物理学思想简史



王德生 著

出版信息

物理学思想简史

从属性纠缠到物理世界的创造

著者　王德生

出版　德麦国际出版社

Demai International Press

德麦国际专著第 376 号

2026 年 10 月 第一版

ISBN 979-8-90690-956-5

定价 US$20

© 2026 王德生。版权所有。

本书未经许可，不得以任何形式复制或传播。学术引用及法律允许的合理使用除外。

在线阅读、学习与问对：

https://sdeuniverses.com/books/m/376/

目录

导论 从一个苹果开始：物理学为何是一部创造史 6

第一编 物如何发生：SDE 与属性纠缠的本体论

第一章 物不是现成的：一块石头如何成为一个“物” 20

第二章 物体、物质、物理：三个层级如何同生 25

第三章 属性没有独自的主人：三属性纠缠 30

第四章 SDE 三方程与六路径、三原理：研究物理创造的共同语法 35

第五章 如何证明“新物已经发生”：创造判据与反例 41

第二编 从自然秩序到可测的运动：前近代物理的多条源流

第六章 五大文明如何生成天、地、声音与重量的秩序 46

第七章 亚里士多德：性质、位置与自然运动 52

第八章 原子与几何：德谟克利特、欧几里得、阿基米德 58

第九章 光与声的实验传统：伊本·海赛姆及多文明实践 65

第十章 哥白尼、开普勒与伽利略：从天空秩序到运动序列 72

第三编 微积分的创造：差异如何成为可计算的物理

第十一章 没有微积分之前：运动为什么难以连续地书写 78

第十二章 牛顿与莱布尼茨：两条路径创造微积分 85

第十三章 微分：局部差异如何凝成新的物理规定性 93

第十四章 积分：怎样从局部规则创造整体运动与守恒 101

第十五章 从常微分到偏微分：数学如何创造波、场与连续体 110

第四编 牛顿世界：新力学物体与经典法则的创造

第十六章 牛顿重造了怎样的物：质量、惯性、运动量 120

第十七章 万有引力：天空与地上为什么能成为同一物理 127

第十八章 测量与单位：经典物理如何创造公共“物理” 135

第十九章 拉格朗日、哈密顿：旧力学为什么能再生新形式 144

第二十章 马赫与经典物理的界限：惯性真是物体自带的吗 152

第五编 热、时间与概率：统计物理的新物质规定

第二十一章 蒸汽机与卡诺：热为什么成为一门新物理 160

第二十二章 能量与熵：守恒之外为何仍有方向 166

第二十三章 玻尔兹曼：微观运动怎样生成宏观物理 173

第二十四章 吉布斯与系综：概率是否只是无知 180

第二十五章 相变与临界：新物态为什么从整体中出现 187

第六编 场之物：法拉第、麦克斯韦与电磁世界

第二十六章 电与磁的相遇：实验怎样把分离的世界联结起来 194

第二十七章 法拉第的铁环：短暂变化怎样成为新的物理对象 201

第二十八章 麦克斯韦：位移电流怎样使场成为共同演化的系统 209

第二十九章 赫兹：怎样使看不见的电磁波取得实验身份 217

第三十章 从作用到场：能量与动量怎样取得局域的历史 225

第七编 时空之物：洛伦兹、爱因斯坦、闵可夫斯基与诺特

第三十一章 迈克耳孙、洛伦兹、庞加莱：旧变换遇到新困难 233

第三十二章 爱因斯坦 1905：同时性使“长度”成为另一种物 240

第三十三章 闵可夫斯基：四维时空怎样重组物理对象 247

第三十四章 广义相对论：时空几何与物质怎样互为条件 254

第三十五章 诺特定理：为什么变换与守恒可以共同规定物理 262

第八编 从辐射到原子：量子物理的早期创造

第三十六章 黑体与普朗克：能量怎样成为可以计数的份额 269

第三十七章 爱因斯坦光量子：光怎样在一次交换中取得身份 276

第三十八章 汤姆孙与卢瑟福：原子怎样从不可分之物成为可追问的结构 283

第三十九章 玻尔原子：稳定状态与谱线怎样共同规定一种物 290

第四十章 德布罗意与康普顿：波和粒子的边界怎样被重新组织 298

第九编 量子态、测量与属性纠缠的最艰难检验

第四十一章 海森堡矩阵力学：经典轨道之外还有什么物 305
第四十二章 薛定谔波动力学：波函数是怎样的“物” 313
第四十三章 玻恩与不确定性：概率、分布与可观测量 322
第四十四章 玻尔、爱因斯坦与量子测量：装置参与了什么 331
第四十五章 海森堡与薛定谔：两套方法如何会合成新的物理 341

第十编 粒子、对称性与量子场：新物身份的持续生成

第四十六章 狄拉克与反物质：方程怎样改变粒子的身份 350
第四十七章 费曼、施温格与朝永：相互作用怎样成为可计算的物理 357
第四十八章 杨振宁与米尔斯：局部比较怎样成为相互作用 365
第四十九章 标准模型：粒子怎样成为场的激发 374
第五十章 物理常数与尺度：什么在变化中保持，什么需要重新规定 383

第十一编 复杂物质与宇宙：整体何以产生新的物理

第五十一章 固体与材料：晶体、能带和“这一类”的新身份 391
第五十二章 超导与超流：整体性究竟创造了什么新物态 397
第五十三章 混沌、湍流与临界：不收敛一定是新本体吗 404
第五十四章 宇宙学：整体宇宙怎样成为可研究的对象 411
第五十五章 黑洞、引力波、暗物质与暗能量：未完成之物 418

第十二编 新物理世界的下一轮创造：科学典范与 SDE 研究纲领

第五十六章 视觉、听觉、触觉的 SDE 怎样在现代仪器里聚合 426
第五十七章 微积分之后：非局域、随机、离散与计算物理 432
第五十八章 科学典范怎样碰撞：库恩、拉卡托斯与 SDE 对话 438
第五十九章 十大物理学家的十种创造：同一张比较图 445
第六十章 第十一位是谁：AI、科研共同体与未发生的物理 453
结语　物理世界为什么仍将不断发生 459

导论 从一个苹果开始：物理学为何是一部创造史

桌上的苹果，书中的问题

请在想象中把一个苹果放在桌上。它不是名人传记中的道具，也不是在某个传奇时刻击中谁的头顶；这里只安排一个普通的、可以由读者继续改变的场景。你可以拿起它，闻到气味，把它放上天平，也可以让它从不高的地方落下，再用摄像记录运动。每一种做法似乎都面对同一个苹果，却未必在回答同一个问题。

如果问题是能否食用，成熟程度和保存状态很重要；如果问题是怎样下落，质量分布、形状、空气作用与释放条件便进入讨论；如果问题是材料内部怎样回应光或电磁探测，日常可见的整个苹果又可能不再是最合适的研究单位。科学并不只是把一个对象看得越来越仔细，它还会改变什么被计为对象，哪些差异应被保留，怎样的回答才算有内容。

把苹果放上天平，并不是创造宇宙中原来没有的全部物质。可是在这一安排里，苹果取得了可与标准比较的质量规定，仪器、操作及校准承担了不可省略的角色。让苹果下落，也不是发明下降本身；但把持续变化组织为位置与时间的对应，就使速度、加速度、误差和预测成为能够共同追问的问题。日常物在研究中获得了新的物理身份。

本书从这个差别开始。它不要求读者怀疑苹果是否存在，而邀请读者暂缓一句过于轻便的话：“它本来就是那样，科学只把它说出来。”有些情况下，这句话足以支持日常行动；有些情况下，它却提前拿走了最重要的问题：所说的“那样”究竟由什么规定，哪些条件使它成立，改变条件以后，它还是不是同一种属性？

创造一词为什么必须冒险使用

把物理学写成创造史，是一个需要承担后果的选择。发现史通常强调自然先在的约束如何被逐渐认识，创造史则进一步追问研究对象与公共判断方式怎样建立。两种说法未必在每一处相互排斥；但若创造只意味着科学家很有想象力，它对思想史就没有增加多少内容。

本书所说的创造，涉及可研究的规定性发生变化：怎样识别一个物，什么状态属于它，哪些操作能够改变或保持它，如何比较不同装置中的结果，以及哪些关系可以被称为规律。一个新符号可能参与这种变化，一项新实验也可能参与；两者都不能仅凭新颖就自动被宣布为新世界。

危险在于把创造说得太强。若说没有理论之前就没有任何相应过程，便混淆历史中的认识形成与自然过程的持存；若说任何命名都产生一种物理存在，又会取消物理对判断的抵抗。另一方面，把创造说得太弱，只承认文字或心理表象改变，也无法充分说明装置、材料、操作与共同体的实际重组。

我们选择在这两种缩减之间工作。创造不是任意制造事实，发生也不是否定持存。它要求展示一种规定怎样从可以忽略的背景变成必须说明的组成，又怎样在反复操作中形成可稳定使用的对象。这个立场应当接受竞争解释，而不能把“创造”预先写进所有结论。

因而，本书标题中的创造不是颁给每位物理学家的统一奖章，而是一项研究任务。有的章节主要展示计算形式的创造，有的展示实验对象的创造，有的涉及存在规定的深刻重建。读者不必因为同意其中一种，就同意所有其他层次；每个层次都应有自己的证据。

让“物理世界”具有可以讨论的范围

一个世界不必等于全部宇宙的一张总照片。对于研究者，它也可以指一组彼此接续的对象、可执行操作、允许的变化和共同检验关系。例如，一套运动模型把位置、速度、受力与边界联系起来，使某些问题能够形成，某些回答能够相互比较。这种组织有范围，却不是私人的幻觉。

本书使用“物理世界”时，关注的正是这种具有公共效力的组织。它可以接触真实材料，产生可重复后果，并容许别人通过不同路径发现原判断的错误。一个物理世界越成熟，越容易把自己的形成条件封装起来，使研究者能够直接使用结果。思想史的工作之一，就是适时打开这种封装。

打开并不意味着每次计算都要从文明起源开始。工程师使用可靠公式时，可以暂时不追问它的全部历史；物理学家也可以把经过检验的量度当作现成条件。问题只在于，当公式失效或对象改变时，原先被省略的条件是否还能被找回。不能找回的便利，可能逐渐变成难以察觉的障碍。

因此，新的物理世界并不总以取代旧世界为标志。有时它把旧规律放进更明确的适用范围，有时建立新旧表达的转换，有时必须舍弃某些旧实体解释。包含、近似、并行与替代都可能出现。全书不会把其中一种预先规定为所有历史的终点。

苹果仍可被拿来食用，也可被当作下降的物体。日常身份没有因为力学出现便全部消失。我们要解释的不是为什么后来人终于只剩一种正确看法，而是不同规定如何共存、冲突、转换，并在某些条件下形成更强的公共组织。

四个词，四种需要区分的工作

“物”“物体”“物质”“物理”在日常语言中常常彼此借用。本书不会假定这些词具有永不改变的词典边界，而将它们安排为四个提问位置。物使我们追问一种稳定聚合如何成立；物体使我们追问个体怎样被持续指认；物质使我们追问类型为何支持替换；物理使我们追问不同人怎样把结果交给同一套检验关系。

这些提问不是逐层叠起来的楼梯。决定把苹果的哪一部分作为样本，已经可能受到材料分类和测量要求影响；改变制备方式，可能使原来被算作同类的样本分开；新的比较标准又可能重新规定什么算一个完整实验对象。层次之间互相进入，不能先把每一层永久做完，再进入下一层。

读者将在第一编看到这些问题的具体展开。导论只提供一种阅读姿势：每遇到一个新名词，先问它承担哪一种工作。电子不是因为名字像一个小东西，就必须具有苹果式边界；场不是因为难以抓握，就只是一种修辞；概率分布不是因为写在纸上，就自动与实验样本同一。对象的资格须由其特定成立方式说明。

这也使思想史不再只是概念目录。我们需要追踪名称背后的辨认、制备、推导与公共接续。一个名称可以长期保持，内部规定却已改变；一个新名称也可能只把旧关系重新包装。用词的连续与物理结构的连续，必须分别检查。

读者若在某一章暂时不接受本书的本体论判断，仍可利用这四种提问检查材料。它们首先要求把研究对象说准确，而不是要求每个人在开始阅读之前已经接受同一哲学结论。

普通经验为什么既是起点又不够用

普通经验给物理思想提供了大量入口：推拉、冷热、声音、明暗、轻重、远近。没有这些差异，科学语言很难与人的行动相接。但普通经验并没有替所有物理对象规定最终形式。以熟悉的固体想象光，以熟悉的流体想象电，能够启发问题，也可能把不适合的限制带进模型。

一个好的类比让读者知道相同在哪里，也知道应当在哪一处停止。把传播放大成水面波纹，可以帮助理解局部变化与传播的区别；它不能单独证明各种传播具有同一种介

质。把量子态比作一张尚未翻开的卡片，可以说明某种知识限制，却不能未经论证便把全部量子现象变成预先决定的隐藏答案。

本书的日常场景因此都带有边界。石头、铁棒、苹果、绳索不是所有自然存在的缩小版，它们只是使一个问题获得可操控形状的入口。进入以后，数学和实验可能迫使我们放弃场景中的某些直觉。读者不应因为一个比喻很好懂，就把它当作证明。

同样，感觉不能被当作三只彼此独立的收件箱。视觉、听觉、触觉各自都有辨认、改变与稳定的组织，又能通过行动互相校正。后文将讨论这种组织怎样参与物的成位。但这里应先保留一条限制：没有实际实施的感觉操作，不能被写成已经获得的证据；想象可以安排检验，不能冒充检验完成。

数学进入之前，变量已经承担了什么

一本面向普通读者的物理思想史，应当帮助读者理解公式为什么出现，而不只是展示它很重要。写下一个字母并不困难，困难在于说明它代表什么，怎样获得数值，在何种变化下仍保持意义。一个变量成为可计算对象之前，常常已经凝结了测量、分类与理想化。

回到苹果。若记下随时间变化的位置，必须先决定追踪中心、表面一点还是图像边缘。三种记录可以都合法，却不能在同一条数据链中未经转换便轮流使用。若苹果转动或变形，差别尤其明显。方程不能自动修复记录对象的不一致，精确求解也可能只是精确地计算了一个混合问题。

再看一个最小的数值例子。假设某个教学模型给出的路程为 s=二 t 平方，距离以米计，时间以秒计。两秒时路程为八米，三秒时为十八米，因此这一秒区间内平均速度为每秒十米。模型在两秒处的瞬时速度则为每秒八米。两个结果不同，没有矛盾，因为它们回答不同问题。

把时间单位改为毫秒时，公式中的系数必须相应变化，不能保留原来的数字二却继续把同一个字母当成同一个单位。把平均速度当成瞬时速度，也不是某种深刻的新物理，只是概念与运算未对齐。全书保留这样的简单算例，是为了让宏大解释始终受到小而具体的检查。

这个模型是人为设定，不是我们在此测得的苹果运动。数学内部的正确与经验上的成立应当分别判断。读者可以完全同意上述计算，同时继续要求证据说明真实装置是否满足模型前提。这样的分层不是降低数学地位，而是使数学的力量使用在准确的位置上。

微积分为何贯穿全书

微积分在本书中不仅是一章技术成就。它使局部变化与整体过程之间建立系统联系，也让初值、边界与连续性成为物理问题不可省略的组成。它提供了比“物体从这里到了那里”更细的追问：怎样经过中间状态，局部变化如何接续，什么条件使整体运动得以确定。

但本书不会把微积分写成凭空产生一切物理的总钥匙。一个微分方程可以有多个解，需要初始或边界条件；一个数值方法可以逼近解，也可能因离散化而失稳；一个连续模型在断裂或突变处可能需要换用新的数学组织。每种能力都带着适用条件，条件并不是可以在故事中删掉的附注。

《科学与微积分》与《科学是什么？》构成本书研究局部差异、整体重建及再检验关系的思想背景。[导 1] 本书将相应提问落实到具体模型，不把早期术语与现行 SDE 记号机械互换。读者无需先通读这些底本，便可以顺着每个例子的变量、运算与条件理解论证。

数学也会反过来改变物理对象。如果一套方程显示原先当作独立的量必须满足共同约束，研究者便需要重新审视自由度；如果不同表示能够严格互相转换，某些看似根本的分歧可能只是表示差别；如果现有状态不足以确定后续变化，则可能需要增加有操作意义的变量。

因此，第三编的阅读重点不是记住谁先写出哪一种符号，而是观察新的数学组织怎样使问题成为问题，又怎样把尚未完成的物理规定暴露出来。能够计算的地方应当算清，不能由计算决定的地方则应留给经验与概念辨析，二者不能互相代办。

史料、推导与解释，不能共用一个证明

物理思想史至少面对三种不同的真实性要求。历史要求我们说明人物在何时、何处提出什么，证据来自原稿、出版文本、器物还是后来回忆。数学要求推导在给定前提下成立。哲学解释则要求说明一个变化为什么涉及对象构成，而非只涉及描述便利。三者相互联系，却不共享同一张合格证。

某位作者确实写过一句话，不证明那句话物理上正确；一个公式今天可以严密推导，不证明历史人物当时使用同一推导；一个实验结果符合现代模型，也不自动证明本书的发生解释是唯一合理说明。我们会尽量把这三种句子分开：史料支持什么，本书怎样重建，重建还可以受到什么质疑。

原典也不是不会出错的圣物。它们可能包含错误观测、时代性猜想、修辞安排与未完成论证。阅读原典的价值，不是把每一句都奉为真理，而是避免用后来教科书的整齐形式替代当时真正存在的问题。看见前人如何困难地工作，比替他们补写完美预言更有历史意义。

器物材料又有另一种边界。一件馆藏能够说明某种制作与使用的可能，却未必代表整个文明或整个年代；一项设计建议不能自动当成普遍实施的证据；一幅后世绘画也不能无条件充当事件现场。多文明比较需要这种材料纪律，才不会在扩大视野时扩大虚构。

本书中的思想实验与数字示例均为解释服务。它们可以显示逻辑关系、揭示反例或帮助复算，却不会伪装成未曾实施的实验。一个例子设计得再合理，也需要等待真实材料决定是否适合所讨论的世界。

多文明比较不应寻找单一冠军

天文、历法、测量、声音与重量的实践，曾在不同社会条件下获得各自组织。把这些工作放在一起，可以显示问题入口的多样性：有的紧贴农业与历法，有的接近建筑、贸易或乐器，有的进入论辩与几何证明。但这些差异不能被直接排成文明优劣的名单。

如果以现代大学物理系的学科边界筛选古代材料，许多重要实践会被排除；如果把一切精巧工艺都追认为完整现代科学，又会抹掉不同知识形式的实际区别。本书希望保留它们的任务、证据与解释限度，再讨论哪些关系能够被后来的研究接续。

相似成果也不自动证明传播。两个地区都处理影长，可能因为都面对太阳照射与测量任务；要建立直接继承，还需要文本、路线、翻译或器物链的证据。反过来，缺乏完整传播证据也不意味着它们彼此绝对隔绝。我们应当把已知联系、合理可能与尚未证明分开书写。

比较的目的不是把所有传统改造成同一理论的预演。伊本·海赛姆不需要先拥有 SDE 术语，才值得进入本书；古代原子论也不需要被宣布提前发现量子粒子，才具有思想力量。历史人物的创造性存在于他们能够提出、改变并交付的问题中，而不在于我们能为他们贴上多少现代标签。

在这样的写法中，多文明不是附在主线前面的礼貌名单，而是使主线接受挑战的材料。若不同实践显示物理对象可以通过不相同的操作组织成立，本书就必须说明这种差异，而不能为了结构整齐把它们改写成同一条必然道路。

人物很重要，人物也不是整个解释

科学家的名字有助于记忆，也能把抽象问题放回具体生命。一个人的训练、判断、坚持与改变，有时确实决定研究能否越过某个困难。将这些作用抹去，会使思想史失去行动者；把一切归给天才，又会使行动发生的条件消失。

一项能够公开检验的成果通常需要文本、工艺、测量传统、数学资源、合作与反对意见。个人并不是站在这些条件之外，偶然使用它们；他或她的能力也在这些关系中形成、加强或受阻。讨论这种构成，不是否定个人创造，而是使创造不再只能由神秘天赋说明。

本书因此不以传奇瞬间替代持续工作。一个醒目的故事可以成为阅读入口，却不能承担全部因果解释。即便某段回忆确有史料根据，它也可能只记录研究者后来怎样理解自己的起点，未必完整保留当时的过程。我们宁愿让证据较少的场景短一点，也不替历史补上戏剧所需要的细节。

人物叙事还要容纳失败。某些后来被放弃的设想曾经合理地组织工作，某些没有立即获胜的方法为后继研究提供了资源。失败的价值需要具体说明，不能一律浪漫化；但只记录最终正确答案，也无法解释为什么人们曾投入不同方向。

最后，谁被记住本身也是历史问题。可留下文字与署名的人，不一定是唯一承担技术工作的群体。我们不可能在一本简史中恢复全部匿名劳动，但可以避免把复杂实践压缩成一个名字的独占，并在有材料处让装置制造、观测与校准工作进入叙述。

一个对象可以有多种有效世界

从经典力学进入热学，再进入相对论和量子研究，读者容易期待一条逐级淘汰的队列：后一个理论更深，所以前一个应当消失。实际研究却常保留不同尺度上的有效组织。保留并不说明它们永远同等正确，而说明任务、范围与可接受误差不同。

描述苹果的整体下降，不一定需要逐个追踪内部微观状态；研究内部热过程，又不能只依靠质心位置。简化是一项有根据的工作：它说明哪些差异在当前范围内不影响所承诺的结果，或者影响小于明确尺度。无根据的省略会误导，有根据的封装则使科学能够行动。

新理论有时改变旧参数的意义，而不是立即改变所有日常数字。同一符号在不同体系中所承担的角色可能发生变化，因此判断继承不能只看方程外形。反过来，方程外形差别很大，也可能经过适当变换给出相同的经验内容。我们将同时检查表示、对象与预测，不用其中一项代替全部。

这种层级意识尤其影响本体论。一个模式在某尺度上具有稳定身份，并不要求它与更小尺度的组成互相绝缘；它也不因依赖组成便失去研究资格。问题是该层次的变量能否承担独立的规律、控制和比较。全书关于热、统计、场与多体系统的讨论，会不断返回这一点。

物理世界的复数性因此不等于真理任意化。不同世界之间若有重叠任务，必须说明它们如何相容或为何冲突；若提出无法区分的竞争解释，也应明确现有证据不足以裁决。保留多种有效组织与放弃判断标准，是两种完全不同的态度。

怎样提出一个真正有力的对手

一本有主张的书最容易犯的错误，是为自己挑选最弱的对手。若把实在论写成科学家不需仪器也能直接看见宇宙本质，把工具主义写成公式只是一场文字游戏，把历史社会研究写成事实完全由权力决定，本书就能很容易获胜，但这种胜利没有多少价值。

强有力的实在论可以承认测量有历史、概念会修订、知识有误差，同时主张成功源于理论越来越接近独立世界的结构。它会追问：发生分析是否只是说明我们怎样获得知识，为什么足以改变关于存在的判断？这是本书必须认真回应的问题，不能用“认识也是发生”一句话结束。

强有力的工具解释则会承认新数学和新仪器具有巨大生产力，却怀疑是否需要额外的本体论语言。它可以指出，相同结果经等价模型获得，说明创造主要在方法，而不是在自然界。对此，本书需要显示对象身份、可实施操作或证据资格怎样发生了不能被纯换算消除的变化；若显示不出来，就应缩小主张。

还有一种社会历史解释，会强调组织、资源、教育和认可机制影响研究方向。SDE 不应把这些解释当作外围杂音，也不能把它们全部吞进一个无边的 E 就声称已完成说明。哪些制度关系经什么通道影响问题选择、装置能力或结果传播，需要实际材料。哲学上的整体性不能代替历史因果分析。

在每一个争论中，最有价值的不是宣布对方只看到一面，而是指出双方会对什么材料作出不同预期。若目前没有可区分预期，就保留争论层级；若一项竞争解释已经足够说明现象，就承认它的解释权。发生学必须通过这种约束取得自己的位置。

SDE 在这本书里承担什么

SDE 中的 S 指显露，D 指差异序列，E 指特征纠缠。它们不是三件先在零件，也不分别等同于主体、互动和客体。主体、互动、客体属于另一条分析轴；视觉、听觉、触

觉同样不能被分派成三个字母。第一编会作专门区分，导论只强调这套语言所要求的研究责任。

研究一个显露，需要说明它经什么差异而来，又有哪些特征关系使这些差异取得意义。研究一个操作序列，需要说明它为何这样切分、怎样被已有结果引导。研究一种纠缠，则需要指出具体承接，而不是列出一份包含所有可能因素的环境名单。

三方程 S=F(D,E)、D=G(S,E)、E=H(S,D)表达上述相互规定。这里的字母不是已经测得的常数，函数也不是可以不加定义便计算的自然律。要进入具体求解，仍须给出对象、变量、范围、参数与条件。一个输出可能由多种组合产生，三式本身不保证逆推唯一。

六路径使研究可以从不同入口展开；三原理则要求两维之间的具体张力推动第三维改变，并返回原问题检验。它们应当改变下一步做什么，而不只给已经发生的事情重新命名。一个方法若只能在事后说明任何结果，却无法帮助排除错误解释，就尚未显示足够区分力。

本书也会检验这套方法。如果在反复研究中，SDE 分类无法复核，不能帮助定位失配，也不能比更简单的说明增加任何可操作内容，那么其方法地位就应受到限制。允许理论被用于物理史，也必须允许物理史使理论变窄。

从苹果记录走一条完整路径

为了使方法不只停在定义，设想两次苹果下降记录出现系统差异。第一次图像显示路程近似按时间平方增长，第二次短时段偏离明显。我们尚不知道原因。可能是释放时带有初速度，可能是计时零点不同，也可能是图像追踪对象不一致。没有必要立即宣布新定律。

沿 S—E—D 进入路径，先把重复可见的差异 S 作为问题，再检查计时、取点与释放之间的关联 E，最后安排区分操作 D。可以固定相机与取点规则，更换释放方式；也可以保持释放装置，改用独立计时。每一次改变应尽量说明保留了什么，使新结果能够返回原差异。

若 S 与 D 之间的失配迫使我们发现原先忽略的初速度，就改变 E 中的起始状态规定，再检验；若 S 与既有 E 不合而不同计时方案能给出区别，就改变 D，再检验；若现有 D 与 E 不能支持把所有记录合称一种下降，就细分 S 的类别，再检验。三种推进都不是由词语保证成功，它们需要结果承担约束。

这里没有完成真实实验。我们只是让读者看到，一条路径的完成不等于依次念出三个字母，而是形成一个能够失败的比较回路。如果检查以后发现全部异常来自计时偏差，就应接受这一解释。发生学的价值可以体现在帮助排除伪异常，而不必每次都收获新的本体革命。

这个例子也指出研究路径与事件时序不同。我们今天先看到异常，再回溯昨天的释放条件，并不意味着异常在物理过程中先于释放发生。方法的入口可以改变，因果承接仍须由实际材料说明。整体性不能使时间方向失去意义。

证伪不是在末尾附一句谦辞

“本书可能有错”是一句容易写的话，真正困难的是说明哪一种材料会使哪一种主张退出。某个历史日期错误，只需修正日期；某个算例错误，需要重算并检查依赖它的论证；某个对象构成判断被竞争模型完整解释，则可能要求缩小创造主张。不同错误的后果不应混在一起。

本书不会要求一个反例自动推翻整套思想，也不会允许每个反例都被解释成无关细节。需要建立依赖关系：哪些局部结论支持哪些较强主张，局部失败以后，较强主张还剩多少根据。这样的记录比一句“整体仍然正确”更诚实，也更有助于继续工作。

关于历史方向的命题尤其需要谨慎。若声称所有重大变革都表现为旧典范被新典范包含，就应允许那些无法形成受控包含、却保持成功预测的新体系构成反对材料；若声称多方法碰撞必然优于单一方法，就应接受相同资源条件下相反结果的限制。历史不应被要求为了理论美观而表现一致。

第一编将提出创造检验，后面的章节则在具体材料中使用和修订它们。导论不把这些检验逐项预演一遍，因为判据只有面对不同案例才显示力量。读者可以把每章当作一次试用：它真正排除了什么，保留了什么，又把什么问题交给下一章？

一部思想史若允许这样阅读，就不需要把不同意见看作破坏整体。相反，读者提出的准确反例可能使它从一种过宽的哲学叙事，变成能够长期工作的问题组织。

全书可以怎样走

第一编从普通物与属性入手，建立研究单位和方法边界。读者将看到，物体身份、材料类型与公共规律并非三个互不相干的问题。第二编回到多文明实践与古代、近代早期自然研究，考察历法、几何、运动、光声实验怎样使新的问题取得可比较的形状。

随后，微积分与经典力学使局部变化、整体轨道、约束和参考结构进入中心。阅读这些章节时，可以持续问：方程接手了哪些已经形成的规定，又在哪些地方迫使研究者重建规定？不要只把每个公式看成需要记住的结论，而要看它怎样组织可行动的问题。

热与统计把视野从少数对象的轨迹推向大量差异中的稳定、转换与方向；场的研究改变作用如何在空间与时间中组织的方式；相对论让时空量度、变换与几何之间的关系重新承担证明责任。它们不是为同一座楼添几个房间，而会改变楼的支撑方式，同时保留部分原有工作。

量子研究进一步使状态、制备、测量、概率和对象身份不能轻易分开。后续关于复杂系统、信息与开放问题的讨论，则考察物理组织怎样跨尺度扩展，哪些成功可以迁移，哪些类比应当停止。最后的科学典范与 AI 科研章节，把提问返回正在形成的研究共同体，而不把历史停在一份已经完成的英雄名单。

这条路线有时间顺序，也有问题回返。书中某些关键词会重新出现，但再次出现时应承担新的材料，而不是重复定义。读者可以顺读，也可以从熟悉的领域进入，再返回第一编检查方法。选择不同入口并不影响共同要求：明确对象、检查计算、保存条件、比较解释。

给不同读者的几种使用方法

对普通读者，最重要的不是一次记住所有名词，而是学会辨认一句论断属于什么层级。它是在讲史实、给数学推导、描述操作，还是提出哲学解释？遇到难懂公式时，可以先读变量与条件，再看算例；若仍无法复算，至少追问这个公式排除了什么可能结果。

对已有物理训练的读者，本书邀请你检查那些最容易被熟练使用掩盖的步骤。变量何时取得身份，近似怎样选择，实验对象如何被封装，理论等价在哪个范围成立？你完全可以不同意本书的创造解释，但请把反对落实到具体条件与映射上，使争论不只停在立场名称。

对科学史和哲学读者，数学部分不是装饰。它经常帮助限制一个过强的论断，例如耦合并不必然导致不收敛，数学改写并不必然产生新经验，单次随机也不排除分布稳定。一个简短反例可能比许多页立场辩论更有效，前提是我们不把模型反例误当成关于所有实际系统的定论。

对教师与研究者，每章可以被改造成讨论任务：保留一个历史材料、一个可复算关系和一个竞争解释，让学习者尝试安排下一步区分。教学的目的不必是让学生复述 SDE，而可以是使他们提出更准确的问题。若另一套语言同样有效，应当允许比较。

对正在使用 AI 的读者，尤其需要区分表达流畅与证据可靠。模型能够帮助整理、计算和提出候选，也可能把不同层次拼接得过于顺畅。把输出交给原典、程序与独立检验，是保持创造能力的条件，而不是创造之后才补做的手续。

简史的省略应当公开

任何简史都必须选择。它不能同时完成每一种文明的完整档案研究、每一项理论的全部证明和每场争论的最终裁决。问题不在于有没有省略，而在于省略是否改变了主张的资格。本书会尽量保留关键条件，使简化不被误读为已经排除所有其他可能。

章节的篇幅也不直接代表历史重要性。有些问题需要较多数学展开，有些人物只能通过有限材料讨论；有些实践缺少容易保存和取得的文本，不能因此被判为没有思想。本书的选择由主轴决定，却不主张主轴就是物理史唯一合理的组织方式。

书目指向实际采用的材料范围，不把资料列表的长度当作证据厚度。读到原典某节，便只据此支持相应判断；读到机构摘要，便不冒称逐页核验全文。对于跨理论比较，我们区分已建立的定理、教学重述和仍需进一步研究的解释。研究的可信度往往取决于这样的边界，而不是措辞有多肯定。

未决问题不是空白的耻辱。它们是接续研究的位置。只要能够说明缺少哪种材料、什么比较可能推进、结果出来以后如何改变结论，未决就具有工作内容。相反，用一个宏大词语把空白填满，可能使后来的研究失去入口。

因此，这本书不以覆盖一切为自我证明。它希望建立一种能够被补充、纠正与重新组织的阅读方式，使简史的有限性成为继续研究的条件，而不是隐藏在完整叙事背后的缺口。

创造的公共性与自然的抵抗

创造并不只发生在提出想法时。一个想法需要被别人理解、转换、重做，有时还需要被别人反驳，才能取得公共生命。共同体不是在外部给成果盖章的观众，而是通过独立操作、错误发现与标准形成，参与规定什么可以稳定称为知识。

公共性也不要求所有人亲手完成所有实验。大型研究需要分工，信任因此不可避免。但可靠信任应当有可追溯的责任与检查通道。谁完成了哪项校准，何种数据支持哪

条推论，哪些结果经过独立复核，这些关系使信任不同于盲目服从。知识规模扩大以后，公开检验的形式会改变，检验责任不会消失。

自然的抵抗始终在场。样品会破裂，预测会失效，独立团队会得到不同结果。人们可以争论差异来自哪里，却不能仅凭愿望把它取消。正因为创造要面对这种抵抗，它才不是主观任意。一个新物理对象的成立，需要经受它所打开的操作，而不只是被一个共同体命名。

同时，抵抗也不直接说出自己的唯一解释。一项异常可以提示新机制，也可以来自测量失误。研究者必须建立区分条件，使回应获得明确方向。发现论强调回应不是人随意决定的，发生论追问回应怎样成为可辨认的证据；两者的争论应发生在这个具体连接处。

在最好的情形中，创造使更多人能够对同一关系提出不同问题，又能把差异交给共同检验。它不是消灭所有歧见，而是使歧见获得可继续推进的形式。这也是本书所理解的物理世界的公共力量。

