借来的
常数

微积分、科学与 AI 的边界

王德生　著

德麦国际出版社
Demai International Press

出版信息

项目	内容
书名	借来的常数——微积分、科学与 AI 的边界
著者	王德生
底本	王德生《科学与微积分：差异、整体与文明的数理内核》（SIO 思想体系文丛）
执笔	AI 写作组（Claude，Anthropic）据底本与作者体系分章扩写
出版	德麦国际出版社（Demai International Press）
编号	德麦国际专著第 498 号
ISBN	979-8-90690-786-8
定价	US$23.00
开本	170 mm × 240 mm
字数	约 23 万字（正文约 21 万字）
版次	2026 年 10 月第 1 版
结构	前言 · 导读 · 导论 · 七编三十五章 · 结语 · 参考书目 · 附录三件 · 后记

AI 协作声明。 本书以王德生《科学与微积分》为底本，在作者的 SDE 本体论体系下，由 AI 写作组分章扩写，经统稿、引文史实核验与三审（一审内容、二审文字、三审终审）成书。书中观点的立论权属于作者；引文与史实的核对依据见附录I，全书主张的推翻条件见附录G。

版权声明。 版权所有，未经出版者书面许可，不得以任何方式复制或转载本书内容。

作者介绍

王德生，计算数学博士（中国科学院），SDE（显露 Show · 差异序列 Difference · 特征纠缠 Entanglement）本体论的创立者，德麦国际（Demai International Pte. Ltd.，新加坡）首席执行官。生长于中国农村，1990 年进入湘潭大学。

曾于中国科学院从事博士后研究，其间与哥伦比亚大学杜强（Qiang Du）教授合作，合著 ACVT（各向异性中心 Voronoi 剖分）奠基性论文；后任职于英国斯旺西大学与新加坡南洋理工大学，在南洋理工大学任博士生导师十三年，长期从事 Voronoi／Delaunay／CVT 网格生成、有限元、自适应算法等计算数学研究。2005 年起定居新加坡。SDE 的最初火种，正是在对“网格如何从散点中生成”这一具体问题的长期追问中点燃的：点还是那些点，规则还是那些规则，网一旦织成，就有了点所没有的东西。

他的思想体系先后经历四站：三视角·知识的结构和体系（2017 年起）、321 智慧系统（2022 年起）、SIO 本体论（2024 年起），以及 SDE 本体论（2026 年起；同年 5 月起，“结构”改读为“显露”）——从“并排放着的三样”，走到“互相生出来的三样”。其核心主张，是把人类的根本追问从“世界是什么”（发现范式）转向“世界如何发生”（发生范式）。

以 SDE 为方法，他已在哲学、数学、人工智能、医学、教育、艺术、文明研究等约二十个学科出版专著二百余部。现居新加坡，创办德麦国际，致力于 SDE 知识创新体系的研究、出版与教学，以及以 SDE 为内核的 AI 超级员工与智能体研发。在他看来，通往真正的通用智能，需要的不是更多“已完成的知识”，而是“知识如何发生”的过程数据。

本书的底本《科学与微积分：差异、整体与文明的数理内核》写于 SIO 阶段，是一位与网格、有限元打了三十年交道的计算数学家，回头追问微积分为什

么能成为科学的心脏。扩写本按 SDE 现行口径把它推进为“现实的微分—理念的积分—落地的再微分”三段，并用它来评估“AI 能解决一切问题”一类的论断：AI 能替人做越来越多的积分，落地时那个常数，仍要从现实里借。

微分忠实抄写差异，积分大胆发明整体。

——王德生《科学与微积分：差异、整体与文明的数理内核》

C 在课本里是一个常数，在现实里是一整片土壤。

——本书第12章

凡例

一、本书正文分七编三十五章，前有导论，后有结语。每章章题下的“这一章的问题”，用一句话交代本章要回答什么；章末的“带走的话”是本章最该记住的判断；“小账”交代本章例子的性质、证据的出处等级、本章不主张什么，以及本章主张的推翻条件在附录G中的编号。

二、正文中带“【方框】”标题的灰底方框，收公式、推导与较严格的论证。跳过方框不影响读懂正文。

三、S、D、E　分别指显露（Show）、差异序列（Difference）、特征纠缠（Entanglement）。S　读作“显露”，不读作“结构”。三方程　S=F(D,E)、D=G(S,E)、E=H(S,D)　在本书中只作关系的占位写法，不代表已经给出了函数的具体形式。新旧术语的对应见附录H。

四、书中引用他人的话，凡见于原文的，注明出处；只见于转述的，注明“据转述”。引文与史实的核对结果见附录I。

五、书中凡以“我”出现的个人经历，属于底本作者王德生。

六、年份、数据、章号用阿拉伯数字；外国人名首次出现时附原文；数学符号按通行写法。

前言

一个常数，为什么只能借

◆ 一、一本五万字的小书

这本书的前身，是王德生的《科学与微积分：差异、整体与文明的数理内核》，收在 SIO 思想体系文丛里，约五万字。那本小书立下了一条很硬的主干：微分忠实抄写差异，积分大胆发明整体，微积分基本定理把二者焊在一起，科学就在这一抄一造之间循环往复。它从主客二分的困境讲起，经过微积分之前那片“黎明前的天空”，讲到牛顿与莱布尼茨，讲到微积分三百年间的几次扩张，一直讲到工程、控制和人工智能。

底本里有一段话，是整本书问题意识的来源。作者写道，从高中第一次接触微积分，到大学里讲授微积分，他用了几十年，才真正明白它的价值与意义。课堂教人怎样求导、怎样换元、怎样解方程，却很少有人说清楚：微积分为什么存在，它在科学与文明里意味着什么。为微积分正名，是底本的初衷，也是这本书仍然保留的初衷。

现在这本书约二十万字，分七编三十五章，前有导论，后有结语和三个附录。从五万字到二十万字，不是把每一句话拉长四倍。本书保留了底本的主干，做了三件事。

第一，换口径。底本写于 SIO 阶段，用的是 ΔSIO、创生律、成就律、演化律这些说法。本书一律改用现行的 SDE 口径，以《SDE 思想体系入门》（第395卷）为准：S 是显露，D 是差异序列，E 是参与发生的纠缠土壤，三方程、六路径、123 原理取代了旧的术语。旧说法在需要时以“底本原作……”一句带出，新旧对照放在附录H。

第二，补一段。底本讲的是“微分—积分”两段，本书把它推成三段：现实的微分，理念的积分，落地的再微分。第三段是新写的，技术、工程、医疗、一

切“应用”都落在这里。随着这一段一起出来的，是本书最要紧、也最容易被检验的一个命题：积分常数 C 必须从现实借。

第三，加一编。本书新设第六编，专门评估哈萨比斯关于“自然界中尚未发现不可计算之物”的论断：先核对原话，再替他把话说足，然后分三刀批判，最后给出判词，并写明这些判词在什么情况下会被推翻。

还有几处是收窄。底本有一些说得太满的话，本书按第395卷的证据责任一一收窄：李约瑟之问的数理解释，改为与制度、经济、地理解释并列的候选之一，并交代它的循环风险与史实风险；伞场微积分只保留在限定核、尺度归一、正则条件下的局部逼近命题；“没有微积分就没有现代科学”，限定为动力学科学。收窄不是退缩。一句话说得越大，能落到实处的部分就越少。

◆ 二、为什么是现在

2024 年的诺贝尔化学奖，有一半颁给了蛋白质结构预测。获奖者之一戴密斯·哈萨比斯在演讲里提出一个猜想，大意是：凡是自然界能够生成或找到的模式，都可以被经典的学习算法高效地发现并建模。后来在访谈里，他又说过，迄今还没有人在宇宙里发现不可计算的东西。这些话在传播中一层层放大，到了短视频的标题里，就成了“没有 AI 不能解决的问题”。

与此同时，“AI for Science”成了一个时代的口号。蛋白质结构、天气预报、新材料筛选，一个个原本要几代人慢慢啃的问题，被机器以惊人的速度推进。这些成绩是真的，本书不打算否认，也不打折扣。

可是那句流行的话，把两个不同的问题捏在了一起。一个问题是“能不能算出来”；另一个问题是“算出来的东西，能不能落到这一个病人、这一块地基、这一个孩子身上”。前一个问题，AI 越来越擅长；后一个问题，从来不只是计算的事。哈萨比斯本人的命题，比流行版要窄得多，也审慎得多，他自己就为量子、混沌和人的提问留了余地。本书分清这两个版本，只评估前者，并且先替它把最强的辩护说足。

这句话之所以值得一本书来回答，是因为它把一个三百多年来一直存在、却很少被正面问出的问题问了出来：科学究竟能抓住什么，抓不住什么？抓不住的那一部分，是暂时没抓住，还是这张网本来就抓不住？在机器越来越会算的今天，这个问题第一次变得人人相关。

◆ 三、一句话的主张

全书的主张可以压成一句话：

科学，是在现实里微分、在理念里积分、再回到现实里微分的一个圈。微分切掉的那一份纠缠，积分找不回来，只能在落地的那一刻从现实借回来。可计算性只属于圈的理念一侧；而 AI 能替人做积分，替不了落点上的那一次再微分。

拿看病来说。一种降压药上市之前，要挑出几千个病人做试验，把孕妇、肝肾功能不全的人、同时吃好几种药的老人排除在外，量出“平均降多少”。这是在现实里截一刀，是微分。成百上千个这样的试验，被累加成一部用药指南，覆盖“一类病人”，这是积分。可是坐在诊室里的这位老人七十八岁，肾功能偏弱，正在吃治痛风的药，常常忘了晚上那一顿。指南里没有这个人。医生把剂量减半、改成早上一次、嘱咐两周后复查，这一步就是再微分：把整体模型落回到具体的一点上。而这一步要用的东西，恰恰是当初为了截得干净而排除掉的。

圈之所以是圈，是因为它会转回来。老人吃了药，血压变了，下一次复查的数值又成了新的差异；成千上万个诊室回传的病例，又会改写下一版指南，甚至逼出新的试验，比如专门给八十岁以上合并肾病的人做一组。落地的动作反过来改写了下一轮微分的条件，SDE 把这叫作回写。回写要被观察，不能被宣布；它也是中性的，有时把底子写得更好，有时写得更坏。科学因此不是一条从数据通往真理的直线，而是一个不断回到现实、又不断被现实改写的循环。

这一句话，导读和结语里还会再出现。全书七编，都是在为它铺路、拆解、检验和划界。

◆ 四、借来的常数

书名“借来的常数”，取自中学课本里最容易被忽略的一个符号。

学不定积分时，答案后面总要添一个“+C”。老师说，别忘了加常数，忘了要扣分。很少有人追问：这个 C 是从哪里来的？道理其实很简单：求导会把常数抹掉，一个函数加上任何常数，导数都一样；所以从导数倒推回去，只能得到一族函数，挑不出“这一个”。要挑出这一个，必须另外给一个条件，比如“出发的时候在哪里”。

大白话说：手机只记下你一路上每一秒的车速，它能算出你开了多远，却算不出你此刻在哪条街。除非你告诉它，你是从家门口出发的。出发点不在速度记录里，速度记得再细也找不回来，只能从现实里另外拿。

本书把这个“+C”读作科学的一般处境。微分为了截得干净，宣布“其他条件不变”，把这一点周围的纠缠都切掉了；积分把切面拼成整体，拼得再漂亮，那切掉的一份也不会自己回来。落地的时候，它必须从现场借：物理里是初始条件和边界条件，工程里是这一块地基的土质，医疗里是这一个病人，天气预报里是今天早上全球的观测。C 在课本里是一个常数，在现实里是一整片土壤。

这个说法不只是一个比喻，它可以被检验。如果它对，科学的失败就应当较多地集中在初始条件与情境依赖处，而不是集中在推导里；越是靠“其他条件不变”切得狠的学科，复现的麻烦就应当越重。第12章把这一点展开，并写明什么样的证据会让它收回。

今天的人工智能能替人做越来越多的积分。它把千亿次梯度累加进权重，这本身就是一次规模空前的积分，只是没有写成人能读的方程。可那个常数，仍要从现场借来。副题里的“AI 的边界”，指的就是这一处：不是 AI 算得不够好，而是有一份东西本来就不在它积分过的数据里。

◆ 五、本书不押什么

一本带批判性的书，最容易被读成它没有说的样子。所以先把本书不押的几样东西说在前面。

本书不押“AI 做不到”。“某某永远做不到某某事”这类不可能性结论，现行 SDE 口径已经明确收回。本书对 AI 的判断是一个结构性的判断：凡是落地，就要从现场借回 C。AI 可以接上传感器、接上现场、接上人，参与这一步；那时承担落点的，是人与机器组成的复合体，而不是孤立的模型。这个判断也写了失败条件：如果有 AI 系统在没有现场初始条件输入的情况下，仍能跨情境稳定地完成落地决策而不出情境性失败，本书关于“再微分不可替代”的说法就必须收窄。

本书也不押 SDE 必然正确。三方程里的 F、G、H 只是关系的占位记号，不是可以代入计算的函数；123 原理是组织分析的框架，不能替代具体科学里的具体机制。如果把 S、D、E 这些词删掉以后，各章的结论一个也不改变，那就说明 SDE 在这些问题上只是在整理语言，本书会如实承认它的增量有限。

本书的主张分五类：概念规定、哲学解释、数学命题、经验判断、实践建议。五类的证明责任各不相同，不能互相豁免。哲学解释不能顶替数学证明，跨领域的相似只帮助提问，不当作证据。每一编末尾都有一节写本编怎样算错，第35章按五类再讲一遍，附录G把全书的失败条件逐条登记，编了号，便于读者一条一条去核对、去打叉。一个解释若能在任何结果出现之后都随意调换说法，它就没有经受过检验。本书宁可让自己有可能输。

◆ 六、写给谁

这本书预想过三种读者。

一种是正在学、或者曾经学过微积分的人。本书希望替你缩短底本作者那几十年：不是把微积分讲得更简单，而是让它从考试里的公式，变回它本来的样子，变回人类在现实与理念之间来回走的那条路。

一种是做工程、做医疗、做一切“应用”的人。你们站在圈上最难的那一段，在一个具体的点上替整个模型还债。本书想把你们每天在做、却很少被写进教科书的那件事说清楚。

还有一种是做 AI、或者正被 AI 改变着工作与生活的人。本书不劝你放慢，只请你多问两句：训练数据是在什么情境下截取的，切掉了什么？模型落到谁身上，谁来借回那份 C，谁来判它错了？

正文尽量讲大白话，每个观点配一个日常的例子；公式与严格论证放进方框，跳过方框不影响读懂主线。怎样按自己的兴趣挑着读，导读里有三条路线。

最后交代一件事：书中以“我”出现的个人经历，属于底本作者王德生；扩写部分多以“本书”行文。引文与史实已逐条核对，核对的结果与依据见附录I；只见于转述、未见原文的，正文都注明“据转述”。

导读

这本书怎么用

全书一句话

科学，是在现实里微分、在理念里积分、再回到现实里微分的一个圈。微分切掉的那一份纠缠，积分找不回来，只能在落地的那一刻从现实借回来。可计算性只属于圈的理念一侧；而 AI 能替人做积分，替不了落点上的那一次再微分。

这句话在导论里第一次完整提出，在第三编被拆成三段逐一论证，在第六编被拿去评估哈萨比斯的论断，在第35章被拆成一条条可以检验的主张，最后在结语里收回。

七编路线

全书按“总—分—总”展开：导论立靶，七编分论，结语收束。

第一编“根基”（第1—4章），回答科学的对象是什么。它先拆掉“先有一个世界摆在那里、再由我们去认识它”的预设，再提出全书的第一个关键区分：科学家口中的“自然界”，是物理界经数学这张筛子过滤后的理念物理界。最后用信息九宫说明，“可计算”只是筛子里面的一句话。

第二编“历史”（第5—9章），讲微积分是怎样发生的：从巴比伦的泥板、托勒密的圆，到阿基米德、刘徽、祖暅，到牛顿与莱布尼茨；再重审李约瑟之问，看无穷小这个“幽灵”怎样逼着微积分把自己重做一遍。

第三编“机理”（第10—16章），是全书的核心。微分在现实里截一刀，积分在理念里拼回整体，积分常数 C 是拼不回的那一份，基本定理把两者焊在一起又留下焊缝，再微分是技术与应用发生的地方；最后用三方程与 123 原理，说明这个圈怎样转、为什么会转不动。

第四编“扩张”（第17—21章），看微积分怎样一次次被新的差异逼着长大：从一个点到一片场，从连续到离散，从光滑到不可微，从局域到非局域，从有限维到无限维。每一次扩张都接回了一部分被筛掉的东西，也在别处切下新的一刀。

第五编“文明”（第22—27章），从“可计算不等于可预测”出发，经工程与控制论，走到 AI：它的微分在两端，被藏起来的是积分；它替得了积分，替不了落点。

第六编“评估与批判”（第28—33章），专门处理哈萨比斯的论断：核对原话，替他辩护，分三刀批判，给出判词，并写明判词怎样会错。

第七编“边界”（第34—35章），把刀口转向自己：数学是理念界的场态，不是科学本身；本书的每一个主张，也要按五类分开，逐条写明怎样算错。

三条读法

第一条，从头读。这是本书设计时预想的读法。七编环环相扣：第一编铺下筛子与信息九宫，第三编的“借　C”才有落脚处；第三编立好圈，第六编的三刀才砍得准。时间充裕的读者，按顺序读下来，收获最完整。

第二条，只关心　AI　的读者：读完导论，直接跳到第五、六编（第22—33章），再回头读第三编（第10—16章）。第五编讲 AI 在圈上的位置，第六编讲哈萨比斯；读到“借　C”“再微分”“回写”这些词心里起疑时，回到第三编去看它们是怎样立起来的。遇到“理念物理界”“信息九宫”等不熟的词，可查第3、4章。

第三条，只关心微积分史的读者：从第二编（第5—9章）读起，接着读第三编，再读第四编（第17—21章）。这条线从泥板上的天文表格一路走到无限维的函数空间，看一门数学怎样在现实的顶撞中一次次改织自己。读完可以再看第34章，那里讲数学与科学的分野。

只有一个晚上的读者，可以读导论、第12章、第14章、第33章和结语。

书里的几样东西

每一章开头，有一行“这一章的问题”，用一句大白话把本章要回答的问题先问出来，常常带着本章里的一个场景。读之前看一眼，读完再回来看一眼，看本章是否答上了。

正文从一个日常场景讲起：开车、看病、买菜、记账、做饭。每个观点都配一个这样的例子，例子是讲法上的通俗，不是思想上的稀释。

公式与严格论证放在“方框”里，以“【方框】”开头。不熟悉微积分的读者可以跳过方框，不影响主线；熟悉的读者请先看方框，因为本书的许多判断，成败就在方框里那几个条件上。

每一章末尾，有一段“带走的话”，把本章最该记住的判断用一段话说出来，口径要窄。只想抓主线的读者，可以把全书的“带走的话”连起来读一遍。

“带走的话”之后是“小账”，交代本章的证据与边界：哪些例子是构想或比喻，不对应真实个案；引文与数据出自哪里，核对到什么程度；本章不主张什么；本章的主张若被推翻，对应附录G的哪一条。

每一编末尾还有一节“这一编怎样算错”，列出几条“若……则……”：若出现什么情形，本编的哪个主张就要收回、收窄或降格。第35章把全书这些条件按概念规定、哲学解释、数学命题、经验判断、实践建议五类重讲一遍。

三个附录怎样用

附录H是 SIO 到 SDE 的术语对照表。本书底本写于 SIO 阶段，正文偶尔以“底本原作……”带出旧说法；读到旧词，或读过底本的读者想知道某个旧概念在本书里换成了什么，就查这张表。

附录G是推翻条件登记表，把各章、各编写下的失败条件汇总在一处，按编排列，编了号。读者若想检验本书，不必先接受本书的任何术语：在表里找到对应

的那一条，看它说“若出现……”的情形是否已经出现。出现了，就在那一条上打叉，并按“则……”一栏要求本书收回、收窄或降格相应的主张。

附录I是引文与史实核验表，登记全书引文与史实的核对情况：哈萨比斯言论的出处与等级，正文所据数学与实证研究的文献与使用条件，以及其他史实的核对结果。凡只见于转述、未见原文的，正文注明“据转述”，表中也写明。对某一句引文或某一个年份有疑问，先查这张表。

随身三问

读完任何一章，合上书，可以问自己三句：

这一刀切在哪里，切掉了什么？

拼出来的这个整体，缺的那个 C 要从哪里借，由谁来借？

借回来之后，现场生出了什么新的差异，谁在看着它？

目　录

作者介绍 v

凡例 viii

前言　一个常数，为什么只能借 ix

导读　这本书怎么用 xv

导论　渔夫的网：一句话把科学的边界问了出来 1

第一编　根基：科学的对象是什么 **15**

第 1 章　主客二分的困境与发生学的回答 16

第 2 章　差异先于实体：特征怎样稳下来 26

第 3 章　科学的“自然界”：理念物理界与物理界 36

第 4 章　信息九宫：可计算性是哪一格的谓词 45

第二编　历史：微积分是怎样发生的 **55**

第 5 章　黎明前的天空 56

第 6 章　古代先声：阿基米德、刘徽、祖暅 65

第 7 章　牛顿与莱布尼茨：发动机与语言 75

第 8 章　李约瑟之问再审 84

第 9 章　无穷小的幽灵与严格化 94

第三编　机理：现实的微分，理念的积分，落地的再微分 **105**

第 10 章　微分：在现实里截一刀 106

第 11 章　积分：在理念里拼回整体 116

第 12 章　积分常数 C：微分丢掉、积分找不回的那一份 125

第 13 章　基本定理：焊接与焊缝 136

第 14 章　再微分：技术与应用的发生　146
第 15 章　三方程下的微积分循环　156
第 16 章　123 原理：反常、危机与回写　165

第四编　扩张：微积分的自我更新，以及它的边　**175**
第 17 章　场：偏微分方程　176
第 18 章　数值：离散化是又一次截取　185
第 19 章　随机：拥抱不可微　194
第 20 章　记忆与非局域：分数阶与伞场微积分　204
第 21 章　无限维：量子与泛函　214

第五编　文明：混沌、工程、控制与 AI　**225**
第 22 章　混沌与复杂性：线性世界的黄昏　226
第 23 章　工程：从经验技艺到预测科学　236
第 24 章　控制论：反馈就是实时的再微分　245
第 25 章　AI 的微积分：梯度、反向传播、神经ODE、扩散模型　255
第 26 章　不成文的积分：AI 藏起了什么　264
第 27 章　AI for Science：替得了积分，替不了落点　273

第六编　评估与批判：哈萨比斯的论断　**283**
第 28 章　他到底说了什么　284
第 29 章　最强辩护　295
第 30 章　第一刀：可计算是筛子的性质　304
第 31 章　第二刀：“从未发现”是回写的战绩　313
第 32 章　第三刀：问题是谁命名的　322
第 33 章　判词与失败条件　331

第七编　边界：数学不是科学，SDE 也要接受检验　**341**
第 34 章　数学的三态与三化　342

第 35 章　本书怎样算错 351

结语　科学、微积分与文明 361

参考书目 370

附录 **381**

附录G　推翻条件登记表 382

附录H　SIO → SDE 术语对照表 388

附录I　引文与史实核验表 393

后记 402

导 论

渔夫的网：一句话把科学的边界问了出来

晚饭后刷手机，一条短视频的标题跳出来：“诺奖得主断言：没有 AI 不能解决的问题。”底下的评论分成两派。一派说，蛋白质结构几十年没人解完，AI 几年就解了，天气预报也被它抢了饭碗，这话有什么不对？另一派说，AI 连我家孩子为什么不肯写作业都解决不了，吹什么牛。两派吵得很热闹，可吵的其实不是同一件事：前一派说的是“能不能算出来”，后一派说的是“算出来的东西能不能落到我这里”。一句话，把两个不同的问题捏在了一起。

本书就从这句话起手。不是因为它说得对或不对，而是因为它把一个三百多年来一直存在、却很少被正面问出的问题问了出来：科学到底能抓住什么，抓不住什么？抓不住的那一部分，是暂时没抓住，还是这张网本来就抓不住？如果是后者，那么“算得越来越快”的机器，又能替我们走多远？

◆ 一、一句话的两个版本

先把靶子立准。

“没有 AI 不能解决的问题”这句话，并不是戴密斯·哈萨比斯（Demis Hassabis）本人的原话，而是传播中的放大版。他本人的命题要窄得多，也有意思得多。据他在 2024 年诺贝尔化学奖演讲中提出、后来在莱克斯·弗里德曼（Lex Fridman）的播客（第475期，2025 年）中再次引述的说法，他的猜想大意是：凡是自然界能够生成或能够找到的模式，都可以被一个经典的学习算法高效地发现并建模（原句见诺贝尔演讲幻灯片，访谈中的复述见该期播客文字稿）。他给的理由是：自然不是随机的，它经过了进化和物理过程长期的“筛选”，留下来的是有结构的东西，所以能被重新学会。访谈中主持人提出，也许

存在一类“可学习的自然系统”，并把 P 与 NP 的关系说成一个物理问题，他表示认同。

他自己也留了余地。在 2025 年与弗里德曼的访谈里，他承认有些问题没有可学习的结构，比如分解大数，只能硬算，那里也许需要量子计算机；他押注的是，自然界里我们关心的大多数东西因演化而带有结构，可以被经典系统（实际上就是图灵机）建模；他也承认今天的 AI 擅长渐进改良，还没有做出范式级的突破，并把“给定 1900 年的知识能否发明相对论”一类的事当作更高的检验。另据转述，他在 2025 年 12 月 Google DeepMind 的一期播客（主持人汉娜·弗莱，Hannah Fry）里说过“迄今宇宙中还没有人发现不可计算的东西”这类话；节目没有公开文字稿，本书所据是汤姆·瓦尔恰克（Walczak）2026 年 1 月的转述，详见第28章。

所以这里有两个命题，必须分开：

一是**本人命题**：自然中的模式大多可学、可建模。这是一个关于“自然有没有结构”的经验判断加一个大胆的猜想。AlphaFold 与新一代天气模型是它的真成绩，本书不打算否认。

二是**传播命题**：没有 AI 不能解决的问题。这一步从“自然的模式可学”跳到了“人的问题都可解”，中间跳过了好几道门槛：谁来决定什么算一个问题？谁来判定解对了没有？算出来的东西怎样落回到一个具体的人、一块具体的地基上？

本书第六编会专门处理哈萨比斯的论断：先把他原话的限定与语境还原（第28章），再替他把话说足（第29章），然后分三刀评估（第30—32章），最后给判词并写明这些判词在什么情况下会被推翻（第33章）。导论只做一件事：说明为什么这句话值得一本书来回答。答案藏在一位天文学家讲过的一个渔夫的故事里。

◆ 二、爱丁顿的渔夫

二十世纪上半叶，英国天文学家阿瑟·爱丁顿（Arthur Eddington）在《物理科学的哲学》（1939）里讲过一个寓言。一位鱼类学家，用一张网眼两英寸的网在海里捕鱼。捞了很多网之后，他总结出一条定律：海里没有短于两英寸的生物。有人提出异议：不是没有，是你的网漏掉了。学者回答：我的网抓不到的东西，不属于鱼类学知识的范围；抓不到的，对鱼类学来说就是不存在的。

这个寓言在不同的人手里有不同的讲法。爱丁顿本人用它来说明，物理学中的一部分“定律”其实是测量手段的性质，而不是世界的性质——网的规格写进了定律。后来的人更常用它来提醒：科学的“世界”是科学的工具捕到的世界。

放到菜市场里也一样。一个只收“能过秤的货”的摊位，账本上记着土豆、白菜、猪肉；一把葱、一句“老板送你的”、一份熟客之间的信任，都不在账上。你不能拿着这本账说“这个菜市场里没有人情”，账本只是没有记录人情的那一栏。

哈萨比斯的论断，放在这个寓言旁边，就露出了它的要害。“自然中能被找到的模式都可以被学会”——这里的“自然”，是海，还是网捞上来的鱼？如果是网捞上来的鱼，那么“网捞上来的东西都能被学会”几乎是对的，而且相当有分量：毕竟网已经把东西变成了数据，数据有结构，机器善于学结构。如果是海本身，那么这个说法就需要另外的论证，而且论证会很难，因为我们对海的了解，全部来自网。

“迄今没人发现不可计算的东西”也一样。网眼两英寸的网，当然从来没有捞上过一英寸的鱼。“从未发现”本身，可能是网的战绩，而不是海的性质。

还有一层更微妙的。真实的渔夫并不死守一张网。网眼漏鱼漏得太厉害，他会换细一点的网；细网又会拖住水草，他就改撒法、改季节。科学的网也是这样一路改织过来的：牛顿的连续量处理不了骰子，于是有了概率论；光滑的函数描述不了花粉在水里的乱跳，于是有了随机微积分；只看“此刻”的导数描述不了

带记忆的材料，于是有了分数阶导数。数学每一次碰到它网不住的东西，都会扩充自己，把原来漏掉的一部分重新网进来。

这像一个裁缝。顾客的身材总有不合尺码的地方，裁缝就不断加尺码：先有大中小，后来有加大、加加大，再后来有量体裁衣。尺码表越来越全，于是裁缝可以说：“我从来没遇到过做不出衣服的人。”这句话是真的，但它首先是裁缝不断改尺码的记录，而不是“人体天生就合尺码”的证明。“从未发现不可计算的东西”，至少有一部分是数学这样不断自我扩充的记录。第31章会把这一点展开，并且指出：已知的几个“不可计算”的数学结果，恰恰都出现在理念的无穷极限处——无穷大的晶格、连续统里的非光滑函数——也就是网织得最细、离海最远的地方。

◆ 三、网和鱼都不是现成的

到这里，读者也许会以为本书要说“科学只是人造的网，真实的海永远不可知”。不是。这种说法本身就站在另一个陷阱里：它预设了海是现成摆在那里的、网也是现成摆在那里的，二者一个在外、一个在内，中间隔着一道永远越不过的沟。这正是第1章要处理的“主客二分”。

本书采用的是 SDE 发生学的看法。SDE 是“显露（Show）—差异（Difference）—纠缠（Entanglement）”的缩写。它的最短表述是一句话：**在 E 中，经 D，成 S**。任何能被认识、被表达、被维持下来的东西，都不是预先以完整的样子摆在那里等人去看，而是在一定的纠缠土壤（E）中，经过一串差异的推进、比较、修正与收敛（D），才显露（S）为可识别、可维持、可表达的样子。

用这句话看渔夫，网和鱼都不是现成的。网是渔民在一代代捕捞中，在某片海、某种船、某种需求里，经过一次次换网眼、换材料、换撒法，才稳定成现在这张网的；“鱼”作为鱼类学的对象，也是在这张网反复打捞、反复分类、反复命名之后，才显露为一份清晰的物种清单的。网与鱼是一起发生的。

这不等于唯心论。海里有没有一英寸的小鱼，不由渔夫说了算——换一张细网，它们就会出现在船舱里，甚至会让渔夫的定律当场作废。SDE 对发生学立了两道防线（第1章细说）：一是区分“存在”与“存在形态的显露”，石头没人看也在，但“石头”在河流、在地质学家、在孩子那里显露成什么样子，各不相同；二是发生受三重约束——土壤约束（什么条件下根本发生不了）、现实顶撞（现实会顶回来）、路径收敛（试探最后要收拢到可重复的结果上）。比起“发现”，发生学更承认创造；比起“发明”，它更承认约束。

这样，渔夫的故事就从一个关于“认识的局限”的悲观寓言，变成了一个关于“认识怎样发生”的工作寓言：网是怎样织出来的？它筛掉了什么？筛掉的东西会不会从哪里顶回来？顶回来之后，网会不会被改织？改织之后，渔夫又要去哪里撒网？本书把这一整套问题，压缩到一个数学家都熟悉的词上：微积分。

◆ 四、三段式主命题

本书的主命题是：

> 科学，是在现实里微分、在理念里积分、再回到现实里微分的一个圈。微分切掉的那一份纠缠，积分找不回来，只能在落地的那一刻从现实借回来。可计算性只属于圈的理念一侧；而 AI 能替人做积分，替不了落点上的那一次再微分。

拆成三段：

段	在哪一界	做了什么	付出的代价
① 微分	现实界	在流动的现实里截取某一点的增量，“其他条件不变”	切掉这一点周围的纠缠
② 积分	理念界	把增量累加成整体模型、定律、预测	积出来的是 $f + C$，C 不知道

段	在哪一界	做了什么	付出的代价
③ 再微分	回到现实界	在一个具体的点上，用整体模型算出“此刻该怎么做”	必须把 C 从现实借回来；同时生出新的差异

用开车来讲最直白。手机里的速度表每一秒记下你的车速，这是微分：它只关心这一瞬间的变化量，路边的风景、前车的脾气、你心里的焦急都被切掉了。把一路上的车速累加起来，可以算出你开了多远，这是积分。可是光凭速度记录，算不出你现在在哪条街上——因为积分只告诉你“走了多远”，不告诉你“从哪里出发”。出发点就是那个积分常数 C。它不在速度记录里，必须从现实里另外拿：你记得自己是从家门口出发的。

导航说“前方两百米右转”，这一句就是再微分：一张整体的地图模型，落到你此刻、此地、这一辆车的这一个点上，变成一个具体的动作。地图做得再好，这一步也离不开现场——卫星定位告诉它你此刻在哪里，路口那棵挡住指示牌的树、临时封掉的半条车道，都要在现场补上。补上之后，你右转了，车流因此有了一点新的变化，下一秒的导航又要据此再算。这一点“新的变化”并不是附带的，它就是 SDE 所说的回写：落地的动作反过来改写了下一轮微分的条件。

> **为什么积分总带着一个 C** 若 $F'(x)=f(x)$，则对任意常数 C，都有 $(F(x)+C)'=f(x)$。所以由导数 f 反求原函数，只能得到一族函数 $F(x)+C$，而不是一个函数。要从这一族里挑出“这一个”，必须另外给条件，例如初值 $F(x_0)=y_0$。在物理里，这就是初始条件与边界条件；在工程里，是这一块地基的土质、这一座桥的跨度；在医疗里，是这一个病人。 这个事实本身是初等的。本书要从它引出的不是新数学，而是一个关于科学怎样落地的判断：**微分时被切掉的那份现实，积分在原则上找不回，只能在落地时重新从现实取得**。这一判断能不能成立，第12章用“初值问题与边值问题”展开，并提出检验把手：如果它对，科学的失败应当较多集中在初始条件与情境依赖处；越是靠“其他条件不变”切得狠的学科，复现的麻烦应当越重。若大规模证据显示失败与此无关，本书的这一主张就须收回。

有了这三段，前面那两派网友的争论就能拆开了。“蛋白质结构能不能算出来”问的是第②段，积分；“孩子为什么不写作业”问的是第③段，再微分，而且是在一个特定的家、特定的孩子身上。AI 在第②段上可以非常强——它把海量的梯度一步步累加进参数，这本身就是一种积分，只不过没有写成人能读的方程（第26章称之为“不成文的积分”）。但第③段要求的东西，不在它积分过的数据里：那个具体的孩子、那个具体的晚上、那份只在现场才拿得到的 C。

◆ 五、三段在一次看病里走一遍

开车的例子太顺，再换一个不那么顺的：看病。

一种降压药上市之前，要做临床试验。研究者挑出几千个病人，按年龄、病程、合并症筛一遍，把不合条件的排除掉，再随机分组，一组吃药、一组吃安慰剂，量血压，记下差值。这是第①段：在流动的现实里截下一个增量——“吃这种药，收缩压平均下降多少”。为了截得干净，试验设计刻意让“其他条件不变”：排除孕妇、排除肝肾功能不全的人、排除同时在吃另外三种药的老人。截得越干净，结论越可信；可截掉的东西也越多。

然后是第②段。成百上千个这样的试验结果，被累加成指南：哪一类病人首选哪一类药，剂量从多少起步，多少周复查。这是积分。指南是一个整体模型，它覆盖的是“一类病人”，而不是任何一个具体的病人。

第③段发生在诊室里。坐在医生面前的这位老人 78 岁，肾功能偏弱，正在吃治痛风的药，女儿说他常常忘记吃晚上那一顿，家住六楼没有电梯，夏天血压会自己往下掉。指南里没有这个人。医生要做的，是把整体模型落到这一个点上：起始剂量减半，改成早上一次，嘱咐两周后复查，夏天另说。这一步需要的东西——肾功能、合用药、生活习惯、记性、楼层——恰恰是第①段为了截得干净而排除掉的。它们没有消失，只是被留在了现场，等着在这里被借回来。

老中医所说的“手感”，老护士说的“这个病人今天不对劲”，在本书里都被看作这一步的功夫：一种没有写成方程、却在无数次落地中积累起来的东西。它不神秘，但也不容易被搬走。后文会借用一个说法，把它叫作“不成文的积

分”——这个词在第26章还会被用来描述 AI，那时我们会看到，人的手感与机器的参数，在“不成文”这一点上相似，在“落在哪里”这一点上不同。

还要补一句。医生开完药，老人吃了，血压变了，下一次复查的数值又成为新的差异；成千上万个这样的诊室回传回来的病例，又会改写下一版指南，甚至逼出新的试验设计——比如专门给八十岁以上合并肾病的人做一组试验。这就是圈：微分、积分、再微分，再微分生出的新差异又成为下一轮微分的材料。科学不是一条从数据通往真理的直线，它是一个不断回到现实、又不断被现实改写的循环。

◆ 六、这张网为什么这么好用

讲到这里，容易给人一个印象：本书要贬低筛子、贬低数学、贬低科学的抽象。恰恰相反。

渔夫的网之所以值得认真对待，是因为它非常好用。牛顿的力学，用几条微分方程就把落地的苹果和绕日的行星放进同一张网里；麦克斯韦的方程组，把电、磁、光收进四行式子；今天的天气预报，把大气切成几千万个小格子，每一格写一组方程，一步步往前推。没有这张网，人类不会有桥梁、电网、卫星、疫苗。这张网的强大，恰恰来自它的“狠”：它敢于丢掉，敢于说“其他条件不变”，敢于只看一个变量对另一个变量的变化率。不丢就抓不住，什么都想留，最后什么也留不下。

所以本书的批判对象不是网，而是两种对网的误读。一种是把网捞上来的世界当成海的全部，说“抓不到的就不存在”——这是爱丁顿笔下那位学者的毛病。另一种是相反的：因为网漏了东西，就说网捞上来的都不可信——这是另一种懒惰。本书走的是第三条路：承认网的力量，同时追问它在哪里丢了什么，丢掉的东西以什么方式回来，回来以后谁负责把它补上。

哈萨比斯的猜想，在这个意义上站在网的力量一边，而且站得相当有道理。本书只是要指出：他所说的“可学”“可建模”“没有发现不可计算的东西”，

都是对网捞上来的东西而言的；从网到海，从模型到落点，中间还有一段路，那一段路不在计算里。

“可计算性只属于圈的理念一侧”这一句，本书会格外小心地论证。在 SDE 现行的说法里，每一界——现实界、理念界、自我界——都有自己的信息，而每一界的信息又各有符号态、逻辑态、数学态三种模态，合起来叫“信息九宫”。“可计算”是一个只在信息上升到数学态、并且已被形式化的那一格里才有确切定义的词。问“物理界本身可不可计算”，就像问“海本身有没有过秤”——不是答案难找，是问题落错了层。第4章专门处理这一点，第30、31章再拿它去评估哈萨比斯。

◆ 七、本书的两条纪律

这样一个主命题，说大了很容易，说准了很难。本书给自己定两条纪律。

第一条，**每个主张先写好怎样算错**。哲学解释不能顶替数学证明，也不能顶替经验证据；跨领域的相似只帮助提问，不当作证明。本书的主张分五类：概念规定、哲学解释、数学命题、经验判断、实践建议，它们的证明责任各不相同，不能互相豁免。每一编末尾都有一节“这一编怎样算错”，第35章再把全书的失败条件按五类汇总。如果读者读完只记得一句话，本书希望是这一句：一个解释若能在任何结果出现后都随意调换说法，它就没有经受过检验。

第二条，**底本的强断言一律收窄**。本书由王德生早年的《科学与微积分：差异、整体与文明的数理内核》扩写而来。底本写于作者使用 SIO（主体—互动—客体）框架的阶段，有不少当时的大胆断言，例如把“没有微积分”说成中国没有现代科学的根因，把某种非局域的“伞模型”说成传统数学的上位理论。这些在本书里都按现行 SDE 口径（以第395卷《SDE 思想体系入门》为准）收窄：李约瑟之问里，数理解释只是候选之一，与制度、经济、地理解释并列，而且要交代它的循环风险与史实风险（第8章）；伞模型只保留一个有严格条件的局部数学命题（第20章）。底本的旧说法在需要时会以“底本原作……”带出，正文一律用现行说法；附录H 给出 SIO 与 SDE 的术语对照。

收窄不是退缩。一个主张的锋芒，来自它说清楚了自己在哪里成立、在哪里不成立；说什么都对的话，割不开任何东西。

◆ 八、七编路线

全书按“总—分—总”展开，分七编。

第一编　根基：科学的对象是什么（第1—4章）。先拆主客二分：科学研究的不是摆在那里等人去看的“实体”，而是已经稳定下来的差异（第1、2章）。再提出本书的一个基本区分：科学家所说的“自然界”，是物理界经过测量、理想化、剔除噪声之后剩下的“理念物理界”，数学是那张筛子，而反常是物理界从筛子底下顶回来的地方（第3章）。最后用信息九宫说明，“可计算”是哪一格的谓词（第4章）。

第二编　历史：微积分是怎样发生的（第5—9章）。从微积分之前的天文与力学，到阿基米德、刘徽、祖暅，到牛顿与莱布尼茨；再审李约瑟之问，最后看无穷小怎样从“幽灵”被严格化。这一编要说明：微积分本身也是在 E 中、经 D、成 S 的，它没有一个现成的诞生时刻。

第三编　机理：现实的微分，理念的积分，落地的再微分（第10—16章）。全书核心。微分是在现实里截一刀，截取即丢弃；积分是在理念里拼回整体；积分常数 C 是拼不回的那一份；基本定理把两者焊在一起，焊缝在“差异已经稳定”的地方；再微分是技术与应用发生的地方。这一编最后用三方程与 123 原理，说明这个圈怎样自己转动起来：反常累积、旧模型撑不住、新模型立起、又回写下一轮的测量与问题。

第四编　扩张：微积分的自我更新，以及它的边（第17—21章）。偏微分方程、数值方法、随机微积分、分数阶与非局域算子、无限维。每一次扩张都在把原来筛掉的东西接回一部分，也都在别处切下新的一刀。

第五编　文明：工程、控制与 AI（第22—27章）。混沌告诉我们“可计算”不等于“可预测”；工程与控制是再微分的大规模制度化；AI 把微积分推到

了一个新位置——微分在两端（训练时的梯度、应用时的落点），被藏起来的是积分。AI for Science 的真成绩与它的边界都在这里。

第六编 评估与批判：哈萨比斯的论断（第28—33章）。导论只立了靶子，这一编动手。三刀分别是：可计算是筛子的性质；“从未发现”是回写的战绩，而已知的不可计算结果恰在积分的无穷极限处；问题是谁命名的——被筛掉的残差，连“问题”的身份都拿不到。

第七编 边界：数学不是科学，SDE 也要接受检验（第34—35章与结语）。数学本身也有符号、逻辑、数学三态的往返；本书的每一个主张都要列出失败条件。最后回到文明：科学、微积分与 AI，各自在圈的哪一段上。

◆ 九、本书不主张什么

一本带批判性的书，最容易被读成它没有说的样子。先把几条容易误读的地方说在前面。

本书**不主张**“科学只是人的建构，没有客观性”。发生学恰恰要说明客观性是怎样发生的：同一个测量，换一台仪器、换一个实验室、换一批人，结果仍然收拢到一处，这种收敛就是客观性的来源。它不是从天上掉下来的，而是被一次次现实顶撞磨出来的，所以才更结实。就像一条被千万人踩出来的山路，它不是谁规定的，可你也不能随便说它不在那里。

本书**不主张**“微积分是全部科学”。底本有一些全称判断，比如“没有微积分，就没有现代科学”。本书把它收窄为：在现代物理意义上的动力学科学里，微积分是核心语言；分类学、考古学、大量的临床与田野研究，并不以微分方程为主干，它们同样是科学。微积分在本书里的位置，是一个特别清楚的样本：科学的“截取—累加—落地”在它身上表现得最干净，所以拿它来讲。

本书**不主张**“AI 永远做不到某某事”。那种“不可能性结论”在现行 SDE 口径里已经被明确收回：说某个 AI 系统“缺少现实的那一界”，只能作为针对具体配置的待检诊断，要分别检查基础模型、工具连接、记忆安排、人的参与，而不能当作一劳永逸的判决。本书对 AI 的判断是一个结构性的判断：凡是落地，就

需要从现场借回 C；AI 可以接上传感器、接上现场、接上人，从而参与这一步，但那时候承担落点的已经是“人—机器复合体”，而不是孤立的模型。这个判断同样写了失败条件：若有 AI 系统在没有现场初始条件输入的情况下，仍能跨情境稳定地完成落地决策而不出情境性失败，本书关于“再微分不可替代”的说法就必须收窄。

本书也**不主张** SDE 已经是一门完成的科学。三方程 $S = F(D, E)$、$D = G(S, E)$、$E = H(S, D)$ 里的 F、G、H，只是生成关系的占位记号，不是可以代数计算的函数；123 原理描述的是一个循环机制——差异与纠缠的矛盾累积，逼出显露的改变，显露的改变再回写差异与纠缠——它是组织分析的框架，不能替代具体科学里的具体机制。如果加入 D、E 之后，分析不能比普通的说法多出可检验的区分，那就说明 SDE 在这个问题上只是在整理语言，本书会如实承认它的增量有限。

说这些，是因为一本讨论“边界”的书，自己首先要有边界。

◆ 十、怎样读这本书

正文尽量讲大白话，严格的论证与公式放在“方框”里。不熟悉微积分的读者可以跳过方框，不影响主线；熟悉的读者请先看方框，因为本书的许多判断，成败就在方框里那几个条件上。

时间不多的读者，可以读导论、第3章、第12章、第14章、第26章、第30—33章和第35章，这是一条最短的主线：筛子、积分常数、再微分、不成文的积分、对哈萨比斯的评估、本书怎样算错。关心历史的读者从第二编读起；关心 AI 的读者可以先读第五、六编，遇到不熟的词再回到第一编。

还要交代一件事。底本序言里，作者写过自己从高中第一次接触微积分、到大学讲授微积分，用了几十年才真正明白它的价值。这是作者本人的经历，本书保留它，因为它恰好是一个关于“回写”的例子：一套公式可以在一个人身上存放几十年，熟练、正确、会考试，却一直与他看世界的方式“并联”而不是“串联”；直到某一天它与人的问题正面相撞，才真正写进了他的底盘。本书希望替

读者缩短这几十年——不是把微积分讲得更简单，而是让它从考试的公式，变回它本来的样子：人类在现实与理念之间来回走的那条路。

◆ 十一、回到那条短视频

现在再看那条短视频的标题。“没有 AI 不能解决的问题”，在本书的框架里可以拆成三问：

它说的“问题”，是网捞上来的问题，还是海里的问题？

它说的“解决”，是在理念里积出一个整体模型，还是在一个具体的点上把事情做成？

它说的“不能”，是现在不能，还是在这张网的规格下永远不能——如果是后者，又是谁织的网、谁有权改网眼？

这三问，分别对应本书的第一刀、第三段与第三刀。它们不是为了驳倒谁。哈萨比斯说出了一个值得认真对待的猜想，而这个猜想最有价值的地方，恰恰是它逼着我们把科学自己的边界说清楚：科学抓住的是哪一部分，这部分是怎样被抓住的，抓不住的那一部分又会以什么方式回来。

爱丁顿的渔夫，最后也许会换一张网。换网不是因为他想通了哲学，而是因为某一天，一条不该存在的小鱼卡在了网眼上。科学史上这样的小鱼并不少：水星近日点多出的一点点进动，黑体辐射曲线上与理论对不上的那一截。后一例常被讲成“紫外灾难逼出了量子”，其实次序要细一些：普朗克 1900 年直接回应的，是维恩公式在长波一端与测量不符；经典理论在高频处发散，是 1900—1905 年间才逐步看清的，“紫外灾难”这个名字则出自 1911 年的埃伦费斯特（第3、16章）。它们都是从筛子底下顶回来的现实。本书要讲的，就是网与鱼、筛子与顶回来的东西之间，那个一圈又一圈的发生过程——而微积分，正是人类为这个过程造出的最精密的一套语言。

下一章，从这个过程最古老的一道障碍讲起：人们习惯以为，先有一个“我”，再有一个“世界”，然后“我”去认识“世界”。

渔夫的网：一句话把科学的边界问了出来

第 一 编

根基：科学的对象是什么

编 序

第一编只做一件事：在谈微积分之前，先把科学研究的对象说清楚。

人们习惯以为，世界先摆在那里，科学家再拿仪器去看它。这一编不急着推翻这个习惯，只问它在哪里够用、在哪里开始沉默。

第1章从主客二分的困境讲起，看“主体认识客体”这张图在哪些问题上说不通，再给出发生学的回答：在 E 中，经 D，成 S。“主体看客体”只是这个过程走到稳定处时的一张截图。第2章讲差异先于实体：实验记录下来的是差异，差异稳住了才成为特征，许多特征缠在一起，才有了我们叫作“东西”的那个聚合体。第3章提出全书的第一个关键区分：科学家口中的“自然界”，是物理界经测量、理想化、剔除噪声之后剩下的理念物理界，数学是那张筛子；而水星近日点、黑体辐射这类反常，是物理界从筛孔里顶回来的地方。第4章用信息九宫给“可计算”找到它的格子：现实、理念、自我三界各有符号态、逻辑态、数学态，可计算性只对上升到数学态的那一格有确切的定义。

读这一编时，可以一直带着一个问题：科学家最后拿在手里的那个“自然界”，和筛之前的那个世界，是什么关系？

01

第 1 章

主客二分的困境与发生学的回答

> 这一章的问题：一个三岁孩子先把橙子、粉毯子都叫作“红色”，几个月后才分清——这个“红”是本来等在那里被发现的，还是在一次次纠正里长出来的？

一个三岁的孩子指着窗外喊了一声“红色”。大人很高兴：孩子认识颜色了。可这一声“红色”是怎么来的？最常见的说法是：外面有一个红色的东西，孩子有一双眼睛和一个脑子，眼睛看到了，脑子记住了大人教的名字，于是“认识”了。一边是客体，一边是主体，中间搭一座叫“感知”的桥。

可带过孩子的人都知道，事情没有这么干净。孩子第一次喊“红色”，指的可能是一辆消防车，第二次指的是一个橙子，第三次指的是妈妈的口红，第四次突然对着一块粉色的毯子也喊“红色”。大人一次次纠正：“这是橙色”“这是粉色”。几个月下来，“红”才慢慢在孩子那里站稳：它从消防车、草莓、灯笼、指甲油里被拎出来，有了边界，明天再问还在，能说出来，还能用来跟小朋友争论“这个苹果到底红不红”。

在这几个月里，到底是谁在“认识”谁？说“红色本来就在那里，孩子去发现了它”，解释不了为什么孩子先把橙子也叫成红色；说“红色是孩子脑子里的概念，被投射到了世界上”，又解释不了为什么不管孩子怎么想，消防车就是比橙子更“红”。这一章要处理的，就是这个看似简单、却卡住了几百年科学哲学的问题。

◆ 一、一个被默认的框架

日常语言里有一个很普通的词：发现。我们说哥伦布发现了美洲，牛顿发现了万有引力，孩子发现了颜色。这个词里藏着一个默认的图景：对象本来就在那里，完整、现成，等着被揭开；认识就是揭开盖子的那一下。

把这个日常用法放大成普遍法则，就得到 SDE 所说的“发现范式”。它有三根支柱，合起来是一个理想类型，用来辨认那些至今仍在起作用的默认设定：

其一，**对象先在**。对象先于认识完整地存在，研究不改变它，认识者像一个抄写员，把封存好的答案抄下来。

其二，**方法中立**。方法像一面理想的镜子，语言、概念框架、仪器都是透明的，可以随时抽离，不在结果上留下指纹。

其三，**知识累积**。知识只增不减，一步步逼近完整的真理；出了错，就用新的换掉旧的。

这三根支柱，正是“主客二分”在科学里的样子：主体在这边，客体在那边，方法是中间那块干净的玻璃。

发现范式不是凭空出现的。十七世纪，伽利略、笛卡尔、牛顿等人把目的与意义从自然里抽出去，只留下可以测量的部分；笛卡尔把世界分成“广延之物”和“思维之物”。机械的世界给了“对象先在”以根据，被隔离的心灵给了“方法中立”以根据，固定的规律给了“知识累积”以根据。这个框架是为了处理无意图、稳定、可测的对象而发生的，在这类对象上，它的成就无可替代。本书不主张抛弃它，只主张看清它在哪里够用、在哪里开始说不通。

> **发现范式的三个默认设定，与发生学的四个要素** 发现范式（理想类型）： ① 对象先在：对象 O 在认识之前已经完整确定； ② 方法中立：测量 M 作用于 O 而不改变被报告的结果，M 可以抽离； ③ 知识累积：知识集合 K_t 随时间只增不减，$K_{t+1} \supseteq K_t$，错误只能整块替换。 发生学的回应（第395卷第4章的四要素）：发生**需要土壤**（E）、**需要路径**（D）、**造就显露**（S）、**会回写土壤**。 两者的差别可以

> 用一个可检验的方式说出来：如果换一种测量安排（例如换仪器、换分类口径）而被报告的“对象”不变，发现范式在这一处够用；如果报告出的“对象”随测量安排系统地变化，就需要把测量安排算作发生的条件之一。前者与后者都在科学里大量存在，所以本书的主张是“发现范式有边界”，而不是“发现范式错了”。

◆ 二、三处说不通

底本把主客二分的困境概括为三重：认识论悖论、存在论分裂、价值虚无化。这三重说法有力，但口气偏大。按现行口径，可以改写成发现范式在三个地方说不通，每一处都对应一个具体的问题。

第一处：同一个对象，为什么被知识化成完全不同的样子？光在牛顿那里是微粒，在惠更斯和菲涅耳那里是波，在麦克斯韦那里是电磁场的振荡，在爱因斯坦那里又带上了一份份的能量子。如果光是一个先在、现成的对象，科学只是在一步步逼近它，那么这几种说法应当一个比一个更接近“光本身”。可事实是，它们各自在不同的实验安排里成立：做双缝，光显露为波；做光电效应，光显露为一份一份的。说“前人错了、后人对了”，解释不了为什么后人仍要同时保留两套说法。

用厨房来讲：同一把盐，炒菜时它是“咸味”，腌菜时它是“防腐剂”，冬天撒在路上它是“化雪的东西”，实验室里它是“氯化钠晶体”。说哪一个是“盐本身”？每一个都是盐在某一类安排里显露出来的样子。

第二处：全新的知识怎样成为可能？如果知识只是把现成的东西揭开，那么“创造”就只能是“更快地揭开”。可微积分不是被揭开的；它是在开普勒的数据、伽利略的运动学、笛卡尔的解析几何、几十年的无穷小争论里，被一点点做出来的。说它“本来就在柏拉图的天上”，等于把创造这件事整个取消了。

家常菜里也有这样的事。一道红烧肉，各家做法不同：有的先焯水，有的先煸糖色，有的最后收汁加一勺醋。说“世上本来就有一道完美的红烧肉，各家只是在逼近它”，听起来很玄；说“红烧肉是在某个灶台、某种酱油、某家人的口味里，一次次试出来、改出来的”，就顺了。试出来不等于随便：糖炒过头会发

苦，火大了会糊锅，这些都会顶回来。可在约束之内，确实长出了原本没有的东西。科学里的新知识，也是在这种“有约束的试出来”里发生的；把它说成“揭开”，丢掉的恰恰是试与改的那一整段过程。

第三处：知识为什么会死？燃素说被氧化理论取代，这还能算“替换”。可还有另一些死法，替换说解释不了：“地球是圆的”没有被推翻，却早就不再让任何人惊讶，它死于成功，变成了不再标示任何东西的常识；一门只剩最后一位传人的手艺，没有被证伪，却随着那个人的去世一起消失了。知识累积的图景里，没有给“老去”和“失传”留位置。

这三处说不通，不是靠思辨挑出来的毛病；它们在当代科学的三个现场都被顶了出来。

◆ 三、三个现场

第一个现场是**量子测量**。在量子力学里，一个粒子的某些属性，在测量之前并没有一个确定的值等着被读出；它在什么样的测量安排里被测，就显露出什么样的结果。这里要小心两点。其一，“测量安排”指的是仪器、装置与相互作用序列构成的物理条件之网，与意识无关——换一台自动记录仪，结果一样。它不是“人的意识创造了世界”的证据。其二，量子力学里的“纠缠”是一个精确的物理概念，SDE 里的“特征纠缠”含义更宽，二者只是名称巧合，不应借此占便宜。量子现场能说明的只是一点：“对象先在、测量透明”这条默认，在微观世界里不再自动成立。

第二个现场是**复杂系统**。蚁群、市场、大脑、气候，它们的秩序不在任何一个部分里，只在运行中出现，而且刻着历史：同样的部件，走过不同的路径，会长出不同的整体。钟表是组装出来的，拆开再装回去，还是那块表；蚁群是生成出来的，拆开就没有了。对生成物，“先拆成部件，再逐个研究，最后加起来”这种方法不再中立——拆的那一刀本身就改变了对象。

第三个现场是**大模型**。它在符号与逻辑层面的信息极厚，能写出流畅的论证，却常常在现实面前出错。这显示：信息的发生不等于知识的发生，知识除

了“说得通”，还要接到与任务相应的检验上、接住错误并修订下去。这条区分保留，但按现行口径，它是一个针对具体配置的待检诊断，不是对一切 AI 的判决（第26、27章展开）。

三个现场各自击穿发现范式的一根支柱：量子测量对应“对象先在”，复杂系统对应“方法中立”，大模型对应“知识累积”。但必须同时说清楚：三个现场合在一起，**不能证明**　SDE。它们只说明旧框架在这些地方需要补，至于补什么、怎么补，还要另行论证。

◆ 四、发生学的回答

SDE　的回答，是把“发现”换成“发生”。总命题是：任何可被认识、可被表达、可被维持的存在形态，都不是预先以完整样子摆在那里等人观看，而是在一定条件、一定差异路径与一定纠缠土壤中，被发生、被稳定、被表达出来的。最短的写法是：

> 在 E 中，经 D，成 S。

E　是特征纠缠，是发生的土壤；D　是差异序列，是试探、比较、修正、收敛构成的路径；S　是显露，是最后站出来、能被识别、能被维持、能被表达的那个样子。三者之间的关系，可以写成三个方程：

$$S = F(D, E), \qquad D = G(S, E), \qquad E = H(S, D).$$

这里要立刻加一条纪律：F、G、H 只是“生成关系”的占位记号，用来提醒三者互相生成、不能单独抽出一个当起点，它们**不是**可以代入数字计算的函数。要把它们变成模型，还得补上变量、映射、时间尺度、解的存在与稳定，以及经验校准。这一点在本书里会反复出现，因为本书讲的恰恰是数学——越是讲数学的书，越不能让一个占位记号冒充一个方程。

回到那个喊“红色”的孩子。E　是什么？是孩子的视觉系统、家里的东西、大人的语言、绘本、幼儿园里小朋友的说法，是一整张已经沉淀下来的网。D　是

什么？是那一串“红色？不对，这是橙色”“这个也是红色？对”的试探与纠正。S　　是什么？是几个月后，“红”终于在孩子那里显露为一个能识别、能维持、能说出来的东西。红不是现成地等在外面，也不是孩子凭空造出来的，它是在这张网里、经这一串纠正、显露出来的。

再回到科学。“温度”是不是一个现成的东西？十七世纪的温度计五花八门，刻度各不相同，有的用酒精、有的用空气，读数互相对不上；经过一百多年的比较、校准、争论，固定点从“健康人的体温”换成了“冰水混合物”和“沸水”，后来又换成更精确的定义，“温度”才稳定成今天这个可以写进教科书的量。它显露的过程有土壤（玻璃工艺、热学理论、需要测温的工业）、有路径（一次次校准与淘汰），最后才成为科学可以放心使用的 S。

发生学说出这样的话，马上会遇到两个质疑。

第一个质疑：这不是唯心论吗？不是。SDE 区分“存在”与“存在形态的显露”。河里的石头没有人看也在那里；但“石头”在河流那里显露为被冲刷的阻碍，在地质学家那里显露为某个年代的沉积岩，在孩子那里显露为打水漂的好材料——显露的样子不同，石头照样在。而且“条件之网”是实在的：仪器、身体、语言、历史，没有一样是心里想出来的。

第二个质疑：这不是说怎么建构都行吗？也不是。发生受三重约束：土壤约束——古埃及不可能有“心脏是泵”这种理解，因为没有泵；现实顶撞——炼金术士怎么努力也点不成金；路径收敛——试探最后必须收拢到可重复的结果上，否则什么也没显露。导论说发生学“比起‘发明’，更承认约束”，指的就是这三重约束。

> **“关系”与“纠缠”的分水岭**　通常说“A 与 B 有关系”，预设 A 和 B 先各自存在，然后再连一条线。SDE 说的“纠缠”方向相反：**缠绕在先，端点在后**。　以温度为例：不是先有一个现成的“温度”，再去和“水银柱高度”“冰点”“人的冷热感”建立关系；而是这几股特征在一段测温史里缠在一起、互相校准，“温度”这个端点才稳定出来。　这一区分的可检验之处在于：若某个科学概念的定义在历史上从未随测量手段与校准方式而改动，它就更接近“关系”的图景；若它的定义多次随测

> 量安排改写（温度、秒、千克都是例子），“缠绕在先”的描述就更贴切。这是一个可以逐个概念核对的历史问题，不是靠宣布就能成立的。

◆ 五、从“主体—互动—客体”到“在 E 中经 D 成 S”

底本原作用的是 SIO 框架：主体（Subject）、互动（Interaction）、客体（Object）三者不可分割，世界运行的最小单元是“差异”（底本写作 ΔSIO）。这一框架的核心洞见——互动不是事后搭上的桥，而是主体与客体得以出现的条件——在现行口径里被保留下来，但换了更精确的说法。

需要说明两点。第一，SIO 与 SDE 不能按字母对齐：SIO 的 S 是“主体”，SDE 的 S 是“显露”，两个 S 不是同一个东西。现行口径里，“主体—互动—客体”被看作显露中特别稳定成形的一种，是 S 维度里的一个分析层面，而不是与 SDE 并列的另一套本体论。第二，底本借量子力学说“不存在脱离观测的客体本身”，口气过满；按现行口径，量子现场只支持“测量安排参与属性的显露”，不支持“意识决定存在”。

底本还把主客二分的后果推到文明层面：知识的空、技术的乱、艺术的迷，最终走向意义的虚无；并用意义三律——特征律、自由律、幸福律——分别对应知识、技术、艺术。这个文明诊断有它的锋芒：它提醒我们，若把知识当作对现成客体的外在抄写，知识就会与人的问题脱节。但现行口径要求加上限定：意义三律与知识、技术、艺术的对应，只是“主要侧重”，不是严格对应；“主客二分导致意义虚无”是一个哲学解释，不是一个已被证明的历史因果。本书在后面谈技术与 AI 时，会使用这一诊断的精神，但不把它当作结论来用。

◆ 六、换一个框架，多出了哪些问题

一个新框架值不值得用，不看它的口号，要看它能不能让人问出旧框架问不出的问题，而且这些问题有办法回答。发生学在科学上至少多出四个问题。

第一问：这个对象是在什么土壤里显露的？发现范式只问“它是什么”，发生学要追问“它在哪种条件网里成了这个样子”。塞麦尔维斯在十九世纪中叶就

发现医生洗手能大幅降低产褥热的死亡率，可他的主张在当时几乎被整个医学界拒绝，他本人也郁郁而终；二十多年后，巴斯德和科赫的细菌学说立起来，“洗手”才成了常识。同一个洞见，一死一生。发现范式只能说“前人没看到真相”，发生学则能说：塞麦尔维斯的洞见落在一块没有细菌理论、没有显微镜证据链、医生地位又极其敏感的土壤里，E 太薄，长不出来。

第二问：它是沿着什么路径稳定下来的？“温度”经过一百多年校准才稳定，这条路径不是可有可无的前史，它决定了今天温度计能测什么、不能测什么。忘了路径，就会把一个有条件的量当成无条件的“本体”。

第三问：它现在处在生长、成熟还是僵化？一个理论刚提出时漏洞百出，成熟后自洽而高效，再往后可能对一切质疑都有现成答案，连一个真正的反例也接不住了。三者看起来都“有理论”，处境完全不同。就像一家馆子，刚开业时菜单天天改，生意好了菜单固定，再过十年菜单还是那张，老客人在走、新客人不来——账面上一切照旧，其实已经停了。

第四问：它会死于哪一维？被遗忘（土壤散了）、被吸收成常识（差异消失了）、还是变得容不下任何新变体（显露僵化了）。

这四个问题都可以落到具体的科学史材料上逐一核对。若在某个具体案例里，加上这四问之后得到的结论与普通的科学史叙述毫无二致，那就应当承认：在这个案例上，发生学的增量有限。

◆ 七、这对科学意味着什么

把发现范式换成发生学，对科学有一个直接的后果：科学研究的对象，不再是“摆在那里的实体”，而是**已经稳定下来的差异**。

实验仪器记下来的，从来不是“热的本体”，而是温度差；秒表记下来的，不是“时间本身”，而是两个时刻之间的间隔。科学现象之所以能被数字化、符号化，正因为它首先是差异，而差异可以被比较、被记录。但仅有差异还不够：一大堆杂乱的读数不是科学现象。读数要在一次次重复中稳定下来，被认作“同一个东西”，才成为可以写进方程的现象。

这就引出了下一章的问题：差异怎样稳定下来，变成我们称为“特征”和“实体”的东西？再往后，第3章还要追问：科学家把这些稳定下来的差异收拢成一个“自然界”，这个自然界与物理世界本身是什么关系？那是渔夫的网第一次被正面打量的地方。

◆ 八、收束

主客二分不是一个错误，而是一个被当成了普遍真理的局部工具。在无意图、稳定、可测的对象上，它几乎不出错；一旦问题变成“对象怎样被划分出来”“方法怎样塑造了结果”“知识为什么会老去”，它就开始沉默。发生学不是要推翻它，而是要把它放回它该在的位置：先有土壤、路径与显露的过程，“主体看客体”只是这个过程走到稳定处时的一张截图。

那个喊“红色”的孩子，最后会长成一个能用分光仪测波长的人。到那时，“红”又会在新的土壤里、经新的路径，显露成“620—750 纳米”这样一个样子。下一章就从这一步讲起：差异先于实体，特征是怎样稳下来的。

带 走 的 话

发现范式在无意图、稳定、可测的对象上够用；它的边界在于，一旦被报告的“对象”随测量安排系统地变化，就得把测量安排算作发生的条件。发生学的回答是“在 E 中，经 D，成 S”，但三方程里的 F、G、H 只是占位记号；量子测量、复杂系统、大模型三个现场只说明旧框架要补，不证明 SDE。发生学值不值得用，要看它能否问出普通科学史问不出、又答得了的问题。

小 账

- 本章喊“红色”的孩子、红烧肉、开了十年菜单不变的馆子，均为构想场景或比喻，无真实原型。
- 本章没有使用任何调查数字；光的学说史、温度计校准史、塞麦尔维斯一案取科学史通行叙述，未单列入附录I。
- 本章不主张“发现范式错了”，不主张量子测量与意识有关，也不主张三个现场能证明SDE；“主客二分导致意义虚无”只作哲学解释，不作历史因果。
- 本章主张的推翻条件见附录G第G8、G9条；“缠绕在先”一条与第2章共用，见第G10条。

02

第 2 章

差异先于实体：特征怎样稳下来

> 这一章的问题：温度计的水银柱先乱跳、后停住，你说“水是 62 度”——从一串变来变去的读数到一个“东西”，中间到底发生了什么？

把一支温度计放进一杯刚倒好的热水里。最开始几秒，水银柱往上蹿，忽高忽低，读不出一个准数；过了半分钟，它慢下来；再过一会儿，停在一个刻度上，不动了。你说：“水是 62 度。”

这句话里有一个不太显眼的跳跃。温度计一开始给你的，是一串不停变化的差异：比刚才高了一点、又高了一点、好像低了一点。你最后说出来的，却是一个名词性的东西：“水的温度”，仿佛它一直都在那里，只是等水银柱爬上去把它读出来。从“一串差异”到“一个东西”，中间发生了什么？

如果你再细心一点，会注意到更多：杯壁在散热，水面在蒸发，你的手握着温度计的那一端也在传热。“62 度”并不是杯子里每一个水分子的属性，单个分子没有温度可言；它是亿万个分子的乱撞在宏观上收拢出来的一个稳定读数。这一章讨论的，就是这件每天都在发生、却很少被正面看待的事：差异怎样稳下来，变成我们称为“特征”和“实体”的东西。

◆ 一、实验记录的是差异

上一章的结论是：科学研究的对象，不是摆在那里的现成实体，而是已经稳定下来的差异。这一章先把“差异”这一半讲实。

翻开任何一份实验记录，里面写的都是差异。温度计记录的是温度差（相对于冰点、相对于刚才）；秒表记录的是两个时刻之间的间隔；尺子记录的是两个刻线之间的距离；电压表记录的是两点之间的电势差。没有一台仪器能直接读出“热本身”“时间本身”“空间本身”。仪器能做的，是在两个状态之间做一次比较，把比较结果变成一个读数。

菜市场的秤也一样。秤不知道“土豆的本质”是什么，它只知道放上土豆之后，弹簧比空着的时候多压下去多少。卖菜的人每天收的钱，全部来自这一个差。

微积分的符号恰好把这一点写在了脸上：dx、dy、dt，都是“差”的写法。导数 dy/dx 是两个差之比的极限；积分 $\int f\,dx$ 是许多小差的累加。数学和实验在这里接上了头：仪器能记录的最小单元是差异，符号能操作的最小单元也是差异。实体不能被直接读出，只能通过差异显露。

但这还只是一半。一堆差异本身，还不是科学现象。温度计刚放进热水那几秒钟，读数一直在变，你什么也说不出来。必须等到它稳下来。

◆ 二、特征：差异出了稳定性

SDE 用“特征”来命名这一步：**特征是在差异序列运行中收敛出来的稳定同一性**。差异原本是流动的、来了就走的；当某些差异在反复出现中表现出一致性，并被认作“同一个”，特征就发生了。

这里有两点要说清楚。

第一，特征不是附着在一个现成事物上的属性。我们习惯说“石头有坚硬这个属性”，仿佛先有石头，再给它贴上“坚硬”的标签。发生学的次序反过来：先是一次次握、敲、踩，坚硬这股差异稳定下来；再是坚硬、沉重、灰色、冰凉这几股稳定的特征缠在一起，“石头”才作为一个东西站出来。

第二，特征天生是分模态的。“红”是在视觉差异里收敛出来的，“硬”是在触觉差异里收敛出来的，“咸”是在味觉差异里收敛出来的。每一种感官通道

都有自己的“生产线”，仪器也是这样：温度计生产温度特征，光谱仪生产波长特征。不同模态的特征各自发生，然后才有可能缠到一起。

差异是怎样稳下来的？底本给出了一个三段式的描述：混沌、自组织、涌现。初看夜空，繁星杂乱无章；长期追踪，某些星的位置开始呈现周期，这是局部的相关性；再往后，行星的轨道作为椭圆整体地显露出来。单摆刚被推动时，空气阻力、推力的偏差让轨迹乱晃；几个周期之后，它稳定在一个固定的周期上。气体里单个分子的速度是乱的，亿万个分子的平均动能却稳定成“温度”。

这个三段式是一个好用的描述，但要加两条限定。其一，“混沌—自组织—涌现”在复杂性科学里各有专门含义，这里用的是宽松的描述义，不是说每一个特征的形成都经历了严格意义上的自组织过程；温度的稳定，主要靠的是大数定律下的统计平均，这和蚁群的自组织不是同一种机制。其二，稳下来的过程不只发生在对象一侧，也发生在测量安排一侧：温度计本身的读数要稳，校准要稳，读数的人也要学会在什么时候读。这两侧是一起稳的。

> **“稳下来”在数学上是什么意思** 设一次测量给出读数序列 $x_1, x_2, \ldots, x_n$。“差异稳下来”至少可以用两种方式表达：① 时间上的收敛：读数随时间趋于一个值，$\lim_{t\to\infty} x(t) = x^*$（温度计在热水中达到平衡）；② 重复上的一致：同一安排下多次独立测量的离散度足够小，例如样本标准差 s 相对均值 $\bar{x}$ 小于某个约定阈值。大量分子的平均动能之所以稳定，背后是大数定律：独立同分布的随机变量，其样本均值几乎必然收敛到期望值。单个分子的速度没有“温度”，但 N 很大时，平均动能的相对涨落约按 $1/\sqrt{N}$ 缩小。要注意：这些数学表述描述的是“稳下来”的**结果**，它们本身并不说明是什么让差异稳了下来。说明这一点，要回到具体的物理、化学、生物机制。SDE 的“特征律”在这里只是一个组织性的名称，不能替代这些机制。

◆ 三、S 的“三个可”

差异稳下来，变成可以被指认的东西，在 SDE 里就是“成 S”，即显露。现行口径给显露加了一个检查表，叫 S 的“三个可”，缺一不可：

可识别：它从背景中站了出来，有轮廓。热水里的温度读数，终于和“刚才的乱跳”区别开了。

可维持：明天再看，它还在。同一杯水换一支温度计、过五分钟再测，读数仍在那附近。

可表达：它能被说出来、写下来、传给别人。“62 度”可以写进实验记录，可以告诉同事，同事可以拿去复核。

三个“可”是三道独立的关口。一个现象可以被识别却维持不住：某个实验室第一次看到的奇异信号，换个时间就消失了——这是科学史上大量“发现”最后被撤回的原因。一个现象可以被识别、也能维持，却表达不出来：老工匠能稳定地听出一把钢刀淬火是否到位，可他说不清自己听的是什么，徒弟只能跟着听上几年。只有三关都过了，一个特征才算作为科学现象显露出来。

现行口径还要求再追问一层：说某物“显露”了，要说明**向谁**显露、**通过什么**表现、在**哪类任务**中、稳定到**什么程度**。否则“显露”会变成一个万能词。老工匠听出的“火候”，向他本人显露了，通过耳朵与手感，在淬火这一类任务中，稳定到他几十年极少失手；但向学术共同体，它还没有显露到可以写成文献的程度。显露有三个尺度：身体与材料的实际变化、参与者能辨认什么、共同体能保存交流检验什么——三者并不保证同步。

这一点对本书很要紧。科学所说的“现象”，是在第三个尺度上显露的东西：共同体能保存、交流、检验的东西。而大量真实存在的差异，停在前两个尺度上。它们不是不存在，只是还没有以科学能接收的方式显露出来。第3章所说的“筛子”，筛的正是这一层。

◆ 四、特征纠缠聚合体：“东西”是怎样来的

单个特征还不是“东西”。“红”不是苹果，“甜”也不是。我们说“这是一个苹果”，指的是红、圆、甜、脆、某种香气、拿在手里的分量，这一大束特征总是一起出现、互相牵动。

SDE 把这样一束东西叫作**特征纠缠聚合体**：不同模态的特征在同一场中共同发生、彼此牵动，多股纠缠再聚合成一个可以当作“一个东西”的整体。世界的组成元素，在发生学看来就是这样的聚合体；我们平常说的“物体”，是这种聚合体被稳定显露出来的表征。

底本对此有一个很直观的描述：要让意识说出“这是一个东西”，至少要三个条件同时具备——**整体性**（各个特征被体验为一个统一体）、**稳定性**（它们在不同时间和场景下持续一致）、**独立性**（这一束能与别的束区分开，有边界）。石头在河里、在地上、白天黑夜都坚硬沉重，又和木块、和水分得开，于是成了“一个东西”。现行口径保留了这三个判据，并把它们用在更广的地方：不只物体，一个概念、一个身份、一个口号，只要它在一个人的经验里同时满足整体、稳定、独立，就会像一颗“粒子”一样被当作一个东西来用。

聚合体有两个容易被忽略的性质。

一是**厚薄随境况而变**。一棵你只远远看过的树，和一棵你爬过、摸过树皮、闻过开花时气味的树，在你那里是厚薄不同的两个聚合体——后者多了几股触觉、嗅觉的纠缠。科学对象也一样：一个只在教科书里见过的“电子”，和一个在实验室里调了十年电子束的人心中的“电子”，厚度差得很远。

二是**缠绕在先，端点在后**。上一章已经说过：不是先有“苹果”这个端点，再去连上红、甜、脆；而是红、甜、脆在一次次吃苹果的经验里缠到一起，“苹果”这个端点才稳定出来。把这个次序弄反，就会以为“苹果本身”是一个藏在所有特征背后的神秘核心——这正是哲学史上“物自体”一类说法的来路。

◆ 五、名是指针

聚合体稳定下来，人们给它起名字。“苹果”“电子”“温度”。名字是怎样和聚合体连在一起的?

SDE 的回答是：**名是指针**。名与聚合体之间是“指向”关系，不是包含，也不是等于。“苹果”这两个字里没有甜，也没有脆，它只是指向那一束纠缠的一个记号。

这条看似平常的说法，有几个不平常的后果。

第一，名字可以比聚合体活得久，也可以比它死得早。“燃素”这个名字还在化学史书里，它指向的那一束纠缠早已被拆散重组成了“氧化”；而某种只有一位老工匠能稳定识别的火候，可能从来没有一个通行的名字，他去世后，聚合体本身也散了。

第二，同一个名字可以在不同时期指向厚薄不同的聚合体。“原子”在德谟克利特那里指向一个思辨的“不可分者”，在道尔顿那里指向化学反应中的定比关系，在卢瑟福之后指向一个有核、有电子的结构，在量子力学之后又带上了波函数。名字没变，指向的东西已经换了几轮。把名字当成东西本身，就会误以为科学一直在研究“同一个原子”，只是越看越清楚；把名字当成指针，就能看出：每一轮，都是在新的土壤里、经新的路径，显露出了一个新的聚合体。

第三，也是对本书最要紧的一点：**名字可以被拿走去单独运算**。一旦“温度”有了名字、有了符号 T，它就可以脱离那杯热水、那支温度计，进入方程，和别的符号一起被推导。这是数学化的巨大力量：指针可以在纸上被操作，比在现实里拖着整束纠缠走轻便得多。但代价是：纸上被操作的是指针，不是聚合体本身。方程里的 T 只保留了“温度”这一束纠缠里被测量、被定义的那一股；你握着杯子时那份烫手的感觉、杯壁上凝的水珠、水面上那层若有若无的油膜，都没有跟着指针一起进入方程。

指针、聚合体与方程　设某个现象 P 在现实中显露为一束特征纠缠 $\{c_1, c_2, \ldots, c_k, \ldots\}$。科学为它定义一个符号 T，并规定一个测量程序 M，使 $T = M(P)$。这一步做了两件事：① 从整束纠缠中只取出可被 M 读出的那几股；② 给取出的部分一个可以进入运算的名字。之后的推导，例如 $\frac{dT}{dt} = -k\,(T - T_{\text{env}})$（牛顿冷却定律），全部在 T 上进行。这条定律在相当多的情形下工作得很好。但它是否适用，取决于被丢掉的那些股纠缠（对流、蒸发、容器形状……）在具体情形里是否真的可以忽略。判断“可以忽略”的那一步，不在方程里，而在现场。这是第12章“积分常数 C 的外借”与第14章“再微分”要展开的话题的种子。

◆ 六、三界平权

名是指针，指向的聚合体在哪里？SDE 的回答是三界：现实界的指针指向感觉特征的聚合（树、石头、热水），理念界的指针指向信息特征的聚合（素数、函数、能量守恒），自我界的指针指向内感与时间连续性的聚合（“我”、我的记忆、我的承诺）。

三界**平权**。这句话针对的是三个历史上常被当作“最终底座”的东西：现实界背后的“物自体”、理念界背后的“理念共相”、自我界背后的“实体性自我”。在发生学里，这三者都不再享有特权：它们都是特征纠缠聚合体在某一界里显露出的样子，都在 E 中、经 D、成 S，也都可以老去、散掉、重组。

对科学来说，三界平权的意思是：素数和石头、“我”和温度，不是一个比另一个更“实在”，而是在不同的界里、以不同的方式显露。理念界的东西并不因为看不见摸不着就不实在——“二是素数”比你家门口那块石头稳定得多；现实界的东西也不因为摸得着就自动进入科学——它要先被测量、被命名、被写成理念界的指针，才能进入方程。科学，恰恰活在现实界与理念界之间的那条通道上。

◆ 七、各门科学里的“稳下来”

底本用五个领域印证差异的稳定化，这些例子仍然好用，只是要把每一个的机制说清楚，不让一个笼统的名称盖过它们之间的区别。

在**物理学**里，单摆的周期是典型。推动时的偏差在几个周期后被阻尼与回复力磨平，留下一个只取决于摆长与重力加速度的周期——这是动力学系统趋于稳定轨道的结果。气体的温度则是另一回事：它靠的是大量分子的统计平均，单个分子从头到尾都在乱撞，稳定的只是平均值。

在**天文学**里，行星的逆行曾令古人困惑。开普勒从第谷留下的大量观测数据中，整理出火星轨道是椭圆。这里稳下来的不是行星本身的运动——它一直就那

样走——而是人类的描述：观测精度、坐标系、计算方法一起稳定，椭圆这一特征才显露出来。

在**生物学**里，神经细胞膜上每个离子通道随机开合，细胞整体却维持着一个大约 −70 毫伏的静息电位（具体数值因细胞类型而异）。这是离子泵与通道在动态平衡中维持的稳态，是一个不停消耗能量才保住的“稳”。种群层面，个体生死难测，种群数量却可能沿着逻辑斯谛曲线趋于环境容量——前提是环境本身相对稳定，这个前提一变，曲线就不再成立。

在**心理学**里，光线、角度、距离一直在变，我们却总能认出“这是同一张脸”。这是知觉系统的稳定化，它常常稳得过头：人会在云彩、插座、汽车前脸上看出脸来。

在**经济学**里，每一笔交易价格各不相同，市场上却会形成相对稳定的价格水平。但这种稳很脆：一则消息、一次恐慌，就能让它在一天之内跌穿。

把这几个例子放在一起看，可以得出比底本更细的一个结论：**“稳下来”有多种机制**——阻尼、统计平均、动态平衡、知觉整合、市场调节——它们各自可以被研究、被检验，也各自有失效的条件。说它们都是“特征律在运行”，只是给它们一个共同的名字；真正的科学工作，在于弄清每一种“稳”靠什么维持、在什么情况下会垮。家里的一碗汤放在桌上会慢慢凉到室温，这是一种稳；一锅一直开着小火的汤保持温热，是另一种稳——前者不需要管，后者一关火就没了。把两者都叫作“稳定”没错，但做饭的人必须知道自己面对的是哪一种。

这一区分，正是后面讲“复现危机”时（第12章）的一把钥匙：一个只在特定条件的动态平衡中才稳住的特征，一旦被当作无条件的常量写进理论，换个实验室就可能复现不出来。

◆ 八、收束：从“实体”到“稳定的差异”

把这一章收拢起来，是一条从差异到科学对象的链：

差异（仪器能读出的比较结果）→ 特征（差异出了稳定性，挂在某个模态上）→ 特征纠缠聚合体（多股特征缠成一个可当作“东西”的整体）→ 名（指向聚合体的指针）→ 科学对象（能被共同体识别、维持、表达的那一部分）。

底本把这条链写成“ΔSIO → 特征律 → 意识同一性 → 科学现象 → 微积分”，并断言“没有特征律，就没有科学现象；没有微积分，就没有科学理论”。现行口径保留这条链的方向，但收窄两处：“特征律”是对差异稳定化的组织性名称，不替代具体机制；“没有微积分就没有科学理论”只在动力学科学里成立，分类学、地质学、大量的临床研究，并不以微分方程为主干。

这条链的每一环都在丢东西。特征只取了某一个模态；聚合体有厚有薄；指针只保留了被测量的那几股；科学对象只收下了第三个尺度上显露的部分。丢东西不是缺点，是代价——不丢就拿不起来。问题是：科学家最后拿在手里的那个“自然界”，和丢之前的那个世界，是什么关系？下一章，把这张网正面摊开来看。

带 走 的 话

科学对象不是现成实体，而是稳定下来的差异：差异稳成特征，特征缠成聚合体，聚合体得名，名是指针。“稳下来”有阻尼、统计平均、动态平衡等多种机制，“特征律”只是它们的共同名字，不替代其中任何一种。指针一旦进入方程，只带走被测量的那一股；判断丢掉的部分“可以忽略”，这一步不在方程里，而在现场。

小 账

- 本章热水与温度计、听火候的老工匠、桌上凉汤与小火温汤，均为构想场景或日常比喻，无真实原型。
- 本章所举静息电位约 −70 毫伏、逻辑斯谛曲线、牛顿冷却定律、样本均值的大数定律，取教科书通行表述，具体数值因对象而异；本章没有使用任何调查数字。
- 本章不主张“特征律”是一种物理机制，也不主张“没有微积分就没有科学理论”——后者只在动力学科学里成立。
- 本章主张的推翻条件见附录G第G10条（与第1章共用）。

03

第 3 章

科学的“自然界”：理念物理界与物理界

> 这一章的问题：一个只见过过筛面粉的人写出一本《面粉学》，条条定律都对，却还是错了——科学家说的“自然界”，是不是也只是过了筛的那一层？

厨房里筛面粉。一勺面粉倒进细筛，轻轻拍几下，落下去的是细细的一层，雪一样匀；留在筛网上的，是几粒结块、一点麸皮，偶尔还有一只不知什么时候钻进去的小虫。做蛋糕的人要的是落下去的那一层。筛子的好处是明显的：没有它，蛋糕会结疙瘩。

可如果有人只看过落下去的面粉，从来没见过筛网上留下的东西，他会以为面粉天生就这么细、这么匀、这么干净。他会写一本《面粉学》，里面所有的定律都对，因为它们说的都是过了筛的面粉；只有一点不对：他以为自己写的是“面粉本身”。

再追问一步。如果某天做出来的蛋糕吃出一股怪味，查来查去，原来是筛网破了一个洞，漏下去一点没筛干净的东西——这时候，被筛掉的世界从筛子底下顶了回来。《面粉学》要么修补，要么改写。

本章要说的就是这三件事：科学家所说的“自然界”，是过了筛的那一层；数学是那张筛子；而反常，是被筛掉的东西从筛子底下顶回来的地方。

◆ 一、两个“自然界”

“自然界”这个词，在日常语言和科学语言里指的不是同一个东西。

日常所说的自然界，是我们身处其中的这个世界：下雨、刮风、花开、人老，它们彼此缠在一起，没有一件是单独发生的。今天下午的这场雨，和早上的气温、远处的山、楼下那棵树的蒸腾、你忘了收的衣服，全都牵连着。

科学家在论文里所说的自然界，是另一回事。它是一组经过测量、理想化、剔除噪声之后留下的对象：质点、刚体、理想气体、无摩擦的斜面、不可压缩的流体、处于热平衡的系统。这些对象在现实里一个也找不到——没有哪个物体真是一个点，没有哪块斜面真无摩擦——但科学的定律，恰恰是关于它们的。

本书把这两个“自然界”分开命名：前者叫**物理界**，指现实界中那个未经筛选、处处纠缠的世界；后者叫**理念物理界**，指物理界经过测量、理想化、去噪之后，以数学形式保存在理念界里的那一部分。科学定律直接描述的，是理念物理界；它对物理界有效，是经由一道道“落地”的手续才实现的。

要注意，理念物理界并不是“物理界的一个子集”，好像从一筐苹果里挑出几个好的那样。质点不是物理界里某个特别小的物体，它是一个在物理界里不存在、只在理念界里才有的对象；它和物理界的关系，是“代表”而不是“包含”。这正呼应上一章的“名是指针”：理念物理界里的每个对象，都是指向物理界中某一束纠缠的一根指针，只是这根指针被磨得极其光滑，光滑到可以放进方程里滑来滑去。

这不是贬低科学。一张城市地图上没有每一棵树、每一块砖，正因为没有，它才能用来找路。理念物理界之于物理界，就像地图之于城市：它是一次有意识的、高度成功的简化。问题只在于，不要把地图当成城市。

◆ 二、筛子论：数学化是一次过滤

从物理界到理念物理界，中间要过三道手续。

第一道，测量。只有能被仪器读出的差异才能进入。上一章说过，仪器读出的是某一个模态上的比较结果：长度、时间、质量、温度。一朵花的颜色可以测成波长，它的香气可以测成挥发物的成分谱，但“这朵花让我想起外婆家的院子”这一股纠缠，在测量这一关就被留在了筛网上。

第二道，理想化。读出来的数总是带着杂质：斜面有摩擦，空气有阻力，桌子在微微震动。科学家的办法是把它们设想为可以减到零的东西：令摩擦系数趋于零，就得到伽利略的理想斜面；令物体大小趋于零，就得到牛顿的质点。这一步是一个极限动作，和微积分里的 $\Delta x \to 0$ 是同一种思路。

第三道，去噪。多次测量的结果总有起伏。科学家用平均、拟合、误差分析，把起伏当作“噪声”分离出去，留下“信号”。至于这些起伏里是否藏着别的东西，在去噪这一关通常不问。

三道手续合起来，就是数学化。数学化不是把物理界“翻译”成数学，而是一次**过滤**：它只让能被测量、能被理想化、能与噪声分开的那部分通过。数学是那张筛子。

这张筛子的筛孔有一个共同的形状：**“其他条件不变”**。研究落体，假设空气阻力不变或可忽略；研究药效，假设病人的其他状况不变；研究价格，假设收入、偏好、替代品都不变。这句话的作用，就是把要研究的那一点从它周围的纠缠里切出来。切得越干净，方程越漂亮；切掉的东西，也就越多。

> **理想化作为极限，筛孔作为假设**　以斜面上的物体为例。真实的运动方程大致是 $m\ddot{x} = mg\sin\theta - \mu mg\cos\theta - \gamma\dot{x} + \xi(t)$，其中 μ 是摩擦系数，γ 是空气阻力系数，$\xi(t)$ 是各种扰动。伽利略的理想斜面相当于令 $\mu \to 0$，$\gamma \to 0$，$\xi \equiv 0$，得到 $\ddot{x} = g\sin\theta$，加速度为常数，位移 $x = \frac{1}{2}g\sin\theta\, t^2$。这个结果在理念物理界里是精确的；在物理界里，它的好坏取决于被置零的那些项实际上有多大。μ、γ、ξ 不是从世界里消失了，而是被留在了筛网上——这就是“筛子”一词在本书里的严格含义：**一组被宣布为可忽略的项，加上把它们置零的极限动作**。失败条件：若某一领域的定律在推导时没有用到任何“可忽略”的假设或极限动作，那么在该领域里，“数学化是过滤”的说法不成立，应当收回。

◆ 三、伽利略的书，胡塞尔的批评

“自然这本大书是用数学语言写成的。”这是后人对伽利略 1623 年《试金者》里一段话的通行概括。原文说的是哲学写在宇宙这本大书上，大意是：这本

书摊开在我们眼前，但只有懂得它的文字的人才能读懂，而它是用数学的语言写的，文字是三角形、圆和其他几何图形。这句话常被看作现代科学的宣言。

从筛子论看，这句话有两种读法。

一种读法是：自然本身就是数学的，数学只是把它本来的样子读了出来。这是发现范式的读法。

另一种读法是：只有被写成数学的那一部分自然，才能被这本书收录。翻开《自然之书》，你读到的每一页都是数学，因为不是数学的东西没有被印进去。这是筛子论的读法。

二十世纪的哲学家胡塞尔在晚年的《欧洲科学的危机与超越论的现象学》（写于 1930 年代）里，对伽利略做过一次有名的诊断。在该书第9节讨论伽利略把自然数学化的那一长节里，他说伽利略是“发现的天才，同时也是遮蔽的天才”（德文作 entdeckender und verdeckender Genius）。伽利略把自然数学化，开辟了现代科学的道路；但在这个过程中，人们渐渐忘了数学化的自然是从哪里来的。胡塞尔把那个源头叫作“生活世界”：我们直接感知、直接生活在其中的世界，有颜色、冷暖、气味、意义的世界。数学化的自然是从生活世界里经过理想化、测量、构造而得来的，就像在生活世界外面披上一件“理念的外衣”；可后来人们把这件外衣当成了真正的存在，反过来把生活世界看成主观的、次要的表象。

胡塞尔的诊断与本书的筛子论有相当深的共鸣：理念物理界是从物理界筛出来的，不能把筛下来的当成全部。本书接住他这一点。

但本书在三处与他不同，也算是对他的补齐。

第一，胡塞尔的“生活世界”主要是一个意识与意义的世界，他走的是回到主体经验的路。本书所说的物理界，不只是被体验到的世界，它包括了仪器、材料、身体、地质这些在现实界里实实在在起作用的纠缠。被筛掉的不只是“意义”，还有摩擦、扰动、初始条件这些硬邦邦的物理项。

第二，胡塞尔把数学化描述为一次历史性的“遗忘”，口气偏于哀叹。本书把它描述为一个循环：筛掉的东西并没有被永久遗忘，它会以反常的形式顶回来，逼着筛子改织。遗忘只是这个循环停在某一刻时的样子。

第三，胡塞尔要人回到生活世界去为科学“奠基”。本书不认为有一个可以一劳永逸地回去的基础。物理界不是一块静止的地基，它是不断与理念物理界来回发生的那一头。科学的落地不是一次性的奠基，而是一次次的“再微分”（第14章）。

◆ 四、反常：回流通道

如果数学是一张筛子，被筛掉的东西就永远被挡在外面了吗？没有。科学史上最激动人心的时刻，往往是被筛掉的东西从筛子底下顶回来的时刻。本书把这种时刻叫作**回流通道**，它们在科学里的名字是：**反常**。

第一个例子是**水星的近日点**。水星绕太阳的椭圆轨道并不完全闭合，它的近日点会缓慢地往前移动（进动）。用牛顿引力计算其他行星的扰动，可以解释大部分进动，但总还剩下一小截对不上：每世纪 40 多角秒。1859 年，法国天文学家勒维耶（Urbain Le Verrier）首先指出这一差额，当时算得约 38 角秒；后来美国天文学家纽康（Simon Newcomb）重算，修正为约 43 角秒，这是今天通常引用的数值。这一截差额，在牛顿力学的筛子下是“噪声”还是“信号”？

有意思的是，勒维耶此前处理过一个几乎一模一样的问题。天王星的轨道也有对不上的地方，他假设外面还有一颗未知行星在拉它，算出了位置，1846 年海王星果然在那附近被找到了。这一次，反常在原来的筛子之内被吸收了：筛子不用改，只是补上一个漏算的对象。于是对于水星，他如法炮制，假设在水星轨道内侧还有一颗行星，甚至给它起了名字“火神星”（Vulcan）。可这颗行星始终没有被可靠地找到。直到 1915 年，爱因斯坦用广义相对论算出，水星那一截多出来的进动正好是时空弯曲的效应。这一次，筛子本身被改织了。

同样是反常，一次在旧筛子里被吸收，一次逼出了新筛子。这个对比很重要：**反常本身不告诉你它属于哪一种**。科学家面对反常时，先试着在原来的筛子

里吸收它——这是合理的，大多数反常最后确实是测量误差或漏算的因素；只有当吸收的努力一次次失败，筛子才会被改织。

第二个例子是**黑体辐射**。十九世纪末，物理学家想算出一个理想的“黑体”在各个频率上辐射多少能量。按照经典理论推出的公式（即后来的瑞利—金斯公式），在长波（低频）一端与实验符合得不错，可到了短波（高频）一端，理论预测的能量越来越大，趋于无穷——后来被称为“紫外灾难”。实验测得的曲线却在高频处掉了下来。这是事后回看的整理，历史的次序要细一些：普朗克 1900 年直接回应的，是当时通行的维恩公式在长波一端与新测量不符；经典理论在高频处发散，到 1900—1905 年间才逐步清楚，“紫外灾难”这个名字则出自 1911 年的埃伦费斯特（第16章还会细说）。1900 年，普朗克为了得到与实验相符的公式，做了一个他自己当时也不太情愿的假设：能量只能一份一份地交换。这一份一份的“量子”，在经典的连续性筛子里是被筛掉的东西，它从曲线与理论对不上的地方顶了回来，最后长成了一整门量子力学。

