

从零开始建宇宙

少年数学家的积木书



Xulei Sun

从零开始建宇宙

少年数学家的积木书

Xulei Sun

2026-07-10

目录

序言	5
1 对与错的游戏 —— 思考的绝对准绳	6
1.1 先学会读规则 —— 符号、变量与原子公式	6
1.2 只能选一个的硬币 —— 命题与真值	7
1.3 思维的魔法阵 —— 逻辑运算（非、且、或）	8
1.4 条件的开关 —— 蕴含与等价	9
1.5 全知之眼 —— 任意与存在	11
1.6 法庭上的利刃 —— 证明的绝对法则	13
1.7 本章结语	15
2 万物的分类与打包 —— 创造世界的容器	16
2.1 朴素的口袋与思维的崩塌	16
2.2 造物主的咒语 —— 公理系统	18
2.3 跨越维度的配对 —— 笛卡尔积与映射	22
2.4 数学家最伟大的魔法 —— 极度抽象的商集	26
2.5 本章结语	28
3 造出数字 —— 从零到无限的阶梯	30
3.1 无中生有的阶梯 —— 皮亚诺性质与自然数 $\mathbb{N}$	30
3.2 赋予运算的灵魂 —— 递归定义与加法	34
3.3 用商集填补漏洞 —— 整数 $\mathbb{Z}$ 的诞生	38
3.4 填平碎片的鸿沟 —— 有理数 $\mathbb{Q}$ 的网	42
3.5 本章结语	45

4	抽象的骨架 —— 抽离运算的灵魂	47
4.1	抽离加法的灵魂 —— “群” 的诞生	47
4.2	万物皆可运算 —— 映射组成的置换群	49
4.3	双重魔法的叠加 —— 环与域	51
4.4	未来的魔法容器 —— 多项式环	52
4.5	建立绝对的阶级 —— 序结构	55
4.6	本章结语	57
5	完美的直线 —— 直面连续统的深渊	58
5.1	致命的一刀 —— 发现深渊的裂缝	58
5.2	造物主的利刃 —— 戴德金分割	60
5.3	完美的实数轴 $\mathbb{R}$ 与确界原理	62
5.4	永恒的追逐 —— 柯西列与完备性	66
5.5	终极的魔法扩展 —— 复数域 $\mathbb{C}$ 的降临	69
5.6	本章结语	71
6	距离与形态 —— 奇异的拓扑魔法	72
6.1	虚拟的尺子 —— 创造度量空间	72
6.2	画出开球与“靠近”的绝对标准	74
6.3	抛弃尺子 —— 纯粹拓扑空间的诞生	75
6.4	撕不断与逃不出 —— 连通性与紧致性	80
6.5	空间之间的舞蹈 —— 连续映射	81
6.6	本章结语	83
7	超越的造物 —— 填满缝隙的无限之舞	84
7.1	在无限的边缘跳舞 —— 级数与收敛	84
7.2	宇宙第一召唤阵 —— 指数魔法 $\exp(x)$	88
7.3	用代数撕裂波纹 —— 召唤波浪的灵魂	90
7.4	捕获幽灵 —— 圆周率 π 的纯代数身份证	91
7.5	大一统的和弦 —— 欧拉公式与复数 $\mathbb{C}$ 的圆满	92
7.6	本章结语	95

8 多维的平直宇宙 —— 阵列与对偶的代数	96
8.1 撑开维度的网 —— 线性空间与基底	96
8.2 操控宇宙的坐标阵列 —— 线性映射与矩阵	99
8.3 捕捉影子的猎人 —— 对偶空间 V^*	101
8.4 定义长度与垂直 —— 内积与几何法则	102
8.5 换座位就变号 —— 外代数与反对称性	105
8.6 本章结语	107
 9 变化的魔法 —— 驯服无穷小的野兽	 109
9.1 捕获瞬间的切线 —— 导数的诞生	109
9.2 碎片的堆积 —— 积分与极限的总和	112
9.3 导数与积分的联系 —— 微积分基本定理	114
9.4 多维的瞬间扭曲 —— 雅可比矩阵	115
9.5 跨越维度的通行证 —— 隐函数定理	118
9.6 本章结语	120
 10 弯曲的宇宙与边界 —— 斯托克斯的终极和弦	 121
10.1 缝合多维的地图 —— 微分流形	121
10.2 在弯曲中寻找平直 —— 切空间	125
10.3 测量切向量 —— 微分形式 dx	127
10.4 选择一致的方向 —— 定向与边界	131
10.5 统一内部与边界 —— 斯托克斯定理	134
10.6 全书结语	141

序言

如果有人问你：“为什么 $1 + 1 = 2$ ？”你也许会拿出两个苹果。这个解释很直观，却把更深的问题留在了背后：“1”“+”和“=”究竟是什么意思？它们为什么能够按照我们熟悉的方式工作？

这本书就从这些问题出发。请想象面前有一张空空的工作台：上面还没有数，没有直线，也没有加减乘除。我们先学习怎样把话说准确、怎样让推理不偷步；再用集合造出数，定义运算，建成实数直线；随后研究距离、连续、向量与变化；最后走进弯曲的空间，看看局部规律怎样汇成整体结论。这就是书名所说的“从零开始建宇宙”。

当然，“从零开始”并不是说什么都不使用。阅读这本书仍然需要日常汉语、纸面符号和按照规则思考的能力；数学本身也必须有公开的起点。本书在第一章说明逻辑规则，在第二章列出集合论公理，后面的对象与结论都从这些起点逐步建立。自然数正式造出以前，页码、章节编号和故事中出现的数字只是讲书用的标记，不会被当作数学证明的根据。

全书遵守一项基本约定：一个数学概念必须先定义，才能正式使用；一个对象的存在必须由公理或此前内容保证；一个结论必须能够从已经得到的知识推出。图画、比喻和例子可以帮助理解，却不能冒充证明。简单而重要的证明会完整写出；过于繁琐或需要较多准备的证明有时会省略，但相应结论仍会准确陈述，并且能够由此前的规则与知识证明。

正文会用【定义】【定理】【直觉】【例子】【思考训练】等标记提醒你：眼前的话是在规定含义、陈述结论，还是帮助想象。读到新内容时，不妨经常问三个问题：它是什么意思？它从哪里来？它为什么成立？

你不必读得很快。数学不是一场背诵比赛，而是一座可以反复拆开、重新拼好的积木城。愿意在“为什么”面前多停一会儿，正是数学思考开始生长的地方。

现在，让我们放下第一块积木，一步一步建起自己的数学宇宙。

1 对与错的游戏 —— 思考的绝对准绳

欢迎来到绝对虚无的边界。

这里没有空间，没有时间，没有积木，没有画笔，没有三角形，甚至连数字都没有。“ $1 + 1 = 2$ ”在这里毫无意义，因为我们还没向宇宙法庭申请创造出“1”和“2”。

但不要害怕，即使在一无所有的混沌中，你依然拥有一种最强大的武器——思考。

在这一章，我们要剥离一切现实的伪装，直接写下思考的“基础法则”。这是创造整个数学宇宙的绝对前提。如果你不能熟练掌握这些法则，未来的宇宙就会充满自相矛盾的漏洞，最终彻底崩塌。

现在，让我们开始制定宇宙的第一批法则。

1.1 先学会读规则 —— 符号、变量与原子公式

【约定】我们先承认可以写出有限长的符号串，并按从左到右的顺序阅读它们。括号只负责标明组合范围；括号本身不会创造新的对象或新的真理。

字母 x, y, z 可以暂时代替论域中的对象，称为**变量**。变量，以及以后明确定义为“指代某个对象”的符号串，统称为**项**。符号 $=$ 表示“是同一个对象”。如果 s, t 是项，那么 $s = t$ 是最简单的一类公式，叫作**原子公式**。在具体理论中还可以有别的关系符号；例如第二章引入 $\in$ 后， $s \in t$ 也会成为原子公式。

【等号规则】每个对象都等于自己；若 $s = t$ ，则 $t = s$ ；若 $s = t$ 且 $t = u$ ，则 $s = u$ 。等号替换还必须服从一条防止符号“误入量词笼子”的条件。等到量词获得定义后，我们会在本章末精确写出这条条件。

原子公式只是造句的种子。下面将依次定义怎样用逻辑联结词和量词把它们拼成更长的公式，并区分“指代对象的项”“关于对象的公式”和“已经能够判断真假的命题”。

【易混点】 $x = y$ 已经是一条公式，却还不一定是命题：只要 x, y 尚未指定，它就留下了空位。相反，在给定解释后，一个不含空位的式子才能直接谈真或假。“是合法公式”和“已经有确定真值”是两层不同的问题。

1.2 只能选一个的硬币 —— 命题与真值

人类的日常语言充满了含糊其辞。比如“这朵花很美”，有人觉得美，有人觉得不美；又比如“今天天气真好”，对于想出去玩的人来说是好天气，对于想在家睡觉的人来说可能就不算好。

但在我们的数学宇宙里，绝不允许出现这种模棱两可的话。

我们引入第一个概念：**命题**。

【初步定义】给定一个非空论域，并给各个非逻辑符号规定解释后，一个没有留下“尚未指定对象”的完整陈述叫作**命题**。本章量词一节会用“自由变量”把这句话写得完全精确。在本书采用的**经典逻辑**中，每个命题恰有【真】和【假】两个真值之一。

因为没有任何具体事物，我们用大写字母 P, Q, R 来代表任意的命题。

同时，我们规定命题只有两种“真值”（真实的状态）：

- 【真】：代表这句话在当前给定的论域和解释下成立。
- 【假】：代表这句话在当前给定的论域和解释下不成立。

下面的练习不依赖外部经验，而是先公开给出一个小型解释：我们约定物体甲被涂成红色，物体乙没有被涂成红色；“漂亮”则没有被赋予固定解释。

【思考训练】

请判断下面这四句话，哪些是命题？如果是命题，它是【真】还是【假】？

1. 物体甲被涂成红色。
2. 物体乙被涂成红色。
3. 物体甲很漂亮。
4. 物体甲被涂成红色吗？

解答：

第 1 句是命题；根据公开给出的解释，它的真值是【真】。

第 2 句是命题；根据同一解释，它的真值是【假】。

第 3 句在当前约定下不是命题，因为“漂亮”没有被赋予固定解释，无法确定这句话的真值。

第 4 句不是命题，它是一个问句，不是一个陈述判断。

1.3 思维的魔法阵 —— 逻辑运算（非、且、或）

有了命题，我们可以像玩积木一样，把简单的句子拼成更复杂的句子。这就需要用到三种最基础的逻辑魔法阵。为了方便理解，我们先设定两个命题：

命题 P ：我今天写完了作业。

命题 Q ：我今天吃了一块糖。

1. 绝对的倒影：【非】($\neg$)

这是最简单的魔法。在命题前面加上符号 $\neg$ （读作“非”），它的真假就会瞬间颠倒。这就像照一面魔镜，镜子里面的世界和外面完全相反。

- 如果 P 是【真】（我确实写完了作业），那么 $\neg P$ 就是【假】（“我今天没有写完作业”这句话是错的）。
- 如果 P 是【假】（我根本没写完作业），那么 $\neg P$ 就是【真】（“我今天没有写完作业”这句话是对的）。

【思考训练】 $\neg(\neg P)$ 是什么？

解答：颠倒两次后的公式与 P 在每种赋值下真值相同；稍后我们会把这种关系命名为“逻辑等价”。这里不是说两个符号串逐字相同。这就好比“我不是没有写完作业”，在经典解释下与“我写完了作业”真假一致。

2. 最严苛的法官：【且】($\wedge$)

把命题 P 和 Q 用 $\wedge$ 连起来，变成 $P \wedge Q$ （读作“ P 且 Q ”）。这意味着两件事情必须同时发生。

这个魔法阵极其严格：只有当 P 和 Q 同时为【真】时，整个句子才算【真】。只要其中混入哪怕一个【假】，整个句子就是【假】。

我们可以列出一个穷尽所有可能的“规则表”（在数学里叫真值表）：

- P 真， Q 真（写了作业，也吃了糖）： $P \wedge Q$ 为【真】
- P 真， Q 假（写了作业，没吃糖）： $P \wedge Q$ 为【假】

- P 假, Q 真 (没写作业, 吃了糖): $P \wedge Q$ 为【假】
- P 假, Q 假 (没写作业, 也没吃糖): $P \wedge Q$ 为【假】

3. 宽容的守门人:【或】($\vee$)

把命题 P 和 Q 用 $\vee$ 连起来, 变成 $P \vee Q$ (读作“ P 或 Q ”)。这就宽容多了。

这个魔法阵的规则是: 只要 P 和 Q 里面至少有一个是【真】, 整个句子就是【真】。只有当两人全都是【假】时, 它才是【假】。

- P 真, Q 真 (写了作业, 也吃了糖): $P \vee Q$ 为【真】
- P 真, Q 假 (写了作业, 没吃糖): $P \vee Q$ 为【真】
- P 假, Q 真 (没写作业, 吃了糖): $P \vee Q$ 为【真】
- P 假, Q 假 (没写作业, 也没吃糖): $P \vee Q$ 为【假】

有了这三个符号 ($\neg, \wedge, \vee$), 我们就可以像搭积木一样制造复合公式。括号标明组合范围; 例如 $\neg(P \wedge Q)$ 与 $(\neg P) \wedge Q$ 是不同的公式。

【经典逻辑法则】现在这些符号已经有了含义, 我们正式写下**排中律** $P \vee \neg P$ 与**矛盾律** $\neg(P \wedge \neg P)$ 。前者说“ P 或非 P ”必真; 后者说同一命题不能同时为真又为假。这是两条不同的法则。

1.4 条件的开关 —— 蕴含与等价

接下来, 我们定义一种极度重要、但也最容易违背直觉的逻辑联结词。请注意: 逻辑蕴含不自动表示现实中的原因和结果。

【定义】我们用箭头 $\rightarrow$ 表示**蕴含命题**, 写成 $P \rightarrow Q$ (读作“如果 P , 那么 Q ”)。这里的 P 叫前件, Q 叫后件。只有把这些公式用于推理时, 我们才把已经假设的公式称为前提, 把推出的公式称为结论。

如何判断 $P \rightarrow Q$ 这句话整体是【真】还是【假】呢? 我们把 $P \rightarrow Q$ 想象成父母和你签的一份“契约”。

命题 P : 你期末考试考了一百分。

命题 Q : 父母给你买一个大玩具。

契约 $P \rightarrow Q$: 如果你考了一百分, 父母就给你买一个大玩具。

我们来看看父母有没有违背契约 (说谎):

- 情况一： P 是真， Q 是真。你考了一百分，父母买了玩具。契约完美履行，父母没说谎，这句话为【真】。
- 情况二： P 是真， Q 是假。你考了一百分，父母却没有买玩具。这是赤裸裸的违约！父母说谎了，这句话为【假】。

真正反直觉的考验来了：

如果前件 P 一开始就是【假】的呢？也就是说，你没有考到一百分。

这个时候，如果父母依然给你买了一个大玩具（ Q 为真），父母违背契约了吗？没有，因为契约只规定了“考一百分”会怎样，没规定“不考一百分”会怎样。父母大发慈悲奖励你，并不算违约。

这个时候，如果父母没有给你买玩具（ Q 为假），父母违背契约了吗？也没有。

在数学宇宙中，我们做出了一个霸道的规定：只要前件 P 为【假】，那么不管后件 Q 是真是假，整个命题 $P \rightarrow Q$ 都算作【真】！因为只要前件没发生，契约就没有被打破。

这被称为“空虚的真”。虽然听起来怪怪的，但请死死记住它。未来当我们面对“什么都没有”的深渊时，这个规定将挽救我们的宇宙。

还有一种更紧密的关系：等价（ $\leftrightarrow$ ）。

$P \leftrightarrow Q$ 意味着 $(P \rightarrow Q) \wedge (Q \rightarrow P)$ 。在一次给定的真假赋值下，只要 P 和 Q 同真或同假， $P \leftrightarrow Q$ 就为真。

但是，一次赋值下同真同假并不表示两个命题“互为因果”或在所有地方都能替换。只有当 $P \leftrightarrow Q$ 在所有允许的解释与变量赋值下都为真时，我们才称 P 与 Q 逻辑等价，记作 $P \equiv Q$ ；此时才能在纯逻辑推理中互换它们。例如两次取非给出 $\neg\neg P \equiv P$ 。

【思考训练】

有一份契约：“如果明天下雨，我就带雨伞。”

假设第二天是大晴天（没有下雨），而且我也没带雨伞。请问，我违背契约了吗？这份契约是【真】还是【假】？

解答：

没有违背契约。因为前提“明天下雨”是假的，所以无论我带不带伞，契约都没有被破坏。根据“空虚的真”，这份契约整体是【真】的。

1.5 全知之眼——任意与存在

前面用 P, Q 讨论的主要是不带变量空位的命题。但在未来的宇宙里，我们会创造出许许多多的成员（比如点和数字）。如果要对一群对象统一陈述规则，就需要两个量词符号。

每次解释量词时，都必须先明确一个非空的**论域**，也就是变量允许取值的范围。假设论域是一袋球， $P(x)$ 表示“球 x 是红色的”。以后写“ $x \in A$ 中任取 x ”时，集合 A 就充当局部范围；空范围会由第二章定义的有界量词缩写处理，而不是把一阶逻辑的总论域设为空。

【公式生成规则】 原子公式是公式；若 P, Q 是公式，则 $\neg P$ 、 $(P \wedge Q)$ 、 $(P \vee Q)$ 、 $(P \rightarrow Q)$ 与 $(P \leftrightarrow Q)$ 都是公式；若 P 是公式而 x 是变量，则 $\forall x P$ 与 $\exists x P$ 也是公式。只有从原子公式出发，有限次使用这些规则得到的符号串，才算本书的公式。

量词只管辖紧跟它的公式范围。在 $\forall x P$ 或 $\exists x P$ 中，该量词约束 P 里那些没有先被更内层同名量词接管的 x ；这些位置叫作 x 的**约束出现**，其余位置叫作**自由出现**。一个变量若在公式中有自由出现，就叫该公式的**自由变量**；没有自由变量的公式叫作**闭公式**。给定论域和各个非逻辑符号的解释后，闭公式才成为前面初步定义的命题。例如 $\forall x (P(x) \rightarrow Q(y))$ 中的 x 被约束，而 y 仍是自由的。

1. 绝对无死角：【任意】($\forall$)

这是一个倒写的字母 A，代表“无一例外”、“所有的”。

写成： $\forall x P(x)$

这读作：“对于任意一个球 x ，它都是红色的。”

要让这句话成立， $P(x)$ 必须对论域中每个球都成立。证明时不必逐个检查；也可以对任意选取的球作统一推理。但只要存在一个反例，整个全称命题就为假。

2. 哪怕只有一个：【存在】($\exists$)

这是一个反写的字母 E，代表“至少有一个”、“能找到”。

写成： $\exists x P(x)$

这读作：“存在至少一个球 x ，它是红色的。”

给出一个满足 P 的具体对象（称为见证）就足以证明存在命题。经典逻辑也允许通过反证等方法间接证明存在，但无论采用哪种证明，命题的含义始终是“论域中至少有一个对象满足 P ”。

量词的阴阳转换（极其重要的法则）：

如果有人大声宣布：“世界上所有的乌鸦都是黑色的！”（任意乌鸦，都是黑色）。你要怎么反驳他？你需要抓尽世界上所有的乌鸦，证明它们都不是黑色吗？不需要！你只需要像个聪明的侦探一样，在森林里找到仅仅一只白色的乌鸦，对方就输了！

这就是量词转换的魔法：要推翻“所有人都是对的”，只需要找到“存在一个人是错的”即可。

用我们刚学的符号写出来，就是无比优雅公式：

- $\neg(\forall x P(x)) \equiv \exists x \neg P(x)$
（“并非任意 x 都满足 P ”等价于“存在一个 x 不满足 P ”）

同理，“不存在会飞的动物”与“所有动物都不会飞”表达同一个条件；只要找到一只会飞的动物，这两种说法就都会被推翻。

- $\neg(\exists x P(x)) \equiv \forall x \neg P(x)$
（“不存在 x 满足 P ”等价于“任意 x 都不满足 P ”）

量词的范围尤其不能凭感觉移动。看一个完全确定的例子：设论域里只有甲球和乙球， $P(x)$ 表示“ x 是红色的”， $Q(x)$ 表示“ x 是蓝色的”；甲球只红不蓝，乙球只蓝不红。那么

$$\forall x (P(x) \vee Q(x))$$

为真，因为每只球至少有一种颜色；但

$$(\forall x P(x)) \vee (\forall x Q(x))$$

为假，因为并非所有球都红，也并非所有球都蓝。因此这两个公式不逻辑等价。括号和量词位置真的会改变意思。

【思考训练】 $\forall x (P(x) \wedge Q(x))$ 与 $(\forall x P(x)) \wedge (\forall x Q(x))$ 是否逻辑等价？

解答：是。若每个对象同时满足 P, Q ，那么每个对象满足 P ，也每个对象满足 Q ；反过来，若每个对象都满足 P 且每个对象都满足 Q ，任取一个对象，它便同时满足两者。两边在每个论域和解释下真值都相同。

掌握了这套魔法，你就拥有了全知之眼。再宏大的宣告、再隐蔽的谎言，在你面前都无所遁形。

1.6 法庭上的利刃 —— 证明的绝对法则

作为未来宇宙的造物主，你不可能每遇到一个问题都去穷举所有真假情况。我们需要先说明什么叫“合法地推出”。符号 $\Gamma \vdash Q$ 表示：从前提清单 Γ 出发，能够用允许的推理规则得到 Q 。它不同于命题联结词 $P \rightarrow Q$ 。

若公式 φ 中有 x 的自由出现， $\varphi[t/x]$ 表示把这些自由出现换成项 t 。只有当替换进去的 t 不会被原公式里的量词意外约束时，才允许这样写；更精确地说， x 的任何被替换自由出现都不能位于一个会约束 t 中变量的量词范围内。此时称“ t 可以自由代入 x ”。例如不能把 y 代入 $\forall y (x = y)$ 中的 x ，否则原本自由的 y 会被 $\forall y$ 捕获。

【公理/推理规则】本书允许使用下列经典推理规则：

1. 肯定前件：由 P 和 $P \rightarrow Q$ 推出 Q 。
2. 合取规则：由 P, Q 推出 $P \wedge Q$ ；由 $P \wedge Q$ 可分别推出 P 与 Q 。
3. 全称实例化：由 $\forall x \varphi$ 推出 $\varphi[t/x]$ ，其中项 t 必须可以自由代入 x 。
4. 全称概括：若已经推出 $\varphi(a)$ ，而新参数 a 不自由出现在任何尚未解除的前提中，则可把 a 看作任意对象，推出 $\forall x \varphi(x)$ 。
5. 存在引入：由 $\varphi[t/x]$ 推出 $\exists x \varphi$ ，其中 t 必须可以自由代入 x 。
6. 等号替换：由 $s = t$ 与 $\varphi[s/x]$ 推出 $\varphi[t/x]$ ，但两次代入都必须是自由代入。也就是说，只替换项的自由出现，绝不允许替换动作造成变量捕获。
7. 蕴含引入：若临时假设 P 后能够推出 Q ，则解除该临时假设并推出 $P \rightarrow Q$ 。
8. 析取规则：由 P 可推出 $P \vee Q$ ；已知 $P \vee Q$ 时，若临时假设 P 能推出 R ，临时假设 Q 也能推出同一个 R ，则解除这两个临时假设并推出 R 。
9. 存在消去：已知 $\exists x P(x)$ 时，可取一个此前未用的新参数 a ，临时假设 $P(a)$ 。若由此推出 R ，并且 a 不自由出现在 R 中，也不自由出现在除 $P(a)$ 外任何尚未解除的前提中，就可解除 $P(a)$ ，推出 R 。这些侧条件保证结论没有偷偷依赖所选见证的名字。
10. 反证与否定引入：若假设 P 导出矛盾，则推出 $\neg P$ ；在经典逻辑中，若假设 $\neg P$ 导出矛盾，则推出 P 。
11. 矛盾消去：若已经推出 P 与 $\neg P$ ，则可以推出任意命题 Q 。这条规则也叫“从矛盾可推出一切”。

完整的形式逻辑可以把这些规则进一步拆成更少的基本规则；这里把它们作为公开的起始规则。现在我们打造两把常用利刃：

一次三步推导

假设前提清单里只有 $P \wedge (P \rightarrow Q)$ 。我们来展示“每一步都报出许可证”的写法：

1. 已知 $P \wedge (P \rightarrow Q)$;
2. 由合取规则的拆分方向, 分别得到 P 与 $P \rightarrow Q$;
3. 由肯定前件得到 Q 。

因此 $P \wedge (P \rightarrow Q) \vdash Q$ 。这里没有因为结论“看起来合理”而跳步, 每一步都能在上方的规则表中找到来源。

【思考训练】 若已知 $s = t$ 与 $R(s)$, 其中 $R(x)$ 是一个含自由变量 x 的公式, 并且代入不会造成变量捕获, 怎样推出 $R(t)$?

解答: 在等号替换规则中取 $\varphi(x) = R(x)$ 。此时 $\varphi[s/x] = R(s)$ 、 $\varphi[t/x] = R(t)$, 所以由 $s = t$ 与 $R(s)$ 推出 $R(t)$ 。

第一把利刃: 逆否之剑

假设你要证明 $P \rightarrow Q$ 。检查真值表可得到一个逻辑定理: $P \rightarrow Q$ 与 $\neg Q \rightarrow \neg P$ 在所有赋值下真假相同。
这叫“逆否命题等价”。

因此, 证明 $P \rightarrow Q$ 时, 可以改为证明逆否命题 $\neg Q \rightarrow \neg P$ 。这是一条纯逻辑等价, 不需要、也通常不可能靠检查世界上所有对象来完成。

第二把利刃: 反证法 (宇宙崩溃大法)

这是数学家最喜欢、也最具破坏力的终极武器。

如果你想证明一个结论 R 是【真】的, 但实在不知道怎么下手。怎么办?

1. 你先假装对方赢了, 假设 R 是【假】的 (也就是假设 $\neg R$ 为【真】)。
2. 你把 $\neg R$ 当作前提, 一步步运用已经公开的逻辑规则继续推理。
3. 突然, 你推导出了一个不可理喻的结论! 比如你推导出了 $P \wedge \neg P$ (某件事既是真又是假)。
4. 根据前面讲过的**不矛盾律**, $P \wedge \neg P$ 不可能为真。
5. 既然推理的过程没有犯规, 那问题出在哪里呢? 只能说明: 最初的假设 $\neg R$ 是个彻头彻尾的错误。
6. 既然 $\neg R$ 是假, 那么 R 就必定是【真】!

这就是经典逻辑中的反证法。它依赖不矛盾律、排中律与双重否定消去; 这些都已被列入本章采用的经典逻辑规则。

【思考训练】

桌子上有一个不透明的盒子，我想证明“盒子里没有红色的球”。请用反证法描述一下你的思考过程。

这里额外给定经验前提：盒中只有有限多个球，我们确实把它们全部取出，并可靠地观察到取出的三个球都是白色。这个练习只演示如何从明确前提推理；后文不会把观察当作数学公理。

解答：

1. 先假设结论是错的，也就是假设“盒子里有红色的球”。
2. 若盒子里有红球，那么把所有球全部取出后，取出的球中必定有红球。
3. 结果我把盒子里的球全倒出来，发现只有三个白球，根本没有红球。
4. 这就产生了矛盾：“盒子里有红球”且“倒出来的全都没红球”。
5. 这个矛盾说明第一步的假设是错的。
6. 最终证明：“盒子里没有红色的球”是对的。

1.7 本章结语

现在，回过头看看你掌握了什么。

你能熟练运用 $\neg, \wedge, \vee, \rightarrow, \forall, \exists$ 这些魔法符号，你还掌握了逆否之剑与反证法。

你没有学任何算术，也没有画任何图形。但你已经拿到了宇宙最高法庭的法官之槌。无论未来我们要创造怎样光怪陆离的数字、怎样奇妙的空间，它们都必须绝对服从今天写下的这些“对与错的基础法则”。

一切准备就绪。在下一章，我们将运用这些逻辑魔法，去创造一个能装下世间万物的容器，并从“什么都没有”的深渊中，召唤出宇宙的第一个实体。

2 万物的分类与打包 —— 创造世界的容器

在上一章，你拿到了宇宙最高法庭的法官之槌，掌握了判断对错的基础法则。

但是，现在的宇宙依然是一片虚无。只有法则，没有受法则约束的实体。就像你定好了完美的交通规则，但世界上连一条马路和一辆车都没有。

这一章，我们要凭空捏造出数学宇宙中最基础的“实体”。我们将发明一种无所不包的“容器”，把虚无打包，变出万事万物；我们还要设定容器之间交流的规则，为下一章创造“数字”搭建好完美的舞台。

2.1 朴素的口袋与思维的崩塌

如果我们要创造事物，最直观的想法就是“打包”。

【公理语言】 集合论使用第一章的经典一阶逻辑与等号，并只增加一个非逻辑关系符号：**属于**（记为 $\in$ ）。变量 $x, y, z, A, B, \dots$ 全都在“所有集合”组成的论域中取值；大小写只帮助阅读，没有逻辑上的身份差别。“元素”和“集合”也不是两种不同材料，而是同一集合在一句话中扮演的不同角色。

严格地说，这门一阶语言的项只有变量，原子公式只有 $x = y$ 与 $x \in y$ ；其余公式由第一章的联结词和量词有限次生成。 $\emptyset, \{x, y\}, \mathcal{P}(A)$ 等符号会在公理保证相应集合唯一存在后，作为方便阅读的缩写引入。以后出现“所有集合构成的类”或 $x \mapsto \dots$ 这样的**类记号**时，它只缩写一条公式，不自动代表某个集合。

想象你拥有一个看不见的魔法口袋，你可以把具有某种相同性质的东西装进去。这个口袋，我们称为**集合**（记为大写字母 A, B, C ）。口袋里的东西，称为**元素**（记为小写字母 x, y, z ）。

如果 x 被装进了口袋 A 里，我们就写成： $x \in A$ （读作： x 属于 A ）。

反之，利用第一章的“非”魔法，如果 x 不在 A 里，就是 $\neg(x \in A)$ ，为了简便，我们写成 $x \notin A$ 。

【记号】对任意两个项 s, t ， $s \neq t$ 是 $\neg(s = t)$ 的缩写；它不引入第二种“不相等”关系。

【定义】若 $\forall x (x \in A \rightarrow x \in B)$ ，就写 $A \subseteq B$ ，称 A 是 B 的子集。若还满足 $A \neq B$ ，写 $A \subsetneq B$ ，称为真子集。

【记号】 $\forall x \in A P(x)$ 是 $\forall x (x \in A \rightarrow P(x))$ 的缩写； $\exists x \in A P(x)$ 是 $\exists x (x \in A \wedge P(x))$ 的缩写。

连续写几个变量时表示重复使用同一种缩写；例如 $\forall x, y \in A P(x, y)$ 表示 $\forall x \in A \forall y \in A P(x, y)$ 。

【记号】 $\exists! y P(y)$ 表示“存在唯一的 y 满足 P ”，它是

$$\exists y (P(y) \wedge \forall z (P(z) \rightarrow z = y))$$

的缩写；其中 z 要选成不会捕获别处变量的新变量。

第一次危机降临：

假如我们朴素地允许“任意规则都能制造一个集合”，那么只要想出一条规则，似乎就能凭空造出一个口袋。比如，“把所有红色的东西装进一个口袋”，或者“把所有长翅膀的东西装进一个口袋”。

这听起来似乎很合理，但这样的许可会立刻引出矛盾。

让我们来讲一个“理发师的故事”。在一个小镇上，有一位理发师；他自己也是镇上的人。他定下一条极其严格的规则：“对于镇上的每个人，我给他剪头发，当且仅当他不给自己剪头发。”

那么问题来了：理发师要不要给自己剪头发？

- 如果他不给自己剪，那他就是一个“不给自己剪头发的人”，根据铁律，他必须给自己剪头发。矛盾！
- 如果他给自己剪，那他就是一个“给自己剪头发的人”，根据铁律，他绝对不能给自己剪头发。再次矛盾！

这个故事是著名的**罗素悖论**的生活化版本。在数学宇宙里，我们可以利用第一章的逻辑法则，设计一个与理发师规则相似的超级口袋 R 。

如果任意性质都能制造集合，我们就可以写出超级口袋 R 的准入规则： $x \in R \leftrightarrow x \notin x$ ，也就是“ x 属于 R ，当且仅当 x 不属于自己”。

致命的问题来了：超级口袋 R 自己，属不属于 R 呢？（即 $R \in R$ 是【真】还是【假】？）我们来推导一下：

- 假设 $R \in R$ 为【真】：既然 R 在 R 里面，说明 R 满足了准入规则。但准入规则明明是“不包含自己”。这就意味着 $R \notin R$ 。矛盾！
- 假设 $R \notin R$ 为【真】：既然 R 不包含自己，那它完全符合准入规则！既然符合规则，它就必须被装进 R 里。这就意味着 $R \in R$ 。再次矛盾！

于是 $R \in R$ 与 $R \notin R$ 都会导出矛盾。问题不在经典逻辑，而在“任意性质都能制造集合”这个未经许可的假设：这样的 R 不能作为集合存在。

【思考训练】

除了理发师，你能再想出一个类似“罗素悖论”的生活例子吗？

解答：

比如一张纸条，上面只写了一句话：“这张纸条上写的话是假的。”

如果这句话是真的，它就为假；如果它为假，它又为真。这叫说谎者悖论。它提醒我们不能不加限制地在语言内部制造“谈论自身真值”的句子；它不是经典命题逻辑本身已经推出了矛盾。这个故事只作【直觉】类比，后文不把它当作集合论证明。

2.2 造物主的咒语 —— 公理系统

罗素悖论告诉我们一个惨痛的教训：绝对不能随便打包！

为了挽救宇宙，我们必须明确列出允许哪些集合存在。这些起始规则不叫“不证自明”，而叫**公理**：我们公开采用它们，并研究能从中推出什么。下面先列出 ZF 公理；等“映射”获得正式定义后，再补上选择公理，合称 ZFC。

1. 平等法则（外延公理）

到底什么才算“同一个口袋”？

宪法规定：只要两个口袋里装的东西一模一样，不管它们外表怎样，它们就是同一

个口袋：

$$\forall A \forall B ((\forall x (x \in A \leftrightarrow x \in B)) \rightarrow A = B).$$

比如，口袋 A 里装了苹果和香蕉，口袋 B 里装了香蕉和苹果。在集合的世界里，顺序根本不重要，因为里面的东西完全一样，所以口袋 A 和口袋 B 是完全相等的 ($A = B$)。

2. 无中生有（空集公理）

现在的宇宙还是空无一物，我们去哪里找第一个元素呢？

空集公理规定：存在一个合法的口袋，里面没有任何元素：

$$\exists E \forall x (x \notin E).$$

根据平等法则，里面什么都没有的口袋只能有一个（因为它们装的东西一模一样——都是“没有”）。我们给这个绝对虚无的口袋赋予一个神圣的符号：**空集**（记为 $\emptyset$ ）。

这个唯一性可以很短地证明。若 E, F 都没有元素，那么对任意 x ， $x \in E$ 与 $x \in F$ 都为假，所以 $x \in E \leftrightarrow x \in F$ 为真。于是两者拥有完全相同的元素，外延公理给出 $E = F$ 。因此符号 $\emptyset$ 不会因我们挑了不同的“空口袋”而指向不同对象。

