化向量** (*normalized vector*) $|e\rangle$ 有 $\langle e|e\rangle = 1$ 。 N 维线性空间上的一组基 $B = \{|e_i\rangle\}_{i=1}^N$ 是**标准正交基** (*orthonormal basis*)如果:

$$\langle e_i|e_j\rangle = \delta_{ij} \equiv \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (2.28)$$

其中 δ_{ij} 称作是**克罗内克符号** (*Kronecker delta*)。

例 2.2.3 正交性的例子:

1. 设 U 和 V 为内积空间并且有 $W = U \oplus V$, 若 $|w_i\rangle = (|u_i\rangle, |v_i\rangle), i = 1, 2$, 则 $\langle w_1|w_2\rangle = \langle u_1|u_2\rangle + \langle v_1|v_2\rangle$ 为 W 上内积的定义, 使得 W 为内积空间。其中任意 U 中的向量与 V 中任意向量正交, $U = \{(|u\rangle, |0\rangle_V) || u\rangle \in U\}; V = \{(|0\rangle_U, |u\rangle) || u\rangle \in U\}$ 。

2. $\mathbb{R}^n, \mathbb{C}^n$ 的标准基 $|e_1\rangle = (1, 0, \dots, 0), |e_2\rangle = (0, 1, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)$ 也是在一般内积下标准正交的。

3. 设 $|e_l\rangle = e^{ikx}/\sqrt{2\pi}$ 为 $\mathbb{C}(0, 2\pi)$ 上的函数且有 $w(x) = 1$, 那么有 $\langle e_j | e_k \rangle = \delta_{jk}$, 因为有以下两个积分成立:

$$\langle e_j | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} e^{ikx} dx = 1, \quad j = k \quad (2.29)$$

$$\langle e_j | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ilx} e^{ikx} dx = 0, \quad j \neq k. \quad (2.30)$$

4. 设 $|\psi_n\rangle = \sqrt{\frac{2}{a}} \sin(\frac{n\pi}{a}x)$ 为 $\mathbb{C}(0, 2\pi)$ 上的函数且有 $w(x) = 1$, 那么有 $\langle \psi_m | \psi_n \rangle = \delta_{mn}$, 证明:

$$\int \psi_m(x)^* \psi_n(x) = \frac{2}{a} \int_0^a \sin(\frac{m\pi}{a}x) \sin(\frac{n\pi}{a}x) dx \quad (2.31)$$

$$= \frac{1}{a} \int_0^a [\cos(\frac{m-n}{a}\pi x) - \cos(\frac{m+n}{a}\pi x)] dx \quad (2.32)$$

$$= \frac{1}{a} \left\{ \frac{\sin[(m-n)\pi]}{m-n} - \frac{\sin[(m+n)\pi]}{m+n} \right\} = \delta_{mn}. \quad (2.33)$$

2.2.2 格拉姆-施密特正交化 (The Gram-Schmidt Process)

在我准大一暑假学微积分的时候, 第一次周末有两个小题—计算点到线和点到面的距离。由于我没有背公式也没有做练习, 虽然能够做出点到线这样初中数学题, 然而卡在了点到面。于是我便开始任意设定连接点与面向量, 从零开始的推导。然而我摆弄了 5 分钟实在想不通该如何把一个向量投影到平面上, 好在鬼使神差地我想

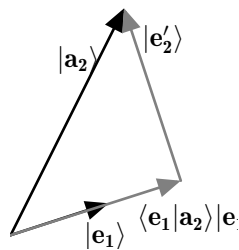


图 2.1: 点到线距离

先去用这套推出点到线距离公式看看。我轻松的将设定连接点与线向量 \mathbf{p} 运用点乘投影到线 \mathbf{s} 上。当然首先我们需要让 \mathbf{s} 是只描述方向的单位向量 $\hat{\mathbf{s}} = \frac{\mathbf{s}}{|\mathbf{s}|}$, $|\mathbf{s}| = \sqrt{\mathbf{s} \cdot \mathbf{s}}$, 这

样我就能将 \mathbf{p} 投影在线段的方向上得到 $\mathbf{p}_1 = (\mathbf{p} \cdot \hat{\mathbf{s}})\hat{\mathbf{s}}$, 使得点到线的距离为**垂直 (正交)**于线的向量 $\mathbf{d} = \mathbf{p} - \mathbf{p}_1$, 其模长就是答案。

接下来我顺着这样的思想, 首先从平面公式上随便生成两个点得到一个向量 \mathbf{a} , 然后我用其中一个点与目标点形成向量 \mathbf{p} 投影在线上 $\mathbf{p}_a = (\mathbf{p} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}}$, 不过 $\mathbf{p} - \mathbf{p}_a$ 是这条线到点的最短距离。不难想象 $\mathbf{p} - \mathbf{p}_a$ 投影在平面上垂直于 \mathbf{a} , 因此我们重复之前操作找到一个垂直于 \mathbf{a} 的向量 \mathbf{o} , 最后我们就能过得到连接目标点到平面的垂直距离为 $\mathbf{d} = \mathbf{p} - (\mathbf{p} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}} - (\mathbf{p} \cdot \hat{\mathbf{o}})\hat{\mathbf{o}}$ 。其实在那时我就以及模糊地觉察到了, 这样的过程或许可以继续的推广下去。不过鉴于我无法想象一个点到体的距离, 因而打消了这样的念头。

定义 2.2.4 格拉姆-施密特正交化是用投影的思想通过适当的线性组合将 V 中的任意一组基转换成一组标准正交基的方式。

假设基底 $B = \{|a_i\rangle\}_{i=1}^N$ 。我们可以对 $|a_i\rangle$ 线性组合得到一组标准正交的向量。首先, 我们直接标准化 $|a_1\rangle$ 得到 $|e_1\rangle = |a_1\rangle/\sqrt{\langle a_1|a_1\rangle}$ 使得 $\langle e_1|e_1\rangle = 1$ 。接下来我们希望让 $|e_2\rangle$ 正交于 $|e_1\rangle$, 那么就需要巧妙利用**点乘投影的思想**。

如图 2.1 所示, 我们将向量 $|a_2\rangle$ 投影在 $|e_1\rangle$ 上面得到 $\langle e_1|a_2\rangle|e_1\rangle$, 然后我们用 $|a_2\rangle$ 减去其在 $|e_1\rangle$ 上的投影就能够得到一个**垂直于** $|e_1\rangle$ 的向量 $|e'_2\rangle = |a_2\rangle - \langle e_1|a_2\rangle|e_1\rangle$ 。我们对其**归一化**就能够得到单位向量 $|e_2\rangle = |e'_2\rangle/\sqrt{\langle e'_2|e'_2\rangle}$ 。我们固然能够从图上直观的看出 $|e_1\rangle$ 垂直于 $|e'_2\rangle$, 但还是需要用“代数”推出其**正交性**有 $\langle e_1|e'_2\rangle = \langle e_1|a_2\rangle - \langle e_1|a_2\rangle\langle e_1|e_1\rangle = 0$ 。

接下来用我们用之前点到面推导的思路, 将 $|a_3\rangle$ 减去其在 $|e_1\rangle, |e_2\rangle$ 方向上投影的向量, 进而得到

$$|e'_3\rangle = |a_3\rangle - |e_1\rangle\langle e_1|a_3\rangle - |e_2\rangle\langle e_2|a_3\rangle \quad (2.34)$$

, 最后我们再对其归一化得到 $|e_3\rangle$ 。不难验证 $|e'_3\rangle$ 是**正交于** $|e_1\rangle$ 和 $|e_2\rangle$ 的, $\langle e_1|e'_3\rangle = \langle e_1|a_3\rangle - \langle e_1|e_1\rangle\langle e_1|a_3\rangle - \langle e_1|e_2\rangle\langle e_2|a_3\rangle = 0$, 同理有 $\langle e_2|e'_3\rangle = 0$ 。这其中的规律很显然, 如果我们已经计算了在该基底上 m 个标准正交向量 $|e_1\rangle, \dots, |e_m\rangle$ 且 $m < N$, 那么可以用如下规则计算下一个向量 $|e_{m+1}\rangle$:

$$|e'_{m+1}\rangle = |a_{m+1}\rangle - \sum_{i=1}^m |e_i\rangle \langle e_i | a_{m+1}\rangle, \quad (2.35)$$

$$|e_{m+1}\rangle = \frac{|e'_{m+1}\rangle}{\sqrt{\langle e'_{m+1} | e'_{m+1}\rangle}}. \quad (2.36)$$

注意：除了第一步之外的任意步骤 i 都是需要利用上一步得出的标准正交向量 $|e_i\rangle$ ，因此计算 $|e_{m+1}\rangle$ 需要我们重复上述操作 m 次。此外，每一步骤都是原有向量 $|a_{m+1}\rangle$ 减去基底的线性组合。♠

2.2.3 施瓦茨不等式 (Schwarz Inequality)

柯西-施瓦茨不等式是一个在数学上有着广泛应用的很重要的不等式，无论是有限维度还是无限维度它都成立。在二维和三维的情况下，我们能够很直观的想象该等式的描述。因为一个向量投影到另一个向量上的模长一定小于或者等于自身，毕竟 $\cos \theta$ 的最大值是 1，其中 θ 为两个向量的夹角。

定理 2.2.1 内积空间 V 上的任意一对向量 $|a\rangle, |b\rangle$ 都有**施瓦茨不等式**：
 $\langle a|a\rangle \langle b|b\rangle \geq |\langle a|b\rangle|^2$ 成立。

证明：首先，我们需要利用施密特正交化建立 $|a\rangle, |b\rangle$ 之间线性组合的关系。我们让第一个正交标准基为 $|e_1\rangle = |a\rangle / \sqrt{\langle a|a\rangle}$ ，那么就有，

$$|e_2\rangle = |b\rangle - \langle e_1|b\rangle |e_1\rangle = |b\rangle - \frac{\langle a|b\rangle}{\langle a|a\rangle} |a\rangle \Rightarrow |b\rangle = \frac{\langle a|b\rangle}{\langle a|a\rangle} |a\rangle + |e_2\rangle \quad (2.37)$$

$$\langle b|b\rangle = \left| \frac{\langle a|b\rangle}{\langle a|a\rangle} \right|^2 \langle a|a\rangle + \langle e_2|e_2\rangle = \frac{|\langle a|b\rangle|^2}{\langle a|a\rangle} + \langle e_2|e_2\rangle \quad (2.38)$$

$$\langle e_2|e_2\rangle \geq 0 \Rightarrow \langle b|b\rangle \geq \frac{|\langle a|b\rangle|^2}{\langle a|a\rangle} \Rightarrow \langle a|a\rangle \langle b|b\rangle \geq |\langle a|b\rangle|^2. (\text{Q.E.D}) \quad (2.39)$$

等式成立当且仅当 $\langle e_2 | e_2 \rangle = 0 \Rightarrow |e_2\rangle = |b\rangle - \frac{\langle a|b\rangle}{\langle a|a\rangle}|a\rangle = 0$, 也就是当 $|a\rangle, |b\rangle$ 成正比 (线性相关夹角为零) 时。这样抽象的推导使得施瓦茨不等式可以被广泛的运用, 因为我们仅仅只从的内积空间的基本假设推导出了施瓦茨不等式, 而**没有**用到内积的特殊性质。因此, 每当我们遇到一个新的内积空间时, 我们不需要再证明这个不等式。

2.2.4 向量的模长

中学在处理平面或空间中的有向线段等物体时, 我们可以用向量长度的直观概念来定义**点乘**。这里我们先引入内积, 然后再来定义向量的长度。

定义 2.2.5 内积空间中的一个向量 $|a\rangle$ 的**范数 (norm)**, 也就是长度定义为 $\|a\| \equiv \sqrt{\langle a|a\rangle}$ 。向量 $\alpha|a\rangle + \beta|b\rangle$ 的模记作 $\|\alpha a + \beta b\|$ 。

向量的范数有以下 3 条性质:

1. **范数正定性:** $\|a\| \geq 0$ 并且 $\|a\| = 0$ 当且仅当 $|a\rangle = |0\rangle$ 。
2. **齐次性:** $\|\alpha a\| = |\alpha| \|a\|$; 3. **三角不等式:** $\|a + b\| \leq \|a\| + \|b\|$ 。

向量空间上任何满足上述性质的函数称为**范数**, 赋范线性空间 (normed linear space) 是定义了范数的向量空间。因此, 我们并不需要内积也能够有范数。在赋范线性空间中我们可以引入两个向量之间的“距离”这样的概念。向量 $|a\rangle, |b\rangle$ 之间的距离记作并定义为 $d(a, b) \equiv \|a - b\|$ 。不难发现它满足之前**度量函数**的全部性质。不过, 我们并不需要用赋范线性空间来定义距离。比如, 我们可以定义球面上的两个点的距离, 但不需要定义两个点的加法, 也就是向量空间所必须有的结构没有定义, 因此球面的点形成了度量空间, 但不是线性空间。

定理 2.2.2 范数线性空间是内积空间当且仅当其范数满足平行四边形定律。

内积空间都自然是赋范线性空间, 不过反之则不成立。也就是满足上述

性质的空间一定是内积空间，但是如果其范数满足**平行四边形法则** (parallelogram law): $\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2$, 那么我们就可以在该空间上定义内积: $\langle a|b \rangle \equiv \frac{1}{4}\{\|a+b\|^2 - \|a-b\|^2 - i(\|a+ib\|^2 - \|a-ib\|^2)\}$

定理 2.2.3 所有有限维度向量空间都可以被转变为一个内积空间。

设 V 为 N 维度向量空间，有一组基 $\{a_i\}_{i=1}^N$ 对于任意向量 $|a\rangle$ 的分量为 $\{\alpha_i\}_{i=1}^N$, 也就是 $|a\rangle = (\alpha_1, \dots, \alpha_N)$ 。我们定义范数: $\|a\|^2 = \sum_{i=1}^N |\alpha_i|^2$, 不难验证其满足平行四边形法则。比如，我们考虑线性空间 \mathbb{C}^n 中两个向量 $|a\rangle, |b\rangle$ 之间的距离:

$$d(a, b) = \|a - b\| = \sqrt{\langle a - b | a - b \rangle} = \sqrt{\sum_{i=1}^n |\alpha_i - \beta_i|^2}. \quad (2.40)$$

我们还可以定义另一个范数, $\|a\|_1 \equiv \sum_{i=1}^n |\alpha_i|$ 并得到距离, $d_1(a, b) = \|a - b\|_1 = \sum_{i=1}^n |\alpha_i - \beta_i|$ 。实际上这两个范数都是以下范数和距离函数在 $p = 2$ 和 $p = 1$ 的特殊情况:

$$\|a\|_p \equiv \left(\sum_{i=1}^n |\alpha_i|^p \right)^{1/p} \quad (2.41)$$

$$d(a, b)_p = \|a - b\|_p = \left(\sum_{i=1}^n |\alpha_i - \beta_i|^p \right)^{1/p}. \quad (2.42)$$

2.3 线性映射 (Linear Maps)

尽管我们以及为线性空间增添了**内积**和**范数**这些结构，但是我们仍旧还是在单一的向量空间进行讨论。因此，若要带来更加广泛和重要的运用，我们需要让一个向量空间能够变换和联系到另一个向量空间且保持向量加法和标量乘法。这样的操作被称之为**线性映射** (linear maps)或者**线性变换**

(linear transformation)。我们先来用几个例子回顾之前学习的映射。

例 2.3.1 1. 设映射 $F: \mathbb{R}^2 \rightarrow \mathbb{C}$ 有 $F(x, y) = U(x, y) + iV(x, y)$ 且 $U: \mathbb{R}^2 \rightarrow \mathbb{R}, V: \mathbb{R}^2 \rightarrow \mathbb{R}$ 。设映射 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 有 $f(x) = (x-2, 3x-4)$ 。
2. 空间中质点的运动可以看作是映射 $M: [a, b] \rightarrow \mathbb{R}^3$, 其中 $[a, b]$ 是描述时间的实闭区间, 有任意时刻 $t \in [a, b]$ 。我们定义 $M(t) = (x(t), y(t), z(t))$, 其中 $x(t), y(t), z(t)$ 都是关于 t 的实值函数。这样 $M(t)$ 就是描述质点运动路径以时间为变量的函数, 其中 a 和 b 分别是开始和结束的时刻。

定义 2.3.1 复向量空间 V, W 之间的**线性映射** $T: V \rightarrow W$ 需要满足:

$$T(\alpha|a\rangle + \beta|b\rangle) = \alpha T(|a\rangle) + \beta T(|b\rangle), \forall |a\rangle, |b\rangle \in V, \alpha, \beta \in \mathbb{C}. \quad (2.43)$$

线性变换 $T: V \rightarrow V$ 称为 V 的**自同态 (endomorphism)**或 V 上的**线性算子 (linear operator)**。对一个向量的记法可省略括号: $T(|a\rangle) \equiv T|a\rangle$ 。同样的定义也适用于实向量空间, 注意两个向量空间一定有同一标量域。

由定义可得**线性映射** $T: V \rightarrow W$ 有以下性质: **1.** $T|0\rangle_V = |0\rangle_W$. 比如 $f(x) = x + 3$ 就不是线性映射。**2.** 若 $|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle \in V$ 线性相关, 则 $T|v_1\rangle, T|v_2\rangle, \dots, T|v_n\rangle$ 线性相关, 然反之则非也。**3.** 若有限维的 V 中的一组基底为 $B = \{|a_i\rangle\}_{i=1}^N$ 则对任意的 $|v\rangle \in V$ 都有 $|v\rangle = \sum_{i=1}^N \alpha_i |a_i\rangle$ 使得 $T|v\rangle = \sum_{i=1}^N \alpha_i T|a_i\rangle$ 。

定义 2.3.2 映射 $f: V \rightarrow W$ 是**反线性 (antilinear)**(或**共轭线性**或**半线性**) 映射则有: $f(\alpha a + \beta b) = \alpha^* f(a) + \beta^* f(b)$ 。