从这个苹果走向整部物理思想史

现在再看桌上的苹果。它没有因为导论结束而变成另一种神秘物件。我们只是比开头多了几项能够追问的区别：日常持存与物理规定，变量与读数，运算与经验，继承与包含，个人判断与公共检验。每一项区别都可能在后文受到新的材料考验。

本书不会要求读者在第一章以前就相信，所有科学创造都只能由 SDE 解释。它提出的是一项较艰难的邀请：把通常当成现成答案的地方重新打开，跟随具体人物、装置、计算和争论，看看新的规定怎样获得稳定，又怎样失去或扩大自己的范围。

有时，结果会支持发生解释；有时，更简单的工具说明已经足够；有时，我们只能确认争论仍未结束。三种结果都应进入书中。若只有第一种被允许，所谓创造史就会变成理论的例证集，而不再是向历史学习的研究。

苹果可以暂时留在桌上。接下来出现的石头、天体、琴弦、热机、场、量子态和几何，将不断要求我们修订“一个物”最初看似容易的含义。读完以后，重要的未必是给世界增加了多少生僻名词，而是能够更准确地问：什么在这里成立，怎样成立，我们凭什么这样说，又准备让什么结果改变这个判断？

物理学的创造史，从这种愿意把判断交给发生过程的态度开始。它也应当把自己的最后一页，交给尚未发生的下一次研究。

参考与说明

[导 1] 王德生：《科学与微积分：差异、整体与文明的数理内核》，专题长文，https://sdeuniverses.com/column/science-and-calculus/；《科学是什么？》，德麦国际专著第 231 号，https://sdeuniverses.com/books/m/231/text/。本导论仅说明问题背景；具体微积分原典、定理与推导见第三编。

[导 2] 王德生：《属性纠缠论》，德麦国际专著第 396 号，https://sdeuniverses.com/books/m/396/；《科学发生的结构》，第 342 号，https://sdeuniverses.com/books/m/342/text/。定义与版本的详细讨论分见第一编及第五十八章。

[导 3] 本书第一至第五章。苹果及下降记录为本书自拟思想案例，不据此重建牛顿传记事件。

第一编 物如何发生：SDE 与属性纠缠的本体论

第一章 物不是现成的
一块石头如何成为一个“物”

1.1　拿起石头之前，发生了什么

请设想一个平常的场景。雨后的路边有一块石头，颜色深浅不一，一面圆钝，一面留下细小的缺口。你把它拿起来，指尖先碰到凉意，随后感到重量；用另一块石头轻轻敲击，短促的声音散开。你说：“就是这一块。”这句话说得太顺，顺得使它内部的问题几乎没有机会出现：颜色、重量、缺口、凉意与声音，凭什么能够归到“这一块”上？

本书并不打算把一个容易的动作故意说得神秘。恰恰相反，我们要研究的，是容易怎样成为容易。一个成熟的动作能够省略许多说明，正如熟练的阅读不再逐笔辨认汉字。但是，省略得熟练，不意味着被省略的发生从来没有参与。手可以不加思索地抓住石头，不能由此推出抓握、辨认与物体边界之间不存在需要说明的关系。

如果从“一个人观察一块已经完整成立的石头”开始，石头作为物的身份已经被放进了前提。随后无论加入多少感官、仪器、心理因素，问题仍然容易变成：这些因素怎样帮助或者妨碍认识它。本书选择另一个起点：什么样的显露、差异序列与特征纠缠，使“这一块”能够作为具有这些规定性的物而成立？这里的研究不是给现成的物再添加环境，而是展开物之规定性的发生。

《何谓物》以视觉、听觉、触觉 SDE 在持续运行中形成的特征纠缠聚合体讨论物；《属性纠缠论》进一步追踪承载者与属性如何共同获得身份。[1，第 15、24—25 章；6] 以下的石头场景，是本书为展开这一思路而设置的思想案例，不是某项感知实验的记录。

1.2　颜色不是重量的标签，声音也不是形状的翻译

把三个感觉通道列出来，还远没有说明纠缠。我们可以同时看见红灯、听见脚步、碰到桌面，却不会因此把三者说成一个物。共同发生在一个时刻，只说明时间上的相邻；要使它们形成“这一块石头”，还需要一种更有约束的承接。

你转动手腕，石头的可见轮廓改变，指腹接触的部位也改变。缺口转到另一侧，手指寻找它的位置随之调整。你抬起石头，原来被遮住的表面显出；放回地面，接触产生新的响声。这里不是一种感觉替另一种感觉提供装饰，而是一次操作同时进入几条可以相互校正的差异序列。看见的转动与触到的边缘若长期不相容，“这是同一个物”的判断就需要重新检查。

视觉、听觉、触觉在本书中不是三个被动接收箱。视觉 SDE 中已有显露、差异序列和特征纠缠；触觉 SDE、听觉 SDE 也如此。将视觉规定为 S、触觉规定为 D、听觉规定为 E，会把完整发生系统误拆成三个零件。将三个通道分别固定为客体、互动、主体，同样会混淆感觉系统与号位侧重。

一个视觉显露可以以轮廓为重，也可以以距离变化或整体分布为重；一次触觉发生既可能突出受力表面，也可能突出按压过程及自身行动的调整。不同通道各自分化，又在新的承接中聚合。所谓聚合不是消除它们的差别，而是使不同差别能够在一个整体中保持关联。没有差别，聚合只剩重复；没有关联，差别仍然散落。

这也意味着，并非每一次辨认都需要三种感觉同时充分到场。远处的山峰不能随手触摸，安静的石头未必正在发声。应当研究已有的跨通道承接怎样沉淀为后续行动的可能，以及哪些可能仍须检验；不能把未实施的触摸写成已经取得的证据。某一现场的通道缺席，与一种规定性完全不需要任何跨通道支持，是两件事。

1.3 “同一个”是变化中的成就

如果一块石头必须从所有角度都呈现完全相同的轮廓，才配叫同一个物，那么转一下手腕就会使它失去身份。日常辨认显然不采取这样的规则。我们允许轮廓改变，同时要求改变之间有某种可以接续的组织。于是，同一性不能简单等同于显露逐项不变。

这里有两种需要区分的稳定。第一种是重复时某项读数近似相同；第二种是读数随操作改变，却依照同一种关系改变。对于后一种稳定，变化不是同一性的敌人，而是检验同一性的材料。转动之后缺口怎样移动，遮挡解除后表面怎样接续，都使“同一个”获得比单张静止图像更丰富的根据。

进一步设想，桌面上的石头被敲掉一小角。它还是原来的石头吗？搬运时，我们可能继续把它算作原物；研究断裂时，脱落碎片与剩余部分则需要分别成位；研究质量变化时，又须说明计量边界是否包括碎片。这些提问不能靠一个未经说明的“当然还是”或“当然不是”同时解决。

这不是任意相对主义。不同任务不能随心改变实际断口、碎片位置和称量记录。任务改变的是需要成立的同一性范围，而发生材料限制这种范围能否成立。可以把它理解为本书的一个工作命题：物体身份既不是任意命名，也不是脱离一切变化而预存的单一答案，而是对特定差异能够持续承担说明责任的稳定组织。

新物理系统的创造，正可能发生在这种组织的重建处。新的运动描述、新的实验操作和新的数学关系，未必增加一块肉眼可见的石头，却可能使石头以此前未能成立的物理规定性成位。我们需要证明的正是这种重建，而不是只给“石头”换一个更响亮的名字。

1.4 看不见时，石头并不因此消失

“物的规定性在发生中成立”，很容易被偷换成“没有人看，物就不存在”。这并不是本书采用的底本立场。《属性纠缠论》明确区分发生层的持存与意识层的属性显影，并以未被触碰的石头说明：分子联系与受力响应可以在无人关注时持存，属性的互相需要不能被用来取消这种持存。[1，第 15 章及“互相需要”相关论述]

这个区分不能在开篇声明一次，然后在后文忘掉。倘若我们一面承认持存，一面又说法拉第之前绝无任何电磁过程、牛顿之前物体从来不会按某种稳定方式运动，便重新混淆了不同层级。关于过去发生的科学判断，可以由后来形成的物理系统回溯建立；这种回溯需要证据，并不要求把过去改写成等待某个个人意识才开始发生。

本书所谓创造新的物理世界，首先要求严格研究新的物理存在规定及其发生系统。它不只是改变一幅内心图像，因为实验安排、测量对象、可执行操作和公共验证也参与改变；它也不是凭空制造全部物质，因为新系统始终面对不能由意愿取消的回应与限制。两种缩减都应拒绝：把创造缩成脑内看法，或者把创造夸成思想能够随意生产宇宙。

因此，我们不以“到底先有物还是先有人”的简单时间竞赛替代本体论问题。一个已经成熟的发生系统可以在研究中把若干条件暂时当作既定，这种局部处理是必要的；关键在于不能把局部处理的便利，直接升级为所有存在都由相互无关的完整实体组成的证明。

1.5 镜中的石头：一次聚合失败的检验

再设想桌上有一面镜子。镜中石头具有清楚轮廓，也随着观看位置改变而呈现相应变化。仅凭“显得很像”，能否确认镜中另有一块与桌上石头具有同样规定性的物体？

伸手这一操作会带来区分。手接触镜面，而不是接触镜后图像所指的位置；敲击产生的是当前接触系统的响应，而不是一块额外石头的独立响应。这个例子不是为了宣布镜像“不存在”。镜像有自己的发生系统和可研究的光学规定性。它说明的是：一种显露真实成立，不等于它以另一种物体的构成方式成立。

《属性纠缠论》以遮挡、替换、扰动与质料检验展开相近问题，并区分聚合、粒子化、维持与解散。[1，第 24 章]本书在此推进的判断是：物的检验不能只问“有没有一个显露”，还要问“这个显露能够承接哪些操作”。图像与石头可以共享某些视觉差异，却不因此共享全部受力、位移和接触关系。

这是一项重要的反例训练。若任意可辨认的图像都被称为同一种物体，发生学就失去区分力；若凡不能直接抓住的都被排除，又会把光学图像等具有明确发生条件的存在一并抹掉。问题不是在“全是真的”与“全是假的”之间选边，而是为不同成立方式作出准确规定。

同样的研究态度将进入后面的场、量子态与粒子身份讨论。不能因为它们不同于日常石头就宣布它们虚无，也不能因为某个公式能够写下就直接授予它们与石头相同的存在方式。需要逐项展示：什么发生了，什么得到稳定，哪些操作能够接续，哪些越界解释应当停止。

1.6　从“这一块”到一条运动曲线

现在把石头的位置写成随时间变化的函数。在纸上，这似乎只是一句简洁的“令位置为 x(t)”。但这句话至少已经接纳了几项规定：我们追踪的是哪个物；位置相对于什么参照确定；一次记录与下一次记录如何属于同一运动；时间怎样比较；在哪个尺度上将断裂、转动或形变暂时排除。

函数并不负责从无到有地替我们完成所有规定。数学形式能够把已经建立的区分组织得更强，也能反过来暴露那些区分尚不充分。例如，所谓同一个位置若有时记录质心、有时记录前端，一条拟合得再光滑的曲线也可能掩盖测量对象的不一致。错误并不一定出在计算，可能出在成为变量之前的那一步。

还可以反过来看：一旦位置与转动得到更精细的规定，先前被当成同一种移动的过程可能被分开。石头的质心平移与表面某一点的移动不再等同。数学问题由此返回物体身份，使我们必须说清被追踪的究竟是哪一种变化。这里不是先把物永久规定完，再让数学进场；物的规定会在数学工作中继续发生。

这就是第一编通向第三编的第一个接口：微积分进入物理学时，接手的不是一堆尚无身份的裸数。它接手的是已经具有某种可比性、可接续性与操作含义的差异；同时，它又通过新的局部关系和整体约束，参与重建这些差异所规定的物。

石头于是成为一个合适的起点。它太普通，以至于我们容易忘记它的成立；它又足够具体，使发生学不能只停留在抽象词汇。让“这一块”重新变成一个问题，不是为了永远怀疑手中的石头，而是为了看见物理学创造史真正开始的地方：那些被我们熟练使用的物之规定，原来仍有可以展开、可以修订、可以重新创造的发生过程。

第二章 物体、物质、物理：三个层级如何同生

2.1 一根铁棒里的三个问题

设想一家小型工坊要用一根铁棒制作支架。工人首先需要找到它：不是旁边较短的那根，也不是已经弯曲的那根。随后要判断它属于哪类材料，适不适合承受计划中的作用。最后还要确定，设计所使用的尺寸、承载要求与检验办法，能否被另一位操作者或另一家工坊可靠地接续。

这三个问题彼此相连，却不相同。“是哪一根”涉及物体的个体成位；“是什么材料”涉及可重复辨认的类型；“按什么关系可以共同计算与检验”涉及物理规定的公共成立。如果把三个问题都缩成“这根棒本来有什么属性”，实际工作的差别就被一个过于宽大的句子遮住了。

《属性纠缠论》第 25 章以“这一个”“这一类”“谁来都一样”分别展开物体、物质与物理。[1]这是一种发生学解码，而不是把三个词在所有语境中的用法规定完毕。本书继承这一区分，并进一步研究：个体、类型与公共规律怎样在物理创造中互相规定，怎样因其中一层的变化而共同重组。

尤其不能把它们写成三道单向工序：先制造物体，再归类物质，最后发现规律。我们选择什么样的棒作为一个实验样本，往往已经受到材料分类和实验规则的影响；新规则形成以后，又可能要求重新切分样本，甚至重新界定材料类型。一个看似稳固的“这一个”，也有来自“这一类”与公共操作的支持。

2.2 成位：边界不是简单画一条线

工人用粉笔在棒上作出记号，并不能因此把记号两侧变成两个完全独立的力学物体。若尚未切断，两侧仍可能共同承受作用；若已经切断，先前整体的受力关系就需要重新说明。可见，命名边界、视觉边界与操作边界不必重合。

本书把“成位”理解为一个研究任务：在某种发生系统内，哪些差异被组织成可以指认、操控和承接的个体。它不是在世界中随意圈一块区域，也不意味着每个个体都必须具有刚硬的外壳。一个可研究的振动模式没有石头式边缘，却可以在相应的制备、激发与辨认中取得身份。不同成位方式应当分别研究，而不是让日常固体替所有物理存在规定资格。

即使研究铁棒，个体性也不只是由外观保证。两根形状相同的棒可以在记录中被混淆，运输中的标记可以脱落，切取试样后也可能失去与原材料批次的联系。因此，编

号、取样、操作顺序与记录连续性具有实际作用。但这些程序不能代替物理检验：编号相同不保证材料未被换掉，标签完好也不保证内部状态不变。

由此可以区分两种失败。第一种是物体发生了变化，记录仍把它当作原来那个；第二种是物体所需的同一性关系仍可保持，记录却把它错误地拆散。发生学既要追踪边界的形成，也要追踪边界何时失去作用。这样，“物体是发生的”才不只是赞美变化的口号，而具有可落实的判断。

2.3 物质：类型为何不能只由外观决定

工坊中的“铁”“钢”“合适的支架材料”，不是完全相同的分类。前两种名称指向材料规定，后一种还包含用途与性能条件。即使一批棒具有相近外形和颜色，能否在同一设计中互换，仍需要相应证据。这里不必先进入复杂材料学，便能看出类型成立的核心：它要求某些差异可以重复识别，并在所声明的条件下支持替换。

“这一类”因此不是把许多“这一个”堆在一起。归类需要规定哪些差异重要，哪些差异在当前问题中可以保留为个体变化。若研究对象是表面磨损，表面处理可能非常重要；若研究另一个经论证可以忽略表面差异的问题，同样差异就可能退出必要说明。类型并不因此变成随意命名，因为替换之后是否仍能完成承诺，会反过来限制分类。

我们可以提出一个具体的思想检验。假设两批材料被归入同一工程类型，理由是它们在既定装置中表现出相近的弹性响应。当试验温度或加载路径改变，两批材料出现系统性差别时，研究可以有几条路线：缩小原分类的适用范围，增加状态说明，或者建立更细的类型。哪条路线合适，要由用途、差异与证据共同决定。

这段分析是本书对底本“可复现的纠缠质地”的展开，不是宣称现实材料已经被某个具体实验判为同类或异类。它的要点在于：类型有形成史，也有可承担后果的运行范围。物质规定的创造，有时就表现为将过去混在一起的响应分开，或者将过去看作不同材料的显露纳入一个更强的共同结构。

因此，材料的发生史不能简化成“先有很多东西，后来人给它们分类”。新的制备、加工和检验可以改变材料本身的可发生方式；新的数学描述又能改变哪些特征被组织成一种可重复的质地。分类、制作与解释在这里相互进入，而不是站在三条永不相交的线上。

2.4 物理：“谁来都一样”不是人人写同一个数

底本用“谁来都一样”概括公共规定性，容易让读者误以为客观性要求任何人、任何仪器、任何参考下都写出完全相同的数字。第 379 卷已经明确指出，经典物理也允许不同坐标描述具有不同数值；重要的是差异能否经有约束的关系相互转换。[2，第 6—7 章]

例如，两个人采用不同长度单位，报告的数字可以不同，而所报告的长度并不因此矛盾。更复杂的情况不是单位变换，而是测量对象或操作规定不同。棒在某种状态下的端点距离，与包含夹持部分的装配长度，可能本来就不是同一个被测量。把它们强行要求为同数，得到的不是更严格的科学，而是对问题的混淆。

本书据此把公共性解释为一种有条件的可接续：更换承担操作的人，保留或者明确转换必要条件，新的结果仍能进入同一套检验关系。人可以互换，承担的操作责任不能凭空消失；装置可以改进，新装置与旧结果之间的联系也不能靠声称“更先进”来保证。

这个意义上的公共性不是削弱主体，而是重新组织主体差异。某些私人偏好可以被排除，因为没有进入所研究的作用通道；某些选择必须被保留，因为它们规定了测什么、怎样测、怎样解释。把二者统称“主观因素”一并驱逐，与把所有个人感受一并放进物理方程，都是过于粗糙的处理。

新物理世界之所以能够成为公共世界，正需要这种精细的分辨。它既不能由单个人的赞同决定，也不能靠宣告所有人都应当同意而成立。它要使他人的不同操作有机会发现错误，并让这些错误能够返回修改原来的规定。

2.5 从手感到刻度：四种操作怎样改变显露

《属性纠缠论》提出放大、过滤、固化与符号压缩四种操作。[1，第 25 章]它们为研究科学对象怎样获得可承接性提供了一条线索。本书不把四者规定为所有实验必须依次经过的固定阶段，而将其看作常会相互配合的构成工作。

放大，是使原本难以辨认的差别能够进入观察或操作范围。过滤，是使某些差异退出当前记录，让所研究的关系获得较清楚的显露。固化，是把可以重复实施的安排稳定下来。符号压缩，则把已经形成的关系转成便于计算、传递与重新操作的表达。每一种工作都可能增加能力，也可能留下新的限制。

在铁棒思想案例中，一个极小的位移可以经装置成为较大的指示变化；读数程序可能排除明显不属于目标的摆动；支撑与加载步骤可以被固定；最后，一组响应被写成表

格或函数。此时出现的数值并不是离开所有操作后剩下的神秘本质，而是受控关系中的显露。不过，这也不意味着数值只反映装置，与材料无关。材料若不按所预期的方式响应，整个安排就必须承担失败。

更重要的是，四种操作会互相限制。过度过滤可能删去真正重要的变化；过度放大也可能把装置噪声推到前台；固化了错误步骤会稳定地生产偏差；符号压缩若丢掉适用范围，公式传播越广，误用也可能越广。因此，一项成熟的物理规定应当保存重新展开压缩的路径，让后来的研究者知道简洁表达依赖什么。

我们可以把这种保存称为“可恢复的简洁”。它不同于把全部现场永远写进每一次计算，也不同于用一句“条件相同”掩盖所有差异。科学的力量不是拒绝省略，而是使省略具有可说明、可检查、可在必要时撤回的根据。

2.6 单位、数值与常量：三件不能混成一件的事

国际单位制把量的数值与单位区分开来，并通过规定基本常数的固定数值建立单位体系。[11]这个具体的计量学安排提醒我们：一个数字的精确性，必须先问它属于哪一层。单位定义中的精确数值、某项测量的估计结果、一个模型在特定条件下采用的参数，并不是同一种“常数”。

若约定用不同单位表示同一长度，数值会变，但不能据此断言长度发生了新的物理变化。若单位定义已经固定，具体实现仍可能具有测量不确定度，也不能因此说定义失败。若模型参数随状态变化，首先应当检查参数的意义和适用域，而不是立刻宣告所有物理常数都不成立。

这一区分尤其重要，因为研究对话曾把多个领域的参数漂移、量子不确定性与相对论读数差别统一放在“不收敛”标题下。[7，第 30—42 页]它提供了寻找共同构成问题的启发，但这些现象在数学上并不是同一个命题。本书保留其问题意识，不将尚未区分的“常数难题”直接写成通用定理。

对铁棒而言，弹性模型中的参数能否稳定，需要实验与模型的具体支持；对单位体系而言，采用哪种定义是公开的计量约定；对某个理论不变量而言，稳定可能表现为变换关系中的保持。三种问题可以互相进入，却不能互相替代。属性纠缠论若要增强研究，就必须帮助我们更准确地区分它们，而不是让“纠缠”成为免除区分的口令。

2.7 三层同生，怎样进入一个方程

回到工坊。一根棒被选为样本，其材料类型被规定，加载和记录形成公共程序。随后，位移、作用与时间进入一个数学模型。这个过程中，个体成位、材料归类和规律表达相互支持：样本边界不清，函数变量就可能混乱；类型未经检验，参数迁移就可能失效；公共程序不稳定，一组精密结果也可能无法被他人承接。

这里的公共检验也不能取代物理关系本身。一个可共同执行的程序，不仅规定大家怎样报告，还要允许材料对安排说不：装配不能完成，结构不能维持，响应不满足承诺，都是能够纠正规则的后果。公共性不是许多人共享一句话，而是不同承担者能够进入同一种受限制的发生。

从这些相互关系出发，本书提出一项连接命题：物理变量的建立，是物体、物质与物理三层共同取得有限稳定的结果；微积分的运用，则会进一步检验这种稳定是否足以支持局部变化与整体重建。一个变量不是随手写下的字母，一条方程也不是脱离发生条件的纯粹命令。

假如后来发现棒的响应不能只由当前位移说明，还需要加载历史，那么被修订的就不只是一项数值。研究可能需要改变样本制备、状态分类和函数形式。原先的“这种材料”与“这条规律”一起获得新的范围，甚至形成新的研究对象。这就是三层同生的实际内容：每一层都不是等其他层全部完成之后才来领取一个名字。

“物体、物质、物理”于是构成一种有方向的追问，而不是三个静止抽屉。我们从个体进入类型，从类型进入公共关系，也可以反过来从新关系进入新的个体划分。方向可以改变，整体却不能被随意拆散。正是在这些方向的交错中，物理学不只是记录已经完成的东西，而不断创造新的成立方式。

第三章 属性没有独自的主人：三属性纠缠

3.1 “石头硬”究竟说了什么

设想你要用一块石头压住纸张，又要用另一块材料制作不会划伤桌面的垫片。“硬”在两个任务中似乎具有不同地位：前一个任务强调稳定支撑，后一个任务却可能把过硬当作麻烦。这里首先变化的不是一句赞美或批评，而是哪些作用关系进入判断，以及哪些后果被要求保留。

这并不说明硬度可以随评价任意改变。人不能仅凭喜欢就使受力响应改变。它说明的是，研究某个属性时，需要辨认属性显露、操作条件与判断位置怎样共同成立。说“硬只是石头自己的”，容易把成立条件放到属性之外；说“硬只是人的感觉”，又会使材料回应失去约束力；说“硬就是二者之间的互动”，如果仍把两端规定为先在完整实体，也没有真正改变问题。

属性纠缠论的起点，是取消一个未经检验的独占：任何一端都不能先把属性据为己有，然后把另外两端贬为附加影响。“属性没有独自的主人”，不是说属性没有可辨认的承载者，而是说承载者不能脱离共同发生而自行担保全部规定性。[1，第 15 章及三属性纠缠相关章节]

本章将这一命题从口号推进为几项区分。只有这些区分得到落实，三属性才能进入科学研究，而不只是把一个句子改写成三个句子。

3.2 三号位不是三维，也不是三件物

SDE 中的 S、D、E 分别表示显露、差异序列与特征纠缠。客体、互动、主体则是号位侧重。二者不是同一组三分法。不能将 S 直接解释为主体，将 D 解释为互动，将 E 解释为客体；也不能把 SDE 三方程直接当成三个号位各自的物理方程。

《属性纠缠论》的组织式，把号位侧重、维度侧重与所在之界共同纳入属性发生。[1，第 15 章]为方便阅读，本书将其写成文字式：属性怎样发生，需要同时说明承载者从哪个号位侧重显影、哪一维发生受到突出，以及相关特征纠缠处于怎样的界别。这里的乘号或并列不是算术乘法，而是共同规定的要求。

在“石头硬”的案例中，客体属性并不等于一组未经过任何操作的分子数据。我们还需要说明这些数据怎样进入“硬”这一属性的辨认。互动属性也不只是按压持续几秒，而要说明按压、接触与响应之间的组织。主体属性则不宜缩成喜欢或不喜欢；选择哪一种响应作为问题、怎样判断达到要求，也属于需要研究的号位侧重。

三种侧重各自都具有完整 SDE。例如，对可重复压痕的客体侧重分析，也需要差异序列和特征纠缠；对按压过程的互动侧重分析，也有显露的身份及其稳定条件；对判断活动的主体侧重分析，同样不能缺少实际操作与可核查的结果。三号位并不是把一次发生切成互不相干的三份劳动。

理念、现实与自我又是另一条分析轴。一个数学定义、一次实际按压与操作者关于自己能力的判断，可以在同一任务中纠缠，却不能靠把其中一项命名为“主体”就全部解释。底本的三界划分要求我们追踪这些不同发生如何互相进入，而不是建立三座彼此封闭的仓库。

3.3 从共变到构成：四种相互关系

底本区分互相关联、互相需要、互为条件与互相影响。[1，第七编]四个词接近，却不应视为同一句话的重复。它们分别把研究推进到不同的说明责任。

互相关联首先要求发现可以比较的变化。假设改变一种加载安排，材料响应和操作者判断都改变，这为进一步研究提供线索。但是，共变并不能自行证明相互作用。两种变化可能受到共同条件驱动，也可能来自记录方式的同一偏差。底本本身提醒，相关强弱不能直接充当纠缠程度的充分证据。

互相需要追问得更深：完全去除某一构成角色，并且不引入承担同样角色的替代，另外的属性能否仍按原方式显露？这里的去除不能被理解为只拿走某一台具体仪器。换一种装置仍能实现原有测量，可能说明实现材料可以替换，而不是说明操作角色不需要存在。反过来，也不能在每次检验失败后都宣称“还有一个看不见的替代”，让需要关系永远无法触及。

互为条件强调彼此的成立范围。一次明确的受力问题使某种接触关系变得相关，接触关系又限制什么响应能够被称为这项属性。条件不是一层无所不包的背景，而是能够指明其作用位置的组织。例如端点如何规定影响长度，而颜色偏好在没有进入取点、筛选或操作时，未必是长度模型必需的条件。

互相影响则要求实际变化通道。一次实验结果促使研究者改变后续操作，这是可以追踪的回写；读数被存进一个从未使用的文件，并不能仅凭“记录也是发生”就宣称它已经改变了下一次物理过程。只有承接发生，影响才有具体意义。三属性纠缠不是对所有关系一律宣布强因果，而是要求把哪些关系构成了什么讲清楚。

这四种关系因此不能用一个统计交叉项统一替代。交叉项可能有助于描述某种联合变化，却不自动证明角色必要、条件相互构成或后果已经回写。它们需要不同材料互相限制，才可能形成有内容的整体解释。

3.4 绳索张力：一项属性如何跨过孤立归属

考虑一个明确标为思想案例的系统：一根绳索连接支撑端与被悬挂的物体，装置在所讨论的范围内近似静止。我们希望研究绳索张力。若只说“张力是绳子的属性”，就会漏掉支撑、载荷与具体状态；若只说“张力是测量的结果”，又会抹掉系统对操作的真实限制。

这个案例适合说明三属性并非三个可任意独立拨动的旋钮。改变支撑方式，可能同时改变几何构型和受力关系；改变读数装置，可能需要重新检查连接；改变研究目标，例如从平均张力转向瞬时变化，记录的时间分辨也要调整。研究者能够分析各个方面，却不能保证每次干预只改变一个方面。

本书由此提出一种处理：先说明希望保持不变的是什么，再说明何种操作有资格实现这种保持，最后用独立记录检查保持是否成功。“固定客体、只改变互动”不能停留在文字上。若操作本身已改变载荷、连接或状态，原先的固定声明就应当撤回。

同样，一个电子的电荷也不应因我们采用属性纠缠论，就被说成由操作者意愿指定。研究电荷的具体测量需要理论与实验条件，关于其公共一致性的物理内容则要接受专门检验。本章不从三属性词汇推导新的电荷数值，也不预先否定既有一致性结果。这里要改变的是说明单位，不是未经计算就改写物理事实。

这使属性纠缠论与“任何条件都会影响任何属性”区分开来。后者几乎不可能形成可用的研究；前者需要辨别特定角色、特定路径、特定范围以及可以省略的差异。整体性越有内容，省略与保留的理由反而应当越明确。

3.5 一个数学反例：有耦合，不等于永不收敛

研究对话把多种不收敛现象与三属性纠缠联系起来，并曾提出“只要互相回写，常数就不能稳定”的强表述。[7，第 15—17、37—42 页]这是一项需要检验的研究猜想，不能直接当作数学定理。为了说明区别，下面构造一个仅用于逻辑检验的两变量模型：

$$\dot{x} = -x + \lambda y,\ \dot{y} = \lambda x - y,\ 0 < \lambda < 1.$$

两个变量确实互相进入对方的变化率。令 u=x+y，w=x-y，直接相加、相减，得到：

$$\dot{u}=-(1-\lambda)u,\dot{w}=-(1+\lambda)w.$$

于是 u 和 w 都按指数衰减，x 与 y 也都趋向零。耦合没有被切断，系统却可以收敛。这个计算没有证明真实物理系统一定如此，更没有证明三属性纠缠论错误；它只否定“出现耦合便必然不收敛”这一未经条件限制的推论。

相反，若某个系统出现持续振荡、多个稳定状态或参数漂移，研究必须进一步说明什么结构造成这种行为。反馈方向、作用强度、时滞、状态范围与外部输入都可能进入具体模型。单凭“有回写”不能预告全部结果。第 379 卷对单次随机性、分布稳定与估计收敛作出的分层，在这里尤为重要。[2，第 8、13 章]

对本书而言，这不是削弱创造命题，而是为它清除一个不必要的负担。新的物理世界既可能产生此前不能组织的变异，也可能产生更强的稳定关系。属性纠缠不应只解释世界为什么散开，还应解释某种稳定怎样形成、能够保持多久、如何跨路径承接。

更深的启示是，属性的非独占并不等于数值的无稳定性。一个量可能在充分说明的条件下表现为常量，也可能成为变换中的不变量。问题不在于预先拥护收敛还是不收敛，而在于哪一种发生使当前稳定或变异成立。

3.6　两组三方程，各自承担什么

属性纠缠的三种相互规定，可以写成主体属性依赖客体属性与互动属性，客体属性依赖主体属性与互动属性，互动属性依赖主体属性与客体属性。底本以这种方式概括三属性的联合发生。[1，第 50—51 章]为避免与 SDE 混淆，本书用 A 客、A 互、A 主指称它们：

$$A_{主}=f(A_{客},A_{互}),A_{客}=g(A_{主},A_{互}),A_{互}=h(A_{主},A_{客}).$$

这组式子与 S=F(D,E)、D=G(S,E)、E=H(S,D)的分析层级不同。前者追问属性号位之间怎样相互规定；后者追问显露、差异序列与特征纠缠怎样相互规定。两者可以联合使用，却不能把字母逐项换名就宣布完成对接。

对接的实际工作，是在每一种属性侧重中展开 SDE，再追踪这些发生如何相互进入。比如讨论主体属性，不只说明一个判断结果，还要说明它经过怎样的比较而形成、

依赖怎样的特征纠缠；讨论客体属性，也不只提供一个测量值，而要说明其同一性和差异运行。如此，三属性不再是三个已经完成的数，而是三个带有不同侧重的发生研究。

尤其应当区别主体号位与完整人格。一个装置可以承载先前形成的选择规则，这为主体侧重分析提供入口，却不能由此推出装置有人的体验、价值生活或自我意识。把主体性全部缩成个人心情是不够的，把一切选择机构都宣布为完整主体同样过度。底本的号位分析提供侧重，不是无条件授予某种心理身份。

同样，客体侧重不等于被动。一个具有客体属性显露的系统能够在实际关系中反过来限制操作，迫使判断改变。三个号位之所以不能被做成三个静止盒子，就因为同一发生中的约束与承接可能穿过不同侧重。研究不是寻找哪个盒子最重要，而是追踪这种穿越的具体结构。

这些函数记号首先承担概念义务，并不自动保证可解、唯一或能够独立测量。要成为某个具体领域的模型，还需要变量域、操作定义、约束条件与检验结果。数学表达可以压缩说明，不能替代说明。这个区别将成为下一章的方法基础。

3.7 从属性之问到微积分之问

至此，第一编已经形成一个清楚的转折。我们研究的不再是把属性从主体搬到客体、再从客体搬到互动，而是研究属性的构成怎样跨越独占归属。不同号位的侧重可以明确，明确并不意味着彼此独立；不同变量可以被量化，量化也不意味着发生已经被穷尽。

微积分即将面对的正是这样的变量。一个变化率不只需要“有两个数”，还需要它们在何种意义上是同一量的变化；一个积分不只需要“把很多小量加起来”，还需要这些小量如何属于同一整体。属性纠缠论为这两个问题提供发生学入口，微积分则为其中一部分已经获得数理规定的关系提供强有力的组织。

二者的桥梁不是一句“纠缠等于交叉项”。它是一个更严格的任务：先说明变量怎样成立，再说明哪些差异能够被局部化，最后说明局部关系如何在更广的约束中组成整体。每一步都允许失败，每一步也都可能创造此前尚未成立的物理规定。

属性没有独自的主人，并不使科学失去对象。它要求科学对对象负起更多责任：不仅报告对象显露了什么，还要能够说明这种显露的身份、条件、变化与可承接性怎样共同形成。新的物理世界，正是在这种责任的扩展中获得更具体的内容。