奇迹就在这里发生了！虽然 $\emptyset$ 里面什么都没有，但 $\emptyset$ 本身是一个合法存在的口袋实体！我们从“无”中，创造出了“有”。

下面的配对公理保证我们可以把 $\emptyset$ 作为元素装进另一个新口袋，得到 $\{\emptyset\}$ 。

注意， $\emptyset$ 和 $\{\emptyset\}$ 绝不是同一个东西！前者是空的，后者里面装着一个叫“空集”的元素。就像“一个空盒子”和“一个装着空盒子的盒子”是截然不同的。

3. 配对公理

对于任意集合 a, b ，存在集合 P ，使

$$\forall x (x \in P \leftrightarrow (x = a \vee x = b)).$$

外延公理保证 P 唯一，我们把它记作 $\{a, b\}$ 。令 $a = b$ 就得到单元素集合 $\{a\}$ 。

【易混点：属于与包含】 $\emptyset \in \{\emptyset\}$ 为真，因为单元素集合的那个元素正是 $\emptyset$ ；同时 $\emptyset \subseteq \{\emptyset\}$ 也为真，因为根本找不到一个属于 $\emptyset$ 却不属于 $\{\emptyset\}$ 的对象。前一句把 $\emptyset$ 当作一个元素，后一句比较两个集合的全部成员，理由完全不同。

【思考训练】 $\emptyset \in \emptyset$ 与 $\emptyset \subseteq \emptyset$ 各是真还是假？

解答：前者为假，因为空集没有任何元素；后者为真，因为“若 $x \in \emptyset$ ，则 $x \in \emptyset$ ”的前件永远为假，所以不存在反例。这正是第一章讲过的“空虚的真”。

4. 并集公理

对于任意集合 A ，存在集合 U ，使

$$\forall x(x \in U \leftrightarrow \exists Y(Y \in A \wedge x \in Y)).$$

这个唯一的 U 记作 $\bigcup A$ 。二元并集定义为 $A \cup B = \bigcup\{A, B\}$ 。

5. 幂集公理

对于任意集合 A ，存在集合 P ，使

$$\forall X(X \in P \leftrightarrow X \subseteq A).$$

这个唯一的 P 记作幂集 $\mathcal{P}(A)$ ，其元素恰好是 A 的全部子集。

6. 分离公理模式

下列两个公理模式中，参数变量 $p_1, \dots, p_k$ 都要在最外层加上全称量词。为了让式子易读，正文省略这串外层量词；正式公理取所显示公式的**全称闭包**，所以每一条公理都是没有自由变量的闭公式。

这里不是说“任何说得通的性质”都可以使用。对集合论一阶语言 $\{\in, =\}$ 的**每一个**公式 $\varphi(x, p_1, \dots, p_k)$ （允许有限多个集合参数 $p_1, \dots, p_k$ ），都有一条公理：

$$\forall A \exists B \forall x(x \in B \leftrightarrow (x \in A \wedge \varphi(x, p_1, \dots, p_k))).$$

这样的“每个公式各有一条”叫作**公理模式**。所得唯一集合记作 $\{x \in A : \varphi(x, p_1, \dots, p_k)\}$ 。它只允许从已有集合中筛选，不能制造“所有满足 φ 的对象”这种无限制口袋，所以不会重新制造罗素集合。

为什么多出边界 A 就躲开了罗素悖论？给定任意集合 A ，分离公理只制造

$$R_A = \{x \in A : x \notin x\}.$$

【引理】 $R_A \notin A$ 。

证明：若反过来假设 $R_A \in A$ ，先把 $x = R_A$ 代入分离定义，严格得到

$$R_A \in R_A \leftrightarrow (R_A \in A \wedge R_A \notin R_A).$$

再使用反设中的 $R_A \in A$ ，这个等价式才化简为

$$R_A \in R_A \leftrightarrow R_A \notin R_A,$$

若左边为真，右边就迫使它为假；若左边为假，则右边为真，反向蕴含又迫使左边为真。排中律保证两种情况必居其一，所以一定出现矛盾。因此最初的假设不成立，故 $R_A \notin A$ 。证毕。

这里没有推出宇宙崩溃；我们只推出这个新集合不在原来的边界 A 中。无界的“所有集合”不能充当 A ，正是限制起作用的地方。

7. 替换公理模式

对集合论一阶语言的每一个公式 $\varphi(x, y, p_1, \dots, p_k)$ ，都有一条公理

$$\forall A \left[(\forall x \in A \exists! y \varphi(x, y, p_1, \dots, p_k)) \rightarrow \exists B \forall y (y \in B \leftrightarrow \exists x \in A \varphi(x, y, p_1, \dots, p_k)) \right].$$

因此，一个公式若给 A 中每个 x 唯一指定一个 y ，这些 y 能被收集成集。这里也允许有限多个集合参数，所以替换同样是一套公理模式。

8. 永恒的许可（无穷公理）

先把由公式唯一描述的集合 $x \cup \{x\}$ 缩写为 $S(x)$ 。此时 S 只是对任意集合都能使用的类操作记号，还不是一个已经给定定义域和陪域的映射。无穷公理规定：存在集合 I ，使得

$$\emptyset \in I \quad \text{且} \quad \forall x \in I (S(x) \in I).$$

它保证至少存在一个对这个操作封闭的集合；下一章会从中筛出最小的那个自然数集合，再把 S 限制成真正的映射 $\mathbb{N} \rightarrow \mathbb{N}$ 。

9. 正则公理（基础公理）

每个非空集合 A 都含有某个 x ，使得没有对象同时属于 x 和 A ：

$$\forall A(A \neq \emptyset \rightarrow \exists x(x \in A \wedge \neg \exists y(y \in x \wedge y \in A))).$$

它可以推出不存在 $x \in x$ ；等自然数与映射构造完毕后，还能推出不存在由映射给出的无穷向下成员链。后文只使用这些可由公理证明的推论。

【定义】交集 $A \cap B = \{x \in A : x \in B\}$ ；集合差 $A \setminus B = \{x \in A : x \notin B\}$ 。它们的存在由分离公理保证。

【思考训练】

请仔细观察，口袋 $A = \{\emptyset\}$ ，口袋 $B = \{\emptyset, \{\emptyset\}\}$ 。它们是同一个口袋吗？口袋 B 里有几个元素？

解答：

它们不是同一个口袋。用讲故事的计数词说，口袋 A 里只有一个元素（一个空盒子）；口袋 B 里有两个元素，分别是“一个空盒子”和“一个装着空盒子的盒子”。第三章会把“一个”“两个”正式构造成自然数，并证明这里的计数。

2.3 跨越维度的配对——笛卡尔积与映射

有了合法的口袋（集合），我们要让不同的口袋之间产生联系。

1. 秩序的诞生：有序对

刚才我们说了，在普通的口袋 $\{x, y\}$ 里，元素是没有顺序的，先抓出 x 和先抓出 y 没有区别。

但是，如果要表达“谁先谁后”，或者排队时的“第一名和第二名”，我们必须发明一种“有顺序的口袋”。

【定义】我们把有顺序的对子 (x, y) 定义为集合 $\{\{x\}, \{x, y\}\}$ 。配对公理保证这些集合存在。

你看懂这个魔法了吗？为了证明 x 是排在前面的，我们专门给 x 发了一个“贵宾包厢” $\{x\}$ 。当我们打开外面这个大口袋时，只要发现谁拥有自己单独的包厢，我们就立刻知道：“哦！原来 x 是排在前面的那一个！”

不需要引入任何新的基础概念，仅仅通过口袋的嵌套，我们就在没有时间概念的宇宙中，硬生生地创造出了“顺序”！

2. 口袋的交叉：笛卡尔积 ($\times$)

假设有两个集合 A 和 B 。我们从 A 里任意抽出一个元素 x ，从 B 里任意抽出一个元素 y ，把它们组成一个有顺序的对子 (x, y) 。

【定义】 A 和 B 的笛卡尔积是

$$A \times B = \{(x, y) : x \in A \text{ 且 } y \in B\}.$$

【定理（笛卡尔积存在）】由配对、并集、幂集和分离公理，上式确实定义了一个集合。

这个存在性证明其实很短。令 $C = A \cup B$ 。若 $x \in A, y \in B$ ，则 $\{x\}$ 与 $\{x, y\}$ 都是 C 的子集，所以有序对

$$(x, y) = \{\{x\}, \{x, y\}\}$$

属于 $\mathcal{P}(\mathcal{P}(C))$ 。于是，只需从已经存在的 $\mathcal{P}(\mathcal{P}(C))$ 中分离出这样的成员：存在某个 $x \in A$ 和某个 $y \in B$ ，使该成员等于 (x, y) 。所得集合就是 $A \times B$ 。

【定理（有序对的唯一性，证明略）】 $(x, y) = (u, v)$ 当且仅当 $x = u$ 且 $y = v$ 。因此这个集合编码确实保存了两个位置。

【例子】取 $A = \{\emptyset\}$ ， $B = \{\emptyset, \{\emptyset\}\}$ ，直接按定义列出

$$A \times B = \{(\emptyset, \emptyset), (\emptyset, \{\emptyset\})\}.$$

第一个位置只能来自 A ，第二个位置只能来自 B ；这就是“有序”真正保存的信息。

请注意！这里的乘号 $\times$ 和算术里的乘法没有任何关系（因为我们还没发明数字呢！）。它只是一个用来“交叉配对”的魔法符号。这是我们在宇宙中第一次将两个独立的集合编织成了一张网。

3. 宇宙中最核心的动词：映射（函数的本源）

有了网，我们就能制定“变换的规则”。

【定义】先把有序三元组 (A, B, G) 约定为嵌套有序对 $((A, B), G)$ 。从集合 A 到集合 B 的**映射**是一份三元数据

$$f = (A, B, G), \quad G \subseteq A \times B, \quad \forall x \in A \exists! y (y \in B \wedge (x, y) \in G).$$

A 叫 f 的**定义域**, 记作 $\text{dom}(f)$; B 叫**陪域**, 记作 $\text{cod}(f)$; G 叫**函数图**, 记作 $\text{Graph}(f)$ 。对每个 $x \in A$, 把上述唯一满足 $(x, y) \in G$ 的 y 定义为 $f(x)$, 并写 $f : A \rightarrow B$ 。

【术语约定】本书把“函数”作为“映射”的同义词。强调集合之间的对应时常说“映射”，强调输入与输出的变化时常说“函数”，但两者指的是同一种三元数据。

把定义域和陪域列入数据十分重要：即使两个映射的函数图含有相同的有序对，它们的陪域不同，也不是同一个映射。例如从空集到 B 的函数图总是空集，但陪域仍必须记录下来。

【定义】映射 $f : A \rightarrow B$ 的**值域**是

$$\text{ran}(f) = \{f(x) : x \in A\}.$$

替换公理保证它是集合，并且 $\text{ran}(f) \subseteq B$ 。陪域 B 是映射数据的一部分，值域则是实际被取到的值；两者不可混淆。

【定义】对映射 $f : A \rightarrow B$:

- 若对任意 $x_1, x_2 \in A$, $f(x_1) = f(x_2)$ 都蕴含 $x_1 = x_2$, 就称 f 为**单射**;
- 若对任意 $y \in B$, 都存在 $x \in A$ 使 $f(x) = y$, 就称 f 为**满射**;
- 既是单射又是满射的映射称为**双射**。

【同一函数图，不同映射】令

$$A = B = \{\emptyset\}, \quad C = \{\emptyset, \{\emptyset\}\}, \quad G = \{(\emptyset, \emptyset)\}.$$

三元数据 $f = (A, B, G)$ 与 $g = (A, C, G)$ 都是映射，而且函数图完全相同；但它们的陪域不同，所以 $f \neq g$ 。此外， $f : A \rightarrow B$ 是满射，而 $g : A \rightarrow C$ 不是满射，因为

陪域中的 $\{\emptyset\}$ 没有被取到。这个例子说明：陪域不仅是包装标签，它会改变映射的性质。

【定义】若 $f : A \rightarrow B$ 、 $g : B \rightarrow C$ ，复合映射 $g \circ f : A \rightarrow C$ 的函数图是

$$\{(x, z) \in A \times C : \exists y \in B ((x, y) \in \text{Graph}(f) \wedge (y, z) \in \text{Graph}(g))\};$$

于是 $(g \circ f)(x) = g(f(x))$ 。恒等映射 $\text{id}_A : A \rightarrow A$ 的函数图是 $\{(x, x) : x \in A\}$ ，于是 $\text{id}_A(x) = x$ 。分离与替换公理保证这些函数图存在。

【定理（复合与逆映射，证明略）】映射复合满足结合律，恒等映射是复合单位元。若 $f : A \rightarrow B$ 是双射，则存在唯一映射 $f^{-1} : B \rightarrow A$ 使 $f^{-1} \circ f = \text{id}_A$ 且 $f \circ f^{-1} = \text{id}_B$ ；反过来，拥有这种逆映射的映射必为双射。

【定理（映射外延性，证明略）】若 $f, g : A \rightarrow B$ ，则 $f = g$ 当且仅当对每个 $x \in A$ 都有 $f(x) = g(x)$ 。若定义域或陪域不同，则即使在共同输入上的值碰巧相同，也不能用这条定理判定为同一个映射。

【公理（选择公理）】若 $\mathcal{A}$ 是由非空集合组成的集合，则存在映射 $f : \mathcal{A} \rightarrow \bigcup \mathcal{A}$ ，使每个 $X \in \mathcal{A}$ 都满足 $f(X) \in X$ 。ZF 加上这条公理就是本书采用的 ZFC。前半册的具体构造通常不需要它；公开列出是为了让后文的一般选择不成为暗中增加的规则。

【定义】以集合 I 为指标的集合族，是某条映射 $F : I \rightarrow B$ ，其中 B 是集合；把 $F(i)$ 记作 A_i 。替换公理先保证值域 $\{A_i : i \in I\}$ 是集合，再定义

$$\bigcup_{i \in I} A_i = \bigcup \{A_i : i \in I\}.$$

由并集公理可知它是集合，并且 x 属于它当且仅当 $\exists i \in I (x \in A_i)$ 。若 $I \neq \emptyset$ ，取 $i_0 \in I$ ，定义

$$\bigcap_{i \in I} A_i = \{x \in A_{i_0} : \forall i \in I (x \in A_i)\}.$$

分离公理保证它是集合，外延公理保证结果不依赖 i_0 的选择。没有指定共同大集合时，我们不定义空指标族的交集，因为“所有集合”不是集合。

映射的定义还意味着：集合 A 里的每个元素都必须恰好射出一支箭，既不能弃权，也不能同时射出两支箭。不过，集合 B 里的靶子可以被多支箭同时射中，也可以完

全没有箭射中。

单射、满射和双射这三个概念可以画成射箭游戏：

- **单射**：不同弓箭手不会射中同一个靶子。
- **满射**：陪域里的每个靶子都被射中。
- **双射**：两条规则同时成立。这意味着集合 A 里的弓箭手和集合 B 里的靶子一个配一个，既不争抢，也不落单。

记住“双射”这个概念！即使我们现在还没有定义怎样数数，只要两个口袋之间能建立双射，就可以明确地说：这两个口袋里装的东西“一样多”。

【思考训练】

集合 A 里有几只不同的小狗，集合 B 里有几根不同的骨头。每只小狗都恰好叼起一根骨头，不同小狗叼的是不同骨头，而且没有骨头剩下。请问，这种配对属于单射、满射还是双射？

解答：

属于双射，也同时属于单射和满射。因为不同小狗叼的是不同骨头（单射），而且所有骨头都被叼走了（满射），所以这种配对是双射。这说明小狗的数量和骨头的数量“一样多”。

2.4 数学家最伟大的魔法——极度抽象的商集

这一节，我们将传授给你未来填补宇宙漏洞的最强法术：**求同存异**。

世界上没有两片完全相同的树叶。但在某些时候，为了建立规则，我们必须“强行”认为它们是相同的。比如在一堆玻璃球里，有大有小，有新有旧，但只要它们都是红色的，我们就强行认为它们“等价”。

在数学中，如何合法地“无视差异”？先要定义什么叫“关系”。

【定义】 集合 A 上的二元关系 R 是 $A \times A$ 的一个子集。写 xRy ，就是写 $(x, y) \in R$ ；因此这里始终默认 $x, y \in A$ 。关系本身仍是一个集合，不是从语言外突然出现的新材料。

【定义】 集合 A 上的二元关系 $\sim$ 若同时满足下列三个公式，就叫作 A 上的**等价关系**：

1. 自反性: $\forall x \in A (x \sim x)$ 。
2. 对称性: $\forall x, y \in A (x \sim y \rightarrow y \sim x)$ 。
3. 传递性: $\forall x, y, z \in A ((x \sim y \wedge y \sim z) \rightarrow x \sim z)$ 。

三个量词的范围 A 都不可省略: 换一个底集, 同一个有序对集合可能不再是那里的等价关系。

终极法术: 商集

请注意, 这里的“商”字和买卖东西的商人无关, 也和算术里的除法无关。它代表的是一种“分组打包”的魔法。

一旦在 A 上确立了等价关系 $\sim$, 对每个 $x \in A$ 定义它的等价类

$$[x] = \{y \in A : y \sim x\}.$$

- 比如我们制定等价关系为“颜色相同”。我们就把所有红色的球装进一个盒子, 把所有蓝色的球装进另一个盒子。这些盒子就是等价类。
- 根据自反、对称和传递三条性质, 每个元素都属于自己的等价类, 并且任意两个等价类要么相等, 要么没有共同元素。

【定义】商集是

$$A/\sim = \{[x] : x \in A\}.$$

替换公理保证这些等价类可以收集成集合。

【最简单的商集例子】由分离公理,

$$\Delta_A = \{(x, y) \in A \times A : x = y\}$$

是 A 上的二元关系。规定 $x \sim y$ 当且仅当 $(x, y) \in \Delta_A$, 也就是当且仅当 $x = y$ 。等号规则立即给出自反、对称和传递性, 所以这是等价关系。对每个 $x \in A$,

$$[x] = \{y \in A : y = x\} = \{x\},$$

商集 $A/\sim$ 就是由各个单元集合 $\{x\}$ 组成的集合。商集不一定真的把不同元素粘在一起; 它只按照给定的等价关系分组。

【定义】 A 的一个**分割**是一个集合 $\mathcal{C} \subseteq \mathcal{P}(A)$, 满足: $\mathcal{C}$ 的每个成员都非空; $\bigcup \mathcal{C} = A$; 而且 $\mathcal{C}$ 中任意两个不同成员不相交。

【定理 (等价类分割定理, 证明略)】若 $\sim$ 是等价关系, 则 $A/\sim$ 是 A 的一个分割。反过来, 若 $\mathcal{C}$ 是 A 的分割, 规定 $x \sim y$ 当且仅当存在 $C \in \mathcal{C}$ 使 $x, y \in C$, 便得到 A 上的等价关系, 其等价类恰是 $\mathcal{C}$ 的成员。

这为什么是伟大的魔法?

因为通过商集, 我们完成了一次“视角的飞跃”。我们强行把盒子内部的所有细微差异彻底抹除, 不再关心里面是哪颗具体的红球, 而是把“装着红球的盒子”变成了一个绝对不可分割的“新实体”。

在未来的章节你就会看到, 当我们因为没有“负数”和“分数”而陷入绝境时, 正是这种把旧概念捆绑打包的“商集魔法”, 帮我们无中生有地造出了完整的数字帝国!

【思考训练】

给每个小朋友**恰好指定一个**水果标签, 并规定: 两人等价, 当且仅当标签相等。由于等号满足自反、对称和传递性, 这确实定义了等价关系。

大口袋里有五个小朋友: 小红 (苹果标签)、小明 (香蕉标签)、小刚 (苹果标签)、小华 (橘子标签)、小丽 (香蕉标签)。

如果用这个等价关系来制造“商集”, 最后这个商集口袋里, 会装有几个硬纸盒 (等价类)?

解答:

会装有三个硬纸盒。

第一个硬纸盒装着: 小红和小刚 (苹果组)。

第二个硬纸盒装着: 小明和小丽 (香蕉组)。

第三个硬纸盒装着: 小华 (橘子组)。

商集里面有这三个独立的盒子。

2.5 本章结语

这一章, 你经历了思维的崩塌与重建。

你亲手颁布了宇宙宪法, 从绝对的虚无中召唤出了空集 $\emptyset$, 利用口袋的嵌套创造了顺序, 利用笛卡尔积和映射编织了联系的网, 最后掌握了求同存异的“商集”大法。

万物的容器已经准备就绪。所有的魔法阵都已经刻画完毕。

深呼吸。在下一章，我们将踏入整本书最令人激动的一步——我们要仅凭这些空荡荡的魔法口袋，去生生地造出整个宇宙的骨架：**数字**。

Xulei Sun

3 造出数字 —— 从零到无限的阶梯

在上一章，你凭借“空集公理”和“无穷公理”，从绝对的虚无中凭空捏造出了合法存在的“口袋”（集合），并学会了用“映射”和“商集”来建立规则。

但这还远远不够。现在的宇宙虽然有了容器，却像是一个无边无际、毫无秩序的大仓库。如果我们要衡量宇宙，如果要指定谁先谁后，我们必须拥有数字。

请注意，在这个纯粹逻辑的宇宙里，不存在“数字本来就在那里”这种前提。我们不能借助手指数数，因为这里没有“手指”；也不能预先使用加减乘除。任何对象或运算，都必须先得到严格定义。

准备好迎接造物主般极致的浪漫了吗？我们将见证，如何仅仅用上一章造出来的“空荡荡的魔法口袋”，像搭积木一样，搭建起通往无限的数字阶梯。

3.1 无中生有的阶梯 —— 皮亚诺性质与自然数 $\mathbb{N}$

我们要造的第一批实体，是用来标记顺序的符号。我们称它们为“自然数”。

但是在我们的宇宙里，目前唯一的合法实体只有口袋（集合）。怎么用口袋来造数字呢？

数学家提出了一个天才的想法：用特定的集合来**编码**自然数。这里不能说“先看口袋里有多少个元素”，因为“多少个”正是我们还在构造的概念。以后我们会证明：编码 n 的集合恰好拥有与 n 对应的元素数目。

第二章已经用并集公理定义了 $A \cup B$ 。现在我们使用后继的类操作记号 $S(x) = x \cup \{x\}$ 。它对每个集合 x 都唯一指定一个集合，但在选定定义域和陪域以前，按照第二章的约定还不能称为映射。

零的诞生：

我们用哪个集合来编码“0”呢？上一章已经构造出一个没有任何元素的集合——空集 $\emptyset$ 。

因此我们作出定义：**数字 0 就是空集 $\emptyset$ 。**

后继魔法：

怎么从 0 变成下一个数呢？我们需要一个“往上爬一格”的动作。利用上一章的集合知识，我们使用已经定义好的**后继操作**（记为 $S(x)$ ）。

它的规则是：把口袋 x 里的所有东西，加上口袋 x 本身（把它当作一个硬纸盒），全部合并倒进一个新口袋里。

用集合的合并符号写出来： $S(x) = x \cup \{x\}$ 。

现在，请瞪大眼睛，见证奇迹的时刻到了：

1. 造出 1：对 0 使用后继魔法。

$$1 = S(0) = 0 \cup \{0\}$$

因为 0 本身就是空集 $\emptyset$ ，所以这就等于把“什么都没有”和“一个装着空集的盒子 $\{\emptyset\}$ ”合并在一起。

结果是： $1 = \{\emptyset\}$ 。

数字 1 的唯一元素是空集。稍后我们会正式定义“恰好有 n 个元素”。

2. 造出 2：对 1 使用后继魔法。

$$2 = S(1) = 1 \cup \{1\}$$

把数字 1（里面装着 $\emptyset$ ）和装着数字 1 的盒子 $\{1\}$ 合并。

结果是： $2 = \{\emptyset, \{\emptyset\}\}$ 。

把它们替换成数字，其实就是 $2 = \{0, 1\}$ 。

数字 2 的元素恰好是 0 和 1。

3. 造出 3：对 2 使用后继魔法。

$$3 = S(2) = 2 \cup \{2\}$$

结果是 $3 = \{0, 1, 2\}$ 。它的元素恰好是已经排在它前面的 0、1、2。

在这个疯狂的俄罗斯套娃般的嵌套中，每一个数字不仅代表了它所在的高度，它本身就包含着之前创造出来的所有数字！

自然数集合的正式定义

【定义】若集合 J 满足 $\emptyset \in J$ ，并且 $\forall x \in J (S(x) \in J)$ ，就称 J 为**归纳集**。无穷公理保证至少有一个归纳集 I 。在这个 I 中定义

$$\mathbb{N} = \{n \in I : \forall J \in \mathcal{P}(I) (J \text{ 是归纳集} \rightarrow n \in J)\}.$$

“ J 是归纳集”是刚才公式的缩写，所以分离公理保证 $\mathbb{N}$ 是集合。

【定理（最小归纳集，证明略）】 $\mathbb{N}$ 是归纳集；对每个归纳集 K 都有 $\mathbb{N} \subseteq K$ ；不同的起始归纳集 I 会给出同一个 $\mathbb{N}$ 。因此它是所有归纳集合中唯一的最小者，只含从 $0 = \emptyset$ 反复取后继所必需的成员。

【思考训练】若 K, L 都是归纳集，为什么 $K \cap L$ 仍是归纳集？

解答：因为 0 同时属于 K, L ，所以 $0 \in K \cap L$ 。若 $x \in K \cap L$ ，则 $x \in K$ 且 $x \in L$ ；两者都对后继封闭，于是 $S(x) \in K$ 且 $S(x) \in L$ ，故 $S(x) \in K \cap L$ 。这说明不断取归纳集的共同部分不会丢掉起点和后继，也解释了“最小归纳集”为什么可以被筛出来。

皮亚诺的五条性质

【定理（皮亚诺性质，证明略）】由第二章的集合公理和上面的定义可以推出：

1. **起点**： $0 \in \mathbb{N}$ 。
2. **后继阶梯**：若 $n \in \mathbb{N}$ ，则唯一确定的 $S(n)$ 也属于 $\mathbb{N}$ ；所以把类操作限制在 $\mathbb{N}$ 上，便得到真正的映射 $S_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{N}$ 。以后在自然数语境中仍简写为 S 。
3. **单行道**：不存在 $n \in \mathbb{N}$ 使 $S(n) = 0$ 。
4. **不撞车**：若 $S(m) = S(n)$ ，则 $m = n$ 。
5. **归纳法**：若 $A \subseteq \mathbb{N}$ ，且 $0 \in A$ ，并且 $n \in A \Rightarrow S(n) \in A$ ，那么 $A = \mathbb{N}$ 。

第五条铁律极其重要。【直觉】你可以把它想象成推倒一列不断延伸的多米诺骨牌：只要你能保证推倒第一块（数字 0 ），并且保证只要前一块倒下，就一定会砸倒下一块（后继），归纳法就断言每一块都会倒下。

这就是我们未来证明许多宇宙真理的终极武器——**数学归纳法**。

因此，上面定义的集合正是由 0 和后继生成的自然数集 $\mathbb{N}$ 。在本书的集合论路线中，皮亚诺性质是定理，而不是另一套额外公理。

【引理（自然数向下封闭，证明略）】若 $n \in \mathbb{N}$ 且 $k \in n$ ，则 $k \in \mathbb{N}$ 。因此每个 $n \in \mathbb{N}$ 都满足 $n \subseteq \mathbb{N}$ ，而 $S(n) = n \cup \{n\}$ 也满足 $S(n) \subseteq \mathbb{N}$ 。这个结论可直接对 n 使用归纳法证明。

【定义】若存在双射 $n \rightarrow A$ ，其中 $n \in \mathbb{N}$ ，就说集合 A **恰好有 n 个元素**。若对某个 n 满足此条件，就称 A **有限**；若 A 不是有限集，就称 A **无限**。

【定理（自然数的大小，证明略）】映射 $\text{id}_n : n \rightarrow n$ 定义为 $\text{id}_n(x) = x$ ，它是双射，所以冯·诺伊曼自然数 n 恰好有 n 个元素；并且 $S(n) = n \cup \{n\}$ 恰好比 n 多出一个元素 n 。这使前面的“口袋大小”观察成为构造后的定理，而不是定义时的循环前提。

【定理（有限基数唯一，证明略）】若 $m, n \in \mathbb{N}$ ，并且同一集合 A 分别与 m, n 双射，则 $m = n$ 。因此“恰好有 n 个元素”中的自然数 n 不会因配对方法不同而改变。

【定理（有限鸽笼原理）】对每个 $n \in \mathbb{N}$ ，不存在从 $S(n)$ 到 n 的单射。

证明：对 n 使用归纳法。当 $n = 0$ 时，不存在从非空集合 $S(0) = \{0\}$ 到空集的映射。现在假设不存在单射 $S(n) \rightarrow n$ ，却存在单射

$$g : S(S(n)) \rightarrow S(n).$$

记 $y = g(S(n))$ 。在 $S(n)$ 上定义映射 τ ：若 $y = n$ ，就取恒等映射；若 $y \neq n$ ，则规定

$$\tau(z) = \begin{cases} n, & z = y, \\ y, & z = n, \\ z, & z \neq y \text{ 且 } z \neq n. \end{cases}$$

第二种情形中直接分情况可得 $\tau \circ \tau = \text{id}_{S(n)}$ ，所以 τ 也是双射。于是 $(\tau \circ g)(S(n)) = n$ 。对每个 $x \in S(n)$ ，第二章正则公理的推论给出 $S(n) \notin S(n)$ ，所以 $x \neq S(n)$ ；再由 $\tau \circ g$ 的单射性可得 $(\tau \circ g)(x) \neq n$ 。因此，把 $\tau \circ g$ 限制在 $S(n)$ 上便得到单射

$$S(n) \rightarrow S(n) \setminus \{n\} = n,$$

其中 $S(n) \setminus \{n\} = n$ 还使用了第二章正则公理的推论 $n \notin n$ 。于是所得单射与归纳假设矛盾。故结论对 $S(n)$ 也成立，由归纳法得证。

【定理（自然数集是无限集）】 $\mathbb{N}$ 不是有限集。

证明：若 $\mathbb{N}$ 有限，就存在某个 $n \in \mathbb{N}$ 和双射 $f : n \rightarrow \mathbb{N}$ 。因为 $S(n) \subseteq \mathbb{N}$ ，由 $\iota(x) = x$ 定义的包含映射 $\iota : S(n) \rightarrow \mathbb{N}$ 是单射；于是

$$f^{-1} \circ \iota : S(n) \rightarrow n$$

也是单射，与有限鸽笼原理矛盾。故 $\mathbb{N}$ 无限。证毕。

【定义】若存在双射 $\mathbb{N} \rightarrow A$ ，就称 A **可数无限**；有限或可数无限的集合统称**至多可数**。因为双射及其逆映射可以复合，而 $\mathbb{N}$ 已证明为无限集，所以可数无限集不可能同时是有限集。

【思考训练】

请按照上面的规则，写出数字 4 的集合表示。

解答：

$4 = S(3) = 3 \cup \{3\}$ 。因为 $3 = \{0, 1, 2\}$ ，所以 $4 = \{0, 1, 2, 3\}$ 。

数字 4 就是一个装着 0、1、2、3 这四个元素的口袋。

3.2 赋予运算的灵魂 —— 递归定义与加法

现在有了自然数集 $\mathbb{N}$ ，但我们得到的还只是一排层层嵌套的空盒子。怎样让它们参与运算？又怎样在这个宇宙里严格定义“加法”？

在公理体系里，加法（符号记为 $+$ ）必须被定义为一个明确的映射。

但我们现在手里只有每次走一步的后继操作 S 。怎样用它定义可以跨越任意步数的加法呢？

这就需要用到一种名为**递归**的技巧：把一个复杂动作拆成不断重复的后继步骤。

【定理（自然数递归定理，证明略）】设 A 是集合， $a \in A$ ，且 $H : \mathbb{N} \times A \rightarrow A$ 是映射。那么存在唯一映射 $f : \mathbb{N} \rightarrow A$ ，满足

$$f(0) = a, \quad f(S(n)) = H(n, f(n)).$$

特别地，若已有映射 $g : A \rightarrow A$ ，下一步只依赖当前值，取 $H(n, u) = g(u)$ ，便得到 $f(S(n)) = g(f(n))$ 。这个定理由替换公理、归纳法和映射的集合定义推出。它保证下面各种递归方程不是愿望，而确实唯一指定了映射。

我们用两条精确的递归规则来定义加法（ $+$ ）：

对于任意的自然数 m 和 n ：

1. **基本规则**： $m + 0 = m$ 。（任何数加上 0，等于原地踏步）。

2. **递归规则：** $m + S(n) = S(m + n)$ 。（任何数加上 n 的后继，就等于“先把这两个数相加，然后再对结果整体取一次后继”）。

对每个固定的 m ，递归定理保证上面两条规则唯一确定映射 $n \mapsto m + n$ ；替换公理再把所有 m 的结果合在一起，得到二元映射 $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ 。我们来看看它如何计算 $1 + 1$ 。

（请牢记：1 是 0 的后继，所以 $1 = S(0)$ ；而 2 是 1 的后继，所以 $2 = S(1)$ ）

- 任务：计算 $1 + 1$
- 机器把第二个 1 替换成它原本的样子： $= 1 + S(0)$
- 机器触发递归规则，把后继符号提到最外面： $= S(1 + 0)$
- 机器触发基本规则，任何数加 0 等于它自己，所以 $1 + 0$ 就是 1： $= S(1)$
- 机器查阅自然数阶梯，1 的后继是什么？是 2： $= 2$

看！没有任何现实中拨算盘的经验，纯粹依靠符号的替换，我们让机器合法地推导出了 $1 + 1 = 2$ 。

【引理（零放在左边，短证明）】对每个 $n \in \mathbb{N}$ ，都有 $0 + n = n$ 。

证明：当 $n = 0$ 时，基本规则给出 $0 + 0 = 0$ 。若已经知道 $0 + n = n$ ，那么递归规则与等号替换给出

$$0 + S(n) = S(0 + n) = S(n).$$

由数学归纳法，结论对每个自然数 n 成立。证毕。

同理，递归定理保证下面两条规则唯一确定乘法映射 $\times: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ ：

1. **基本规则：** $m \times 0 = 0$ 。（任何数乘以 0 都是 0）。
2. **递归规则：** $m \times S(n) = (m \times n) + m$ 。（乘以 n 的后继，就等于先乘以 n ，然后再用加法加上一个 m ）。

【记号】从现在起，自然数乘法 $a \times b$ 也可写成 $a \cdot b$ 或省略乘号写成 ab ；三种写法表示同一个刚刚递归定义的映射。此后在整数、有理数以及其他数系中定义了乘法后，也沿用点号和省略乘号这两种写法。

刚才的引理立刻完成一次乘法计算。因为 $1 = S(0)$ ，所以

$$m \times 1 = m \times S(0) = (m \times 0) + m = 0 + m = m.$$

这里每个等号依次来自 1 的定义、乘法递归规则、乘法基本规则和“零放在左边”引理。

【定理（自然数算术律，证明略）】由递归定义和归纳法可以证明：加法、乘法满足结合律与交换律，乘法对加法满足分配律；加法可消去，即 $a + c = b + c \Rightarrow a = b$ ；且 $ab = 0 \Rightarrow a = 0$ 或 $b = 0$ 。

【定义】幂由递归给出： $a^0 = 1$ ， $a^{S(n)} = a^n \cdot a$ 。若 $a : \mathbb{N} \rightarrow \mathbb{N}$ 是映射，则有限和由递归给出：

$$\sum_{k=0}^0 a_k = a_0, \quad \sum_{k=0}^{S(n)} a_k = \left(\sum_{k=0}^n a_k \right) + a_{S(n)};$$