从 V 到 W 的不同线性映射组成的集合记作 $L(V, W)$, 集合 $L(V, W)$ 是一个向量空间。**零变换 0** 是对每一个 V 中向量都映射到 W 中的零向量上。

线性变换 \mathbf{T}, \mathbf{U} 的**加法**为线性变换 $\mathbf{T} + \mathbf{U}$, 其对向量 $|a\rangle \in V$ 有操作定义为 $(\mathbf{T} + \mathbf{U})|a\rangle \equiv \mathbf{T}|a\rangle + \mathbf{U}|a\rangle$ 。类似地, 定义**标量乘法** $\alpha\mathbf{T}$ 有 $(\alpha\mathbf{T})|a\rangle \equiv \alpha(\mathbf{T}|a\rangle) = \alpha\mathbf{T}|a\rangle$ 。 V 的**自同态**的集合用 $L(V)$ 或 $\text{End}(V)$ 表示, 而不是 $L(V, V)$ 。

定义 2.3.3 设 V 和 U 为内积空间, 线性映射 $\mathbf{T}: V \rightarrow U$ 称为**等距映射 (isometric map)**如有: $\langle \mathbf{T}a | \mathbf{T}b \rangle = \langle a | b \rangle, \forall |a\rangle, |b\rangle \in V$ 。

如果有 $U = V$, 则 \mathbf{T} 称为**线性等距 (linear isometry)**或 V 的等距。我们通常将复 (实) 内积空间 V 的等距称为**酉 (unitary)**(正交) 线性算子。

例 2.3.2 线性映射的例子:

1. 设 V 是一维空间 (如 $V = \mathbb{C}$), 则 V 的任意自同态 \mathbf{T} 的形式为 $\mathbf{T}|x\rangle = \alpha|x\rangle$, 其中 α 为标量。如果 \mathbf{T} 是等距, 则有 $|\alpha|^2 = 1$ 。如果 $V = \mathbb{R}$ 且 \mathbf{T} 是等距, 则有 $\mathbf{T}|x\rangle = \pm|x\rangle$ 。
2. 设双射函数 $\pi: \mathbb{N} \rightarrow S$ 为 $S = \{1, 2, \dots, n\}$ 上的任意一个**排列 (permutation)**。也就是说 π 会任意打乱 (赋予) S 一个顺序, 那么对于 $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{C}^n$ 就有**线性映射** $\mathbf{A}_\pi|a\rangle = (\alpha_{\pi(1)}, \alpha_{\pi(2)}, \dots, \alpha_{\pi(n)})$ 。
3. 任意的多项式向量 $|x\rangle \in \mathcal{P}^c[t]$ 有 $x(t) = \sum_{j=0}^n \alpha_j t^j$, 定义 $\mathbf{D}|x\rangle = |x_d\rangle$, $x_d(t) = \sum_{j=1}^n k \alpha_j t^{j-1}$, 那么线性映射 \mathbf{D} 称作**导数算子 (derivative operator)**。若定义 $\mathbf{S}|x\rangle = |x_s\rangle$, $x_s(t) = \sum_{j=0}^n [\alpha_j / (k+1)] t^{j+1}$, 那么线性映射 \mathbf{S} 称作**积分算子 (integration operator)**。
4. 设 $C^n(a, b)$ 为闭区间 $[a, b]$ 上 n 阶可导且连续的所有实值函数, 对于任意 $|f\rangle \in C^n(a, b)$ 定义 $\mathbf{G}|f\rangle = |u\rangle$, $u(t) = g(t)f(t)$, 其中 $g(t)$ 为 $C^n(a, b)$ 中一个固定的函数, 那么 \mathbf{G} 为线性映射。

由之前的性质 3 可知: 两个线性变换 $\mathbf{T}: V \rightarrow W$ 和 $\mathbf{U}: V \rightarrow W$ 是相

等的当且仅当对于所有 V 基底中的 $|a_i\rangle$ 都有 $\mathbf{T}|a_i\rangle = \mathbf{U}|a_i\rangle$ 成立。因此，一个线性变换是由它在其定义域上的某一组基底所唯一决定的。对于内积空间来说我们还可以用其他更方便的方法来定义线性算子的相等。

定理 2.3.1 内积空间的自同态 \mathbf{T} 为零变换 $\mathbf{0}$ 当且仅当 $\langle b|\mathbf{T}a\rangle = 0, \forall |a\rangle, |b\rangle$ 。

证明：如果 $\mathbf{T} = \mathbf{0}$ 那么显然就有 $\langle b|\mathbf{T}a\rangle = 0$ ，当且仅当需反之亦然。因此，如果 $\langle b|\mathbf{T}a\rangle = 0$ 并且定义映射 $|b\rangle = \mathbf{T}|a\rangle$ ，那么根据内积的正定性就有， $\langle \mathbf{T}a|\mathbf{T}a\rangle = 0 \Leftrightarrow \mathbf{T}|a\rangle = 0 \Leftrightarrow \mathbf{T} = 0, \forall |a\rangle$ 。♣

定理 2.3.2 内积空间的自同态 \mathbf{T} 为零变换 $\mathbf{0}$ 当且仅当 $\langle a|\mathbf{T}a\rangle = 0, \forall |a\rangle$ 。

证明：如果 $\mathbf{T} = \mathbf{0}$ 那么 $\langle a|\mathbf{T}a\rangle = 0$ ，反之如果 $\langle a|\mathbf{T}a\rangle = 0, \forall |a\rangle$ 且设 $|c\rangle = \alpha|a\rangle + \beta|b\rangle$ 则有 $\langle c|\mathbf{T}c\rangle = \langle \alpha a + \beta b|\mathbf{T}|\alpha a + \beta b\rangle = 0$ ，引入极化恒等式 (polarization identity):

$$\alpha^* \beta \langle a|\mathbf{T}b\rangle + \alpha \beta^* \langle b|\mathbf{T}a\rangle = \langle c|\mathbf{T}c\rangle - |\alpha|^2 \langle a|\mathbf{T}a\rangle - |\beta|^2 \langle b|\mathbf{T}b\rangle \quad (2.44)$$

根据定理 $\langle a|\mathbf{T}a\rangle = \langle b|\mathbf{T}b\rangle = \langle c|\mathbf{T}c\rangle = 0$ 使得 $\alpha^* \beta \langle a|\mathbf{T}b\rangle + \alpha \beta^* \langle b|\mathbf{T}a\rangle = 0$ 。引用定理 2.3.1: $\langle b|\mathbf{T}a\rangle = 0, \forall |a\rangle, |b\rangle \Leftrightarrow \mathbf{T} = 0, \forall |a\rangle$ 使得前后一致。♣

为了证明内积空间上的两个算子 \mathbf{T} 和 \mathbf{U} 是相等的，我们可以让它们作用于任意向量并证明它们会得到同样的结果 $\langle a|\mathbf{T}b\rangle = \langle a|\mathbf{U}b\rangle, \forall |a\rangle, |b\rangle$ ，或者用上述定理来证明 $\mathbf{T} - \mathbf{U}$ 是零变换 $\mathbf{0}$ ， $\langle a|\mathbf{T} - \mathbf{U}|b\rangle = 0, \forall |a\rangle, |b\rangle$ 。

定义 2.3.4 极化恒等式 用赋范向量空间的范数表示两个向量的内积。

首先，考虑实向量空间有满足平行四边形法则的范数，

$$\|a + b\|^2 = \|a\|^2 + 2\langle a|b\rangle + \|b\|^2 \quad (2.45)$$

$$\langle a|b \rangle = \frac{1}{2}(\|a+b\|^2 - \|a\|^2 - \|b\|^2) \quad (2.46)$$

$$= \frac{1}{2}(\|a\|^2 + \|b\|^2 - \|a-b\|^2) \quad (2.47)$$

$$= \frac{1}{4}(\|a+b\|^2 - \|a-b\|^2) \quad (2.48)$$

式 2.43 就是我们常见的一个极化恒等式, 实际上我们还可以用点乘定义范数 $\|a+b\|^2 = \|a\|^2 + 2\langle a|b \rangle + \|b\|^2$, $\|a-b\|^2 = \|a\|^2 - 2\langle a|b \rangle + \|b\|^2$ **两式相减**得到极化恒等式, 或者**两式相加**得到平行四边形法则。

我们顺着之前的结果得到, $\langle a|b \rangle + \langle b|a \rangle = \frac{1}{2}(\|a+b\|^2 - \|a-b\|^2)$ 。注意根据复内积的定义 $\langle a|b \rangle \in \mathbb{C}$, $\langle a|b \rangle$ 有虚数部分, 由于 $\langle a|b \rangle = \langle b|a \rangle^*$ 使得 $\langle a|b \rangle + \langle b|a \rangle$ 正好抵消了各自的虚数部分, 好在这也告诉我们它们的实数部分相等, 也就是有 $\Re\langle a|b \rangle = \frac{1}{4}(\|a+b\|^2 - \|a-b\|^2)$ 。接下来, 我们还需要找到虚数部分。既然 $\Im(\langle a|b \rangle + \langle b|a \rangle) = 0$, 那么 $\Im(\langle a|b \rangle - \langle b|a \rangle) = 2\Im(\langle a|b \rangle)$ 。找到 $\langle a|b \rangle - \langle b|a \rangle$ 只需用对之前**两式相减**得到极化恒等式稍作修改:

$$\langle a+ib|a+ib \rangle = \langle a|a \rangle + \langle ib|a \rangle + \langle a|ib \rangle + \langle b|b \rangle \quad (2.49)$$

$$\|a+ib\|^2 = \|a\|^2 - i\langle b|a \rangle + i\langle a|b \rangle + \|b\|^2 \quad (2.50)$$

$$\|a-ib\|^2 = \|a\|^2 + i\langle b|a \rangle - i\langle a|b \rangle + \|b\|^2 \quad (2.51)$$

$$\|a+ib\|^2 - \|a-ib\|^2 = 2i(\langle a|b \rangle - \langle b|a \rangle) \quad (2.52)$$

$$\langle a|b \rangle - \langle b|a \rangle = -\frac{1}{2}i(\|a+ib\|^2 - \|a-ib\|^2) \quad (2.53)$$

这样我们找到了 $\Im(\langle a|b \rangle) = -\frac{1}{4}(\|a+ib\|^2 - \|a-ib\|^2)$, 最后结合实部:

$$\langle a|b \rangle \equiv \frac{1}{4}\{\|a+b\|^2 - \|a-b\|^2 - i(\|a+ib\|^2 - \|a-ib\|^2)\} \quad (2.54)$$

这里的符号确实有些歧义, a, b 在这里都是向量, 而不是我们所习惯的

实数, 所以 $\langle a + ib | a + ib \rangle = \|a + ib\|^2$ 而不能取共轭 $\langle a + ib | a - ib \rangle$ 。此外, 或许大家注意到我们推导出来的式 2.49 与大多数数学书上的极化恒等式不同, 并且恰好是它们的共轭。这是由于物理约定左矢, 也就是第一位自变量 (argument) 取共轭, 数学上一般在第二位取共轭。

2.3.1 线性映射的核 (kernel)

从线性映射的定义中可以直接得出性质 1: V 中的零向量的像也是 W 中的零向量 $\mathbf{T}|0\rangle_V = |0\rangle_W$ 。这对一般映射并不成立, 不过对线性映射是一定成立的。 V 中的其他向量也能够映射到 W 的零向量上, 有如下定理:

定理 2.3.3 若 V 中的向量组成的一个集合在线性变换 $\mathbf{T}: V \rightarrow W$ 下都映射到了 W 的零向量上, 那么该集合形成了 V 的子空间称为 \mathbf{T} 的核 (kernel), 或零空间 (null space), 记作 $\ker \mathbf{T} = \{|v\rangle \in V : \mathbf{T}|v\rangle = |0\rangle_W\}$ 。 $\ker \mathbf{T}$ 的维度称作 V 的零化度 (nullity)。

证明: 设 $|a\rangle, |b\rangle \in \ker \mathbf{T}$ 则 $\mathbf{T}|a\rangle = \mathbf{T}|b\rangle = |0\rangle_W$ 使 $\mathbf{T}(\alpha|a\rangle + \beta|b\rangle) = \alpha\mathbf{T}(|a\rangle) + \beta\mathbf{T}(|b\rangle) = |0\rangle_W$, 有 $\alpha|a\rangle + \beta|b\rangle \in \ker \mathbf{T}$ 。♣

定理 2.3.4 线性映射 $\mathbf{T}: V \rightarrow W$ 的值域 $\mathbf{T}(V)$ 是 W 的子空间 (证明思路如上)。我们将 $\mathbf{T}(V)$ 的维数称为 \mathbf{T} 的秩 (rank)。

定理 2.3.5 线性映射 \mathbf{T} 为单射当且仅当其核为零向量 $\ker \mathbf{T} = \{|0\rangle\}$ 。

证明: 假设 $|v_1\rangle, |v_2\rangle \in V$ 有同一个像 $\mathbf{T}|v_1\rangle = \mathbf{T}|v_2\rangle$, 那么意味 $\mathbf{T}(|v_1\rangle - |v_2\rangle) = |0\rangle_W$ 如果核为零向量, 那么 $|v_1\rangle = |v_2\rangle$ 使得该映射一定为单射。♣

定理 2.3.6 线性等距映射 \mathbf{T} 一定是单射的。

证明: 设 \mathbf{T} 为等距映射且 $\forall |a\rangle, |b\rangle \in \ker \mathbf{T}$ 则有 $\langle a|b\rangle = \langle \mathbf{T}a|\mathbf{T}b\rangle = \langle 0|0\rangle = 0$ 。由内积性质可得 $|a\rangle = |b\rangle = |0\rangle_V$, 由定理 2.3.5 得 \mathbf{T} 为单射。♣

定理 2.3.7 对于任意线性映射 $\mathbf{T} : V \rightarrow W$ 都有**秩-零化度定理** (*Rank-nullity theorem*)或**维度定理** (*dimension theorem*):

$$\dim V = \dim \ker \mathbf{T} + \dim \mathbf{T}(V) \quad (2.55)$$

证明: 设 V 的一组基为 $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle\}$ 且 $\ker \mathbf{T}$ 的一组基为 $B' = \{|a_1\rangle, |a_2\rangle, \dots, |a_k\rangle\}$ 。对于任意 $|v\rangle \in V$ 都有 $\mathbf{T}|v\rangle = \mathbf{T} \sum_{i=1}^n \alpha_i |a_i\rangle$ 。其求和中的前 k 项都是核中的基, 因此 $\mathbf{T}\alpha_1|a_1\rangle = \mathbf{T}\alpha_2|a_2\rangle = \dots = \mathbf{T}\alpha_k|a_k\rangle = 0$ 。也就是意味着, $\mathbf{T}(V)$ 的一组基为 $\{\mathbf{T}|a_{k+1}\rangle, \dots, \mathbf{T}|a_n\rangle\}$ 使得 $\dim \mathbf{T}(V) = n - k, \dim V = n, \dim \ker \mathbf{T} = k$ 。♣

定理 2.3.8 有限维向量空间的**自同态**是双射, 如果其为单射或满射。

证明: 引理 2.3.5 单射映射当且仅当核为零向量使得引理 2.3.7 有 $\dim V = \dim \mathbf{T}(V)$, 单射且基数相同则等价于双射。然后自同态若是满射则直接可得 $\dim V = \dim \mathbf{T}(V)$ 。♣ **注意:** 维度定理只对有限维向量空间有效。♠

例 2.3.3 设 $\mathbf{T} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ 给定 $\mathbf{T}(x_1, x_2, x_3, x_4) = (2x_1 + x_2 + x_3 - 3x_4, x_1 + x_2 + 2x_3 + 2x_4, x_1 - x_3 - 3x_4)$ 。首先, 我们先来计算 $\ker \mathbf{T}$:

$$\mathbf{T}(x_1, x_2, x_3, x_4) = (0, 0, 0) \quad (2.56)$$

$$2x_1 + x_2 + x_3 - 3x_4 = 0 \quad (2.57)$$

$$x_1 + x_2 + 2x_3 + 2x_4 = 0 \quad (2.58)$$

$$x_1 - x_3 - 3x_4 = 0 \quad (2.59)$$

该方程的“解”为: $x_1 = x_3 + 3x_4; x_2 = -3x_3 - 5x_4$ 。因此, 若要在 $\ker \mathbf{T}$ 中, 其向量必须是 \mathbb{R}^4 的形式: $(x_3 + 3x_4, -3x_3 - 5x_4, x_3, x_4) = x_3(1, -3, 1, 0) + x_4(3, -5, 0, 1)$ 。其中 x_3, x_4 为任意实数, 使得任意 $\ker \mathbf{T}$

中的向量都是两个线性无关的向量 $(1, -3, 1, 0), (3, -5, 0, 1)$ 的线性组合。因此有 $\dim \ker \mathbf{T} = 2$ 。根据定理 2.3.7 有 $\dim \mathbf{T}(\mathbb{R}^4) = 2$, 注意到: $\mathbf{T}(x_1, x_2, x_3, x_4) = (2x_1 + x_2 + x_3 - 3x_4)(1, 0, 1) + (x_1 + x_2 + 2x_3 + 2x_4)(0, 1, -1)$ 。因此, 任意 $\mathbf{T}(x_1, x_2, x_3, x_4) \in \mathbf{T}(\mathbb{R}^4)$ 都可以写作两个线性无关的向量 $(1, 0, 1), (0, 1, -1)$ 。

2.3.2 线性同构 (Linear Isomorphism)

定义 2.3.5 向量空间 V 是**同构的** (*isomorphic*) W 记作 $V \cong W$, \mathbf{T} 是**同构** (*isomorphism*) 若存在一个双射线性映射 $\mathbf{T}: V \rightarrow W$ 或者说**可逆** (*invertible*) 线性映射。一个双射的自同态 $\mathbf{T}: V \rightarrow V$ 称之为 V 的**自同构** (*automorphism*)。 V 的自同构集和记作 $GL(V)$ 。