第四章 SDE 三方程与六路径、三原理
研究物理创造的共同语法

4.1 一条曲线提出了新的问题

设想一个用于教学研究的弹性装置。第一次逐步增加载荷，位移形成一条平滑曲线；第二次先加到较大载荷，再逐步卸载，返回路径与第一次不完全重合。这个场景在本章中只是研究设计的思想案例，不代表我们已经完成某种材料实验。它的用途是把方法放到一个有具体分歧的位置上：差别来自装置偏差、材料状态、操作次序，还是原先对研究对象的规定不够充分？

如果回答只是“存在 S、D、E”，研究不会因此前进。如果回答只是“整体大于部分”，下一步也仍不清楚。第 379 卷强调，三方程、六路径与三原理联合使用，应当改变问题、证据安排和修订方向，而不是只提供一套解释性词汇。[2，第 3 章]

我们据此提出一个朴素要求：每当方法被使用，它应当使一项原来不明确的选择变得更明确。是先检查零点，还是增加一种加载次序？是继续拟合同一函数，还是重新定义状态变量？什么材料能使我们放弃最喜欢的解释？这些问题有了进展，方法才真正进入研究。

4.2 三方程不是三条已经求解的自然律

SDE 三方程写为：

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

本书依照底本，将它们首先理解为三维相互规定的表达。[2，第 3 章]显露需要差异序列与特征纠缠的说明；差异序列不是独立流过系统的空时间，它受到已成显露和纠缠组织的规定；特征纠缠也不是永远给定的背景，它会在显露与差异运行中维持、变化或重组。

在弹性装置案例中，S 不是屏幕上任意一个数字，而是经过判断后被认定为目标响应的显露。D 不是仅有先后日期，而包括加载、保持、卸载与重复之间的可比较差异。E 不是装置、样本、操作者三个名词的总和，而是哪些连接、校准、状态与判断使这组差异能够共同形成一个可研究的响应。

同样的数值可以来自不同过程，这会迫使研究者重新辨认 S；同样名称的实验程序可能包含不同操作，这会迫使研究者重建 D；新结果持续无法被既有连接关系解释，则可能要求重建 E。三方程的作用，是不允许其中某一维因研究方便而永久退出说明。

但是，写出三个等号并没有获得一组可直接数值求解的动力学方程。要进行求解，还需要定义域、状态空间、时间尺度、边界与参数。一个概念上的相互规定，也不自动保证任意两维能唯一反演第三维。例如，同一个输出可以对应多个输入组合；若不说明分支或附加约束，逆问题可能没有唯一解。这是数学模型的责任，不应由发生学术语替它承担。

因此，本书将“同时成立”与“同一时刻发生双向因果”分开。三维在整体解释中相互需要，不意味着每一条具体作用都没有时间方向。某次加载先改变响应，新的响应被记录后才改变后续安排，这一先后必须保留。方法不能为了强调整体而抹掉实际过程。

4.3　从 E 进入：沉淀与锚定

六路径中的 E—D—S 称为沉淀型，E—S—D 称为锚定型。[2，第 3 章]两者从特征纠缠进入，却并不做同一件事。

沉淀型适合追踪一个已经存在的组织怎样形成稳定显露。研究者从装置的连接、校准和样本制备进入，查看这些条件怎样通过反复操作形成某种可重复响应。重点不是列出一份条件表，而是追踪条件怎样进入差异。例如支撑方式的改变，是否改变了有效加载路径；静置过程是否使下一次响应能够比较；某项校准是否实际上承担了此前未被注意的判断。

锚定型则可以从复杂组织中先固定一个值得追踪的显露。假如多次试验中某一响应特征相对稳定，研究者可以把它作为临时锚点，再组织新的差异来辨认其成立范围。锚点不是已被证明的本质，而是能够让后续比较共享参照的研究支点。

这两条路线都可能失败。沉淀型容易把“做过这些操作”误写成“这些操作已经解释结果”；锚定型容易只寻找支持鲜明显露的材料。为了避免这种偏向，应当比较条件近似而结果不同、结果近似而过程不同的情形。两类材料能够迫使研究者分清真正承接了结果的关系，与仅仅同时出现的因素。

沿这两条路径获得的 S，也不是研究终点。它要返回检验起初对 E 和 D 的理解。若稳定显露只能在一组狭窄条件下保持，研究应当明确范围，而不是为了保住结论把所有其他条件都称为污染。

4.4 从 D 进入：倒逼与原型

D—E—S 称为倒逼型，D—S—E 称为原型型。[2，第 3 章]弹性装置的两条不同响应路径，正提供了从 D 进入的机会。差别先出现，解释尚未确定。

倒逼型要求从不能被当前说明安置的差异，追查特征纠缠，并据此重建显露或研究对象。这里必须先排除比较错误：两次试验的坐标、单位、目标状态是否相同？差别是否超出已有不确定度？如果连比较对象都没有对齐，就不能把不同曲线直接宣布为新物理。

比较成立以后，研究者可以提出相互竞争的解释。一个解释认为差别来自装置；另一个认为样本状态受到先前操作影响。它们要对下一次试验作出不同承诺，例如更换装置后差别是否保持，改变操作次序后是否出现预告的变化。能够引出区分性操作，才是倒逼型的有效推进。

原型型不要求先完成全部解释。研究者可以依据已有差异制作一个简化装置或最小模型，使问题获得可操作形式。例如保留某种加载与卸载关系，去除暂时无关的复杂部件，观察原型仍能产生哪些结构。这里的试作不是证明新理论已经正确，而是让此前混在一起的条件显出更清楚的分工。

原型失败也可能富有价值。若简化以后关键差异消失，它提示我们删掉的某个关系可能重要；若简化以后差异仍在，研究就可以缩小解释范围。但是，无论成功还是失败，都不能自动兑换成“范式革命”。原型的价值取决于它怎样改变问题和证据，而不是它被赋予了多宏大的名称。

4.5 从 S 进入：进入与再创

S—E—D 称为进入型，S—D—E 称为再创型。[2，第 3 章]前者从既成显露回到其构成，再安排新的差异；后者让既成显露先接受新的比较与操作，借此重组特征纠缠。

例如，我们已经有一条被许多次使用支持的响应关系。进入型不急于怀疑它，也不急于把它推广到所有材料，而是查看它依靠什么样的制备、量程与时间条件成立，随后安排能够区分这些条件作用的试验。稳定成功本身值得解释，发生学不能只在失败和危机处才出现。

再创型则可以把这条关系带入新用途。原来缓慢加载时有效的描述，在改变节奏后是否仍可使用？原来一种几何形状下的结果，在另一形状中需要怎样变换？新差异可能扩大原关系的作用范围，也可能迫使我们缩小其名称，承认它只是更广系统中的一个有效近似。

这样的再创并非否定原成果。一个理论被明确限定范围，可能比它未经限制地传播时更可靠。知识的继续发生，不只有推翻一种形式，也包括建立新的转换、重新解释参数以及把先前成功的部分保留下来。第 342 卷关于典范碰撞、继承和再生的讨论，为这种非单线的研究提供了背景。[3]

六路径也不是六个互不往来的学派。一项工作可能先从 S 进入，遇到异常后改走 D—E—S，再通过原型探索条件。需要记录的不是研究者一开始选了哪个漂亮名字，而是什么材料使原入口不再足够，为什么新的路径能够承担下一步的说明。

4.6　三原理：张力必须有具体位置

三原理可以按维度张力表述为：S 与 D 之间的张力推动 E 改变，D 与 E 之间的张力推动 S 改变，S 与 E 之间的张力推动 D 改变；第三维改变后，还要返回检验前两维。[2，第 3 章]这里的“推动”不是历史必然，也不是一句保证任何冲突都会产生进步的许诺。

在思想案例中，如果我们把稳定曲线解释为与操作次序无关，而 D 明确显示结果取决于先前步骤，S 与 D 便在这一具体解释中发生张力。研究可能重建 E，把状态承接纳入条件组织。张力不是稳定与变化天然相反，而是某个现有说明无法同时容纳已有材料。

如果当前 E 规定声称一切相关条件已经控制，D 却持续产生超出模型范围的分化，也可能需要改变 S 的分类：原来被称为同一种响应的结果，是否混合了不同状态？这种新分类必须帮助解释未参与分类的材料，不能每遇到一个异常就新造一个名称。

如果既成 S 被称为“可普遍迁移的材料规律”，而 E 只包含狭窄的装置和状态，张力首先发生在证据承诺上。需要改变 D，安排真正不同的条件，让普遍性接受检验。这里不能把研究者的证据不足直接写成材料内部的动力冲突。张力在哪一层，必须说清楚。

这些原理与三属性号位之间的张力分析可以配合，却不应混用编号。主体、客体、互动之间的改变，与 S、D、E 之间的改变属于不同分析轴。用其中一组词替换另一组词，可能让表述流畅，却会使真正需要说明的连接消失。

4.7 回写不是对“有关系”的另一种称呼

本书特别保留“回写”的强含义：一次显露或操作的后果，通过可辨认的承接，改变后续的差异与纠缠组织。反馈需要载体、通道与时间。它可以通过材料状态的持续改变发生，也可以通过研究者改变下一轮安排发生，但两者不是同一种机制。

在弹性装置中，一次加载若改变了下一次试验的起始状态，属于对象过程的历史承接；一次读数若使研究者更换测量范围，属于研究组织的承接；一篇论文若改变另一实验室的操作规则，则又涉及公共知识的承接。三者可能连接，却不能因为都叫回写就被视为相同的物理作用。

由此还需要区分三种时间：事情发生的时间、记录获得的时间、解释形成的时间。研究者今天先读到结果，再回溯昨天的操作，不意味着结果在昨天的事件中先于操作发生。六路径记录研究入口，不能代替事件时序。三方程记录相互规定，也不能取消因果辨识。

这种区分使三原理不至于成为万能的事后故事。若说某次新结果改变了条件，就应指出哪项条件在什么时候、经何种承接改变；若找不到承接，只能把它保留为候选解释。没有材料时说“尚待确认”，比用整体性语言填满空白更能保护理论的创造能力。

4.8 方法怎样接到微积分上

在一项物理研究中，SDE 方法可以帮助我们辨认对象、差异和纠缠；微积分则能够在部分关系已经获得明确数理规定以后，组织局部变化与整体重建。二者既不相互取代，也不彼此隔绝。

一条微分方程需要变量与变化率，其物理含义可以返回 SDE 分析；一个积分解需要初值、边界与相容关系，其成立范围同样可以返回 SDE 分析。反过来，数学推导可能揭示原来没有注意的约束，使研究者重新划分 S、重新安排 D、重新理解 E。桥梁不是把三方程“求导”，而是使不同层级的问题相互提供实质限制。

例如，若同一组局部方程容许多个整体解，发生学不能用“整体纠缠”替我们选定其中一个；它应当推动对实际初始与边界关系的研究。若一项微分模型只能在光滑变化中使用，发生学也不能把断裂事件强行塞进现有导数；它应当帮助定位模型需要换一种组织方式的地方。

因此，可以为弹性思想案例形成一份很小但有用的研究记录：哪些响应已经取得，哪些加载次序已经比较，哪些条件确实受控，哪些解释仍然竞争，以及下一次操作将区

分什么。每个字段都要有资料对应，不能只写一个 S、D 或 E。一次修改以后，还须保留旧解释为何曾经合理、哪项新材料使它不足。

这样的记录也保存创造，而不只保存正确答案。研究者曾经选择过哪一种变量、放弃过哪一种近似、为何改变了比较范围，都可能成为后来者的发生资源。它们不是应当全部塞入最终方程的外部细节，而是使方程能够被理解、修订与再创的承接路径。科学知识的公共生命，正需要这种能够返回发生过程的能力。

方法还需要衡量研究代价。若一种更复杂的描述没有增加区分能力，却要求保存全部无关资料，它未必比有根据的简化更好。三维联合并不意味着每次都最大限度展开，而是使展开与封装有可说明的理由。准确、可行动与整体协调应当在研究中相互限制，不能用形式丰满代替真正的解释增量。

这就是共同语法的意义：并不要求所有物理创造长成相同形状，却要求每一种创造说清自己怎样成立、怎样接受改变、怎样被后来者接续。方法的成熟，不在于能够解释所有故事，而在于它使新的研究问题获得更精确的形状。

第五章 如何证明“新物已经发生”
创造判据与反例

5.1 新名字可以来得很快，新物不能靠宣布成立

设想有人为一根普通绳索起了一个新名称，随后宣告发现一种新的物理存在。另一个人没有改名字，却设计出此前没有实现的操作，使某种稳定响应能够被重复产生。还有人只改变数学表示，却由此打开一大类过去难以处理的问题。这三种变化，显然不应由“都是创新”一句话统一结算。

本书提出“物理学思想史是一部创造史”，就有责任辨认创造具体发生在哪里。若任何措辞变化都可以证明新物发生，命题没有区分力；若只有制造一件肉眼可见的新器具才算创造，又会把状态、场和新的公共规律排除在外。我们需要一种能够容纳不同创造，又不取消差别的判据。

这一要求来自两部底本的共同约束。《属性纠缠论》追问规定性与承载者怎样共同成立，第 379 卷要求新方法改变可研究的对象、证据与行动，而不是只给旧材料贴标签。[1；2，第 3、10、15 章]本章在此基础上提出三项工作检验。它们是本书的研究工具，不是已经被底本证明的普遍定理。

5.2 五个层面的新，不能自动相互兑换

第一层是名称的新。名称可能很重要，因为它使某种此前难以交流的区分获得公共入口。但是，名称的出现本身不证明所指的物理规定已经改变。一个名称能否承担新研究，要看它怎样连接判断、操作和证据。

第二层是数学表示的新。新坐标、新符号、新计算形式可能显著改变处理问题的能力，即使它们在一定范围内与旧形式严格等价。这样的创造不必因等价而被贬低，却也不能仅凭符号不同就宣布经验世界发生断裂。表示的生产力与新的物理约束，需要分别展示。

第三层是操作对象的新。研究者能够制备、区分、操控或迁移此前未能以同样方式处理的系统。这里的创造已经超出名称，但仍需要辨认：得到的是旧规律下的新实现，还是要求改变对象构成的现象？二者都有价值，证明责任却不同。

第四层是经验约束的新。新系统对独立观察作出此前没有得到的限制，使某些可能结果被排除，另一些关系得到支持。一条漂亮曲线若仅重述已经用于拟合的资料，尚不足以承担这一层的承诺。需要有未参与构造的材料，或者真正能区分竞争解释的操作。

第五层是本体规定的新。何者可以作为物、怎样成位、哪些变化仍属于同一对象，被重新组织。这一层不能由前四层中的某一项机械推出，却可以受到它们共同支持。它要求说明新的构成关系为什么不只是外部认识条件，以及删去它以后原有规定性是否仍然完整。

五层不是简单的荣誉阶梯。一次重大的数学重写可能主要发生在第二层，却为后续物理创造铺路；一次精确测量可能主要强化第四层，而不立即要求第五层改变。把所有成果都包装成本体革命，会让真正的本体问题失去重量。

5.3　构成条件删除检验：究竟删去什么

第一项工作检验是：删去被宣称为新增构成的角色，而且不偷偷引入等效替代，原有物理规定能否仍以同样方式成立？这里关键在于“角色”，而不是具体零件。

例如，某项长度测量使用一把特定尺子，换另一把经校准的尺子仍然完成同样测量，不能据此证明长度不需要比较结构。它可能只说明具体实现可以替换。反过来，若拿走尺子便测不到数，也不能立即证明“尺子制造了全部长度”。任何测量工具被移除都可能造成认识困难，认识困难与物之构成被撤销之间还需要论证。

因此，删除检验必须问得更细：原来需要成立的属性是什么？新增关系承担什么作用？替代是否保留该作用？失败发生在信息获取、操作实现，还是属性定义本身？只有区分这些层次，才能避免把“我不知道”误写成“它没有”，或把“我能换工具”误写成“任何构成条件都不需要”。

以变化率为例，删去时间比较或其他参数化关系，仍可以谈物体位置，却不能完整规定同一个速度问题。这里失去的不是石头的全部持存，而是某一类运动规定。第一编已经保留发生层持存与属性显影的区别，删除检验必须遵守这个边界。[1，第 15 章及“互相需要”相关论述]

删除检验也允许反驳。如果所谓新增构成被删去后，旧系统通过已经具有的关系完整地重建全部相关操作与约束，那么新解释就可能只是等价重述。这不必否定其表达价值，但会限制它宣称新物发生的范围。

5.4　跨路径迁移检验：新系统能否离开自己的出生现场

第二项检验，是改变实现路径，保留或者明确变换所需关系，观察新规定能否被重新发生。一个只在单一叙述中成立的对象，与一个能由不同操作接续的物理系统，承担的公共责任不同。

这里并不要求所有实验得到同样数字。应当先说明跨路径保持的是什么：某个量、某种分布、某条变换关系，还是一组允许与禁止的过程。若保持内容不明确，任何相似都可能被赞为成功，任何差异也都可能被说成“纠缠不同”，检验便没有实际作用。

设想一项弹性描述在装置甲中形成，随后进入装置乙。两套装置的读数可以不同，研究者需要说明这种不同是否来自单位、几何、校准或真正不同的状态。只有转换关系能够提前规定，迁移才不是事后寻找相似。如果每次出现新资料都任意改动转换函数，所谓跨路径稳定便失去证据力量。

迁移失败也不一定证明对象不存在。它可能揭示原来对象的适用范围较窄，可能说明实现条件尚未成熟，也可能要求重建对象身份。方法的价值在于帮助区分这些可能，而不是预先规定失败只能有一个哲学答案。

这一检验特别适合连接视觉、听觉、触觉 SDE。跨感觉转换能够增加辨认与操作能力，但不能把不同通道的显示方式当作互相自动证明。例如把数列转成声音，可以帮助发现节奏，却不意味着声音已经独立验证数列的物理来源。转换带来新的可显露性，独立验证仍需要不同约束。

5.5　继承与不可还原：新物不是与旧物彻底绝缘

第三项检验，是追踪新系统继承了什么，又在哪些明确意义上超出旧系统。第 342 卷用典范三组合讨论生成组织；本书据此不把一组新结果与一个新典范直接等同。[3]一种稳定的生成组织，应当能够支持多个问题，而不只解释一项孤立的成功。

所谓不可还原，也必须限定层级。数学上不能由某种变换获得，与实验上出现新约束不是同一件事；某种表示难以计算，与原理上无法表达也不同；一个新概念改变解释方向，更不能由此直接推出旧方程在原范围内失效。

可以设想两种数学形式，经明确映射后给出同样可检验结果。它们的发生路径可能不同，计算便利也可能不同，但我们不能仅凭路径不同就宣告经验不等价。相反，这种贯通本身可能是一项创造：它让不同问题的资源进入同一研究组织，扩大可操作范围。

还可以设想一个新模型在某个极限下回到旧模型。这并不证明新系统只是换皮，因为旧模型的适用边界、参数意义和整体位置可能被重新规定；但它也不许可我们把每个新旧关系都自动写成严格包含。有的继承是精确映射，有的是近似，有的只保留部分问题上的成功。每一种都要分别说明。

因此，本书不以“旧世界全部死亡”证明新世界诞生。新的物理世界可以使旧系统成为一个有范围的组成，也可以与旧系统在不同层级并行。发生史需要同时写创造、保持、失效与再组织，否则它只剩一串激烈的替代故事。

5.6 一张曲线的三种不同命运

为了把三项检验连起来，再看一个思想案例。研究者得到一条过去未曾报告的响应曲线。有三种完全不同的可能。

第一种，曲线来自读数零点的漂移。修复校准以后，异常消失。这里有真实的研究与技术进步，但没有必要宣布新物理本体。若发生学只允许异常支持自己，不允许校准解释获胜，它就不是研究方法。

第二种，曲线来自旧理论已经容许、但此前未实现的新操作条件。通过制备与控制，研究者稳定地做出了这种响应。它可以构成新操作对象或新材料能力，而不要求已经有一条新的基本自然律。把这种创造贬为“只是应用”，与把它夸成“推翻所有旧物理”，都不准确。

第三种，曲线在独立装置、不同操作路径与事先规定的检验下持续出现，旧模型无论怎样使用其合法参数都不能同时满足相关约束。此时有理由重建模型，甚至进一步重建对象身份。但是，哪一种新解释最合适，仍需比较，而不是由“旧模型失败”直接推出 SDE 解释唯一正确。

这个案例说明，创造判据不能只奖励肯定答案。它要帮助我们把不同层面的成功与失败放在准确位置。新的经验并不天然带着自己的解释，新理论也不天然拥有所有经验。二者必须通过具体关系互相限制。

5.7 数学闭合不是物理闭合

一条关系经过微分又积分，能够返回原形式，首先说明所使用的运算相容。它并不自动证明这条关系描述了某个实际系统。我们完全可以对一个虚构函数完成正确运算，而没有产生任何新的物理知识。

同样，一组方程具有解，不意味着实验中一定能够实现这个解；方程有唯一解，也不意味着建模时没有遗漏重要变量。数学闭合与物理闭合之间，需要变量的操作含义、成立条件以及独立的经验约束。微积分的创造性，恰恰要在这条连接上展开，而不能借数学正确直接替换物理成立。

反过来，数学上的困难也不应被立即翻译成本体论错误。方程难求解，可能是计算能力不足；数值结果发散，可能来自算法不稳定；解不光滑，可能要求不同的函数空间或弱形式。每一次困难都值得研究，但不能全部被收编成同一种“旧本体论失败”。

这并没有把哲学赶出科学。哲学可以帮助我们辨认何时把一种表达习惯误当作存在必然，何时把可解性误当作真实性，也可以推动新的变量与研究单位出现。它的力量来自精确地进入这些区别，而不是提前宣布哪个学科拥有最后裁判权。

5.8 第一编的结论：物的创造有了可承担的内容

第一编从石头出发，经过物体、物质与物理的共同形成，进入三属性纠缠与 SDE 方法，最后来到创造判据。我们尚未通过这几章证明所有物理学史都必须被同一种模式解释。我们已经获得的是一组可以携带到历史与数学中的问题：一种规定怎样成立，哪些差异构成它，什么角色不可省略，什么关系能够迁移，什么内容不能被旧组织完整承担。

由此，“创造新的物理世界”不再只是比“发现新知识”更有气势的说法。它要求具体展示新的发生系统，而不是只展示新结论。物理学家的创造可以表现在新的物体身份、物质类型、局部作用、整体约束与公共操作中；这些层级互相进入，却不能未经论证就视为同一件事。

接下来的第二编先回到多种文明的测量、天文、运动与实验传统，考察新的研究对象怎样逐渐获得可比较形式。第三编再把这些要求带到微积分的形成与使用中：微分怎样使局部变化取得新的规定，积分怎样在初值、边界与整体相容中形成可检验的运动，哪些数学创造同时重组了物理对象。这些问题需要历史材料与实际计算共同回答。

本体论的工作，至此不是替数学预先写好胜利结论，而是把数学即将接手的对象打开。数学的工作，也不是为本体论提供华丽符号，而是让它的桥梁接受约束：能够证明的就证明，能够计算的就计算，尚需经验材料的则留下明确的位置。只有这样的连接，才能使这部思想史的创造主张获得比口号更持久的生命。

本编参考资料与第三编合列于第十五章后。

第二编 从自然秩序到可测的运动
前近代物理的多条源流

第六章
五大文明如何生成天、地、声音与重量的秩序

6.1 天空何时成为能够共同使用的天空

一个人仰望天空，看见亮点出现、移动、消失；另一个人也抬头，却未必取得同一个可以传递的天象。要让两次观看成为“同一颗星在不同日期的位置”，必须先形成辨认、定位、命名和记录的连续工序。本章不把生成理解为书写凭空发出了星光。能够被多人追踪、被历表安排、被后来的观察纠正的天象系统，不能离开这些工序而成立。本章所谓文明生成物理世界，首先指这种可重复显露、可共同操作的存在系统的生成。

中国、两河、埃及、印度与希腊不是五个同时出发、参加同一考试的选手。“文明”在这里是整理材料的尺度，不是拥有固定性格的超级主体。每一区域内部都有跨越多个世纪的断裂、迁移、职业差异和知识争论；区域之间也有接触。我们选择若干可定位的文物和文本，比较它们怎样使天、地、声音、重量取得稳定规定性，不据此宣布某种民族天生擅长抽象或只能积累经验。

对比的单位因而不是孤立的观念，而是一次能够继续发生的实践。例如，某块石头成为砝码，需要它进入与其他物件的比较序列；一次声响成为某个音，需要击打、辨听、再击打和校正的连续性；一片土地成为可分配的面积，需要界线、单位、操作规则与接受该结果的共同体。三者都包含可见或可听的显露，也包含不能从那一次显露中直接读出的构成条件。

本书以 S 表示显露，D 表示差异序列，E 表示特征纠缠。这里的 E 不是把王权、宗教、农业和贸易全部装进去的背景袋，而是追问：究竟哪些特征通过怎样的相互制约，构成了这个具体结果的成立条件？如果删去某一联系，结果仍然以相同方式成立，就不能仅凭“它也在当时存在”而把它宣布为必要纠缠。

6.2 两河：从一次观看到可以继续核对的天象

大英博物馆对《MUL.APIN》第一泥版的馆藏著录，将它列为天文汇编，涉及天空分区、星体升起与同升同落以及月亮路径中的星群。[1] 馆藏著录能够确认材料的性质，

却不能代替泥版逐行校勘。仅凭这些信息，不能推断某个算法的精度，更不能把保存下来的抄本日期当作知识的唯一诞生日期。

这个有限证据仍足以打开一个关键问题：清单为何具有思想史意义？如果星名只是一串声音，它不能把今天的观看接到明天；如果日期只是一串数字，它也不能保证同一观察时段得到重复。只有星体识别、时序划分与书写位置能够彼此对应，表中的项目才成为后续观察的入口。清单使差异具有可以重返的位置，使一次经验能够被另一次经验质询。

用三方程考察，S=F(D,E)要求说明被认作稳定天象的显露怎样形成；D=G(S,E)要求说明已经形成的星群与时序规定怎样重排下一次观察；E=H(S,D)则追问反复观察和修正怎样使某些对应关系稳定下来，另一些关系被放弃。这里的函数符号表示研究义务，不是已有参数、能够计算古人行为的经验方程。若缺少泥版系列和具体年代资料，不能伪装成已经完成了三项历史证明。

尤其不能把制度的力量误写成任意创造天象。记录者可以规定从何处开始计时，却不能仅凭宣布，让预期的星体在被云遮蔽的夜晚显现。反过来，天象也不会自动替人决定应当记录什么。这个相互限制的区间，才是物理世界创造命题必须面对的地方：创造是一套新的规定性及其检验方式成立，不是意愿取消抵抗。

6.3 埃及：尺不是缩小了的土地

大都会艺术博物馆收藏的一件木制度量尺，编号 15.3.1128，来自利什特北部，馆方将其定为中王国时期。[2] 这件物品首先证明一种量具实践留下了物质见证；它本身不能证明所有埃及工程都使用同一规格，也不能独自证明某一洪水事件直接造成几何学。

把尺放到土地旁边，是两个此前可以各自存在的实践发生结合：人的身体和器具能够逐段移动，土地上的边界能够按某种规则被辨认。两者结合以后，“多长”才不只是视觉上的远近印象，而成为可以搬运的比较结果。真正搬运的不是整块土地，而是一套在另一处能够重新实施的操作。它要求端点可确认、单位可复现、接续不漏段；任何一项失效，都可能使数字看起来确定而结果并不可靠。

尺上有刻线并不意味着测量已经完成。斜置和水平放置的差别、弯曲地面与所取直线的差别、身体到达不了的边界与想象延长线的差别，都要在实践中获得处理。这里可以看见日后数学物理问题的雏形，但不应将古代量地直接命名为微积分。微积分需要它自己的变量、极限与运算条件。前者提供的是一项更早的构成性问题：什么东西可以作为同一单位在不同位置继续使用？

一个现代教学例子可以把这一点说清。在平面欧氏模型中，若长宽分别由同一种长度规范确定，矩形面积可写作 A=ab。这个式子要求矩形边界已经成立，要求单位平方可以铺展，并非任意四边形都可把两个边长相乘。公式的正确推演不能替代对象条件的成立。我们借此检验自己的叙述，而不把这个现代说明反投为某位古代测量者留下的自述。

6.4 中国：编钟使声音取得可复返的身份

湖北省博物馆发布的曾侯乙编钟资料记载，这组出土器物共有六十五件；钟体、钟架及挂钩保存乐律铭文，部分说明涉及音名及其组织。合瓦形钟体能够在不同击打位置产生两个不同乐音。[3] 本章只采用这些有限事实，不由现代试奏效果推断每位古代听者具有相同听觉经验。

编钟的思想史价值不只在于“古人已经知道音高”。更深的变化在于：声音能够取得一个既可辨认、又可再次引出的身份。钟尚未受到击打时，铭文并不等于实际声音；一次击打以后，余响消失，名称却能引导后来的演奏。听觉中的稳定辨认、触觉中的击打差异、视觉中的器形与标记，并不是三块零件各占 S、D、E 一个位置。每一种感知活动都有自己的完整发生结构，它们又在演奏中相互纠缠。

由此，“同一个钟”与“同一个音”并不是同一种同一性。如果只按物件身份辨认，就可能忽略击打位置不同所生成的音差；如果只按一次听见的音辨认，又可能忽略它怎样依赖支撑、击打和器形。属性纠缠分析所追问的，正是这些同一性何时可以分开、何时必须联合。物件并没有因为名称固定，就拥有一个不问操作条件的单一声学身份。

我们也不能把礼乐直接降格为声学知识的外包装。共同演奏需要什么时候发声、与谁配合、什么差别值得纠正，这些要求能够组织练习和制造。可是，礼制认可某个声音，并不能证明关于振动的一切解释都正确。要区分社会上被接受的乐音、操作上能够复现的声响以及理论上充分解释的振动。三者可以相互推进，也可能长期不重合。物理学思想史应当保存这种不重合，而不是把前两者自动写成现代声学的未完成版本。

6.5 印度：重量与祭坛不能压缩成同一条起源链

大英博物馆著录的一件印度河文明燧石砝码为立方形，编号 1939,0619.231，年代范围列为公元前 2500 至 2000 年。[4] 《百道梵书》另一处文本则把祭坛中的方向、土块与仪式操作联系起来。[5] 两类材料属于不同问题与历史层次；将它们并列，不意味着已经证明印度河砝码制度连续传承为后来的祭坛知识。

砝码提示我们把“重”分成两个层次。举起石块时的费力，与用比较装置取得的重量关系，不是同一种显露。前者随身体状态和动作而改变，后者通过规定比较程序来约束这种变化。因此，可转移的重量知识不是把个人感觉放大，而是使不同人的操作在某些条件下能够接续。不过，没有具体器具组合和出土关联，我们不能仅从一块石头复原所有交易程序。

祭坛文本则显示另一种值得研究的可能：空间操作可以同时具有仪式意义和可实施的顺序。方向不是脱离身体行动而悬浮的名词，布置也不是无条件服从几何定理。历史研究需要继续问，哪些位置必须按规定安排，哪些数值可以调整，错误由谁判定。原典中确有仪式对应，并不等于这些对应已被现代自然科学证实；但把它们一概删去，也会丢失实践为什么持续、为什么需要精确重复的问题。

这两类材料共同迫使我们拒绝一个过于省力的叙述：先有迷信，后来理性清除迷信，科学便出现。操作的精确性可以在非现代宇宙论中成长；宏大宇宙论也可能保护某些不接受操作反驳的解释。需要逐项考察的是纠正渠道。如果重量比较不符，可以重做比较；如果祭仪解释遭遇失败却总能另换说辞，它的可修正条件就不同。比较发生机制，不能用笼统的“东方智慧”把差异消融。

6.6 希腊：“自然物”本身成为待规定的类别

亚里士多德《物理学》第二卷第一章，以自身具有运动与静止的起因讨论自然物，并以床等人工物作对照；第二章又区分数学与自然研究的处理方式。[6] 这不是现代“自然科学”学科目录的提前发表，而是对于哪些对象可以一起研究、凭什么一起研究的概念工作。

在本书视角下，分类不是观察结束以后的贴标签。把某类变化认作来自事物自身，会引导人去寻找某种原因；把另一些变化放在制作活动中，则会组织另一套问题。因此，分类能够回写差异序列：什么需要比较，什么属于反例，什么才算解释完成，都会随之改变。但我们不能据此断言亚里士多德已经采用 SDE 本体论。他的论证具有自己的概念前提，正应成为本书对话与批判的对象。

一个竞争解释会说：分类只改变了关于世界的说法，世界本来就在那儿。本书并不靠否定这个句子中的一切来回答。应当具体比较两套分类下，哪些物能够被识别为同类，哪些操作成为有效提问，哪些新显露取得可重复身份。如果全都不变，仅换名称，便不能称为生成了新的物理规定性；如果对象边界、操作和检验同时改变，变化就超过词汇替换。创造命题因此必须付出举证责任，而不能通过宣言免于检验。

6.7 六条路径如何防止文明比较成为贴标签

三方程给出相互依赖，六路径则要求我们改变进入问题的次序。本章把S、D、E的六种顺序作为研究进路，而不把箭头看作古代事件已经得到证明的单向因果。由S进入D再到E，是从一个稳定读数追查操作与构成联系；由S进入E再到D，是从一个被接受的乐音追查它的配合关系，再询问怎样重做；由D进入S再到E，是从连续观察辨认显露，再检验其相互约束。

另三条同样必要。由D进入E再到S，可以先考察操作差别怎样改变联系结构，随后看新显露能否成立；由E进入S再到D，可以从已有对应网络辨识它允许什么显露，再寻找复验办法；由E进入D再到S，则追问改变已有联系后，操作怎样调整、结果怎样变化。这六种阅读不能都停在文字上。每一条都要有材料入口，缺少入口时应明确留空。

例如，从编钟音名出发与从击打差别出发，可能走向同一组器物，却不保证得到相同历史解释。前者容易突出共同体的命名，后者容易突出制造与演奏。如果两路相互校验，便会发现仅有命名不足以造成声响，仅有声响也不足以形成可教学的乐律。六路径的意义在于迫使解释接受其他入口的限制，而不是把每个史实重复讲六遍。

6.8 张力所推动的第三维改变，必须回到证据

三原理在本章以两维之间的张力推动第三维改变并回检的方式使用。既有显露与操作序列不相容时，要追问纠缠结构是否需要重组；既有操作与联系结构相互牵制时，要检验能否出现新的显露；既有显露与联系结构要求不一致时，要尝试重排差异序列。这是研究者组织证据与提出问题的方法，不是已经替五种文明找到了统一历史发动机。