有限积则精确定义为

$$\prod_{k=0}^0 a_k = a_0, \quad \prod_{k=0}^{S(n)} a_k = \left(\prod_{k=0}^n a_k \right) \cdot a_{S(n)}.$$

自然数递归定理保证这些记号唯一确定。

【约定（自然数的十进制记数）】先依次给已经构造出的自然数取短名字：

$$\begin{aligned} 1 &:= S(0), & 2 &:= S(1), & 3 &:= S(2), & 4 &:= S(3), & 5 &:= S(4), \\ 6 &:= S(5), & 7 &:= S(6), & 8 &:= S(7), & 9 &:= S(8), & 10 &:= S(9). \end{aligned}$$

符号 $0, 1, \dots, 9$ 叫作十进制数字。对非空的有限数字串定义值 V ：单个数字满足 $V(d) = d$ ；若 w 已是数字串而 d 是一个数字，则

$$V(wd) = V(w) \cdot 10 + d.$$

写下数字串时，约定它表示 V 的值。例如

$$352 = V(352) = (3 \cdot 10 + 5) \cdot 10 + 2.$$

特别地，数字串“10”的新解释与先前短名字 $10 := S(9)$ 相容，因为

$$V(10) = 1 \cdot 10 + 0 = 10.$$

这是在自然数乘法和加法已经构造完成后引入的缩写，因此没有用“十进制”反过来定义自然数。小数点尚未定义；它要等有理数构造完成后才会出现。

【思考训练】按照 V 的递归规则展开数字串 204；前面添上零得到 007 时，值会改变吗？

解答：

$$V(204) = (2 \cdot 10 + 0) \cdot 10 + 4,$$

而

$$V(007) = (0 \cdot 10 + 0) \cdot 10 + 7 = 7.$$

在任意数字串前添上一个零，都不会改变它的值。事实上，先对单个数字 d 计算

$$V(0d) = 0 \cdot 10 + d = d = V(d).$$

若已经知道 $V(0w) = V(w)$ ，再在末尾接一个数字 d ，便有

$$V(0wd) = V(0w) \cdot 10 + d = V(w) \cdot 10 + d = V(wd).$$

所以对数字串的长度作归纳可得：前置零不会改变值。这里没有借助“位值经验”，只使用了 V 的递归规则和此前的自然数算术。

【定义】对 $m, n \in \mathbb{N}$ ，若存在 $k \in \mathbb{N}$ 使 $m + k = n$ ，定义 $m \leq n$ ；若还满足 $m \neq n$ ，定义 $m < n$ 。

【记号】 $a \geq b$ 表示 $b \leq a$ ，而 $a > b$ 表示 $b < a$ 。这些只是把已经定义的顺序符号反向书写。

【定义】集合 A 上的关系 $\leq$ 若满足自反性 $a \leq a$ 、反对称性（ $a \leq b$ 且 $b \leq a$ 蕴含 $a = b$ ）、传递性，以及任意 a, b 都可比较（ $a \leq b$ 或 $b \leq a$ ），就称为**全序**。

【定理（自然数顺序与良序，证明略）】 $\leq$ 是全序；若 $a \leq b$ ，则对每个 $c \in \mathbb{N}$ 有 $a + c \leq b + c$ 与 $ac \leq bc$ ；若 $a < b$ 且 $c \neq 0$ ，则 $ac < bc$ 。每个非空自然数子集都有最小元素。后面关于“更大”“最小”的说法都使用这个定理。

【定义】自然数 n 若能写成 $n = 2k$ ($k \in \mathbb{N}$)，就称为**偶数**；若能写成 $n = 2k + 1$ ，就称为**奇数**。

【定理（自然数奇偶分解，证明略）】每个自然数恰为偶数或奇数之一。因此 n^2 是偶数当且仅当 n 是偶数。证明只使用归纳法、分配律和消去律。

【思考训练】

请模仿机器的步骤，利用加法的两条规矩，一步步计算 $2 + 2$ 。

（提示：先把第二个 2 写成 $S(1)$ ，然后利用递归规则一直往下拆解）

解答：

- $2 + 2$
- $= 2 + S(1)$ （因为 2 是 1 的后继）
- $= S(2 + 1)$ （运用递归规则）
- $= S(2 + S(0))$ （因为 1 是 0 的后继）
- $= S(S(2 + 0))$ （再次运用递归规则）
- $= S(S(2))$ （运用基本规则 $m + 0 = m$ ，消灭掉 0）
- $= S(3)$ （因为 2 的后继是 3）
- $= 4$ （因为 3 的后继是 4）

3.3 用商集填补漏洞——整数 $\mathbb{Z}$ 的诞生

宇宙似乎很完美。但当你试图寻找一种“逆向”的操作时，灾难降临了。

假设我们有一个起点的自然数 5，和一个终点的自然数 3。我们问：能不能在自然数集 $\mathbb{N}$ 里找到一个未知的自然数 x ，使得 $5 + x = 3$ 成为【真】命题？

答案是：绝对找不到。因为在 $\mathbb{N}$ 里，无论 5 加上什么自然数，结果都会变得更大或者不变，绝不可能退回 3。自然数集 $\mathbb{N}$ 被撕开了一个巨大的漏洞——它只能前进，不能后退。

为了填补漏洞，我们要凭空造出一种可以“后退”的新数字。怎么造？因为我们还不会减法，我们必须动用第二章学过的最强法术——**商集（求同存异）**！

思路极其绝妙：用“一对”自然数来代表一个新数！

我们利用第二章学的笛卡尔积，制造出形如 (a, b) 的自然数对子。

我们把对子 (a, b) 的意义设定为一种组合动作：“先赢得 a 分，再输掉 b 分”。

- $(5, 3)$ 代表赢 5 分，输 3 分。最终效果等同于赢了 2 分。
- $(3, 5)$ 代表赢 3 分，输 5 分。最终效果是输了 2 分！这正是我们梦寐以求的倒退！

但是问题来了： $(5, 3)$ 和 $(4, 2)$ 的最终效果都是“赢 2 分”； $(3, 5)$ 和 $(0, 2)$ 的最终效果都是“输 2 分”。如果有很多对子的效果一样，我们要怎么处理？

我们可以用等价关系把效果相同的对子归为一类。

我们先在 $\mathbb{N} \times \mathbb{N}$ 上定义一个二元关系 $\sim$ ：

对于任意的两个对子 (a, b) 和 (c, d) ，怎么判断它们是不是等价呢？我们目前没有减法，绝对不能用 $a - b = c - d$ 来判断。

可以把原本想表达的相减关系改写成只含加法的等式：

【定义】若 $a + d = b + c$ ，就规定 $(a, b) \sim (c, d)$ 。

【定理（整数等价关系，证明略）】利用自然数加法的交换律、结合律和消去律，可以证明 $\sim$ 满足自反、对称和传递性。

注意：判断等价的依据只有上一节已经定义的自然数加法 $(+)$ 。

我们来验证一下： $(5, 3)$ 和 $(4, 2)$ 等价吗？

我们让第一个对子的赢（5）加上第二个对子的输（2），得到 $5 + 2 = 7$ 。

让第一个对子的输（3）加上第二个对子的赢（4），得到 $3 + 4 = 7$ 。

两个结果相等，所以这两个对子等价。

【定义】整数集是商集

$$\mathbb{Z} = (\mathbb{N} \times \mathbb{N}) / \sim.$$

把 (a, b) 所在的等价类记作 $[(a, b)]$ 。先定义候选映射

$$\iota_{\mathbb{N}\mathbb{Z}} : \mathbb{N} \rightarrow \mathbb{Z}, \quad \iota_{\mathbb{N}\mathbb{Z}}(n) = [(n, 0)].$$

替换公理保证它确实是一条映射；稍后会证明它是保持运算与顺序的单射，届时才有资格叫作“嵌入”。

在等价类上定义

$$\begin{aligned} [(a, b)] + [(c, d)] &= [(a + c, b + d)], \\ -[(a, b)] &= [(b, a)], \\ [(a, b)] \cdot [(c, d)] &= [(ac + bd, ad + bc)]. \end{aligned}$$

最后定义 $x - y = x + (-y)$ 。

为什么一定要检查“良定义”？我们只展示加法这一项。若

$$(a, b) \sim (a', b'), \quad (c, d) \sim (c', d'),$$

那么 $a + b' = b + a'$ 且 $c + d' = d + c'$ 。利用自然数加法的结合律与交换律，

$$\begin{aligned} (a + c) + (b' + d') &= (a + b') + (c + d') \\ &= (b + a') + (d + c') \\ &= (b + d) + (a' + c'). \end{aligned}$$

这正说明 $(a + c, b + d) \sim (a' + c', b' + d')$ 。所以换掉输入代表元，不会改变相加后所在的等价类。相反元与乘法的检查更长，交给下面的良定义定理统一保证。

【定义】记

$$0_{\mathbb{Z}} := [(0, 0)], \quad 1_{\mathbb{Z}} := [(1, 0)].$$

【定理（整数运算良定义，证明略）】如果换用同一等价类里的其他代表，上述三个结果的等价类不变，并且：

- 加法满足结合律、交换律，以 $0_{\mathbb{Z}}$ 为零元，每个元素都有刚才定义的相反元；
- 乘法满足结合律、交换律，以 $1_{\mathbb{Z}}$ 为单位元，并对加法满足分配律；
- $xy = 0_{\mathbb{Z}}$ 蕴含 $x = 0_{\mathbb{Z}}$ 或 $y = 0_{\mathbb{Z}}$ 。

所以，若 $c \neq 0_{\mathbb{Z}}$ 且 $cx = cy$ ，就有 $x = y$ 。第四章会把这些规律分别命名为交换环与“没有零因子”。

【定义】对 $x = [(a, b)]$ 与 $y = [(c, d)]$ ，规定

$$x \leq_{\mathbb{Z}} y \iff a + d \leq b + c,$$

并规定 $x <_{\mathbb{Z}} y$ 当且仅当 $x \leq_{\mathbb{Z}} y$ 且 $x \neq y$ 。

【定理（整数顺序，证明略）】上述定义不依赖代表元； $\leq_{\mathbb{Z}}$ 是全序，并满足：

- 若 $x \leq_{\mathbb{Z}} y$ ，则 $x + z \leq_{\mathbb{Z}} y + z$ ；
- 若 $0_{\mathbb{Z}} \leq_{\mathbb{Z}} x$ 且 $0_{\mathbb{Z}} \leq_{\mathbb{Z}} y$ ，则 $0_{\mathbb{Z}} \leq_{\mathbb{Z}} xy$ ；
- 若 $x <_{\mathbb{Z}} y$ 且 $0_{\mathbb{Z}} <_{\mathbb{Z}} z$ ，则 $xz <_{\mathbb{Z}} yz$ ；
- 若 $x <_{\mathbb{Z}} y$ 且 $z <_{\mathbb{Z}} 0_{\mathbb{Z}}$ ，则 $yz <_{\mathbb{Z}} xz$ 。

于是“正、负、较大、较小”都有了正式含义。

【定理（自然数嵌入整数，证明略）】 $\iota_{\mathbb{NZ}}$ 是单射，并且对 $m, n \in \mathbb{N}$ 有

$$\begin{aligned} \iota_{\mathbb{NZ}}(0) &= 0_{\mathbb{Z}}, & \iota_{\mathbb{NZ}}(1) &= 1_{\mathbb{Z}}, \\ \iota_{\mathbb{NZ}}(m+n) &= \iota_{\mathbb{NZ}}(m) + \iota_{\mathbb{NZ}}(n), & \iota_{\mathbb{NZ}}(mn) &= \iota_{\mathbb{NZ}}(m)\iota_{\mathbb{NZ}}(n), \\ m \leq n &\iff \iota_{\mathbb{NZ}}(m) \leq_{\mathbb{Z}} \iota_{\mathbb{NZ}}(n). \end{aligned}$$

因此不会把两个不同自然数粘成同一个整数。以后把自然数与它的像认作同一个符号，省略 $\iota_{\mathbb{NZ}}$ ，并把 $0_{\mathbb{Z}}, 1_{\mathbb{Z}}, \leq_{\mathbb{Z}}, <_{\mathbb{Z}}$ 简写为 $0, 1, \leq, <$ ；这只是受上述单射保障的记号约定。

【定义】整数 z 若能写成 $z = 2k$ ($k \in \mathbb{Z}$)，就称为**偶数**；若能写成 $z = 2k + 1$ ，就称为**奇数**。

【定理（整数奇偶桥梁，证明略）】每个整数恰为偶数或奇数之一；嵌入后的自然数奇偶性与原定义一致；偶数不可能等于奇数。特别地，对任何 $x \in \mathbb{Z}$ ， $2x$ 是偶数，而 $3 = 2 \cdot 1 + 1$ 是奇数。

用盒子标签观察几个例子：

- 装满 $(5, 3), (4, 2), (2, 0)$ 等等对子的盒子，我们就给盒子贴个标签，叫作 **正二**（记为 **+2**）。
- 装满 $(3, 5), (2, 4), (0, 2)$ 等等对子的盒子，我们就给盒子贴个标签，叫作 **负二**（记为 **-2**）。
- 装满 $(1, 1), (5, 5), (0, 0)$ 等等对子的盒子，就叫作 **零**（记为 **0**）。

在这个新世界 $\mathbb{Z}$ 里，每一个数字都有加法逆元。减法不再是一种尚未解释的动作，而是“加上加法逆元”的简称。

【思考训练】

请问对子 $(1, 4)$ 应该被装进哪个标签的盒子里？对子 $(2, 5)$ 呢？它们在同一个盒子里吗？

解答：

$(1, 4)$ 代表赢 1 分输 4 分，最终是输 3 分，所以装进“负三 (-3) ”的盒子里。

$(2, 5)$ 代表赢 2 分输 5 分，最终也是输 3 分，所以也装进“负三 (-3) ”的盒子里。

我们可以按定义验证： $1 + 5 = 6$ ， $4 + 2 = 6$ ，两者相等，所以这两个对子等价，必须装在同一个盒子里。

3.4 填平碎片的鸿沟 —— 有理数 $\mathbb{Q}$ 的网

整数集 $\mathbb{Z}$ 让我们拥有了减法。但是，当我们用乘法提出类似的问题时，新的漏洞又出现了。

我们问：能不能在整数集 $\mathbb{Z}$ 里找到一个未知的整数 x ，使得 $2 \times x = 3$ 成为【真】命题？

答案再次是：找不到。若 $2x = 3$ ，刚才的整数奇偶桥梁说明左边是偶数、右边是奇数，而整数不可能同时是奇数和偶数，矛盾。

这一次，你已经轻车熟路了。我们要再次施展“商集”魔法，用一对整数，造出能够填满鸿沟的“碎片数字”！

1. **准备对子：**令 $D = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ 。只允许第二项 $q \neq 0$ 。蛋糕故事只是直觉；正式对象是整数对 (p, q) 。

2. **规定等价：**什么时候两个对子代表的“每个人分到的蛋糕”一样多呢？我们再次定义一条二元关系 $\sim$ ：

对于任意的 (p, q) 和 (r, s) ，既然我们现在还不会除法，我们就用合法的乘法交叉相乘来判断！

【定义】若 $ps = qr$ ，就规定 $(p, q) \sim (r, s)$ 。

【定理（有理数等价关系）】这个关系自反、对称且传递。自反性来自 $pq = qp$ ；对称性来自等号的对称性与整数乘法交换律。

其中最值得检查的是传递性。设

$$(p, q) \sim (r, s), \quad (r, s) \sim (u, v).$$

于是 $ps = qr$ 、 $rv = su$ ；又因为 $(r, s) \in D$ ，所以 $s \neq 0$ 。由整数乘法的结合律、交换律与等号替换，

$$s(pv) = psv = qrv = qsu = s(qu).$$

整数的非零乘法消去律给出 $pv = qu$ ，所以 $(p, q) \sim (u, v)$ 。这个证明没有偷偷做除法；它只在最后使用了此前已经证明的消去律。至此三条性质都已验证，故 $\sim$ 是等价关系。证毕。

我们来验证一下：(1, 2) 代表 1 个蛋糕分给 2 个人；(2, 4) 代表 2 个蛋糕分给 4 个人。它们等价吗？

因为 $1 \times 4 = 4$ ，且 $2 \times 2 = 4$ ，两个结果相等，所以这两个对子等价。

3. 打包成盒：【定义】有理数集是商集 $\mathbb{Q} = D/\sim$ ，先把 (p, q) 的等价类记作 $[p, q]$ 。“分数线”要等除法定义完成后再作为简写引入。

4. 定义运算：

$$[p, q] + [r, s] = [ps + rq, qs],$$

$$[p, q] \cdot [r, s] = [pr, qs],$$

$$-[p, q] = [-p, q].$$

【定义】记

$$0_{\mathbb{Q}} := [0, 1], \quad 1_{\mathbb{Q}} := [1, 1].$$

若 $p \neq 0$ ，定义 $[p, q]^{-1} = [q, p]$ ；除法定义为 $x/y = x \cdot y^{-1}$ ，其中 $y \neq 0_{\mathbb{Q}}$ 。再定义候选映射

$$\iota_{\mathbb{Z}\mathbb{Q}} : \mathbb{Z} \rightarrow \mathbb{Q}, \quad \iota_{\mathbb{Z}\mathbb{Q}}(z) = [z, 1].$$

【定理（有理数运算良定义，证明略）】这些结果不依赖代表元的选择，并且：

- 加法满足结合律、交换律，以 $0_{\mathbb{Q}}$ 为零元，每个元素都有相反元；
- 乘法满足结合律、交换律，以 $1_{\mathbb{Q}}$ 为单位元，并对加法满足分配律；
- 每个非零有理数都有唯一乘法逆元，而且 $0_{\mathbb{Q}} \neq 1_{\mathbb{Q}}$ 。

若 $[p, q] \neq 0_{\mathbb{Q}}$, 则 $p \neq 0$, 所以前面写出的 $[q, p]$ 确实合法且不依赖代表元。第四章会把这样的结构命名为域。

$0_{\mathbb{Q}}$ 不可能有逆元。否则, 若 $0_{\mathbb{Q}}u = 1_{\mathbb{Q}}$, 而分配律和加法消去律又给出 $0_{\mathbb{Q}}u = 0_{\mathbb{Q}}$, 便会推出 $0_{\mathbb{Q}} = 1_{\mathbb{Q}}$ 。

【定理 (整数嵌入有理数, 证明略)】 $\iota_{\mathbb{Z}\mathbb{Q}}$ 是单射, 保持 $0, 1, +, \cdot$ 与相反元。对 $p, q \in \mathbb{Z}$ 且 $q \neq 0$,

$$\frac{\iota_{\mathbb{Z}\mathbb{Q}}(p)}{\iota_{\mathbb{Z}\mathbb{Q}}(q)} = [p, q].$$

以后把整数与它的像认作同一个符号, 省略 $\iota_{\mathbb{Z}\mathbb{Q}}$, 并把 $0_{\mathbb{Q}}, 1_{\mathbb{Q}}$ 简写为 $0, 1$; 同时把 $[p, q]$ 写成 p/q 。自然数进入 $\mathbb{Q}$ 的映射是两个嵌入的复合

$$\iota_{\mathbb{N}\mathbb{Q}} = \iota_{\mathbb{Z}\mathbb{Q}} \circ \iota_{\mathbb{N}\mathbb{Z}},$$

所以 $\mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q}$ 的每一步都不会粘合原有的不同数字, 并且都保持已经定义的算术运算。

【约定 (有限小数)】设 w 是小数点前的非空数字串, 小数点后依次有 $k \geq 1$ 个数字 $d_1, \dots, d_k$ 。令

$$N = V(wd_1 \cdots d_k) \in \mathbb{N},$$

并定义

$$w.d_1 \cdots d_k := \frac{N}{10^k} \in \mathbb{Q}.$$

自然数算术律保证 $10^k \neq 0$, 所以这个分母合法。例如 $1.4 = 14/10$, 而 $0.01 = 1/100$; 前置负号表示取加法逆元。可见有限小数只是已构造有理数的缩写, 不是新的数字。

【思考训练】只用有理数的等价关系证明 $0.50 = 0.5$ 。

解答: 按照有限小数定义, $0.50 = 50/100$, $0.5 = 5/10$ 。交叉相乘得到

$$50 \cdot 10 = 500 = 100 \cdot 5,$$

所以 $(50, 100) \sim (5, 10)$ ，两种小数写法表示同一个有理数。

【前瞻】下一章会先定义有理数的顺序，再证明：每当一个有理数排在另一个之前时，中点 $(a + b)/2$ 还排在二者之间。那时才把这项性质命名为**稠密性**。本章尚未定义顺序符号，因此这句前瞻不参与本章的任何证明。

数字的网，似乎已经无限细密，天衣无缝了。

【思考训练】

有两堆对子，第一堆是 $(2, 3)$ ，代表 2 个蛋糕分给 3 个人；第二堆是 $(4, 6)$ ，代表 4 个蛋糕分给 6 个人。请用上面刚刚学的“交叉相乘验证法”，证明它们必须被装进同一个等价类的盒子里。

解答：

把第一堆的蛋糕（2）乘以第二堆的人数（6）： $2 \times 6 = 12$ 。

把第一堆的人数（3）乘以第二堆的蛋糕（4）： $3 \times 4 = 12$ 。

因为 $12 = 12$ ，所以 $(2, 3) \sim (4, 6)$ 是【真】命题。这两个对子等价，属于同一个名为“三分之二”的盒子。

3.5 本章结语

这一章，你完成了创世法则中最惊心动魄的几步跳跃。

你没有借助任何现实世界里的苹果、算盘或者手指，仅仅利用第一章的对错逻辑和第二章的空口袋：

- 用空集 $\emptyset$ 和后继映射，堆叠出了自然数阶梯 $\mathbb{N}$ 。
- 用机器般的递归法则，赋予了它们加法和乘法的生命。
- 当自然数的加法无法逆转时，你用笛卡尔积和商集魔法，将自然数对子强行打包，无中生有地造出了拥有减法的整数王国 $\mathbb{Z}$ 。
- 当整数的乘法留下鸿沟时，你故技重施，用整数对子造出了拥有除法的有理数大网 $\mathbb{Q}$ 。

这就是数学家对秩序的绝对狂热：没有任何东西是天上掉下来的。每一个漏洞，都必须用更高级的逻辑结构去合法地缝补。

有了加法、减法、乘法和“除以非零数”都畅通无阻的有理数 $\mathbb{Q}$ ，我们似乎已经掌控了强大的计算工具。

但在下一章，我们要暂时停下制造数字的脚步。我们要把数字本身全部抛弃，去审视加减乘除背后隐藏的“纯粹骨架”。我们将看到，真正的魔法不仅存在于数字之中，更存在于那些看不见的旋转与对称之中。

Xulei Sun

4 抽象的骨架 —— 抽离运算的灵魂

在上一章，我们像搭积木一样，用一个个口袋（集合）一步步拼出了自然数、整数和有理数，并且赋予了它们加、减、乘、除的运转规则。

现在，你可能会觉得：数学，就是研究数字的学问。

大错特错。

在纯粹的数学宇宙里，元素是不是数字并非关键；自然数、整数和有理数只是不同的研究对象。真正把它们组织起来的，是作用在这些对象上的“规则”。

如果我们在这一刻，施展魔法把所有的“数字”全部消灭，宇宙机器会停止运转吗？

不会的。因为这一章，我们要把加减乘除的共同结构从数字中抽离出来。你会惊奇地发现，只要满足特定的“骨架规则”，即使元素不是数字，也能进行严格定义的运算！

欢迎来到代数学的最高殿堂。

4.1 抽离加法的灵魂 —— “群”的诞生

让我们盯着上一章在整数中合法的“加法”看。抛开那些枯燥的数字，加法究竟是个什么东西？

利用第二章的知识，所谓的“加法”，其实是一种映射规则：它把两个输入元素组合起来，得到一个新的元素。

现在，取一个集合 G 。

【定义】 G 上的二元操作是映射 $\star : G \times G \rightarrow G$ 。它已经把“封闭”包含在映射的陪域里。

只要这个操作 $\star$ 能通过宇宙法庭的四项绝对测试，我们就赋予口袋 G 一个无比尊贵的头衔——群。

哪四项测试呢？这就是加法的“灵魂骨架”：

1. **封闭法则**：已经由 $\star : G \times G \rightarrow G$ 保证。
2. **结合法则**：对任意 $a, b, c \in G$, $(a \star b) \star c = a \star (b \star c)$ 。
3. **单位元**：存在 $e \in G$, 使任意 $a \in G$ 都满足 $e \star a = a \star e = a$ 。
4. **逆元**：对任意 $a \in G$, 存在 $a^{-1} \in G$, 使 $a \star a^{-1} = a^{-1} \star a = e$ 。

【定义】满足这四条的 $(G, \star)$ 是一个群。若对任意 $a, b \in G$ 还满足 $a \star b = b \star a$, 就称为交换群。

【定理（唯一性）】群的单位元唯一，每个元素的逆元也唯一。因此，“这个单位元”和“ a 的逆元”都不需要另作选择。

这个证明很短，值得完整看一次。若 e 与 e' 都是单位元，那么

$$e = e \star e' = e'.$$

第一个等号使用“ e' 是右单位元”，第二个等号使用“ e 是左单位元”。单位元因而只能有一个。再设 b, c 都是 a 的逆元，则利用结合律，

$$b = b \star e = b \star (a \star c) = (b \star a) \star c = e \star c = c.$$

所以逆元也只能有一个。这个证明展示了公理的作用：结论不是靠观察许多例子猜出，而是由四条群法则必然推出的。

【例子】整数加法群 $(\mathbb{Z}, +)$ 是交换群：单位元是 0，整数 a 的逆元是 $-a$ 。非零有理数在乘法下也组成交换群 $(\mathbb{Q} \setminus \{0\}, \cdot)$ ：单位元是 1， q 的逆元是 q^{-1} 。同一个“群”字可以描述完全不同的元素与操作；真正共同的是四条法则。

【思考训练】

请你当一次法官。口袋里装的是我们上一章造出来的自然数集（包含 0, 1, 2, 3, ...），操作规则是加法。

请问，自然数在加法操作下，能构成一个“群”吗？它倒在了哪一项测试上？

解答：

它不能构成群，因为它不满足第四项“逆元”条件。

例如，对自然数 3，我们需要在自然数集中找到一个数，使它与 3 相加等于单位元 0。但是自然数集中没有 -3 （它属于整数集），所以 3 在自然数加法下没有逆元，自然数集也就不能构成群。

4.2 万物皆可运算 —— 映射组成的置换群

为了证明“群”的骨架根本不需要数字，我们来玩一个纯粹的“换位游戏”。

【定义】 设 A 是集合。从 A 到 A 的双射称为 A 的一个**置换**。所有置换组成的集合记作

$$\text{Sym}(A) = \{\sigma : A \rightarrow A : \sigma \text{ 是双射}\}.$$

若 $\sigma, \tau \in \text{Sym}(A)$ ，定义

$$\sigma * \tau = \tau \circ \sigma.$$

因此 $\sigma * \tau$ 表示“先做 σ ，再做 τ ”；这个次序约定会在下文一直保持。

【记号】 若有限集合 C 恰有 m 个元素，就写作 $|C| = m$ 。第三章关于有限集合的定义保证这个自然数 m 唯一。

【定理（有限乘法计数原理，证明略）】 设 S 是一个集合， C 是第一步选择组成的有限集合。若映射 $p : S \rightarrow C$ 把每个完整方案送到它的第一步选择，且 $|C| = m$ ，并且对每个 $c \in C$ ，集合 $\{s \in S : p(s) = c\}$ 都恰有 n 个元素，那么 S 是有限集，并且 $|S| = m \times n$ 。因此这条定理不仅计算数目，也证明完整方案确实只有有限多个；它可由有限集合、双射与自然数乘法推出，有限多个步骤可反复使用。

现在令 $A = \{1, 2, 3\}$ 表示桌上的三个位置，并在三个位置上分别放苹果卡、香蕉卡、橘子卡。一次换位规则给每个原位置指定唯一的新位置，并且不漏掉任何位置；因此它恰好是集合 A 的一个置换。

比如，我们规定一个置换规则，取名叫作“规则一”：把第一张和第二张卡片互换位置，第三张不动。

再规定一个规则，取名叫作“规则二”：把所有的卡片统一向右挪动一个位置（最后一张挪到最前面）。

一个置换由三张卡片各自的去处唯一确定。第一张卡片的去处有 3 种；确定后，第二张为了不与它相撞有 2 种；最后一张只剩 1 种。反复使用有限乘法计数原理，

$$|\text{Sym}(A)| = 3 \times 2 \times 1 = 6.$$

我们现在把这 6 个置换规则本身当作新集合的元素。

这里的操作正是上面已经定义的 $\star$ 。比如“规则一 $\star$ 规则二”，意思就是先执行规则一，再执行规则二。

【定理（对称群，证明略）】 $(\text{Sym}(A), \star)$ 是群：

1. 两个双射的复合仍是双射，所以运算封闭；
2. 映射复合满足结合律，所以 $\star$ 也满足结合律；
3. 恒等映射 id_A 是单位元；
4. 每个双射 σ 都有逆映射 σ^{-1} ，它就是群中的逆元。

置换群还告诉我们：群运算不一定能交换次序。令 s 交换 1 与 2、固定 3；令 r 把 $1 \mapsto 2$ 、 $2 \mapsto 3$ 、 $3 \mapsto 1$ 。按照“先做左边、再做右边”的约定，

$$(s \star r)(1) = r(s(1)) = r(2) = 3,$$

而

$$(r \star s)(1) = s(r(1)) = s(2) = 1.$$

两个映射在输入 1 时给出不同结果，所以 $s \star r \neq r \star s$ 。因此 $\text{Sym}(A)$ 是群，却不是交换群。这也是为什么群定义不能擅自加入交换律。

计数六种规则时，我们使用了自然数乘法；但这个群自身的元素不是数字，群运算也不是数的加减乘除，而只是置换规则的依次复合。它依然通过了群的全部测试！

这个群称为 A 的**对称群**。

【定义】若 $H \subseteq \text{Sym}(A)$ ，并且把 $\star$ 限制在 H 上以后， $(H, \star)$ 仍是群，就称 H 是 $\text{Sym}(A)$ 的一个**子群**，也称为 A 上的一个**置换群**。这个例子揭示了一个重要事实：运算不一定是在计算数字，也可以是在组合规则。

4.3 双重魔法的叠加 —— 环与域

有了群的骨架，我们可以处理一种操作（比如加法）。但有理数拥有加法和乘法两种操作。我们必须把乘法的骨架也抽离出来。

【定义】若集合 R 带有两种二元操作 $\oplus, \otimes$ ，并对任意 $a, b, c \in R$ 满足下列法则，就称 $(R, \oplus, \otimes)$ 为环：

1. 加法群： $(R, \oplus)$ 是交换群，加法单位元记为 0 。
2. 乘法结合与单位元： $(a \otimes b) \otimes c = a \otimes (b \otimes c)$ ，并存在 $1 \in R$ ，使每个 $a \in R$ 都满足 $1 \otimes a = a \otimes 1 = a$ 。
3. 左右分配律：

$$a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c),$$

$$(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c).$$

本书采用“环含乘法单位元”的约定。若对任意 $a, b \in R$ 还有 $a \otimes b = b \otimes a$ ，就称为交换环。

整数就是这样一个环。但在上一章我们知道，并非每个非零整数都有整数范围内的乘法逆元。例如，整数集中没有 2 的乘法逆元，因为 $1/2$ 不是整数。

为了让每个非零元素都能进行乘法逆运算，我们引出一种更强的代数结构——域。

【定义】若交换环满足 $0 \neq 1$ ，且 $R \setminus \{0\}$ 在乘法下构成交换群，就称为域。

为什么 0 不可能有乘法逆元？先由环公理证明一个短结论。这里仍使用环定义中的符号 $\oplus, \otimes$ 。对任意 $a \in R$ ，分配律给出

$$0 \otimes a = (0 \oplus 0) \otimes a = (0 \otimes a) \oplus (0 \otimes a).$$

令 c 是元素 $0 \otimes a$ 在加法群中的逆元。在等式两边同时加上 c ，便得到 $0 \otimes a = 0$ 。若域中的 0 有乘法逆元 u ，就会有 $1 = 0 \otimes u = 0$ ，与域所要求的 $0 \neq 1$ 矛盾。因此，“不能除以 0”不是额外禁令，而是由环与域的法则共同推出的结论。

在域中总能进行加法、减法和乘法；只要除数不为 0，也总能进行除法，而且结果仍留在域中。

上一章我们千辛万苦利用商集魔法造出来的有理数网，正是我们目前拥有的第一个绝对合法的“域”。

【思考训练】

请问，在有理数域的乘法下，保持其他元素不变的单位元是哪个数？数字 5 的乘法逆元是谁？

解答：

乘法单位元是 1，因为任何有理数乘以 1 都等于它自己。

数字 5 的乘法逆元是五分之一 ($1/5$)，因为 $5 \times (1/5) = 1$ ，乘积恰好是乘法单位元。

4.4 未来的魔法容器 —— 多项式环

在进入下一章之前，我们用刚建立的代数骨架锻造一件重要工具：多项式环。

平时，你可能听过“未知数 X ”这种说法。但是，在严密的定义中，只说一个对象“未知”还不够精确；我们必须用集合和映射明确说明它是什么。

设 R 是任意交换环。以下把环运算 $\oplus$ 简写为 $+$ ，把 $\otimes$ 简写为 $\cdot$ ，并允许像普通乘法那样省略乘号；加法零元与乘法单位元仍记作 $0, 1$ 。利用第二章的映射法则，一个取值于 R 的序列就是映射 $a: \mathbb{N} \rightarrow R$ ，把 $a(n)$ 简写为 a_n 。

【定义】 若存在 $N \in \mathbb{N}$ ，使每个 $n > N$ 都有 $a_n = 0$ ，就称 a 具有有限支撑。

我们在这些盒子里放入 R 的元素。取 $R = \mathbb{Q}$ 时， $(2, 3, 5, 0, 0, \dots)$ 表示第零号盒子装 2，第一号盒子装 3，第二号盒子装 5，后面的盒子全装 0。

为了防止它失控，我们下达一道铁律：这排无限的盒子里，只有有限个盒子里可以装非零的数字，后面的无穷多个盒子必须全都是 0。

卷积公式需要把 R 中有限多个元素相加，所以先把这个记号正式造出来。对任意序列 $c: \mathbb{N} \rightarrow R$ ，第三章的递归定理给出唯一序列 $S: \mathbb{N} \rightarrow R$ ，满足

$$S_0 = c_0, \quad S_{n+1} = S_n + c_{n+1}.$$

【定义】 记

$$\sum_{k=0}^n c_k = S_n.$$

因此这里的有限和使用的是环 R 的加法，而不是把系数误当成自然数。

卷积公式还会用到自然数的差。若 $k, n \in \mathbb{N}$ 且 $k \leq n$ ，第三章对 $\leq$ 的定义保证存在 $r \in \mathbb{N}$ 使 $k + r = n$ ，加法消去律又保证这个 r 唯一。

【定义（自然数差）】把这个唯一的 r 记作 $n - k$ 。这里仅在 $k \leq n$ 时使用自然数差。

接下来，我们在这排长长的盒子上定义两套操作：

- 加法： $(a + b)_n = a_n + b_n$ 。
- 乘法（卷积）：

$$(ab)_n = \sum_{k=0}^n a_k b_{n-k}.$$

右边正是刚刚递归定义的、使用 R 中加法的有限和。所有下标之和为 n 的系数乘积都要相加，不能只取其中一项。

【定理（多项式环，证明略）】所有从 $\mathbb{N}$ 到 R 的有限支撑序列在上述运算下封闭，并构成交换环，记为 $R[X]$ 。它的加法零元是全零序列，乘法单位元是 $(1, 0, 0, \dots)$ 。证明只需环公理、有限和与自然数归纳法。