> **一个反常怎样从“噪声”变成“信号”** 设理论预测为 $y_{\text{th}}(x)$，测量值为 $y_{\text{obs}}(x)$，残差 $r(x) = y_{\text{obs}}(x) - y_{\text{th}}(x)$。 在筛子论里，残差有三种去向： ① 被判为随机噪声：r 在误差范围内无规律起伏，不处理； ② 在原筛子内被吸收：补上一个漏算的项（海王星），y_{th} 改为 $y_{\text{th}} + \delta y$，筛孔形状不变； ③ 逼出筛子改织：r 有稳定的结构（在某一区段系统偏离，如黑体辐射曲线的一端），且多次在原框架内补项都失败，最后改写理论的基本假设。 判断走哪一条，不能在数学内部完成：同一个 r，是测量误差、漏算对象还是新物理，要靠更多观测、更好仪器、独立实验来定。这正是本书说“反常是回流通道”的意思——回流的东西来自筛子之外。 失败条件：若能找到一类领域，其理论改写从不由与观测的系统残差触发，而纯由内部形式考虑推动且事后从未遇到经验修正，则“反常是主要回流通道”的说法须收窄为“回流通道之一”。

◆ 五、回流之后：筛子会改织

反常顶回来之后，发生了什么？用 SDE 的语言说，这正是 123 原理的一次完整运转。

① **差异与纠缠的矛盾累积**。水星那一截进动、黑体那一段对不上的曲线，是新的差异（D）；牛顿力学、经典电磁学和热力学，是已经沉淀下来、有很大惯性的土壤（E）。新差异一再出现，旧土壤一再吸收不了，张力越积越大。

② **矛盾逼出显露的改变**。广义相对论、量子假设，是理论层面新的显露（S）：一套新的、可识别、可维持、可表达的理论形态。

③ **新显露回写差异与纠缠**。新理论不只是多了一条定律。它重新规定了“什么算差异、往哪里看”：相对论之后，引力红移、光线偏折、引力波都成了值得测的东西；量子力学之后，原子光谱、半导体、激光都进了视野。它也沉淀成新的土壤：新的仪器、新的教科书、新的一代人的直觉。新的筛子织好了，下一轮的反常又会在新筛子的边缘出现。

这一段要格外强调两点纪律。一是回写要被观察，不能被宣布：说“相对论回写了物理学”，要能指出具体的痕迹——教学大纲的改动、新仪器的设计、新实验的立项，而不能只是一句口号。二是回写是中性的：新筛子不一定比旧筛子“更接近海”，它只是在新的地方筛，筛掉的东西换了一批。量子力学在接回了“一份一份”之后，也在别处切下了新的一刀——比如它对单次测量结果只给概率，这在某些人看来本身就是一个新的残差。

这样看，筛子不是一张死网，而是一张一直在被改织的网。科学的历史，在很大程度上就是这张网被一次次改织的历史。

◆ 六、筛子之外的东西不都会回流

讲了两个回流成功的故事，容易让人以为：只要耐心等，被筛掉的东西迟早会顶回来。这是一种过于乐观的想象，需要泼一点冷水。

有一个老笑话：深夜，一个人在路灯下弯着腰找东西。路人问他丢了什么，他说钥匙。路人帮他找了半天没找着，问：“你确定是丢在这儿的？”他说：“不是，丢在那边草丛里了，可这儿有灯。”

科学的筛子也带着一盏路灯。能被测量的地方亮，测不到的地方黑。反常要顶回来，前提是它出现在灯光照得到的地方：水星的进动能被望远镜一年年记下来，黑体辐射能在实验室里一段段测出来。如果被筛掉的东西恰好落在仪器够不着的地方、落在没有人想到去测的变量上，它就不会以“残差”的形式出现——它连残差都不是，因为根本没有对应的那一栏。

这类情形在医学和社会科学里尤其常见。一种药的试验只登记了血压和几项化验指标，病人吃药后夜里睡得更差、白天更容易摔跤，如果表格里没有“睡眠”和“跌倒”这两栏，这些变化就不会作为残差出现在任何一张图上；它们要等到药物大规模使用、跌倒骨折的老人多到足以被另一套统计系统注意到，才有机会回流。即使回流，也常常走了很长的弯路。

所以回流通道有三个条件：被筛掉的东西要在某处留下可测量的痕迹；有人在那里设了观测；观测到的偏差足够稳定，大到不能被当作噪声打发掉。三者缺一，回流就不发生，被筛掉的东西就一直待在筛网上，甚至被扫进垃圾桶。第32章会把这一点推到最尖锐处：被筛掉的残差，有时连“问题”的身份都拿不到，因为它从来没有被命名过。

这也提示了一个与后文相连的检验方向。如果筛子论对，那么在“其他条件不变”切得最狠、而测量栏目又最窄的领域，没有回流的残差应当最多，它们会以“换一批人、换一个地方就做不出来”的形式表现出来——这就是近年在心理学、医学等领域引起广泛讨论的复现问题。第12章会把这一点当作“积分常数外借”的检验把手正式提出来；这里先记下：筛子的问题，不只是筛掉了什么，还在于筛掉的东西有没有路回来。

◆ 七、收束：科学抓住的是什么

回到筛面粉。筛子不是骗人的东西，没有它做不出好蛋糕；筛网上留下的东西也不是垃圾，有时候里面有一只能让你改掉整个配方的虫子。科学抓住的，是过了筛的那一层——理念物理界；它对物理界有效，靠的是两件事：一是筛孔开

得恰当，被丢掉的项在那类情境里确实可以忽略；二是筛子留着回流通道，反常能顶回来、能被认真对待。

这给后面各章留下了几个问题。如果科学的“自然界”是过了筛的那一层，那么“可计算”这个词，是在说过了筛的这一层，还是在说筛子之外的物理界？第4章用信息九宫回答：可计算性是只在信息上升到数学态的那一格里才有确切定义的谓词。微分是在筛孔处切下的那一刀，积分是把切下来的东西重新拼成整体，拼不回的那一份——筛网上留下的东西——要在落地时从现实借回来，这就是第三编的主题。至于哈萨比斯说的“自然中能被找到的模式都可以被学会”，第30章会问：这里的“自然”，是海，还是网捞上来的鱼？

下一章，先把“可计算”这个词放进它该在的格子里。

带 走 的 话

科学定律直接描述的是理念物理界：物理界经测量、理想化、去噪三道手续之后留下的那一层，筛孔的形状就是“其他条件不变”。被筛掉的项没有消失，可能以反常的形式顶回来；但反常本身不说明自己是误差、漏算还是新物理，而且只有留下可测痕迹、有人设了观测、偏差稳定到不能当噪声打发时，回流才会发生。

小 账

- 本章筛面粉与《面粉学》、路灯下找钥匙是比喻与笑话；降压药试验漏登“睡眠”“跌倒”一段是构想情形，不指任何具体药物。
- 本章胡塞尔引文的出处等级见附录I第I18条；伽利略一句是通行概括，正文已交代原文大意，附录I未单列；水星近日点进动的两个数值（勒维耶约 38 角秒、纽康约 43 角秒）与海王星的发现见附录I第I53条；黑体辐射的前史取物理学史通行叙述。
- 本章不主张数学化是错误或遮眼法，也不主张改织后的新筛子一定“更接近海”；回写是中性的，回流也不是迟早必然发生的。
- 本章主张的推翻条件见附录G第G11、G12条。

04

第 4 章

信息九宫：可计算性是哪一格的谓词

> 这一章的问题：老木匠敲三下就知道板子里有暗裂，却说这本事“写不成程序”——问一样本事能不能算之前，是不是得先问它住在哪一格？

县城老街上有一家开了四十年的木器铺。老师傅挑料，从不带仪器：一块松木拎起来掂一掂，指节在板面上敲三下，听声音发闷还是发脆，再凑近闻一闻，就说“这块心里有暗裂，做桌面不行，做凳腿还凑合”。十有八九是对的。他儿子从职校学了数控，铺子里添了一台雕刻机。机器能把一张图纸刻得分毫不差，可它不会挑料——你给它一块有暗裂的板，它照刻不误，刻完一上漆，裂纹就出来了。

儿子有一回半开玩笑地说：“爸，您那套本事，要是能写成程序就好了。”老师傅摆摆手：“写不出来，这得手上有数。”

这句“手上有数”，值得停下来想一想。老师傅心里明明有“数”，有轻重、有先后、有“闷声加上沉手就是暗裂”的推断，可这个“数”偏偏进不了雕刻机。不是机器笨，也不是老师傅糊涂，而是他的“数”和机器能吃的“数”，不在同一个格子里。

本章要做的，就是把这些格子画出来。上一章说，科学家口中的“自然界”，是物理界经过数学这张筛子筛过之后留下的部分（见第3章）。这一章再往前一步，问：筛子筛的是什么？筛出来的东西放在哪里？“能不能算”这个问题，到底是对哪一种东西提的？

◆ 一、每一界都有信息

先从一个容易被忽略的前提说起：信息不是只存在于书本和电脑里。

第395卷把人所面对、所参与的世界分为三界。**现实界**：身体、器具、土地、木头、仪器。**理念界**：概念、规则、制度、定理、语法。**自我界**：记忆、身份、情绪、意志。早先的写法把信息只放在理念界内部，好像现实界只有“物”，自我界只有“感受”，信息是理念界的专利。现行口径纠正了这一点：**每一界都有信息**。

老师傅掂木头、敲木头时，木头确实在“说话”：重量、回声、气味，是现实界的信息。他心里“这块不对劲”的那一下警觉，是自我界的信息。他将来要是带徒弟，说出“声闷手沉，必有暗裂”这八个字，这八个字进了理念界。三份信息说的是同一块木头，却住在三个地方。

大白话的例子到处都是。你开车，方向盘传来轻微抖动，这是现实界的信息；你心里一紧，“是不是胎压不对”，这是自我界的信息；修车师傅告诉你“四轮定位偏了会让方向盘在某个车速区间抖”，这是理念界的信息。三者可以一致，也可以打架——你心里慌，车其实没事；或者你心里不慌，车已经快出事了。

“每一界都有信息”这句话的分量在于：它不允许把理念界当成信息的唯一合法住处。这一点对本书很要紧。后面我们会反复说，微积分是在理念界里拼回整体的；如果信息只在理念界，那么“理念界的模型”就会被误当成信息的全部，而现实界那份没有被收进模型的信息，就连“被遗漏”的身份都没有了。

◆ 二、三态：符号、逻辑、数学

每一界的信息，又可以有三种模态。第395卷借了三个比喻性的名称：符号态叫“粒子”，逻辑态叫“波”，数学态叫“场”。这三个名称只是帮助记忆的类比，不是说信息真的是物理粒子、物理波、物理场。

先提醒一句，免得后面混读：本书还会用到另一组“三态”——秩序、混沌、介生，那是描述一个系统处在稳定、紊乱还是能承接新差异的分析语言（第8、16、22章会用到）。它与这里说的符号、逻辑、数学三态不是一回事，两组说法的对照见附录H。

符号态：一个可以被指认、保存、传递的单位。“暗裂”“松木”“胎压”都是符号。一个好的符号有三重性质：整体性（一个名字能调动一整组线索，说到“暗裂”，老师傅脑子里同时出现声音、手感、裂纹走向、后果）；稳定性（换一块木头、换一天，仍然认得出）；相对独立性（可以离开那块木头单独谈论它，回头又能接回木头上）。符号不是低级的碎片，第395卷说它是“饱满的一颗”。

逻辑态：信息怎样一步步推进、怎样获得理由。“声闷、手沉，所以有暗裂，所以不能做桌面，但可以做凳腿”，这是一条推进的链。逻辑态有两个分析维度：速度（推得快不快）和黏度（推进中卡不卡）。老师傅三秒钟就推完，徒弟要推三分钟，还常常卡在“闷到什么程度算闷”这一步。

数学态：信息之间的连接、次序与共同约束所构成的结构。“木材含水率每升高若干个百分点，声速下降若干”，“板厚、跨度与允许挠度之间满足某个关系”，这些不是一条一条的推断，而是许多量之间同时成立的约束网。

三态之间是**往返，不是阶梯**。这一点第395卷特别强调。冰、流水、蒸汽是同一种水的三种样子，谁也不比谁高贵；符号态可以展开成逻辑链，逻辑链又可以凝结成新符号，结构关系也可以被压缩回一个名字（“胡克定律”四个字就是一整张约束网压成的一颗粒子）。说“符号态是初级、数学态是高级”，是把一张平面地图读成了一架梯子。

厨房里最容易看清这一点。奶奶的菜谱写着“盐少许”，这是符号态，而且是饱满的符号——“少许”里装着她几十年对咸淡的拿捏。她教你时会说“先尝汤，汤淡了再加，加了要等滚一下再尝”，这是逻辑态。食品厂的配方写“每千克汤底加盐 12 克，误差正负 0.5 克”，这是数学态。三份信息各有用处：食品厂

的配方能让一万锅汤咸淡一致，却做不出奶奶那一锅；奶奶那一锅好吃，却无法让一万锅一致。

◆ 三、九宫：三界乘三态

把三界和三态交叉，得到第395卷所说的“信息九宫”：

	现实界	理念界	自我界
符号态（粒子）	现实符号：这棵树、这块板	理念符号：素数、导数	自我符号：“我”
逻辑态（波）	现实逻辑：声闷则有裂	理念逻辑：由定义推出定理	自我逻辑：一想到考试就心慌
数学态（场）	现实数学：量与量的实测约束	理念数学：公理化的形式结构	自我数学：情绪与行为模式的结构

逐格说几句，免得九宫成了一张好看的表。

现实符号最熟悉：我们指着一样东西叫出它的名字。理念符号是“素数”“导数”“函数”这一类，它们不指着任何一样摸得着的东西。自我符号最微妙，“我是个急性子”“我是个失败者”，这些名字一旦被当成自己的本性，就会反过来驾驭一个人。

现实逻辑是“天阴了要下雨”“声闷有暗裂”这类在现实里推进的链，它不靠公理，靠千百次的对照。理念逻辑是数学证明、法律推理那一类，在定义和规则之内推进。自我逻辑是“知道该早睡却总熬夜”那一类——第395卷把“知道却做不到”解释为理念逻辑与自我逻辑的顶撞：理念那条链推到了“该睡”，自我那条链推到了“再刷一会儿”，两条链不在一个格子里，谁也说服不了谁。

数学态的三格要特别小心。理念数学最清楚：有对象、有定义、有公理、有可检查的推导，这是数学家日常的工作台。现实数学指现实中量与量之间的实测约束，但第395卷提醒，要谈“现实数学”，必须交代变量是什么、映射如何给定、尺度多大、误差多少，否则只是一句漂亮话。自我数学是最具扩展性的命

名，它不等于已经形式化的数学，更不能拿来宣布疗效。九宫是提问查漏的工具，不保证每一格都已经被同样充分地建立起来。

这里还要补一条第395卷自己承认的边界：三态是否穷尽一切信息模态，尚待论证；九宫是“便于分析的互补维度”，不是已证明的完备分类。遇到第四种候选，不能硬塞回三格里去。

◆ 四、老师傅停在哪一格

现在回到木器铺。

老师傅的本事，在九宫里大致落在“现实逻辑”这一格，并且是饱满的现实逻辑：推进快，黏度低，很少出错。它也有现实符号的部分（“暗裂”“走纹”“油性大”这些行话），也有自我界的部分（那种手上一沉、心里一紧的警觉）。但它没有被提取成理念逻辑——没有写成一套徒弟拿去就能照着推的规则；更没有上到数学态——没有一组含水率、密度、声速、裂纹深度之间的定量关系。

第395卷给这种情况取了个名字叫“封顶”：某一界的信息停在符号态或逻辑态，上不到数学态。老中医号脉、老工匠挑料，是该卷举的两个例子。

但这个词必须立刻加上三道限制，否则就会从诊断滑向审判。

第一，封顶是**局部的**。它说的是“这一份信息、在这一项任务上”停在哪一格，不是说老师傅这个人、木匠这个行当、更不是说某一个文明“只到逻辑态”。第395卷明确禁止把“符号态封顶”用于整体评判一个文明、宗教或思想传统。一个文明有一千个行当，每一行都有自己的格子分布，拿一个“封顶”去概括整个文明，等于用一块木头的回声去判断一座森林。本书第8章讨论李约瑟之问时，会再回到这条禁令。

第二，封顶是**可撤回的**。工匠说不出来，不等于他心里没有逻辑结构；他的逻辑可能是靠示范、靠手把手传下去的，那是另一条传递通道，不是缺陷。徒弟

跟三年，照样学会挑料，只是这条路慢、贵、容易失传。所以“停在逻辑态”首先是对“传递方式”的描述，其次才可能是对“缺什么”的诊断。

第三，修复**不必都是公式化**。把老师傅的本事上升，可以是写成一本带图的手册（理念逻辑），可以是录一百段敲击声给徒弟听（加厚现实符号），也可以是测一千块木头，拟合出声速与暗裂深度的关系（现实数学）。哪一种合适，看任务。如果铺子只想把手艺传给儿子，后两种未必划算。

这三道限制，是本书和底本的一处重要分歧。底本在讨论古代科学时，倾向于把“没有上升到差异的数学表征”当作整个时代、整个文明的根本缺陷。按现行口径，这只能是局部诊断，而且只是若干候选解释之一。

◆ 五、可计算性：谓词有它的论域

儿子那句“要是能写成程序就好了”，其实问的是：老师傅的本事可不可计算？

这个问题初听起来很自然，细想却是一个错位的问题。“可计算”是一个谓词，就像“是偶数”是一个谓词。问“七是不是偶数”有意义；问“这块松木是不是偶数”，不是答案为否，而是问题本身没有落在谓词的论域里。

可计算性只对上升到数学态的那一格有定义。 严格地说，它是理念数学这一格里的谓词：一个函数、一个判定问题，只有在它的输入、输出、对应规则都被精确规定之后，才能问它是否可计算。老师傅那份现实逻辑，在被送进数学态之前，问它“可不可计算”，就像问松木是不是偶数。

> **可计算性是对什么下的判断** 在可计算性理论里，被判断的对象是函数 $f:\mathbb{N}\to\mathbb{N}$（或可以编码为自然数的有限对象之间的映射）。称 f 可计算，是说存在一台图灵机（或等价的计算模型），对每个输入 n 都在有限步内停机并输出 $f(n)$。判定问题是其特例：f 只取 0 或 1。 要对一个现实问题使用这个谓词，必须先做三件事：①选定什么算“输入”，什么算“输出”；②把它们编码为有限符号串；③规定输入到输出的对应规则，使每个输入恰好对应一个输出。 这三步本身不是计算，而是**送入数学态**的工作。可计算性理论回答的是③之后的问题；①②③做得对不对、丢掉了什么，

> 它不回答，也无法回答。　由此有一个朴素但常被忽略的推论：一个现实问题“可计算”，严格的说法应当是“它的某一个数学化版本可计算”。同一个现实问题，可以有多个数学化版本，其中一些可计算、一些不可计算，还有一些根本没抓住原问题。

这个方框里的三步，和本书的主命题正好接上。第①步“选定输入”，就是在流动的现实里截一刀，规定“其他条件不变”——这就是第10章要细讲的**微分**。老师傅的回声，在数学化时会被截成“某频段的振幅”；他手上那一沉、心里那一紧，多半截不进去。截取即丢弃。

截完之后，在数学态里把许多截口累加成模型，这是**积分**；模型可不可计算、算得多快，都是在这一侧发生的事。等模型要回到铺子里，告诉儿子“这一块板该不该用”时，又得在具体一块木头上**再微分**：这块板的含水率是多少，今天的湿度是多少，客人要做的是桌面还是凳腿。这些，模型里没有，要从现实里借回来。

所以“可计算性只属于圈的理念一侧”，不是一句玄话，而是可计算性理论自身的定义所带来的边界：它对数学态的那一格说话，对其他八格不说话。不说话，不等于其他八格“不可计算”，而是“可计算”这个词在那里尚未获得意义。

再举一个看病的例子。医生说“这个病人看着不对劲”，这是现实逻辑加自我逻辑，常常极准。医院把“不对劲”拆成心率、血压、乳酸、呼吸频率，做成一个预警评分，这是送入数学态；评分可以计算，可以自动报警。但评分漏掉的那部分“不对劲”，不会因为评分可计算就自动变得可计算——它根本没有进来。评分报警了，医生仍要走到床边，看这一个人，这一刻，再判断一次。

◆ 六、上升与下降：一张地图，不是一架梯子

最后要防一个误读：既然可计算性只在数学态有定义，那是不是一切信息都该“上升”到数学态，才算进步？

不是。三态是往返，不是阶梯。上升有收益：可传、可比、可算、可大规模复制，雕刻机能一天做一百张桌子。上升也有代价：每一次送入数学态，都要先截取、先编码，而被截掉的部分并不会消失，它会在落地时作为“意外”回来——这正是第12章要讲的积分常数 C 的外借。

看天的农谚是个好例子。“朝霞不出门，晚霞行千里”，是现实逻辑凝成的符号，世代相传，在某些地区相当灵验。气象台的数值预报把大气写成一组偏微分方程，在网格上一步步算出明天的云和雨，这是理念数学，而且是可计算的。数值预报的整体准确率早已远超农谚，这是上升的收益。可是网格之间的那一小片山谷、那一条河湾，方程里并没有，老农对自家那块坡地的判断，有时仍比预报更准。这不是农谚胜过了方程，而是方程在截取时把那片山谷切掉了，要在落地时从现实里借回来。上升把信息变得可算，同时也决定了哪些信息从此不在被算之列。

下降同样是正当的工作。一个公式被老工人用久了，会凝结成一句口诀（“三角形最稳”），从数学态回到符号态；这不是退化，而是让结构重新获得可以随身携带的“脸”。好的符号是“可返回的压缩”：口诀省去了推导，但必要时能沿着它找回推导。坏的符号是不可返回的压缩：只剩一句话，回不到它曾压缩的那张网。

这张九宫地图对本书的用处，可以概括为三点。第一，它让我们在谈“科学”时，说清楚自己在哪一格：大多数现代科学定律住在理念数学，但它们的数据来自现实界，它们的使用落在现实界，它们的意义被自我界承受。第二，它让“可计算”回到自己的论域：凡是听到“这个问题是可计算的”，先问一句“是它的哪一个数学化版本”。第三，它让我们对“封顶”保持克制：停在逻辑态的信息，可能正握着数学态里没有的那一份。

下一编要讲历史。我们会看到，微积分的发生，正是人类把“变化本身”从现实逻辑送进理念数学的那一次大规模上升；而在它之前，各个文明在九宫的不同格子里，早已积攒了深厚的信息。

◆ 七、这一编怎样算错

第一编（第1—4章）提出了四个主要主张：发生学可以作为主客二分困境的一个回答；科学对象是稳定下来的差异特征，而非先在实体；科学家所说的“自然界”是理念物理界，是物理界经数学筛选的部分；可计算性只对信息九宫中的数学态有定义。下列情形出现时，本编须收回或收窄相应主张。

1. 若能给出一种重要的科学对象，它在任何意义上都不依赖差异的稳定化（不需要可识别、可维持、可表达这“三个可”中的任何一项）即可被科学研究，则第2章“差异先于实体”须收窄为“适用于某些学科”。
2. 若发现某门成熟的定量科学，其研究对象与经测量、理想化后的模型之间不存在系统性的残差，也从未出现由残差引发的理论修订（即筛子从无回流），则第3章“理念物理界与物理界之分”须降为方法论提醒，不能作为本体论区分。
3. 若可以构造一种不经任何形式化编码（不选定输入、输出与对应规则）就能有意义地判定“可计算与否”的程序，则第4章“可计算性只对数学态有定义”须收回。
4. 若九宫在具体案例中无法比普通的“隐性知识／显性知识”二分多做出任何可检验的区分（例如无法预测哪一类经验在形式化时损失最大），则须承认九宫只是整理语言，其增量有限。
5. 若在具体任务中，被诊断为“停在逻辑态”的工匠知识，在系统形式化之后预测与操作表现均不低于原工匠，且没有可辨认的损失，则“截取即丢弃”在该领域须收窄为“可能丢弃”。
6. 若本编任何一处被用来整体判定一个文明“只有符号”“只到逻辑态”，则该用法本身即违反现行口径，应当删除，而不是修补。

带 走 的 话

“可计算”是一个有论域的谓词：只有输入、输出与对应规则都被规定之后，才能问一个问题可不可算。所以“这个现实问题可计算”，准确的说法是“它的某一个数学化版本可计算”；送进数学态时截掉了什么，可计算性理论不回答。九宫只是查漏的地图，三态往返而不分高低；“封顶”只能说某一份信息在某一项任务上停在哪一格，不能拿来判定一个人、一个行当或一个文明。

小 账

- 本章木器铺父子、开车、看病、奶奶菜谱、看天老农等例子均为构想场景或通俗比喻，不对应真实个案。
- 本章没有使用任何调查数字；信息九宫与三态的说法依据第395卷，可计算性的定义采用可计算性理论的标准表述。
- 本章不主张九宫是已证明的完备分类，也不主张停在逻辑态的知识低于数学态的知识。
- 本章主张的推翻条件见附录G第G13—G16条（第一编其余各章的条件见第G8—G12条）。

第 二 编

历史：微积分是怎样发生的

编 序

第二编只做一件事：把微积分放回它发生的过程里，看它不是天才的灵光一闪，而是在一片土壤里、经一串差异、一步步显露出来的。

第5章讲微积分之前那片“黎明前的天空”：巴比伦人会算而且算得准，托勒密的圆拟合得越好越说不出为什么，开普勒找对了形状却还没有力，从亚里士多德到伽利略，运动的观念换了底子。满天是光，只是还没汇聚到一个焦点上。第6章讲阿基米德、刘徽、祖暅：两千年前，在地中海的岛上和长江北岸的城里，有人各自想到了“切细再加”，并诚实地写下了做不到的地方。第7章讲牛顿与莱布尼茨：一个造出了发动机，一个造出了语言，以及那场旷日持久的优先权之争告诉了我们什么。第8章重审李约瑟之问：数理解释只是桌面上几种候选之一，要与制度、经济、地理解释并列，还要过循环与史实两道风险。第9章讲贝克莱讥讽的“消逝量的幽灵”，看柯西与魏尔斯特拉斯怎样把“趋近”改写成一句承诺，以及严格化本身又逼出了哪些新东西。

读这一编时，可以带着一个问题：微积分自己，是怎样被反常逼着走完一圈的?

05

第 5 章

黎明前的天空

> 这一章的问题：巴比伦人算得准木星，开普勒算得准火星，却都说不出行星为什么这样走——微积分出现之前，人类到底缺的是哪一样东西？

菜市场里卖了三十年菜的大姐，不看天气预报也知道明天进多少货。她心里有一本账：周五人多，月底人少，下雨天菜价涨、来的人少，端午前三天粽叶要多备。她能说出“下个月十五号前后，冬瓜要多进”，而且多半说中。可你要是问她“为什么下雨天来的人少”，她会说“下雨了谁出门啊”——这句话对，但它只是把现象再说了一遍。

她的账本有两种本事：记得住，算得准。它缺的是第三种本事：说清楚这一切是怎么一步步变过来的。她能预测，却说不出变化本身的机制；她能把一年的起落记成一张表，却写不出一条能推到她没见过的情形里去的规律。

人类在微积分出现之前的几千年里，面对天空、面对物体的下落、面对身体和物质，处境和这位大姐很像。他们记得住，有时也算得极准，但“变化本身”——每一刻变多少，为什么这样变——一直没有一个能在理念里反复推演的落脚处。本章就沿着天文、物理、化学、医学与工程，看一看这片黎明前的天空：光已经有了，而且有时很亮；差的是什么。

这里先立一条纪律。说“差了什么”，不是说古人笨，也不是说某个文明有根本缺陷。用上一章的话说，那时各界信息都很丰富，只是“变化本身”还没有被大规模送进理念数学这一格。这是一个关于信息位置的判断，不是一份关于文明高下的判决书。

◆ 一、泥板上的天空：会算，而且算得准

先说巴比伦。

至迟从公元前七世纪中叶起，两河流域的观星者就在泥板上逐日记录行星、月亮的位置与异象，留下了被今人称作“天文日志”的大量记录；现存可定年的日志，大约从公元前 652 年延续到公元前 61 年。到了公元前最后几个世纪，他们已经发展出相当精密的算术方案：用“之字形函数”和“阶梯函数”去推算月亮和行星在下一个周期的位置，用以预报月食、新月初见等现象。

底本说巴比伦人“只能说某个现象会再现，却不能解释为什么”，这话前半句要修正。2016 年，科学史家奥森德里弗（Mathieu Ossendrijver）在《科学》杂志上发表研究，释读了几块公元前 350 年至前 50 年之间的泥板：巴比伦天文学家把木星每天移动的度数（相当于速度）画在一边，把天数画在另一边，用一个梯形的面积去求木星在一段时间里走过的总距离，甚至去找把这个梯形分成两半的那个时刻。这已经是“速度—时间图下的面积等于位移”的思路，比欧洲十四世纪英、法学者（牛津默顿学派与巴黎的奥雷姆）的类似做法早了约一千四百年。

这件事本身就是一个回流的例子：我们以为“古人没有”的东西，被一块泥板顶了回来。所以，巴比伦人并非只会抄表格。他们在一个具体问题上，触到了“把变化率累加成总量”的门槛。

那么还差什么？差在这套方法是**针对木星的**，是一个算法，而不是一种语言。它没有被抽出来成为“任何一个随时间变化的量，都可以这样从变化率累加回总量”。用九宫的说法：巴比伦天文学在理念界已经有相当强的数学态，但那是一张为某几颗星定制的表，不是一台可以搬到别的问题上去的发动机。

大白话说，这就像菜市场大姐不仅记账，还发明了一个很聪明的算冬瓜的方法：把每天卖出的斤数加起来，按天气打个折，就能算出月底剩多少。方法好用，可它只为冬瓜而设，没人想到它也能用来算猪肉、算存钱、算孩子长个子。

梯形的面积就是走过的路 设某天体的日运动速度在时段 $[0,T]$ 内从 v_0 线性变到 v_1。按今天的写法，走过的总角距离为

$$s=\int_0^T v(t)\,dt=\frac{v_0+v_1}{2}\,T,$$

恰是速度—时间图中梯形的面积。巴比伦泥板做的就是这件事的具体数值版本。 但要注意两点限制：①它只适用于速度线性变化的情形，梯形公式是精确的；速度若是曲线变化，就需要把时段切得更细再求和，这一步泥板上没有；②泥板没有把“速度”与“位置”之间的关系写成一般性的互逆关系——从位置求速度（微分）与从速度求位置（积分）之间的对偶，要等到十七世纪才被普遍化。 因此，这是“积分思想的一次具体使用”，不是“微积分的发明”。两者的区别，在于是否成为一种对任何变化量都适用的通用语言。

◆ 二、托勒密的圆：拟合得越好，越说不出为什么

再说希腊天文学的巅峰：托勒密。

公元二世纪，托勒密在亚历山大里亚写成《天文学大成》（Almagest，约公元 150 年前后）。他继承了“天体做匀速圆周运动”的古老前提，用本轮、均轮、偏心圆，再加上一个叫“偏心等距点”（equant）的巧妙装置，去拟合行星忽快忽慢、时而逆行的观测轨迹。

底本称之为“补丁式几何”。这个说法有道理，但需要公道。托勒密体系在一千多年里能相当好地预报行星位置，它不是胡乱拼凑，而是一套严密、可计算、可校验的数学模型。它的问题不在于不准，而在于它的“准”从哪里来。

托勒密的模型是**纯运动学**的：它描述行星在哪里，却不问是什么让它到那里。每一个本轮，都是为了让模型贴合某一组观测而加上的；模型不准了，就加一个圆或调一个参数。这像一个裁缝，客人哪里不合身就在哪里打个褶，最后衣服确实合身了，可你看不出这个人的骨架。

用本书的主命题来看：托勒密在理念界里做了一次非常漂亮的积分——把观测拼成一个整体模型；但他积分的对象，是位置的记录，不是变化的机制。他没

有一个“此刻受了什么作用，所以下一刻怎样变”的微分环节。所以他的模型可以被调得很准，却无法回答一个新问题：如果多出一颗行星，或者行星的距离变了，会怎样？

◆ 三、开普勒：形状对了，力还没来

托勒密之后一千四百年，事情在一位脾气古怪的丹麦贵族和一位穷困的德国数学家手里起了变化。

1543 年，哥白尼的《天体运行论》出版，把太阳放到中心。这让本轮少了许多，但哥白尼仍然坚持圆周运动，所以他的体系在预报精度上并没有明显胜过托勒密。真正的转折来自数据。第谷·布拉赫在丹麦汶岛上建起天文台，用肉眼和巨大的仪器，积累了二十多年空前精确的行星观测，最好的测量精度接近 1 角分（约为满月直径的三十分之一），一般估计在 1 到 2 角分之间。第谷去世后，开普勒接手了这批数据。

开普勒与火星缠斗了多年。他先按圆轨道去拟合，结果与第谷的数据差了大约 8 角分。这 8 角分大约是满月直径的四分之一，放在托勒密的年代，完全可以当作观测误差打发掉。开普勒没有打发它。他相信第谷的数据，于是放弃了两千年来的圆。1609 年，他在《新天文学》中发表了前两条定律：行星沿椭圆轨道运行，太阳位于一个焦点上；行星与太阳的连线在相等时间内扫过相等的面积。1619 年，在《世界的和谐》中，他又给出第三条：行星公转周期的平方，与轨道半长轴的立方成正比。

这 8 角分，是本书反复要讲的那种东西：物理界从筛子里顶回来的残差。托勒密的筛子把它当噪声筛掉，开普勒把它当信号留下，于是天文学换了一张筛子。这可以看作 123 原理前两步的一个清楚的样本：新的差异（更精的数据）与旧的土壤（圆周运动的信念、托勒密的计算传统）顶撞，矛盾累积，最后逼出了新的显露（椭圆）。

值得一提的是第二定律。“相等时间扫过相等面积”，是一句关于**变化率**的话：行星离太阳近时跑得快，远时跑得慢，但单位时间扫过的面积不变。开普勒

为了得到它，实际上做了把轨道切成许多小段、把每一小段的面积累加起来的工作——又一次碰到了积分的门槛。

可开普勒说不出为什么是椭圆，为什么面积相等，为什么平方对立方。他曾猜想太阳发出某种“动力”推着行星转，但这个猜想没有能落成方程。他的三定律像三张极其准确的照片：形状对了，节奏对了，可照片里看不见那只推动的手。这只手，要等牛顿用万有引力和一套关于“每一刻变多少”的新数学，才被写下来（见第7章）。

用菜市场的话说，开普勒就像终于算准了冬瓜每个月卖多少、哪天卖得快，而且发现了一个很漂亮的规律，可他还说不出客人为什么那几天来得多。

◆ 四、从本性到惯性：亚里士多德、冲力与伽利略

天空之外，是地上的运动。

亚里士多德把运动分为“自然运动”与“受迫运动”：石头下落，是因为它的本性要回到宇宙中心；火焰上升，是因为它的本性在上方。受迫运动则需要推动者持续作用，推动一停，运动也就停下。

这套说法不是愚蠢的。它和日常经验很合：推车，手一松，车就停了。它在理念界有完整的逻辑——从“本性”“目的”推出“该往哪里去”。用九宫的话说，它是理念逻辑非常饱满的体系，只是它的推进方式是“为了什么”，而不是“每一刻变多少”。它解释“为什么往下”，却不问“下落得多快、快得多快”。

底本把中世纪描述为科学与神学捆在一起、学者只依靠亚里士多德和《圣经》的权威。这个印象流传很广，却需要修正。中世纪的大学恰恰是讨论自然哲学的主要场所，“自然是上帝写的另一本书”这种信念，在许多学者那里反而成了认真研究自然的理由。神学的框架确实规定了哪些问题值得问、哪些答案可以接受，这是土壤对差异的约束；但土壤里也长出了新东西。

中世纪的学者并没有原地踏步。十四世纪巴黎的布里丹提出“冲力”（impetus）说：抛出去的物体获得了一种内在的冲力，冲力会被空气阻力和重量慢慢消耗，所以物体在离手之后还能继续飞。这已经很接近惯性的思想。同一时期，牛津默顿学院的学者们讨论了“匀加速运动”的平均速度规则，巴黎的奥雷姆则用图形——横轴表示时间，纵轴表示速度——把这个规则画了出来。对照巴比伦泥板，这是同一种思路在欧洲的又一次独立出现。

然后是伽利略。他用斜面放慢下落，测出位移与时间的平方成正比；他论证水平方向上不受阻碍的运动会一直保持下去。1623 年他在《试金者》中写下那个著名的比喻，后人常把它概括为：自然这本大书，是用数学的语言写成的（原文大意见第3章）。1638 年的《关于两门新科学的对话》系统整理了他的运动学成果。

可是，伽利略写这本书的语言，仍然是几何与比例。他能证明“位移与时间平方成正比”，却是用三角形面积与相似比去证，而不是从“加速度是常数”出发、经过一次通用的运算得到位移。每一个新问题，他都要重新画一张图、重新找一组比例。

这就像一个手艺极好的家常菜厨子：红烧肉、糖醋鱼、宫保鸡丁，每一道都做得好，可每一道菜的火候都在他脑子里单独存着。他没有一张“火力、时间、食材厚度”之间的通用关系表，所以换一种没做过的食材，他得从头试。

从亚里士多德到伽利略，两千年里，“变化”一步一步从“本性”里被剥出来，从“为了什么”转向“变了多少”。剥到伽利略手里，已经剥得很干净了。剩下的那一步，是给“每一刻的变化”一个能独立运算的名字和一套通用的算法。

◆ 五、化学、医学与工程：现实逻辑的深厚积累

天文与物理之外，还有更多领域走的是另一条路。

化学。古希腊有恩培多克勒的四元素说，亚里士多德把它纳入冷热干湿的性质框架。中世纪的炼金术士在象征与神秘的外衣下，实际积累了蒸馏、升华、溶

解、结晶等大量操作技艺。这是厚厚一层现实逻辑：什么东西加热会怎样，什么溶液混合会冒泡沉淀。它缺的是一个能把“变化前后”精确对账的量。到了十八世纪后期，拉瓦锡坚持用天平称量反应前后的物质，并在 1789 年的《化学基础论》中系统阐述了质量守恒：反应前后总质量之差为零。这是一个关于变化的定量约束，化学由此开始进入数学态。值得一提的是，这一步主要靠的是精确称量与守恒式，而不是微积分；化学反应速率的微分方程化要更晚。所以，“每门科学的现代化都是由微积分推动的”这种全称说法，要收窄。

医学。希波克拉底学派的四体液说，中医的阴阳五行，都是以“平衡”为核心的解释体系，在经验上积累了大量有效的诊疗知识。文艺复兴时，维萨里在 1543 年出版《人体构造》，以前所未有的精度画出了人体的解剖结构。1628 年，哈维出版《心血运动论》，他做了一个关键的估算：心脏每搏动一次推出的血量乘以每小时的搏动次数，远远超过身体能在同一时间里造出的血量，所以血液必须循环。这个论证的力量，恰恰来自一次粗略的“速率乘以时间”的累加。医学开始从“描绘形态”转向“追问过程”，但过程的定量方程化——血流动力学、药物代谢动力学——是很久以后的事。

工程。埃及金字塔、罗马的拱券和高架渠、中国的都江堰与大运河，都是现实逻辑的伟大结晶。都江堰相传由李冰主持修建于公元前三世纪，鱼嘴分水、飞沙堰泄洪排沙、宝瓶口控制进水，千百年来一直在用。它背后没有流体力学方程，却有一套极精的现实逻辑与世代维护的制度（“深淘滩，低作堰”）。这些工程不是“缺乏科学”的证明，而是在另一格里达到了高度成熟的证明。它们的局限在于：方法与这一条河、这一片地紧紧缠在一起，很难搬到另一条河去；遇到前所未见的条件，只能再试错。

这三个领域的共同点是：现实逻辑极其深厚，理念逻辑各有体系，但关于“变化过程本身”的通用数学态，尚未形成。

◆ 六、黎明前缺的是哪一格

把这些领域放在一起看，可以得出一个比底本收窄、但更站得住的判断。

底本的说法是：在微积分之前，科学“始终没有抓住真正的对象”，“只能停留在经验与哲学解释阶段”。按现行口径，这需要改写为三点。

第一，**古代各界信息都很丰富**。巴比伦有精密的理念数学（而且碰到过积分的门槛），希腊有严密的几何与理念逻辑，工匠与医者有深厚的现实逻辑。说他们“没有抓住对象”，是用今天的格子分布去要求一个格子分布不同的时代。

第二，**缺的是一条特定的通道**：把“每一刻的变化率”与“累积起来的总量”之间的互逆关系，做成一种对任何变化量都通用的理念数学。巴比伦人在木星上用过它的一半，默顿学派与奥雷姆又用过一次，下一章要讲的阿基米德、刘徽、祖暅在面积与体积上用过它的积分一半。但没有人把两半焊在一起，也没有人让它成为一门可以教给任何人、用在任何问题上的演算。

第三，**这个判断限于动力学科学**。在天文与力学中，缺了这条通道，模型就只能是运动学的拼贴（托勒密）或个案的几何证明（伽利略）。在化学、工程中，通往现代形态的路有好几条，精确称量、守恒律、材料试验、统计方法，都各自起了作用。

还要补一句，这条通道的“前半截”在许多文明里都被走过，而且走得很远。中国的历法家为了推算日月五星在不等速运动中的位置，发展出精细的插值算法；元代郭守敬等人编《授时历》（1281 年起施行），用到了相当于三次差分的“招差术”，去处理太阳运行速度的逐日变化。这是在理念数学里对“变化的变化”做逐级差分，而且是为了预报而做的、可以实际运算的方法。伊斯兰世界的天文学家与数学家保存、翻译并大大发展了希腊天文学与印度数码，花拉子米在九世纪写下的代数著作，后来经十二世纪的拉丁译本进入欧洲，成为欧洲代数的源头之一。这些都提醒我们：十七世纪欧洲的那一步，是站在一个跨越多个文明、积累了两千年的底盘之上迈出的。底盘不属于某一个文明，迈出那一步的地方，也不等于底盘的全部（中国算学的这条路为什么没有走向微分，见第8章）。

这三点合起来，才是“黎明前的天空”的确切含义：不是一片漆黑，而是满天的光，只是这些光还没有被汇聚到同一个焦点上。

用菜市场大姐来收尾。她的账本已经极好；如果有一天，她发现“每天的进出”与“月底的结余”之间，有一种对冬瓜、猪肉、存款、孩子身高都一样的关系，并且能用几个符号把它写下来、教给隔壁摊位的小伙子，那就是另一个时代了。下一章，我们回到更早的古代，看看有哪几个人，已经在面积与体积里，摸到了这扇门的门槛。

带 走 的 话

微积分出现之前，各界信息都很丰富：巴比伦有精密的天文算法，希腊有严密的几何，工匠与医者有深厚的现实逻辑，中国历法家已会逐级差分。缺的不是“对象”，而是一条特定的通道——把每一刻的变化率与累积起来的总量做成一种对任何变化量都通用、可以互逆运算的理念数学。这个判断限于动力学科学，是关于信息位置的诊断，不是关于文明高下的判决。

小 账

- 本章菜市场大姐、家常菜厨子等均为构想的比喻人物，无真实原型。
- 本章所用巴比伦天文日志的起止年代、第谷观测精度等史实已核，出处见附录I第I19、I20条；奥森德里弗 2016 年的泥板释读据其在《科学》杂志发表的研究转述。
- 本章不主张古代“没有科学”，也不主张化学、医学、工程的现代化都由微积分推动。
- 本章主张的推翻条件见附录G第G17条；微积分是动力学科学“发动机”一说（与第7章共用）见第G53条。

06

第 6 章

古代先声：阿基米德、刘徽、祖暅

> 这一章的问题：阿基米德、刘徽、祖暅早在一千五百多年前就会“切细了再加起来”，把球的体积算得分毫不差——他们离微积分只差多远，又停在了哪里？

切过西瓜的人都知道一个窍门：想知道一个圆滚滚的西瓜大概有多少瓤，不必真去量那个曲面。把它切成一片一片，每一片近似一块薄圆饼，饼的大小一眼看得出，把所有饼加起来，就差不多是整个瓜。切得越薄，误差越小。

再想一个更细的问题：两摞硬币，一摞码得笔直，一摞歪歪扭扭像斜塔，哪一摞的金属多？只要硬币一样、枚数一样，当然一样多——每一层都是一枚同样的硬币，层数相同，总量就相同，跟摞得直不直无关。

这两个日常的念头，一个是“切细了再加起来”，一个是“每一层都一样，总量就一样”。它们看起来朴素，却是积分思想最核心的两块砖。从两千多年前到一千五百多年前，已经先后有人把这两块砖砌进了严格的数学里：希腊的阿基米德，中国的刘徽与祖暅。

上一章说，黎明前的天空满是光，差的是一条把“变化率”与“累积量”焊在一起、对任何变化都通用的通道。本章要看的是这条通道的一半——积分的那一半——在古代是怎样被走通的，走到了哪里，又停在了哪里。停在哪里，是本章最要说清楚的事；而说清楚它，首先要把史实摆准。

◆ 一、穷竭：用可以算的，去夹住不能直接算的

先说希腊。

在阿基米德之前，柏拉图学园的欧多克索斯（约公元前四世纪）已经发展出一种严格处理曲边图形的方法，后人称之为“穷竭法”。它的思路是：要证明一个曲边图形的面积等于某个值，就构造一系列直边图形从里面去逼近它，证明二者之差可以小于任何给定的量；再用反证法排除“大于”与“小于”两种可能，于是只能“等于”。欧几里得《几何原本》第十二卷用这种方法证明了“圆面积之比等于直径平方之比”。

这个方法极其谨慎。它从不说“无穷多块加起来”，只说“无论你要求误差多小，我都能做到更小”。这份谨慎有它的道理：希腊人已经被芝诺的悖论吓过一回，知道一旦碰“无穷”，推理就可能出事。

阿基米德（约公元前 287—前 212 年）把这套方法用到了极致。在《圆的度量》里，他用圆内接与外切的正九十六边形，把圆周率夹在 $3\frac{10}{71}$ 与 $3\frac{1}{7}$ 之间。在《抛物线弓形求积》里，他证明了抛物线弓形的面积，等于同底同高内接三角形面积的三分之四。在《论球与圆柱》里，他证明球的体积是外切圆柱体积的三分之二，球的表面积是圆柱侧面积——据说他希望把球与圆柱刻在自己的墓碑上。

大白话说，穷竭法像是给一个形状古怪的池塘估水量：先在池塘里铺满一块块规整的方砖，砖的总面积肯定比池塘小；再在池塘外围一圈方砖，总面积肯定比池塘大；砖越铺越小，里外两个数越夹越紧，最后能断言池塘面积就是两者共同逼近的那个数。

> **抛物线弓形：一个几何级数的“穷竭”** 设抛物线弓形内接三角形面积为 T。阿基米德证明：在剩下的两个小弓形中各作同样的内接三角形，它们的面积和为 $T/4$；再往下，下一层的总面积为 $T/16$，依此类推。于是弓形面积 A 满足
>
> $$A = T\left(1 + \tfrac{1}{4} + \tfrac{1}{16} + \cdots\right).$$

> 今天我们直接写 $A=\frac{4}{3}T$。阿基米德不这样写。他证明的是：有限项之和 $T\left(1+\frac{1}{4}+\cdots+\frac{1}{4^n}\right)$ 加上最后一项的三分之一，恰好等于 $\frac{4}{3}T$；再分别论证 A 不能大于、也不能小于 $\frac{4}{3}T$。 这个证明里没有“无穷项相加”，只有“任意有限次之后的精确余项”加上双重反证。它是严格的；代价是：每一个新图形，都要重新找一套构造与余项。

这个代价，就是穷竭法的局限所在。它是一种**证明方法**，不是一种**求法**。你得先知道答案，再用它去证明答案对。那么答案从哪里来?

◆ 二、《方法》：先用杠杆“称”出答案

阿基米德自己回答了这个问题，只是答案在一千多年里不见天日。

他写给埃拉托色尼的一封长信，后人称为《方法》（全名大意为“关于力学定理的方法”）。这部著作长期失传，1906 年，丹麦学者海贝格在君士坦丁堡见到一部中世纪的羊皮纸祈祷书，发现它是用刮去旧字的羊皮重写的，底层隐约透出的正是阿基米德的文字——这就是著名的“阿基米德重写本”。二十世纪末以后，研究者借助多光谱成像等技术，又从中读出了许多海贝格当年未能辨认的内容。

《方法》里，阿基米德坦白了他的发现过程：把一个图形想象成由无数条平行的线（或薄片）组成，把这些线一条条挂到一根想象的杠杆上，和另一个已知图形的对应线逐一平衡；既然每一条都平衡，整体也就平衡，面积或体积之比就“称”出来了。他明说这种方法不算证明，得出结果之后，还要用穷竭法严格证一遍。

这是一个非常动人的场景：一个人先在理念里做了一次“物理实验”，把图形切成无数薄片、逐片称重，得到一个猜想；再回过头来，用当时最严格的语言去证明它。按本书的语言，前一步是在理念界里借了现实界的逻辑（杠杆平衡），后一步是把它送进理念数学。

这也是底本说“先发现差异关系、再积分性重建整体”的出处。但这里要收窄一点：阿基米德的“切片”切的是静止的几何图形，不是随时间变化的量。他的薄片并不是“变化率”，而是“截面”。这一区分，后面会变得很要紧。

◆ 三、割圆：割之又割，以至于不可割

在中国，同一种智慧沿着另一条路生长。

《九章算术》大约在汉代成书，是一部以问题与算法组织起来的数学经典，涉及田亩面积、粮食比例、工程土方、方程组等。书中计算圆面积用的是“周三径一”，即取圆周率为三。三国时期魏人刘徽在公元 263 年为《九章算术》作注，对这一粗率很不满意，于是创立了“割圆术”。

他从圆内接正六边形开始，边数逐次加倍：十二、二十四、四十八、九十六，一直算到一百九十二边形，求得圆周率约为 3.14，即分数 $\frac{157}{50}$，后世称“徽率”。注文中还有一段算到三千零七十二边形、得出 3.1416 的内容，这部分究竟出自刘徽本人还是后人，学界有争议：有学者认为出自祖冲之，三上义夫、李约瑟等则认为仍是刘徽的手笔。

刘徽在注中写下了那段著名的话：“割之弥细，所失弥少。割之又割，以至于不可割，则与圆周合体而无所失矣。”（底本引作“与圆合体”，原文作“与圆周合体”。）意思是：边数越多，多边形与圆之间的差距越小；一直割下去，割到不能再割，多边形就和圆周合为一体，不再有任何差失。

和阿基米德相比，刘徽有两处值得特别注意。

第一，他不止于内接多边形，还用到了一个巧妙的上界：在内接多边形的每条边外，补上一个以边为底、以“余径”为高的小矩形，得到的面积一定大于圆。这样，圆面积被夹在两个可算的数之间。这同样是一种“夹逼”，虽然没有希腊式的双重反证，却清楚地意识到了误差的上下界。

第二，他的话里出现了“以至于不可割”这样直接面对极限的表述。希腊人对无穷小心翼翼地绕开，刘徽则正面说出“合体而无所失”。这是不是更大胆？

是的。是不是更严格？按今天的标准，“割到不可割”需要一个关于极限的精确定义，这在刘徽那里还没有，在阿基米德那里也没有，要等到十九世纪（见第9章）。

刘徽的工作远不止割圆。他在讨论“阳马”“鳖臑”等立体的体积关系时，也使用了把立体无限二分、使剩余部分趋于零的论证，有学者称之为“刘徽原理”。他还研究了球的体积：他构造了一个叫“牟合方盖”的立体——两个圆柱垂直相贯所围成的那块，形状像两把合起来的方伞——并正确指出，球与牟合方盖在每一高度上的截面积之比都是 $\pi:4$，所以二者体积之比也是 $\pi:4$。可他算不出牟合方盖本身的体积，于是坦然写下：“敢不阙疑，以俟能言者。”——不敢强作解人，把疑问空着，等待能说清楚的人。

这个“以俟能言者”，是数学史上最好的一种诚实：把自己做到了哪里、卡在了哪里，清清楚楚地留给后人。

◆ 四、祖暅：幂势既同，则积不容异

“能言者”在两百年后出现了。

南朝的祖冲之（429—500）把圆周率推算到 3.1415926 与 3.1415927 之间，并给出约率 $\frac{22}{7}$ 与密率 $\frac{355}{113}$。在分母小于一万六千的分数里，密率是最接近圆周率的一个；他的圆周率精度在世界范围内保持了九百多年的领先，直到十五世纪初，才先后被印度喀拉拉学派的摩陀婆（约 1400 年）和中亚的阿尔·卡西（1424 年）超过。祖冲之父子的数学著作《缀术》后来失传，今人只能从其他文献的引述中窥见一二。

他的儿子祖暅（一作祖暅之），解决了刘徽留下的难题。唐代李淳风等人为《九章算术》“开立圆术”作注时，记载了祖暅的方法：他取牟合方盖的八分之一，考察它与外切正方体之间的“剩余”部分，发现这块剩余在每一高度 h 上的截面积都等于 h^2，与一个倒立的方锥的截面积相同。于是剩余部分的体积等于方锥的体积，由此算出牟合方盖的体积，再按刘徽的 $\pi:4$，得到球体积公式。

支撑这一步的，就是那句著名的原理：“缘幂势既同，则积不容异。”“幂”指截面积，“势”指高；两个立体在等高处的截面积处处相等，体积就不可能不同。这正是本章开头那两摞硬币：每一层都一样，总量就一样，跟摞得直不直无关。

> **祖暅原理与卡瓦列里原理**　设两个立体夹在两平行平面之间，用任一与之平行、距底面为 h 的平面去截，所得截面积分别为 $S_1(h)$、$S_2(h)$。若对一切 h 都有 $S_1(h)=S_2(h)$，则两立体体积相等。 用今天的话说，体积 $V=\int_0^H S(h)\,dh$，被积函数处处相等，积分当然相等。 欧洲方面，意大利数学家卡瓦列里在 1635 年出版的《用新方法推进的连续不可分量几何学》中提出了相同的原理，后世称“卡瓦列里原理”。祖暅比他早了一千一百年左右。　但要看清这个原理**没有**说什么：它只比较两个立体，告诉我们“相等”，并不给出一个从 $S(h)$ 直接求出 V 的一般算法；求出具体数值，仍要找到一个体积已知的参照立体（祖暅找的是方锥）。从“比较原理”到“通用算法”，中间还隔着一个基本定理（见第13章）。

◆ 五、门槛之外的回声

阿基米德、刘徽、祖暅不是孤例。在他们之后、牛顿之前，“切细再加”的念头又在好几个地方独立冒了出来，每一次都往前挪了一点。

十四至十六世纪，印度南部喀拉拉地区有一个数学与天文学派，以摩陀婆（Madhava，约 1340—约 1425，生卒年均为学者推定）为开端。这个学派得到了正弦、余弦、反正切的无穷级数展开，以及圆周率的无穷级数表示，比欧洲的格雷戈里、莱布尼茨早了两百多年。他们的推导中，有把弧切成许多小段、逐段求和、再让段数无限增加的步骤。这已经很接近把“无穷求和”当作一种可操作的工具了。不过，这些成果主要服务于天文计算，以梵文诗体写在注疏里，长期没有被外界知道；是否传入欧洲，学界目前没有确证。

欧洲方面，开普勒在 1615 年写过一本小书，起因很生活化：他再婚时买了几桶葡萄酒，发现酒商只用一根量杆从桶口斜插到底，读一个数就算价钱，不管桶的胖瘦。他觉得不对，于是研究起各种旋转体的体积，书名大意为《酒桶的新立

体几何》。他的方法相当随意：把一个立体想成无数个极薄的圆片或无数个以中心为顶点的小锥体，直接把它们加起来。这种做法在严格性上远不如阿基米德，却很好用，算出了九十多种立体的体积。

然后是卡瓦列里的“不可分量”（1635），以及 1630—1640 年代费马、罗贝瓦尔、托里拆利、沃利斯等人一连串求面积、求切线的工作。到这时，“切细再加”已经从少数天才的绝技，变成了一批数学家都会用的家常手艺，只差有人把它收拾成一套可以教、可以传、可以在任何问题上照着做的规则。

这些回声说明了同一件事：积分的这一半，是人类在很多地方、很多次独立走到的。它像一条许多人各自从不同方向走到的山口，山口本身不属于任何一个人。

◆ 六、先声停在了哪里

把阿基米德、刘徽、祖暅放在一起，可以看到一幅清楚的图。

他们都在理念数学这一格里做到了非常高的水平。他们都会“切细再加”，都知道误差可以夹住，都触及了“无穷分割”的边界。在面积与体积这件事上，他们的结果是精确的、可检验的，至今仍是对的。

底本说，这些先声“停留在积分性直觉的层面”。这个说法要改。它们不是“直觉”，而是严格的、有证明或有清楚算法的数学成果；说它们只是直觉，是低估了它们。

那么，它们停在了哪里？可以分三层说。

第一层：对象是静止的。 三个人处理的都是图形——圆、抛物弓形、球。被切开的是空间，不是时间；薄片是截面，不是“某一时刻的变化率”。没有一个“随时间变化的量”被当作对象。所以积分的这一半，还没有和“运动”“变化”接上。

第二层：方法是逐题的。 每一个新图形都要重新构造：阿基米德要为抛物线找几何级数，刘徽要为圆找多边形，祖暅要为牟合方盖找方锥。这些构造极其巧

妙，但巧妙本身就说明它们不能被机械地搬到下一题。用上一章的说法，这是一台台为某个问题量身定做的精密仪器，还不是一台通用的发动机。

第三层：没有反过来的那一半。 “从截面求体积”是积分；“从体积求截面”、“从位置求速度”是微分。两者互逆，这一点在古代没有被说出来，也没有被当作方法使用。缺了这一半，就没有那个可以来回走的圈。

打个记账的比方。古代先声相当于已经学会了“把一年里每一天的收入加起来，得到全年总收入”，而且加得又准又严格；可他们记的是一本已经写完的旧账，账上的每一页是静止的。他们还没有反过来问：“如果我知道这一年里每一时刻的余额，能不能算出此刻钱是在流进来还是流出去、流得多快？”更没有想到：这两个问题是同一个问题的两面，会了其中一个，就自动会了另一个。

这三层合起来，就是古代先声的确切位置：积分的一半已经被严格地走通，但它停在静态图形上、停在逐题构造上、停在单向上。

◆ 七、不要把它当成判决

最后要守住一条边界。

底本在这一节末尾写道，这些先声“恰恰印证了李约瑟之问的答案”，暗示中国（以及希腊）因为没有迈出微分这一步，所以长不成现代科学。按本书的证据责任，这个推论不能这样下。

理由很简单：希腊和中国同样停在了上面那三层，可我们并不会因此说“希腊文明缺乏现代科学的数理内核”。阿基米德之后，希腊的数学传统在亚历山大里亚又延续了几百年，同样没有走向微分。这说明，“停在积分一半”是古代数学普遍的位置，而不是某一个文明的特殊病征。一个所有文明都共享的状态，不能用来解释一个文明与另一个文明之间的差别。

至于中国算学为什么在宋元达到高峰之后没有走向微积分、欧洲为什么在十七世纪走了出去，那是一个牵涉数学性格、符号系统、制度、经济、地理等多种因素的问题，本书第8章会把它们并列讨论，并交代其中的循环风险与史实风险。

不过，有一个现象值得在这里先记一笔，因为它和本书后面要讲的“回写”有关。阿基米德的《方法》失传了一千多年，差一点就永远压在祈祷文底下；祖冲之父子的《缀术》在唐代曾列为国子监算学的教材，《隋书·律历志》记它太难，“学官莫能究其深奥，是故废而不理”，后来终于亡佚。两份最接近“方法”而非“结果”的文献，都没能留下来。留下来的，多是结果：圆周率的数值、球的体积公式。

这像一位厨师把一道菜做成了，菜谱上只记下“出锅时撒葱花”，却没记下他是怎么想出这道菜的。下一代人能照做这道菜，却很难照着他的思路做出下一道新菜。结果能传，方法难传；结果入了库，方法没有回写进后人的底盘。古代先声之所以只是“先声”，原因之一或许正在这里：它们的方法没有被一个持续的共同体反复使用、反复改造，从而一代代沉淀下来。这是一个可以拿史料去检验的猜想，不是结论。

在这里，我们只需要记住一个更朴素的事实：两千多年前在地中海的西西里岛上，一千七百多年前在三国时的魏地，一千五百多年前在南朝的都城建康，先后有人各自独立地想到了“切细再加”；他们各自把它做到了自己时代的极限，并且诚实地写下了做不到的地方。这一份积累，是人类共同的底盘。

那么，是谁把“切细再加”和“每一刻变多少”焊到一起，并让它驾驭了运动本身？下一章，我们来到十七世纪的英格兰与德意志。

带走的话

阿基米德的穷竭与杠杆、刘徽的割圆、祖暅的“幂势既同，则积不容异”，是严格的数学成果，不是“积分性直觉”。它们停在三处：对象是静止的图形而非随时间变化的量，方法是逐题构造而非通用算法，只走通了积分一半而没有说出与微分的互逆。这是古代数学普遍的位置，不能拿来解释某一个文明与另一个文明的差别。

小 账

- 本章切西瓜、两摞硬币、铺砖量池塘、记账、厨师菜谱等均为构想的比喻，不对应真实个案。
- 本章所涉刘徽注中圆周率一段的归属、祖冲之领先年数、摩陀婆生卒年、《隋书·律历志》对《缀术》的记载，出处见附录I第I21—I24条。
- 本章所说“方法失传而结果留存，是先声只成先声的原因之一”，只是待史料检验的猜想，不当结论。
- 本章不主张古代先声证明了李约瑟之问的任何答案。
- 本章主张的推翻条件见附录G第G18条。

07

第 7 章

牛顿与莱布尼茨：发动机与语言

> 这一章的问题：一个人最会“开”，把新数学装进了天上地下的运动；一个人最会“说”，给它造了人人能照着算的符号——微积分为什么要这两半合起来才算诞生？

学开车的人都有过这样的体验。教练坐在旁边，说“松一点离合，脚下给一点油”，你照做，车一抖就熄火了。教练说的每个字你都懂，可“一点”到底是多少，只有脚知道。练了几十回，有一天车平平稳稳地起步了，你却说不清自己做对了什么。这是一种本事：会开。

后来你给侄子讲怎么开车，发现自己必须找一套说法：“先把离合抬到车身微微发抖的那个位置，停住，再慢慢给油。”“半联动”这个词，是你从教练那里学来的，你一说，侄子就有了一个抓手。这是另一种本事：会说。

会开不一定会说，会说也不一定会开。最好的情况，是一个人既会开，又能把开车变成一套任何人都学得会的说法。

十七世纪后半叶，微积分的诞生恰好是由这样两个人完成的：一个最会“开”，一个最会“说”。牛顿把这门新数学装进了自然哲学，让它拉动了天上地下的一切运动；莱布尼茨为它造了一套符号，让任何受过训练的人都能照着规则去算。底本的比喻很贴切：牛顿的微积分像“科学的发动机”，莱布尼茨的微积分像“科学的语言”。本章就从这两半讲起，最后讲两人之间那场旷日持久的争吵，以及这场争吵本身告诉我们什么。

◆ 一、底盘已经备好

先看两人登场时的底盘。

上两章讲过，“切细再加”的积分一半，从阿基米德、刘徽到卡瓦列里、费马，已经被反复走通。十七世纪上半叶，另一半也在快速成形：笛卡尔在 1637 年的《几何学》中把曲线写成方程，费马几乎同时独立做到了这一点，曲线从此可以被代数地操作；费马用“让变量增加一个很小的量、比较前后差值”的办法求极大极小值和切线；牛顿在剑桥的前辈巴罗讲授几何，他的讲义里已经出现了切线问题与面积问题之间的某种对应。沃利斯在 1656 年的《无穷算术》中大胆地处理无穷级数与无穷乘积。

换句话说，到 1660 年代，求切线（微分的雏形）和求面积（积分的雏形）都已有一批零散的技巧，但它们像散落在工具箱里的一堆零件：每一个都能用，却没有被装成一台机器，更没有一份说明书说明它们之间的关系。

这就像一个家里攒了许多年的厨具：有一口好锅，一把好刀，一杆秤，一个计时的沙漏，每一样都有人用过、用得很好。可是没有人把它们按顺序摆上灶台，写出一张“先称、再切、再下锅、看着沙漏翻面”的流程，所以每一顿饭都要从头琢磨。零件齐备，与一套能反复运转的做法，是两回事。从零件到做法，需要有人看清各个零件之间的关系——尤其是那条最关键的关系：求切线与求面积，是同一件事的正反两面。

两人各自完成的，正是装配。

◆ 二、牛顿：流动的量，和量的流速

艾萨克·牛顿生于 1642 年（按当时英格兰仍用的旧历为 1642 年圣诞节，按新历为 1643 年 1 月），1661 年进入剑桥三一学院。1665 年至 1666 年间，伦敦大瘟疫蔓延，大学关闭，他回到林肯郡伍尔索普的家中。后来他自己回忆，正是在这段时间里，他想出了流数法、颜色理论与万有引力的初步思想。“奇迹之

年”的说法由此而来；但史家也指出，这份回忆是牛顿在优先权之争中写下的，有替自己争先的成分，许多思想实际是在此后多年里逐步完成的。

牛顿的出发点是运动。他把一个量想象成随时间“流动”的东西，称为“流量”（fluent）；流量变化的速度，称为“流数”（fluxion）。一个点在曲线上移动，它的横坐标和纵坐标都是流量，它们各自的流速就是流数。求切线，就是求纵坐标流速与横坐标流速之比；求面积，就是反过来，已知流数求流量。

这里最关键的一步，是牛顿看清了：**这两个问题是互逆的。**　知道流量求流数，与知道流数求流量，是同一件事的两个方向。于是，所有求面积的问题，都可以转化为“找一个量，使它的流数恰好是给定的那个”；而这类“反求”，可以借助已经做好的“正求”表格来完成。从这一刻起，过去需要逐题构造的面积问题，变成了查表与运算。

1669 年，牛顿写成《运用无穷多项方程的分析》（De analysi），经巴罗交给伦敦的数学通讯人柯林斯，在少数人之间传阅，但没有出版。1671 年前后，他写成更系统的《流数法与无穷级数》，同样没有出版，直到他去世后的 1736 年才以英译本面世。1704 年，他把《论曲线求积》作为《光学》的附录发表。至于后来通行的“字母上加点”的流数记号（如 $\dot{x}$），是他在 1690 年前后才开始使用的，并非 1665 年就有；1693 年沃利斯的著作里首次印出了牛顿流数法的一些内容。

流数的算法：一个“无穷小时段”　牛顿早期的做法是：设时间流过一个无穷小的时段 o，流量 x 增加 $\dot{x}\,o$，流量 y 增加 $\dot{y}\,o$。若 $y = x^n$，则

$$y + \dot{y}\,o = (x + \dot{x}\,o)^n = x^n + nx^{n-1}\dot{x}\,o + \frac{n(n-1)}{2}x^{n-2}\dot{x}^2o^2 + \cdots$$

两边减去 $y = x^n$，再除以 o，得 $\dot{y} = nx^{n-1}\dot{x} + (\text{含 } o \text{ 的项})$；最后令含 o 的项“消失”，得 $\dot{y}/\dot{x} = nx^{n-1}$。　这一步的困难也在这里：先用 o 去除（所以 $o \neq 0$），后又把含 o 的项丢掉（好像 $o = 0$）。1734 年，贝克莱主教在《分析学家》中正是抓住这一点，讥讽这些量是“消逝量的幽灵”。牛顿本人后来改用“最初比与最终比”的说法，接近极限思想，但没有给出今天意义上的严格定义。这个“幽灵”要到十九世纪才被驱散（见第9章）。

牛顿最大的不同，不在算法，而在用途。1687 年，他出版《自然哲学的数学原理》。书中的论证大量采用几何形式，而不是直接写流数演算——这是当时的学术习惯，也可能是牛顿有意让论证显得更“古典”、更难被挑剔。但在几何外衣下，推动全书的是一个新的思路：力决定的是运动的**变化**，而不是运动本身。用今天的写法，就是质量乘以加速度等于力，加速度是位置对时间的二阶变化率。

把这个关系与“引力与距离平方成反比”放在一起，牛顿推出了开普勒的三条定律：椭圆轨道、等面积、周期平方与半长轴立方成正比。第5章那张“形状对了、看不见那只手”的照片，第一次有了那只手。而且同一组关系，还能用来算炮弹、潮汐、地球的扁率、月球运动的不规则。

用开车来比喻：此前的天文学，像是一个人站在路边，把一辆车每天经过的时间和位置记得清清楚楚；牛顿则坐进了驾驶室，看见了油门和方向盘——只要知道此刻踩了多少油门、打了多少方向，就能推出车下一刻在哪里。这正是“发动机”的意思：它不是对轨迹的又一次描述，而是生成轨迹的那个机制。

按本书的主命题来看，牛顿完成了圈的前两段：在现实里截取“此刻的受力与加速度”（微分），在理念里把它们累加成完整的轨道（积分）。而他用这个模型去预报彗星回归、计算潮汐时，又必须把具体的初始位置与速度从观测里借回来——这是第三段，再微分。牛顿的体系第一次让这三段连成一个能反复运转的圈。

◆ 三、莱布尼茨：给差异造一个名字

莱布尼茨的道路完全不同。

戈特弗里德·威廉·莱布尼茨 1646 年生于莱比锡，学的是法律与哲学，一生为宫廷做外交、法律、图书馆与家谱的工作。他对数学的深入，晚于牛顿：1672 年他以外交使命到巴黎，在那里结识惠更斯，在惠更斯的指导下才系统研究当时的前沿数学。1672 至 1676 年的巴黎时期，是他的微积分发生的时期。

莱布尼茨的出发点不是运动，而是符号与组合。年轻时他就设想过一种“普遍文字”：如果能为一切概念找到恰当的符号，那么争论就可以化为计算——两个人有分歧，不必吵，坐下来说一句“让我们算一算”。这个梦想后来没有实现，但它决定了他看待数学的方式：好的数学，应当有好的记号，让思考变成照规则操作。

他从数列的差与和入手。一列数，相邻两项相减得到“差”的数列；把差逐项加起来，又回到原来的数列。差与和是互逆的。他把这个观察推到连续的情形：曲线上相邻两点的纵坐标之差，就是一个无穷小的“差”；把无穷多个这样的小量加起来，就是面积。在研究帕斯卡关于圆的一篇文章时，他注意到一个由 dx、dy 与弧长元素组成的“特征三角形”，由此看清了切线问题与面积问题之间的普遍联系。

1675 年 10 月 29 日的一份手稿里，莱布尼茨第一次写下拉长的 S 形积分号 $\int$，取拉丁文 summa（和）的首字母；约两周之后（同年 11 月 11 日的手稿），又用 d 表示差（differentia）。这两个符号一出现，许多过去需要绞尽脑汁的推导，变成了几乎机械的操作。例如乘积的微分法则 $d(xy) = x\,dy + y\,dx$，链式法则 $\dfrac{dy}{dx} = \dfrac{dy}{du}\cdot\dfrac{du}{dx}$——后者看上去就像分数约分，虽然严格地说它不是约分，但这个“看上去像”本身极有用：它让人凭符号的形状就能猜对结果。

1684 年，莱布尼茨在莱比锡的《博学者学报》（Acta Eruditorum）上发表了一篇只有几页的论文，题目大意是《求极大值、极小值与切线的新方法》，这是世界上第一篇公开发表的微分学论文。1686 年，他在同一刊物上发表关于积分的文章，$\int$ 首次出现在印刷品中。

此后的传播非常快。瑞士的雅各布·伯努利与约翰·伯努利兄弟迅速掌握了这套方法，并用它解决了悬链线、最速降线等一系列难题；1696 年，洛必达侯爵出版了第一本微积分教科书《无穷小分析》，内容大量来自约翰·伯努利的讲授。一门学问，从一个人的手稿变成一群人的日常工具，只用了十几年。

这就是“语言”的意思。牛顿的流数法是一位大师手里的发动机，旁人看得到它的威力，却很难拆开来学；莱布尼茨的符号是一本说明书，任何人照着规则

都能把发动机装起来、开起来。用九宫的话说，莱布尼茨的贡献，很大一部分是在理念符号这一格：他为“差”与“和”造了两颗饱满的粒子，让它们可以被调用、组合、传递，并且能随时展开回背后的运算——这正是第395卷所说的“可返回的压缩”。

莱布尼茨的梦想，还有一层要在这里先标出来，因为本书的后半部会一再回到它。“让我们算一算”，意思是：只要把问题写成恰当的符号，分歧就可以交给计算来裁决。这是“科学即计算”这一观念最早、最清楚的表达之一，后来的数理逻辑、图灵机、计算机都可以追溯到这条线上。

按第4章的九宫来看，这个梦想有一个默认前提：问题已经被送进了数学态。可是现实中的分歧，大多卡在送进去之前——什么算输入，什么算相关，哪些条件可以“不变”。两个邻居为院墙的位置吵架，拿出尺子量一量，确实能平息一部分争执；可如果他们吵的是“这堵墙挡了我家的光”，尺子就帮不上多少忙了。“算一算”只能裁决已经被写成符号的那一部分；写成符号的那一刻，截掉了什么、由谁来截，是计算之前的事。莱布尼茨的伟大之处在于给“算”造了最好用的语言；他的梦想过界的地方，在于把一切争论都当成了可以先写成符号的东西。本书第六编讨论“没有 AI 不能解决的问题”时，会看到同一个梦想的当代版本。

◆ 四、优先权之争：一场关于“谁先”的漫长官司

两个人几乎同时、彼此独立地走到了同一个地方，争吵几乎是不可避免的。

事情的时间线大致是这样。牛顿的流数思想形成于 1660 年代中后期，但一直没有发表。1676　年，莱布尼茨第二次访问伦敦时，在柯林斯处看到过牛顿《分析》的部分内容。同一年，牛顿应请求先后写了两封长信寄给莱布尼茨，介绍他的级数方法，但对流数法的核心只用一个字母打乱的“字谜”（anagram）带过，等于把发明“锁”了起来——既声明了自己有，又不让对方看见。莱布尼茨的回信则相当坦率地介绍了自己的微分方法。

真正的争吵在 1690 年代末才爆发。1699 年，旅居英国的瑞士数学家法蒂奥公开暗示莱布尼茨抄袭；1708 年前后，牛顿的追随者凯尔在皇家学会的刊物上更直接地指控。莱布尼茨向皇家学会申诉，要求主持公道。1712 年，皇家学会成立了一个委员会调查此事，次年初刊行报告《通信集》（Commercium Epistolicum），判定牛顿为第一发明人，并暗示莱布尼茨剽窃。后来的史家查明，这份报告的主要内容实际上由时任皇家学会会长的牛顿本人起草；1715 年刊物上那篇为报告作总结的匿名长文，也出自牛顿之手。1716 年莱布尼茨去世，争吵并未随之结束。

今天的科学史界基本已有共识：两人是独立发明的。牛顿在时间上更早，莱布尼茨在发表上更早；莱布尼茨在伦敦看到的材料，不足以让他从中“抄”出自己那套完全不同的符号与思路。

◆ 五、争吵的代价，与它告诉我们的事

这场争吵对英国数学的影响，常被讲得很戏剧化。底本说，它使英国数学在十八世纪相对落后，直到十九世纪才在巴贝奇等人的推动下“翻译回”莱布尼茨体系。这个说法大体可以成立，但要稍加限定：十八世纪英国并非没有出色的数学家（泰勒、麦克劳林都在这一时期作出了重要工作），“落后”主要是指在分析学的大规模发展上，英国与以伯努利家族、欧拉、拉格朗日为代表的欧陆主流隔开了。1812 年前后，剑桥的几位青年学生——巴贝奇、约翰·赫歇尔、皮科克等——组织了“分析学会”，1816 年翻译出版了法国拉克鲁瓦的微积分教程，把莱布尼茨的 d 记号重新带回英国大学。

这件事可以从本书的视角读出一层意思。

符号不是中性的外衣。同一个数学内容，用牛顿的点号写，与用莱布尼茨的 d 写，算起来没有区别；可是当问题变复杂——多个变量、高阶导数、变量替换——莱布尼茨的记号会“替你思考”，牛顿的记号则要你自己记住每一步在做什么。就像两本同样内容的菜谱，一本写“炒至断生”，一本写“大火翻炒 90 秒至表面变色”，老师傅看哪本都一样，新手只能靠后一本。

更深的一层是：英国的数学家把记号当成了忠诚的标志。坚持点号，是为了捍卫牛顿；改用 d，就像背叛。于是一个本该被检验、被改进的工具，变成了一个身份符号。按 123 原理来看，这可以读作“回写”被堵住的一个样本：欧陆那边，新符号不断被使用、被扩展，回写进一代代数学家的底盘，演化出常微分方程、变分法、偏微分方程；英国这边，旧符号被当作需要守护的成就，矛盾被压平，循环慢了下来。

还有一层容易被忽略。优先权之争之所以变得这样激烈，是因为“谁先发现”开始变得重要了。十七世纪下半叶，皇家学会（1660 年成立）、巴黎科学院（1666 年成立）以及《哲学汇刊》《博学者学报》这样的刊物先后出现，知识开始需要被公开、被登记、被裁决。底本说，这场争论体现了科学从“个体发现”走向“制度化知识”的过程。这话对，但不全是好消息：制度化让知识可以积累、可以检验，同时也带来了新的问题——谁有权裁决？那个委员会的报告，正是由当事人之一起草的。命名权与判错权落在谁的手里，从那时起就是科学的一部分（本书第32章还会回到这个问题）。

◆ 六、发动机加语言

回头看，牛顿与莱布尼茨各自完成了什么，又一起完成了什么？

牛顿完成的是：把“变化率”与“累积量”的互逆关系，接到了运动与力上，让它第一次驾驭了一个完整的自然领域。没有这一步，微积分可能只是一套精巧的求面积技巧。

莱布尼茨完成的是：为这一互逆关系造了一套可以脱离发明者而运转的符号，让它变成一门可以教、可以传、可以被许多人共同推进的演算。没有这一步，微积分可能只是一位天才手里的绝技。

两者合起来，才有了本书所说的那个圈：在现实里截取变化，在理念里累加成整体，再回到现实的一点上去算“此刻该如何”。这个圈从十七世纪末开始高速运转，在十八世纪由欧拉、拉格朗日、拉普拉斯等人推到天体力学、流体、振动、热的各个角落。

但这里仍要守住限定。“没有微积分就没有现代科学”的说法，应当收窄为“没有微积分，就没有以动力学方程为核心的那一类现代物理科学”。化学的质量守恒、生物学的分类与进化、统计学的兴起，都有各自的路径（化学一例见第5章）。

还有一件事没有解决：那个在分母里时而存在、时而消失的无穷小，究竟是什么？牛顿与莱布尼茨都没能回答。圈已经转起来了，焊缝却还没有焊牢。在讨论那个“幽灵”之前，下一章要先回答一个长久以来缠绕中国读者的问题：既然中国算学曾经那样发达，这一步为什么不是在中国迈出的？

带走的话

牛顿看清了求切线与求面积互逆，并把这层关系接到力与运动上，让它第一次驾驭一个完整的自然领域；莱布尼茨为同一层关系造了 d 号与积分号，让它脱离发明者、变成人人可学的演算。两半合起来，那个“截取—累加—落回一点”的圈才转起来。但“没有微积分就没有现代科学”只对以动力学方程为核心的那一类现代物理科学成立，而“让我们算一算”只能裁决已经写成符号的那一部分分歧。

小账

- 本章学开车、厨具、菜谱、邻居争院墙等均为构想的比喻，不对应真实个案。
- 本章所述牛顿生年的新旧历、各著作的写作与出版年份、优先权之争的时间线，依据通行的科学史研究；莱布尼茨两个记号的手稿日期与《博学者学报》《通信集》的译名见附录I第I54条；“奇迹之年”出自牛顿事后的回忆，本章已注明其有争先成分。
- 本章采用科学史界“两人独立发明”的共识，不判定两人人品高下；对英国十八世纪数学“落后”的说法已作限定。
- 本章主张的推翻条件见附录G第G19、G20条；“发动机”一说（与第5章共用）见第G53条。

08

第 8 章

李约瑟之问再审

> 这一章的问题：中国有秦九韶、李冶的算学，徐光启也译出了《几何原本》，数学化的自然研究却没有在这里长成——“缺微积分”能算答案吗？

同一条街上开着两家面馆，二十年前一般大，一般热闹。如今东头那家开了十几家分店，西头那家还是老样子。街坊们聊起来，各有各的说法：有人说东头老板舍得在汤底上下功夫；有人说东头那块地正好挨着后来修的地铁口；有人说西头老板的儿子去考了公务员，没人接班；还有人说西头那家其实一直赚钱，只是不想扩张，“落后”这两个字本身就说错了。每一种说法都有几分道理，每一种说法单拿出来又都不够。最容易出错的，是那个一开口就说“根本原因只有一个”的人。

李约瑟之问，就是放大到文明尺度上的这条街。问题的分量太重，答案的诱惑也就太大：谁能给出“那一个根本原因”，谁就像握住了历史的钥匙。本书的底本在这里也伸过手。它在第三章末尾写道，中国长期科技领先却没有诞生现代科学，答案“指向一个更深的数理事实”——缺了把差异符号化、再从差异重建整体的那个内核，也就是微积分。这个判断有锋芒，也有思想史的来路；但照原样写下去，它会在史学上失守，在逻辑上打转。本章要做的，是把它从“根因”收窄为“候选”，同时说清楚收窄之后它还剩下什么、靠什么来检验。

需要交代一句：底本原作是在 SIO 框架下写的，它的原话是“中国文明重整体，但缺乏差异的数学表征”，并把“中庸”的思维习惯也算进原因里。这类说法把一种文明的思维性格当作科学史的解释变量，很难被独立测量，也很容易在

任何结果出现之后被拿来自圆其说。本书不再沿用这种说法。底本自己在末尾的“补篇”里已经做了收窄的工作，本章就以补篇的材料为骨架往下写。

◆ 一、问题本身是怎么提出来的

先回到提问的人。李约瑟（Joseph Needham）在《大滴定：东西方的科学与社会》（1969）里把问题写成两句：一是为什么从公元前一世纪到公元十五世纪，中国文明在把自然知识用于实际需要方面比西方更有效；二是为什么现代科学只在欧洲兴起，而没有在中国（或印度）文明中发展出来。值得注意的是他给“现代科学”下的定义：关于自然之假说的数学化（the mathematization of hypotheses about Nature），以及与之相伴的技术。

这个定义很要紧。它说明底本强调数理内核，并不是凭空独创，而是和提问者本人站在同一条线上；科学革命史的经典叙述（柯瓦雷、戴克斯特豪斯《世界图景的机械化》等）也都把“自然的数学化”当作那场革命的枢纽。底本在这一步站得住。

可同一个定义也埋着一个陷阱，后面会专门拆开。这里先记下一个事实：李约瑟自己并没有把答案放在数学上。他给出的主要解释是社会经济的——他称为“官僚封建制”的帝制结构压住了商人阶层的独立兴起，近代资本主义与近代科学相伴而生的那种土壤，在中国没有长成。换句话说，提问的人把问题定义成数学的，给的答案却是制度的。这本身就提示我们：问题里有数学，不等于答案就是数学。

◆ 二、桌面上不止一种解释

把几十年来的主要解释摆开，大致是这样一张桌面（以下是转述，各家原著细节以原文为准）：

一是**制度—经济解释**，李约瑟本人即属此类。

二是伊懋可（Mark Elvin）在《中国过去的模式》（1973）里提出的**“高水平均衡陷阱”**。他的意思不是中国落后，恰恰相反：前现代中国的农业、商业、

运输效率已经高到一定程度，耕地几乎用尽，人口持续增长，劳动力极其便宜，于是投资于省力机械反而无利可图。他举过一个著名的例子：十四世纪的农书里已有水力驱动的大纺车，后来却被弃置不用。伊懋可同时承认，那个时期中国的科学、数学与工程处在世界前列。

三是林毅夫在《李约瑟之谜》（1995，《经济发展与文化变迁》）里给出的**供给侧解释**。他区分两种发明方式：前现代靠工匠和农人的经验试错，人口越多，试的人越多，发明也越多——这正是中国早期领先的来由；现代则要靠受过数学训练、做受控实验的少数人。中国没能从前一种转到后一种，关键在科举：它把最聪明的头脑引向经义与仕途，不奖励数学，也不奖励实验。

四是**国家竞争解释**：欧洲列国并立、彼此相争，任何技术与知识优势都关乎存亡，学者与工匠也可以“用脚投票”换个君主；大一统帝国没有这种压力。琼斯（E. L. Jones）《欧洲奇迹》（1981）等著作常被归入这一路，“国家体系”（states system）这一说法也常被归到琼斯名下；不过琼斯的书同时讲环境、经济与地缘，不是单一的国家竞争论。

五是托比·胡弗（Toby Huff）在《近代早期科学的兴起》（1993）里的**制度—法律解释**：欧洲中世纪发展出法人观念，大学作为自治法人可以合法地保存、讲授、争论自然哲学，怀疑与辩驳被制度化了。

六是彭慕兰（Kenneth Pomeranz）《大分流》（2000）的**地理—资源解释**。他用经济史材料论证：到约 1800 年，中国最发达的江南与欧洲最发达的英格兰在生活水平、市场发育上相差不大；真正拉开差距的，是英格兰的煤矿恰好离需求中心近，以及新大陆提供的大片“幽灵土地”（ghost acres）。如果这一说法成立，那么“中国很早就因某种深层缺陷落后”的前提本身就先被动摇了。

七是席文（Nathan Sivin）的**问题批评**。他在 1982 年的文章《为什么科学革命没有在中国发生——或者它发生了？》（刊于《中国科学》第5期，原刊名 Chinese Science）里指出，“为什么科学革命没有在中国发生”不是历史研究能回答的问题，它往往是拿欧洲的路径当标尺，再去清点别人缺了哪一块；他还认为，按通常的思想史标准，中国在十七世纪有过自己的科学革命：西方数理天

文学传入后，一批中国学者（文中举的是王锡阐）改变了理解天体运动的方式，几何与三角取代了传统的数值—代数程序，他称之为天文学里的一场“概念革命”，只是它没有带来我们以为科学革命会带来的那种社会后果。

八是弗洛里斯·科恩（H. Floris Cohen）《现代科学如何进入世界》（2010）的**多文明汇流解释**：希腊的数学化自然、伊斯兰世界的实验与代数、十七世纪欧洲的特定汇合，谁都不是单独的钥匙。

摆完这张桌面，结论已经很清楚：学界不是独尊一因，而是制度、经济、地理、认知—数理多种解释并存且相互竞争，连“中国何时落后、是否落后”都还在争论。底本把答案锁定在“一个更深的数理事实”，是这张桌面上一个有力的位置，但不是已经胜出的那一个。就像那两家面馆，“汤底”可能真的重要，但谁也没资格说地铁口不重要。

◆ 三、两道风险：循环与史实

收窄不只是谦虚，它有两条硬理由。

第一道是循环风险。 回到李约瑟的定义：现代科学＝自然之假说的数学化。如果这就是定义，那么“中国缺那种数学，所以没有现代科学”几乎是在重述定义：没有数学化的自然研究，所以没有数学化的自然研究。这像是说“这家面馆没扩张，是因为它没有开分店”。话不错，但没有解释任何东西。

> **循环是怎样发生的** 设 M＝“出现了以连续量数学为骨架的自然研究”，S＝“出现了现代科学”。若按定义 $S \Rightarrow M$ 且 M 几乎就是 S 的核心内容，那么由 $\neg M$ 推出 $\neg S$ 只是逻辑上的逆否，不提供历史因果。 要让“缺微积分”成为解释而不是同义反复，至少要满足两条中的一条： （1）独立识别 M 的缺席，并说明它为什么缺席——而这一步追问下去，往往又落回制度、经济、教育； （2）找到“M 的组件已经具备、而 S 仍未出现”或“M 组件不全、S 却部分出现”的对照案例，看 M 是否真有区分力。 这正是第395卷所说的方法层失败条件：多加一个解释项，若不能多出可检验的区分，就只是在整理语言。

第二道是史实风险。 底本有两处说法需要改写。一处说中国文明“缺乏差异的数学表征”，另一处说中国算学“长于处理具体的、静态的量，却不习惯把变化本身抽象为研究对象”，并把汉字的表意性当作不便运算的原因。这几句话碰到史料就站不稳。

刘徽在 263 年注《九章算术》，割圆术里“割之弥细，所失弥少”的思路，是对极限过程的清楚自觉（见第6章）。祖冲之把圆周率夹在 3.1415926 与 3.1415927 之间，并给出密率 355/113，这一精度在世界范围内领先了九百多年（见第6章）。秦九韶《数书九章》（1247）用“正负开方术”数值求解高次方程，其程序与欧洲十九世纪初的鲁菲尼—霍纳方法同构。李冶《测圆海镜》（1248）的天元术、朱世杰《四元玉鉴》（1303）的四元术，是用算筹位置来记未知数的代数方法——这本身就是一套可机械操作的符号系统，它不靠汉字的表意，而靠位置。再说历法：中国历算处理日月五星不均匀运动，用到了高阶内插法（如郭守敬等《授时历》的招差术），处理的恰恰是“变化中的变化”。

所以更站得住的说法是：中国数学不是没有，也不是不处理变化，而是**发达在算法—代数一路**——重程序、重数值、重可执行的算法；它没有走上欧几里得式公理演绎、代数与几何在坐标里焊接、再把连续变化一般化为一门演算的那条路。不是“缺乏数学表征”，而是“走了另一种数学”。

还有一个旁证，值得单独摆出来，因为它恰好能检验“数学组件是否足够”。十四至十六世纪印度喀拉拉学派（以摩陀婆为开端，见第6章）已经得到正弦、余弦、反正切的无穷级数展开，这些属于今天微积分的核心内容；伊斯兰世界同时继承了希腊几何与花拉子米以来的代数。可这两处都没有随之出现欧洲那样的科学革命。如果“有了那种数学就会有现代科学”，这两个案例就是反例；它们至少说明，数学组件哪怕已经出现，也不足以单独推出现代科学。数理条件更像是必要条件的候选，而不是充分条件。

◆ 四、收窄之后，数理解释还剩什么

把“根因”降为“候选”，不等于把它扔掉。它还剩下三样东西，而且这三样都比原来的说法更结实。

第一，它剩下一个**更精确的对象**。不是“中国缺数学”，而是“缺一条特定的谱系”：公理演绎的证明文化，加上代数与几何的融合（笛卡尔解析几何），再加上把连续变化当作普遍可运算对象的符号演算。这三股在十七世纪欧洲汇在一起，微积分才长出来。打个比方：不是说西头面馆不会做面，而是说它没有走“中央厨房＋标准化配方＋连锁”的那条路。这条路本身是什么、要哪些部件，可以讲清楚。

第二，它剩下一个**更合适的位置**。底本补篇已经给出了很好的改写：微积分之所以关键，也许不在于它是现代科学的原因，而在于它是现代科学这种看世界方式——在差异中读关系、再把关系积回整体——最集中的显露。它是这套认知方式的结晶，未必是第一推动。用本书的主命题说，微积分把“在现实里微分、在理念里积分、再回到现实里微分”这个圈第一次写成了可传授、可运算、可检验的形式。圈能被写下来，与圈怎样在一个社会里转起来，是两件事；后者要靠制度、经济、地理一起回答。

第三，它剩下一个**和其他解释的接口**。制度解释说科举不奖励数学，这正好解释了算法传统为什么没有被推向演绎化；经济解释说省力创新无利可图，这正好解释了为什么精密的数学没有被大规模接到工程与机械上；地理解释说煤和新大陆改变了欧洲的约束，这正好解释了为什么欧洲的数学化自然研究很快就有了需要它的现实落点。数理解释与这些解释不是你死我活，而是可以咬合的齿轮。

用 SDE 的话说，这正是一个“理解一个场域”的任务，起手宜走 E→S→D：先铺开那个时代的纠缠土壤——现实界的煤与水运、理念界的科举与法人制度、自我界的士人抱负与工匠身份；再辨认在这片土壤里显露出来、并且稳定下来的知识形态——算法型的数学、经验试错型的发明；最后才看差异序列往哪里推、在哪里被吸收。这样做的好处，是不把“数学”从土壤里单独拔出来当作第一

因。123 原理提醒过：S 是矛盾的结算点，不是起点。把“数学形态”抬成自己会动的第一因，就是把结算点误当成了发动机。

这里还要重申第4章已经立下的一条禁令。第395卷已经明确：“符号态封顶”只能作为针对具体文本、课程、学习任务的局部、可撤回诊断，**禁止用于整体评判一个文明**，也不能单独解释科学在不同社会的发展差异。本书遵守这一条。说“中国文明的信息停在符号态、上不到数学态”，既违背史实（上面列的秦九韶、李冶就是反例），也违背口径。

◆ 五、一个现成的对照：《几何原本》来了以后

历史给过一次近乎实验的机会。1607 年，徐光启与利玛窦合译《几何原本》前六卷刊行；此后西方天文历算陆续传入，崇祯年间修成《崇祯历书》，清初又有梅文鼎、王锡阐等人会通中西，康熙朝编成《数理精蕴》。演绎几何这块“缺的部件”，在十七世纪就已经摆上了中国学者的书案。

如果“缺那条演绎谱系”是充分的解释，那么部件到了，事情就该接着往下走。可实际情形是：会通工作做得相当出色，徐光启本人也看重《几何原本》“不必疑、不必揣”的推理品格，但它主要被接进了历算这个既有的土壤——为朝廷修历服务，围绕钦天监展开；它没有长成一个独立的、以自然哲学为对象、靠同行争论推进的共同体。徐光启在崇祯二年（1629）的修历奏疏里提出“度数旁通十事”，设想把度数之学推广到历法之外的十个方面——一般概括为天文气象、水利、音律、兵法与军事工程、会计理财、营建、机械、舆地测量、医药、钟漏计时，这个设想后来并没有在制度上落地。

这个对照说明了两件事。其一，数学部件的到位不等于数学化自然研究的发生；部件要被接住，得有接住它的场。第395卷讲介生态时说过，介生的要害是“承接”：差异能进入，后果能保留，局部结构能形成，结构能被后续证据修订。十七世纪的中国，差异是进来了，但能保存、讲授、争论它的那种建制没有跟上，后果保留不下来，局部结构也难以被反复修订。其二，这恰好把数理解释和制度解释连在了一起：问题不只在“有没有那种数学”，还在“那种数学进来

之后被放进了哪一格”。好比一家面馆买回了中央厨房的设备，却仍按老规矩只给老主顾一锅一锅地煮，设备再好，也长不成连锁。

当然，这只是一个对照，不是定论。明清易代的震荡、礼仪之争后传教渠道的收缩、乾嘉学术的转向，都可能在其中起作用；各因素的分量要靠专门的史学研究来定。本书只借它说明：数理候选需要被放进具体的历史土壤里去检验，不能悬空成立。

◆ 六、怎样检验一个“候选”

把数理解释降为候选，就得给它配一份检验单。一个解释要想留在桌面上，至少要能回答下面几种对照：

其一，**组件对照**。在有算法—代数传统、也接触到演绎几何的社会（如明末清初传入《几何原本》前六卷之后的中国、伊斯兰世界、喀拉拉学派所在的印度南部），数学化自然研究的推进是否与“演绎—代数—连续量”三股的汇合程度相关？如果相关性很弱，数理候选的分量就要下调。

其二，**时间对照**。如果彭慕兰等人的经济史结论成立，即到十八世纪末差距才明显拉开，那么任何“很早就注定”的解释都要面对时间上的不吻合。数理解释必须说明，它所指的那种缺席，是在哪一段时间里起作用的。

其三，**制度可剥离性**。如果在控制了科举、法人制度、国家竞争等变量之后，数学传统的差别不再解释任何剩余差异，那么数理解释就只是制度解释的一个中间环节，而不是独立的一格。这不丢人——承认它是中间环节，本身就是一个可检验的结论。

这三条不是要把数理解释判死，而是给它一个被判错的机会。一个永远不可能被判错的解释，在第395卷的口径里，恰恰是最没有分量的解释。

◆ 七、回到那条街

那两家面馆，后来街坊们达成了一种比较诚实的共识：汤底确实重要，但汤底不是全部；地铁口、接班人、老板想不想扩张，样样都在起作用；而且“落后”这个词，得先说清楚是拿谁的标准在量。

李约瑟之问也该这样读。数理的视角留得下来，而且很有价值：它让我们看见，现代科学之所以是那个样子，和一种把变化写成可运算关系、再积回整体的数学形态分不开。但它只是诸多解释里的一种，它要和制度、经济、地理的解释咬合起来，才能讲出一个完整的故事；它还得自带失败条件，才配叫解释。