在许多情况下, 两个向量空间虽然“看似”不同, 但实际上它们非常相似。例如, 实数域上的二维向量空间 \mathbb{C} 就像是线性空间 \mathbb{R}^2 。广义上来说就是, 域 \mathbb{F} 上的 n 维向量空间类似于向量空间 \mathbb{F}^n 。上述定义的**同构**就是描述相似性的普遍抽象性质。

映射 $\mathbf{T}: \mathbb{C} \rightarrow \mathbb{R}^2, \mathbf{T}(x + iy) = (x, y)$ 是同构, 不难验证首先 \mathbf{T} 是线性映射, 然后其为单射, 最后是满射。不过我们需要**特别注意**: 集合 \mathbb{C} 和 \mathbb{R}^2 是同构的, 仅在其**结构** (*structure*) 都为线性空间下成立, 如果我们考虑其他结构, 则它们可能不是同构的。比如, \mathbb{C} 的元素有乘法, 但是 \mathbb{R}^2 没有。

定理 2.3.9 有限维向量空间 V 的**等距**是该向量空间的**自同构**。

证明: 根据定义 2.3.3, V 的**等距** \mathbf{T} 既是自同态也是等距映射。定理 2.3.6 得到 \mathbf{T} 作为等距映射是单射, 定理 2.3.8 得到自同态 \mathbf{T} 是双射的, 因而根据定义 2.3.5 可得 \mathbf{T} 也是**自同构**。♣

定理 2.3.10 **满射**线性映射 $\mathbf{T}: V \rightarrow W$ 是**同构**当且仅当其**零化度**为零。

证明: 因为 $\ker \mathbf{T} = \{|0\rangle_V\}$, 定理 2.3.5 可得 \mathbf{T} 是单射, 使得 \mathbf{T} 为既是单射又是满射的双射线性映射也就是**同构**。♣

定理 2.3.11 单射线性映射 $\mathbf{T}: V \rightarrow W$ 将一组线性无关的向量对应到另一组线性无关的向量上。

证明: 由单射可得如果 $\mathbf{T}|a\rangle = \mathbf{T}|b\rangle$, 则 $|a\rangle = |b\rangle$ 。因此有 $\alpha_1|a_1\rangle + \alpha_2|a_2\rangle + \cdots + \alpha_n|a_n\rangle = 0 \Leftrightarrow \alpha_1\mathbf{T}|a_1\rangle + \alpha_2\mathbf{T}|a_2\rangle + \cdots + \alpha_n\mathbf{T}|a_n\rangle = 0, \alpha_1 = \cdots = \alpha_n = 0$, 使得 $|a_1\rangle, |a_2\rangle, \cdots, |a_n\rangle$ 线性无关当且仅当 $\mathbf{T}|a_1\rangle, \mathbf{T}|a_2\rangle, \cdots, \mathbf{T}|a_n\rangle$ 。♣

定理 2.3.12 两个有限维向量空间 V, W 是同构的, 当且仅当它们具有相同的维数 $\dim V = \dim W$ 。

证明: 设有限维空间 V, W 有 $\dim V = \dim W = n$, 则 V 的一组基为 $B_V = \{|a_i\rangle\}_{i=1}^n$, W 的一组基为 $B_W = \{|b_i\rangle\}_{i=1}^n$ 。使得 V 中不同的向量 $|v\rangle$ 都能唯一的对应到 W 中的不同向量。因此 $\mathbf{T}: V \rightarrow W$ 一定是**双射**。♣

到这里, 我们终于验证了一开始所说的相似性就是同构。因此, **域 \mathbb{F} 上的 n 维向量空间与向量空间 \mathbb{F}^n 是同构的**。也就是, 所有的 \mathbb{R} 上的 n 为向量空间和 \mathbb{R}^n 都是同构的。

定理 2.3.13 设 $V = V_1 \oplus V_2$ 且 \mathbf{T} 是 V 的自同构, 使 V_1 为**不变量 (invariant)**, 即 $\mathbf{T}(V_1) = V_1$, 那么 V_2 也是不变量。实际上, 只要有一个是不变量, 则另一个也是不变量。

证明: $V_1 \oplus V_2 = \mathbf{T}(V) = \mathbf{T}(V_1 \oplus V_2) = \mathbf{T}(V_1) \oplus \mathbf{T}(V_2) = V_1 \oplus \mathbf{T}(V_2)$ 由定理 2.1.11 可得 $\mathbf{T}(V_2) = V_2$ 。♣

例 2.3.4 维度定理的另一个证明:

设 $\mathbf{T}': V \rightarrow W$ 为线性映射且 $\mathbf{T}': V/\ker \mathbf{T} \rightarrow \mathbf{T}(V)$ 。如果 $|a\rangle$ 为等价类 $[a]$ 的代表, 则 $\mathbf{T}'([a]) = \mathbf{T}|a\rangle$ 。

定义 2.3.6 若 f 定义良好 (*well defined*) 则有 $x = y \Rightarrow (x) = f(y)$ 。

首先, 该映射是有定义良好的, 因为如果 $[a'] = [a]$ 则有 $|a'\rangle = |a\rangle + |z\rangle$ 其中 $|z\rangle \in \ker \mathbf{T}$ 使得 $\mathbf{T}'([a']) \equiv \mathbf{T}|a'\rangle = \mathbf{T}(|a\rangle + |z\rangle) = \mathbf{T}|a\rangle + \mathbf{T}|z\rangle = \mathbf{T}|a\rangle$ 。

接下来, 我们证明 \mathbf{T}' 是一个同构。首先, 因为 \mathbf{T} 是线性映射, 所以有 $\mathbf{T}'(\alpha[a] + \beta[b]) \equiv \mathbf{T}(\alpha|a\rangle + \beta|b\rangle) = \alpha\mathbf{T}(|a\rangle) + \beta\mathbf{T}(|b\rangle) = \alpha\mathbf{T}'([a]) + \beta\mathbf{T}'([b])$ 使得 \mathbf{T}' 是线性映射。然后, 对于任意的 $|x\rangle \in \mathbf{T}(V)$, 都存在 $|y\rangle \in V$ 使得 $|x\rangle = \mathbf{T}|y\rangle = \mathbf{T}'([y])$, 因此 \mathbf{T}' 为满射。最后, 如果 $\mathbf{T}'([x]) = \mathbf{T}'([y])$, 则意味着 $\mathbf{T}|x\rangle = \mathbf{T}|y\rangle$, 也就是 $\mathbf{T}(|x\rangle - |y\rangle) = 0$, 那么有 $|x\rangle - |y\rangle \in \ker \mathbf{T}$ 使得 $[x] = [y]$, 因此 \mathbf{T}' 为单射。因此得到了 \mathbf{T}' 是一个同构, 根据定理 2.4.4: $\dim(V/\ker \mathbf{T}) = \dim \mathbf{T}(V) \Leftrightarrow \dim(V) = \dim \ker \mathbf{T} + \dim \mathbf{T}(V)$ (Q.E.D)。

定理 2.3.14 设 V, W 为线性空间且有线性映射 $\mathbf{T}: V \rightarrow W$ 。若 U 为 V 的子空间, 定义 $\mathbf{T}': V/U \rightarrow \mathbf{T}(V)$ 有 $\mathbf{T}'([a]) = \mathbf{T}|a\rangle$, 其中 $|a\rangle$ 为 $[a]$ 的代表, 则 \mathbf{T}' 是一个定义良好的同构。

设 U, V, W 为复向量空间且定义线性映射如下:

$$\mathbf{T}: (U \oplus V) \otimes W \rightarrow (U \otimes W) \oplus (V \otimes W) \quad (2.60)$$

$$\mathbf{T}((|u\rangle + |v\rangle) \otimes |w\rangle) = |u\rangle \otimes |w\rangle + |v\rangle \otimes |w\rangle. \quad (2.61)$$

不难看出 \mathbf{T} 是同构, 因此有 $(U \oplus V) \otimes W \cong (U \otimes W) \oplus (V \otimes W)$ 。这里我们固然可以一如既往繁琐的证明, 但用定理 2.4.4 就要容易的多。根据定理 2.1.8 和 2.1.12 可得:

$$\dim((U \oplus V) \otimes W) = \dim(U \oplus V) \dim W \quad (2.62)$$

$$= (\dim U + \dim V) \dim W \quad (2.63)$$

$$= \dim U \dim W + \dim V \dim W \quad (2.64)$$

$$= \dim((U \otimes W) \oplus (V \otimes W)) \quad (2.65)$$

2.4 复结构 (Complex Structures)

到目前为止, 我们处理向量空间都避免改变其标量的性质。当向量空间是复数时, 我们始终保持这个向量空间的标量作为复数。在这部分中, 我们将探讨改变标量的可能性, 以及随之而来的在向量空间的其他结构中可能发生的相应变化。推广内积的概念能够方便我们讨论关于改变标量以及其他的形式化内容。虽然正确定性的概念对于内积的物理应用是至关重要的, 但它的限制性太大, 所以我们放宽这个要求, 重新定义内积。

定义 2.4.1 设 \mathbb{F} 非 \mathbb{C} 即 \mathbb{R} , 则一个域 \mathbb{F} 下的线性空间 V 之上的**内积**为线性映射 $g: V \times V \rightarrow \mathbb{F}$ 满足如下性质:

1. **对称性 (symmetry)**: $g(|a\rangle, |b\rangle) = g(|b\rangle, |a\rangle)$
2. **双线性 (bilinearity)**: $g(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha g(|a\rangle, |c\rangle) + \beta g(|b\rangle, |c\rangle);$
 $g(|c\rangle, \alpha|a\rangle + \beta|b\rangle) = \alpha g(|c\rangle, |a\rangle) + \beta g(|c\rangle, |b\rangle)$
3. **非退化 (non-degeneracy)**: $g(|a\rangle, |b\rangle) = 0, \forall |a\rangle \in V \Rightarrow |a\rangle = |0\rangle$

我们也可以用狄拉克符号 $\langle \cdot | \cdot \rangle_{\mathbb{F}}$ 来表示域 \mathbb{F} 上的内积, 当 $\mathbb{F} = \mathbb{R}$ 时, 则 $\langle \cdot | \cdot \rangle_{\mathbb{F}} = \langle \cdot | \cdot \rangle$ 。**非退化**也就是, 对于任意的非零向量 $|a\rangle \in V$ 一定都至少存在一个向量 $|b\rangle \in V$ 使得 $g(|a\rangle, |b\rangle) \neq 0$ 。也就是说, 仅只有零向量正交于内积空间的所有向量。

定义 2.4.2 线性算子 $A \in \text{End}(V)$ 的**伴随 (adjoint)**记作 A^T , 满足:

$$\langle Aa|b\rangle_{\mathbb{F}} = \langle a|A^T b\rangle_{\mathbb{F}} \text{ 或者 } \langle a|A^T b\rangle_{\mathbb{F}} = \langle b|Aa\rangle_{\mathbb{F}} \quad (2.66)$$

\mathbf{A} 是自伴算子 (*self-adjoint*) 若 $\mathbf{A}^T = \mathbf{A}$, 是反伴算子 (*skew*) 若:

$$\boxed{\mathbf{A}^T = -\mathbf{A}}. \quad (2.67)$$

命题 2.4.1 非退化使得 $\langle a | (\mathbf{A}^T)^T b \rangle_{\mathbb{F}} = \langle b | \mathbf{A}^T a \rangle_{\mathbb{F}} = \langle a | \mathbf{A} b \rangle_{\mathbb{F}}$, 因此有:

$$\boxed{(\mathbf{A}^T)^T = \mathbf{A}}. \quad (2.68)$$

命题 2.4.2 \mathbf{A} 是一个反伴算子当且仅当 $\langle x | \mathbf{A} x \rangle_{\mathbb{F}} = 0, \forall |x\rangle \in V$.

证明: 若 \mathbf{A} 为反伴算子, 则 $\langle x | \mathbf{A} x \rangle_{\mathbb{F}} = \langle x | \mathbf{A}^T x \rangle_{\mathbb{F}} = -\langle x | \mathbf{A} x \rangle_{\mathbb{F}} \Rightarrow \langle x | \mathbf{A} x \rangle_{\mathbb{F}} = 0$. 反之, 假设 $\langle x | \mathbf{A} x \rangle_{\mathbb{F}} = 0, \forall |x\rangle \in V$, 则对于非零的 $\alpha, \beta \in \mathbb{F}, |a\rangle, |b\rangle \in V$ 有,

$$0 = \langle \alpha a + \beta | \mathbf{A} \alpha a + \beta b \rangle_{\mathbb{F}} \quad (2.69)$$

$$= \alpha^2 \langle a | \mathbf{A} a \rangle_{\mathbb{F}} + \alpha \beta \langle a | \mathbf{A} b \rangle_{\mathbb{F}} + \alpha \beta \langle b | \mathbf{A} a \rangle_{\mathbb{F}} + \beta^2 \langle b | \mathbf{A} b \rangle_{\mathbb{F}} \quad (2.70)$$

$$= \alpha \beta (\langle b | \mathbf{A}^T a \rangle_{\mathbb{F}} + \langle b | \mathbf{A} a \rangle_{\mathbb{F}}). \quad (2.71)$$

既然 $\alpha \beta \neq 0$, 则 $\langle b | (\mathbf{A} + \mathbf{A}^T) a \rangle_{\mathbb{F}} = 0, \forall |a\rangle, |b\rangle \in V$. 根据非退化可得若要 $(\mathbf{A} + \mathbf{A}^T)|a\rangle = |0\rangle$ 对于任意 $|a\rangle$ 都成立, 则一定有 $\mathbf{A}^T = -\mathbf{A}$. ♣ 回顾定理 2.3.2 可见正定性带来的限制。

定义 2.4.3 实向量空间上的**复结构** (*complex structure*) \mathbf{J} 是满足 $\mathbf{J}^2 = -1$ 并且 $\langle \mathbf{J}a | \mathbf{J}b \rangle = \langle a | b \rangle, \forall |a\rangle, |b\rangle \in V$ 的一个线性算子。

命题 2.4.3 复结构 \mathbf{J} 也是反伴算子。

证明: 设 $|a\rangle \in V$ 且 $|b\rangle = \mathbf{J}|a\rangle$, 那么我们可以推出两个等式:

$$\langle a | \mathbf{J}a \rangle = \langle a | b \rangle = \langle \mathbf{J}a | \mathbf{J}b \rangle = \langle \mathbf{J}a | \mathbf{J}^2 a \rangle = -\langle \mathbf{J}a | a \rangle \quad (2.72)$$

$$\langle a | \mathbf{J}a \rangle = \langle a | b \rangle = \langle b | a \rangle = \langle \mathbf{J}a | a \rangle \quad (2.73)$$

使得 $-\langle \mathbf{J}a | a \rangle = \langle \mathbf{J}a | a \rangle$ 则 $\langle a | \mathbf{J}a \rangle = 0$ ，由命题 2.4.2 得 \mathbf{J} 为反伴。♣

定理 2.4.1 向量集合 $\{|e_i\rangle, \mathbf{J}|e_i\rangle\}_{i=1}^m$ 是 $n = 2m$ 维度实内积空间 V 的一组正交基，特别是 V 的维度必须是偶数才能有复结构 \mathbf{J} 。

证明： 设 $|a\rangle$ 为 n 维实内积空间中的任意向量，并对其归一化得到 $|e_1\rangle$ 。根据命题 2.4.2 和 2.4.3 可知 $\mathbf{J}|e_1\rangle$ 正交于 $|e_1\rangle$ ，接下来我们归一化 $\mathbf{J}|e_1\rangle$ 得到 $|e_2\rangle$ 。如果 $n > 2$ 设 $|e_3\rangle$ 为任意正交于 $|e_1\rangle$ 和 $|e_2\rangle$ 的单位向量，那么 $|a_3\rangle \equiv \mathbf{J}|e_3\rangle$ 显然正交于 $|e_3\rangle$ ，并且同样正交于 $|e_1\rangle$ 和 $|e_2\rangle$ ：

$$\langle e_1 | a_3 \rangle = \langle \mathbf{J}e_1 | \mathbf{J}a_3 \rangle = \mathbf{J}e_1 | \mathbf{J}^2 e_3 \rangle = -\mathbf{J}e_1 | e_3 \rangle = -\langle e_2 | e_3 \rangle = 0 \quad (2.74)$$

$$\langle e_2 | a_3 \rangle = \langle \mathbf{J}e_1 | \mathbf{J}e_3 \rangle = \langle e_1 | e_3 \rangle = 0 \quad (2.75)$$

重复上面的步骤就能够得到定理 2.4.1。♣

定义 2.4.4 复数集 \mathbb{C} 和实向量空间 V 的张量积 $\mathbb{C} \otimes V$ 若同时满足复数乘积：

$$\alpha(\beta \otimes |a\rangle) = (\alpha\beta) \otimes |a\rangle, \quad \forall \alpha, \beta \in \mathbb{C} \quad (2.76)$$

则 $\mathbb{C} \otimes V$ 为复向量空间并称作 V 的**复化** (*complexification*) 记作 $V^{\mathbb{C}}$ ，特别有 $(\mathbb{R}^n)^{\mathbb{C}} \equiv \mathbb{C} \otimes \mathbb{R}^n \cong \mathbb{C}^n$ 。

注意： $V^{\mathbb{C}}$ 在复数域下的维度等于 V 在实数域下的维度 $\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V$ 并且还有 $\dim_{\mathbb{R}} V^{\mathbb{C}} = 2 \dim_{\mathbb{R}} V$ 。若 V 的一组基为 $\{|a_k\rangle\}_{k=1}^N$ ，则它同样也是 $V^{\mathbb{C}}$ 作为复向量空间的一组基，毕竟 $\{|a_k\rangle, ia_k\rangle\}_{k=1}^N$ 是 $V^{\mathbb{C}}$ 作为实

向量空间的一组基。♠ 在复化了实向量空间后, 我们还可以在此基础上定义一个符合双线性 (厄密) 的内积: $\langle \alpha \otimes a | \beta \otimes b \rangle \equiv \alpha^* \beta \langle a | b \rangle, \forall |a\rangle, |b\rangle \in V$ 。

复化实向量空间 V , 我们必须将它与一组复数“相乘” $\mathbb{C} \otimes V$ 得到一个为原先 V 的维数两倍的实向量空间。那么我们不禁疑惑, 是否有一个反向“分割”(必然是偶数维)的实向量空间过程?