此时，我们从这个新环里揪出一个极其特殊的盒子序列，它只有第一号盒子里装着 1，其他所有的盒子里全都是 0： $(0, 1, 0, 0, \dots)$ 。

我们给这个特殊的序列赋予一个神圣的符号：大写字母 X 。

利用我们刚定义的盒子编号相加的乘法规则，你会发现一个惊天奇迹：

当序列 X 乘以序列 X （即第一号盒子乘第一号盒子），结果会跑到 $1 + 1 = 2$ 的位置。于是 $X \cdot X$ 的结果变成了 $(0, 0, 1, 0, \dots)$ ，也就是第二号盒子装了 1。

如果再乘一次 X ，结果就会跑到第三号盒子： $(0, 0, 0, 1, 0, \dots)$ 。

因此，与 X 相乘会把每个系数向后移动一格；继续乘以 X ，就会继续向后移动。

把环元素 $c \in R$ 识别为常数序列 $(c, 0, 0, \dots)$ ，并递归定义 $X^0 = 1$ 、 $X^{n+1} = X^n X$ 。若 $a_n = 0$ 对所有 $n > N$ 成立，则相应序列可写成

$$a_0 + a_1 X + \dots + a_N X^N,$$

而每个位置上的系数 a_n 由序列唯一确定；在末尾添上系数为 0 的项不会产生新的序列。

特别地，在 $\mathbb{Q}[X]$ 中，序列 $(2, 3, 5, 0, \dots)$ 可以写成 $2 + 3X + 5X^2$ 。

【例子（逐格计算多项式乘法）】在 $\mathbb{Q}[X]$ 中，令

$$a = 1 + 2X = (1, 2, 0, \dots), \quad b = 3 + X = (3, 1, 0, \dots).$$

卷积公式逐格给出

$$(ab)_0 = 1 \cdot 3 = 3,$$

$$(ab)_1 = 1 \cdot 1 + 2 \cdot 3 = 7,$$

$$(ab)_2 = 1 \cdot 0 + 2 \cdot 1 + 0 \cdot 3 = 2.$$

更高位置的系数全为 0，所以

$$(1 + 2X)(3 + X) = 3 + 7X + 2X^2.$$

熟悉的“展开括号”并不是另一套规则，而只是把盒子序列的卷积公式写成简洁符号后的结果。

我们把这个由无限盒子构成的全新代数结构命名为**多项式环**。在这个结构里， X 不是“某个尚不知道数值的数”，而是由序列 $(0, 1, 0, \dots)$ 严格定义的元素。因为 R 可以是任意交换环，所以第五章构造出 $\mathbb{R}$ 与 $\mathbb{C}$ 后，同一份定义会立即给出 $\mathbb{R}[X]$ 与 $\mathbb{C}[X]$ ，不需要另造一套规则。

4.5 建立绝对的阶级——序结构

现在，有理数已经构成一个域，可以在相应条件下进行加、减、乘、除。

但是，我们的宇宙还缺失了一种极其重要的能力：比较大小。

如果没有大小之分，所有的有理数在网里就只是一团混沌。我们必须给它们制定“排队法则”。

请记住，我们不能依靠人类的生活经验去说“3 比 2 大”。所有的比较，都必须转化为合法的“集合与逻辑”。

利用第三章的整数顺序，我们正式定义有理数的正子集：

$$P = \{x \in \mathbb{Q} : \text{存在 } p, q \in \mathbb{Z}, \text{ 使 } q \neq 0, x = [p, q] \text{ 且 } ((p > 0 \wedge q > 0) \vee (p < 0 \wedge q < 0))\}.$$

也就是说，分子与分母同号时，这个有理数为正。

【定理（正性良定义，证明略）】若 $[p, q] = [r, s]$ ，其中 $q, s \neq 0$ ，则 p, q 同号当且仅当 r, s 同号。因此上面对 P 的定义不依赖代表元的选择。并且对任意 $a, b \in \mathbb{Q}$ ， P 满足下面两条法则：

1. **绝对的分裂（三歧性）**： $a \in P$ 、 $a = 0$ 、 $-a \in P$ 三个命题中恰有一个为真；
2. **绝对的封闭**：若 $a, b \in P$ ，则 $a + b \in P$ 且 $ab \in P$ 。

因此，有理数中确实存在这样一个正子集；下面的顺序不是未经证明的任意选择。我们终于可以利用 P ，严格定义表示大小关系的符号：大于（记为 $>$ ）。

究竟什么是 $a > b$ ？

【定义】 $a > b$ 当且仅当 $a - b \in P$ ；定义 $a < b$ 为 $b > a$ ，并用 $a \leq b$ 表示 $a < b$ 或 $a = b$ ，用 $a \geq b$ 表示 $b \leq a$ 。

举个例子：为什么 5 大于 3？因为 5 减去 3 等于 2，而 2 这个元素，恰好躺在我们定义好的正子集 P 这个口袋里！

所有的大小比较，最终都被转化为“某个元素是否属于集合 P ”这一逻辑判断。

【推论（顺序运算律，证明略）】利用 P 的三歧性和封闭性可推出：

- 如果 $a > b$ 且 $b > c$ ，那么必定有 $a > c$ 。
- 如果 $a > b$ ，那么对任意有理数 c ，两边同时加上 c 后，仍有 $a + c > b + c$ 。

【定义】设 F 是域， $\leq$ 是 F 上的全序。若对任意 $a, b, c \in F$ 都满足

$$a \leq b \implies a + c \leq b + c,$$

并且对任意 $a, b \in F$ 都满足

$$0 \leq a \wedge 0 \leq b \implies 0 \leq ab,$$

就称 $(F, \leq)$ 为**有序域**。

【定义】全序集称为**稠密的**，当且仅当对任意 $a < b$ ，都存在元素 c 使 $a < c < b$ 。

【定理（有理数有序域，证明略）】上述关系是全序，并与加法、乘法相容；若 $a < b$ ，则 $a < (a + b)/2 < b$ 。因此 $\mathbb{Q}$ 是稠密有序域。

最后一个不等式有一个很短的证明。令 $c = (a + b)/2$ 。因为 $b - a > 0$ ，并且正数除以正数 2 仍为正数，所以

$$c - a = \frac{b - a}{2} > 0, \quad b - c = \frac{b - a}{2} > 0.$$

按照“大于”的定义，这正是 $a < c$ 与 $c < b$ 。因此，任意两个不同的有理数之间都能明确构造出另一个有理数，而不是只凭“看起来很密”作判断。

【定理（有理数的阿基米德性质，证明略）】对每个 $q \in \mathbb{Q}$ ，都存在 $n \in \mathbb{N}$ 使 $n > q$ 。等价地，对每个正有理数 ϵ ，都存在 $n \in \mathbb{N}$ 且 $n \geq 1$ ，使

$$0 < \frac{1}{n} < \epsilon.$$

把 q 写成整数之比后，利用自然数的后继与整数顺序即可证明；这里不需要实数。第五章验证戴德金分割的乘法逆元时会使用这条定理。

【推论（有序域的符号法则，证明略）】在任意有序域中， $x^2 \geq 0$ ，并且 $x^2 = 0$ 当且仅当 $x = 0$ ；若 $x \neq 0$ ，则 $x^2 > 0$ 。此外， $a^2 + b^2 = 0$ 当且仅当 $a = b = 0$ 。这些结论都可由三歧性、乘法封闭性和域中无零因子推出。

【定义】在任意有序域中， $(a, b) = \{x : a < x < b\}$ ， $[a, b] = \{x : a \leq x \leq b\}$ ；半开区间用同样方式定义。以后“区间”都指由这些顺序条件指定的集合。

至此，原本混沌一团的有理数网，被强行拉扯成了一条有着绝对先后顺序的单行道。所有的元素规规矩矩地排成了一列长长的队伍。

我们给带有这种完美排队规则的域，颁发了最终的称号：**有序域**。

【思考训练】

已知 $8 > 5$ ，也就是说 $(8 - 5)$ 这个结果属于正子集 P 。

请利用这个定义证明：在不等式两边同时加上 2 后，依然有 $10 > 7$ 。

解答：

要证明 $10 > 7$ ，只需判断 $10 - 7$ 是否属于 P 。由

$$10 - 7 = 3 = 8 - 5$$

以及已知的 $8 > 5$ ，可知 $8 - 5 \in P$ 。根据等号替换， $10 - 7$ 与 $8 - 5$ 是同一个有理数，所以 $10 - 7 \in P$ 。按照“大于”的定义，这就证明了 $10 > 7$ 。

4.6 本章结语

这一章，你体会到了剥离现实后，数学最深层的恐怖与美丽。

你发现，代数运算不一定作用在数字上。“群”是单一运算的骨架，“环”协调两种运算，而“域”还保证每个非零元素都有乘法逆元。

你用有限支撑的盒子序列构造出了多项式环。

最后，你用口袋分类的逻辑，给有理数刻下了绝对的“大于”秩序，将其化为完美的“有序域”。

现在，我们的有理数大军不仅可以进行四则运算，而且整齐地排好了长队，仿佛组成了一条无限延伸、看似密不透风的直线。

但这条直线真的坚不可摧吗？

深呼吸。在下一章，我们将在这条看似完美的直线上，残忍地劈下一刀。这一刀，将划破有理数的伪装，让你直视一个深不见底的恐怖深渊——连续统。在那里，你将见证宇宙中最宏大的一场造物奇迹。

5 完美的直线 —— 直面连续统的深渊

在上一章的结尾，我们的宇宙中诞生了“有序域”。所有的有理数都根据“大于”和“小于”的排队法则，整齐划一地排成了一条前后无限延伸的队伍。

当你在这条队伍中随便挑出两个有理数（比如排在前面的甲和排在后面的乙），你总能用上一章合法的加法和除法，把它们加起来除以 2，在它们正中间精确地找到一个新的有理数。

由于你可以无限次地执行这个“找中间”的操作，有理数仿佛已经像沙子一样密密麻麻，彻底填满了整个宇宙，连一丝缝隙都不剩。

但这是最大的幻觉。

这一章，我们要拿起纯逻辑的利刃，狠狠劈向这条看似密不透风的有理数队伍。我们将撕开它的伪装，直视里面深不见底的恐怖深渊。

随后，为了拯救宇宙免于破碎，我们必须发动整个数学史上最宏大、最极致的“造物魔法”，去铸造一条真正完美、没有任何破绽的绝对直线。

欢迎来到实数与连续统的世界。

5.1 致命的一刀 —— 发现深渊的裂缝

我们要如何证明密不透风的有理数队伍里，其实至少存在一个空洞？

答案是：用一个猜谜游戏来寻找幽灵。

回想我们在第三章寻找负数和分数时使用的手段。当时我们问：

- 在自然数里，能不能找到一个未知的数 x ，使得 $x + 5 = 3$ 为【真】？找不到，于是诞生了可以后退的整数。
- 在整数里，能不能找到一个未知的数 x ，使得 $2 \times x = 3$ 为【真】？找不到，于是诞生了可以填补碎片的有理数。

现在，我们在有理数里，提出一个极其简单的挑战：
能不能在有理数里找到一个正数 x ，使得 $x \times x = 2$ 为【真】？

让我们用第一章的反证法审判这个问题。

【定义】对整数 n ，若存在 $k \in \mathbb{Z}$ 使 $n = 2k$ ，就称 n 为偶数；若存在 $k \in \mathbb{Z}$ 使 $n = 2k + 1$ ，就称 n 为奇数。

【定义】对整数 d, n ，若存在 $k \in \mathbb{Z}$ 使 $n = dk$ ，就说 d 整除 n 。同时整除整数 p, q 的正整数叫共同正因子；若这些共同正因子中有最大的一个，就称它为 p, q 的最大公因子。

【引理（证明略）】第三章的良序与整数运算可以构造最大公因子，并结合奇偶分解推出：

1. 每个非零有理数都可写成 p/q ，其中 $q > 0$ ，且 p, q 没有大于 1 的共同正因子；这样的表示叫最简表示。
2. 若整数 n^2 是偶数，则 n 是偶数。它的逆否证明是：若 $n = 2k + 1$ ，则 $n^2 = 2(2k^2 + 2k) + 1$ 是奇数。
3. 假设：对方赢了，假设真的存在这样一个合法的有理数 x ，满足 $x \times x = 2$ 。
4. 推导：既然 x 是有理数，根据第三章的定义，它必定是一个分数。为了方便，我们把它写成上下两个整数相除的形式： $x = \frac{p}{q}$ 。
取上面引理保证存在的最简表示，所以 p 与 q 不可能同时为偶数。
5. 代入：既然 $x \times x = 2$ ，那么 $\frac{p \times p}{q \times q} = 2$ 。由于 $q \neq 0$ ，两边同乘 $q \times q$ ，得到 $p \times p = 2 \times (q \times q)$ 。
6. 奇偶的审判：

- 观察上面的式子， $p \times p$ 等于 2 乘以某个数。这意味着 $p \times p$ 必定是一个“偶数”（因为它包含 2 的倍数）。
- 由“ n^2 为偶数则 n 为偶数”的引理， p 必定是偶数。
- 既然 p 是偶数，我们把 p 写成 $2 \times k$ （ k 是某个未知的整数）。
- 把 p 替换掉，代回刚才的式子： $(2 \times k) \times (2 \times k) = 2 \times (q \times q)$ 。
- 稍微整理一下： $4 \times (k \times k) = 2 \times (q \times q)$ 。
- 把两边都写成含因子 2 的乘积，并用整数的乘法消去律消去非零因子 2，得到 $2(k \times k) = q \times q$ 。

- 看！现在轮到 $q \times q$ 等于 2 乘以某个数了！所以 $q \times q$ 必定也是偶数，从而得出结论： q 本身必定也是一个偶数！

7. **出现矛盾：**第 2 步选取的是最简表示，所以 p 和 q 不能同时为偶数；第 4 步却推出了它们同时为偶数。

8. **结论：**这与最简表示矛盾。根据第一章的反证法，最初的存在假设为假。

法庭最终宣判：不存在正有理数使 $x^2 = 2$ 。若有负有理数满足这个等式，它的相反数也会满足，而且是正有理数；0 也不是解。因此，不存在任何有理数使 $x^2 = 2$ 。

这意味着什么？

想象你把有理数队伍按照大小排好：

若只观察非负有理数，可以把平方小于 2 的数放在左侧，把平方大于 2 的数放在右侧。若要切分全部有理数，还必须把所有负有理数放入左侧；下一节会写出精确定义。

当你在这两支队伍交界的地方寻找那个“刚刚好等于 2”的界碑时……

那里什么都没有！那里是一个无法名状的黑洞，是一个将有理数队伍彻底切断的深渊裂缝。

仅仅这一道裂缝，就已经足够证明有理数的稠密性不等于“没有缺口”。至于还能构造多少道不同裂缝，需要另外的定理；在那些定理出现之前，我们不凭直觉断言它们的数量。

【思考训练】

请问，在有理数队伍里，能不能找到一个正数 y ，使得 $y \times y = 9$ ？这个方程是否也暴露出一道有理数裂缝？

解答：

能找到，因为 $3 \times 3 = 9$ ，而 $3 = 3/1$ 是有理数。因此这个方程没有暴露新的有理数裂缝； $x^2 = 2$ 所确定的分割才没有有理数位于刀口。

5.2 造物主的利刃 —— 戴德金分割

宇宙绝不允许破碎。为了填补这些深不见底的裂缝，我们必须造出一种全新的数字。

但这次和第三章造分数不一样，我们甚至无法用一对整数来定位它，因为它根本就不在比例的网格上！

怎么无中生有地造出深渊？

德国数学家戴德金拔出了一把名为“纯逻辑”的极致利刃。他提出了数学史上最狂野的造物思想：

既然裂缝本身什么都没有，那我们就把“裂缝左边所有的有理数”打包起来装进口袋，把这个巨大的口袋本身，当成那个新数字！

我们要向有理数队伍挥下一刀（这被称为**戴德金分割**）。

只记录左口袋就足够，因为右口袋必定是 $\mathbb{Q} \setminus L$ 。

【定义】 集合 $L \subseteq \mathbb{Q}$ 称为一个**戴德金分割**，若满足：

1. $L \neq \emptyset$ 且 $L \neq \mathbb{Q}$;
2. **向下封闭**：若 $x \in L$ 且 $y < x$ ，则 $y \in L$;
3. **没有最大元**：对每个 $x \in L$ ，存在 $y \in L$ 使 $x < y$ 。

令 $U = \mathbb{Q} \setminus L$ 。由定义可证明：若 $l \in L$ 且 $u \in U$ ，则 $l < u$ ；而且若 $u \in U$ 且 $u < v$ ，则 $v \in U$ 。所以这一刀不会在队伍里乱切。

【定义】 对每个 $r \in \mathbb{Q}$ ，记

$$r^* = \{q \in \mathbb{Q} : q < r\}.$$

【定理（有理分割判别，证明略）】 r^* 是戴德金分割，并且 $\mathbb{Q} \setminus r^*$ 有最小元 r 。反过来，若 L 是戴德金分割且 $\mathbb{Q} \setminus L$ 有最小元 r ，那么 $L = r^*$ 。因此，一个分割来自有理数，当且仅当它的右口袋有最小元；这个有理数也由最小元唯一确定。

现在，见证造物的奇迹。让我们观察这干净利落的一刀劈下去后，右口袋 U 会发生什么？

右口袋只有两种可能：

情况一：右口袋 U 里面有一个最小元 r 。

由有理分割判别，此时 $L = r^*$ 。比如刀刚好劈在有理数 3 的位置，左口袋是 3^* ，而 3 是右口袋的最小元。这个分割精确地表示已经存在的有理数 3。

情况二：右口袋 U 没有最小元。

对于平方等于 2 的裂缝，先把下面这个左口袋记作 L_2 ：

$$L_2 = \{q \in \mathbb{Q} : q < 0 \text{ 或 } q^2 < 2\}.$$

其中“ $q < 0$ ”保证所有负有理数都在裂缝左边。

【引理（平方根分割，证明略）】利用有理数有序域的运算与稠密性，可以证明 L_2 非空、不是全体 $\mathbb{Q}$ 、向下封闭且没有最大元；其补集也没有最小元。因此它是一个合法分割，并由有理分割判别知道它不等于任何 r^* 。

在情况二中，左口袋没有最大元，右口袋也没有最小元。刀刃仿佛悬停在一道深渊之上。

就在这最令人绝望的时刻，戴德金宣告了神谕：

不要去深渊里寻找数字！这个把有理数队伍一分为二的“左口袋 L ”本身，就是一个绝对合法的全新实体！

【思考训练】

假设我们往左口袋 L 里装入所有小于 5 的有理数，往右口袋 U 里装入所有大于或等于 5 的有理数。

请问：左口袋里有最大的数吗？右口袋里有最小的数吗？这属于上面的情况一还是情况二？

解答：

左口袋里没有最大的数：任取 $x < 5$ ，有理数的稠密性给出

$$x < \frac{x+5}{2} < 5,$$

所以还有一个比 x 大、同时仍小于 5 的有理数属于左口袋。

右口袋里有最小的数，那就是 5 本身。

这属于“情况一”，并且左口袋正是 5^* 。

5.3 完美的实数轴 $\mathbb{R}$ 与确界原理

利用幂集和分离公理，我们把所有满足分割条件的 $L \subseteq \mathbb{Q}$ 收集起来。

- 那些对应着“情况一”的左口袋，我们就给它们贴上旧标签，比如代表有理数 3、有理数 5。
- 那些对应着“情况二（深渊裂缝）”的左口袋，我们就赋予它们全新的合法身份。比如 L_2 会代表一个全新的实数；下面的域结构定理将证明它的平方等于 2，届时才正式称它为“2 的平方根”。

【定义】实数集是所有戴德金分割组成的集合：

$$\mathbb{R} = \{L \in \mathcal{P}(\mathbb{Q}) : L \text{ 是戴德金分割}\}.$$

把前面定义的分割 r^* 作为有理数 r 在 $\mathbb{R}$ 中的像。

【定理（有理数嵌入，证明略）】映射 $r \mapsto r^*$ 是单射。因此，以后可以把 r 与 r^* 视作同一个数。

【定义】对 $A, B \in \mathbb{R}$ ，规定 $A < B$ 当且仅当 $A \subsetneq B$ ，规定 $A \leq B$ 当且仅当 $A \subseteq B$ 。

先用旧有理数检查这个新顺序。若 $r < s$ ，则每个 $q < r$ 也满足 $q < s$ ，所以 $r^* \subseteq s^*$ ；而中点 $(r + s)/2$ 属于 s^* 却不属于 r^* ，所以包含是严格的。反过来，若 $r^* \subsetneq s^*$ ，取 $q \in s^* \setminus r^*$ ，便有 $q < s$ 且 $q \not< r$ ，从而 $r \leq q < s$ ，故 $r < s$ 。因此

$$r < s \iff r^* \subsetneq s^*.$$

例如 $2^* \subsetneq 3^*$ ，其中 $\frac{5}{2}$ 就是“属于 3^* 、不属于 2^* ”的明确见证。把分割按包含关系排队，确实延续了原有的有理数顺序。

【定义】对分割 $A, B \in \mathbb{R}$ ，定义

$$A + B = \{a + b : a \in A, b \in B\},$$

$$-A = \{q \in \mathbb{Q} : \text{存在 } r \in \mathbb{Q} \setminus A \text{ 使 } q < -r\}.$$

零元是 $0^* = \{q \in \mathbb{Q} : q < 0\}$ 。

若 $A, B \geq 0^*$ ，定义

$$A \cdot B = \{q \in \mathbb{Q} : q < 0, \text{ 或 } \exists a \in A \exists b \in B (0 < a \wedge 0 < b \wedge q < ab)\}.$$

其余符号组合分情况定义：若 $A < 0^* \leq B$ ，令 $A \cdot B = -((-A) \cdot B)$ ；若 $B < 0^* \leq A$ ，令 $A \cdot B = -(A \cdot (-B))$ ；若 $A, B < 0^*$ ，令 $A \cdot B = (-A) \cdot (-B)$ 。每个等式右侧出现的乘法都已归入非负情形。

【定理（戴德金分割的域结构，证明略）】上述集合仍是分割，运算与顺序满足有序域的全部公理；嵌入 $r \mapsto r^*$ 保持 $0, 1, +, \cdot$ 和顺序，并且 $L_2 \cdot L_2 = 2^*$ 。逐项验证使用前四章的集合、有序域法则以及第四章“有理数的阿基米德性质”；后者尤其用于构造并验证非零分割的乘法逆元。

这条定理也让旧算术完整地住进了新世界。例如

$$1^* + 2^* = 3^*, \quad -(2^*) = (-2)^*, \quad \left(\frac{3}{2}\right)^* \cdot 2^* = 3^*.$$

这些等式不是把新运算“碰巧解释成”旧运算，而是嵌入保持运算这一结论的具体实例。

【定义】现在定义 $\sqrt{2} = L_2$ 。由刚才的定理， $\sqrt{2} \cdot \sqrt{2} = 2$ ；这个符号到这里才获得正式含义。

你看，我们不需要预先知道 $\sqrt{2}$ 的任何小数展开，就能用“左口袋”这个集合实体极其严密地捕获它！深渊不再是虚无，深渊被打包成了一个合法的盒子！

从分割的定义可以证明，实数集拥有有理数不具备的**确界性质**。在这里，确界性质是定理，而不是另加的一条公理。

什么是确界原理？

想象在实数 $\mathbb{R}$ 的队伍里，有一群数字被关在一个栅栏里，它们向右边走，但右边有一堵绝对无法跨越的墙（这堵墙叫作“上界”）。

【定义】若集合 $A \subseteq \mathbb{R}$ ，实数 u 满足每个 $a \in A$ 都有 $a \leq u$ ，则 u 是上界。若上界 s 还满足每个上界 u 都有 $s \leq u$ ，则 s 是上确界，记作 $\sup A$ 。

【定理（确界定理）】每个非空且有上界的实数子集都有唯一上确界。更明确地，若 $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}$ ，并且存在 $U \in \mathbb{R}$ 使每个 $L \in \mathcal{A}$ 都满足 $L \subseteq U$ ，那么

$$\sup \mathcal{A} = \bigcup_{L \in \mathcal{A}} L.$$

这里“有上界”不可省略。

这个定理的证明正好展示了分割构造的力量。记

$$S = \bigcup_{L \in \mathcal{A}} L.$$

因为 $\mathcal{A}$ 非空，而且它的每个成员都是非空分割，所以 $S \neq \emptyset$ 。共同上界 U 使 $S \subseteq U$ ；又因为 $U \neq \mathbb{Q}$ ，所以 $S \neq \mathbb{Q}$ 。

若 $x \in S$ ，则 x 属于某个 $L \in \mathcal{A}$ 。由 L 向下封闭可知， S 也向下封闭；由 L 没有最大元可知， S 也没有最大元。因此， S 本身是一个戴德金分割。

每个 $L \in \mathcal{A}$ 都包含于 S ，所以 S 是 $\mathcal{A}$ 的上界。若 V 是另一个上界，则对每个 $L \in \mathcal{A}$ 都有 $L \subseteq V$ ，从而

$$S = \bigcup_{L \in \mathcal{A}} L \subseteq V.$$

所以 S 是最小上界。若有两个最小上界，它们彼此不大于对方，由全序的反对称性可知二者相等，唯一性也就得到证明。

【例子】令

$$E = \{x \in \mathbb{R} : x < 1\}.$$

由于 $0 \in E$ ，集合 E 非空；实数 1 是 E 的上界。任何 $u < 1$ 都不是上界：令 $c = (u + 1)/2$ ，则

$$c - u = \frac{1 - u}{2} > 0, \quad 1 - c = \frac{1 - u}{2} > 0,$$

所以 $u < c < 1$ ，即 $c \in E$ 且 $c > u$ 。因此， u 不是 E 的上界。任取 E 的上界 v ，上面的论证说明 v 不可能小于 1；由全序性， $1 \leq v$ 。故 $\sup E = 1$ 。注意，1 不属于 E ；上确界不必是集合中的最大元。

【定理（非负平方根，证明略）】对每个 $a \geq 0$ ，集合 $\{x \in \mathbb{R} : x \geq 0 \text{ 且 } x^2 \leq a\}$ 非空且有上界；它的上确界是唯一满足 $r \geq 0$ 、 $r^2 = a$ 的实数，记为 $\sqrt{a}$ 。

【推论（有理数不具确界性，证明略）】集合

$$S_2 = \{q \in \mathbb{Q} : q \geq 0 \text{ 且 } q^2 < 2\}$$

非空并且在 $\mathbb{Q}$ 中有上界，却没有属于 $\mathbb{Q}$ 的上确界。否则若 $s \in \mathbb{Q}$ 是其上确界，便可由稠密性与有理数的阿基米德性质推出 $L_2 = s^*$ ，这与有理分割判别和“ L_2 不等于任何 r^* ”矛盾。

但在实数 $\mathbb{R}$ 中，每个非空且有上界的集合都有一个实数上确界。例如，与平方等于 2 的裂缝对应的实数就是 $\sqrt{2}$ 。

至此，实数 $\mathbb{R}$ 变成了一条具有确界性质的实数队伍。我们称它为**实数轴**；“连续映射”将在下一章另行定义，这里不借用那个尚未出现的概念。

5.4 永恒的追逐——柯西列与完备性

为了让你更深刻地感受到这种“没有次序缺口”的性质，我们换一种动态的视角来审视实数宇宙。

我们在上一章的多项式环里学过“序列”的概念：它像一排永远写不到头的盒子，第 0 号、第 1 号、第 2 号，依次延伸下去；现在，我们在里面装入实数。

法国数学家柯西提出了一种极度严苛的追逐游戏：

在这排无尽的盒子里，随着编号越来越大，里面的数字会不会“互相挤作一团”？

【定义】对 $x \in \mathbb{R}$ ，定义它的**绝对值**为

$$|x| = \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

【引理（绝对值法则）】对任意 $x, y \in \mathbb{R}$ ：

1. $|x| \geq 0$ ，且 $|x| = 0$ 当且仅当 $x = 0$ ；
2. $|-x| = |x|$ ；
3. $|xy| = |x||y|$ ；
4. $|x + y| \leq |x| + |y|$ （三角不等式）；
5. $||x| - |y|| \leq |x - y|$ （反三角不等式）；

6. 若 $r \geq 0$, 则 $|x| \leq r$ 当且仅当 $-r \leq x \leq r$; 若 $r > 0$, 则 $|x| < r$ 当且仅当 $-r < x < r$ 。

这些结论由第四章的有序域符号法则分情况推出。其中三角不等式有一个能看见结构的短证明：绝对值定义给出

$$-|x| \leq x \leq |x|, \quad -|y| \leq y \leq |y|.$$

把同方向的不等式相加，得到

$$-(|x| + |y|) \leq x + y \leq |x| + |y|.$$

因为 $|x| + |y| \geq 0$ ，第 6 条中的非严格不等式形式便给出 $|x + y| \leq |x| + |y|$ 。

【定义】 两实数 x, y 的**数值距离**是 $|x - y|$ ；下一章会验证它满足抽象度量的全部公理。

【定义】 实数序列是映射 $a : \mathbb{N} \rightarrow \mathbb{R}$ ，记 $a_n = a(n)$ 。

【定义】 若对每个 $\epsilon > 0$ ，都存在 $N \in \mathbb{N}$ ，使任意 $m, n \geq N$ 都满足 $|a_m - a_n| < \epsilon$ ，就称 (a_n) 为**柯西列**。

【定义】 若存在 $L \in \mathbb{R}$ ，使对每个 $\epsilon > 0$ 都存在 $N \in \mathbb{N}$ ，并使每个 $n \geq N$ 都满足 $|a_n - L| < \epsilon$ ，就称 a_n 收敛到 L ，记作 $a_n \rightarrow L$ 。

最简单的例子是常值序列：固定 $c \in \mathbb{R}$ ，令每个 $a_n = c$ 。对任意 $\epsilon > 0$ ，取 $N = 0$ 。于是对每个 $n \geq N$ ，都有 $|a_n - c| = 0 < \epsilon$ ，所以 $a_n \rightarrow c$ ；同样，因为 $|a_m - a_n| = 0$ ，它也是柯西列。

【引理（极限的基本性质）】 实数序列的极限若存在就唯一；每个收敛实数序列都是柯西列。

先证唯一性。假设同一序列同时收敛到 L 和 M ，且 $L \neq M$ 。令 $\delta = |L - M| > 0$ ，并在两个收敛定义中都取 $\epsilon = \delta/3$ 。分别得到 N_1, N_2 后，取 $N = N_1 + N_2$ ，则 $N \geq N_1, N_2$ ，从而

$$|L - M| \leq |L - a_N| + |a_N - M| < \frac{2\delta}{3},$$

即 $\delta < 2\delta/3$ ，与 $\delta > 0$ 矛盾。所以 $L = M$ 。

再设 $a_n \rightarrow L$ 。给定 $\epsilon > 0$ ，在收敛定义中把容许误差取为 $\epsilon/2$ ，得到 N 。若 $m, n \geq N$ ，则

$$|a_m - a_n| \leq |a_m - L| + |L - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

这正是柯西条件。注意这个方向只用三角不等式；反方向“每个柯西列都收敛”才需要实数的完备性。

柯西列就像一群在黑暗中不断互相靠近的萤火虫。它们越往后飞，就挤得越紧，仿佛在疯狂地逼近某个神秘的目标点。

这里，宇宙法庭展现了实数集 $\mathbb{R}$ 的终极威力：

【定理（柯西完备性，证明略）】 每个实数柯西列都收敛到唯一实数。这个定理可以由确界定理推出。

【定理（阿基米德性质，证明略）】 对每个实数 x ，存在自然数 n 使 $n > x$ 。等价地，对每个 $\epsilon > 0$ ，存在 $N \in \mathbb{N}$ 且 $N \geq 1$ ，使每个 $n \geq N$ 都满足 $0 < 1/n < \epsilon$ ；特别地， $1/10^n \rightarrow 0$ 。

【引理（整数夹逼）】 对每个实数 t ，都存在整数 m 使

$$m - 1 \leq t < m.$$

证明：由阿基米德性质取 $N \in \mathbb{N}$ 使 $N > -t$ ，于是 $s = t + N > 0$ 。自然数集合

$$C = \{k \in \mathbb{N} : k > s\}$$

非空；由自然数的良序性，它有最小元 k 。因为 $s > 0$ ，所以 $k \geq 1$ ；最小性又说明 $k - 1 \leq s < k$ 。令 $m = k - N \in \mathbb{Z}$ ，把这串不等式同时减去 N ，便得到 $m - 1 \leq t < m$ 。

【定理（有理数在实数中稠密）】 若 $x < y$ 是实数，则存在 $q \in \mathbb{Q}$ 使 $x < q < y$ 。

证明：由阿基米德性质取 $n \in \mathbb{N}$ 且 $n \geq 1$ ，使 $1/n < y - x$ 。对 $t = nx$ 使用整数夹逼，取 $m \in \mathbb{Z}$ 使

$$m - 1 \leq nx < m.$$

令 $q = m/n \in \mathbb{Q}$ 。由 $nx < m$ 得 $x < q$ ；由 $m \leq nx + 1$ 得

$$q = \frac{m}{n} \leq x + \frac{1}{n} < y.$$

所以 $x < q < y$ 。这条定理以后会让我们只用可数多个有理坐标的小球描述多维实数空间的全部开集。

不管萤火虫们怎样飞，只要它们满足柯西条件，实数轴上就一定有一个点作为它们的极限。绝不会出现“序列满足柯西条件，却没有实数极限”的情况。

正是这极其霸道的“完备性”，赋予了我们未来去研究无穷无尽的微小变化的合法权力。

【思考训练】

假设有一群萤火虫组成一个序列，它们依次停在 $1, 0.1, 0.01, 0.001, 0.0001, \dots$ 上。请问，随着编号越来越大，它们互相之间的距离是不是越来越小？它们在逼近实数轴上的哪一个目标点？

解答：

把第一项编号为 0，这个序列就是 $a_n = 1/10^n$ 。上面的阿基米德性质已经给出 $1/10^n \rightarrow 0$ ，所以目标点是 0；再由“收敛序列必为柯西列”，对任意 $\epsilon > 0$ ，足够靠后的任意两项之间距离都小于 ϵ 。这才是“互相之间的距离越来越小”的严格含义。

5.5 终极魔法扩展——复数域 $\mathbb{C}$ 的降临

实数轴 $\mathbb{R}$ 在有序与完备性方面已经十分完整。在实数域里，可以进行加、减、乘以及除数非零的除法，原先的次序缺口也已被填补。

但是，当我们在这个完美的域里，尝试用上一章的代数骨架去制造一种特殊的挑战时，最后一道封印出现了。

在第四章的通用多项式构造中取系数环 $R = \mathbb{R}$ ，便得到 $\mathbb{R}[X]$ 。现在考察其中的多项式 $X^2 + 1$ 。

我们问：在完美的实数 $\mathbb{R}$ 里，能不能找到一个数 x ，让 $x \times x + 1 = 0$ 成为【真】命题？

根据第四章有序域的平方正性法则：

- 如果 x 是正数，两个正数相乘必定还是正数， $x \times x$ 是正的。正数加上 1，绝对不等于 0。
- 如果 x 是负数，“负负得正”的绝对代数法则规定，它乘以它自己依然是正数。正数加上 1，绝对不等于 0。
- 如果 x 就是 0，0 乘以 0 等于 0，加上 1 等于 1，也不等于 0。