不过，就在欧洲那条路上，微积分本身也并不干净。它一出生，就带着一个说不清的东西——无穷小。那个东西既是零又不是零，被一位主教讥为“消逝量的幽灵”。下一章要讲的，就是这个幽灵怎样逼着微积分把自己重新做了一遍。

带 走 的 话

“缺微积分”不是中国没有现代科学的根因，只是桌面上的候选之一，与制度、经济、地理解释并列。它带着两道风险：按李约瑟的定义，“缺那种数学所以没有现代科学”近乎同义反复；中国算学在算法—代数一路极为发达，“缺乏数学表征”不合史实。收窄之后它仍有用处，但必须经得起组件、时间、制度可剥离性三种对照，才配留在桌面上。

小 账

- 本章开头与结尾的两家面馆是构想的比喻场景，不对应真实店铺。
- 本章对李约瑟、伊懋可、林毅夫、琼斯、胡弗、彭慕兰、席文、科恩诸家观点的介绍均为转述，不是原文引语；琼斯的归类、席文的论点与徐光启“度数旁通十事”的核验见附录I第I25—I27条。
- 本章不裁定各解释的分量，也不主张中国何时落后、是否落后已有定论；不用“符号态封顶”评判任何文明。
- 本章主张的推翻条件见附录G第G21条；组件对照意义上的失败条件另见第9章“这一编怎样算错”第2、3条。

09

第 9 章

无穷小的幽灵与严格化

> 这一章的问题：师傅说窗框“差一丢丢”，厂家却要一句“误差不超过两毫米”——微积分那个说不清是不是零的“一丢丢”，是怎样被改成可以核对的承诺的？

家里装修，师傅量窗框，说“差不多一米二”。你问到底是多少，他说“一米二差一丢丢”。你再问一丢丢是多少，他有点不耐烦：“一丢丢就是一丢丢，小到不用管。”可做窗的厂家不吃这一套。厂家要的是一个数，外加一个说法：“误差不超过两毫米。”有了这句话，“一丢丢”才算有了着落——它不是一个神秘的小东西，而是一份大家都能核对的承诺。

微积分诞生后的一百多年，就一直靠着那个“一丢丢”过日子。牛顿叫它“正在消失的增量”，莱布尼茨写成 dx。用起来极其顺手，算行星、算抛物线、算悬链线，处处灵验；可只要有人追问“它到底是零还是不是零”，回答就开始支支吾吾。本章讲的，是这个“一丢丢”怎样被一位主教当众点破，又怎样在一百多年后被改写成一份可以核对的承诺——以及在这个过程中，微积分自己被重新做了一遍。

◆ 一、主教的那一刀

1734 年，爱尔兰克洛因主教乔治·贝克莱出版了一本小书，书名很长，通常简称《分析学家》。副题说它是写给“一位不信神的数学家”的——一般认为指的是哈雷，也有人认为是牛顿本人。贝克莱的用意并不全在数学：当时有人嘲笑神

学充满说不清的神秘，他就反问，你们数学家的流数和无穷小，难道不也是靠信仰撑着？

他的批评集中在一个计算上。求 x^2 的流数时，人们先让 x 增加一个小量 o，得到 $(x+o)^2 - x^2 = 2xo + o^2$，两边除以 o，得 $2x + o$；然后说 o 消失了，于是结果是 $2x$。贝克莱指出：除以 o 的时候，你假定它不是零；扔掉它的时候，你又假定它是零。同一个量在同一段推理里先不为零、后为零，这在逻辑上是不允许的。至于牛顿说的“最终比”——量在消失的那一刻的比——贝克莱讥讽道，这些既不是有限量、也不是无穷小、又不是零的东西，岂不就是“消逝量的幽灵”（the ghosts of departed quantities）？

最尖锐的一点是：贝克莱并不否认结果是对的。他承认微积分得出的切线、面积都正确，但他认为这是“两个错误相互抵消”的结果。这就像一个记账的人，两笔账都记错了，月底一对，居然平了。账是平的，可你敢把公司交给他吗？

用本书的话说，这是一次典型的“反常”。微积分在理念界已经积成了一个运转良好的整体，可在它最基础的那一点上——“此刻这个 dx 是什么”——它交不出一个站得住的说法。贝克莱做的事，正是把这个整体按回到一个具体的点上，逼它在这一点上说清楚。

◆ 二、一百年的修补

贝克莱的书引来了不少回应。苏格兰数学家麦克劳林在 1742 年出版了厚厚的《流数论》，试图用古希腊穷竭法式的几何论证给流数打地基，严谨是严谨了，却笨重得几乎没人照着用。欧拉干脆另辟蹊径：他把无穷小看作“真正的零”，导数是两个零的比，比值可以是确定的数。这在今天看来不够严格，但欧拉凭着惊人的直觉，用这套办法算出了海量正确的结果，也把微积分推进到无穷级数、变分法、复数函数的广阔天地。

拉格朗日在 1797 年的《解析函数论》里换了思路：既然无穷小说不清，那就不要它，把每个函数都写成幂级数，导数就定义为级数里一次项的系数。这条路

的问题在于，并不是每个函数都能展开成幂级数——他把一个特殊的情形当作了普遍的前提。

这一百年很像一家生意红火、账目一团糟的小店：店主忙着进货卖货（欧拉），会计想按老派复式记账重做一遍（麦克劳林），另一个会计干脆规定“只收整钱”来回避零头（拉格朗日）。生意没停，账却一直没理清。

这里值得停一下。为什么一门基础不清的学问能繁荣一百年？因为它在应用的那一端一再得到印证——行星的位置、炮弹的落点、钟摆的周期，每一次落回现实的具体点，结果都对。这正是第14章要讲的“再微分”：模型在一个个落点上被检验。但应用端的成功不能替代基础端的证明。按第395卷的证据责任，经验判断与数学命题是两类主张，彼此不能豁免。微积分在物理上“管用”，并不等于它在数学上“成立”；贝克莱抓住的，正是这两者之间的缝。

反过来也要说一句：贝克莱的批评之所以能刺痛人，恰恰因为数学家自己也认这个账。如果当时的数学共同体只看结果、不问理由，主教的书不过是一本护教小册子，翻过就算。正是因为“证明要讲得通”已经是数学这片土壤里沉淀了两千年的规矩——从欧几里得那里一路传下来——一个外行的逻辑质疑才会被当成必须回应的问题。这就像一家老字号，有人挑剔它的账目，它之所以坐不住，是因为“账要对得上”本来就是它自己立的招牌。批评能起作用，靠的是被批评者土壤里早有接住它的东西。

◆ 三、柯西与魏尔斯特拉斯：把“趋近”改写成一句承诺

转机出现在十九世纪。捷克数学家波尔查诺在 1817 年已有接近现代的连续性定义，但影响有限。法国的柯西在 1821 年的《分析教程》里系统地以“极限”为核心重建微积分：导数是差商的极限，积分是和式的极限，连续是“自变量的无穷小增量引起函数的无穷小增量”。柯西的文字里仍然保留着“无穷小”这个词，但他把它解释为“以零为极限的变量”——不再是一个神秘的静止的小东西，而是一个过程。

真正把“一丢丢”改写成承诺的，是十九世纪中叶在柏林讲课的魏尔斯特拉斯。他把“趋近”彻底换成了不等式。

> **ε-δ：一份可以核对的承诺**　“当 x 趋近 a 时，$f(x)$ 趋近 L”，写作 $\lim_{x\to a} f(x)=L$，其严格含义是：
> $\forall \varepsilon>0,\ \exists \delta>0,$ 使得当 $0<|x-a|<\delta$ 时，$|f(x)-L|<\varepsilon$. 读成大白话：你随便提一个误差要求 ε，多严都行；我总能给出一个范围 δ，只要 x 离 a 不超过这个范围（又不等于 a），$f(x)$ 离 L 就不超过你的要求。　回到贝克莱的例子：$\frac{(x+h)^2-x^2}{h}=2x+h$（$h\neq 0$）。要使 $|(2x+h)-2x|<\varepsilon$，只需取 $\delta=\varepsilon$。整个过程中 h 从未等于零，也从未被“扔掉”；我们只是证明了差商与 $2x$ 的距离可以被压到任意小。贝克莱指出的矛盾，在这个写法里根本不出现。

这就是开头那个做窗厂家的逻辑。厂家不问“一丢丢到底是什么东西”，它问的是：你要多准，我就能做到多准吗？极限从此不再是一个“东西”，而是一场你出题、我应答的博弈，一份对任意误差都能兑现的承诺。

这次改写有一个常被忽略的特点：**它把运动从定义里拿掉了。**　牛顿的流数是在时间里流动的量，“趋近”本来是一个动词；可在 ε-δ 的定义里，没有任何东西在动，只有一组静态的量词和不等式。“当 x 趋近 a”这句话，被翻译成了“对每一个 ε 都存在一个 δ”。这是一次彻底的截取：把直觉里那个“正在消失”的过程整个切掉，只保留一个可以逐条核对的逻辑结构。

用第10章将要细讲的说法，截取就是丢弃。ε-δ 换来了严格，丢掉的是直观。此后一百多年，每一代学微积分的学生都要在这里摔一跤：他们心里想的是“越来越近”，纸上写的却是“对任意……存在……”。这道沟，就是这一刀切出来的。

◆ 四、严格化逼出的新东西

严格化不只是给旧结果补上证明，它还逼出了旧框架里看不见的东西。

一个例子是“连续却处处不可导”的函数。直觉告诉人们：一条连续的曲线，大部分地方总该有切线吧？1872 年，魏尔斯特拉斯在柏林科学院给出一个函数——由无穷多个越来越密、越来越小的余弦波叠加而成——它处处连续，却在任何一点都没有导数。这条曲线画不出来，因为它在任何尺度上都布满锯齿。当时有数学家称这类函数为“怪物”。可半个多世纪后，人们发现布朗运动的轨迹几乎必然就是这样的曲线（见第19章）。被当作怪物的东西，后来成了一类自然现象的标准模型。

另一个例子是实数本身。ε-δ 要成立，需要“数轴上没有洞”：一个越来越收敛的序列，必须有一个实数作为它的极限。这件事过去被当作不言自明，严格化以后才发现它需要被构造出来。1872 年前后，戴德金用“分割”、康托尔用收敛序列分别给出了实数的构造。于是，微积分的地基从“几何直观中的连续量”换成了“由有理数构造出的实数系统”。

还有一个例子，揭示了严格化的必要。柯西在《分析教程》里曾断言：连续函数构成的级数若收敛，其和也连续。1826 年，阿贝尔在一篇论二项式级数的论文里指出这个说法有例外，举的是级数 $\sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \cdots$：每一项都连续，和在 $-\pi < x < \pi$ 内等于 $x/2$，到 $x = \pi$ 处却跳回零，在 π 的奇数倍处断开。后来人们分清了“逐点收敛”与“一致收敛”，才把这个缺口补上。这说明，没有严格的语言，连大师也会在直觉上摔跟头。就像菜谱上写“盐少许”，老厨师照着做没问题，换一个新手、换一口锅、换一种盐，味道就可能全变；严格的语言，相当于把“少许”换成“3 克，适用于两人份”，并且注明在哪些情形下这个数字不再适用。

这三件事放在一起看，就是 123 原理的一次完整运转。① 差异在累积：微积分的推广（D）不断碰到旧的直观土壤（E）承接不了的问题——无穷级数的怪异行为、傅里叶级数能否表示任意函数的争论、贝克莱那样的逻辑质疑；② 张力逼出显露态的改变：分析学的基础被重写，极限、实数、一致收敛成为新的稳定显露；③ 回写：新的基础反过来重置了“什么算一个合格的问题”——从此“这个函数可导吗”“这个级数一致收敛吗”成了必须先问的事；它也沉淀成新的土壤，此后的教科书、课程、研究规范都在这片土壤上长出来。回写也有负向的一

面：它把“直观上的流动”从合法推理里驱逐出去，很多学生从此对微积分只剩下符号操作，丢掉了牛顿那种“看见运动”的能力。

◆ 五、极限从哪里来：哲学解释的分寸

底本在附录里对极限有一段发生学解读，大意是：极限是差异序列在稳定化之后的收敛；ε-δ 规定了“要多稳定才算收敛”，却没有创造稳定性本身，稳定性来自存在的差异结构。底本还说，芝诺的“阿基里斯追不上乌龟”之所以是悖论，是因为只看见差异的无限分割，没看见无穷多段可以累积成一个有限的整体。

这段解读有它的价值，但要放在合适的位置上。

先说它说对了什么。ε-δ　定义确实不保证任何一个具体的极限存在，它只给出“存在”是什么意思。一个具体的差商序列收不收敛，要看那个函数本身。比如测量一辆车的速度，你把时间间隔从一秒缩到十分之一秒、百分之一秒，读数如果越来越稳，你就有理由谈“此刻的速度”；如果读数越缩越乱——像测一粒花粉在水里的抖动——那“此刻的速度”就不存在。从这个角度说，“可以被微分”是对象的一种性质，不是定义送给你的礼物。这与本书第2章讲的 S 的“三个可”同一个思路：一个特征要显露，得可识别、可维持、可表达；导数存在，正是差商这一特征在越来越细的尺度上“可维持”。

再说它不能做什么。第一，它不能替代证明。一个函数在某点是否可导，只能用数学的办法证明，不能用“差异已经稳定”这句话宣布。第二，它不能裁决芝诺。芝诺悖论在数学上的标准处理，靠的是收敛级数与实数完备性，例如 $\sum_{n=1}^{\infty} 2^{-n} = 1$；哲学上关于“完成无穷多步是否可能”的争论，至今仍有人在讨论。说“微积分一举解决了困扰两千年的难题”，是把数学处理当成了哲学终审。第三，它不能把“潜无穷”与“实无穷”的争论简单地调解掉。直觉主义者与形式主义者在这个问题上的分歧，涉及什么样的数学对象可以被承认存在，不是换一套说法就能化解的。

按第395卷的证据责任，概念规定、哲学解释、数学命题是不同类型的主张。发生学的解读可以帮我们问出好问题——例如“为什么有些自然对象可微，有些不可微”——但答案要由数学和经验各自去给。这是本书在这里给自己划下的线。

◆ 六、后话：幽灵领到了户口

故事还有一个尾声。1960 年代，美国逻辑学家亚伯拉罕·罗宾逊（Abraham Robinson）用模型论构造出一种扩充的数系——超实数。在这个数系里，真的存在比任何正实数都小、却又不等于零的正数，也就是严格意义上的无穷小。罗宾逊证明，只要遵守一条“转移原理”，在超实数里做的推理与在实数里做的推理结论一致。于是，莱布尼茨式的 dx 演算，可以在严格的意义上被重新合法化。这套理论叫非标准分析。

> **超实数里的导数**　在超实数系 $^*\mathbb{R}$ 中取一个非零无穷小 h，导数可定义为 $f'(x) = \mathrm{st}\left(\frac{f(x+h)-f(x)}{h}\right)$，其中 st（标准部分）把一个有限超实数映到与它无限接近的唯一实数。对 x^2：差商为 $2x + h$，其标准部分为 $2x$。h 始终不为零，“扔掉 h”这一步被替换为“取标准部分”这一个明确定义的操作。贝克莱的“既为零又不为零”，在这里被拆成了两个不同的东西：h 本身不为零；h 的标准部分为零。

这个尾声很有意思。它表明，十九世纪的那次严格化并不是唯一可能的出路：同一个“幽灵”，可以被驱逐（ε-δ），也可以被安置（超实数）。两条路给出的定理在标准范围内是一致的，选择哪一条，部分取决于教学和研究的方便，部分取决于数学共同体的习惯。这不是说数学可以随意，而是说：**同一组差异，可以在不同的理念土壤里稳定成不同的显露**。这与本书第2章讲的“差异先于实体”是相通的。

但也要说清楚，非标准分析的出现，并不意味着莱布尼茨当年就已经是对的。莱布尼茨没有转移原理，也没有超实数的构造；他用的是一种富于启发的直

觉，而不是证明。说“罗宾逊为莱布尼茨平反”，是一种常见但需要打折扣的说法。直觉被后来的严格理论接住，与直觉本身已经严格，是两回事。

◆ 七、收束：微积分也要被再微分

回看这一百多年，可以看到一个反复出现的结构。微积分作为一个整体，在理念界越积越大；可是每当有人把它按到一个具体的点上——此刻的 dx 是什么，这一个级数收不收敛，这一条曲线有没有切线——它就得在那一点上重新回答问题。回答不了的地方，就是反常；反常累积到一定程度，基础就被重写。

本书在后面把“模型落回现实具体一点”称为“再微分”。严格化不是一次物理意义上的落地，而是数学内部的事；把它比作“微积分自己被再微分了一次”，是一个类比，帮助我们看见同一个形态：整体模型必须在具体点上兑现，而兑现的那一刻，总要借来一些整体里没有的东西——对贝克莱的问题来说，借来的是实数的完备性、量词的严格语言，以及一个数学共同体愿意接受的证明标准。

下一章，我们离开历史，转入机理。先从微分讲起：它在现实里截下一刀，截下的是什么，切掉的又是什么。

◆ 八、这一编怎样算错

第二编（第5—9章）讲了微积分怎样发生。以下各条一旦成立，本编相应的主张须收回或收窄：

1. **若**史学研究表明，十七世纪欧洲动力学科学的关键进展（如开普勒问题的解决、地面力学与天体力学的统一）在微积分形成之前已经以等价的几何方法完成，且后来的推进并不依赖微积分的符号演算，**则**第5、7章把微积分视为动力学科学“发动机”的判断，须收窄为“重要工具之一”。
2. **若**能找到一个文明，它具备了演绎证明、代数—几何融合与连续量符号演算这三股传统，并有相应的承接建制，却仍未出现数学化的自然研究；或反

之，三股俱缺却出现了，**则**第8章所保留的“数理候选”须进一步下调，甚至撤出桌面。

3. **若**控制科举、法人制度、国家竞争、资源地理等变量之后，数学传统的差别不再解释任何剩余差异，**则**第8章须承认数理解释只是制度解释的中间环节，不能作为独立一格。
4. **若**对刘徽、祖暅、阿基米德等人的工作的细读表明，他们已经系统处理了瞬时变化率，**则**第6章“积分的一半已经被严格地走通，但它停在静态图形上、停在逐题构造上、停在单向上”的判断须撤回或收窄。
5. **若**数学史表明，十九世纪分析严格化的主要动力来自教学与学科建制的需要，而与应用中出现的反常（级数、傅里叶分析、不可导函数等）关系不大，**则**第9章用 123 原理解释严格化的叙述须降为“一种可能的组织方式”，不作机制断言。
6. **若**“再微分”这一类比在第9章的用法中，不能比纯粹的史学叙述多给出任何可检验的区分（例如预测哪些地方会先出现基础危机），**则**按第395卷方法层失败条件，这个类比应从论证中撤下，只作修辞保留。

带 走 的 话

微积分在应用端处处灵验，不等于它在基础上成立；贝克莱抓住的正是这道缝。ε-δ 把“趋近”改写成对任意误差都能兑现的不等式承诺，换来严格，丢掉直观；它只规定“极限存在”是什么意思，不保证哪一个极限存在。超实数表明，同一个“幽灵”既可以被驱逐，也可以被安置。发生学的解读能帮人问出好问题，但替代不了证明，也裁决不了芝诺。

小 账

· 本章的量窗框、小店记账、老字号对账、菜谱“盐少许”均为说明用的比喻，不对应真实个案。

· 本章所述年份、著作与数学史实（贝克莱、麦克劳林、拉格朗日、柯西、阿贝尔、魏尔斯特拉斯、戴德金、康托尔、罗宾逊）属数学史通行记载，附录I未单列条目；方框中的数学是标准教科书内容。

· 本章不主张“罗宾逊为莱布尼茨平反”，不主张微积分了结了芝诺悖论的哲学争论；把严格化称作“微积分被再微分”只是类比，不作机制断言。

· 本章主张的推翻条件见附录G第G22条；第二编的失败条件见本章“这一编怎样算错”，已登记为附录G第G17—G22条，以及第G53条（本节第1条，“发动机”一说）、第G54条（本节第6条，再微分类比）。

第 三 编

机理：现实的微分，理念的积分，落地的再微分

编 序

第三编是全书的核心，只讲一个圈：现实的微分，理念的积分，落地的再微分。

第10章讲微分：在现实里截一刀。偏导数是“其他条件不变”的数学形式，截下的那一刀很忠实，也必然切掉了这一点周围的纠缠。第11章讲积分：在理念里拼回整体。整体从未被直接经验，只被积分出来；斯托克斯定理说边界管得住内部的总账，管不住内部的分布。第12章讲积分常数 C：初值问题、边值问题和阿达马的适定性三问，说明微分丢掉的那一份积分找不回来，只能落地时外借；并以复现危机为检验把手，写明这个主张怎样会错。第13章讲基本定理：它是一次真焊接，焊住了变化，没焊住起点，条件一破，焊缝就开。第14章是全书的枢纽，讲再微分：技术与应用就发生在整体模型落回具体一点的那一刻，这一刻要借回 C，也会生出新的差异，回写有正有负。第15章用三方程看这个圈的静态关系：模型由切法与土壤共同生成，模型又规定下一次怎么切；并交代加上 D 和 E 之后多出了哪些可检验的区分。第16章讲让圈转起来的 123 原理：从水星近日点与黑体辐射，看反常怎样累积、逼出改变、再回写测量与问题，并接住库恩、波普尔、费耶阿本德。

读这一编时，可以带着一个问题：那个 C，在每一个现场分别是什么，由谁去借？

10

第 10 章

微分：在现实里截一刀

> 这一章的问题：厨师按住火候和水量只动盐，试出“加一撮咸一点”，换一锅汤这句话却可能不灵——微分截下了什么，又切掉了什么？

炖一锅汤，尝一口，觉得淡了，加一小撮盐，再尝。这一撮盐带来的“咸了一点”，就是厨房里最朴素的微分：在一个具体的时刻，问一个具体的变量，多一点点会怎样。厨师不会同时去改火候、换锅、加水、调料一起动——那样尝出来的变化，说不清是谁带来的。他把别的都先按住，只动盐。

可是锅里的东西并不配合。盐多了，汤的鲜味会被吊起来，同一撮盐在大火收汁时和小火慢炖时咸淡不一样，番茄多的汤和骨头汤对盐的反应也不同。厨师按住火候、按住水量去试盐，得到的“加一撮咸一点”，是在“这锅汤、此刻、这个火候”下才成立的一句话。换一锅，这句话可能就不灵了。

本章讲微分。底本把微分称作“差异的忠实抄写”，这个说法抓住了一半：微分确实紧贴着可测的差异，是科学里最硬、最可检验的一段。但它漏了另一半：微分之所以能“抄”下来，是因为它先在现实里截了一刀，把其他一切都按住了。截下来的那一份抄得很忠实，被截掉的那一份，从此不在账上。

◆ 一、微分做了什么：从差商到导数

先把动作看清楚。一辆车在路上走，位置是 $x(t)$。我们能测的，只是两个时刻之间的位置差 Δx 和时间差 Δt。它们的比值 $\Delta x/\Delta t$ 是这段时间里的平均速

度。把时间间隔一缩再缩，如果这个比值越来越稳，趋向一个确定的数，这个数就叫此刻的速度，即导数 dx/dt。

这个过程里有三个动作，每一个都值得单独看。

第一，**选点**。导数总是“在某一点”的导数。它回答的不是“这辆车跑得怎么样”，而是“在 t_0 这一刻，它的位置对时间变化得多快”。这就像医生量血压，量的是此刻、这只胳膊、坐着休息五分钟以后的读数。

第二，**选变量**。导数总是“对某个变量”的导数。dx/dt 问的是位置对时间；如果问位置对油门开度，那是另一个导数。选哪个变量，决定了你把世界的哪一个方向当成“动的那一个”。

第三，**取极限**。差商的极限存在，需要那个差异在越来越细的尺度上保持稳定。第9章已经说过，ε-δ 只告诉你“存在”是什么意思，不保证它存在。测一辆车，越缩越稳；测水中花粉的抖动，越缩越乱，导数就不存在。用第2章讲 S 的“三个可”来说，导数存在，意味着差商这个特征在越来越细的尺度上可识别、可维持、可表达。

底本在这里写的是“特征律保证差异的稳定性，微分把差异的稳定性符号化”。按现行口径，本书不说“保证”：差异是否稳定，是对象自己的事，要一个个去测、去证；说某个对象可微，要说明是在哪个尺度、哪类测量、多大精度下可微。“可微”不是世界送给科学的通行证，是科学在每个具体对象上要挣来的东西。

◆ 二、偏导数：“其他条件不变”的数学形式

真实世界里的量，很少只跟一个变量有关。汤的咸淡，跟盐有关，也跟水量、火候、时间、食材有关。写成函数，就是 $u = f(x_1, x_2, \dots, x_n)$。这时候问“盐多一点会怎样”，数学上就是求偏导数 $\partial u/\partial x_1$：让 x_1 动一点，其他 $x_2, \dots, x_n$ 全部固定。

经济学家有一句拉丁文行话，叫 ceteris paribus，“其他条件不变”。偏导数正是这句话的精确形式。它是现代科学最有力的工具之一：没有它，就没法从一团纠缠的现象里分离出单个因素的作用。实验室里的“控制变量”、临床试验里的“随机分组”，都是在现实里尽力造出一个“其他条件不变”的局面，好让偏导数有东西可量。

但这里有一个代价，底本没有算进去。

> **截取即丢弃：偏导数丢了什么** 设 $u = f(x, y)$。在点 (x_0, y_0) 附近，全增量可以写成 $\Delta u \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{1}{2}\left(f_{xx}\Delta x^2 + 2f_{xy}\Delta x\Delta y + f_{yy}\Delta y^2\right) + \cdots$ 其中 $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$ 是交叉项，描述“x 的作用怎样随 y 而变”。只求 $\partial f/\partial x$，就是令 $\Delta y = 0$。这时交叉项 $f_{xy}\Delta x\Delta y$ 恒为零，不进入任何测量。换言之：**固定 y 求出的 x 的效应，在原理上看不到 x 与 y 的交互**。交互越强（$|f_{xy}|$ 越大），在 (x_0, y_0) 测得的 $\partial f/\partial x$ 搬到另一个 y 值下就越不准：$\frac{\partial f}{\partial x}(x_0, y_1) - \frac{\partial f}{\partial x}(x_0, y_0) \approx f_{xy}\,(y_1 - y_0)$。

用大白话说：你在骨头汤里试出了“加一撮盐咸多少”，这个数换到番茄汤里不一定还对，因为盐和番茄的酸之间有交互。偏导数把这个交互按成了零，按住的那一刻，交互就从你的数据里消失了。

这就是本书说的“截取即丢弃”。微分在现实里截下一刀，截下来的那一面——此刻、此处、沿这个方向的变化率——被忠实地记录了；被切掉的，是这一点周围的纠缠：其他变量怎样参与、它们之间怎样相互改变彼此的作用。用 SDE 的话说，E 是参与发生的土壤，不是背景；偏导数把土壤当成了背景，固定住，于是它就不再出现在结果里。

◆ 三、微分方程：把一刀变成一条规矩

单个导数只是一张切面。科学真正有力的地方，是把切面之间的关系写成方程，说“在任何一点上，这个变化率都等于那样一个量”。这就是微分方程。

底本列了好几个典范，这里挑三个，各看一眼它们按住了什么。

牛顿第二定律 $m\frac{d^2x}{dt^2}=F$ 说：任何时刻，物体加速度乘以质量等于它受的合力。这条规矩极其干净，可它默认了质量不变（火箭烧燃料时就得改写），默认了参照系是惯性的（在转弯的车里就得加一个“离心力”），默认了速度远小于光速。这些默认，就是牛顿在现实里截那一刀时按住的东西。

化学反应的一级速率方程 $\frac{dC}{dt}=-kC$ 说：浓度下降的快慢与当时的浓度成正比。速率常数 k 看起来是一个数，其实它悄悄装着温度、催化剂、溶剂、容器表面这些条件。温度一变，k 就变；同一个反应，在实验室的小烧杯里测出来的 k，放大到工厂的大反应釜里常常不灵，因为传热、搅拌这些在小烧杯里可以忽略的东西，在大釜里全回来了。化工上专门有一门功夫叫“放大”，干的就是把被按住的东西重新请回来。

种群的逻辑斯谛方程 $\frac{dN}{dt}=rN\left(1-\frac{N}{K}\right)$ 说：种群增长先快后慢，最后停在环境容量 K 附近。K 是一个数，但它其实是食物、天敌、气候、疾病缠在一起的结果。把它写成一个常数，就是假定这一团纠缠不变。一场干旱、一种外来天敌，都会让 K 不再是那个 K。

这三个方程有一个共同的结构：**每一个常数，都是一捆被截下来、压扁了的纠缠。** m、k、r、K 看起来是简单的数，其实是“在这些条件下”才成立的数。方程本身写得越干净，这一捆东西就压得越紧，也越容易被忘掉。就像菜谱上写“中火十分钟”，你家的灶和写菜谱那个人家的灶不是一个灶，“中火”这两个字里压着的东西，菜谱上不写。

这并不贬低微分方程。恰恰相反，把一团纠缠压成一个常数，是科学能运转的前提：压不下去，就没有方程，也就没有预测。要警惕的只是一件事——忘了这个常数是压出来的，以为它是世界本来就有的一个刻度。

◆ 四、刀从哪里下：选变量的那一刻

还有一个比“按住什么”更早的问题：这一刀从哪个方向下。

求导之前，得先决定对什么求导。这件事看起来理所当然——速度当然是对时间求导——可在很多学科里，它一点也不理所当然。研究一个孩子的学习，你

是对“学习时长”求导，还是对“睡眠时间”求导，还是对“和老师说话的次数”求导？研究一座城市的拥堵，你是对“车辆数”求导，还是对“红绿灯配时”求导，还是对“地铁站之间的距离”求导？选了哪个变量，就等于认定“这个是值得单独拿出来动一动的东西”，其余的都归入“其他条件”。

这一步，在第395卷的语言里，是命名权的行使：先有人把某个东西命名为一个变量，它才有资格出现在导数的分母上。没有被命名为变量的东西，连“被按住”的资格都没有——它根本不在方程里，既不动，也不被固定，只是不存在。

用一个日常的例子说。家长发现孩子成绩下滑，第一反应常常是“学习时间不够”，于是对“学习时间”求导：多学一小时，成绩能涨几分？这个导数也许真能测出来。可如果真正的原因是孩子和同桌闹了别扭、上课坐立不安，那么这个因素既没被选成变量，也没被列入“其他条件”，它压根不在家长的账上。家长测出的导数没有错，只是问错了方向。

所以，微分的“忠实”还有一个前提：它忠实于被选中的那个方向。方向怎么选，不是微分自己能决定的。它来自研究者的问题、学科的传统、仪器能测什么，以及谁有权说“这个算一个变量”。第32章会把这个问题推到更远的地方：在一个由计算主导的时代，那些从未被命名为变量的东西，连“问题”的身份都很难拿到。

◆ 五、看得见的代价：药、教室与市场

截取的代价不只在厨房里。

吃药就是一个例子。一种降压药在临床试验里表现很好，试验的受试者是经过筛选的：年龄在一定范围，没有严重的肝肾问题，不同时服用别的某些药物。这些筛选条件，就是在现实里造出的“其他条件不变”。药效是在这个被按住的局面里测出来的。到了门诊，来了一位八十岁、同时吃着五种药、肾功能偏弱的老人，那个在试验里测得的“药效”，要搬到他身上就得打问号——不是试验错了，而是试验故意切掉的那些纠缠，在这个病人身上全都回来了。

教育研究也一样。一项教学方法在几所实验学校里提高了成绩，研究者控制了班级人数、师资、课时。推广到别的地区，效果常常打折扣，甚至消失。原因往往不在方法本身，而在被按住的那些条件：那几所学校的老师恰好有热情，家长恰好配合，课时恰好够用。

经济学里的“边际”分析，同样是偏导数。底本举过边际效用的例子 $MU = dU/dq$。可人对一样东西多一份的满足感，跟他手里别的东西有关，跟周围的人有没有、价格是不是在变也有关。把这些都按住求出的边际效用，干净、可算，但它只在被按住的那一圈里有效。

这三个例子有一个共同点：**截取越狠的地方，搬家时越容易出问题**。 一个结论，在原来那个被按住的局面里越精确，搬到一个没被按住的局面里，越可能失灵。这一点会在第12章变成一个可以检验的预测：在依赖“其他条件不变”切得最狠的领域，研究结果的复现应当最困难。这里先把它记下，详细的讨论留给那一章。

◆ 六、泰勒展开：一点上能看到多远

微分只看一点，可人们总想知道这一点附近的整个样子。泰勒展开就是从一点出发往外看的工具。

> **泰勒展开与余项** 若 f 在 a 附近有足够多阶导数，则 $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$，其中余项可写为 $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$，$\xi$ 介于 a 与 x 之间。 两个关键事实：（1）余项 R_n 里的 ξ 不在点 a 上，而在 a 与 x 之间的某处；所以误差的大小，要靠点 a 以外的信息来估计。（2）即使 f 在 a 点有任意阶导数，泰勒级数也可能不等于 f。经典例子：$f(x) = e^{-1/x^2}$（$x \neq 0$），$f(0) = 0$。它在 0 点的各阶导数全为零，泰勒级数恒为零，而 f 在 0 以外处处为正。

底本对泰勒展开评价极高，说“一个函数在某点邻域内的整体行为，竟然可以被它在那一个点上的全部各阶差异完全决定”。这句话需要收窄：它只对一类

特殊的函数成立，数学上叫“解析函数”，并且只在一定的收敛半径之内成立。方框里那个例子说明，即使你在一点上知道了全部各阶导数，仍然可能对这一点以外的样子一无所知。一点上的信息再多，也替不了点外的信息。

这件事用开车来比，很直观。你此刻知道自己的位置、速度、加速度，甚至加速度的变化率，你就能比较准确地预测接下来一两秒车在哪里。可是要预测一分钟以后在哪，这些信息就不够了——前面有没有红灯，旁边有没有车并线，这些不在你此刻的导数里。泰勒展开的前两项 $f(a)+f'(a)(x-a)$ 叫线性近似，工程上用得最多：小角度时 $\sin\theta\approx\theta$，弹簧在小形变时力与伸长成正比，都是这一步。它好用，前提是你不走远。

今天训练大模型用的梯度下降，每一步也是这样一次线性近似：在当前参数这一点上算出梯度，朝下坡方向迈一小步，再在新的点上重新算。步子迈得太大，线性近似失效，训练就会震荡甚至发散；所以工程师要小心地调“学习率”，实际上就是在调“离开这一点能走多远而近似还算数”（见第25章）。

所以，泰勒展开不是“从一点重建整体”，而是**用一点上的局部信息，在一个有限的范围里近似整体，同时自带一张误差的账单**。账单的大小取决于点外的东西。这正是本书主命题的一个缩影：微分在一点上能拿到的，再精细，也有一部分要靠点外来补。

◆ 七、为什么非截不可

读到这里，也许有人会问：既然截取要丢东西，为什么不干脆别截，直接研究整团纠缠？

答案很实际：不截，就什么也抄不下来。一团纠缠的现象，同时有几十个量在一起变，你记下的任何一个变化，都说不清是谁带来的。就像一锅汤里同时加了盐、加了水、开了大火，尝出来“味道变了”，你没法从这一口里学到任何可以带走的东西。下次再做汤，你还是不知道该动哪一样。

截取的好处，是把一个说不清的整体变化，拆成一个个说得清的局部变化。每一个局部变化都可以被重复测量、被别人核对、被写下来传给下一个人。这正

是第3章讲的那张筛子：理念物理界是物理界经过测量、理想化、剔除噪声之后剩下的部分。微分就是筛子上最细的那几个网眼——它只让“一点、一个方向、一个变化率”通过，其余的都拦在外面。拦下来的东西没有消失，它们还在现实里，只是暂时不进账。

而且，截取的代价并不总是很大。在很多物理问题里，交互项确实很小，按住了也无妨；行星之间的引力相互影响很弱，所以把每颗行星单独对太阳算，就已经相当准。截取在这些地方几乎是免费的，这也是经典力学能取得巨大成功的原因之一。问题出在交互项大的地方：生命、社会、心理、生态。在那里，截取的代价陡然上升，同样的刀法，切下来的东西就不那么可靠了。

所以要紧的不是“截不截”，而是“截的时候知不知道自己丢了什么、丢得多不多”。一个好的厨师和一个新手的差别，不在于谁更敢只动盐，而在于好厨师知道这锅汤里盐和别的东西缠得有多紧，知道“只动盐”这个判断在哪里会失灵。

◆ 八、微分的忠实，到底忠实于什么

底本说微分是科学的“忠实部分”：不虚构、不添加，只是把已经稳定的差异关系如实抄写。这一点本书保留，而且认为它是对的——微分方程之所以是科学里最可检验的部分，正因为它紧贴着仪器能记录的差异。

但要补上一句：**微分忠实于它截下的那一面，不忠实于整体。** 它像一张证件照，五官清清楚楚，可它不告诉你这个人走路的样子、笑的样子、生气的样子。证件照没有说谎，它只是在一个被规定好的角度、光线、表情下拍的。拿证件照去认人，大多数时候够用；可要是有人问“这个人是什么样的人”，证件照就答不上来。

还要澄清一个底本里的旧说法。底本把微分对应于 SIO 意义三律中的“创生律”，说微分是“创生律的数学化”。现行口径不再以创生律作主术语，本书也不建立这种一一对应。比较稳妥的说法是：微分在三方程里主要落在“差异”一侧——它抓的是某个方向上的差异序列，在一个点上被稳定下来、被写成符号的

那一刻；而它付出的代价落在“纠缠”一侧——它把这一点周围的 E 当作固定背景处理掉了。这只是一种组织方式，不是严格对应。它的用处在于提醒：每用一次微分，都要问一句“我按住了什么”。

这句追问可以写成一张很短的检查单。用一个导数、一个偏导数、一个边际效应之前，问：

- 这是在哪一点上测的？那一点有什么特殊？
- 哪些变量被固定了？固定在什么值？
- 被固定的变量与我关心的变量之间，有没有交互？大不大？
- 我打算把这个结论用到哪里？那里的被固定变量还是原来的值吗？

四个问题答不上来的时候，微分仍然是忠实的，但你已经不知道它忠实于什么了。

◆ 九、收束

微分在现实里截下一刀。它截得准、截得细，把一个点上、一个方向上的变化忠实地写成了符号；它也因此切掉了这一点周围的纠缠。这不是微分的缺陷，而是它的本性：不截，就什么都抄不下来；截了，就有东西留在刀外。

截下来的一片片切面，单独看都很清楚。可是科学要的不只是切面，它要整体：行星的整条轨道，汤的整锅味道，病人整个疗程的走向。下一章讲积分——人怎样把这些切面在理念里拼回一个整体，以及拼回来的那个整体，究竟是什么。

带走的话

偏导数是“其他条件不变”的精确形式：令其余增量为零，交叉项就进不了任何测量，所以固定一个变量测出的效应，在原理上看不到交互。方程里的常数是压扁的纠缠，泰勒展开的误差要靠点外的信息来估。交互越强，结论搬家越容易失灵。用一个导数之前，先问清楚：在哪一点测的，按住了什么，有没有交互，要搬到哪里去用。

小账

- 本章的炖汤、量血压、开车、证件照、家长为孩子成绩“求导”均为说明用的日常比喻，不对应真实个案；降压药试验与教学方法推广两例是对常见情形的概括，不指某一项具体研究。
- 本章没有使用调查数字；方框中的偏导数展开、泰勒余项与 e^{-1/x^2} 反例是标准教科书内容。
- 本章把微分归在三方程的“差异”一侧、把代价归在“纠缠”一侧，只是一种组织方式，不与“创生律”或任何意义律建立对应；本章也不主张截取处处代价很大，交互项小的物理问题里截取几乎是免费的。
- 本章主张的推翻条件见附录G第G23条；“截取最狠处复现最难”的检验在第12章展开，相关条件见附录G第G1—G3条。

11

第 11 章

积分：在理念里拼回整体

> 这一章的问题：月底把一笔笔流水加起来，得到一个谁也没亲眼见过的“总支出”——积分拼回来的整体算不算真的，它为什么总缺月初那一块？

月底对账，是很多家庭都熟悉的场景。一个月里，钱是一笔一笔花出去的：早上一个包子，中午一份盒饭，下午给孩子交了个兴趣班，晚上网购了一箱水果。每一笔都很小，单拿出来谁都不在意。到了月底，把账单一条条加起来，才发现“这个月居然花了这么多”。

这个“这么多”很有意思。它是真的——银行卡上的余额不会骗人；可它又不是你在任何一个时刻亲眼看见的东西。你看见的只有一笔笔流水，那个“月度总支出”，是你坐下来、拿起计算器、按一定规则加出来的。换一种加法，比如把房贷算进去还是不算，把给父母的钱算作“支出”还是算作“转移”，得出的“总数”就不一样。总数是真的，但总数是加出来的。

本章讲积分。底本说积分是“整体的发明创造”，这个说法抓住了要害：我们从来没有直接经验过整体，整体是被积分出来的。本章沿着这条线往下走，再补上两件底本没有说透的事：积分要成立需要什么条件；积分出来的整体，为什么总缺一块。

◆ 一、整体从未被经验，只被积分出来

先看一个最基本的事实。科学里，我们能直接测的，总是局部的、此刻的量：此刻的速度、此处的温度、这一秒流过的电流。没有任何仪器能一下子“看

见”一颗行星的整条轨道、一座城市一年的用水总量、一个病人整个疗程的身体变化。这些整体，都是从局部的测量累加出来的。

积分就是这个累加的数学形式。知道每一刻的速度 $v(t)$，把它们乘上各自的那一小段时间，再全部加起来，就得到走过的总路程：$\int_{t_0}^{t_1} v(t)\,dt$。知道每一处的密度，在整个体积上加起来，就得到总质量。

底本在这里说得很好：整体“永远不在我们的感官之内”，它是被理念地重建出来的。这正是本书说积分发生在理念界的意思。微分在现实里截一刀，截下来的切面还紧贴着可测的东西；积分则把这些切面搬到纸上、搬到脑子里、搬到计算机里，按一套规则拼起来。拼成的那个整体——一条轨道、一个总量、一条定律——住在理念界，不住在现实界。

这不是说整体是假的。月底的总支出是真的，行星的轨道也是真的。说它“住在理念界”，是说它是经过一套规则加工以后才出现的东西，它的样子取决于规则。就像同一个月的流水，按不同口径加，得出不同的“总支出”。这些总数各自都对，但它们回答的是不同的问题。

◆ 二、积分要成立，需要什么

积分不是随便加一加就行。数学上，一个函数能不能积分、积出来是不是有限的数，都有条件。

> **积分是一个极限，而极限可能不存在** 黎曼积分的定义：把区间 $[a,b]$ 切成 n 小段，在每段 $[x_{i-1},x_i]$ 里任取一点 ξ_i，作和 $S_n=\sum_{i=1}^{n} f(\xi_i)\,(x_i-x_{i-1})$，当最大段长趋于零时，若 S_n 趋于同一个数（与怎样切、怎样取点无关），这个数就是 $\int_a^b f(x)\,dx$。两个要点：（1）“与怎样取点无关”是一个实打实的条件。例如在 $[0,1]$ 上，有理点取 1、无理点取 0 的狄利克雷函数，取点不同和式就不同，黎曼积分不存在（需要换用勒贝格积分才能处理）。（2）区间无限或函数无界时，积分可能发散：$\int_1^{\infty} \frac{1}{x}\,dx=\infty$，而 $\int_1^{\infty} \frac{1}{x^2}\,dx=1$。尾巴怎样衰减，决定了整体是不是一个有限的数。

用记账来说：如果每一笔流水都清楚，加起来自然有个总数；如果有些流水本身就说不清（这笔到底是 50 还是 500？），那么取不同的值，总数就跟着变，你得先换一套更精细的记账规则，才能给出一个站得住的总数。再比如，一笔无限期的小额扣费，每个月扣一点，永远不停——它加起来是不是有限，取决于这笔钱是不是越扣越少、少得够快。

“平均”也是一种积分，同样会撞上这条尾巴。算一个小区住户的平均身高，多来一个人、少来一个人，平均数几乎不动；算同一批人的平均财富，只要搬进来一位富豪，平均数就被拉上去一大截，样本再多也稳不住。数学上，有些分布的尾巴衰减得太慢，连“平均值”这个积分都不收敛——柯西分布就是标准的例子，它的样本均值并不随样本增多而趋于任何确定的数。保险公司对巨灾、财务人员对极端行情格外小心，就是因为他们知道，有些整体的“平均”根本积不出一个可靠的数。哪些量能放心地取平均，哪些不能，不是记账规则说了算，而是被加的那些东西自己的尾巴说了算。

底本把积分能成立的原因归于“特征律保证了差异的稳定性”。按现行口径，这里不说“保证”，而说：积分能不能收敛，是被积对象自己的性质，要逐个去查、去证。第395卷的态度是，哲学解释可以帮人问“为什么有些整体能被积出来、有些不能”，但具体某个积分是否存在，只能由数学来回答。

◆ 三、从一条线到一片场

积分的拼合能力，随着维度往上走，越来越惊人。

一元积分是沿一条线加：曲线下的面积，一段时间里走过的路程。二重积分、三重积分是在一片面积、一块体积上加：一块不均匀的钢板有多重，一个房间里的空气总共含多少热量。再往上，研究的对象变成了“场”——空间每一点都有一个量，比如每一点的温度、每一点的电场方向——积分就发展成沿一条路径的曲线积分、穿过一张曲面的曲面积分。

用生活来比：一元积分像数一条街上每家店的营业额，加成这条街的总营业额；二重积分像在一张城市地图上，把每一个街区的人口加成全城人口；曲面积

分像站在一座城市的城墙上，数一整天从各个城门进出的人流，算出城里净多了多少人。

维度越往上，积分拼出来的整体就越抽象、越远离任何人能直接经验的东西。没有人能“看见”一座城市一天的净人流，也没有人能“看见”地球磁场穿过某一片区域的总通量。这些整体，完全是在理念里按规则拼出来的。它们极其有用——电磁学、流体力学、热力学的很多核心定律，正是写成这种积分形式的——但它们离开现实一步比一步远。离得越远，回到现实时要补的东西就越多。这一点，会在第14章讲“再微分”时变成一个具体的工程问题。

◆ 四、同一组流水，几种总数

积分的规则由谁来定？这个问题听起来很哲学，其实很日常。

还是记账。同样一个月的流水，财务人员会给出好几种“总数”。现金口径：这个月实际流出了多少钱。权责口径：这个月该算作花掉的有多少——年初一次交了一年的保险，按月分摊，这个月只算十二分之一。还有“可支配”口径：扣掉房贷、扣掉固定开销，剩下真正能随意花的是多少。三个总数都是从同一组流水积出来的，它们都没错，但回答的问题不同，被拿去做的决定也不同。

科学里同样如此。同一批气象站的温度读数，可以积成“全球平均气温”，也可以积成“陆地平均气温”“海洋表面平均气温”，还可以按面积加权或按站点加权。每一种积法都有道理，但积出来的整体不一样，得出的趋势强弱也可能不一样。这正是为什么同一组数据经过不同团队处理，结论会有差别；也是为什么科学共同体要花大量功夫去比较、校准、公开积分的规则。

这里有一个第395卷的提醒值得放进来：谁有权定积分的规则，本身就是一种权力——第395卷称为“计算权”，与命名权、判错权并列。月底的账按哪种口径算，常常决定了家里谁该省钱；一项公共政策的效果按哪种口径积，常常决定了这项政策是被延续还是被取消。规则不是中立的技术细节，它决定了哪一种整体会被看见。第32章会专门讨论这个问题；这里只记下一点：积分在理念里拼回的整体，带着拼它的那双手的印记。

◆ 五、斯托克斯：边界决定内部的累积

积分有一组极为优雅的定理：格林定理、高斯散度定理、斯托克斯定理。它们可以统一写成一个式子，通常称为广义斯托克斯定理。

> **广义斯托克斯定理** 对一个紧致、有边界的光滑区域 M 及其上的光滑微分形式 ω：$\int_{\partial M}\omega=\int_M d\omega$。 左边是在边界 ∂M 上的积分，右边是在内部 M 上对 $d\omega$（ω 的“外微分”）的积分。 一维特例就是微积分基本定理：$\int_a^b f'(x)\,dx=f(b)-f(a)$，其中“边界”是两个端点 $\{a,b\}$。 三维特例是高斯散度定理：$\oint_{\partial V}\mathbf{F}\cdot d\mathbf{S}=\int_V\nabla\cdot\mathbf{F}\,dV$。 注意：边界上的积分决定的是**内部 $d\omega$ 的总和**，不决定内部每一点的 $d\omega$ 是多少。

这个定理可以用一个日常例子说清楚。一个仓库，你不想挨个去数里面的货，只在门口站着，记下一个月里搬进来多少、搬出去多少。进出相减，就知道这个月仓库里的货一共增加了多少。门就是“边界”，仓库内部每一个角落里货的增减就是“内部的微分”，定理说：门口的总账，等于内部所有增减的总和。

底本把这一点概括为“整体内部的差异之和，完全由整体的边界所决定”，并称之为惊人的普遍规律。这句话要加一个关键的限定：边界决定的是内部的**总和**，不是内部的**分布**。站在门口记账，你知道仓库一共多了一百箱，但你不知道这一百箱放在哪个货架上，也不知道是不是某个角落多了两百箱、另一个角落丢了一百箱。总账对得上，分布可能千差万别。

这个限定不是挑刺，它恰好说出了积分的本性：**积分是一种有损的压缩**。 它把内部无数点上的信息，压成一个数或一个边界上的量。压缩以后，总量保住了，内部的细节丢了。这和第10章讲的微分正好对称：微分在一点上截取，丢了这一点周围的纠缠；积分在一个区域上累加，丢了这个区域内部的分布。医学研究里的“平均疗效”就是这样一种压缩：一种药让多数人略有好转、让少数人明显受损，平均下来仍可能显得“有效”；可落到任何一个具体的病人身上，这个平均数都不告诉他，自己会落在哪一头。

斯托克斯定理还有一个前提常被忽略：被积的东西要在整个区域上处处有定义、处处光滑；区域里若挖掉了一个洞，洞口也算边界，账上必须把它记进去。如果仓库里有一个暗门，货从那里进出不经过大门，门口的账就对不上了。数学上，区域里有一个“洞”（比如挖掉一个点的平面），某些看上去处处平衡的场，绕洞一圈的积分却不为零。

物理上有两个现成的例子。一是阿哈罗诺夫—玻姆效应：一根通电的长螺线管，磁场全关在管内，管外磁场处处为零，可管外的矢势绕管一圈的积分却等于管内的磁通，不为零；电子从管的两侧绕过，会因此出现可观测的相位差。管外区域少了一根“柱子”，不是单连通的，那里的矢势“局部无旋”却不能写成一个全局函数的梯度——数学上叫“闭而不恰当”。二是狄拉克磁单极子：挖掉单极子所在的那一点，周围的磁场处处无源，穿过包住它的球面的总通量却不为零，于是找不到一个处处光滑的全局矢势。两者说的都是同一件事：区域里有洞，局部处处平衡，整体的账仍可能对不上（数学上由上同调来记这种“洞”）。边界决定内部，要以“没有暗门”为前提。

◆ 六、科学理论是积分性的

把视野放大，底本有一个大胆的判断：科学理论的建构，本质上是积分性的。理论不是对客观整体的直接发现，而是从大量局部数据出发，通过数学整合，形成一个整体模型。

这个判断本书保留，并把它说得更具体一点。一条科学定律，比如“行星沿椭圆轨道运行，太阳在一个焦点上”，它从哪里来？开普勒手里有的，是第谷记录的一个个时刻、一个个方位的观测——一笔笔流水。牛顿给出了每一刻的加速度规律，把它积分，才得到整条轨道。轨道这个整体，没有人在天上看见过；它是从一笔笔局部关系里，按规则积出来的。

医学上也一样。“某种药能把高血压患者的中风风险降低多少”，这是一个整体判断。它来自成千上万个病人、成千上万次随访的记录，按统计规则加总出

来。没有哪个医生看见过“降低的风险”本身；他看见的是一个个具体的病人，有的中风了，有的没有。

把科学理论看成积分，有一个好处：它提醒我们，理论的样子取决于积分的规则。选哪些数据进来，怎样加权，在什么范围内加，都会影响最后积出来的整体。这不是相对主义——规则是可以被比较、被检验、被批评的——而是一种清醒：一个理论再普遍，也是在某套规则下积出来的整体，不是世界自己摊开的一张全图。

底本把积分对应于“成就律”，称之为“成就律的数学化”。按现行口径，这种对应不再作为主张。本书只取一个较弱的说法：在三方程的语言里，积分主要落在“显露”一侧——许多局部差异在理念里被拼成一个可识别、可维持、可表达的整体样子；它付出的代价，是内部分布被压扁、被拼合的那部分纠缠不再可见。这是组织方式，不是严格对应。

◆ 七、发现与发明：一个要收窄的二分

底本有一句被反复申明的话：“科学的发现性来自微分，科学的发明性来自积分。”微分忠实地抄下差异，所以是发现；积分大胆地重建整体，所以是发明。底本还认为，这一区分一举化解了经验主义与理性主义关于“科学是发现还是发明”的长期争论。

这个二分有启发性，本书保留它的核心直觉：积分出来的整体确实带有建构的成分，它不是世界自己摊开的一张全图。但它需要从两个方向收窄。

第一，微分也不是纯粹的发现。上一章已经说明，微分之前要选点、选变量、决定按住什么，这些都是研究者的选择，带着建构的痕迹。证件照是真实的，可拍哪个角度、用什么光，是摄影师定的。

第二，积分也不是纯粹的发明。积分的结果要受到严格约束：同一组局部关系，按同一套规则积分，答案是唯一的（至多差一个常数）。牛顿从引力定律积出椭圆轨道，不是想要什么形状就能积出什么形状。月底的总支出可以按不同口

径算，但一旦选定口径，总数就由流水决定，不由记账的人随意。积分里有选择，但选择之后没有随意。

所以更准确的说法是：微分和积分**都**是发现与建构的混合，只是比重和位置不同。微分的建构主要在下刀之前——选哪一点、哪个方向、按住什么；下刀之后，它紧贴着可测的差异。积分的建构主要在规则上——加哪些、怎么加、在什么范围加；规则定下之后，结果受数学严格约束。至于这一区分能否“一举化解”哲学史上的长期争论，本书不作此主张：休谟的归纳问题、康德的先天综合判断问题，都有各自的论证结构，不能靠一个数学类比来裁决。按第395卷的证据责任，哲学解释可以提供新的看法，但不能顶替对原论证的逐条回应。

用一个家常的例子收一下：做一桌菜，买菜是“发现”——市场上有什么，你只能从中挑；配菜做菜是“发明”——同样的材料，可以做出很不一样的菜。可买菜时你挑什么已经带着你想做什么菜的打算，做菜时你也不能把白菜炒成牛肉。两头都有发现，两头都有发明。

◆ 八、积出来的，总缺一块

最后，要为下一章埋一个伏笔。

学过微积分的人都记得，求不定积分时，答案后面总要加一个 $+C$。已知速度 $v(t) = gt$，积分得位置 $x(t) = \frac{1}{2}gt^2 + C$。这个 C 是积分本身无论如何算不出来的那个数——出发时的位置。月底对账也一样：把一个月的流水全部加起来，得到的是“这个月净增加了多少”，不是“月底卡里有多少钱”；月初余额不在任何一笔流水里，只能去查银行卡，也就是去现实里借。看病同理：一个月的血糖升降记得再细，也替代不了今天早上空腹那一针——起点是 5 还是 9，意味着完全不同的处置。

这个小尾巴连着全书的主命题。微分截取时按成零的东西——起点、背景、被固定的条件——积分只能拼回截下来的部分，按掉的那些它找不回来，只有在模型落回现实某一个具体点时才能补回。它是什么、为什么找不回、科学的许多失败为何恰恰集中在它身上，是下一章的题目。

◆ 九、收束

积分在理念里把一片片切面拼回整体。整体是真的，但它从未被直接经验，只被积分出来；它的样子取决于积分的规则；它保住了总量，压扁了分布；它总是差一个常数。斯托克斯定理说，边界管得住内部的总账，但管不住内部的分布，也管不住暗门。

这些不是积分的缺点，而是它的本性：不压缩，就拼不出整体；压缩了，就有东西留在外面。微分丢了一点周围的纠缠，积分丢了区域内部的分布，两处丢掉的东西，会在下一章汇成同一个问题——那个加在每个不定积分末尾的 C。

带 走 的 话

积分是一种有损的压缩：保住总量，压扁分布。一个积分能不能成立、结果是否有限，是被积对象自己的性质，要逐个去证；按哪种口径积，决定哪一种整体被看见。斯托克斯定理里，边界决定的是内部的总和，不是内部每一点，而且以区域“没有洞”为前提。微分与积分都混着发现与建构，只是建构落的位置不同：微分在下刀之前，积分在定规则之时。

小 账

- 本章的月底对账、仓库进出、城门人流、做一桌菜均为说明用的比喻，不对应真实个案；“一种药让多数人略好、少数人受损”是说明平均数的假设情形，不指某一种药。
- 本章没有使用调查数字；黎曼积分、广义斯托克斯定理、柯西分布是标准教科书内容，阿哈罗诺夫—玻姆效应与狄拉克磁单极子的表述核验见附录I第I28条。
- 本章不主张“发现—发明”二分能化解经验主义与理性主义之争；把积分归在三方程的“显露”一侧只是组织方式，不与“成就律”对应。
- 本章主张的推翻条件见附录G第G24条。

12

第 12 章

积分常数 C：微分丢掉、积分找不回的那一份

> 这一章的问题：一本记了三十年流水的账，第一页被撕掉了，为什么怎么加也算不出老人现在有多少钱——科学的定律又缺了哪一页？

老人走后，儿女在抽屉里翻出一本账。账记得极细：哪天买了米，哪天交了电费，哪天收到退休金，一笔一笔，进出分明，三十年没断过。可账本的第一页被撕掉了。第一页上写的是什么，谁也不知道——也许是当年的存款，也许是一笔欠账。

于是出现了一件奇怪的事：三十年里每一天的“变化”都清清楚楚，可老人现在到底有多少钱，从这本账里算不出来。把所有流水加起来，只能得到“比开头多了多少”或者“比开头少了多少”。要知道余额，得去银行查一次——也就是说，得从账本之外，从现实里，再借回一个数。

这本账就是本章的全部问题。流水是导数，余额是原函数，撕掉的第一页，就是数学课本里那个总被轻轻带过的“+C”。本章要说的是：这个 C 不是一个技术性的尾巴，而是整个科学之圈里最要紧的一道缝。微分切下来的时候，它被丢掉了；积分拼回去的时候，它找不回来；科学落回现实的时候，它必须从现实那里借。

◆ 一、一个被课本轻轻带过的“+C”

学过微积分的人都写过这样一行字：

> **不定积分为什么必须带一个 C** 若 $F'(x)=f(x)$，则对任意常数 C，都有 $(F(x)+C)'=f(x)$。因此 $\int f(x)\,dx=F(x)+C$。 反过来，在一个连通区间上，若 $G'=F'$，则 $(G-F)'=0$，由中值定理 $G-F$ 是常数。所以“+C”不多不少，恰好描述了从导数回到原函数时丢失的全部信息：一个常数，不多也不少。
> 写成一句话：$\int f'(x)\,dx=f(x)+C$。求导把 f 的“高度”抹掉，只留下“坡度”；积分只能还原坡度，还原不了高度。

老师通常会提醒一句：“别忘了加 C，否则扣分。”学生记住的往往也只是扣分。可这一行字里藏着一个很硬的事实：导数这件工具，从一出生就是有损的。它记下了每一处的变化，却把“从哪里变起”这件事扔掉了。原因很简单——常数的导数是零。一切“不变的部分”，在微分眼里都看不见。

日常的例子比比皆是。汽车的车速表一路记着每一刻的速度，可只看车速表，你永远不知道车在哪个城市；你得知道它从哪儿出发。体重秤告诉你这个月胖了两斤，可“胖了两斤”既可能发生在一个瘦子身上，也可能发生在一个已经超重的人身上——两斤的意义，取决于那个没被记录的起点。

第10章说过，微分是在现实里截一刀，“其他条件不变”。本章要补的是这一刀的另一面：被宣布为“不变”的那些东西，恰恰是微分记不下来的东西。它们并没有消失，只是从记录里退了出去，变成了一个空着的位置，等着在别处被填上。这个空位，在最简单的情形下，就是 C。

◆ 二、初值问题：从哪里出发，现实说了算

把 C 从“扣分点”变成“科学问题”的，是微分方程。

物理学写下的定律，几乎都是微分方程：它说的不是“东西在哪里”，而是“东西怎样变”。牛顿第二定律说加速度与力成正比；放射性衰变说每一刻衰变掉的量与剩下的量成正比；热传导说温度的变化取决于周围温度的弯曲程度。定律只管“怎样变”，不管“从哪儿变起”。

初值问题：定律给形状，现实给起点　以衰变为例：$\frac{dN}{dt} = -\lambda N$。它的通解是 $N(t) = C\,e^{-\lambda t}$。定律（λ）决定了曲线的“形状”——每过一个半衰期减半；但曲线停在哪一条上，由 C 决定。只有给出 $N(0) = N_0$，才得到 $C = N_0$，曲线才唯一。一般地，$y' = g(t, y)$ 加上初值 $y(t_0) = y_0$，称为初值问题。皮卡—林德勒夫定理说：若 g 连续且对 y 满足利普希茨条件，则在 t_0 附近解存在且唯一。注意：唯一性是“定律＋初值”一起给出的，单靠定律给不出。对 n 阶方程，需要 n 个初值——牛顿力学要同时知道初始位置与初始速度，正是因为它是二阶方程，积两次分，丢两个常数。

这里值得停一下。牛顿力学常被说成“决定论”的典范：知道了定律，一切都定了。可准确的说法是：知道了定律，**再加上**某一时刻所有粒子的位置和速度，一切才定了。拉普拉斯那个著名的“全知者”，之所以需要知道“某一时刻宇宙中一切的位置与受力”，就是因为定律本身是一张没有填起点的表格。决定论的全部力量，一半在方程里，另一半在那个从现实中读取的初始状态里。而后者，从来不是理论能自己给出的。

厨房里的例子：一本菜谱写“每分钟升温约十度，炖四十分钟”。这是“定律”。可同一本菜谱，用冰箱里刚拿出来的排骨和用室温放了一小时的排骨，结果不同。菜谱没错，只是它默认了一个起点，而这个起点在你家厨房里才被确定。

◆ 三、边值问题：外面决定里面

初值问题是“从时间的起点借 C”，边值问题则是“从空间的边界借 C”。

一根两头架着的梁，梁上每一点的弯曲都服从同一条微分方程。可这根梁两头是焊死的、是搁着的、还是一头悬空，方程本身不知道。不同的“两头”，给出完全不同的挠度曲线。一间屋子里的温度分布，服从热方程；可墙外是零下十度还是三十度，窗户开没开，方程同样不知道。

第11章讲过斯托克斯定理的一种读法：边界决定内部。在那里，它是积分的一种美丽对称；在这里，它变成一种依赖：内部的解要等边界交出数据，才能被写出来。偏微分方程的解，往往不是“丢一个常数”，而是“丢一整条函数”——边界上每一点的值都得从外面给。微分方程的阶数越高、维数越多，它丢掉的东西就越多，要向现实借的也就越多。

大白话的说法：你可以知道一锅水里热量怎样流动，但要知道锅里哪儿烫手，得先知道锅底下的火有多大、锅盖盖没盖。定律是锅里的规矩，火和锅盖是锅外的现实。

◆ 四、阿达马的三问：适定性的发生学读法

1902 年前后，法国数学家阿达马（Jacques Hadamard）提出了一个看似朴素的要求：一个数学物理问题要算“提得好”（适定），必须同时满足三条——解存在；解唯一；解连续地依赖于给定的数据。第三条最容易被忽视，也最要紧：初值或边值稍微变一点，解也只应稍微变一点。

> **适定与不适定** 阿达马的三条：（1）存在；（2）唯一；（3）稳定——数据的小扰动只引起解的小扰动。 典型的不适定例子：拉普拉斯方程 $\Delta u = 0$ 的柯西问题。在半平面上给定边界值 $u(x,0) = 0$、法向导数 $u_y(x,0) = \frac{1}{n}\sin(nx)$，解是 $u = \frac{1}{n^2}\sin(nx)\sinh(ny)$。当 $n \to \infty$，数据趋于零，但对任意固定的 $y > 0$，解的振幅 $\frac{1}{n^2}\sinh(ny)$ 趋于无穷。 另一个例子：热方程倒着解。从今天的温度分布推算一小时前的分布，数据里极小的误差会被指数放大。 要点：不适定不等于“没有解”，而是“解对借来的数据过分敏感”——数据上的一丝偏差，在解上变成一场灾难。

阿达马本人一度认为，不适定的问题在物理上没有意义。后来的历史表明他看窄了：医学成像、地球物理反演、从模糊照片恢复清晰图像，都是不适定问题，人们发展出正则化等办法去驯服它们。但他的三问留了下来，成为判断一个模型能不能用的基本检查。

从本书的角度看，阿达马三问可以读成一份关于 C 的“借贷合同”。

第一问“存在”，问的是：借来的数据和定律能不能相容？如果你给的初值本身就违背方程（比如给一根梁规定了不可能同时满足的两端条件），那就借不成。第二问“唯一”，问的是：借来的数据够不够？少借了，解就有一族，C 没被填满。第三问“稳定”，问的是：借来的数据有误差时，结果能不能扛得住？现实里借来的 C 从来不是精确的——量具有刻度，病人报的体重有出入，地质勘察只打了几个钻孔。一个适定的问题，是对借来的 C 有一定宽容度的问题；一个不适定的问题，是对借来的 C 苛刻到近乎不可能满足的问题。

这里要守一条边界：这只是一种读法，属于哲学解释，不是数学命题。阿达马的条件有严格定义，用不着 SDE 来证明；SDE 能做的，是把它放进“微分—积分—再微分”这个圈里，问它在圈上的哪个位置起作用。答案是：它守在积分回到现实的那个关口上。

天气预报是最贴身的例子。大气方程本身是确定的，可今天早上全球观测网量到的气温、气压、湿度，总有误差和空白。洛伦兹在 1960 年代发现，某些方程对初值极端敏感，这使天气预报有一个可预报上限：张福青（Fuqing Zhang）等人 2019 年在《大气科学杂志》上的研究估计，中纬度逐日天气的内在可预报上限平均约为两周，而当时实际的有效预报时效约为十天；具体天数随模式与变量而异。不是方程不好，是现实借给我们的 C 不够精确，而这个问题会把不精确放大。第22章还会回到这里。

◆ 五、C 是什么：被切掉的纠缠，落地时外借

现在可以把本章的主张说出来了。

微分切下来的时候，把这一点周围的纠缠宣布为“不变”，于是它们从记录里退场；积分拼回整体的时候，只能拼回“怎样变”的形状，拼不回“在何处、从何时、在何种周遭里变”。这些拼不回的东西，在最简单的数学里表现为一个常数 C，在微分方程里表现为初始条件、边界条件和参数，在现实里表现为这一个病人、这一块地基、这一批原料、这一群受试者。它们必须在科学落回现实的那一刻，从现实里借回来。

用 SDE 的话说，微分是在一个具体的 E——一团特征纠缠的土壤——里取出一段差异序列 D，把它写成一条可传的规律；积分在理念界把许多这样的规律累加成一个显露得很清楚的模型 S。模型是可识别、可维持、可表达的，可它的清楚是有代价的：它清楚，恰恰因为它把那片土壤剥掉了。C 就是被剥掉的那片土壤在模型里留下的一个空位。

这个空位有三种常见的形态。其一是起点：牛顿力学要初始位置与速度，药代动力学要给药时的血药浓度。其二是边界：结构力学要支座条件，热工设计要外墙温度。其三是参数：同一个方程里的“常数”——摩擦系数、渗透率、病人的肝肾清除率——往往并不是自然界的普适常数，而是这一处现实的局部读数。第三种最隐蔽，因为它被写进了方程，看起来像定律的一部分，其实也是借来的。

一个日常对照：同一份家常红烧肉的做法，在你妈手里好吃，在你手里不好吃。菜谱（模型）一字不差，差在“锅是哪口锅、火是哪种火、肉是哪块肉”——这些是菜谱默认了却没写下的东西。你妈的手艺，很大程度上就是一套在她自己厨房里反复借过、借准了的 C。

这里必须说清楚这个主张的性质。它包含两层：一层是对数学事实的解释——积分常数、初值、边值、适定性，这些都是现成的数学，本章只是给它们一个位置；另一层是一个关于经验科学的主张——**科学在现实中的失败，有相当一部分集中在 C 借不准的地方，而不是集中在规律本身写错的地方**。前一层是解释，用不着检验；后一层是经验主张，必须能被检验，也必须可能是错的。

◆ 六、谁来借，怎样借

C 既然必须从现实里借，就要问：谁在借，用什么办法借。粗略看，有三种办法。

第一种是量。工程师去现场打钻孔、测土的承载力；医生给病人抽血、量肾功能；气象站每天定时放探空气球。这是最直接的借法，代价是钱、时间和仪器的精度。量永远是在有限的几个点上量，点与点之间，仍然要靠模型去补。

第二种是校。先用模型算，再拿现实的反馈去调整参数，直到模型的输出与这一处现实对得上。导航软件根据你过去几次的实际车速，修正它对某条路堵车程度的估计；老中医摸几次脉、看几次舌苔，再调方子。校的办法不要求一次量准，但要求有回路：借了，看结果，再借。

第三种是默认。不量也不校，直接拿一个“通常值”填进去：成年人平均体重、标准大气压、“正常人群”的参数。这是成本最低的借法，也是最容易出事的借法——它把一个本该在这里借的 C，换成了在别处借来、平均过的 C。默认值对“平均的人”最准，对任何一个具体的人都可能不准。

这三种办法在现实里常常混用。值得注意的是，它们都不在理论之内：没有一条定律能告诉你这一位病人的肝脏代谢有多快，只能去量、去校，或者冒险用默认值。这正是第14章要展开的事：科学落回现实的那一步，本身就是一门手艺，而且是一门无法完全写进理论里的手艺。

◆ 七、检验把手：复现危机集中在哪里

如果上面的经验主张有道理，它应当留下可以观察的痕迹。最直接的一条推论是：**一个学科的发现越依赖未被写明的情境，即越依赖没被交代清楚的 C，它在别的地方、别的时间、别的人群中复现时就越容易失败**。换句话说，复现失败不应均匀分布，而应在“情境依赖”高的地方扎堆。

近十余年的复现危机，提供了检验这条推论的材料。下面只列有出处的事实，并把相反的证据一并摆上。

2015 年，“开放科学协作组”在《科学》杂志上报告了对 100 项心理学研究的重复：原研究中约 97% 报告了显著结果，重复后只有约 36% 达到显著。这是复现危机最常被引用的数字。但这个数字本身并不说明失败的原因——统计功效不足、发表偏倚、研究者在分析上的灵活性，都是公认的主要解释，而且它们与 C 无关。

2016 年，范·巴维尔（Van Bavel）等人在《美国国家科学院院刊》上做了一件正对本章问题的事：请不知道复现结果的评分者，给这 100 项原研究的“情境

敏感性”打分——即该效应在多大程度上可能随时间、文化、地点、人群而变。结果是，情境敏感性越高，复现成功率越低；作者报告这一关联在控制了统计功效、效应量等因素后仍然存在。

这看起来像是对本章主张的直接支持。可几乎同时，因巴尔（Inbar）发表评论指出：这一关联可能完全来自一个被遗漏的第三变量——子领域。社会心理学的研究既更“情境敏感”，又复现得更差；认知心理学则相反。他在加入子领域作为控制变量后，情境敏感性的效应不再显著；在两个子领域内部分别看，关联也接近于零。可是子领域之间还有许多别的差异（例如被试内与被试间设计），所以这一争论没有定论。

2018 年发表的“多实验室项目第二轮”（Many Labs 2）则提供了另一种材料：28 个效应，在 36 个国家和地区、125 个样本、1.5 万余名被试中按预注册方案重复。结果约一半效应得到复现；而对本章更要紧的是：同一效应在不同样本、不同场合之间的差异总体上不大，作者的结论是效应大小的变异更多来自“研究的是哪个效应”，而不是“在哪里、对谁做”。

这是一条对本章不利的证据，必须正面对待。它说明：至少在这一批效应、这一种跨样本设计里，“换一群人、换一个国家”并没有像本章推论所期待的那样大幅改变结果。一种可能的辩护是：这些效应本来就是挑选来做跨文化比较的，换样本的维度（国家、线上线下）未必是那些真正被切掉的纠缠；真正的 C 可能藏在更细的地方——实验的措辞、当年的社会气氛、主试的态度。但这种辩护必须小心：如果每一次反面证据出现，都可以说“真正的 C 在别处”，那这个主张就变得无法被驳倒，也就失去了经验内容。

医学方面，常被引用的是 2012 年贝格利（Begley）与埃利斯（Ellis）在《自然》第483卷上的评论《提高临床前癌症研究的标准》：安进（Amgen）公司的科学家在过去十年里尝试重复 53 项“里程碑式”的临床前癌症研究，只确认了其中 6 项（约 11%）。所谓“没能确认”，标准是结果不够稳健、撑不起一个药物研发项目；重复失败时，他们还联系原作者、交换试剂，甚至在原作者指导下重做。临床前研究依赖细胞系、动物品系、实验室条件——这些都是典型

的“未写明的 C”。但这一报告没有公开逐项数据，本身也受过批评，因此它只能作为提示，不能作为证据的主体。临床医学里更扎实的现象是“疗效在试验人群与真实病人之间打折”：试验排除了合并症、高龄、多药合用的病人，而真实世界的病人恰恰带着这些。这是 C 借不准的教科书情形，但要把它量化成对本章主张的检验，还需要专门的研究设计。

把这些放在一起，可以诚实地说：**复现失败与情境依赖之间存在一种被反复提出、部分得到支持、也受到有力质疑的关联**。它足以让本章的主张成为一个值得检验的假说，不足以让它成为一个已被证实的结论。

◆ 八、这一主张怎样会错

为了让主张保持经验内容，先把失败条件写在这里。

第一，若在控制了统计功效、发表偏倚、分析灵活性与学科差异之后，“未写明的情境条件数量”对复现成功率不再有任何预测力，那么“科学失败集中在 C 上”这一经验主张就须收回，只保留积分常数作为数学解释的那一层。

第二，若有设计良好的大规模研究表明，同一效应在系统性改变情境（而不只是换国家、换线上线下）后依然稳定，那么本章须承认：至少在该领域，被微分切掉的纠缠并不构成主要的失败来源。

第三，若“情境依赖”只能在复现失败之后才被识别——即每次失败后都能找出一个“没交代的情境”，而事先无法评定——那么它就只是事后解释，本章的主张须降级为启发性的说法。可操作的做法是：在复现之前，由不知道结果的人对每项研究“需要从现实借回多少东西”做出评定，然后看评定能不能预测结果。范·巴维尔等人的设计正是这一路的雏形。

第四，若那些情境依赖极低的学科（如经典力学的教学实验）也大规模复现失败，那么失败的主因就不在 C，本章的检验把手就选错了地方。

第五，在方法上，本章的读法如果不能比“统计功效不足”“发表偏倚”这些现成解释多出任何可检验的区分，它就只是在给现成的东西换名字，按第395卷的自检纪律，应当承认增量有限。

◆ 九、收束

那本撕了第一页的账，提醒我们一件小事：知道一切变化，并不等于知道一切。微分让科学变得锋利，因为它敢于把“其他条件”放下；积分让科学变得宏大，因为它能把无数锋利的片段拼成整体。可放下的东西并不会因此不存在，拼成的整体也始终带着一个空位。科学的诚实，不在于假装这个空位不存在，而在于知道它在哪里、知道它要从哪里借回来、知道借错了会怎样。

C　在课本里是一个常数，在现实里是一整片土壤。它在数学上的确切位置，要到微积分基本定理里才看得最清楚——那条把微分与积分焊在一起的定理，在焊接的同时，也留下了一道焊缝。下一章就去看这道焊缝。

带 走 的 话

“+C”是数学事实：求导抹掉了起点，积分只能还原“怎样变”，还原不了“从哪里变起”。本书把这个空位读作被微分切掉的纠缠在模型里留下的位置，在微分方程里即初始条件、边界条件和局部参数，理论给不出，须在落地时从现实借回。至于“科学的失败有相当一部分集中在 C 借不准的地方”，那是待检验的经验假说，复现危机的材料部分支持它，也有子领域混淆与多实验室重复的有力反证。

小 账

- 本章的撕页账本、菜谱与排骨、红烧肉等均为构想场景或比喻，不对应真实个案。
- 本章所用复现研究（开放科学协作组 2015、范·巴维尔等 2016 与因巴尔的评论、多实验室项目第二轮 2018、贝格利与埃利斯 2012）及张福青等 2019 年的可预报性估计，出处见附录I（I13—I15、I29、I30）。
- 本章把阿达马三问读作 C 的“借贷合同”，属哲学解释，不是数学命题；积分常数、初值、边值与适定性本身是现成的数学，用不着 SDE 来证明。
- 本章不主张科学的失败全部来自 C，也不否认统计功效不足、发表偏倚、分析灵活性是复现失败的主要解释。
- 本章“这一主张怎样会错”五条登记为附录G第G1—G3条与第G55、G56条；第五条（增量有限）并入第G6条。

13

第 13 章

基本定理：焊接与焊缝

> 这一章的问题：流量计记下了一整个月每一刻的水流，为什么只算得出这个月用了 15 吨，却说不出水表上此刻显示的是多少？

小区门口的水表，一个月抄一次。抄表员不管你哪天洗了澡、哪天浇了花，他只看表盘上的读数：月初是 1203 吨，月底是 1218 吨，这个月用了 15 吨。另一边，如果你在水管上装一个流量计，它会告诉你每一刻水流得多快——早上七点猛一阵，半夜几乎为零。

这两样东西看起来是两回事。一个记的是“总共多少”，一个记的是“此刻多快”。可只要你稍想一下就明白：把流量计一整个月每一刻的读数累加起来，应该正好等于水表上的 15 吨；反过来，水表读数每一刻涨得多快，正好就是流量计上的数。“累计”与“速率”是同一件事的两种看法。

这件在水表上一眼就能看穿的事，在数学史上却花了将近两千年才被说清楚。求曲线下的面积（累计）与求曲线的切线（速率），在古希腊以来一直被当作两个不相干的难题。直到十七世纪，巴罗、牛顿、莱布尼茨才把它们连在一起，这条连接后来被称为“微积分基本定理”。底本把它称为“忠实与发明的焊接”。本章接过这个比喻，并补上底本没有说的一半：凡是焊接，就有焊缝。

◆ 一、定理的两半

先把定理说清楚。

> **微积分基本定理**　**第一部分**：设 f 在 $[a,b]$ 上连续，定义累积函数 $F(x)=\int_a^x f(t)\,dt$。则 F 在 $[a,b]$ 上可导，且 $F'(x)=f(x)$。——先积分再微分，回到原处。**第二部分**：设 f 在 $[a,b]$ 上黎曼可积，G 是 f 在 $[a,b]$ 上的一个原函数（$G'=f$ 处处成立）。则 $\int_a^b f(x)\,dx=G(b)-G(a)$。——先微分再积分，回到原处，但只回到“差值”。　两部分合起来说：在条件满足时，微分与积分互为逆运算。

第一部分可以用水表来懂：从月初开始累计的用水量，它每一刻的增长速度，就是此刻的流量。第二部分也可以用水表来懂：要知道这个月用了多少水，用不着把流量计每一刻的读数都加一遍，只要看月底读数减月初读数。

后者在计算上的意义是巨大的。在基本定理之前，求一块曲边图形的面积，要像阿基米德、刘徽那样，用越来越细的分割去逼近，每一个问题都要单独想办法（见第6章）。有了基本定理，求面积变成了“找原函数”，而找原函数可以靠求导公式倒着查表。一大片原本各自为战的难题，突然变成了同一种演算。这就是“微积分”之所以成为一门学科、而不只是一堆技巧的原因（见第7章）。

◆ 二、一条定理，三代人才焊牢

基本定理并不是一次写成的。它的历史本身就是一部“焊—裂—再焊”的小史。

牛顿在剑桥的前辈巴罗，在 1670 年前后的几何讲义里，已经用几何图形表达了“求面积与求切线互逆”的关系，但他是一个一个图形去论证的，没有把它变成一种通用的演算。牛顿和莱布尼茨的功劳，是把这种互逆变成了规则：牛顿用流数（变化率）与流量（累积量）说话，莱布尼茨发明了 dx 与 $\int$ 这套至今沿用的记号，让“积分是求导的逆”可以像代数一样操作。

可在他们手里，这条定理的地基是摇晃的。所谓“无穷小”究竟是零还是非零，说不清楚（第9章已讲过贝克莱的讥讽）。定理是被大量成功的计算撑着的，而不是被证明撑着的。十九世纪，柯西用极限重新定义了导数与积分，并在连续

函数的条件下给出了基本定理的证明；黎曼把积分定义为分割求和的极限，划出了“黎曼可积”的范围；到了二十世纪初，勒贝格又发现，要使第二部分在最一般的情形下成立，连续还不够，需要更强的条件。

这段历史对本书的意义在于：**定理的“条件”不是一开始就摆在那里的，而是在一次次反例的逼迫下被逐渐写清楚的**。牛顿和莱布尼茨以为他们焊住了一切；柯西和黎曼告诉人们焊住的是哪一部分；勒贝格和后来的人告诉人们焊缝在哪里。每一代人都没有推翻前一代的计算——那些计算至今照用——但都改写了“这条焊接在什么范围内是牢的”。

日常的对照：一座老桥，最初的建造者只知道“能过车”；后来的工程师算出了它的承重上限；再后来的人发现，在某种共振频率下，它即使承重不超也会出问题。桥一直是那座桥，可人们对它“在什么条件下可靠”的知识，一代比一代清楚。

◆ 三、焊接：差异与整体之间的一条通道

底本对这条定理的读法是：它证明了差异与整体之间存在一条严格互通的数理通道。整体不是在差异之外的神秘实体，而是差异的累积；差异也不是整体的残片，而是整体的分解。底本原作用 SIO 的说法，把微分对应“创生律”、积分对应“成就律”，说基本定理把二者焊为一个循环。按本书现行的说法，可以这样讲：微分在现实界截取局部的差异，积分在理念界把差异拼成整体的显露；基本定理保证，在一定条件下，这两个方向的操作可以互相还原。

这个读法抓住了一件真实的事：科学之所以既能“贴着事实”又能“建立整体”，在数学层面，是因为这两个动作之间有一条现成的通道。一个物理学家可以从局部的定律（加速度与力成正比）出发，积分出行星的整条轨道；也可以从观测到的轨道出发，求导回去，检查局部的定律是否成立。往返之间，“怎样变”这一层什么也不丢——在条件满足的时候。

日常的对照：记账。每天的流水是“差异”，月底的余额是“整体”。好的账本，流水加起来正好等于余额的变化；从余额的变化也能核对流水有没有漏。会计里叫“平账”。基本定理，就是数学里的平账原理。

但底本的读法里有一句话需要收窄：“这一‘拆解’与‘重建’是完全可逆的，没有信息在往返中丢失。”这句话在第一部分上成立，在第二部分上并不完全成立。丢掉的那一点，正是上一章的主角。

◆ 四、C 在定理里的确切位置

仔细看第二部分的写法：$\int_a^b f\,dx = G(b) - G(a)$。它给出的是一个**差值**。$G$ 可以是 f 的任何一个原函数——G 也好，$G+5$ 也好，$G-100$ 也好，减法一做，常数就消掉了。

这意味着什么？意味着从速率回到累计，你能得到的只有“这一段时间里增加了多少”，得不到“现在是多少”。水表的例子说得很直白：流量计一个月的记录，可以告诉你这个月用了 15 吨，但不能告诉你表盘上现在显示的是 1218——除非你在月初亲眼看了一次表。

再看第一部分：$F(x) = \int_a^x f(t)\,dt$。这里有一个隐蔽的选择——起点 a。换一个起点，F 就差一个常数。基本定理第一部分给出的累积函数，天生就“从 a 那里开始算”，并且在 a 处取值为零。如果真实的量在 a 处不是零，就得把那个真实的值加上去：

> **两部分合起来，就是“+C”**　若 $G' = f$，则由第一部分与第二部分，$G(x) = G(a) + \int_a^x f(t)\,dt$。右边第一项 $G(a)$ 是**起点处的真实值**，不能由 f 决定；第二项是**从起点起的累积**，完全由 f 决定。所以不定积分的“+C”不是记号上的习惯，而正是这里的 $G(a)$：定律（f）只给出累积，起点必须单独给出。定积分（差值）可以不借 C；“某一点处的值”必须借 C。

这就是上一章那句话的数学底子：微分丢掉的，积分找不回。更精确地说，**积分能找回的是“差值”，找不回的是“位置”**。科学里许多最漂亮的结果——

能量守恒、动量守恒、某一过程中做的功——都是差值，所以理论可以不借现实就给出；而工程师、医生、导航员真正要的，往往是位置：这根梁此刻承受多大的力，这位病人此刻血药浓度是多少，这艘船此刻在哪里。位置要借。

物理学里有一个现成的例子，把这件事说得更透。中学物理就教：重力势能的“零点”可以任选——你可以把地面定为零，也可以把桌面定为零，算出来的势能数值不同，但两点之间的势能差一样，由此推出的运动也一样。为什么可以任选？因为真正进入运动方程的是势能的**变化**（它的导数给出力），而不是势能本身。势能本身带着一个 C，物理学干脆承认这个 C 不可观测，于是“零点任选”。电学里的电势也是这样：只有电压（电势差）可测，“这一点的电势是多少”要先约定一个参照。

这是一个很有意思的情形：在某些问题里，C 不仅借不回来，而且根本不需要借，因为我们关心的只是差值。可一旦问题变成“这块石头会不会砸穿屋顶”，你就不能只说势能差，得知道它从多高落下——这时 C 又回来了，而且必须从现实里量。哪些问题可以不借 C，哪些问题必须借，这本身就是判断一个模型能不能直接拿去用的第一道检查。

这里可以看出焊接与焊缝的关系。焊接是真的：在条件满足时，速率与累计之间确实可以严格地互相推出。焊缝也是真的：这条通道只通“变化”，不通“起点”。说得再大白话一点：基本定理是一座桥，桥修得极好，但它只连接两岸的“相对位置”；你站在桥上哪里，桥自己不知道。

◆ 五、条件：什么时候能焊

底本强调，基本定理不是无条件成立的。它要求连续、可积、极限存在。底本把这些条件读作“特征律在数学层面的显影”。按现行口径，可以说得更朴素也更可检验：**这些条件要求差异已经稳定下来了**——在相邻的尺度上可以过渡，在累加时可以收敛。

“连续”在大白话里是：此刻和下一刻之间没有突然的跳。水龙头的流量可以变，但不会在一瞬间从零跳到无穷大。“可积”在大白话里是：把每一刻加起来，加得出一个确定的数，不会因为你切分的方法不同而得到不同的答案。

这两个条件和第2章讲的“特征”有关联：一个东西要成为科学的对象，它的差异要能被识别、被维持、被表达（S 的“三个可”）。在数学上，连续和可积可以看作“可维持”的一种最低形式——差异在相邻处可以接得上，累加起来不发散。但这只是一种对应的提法，属于解释层，不是说 SDE 推出了连续性的定义，也不是说凡满足 S 的“三个可”的东西都一定连续。数学条件有自己独立的定义与证明，不需要也不应当借哲学来撑腰。

这一节的要点只有一个：**焊接是有条件的；条件一旦破了，焊缝就开**。下面看三个开焊的例子。

◆ 六、焊缝开了：三个反例

第一个例子：魔鬼的阶梯。康托尔函数（俗称“魔鬼阶梯”）是一个从 0 连续地爬升到 1 的函数，但它在几乎所有点上的导数都是零——它的全部上升，都发生在一个“长度为零”的点集上。于是：

> **焊缝开裂的数学样本** **康托尔函数** $c(x)$：在 $[0,1]$ 上连续、单调、$c(0)=0$、$c(1)=1$；几乎处处 $c'(x)=0$。 于是 $\int_0^1 c'(x)\,dx = 0$（在勒贝格意义下），而 $c(1)-c(0)=1$。导数积回去，丢掉了全部的增长。 补焊：勒贝格理论指出，要使 $\int_a^b G' = G(b)-G(a)$ 成立，需要 G **绝对连续**——一个比“连续”更强的条件。 **沃尔泰拉函数**：处处可导、导数有界，但导数不是黎曼可积的。在黎曼框架下，第二部分的左边根本无从谈起；改用勒贝格积分后，焊缝才重新合上。 **布朗运动**：几乎每条路径处处连续、处处不可导。第一部分的“$F'=f$”在这里失去意义，必须另起炉灶（伊藤积分，见第19章）。

用水表来翻译魔鬼阶梯：设想一块怪表，它的流量计几乎每一刻都显示零，月底读数却实实在在涨了一吨。用水全都发生在流量计“看不见”的那些瞬间里。你如果只信流量计，就会以为这个月一滴水没用。

这就是焊缝开裂的样子：差异没有消失，只是它发生在微分捕捉不到的地方。微分只看“每一处的斜率”，而这个函数的变化恰好不在任何“处”上，而在一种过于稀碎、过于聚集的方式里。

第二个例子：沃尔泰拉函数说的是另一件事。它处处有导数，导数也不大，可这个导数乱到黎曼的办法加不起来。也就是说，差异每一处都记下了，但它们没有稳定到可以累加的程度。焊接失败不是因为微分漏了东西，而是因为积分接不住。

第三个例子：布朗运动。水里的花粉颗粒被水分子从四面八方撞击，走出一条极其曲折的路径。数学上的布朗运动（维纳过程）几乎每条路径都连续，却处处不可导：在任意小的时间段里，它都在剧烈地来回抖动，没有一个确定的“此刻速度”。经典基本定理的第一部分在这里落空，因为根本没有 f 可言。

二十世纪，伊藤清发展了一套随机积分，让这样的过程重新可以“微分—积分”。代价是链式法则多出一项——随机抖动在二阶上不能忽略。这一点留给第19章展开。这里只记住：当差异尚未稳定到“连续可导”，经典的焊接就开了；数学并没有因此放弃，而是发明了新的焊法。

三个例子放在一起，有一个共同的教训：**基本定理的条件不是数学家的洁癖，而是对“差异稳定到什么程度”的精确要求**。每一次焊缝开裂，都逼出了一种更细的积分理论——勒贝格积分、亨斯托克—库尔茨韦尔积分、伊藤积分。这正是第16章要讲的那个循环在数学自身里的一次小规模演出：新出现的病态函数（差异 D）与已沉淀的积分理论及其习惯用法（E）之间的矛盾累积，逼出积分概念（S）的改变，新的积分概念又回写了“什么算一个函数”“什么算可积”的标准。这一回写是可以在数学史里被观察到的，而不是被宣布的。

◆ 七、焊缝的另一侧：现实没有义务满足条件

数学上的焊缝，只要发明更好的积分就能补上。可现实里的焊缝，是另一回事。

模型能用基本定理，是因为模型里的函数是我们规定为连续、可积的。可现实中的量，是否连续、是否可积，并不由我们规定。股价在开盘时会跳空；地震前后的地应力会突变；一个病人的病情可能在一夜之间转折。把这些量当作光滑函数来处理，等于假定它们的差异已经稳定——而这个假定本身，需要现实来批准。

这就把本章与上一章连了起来。上一章说，从模型回到现实，必须借回 C。本章补上一句：借 C 之前，还要先确认现实这一段确实满足“可以焊”的条件。在可以焊的地方，借一个起点就够了；在不能焊的地方——跳变、突转、处处抖动——光借一个起点也不够，还得换一种积分，甚至承认这一段不能被积分模型覆盖。

一个开车的例子：导航软件靠“速度 × 时间 = 路程”来推算你的位置（这在航海里叫航位推算）。在平直的高速公路上，这条焊接几乎完美。可一旦进了隧道、绕了立交、GPS 信号跳了一下，推算出来的位置就开始漂移。导航软件的办法不是改进积分公式，而是隔一会儿重新从卫星借一次位置——重新借 C。这个“隔一会儿再借”，已经是下一章的主题了。

◆ 八、两个方向并不对等

最后还要指出一个常被忽略的不对称。

基本定理的第一部分，“先积分再微分”，是相当宽容的：只要被积函数连续，累积函数就一定可导，导数也一定回到原函数。可以说，从整体回到差异，这条路比较好走。第二部分，“先微分再积分”，则苛刻得多：它要求函数的导数存在、可积，在最一般的情形下还要求绝对连续。从差异回到整体，这条路窄得多，而且即便走通了，也只回到差值，不回到位置。

翻成大白话：拆一台钟容易，装回去难；拆的时候每个零件都看得清，装的时候还得知道它原来摆在什么位置。科学里也是这样。把一个已有的整体模型拆成局部的变化率，几乎总能做到；从零散的局部变化率拼回一个整体，则要求那些局部已经足够稳定、足够整齐，而且还要从外面补上一个起点。

这一不对称，和全书的主命题是一致的：在“微分—积分—再微分”这个圈上，微分与积分不是两件等价的事。微分是截取，几乎总能截；积分是拼合，需要条件、需要起点。于是，一个模型越是靠大量局部的微分拼出来，它就越依赖那些没被写进方程的条件与起点。这也是为什么 AI 把海量的梯度（微分）累加进权重（积分）之后，它在何种条件下可靠、它的“起点”藏在哪里，会成为一个真问题（见第26章）。

◆ 九、这一章的读法怎样算错

本章对基本定理的读法有两层，各有失败条件。

其一，“定积分只给差值、位置必须外借”是数学事实，不会错；但“现实中科学应用的困难主要来自位置（C）而非差值”是经验主张，与第12章同一个检验把手共用：若应用失败与起点、边界数据的质量无关，这个说法须收窄。

其二，“基本定理的条件对应于差异已经稳定”是一种对应的提法。若有人能举出大量“差异明显没稳定、却完全满足连续可积”的现实量，或反之，那么这个对应只能保留为局部比喻，不能当作结构上的说明。

◆ 十、收束

微积分基本定理是一次真焊接：它让局部的差异与整体的累积，在条件满足时可以严格地互相推出。底本说它是全书的数学基石，这句话可以保留。但基石上刻着两行小字：第一行，它只焊住了“变化”，没焊住“起点”——起点要从现实借，这就是 C 的数学身份；第二行，它要求差异已经稳定，条件破了焊缝就开，开焊之处往往正是新数学与新科学的生长点。

焊好的东西，终究要用。一个积分出来的模型，要在现实中派上用场，就得回到某一个具体的点上，借回 C，再算一次“此刻该怎样”。这一步，本书叫它再微分。

带走的话

基本定理是一次真焊接：在条件满足时，速率与累计可以严格地互相推出。但它焊住的是“差值”，不是“位置”：定积分不借 C 也算得出，某一点处的值必须单独给出起点。它的条件也不是洁癖，连续、可积乃至绝对连续一旦不满足，焊缝就开。本书把这些条件读作“差异已经稳定”，这是解释层的对应提法，不是由 SDE 推出的数学结论。

小账

- 本章的水表、老桥、拆钟、怪表与航位推算的开车例子均为构想或比喻，不对应真实个案。
- 本章所述定理的两部分及康托尔函数、沃尔泰拉函数、布朗运动诸反例是标准数学结论；巴罗、柯西、黎曼、勒贝格的贡献按通行数学史叙述。本章没有使用任何调查数字。
- 本章不主张连续、可积的定义可由“三个可”推出，也不主张满足“三个可”的东西都一定连续。
- 本章主张的推翻条件见附录G第G25、G26条。

14

第 14 章

再微分：技术与应用的发生

> 这一章的问题：药理书写得清清楚楚，医生为什么还要让七十岁的老太太抽血、调量、再抽血——书上缺的那几个数从哪里来，用药又会改写什么？

一位七十岁的老太太，房颤，医生给她开了一种抗凝药。药理书上写得清清楚楚：这种药在体内怎样吸收、怎样代谢、多大剂量把凝血指标压到什么范围。可医生并不照着书上的“标准剂量”一次开定。他先开一个起始量，叮嘱她过几天回来抽血；化验单一出来，看凝血指标偏高还是偏低，再把剂量往下调半片或往上加一点；再过一周，再抽，再调。几个来回之后，剂量才稳定下来。旁边一位同龄的老先生，吃同一种药，最后稳定在完全不同的剂量上。