以历表为例，表中预期与后续观看冲突，可以推动重新组织观测时间、辨认规则或记录关系。但若只改解释来保住表格，张力并没有取得可检验的解答。以测量为例，反复量得不同结果，可以促成单位和接续程序的重构；然而宣称统一单位，并不等于器具已经一致。所谓回检，必须再走向独立操作，看原来的冲突是否减少，以及新的方法是否引入另一种误差。

五组材料于是呈现出不同的物理世界生成入口：星表让天象获得可接续的位置，量具让长度取得可搬运的比较规范，编钟让声音身份进入制造和复演，砝码让轻重比较获得器物支撑，概念分类则重新规定什么对象与什么解释相配。它们不是彼此隔绝的五条本质道路，也不是注定汇入同一个终点的五级台阶。

下一步要研究的是，这些局部稳定化如何形成更广的物理体系。这里仍有一个尚未解决的问题：当一种操作在小范围内可靠，它凭什么能够越过器具、地点和人的差别而推广？后来的物理思想不断重新回答这个问题。任何宏大理论都不能跳过这一关，因为物理世界的创造，既要使新规定性能够发生，也要说明它在何处持续、在何处失效、又如何接受重建。

本章参考资料

[1] British Museum, tablet 1899,0610.108，https://www.britishmuseum.org/collection/object/W_1899-0610-108。核验范围为馆藏著录，不涉及泥版全译。

[2] Metropolitan Museum of Art, Cubit measuring rod, 15.3.1128，https://www.metmuseum.org/art/collection/search/557017。已读馆藏元数据。

[3] 湖北省博物馆，《曾侯乙编钟—甬钟—局部 1》，Google Arts & Culture，https://artsandculture.google.com/asset/the-chimes-of-zeng-hou-yi-part5/oAFUFT9J-ZAoJQ?hl=zh-cn。已读馆方展示说明。

[4] British Museum, weight 1939,0619.231，https://www.britishmuseum.org/collection/object/A_1939-0619-231。核验范围为馆藏著录。

[5] Sátapatha Brāhmaṇa VII.3.1，Julius Eggeling 英译，https://sacred-texts.com/hin/sbr/sbe41/sbe4161.htm。已读第 20—27、35—47 段及注释，未校梵文。

[6] Aristotle, Physics II.1–2, 192b8–194a18，希腊文／英译对照，https://www.logicmuseum.com/wiki/Authors/Aristotle/physics/liber2。已读本章所引段落，未独立校勘希腊文。

第七章 亚里士多德：性质、位置与自然运动

7.1 先问什么变化才属于这个物

把一块石头放到火旁边，可以观察它的位置、温度和表面的变化。把它移开，位置改变而石头仍被当作同一块石头；把它打碎，又会出现“还是不是原来那个物”的问题。日常经验似乎知道如何回答，可一旦追问判据，困难便显现：为什么某种变化只改变一个物的状态，另一种变化却使原来的物不再成立？物理学不仅需要描述变化速度，还必须规定变化着的是什么。

亚里士多德的重要性正在于，他把这种规定工作放进一个广阔的自然研究体系。上一章讨论分类如何参与对象生成，本章则进一步考察：当性质、位置与运动被组织成相互支持的解释时，一种物理世界如何获得连续性，又在什么地方把自己的起点当成了不再需要说明的东西？我们不把他写成等待牛顿纠正的错误清单，也不把他改装成 SDE 的古代代言人。

《物理学》第二卷区分自然物与人工制品，并从自然物自身的运动与静止根源说明其自然。第三卷讨论运动时，则把可能性与实现之间的关系放在中心位置。[7-1][7-2] 这些概念不能直接替换为今天的位移、速度和加速度。“运动”在相应论述中处理的范围，比一个质点从甲地到乙地宽得多。

本书的分析从一个现代读者能够实际辨认的难题进入：要判定发生了变化，必须已有某种保持；要判定保持的是同一物，又必须允许某些差异出现。如果把保持与变化各自封闭，前者无法说明新状态，后者无法说明究竟是谁改变。亚里士多德提供了一套回答这一难题的概念结构，而 SDE 研究要进一步追问这套结构自身怎样获得显露与可复用的效力。

7.2 四种为什么，不是四只看不见的手

《物理学》第二卷第三章区分材料、形式、变化的起源以及所为的目的等不同原因；同一事物可以在不同意义上具有多种原因。[7-1] 读者不宜把四因理解成四个依次推撞物体的力量，也不能把其中某一项不符合现代用语，就宣布整个问题没有价值。

以制作一只陶杯作教学例子，泥料回答它由什么形成，杯形规定它为何能够盛水，制作操作使这个形状取得现实的稳定，使用目的则影响工匠为何选择这种口径与把手。这个例子是本章为说明不同提问层次而设，并非某段原典的逐句重述。四个答案并不相互排斥，因为它们回答的不是同一个缺口。

然而，答案可以互补，不代表它们已经具有同样的检验资格。泥料是否适合烧制，可以通过材料与工艺比较来检验；使用者是否喜欢杯形，需要另一套比较；杯子为何存在于宇宙之中，则不能只凭“它适合喝水”得到同样确定的结论。把解释层次展开，是一种能力；辨明哪些层次之间可以迁移，哪些不能直接迁移，是另一种能力。

四因在本章的价值不在于提供现成的 SDE 映射。材料不是天然等于 E，制作不是天然等于 D，形式也不是天然等于 S。若讨论泥料的可塑性，它自己就需要显露、差异序列与特征纠缠的完整分析；若讨论目的，目的也通过选择、表达和行动的相互约束而成立。将四因压进三个格子，既损坏原有问题，也使本书方法沦为名称替换。

真正有生产力的对话是：同一个陶杯的发生，为什么不能只靠任何一类答案完成？进一步还要问，材料、形状、操作和用途的划分本身，是否在不同实践中保持不变。一只破杯可以变成花盆，碎片可以成为修补材料。用途改变未必立即改变物质组成，却可能重新规定什么损坏可以容忍、哪些边界必须修复。这种对象身份的变化，为属性纠缠研究提供了比简单对应更具体的入口。

7.3 元素不只是四个熟悉名词

《生成与消灭》第二卷以冷热、干湿等性质讨论元素，给出火、气、水、土与四种性质组合的对应，同时讨论元素之间的转变。[7-3] 这些“元素”不等于现代元素周期表中的元素；把“水”直接替换为水分子，或者把“火”替换为某一种粒子，都会改写原来的论证。

这种体系值得关注的地方，是它尝试让物的性质不再只是无序堆积。只要少数性质之间的组合与转变能够组织更多差异，自然研究便获得一种生成语法：读者不必为每个新物另造一个完全无关的解释。一个简洁体系能够把大量经验放进共同讨论，这正是它具有吸引力的原因之一。

但简洁同时带来风险。把某种性质宣布为基本性质之后，就可能以它来解释大量其他性质，却不再追问它自身如何被识别。冷热来自比较还是来自不可再问的本性？干湿是在什么操作下成立的区分？当一块材料在表面显干、内部含液时，研究者究竟依据何种尺度决定其归类？这些问题不是用现代仪器嘲笑古人，而是考察一个性质体系必须承担的识别责任。

从 SDE 看，“热”如果是某个具体互动中的显露，就需要与接触时长、比较次序、材料状态等共同考察。把它当作某个独立成品身上的固定标签，可能遮蔽它的发生条件。但这不等于任何冷热判断都随心所欲。同一套操作可以受到重复结果的限制，不同

操作也可能通过校准形成可转换关系。问题在于建立这种限制与转换，而不是在“纯客观”和“纯主观”之间选择一端。

这里还应区分理论批评和历史判决。我们可以指出某种基本性质不足以支撑现代材料分析，却不能据此说使用这些性质的人从未取得可靠判断。厨房、窑炉和手工业中的局部可靠性，与一个体系是否能够跨全部场景推广，是不同层次的问题。后来物理学的创造，并不只是用正确词语取代错误词语，而是重新生成哪些性质能够作为公共量、怎样转换、何处失效的组织方式。

7.4 位置为何带着方向与意义

《论天》第四卷把趋向中心与离开中心的运动联系于重与轻的区分，并在这个框架中讨论土、水、气、火。[7-4] 这是一种带有中心与层次的空间秩序。它不能直接等同于现代均匀引力场，也不能只画出“向上”“向下”两支箭头，就认为已经解释了原体系。

现代坐标纸上的上方，可以因为纸张转动而成为下方。可是一个由自然位置组织的世界，并不把方向完全交给画图者的方便。某种方向与某类物的运动关联，运动又支持这类物的身份。于是，物是什么与它向哪里去，在解释上发生紧密连接。这里需要注意的不是古人缺少坐标，而是他们让空间承担了不同的本体论工作。

假设一个解释把某类物的落下归结为它趋向自己的自然位置。这个回答能够把多次落下组织成一个可理解的类，却未必能够进一步给出到达所需时间。类别解释与定量预测的差别在此显现。前者可能回答“为什么朝这个方向”，后者还必须说明“在这些条件下，以怎样的时间过程前进”。不能因为前者没有完成后者，就否认它曾组织问题；也不能因为它组织了问题，就替它宣布所有后续预测已经包含其中。

本书的批判落在一个更准确的位置：如果物的身份由运动归类，又以同一身份解释该运动，而没有独立区分条件，就可能形成解释闭环。看见它下降，于是称它具有下降的自然；它为什么下降，因为具有这种自然。要让闭环打开，必须增加能够区分解释的材料，例如同一对象在不同介质中的变化，或同类对象在不同准备条件下的变化。

这些比较在现代读者的实验台上容易提出，但不能倒推成历史必然。什么值得比较，需要已有问题引导；器具能否支持稳定比较，又限制了哪些反例能够清晰显露。一个理论的边界，并不总是被一个早就摆在眼前的孤立事实击穿。经常需要新的差异序列，原本可以被忽略的偏差才会成为无法回避的问题。

7.5 运动分类也在塑造观察者

为了进一步分析，设想三个当代教学场景：球滚过桌面，金属被加热，种子长成幼苗。我们可以分别记录位置、温度和形态，但若只说“三者都在变化”，并没有完成解释。它们需要不同的保持判据、不同的时间尺度和不同的操作。把它们完全分开，又可能错过加热与形变、位移与环境交换之间的联系。

分类首先提供一种节约：它帮助研究者决定先记录什么，暂时不记录什么。这个节约不是无代价的。只追踪球心位置，会把球的转动、形变和摩擦热压到别处；只记录材料温度，会暂时隐去温度读数怎样受接触方式影响。分类因此不是对完成世界进行最后归档，而是预先安排后续显露的通道。

亚里士多德的体系让多种变化进入有组织的讨论，现代物理则不断重组这些讨论边界。SDE 分析并不因此把前者判作低级、后者判作终点。它要求逐次说明：新增的边界在哪里，旧有的边界为何失效，哪一种观察能力因此获得扩大，哪一种差异又被新的简化暂时隐藏。创造物理世界的命题，只有落实到这些变化，才不只是宏大修辞。

观察者也在这个过程中改变。一个熟悉分类的人能够注意到初学者忽略的差别，却也可能太快把异常归入既有类别。知识既能打开显露，又能提前封装显露。因此，研究者需要保留一种延迟归类的能力：先记下与预期不一致的细节，不急于把它称为杂质、偶然或操作失误。延迟并不意味着停止判断，而是给可能改变分类的证据留出位置。

7.6 不能把目的从制品无条件推广到自然

我们知道陶杯有用途，往往因为能够追踪设计、制作和使用之间的实际联系。若把同样的说明用于自然过程，则需要另行论证，不能只因为结果有序，就断定它事先以这个结果为目的。反过来，仅仅拒绝目的语言，也不会自动得到充分的机制说明；说“没有目的”不能代替说明一个过程为何稳定出现。

这里可以提出两个互相竞争的研究任务。一种任务考察某结构怎样有助于某种持续活动；另一种任务考察这个结构怎样在历史中形成。前者揭示功能关系，后者揭示形成过程。二者可以关联，但不能相互代替。用今日有用解释最初必然如此，容易把结果投射到起点；只列起点事件而不说明持续条件，又可能无法解释结构为何保留。

SDE 发生分析的纪律是让方向性获得具体来源。所谓趋向某个结果，需要指出哪些差异被保留、哪些反馈造成下一次操作改变、哪些特征纠缠使结果持续。若没有这些环节，不能靠给系统一个拟人化愿望来完成论证。若确有这些环节，也不必因为出现方向性就取消研究，把它视为不可讨论的神秘力量。

因此，本书不会把亚里士多德的目的论称作三原理的原型。相似的词语不构成相同理论。两维张力推动第三维改变并回检，是本书用于研究的具体要求；它是否在某段历史材料中获得支持，要由可定位证据说明，不能借古代大师的权威反向保证。

7.7 从三方程到六条完整的研究进路

在一个“材料加热后状态改变”的现代教学例中，可以把可重复辨认的状态规定记为 S，把准备、接触、等待、比较与复验的序列记为 D，把材料响应、器具标记与分类标准之间的构成联系记为 E。S=F(D,E)要求说明状态显露怎样成立；D=G(S,E)要求说明新显露怎样改变下一轮比较；E=H(S,D)要求说明这些比较怎样改变特征之间原有的连接。三式在这里不是已标定的物理方程。

在本章最有解释力的进路，是 S→E→D：先从“同一材料变热”这个显露出发，检查材料身份、冷热分类和接触响应之间的联系，再安排能够区分“材料改变”与“读数条件改变”的操作。比如先固定比较材料与接触次序，只改变等待时间；若分类边界随等待时间系统移动，原先被当作稳定性质的判断就需要重新限定。这不是复原亚里士多德做过的实验，而是用现代教学安排说明他的性质问题还能怎样被继续追问。

另一条 E→D→S 从既定自然分类进入比较：先明确分类预计哪些结果，再让同类材料接受不同操作，观察原来同一名称是否仍指向足够一致的显露。若结果分裂，不能只在词典里给名称增加例外，必须判断是操作没有可比性，还是分类本身需要重构。六路径的其他入口作为反向核查，防止我们既以操作定义性质，又以该性质证明操作正确。这里用的是完整三维次序，不是从两维之间画一条箭头便宣布发生已经得到解释。

三原理在此表现为三个具体追问。S 与 D 发生张力，例如同样命名的状态在不同程序中不稳定，就推动修订 E，再用原程序回检。S 与 E 发生张力，例如既有连接解释不了反复出现的显露，就推动改造 D，寻找能够排除解释的比较。D 与 E 发生张力，例如操作与构成联系无法共同维持预期结果，就推动修订 S 的对象规定，再看新规定是否改善预测。任何一项改变都不是成功的保证。

7.8 一种世界的力量，也是一种世界的边界

亚里士多德式研究所提供的，不是几个可以从体系中拔下的错误答案。它把性质、变化、原因与位置组织成能够持续提出问题的世界。这样的组织具有生产力，因为同一套区分可以进入多个场景；它也具有封闭危险，因为对象一旦由体系定义，体系就可能把不合的显露排除为不纯粹或偶然。

一个公平的比较不能只计算后来保留了多少词语。应当考察新体系究竟解决了旧体系中的哪些具体困难，以及它为解决困难付出了什么选择代价。将运动缩小为可计算轨迹，能够创造强大的预测能力，也会把另外一些变化问题留给其他研究。扩大某个方向的自由，不等于一次性获得所有方向的完整解释。

本章因而保留两个判断。作为史实，原典展示了与现代动力学不同的自然解释结构；作为本书推论，这种结构参与生成了可识别、可分类、可连续追问的物理规定性。后一个判断并不由前一个自动证明，需要不断与其他哲学解释竞争。下一章转向原子、几何与平衡，考察另一种创造路线：当性质被追问为组合，当物被规定为可比较的量，世界的公共性又怎样发生。

本章参考资料

[7-1] Aristotle, Physics II.1、II.3、II.7，Hardie 与 Gaye 英译，https://classics.mit.edu/Aristotle/physics.2.ii.html。已读所引段，未校希腊文。全章对该网页的历史内容简述控制于两处短段，其余为独立分析。

[7-2] Aristotle, Physics III.1，https://classics.mit.edu/Aristotle/physics.3.iii.html。已读运动论开篇，正文仅作简述。

[7-3] Aristotle, On Generation and Corruption II.2–4，Joachim 英译，https://classics.mit.edu/Aristotle/gener_corr.2.ii.html。已读性质组合段，未将其当作现代化学。

[7-4] Aristotle, On the Heavens IV.1、IV.4–5，https://classics.mit.edu/Aristotle/heavens.4.iv.html。已读轻重与中心段，正文仅简述框架。

第八章 原子与几何
德谟克利特、欧几里得、阿基米德

8.1 把不同的东西放进同一种比较

一块木头、一块金属和一块石头，拿在手里显得不同。如果把它们挂在一根能够绕支点转动的杆上，问题便发生改变：材料的名称还在，但眼下需要辨别的是重量与悬挂位置能否互相补偿。再把同样的物件放进水中，重量没有因此成为无意义的概念，排开多少水却开始进入能否浮起的判断。新的比较没有消除原来的物，而是重新组织它能够以什么身份进入共同推理。

本章把原子论、几何与平衡放在一起，不是因为德谟克利特、欧几里得和阿基米德共享一套理论，更不是要把三人的工作连成未经证实的直接影响链。它们提出了互不相同而又能够相互照亮的问题：可感性质能否通过更基本的构成来解释？图形关系怎样取得可公开检查的稳定性？物体的某些行为又怎样进入这种稳定性？

这些问题都有一个容易被忽略的前提：要进行公共比较，先要形成能够被不同人重复使用的单位与对象。一个人说“很重”，另一个人未必取得相同判断；一个人说“两段线相等”，若没有共有的判定规则，争论也不会自动停止。公共性不是从个人感觉中删除个人之后剩下的空白，而是一套差异能够被组织、纠正和接续的能力。

8.2 原子不是缩小版的现代粒子

亚里士多德《形而上学》第一卷第四章转述留基伯与德谟克利特，以充实与虚空为原则，并把形状、次序和位置的差别用于说明事物差异。[8-1] 这是带有解释立场的古代见证，并不是德谟克利特完整著作的原稿。本章据此讨论一种构成思路，不据此复原所有原子论细节。

这一思路的重要推进，是不再把每一种可感差异都当作不可继续追问的终点。如果组成部分相同，排列不同仍可能出现不同结果；如果整体看似连续，构成解释仍可以提出不可见的区分。这给自然研究打开一个空间：眼前的显露不必与其构成方式长得一样。一个解释可以越过直接观看，但必须说明越过以后如何返回可讨论的结果。

古代原子不能被称作量子粒子的早期名称。两种概念的成立条件不同：后一种需要特定实验、数学与预测结构，不能由“东西很小”这个共同印象推出。若把后来的成功倒赠给古代学说，实际上取消了后来物理学创造新对象的工作，也抹去了古代思想自身真正面对的问题。

原子论在本书中的核心问题，因而不是“猜中了多少现代知识”，而是怎样重新安排性质的解释地位。设想两套相同的拼片，一套形成封闭环，一套形成直线。数量相同，不代表结构相同；单片的材料性质相同，也不保证整体的可操作性相同。这个现代教学例只说明构成与组织之间的区分，不证明古代原子像拼片，更不证明所有性质都能由几何拼接解释。

SDE 会继续追问：即便接受基本单元，基本单元的同一性又由什么支持？如果一种解释把整体的发生归给原子组合，却把原子的性质全部放在解释以外，那么它解决了一层问题，同时留下另一层边界。边界并不自动使解释无效，但它不应被伪装成无边界的完成。理论强度取决于它能否说明自己的构成规则与适用范围，而不只取决于基本对象有多小。

8.3 点与线为什么能被不同人共同使用

《几何原本》第一卷以点、线等定义展开，并在公设中规定从一点到另一点作直线；第一命题给出在已知线段上构造等边三角形的程序。[8-2] 这些文字为几何活动规定对象与可用操作。本章不把书页上的图形当作一幅被动复制的自然照片，也不把定义本身说成经验实验的结果。

在纸上画一个点，总有墨迹的大小；画一条线，总有笔画的宽度。几何对象却可以在推理中不以这一次笔画的粗细作为决定条件。这样一种剥离不是简单“看不见细节”，而是建立规则：哪些差异在本次推理中可以被忽略，哪些差异会使论证失效。若把两点连错，粗细无关的约定无法挽救它；若只是笔迹略宽，结论未必受到影响。

几何的公共性因而具有双重结构。图形帮助我们指认对象，论证使结果不只依赖某个人眼力好坏。观看和论证不是互相取消，而是在合适的范围内分工。一个三角形画得歪斜，仍可能支持某种有效证明；一个三角形画得漂亮，却不保证推理成立。读者必须学习如何从实际图形进入被规定的关系，再从推理返回图形检查有没有暗中使用未给出的条件。

这也说明“理念世界”不能简单等同于脑内影像。公共几何中的一个点通过名称、位置关系和允许的构造操作取得身份。另一个人能够接续，不需要复制前一个人的全部心理体验，但需要进入同一套约束。这里形成的是可以跨个体持续的活动结构，而不是某个孤立主体将私人印象搬到他人脑中。

在 SDE 解释下，定理的稳定显露依赖证明与构造的差异序列，也依赖对象定义和推理许可之间的特征纠缠。可是这个解释不把逻辑有效性降格为社会投票。一个共同体一

致接受某个证明，仍然可能漏掉条件；正因为对象与操作具有公共约束，后来的读者才可能指出错误。公共性不仅使人同意，也使不同意能够准确找到位置。

8.4 离散计数与连续量不能任意互换

原子式构成容易把人引向计数：有多少个，怎样排列。几何量却不总能用某个共同小单位有限次铺满。《几何原本》第五卷的比例定义通过各倍量之间的比较关系处理比例相等。[8-3] 它提醒我们，比例关系不必以先找到一个共同单位的整数个数为前提。

为了具体理解这个区别，可以考虑边长为一的正方形。在通常欧氏条件下，它的对角线长度满足平方等于二。若假设对角线与边的比是最简分数 p/q，就有 p 的平方等于二倍 q 的平方；由此 p 为偶数，写成二倍某整数后，又推出 q 为偶数，与最简假设矛盾。这是现代记号下的教学证明，不是对某个古代发现现场的复述。

这个例子说明，离散计数取得的清晰性，不能无条件推广为所有量的构成方式。可以不断取更细刻度，得到越来越精细的近似；但近似程序的延续，与存在一个最终共同刻度，不是同一个命题。把二者混同，会让测量的成功遮蔽数学对象的差别。反过来，知道不可公度也不意味着不能测量；测量需要的是满足用途的误差约束，而不必等同于有限步骤取得绝对精确比值。

本书在这里区分三种公共性：整数计数要求单位身份稳定，几何比较要求关系与构造稳定，经验测量要求误差与校准可说明。三者能够纠缠，却不能相互冒名。某个装置显示六位小数，不会自动证明其对象由百万分之一的离散块组成；某个数学证明具有精确结论，也不会自动保证现实器物满足全部前提。

8.5 杠杆：重量的意义进入距离

《论平面图形的平衡》第一卷从等重等距平衡等假定展开，命题六与七分别处理可公度和不可公度情形下重量与距离的反比例平衡。[8-4] 这里使用的是原典节录所能支持的内容；本章以下的代数式与算例均为现代重述。

设一根刚性杆绕固定支点转动，杆的自重可以忽略或其力矩已经补偿，两侧悬挂重物，重力方向平行，支点摩擦可忽略。令两侧重量为 W_1 与 W_2，支点到两条重力作用线的垂直距离为 l_1 与 l_2。静力矩平衡要求 $W_1 l_1 = W_2 l_2$。距离应取力臂，而不是不问方向地读取杆面上的刻度；若杆或力的方向改变，几何关系也要重新说明。

例如左侧重二十牛顿，力臂三十厘米，右侧重十牛顿，则右侧力臂为六十厘米时，两侧重力矩均为六牛顿米。这个结论没有要求两块物体具有相同材料，也没有要求手提

它们时具有相同感觉。它建立的是一种允许重量与位置相互补偿的比较关系。所谓杠杆“省力”，并不是凭空取消负载，而是改变操作中的配合。

若右侧力臂读成五十厘米，仍能看到杆暂时不动，不能立即认为方程被推翻，也不能立即把现象排除。应当检查支点摩擦是否足以维持暂时静止，杆重是否产生附加力矩，悬挂物是否接触别处。理论的使用包含一次回到构成条件的工作。但如果每次失配都随意假设一个看不见的补偿，就使理论失去检验性。合理修正必须带出新的可区分结果，例如改变负载后同一摩擦估计能否继续解释。

由此可以实际走通 S→D→E：先面对“杆保持平衡”这个显露，再改变悬挂位置与载荷的序列，最后辨认刚性、支撑和重力作用线之间哪些联系是必要条件。另一条 E→D→S 则先改变支点或杆的结构，再按计划调整载荷，观察原有平衡规则能否继续成立。六路径中的不同入口在这里不是装饰；它们使同一个静止外观不再垄断解释。

8.6 浮力：没有拿走材料，为什么称量改变

《浮体论》第一卷命题五把浮起固体的重量与相应浸入体积所排液体的重量联系起来；命题七讨论较重固体浸入液体后称量减轻的情形。[8-5] 这些关系把几何体积、重量与液体状态放进同一个问题。现代常见的符号表达，需要与原典的表述方式分开。

考虑静止、均匀密度的液体，在尺寸范围内重力加速度近似恒定，忽略表面张力及与容器接触的支撑。设固体完全浸没，顶部与底部水平，横截面积为 A，高度为 h。由静水压强随深度的变化，上下表面压强差为 ρgh；两侧水平作用互相抵消，于是向上的合力为 ρgAh，也就是 ρgV。这个现代推导展示的是所列条件下的模型关系，不冒称阿基米德曾使用这套符号。

更一般的形状需要对表面压力进行合成。在满足同样静水条件时，合力仍可与所排开液体的重量联系。这里“排开”不是一定要让水溢出容器，而是指物体占据了原本可由液体占据的体积。语言若不加澄清，读者可能误以为没有溢水就没有浮力，把一种演示安排误作现象成立的必要条件。

举一个数量例子：某固体体积为一千立方厘米、重量为十五牛顿，完全浸入密度取每立方米一千千克的水中，取 g 为每二次方秒九点八米，则浮力为九点八牛顿。若用绳子从上方静止悬挂它，且绳重等影响可忽略，拉力为五点二牛顿。材料并未因此消失；改变的是称量互动中绳子承担的份额。

这个例子尤其适合纠正“重量只是物体里面的一项存货”的直觉。物体所受重力与测力计读数应当区分。读数可以因为液体承担了一部分支撑而改变。属性纠缠分析所揭

示的不是重力随名称任意变化，而是某种“轻了”的显露必须在具体互动结构里解释。若一方面用水中读数定义物体全部重量，另一方面又拿它与空气中读数直接比较，就把不同构成条件误当成同一问题。

8.7 整体浮起并不保证姿态稳定

浮力与重力大小相等，只是静止条件的一部分。一个漂浮体还可能倾斜或翻转。若两种力的作用线不重合，便会产生转动趋势；姿态改变以后，浸入形状又会改变，浮力作用线也随之移动。用一条标量等式判断一切浮体行为，会遗漏几何与转动之间的相互构成。

可以设想两个总质量和总体积相同的密封盒，一个把附加重物固定在底部，一个固定在顶部。它们在给定液体中排开的体积可以相同，但遭受小扰动后的表现未必相同。这个思想实验不是宣告所有低重心设计必然稳定，而是让读者看到总量相同并不保证分布相同，分布不同可以改变回复或翻转的条件。进一步预测还要指定外形、重心和浸入几何。

这正是从离散拼合到几何量、再到整体平衡的连续问题：单元怎样分布，决定整体可以显露什么。属性纠缠不是“每个因素都重要”的泛泛提醒，而是指出具体的依赖关系。改变重量分布会改变重心；倾斜改变浸入体积的形状；二者共同改变力矩。只有把这些联系写清，才知道需要测什么，也才知道哪项变化可以忽略。

8.8 三方程在一个平衡台上获得实际工作

在杠杆或浮体实验中，S 可取在规定误差内重现的平衡关系，D 是改变位置、负载或浸入状态并重复观察的序列，E 是几何边界、作用线、支撑与测量规范的特征纠缠。S=F(D,E)要求给出平衡为何能够稳定显露；D=G(S,E)使偏差反过来安排新的比较；E=H(S,D)则要求依据比较修订哪些联系真正参与结果。这三式不取代力矩方程与静水关系，而是检查这些关系怎样成为有效知识。

三原理的使用可以直接落在失配上。若同一读数与不同操作之间存在张力，应修改对支撑或摩擦的 E 判断，再回到原操作检查；若平衡显露与既定结构解释之间冲突，应改造 D，安排能够区别接触支撑与流体支撑的测试；若操作与结构始终不能产生预期静止，则可能需要改变 S，把“稳定平衡”改为“暂时停留”或“持续振动”，再检验新对象规定是否有预测价值。

最后这一种改变尤其重要。研究不一定以拯救原来期待的显露为目标。发现某个状态根本不能稳定成立，也是一项结果。若所有偏差都被解释成尚未找到正确操作，理论

就永远不会面对失败。允许对象规定本身改变，才使三维之间的循环具有真正的开放性，而不是无限延迟裁判。

8.9 公共物理量是一项创造，也是一项责任

原子、几何和静力学展示三种不同的创造：越过直接显露去设想构成，规定可公开复查的对象关系，将物体行为接入能够计算的比较。它们之间有亲缘问题，却没有因此成为一条无缝的思想流水线。原子解释不自动包含几何证明，几何证明不自动成为经验规律，一次平衡实验也不能直接决定全部本体论。

本书称这些工作参与创造物理世界，意思是它们重组了物的身份、差异的资格与共同检验的方式。这个命题应面对一种竞争解释：也许这些工作只是逐渐发现原有关系。两者的争论不能只靠换动词结束。需要比较的是，发生分析是否揭示了那些关系取得意义、公共性与可操作性的具体条件，以及这些条件能否被其他解释充分说明。

从这个角度看，一条杠杆公式的伟大不在于把世界永远封闭在简短等式里，而在于使更多人能够用明确条件提出同一问题，又能够指出等式在哪些对象与操作中不够用。未来的动力学、连续介质理论与微积分会进一步扩展这项工作，但不应被提前塞入古代文字。真正连续的是不断追问对象怎样成立的责任，不是所有时代从一开始就拥有同一套答案。

本章参考资料

[8-1] Aristotle, Metaphysics I.4，W. D. Ross 英译，https://classics.mit.edu/Aristotle/metaphysics.1.i.html。已读原子论见证段；这是亚里士多德转述，不是德谟克利特原稿。

[8-2] Euclid, Elements I，定义 1、公设 1、命题 1，David Joyce 网页版本：https://mathcs.clarku.edu/~djoyce/elements/bookI/defI1.html；https://mathcs.clarku.edu/~djoyce/elements/bookI/post1.html；https://mathcs.clarku.edu/~djoyce/elements/bookI/propI1.html。已读页面原典条目，不把编辑评注归于欧几里得。

[8-3] Euclid, Elements V，定义 5—6，https://mathcs.clarku.edu/~djoyce/elements/bookV/defV5.html。已读比例定义；平方根证明为本书现代记号教学推导。

[8-4] Archimedes, On the Equilibrium of Planes I，假定及命题节录，University of St Andrews MacTutor：https://mathshistory.st-andrews.ac.uk/Extras/Archimedes_on_statics/。已读原典节录，未读完整古希腊版本。

[8-5] Archimède, Des corps flottants, Livre I，Adrien Legrand 法译，1891，命题 V、VII，https://fr.wikisource.org/wiki/Le_Traité_des_Corps_flottants_d’Archimède/Livre_I。已读法译相应命题及前提，未校希腊文；正文浮力算例和压强推导是现代重述。

第九章 光与声的实验传统
伊本·海赛姆及多文明实践

9.1 从可比较的量到能够质问感觉的装置

上一章的杠杆和浮体，让重量、距离和体积形成公共比较。本章转向另一种困难：如果知识依靠看见与听见，我们又怎样检查看见与听见？镜中的人似乎站在镜后，回声似乎来自远处另一位说话者。直接经验并非毫无秩序，但它所呈现的位置、来源与大小，不必与造成这种呈现的过程相同。研究感觉不能只要求人更加认真，还需要安排一种能够区分解释的互动。

先做一个思想实验。屋外有一盏灯，暗室墙上开一个小孔，孔后放一面白屏。移动灯，屏上的亮斑改变位置；遮住灯，亮斑消失。如果只让观察者闭眼，屏的变化不会因此停止，只是这位观察者不再取得相应视觉显露。这一差别帮助我们把光的传播条件、屏上的图像和人的看见分别提出，再研究它们怎样相互支持。

这里并没有一个完全退出过程的“纯观察者”。观察者需要识别亮斑、决定移动距离，并在下一次观看中辨认是否发生同类改变。但这也不意味着他能够任意规定结果。装置的意义恰在于把一部分判断转交给可重复操作：挡板放在何处，光是否到达；开口改变多少，像的清晰度怎样变化。新的视觉世界不是更多印象的简单累积，而是印象能够接受构成条件的检验。

本章中的多文明比较以材料为边界。中国文献、阿拉伯语光学传统、希腊罗马建筑论及印度乐器实践，在时间和任务上并不相同。并列它们不是举行“谁最早发明科学”的竞赛，更不依据相似操作就宣布存在传播路线。我们要辨认的是：不同实践怎样让光、像、声与共鸣从一次感觉，转化为可以继续追问的对象。

9.2 墨家之“景”：上下交换需要一个过程

《墨子·经说下》保存了光与影像的讨论，其中一段把上下位置的交错与成景过程相联系，另有关于镜与远近大小的条目。[9-1] 古字、篇章分合与解释存在研究问题；本章只据实际可读文字确认它讨论了相应现象，不把现代完整摄影装置的所有细节倒填进去。

上下为什么交换，是比“这里出现了一个奇怪影像”更深入的问题。若成像被当作整个物体一次性投下的复制品，倒立就容易成为额外的神秘性质；若逐点追踪可以通过

开口的路径，倒立则进入空间关系。重要的转变不只是取得一个答案，而是把整体图像拆成可以分别改变、又能重新合成的贡献。

在今天的教学重建中，可以把一根明亮竖杆放在孔前，分别遮住它的上半部与下半部，查看屏上哪部分消失。这里的“重建”是我们设计的区分操作，不是关于墨家实验现场的考古结论。它使两种说法承担不同后果：若屏像各部分依赖相应物点的路径，局部遮挡应造成有规则的局部缺失；若成像只是观察者对某个整体符号的联想，则这些遮挡关系不必如此稳定。