这意味着，在实数字宙里，任何数自己乘自己，都不可能等于 -1！

完美的实数直线，在纯粹的代数挑战面前，依然存在着一个死胡同！

为了打破这道封印，我们不必提前引入更高级的商环工具，直接利用有序对即可完成构造。

【定义】 复数集是笛卡尔积 $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ 。在实数对上定义

$$(a, b) + (c, d) = (a + c, b + d),$$

$$(a, b)(c, d) = (ac - bd, ad + bc).$$

定义 $0 = (0, 0)$ 、 $1 = (1, 0)$ ，并把实数 a 识别为 $(a, 0)$ 。定义虚数单位 $i = (0, 1)$ 。

直接计算得到

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1.$$

每个 (a, b) 都唯一写成 $a + bi$ 。

【推论】 对 $a, b \in \mathbb{R}$ ，第四章的有序域符号法则给出 $a^2, b^2 \geq 0$ ，并且 $a^2 + b^2 = 0$ 当且仅当 $a = b = 0$ 。

【定理（复数域，证明略）】 上述加法和乘法使 $\mathbb{C}$ 成为域。若 $(a, b) \neq (0, 0)$ ，其乘法逆元是

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right),$$

因为上述推论在 $(a, b) \neq (0, 0)$ 时给出 $a^2 + b^2 > 0$ 。这证明复数除法确实可行。

【说明】由第四章的通用构造，现在也有了多项式环 $\mathbb{C}[X]$ 。“每个非常数复系数多项式都有复根”叫代数基本定理，是远比“ $i^2 = -1$ ”更深的结论，本书后文不会在未证明时把它当作工具。

【思考训练】

在刚刚诞生的复数域里，神圣的虚数单位 i 自己乘以自己等于 -1 （即 $i \times i = -1$ ）。请问，如果把四个 i 乘在一起（ $i \times i \times i \times i$ ），结果等于多少？

解答：

我们把它们两个两个组合起来。

前两个乘起来是 -1 ，后两个乘起来也是 -1 。

然后把这两个结果相乘： $(-1) \times (-1) = 1$ 。

这就是虚数的魔法流转。

5.6 本章结语

这一章，你亲历了宇宙诞生以来最惊心动魄的几次造物奇迹。

你用纯逻辑的猜谜游戏说明了有理数的稠密不等于完备，并证明了有理数中不存在平方等于 2 的数。

你挥舞戴德金分割的利刃，将深渊本身打包成合法的口袋，铸造出了具有确界性质并且柯西完备的实数轴 $\mathbb{R}$ 。

你用柯西的萤火虫序列，验证了实数的柯西完备性。

最后，你用实数有序对和明确的两条运算规则构造了虚数单位 i ，并证明这些有序对组成复数域 $\mathbb{C}$ 。

我们的数字大厦，至此已经完美封顶。它们不再是简单的符号，而是由纯粹逻辑、集合嵌套、代数骨架和完备性法则共同编织的终极神器。

但是，只有直线和符号的世界未免太过孤单。

在下一章，我们要暂时从数轴走向更广阔的空间。我们将利用刚刚构造的实数 $\mathbb{R}$ 建立严格的距离概念，并进一步定义什么是“靠近”、什么是“连通”、什么是拓扑意义下的“形状”。

准备好进入拓扑与距离的奇异世界吧。

6 距离与形态 —— 奇异的拓扑魔法

在上一章，我们挥舞戴德金之刀，填平了有理数的次序缺口。我们终于拥有了具有确界性质并且柯西完备的实数轴 $\mathbb{R}$ 。

现在，我们的宇宙里充满了坚实的数字点。但如果我们要在这广袤的宇宙中航行，如果我们要研究事物是如何“变化”和“运动”的，我们还缺少一样极其关键的东西。

想象一下：如果你是一只生活在纯逻辑宇宙里的蚂蚁，你身边只有密密麻麻的数字，没有任何现实世界里的尺子。

你怎么知道点 A 离点 B 有多远？

你怎么判断自己是在“靠近”目标，还是在“远离”目标？

如果你面前有一团扭曲的数字集合，你怎么判断它是不是“连在一起”的？

这一章，我们要用集合与逻辑的语言，严格定义“距离”、“靠近”和“形状”。

更疯狂的是，我们最终会抛弃“距离”本身，提炼出宇宙中最纯粹、最奇异的几何法则——**拓扑**。

欢迎来到形态的绝对真理世界。

6.1 虚拟的尺子 —— 创造度量空间

什么是“距离”？在公理化的宇宙里，我们不能用真实的尺子去量，我们必须把它定义为一个合法的规则（映射）。

设 X 是集合。

【定义】 X 上的**度量**是映射 $d : X \times X \rightarrow \mathbb{R}$ ，并且对任意 $x, y, z \in X$ 满足下列条件。

但法庭不会随便批准。为了让这个规则真正配得上“距离”这两个字，它必须死死遵守三条铁律：

1. **非负与正定性**: $d(x, y) \geq 0$, 并且 $d(x, y) = 0$ 当且仅当 $x = y$ 。
2. **对称性**: 交换两点的次序不会改变距离, 即 $d(x, y) = d(y, x)$ 。
3. **三角不等式**: $d(x, z) \leq d(x, y) + d(y, z)$ 。把它想成“经 y 绕道不会更短”只是直觉; 一般度量空间未必真的存在道路或“最快路线”。

只要一个集合配备了这样一个合法的距离规则 d , 就称有序对 (X, d) 为一个**度量空间**。

在上一章的实数轴上已经定义了绝对值。令 $d(x, y) = |x - y|$ 。

【定理 (实数的通常度量, 证明略)】上一章的绝对值法则分别给出非负与正定性、对称性和三角不等式, 所以这个 d 满足全部度量公理。

例如, 点 2 和点 5 的距离是 3; 点 -1 和点 3 的距离是 4。

度量不一定来自直尺。

【例子 (离散度量)】在任意集合 X 上定义

$$d_{\text{disc}}(x, y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

非负、正定与对称性可直接从定义读出。验证三角不等式只需分两种情况。若 $x = z$, 左边为 0; 若 $x \neq z$, 则 $x = y$ 与 $y = z$ 不可能同时成立, 所以 $d_{\text{disc}}(x, y)$ 、 $d_{\text{disc}}(y, z)$ 至少有一个等于 1, 于是

$$d_{\text{disc}}(x, z) = 1 \leq d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z).$$

因此任何集合, 不论其中的元素是什么, 都至少能配备这样一种合法度量。

现在, 我们终于在虚拟的宇宙中, 拥有了一把合法的、无形的光束尺。

【思考训练】

假设有一条神奇的街道, 从你家 (点 A) 到学校 (点 B) 的距离是 500 米, 但是从学校回家时, 距离却变成了 800 米。

请问, 这个街道的距离规则, 符合度量空间的铁律吗? 它违背了哪一条?

解答:

不符合。它违背了第二条“对称性”: 交换起点与终点后, 距离应当保持不变。

6.2 画出开球与“靠近”的绝对标准

有了这把尺子，我们终于可以严格定义“某个点附近”。

在上一章，我们定义了实数序列的收敛。稍后我们会把同一个想法推广到任意度量空间；现在先锻造它所需的开球。

为了把这种“逼近”具象化，我们需要在空间里画出一个极其重要的几何结构：**开球**。

把与点 a 的距离小于 r 的所有点收集起来，就得到下面的集合。

取实数 $r > 0$ 。

【定义】 以 a 为中心、 r 为半径的开球是

$$B_r(a) = \{x \in X : d(x, a) < r\}.$$

距离恰好等于 r 的点不属于 $B_r(a)$ 。

【定义】 集合 $\{x \in X : d(x, a) = r\}$ 称为度量球面。现在只记录“球面被排除”这一事实，不先把它叫作边界；“皮”只是帮助想象的比喻。

为什么定义中使用严格不等式？因为我们要用开球来描述一种严格的局部邻近关系。

【定义】 子集 $U \subseteq X$ 是开集，当且仅当对每个 $x \in U$ ，都存在 $r > 0$ 使 $B_r(x) \subseteq U$ 。

这就是度量空间中的开集。这个条件只保证每个点附近有一个安全小球，并不表示反复迈小步永远不能走出 U 。

离散度量给出一个可以完全算清的例子。若 $0 < r \leq 1$ ，那么

$$B_r(a) = \{a\};$$

若 $r > 1$ ，则 $B_r(a) = X$ 。因此任取子集 $U \subseteq X$ 和点 $a \in U$ ，半径为 1 的开球 $B_1(a) = \{a\}$ 都包含于 U 。所以离散度量下的**每一个**子集都是开集。这里没有凭图形判断；结论只来自距离的两个可能值 0 与 1。

【引理（开球是开集）】在任意度量空间中，每个开球 $B_r(a)$ 都是开集。事实上，若 $y \in B_r(a)$ ，则

$$s = r - d(y, a) > 0.$$

对任意 $z \in B_s(y)$ ，三角不等式给出

$$d(z, a) \leq d(z, y) + d(y, a) < s + d(y, a) = r,$$

所以 $B_s(y) \subseteq B_r(a)$ 。这正好验证了开集定义对球内每个点的要求。

【直觉】开集中的每个点都有一个仍留在集合内的小开球；这描述的是局部性质，不保证经过许多小步后永远不会走出集合。

【思考训练】

请问，在实数轴上，开区间 $(0, 1)$ 是开集吗？如果站在 0.0001 这个离端点很近的位置，还能找到完全落在 $(0, 1)$ 内的小开球吗？

解答：

它是开集。严格地说，任取 $x \in (0, 1)$ ，令

$$r = \frac{x(1-x)}{2}.$$

有序域法则给出 $r > 0$ 、 $r < x$ 且 $r < 1 - x$ 。若 $|y - x| < r$ ，由绝对值定义得到 $x - r < y < x + r$ ，从而 $0 < y < 1$ 。因此 $B_r(x) \subseteq (0, 1)$ 。

站在 0.0001 时也能按同一公式选到正半径。关键是区间内的每个点都有自己的小球；0 和 1 不属于这个区间，并不意味着开集没有边界。正式的边界概念要等拓扑定义完成后再引入。

6.3 抛弃尺子——纯粹拓扑空间的诞生

现在，我们要进行全书最疯狂的一次思想跳跃。

在第一节，我们费尽心机定义了“距离尺子 d ”。

现在，我们要把这把尺子彻底砸碎。

为什么？因为在未来的弯曲宇宙里，空间可能被像橡胶一样任意拉伸、扭曲。在被拉伸的空间里，“距离”这种东西是极其脆弱和善变的。今天两点之间的距离是 1，明天宇宙一膨胀，距离就变成了 100。

如果连尺子都在变形，我们怎么去寻找宇宙中永恒不变的几何真理？

数学家发现了一个惊天秘密：许多重要性质只需要知道哪些集合是开集，不需要知道距离的具体数值。于是我们暂时忘掉尺子的刻度，只保留开集之间的包含、并与交。

于是，宇宙法庭颁布了凌驾于距离之上的绝对形态法则——**拓扑**。

【定义】 集合 X 上的一个拓扑是子集族 $\mathcal{T} \subseteq \mathcal{P}(X)$ ，满足：

1. **包含空集与全集**： $\emptyset \in \mathcal{T}$ 且 $X \in \mathcal{T}$ 。
2. **任意并封闭**：若 I 是任意集合，并且对每个 $i \in I$ 都有 $U_i \in \mathcal{T}$ ，则 $\bigcup_{i \in I} U_i \in \mathcal{T}$ 。
3. **有限交封闭**：若 $n \geq 1$ 且 $U_1, \dots, U_n \in \mathcal{T}$ ，则 $U_1 \cap \dots \cap U_n \in \mathcal{T}$ 。

满足这些条件时， $(X, \mathcal{T})$ 称为**拓扑空间**， $\mathcal{T}$ 的成员称为**开集**。

【术语约定】 若集合 $N \subseteq X$ 包含某个含点 x 的开集，就称 N 是 x 的一个**邻域**；若 N 本身还是开集，就称它是 x 的一个**开邻域**。

【定理（度量诱导拓扑）】 度量空间中按开球定义的全部开集满足这三条公理，所以每个度量空间都自然产生一个拓扑，称为该度量**诱导的拓扑**。实数的通常度量 $d(x, y) = |x - y|$ 所诱导的拓扑称为实数的**通常拓扑**。

这条定理有一个简短的证明。空集没有元素，因此开集条件自然成立；对每个 $x \in X$ ，都有 $B_1(x) \subseteq X$ ，所以 $\emptyset$ 与 X 都是开集。

若 $x \in \bigcup_{i \in I} U_i$ ，则 x 属于某个 U_i 。这个 U_i 中包含 x 的开球也包含于整个并集，所以任意并仍是开集。

若 $x \in U \cap V$ ，分别取 $r, s > 0$ ，使 $B_r(x) \subseteq U$ 且 $B_s(x) \subseteq V$ 。令 $t = \min\{r, s\} > 0$ ，便有

$$B_t(x) \subseteq U \cap V.$$

所以两个开集的交仍是开集；反复使用这个结论便得到有限交封闭。

【例子（两种极端拓扑）】对任意集合 X ，幂集 $\mathcal{P}(X)$ 满足拓扑公理，称为 X 上的**离散拓扑**；上面的离散度量恰好诱导它。另一方面，只含空集与整个空间的集合族

$$\{\emptyset, X\}$$

也满足拓扑公理，称为**平凡拓扑**。前者宣布每个子集都开，后者只宣布空集与整个空间开。拓扑公理允许这两个极端，也允许介于它们之间的许多选择；当 $X = \emptyset$ 时，这两个拓扑恰好相同。

【定义】子集 $F \subseteq X$ 是闭集，当且仅当补集 $X \setminus F$ 是开集。集合 $A \subseteq X$ 的闭包 $\bar{A}$ 是包含 A 的所有闭集的交集。

闭超集族非空，因为 X 本身是闭集并且包含 A 。任意非空族闭集之交仍闭：这个交集的补集是各闭集补集的并集，因而是开集。最后，交集包含 A ，并且包含于每个含 A 的闭集，所以 $\bar{A}$ 的确是包含 A 的最小闭集。

【定义】集合 $A \subseteq X$ 的**边界**是

$$\partial A = \bar{A} \cap \overline{X \setminus A}.$$

边界点不一定属于 A 。

【定理（开区间的边界，证明略）】在实数的通常拓扑中， $\partial(0, 1) = \{0, 1\}$ 。所以，“开集”不等于“没有边界”；它只要求集合内每个点都有一个留在集合内的开球。

【定义】若 $A \subseteq X$ ，则 A 的子空间拓扑由全部形如 $A \cap U$ 的集合组成，其中 U 是 X 的开集。

【定理（证明略）】这些集合满足拓扑公理。谈论一个子集自身是否连通或紧致时，默认使用这个拓扑。

【定义】子族 $\mathcal{B} \subseteq \mathcal{T}$ 是拓扑的一个基，当且仅当对每个开集 U 和每个 $x \in U$ ，都存在 $B \in \mathcal{B}$ 使 $x \in B \subseteq U$ 。若存在至多可数的基，就称空间第二可数。

【定义】拓扑空间 X 称为**Hausdorff 空间**，当且仅当对任意不同的 $x, y \in X$ ，都存在互不相交的开集 U, V ，使 $x \in U$ 且 $y \in V$ 。这条性质保证不同的点能被开集分开。

【定理（度量空间是 Hausdorff 空间）】若 $x \neq y$, 令 $\delta = d(x, y) > 0$ 。开球 $B_{\delta/3}(x)$ 与 $B_{\delta/3}(y)$ 不可能相交；否则若 z 同时属于二者，三角不等式会给出

$$\delta = d(x, y) \leq d(x, z) + d(z, y) < \frac{2\delta}{3},$$

矛盾。由“开球是开集”引理，这两个开球就是分开 x, y 所需的开集。

为了给后文的多维空间准备拓扑基础,先明确有限坐标元组的集合含义。对 $n \in \mathbb{N}$, 令

$$[n] = \{k \in \mathbb{N} : 1 \leq k \leq n\},$$

并定义 $\mathbb{R}^n$ 为所有映射 $x : [n] \rightarrow \mathbb{R}$ 组成的集合；当 $n \geq 1$ 时写作 $x = (x_1, \dots, x_n)$ 。沿用第五章把 $\mathbb{Q}$ 认作 $\mathbb{R}$ 子集的嵌入约定，定义

$$\mathbb{Q}^n = \{q \in \mathbb{R}^n : \text{对每个 } j \in [n] \text{ 都有 } q_j \in \mathbb{Q}\}.$$

所以，这里的 $\mathbb{Q}^n$ 明确是 $\mathbb{R}^n$ 的子集。它与“从 $[n]$ 到 $\mathbb{Q}$ 的映射”逐坐标一一对应，但按照第二章严格的映射三元数据定义，二者的陪域不同，不能不加说明地视为同一个对象。

特别地， $[0] = \emptyset$ ，所以 $\mathbb{R}^0$ 只有一个空坐标元组。第八章会在这些元组上加入向量运算；这里暂时只研究它们的开集。

实数全序与归纳法保证有限个实数中有最大者。对 $n \geq 1$ ，定义**最大距离**

$$d_\infty(x, y) = \max_{1 \leq j \leq n} |x_j - y_j|.$$

【定理(最大距离)】 d_∞ 是 $\mathbb{R}^n$ 上的度量。非负、正定与对称性可逐坐标得到；而对每个 j ,

$$|x_j - z_j| \leq |x_j - y_j| + |y_j - z_j| \leq d_\infty(x, y) + d_\infty(y, z),$$

再对 j 取最大值，便得到三角不等式。

【定义】对 $n \geq 1$ ，把最大距离 d_∞ 诱导的拓扑称为 $\mathbb{R}^n$ 的**通常拓扑**； $\mathbb{R}^0$ 的通常拓扑是其唯一的拓扑。第八章定义欧氏距离后，才会证明欧氏距离诱导同一个拓扑；本章不提前使用那个结论。

【可数性引理（证明略）】至多可数集合的子集和映射像仍至多可数，有限个至多可数集合的笛卡尔积仍至多可数。核心工具是按 $m + n$ 的大小逐条对角线枚举 $\mathbb{N} \times \mathbb{N}$ 。整数可按

$$0, 1, -1, 2, -2, 3, -3, \dots$$

枚举，而每个有理数都有分母非零的整数对作为代表，所以 $\mathbb{Z}$ 、 $\mathbb{Q}$ 都至多可数。从 $[n]$ 到 $\mathbb{Q}$ 的映射集与有限个 $\mathbb{Q}$ 的笛卡尔积一一对应，因而至多可数；再用前面说明的逐坐标双射，可知作为 $\mathbb{R}^n$ 子集定义的 $\mathbb{Q}^n$ 也至多可数。

【定理 ($\mathbb{R}^n$ 的分离性与第二可数性)】对每个 $n \in \mathbb{N}$ ，通常拓扑下的 $\mathbb{R}^n$ 是 Hausdorff 且第二可数。

当 $n = 0$ 时空间只有一个点，结论直接成立。设 $n \geq 1$ 。Hausdorff 性由最大距离是度量立即得到。令 $B_\rho^\infty(q)$ 表示最大距离下以 q 为中心、 ρ 为半径的开球，并令

$$\mathcal{B}_n = \{B_\rho^\infty(q) : q \in \mathbb{Q}^n, \rho \in \mathbb{Q}, \rho > 0\}.$$

可数性引理说明 $\mathcal{B}_n$ 至多可数。现在设 U 是通常开集且 $x \in U$ 。可取 $\epsilon > 0$ 使 $B_\epsilon^\infty(x) \subseteq U$ 。由第五章的有理数稠密性，选取正有理数 $\rho < \epsilon/3$ ，并对每个坐标选取 $q_j \in \mathbb{Q}$ 使

$$|x_j - q_j| < \rho.$$

令 $q = (q_1, \dots, q_n)$ 。按照 $\mathbb{Q}^n$ 的定义， $q \in \mathbb{Q}^n$ ；又因为各坐标距离都小于 ρ ，所以 $d_\infty(x, q) < \rho$ ，即 $x \in B_\rho^\infty(q)$ 。若 $y \in B_\rho^\infty(q)$ ，则

$$d_\infty(y, x) \leq d_\infty(y, q) + d_\infty(q, x) < 2\rho < \epsilon,$$

所以

$$x \in B_\rho^\infty(q) \subseteq B_\epsilon^\infty(x) \subseteq U.$$

这证明 $\mathcal{B}_n$ 是一个至多可数的基。

【定理（子空间继承）】Hausdorff 性与第二可数性都由子空间继承。确切地说，若 $A \subseteq X$ 且 X 是 Hausdorff 空间，那么对任意不同的 $x, y \in A$ ，在 X 中分开它们的开集与 A 相交后，仍是在 A 中分开 x, y 的开集。若 $\mathcal{B}$ 是 X 的至多可数基，则

$$\mathcal{B}|_A = \{A \cap B : B \in \mathcal{B}\}$$

是 A 的至多可数基：任取子空间开集 $A \cap U$ 及其中一点，先在 X 中用某个 $B \in \mathcal{B}$ 缩进 U ，再与 A 相交即可。因此 $\mathbb{R}^n$ 的每个子空间也都是 Hausdorff 且第二可数。

在只给出拓扑 $\mathcal{T}$ 的空间里，数据中没有距离数值。

【定义】 X 中的序列是映射 $x : \mathbb{N} \rightarrow X$ ，记作 (x_n) 。对序列 (x_n) 和点 $a \in X$ ，若对每个满足 $a \in U$ 的开集 U ，都存在 $N \in \mathbb{N}$ ，使每个 $n \geq N$ 都有 $x_n \in U$ ，就称 x_n 收敛到 a 。

【定理（度量收敛的等价形式，证明略）】在度量诱导的拓扑中， $x_n \rightarrow a$ 当且仅当对每个 $\epsilon > 0$ ，最终都有 $d(x_n, a) < \epsilon$ 。

这就是拓扑魔法：即使数据中没有距离数值，仍能用开集的包含关系表达序列收敛。

在离散拓扑中，收敛还有一个非常具体的形式。

【命题】 $x_n \rightarrow a$ 当且仅当序列从某一项起恒等于 a 。若 $x_n \rightarrow a$ ，把开集取成单点集 $\{a\}$ ，收敛定义就给出某个 N ，使 $n \geq N$ 时 $x_n = a$ 。反过来，若序列从某一项起恒等于 a ，那么任何包含 a 的开集最终都包含 x_n ，所以序列收敛到 a 。

6.4 撕不断与逃不出——连通性与紧致性

有了拓扑开集，我们终于可以严格定义形状的两种重要属性。

1. 撕不断：连通性

如果你面前有一团复杂的空间，你怎么证明它是“完整的一块”，而不是“断开的两块”？

直觉上，你会说“它们中间没有缝”，但这句话还不是严格的定义或证明。

我们用开集给出严格定义：

【定义】空间 X 是断开的，当且仅当存在两个非空开集 U, V ，使 $U \cap V = \emptyset$ 且

$$X = U \cup V。$$

换句话说，你能用两个互不相交的非空开集，把整个空间恰好分成两块。

若不存在这样的 U, V ，就称 X **连通**。非空条件不可省略，否则 $\emptyset$ 与 X 会把每个空间错误地判成断开。

2. 逃不出：紧致性

紧致性的核心是：无论用多少个开集覆盖整个空间，总能从中选出有限多个，仍然覆盖整个空间。

想象许多开集口袋盖住了空间 X ，而且每个点都至少落在一个口袋里；这就是开覆盖。紧致性要求，这样的覆盖中总有有限多个口袋已经足够。

【定义】若 $\{U_i\}_{i \in I}$ 是 X 的开集族，并满足 $X = \bigcup_{i \in I} U_i$ ，就称它为 X 的**开覆盖**。若每个开覆盖都含有仍能覆盖 X 的有限子族，就称 X **紧致**。

从“无穷”的混沌中，强制提取出“有限”的秩序！这就是“紧致性”。

【定理（闭区间紧致性，证明略）】若 $a, b \in \mathbb{R}$ 且 $a \leq b$ ，则闭区间 $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ 紧致。证明可由实数确界定理推出。紧致性以后允许我们把局部信息缩减到有限多个局部区域，但它本身不等于“有限求和”。

一个最小例子就能区分连通与紧致。设 $X = \{p, q\}$ 且 $p \neq q$ ，赋予离散拓扑。单点集 $\{p\}$ 与 $\{q\}$ 都是非空开集，互不相交且并为 X ，所以 X 不连通。可是 X 紧致：给定任意开覆盖，只需从覆盖中取一个盖住 p 的开集和一个盖住 q 的开集，至多两个就已覆盖 X 。

同样的理由给出更一般的结论。

【命题（有限空间紧致）】任意有限拓扑空间都紧致。若 X 为空，空子族已经覆盖它；若 X 非空，就给每个点从开覆盖中选一个盖住它的成员。因为点只有有限多个，选出的覆盖成员也只有有限多个。可见，“有限”保证紧致，却不保证连通。

6.5 空间之间的舞蹈——连续映射

现在，设 X 和 Y 是两个拓扑空间；它们也可以是同一个空间。

利用第二章的映射法则，我们建立一个从 X 指向 Y 的配对规则 f （也就是说，当点在 X 宇宙里移动时，它在 Y 宇宙里的影子也会跟着移动）。

我们怎样用已经定义的开集，判定这个映射是否保留空间的局部结构？

在高贵的拓扑学里，我们只需要一句话，一句没有任何数字公式的纯魔法咒语：

【定义】 对 $V \subseteq Y$ ，原像是 $f^{-1}(V) = \{x \in X : f(x) \in V\}$ 。映射 $f : X \rightarrow Y$ 连续，当且仅当 Y 中每个开集 V 的原像 $f^{-1}(V)$ 都是 X 中的开集。

只要满足这个条件，就称映射 f 连续。

“开集的原像仍开”就是连续性的拓扑定义。橡胶不撕裂只是直觉图像，正式推理只使用原像条件。

【例子】 固定 $y_0 \in Y$ ，令 $f(x) = y_0$ 。对 Y 中任意开集 V ，若 $y_0 \in V$ ，则 $f^{-1}(V) = X$ ；若 $y_0 \notin V$ ，则 $f^{-1}(V) = \emptyset$ 。两种原像都开，所以每个常值映射都连续。再者，若定义域 X 取离散拓扑，那么任意映射 $f : X \rightarrow Y$ 都连续，因为每个原像只是 X 的一个子集，而离散拓扑中的所有子集都开。

【定理（连续映射保持连通与紧致，证明略）】 若 $f : X \rightarrow Y$ 连续且 X 连通，则像集 $f(X)$ 在子空间拓扑下连通；若 X 紧致，则 $f(X)$ 在子空间拓扑下紧致。这两条结论直接由开集原像、开分割和开覆盖的定义推出。

【思考训练】 设 X 非空且连通， $D = \{p, q\}$ 带离散拓扑且 $p \neq q$ 。连续映射 $f : X \rightarrow D$ 能同时取到 p 与 q 吗？

解答：不能。若两个值都能取到，就有 $f(X) = D$ 。连续映射保持连通，所以 D 应当连通；但 $\{p\}$ 与 $\{q\}$ 已把 D 分成两个互不相交的非空开集，矛盾。因此，这样的连续映射只能取到一个值，也就是常值映射。

【定理（连续映射的组合，证明略）】 恒等映射连续；若 $f : X \rightarrow Y$ 与 $g : Y \rightarrow Z$ 连续，则 $g \circ f : X \rightarrow Z$ 连续。这是因为对任意开集 $W \subseteq Z$ ，都有 $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$ 。

【定义】 若 $f : X \rightarrow Y$ 是双射， f 连续，且逆映射 $f^{-1} : Y \rightarrow X$ 也连续，就称 f 为同胚。

只要两个空间之间存在同胚，拓扑学就把它们视为同一种拓扑形状。

【定理（证明略）】 恒等映射是同胚；同胚的逆映射仍是同胚；两个同胚的复合仍是同胚。因此，“同胚”确实是一种等价关系。严格判定同胚必须给出双射并验证两个方向的连续性，不能只凭外观判断。

【直觉练习（不作为证明）】

把气球皮理想化为一张没有粘合、自相交或破洞的橡胶曲面。它在轻微放气和吹圆后是否具有同一种拓扑形状？

解答：

这个想象只能帮助建立直觉，不能单独证明同胚。若“轻微形变”被精确定义成一个双射，并且形变与反向形变都连续，那么它按定义就是同胚；若形变把不同点粘成同一点，它便不是双射，因而不是同胚。后文若要正式使用某个具体同胚，必须给出映射或引用已经陈述的定理。

6.6 本章结语

这一章，你在绝对的虚空中建立了几何的至高准则。

你用距离尺子确立了度量的三条公理，以严格不等式定义开球，并进一步定义了度量空间中的开集。

随后，你暂时放下距离数值，只用开集及其并、交关系定义了拓扑空间。

你用开集严格定义了连通性，又用“每个开覆盖都有有限子覆盖”定义了紧致性。

最后，你用“开集的原像仍是开集”的规则，严格定义了连续映射。

有了完备的实数轴（第五章）和开集、连通、紧致、连续这些拓扑法则（第六章），我们的宇宙舞台终于搭建完毕了。

但是，舞台上还缺少真正的主角——那些跳跃着、旋转着、能够描绘世间万物轨迹的复杂规则。

在下一章，我们将在实数完备性的护航下，研究怎样严格处理无穷级数。我们将用代数方法构造圆周率 π ，并逐步写下描述波动的公式。

准备好迎接无穷级数与超越魔法的降临吧！

7 超越的造物 —— 填满缝隙的无限之舞

在上一章，我们在没有任何尺子的地方，用“拓扑开集”的集合魔法，严格定义了什么是“靠近（逼近）”，什么是“连在一起（连通）”，以及什么是“不撕裂的动作（连续映射）”。

结合第五章构造的完备实数轴 $\mathbb{R}$ ，我们已经有了严密描述极限与连续性的舞台。

但目前明确构造的函数还很少。第四章虽然构造了多项式环，我们却还没有借助无穷过程构造出本章要研究的指数函数与波动函数。

这一章，我们要越过“有限次运算”的边界，借助实数的完备性研究**无穷多个**数字相加的含义。

我们将看到，无穷级数可以构造出指数函数和波动函数，还能在不预先使用圆的几何定义时引出圆周率 π 。

欢迎来到无限之舞。

7.1 在无限的边缘跳舞 —— 级数与收敛

假设实数轴上排着一列数字：第 0 项、第 1 项、第 2 项……一直延续下去。

现在，我们提出一个疯狂的想法：把它们全部加起来！

令 $a : \mathbb{N} \rightarrow \mathbb{R}$ 是实数序列。对每个 n ，用递归定义有限部分和

$$s_n = \sum_{k=0}^n a_k.$$

符号 $\sum_{k=0}^{\infty} a_k$ 称为一个无穷级数，但它的含义只由部分和序列决定。

但是，“无限相加”不能照搬有限加法。如果考察 $1 + 1 + 1 + \cdots$ ，第 n 个部分和是 $s_n = n + 1$ 。若 $s_n \rightarrow S$ ，则存在 $N \in \mathbb{N}$ ，使 $n \geq N$ 时都有 $s_n < S + 1$ 。另一方面，阿基米德性质允许选取 $n \geq N$ 且 $n > S$ ，于是 $s_n = n + 1 > S + 1$ ，矛盾。所以这个级数发散。

那么，怎样判断一个级数的部分和是否会逼近某个实数呢？

我们要使用第五章的关键结论——**实数 $\mathbb{R}$ 的完备性**：只要一个实数序列满足柯西条件，实数轴上就有它的极限。

我们不直接看无穷个相加的结果。我们玩一个叫“部分和”的游戏，一步一步来：

- 第 0 步结果：只算第 0 个数字。
- 第 1 步结果：前两个数字相加。
- 第 2 步结果：前三个数字相加。
-

这样，“无限相加”的问题就转化为研究一个由有限加法逐步产生的新序列。

【定义】若部分和序列 s_n 收敛到 $S \in \mathbb{R}$ ，就称级数收敛，并定义 $\sum_{k=0}^{\infty} a_k = S$ 。若部分和不收敛，就称级数发散。

【定理（级数的柯西判别，证明略）】级数收敛，当且仅当对每个 $\epsilon > 0$ 都存在 $N \in \mathbb{N}$ ，使所有满足 $n > m \geq N$ 的 $m, n \in \mathbb{N}$ 都有

$$|a_{m+1} + \cdots + a_n| < \epsilon.$$

这是实数柯西完备性对部分和序列的直接应用。

这里有三个很容易混淆的对象： a_n 是第 n 个**单项**， s_n 是前 $n + 1$ 项组成的**部分和**，而 S 是部分和最终逼近的**级数之和**。所谓“无穷相加”，始终只是在研究有限部分和 s_n 的去向；我们没有在某一步真的完成无穷多个加法。

【命题（收敛级数的单项必须趋于零）】若 $\sum_{n=0}^{\infty} a_n$ 收敛，则 $a_n \rightarrow 0$ 。

短证明：设 $s_n \rightarrow S$ 。给定 $\epsilon > 0$ ，存在 $M \in \mathbb{N}$ ，使每个 $k \geq M$ 都满足 $|s_k - S| < \epsilon/2$ 。于是对每个 $n \geq M + 1$ ，

$$|a_n| = |s_n - s_{n-1}| \leq |s_n - S| + |s_{n-1} - S| < \epsilon.$$

所以单项必须趋于 0。这只是**必要条件**，不是充分条件：级数收敛一定能推出单项趋于 0，但单项趋于 0 不一定能反推出级数收敛。

【反例（单项趋于零仍可能发散）】调和级数

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

的单项由阿基米德性质可知趋于 0，但它发散。事实上，对 $j \geq 1$ ，把从第 $2^{j-1} + 1$ 项到第 2^j 项归为第 j 块；这一块共有 2^{j-1} 项，每项至少为 $1/2^j$ ，所以每块之和至少为 $1/2$ 。因此

$$\sum_{n=1}^{2^m} \frac{1}{n} \geq 1 + \frac{m}{2}.$$

这里称序列 (u_n) **有界**，是指存在 $M > 0$ 使每个 n 都有 $|u_n| \leq M$ ；若不存在这样的 M ，就称它无界。阿基米德性质说明上式右侧可以超过任意给定实数，所以调和级数的部分和序列无界。另一方面，收敛序列一定有界：若 $s_n \rightarrow S$ ，在收敛定义中取 $\epsilon = 1$ ，则存在 $N \in \mathbb{N}$ ，使每个 $n \geq N$ 都满足 $|s_n| < |S| + 1$ ；把 $|S| + 1$ 与前面有限项的绝对值一起取最大值，就得到整个序列的上界。故部分和不可能收敛。

【可算例子（几何级数）】令

$$s_N = 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^N}.$$

有限运算给出

$$2s_N - s_N = 2 - \frac{1}{2^N},$$

所以 $s_N = 2 - 1/2^N$ 。归纳法给出 $2^N \geq N + 1$ ： $N = 0$ 时等号成立；若对 N 成立，则

$$2^{N+1} = 2 \cdot 2^N \geq 2(N + 1) \geq N + 2.$$

给定 $\epsilon > 0$ ，阿基米德性质允许选取 N_0 使 $N_0 + 1 > 1/\epsilon$ 。于是 $N \geq N_0$ 时

$$0 < \frac{1}{2^N} \leq \frac{1}{N+1} < \epsilon.$$

因此 $1/2^N \rightarrow 0$, 于是

$$\sum_{n=0}^{\infty} \frac{1}{2^n} = 2.$$

这不是猜出“最后一项”，而是算出了每个有限部分和，再求它们的极限。

【定义】 若 $\sum_{n=0}^{\infty} |a_n|$ 收敛，就称 $\sum_{n=0}^{\infty} a_n$ **绝对收敛**。

【定理（绝对收敛与重排，证明略）】 绝对收敛蕴含收敛；对任意双射 $\sigma: \mathbb{N} \rightarrow \mathbb{N}$ ，重排级数 $\sum_{n=0}^{\infty} a_{\sigma(n)}$ 收敛到同一个和。这里的“随意打乱”严格指这样的双射重排。