药理书没有错，医生也没有不信药理书。他只是知道一件书上写不了的事：书上的模型是从成千上万个病人身上积分出来的，而眼前这位老太太，是现实里的一个点。要把书用到这个点上，得先从她身上借回几个数——她的体重、肝肾功能、正在吃的其他药、甚至她平时吃不吃绿叶菜——然后在这个点上，重新算一次“此刻该给多少”。算完，看她的身体怎样回应，再算一次。

本书把这一步叫作**再微分**。前两章说，微分切掉的纠缠，积分找不回，只能在落地时从现实借回来。本章要说：借回来的那一刻，科学变成了技术；而技术一旦落地，又会在现实里引出新的差异，反过来冲击科学原来的模型。这是全书主命题的第三段，也是“科学之圈”闭合的地方。

◆ 一、圈的第三段：什么叫再微分

先把词说清楚。

第10章的微分，是在流动的现实里截取一个局部：在“其他条件不变”的前提下，看一个量对另一个量的变化率。第11章的积分，是在理念界把许多局部拼成整体：定律、模型、预测。再微分，是把这个整体模型**带回现实里的某一个具体点**，在那里求出局部的、可执行的答案——这根梁在这个荷载下会弯多少，这位病人此刻该吃多少药，这艘船下一分钟该往哪个方向偏几度。

为什么叫“再微分”而不只叫“应用”？因为这一步在数学形式上确实是一次微分：在一个点附近，用切线、梯度、局部线性化去近似整体。而它与第一次微分的方向相反：第一次微分是从现实走向理念——截取差异，准备积分；再微分是从理念走回现实——拿着积分好的整体，到一个具体点上取局部。

> **再微分的最小形式：在一个点上取切线** 设积分得到的整体模型是 $y = f(x)$，它的通解里含有待定的常数（或参数）。落到现实中的点 x_0 时，要做两件事：（1）**借C**：从现实测出 $f(x_0)$，或测出模型里的局部参数，把通解定成这一个解；（2）**取局部**：在 x_0 附近用泰勒展开 $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$，得到“此刻每改动一点，结果会变多少”，据此做出决定。 若模型是多变量的，第二步就是求梯度 $\nabla f(x_0)$ 或雅可比矩阵：工程上的灵敏度分析、控制上的线性化、医学上的剂量—反应斜率，都是这一步。 关键：第（1）步不在模型里，必须来自现实；第（2）步的准确性只在 x_0 附近成立，离开这个点就要重做。

方框里的第二条值得强调：再微分得出的答案是**局部**的。它只在这一个点附近有效。医生给老太太调好的剂量，对老先生不适用；一个月后老太太换了别的药，也可能不再适用。技术的每一个答案，都是在一个点上、在一个时间窗里成立的答案。这不是技术的缺陷，而是它的本性。

大白话的对照：裁缝。服装学有一整套人体比例的理论，可裁缝做衣服，先要在你身上量一遍尺寸。量完了，按理论在你这个“点”上画样、剪裁；试穿时

哪里紧了、哪里垮了，再改。成衣店卖的是积分的产物——按“平均身材”做的衣服；裁缝做的是再微分——按“这一个身材”做的衣服。

◆ 二、四个现场：在哪一点上借，借的是什么

再微分不是抽象的说法，它发生在每一个技术现场。下面看四个。

工程设计。结构力学的方程对所有钢筋混凝土梁都一样，可每一栋楼都要单独计算。设计师要借回的东西包括：这块地的地质勘察数据（几个钻孔的土层和承载力）、当地的风荷载与地震设防等级、这批钢材的实测强度。地质勘察就是典型的借 C：只在几个点上打孔，点与点之间靠模型推断。工程史上不少事故，不是因为力学错了，而是因为借来的 C 错了——地下有一层没勘察到的软土，或者实际使用的荷载远超设计时的假定。规范里的“安全系数”，可以读作一种对 C 借不准的预先承认：我知道我借来的数不精确，所以多留出几倍的余量。

医疗用药。开头的例子已经说过。这里补一个更细的层次：同一种药，在不同人体内代谢速度可以差得很远，其中一部分差异与基因有关——有些人携带的代谢酶基因变体会让药物在体内停留得更久。开头那类抗凝药华法林就是一例：携带 CYP2C9 功能降低变体的人清除它更慢，再加上药物靶点基因 VKORC1 的差异，所需剂量可以相差数倍。临床药物基因组学实施联盟（CPIC）2017 年更新的指南，已按这两个基因的分型给出起始剂量建议；不过，按基因定剂量能否改善出血、血栓这些临床结局，几项大型随机试验的结论并不一致。所谓“精准医疗”的一条主线，就是想在用药之前，就从这个人身上多借回几个 C。但无论借得多细，最后一步仍然是开头那位医生做的事：给一个量，看反应，再调。剂量—反应曲线是积分出来的；“这个人此刻该给多少”，永远要在这个人身上再微分。

导航。上一章提到航位推算：知道速度和航向，积分出位置。可积分会累积误差，越走越偏。1960 年代阿波罗登月的导航系统里，用上了卡尔曼（Rudolf E. Kálmán）在 1960 年提出的一种递推估计方法，即后来所称的卡尔曼滤波：一边用运动方程推算位置，一边隔一段时间用新的观测去修正推算——每一次修

正，都是重新从现实借一次 C，并据此把误差的估计也一起更新（第24章讲控制论时还会回到它）。今天手机导航里的定位，基本是同一种思路：卫星信号、车速、陀螺仪，几路数据不断地互相校正。导航里没有一劳永逸的积分，只有一刻不停的再微分。

制造工艺。一条注塑生产线，理论上，温度、压力、冷却时间都可以从材料的流变学方程算出来。可实际开机的第一批产品，往往有缩痕、有飞边、尺寸偏差。工艺工程师要做的是“调机”：在这一台机器、这一批原料、这一天的车间温湿度下，把参数一点点调到合格。调好之后，换一批原料，可能又要重调。工厂里所谓“老师傅的手感”，很大程度上就是一套积累起来的、在具体机器上借 C 的本事。统计过程控制则把这件事系统化：持续地抽样测量，一旦数据偏出控制界限，就再调一次。

四个现场有同一个结构：**整体模型由科学提供，具体的点由现实提供，答案在两者相遇的地方被算出来，并且只在那里有效**。

◆ 三、技术是“积分模型在现实一点上的再微分”

现在可以给技术一个本书意义上的说法：**技术，是积分模型在现实一点上的再微分**。

这个说法有几层意思。

第一，技术离不开科学的积分。没有结构力学，盖楼只能靠经验与试错；没有药代动力学，用药只能靠师徒相传的“大约几钱”。底本第十章讲得很充分：微积分让工程从经验技艺变成预测科学。这一层不必重复（第23章会展开）。

第二，技术不等于科学的积分。科学给出的是一族解，技术要的是这一个解。从一族到这一个，中间隔着 C。所以技术永远有一部分不在理论里：现场勘察、临床观察、调机、试飞。这一部分常被看作“只是操作”，可它恰恰是整个圈里唯一与现实直接接触的那一段。

第三，技术的知识有两种。一种是可以写进教科书的“积分知识”——公式、规范、手册；另一种是在具体点上反复借 C 积累下来的“再微分知识”——这块地的脾气、这台机器的毛病、这类病人的反应。后一种知识难以完全写出，常以经验、手艺、直觉的样子存在。它不是比科学低一等的东西，而是科学落地必须付出的另一半。

和底本的说法比较一下。底本原作把技术对应“自由律”，说技术不断扩展差异的选择空间。这个说法抓住了技术的一个后果：有了可算的模型，人可以在动工之前反复比较多种方案，选择的余地大了。按现行口径，意义三律与具体对象之间只是“主要侧重”，不是严格对应，所以本书不把技术等同于自由律。本章补的是另一面：选择空间的扩大，是在理念界发生的；而每一次选择要落地，都得回到现实的一个点上，付出借 C 的代价。可以在图纸上画一百座桥，但每一座要建的桥，都要在它那块地上重新算一次。

◆ 四、再微分引出回写：新差异冲击旧模型

到这里，再微分还只是“把模型用出去”。可它还有另一面，而这一面才使它成为全书的枢纽：**再微分会在现实里引出新的差异，这些新差异反过来冲击原来的模型**。

用第395卷的 123 原理来说：① 现实中不断冒出的差异（D）与已沉淀的纠缠土壤（E：既有的规范、做法、默认条件）之间存在张力，张力累积；② 张力逼出显露的改变——新的模型、新的技术方案立了起来；③ 新显露回写差异与纠缠——它重新规定什么算差异、往哪里走，并沉淀为新的土壤。这里要把“再微分”的两种用法分清（第16章还会再提）：本章主讲的是**应用的再微分**——拿既有模型落到一个具体点上、借回 C，它单次本身不改模型；另一种是**回写的再微分**——模型被逼改之后，用新模型重新规定从现实里切什么，这才是第③步的科学形态。两者的接口在于：应用的再微分在现实里引起的新差异，若系统性地累积起来，就成为第①步里 D 与 E 拉扯的材料；一旦逼出模型改变，随之而来的就是回写的再微分。每一次技术落地，都可能是这条链的起点。

再微分为什么会回写 设模型 M 在点 x_0 处再微分，得到决策 u_0，并在现实中执行。执行改变了现实：$x_0 \to x_1$。若 M 的建立依赖“其他条件不变”，而 u_0 恰恰改变了那些“其他条件”（例如：用药改变了病原体的选择压力；修路改变了出行者的选择），则在 x_1 处，M 原先依赖的切分已不再成立。于是出现三种可能：（a）偏差落在模型的误差范围内，模型不变，只需重新借 C；（b）偏差系统性地超出范围，形成反常，累积到一定程度，逼出模型修改；（c）偏差改变了现实本身的结构，使原来的问题变成了另一个问题。只有（b）和（c）会进入 123 原理的循环、引出本书意义上的回写。（a）只是普通的再校准，不能被夸大为回写。

方框里最后一条很重要。第395卷强调，回写要被观察，不能被宣布。不是每一次技术落地都回写了科学；绝大多数时候，只是借了一次 C，用完就完了。要说某次再微分引出了回写，必须能指出：原来的模型在哪一处被改了，或者原来的问题在哪一处变了。

正向回写的历史例子并不少。蒸汽机是在热力学成熟之前造出来的；工程师们为了提高它的效率，在一台台具体的机器上反复试验、测量、比较。卡诺在 1820 年代研究“热机的最大效率”，由此打开了一条通向热力学第二定律的道路。在这里，是技术在现实中一个个点上的再微分，逼出了科学的一次新积分。半导体器件也有类似的故事：为了让晶体管工作得更稳定，人们不得不去研究材料表面和缺陷的物理，这些研究反过来丰富了固体物理。

◆ 五、回写是中性的：负向回写

回写不等于进步。第395卷明确说过：回写是中性的，存在负向回写；发生不等于改善。技术领域里，负向回写的例子同样清楚。

抗生素。抗生素的药理模型——它怎样杀死细菌——在临床上一次次被成功地再微分到一个个病人身上。可每一次用药，也是对细菌群体的一次选择：敏感的死了，耐药的留下来。用得越多、越不规范，耐药菌就越多。最后，原来的模型（“这种药能杀死这种菌”）在现实中失效了。药理学并没有错，错在模型建

立时被切掉的那一份纠缠——细菌群体会演化——在大规模再微分之后，以耐药的形式顶了回来。

修路。交通工程里有一个被反复讨论的现象：为了缓解拥堵而拓宽道路，过一段时间后，拥堵往往又回来了，因为更好走的路吸引了更多的车——人们改变了出行的时间、路线乃至居住地点。这常被称为“诱发需求”。迪朗东（Duranton）与特纳（Turner）2011 年在《美国经济评论》上用美国城市数据估计，州际高速公路的车道里程每增加一成，车辆行驶里程也大约增加一成，他们称之为“道路拥堵基本定律”；在其他等级的道路和其他国家，这个比例是否同样成立，研究中仍有争议。原先的流量模型把“需求”当作给定的 C；可修路这次再微分，改变了需求本身。

考核指标。经济学里有一条以古德哈特（Charles Goodhart）命名的说法，大意是：一项指标一旦被当作目标，就不再是好的指标。在管理中，一个衡量工作表现的模型被再微分到每个员工身上，变成了考核标准；员工随之改变行为去迎合指标，于是指标与它原本要衡量的东西脱了钩。

三个例子有同一个结构：**再微分改变了被模型默认为“不变”的那些条件**——细菌的构成、人的出行选择、员工的行为方式。这些条件原本被第一次微分切掉了，藏在 C 里；再微分把模型推回现实，它们就以新差异的形式回来，冲击模型。

日常的对照：家长给孩子定了“每天背二十个单词就可以玩游戏”的规矩。头几天很有效。过了一阵，孩子学会了挑最短的单词背，背完就忘。规矩（模型）没变，孩子（现实）变了；规矩当初默认的“孩子会认真背”，被规矩本身改写了。这是一次小小的负向回写。

负向回写并不意味着不该再微分。抗生素救了无数人，道路也确实连接了城市。它意味着两件事：一，每一次再微分都要预留观察回写的通道——耐药监测、交通流量的长期跟踪、对考核指标的定期审查；二，回写的方向不能被预设，它可能把底盘写得更好，也可能写得更坏。第395卷那句口诀在这里很贴切：底盘每天都在被写，看好你的笔。

◆ 六、AI 也绕不过这一步

现在可以回到全书的另一条线：人工智能。

一种流行的想象是：只要模型足够强大，它就能直接“算出”现实中的答案，不再需要人去现场、去试、去调。这种想象，把科学之圈的三段压成了一段——仿佛有了积分，就不再需要再微分。

本书的判断是：**AI 能替人做积分，甚至能把积分做得远超人力；但落在现实某一点上的那一次再微分，它绕不过去。**

这句话需要拆开说，否则容易被误读。

第一，再微分的计算部分，AI 当然可以做，而且已经在做。自动驾驶系统在每一刻根据传感器数据计算方向盘角度，飞机的飞控系统在每秒内多次修正舵面，推荐系统针对每一个用户算出推送内容——这些都是在一个点上取局部的计算。说 AI 做不了再微分的计算，是不对的。

第二，绕不过去的是**借 C**。再微分要成立，必须从这一个现实点上借回数据：这条路此刻的路况、这位病人此刻的指标、这批材料此刻的性质。这些数据不在模型的权重里，只在现实里。AI 可以帮忙借——更好的传感器、更快的处理——但它借来的东西，与人借来的东西一样，受测量精度、覆盖范围、偏差的限制。一个在别处的数据上训练出来的模型，用在一个从未见过的现实点上，就是用别处借来的 C 冒充此处的 C，这正是上面说的“默认值”借法，风险也一样。

第三，绕不过去的是**回写的观察与承担**。再微分在现实里引起的新差异，要有人去观察、去判断它是可以忽略的误差、需要警惕的反常，还是已经改变了问题本身的结构性变化；而当回写是负向的，要有人承担后果、决定是否停下来。这一点与 AI 的计算能力强弱无关，而与“谁在现实中站在那个点上”有关。第27章会用 AlphaFold 和天气模型的例子说明：AI 在积分一侧取得的成就是真实的，可那些成就要变成药、变成防灾决策，仍要经过实验室、临床、现场的一次次再微分。

用一个日常的比方：一位熟读万卷医书、记性极好的助手，可以在你看病前把所有相关文献整理好，甚至给出建议的剂量。但药是吃进你的身体里的，抽血化验的结果是你身上的，吃了药之后的反应，也只在你身上发生。助手再强，那一次“在你身上”的调整，绕不过你、绕不过你的身体、也绕不过替你作决定并为此负责的那个人。

这个判断有它的边界，下一节写清。

◆ 七、本章主张怎样算错

第一，若存在一类技术，它在落地时完全不需要从具体现实点借回任何数据（初始条件、边界条件、局部参数），仅凭通用模型就能在各种具体情形下稳定达到要求，则“技术是在一点上的再微分”须收窄为“多数技术”。数学里的纯算法、密码学这类不与物理现实接触的技术，可能本来就在这个说法的范围之外，应当明确排除。

第二，若技术落地后的偏差在大量案例中都只是普通的再校准（方框中的情形 a），从未系统性地逼出模型修改，那么“应用的再微分会引出回写”就须降级为“偶尔引出回写”，并且要给出区分这两种情形的判据。

第三，若在把回写说成“中性”之后，任何结果——变好、变坏、不变——都能被说成是“回写”，这个概念就失去了检验力。本章要求的最低标准是：指出被改写的是哪一个模型、哪一项默认条件，以及改写前后可以观察到的差别。

第四，若将来出现 AI 系统，能在没有人类参与的情况下，自己从现实中借回足够精确的 C、自己识别回写并据此修改模型，且在长期的现实运行中表现稳定，那么“AI 绕不过再微分”中关于“观察与承担”的部分，就须收回到“目前尚未”的程度。本书对此保持开放：主命题说的是“替不了落点上的那一次再微分”，若这一点被经验推翻，主命题须相应修改。

第五，若“再微分”这个词在分析具体案例时，不能比“应用”“校准”“实施”这些现成说法多出可检验的区分，它就只是换了一个名字，应当承认其增量有限。本章认为它多出的区分在于：它把“必须借回什么（C）”与“会

不会改写模型（回写）”这两个问题同时摆上桌面。这一点需要在后续各章的案例中接受检验。

◆ 八、收束

现在，科学之圈闭合了。在现实里微分，切下一段差异；在理念里积分，拼出整体模型；再回到现实的某一点上微分，借回被切掉的那一份，算出此刻该怎么做。这第三段，是科学变成技术的地方，也是科学被现实再次考验的地方：每一次落地都会引出新的差异，有时它们只是噪声，有时它们累积成反常，逼出新的模型——有时把底盘写得更好，有时写得更坏。

这个圈怎样用 SDE 的三方程来描述，在三方程的静态关系里，微分、积分、再微分各自落在哪里，是下一章要做的事。

带 走 的 话

再微分，是把积分出的整体模型带回现实的某一点，借回 C，在那里求出局部、可执行的答案；技术就发生在这一步，答案也只在这一点、这一时间窗里成立。单次的再微分不改模型，多数时候只是再校准；只有偏差系统性累积成反常，或改变了问题本身，才谈得上回写，而回写可正可负，必须指出改了哪一处。AI 能做再微分的计算，绕不过的是借 C，以及对回写的观察与承担。

小 账

- 本章的老太太与老先生用药、裁缝、背单词的规矩、熟读医书的助手均为构想场景或比喻，无真实原型。
- 本章所用华法林与 CPIC 2017 年指南、迪朗东与特纳 2011 年的估计，出处见附录I（I31、I32）；卡尔曼滤波、卡诺、古德哈特的说法按通行史实与转述。
- 本章不构成用药或医疗建议；不主张 AI 做不了再微分中的计算，“观察与承担”一条对未来的 AI 系统保持开放。
- 本章不把每一次落地都算作回写；方框中的情形（a）只是再校准。
- 本章主张的推翻条件见附录G第G4、G5条。

15

第 15 章

三方程下的微积分循环

> 这一章的问题：卖菜老人的一本小账里，算出的结论、记账的老规矩、手边的秤和账本，谁先谁后说不清——科学里的模型、切法和土壤，是不是也这样互相生出来？

傍晚的菜市场，一个卖菜的老人收摊前要算今天赚了多少。他不会把每一棵白菜的来龙去脉都记下来，只记三样东西：进了多少斤，卖了多少斤，各是什么价。斤数和价钱，是他从一整天流动的买卖里“切”出来的增量；切法不是他临时想的，是他跟着父亲摆摊那些年就定下来的老规矩——零头抹不抹，烂叶算不算损耗，赊账记在哪一页。他手里那杆秤、那个旧账本、隔壁摊位之间互相通报的批发价，都参与了这本账的成形。账算完了，结论是“今天赚了八十块”；而这个结论第二天又会反过来影响他：赚得少，他会改进货的量，也会改记账的格子，比如单独给“隔夜菜”开一栏。

这本小账里藏着三样东西的相互牵扯：一个算出来的整体结论，一套“切什么、怎么切”的记账习惯，一堆秤、账本、同行之间的往来。结论离不开切法和工具；切法是被以往的结论和工具规定的；工具和习惯又是一次次结论沉淀下来的。谁先谁后？说不清，也不必硬说清。

前面几章把科学拆成了三段：在现实里微分，在理念里积分，落地时再微分（第10—14章）。拆开讲是为了看清每一段的手艺与代价，但真实的科学从来不是三段依次排队走完。本章要把拆开的三段重新放回一个互相牵扯的整体里看——用 SDE 的三大方程。

◆ 一、三个字母各指什么

先把三个字母落到科学上。

S（显露），在这里指一个科学模型或定律被共同体看见、维持、表达的那个样子：伽利略的落体规律、开普勒的三定律、牛顿的运动方程、一张标准的元素周期表。按第395卷的“三个可”，一个模型要算显露，须可识别（从一堆数据里站出来，有轮廓）、可维持（换一批人、隔十年再做，它还在）、可表达（能写下、能教、能传）。S 是显露，不是“结构”；一个尚在争论、只有少数人看得出轮廓的假说，也是显露，只是显露得弱。

D（差异序列），在这里指“允许从现实里切出哪些增量、按什么次序切、哪些切下来的差异被放大、哪些被当作噪声扔掉”的那条推进链。它不是“步骤”，也不是笼统的“变化”。第谷二十年记下的火星位置是一串差异；开普勒决定拿哪几组数据比、拿什么曲线去套、8 角分的偏差算不算数，这才使这串差异成了序列。

E（特征纠缠网络），在这里指科学得以发生的土壤：理念工具（欧几里得的比例论、坐标系、微积分本身）、仪器（望远镜、摆钟、温度计）、共同体（学会、期刊、师承、同行评审），以及它们之间缠在一起的方式。E 不是“背景”或“环境”的好听名字：背景可以替换而不影响剧情，E 抽掉了，发生就没有了。没有精确到秒的计时手段，伽利略的斜面实验连“时间的平方”这个差异都切不出来。

> **三大方程在本章的用法** 三大方程写作 $S = F(D, E)$，$D = G(S, E)$，$E = H(S, D)$。三个字母 F、G、H 不同，是为了提醒这是三种不同的生成关系。必须说清：F、G、H 只是**关系的占位记号**，不是函数。本书没有、也不打算给出它们的变量、定义域、函数形式和时间尺度；任何把它们写成可代入数值的计算式的做法，都是误用。 第395卷区分三层用法：**概念层**——提醒不要漏掉互生关系；**模型层**——须补出变量、时间尺度、更新规则、竞争解释；**定理层**——列明假设后给出证

> 明。本章只在概念层使用三方程，并在末节给出它要多交出什么可检验的区分，否则它只是在整理语言。

◆ 二、S＝F(D,E)：模型由切法和土壤共同生成

第一个方程说：一个科学模型，是由“允许切出哪些增量”和“理念工具、仪器、共同体”共同生成的。

拿伽利略的落体定律来看。他要的那个　S——下落距离与时间的平方成正比——在今天的教科书里只占一行。可它的发生，需要 D 和 E 两边同时到位。D 一侧：他决定不去追问“物体为什么下落”（亚里士多德一路问的是这个），而是只切“经过的距离”和“所用的时间”这两种增量，并且把斜面的倾角作为可调的一格，用来把下落“放慢”到可测的程度。这是一个大胆的切法：它主动扔掉了颜色、材质、目的这些在当时看来同样重要的差异。E　一侧：他手里有水钟（用流出的水的重量计时）、有打磨过的斜槽、有欧几里得比例论这一整套“按比例说话”的理念工具——他的结论最初正是用比例而不是用代数式写下的。如果只有切法而没有水钟，距离—时间的差异切不精；如果只有水钟而没有那种切法，量得再准也不会想到去比“时间的平方”。

这正是本书在第10章说的“截取即丢弃”在发生学上的位置：切法本身就是 D 的一部分，它决定了这一次积分出的 S 会长成什么样。厨房里也是这样：同一锅汤，一个人尝咸淡，一个人量温度，一个人看颜色，三人各自“积分”出来的“这锅汤怎么样”，是三个不同的结论。不是汤有三个，而是切法不同；而能切成什么，又看各人手边有勺子、温度计还是一双好眼睛。

这一方程的锋芒在于：它不允许把模型当作“自然本身的样子”被直接读出来。第3章说过，科学家所说的“自然界”是理念物理界——经过测量、理想化、剔除噪声之后剩下的部分。现在可以补一句：那张筛子的网眼，一半由 D 定（切哪些增量），一半由 E 定（手里有什么工具、共同体认什么）。

◆ 三、D＝G(S,E)：已有理论规定下一次怎么微分

第二个方程常常被漏掉：已经显露的模型，会反过来规定下一次怎么微分。

牛顿之后的天文学家观测行星，已经不是“从零开始看天”。他们先用牛顿引力算出一颗行星“应该”在哪里，再测它“实际”在哪里，记下的是两者之差——残差。残差这种差异，在牛顿之前根本不存在：没有那个先算出来的“应该”，就没有那个“差”。换句话说，S 生出了一种新的 D。十九世纪天王星的轨道偏差、水星近日点多出来的那一点进动，全都是以这种“相对于理论的残差”的形式第一次被切出来的（第16章细讲）。

E 也参与这种规定。望远镜的口径、照相底片的感光度、计算表格的精度，决定了残差能切到几位小数。共同体的规矩决定了多大的残差“值得报告”、多大的残差“应当归入观测误差”。一位天文学家看到偏差，先问的不是“牛顿错了吗”，而是“是不是我的仪器没校准”——这是 E 在替 D 把关。

带孩子的人都熟悉这一点。家长心里已经有一套“孩子该怎样长大”的模型（S），这套模型决定了他每天盯着看什么：身高体重、考试分数、有没有礼貌。看的东西又依赖家里有什么（E）：有体重秤就天天称，有成绩单就天天比。至于孩子某天对着蚂蚁看了一个下午这种差异，因为不在模型的格子里，多半不会被“切”下来。不是这差异不重要，而是 D 已经被 S 和 E 规定好了。

这一方程解释了一件看似奇怪的事：科学越成熟，它看见的东西越精确，也越窄。精确，是因为 S 和 E 把切法磨得锋利；窄，是因为同一把刀只切它规定要切的那一处。

◆ 四、E＝H(S,D)：成功的模型沉淀为土壤

第三个方程说：成功的模型和反复走通的切法，会沉淀下来，成为下一代科学的土壤。

沉淀有好几种形态。最显眼的是**教科书**：一个模型一旦进了教科书，它就不再以“某人的主张”出现，而是以“事实”出现——学生学的是牛顿定律，而不是“牛顿认为”。其次是**仪器**：温度计的刻度是按某种热学理论定的，示波器的设计里写满了电磁理论，卫星导航系统的钟必须按相对论修正才能定位准确——理论被铸进了铜和硅里，用仪器的人不必再懂那套理论。再次是**数据库与标准**：国际单位制、标准参考数据、各种公共数据库，把无数次切出来的差异整理成后人可以直接调用的底料。到了今天，又多出一种：**AI 训练集**。几十年间积攒下的结构数据、气象再分析数据、文献语料，成了机器学习模型的“土壤”；某种意义上，过去的 S 和 D 被压进了 AI 的权重（第26、27章细谈）。

厨房里的例子：一道菜做成功了十次，菜谱就写下来了；写下来之后，家里会为它专门买一口锅、备一种调料；锅和调料又反过来让这道菜更容易做成，也让别的菜不知不觉往这道菜的口味靠。菜谱是 S，十次的火候与试错是 D，锅与调料就是沉淀出的 E。

沉淀是科学能累积的原因，也是科学会板结的原因。第395卷说 E 有四相——建立、巩固、激活、释放；只建立巩固、不激活释放，土壤就板结。一台仪器的设计里铸进了某个理论，它测出来的东西就天然与那个理论相容；与理论不相容的信号，可能在仪器的滤波电路里就被当作噪声滤掉了。

◆ 五、微积分自己也在这三个方程里

把三个方程放在一起看，会发现微积分在其中的位置很特别：它同时出现在三个字母里。

在 S＝F(D,E) 里，它是 E 的一部分——是理念工具，是那把切增量、拼整体的刀和胶。没有它，“瞬时速度”这种增量连切都切不出来。在 D＝G(S,E) 里，它规定了“什么算一次合格的切”：一个现象若能写成微分方程，它就被认为“被理解了”；写不成的，往往被归到“经验规律”“唯象描述”一类，低一等。在 E＝H(S,D) 里，它本身又是沉淀的产物：牛顿和莱布尼茨那一代人在力

学、几何问题上反复走通的手法，经过欧拉、拉格朗日、柯西一个多世纪的打磨（第7、9章），才沉淀成今天每个理工科学生必修的那门课。

也就是说，微积分既是科学循环的土壤，又是科学循环的果实。这本身就是三方程“同时互生”的一个样本：工具塑造了模型，模型又反过来塑造了工具。这里要守一条边界：说“同时互生”，不等于宣布了就解除了循环论证的嫌疑。它能不能站得住，要看在具体的历史材料里，能否把“工具”“切法”“模型”三者分别指认出来，并且看它们的先后牵动能否被独立的史料核对。比如，可以查某一种新的数学工具（傅里叶级数、张量分析）进入某一学科之前与之后，该学科报告的问题类型是否发生了可见变化。

用开车来比方：一条路修好了，车就多了；车多了，路又被拓宽；拓宽了，沿路又长出新的镇子，镇子又需要新的路。你说不清是路先还是车先，但每一段的修路记录和车流记录都在，可以一段一段核对。三方程的用处，正在于提醒研究者把这三本账分开记，而不是混成一句“相互作用”。

◆ 六、六路径：科学从哪一维起手

三大方程是一张静态的关系图：三维同时互生，没有谁先谁后。可是一个活的研究者、一个活的研究共同体，注意力一次只能落在一处，必须从某一维起手。六路径讲的就是起手次序。本书只挑三条与科学最相关的来说，并且先声明：**没有哪条路径天生更高级，只有对不对得上任务 DNA**；六路径是提问助手，不替代研究设计。

健康的科学的起手：D→S→E。 这里的“健康”说的是一门科学的状态——它还会被惊讶、还接得住真反例（第16章的“验尸三问”），与下面的“固化”相对；不是医学意义上的“健康科学”。一个新领域、一个真正的反常面前，最吃紧的是“它是怎么一步步走出来的”——跟住差异序列，让模式自己显露，然后再让它沉淀为土壤。开普勒对着第谷的火星数据，试了一种又一种圆形组合，最好的一种仍与观测差了约 8 角分；他没有把这 8 角分归入误差，而是跟着这串差异继续走，最终放弃了圆，显露出椭圆。之后，椭圆轨道进入星表编制

与天文计算，沉淀为 E。看病也一样：一个高明的医生面对一个说不清的病人，先跟症状的来龙去脉（D），让病的轮廓自己显出来（S），再追问生活与家族（E），而不是先按一张检查单套。

常规科学的固化：E→D→S。 一个领域成熟后，最自然的路径是从已有土壤起手：先看手里有什么方法、仪器和公认的模型（E），再在这些方法能处理的范围里选问题、切增量（D），最后得出与已有框架相容的结论（S）。这条路径本身并不错——第395卷把 E→D→S 列为“论证一个结论”的最优起手，库恩也说过常规科学是极有成效的解谜活动。问题出在它被用成**唯一**路径的时候：研究者只去微分自己积分得了的东西。只研究方程能解出来的情形，只测量仪器能测的量，只问数据库里有字段的问题。老笑话说，一个人夜里在路灯下找钥匙，别人问他钥匙是在这儿丢的吗，他说不是，但只有这儿有光。路灯就是 E：理念工具与仪器照亮了一小块地，D 只在亮处切，S 自然只长在亮处。线性化、可积系统、“代表性个体”这类工具都极有价值，但当一个领域的问题清单几乎完全由“能不能解”来决定时，它就落进了路灯效应。

工程应用的常路：S→E→D。 工程的任务 DNA 是“达成一个目标结构”：一座桥要立住，一台发动机要达到某个功率。所以它先立靶子（S），再盘家底——有什么材料、什么规范、什么计算软件、什么施工队（E），最后才定路线——在哪里取样、做哪些荷载试验、哪几处要实测（D）。注意最后这一步：工程的 D 恰恰发生在第14章所说的“再微分”处，是在这一个桥址、这一块地基上把积分常数 C 借回来。工程师不重新推导材料力学，但他一定要钻孔取土。

六路径与科学的三段 | 起手 | 对应的任务 DNA | 在微积分循环中的位置 | 被误用时的症状 | |---|---|---|---| | $D \to S \to E$ | 理解一个过程 | 从现实的差异起手，积分后再沉淀 | 跟得太久不肯收敛，只有数据没有模型 | | $E \to D \to S$ | 论证一个结论 | 从既有理念工具起手，只切工具能处理的增量 | 路灯效应：只微分积分得了的东西 | | $S \to E \to D$ | 达成一个目标结构 | 从模型起手，落到具体点上再微分 | 拿模型硬套现场，不借 C | 其余三条（$S \to D \to E$ 讲解体系，$E \to S \to D$ 描绘一个学科的全景，$D \to E \to S$ 干预一个卡住的研究进程）同样会在科学中出现，例如写教科

> 书多走 $S \to D \to E$。路径是起手的纪律，不是全程的枷锁；六条都含三维，区别只在次序。

起点错配的代价最高，而且不报错。一个研究“某种新现象为何出现”的课题组，若从既有工具起手、只做工具能做的测量，可以多年发表、指标漂亮，却始终没碰到那个现象的来路。这不是不勤奋，是勤奋被放大了起点的错误。

◆ 七、加了 D 和 E，多出了什么可检验的区分

按第395卷的方法层失败条件，必须回答：把科学写成三方程，比通常的“理论—观测—修正”说法多出了什么？如果什么也没多，三方程就只是给熟知的科学史换了一套名字。

本书认为至少多出三处可以拿材料去查的区分。

第一，**D＝G(S,E) 预言了反常的“形状”**。如果下一次怎么微分是被已有理论规定的，那么一个领域里被正式报告的反常，大多应当以“相对于现行模型的残差”这种形式出现，而不是以与现行模型无关的新现象出现；并且这一比例在仪器更新前后会发生可见的变化。这可以用某一学科的期刊档案去查。若查出来反常的报告形态与现行模型、现行仪器毫无关系，这一方程在科学上就只剩比喻。

第二，**E＝H(S,D) 预言了沉淀的偏向**。仪器、数据库、训练集既然是成功模型的沉淀，它们就应当系统性地偏向产生与那些模型相容的数据。这一点同样可查：在同一对象上，用依据不同理论设计的仪器去测，若不存在任何与设计理论相关的系统偏差，这一方程的“沉淀”说就要收窄。

第三，**六路径预言了起点与成果类型的关联**。若一个领域长期以 E→D→S 起手为主，它产出的应当多是对既有模型的精化，少有新对象的显露；而真正的新对象，应更多出现在以 D 起手的工作里。这一区分可以拿科学史的个案组去比，但要预先定好什么算“新对象”，不能事后挑例子。

这三条都还没有被系统检验过。本书把它们写在这里，正是为了让三方程从概念层往模型层挪一步，而不是停在宣布。

◆ 八、收束

三方程把前面拆开的三段重新缠回一处：模型由切法与土壤共同生成，模型又规定下一次怎么切，切成功了又沉淀为土壤。它是一张照片，画的是三者怎样互相支撑，却看不出这个圈为什么会转、什么时候会断。卖菜的老人哪天发现账怎么也对不上——这时候，圈就要动了。让这个圈动起来、甚至让它断裂重组的引擎，是下一章的 123 原理。

带 走 的 话

三方程在本书只作概念层的占位，F、G、H 不是可代入数值的函数。它提醒的是三重牵扯：模型由切法与土壤共同生成，模型又规定下一次切什么，走通的模型与切法再沉淀为教科书、仪器、数据库和训练集。这套说法要算比“理论—观测—修正”多出东西，就得交出可查的区分：反常以相对现行模型的残差出现，仪器偏向与设计理论相容的数据，起手路径与成果类型相关。三条都还没有被系统检验。

小 账

- 本章的卖菜老人、带孩子的家长、尝汤与菜谱、修路与车流、路灯下找钥匙均为构想场景或比喻，不对应真实个案。
- 本章所述伽利略、开普勒、天王星的史实按通行科学史叙述，水星近日点一事见附录I第I53条；本章没有使用任何调查数字。
- 本章只在概念层使用三方程，不主张它已达模型层或定理层；六路径只讲了三条，不主张哪条路径天生更高级。
- 本章主张的推翻条件见附录G第G27条；SDE 词汇的整体增量另见第G6条。

16

第 16 章

123 原理：反常、危机与回写

> 这一章的问题：水星近日点多出的那一小份，人们为什么先去找一颗看不见的行星，过了半个多世纪才肯动引力本身——一门科学究竟在什么时候、怎样改掉自己？

家里有一本记了很多年的家用账。每月月底，按收入、房贷、伙食、水电几栏一加，和存折上的数一对，差个几块钱，大家都不在意，记一笔“零星”了事。有一年开始，差额越来越大，先是几十，后是几百。头几个月，记账的人照旧把它塞进“零星”，再后来，“零星”一栏比伙食还大。家里于是开了一次会：是有人乱花钱，是银行扣了什么费，还是账本的栏目本身已经不对了？一查才发现，孩子上了网课，各种自动续费的小额订阅散落在好几张卡上，原来的几栏根本没有它们的位置。账本被重新设计了：多了“订阅”一栏，每张卡的流水单独对一遍，“零星”这一栏的上限也被规定下来，超过就必须追查。

这件小事里有三步：越来越大的差额，和一本老账本之间的拉扯；拉扯到某一刻，账本的样子变了；新账本反过来改了全家每个月对账的方式，也改了大家对“钱去哪了”的感觉。上一章的三方程是一张照片，这一章要讲的，是让照片里的东西动起来的那台引擎——123 原理。

◆ 一、123 原理只讲一件事

先把原理摆正。第395卷的 123 原理是一个循环，只讲一件事：

> ① D 与 E 相互矛盾，张力累积；② 矛盾逼出 S 改变；③ S 的改变回写 D 与 E。被重置过的 D、E 之间又生出新的矛盾，回到①。

放到科学里：D 是一次次切出来的增量，特别是第15章说的“相对于现行模型的残差”；E 是已沉淀的理念工具、仪器、教科书和共同体的规矩。D 总在往前推——仪器更精，样本更多，测得更远；E 总有惯性——教科书不会每年重写，共同体不会因为一个怪数据就改口。一个要变，一个要稳，张力就累积起来。累积到旧的显露再也装不下，模型就得改，S 位移。改了之后，新模型又回过头来，重新规定什么算一次合格的切（重置 D），并沉淀为新的教科书、仪器和训练方式（重置 E）。

这里要守三条纪律。第一，**矛盾是引擎，不是故障**：一门再也没有反常、再也没有争论的学科，不是完成了，是熄火了。第二，**S 是矛盾的结算点，不是起点**：新理论不是某个天才凭空先想出来，再拿去套现实的；它是被逼出来的。第三，**回写最容易被漏掉**：科学史书写到“新理论诞生了”往往就停笔，可正是新理论怎样改了下一代人的切法和工具，决定了这次突破是不是只“发生了一次”。

还要守一条边界：123 原理不是“任何三步都算”。“提出假说—实验—修正”是一种三步，但它不必然是 D—E 矛盾逼出 S 改变；本书只在能指认出 D、E 各是什么、矛盾累积在哪里、S 改了什么、又回写了什么的时候，才用这个名字。

> **把底本的“八步循环”换成一个原理** 底本原作把科学知识写成八步：“新差异—特征律—微分—积分—知识—异常—再微分—再积分”，并以“创生律、成就律、演化律”分别对应微分、积分与微积分形式的扩张。本书改用现行口径： - 前四步对应全书三段中的①微分、②积分（第10—11章），不另立原理； - “异常”是 123 原理第①步的显影：D（残差）与 E（现行理论、仪器、共同体）的矛盾； - “再微分、再积分”是第②步（S 改变）与第③步（回写）的科学形态。 其中“再微分”一词在本书有两种用法，须分清：第14章的再微分是**应用的再微分**——用既有模型落到一个

具体点上，借回积分常数 C；本章的再微分是**回写的再微分**——用改过的模型重新规定从现实里切什么。前者单次不改模型（第14章方框中的情形 a），但它在现实中引出的新差异若系统性累积，就成为第①步的材料；后者改了切法。二者都回到现实界，所以同名；但只有后者是 123 原理的第③步。

◆ 二、水星近日点：同一个动作，两种结局

十九世纪中叶，天文学家用牛顿引力计算行星轨道，已经能算到很细。天王星的位置总是对不上，残差一年年变大。这是 D 在推：观测精度上来了，切出来的残差越来越清楚。E 在稳：牛顿力学是当时最成功的理论，没有人愿意为一颗行星动它。张力的出口，是勒维耶等人在 1846 年前后的做法：不改牛顿，而是假设外面还有一颗没看见的行星在拉天王星，并据此算出它该在哪里。望远镜一指，海王星果然在那附近。

这是科学史上最漂亮的胜利之一。但请注意它的结构：矛盾累积了，可是 S 没有改——牛顿理论原封不动，只是在模型里添了一个天体。E 赢了，而且因为赢了，它更稳了。

同一个勒维耶，后来又盯上了水星。水星近日点每世纪都在缓慢挪动，扣掉其他行星的影响之后，还多出一小份——今天的公认数值约为每世纪 43 角秒；勒维耶 1859 年给出的是约 38 角秒。他用了同一招：假设水星轨道内侧还有一颗行星，人们甚至给它起了名字“火神星”。此后几十年，不少人报告看到过它，又都没能被确认。这一次，同样的动作没有带来同样的结局。残差不肯走，共同体又试了别的补丁：修改引力的平方反比律、假设太阳周围有物质环……每一个补丁都只能解决这一个残差，又会在别处引出新的不对。

矛盾就这样拖了半个多世纪。直到 1915 年，爱因斯坦用广义相对论算出的近日点进动，与那多出来的一小份对上了。这一次，S 改了：不是加一颗星，而是引力被重新理解为时空的弯曲。

两个故事放在一起，最能看清 123 原理的第①步不是“出现了反常”那么简单。反常是 D 与 E 之间的拉扯，拉扯的结果要么被现有 S 吸收（海王星），要么

吸收不了、张力越积越大（水星）。事前，没有人能仅凭“出现了残差”就判断它属于哪一种。这也是本书给自己的限定：123　原理是组织框架，它告诉你该去看哪里，却不能替代具体机制，也不能替你预言某个残差是下一个海王星还是下一个水星。

买菜的经验里也有这一课。秤上的数和你估的对不上，第一次你会怀疑自己眼力；第二次你会怀疑那棵菜沾了水；第五次、第十次，你才会怀疑那杆秤。先找补丁，补丁不灵了再动秤——这不是保守，而是合理的次序；只是要记得，补丁有用完的时候。

◆ 三、黑体辐射：连“怎么切”都被改了

如果说水星的故事改的是“引力是什么”，黑体辐射的故事改得更深：它改到了微分本身的切法。

十九世纪末，人们测量一个加热的空腔向外辐射的能量怎样随频率分布。实验做得越来越精，形成了一条清楚的曲线。理论这边，按照当时公认的统计物理与电磁理论推下去，高频一端的辐射应当无限增长——后来人们把这个困境叫作“紫外灾难”。实验的 D 说曲线在高频处回落，理论的 E 说它应当一路往上。这是一个教科书式的 D—E 矛盾。严格说，普朗克 1900 年直接面对的，是当时通行的维恩公式在低频红外一端与新测量的偏离；高频发散的那一面，是此后几年经瑞利、金斯等人才讲清的。两面合起来，才是这一处矛盾的全貌。

1900 年，普朗克找到了一个与实验完美吻合的公式。为了从理论上推出它，他不得不做一个当时看来很别扭的假设：振子的能量不能连续地取任意值，只能一份一份地取，每一份的大小与频率成正比。

> **一个被改掉的切法**　普朗克假设振子能量只能取 $E = nh\nu$（$n = 0, 1, 2, \ldots$），其中 h 后来称为普朗克常数。由此得到的辐射公式在低频端趋近经典结果，在高频端因指数因子而回落，与实验相合。　从本书的角度看，要紧的不在公式本身，而在它动了哪一处：经典处理里，能量被当作可以无限细分的连续量，正是这一前提允许对它作 dE 那样的微分。普朗克的假设等于宣布：在这一类对象上，能量这个量不允许被切

> 得无限细。微积分“截一刀、再取极限”的基本动作（第10章），在这里撞上了现实的一个硬颗粒。　这并不意味着微积分失效——量子力学照样大量使用微分方程。改掉的是“对哪个量、在哪个尺度上可以假定连续”的切法，即 D 的规定。

普朗克本人起初并不把这当作物理实在，更多当作一个让公式成立的计算手段。三十年后，他在 1931 年 10 月 7 日写给美国物理学家伍德（R. W. Wood）的信里回顾，把当年那一步概括为“一次绝望之举”（ein Akt der Verzweiflung）：理论解释必须不惜代价找到。真正把“能量子”当作实在的，是 1905 年爱因斯坦对光电效应的解释，以及此后十多年玻尔、索末菲等人的一连串工作。这正好显出 123 原理第②步的样子：S 的改变不是一个瞬间，而是一段被矛盾一点点逼出来的过程，起初连改动者本人都不确定自己改了什么。

◆ 四、回写：被观察，而不是被宣布

现在来看最容易漏掉的第③步。新理论立起来之后，反过来重置了什么？

先说　D。广义相对论之后，“去测光线经过太阳时偏折多少”“去测引力场中的钟走得慢多少”成了合格的、值得做的切法；而在牛顿框架里，这些差异要么切不出来，要么不会有人去切。量子论之后，光谱线的分裂、原子的稳定性这些原先被当作化学细节的差异，被重新切成物理学的头等问题。每一次　S　的改变，都重新划定了现实里哪些增量值得截取。

再说　E。新理论进了教科书，培养出一代只会用新方式思考的学生；新理论被铸进仪器，例如卫星导航系统里的时钟修正，用户根本不需要懂相对论，却每天依赖它；新的共同体、新的期刊、新的实验室分工也随之出现。

这些回写都是可以被观察到的：教科书的目录变了，仪器的设计规格变了，期刊里报告的问题类型变了。第395卷要求回写“要被观察，不能被宣布”。说一个理论“回写了整个文明”而拿不出这样的痕迹，就是宣布，不是观察。

还要记得回写是中性的，**也有负向回写**。一个成功的理论沉淀为仪器之后，仪器会把不合理论的信号滤成噪声；一种成功的切法成为规矩之后，会把切不进

去的现象挡在门外。第15章的路灯效应，就是回写的副作用：昨天的突破，变成了今天照亮一小块地、也遮住了其余地方的那盏灯。家里那本账也一样：新设了“订阅”一栏后，大家对订阅格外警惕，却可能因此不再细看伙食一栏里慢慢多出来的外卖。

> **123 原理在科学史上的检查单** 用 123 原理分析一段科学史，至少要逐项写出： 1. D 是什么：哪一种增量、由什么仪器切出、精度到哪里； 2. E 是什么：哪个理论、哪些工具、哪个共同体的什么规矩在抵抗； 3. 矛盾怎样累积：有哪些补丁，各失败在哪里，持续多久； 4. S 改了什么：改的是对象、定律，还是切法本身； 5. 回写的痕迹：哪些教科书、仪器、问题清单在之后可见地改变了；有没有负向回写。 写不出第5项的，不能说“发生了回写”；写不出第1、2项的，不能说“这是 123 原理”。

◆ 五、接住库恩、波普尔、费耶阿本德

科学怎样在反常中更新，二十世纪的科学哲学已经说了很多。底本原作对三位哲学家的态度不一：吸纳库恩，修正波普尔，批判费耶阿本德。本书改用第395卷“先接住、再补齐”的方式：先说清他们看到了什么，而且看对了，再说 SDE 想补上哪一块。补齐不是比他们高明，而是站在他们打下的地基上多盖一层。

库恩看到，科学大部分时间处在“常规科学”之中，在一个范式里解谜；反常起初被搁置或打补丁，积累到一定程度引发危机，最后经由一场不能用旧规则裁决的革命，换成新范式。这幅图景对得极好，水星与黑体辐射两个故事几乎就是它的注脚。第395卷与之的对应是：常规科学大致相当于成熟态，反常累积大致相当于 D—E 矛盾的累积，危机近于退化，革命近于重组。SDE 想补齐的有两处。其一，库恩是在事后描述，SDE 希望给出一些事前可用的检查线索，例如上面的检查单、第395卷的“验尸三问”（共同体还活着吗？还会被惊讶吗？还接得住真反例吗？）。其二，把“回写”单独提出来：库恩讲了新范式如何取代旧范式，SDE 更关心新范式如何改写下一轮的切法与土壤，包括负向的那一面。第395卷有一句判词：库恩描述天气，SDE 想做气象学。这是志向，不是已完成的成绩；气象学要靠预报的准确率来证明自己。

波普尔看到，科学理论不能被经验最终证实，却可能被经验反驳；科学的尊严在于它把自己暴露在被驳倒的风险里。这一条，本书全盘接住，而且当作纪律在用——每一编都写“怎样算错”，正是波普尔精神。SDE 想补齐的是：被反驳之后的那一段。一个残差出现时，被打掉的到底是理论的核心，还是某个辅助假设？勒维耶对天王星的处理，等于保住核心、修改辅助假设（天体清单），结果成功了；对水星用同一招，失败了。只看“证伪”这一刻，分不出两者；要看 D—E 张力在之后几十年里怎样演化，补丁是越打越少还是越打越多。这层意思，波普尔之后的拉卡托斯在 1970 年的《证伪与科学研究纲领方法论》中已经讲过“进步的”与“退化的”研究纲领（progressive / degenerating research programmes），SDE 在此并不宣称首创，只是把它放进 D—E—S 的三维里，并补上回写一笔。

费耶阿本德看到，真实的科学史远比任何方法论规则凌乱：伽利略并没有严格遵守后人总结的“科学方法”，许多突破恰恰是违反当时规则得来的；任何一条固定的方法规定，都可能在某个关键时刻阻碍进步。他那句著名的口号常被简化为“怎么都行”。底本原作对此作了严厉批判，认为他把自由误解为无序。本书收回这种批判的口气，改为接住：费耶阿本德看到的，正是第15章所说的路径固化之害——一个共同体若只允许一种起手路径，就会把切不进那条路的现象全部挡在门外。他为方法的多元辩护，在 SDE 里对应着六路径的纪律：起手要按任务 DNA 选，而不是按惯例选。SDE 想补齐的是：多元不等于没有承接。第395卷讲介生态，要害在“承接”——差异能够进入，后果能够保留，局部结构能够形成，结构能被后续证据修订。“怎么都行”保住了差异的进入，SDE 补上后面三条：进来的东西要能被记录、被检验、被修订。“猜想要野，承接要严”，费耶阿本德守住了前半句，后半句由共同体的 E 来承担。

三位哲学家放在一起，正好各守了 123 原理的一处：库恩讲了矛盾怎样累积成范式的改变，波普尔讲了矛盾怎样被认真对待而不被抹平，费耶阿本德讲了 D 不能被一种路径锁死。SDE 想做的，是把这三处放进同一个循环，并补上回写。

◆ 六、收束

科学没有真正的终点，因为 123 原理的循环在每一次成功之后都会被重新点着：新的 S 规定了新的切法，新的切法会切出新的残差，新的残差又会与新沉淀下来的土壤拉扯。熄火的方式也有三种，恰好对应本章开头那三条纪律被违反——矛盾被压平、某一维被抬成不容质疑的第一因、回写丢失——每一种都在科学史上出现过。到这里，第三编的核心机理已经讲完。下一编要看这个圈怎样把微积分自己推向新的地盘，第一步是从“一点在变”走到“一片在变”的偏微分方程。在此之前，这一编还要交出一份清单：它的哪些主张，在什么情况下应当收回。

◆ 七、这一编怎样算错

第三编（第10—16章）的主张：科学是“现实里微分、理念里积分、落地时再微分”的一个圈；微分切掉纠缠，积分只得 $f+C$，C 须在落地时从现实借回；三段由三方程互相牵扯，由 123 原理推动。下列任一条成立，相应主张须收回或收窄：

1. **若**能找到大量情形，模型在新的具体点上应用时完全不需要补充现场信息（初始条件、边界条件、个体参数）即可达到与原研究相当的精度，**则**“积分常数 C 必须外借”（第12、14章）须收窄为只适用于部分学科的命题。
2. **若**复现失败率与学科“其他条件不变”的切割强度、与初始条件和情境依赖程度均无可见关联，**则**第12章以复现危机作检验把手的说法须撤回，“切得越狠、C 越难借”只剩比喻。
3. **若**互逆的基本定理在差异尚未稳定（不连续、不可积）的对象上照常成立且无需改写，**则**第13章“互逆的条件即差异已稳定”须收回。
4. **若**技术与应用史上的关键进展大多不经过“在具体点上再微分”这一步，而是直接由理论推导完成，**则**第14章把再微分作为技术发生之处的枢纽地位须降格。

5. **若**加入 D、E 之后，对科学史个案（如第15章所列的反常报告形态、仪器沉淀偏向、起手路径与成果类型）给不出任何普通“理论—观测—修正”叙述所没有的可检验区分，**则**第15章的三方程只算整理语言，不算解释。
6. **若**用本章检查单分析的科学史个案中，写不出可观察的回写痕迹的占多数，或同一个案在结果出来之后可以随意调换 D、E 的指认而都“说得通”，**则**第16章用 123 原理解释科学更新的主张须收回为单纯的叙事框架。

带走的话

反常不等于理论该换。它是新切出的残差（D）与已沉淀的理论、仪器和共同体规矩（E）之间的拉扯：可能被旧的显露吸收，如海王星；也可能吸收不了，逼出 S 改变，如水星近日点与黑体辐射。事前单凭残差分不出是哪一种。123 原理最容易漏掉的是第③步：新理论怎样改了下一轮切什么、用什么工具，要拿得出教科书、仪器、问题清单上的痕迹；拿不出，就只是宣布。

小账

- 本章开头的家用账、买菜看秤均为构想的日常场景，无真实原型。
- 本章所引普朗克 1931 年致伍德信、拉卡托斯 1970 年论文的出处见附录I第I33、I34条；勒维耶与海王星、水星近日点进动数值见第I53条；火神星与黑体辐射的来龙去脉，取科学史通行叙述。
- 本章对库恩、波普尔、费耶阿本德的转述取其通行要义，不是逐字引文；“验尸三问”与“库恩描述天气，SDE 想做气象学”出自第395卷。
- 本章不主张 123 原理能事前预言某个残差是下一个海王星还是下一个水星；它只指示该去看哪里，不替代具体机制。
- 本章主张的推翻条件见附录G第G28条；第三编整体的失败条件见本章末节。

第 四 编

扩张：微积分的自我更新，以及它的边

编 序

第四编只做一件事：看微积分在三百年里怎样被新的差异逼着长大，以及每长大一次，它的边划在了哪里。

第17章讲场：偏微分方程把一点的变化推到一片的变化，傅里叶把一片拆成许多简单的样子；内部的规律越普遍，越依赖边界，而边界条件又是一次从现实借来的C。第18章讲数值方法：把连续的场切成有限的格子，是又一次截取，被切掉的东西有名字，切法本身还会造反；数值分析是微积分家族里最懂自我审计的一支。第19章讲随机微积分：布朗运动处处没有切线，伊藤积分把焊缝从逐点焊死改为按统计焊牢，布莱克—舒尔斯公式则示范了焊得太牢的代价。第20章讲记忆与非局域：分数阶导数让“此刻”带上了过去，非局域算子让影响来自远处；伞场微积分在这里只保留限定条件下的局部逼近命题。第21章讲无限维：从变分法到希尔伯特空间，从路径积分到泛函分析，被微分与积分的对象变成了“所有可能的曲线”，并由此接上现代 AI 的数学。

读这一编时，可以带着一个问题：每一次扩张，接回了什么，又在别处切下了什么？

17

第 17 章

场：偏微分方程

> 这一章的问题：同一块冻肉，放在木案板上和放进铝盆里，化冻的快慢与次序就全变了——热传导方程一字未改，变了的那一部分是从哪里来的？

冬天早上，把一块冻得硬邦邦的肉从冰箱拿出来，放在案板上化冻。过了半小时去摸，边上已经软了，中间还是硬的；再过一会儿，靠近案板那一面比朝天那一面化得快——案板是木头的，还好；要是放在一只铝盆里，贴着盆底的那一面会化得更快。有经验的人还知道，切成薄片的肉比一整块化得快得多，用一碗冷水泡着又比干放着快。

这块肉没有“一个”温度。它每一处都有自己的温度，每一处的温度又都在变；而且每一处变得多快，取决于它周围比它热还是比它冷。化冻不是一个点的事，是一整片的事。

前几章说的微积分，大多是一个点的事：一块石头从多高落下、一颗行星在轨道的哪一点、一个病人体内药物浓度怎样随时间降。它们都只追踪一个量随一个变量（多半是时间）的变化，用的是常微分方程。可现实里大量的东西像这块肉：热在固体里散开，声音在空气里传，电和磁在空间里交织，风和水在流。它们要求我们追踪“一片”在变。这就是偏微分方程登场的地方。

◆ 一、从一点的变化到一片的变化

常微分方程只有一个自变量。“偏”字的意思是：现在自变量不止一个了——既有时间，又有空间里的位置。一块肉里某一点的温度 T，同时依赖它在哪

儿（x, y, z）和现在是什么时候（t）。于是要问两种不同的变化：守在一处不动，看温度随时间怎么变；在同一时刻沿着某个方向挪一步，看温度在空间上怎么变。前者叫对时间的偏导数，后者叫对空间的偏导数。“偏”，就是一次只让一个变量动、其余按住不动。

这一点用第10章的话说，就是“截一刀”时多了几个方向可以截。常微分是一刀，偏微分是几刀同时截，每一刀都按住其他方向。“其他条件不变”这个动作，在这里被做得更彻底，也更显眼：每一个偏导数，都是一次明确的“只看这一个方向，其余暂且冻住”。

偏微分方程，就是把这几种变化之间的关系写成一个等式：某一点在时间上的变化，等于它在空间上和周围的某种关系。说得更平白一点：**一点怎么变，取决于它的邻居**。

这句话里藏着一个重要的转变。经典力学里，一颗行星怎么动，取决于远处的太阳有多重、离它多远——力是隔空作用的。而场的方程说：一点只和紧挨着它的邻居打交道，影响是一点一点传过去的。菜市场里的消息就是这样走的：东头鱼摊降价，先是隔壁肉摊知道，再是对面菜摊，半小时后西头卖豆腐的也知道了。没有谁拿大喇叭向全场广播，消息是靠邻居传邻居走完全场的。

“场”这个概念的核心正在这里：空间的每一处都有一个量（温度、位移、电场强度、流速），这些量连成一片，而且彼此只靠近邻牵动。底本原作称之为“差异场”：不再追踪一个差异单元，而是追踪整片差异网络的协同演化。用现行口径讲，场是把现实里一片连成一体的纠缠，用“相邻点之间的差异”这种方式切下来，再在理念里拼回整体。

◆ 二、四个方程，四种一片

科学史上，有四个偏微分方程最能说明“一片在变”的几种样子。

热传导方程。 傅里叶 1807 年向巴黎科学院宣读关于固体中热传播的论文，1822 年出版《热的解析理论》，系统处理了固体里热怎样扩散。它说的是：一点温度升降的快慢，取决于它比周围的平均温度高还是低——比邻居冷，就会被邻

居焐热；比邻居热，就会把热散给邻居。化冻的肉、冬天的墙、刚出锅的馒头放凉，全是这个方程。它的脾气是“抹平”：任何高低不平的温度分布，随着时间都会被摊平，尖角会变圆，热点会变温。而且它不可倒放：你看不到一块温吞的肉自己变得一头冰一头烫。

波动方程。 拨一下吉他弦，弦上每一点的位移随时间来回振动，振动又顺着弦传开。波动方程说：一点的“加速度”取决于它与邻居的高低差。和热传导不同，它不抹平，而是把形状完整地往外送——你说话的声音能被对面的人认出是你，正因为声波在空气里传播时大体保住了形状。热传导的脾气是“忘记”，波动的脾气是“传递”。

麦克斯韦方程组。 1860 年代，麦克斯韦把当时已知的电学与磁学规律整理成一组偏微分方程，并补上了一项“位移电流”。补上这一项之后，方程组推出一个惊人的结论：变化的电场生出磁场，变化的磁场又生出电场，两者互相推着在空间里传播，速度算出来恰好与当时测得的光速相近。光由此被理解为电磁波。二十多年后，赫兹在实验室里造出并探测到了电磁波。这是第16章所说回写的一个清楚样本：一个在理念一侧积分出来的预言，重新规定了现实里该去切什么——去找一种此前没人想到要找的波；找到之后，又沉淀为无线电、雷达、手机这一整片新的土壤。

纳维—斯托克斯方程。 风、水、血液在血管里、空气绕过机翼——这些流动由纳维—斯托克斯方程描述。它和前面三个有一处根本不同：它是非线性的，流动的速度场会自己“搬运”自己（方程里那个速度乘以速度梯度的项）。所以一缕平稳的烟会在某个高度突然卷成乱涡，一条平静的河在桥墩后面会翻出漩涡。这个方程至今最出名的一点是：在三维情形下，人们还不知道它的光滑解是否对一切合理的初始条件都在全部时间里存在。这是克雷数学研究所 2000 年公布的“千禧年大奖难题”之一。

> **四个方程** 热传导：$\dfrac{\partial T}{\partial t} = k\nabla^2 T$。其中 $\nabla^2 T$（拉普拉斯算子）在某点的值，大致正比于“周围的平均温度减去这一点的温度”。 波动：$\dfrac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u$。同一个

∇^2，但放在二阶时间导数旁边，脾气从“抹平”变成“传送”。　麦克斯韦方程组（真空中，含电荷与电流源）：$\nabla\cdot\mathbf{E}=\rho/\varepsilon_0$，$\nabla\cdot\mathbf{B}=0$，$\nabla\times\mathbf{E}=-\partial\mathbf{B}/\partial t$，$\nabla\times\mathbf{B}=\mu_0\mathbf{J}+\mu_0\varepsilon_0\,\partial\mathbf{E}/\partial t$。在无源区域可推出　$\mathbf{E}$、$\mathbf{B}$ 各自满足波速为　$c=1/\sqrt{\mu_0\varepsilon_0}$　的波动方程。　纳维—斯托克斯（不可压缩）：$\rho\left(\dfrac{\partial\mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla p+\mu\nabla^2\mathbf{v}+\mathbf{f}$，$\nabla\cdot\mathbf{v}=0$。非线性来自 $\mathbf{v}\cdot\nabla\mathbf{v}$。 四个方程的共同之处：都只联系一点与它无穷小邻域里的量。“远处”的影响，全靠这种近邻关系一层层接力。

◆ 三、傅里叶的办法：把一片拆成许多简单的样子

方程写下来了，怎么解？傅里叶在研究热传导时给出的办法，影响远远超出了热学。

他的想法是：一条杆上任意复杂的初始温度分布，都可以看作许多条简单正弦曲线的叠加——一条起伏一次的，一条起伏两次的，一条起伏三次的……每一条简单曲线在热传导方程下怎样变化，是很容易算的：它保持形状，只是幅度按自己的速率衰减，起伏越密的衰减得越快。于是，把初始分布拆成这些简单成分，让每一个成分各自衰减，再把它们加回去，就得到了任意时刻的温度分布。这也顺便解释了热传导“抹平”的脾气：细密的起伏最先消失，剩下的总是越来越平缓的大轮廓。

这个办法有一个前提：方程是**线性**的，也就是两个解相加仍然是解。热传导方程、波动方程、真空中的麦克斯韦方程组都是线性的，所以叠加之法在它们身上大显身手；纳维—斯托克斯方程是非线性的，叠加不再成立，这正是它难的根源之一。

厨房里有个现成的比方。一锅复杂的老汤，滋味可以看作咸、鲜、酸、甜、香几种基本味道按不同分量叠在一起；如果每一种味道放久了各自按自己的速度淡下去，互不干扰，那么你只要知道每种味道一开始有多少、各自淡得多快，就能算出放了三小时的汤是什么味。可真实的汤不是这样：酸会让鲜显得更鲜，久

煮的香味会生出新的味道——味道之间相互牵动，就“不线性”了，拆开再加回去，加不回原来那锅汤。

傅里叶的拆法在发生学上很有意味：它是又一次截取。它把一片纠缠在一起的场，切成了一组彼此独立的简单样子。在线性的世界里，这一刀切掉的东西可以原样拼回；在非线性的世界里，切掉的正是各成分之间的牵动，而那一部分，叠加拼不回来。第22章讲“线性世界的黄昏”，起点就在这里。

◆ 四、边界条件：现实又一次借给理念的 C

到这里，要回到全书的主命题。第12章讲过，积分出来的只是 $f + C$，那个 C 不知道，必须在落地时从现实借回来。对一个常微分方程来说，借回来的东西不多：一个起点，有时再加一个起始速度——石头从多高、以多快的速度出手。

偏微分方程要借的，多得多。

先是**初始条件**：不再是一个数，而是一整片——开始那一刻，肉里每一处的温度各是多少。再是**边界条件**：在这一片的边上，发生了什么。肉放在木案板上还是铝盆里，泡在冷水里还是摆在空气中，表面裹没裹保鲜膜——方程本身一字不改，变的全是边界。同一个热传导方程，边界一换，化冻的快慢和次序就全变了。

这就是为什么说边界条件是现实又一次借给理念的 C，而且是借得最多的一次。方程只管“内部每一点怎样和邻居打交道”，这是理念一侧积分得出的普遍规律；它对“这一块肉在哪儿、碰着什么”一无所知。这些只能从现实里一个一个量来：这一个房间的墙外是什么温度，这一根琴弦两端是不是钉死，这一架飞机的机翼是什么形状，这一条河的河床是什么样子。

> **为什么内部规律必须向边界借**　以一维热传导为例：长为 L 的细杆，$\partial_t T = k\,\partial_{xx} T$，$0 < x < L$。只给方程，解有无穷多个。要得到唯一解，须另给：初始条件 $T(x,0) = T_0(x)$（整条杆的初始温度分布），以及两端的边界条件，例如 $T(0,t) = a$，$T(L,t) = b$（两端各接一个恒温源），或 $\partial_x T(L,t) = 0$（一端绝热）。换一种边界条件，解就是另一个函数。　这类问题称为初边值问题。数学上还

> 要求问题“适定”：解存在、唯一，且对初始与边界数据连续依赖（阿达马的提法，见第12章）。适定性本身是条件性的：同一个方程，配上不合适的数据，可能无解、多解，或对数据的微小扰动极度敏感。与第11章斯托克斯定理“边界决定内部”的说法对照：在那里，边界上的积分等于内部微分的积分；在这里，边界上给定的数据决定内部的解。两处都显出同一件事——在理念一侧，内部的规律是普遍的，而这一个具体的内部长什么样，要看现实给了它什么样的边。

有一种极端情形，把这一点显露得最彻底。肉放久了，化冻完了，整块肉的温度不再随时间变化，达到稳态；这时热传导方程里对时间的那一项变成零，只剩下 $\nabla^2 T = 0$——拉普拉斯方程。它的解有一个著名的性质：内部任何一点的值，都等于它周围一圈的平均；内部既没有最高点也没有最低点，最高与最低都只能出现在边界上。换句话说，在稳态里，**内部自己什么也不决定，一切由边界说了算**。房间里的温度一旦稳定，哪儿暖哪儿冷，完全看暖气片在哪、窗户在哪、外墙朝哪。数学上这类方程叫椭圆型，热传导叫抛物型，波动叫双曲型；三类方程向现实借的东西不一样：椭圆型主要借边界，抛物型借边界加一个初始时刻，双曲型除了初始位移还要借初始速度，并且它的信息沿着有限速度传播，远处的边界要过一会儿才“说得上话”。

这三种借法，在日常里各有其样：布置一个房间的冷暖，看的是墙和窗（只借边界）；一锅汤放凉要多久，看锅怎么放、汤一开始多烫（边界加初值）；一根拨动的琴弦发什么音，不只看它一开始被拉到哪，还要看手指放开那一刻是怎么放的（初始位移加初始速度）。借什么、借多少，是方程的类型规定的；借来的具体数值，只能现场量。

再微分（第14章）在这里有了最具体的样子。工程师手里有了经典的偏微分方程，还解决不了任何一个具体的工程问题；他必须为这一座建筑、这一条管道、这一片海域确定边界：哪里固定、哪里自由、哪里有多少热流进来。这一步往往比解方程本身更难，出错也最多。天气预报的难处，大半不在方程——大气运动的基本方程早已写下——而在初始条件：此刻全球大气每一处的温度、气压、风速、湿度，只能从有限的观测站、气球、卫星里一点一点估出来，再填进模型。观测不到的地方，就只能靠估计。这就是为什么天气预报越往后越不准：

方程把初始那一刻借来的误差，一层层传到了后面（第22章讲混沌时还要回到这里）。

看病也是这样。人体内药物的扩散、血液在血管里的流动，都可以写成偏微分方程；但同一套方程，用在这一个病人身上，血管的形状、管壁的弹性、哪里有斑块，全是他这一个人的边界条件。医学影像越来越被用来给这类模型提供病人自己的几何形状。一个已进入临床的例子是冠状动脉的“CT　　血流储备分数”（FFR-CT）：从病人的冠脉 CT 重建出他自己的血管树，再解血流方程，估算狭窄处的压力落差。美国 HeartFlow 公司的这类软件 2014 年获美国食品药品监督管理局（FDA）准许上市，2021 年美国心脏学会与美国心脏病学会的胸痛诊疗指南也已把它列为可选的评估手段。这本质上就是在大规模地从现实借 C。

◆ 五、场的方程切掉了什么

每一次截取都有代价，偏微分方程也不例外。

第一个代价，是**连续介质的假设**。写热传导方程、纳维—斯托克斯方程时，物质被当作可以无限细分的连续体，每一点都有确定的温度、速度、密度。实际上，肉是细胞组成的，水是分子组成的。这个假设在比分子大得多的尺度上极其好用，在稀薄气体、纳米尺度的流动里就要打折扣。厨房里的面粉看上去像一种连续的白色粉末，倒进量杯里可以当液体来量；可你要是想知道面粉怎样结块，就得回到一颗一颗的颗粒。

第二个代价，是**只认近邻**。场方程假定一点只和无穷小邻域打交道。可是有些材料、有些过程带着长程的牵连或长时间的记忆——一块受过力的橡胶记得它被拉过，某些介质里的扩散不服从“只看邻居”的规矩。第20章讲分数阶与非局域算子时，会看到人们怎样在这一处做修补，以及这种修补本身的适用条件。

第三个代价，在纳维—斯托克斯那里最显眼：**方程写得出，未必解得出，解得出也未必算得准**。湍流涉及从大漩涡到小漩涡的极宽尺度范围，在工程上几乎总要借助数值计算和经验性的湍流模型。也就是说，理念一侧的积分（方程）在这里并不完备，它要额外向现实借一些东西——不是边界，而是经验系数、模型

参数，这些系数又要靠风洞、水槽里的实验来定。第18章会细说，数值离散化本身又是一次截取。

> **三方程看偏微分方程的兴起（概念层）** 用第15章的语言，场论的兴起可以这样记：E 一侧，有了多元函数、傅里叶级数等理念工具，有了更精的温度计、检流计；D 一侧，人们开始切“同一时刻不同位置的差异”而不只是“同一物体不同时刻的差异”；S 一侧，“场”作为一种对象显露出来。之后，场论回写：法拉第的力线图像经麦克斯韦数学化之后，“场”沉淀为物理学的基本对象，后来的引力、量子理论都在场的语言里展开。 这里只是概念层的提醒，不是对这段历史的因果解释；具体哪项工具、哪次测量在多大程度上促成了哪一步，须由科学史研究分别核对。

◆ 六、内部越普遍，越依赖边界

偏微分方程给科学带来的扩张极大：天气、海洋、航空、电磁通信、建筑结构、医学影像，几乎所有“一片在变”的现代技术都建在它上面。它也把一个问题带得更尖锐：内部规律越普遍、越漂亮，落到现实里就越依赖那一圈边界和那一刻初值。

这对全书的主命题是一种加强。麦克斯韦方程组是理念一侧最成功的积分之一，四个式子统一了电、磁与光；但没有一根天线是靠这四个式子单独造出来的——每一根天线都要在这一个形状、这一种材料、这一个安装位置上再微分。纳维—斯托克斯方程写下近两百年，人们仍在问它的解在数学上是否总是好好地存在；而飞机早已飞了一百多年，靠的是风洞、试飞和一次次从现实里借回来的系数。

把冻肉放进铝盆，是你从现实里给它选了一个边界；把它切成薄片，是你改了它的几何。没有人需要会解热传导方程才会化冻，可正是这些日常的“边界选择”，在方程里占着方程自己给不出的那一部分。

偏微分方程的解，大多写不成公式，只能交给计算机一格一格去算。把连续的场切成有限的格子，这又是一刀。下一章，讲这一刀切掉了什么。

带 走 的 话

偏微分方程把“一点怎么变”写成“一点怎么变取决于它的邻居”，于是热、声、电磁与流动这些“一片在变”的现象有了理念里的写法。但方程只管内部的近邻关系，对这一块肉碰着什么、这一根弦怎样被放开一无所知。初始条件与边界条件只能现场量来，是现实借给理念最多的一次　C：内部规律越普遍，落地时越依赖那一圈边界和那一刻初值。

小 账

- 本章的化冻冻肉、菜市场传消息、一锅老汤、量杯里的面粉，均为说明用的日常比方，不对应真实个案。
- 本章所述傅里叶、麦克斯韦、赫兹、纳维—斯托克斯方程与千禧年大奖难题取数学史、科学史通行事实；FFR-CT 一例的出处见附录I（I35）。
- 本章“三方程看偏微分方程的兴起”方框只是概念层的提醒，不构成对场论历史的因果解释。
- 本章不主张方程可有可无；只主张落到一个具体对象上时，方程自己给不出边界与初值。
- 本章主张的推翻条件见附录G第G29条。

18

第 18 章

数值：离散化是又一次截取

> 这一章的问题：护士每四小时量一次体温，每个点都量得准，凌晨那次高烧却不在折线图上——把连续的方程切成格子交给计算机，切掉的又是什么？

住过院的人大概都记得这样的夜晚：护士每隔四个小时推门进来，量一次体温，在表格上点一个点。第二天早上医生翻看那张折线图，八点 37.2℃，十二点 37.5℃，十六点 37.3℃，一条平缓的小山包，看上去一切平稳。可是陪床的家属知道，凌晨两点病人曾经烧到 39℃，出了一身大汗又退了下去——那一次高烧恰好落在两个点之间，折线图上没有它。

这张表格不是错的。它的每一个点都量得很准。问题在于，一条连续的体温曲线被切成了六个点，点与点之间发生过什么，表格不知道，也没有办法知道。医生如果只看表，就会把一条被切过的曲线当成曲线本身。

第10章讲过，微分是在流动的现实里截一刀：取一点，看它附近的增量，假定“其他条件不变”，于是得到一个导数、一条微分方程。那一刀切掉了这一点周围的纠缠。可是故事到这里并没有结束。微分方程写好以后，绝大多数时候没有人能把它解成一个漂亮的公式。要得到答案，就得把它交给计算机，而计算机只认识有限个数。于是方程被再切一次：时间切成一格一格，空间切成一块一块，连续的导数退回成有限的差商。这第二次切片，就是本章的题目。

◆ 一、第二刀切在哪里

先把两次截取分清楚。第一次截取发生在现实界：工程师面对一根正在发热的钢梁，决定只关心温度，不关心钢梁内部晶粒怎样排列、车间里谁开了窗；他把这一份现实压进热传导方程。这一刀的代价，是现实里的纠缠被留在了方程外面。

第二次截取发生在理念界内部：热传导方程已经是一个理念对象了，它里面的导数是极限，是“无限细”的东西。计算机做不了极限，它只能取一个具体的步长，比如每隔一毫米取一个点、每隔一毫秒算一步，用这些点上的差商去顶替导数。这一刀切的不再是现实，而是那个已经被切过一次的理念方程。

底本对此有一段耐人寻味的说法：微分本来是差商取极限得到的，数值方法又把微分“退回”有限差商，这不是倒退，而是回到差异的本来面目。这个直觉抓住了一点：计算机里那个 $\Delta y/\Delta x$ 确实比 dy/dx 更像我们手里能拿到的东西。但是需要把话收窄。计算机里的差商并不是现实里原来的那个差异。现实里的温差是钢梁自己的；计算机里的步长 Δx 是人选的，选一毫米还是一厘米，取决于电脑有多快、预算有多少、交稿日期有多紧。数值差商是从理念方程里再截出来的，它离现实隔着两刀，而不是一刀。

这个区分有实际后果。打个比方：一个厨师先把一锅老汤的味道写成菜谱（第一刀），学徒再按菜谱把“小火慢炖两个小时”改成“每十分钟看一次火、加一次水”（第二刀）。学徒做出来的汤和菜谱不一样，可能是菜谱本来就漏了东西，也可能是学徒的十分钟一看切得太粗。两种失误的修法完全不同：前者要回厨房重新尝那锅老汤，后者只要把间隔改成五分钟。数值计算出了问题，第一件事就是分清是哪一刀的错。

◆ 二、三种切法

数值方法的家族庞大，这里只看三位最有代表性的成员。它们代表三种不同的切法。

第一种是有限差分。它最直白：在网格点上，用相邻点的函数值之差去顶替导数。一阶导数用前后两点的差除以间距，二阶导数用左右两点之和减去中间点的两倍，再除以间距的平方。热传导方程就这样变成了一张巨大的加减法表：下一时刻每个点的温度，等于这一时刻它自己和左右邻居温度的某种加权组合。这像菜市场里估一个摊位的行情——不去追问每一笔成交价，只看它和左右两家比贵了还是便宜了。

第二种是有限元。它不在点上切，而在块上切。把一座桥、一个机翼、一颗心脏的形状切成许多小三角形或小四面体，在每一小块里用最简单的函数（比如一次函数）近似真实的解，再按照“整体能量最小”或“在每个检验函数上平均成立”的原则把各块拼起来。这个想法可以追溯到柯朗 1943 年的一篇论文；1950 年代飞机工业里特纳、克拉夫等人把它做成了结构分析的工具，“有限元”这个名字据说是克拉夫在 1960 年前后起的。有限元的长处在于能贴合复杂的边界：桥墩是斜的、骨头是弯的，三角形都能铺上去。它像给一件不规则的家具做布罩，不量每一处曲率，只把布裁成许多小块，一块块缝起来。

第三种是龙格—库塔方法，专门处理随时间演化的常微分方程。最朴素的办法是欧拉法：站在此刻，看此刻的斜率，沿着这个斜率往前走一步。问题是斜率在这一步之内会变，欧拉法等于假定“这一步里一切照旧”，越走越偏。龙格（1895 年）和库塔（1901 年）的办法是在一步之内多看几眼：先看起点的斜率，往前探半步再看一次，换个方向再探一次，最后看终点附近的斜率，把这几次看到的斜率按权重平均，再迈这一步。最常用的经典四阶格式，每一步看四眼。好比开车下山，不是盯着车头前一米决定整段山路的方向盘，而是在这一段里扫几眼前方的弯道，取个综合判断。

三种切法各有偏爱：有限差分偏爱规整的网格，有限元偏爱复杂的形状，龙格—库塔偏爱时间方向上的精度。但它们共享一个动作——把无限细的东西换成有限粗的东西，并且必须为这个“粗”付账。

◆ 三、被切掉的东西有名字

数值分析这门学问最诚实的地方，是它给被切掉的东西起了名字，而且算得出大小。主要有两笔。

第一笔叫截断误差。用差商顶替导数，等于把泰勒展开式在某一项之后截断了，后面那些项就是被扔掉的部分。步长越小，扔掉的越少。欧拉法每一步扔掉的量大约与步长的平方成正比，累积到底，总误差与步长本身成正比——步长减半，误差减半。经典四阶龙格—库塔每一步扔掉的量与步长的五次方成正比，总误差与步长的四次方成正比——步长减半，误差降到原来的十六分之一。这就是“阶”的含义。

第二笔叫舍入误差。计算机里的数不是实数，而是有限位的二进制小数。常用的双精度浮点数，有效数字大约 16 位；它能分辨的相对差别有一个下限，称为机器精度，约为 2^{-52}，也就是一亿亿分之二左右。超过这个精度的差异，计算机根本看不见。十进制的 0.1 在二进制里是无限循环小数，存进计算机就被截掉了尾巴，所以在许多编程语言里，0.1 加 0.2 不等于 0.3。

两笔误差的脾气相反。步长越小，截断误差越小；可步长越小，计算步数越多，并且差商的分子是两个几乎相等的数相减，有效数字被大量抵消，舍入误差反而越大。于是误差随步长变化的曲线是一个 U 形：太粗不行，太细也不行，中间有一个最好的步长。这一点颠覆了一个朴素的直觉——“切得越细越接近真理”。在计算机里，切到某个程度以后，再细就是在噪声里打转。

> **两笔误差与 U 形曲线** 用前向差商近似导数：$f'(x) \approx \dfrac{f(x+h)-f(x)}{h}$。由泰勒展开 $f(x+h)=f(x)+hf'(x)+\frac{h^2}{2}f''(\xi)$，截断误差为 $\frac{h}{2}|f''(\xi)|$，与 h 同阶。计算机存储 f 的值时有相对误差约 ε（双精度 $\varepsilon \approx 2.2\times10^{-16}$）。两个值相减再除以 h，舍入误差约为 $\dfrac{2\varepsilon|f|}{h}$，$h$ 越小越大。总误差约 $E(h)\approx \dfrac{h}{2}|f''|+\dfrac{2\varepsilon|f|}{h}$。对 h 求最小，得最优步长 $h^{*}\sim\sqrt{\varepsilon}\approx10^{-8}$（设 f 与 f'' 同量级），此时最好的误差也只有 10^{-8} 左右——16 位有效数字只剩下一半。对

> 常微分方程 $y' = f(t, y)$，欧拉法全局误差为 $O(h)$；经典四阶龙格—库塔格式每步局部误差 $O(h^5)$，全局误差 $O(h^4)$。

这两笔误差不是书斋里的问题。1991 年海湾战争期间，美军在沙特达兰的一套“爱国者”导弹防御系统连续运行了大约 100 个小时。系统内部以十分之一秒为单位计时，而 0.1 存入 24 位寄存器时被截掉了尾巴，每一次计时都少了一点点。100 个小时累积下来，时钟偏差约三分之一秒；对一枚每秒飞行一千多米的飞毛腿导弹来说，这意味着跟踪窗口偏出了半公里以上。系统没能拦截来袭导弹，28 名士兵死亡。一个在单次计算里小到可以忽略的舍入，在 100 个小时的积分里长成了致命的偏差（事后美国审计总署 1992 年的调查报告 GAO/IMTEC-92-26 对此有详细说明）。

这正是全书主命题在数值层面的样子：被切掉的那一份，在单步里看不见，在累加里会长大；积分找不回它，只能在落地时被现实撞出来。

◆ 四、切法本身会造反：稳定性

截断与舍入还算是“少了一点”。数值计算里还有一种更凶的失败：方程本身没问题，误差也每步都很小，算着算着结果却炸了，数值在几十步之内从 30 度跳到 10^{20} 度。这叫数值不稳定。

拿开车做比方。一位新手在高速上发现车往左偏了一点，猛打一把方向往右；车冲到右边，他又猛打回左。每次纠正的方向都对，可是纠正的幅度大于偏差本身，于是越纠越偏，最后冲出路面。不稳定的数值格式就是这位新手：每一步的局部判断都“对”，但步子跨得太大，误差被一步步放大而不是消化。

1928 年，柯朗、弗里德里希斯和莱维在一篇讨论偏微分方程差分格式的论文里，指出了一个条件，后来以三人姓氏首字母称为 CFL 条件。它的道理朴素得惊人：信息在真实的方程里以某个速度传播，数值格式在一个时间步内只能让信息走过有限几个网格；如果一个时间步里真实信息已经跑出了数值格式能“看见”的范围，格式就在用它根本不知道的东西作判断，于是必然出错。

CFL 条件与拉克斯等价定理 对一维平流方程 $u_t + c\,u_x = 0$（$c > 0$），用迎风格式 $u_j^{n+1} = u_j^n - \lambda\,(u_j^n - u_{j-1}^n)$，其中 $\lambda = c\,\Delta t/\Delta x$ 称为柯朗数。可以证明：该格式稳定当且仅当 $0 \le \lambda \le 1$，即一个时间步内信息走过的距离 $c\,\Delta t$ 不超过一个网格 Δx。 对热传导方程 $u_t = k\,u_{xx}$ 的显式格式，稳定条件为 $k\,\Delta t/\Delta x^2 \le \frac{1}{2}$：空间网格加密一倍，时间步必须缩小到四分之一。 拉克斯等价定理（拉克斯与里希特迈尔，1956）：对适定的线性初值问题，一个相容的差分格式收敛，当且仅当它稳定。相容管“每一步切得对”，稳定管“误差不被放大”，两者合起来才保证切得越细越接近方程的真解。

这个定理值得停下来想一想。它说：仅仅每一步都近似得好（相容）是不够的，还得保证误差在传递中不被放大（稳定）。这和第13章所说的“焊接的条件是差异已稳定”是同一个结构在数值层的回声。微分与积分能互逆，前提是函数足够规矩；离散与连续能互通，前提是格式足够稳。没有这个前提，切得再细也接不回去。

CFL 条件还告诉我们一件更深的事：第二刀怎么切，不是计算者可以随意定的，它受到被切对象本身的约束。信息传播多快，是方程里的东西，最终追溯到现实里的东西（声速、热扩散率）。数值方法自以为在理念界内部自由地切，可是现实的速度从第一刀一路传到了第二刀，规定了第二刀的下刀分寸。这是理念物理界的一条回流通道：数值格式炸掉的时候，常常是在告诉你，你忘了现实里有多快。

◆ 五、数值解的 C 从哪里借

全书反复讲积分常数 C 的外借：微分丢掉的那一份，积分找不回来，只能在落地时从现实借回来。数值计算让这件事变得格外具体。

任何一次数值模拟，都要输入初始条件、边界条件、材料参数、网格。初值来自测量，边界来自对现场的判断，参数来自实验或手册。程序本身只负责“积分”；凡是该从现实借的，一样也少不了。而且数值计算还多出一笔新的外借：

网格与步长。真实世界没有网格，网格是人铺上去的。铺在哪里密一点、哪里疏一点，靠的是工程师对现场的理解——哪里应力集中、哪里有尖角、哪里材料突变。这份理解不在方程里，在人对这一个具体对象的经验里。

1991 年 8 月，挪威北海的斯莱普内尔 A 号海上平台，在峡湾里做压载下沉试验时混凝土基座破裂，整座结构沉入海底，损失约 7 亿美元。事后调查发现，设计中一处关键构件的有限元分析把剪应力低估了约 47%，原因之一是那个部位的单元形状划分不当，另一原因是同一部位的钢筋锚固不足。方程没有错，程序也照规则运行了；数值这一侧的错，出在第二刀切的位置，恰好切掉了那一处应力最集中的东西。

工程界后来形成的一整套“验证与确认”规程，可以看作对两刀的分别检查。“验证”问的是：程序是否正确地解了方程？办法是加密网格看结果是否收敛、用已知精确解的问题来考程序——这是检查第二刀。“确认”问的是：方程是否正确地代表了现实？办法是拿模拟结果去和实验、现场数据对照——这是检查第一刀，也就是把 C 从现实借回来对账。两者不能互相代替：网格再收敛，也不能证明方程选对了；实验对上了，也可能是两个错误恰好抵消。

天气预报是另一个看得见两刀的地方。气象中心的模式把大气切成水平几公里到十几公里的网格，比一朵积雨云还粗。云里的上升气流、雨滴的凝结，都发生在网格看不见的尺度上，只能用一组经验公式“参数化”塞回每个格子里——这是第二刀切掉的东西被人用经验补回来。而每天的预报都要先把全球观测站、卫星、探空气球的数据同化进模式，给出当天的初值——这是从现实借 C，一天借好几次。预报员常说“模式是工具，不是答案”，说的就是这两笔外借永远不能省。

这就引向一个对后文很重要的判断。计算机极大地扩展了积分的能力：过去解不出的方程现在能算，过去算不完的步数现在一秒钟就走完。但它扩展的是圈的理念一侧。在第二刀的下刀位置上、在初值与边界的选取上、在最后把数字交给工地或病床的那一刻，现实仍然要被重新借进来。数值方法没有消除再微分，

它只是把再微分的位置挪到了更多的地方：每一次铺网格、每一次选步长，都是一次小小的、需要对现实负责的落点。

◆ 六、计算机里的数，住在哪一格

最后用第4章的信息九宫来定个位。现实界的钢梁有它的温度，那是现实信息；热传导方程是理念界的数学态；计算机里的浮点数组又是什么？它仍在理念界，仍是数学态，但它是数学态中一个被再次筛过的子集：一个有限的、离散的、带有舍入规则的数的集合。实数轴上的绝大多数点，在这个集合里没有位置。

这意味着“可计算”这个谓词在数值层面有了更细的含义。一个问题在实数的意义上有解，不等于在浮点数的意义上能被稳定地算出；一个格式在理论上收敛，不等于在有限精度下能走到收敛。可计算性不仅是理念界的性质，而且是理念界里某一种特定的、有限的数学态的性质。每多一次截取，可计算的地盘就多一层条件。

数值分析自身的成长，也可以用 123 原理来看一段，但要看清它只适用到哪里。1940—1950 年代，电子计算机刚刚可用，人们满怀信心地把差分格式搬上机器（这是新的差异推进，D），而当时的数学习惯是“方程对了，算法自然对”（这是已沉淀的土壤，E）。两者的矛盾以一种非常具体的方式累积：格式莫名其妙地发散，同一个方程换个步长就给出完全不同的答案。这些反常逼出了新的显露（S）——稳定性分析、冯·诺伊曼的傅里叶稳定性判据、拉克斯等价定理，数值分析由此成了一门有自己定理的学科。这一改变又回写了土壤：此后任何一位工程师拿到一个新格式，第一件事不是算，而是先问“它稳定吗、几阶收敛”；大学的课程表、软件的默认设置、期刊的审稿标准都随之改了。这一段回写是有痕迹可查的，不是事后宣布的。至于这段历史能否被别的解释同样好地说明（比如单纯的工业需求推动），本书不作排他性主张，只指出这里确实走完了一圈可观察的循环。

这并不是贬低数值方法。恰恰相反，数值分析是整个微积分家族里最懂得自我审计的一支：它不假装切掉的东西不存在，而是把它命名、估算、设限，写进收敛阶、稳定条件和误差界。它示范了一种对待截取的成熟态度——承认切，算出切掉多少，规定在什么条件下还能接回去。

可是有一类对象，连这样诚实的切法都难以对付。数值方法再粗，它预设的仍是一条足够光滑的曲线：放大到足够细，曲线总会变得像直线，差商总会趋于一个确定的值。如果有一条路，放大一千倍还是锯齿，放大一百万倍仍是锯齿，处处没有切线，那么第一刀本身就切不下去。下一章要处理的，正是这样一条路。

带走的话

数值计算是第二刀：它切的不是现实，而是已经切过一次的理念方程，所以数值结果离现实隔着两刀。数值分析的长处是给切掉的东西起名、算账、设限——截断误差、舍入误差、稳定条件。但它扩展的只是圈的理念一侧：初值、边界、参数要从现实借，网格与步长也要凭对这一个对象的理解来铺。计算机没有取消再微分，只是把落点分散到了每一次铺网格、选步长上。

小账

- 本章开头的住院量体温，以及老汤菜谱、开车纠偏等，均为说明用的构想场景或比方，不对应真实个案。
- 本章所引“爱国者”导弹拦截失败与斯莱普内尔 A 号平台沉没两例，据事故调查报告及其通行转述，核验见附录I第I63、I64条。
- 本章所用截断误差、舍入误差、CFL 条件与拉克斯等价定理均为数值分析的标准结论，方框中只给最简情形。
- 本章用 123 原理读数值分析的成长，只指出走完了一圈可观察的循环，不排斥工业需求等其他解释。
- 本章主张的推翻条件见附录G第G30条。

19

第 19 章

随机：拥抱不可微

> 这一章的问题：咖啡里一颗脂肪小球下一秒往哪儿走，谁也说不出，整杯奶怎样散开却算得出——一条处处没有切线的路，微积分还能怎样接住它？

早上冲一杯速溶咖啡，往里倒一勺牛奶，先别搅。白色的奶在褐色的咖啡里慢慢洇开，边缘毛茸茸的，像一团正在长大的云。等五分钟再看，整杯都变成了浅棕色。

你说得出这团奶是怎么散开的吗？说得出大概：奶从浓的地方往淡的地方走，越来越均匀，五分钟后基本混好。这个“大概”可以写成方程，算出几分钟后哪里有多浓。可是如果问：此刻杯子正中那一颗脂肪小球，下一秒往哪边走？谁也说不出。它被成千上万个水分子从四面八方乱撞，往左一点，往右一点，毫无章法。

奇妙的是，两个问题的答案都对，而且彼此不矛盾：每一颗小球的路是乱的，所有小球合起来的扩散却是规矩的。一个人说不清自己下一秒往哪儿走，一座城市早高峰的人流却年年相似。本章要问的是：微积分怎样处理这种“单个不可说、整体可以说”的东西？更尖锐地问：当一条路径处处没有切线，第一刀切不下去的时候，微分还剩下什么？

◆ 一、一条处处没有切线的路

1827 年，植物学家罗伯特·布朗在显微镜下看到，悬浮在水里的花粉微粒（准确说是花粉里释放出的更小颗粒）在不停地乱颤。他起初怀疑是某种生命力，后来发现无机物的微粒也一样颤，于是这件事成了一个悬了几十年的反常。

1905 年，爱因斯坦给出了解释：微粒被看不见的水分子不断碰撞，每一次碰撞的方向是随机的，累积起来就是乱颤。他还算出了一个可以检验的结果：微粒偏离原位的平均平方距离与时间成正比。斯莫卢霍夫斯基在 1906 年独立得到了类似的结论。几年后，佩兰用精密的实验验证了这一关系，并据此测出了阿伏伽德罗常数，这在当时成了原子真实存在的有力证据之一。一个显微镜下的小反常，回流到了物理学最根本的问题上。

还有一段常被忽略的前史。1900 年，法国人路易·巴舍利耶在博士论文《投机理论》里，为巴黎交易所的债券价格写下了一种随机游走模型，其中已经包含了后来布朗运动的许多数学性质，比爱因斯坦还早五年。论文在当时几乎无人问津，直到半个世纪后才被经济学家重新发现。同一套数学，先在交易所里、后在显微镜下各自发生了一次，彼此并不知道。这提醒我们，跨领域的相似可以帮助提问，但两处现象是否真由同一种机制驱动，要各自取证。

但物理学家的解释还没有回答数学家的问题：这条乱颤的路，作为一条曲线，究竟是什么样子？1923 年，诺伯特·维纳给出了严格的数学构造，后人称为维纳过程，也就是数学上的布朗运动。它有三条基本性质：从零出发；任意两段不重叠的时间里的增量彼此独立；一段长为 t 的时间里的增量服从均值为零、方差为 t 的正态分布。由这些性质可以证明，几乎每一条这样的路径都是连续的，却处处不可微（佩利、维纳与齐格蒙德在 1933 年给出了这一结论的证明）。

“连续而处处不可微”在十九世纪曾被当作数学怪物。魏尔斯特拉斯 1872 年前后公开的那条函数让许多数学家不舒服，觉得它是人工造出来的病态例子。布朗运动让这个怪物从病理标本变成了自然界最常见的东西之一：花粉的路、扩散的分子、股价的走势，都长得像它。

用一句话说它为什么没有切线：在一小段时间 Δt 里，路径的增量大约有 $\sqrt{\Delta t}$ 那么大，而不是 Δt 那么大。于是差商 $\Delta W/\Delta t$ 大约是 $1/\sqrt{\Delta t}$，时间切得越细，差商越大，趋向无穷，没有极限。放大一千倍，它还是锯齿；放大一百万倍，还是锯齿。第10章所说的“截一刀，看局部像什么直线”，在这里切不下去：局部根本不像任何直线。

拿体重秤作个比方。每天早上称体重，读数总在上下跳，今天多四两，明天少三两。如果你拿相邻两天的差去除以一天，算“每天的增重速度”，会得到一串乱跳的数；如果改成每小时称一次，再算“每小时的增重速度”，跳得更凶——喝口水就是半斤。时间切得越细，“速度”越没意义。这不是秤坏了，而是体重在短时间里本来就主要由涨落构成，平稳的趋势要拉长时间才看得出来。