一旦这个差异能够反复重现，视觉对象就获得新的构成：杆不仅是一个整体，还是能够在操作中被分段选择的光学对象；孔不仅是墙上的缺口，还是筛选传播路径的条件；屏不仅承接一幅画，还是把空间对应显露出来的界面。它们的作用不能从各自孤立的外观全部读出。

9.3 伊本·海赛姆：几何与视觉为什么必须同时工作

伦敦大学出版社对《光学》校译本的说明指出，该书将光线的数学研究与眼的物理问题共同纳入七卷结构；其拉丁译传又参与后续光学知识传播。[9-2] UNESCO 所载专业介绍强调伊本·海赛姆以安排实验检验光传播等假设的方法。[9-3] 上述材料能够支持这一研究组合的历史定位；对具体实验的讨论仍需以原典校译为依据。

尽管证据边界有限，这一研究组合已经提出一个不可省略的问题：几何线条怎样有资格说明视觉？画一条线很容易，但如果线的位置由实验结果随意决定，几何就只是事后描摹；如果线条先给出明确限制，要求光源、障碍、眼与屏在某种排列下产生不同结果，它才成为能够受检验的模型。

同样，研究眼并不等于把光学全部还原为眼的内部。若只讨论眼的状态，而不说明光从哪里来、经过什么，便无法区分眼前放着暗物与亮物的差别；若只追踪外部路径，而把所有视觉判断视为自动产生，又会遗漏距离辨认、遮挡推断和适应条件。因此，一个更有力量的光学世界需要同时处理传播、接收与辨识，而不是在“物造成一切”与“心造成一切”之间选择。

这与 SDE 的对话必须保持准确。伊本·海赛姆不因研究眼、光与图像就成为 SDE 理论的先行版本。我们的任务是用发生分析检查具体构成：哪些显露被当作稳定对象，哪些操作允许区别解释，哪些特征纠缠让数学线条与视觉结果相互约束。历史人物提供的是待研究的工作，而不是为当代方法颁发资格的权威。

9.4 小孔像的大小：一个可以复算的模型

设物高为 H，物到孔的距离为 u，屏到孔的距离为 v。假定光在所讨论区域沿直线传播，孔足够小，物与屏近似平行，并忽略衍射与有限孔径模糊。由两个相似三角形得到像高的绝对值 h=Hv/u；上下倒置对应放大率的负号。这个式子是现代教学表达，不是前述古籍采用的原始记号。

例如物高六十厘米，物距三米，屏距五十厘米，则像高为十厘米。把物移到六米处，屏距不变，像高应降为五厘米；把屏移到孔后一米处，物仍在三米处，像高则变为二十厘米。只需统一长度单位，三项计算便可复核。模型不只解释已经出现的像，还提前规定改变装置后应出现什么。

这些比例关系也使模型可败。若在满足所列条件、已检查距离与倾斜的范围内，像高持续不随 v 成比例，便不能只说观察者没有正确理解。需要检查孔径、介质、屏位置或模型本身。若每次偏差都通过另选一个未经测量的“有效距离”消去，就已经失去独立预测。

孔越小是否永远越好？仅在几何模型中，缩小孔径可以减轻不同路径造成的展宽；实际装置的光量也会减少，足够小的孔又不能继续忽略波动效应。因此，“更小”同时改变可辨亮度与空间清晰度的条件。一个成功的局部封装可以在某个范围内可靠，却不能把自身遗漏的机制宣布为不存在。

这也解释为何本章不把光线等同于视觉 SDE 的全部。光线是模型中的组织方式；看见一个清晰倒像，还需要光源强度、屏的响应、观察距离与辨认条件共同成立。实验所创造的是这一套可重复显露的组织，而不是单独发明一条画在纸上的直线。

9.5 镜中的位置与观看者的身体

考虑一面理想平面镜，物体在镜前一米，观察者借反射光看到的虚像位于镜后相应一米的位置。这里说的是几何光学的虚像位置，不是有一件复制物体真的站到墙后。若把一张纸放在所指的镜后位置，通常不能像接住小孔实像那样接住这个虚像。两种“像”必须通过操作区分。

一个简单检验是让两位观察者从不同位置观看，分别记录使目标像与标记对齐的位置，再根据反射几何重建光线。能否让不同视线的反向延长相容，提供了对虚像位置的约束。这个过程也揭示观看者并非可随意删除的参数：身体移动改变进入眼中的光线，却不必使所重建的虚像位置任意移动。

于是，主体参与和公共结果可以同时成立。不同人沿不同路径看见，不等于他们只能拥有无法比较的私人世界；他们可以通过公开的移动、标记和几何规则取得相容判断。反过来，结果相容也不意味着参与过程从未存在。公共性的工作正是在保留差异的条件下，使差异能够转换与回检。

本章由此走通 S→D→E 的一条路径：先面对“像似乎在镜后”的显露，接着安排移动、遮挡和屏的试验，再检查反射路径、观察位置与位置判据的特征纠缠。换用 E→D→S，则可以先改变镜面形状或光路安排，再设计比较，预测像的大小与位置怎样改变。两条完整路径都要求回到显露，而不能止于重新解释一个词。

9.6 沈括的纸人：让听觉问题获得可见标记

《补笔谈》“乐律”记录把纸人置于琴弦上，拨动相应琴弦，观察纸人跳动；文字还讨论另一张琴上声音相应时弦会振动。[9-4] 这一记载的价值在于把原本主要靠听辨的响应，接到可见的运动标记上。本章不据此宣称作者已经具有现代振动方程，也不把所有传统乐律判断都追认为现代声学结论。

从发生角度看，纸人改变了什么？它使一个“听起来相应”的关系获得第二种检验入口。若所有弦都被桌面震动带动，纸人的跳动不能单独证明选择性共鸣；若只有与某种调弦条件相应的弦明显响应，解释就需要说明为何发生选择。问题因此从“有没有动”进一步变成“哪些条件下谁会动”。

纸人也不是无代价的观察工具。附加物会改变弦的负载和阻尼；放置位置不同，响应可能不同。标记过重时，它可能压低本来想要显示的振动；标记过轻时，又可能被空气流动扰动。让不可见变得可见，不是把透明窗口轻轻摆上去，而是引入新的互动。可靠装置需要认识并控制这种介入。

可以设计一种现代区分试验：保持拨弦位置和力度的可比性，改变另一根弦的张紧状态，记录响应是否在某个区间增强；再把两件乐器的机械支撑分开，检查响应是否仍存在。前一项区分选择性响应，后一项检查通过共同支架传播的竞争解释。两项都不能只凭一次成功演示完成，必须给失败结果留下记录位置。

9.7 剧场与乐器：声音从来不是孤立的震动

维特鲁威《建筑十书》第五卷把声音传播、反射与剧场听辨联系起来，并提出声学器皿等设计安排。[9-5] 这些文字确证古代建筑论处理了声音与空间的问题，却不能独自证明每一项建议都曾普遍实施，或达到作者期待的声学效果。文本方案与遗址实际成效需要分别核验。

对比一间空房与一间坐满听众的房间，读者容易理解声源相同不等于听见相同。传播路径和边界响应参与了到达耳边的时间结构。对语言而言，某些迟到的声响可能使音节相互覆盖；对音乐而言，较长的衰减又可能改变人对延续感的判断。这里同时存在物理量和用途评判，不能把“悦耳”直接当成某个唯一客观数值。

印度乐器实践提供另一种入口。大都会艺术博物馆对一件南印度维纳琴的馆藏说明，指出其空腔琴身、琴颈及琴桥与特有振弦发声方式的关联。[9-6] 该器物记录属于具体年代的见证，不能未经文献论证就前推数千年。我们借它讨论的是：一种被辨认为“这个乐器的声音”，由多项构成条件共同支持。

同一根弦换到不同琴体上，响度、衰减和音色可能改变；同一琴体换用不同拨奏位置，也可能改变激发方式。器物名称并不自动囊括全部声学规定。音色不是藏在木头内部等待取出的孤立标签，也不是演奏者主观宣布的效果。它在材料、几何、激发、传播和听辨中形成稳定差异，这种稳定还可以通过制作与训练继续得到强化。

9.8 理想弦的计算及其失效入口

用现代理想弦模型说明上述问题。设弦长 L，张力 T，单位长度质量 μ，两端固定，弦近似柔软、均匀，振幅足够小。基频为 f=√(T/μ)/(2L)。这个模型不是对任何古琴、维纳琴或古代作者的逐字复原，而是用于看清“同一个音高”需要哪些条件联合成立。

取 L 为零点六米，T 为七十二牛顿，μ 为每米零点零零二千克，则 T/μ 等于三万六千平方米每二次方秒，波速约每秒一百八十九点七米，基频约一百五十八点一赫兹。弦长减半而张力与线密度保持，基频加倍；张力增为四倍而其余条件不变，基频也加倍。两种操作能够得到相同基频，却不意味着所有振动细节相同。

这个例子尤其说明同一 S 不唯一决定 D 与 E。音高相同，可以来自不同长度、张力和材料组合；如果只按基频辨认，某些差异被暂时封装。要研究音色或耐久性，那些差异可能重新进入。三方程中的回写不是从一个读数魔术般恢复所有条件，而是在其他约束参与下重组研究，辨认哪些解释仍有竞争力。

模型的可败条件也应提前说清。若弦具有明显弯曲刚度，或振幅很大，或琴桥的边界不再近似固定，整数倍频结构可能出现偏离。不能把这种偏离一概称为调音不准；它可能是对象模型需要改变的证据。反过来，偶然一次频率读数异常也不能立即推翻模型，必须先检查测量、激发与边界条件是否保持。

9.9 光与声之间可以借用什么，不能借用什么

两者都能借助遮挡、路径、反射和接收条件来组织问题，因此实验者可以迁移某些比较习惯。但不能因为声音可见地引起振动，就直接宣布光必有完全相同的介质运动；也不能因为光线图便于画直路，就把所有声音传播限制为直线。类比能够提出问题，不承担最后证明。

一个强竞争解释会说，这些装置没有创造任何物理世界，只是把已有现象测得更清楚。对本书来说，回应必须落到具体改变：没有屏、遮挡、调弦和标记的组织，哪些对象规定尚不能成立？引入它们后，是否出现了可重复、可迁移并能排除旧解释的关系？如果只是把既有数据换一种好看的展示，便不足以支持创造命题；如果“像的位置”“选择性响应”“同一音高”因此获得新的构成与检验规则，发生分析就提出了额外解释内容。

这个论证仍可被反驳。若有人能够完整说明上述对象的意义与比较资格，而不需要这些构成过程，我们必须比较两种解释的覆盖与代价，不能把异议先定义成“不懂发生学”。科学史中材料支持的结果与哲学上的本体论判断，应当互相约束，但不能互相冒充。

9.10 张力怎样把实验推向下一次

把小孔像作为一个具体回路，S 是稳定的倒像比例，D 是移动光源、改变屏距及局部遮挡的序列，E 是路径选择、物屏几何与辨认条件的特征纠缠。S=F(D,E)要求说明比例显露怎样发生；D=G(S,E)要求让失配指导新的比较；E=H(S,D)要求实验结果能够修订原有构成解释，而不是只增添例外。

三原理在光声实践中各有实际入口。若同一操作不能保持像的比例，S 与 D 的张力推动检查 E，例如开孔或屏面是否改变；若共鸣显露与既定机械联系不合，S 与 E 的张力推动改变 D，把共同支撑分离再试；若现有操作与构成联系无法维持预期结果，D 与 E 的张力则推动修订 S，将笼统“相应”细分为不同机制的响应。任何第三维变化都必须返回先前冲突处回检。

不同文明留下的这些工作，使物理思想史不再只是公式之间的继承。一个小孔、一面镜、一张琴与一处剧场，都能成为新规定性发生的场所。但不能因为它们具有发生力量，就把每个工匠和作者的工作抹成同一种方法。共同问题在不同条件下取得不同入口，正是这部历史的多样性所在。

下一章中的天空观察将进一步放大这个问题：当眼睛通过新的仪器观看远方，当连续记录与几何模型彼此质问，原来稳定的天空秩序怎样变成可重建的运动序列？光学在这里不只是提供一件工具，它已经改变了什么能够被看见、怎样被信任，以及看见以后什么解释必须重新承担证明责任。

本章参考资料

[9-1] 《墨子·经说下》，光、景与鉴相关条目。https://zh.wikisource.org/wiki/墨子/經說下。已读文本，未完成版本校勘；不据此重建完整古代暗箱。

[9-2] University of London Press，The Optics of Ibn al-Haytham Books IV–V，校译本发布说明。https://uolpress.co.uk/book/the-optics-of-ibn-al-haytham-books-iv-v/。已读机构书介，未取得全部译正文。

[9-3] S. M. R. Ansari，Ibn al-Haytham’s scientific method，UNESCO Courier。https://courier.unesco.org/en/articles/ibn-al-haythams-scientific-method。仅据其机构介绍简述实验与假设检验，不采用其所有“先驱”判断。

[9-4] 沈括，《补笔谈》卷一“乐律”，纸人及琴弦相应条。https://zh.wikisource.org/wiki/夢溪筆談/補筆談。已读原文对应段。

[9-5] Vitruvius, De Architectura V.3、V.5、V.8，LacusCurtius 英译。https://penelope.uchicago.edu/Thayer/E/Roman/Texts/Vitruvius/5%2A.html。已读声音/剧场相关段，不将设计建议当作普遍实施证据。

[9-6] Metropolitan Museum of Art，Vina，馆藏对象页。https://www.metmuseum.org/art/collection/search/506151。已读器物说明，未主张其代表所有年代印度音乐。

第十章 哥白尼、开普勒与伽利略
从天空秩序到运动序列

10.1 天空并没有在一个早晨突然改变

上一章的光学实践提出了一个难题：看见不是判断的终点，必须追问什么安排使某个形象能够稳定出现。把这个问题带到天空，困难立刻增大。人在地面上移动屏幕与镜子，却不能随意把火星搬到另一条轨道；观察者所在的地球是否运动，本身也在争论之中。于是，天空的秩序不是直接摆在任何一只眼前的完整图画，而要经过连续记录、时间尺度、空间构图和仪器校验才能取得可比较的形式。

哥白尼、开普勒与伽利略的工作不能合写成一个英雄的连续胜利。三人所面对的对象、所容许的理想化和证据入口并不相同。把地球放入运动系统，改变行星路径的几何形状，把加速过程组织成距离与时间的序列，是三种相互关联却不可相互替代的变革。任何一种变革都没有自动提供另外两种变革所需的全部理由。

本章关心的不是给旧宇宙观宣布一个统一的死亡日期，而是考察新的规定性怎样取得解释权。当天体不再只是一个等级位置，而成为在明确时刻占据某个位置的物体；当运动不再仅以起点与终点分类，而需要说明中间每一段的变化；物理世界的构成已经发生深刻重组。这是本书的发生解释，不能冒充三位作者共同说过的历史宣言。

10.2 让观察地点进入运动关系

哥白尼在一五四三年出版的体系中把地球纳入行星运动，仍保留圆周组合的数学要求；这不是把后来开普勒的椭圆直接放进早期日心模型。[10-1] 对这一改变，可以先问一个比“太阳在哪里”更基础的问题：观察地点的运动，是否也参与生成所见的天体路径？

设想两名沿同一方向行进的旅人。前方旅人一直向东，后方较快的旅人从旁边经过。对于后者，前者可能暂时显得向后移。这里无需假定前方旅人真的掉头。可是把这个比方用于行星时，不能仅凭一句“地球超过火星”结束证明。视线方向取决于两个位置之差；两个绕行者的速度、距离及在轨道上的相位，都要进入计算。

用现代平面记号作一个教学模型。令地球的位置为向量 $r_{地}$，外行星的位置为 $r_{行}$，则从地球望去的视线向量为 $q=r_{行}-r_{地}$。这个减法的思想意义在于，显露的方向不属于行星单独拥有的特征，也不能只归给观察者。它由两个运动位置与共同空间规定的联系共同产生。地球自身的改变因此可以表现为远方物体的方向改变。

若两者作同向匀速圆周运动，示例周期分别取一年与两年，则平均角速度分别是每年三百六十度与一百八十度，日心相位差一年增加一百八十度，两年回到原值。同向再次会合的周期为二年。这只是人为选数的圆轨道模型，不是火星的实测周期；而且相位差的变化率也不是地面所见经度的变化率，后者还需要从 q 计算方向。

这个细节十分重要。错误的科普常把不同角度偷换为同一个角度，得到貌似清楚的“证明”。本书所说的特征纠缠，恰恰要求保留角度属于哪一个顶点、方向相对于哪一个参照、时间由什么尺度标定。缺失这些规定，再漂亮的运动图也只能帮助想象，不能承担检验。

10.3 坐标改写与世界改造之间

同一组相对位置可以在不同坐标中表达。把所有位置同时减去地球位置，地球便在新坐标里停在原点。这种数学变换本身既不能证明地球物理上绝对静止，也不能证明日心说已被取消。坐标是否方便，与何种运动关系能得到统一动力解释，是两个需要联系却不能混淆的问题。

因此，强有力的竞争解释可以这样提出：哥白尼所做的也许只是重新整理计算，天空本身未发生任何构成变化。这个看法能够解释为什么相近的观测结果可能被不同图式表达，也提醒我们不要把图上居中直接当成宇宙地位。但是，它还需要解释新的整理为什么改变了问题的结构：地球与其他行星开始承担同类运动问题，地上观察者不再享有不必说明自身状态的豁免。

本书的创造论述在此有一个明确限度。它不声称写出新书使地球首次开始绕行，而是说“地球作为行星”“观测作为相对运动的结果”等规定，进入了一个可共同推演和争辩的物理世界。假如这种规定变化并未改变推理、操作或证据要求，只是名词替换，那么创造论述在这个案例上便应撤回。它必须证明重组的实际内容，不能把一切改称为创造后便自认为已经解释历史。

10.4 椭圆不是把圆轻轻压扁

开普勒的工作使轨道形状与运动快慢不能再各自孤立。研究机构对《新天文学》的介绍强调，其火星研究包含长期反复构模；椭圆也不是把若干观测点连起来便自然显出的答案。[10-2] 在成熟的几何表述中，太阳占据椭圆的一个焦点，行星与太阳的连线在相等时间扫过相等面积。[10-3]

为了理解这一改变，可以把一条轨道画在纸上，再分别规定两种走法：一种每秒走过相等弧长，另一种每秒扫过相等面积。轨道完全相同，两种序列通常仍不相同。形状

不足以规定运动。反过来，只写下快慢比例而不说明距离从哪个点量起，也不足以规定可计算的轨迹。新的天体对象因而同时包含路径与时间规则。

作一个可复算的现代模型：椭圆半长轴取十个任意长度单位，离心率取零点二。近日点距离为十乘以零点八，等于八；远日点距离为十乘以一点二，等于十二。在两端，速度方向都与焦点连线垂直。若单位时间扫过的面积相同，便有八乘近日点速度，等于十二乘远日点速度。因此近日点速度是远日点速度的一点五倍。

这里用到的是小时间内扫过面积近似等于二分之一乘距离、横向速度与时间；取极限即可得到瞬时面积速率。这个现代微小量推导说明了公式的含义，并不意味着开普勒已经以现代微积分课本的语言工作。还要注意，离开两端后，速度不总与焦点连线垂直，不能在任意位置都把总速率直接替换为横向速率。

若有人提出“椭圆上匀速运动也一样好”，这一竞争模型可以被具体比较。在上述椭圆的两端，匀速模型给出相同速率；面积规则却给出一点五的比值。通过跨越不同位置的时间记录，可以考察哪一种更能保持拟合后的预测。若只检验轨道形状，两者不能被区分；若只用同一批数据调参再宣布胜出，也会高估证据强度。保留未参与拟合的时段，是现代重建中使模型承担额外风险的一种方式。

10.5 小偏差为什么有时具有大力量

观测与模型之间有偏差，不等于模型已经被推翻。偏差可能来自仪器、折射修正、时刻记录、计算舍入，或路径假定。它也可能来自这些因素的组合。真正困难的工作是把偏差变成有结构的问题：它是否随轨道位置规律变化，是否在不同观测中重复，是否大于已经查明的误差范围？

一个例子可以说明差别。若四次误差依次为正一、负一、正一、负一，而每次测量本来就可能误差两个单位，那么这些数字不足以迫使人改换轨道。但若某段轨道上多次偏差同为正十，另一段同为负十，而独立校验仍把仪器误差限制在两个单位内，就有理由怀疑模型遗漏了一项有组织的关系。这里的数字只是辨别逻辑的示例，不是对历史观测表的复制。

沿 S—E—D 路径，先把稳定而有方向性的失配确认为需要解释的显露 S，再追问既有几何与位置联系 E 是否允许这种失配，继而改变求解与观测安排 D，例如比较新的轨道形式，并把预测送回尚未使用的记录。整条路径不能停在“旧模型不够好”的感叹；它必须产生能够再次失败的新比较。

而且一次成功修订不能使其余维度免于检查。如果椭圆模型与某些数据仍有出入，不能为了维护几何美感便随意删掉它们。也不能反过来因为现代行星轨道并非永远严格的孤立椭圆，便说椭圆方法没有效力。模型的成立范围要由对象之间哪些联系被保留、哪些联系被忽略来说明。近似有条件，才有可扩展的知识。

10.6 望远镜带来的并非不需说明的真相

伽利略观察到金星具有不同照明相位。伽利略博物馆指出，这与把金星始终安排在地球与太阳之间的标准托勒密方案不合；但第谷的地心—日心体系同样能够容纳这些相位。[10-4] 所以，“看到金星相位”不能被写成一次观察就唯一证明地球绕日运动。

这项限制并不降低观察的重要性。一个结果可以有力排除某个模型，却不能排除所有竞争者。证据的力量在于它改变了候选解释之间的关系，不在于每次都必须把历史带到最终结论。若要求一次观察立即解决整个宇宙论，就会同时误解科学的强处与限度。

还有更贴近仪器的竞争解释：新出现的形状会不会来自镜片、眩光或辨认方式？解决它需要重复观察、比较不同器具、考察形态是否按照连续时序变化，并检查同一仪器面对不同对象时是否制造相同伪影。这样的比较不是在观察之后额外添加的礼仪，而参与构成什么可以算作天体的一项稳定特征。

从上一章延续而来的光学问题，因此进入了天文学核心。可见的形状不再是没有条件的 S，望远镜也不是随手放在世界外面的 D。成像、照明、视线以及观测时刻之间的 E，使连续形态具有可解释性。若同样的形态随镜筒旋转而旋转，便要检查仪器来源；若它随天体相对位置连续变化，则支持另一类构成说明。这里提出的是一般检验方案，不是声称伽利略逐项实施了这个现代清单。

10.7 把下降变成能够放慢的序列

天文学不能任意重新安排天体，地面运动却可以通过装置获得新的比较入口。《两门新科学》第三日叙述以斜槽中的青铜球研究下降，并用收集水的重量比较持续时间；书中把数学推演与自然运动是否符合推演分开提出。[10-5] 这比把整场变革压缩成斜塔上投球的戏剧更能说明问题。

关键不是球终于肯听从数字，而是下降过程可以被拆成能够分别比较的路段。直接从高处落下时，变化太快，中间过程不容易区分。斜面的作用在于重新安排过程，使时间差与路程差进入测量能力之内。不过，放慢不是无条件复制：滚动球还涉及转动与接触，不能把斜槽滚动的系数不经说明就等同于理想质点竖直下落。

用一个现代教学例子取沿程恒定加速度为每平方秒零点四米，初速度为零。则 s 等于二分之一 at 平方。第一秒末的总路程为零点二米，第二秒末为零点八米，第三秒末为一点八米。三段相继的一秒区间内，路程分别是零点二、零点六、一米，比为一比三比五。总路程比则为一比四比九。把这两个比例混为一谈，会毁掉本来清楚的证据。

这组数字并未测出自然界的真实加速度；它让读者检验一个假定的内部后果。要把后果转为经验命题，必须检查起动是否近于静止，计时零点是否一致，轨道是否足够均匀。若初速度为每秒零点一米，同样的加速度会给出零点三、一、二点一米的累计路程，原来的平方比例便不再成立，但这不等于恒加速度本身被否定。失败究竟属于哪个条件，必须逐一定位。

10.8 反例不是从数据表中随手挑出的数字

对于下降实验，强竞争模型不是一句“古人不懂科学”，而可以是恒速运动、随路程增长的速度规则，或具有系统计时偏差的替代拟合。恒速模型对等时段给出相等路程，因而与一比三比五的区间关系不同。若可靠测量在宽广范围内维持等路程，恒加速度解释就不能凭权威保留。

另一种困难更隐蔽：计时存在固定延迟。若实际运动已经开始而记录稍后才启动，短时间区间受影响尤其大，数据可能显示一种假曲率。为识别这一点，可以改变总路程、重复释放、采用独立计时方式，并比较偏差是否随着时间长度按同一模式缩小。可败条件必须涉及校准后的结果，否则任何模型都可以把反例推给未说明的误差。

这里走的是 D—S—E 路径。先通过斜面、释放与计时安排 D，使原本难分辨的变化形成序列 S；再分析哪些接触、转动、起始状态与尺度关系 E 使这一序列成立；随后返回装置，改变这些关系并检验结果。它与前述从轨道失配进入几何重建的路径不同，但同样不能删去第三维。

S=F(D,E)在此要求说明读数的生成条件，D=G(S,E)要求让已有结果指导新的区间与校准选择，E=H(S,D)则要求根据重复操作修订对过程联系的判断。这三个式子不是可直接代入数字的运动方程。上面的路程公式才是已规定对象与条件后的具体模型；两种层次各自承担任务，不能互相冒名。

10.9 新运动世界怎样保持开放

三原理在本章并不构成另一套天文学定律。若同一观测安排 D 不断产生与模型预期不合的 S，两维张力推动检查并改变 E，然后返回原记录核对；若某种显露 S 与既定构成关系 E 难以相容，则推动新的 D，例如延长观测时段或独立校准，再检验冲突是否消

失；若操作 D 与构成解释 E 不能共同维持原有对象规定，则推动 S 的重新区分，例如把总路程与逐段路程分开，再回到原操作检验。第三维的改变必须承担回检，才不是任意修辞。

这也解释为什么天空与地面问题能够开始接近，却没有在本章自动完成统一。开普勒的路径与时间关系没有单独给出一切力的来源；伽利略的加速序列也没有直接推导行星轨道。若把后来的引力体系提前放入他们的研究，会看不见那些尚未解决的问题，甚至把真正需要下一代创造的工作当成已经完成。

从历史的另一端看，旧问题也没有全部消失。相对位置如何获得绝对意义，数学上相容的不同模型怎样选择，局部装置中的规律怎样推广，仍会以新形式返回。成熟物理的力量不是不再遭遇这些问题，而是把它们变成可以分工、比较和继续推进的问题。

本书据此把这一时期理解为物理世界构成的一次转换：运动开始拥有可逐时刻追问的序列，观察地点进入被说明的关系，仪器读数进入证明责任，形状和时间规则被迫共同接受检验。它保留了自然独立施加约束这一事实，也强调约束必须通过明确规定才能成为公共知识。接下来的近代物理，将在这些新对象之间寻找更统一的联系，同时再次改变它们各自的身份。

本章参考资料

[10-1] Pisgah Astronomical Research Institute，“Three Models,” Kepler’s Discovery，含《新天文学》第一章节译及三体系说明。https://www.keplersdiscovery.com/threeModels.html。阅读对应机构页面；仅据以简述哥白尼体系与圆周组合。

[10-2] PARI，“Kepler’s New Astronomy”及“The Ellipse”。https://www.keplersdiscovery.com/keplersNewAstronomy.html；https://www.keplersdiscovery.com/theEllipse.html。阅读研究过程与椭圆主题页，未据此声称通读《新天文学》。

[10-3] PARI，“The Area-Time Principle”。https://www.keplersdiscovery.com/areaTimePrinciple.html。阅读面积与时间关系的说明；正文椭圆数值和极限推导为现代教学重述。

[10-4] Museo Galileo，“Phases of Venus”。https://catalogue.museogalileo.it/indepth/PhasesVenus.html。阅读金星相位与竞争体系说明。

[10-5] Galileo Galilei, Dialogues Concerning Two New Sciences, Third Day，Henry Crew、Alfonso de Salvio 英译，1914，均匀加速运动及斜槽、水计时相关段。https://en.wikisource.org/wiki/Dialogues_Concerning_Two_New_Sciences/Third_Day。阅读相应原典译文；现代数值、误差设计与 SDE 解释不归于伽利略原话。

第三编 微积分的创造：差异如何成为可计算的物理

第十一章 没有微积分之前
运动为什么难以连续地书写

11.1 连续运动，不会自动变成连续知识

一颗球从斜面上滚下，人看见它一直在动。看见连续运动，并不等于已经能够说明每一时刻怎样运动。我们可以记录起点与终点，可以比较整个过程用了多久，也可以在若干位置留下标记；但两个标记之间发生的变化，仍然提出新的问题。

这不是古人才遇到的困难。今天，即使用精密仪器取得大量离散记录，从有限资料建立连续关系也仍然需要模型。记录再密，仍要说明哪些记录属于同一运动、如何比较时间、怎样处理误差，以及为什么可以采用某一类函数。连续运动的日常显露，与连续关系的数学成立之间，并没有一条不经思考便能通过的道路。

第一编已经说明，一个变量不是裸数的容器。现在需要进一步追问：使变量能够变化，还需要什么？差异必须可以比较，比较必须保持某种同一性，局部变化还要能够进入更大的组织。微积分的重要性正在这里。它不仅让计算更快，而且使一大类局部变化与整体关系取得新的可操作形式。

《科学与微积分》把这一转变概括为差异的表征与整体的重构。[4，第 3—8 章]本编继承这一方向，但不将其写成“微积分之前没有科学”。有能力处理某些面积、体积、平衡或运动问题，与拥有一套可以广泛迁移的微分积分算法，是不同的历史成就。

11.2 阿基米德留下的不是一片空白

阿基米德《方法》的开头，向埃拉托色尼说明机械研究怎样帮助发现数学关系，同时明确区分这种发现途径与随后需要完成的几何证明。[10，致埃拉托色尼的开篇说明]这一原典材料足以否定一种粗糙叙述：古人在牛顿之前只会静止地看图，从不让力学和数学互相启发。

更有价值的是，原典把发生过程保留下来。一个结论怎样首先变得可见，与它怎样获得证明，并不是完全相同的工作。机械方式可以引导猜想，几何论证则承担另外的约束。这种区别与本书正在建立的层次十分接近：创造的入口、形式的正确性与共同接受的理由，可以相互支持，却不能直接画等号。

本书在这里作的是方法上的对读，而不是声称阿基米德已经使用 SDE 术语，也不是把他的证明偷偷翻译成现代极限语言后，宣布古代微积分与现代体系没有区别。历史的丰富性正在于：某些关系已经被深入研究，而符号、一般算法、问题范围与证明标准尚有不同。

因此，书写微积分的创造，应当既看见前史，也看见新组织的形成。没有前史，创造像无源奇迹；没有组织差别，创造又被说成旧知识换了名字。我们要追踪的，是零散而深刻的做法怎样进入能够反复生成新问题与新解答的共同系统。

11.3　三个问题，在同一处相遇

第一个问题是速度。知道一段时间走了多远，只能直接得到这段时间的平均变化率；要给某一时刻的速度一个确定含义，需要进一步处理时间间隔缩小的关系。

第二个问题是切线。曲线整体弯曲，而切线在一点给出一种局部方向。两点连成割线很容易，一点的方向怎样由附近关系确定，却需要说明。单凭把尺子贴在图上，无法替所有曲线给出一般算法。

第三个问题是面积或其他累积量。矩形面积可以直接相乘，弯曲边界围成的区域则需要比较、分割和控制剩余部分。将很多小量拼接成整体，必须说明拼接误差怎样受到限制，以及不同分割是否趋向同一个结果。

这三个问题并不是毫无关联的计算技巧。速度与切线都在追问局部变化，面积则追问局部贡献怎样组成整体。微积分把这些方向之间的联系发展成可操作的结构，使原来必须为不同问题反复设计的方法，能够获得更广的迁移能力。这样的统一不是取消具体问题，而是让一个问题的解法能够进入另一个问题。

以下用现代记号作三个推导。这些推导用于说明数学内容，不应被误认作古代或十七世纪作者原来的表达。历史原典与现代理解的差别，必须保持可见。

11.4　从平均速度到瞬时速度

假设在某个明确的模型中，位置由 $x(t)=at^2$ 给出，其中 a 是具有相应量纲的常量。时间从 t 增加到 t+h，平均速度为：

$$\frac{x(t+h)-x(t)}{h}=\frac{a(t+h)^2-at^2}{h}=2at+ah,\quad h\neq 0.$$

随着 h 趋向零，右端趋向 2at。我们于是将该模型在时刻 t 的瞬时速度定义为这个极限：

$$v(t)=\lim_{h\to 0}\left(\frac{x(t+h)-x(t)}{h}\right)=2at.$$

这里没有把零除以零当作普通除法。我们先对非零间隔形成差商，再研究它在间隔缩小时是否趋向确定值。瞬时速度并不是在一个没有时间厚度的照片中直接读出的东西，而是通过邻近关系获得的局部规定。

这个推导同时显示，局部并不等于孤立。x(t)的一个单独数值不能决定 v(t)；速度需要附近位置之间的组织。它也不是任意选择，因为一旦函数及参数含义确定，极限便受到严格约束。微分把“附近怎样变化”转成一个新的量，却没有许可研究者随意决定这个量。

作为物理模型，$x(t)=at^2$ 还需要经验根据。代数计算完全正确，并不自动保证某一颗真实小球满足该式。要检验它，必须说明位置记录、时间比较和适用范围。第一编的发生问题没有被数学消灭，而是以变量与模型条件的方式继续在场。

11.5　从一条割线到一个局部方向

把同一个函数看成平面曲线 $y=ax^2$。连接横坐标 x 与 x+h 处的两点，割线斜率同样是 2ax+ah。若极限存在，切线斜率就是 2ax。速度问题与切线问题使用了相同结构，却不意味着运动和几何已经成为同一种经验事件。