设 W 为 $2m$ 维实向量空间且 \mathbf{J} 是 W 上的一个复结构, 使得 $\{|e_i\rangle \mathbf{J}|e_i\rangle\}_{i=1}^m$ 是 W 的一组基, 并在子空间上 $W_1 \equiv \text{Span}\{|e_i\rangle\}_{i=1}^m$ 上定义复数的乘积: $(\alpha + i\beta) \otimes |w_1\rangle \equiv (\mathbf{1}\alpha + \mathbf{J}\beta)|w_1\rangle, \forall \alpha, \beta \in \mathbb{R}, |w_1\rangle \in W$. 就能够将一个 $2m$ 维实向量空间 W 转变为一个 m 维的复向量空间 $W_1^{\mathbb{C}}$ 。

2.5 线性泛函 (Linear Functionals)

定义 2.5.1 设 V 为域 \mathbb{F} 上的向量空间, 则**线性型 (linear form)**是从 V 对应到其标量域 \mathbb{F} 上的一个线性映射 $\mathbf{f}: V \rightarrow \mathbb{F}$ 。 V 的所有线性泛函 $\mathcal{L}(V, \mathbb{F})$ 构成的集合称之为 V 的**对偶空间 (dual space)**记作 V^* 。

注意:(表明某映射为线性泛函) 首先, 确保该映射是从一个向量空间对应到其标量域, 然后证明这个映射是一个线性映射。♠

例 2.5.1 线性泛函的例子:

1. 设 $|a\rangle = (\alpha_1, \dots, \alpha_n) \in \mathbb{C}$, $\mathbf{f}_1: \mathbb{C}^n \rightarrow \mathbb{C}$, $\mathbf{f}_1(|a\rangle) = \sum_{k=1}^n \alpha_k$ 。
2. 设 α_{ij} 为 $m \times n$ 矩阵 \mathbf{A} 的元素, $\mathbf{f}_2: \mathcal{A}^{m \times n} \rightarrow \mathbb{C}$, $\mathbf{f}_2(\mathbf{A}) = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij}$. $\mathbf{f}_2(\mathbf{A}) = \sum_{i=1}^m \alpha_{ii}$. 其结果为对角元素之和。
3. 连续函数的定积分是线性泛函, $\text{int}: \mathcal{C}^0(a, b) \rightarrow \mathbb{R}$, $\text{int}(f) = \int_a^b f(t)dt$ 。
4. 固定复内积空间 V 中向量 $|a\rangle$, $\mathbf{f}_a: V \rightarrow \mathbb{C}$, $\mathbf{f}_a(|b\rangle) = \langle a | b \rangle$ 为线性泛函。
5. 设 $\{|a_i\rangle\}_{i=1}^m$ 为 V 中任意向量集, $\{\mathbf{f}_j\}_{j=1}^m$ 为 V 的任意一组线性泛函, 那么我们可以将 V 上的一个**线性算子 \mathbf{A}** 定义为:

$$\mathbf{A} \equiv \sum_{k=1}^m |a_k\rangle \mathbf{f}_k \in \text{End}(V), \quad \mathbf{A}|x\rangle = \sum_{k=1}^m |a_k\rangle \mathbf{f}_k(|x\rangle) = \sum_{k=1}^m \mathbf{f}_k(|x\rangle) |a_k\rangle. \quad (2.77)$$

接下来我们讨论一个向量空间对应到其对偶空间的线性同构。设 n 维向量空间 V 的一组基为 $B = \{|a_i\rangle\}_{i=1}^n$ 且对于任意标量集合 $\{\alpha_i\}_{i=1}^n$ ，定义线性泛函 $\mathbf{f}_\alpha|a_i\rangle = \alpha_i$ 。那么对于任意 V 中向量 $|b\rangle = \sum_{i=1}^n \beta_i|a_i\rangle$ 都有：

$$\mathbf{f}_\alpha|b\rangle = \mathbf{f}_\alpha\left(\sum_{i=1}^n \beta_i|a_i\rangle\right) = \sum_{i=1}^n \beta_i \mathbf{f}_\alpha|a_i\rangle = \sum_{i=1}^n \beta_i \alpha_i. \quad (2.78)$$

定义 2.5.2 该等式表明向量 $|b\rangle$ 可以被表示为一个由 $\beta_1, \beta_2, \dots, \beta_n$ 所构成的**列向量 (column vector)**并且 \mathbf{f}_α 为一个由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 所构成的**行向量 (row vector)**。其中行向量的**转置 (transpose)**为列向量，列向量的转置为行向量并记作：

$$[\beta_1, \dots, \beta_n]^T = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}; \quad \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}^T = [\beta_1, \dots, \beta_n] \quad (2.79)$$

那么， $\mathbf{f}_\alpha|b\rangle$ 不过就是行向量和列向量的**矩阵乘积 (matrix product)**：

$$\mathbf{f}_\alpha|b\rangle = [\alpha_1, \dots, \alpha_n] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \alpha_1\beta_1 + \dots + \alpha_n\beta_n = \sum_{i=1}^n \beta_i \alpha_i. \quad (2.80)$$

线性泛函 \mathbf{f}_α 是被标量集合 $\{\alpha_i\}_{i=1}^n$ 所唯一确定的，使得有一组特别的线性泛函 $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ 分别对应标量集 $\{1, 0, \dots, 0\}, \{0, 1, \dots, 0\}, \dots, \{0, 0, \dots, 1\}$ 。

也就有, $\mathbf{f}_i|a_j\rangle = \delta_{ij}$, 其中 δ_{ij} 为克罗内克符号。

定理 2.5.1 对于 V 的任意一组基 $B = \{|a_i\rangle\}_{i=1}^n$ 都唯一有其对偶空间 V^* 上的基 $B^* = \{\mathbf{f}_i\}_{i=1}^n$ 与之对应且满足 $\mathbf{f}_i|a_j\rangle = \delta_{ij}$ 。

证明: 设任意 $\gamma \in V^*$, V 的基为 $B = \{|a_i\rangle\}_{i=1}^n$ 并且 $\gamma|a_i\rangle = \gamma_i \in \mathbb{C}$, 若要证明 $\gamma = \sum_{i=1}^n \gamma_i \mathbf{f}_i$, 则需注意到对应任意 $|a\rangle = \sum_{i=1}^n \alpha_i |a_i\rangle$ 都有:

$$\gamma|a\rangle = \gamma \sum_{i=1}^n \alpha_i |a_i\rangle = \sum_{i=1}^n \alpha_i \gamma|a_i\rangle = \sum_{i=1}^n \alpha_i \gamma_i, \text{ 同时} \quad (2.81)$$

$$\left(\sum_{i=1}^n \gamma_i \mathbf{f}_i\right)|a\rangle = \left(\sum_{i=1}^n \gamma_i \mathbf{f}_i\right)\left(\sum_{j=1}^n \alpha_j |a_j\rangle\right) = \sum_{i=1}^n \gamma_i \sum_{j=1}^n \alpha_j \mathbf{f}_i|a_j\rangle \quad (2.82)$$

$$= \sum_{i=1}^n \gamma_i \sum_{j=1}^n \alpha_j \delta_{ij} = \sum_{i=1}^n \gamma_i \alpha_i \quad (2.83)$$

既然 γ 和 $\sum_{i=1}^n \gamma_i \mathbf{f}_i$ 的确对任意 $|a\rangle$ 都有相同的结果, 因此有 $\gamma = \sum_{i=1}^n \gamma_i \mathbf{f}_i$, 即 $V^* = \text{Span}\{\mathbf{f}_i\}_{i=1}^n$ 。♣

注意: 根据定理 2.3.12 和 2.5.1 可知 n 维向量空间的^{对偶空间}也是 n 维的, 因此它们是同构的。设 V 的一组基为 B , 则其对偶空间 V^* 的一组基 B^* 称作 B 的^{对偶基} (dual basis)。定理 2.5.1 可^{推论}出 $\forall |a\rangle \in V$ 都唯一存在与之对应的一个线性泛函 $\mathbf{f}_a \in V^*$, 使得 $\mathbf{f}_a|a\rangle = \sum_{i=1}^n \alpha_i \mathbf{f}_i$ 。这里, 我们称线性泛函 \mathbf{f}_a 为 $|a\rangle$ 的^{对偶} (dual)。

定义 2.5.3 向量 $|a\rangle \in V$ 的^{零化子} (annihilator)为其对偶空间的一个线性泛函 $\mathbf{f} \in V^*$, 使得 $\mathbf{f}|a\rangle = 0$ 。设 W 是 V 的一个子空间, 在 V^* 中的一组^{零化} (annihilate)在 W 中所有向量的线性泛函集合记作 W^0

注意: 根据线性泛函的性质, W^0 是 V^* 的一个子空间。此外, 如果我们扩充 W 的一组基 $\{|a_i\rangle\}_{i=1}^k$ 至 $\{|a_i\rangle\}_{i=1}^{n+k}$, 那么从对偶基 $B^* = \{\mathbf{f}_i\}_{i=1}^{n+k}$ 中

截取的 $\{\mathbf{f}_i\}_{i=k+1}^{n+k}$ 能够生成 $W^0 = \text{Span}\{\mathbf{f}_i\}_{i=k+1}^{n+k}$, 也就意味着有如下等式:

$$\dim V = \dim W + \dim W^0 \quad (2.84)$$

在以后讨论辛几何时, 我们将有机会使用零化子。♠

定义 2.5.4 设 $\mathbf{T}: V \rightarrow W$ 为线性映射, 并将 $\mathbf{T}^*: W^* \rightarrow V^*$ 定义为:

$$[\mathbf{T}^*(\gamma)]|a\rangle = \gamma(\mathbf{T}|a\rangle) \Leftrightarrow \mathbf{T}^*(\gamma) = \gamma \circ \mathbf{T}, \forall |a\rangle \in V, \gamma \in W^* \quad (2.85)$$

那么我们将 \mathbf{T}^* 称作 \mathbf{T} 的**对偶 (dual)**, **转置**, 或者**拉回 (pullback)**。

$\mathbf{T}^* \in \mathcal{L}(W^*, V^*)$ 是一个 W^* 上的线性算子, 并且它的映射性质于 \mathbf{T} 紧密相连。为了理解这一点, 我们首先考虑 \mathbf{T}^* 的**核**。显然, γ 在 \mathbf{T}^* 的**核**中, 当且仅当 γ **零化**了所有的向量 $\mathbf{T}|a\rangle$, 因此有 $\gamma \in \mathbf{T}(V)^0$ 。**注意:** 如果 \mathbf{T} 是满射, $\mathbf{T}(V) = W$, 并且 γ 零化了所有 W 中的向量, 那么就有 $\ker \mathbf{T}^* = 0$, 并因此有 \mathbf{T}^* 为单射。类似的, 如果 \mathbf{T} 是单射, 则 \mathbf{T}^* 为满射。♠

命题 2.5.1 设 \mathbf{T}^* 为线性映射 \mathbf{T} 的**对偶**, 那么 $\ker \mathbf{T}^* = \mathbf{T}(V)^0$ 。如果 \mathbf{T} 是单射 (满射), 则 \mathbf{T}^* 为满射 (单射)。如果 \mathbf{T} 是同构则, \mathbf{T}^* 也是同构。

接下来, 我们需要联系**内积和线性泛函**。考虑一组基 $B = \{|a_i\rangle\}_{i=1}^n$ 并且设 $\alpha_i = \langle a|a_i\rangle$ 。因为标量集 $\{\alpha_i\}_{i=1}^n$ 唯一确定线性泛函 γ_a , 使得 $\gamma_a|a_i\rangle = \alpha_i$, 所以我们可以认为 γ_a 等同于 $\langle a|$, 记作 $\gamma_a \mapsto \langle a|$ 。

左矢和右矢之间, 我们用符号 $(|a\rangle)^\dagger \equiv \langle a|$ **剑标 (dagger)**, 或**共轭转置 (conjugate transpose)**来联系, 由内积的**对称和双线性 (sesqui-symmetry)**可推广至任意向量:

$$(\alpha|a\rangle + \beta|b\rangle)^\dagger = \alpha^*\langle a| + \beta^*\langle b|; \quad \gamma_{\alpha a + \beta b} \mapsto \alpha^*\langle a| + \beta^*\langle b|. \quad (2.86)$$

例 2.5.2 比如, 设 $|a\rangle \in \mathbb{C}^n$ 为列向量 $(\alpha_1, \dots, \alpha_n)^T$ 则有:

$$\gamma_a|b\rangle = \langle a|b\rangle = (\alpha_1^*, \dots, \alpha_n^*) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \alpha_1\beta_1 + \dots + \alpha_n\beta_n = \sum_{i=1}^n \alpha_i^* \beta_i. \quad (2.87)$$

线性泛函可运用在行列式的构造上以及后半部分在对张量的处理上。

2.6 多重线性映射 (Multilinear Maps)

定义 2.6.1 设 V 和 W 为线性空间, 并将 V 的 p 重 (p -fold)笛卡尔积记作 V^p , 那么一个从 V 对应到 W 的 p 重线性映射 (p -linear map)是对于每个参数都有线性的映射 $\mathbf{m}: V^p \rightarrow W$, 使得有:

$$\mathbf{m}(|a_1\rangle, \dots, \alpha|a_j\rangle + \beta|b_j\rangle, \dots, |a_p\rangle) = \quad (2.88)$$

$$\alpha\mathbf{m}(|a_1\rangle, \dots, |a_j\rangle, \dots, |a_p\rangle) + \beta\mathbf{m}(|a_1\rangle, \dots, |b_j\rangle, \dots, |a_p\rangle) \quad (2.89)$$

从 V 对应到 \mathbb{R} 或 \mathbb{C} 的 p 重线性映射称为 V 中的 p 重线性函数。

例 2.6.1 设 $\{\mathbf{f}_i\}_{i=1}^p$ 为 V 的线性泛函, 那么有 p 重线性映射: $\mathbf{m}(|a_1\rangle, \dots, |a_p\rangle) = \mathbf{f}_1(|a_1\rangle) \cdots \mathbf{f}_p(|a_p\rangle)$, $|a_i\rangle \in V$ 。

定义 2.6.2 一些数学物理方法的回顾:

1. 设 $i < j$ 且 $\sigma(i) > \sigma(j)$, $\forall i, j \in s$, 则 (ij) 或 $(\sigma(i), \sigma(j))$ 称为 σ 的逆序 (inversion)。逆序集是所有逆序的集合, 逆序数是逆序集的基数。
2. 设 σ 为集合 $s = \{1, 2, \dots, p\}$ 上的一个置换 (permutation), 那么 σ 是改变 s 循序的映射, 并且用逆序或对换 (transposition)(ij)

的乘积依次表示对换的顺序。逆序数为偶数的排列称为**偶排列** (*even permutation*); 逆序数为奇数的排列称为**奇排列** (*odd permutation*)。

3. **列维-奇维塔符号** (*Levi-Civita*) ϵ_σ 表现了排列的**反对称性** (*anti-symmetric*), 若排列 σ 为**奇排列**, 则 $\epsilon_\sigma = -1$; 若为**偶排列**, 则 $\epsilon_\sigma = 1$; 其余情况, 也就是如果排列 σ 中**任意两个指标相等**, 则 $\epsilon_\sigma = 0$ 。那么就有定义式如下:

$$\epsilon_\sigma = \prod_{1 \leq i < j \leq p} \text{sgn}(\sigma_j - \sigma_i). \quad (2.90)$$

其中 sgn 为**符号函数** (*sign function*), 定义为:

$$\text{sgn}(x) = \begin{cases} -1, & \text{for } x < 0 \\ 0, & \text{for } x = 0 \\ 1, & \text{for } x > 0 \end{cases} \quad (2.91)$$

4. **反对称**或**斜对称** (*skew-symmetric*) 矩阵是一个转置和自身负数相等的方形矩阵, 也就是满足: $\mathbf{M}^T = -\mathbf{M}$ 。

例 2.6.2 置换和列维-奇维塔符号操作规则的例子:

1. 设 $s = \{a, b, c, d\}$ 且排列 $\sigma(1) = b, \sigma(2) = a, \sigma(3) = d, \sigma(4) = c$ 则, 我们将其写作 $\sigma = (a, b)(c, d)$ 表示我们将原有的 $abcd$ 的顺序通过对换 a, b 然后对换 c, d 得到 $badc$ 。因为 σ 的逆序为二是偶排列, 因此 $\epsilon_\sigma = 1$ 。

2. 设 $s = \{1, 2\}$, 则 ϵ_{ij} 的值组成一个 2×2 的**反对称矩阵**:

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (2.92)$$

反对称性还体现在任意一对指标索引交换顺序都会带来负号, 比如:

$\epsilon_{1432} = -\epsilon_{1234} = -(-\epsilon_{2134}) = -1$, 任意指标重复: $\epsilon_{1434} = \epsilon_{1232} = 0$ 。

由上述例子我们不难看出, ϵ_σ 的取值由排列 σ 的逆序数决定。由于排列以**对换的乘积** (product of transpositions) 所表示, 因此设 π 也是排列 $\sigma\pi$ 的逆序数为 π 和 σ 的逆序数之和, 使得对于任意指标不重复的排列都有 $\epsilon_\sigma\epsilon_\pi = \epsilon_{\sigma\pi}$ 和 $(\epsilon_\sigma)^2 = 1$ 成立。我们还约定符号 $\epsilon_\sigma = \epsilon_{\sigma(1), \dots, \sigma(p)} \equiv \epsilon_{i_1, \dots, i_p} \circ$

定义 2.6.3 若要考虑 p 元组映射 \mathbf{w} 上的一个排列 σ , 则必须使得有 p 重线性映射 $\sigma\mathbf{w}$ 为: $\boxed{\sigma\mathbf{w} = \mathbf{w} \circ \sigma}$, 具体上也就是,

$$\sigma\mathbf{w}(|a_1\rangle, \dots, |a_p\rangle) = \mathbf{w}(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle). \quad (2.93)$$