【定理（交错级数与余项估计，证明略）】 设 $b: \mathbb{N} \rightarrow \mathbb{R}$ 是实数序列。若对每个 $n \in \mathbb{N}$ 都有 $b_n \geq 0$ 、 $b_{n+1} \leq b_n$ ，并且 $b_n \rightarrow 0$ ，则

$$S = \sum_{n=0}^{\infty} (-1)^n b_n$$

收敛。对部分和 $S_N = \sum_{n=0}^N (-1)^n b_n$ ，余项 $R_N = S - S_N$ 满足

$$0 \leq (-1)^{N+1} R_N \leq b_{N+1}.$$

因此 $|R_N| \leq b_{N+1}$ 。若 $R_N \neq 0$ ，它的符号与第一个被省略的项 $(-1)^{N+1} b_{N+1}$ 相同；若 $R_N = 0$ ，就不对零谈正负符号。特别地， $b_{N+1} = 0$ 时，单调性迫使以后各项全为 0，因而必有 $R_N = 0$ ；即使 $b_{N+1} > 0$ ，后续项也可能恰好抵消，所以不能仅凭 $b_{N+1} > 0$ 断言余项非零。

【思考训练】

考察无穷级数 $1 - 1 + 1 - 1 + \dots$ 。

它的前几个部分和分别是什么？这个级数收敛吗？

解答：

第 0 步：1；

第 1 步： $1 - 1 = 0$ ；

第 2 步: $1 - 1 + 1 = 1$;

第 3 步: $1 - 1 + 1 - 1 = 0$ 。

部分和始终在 1 与 0 之间交替, 不能趋于同一个实数, 所以级数发散。

7.2 宇宙第一召唤阵 —— 指数魔法 $\exp(x)$

有了绝对收敛这一工具, 我们可以开始构造重要的非线性函数。

【定义】对 $n \in \mathbb{N}$, 阶乘由递归给出: $0! = 1$, $(n+1)! = (n+1)n!$ 。对 $x \in \mathbb{R}$, 幂由递归定义为 $x^0 = 1$, $x^{n+1} = x^n x$ 。归纳法给出 $n! \geq 1$, 所以后文以 $n!$ 为分母的式子都有定义。

【定义】对任意实数 x , 定义

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

【定义】设 $E \subseteq \mathbb{R}$, (f_n) 是函数序列, 每个 $f_n : E \rightarrow \mathbb{R}$, 并设 $f : E \rightarrow \mathbb{R}$ 。若对每个 $\epsilon > 0$, 都存在一个与 x 无关的 $N \in \mathbb{N}$, 使所有 $n \geq N$ 和 $x \in E$ 都有 $|f_n(x) - f(x)| < \epsilon$, 就称 f_n 在 E 上一致收敛到 f 。

【引理 (连续函数的有限代数运算, 证明略)】设 $E \subseteq \mathbb{R}$ 。在 E 的子空间拓扑下, 常值函数与函数 $x \mapsto x$ 连续; 若 $f, g : E \rightarrow \mathbb{R}$ 连续, 则 $f + g$ 、 $f - g$ 与 fg 都连续。因此, 每个有限多项式表达式所定义的函数都连续。

【定理 (一致极限定理, 证明略)】设 $E \subseteq \mathbb{R}$ 。若每个 $f_n : E \rightarrow \mathbb{R}$ 都连续, 并且 f_n 在 E 上一致收敛到 $f : E \rightarrow \mathbb{R}$, 则 f 连续。

【定理 (指数级数, 证明略)】对每个实数 x , 上述级数绝对收敛; 它的部分和在每个闭区间上一致收敛, 所以 $\exp$ 连续。证明使用级数柯西判别、阶乘估计、连续函数的有限代数运算和一致极限定理。

【定义】对 $n, k \in \mathbb{N}$, 当 $0 \leq k \leq n$ 时, 沿用第四章定义的自然数差 $n - k$, 并定义二项式系数

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

并约定 $k > n$ 时 $\binom{n}{k} = 0$ 。

【定理（整性与二项式定理，证明略）】这些系数其实都是自然数。对任意 $a, b \in \mathbb{R}$ 与 $n \in \mathbb{N}$,

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

而对任意 $n, k \in \mathbb{N}$, 帕斯卡恒等式是

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

它与二项式定理都可由自然数归纳法证明。

【定理（绝对收敛级数的柯西乘积，证明略）】两个绝对收敛级数相乘时，可以按总次数分组，其和等于两个级数之和的乘积。结合二项式定理可推出：对任意 $a, b \in \mathbb{R}$,

$$\exp(a) \exp(b) = \exp(a + b).$$

【推论（指数函数的正值与倒数公式）】对每个实数 x , $\exp(x) > 0$, 并且

$$\exp(-x) = \frac{1}{\exp(x)}.$$

先由级数定义得 $\exp(0) = 1$ 。再把 $a = x, b = -x$ 代入加法公式，得到

$$\exp(x) \exp(-x) = 1,$$

所以两因子都不为 0。最后

$$\exp(x) = \exp(x/2)^2 > 0,$$

因为非零实数的平方为正。

【思考训练】只使用加法公式和 $e = \exp(1)$, 求 $\exp(2)$ 与 $\exp(-1)$ 。

提示与简解：

$$\exp(2) = \exp(1 + 1) = e^2, \quad \exp(-1) = \frac{1}{e}.$$

这个无穷级数把输入的加法转化成了函数值的乘法。取 $x = 1$ ，便得到重要常数自然底数 e ：

$$e = \exp(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots.$$

【定理（ e 的数值界，证明略）】用部分和计算前若干项，再用阶乘估计控制所有未计算项，可以证明

$$2.718 < e < 2.719.$$

7.3 用代数撕裂波纹——召唤波浪的灵魂

传统上，正弦和余弦常由圆与角度引入。

这里我们不预先使用圆规、量角器或角度测量，而是直接从级数出发。

仿照指数级数，我们把偶次幂与奇次幂分别排列，并赋予交替的正负号。

【定义】对每个 $x \in \mathbb{R}$ ，

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

【命题（奇偶性，短证明）】对每个实数 x ，

$$\sin(-x) = -\sin x, \quad \cos(-x) = \cos x.$$

因为 $(-x)^{2n} = x^{2n}$ ，而 $(-x)^{2n+1} = -x^{2n+1}$ ；代入定义级数并逐项化简即可得到结论。

【定理（三角级数，证明略）】这两个级数对每个实数绝对收敛，并在每个闭区间上一致收敛，所以 $\sin$ 与 $\cos$ 连续。对任意 $a, b \in \mathbb{R}$ ，利用柯西乘积可证明加法公式

$$\sin(a+b) = \sin a \cos b + \cos a \sin b,$$

$$\cos(a+b) = \cos a \cos b - \sin a \sin b.$$

令 $b = -a$ ，结合级数的奇偶性，可推出

$$\sin^2 x + \cos^2 x = 1$$

(对任意实数 x 成立)。

【可算例子（带保证的近似）】取 $x = 1$ 。各项绝对值递减，因此交错级数余项估计给出

$$\frac{5}{6} \leq \sin 1 \leq \frac{5}{6} + \frac{1}{120} = \frac{101}{120},$$

$$\frac{1}{2} \leq \cos 1 \leq \frac{1}{2} + \frac{1}{24} = \frac{13}{24}.$$

这两行不依赖预先画出的波形：只计算前两项，就同时得到近似值和严格的误差界。

这里没有使用三角形或角度；三角函数的基本规律已经蕴含在这些级数之中。

7.4 捕获幽灵——圆周率 π 的纯代数身份证

下面结合上一章的拓扑与本章的级数，在不预先借助圆的几何定义时构造 π 。

先考察刚刚定义的余弦函数 $\cos x$ 。

1. **起点**：由定义 $\cos 0 = 1$ 。
2. **终点**：从第三项起， $\cos 2$ 的交错项绝对值严格递减。交错级数余项估计给出

$$\cos 2 = -1 + \left(\frac{2}{3} - \frac{4}{45} + \cdots \right) < -1 + \frac{2}{3} < 0.$$

现在, $\cos$ 在区间的两个端点分别取正值和负值。接下来使用第六章的**连通性与连续映射**。

- 【定理（区间连通性，证明略）】实数的每个区间都是连通空间；证明使用确界定理。
- 上一节已经由闭区间上一致收敛证明了 $\cos$ 连续。仅有逐点绝对收敛本身并不足以推出连续。

【定理（介值定理，证明略）】设 $a, b, y \in \mathbb{R}$ 且 $a \leq b$ 。若 $f : [a, b] \rightarrow \mathbb{R}$ 连续, 且 $f(a) \leq y \leq f(b)$ 或 $f(b) \leq y \leq f(a)$, 则存在 $c \in [a, b]$ 使 $f(c) = y$ 。它由“连续映射保持连通”和“实数连通子集是区间”推出。

因此在 0 与 2 之间至少有一个 α 满足 $\cos \alpha = 0$ 。

直观地说, 从正值连续变化到负值, 途中必定经过零值。

即使还不知道 $\cos x$ 的完整图像, 连续性也已经迫使它在区间内经过零点。

【引理（第一个余弦零点，证明略）】利用级数估计、加法公式和 $\sin^2 x + \cos^2 x = 1$, 可以证明 $\cos x$ 在 $[0, 2]$ 只有一个零点 α , 并且 $\sin \alpha = 1$ 。

由此作出定义：

【定义】 $\pi = 2\alpha$ 。

【推论】由加法公式, $\cos \pi = \cos(2\alpha) = -1$, $\sin \pi = \sin(2\alpha) = 0$ 。

按照本书的构造顺序, π 不是预先用“圆周长与直径之比”定义的, 而是由交错级数与拓扑连通性构造出来的。以后还可以证明, 这个数与几何中的圆周率一致。

7.5 大一统的和弦 —— 欧拉公式与复数 $\mathbb{C}$ 的圆满

在第五章结尾, 我们构造了满足 $i^2 = -1$ 的虚数单位 i , 并建立了复数域 $\mathbb{C}$ 。

【定义】对 $z = a + bi \in \mathbb{C}$, 定义模 $|z| = \sqrt{a^2 + b^2}$ 。

【定理（复数模，证明略）】对任意 $z, w \in \mathbb{C}$, 都有 $|z| \geq 0$, 且 $|z| = 0$ 当且仅当 $z = 0$; 还有 $|-z| = |z|$ 、 $|zw| = |z||w|$ 与 $|z + w| \leq |z| + |w|$ 。因此 $d(z, w) = |z - w|$ 是 $\mathbb{C}$ 上的度量。

【定义】设 $z : \mathbb{N} \rightarrow \mathbb{C}$ 是复数序列, 记 $z_n = z(n)$ 。给定 $w \in \mathbb{C}$, 若对每个 $\epsilon > 0$, 都存在 $N \in \mathbb{N}$, 使每个 $n \geq N$ 都满足 $|z_n - w| < \epsilon$, 就称 z_n 收敛到 w 。若对每个

$\epsilon > 0$, 都存在 $N \in \mathbb{N}$, 使任意 $m, n \geq N$ 都满足 $|z_m - z_n| < \epsilon$, 就称 (z_n) 为复数柯西列。

【定义】复级数通过部分和定义。若实级数 $\sum_{n=0}^{\infty} |z_n|$ 收敛, 就称 $\sum_{n=0}^{\infty} z_n$ 绝对收敛。

字母 i 并没有遮住第五章的构造: $a + bi$ 只是实数有序对 (a, b) 的常用写法。运算规则是

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

例如

$$(2 + i)(1 - 3i) = 2 - 6i + i - 3i^2 = 5 - 5i,$$

而 $|3 + 4i| = \sqrt{3^2 + 4^2} = 5$ 。复数运算因此仍然归结为实数的四则运算与平方根。

【思考训练】计算 $(1 + i)^2$ 与 $|1 + i|$, 再用复数模定理检查 $|(1 + i)^2| = |1 + i|^2$ 。

简解:

$$(1 + i)^2 = 1 + 2i + i^2 = 2i, \quad |1 + i| = \sqrt{2}.$$

所以左侧 $|2i| = 2$, 右侧 $(\sqrt{2})^2 = 2$ 。

【定理（复数完备性, 证明略）】复数序列 (z_n) 是柯西列, 当且仅当其实部序列与虚部序列都是实数柯西列。因此, 每个复数柯西列都收敛。

【定义】对 $z \in \mathbb{C}$ 与 $n \in \mathbb{N}$, 自然数次幂沿用递归 $z^0 = 1$ 、 $z^{n+1} = z^n z$ 。

【定理（复指数级数, 证明略）】对每个 $z \in \mathbb{C}$, 级数

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

绝对收敛；对任意 $z, w \in \mathbb{C}$ ，还满足 $\exp(z + w) = \exp z \exp w$ 。这把原先只在实数上定义的指数合法地扩展到了复数。

【记号】因为 $e = \exp(1)$ ，对复数 z 约定 e^z 是 $\exp z$ 的另一种写法；这不是把尚未定义的“复数次乘方”偷偷带进来，而只是给已经构造好的复指数函数换一个名字。

【引理（绝对收敛复级数的分组，证明略）】设 (z_n) 是复数序列。若复级数 $\sum_{n=0}^{\infty} z_n$ 绝对收敛，则它的偶数项子级数与奇数项子级数都绝对收敛，而且

$$\sum_{n=0}^{\infty} z_n = \sum_{k=0}^{\infty} z_{2k} + \sum_{k=0}^{\infty} z_{2k+1}.$$

更一般的双射重排也不改变和。这个结论由复数三角不等式、柯西判别以及上面的复数完备性推出。

现在取 $x \in \mathbb{R}$ ，把 ix 代入复指数级数。由

$$i^2 = -1, \quad i^3 = -i, \quad i^4 = 1,$$

i 的幂每四次循环一次。

不过，无穷级数不能仅凭有限和的经验随意拆开。这里 $\exp(ix)$ 的级数绝对收敛，所以上面的分组引理保证：把偶数项和奇数项分别收集不会改变总和。

$$\begin{aligned} \exp(ix) &= \sum_{k=0}^{\infty} \frac{(ix)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(ix)^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

第一部分是

$1 - x^2/2! + \dots$ ，第二部分在提出因子 i 后是 $x - x^3/3! + \dots$ ；它们恰好分别是前面定义的 $\cos x$ 与 $\sin x$ 。因此，对每个 $x \in \mathbb{R}$ ，

$$\exp(ix) = \cos x + i \sin x.$$

还有一个立刻算得出的特殊值。因为 $\pi = 2\alpha$ ，所以 $\alpha = \pi/2$ ；前面已经证明 $\cos \alpha = 0$ 、 $\sin \alpha = 1$ ，从而

$$\exp\left(i\frac{\pi}{2}\right) = i.$$

利用复指数的加法公式，把这个等式两边平方，就得到 $\exp(i\pi) = i^2 = -1$ 。

也可以直接把上一节定义的 π 代入欧拉公式；前面的推论已经给出 $\cos \pi = -1$ 、 $\sin \pi = 0$ ，于是

$$\exp(i\pi) + 1 = 0.$$

按照刚才的记号，这也可写成更常见的形式

$$e^{i\pi} + 1 = 0.$$

加法单位元 0、乘法单位元 1、虚数单位 i 、圆周率 π 与自然底数 $e = \exp(1)$ ，由此出现在同一个简洁的公式中。

7.6 本章结语

这一章，你在无限的边缘完成了最华丽的舞蹈。

你借助实数完备性定义了无穷级数，并用级数构造出指数函数 $\exp$ 与三角函数 $\sin$ 、 $\cos$ 。

随后，你没有预先使用圆的几何定义，而是结合拓扑法则构造了 π ；最后又在复数域中推出了欧拉公式。

至此，我们已经从实数序列走到级数，再由级数构造出重要函数。

但宇宙远不止是一条单薄的实数线。在下一章，我们将踏入广袤的多维宇宙。我们要把单一的数字组合成庞大的阵列，去探寻多维空间中隐藏的对称与旋转。

下一章将进入矩阵与多维几何。

8 多维的平直宇宙 —— 阵列与对偶的代数

第五章构造了完备实数轴，上一章又在实数轴上定义了指数函数和三角函数。

但这些对象主要生活在一维空间中。下面要把一个实坐标推广为多个实坐标。

这一章，我们利用第二章的笛卡尔积，把若干条实数轴组合成多维的平坦空间。这个空间中的元素不再是单个实数，而是由若干实数组成的阵列。

我们将为这些阵列定义向量空间、线性映射、对偶空间与内积，并用外代数解释换序变号和行列式。

这就是本章要建立的线性代数框架。

8.1 撑开维度的网 —— 线性空间与基底

怎么把一维的实数变成多维？

【沿用第六章的集合，并加入新结构】第六章已经定义

$$[n] = \{k \in \mathbb{N} : 1 \leq k \leq n\}, \quad \mathbb{R}^n = \{x : x \text{ 是从 } [n] \text{ 到 } \mathbb{R} \text{ 的映射}\}.$$

本章沿用完全相同的集合，不另造一个同名对象；新工作是在这些有限坐标映射上加入运算。对 $x, y \in \mathbb{R}^n$ 、 $c \in \mathbb{R}$ 和 $j \in [n]$ ，定义

$$(x + y)_j = x_j + y_j, \quad (cx)_j = cx_j.$$

零元是每个坐标都为 0 的元组, $-x$ 的第 j 个坐标是 $-x_j$ 。当 $n = 0$ 时, 这些规则在唯一的空元组上仍成立。配备这些逐坐标运算后, $\mathbb{R}^n$ 的元素称为**向量**。

第四章已经从具体运算中抽象出群、环和域。现在再为向量及其运算定义**线性空间**:

【定义】 域 F 上的线性空间 (向量空间) 是集合 V , 带有加法 $V \times V \rightarrow V$ 和标量乘法 $F \times V \rightarrow V$, 满足:

1. $(V, +)$ 是交换群, 零向量记为 0_V ;
2. $1v = v$;
3. $(ab)v = a(bv)$;
4. $a(u + v) = au + av$;
5. $(a + b)v = av + bv$ 。

这些等式对任意 $a, b \in F$ 和 $u, v \in V$ 成立。本章主要取 $F = \mathbb{R}$ 。逐坐标运算的 $\mathbb{R}^n$ 满足这些公理, 验证只使用实数域的运算律。

要有效描述整个空间, 还需要选出一组能够生成全部向量的“骨架”。

【定义】 向量 $v_1, \dots, v_k$ 的线性组合是 $a_1v_1 + \dots + a_kv_k$ 。允许 $k = 0$, 并约定空线性组合等于 0_V 。

若对任意正自然数 k 、任意互不相同的 $v_1, \dots, v_k \in B$ 以及任意标量 $a_1, \dots, a_k \in F$, 等式 $a_1v_1 + \dots + a_kv_k = 0_V$ 都必然推出 $a_1 = \dots = a_k = 0$, 就称向量族 B **线性无关**。若每个 $v \in V$ 都是 B 中有限多个向量的线性组合, 就说 B 张成 V 。因此, 空集恰好张成零向量空间, 并构成它的一组基底。

若 B 既线性无关又张成 V , 就称 B 为一组**基底**。

若 V 拥有一组有限基底, 就称 V 是**有限维空间**。

- **线性无关**: 基底向量之间没有非平凡的线性关系。
- **张成空间**: 每个向量都能写成有限个基底向量的线性组合。

【定理 (坐标唯一性与维数不变性, 证明略)】 相对于一组基底, 每个向量的线性组合表达唯一; 有限维空间的任意两组基底含有相同数目的向量。这个共同的自然数定义为 $\dim V$ 。

例如, 一组基底若含有 3 个向量, 这个空间就是 3 维的。

【定义】若有序基底是 $B = (v_1, \dots, v_n)$ ，并且

$$v = a_1 v_1 + \dots + a_n v_n,$$

就把唯一的数字阵列

$$[v]_B = (a_1, \dots, a_n)$$

称为 v 在基底 B 下的坐标。向量 v 本身没有改变；坐标只是选定基底后描述它的一组数字。

【定义与命题（标准基底，短证明）】当 $n \geq 1$ 且 $j \in [n]$ 时，定义 $e_j \in \mathbb{R}^n$ ：它的第 j 个坐标为 1，其余坐标为 0。对每个 $x = (x_1, \dots, x_n)$ ，逐坐标检查可得

$$x = \sum_{j=1}^n x_j e_j.$$

若 $\sum_{j=1}^n a_j e_j = 0$ ，比较第 j 个坐标便得 $a_j = 0$ 。所以 $e_1, \dots, e_n$ 张成 $\mathbb{R}^n$ 且线性无关，组成一组基底，称为 $\mathbb{R}^n$ 的标准基底。

【思考训练】

在二维空间里取 $e_1 = (1, 0)$ 和 $e_2 = (0, 1)$ 。向量 $v = (3, 5)$ 能否由它们线性组合得到？

解答：

可以。把 e_1 乘以 3 得到 $(3, 0)$ ，把 e_2 乘以 5 得到 $(0, 5)$ ，再相加便有

$$3e_1 + 5e_2 = (3, 0) + (0, 5) = (3, 5) = v.$$

同样的方法证明每个 (x, y) 都等于 $xe_1 + ye_2$ ；而 $ae_1 + be_2 = 0$ 会逐坐标推出 $a = b = 0$ 。因此 e_1, e_2 确实构成基底。只验证一个例子本身还不足以证明“张成全部空间”。

【可算例子（同一向量，不同坐标）】在 $\mathbb{R}^2$ 中另取

$$b_1 = (1, 1), \quad b_2 = (1, -1).$$

对任意 (x, y) , 都有

$$(x, y) = \frac{x+y}{2}b_1 + \frac{x-y}{2}b_2,$$

所以它们张成 $\mathbb{R}^2$ 。若 $ab_1 + bb_2 = 0$, 逐坐标得到 $a + b = 0$ 与 $a - b = 0$, 从而 $a = b = 0$; 所以它们也线性无关, 确实组成基底 $B = (b_1, b_2)$ 。又因为

$$(3, 1) = 2b_1 + b_2,$$

同一个向量在标准基底下的坐标是 $(3, 1)$, 在 B 下却是

$$[(3, 1)]_B = (2, 1).$$

【思考训练】 求向量 $(5, 1)$ 在基底 $B = (b_1, b_2)$ 下的坐标。

提示与简解: 设 $(5, 1) = ab_1 + bb_2 = (a + b, a - b)$, 于是

$$a + b = 5, \quad a - b = 1.$$

两式相加得 $a = 3$, 再得 $b = 2$, 所以 $[(5, 1)]_B = (3, 2)$ 。

8.2 操控宇宙的坐标阵列 —— 线性映射与矩阵

在第二章, 我们学过“映射”(把一个集合里的东西变成另一个集合里的东西的规则)。

现在考虑线性空间之间的映射, 例如从三维空间映到二维空间。保持加法与数乘的映射称为**线性映射**。

【定义】 设 V, W 是同一域 F 上的向量空间。映射 $T: V \rightarrow W$ 称为线性映射, 若对所有 $u, v \in V$ 与 $c \in F$ 都满足:

1. 尊重加法: $T(u + v) = T(u) + T(v)$ 。
2. 尊重缩放: $T(cv) = cT(v)$ 。

这两个条件保证线性组合得到保留。线性映射可以把某些方向压扁，却不会破坏“先线性组合再变换”等于“先变换再线性组合”的规则。

怎样用有限数据记录线性映射？由于线性映射保持线性组合，只要知道每个基向量的像，就能计算任意向量的像。

若 V 的基底是 $v_1, \dots, v_n$ ， W 的基底是 $w_1, \dots, w_m$ ，则每个 $T(v_j)$ 有唯一表达

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

【定义】 设 $m, n \in \mathbb{N}$ 。域 F 上的一个 $m \times n$ 矩阵，严格地说是映射

$$A : [m] \times [n] \rightarrow F.$$

把 $A(i, j)$ 简写成 a_{ij} 。上面由线性映射 T 得到的唯一系数因此确定矩阵 $A(i, j) = a_{ij}$ ，称为 T 在这两组基底下的矩阵。

【记号】 为了阅读方便，我们把这个映射的值按第一个指标分行、第二个指标分列，排成 m 行 n 列的表格。表格只是矩阵的书写格式；矩阵这个数学对象仍是刚才定义的映射。

【定义】 若向量坐标列为 $x = (x_j)$ ，矩阵向量乘法定义为

$$(Ax)_i = \sum_{j=1}^n a_{ij} x_j.$$

若 A 是 $m \times n$ 矩阵， B 是 $n \times r$ 矩阵，定义

$$(AB)_{ik} = \sum_{j=1}^n a_{ij} b_{jk}.$$

这两个公式在 $n = 0$ 时使用空和约定：没有任何加数的和等于域 F 的零元。因此矩阵定义也覆盖零维空间，不需要临时排除空坐标。

【定理（矩阵表示，证明略）】 $T(v)$ 的坐标等于 A 乘以 v 的坐标；复合映射的矩阵等于矩阵乘积。矩阵依赖所选基底，线性映射本身不依赖。

【可算例子（为什么基向量的像要按列排放）】定义线性映射

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T(x, y) = (x + 2y, 3x - y).$$

实数的分配律逐坐标验证它满足线性映射的两条定义。

标准基向量的像是

$$T(e_1) = (1, 3), \quad T(e_2) = (2, -1),$$

所以矩阵的两列正是这两个坐标列：

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}.$$

例如

$$A \begin{pmatrix} 4 \\ 5 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 5 \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 14 \\ 7 \end{pmatrix},$$

这恰好就是 $T(4, 5)$ 。矩阵乘向量，本质上是在用输入坐标给各列加权，再把结果相加。

【思考训练】某个线性映射把 e_1 送到 $(2, 0)$ ，把 e_2 送到 $(1, -1)$ 。不用先写出一般公式，直接求它把 $(3, 4)$ 送到哪里。

简解：

$$T(3, 4) = 3T(e_1) + 4T(e_2) = 3(2, 0) + 4(1, -1) = (10, -4).$$

8.3 捕捉影子的猎人 —— 对偶空间 V^*

对偶空间是第十章定义微分形式的关键工具。一般线性映射可以把向量送到另一个向量空间；其中，陪域为 $\mathbb{R}$ 的线性映射称为**线性泛函**。可以把它想成一个“评分员”：输入向量，输出实数，并且保持线性组合。

【定义】对实向量空间 V ，对偶空间是

$$V^* = \{\varphi : V \rightarrow \mathbb{R} : \varphi \text{ 是线性映射}\}.$$

在其中逐点定义 $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$ 和 $(c\varphi)(v) = c\varphi(v)$ 。

【定理（对偶空间，证明略）】这些运算满足全部向量空间公理。若 V 有有限基底 $v_1, \dots, v_n$ ，则存在对偶基底 $v^1, \dots, v^n$ ，满足 $v^i(v_j) = 1$ （当 $i = j$ ）而在 $i \neq j$ 时等于 0。

【可算例子（评分员也由基底决定）】在 $\mathbb{R}^2$ 的标准基底，定义

$$e^1(x, y) = x, \quad e^2(x, y) = y.$$

那么 $e^1(e_1) = 1, e^1(e_2) = 0$ ，而 $e^2(e_1) = 0, e^2(e_2) = 1$ ，所以它们正是对偶基底。任意线性泛函 φ 都满足

$$\varphi(x, y) = \varphi(xe_1 + ye_2) = x\varphi(e_1) + y\varphi(e_2).$$

因此只知道评分员给两个基向量打多少分，就知道它给所有向量打多少分。例如 $\varphi(e_1) = 2, \varphi(e_2) = -1$ 时，

$$\varphi(x, y) = 2x - y.$$

【思考训练】对上面的 φ ，计算 $\varphi(3, 5)$ ，并直接检查 $\varphi((3, 5) + (1, -2)) = \varphi(3, 5) + \varphi(1, -2)$ 。

简解： $\varphi(3, 5) = 1$ ；左侧为 $\varphi(4, 3) = 5$ ，右侧为 $1 + 4 = 5$ 。

因此， V^* 的元素不是原空间中的向量，而是作用于向量并返回实数的线性规则。把 $\varphi \in V^*$ 作用于 $v \in V$ ，便得到实数 $\varphi(v)$ 。

8.4 定义长度与垂直 —— 内积与几何法则

我们的多维空间虽然平直，但目前它只有加法和缩放。它没有角度，没有长短。怎样只用代数规则定义长度与垂直？为此引入**内积**； $\mathbb{R}^n$ 上的标准内积也常称为点

积。

【定义】实向量空间 V 上的内积是映射 $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ ，满足：

1. **双线性**： $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$ ，并且对第二个位置也成立；
2. **对称**： $\langle u, v \rangle = \langle v, u \rangle$ ；
3. **正定**： $\langle v, v \rangle \geq 0$ ，且等号成立当且仅当 $v = 0_V$ 。

【定义】在 $\mathbb{R}^n$ 上，**标准内积**定义为

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j \quad (x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)).$$

当 $n = 0$ 时，右边是空和，约定其值为 0。它把两个向量阵列对应位置上的数字两两相乘，再把所有乘积相加。实数运算律保证它满足刚才的三条内积公理。

有了内积，就可以定义长度与垂直：

1. **长度**： $\|v\| = \sqrt{\langle v, v \rangle}$ 。第五章的确界性质保证每个非负实数有唯一非负平方根。
2. **垂直**： $u \perp v$ 当且仅当 $\langle u, v \rangle = 0$ ；零向量也允许与所有向量垂直。

【定义】实向量空间 V 上的**范数**是映射 $N : V \rightarrow \mathbb{R}$ ，通常写作 $N(v) = \|v\|$ ，并满足： $\|v\| \geq 0$ ，且等号成立当且仅当 $v = 0_V$ ； $\|cv\| = |c| \|v\|$ ； $\|u + v\| \leq \|u\| + \|v\|$ 。上面由内积给出的长度函数称为该内积**诱导的范数**； $\mathbb{R}^n$ 的标准内积诱导的范数简称**标准范数**。

【定理（柯西—施瓦茨与范数，证明略）】 $|\langle u, v \rangle| \leq \|u\| \|v\|$ 。内积的双线性还给出 $\|cv\| = |c| \|v\|$ ，柯西—施瓦茨不等式则推出 $\|u + v\| \leq \|u\| + \|v\|$ ；因此上面的长度函数确实是范数，并且 $d(u, v) = \|u - v\|$ 是度量。

【命题（兑现第六章的距离比较，短证明）】在 $\mathbb{R}^n$ 上，把内积范数给出的距离记为

$$d_2(x, y) = \|x - y\| = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}.$$

当 $n \geq 1$ 时，对第六章定义的最大距离 d_∞ 有

$$d_\infty(x, y) \leq d_2(x, y) \leq \sqrt{2} d_\infty(x, y).$$

第一边是因为每个 $|x_j - y_j|$ 都不超过所有坐标平方和的平方根。再写

$$x - y = \sum_{j=1}^n (x_j - y_j) e_j,$$

由范数的三角不等式与 $\|ce_j\| = |c|$, 得到

$$d_2(x, y) \leq \sum_{j=1}^n |x_j - y_j| \leq n d_\infty(x, y).$$

因此, 一个半径为 r 的 d_2 开球包含一个半径为 r/n 的 d_∞ 开球, 并被半径为 r 的 d_∞ 开球包含。每一种距离下的开球都能在另一种距离下找到更小的开球, 所以两种距离诱导相同拓扑; 这正是第六章已经命名的 $\mathbb{R}^n$ 的通常拓扑。 $n = 0$ 时空间只有一个点, 结论直接成立。

【命题（垂直时的勾股等式，短证明）】若 $u \perp v$, 则

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

事实上, 由双线性、对称性和 $\langle u, v \rangle = 0$,

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

所以熟悉的勾股关系不是额外画图得到的, 而是内积公理的一个短推论。

【定义】 $K \subseteq \mathbb{R}^n$ 有界, 若存在 $R > 0$ 使每个 $x \in K$ 都有 $\|x\| \leq R$ 。

【定理（有限维 Heine–Borel 定理，证明略）】 $K \subseteq \mathbb{R}^n$ 紧致, 当且仅当它既闭又有界。证明从第六章的闭区间紧致性出发, 使用有限笛卡尔积和闭子集保持紧致。

由此, 长度与垂直都被转化成了实数的乘法与加法。

【思考训练】

有两个二维向量 $A = (3, 4)$ 与 $B = (-4, 3)$ 。请用标准内积计算 $\langle A, B \rangle$, 并判断它们是否垂直。

解答：

两个坐标位置分别贡献 $3 \cdot (-4) = -12$ 与 $4 \cdot 3 = 12$ ，所以

$$\langle A, B \rangle = -12 + 12 = 0.$$

因此 $A \perp B$ 。

8.5 换座位就变号 —— 外代数与反对称性

内积解决了长度与垂直。要进一步记录若干向量的次序、线性相关性，并为后面的有向面积与体积做准备，还需要外代数。此时尚未定义面积和体积，因此下面先建立纯代数结构。

【构造定理兼定义（外代数，证明略）】对每个实向量空间 V ，可以构造一个实向量空间 $\Lambda(V)$ ，并用一个单射线性映射把 V 识别为它的子空间；还可以同时构造子空间 $\Lambda^k(V)$ ($k \in \mathbb{N}$)、元素 $1 \in \Lambda^0(V)$ 和满足结合律、双线性的乘法 $\wedge$ ，使下列性质同时成立：

1. $\Lambda(V) = \bigoplus_{k=0}^{\infty} \Lambda^k(V)$ 。这里的含义是：每个元素都能唯一写成有限个不同阶成分的和。
2. $\Lambda^0(V) = \mathbb{R} \cdot 1$ ，并借助 $a \leftrightarrow a1$ 把它与 $\mathbb{R}$ 视为同一空间；在这个识别下可简写为 $\Lambda^0(V) = \mathbb{R}$ 。同时， $\Lambda^1(V) = V$ 。
3. 对任意 $p, q \in \mathbb{N}$ ，楔积限制为双线性映射

$$\Lambda^p(V) \times \Lambda^q(V) \rightarrow \Lambda^{p+q}(V).$$

4. $1 \wedge \omega = \omega = \omega \wedge 1$ 。因此零个向量的空楔积约定为 1。
5. 对每个 $v \in V$ ，有 $v \wedge v = 0$ ； $\Lambda^k(V)$ 由所有 k 重楔积 $v_1 \wedge \cdots \wedge v_k$ 的线性组合组成。

若 $e_1, \dots, e_n$ 是有限维空间 V 的基底，那么 $\{1\}$ 是 $\Lambda^0(V)$ 的基底；对 $1 \leq k \leq n$,

$$e_{i_1} \wedge \cdots \wedge e_{i_k} \quad (1 \leq i_1 < \cdots < i_k \leq n)$$