它们共享的是一种可迁移的数学组织：比较相邻输入造成的输出差异，再研究比值的局部极限。这个组织可以在不同物理与几何问题中复用，但每次复用都需要重新说明输入、输出和单位。数学迁移的力量与物理解释的责任，从一开始就应当一起看见。

切线还提供一种非常重要的近似。对于可微函数，小位移 h 所引起的变化，可以写成 f′(x)h 加上比 h 更小的余项。这使一条弯曲关系在局部获得可计算的线性形式。线性不是被宣布为世界本来如此，而是在一定邻域和误差要求下形成的近似。

因此，微分并不把曲线“切成无数小直线”以后宣布曲线消失。它保持局部与整体的区别：局部线性关系有适用范围，多个局部之间怎样连接还需要进一步研究。将局部近似推广到大范围而不检查误差，恰恰是微分使用中的一种典型越界。

11.6　从有限求和到一个整体量

考虑区间零到一上曲线 $y=x^2$ 之下的面积。这是一个纯数学例子。把区间等分为 n 段，取每段右端的函数值构成矩形，总面积为：

$$S_n = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{k}{n}\right)^2 = \frac{n(n+1)(2n+1)}{6n^3}.$$

于是 S_n 趋向三分之一。若取每段左端形成另一组矩形，其结果也趋向三分之一；两组估计之间的差距逐渐缩小。这种控制说明，整体值不是从某一幅看起来足够细的图里猜出来的，而是在分割继续细化时取得稳定。

这个例子有一个不能省略的地方：我们已经规定了区域、函数与分割。所谓“把小量加起来”，不是随意收集很多数。每一项必须属于同一个整体问题，组合方式必须与所求的量相容。若把面积元素、时间间隔和位移直接相加，符号再相似，也不会自动形成有意义的积分。

从 SDE 角度，本书把这一点解释为整体重建需要构成关系。这个解释不能反过来被说成积分定义已经证明 SDE 本体论。数学给出的是一组明确条件下的整体极限；发生学追问这些量及其组合何以进入物理研究。两种说明相接，却不互相冒充。

有限分割与无限极限也不能被当作两种现实仪器。实际计算总在有限步骤中进行，精确数学结果则使我们能够控制近似的方向和误差。理想化的价值不是要求世界执行无限次操作，而是让有限操作获得可以检查的标准。

11.7　位置表不能独自决定连续世界

假设我们在不同时间 t_1 到 t_n 测得位置，某个多项式 $p(t)$ 穿过所有记录点。是否已经唯一确定运动？没有。对任意常数 α，构造：

$$q_\alpha(t) = p(t) + \alpha\prod_{i=1}^{n}(t - t_i).$$

在每个记录时刻，乘积中的一项为零，所以 $q_\alpha(t_i)=p(t_i)$。但是，在记录点之间，不同 α 通常给出不同曲线；甚至在记录点处，它们的导数也可能不同。全部已记录位置相同，并不保证速度与加速度相同。

这个简单构造不是说实验永远无法认识运动，而是说明有限资料必须与其他约束联合。动力学、允许状态、平滑性、边界条件以及新的记录，都可能排除大量候选。科学的工作不是在数据之后随意涂一条线，而是建立能够承担更多独立约束的组织。

这也让“微分忠实表征差异”得到更严格的含义。忠实不是未经模型便直接抄写世界，而是明确交代记录、近似和构成关系，并让新结果能够纠正模型。[4，第 5、8 章；

5，第 21 章]如果模型是任意挑选，求导只会精确地显示这个选择的后果，不会自动证明选择正确。

这里正连接到第一编的创造判据。新的函数形式可以是重要创造，但它只有与操作对象及独立证据接上，才成为新的物理规定。数学的丰富性提供可能空间，物理研究负责使部分可能受到世界回应的限制。

11.8 间隔越小，测量就一定越好吗

在理想定义中，我们让 h 趋向零；在实际测量中，却不能因此直接得出“取样间隔越小越准确”。为了看清这一点，再构造一个误差分析。假定位置记录为真实模型值加误差，每次误差的绝对值不超过 ε；又假定所研究区间内二阶导数绝对值不超过 M。

用前向差商估计速度。由两次读数误差造成的影响，绝对值最多为 2ε/h。由有限间隔代替导数造成的截断误差，最多为 Mh/2。因此，对正的 h，有一个上界：

$$\left|\hat{v}_h(t)-x'(t)\right|\le\frac{2\varepsilon}{h}+\frac{M}{2}h.$$

这不是某台仪器的实测性能，而是上述假设下的数学估计。第一项随着 h 缩小而增大，第二项随着 h 缩小而减小。在这个简化上界中，平衡两者得到 h=2√(ε/M)，前提是 ε 与 M 均为正，且该间隔仍位于模型适用范围内。

这个计算说明，极限定义与有限测量承担不同工作。定义告诉我们所求量是什么，误差分析则告诉我们有限资料怎样有理由地逼近它。若忽略后者，越接近零的差分可能越受到读数误差支配；若忽略前者，又不知道有限差商应当逼近什么。

对本书的桥梁而言，这是一项积极结果：微积分不是脱离仪器的纯思辨，仪器也不是只要足够精密就不需要数学。二者在误差、尺度与操作中互相限制。新的物理量，正通过这种组织变得可以被可靠地使用，而不是在一个抽象极限符号出现时就自动完成全部发生。

11.9 差异需要单位、方向与可比较性

速度的量纲不同于位置，面积的量纲不同于长度。单位不是在最后给数字加一个装饰；它帮助辨认哪些运算有意义，哪些关系可能发生混淆。若时间单位改了，变化率的数值会相应改变；若坐标方向改了，速度符号也可能改变。这些改变可以被明确换算，不应被解释为物理规律失去公共性。

同样，微分所需要的邻近关系也不是由字母自己产生。我们必须知道在何种参数上趋近，哪些变化被允许，哪些状态被排除。对于一个受约束的系统，若只允许沿某一关系运动，就不能把所有变量都当成可以独立变化。这些限制不是对自由计算的妨碍，而是物理问题本身的一部分。

因此，微分之前已经有构成工作，微分之后仍有构成工作。前者使差异能够比较，后者检验局部关系能否被迁移和组合。数学并没有把 SDE 中的 D 全部装进一个 dx；它是在某种规范化之后，对 D 中可以数量化的一部分组织提供表达。

这项有限性不削弱数学。正因为它说明自己处理什么、不处理什么，数学的结论才足够强。把所有差异都叫微分，反而会失去微分之所以能够严格运作的条件。

11.10　临界的一步：把比较组织成可以继续的规则

设想两种研究能力。第一种能够准确回答一个已经画好的图形面积，第二种能够在图形的某个边界参数改变时，继续回答面积怎样变化。两种能力都有数学价值，后者却增加了新的生成范围：研究者不必每次从头建立一个彼此无关的问题。

用现代例子说明，令 $A(b)=\int_0^b x^2dx$，则 $A(b)=b^3/3$。一个固定的 b 只给出一个面积，整个函数却组织了一族面积；再求导得到 $A'(b)=b^2$，又把整体面积的变化连回边界处的局部值。我们不只是知道一个结果，还知道怎样在一组允许变化中继续产生结果。

这正是本书区分结果与生成组织的一个小型模型。孤立面积值可以是 S 的入口，一族参数化关系则使新的 D 能够被安排；区间、函数与边界之间的组织参与 E 的说明。不过，这些对应只是本书对具体数学工作的分析，不是宣称 A、b 与 x^2 分别天然等于 S、D、E。

这个例子还说明，统一不会消灭不同。不同 b 的面积并不相同，却可以被同一种关系承接；一个共同算法并不要求所有问题同数，只要求问题满足相应构成条件。公共数学由此不再只是大家记住同一个答案，而是能够共同实施一套有约束的变化。

它与物理学的连接也就更加具体。若某一累积量随时间或空间范围变化，研究者可以追问变化与局部贡献之间是否存在这种关系，而不是预先假设存在。模型一旦建立，新的测量便可以检验它。数学给出了可提出的问题形式，物理研究使这种形式进入具体发生。

所以，微积分的新能力不只是无穷小更小、求和更多，而是把过去容易分散的比较组织成能够继续运行的规则。这种规则既使研究者节省重复劳动，也改变了他能够走向

哪些尚未经历的情形。创造的价值，由此从一个漂亮结果扩展为持续生成新差异的能力。

11.11 微积分带来的，是新的往返能力

回顾本章三个问题，速度和切线把局部变化组织起来，面积把局部贡献连到整体。真正具有革命意义的，不只是各自求出几个新数，而是局部与整体之间开始获得可以反复往返的共同方法。下一章将通过牛顿与莱布尼茨的不同路线，考察这种能力怎样取得明确形式。

从本书的发生学立场看，这种往返能力参与创造新的物理世界。它使运动不只以起止或粗略快慢成立，也能够以连续变化率、轨道约束及可检验的局部响应成立；使整体不只是一幅完成的图，也能够由可说明的局部关系和边界组织起来。

但这里不能倒置论证。我们尚不能因为微积分能够表达运动，就说运动完全由微积分制造；也不能因为实际运动可以被某种公式描述，就说数学只是被动记录。需要研究的是，两者怎样共同形成新的物理规定，以及这些规定怎样通过操作与证据获得稳定。

物理学思想史的这一页，于是既不是空白突然被天才填满，也不是一个永远既定的世界换了一套方便符号。它是一种新的发生能力正在形成：局部差异变得可以精确追问，整体关系变得可以重新构造，而每一次构造又能够返回局部，提出此前尚未以这种方式成立的问题。

第十二章 牛顿与莱布尼茨：两条路径创造微积分

12.1 先把两份文本放到桌上

讨论牛顿与莱布尼茨，最容易把思想史写成一场争夺奖杯的比赛：谁先想出来，谁先发表，谁更值得成为封面人物。优先权当然属于需要史料研究的问题，但它不能替代另一个问题：两条研究路线究竟使什么新的数学能力成立？

本章选择两个具体入口。牛顿 1666 年 10 月的流数手稿，一方面讨论由量之间的关系寻找变化速率之间的关系，另一方面讨论从变化速率关系返回量的关系。[8，命题 7—8]莱布尼茨 1684 年的《求极大极小及切线的新方法》则公开展示一种微分算法，并将其用于包括折射路径在内的问题。[9]前者让“量怎样流变”与逆向重建并置，后者使微分规则成为能够跨问题操作的表达。

这里的比较不是把牛顿全部思想压成“动态”，把莱布尼茨全部思想压成“符号”。两人都处理丰富的数学问题，后来的体系也不能被还原成两种性格。本章只根据所选原典，追踪两个入口怎样通向局部与整体之间的新组织；下文的公式使用现代记号，明确属于本书的数学重述。

12.2 牛顿的问题：量的关系怎样规定变化

如果两个量受到同一个关系约束，它们就不能任意变化。牛顿手稿中的一个重要方向，是把量之间的方程转成变化速率之间的关系。[8]这个方向看似朴素，却使“变化”不再只是一种口头附加：它有了可以从约束中推出的组织。

以一个现代例子说明。设某段运动满足 $x^2+y^2=R^2$，R 为正常量。对时间求导，得到：

$$2x\dot{x}+2y\dot{y}=0.$$

这个式子不是额外发现了一条关于圆的神秘定律。它说明，当位置受圆周约束时，两个坐标的变化率必须相容。若 y 不为零，可以写成 $\dot{y}=-(x/y)\dot{x}$；若 $\dot{x}$不为零，也可以讨论局部斜率 $\dot{y}/\dot{x}$。除法所需要的条件，不能在符号顺手时省掉。

这个例子把几何与运动接在一起。位置不是静止的点名单，变化率也不是点之外的一股独立流动；它们通过约束共同取得含义。用本书的语言说，某些可比较差异被组织成局部规定，新的数学表达使这种规定能够反过来限制允许的运动。

但是，圆周方程及其导数还没有给出全部动力学。沿圆周缓慢移动、快速移动或者改变速度，都可以满足位置约束。要确定实际运动，还需要时间规律或作用条件。一个正确的微分关系，不应被误认为已经说明“为什么这个物体如此运动”。微分增加了约束，并没有自动消除其余问题。

12.3　逆问题：从变化怎样返回量

牛顿手稿还把逆向问题明确放在前台：给定运动速率或它们之间的关系，怎样恢复量的关系？原文也没有声称所有这类问题都能轻易解决。[8，命题 8]这项限制非常重要，因为它使创造的现场保留了真正的困难，而不是被后来的教科书光滑化。

看一个现代的简单例子。若 $\dot{x}=2at$，则积分得到 $x=at^2+C$。同一个变化率对应一族相差常数的位置函数。只知道局部变化，还不知道整体位于哪里；必须再给出某个位置条件，才可能确定 C。

这不是积分的失败，而是逆问题正确保存了信息边界。求导时，常数平移对局部变化率没有贡献；积分若不增加其他条件，就没有理由替我们把这个信息找回来。数学没有遗忘应当知道的东西，它是在提醒我们：原来的局部资料并不包含全部整体规定。

这个结构连接到第一编的属性问题。一个属性的局部显露，未必独自携带整个发生系统。恢复整体需要其他关系参加，而不是把局部结果反复计算，就能从其中榨出不存在的信息。积分的创造性不在于任意补上缺失，而在于把需要补充什么显题化，使初始、边界与相容关系进入明确研究。

因此，正向求变化与逆向重建不是两项彼此孤立的技术。它们共同构成一个有张力的组织：局部使整体变得可分辨，整体又限制局部能够怎样发生。微积分的力量，正在于使这种相互限制成为可以不断继续的计算。

12.4　莱布尼茨：关系怎样获得可迁移的书写

莱布尼茨的微分表达提供另一个入口：不只为某条曲线解决一个问题，而是组织可以反复使用的运算规则。[9]当一个复杂表达能够被分解、微分再重新组合时，符号就不只是结果的缩写，它也参与了下一步研究的组织。

下面以现代形式说明乘积法则。设 u 与 v 在所讨论点可微。对小变化 h，写出 u 的变化 Δu 和 v 的变化 Δv，乘积的变化满足：

$$\Delta(uv)=u\Delta v+v\Delta u+\Delta u\Delta v.$$

除以 h 后，在可微条件下令 h 趋向零，最后一项趋向零，于是得到(uv)′=u′v+uv′。这里的消去不是随意扔掉不喜欢的项，而依靠明确的极限阶次。用微分记号可以写成 d(uv)=u dv+v du。

这种书写的价值，是让关系可以组合。复杂函数不必每次从几何图形重新猜测切线，而可以通过明确规则处理。可是规则越方便，越需要保留它成立的条件：函数是否可微，某个分母是否为零，变量是否按原来意义独立。算法不会替我们自动检查每一次应用的物理合法性。

本书据此将符号看作一种发生能力。它把过去需要漫长重建的操作压缩成可承接规则，使新的研究者能够从已经稳定的结构起步。但符号也可以遮蔽操作含义。一个人熟练地写出 dx，并不保证他已经说明在当前实验中什么算作 x 的变化。

12.5　折射路径：一个小问题怎样打开大方法

莱布尼茨 1684 年文章中的折射讨论，以两点经界面上一点连接的路径为问题，通过极值条件组织推导。[9，原文第 467 页的译文节选]以下不用原文中含有排印差异的字母关系，而设置一个透明的现代模型，展示这一类思想的数学内容。

设点 A 位于界面上方，坐标为(0,a)，点 B 位于下方，坐标为(L,−b)，其中 a、b、L 均为正。路径在界面上的交点横坐标为 z。两段直线的长度分别为 $\sqrt{a^2+z^2}$ 和 $\sqrt{b^2+(L-z)^2}$。给两段长度赋予正权重 n_1、n_2，定义总量：

$$\Phi(z)=n_1\sqrt{a^2+z^2}+n_2\sqrt{b^2+(L-z)^2}.$$

若寻找使 Φ 最小的交点，先计算导数：

$$\Phi'(z)=\frac{n_1 z}{\sqrt{a^2+z^2}}-\frac{n_2(L-z)}{\sqrt{b^2+(L-z)^2}}.$$

在驻点处，两项相等。若 θ_1、θ_2 分别为两段路径与界面法线的夹角，就有 $n_1\sin\theta_1=n_2\sin\theta_2$。第二阶导数是两个正项之和，所以在这个模型内，驻点确实给出唯一极小点。

这个推导最有价值的地方，不在于光线被拟人化为会选择“最省力”的路线。它显示的是：一个整体路径问题可以通过局部变化条件获得约束。改变交点，使两段长度一起改变；最小条件不是只在某一段内部决定，而是两段关系共同限制的结果。

把 n_1、n_2 解释为适当的光学量，使 Φ 进入光程模型，还需要几何光学的物理条件。数学证明只告诉我们该模型的极小路径满足什么关系，不能单独证明现实中的光在一切尺度、一切介质中都服从这一模型。光学解释与数学推导应当彼此支持，却不能彼此替代。

12.6 局部条件背后的整体选择

折射例子还有一个容易被忽略的层次：微分之前，Φ 本身已经被构造出来。为什么选择两段长度的加权和？为什么允许路径在界面处改变方向？为什么固定端点？这些不是求导以后才附加的说明，而是求导所处理的问题结构。

所以，新方法不能被描述成“世界先给一个现成函数，数学只负责算”。在实际研究中，选择变量、建立总量、规定可变路径与固定条件，都可能是创造的一部分。微分把这些选择的后果显露得更清楚，也可能揭示它们不能同时成立。

但是，这也不意味着研究者只要构造一个函数，就已经创造了相应的物理世界。选择必须进入独立的经验约束。若另一个总量同样能拟合旧资料，却对新的路径条件作出不同预告，研究就需要设计可以区分二者的检验。创造的自由与回应的限制，恰恰在这里相遇。

从属性纠缠的角度，本书将这段工作解释为对象、操作与判断的共同规定。数学变量的客体侧重、允许变化的互动侧重、问题选择的主体侧重，不是三件先在物的简单相加。它们在同一研究中相互改变，而每一侧重内部仍有自己的 S、D、E。

这项解释属于本书的 SDE 重构，不是莱布尼茨文章中的原话。保留这条界线，不会减少对历史的理解；相反，它让读者可以分别判断原典做了什么，以及我们的发生学解释是否确实增加了问题的清晰度。

12.7 微积分基本定理：一座可以证明的桥

局部与整体的联系，并不只是一种哲学愿望。在一组明确条件下，微积分基本定理给出真正的数学桥梁。下面使用现代连续函数的形式陈述和证明，不将此证明冒充为十七世纪原典的逐句翻译。

设 f 在闭区间[a,b]上连续，定义 A(x)为从 a 到 x 的定积分。对内部点 x 和足够小的非零 h，有：

$$\frac{A(x+h)-A(x)}{h}=\frac{1}{h}\int_{x}^{x+h} f(s)\,ds.$$

右端是小区间上函数的平均值。由于 f 在 x 连续，这个平均值在 h 趋零时趋向 f(x)。因此 A′(x)=f(x)。整体累积所定义的函数，局部变化率恰好返回被累积的函数。

再设 F 是 f 的一个原函数，即 F′=f。因为(F−A)′=0，在该区间上 F−A 为常数。比较端点便得到：

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

这两段证明使一座具体桥梁成立：累积与变化率在相应条件下互相联系。它不是说任何整体都能由任意局部无条件恢复，也不是说每一项物理发生都必然可微。连续性、区间和运算定义，共同限定了结论的范围。

桥梁的严格性，恰恰来自它没有承诺一切。一个哲学解释若借基本定理声称“已经证明全部世界由 SDE 生成”，就越过了定理的变量域；一个读者若因定理具有条件而说它没有普遍价值，也误解了数学普遍性。它的普遍，在于所有满足条件的对象都受同一证明约束。

12.8　两个方向，为何不只是相反的按钮

微分与积分之间存在深刻联系，却不能简单想成按下一个按钮再按另一个按钮，一切都毫无损失地回到原处。求导可能失去常数信息，整体重建可能需要初值与边界；局部形式还可能遇到定义域和整体相容问题。这些内容将在第 14 章展开。

在物理研究中，往返更加丰富。一组实验差异可以促成微分关系，积分关系可以组织整体运动，再从整体模型推出新的差异安排。第三步不一定是对某个函数再作一次字面求导，也可以是推导新的可检验后果。《科学是什么？》用“微分—积分—再微分”讨论这种科学循环。[5，第 21 章]

因此，需要同时保留数学操作与方法论转用的区别。字面微分有精确定义，方法论上的再微分则指向新差异的产生和检验。若二者混为一谈，科学循环会变成术语游戏；若完全隔绝，又会错过数学怎样参与研究组织的实际作用。

本书所要建立的桥梁，正位于这个交界处。微积分把局部与整体的某些关系做成严格操作；SDE 分析追踪这些操作怎样进入对象规定、实验安排与新问题发生。二者共同工作，却不争夺一个含糊的“唯一解释全部世界”的位置。

12.9　参数化改变，不等于换了一个运动

两种表达如何贯通，可以用一个现代计算看清。假定一条曲线写成 x=x(t)、y=y(t)，又用另一个单调、可微的参数 τ 描述同一曲线，且 t=t(τ)。在导数存在且 dt/dτ 不为零时，有 dx/dτ=(dx/dt)(dt/dτ)，y 也有相应关系。因此，在 dx/dt 不为零的地方，两种参数给出的 dy/dx 相同。

这个结果同时保留了不同与共同：各量对参数的变化率可以不同，局部曲线关系却按明确方式保持。它不是说任何时间参数在物理上都完全等价。若参数代表实际时钟，怎样重新参数化、怎样保留动力学含义，还需要说明。数学上的曲线重描，与物理时间的测量不能直接混为一件事。

从流数到微分关系的现代对读，正需要这样的细致连接。我们不能因为表达不同就宣布两套数学毫不相通，也不能因为存在转换就把其问题入口和操作组织完全抹掉。转换本身需要条件，而条件可能在某些点失效。例如 dx/dt 为零时，用 x 作为局部参数未必合适；曲线依然存在，失效的是一种表达安排。

本书据此提出一个发生学判断：公共结构的形成，不必靠消除所有路径差异；它也可以通过建立路径之间的合法转换实现。转换使差异被保留为可计算内容，而不是被视为必须压平的错误。这一判断将贯穿后面的力学、场论与相对论讨论，但每个领域的具体转换必须重新证明。

另一项变化发生在推理劳动上。若一个问题的不同表达能够被翻译到共同规则，研究者便可以借用另一条路径形成的工具。一个几何问题可以接受运动参数的启发，一个运动问题也可以借助曲线结构来研究。这里的创造不只在某个结果，而在新的相互进入方式。

不过，表达贯通没有免除现实检验。两套方程若只是形式等价，它们对同一实验的支持并不是两份完全独立证据；把同一资料以两种记号算对，也不能把置信程度任意翻倍。真正的独立支持要来自不同资料或不同操作约束，而不是同一推导的重复包装。

因此，微积分的共同语言同时带来两种责任：承认不同路径能够共享严格结构，也承认严格结构不能代替路径各自的发生证据。把这两种责任放在一起，才能既看见牛顿与莱布尼茨之间的数学联系，又保留科学思想史不是一份符号对照表的深度。

12.10 两条路径的贯通，不取消它们的不同

牛顿选取的流变问题，与莱布尼茨所展示的微分规则，为同一大类联系提供不同入口。我们可以在现代数学中把许多表达互相转换，但这种转换不意味着历史发生过程完全相同。

一个共同结果可能由不同问题组织形成；不同的符号安排也可能改变方法传播、组合和使用的难易。第 342 卷将典范放在生成组织，而不只是最终显露上讨论。[3]借助这一区分，本书可以同时承认数学联系与发生差异，而不必在“两人完全一样”与“两人创造两个毫不相通的世界”之间选择。

但这种发生学解释也要避免过度。我们不能根据一份手稿的成文顺序，直接断定作者心中最先出现的念头；不能仅凭一个符号就推断全部思想来源；也不能把后来形成的标准体系完整投射回早期文本。历史解释需要手稿、发表物、通信与实际使用共同支持。

本章因此不对全部优先权争论作出裁决。它已经能从所选原典确认两种有内容的工作：使变化速率关系成为可研究对象，以及使微分规则成为可跨问题的组织。对本书主线而言，这比一次简单排名更接近微积分创造的实质。

12.11 创造的数学，怎样成为公共能力

一种方法只有作者自己能够使用，已经可能包含创新；但当它能够被其他人接续、修订并用于不同问题时，其公共生命才获得更广的展开。记号、规则、范例与证明在这里互相需要，任何一项都不能独自担保全部传播。

清楚的记号可以减少重复劳动，也可能让操作者忘记其含义；一般规则可以扩大迁移，也可能诱发超出条件的使用；范例能够打开理解，却不能代替一般证明；证明能够建立约束，却不自动提供物理解释。因此，数学体系的成熟不是一个单项能力不断放大，而是这些工作逐渐形成可接续的组织。

本书据此提出一个进一步判断：微积分对物理世界的创造，不只发生在第一条公式被写下的时刻，也发生在它形成可共同执行的研究能力之时。不同研究者能够根据同样条件展开局部差异、重建整体并互相发现错误，新的物理规定才获得持续发生的公共基础。

这不意味着共同体同意就创造真理。共同体也可能共享错误，标准化也可能固化误用。正因如此，公共能力必须包含返回材料、恢复条件与检查失败的通道。压缩与展开共同存在，微积分才不是一套只会给出答案的封闭仪式。

12.12 “微分发现，积分创造”应当怎样读

《科学与微积分》用鲜明表述强调微分对差异的忠实与积分对整体的创造。[4]本书继承其反对把科学缩成零散观察的方向，但在具体分析中把这句话展开，而不把它僵化为两种互斥分工。

选择新的局部变量、构造新的变化率，微分研究本身可以具有创造性；求出某个已定积分，也可能只是既有算法的直接执行。反过来，积分模型的提出、初始与边界关系的组织，能够产生真正的新整体。创造不由运算名称预先担保，而要看什么新的规定与能力实际成立。

因此，本书采用一种更严格的表述：微分使局部差异获得可计算的关系，积分使局部关系在适当条件下进入整体组织；两者是否构成物理创造，取决于它们怎样重建对象、操作与可检验约束。这里是对底本方法口号的展开性修订，不是把新表述冒充原文。

牛顿与莱布尼茨的思想价值，正在于他们的工作使这种双向组织获得强大的形式。从此，运动、曲线、累积与路径能够以新的方式互相进入。接下来的问题便更具体了：一个导数究竟使物理属性增加了什么规定，一次积分又需要哪些不能省略的整体条件？这两问，将把数学的创造推进到物的创造内部。

第十三章 微分：局部差异如何凝成新的物理规定性

13.1 速度不是贴在物体上的第二张标签

在日常语言中，我们说一辆车是红色的，也说它很快。两句话都像在给一个物体添加属性，但“快”的规定不能只由某一时刻的一幅轮廓决定。它涉及相对于什么参照、经过怎样的时间比较、在何种范围内讨论变化。微分的进入，使其中一类关系得到精确组织。

这并不是把速度从物体身上抢走，改判给时钟或观察者。它是改变属性的说明方式：物体身份、位置差异、时间比较与局部极限共同参与速度的成立。任何一项缺少必要规定，都可能使“同一个速度”变成含糊表达。

《科学是什么？》把微分放在对局部差异、梯度和敏感关系的研究中，又要求这些表达返回整体模型和新的检验。[5，第 21 章]本章沿这一线索，以现代数学例子展示微分如何参与物理规定。所有演算都是在明确模型条件下的说明，不把它们当作已经证明 SDE 本体论的实验。

13.2 导数与微分：变化率和局部线性关系

设 y=f(x)，函数在 x 处可微。导数 f′(x)给出局部变化率；微分则可理解为由这一变化率规定的线性映射，把输入增量 h 映到 f′(x)h。有限变化满足：

$$f(x+h)-f(x)=f'(x)h+r(h),\frac{r(h)}{h}\to 0.$$

余项 r(h)不能随意删除。它的含义是，当 h 足够小时，余项相对于 h 变得更小。对一个具体有限 h，近似误差仍然需要判断。这使微分既足够强，又不过度承诺：它提取局部主要关系，不把全部有限变化压成严格直线。

以 $f(x)=x^2$ 为例，有限变化为 $2xh+h^2$，微分为 2xh。若 x 固定而 h 缩小，二次项相对一次尺度退居次要；若 h 并不小，二次项可能不能忽略。在 x=0 处，微分为零，但有限变化仍有 h^2。这不是函数“没有变化”，而是一次线性层次没有捕捉到主要的二次变化。

这项区别具有本体论启发，却首先是一项数学事实。一个侧重下的显露为零，不等于一切发生都不存在；但是，我们不能仅凭这个形式相似，就把所有零读数都解释为被隐藏的高阶属性。是否需要更高阶关系，仍要由具体问题和证据决定。

微分因此不只是缩小观察窗口。它改变了我们能够提出的问题：在当前状态附近，什么变化占主要地位，哪些变化以更高阶方式进入，什么范围内这种区分有效？这种新的局部组织，可以成为新物理规定的组成。

13.3 运动参照改变，什么保持

考虑一种现代的简化运动描述。两个惯性坐标采用相同时间，彼此以恒定速度 U 相对平移，关系写为 x′=x−Ut。对时间求导，得到 v′=v−U；再次求导，得到 a′=a。这里讨论的是经典伽利略型变换，不是相对论变换。

这组三行计算给出一种清楚的稳定：位置与速度数值可以改变，加速度在这一变换下保持。由此可见，即使在经典范围内，公共物理也不等于所有观察位置都写同一个数。规律的共同性可以表现为不同描述之间的相容关系。

对于本书，这提供一项具体反例：不能把经典物理概括成“从来不承认观察位置与过程有作用”。第 379 卷已经保留这一限制。[2，第 6 章]我们可以批判某种解释把条件过度省略，却不能删掉经典理论本来拥有的变换结构，再把缺失当作其必然本质。

不过，形式变换也不自动证明任何两个实验都可互换。若两次实验实际改变了作用条件，或者并不满足恒定相对速度与共同时间的假设，就不能套用这组关系。数学的清楚在于它说明什么改变，以及在什么条件下什么不变。

发生学在这里要研究的，不是把每个坐标差别都叫新物，而是这些变换怎样使一个运动对象能够跨描述保持身份。新的物理世界可以包含新不变量，也可以包含新转换规则；二者都超出“一个数属于一个东西”的简单图景。

13.4 速率不变，为什么仍然有加速度

再看一个纯粹的运动学模型：平面位置为 r(t)=(Rcosωt，Rsinωt)，R 与 ω 为正常量。对时间求导，速度为 Rω(−sinωt，cosωt)，其大小恒为 Rω；再求导，加速度为 $-\omega^2 r(t)$。速度大小没有改变，速度向量的方向却持续改变。

这说明“差异”并不只是在一个标量数值上变大或变小。如果把速度全部压成速率，就会丢掉方向变化；如果据此认为没有加速度，错误并非发生在求导规则，而是发生在变量的构成层级。一个过于粗的变量不能承担更丰富的问题。

本书把这个例子看作微分参与属性重建的一个清楚入口。物体的运动规定从单一快慢扩展为具有方向的变化，新的数学对象使此前被压缩掉的差异能够进入计算。但是，这一运动学构造并没有给出物体为何沿圆周运动的实际原因。作用模型还要另行建立。

因此，数学形式的增强与物理解释的完成应当分开。一种更丰富的表示可能是必要步骤，却不保证全部因果问题已经解决。微分帮助我们发现原变量不够用，也提醒我们，补足运动学不等于补足动力学。

13.5 弹性响应：同一个“硬”可以有不同局部规定

现在设置一个明确的构成模型。设某种一维可回复响应，在当前研究范围内由外加载荷与位移的关系 $P(x)=kx+\beta x^3$ 描述，其中 $k>0$、$\beta\geqslant 0$，位移与作用方向已经约定。这里的 P 是维持位移所需的外加载荷；恢复作用的符号应与之区分。

若用 P/x 表示从零状态到当前状态的平均比例，得到 $k+\beta x^2$；若研究当前状态附近的微小加载，局部斜率为：

$$\frac{dP}{dx}=k+3\beta x^2.$$

这两种量一般不相同。一个是跨越有限区间的比值，一个是当前邻域的微分响应。若只给二者共同起名“硬度”，就可能把不同属性规定混在一起。它们不是谁更真实的问题，而是分别回答了什么问题。

这个例子也显示，局部属性可以随工作点改变。即使 k 和 β 在模型中固定，dP/dx 也依赖 x。这种变化并不证明材料规律不稳定；恰好相反，变化受到一个明确函数的约束。稳定的可以是关系，而不是每个工作点的数值。

在物理应用中，k 与 β 怎样取得、几何与制备是否影响它们、该模型是否容许卸载返回，都需要实验支持。本章不把假设多项式当作某类真实材料的普遍定律。它用于说明微分怎样把一个笼统属性拆成更精确、可比较的局部规定，而这种精确又怎样提出新的测量任务。

从三属性角度，客体侧重的响应数值、互动侧重的加载邻域与主体侧重的问题选择共同进入这一规定。三者不是相加为一个数，而是共同决定“我们正在测哪一种硬”。这种共同规定仍然要面对材料实际回应，不能由定义任意兑现。

13.6 偏导数与全导数：沿哪条路变化

如果响应同时依赖位移 x、温度 T 和内部状态 z，可以在一个可微模型中写成 $P=P(x,T,z)$。沿实际过程对时间求导，有：

$$\frac{dP}{dt}=\frac{\partial P}{\partial x}\dot{x}+\frac{\partial P}{\partial T}\dot{T}+\frac{\partial P}{\partial z}\dot{z}.$$

偏导数描述在模型允许的意义下固定其他变量时的局部变化；全导数则组合实际路径上各变量的变化。两者不能相互替代。若实际实验升温与加载同时发生，就不能把观测到的总变化率直接当作固定温度的偏导数。

这正是“固定其他条件”需要被认真研究的地方。固定并非一句口头承诺，它需要可执行的控制，并需要检查控制是否改变了其他关系。若固定某一变量会使系统离开原来讨论的状态范围，原偏导数可能不再对应所设想的实验。

如果变量还受到约束，例如 $T=\phi(x)$，并保持 z 不变，实际路径便不能在 x 与 T 上任意独立移动。沿约束的响应变化为 $\partial P/\partial x+(\partial P/\partial T)\phi'(x)$，而不是单独的 $\partial P/\partial x$。这个区别没有取消偏导数，它说明哪一种导数回答哪一种问题。