\mathbf{w} 是从 V 到 U 的**反对称性** p 重线性映射, 如果有 $\boxed{\sigma\mathbf{w} = \epsilon_\sigma \cdot \mathbf{w}}$:

$$\mathbf{w}(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) = \epsilon_\sigma \mathbf{w}(|a_1\rangle, \dots, |a_p\rangle) \quad (2.94)$$

从 V 到 U 的 p 重线性反对称映射集合记作 $\Lambda^p(V, U)$; V 中的 p 重线性反对称映射集合记作 $\Lambda^p(V)$ 。

任意 p 重线性映射都可以被转变为**反对称** p 重线性映射, 设 \mathbf{m} 为一个 p 重线性映射, 并且定义映射 \mathbf{w} 为: $\mathbf{w} \equiv \sum_\pi \epsilon_\pi \cdot \pi \mathbf{m}$ 其中 π 可以是对任意有限大小的排列, 那么 \mathbf{w} 为**反对称** p 重线性映射, 因为:

$$\sigma\mathbf{w} = \sigma \sum_\pi \epsilon_\pi \cdot \pi \mathbf{m} = \sum_\pi \epsilon_\pi \cdot (\sigma\pi) \mathbf{m} = (\epsilon_\sigma)^2 \sum_\pi \epsilon_\pi \cdot (\sigma\pi) \mathbf{m} \quad (2.95)$$

$$= \epsilon_\sigma \sum_\pi (\epsilon_\sigma \epsilon_\pi) \cdot (\sigma\pi) \mathbf{m} = \epsilon_\sigma \sum_{\sigma\pi} (\epsilon_{\sigma\pi}) \cdot (\sigma\pi) \mathbf{m} = \epsilon_\sigma \cdot \mathbf{w}. \quad (2.96)$$

假设 π 是对 $1, 2, \dots, p$ 的排列, 我们约定 \sum_π 为对 ϵ_π 的下标求和:

$$\sum_\pi \equiv \sum_{i_1, i_2, \dots, i_p=1}^p = \sum_{i_1=1}^p \sum_{i_2=1}^p \cdots \sum_{i_p=1}^p = \sum_{1 \leq i_1, i_2, \dots, i_p \leq p}. \quad (2.97)$$

定理 2.6.1 设 $\mathbf{w} \in \Lambda^p(V, U)$ 则有以下三个互相等价的表述:

1. **指标重复:** $\mathbf{w}(|a_1\rangle, \dots, |a_p\rangle) = 0$, 每当 $|a_i\rangle = |a_j\rangle, i \neq j$.
2. **置换完全反对称性:** $\mathbf{w}(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) = \epsilon_{\sigma} \mathbf{w}(|a_1\rangle, \dots, |a_p\rangle)$, 对于任意 $1, 2, \dots, p$ 的排列 σ , 以及 $|a_i\rangle, \dots, |a_p\rangle \in V$.
3. $\mathbf{w}(|a_1\rangle, \dots, |a_p\rangle) = 0$, 每当 $\{|a_i\rangle\}_{k=1}^p$ **线性无关**.

定义 2.6.4 一个值为零的量被称为**消失 (vanish)**。比如, $f(x) = x^3$ 消失在 $x = 0$ 这一量上。

命题 2.6.1 设 $N = \dim V$ 并且 $\mathbf{w} \in \Lambda^N(V, U)$, 也就是反对称 N 重线性映射 \mathbf{w} 可以有 N 个参数, 并且每一参数上的向量 $|v\rangle \in V$ 的**维数**也都等于 N , 那么 \mathbf{w} 是被其自身在 V 上的一组基的值所唯一**决定 (determined)** 的。换句话说, 就是 \mathbf{w} 对于 V 中任意基数为 N 的向量集合, 都能够唯一的被 \mathbf{w} 作用于 V 的一组基所决定。特别是, 当 \mathbf{w} 在一组基上的值**消失**, 则有 $\mathbf{w} = 0$ 。

证明: 设 $\{|e_k\rangle\}_{k=1}^N$ 为 V 的一组基。让 $\{|a_j\rangle\}_{j=1}^N$ 作为 V 中任意向量集合, 并且有 $|a_j\rangle = \sum_{k=1}^N \alpha_{jk} |e_k\rangle$, $j = 1, \dots, N$, 那么就有:

$$\mathbf{w}(|a_1\rangle, \dots, |a_N\rangle) = \sum_{k_1, \dots, k_N=1}^N \alpha_{1k_1} \cdots \alpha_{Nk_N} \mathbf{w}(|e_{k_1}\rangle, \dots, |e_{k_N}\rangle) \quad (2.98)$$

$$\equiv \sum_{\pi} \alpha_{1\pi(1)} \cdots \alpha_{N\pi(N)} \mathbf{w}(|e_{\pi(1)}\rangle, \dots, |e_{\pi(N)}\rangle) \quad (2.99)$$

$$= \left(\sum_{\pi} \epsilon_{\pi} \alpha_{1\pi(1)} \cdots \alpha_{N\pi(N)} \right) \mathbf{w}(|e_1\rangle, \dots, |e_N\rangle). \quad (2.100)$$

注意到括号内每一项都是常数, 这就证明了 \mathbf{w} 在 V 的一组基上的值 $\mathbf{w}(|e_1\rangle, \dots, |e_N\rangle)$ 通过与一个常数 $(\sum_{\pi} \epsilon_{\pi} \alpha_{1\pi(1)} \cdots \alpha_{N\pi(N)})$ 相乘能够唯一地

确定 \mathbf{w} 对于 V 中任意基数为 N 的向量集合的值 $\mathbf{w}(|a_1\rangle, \dots, |a_N\rangle)$ 。

定义 2.6.5 V 中的斜对称 N 重线性函数 $\Delta \in \Lambda^N(V)$ 称作 V 中的**行列式函数** (*determinant function*)。

设 $B = \{e_k\}_{k=1}^N$ 为 V 的基, 并且 $B^* = \{f_j\}_{j=1}^N$ 为对偶空间 V^* 的基。在 V 中任意基数为 N 的向量集合 $\{a_j\}_{j=1}^N$ 上定义 N 重线性函数 \mathbf{m} 为:

$$\mathbf{m}(|a_1\rangle, \dots, |a_N\rangle) = \mathbf{f}_1(|a_1\rangle) \cdots \mathbf{f}_N(|a_N\rangle), \quad (2.101)$$

并且回顾例 2.6.1 和式 2.93 后根据定理 2.5.1 有:

$$\pi \mathbf{m}(|e_1\rangle, \dots, |e_N\rangle) \equiv \mathbf{m}(|e_{\pi(1)}\rangle, \dots, |e_{\pi(N)}\rangle) = \delta_{\iota\pi}. \quad (2.102)$$

其中 ι 是**恒元排列** (identity permutation) 并且如果 $\pi = \iota$ 则 $\delta_{\iota\pi} = 1$, 反之 $\pi \neq \iota$ 则 $\delta_{\iota\pi} = 0$ 。现在我们将 Δ 定义为:

$$\Delta \equiv \sum_{\pi} \epsilon_{\pi} \cdot \pi \mathbf{m} \quad (2.103)$$

根据式 2.96 可得 Δ 是反对称多重线性映射, 并且因为在式 2.101 中我们是在 V 上定义了 N 重线性函数 \mathbf{m} , 因此我们由式 2.103 所定义的 Δ 是一个**行列式函数**。

定理 2.6.2 V 中的基 $B = \{e_k\}_{k=1}^N$ 上行列式 Δ 的值一定等于 1。

证明: 考虑式 2.102 和 2.103 可得:

$$\Delta(|e_1\rangle, \dots, |e_N\rangle) = \sum_{\pi} \epsilon_{\pi} \pi \mathbf{m}(|e_1\rangle, \dots, |e_N\rangle) \quad (2.104)$$

$$= \sum_{\pi} \epsilon_{\pi} \delta_{\iota\pi} = \epsilon_{\iota} = 1. \quad (2.105)$$

上述讨论引申出, 在每个有限维向量空间中, 都存在一个**不恒为零** (not identically zero), 也就是不在 V 的基上**消失**的行列式函数。

注意: 一个函数是对应到标量域 \mathbb{F} 的映射, 因此行列式函数集合准确的来说记作: $\Lambda^N(V, \mathbb{F})$ 。

命题 2.6.2 设 $\mathbf{w} \in \Lambda^N(V, U)$, 并让 Δ 为一个固定的非零行列式函数, 则 \mathbf{w} 可唯一决定一个 $|u_\Delta\rangle \in U$ 使得:

$$\mathbf{w}(|v_1\rangle, \dots, |v_N\rangle) = \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |u_\Delta\rangle. \quad (2.106)$$

证明: 设 $\{e_k\}_{k=1}^N$ 为 V 的一组基, 让 $\{|v_j\rangle\}_{j=1}^N$ 作为 V 中任意向量集合, 由定理 2.6.2 得到 $\Delta(|e_1\rangle, \dots, |e_N\rangle) = 1$ 。这样我们就唯一地确立了在基底上 \mathbf{w} 所决定的 $|u_\Delta\rangle = \mathbf{w}(|e_1\rangle, \dots, |e_N\rangle)$ 。

由命题 2.6.1 可 \mathbf{w} 是被其自身在 V 上的一组基的值所唯一决定的。设 $|v_i\rangle = \sum_{k=1}^N \alpha_{jk} |e_k\rangle$, $j = 1, \dots, N$ 使得有:

$$\mathbf{w}(|v_1\rangle, \dots, |v_N\rangle) = \left(\sum_{\pi} \epsilon_{\pi} \alpha_{1\pi(1)} \cdots \alpha_{N\pi(N)} \right) |u_\Delta\rangle \quad (2.107)$$

回顾式 2.101 中我们对 N 重线性映射 \mathbf{m} 定义, 其中的线性泛函 \mathbf{f}_j 我们采用类似例 2.5.1 中的第一个小例 1. $\mathbf{f}(|a_j\rangle) = \sum_{k=1}^N \alpha_{jk}$ 从而得到:

$$\sum_{\pi} \epsilon_{\pi} \alpha_{1\pi(1)} \cdots \alpha_{N\pi(N)} = \sum_{\pi} \epsilon_{\pi} \mathbf{f}_1(|a_{\pi(1)}\rangle) \cdots \mathbf{f}_N(|a_{\pi(N)}\rangle) \quad (2.108)$$

$$= \sum_{\pi} \epsilon_{\pi} \mathbf{m}(|a_{\pi(1)}\rangle, \dots, |a_{\pi(N)}\rangle) \quad (2.109)$$

$$= \sum_{\pi} \epsilon_{\pi} \pi \mathbf{m}(|a_1\rangle, \dots, |a_N\rangle) \quad (2.110)$$

$$= \Delta(|v_1\rangle, \dots, |v_N\rangle) \quad (2.111)$$

推论 2.6.1 每个 V 中固定的非零行列式函数都是 Δ 的标量倍数。也就是，任意两个非零行列式函数都是互相成比例的。

证明：将命题 2.6.2 中的 U 设置为 \mathbb{F} 使得 $|u_\Delta\rangle \in \mathbb{F}$ 。♣

命题 2.6.3 设 Δ 为 N 维向量空间 V 的一个行列式函数，并且 $|v\rangle$ 和 $\{|v_k\rangle\}_{k=1}^N$ 都是 V 中的向量，那么有：

$$\sum_{j=1}^N (-1)^{j-1} \Delta(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) |v_j\rangle = \Delta(|v_1\rangle, \dots, |v_N\rangle) |v\rangle \quad (2.112)$$

其中 $\widehat{|v_j\rangle}$ 上的帽表示该向量**不存在**于此。

证明：注意到等式左侧为 $\{|v_k\rangle\}_{k=1}^N$ 的线性组合，

推论 2.6.2 根据定义 2.1.2 可得，若一向量组线性相关，则在其中加上任意个向量后，仍然保持线性相关。

因此如果 $\{|v_k\rangle\}_{k=1}^N$ 是线性相关的，则等式两边行列式函数中的向量都是线性相关的。同时，根据定理 2.6.1 第三条有等式两侧为 0。如果 $\{|v_k\rangle\}_{k=1}^N$ 是线性无关的，由于其基数等于 V 的维数，因此它们可以生成线性空间 V 并被看作 V 的一组基，使得 $|v\rangle$ 允许为 $\{|v_k\rangle\}_{k=1}^N$ 的线性组合 $|v\rangle = \sum_{i=1}^N \alpha_i |v_i\rangle$ 。由定义 2.6.5 的行列式函数的线性性质可得：

$$\begin{aligned} \Delta\left(\sum_{i=1}^N \alpha_i |v_i\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle\right) &= \sum_{i=1}^N \alpha_i \Delta(|v_i\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) \\ &= \sum_{i=1}^N \sum_{j=1}^N (-1)^{j-1} \alpha_i \Delta(|v_i\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) |v_j\rangle \end{aligned} \quad (2.113)$$

我们注意到 $\Delta(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) = 0$ 除非有 $i = j$ 使得:

$$\sum_{j=1}^N (-1)^{j-1} \alpha_j \Delta(|v_j\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) |v_j\rangle \quad (2.114)$$

$$= \sum_{j=1}^N \alpha_j |v_j\rangle (-1)^{j-1} \Delta(|v_j\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) \quad (2.115)$$

$$= \Delta(|v_1\rangle, \dots, |v_N\rangle) |v\rangle \quad (2.116)$$

注意到式子中的 $(-1)^{j-1}$ 是因为行列式函数定义中列维-奇维塔符号的完全反对称性, 使得每次在一个置换中交换任意位置的循序都会带来**负号**。也就是当 $j = 1$ 时不对换, $(-1)^0 = 1$, 当 $j = 2$ 时对换 $|v_j\rangle$ 和 $|v_{j-1}\rangle$ 得到 $-\Delta(|v_1\rangle, |v_j\rangle, \dots, |v_N\rangle)$, 当 $j = 3$ 时对换 $|v_j\rangle$ 和 $|v_{j-1}\rangle$ 得到 $\Delta(|v_1\rangle, |v_2\rangle, |v_j\rangle, \dots, |v_N\rangle)$, 以此类推。♣

2.6.1 线性算子的行列式

设 \mathbf{A} 为 N 维向量空间 V 的一个线性算子, 给定一个非零行列式函数 Δ 。在 v 中一组基 $\{|v_k\rangle\}_{k=1}^N$ 上定义函数 $\Delta_{\mathbf{A}}$ 为:

$$\Delta_{\mathbf{A}}(|v_1\rangle, \dots, |v_N\rangle) \equiv \Delta(\mathbf{A}|v_1\rangle, \dots, \mathbf{A}|v_N\rangle). \quad (2.117)$$

从等式右侧看出 $\Delta_{\mathbf{A}}$ 也是**行列式函数**, 根据推论 2.6.1, $\Delta_{\mathbf{A}}$ 和 Δ 成比例 $\Delta_{\mathbf{A}} = \alpha \Delta$ 。此外, $\Delta_{\mathbf{A}}$ 与其所选的非零行列式函数无关, 因为如果 $\Delta' = \lambda \Delta$ 是另一个非零行列式函数则有:

$$\Delta'_{\mathbf{A}} = \lambda \Delta_{\mathbf{A}} = \lambda \alpha \Delta = \alpha \Delta' \quad (2.118)$$

式 2.114 还意味着, 常数 α 仅只被 \mathbf{A} 所决定, 和其他非零行列式函数以及基底的选取是无关紧要的。

定义 2.6.6 设 $\mathbf{A} \in \text{End}(V)$, 并且 Δ 为 V 中的一个非零行列式函数, 根据式 2.113 中对 $\Delta_{\mathbf{A}}$ 的定义, 那么就有 \mathbf{A} 的**行列式 (determinant)** 记作 $\det \mathbf{A}$ 定义为:

$$\Delta_{\mathbf{A}} = \det \mathbf{A} \cdot \Delta \quad (2.119)$$

定理 2.6.3 线性算子 \mathbf{A} 的行列式满足性质:

1. 如果 $\mathbf{A} = \lambda \mathbf{1}$, 那么 $\det \mathbf{A} = \lambda^N$ 。
2. \mathbf{A} 可逆, 当且仅当 $\det \neq 0$ 。
3. $\det(\mathbf{A} \circ \mathbf{B}) = \det \mathbf{A} \det \mathbf{B}$ 。

2.6.2 经典伴随 (Classical Adjoint)

命题 2.6.4 设 Δ 为 N 维向量空间 V 的一个**行列式函数**, 并且 $\mathbf{A} \in \text{End}(V)$ 。设 $|v\rangle$ 和 $\{|v_k\rangle\}_{k=1}^N$ 都是 V 中的向量, 那么可以定义 $\mathbf{O} : V^N \rightarrow \text{End}(V)$ 为:

$$\mathbf{O}(|v_1\rangle, \dots, |v_N\rangle)|v\rangle \quad (2.120)$$

$$= \sum_{j=1}^N (-1)^{j-1} \Delta(|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle)|v_j\rangle \quad (2.121)$$

该多重线性映射 \mathbf{O} 是**反对称**的。

定义 2.6.7 由命题 2.6.2 可得, 对于该**斜对称**多重线性映射 \mathbf{O} 一定唯一地存在一个**线性算子**记作 $\text{adj}(\mathbf{A}) \in \text{End}(V)$ 使得:

$$\mathbf{O}(|v_1\rangle, \dots, |v_N\rangle) = \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{adj}(\mathbf{A}). \quad (2.122)$$

我们将 $\text{adj}(\mathbf{A})$ 称作 \mathbf{A} 的**经典伴随** (*classical adjoint*) 并且该等式表明经典伴随与行列式函数的选择是无关的。

命题 2.6.5 经典伴随满足以下关系式：

$$\text{adj}(\mathbf{A}) \circ \mathbf{A} = \det \mathbf{A} \cdot \mathbf{1} = \mathbf{A} \circ \text{adj}(\mathbf{A}) \quad (2.123)$$