◆ 二、焊缝怎样重焊：伊藤积分

第13章把微积分基本定理比作一道焊缝：微分与积分能够互逆，条件是“差异已稳定”——函数足够光滑，局部足够像直线。布朗运动恰恰拆掉了这个条件。那么积分还能不能做？

普通的积分（黎曼—斯蒂尔切斯积分）对付 $\int f\,dW$ 这种东西要求 W 的“总变差”有限，也就是把路径上所有上上下下的距离加起来是个有限的数。布朗路径不满足：它在任何一小段时间里来回走过的总路程都是无穷大。于是旧焊条用不上了。

1944 年，伊藤清给出了一种新的焊法。他的办法有两个要点。第一，求和时，每一小段上的被积函数只取这一小段**开头**的值，而不是中间或末尾的值。这听起来是技术细节，实际上是一个关于时间的约定：此刻的决定只能依据此刻之前已经知道的东西，不能偷看下一刻的随机冲击。做股票交易的人最懂这一点：你今天按今天的价格决定买多少，明天涨跌与你今天的决定无关。第二，极限不再是逐条路径取的，而是在“均方”的意义上取的：不要求每一条路径上的和都收敛，只要求误差的平方的平均值趋于零。

这两个要点合起来，等于把焊缝从“逐点焊死”改成了“按统计焊牢”。焊口不再要求每一点都严丝合缝，只要求平均起来的偏差趋于零。旧的焊条要求差异稳定；新的焊条接受差异不稳，但要求不稳的方式本身是稳定的——增量独立、方差与时间成正比。可以说，伊藤把稳定性从路径挪到了分布上。

> **为什么 $\int_0^t W\,dW$ 不等于 $\frac{1}{2}W_t^2$**　若 W 是光滑函数且 $W_0=0$，则 $\int_0^t W\,dW=\frac{1}{2}W_t^2$。对布朗运动，把 $[0,t]$ 分成 n 段，取左端点：$\sum W_{t_i}(W_{t_{i+1}}-W_{t_i})=\frac{1}{2}W_t^2-\frac{1}{2}\sum(W_{t_{i+1}}-W_{t_i})^2$。最后一项是增量平方和。每个增量平方的期望是 Δt，加起来期望为 t；其方差随分割加细趋于零，所以这一项在均方意义下收敛到 t（称为二次变差 $[W]_t=t$）。于是 $\int_0^t W\,dW=\frac{1}{2}W_t^2-\frac{1}{2}t$。多出来的 $-\frac{1}{2}t$ 不是误差，是锯齿本身的累积：对光滑曲线，增量平方和随分割加细趋于零；对布朗路径，它趋于时间 t。若改取每段中点（斯特拉托诺维奇积分），结果恢复为 $\frac{1}{2}W_t^2$，但代价是失去“不偷看未来”的性质。两种积分都是正确的数学，选哪一种取决于被建模的对象。

这个方框里的最后一句值得强调。伊藤积分与斯特拉托诺维奇积分是两种不同的焊法，数学上都成立，换算公式也清楚。物理学家建模有色噪声的极限时常用后者，金融与生物里要求“决定只依赖过去”时常用前者。选哪一种不是数学自己能决定的，要回到被建模的现实里看：噪声是怎样进入系统的？这又是一处 C 的外借——这一次借的不是一个数，而是一种约定。

◆ 三、伊藤引理：二阶项不能扔

普通微积分里有一条链式法则：若 $y=f(x)$，则 $dy=f'(x)\,dx$。泰勒展开里的二阶项 $\frac{1}{2}f''(x)(dx)^2$ 被扔掉，理由是 $(dx)^2$ 比 dx 小得多。第10章讲泰勒展开时说过，截在一阶就是“局部像直线”的数学表述。

布朗运动拆掉了这个理由。因为增量 dW 的大小约是 $\sqrt{dt}$，它的平方约是 dt，和一阶项同样大小，扔不得。于是链式法则多出一项。这就是伊藤引理。

伊藤引理与布莱克—舒尔斯方程 设 $dX_t = \mu\,dt + \sigma\,dW_t$，$f(t,x)$ 足够光滑。则 $df(t,X_t) = \left(\partial_t f + \mu\,\partial_x f + \tfrac{1}{2}\sigma^2\partial_{xx} f\right)dt + \sigma\,\partial_x f\,dW_t$。 记忆口诀：$(dW)^2 = dt$，$dt\cdot dW = 0$，$(dt)^2 = 0$。 例：股价模型 $dS_t = \mu S_t\,dt + \sigma S_t\,dW_t$（几何布朗运动）。取 $f = \ln S$，得 $d\ln S_t = (\mu - \tfrac{1}{2}\sigma^2)dt + \sigma\,dW_t$，故 $S_t = S_0\exp\left((\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t\right)$。漂移里多出的 $-\tfrac{1}{2}\sigma^2$ 正是二阶项的贡献：波动本身会压低典型路径的增长。 布莱克—舒尔斯（1973）：在无摩擦、可连续交易、无风险利率 r 与波动率 σ 为常数等假设下，期权价格 $V(t,S)$ 满足 $\partial_t V + \tfrac{1}{2}\sigma^2 S^2\partial_{SS} V + rS\,\partial_S V - rV = 0$。 方程里不出现 μ：通过连续调整股票持仓（“对冲”），随机项 dW 被抵消，期权价格不再依赖股票预期涨多少。

漂移里那个 $-\frac{1}{2}\sigma^2$ 有一个大白话的版本。假设你的投资第一年涨 50%，第二年跌 50%，平均“收益率”是零，可你的钱从 100 变成 150，又变成 75，亏了四分之一。波动越大，这种“平均不赚、实际亏损”越严重。二阶项讲的就是这件事：在涨落占主导的世界里，波动本身是一种成本，不是可以平均掉的噪音。

伊藤引理是底本所说“多出来的那一项，正是随机涨落在二阶上的顽强显影”的严格形式。从全书主命题看，它说明了一件事：在经典微积分里，截在一阶被允许，是因为被扔掉的二阶纠缠真的小；当现实里的纠缠以涨落的形式大到和一阶同级时，第一刀就不能那样切了，必须把二阶留下来。不是微分失效了，而是“切掉多少”的规矩随对象改变了。

◆ 四、布莱克—舒尔斯：一个焊得太牢的公式

1973 年，费希尔·布莱克与迈伦·舒尔斯发表了期权定价公式，罗伯特·默顿同年给出了更一般的推导。公式的核心想法很漂亮：期权的风险可以通过不断买卖股票来对冲掉，对冲之后剩下的组合是无风险的，它只能赚无风险利率，否则就有人能白捡钱。从这个“不能白捡钱”的论证出发，期权价格就被确定了。1997 年，舒尔斯与默顿获得诺贝尔经济学奖，布莱克已于 1995 年去世。

公式出来后的几年里，芝加哥期权交易所的成交量迅速增长，交易员手持计算器按公式报价。社会学家麦肯齐与米洛 2003 年对芝加哥期权交易所的历史研究指出，市场价格在一段时间里越来越贴近公式的预测，部分原因是交易员都在用这个公式——模型在一定程度上塑造了它所描述的市场。这是一种可以观察的回写：一个理念界的积分结果，落回现实，改变了现实里的价格形成方式，进而改变了下一轮数据。

但回写是中性的，也有负向的。1987 年 10 月的股灾之后，期权市场出现了所谓“波动率微笑”：按公式反推出来的波动率不再是常数，而随行权价弯成一条曲线——市场开始给极端下跌定一个比正态分布高得多的价。这是现实顶回来的反常：股价的涨落并不是正态的，大跌比模型认为的更常见，而且在恐慌中，涨落之间不再独立。1998 年，以默顿、舒尔斯为合伙人的长期资本管理公司在俄罗斯债务违约引发的动荡中巨亏，最终由纽约联储出面协调多家银行注资接管。需要说清：它的亏损主要来自高杠杆下的收敛套利（如利率互换价差）与做空股票波动率等头寸，并非直接源于布莱克—舒尔斯公式本身；这里只把它引作“平静时期借来的模型假设在极端情境下一齐失效”的例子，不能读成“公式导致倒闭”。

“波动率微笑”这个反常没有让行业抛弃随机微积分，而是逼出了新一轮模型：1990 年代出现的局部波动率模型（杜皮尔等）与随机波动率模型（赫斯顿等），前者让波动率随价格与时间变化，后者让波动率本身也随机变动；后来又有带跳跃的模型，专门给“突然的大跌”留位置。交易台的做法也随之改变：公式不再被当作给出“正确价格”的机器，而是当作一种报价的语言——交易员报的是“隐含波动率”，用公式把它换算成价格。这是一次可观察的回写：反常没有推翻工具，却改变了人们使用工具的方式。

用本书的语言说：布莱克—舒尔斯在理念界积得很漂亮，而它所需的那些 C——波动率是常数、交易可以连续进行、市场随时有对手方——都是在平静时期从现实借来的。到了危机时刻，现实把这些借款一次性收回。最先断掉的，正是“可以连续对冲”这一条：当所有人都要同时卖出，流动性消失，对冲的那只手伸不出去。

这说明随机微积分并没有把不确定性“消灭”掉。它把不确定性的一种特定形状——独立、正态、方差稳定——收编进了数学；而现实中不属于这种形状的那部分不确定性，依然留在筛子外面，等着在某个具体的日子里回流。

◆ 五、“统计意义上可重复”是什么意思

底本说，随机微积分通过“概率性特征律”建立稳定性：不是每一次都相同，而是统计意义上可重复。这句话需要放到现行口径里重新说清楚。

第2章讲 S 的“三个可”：可识别、可维持、可表达。对一条布朗路径来说，具体哪一条显露出来，不可维持——下一次实验就是另一条。可维持的是更高一层的东西：增量的分布、方差与时间的比例、扩散系数。一杯咖啡里奶的扩散，每一次的分子路径都不同，可五分钟后的浓度分布每一次都差不多。被稳下来的不是路径，是路径的统计特征。

数学上，这两层之间有一座精确的桥。单个微粒的路径满足一个随机微分方程；而大量微粒的概率密度，满足一个确定性的偏微分方程，称为福克—普朗克方程（在数学里也叫科尔莫戈罗夫前向方程）。对最简单的布朗运动，它就是第17章讲过的热传导方程。开头那杯咖啡的两个问题，正对应这座桥的两端：问一颗脂肪球下一秒去哪，是在桥的随机一端提问，答不出；问五分钟后哪里多浓，是在桥的确定一端提问，答得出。随机微积分的贡献不只是让随机的一端有了微积分，更是说清了从一端走到另一端时，什么被保留（分布的演化），什么被丢掉（每一条具体的路）。

用开车作比：交通部门能相当准确地预报明早七点半某条环路的平均车速，却说不出你那辆车会在哪个路口被加塞。平均车速是分布那一端的事，你被加塞是路径那一端的事。把两者混为一谈，要么会责怪预报“不准”，要么会以为自己能逃过统计规律。

这对“科学的可重复性”有一个直接的推论。第12章用复现危机作检验：越靠“其他条件不变”切得狠的学科，复现危机越重。随机微积分提供了另一种看法：有些研究对象的显露，本来就只在分布层面可重复，在个体层面不可重复。

一项心理学实验，如果它的效应只是总体分布上的微小偏移，那么任何一次具体的实验都可能看不见它；正确的复现不是“再做一次得到同样的数”，而是“做足够多次，分布落在预先说好的区间里”。把分布层面的规律当成个体层面的规律去复现，失败几乎是必然的。

但这个看法不能被反过来滥用。“统计意义上可重复”是一个可检验的主张：要预先说明是哪个统计量、在多大样本下、落在什么区间。如果一个解释在每次失败后都说“这本来就是随机的”，它就什么也没说。概率的语言是对证据责任的加码，不是豁免。

还有一层区分要守住。物理学家相信布朗运动的随机性来自我们无法追踪的大量分子碰撞，即“认识上的随机”；量子力学里的随机性则被许多物理学家认为是更根本的。随机微积分作为数学工具对这两种来源一视同仁，它只关心涨落的统计形状。用它建模，不等于对世界是否“本来随机”作了表态。

◆ 六、蒙特卡罗：用骰子做积分

随机微积分还反过来给了计算一件新工具。前一章讲数值方法时，积分是在网格上一格一格加起来的。维数一高，网格就爆炸：10 维空间里每一维取 10 个点，就是 10^{10} 个、也就是一百亿个格点。1940 年代，乌拉姆、冯·诺伊曼等人在洛斯阿拉莫斯想出了另一个办法：不铺网格，而是随机撒点，用撒中的点的平均值去估计积分。这个办法以摩纳哥的赌城命名，叫蒙特卡罗方法。它的误差大约随撒点数的平方根倒数下降，与维数基本无关——这正是布朗运动“增量约为 $\sqrt{t}$”的那个平方根。

期权定价、核反应堆中子输运、蛋白质构象采样、今天大模型训练中的随机梯度，都在用这一思路。生活里的类比是民意调查：不必问遍全国每一个人，随机问 1000 人，误差就能控制在几个百分点。前提是“随机”真的随机——如果只在一个小区里问，样本再大也没用。蒙特卡罗方法同样要向现实借一样东西：抽样的方式要与被估计的对象相称。

◆ 七、小结：把不可微收编进数学的代价

随机微积分的贡献，是把一类原先被当作“噪声”扔掉的东西收进了数学：处处不可微的路径有了积分，有了链式法则，有了方程。它改造了焊缝——从逐点焊死改为按统计焊牢；它改造了截取的规矩——二阶项不能扔。这些都是真正的数学发明，而不是换了说法。

代价同样清楚。它收编的只是不确定性里形状规矩的那一部分：独立的增量、稳定的方差、短促而健忘的冲击。布朗运动的每一步都不记得上一步。可是现实里有许多东西是记得的：一块橡胶被压过之后回弹得慢，一个人受过伤之后对同样的刺激反应不同，一条河的今年水位与几十年前的旱涝有关。这些系统的今天不只取决于刚才，还取决于很久以前。下一章要问的是：当差异带着记忆，当影响来自远处，微分这把只看“此处此刻”的刀，还能怎样改造。

带 走 的 话

布朗路径处处没有切线，第一刀切不下去；伊藤积分没有放弃微积分，而是把稳定性从路径挪到了分布上，并让二阶项留了下来。但随机微积分收编的只是一种形状规矩的不确定性：增量独立、近于正态、方差稳定。不属于这种形状的那部分——大跌扎堆、恐慌中涨落彼此牵连——仍留在筛子外面，会在某个具体的日子回流。

小 账

- 本章的咖啡加奶、体重秤、早高峰车流、民意调查，均为说明用的日常比方，不对应真实个案。
- 本章所述布朗、爱因斯坦、巴舍利耶、维纳、伊藤清、布莱克—舒尔斯与默顿、长期资本管理公司等史实取通行叙述；“模型塑造市场”一说据麦肯齐与米洛 2003 年的研究，出处见附录I第I62条。
- 本章不主张布莱克—舒尔斯公式导致了长期资本管理公司的倒闭，只把它引作借来的假设在极端情境下一齐失效的例子；本章也不构成投资建议。
- 本章不主张世界“本来随机”或“本来确定”；随机微积分只关心涨落的统计形状。
- 本章主张的推翻条件见附录G第G31条。

20

第 20 章

记忆与非局域：分数阶与伞场微积分

> 这一章的问题：老沙发的坐垫记得父亲坐了十几年的位置，可导数只看此刻这一点——微积分怎样学会记住过去、够到远处，又能走多远？

家里的老沙发，坐垫上有一个凹坑。那是父亲坐了十几年的位置。你坐上去，起身，海绵慢慢地鼓回来，可鼓不到原样；你坐别的位置，起身，海绵很快就平了。同一块海绵，同样的一次坐压，反应却不一样——它记得过去。

再想想看病。医生问的从来不只是“现在哪里疼”，还要问“疼了多久”“以前有没有过”“小时候得过什么病”。同样是此刻三十八度的体温，一个三天前刚做过手术的人和一个一直健康的人，意味着完全不同的事。现在的状态，不足以决定接下来会怎样。

经典微积分有一个很少被说出口的假定：要知道一个系统下一刻怎样变，只需知道它此刻在哪里、此刻的变化率是多少。导数是一个点上的东西，它只看这一点无限小的邻域；微分方程把下一刻交给此刻。这个假定在行星运动里极为成功，可在老沙发、病历本、河流的旱涝记录面前，它切掉的东西太多了。本章讨论两种修补：一种在时间上拉长记忆，叫分数阶微积分；一种在空间上伸长手臂，叫非局域算子。本书作者王德生提出的伞场微积分属于后一类。我们会看到，它能被严格说出的部分，比底本当初设想的要小，也更结实。

◆ 一、半阶导数：一个三百年的问题

1695 年，洛必达在给莱布尼茨的信里问：既然有一阶导数、二阶导数，那么“二分之一阶”的导数是什么意思？莱布尼茨的回信说，这看起来是个悖论，但总有一天会从中得出有用的结论。这段通信常被当作分数阶微积分的起点；莱布尼茨那封回信，通常系于 1695 年 9 月 30 日，原话措辞各家转述不一，这里只取大意。

这个问题之所以难，是因为整数阶导数是“局部”的：一阶导数只看一点附近，二阶导数也只看一点附近。可是“半阶”该看多远？十九世纪，阿贝尔在处理等时降线问题时（1823 年）、刘维尔与黎曼在各自的工作里，逐步给出了答案。关键的线索来自积分。对一个函数反复积分 n 次，柯西公式能把它写成一次积分，被积函数里有一个因子 $(t-\tau)^{n-1}$；把整数 n 换成任意正数 α，用伽马函数替代阶乘，就得到了“分数次积分”。分数阶导数则可以定义为：先做分数次积分，再做整数次求导。

这样定义出来的东西有一个惊人的性质：它不再是局部的。要算一个函数在时刻 t 的半阶导数，必须知道这个函数从起点到 t 的全部历史，每一个过去时刻都以幂律衰减的权重参与进来。越远的过去，权重越小，但从不为零。

分数阶积分与导数　黎曼—刘维尔分数次积分（$\alpha>0$）：$I^{\alpha}f(t)=\dfrac{1}{\Gamma(\alpha)}\displaystyle\int_0^t (t-\tau)^{\alpha-1}f(\tau)\,d\tau$。当 $\alpha=n$ 为正整数时，恰为 n 重积分。黎曼—刘维尔分数阶导数（$n-1<\alpha<n$）：$D^{\alpha}f=\dfrac{d^n}{dt^n}I^{\,n-\alpha}f$。　卡普托导数：${}^{C}D^{\alpha}f=I^{\,n-\alpha}\dfrac{d^n f}{dt^n}$，即先求导、后做分数次积分。它的好处是初始条件仍可用 $f(0),f'(0),\dots$ 这类有物理意义的量给出；卡普托在 1967 年前后把它用于黏弹性与地球物理问题。　例：$D^{1/2}\,t=\dfrac{2}{\sqrt{\pi}}\,t^{1/2}$，再求一次半阶导数得 $D^{1/2}D^{1/2}t=1=\dfrac{d}{dt}t$。两次半阶连起来等于一阶。　注意：常数的黎曼—刘维尔半

> 阶导数不为零（$D^{1/2}1 = 1/\sqrt{\pi t}$），卡普托导数则为零。两种定义在同一函数上给出不同结果，选哪一种要看初始条件在被建模的对象里是什么意思。

方框最后一条又是一处“约定的外借”，与上一章伊藤与斯特拉托诺维奇的选择同一类型：数学给出几种都正确的定义，哪一种合用，要回到现实里看初值是什么。

◆ 二、记忆为什么是幂律

分数阶微积分在二十世纪后半叶从数学的角落走进工程，主要是因为一类材料和一类扩散。

先说材料。一块弹簧，你拉它，它立刻产生与伸长成正比的力，松手就回去，不记得任何事——这是胡克定律，力只看此刻的形变。一桶蜂蜜，你搅它，阻力与搅动的速度成正比，停手就不再有力——这是牛顿黏性，力只看此刻的变化率。橡胶、沥青、生物软组织、那块老沙发的海绵，介于两者之间，称为黏弹性材料：力既与形变有关，又与形变的历史有关。工程师早就用弹簧与阻尼器的各种组合去模拟它们，可组合越来越复杂，参数越来越多。

1930—1940 年代起，斯科特·布莱尔等人发现，许多材料的应力松弛曲线近似是幂律，而幂律恰好就是分数阶导数的“指纹”：力与形变的 α 阶导数成正比，$\alpha = 0$ 是弹簧，$\alpha = 1$ 是蜂蜜，中间的 α 给出介于两者之间的材料，只用两个参数就能拟合许多实验曲线。1983 年，巴格利与托维克在流变学期刊上从分子理论的角度论证了这类模型在某些聚合物中的合理性。

再说扩散。上一章讲布朗运动，平均平方位移与时间成正比。可是在细胞内部、多孔岩石、某些湍流中，测到的平均平方位移与时间的某个非整数次幂成正比，称为反常扩散。若粒子常常被困住很久才再动一下（等待时间的分布有很长的尾巴），扩散就会变慢，描述它的方程里时间导数变成了分数阶。

两个例子共享一个结构：系统的“今天”是过去所有时刻的加权和，权重按幂律衰减。生活里也到处是这种结构。一个人的信用记录，昨天的一次逾期影响

很大，五年前的一次影响变小，却一直还在；一座城市的交通，今天的拥堵与几十年前的规划有关。用指数衰减来描述记忆，等于说过去会在一个特征时间之后基本被忘掉；用幂律来描述，等于说没有这样一个特征时间，每一个尺度上都还有一点残留。

从全书主命题看，分数阶导数做的事情是：在第一刀落下时，不再只取此刻这一点，而是把这一点身后的整条历史按某种权重一起截进来。它少切掉了一部分纠缠——时间方向上的纠缠。代价是：要算此刻，必须带着全部过去，计算量随时间增长；而那个幂律的指数 α，几乎从不能从第一原理推出，只能从实验数据拟合，又是一笔从现实借来的 C。

工程里还有第三条入口：控制。第24章会讲，经典的 PID 控制器根据误差的现在（比例）、过去的总和（积分）、变化的趋势（微分）来调节；1999 年前后，波德卢布尼提出把其中的积分与微分换成分数阶，称为 $PI^{\lambda}D^{\mu}$ 控制器，多了两个可调的阶数。它在一些对象上调得更柔顺，代价是参数更难整定，实现时又要把分数阶算子离散化——第18章那第二刀在这里又落了一次。

分数阶的计算代价值得单独说一句，因为它与“记忆”这件事是同一枚硬币的两面。普通微分方程往前算一步，只需要上一步的值；分数阶方程往前算一步，要把从起点到此刻的所有值都翻出来加权求和。算一千步，每一步都要回看前面所有步，工作量大约是步数的平方。记账的人对此有体会：只记“今天余额”的账本，每天加减一笔就行；要按“每一笔旧账都按年头打折计入今天信用”的规矩记账，每天都得把整本旧账重翻一遍。工程上于是又发明了“短记忆”原则——只回看最近一段，远处的历史截掉。可见，分数阶少切掉的那一段过去，到了计算的时候常常又被切回去一部分，切多少，要看精度要求与算力预算。

还要说一句收窄的话。底本称分数阶微积分对应“复杂系统的多层纠缠”。这可以作为提问的线索，但不能当作结论：一条实验曲线能用分数阶模型拟合，不等于系统里真有某种“多层纠缠”机制；同样的数据往往也能用多个指数衰减

之和拟合。分数阶模型是否比别的模型更好，要看它在新数据上的预测，而不是看它的名字与 SDE 的词汇是否相近。

◆ 三、伸向远处的手：非局域算子

分数阶在时间上记住过去；另一类改造在空间上看向远处。

经典的扩散与弹性方程里，一点只与它无限小的邻域相互作用。拉普拉斯算子 $\nabla^2 u$ 衡量的是一点的值与它紧邻处平均值之差。可是有些现象里，一点直接受到有限距离之外的影响。一块布被撕开时，裂缝尖端附近的纤维同时受到一片区域的牵拉；一片森林里，种子被风和鸟带到几十米外；一个人的情绪不只受隔壁座位的人影响，也受整个办公室的气氛影响。

数学上，处理这类现象的办法是用一个积分代替导数：一点的变化，等于它与周围一片区域内所有点之差的加权和。权重由一个函数给出，称为核。核决定了“多远算邻居、邻居各占多重”。这样的算子称为非局域算子。

非局域差分算子与它的局域极限　非局域差分算子：$(\mathcal{L}u)(x)=\int K(x,y)\,[u(y)-u(x)]\,dy$。使用前须给定核 K、函数空间与积分存在的条件。离散形式为 $\sum_j w_{ij}(u_j-u_i)$。　一维局域极限的设定：$K_r(x,y)=r^{-1}\varphi\big((y-x)/r\big)$，其中 φ 非负、偶对称、可积、积分为 1、支撑在有界区间 $[-R,R]$ 内，二阶矩 $M_2=\int z^2\varphi(z)\,dz>0$；$u$ 在 x 附近二阶连续可导。命题：$\lim_{r\to 0} r^{-2}\,\mathcal{L}_r u(x)=\frac{M_2}{2}\,u''(x)$。　证明要点：令 $z=(y-x)/r$，得 $\mathcal{L}_r u(x)=\int\varphi(z)[u(x+rz)-u(x)]\,dz$。泰勒展开，一阶项因 φ 偶对称而积分为零，二阶项给出 $\frac{r^2}{2}u''(x)M_2$，余项为 $o(r^2)$。三点限制：①不乘 r^{-2}，$\mathcal{L}_r u$ 本身趋于零，尺度归一不能省；②多维时极限由二阶矩矩阵决定，只有各向同性时才是拉普拉斯算子的常数倍；③若核的支撑无界，任意二阶可导函数未必使积分存在，若只有更低的正则性，也未必得到 $O(r^2)$ 级的误差。

一个人人手机里都有的例子：照片降噪。最朴素的做法是让每个像素取它上下左右邻居的平均值，噪点被抹平了，可边缘也糊了——这相当于经典的局部扩

散。2005 年前后，布阿德斯、科尔与莫雷尔提出“非局部均值”法：一个像素不只向紧邻取平均，而是在整张图里寻找“周围纹理和它相像”的像素，按相像程度加权平均。砖墙上的一块砖可以向远处另一块砖借信息，边缘因此保住了。这里的核不是按距离、而是按相似度定的，说明“非局域”的“邻居”可以由任务重新定义。这又是一笔向现实借来的东西：什么算相像，要看这张照片拍的是什么。

这个命题说的是：在严格限定的条件下，当核的作用范围缩到零、并且按适当的比例放大时，一类非局域算子在一点上逼近二阶导数的常数倍。它是一座桥：从“看远处”的算子走回“只看邻近”的导数。这类结果在非局域扩散与近场动力学（西林 2000 年提出的一种不用导数描述裂纹的连续介质力学）的文献里早有系统研究，德利亚与贡兹伯格、杜强等人的工作讨论了非局域模型的适定性及其局域极限。

◆ 四、伞场微积分：能说的与不能说的

本书作者王德生把这类“以一点为中心、向四周张开、可以与他处重叠的影响范围”称为“伞”，并以此为出发点设想了一种伞场微积分。在第395卷《SDE思想体系入门》的现行口径里，伞被定义为“带尺度的核场”：用中心、尺度与权重核描述的、带影响范围的相互作用结构。上面方框里的算子与命题，正是第395卷第84章给出的形式化起点。

这里必须把能说的与不能说的分开写清楚，因为底本及更早的网站文字曾有过更强的说法。

能说的，只有一个局部命题：在限定核（非负、偶对称、归一、紧支撑、二阶矩为正）、尺度归一（乘以 r^{-2}）、正则条件（被作用的函数在该点附近二阶连续可导）之下，一类非局域差分算子 $(\mathcal{L}u)(x)=\int K(x,y)[u(y)-u(x)]\,dy$ 在点态意义上逼近特定的二阶微分算子。这是一个可以复核的数学命题，条件、结论、反例边界都写得出来。

不能说的，有以下几条，第395卷已明确撤回：

第一，“传统数学是伞模型的极限或特例”（或写作“当某个参数 $N \to 0$ 时伞模型退化为传统数学”）。这一说法已被作者在第395卷（第52、53、84章及撤回清单）撤回。理由很具体：上面那个命题只说明一类核算子在一点上与二阶导数的渐近关系，远远推不出全部传统数学——代数、数论、拓扑、离散数学——都从伞模型退化而来；即便在微分算子内部，它也只对应二阶、对称的那一类。另外，作者曾提出的指标 $N = r\omega/(vT)$（半径、重叠度、传播速度、特征时间之比）趋于零，并不推出作用范围趋于零：固定半径与重叠度，只让传播速度变大，N 就趋于零，可伞一点也没收拢。

第二，“相容性是数学的普遍公理”“传统数学禁止重叠，因此无法处理非局域”。这一说法同样被撤回。有限元的基函数本来就跨单元重叠，非局域算子早已可以用积分核、图上的权重表达；矩阵变稠只增加计算成本，不等于原则上无法表达。几何单元不重叠、函数空间相容、矩阵稀疏、物理作用局域，是四种不同的性质，不能混为一谈。

第三，“三态求解器已完成”“地震波已验证伞模型”。第395卷说明：伞模型尚未定义一个能区分秩序、混沌、介生三态的动力学方程，也没有完成地震数据的可证伪检验。把已知的地震波频散改写成伞的语言，只是解释性的映射；要成为证据，必须给出与已有波动理论不同、且能被数据否定的预测。

那么，伞场微积分还剩下什么值得做？第395卷把它重新定位为一个研究纲领，列出了五项形式化任务：带尺度核场的定义、算子性质、限定条件下的局域极限、三态判据、与已有算法的比较。其中只有局域极限一项在限定条件下有了证明，其余都是开放的。还有一个具体而重要的技术问题：连续模型里的伞缩小（$r \to 0$）与计算机里的网格加密（$h \to 0$）是两件不同的事。第18章讲过，数值格式有自己的误差；而乘上 r^{-2} 的尺度归一会把离散误差放大。若离散误差约为 $\varepsilon(h, r)$，总误差的上界大致是 $\varepsilon(h, r)/r^2$ 加上 r^2 量级的连续误差，只有当 $\varepsilon(h, r)/r^2$ 也趋于零时，这个上界才能保证收敛。先缩网格再缩伞，与两者一起缩，结果未必相同。这正是第18章所讲“两刀分别检查”在伞场上的具体形态。

用一个大白话的比喻收住这一节。一个新厨师说：“我发明了一种炖法，在火候特别小、时间特别短的极限下，它跟传统的焯水一模一样。”这句话如果他能在厨房里演示出来，就是一个有价值的发现，说明新炖法与旧做法之间有一座可以走通的桥。但他不能接着说：“所以所有中餐都是我这种炖法的特例。”炒、蒸、腌、生拌都没有经过那座桥。伞场微积分目前的处境，就是这位厨师刚刚演示完一座桥。

◆ 五、从 SDE 看记忆与非局域，以及这样看的边界

分数阶与非局域算子，都是对经典微分“只看此处此刻”这一刀的放宽：前者在时间上不切断过去，后者在空间上不切断远处。用全书的语言说，它们让第一刀少切掉一部分纠缠。

这里可以用第15章的三方程作一次提问，而不是作一次断言。若把“此刻显露的状态”看作 S，把“状态的推进规则”看作 D，把“过去的历史与远处的影响”看作 E，那么经典微分方程相当于假定 E 只以“初值”和“紧邻”两种最薄的方式参与；分数阶与非局域模型则让 E 以更厚的方式进入每一步的推进。这个对应只是建模启发：E 在 SDE 里指特征纠缠网络，不能仅凭名称就认定它与物理上的非局域作用是同一件事。要让这种对应有用，必须说明在一个具体问题里，哪个数学对象代表 E、它的经验含义是什么、加进来之后能否多出一个可检验的区分。如果不能，三方程在这里就只是在整理语言。

一个检验的例子：对一块黏弹性材料，经典模型与分数阶模型都能拟合同一组松弛实验。如果“历史以幂律方式参与”是对的，那么换一种加载历史（比如先压一下、停一会儿、再压），分数阶模型应当在不重新调参数的情况下给出更准的预测。这样的检验做得出来，也确实有人在做。哲学的解释可以指出往哪里看，可是答案要由这类实验给出。

◆ 六、收束

本章的两种改造，使微积分能够处理记得过去的材料和够得着远处的作用。它们的成功是真实的，它们的代价也是清楚的：要带着全部历史计算，要从数据里借来幂律指数和核的形状，要另外证明非局域模型在什么意义下收敛。

伞场微积分在这里的位置，经过第395卷的收窄之后，反而比底本更站得住：它不是要取代传统数学的一个总框架，而是一个带着一条已证明的桥、几项待证任务的研究纲领。一种思想敢于收回自己说过头的话，可以看作第16章所说的一次回写——反常与批评回到理论内部，改写了它自己的主张。这次回写不是宣布出来的，它留有记录：第395卷的撤回清单写明了收回哪些说法、为什么收回。

到这里，微积分已经从一个点走到了一段历史、一片区域。可是还有一种扩张更大：被微分和积分的对象，不再是一个数、一条曲线，而是“所有可能的曲线”构成的整个空间。下一章进入无限维。

带 走 的 话

分数阶导数在时间上不切断过去，非局域算子在空间上不切断远处，二者让第一刀少切掉一部分纠缠，可幂律指数与核的形状仍要从数据里借。分数阶拟合得好，不等于系统里真有某种“多层纠缠”。伞场微积分能严格说出的只有一条：在限定核、尺度归一、正则条件下，一类非局域差分算子在一点上逼近二阶导数的常数倍；“传统数学是伞模型的极限或特例”已由作者撤回。

小 账

- 本章的老沙发、看病、信用记录、记账、新厨师的炖法都是帮助理解的日常比喻，不对应真实个案；照片降噪所用的非局部均值法则是真实的方法。
- 本章没有使用任何调查数字；洛必达与莱布尼茨 1695 年的通信只取大意，核验见附录I第I36条，其余史实据通行文献，附录I未单列。
- 本章唯一的正面数学主张是方框中的一维点态局域极限命题，条件与证明要点已列出，可独立复核。
- 本章不主张伞场微积分是取代传统数学的总框架，也不主张“三态求解器已完成”或地震数据已验证伞模型。
- 本章主张的推翻条件见附录G第G32、G33条。

21

第 21 章

无限维：量子与泛函

> 这一章的问题：导航要在无数条回家的路里挑最快的一条，当微积分的对象从一个数变成“所有可能的路”，它还要向现实借什么？

周末开车去城郊看父母，手机导航跳出三条路线：走高速四十分钟，走国道五十分钟，走老城区三十八分钟但“路况复杂”。你选了高速。其实这三条只是导航挑出来给你看的；从你家门口到父母家门口，理论上能走的路有无数条——每一个路口左转还是右转，每一段走快一点还是慢一点，都是一条不同的路。导航在后台做的事情，是在这无数条路里找一条“总耗时最短”的。

这件事和前面几章有一个根本不同。前面无论是求导数还是解方程，被问的对象都是一个数或一条曲线：此刻速度多少、温度怎样随时间变。导航问的却是：在所有可能的路线里，哪一条最好？被比较的东西不再是一个数，而是一整条路线。每条路线都对应一个数（总耗时），于是有了一种“吃进一整条曲线、吐出一个数”的东西。数学上管它叫泛函。

一旦对象变成“所有可能的曲线”，微积分就被推进了无限维：一条曲线由无数个点上的值决定，所有曲线构成的空间，每一个“方向”都是一种改变曲线的方式，方向有无穷多个。本章看三次这样的扩张：变分法问“哪条路最好”，希尔伯特空间为量子力学提供了舞台，路径积分则干脆把“所有可能的路”都加起来。积分的对象，从“量”变成了“路”。

◆ 一、哪条路最快：变分法

1696　年，约翰·伯努利向欧洲数学界提出一个挑战：一颗小珠子在重力作用下，沿一条光滑的轨道从高处一点滑到低处另一点（不在同一竖直线上），轨道是什么形状时用时最短？直觉说是直线，因为直线最短。答案却是一段摆线——先陡后缓的一条弧。先陡，珠子很快获得速度；速度大了，后面多走一点路也划算。这个问题称为最速降线问题，牛顿、莱布尼茨、雅各布·伯努利、洛必达都给出了解答。

这像一个老司机的经验：去机场不一定走直线距离最近的那条小路，先绕上高速，虽然路长了，总时间反而短。最好的路线不是每一段都最好，而是整条路合起来最好。

最速降线开启了变分法。半个多世纪后，欧拉与拉格朗日把它系统化：要找使某个泛函取极值的曲线，就看曲线被轻轻“扰动”一下时泛函怎样变化；若曲线已是最优，那么任何微小扰动在一阶上都不改变泛函的值。由此可以推出曲线必须满足一个微分方程，称为欧拉—拉格朗日方程。

这里藏着一个极为巧妙的转换。原问题是全局的：在所有曲线里挑最好的一条，这是一个对整条路线的判断。欧拉—拉格朗日方程却是局部的：它是曲线上每一点都必须满足的一条微分方程。一个关于整体的“积分式”问题，被翻译成了一个处处成立的“微分式”条件。这正是第11章所讲斯托克斯式“边界决定内部”、第13章所讲焊接的又一个版本：整体与局部，在一定条件下可以互相改写。

物理学很快抓住了这个工具。莫佩尔蒂、欧拉、拉格朗日、哈密顿先后提出，力学运动可以表述为某个量——作用量——取极值（更准确地说，取驻值）。一块石头从扔出到落地，走的那条抛物线，是所有可能路线中使作用量取驻值的那一条。牛顿定律可以从这个原理推出来。

作用量原理与欧拉—拉格朗日方程　对路径 $x(t)$，$t\in[t_1,t_2]$，定义作用量 $S[x]=\int_{t_1}^{t_2}L(x,\dot{x},t)\,dt$，其中拉格朗日量 $L=T-V$（动能减势能）。端点固定，考虑扰动 $x+\varepsilon\eta$（$\eta(t_1)=\eta(t_2)=0$）。要求 $\frac{d}{d\varepsilon}S[x+\varepsilon\eta]\Big|_{\varepsilon=0}=0$ 对一切 η 成立，分部积分得 $\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial\dot{x}}=0$。取 $L=\frac{1}{2}m\dot{x}^2-V(x)$，得 $m\ddot{x}=-V'(x)$，即牛顿第二定律。注意：驻值不一定是最小值；“最小作用量原理”的名字在一般情形下并不准确。并且这个原理要求事先给定两个端点——起点和终点都要从现实里借来。

方框最后一句是本书关心的。作用量原理看起来比牛顿定律更“整体”，它一下子就看到了整条路。可是它要求事先知道终点。牛顿定律要的是起点和初速度，作用量原理要的是起点和终点。两种表述数学上等价，但向现实借的 C 换了一种形状：一个是初值问题，一个是边值问题（见第12章）。

◆ 二、量子力学的舞台：希尔伯特空间

二十世纪初，物理学碰上了一个更彻底的反常。原子发出的光只有一些特定的颜色，电子的能量只能取一些特定的值，一束电子穿过两条缝会像水波一样形成干涉条纹。这些都在经典力学的筛子之外。

1925 至 1926 年，海森堡的矩阵力学与薛定谔的波动力学先后出现。前者用无限大的数表（矩阵）描述可观测量，后者用一个满足偏微分方程的“波函数”描述状态。两者看起来完全不同，后来被证明是同一套理论的两种写法。把它们统一起来的数学框架，主要由希尔伯特一派的数学与约翰·冯·诺伊曼 1932 年的《量子力学的数学基础》完成：量子系统的状态，是一个无限维空间里的向量；可观测的物理量，是作用在这个空间上的算子；测量可能得到的数值，是算子的本征值。这个空间就是希尔伯特空间。

“无限维空间里的向量”听起来很玄，其实可以从一个熟悉的东西出发理解。录音棚的调音台上有一排推子，低音、中低音、中音、中高音、高音，每个

推子的位置合起来，就决定了一段声音的音色。如果推子不是五个而是无穷多个，每一个对应一个频率，那么一组推子的位置就能描述任何一段声音——这就是傅里叶在研究热传导时（见第17章）发现的事情。一个函数可以看作无穷多个分量组成的向量，所有这样的函数构成一个无限维空间。为了让“长度”“角度”“收敛”在这里还有意义，需要附加一些条件（例如各分量的平方和有限），满足这些条件的空间就是希尔伯特空间。

在这个舞台上，微积分要重新出场。对函数求导，变成了一个算子作用在无限维向量上；动量是一个求导算子，能量是哈密顿算子。薛定谔方程说的是：状态向量随时间的变化率，等于哈密顿算子作用在它上面。这仍然是一条微分方程，只不过未知量不再是一个数，而是无限维空间里的一个点。

> **为什么无限维需要额外的小心**　在有限维里，一个线性映射总是“连续的”，所有范数彼此等价，有界闭集一定紧致（序列总有收敛的子序列）。在无限维希尔伯特空间里，这三条都可能失败。下面①②是两种失败的例子，③是补救的一条路：　① 求导算子 $\frac{d}{dx}$ 是无界的：$\sin(nx)$ 的大小不变，求导后大小乘 n，可以任意大。所以量子力学里的动量、能量算子不能定义在整个空间上，只能定义在一个稠密的子集上，“自伴性”需要仔细检查。　② 单位球不再紧致：$\{\sin(nx)\}$ 这串向量长度都相同，彼此距离却不缩小，没有收敛的子序列。变分法里“极小值是否存在”因此成了真问题，需要索伯列夫空间、弱收敛等工具。　③ 冯·诺伊曼与斯通的工作表明，在适当条件下，自伴算子有谱分解，测量值的概率由此给出。这是把“微分算子”与“可观测量”连接起来的那道焊缝。

方框里第二条有一段著名的历史。十九世纪中叶，黎曼在证明一个重要定理时，用了所谓“狄利克雷原理”：某个泛函一定能取到最小值。魏尔斯特拉斯 1870 年前后指出，这个推理不严格，并给出了一个泛函有下界却取不到最小值的例子。“最优的那一条”可能根本不在候选之中。直到 1900 年前后，希尔伯特才在适当条件下挽救了这一原理。这件事对后文很要紧：在无限维里，“存在最优解”本身需要证明，不能因为问题提得出来就假定答案在那里。

◆ 三、对所有可能的路求和：路径积分

量子力学最令人困惑的地方之一，是双缝实验：一个电子穿过两条缝时，似乎“同时走了两条路”，两条路的贡献相互干涉。1948 年，理查德·费曼把这个想法推到极致，给出了量子力学的路径积分表述（狄拉克 1933 年的一篇论文已有萌芽）。

费曼的说法是：一个粒子从起点到终点的概率幅，等于对所有可能路径的贡献求和。每一条路径——直线、弯曲的、绕远的、折返的——都贡献一个大小相同、相位由该路径的作用量决定的复数。把所有路径的贡献加起来，就得到粒子到达终点的概率幅。经典路径附近，作用量变化缓慢，相邻路径的相位差不多，贡献彼此加强；远离经典路径的地方，相位快速旋转，贡献彼此抵消。于是在宏观世界里，只有经典那一条路“显露”出来——这正是本章第一节所讲作用量原理的来源。

回到开头的导航。经典力学像一个只给你一条最优路线的导航；路径积分则像一个把每条可能路线都“走一遍”、按某种规则加权、再把结果叠起来的导航。在城市尺度上，它们给出几乎相同的答案；在原子尺度上，被抵消得不完全的那些路线开始显露，粒子表现出干涉。

底本把路径积分称为“积分性发明的极致表达”。这个说法有它的道理：积分的对象从一个区间上的数，变成了一个无限维空间里的所有路径。但需要补上一个关键的事实，它恰恰是本书主命题的一个强证据。

> **路径积分的数学地位**　形式上写作 $\langle x_f|e^{-iHt/\hbar}|x_i\rangle = \int \mathcal{D}[x(t)]\, e^{iS[x]/\hbar}$。　对应的“路径测度” $\mathcal{D}[x(t)]$ 在通常意义下并不存在：卡梅伦 1960 年证明，不存在一个可数可加的复测度能使这一表达式按字面成立。　物理学家的实际做法：把时间切成 N 段，对每个中间时刻的位置做普通积分，再令 $N \to \infty$——这是第18章那种离散化的又一次运用。　一条严格的路：把时间转到虚轴（“维克转动”，$t \to -i\tau$），$e^{iS/\hbar}$ 变成实的衰减因子 $e^{-S_E/\hbar}$，此时可以用维纳测度（第19章的布朗运动）严格定义，即费曼—卡茨公式（卡茨，1949）。　对实时间、相互作用的量子场论，严格的

> 数学构造在一般情形下仍未完成；四维杨—米尔斯理论的存在性与质量间隙是克雷研究所千禧年难题之一。

这个方框说明了一件很少在科普里讲清的事：二十世纪最成功的物理理论之一，它的核心积分在数学上至今没有一般的严格定义。物理学家用它算出的数，与实验在十位以上的有效数字上吻合（例如电子的磁矩：2023 年的测量把电子 g 因子测到约万亿分之零点一的精度，理论值与之相符到约十二位有效数字，其反常部分本身也吻合到约十亿分之一），靠的不是一个完成了的积分，而是一套切片、重整化、与实验对账的程序。

用全书的语言说：路径积分把积分推到了理念界最远的地方——对所有可能的路求和。可是在那个地方，积分自己合不拢了。它要靠离散化（再一次截取）、靠转到虚时间借用布朗运动的测度、靠重整化里从实验借来的物理参数，才能给出数。被积的东西是“所有可能”，可是算出数来的每一步，都要从现实借回一点东西。

这里要收窄一句。这不等于说路径积分“不科学”或“不严格就不可信”。它的预言被反复检验，是物理学最可靠的工具之一；它的数学严格化是一个活跃而有成果的研究领域。要说的是：在这里，“能算出来、算得很准”与“作为积分对象在数学上完整存在”是两件事，前者并不依赖后者全部完成。这对第六编讨论“可计算性”时是一个有用的提醒。

◆ 四、泛函分析：把函数当成点

变分法、希尔伯特空间、路径积分有一个共同的数学底座，叫泛函分析。二十世纪上半叶，希尔伯特、里斯、巴拿赫、冯·诺伊曼等人建立了它。它的核心想法一句话就能说清：把函数当作一个空间里的点，把作用在函数上的运算（求导、积分、求解方程）当作这个空间里的映射，然后用几何的直觉——距离、角度、投影、收敛——来研究它们。

这个想法的力量，在于它把许多看似无关的问题变成了同一类问题。解一个偏微分方程，变成了在函数空间里找一个点；求一个变分问题的最优解，变成了在函数空间里找最低点；第18章讲的有限元，变成了在一个有限维子空间里做投影。有限元为什么收敛、误差有多大，正是靠泛函分析里的定理（如拉克斯—米尔格拉姆定理、塞阿引理）说清楚的。

生活里有一个类似的转换。装修房子时，与其一件一件考虑每样家具摆在哪里，不如把“整套摆法”看成一个整体方案，在方案与方案之间比较：这个方案比那个方案离理想多远，从这个方案改到那个方案要挪动多少。泛函分析就是给“方案的空间”配上一把尺子。

◆ 五、与现代 AI 的衔接

这一编最后需要接上第五编。今天的深度学习，在数学上可以看作泛函分析与变分法的一个后裔：一个神经网络是由几亿到上万亿个参数决定的函数，所有参数的取值构成一个极高维的空间；训练就是在这个空间里找一个使“损失泛函”尽量小的点。梯度下降是在这个空间里沿着“下降最快的方向”一步一步走，正是第18章所讲的离散化迭代；有限维之外，理论研究者还用无限宽网络的极限、函数空间上的梯度流来理解训练的行为。

底本称梯度下降、反向传播等“本质上都是在高维乃至无限维的参数空间中做微分与积分”。这句话的前半句是准确的描述；“本质上”一词需要收窄：这一框架描述了 AI 训练的数学形式，但它是否抓住了 AI 能力的全部来源，是一个开放的研究问题。至于 AI 把千亿次梯度累加进权重这件事，在本书的意义上算不算一种“不成文的积分”、被藏起来的是什么，留给第25至27章细说。这里只指出一点：本章看到的每一个无限维问题——最优解是否存在、积分是否有严格定义、离散化是否收敛——在 AI 的参数空间里都以新的面目重新出现。

◆ 六、收束

从最速降线到路径积分，微积分的对象从一个数，扩张成了一条路，再扩张成所有可能的路。这是理念界的一次巨大扩张：人不再只问“这一点怎样变”，而能问“在所有可能里哪一个最好”“所有可能合起来是什么”。

可这次扩张越远，向现实借的东西不是越少，而是越精细。变分法要借端点；希尔伯特空间要借“哪些算子对应可观测量”的物理判断；路径积分在数学上合不拢的地方，要借离散化、虚时间和实验参数补上。积分对象从量变成了路，积分常数 C 也从一个数变成了整套边界与约定。第四编的五章由此可以合起来看：偏微分方程、数值、随机、记忆与非局域、无限维，每一次扩张都让第一刀少切掉一些纠缠，或者把积分推到更远；而每一次扩张，都在落点上留下了新的、必须从现实借回来的东西。下一编先看混沌怎样让这笔借款利滚利，再看人类怎样在工程、控制与 AI 里，把这些借款一笔一笔还上。

◆ 七、这一编怎样算错

第四编（第17—21章）的总主张是：微积分的每一次扩张，都改变了“第一刀切掉多少纠缠”与“落点上要从现实借回什么”，但没有取消后者。下列情形出现时，相应主张须收回或收窄。

1. 若能找到一类重要的工程或科学计算，其结果的可靠性被证明与初值、边界、网格、参数等外部输入的选取基本无关（即不需要“验证与确认”中的确认一步），则第18章“数值计算只扩展了圈的理念一侧、落点仍须向现实借 C”的主张须收窄为“对多数而非全部计算成立”。
2. 若在主要应用领域（金融、物理、生物）里，以布朗运动为基础的随机模型在独立的样本外检验中，与不假定增量独立、方差稳定的模型相比，系统地同样好或更好，以致“现实中不属于这种形状的不确定性会在极端情境回流”找不到可检验的例证，则第19章关于“随机微积分只收编了形状规矩的那部分不确定性”的论断须收回。

3. 若分数阶模型在换一种加载或刺激历史后，不能比参数个数相当的整数阶模型给出更好的样本外预测，则第20章“分数阶导数少切掉了时间方向上的纠缠”应降为“一种有效的曲线拟合工具”，不再赋予它记忆机制的解释地位。
4. 第20章关于伞场微积分的唯一正面主张，是限定核、尺度归一、正则条件下的一维点态局域极限。若该命题的条件被发现不足以支撑结论（例如方框中的推导有漏洞），须收回；凡出现“传统数学是伞模型极限或特例”“相容性是普遍公理”“三态求解器已完成”等表述，无论出于哪一章，都应视为违反本书口径而删除。
5. 若路径积分在实时间、相互作用量子场论中获得了一般的严格数学构造，第21章关于“积分自己合不拢、要借离散化与实验参数补上”的例证须改写为历史陈述；但如果构造完成后实际计算仍依赖重整化中从实验借来的参数，“落点借 C”的论断仍可保留。
6. 若用 S、D、E 的语言重述本编各章的数学扩张，不能比通常的数学史与科学哲学叙述多出任何可检验的区分（例如对何时会出现反常、何处会出现失败的预期），则本编中的 SDE 解读只算“整理语言”，应如实降格，不得作为额外的解释力主张。

带 走 的 话

积分的对象从一个数推进到“所有可能的路”之后，最该记住的是一个分别：算得很准，不等于被积对象在数学上已经完整存在。实时间路径积分至今没有一般的严格定义，却给出了与实验吻合到十位以上有效数字的结果；它靠的是切片、虚时间与重整化这一整套程序，程序的每一步都要从现实借回一点东西。在无限维里，“存在最优解”“积分有定义”本身都需要证明，不能因为问题提得出来就假定答案在那里。

小 账

- 本章的导航选路、老司机去机场、调音台推子、装修方案都是帮助理解的比喻，不对应真实个案。
- 本章没有使用任何调查数字；卡茨 1949 年、卡梅伦 1960 年的结果与电子磁矩的测量精度见附录I第I58、I59条，最速降线、冯·诺伊曼 1932 年专著、魏尔斯特拉斯的反例、费曼 1948 年等史实据通行文献。
- 本章不主张路径积分“不科学”或不可信，只主张“能算准”与“作为积分对象完整存在”是两件事。
- 本章不主张参数空间上的微分与积分抓住了 AI 能力的全部来源。
- 本章主张的推翻条件见附录G第G34条；第四编其余各条见第G29—G33条与第G6条，本章末“这一编怎样算错”逐条写出了其中第18—21章与 SDE 解读的部分。

第 五 编

文明：混沌、工程、控制与 AI

编 序

第五编只做一件事：看微积分走出教科书以后，怎样变成工程、机器与 AI 的日常工作，以及在每一处，C 由谁来借。

第22章讲混沌：洛伦兹的一次四舍五入告诉我们，可计算不等于可预测；初值敏感，就是积分常数 C 的放大，集合预报则是承认借不准、于是多借几份。第23章讲工程：从都江堰的经验技艺，到图纸上做完积分的预测科学；纳维的桥、塔科马与千禧桥说明，纸上算对了，现场仍会把账顶回来。第24章讲控制论：从调速器、PID 到卡尔曼滤波，反馈就是每秒成百上千次的再微分，C 不必一次借准，但不能不借。第25章讲 AI 的微积分：梯度下降、反向传播、神经 ODE、扩散模型，训练是一场巨型的微分，产物是一次不成文的积分。第26章追问 AI 藏起了什么：微分在两端，积分写进了权重而没写成方程，这一点与老中医的手感相似，可解释性研究正试着把它写成文。第27章讲 AI for Science：AlphaFold、天气模型、材料发现，替得了积分；药效、剂量、临床与实验台上的那一步，仍要在现场落下。

读这一编时，可以带着一个问题：机器把积分做得越来越好之后，圈的两端站着谁？

22

第 22 章

混沌与复杂性：线性世界的黄昏

> 这一章的问题：天气的方程早就写下，超级计算机也算得动，为什么明天下不下雨能信，下周三晒不晒被子却说不准？

星期天晚上，一位母亲打开手机上的天气软件，想决定下周三要不要把被子拿出去晒。软件给出的是“周三，多云转小雨，降水概率百分之四十”。她把这句话读了两遍，还是拿不定主意：四十是多还是少？她又去翻明天的预报——“晴，降水概率百分之五”。明天的话她信，下周三的话她半信半疑。

这个犹豫里藏着现代科学最深的一道裂缝。气象台手里的方程并不神秘：大气是流体，流体服从质量守恒、动量守恒、能量守恒，这些都能写成偏微分方程（见第17章），全世界最好的超级计算机日夜不停地在解它们。方程写得出，计算机也算得动，可是“下周三下不下雨”就是算不准。明天的预报可以精确到几点开始下，七天以后只能给一个百分数，半个月以后，气象学家干脆告诉你：只能看气候平均值了。

问题出在哪里？为什么一个写得出、算得动的确定性系统，仍然预测不了？本书的回答是：**可计算不等于可预测**。中间差的那一截，正是第12章讲过的积分常数 C——微分丢掉、积分找不回、只能在落地时从现实借回来的那一份。在线性世界里，借得不准的 C 只会带来不准的结果；在混沌世界里，借得不准的 C 会被指数地放大，直到结果与原来毫不相干。

◆ 一、线性的好脾气

先说清楚“线性世界”是什么样子，才知道它的黄昏意味着什么。

线性系统有一种好脾气：原因加倍，结果也加倍；两个原因一起作用，结果就是各自结果的相加。菜市场的秤就是线性的——一斤白菜加一斤萝卜，秤上显示两斤，不会因为白菜和萝卜放在一起而变成三斤。弹簧也是线性的（在不拉坏的范围内）：挂一个砝码伸长一厘米，挂两个伸长两厘米。

这个好脾气在数学上叫**叠加原理**。它的价值怎么夸都不过分：一个复杂的输入，可以拆成许多简单输入之和；每个简单输入单独求解，再把结果加起来，就是总的答案。傅里叶把任意信号拆成正弦波之和，工程师把复杂荷载拆成若干简单荷载，量子力学把任意状态拆成本征态的叠加——三百年的数理科学，很大一部分是靠这一招打天下的。

> **叠加原理与它的失效** 设 L 是一个线性微分算子，即对任意函数 u, v 与常数 a, b 有 $L(au+bv) = aL(u) + bL(v)$。若 u_1 是 $L(u)=f_1$ 的解，u_2 是 $L(u)=f_2$ 的解，则 au_1+bu_2 就是 $L(u)=af_1+bf_2$ 的解。非线性项打破这一点。以 $\dot{x}=x^2$ 为例：x_1、x_2 各自是解，x_1+x_2 一般不是解，因为 $(x_1+x_2)^2 \neq x_1^2+x_2^2$，多出的交叉项 $2x_1x_2$ 正是“两个原因互相纠缠”在数学上留下的痕迹。

线性世界好算，是因为它允许把原因一个一个分开处理——这正是第10章“其他条件不变”的数学理想；非线性世界里切开分别算再相加，就漏掉了交叉项。用本书的话说：**线性是微分那一刀最舒服的世界，因为在那里，被切掉的纠缠恰好等于零**。

问题在于，真实世界的大部分并不是这样。天气、湍流、心跳、生态、市场、交通——这些系统里，原因之间的交叉项不但不为零，而且常常是主角。三百年来科学的好运气，一部分来自它挑选了线性的、或者在小范围内可以近似为线性的问题；另一部分来自泰勒展开（见第10章）的承诺——任何光滑的函数，

在足够小的邻域里看都像一条直线。这个承诺是真的，但它有一个前提：“足够小的邻域”。混沌正是那种邻域根本守不住的系统。

◆ 二、洛伦兹的那一次四舍五入

1961 年冬天，麻省理工学院的气象学家爱德华·洛伦兹在一台简陋的计算机上跑一个十二个变量的简化大气模型。他想把一段模拟重看一遍，为了省时间，没有从头开始，而是从中途打印出来的一个数值接着算。打印纸上的数是 0.506，机器内存里存的是 0.506127。差别在小数点后第四位，相对误差不到万分之三。他去喝了杯咖啡回来，发现新算出的“天气”在头几天还和原来差不多，随后越走越远，两个月的模拟后已经毫无相似之处。

他先怀疑机器坏了，确认没坏后，意识到撞上了原理性的东西。两年后，他把模型压缩成只有三个变量的方程组，发表了那篇后来被引用无数次的论文《确定性的非周期流》。1972 年，他在美国科学促进会年会上的报告用了这样一个题目（据记载，题目是会议召集人代拟的）：巴西的一只蝴蝶扇动翅膀，会不会在得克萨斯引起一场龙卷风？“蝴蝶效应”由此得名。

> **洛伦兹方程与误差的指数增长**
>
> $\frac{dx}{dt}=\sigma(y-x),\quad \frac{dy}{dt}=x(\rho-z)-y,\quad \frac{dz}{dt}=xy-\beta z$。洛伦兹取 $\sigma=10,\ \rho=28,\ \beta=8/3$。方程里没有任何随机项，非线性只来自 xz 与 xy 两个交叉项。 两条从相距 δ_0 的初值出发的轨道，其距离在一段时间内近似按 $\delta(t)\approx\delta_0 e^{\lambda t}$ 增长，$\lambda>0$ 称为最大李雅普诺夫指数（洛伦兹系统在上述参数下约为 0.9）。若允许的最大误差为 Δ，则可预报时长约为 $T\approx\frac{1}{\lambda}\ln\frac{\Delta}{\delta_0}$。 要点在对数：把初值误差 δ_0 缩小到十分之一，可预报时长只增加一个常数 $\frac{\ln 10}{\lambda}$，而不是增加十倍。测量精度的指数级投入，只换来预报时长的线性增长。

想把预报从七天延长到十四天，不是把仪器精度提高一倍就够了，而是要把初值误差再缩小若干个数量级。而初值误差有一个绕不开的来源：我们不可能在

地球的每一立方米空气里都放一支温度计，更不可能把每一支温度计读到无限多位小数。

大白话说：这像打台球。高手打一杆，母球碰一颗、两颗彩球，落点可以算得很准；如果要求母球连续碰撞二十颗球后的位置，哪怕球杆的角度只偏了头发丝那么一点，最后那颗球也不知道会滚到哪个袋口去。每一次碰撞都把前一次的小偏差放大一些，二十次放大之后，头发丝就变成了整张台面。

◆ 三、初值敏感，就是积分常数 C 的放大

现在可以把混沌放进本书的主命题里看了。

第12章说过：微分方程只告诉你“此刻怎样变”，要知道“此刻在哪里”，还必须补上初始条件——也就是积分常数　C。方程属于理念一侧，可以写在纸上、放进计算机；C　属于现实一侧，必须到现场去量。牛顿力学的辉煌，部分建立在一个心照不宣的前提上：C　量得大致准就行，因为量错一点，结果也只错一点。拉普拉斯那个著名的“妖”——一个知道宇宙此刻全部位置与速度的智者，能算出全部过去与未来——正是把这个前提推到了极限：只要 C 无限精确，预测就无限精确。

混沌没有推翻方程，它推翻的是这个前提。在洛伦兹系统里，方程是完全确定的，计算机可以一步一步算下去；但算出来的结果有多少意义，完全取决于借回来的 C 有多准，而 C 的每一点误差都会被按 $e^{\lambda t}$ 放大。于是本书可以把这件事说得更直白一些：

- **可计算**，说的是理念一侧的事：给定方程和一个精确的初值，能不能按有限步骤得出结果。洛伦兹方程是可计算的。
- **可预测**，说的是理念与现实交界处的事：用从现实借回来的、必然带误差的初值，能不能得出对现实有用的结果。超过某个时间之后，洛伦兹方程不可预测。

两者的差别不在方程里，而在 C 的借贷关系里。拉普拉斯妖之所以无所不能，是因为它被设定为不需要向现实借 C——它“知道”一切。可是任何真实的观测者都必须借，借来的东西都有误差。混沌告诉我们的是：在一大类系统中，这笔借贷的利息是复利，而且利率不低。

这个看法有一个可以检验的推论，本书把它当作检验把手写出来：**如果“可计算≠可预测”的差距确实来自 C 的外借，那么预报技巧的提高，应当主要来自借 C 的手段的改进（观测网、卫星、资料同化），而不是方程形式的改变。** 气象学的历史大致支持这个推论。2015 年，《自然》杂志上鲍尔等人的一篇综述把数值天气预报的进步称为“静悄悄的革命”，三到十天的预报技巧大约每十年提前一天；作者们列举的来源包括卫星观测与资料同化、集合方法、计算能力，也包括模式分辨率与物理过程参数化的改进，而描述大气的基本方程，从十九世纪纳维和斯托克斯写下之后并没有本质改变。反过来说，**若有一天，人们发现预报技巧的主要提升来自对基本方程的修改，而与初值获取无关，本章“初值敏感即 C 的放大”这一解读就需要收窄。**

还有一点：计算机每一步的舍入误差（见第18章）在混沌系统里同样被放大，所以算出的长时间轨道严格说来只是“某个附近的 C”的答案——在混沌面前，连理念一侧的“精确”也要打折扣。

◆ 四、集合预报：承认借不准，就多借几份

既然 C 借不准，气象学家做了什么？他们没有放弃，而是改变了提问的方式。

1992 年前后，欧洲中期天气预报中心和美国国家气象中心先后把**集合预报**投入业务。做法很朴素：既然不知道今天的大气状态精确是多少，那就在观测允许的误差范围内，故意取几十个略有不同的初值，每个都算一遍，得到几十种“未来”。如果这几十种未来在下周三大多下雨，就说降水概率高；如果一半下一半不下，就说概率百分之五十；如果各说各的，就承认这一天的天气此刻还说不清。

那位母亲在手机上看到的“降水概率百分之四十”，就是这样来的。它不是气象台偷懒，而是一句诚实的话：我们借来的 C 有一个误差范围，在这个范围之内，有四成的可能会下雨。

从本书的角度看，集合预报是一次可以观察到的**回写**。按第16章讲过的 123 原理：洛伦兹的发现是 D 与 E 之间的矛盾——差异序列（数值试验一次次给出发散的结果）与气象学既有的土壤（“把方程算得更准就能把天气预报得更远”这一沉淀多年的信念和工作方式）相互冲撞；矛盾累积到一定程度，逼出了显露的改变：预报的产品从“一个答案”变成了“一组答案加一个概率”。这个改变又回写了差异与土壤：什么算“预报得好”被重新定义（不再看单一预报命中与否，而看概率是否校准得当），观测网的设计、计算资源的分配、公众理解天气的方式，都跟着改了。这一回写是有记录可查的：业务规程变了，评价指标变了，公众看到的界面也变了。它不是我们宣布出来的，而是可以在气象机构的文件和产品里看到的。

大白话：家长说不准孩子放学几点到家，就不再问“几点到”，而问“最早、最晚、大概率几点”，按最晚的安排晚饭——不是放弃预测，而是换一种对现实诚实的预测。

◆ 五、混沌不是无序：天气与气候

讲到这里，容易滑向另一个极端：既然连天气都算不准，科学是不是对复杂系统无能为力？不是。

洛伦兹方程的解虽然对初值敏感，却永远不会跑出相空间里一个有界的区域，而且最终会贴着一个形状奇特的集合运动——那就是著名的洛伦兹吸引子，画出来像一对蝴蝶翅膀。轨道在两只翅膀之间来回跳，什么时候跳、跳几圈，无法长期预测；但它总在这对翅膀上，不会跑到别处去。顺带一提，“洛伦兹方程确实有一个奇异吸引子”这件事，直到 1999 年才由数学家沃里克·塔克借助严格的计算机辅助证明完成，此前它是一个数值上显然、数学上未证的命题——这本身就是理念一侧追赶现实一侧的一个小例子。

这就是**天气与气候**的分别。下周三下不下雨，属于轨道在吸引子上的具体位置，长期不可预测；明年七月的平均气温大概多少、这座城市一年大约下多少天雨，属于吸引子本身的统计性质，可以有相当好的估计。大白话：你没法预测河口某个漩涡下一秒会出现在哪块石头后面，但你知道这条河往东流，夏天水大，冬天水小。

所以混沌划出的是一条**可预测性的边界**，不是一片“不可知”的荒原。边界这一边，C 的误差还没有被放大到吞没结果，可以预测具体轨迹；边界那一边，只能预测统计性质。气象学在这条边界两侧都做了大量有用的工作。底本说混沌是“有结构的不可预测”，这个说法本书保留，但要加一句限定：说它“有结构”，是因为吸引子及其统计性质可以被识别、被维持、被表达——这正是第2章所说 S 的“三个可”；而说它“不可预测”，是对具体轨迹、在超过某个时间尺度之后而言的。两句话各有其成立的范围，不能混用。

◆ 六、自组织与涌现：秩序从哪里来

混沌讲的是差异如何被放大，另一类现象讲的是差异如何自己聚成秩序。

把一锅薄薄的油从底下均匀加热，温差小的时候，热量只是一层层往上传，油面平静；温差超过某个临界值，油面上会突然出现一格一格近似六边形的对流花纹。这是贝纳德对流。1950 年代，苏联化学家别洛乌索夫发现一种化学反应的颜色会周期性地来回变，此后扎博廷斯基等人继续研究，人们看到溶液里自发地出现一圈圈向外扩散的化学波。1952 年，图灵写了一篇论文《形态发生的化学基础》，用一组反应扩散方程说明：两种化学物质如果扩散速度不同，又互相促进或抑制，原本均匀的分布可以自发地失稳，长出斑点或条纹。普里高津把这类远离平衡、靠持续输入能量维持的有序结构称为**耗散结构**，并因此获得 1977 年诺贝尔化学奖。

> **图灵失稳：均匀为什么会长出花纹**　反应扩散方程组：$\partial_t u = D_u \nabla^2 u + f(u,v)$，$\partial_t v = D_v \nabla^2 v + g(u,v)$。设 (u_0, v_0) 是均匀稳态，且在没有扩散时是稳定的。对它加上波数为 k 的小扰动，线性化后扰动的增长率由矩阵 $J - k^2\,\mathrm{diag}(D_u, D_v)$ 的

> 特征值决定（J 为反应项在稳态处的雅可比矩阵）。图灵指出：当两种物质的扩散系数相差足够大时，存在一段波数 k，使这个矩阵出现正实部特征值——扩散本来是“抹平差异”的，在这里反而让某一尺度的差异增长起来，于是长出有特定间距的斑图。 注意：这是线性稳定性分析，只说明哪一种花纹会开始长；花纹长成什么最终形状，要靠非线性项决定。

图灵机制的迷人之处在于反直觉：扩散本来抹平差异，墨水滴进水里终会均匀；可两种扩散快慢不同的物质互相纠缠，扩散反成了差异的放大器。大白话：厨房里油烟往外散得快，香味只在灶台边浓，结果屋里不是一片均匀的气味，而是一块一块的“味区”——比喻只帮助理解，不是化学结论。

这些现象常被归到“涌现”名下：单个油分子没有“六边形”，单个细胞没有“条纹”。安德森 1972 年的短文《多者异也》说：能还原到基本定律，不等于能从基本定律构造出一切。

底本在这里说得很重：“复杂性科学用数学证明了本书反复申明的命题”。按现行口径，这句话必须收窄。耗散结构、图灵斑图与本书所说“在 E 中，经 D，成 S”之间，确有可以对照的地方：花纹（显露）是在持续的能量流与物质交换（土壤）中、经由扰动的放大与筛选（差异序列）而稳定下来的；而且一旦稳定，花纹本身又改变了热量与物质如何流动——这是一个可以测量的“回写”。但本书的证据责任说得明白：**跨领域相似只帮助提问，不证明同构**。图灵方程有明确的变量、参数和稳定性条件，可以算、可以证、可以被实验否证；SDE 的三方程 $S = F(D, E)$ 等只是关系的占位记号，没有给出这些东西。所以正确的说法是：复杂性科学提供了一批具体的、可检验的“秩序从差异中长出来”的机制，它们与发生学的问题意识相互印证，可以帮助 SDE 提出更好的问题；但它们证明的是自己的定理，不是 SDE。

同样，第16章用到的秩序、混沌、介生三态，是分析语言，不是物理量。洛伦兹系统的“混沌”是一个有严格定义的数学性质（正的李雅普诺夫指数、拓扑传递性等），与 D2 三态里的“混沌态”只是同名。把两者混为一谈，是本书要避免的典型错误。如果一定要从发生学里借一个词来看自组织，最贴切的或许

是“介生的要害在承接”：对流花纹能长成，是因为扰动能够进入（差异能进来）、增长的模式能够被保留（后果能保留）、局部形成稳定的格子（局部结构能形成）、条件变了花纹也跟着重组（结构能被修订）。这四条是可以在实验里逐一检查的，也就是说，它们可以被检验，从而可以被证伪。

◆ 七、黄昏不是黑夜

本章的标题叫“线性世界的黄昏”，需要说明它不是什么意思。

它不是说线性方法过时了。分析非线性系统的第一步几乎总是在平衡点附近线性化——图灵失稳的分析就是这样做的。黄昏是说：线性不再能被默认为常态，它退回成一个有适用范围的特例。线性与非线性是要分别论证的描述维度，不能凭一张闭环图判定非线性；反过来，方程写成线性也不证明现实是线性的——那可能只是那一刀切掉了交叉项。

它也不是底本所说的“复杂性科学取代经典科学”：复杂性科学用的仍是微分方程、相空间、稳定性分析，只是用在了此前回避的问题上。真正改变的是期待：从“给我初值，我还你未来”，变成“给我初值的误差范围，我还你未来的概率分布，并告诉你这张分布能管多久”。

把这一章收拢，可以得到全书主命题在混沌问题上的一个具体版本：**微分方程属于理念一侧，可以写、可以算；预测则必须把积分常数 C 从现实借回来。在线性世界里，借贷的误差是单利；在混沌世界里，是复利。可计算性管不到这笔利息，因为利息长在 C 上，而 C 在现实一侧。**

这个判断对第六编讨论的那个论断也有直接意义。哈萨比斯的猜想大意是：自然界中的模式可以被经典学习算法高效地发现和建模（原话与出处见第28章）。这在吸引子与统计性质的层面上，有大量事实支持；但“下周三下不下雨”这类问题，卡住它的不是学习算法的能力，而是现场初值的精度。不管模型是人写的方程，还是机器从数据中学到的，只要系统是混沌的，它在具体轨迹上的预报期限都受同一笔复利的约束。近年来以机器学习为基础的天气模型在中期预报上取得了与传统数值预报相当甚至更好的成绩（见第27章），它们依然要从

同一张观测网里读取初始场，依然要做集合。这是检验本书判断的一个好场地：**如果有一天，某个模型在不增加现场初值信息的情况下，系统性地突破了由李雅普诺夫指数决定的那条预报期限，本章的核心论断就须收回。**

混沌把“落地时借回　C”这件事暴露得最为彻底：它让人看到，预测的成败往往不在方程有多漂亮，而在你从现场拿回了什么。接下来的一章转向一个比天气更需要“动工之前就知道答案”的领域——工程。一座桥不能先建起来、塌了再说；工程师是怎样在图纸上算出“会不会塌”，又在哪里必须向那一块具体的地基低头的？

带 走 的 话

可计算不等于可预测。洛伦兹方程完全确定，可以一步一步算下去，但预报所用的初值是从现场借回来的，必然带误差，混沌把这份误差按指数放大：初值精度提高十倍，预报期限只增加一个常数。所以卡住具体轨迹预报的，是现场初值的精度，不是方程或算法；而吸引子与气候统计仍然可以预测。混沌划出的是一条可预测性的边界，不是一片不可知的荒原。

小 账

- 本章母亲看天气软件、打台球、等孩子放学、厨房“味区”都是构想场景或比喻，不对应真实个案。
- 本章所述 1992 年集合预报投入业务、2015 年《自然》综述、塔克 1999 年的证明，核验见附录I第I65、I60、I61条；洛伦兹 1961 年的四舍五入、1972 年的蝴蝶报告等史实据通行文献；本章没有使用调查数字。
- 本章不主张复杂性科学证明了 SDE：耗散结构、图灵斑图与发生学的对照只帮助提问；洛伦兹系统的“混沌”与 D2 三态的“混沌态”只是同名。
- 本章不主张线性方法过时，也不主张复杂性科学取代经典科学。
- 本章主张的推翻条件见附录G第G35、G36条。

23

第 23 章

工程：从经验技艺到预测科学

> 这一章的问题：老泥瓦匠靠规矩盖房，工程师在动工前就算出梁会弯几毫米，可纳维算对了的桥为什么还是没能通车？

农村盖房子，老一辈的泥瓦匠有一套不成文的规矩：墙多厚、梁多粗、檩条隔多远，全凭师傅一句话。问他为什么这么定，他会说“一直都这么盖”“我师傅就是这么教的”。这些规矩大多管用，因为它们是几代人在本地的土、本地的木料、本地的雨水和风里，用倒掉的房子换来的。可是你要他盖一栋从没盖过的东西——比如一座跨过大河的桥、一栋三十层的楼——他就没底了。规矩只在他见过的情形里管用，出了这个范围，只能试；而桥和楼，是不能先盖起来、塌了再说的。

现代的结构工程师不一样。在一根钢筋都还没进场之前，他就能在图纸上告诉你：这根梁在设计荷载下会弯下去几毫米，最大应力在哪个截面，离屈服还有多少余量。他不是比老师傅更有经验，而是手里多了一样东西：一套能从局部推出整体、从此刻推出将来的方程。

本章的问题是：工程是怎样从“盖了才知道”变成“盖之前就知道”的？这个转变靠的是什么，又在哪里碰到了它必须低头的地方？本书的回答是：微积分让工程第一次能够在理念一侧做完积分——把材料、荷载、几何拼成一个整体模型，预先算出会不会塌；但每一座具体的桥、每一栋具体的楼，仍然要在落地时把那一份积分常数 C——这一块地基、这一批钢材、这一群走在桥上的人——从现实里借回来。工程是全书所说“再微分”最典型、也最不容出错的场地。

◆ 一、都江堰：经验技艺的高峰

先说清楚，经验技艺并不低级。

公元前三世纪中叶，秦国蜀郡守李冰主持修建都江堰。它把岷江一分为二：鱼嘴分水，内江灌溉、外江泄洪；飞沙堰在洪水时把多余的水和泥沙甩回外江；宝瓶口卡住进入成都平原的水量。整套系统不用一道大坝，却在两千多年里大体保持运转，至今仍灌溉着大片农田。后人把它的维护经验总结成六个字：“深淘滩，低作堰”——每年枯水期把河床淘到一定深度，堰不要筑得太高。

这里面有大量的知识。分水比例随水位自动调节（民间说“分四六，平潦旱”），泥沙借弯道环流被甩向一侧，每年岁修的深度有石刻标记可依。可这些知识是什么形态的？它们是在现场一年一年试、一年一年修、一代一代传下来的。用本书第4章的信息九宫来说，都江堰的知识丰厚地存在于现实界与理念界的符号态、逻辑态里——有口诀、有规矩、有因果推理——但没有上升到“数学态”，即没有一组方程能让人离开岷江、在另一条河上预先算出该怎么修。

这里要守住第395卷对“封顶”说法的限定：说某种知识停在符号态或逻辑态，只是针对具体任务的局部诊断，不能拿来整体评判一个文明。都江堰的工匠不是“缺乏逻辑”，他们的本地逻辑极其严密；工匠难以口述的东西，靠示范和岁修制度传下来，同样是严格的知识传递。差别只在一件事上：**这种知识的适用范围，与它被试出来的那片现场绑在一起**。都江堰的六字诀离开了岷江，就要重新试。这不是缺点，而是经验知识的边界。

大白话例子：一位做了三十年红烧肉的奶奶，火候、糖色、收汁的时机拿捏得分毫不差，但你让她把同样的味道在高原上的高压锅里做出来，她也得先失败两锅。她的知识是真的，只是和她那口锅、那个灶、那个海拔缠在一起。

◆ 二、方程进场：在图纸上做完积分

工程变成“工程科学”，转折点在十八到十九世纪。几个关键的方程在这一时期落定：伯努利和欧拉研究梁的弯曲，给出了今天叫作欧拉—伯努利梁的方

程；纳维在 1820 年代把弹性力学的一般方程写了出来，又提出了黏性流体的运动方程，二十多年后斯托克斯独立地给出了更完善的推导，这组方程因此叫纳维—斯托克斯方程。

> **一根梁怎样在纸上“弯下去”**　欧拉—伯努利梁方程：$\dfrac{d^2}{dx^2}\left(EI\,\dfrac{d^2w}{dx^2}\right)=q(x)$，其中 $w(x)$ 是梁在位置 x 处的挠度，$q(x)$ 是单位长度上的荷载，E 是材料的弹性模量，I 是截面惯性矩。一根长 L、两端简支、受均布荷载 q 的等截面梁，积分四次并代入边界条件 $w(0)=w(L)=0$、$w''(0)=w''(L)=0$，得到跨中最大挠度 $w_{\max}=\dfrac{5qL^4}{384EI}$。这里的四个积分常数，全部由“两端怎么支着”这一现实条件定下。同一个方程，两端改成固定，结果就是 $\dfrac{qL^4}{384EI}$，只有简支的五分之一。

这个方框很能说明微积分在工程里干了什么。方程本身来自对一小段梁的微分分析：假定截面变形后仍是平面、材料在弹性范围内、变形很小——这是在现实里截一刀，切掉了剪切变形、材料不均、大变形等纠缠。然后在理念一侧积分四次，把小段的关系拼成整根梁的形状。但积分出来的东西还不是答案，它带着四个常数。这四个常数只能从现实里借：这根梁的两头到底是搁着、焊着，还是半搁半焊？工程师说的“边界条件”，就是本书说的 C。

大白话：裁缝手里有一套从人体比例推出的版型公式，这是理念一侧的积分；可每做一件衣服，还是要量这一个人的肩宽和腰围。同一个版型公式，肩宽量错两厘米，衣服就不合身。

方程进场带来的变化是根本的。工程师第一次可以在动工之前回答“如果”：如果跨度加大一倍，挠度会增加多少？（四次方，十六倍。）如果把实心截面换成工字形，省一半材料，刚度损失多少？这些问题过去只能靠试，现在可以在纸上算。这就是底本所说的“从经验技艺变为预测科学”，本书保留这个说法，但要补上它的限定：**能在纸上预测的，只是那个被截出来、被理想化之后的结构**。纸上的梁和工地上的梁之间，隔着一层被切掉的纠缠。

◆ 三、纳维的桥：纸上算对了，现场塌了

最能说明这层纠缠的，恰好是写下那些方程的人自己。

1820 年代，纳维在巴黎塞纳河上主持建造一座悬索桥，位置在荣军院对面。他是那个时代最懂力学的工程师之一，桥的设计建立在严格的计算之上，样式也极其优雅。可是，1826 年夏天桥快完工时，锚碇（拉住主缆的那一端基础）出现了开裂；同年 9 月，附近一条输水总管破裂，水渗进地基，锚碇处的土体松动，裂缝扩大，一侧桥塔向河面倾斜。桥最终没有通车，1827 年工程撤销，桥被拆掉了。据记载，事后有政府委员会批评纳维过于依赖数学计算，没有在计算结果之上留出足够的安全余量。至于地基本身的沉降与水管破裂各占多大责任，各家记载的侧重有出入。

把纳维与都江堰的工匠放在一起看，会发现两种知识各自把 C 放在了不同的地方。工匠的知识里，C 从来没有被拿出去过：岷江的水量、河床的砂石、每年的洪峰，都已经缠在“深淘滩，低作堰”这六个字里，所以口诀在岷江上几乎不会错，搬到别处却要重试。纳维的知识把 C 拿了出去，方程因此可以带到任何一条河上；代价是，每到一处都必须把 C 重新借回来，借错了，再漂亮的方程也接不住。前者的风险在“搬不走”，后者的风险在“借不准”。现代工程的全部制度，大体就是在这两种风险之间找平衡。

纳维的桥常被当作“理论派不如经验派”的笑话。本书不这么读。纳维的计算没有错，错在他借回来的 C：他假定锚碇所在的土是稳的，而那一块具体的土不稳；他没有想到旁边会有一条水管破裂。这两件事都不在方程里，也不可能在方程里——它们属于这一个现场、这一段时间。

这个故事给出了工程里一条朴素而深刻的规矩的来历：**安全系数**。今天的结构设计规范，不会让你按计算刚好够用来设计，而是要求荷载乘以大于一的分项系数、材料强度除以大于一的分项系数，两头都留余量。从本书的角度看，安全系数不是对理论的不信任，而是对“C 借不准”的制度化承认：方程在理念一侧可以算得很精确，但从现实借回来的地基承载力、钢材强度、实际荷载，都有一

个误差范围；安全系数就是给这个误差范围预留的位置。它是经验技艺在预测科学内部留下来的一份遗产。

◆ 四、现场把账顶回来：塔科马与千禧桥

纳维的桥是 C 借错了。还有一类失败更深一层：不是现场的某个参数借错了，而是方程在截取时切掉了一整种现象，而这种现象在现场把账顶了回来。

1940 年 11 月，美国华盛顿州的塔科马海峡大桥通车才四个月，在一阵每小时六十多公里、并不算特别猛烈的风中剧烈扭转，最终坍塌。这座桥的桥面又窄又薄，按当时的设计方法，它完全能扛住静力风荷载。问题在于，当时的设计主要把风看成一个静止的推力，而真实的风与桥面之间存在一种动态的相互作用：桥面一扭，改变了气流；气流一变，又推着桥面扭得更厉害。后来的研究把这类现象称为气动弹性颤振。注意，流行的说法常把它讲成“共振”，这并不准确：它不是风恰好以桥的固有频率周期性地推，而是桥与风的相互作用让振动自激地放大。

六十年后，伦敦。2000 年 6 月，横跨泰晤士河的千禧桥开放，当天成千上万的人走上去，桥身出现了明显的左右晃动，开放两天后就被关闭。原因找到了：人走路时身体会微微左右摆，平时各摆各的，相互抵消；桥一旦开始轻微横晃，行人为了站稳，会不自觉地顺着桥的节奏迈步，于是大家的步伐渐渐同步，推得桥晃得更厉害。工程界称之为同步横向激励。桥后来加装了阻尼器，于 2002 年重新开放。

这两座桥的设计者都不是外行，他们的计算在当时的模型里都是对的。失败之处在于：**方程截取时，把“结构会反过来改变作用在它身上的力”这一份纠缠切掉了**。桥面会改变风，桥的晃动会改变人走路的方式——这些都是结构与环境之间双向的、非线性的交叉项（回忆第22章那个方框里的 $2x_1x_2$）。切掉它们，方程就好算了；切掉它们，现实就会在某一天把账顶回来。

这正是第3章所说理念物理界的“回流通道”在工程里的样子，也是第16章 123 原理在工程史上可以观察到的一次完整循环：桥塌了、桥晃了，一次次失事

（差异序列）与既有的设计规范和工程信念（土壤）之间的矛盾累积，逼出了显露的改变——风洞试验成为大跨桥梁的必经环节，气动稳定性写进了设计规范，人行桥的横向动力性能也被列入检查项目。这些改变又回写了差异与土壤：此后什么算“设计合格”被重新定义，工程师的训练内容、规范的条文、甚至桥梁的造型（流线型箱梁取代窄薄的板梁）都变了。这一串改变都留有可辨认的痕迹——规范文本、试验标准、教科书目录。按现行口径，回写要被观察，不能被宣布；工程史是少数能把回写观察得这么清楚的领域之一。

大白话：一个家庭做年度预算，算得很细，但没有算“孩子上了初中，补习班会突然多出来”。不是不会算，是当初根本没把这一项放进表格。第一年超支了，第二年的表格里就多了一行。规范的每一条增补，几乎都是这样多出来的一行。

◆ 五、有限元：把再微分做成日常工作

二十世纪中叶以后，工程进入了一个新阶段。真实的结构——一架飞机的机翼、一座核电站的安全壳、一辆车的车身——形状太复杂，梁方程那样的解析解根本写不出来。工程师需要一种办法，把任意形状的结构都交给方程来算。

答案是**有限元法**。思路说起来很简单：把一个复杂的结构切成成千上万个形状简单的小块（单元），在每个小块里用简单的函数去近似位移，再把所有小块的关系拼成一个巨大的代数方程组，交给计算机求解。这个方法在 1950 到 1960 年代由航空与土木领域的工程师和数学家逐步发展起来（克拉夫在 1960 年的一篇论文里用了“有限元”这个名称），此后成为几乎一切工程设计的标配。

从本书的角度看，有限元是又一次截取（见第18章“离散化是又一次截取”）：连续的结构被切成离散的小块，每一块里的变化被近似为简单函数。但它还有一个更重要的意义：**它把“再微分”变成了工程师的日常工作**。在有限元之前，工程师手里的是一般性的方程；面对一座具体的桥，他要做大量简化才能把这座桥塞进某个已知解的模板里。有了有限元，他可以把这一座桥的真实几何、真实材料分布、真实的支座条件、甚至地质勘察报告里的土层参数，直接输

入模型。理念一侧的一般方程，被一次又一次地落回到具体的这一个结构上去算。

> **有限元：边界条件在哪里进场**　弹性体的平衡可写成变分形式：在所有满足位移边界条件的位移场 u 中，真实解使总势能 $\Pi(u)=\frac{1}{2}\int_\Omega \varepsilon(u):\sigma(u)\,d\Omega-\int_\Omega f\cdot u\,d\Omega-\int_{\Gamma_t} t\cdot u\,d\Gamma$ 取最小。把区域 Ω 分成单元，在每个单元上用节点位移插值，得到线性方程组 $K\mathbf{d}=\mathbf{F}$，K 为总刚度矩阵。未施加位移边界条件时，K 是奇异的——结构可以整体平移、转动而不产生应力，方程组没有唯一解。只有把“哪些点被固定住”这一现实信息代入，K 才变得可逆。这是积分常数 C 在离散世界里的样子：没有从现场借回的边界条件，再大的计算机也解不出一个确定的答案。

这个方框里最值得记住的是那个“奇异”。一个没有约束的结构，在数学上就是一个解不唯一的系统，计算机会报错。这不是数值技巧问题，而是一个原理：**模型自己不知道它站在哪里**。它必须被告知——由工程师，由勘察报告，由那一块具体的地基。

有限元让工程走得非常远，也让一种新的错误变得常见：模型越来越精细，工程师越来越容易相信屏幕上的彩色应力云图就是现实。行业里有一句老话：输入垃圾，输出垃圾。网格可以划到毫米，如果输入的土层参数是从隔壁工地抄来的，输出的漂亮云图就只是对一个不存在的结构的精确计算。

◆ 六、监测与数字孪生：落地之后继续借

近二十年，工程又往前走了一步。大桥、大坝、高楼上装满了传感器：应变计、加速度计、位移计、温度计，日夜不停地把结构的实际状态传回来。工程师用这些数据去校准模型，修正最初借来的那些 C——地基的实际刚度、混凝土的实际弹性模量、支座的实际摩擦。这套做法有个时髦的名字，叫“数字孪生”。

本书对它的看法很简单：数字孪生不是让模型摆脱了现场，恰恰相反，它把“从现场借　C”从一次性的勘察，变成了持续不断的借贷。动工前借一次，竣

工时借一次，此后每天每时都在借。这说明，越是成熟的工程科学，越承认理念一侧的积分离不开现实一侧的补给。

大白话：一个慢性病人，医生开药时按体重和化验单算剂量（一次借 C），但好医生不会就此不管，而是让他定期复查、根据新的化验结果调剂量。数字孪生对一座桥做的，就是这样的定期复查——只是频率高到了每秒钟。

这里可以写下一个检验把手。如果本章的判断成立——工程失败集中在 C 的外借上，而不是方程本身——那么对重大工程事故的调查结论做系统统计，主要成因应当集中在地基与地质条件、材料实际性能、施工偏差、荷载与环境作用的低估这几类“现场信息”上，而不是集中在力学基本方程的错误上。**若系统统计显示，工程事故的主要成因是基本力学方程本身失效，而非现场条件与截取时遗漏的作用，本章“工程失败集中于 C 的外借与被切掉的纠缠”这一主张须收回。**本书没有做这项统计；各国事故数据库口径不一，留待检验。

◆ 七、小结：工程是再微分的极端形式

把这一章收拢：工程从经验技艺变为预测科学，是因为微积分让它能够在理念一侧做完积分，在动工之前算出整体行为。但是工程比任何其他领域都更清楚地显示了第14章所说的那件事：**应用必须在一个具体的点上再微分**，而这一刻必须把 C 从现实借回来。都江堰的工匠把 C 和知识一起缠在岷江上，所以知识好用而难以搬走；纳维把知识从现场里提纯出来，所以可以搬走，却在借 C 时栽了跟头；塔科马和千禧桥告诉我们，连截取本身都可能切掉一整种会回头找你算账的纠缠；安全系数、风洞试验、有限元、监测，都是工程在“预测科学”的身份之下，为这份不可省略的外借所建立的制度。

工程的预测是在动工之前一次算好的。可是有一类机器，没法在出厂之前就把所有情况算好——蒸汽机要在负荷忽大忽小时保持转速，飞机要在阵风中保持航向。它们需要的不是一次再微分，而是每时每刻都在做的再微分。下一章讲的，就是人类怎样把这件事交给了机器自己。

带 走 的 话

微积分让工程能在理念一侧做完积分，动工之前就算出整体行为；但方程积分出来带着常数，边界条件、地基、钢材、荷载只能从这一个现场借回来。本书的判断是，工程的失败主要不在方程算错，而在 C 借错，或者截取时切掉了一整种会回头算账的作用；这一判断还等着事故成因的系统统计来检验。安全系数不是对理论的不信任，而是对“C 借不准”的制度化承认；越成熟的工程，越是持续地向现场借。

小 账

- 本章的老泥瓦匠、做红烧肉的奶奶、裁缝、家庭预算、慢性病人复查都是帮助理解的比喻，不对应真实个案；都江堰、纳维的桥、塔科马海峡大桥、千禧桥是真实工程。
- 本章所述纳维荣军院吊桥的经过，核验见附录I第I37条，事故主因各家记载侧重不一；其余史实据通行文献，附录I未单列条目。
- 本章没有做工程事故成因的统计，也没有使用调查数字，“工程失败集中于 C 的外借”是待检验的判断（附录I第I38条）。
- 本章不主张经验技艺低级，也不拿“封顶”整体评判一个文明。
- 本章主张的推翻条件见附录G第G37条。

24

第 24 章

控制论：反馈就是实时的再微分

> 这一章的问题：洗澡时一边看水温一边拧阀门，几下就能调好——这种边做边改的办法，为什么比事先把账算准更靠得住，它又靠不住在哪里？

冬天早上洗澡，打开热水器，水先是凉的。你把阀门往热的方向拧一大截，过了几秒，水突然烫得人一哆嗦，赶紧往回拧；拧多了，又凉了，再往热的方向拧一点……来回两三次，水温才稳在合适的地方。如果你性子急，每次都拧得很猛，水温会忽冷忽热好一阵子；如果你学乖了，看水温快到了就提前收手，几下就调好。

这个人人都做过的小动作里，有一整门学问。你在做的事情是：不停地看“现在的水温”和“想要的水温”差多少，根据这个差去拧阀门；你还会留意水温变化得多快，快到了就提前收手；如果水温总是比想要的低一点点，你会再多拧一丝。看差、看差变化得多快、看差积累了多久——这三件事，就是工程师写进几乎每一台自动机器里的那条控制律。

上一章讲的工程，是在动工之前把账算好。可是有很多事，没法在事先算好：负荷忽大忽小，风向忽东忽西，路面忽平忽陡。对这类事，人类找到的办法不是算得更准，而是**边做边看、看了就改**。本章要说的是：这个办法在数学上就是把微积分做成一个闭环，在本书的框架里，它是“再微分”的实时化——不是落地一次，而是每秒钟落地成百上千次，每一次都从现场重新借回 C。

◆ 一、调速器：反馈的机器化

十八世纪末，瓦特给他的蒸汽机装上了离心调速器。一根竖轴随发动机转动，轴上挂着两个铁球；转得快了，铁球被甩开、升高，通过连杆把进汽阀关小一些；转得慢了，铁球落下，阀门开大。不需要人守着，发动机自己就能大致保持转速。这类装置并不是瓦特首创，风车和磨坊里早有类似的东西，但瓦特把它用到了蒸汽机上，随着工业革命传遍了世界。

调速器有一个毛病：有时候它不但不稳，反而让发动机转速来回振荡，越振越凶。工程师们一度只能凭经验调。1868 年，麦克斯韦写了一篇论文《论调速器》，把调速器和发动机写成一组微分方程，指出系统是否稳定，取决于方程特征根的位置：如果所有特征根的实部都是负的，扰动会衰减；只要有一个特征根实部为正，扰动就会放大。这是控制理论的开端：**稳定与否，第一次可以在纸上判断**。

大白话：调速器就像一个自动的“看水温的人”。麦克斯韦做的事，是告诉你这个人什么情况下会把水温调稳，什么情况下会越调越乱——不用真的去洗那个澡。

◆ 二、PID：现在、过去和趋势

二十世纪初，自动控制进入了船舶。舵手掌舵，靠的是长期练出来的手感：船头偏了就打舵，偏得快就多打，快回来了就提前回舵。1922 年，俄裔美国工程师米诺尔斯基观察舵手的操作，发表了一篇关于自动操舵稳定性的论文，把舵手的手感写成了一条控制律。这就是后来遍布工业界的 PID 控制器的一个重要源头。

> **PID 控制律** 设目标值为 $r(t)$，实际测量值为 $y(t)$，误差 $e(t) = r(t) - y(t)$。控制器输出 $u(t) = K_p\, e(t) + K_i \int_0^t e(\tau)\, d\tau + K_d \dfrac{de(t)}{dt}$。 比例项（P）对此刻的误差作出反应；积分项（I）把过去的误差累加起来，只要误差长期不为零，它就会越

> 积越大，直到把稳态偏差消除；微分项（D）看误差变化的速度，相当于根据趋势提前收手，抑制超调。 三个系数 K_p, K_i, K_d 需要针对具体对象整定。整定方法有理论推导，也有经验法则（如齐格勒—尼科尔斯法），而在大量工业现场，最后一步往往是在设备上试出来的。

底本把这三项说成差异的“现在、过去、未来”，这个说法很好记，本书保留，但要加一个限定：微分项看的是此刻误差变化的速度，它只是按当前趋势外推，并不知道未来。说它“预判”，是在“趋势会持续一小会儿”这个假定下才成立的。