本书据此把实验中的控制理解为构成工作，而不是从外面抹掉干扰。它规定了允许的差异路径，也使某些局部属性取得清楚含义。微分的严格性要求这种规定，属性纠缠的研究则追踪它如何成立、能否保持、何时需要重组。

13.7　状态变量不是把历史直接塞进一个字母

在上式中加入 z，看起来便容纳了历史。但若 z 没有独立规定，它可能只是一个“未知因素”的新名称。把每次异常都解释为 z 不同，模型能够适配很多结果，却未必增加知识。

一个有内容的状态变量，应当能够通过操作、记录或其他约束取得可比较的含义。它要说明什么历史差别被保留为当前状态，以及什么新的变化会随之出现。如果两个被称为相同 z 的系统在相关条件下持续不同，就需要检查状态划分是否充分，而不是不断添加不可见参数保护公式。

这项工作与属性发生直接相连。某个属性在当前时刻怎样显露，可能不能只由外观或一个瞬时读数说明；但也不能从这一点推出“所有历史都必须完整保留”。研究需要寻找足以承担当前问题的状态组织，并说明丢掉哪些历史不会改变所声明的预测。

因此，状态压缩既是一种能力，也是一种需要验证的省略。它使复杂过程可以被有限变量处理，同时保留在迁移失败时重新展开的责任。第 379 卷对封装与省略条件的分析，可以在这一层与微分模型接通。[2，第 6、12 章]

13.8 交叉项为什么不能直接等于纠缠

考虑一个数学关系 z=xy，在 x、y 均为正的范围内，它具有显式乘积项；改用 u=logx、v=logy、w=logz，则关系变成 w=u+v。同一个可逆变换下的关系，在一种表示中出现乘积，在另一种表示中表现为加法。

这个构造说明，单凭是否出现某种交叉项，不能获得不依赖表示的本体论判据。交叉项可能对某个模型和量表非常重要，但它的意义必须连同变量定义、允许变换与检验对象说明。若变量换一种合法表达，形式特征就消失，我们就不能把这个形式特征未经限制地宣布为纠缠的全部本质。

同样，某个混合偏导数不为零，可以说明在当前坐标表达下局部敏感关系相互依赖；它不自动证明三号位共同构成，也不自动证明实际发生了双向回写。关于角色必要、条件成立与因果承接，仍需其他材料。

本书并不因此排斥交叉项研究。相反，明确其层级以后，交叉项能够成为有用工具：它可以比较某个模型是否遗漏联合效应，帮助设计改变一个条件时另一响应如何变化的检验。其力量来自具体含义，而不是被赋予一个无需证明的哲学身份。

这也是对研究对话中若干强表述的明确修订：统计交互、动力耦合、量子纠缠与 SDE 特征纠缠，不能因共享“不是简单相加”的外观就被视为同一定理。[7，第 5—17、30—37 页]它们可以对话，但必须各自保留定义与证明责任。

13.9 不光滑的地方，不是没有世界

函数 f(x)=|x|在零点连续，却没有通常的双侧导数。左侧差商趋向−1，右侧差商趋向 1。这个局部变化并不因为导数不存在就不成立；不存在的是某一种特定的光滑表达。

在物理建模中，接触切换、理想化冲击或分段操作可能使某个光滑模型不再合适。这里需要检查是现实过程本身的尺度问题，还是建模采用了尖锐理想化；可能需要分段处理、有限时间模型或其他数学形式。不能一见不可微就宣布科学失效，也不能为了保住导数而强行把所有尖点抹平。

发生学在此具有一个明确任务：辨认哪种差异超出了当前表达的资格。原来允许的邻近变化是否还存在？对象边界是否改变？某个新过程是否已经进入？这些问题可以推动数学重建，而不是替代数学重建。

同时，新的数学工具并不自动带来新的物理对象。有时它只是更准确地处理一个已知现象；有时它使此前不能组织的过程获得研究形式。应当返回第一编的层级判据，分别说明表示、操作与物理规定的改变。

13.10　从运动到作用：不能跳过的物理假设

如果我们从位置记录得到加速度，又把质量乘上加速度，是否已经独立测得了所有作用？这取决于所采用的动力学结构及其适用条件。加速度的定义是运动学关系；将它与作用联系，是另一个需要物理支持的规定。

一个常见的逆向难题由此出现：不同作用模型可能在一段有限运动中给出相近轨迹。仅凭这一段轨迹，未必能唯一辨认作用机制。研究可以通过改变初始条件、增加其他可观测量或使用独立的作用测量来区分。不能把“存在一个能够拟合的力”直接当作“已经找到了唯一真实原因”。

这与第 11 章的位置插值例子相呼应。资料可以非常准确，而解释仍不唯一。问题不是再增加几位小数就一定解决，而是需要新的差异安排，使不同模型作出不同承诺。微分在这里揭示限制，也参与创造更好的问题。

所以，微分不是自动的因果发生器。它能够建立局部变化关系，却不能仅靠运算决定什么因素实际造成变化。SDE 方法帮助追踪构成与回写，具体物理模型仍要承担动力学内容。两者的紧密联系，正因为分工清楚而更有力量。

13.11　一个局部实验怎样被数学重新设计

进一步构造一个有两个待辨参数的模型 $P(x)=kx+\beta x^3$。假定我们只在一个非零位移 x_1 测得载荷 P_1，则有 $k+\beta x_1^2=P_1/x_1$。这个方程容许许多参数组合，不能仅凭这一点分别确定 k 与 β。读数即使没有误差，参数仍可能不唯一。

若再选择不同大小的非零位移 x_2，得到第二个读数，则两条关系可以在 x_1^2 不等于 x_2^2 时区分参数。若第二次只取 $x_2=-x_1$，在这个奇函数模型中，它未必增加所需的独立约束。看似多做了一次实验，实际上可能沿同一种参数方向重复。

这个小计算显示，实验次数与有效差异不是同一个东西。新读数是否提供新信息，要看它在模型中怎样限制候选。微分与代数敏感关系可以帮助研究者选择更有区分力的工作点，从而组织真正新的 D，而不是盲目积累更多同类记录。

还可以换一种设计：在某个工作点独立测量局部斜率，与该点载荷一起使用。载荷给出 $k+\beta x^2$，斜率给出 $k+3\beta x^2$，在 x 不为零时即可分开两参数。但“独立测量斜率”不

能只是对同一条已拟合曲线求导；否则新方程仍来自同一假设，没有新增经验约束。它需要新的邻域操作及误差说明。

在这个意义上，数学不仅组织 S，也组织下一轮比较；新的比较又可能迫使我们撤回多项式模型。若新数据无法由同一组参数承接，首先需要判断是测量不确定度、状态范围还是函数形式出了问题。模型不能一边安排所有检验，一边免于检验。

把这段工作放到五部底本之间，本书看到一条具体连接：属性纠缠论追问变量与响应怎样共同成立，科学建构论追问符号、逻辑和数学怎样相互支持，微积分则使局部敏感关系能够反过来设计新的差异。这一连接是本书的联合阐释，而不是把底本的概念逐字替换成偏导数。

尤其要注意，在既定函数内推导出一个斜率，与建立这个函数作为新的物理构成，是两个层级。前者可以完全由规则完成，后者还需要对象、操作和独立证据的重组。二者都重要，但只有将它们分清，才能说明数学什么时候在执行已有规定，什么时候参与新的知识发生。

13.12 局部敏感性如何成为研究工具

在一个多变量模型中，导数矩阵可以组织各输入方向怎样改变输出。例如两个输出分别依赖两个操作参数，小变化可以近似写成 $\Delta y \approx J\Delta x$，其中 J 的各项是相应偏导数。这个表达不必被抬高为完整世界模型，已经能帮助研究者判断哪些方向容易被区分，哪些变化在记录中可能混在一起。

若两列局部响应几乎相同，仅观察这些输出就可能难以区分两种输入变化。改进研究不一定只是提高读数精度，也可能是增加一个新的响应量，或选择另一个工作点，使两种作用方向在资料中分开。数学由此直接参与实验组织。

但是，这种局部辨认能力受到模型和尺度约束。导数矩阵在一个点有用，不意味着全范围都相同；局部可辨识也不自动保证全局唯一；参数之间的相关，更不能未经证明就解释为本体纠缠。研究需要清楚声明：我们正在得到哪一种有限能力。

本书将这一点看作微分创造的积极一面。它不只是把已经完成的实验压成公式，也可以让研究者看见应当怎样改变实验，才能获得此前没有的区分。新差异的发生因此不是外在于数学，而可能由数学关系组织出来。

13.13 局部规律怎样改变“物”的规定

经过这些例子，微分对物理对象的作用已不再只是一句“方便计算”。它使速度区别于位置，使方向变化区别于速率变化，使局部响应区别于有限比值，使实际路径区别于固定其他变量的偏导，也使状态与历史的压缩承担新的检查。

还要保留一个积极判断：局部近似虽有限，却能构成可靠行动。知道误差怎样随邻域缩小，可以据此规定安全的计算范围；知道哪种敏感方向尚未辨认，可以据此改变实验。有限不是失败的委婉说法，而是某种能力已经明确成立、同时知道自己怎样继续发生的状态。

在本书的 SDE 解释中，这些改变可以参与新物的发生。物体不再只按外观成位，也可以按某一类局部响应、允许变化与可迁移关系获得规定。但是，每一项改变仍需看它是否形成新的操作与约束；并非每求出一个导数，都诞生一个独立物理世界。

《科学与微积分》的“忠实表征”，在这里可以被具体理解为对变量、局部范围、余项与操作含义的负责。[4，第 5 章]这种忠实不是取消创造，而是使创造不失去限制。我们可以选择新的变量与局部形式，却必须承担选择带来的后果。

下一章将进入另一半：局部关系怎样形成整体。微分把变化组织得如此精细，却不保证这些局部能够自动拼成唯一的运动、唯一的场或唯一的世界。初始、边界、整体相容与守恒，将使属性纠缠的另一些构成要求显现出来。

第十四章 积分
怎样从局部规则创造整体运动与守恒

14.1 把每一步写对，还没有得到整段旅程

假设我们知道一辆车在某些时刻的速度。要计算它在一段时间内的位移，不能把速度数字直接相加；需要知道各段持续多久、速度在段内怎样变化，以及所用方向是否一致。若速度是连续函数，积分便把这些局部贡献组织成整体位移。

这是一种强大的能力，但不应被说成“整体只是部分相加”。积分中的部分已经按照某种量的含义、某种参数与某种相容关系组织起来。没有这些规定，小量再多，也不自动属于同一个整体。数学中的求和与极限，是在明确结构内发生的运算。

《科学与微积分》强调积分对于整体重构的作用，《科学是什么？》又把整体模型接到新的差异预测与检验。[4，第 6—8 章；5，第 21 章]本章将这一思路落实为几组可以完整检查的计算，同时说明：整体重建需要哪些条件，什么情况下不唯一，什么情况下局部规则甚至不能组成所要求的整体。

14.2 积累量与状态量，先要分清

若 v(t)是有向速度，积分给出位移；若取速度大小的积分，则给出相应路径长度。两者可以不同：先前进后返回时，最终位移可能为零，走过的路却不为零。它们不是两个相互矛盾的测量，而是两个不同整体量。

这个最简单的区分已经说明，积分不是对所有过程采用同一种“总量”。我们必须知道被积函数是什么，正负号怎样规定，积分所形成的量回答什么问题。若把有向变化与总活动混同，正确计算也会得到错误解释。

对于能写为某个状态函数差值的量，整体结果可能只依赖端点；对于一般过程量，路径可能不能省略。是否能够省略路径，不应由“经典”或“现代”这样的时代标签决定，而应由具体数学与物理条件决定。

这也回应研究对话中的一个过强表述：经典物理不是任何路径都无关。[7，第 57、75 页]本书不沿用这一说法。积分理论本身就清楚地区分端点关系与路径关系。属性纠缠的研究应当解释何时路径被有根据地压缩，而不是先把整个经典理论写成没有过程的体系。

14.3　一个可以算到底的落体模型

考虑一个明确限定的模型：质量 m 为正常量，运动沿竖直方向，向下为正；重力加速度 g 在研究范围内取常量；阻力假设为与速度成正比的−bv，其中 b>0。忽略其他效应。在这些假设下，方程为：

$$m\dot{v} = mg - bv,\quad v(0) = v_0.$$

这不是关于所有落体的普遍断言，也不是对某段历史实验的复述。我们选择它，是因为可以同时看见局部规律、整体解、初始条件与模型范围。

令 τ=m/b，终端速度记为 v∞=mg/b。方程改写为 $\dot{v}+(1/\tau)v=g$。乘以积分因子 $\exp(t/\tau)$，左边成为乘积的导数：

$$\frac{d}{dt}\left(e^{t/\tau}v\right) = g\,e^{t/\tau}.$$

从零积分到 t，并代入初始条件，得到：

$$v(t) = v_\infty + \left(v_0 - v_\infty\right)e^{-t/\tau}.$$

局部变化率由当前速度决定；整体速度则同时保留初始值和时间尺度。即使不同初始速度最终趋向同一个 v∞，其有限时间过程仍然不同。稳定极限与历史差异可以共存，不必在二者之间选一个。

14.4　再积分一次，位置为何仍需要一个条件

知道速度还不能确定位置原点。设 $x(0)=x_0$，再积分得到：

$$x(t) = x_0 + v_\infty t + \tau\left(v_0 - v_\infty\right)\left(1 - e^{-t/\tau}\right).$$

这个解把原来分散的内容组织成整体运动：质量与阻力决定 τ，重力与阻力决定 v∞，初始速度决定暂态，初始位置决定整体平移。删去其中一项条件，通常会失去相应的唯一规定。

解写出来以后，仍应返回检查。对 x 求导，是否得到 v？把 v 代入方程，是否满足局部规律？t=0 时，位置与速度是否正确？这些检查确认数学相容，但不能替代对线性阻力假设的经验检验。

进一步让 b 趋向零，应当恢复无阻力模型的有限时间形式。对指数项展开并保留相应阶次，可得 v(t)趋向 v_0+gt，x(t)趋向 $x_0+v_0t+gt^2/2$。这里必须连同 v_∞ 与 τ 一起处理，不能分别把它们的发散当作整体解发散。参数极限需要检查完整表达，而不是盯住单个符号。

这组计算提供了一个清楚的发生学接口：局部关系经积分成为整体，不是因为所有条件突然消失，而是因为条件在整体表达中取得了明确位置。积分的组织力越强，我们越能看到初值、作用形式与尺度怎样共同规定运动。

14.5　整体解能预测什么，不能预测什么

上述模型可以预告速度向 v_∞ 接近，也可以预告改变 b 之后暂态时间尺度怎样变化。若这些参数由独立程序取得，就可以在新的时间段或新的初始条件下检验预测。若所有参数都是为了贴合同一条轨迹而事后调整，证明责任则不同。

一个模型拟合得好，并不自动说明其机制唯一。不同阻力函数在较小速度范围内可能给出相近曲线。为了区分，研究需要扩大或改变可合法比较的条件，使候选模型作出不同预告。这不是放弃积分，而是让积分产生的整体后果进入新的差异安排。

这正对应底本“再微分”的方法论意义：从整体模型返回新的可检验差异。[5，第 21 章]再微分在此不必字面等于再求一个导数，而是使模型能够生成独立问题。若只对自己假设的方程求解，再把解求导返回同一方程，完成的是内部一致性检查，不是新增经验支持。

因此，积分的创造性包含两种需要分开的工作：提出能够组织问题的整体结构，以及在既定结构中计算其后果。前者可能改变物理对象，后者可以是必需而严密的推演。只有看到二者的联动，才不会把科学创造全部归给灵感，也不会把所有计算都称为新世界诞生。

14.6　振子：整体可以稳定地不收敛为一点

再考虑 $m\ddot{x}+kx=0$，m 与 k 均为正常量，初始位置为 x_0，初始速度为 v_0。令 $\omega=\sqrt{k/m}$，其解为：

$$x(t)=x_0\cos\omega t+\frac{v_0}{\omega}\sin\omega t.$$

这个系统一般不会趋向一个固定位置。它持续振荡，但这种不趋于一点并不等于没有稳定结构。定义能量函数 $H=m\,\dot{x}^2/2+kx^2/2$，求导得到 $\dot{H}=\dot{x}(m\,\ddot{x}+kx)=0$，所以 H 保持。

为了避免符号混淆，这里的 H 是模型的能量函数，不是 SDE 三方程中表示特征纠缠关系的函数 H。相同字母在不同数学语境中不能自动承担同一种本体含义。

这个例子再次提醒我们，收敛、周期行为、分布稳定与守恒是不同概念。一个系统不趋向固定点，却可以有严格守恒；一个系统趋向稳定状态，也可能在此前经历丰富耦合。不能把任何“不定格”都称为同一种不收敛，更不能把它统一解释为某个号位复活的充分证据。

从物理思想的创造看，能量层面的稳定提供一种不同于位置不动的整体身份。物体一直运动，某个整体关系却保持。这使“稳定是什么”获得更丰富的规定，也使积分与微分之间的联系超出简单的静止与变化对立。

14.7　守恒来自条件，不来自积分符号的祝福

在振子模型中加入线性阻尼，方程成为 $m\,\ddot{x}+b\,\dot{x}+kx=0$，$b>0$。对同一个 H 求导，得到：

$$\dot{H} = -\,b\,\dot{x}^2 \le 0\,.$$

因此 $H(t)-H(0)=-\int_0^t b\,\dot{x}(s)^2 ds$。整体能量变化由过程积分给出，先前的守恒不再成立。改变的是模型中的作用条件，不是积分运算突然失去可靠性。

这个计算有助于澄清“积分创造守恒”的含义。积分可以组织平衡关系，某些条件下这个关系成为守恒；它本身不会保证任何量永远不变。若有源项、耗散或边界通量，整体关系必须把它们保留下来。

同样，若恢复力来自势函数 U，沿一维路径的作用功可以在适当条件下转成势能差。对第 13 章的模型，$U(x)=kx^2/2+\beta x^4/4$，其导数是维持位移所需的载荷 $P(x)$。恢复作用为 $-U'(x)$。这里的正负与角色必须保持，否则一个形式正确的积分会被赋予相反物理意义。

属性纠缠的研究可以追踪这些条件如何在具体系统中成立，而不能把“有整体”直接翻译成“有守恒”。整体可能以流入、流出和内部变化共同成立。更加完整的发生解释，应当能够容纳守恒与非守恒，而不是让其中一种成为预先规定的正确世界。

14.8 同样局部规则，为什么会有多个未来

考虑一个无量纲数学例子：$\dot{x}=2\sqrt{x}$，$x\geq 0$，初值 $x(0)=0$。对任何 $\tau\geq 0$，都可以构造一条解：在 $0\leq t\leq\tau$ 时令 $x(t)=0$，在 $t>\tau$ 时令 $x(t)=(t-\tau)^2$。

逐段检查即可发现，两段都满足方程，在连接时刻导数也衔接为零。于是，同一个初值对应许多不同解：它们可以停留任意一段时间，再开始增长。这个例子中的右端在零点附近不满足保证通常局部唯一性的那种利普希茨条件。

它不证明现实物体能够凭意志决定何时启动，也不证明经典动力学普遍没有预测力。它准确说明的是：方程、初值与唯一性之间存在条件关系。仅仅写成“变化率由当前状态决定”，还不足以在所有函数形式下保证唯一未来。

从本书的桥梁看，积分重建整体必须检查存在性、唯一性与稳定依赖，而不能认为“已经有微分方程”就等于“世界已经被唯一规定”。不同模型的数学性质，需要具体分析。若经验过程需要唯一预测，而模型没有相应条件，就应重新审查模型或补充有效规定。

这项反例还使“选择”一词获得边界。数学上出现多个解，不等于已经存在一个现实的选择机制。要解释某个实际结果，仍需说明哪些构成关系使其发生。不能用解的不唯一直接证明主体性，也不能把尚未规定的分支当作任意创造的许可证。

14.9 边界可以少给、给足，也可以彼此冲突

看一个更简单的边值问题：在零到一的区间内，要求 $y''=0$。一般解是 $y=Ax+B$。若给定两端值 $y(0)=a$、$y(1)=b$，就得到唯一的线性解 $y=a+(b-a)x$。

若只给两端导数都为零，那么 $A=0$，B 却可以任意，整体仍不唯一。若要求 $y'(0)=1$ 而 $y'(1)=2$，则没有满足全部要求的解，因为 y' 在这个方程下必须恒定。边界资料不是给得越多越好，它们需要彼此相容，也需要与方程相容。

这个例子将“整体条件”从抽象背景变成可以检查的内容。局部规律允许什么，边界选择什么，二者能否共同成立，都有明确答案。任意一边都不能单方面决定全部结果。

在 SDE 解释中，本书把这种相容研究视为整体构成的重要接口。但必须保留一条边界：数学边界条件只是 E 分析中可能涉及的一部分规定，不等于完整的特征纠缠。它没有自动包含对象的识别、装置的制备、条件的形成与后续回写。数学可以准确处理其中已形式化的关系，发生学继续追踪这些关系怎样进入研究。

14.10 局部看起来可以积，整体却未必如此

现在把问题从一条线移到平面。考虑去掉原点的平面上一个无量纲向量场：

$$F(x,y)=\left(-\frac{y}{x^2+y^2},\frac{x}{x^2+y^2}\right).$$

在每个不含原点的小邻域内，可以检查它的旋度为零。但是，沿单位圆作一周积分，参数化为 x=cosθ、y=sinθ，代入后得到：

$$\oint F\cdot dr=\int_0^{2\pi}1\,d\theta=2\pi.$$

若在整个去心平面上存在单值势函数，其梯度等于 F，沿闭合曲线的积分应为零；上式却不为零。所以，局部的条件并未在这个定义域上保证全局单值势。

这里没有发现一种新的实际力场。本例是数学构造，用来说明整体连接方式的重要性。局部微分检查、整体闭合路径与定义域拓扑，必须一起考虑。把局部关系逐点写全，并不等于已经解决全局组织。

这正是属性纠缠与微积分之间一个很有内容的桥梁：整体不能被理解为未经条件审查的局部相加。哪些局部能够共同组成单值整体，需要额外的相容结构。但是，这个数学结论仍不能直接证明所有物理属性都具有某种特定纠缠机制。它提供的是一项严格的限制，使本体论解释不得把整体组织说得过于轻易。

14.11 路径依赖不是旧科学的意外漏洞

上一节已经说明，闭合路径本身可以携带不能由端点抹掉的信息。对于有历史依赖的响应，问题也可能表现为：相同当前位移对应不同作用值。此时，若仍坚持用一个单值函数 P(x)描述全部过程，就可能混合不同状态。

一个合理处理是明确增加足以辨认状态的变量，或者将过程分支及转换条件写出来。随后，积分沿实际路径计算，而不是假定只要端点相同，所有过程贡献就自动相消。新的变量是否足够，要由独立过程检验；不能无限增添隐含状态，使模型永远免于失败。

本书在此反对两种极端。一种认为只要出现路径依赖，就证明此前全部物理都是错误的；另一种认为只要可以形式上增加一个变量，便不需要讨论对象规定如何改变。前

者抹掉了既有理论的丰富性，后者则可能忽视新增状态对物的身份与操作产生的实际改变。

发生学的工作，是研究这种改变究竟到哪一层。它可能只是同一模型中的参数更新，也可能需要新的状态空间、新的实验单位与新的公共承接方式。积分帮助我们把过程后果组织起来，属性纠缠论则追问这些过程何以共同构成当前的物。

14.12　记忆也可以被积分组织，而不必被积分抹掉

积分不只从当前局部规律重建轨道，也可以把历史如何进入当前状态写清。为此设置一个简化响应模型：$P=kx+z$，内部状态满足 $\dot{z}=-z/\tau+\eta\dot{x}$，其中 $\tau>0$，η 为正常量，所有参数量纲与相应变量一致。这个模型是本书用于逻辑分析的构造，并非某种材料的实测定律。

用积分因子求解，得到：

$$z(t)=e^{-t/\tau}z(0)+\eta\int_0^t e^{-(t-s)/\tau}\dot{x}(s)\,ds.$$

当前状态于是包含过去变化率的加权积分。距离当前越远的变化，在这个特定模型中权重越小；这来自所设方程的指数核，不应被说成所有记忆都天然如此。不同的实际机制可能需要不同模型。

现在比较两条连续加载路径。两者都从零位移开始，并在时刻 $t_0>0$ 达到相同终值 $X>0$；初始 z 均为零。第一条路径以恒定速度 $v=X/t_0$ 加载；第二条在前半段保持零位移，后半段以 $2v$ 加载。两条路径的终点相同，经历不同。

由上式计算，第一条路径在 t_0 的内部状态为 $\eta v\tau[1-\exp(-t_0/\tau)]$；第二条为 $2\eta v\tau[1-\exp(-t_0/(2\tau))]$。令 $r=\exp[-t_0/(2\tau)]$，二者差值可以化成：

$$z_{\text{后段加载}}(t_0)-z_{\text{均匀加载}}(t_0)=\eta v\tau(1-r)^2>0.$$

因此，同样最终位移对应不同当前载荷。差别不是读数错误，而是这个模型所预告的路径后果。若研究者只用 $P(x)$ 拟合，就会把同一 x 下的不同 P 视为散布；加入有明确动力规定的 z 以后，散布取得了可计算的组织。

但是，这并不证明凡有历史便永不稳定。若两条路径到达 X 后都保持不动，$\dot{x}$ 归零，z 便按指数衰减，P 趋向同一个 kX。路径差异可以在有限时间内真实存在，又在后

续条件下逐渐减弱。稳定与历史依赖在同一模型中并存，不能被拆成两种互不相容的本体立场。

这个例子为属性纠缠提供一项正面的数学接口。D 中的加载差异经具体承接形成当前状态，当前状态参与 S 的响应规定；E 的分析则追问这种承接关系怎样由材料、装置和模型条件获得支持。我们并不把 τ 直接等同于全部 E，也不把积分核宣布为三属性纠缠的唯一数学形式。

真正的研究将继续追问：这个内部状态能否被独立约束，指数衰减是否与新资料相容，改变路径是否出现预告的差别，哪些模型可以与它竞争。只要这些问题仍然开放，积分记忆模型就不是一个万能解释，而是一个能够提出有限承诺的研究原型。

它还让“属性是历史函数”取得可检验的含义。历史不再是一团无法处理的背景，而是通过某种有明确后果的组织进入当前。属性纠缠论不需要反对数学封装；它需要区分有根据的历史压缩与未经检查的历史抹除。积分可以服务前者，不能被一概描写成后者。

最后，若模型在不同装置中失败，可能需要修改核函数、状态维数或变量划分。哪种修改成立，必须返回资料，而不能以“真实纠缠太复杂”为由永远保护原式。一个能够认真失败的积分模型，比一个解释所有结果却无法失效的整体宣言，更接近本书所说的创造。

14.13　从一条解到一个物理世界，还差什么

至此，我们已经求出了运动，辨认了守恒，展示了不唯一与全局障碍。数学成果丰富，却不能直接宣布一个实际物理世界已经完全建立。仍需要变量与操作之间的对应，需要参数的独立来源，需要成立范围，也需要面对与模型不相容的资料。

在落体模型中，线性阻力是一项物理假设；在振子中，回复关系与质量取值受到模型限制；在去心平面例子中，向量场只是一项数学构造。这些不同地位必须保留。它们都能启发发生学，却不是同一种经验支持。

不同积分常数也不必一律代表不同物理世界。例如势函数 U 与 U 加一个常数，给出相同的导数，因此在只由其梯度决定作用的模型中，不改变相应作用关系。相反，振子初始振幅的改变，会改变当前模型中的运动与能量。数学解族里的自由参数，哪些只是表示选择，哪些代表可区分物理状态，需要通过操作含义来辨认。

这使非唯一性分析更精确了。我们既不能忽略缺失条件，也不能把所有数学差异都当作新增物理存在。先建立哪些表示彼此等价，再讨论哪些差异具有经验后果，才能从

解空间走到物理状态空间。一个模型在数学上有多种写法，可能在物理上对应同一关系；一个数值相同的记录，也可能背后有不同状态。两种方向都需要研究。

由此可见，本体论桥梁不是在数学差异上贴上存在标签，而是辨认差异怎样取得物理资格。这一资格来自变量、操作与约束的共同规定，不能仅由公式数量或解的数量决定。

《科学是什么？》要求符号、逻辑与数学共同进入实体科学。[5，第 19—22 章]本书据此把整体解视为一个关键环节，而不是全部科学。符号明确所指，逻辑组织承诺与反驳，数学形成局部与整体约束；独立实验则检验这些约束怎样在实际发生中被实现。

这也使“数学创造物理世界”获得准确含义。它不是数学在纸上宣布，物质便依令出现；也不是数学只在物理完成后提供计算。数学参与规定什么可以作为状态、什么变化可被比较、什么整体能够成立，而这些规定必须与实验、操作和对象身份共同发生。

14.14　积分的创造性，在于组织有约束的整体

积分最令人惊叹的地方，不是把无数小东西神奇地堆成大东西，而是使局部规律、初始条件、边界结构与全局相容形成可以检查的整体。它允许我们预测未逐点观察的过程，也允许我们清楚指出哪些条件不足以支持唯一结论。

对本书而言，这种能力直接扩展了“物”的规定。一个运动系统不只由此刻位置成位，还可以由允许的整体轨道、可保持的关系以及跨条件的变化方式获得身份。一个物理系统的整体，不再只是日常感觉给出的外壳，也可以是经严格重建而成立的运行组织。

但整体不是终点。新轨道会提出新测量，新守恒会要求检查新的边界，新失败会暴露原来的状态不足。积分所得的 S 可以重新组织 D，也可以要求重建 E；这项回写需要实际研究承接，而不能只在图上画一个返回箭头。

接下来，整体将不再只是一个物体随时间的运动。空间中的每一处都可能具有状态，邻近处的变化又互相进入。波、热与连续体将使“局部—整体”的问题获得新的规模，也让微积分参与物理世界创造的力量进一步展开。

第十五章 从常微分到偏微分
数学如何创造波、场与连续体

15.1　一根弦，不能只用一个位置说明

把一根绷紧的弦轻轻拨开。若只记录弦中点的位置，当然可以得到一条随时间变化的曲线；但是，中点的运动不足以说明整根弦。不同位置可以在同一时刻向不同方向移动，端点的固定方式也会限制整体形状。

这时，研究对象从一个随时间变化的量，扩展成依赖位置与时间的函数 $u(x,t)$。变量 x 标记弦上的位置，t 标记时间，u 表示相对于基准的横向位移。一个新的问题单位正在出现：不仅某个点怎样运动，还包括空间分布怎样共同变化。

欧拉关于弦振动的论文写于 1748 年，发表于 1750 年，属于这一问题发展的原始文献；傅里叶 1822 年的《热的解析理论》则将热传播与解析方法组织为系统研究。[12; 13]下文不把现代方程逐字归给某一位作者，而通过明确的模型推导，展示这些历史方向所打开的数学与物理能力。

15.2　偏导数增加的，不只是一个符号

对 $u(x,t)$，固定位置考察时间变化，得到 u_t；固定时刻考察空间变化，得到 u_x。两者对应不同的比较。前者记录同一位置怎样随时间运动，后者记录同一时刻邻近位置怎样不同。

这种区分依赖空间标记、时间比较和变量身份。若实验中追踪的是随材料移动的标记点，而数学中使用固定空间位置，两种变化率就不能未经转换视为相同。偏微分方程不是在普通导数上多加一个弯曲符号，而是要求我们把不同方向的差异组织清楚。

第一编讨论的同一性问题因此重新出现。什么叫同一位置，什么叫同一时刻，哪些读数构成同一空间分布，都需要规定。连续场不只是很多点的名字集合，而是这些点之间具有可比较、可接续关系的状态组织。

本书由此把场之物的创造与微积分联系起来：当空间分布及其邻近关系能够进入局部规律，物理对象便不再只能以一个整体位置来描述。不过，场的数学表达与任何具体场的物理存在仍须分开，写出 u 并不自动证明世界中有它。

15.3　从弦的一小段，推出波动方程

设弦的单位长度质量为 μ，基准张力为正常量 T_0，横向斜率足够小，暂时忽略张力随运动的变化和其他作用。取从 x 到 x+Δx 的一小段。在线性近似下，两端张力的横向分量之差，约为 T_0 乘以两端斜率之差。

这一小段的质量约为 μΔx，横向加速度为 u_{tt}。于是局部平衡给出 μΔx $u_{tt}\approx T_0[u_x(x+\Delta x,t)-u_x(x,t)]$。除以 Δx 并在模型允许的光滑条件下取极限，得到：

$$\mu u_{tt}=T_0u_{xx},c=\sqrt{T_0/\mu}.$$

这个方程把两种差异连起来：空间弯曲关系与时间加速度共同受约束。传播速度 c 不是由弦上某一个孤立点独自携带的数字，而是在模型的张力与质量分布条件中取得规定。

但不能把推导中的近似藏掉。大幅运动、非均匀材料、显著耗散或张力变化，都可能要求改变方程。微分的极限步骤只有在既定模型中合法，不能通过“把小段取得足够小”自动消除全部物理近似。

这个推导是本书对局部构成的一个工作展示。它使模型角色明确：质量、连接、张力、位移与空间比较不是互不相关的清单，而共同支持一种允许的变化关系。哪些关系进入方程，可以检查；哪些没有进入，也可以明确提出。

15.4　整体解需要一整个初始形状

一个质点的二阶运动方程通常需要位置和速度初值。弦的波动方程则需要初始形状 f(x)和初始速度分布 g(x)。在无限直线的理想模型中，对足够光滑的初始资料，可写出：

$$u(x,t)=\frac{f(x-ct)+f(x+ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct}g(s)\,ds.$$

令 t=0，直接得到 u(x,0)=f(x)；对时间求导后取 t=0，得到 $u_t(x,0)=g(x)$。进一步求导可以检查波动方程。这个整体形式清楚地显示：当前位置的后续状态，与一定范围内的初始分布相联系。

新的整体不是由一个初始数字全部决定，而是由空间函数和传播关系共同组织。研究对象因此从单点状态走向具有内部差异的连续分布。数学的扩展同时改变需要怎样制备、测量和验证系统。

在 SDE 重构中，本书把这称为新的整体规定能力。它不意味着每个位置拥有私人意志，而是意味着各位置的状态不能在忽略邻近关系的情况下任意组合。局部与整体之间已经具有明确约束，不再只是“很多东西在一起”的日常集合。