证明：结合式 2.117 和 2.118 并且将 $|v\rangle$ 替换为 $\mathbf{A}|v\rangle$ 可得：

$$\Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{adj}(\mathbf{A}) \circ \mathbf{A}|v\rangle \quad (2.124)$$

$$= \sum_{j=1}^N (-1)^{j-1} \Delta(\mathbf{A}|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle) |v_j\rangle \quad (2.125)$$

同时根据命题 2.6.3 可得等式右侧 2.121 为：

$$\text{RHS} = \det \mathbf{A} \sum_{j=1}^N (-1)^{j-1} \Delta(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) |v_j\rangle \quad (2.126)$$

$$= \det \mathbf{A} \cdot \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle \quad (2.127)$$

与上述步骤平行的我们接下来，结合式 2.117 和 2.118 并且直接对其施加 \mathbf{A} ，根据命题 2.6.3 可得：

$$\Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \mathbf{A} \circ \text{adj}(\mathbf{A})|v\rangle \quad (2.128)$$

$$= \sum_{j=1}^N (-1)^{j-1} \Delta(|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle) \mathbf{A}|v_j\rangle \quad (2.129)$$

$$= \Delta(\mathbf{A}|v_1\rangle, \dots, \mathbf{A}|v_N\rangle) = \det \mathbf{A} \cdot \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle \quad (2.130)$$

鉴于式 2.123 和 2.126 是完全一样的，我们就可以得到：

$$\begin{aligned}
 \det \mathbf{A} \cdot \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle &= \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{adj}(\mathbf{A}) \circ \mathbf{A}|v\rangle \\
 &= \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \mathbf{A} \circ \text{adj}(\mathbf{A})|v\rangle
 \end{aligned}
 \tag{2.131}$$

推论 2.6.3 如果 $\det \mathbf{A} \neq 0$, 那么 \mathbf{A} 是可逆的, 并且有:

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \cdot \text{adj}(\mathbf{A}) \tag{2.132}$$

2.7 习题

习题 2.1: 设 \mathbb{R}^+ 为正实数的集合。定义 \mathbb{R}^+ 的两个元素的“和”为它们通常的乘积, 并定义**标量乘法**为 $r \cdot p = p^r$, 其中 $r \in \mathbb{R}, p \in \mathbb{R}^+$ 。证明 \mathbb{R}^+ 是实数域 \mathbb{R} 上的一个向量空间。

习题 2.2: 证明两个子空间的交集也是子空间。

习题 2.3: 判断下列 \mathbb{R}^3 的子集是否也是 \mathbb{R}^3 的子空间:

$$(a) \{(x, y, z) \in \mathbb{R}^3 | x + y - 2z = 0\}; \tag{2.133}$$

$$(b) \{(x, y, z) \in \mathbb{R}^3 | x + y - 2z = 3\}; \tag{2.134}$$

$$(c) \{(x, y, z) \in \mathbb{R}^3 | xyz = 0\}. \tag{2.135}$$

习题 2.4: 证明一个向量在给定一组基上的每一个分量是唯一的。

习题 2.5: 表明下列向量是 $\mathbb{C}^n \vee \mathbb{R}^n$ 的一组基。

$$|a_1\rangle = (1, 1, \dots, 1, 1), |a_2\rangle = (1, 1, \dots, 1, 0), \dots, |a_n\rangle = (1, 0, \dots, 0, 0)$$

第三章 代数 (Algebra)

在许多物理应用中, 向量空间 V 有一个自然的“乘积”, 即二元运算 $V \times V \rightarrow V$ 。这种向量空间的主要例子是矩阵的向量空间。因此, 讨论这种向量空间上的乘积是很有用的。

3.1 从向量空间到代数

我们在这部分中将会定义代数, 学习例子, 并讨论代数的基本性质。在这一章我们放弃使用狄拉克符号, 而是用粗体英文字母 $\mathbf{a}, \mathbf{b}, \dots$ 表示向量。

定义 3.1.1 数域 \mathbb{F} 上的一个**代数 (algebra)** A 是在数域 \mathbb{F} 上的一个向量空间, 并且有二元运算 $f: A \times A \rightarrow A$ 称作**乘法 (multiplication)**。设向量 $\mathbf{a}, \mathbf{b} \in A$, 并且 $(\mathbf{a}, \mathbf{b}) \in A \times A$ 的像记作 $\mathbf{ab} \in A$ 也就是 $f(\mathbf{a}, \mathbf{b}) = \mathbf{ab}$, 并且它满足以下双线性关系:

$$\mathbf{a}(\beta\mathbf{b} + \gamma\mathbf{c}) = \beta\mathbf{ab} + \gamma\mathbf{ac} \quad (3.1)$$

$$(\beta\mathbf{b} + \gamma\mathbf{c})\mathbf{a} = \beta\mathbf{ba} + \gamma\mathbf{ca} \quad (3.2)$$

$$\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in A; \beta, \gamma \in \mathbb{F}. \quad (3.3)$$

该向量空间 A 的维数称作**代数维数**。如果乘积满足: $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$,

则称其为**结合** (*associative*) 代数。如果乘积满足: $ab = ba$, 则称其为**交换** (*commutative*) 代数。具有乘法**恒元** (*identity*) 的代数满足 $a1 = 1a = a$ 。若代数具有恒元且有 $ab = 1$, 则 a 是 b 的**左逆** (*left inverse*); b 是 a 的**右逆** (*right inverse*)。恒元又称**单位元** (*unit*); 具有恒元的代数又称**含么代数** (*unital algebra*)。

我们有时需要使用不同的符号来表示代数的**恒元**, 特别是当我们在同一时间讨论多个代数的时候。因此, 除了 1 之外, 还有一种常见的符号是 e 被用作表示恒元。“么”指“一”读作“妖”。

3.1.1 广义性质 (General Properties)

在上述定义中取 $\beta = 1 = -\gamma$ 和 $b = c$ 直接可得:

$$a0 = 0a = 0, \forall a \in A. \quad (3.4)$$

一个代数的恒元是**唯一**的, 如果有两个恒元 $1, e$, 那么 $1e = e$ 因为 1 为恒元, 并且因为 e 是恒元, 所以 $1e = 1$ 。如果 A 是一个**结合代数**并且 $a \in A$ 的**左逆**为 b 且**右逆**为 c , 那么就有 $b = c$, 因为:

$$bac = (ba)c = 1c = c, \quad (3.5)$$

$$bac = b(ac) = b1 = b. \quad (3.6)$$

因此, 在结合代数的讨论中既然左逆和右逆是相等的, 因此我们没有必要区分一个逆的左或右, 这也意味着这个逆是**唯一**的。

定理 3.1.1 设 A 为一个有恒元的结合代数, 则如果 $a \in A$ 有一个左逆和右逆, 那么它们彼此**相等**并且这一个逆是**唯一**的, 我们将其记作 a^{-1} 。如果 $a, b \in A$ 都是可逆的, 那么 ab 也是可逆的, 并且有:

$$(\mathbf{ab})^{-1} = \mathbf{b}^{-1}\mathbf{a}^{-1}. \quad (3.7)$$

证明: $(\mathbf{ab})^{-1}\mathbf{ab} = \mathbf{ab}(\mathbf{ab})^{-1} = \mathbf{b}^{-1}(\mathbf{a}^{-1}\mathbf{a})\mathbf{b} = \mathbf{a}(\mathbf{bb}^{-1})\mathbf{a}^{-1} = \mathbf{1}$. ♣

定义 3.1.2 设 A 为一个代数, 如果 A' 是 A 的线性子空间并且是在乘法封闭下的, 也就是对于任意的 $\mathbf{a} \in A'$ 和 $\mathbf{b} \in A'$ 都有 $\mathbf{ab} \in A'$, 那么 A' 就是 A 的**子代数** (*subalgebra*)。

显然, 一个结合 (交换) 代数的**子代数**也是结合 (交换) 的。

定义 3.1.3 设 A 为一个结合代数并且 S 为 A 的一个子集, 由 S **生成** (*generate*) 的**子代数** A_S 是 $\mathbf{s}_1\mathbf{s}_2\cdots\mathbf{s}_k$, $\mathbf{s}_i \in S$ 全部线性组合的集合:

$$A_S \equiv \left\{ \sum_{i=1}^n \alpha_i \prod_{j=1}^m \mathbf{s}_{\sigma_i(j)} \mid \forall n, m \in \mathbb{N}, \alpha_i \in \mathbb{F}, \mathbf{s}_{\sigma_i(j)} \in S \right\} \quad (3.8)$$

其中 σ_i 是 S 的任意一个数组 (序列)。

例 3.1.1 如果 S 为 \mathbf{s} 的单元素集合, 则由 \mathbf{s} 产生的**子代数**为 \mathbf{s} 的所有多项式的集合, 其中的元素比如: $\alpha_{11}\mathbf{s} + \alpha_{12}\mathbf{s}^3, \alpha_{21}\mathbf{s}, \cdots$ 。如果 S 为 $\mathbf{s}_1, \mathbf{s}_2$, 其中的元素比如: $\alpha_{11}\mathbf{s}_1\mathbf{s}_2\mathbf{s}_1\mathbf{s}_2^2\mathbf{s}_1 + \alpha_{12}\mathbf{s}_1\mathbf{s}_2\mathbf{s}_1, \cdots$, 其中 $\mathbf{s}_1\mathbf{s}_2\mathbf{s}_1$ 所对应的数组为 $\sigma = [1, 2, 1]$ 。

命题 3.1.1 \mathbb{F} 是任意**含幺代数**的**子代数**。

例 3.1.2 证明: 设 A 为一个含幺代数, 则由恒元 $\mathbf{1}$ 张成的向量空间为:

$$\text{Span}\{\mathbf{1}\} = \{\alpha\mathbf{1} \mid \alpha \in \mathbb{F}\} \quad (3.9)$$

注意到 $\text{Span}\{1\}$ 和数域 \mathbb{F} 没有任何的区别, 同样它也是由 A 的一个子集 $\{1\}$ 生成的 A 的**子代数**。♣

定义 3.1.4 设 A 为一个代数, 与 A 中任意元素都可**交换** (*commute*) 并属于 A 的元素所组成的集合, 我们称其为 A 的**中心** (*center*) 记作

$$\mathcal{Z}(A) = \{z \in A | za = az, \forall a \in A\} \quad (3.10)$$

命题 3.1.2 由代数的**双线性**可得, $\mathcal{Z}(A)$ 为 A 的**子空间**, 并且若 A 是**结合代数**, 则 $\mathcal{Z}(A)$ 为 A 的**子代数**。

定义 3.1.5 含幺代数 A 被称为**中心代数** (*central algebra*), 如果有:

$$\mathcal{Z}(A) = \text{Span}\{1\} \quad (3.11)$$

设含幺代数 (向量空间) S 的基底为 $\{e_i\}_{i=0}^3$ 并且给定乘法表 3.1, 其中 e_0 为恒元。

例 3.1.3 中心代数的例子:

为了找到属于 S 的中心的元素, 首先设 $a \in \mathcal{Z}(S)$, 那么对于任意的元素 $b \in S$, 我们必须满足 $ab = ba$ 。接下来, 我们用基底来表示 $a = \sum_i^3 \alpha_i e_i$ 和 $b = \sum_i^3 \beta_i e_i$ 。那么就有: $ab = (\alpha_0\beta_0 + \alpha_1\beta_1 - \alpha_2\beta_2 + \alpha_3\beta_3)e_0 + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_0\beta_2 + \alpha_1\beta_3 + \alpha_2\beta_0 - \alpha_3\beta_1)e_2 + (\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)e_3$ 与之相似地, 我们也能得到 ba 表达式。然后, 经过简单计算我们就能得到, 有 $ab = ba$ 成立, 当且仅当 $\alpha_2\beta_3 = \alpha_3\beta_2$ 和 $\alpha_1\beta_3 = \alpha_3\beta_1$ 以及 $\alpha_1\beta_2 = \alpha_2\beta_1$ 成立。对于任意的 b 上述三个等式都成立, 当且仅当有 $\alpha_1 = \alpha_2 = \alpha_3 = 0$ 成立且 α_0 可任

	e_0	e_1	e_2	e_3
e_0	e_0	e_1	e_2	e_3
e_1	e_1	e_0	e_3	e_2
e_2	e_2	$-e_3$	$-e_0$	e_1
e_3	e_3	$-e_2$	$-e_1$	e_0

表 3.1: S 的乘法表

取。因此, $\mathbf{a} \in \mathcal{Z}(S)$ 当且仅当 \mathbf{a} 为恒元 \mathbf{e}_0 的倍数, 也就是 $\mathcal{Z}(S) = \text{Span}\{\mathbf{e}_0\}$ 。由定义 3.1.5 可知, S 为**中心代数**。

定义 3.1.6 设 A 和 B 为代数 C 的子集, 那么我们将 AB 定义为: 任意元素 $\mathbf{x} \in C$ 并且以写成 A 中元素与 B 中元素乘积的和:

$$AB \equiv \left\{ \mathbf{x} \in C \mid \mathbf{x} = \sum_{i=1}^n \mathbf{a}_i \mathbf{b}_i, \forall n \in \mathbb{N}, \mathbf{a}_i \in A, \mathbf{b}_i \in B \right\} \quad (3.12)$$

该定义下的一个特殊情况称作 C 的**导出代数 (derived algebra)** 记作:

$$C^2 \equiv \left\{ \mathbf{x} \in C \mid \mathbf{x} = \sum_{i=1}^n \mathbf{a}_i \mathbf{b}_i, \forall n \in \mathbb{N}; \mathbf{a}_i, \mathbf{b}_i \in C \right\} \quad (3.13)$$

定义 3.1.7 设代数 A 的“乘法”二元运算为 $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{ab}$, 那么就有另一个同线性空间不同乘法规则 $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{ba}$ 的一个 A 的**反代数 (algebra opposite to A)** 记作 A^{op} , 并且有记法 $(\mathbf{ab})^{op} = \mathbf{ba}$ 。

若 A 为**结合代数**, 则 A^{op} 也同样是; 若 A 为**交换代数**, 则有 $A = A^{op}$ 。

例 3.1.4 代数的一些例子:

1. 在 \mathbb{R}^2 上定义乘积规则: $(x_1, x_2)(y_1, y_2) = (x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1)$ 可使得 \mathbb{R}^2 成为**交换代数**。在 \mathbb{R}^3 上定义以**叉乘**为乘积规则, 可使得 \mathbb{R}^3 成为**非交换**和**非结合**的代数。
2. 所有代数的范式都是**矩阵代数 (matrix algebra)**, 其二元运算是 $n \times n$ 矩阵的普通乘法, 该代数是**结合的**, 但**不是交换**的。
3. 设代数 S 为 $n \times n$ 矩阵的集合, 并且在其之上定义二元运算 \bullet 为: $A \bullet B = AB - BA$ 其中 A, B 为 $n \times n$ 矩阵, AB, BA 为矩阵的普通乘法, 这样的代数既不**结合**也不**交换**。

定义 3.1.8 设 A 和 B 是代数 (向量空间), A 和 B 向量上的直和 $A \oplus B$ 称为**代数直和** (*algebra direct sum*), 若 $A \oplus B$ 上有如下乘积定义:

$$(\mathbf{a}_1 \oplus \mathbf{b}_1)(\mathbf{a}_2 \oplus \mathbf{b}_2) = (\mathbf{a}_1 \mathbf{a}_2 \oplus \mathbf{b}_1 \mathbf{b}_2) \quad (3.14)$$

注意到对于元素 $\mathbf{a} \in A$, 我们可以将其表示为 $A \oplus B$ 中的一个元素 $\mathbf{a} \oplus \mathbf{0} \in A \oplus B$, 同样元素 $\mathbf{b} \in B$ 也是 $\mathbf{0} \oplus \mathbf{b} \in A \oplus B$. 后续我们会看到, 如果给定的代数是其子代数们的直和, 那么这个条件是**必要的**. 因此有如下命题:

命题 3.1.3 任意 A 中元素与 B 中任意元素的 (代数直和) **乘积是零**:

$$AB = BA = \{\mathbf{0}\} \quad (3.15)$$

根据定义 3.1.4, $\mathbf{a} \oplus \mathbf{b}$ 在 $A \oplus B$ 的**中心**里当且仅当:

$$(\mathbf{a} \oplus \mathbf{b})(\mathbf{x} \oplus \mathbf{y}) = (\mathbf{x} \oplus \mathbf{y})(\mathbf{a} \oplus \mathbf{b}) \Rightarrow \mathbf{ax} \oplus \mathbf{by} = \mathbf{xa} \oplus \mathbf{yb} \quad (3.16)$$

$$(\mathbf{ax} - \mathbf{xa}) \oplus (\mathbf{by} - \mathbf{yb}) = \mathbf{0}, \forall \mathbf{x} \in A, \mathbf{y} \in B \Leftrightarrow \quad (3.17)$$

$$\mathbf{ax} - \mathbf{xa} = \mathbf{by} - \mathbf{yb} = \mathbf{0} \Rightarrow \mathbf{a} \in \mathcal{Z}(A), \mathbf{b} \in \mathcal{Z}(B) \quad (3.18)$$

命题 3.1.4 A 和 B 的代数直和的**中心**等于 A 与 B 的**中心**的直和。

$$\mathcal{Z}(A \oplus B) = \mathcal{Z}(A) \oplus \mathcal{Z}(B) \quad (3.19)$$

定义 3.1.9 设 A 和 B 是代数, 那么向量空间上的张量积 $A \otimes B$ 就被称为一个**代数张量积** (*algebra tensor product*), 如果有 $A \otimes B$ 上的乘积规则定义:

$$(\mathbf{a}_1 \otimes \mathbf{b}_1)(\mathbf{a}_2 \otimes \mathbf{b}_2) = (\mathbf{a}_1 \mathbf{a}_2 \otimes \mathbf{b}_1 \mathbf{b}_2) \quad (3.20)$$

根据定理 2.1.12 和 2.3.12 可得同构 $A \otimes B \cong B \otimes A$, 因此我们需要有 $\mathbf{a} \otimes \mathbf{b} = \mathbf{b} \otimes \mathbf{a}, \forall \mathbf{a} \in A, \mathbf{b} \in B$. 这个条件变得很重要, 当我们把一个代数 A 写作它的两个子代数 B 和 C 的代数张量积时。在这种情况下, \otimes 与 A 中的乘法重合, 该条件变为 B 的所有元素与 C 的所有元素交换, 即 $BC = CB$.