回到洗澡的例子：P 是“现在差多少就拧多少”；I 是“水温已经偏凉好一阵了，再多拧一点”；D 是“水温涨得很快，先收手”。只用 P，常常会留下一点永远消不掉的偏差（水温总是差半度）；加上 I，偏差能消掉，但容易过头；再加上 D，过头的幅度能压下来。

PID 不需要知道被控对象的详细方程。这是它最了不起的地方，也是它在工业界独占鳌头的原因：一台化工反应釜、一个空调房间、一个无人机的电机，内部机理各不相同，PID 只看误差这一个量，就能把它们大致管住。据工业界的常见估计，今天过程控制回路里的绝大部分仍是 PID 或其变形：德斯伯勒（Desborough）与米勒（Miller）2002 年报告的一项调查覆盖炼油、化工、造纸行业的一万一千多个控制器，其中 97% 的基础调节回路用的是 PID 算法；阿斯特伦（Åström）与哈格隆德（Hägglund）的教科书给出的估计也在九成以上。具体比例随行业与统计口径而异。

◆ 三、拉普拉斯变换：把微分方程变成乘法

控制系统一复杂，稳定性的分析就困难起来。一个闭环里串着放大器、电机、齿轮、传感器，每一样都有自己的微分方程，连在一起，方程组越来越难解。二十世纪三四十年代，电话、雷达、火炮控制的工程师们普遍采用了一套工具：拉普拉斯变换和传递函数。

传递函数：输入怎样变成输出 拉普拉斯变换把时间函数 $f(t)$ 变成复变量 s 的函数：$F(s)=\int_0^\infty f(t)\,e^{-st}\,dt$。它的关键性质是：在初始条件为零时，求导变成乘以 s，积分变成除以 s。于是一个线性常系数微分方程，变成了一个代数式。例如质量—弹簧—阻尼系统 $m\ddot{y}+c\dot{y}+ky=u$，在零初值下变为 $(ms^2+cs+k)Y(s)=U(s)$，传递函数为 $G(s)=\dfrac{Y(s)}{U(s)}=\dfrac{1}{ms^2+cs+k}$。控制器 $C(s)$ 与对象 $G(s)$ 接成单位负反馈闭环，闭环传递函数为 $\dfrac{C(s)G(s)}{1+C(s)G(s)}$。系统稳定的条件是：$1+C(s)G(s)=0$ 的所有根都在复平面左半边。PID 控制器的传递函数是 $C(s)=K_p+\dfrac{K_i}{s}+K_d s$——积分是 $1/s$，微分是 s。注意“零初值”这一前提：传递函数刻画的是系统对输入的响应，把初值的影响单独放在一边。

这个方框里藏着一件与本书主题直接相关的事。拉普拉斯变换之所以好用，是因为它把“求导”变成了“乘 s”——但这一步有个条件：初始条件为零。如果初值不为零，变换后会多出与初值有关的项。也就是说，传递函数这个漂亮工具，是在把积分常数 C 暂时请出门之后才变得漂亮的。

控制工程师怎么处理被请出门的 C？答案恰恰是反馈。开环系统（不看结果、只按事先算好的指令执行）的输出完全依赖初值和模型是否准确；闭环系统不断测量实际输出，把误差送回来修正，于是初值偏了、模型错了、外界扰动了，都会在误差里显露出来，再被控制器一点一点吃掉。**反馈的本质，就是不再指望一次把 C 借准，而是每时每刻重新借。**

1927 年，贝尔实验室的布莱克为了解决长途电话放大器失真的问题，想到了负反馈放大器：把输出的一部分反向送回输入。代价是放大倍数降低了，换来的是放大器对自身元件参数的变化变得不敏感——电子管老化了、温度变了，输出依然稳定。几年后，奈奎斯特给出了判断这类反馈放大器是否稳定的判据。这是反馈思想最清楚的一次展示：**用一部分性能，换取对“模型不准”的免疫。**

◆ 四、维纳：控制与通信是同一件事

1948 年，数学家维纳出版了《控制论》，副标题是“关于在动物和机器中控制和通信的科学”。二战期间，他研究高射炮的火控问题：飞机在天上飞，炮弹要飞几秒才到，必须预测飞机几秒后的位置；而飞行员看到炮火会规避，规避又改变了飞机的轨迹。维纳发现，这个问题里，预测、反馈、信息、噪声纠缠在一起，不能分开处理。

他从中提炼出一个大胆的看法：无论是机器、动物还是人的组织，凡是有目的的行为，都依赖于把“结果与目标的差距”作为信息送回来，再据此修正行动。伸手拿杯子，眼睛不停地把手与杯子的距离报告给大脑，大脑不停地修正手的运动——一个小脑受损的病人，会出现手越接近杯子越抖的“意向性震颤”，维纳认为这正是反馈过度、振荡失稳的表现。

维纳的视野后来产生了巨大的影响，也引来了大量的泛化与滥用：有一段时间，人们几乎把一切都说成“反馈系统”。这里正好用得上第395卷的一条纪律：**线性与非线性、单向与反馈是不同的描述维度，不能凭一张闭环图就下结论**。说某个社会现象是“反馈”，必须指出被测量的是什么、比较的目标是什么、修正动作是什么、时间延迟有多长；否则“反馈”就成了万能词。控制论之所以是一门科学，正因为它的每一个闭环都写得出方程、算得出稳定性、可以在实验台上被证伪。

◆ 五、反馈是实时的再微分

现在可以把控制论放进全书的主命题里了。

第14章说过：再微分，是把理念一侧积分出来的整体模型，落回现实的一个具体点上，算出“此刻该怎么做”；这一步必须把积分常数 C 从现实借回来。工

程（第23章）是在动工之前做一次再微分。控制系统做的是同一件事，只是做法完全不同：

- **模型（积分）在理念一侧**：工程师对被控对象有一个大致的模型（传递函数、状态方程），这是从物理定律和实验数据积分出来的整体知识。PID 甚至可以只用一个非常粗糙的模型。
- **测量（微分）在现实一侧**：传感器每隔一小段时间读一次实际值，求出误差——这是在流动的现实里截取此刻的增量。
- **控制输出（再微分）在交界处**：控制器把此刻的误差、误差的积累、误差的变化率合成一个动作，落回到执行机构上，改变现实。然后，现实变了，下一次测量又开始。

一台现代工业控制器，采样周期常在毫秒量级，也就是每秒一千次。这意味着：**每秒一千次在现实里截一刀，每秒一千次落回现实，每秒一千次从现场借回 C**。控制论没有消灭积分常数 C 的问题，它是把这个问题从“一次借准”改成了“永远在借”。

大白话：开车用导航。导航有全城的地图（积分好的整体模型），但它不会在你出发前一次性告诉你“方向盘在第几秒转几度”。它靠 GPS 每秒报告你在哪（测量），一偏离路线就重新规划（反馈），一路都在说“前方两百米右转”（落点上的再微分）。而真正把方向盘转到位、躲开突然冲出来的电动车的，是你的手和眼睛——那是一个更快、更细的反馈环。

这个说法有一个可以检验的推论：**如果反馈确实是“持续从现场借 C”，那么一个控制系统对模型误差和初值误差的容忍度，应当随反馈的速度与质量（采样频率、测量精度、回路延迟）而提高；反过来，当反馈中断或严重延迟时，系统的表现应当退化为开环，即完全依赖模型与初值的准确度。** 控制工程的大量实践与这一推论一致：回路延迟是稳定性的大敌，传感器失效是事故的常见起因。若有足够多的例证表明，某类系统在反馈中断后仍能长期保持与闭环同等的表现，而不依赖更准确的模型或初值，本章的这一解读就须收窄。

◆ 六、卡尔曼滤波：把“借 C”本身做成算法

PID 只看误差，不管系统内部是什么状态。二十世纪五六十年代，随着导弹与航天的需要，控制理论转向了“状态空间”的写法：把系统写成一组一阶微分方程 $\dot{x} = Ax + Bu$，其中 x 是描述系统此刻全部状况的状态向量。这样做的好处是可以处理多输入多输出、可以做最优设计；但它立刻带来一个问题：控制律需要知道此刻的状态 x，而传感器往往只能测到其中一部分，并且读数带噪声。

换成本书的话说：状态空间方法把积分常数 C 摆到了桌面中央。要算出“此刻该怎么做”，必须先知道“此刻在哪里”，而“此刻在哪里”正是需要从现实借回来、又借不准的那一份。

1960 年，卡尔曼发表了一篇论文，给出了在线性系统与高斯噪声假定下估计状态的最优递推算法，后来称为**卡尔曼滤波**。它的思路可以用两步说清：第一步“预测”，用模型把上一时刻的状态估计往前推一步，同时承认推出来的东西有误差，误差还会变大；第二步“校正”，拿到新的测量值，把预测与测量按各自的可信程度加权合并——模型越可信，越听模型的；测量越可信，越听测量的。然后进入下一轮。

> **卡尔曼滤波的一轮**　预测：$\hat{x}_{k|k-1} = A\hat{x}_{k-1} + Bu_{k-1}$，$P_{k|k-1} = AP_{k-1}A^{\top} + Q$。　校正：$K_k = P_{k|k-1}H^{\top}(HP_{k|k-1}H^{\top} + R)^{-1}$，$\hat{x}_k = \hat{x}_{k|k-1} + K_k\,(z_k - H\hat{x}_{k|k-1})$，$P_k = (I - K_kH)P_{k|k-1}$。　其中 z_k 为测量，H 为测量矩阵，Q、R 分别是模型噪声与测量噪声的协方差，P 是状态估计的误差协方差。读法：P 就是“C 借得有多不准”的账本。预测一步，P 因 Q 而变大——只靠理念一侧往前推，误差会积累；校正一步，P 因测量而变小——从现实借回的信息把误差压下去。增益 K_k 决定这一轮向现实借多少。

卡尔曼滤波很快被用于阿波罗登月计划的导航，此后进入卫星定位、飞行控制、手机里的运动传感器融合，以及第22章提到的天气预报资料同化（集合卡尔曼滤波是其中一类重要方法）。大白话：你在陌生小区里找朋友家，心里按“走

了几步、拐了几个弯”推算自己的位置（预测），越走越没底；看见一个门牌号（测量），马上修正。门牌号越清楚，你越信门牌；门牌被树挡了一半，你就更多地信自己的推算。

卡尔曼滤波让本书的判断多了一层清晰度：**最成熟的控制理论，并没有试图摆脱现实，而是把“理念一侧往前推会越推越没底、必须周期性地从现实借回信息”写成了一条可计算、可证明最优的规则**。它的账本 P 明明白白地记着：只积分不再微分，误差只会增长。

◆ 七、反馈的边界：控制不了的东西

最后要说清楚，反馈不是万能的。

第一，**反馈只能纠正已经显露在测量里的误差**。传感器测不到的东西，控制器不知道。一架飞机的空速管被冰堵住，读数错误，自动驾驶仪就会按错误的读数忠实地“纠偏”。反馈系统的眼睛就是它的传感器，眼睛看不到的地方，就是它的盲区。这与第3章说的“筛子”是同一个道理：能进入控制回路的，只是被测量筛过之后的那一部分现实。

反过来，传感器本身也会说谎：老化、漂移、被污染。成熟的系统因此常用多套不同原理的传感器互相印证，表决或交叉检查——这是在“借　C”这件事上再加一道保险。可是要注意，几套传感器如果装在同一个位置、受同一种冰冻，它们会一起出错；同源的测量不能互作独立证据。这条纪律在讨论多个 AI 系统互相校验时同样适用。

第二，**反馈会被延迟毁掉**。洗澡的例子里，如果热水管很长，从你拧阀门到水温变化要等十秒，你几乎一定会拧过头，水温在冷热之间来回振荡。控制理论精确地刻画了这一点：回路延迟会侵蚀稳定裕度，延迟过大，任何激进的控制都会失稳。

第三，**反馈会被混沌限制**（见第22章）。对一个混沌系统，小扰动会被指数放大，控制必须在扰动长大之前介入，这对测量频率和执行速度提出了极高的要求。有趣的是，混沌系统也有一个可以利用的性质：它对小扰动敏感，因此在适

当的时机给一个很小的推动，就能把它引到想要的轨道附近。这类“混沌控制”的研究在 1990 年前后兴起，它把混沌的弱点变成了控制的杠杆——但前提仍是能及时测到系统的状态。

第四，**反馈决定不了目标**。控制器只管把误差压到零，至于“想要的水温”是多少，是人定的。一个工厂的控制系统可以把产量精确地维持在设定值，却不会问这个设定值是否合理。这一点在下一章会变得更重要：当人们用梯度下降去训练神经网络时，网络同样在忠实地把“误差”压小——而误差是谁定义的、用什么数据定义的，不在优化过程之内。

◆ 八、小结

从调速器到 PID，从传递函数到卡尔曼滤波，可以看到一条清楚的线：每一次进步，都不是让控制系统更少地依赖现场，而是让它更聪明、更有分寸地依赖现场。麦克斯韦告诉人们闭环何时稳；布莱克用负反馈换来对元件老化的免疫；卡尔曼把“该向现实借多少”算成了最优增益。没有一步是“模型足够好，从此不用测”。

控制论把微积分做成了一个闭环。微分在测量端，积分在模型与控制器的积分项里，再微分在执行机构落回现实的那一刻，而这一刻每秒发生成百上千次。它对本书主命题的贡献是：积分常数 C 不必一次借准，可以持续地借；但它不能不借——一旦反馈中断，系统就回到开环，完全听凭模型与初值的准确度。反馈的眼睛是传感器，反馈的目标是人给的，这两件事都在闭环之外。

下一章转向一种看起来全新的机器：神经网络。它的训练过程，本质上也是一个巨大的误差反馈回路，只不过被调整的不是阀门和舵，而是千亿个参数。在那里，微分被做到了前所未有的规模。

带 走 的 话

按本书的读法，闭环控制把再微分做成了实时的：传感器在现实一侧截取误差，控制器把误差合成动作落回现实，每一轮都重新借回积分常数 C。卡尔曼滤波的误差账本 P 把这件事写成了算式：只靠模型往前推、不用测量校正，误差只会增长。这个读法也有能输的地方：若足够多的系统断开反馈后仍长期保持闭环表现，又不靠更准的模型或初值，它就要收窄。

小 账

- 本章洗澡调水温、开车用导航、在陌生小区找门牌等均为日常比喻，不对应具体个案。
- 本章所述控制史（瓦特调速器、麦克斯韦 1868 年《论调速器》、米诺尔斯基 1922 年、布莱克 1927 年、卡尔曼 1960 年）按通行文献记载；PID 占比出自德斯伯勒与米勒 2002 年的调查，核验见附录I第I39条。
- 本章不主张凡有目的的行为都是反馈，也不主张反馈能替人设定目标。
- 本章“反馈就是持续借 C”这一解读的推翻条件见附录G第G38条。

25

第 25 章

AI 的微积分：梯度、反向传播、神经ODE、扩散模型

> 这一章的问题：雾夜里只看得见脚下一两米的人，一步步试着也能摸下山——今天的 AI 是不是也这样“学”出来的，学完之后留下的又是什么？

起雾的夜里，一个人在山坡上迷了路，手电只照得见脚下一两米。他想下到山谷里的村子，可是看不见村子在哪，也看不见整座山的形状。他能做的只有一件事：用脚试一试四周，哪个方向地势往下倾得最陡，就往那个方向迈一小步；站定了，再试，再迈。一步一步，他未必走得最快，也可能被困在半山腰的一个小洼里，但大多数时候，他最终会走到低处。

今天几乎所有的人工智能模型，都是这样“学”出来的。只不过那座山不是三维的，而是有几十亿、几千亿个方向；那个人每一步要试的，不是脚下四周，而是在每一个方向上地势倾斜多少。这个“倾斜多少”，就是导数；在所有方向上的导数合在一起，就是梯度。

底本把神经网络的训练称为“一个巨型的微积分问题”，这个说法本书保留，并且要把它说得更细。本章只讲 AI 的“微分面”：梯度是什么，千亿个梯度怎样算得出来（反向传播），网络的深度怎样被看成连续时间（神经 ODE），图像又怎样从噪声里被一步步“倒推”出来（扩散模型）。在本章末尾，本书会指出：这一场巨型微分的产物——千亿次梯度累加进权重——其实是一次积分，而且是一次没有写成方程的积分。这件事留给第26章展开。

◆ 一、梯度下降：雾中下山的数学

先把“学习”说成数学。

一个神经网络，说到底是一个带着大量可调参数的函数。给它一张猫的图片，它输出“猫”的概率；给它一句话的前半截，它输出下一个词的概率分布。参数怎么定？先随便定一组，然后拿大量已知答案的例子去试：网络答对多少、错多少，用一个数来衡量，叫“损失”。损失越小，网络越好。学习就是：调整参数，让损失变小。

问题是参数有几十亿个，不可能一个一个去试。梯度下降的办法是：算出损失对每一个参数的偏导数——如果这个参数稍微增大一点，损失会增加还是减少，变化多快——然后让每个参数都朝着让损失减少的方向挪一小步。所有参数同时挪，这就是沿着梯度的反方向走一步。

> **梯度下降与随机梯度下降**　设参数为 $\theta \in \mathbb{R}^n$，在全部训练数据上的平均损失为 $L(\theta)$。梯度下降的一步：$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t)$，$\eta > 0$ 为学习率（步长）。 当 L 光滑、η 足够小时，每一步都使 L 下降，因为一阶泰勒展开给出 $L(\theta - \eta \nabla L) \approx L(\theta) - \eta \|\nabla L\|^2$。 实际训练中，算一次全数据梯度太贵，于是每步只取一小批样本 B_t 估计梯度：$\theta_{t+1} = \theta_t - \eta \nabla_\theta L_{B_t}(\theta_t)$，这叫随机梯度下降（SGD）。它的思想可追溯到 1951 年罗宾斯与门罗的随机逼近方法。 注意：对神经网络这类非凸函数，梯度下降一般只保证走到梯度很小的地方，不保证找到全局最低点。

这个方框里最关键的是那个一阶泰勒展开。梯度下降之所以有效，是因为在足够小的邻域里，任何光滑的地形都像一块斜面（见第10章“泰勒展开”）——这是微分最古老的承诺。雾中下山的人只看得见脚下一两米，他之所以敢迈步，正因为在一两米之内，山坡就是一块斜面。

这里有一处与本书主命题直接相关的地方：**每一次梯度计算，都是一次典型的“在现实里截一刀”**。取一小批数据，计算此刻参数下的损失变化率，假定其

余一切暂时不变。被切掉的是什么？是这一小批之外的全部数据，是这一步之后地形的弯曲，是参数之间的高阶交叉项。随机梯度下降之所以要“小步走”，正是因为它知道自己截下来的只是一个局部，步子迈大了，局部的斜面就不再代表真实的地形。

大白话：一个新手学炒菜，每炒一盘尝一口，咸了下次少放半勺盐，淡了下次多放半勺。他不会因为这一盘咸了，下次就一点盐都不放——那就是步子迈得太大。每一盘菜只告诉他一个局部的信息：这一锅、这一回、这个火候下，盐多了。

◆ 二、反向传播：链式法则的大规模执行

梯度下降的思想很早就有，柯西在 1847 年就用过类似的方法。困难在于算梯度。一个有几十亿参数的网络，如果对每个参数分别做一次“稍微动一动、看看损失变多少”的试验，要把整个网络跑几十亿遍，根本不可行。

解决办法是**反向传播**。它的数学内核，是中学就学过的链式法则：如果 y 依赖于 z，z 又依赖于 θ，那么 y 对 θ 的导数等于 y 对 z 的导数乘以 z 对 θ 的导数。神经网络是一层套一层的复合函数，输入经过第一层变成中间结果，中间结果经过第二层再变……最后变成输出和损失。反向传播做的事情是：先从输入往前算一遍，记下每一层的中间结果；然后从损失出发，往回一层一层地乘导数，每往回走一层，就顺手把这一层所有参数的梯度都算出来。

反向传播为什么省 设网络为复合函数 $L = \ell(f_K(\cdots f_2(f_1(x;\theta_1);\theta_2)\cdots;\theta_K))$，记第 k 层输出为 h_k。由链式法则，$\dfrac{\partial L}{\partial \theta_k} = \dfrac{\partial L}{\partial h_K}\cdot\dfrac{\partial h_K}{\partial h_{K-1}}\cdots\dfrac{\partial h_{k+1}}{\partial h_k}\cdot\dfrac{\partial h_k}{\partial \theta_k}$。关键在于前面那一串乘积 $\delta_k = \dfrac{\partial L}{\partial h_k}$ 可以从后往前递推：$\delta_k = \delta_{k+1}\cdot\dfrac{\partial h_{k+1}}{\partial h_k}$。每层只做一次局部运算，就把这一层所有参数的梯度都算出来。　结果：求出全部参数的梯度，计算量只是前向计算的一个小常数倍，与参数个数无关。这在自动微分理论里叫“反向模式”，芬兰学者林纳因马在 1970 年的硕士论文里就给出过一般表述；1986 年鲁梅尔哈特、辛顿、威廉姆斯的论文让它在神经网络界广为人知。

这个“只是一个小常数倍”是整个深度学习时代的计算基础。没有它，训练一个千亿参数的模型，一步梯度就要跑千亿遍网络；有了它，一步梯度大约只要跑两三遍。

大白话：一家连锁餐馆这个月亏了钱，老板想知道每一家分店、每一个部门、每一道工序各该负多少责任。笨办法是把每个环节单独改一改，看看总账变多少——几千个环节就要重算几千遍账。聪明的会计从总账出发往回倒：总亏损里，哪个区域占多少；每个区域里，哪家店占多少；每家店里，哪个部门占多少——一路往回分摊，一遍下来，每个环节的“责任份额”就都有了。反向传播就是这位会计。

但这里要补一句底本没有说的话：**链式法则能把“责任”分得清清楚楚，前提是损失函数本身被定义好了**。会计能分摊亏损，是因为“亏损”这个数是确定的；如果老板换一个指标——比如不看利润，看顾客满意度——同样的分摊程序会给出完全不同的责任表。反向传播是一台极其忠实的分摊机器，它不过问分摊的那个数是谁定的、为什么定这个。这一点与第24章所说“反馈决定不了目标”是同一件事：梯度下降是一个巨大的误差反馈回路，误差的定义在回路之外。

◆ 三、神经 ODE：把深度变成时间

2015 年，何恺明等人提出了“残差网络”，让网络可以叠到上百层而仍然训练得动。它的关键改动很简单：每一层不再直接输出新的结果，而是输出“在原来基础上的增量”，再加回到原来的结果上：$h_{k+1} = h_k + f(h_k)$。

数学家一看这个式子就会心一笑：这不就是解常微分方程的欧拉法吗？把 h 看成随“时间”变化的状态，f 看成变化率，每一层就是往前走一个时间步。2018 年，陈天琦等人在一篇获得当年神经信息处理系统大会最佳论文奖的工作里，把这个观察推到底：既然残差网络是欧拉法的离散步，那干脆不要离散的层，直接让网络学一个微分方程的右端，用现成的 ODE 求解器把状态从“时间 0”积分到“时间 1”。这就是**神经常微分方程**（Neural ODE）。

神经 ODE 与伴随方法 残差网络：$h_{k+1} = h_k + f(h_k;\theta)$。令层间隔趋于零，得 $\dfrac{dh(t)}{dt} = f(h(t),t;\theta)$，输出为 $h(T) = h(0) + \displaystyle\int_0^T f(h(t),t;\theta)\,dt$。训练时需要损失 $L(h(T))$ 对 θ 的梯度。若把求解器的每一步都存下来做反向传播，内存开销巨大。陈天琦等人采用控制论中早已成熟的**伴随方法**：定义伴随变量 $a(t) = \partial L/\partial h(t)$，它满足一个反向时间的微分方程 $\dfrac{da}{dt} = -a^\top \dfrac{\partial f}{\partial h}$，从 $t=T$ 往回积分到 $t=0$，顺路累积得到 $\dfrac{dL}{d\theta} = -\displaystyle\int_T^0 a(t)^\top \frac{\partial f}{\partial \theta}\,dt$。伴随方法是反向传播在连续时间里的对应物；它的根源可追溯到 1950—1960 年代的最优控制理论（庞特里亚金极大值原理）。

神经 ODE 的意义不在于它在所有任务上都比普通网络好——实际上，在很多任务上它更慢、更难训练。它的意义在于**把深度学习接回了微积分的主干**。过去，神经网络被看成一种“黑箱拼装”，与三百年的微分方程传统各走各的；神经 ODE 表明，网络的深度可以看成时间，层可以看成积分步，训练可以看成最优控制。前面三章讲的东西——连续动力系统、稳定性、反馈与控制——在这里重新相遇。

大白话：过去的做法像爬楼梯，一共一百级台阶，每一级的高度可以单独调；神经 ODE 把楼梯换成了一条斜坡，你只需要说清楚“每一处坡度多陡”，至于从坡底走到坡顶，交给专门算路程的仪器去算。台阶也好，斜坡也好，走的都是从输入到输出的那段路。

◆ 四、扩散模型：从噪声里倒着走回来

近几年最引人注目的图像生成模型，大多基于一个看起来很绕的想法：先学会怎样把一张图一点点弄脏，再学会怎样把脏图一点点变干净。

具体说，给一张清晰的照片，每一步往上加一点点随机噪声，加上几百上千步，照片就变成了一片毫无意义的雪花。这个“加噪”过程是事先规定好的，很简单。难的是反过来：给一片雪花，怎样一步步去掉噪声，变回一张像样的照

片？扩散模型训练一个神经网络，专门学“此刻这张脏图，噪声大概是往哪个方向加上去的”。学会了，就可以从一片随机雪花出发，每一步往“更干净”的方向挪一点，最后得到一张从未存在过、却像模像样的图片。

这个思路由索尔—迪克斯坦等人在 2015 年提出，2020 年乔纳森·何（Jonathan Ho）等人的“去噪扩散概率模型”让它的效果大幅提升；2021 年，宋飏等人把整件事统一写成了随机微分方程的语言——这就接上了第19章讲过的伊藤积分。

扩散模型的正向与逆向随机微分方程 正向（加噪）：$dx = f(x,t)\,dt + g(t)\,dW_t$，$W_t$ 为布朗运动。从数据分布 p_0 出发，到 $t = T$ 时变成近似的纯噪声分布。 逆向（去噪）：控制学者布赖恩·安德森在 1982 年证明，这个过程在时间上反过来，仍是一个扩散过程：$dx = \left[f(x,t) - g(t)^2\,\nabla_x \log p_t(x)\right] dt + g(t)\,d\bar{W}_t$，其中时间从 T 倒流到 0。唯一不知道的是 $\nabla_x \log p_t(x)$，即“得分函数”——在噪声水平为 t 时，概率密度上升最快的方向。扩散模型用神经网络 $s_\theta(x,t)$ 去学它。 于是，生成一张图像，就是从纯噪声出发，数值求解这个逆向随机微分方程。每一步都在问同一个问题：此刻往哪个方向走，概率密度升得最快？——这又是一个梯度。

请注意，扩散模型在两个层面上都是微分的：训练时，用梯度下降去学一个网络；这个网络学的东西本身，又是一个梯度（概率密度对位置的梯度）。生成图像的过程，就是沿着这个学来的梯度场，从混沌的雪花一步步“爬”到有形象的图片。

大白话：一张老照片泡了水、褪了色、长了霉，修复师傅一点一点去霉、补色、描线。他之所以知道“这里原来大概是什么样”，是因为他修过成千上万张老照片，知道照片“通常长什么样”。扩散模型做的事情类似，只是它从一片完全的雪花开始修，修出来的是一张从未存在过的“老照片”。

这里要守一个限定：说扩散模型“从噪声中让图像显露出来”，这个“显露”是在模型所学的那个数据分布之内的显露——它生成的东西，再新，也落在训练数据所勾勒的那片概率地形之中。这不等于说它不能产生新东西（第395卷已

撤回“演绎产出零新知识”一类说法），而是说，它的“新”的范围，被那片地形划定。地形之外的东西，它的梯度场指不过去。

◆ 五、梯度的病：消失、爆炸与第22章的老朋友

微分做到这么大的规模，也会生出微分特有的病。

回到反向传播那个方框：一个参数的梯度，是一长串导数连乘的结果。网络有几十层，就要连乘几十次。如果每一次乘的数都略小于一，乘几十次后，梯度小得几乎为零，前面几层的参数根本学不动——这叫**梯度消失**；如果每一次乘的数都略大于一，乘几十次后，梯度大得失控，参数一步就被甩到九霄云外——这叫**梯度爆炸**。1991 年前后，霍赫赖特在他的学位论文里系统分析了这个问题，1994 年本吉奥等人也给出了论证。

读者也许已经认出来了：这正是第22章混沌那个方框的翻版。在洛伦兹系统里，初值的微小误差沿时间被一次次乘以局部的放大倍数，平均放大率由李雅普诺夫指数刻画；在深度网络里，误差信号沿层数被一次次乘以每层的雅可比矩阵。前者让天气预报超过某个期限就失效，后者让过深的网络学不动。两件事在数学上是同一种结构：**一长串局部微分连乘，结果对每一个因子都极度敏感**。有研究者正是借用动力系统的语言，把“让网络处在既不消失也不爆炸的临界状态”作为初始化的设计原则。

工程师们找到了一系列对策：残差连接（让信号有一条可以绕过连乘的捷径，这正是神经 ODE 的出发点）、长短期记忆单元（霍赫赖特与施密德胡伯 1997 年提出，用“门”控制信息的保留与遗忘）、归一化层、梯度裁剪（爆炸时强行把梯度截到一个上限）、精心设计的初始化。这些办法没有一个是从“智能”的哲学里推出来的，它们全都是在微积分层面对连乘结构的修补。

大白话：传话游戏。一句话从第一个人传到第三十个人，每个人都稍微走一点样，最后面目全非。要么每个人都把话说轻了一点，传到最后听不见（消失）；要么每个人都添油加醋一点，传到最后变成了谣言（爆炸）。残差连接就像允许每个人除了听前一个人的话，还能直接看一眼最初那张写着原话的纸条。

这一节还有一层意思值得点出。混沌一章说，可计算不等于可预测；这里看到的是，**可微也不等于可学**。网络里每一个运算都可微，链式法则对每一层都成立，梯度在理念一侧算得出来；但算出来的梯度能否真正把参数带到好地方，取决于连乘结构是否良态、步长是否合适、数据是否覆盖了要用的地方。这些都不是“可微”这个谓词能保证的。

◆ 六、一场巨型微分，一次不成文的积分

把这一章收拢，可以看到一幅统一的图景：现代人工智能，从训练到生成，从头到尾是一场规模空前的微分。

- 梯度下降：每一步在一小批数据上截取一个局部斜率，迈一小步。
- 反向传播：用链式法则，把一个总误差的“责任”分摊到千亿个参数上。
- 神经 ODE：把网络的深度看成时间，层看成积分步，训练看成最优控制。
- 扩散模型：学一个梯度场，再沿着它从噪声走回数据。

这些都是第10章所说“在现实里截一刀”的大规模工业化执行。一次大模型的训练，梯度要算上百万步，每一步涉及千亿个参数——人类历史上从没有在任何一件事上做过这么多次微分。

可是，一场微分做完，留下来的是什么？是一组权重：最终的参数，等于初始的参数，加上每一步挪动的总和。在数学形式上，这是一次积分。**训练是一场巨型微分，它的产物却是一次积分。** 这次积分的严格写法与限定、它为什么没有写成人能读的方程、它与牛顿那种积分差在哪里，是第26章的主题，这里不展开。

本章只补一个从微分这一面看得见的细节：那个“初始的参数”是训练开始前随机撒下的一组数。它在形式上正占着积分常数 C 的位置——换一个随机种子重训，积出来的权重逐个数字都不同，模型的本事却常常相近。可见这次积分的产物，并不唯一地由数据与梯度决定；它带着一份起点上的任意。

本章还要留下一个检验把手。若本书“AI 的产物是一次不成文的积分”这一判断成立，那么可解释性研究的进展应当表现为：从权重中逐步提取出可被人表达、可脱离原模型独立检验的规律，而且被提取出的部分越多，人对模型在新情境下的行为就越能事先预测。**若有朝一日，人们能从一个大模型的权重里读出一组简洁的、可独立检验的方程，用它们重现模型的主要能力，并通过第26章所说的减法检验，则“不成文”这一定性就须收回，AI 的积分将不再是“藏起来的”，而是“写出来的”。** 至于“AI 能否替人做再微分”，本章不作结论，留给第27章与第六编。

带走的话

现代 AI 的每一次梯度计算，都是一次地道的截取：取一小批数据，看此刻的局部斜率，假定其余暂时不变。但可微不等于可学，梯度能否把参数带到好地方，要看连乘结构是否良态、数据是否覆盖要用的地方；损失由谁定义，也不在梯度回路之内。这场微分留下的权重，形式上是一次积分，带着随机起点的任意，也没有写成人能读的方程。

小账

- 本章雾中下山、学炒菜、连锁餐馆的会计、传话游戏、修老照片均为比喻，不对应真实个案。
- 本章所述方法的人物与年份（柯西 1847 年、罗宾斯与门罗 1951 年、林纳因马 1970 年、鲁梅尔哈特等 1986 年、安德森 1982 年、何恺明等 2015 年、陈天琦等 2018 年、宋飏等 2021 年）按通行文献记载，附录I未单列条目。
- 本章说“权重是一次积分”，只指数学形式上的累加关系；不主张权重里装着可读定律，也不主张扩散模型不能产生新东西。
- 本章“不成文的积分”这一定性的推翻条件见附录G第G39条。

26

第 26 章

不成文的积分：AI 藏起了什么

> 这一章的问题：老中医说“这个人的脉压不住十克”却讲不出为什么，大模型张口答对也讲不出为什么——两者的本事累在了哪里，藏起来的又是什么？

镇上有位看了四十年病的老中医。徒弟跟了他三年，每天抄方、煎药、听他搭脉。有一回来了个面色发黄的中年人，老先生三根手指搭上去，停了半分钟，开了方子，又把其中一味药的分量从十克改成六克。徒弟问：师父，书上写的是十克，您为什么改？老先生想了想，说：“这个人的脉，压不住十克。”再追问“怎么就压不住”，他摆摆手：“你摸多了就知道。”

徒弟并不觉得师父在故弄玄虚。他知道师父的手底下压着几万个病人：每一个病人都是一次细小的比较——这个脉配这个方子，三天后回来是好了还是没好，好到什么程度。几万次这样的比较，一点一点累进了那三根手指里。累出来的东西是真的，它让老先生在大多数病人身上改对了分量；可它没有被写下来，书上找不到，徒弟也抄不走。

今天的大模型，在一个要紧的地方和这位老先生很像：它也把海量的细小比较累加成了一个能用的整体，而这个整体同样没有写成别人能读的样子。本章要把这件事讲清楚——AI 的微分在哪里，它的积分在哪里，被藏起来的到底是什么，以及这对科学这个“微分—积分—再微分”的圈意味着什么。

◆ 一、微分在两端：训练的梯度与应用的落点

第25章已经讲过，神经网络的训练本质上是一个巨大的微积分过程：模型对每一批样本算出损失，再对几亿乃至上千亿个参数各求一次偏导，沿着梯度的反方向挪一小步。这里只取那一章的结论，换一个角度看它在全书主命题里的位置。

先看训练这一端。每一次梯度计算，都是一次地道的“微分”：在流动的数据里截取一小批，问“此刻参数往哪个方向挪一点，误差会减少多少”，其余一切都当作暂时不变。这和第10章所说的“在现实里截一刀”同构——截下来的是一个局部增量，切掉的是这一小批数据之外的全部语境。一个孩子学认字，老师指着“日”字说“这念 rì”，孩子在那一刻只记住了这一个字形与这一个读音的对应，至于这个字在诗里、在姓氏里、在“日子”里还有什么讲究，这一刀都切掉了。

再看应用这一端。训练结束以后，模型被拿到一个具体的地方去用：一位用户问一个具体的问题，一张具体的片子要判读，一条具体的蛋白质序列要预测结构。这一刻，模型要把它学到的整体落回一个具体的点上，给出“此时此地”的输出。这正是第14章所说的“再微分”：用整体模型在具体一点上算出“此刻该怎么做”。你在手机导航里输入目的地，导航说“前方两百米右转”——它背后的整张路网是整体，这句“右转”是落在你此刻所在路口上的再微分。

于是 AI 的微分分在两端：入口处，千亿次梯度一刀一刀地截取训练数据；出口处，一次次推理把整体落到具体的提问上。两端都看得见——训练日志记着每一步的损失，推理时我们看得见输入和输出。真正看不见的，是夹在两端中间的那一段。

◆ 二、积分写进了权重，没写成方程

夹在中间的那一段，是积分。

这不是比喻，至少在一个可以说清楚的意义上不是比喻。梯度下降每一步都把一个小小的修正量加到参数上，训练多少步，就加了多少次。训练结束时的权重，等于初始权重加上全部修正量之和。当学习率足够小，这个求和就逼近一个积分：权重是梯度沿训练轨迹的累加。

> **权重是梯度的积分**　设参数为 θ，损失为 $L(\theta)$，学习率为 η。梯度下降每一步为 $\theta_{k+1} = \theta_k - \eta \nabla L_k(\theta_k)$，其中 L_k 是第 k 批样本上的损失。走完 T 步，$\theta_T = \theta_0 - \eta \sum_{k=0}^{T-1} \nabla L_k(\theta_k)$。若把 η 看成时间步长、令其趋于零，并把批次间的随机性平均掉，这一和式逼近“梯度流” $\dfrac{d\theta}{dt} = -\nabla L(\theta)$ 的解，即 $\theta(t) = \theta(0) - \int_0^t \nabla L(\theta(s))\, ds$。　两点限定：其一，真实训练用的是随机梯度、自适应学习率、权重衰减等，上式只是理想化的骨架，严格的逼近需要光滑性与步长条件；其二，这里说“权重是积分”，指的是数学形式上的累加关系，不是说权重里装着一条可读的定律。恰恰相反——本章要指出的正是：这一积分的结果不以可读定律的形式存在。

请注意这个积分与牛顿那种积分的差别。牛顿也把无数瞬时的差异累加成整体，但他累加出来的东西写成了寥寥几个符号：$F = ma$，万有引力与距离平方成反比。一个中学生可以把它抄在本子上，带到考场，带到另一个星球上去算。这种积分的产物是一条“理念方程”：它住在理念界，上升到了数学态，而且对人“可表达”。

AI 的积分产物是一大张数表。几千亿个浮点数，每一个单独拿出来都毫无意义，合在一起却能写诗、译文、预测蛋白质结构。它也住在某种意义上的理念一侧——它不是任何一块具体的石头、一个具体的病人，而是从海量记录中抽出的整体规律；它的形式也确实是数——每一个参数都是数学对象。但它缺了一样东西：对人可读。

用第395卷“显露”的“三个可”来检查最清楚。说某物显露为 S，要看它是否可识别、可维持、可表达。AI　积分出的那个整体：可维持——权重存在硬盘上，明天再调用，还是同样的行为；在功能上可识别——它确实稳定地完成某类任务，我们能分辨它“会”什么、“不会”什么；但它对人基本不可表达——没

有人能把“模型为什么把这个蛋白折成这样”说成一句话、写成一条方程、交给另一个人去检验。“三个可”缺了第三个。缺了这一个，它就是一种“不成文的积分”。

用记账来打比方。一家小铺子每天流水进出，月底累计出余额。普通的积分像一本好账：余额是多少，看得见；为什么是这个数，翻流水能一笔笔对上；明年要做预算，可以按科目汇总出“房租占三成、进货占五成”这种规律。不成文的积分则像一个记性极好却不会记账的掌柜：你问他下个月能剩多少钱，他张口就报一个数，十有八九报得准；你问他凭什么，他说不出科目，也拿不出账本。他报得准是真本事，但这本事锁在他脑子里。

◆ 三、老中医的手感：另一种不成文的积分

回到开头那位老先生。第395卷在讲 E 的解剖学时有一个说法叫“封顶”：某一界的信息停在符号态或逻辑态，上不到数学态；书里举的正是老中医和老工匠——本地逻辑饱满，却没有被提取为可传授的理念逻辑。老先生那句“压不住十克”就是封顶的样子：判断是准的，逻辑在他手里是连贯的，可它停在他的身体与经验里，没有被写成方剂学教科书上的一条规则。

把老先生的手感和大模型的权重并排放，可以看见两种不成文的积分有同有异。

相同处在于结构：两者都是海量细小比较的累加，都在应用时表现为“张口就对”的整体判断，都对旁人不可表达。徒弟抄不走师父的手感，工程师也读不出权重里的规律。

不同处有三层，每一层都关系到后面两章的论证。

第一层，积分落在哪一界。老先生的积分落在自我界——和他的身体、记忆、几十年的行医身份缠在一起；他的手指本身就是测量仪器，积分与测量长在同一具身体里。大模型的积分落在一组参数里，而参数背后的训练数据，是早已被记录、被数字化的材料：文本、图像、已解析的蛋白结构。换句话说，老先生

的积分直接从现实界一个个病人身上来；大模型的积分从“已被写下来的现实”里来。前者少过了一道筛，后者已经过了第3章所说的那张“筛子”。

第二层，积分还活不活。老先生每看一个新病人，就是一次新的再微分，而这次再微分的结果——三天后病人回来说好了还是没好——又会回写进他的手感。他的积分是一边用一边长的。大模型的基础权重在训练结束时就定住了：部署后的每一次推理都是再微分，可这些再微分的后果，默认并不回到权重里。这正是第395卷所说的“冻结态”。但必须照第395卷的口径把话说准：冻结态只是“缺少有效回写”的功能性诊断，不是不可能性结论。一个具体的 AI 系统能不能让落点的后果回写进来，要按四项分别体检——基础模型、工具连接、记忆安排、人的参与。基础模型的权重不动，不等于整个系统不学：接上检索工具，挂上长期记忆，再让一位医生把每一次误判标回去、定期微调，这个系统的“手感”也可以一边用一边长。是否如此，要看配置，要看记录，不能一句话判死。

第三层，C　由谁补。老先生改分量之前，要先搭脉、看舌、问饮食起居——这是在落点上从现实借回积分常数　　C。他的手感再好，不搭这一下脉也不敢开方。大模型同样需要 C，它的 C 来自输入：用户写进对话框的那几句话、上传的那张化验单。问题在于，用户写什么、漏什么，模型无从追问到底；化验单之外这个病人昨晚喝没喝酒，模型不知道。老先生的 C 是他亲手借的，大模型的 C 是别人递过来的。

这三层差异不是在给老中医加分、给 AI 减分。老先生的手感也有它的病：传不下去，他一走就散了；看错了也没人能查，因为没有可以对照的明文。不成文的积分，无论装在人身上还是装在机器里，都有同一个软肋——它无法被别人检查。

◆ 四、想把积分写成文：可解释性研究在做什么

既然不成文是软肋，那就去把它写成文。这就是近十年“可解释性”研究的基本冲动。

最早的做法像给老中医录像：不问他为什么，只看他在什么情况下做什么，再从外面总结规律。对模型来说，这叫行为测试与归因——改动输入的某一处，看输出怎么变，从而推断模型“在看哪里”。孩子做错题，家长不去拆孩子的脑子，而是换几道相似的题看他在哪一步出错，道理一样。

更进一步的做法像解剖。研究者把手伸进权重内部，去找那些稳定对应某种概念的内部方向或“特征”：有的研究在下棋文本训练出来的小模型里，找到了能读出棋盘局面的内部表示（如 Kenneth Li 等人 2023 年对黑白棋模型的研究，Adam Karvonen 2024 年对国际象棋模型的研究）；有的研究用稀疏字典的方法，把大模型某一层的活动拆成上百万个特征，其中不少可以贴上人能理解的标签（Anthropic 公司 2023 年的“走向单义性”在单层小模型上试出这一方法，2024 年的“单义性的规模化”把它用到 Claude 3 Sonnet 上，拆出多达约三千四百万个特征）；还有的研究一路追踪某种行为在网络内部经过哪些组件，画出所谓“电路”（这一路的纲领见奥拉等人 2020 年的“放大看”，Anthropic 公司 2025 年的“电路追踪”又用“归因图”把它推到了大模型上）。

这些工作做的事情，翻译成本书的语言，就是把不成文的积分往“成文”那边推：把存在参数里的数学对象，提取成人能读的符号与逻辑，补上“三个可”里缺的那个“可表达”。这和中医界把名老中医的经验整理成医案、再从医案里归纳方剂规律，是同一类工作。

但写成文有它自己的门槛。一个特征被贴上“这是‘桥’的概念”的标签，未必真的是模型用来做判断的那个东西——它可能只是碰巧和“桥”一起出现。贴标签容易，验证难。

> **什么才算“写成了文”**　借用第395卷的两条纪律，可以给“把一段不成文的积分写成文”定三条最低标准：　一、可识别：这个内部特征在大量不同输入上，稳定地与某个人能命名的差异共变，而不是只在研究者挑选的几个例子上共变。　二、可维持：换一批数据、换一个随机种子重训，类似的特征仍能被找到；否则它只是这一次训练的偶然产物。　三、可表达并经得起减法检验：把这个特征“去掉”或“调高”，模型的行为要按我们的文字解释所预测的方向变化。第395卷讲人—机器复合体时称之为“减法检验”：去掉某环节，结果是否明显不同。贴了标签却通不过减法

> 检验的解释，只是事后讲故事。　按这三条，目前的可解释性研究已在若干局部通过检验，但离“把一个大模型的整体积分写成一套可读的理念方程”还很远；这是对研究现状的描述，不是对其前途的断言。

还有一层更深的难处。牛顿的方程之所以可读，不只是因为短，还因为它是用人已有的概念写的：力、质量、加速度。可模型积分出来的整体，未必是按人已有的概念切分的。它可能学到了某种有用的区分，而人类语言里根本没有现成的词。这时，“写成文”就不只是翻译，还得先命名——先为那个区分发明一个名字，再看这个名字能不能在别的地方站住。这一点留到第32章讨论“问题是谁命名的”时再展开。

还要补一句反方向的提醒：写成文本身也是一次截取。名老中医的医案整理出来，留下的是“某证用某方，加减某味”，丢掉的是他搭脉时手指下那一点说不出的轻重。把权重里的特征提取成一句“这个方向代表‘桥’”，同样丢掉了这个方向在其余上千种语境里的细微作用。成文的积分可读、可查、可迁移，代价是比不成文的积分更粗。经典科学一直在做这笔交易：用一部分精度，换取公开与可判错。可解释性研究面对的，正是同一笔交易的新版本——在模型这里，不成文的那一份太大了，成文能换回来的那一份，眼下还只是零头。因此，判断一项可解释性工作的价值，不该只问“它解释了多少”，还该问“它让哪一类错误变得可以被人指出来”。后者才是成文真正要买的东西。

◆ 五、这对科学意味着什么

现在可以回到全书的主命题。科学是在现实里微分、在理念里积分、再回到现实里再微分的一个圈。在牛顿以来的经典科学里，这个圈中段的积分是公开的：它写成方程，印在论文里，任何人都可以拿去推演、拿去挑错。积分公开，有三样好处。

第一，可传。一条方程可以教给下一代，不必每个人重新摸几万次脉。

第二，可判错。方程写在纸上，反常出现时，人们能指着它说“是这一项错了”。第16章讲的 123 原理——D 与 E 的矛盾累积，逼出 S 的改变，再回写 D 与 E——在科学史上之所以能一轮轮转下去，靠的就是反常能被定位。水星近日点多出那一点进动，人们知道该去怀疑引力定律的哪一部分。

第三，可迁移。写成文的积分可以搬到它从未见过的情境里去，因为人知道它成立的条件是什么，知道哪一步用到了“其他条件不变”。

当积分不成文时，这三样好处都打了折扣，又各自换来了一样新东西。

可传，变成了“可复制”：权重可以一字不差地拷贝给全世界，比任何教科书传得都快；但被拷贝走的是能力，不是理解，接手的人仍然不知道它为什么对。

可判错，变成了“可再训练”：模型在某类情境下系统地出错，人们未必能指着某一项说“是它错了”，常见的做法是补数据、再训练，让新的梯度把错误冲掉。这也是一种回写，而且常常管用；但它回写进的是另一张同样不成文的数表，而不是一条更好的方程。123　原理的循环还在转，只是转的地方从公开的理念界挪进了一个对人不透明的地方。反常还在被吸收，却越来越少被理解。

可迁移，变成了“看分布”：模型在与训练数据相似的情境里表现极好，离开得越远，越难事先知道它会不会失灵——因为它成立的条件没有写出来。导航软件在大路上几乎从不出错，到了一条刚改成单行道、地图还没更新的小巷，它就会信誓旦旦地让你逆行。

还有一层容易被忽视的回写，是落在人身上的。导航普及以后，许多司机不再记路，地图上没有的那些小巷，从人的经验里一条条消失了。这是负向回写的一个日常样子：一个外部的不成文积分好用到一定程度，人自己身上那份不成文积分——对街巷的手感——就不再被练习，慢慢退化。科学共同体也可能如此：当越来越多的预测交给一个不成文的积分去做，研究者自己“摸脉”的那份手感，可能也在悄悄变薄。是否真的如此，要被观察、被记录，不能被宣布；但它值得被列为一个待检的问题。

这一切并不推出“AI 对科学有害”或“不可解释的东西不算知识”。它推出的是一个更窄、也更可检验的判断：AI 把科学之圈中段的积分做得比人更快、更大，同时把这段积分从公开的理念界搬进了一张不可读的数表；圈的两端——入口处的数据从哪里来，出口处的落点由谁来补——反而因此更加吃重。

◆ 六、收束

老中医那句“摸多了就知道”，说的是一种真实的知识，也是一种锁住的知识。大模型把这种锁住的知识做到了前所未有的规模：千亿次梯度累加进权重，积分完成了，却没写成人能读的方程。可解释性研究想把它写成文，已经在局部写出了几行，离成篇还远。在它被写成文之前，科学要用它，就得格外看好圈的两端。下一章就走到出口那一端：AlphaFold、天气模型、材料发现这些最亮眼的成绩，替人做了哪一段，又把哪一段原封不动地留给了现场。

带 走 的 话

训练把千亿次梯度累加进权重，在数学形式上是一次积分；这个整体可维持、在功能上可识别，却对人基本不可表达，所以是“不成文的积分”。这不推出“AI 对科学有害”，也不推出“不可解释的不算知识”，只推出：圈中段的积分从公开的理念界搬进了不可读的数表，可传、可判错、可迁移都打了折扣，圈的两端因此更吃重。

小 账

- 本章老中医与徒弟、小铺掌柜、手机导航与小巷均为构想场景或比喻，无真实原型。
- 本章所引可解释性研究（Li 等 2023，Karvonen 2024，Anthropic 2023、2024、2025，奥拉等 2020）的出处见附录I第I40条；本章没有使用调查或统计数字。
- 本章所说“冻结态”只是缺少有效回写的功能性诊断，不是不可能性结论；所说负向回写是一个待检问题，不是已观察到的结论。
- 本章“不成文的积分”的推翻条件见附录G第G39条；负向回写的推翻条件见附录G第G41条。

27

第 27 章

AI for Science：替得了积分，替不了落点

> 这一章的问题：医生查指南就能定下用哪种抗生素，孩子每次喝几毫升却得当场算——AlphaFold、天气模型这些 AI 成绩，替人做的究竟是哪一段？

一位母亲带着发烧的孩子去社区医院。医生看了一眼化验单，说是细菌感染，要用抗生素。药名定了，医生却没有马上写剂量，而是先问：孩子多重？最近吃过什么药？对什么过敏？吃不吃得下药片？问完，他拿计算器按了几下，写下“每次几毫升，一天两次”。

药名，是成千上万次临床试验积分出来的结论：这一类细菌，对这一类药敏感。这份结论写在指南里，医生一查就有。可“每次几毫升”不在指南里，至少不在指南的正文里——它要在这一个孩子身上重新算一次：体重、肝肾、过敏史、能不能吞咽。母亲回家后还要盯着：孩子吐了没有，烧退了没有，第三天要不要复诊。药名是积分，剂量和那三天的观察是再微分。

本书第14章把这件事立为全书枢纽：应用必须在具体的点上再微分，而再微分要把积分常数 C 从现实里借回来。第26章讲到，AI 把科学之圈中段的积分做得又快又大，只是没写成文。本章看几件 AI 用于科学的代表性成绩——AlphaFold、机器学习天气模型、材料发现——逐一问同一个问题：它替人做了圈上的哪一段，又把哪一段留给了现场。

◆ 一、AlphaFold：一次漂亮的积分

蛋白质由一串氨基酸连成，这串珠子会在细胞里自己折叠成特定的三维形状，形状决定功能。从序列推结构，是困扰结构生物学半个世纪的难题。传统办法是做实验：X 射线晶体学、核磁共振、冷冻电镜，解一个结构往往要数月乃至数年。

2020 年，DeepMind 的 AlphaFold 2 在蛋白质结构预测竞赛 CASP14 上，对大量目标给出了接近实验精度的预测。此后 DeepMind 与欧洲生物信息研究所合作，把预测扩展到两亿多个已知蛋白序列，并公开数据库。哈萨比斯在 2024 年诺贝尔演讲中给出的数字是：两亿多个蛋白已被“折叠”，190 个国家的两百多万研究者用过它。2024 年他与约翰·江珀因此分享了诺贝尔化学奖的一半。

用本书的语言，AlphaFold 做了一件极其典型的积分。它的训练材料是蛋白质数据库里几十年间由实验一个个解出的结构——每一个结构，都是某个实验室在现实里截下的一刀：特定的温度、特定的结晶条件、特定的那一批蛋白。几十万次这样的截取被累加成一个整体：一种从序列到结构的映射。这个整体能推广到从未被实验解过的序列上，这正是积分的力量——第11章称之为“在理念里拼回整体”。

而且这个积分是不成文的。AlphaFold 没有给出一条“蛋白质折叠定律”，它给出的是一张网络和一个置信度分数。研究者知道它准，却很难说清它为什么准。第26章说的那种“记性极好却不会记账的掌柜”，在这里有了最耀眼的一位。

值得称道的是，这位掌柜会说自己哪里没把握。AlphaFold 对每一段预测都附一个置信度分数，研究者很快发现：分数低的地方，往往是蛋白里本来就没有固定形状的“无序区”，或者是要与别的分子结合才会成形的部分。换句话说，积分在它自己的边界上留下了记号——这里训练数据里的“截取”本来就少，或者这里的形状本来就取决于环境。这个置信度分数本身就是一份可贵的“成文”：它没有说出为什么，却说出了哪里不该太信。而那些低分区，恰恰是生物

学里最活跃、最依赖情境的地方——信号传递、调控开关，许多疾病相关的突变也在这里出没。积分最没把握的地方，往往就是 C 最重的地方。这并非巧合：一类对象的共同规律越弱、越依赖具体环境，从“一类”出发的积分就越难把它拼回来。

◆ 二、积分之后：药效、剂量、临床仍在病人身上

AlphaFold 给出了结构，然后呢?

第一步，结构是静态的。哈萨比斯本人在《代达罗斯》2026 年冬春合刊的对谈里说得很清楚：AlphaFold 2 预测的本质上是蛋白质的静态快照，而真实的生物学是动态的——蛋白会摆动、会与别的分子结合、会随环境改变构象。他说 AlphaFold 3 是“我们在梯子上的下一步”（our next step on the ladder），又把“虚拟细胞”当作更远的目标：在他看来，一个虚拟细胞对基础生物学和药物发现都将大有裨益。也就是说，在他本人的叙述里，结构预测是梯子的一级，不是终点。

第二步，从结构到药物，中间隔着化学、隔着生物、隔着人。知道一个靶点长什么样，可以帮助设计能卡进去的小分子；但那个分子在人体里能不能被吸收、会不会被肝脏迅速分解、会不会顺带卡进别的蛋白造成毒性，结构图上都看不出来。这些问题要在细胞上、动物身上、最后在人身上一项一项地试。

第三步，即便一种药被证明有效，它落到每一个病人身上时，仍要重新算剂量、看相互作用、盯不良反应。开头那位医生在计算器上按的那几下，没有任何结构预测能替他按掉。

波士顿咨询公司的贾亚通加（Jayatunga）等人 2024 年在《今日药物发现》（Drug Discovery Today）上做过一次初步统计：由 AI 辅助发现的候选分子，一期临床试验成功率约为八到九成，明显高于约四到六成五的历史平均；但到了二期——也就是第一次真正检验“对病人有没有用”的阶段——成功率约为四成，与行业历史水平相当。作者也说明，进入二期的样本还很少。如果这一观察站得住，它恰好落在本章的论点上：一期主要检验安全性与药物在体内的去向，

这些与分子本身的性质关系更密，积分能帮上大忙；二期开始问“对这一群病人有没有用”，这里有疾病的异质性、病人的个体差异、诊断标准的分歧——全是积分常数 C 的地盘。

用一个菜市场的比方：AlphaFold 像一位能一眼认出每一棵菜品种的老师傅，认得又快又准；可这棵菜炒给谁吃、放多少盐、那位顾客是不是忌口，认菜的本事管不到。认菜是了不起的本事，但做菜的人仍然得站在灶台前。

> **积分常数 C 为什么在落点处才付账** 设某个科学模型给出的是一类对象的变化规律 f，例如“这一类分子在这一类靶点上的结合倾向”。积分 $\int f$ 得到的是 $F + C$：F 是对整类对象都成立的部分，C 是对这一个具体对象才确定的部分。 在 AlphaFold 的情形里：F 是从序列到典型结构的映射；C 包括这个蛋白在这个细胞环境里的实际构象、与这个病人体内其他分子的相互作用、这个病人的代谢速度。 积分阶段可以不知道 C，因为它只需对一类对象成立。落点阶段必须知道 C，因为“此刻该怎么做”是对这一个对象而言的。AI 把 F 求得再准，也不改变一个形式事实：C 不在 F 里。第12章所说的复现危机集中出现在“情境依赖”与“初始条件”上，就是 C 在账面上的表现。

◆ 三、天气模型：最依赖初值的那门科学

天气预报是检验这一论点的极好场地，因为它是全部科学里对初始条件最敏感的一门。第22章讲过，洛伦兹正是在天气模型里发现了混沌，也讲过初值误差怎样被指数放大、集合预报怎样应运而生。本章不再重复那套道理，只问一件事：机器学习接手之后，C 从哪里来。

近几年，机器学习天气模型取得了惊人进展。DeepMind 的 GraphCast 在 2023 年发表于《科学》杂志：在 0.25 度分辨率下做 10 天全球预报，在研究者设定的 1380 个检验目标（变量与预报时效的组合）中，九成以上超过了欧洲中期天气预报中心的高分辨率确定性预报（HRES）；一次十天预报在一台谷歌 TPU v4 机器上不到一分钟，而 HRES 要在由数百台机器组成的超级计算机上算数小时。后续的 GenCast 进一步给出概率集合预报。哈萨比斯在莱克斯·弗里德曼（Lex

Fridman）的播客访谈中说，他们造出了“世界上最好的天气预报系统”（据播客官方网站公布的文字稿，原文是复数的 systems）。

这是一项真成绩，也是一次积分：模型从几十年的再分析资料里学到了大气从此刻到六小时之后的演变规律，再一步步滚动出十天。

但请看它的入口。GraphCast 每次预报，需要输入此刻和六小时前两个时刻的全球大气状态。这个“此刻的全球大气状态”从哪里来？它不是任何一颗卫星直接拍下的，而是“资料同化”的产物：把全球的气象站、探空气球、卫星、飞机、浮标的观测，与一个传统的物理数值模式结合，经过复杂的计算，估出一张最可能的、自洽的全球大气图。学界称之为“分析场”。

也就是说，机器学习天气模型替掉了传统数值模式的“积分”那一段——从此刻推到未来；却没有替掉“借 C”的那一段——估出此刻。它的初值，仍然来自遍布全球的观测网络与传统模式的同化系统。近年已有研究尝试连这一段也交给机器学习：剑桥大学与艾伦·图灵研究所等团队 2025 年在《自然》发表的 Aardvark Weather，直接吃进原始观测、输出预报，跳过传统同化；但它仍是研究系统，眼下的业务预报依旧依赖传统同化来定初值。即便如此，它也只是把“借 C”的手换了一只——C 仍然要从观测网络这块现实里借来。

再看它的出口。一张十天后的降水预报图落到某个县城的防汛指挥部，指挥部要决定今晚要不要转移河边的住户。这一决定要把预报与当地的河道、堤防、老人有没有手机、上一次转移时谁拒绝离开这些情况放在一起——这是在一个具体的点上再微分，而这些“C”没有一项在模型里。

GenCast 一类模型给出的是概率集合预报，其道理与第22章所讲的传统集合预报相同：此刻估不准，就把初值撒开一片，看结果散成多宽。机器学习让撒开几十份变得便宜，却没有改变这个结构。概率预报没有消灭 C，它把 C 的不确定摊开给用户看——至于七成的雨要不要取消明天的露天婚礼，仍要办婚礼的那家人自己拿主意。积分可以把风险算成数，算不出这家人能承受多大的风险。

天气的例子把 AI for Science 的结构照得特别清楚：圈的两端——估出此刻、决定此地——都牢牢扎在现实界，中间那一段积分被 AI 做得更快更好。就像

手机地图把“怎么从甲地到乙地”算得比任何老司机都准，但“你此刻在哪里”要靠卫星定位，“这条路今天封没封”要靠路上的人上报。

◆ 四、材料发现：积分的产出要回到实验台

第三个例子是材料。2023 年，DeepMind 在《自然》上发表 GNoME 项目，宣布用图神经网络预测出约两百二十万种新晶体结构，其中约三十八万种被判断为热力学稳定，并报告外部实验室已独立合成其中数百种。同期还有一项配套工作，用机器人自动化实验室依据预测去合成新材料。

这项工作发表后不久，固体化学家撰文提出质疑。安东尼·奇塔姆与拉姆·塞沙德里 2024 年在《材料化学》（Chemistry of Materials）上撰文指出：所列“稳定”化合物中很大一部分采用的是已知结构，或是已知化合物的平凡变体；计算在近乎绝对零度下进行，忽略了构型熵，而在实际合成温度下，许多预测的有序排布更可能以无序形式出现；这些化合物在显示出功能之前，还谈不上“材料”——他们的结论是，同时满足新颖、可信、有用三条的证据“寥寥”。另一组化学家（Leeman、Palgrave、Schoop 等，2024 年，《PRX 能源》）重新审视了自动化实验室公布的 X 射线衍射数据，认为其报告的“新合成”大多是把已知的无序化合物误判为预测的有序结构，在他们看来几乎没有一种可以确认是新材料。被质疑的一方也做了回应，争论的细节至今仍有分歧。

本书不裁判这场争论的胜负，只看它的形状。GNoME 的预测，本质上是在理念一侧——在第一性原理计算所定义的那个能量地形上——做了一次巨大的积分：它告诉你，在这张理想化的能量图上，哪些原子排布是低谷。争论发生在哪里？发生在“这个低谷在现实里能不能被做出来、做出来是不是那个东西、有没有用”。也就是发生在落点上。计算判稳定，用的是绝对零度、完美晶体的理想化模型；而真实的合成要面对温度、杂质、动力学路径、前驱体的选择——这些全是 C。

厨房里也有这样的事：一本菜谱用营养学原理算出一千道“理论上又香又健康”的新菜。这本菜谱当然有价值，它大大缩小了厨师试菜的范围；但哪一道真

做得出、做出来好不好吃、是不是跟老菜换了个名字，要厨师一道一道下锅才知道。下锅这一步，菜谱替不了。

◆ 五、C 的外借与现场责任

三个例子摆在一起，可以把本章的判断说得更确切。

AI 在这三处替人做的，都是科学之圈的中段：把大量已被测量、已被数学化的记录累加成一个能推广的整体。在这一段，AI 已经在若干领域做得比传统方法更快、更准，这是真成绩，不需要打折扣。

AI 在这三处没有替掉的，是圈的两端。入口端：训练所用的数据，是几十年实验、观测、标注的沉积；每一次使用时的初值，仍要从现实里估出来。出口端：每一次落点，都要把这一个病人、这一个县城、这一炉材料的具体情况借回来。

这里有一个容易被一笔带过的问题：借回 C 的那一步，责任落在谁身上？

在经典科学里，积分是公开的，落点由工程师、医生、技术员负责；他们出了错，可以对照公开的模型查出是模型错了还是落点时 C 借错了。当积分变得不成文，这条查错的路变窄了：一个用 AI 辅助的诊断出了错，究竟是模型在这类病人身上本就不准，还是医生输入时漏了关键信息，还是这个病人恰好落在训练数据很少覆盖的地方？三种可能都在，而且越来越难分开。

第395卷讲人—机器复合体时有一句要紧的话：能力分布不等于责任分布。AI 在复合体里贡献了大部分的计算能力，并不意味着它承担了相应的责任；责任仍落在那个在现场借 C、签字落地的人身上。可是，如果落地的人对积分的内容越来越不了解，他借 C 的能力也可能在退化——第26章提到的负向回写，在这里有了一个具体的样子：医生越来越依赖一个他看不透的判断，他自己在病人身上“摸脉”的那份手感就越来越少被练习。

还要看入口端的一笔账。AlphaFold 吃的是蛋白质数据库里半个世纪的实验结构，GraphCast 吃的是几十年的再分析资料，GNoME 吃的是材料数据库里海

量的第一性原理计算。这些数据不是从天上掉下来的，而是无数实验员、观测员在现实里一刀一刀截下来、再按统一格式记录在案的。AI 越成功，这些沉积就越值钱；可它们的生产方式——申请经费、守着仪器、一年解几个结构——并不因 AI 而变得更体面。如果一门学科把力气都转去用积分、而没人再去做新的截取，下一代模型就只能在旧沉积里打转。用菜市场的话说：大家都去学怎么挑菜，没人去种菜，菜摊迟早会空。这也是一种回写，它是正是负，要看共同体怎样安排资源，不能预先宣布。

所以本章的结论不是“AI 只能做积分”，而是一个带条件的判断：在目前已知的 AI for Science 成绩里，被替代的是积分段；落点上的再微分——借回 C、承担后果、并把后果回写给下一轮——仍须在现场完成，而且因为积分不成文，这一步比以往更吃重，不是更轻松。

◆ 六、收束

AlphaFold 认出了蛋白的形状，GraphCast 推出了十天后的云，GNoME 列出了几十万个可能的晶体。三件事都是真本事，都是人类积分能力的一次扩张。可那个发烧的孩子每次喝几毫升药，那个县城今晚转不转移，那一炉粉末烧出来是不是新东西——这些仍要有人站在现场，把 C 借回来，再微分一次。至此，第五编从混沌、工程、控制走到 AI，把“替得了积分，替不了落点”这个判断交给了下一编：哈萨比斯说的，恰恰是关于这个圈的另一种看法。在评估之前，先要把他到底说了什么弄清楚。

◆ 七、这一编怎样算错

1. 若出现这样一种 AI 系统：在没有现场初始条件输入（不依赖资料同化、不依赖病人个体信息）的情况下，仍能跨情境稳定完成落地决策，且在独立评估中没有系统性的情境性失败——则本编“再微分不可替代”的主张须收窄为“在当前配置下尚未被替代”。

2. 若对 AI 辅助研发药物的追踪表明，其二期及以后的临床成功率显著且稳定地高于传统方法，并且这一优势不能用靶点选择、适应症偏向等因素解释——则本编“积分帮得上前段、帮不上落点”的判断须修正：AI 已在一部分 C 上起作用，须说明是哪一部分。
3. 若可解释性研究能把一个前沿模型在某一科学领域里的主要判断，系统地提取为一组人能读、能检验、能迁移的规则，且这些规则通过减法检验——则第26章“不成文的积分”须改为“正在成文的积分”，相应论证降格为对某一阶段的描述。
4. 若长期追踪证明，深度使用 AI 辅助的工程师、医生在现场判断能力上没有退化，甚至提升——则本编关于“负向回写”的担忧须撤回，或限定到具体配置。
5. 若足够多的控制系统断开反馈后仍长期保持闭环表现，且不靠更准的模型或初值——则第24章“反馈就是持续借 C”的解读须收窄（与附录G第G38条相同）。
6. 若混沌系统的长期预测被证明可以在不依赖精确初值的条件下稳定做出，且不只是统计气候意义上的预测——则第22章“可计算≠可预测”的论断须重新限定其适用范围。

带 走 的 话

在目前已知的 AI for Science 成绩里，AI 替下的是科学之圈的中段积分：从序列到结构，从此刻到十天之后，从能量地形到候选晶体。圈的两端仍扎在现实里——初值要从观测网络估出，落点要把这一个病人、这一个县城、这一炉材料借回来。积分不成文，查错的路就变窄，借回 C 并承担后果的那一步因此更吃重。

小 账

- 本章发烧的孩子、县城防汛、露天婚礼、菜市场与菜谱均为构想场景或比喻，无真实原型。
- 本章所引哈萨比斯的诺贝尔演讲与《代达罗斯》对谈出处见附录I第I1、I3条，“梯子上的下一步”“虚拟细胞”“最好的天气预报系统”三处原话的核对见第I66、I67条；天气模型初值仍靠传统同化的限定见第I41条；对自动化实验室“新材料”的复核见第I42条；AI 辅助药物临床成功率出自贾亚通加等 2024 年的初步统计，样本尚少，见第I55条。
- 本章不主张“AI 只能做积分”，也不裁判 GNoME 争论的胜负。
- 本章主张的推翻条件见附录G第G4条（再微分不可替代）、第G40条（替不了落点）、第G41条（负向回写）；本编末第5条即第G38条。

第六编

评估与批判：哈萨比斯的论断

编序

第六编只做一件事：认真对待一句话。哈萨比斯说，自然中能被找到的模式都可以被学会，迄今没人发现不可计算的东西。这一编先还原它，再替它辩护，然后才动刀。

第28章核对他到底说了什么：诺贝尔演讲里的猜想，播客与对谈里的展开和限定，证据分三个等级，并看流传版“没有 AI 不能解决的问题”是怎样一步步长出来的。第29章替他把最强的辩护说足：自然不是随机的，能学而且能高效地学，成绩单摆在那里，P 与 NP 可以当成物理问题来问，他自己也留了余地；并写下批评者必须过的三道门槛。第30章是第一刀：可计算是筛子的性质，他所说的“自然”，是海，还是网捞上来的鱼？第31章是第二刀：“从未发现”同样可以读作数学每次碰壁就改尺码表的战绩，而已知的不可计算结果，恰恰出现在理念的无穷极限处。第32章是第三刀：问题是谁命名的，命名权、判错权、计算权三者之中，计算权排在最后，被筛掉的残差连“问题”的身份都拿不到。第33章分经验、概念、历史、外推、应用五层给出判词，并写明这些判词自己在什么情况下会错。

读这一编时，可以带着一个问题：一项成绩是真的，从它能推出多远？

28

第 28 章

他到底说了什么

> 这一章的问题：家族群里转来“诺奖得主说没有 AI 解决不了的问题”，哈萨比斯本人到底说过什么、没说过什么，又是在什么场合、带着哪些保留说的？

亲戚群里转来一条短视频，标题是：“诺奖得主：十年内，AI 能治好所有的病”。底下一片点赞，也有人留言：“那还上什么医学院。”一位在医院做护士的表妹回了一句：“他原话不是这么说的吧？”没人理她。过了两天，又有人转来另一条：“谷歌 AI 掌门人：没有 AI 解决不了的问题。”

表妹那句话问得对。一句话从说出口到传进家族群，中间要经过记者的标题、自媒体的剪辑、转发者的添油加醋。每经一手，限定词就掉一个，语气就加重一分。到了最后，一个科学家在特定场合、带着若干保留说出的猜想，变成了一句谁也没说过的口号。

本书第六编要评估德米斯·哈萨比斯关于“自然、计算与 AI”的论断。评估之前，第一件事是弄清他到底说了什么。这不是礼貌，是方法：批评一个没人说过的口号，赢了也不算数；批评一个人真正说过的话，才有可能学到东西。本章只做一件事——把原话、语境和他自己的限定摆出来，并把他本人的命题与流传版本分开。引外文原话只取短句，其余转述。

◆ 一、一个人与一张成绩单

先交代说话的人。哈萨比斯生于 1976 年的伦敦，少年时是国际象棋神童，后来做电子游戏设计，再读计算机、读认知神经科学博士，2010 年与人合创

DeepMind。这家公司 2014 年被谷歌收购，后与谷歌大脑合并为今天的 Google DeepMind，他任首席执行官。

他的成绩单上，有三件事与本书直接相关。一是 AlphaGo（2016 年击败李世石），那场比赛第二局的第37手，他在诺贝尔演讲幻灯片里列为 AlphaGo 想出“新颖的、有创造性的策略”的例子，在后来的播客里又说它“开始触及创造力的问题”。二是 AlphaFold，第27章已经讲过，他因此在 2024 年与约翰·江珀分享诺贝尔化学奖的一半。三是一系列“AI 用于科学”的项目：天气、材料、数学、核聚变控制。这三件事让他不只是一位企业家，也是一位用成绩说话的科学家；他关于自然与计算的说法，不是坐在书斋里想出来的，而是从这些成绩里长出来的。这一点在下一章替他辩护时很重要。

◆ 二、诺贝尔演讲里的那个猜想

本编要评估的核心命题，最早的正式出处是他 2024 年 12 月 8 日在斯德哥尔摩的诺贝尔演讲，题为《用 AI 加速科学发现》（Accelerating scientific discovery with AI）。演讲的大部分在讲 AlphaFold 的来龙去脉：两亿多个蛋白结构、190 个国家的两百多万研究者。临近结尾，他提出一个猜想，原文以“Any pattern that can be generated or found in nature” 起句，意思是：

> 凡是自然界能够生成或能够找到的模式，都可以被一种经典的学习算法高效地发现和建模。

这句话有几个词要逐一看清。

“模式”（pattern）——他说的是模式，不是“问题”，也不是“一切”。一个模式，是自然中反复出现、有规律的东西：蛋白如何折叠，大气如何流动，晶体如何排布。

“自然界能够生成或找到的”——这是限定范围。人造的谜题、纯数学里的构造、社会生活中的价值冲突，都不在这句话的字面范围内。

“经典的学习算法”——“经典”是相对于量子计算而言的。他说的是今天这类跑在普通芯片上的神经网络，也就是图灵机意义上的计算。

“高效地”——这是计算复杂性的用语，大意是所需的计算量不随问题规模爆炸式增长。

演讲中他还说，如果经典系统能模拟某些量子系统，这可能对计算复杂性理论——包括 P 是否等于 NP 的问题——“甚至基础物理”都有重大影响。在稍前的一页里，他把生物学在最基本的层面上看作一个信息处理系统，“尽管是一个极其复杂而涌现的系统”。（至于“信息比物质和能量更基本”这一说法，演讲幻灯片里没有，出自下文莱克斯·弗里德曼（Lex Fridman）的播客。）

到这里为止，这是一个有明确范围的猜想：关于自然模式，关于经典学习，关于效率。它是一个“猜想”，他自己就是这样称呼的。

◆ 三、播客里的展开：弗里德曼第475期

2025 年 7 月，哈萨比斯第二次上弗里德曼的播客（第475期）。主持人开场就引了诺贝尔演讲里的那个猜想，请他展开。播客官方网站公布了这一期的文字稿，它是本章最详细的材料来源。（网站注明文字稿由人工整理、可能有误，个别措辞以音频为准。）

他给出的理由，可以归成三层。

第一层：自然不是随机的。他说，自然系统之所以有结构，是因为它们经受过进化过程的塑造。这里的“进化”不只是生物进化：山的形状、行星的轨道、蛋白的折叠，都是在漫长的物理过程中“存活”下来的结果。不稳定的构型早被淘汰了，留下来的都有规律。他的原话里有一个短句：“nature’s not random”。因为不随机，所以可以被高效地重新发现。

第二层：可能存在一类新的复杂度类。他说自己和同事在想：是不是应该有一类问题，专门是这种神经网络过程能高效解决的。主持人顺口给它起了个名字叫“可学习的自然系统”，他没有接这个名字，但同意问题的方向。

第三层：P 与 NP 可能是一个物理问题。他说信息可能是宇宙最基本的单位；主持人接过话头：如果把宇宙看成一个信息系统，P 是否等于 NP 就成了一个物理问题。他表示同意。

同时，他在这一期里也明说了几处限定，要一并记下：

其一，量子。他说对那些没有可学结构的问题——例如大数分解——没有模型可学来帮你搜索，“那样的话你也许需要一台量子计算机”。

其二，混沌。他提到“初始条件真正起作用”的混沌系统，承认它们“可能很难建模”。

其三，理解的程度。谈到视频生成模型 Veo 能逼真地画出流体与材料，他说能连贯地预测下一帧，是一种理解；但他也说明，这不是人那种深层的理解，更像小孩子对物理的直觉。

其四，能力的现状。他说眼下 AI 带来的“更多是渐进的改进”，若要理解上的大飞跃，还需要大得多的进展；又说提出一个真正好的猜想，比证明它更难。他列举的“灯塔式”检验，是用截至 1900 年的全部文献训练一个系统，看它能否像爱因斯坦那样提出狭义与广义相对论，或者让它发明一种像围棋一样深的游戏。（这一检验在《代达罗斯》对谈里换成了“截止于 1910 年、推出广义相对论”的版本，见下文；本书后文提到“灯塔”时，两个年份各随其出处。）

这一期还有一处常被引用：他说他们造出了“世界上最好的天气预报系统”。这是一个关于具体产品的经验陈述，第27章已经讨论过它的分量。

◆ 四、DeepMind 播客：“迄今没人发现不可计算的东西”

另一句被广泛引用的话，出自 Google DeepMind 官方播客 2025 年 12 月 16 日发布的一期，题为《智能的未来》（The future of intelligence），由数学家汉娜·弗莱主持，是她与哈萨比斯的年度对谈。这句话的流传主要经由一篇 2026 年 1 月的网络评论文章《哈萨比斯在可计算性问题上错了》（作者 Tom Walczak）。据该文转述，哈萨比斯说的大意是：迄今还没有人在宇宙中发现任

何不可计算的东西；DeepMind 所做的事情，归结起来就是去找那个极限。该文脚注还转述了他的两句补充：也许宇宙中的一切，只要用对了角度看，都是计算上可处理的；以及他“搞量子计算的朋友们”会说有极限，量子系统需要量子计算机来算。

这里要标明证据的等级：本书核对到了这期节目的标题、发布方与发布日期（Google DeepMind 的视频频道同日上线，节目分段目录里在约 44 分钟处有一段以“图灵机”为题），但未能找到公开文字稿，也未能逐字核对音频中的原句；以上内容据转述。值得一提的是，那篇批评文章本身也承认：哈萨比斯是在提一个物理命题，而不是严格意义上的计算机科学命题。

请注意这句话的形状。“迄今没人发现”是一个历史陈述；“去找那个极限”是一个研究纲领。两者都不等于“没有极限”。相反，“去找那个极限”预设了极限可能存在——只是他认为它比大多数人想的远。

◆ 五、《代达罗斯》对谈：最审慎的版本

2026 年冬春合刊的《代达罗斯》（美国艺术与科学院院刊，第155卷第1—2期）刊出詹姆斯·马尼卡与哈萨比斯的对谈，题为《作为科学终极工具的 AI》（AI as the Ultimate Tool for Science）。“终极工具”一词也出自他本人，而且带着限定：他说自己一直把 AI 看作“也许是”推进知识前沿的终极工具（potentially the ultimate tool）；在弗里德曼的播客里，他也说过 AI 可以成为帮人回答这类根本问题的终极工具。这是本章材料中经过编辑、形式最正式的一份，也是他的说法最审慎的一个版本。

在这份对谈里，他把同一个思路说成：自然中的大多数事物都有一种稳定结构，这种结构是经过时间演化并存活下来的；因此应当存在某种可以被神经网络学到的模式或过程。注意这里的量词——“大多数”（most），而不是“一切”。

关于计算的界限，他说了一个短句：“I suspect P is not equal to NP”——他猜测 P 不等于 NP。他又说，有些问题若不借助例如量子计算机的帮助，可能是难以处理的。他甚至对现有量子理论是否完整也表示了疑问。

关于 AI 在科学中的位置，他说未来十年左右，应当把 AI 看作帮助科学家的了不起的工具；猜想和假说将来自人类专家；再往后，这些系统也许会更像合作者。他还说，提出正确的问题是科学中最难的部分。

关于 AGI 的检验，他提出：用知识截止于 1910 年的数据训练一个 AI，看它能否提出广义相对论。他自己的回答是：“For now, the answer is clearly no.”——目前，答案显然是否定的。关于规模扩张能否单独通向 AGI，他称之为一个开放的经验问题。

关于 AlphaFold，他说它预测的本质上是蛋白质的静态快照，并说精度“也许不完美”，但对实际用途已足够好。

◆ 六、对手是谁：彭罗斯的影子

要读懂一个论断，还要知道它在跟谁争。2026 年出版的塞巴斯蒂安·马拉比所著哈萨比斯传记《无限机器》（The Infinity Machine），据一篇网络书评（约翰·巴林顿发表于其 Substack 专栏）转述，把哈萨比斯的这一猜想放在与罗杰·彭罗斯的对照中来讲；本书未能核对传记原文。彭罗斯是 2020 年诺贝尔物理学奖得主，他长期主张：人的数学洞见不能被任何算法穷尽，背后可能有某种尚未被理解的量子引力过程，因此意识在原理上是不可计算的。

按书评的概括，哈萨比斯的赌注正押在另一边：心智的核心是找模式，而找模式这件事，经典计算机可以做得很远。据同一书评转述，他也承认，有些现象可能需要巨量计算，或者无法被经典方法高效还原，在那里“也许需要量子系统”。书评作者认为他低估了这道量子的缝，这是书评作者的看法，不是他的。

这一层对照很要紧。它说明哈萨比斯的命题首先是在“心智与计算”这场老争论里站队：他站在“不需要神秘的不可计算成分”这一边。按这层对照来读（这是本书的读法，不是他本人的交代），他那句据转述的“迄今没人发现不可

计算的东西”，矛头所指主要是彭罗斯式的“意识不可计算”论，而不是在宣称“一切人类的问题都能被算掉”。把这句话从那场争论里抽出来，放进“AI 万能”的语境，它的意思就整个变了。

用一个家常的例子：两个厨师争论“好味道是不是只能靠天赋”。一位说是，好舌头是天生的，学不来；另一位说不是，味道有规律，练得出来。后者说“我还没见过哪种味道是练不出来的”，是在反驳“只能靠天赋”，并不是在宣称“我能做出每一位客人都满意的菜”。前一句关乎味道能不能被学会，后一句关乎每一个具体的人想吃什么——那是另一个问题。

◆ 七、证据的三个等级

把本章用到的材料按可靠程度排一下，便于后面几章引用时有分寸：

第一等，经过编辑、有正式出版信息的文本：诺贝尔演讲幻灯片（诺贝尔基金会网站公开的演讲幻灯片）、《代达罗斯》对谈（有卷期、页码与数字对象标识）。这一等的话可以直接作为“他说过”。

第二等，公开的访谈文字稿：弗里德曼播客第475期，用的是播客官方网站公布的文字稿。内容可信，但口语即兴，措辞不如书面精确，网站也注明文字稿可能有误。引用时要照顾上下文，不能把他顺口接主持人的一句话当成深思熟虑的定论。

第三等，转述：DeepMind 播客那句“迄今没人发现不可计算的东西”，以及传记书评里关于量子与彭罗斯的说法。这一等只能写成“据转述”，并在原始音频或原书核对之前，不承担论证的主要重量。

这样分级，不是吹毛求疵。第六编的批评若主要压在第三等材料上，那么一旦原话与转述不符，整编就要返工。所以第30—32章的论证，应当以第一等材料中的猜想为主靶，第三等材料只作佐证。

◆ 八、流传版是怎样长出来的

现在可以把他本人的命题与流传版对照了。

> **本人命题与流传版的对照** 甲、本人命题（猜想）：凡自然界能生成或找到的模式，都可被经典学习算法高效发现与建模（诺贝尔演讲，2024）。在《代达罗斯》中收为：自然中的大多数事物有稳定结构，应当可被神经网络学到。 乙、本人的历史陈述（据转述）：迄今没人在宇宙中发现不可计算的东西。 丙、本人的限定：无可学结构的问题可能要量子计算机；混沌系统可能难以建模；他猜 P≠NP；今天的 AI 以渐进改进为主；未来十年 AI 是工具，猜想来自人；“1910 年数据推出广义相对论”目前做不到。 丁、另一句常被串联的话：在 2025 年 4 月美国 CBS《60 分钟》节目（斯科特·佩利采访）中，他先说“也许有一天，我们能在 AI 的帮助下治愈所有疾病”；主持人追问“疾病的终结？”，他答：“我觉得那是够得着的，也许在未来十年左右，我看不出为什么不能。”（据节目文字稿与当时报道；“终结疾病”这个说法出自主持人的追问，他接下了。）这是关于医学前景的乐观预测，不是关于可计算性的命题。 戊、流传版：“没有 AI 不能解决的问题。” 从甲到戊，至少发生了四次替换：“自然模式”换成了“问题”；“可被发现与建模”换成了“能解决”；“经典学习算法”换成了泛指的“AI”；“猜想”换成了断言，丙中的全部限定消失。丁则被当成了戊的佐证，尽管两者谈的不是一件事。

这四次替换，每一次都不是小事。

“模式”变成“问题”，跳过的是：谁来命名问题。自然中的模式是自然显露出来的；问题是人提出来的——“这个病怎么治”“这条河要不要筑坝”“这个孩子该不该留级”。问题里装着人的目的与价值，模式里没有。

“可建模”变成“能解决”，跳过的是：模型与落点之间的距离。第27章已经看到，能预测一个蛋白的结构，离治好一个病人还隔着药效、剂量与临床。

“经典学习算法”变成“AI”，跳过的是：他本人留给量子的那一块。

“猜想”变成“断言”，跳过的是：他自己说的“也许”“大多数”“目前显然不能”。

用一个家常的比方：一位老农说“这片地的土性，我种几年就能摸清”。传到村口，成了“他说什么地都能种”；传到镇上，成了“他说什么庄稼都能种活”；传到县里，成了“他说种地的事他全能解决”。老农的原话是关于“摸清土性”，而且限于“这片地”；最后的版本连“种不种”这个决定也算到了他头上。

◆ 九、同一个猜想的三种读法

即便只看第一等材料，那个猜想也可以有三种不同强度的读法，后面的评估要分开对付。

强读法：自然中每一个模式，都能被经典学习算法高效地学到。这是一个全称命题，一个反例就能推翻它。

中读法：自然中大多数有用的模式，能被经典学习算法高效地学到。这接近《代达罗斯》里的说法（那里是“大多数”，“有用”二字是本书加的），也接近他在访谈中实际的口气。它不怕个别反例，要看的是“大多数”与“有用”指的是什么、由谁来算。

弱读法，或者说纲领读法：我们应当假定自然模式可学，并按这个假定去做研究，直到撞上界限为止；DeepMind 做的事就是去找那个界限（这一说法据转述）。这不是一个能被证伪的命题，而是一种研究方向的押注。押注本身无所谓对错，只有成败与代价。

这三种读法在日常里也分得清。一个修车师傅说“车的毛病我都能查出来”，强读是“世上每一种故障我都能查出来”；中读是“常见的毛病我基本都能查”；弱读是“来一辆修一辆，查不出来的算我学艺不精，下回接着查”。对强读，一个反例就够；对中读，要看他查出来的比例；对弱读，只能看他这些年修好了多少车、又在哪些车上栽了跟头。

公道地说，诺贝尔演讲那句的字面措辞（“any pattern”）是强读；但他当场就称之为猜想，在《代达罗斯》里又收为“大多数”。本书倾向认为，他在不

同场合实际游走于中读与弱读之间，很少把强读当作已经成立的断言；而流传版恰恰是把强读当成了断言，还又往外推了一步。

◆ 十、本编评估的是哪一个

把材料摆清以后，本编的评估对象需要明确：

第一，第29章“最强辩护”，替他辩护的是甲、乙两条，加上他在丙中的限定——也就是他本人最强、也最审慎的版本。

第二，第30—32章的“三刀”，批评的主要对象也是甲与乙，而不是戊。对戊，本书只需一句话：它不是哈萨比斯的命题。但本书也会追问：为什么甲会如此顺滑地被传成戊？这种顺滑本身是否透露出甲在概念上留着一道缝——它在“模式”与“问题”之间、在“可建模”与“可落地”之间，没有划下足够清楚的界线？

第三，凡本书在后面几章提到“他说”，都指向本章核实到的出处；凡据转述的，照本章的标注处理。

这样做的好处，是让后面的批评有一个可以被检查的靶子。若读者认为本书曲解了他，可以拿本章的原话去对。

◆ 十一、收束

表妹那句“他原话不是这么说的吧”，问出的是任何评估的第一步。哈萨比斯真正说过的，是一个关于自然模式可学的猜想，加上一句“迄今没人发现不可计算之物”的历史陈述，以及量子、混沌、P≠NP、“目前显然不能”等一串他自己设下的限定。流传版“没有 AI 不能解决的问题”，是在四次替换之后长出来的。下一章要做的，是一件看似反常的事：在动刀之前，先替他把话说足——把他的猜想放到最有力的位置上，看它到底有多强。

带 走 的 话

哈萨比斯本人提出的是一个猜想：自然界能生成或找到的模式，可被经典学习算法高效发现与建模。他在《代达罗斯》里把它收为“大多数”，又亲自给量子、混沌、P≠NP 和“目前显然不能”设了限定。“迄今没人发现不可计算之物”只见于转述，不承担论证的主要重量。“没有 AI 不能解决的问题”不是他的话，批评它不等于批评他。

小 账

- 本章开头的家族群与表妹，以及厨师、老农、修车师傅，均为构想场景或比喻，不对应真实个案。
- 哈萨比斯言论的出处与核验状态见附录I第I1—I7条：诺贝尔演讲与《代达罗斯》对谈已对照原文，弗里德曼播客第475期据播客官方网站公布的文字稿，DeepMind 播客那句与传记的相关说法仅据转述；第37手、“终极工具”、刊物卷期与题目的核对见第I68、I69、I71、I72条。
- “三种读法”与“四次替换”是本书的分析框架，不是他本人的划分；本章不评判他的猜想对错，只确定评估对象。
- 若原话核定后与本书所批的命题不符，判词随原文修改，依据转述的论点作废，见附录G第G46条。

29

第 29 章

最强辩护

> 这一章的问题：河滩上的卵石为什么一眼就能看懂？哈萨比斯“自然模式可学”的猜想，在替他把话说到最足的时候，究竟有多强？

河边长大的孩子都知道，河滩上的石头是圆的。山上崩下来的石块棱角分明，滚进河里，被水冲、被别的石头磕，一年一年，棱角磨掉了，剩下的都是光滑的卵石。你若在河滩上随手捡十块，它们大小不一、颜色各异，可形状却出奇地相像：扁圆、两头略尖、表面光滑。为什么？因为那些形状不稳的——薄片会断，尖角会碎——早在流水里被淘汰了。留在河滩上的，是经得住流水的那几种形状。

一个从没见过这条河的人，只要给他看几百块卵石，他大概就能说出这条河里的石头“应该长什么样”，甚至能大致估出哪一段水急、哪一段水缓。他用不着重演几万年的冲刷，因为冲刷已经替他把可能的形状筛过一遍了。

这个画面，几乎就是哈萨比斯那个猜想的全部直觉。本章做一件评估者常常省略的事：在批评之前，替他把话说足。第395卷讲思想史的“接住／补齐”，要求先接住对方真正有力的东西，再说补齐什么；也要求自己不能靠宣布别人停在流沙上来取得立足点。这一章只接住，不批判。批判是第30章以后的事，而且它必须对付的是本章这个最强的版本，而不是第28章里那个流传版。还要说明一点：下面五根柱子里，注明出处的是他说过的话；其余的推演与补充，是本书替他补上的，不能算作他的原话。

◆ 一、第一根柱子：自然不是随机的

哈萨比斯最核心的理由，是一个关于自然本身的经验判断：自然系统之所以有结构，是因为它们经受了进化过程的塑造；所以自然“不是随机的”，它的模式可以被高效地重新发现。

把这句话说足，要看到“进化”在他那里用得很宽。生物进化只是其中一种：不适应环境的基因型被淘汰，留下的有规律。物理与化学也在做同样的事：不稳定的原子核衰变了，不稳定的分子分解了，不稳定的轨道上的天体撞掉了或被甩出去了。宇宙一百多亿年的历史，就是一部持续不断的“淘汰不稳定构型”的历史。今天我们能看到、能测到的一切，都是这场淘汰的幸存者。

这个判断有一层很深的道理：可能性的空间极大，实际存在的东西却只占其中极小、极有规律的一角。一副扑克牌洗一次，可能的排列（约 8×10^{67} 种）比整个地球上的原子还多；可是一位老牌手看几十局就知道对手的习惯，因为打牌的人不是随机出牌的。自然也一样：原子可能的排布方式多到无法枚举，实际出现的晶体却集中在为数不多的几百种对称类型里；氨基酸可能的序列多到无法枚举，实际存在的蛋白质折叠方式，也只有一千余种基本类型——结构生物学家乔西亚（Cyrus Chothia）1992 年在《自然》上发表《给分子生物学家的一千个家族》，估计自然界的蛋白质家族不过一千左右；三十年后的结构分类数据库 SCOPe（2.08 版）收录的折叠类型为 1257 种，与这个估计相去不远。

这一点，本书自己的理论其实是同意的。第2章讲“差异先于实体：特征怎样稳下来”，第395卷所说的特征律，正是讲差异怎样稳定下来，凝成可被指向的同一性。一个特征之所以是特征，就在于它“可维持”——明天再看还在。哈萨比斯说自然有可学的结构，用本书的话说，就是：自然界里大量的差异已经稳定成了特征，而稳定的东西有规律可循。在这一层上，他与本书站在同一边。

◆ 二、第二根柱子：能学，而且能高效地学

自然有规律，并不自动等于能被高效地学到。这里需要第二根柱子，而它最好的证据，是一个半世纪前就被提出的悖论。

> **列文索尔悖论：自然自己走了捷径** 1969 年前后，分子生物学家列文索尔指出：一条中等长度的蛋白质，假设每个氨基酸残基只有几种可能的局部构象，整条链可能的构象数就是天文数字。若蛋白在这些构象之间随机试探、逐一尝试，找到正确折叠所需的时间将远超宇宙年龄。可实际上，蛋白在细胞里折叠只要毫秒到秒的量级。 粗略估算：设有 $n = 100$ 个残基，每个残基 3 种构象，总数为 $3^{100} \approx 5 \times 10^{47}$；即使每秒试 10^{13} 次，穷举也要约 10^{27} 年量级（具体数字随假设而变，此处只示量级）。 结论不是“蛋白折叠不可能”，而是“蛋白不是在穷举”：能量地形有漏斗形状，局部的正确接触会引导整体走向正确结构。自然自己找到了捷径。 哈萨比斯在诺贝尔演讲幻灯片第10页亲自引了这一悖论（幻灯片给出的估计是：一个典型蛋白可能的构象约有 10^{300} 种）。下面的接续是本书替他重建的：既然自然能在多项式般的时间里“算出”结果，这个过程里就有可学的结构；一个学习系统若能学到这个漏斗的形状，就能同样高效地给出答案。AlphaFold 的成功，可以看作这一论证的一次实证。

这一论证的力量在于，它把“可学”从一种乐观的信念，变成了一个有具体机制支撑的推断：自然本身在解一个看似极难的搜索问题，而且解得很快；自然能快，说明问题的实际结构远比最坏情况简单；学习系统所学的，正是这个实际结构。

用家常的话说：一个城市的街道可以有无数种走法，但出租车司机不会去穷举，因为城市不是迷宫，它有主干道、有单行线、有早晚高峰的规律。一个新司机跟车跑几个月，就能学到老司机的大部分路感。路感之所以学得会，是因为城市本身是“被规划过的”。自然没有规划者，但淘汰替它做了类似的工作。

在这个基础上，他提出也许应当有一类新的复杂度类：一类问题，专门适合这种从数据中学习结构、再用学到的模型引导搜索的方法去解。计算复杂性理论

传统上按“最坏情况”给问题分类：一个问题只要存在极难的实例，就被归入难的一类。可自然界给出的实例，可能从来不是最坏情况。若真能刻画“自然给出的实例”这一类，许多被判为“难”的问题，在实际中也许并不难。这是一个严肃的、可以被数学化的研究方向，不是一句口号。

◆ 三、第三根柱子：成绩单

一个猜想的分量，要看它背后的成绩。哈萨比斯的猜想不是空想，它背后有一串实打实的结果，每一个都可以被看作这个猜想的局部验证。

蛋白质结构：AlphaFold 在 CASP14 上达到接近实验精度，并扩展到两亿多个已知蛋白（见第27章）。这是“自然模式可学”最响亮的一次证明。

天气：GraphCast 一类模型在中期预报的多数指标上超过了基于物理方程的传统数值模式，计算量却小了几个数量级（见第27章）。天气本是混沌系统的代名词，这一成绩表明，即便在混沌系统里，在几天到十来天的尺度上，仍有大量可学的规律。