构成 $\Lambda^k(V)$ 的基底。当 $k > n$ 时， $\Lambda^k(V) = \{0\}$ 。这个对象可以只用有限向量序

列、线性组合和商集构造；证明虽长，但不需要增加新公理。后文只使用这里明确列出的性质。

由双线性 $(u+v) \wedge (u+v) = 0$ 得

$$u \wedge v = -v \wedge u.$$

所以交换两个座位会变号。更一般地，只要 $v_1, \dots, v_k$ 线性相关，就有 $v_1 \wedge \dots \wedge v_k = 0$ ；仅仅“写成几个不同向量”还不足以产生非零楔积。

【可算例子（二维楔积）】在 $\mathbb{R}^2$ 中写

$$u = ae_1 + be_2, \quad v = ce_1 + de_2.$$

只用双线性、 $e_i \wedge e_i = 0$ 和 $e_2 \wedge e_1 = -e_1 \wedge e_2$ ，便有

$$\begin{aligned} u \wedge v &= ac e_1 \wedge e_1 + ad e_1 \wedge e_2 + bc e_2 \wedge e_1 + bd e_2 \wedge e_2 \\ &= (ad - bc)e_1 \wedge e_2. \end{aligned}$$

例如 $u = (2, 1), v = (1, 3)$ 时，

$$u \wedge v = (2 \cdot 3 - 1 \cdot 1)e_1 \wedge e_2 = 5e_1 \wedge e_2.$$

这里的楔积不是另一个普通二维向量，而是 $\Lambda^2(\mathbb{R}^2)$ 中的二阶对象；数字 5 是它相对于基底 $e_1 \wedge e_2$ 的坐标。

【思考训练】展开计算

$$(e_1 + e_2) \wedge (2e_1 - e_2).$$

简解：含 $e_1 \wedge e_1$ 或 $e_2 \wedge e_2$ 的项为 0，其余两项之和为

$$-e_1 \wedge e_2 + 2e_2 \wedge e_1 = -3e_1 \wedge e_2.$$

【定理（顶次外幂与行列式，证明略）】若 $\dim V = n \geq 1$ ，则 $\Lambda^n(V)$ 是一维空间。对每个线性映射 $T: V \rightarrow V$ ，存在唯一标量 $\det T \in \mathbb{R}$ ，使得对所有 $v_1, \dots, v_n \in V$ 都有

$$T(v_1) \wedge \cdots \wedge T(v_n) = (\det T)(v_1 \wedge \cdots \wedge v_n).$$

这个标量称为行列式。它满足

$$\det(\text{id}_V) = 1, \quad \det(S \circ T) = (\det S)(\det T),$$

其中第二个等式对任意线性映射 $S, T : V \rightarrow V$ 成立；并且 $\det T \neq 0$ 当且仅当 T 可逆。一个 $n \times n$ 方阵 A 的行列式，定义为标准基底下线性映射 $x \mapsto Ax$ 的行列式。

刚才的二维计算立刻解释了熟悉的二阶公式：若

$$T(e_1) = ae_1 + be_2, \quad T(e_2) = ce_1 + de_2,$$

则

$$T(e_1) \wedge T(e_2) = (ad - bc)e_1 \wedge e_2,$$

所以矩阵

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

的行列式就是 $ad - bc$ 。这个公式正是两个像向量的楔积系数。

【直觉】 在通常的欧氏坐标中，一旦以后严格定义有向体积， $|\det T|$ 会成为体积的放大倍数，而 $\det T$ 的正负会记录方向是否颠倒。本章尚未定义体积与方向，因此这句话只作直觉说明，不参与后面的证明。

8.6 本章结语

这一章，我们为多维平坦空间建立了线性代数骨架。

我们用笛卡尔积构造出由数字阵列组成的向量空间。

我们用基底和矩阵，把线性变换编码成由坐标对索引的有限映射，并可用行列表格方便地书写。

我们又引入对偶空间，使线性泛函能够把向量映成实数；内积则给出了长度与垂直。最后，外代数的反对称法则导出了行列式，为以后严格解释多维方向与体积变化做好准备；本章没有把尚未定义的体积直觉当作证明。

至此，多维空间的代数骨架已经建立，但目前讨论的空间仍是平坦的。

下一章将把第六章的连续与极限同本章的线性映射结合起来，研究非线性映射在一点附近的变化。

下一步就是微积分。

9 变化的魔法 —— 驯服无穷小的野兽

前面已经建立了两套工具：第六章的拓扑语言严格描述逼近与连续，第八章的线性映射和矩阵则描述多维空间中的线性变化。

如果变化不是线性的，例如第七章定义的 $\sin x$ ，怎样描述它在某一点的瞬时变化率？

变化量不能直接取 0，否则差商会出现除以 0；变化量只要非零，差商描述的仍是两个时刻之间的平均变化率。关键是研究变化量趋于 0 时，差商是否趋于一个确定的数。

本章把极限与线性映射结合起来，建立导数、积分和多元微分的基本框架。

先从一维导数开始。

9.1 捕获瞬间的切线 —— 导数的诞生

回到一维实数轴。设函数 f 记录汽车的位置：输入时刻 x ，输出位置 $f(x)$ 。

当时刻从 x 变为 $x + h$ 时，汽车的位置变为 $f(x + h)$ 。

这段时间内的位置变化量是 $f(x + h) - f(x)$ ，平均速度是

$$\frac{f(x + h) - f(x)}{h}.$$

但这只是时间段上的平均速度。要描述时刻 x 的瞬时速度，不能把 h 直接取成 0，而要考察非零的 h 趋于 0 时差商的极限。

【定义】设 $U \subseteq \mathbb{R}$ 是开集， $f : U \rightarrow \mathbb{R}$ ，并且 $x \in U$ 。写

$$\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = L$$

表示：对每个 $\epsilon > 0$ ，存在 $\delta > 0$ ，只要 $h \in \mathbb{R}$ 、 $0 < |h| < \delta$ 且 $x + h \in U$ ，就有

$$\left| \frac{f(x+h) - f(x)}{h} - L \right| < \epsilon.$$

若这样的 L 存在，就称 f 在 x 可导，并定义 $f'(x) = L$ 。

【定理（极限唯一性，证明略）】若上述 L 存在，则它唯一；因此导数确实是一个确定的实数。可导还蕴含连续。

【可算例子（从定义求 x^2 的导数）】令 $f(x) = x^2$ 。当 $h \neq 0$ 时，

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = 2x + h.$$

我们猜导数是 $2x$ ，但还要检查量词。给定 $\epsilon > 0$ ，取 $\delta = \epsilon$ ；只要 $0 < |h| < \delta$ ，就有

$$\left| \frac{f(x+h) - f(x)}{h} - 2x \right| = |h| < \epsilon.$$

所以严格按照定义， $f'(x) = 2x$ 。这里从未把 h 取成 0；我们只是让所有足够小的非零 h 都给出足够接近 $2x$ 的差商。

【思考训练】用同一方法求 $f(x) = 3x + 2$ 的导数。

简解：对每个 $h \neq 0$ ，

$$\frac{f(x+h) - f(x)}{h} = \frac{3h}{h} = 3.$$

差商已经恒等于 3，所以 $f'(x) = 3$ 。

【定理（基本求导法则，证明略）】若 f, g 在 x 可导， $c \in \mathbb{R}$ ，则

$$(f+g)'(x) = f'(x) + g'(x), \quad (cf)'(x) = cf'(x),$$

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

若另有 $g(x) \neq 0$, 则

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

常值函数的导数为 0, 恒等函数 $x \mapsto x$ 的导数为 1; 因此对每个正自然数 n , 有 $(x^n)' = nx^{n-1}$. 这些结论都可直接把差商通分、展开后用本节的极限定义证明。

【定理（一维中值定理，证明略）】若 $A < B$, 函数 $f : [A, B] \rightarrow \mathbb{R}$ 连续, 并在 (A, B) 上可导, 则存在 $c \in (A, B)$ 使

$$f(B) - f(A) = f'(c)(B - A).$$

它说“整段的平均变化率, 等于某个中间点的瞬时变化率”。证明可由闭区间紧致性保证最大值、最小值存在, 再从导数定义证明内部极值点的导数为 0。后面的微积分基本定理和混合偏导定理都可使用它。

【定理（本书三个幂级数的逐项求导，证明略）】第七章中 $\exp$ 、 $\sin$ 、 $\cos$ 的导数级数在每个闭区间上一致收敛, 因此可以逐项求导, 并得到

$$(\exp x)' = \exp x, \quad (\sin x)' = \cos x, \quad (\cos x)' = -\sin x,$$

并且它们拥有任意阶连续导数。

导数 L 不是把 h 代成 0 后相除, 而是非零差商的极限。几何上, 它给出曲线在这一点切线的斜率。

【思考训练】

假设一辆汽车的位置函数恒为 $f(x) = 5$ 。请用差商计算它在非零时间增量 h 上的平均速度, 并求导数。

解答:

它在 $x + h$ 与 x 时的位置都是 5, 所以当 $h \neq 0$ 时,

$$\frac{f(x+h) - f(x)}{h} = \frac{5 - 5}{h} = 0.$$

按极限定义, 这个恒为 0 的差商趋于 0; 由极限唯一性, $f'(x) = 0$ 。

9.2 碎片的堆积 —— 积分与极限的总和

导数描述瞬时变化。下面研究如何用越来越细的有限分割定义累积总量。

先设 $A < B$ 。如果在区间 $[A, B]$ 上有函数 f ，我们想计算它的累积总量。若 $f \geq 0$ ，这个数可解释为图像下的面积；若 f 有负值，积分计算的是带正负号的面积，不能直接叫普通几何面积。

对一般函数，一次有限分割得到的矩形和通常只给出近似值。

这就引出黎曼积分：

1. **有限切割**：取有限分点 $A = x_0 < x_1 < \cdots < x_n = B$ ，这叫分割 P 。每个小区间选一点 $t_i \in [x_{i-1}, x_i]$ 。
2. **局部取样**：在每个小区间任选一点，以该点的函数值代表这段区间的高度。
3. **相加**：定义黎曼和

$$R(f, P, t) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1}),$$

并定义网格宽度 $\|P\| = \max_i(x_i - x_{i-1})$ 。

即使函数可积，单个黎曼和也不保证恰好等于积分值；但网格变细后，这些和可能稳定地逼近同一个实数。这里需要第五章的实数完备性。

【定义】 设 $A < B$ 且 $f : [A, B] \rightarrow \mathbb{R}$ 。若存在 $S \in \mathbb{R}$ ，使对每个 $\epsilon > 0$ 都存在 $\delta > 0$ ，并且对任意带标记的有限分割，只要 $\|P\| < \delta$ 就有 $|R(f, P, t) - S| < \epsilon$ ，则称 f 在 $[A, B]$ 上黎曼可积，并定义

$$\int_A^B f(x) dx = S.$$

当上下限相等时，约定

$$\int_A^A f(x) dx = 0.$$

当 $A > B$ 且 f 在 $[B, A]$ 上可积时，定义反向积分

$$\int_A^B f(x) dx = - \int_B^A f(x) dx.$$

因此积分记号对任意次序的两个端点都有明确含义。

此处 dx 只是标明积分变量是 x 的记号，并不表示一个“无穷小实数”；第十章会把 dx 重新解释为对偶空间中的微分形式。

【命题（常值函数的积分，短证明）】若 $f(x) = c$ ，则对任意端点 A, B ,

$$\int_A^B c \, dx = c(B - A).$$

先设 $A < B$ 。对任意带标记分割，

$$\begin{aligned} R(f, P, t) &= \sum_{i=1}^n c(x_i - x_{i-1}) \\ &= c((x_1 - x_0) + \cdots + (x_n - x_{n-1})) \\ &= c(B - A). \end{aligned}$$

当 $A < B$ 时，所有黎曼和已经精确相同，无需等待网格变小。若 $A = B$ ，等式两边都为 0；若 $A > B$ ，由反向积分约定和刚证完的正向情形，

$$\int_A^B c \, dx = -c(A - B) = c(B - A).$$

所以公式对三种端点次序都成立。

【定理（积分值唯一性，证明略）】若两个实数 S, T 都满足上面黎曼和的 ϵ - δ 条件，则 $S = T$ 。因此 $\int_A^B f(x) \, dx$ 是一个确定的实数。

【定理（闭区间一致连续性，证明略）】闭区间上的连续函数满足：对每个 $\epsilon > 0$ ，存在统一的 $\delta > 0$ ，使区间内任意 $|x - y| < \delta$ 都有 $|f(x) - f(y)| < \epsilon$ 。

【定理（连续函数可积，证明略）】每个闭区间上的连续函数都黎曼可积。证明使用上面的一致连续性控制黎曼和，再用实数完备性锁定极限。

因此，黎曼积分的定义不依赖图形直觉；它是所有足够细的带标记分割所共同逼近的实数。

9.3 导数与积分的联系 —— 微积分基本定理

现在有了两种极限过程：

- **导数**：研究有限差商在步长趋于 0 时的极限。
- **积分**：对越来越细的有限分割分别计算有限黎曼和；若所有足够细的分割都逼近同一个实数，就取这个共同极限。

两者表面上方向相反：一个从函数提取局部变化率，另一个把局部取样累积成整体数值。微积分基本定理精确说明了它们之间的联系。

【定理（微积分基本定理，证明略）】设 $A < B$, $f : [A, B] \rightarrow \mathbb{R}$ 连续，并定义

$$G(x) = \int_A^x f(t) dt.$$

由于已经约定 $\int_A^A f = 0$ ，这个定义也包含端点 $x = A$ 。函数 G 在 $[A, B]$ 上连续，在 (A, B) 上可导，并且 $G'(x) = f(x)$ 。反过来，若 F 在 $[A, B]$ 上连续、在 (A, B) 上可导，并且对每个 $x \in (A, B)$ 都有 $F'(x) = f(x)$ ，则

$$\int_A^B f(x) dx = F(B) - F(A).$$

证明使用连续函数在小区间上变化很小、黎曼和与有限差分的望远镜式相消。连续性等条件不能省略。

【可算例子（导数把积分算出来）】因为

$$\left(\frac{x^3}{3}\right)' = x^2,$$

微积分基本定理给出

$$\int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

这里的方括号是简写： $[F(x)]_A^B$ 表示 $F(B) - F(A)$ 。

【思考训练】计算

$$\int_0^2 (3x^2 + 2) dx.$$

提示与简解: $F(x) = x^3 + 2x$ 满足 $F'(x) = 3x^2 + 2$, 所以

$$\int_0^2 (3x^2 + 2) dx = F(2) - F(0) = 8 + 4 = 12.$$

当被积函数满足定理条件, 并存在满足 $F' = f$ 的函数 F (称为 f 的**原函数**) 时, 积分只由原函数在两个端点的值决定. 微积分基本定理由此把一维导数与积分联系起来.

9.4 多维的瞬间扭曲 —— 雅可比矩阵

一维导数是表示斜率的实数. 对于从 n 维空间到 m 维空间的映射, 一个实数通常不足以描述所有输入方向上的变化; 合适的一阶近似应当是第八章定义的**线性映射**, 选定基底后则由矩阵表示.

设 m, n 是正自然数, $F: U \rightarrow \mathbb{R}^m$, 其中 $U \subseteq \mathbb{R}^n$ 是开集, $P \in U$. 以下距离和长度都使用第八章由标准内积得到的范数.

【定义 (多元可微的 ϵ - δ 形式)】若存在一个线性映射 $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$, 使得对每个 $\epsilon > 0$, 都存在 $\delta > 0$, 并且对每个满足

$$h \in \mathbb{R}^n, \quad 0 < \|h\| < \delta, \quad P + h \in U$$

的 h , 都有

$$\frac{\|F(P + h) - F(P) - Lh\|}{\|h\|} < \epsilon,$$

就称 F 在 P 可微. 这也简写为

$$\lim_{h \rightarrow 0} \frac{\|F(P + h) - F(P) - Lh\|}{\|h\|} = 0.$$

这里的极限符号只是在缩写前面的量词陈述，并没有引入“无穷小向量”。若这样的线性映射存在，把它记为 $DF(P)$ 。这准确表达了“误差比移动距离缩小得更快”。

【定理（多元导数唯一性，证明略）】满足上述极限条件的线性映射至多有一个，所以记号 $DF(P)$ 没有歧义。

【命题（一维导数正是线性近似，短证明）】当 $n = m = 1$ 时，每个线性映射 $L: \mathbb{R} \rightarrow \mathbb{R}$ 都有形式

$$L(h) = \lambda h, \quad \lambda = L(1).$$

于是对 $h \neq 0$,

$$\frac{|f(x+h) - f(x) - L(h)|}{|h|} = \left| \frac{f(x+h) - f(x)}{h} - \lambda \right|.$$

因此，多元可微定义在一维时恰好就是本章开头的导数定义，而且 $Df(x)$ 是线性映射 $h \mapsto f'(x)h$ 。多元导数正是一维斜率的线性映射版本。

【定义】写 $F = (F_1, \dots, F_m)$ ，并令 e_j 是 $\mathbb{R}^n$ 的第 j 个标准基向量。若一维极限

$$\frac{\partial F_i}{\partial x_j}(P) = \lim_{t \rightarrow 0} \frac{F_i(P + te_j) - F_i(P)}{t}$$

存在，就称它为 F_i 在 P 对第 j 个坐标的偏导数；其中只考虑使 $P + te_j \in U$ 的非零 t 。

【定理（雅可比矩阵，证明略）】若 F 在 P 可微，则上述一阶偏导数全都存在。选定标准基底后， $DF(P)$ 的矩阵称为雅可比矩阵 $J_F(P)$ ，它的第 (i, j) 项等于 $\partial F_i / \partial x_j(P)$ 。

【定义】设 $r \geq 1$ 是自然数。称 F 在开集 U 上属于 C^r ，若 F 连续，并且对每个 $1 \leq k \leq r$ 、每个分量 F_i 以及每组坐标指标 $j_1, \dots, j_k \in \{1, \dots, n\}$ ，迭代偏导数

$$\frac{\partial^k F_i}{\partial x_{j_k} \cdots \partial x_{j_1}}$$

在 U 的每一点都存在且作为 $U \rightarrow \mathbb{R}$ 的函数连续。另定义 C^0 就是连续， C^∞ 表示对每个 $r \in \mathbb{N}$ 都属于 C^r 。特别地，“连续可微”与“属于 C^1 ”含义相同。

【定理 (C^2 函数的混合偏导可交换, 证明略)】若 $U \subseteq \mathbb{R}^n$ 是开集, $f : U \rightarrow \mathbb{R}$ 属于 C^2 , 则对任意 $P \in U$ 及 $1 \leq i, j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(P) = \frac{\partial^2 f}{\partial x_j \partial x_i}(P).$$

证明可在 P 附近取一个足够小的坐标长方形, 对两个坐标方向先后各使用一次一维中值定理, 再利用二阶偏导的连续性令长方形缩小。第十章证明 $d^2 = 0$ 时会引用这个定理。

【定理 (可微性判别, 证明略)】若各一阶偏导数在 P 的某邻域存在并连续, 则 F 在 P 可微, 雅可比矩阵就是上述偏导数组成的矩阵。

【定理 (链式法则, 证明略)】设 $U \subseteq \mathbb{R}^n$ 、 $V \subseteq \mathbb{R}^m$ 是开集, $F : U \rightarrow V$, $G : V \rightarrow \mathbb{R}^p$, 并且 $P \in U$ 。若 F 在 P 可微, G 在 $F(P)$ 可微, 则复合映射 $G \circ F : U \rightarrow \mathbb{R}^p$ 在 P 可微, 并且

$$D(G \circ F)(P) = DG(F(P)) \circ DF(P).$$

等式右侧依次复合 $DF(P) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 与 $DG(F(P)) : \mathbb{R}^m \rightarrow \mathbb{R}^p$; 在标准基底, 这就是一个 $p \times m$ 矩阵乘一个 $m \times n$ 矩阵。

【可算例子 (雅可比如何控制非线性误差)】定义

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad F(x, y) = (x^2 + y, xy).$$

逐个坐标求偏导, 得到连续的偏导函数; 由可微性判别, F 处处可微, 并且

$$J_F(x, y) = \begin{pmatrix} 2x & 1 \\ y & x \end{pmatrix}.$$

在 $P = (1, 2)$, 候选线性近似为

$$L(r, s) = J_F(1, 2) \begin{pmatrix} r \\ s \end{pmatrix} = (2r + s, 2r + s).$$

直接展开可得

$$F(P + (r, s)) - F(P) - L(r, s) = (r^2, rs).$$

若记 $h = (r, s)$, 则

$$\|(r^2, rs)\| = |r|\sqrt{r^2 + s^2} \leq \|h\|^2.$$

所以当 $h \neq 0$ 时, 误差与移动长度之比至多为 $\|h\|$, 它随 $h \rightarrow 0$ 而趋于 0。这不只算出了雅可比矩阵, 也直接验证了它确实满足多元可微定义。

矩阵的第 j 列记录只沿第 j 个标准基方向移动时的一阶响应, 整个矩阵再通过线性组合处理任意小位移。偏导数分别考察各坐标方向, 可微性则保证这些信息确实能合成同一个线性近似。

因此, 在可微点附近, 函数能被雅可比矩阵所表示的线性映射作一阶近似; 近似不是完全相等, 也不是每个函数在每一点都自动可微。

9.5 跨越维度的通行证 —— 隐函数定理

隐函数定理是下一章构造流形实例的重要工具。

先看非线性方程 $xy^2 + zx = 5$ 。令输入变量块 $u = (x, z) \in \mathbb{R}^2$, 把坐标按 $(u, y) = ((x, z), y)$ 的次序重排, 并写成

$$H(u, y) = xy^2 + zx - 5 = 0.$$

前面的基本求导法则给出 $\partial H / \partial y = 2xy$ 。若要用下面的隐函数定理保证在某个解 $((x_0, z_0), y_0)$ 附近能把 y 写成 x, z 的函数, 就需要检查 $2x_0y_0 \neq 0$ 。

一般的非线性方程未必能方便地显式解出某一组变量。隐函数定理改用雅可比矩阵的可逆性, 判断局部能否把待解变量表示成其他变量的函数。下面把变量一般地分成 $x \in \mathbb{R}^n$ 与待解的 $y \in \mathbb{R}^m$ 。

【约定】 把一个 $(n + m)$ 元组的前 n 个坐标与后 m 个坐标分开, 会给出 $\mathbb{R}^{n+m}$ 与 $\mathbb{R}^n \times \mathbb{R}^m$ 之间的标准双射。以下借助这个双射识别两者, 并把相应的点写成 (a, b) 。

【定义】若 F 在 (a, b) 可微，就定义两个线性映射

$$D_x F(a, b)(u) = DF(a, b)(u, 0), \quad D_y F(a, b)(v) = DF(a, b)(0, v).$$

它们分别只记录输入变量 x 与待解变量 y 的一阶变化。

【定理(隐函数定理, 证明略)】设 n, m 是正自然数, $U \subseteq \mathbb{R}^{n+m}$ 是开集, $F : U \rightarrow \mathbb{R}^m$ 属于 C^1 , $(a, b) \in U$ 且 $F(a, b) = 0$ 。若线性映射

$$D_y F(a, b) : \mathbb{R}^m \rightarrow \mathbb{R}^m$$

可逆——等价地, $\det D_y F(a, b) \neq 0$ ——那么存在开邻域

$$a \in U_a \subseteq \mathbb{R}^n, \quad b \in U_b \subseteq \mathbb{R}^m, \quad U_a \times U_b \subseteq U,$$

以及唯一的 C^1 映射 $g : U_a \rightarrow U_b$, 满足 $g(a) = b$, 并且对每个 $(x, y) \in U_a \times U_b$ 都有

$$F(x, y) = 0 \iff y = g(x).$$

这些邻域还可选得使每个 $x \in U_a$ 对应的 $D_y F(x, g(x))$ 都可逆, 并有

$$Dg(x) = -[D_y F(x, g(x))]^{-1} \circ D_x F(x, g(x)).$$

这里检查的是关于**待解变量**的方阵, 而不是任意一个可能为长方形的完整雅可比矩阵。

这个定理不一定给出可写成初等公式的答案, 但保证局部解映射存在、唯一且属于 C^1 。

【可算例子 (只在选定的分支附近解开)】令

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad F(x, y) = x + y^2 - 1.$$

基本求导法则说明这个多项式函数属于 C^∞ 。在点 $(a, b) = (0, 1)$, 有 $F(0, 1) = 0$, 并且

$$D_y F(0, 1)(v) = 2v, \quad D_x F(0, 1)(u) = u.$$

$D_y F(0, 1)$ 可逆，所以隐函数定理保证：在 $(0, 1)$ 附近，方程 $x + y^2 - 1 = 0$ 的解恰好写成唯一的 $y = g(x)$ 。导数公式还立刻给出

$$g'(0) = -\frac{1}{2}.$$

同一方程还有靠近 $(0, -1)$ 的另一支解，但只要把 U_b 选在 1 附近，它就不在当前邻域内；这里的“唯一”始终是指定邻域内的唯一。

【思考训练】 同一个方程在解 $(1, 0)$ 处为什么不能直接使用隐函数定理理解出 y ？这里实际发生了什么？

提示与简解： $D_y F(1, 0)$ 是零映射，不可逆。对略小于 1 的 x ，有两个靠近 0 的解

$$y = \sqrt{1-x} \quad \text{与} \quad y = -\sqrt{1-x},$$

而对略大于 1 的 x 没有实数解，所以这里确实不能在 $x = 1$ 附近得到一个对每个 x 都存在且唯一的解函数。一般而言，条件失败只表示定理不再保证结论；这个例子额外展示了结论确实失败。

因此，只要关于待解变量的线性映射 $D_y F$ 可逆，零点集合在足够小的邻域内就能写成一张函数图像。

9.6 本章结语

这一章先用差商极限定义导数，再用有限分割和实数完备性定义黎曼积分；微积分基本定理说明了两者的联系。进入多维后，雅可比矩阵给出一阶线性近似，隐函数定理则把雅可比的可逆性转化为局部解函数的存在与唯一性。

下一章将不再要求整个空间拥有一套全局平直坐标，而只要求每个点附近都能使用欧氏坐标。这将引出**微分流形**，并最终导向统一内部积分与边界积分的斯托克斯定理。

10 弯曲的宇宙与边界 —— 斯托克斯的终极和弦

前九章依次建立了逻辑与集合的语言、数系与代数结构、完备实数、拓扑空间、线性代数和微积分。上一章又用线性映射描述了非线性映射的局部一阶变化，并给出了隐函数定理。

但是，我们至今进行微积分计算时使用的模型空间都是 $\mathbb{R}^n$ 这样的平坦空间。第六章允许抽象拓扑空间存在，却还没有给它们光滑坐标和微分结构。

许多空间却没有一套覆盖全体的欧氏坐标。例如，圆周和球面需要多张局部地图才能完整描述。

本章不再要求全局平坦，而是用相容的局部坐标拼接出“局部欧氏”的光滑流形。

随后把第八章的对偶空间与外代数应用到每个切空间，定义微分形式、定向与积分，最后得到斯托克斯定理。

10.1 缝合多维的地图 —— 微分流形

怎样用集合与映射严格描述圆周、球面等空间？核心做法是使用多张局部坐标图：每张图把空间的一部分对应到欧氏空间的开集，图与图之间再用坐标转换衔接。

【定义】设 $n \in \mathbb{N}$ 。一个 n 维拓扑流形是拓扑空间 M ，满足：

1. **Hausdorff**：任意两个不同点都分别含在两个不相交的开集中；
2. **第二可数**：存在至多可数的开集基，使每个开集都能写成这些基开集的并；
3. **局部欧几里得**：每个 $p \in M$ 都有开邻域 U 和同胚 $\varphi : U \rightarrow \varphi(U)$ ，其中 $\varphi(U)$ 是 $\mathbb{R}^n$ 的开子集。

【零维微分约定】 $\mathbb{R}^0$ 是只含空元组的单点空间。若 $O \subseteq \mathbb{R}^r$ 、 $O' \subseteq \mathbb{R}^s$ 是开集，并且 $r = 0$ 或 $s = 0$ ，则映射 $H : O \rightarrow O'$ 自动视为 C^∞ ；在每个有定义的点， DH 是唯一的线性映射 $\mathbb{R}^r \rightarrow \mathbb{R}^s$ 。高阶导数沿用同样的空指标约定。

二元组 (U, φ) 称为坐标图；覆盖 M 的一族坐标图称为图册。“局部拍扁”严格指同胚到 $\mathbb{R}^n$ 的开子集，不是整个 $\mathbb{R}^n$ 。

因此，在每个点附近都可以使用 n 个实数作局部坐标。

当两张坐标图 (U, φ) 、 (V, ψ) 重叠时，定义坐标转换

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \rightarrow \psi(U \cap V).$$

坐标转换把第一套坐标改写为第二套坐标；它的定义域和陪域都是欧氏开集，因此可以使用上一章的微积分。

【定义】若所有坐标转换都具有任意阶连续偏导数，就称图册光滑；这里的 C^∞ 使用第九章已经给出的递归定义。与该图册光滑相容的所有坐标图共同确定一个光滑结构，带此结构的 M 称为光滑流形。

【例子：平坦空间本身】

任取开集 $O \subseteq \mathbb{R}^n$ 。只用一张坐标图

$$(O, \text{id}_O)$$

就能把 O 变成 n 维光滑流形。第六章的子空间继承定理保证 O 是 Hausdorff 且第二可数；恒等图把它映到 $\mathbb{R}^n$ 的开子集，而唯一的坐标转换是恒等映射。

因此，流形不必在外观上弯曲；定义真正要求的是每个点附近都能使用 n 个实数作光滑坐标。

【说明：点不是坐标】

同一个点 P 在两张地图上可能分别写成 $x = \varphi(P)$ 与 $y = \psi(P)$ 。 P 是 M 的元素， x, y 只是描述它的两串实数；公式 $y = (\psi \circ \varphi^{-1})(x)$ 才说明两种描述怎样互译。以后所有与坐标无关的对象，都必须在这种互译下保持一致。

【定义】映射 $F : M \rightarrow N$ 称为光滑，若对每个 $P \in M$ ，任取含 P 的坐标图 (U, φ) 和含 $F(P)$ 的坐标图 (V, ψ) ，都存在 M 中的开邻域 W ，使

$$P \in W \subseteq U \cap F^{-1}(V),$$

并且局部坐标表达

$$\psi \circ F \circ \varphi^{-1}|_{\varphi(W)} : \varphi(W) \rightarrow \psi(V)$$

属于 C^∞ 。这里 $\varphi(W)$ 是欧氏开集，而 $W \subseteq F^{-1}(V)$ 保证复合式有定义，因此不需要假设整个 $\varphi(U \cap F^{-1}(V))$ 是开集。坐标转换与链式法则还说明：对每个 P 只找一对满足上述条件的坐标图，便可推出任意坐标图下的同一结论。

【说明】若方程满足第九章隐函数定理的雅可比可逆条件，其局部解集可以由该定理证明是流形；但隐函数定理不是流形定义本身所必需的。

这就是流形的本质：整体可能复杂，但每个点都有与欧几里得开集同胚的邻域；“等价于切平面”只是直觉，正式等价发生在坐标图之间。

【定义】圆周是 $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ，取自 $\mathbb{R}^2$ 的子空间拓扑。

【定理（圆周图册）】 S^1 是紧致的一维光滑流形；删去北极点或南极点后，各自都可用一张光滑坐标图映到 $\mathbb{R}$ 。因此两个这样的坐标图覆盖 S^1 。

证明：记北极点 $N = (0, 1)$ 、南极点 $S = (0, -1)$ ，并令

$$U_N = S^1 \setminus \{N\}, \quad U_S = S^1 \setminus \{S\}.$$

这两个集合在 S^1 中开。以 U_N 为例：若 $P \neq N$ ，则

$$B_{\|P-N\|/2}^{\mathbb{R}^2}(P) \cap S^1$$

是 S^1 中含 P 的开邻域，而且不含 N ，所以它包含于 U_N ； U_S 同理。

定义

$$\varphi_N(x, y) = \frac{x}{1-y}, \quad \varphi_S(x, y) = \frac{x}{1+y}.$$

它们都映到 $\mathbb{R}$ ，逆映射分别是

$$\varphi_N^{-1}(t) = \left(\frac{2t}{1+t^2}, \frac{t^2-1}{1+t^2} \right), \quad \varphi_S^{-1}(t) = \left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2} \right).$$

直接代入可以验证两边复合都是恒等映射。原映射的分母在相应定义域内非零，逆映射的分母 $1 + t^2$ 始终为正。第九章的求导法则与可微性判别说明这些环境空间中的有理式映射属于 C^∞ ，因而连续；把定义域或陪域换成相应子空间后，连续性由开集原像定义直接保持。因此原映射与逆映射都连续，它们确实是同胚坐标图。

在重叠部分， $t \neq 0$ ，而且

$$(\varphi_S \circ \varphi_N^{-1})(t) = \frac{1}{t}.$$

这个函数及其逆函数在 $\mathbb{R} \setminus \{0\}$ 上都是光滑的，故两张图组成光滑图册。Hausdorff 与第二可数性质由 $\mathbb{R}^2$ 的子空间继承。

最后，第九章的可微性判别说明多项式函数 $q(x, y) = x^2 + y^2$ 连续，并且 $S^1 = q^{-1}(\{1\})$ 。因为 $\mathbb{R} \setminus \{1\}$ 开，连续性使 $q^{-1}(\mathbb{R} \setminus \{1\})$ 开，所以其补集 S^1 闭。

又因 S^1 每一点的标准范数都等于 1，它有界。由 Heine–Borel 定理， S^1 紧致。证毕。

这个例子还展示了“整体弯曲，局部平直”的准确含义：我们没有把整个圆周压成一条直线，而是允许两张直线地图在 $t \mapsto 1/t$ 的规则下重叠。

【思考训练】

一个坐标图能否覆盖整个圆周 S^1 ？至少需要多少张坐标图才能覆盖它？

解答：

一个坐标图不能覆盖整个圆周：否则，圆周会同胚于 $\mathbb{R}$ 的某个非空开子集 O 。但这样的 O 不可能紧致。

事实上，若 O 紧致，Heine–Borel 定理使它有界且闭。令 $s = \sup O$ 。若 $s \notin O$ ，则 s 属于开集 $\mathbb{R} \setminus O$ ，故某个以 s 为中心的开区间与 O 不相交。可是上确界定义保证该区间的左半边含有足够接近 s 的 O 中元素，矛盾。因此 $s \in O$ 。

又因为 O 开，存在某个 $\delta > 0$ 使 $s + \delta \in O$ ，这与 s 是上界矛盾。

同胚保持紧致性，所以一个坐标图的假设不成立。前面构造的两张坐标图确实覆盖圆周，因此最少需要两张。

10.2 在弯曲中寻找平直 —— 切空间

流形未必拥有全局坐标，但在每个点附近都有欧氏坐标。固定一点 P ，曲线经过 P 时的一阶速度为切向量提供了直观模型。

曲线速度很适合想象切向量，但到了带边流形的边界，它会漏掉向外方向。为了让同一个定义从现在一直适用到斯托克斯定理，我们直接用坐标转换来粘合向量。

【定义】 取一个含 P 的坐标图 (U, φ) 和一个向量 $v \in \mathbb{R}^n$ ，把二者组成一个代表 (φ, v) 。若 (V, ψ) 是另一张含 P 的坐标图，规定

$$(\varphi, v) \sim (\psi, w) \iff w = D(\psi \circ \varphi^{-1})(\varphi(P))v.$$

这里右侧是第九章定义的导数线性映射。一个切向量就是这种代表的等价类，所有等价类组成 $T_P M$ 。

在任一固定坐标图中定义

$$[(\varphi, v)] + [(\varphi, w)] = [(\varphi, v + w)], \quad c[(\varphi, v)] = [(\varphi, cv)].$$