定义 3.1.10 给定一个代数 A 和其一组基 $B = \{\mathbf{e}_i\}_{i=1}^n$ 可得:

$$\mathbf{e}_i \mathbf{e}_j = \sum_{k=1}^N c_{ij}^k \mathbf{e}_k, \quad \forall c_{ij}^k \in \mathbb{F} \quad (3.21)$$

在这个式子中, 表示基底 B 中的向量 $\mathbf{e}_i \mathbf{e}_j$ 的分量 c_{ij}^k 为 A 的**结构常数** (*structure constants*).

结构常数通过给出基底 B 中基向量的乘积而决定了 A 中任意两个向量的乘积规则。反过来我们也可给定一个 N 维度向量空间 V , 选择一组基底 B 和基数为 N^3 的结构常数集合 $\{c_{ij}^k\}_{i,j,k=1}^N$, 最后根据式 3.21 定义基向量上的乘积规则。

例 3.1.5 设代数 A 和 B 的代数结构分别为 $\{a_{ij}^k\}_{i,j,k=1}^M$ 和 $\{b_{mn}^l\}_{l,m,n=1}^M$, 它们对应的基底分别为 $\{\mathbf{e}_i\}_{i=1}^M$ 和 $\{\mathbf{f}_i\}_{i=1}^M$ 使得有 $\mathbf{e}_i \mathbf{e}_j = \sum_{k=1}^M a_{ij}^k \mathbf{e}_k$ 和 $\mathbf{f}_m \mathbf{f}_n = \sum_{l=1}^M b_{mn}^l \mathbf{f}_l$, 接下来定义基底 $\{\mathbf{v}_{kl}\}_{k,l=1}^{M,N}$ 上的结构常数 $c_{im,jn}^{kl} = a_{ij}^k b_{mn}^l$ 构造 MN 维度代数 C , 使得有

$$\mathbf{v}_{im} \mathbf{v}_{jn} = \sum_{i,j=1}^M \sum_{m,n=1}^N c_{im,jn}^{kl} \mathbf{v}_{kl} \quad (3.22)$$

由定理 1.2.12 可得 C 同构于代数 $A \otimes B$, 若设 $\mathbf{v}_{kl} \equiv \mathbf{e}_k \otimes \mathbf{f}_l$, 那么就有

$$\mathbf{v}_{im}\mathbf{v}_{jn} = \sum_{i,j=1}^M \sum_{m,n=1}^N c_{im,jn}^{kl} \mathbf{e}_k \otimes \mathbf{f}_l = \sum_{i,j=1}^M \sum_{m,n=1}^N a_{ij}^k b_{mn}^l \mathbf{e}_k \otimes \mathbf{f}_l \quad (3.23)$$

$$= \left(\sum_{i,j=1}^M a_{ij}^k \mathbf{e}_k \right) \otimes \left(\sum_{m,n=1}^N b_{mn}^l \mathbf{f}_l \right) = (\mathbf{e}_i \mathbf{e}_j) \otimes (\mathbf{f}_m \mathbf{f}_n) \quad (3.24)$$

使得该结果符合设定 $\mathbf{v}_{im} \equiv \mathbf{e}_i \otimes \mathbf{f}_m$ 和 $\mathbf{v}_{jn} \equiv \mathbf{e}_j \otimes \mathbf{f}_n$ ，并且与式 3.20 中的乘积规则保持了一致性。

定义 3.1.11 如果一个含么代数中的所有非零元素都有逆，那么我们称为其为**除法代数** (division algebra)。

例 3.1.6 设 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 为 \mathbb{R}^2 的一组基，其乘积为 $\mathbf{e}_1^2 = -\mathbf{e}_2^2 = \mathbf{e}_1$ ， $\mathbf{e}_1\mathbf{e}_2 = \mathbf{e}_2\mathbf{e}_1 = \mathbf{e}_2$ 且有**结构常数**： $c_{11}^1 = -c_{22}^1 = c_{12}^2 = c_{21}^2 = 1$ 和 $c_{12}^1 = -c_{21}^1 = c_{11}^2 = c_{22}^2 = 0$ 。我们将 \mathbf{e}_1 认定为 1；将 \mathbf{e}_2 认定为 $\sqrt{-1}$ ，那么很轻易就能看出，这样构造的代数就是 \mathbb{C} 。由于任意 \mathbb{C} 中的非零元素 α 都唯一存在一个 α^{-1} 使得有 $\alpha\alpha^{-1} = 1$ ，因此 \mathbb{C} 也是一个**除法代数**。

例 3.1.7 四元数代数：

在 \mathbb{R}^4 的标准基 $\{\mathbf{e}_i\}_{i=0}^3$ 中，选择结构常数如下：

$$\mathbf{e}_0^2 = -\mathbf{e}_1^2 = -\mathbf{e}_2^2 = -\mathbf{e}_3^2 = \mathbf{e}_0 \quad (3.25)$$

$$\mathbf{e}_0\mathbf{e}_i = \mathbf{e}_i\mathbf{e}_0 = \mathbf{e}_i, \quad \forall i = 1, 2, 3 \quad (3.26)$$

$$\mathbf{e}_i\mathbf{e}_j = \sum_{k=1}^3 \epsilon_{ijk} \mathbf{e}_k, \quad \forall i, j = 1, 2, 3, \quad i \neq j \quad (3.27)$$

其中 ϵ_{ijk} 在它的所有指标上都是完全反对称的并且 $\epsilon_{123} = 1$ 。这样的

结构使得 \mathbb{R}^4 转变为一个结合性但非交换性的代数。这个代数称为**四元数** (quaternions) 代数并记作 \mathbb{H} 。这里我们通常将 \mathbf{e}_0 记作 1；将 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 分别记作 i, j, k ；将 \mathbb{H} 其中一个元素写作 $q = x + iy + jz + kw$ 。

实际上四元数 \mathbb{H} 是对复数域 \mathbb{C} 的一个广义推广，和复数域类似地我们将 x 称作四元数 q 的实部，将 (y, z, w) 称作四元数 q 的**纯部** (pure part)，以及 q 的共轭 $q^* = x - iy - jz - kw$ 。方便起见我们将 q 写作 $q = x_0 + \mathbf{x}$ ，其中 \mathbf{x} 为三维向量。设 $q = x_0 + \mathbf{x}$ 和 $p = y_0 + \mathbf{y}$ ，那么就有

$$qp = x_0y_0 - \mathbf{x} \cdot \mathbf{y} + (x_0\mathbf{y} + y_0\mathbf{x} + \mathbf{x} \times \mathbf{y}) \quad (3.28)$$

$$q^*p^* = x_0y_0 - \mathbf{x} \cdot \mathbf{y} + (-x_0\mathbf{y} - y_0\mathbf{x} + \mathbf{x} \times \mathbf{y}) \quad (3.29)$$

注意：其中括号中的部分为 qp 的纯部，其余为实部。这里我们分别改变 \mathbf{x}, \mathbf{y} 成为 $-\mathbf{x}, -\mathbf{y}$ 得到的是 q^*p^* ，这并不是 qp 的共轭 $(qp)^* \neq q^*p^*$ ，这里我们直接变更括号前的符号就能得到共轭，因为 $\mathbf{y} \times \mathbf{x} = -\mathbf{x} \times \mathbf{y}$ 其余交换不变，因此有 $(qp)^* = p^*q^*$ 。此外与复数类似 q 的**绝对值** (absolute value) $|q| = \sqrt{qq^*} = \sqrt{x_0^2 + |\mathbf{x}|^2}$ ，如果 $q \neq 0$ ，那么 q 的逆为 $q^{-1} = q^*/(x_0^2 + |\mathbf{x}|^2)$ 。这说明四元数代数也是一个除法代数。此外，还不难验证如下等式：

$$q^n = \sum_{k=0}^{[n]/2} (-1)^k \binom{n}{2k} x_0^{n-2k} |\mathbf{x}|^{2k} + \sum_{k=0}^{[n]/2} (-1)^k \binom{n}{2k+1} x_0^{n-2k-1} |\mathbf{x}|^{2k} \mathbf{x} \quad (3.30)$$

其中若 n 是偶数，则 $[n] = n$ ，若 n 为奇数，则 $[n] = n - 1$ 。为了使 $\mathbf{a} \otimes \mathbf{b}$ 在 $A \otimes B$ 的中心之中，那么有

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{x} \otimes \mathbf{y}) = (\mathbf{x} \otimes \mathbf{y})(\mathbf{a} \otimes \mathbf{b}) \Rightarrow \mathbf{ax} \otimes \mathbf{by} = \mathbf{xa} \otimes \mathbf{yb} \quad (3.31)$$

$$(\mathbf{ax} - \mathbf{xa}) \otimes (\mathbf{by} - \mathbf{yb}) = \mathbf{0}, \forall \mathbf{x} \in A, \mathbf{y} \in B \Leftrightarrow \quad (3.32)$$

$$\mathbf{ax} - \mathbf{xa} = \mathbf{by} - \mathbf{yb} = \mathbf{0} \Rightarrow \mathbf{a} \in \mathcal{Z}(A), \mathbf{b} \in \mathcal{Z}(B) \quad (3.33)$$

命题 3.1.5 A 和 B 代数张量积的**中心**等于 A 与 B 的**中心**的张量积。

$$\mathcal{Z}(A \otimes B) = \mathcal{Z}(A) \otimes \mathcal{Z}(B) \quad (3.34)$$

设 A 是一个结合代数, 那么 A 的子集被称作代数 A 的**生成元** (generator), 如果 A 中的每个元素都可以表示为 S 中元素乘积的线性组合, 例如向量空间 A 的一组基显然是 A 的一个生成元。然而, 它不是最小的一个生成元, 因为我们有可能通过基向量的子集之间的乘法来得到整个基向量。比如, 叉乘下的代数 \mathbb{R}^3 的基地为 $\{\hat{e}_x, \hat{e}_y, \hat{e}_z\}$, 但是任意一对单位向量如 $\{\hat{e}_x, \hat{e}_y\}$ 都是一个生成元, 因为有 $\hat{e}_z = \hat{e}_x \times \hat{e}_y$ 。

3.1.2 同态 (Homomorphisms)

定义 3.1.12 设 A 和 B 为代数, 被称作**代数同态** (algebra homomorphism) 的线性映射为

$$f: A \rightarrow B; f(\mathbf{ab}) = f(\mathbf{a})f(\mathbf{b}), \mathbf{a} \in A, \mathbf{b} \in B \quad (3.35)$$

单射, 满射, 和双射的代数同态分别称为**单态射** (monomorphism), **满态射** (epimorphism), 和**代数同构** (isomorphism)。将一个代数对应到其自身的代数同构被称作**自同构** (automorphism)。

例 3.1.8 设 $A \equiv \mathbb{R}^3$, 映射 $f: A \rightarrow B$ 的定义为,

$$f(\mathbf{a}) = f(a_1, a_2, a_3) = \begin{pmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_1 \\ a_2 & -a_3 & 0 \end{pmatrix} \quad (3.36)$$

其中 B 就是上面这样类型的矩阵所构成的集合, 不难得出该映射为线

性同构。设 A 上的二元运算为叉乘, 那么 A 就转变成为一个代数。设 B 上的二元运算为 $\mathbf{b}_1 \cdot \mathbf{b}_2 = \mathbf{b}_1 \mathbf{b}_2 - \mathbf{b}_2 \mathbf{b}_1$, 其中 \mathbf{b} 是 B 中的一个 3×3 的矩阵。这样一来就使得, f 扩展成了一个代数同构。

由此可见, 两个向量空间之间的线性变换可以稍加修改, 以适应相应的代数乘法的二元运算, 这样就使得原本的线性变换可作用于代数之间。

命题 3.1.6 设 A 和 B 是代数。设 $\{\mathbf{e}_i\}$ 是 A 的基, 并且有 $f: A \rightarrow B$ 是一个线性变换。那么 f 是代数同态当且仅当

$$f(\mathbf{e}_i \mathbf{e}_j) = f(\mathbf{e}_i) f(\mathbf{e}_j) \quad (3.37)$$

证明: (反之亦然) 设 $\mathbf{a} = \sum_i \alpha_i \mathbf{e}_i$ 和 $\mathbf{b} = \sum_j \beta_j \mathbf{e}_j$, 那么

$$f(\mathbf{ab}) = f\left(\left(\sum_i \alpha_i \mathbf{e}_i\right) \left(\sum_j \beta_j \mathbf{e}_j\right)\right) = f\left(\sum_i \alpha_i \sum_j \beta_j \mathbf{e}_i \mathbf{e}_j\right) \quad (3.38)$$

$$= \sum_i \alpha_i \sum_j \beta_j f(\mathbf{e}_i \mathbf{e}_j) = \sum_i \alpha_i \sum_j \beta_j f(\mathbf{e}_i) f(\mathbf{e}_j) \quad (3.39)$$

$$= f\left(\sum_i \alpha_i \mathbf{e}_i\right) f\left(\sum_j \beta_j \mathbf{e}_j\right) = f(\mathbf{a}) f(\mathbf{b}) \quad (3.40)$$

推论 3.1.1 设 $f: A \rightarrow B$ 是代数同态, 则 $f(A)$ 为 B 的一个子代数。

根据定理 2.3.4, $f(A)$ 是 B 的子空间。给定 $\mathbf{b}_1, \mathbf{b}_2 \in f(A)$, 那么存在 $\mathbf{a}_1, \mathbf{a}_2 \in A$ 使得 $\mathbf{b}_1 = f(\mathbf{a}_1)$ 和 $\mathbf{b}_2 = f(\mathbf{a}_2)$, 那么就有:

$$\mathbf{b}_1, \mathbf{b}_2 = f(\mathbf{a}_1) f(\mathbf{a}_2) = f(\mathbf{a}_1 \mathbf{a}_2), \Rightarrow \mathbf{b}_1 \mathbf{b}_2 \in f(A) \quad (3.41)$$

例 3.1.9 \mathbb{F} 是任意含幺代数的子代数的另一证明:

回顾例 3.1.2, 我们还可以从另一个角度阐述。设 A 为一个单位元是 1 的代数, 有线性映射 $f: \mathbb{F} \rightarrow A$, $f(\alpha) = \alpha 1$, 考虑 \mathbb{F} 为一个在其自身之上的代数, 那么我们有:

$$f(\alpha\beta) = \alpha\beta 1 = (\alpha 1)(\beta 1) = f(\alpha)f(\beta) \quad (3.42)$$

这表明 f 是一个代数同态, 进一步也是单态射因为:

$$f(\alpha) = f(\beta) \Rightarrow \alpha 1 = \beta 1 \Rightarrow (\alpha - \beta)1 = 0 \Rightarrow \alpha = \beta. \quad (3.43)$$

因此, \mathbb{F} 与 $f(\mathbb{F})$ 是一致的。根据推论 3.1.1 可得 \mathbb{F} 是 A 的子代数。

定义 3.1.13 设 A 和 B 是含幺代数。如果代数同态 $f: A \rightarrow B$ 有 $f(1_A) = 1_B$, 则 f 是含幺的 (*unital*)。

命题 3.1.7 满态射 $f: A \rightarrow B$ 是含幺的, 如果 A, B 为含幺代数。

我们已经建立了从 V 到 W 的线性变换所组成的集合 $L(V, W)$ 是一个向量空间。让我们尝试在线性空间 $L(V, W)$ 上定义一个乘法规则, 使其成为一个代数。我们自然的就会想到让线性映射的复合成为其乘法。如果 $T: V \rightarrow W$ 和 $S: W \rightarrow U$ 都是线性算子, 那么它们的复合 $S \circ T: V \rightarrow U$ 也是一个线性算子。我们知道一个代数必须是在单个向量空间上的, 然而这样的乘法并非在单个向量空间上被定义的, 而是有 $T \in L(V, W)$ 和 $S \in L(W, U)$ 乘积而得出了 $S \circ T \in L(V, U)$ 。那么, 为了使其成为一个代数, 我们就需要让 $V = W = U$, 使得这三个不同的向量空间缩减成一个向量空间 $L(V, V)$, 注意到 $L(V, V)$ 也是 V 的自同态的集合 $\text{End}(V)$, 其中包含元素 $T, S, ST \equiv S \circ T, TS \equiv T \circ S$ 。因此, V 上的自同态 $L(V, V)$ 是一个代数, 并且有恒等映射 1 作为其他映射的恒元。

定义 3.1.14 代数 V 的自同态 $w: V \rightarrow V$ 是一个**对合** (*involution*), 如果 w 的**平方**为恒等映射 $w \circ w \equiv ww = \mathbf{1}$ 。也就是说, 对合 w 在作用两次之后回到起点。由此可见, 对合一定是双射。

注意: 尤其有 $\mathbf{1} \in \text{End}(V)$ 为一个对合。如果 w_1 和 w_2 都是对合, 并且有 $w_1 \circ w_2 = w_2 \circ w_1$, 那么 $w_1 \circ w_2$ 也是一个对合。之前的共轭复数 $*$ 和转置 T 都是对合, 并且先共轭再转置和先转置再共轭没有区别, 因此共轭转置也是对合。

扩展 3.1.1 双重否定表肯定, 因此数理逻辑的非 \neg 和集合论的补集 \sim 都是对合。在欧几里得几何中, 镜像是对合, 因为原像的镜像的镜像是原像。如果 $\mathbf{A}^2 = \mathbf{I}$, 那么 \mathbf{A} 是对合矩阵。