围棋：AlphaGo 的第37手是人类职业棋手几乎不会下的一步，事后却被证明极有远见；他在诺贝尔演讲幻灯片第3页把它列为 AlphaGo 想出“新颖的、有创造性的策略”的例子。它表明，一个学习系统不只能复现人类已知的模式，还能找到人类没找到的模式。

数学与算法：AlphaTensor（Fawzi 等，2022 年 10 月发表于《自然》）找到了某些规模下比已知方法更少乘法次数的矩阵乘法算法，例如在二元域上做 4×4 矩阵乘法只需 47 次乘法，少于反复套用施特拉森算法所需的 49 次；FunSearch（Romera-Paredes 等，2023 年 12 月在线发表、2024 年刊于《自然》）在帽集（cap set）问题上给出了新的构造，在 8 维情形找到了大小为 512 的帽集，超过此前已知的最好结果，并改进了帽集容量的下界。这说明“可学”不只限于物理与生物，也延伸到了数学结构的搜索。

核聚变：DeepMind 与瑞士洛桑联邦理工学院合作，用强化学习控制托卡马克装置中的等离子体形状，在 2022 年 2 月发表于《自然》（Degrave 等，《通

过深度强化学习对托卡马克等离子体进行磁控制》）。这是一个控制问题，也是一个高度非线性的物理系统。

视频模型：他在访谈中提到，视频生成模型从大量视频中“被动观察”，就学会了流体、材料大体上该怎样运动。他自己强调这不是深层的理解，但它提示：物理规律中的相当一部分，可以不经方程、仅从观察中被学到。

这一串成绩有一个共同点：每一项都发生在过去被认为需要专门理论、专门计算、或专门人类直觉的地方。若说“自然模式可被高效学到”只是一个猜想，那么它至少是一个被反复局部验证、迄今尚未遭遇决定性反例的猜想。

◆ 四、第四根柱子：把 P 与 NP 当成物理问题

哈萨比斯思想里最大胆、也最容易被误解的一步，是同意“P 是否等于 NP 可能是一个物理问题”。把这一步说足，要回到一个在物理学与计算机科学交界处有悠久传统的思路。

> **物理的丘奇—图灵论题**　原始的丘奇—图灵论题是关于“有效计算”这一直观概念的：凡是一个人按规则一步步能算的，图灵机都能算。这是一个关于数学概念的论题。　后来有人把它推进为一个关于物理世界的论题：凡是物理过程能“算出”的，图灵机都能模拟（物理版）；进一步，凡是物理过程能高效算出的，概率图灵机都能高效模拟（强物理版，量子计算对它构成公认的挑战）。物理学家多伊奇在 1985 年的论文中，正是从这一方向出发提出了量子图灵机的构想。　哈萨比斯的立场可以放在这条线上理解：如果信息是宇宙的基本单位，物理过程就是信息处理过程；那么“什么能被高效计算”就不只是数学家的问题，也是一个关于宇宙实际如何运作的问题。在这个意义上，P 与 NP 的关系可能有物理内容。他本人猜测 P 不等于 NP（《代达罗斯》对谈）。

这一步的力量，在于它拒绝把计算看成一个与世界无关的纯形式游戏。一台计算机本身就是一块物理设备：电子在芯片里流动，热量从散热片散出。计算能做到什么，最终受物理定律约束。反过来，若物理世界的每一个过程都能被计算机高效模拟，那么“世界是什么”与“计算能做什么”就紧紧缠在了一起。物理

学家惠勒有一句著名的短语“it from bit”（万物源于比特），表达的正是这种把信息放在物质之前的直觉。哈萨比斯并非第一个这样想的人，但少有人像他这样手握 AlphaFold 这样的成绩来这样想。

在这一框架里，他那句“迄今没人在宇宙中发现不可计算的东西”（据转述）就有了明确含义：在实验室与观测站里，人类迄今没有发现任何一个物理过程，其结果被证明无法被计算机模拟出来。这是一个关于经验记录的陈述，就目前所知，它大体上站得住：从行星运动到粒子碰撞，从天气到蛋白质，物理学家的日常工作正是用计算机去模拟它们。

◆ 五、第五根柱子：他自己留了余地

最后一根柱子，是他本人的审慎。最强的辩护，不是把对手的话说得最满，而是把他自己设下的限定也算进去——因为这些限定让他的猜想更难被驳倒。

他承认量子：对那些没有可学结构的问题，比如大数分解，他说也许需要量子计算机；他甚至对现有量子理论是否完整表示了疑问。这意味着他的猜想从一开始就没有覆盖全部计算问题，它覆盖的是“自然中有结构的那一大块”。

他承认混沌：初始条件真正起作用的系统可能很难建模。这意味着他并不认为“可学”就等于“可长期预测”。

他承认现状：今天的 AI 以渐进改进为主；让 AI 从 1910 年的知识推出广义相对论，目前显然做不到；未来十年，AI 是帮助科学家的工具，猜想和假说来自人类专家；提出正确的问题是科学中最难的部分。

他承认这是一个猜想：诺贝尔演讲里，他用的正是“猜想”这个词；在 DeepMind 播客里，据转述，他把公司的工作说成是去“找那个极限”。一个说“我去找极限”的人，并没有说极限不存在。

把这五根柱子立在一起，他的立场可以这样写成最强的形式：

自然经过长期的物理与进化淘汰，绝大多数存留下来的模式都有稳定结构；这些结构使得自然界给出的问题实例远比最坏情况简单；因此，经典学习系统有

望高效地学到其中大部分；这一点已被蛋白质、天气、围棋、数学、等离子体控制等一系列成绩局部证实；若这一点普遍成立，它将意味着“可高效计算”是关于宇宙的一个物理事实，而非仅是关于形式系统的数学事实；量子系统、无结构问题、混沌系统的长期行为是已知的可能例外；眼下 AI 的角色是科学家的工具，提出问题仍是人的事。

这是一个强有力的、有经验支撑的、自带边界的科学纲领。它远比“没有 AI 不能解决的问题”精细，也远比后者难以反驳。

◆ 六、替他预先回答本书

最强的辩护还要再走一步：设想他读了本书前几编，会怎样回答。把这几条回答也替他说出来，第30章以后的批评才不至于打在稻草人身上。

本书说，AI 学到的是已被测量、已被数学化的数据，不是物理界本身。他大概会回答：这正是全部科学的处境。牛顿学到的也不是“物理界本身”，而是第谷和开普勒记下的行星位置；任何科学家都只能通过测量接触世界。若这算缺陷，它是科学的缺陷，不是 AI 的缺陷。况且，AI 正在扩大测量本身：自动化实验室能日夜不停地合成、表征新材料；机器人可以在人去不了的地方采样。测量的边界在 AI 的帮助下向外推，“被筛掉的那一份”会越来越少。

本书说，积分常数 C 必须在落点处从现实借回，AI 替不了落点。他大概会回答：借 C 本身也是一种可学的技能。天气预报里的资料同化，正在被机器学习方法逐步改进；医疗里的个体化剂量，可以从大量病人的记录中学出规律；机器人在真实环境里的控制，正是在现场一边感知一边调整。今天的落点要靠人，不等于明天的落点仍要靠人；他自己说的是“未来十年左右”AI 是工具，并没有说永远是工具。

本书说，AI 的积分不成文，对人不可读。他大概会回答：可读性是人的要求，不是自然的要求。自然本身也不“写成文”——蛋白质折叠时并不先写下一条定律。若一个系统能稳定、准确地预测，并通过了独立检验，那么它“理

解”了多少，可以留给哲学家去争；对一个要治病的病人来说，药有没有效才是要紧的。何况可解释性研究也在进步，DeepMind 自己就在做。

本书将说，问题是人命名的，AI 没有命名权。他大概会回答：他从未声称 AI 现在能提出好问题；他明说“提出正确的问题是科学中最难的部分”，并且把“发明一个像围棋一样深的游戏”、“从 1910 年的知识推出相对论”当作尚未通过的检验。这恰恰说明他清楚地知道命名是一道门槛，并且把跨过这道门槛设为研究目标，而不是假装已经跨过。

这四条回答，每一条都有分量。它们不在本章接受审查；本章只是把它们摆上桌，让后面的批评必须正面对付它们。一位好的辩护律师，不会只替当事人陈述事实，还会预先回应对方律师可能提出的每一项指控。这样，法庭上的交锋才有意义。

◆ 七、批评者要过的三道门槛

既然替他把话说足了，就该同时写下：任何想反驳他的人，至少要过哪几道门槛。这也是本书给自己的第六编立的规矩。

第一道，不能拿流传版当靶子。凡是批评“没有 AI 不能解决的问题”的论证，对哈萨比斯本人一概不构成反驳。批评必须指向第28章核实过的那几句话。

第二道，不能拿他自己已承认的例外当反驳。指出“量子系统也许需要量子计算机”“混沌系统长期难以预测”“今天的 AI 还提不出相对论”，都只是在重复他的话。批评必须找到他的限定没有覆盖到的地方。

第三道，不能否认成绩。AlphaFold 与天气模型的成绩是被独立检验过的；任何把它们说成“只是拟合”“只是统计”的批评，都要先说明：为什么同样的话不能用来贬低牛顿对开普勒数据的拟合。

过不了这三道门槛的批评，再响亮也是自说自话。第30—32章的三刀，要逐一接受这三道门槛的检查；第33章的判词，还要写明自己在什么条件下算错。这

就像菜市场里退货：说“这菜不新鲜”不够，得指出是哪一棵、哪里坏了，而且不能是卖菜的人事先已经提醒过你的那一处。

◆ 八、收束

河滩上的卵石之所以能被一眼看懂，是因为流水已经替我们筛过了。哈萨比斯的全部论证，可以收进这一个画面里：自然不是随机的，淘汰替学习者做了大半的工作；AlphaFold 和天气模型证明了这一点；如果这一点普遍成立，可计算性就成了一个关于世界的物理事实；而他自己也给量子、混沌和人的提问留了位置。这个版本值得认真对待。正因为它强，下一章的第一刀才必须找准位置：问题也许不在“自然是否有可学的结构”，而在“可学”这个词所指的，究竟是哪一个世界——是流水中的河滩，还是被捡起来、洗干净、摆进标本柜里的那几百块卵石。

带走的话

在说得最足的形态下，哈萨比斯的猜想是一个有经验支撑、自带边界的科学纲领：自然经长期淘汰而多有稳定结构，自然给出的实例远比最坏情况简单，AlphaFold、天气模型等成绩局部证实了这一点；他本人给量子、混沌和人的提问留了位置。批评若打在流传版上、只重复他已承认的例外、或否认被独立检验过的成绩，都不算反驳。

小账

- 本章的卵石、扑克牌、出租车司机、退货等均为比喻；“替他预先回答本书”的四段回答是本书的设想，不是他的原话。
- 本章五根柱子中注明出处的说法，依据见第28章与附录I第I1—I6条；他在演讲幻灯片里引列文索尔悖论、举第37手为例，核对见第I70、I68条，方框里的估算只示量级；乔西亚、SCOPe、AlphaTensor、FunSearch 的数字见附录I第I47、I48条。
- 本章只接住、不批判；替他把话说足，不表示本书同意“可高效计算是关于宇宙的物理事实”。
- 本章立下的“三道门槛”对后面各章的约束，见附录G第G47条。

30

第 30 章

第一刀：可计算是筛子的性质

> 这一章的问题：鱼铺老板的本子算得清每一斤上过秤的鱼，他能不能因此说“海里的东西我都能算”？哈萨比斯说的“自然”，是哪一个自然？

菜市场东头有一家收鱼的铺子，老板只认秤。渔船靠岸，他把鱼一筐一筐倒上电子秤，屏幕跳出数字，按斤论价，记进本子。一天下来，本子上密密麻麻全是数，加起来清清楚楚：今天收了多少斤，花了多少钱，哪条船的货最多。老板的账从来不出错。

可那天有个老渔民拎来半桶东西：几只刚蜕壳、软得拿不住的蟹，一把还在跳的小虾，一团缠在网上的海草，还有一句话——“今早南边的水颜色不对，怕是要起风”。老板摆摆手：秤不了，不收。老渔民也不恼，把桶拎回去了。第二天果然起风，三条船没回来。

老板的本子上，这一天和别的日子没什么两样，数字照样整齐。他若有一天说“凡是海里有的东西，我的本子都能记、都能算”，你很难驳他：凡是上过秤的，他确实都算得清。问题不在算得准不准，而在那句话里的“凡是海里有的”。他的本子里装的，从来只是“上过秤的”。

第28章已经摆出哈萨比斯的原话，第29章已经替他把理由说足。本章开始下刀。第一刀砍的不是他的结论，而是他的结论所站的那块地：他说的“自然”，是哪一个自然？

◆ 一、他的“自然”是哪一个自然

先把要砍的命题写准。哈萨比斯在诺贝尔演讲里提出、此后在访谈中反复申说的，是一个猜想：自然中能被生成或找到的模式，都能被经典学习算法高效地发现和建模（原话与出处见第28章）。他的理由是，自然不是随机的，它经受了进化与物理过程的长期塑造，留下来的是“存活”下来的东西（本书把这一层概括为“筛选”，这不是他的原词），所以有结构，有结构就能学。

这个猜想有两层。一层是经验层：大量自然模式确实可学，AlphaFold 学到了蛋白质折叠的规律，天气模型学到了大气演化的规律，这些是真成绩，本书不打折扣（见第27章）。另一层是这句话没有明说、却必须预设的东西：“自然中的模式”是一个现成的、摆在那里等待被学的对象集合。第一刀就砍在这个预设上。

本书第3章已经区分了两个“自然”。一个是物理界：流动的、纠缠的、没有被谁切过的现实，鱼在水里游，风在海上起，蟹在蜕壳。另一个是理念物理界：物理界经过测量、理想化、剔除噪声之后剩下的那部分——它是科学家口中的“自然界”，是写进教科书、存进数据库、可以被公式描述的那个自然。连接两者的是一张筛子，筛子就是数学化：测量把连续的现实变成数值，理想化把数值变成变量，模型把变量变成方程。

哈萨比斯说的“自然中的模式”，落到他实际做事的地方，指的是哪一个？看他的证据就知道。AlphaFold 学的不是蛋白质本身，是蛋白质数据库里几十年积累的实验结构——每一个结构都是 X 射线晶体衍射、核磁共振或冷冻电镜测出来、经过精修、写成原子坐标的数据。天气模型学的不是大气本身，是再分析资料——观测站、卫星、探空气球的读数，经过数据同化，插值成规则网格上的温度、气压、风速。视频模型从“被动观察”里学物理，它看到的也不是世界，是像素：光被镜头聚焦、被传感器量化、被编码压缩之后的阵列。

这三样东西有一个共同点：它们都已经过了秤。它们是理念物理界里的居民，不是物理界的原住民。

大白话说：鱼铺老板的本子能把“上过秤的鱼”算得清清楚楚，这是本事；但他若因此说“海里的东西我都能算”，就把“上过秤的”偷换成了“海里有的”。

◆ 二、爱丁顿的渔夫

这个偷换不是哈萨比斯的发明，它是近代科学的老习惯。爱丁顿那位用两英寸网眼捕鱼、宣布“海里没有短于两英寸的生物”的学者，导论已经讲过，这里不复述。导论用它提问；本章要用它做一件更具体的事：学习算法是一张网，而且是一张建在另一张网之上的网。第一张网是测量与数学化，它决定什么能进入数据；第二张网是学习算法的归纳偏置，它决定数据里什么能被学成模式。哈萨比斯的猜想讲的是第二张网的能力：凡是第一张网捞上来的东西里有模式，第二张网都能高效学到。这个猜想也许是对的，甚至很可能在相当大的范围内是对的。但它对第一张网没说一个字，而“自然中能被生成或找到的模式”这个说法，恰恰把第一张网的边界抹掉了。

> **两张网的叠加** 设物理界为 W，测量—数学化映射为 $M : W \to D$，其中 D 是数据空间（有限精度的数值、坐标、像素）；学习算法在 D 上拟合一个假设 $h \in \mathcal{H}$。哈萨比斯猜想的可检验形式大致是：对 D 中一大类“非随机”的分布，存在多项式资源的经典算法使 h 逼近其生成规律。 这一形式的全部量词都落在 D 和 $\mathcal{H}$ 上。M 的像集 $M(W)$ 之外的东西，即 W 中被 M 丢弃或无法编码的部分，不在任何量词的辖域之内。 因此，“自然中的一切模式可学”若要成立，至少还需第二个命题：M 不丢失与问题相关的结构。这个命题不是学习理论的命题，而是关于测量与数学化的命题，学习理论既不能证明它，也不能证伪它。

这个方框不是在说哈萨比斯算错了什么。他的猜想作为学习理论的猜想，可以在数据空间里被严肃讨论，第29章替他把这一层说得很足。方框指出的是：从“数据空间里的模式可学”到“自然中的模式可学”，中间隔着一道 M。他的表述把 M 当成了透明的玻璃，而本书认为它是一张有网眼的网。

◆ 三、信息九宫：可计算只对哪一格有定义

为什么说 M 不透明？这要回到第4章的信息九宫。第395卷的现行口径是：每一界都有信息——现实界、理念界、自我界各有其信息；每一界的信息又有三种模态——符号态、逻辑态、数学态。三界乘三态，九格。

老渔民那半桶东西，正好分布在不同的格子里。“水颜色不对”是现实界的符号态：一个饱满的、直接被认出的特征，说不出公式，但一眼就知道。“这种颜色之后常起风”是现实界的逻辑态：他几十年出海，把颜色、风向、云脚、鱼群的走向缠成了一条本地的推理链，这条链有对错，能被下一次出海检验，但没有被提取成可以写下来教给陌生人的理念逻辑。至于数学态，在这件事上它可以有——气象学可以把海水颜色换成叶绿素浓度或悬浮物浓度，把风换成气压梯度，写进方程。但请注意，一旦写进方程，老渔民的那条逻辑链就被换掉了：方程里没有“他看了几十年”的那一份，只有“某波段反射率”。

第4章已经论证过，可计算性是一个有论域的谓词：只有输入、输出、对应规则都被精确规定之后，才能问一个对象是否可计算。对还停在符号态的“颜色不对”问“可不可计算”，就像问一首歌“有多少公斤”——不是答案为否，而是问题没有定义。第4章把这一点说成可计算性理论的自我边界；本章要把它变成一把指向哈萨比斯的刀。

所以第一刀的核心命题是：**可计算性只对信息九宫中上升到数学态的那些格子有定义；而且在本书的读法里，严格说只对理念界的数学态有定义。** 现实界的数学态是借来的说法，它指的是现实被测量、被赋值之后，以理念数学的形式被表示出来的那一份；它“可计算”，是因为它已经被翻译进了理念界。说“物理界是可计算的”或“物理界是不可计算的”，都犯了同一个错：把一个只在理念层有定义的谓词，安到了理念层之外。

大白话说：说“这道菜可计算”，意思只能是“这道菜的菜谱、配比、火候曲线可以被写成程序”；菜本身在锅里，它既不可计算也不是不可计算，它在被吃。

这一点也可以替哈萨比斯想一想。他在访谈中同意主持人的说法：P 与 NP 可能是一个物理问题（见第28章），本书认为这个方向有启发：它逼人去问，计算复杂性的界限是否受物理定律约束。但“受约束”和“是物理界的性质”是两回事。物理定律本身已经是理念物理界的居民；问“物理定律允许多快的计算”，是在理念层内部，拿一个数学模型去约束另一个数学模型。这个问题很好，但它从头到尾没有离开筛子之上。

◆ 四、“筛选”是谁的筛选

哈萨比斯为他的猜想给过一个漂亮的理由，第29章替他展开过：自然不是随机的；宇宙里留存下来的结构，都是经过进化、经过物理过程长期“筛选”的幸存者；幸存者有稳定的结构，稳定的结构就能被高效地逆向工程。他甚至设想过，也许该有一类专属于这种学习过程的复杂度类（“可学习的自然系统”这个名字是访谈主持人起的，见第28章）。

这个理由里藏着两次筛选，他只说了一次。

第一次是自然的筛选：不稳定的分子分解了，不适应的物种灭绝了，剩下的东西确有结构。这一次，本书承认。

第二次是人的筛选：在所有现存之物里，科学先挑那些稳定到可以被反复测量的东西去测。一个蛋白质要能结晶才能被衍射测出结构；一个天气现象要在足够大的尺度上重复出现，才能被网格捕到；一种行为要在实验室里可重复，才能进入心理学的数据集。第395卷说，S 要有“三个可”——可识别、可维持、可表达；一个特征必须先稳下来，才能显露为可以指认的对象。科学选择测量对象时，偏偏就选那些“三个可”最充足的东西。

于是一个循环出现了。哈萨比斯说：自然中的模式之所以可学，是因为它们稳定。本书说：进入数据的模式之所以稳定，有一部分是因为不稳定的东西从一开始就没有进入数据。前一句说的是自然的性质，后一句说的是筛子的性质。他把两次筛选的成绩，都记在了第一次的账上。

大白话说：一所重点中学说“我们的学生个个成绩稳定，可见本地孩子都很好教”。它没提的是，招生考试已经把成绩不稳定的孩子筛在了门外。孩子们的稳定，一部分是孩子的，一部分是门槛的。

这并不推翻他的理由。自然的筛选是真的，第二次筛选也不是作弊——不先挑稳定的东西测，科学根本起不了步。但它要求他的猜想收窄一句：被测量、被数学化的那部分自然，稳定到足以被高效学习；至于没有被测的部分，它的稳定与否，学习的成功不能作证。一个“可学习的自然系统”的复杂度类若要定义，它的输入也只能是数据，而数据已经是第二次筛选的产物。

◆ 五、“可学”的对象是被测过的数据

把第一刀落到他最有力的证据上。AlphaFold 是这个时代最漂亮的科学成就之一。它证明的是什么？它证明，在蛋白质结构数据库这片已经数学化的土地上，序列到结构的映射有足够的规律，可以被一个深度网络高效学到，而且能推广到没测过的序列上。这是一个关于理念物理界的、极强的经验结论。

它没有证明的是：蛋白质在细胞里的全部行为都在这片土地上。实验结构多是在晶体里、低温下、纯化后测的；细胞里的蛋白质在拥挤、流动、与成百上千种分子纠缠的环境中，有一部分天然无序区根本不折叠成固定结构，有一部分功能取决于什么时候、和谁结合、被怎样修饰。结构生物学界对此并无隐讳，AlphaFold 团队自己也提示过对无序区的预测要谨慎：Tunyasuvunakool 等 2021 年在《自然》上发表的人类蛋白质组预测论文，专门把低置信度区段当作“可能无序”的区域来标出；欧洲生物信息研究所（EMBL-EBI）的 AlphaFold 官方教程也说明，pLDDT 低于 50 的区段通常意味着该处天然高度柔性或内在无序，而对那些只在结合或修饰后才折叠的区段，AlphaFold2 倾向于给出其折叠态，即便天然状态下它并不如此。这不是 AlphaFold 的失败，这是筛子的网眼：晶体衍射这张网，捕到的是能结晶的、静止的那一份。

再看天气。数值天气预报与 AI 天气模型的进步，是理念物理界中一场接一场的胜仗。但天气模型学的再分析资料，在海洋上、在高原上、在没有观测站的地

方，本身就是数据同化模型“补”出来的。模型学到的是“被测量并被补齐之后的大气”。它预报的准确度提高了，这是真的；而“大气中一切与人相关的模式都在那里”，这一句并没有被证明——城市里一条街巷的热岛、一片山谷里的局地雾，在网格之下，被平均掉了。

第三个证据最耐人寻味。哈萨比斯在访谈里谈到，视频生成模型仅凭“被动观察”大量视频，就学会了液体怎样流、物体怎样落、光怎样反射，这让他相信物理世界的许多规律可以从观察中重新学出来；他同时说明，这不是人那种深层的理解（见第28章）。这确实令人惊叹。可是“观察”二字要拆开看。视频是什么？是摄像机在某个焦距、某个帧率、某个动态范围下，把光强量化成八位或十位整数的阵列。人拍下来的，是人觉得值得拍、拍得清的东西：倒水、扔球、海浪拍岸。温度没有被拍下来，气压没有被拍下来，比一个像素更小的涡旋没有被拍下来，快过帧率的碎裂没有被拍下来，镜头之外的世界更没有被拍下来。模型学会的，是“被拍成视频的物理”，是物理界穿过镜头、传感器与压缩算法三道网之后的投影。它在这个投影上学得越好，越说明投影里有结构；但投影外有没有、有什么，它一个字都没学到，也不可能学到。

大白话说：一个只看过家庭录像的孩子，能把“生日蛋糕上的蜡烛一吹就灭”学得极准；可他不会从录像里学到厨房里煤气的味道，因为录像没有味道这一格。

这里有一个常见的反驳，必须正面接住：数据会越来越多，网眼会越来越细，今天漏掉的，明天就捞上来了。这个反驳在方向上是对的，科学史就是网眼不断变细的历史，第3章讲的“回流通道”正是这件事：水星近日点的反常、黑体辐射的反常，都是物理界从网眼里顶回来、逼着数学重新织网的地方。但这个反驳恰恰承认了本章的论点：网眼变细，靠的不是学习算法，而是测量与数学化的更新——新仪器、新变量、新概念。网眼每变细一次，都是一次新的命名、新的切割；而在下一次之前，网眼之下的东西照样不在“自然中的模式”之列。即便像第29章替他设想的那样，自动化实验室让 AI 参与了测量，决定测什么、立哪些变量的那一步，仍属于测量与数学化的更新，不是学习算法本身的能力。哈萨比斯的猜想对网眼的更新不负责，它只对已经捞上来的负责。

◆ 六、这一刀砍到哪里为止

锋利必须配上公允。第一刀砍到哪里，不砍到哪里，要说清楚。

第一，这一刀不否认他的经验层。自然确实不是随机的，大量模式确实可学，这在理念物理界内部是被一再证实的。本书第27章已经说过：AI 能替人做积分，而且做得极好。

第二，这一刀不是相对主义。说“可计算是理念层的谓词”，不等于说理念层的判断随意。在数学态那一格里，对错是严格的：一个函数可计算就是可计算，一个蛋白质结构预测得准就是准，这些都有本地逻辑域内的硬标准。撤掉的是“物理界自己可计算”这个普遍法庭，不是严格性。

第三，这一刀不预设网眼之下必有宝藏。被筛掉的残差里，可能大部分真的是噪声，对人的任何问题都无关紧要。本章只主张：残差里是否含有与人的问题相关的结构化差异，是一个需要逐案检查的经验问题，不能由学习算法的成功来预先回答。若有人能系统地证明残差里不含这样的差异，第一刀就失效——这条失败条件第33章会正式写出。

第四，这一刀对传播版和对本人有不同的分量。传播版说“没有 AI 不能解决的问题”，这一刀对它是致命的：它连“问题”的范围都默认成了“已经上秤的”。对哈萨比斯本人，这一刀只是指出他的猜想有一个未声明的前提；他若补上一句“在已被测量与数学化的范围内”，他的猜想就安全了——但也就小了。小了多少？小到不能再支撑《代达罗斯》对谈题目“AI：科学的终极工具”那样的口气，因为科学最要紧的那一步，恰恰是决定什么东西上秤。

第五，这一刀也砍向本书自己。本书说“可计算是理念层的谓词”，这本身是一个概念规定，不是经验发现；它的价值只在于能不能多出可检验的区分。它确实多出了一条：凡是遇到“AI 能解决某个领域的一切问题”的说法，先问那个领域的数据是怎样测出来的、测量丢了什么、丢掉的部分有没有人在承担后果。若这一问在实践中问不出任何新东西，这把刀就只是在整理语言。

这正是三个“自然”之间的那条缝：物理界、理念物理界、学习算法的数据空间。哈萨比斯的猜想住在第三层，他的口气却说的是第一层。

回到鱼铺。老板的本子没有错，错在那句“海里的都算得清”。而更让人不安的是另一件事：老渔民的那半桶东西，并不是一开始就在秤外的。几十年前，鱼铺也收蟹、收虾，按只按把论价；后来收鱼的规矩变了，只认秤，这些东西才慢慢退出了本子。网眼不是天生的，它是一次次被织出来的；而每一次织网，数学都在扩张，同时都在丢弃。下一章要问的正是这件事：哈萨比斯说“迄今没人在宇宙里发现不可计算的东西”（据转述），这句话记录的，究竟是自然的性质，还是网一次次被重织的战绩？

带 走 的 话

哈萨比斯猜想的全部量词，都落在已被测量、已被数学化的数据上；可计算性只对上升到数学态的那一格有定义。从“数据里的模式可学”到“自然中的模式可学”，中间隔着测量与数学化这张网。这一刀不否认他的成绩，只要求他的猜想补上“在已被测量与数学化的范围内”这一句：补上之后，它安全了，也小了。

小 账

- 本章的鱼铺、老渔民、重点中学、家庭录像均为构想场景或比喻，无真实原型。
- 哈萨比斯的原话以第28章为准，DeepMind 播客那句仍据转述；AlphaFold 对无序区的提示，出处见附录I第I49条。
- 本章不主张网眼之下必有与人的问题相关的结构，也不主张理念层的判断随意；“可计算只对数学态有定义”是概念规定，不是经验发现。
- 这一规定的推翻条件见附录G第G13条；第一刀整体的失败条件见第G42条。

31

第 31 章

第二刀：“从未发现”是回写的战绩

> 这一章的问题：老裁缝说“三十年没碰到做不了的人”，而他的尺码表从五栏长到了十七栏——“从未发现不可计算之物”，说的是自然，还是那张表？

街口有一家老裁缝铺。几十年来，师傅只做一种尺码表：胸围、腰围、肩宽、袖长、衣长，五个数。来一个客人量五个数，照表裁剪，从没出过大错。有人问他：“师傅，您这五个数，能不能做所有人的衣服？”师傅很笃定：“干了三十年，还没碰到哪个人是这五个数做不了的。”

可徒弟知道这句话是怎么来的。头几年，碰上驼背的老人，五个数做出来的衣服前短后长，师傅就在表上加了一栏“背长”；碰上一个肩膀一高一低的木匠，又加了“左右肩差”；后来做旗袍，又加了“臀围”“领围”。每一次碰壁，表就多一栏。三十年下来，表从五栏变成了十七栏。“还没碰到做不了的人”，说的其实是：每次碰到做不了的人，表就改了，改完就做得了了。

还有一句徒弟没说出口：真有一个客人，是任何尺码表都做不了的——一个正在长个子、每个月都变的孩子。师傅的办法是，“过半年再来”。于是那个孩子从来没有出现在“做不了的人”的名单里，他只是被推迟了。

第28章录下了哈萨比斯另一层意思：迄今为止，宇宙里还没有人发现过不可计算的东西。这句话目前据转述（见第28章对出处的交代），按第28章的分级属第三等材料，所以本章的论证不压在原话的字面上，而压在这类说法共有的推理形状上；本书按其最强读法来对待：它是一个关于物理学史的经验陈述，用来支撑“自然可被经典算法高效建模”的猜想。本章要砍的就是它。不是说这句话是

假的——在严格意义上它大概是真的——而是要问：这个“从未”，是自然的性质，还是裁缝铺尺码表的性质？

◆ 一、数学每次碰壁，就改尺码表

先看历史。第四编与第22章已经一章一章地讲过微积分的自我更新，这里只把它们串成一条线，看清它们的共同形状。

牛顿的微积分是为光滑的、连续的、确定的运动打造的。行星沿着光滑轨道走，炮弹沿着抛物线飞，一切都可以求导。它第一次显出边界，是在不确定性面前：骰子、人口、观测误差。数学的回应不是承认“这些东西算不了”，而是让概率论（其源头比微积分还早，在帕斯卡与费马的通信里）与分析学合流，到拉普拉斯、高斯手里成了处理误差与统计的一整套工具，把不确定性本身变成可以计算的对象。

第二次碰壁，是在不可微面前。花粉在水里的乱跳，股价在交易日里的抖动，这些轨迹处处连续、处处不可导，牛顿的导数在这里失灵。数学的回应是伊藤积分：既然普通的导数不存在，就重新定义积分，让随机微分方程成为合法的对象（第19章）。

第三次碰壁，是在微观世界面前。黑体辐射、原子光谱、电子衍射，经典力学的连续轨道一概不成立。数学的回应是把希尔伯特一派已有的无限维空间理论改造成量子力学的舞台，发展出算子理论：把“状态”从相空间里的一个点，改写为无限维空间里的一个向量，把“可观测量”改写为算子（第21章）。

第四次碰壁，是在混沌面前。三体问题没有一般的封闭解，天气对初值敏感到一只蝴蝶也能搅乱预报。数学的回应是动力系统理论：放弃预测单条轨道，转而描述吸引子、李雅普诺夫指数、不变测度，把“不可预测”本身变成一个可以严格陈述、可以计算的性质（第22章）。

四次碰壁，四次扩表。这里要说清一点：并非每次扩充都由物理反常引起，概率论起于赌局与人口统计，它与分析学合流时受的是物理与统计的双重推动；上面四例取的是与物理顶撞最直接的几次。形状是一样的：现实从网眼里顶回

来，旧的数学说不清；数学不认输，它扩充自己，长出新的对象、新的运算、新的定义，然后在新的框架里，原来“说不清”的东西变得说得清、算得出了。

用第395卷的 123 原理来讲，这正是一个完整的循环：① 物理界顶回来的反常，与已沉淀的数学框架发生 D—E 矛盾；② 矛盾累积到旧框架装不下，逼出数学显露形态的改变——新的分支、新的概念；③ 新的数学回写：它重新定义了“什么算一个合法的问题”（重置 D），沉淀为下一代科学家训练的土壤（重置 E）。回写之后，原来那个反常不再是反常，它变成了新框架里的一道习题。

还要补上回写的另一面。第395卷提醒，回写是中性的，不等于改善。每一次扩表，数学都网进了一些东西，也放走了一些东西。概率论让骰子可算，代价是放弃追问“这一次为什么是六点”；统计力学让气体可算，代价是不再追踪任何一个分子；混沌理论让不可预测变得可陈述，代价是承认单条轨道不再是研究对象。新框架里“算得出”的，常常是换了一个问题之后的答案。这不是缺点，这是科学的成熟方式；但它意味着“从未发现不可计算之物”，有一部分是靠“把算不出的那个问题换掉”来维持的。

所以，当有人说“迄今没人在宇宙里发现不可计算的东西”时，本书要追问一句：每一次差点发现的时候，发生了什么？答案是：数学改了表。

◆ 二、“从未发现”的两种读法

这里必须小心，不能把话说过头。

“数学每次碰壁就扩充”，是否推出“自然其实不可计算，只是被数学掩盖了”？不推出。本书没有、也不能主张这一点。扩充后的数学确实算得出来——伊藤积分有数值方法，量子力学可以在计算机上模拟（虽然代价可能是指数级的），混沌系统的统计性质可以被可靠地估计。扩表之后，做得了就是做得了，不是假装做得了。

本书主张的是另一件事：“从未发现不可计算的东西”有两种读法，而这两种读法得出的推论完全不同。

第一种读法，是哈萨比斯需要的读法：自然本身是一个可计算的系统，所以我们从未碰到不可计算的东西，将来也不会；“从未”是自然性质的证据。

第二种读法，是本书的读法：可计算与否，是数学化之后才有定义的判词（第30章）；每次碰到现有数学处理不了的东西，科学的习惯动作不是宣布“此物不可计算”，而是改造数学，直到能处理为止；于是“从未发现”主要是一个关于科学习惯的记录，而不是关于自然本身的证据。

两种读法对同一份历史记录都说得通。这就是要害：一份同时支持两种读法的记录，不能用来在两者之间裁决。哈萨比斯若拿它作为自然可计算的证据，就必须先排除第二种读法；而他没有排除。

大白话说：一家医院说“我们科室十年没有出过一例死亡”，可能是因为医术高明，也可能是因为危重病人一来就被转到别的科室。只看“十年零死亡”这一条记录，分不出是哪一种。要分清，得去看转科记录。

那么，科学的“转科记录”在哪里？就在那些已经被严格证明的“不可计算”结果里。如果自然真的处处可计算，这些结果应该指向某种物理上的错误；如果本书的读法对，这些结果应该集中出现在数学化最狠、理想化最深的地方。下面就去看。

◆ 三、已知的不可计算：都在理念的最深处

数学物理里确实有几项被严格证明的“不可计算”或“不可判定”结果。它们常被拿来反驳哈萨比斯。本书的用法正相反：它们不是“自然不可计算”的证据，而是“不可计算性住在理念层”的证据。

第一项是波动方程。1981　年，玛丽安·普尔-埃尔与伊恩·理查兹在《数学进展》上发表了一个结果：三维空间中的波动方程，可以构造出可计算的初始数据，而它在时刻 $t = 1$ 的唯一解却不可计算。这个结果一度被解读为“物理定律能产生不可计算的东西”。但条件要写准。

普尔-埃尔—理查兹的波动方程例子：条件在哪里 方程：$\dfrac{\partial^2 u}{\partial t^2} = \Delta u,\ x \in \mathbb{R}^3$，初值 $u(x,0) = f(x),\ \partial_t u(x,0) = 0$。结果：存在可计算的 f，f 只要求一阶连续可微（C^1），使得 $u(x,1)$ 是连续函数，但在一致范数（上确界范数）意义下不可计算。条件一：初值的光滑度很低。C^1 的初值经基尔霍夫公式给出的解，一般不必二阶可微，只在广义意义上满足波动方程。条件二：“不可计算”依赖所选的函数空间与范数。二人自己的“第一主定理”说，有界线性算子保持可计算性，无界算子则未必；波的传播算子在一致范数下无界，在能量范数下有界——在能量范数下，可计算初值永远给出可计算的解。条件三：此后魏劳赫与钟宁（2002，《伦敦数学会会刊》）证明，若用连带导数信息的拓扑来表示函数，波的传播是图灵可计算的；在那种表示下，上面那个 f 本身就不再算“可计算的输入”。结论：不可计算性出现在“非光滑初值 $\times$ 一致范数”这一特定的数学化选择处；换一种同样合理的数学化，它就消失。

看清楚了吗？这个“不可计算”，不是从物理界里捞出来的，而是在一个特定的数学建模选择——用什么空间、什么范数来描述波——里出现的。物理学家在实验室里激发一个波，不可能制备出一个精确等于 f 的初态：f 本身虽是可计算函数，却是借一个递归可枚举而不递归的整数集精心构造出来的，它的导数信息并不可计算；任何有限精度的测量都只能逼近它，而恰恰在一致范数下，逼近不能保证解也逼近。不可计算性住在“无限精度的初值”与“不连续的传播算子”交汇处，那是实数连续统的最深处。

第二项是量子多体系统的谱隙。2015 年，库比特、佩雷斯-加西亚、沃尔夫在《自然》上证明：判定一个量子多体系统是否有谱隙，是不可判定的。谱隙是凝聚态物理的基本问题之一，它的有无与材料在低温下处于哪一种量子相、会不会发生量子相变密切相关。这个结果分量很重，它说的是：对这一类平移不变、最近邻相互作用的哈密顿量，不存在一个通用算法，能对每一个给定的哈密顿量判断它是“有隙”还是“无隙”。条件同样要写准。

谱隙不可判定：条件在哪里 系统：二维方格晶格，每格点上是 d 维量子系统（d 固定，但在构造中非常大）；哈密顿量平移不变，只有最近邻相互作用。问题：在热

> 力学极限（晶格边长 $L \to \infty$）下，系统是“有隙”（谱隙有一个不随 L 消失的正下界）还是“无隙”（基态之上的谱在极限中连续）。即使附加“必居其一”的承诺，问题仍不可判定。　证明思路：把图灵机的停机问题编码进哈密顿量的基态；机器停机与否，决定极限下有隙与否。　关键：对任何有限的 L，哈密顿量是一个有限矩阵，它的谱原则上可以计算；不可判定性只在 $L \to \infty$ 的极限里出现。构造中，谱隙可以在一个不可计算地大的尺寸上才突然闭合。　后续工作（鲍施、库比特、卢怡、佩雷斯-加西亚，2020，《物理评论　X》第10卷）把类似结论推广到一维链，但仍依赖热力学极限。

这里的不可判定，住在哪里？住在“热力学极限”这个理想化里。真实的材料没有无穷多个原子；一块晶体，原子数是有限的，哪怕是 10^{23}。热力学极限是物理学家为了让问题干净而发明的数学装置：它把“很大很大”替换成“无穷”，以便用连续的概念（相、相变、谱的连续部分）来描述。这个装置极其成功，但它是理念的产物。在理念里，无穷晶格的谱隙不可判定；在现实里，一块有限的材料，原则上可以被计算，只不过可能大到在实践中无从下手。

大白话说：问“一条无限长的队伍里，最后一个人戴不戴帽子”，没人答得上来，因为根本没有最后一个人；可你问“这条一百万人的队伍里最后一个人戴不戴帽子”，跑过去看就行，只是腿要跑断。

还可以再加一项。1990　年，克里斯托弗·穆尔在《物理评论快报》第64卷上构造了一类低维动力系统（二维的分段线性映射，可以嵌入三维的光滑流，也可以由在一组抛物面镜之间反弹的粒子实现），其长期行为等价于一台通用图灵机，因而某些关于轨道的问题不可判定。这个结果的条件，同样是要求以无限精度给定初始位置：把整台图灵机的纸带编码进一个实数的无穷位小数里。一旦初值只能以有限精度给出——这是任何实验的处境——不可判定的问题就退化成“在有限时间内的有限精度预测”，后者是可计算的，虽然可能很贵。

第四类情况尤其值得哈萨比斯本人留意，因为它就在他拿来举例的流体力学里。三维纳维—斯托克斯方程的光滑解是否会在有限时间内“爆破”，至今是一个悬而未决的数学难题。陶哲轩在 2016 年构造了一个“平均化”的纳维—斯托克斯方程，证明它的解会在有限时间爆破，并设想过用理想流体搭建一台“流体计

算机”，让方程的解自己去执行计算；此后卡尔多纳、米兰达、佩拉尔塔-萨拉斯、普雷萨斯（2021，《美国国家科学院院刊》）在一个带特定黎曼度量的三维球面上构造了图灵完备的定常欧拉流，由此，关于流体粒子轨线的某些问题（例如从某点出发的轨线会不会进入某个区域）是不可判定的。这一方向还在推进，未来或许会延伸到真正的纳维—斯托克斯方程。

但请看它们站在哪里：经过“平均化”改造的方程、有限时间里梯度趋于无穷的爆破解、无黏性的理想流体与特意挑选的几何、要求无限精度的轨线问题。它们都在数学化最深、离任何一杯真实的水最远的地方。哈萨比斯说 AI 能学会流体，是在网格尺度、有限精度的数据上说的；那里本来就没有这些不可判定性。两边说的是不同楼层的事。

三项结果，加上这一类正在推进的工作，有三个共同点：它们都在数学化之后才出现；它们都出现在积分的无穷极限处——无穷精度的实数、无穷大的晶格、无穷长的时间；它们都依赖理想化选择，换一种同样合理的理想化，往往就消失或转形。

◆ 四、可计算与不可计算，都是理念层的判词

现在可以把第二刀合起来了。

哈萨比斯说：至今没人在宇宙中发现不可计算的东西。本书说：这句话在一种意义上是真的——没有人在实验室里直接撞见过一个“不可计算的物理过程”；但它不是自然可计算的证据，原因有二。

其一，历史上每一次差点撞见，数学都改了尺码表；“从未发现”是 123 原理第③步——回写——一次次成功的战绩，是科学的习惯动作留下的痕迹，不能直接记作自然的性质。

其二，那些被严格证明的“不可计算”，没有一项是从物理界里直接捞出来的；它们全都住在理念物理界的最深处，在无穷极限、实数连续统、非光滑解这些理想化的地方。也就是说，“不可计算”和“可计算”一样，都是对数学化之后的对象下的判词。

这两条合起来，结论是对称的：**物理界既没有被证明可计算，也没有被证明不可计算；这个问题在物理界那一层没有定义。** 哈萨比斯倾向于自然可计算，彭罗斯主张意识涉及不可计算的过程，在本书看来，两人都把一个理念层的谓词，搬到了理念层之外。本书不站在任何一边，而是指出两边共同站错了层。

这不是和稀泥。它有一个可检验的推论：如果本书对，那么将来再出现新的“不可计算”结果，它们仍会集中在理想化最深的地方；同时，每当这些结果被逼近到可实验的情境（有限尺寸、有限精度、有限时间），不可计算性就应当消退为“计算代价极高”。反之，如果有一天，某个不可计算结果被证明不依赖任何无穷极限理想化——比如一块有限尺寸的真实材料，其可测量的性质被证明不可判定——那么第二刀的这一论据就作废，第33章会把这一条写进失败条件。

这里也要防一种误用。有人会说：既然不可计算性只在无穷极限处出现，那它与现实无关，可以忽略。不对。热力学极限、实数连续统、光滑解，正是物理学最有力的积分工具；物理学家之所以要走到无穷，是因为只有在那里，“相”“相变”“守恒”这些整体概念才干净地成立。不可计算性恰恰在积分最成功的地方冒出来，这说明它是积分本身的代价：要把无数局部累加成一个干净的整体，就得付出“这个整体未必能被有限程序算出”的价钱。第11、12章讲过，积分出来的是 $f+C$，C 要从现实借回；这里看到的是另一笔账——积分走到无穷，换来了整体的清晰，也换来了可判定性的丧失。两笔账都记在理念那一侧。

还要替哈萨比斯说一句公道话。他本人并不是把“自然可计算”当成定理宣布的。据第28章整理，他在 莱克斯·弗里德曼（Lex Fridman）访谈中说，对没有可学结构的问题“也许需要一台量子计算机”；另据传记书评转述，他承认有些现象可能需要巨量计算，或者无法被经典方法高效还原，那里“也许需要量子系统”。这说明他意识到了边界，只是把边界画在了计算资源上（经典与量子），而没有画在数学化上。本书的第二刀，正是把边界从“用什么机器算”挪到“算的是什么”。前者是工程问题，后者是认识论问题。他的猜想对前者负责，对后者沉默。

回到裁缝铺。师傅说“三十年没碰到做不了的人”，这句话是真的；他的手艺也是真的。但徒弟知道，这三十年里尺码表加了十二栏，还有一个长个子的孩子被推迟了。徒弟不会因此说师傅骗人，他只会说：“没碰到做不了的人”，记录的是这张表一次次被改好的历史，不是人的身体天生就能被一张表装下。

可是表的每一栏，是谁加上去的？谁决定驼背值得加一栏、而某种别的差异不值得？那个长个子的孩子，凭什么没有资格进“做不了的人”的名单？这些问题，已经不是计算的问题了。下一章去问：问题本身，是谁命名的。

带走的话

“迄今没人发现不可计算之物”同时容得下两种读法：自然本来可算，或数学每次碰壁就扩充自己。一份同时支持两种读法的记录，不能用来裁决。已被严格证明的不可计算结果，波动方程、谱隙、穆尔的动力系统，都只在无穷精度、无穷尺寸、非光滑解这些理想化处成立。所以“可计算”与“不可计算”都是数学化之后的判词；物理界本身既没被证明可算，也没被证明不可算。

小账

- 本章裁缝铺、医院转科、无限长队伍等均为构想场景或比喻，无真实原型。
- 哈萨比斯“迄今没人发现不可计算的东西”一语据转述（第28章所列第三等材料），本章只把它当作一种推理形状来检验，不当作已核定的原话。
- 普尔-埃尔一理查兹、库比特等、穆尔、卡尔多纳等的结果出处见附录I（I8—I12）；方框中所写的适用条件以原论文为准。
- 本章不主张自然不可计算，也不主张扩充后的数学是在“假装算得出”。
- 本章主张的推翻条件见附录G第43、44条。

32

第 32 章

第三刀：问题是谁命名的

> 这一章的问题：一条导航地图上没有的近路，骑手投诉也只换来“已是最优解”——当一个问题连被提出的资格都没有，算得再好又能解决什么？

老城区有一片巷子，没有名字。说没有名字也不准确：住在里面的人叫它“后头”，送煤的叫它“三拐”，孩子们叫它“猫道”。它窄得只容一个人推一辆自行车，从菜市场后门斜穿过去，能省下绕大路的十分钟。三十年里，它是这一片人的常识。

导航普及以后，情况慢慢变了。导航的地图上没有这条巷子——它太窄，没有门牌，测绘车开不进去。新搬来的人只走导航指的大路；送外卖的骑手按导航跑，绕远十分钟，超时扣钱，但系统里“最优路线”就是那条大路。几年过去，老住户搬走了一些，孩子们长大了，巷子里堆起了杂物，一头被新盖的车棚封上了。没有人决定废掉这条巷子。它只是从地图上没有，变成了从人的经验里也没有。

后来有个骑手跟平台投诉：这片区域的配送时间算得不合理。平台的回复很客气：系统的路径规划已经是最优解。这句话没错。在系统认得的路网里，它确实是最优解。问题是，“这里有一条更近的路”这件事，在系统里连提出的资格都没有：它不是一个算错了的问题，它根本不是一个问题。

前两刀砍的是“可计算”：第一刀说它是筛子的性质，第二刀说“从未发现不可计算”是回写的战绩。第三刀换一个方向，不再问“算不算得出”，而

问“算的是什么”。从“自然模式可学”走到传播版的“没有 AI 不能解决的问题”，中间有一道常被略过的桥：问题是从哪里来的?

◆ 一、三大主权：计算权排在最后

第395卷第11章讲“E 的临床学”时，提出过三大主权：命名权、判错权、计算权。三者分别对应符号学、逻辑学、数学，对应信息的粒子态、波态、场态。顺序不是随意的：命名权居首。

命名权，是决定“什么东西被叫出名字、成为一个可以被指认的对象”的权力。巷子有没有名字，决定它能不能进入地图；一种症状有没有病名，决定它能不能进入医保目录；一种劳动有没有岗位名，决定它能不能进入工资表。名是指针（第395卷第10章），它指向一束特征纠缠；没有指针，那束纠缠就在符号库里没有地址。

判错权，是决定“什么算错、错了以后怎么办”的权力。导航说“您已偏离路线”，这是判错；它判的标准是它自己的路网。骑手走了巷子，在系统里是“偏离”；系统的这条判错条款，没有一个入口让“偏离”被重新解读为“发现了更近的路”。判错权握在谁手里，谁就决定了哪些失败会被当成信号，哪些只会被当成噪声。

计算权，是在已经被命名、已经有判错标准的对象上，决定“怎样算、按什么代价函数优化”的权力。导航的最短路径算法，是计算权最漂亮的展示：在给定的路网和给定的代价（时间、距离）上，它算得又快又准。

三者的关系是层层托付的：计算权站在判错权上，判错权站在命名权上。一个问题要被计算，先要被命名为问题，再要有判断解答对错的标准，最后才轮到计算。哈萨比斯的猜想，以及 AI for Science 的全部成就，讲的是计算权：一旦问题被说清、对错有标准，AI 能以前所未有的效率去算。这是真本事。但从“计算权极强”到“没有解决不了的问题”，要跨过另外两权，而这两权，他的猜想本身并不处理。（他在别处承认提出正确的问题最难，见下文与第28章；但这一承认没有进入猜想的论证。）

> **问题的三层构成** 把一个“可被 AI 求解的问题”写成三元组 $P=(N,J,C)$：N：命名——对象与变量的指定，即从现实中截出哪些特征、以什么符号表示（输入空间 X、输出空间 Y）；J：判错——成败的标准，即损失函数 $\ell: Y\times Y\to\mathbb{R}$、验证集、何时算“解决”；$C$：计算——在 N、J 给定后，寻找 $h: X\to Y$ 使期望损失最小。学习理论与算法的全部定理都以 N、J 为前提，结论只关乎 C。一个现实困境 R 被“解决”，严格说是：存在某个 (N,J)，使 R 在其下的解 h 被现场接受。这一步不是 C 能证明的；同一个 R 可有多种 (N,J)，在一种下“最优”，在另一种下可能根本没有提及 R 的要害。

大白话说：一个孩子数学考了六十分。“提高分数”是一种命名；“他到底卡在哪个概念上”是另一种命名；“他是不是因为家里出了事、最近睡不着”又是一种命名。AI 在第一种命名下能给出最优的刷题计划。可如果真正的问题在第三种，那份最优计划越精确，离孩子越远。

◆ 二、被筛掉的残差连“问题”的身份都拿不到

第30章说，数学化是一张筛子；第31章说，筛子每次碰壁会改网眼。本章要补上一层更冷的东西：那些从未碰壁、只是静静地漏下去的东西，命运是什么？

它们的命运不是“没被解决”，而是“没被提出”。这两者有天壤之别。“没被解决”的问题，还在问题清单上，还有人记得它、还有资源会流向它、还可能有一天被解决。“没被提出”的东西，不在任何清单上；它不占预算，不进论文，不出现在任何基准测试里；它对系统来说不存在。

巷子就是这样。它不是导航“没算对”的问题——导航从没把它当问题算过。骑手的十分钟、老住户的三十年经验、孩子们的“猫道”，全部落在了命名权的门外。最残酷的是，系统越成功，这扇门关得越紧：大家都用导航，就没人再走巷子；没人走，巷子就荒；荒了，“这里有路”就连在人的记忆里也不成立了。系统的最优解，回过头来抹掉了它从没看见过的东西。第26章把这件事记作落在人身上的负向回写：外部的不成文积分好用到一定程度，人身上那份对街巷的手感就不再被练习。本章要追加的是它在“问题清单”上的后果：手感退化之

后，连“这里本来有条路”的申诉也失去了提出者，命名权的门外从此空无一人。

这一层，哈萨比斯的猜想并不涉及，因为他的论证起点就是数据。“自然中能被找到的模式”——被谁找到？被数据找到。数据之外的残差，不是“尚待学习的模式”，而是“没有资格成为待学习对象的东西”。传播版的“没有 AI 不能解决的问题”，在这个意义上是同义反复：凡是被 AI 系统命名为问题的，AI 系统当然可能解决；没被命名的，压根不在“问题”之列，也就不会出现在“不能解决的问题”之列。一个永远不会出现反例的命题，在经验上是空的。

大白话说：一家公司宣布“我们的客服系统解决了百分之百的工单”。你去查，发现系统只接受下拉菜单里的十二种问题类型；你的问题不在十二种里，就提交不了工单。百分之百，是在这十二种里的百分之百。

这里必须守住公允。不能说凡是被筛掉的都是珍宝。巷子的例子容易煽情，但城市里也有无数被导航“遗忘”的路，是真该被忘掉的死胡同。命名权的筛选本身不是罪，没有筛选就没有科学，也没有城市。本书要说的只是：被筛掉的东西里是否有与人相关的结构化差异，这件事不能交给执行筛选的系统自己来判断，因为它在结构上看不见筛子之下。判断这件事的，必须是一个握有另外两权、并且站在现场的主体。

◆ 三、基准测试：判错权藏在哪里

把这一层放到哈萨比斯最得意的成就上看，会看得更清楚。

AlphaFold 的成功，有一个常被略过的前提：蛋白质结构预测这个问题，在它出场之前，已经被一个科学共同体花了二十多年命名好、判错好了。从 1994 年起，结构预测领域每两年举办一次盲测评比（CASP）：组织者收集一批刚刚测出、尚未公开的结构，让各路方法去预测，再用事先约定的打分标准（如 GDT 分数）比较预测与实验结构的接近程度。什么叫“一个蛋白质的结构”（晶体或冷冻电镜里测到的原子坐标），什么叫“预测得好”（分数怎样算、高到多少才算接近实验精度），这些都是人在几十年里一点点定下来的。AlphaFold 2 在 2020

年那一届取得压倒性成绩，它赢在计算权；而那场比赛的命名和判错，是人类几十年工作的沉淀（蛋白质结构数据库更早，始于 1971 年）。

这不是贬低。恰恰相反，这说明 AI for Science 最成功的案例，都发生在命名权和判错权已经高度成熟的地方。围棋有明确的规则和胜负，蛋白质结构有数据库和盲测，天气有观测和评分。哈萨比斯本人对此并不讳言，他在 2024 年的诺贝尔演讲中列过适合用 AI 攻克的问题的三个条件：巨大的组合搜索空间；有一个清晰的目标函数（度量）可供优化；有大量数据，和／或一个准确而高效的模拟器；第28章核实过的《代达罗斯》对谈里，他也说提出正确的问题是科学中最难的部分。这是一个高明的选题标准。但这个选题标准本身就是一张筛子：它挑出的，是三权中前两权已经被人解决了的那些问题。在这类问题上屡战屡胜，不能外推到前两权尚未落定的问题上去。

那些前两权尚未落定的问题是什么样子？看一个医院的例子。一种慢性疼痛，病人说疼，各项检查都正常。它是什么病？在某个时期它被叫作“心因性”，在另一个时期它有了一个综合征的名字，在不同的医生那里它有不同的判法：有人看疼痛评分，有人看能不能上班，有人看病人睡得好不好。这里没有公认的“结构数据库”，没有公认的“盲测标准”；问题的名字和对错标准本身，还在病人、医生、家属、保险公司之间拉扯。把 AI 放进来，它能做的第一件事，就是按某一种现成的命名和判错去训练——而那一种命名，会因为被 AI 放大而迅速成为“标准”，其余的命名就被挤出了数据。

大白话说：一个班要评“最守纪律的学生”，老师若先说清楚标准是“上课不说话、作业按时交”，那么用软件统计、排名，又快又公平。可如果全班还在争论“守纪律到底是什么意思”，这时候先把软件装上，软件用的那一版定义就赢了，争论也就结束了——不是因为争论有了结果，而是因为只剩一版定义有数据。

所以第三刀在 AI for Science 的语境里有一个更具体的形态：**AI 的成功清单，是一份“前两权已被人类结算”的问题清单；把它当作“所有问题”的样本，是选择偏差。**

◆ 四、AI 成为主要执行者之后：矛盾被压平

再往前推一步。到今天为止，数学化——把现实截成变量、写成方程、累加成模型——主要是由人做的。人做数学化，有一个常被当成缺点的特点：慢，而且做得不彻底。一个工程师把一座桥写进有限元模型，他自己知道模型里没有哪些东西；一个医生用评分系统，他自己还看过病人的脸色。数学化的执行者同时也是现实的经验者，模型的边界就在他身上，筛子之下的东西会通过他的不安、他的直觉、他的“总觉得哪里不对”，顶回来。

第3章所说的“回流通道”，在很多时候就是这样一个人。水星近日点每世纪那约四十三角秒的多余进动，不是数据自己跳出来抗议的；是天文学家不肯把它当误差抹掉，一代代地记着、追着、为它焦虑，最后逼出了广义相对论。用 123 原理来说：D—E 矛盾的累积，需要有人承受。新的差异（D）冒出来，与已沉淀的框架（E）冲突，这个冲突要有一个地方积存，直到它大到逼出 S 的改变。

现在设想 AI 成为数学化的主要执行者。数据进来，模型学到，预测给出，结果被现场采用；下一轮数据又按同样的口径进来。这里面，谁来承受矛盾？模型拟合的目标就是把残差压到最小；对于不符合模式的数据点，它的标准处理是当成噪声平均掉，或者调整参数把它吸收进去。它很少会“焦虑”，也不会因为一个点对不上而拒绝收工。更关键的是，当它的输出被大规模采用之后，现实本身会被它的输出塑造——骑手都走大路，巷子就荒了；于是下一轮的数据里，巷子连作为残差都不再出现。

这正是第395卷谈 123 原理时所说的第一种病：**矛盾被压平。** 井然有序、再无争论，看上去是成熟，实际上是停滞：D—E 之间那台引擎熄了火，因为新的差异要么被平均掉，要么在出现之前就被系统塑造的现实消灭了。第二步（S 的改变）没有了原料，第三步（回写）也就只剩下系统对自身的回写——一个越来越自洽、越来越与它看不见的东西无关的闭环。

要强调，这是一个风险判断，不是事实宣判。第395卷要求回写“要被观察，不能被宣布”，压平也一样。本书不主张“AI 必然压平矛盾”；本书主张：当 AI

同时握有计算权、并通过默认口径事实上握有命名权与判错权时，压平是一个需要被监测的结构性风险，而它的症状是可以观察的——残差的多样性在下降，异常报告在减少，现场人员的“不对劲”越来越少被记录，系统的自评分越来越高而现场的投诉越来越无处投递。

◆ 五、最强的反驳：AI 也能命名

第29章替哈萨比斯摆过一条回答：他从未声称 AI 现在就能提出好问题，而是把命名当作一道尚待跨过的门槛。这里要替他再往前推一步，接住一个更有力的反问：凭什么认定命名权只属于人？大模型已经能从文献里提出新的假说，能在数据里发现人没注意到的聚类，能给出“这里可能有一个新的问题”的提示。AlphaGo 的第37手，就是一次在人类棋理里没有名字的落子。AI 难道不能命名吗？

本书的回答分两层。

第一层，承认：AI 能产生新的符号组合，能在数据空间里指出新的簇，这些都可以算作命名的候选。第395卷已经撤回“大模型创造力几乎为零”一类旧说，本书不重拾它。

第二层，限定：AI 的命名候选，仍然只能从已进入数据的东西里长出来。数据里的一个新簇，是筛子之上的新发现；巷子不在数据里，它就不可能成为一个簇。更要紧的是，一个命名候选要成为“问题”，还要过判错这一关：谁来说“这个新名字指着的东西是真的、值得追”？第37手之所以被认作神来之笔，不是因为 AlphaGo 宣布它好，而是因为它赢了，而“赢”是围棋规则这一条早已存在的判错条款给出的。离开现成的判错条款，AI 的命名候选就只是候选；把它们中哪一个升格为问题，仍要有人在现场承担后果。命名权与判错权可以有 AI 参与，但它们的最后一环——为这个名字、这条判错标准付代价的那个位置——是一个现场位置，不是一个计算位置。

大白话说：孩子的作文里冒出一个新词，老师可以觉得它妙；但这个词能不能进字典，要看它在往后的日子里有没有人用、用了会不会闹误会、误会了有没有人负责。提出新词和让新词站住，是两件事。

◆ 六、把另外两权放回现场

那么，出路是让 AI 少做吗？不是。本书第27章已经说过，AI 替人做积分是一件好事，而且会越做越好。问题不在 AI 算得太多，而在三权被默认地打包交了出去。

第395卷第11章给过治理动作，搬到这里正好合用。第一，命名权要可改：问题清单必须有一个入口，让现场的人能把“这里有一条路”“这个病人不对劲”提交为一个新的问题类型，而不只是在旧类型里挑一个最接近的。第二，判错权要写成条款：什么情况下系统的“最优”应被判为错——比如同一片区连续出现超时投诉，就触发人工复核路网，而不是继续扣骑手的钱；这条条款要有改轨策略和复位规则。第三，计算权要多目标：代价函数里不能只有一个“时间最短”，还要有修复裕度和停止条件——当优化开始吞掉它不认得的东西时，要有一个停下来的规矩。

大白话说：一家人用记账软件管钱，软件算得清清楚楚，这很好。但每个月月底，全家要坐下来一次，问三个问题：有没有什么开销是软件分类里没有、我们却在花的？有没有什么“省下来的钱”其实是该花没花、以后要加倍还的？我们这个月想要的，是不是还只有“结余最多”这一件事？这三个问题，软件替不了，也不该替。

这一刀砍到这里为止。它不否认 AI 的计算权，也不预言 AI 必然压平矛盾。它只指出：从“自然模式可学”到“没有不能解决的问题”，中间缺了“问题是谁命名的、错是谁判的”这两层；而这两层一旦被默认交给计算方，被筛掉的东西就连“问题”的身份都拿不到，系统越成功，它们越彻底地消失。哈萨比斯本人把“给定 1900 年的知识发明相对论”当作检验 AGI 的灯塔（见第28章），本

书认为这个检验选得好，因为它恰恰是一个命名权的检验：相对论的发生，首先是把“同时性”重新命名为一个问题，其次才是计算。

回到巷子。如果有一天导航系统真的学会了那条巷子，那会是因为什么？多半是因为有足够多的骑手违抗了导航，走了巷子，留下了轨迹数据。也就是说，巷子要重新进入计算权的辖区，先得有人在命名权和判错权上冒险。下一章，把三刀合起来，逐层给出判词，并写明本编在什么情况下算错。

带 走 的 话

一个问题要被计算，先要被命名、再要有判错标准，计算权排在最后。AI 的成功清单，几乎都是命名与判错已由人结算好的问题，拿它推断“没有解决不了的问题”是选择偏差。被筛掉的东西连“问题”的身份都拿不到；当 AI 经由默认口径同时握住三权，矛盾被压平便是一个需要监测的结构性风险，而不是已被证实的事实。

小 账

- 本章的无名巷子、骑手投诉、客服工单、考六十分的孩子、慢性疼痛病人、家庭记账等均为构想场景，无真实原型。
- 哈萨比斯诺贝尔演讲中的选题三条件据演讲幻灯片（附录I第I50条）；《代达罗斯》对谈与 莱克斯·弗里德曼（Lex Fridman）访谈的出处见第28章。
- 三大主权、矛盾被压平等用语取自第395卷，在本书只作分析语言，不作可计算的模型。
- 本章不主张 AI 不能参与命名，也不主张 AI 必然压平矛盾。
- 本章主张的推翻条件见附录G第45条。

33

第 33 章

判词与失败条件

> 这一章的问题：修车铺老板在单据背面写明“出现什么毛病算我修错”——三刀砍完，本编对哈萨比斯的判词是什么，又在什么情况下算本编错？

小区门口有一家修车铺，老板给自己定过一条规矩：每修好一辆车，在单据背面写两行字。第一行写“修了什么、为什么这么修”；第二行写“若三个月内出现哪种毛病，算我修错，免费返工”。比如换了刹车片，第二行就写“若三个月内刹车异响或制动距离变长，算我修错”。

街坊起初觉得他多事，后来都只认他家。原因不复杂：别家师傅说“修好了”，这话永远不会错，因为车再坏，总能说是“又坏了别的地方”；他家的单子，写明了自己在什么情况下算错。一句永远不会错的话，等于什么也没保证；一句写明了怎样算错的话，才是一句真话。

第30至32章砍了三刀。刀砍完了，要像修车铺老板那样，把单据翻过来，写两行字：第一行，本编对哈萨比斯的判断，逐层是什么；第二行，本编在什么情况下算错。

◆ 一、先分清：本人命题与传播版

判词之前，先把被判的对象再钉一次。第28章已经区分清楚，这里只做摘要。

本人命题，是哈萨比斯在诺贝尔演讲与此后访谈中提出的猜想：自然中能被生成或找到的模式，能被经典学习算法高效发现与建模。附带的理由是自然经过筛选、不是随机的；附带的限定是，有些现象可能需要巨量计算或量子方法；附带的历史陈述（据转述）是，迄今宇宙中没人发现不可计算的东西；附带的自我检验是，今天的 AI 还做不到范式级突破，“用 1900 年的知识发明相对论”是 AGI 的灯塔。

传播版，是媒体与用户在转述中放大出的口号：“没有 AI 不能解决的问题。”它不是他的原话，也不应由他负全责。

两者在本编中的处境不同。传播版对三刀全无招架之力：它把“已上秤的”当成“海里的”（第一刀），把扩表的战绩当成自然的性质（第二刀），把“被命名的问题”当成“所有问题”（第三刀）。对传播版，本编的判词只有一句：它在经验上是空的，因为它在结构上不可能遇到反例。

本人命题则复杂得多。它有真内容、有自我限定、有可检验的部分。下面逐层下判词。

◆ 二、逐层判词

本编把对他的评估分为五层：经验层、概念层、历史层、外推层、应用层。每一层分别写“他对在哪里”与“他的问题在哪里”。

经验层。 他对在：自然不是随机的，大量自然模式可学；AlphaFold、AI 天气模型、材料筛选，是这个时代真实的科学成绩，不是营销。本书第27章已经把这些成绩当成“积分能力的一次扩张”来承认，本编不打折扣。他的问题在：“可学”的对象，是已被测量、已被数学化的数据，不是物理界本身。他的证据全部来自理念物理界，他的结论却说的是“自然”。判词：**经验层成立，但须加上“在已被测量与数学化的范围内”这一限定；去掉这个限定，经验证据就不再支持结论。**

概念层。 他对在：把 P 与 NP 当作可能的物理问题来问，追问计算复杂性与物理定律的关系，方向有启发，也确实在推动研究。他的问题在：可计算与不可

计算，是信息九宫中数学态那一格的谓词，严格说是理念数学的谓词；问物理界本身“可不可计算”，是把谓词搬出了它的论域。判词：**概念层问错了层；在本书看来，这个错不是他独有的，彭罗斯式的“不可计算”论证犯的是同一个层次错置，只是方向相反。**

历史层。 （这一层所据的原话目前只有转述，见第28章；判词按下文失败条件第7条随原文核定而改。）他对在：至今确实没有在实验中直接撞见一个“不可计算的物理过程”。这句话在严格意义上大概为真。他的问题在：“从未发现”同时支持两种读法——“自然可计算”与“数学每次碰壁就扩充自己”——而他没有排除后一种；同时，已被严格证明的不可计算结果（普尔-埃尔—理查兹的波动方程、库比特等的谱隙、穆尔的动力系统）都出现在无穷精度、无穷尺寸、非光滑解这些理想化处。判词：**历史陈述为真，但不能作为自然可计算的证据；它主要是回写——123 原理第③步——一次次成功的记录。**

外推层。 这一层他本人没有完全走过去，传播版走过去了。从“自然模式可学”推到“没有 AI 不能解决的问题”，要跨过命名权与判错权两道门；跨过去的方式，是默认问题已被命名、对错已有标准。判词：**外推不成立；AI 的成功清单是一份“前两权已由人类结算”的问题清单，以它为样本推断“一切问题”，是选择偏差。** 值得记下的是，他本人用“发明相对论”作为灯塔，恰恰选中了一个命名权的检验——这说明他至少在直觉上知道外推层的门在哪里。

应用层。 他对在：AI 能替人做积分，而且做得越来越好，这一点本书全书都承认。他的问题在：应用要在具体一点上再微分，要把积分常数 C 从现场借回来——这一个病人、这一块地基、这一个路口；这一步的输入在现场，不在数据里。判词：**AI 替得了积分，替不了落点上的再微分；它能大幅缩短从现场到落点的距离，但最后那一段，要有人站在现场补上 C，并为它负责。**

> **五层判词一览** | 层 | 他对在哪里 | 问题在哪里 | 判词 | |—|—|—|—| | 经验 | 自然非随机，大量模式可学；AlphaFold、天气模型是真成绩 | 可学对象是已测量、已数学化的数据 | 成立，须加“在已数学化范围内” | | 概念 | 把 P/NP 当物理问题，方向有启发 | 可计算性是数学态的谓词，问物理界问错了层 | 问错层（与彭罗斯同错，方向相

> 反）| | 历史 | 实验中确未直接撞见不可计算过程 | “从未发现”可读作回写的战绩；已知不可计算结果都在理想化处 | 陈述为真，不能作证据 | | 外推 | （本人未完全走到）| 跳过命名权、判错权 | 不成立；成功清单有选择偏差 | | 应用 | AI 能替人做积分 | 再微分与 C 须在现场补上 | 替得了积分，替不了落点 |

◆ 三、替得了积分，替不了再微分

五层判词里，应用层的判词是全书主命题落在这一编的样子，值得多说几句，因为它最容易被误读成“AI 不行”。

不是 AI 不行。积分这一侧，AI 的能力已经超出了人。一个天气模型在几分钟里给出全球十天的预报，一个结构预测模型在几秒钟里给出一个蛋白质的三维结构，这些都是人类积分史上的跨越。说“AI 替得了积分”，是一个很高的评价。

问题在圈的另一端。积分给出的是 $f + C$；f 是整体的形状，C 是这一个点上那份积分找不回来的东西（第12章）。AI 把 f 算得越来越好，但 C 不在它的训练数据里，因为 C 就是“这一次”的那一份：这一个病人的过敏史和他今天早上没吃饭，这一片山谷今晚的风，这一块地基下面那一层没被勘探到的软土。再微分，就是在落点上把 C 补进来，得出“此刻该怎么做”。

大白话说：菜谱可以越写越好，好到全世界的厨师都照着做；但今天这锅汤咸不咸，要尝一口才知道。尝这一口，菜谱替不了。

有人会说：那就把现场数据也喂给 AI，让它连 C 一起算。这在很多地方正在发生，也是好事：天气模型接入实时观测，临床模型接入病人的个体数据。但请注意，这样做的每一步，都是在现场多做了一次测量、多截了一刀；被喂进去的 C，是被数学化之后的 C，筛子之下的那一份依然在现场。只要落点是一个具体的、不可重复的此刻，就总有一份东西是在这一刻才发生的，它要由站在这一刻的人去接。本书不主张这一份永远大，也不主张它永远重要；只主张它不会是零，而且由谁接、接错了谁负责，不是一个计算问题。

还要补上第三刀留下的一笔。再微分不仅要补 C，还会产生新的差异：病人对药的反应、地基在施工中的变形、骑手在巷子里留下的轨迹。这些新差异是下一轮微分的原料，是123 原理里推动 D—E 矛盾的那份新的 D。若落点上的再微分也被打包交给系统，系统只按自己的口径记录结果，那么新差异就会被预先过滤，循环就有被压平的危险（第32章）。所以“替不了落点”有两层意思：一层是做不到——C 在现场；一层是不该全交——落点是新差异进入循环的入口，入口要有人守。

◆ 四、三刀落在圈的哪一段

把三刀放回全书的那个圈里看，它们各自砍在不同的位置，合起来正好绕圈一周。

全书主命题说：科学是在现实里微分、在理念里积分、再回到现实里微分的一个圈。第一刀砍在第①段与第②段的交界：微分把现实截成数据，数据才进入理念；哈萨比斯的“可学”，只对截下来之后的东西有效。第二刀砍在第②段的内部：积分走向无穷，带来整体的清晰，也带来可判定性的丧失；“可计算”与“不可计算”都是这一段里的事，而这一段每次碰壁就扩充自己，留下“从未发现”的记录。第三刀砍在第③段与下一轮第①段的交界：落点上的再微分产生新差异，新差异要被命名、被判错，才能成为下一轮的原料；若这一环被默认交出，圈就有转成闭环空转的危险。

这样排开，就能看出一件事：三刀没有一刀砍在“AI 算得好不好”上。它们砍的都是圈上 AI 不在场、或不该独占的那几段。AI 住在第②段，而且住得越来越好；本编对他的批评，从头到尾不是“积分做得不够”，而是“只看见了积分这一段，就为整个圈下了结论”。

大白话说：一家饭馆的后厨换了一台极好的自动炒菜机。说它炒得好，是对的；说有了它“饭馆再没有解决不了的问题”，就忘了采购要到菜市场挑菜（第①段），忘了上菜前要有人尝一口、看看这桌客人是老人还是孩子（第③段），

也忘了客人吃完说“今天有点咸”时，要有人把这句话记下来、改进明天的菜单（下一轮）。

◆ 五、判词不能怎样用

判词写出来，也可能被误用。有三种误用要预先堵住。

第一种，是把本编当成“AI 无用论”。本编的每一层判词，都先写了“他对在哪里”。经验层的成立、应用层的“替得了积分”，是本编真心的判断。拿本编去否定 AlphaFold、否定 AI 天气模型，是读反了。

第二种，是把本编当成“人永远高于机器”的宣言。本编说的“现场”，不是人的特权，而是一个位置：站在具体一点上、承担这一点的后果、把这一点的新差异接住的位置。今天占据这个位置的主要是人；将来若有一种系统真能在这个位置上稳定地接住 C、接住新差异、并为错误负责，那么本章末“这一编怎样算错”的第6条就被触发了，本编会照单收窄。本编守的是一个位置的结构，不是一个物种的身份。

第三种，是把“可计算与不可计算都是理念层的判词”读成“科学说的都不算数”。恰恰相反：在理念层内部，可计算性理论的定理是严格的，谱隙问题的不可判定是一个真定理，AlphaFold 的准确率是一个真数据。本编撤掉的是把这些判词直接搬到物理界本身的那一步，不是这些判词本身。网捞上来的鱼是真鱼，称出来的重量是真重量；只是不能说“海里就是这些”。

这三种误用的共同点，是把一个分层的判断压成一个单层的态度。本编之所以坚持逐层下判词，正是为了不让它被压平。

◆ 六、他自己的灯塔，可以拿来对表

还有一件值得做的事：把他自己设下的检验，与本编的三刀对一对表。这样做，是让判词不只是本书单方面的意见，而能与他本人认可的标准接上口。

他在 莱克斯·弗里德曼（Lex Fridman）访谈中说，用截至 1900 年的知识训练一个系统，看它能否提出相对论，可以作为 AGI 的灯塔时刻（《代达罗斯》对谈中的版本是截止于 1910 年、推出广义相对论，见第28章）。这个检验好在哪里？好在 1900 年的数据里没有相对论。当时物理学摆在桌上的，是迈克耳孙—莫雷实验的零结果、麦克斯韦方程与伽利略变换的不相容、水星近日点的那一点反常——这些都是筛子边缘顶回来的东西（这几项在爱因斯坦本人思路里各占多大分量，科学史家说法不一）。按本书的三权来拆——这是一种理性重构，不是逐年的史实次序——狭义相对论的发生，第一步是把“同时性”重新命名为一个需要定义的问题（命名权），第二步是判定“以太漂移测不到”不是实验误差而是原理信号（判错权），第三步才是从两条原理推出洛伦兹变换的数学（计算权；变换式本身洛伦兹已先写出）。

所以，如果有一天真有系统通过了这座灯塔，它通过的方式本身就会回答本编的三个问题：它是从数据里学出了相对论，还是在数据的边缘发现了“这里该有一个新问题”？它是把反常当噪声平均掉了，还是把反常升格为原理？它是在现成的判错标准下赢了，还是自己立了一条新的判错标准并让共同体接受？若是后者，本编第三刀的失败条件（下文第5条后半）就被触发，判词应当照改。

大白话说：考一个学徒能不能出师，不是看他能不能把师傅教过的菜做好，而是看他碰到一道从没见过的食材时，能不能先说出“这东西该怎么对待”。他设的灯塔，本书完全同意；本书只是提醒，这座灯塔照亮的，恰好是计算权以外的那两块地。

◆ 七、本编的判词，自己也可能错

判词写完，轮到单据的第二行。

第395卷要求每个主张先写好怎样算错。本编对哈萨比斯下了五层判词，这些判词本身也是主张，而且分属不同类型：第一刀主要是概念规定加经验判断，第二刀主要是数学事实的解读加历史判断，第三刀主要是哲学解释加风险判断。类型不同，错法也不同；下面逐条写，不互相豁免。

还有一点必须申明：本编的锋芒针对的是论证，不是人。哈萨比斯是这个时代最认真对待“AI 与科学”关系的人之一，他提出猜想、给出理由、写下自我限定、设立灯塔检验，这本身就是可靠秩序的姿态——明确承诺、接受异常记录。本编若有一处把他没说过的话安到他头上，那一处就应当删去；第28章的原话整理，是本编所有判词的唯一合法靶子。

◆ 八、这一编怎样算错

1. **第一刀（第30章）**：若能系统地证明，被测量与数学化筛掉的残差中，不含任何与人的问题相关的结构化差异——即在一批覆盖不同领域的案例中，凡是现场人员报告“模型漏掉了什么”，事后核查都被证明是噪声或误判——则“可学的只是已被数学化的数据”这一刀失效，经验层判词应改为“成立，无须限定”。
2. **第一刀的概念部分（第30章）**：若可以构造一种不经任何形式化编码就能有意义地判定物理界“可计算与否”的程序（与第4章失败条件相同），则“可计算性只对数学态有定义”须收回，概念层判词随之作废。
3. **第二刀（第31章）**：若谱隙不可判定、波动方程不可计算或同类结果，被证明不依赖热力学极限、无穷精度初值、非光滑解等理想化——例如一块有限尺寸的真实材料，其可测量性质被证明在实验可达精度内不可判定——则“已知不可计算结果都住在理念最深处”这一论据作废；历史层判词须改写，本书须承认物理界层面可能存在不可计算性。
4. **第二刀的历史部分（第31章）**：若对数学物理史的系统梳理表明，主要的数学扩充（概率论、随机分析、算子理论、动力系统）大多并非由“现有数学处理不了的物理反常”触发，而是由数学内部问题独立驱动，则“从未发现是回写的战绩”须收窄为“部分是”。
5. **第三刀（第32章）**：若长期观察表明，AI 成为某领域主要的数学化执行者之后，该领域的残差多样性、异常报告数量、新问题类型的提出频率没有下降，甚至上升——则“矛盾被压平”的风险判断须撤回，或限定到具体配

置；若 AI 系统被证明能在没有现成判错条款的领域，独立提出并被现场共同体接受新的问题类型，则“命名权与判错权的最后一环在现场”须收窄。

6. **应用层（本章）**：若出现在没有现场初始条件输入时，仍能跨情境稳定完成落地决策、且在独立评估中没有系统性情境失败的 AI 系统（与第27章失败条件1相同），则“替不了落点上的再微分”须收窄为“在当前配置下尚未被替代”。
7. **对靶子本身**：若查得原始出处表明，“迄今没人发现不可计算的东西”一语并非哈萨比斯所说，或他说时另有本编未顾及的限定，则第31章对他的历史层判词须改为只针对这一说法本身，不再归于他名下；凡本编判词所据的原话，若与第28章核定的原文有出入，以原文为准，判词随之修改。

六条之外再加这一条，是因为本编所有判词都系在原话上；原话一松，判词就不该还挂在他身上。七条里，第3、6两条最接近能被近期研究检验；第1、5两条需要长时间、跨领域的记录才能判定，本书承认它们的检验成本很高，也承认在检验完成之前，相应判词只是有理由的判断，不是定论。

修车铺的单据写完了。下一编不再评他人，而是把同样的单据，递给本书自己和 SDE 本身：数学不是科学，SDE 也要接受检验。

带走的话

对哈萨比斯本人的猜想，本编的判词是分层的：经验层成立，但限于已被测量与数学化的范围；概念层把数学态的谓词搬到了物理界；历史陈述大概为真，却不能作自然可计算的证据；外推到“没有 AI 不能解决的问题”不成立，那是传播版的事；应用层上，AI 替得了积分，替不了落点上的再微分。每一层判词都写明了怎样算错。

小 账

- 本章修车铺、炒菜机饭馆、学徒出师等均为构想场景，无真实原型。
- 本章所判的原话以第28章核定的出处为准；“迄今没人发现不可计算的东西”一语据转述，相应判词随原文核定而改。
- 本章不主张 AI 无用，不主张人永远高于机器，也不否认可计算性理论各定理与 AlphaFold 等成绩本身的真实性。
- 本章判词针对论证，不针对哈萨比斯其人。
- 本章七条失败条件依次对应附录G第42、13、43、44、45、4、46条。

第 七 编

边界：数学不是科学，SDE 也要接受检验

编 序

第七编只做一件事：把前六编用来评估别人的那张单子，递给本书自己。

第六编要求哈萨比斯的论断写明怎样算错，这一编要求本书和 SDE 做同样的事。一本讨论边界的书，自己首先要有边界；一套讲“发生”的框架，也要说清楚它在哪里只是在整理语言。

第34章讲数学的三态与三化：符号、逻辑、数学三态是往返而不是阶梯；数学是理念界的场态，受定义与证明约束，不受实验裁决，所以它不是科学本身；标准化、优化、简化三化各有主要侧重。微积分在这里的位置很特殊：不定积分末尾那个“+C”，把“落点上必须从现实借回一份”原原本本写进了符号里。第35章讲本书怎样算错：把全书主张分为概念规定、哲学解释、数学命题、经验判断、实践建议五类，每一类写明各自的错法，连 SDE 框架本身也要交一张保修单；已经错过、已经收回的地方，也一并登记。

读这一编时，可以带着一个问题：一个不可能输的主张，还说了什么？

34

第 34 章

数学的三态与三化

> 这一章的问题：奶奶的“枣皮色”、菜谱的“30 克”、博主的焦化曲线说的是同一锅糖，哪一种才算数学，数学又凭什么不等于科学？

周末学做红烧肉。奶奶在电话里说：“糖下锅，小火炒，炒到冒小泡、颜色像枣皮那样，赶紧下肉。”菜谱 App 上写的是：“冰糖 30 克，中小火 3 分钟。”一位美食博主更进一步，画了一条曲线：蔗糖在约 160 至 170 摄氏度开始焦化，温度每高十度，颜色变化和苦味都会明显加快，所以火候的关键不在时间，而在锅里那几十度的温度窗口。

三种说法，说的是同一锅糖。奶奶那句话最“土”，却最好用：她看的是泡和颜色，锅厚锅薄、火大火小都不怕，她的眼睛会自己校准。菜谱那句最整齐，换一口薄锅就可能炒糊。博主那条曲线最“数学”，可真到了灶台前，没有温度计的人还得把它翻译回“冒小泡、枣皮色”。

那么，哪一种才是“数学”？数学和“会做菜”这门实打实的本事是什么关系？再推一步：既然本书前面三十多章都在说科学离不开微积分，数学是不是就等于科学本身？

底本第十三章对最后一问的回答斩钉截铁：数学不是科学。本章接着这句话往下走，但要先做两件事：把这句话里站得住的部分留下，把底本当年的说法升级到现行口径；然后把它放回本书的主命题里，看它在“微分—积分—再微分”的圈里到底站在哪个位置。

◆ 一、底本说了什么，本书保留什么

底本的论证可以压成三句。第一，科学研究“存在如何发生”，数学研究“发生之后差异之间的连接、次序与结构”；科学追问发生，数学刻画结构。第二，数学对它的对象做三件事：标准化、优化、简化，底本合称“三化”。第三，三化不是随意的形式游戏，它受意义三律约束。底本还用了一个好比喻：科学研究一条河怎样流、为什么改道；数学研究河道的几何、水系的拓扑、流量的分布。

这三句里，第一句的方向本书保留，但要换掉它的底座。底本的底座是 SIO 整体与差异单元 ΔSIO，现行口径不再以此作主术语（对照见附录H）。第二句的“三化”本书保留，它确实描出了数学工作的三种常见动作。第三句要收窄最多：底本把三化与“创生律、演化律、成就律”一一对应，说数学的三化就是意义三律在形式领域的投影。现行口径下，意义三律是特征律、自由律、幸福律，它们与任何一组三分对象之间，至多只有“主要侧重”，严格对应没有证明过。本章第四节会把这一点落到实处。

还有一处要先说明。底本把数学叫作 SIO 的“场”研究，这个直觉是对的，现行口径把它放进了更细的格子里：数学态是信息的一种模态，而且每一界都有。要讲清“数学不是科学”，得先讲清“数学态”是什么。

◆ 二、三态：符号、逻辑、数学，往返而非阶梯

第4章已经铺过信息九宫：现实界、理念界、自我界，每一界都有信息；每一界的信息又有三种模态，符号态、逻辑态、数学态。这里只取本章要用的部分。

符号态像一颗攥紧的粒子：一个可以被指认、保存、传递的单位。“枣皮色”就是一颗符号，它把颜色、气味、泡的大小、锅铲手感一并攥进一个词里。好的符号是可返回的压缩：省掉重复劳动，又留着通往细节的入口。奶奶说“枣皮色”，你下次站在灶台前，能把这个词重新展开成眼前那锅糖。

逻辑态像一道推进的波：一个信息怎样激发下一个，怎样获得理由。“冒小泡了，说明水分快炒干了，再等就要发苦”——这是一串推移，有快慢，也有卡住的时候。现实界有现实逻辑（锅里的事一件推一件），理念界有理念逻辑（从定义推到定理），自我界有自我逻辑（“知道该关火，手却慢了半拍”）。

数学态像一片场：许多信息之间的连接、次序与共同约束，被当作一个整体把握。博主那条温度—颜色—苦味曲线，就是把一大片经验之间的关系，摊成一张同时成立的结构图。

现行口径对这三态有两条硬纪律。

第一条，**三态不构成单线等级**。底本和早期材料里常有“从符号到逻辑再到数学，一级比一级高”的读法，已被收回。符号态不是低级碎片，而是“饱满的一颗”；奶奶那句“枣皮色”在灶台前比曲线更好用，这不是因为奶奶落后，而是因为在“此刻把这锅糖炒好”这项任务里，那颗符号恰好最合用。

第二条，**三态之间是往返**。成熟的认识在凝结与展开之间来回走：曲线要落到灶台，得重新压回一颗符号（“看到枣皮色就停”）；一条定理要被检查，得重新展开成一步一步的逻辑推移；一串零散的经验要被传授，又得被提到数学态，找出它们之间的共同约束。没有哪一态是终点。

> **信息九宫与三态的判别** 三界 × 三态构成九格：现实符号（如“这棵树”）、理念符号（如“素数”）、自我符号（如“我”）；现实逻辑、理念逻辑、自我逻辑；现实数学、理念数学、自我数学。　三态的判别不看内容“难不难”，而看此刻怎样被处理：被当作一个单位调用，是符号态；被当作一段推移来走，是逻辑态；被当作一组同时成立的连接、次序与约束来把握，是数学态。同一个公式 $F = ma$，背下来当口诀用时是符号态，用它一步步推出落体时间时是逻辑态，把它放进全部力学量的相互约束里看时是数学态。　两条限定：其一，“有且只有三态”是便于分析的约定，穷尽性尚未论证；遇到第四种候选，不能改个名字塞回三格。其二，“数学态”是扩展用语，不等于形式数学；进入形式数学，还需要对象、定义、公理和可检查的推导。

这条“往返”纪律，对理解微积分很要紧。教科书里的微积分，常被当成纯粹的理念数学，学生学的是一张场图。可牛顿那一代人手里的微积分，一直在三态之间来回：“流数”是一颗新符号，割线变切线是一段推移，行星轨道是一片约束。第7章讲过，莱布尼茨的 dx 与 $\int$ 之所以好用，正在于它们是好符号，可返回的压缩做得极好。

◆ 三、数学是理念界的场态，不是科学本身

有了三态，再来看“数学不是科学”。

本书说的“数学”，若指数学家在做的那门学问，它的主要栖身之处是理念界的数学态：在明确的定义与公理之下，研究连接、次序、结构，并给出可检查的推导。它当然也离不开理念符号与理念逻辑，往返是常态；但它的产品，最终以“在这些假设下，这些关系必然成立”的形式交付。

科学则不在任何一格里。按本书的主命题，科学是一个圈：在现实界微分（截取一点的增量，切掉周围的纠缠），在理念界积分（把增量拼成整体模型），再回到现实界一个具体的点上再微分（借回积分常数 C，算出此刻该怎么做）。这个圈横跨两界，至少在“截取”和“落点”两处必须碰到现实。数学只占了圈的理念一侧，而且只是那一侧里的一种模态。

所以“数学不是科学”，在本书里有一个比底本更精确的说法：**数学是科学之圈在理念一侧借用的场态，它本身不负责与现实相撞。**

这句话可以用一件小事验证。欧几里得几何与非欧几何，作为数学，至今都“对”：各自在自己的公理下成立，没有哪个实验能推翻其中一个。但“我们所在的物理空间在大尺度上是哪一种几何”，是一个科学问题，要靠测量来回答，而广义相对论的回答是：在引力场附近，不是欧几里得的那一种。被实验修正的是物理理论，不是几何定理。数学只给出“如果……那么……”，科学要回答“现实里是不是这个如果”。

第395卷把这个分别写成三层：一是现实活动中的关系；二是选来表示这些关系的模型；三是明确公理与规则下的形式结构。三层一混，就会出两种错。

> **三层不混与四条原则**　例：三个小组、两间教室，规定任意两组不得同时同室。作为形式结构，这是一个无解的分配问题，三组无法放进两间且互不相同。　错误一：把模型的不可能当成现实的不可能。现实里也许可以错开时段、借用走廊，规定本身可以改。数学说明“在这些约束下不可启动”，不决定删哪条规定。　错误二：把现实的可变当成定理的失效。规定改了，那是换了一个模型，原来的定理依旧成立。　四条原则：定理问假设与证明；模型问对应与误差；方法问适用与代价；哲学主张问概念与论证。跨层时，必须写出桥梁。

这个三层框架也解释了本书最关心的那个谓词。“可计算”“不可计算”“可判定”“不可判定”，都是在形式结构那一层才有定义的谓词，它们属于理念界数学态那一格。第30章对哈萨比斯的第一刀，正是从这里下的：问“物理界本身可不可计算”，是把一个只在一格里有定义的谓词，搬到了整张九宫之外去用。

还要防一个反方向的误读。说数学不是科学，不是说数学低于科学，更不是说数学可以不顾现实。数学的独立性恰恰是它对科学最大的贡献：正因为几何定理不随一次测量而倒，物理学家才能拿它当一把稳定的尺，去比较“现实更像哪一种”。一把会跟着被量的东西一起伸缩的尺，什么也量不出来。

◆ 四、三化：标准化、优化、简化

底本说数学对它的对象做三件事，本书把它们放回发生学里重说一遍，并且给每一化都记上代价。

标准化：把千差万别的差异，稳定成可以共同使用的单位与规则。结绳、刻痕到阿拉伯数字，“一把”“一捧”到“克”与“毫升”，都是标准化。菜谱的“30 克”就是标准化的产物，它让一千个厨房能对同一句话。代价是：标准化在稳定同一性的同时，划掉了它认定“不影响同样处理”的差异。30 克没写你家锅的厚度，因为在标准化的那一刻，锅厚被判为不相干。

优化：在已有的表达里，找更简洁、更普遍的那一种。罗马数字到位值记数，几何作图到代数方程，一堆特殊公式收拢成一个通式。优化让人省力，也常给人一种“走通了”的畅快，数学家称之为优美。代价是：越简洁的表达，对使用者的预备知识要求越高；一条优美的通式，在不懂它的人手里比十条笨公式更难用。第395卷提醒，优化不是越省越好：省掉的冗余，可能正是必要的备份；还要问省掉的是谁的成本。

简化：剥离具体经验，只留下关系本身。几何从田亩走到公理，代数从具体的数走到群、环、域，拓扑连长短都不要了，只留下“连着还是断开”。简化让数学能够走出它出身的那块地，被搬去回答原本没有的问题。代价最大，也最隐蔽：被剥离的经验，从此不在数学里了；要用回现实，就得在落点上一项项补回来。这正是本书所说积分常数 C 的另一张面孔。

三化放在一起看，会发现它们与本书第三编讲的“微分”是同一类动作：每一化都是一次截取，每一次截取都在稳定某些东西的同时，丢下另一些东西。数学的威力和数学的边界，出自同一把刀。

> **三化与意义三律：只作“主要侧重”** 底本：创生律为标准化提供原料，演化律推动形式演进，成就律使数学体系化、稳定化，三化因此是“意义三律在形式领域的投影”。现行口径下改写为一个较弱的说法： 标准化，主要侧重特征律，即在差异中稳定出可以被指向的同一性； 简化，主要侧重自由律，即剥离原有经验之后，为原本不存在的新路腾出空间，非欧几何、抽象代数都是这样长出来的； 优化，主要侧重幸福律，即一次表达终于走通时的畅快，数学家所谓“优美”。 限定：这只是“主要侧重”，不是对应。反例俯拾即是：阿拉伯数字的标准化同时开出了大量新算法，这是自由律的事；简化之后一套新理论的诞生，也会带来走通的畅快，这是幸福律的事。幸福律“张力—转换—释放”本身只是机制候选。若在一批数学史案例中，“非主要侧重”的情形与“主要侧重”的情形一样多，这组侧重就应撤回。

把三化交给意义三律去约束，真正有用的不是那张对照表，而是一个提问方向：一次标准化稳定了谁的同一性？一次简化打开了哪条新路，又关掉了哪些旧

路？一次优化让谁走通了，让谁被挡在了门外？这三问可以拿到任何一次数学化面前去问，包括哈萨比斯所说的“把自然建模”。

◆ 五、独立而不自足

底本在“数学不是科学”之后，紧接着补了一句：数学也不能完全独立，它的原料来自差异的发生，它的发展靠科学与文明提出新需求。这句话本书保留，但要拆成两层说，因为两层的证明责任不同。

第一层是**发生层**：数学的概念是怎样在人那里发生的。孩子学数，先摸苹果、掰手指、一个个点过去，“三”才慢慢从三种感官的纠缠里被攥成一颗符号；几何的点线面，也有它在视觉与动作里稳定下来的经过。底本附录C、D、E、F 对自然数、几何、概率、极限的发生学描述，属于这一层。这一层是哲学解释加经验判断，要靠儿童认知研究、数学史材料去检验，不能单凭讲得通就算成立。