【定理（切空间良定义）】 关系 $\sim$ 是等价关系，上述运算与代表坐标图无关； $T_P M$ 是 n 维实向量空间，称为点 P 的切空间。对抽象流形，它不必来自某个预先给定的环境向量空间。

证明：同一坐标图到自身的转换是恒等映射，其导数是恒等线性映射，所以 $(\varphi, v) \sim (\varphi, v)$ 。若 $(\varphi, v) \sim (\psi, w)$ ，坐标转换 $\varphi \circ \psi^{-1}$ 是 $\psi \circ \varphi^{-1}$ 的逆；链式法则给出

$$D(\varphi \circ \psi^{-1})(\psi(P))w = v,$$

所以关系可反向。若再有 $(\psi, w) \sim (\theta, z)$ ，链式法则又给出

$$z = D(\theta \circ \psi^{-1})(\psi(P))D(\psi \circ \varphi^{-1})(\varphi(P))v = D(\theta \circ \varphi^{-1})(\varphi(P))v,$$

所以关系可传递。这证明了它是等价关系。

坐标转换的导数是线性映射，因此它同时保持向量加法与数乘；这说明换一张坐标图再做运算，与先做运算再换图结果相同。最后，固定任意一张含 P 的坐标图 (U, φ) ，映射

$$T_P M \rightarrow \mathbb{R}^n, \quad [(\varphi, v)] \mapsto v$$

是线性双射：每个等价类在这张图中有且只有一个代表向量。故 $T_P M$ 与 $\mathbb{R}^n$ 同构，维数为 n 。该论证也包含 $n = 0$ ；此时 $T_P M$ 是零向量空间。证毕。

【例子： $\mathbb{R}^n$ 中的切向量】

把 $M = \mathbb{R}^n$ 配上恒等坐标图。此时 $T_P \mathbb{R}^n$ 中的每个切向量唯一写成 $[(\text{id}, v)]$ ，所以可以自然地把它记成普通向量 $v \in \mathbb{R}^n$ 。要注意，这个识别来自全局恒等坐标；一般流形上并没有一张同时覆盖所有点的恒等地图。

【定理（内部点的曲线解释，证明略）】 若 P 是无边界流形中的点，或以后所说带边流形的内部点，那么每条满足 $\gamma(0) = P$ 的光滑曲线 $\gamma : (-\epsilon, \epsilon) \rightarrow M$ 定义切向量

$$\gamma'(0) = [(\varphi, (\varphi \circ \gamma)'(0))],$$

且每个切向量都可这样得到。这个等式与坐标选择无关。边界点仍使用上面的坐标代表定义，不能只靠双侧曲线速度。

【定义】 若 $F : M \rightarrow N$ 光滑，取 P 附近的坐标图 (U, φ) 和 $F(P)$ 附近的坐标图 (V, ψ) ，定义切映射

$$DF_P : T_P M \rightarrow T_{F(P)} N,$$

$$DF_P([(\varphi, v)]) = [(\psi, D(\psi \circ F \circ \varphi^{-1})(\varphi(P))v)].$$

【定理（切映射良定义，证明略）】 链式法则保证这个定义不依赖坐标图或代表，并且 DF_P 是线性映射。对内部点的曲线表示，它等价于

$$DF_P(\gamma'(0)) = (F \circ \gamma)'(0).$$

【思考训练】

设 $F : M \rightarrow N$ 、 $G : N \rightarrow L$ 都光滑。猜一猜恒等映射和复合映射的切映射应满足什么公式。

解答：链式法则直接给出

$$D(\text{id}_M)_P = \text{id}_{T_P M}, \quad D(G \circ F)_P = DG_{F(P)} \circ DF_P.$$

这说明“先走过 F 再走过 G ”的瞬时变化，等于先用 DF_P 变换切向量，再用 $DG_{F(P)}$ 变换一次。公式两侧的定义域和陪域完全相同，因而复合是合法的。

必须注意：不同点的切向量属于不同的切空间，一般没有天然的加法。若要比较它们，需要另行选择额外结构；本章只在同一点的切空间内运算。

10.3 测量切向量 —— 微分形式 dx

每个切空间 $T_P M$ 都是向量空间，因此第八章的对偶空间构造可以应用于它。对偶空间 $(T_P M)^*$ 称为点 P 的余切空间；它的元素是线性泛函：输入切向量，输出实数。

【定义】 在每个 $P \in M$ 选取一个线性泛函 $\omega_P \in (T_P M)^*$ ，称为一个一阶形式场。它的光滑性将在引入坐标余切向量 dx^i 后，用坐标系数函数精确定义。

若坐标图 φ 的坐标函数为 $x^1, \dots, x^n$ ，并且切向量在这张图中写成 $v = [(\varphi, \xi)]$ ，其中 $\xi = (\xi_1, \dots, \xi_n)$ ，定义

$$(dx^i)_P(v) = \xi_i.$$

它们是坐标切向量基底的对偶基底。对内部点的曲线表示，这等于 $\left. \frac{d}{dt} x^i(\gamma(t)) \right|_{t=0}$ 。 dx^i 通常只在该坐标图内定义，换坐标时会按链式法则变换。

dx 不是“无限小的长度”，而是作用于切向量的线性泛函。在每个点 P ， dx 与 dy 的值都属于同一个余切空间 $(T_P M)^*$ ，所以可以把第八章的外代数构造应用于该空间，得到楔积 $dx \wedge dy$ 。下面把这种构造写成微分形式的正式定义。

【定义】 设 $k \geq 1$ 。映射

$$\omega_P : (T_P M)^k \rightarrow \mathbb{R}$$

这里 $(T_P M)^k$ 表示从 $[k]$ 到 $T_P M$ 的映射组成的集合，也就是有序 k 元组的集合。若 ω_P 在每一个输入位置分别都是线性的，就称为 k 重线性；若交换任意两个输入都会使值变号，就称为交替的。等价地，只要两个输入相同，值就必须为 0。点 P 的一个 k 阶形式就是这样的交替 k 重线性映射；在每个 $P \in M$ 选取一个 ω_P ，称为一个 k 阶形式场。

零阶微分形式另行定义为光滑函数，因此

$$\Omega^0(M) = C^\infty(M).$$

空楔积约定为标量 1。若 f 是零阶形式，则 $f \wedge \omega = \omega \wedge f = f\omega$ 。

对一阶形式，楔积由

$$(\alpha_1 \wedge \cdots \wedge \alpha_k)(v_1, \dots, v_k) = \det(\alpha_i(v_j))$$

定义，再用双线性扩展；一般的 r 阶与 s 阶形式的楔积是 $r + s$ 阶形式。

$dx \wedge dy$ 作用于两个切向量时给出带符号的坐标平行四边形值；它可能为负，也不等于在缺少度量时的“绝对面积大小”。

【可算例子】

在 $\mathbb{R}^2$ 的恒等坐标中，令

$$u = (a, b), \quad v = (c, d).$$

那么

$$(dx \wedge dy)(u, v) = \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc.$$

若交换 u, v ，行列式变号；若 $v = \lambda u$ ，结果为 0。因此一个二阶形式不是“一个很小的面积块”，而是一条规则：它接收两个切向量，并返回一个带符号的实数。

【定理兼定义（形式的坐标表达与光滑性，证明略）】当 $k \geq 1$ 时，一个 k 阶形式场在每张坐标图中唯一写成

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k},$$

其中 $a_{i_1 \dots i_k}$ 是实值函数。据此定义：这个形式场是光滑 k 阶微分形式，当且仅当在每张相容坐标图中，所有系数函数 $a_{i_1 \dots i_k}$ 都属于 C^∞ 。坐标转换与链式法则保证，只在一份光滑图册中检查就足够，而且这个条件不依赖图册选择。一阶形式场也按这条规则判定是否光滑。

【定义】 若 $F : N \rightarrow M$ 光滑， ω 是 M 上的 k 形式，则拉回形式 $F^*\omega$ 定义为

$$(F^*\omega)_P(v_1, \dots, v_k) = \omega_{F(P)}(DF_P v_1, \dots, DF_P v_k).$$

对零阶形式 f ，同一定义就是 $F^*f = f \circ F$ 。

【例子：把平面形式拉回到一条曲线】

令 $F : \mathbb{R} \rightarrow \mathbb{R}^2$ 为 $F(t) = (t, t^2)$ 。它的导数把一维切向量 s 送到 $(s, 2ts)$ ，所以

$$F^*(dx) = dt, \quad F^*(dy) = 2t dt.$$

再取两个一维切向量 s_1, s_2 ，直接使用拉回定义可得

$$F^*(dx \wedge dy)(s_1, s_2) = \det \begin{pmatrix} s_1 & s_2 \\ 2ts_1 & 2ts_2 \end{pmatrix} = 0.$$

一条曲线只有一个独立方向，两个切向量必定线性相关，所以任何二阶形式拉回到曲线上都为零。这严格表达了“维数不足以承载二阶交替形式”。

【定义】 对零阶形式 f ，先定义

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x^j} dx^j.$$

对 $k \geq 1$ 的形式，外微分 d 在坐标中定义为

$$d\omega = \sum_{i_1 < \dots < i_k} \sum_{j=1}^n \frac{\partial a_{i_1 \dots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

【定理（外微分良定义，证明略）】该公式与坐标选择无关，并满足线性、 $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^r \alpha \wedge d\beta$ （ α 为 r 阶形式）以及 $d^2 = 0$ 。

【短证明：为什么对函数有 $d^2 f = 0$ 】

若 f 光滑，则在坐标中

$$d(df) = \sum_{i,j} \frac{\partial^2 f}{\partial x^j \partial x^i} dx^j \wedge dx^i.$$

当 $i = j$ 时， $dx^i \wedge dx^i = 0$ 。当 $i \neq j$ 时，第九章的混合偏导交换定理说明 (i, j) 项与 (j, i) 项的二阶偏导数相等，但楔积互为相反数，所以两项抵消。

全部项都消失，故 $d(df) = 0$ 。一般 k 形式的证明也是同一组“对称偏导数与反对称楔积相抵消”的计算，只是下标更多。

【思考训练】

在 $\mathbb{R}^2$ 上令

$$\omega = x \, dy - y \, dx.$$

求 $d\omega$ 。

解答：使用外微分的乘积法则，并注意 $d(dx) = d(dy) = 0$ ，得到

$$d\omega = dx \wedge dy - dy \wedge dx = 2 \, dx \wedge dy.$$

第二项不是把第一项消掉，而是因为换序变号后又贡献了同一个 $dx \wedge dy$ 。

【思考训练】

根据第八章的反对称法则， $dy \wedge dx$ 与 $dx \wedge dy$ 有什么关系？ $dx \wedge dx$ 又等于什么？

解答：

交换次序会变号，所以

$$dy \wedge dx = -dx \wedge dy.$$

同一个一阶形式与自身作楔积则为零： $dx \wedge dx = 0$ 。

10.4 选择一致的方向 —— 定向与边界

上一章的积分从越来越细的有限分割出发，计算有限黎曼和；当这些和按定义共同逼近同一实数时，那个极限就是积分。要在流形上积分微分形式 ω ，还必须先处理方向的一致性。

【直觉】 想象把纸条扭转半圈再粘合，并沿纸条绕行一周：起初选定的局部方向看起来会被翻转。这幅图只提示“局部选择未必能在全局保持一致”；下面的正式定义不以这幅图为证明。

【定义】 设 V 是 n 维实向量空间， $n \geq 1$ 。对 V 的两组有序基底

$$(v_1, \dots, v_n), \quad (w_1, \dots, w_n),$$

由第八章的基底坐标定理，存在唯一线性同构 $T: V \rightarrow V$ 满足 $T(v_i) = w_i$ ($1 \leq i \leq n$)。若 $\det T > 0$ ，就说这两组基底**同向**；若 $\det T < 0$ ，就说它们**反向**。

【引理】 “同向”是有序基底之间的等价关系；当 $n \geq 1$ 时，它恰好把全部有序基底分成两个等价类。任选其中一类称为**正基底**，另一类便称为**负基底**。

证明：恒等换基映射的行列式是 1，所以每组基底与自己同向。若 T 把第一组基底送到第二组，那么 T^{-1} 把第二组送回第一组，而且

$$\det(T^{-1}) = \frac{1}{\det T},$$

故二者符号相同。依次换两次基底对应复合映射，而

$$\det(S \circ T) = \det S \cdot \det T;$$

两个正数的乘积仍为正，所以“同向”可传递。

现在固定一组基底 $\mathcal{B}$ 。任意其他基底相对于 $\mathcal{B}$ 的换基行列式非零，因而不是正就是负；两个正号或两个负号相除都得到正号，一正一负相除则得到负号，所以恰有两类。又因为 $n \geq 1$ ，把 $\mathcal{B}$ 的第一个向量换成其相反向量便得到行列式为 -1 的基底，故两类都非空。证毕。

【例子：平面的标准定向】

在 $\mathbb{R}^2$ 中规定 (e_1, e_2) 为正基底，就得到标准定向。若

$$u = (1, 1), \quad v = (-1, 1),$$

则以 u, v 为列的矩阵行列式为 $2 > 0$ ，所以 (u, v) 是正基底；交换次序后的 (v, u) 行列式为 -2 ，是负基底。这里，换序变号正好记录了定向的反转。

【定义】当 $n \geq 1$ 时， n 维流形的一个定向，是在每个切空间中选定上述两个等价类中的哪一类为“正”，并满足下面的局部一致性：对任意定义域连通的坐标图 (U, φ) ，坐标基底

$$[(\varphi, e_1)], \dots, [(\varphi, e_n)]$$

在整个 U 上要么处处为正，要么处处为负。能这样选择的流形称为可定向；一次具体选择称为一个定向。

对没有边界的流形，这等价于可以选取一份所有坐标转换雅可比行列式都为正的图册。带边流形固定使用 $\mathbb{H}^n = \{x_n \geq 0\}$ 作模型时，边界图未必都能改成正向；因此，带边情形在下文使用每张图相对于定向的正负符号，不能假定存在覆盖全体的正向边界图。零维定向将在积分前用单独的符号约定处理。

“里面/外面”并不是一般定向的定义。例如一维圆周的定向只是选择顺时针或逆时针。前面的纸带图景只帮助想象局部方向可能遇到的全局障碍，不参与后文推理。

【定义】设 $n \geq 1$ 。带边 n 维流形仍要求 Hausdorff 与第二可数，但把局部欧几里得模型改成半空间

$$\mathbb{H}^n = \{(x_1, \dots, x_n) : x_n \geq 0\}.$$

坐标图映到 $\mathbb{H}^n$ 的相对开子集；坐标转换“光滑”表示它在每一点附近可扩展为普通欧几里得开集上的光滑映射。

这里“相对开”定义为 $O \cap \mathbb{H}^n$ ，其中 O 是 $\mathbb{R}^n$ 的开集。

带边流形之间的映射也按同一规则定义为光滑：它的每个坐标表达都必须在相应点附近拥有普通欧几里得开集上的 C^∞ 延拓。

能在坐标中落到 $x_n = 0$ 的点组成边界 ∂M ；其余点组成内部。

这里的 ∂M 是由半空间坐标定义的**流形边界**，不要同第六章“某个子集在给定环境空间中的拓扑边界”混淆。若把 M 只看成自身的子集，它的拓扑边界反而是空集。

【定理（边界与切空间，证明略）】坐标转换和半空间的拓扑性质保证“边界点”不依赖坐标选择； ∂M 是 M 的闭子集，并且本身是 $(n-1)$ 维光滑流形。

前面用坐标代表定义的 $T_P M$ 原样适用于边界点，仍是 n 维；光滑扩展的导数在边界处唯一确定。若 $P \in \partial M$ ，则 $T_P \partial M$ 自然对应于边界坐标中满足第 n 个分量为 0 的子空间。

【定义】包含映射 $\iota : \partial M \hookrightarrow M$ 是光滑映射。一个形式在边界上的“限制”，严格定义为拉回 $\iota^* \omega$ ，而不是把符号中的变量随意删掉。

【定义】对 $P \in \partial M$ ，若 $v \in T_P M$ 在某张边界坐标中的第 n 个分量小于 0，就称 v **向外且横截边界**。这里没有给流形选择内积，所以不要求 v 与边界“垂直”；正式条件只有刚才规定的坐标分量符号。

【引理（向外方向良定义，证明略）】边界坐标转换的导数保持“第 n 个分量为负”这一侧，因此向外条件与所选边界坐标图无关。

【定义】若 M 已定向，则边界采用“向外横截向量在前”的约定： $(v_1, \dots, v_{n-1})$ 在边界为正，当且仅当 $(v, v_1, \dots, v_{n-1})$ 在 M 中为正，其中 v 是任一向外横截向量。

【引理（边界定向良定义，证明略）】不同的向外横截向量给出同一类正基底，因此上述约定唯一确定斯托克斯公式中的符号。

【一维边界的符号从哪里来】

给区间 $[A, B]$ 选择从左到右的定向，也就是规定正的切向量指向数值增大的方向。在右端点 B ，向外向量也指向右方，所以右端点取得正号；在左端点 A ，向外向量指向左方，与区间定向相反，所以左端点取得负号。于是边界不是无符号的集合 $\{A, B\}$ ，而是带符号的形式

$$\partial[A, B] = \{B\} - \{A\}.$$

这不是额外记忆规则，而是“向外横截向量在前”在一维中的直接结果。

10.5 统一内部与边界 —— 斯托克斯定理

现在设 M 是定向光滑 n 维带边流形, 其中 $n \geq 1$, 边界为 ∂M 。下面只在积分对象具有紧支撑时定义积分。

【定义】函数或微分形式的**支撑**, 是它不为零的点集在所在空间中的闭包, 记为 supp 。若这个闭包紧致, 就称它具有**紧支撑**。

若 M 紧致, 每个光滑形式都自动具有紧支撑。事实上, 记支撑为 K ; 它按定义是 M 的闭子集。给定 K 在子空间拓扑下的开覆盖, 把每个覆盖成员写成 $K \cap V_i$, 其中 V_i 在 M 中开。于是这些 V_i 与开集 $M \setminus K$ 一起覆盖 M 。紧致性给出有限子覆盖, 删去 $M \setminus K$ 后, 相应的有限多个 $K \cap V_i$ 仍覆盖 K , 所以 K 紧致。

【零维积分约定】沿用前面的零维微分约定。

为了让 $n = 1$ 的斯托克斯定理包含一维微积分基本定理, 我们还约定: 定向零维流形的每个点 P 带一个符号 $\varepsilon(P) \in \{+1, -1\}$ 。若零阶形式(函数) f 具有紧支撑, 则

$$\int_N f = \sum_{P \in N} \varepsilon(P) f(P).$$

零维流形是离散空间: 每个点的坐标邻域同胚于 $\mathbb{R}^0$ 的单点开集, 因而该点本身是开集。离散空间的紧致子集只能是有限集, 因为它的全部单点集组成开覆盖, 而有限子覆盖只能覆盖有限多个点。因此上式只有有限多个非零项。作为一维流形边界时, 前面的“向外横截向量在前”约定决定这些符号: 对采用从左到右定向的区间, 右端点为正、左端点为负; 区间反向定向时, 两端符号同时反转。

在写积分前还缺一个拼接工具。

【定理(光滑单位分解, 证明略)】设 M 是第二可数光滑流形, 允许带边; $\{U_i\}_{i \in I}$ 是它的开覆盖。存在非负光滑函数 $\{\rho_i\}_{i \in I}$, 满足

1. $\text{supp } \rho_i \subseteq U_i$;
2. 每个点都有一个邻域, 只与有限多个 $\text{supp } \rho_i$ 相交; 这项性质称为**局部有限**;
3. 对每个 $P \in M$, 只有有限多个 $\rho_i(P)$ 非零, 且 $\sum_{i \in I} \rho_i(P) = 1$ 。

这样的函数族称为**从属于该开覆盖的光滑单位分解**。局部有限性还保证: 任意紧集只与有限多个 $\text{supp } \rho_i$ 相交。

这里对 I 的求和表示: 只把非零的有限多个实数相加。实数加法的交换律与结合律保证结果不依赖排列次序; 后面单位分解产生的求和都沿用这个约定。

【定义】闭长方体是

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n,$$

其中 $a_j \leq b_j$ ($1 \leq j \leq n$), 体积定义为 $\text{vol}(Q) = \prod_{j=1}^n (b_j - a_j)$ 。

一个带标记长方体分割是有限族

$$\mathcal{P} = \{(Q_1, t_1), \dots, (Q_N, t_N)\},$$

并满足以下条件:

1. 每个 $Q_j = \prod_{r=1}^n [\alpha_{jr}, \beta_{jr}]$ 都是坐标轴平行的闭长方体;
2. $Q = \bigcup_{j=1}^N Q_j$, 且不同 Q_j 的开内部 $\prod_{r=1}^n (\alpha_{jr}, \beta_{jr})$ 互不相交;
3. 标记点 $t_j \in Q_j$ 。

使用第八章的标准范数, 定义

$$\text{diam}(Q_j) = \sup\{\|x - y\| : x, y \in Q_j\}, \quad \mu(\mathcal{P}) = \max_{1 \leq j \leq N} \text{diam}(Q_j).$$

这里上确界由实数完备性存在, 有限最大值由实数全序存在。对函数 $f : Q \rightarrow \mathbb{R}$, 相应黎曼和是

$$R(f, \mathcal{P}) = \sum_{j=1}^N f(t_j) \text{vol}(Q_j).$$

【定义】若存在 $I \in \mathbb{R}$, 使对每个 $\epsilon > 0$ 都存在 $\delta > 0$, 并且每个满足 $\mu(\mathcal{P}) < \delta$ 的带标记长方体分割都有

$$|R(f, \mathcal{P}) - I| < \epsilon,$$

就称 f 在 Q 上黎曼可积, 并定义 $\int_Q f = I$ 。

由阿基米德性质反复等分各边, 可以得到网格任意小的分割; 再与第九章一样使用三角不等式, 便可证明积分值唯一。

【例子兼短证明： $\int_0^1 x dx = \frac{1}{2}$ 】

取任意一维带标记分割；删去长度为 0、对黎曼和没有贡献的退化小段后，按从左到右把端点写成

$$0 = x_0 < x_1 < \cdots < x_N = 1,$$

并在第 j 段选标记 $t_j \in [x_{j-1}, x_j]$ 。记 $\Delta x_j = x_j - x_{j-1}$ 。恒等式

$$\frac{1}{2} = \sum_{j=1}^N \frac{x_j + x_{j-1}}{2} \Delta x_j$$

来自右侧的每一项都是 $(x_j^2 - x_{j-1}^2)/2$ ，求和后中间项依次消去。于是

$$\begin{aligned} \left| R(f, \mathcal{P}) - \frac{1}{2} \right| &\leq \sum_{j=1}^N \left| t_j - \frac{x_j + x_{j-1}}{2} \right| \Delta x_j \\ &\leq \frac{1}{2} \sum_{j=1}^N (\Delta x_j)^2 \leq \frac{\mu(\mathcal{P})}{2} \sum_{j=1}^N \Delta x_j = \frac{\mu(\mathcal{P})}{2}. \end{aligned}$$

给定 $\epsilon > 0$ ，取 $\delta = 2\epsilon$ ，便完全满足积分定义。这不只算了一串特殊黎曼和，而是同时控制了所有足够细、标记任意的分割，所以确实构成证明。

半空间中的**闭半长方体**是

$$Q^+ = [a_1, b_1] \times \cdots \times [a_{n-1}, b_{n-1}] \times [0, b_n] \subseteq \mathbb{H}^n,$$

其中 $a_j \leq b_j$ ($1 \leq j \leq n-1$) 且 $b_n \geq 0$ ，并用完全相同的带标记分割、网格与 ϵ - δ 条件定义 $\int_{Q^+} f$ 。

【定理（长方体上的连续函数可积，证明略）】闭长方体和闭半长方体上的连续函数都黎曼可积。证明利用 Heine–Borel 紧致性把逐点连续给出的局部误差控制化为有限个统一控制，再用实数完备性处理所有足够细的带标记分割。

【引理（零延拓与积分区域无关，证明略）】设 E 是 $\mathbb{R}^n$ 的开集，或者是 $\mathbb{H}^n$ 的相对开集，并且连续函数 $f : E \rightarrow \mathbb{R}$ 具有紧支撑。把 f 在相应环境空间 $\mathbb{R}^n$ 或 $\mathbb{H}^n$ 内延拓为零，所得函数仍连续。任取一个包含其支撑的大长方体或大半长方体来积分，结果都相同；扩大积分区域只增加函数恒为零的部分。

【定义】在上述条件下，把零延拓在任一这样的大长方体或大半长方体上的积分定义为 $\int_E f$ 。前面的定理与引理分别保证这个积分存在并且不依赖长方体的选择。

【定义】设 U, V 同为 $\mathbb{R}^n$ 的开集，或同为 $\mathbb{H}^n$ 的相对开集。 C^1 微分同胚是双射 $\Phi: U \rightarrow V$ ，并且 Φ 与逆映射 Φ^{-1} 都属于 C^1 ；在半空间情形， C^1 按前述“可延拓”意义理解。

【定理（多重积分与换元，证明略）】若 $U, V \subseteq \mathbb{R}^n$ 是开集， $\Phi: U \rightarrow V$ 是 C^1 微分同胚，且 $f: V \rightarrow \mathbb{R}$ 连续并具有紧支撑，则

$$\int_V f(y) dy = \int_U f(\Phi(x)) |\det D\Phi(x)| dx.$$

若 U, V 是 $\mathbb{H}^n$ 的相对开集， Φ 及其逆映射按前述“可光滑延拓”意义为 C^1 ，同一公式对半空间积分成立。这个定理由一维积分、紧致性和线性近似推出。

【定义兼定理（微分形式的积分，证明略）】取一族定义域连通且覆盖 M 的坐标图 (U_i, φ_i) ；这样的覆盖可由任意图册在每个点附近选取连通的较小坐标邻域得到。再取从属于该覆盖的单位分解。由定向的局部一致性，坐标基底在整个 U_i 上有一个不变的符号，定义

$$\varepsilon_i = \begin{cases} +1, & [(\varphi_i, e_1)], \dots, [(\varphi_i, e_n)] \text{ 为正基底,} \\ -1, & [(\varphi_i, e_1)], \dots, [(\varphi_i, e_n)] \text{ 为负基底.} \end{cases}$$

若 η 是紧支撑光滑 n 形式，则在第 i 张图中唯一写成

$$\rho_i \eta = f_i(x) dx^1 \wedge \dots \wedge dx^n.$$

若图像是欧氏开集，就用欧氏积分；若图像是半空间的相对开集，就用半空间积分，并定义

$$\int_M \eta = \sum_i \varepsilon_i \int_{\varphi_i(U_i)} f_i(x) dx.$$

因为 η 紧支撑且单位分解局部有限，这个和实际上只有有限个非零项。又因为 $\text{supp } \rho_i \subseteq U_i$ ，每一项都在一张坐标图内具有紧支撑。

对非零实数 t ，定义 $\text{sgn}(t) = 1$ （当 $t > 0$ ）和 $\text{sgn}(t) = -1$ （当 $t < 0$ ）。在两张图重叠处，若 $\Phi = \varphi_j \circ \varphi_i^{-1}$ ，则对每个 $x \in \varphi_i(U_i \cap U_j)$ 都有

$$\varepsilon_j = \varepsilon_i \operatorname{sgn}(\det D\Phi(x)).$$

这个符号关系与含绝对值的换元公式恰好抵消坐标反向造成的负号，从而证明积分与坐标图及单位分解的选择无关。

特别地，对采用从左到右定向的区间，固定使用 $\mathbb{H}^1 = [0, \infty)$ 时，右端的边界图虽然是反向图，因子 $\varepsilon_i = -1$ 会给出正确符号。

【引理（区间形式积分与普通积分一致）】设 $A < B$ ，并给 $[A, B]$ 采用从左到右的定向。若 h 光滑，则

$$\int_{[A,B]} h(x) \, dx = \int_A^B h(x) \, dx,$$

右侧是第九章的一维黎曼积分。

证明：区间内部可使用坐标 x ；左端使用 $u = x - A$ ，它是正向边界图；右端使用 $v = B - x$ ，它是反向边界图，所以该图的符号因子是 -1 。取从属于这三类图的单位分解。

在右端图中

$$h(x) \, dx = -h(B - v) \, dv;$$

系数里的负号与图的符号因子 -1 相消，左端图和内部图则没有额外负号。分别使用一维换元公式后，各项都成为普通 x 坐标中的积分。单位分解系数之和为 1，所以相加正好得到 $\int_A^B h(x) \, dx$ 。证毕。

【例子：形式积分会记住方向】

给 $[0, 1]$ 采用从左到右的定向，并取一阶形式 $\eta = x \, dx$ 。由刚才的一致性引理和前面的黎曼和计算，

$$\int_{[0,1]} \eta = \frac{1}{2}.$$

若把区间的定向反转，每张坐标图的符号因子都同时乘以 -1 ，因而整个积分变号：

$$\int_{-[0,1]} \eta = -\frac{1}{2}.$$

这里 $-[0, 1]$ 只表示“同一流形配上相反定向”，并不是把点集中的每个数变成负数。

【引理(支撑与拉回)】对任意光滑形式 ω , 有 $\text{supp}(\mathrm{d}\omega) \subseteq \text{supp } \omega$; 并且 $\text{supp}(\iota^*\omega) \subseteq \iota^{-1}(\text{supp } \omega)$ 。

证明: 若 $P \notin \text{supp } \omega$, 就有一个含 P 的开邻域, 使 ω 在其中恒为零。其所有坐标系数及偏导数也都为零, 所以 $\mathrm{d}\omega$ 在该邻域为零, 第一项包含关系成立。

若 $P \notin \iota^{-1}(\text{supp } \omega)$, 则 $\iota(P)$ 有一个邻域使 ω 恒为零; 在其逆像邻域上, $\iota^*\omega$ 也恒为零, 第二项包含关系成立。证毕。

现在取一个光滑 $(n-1)$ 形式 ω , 并假设 M 紧致, 或者假设 ω 具有紧支撑。结合上面引理以及 ∂M 在 M 中闭, 可知 $\mathrm{d}\omega$ 与 $\iota^*\omega$ 都具有紧支撑。

现在有两种积分:

边界积分。

把 ω 沿包含映射 $\iota: \partial M \hookrightarrow M$ 拉回到带诱导定向的边界, 得到可积分的 $(n-1)$ 形式。简写为 $\int_{\partial M} \omega$, 严格含义是 $\int_{\partial M} \iota^*\omega$ 。

内部积分。

外微分 $\mathrm{d}\omega$ 是 M 上可积分的 n 形式, 记为 $\int_M \mathrm{d}\omega$ 。

从定义表面看, 边界积分与内部积分是两种不同的计算。斯托克斯定理说明, 它们通过外微分精确联系起来。

【直觉】把区域分成相邻小块后, 一条内部公共边在两块中取得相反的诱导方向, 因此对应贡献相消。真正的外边界只出现一次, 所以最终保留下来。

【定理 (广义斯托克斯定理, 证明略)】设 $n \geq 1$ 。若 M 是紧致定向光滑 n 维带边流形, ω 是定义在 M 上的光滑 $(n-1)$ 形式, 则

$$\int_{\partial M} \omega = \int_M \mathrm{d}\omega.$$

也就是说, 边界上的形式积分等于外微分在内部的积分。

若 M 不紧致, 只要 ω 具有紧支撑, 同一结论仍成立。证明先在半空间坐标盒中用一维微积分基本定理验证局部公式, 再用单位分解求和。内部边界项成对反号相消, 只留下 ∂M 。因此, “内部相消”是正式证明的最后一步, 而不是定理的全部依据。

【二维局部算例：四条边怎样只留下一个总数】

取 $a, b > 0$ 和长方形 $Q = [0, a] \times [0, b]$, 并按“下、右、上、左”的次序给四条边参数化:

$$\begin{aligned}\gamma_1(t) &= (t, 0) & (0 \leq t \leq a), \\ \gamma_2(t) &= (a, t) & (0 \leq t \leq b), \\ \gamma_3(t) &= (a - t, b) & (0 \leq t \leq a), \\ \gamma_4(t) &= (0, b - t) & (0 \leq t \leq b).\end{aligned}$$

这四个方向合起来是平面标准定向诱导的逆时针边界方向。对 $\omega = x \, dy$, 逐条拉回可得

$$\gamma_1^* \omega = 0, \quad \gamma_2^* \omega = a \, dt, \quad \gamma_3^* \omega = 0, \quad \gamma_4^* \omega = 0.$$

所以四个一维积分之和是 ab 。另一方面,

$$d\omega = dx \wedge dy,$$

常数系数 1 的每个黎曼和都是各小长方形的体积之和。把分割中出现的全部横、纵边界线汇成一个有限网格, 再逐行、逐列使用区间长度的可加性, 使得这个体积之和恒为 $\text{vol}(Q) = ab$ 。所以它在 Q 上的二重积分就是 ab , 与四条边上的积分之和相等。

严格地说, 长方形的四个顶点是“角”, 而本章定义的带边流形只允许半空间型边界, 没有另外定义带角流形。因此, 这里不把 Q 当作广义斯托克斯定理的直接实例, 而只把它作为由四个一维积分组成的局部计算。

若要把长方形本身纳入定理, 需要另行定义“带角流形”。

【推论(一维检验)】给区间 $[A, B]$ 采用从左到右的定向, 并令零阶形式为光滑函数 f 。此时 $df = f'(x) \, dx$, 边界符号是右端点为正、左端点为负, 所以斯托克斯公式正好化为

$$f(B) - f(A) = \int_A^B f'(x) \, dx,$$

即第九章微积分基本定理。

【思考训练】

若紧致定向流形 M 没有边界, ω 是一个光滑 $(n-1)$ 形式, $\int_M d\omega$ 等于多少?

解答：此时 $\partial M = \emptyset$ 。空集合上的积分是空和，按有限和约定等于 0，所以斯托克斯定理立即给出

$$\int_M d\omega = \int_{\partial M} \omega = 0.$$

这条结论说的是“外微分得到的最高阶形式”在无边界紧致流形上的总积分为零；不能把它误读为任意最高阶形式的积分都为零。

10.6 全书结语

我们从“什么算命题，什么算证明”开始，随后用集合构造数，用戴德金分割补成实数直线，再依次建立距离、拓扑、级数、向量、导数与积分。到了最后，坐标图把这些工具带入弯曲的空间，外微分与流形积分则把局部计算连接起来。

于是，斯托克斯公式

$$\int_M d\omega = \int_{\partial M} \omega$$

不再是一条凭空出现的神奇公式。式中的流形、边界、方向、微分形式、外微分与积分都已经获得定义；公式成立所需的条件也已经明确写出。它告诉我们：分散在空间内部的局部变化，可以汇成边界上的一个整体结果。

这条漫长的建造路线体现了本书最重要的方法：遇到一个新对象，先问它是什么意思，再问它是否存在；若它可以有不同写法，就检查结果是否与写法无关；面对一个结论，则追问它究竟由哪些已经建立的规则推出。例子和图景帮助我们发现方向，定义和证明才负责把道路铺稳。

“从零开始”并不等于从绝对虚无开始。本书公开采用经典一阶逻辑和第二章列出的集合论公理，正文中的数学结论都相对于这些起点成立。承认起点，说明依赖，并让每一步都可以回头检查，正是严谨的一部分。

这本书的建造到这里告一段落，但提问不会停止。以后无论遇到怎样的新对象，都可以继续沿用同一种办法：放下一块有明确定义的积木，再用可靠的推理接上下一块。一个可以不断扩展、也经得起逐步检查的数学宇宙，正是这样建立起来的。