不过对于代数来说, 对合不仅仅是一个线性映射还必须是一个代数同态。设 A 为一个代数, 并且将 A 的代数同态集合记作 $H(A)$ 。对合 $w \in H(A)$ 满足 $w \circ w = \iota \in H(A)$, 其中 ι 表示恒等同态。如果 A 有恒元 \mathbf{e} , 那么就有 $w(\mathbf{e}) = \mathbf{e}$ 。若设 $w(\mathbf{e}) = \mathbf{a}$, 那么由于 $w \circ w = \iota$ 使得我们必须要有:

$$w(\mathbf{a}) = \mathbf{e} \wedge w(\mathbf{ea}) = w(\mathbf{e})w(\mathbf{a}) = w(\mathbf{e})\mathbf{e} = w(\mathbf{e}) \quad (3.44)$$

在其两侧“乘以”对合 w , 可得 $\mathbf{ea} = \mathbf{e} \Rightarrow \mathbf{a} = \mathbf{e}$ 。

定理 3.1.2 设 U 和 V 都是同构的向量空间, 那么 $L(U)$ 和 $L(V)$ 作为代数都是同构的。

证明: 设 $f: U \rightarrow V$ 为同构 (双射), 映射 $F: L(U) \leftarrow L(V)$ 为:

$$F: L(U) \rightarrow L(V); F(\mathbf{T}) = f \circ \mathbf{T} \circ f^{-1} \quad (3.45)$$

线性映射 F 是一个代数同态, 因为: $F(\mathbf{T}_1)F(\mathbf{T}_2) = f \circ \mathbf{T}_1 \circ f^{-1} \circ f \circ \mathbf{T}_2 \circ f^{-1} = f \circ \mathbf{T}_1\mathbf{T}_2 \circ f^{-1} = F(\mathbf{T}_1\mathbf{T}_2)$, 同时 F 是双射的代数同态, 也就是代数同构, 因为: $F^{-1}(\mathbf{T}) = (f \circ \mathbf{T} \circ f^{-1})^{-1} = f \circ \mathbf{T} \circ f^{-1} = F(\mathbf{T})$ 。

根据定理 2.3.12 两向量空间维度相同当且仅当其为同构的, 可得对于任意数域 \mathbb{F} 上的 N 向量空间 V 都有代数 $L(V)$ 和 $L(\mathbb{F}^N)$ 。

3.2 理想 (Ideals)

子代数是在其元素相乘下是稳定的子空间; 也就是说, 子代数中元素的乘积仍在子代数中。在代数理论中更重要的一类子集, 该子集中的元素与整个代数任意元素乘积的结果任在这一子集中。

定义 3.2.1 代数 A 的子空间 B 是一个**左理想** (*left ideal*) 记作 $AB \subset B$, 如果 B 包含式 3.12 定义的 AB , 并且 $ab \in AB, \forall a \in A, b \in B$ 。类似的子空间 B 是**右理想** (*right ideal*) 记作 $BA \subset B$, 如果 $ba \in BA, \forall a \in A, b \in B$ 。一个子空间如果既是左理想也是右理想, 那么我们称其为一个**双边理想** (*two-sided ideal*) 并且简称为**理想** (*ideal*)。

推论 3.2.1 理想是子代数; 包含恒元或可逆元素的含么代数的唯一理想就是代数本身。

例 3.2.1 左 (右) 零化子是左 (右) 理想:

设 a 为结合代数 A 中的元素, $l(a)$ 为元素 $x \in A$ 的集合, 且有 $xa = 0$ 。对于任意 $x \in l(a)$ 和 $y \in A$, 有 $(yx)a = y(xa) = 0$ 使得 $yx \in l(a)$ 。因此, **左零化子** (*left annihilator*) $l(a) \equiv \{x | xa = 0, x \in A, a \in A\}$ 为 A 的左理想。类似的, **右零化子** $r(a) \equiv \{x | ax = 0, x \in A, a \in A\}$ 也是右理想。

例 3.2.2 $C^r(a, b)$ 无穷维代数:

$C^r(a, b)$ 是一个 r 阶可导的实值函数所组成的一个向量空间, 我们对其定义乘积规则如下:

$$(fg)(t) \equiv f(t)g(t), \forall t \in (a, b), f \in C^r(a, b), g \in C^r(a, b) \quad (3.46)$$

不难验证该乘积下的代数 $C^r(a, b)$ 既是交换也是结合的, 并且有恒元 $f(t) = 1$ 。所有在 $c \in (a, b)$ 这一个点上被零化的函数所构成的集合为 $C^r(a, b)$ 中的一个双边理想, 因为该代数交换。更一般地, 我们设 M_n 为矩阵构成的 (非交换) 代数, 其每一项为 $f_{ij} \in C^r(a, b)$, 则其项在给定的点 $c \in (a, b)$ 上零化的矩阵所构成的集合也是 M_n 中的一个双边理想。

定理 3.2.1 代数同态 $f: A \rightarrow B$ 的核 $\ker f$ 为 A 的双边理想。

证明: 由定理 2.3.3 可得, 同态 $f: A \rightarrow B$ 的核 (零空间) $\ker f$ 是 A 的子空间。因此若设 $\mathbf{x} \in \ker f$ 和 $\mathbf{a} \in A$ 就有:

$$f(\mathbf{x}\mathbf{a}) = f(\mathbf{x})f(\mathbf{a}) = \mathbf{0} \Rightarrow \mathbf{x}\mathbf{a} \in \ker f \quad (3.47)$$

这说明 $\ker f$ 是 A 的一个右理想, 类似的 $\ker f$ 也是 A 的一个左理想。

结合代数 A 的元素 \mathbf{x} 可以**生成理想** (generated ideals)。 A 的左理想为 $A\mathbf{x} \equiv \{\mathbf{a}\mathbf{x} | \mathbf{a} \in A\}$; 右理想为 $\mathbf{x}A \equiv \{\mathbf{x}\mathbf{a} | \mathbf{a} \in A\}$; 双边理想为 $A\mathbf{x}A \equiv \{\mathbf{a}\mathbf{x}\mathbf{b} | \mathbf{a}, \mathbf{b} \in A\}$ 。

定义 3.2.2 代数 A 的左 (右, 双边) 理想 M 称作**极小的** (minimal), 如果所有包含于 M 中 A 的左 (右, 双边) 理想 $K \subset M$ 都和 M 是一致的 $K = M$ 。

定理 3.2.2 l 是 A 的一个极小的左理想, 等价于

$$A\mathbf{x} = l \Leftrightarrow l\mathbf{x} = l, \forall \mathbf{x} \in l \quad (3.48)$$

极小的右理想也是这样由定义直接产生的性质。

定理 3.2.3 若 l 为代数 A 的极小左理想, 那么经过满态射 $f: A \rightarrow B$ 的 $f(l)$ 也是代数 B 的极小左理想。特别有, 对于任意代数的自同态都是其自身的极小理想的同构。

证明: 对于任意的 $b \in B, y \in f(l)$, 都分别存在元素 $a \in A, x \in l$, 使得 $b = f(a), y = f(x)$ 。由于 $ax \in l$,

$$by = f(a)f(x) = f(ax \in f(l)) \quad (3.49)$$

因此 $f(l)$ 是 B 中的一个左理想。现在假设 l 是极小的。为了证明 $f(l)$ 也是极小的, 我们需要用到定理 3.2.2。由于 f 为满态射, 因此 $B = f(A)$ 。设 $u \in f(l)$, 则存在 $t \in l$ 使得 $u = f(t)$ 并且:

$$Bu = f(A)f(t) = f(At) = f(l) \quad (3.50)$$

该定理前半段得证。设自同态 $g: A \rightarrow A$ 对于任意 $t \in l$ 有:

$$l = At = g(l) = g(At) = g(A)g(t) = Ag(t) \Rightarrow g(t) = t \quad (3.51)$$

此外, 我们还能利用定理 $\ker f$ 是 A 的一个理想证明后半段。

定义 3.2.3 A 是其子代数 B, C 的直和, 若有类似向量空间上的 $A = B \oplus C$ 并且 $BC = CB = \{0\}$ 。我们称 A 为其分量 B, C 的直和。一个由其子代数的直和构成的代数被称作**约化 (reducible)** 代数。

尽管 $BC = CB = \{0\}$ 是我们上面定义中的一个条件, 实际上这也是由之前我们对代数直和的定义所带来的必然结果, 因为 B, C 本身就也分别等于 $B \oplus \{0\}, C \oplus \{0\}$ 。

命题 3.2.1 中心代数是不可约化的。

证明: 假若代数 $A = B \oplus C$ 是可约的, 并且是一个中心 (含么) 代

数 $Z(A) = \text{Span}\{\mathbf{1}\} = \mathbb{F}$, 那么根据定理 2.1.6 其恒元能够唯一的表示为 $\mathbf{1} = \mathbf{b} + \mathbf{c}$, 显然 $\mathbf{b}, \mathbf{c} \in \mathbb{F}$ 并且 \mathbf{b}, \mathbf{c} 根据定理 2.1.7, 作为 A 的子代数 (子空间) B, C 中的元素还得是线性无关的, 这是相互矛盾的。回顾例 3.1.3 介绍的中心代数 S , 现在构造一个新的基底 $\{\mathbf{f}_i\}_{i=1}^4$:

$$\mathbf{f}_1 = \frac{1}{2}(\mathbf{e}_0 + \mathbf{e}_3), \mathbf{f}_2 = \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_2), \mathbf{f}_3 = \frac{1}{2}(\mathbf{e}_0 - \mathbf{e}_3), \mathbf{f}_4 = \frac{1}{2}(\mathbf{e}_1 + \mathbf{e}_2). \quad (3.52)$$

由上述公式我们可以得出乘法表 3.2。在么元 $\mathbf{e}_0 = \mathbf{f}_1 + \mathbf{f}_3$ 的两侧乘以任意 S 中的元素, 对于这样的元素, 我们可以将其写作左理想 $L_1 \equiv S\mathbf{f}_1$ 中的一个向量加上左理想 $L_3 \equiv S\mathbf{f}_3$ 。注意到任意 L_1 中的向量都可以写作 S 中的一个向量和 \mathbf{f}_1 的乘积, 也就是设 $\mathbf{a} = \sum_{i=1}^4 \alpha_i \mathbf{f}_i$ 为 S 中的任意一个元素, 那么 L_1 中的任意元素写作如下形式:

$$\mathbf{a}\mathbf{f}_1 = \alpha_1 \mathbf{f}_1 + \alpha_4 \mathbf{f}_4 \quad (3.53)$$

	\mathbf{f}_1	\mathbf{f}_2	\mathbf{f}_3	\mathbf{f}_4
\mathbf{f}_1	\mathbf{f}_1	\mathbf{f}_2	$\mathbf{0}$	$\mathbf{0}$
\mathbf{f}_2	$\mathbf{0}$	$\mathbf{0}$	\mathbf{f}_2	\mathbf{f}_1
\mathbf{f}_3	$\mathbf{0}$	$\mathbf{0}$	\mathbf{f}_3	\mathbf{f}_4
\mathbf{f}_4	\mathbf{f}_1	\mathbf{f}_3	$\mathbf{0}$	$\mathbf{0}$

表 3.2: S 的乘法表

索引

- n 元映射 (n -ary map), 23
- n 元组 (n -tuples), 12
- p 重 (p -fold), 72
- p 重线性映射 (p -linear map), 72
- 不变量 (invariant), 63
- 不可数的 (uncountable), 20
- 乘法恒等 (multiplicative identity), 36
- 二元运算 (binary operation), 24
- 二项式定理 (Binomial theorem), 31
- 交集 (Intersection), 9
- 代数 (algebra), 25
- 代表 (representative), 14
- 伴随 (adjoint), 65
- 值域 (range), 17
- 偶排列 (even permutation), 73
- 像 (image), 17
- 元素 (elements), 6
- 光滑 (smooth), 37
- 克罗内克符号 (Kronecker delta), 50
- 共轭复数 (complex conjugate), 48
- 共轭转置 (conjugate transpose), 71
- 关系 (Relations), 12
- 内积空间 (inner product space), 49
- 内部 (interior), 30
- 函数 (Function), 16
- 分析 (analysis), 25
- 分量 (components), 35, 39
- 划分 (partition), 15
- 列向量 (column vector), 69
- 列维-奇维塔符号 (Levi-Civita), 73
- 剑标 (dagger), 71
- 加权函数 (weight function), 49
- 加法恒等 (additive identity), 36
- 加法的逆 (additive inverse), 36
- 区间 (Intervals), 23
- 半双线性 (sesquilinear), 49
- 单位向量 (unit vector), 40
- 单位圆 (unit circle), 24
- 单射 (injective or one-to-one), 19
- 厄米共轭 (Hermitian adjoint), 35
- 厄米共轭 (hermitian), 49
- 原像 (Preimage), 19
- 双射 (bijective), 20

- 双线性 (bilinearity), 65
- 反伴算子 (skew), 66
- 反函数 (inverse function), 21
- 反对称性 (anti-symmetric), 73
- 反线性 (antilinear), 56
- 发散 (divergence), 28
- 可数的 (countably infinite), 20
- 可逆 (invertible), 62
- 可逆的 (invertible), 21
- 右矢 (ket), 35
- 同余关系 (Congruence Relation), 15
- 同构 (isomorphism), 62
- 同构的 (isomorphic), 62
- 同胚 (homeomorphic), 30
- 同胚映射 (homeomorphism), 30
- 向量 (vectors), 35
- 向量势 (vector potential), 14
- 商映射 (Quotient Map), 22
- 商集 (Quotient Set), 15
- 图像 (Graph), 17
- 域 (field), 36
- 基 (basis), 39
- 基数 (cardinal), 20
- 基数 (cardinality), 39
- 复化 (complexification), 67
- 复合 (composition), 19
- 复向量空间, 36
- 复结构 (complex structure), 66
- 奇排列 (odd permutation), 73
- 子覆盖 (subcover), 30
- 子集 (Subset), 7
- 完备度量空间, 28
- 定义域 (Domain), 17
- 定义良好 (well defined), 64
- 实向量空间 (real vector space), 36
- 对偶 (dual), 70, 71
- 对偶基 (dual basis), 70
- 对偶空间 (dual space), 68
- 对换 (transposition), 72
- 对称和双线性 (sesqui-symmetry), 71
- 对称性 (symmetry), 65
- 导数算子 (derivative operator), 57
- 封闭的 (closed), 25
- 射影映射 (projection map), 23
- 射影空间 (Projective Space), 16
- 左矢 (bra), 35
- 差集 (difference), 10
- 幂集合 (Power Set), 8
- 平行四边形法则, 55
- 并集 (Union), 9
- 序列 (Sequence), 18
- 康托尔集 (Cantor set), 20
- 开球 (Open ball), 27
- 开覆盖 (Open cover), 30
- 开集 (open sets), 29
- 弦长 (chord length), 26
- 归一化向量, 50
- 恒等映射 (Identity map), 16
- 拉回 (pullback), 71
- 拓扑 (Topology), 30
- 拓扑空间 (topological space), 30
- 排列 (permutation), 57
- 收敛 (converges), 28
- 散度 (divergence), 14
- 斜对称 (skew-symmetric), 73

- 施瓦茨不等式, 53
 无界的 (unbounded), 29
 无穷幂级数 (infinite power series), 40
 无限的 (infinite), 20
 映射 (Maps), 16
 有序对 (Ordered Pairs), 12
 有界 (bounded), 26
 有界的 (bounded), 29
 有限的 (finite), 20
 条件比较判断 (comparison test), 12
 极化恒等式 (polarization identity), 58
 极限点 (Limit point), 28
 柯西序列 (Cauchy sequence), 28
 标准 (normal), 50
 标准内积 (standard inner product), 50
 标准化向量 (normalized vector), 50
 标准基 (standard basis), 40
 标准正交基 (orthonormal basis), 50
 标量 (scalar), 36
 标量积 (scalar(dot) product), 47
 核 (kernel), 60
 格拉姆-施密特正交化, 52
 欧几里得 (Euclidean), 49
 欧氏空间 (Euclidean space), 26
 正交 (orthogonal), 50
 正则射影 (canonical projection), 23
 正定实 (positive definite real), 49
 正定性 (positive definite), 49
 泰勒定理 (Taylor' s theorem), 40
 满射 (onto or surjective), 20
 狄拉克符号 (Dirac notation), 35
 目标域 (target space), 17
 真子集 (proper subset), 7
 矩阵乘积 (matrix product), 69
 磁场 (Magnetic Field), 14
 秩 (rank), 60
 秩-零化度定理, 61
 积分算子 (integration operator), 57
 空集 (empty set), 8
 笛卡儿积 (cartesian product), 12
 笛卡尔向量 (Cartesian vectors), 35
 符号函数 (sign function), 73
 等价类 (Equivalence Classes), 13
 等距映射 (isometric map), 57
 类 (class), 15
 紧致的 (compact), 30
 纤维 (Fiber), 19
 线性不相关 (linearly independent), 36
 线性变换 (linear transformation), 56
 线性型 (linear form), 68
 线性映射 (linear maps), 55
 线性生成 (linear span), 38
 线性等距 (linear isometry), 57
 线性算子 (linear operator), 56
 线性组合 (linear combination), 36
 组成方向 (components), 14
 结构 (structure), 62
 维度定理 (dimension theorem), 61
 置换 (permutation), 72
 群 (group), 25
 聚焦点 (cluster point), 28
 自伴算子 (self-adjoint), 66
 自变量 (argument), 60
 自同态 (endomorphism), 56

- 自同构 (automorphism), 62
- 范数 (norm), 54
- 螺线矢量场 (solenoid vector field), 14
- 行列式 (determinant), 24
- 行列式函数, 76
- 行向量 (row vector), 69
- 补集 (complement), 10
- 赋范线性空间, 54
- 转置 (transpose), 69
- 边界 (Boundary), 30
- 逆序 (inversion), 72
- 酉 (unitary), 57
- 闭包 (Closure), 30
- 闭球 (Closed ball), 27
- 闵可夫斯基度量, 27
- 闵可夫斯基度量空间, 27
- 陪域 (Co-domain), 17
- 集合 (sets), 6
- 零化 (annihilate), 70
- 零化子 (annihilator), 70
- 零化度 (nullity), 60
- 零空间 (null space), 60
- 非退化 (non-degeneracy), 65