第二层是**身份层**：一个数学对象在形式系统里是什么、一条定理凭什么成立。这一层只认定义、公理与推导。“1 + 1 = 2”在皮亚诺公理下成立，与哪个孩子怎样学会它无关；一个证明是否正确，与证明者是否有“豁然贯通”的感受无关。第395卷明确收回了“证明共振即成立”的旧说：理解的体验可以是一次真实的发生，但不是证明成立的充分条件。

两层分开，“独立而不自足”就有了确切的意思。数学在身份层上独立：它的定理不随实验倒下。数学在发生层上不自足：它的概念从人与世界的纠缠里长出，它的新分支常常被现实的难题逼出来。傅里叶为了算热传导，逼出了三角级数，又逼出了后来整整一门分析学的严格化（见第9章、第17章）；概率论最早是被赌桌和保险账本逼出来的。

把这两层混在一起，会出两种相反的错。一种是把发生层当身份层：以为讲清了“数”在孩子那里怎样发生，就等于说清了数是什么，于是数学定理的有效性好像也要跟着人的经验走。第395卷对它自己早先（该卷第50章）那套“1 是连接性、0 是断”的说法，就作了这样的提醒：认知形成、符号约定、公理设定、

形式证明，是四件事。另一种是把身份层当发生层：以为定理是永恒的，于是数学的概念好像从来就在那里，与人怎样学会、为什么需要它毫无关系，数学教育也就只剩“记住、套用、刷题”。底本后记里那场持续几十年的困惑，有一部分正出在这里。

这里还要提一处与本书主题相连的地方：**封顶**。某一界的信息停在符号态或逻辑态，没有被提到数学态。老中医的手感、老工匠的分寸，本地逻辑饱满，却没有被提取成可传授的理念结构。现行口径对“封顶”有严格限定：它是一个局部、可撤回的诊断，只能用于具体的技艺、课程或知识传递过程，禁止用来整体评判一个文明。工匠难以口述，可以靠示范传下去，不能因此判他缺逻辑；修复的办法可以是建模，也可以是示范、练习、制度调整，不必都改写成公式。这条限定在第8章讨论李约瑟之问时已经用过，本章只是把它放回它出身的地方。

◆ 六、微积分站在哪里

现在可以回答开头那锅糖。

奶奶的“枣皮色”，是自我界与现实界交界处的一颗饱满的符号，背后是几十年的现实逻辑；菜谱的“30 克、3 分钟”，是一次标准化，把奶奶的经验压成了可共享的单位，也划掉了锅厚；博主的曲线，是把一大片经验提到了数学态，它最普遍，也离灶台最远。三者没有高下，只有往返：曲线要落地，就得重新变回“看颜色”；奶奶的手艺要传下去，最好有人把它提成曲线。这一来一回，就是科学之圈在一口锅里的缩影。

微积分在这个圈里的位置很特殊。作为数学，它是理念界数学态里的一门学问，受定义与证明约束，不受实验裁决；作为科学的工具，它又恰好是那个最贴近“截取—累加—落点”节奏的数学：导数就是一次形式化的截取，积分就是一次形式化的累加，而不定积分末尾那个 $+C$，把“落点上必须从现实借回一份”这件事，原原本本写进了符号里。底本称微积分为“科学的心脏”，本书把这句话收窄为：在动力学科学中，微积分是“微分—积分—再微分”这套看世界

方式最集中、最纯粹的数学器官。它不等于科学，它让科学的那个圈在理念一侧转得最顺。

本章的几处判断，都有可能错：三态的“往返”可能只是一个好听的描述，三化的“侧重”可能只是事后对号，“数学是理念界场态”可能把数学家的实际工作说得太窄。哪些情况算它们错，下一章连同全书的其他主张一起，逐条写出来。

带 走 的 话

数学的主要栖身处是理念界的数学态：在定义与公理之下研究连接、次序与约束，交付“如果……那么……”。它在身份层上独立，定理不随实验倒下；在发生层上不自足，概念从人与世界的纠缠里长出。科学则是横跨两界的圈，必须在截取与落点两处碰到现实。三化的每一化都是一次截取，稳定一些东西，也丢下一些东西；微积分把这件事写得最清楚，那个 $+C$ 就是证据。

小 账

- 本章奶奶、菜谱、美食博主炒糖的场景，以及“三个小组、两间教室”的分配问题，均为构想的说明性例子，无真实原型。
- 本章唯一的数据是蔗糖焦化温度（约 160 至 170 摄氏度），各来源不一，已作限定，出处等级见附录I第 I51 条；底本与第395卷的说法均据原文转述。
- 三化与意义三律只作“主要侧重”，不主张严格对应；本章也不主张数学低于科学，更不用“封顶”评判任何文明。
- 本章主张的推翻条件见附录G第 G48、G49、G50 条（三态往返、三化侧重、“数学是理念界的场态”）。

35

第 35 章

本书怎样算错

> 这一章的问题：热水器的保修单事先写明哪种坏法算谁的错，一本说了不少重话的书，能不能也写出这样一张单子，写清哪种情况下它算错？

家里换热水器，安装师傅干完活，递过来一张保修单。好的保修单不写“质量保证、放心使用”这种话，它写的是：三年内内胆漏水，免费换；电路板烧了，免费修；但若是水压超过多少、没装减压阀，或者自己拆过，不在保修范围。这张纸的分量，不在它夸了热水器多好，而在它事先写清楚了：出了哪种事，算谁的错。

一本书也该有这样一张保修单。本书提出了不少判断，有些说得很重：科学的“自然界”是筛过的；微分切掉的纠缠积分找不回来；AI 替得了积分、替不了落点。读到这里，读者有权问一句：这些话，哪种情况下算错？

各编末尾已有“这一编怎样算错”的短节，本章把它们收拢，并补上贯穿全书、却没有在哪一编单独交代的那几条。全书主张的总清单，已收进附录G；本章做的，是把它们按性质分开，因为不同性质的话，错法不一样。

◆ 一、五类主张，五种错法

第395卷立过一条总纪律：概念规定、哲学解释、数学命题、经验判断、实践建议，五类主张不能互相豁免。哲学的气势替代不了证明，形式的漂亮替代不了验证，个案的成功也证不成普遍的本体论。

这条纪律用大白话讲，就是看病时分清几种话。“我们把这种情况叫作二型糖尿病”是命名；“你这病主要是吃出来的”是解释；“空腹血糖高于某个数值就算”是一条按定义成立的判定规则；“这个药能把血糖降一成”是经验判断；“你每天走一万步”是建议。命名不存在真假，只有好不好用；解释要能比别的解释多说明点什么；判定规则在定义内必然成立，却不保证定义选得对；经验判断要靠数据；建议要看照做以后有没有用、代价谁出。把建议说成定理，或者拿一个漂亮的解释去冒充数据，都是常见的毛病。

> **五类主张的证明责任**　概念规定：问能否多出可检验的区分。失败形式：只换了名字，判断与旧说法完全相同。　哲学解释：问能否比竞争解释多说明什么，能否面对独立材料。失败形式：任何结果出现后都能事后调换说法，从而永远不错。　数学命题：问假设与证明。失败形式：证明有误，或条件被悄悄放宽。　经验判断：问数据、测量与误差。失败形式：预言的相关不出现，或出现了但能被更简单的因素解释。　实践建议：问照做后是否改善、代价由谁承担。失败形式：照做无改善，或改善由别人付出代价。　时间纪律：失败条件须在观察前写出；事后补充的条件，一律标为“修订”，不得悄悄并入原主张。

下面逐类写。每条都按“若……则……”的格式，写清哪种情况出现时，本书哪句话要收回或收窄。

◆ 二、概念规定：若只是改名，就承认增量有限

本书造了或借了几个词：理念物理界、积分常数 C 的外借、再微分、不成文的积分、信息九宫。概念规定本身无所谓真假，它的错法是“没有用”。

理念物理界（第3章）。它说科学家口中的“自然界”，是物理界经过测量、理想化、剔除噪声后剩下的那部分，数学是筛子。这个想法并不新。爱丁顿的渔网已经说过一半；南希·卡特赖特在《物理定律是怎样说谎的》（1983）里也论证过，基础定律只在被理想化的条件下严格成立；“理想化”“模型”“其他条件不变”在科学哲学里各有成熟的讨论。若本书用到“理念物理界”的每一处，换成“经过理想化的模型世界”都不损失任何区分，那么这个词只是换了个说法，

本书应承认它的增量有限，不宣称首创。本书认为它多出了一处区分：筛子有**回流通道**，反常是物理界顶回来的地方，而且只顶在筛孔的位置上。这个“只顶在筛孔上”能否被独立检验，是这个概念的生死线。

再微分（第14章）。若它与工程学里早已成熟的“标定”“现场调试”“个体化”说的是同一回事，没有多出任何判断，本书就把它降为一个统称，并在第14章注明各领域已有的叫法。

不成文的积分（第26章）。它说 AI 把千亿次梯度累加进权重，这是积分，只是没有写成人能读的理念方程。若这个说法所能说出的一切，“黑箱”“缺乏可解释性”都已说过，它就只是一个比喻。本书认为它多出的东西是位置：被藏起来的是积分，而两端的微分（训练的梯度、应用的落点）仍然暴露在外。这个位置判断若推不出任何与“黑箱”说法不同的预期，就应收回。

信息九宫（第4章）。它的风险在“九”这个数。若出现第四种信息模态的候选，或某些信息无法安放进任何一界，本书不得把它改个名字塞回九格，而应承认三界与三态都只是便于分析的约定。

◆ 三、哲学解释：若永远不会错，就是没说什么

全书主命题是一个哲学解释：科学是在现实里微分、在理念里积分、再回到现实里微分的圈；微分切掉的纠缠，积分找不回，只能在落地时从现实借回。

它有一个危险的优点：太好讲了。什么科学成功都可以说“积分做得好”，什么失败都可以说“C　没借回来”。一个解释若能在任何结果出现后都讲得通，它就没有经受检验。所以这一段只写那些让它有可能输的地方。

若一门学科的成功主要发生在“其他条件不变”切得最狠的地方，而它的失败并不集中在初始条件、边界条件与情境依赖上，比如主要败在纯粹的推导错误、或测量仪器的系统误差上，那么“C　外借律”就失去了它最主要的支撑，应降为一个修辞。

若能找到一类成功的应用，它在落点上没有借用任何来自现场的信息，既没有初始条件、也没有边界条件、也不针对这一个对象作任何调整，却跨情境稳定地奏效，那么“再微分不可替代”就要收窄为“多数应用需要再微分”。本书目前想到的最接近的候选是纯数学的应用，如密码算法；但密码系统部署时仍要选参数、配密钥、防侧信道，本书倾向于认为它们仍有落点。这一判断是否成立，留给读者核查。

若回写只能被宣布、不能被观察，即本书所说的每一次“123 原理的回写”（第16章：反常逼出理论改变，改变回写测量与问题），都找不到可辨认的痕迹，例如仪器、课程、期刊栏目、测量规程的具体变化，那么第16章应删去“回写”一词，只保留库恩式的描述。回写也可以是负向的，本书若只挑正向的例子，读者应把这当作挑样本。

筛子论还有一处可能失守：若理念物理界与物理界之间的“回流”，事实上从不受筛孔形状影响，而是均匀地从各处涌现，那么“反常是从筛孔顶回来的”就说错了。

◆ 四、框架本身的错法：SDE 也要交保修单

上面几条针对本书的具体判断。还有一层更深的：本书用来组织这些判断的 SDE 框架，三方程、六路径、123 原理，本身也可能没有用。

这里先说清楚本书对框架的使用层次。三方程 $S = F(D, E)$、$D = G(S, E)$、$E = H(S, D)$ 在本书里只在概念层使用：F、G、H 是关系的占位记号，本书从未给出它们的变量、函数形式和时间尺度，也没有用它们推出过任何数值。第15章把微积分循环放进三方程下讨论，是一种组织方式，不是建模。

在这个层次上，框架的错法是“多余”。

若把全书的 S、D、E 词汇统统删掉，只用“模型”“数据”“情境”“反馈”这些普通科学哲学的词，每一章的结论都原封不动，那么 SDE 在本书中的增量有限，本书应如实承认：它帮作者组织了提问，但没有为读者多出可检验的区

分。按第395卷的说法，SDE 分析与普通清单结论相同时，就承认增量有限；已有成熟名称的机制，不换名，不宣称首创。

若同一个现象，本书在这一章说它是“D 的推进”，在另一章说它是“E 的沉淀”，而两种说法引出的预期毫无差别，那就是在随结果调换 S、D、E 的位置。这样的段落，读者可以直接划掉。

若 123 原理在本书中被用于“反常—危机—回写”之外的任何三步流程，那是本书违反了自己的纪律。123 原理只讲 D—E 矛盾逼出 S 改变、S 的改变回写 D 与 E 这一个循环；把“微分、积分、再微分”这三段说成 123 原理，是一个容易犯的错，本书无意如此，凡出现这种写法，以本条为准更正。

还有一句底本的大话，本书已经收窄，这里再写一次它的失效条件。底本说“微积分是科学的心脏”。本书改为：在动力学科学中，微积分是“微分—积分—再微分”这套看世界方式最集中的数学器官。**若能指出一门成熟的、以预测和干预见长的现代科学，其核心理论完全不依赖连续变化的数学，而依赖的是离散、组合或统计的工具**，例如某些分子生物学、遗传学或计算机科学的分支，那么“心脏”一说须限定在动力学科学之内，不得外推到“全部科学”。本书倾向于认为这样的学科确实存在，所以这条限定从一开始就该写着。

◆ 五、数学命题：只问假设与证明

本书的数学命题不多，而且大多是借来的。它们的错法最清楚，也最不容含糊。

积分常数。在一个连通区间上，两个原函数至多相差一个常数，这是定理。本书在它之上所做的，“C 是微分切掉的那一份纠缠”，不是数学命题，而是哲学解释；它若被驳倒，定理分毫不损。读者不应把定理的确定性，转借给本书的解释。

伞场的局部命题（第20章）。本书只主张：在核非负、偶对称、可积、紧支集、尺度归一、二阶矩为正、被作用函数在该点附近二阶连续可微这些条件下，一类非局域算子 $A_r u(x) = r^{-2}\int \varphi(z)[u(x+rz)-u(x)]dz$ 在 $r \to 0$ 时趋于

$\frac{M_2}{2}u''(x)$。这是一条可以独立复核的局部命题。若复核发现证明有误，此条撤回；若有人把它扩大为“传统数学是伞模型的极限或特例”，那不是本书的主张，第395卷已明确撤回。条件少一条，结论就可能改变，第20章方框里的三点限制已经说明。

不可计算结果（第31章）。本书引用波动方程可计算初值导出不可计算解（Pour-El 与 Richards）、量子多体谱隙不可判定（Cubitt 等）一类结果，并作了一个读法：它们出现在无穷极限、实数连续统、非光滑解这些理想化之处。读法不是定理，它的失效条件见第33章第3条。

◆ 六、经验判断：数据说了算

本书中有些话，归根到底是关于世界的事实判断。它们只能由数据来驳。

复现危机的检验（第12章）。“C　外借律”给出了一个可检验的预言：越是靠“其他条件不变”切得狠、越依赖情境的研究，复现越困难。已有研究报告过，专家评定的情境敏感度与复现失败存在关联：范·巴维尔（Van　Bavel）等2016 年在《美国国家科学院院刊》上，对“心理学复现项目”的一百项研究重新评分，情境敏感度与复现成功的相关约为　−0.23，在控制了原始效应量、样本量、统计功效等因素后仍然显著。

但这个结果随即受到挑战：因巴尔（Inbar）同年在同一刊物上指出，社会心理学研究的情境敏感度普遍高于认知心理学，复现率也普遍更低；把子学科这一项控制住之后，关联基本消失，两个子学科内部的相关都接近于零。原作者回应说情境敏感度与子学科的交互并不显著，争论尚无定论。此外，这条检验还有别的对手：样本量小、发表偏倚、激励扭曲，同样能解释复现率的高低。**若控制了样本量与发表偏倚之后，情境依赖与复现失败的关联消失**，本书应承认 C 外借律在这个检验上没有增量。**若在情境依赖很低的领域里同样大规模复现失败**，此律应收窄。

AI 的再微分（第27章）。失效条件已见第27章第1条与第33章第6条，这里不重写，只补一个检验上的陷阱：同一底座训练出的多个系统彼此一致，不构成独立证据。

AI 积分是否“不成文”。若可解释性研究或符号回归能够常规地从训练好的模型里读出人能检查的方程，“不成文”就只是暂时状态，第26章要改写为“尚未成文”。

李约瑟之问（第8章）。本书只把数理解释当作候选之一，与制度、经济、地理解释并列。**若能独立说明，在制度、经济、地理条件相近的社会之间，数理传统的差别仍能预测科学发展的差别**，数理解释的权重可以上调；**若一旦控制制度与经济因素，数理解释不再多说明任何东西**，它就应降为对结果的重新描述，也就是第8章说的循环风险。

对哈萨比斯的判词（第30—33章）。三刀与靶子本身的七条失效条件，第33章已逐条写出，这里不重抄，只按本章的五类把它们重新归队，因为归队之后才看得出各条该由谁来判。第1条（残差里有无与人相关的结构化差异）和第6条（有无不借现场初值的稳定落地）是经验判断，要靠跨领域的案例记录；第2条（不经形式化能否判定可计算）是对概念规定的挑战，要看能否构造出那样的程序；第3条针对的是对数学命题的读法，由数学物理学家判；第4条是史学上的经验判断，由数学史的系统梳理判；第5条前半是经验判断，后半（AI 能否独立立题并被现场接受）牵涉哲学解释，最容易被事后调换说法，读者应格外盯住；第7条针对的是文献事实，只需核对原始音频与原书。分类的用处在于：一条失效条件若被拿错了裁判，比如用一个漂亮的解释去回应一条经验判断，那就是在豁免它。

◆ 七、实践建议：照做以后呢

本书也提了建议，而且有分量：微积分教学应先讲它为什么存在，再讲怎样算；AI 时代，人要守住落点上那次再微分、那份 C。

建议的错法是：照做了，没变好；或者变好了，代价落到别人身上。

教学建议。若在一段公开声明的时间窗口内，比如一个学年，先讲意义、后讲算法的微积分课，学生在新情境中运用导数与积分的能力，并不比常规课程更好，反而占用了练习时间，那么这条建议应收回或改写。只增加术语负担的课程，按第395卷的说法，应修改或撤回。检验的时间窗口须事前定好，不得在结果不合预期时延长。

“守住落点”。若人坚持亲自做再微分，结果比交给系统更差，例如在某类医疗或工程决策里人的现场判断反而系统性地拖低结果，那么这条建议应限定在“系统确已在该类情境中经过独立检验之前”。守住落点不是守住面子；守住的是对后果的承担与判错的权利，而不是每一次亲手按下按钮。

◆ 八、已经错过的地方

保修单写的是将来；还有一份更难写的清单，是已经错过的。本书由底本扩写而来，底本里有几句话，在本书里已经收回或收窄。把它们列在这里，不是为了表态，而是为了让读者知道这本书的口径是怎样一步步改出来的。

- 底本序言说，没有微积分，就没有把差异与整体统一起来的机制，也就没有现代科学革命。本书收窄为：在动力学科学中如此；作为李约瑟之问的答案，它只是候选之一（第8章）。
- 底本说微分是差异的“忠实抄写”。本书改为：微分是一次截取，截取即丢弃；忠实只是在被截取的那一点上忠实（第10章）。
- 底本把微分、积分、扩张对应创生律、成就律、演化律，说微积分就是这三律的数学化。本书不再把这一对应当结论用，只在附录H里作术语对照。
- 底本说积分“必然在新的差异中崩塌”。“必然”一词本书不取：许多积分模型在其适用范围内长期稳定，崩塌的只是被越界使用的那部分。
- 底本说数学的三化“是意义三律在形式领域的投影”。本书降为“主要侧重”（第34章）。

这份清单的作用，在于给出一个比较的基线：下一版若还要再改，读者可以对照着看，改的是哪一类主张，改的理由是新证据、新论证，还是只是换了说法。

◆ 九、这一编怎样算错

第七编自己也要接受检验。下面几条针对第34章与本章：

- 若“符号、逻辑、数学三态往返”在数学史案例中找不出任何与“单线上升”读法不同的预期，第34章的“往返”只是换了一种说法，应降为描述性用语。
- 若在一批数学史案例中，三化与意义三律的“非主要侧重”情形与“主要侧重”情形一样常见，第34章的侧重表应撤回。
- 若“数学是理念界的场态”把数学家的实际工作说得过窄，例如实验数学、计算机辅助证明中，现实界的机器运行已是证明不可分割的一部分，第34章须承认数学与现实的接触不止在“发生层”。
- 若本章列出的失败条件，事后被一条条解释为“条件不同所以不算”，那么本章本身就是一张不兑现的保修单，读者应据此降低对全书的信任。
- 本书由 AI 写作组据底本分章扩写，经统稿、核验与三审成书（见出版信息中的 AI 协作声明）。写作过程的复杂与篇幅的完整，不证明任何命题正确；同一套口径下写出的各章彼此一致，也不构成相互独立的证据。

这张单子不会完整。真正的保修单之外，总有师傅没想到的坏法。写下它的意义，是让下一个发现坏法的人，知道该在哪一条上打叉。

最后一章不再立新论，只把这个圈收回到起点：科学、微积分，与生活在其中的人。

带 走 的 话

本书的主张分五类，各有各的错法：概念规定错在只是改了名，哲学解释错在永远讲得通，数学命题错在证明或条件，经验判断错在数据，实践建议错在照做无益或代价转嫁。失败条件必须在观察之前写下，事后补的一律标为修订；一条条事后解释掉，这张单子就不兑现。框架本身也在其列：删去 S、D、E 而结论不变，就承认增量有限。

小 账

- 本章热水器保修单、看病时的几种话，均为说明性的日常比方，不对应真实个案。
- 本章唯一的实证数字是范·巴维尔等 2016 年报告的相关约 −0.23，并列了因巴尔的反驳与原作者回应，出处见附录I第 I15、I52 条。
- 本章不新立论点，只把各章、各编已写的失败条件按五类归队；全部条目登记于附录G。
- 本章自身的推翻条件见附录G第 G52 条（失败条件被事后逐条解释掉）；第七编的其余条件见 G48 至 G51 条。

结语　科学、微积分与文明

一半来自积分，一半要靠路口

傍晚开车回家，导航说：“前方两百米右转。”你右转了。导论用开车讲过那个圈：速度记录是微分，累加成路程是积分，出发点是积分常数　C，导航落到路口的那一句是再微分。三十多章走下来再听这句话，能听出它的两半：一半来自千万辆车被截取、被累加的增量，一半要靠此刻的路口。导航给出的是　f，路口给的是 C。

全书的话，说到底就是这一句：科学是在现实里微分、在理念里积分、再回到现实里微分的一个圈；微分切掉的那一份纠缠，积分找不回来，只能在落地的那一刻从现实借回来；可计算性只属于圈的理念一侧；而 AI 能替人做积分，替不了落点上的那一次再微分。

结语不立新论，只做三件事：把这个圈收回来，看它说了什么、没说什么；把作者为微积分正名的那段心事交代清楚；再问一句，在 AI 越来越会做积分的年代，人站在哪里。

收回那个圈

第一段，**微分**，发生在现实界。它在流动的现实里截一刀，取出某一点的增量，并宣布“其他条件不变”。这一刀是科学的入口，也是科学的第一笔代价：被截下的那一点周围，原本缠绕着的千丝万缕，一刀下去都不在了。第10章说“截取即丢弃”，说的就是这个。买菜时只看斤两和单价，账算得清清楚楚，代价是这把菜是不是早上刚摘的、摊主是不是老熟人，都不在账上。

第二段，**积分**，发生在理念界。它把一个个增量累加起来，拼成整体：轨道、定律、模型、预测。这是科学最有创造力的一段，牛顿与莱布尼茨以来三百多年的辉煌，大半在这里。但积分有一个从课本第一页就写着、却很少被认真对

待的特点：不定积分的结果是 $f + C$。C 是什么，积分本身说不出来。被微分切掉的那一份，积分无论多精巧都找不回来（第11、12章）。

第三段，**再微分**，又回到现实界。整体模型要落到一个具体的点上，算出“此刻该怎么做”：这座桥用多粗的钢，这个病人用多大的剂量，这块地基打多深的桩。工程、医疗、技术、应用，都发生在这里。也只有在这里，那份 C 才被补上：初始条件、边界条件、这一个病人、这一块地（第14章）。而且再微分不是把模型原样搬下来就完事，它在现实里留下新的差异，新的差异又成为下一轮微分的原料。123　原理所讲的回写，在这个圈上就是这样发生的：反常累积，逼出理论改变，改变又重置了测量什么、追问什么（第16章）。

这个圈还有一个底座，第一编花了四章去铺：科学家口中的“自然界”，并不是物理界本身，而是物理界经过测量、理想化、剔除噪声之后剩下的那一部分，本书称之为理念物理界。数学是那张筛子。爱丁顿的渔夫说海里没有短于两英寸的鱼，错不在他的网，而在他把网的规格写进了海的定律。筛子并不是封死的：水星近日点多出来的那一点进动、黑体辐射在紫外端的那一段曲线，都是物理界从筛孔里顶回来的地方，每一次顶回，都逼着筛子改织一次。所以本书从不说“科学抓不住现实”，只说科学抓住的是筛过的现实，而现实会从筛孔里顶回来。微分是下筛的那一刻，积分是在筛上的世界里拼图，再微分是带着拼好的图回到筛下，那份 C，就是筛子当初漏掉、如今必须当场补回的东西。

这三段放在一起，本书给出了几个检验得了的推论。科学的失败，应当集中在初始条件与情境依赖上，而不是集中在推导里；越是靠“其他条件不变”切得狠的学科，复现越难。“可计算”“不可计算”是理念一侧形式结构里的谓词，拿它去问整个物理界，问错了层。AI 把千亿次梯度累加进权重，做的是积分，只是没有写成人能读的方程；它两端的微分，训练时的梯度和应用时的落点，仍然暴露在现实里。上一章已经把这些推论的失效条件一条条写下。一个圈若永远不会错，就什么也没说；本书宁可让它有可能输。

还有几句话，要从底本那里收回。底本说“没有微积分，就没有现代科学”，本书限定为动力学科学；底本说李约瑟之问的答案在一个数理事实，本书

只把它当作与制度、经济、地理并列的候选，并承认它有循环的风险，也承认中国算学在算法一路上的极高成就（第8章）；底本把微积分与创生律、成就律、演化律一一对应，本书不再把这当结论。收窄不是退却。一句话说得越大，能落到实处的部分就越少；收窄之后，留下来的才是能扛事的。

三百年走过的路

用这个圈回头看全书走过的路，会发现每一段历史都可以读作圈上某一段的故事。

阿基米德用穷竭法逼近抛物线弓形的面积，刘徽割圆，祖暅说“幂势既同，则积不容异”，这些是积分的先声：在理念里把无数小块拼回整体（第6章）。牛顿与莱布尼茨的贡献，不只是发明了两种运算，而是让微分与积分在一个基本定理里焊在了一起（第7、13章）。焊接有条件，焊缝也一直在：被积函数要足够规整，差异要已经稳定。贝克莱讥笑无穷小是“消逝量的幽灵”，柯西与魏尔斯特拉斯用极限语言把它重新立住（第9章）；这是理念一侧的一次自我修补，回答了“怎样算才算严格”，没有回答“切掉的那一份去了哪里”。

此后三百年，微积分不断扩张：从一个点的运动到一整片场的偏微分方程，从解析解到离散的数值计算，从光滑的轨道到布朗运动那样处处不可微的路径，从只看此刻的导数到带着记忆的分数阶，从有限个变量到无穷维的函数空间（第17—21章）。每一次扩张，都是现实在某处顶回来，逼着理念一侧长出新的形式。用 123 原理的话说，这是反常累积、逼出改变、改变又回写测量与问题的一轮轮循环；用本书的话说，这是筛子一次次被现实顶出新的网眼，又一次次被织补。混沌理论则提醒了另一件事：即使方程完全确定，可计算也不等于可预测，初值那份 C 上差之毫厘，轨道就谬以千里（第22章）。

到了工程、控制与 AI，这个圈的第三段终于站到了台前。工程是把积分好的模型落到一根梁、一座坝上；控制论把再微分做成了实时的反馈，每一刻都在现场重新借一次 C；而 AI 把积分做到了前所未有的规模，却把积分本身藏进了权重

（第23—27章）。三百年的故事，从“怎样把差异拼成整体”，走到了“拼成的整体怎样落回现实、由谁来落”。这正是本书问题意识的落点。

为微积分正名

底本的后记，是全书最切身的一段。底本作者王德生在后记里写道：从高中第一次接触微积分，到大学里讲授微积分，他竟用了几十年，才真正明白微积分的价值与意义。他自己说，这是一件近乎滑稽的事，却也是一整代教育的深刻悲剧。

这段经历值得在这里完整地留下，因为它不是一句谦辞，而是本书问题意识的来源。课堂上的微积分是导数的定义、换元和分部积分、微分方程的解法、泰勒展开和傅里叶级数。这些都重要，但都停在“怎样算”的层面。学生一遍遍被要求解题，却很少有人告诉他们：微积分为什么存在，它在科学和文明里意味着什么。底本说得更重：老师自己也不清楚，因为整个教育把微积分降成了一种训练，而不是一种理解世界的语言。

本书要为这段话补上一层理解，也加一个限定。

补上的一层是：为什么“几十年才明白”几乎是必然的。微积分的符号是极好的压缩。$\int$　一个记号，攥紧了“切成无数小段、一段段累加、取极限”这一整套动作；$+C$　两个字符，攥紧了“积分找不回微分丢掉的那一份”这一整个认识论问题。好的符号是可返回的压缩，可是返回的入口，要有人带着走一遍才打得开。课堂只教人使用这颗符号，不带人返回，学生就会熟练地写　$+C$，却从没想过它在说什么。用第34章的话说，符号、逻辑、数学三态之间本该有的往返，在课堂上被切成了单程。

加的限定是：本书不主张“先讲哲学、后讲计算”就一定教得更好。会算是明白的底子，没有亲手算过一百道题的人，对导数的“截取”与积分的“累加”很难有手感。上一章已经写明，这条教学建议要在事先定好的时间窗口里接受检验；若只增加了术语负担，就应修改或撤回。为微积分正名，正的是它的名分，而不是要把学生从练习里拉出来听讲座。

那么，这个名分是什么？本书的回答是：微积分是人类把“截取—累加—落点”这个认识世界的节奏写成符号的第一次成功。它不是全部科学，但它让科学这个圈在理念一侧转得最顺；它不等于理解世界，但它的每一个符号都提醒着理解世界要付出什么代价。导数提醒你切掉了什么，积分提醒你拼出的是你自己的整体，$+C$ 提醒你：整体落回现实时，还欠着现实一份。

底本后记的最后一句，是一份献给后来学子的邀请：愿你们比我更早地看见那轮太阳。本书愿意把这句话原样接过来，只把“太阳”换成一个更朴素的东西：愿你们早一点看见那个 C。

几十年与几分钟

今天的学生，也许会对这段经历有一个不同的反应：几十年？现在问一句 AI，几分钟就能得到一篇“微积分的哲学意义”，讲得头头是道。

这正是本书要在这里多说几句的原因。一篇讲得头头是道的解释，是别人做好的积分：把三百年的科学史、几代人的反思，累加成一段流畅的文字。它有用，也可能比本书讲得更清楚。但它落到你身上时，仍然缺一份 C。

那份 C 是什么？是你自己算过的那一百道题里，第一次发现导数“在这一点上不存在”时的那一愣；是你第一次用积分算出一个真实物体的体积，和用水去量的结果对上了、或者没对上时的那一下；是你在物理课上写完一个微分方程，老师问“初始条件是什么”，你忽然意识到前面所有推导都悬在半空的那一刻。这些都是学习者自己身上的再微分，别人替不了，AI 也替不了。第395卷有一句话说得很准：入库不算发生，回写才算；要回写，先相撞。

所以，底本作者的几十年，不能被几分钟替代；能被缩短的，是他当年缺少的那个“有人带着返回”的入口。好的老师、好的书、好的 AI 助手，都可以是这个入口。但入口之后的那段路，仍然要学生自己的手、自己的错，一步步走过去。若 AI 只是把现成的理解递过来，学生收下了一段漂亮的文字，却从没有与自己的底盘相撞，那么这几分钟得到的东西，和当年课堂上背下的公式一样，是一本放进图书馆的新书：馆藏多了一册，图书馆还是那个图书馆。

这也是本书对“为微积分正名”最终的理解：正名不是换一套更动听的说法，而是让每一个学习者有机会在自己身上，把那个圈完整地走一遍。

文明：一个圈怎样变成生活方式

底本的结语，把微积分从科学推到文明：微分的忠实让人类对差异保持敏感，积分的发明让人类能创造整体的世界图景，微积分的扩张让文明不断突破边界。这个方向本书保留，但要换一种更落地的讲法，因为“文明”这个词太大，大到什么都能装进去，也就什么都没说。

本书说的文明，指的是一个圈从实验室里走出来，沉淀成千万人日常生活的样子。

先看微分怎样进入生活。现代人习惯了“单位时间的变化”：车速是每小时多少公里，工资涨幅是每年百分之几，孩子的身高曲线每个月体检都要描一个点，股票软件上满屏都是涨跌的斜率。三百年前，绝大多数人并不这样看世界。“变化率”从牛顿的草稿纸走进菜市场和体检表，本身就是一次漫长的回写：一种数学，重置了一个社会测量什么、在意什么。

再看积分怎样进入生活。养老金是一辈子缴费的累积，碳排放总量是千万根烟囱的累积，一张信用评分是你几百次还款记录的累积。现代社会的许多制度，本质上都是一个积分装置：把零散的增量拼成一个整体数字，再拿这个数字去决定资源怎样分配。

然后是再微分，以及它的代价。一个整体数字落回到一个具体的人身上时，那份 C 常常没有被借回来。信用分不知道你那几个月逾期，是因为家里有人住院；一个地区的平均收入涨了，不等于街口那家修鞋铺过得更好了。文明越依赖积分装置，就越需要在落点上有人肯看一眼现场。这件事不是数学能替我们做的，也不是数学的错；它是那个圈对一个文明提出的要求。

所以，本书对“微积分与文明”的判断，比底本谦逊，也更具体：微积分不只是科学的器官，它的节奏，截取、累加、落点，已经沉淀成现代文明的一种生

活方式。这种生活方式的好处有目共睹，它的病也有一个统一的样子：积分做得越来越好，落点上的 C 越来越没人借。

底本曾说，这三者对应意义三律在文明层面的显现：特征律让人把握世界的秩序，自由律让人扩展可能的空间，幸福律让人在张力的化解中获得意义。本书只把这当作“主要侧重”来读，并把它变成三个问题，留给每一个积分装置：它稳定了谁的同一性？它为谁开出了新路，又关掉了谁的旧路？它让谁走通了，又让谁的那份 C 被永远记成了零？

底本序言和结语都用了三句话收束：微积分是太阳，特征纠缠是土壤，SIO 三原理是大地的脉动。这三句话写得有气势，本书不再沿用，因为它们把一个需要检验的主张，写成了一幅不需要检验的画。如果非要留一句，本书愿意换成一句朴素的：微积分是人类第一次把“看世界要付出什么代价”写进了符号里。

AI 时代人的位置

最后回到导论里那条短视频：“没有 AI 不能解决的问题。”

本书第六编已经把哈萨比斯本人的命题与这句传播版分开，替他把话说足，又分三刀评估。这里不再重复，只把三刀合起来，落到一个普通人的问题上：如果 AI 越来越会做积分，我还剩下什么？

剩下的，恰恰是圈的两端。

一端是**落点**。AI 可以把千万份病历积分成一个诊断模型，但眼前这一个病人，他今早吃过什么，他说“还好”时眼神里是什么，他家里有没有人能照看他吃药，这些是现场的 C，只有在现场的人才借得到。AI 可以替你规划一条路线，但前方那辆车突然打了双闪，是你的脚踩下刹车。守住落点，不是守住每一次亲手按按钮；守住的是：在落点上，有一个人看见了现场，借回了那份 C，并且为后果承担。

另一端是**命名**。什么算一个问题，由谁来说？被数学化筛掉的那些残差，很多时候连“问题”的身份都拿不到。一个孩子为什么不肯写作业，一片老街区拆

掉以后少了什么，一种手艺失传了丢的是什么，这些东西若从来没被命名，就永远不会出现在 AI 的训练数据里，也就永远不在“可解决的问题”之列。第32章称之为命名权、判错权与计算权的分离。计算权可以越来越多地交给机器；命名权和判错权，若也悄悄交出去，人就会生活在一个只剩“可算的问题”的世界里，而且不会察觉。导航普及以后，地图上没有的小巷，正在从人的经验里慢慢消失。

所以本书对 AI 时代人的位置的回答，可以压成一句话：**守住落点上的那一次再微分，守住那一份 C，也守住为问题命名、判定对错的那份权利。** 这不是对 AI 的敌意。AI 替人做积分，是三百多年来这个圈转得最快的一次，本书真心为它高兴。但圈之所以是圈，是因为它要回到现实；回不到现实的积分，再精巧，也只是理念界里的一张图。

这句话也有它的失效条件，上一章已写：若有一天，某个系统在没有任何现场输入的情况下，跨情境稳定地完成落地决策，没有情境性失败，本书这句话就要收窄。本书不知道那一天会不会来。本书只知道，在它来之前，把落点交出去的代价，由落点上的人承担。

给三种读者的一句话

这本书预想过三种读者，各留一句话。

给正在学微积分的学生：每写一次 $+C$，停一秒，问自己它代表什么。多数时候它只是一个习题里的常数；但它也是整本微积分里最诚实的那个符号，它承认自己不知道。

给做工程、做医疗、做一切“应用”的人：你们做的不是“把理论用一下”。你们站在圈上最难的那一段，在一个具体的点上替整个模型还债。模型给不出的那份 C，是你们的手艺。

给做 AI 的人：你们让积分快了几个数量级，这是真的成就。请把两端的微分也当回事：训练数据是谁、在什么情境下截取的，切掉了什么；模型落到谁身上，谁来借回那份 C，谁来判它错了。积分藏起来不要紧，两端不能也藏起来。

余下的路

一本书写到最后，最诚实的做法，是指出它没走完的路。

“理念物理界”与“回流通道”需要更具体的科学史案例，去检验反常是否真的只从筛孔顶回来。“C　外借律”需要与复现研究的数据正面相撞，而不是只引一两篇文献。“不成文的积分”需要与可解释性研究对话：如果有一天，模型的积分可以被系统地写成人能读的方程，本书第26章要改写。信息九宫里的“数学态”，离形式数学还有很长的距离；伞场的局部命题，离一个成熟的建模框架也还远。SDE　框架本身，要在具体研究里证明它比普通的科学哲学词汇多出可检验的区分，否则它只是帮作者想清楚了问题，并没有替读者多回答什么。

这些未完成的地方，是本书的边界，也是它留给读者的入口。科学、微积分与文明的关系，不会在一本书里结算完毕；它本身就是那个圈的一部分：截取，累加，落回，再截取。

明天早上出门，导航还会说“前方两百米右转”。那时候，也许你会想起，这句话一半来自千万次的积分，一半要靠你此刻看一眼路口。

参考书目

以下列出本书论证中承担证据或边界作用的著作、论文、演讲与访谈，分七组，组内按作者姓氏字母（中文按拼音）排列；正文仅作线索性提及、未承担关键论证的材料不在此表中。题录信息凡未能核全者，只录可确认的部分，并以〔〕注明。

一、哈萨比斯言论与相关报道

1. Barrington, J. (2026, May 23). The Infinity Machine by Sebastian Mallaby (2026) [Book review]. Substack. https://johnbarrington.substack.com/p/the-infinity-machine-by-sebastian
2. Fridman, L. (Host). (2025, July). #475 – Demis Hassabis: Future of AI, simulating reality, physics and video games [Audio podcast episode]. Lex Fridman Podcast. Transcript: https://lexfridman.com/demis-hassabis-2-transcript
3. Fry, H. (Host). (2025, December 16). The future of intelligence [Podcast episode, with Demis Hassabis]. Google DeepMind: The Podcast. 〔无公开文字稿；正文所引原话据 Walczak（2026）转述〕
4. Hassabis, D. (2024, December 8). Accelerating scientific discovery with AI [Nobel lecture and slides]. Aula Magna, Stockholm University. Nobel Prize Outreach. https://www.nobelprize.org/prizes/chemistry/2024/hassabis/lecture/
5. Hassabis, D., & Manyika, J. M. (2026). AI as the ultimate tool for science: A conversation with Demis Hassabis. Dædalus, 155(1–2), 34–44. https://doi.org/10.1162/DAED.a.971
6. Mallaby, S. (2026). The Infinity Machine: Demis Hassabis, DeepMind, and the Quest for Superintelligence. Penguin Press. 〔本书未核原书，相关内容据 Barrington（2026）书评转述〕
7. Pelley, S. (Correspondent). (2025, April 20). [Interview with Demis Hassabis]. 60 Minutes. CBS. 〔节目段落题名未核；CBS News 网站配发文章题为 “Artificial intelligence could end disease, lead to ‘radical abundance,’ Google DeepMind CEO Demis Hassabis says”〕

8. Walczak, T. (2026, January 9). Demis Hassabis is wrong about computability. Substack. https://tomwalczak.substack.com/p/demis-hassabis-is-wrong-about-computability

二、数学与物理

1. Abel, N. H. (1826). Untersuchungen über die Reihe: $1 + (m/1)x + (m(m-1)/(1\cdot2))x^2 + \ldots$. Journal für die reine und angewandte Mathematik, 1, 311–339.
2. Aharonov, Y., & Bohm, D. (1959). Significance of electromagnetic potentials in the quantum theory. Physical Review, 115(3), 485–491.
3. Anderson, B. D. O. (1982). Reverse-time diffusion equation models. Stochastic Processes and their Applications, 12(3), 313–326.
4. Anderson, P. W. (1972). More is different. Science, 177(4047), 393–396.
5. Bachelier, L. (1900). Théorie de la spéculation. Annales scientifiques de l’ École Normale Supérieure, 3e série, 17, 21–86.
6. Bagley, R. L., & Torvik, P. J. (1983). A theoretical basis for the application of fractional calculus to viscoelasticity. Journal of Rheology, 27(3), 201–210.
7. Barrow, I. (1670). Lectiones geometricae. London: William Godbid.
8. Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. Nature, 525(7567), 47–55.
9. Bausch, J., Cubitt, T. S., Lucia, A., & Pérez-García, D. (2020). Undecidability of the spectral gap in one dimension. Physical Review X, 10, 031038.
10. Berkeley, G. (1734). The Analyst; or, A Discourse Addressed to an Infidel Mathematician. London: J. Tonson.
11. Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637–654.
12. Cameron, R. H. (1960). A family of integrals serving to connect the Wiener and Feynman integrals. Journal of Mathematics and Physics, 39, 126–140.
13. Caputo, M. (1967). Linear models of dissipation whose Q is almost frequency independent—II. Geophysical Journal of the Royal Astronomical Society, 13(5), 529–539.
14. Cardona, R., Miranda, E., Peralta-Salas, D., & Presas, F. (2021). Constructing Turing complete Euler flows in dimension 3. Proceedings of the National Academy of Sciences, 118(19), e2026818118.
15. Cauchy, A.-L. (1821). Cours d’ analyse de l’ École royale polytechnique. 1re partie: Analyse algébrique. Paris: Debure.

16. Cavalieri, B. (1635). Geometria indivisibilibus continuorum nova quadam ratione promota. Bologna: Clemente Ferroni.
17. Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibrations. Bulletin of the American Mathematical Society, 49(1), 1–23.
18. Courant, R., Friedrichs, K., & Lewy, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. Mathematische Annalen, 100, 32–74.
19. Cubitt, T. S., Pérez-García, D., & Wolf, M. M. (2015). Undecidability of the spectral gap. Nature, 528(7581), 207–211.
20. Deutsch, D. (1985). Quantum theory, the Church–Turing principle and the universal quantum computer. Proceedings of the Royal Society of London A, 400(1818), 97–117.
21. Du, Q., Gunzburger, M., Lehoucq, R. B., & Zhou, K. (2012). Analysis and approximation of nonlocal diffusion problems with volume constraints. SIAM Review, 54(4), 667–696.
22. Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik, 17(8), 549–560.
23. Einstein, A. (1915). Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften, 831–839.
24. Fefferman, C. L. (2006). Existence and smoothness of the Navier–Stokes equation. In J. Carlson, A. Jaffe, & A. Wiles (Eds.), The Millennium Prize Problems (pp. 57–67). Clay Mathematics Institute & American Mathematical Society.
25. Feynman, R. P. (1948). Space-time approach to non-relativistic quantum mechanics. Reviews of Modern Physics, 20(2), 367–387.
26. Fourier, J. (1822). Théorie analytique de la chaleur. Paris: Firmin Didot.
27. Galilei, G. (1623). Il Saggiatore. Roma: Giacomo Mascardi.
28. Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. Princeton University Bulletin, 13, 49–52.
29. Heath, T. L. (Ed. & Trans.). (1897). The Works of Archimedes. Cambridge University Press. 〔另见同编者 1912 年增补本 The Method of Archimedes〕
30. Heath, T. L. (Ed. & Trans.). (1908). The Thirteen Books of Euclid’s Elements (3 vols.). Cambridge University Press.
31. Itô, K. (1944). Stochastic integral. Proceedings of the Imperial Academy, Tokyo, 20(8), 519–524.

32. Jaffe, A., & Witten, E. (2006). Quantum Yang–Mills theory. In J. Carlson, A. Jaffe, & A. Wiles (Eds.), The Millennium Prize Problems (pp. 129–152). Clay Mathematics Institute & American Mathematical Society.
33. Kac, M. (1949). On distributions of certain Wiener functionals. Transactions of the American Mathematical Society, 65(1), 1–13.
34. Kepler, J. (1609). Astronomia nova. Heidelberg.
35. Laplace, P.-S. (1814). Essai philosophique sur les probabilités. Paris: Courcier.
36. Lax, P. D., & Richtmyer, R. D. (1956). Survey of the stability of linear finite difference equations. Communications on Pure and Applied Mathematics, 9(2), 267–293.
37. Lebesgue, H. (1902). Intégrale, longueur, aire. Annali di Matematica Pura ed Applicata, Serie III, 7, 231–359.
38. Leibniz, G. W. (1684). Nova methodus pro maximis et minimis, itemque tangentibus. Acta Eruditorum, 3, 467–473.
39. Le Verrier, U. J. J. (1859). Lettre de M. Le Verrier à M. Faye sur la théorie de Mercure et sur le mouvement du périhélie de cette planète. Comptes rendus hebdomadaires des séances de l’ Académie des Sciences, 49, 379–383.
40. Lorenz, E. N. (1963). Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 20(2), 130–141.
41. Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. Journal of the American Statistical Association, 44(247), 335–341.
42. Moore, C. (1990). Unpredictability and undecidability in dynamical systems. Physical Review Letters, 64(20), 2354–2357.
43. Neumann, J. von. (1932). Mathematische Grundlagen der Quantenmechanik. Berlin: Springer.
44. Newton, I. (1687). Philosophiae naturalis principia mathematica. London: Joseph Streater (for the Royal Society).
45. Paley, R. E. A. C., Wiener, N., & Zygmund, A. (1933). Notes on random functions. Mathematische Zeitschrift, 37, 647–668.
46. Planck, M. (1900). Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum. Verhandlungen der Deutschen Physikalischen Gesellschaft, 2, 237–245.
47. Pour-El, M. B., & Richards, I. (1981). The wave equation with computable initial data such that its unique solution is not computable. Advances in Mathematics, 39(3), 215–239.
48. Pour-El, M. B., & Richards, J. I. (1989). Computability in Analysis and Physics. Berlin: Springer.
49. Robinson, A. (1966). Non-standard Analysis. Amsterdam: North-Holland.

50. Silling, S. A. (2000). Reformulation of elasticity theory for discontinuities and long-range forces. Journal of the Mechanics and Physics of Solids, 48(1), 175–209.
51. Tao, T. (2016). Finite time blowup for an averaged three-dimensional Navier–Stokes equation. Journal of the American Mathematical Society, 29(3), 601–674.
52. Toomer, G. J. (Trans.). (1984). Ptolemy’s Almagest. London: Duckworth.
53. Tucker, W. (1999). The Lorenz attractor exists. Comptes Rendus de l’Académie des Sciences, Série I, Mathématique, 328(12), 1197–1202.
54. Turing, A. M. (1952). The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London B, 237(641), 37–72.
55. Weierstrass, K. (1895). Über continuirliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen. In Mathematische Werke (Vol. 2, pp. 71–74). Berlin: Mayer & Müller. (Original lecture delivered 1872)
56. Weihrauch, K., & Zhong, N. (2002). Is wave propagation computable or can wave computers beat the Turing machine? Proceedings of the London Mathematical Society, 85(2), 312–332.
57. Wheeler, J. A. (1990). Information, physics, quantum: The search for links. In W. H. Zurek (Ed.), Complexity, Entropy, and the Physics of Information (pp. 3–28). Addison-Wesley.
58. Wiener, N. (1923). Differential-space. Journal of Mathematics and Physics, 2, 131–174.
59. Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H., & Emanuel, K. (2019). What is the predictability limit of midlatitude weather? Journal of the Atmospheric Sciences, 76(4), 1077–1091.

三、科学史与科学哲学

1. Cartwright, N. (1983). How the Laws of Physics Lie. Oxford: Clarendon Press.
2. Cohen, H. F. (2010). How Modern Science Came into the World: Four Civilizations, One 17th-Century Breakthrough. Amsterdam University Press.
3. Dijksterhuis, E. J. (1961). The Mechanization of the World Picture (C. Dikshoorn, Trans.). Oxford: Clarendon Press. (Original work published 1950)
4. Eddington, A. S. (1939). The Philosophy of Physical Science. Cambridge University Press.
5. Elvin, M. (1973). The Pattern of the Chinese Past. Stanford University Press.
6. Feyerabend, P. (1975). Against Method: Outline of an Anarchistic Theory of Knowledge. London: New Left Books.

7. Huff, T. E. (1993). The Rise of Early Modern Science: Islam, China, and the West. Cambridge University Press.
8. Husserl, E. (1954). Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie (W. Biemel, Ed.; Husserliana VI). Den Haag: Martinus Nijhoff. 中译本：胡塞尔：《欧洲科学的危机与超越论的现象学》，王炳文译，北京：商务印书馆，2001 年。
9. Jones, E. L. (1981). The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia. Cambridge University Press.
10. Kuhn, T. S. (1962). The Structure of Scientific Revolutions. University of Chicago Press.
11. Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), Criticism and the Growth of Knowledge (pp. 91–196). Cambridge University Press.
12. Lavoisier, A.-L. (1789). Traité élémentaire de chimie. Paris: Cuchet.
13. Lin, J. Y. (1995). The Needham puzzle: Why the Industrial Revolution did not originate in China. Economic Development and Cultural Change, 43(2), 269–292.
14. Mikami, Y. (1913). The Development of Mathematics in China and Japan. Leipzig: B. G. Teubner.
15. Needham, J. (1959). Science and Civilisation in China, Vol. 3: Mathematics and the Sciences of the Heavens and the Earth. Cambridge University Press.
16. Needham, J. (1969). The Grand Titration: Science and Society in East and West. London: George Allen & Unwin.
17. Ossendrijver, M. (2016). Ancient Babylonian astronomers calculated Jupiter’s position from the area under a time-velocity graph. Science, 351(6272), 482–484.
18. Penrose, R. (1989). The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford University Press.
19. Planck, M. (1931, October 7). [Letter to R. W. Wood]. 〔原信藏处与刊印出处未核；正文所引“一次绝望之举”（ein Akt der Verzweiflung）为通行引语〕
20. Pomeranz, K. (2000). The Great Divergence: China, Europe, and the Making of the Modern World Economy. Princeton University Press.
21. Popper, K. R. (1959). The Logic of Scientific Discovery. London: Hutchinson. (Original work published 1934)
22. Royal Society. (1712). Commercium epistolicum D. Johannis Collins, et aliorum de analysi promota. London: J. Pearson. 〔扉页署 1712 年，1713 年初分发〕
23. Sachs, A. J., & Hunger, H. (1988–1996). Astronomical Diaries and Related Texts from Babylonia (Vols. 1–3). Wien: Verlag der Österreichischen Akademie der Wissenschaften.

24. Sivin, N. (1982). Why the Scientific Revolution did not take place in China—or didn’t it? Chinese Science, 5, 45–66.

四、人工智能与 AI for Science

1. Allen, A., Markou, S., Tebbutt, W., Requeima, J., Bruinsma, W. P., Andersson, T. R., Herzog, M., Lane, N. D., Chantry, M., Hosking, J. S., & Turner, R. E. (2025). End-to-end data-driven weather prediction. Nature, 641(8065), 1172–1179. https://doi.org/10.1038/s41586-025-08897-0
2. Ameisen, E., Lindsey, J., Pearce, A., Gurnee, W., et al. (2025, March 27). Circuit tracing: Revealing computational graphs in language models. Transformer Circuits Thread. https://transformer-circuits.pub/2025/attribution-graphs/methods.html
3. Bricken, T., Templeton, A., Batson, J., et al. (2023). Towards monosemanticity: Decomposing language models with dictionary learning. Transformer Circuits Thread.
4. Cheetham, A. K., & Seshadri, R. (2024). Artificial intelligence driving materials discovery? Perspective on the article: Scaling deep learning for materials discovery. Chemistry of Materials, 36(8), 3490–3495. https://doi.org/10.1021/acs.chemmater.4c00643
5. Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2018). Neural ordinary differential equations. In Advances in Neural Information Processing Systems 31 (NeurIPS 2018).
6. Degrave, J., Felici, F., Buchli, J., et al. (2022). Magnetic control of tokamak plasmas through deep reinforcement learning. Nature, 602(7897), 414–419.
7. EMBL-EBI. (n.d.). pLDDT: Understanding local confidence. In AlphaFold: A practical guide [Online training course]. https://www.ebi.ac.uk/training/online/courses/alphafold/ 〔无发布日期〕
8. Fawzi, A., Balog, M., Huang, A., et al. (2022). Discovering faster matrix multiplication algorithms with reinforcement learning. Nature, 610(7930), 47–53.
9. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778). (Preprint arXiv:1512.03385, 2015)
10. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In Advances in Neural Information Processing Systems 33 (pp. 6840–6851).
11. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

12. Jayatunga, M. K. P., Ayers, M., Bruens, L., Jayanth, D., & Meier, C. (2024). How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons. Drug Discovery Today, 29(6), 104009. https://doi.org/10.1016/j.drudis.2024.104009
13. Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589.
14. Karvonen, A. (2024). Emergent world models and latent variable estimation in chess-playing language models. arXiv:2403.15498. 〔正式发表信息未核〕
15. Lam, R., Sanchez-Gonzalez, A., Willson, M., et al. (2023). Learning skillful medium-range global weather forecasting. Science, 382(6677), 1416–1421.
16. Leeman, J., Liu, Y., Stiles, J., Lee, S. B., Bhatt, P., Schoop, L. M., & Palgrave, R. G. (2024). Challenges in high-throughput inorganic materials prediction and autonomous synthesis. PRX Energy, 3(1), 011002. https://doi.org/10.1103/PRXEnergy.3.011002
17. Li, K., Hopkins, A. K., Bau, D., Viégas, F., Pfister, H., & Wattenberg, M. (2023). Emergent world representations: Exploring a sequence model trained on a synthetic task. In International Conference on Learning Representations (ICLR 2023).
18. Linnainmaa, S. (1970). Algoritmin kumulatiivinen pyöristysvirhe yksittäisten pyöristysvirheiden Taylor-kehitelmänä [Master’s thesis]. University of Helsinki.
19. Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. Nature, 624(7990), 80–85.
20. Moult, J., Pedersen, J. T., Judson, R., & Fidelis, K. (1995). A large-scale experiment to assess protein structure prediction methods. Proteins: Structure, Function, and Bioinformatics, 23(3), ii–iv.
21. Olah, C., Cammarata, N., Schubert, L., Goh, G., Petrov, M., & Carter, S. (2020). Zoom in: An introduction to circuits. Distill, 5(3), e00024.001.
22. Price, I., Sanchez-Gonzalez, A., Alet, F., et al. (2025). Probabilistic weather forecasting with machine learning. Nature, 637(8044), 84–90.
23. Robbins, H., & Monro, S. (1951). A stochastic approximation method. The Annals of Mathematical Statistics, 22(3), 400–407.
24. Romera-Paredes, B., Barekatain, M., Novikov, A., et al. (2024). Mathematical discoveries from program search with large language models. Nature, 625(7995), 468–475.
25. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536.
26. Silver, D., Huang, A., Maddison, C. J., et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484–489.

27. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. In Proceedings of the 32nd International Conference on Machine Learning (PMLR 37, pp. 2256–2265).
28. Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2021). Score-based generative modeling through stochastic differential equations. In International Conference on Learning Representations (ICLR 2021).
29. Szymanski, N. J., Rendy, B., Fei, Y., et al. (2023). An autonomous laboratory for the accelerated synthesis of novel materials. Nature, 624(7990), 86–91.
30. Templeton, A., Conerly, T., Marcus, J., et al. (2024). Scaling monosemanticity: Extracting interpretable features from Claude 3 Sonnet. Transformer Circuits Thread.
31. Tunyasuvunakool, K., Adler, J., Wu, Z., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature, 596(7873), 590–596.

五、复现与方法论

1. Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. Nature, 483(7391), 531–533.
2. Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. Proceedings of the National Academy of Sciences, 113(34), E4933–E4934.
3. Klein, R. A., Vianello, M., Hasselman, F., et al. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. Advances in Methods and Practices in Psychological Science, 1(4), 443–490.
4. Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science, 349(6251), aac4716.
5. Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016a). Contextual sensitivity in scientific reproducibility. Proceedings of the National Academy of Sciences, 113(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113
6. Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016b). Reply to Inbar: Contextual sensitivity helps explain the reproducibility gap between social and cognitive psychology. Proceedings of the National Academy of Sciences, 113(34), E4935–E4936. https://doi.org/10.1073/pnas.1609700113

六、中国古代数学与天文文献

1. 利玛窦口译，徐光启笔受：《几何原本》（前六卷），万历三十五年（1607）刊。

2. 李冶：《测圆海镜》，1248 年成书。
3. 刘徽注，李淳风等注释：《九章算术》（约汉代成书；刘徽注成于魏景元四年，263 年；李淳风等注释成于唐初）。今据郭书春汇校：《汇校九章筭术》，沈阳：辽宁教育出版社，1990 年。
4. 秦九韶：《数书九章》，1247 年成书。
5. 宋濂等：《元史·历志》（载王恂、郭守敬等编《授时历》，1281 年起施行），北京：中华书局点校本，1976 年。
6. 魏徵等：《隋书·律历志》，唐显庆元年（656）成书；北京：中华书局点校本，1973 年。
7. 徐光启：《条议历法修正岁差疏》（崇祯二年，1629，内列“度数旁通十事”），载王重民辑校：《徐光启集》，北京：中华书局，1963 年。
8. 朱世杰：《四元玉鉴》，1303 年成书。
9. 祖冲之、祖暅：《缀术》（已佚；书名与“学官莫能究其深奥，是故废而不理”一语见《隋书·律历志》）。

七、其他

1. Åström, K. J., & Hägglund, T. (2006). Advanced PID Control. Research Triangle Park, NC: ISA.
2. Black, H. S. (1934). Stabilized feedback amplifiers. Bell System Technical Journal, 13(1), 1–18.
3. Chandonia, J.-M., Guan, L., Lin, S., Yu, C., Fox, N. K., & Brenner, S. E. (2022). SCOPe: Improvements to the structural classification of proteins – extended database to facilitate variant interpretation and machine learning. Nucleic Acids Research, 50(D1), D553–D559.
4. Chothia, C. (1992). One thousand families for the molecular biologist. Nature, 357(6379), 543–544.
5. Clough, R. W. (1960). The finite element method in plane stress analysis. In Proceedings of the 2nd ASCE Conference on Electronic Computation. Pittsburgh, PA. 〔页码未核〕
6. Desborough, L., & Miller, R. (2002). Increasing customer value of industrial control performance monitoring—Honeywell’s experience. In J. B. Rawlings, B. A. Ogunnaike, & J. W. Eaton (Eds.), Chemical Process Control VI (AIChE Symposium Series, 98(326), pp. 169–189).
7. Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. American Economic Review, 101(6), 2616–2652.
8. Goodhart, C. A. E. (1975). Problems of monetary management: The U.K. experience. In Papers in Monetary Economics (Vol. 1). Sydney: Reserve Bank of Australia.

9. Gulati, M., Levy, P. D., Mukherjee, D., et al. (2021). 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain. Circulation, 144(22), e368–e454.
10. Jakobsen, B., & Rosendahl, F. (1994). The Sleipner platform accident. Structural Engineering International, 4(3), 190–193. https://doi.org/10.2749/101686694780601971
11. Johnson, J. A., Caudle, K. E., Gong, L., et al. (2017). Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. Clinical Pharmacology & Therapeutics, 102(3), 397–404.
12. Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1), 35–45.
13. Levinthal, C. (1969). How to fold graciously. In P. Debrunner, J. C. M. Tsibris, & E. Münck (Eds.), Mössbauer Spectroscopy in Biological Systems: Proceedings of a Meeting Held at Allerton House, Monticello, Illinois (pp. 22–24). University of Illinois Press.
14. MacKenzie, D., & Millo, Y. (2003). Constructing a market, performing theory: The historical sociology of a financial derivatives exchange. American Journal of Sociology, 109(1), 107–145.
15. Maxwell, J. C. (1868). On governors. Proceedings of the Royal Society of London, 16, 270–283.
16. Minorsky, N. (1922). Directional stability of automatically steered bodies. Journal of the American Society for Naval Engineers, 34(2), 280–309.
17. Podlubny, I. (1999). Fractional-order systems and $PI^{\lambda D}\mu$-controllers. IEEE Transactions on Automatic Control, 44(1), 208–214.
18. U.S. Food and Drug Administration. (2014). De Novo classification: HeartFlow FFRCT, coronary vascular physiologic simulation software (DEN130045; granted November 26, 2014).
19. U.S. General Accounting Office. (1992). Patriot missile defense: Software problem led to system failure at Dhahran, Saudi Arabia (GAO/IMTEC-92-26).
20. 王德生：《科学与微积分：差异、整体与文明的数理内核》，SIO 思想体系文丛，SDE Universes 网站（德麦国际），网络版。〔本书底本；发布年份未详〕
21. 王德生：《SDE 思想体系入门》，德麦国际专著第 395 号，德麦国际出版社，2026 年，ISBN 979-8-90690-869-8。
22. Wiener, N. (1948). Cybernetics: Or Control and Communication in the Animal and the Machine. Paris: Hermann; Cambridge, MA: Technology Press; New York: Wiley.

附录

附录G　推翻条件登记表

本表把各章、各编写下的失败条件汇总在一处。每一条的读法相同：若前半句所说的情形被确认，后半句所说的主张就要收回、收窄或降格。第35章已按五类主张讨论过这些条件的性质，这里只做登记，便于读者逐条核对。

一、全书主命题

编号	若出现……	则……	见
G1	科学的失败并不集中在初始条件与情境依赖上	收回“积分常数 C 须从现实外借”	导论、第12章
G2	控制统计功效、发表偏倚与子领域之后，情境依赖对复现失败没有预测力	收回“失败集中于 C”；该检验上 C 外借律没有增量	第12、35章
G3	情境依赖只能事后认定，不能在盲评中事先预测	“C 外借律”降为启发性说法	第12章
G4	存在不借现场信息、跨情境长期稳定的落地技术或 AI 应用	收窄“技术即再微分”“再微分不可替代”	第14、27、33、35章
G5	应用中的偏差只引起再校准，从不改写模型	“再微分即回写”降级	第14章
G6	把 S、D、E 词汇删去后，各章结论都不改变	承认 SDE 解读的增量有限，降为“整理语言”	第12、15、21、35章
G7	有现代预测型科学的核心不依赖连续数学	“微积分是科学的心脏”限于动力学科学	第35章、结语
G55	设计良好的大规模研究表明，同一效应在系统性改变情境（不只是换国家、换线上线下）之后依然稳定	承认在该领域被微分切掉的纠缠不是主要的失败来源；“失败集中于 C”在该领域收回	第12章

编号	若出现……	则……	见
G56	情境依赖极低的学科（如经典力学的教学实验）也大规模复现失败	失败的主因不在 C；复现危机不再作为“C 外借律”的检验把手	第12章

二、第一编　根基

编号	若出现……	则……	见
G8	换一种测量安排，研究“对象”从不跟着变	发生学在这一处没有增量	第1章
G9	发生学的追问与普通科学史叙述得出相同结论	承认该处增量有限	第1章
G10	某个科学概念的定义从未随测量与校准改动	“缠绕在先”在该概念上不成立	第1、2章
G11	某领域定律的推导不用任何“可忽略”假设或极限	收回“数学化是一次过滤”	第3章
G12	理论改写从不由系统残差触发	“反常是回流通道”收窄为“回流通道之一”	第3章
G13	不经形式化编码就能判定“可计算”	收回“可计算性只对数学态有定义”	第4、30章
G14	信息九宫比“隐性/显性知识”二分多不出可检验的区分	承认九宫只是在整理语言	第4章
G15	被诊断为“停在逻辑态”的手艺知识形式化后没有可辨认的损失	“截取即丢弃”收窄	第4章
G16	“封顶”被用来整体判定一个文明	这种用法应当删除	第4章

三、第二编　历史

编号	若出现……	则……	见
G17	某古代文明已有通用的“变化率与累积量互逆”的演算	收回“缺的是这条通道”	第5章

编号	若出现……	则……	见
G18	古代先声已用于随时间变化的量，或已有通用算法	“停在静态、逐题、单向”收窄	第6章
G19	史料表明莱布尼茨依赖牛顿的材料推出微积分	改写“独立发明”	第7章
G20	点号与 d 号对后续发展速度没有可辨认的影响	“符号即语言”降为修辞	第7章
G21	控制制度、经济、地理变量后，数学传统不再解释剩余差异	李约瑟之问的数理候选降为中间环节	第8章
G22	微积分的严格化主要由教学建制推动、与反常无关	第9章的 123 循环解读降为叙述方式	第9章
G53	十七世纪动力学科学的关键进展在微积分形成之前已由等价的几何方法完成，后来的推进也不依赖微积分的符号演算	把微积分视为动力学科学“发动机”的判断收窄为“重要工具之一”	第5、7、9章
G54	“再微分”类比不能比纯粹的史学叙述多给出可检验的区分（例如预测哪些地方会先出现基础危机）	该类比从论证中撤下，只作修辞保留	第9章

四、第三编　机理

编号	若出现……	则……	见
G23	截取最狠（最依赖“其他条件不变”）的学科，复现率并不更低	“截取即丢弃”收窄	第10章
G24	边界与内部分布的区分不能带来新的预测	第11章的“边界决定内部”读法降为修辞	第11章
G25	应用失败与起点、边界数据的质量无关	收窄“位置须外借”的经验含义	第13章
G26	“差异已稳定”与“连续、可积”系统性地不一致	二者的对应只保留为比喻	第13章

编号	若出现……	则……	见
G27	加入 D、E 后给不出三项可检验的区分（反常以相对现行模型的残差出现；仪器偏向产生与其设计理论相容的数据；起手路径与成果类型相关）	三方程在本书的用法只算整理语言	第15章
G28	个案写不出可观察的回写痕迹，或 D、E 的指认可以事后随意调换	123 原理在本书的用法退为叙事框架	第16章

五、第四编　扩张

编号	若出现……	则……	见
G29	大量场模型的应用不需要现场的边界或初值数据	“边界是借来的 C”收窄	第17章
G30	某类重要数值计算的可靠性与初值、边界、网格的选取无关	“落点借 C”收窄	第18章
G31	布朗类模型在样本外检验中不逊于非独立、非正态模型	收回“随机微积分只收编规矩形状的不确定性”	第19章
G32	分数阶模型在换一种加载历史后不胜过整数阶模型	分数阶降为拟合工具	第20章
G33	——	伞场微积分只保留限定条件下的局部逼近命题；“传统数学是伞模型极限”等说法已由作者撤回	第20章
G34	实时间路径积分获得一般的严格构造	第21章相应例证改写为历史陈述	第21章

六、第五编　文明

编号	若出现……	则……	见
G35	预报水平的提升主要来自改基本方程，而非初值获取	“初值敏感就是 C 的放大”收窄	第22章

编号	若出现……	则……	见
G36	某模型在不增加现场初值信息的情况下，系统性地突破李雅普诺夫预报期限	收回第22章的核心论断	第22章
G37	统计显示工程事故主要由基本力学方程失效造成	收回“失败集中在 C 的外借”	第23章
G38	足够多的系统断开反馈后仍长期保持闭环表现，且不靠更准的模型或初值	“反馈就是持续借 C”收窄	第24章
G39	能从权重里读出简洁、可独立检验的方程，并重现模型的主要能力（通过减法检验）	“不成文的积分”降为阶段性描述或收回	第25、26章
G40	AI 研发药物在二期以后的成功率稳定、显著地更高	“替不了落点”收窄	第27章
G41	长期追踪发现使用者的现场判断没有退化	撤回对“负向回写”的担忧	第27章

七、第六编　评估与批判

编号	若出现……	则……	见
G42	能系统证明：被数学化筛掉的残差里，不含任何与人的问题相关的结构化差异	第一刀失效	第30、33章
G43	谱隙不可判定、波动方程不可计算解等结果被证明不依赖无穷极限、非光滑解等理想化	第二刀中这一论据作废	第31、33章
G44	数学的几次大扩充主要由数学内部驱动，而非被物理反常逼出	第二刀的史学部分收窄	第31章
G45	AI 成为主要的数学化执行者后，残差多样性与新问题类型不降反升；或 AI 能独立立出问题类型并被现场接受	第三刀收回	第32、33章

编号	若出现……	则……	见
G46	哈萨比斯的原话核定后与本书所批的命题不符	判词随原文修改；依据转述的论点作废	第28、33章
G47	对哈萨比斯的批评过不了第29章的“三道门槛”	该批评不算反驳	第29章

八、第七编　边界

编号	若出现……	则……	见
G48	三态“往返”引不出与单线上升不同的预期	降为描述用语	第34章
G49	三化的“主要侧重”在案例中与非侧重情形一样常见	撤回侧重对应	第34章
G50	计算机辅助证明表明数学在身份层也须碰现实	“数学是理念界的场态”放宽	第34章
G51	先讲意义的微积分课，在事先定好的时间窗内不见效	撤回结语中的教学建议	第35章、结语
G52	以上失败条件在事后被逐条解释掉	第35章本身失效	第35章

附录H　SIO → SDE 术语对照表

本书的底本《科学与微积分：差异、整体与文明的数理内核》写于 SIO 阶段，用的是“SIO 本体论”“ΔSIO”“创生律、成就律、演化律”一类说法。本书正文已统一改用现行 SDE 口径（以第395卷《SDE思想体系入门》为准）。本附录供读者对照底本时查用。

使用前先记住三条：

一、**SIO 的 S 与 SDE 的 S 不是同一个字**。SIO 的 S 是主体（Subject），I 是互动，O 是客体；SDE 的 S 是显露（Show），不是结构（Structure），D 是差异序列，E 是特征纠缠。两组三项不能按字母逐项对齐。

二、**对照不等于对应**。下表“现行说法”一栏，是本书在相应语境中改用的说法，不表示旧词与新词意思完全相同；有些旧说已被收回，只能作为“旧说”提及。

三、凡标“主要侧重”的地方，都只表示某一项与另一项在多数情形下关系较近，**严格对应未经证明**，不能据此推出任何必然结论。

一、本体与对象

底本说法	现行说法	说明
SIO 本体论；存在 = SIO 整体	发生范式；“在 E 中，经 D，成 S”	现行口径中，SIO 降为 S 维度里“稳定成形”的一种显露，不再作本源整体。“在 E 中，经 D，成 S”是关系速写，不是操作顺序，也不是已完成的普遍定律。

底本说法	现行说法	说明
SIO（主体、互动、客体三大意识）	SIO 作为三大意识发生的旧框架，现行口径中降为 S 维度里“稳定成形”的一种显露（同上行）	“SIO 属 S 的特例”是第395卷已撤回的旧说，不再使用；保留 SIO 一词时须说明理论阶段；三号位“先后”主要指讲解次序与结构依赖，不是个体发展时间表。
ΔSIO（差异单元）	差异序列 D；特征	底本把差异当作最小单元；现行口径中 D 是差异的推进、承接、修正与收敛所形成的路径，不是“单元”，也不是“步骤”或“变化”。在本书的微积分语境里，被微分截取的“增量”对应的是现实界中被截出的那一点差异。
特征纠缠（序言“土壤”）	特征纠缠 E；特征纠缠聚合体	基本保留并加细：纠缠是“缠绕在先，端点在后”；物体是特征纠缠聚合体被稳定显露的表征；名是指针。E 是参与发生的土壤，不是背景或环境。
SIO 的内场／外场（附录B）	信息九宫中的“数学态”（场态）	底本说数学研究 SIO 的场；现行口径：每一界（现实、理念、自我）都有信息，各有符号态、逻辑态、数学态；数学态指连接、次序与共同约束，是扩展用语，不等于形式数学。
概念的 SIO 化（附录A）	概念的发生学追问；认知形成与形式身份分层	保留“追问概念从哪里来”的方向；但须区分认知形成、符号约定、公理设定、形式证明四层，讲清发生不等于证成。

二、三原理与意义三律

底本说法	现行说法	说明
SIO 三原理（创生律、演化律、成就律）	不再作主术语；动力问题交给 123 原理，意义问题交给意义三律	底本把三原理称为存在的本体根基。现行体系中的“三原理”指 123 原理这一个循环，书中没有另立三条原理。
创生律（差异不断生成）	无直接替代；相关内容分入 D 的推进与特征律	底本称“微分是创生律的显影”。本书改说：微分是在现实界的一次截取，截取即丢弃。
成就律（差异趋向稳定、形成整体）	部分对应特征律（差异稳定出同一性）；“形成整体”在本书中由积分与 S 的显露承担	底本称“积分是成就律的数学化”，本书不再把这当结论。
演化律（模态张力推动形式升级）	123 原理：D—E 矛盾累积 → 逼出 S 改变 → S 的改变回写 D、E	底本称微积分的扩张是演化律的显影。本书在第16章以 123 原理改写，并强调回写须被观察、不能被宣布，也有负向回写。底本“螺旋上升”的说法不取：重组需要条件，不是必然上升。
意义三律（序言、第十二章：创生律、演化律、成就律）	意义三律：特征律、自由律、幸福律	名同而实异。底本有时把“三原理”也叫意义三律；现行意义三律是特征律（发生目标）、自由律（发生路径）、幸福律（发生动力）。另有 D1 意义目标“创造、自由、幸福”，与三律层次不同。
特征律	特征律	保留。要点：同一不等于完全相同，差异之间有可承接的连接即保持同一；稳定也可能错误。

底本说法	现行说法	说明
三化受意义三律约束（第十三章）	三化与意义三律只有“主要侧重”	本书第34章：标准化主要侧重特征律，简化主要侧重自由律，优化主要侧重幸福律；均非严格对应，且附有撤回条件。幸福律“张力—转换—释放”本身是机制候选。

三、方法与诊断

底本说法	现行说法	说明
科学知识循环：差异—特征—微分—积分—模型—异常—再微分	微分—积分—再微分之圈；反常处用 123 原理	本书主命题。注意：“微分、积分、再微分”三段不是 123 原理，不能把任何三步流程都叫 123 原理。
微分是“差异的忠实抄写”	微分是现实界中的截取，代价是切掉周围纠缠	忠实只在被截取的那一点上成立。
积分是“整体的发明创造”	积分是理念界中的拼合，结果为 $f+C$	保留“发明性”，补上积分常数 C 的外借：C 须在落点上从现实借回。
模态（底本泛称）	E2 信息三模态：符号（粒子态）、逻辑（波态）、数学（场态）	三态不构成单线等级，往返而非阶梯；“有且只有三种”是便于分析的约定，穷尽性待论证。
混沌—自组织—涌现（复杂性机制）	秩序、混沌、介生三态（分析语言）	三态不是物理量，不是好坏排名，没有可测的最佳中点；介生的要害是“承接”。不得把方程的自身项、耦合项、随机项分别对号为秩序、介生、混沌。
中国“缺乏对差异的数学表征”	中国算学走算法—代数一路；数理解释是李约瑟之问的候选之一	见第8章。须同时交代循环风险与史实风险；“符号态封顶”禁止用来整体评判一个文明。

底本说法	现行说法	说明
三号位显影（O→I→S）	解释号位；“先后”主要指讲解次序与结构依赖	不作个体发展时间表，不据此裁决笛卡尔；发育次序是有待检验的经验命题。
27 宫（C⊗M⊗V）	三轴组合图	在未定义向量空间与线性运算前，不是已建立的张量积模型；27 不证明分类完备。
微积分是“科学的心脏”	在动力学科学中，微积分是“微分—积分—再微分”这套看世界方式最集中的数学器官	不作“没有微积分就没有现代科学”的全称判断。

四、怎样使用这张表

读底本时遇到旧词，先查本表，再看正文相应章节。引用底本原句时，请写明“底本原作……”，不要把旧词悄悄换成新词，也不要把新词的限定条件悄悄去掉。表中凡是标作“不再作结论”“不取”的说法，只能作为思想史材料提及；凡是标作“主要侧重”的对应，只能用来提问，不能用来推论。

最后，这张表本身也会过时。现行口径若再修订，本表应随之更新，并保留修订记录：一个体系的术语怎样一步步改过来，本身就是它是否经得起检验的一份证据。

附录I　引文与史实核验表

本表登记全书引文与史实的核验情况，分三部分：哈萨比斯本人言论的出处，正文所据的数学与实证研究，以及成稿后逐条核对过的其他史实与引文。核对的原则是：能找到原始文本的，以原始文本为准；只见于转述的，正文注明“据转述”；查不到可靠出处的细节，改为限定说法或删去。

一、哈萨比斯言论出处

编号	出处	内容要点	核验状态
I1	诺贝尔化学奖演讲，2024 年 12 月 8 日，斯德哥尔摩（诺贝尔基金会网站公开的演讲幻灯片）	猜想原句：自然界中能被生成或找到的任何模式，都能被经典学习算法高效发现和建模（第 23 页）；选题三条件：巨大组合搜索空间、清晰的目标函数、大量数据或精确高效的模拟器（第 7 页）	已对照幻灯片原文核对
I2	莱克斯·弗里德曼播客第 475 期，2025 年 7 月	复述上述猜想；自然经进化与物理过程筛选而有结构；分解大数类问题无可学结构，也许需要量子计算机；认同主持人“可学习的自然系统”“P 与 NP 是物理问题”的说法；“信息比物质和能量更基本”；以 1900 年的知识发明相对论作为检验	已对照播客网站所载文字稿（人工整理）核对
I3	《代达罗斯》第 155 卷第 1–2 期（2026），第 34–44 页，与詹姆斯·马尼卡对谈，DOI 10.1162/DAED.a.971	自然中多数事物有稳定结构、可被经典系统逆向工程；“提出正确的问题是科学中最难的部分”；以 1910 年的知识推出广义相对论作为检验	已对照全文核对

编号	出处	内容要点	核验状态
I4	Google DeepMind 播客“The future of intelligence”，汉娜·弗莱主持，2025 年 12 月 16 日	大意为：迄今宇宙中没人发现不可计算的东西（节目 44:44 起“Turing machines”一段）	节目信息已核；无公开文字稿，原话据汤姆·瓦尔恰克 2026 年 1 月 9 日的文章转述
I5	CBS《60 分钟》，斯科特·佩利采访，2025 年 4 月	“也许有一天我们能在 AI 帮助下治愈所有疾病”；“终结疾病”是主持人的追问，他答“我想那是够得着的，也许在接下来十年左右，我看不出为什么不能”	已对照节目文字稿核对
I6	塞巴斯蒂安·马拉比《无限机器》（The Infinity Machine，2026）	把哈萨比斯的猜想与彭罗斯的量子意识论对照	据约翰·巴林顿的书评转述，未核传记原文
I7	流行转述“没有 AI 不能解决的问题”	——	不是哈萨比斯原话，是传播中的放大；第28章已将二者分开

二、数学与实证结果

编号	文献	正文用途	须注意的条件
I8	T. S. Cubitt, D. Pérez-García, M. M. Wolf, “Undecidability of the spectral gap”, Nature 528 (2015): 207–211	第31章：量子多体系统有无谱隙不可判定	结论建立在热力学极限（无穷大晶格）上；有限晶格的谱隙原则上可算
I9	J. Bausch, T. S. Cubitt, A. Lucia, D. Pérez-García, “Undecidability of the spectral gap in one dimension”, Physical Review X 10 (2020): 031038	第31章：结果推广到一维	仍依赖无穷极限

编号	文献	正文用途	须注意的条件
I10	M. B. Pour-El, I. Richards, “The wave equation with computable initial data such that its unique solution is not computable”, Advances in Mathematics 39 (1981): 215–239；另见二人专著 Computability in Analysis and Physics（1989）	第31章：可计算初值、唯一解不可计算	依赖所选范数与解的光滑程度；在能量范数下解保持可计算
I11	C. Moore, “Unpredictability and undecidability in dynamical systems”, Physical Review Letters 64 (1990): 2354–2357	第31章：粒子运动的长期行为可编码停机问题	依赖无限精度的初始条件
I12	R. Cardona, E. Miranda, D. Peralta-Salas, F. Presas, “Constructing Turing complete Euler flows in dimension 3”, PNAS 118 (2021): e2026818118	第31章：理想流体中可嵌入通用图灵机	针对特定几何上的稳态欧拉流
I13	Open Science Collaboration, “Estimating the reproducibility of psychological science”, Science 349 (2015): aac4716	第12、35章：复现危机的基本数据	复现率随判据不同而不同
I14	R. A. Klein et al., “Many Labs 2”, Advances in Methods and Practices in Psychological Science 1 (2018): 443–490	第12章：效应在多样本间的异质性	异质性总体偏小，正文不夸大情境依赖
I15	J. J. Van Bavel et al., “Contextual sensitivity in scientific reproducibility”, PNAS 113 (2016): 6454–6459；Y. Inbar 同年在 PNAS 的反驳性评论	第12、35章：情境敏感度与复现成功负相关（约 −0.23）	控制子学科后关联基本消失（Inbar），争论未定；正文并列两说

三、其他史实与引文的核对结果

编号	位置	条目	处理
I16	导论	哈萨比斯论量子系统	改写：改据 Lex Fridman #475 访谈原意（分解大数类问题无可学结构，“也许需要量子计算机”）
I17	导论	哈萨比斯猜想原句	核实：诺贝尔演讲幻灯片第 23 页；访谈复述见 Lex Fridman #475
I18	第3章	胡塞尔论伽利略“发现的天才，同时也是遮蔽的天才”	限定：《欧洲科学的危机与超越论的现象学》第 9 节，德文 entdeckender und verdeckender Genius；未注页码
I19	第5章	巴比伦天文日志起止年代	核实：现存可定年者约公元前 652 年至公元前 61 年
I20	第5章	第谷观测精度	核实：最好接近一角分，一般估计一到两角分
I21	第6章	刘徽注中圆周率一段的归属	核实并点名双方：有学者归祖冲之，三上义夫、李约瑟等归刘徽
I22	第6章	祖冲之圆周率领先年数	改正：“近千年”改为“九百多年”，十五世纪初先后被摩陀婆、阿尔·卡西（1424）超过
I23	第6章	摩陀婆生卒年	核实：约 1340—约 1425，均为推定
I24	第6章	《缀术》“学官莫能究其深奥”	出处为《隋书·律历志》
I25	第8章	琼斯《欧洲奇迹》归类	核实并限定：常被归入“国家体系竞争”一路，但该书兼论环境、经济与地缘
I26	第8章	席文论十七世纪中国天文学革命	改正：Sivin, Chinese Science 5 (1982): 45–66；文中点名王锡阐，删去梅文鼎；“没有带来相应社会后果”
I27	第8章	徐光启“度数旁通十事”	改正补全：崇祯二年（1629）修历奏疏，十项按通行概括列出

编号	位置	条目	处理
I28	第11章	磁单极子、阿哈罗诺夫—玻姆效应	改正措辞：AB 效应为非单连通区域上“闭而不恰当”的矢势；狄拉克单极子无全局光滑矢势；由上同调记录
I29	第12章	Begley & Ellis 2012	核实补细节：Nature 483:531–533；安进重复 53 项确认 6 项（约 11%）
I30	第12章	天气可预报上限	核实：Zhang 等 2019, J. Atmos. Sci.，中纬度内在可预报上限约两周
I31	第14章	药物基因组学例子	核实补全：华法林与 CYP2C9/VKORC1，CPIC 2017 指南（Johnson 等, Clin Pharmacol Ther 102:397–404）；临床结局试验结论不一
I32	第14章	道路“诱发需求”	核实：Duranton & Turner 2011, AER，“道路拥堵基本定律”；保留争议限定
I33	第16章	普朗克视量子化为计算手段	核实补出处：普朗克 1931 年 10 月 7 日致伍德信，“一次绝望之举”
I34	第16章	拉卡托斯“进步的/退化的研究纲领”	核实：通行中译；出处 Lakatos 1970《证伪与科学研究纲领方法论》
I35	第17章	个体化血流动力学建模	据实改写：冠脉 FFR-CT，HeartFlow 2014 年获 FDA De Novo 准许；2021 年 AHA/ACC 胸痛指南列为可选评估手段
I36	第20章	洛必达—莱布尼茨通信	核实：回信通常系于 1695 年 9 月 30 日，取大意
I37	第23章	纳维的荣军院吊桥	核实补细节：1826 年锚碇开裂、输水管破裂，1827 年撤销拆除；主因各家侧重不一
I38	第23章	工程事故归因统计	明确写为“本书没有做这项统计，留待检验”
I39	第24章	PID 占比	核实：Desborough & Miller 2002 调查，97% 的基础调节回路用 PID

编号	位置	条目	处理
I40	第26章	可解释性研究	补全：Li 等 2023（ICLR）；Karvonen 2024；Anthropic 2023、2024（Towards / Scaling Monosemanticity）；Olah 等 2020（Zoom In）；Anthropic 2025（电路追踪）
I41	第27章	初值仍来自同化系统	补进展并限定：Aardvark Weather，Allen 等 2025, Nature 641:1172–1179；仍为研究系统，业务预报仍靠传统同化
I42	第27章	对自动化实验室“新材料”的复核	核实补全：Cheetham & Seshadri, Chemistry of Materials 36(8), 2024；Leeman, Palgrave, Schoop 等, PRX Energy 3 (2024): 011002
I43	第28章	DeepMind 播客“没人发现不可计算的东西”	节目信息已核（Google DeepMind, 2025-12-16，Hannah Fry 主持，44:44 起“Turing machines”一段）；无公开文字稿，原话仍“据转述”（Walczak 2026-01-09）
I44	第28章	马拉比《无限机器》	改为限定说法：据 John Barrington 的 Substack 书评转述，未核传记原文
I45	第28章	CBS《60 分钟》2025 年 4 月	核实并改正归属：“终结疾病”出自主持人佩利追问，他答“within reach, maybe within the next decade or so. I don’t see why not.”
I46	第28章	诺贝尔演讲引文	已对照演讲幻灯片（nobelprize.org）核对；“信息比物质和能量更基本”一句改注出自 Lex Fridman #475
I47	第29章	蛋白质折叠类型数	改正：“一两千种”改为“一千余种”；Chothia 1992, Nature 357:543–544；SCOPe 2.08 收 1257 种折叠
I48	第29章	AlphaTensor、FunSearch	核实补全：Fawzi 等 2022, Nature 610:47–53（GF(2) 上 4×4 矩阵 47 次乘法）；Romera-Paredes 等 2024, Nature 625:468–475（8 维帽集 512）

编号	位置	条目	处理
I49	第30章	AlphaFold 对无序区的提示	据实改写：Tunyasuvunakool 等 2021, Nature 596:590；EMBL-EBI AlphaFold 教程（pLDDT<50）
I50	第32章	选题三条件	改为准确引述：诺贝尔演讲幻灯片第 7 页（巨大组合搜索空间、清晰目标函数、大量数据或精确高效的模拟器）
I51	第34章	焦糖化温度	改为限定说法：约 160—170°C（各来源不一）
I52	第35章	Van Bavel 等 2016	核实写入：相关约 −0.23；并列 Inbar 2016 的反驳（控制子学科后关联消失）与原作者回应
I53	第3、16、32章	水星近日点进动与海王星的发现	核实：勒维耶 1859 年算得每世纪约 38 角秒的多余进动，纽康后修正为约 43 角秒；海王星 1846 年由伽勒依勒维耶的预言观测到
I54	第7章	莱布尼茨的记号与刊物	核实：∫ 号见 1675 年 10 月 29 日手稿，d 号见同年 11 月 11 日手稿（据 Cajori《数学记号史》）；Acta Eruditorum 译《博学者学报》，Commercium Epistolicum（1712）译《通信集》
I55	第27章	AI 研发药物的临床成功率	核实：Jayatunga 等 2024, Drug Discovery Today：一期成功率约八到九成，二期约四成，与行业历史水平相当
I56	第31章	Weihrauch 与 Zhong 论波动方程的可计算性	补登：K. Weihrauch & N. Zhong，“Is wave propagation computable or can wave computers beat the Turing machine?”，Proceedings of the London Mathematical Society 85 (2002)：在恰当的函数空间与范数下解算子可计算，与 I10 的结论互为条件

编号	位置	条目	处理
I57	第31章	陶哲轩的平均化纳维—斯托克斯方程	补登：T. Tao，“Finite time blowup for an averaged three-dimensional Navier-Stokes equation”，Journal of the American Mathematical Society 29 (2016): 601–674；构造针对平均化方程，不是原方程
I58	第21章	卡茨 1949、卡梅伦 1960	核实：M. Kac，“On distributions of certain Wiener functionals”，Trans. AMS 65 (1949)（费曼—卡茨公式）；R. H. Cameron, J. Math. Phys. 39 (1960): 126–140：复扩散常数的维纳测度只有限可加，实时间路径积分没有通常意义的可数可加测度
I59	第21章	电子反常磁矩的测量精度	改正措辞：Fan、Gabrielse 等 2023 年测得 g/2 = 1.00115965218059(13)，相对精度约 0.13 ppt，与四圈 QED 计算一致
I60	第22章	数值天气预报的“静悄悄的革命”	核实：P. Bauer, A. Thorpe, G. Brunet, Nature 525 (2015): 47–55；预报技巧约每十年提高一天
I61	第22章	洛伦兹吸引子的存在性	核实：W. Tucker，“The Lorenz attractor exists”，C. R. Acad. Sci. Paris 328 (1999): 1197–1202，计算机辅助证明
I62	第19章	“模型塑造市场”	补出处：D. MacKenzie & Y. Millo，“Constructing a Market, Performing Theory”，American Journal of Sociology 109(1) (2003): 107–145
I63	第18章	“爱国者”导弹拦截失败	核实：GAO/IMTEC-92-26（1992）；1991 年 2 月 25 日宰赫兰，系统连续运行约 100 小时，24 位定点寄存器累积时间误差约 0.34 秒，28 名美军士兵死亡
I64	第18章	斯莱普内尔 A 号平台沉没	核实：1991 年 8 月 23 日；有限元分析低估剪应力约 47%，加之关键部位钢筋锚固不足；损失约 7 亿美元（SINTEF 调查；Jakobsen & Rosendahl, Structural Engineering International 4(3) (1994): 190–193）

编号	位置	条目	处理
I65	第22章	集合预报投入业务	核实：1992 年 12 月，欧洲中期天气预报中心与美国国家气象中心先后投入业务
I66	第27章	“梯子上的下一步”“虚拟细胞”	核实：《代达罗斯》对谈原文 “AlphaFold 3 is our next step on the ladder”“A virtual cell would be an incredible boon…”
I67	第27章	“世界上最好的天气预报系统”	核实：莱克斯·弗里德曼第475期文字稿（00:50:22）“we’ve created the best weather prediction systems in the world”（原文为复数）；“虚拟细胞”另见 00:42:11
I68	第28、29章	第37手	核实：诺贝尔演讲幻灯片第 3 页（“Game 2 - Move 37”“novel creative strategies”）；莱克斯·弗里德曼第475期 00:32:36
I69	第28章	“终极工具”	核实：《代达罗斯》对谈 “I’ve always seen AI as potentially the ultimate tool to advance the frontier of knowledge”；弗里德曼第475期 00:48:03 亦有此说
I70	第29章	列文索尔悖论	核实：诺贝尔演讲幻灯片第 10 页（“Levinthal’s Paradox: estimated ~10^300 possible conformations”）
I71	第28章	《代达罗斯》卷期	核实：Dædalus, Winter/Spring 2026, 155(1–2): 34–44, DOI 10.1162/DAED.a.971
I72	第28章	对谈题目与访谈文字稿	核实：对谈题为 “AI as the Ultimate Tool for Science: A Conversation with Demis Hassabis”；弗里德曼第475期 2025 年 7 月 23 日发布，播客网站所载文字稿注明为人工整理

成稿后共核对上表 57 条；正文中已不留尚未核实的标记。

后记

这本书的底本，是王德生的《科学与微积分：差异、整体与文明的数理内核》。底本约五万字，写于作者使用 SIO 框架的阶段。全书的主干、问题意识和许多关键的例子，都来自那本小书：微分忠实抄写差异，积分大胆发明整体，基本定理把二者焊成科学的循环；为微积分正名的那份心事，也是底本作者自己的。书中以“我”出现的个人经历，都属于他。

从五万字到二十万字，是这样做成的。

先定口径和骨架。动笔之前，先写下一份全书共用的写作纲要：书的主命题，三段式的圈，七编三十五章的目录，现行 SDE 口径的要点，必须收窄或不得再用的说法，每一章的写法与格式。底本哪一章对应扩写版哪一章，也在纲要里一一对应好。这一步决定了这本书不是五万字的注释，也不是另起炉灶的新书，而是在底本主干上的一次扩写与提升。

再分章起草。扩写由一个 AI 写作组完成，在同一份纲要下分章写成初稿，各写若干章。分头写的好处是快，代价是容易各说各话：同一个例子在两章里各讲一遍，同一个词有两种写法，前一章的结尾和后一章的开头接不上。所以每一位起草者交稿时，都要报告本章可能与哪些章重复、本章的主张怎样会错。

然后统稿。统稿做的是让三十几章读起来像一本书：核对每一处“见第N章”是否对得上，统一“理念物理界”“积分常数　C”“再微分”“不成文的积分”这些关键词的写法，把被两章以上展开的长段压缩为一处，修好章与章之间的衔接，并按纲要把越出口径的说法改回来。各章交来的失败条件，汇总进了附录G。

接着核验。全书的引文与史实，做过逐条核对：能找到原始文本的，以原始文本为准；只见于转述的，正文注明“据转述”；核对中发现记错的年份、出处、数字，已在正文里改正或改为限定的说法。哈萨比斯的几处言论，出处等级

不同，有的见于演讲幻灯片和访谈文字稿，有的只见于他人转述，本书在第28章分开写明，在附录I逐条登记。附录I还列出了正文所据的数学与实证文献，以及每一项结果须注意的条件。

最后是三审：一审读内容，查论证是否前后一贯、是否越出口径、对他人的转述是否公道；二审读文字，查错字、病句与标点体例，并补齐每章的“这一章的问题”“带走的话”“小账”等体例要件；复审把全书分成前后两半交叉核查，执行统稿的各项决定；终审通读全稿，定稿付印。

交代这些，不是为了说这本书做得有多细，而是为了让读者知道它可能在哪里出错。分章起草、统一纲要，可能让某些章节读起来比底本更整齐，也更平；核验只能核到能找到的材料为止，附录I里写着“据转述”“未核原文”的条目，就是还没有走到底的地方；三审能挑出的，主要是前后矛盾和越界的说法，挑不出一个整体判断本身的错。

所以，读者若发现本书有错，请不必客气。引文和史实的错，可以对照附录I去查，告诉我们哪一条记错了，依据是什么。论证的错，可以按附录G去检验：找到对应的那一条，看它所说“若出现……”的情形是否已经出现；出现了，本书就应当按“则……”一栏收回、收窄或降格相应的主张。附录G之外的错法，也一定还有。第35章说过，保修单之外，总有师傅没想到的坏法；写下这张单子，是为了让下一个发现坏法的人，知道该在哪一条上打叉。

一本讲积分常数要从现实去借的书，自己的那个常数，也只能从读者那里借。