15.5 端点一变，同一根弦的整体就不同

若弦的长度为 L，两端固定，条件为 u(0,t)=u(L,t)=0。一个简单模式可以写成：

$$u_n(x,t)=A_n\sin\left(\frac{n\pi x}{L}\right)\cos(\omega_n t+\phi_n),\omega_n=\frac{n\pi c}{L}.$$

其中 n 为正整数。代入可检查它满足方程和端点条件。这里的允许频率由局部传播关系与整体边界共同规定，不能只向某一小段弦索取。

若改为一端固定、另一端在相应线性模型下满足零横向斜率，则允许的波数改成 (n+1/2)π/L，n 从零开始。材料和基准传播速度可以不变，整体模式却改变。边界不是最后加在成品波上的装饰，而参与规定什么样的整体振动能够成立。

这提供了“新物”问题的一种具体形状。同一材料支撑不同模式，不等于材料消失；新的模式具有可区分的空间与时间规定，也不等于只换了名字。研究应当说明这些模式能否被激发、识别和迁移，以及它们在什么意义上构成新的操作对象。

从视觉上，位移分布可以被绘出；在实际研究中，还可以通过不同传感与转换形成记录。若把记录转成声音或图像，新的显示能够帮助辨认模式，但转换本身不是独立物理验证。视觉、听觉与触觉 SDE 的跨通道承接，仍需分清发生、显示与证据三个层次。

15.6 热传播：另一种局部关系，另一种整体时间

再设一个一维均匀杆的热模型。令 Θ(x,t)为温度，C 为常量体积热容，κ 为常量热导率，横截面积 A 固定。采用局部热流 Q=−κAΘₓ 的构成假设。对一小段杆，内部能量变化率等于流入减流出，得到：

$$C\Theta_t=\kappa\Theta_{xx},\chi=\kappa/C.$$

这里必须区分两项内容：能量平衡与热流的构成关系。平衡本身不自动决定热流如何依赖温度分布，后者需要物理根据。将两项共同组织，才得到当前热方程。

这与弦方程表面相似，都包含空间二阶导数；时间结构却不同。弦方程是二阶时间变化，热方程是一阶时间变化。在对应模型中，同样的空间正弦模式可以表现为振荡，

也可以表现为指数衰减。数学形式不是彼此随意替换的标签，它们对整体时间作出不同承诺。

傅里叶的原著把热运动方程与不同几何区域中的传播分别展开。[13，第二章及后续传播问题]本书从这一原典组织获得的启发是：方程与区域不能割裂。一个共同局部形式，在不同边界与几何中需要不同整体解；整体构成不是局部规律的附注。

15.7　平均量保持，内部仍可以变化

在绝热端点的理想热模型中，Θ_x 在两端为零。对整个杆积分，有：

$$\frac{d}{dt}\left(C A \int_0^L \Theta \, dx\right) = \kappa A \left[\Theta_x\right]_0^L = 0 .$$

这里相对于固定温度基准的热能表达保持，但内部温度分布可以继续变化。一个整体量不变，并不意味着各处没有发生；局部状态不断调整，也不意味着所有公共常量都失效。

若两端维持同一个温度 Θ_0，令 $w=\Theta-\Theta_0$，一个正弦模式则可以写成 $A_n \sin(n\pi x/L)\exp[-\chi(n\pi/L)^2 t]$。与波方程中的振荡模式比较，局部空间形式相近，时间后果却明确不同。这种差别可以进入模型检验。

这个例子有助于纠正“整体越纠缠，读数必然越散”的简单推论。相互耦合的局部可以支持稳定总量，也可以支持逐渐平滑的分布。究竟发生哪一种结构，要由方程、边界与实际条件说明，而不是由一个关于纠缠厚薄的比喻决定。

属性纠缠论在这里的积极任务，是解释这些稳定如何在共同发生中成立，以及为什么它们不能被还原成每个局部独自拥有同样数字。稳定与差异不是敌对阵营，它们可以在同一整体中相互规定。

15.8　从一条曲线到函数空间

当状态是一整个函数时，变化的对象也扩展了。研究者不只比较两个数字，而比较两个形状、两组初始分布或两种边界组织。此时，“很小的变化”必须进一步说明：是在最大偏差上小，在平均意义上小，还是在某种包含导数的意义上小？不同选择可能对应不同的模型问题。

这不是数学家的多余挑剔。一个尖锐但很窄的变化，可能在平均量上很小，在局部最大幅度上却很大。若研究目标与比较尺度不一致，所谓“初始条件几乎相同”就可能掩盖关键差别。

欧拉的弦振动原典所处的研究方向，以及傅里叶对不同初始分布的解析处理，提示我们：函数不只是某条公式的名字，也成为整体状态的承载形式。[12；13]这里不把关于一般函数的后世定理全部归给某一位早期作者，而强调问题对象发生了扩展。

从 SDE 角度，这意味着 D 的比较范围与 E 的结构组织一起改变。局部函数值仍然重要，但函数之间的差异及其可接续关系也进入研究。新的整体不是比旧整体更大的容器，而是具有不同构成规则的研究对象。

15.9　变分：整体条件怎样返回局部方程

对理想弦模型，考虑一个作用量形式：

$$A[u]=\int\int\left(\frac{\mu}{2}u_t^2-\frac{T_0}{2}u_x^2\right)dx\,dt.$$

让 u 发生一个小的允许变化 u+εη，η 在所指定的时空边界上为零。对 ε 求导并在零处取值，得到一阶变化 $\int\int(\mu u_t\eta_t-T_0u_x\eta_x)dxdt$。分部积分并使用边界条件，变成：

$$\delta A=\int\int\left(-\mu u_{tt}+T_0u_{xx}\right)\eta\,dx\,dt.$$

若对所有允许的 η 一阶变化都为零，就得到波动方程。这展示了一种新的往返：整体泛函的驻定条件，可以返回局部方程；局部方程又使相应整体形式具有结构。

必须注意，驻定不自动等于最小，更不意味着弦会预先知道未来并选择道德上最好的路径。这里的“允许变化”是一种数学比较，不应与 SDE 的 D 维度直接等同。它可以为差异研究提供具体形式，却不包含全部发生过程。

同时，变分式为何适用于某个实际系统，不能仅由分部积分证明。其组成、边界和物理解释需要独立支持。数学往返越严密，越应该清楚区分形式等价与经验根据，而不是让漂亮的整体表达使这一区分消失。

15.10　数值发散，有时是算法在说话

现代计算把偏微分方程转成有限网格上的更新。为了说明数值结果与物理模型的区别，考虑热方程的一个显式离散形式：

$$U_j^{n+1}=U_j^n+r\left(U_{j+1}^n-2U_j^n+U_{j-1}^n\right),r=\chi\,\Delta t/(\Delta x)^2.$$

在无限或周期网格上，取一种交错空间模式 $U_j=(-1)^j$。一次更新后，它被乘以 $1-4r$。若 r 大于二分之一，这个因子的绝对值大于一，模式会在反复更新中放大。物理热模型中相应高频模式趋于衰减，数值程序却可以产生增长。

这个反例由更新式直接推导，不依赖某台计算机的实测故障。它说明，当模拟发散时，首先需要检查离散方法与步长条件，不能立即宣布发现新的物理不稳定，更不能直接把它当作某种本体论猜想的证据。

当零到二分之一之间的 r 使更新成为邻近值的非负加权平均时，还可以在适当边界条件下理解其保持数值范围的作用。这种数理约束帮助研究者区分模型、算法与实现三个层次。计算不是透明通道，它有自己的可检查构成。

因此，新的物理世界可以通过计算获得研究能力，但并非每一幅模拟图都代表一个实际物理存在。模型形式、数值近似、初始资料和独立验证必须共同说明。SDE 分析若忽略算法层，反而会把一个新的中介错误地当成透明观看。

15.11　同一个波动整体，也能携带一个守恒量

波的空间分布不断变化，仍可以在适当条件下具有稳定的整体能量。对前面的理想弦，定义：

$$E(t)=\int_0^L\left(\frac{\mu}{2}u_t^2+\frac{T_0}{2}u_x^2\right)dx.$$

对时间求导，使用波动方程，再对空间分部积分，可以得到能量的变化率只剩边界通量项。两端固定时，端点速度为零，这个项消失，所以整体能量保持。这里的花体能量记号，不是 SDE 的特征纠缠维度。

这个推导有三层共同承担结果。局部方程将空间曲率与加速度联系，积分把局部贡献组织为总量，边界条件决定是否有相应通量穿过系统。缺少其中一层，不能仅凭“波是整体”就宣布守恒。

若端点由外部驱动，边界项可以不为零，整体能量便能够改变。并不是原来的推导突然失去数学正确性，而是曾经使边界项消失的构成条件已经改变。对物理研究来说，识别系统边界及其作用，正是区分封闭近似与开放发生的关键工作。

这个结果也帮助我们分清三种不同的整体。空间积分形成一个总量；场函数描述各处状态的组织；边值问题规定允许的整体模式。三者相互联系，却不是同一个对象。若把它们都简称为“积分后的整体”，就会丢掉新物理系统内部已经形成的层级。

本书因此不把积分看成只能把差异抹平的平均机器。积分也可以准确保存通量、作用和能量之间的关系，使局部差异在整体中获得新的限制。失去差异往往不是积分本身的罪，而是研究者在选择变量、区域或解释时作了未经检查的省略。

对属性纠缠的讨论而言，最重要的不是宣布守恒或变动哪一种更先进，而是说明它们怎样共同出现。一个场可以在局部持续变化，在整体满足守恒；一个总量也可以因边界发生改变而变化。这样的分层，比“经典静止、现代流动”的对立更能够进入真实的物理数学。

15.12 波与场，怎样改变个体的边界

在质点模型中，我们容易用一个位置和速度表示一个物体；在连续体中，一个系统的状态可能是一组函数，某个稳定模式又可以在这些函数中成位。这种变化使物体边界不再只由日常可触外壳规定。

同一根弦可以支持不同模式，一种模式可以在不同具体装置中获得相应实现。由此可以形成新的问题：模式的身份依靠什么保持？更换材料、长度或端点之后，它是同一种模式的变换，还是另一个研究对象？这些问题不能只靠视觉形状相似回答。

本书据此把连续体与模式视为物理世界创造的重要环节。数学使空间差异与整体约束成为可计算对象，实验与操作则使相应规定能够被实现、比较和迁移。新物的发生不是从公式直接跳到存在，而是在这些关系的共同组织中获得内容。

同时，场不能因为在这里是物理术语，就被等同于 SDE 九宫中的所有场模态；数学函数也不能因为依赖空间，就自动成为完整 E。相同词语在不同理论层级中具有不同责任。建立桥梁，需要说明连接，而不是依靠名称相同省掉连接。

15.13 两编之间的桥梁终于落在具体工作上

第一编问：物怎样成位，属性怎样共同发生，什么足以支持新物的判断。第三编现在给出了一组具体数学工作：建立可比较变量，定义局部变化，求解整体关系，检查初值与边界，辨认全局相容，再安排独立检验。

这些工作不是 SDE 真实发生的固定时间流水线。研究可以从既成整体返回局部，也可以从异常差异重建变量，还可以从新的装置关系产生原型。六路径为这种多入口提供

组织；三方程要求三维互相说明；三原理则帮助定位为什么必须改变原来的一项规定。[2，第 3 章]

微积分在其中不是附属计算器。它能够改变哪些差异成为变量、哪些关系被认为相容、哪些整体可以被构造、哪些新结果能够被预告。但它也不是独立的造物者。变量的物理身份、模型的经验根据、装置的可执行性与判断的公共性，需要在同一研究中共同成立。

本书由此提出一个中心连接命题：微积分参与物理世界的创造，是通过重组局部差异与整体约束，进入物、物体、物质与物理的共同发生；这种参与既包含严格演算，也包含演算不能自行替代的对象构成与经验检验。

15.14　不是用数学装饰本体论，也不是用本体论取消数学

如果只把 S、D、E 写在波动方程旁边，理论桥梁并未建立。如果只说积分体现整体、微分体现差异，也仍然停留在词语相似。两编真正需要承担的，是在每一个模型中辨认构成选择、给出数学推导，并允许独立材料改变原来的解释。

弦的边界使模式得到规定，热的通量使总量得到平衡，变分条件使整体与局部相互转换，数值稳定性使计算结果获得可检查范围。这些具体关系，比一个无所不包的比喻更能展示属性纠缠论的研究能力。

在接下来的物理学思想史中，牛顿的运动物体、法拉第与麦克斯韦的场、相对论的时空规定和量子理论的状态组织，都将面对同样的责任：展示新的物理系统怎样发生，而不是仅仅宣布某种哲学已经胜利。

这两编也为后续历史保留了一条可实际使用的比较尺度：某位物理学家究竟重新规定了变量、局部律、整体条件还是跨系统的转换？新增内容怎样进入独立验证，又怎样改变已有物的身份？不同创造可以有不同答案，不必被迫排列成单调增加某一号位的历史。

物理学的创造，因而既有大胆的一面，也有受约束的一面。大胆使原来的对象边界可以重新打开，约束使新的边界不由任意想象决定。属性纠缠与微积分的桥梁，就建立在这两种力量不能彼此分离的地方。

第一编与第三编参考资料

以下资料分为五部理论底本、补充理论材料与物理数学原典。引注中的章号指向底本网页的章节，不对应本书页码。网页核对日期为 2026 年 10 月 16 日。本书所列思想案例、数学反例和模型计算，除明确注明原典来源者外，均为本次写作的说明性构造，不是实测数据。

五部理论底本

[1] 王德生：《属性纠缠论——实体属性发生学导论》，德麦国际专著第 396 号。重点参读第 15 章的属性发生组织式与发生层／意识层区分，第 24—25 章的聚合、物体／物质／物理分析，第七编的三属性纠缠及相应方法论论述。网址：https://sdeuniverses.com/books/m/396/text/。

[2] 王德生：《属性纠缠与科研新典范——从不收敛病理学到科学发生学》，德麦国际专著第 379 卷，统稿版 v1.1。重点参读第 3 章“三方程六路径三原理的联合方法”、第 6—10 章的物理学分析、第 11—15 章的读数事件、测量封装、不收敛及模型比较。网址：https://sdeuniverses.com/books/m/379/text/。阅读范围包括第 3 章正文。

[3] 王德生：《科学发生的结构——超越库恩》，德麦国际专著第 342 号。重点参读科学典范三组合（Function，D，E）、三方程与三组合的区别、典范碰撞与再生相关论述。网址：https://sdeuniverses.com/books/m/342/text/。

[4] 王德生：《科学与微积分——差异、整体与文明的数理内核》，SDEUniverses 专题长文。重点参读第 3—8 章关于微积分前史、牛顿与莱布尼茨、微分表征、积分重构和科学循环的论述。网址：https://sdeuniverses.com/column/science-and-calculus/。该底本采用较早的 SIO／ΔSIO 表述；本书与现行 SDE 的连接属于明确标出的再阐释，不作机械符号替换。

[5] 王德生：《科学是什么？》，德麦国际专著第 231 号。重点参读第 19—22 章的符号、逻辑、数学与实体科学建构，尤其第 21 章“微分—积分—再微分”的方法论循环。网址：https://sdeuniverses.com/books/m/231/text/。

补充理论材料

[6] 王德生：《何谓物》，SDEUniverses 专题文章。用于第 1 章视觉、听觉、触觉 SDE 特征纠缠聚合体的理论入口。网址：https://sdeuniverses.com/column/what-is-a-thing/。

[7] 王德生与 ChatSDE 智能体 WDS：《与 WDS 的对话》，2026 年 10 月 16 日 08:08 导出，用户提供文件《物理学思想简史.pdf》，153 页、51 轮。重点使用第 150—153 页对“记账”比喻的修正及物／物体／物理发生系统的研究命题；较早的不收敛、量子测量与科学史强断言，作为待检验研究材料而非已证史实。

物理数学原典与计量资料

[8] Isaac Newton, “The October 1666 Tract on Fluxions.” Cambridge University Library, MS Add. 3958.3, ff.48v–63v. 本书核对 Newton Project 规范转录本，尤其命题 7—8，分别讨论量的关系与速率关系、由速率关系返回量的关系。网址：https://www.newtonproject.ox.ac.uk/view/texts/normalized/NATP00100。正文的圆周约束和初值演算采用现代记号，不是手稿逐字翻译。

[9] Gottfried Wilhelm Leibniz, “Nova methodus pro maximis et minimis...”, Acta Eruditorum, October 1684. 本书核对第 467 页折射段的英译节选“Leibniz on Refraction”，Stony Brook University 页面，译文改编自 D. J. Struik 编《A Source Book in Mathematics, 1200–1800》。网址：https://www.math.stonybrook.edu/~tony/archive/336s20/documents/Leibniz-translation.html。正文另以现代变量独立重推加权折线路径，不照抄节选页面的排印公式。此处引证范围限于上述折射段，不覆盖全部手稿及优先权争论。

[10] Archimedes, Geometrical Solutions Derived from Mechanics: A Treatise of Archimedes. J. L. Heiberg 德译；Lydia G. Robinson 由德文译为英文；David Eugene Smith 撰引言。Chicago: Open Court Publishing Company, 1909. 本书使用阿基米德致埃拉托色尼的开篇方法说明，区分机械启发与几何证明。Project Gutenberg 版本：https://www.gutenberg.org/ebooks/7825。

[11] Bureau International des Poids et Mesures, The International System of Units (SI), 9th edition，官方在线版本。用于量、数值、单位以及定义与具体实现的区分。网址：https://www.bipm.org/en/measurement-units；手册：https://www.bipm.org/documents/20126/41483022/SI-Brochure-9-EN.pdf。

[12] Leonhard Euler，“Sur la vibration des cordes,” Mémoires de l’Académie des Sciences de Berlin, vol.4, pp.69–85，1750；写于 1748 年，Enestr̈om 编号 E140。Euler Archive 文献记录：https://scholarlycommons.pacific.edu/euler-works/140/。本书使用原始论文的题名、年代与研究对象定位；现代波动方程的模型推导在正文中独立给出。

[13] Joseph Fourier, Théorie analytique de la chaleur. Paris: Firmin Didot, 1822. 用于热运动方程与不同几何区域的整体传播问题之原典定位；本书的现代热方程和模态计算为独立说明。影印／转录入口：https://fr.wikisource.org/wiki/Livre:Fourier_-_Théorie_analytique_de_la_chaleur,_1822.djvu。

第四编 牛顿世界：新力学物体与经典法则的创造

第十六章 牛顿重造了怎样的物
质量、惯性、运动量

16.1 一块石头如何进入一个新世界

桌面上放着一块石头。人可以看它的灰色，摸它的棱角，掂量它的沉重，也可以把它当作镇纸、纪念品或者建筑材料。现在把它移到一条水平轨道上，让它接受牵引并记录位置。石头没有因为进入实验而失去颜色，但颜色暂时不再决定方程中的位置；它的来历没有被消除，却可能不参与此次运动的预测。相反，原先不直接显眼的质量、速度变化和外部作用，开始共同规定这究竟是怎样一个物。

本章所问的“牛顿重造了怎样的物”，不是问牛顿是否制造了这块石头，而是问：一种能够接受统一动力学处理的物体，经过怎样的选择、连接和检验才得以成立？如果把全部工作概括为给既有石头增加几个准确描述，就会遗漏最重要的改变。此前尚未在这种统一动力学结构中取得共同位置的石头、滑块、月球，在新的发生系统中能够占据相同类型的动力学位置；以前紧密相连的沉重感、落地速度和颜色，则被重新分离。什么差异具有解释资格，什么相同允许跨场景迁移，都发生了变化。

在本书的 SDE 解释中，牛顿力学创造的物理世界，是能够持续重现某类显露的视觉、触觉及其他 SDE 特征纠缠聚合体系统。轨迹记录、接触阻抗和数学关系不是给一个已经完备的世界分别加上附件，它们参与规定力学物体的成位。但这一论断不意味着书写符号就能任意产生运动。新世界必须承受失败：预测的石头如果没有抵达给定位置，解释系统需要查明失配发生在哪里。创造具有构成性，也具有可被实际结果纠正的约束。

16.2 定义之间的相互支撑

《原理》开篇把物质量联系于密度与体积，把运动量联系于速度与物质量；随后区分惯性与外加作用，并进一步区分向心作用的不同量度。这里用的是牛顿原典的概念顺序；以下使用的向量和导数则是现代重述，并非其原文排式。[16-1]

这些定义的意义不在于提供一套孤立词典。假设有人说，质量就是一个物体含有多少物质。紧接着就必须追问：怎样比较“多少”？如果用体积比较，压缩会改变体积；如果用手感比较，姿势和支撑会改变感受；如果用重量比较，又要辨别地点与测量条

件。一个概念成为科学量，需要形成可交叉检查的比较组织。每种操作都可能有偏差，因而不能把某一次读数直接提升为不受条件影响的最后实体。

现代教学中常写质量等于密度乘体积。这在密度均匀时成立，却不能仅凭这个式子宣布已经解决质量的发生问题，因为密度本身又常被定义为单位体积的质量。数学上的关系可以正确，经验上的定量程序仍需另外建立。对于不均匀分布，需要对体积分布求和或积分；但在积分之前，单位、局部量度、边界和可加性已经承担工作。积分能整理这些规定，不能凭空替代它们。

由此可以看出，“物体有质量”比日常语言中的“东西有多重”多出一个层次。它要求某项量在特定的操作变换中保持可识别性：换一种运动状态，换一组测量组合，仍能把不同记录关联到同一个动力学参数。这个稳定性不是第一眼就全部给出的。反复实验、误差处理和预测回检，使稳定关系获得可靠的适用范围。在这一意义上，质量的稳定与物体的同一性共同成立，而不是先有绝对确定的同一物体，再随意给它挂上一个数字。

16.3 惯性改变了需要解释的事情

三条运动定律的最低限度内容可以压缩为：无相应外作用时保持静止或匀速直线运动；运动的改变联系于所受作用及其方向；两物体的相互作用大小相等且方向相反。牛顿还在推论中讨论合成、总运动量和整体运动。[16-2]

其中第一项改变了问题的重心。日常推车经验容易让人觉得，运动必须由持续推力维持：手一松，车便停下来。惯性结构则要求把“继续运动”和“速度改变”分别处理。松手以后逐渐停止不是第一定律的直接反例，因为轨道、轴承和空气的作用并未自动消失。真正需要检验的是：随着这些作用受控地减小，速度变化是否也按可说明的方式减小？这样，理想化就不能只是口头宣布“忽略摩擦”，它必须提供一个可以逐步接近并检验后果的比较方向。

更微妙的变化是静止失去了独占的正常地位。在一个适当参照描述中，某个物体不移动；在另一个匀速平移的描述中，它一直移动。二者不应仅凭一个速度数值是否为零而被划为两种不同的受力状态。速度取决于描述关系，而速度变化的判定还需要明确参照框架的性质。把第一定律缩减为“物体很懒”会造成误解：它不是给物体添加人格，而是在给哪些状态可以作为无外力基准建立约束。

如果把惯性直接说成藏在物体内部的某种推力，也会混淆概念。现代惯性系中的匀速运动不需要额外的维持力。改变观察框架后，为了保留某种方程形式引入的惯性力，

又属于另一种建模安排。它不能在没有说明参照条件时，与物体间的接触作用混写为一类。概念的历史沿革可以保留，但方程的使用必须让读者知道自己正在采用哪套语法。

因此，牛顿式物体的一个核心特征，是它能在“保持状态”与“改变状态”的比较中得到识别。这个物并不是凭某个瞬时外观成立，而是在一段可比较的序列中成立。仅凭一张照片，我们不能判断一辆车是否加速，也不能判断一个静止物体是否没有受力。现象的意义需要时间组织和作用关系共同参与。

16.4 运动量让碰撞取得共同尺度

设两个沿同一直线运动的小车，质量分别为一千克与二千克，碰撞前速度分别为每秒三米和零。取前进为正，初始总动量为三千克米每秒。若碰撞后两车粘连，且碰撞期间外界冲量可忽略，那么共同速度为每秒一米。这里的数值是本章为了说明而设置的理想算例，不是对某次历史实验的复述。

这个计算为什么值得注意？因为它表明，物体是否“动得多”不能只看速度。较大的质量能够以较小的速度承担相同动量；两部分运动还必须带着方向相加。迎面运动的两个物体，即使各自速度很大，总动量也可能为零。总动量为零不等于系统内部没有运动，更不等于没有碰撞效应。量度的引入没有把所有差异压成一个数字，而是明确了某一种比较究竟保存什么、允许什么不同。

在上述粘连算例中，初始平动动能为四点五焦耳，碰后整体平动动能为一点五焦耳。少掉的三焦耳不能解释为能量凭空消失，而是提醒我们：仅仅追踪两车整体平动，没有完整追踪形变、内能等变化。这个算例同时展示了同一个过程可以接受不同量的组织。动量守恒不能代替能量分析，动量与能量也不能因都含有质量和速度就相互混同。

进一步说，碰撞前后动量相等，本身不足以唯一确定任意碰撞的结果。如果两车不粘连，我们还需要接触性质或其他补充条件。方程的强大并不等于一条守恒关系包办一切。一个理论为物体建立公共尺度之后，仍会遇到不能由公共尺度唯一决定的差异。承认这些差异，是模型保持诚实的条件，也是后续研究能够继续发生的入口。

这使“物”的统一显出两个面向。一方面，木块和金属块可以作为相同质量的动力学单元进行比较；另一方面，它们的形变和恢复过程可能大不相同。新物理世界的创造并非把全部材料差异消灭，而是建立一个允许统一预测、同时标明差异重新进入位置的组织体系。

16.5 不能把三个符号当作三件孤立的东西

用现代记号，对于经典力学范围内、由固定物质成员构成的系统，在惯性参考系中，总外力等于总动量对时间的变化率。对于质量恒定的质点，可以写为：

$\Sigma F_{外} = dp/dt = ma$，$p = mv$（F、p、a、v 均为向量）。

这里写的是合外力，方向不能省略；质量恒定也不是装饰性条件。如果使用扩展物体的整体平动描述，右侧的加速度应是质心加速度，转动和形变通常还要补充其他关系。若选择一个不断有物质进入或离开的开放系统，则需要处理跨边界的动量通量，不能不加分析地把该系统瞬时质量乘瞬时速度后求导，再把所得结果直接等同于外力。系统边界是什么，决定了哪一项属于内部改变，哪一项属于穿越边界的传输。

这个公式至少包含三种相互约束的角色。加速度把位置差异组织为速度变化；质量规定同类外作用与响应之间的比例；力的模型说明什么相互作用产生何种定向影响。若把力定义为质量乘加速度，再用同一个等式宣称验证了第二定律，就只能得到形式上的自洽。真正的检验需要额外约束，例如用不同装置建立的作用关系，能否预测另一组运动记录；或者在质量已经通过其他过程校准后，响应是否符合预定比例。

这不要求每个量在所有方面都完全独立于其他量。科学测量经常依靠相互校准。关键是不能只有一个可以任意调节的封闭圈：遇到任何结果都重新命名力，遇到任何失配都重新指定质量。至少有一部分量和操作必须先被锁定，再让预测面对新的记录。约束越能跨装置、跨状态保持，理论越不只是为已有曲线寻找一种写法。

从 SDE 看，质量、力和加速度之间的意义不能还原为三件各自完整之物的外部联系。它们的可测性、可比性和适用性在特征纠缠中共同成立。不过，也不能反过来把三个符号分别贴成 S、D、E。一个质量测量过程本身可以接受完整 SDE 分析；一个加速度显露同样需要差异序列和特征纠缠；力的推定也不例外。三维所分析的是发生结构，不是给物理词汇分配三个互斥抽屉。

16.6 微积分进入之前，谁已经规定了变量

为了说明局部与整体的连接，设一维运动中合外力恒为二牛顿，质点质量为一千克，初始位置和初始速度都为零。上述条件给出加速度为每二次方秒二米；一次积分得到速度，二次积分得到位置：

$v(t) = 2t$，$x(t) = t^2$。

这里时间 t 以秒计，v 和 x 分别取米每秒和米的数值。两秒后速度为每秒四米，位置为四米。相同的恒力与质量，如果初始速度改成每秒三米，则位置成为三倍时间加时间平方。因而微分方程规定运动如何继续，初始条件参与选择实际追踪哪一条轨迹。不能只展示漂亮的加速度关系，却把初始准备当成理论外面不重要的杂项。

更早一步的问题是：何谓同一质点的位置？从一张图像取前端、下一张取后端，计算出的加速度可能主要来自识别规则变化。时钟采样也会影响差分估计。时间间隔缩小并不保证读数越来越好，因为噪声可能在求导时被放大。由此可见，数学极限与实际测量的收敛不是自动重合的，需要说明定位方法、采样尺度以及误差对结果的影响。

积分同样不只是把细碎信息装在一起。一个有效的整体轨迹，需要局部关系在连续过程中仍然适用。如果物体撞上障碍，恒力假设可能失效；如果途中脱落一部分，固定成员的选择可能改变；如果转动使受力方向发生变化，一维描述可能不够。每一次积分都把某套连续性承诺带到更大范围，而预测回检负责审查这种承诺是否越界。

本章因此不把微分直接认作 D，也不把积分直接认作 E。微分与积分是数学操作，D 与 E 承担的是发生分析。二者可以发生密切联系，但必须通过变量成立、差异组织、局部关系和整体检验这些具体环节联系起来。否则，符号上的对照很容易掩盖物理工作的真正困难。

16.7 日常物没有被取消，模型省略需要负责

现在回到最初的石头。如果问题只是它在短时间内沿光滑轨道移动多远，颜色通常不进入运动方程。如果问题改为阳光照射后表面温度如何改变，颜色所对应的光学性质可能重要；如果石头在高速气流中运动，形状与表面状况就可能重新影响轨迹。“不重要”因此不是对某个属性的永久判决，而是一个依赖于问题、尺度和容许误差的选择。

同一块石头也可能在使用者的生活中具有不可替代的纪念意义。力学模型不负责从质量和速度中推出这种意义；但这不意味着意义是一种不真实的附加幻觉。它意味着力学世界以特定的特征纠缠组织对象，其适用范围不能自动覆盖全部现实、自我与理念发生。只因为力学能够准确预测落地时刻，就断言它已经穷尽这块石头的一切存在规定，是把成功的局部封装误作无条件的总体完成。

模型省略还必须接受反向检查。若两个外形相似的对象，在同样的牵引安排下表现出系统差异，研究者不能仅说“我们已经忽略了这项因素”。省略是一项有条件的承诺，不是一张免于修正的许可证。可以先检查摩擦，再检查转动、连接松动和质量分

布，最后判断是否需要改变对象模型。每一步都应产生能够区分解释的观察安排，而不是把所有可能原因堆成背景清单。

这里的 E 正好区别于这种清单。温度、地面、观察者和仪器的名字相加，不会自动成为特征纠缠分析。必须说明某项特征怎样改变另一项特征的成立条件：温度改变接触材料的响应，接触响应改变牵引与滑动之间的对应，新的对应又影响从轨迹推定作用量的程序。只有进入这种构成关系，环境因素才成为当前发生的解释部分。

16.8 三方程怎样共同约束力学物体的成立

以一个可重复的牵引实验为例，S 可以取“某一质量参数与轨迹规律在多次试验中稳定显露”；D 是释放、定位、计时、改变牵引条件和进行对照的差异序列；E 是位置识别、时钟读数、接触响应与数学比较之间形成的特征纠缠。在这一具体划分下，S＝F(D,E)要求解释稳定规律如何产生，而不满足于给规律命名。

但研究不会停在第一次显露。已经得到的 S 与已有 E，共同要求改造 D：如果低速段偏差大，就增加相应采样并改变释放方式；如果不同方向结果不对称，就安排反向试验。这里体现 D＝G(S,E)。新的 S 又与新 D 共同改变 E：某些原先被混合的接触效应被分离，时钟与定位之间的校准关系获得修正，原先的“一个物体”可能需要拆成平动与转动相互约束的模型。这体现 E＝H(S,D)。三式中的函数只是结构性表达；没有指定可测变量和运算规则前，不能把它们冒充已经拟合出的物理定律。

六路径要求完整经过三维，而不是只在两个维度之间画箭头。E—D—S 的沉淀路径，由已建立的测量纠缠组织试验，形成稳定显露；E—S—D 的锚定路径，先由纠缠明确可期待的显露，再组织取得它的序列；D—E—S 的倒逼路径，从实际操作暴露的问题改组关系，取得新显露；D—S—E 的原型路径，由一次操作所得的初步结果推动关系重建；S—E—D 的进入路径，从异常显露追查构成条件，再安排区分性试验；S—D—E 的再创路径，从已有显露重做差异序列，使原有特征纠缠获得改造。这里给出的是六种研究进入顺序，不能把名称当作已经完成的论证。例如增加采样若不检查定位噪声，未必改善加速度显露；更换模型若不产生可以区分模型的新操作，也未必推进知识。

三原理在本章的工作方式，是检查两维之间的张力如何推动第三维改变，并且让改变返回原先矛盾处接受检验。S 与 D 不相容时，要求检视 E 是否错误捆绑了摩擦与惯性；S 与 E 不相容时，要求重组 D 以分辨竞争解释；D 与 E 不能共同支持稳定测量时，则要求修订 S 的对象规定和适用范围。推动并不保证成功，回检也不能省略。若每次修改都只增加自由参数而没有增加可区分预测，发生分析就失去了约束。

上述划分不等同于主体、互动、客体三号位。选择实验问题涉及主体性规定，牵引和测量涉及互动性规定，稳定的动力学物体涉及客体性规定；三者都可以拥有 S、D、E 的完整结构。把主体等于 S、互动等于 D、客体等于 E，会同时损害两个分析层次。

16.9 创造论必须面对自己的竞争解释

有一种更节制的解释会说：牛顿只是发明了有效模型，物理对象本身从未被创造。这个竞争解释指出了一项真实约束：理论改变不意味着石头能随思想任意改道。本书接受这一约束，却进一步追问，作为具有质量、动量和定量受力关系的“力学物体”，能否在抽掉全部测量、差异组织和特征纠缠以后，仍保有完全相同的规定？如果不能，那么新物理世界的成立就不仅是给独立成品换一套描述。

这一追问不是单凭物理方程就完成的本体论证明。方程的预测成功支持它所组织的关系，却不会自动排除全部其他哲学解释。本书的贡献在于提出并贯彻一种发生性分析：追踪对象规定怎样获得稳定、跨操作迁移与可修正能力，并比较这种分析是否揭示了单纯描述论容易遗漏的环节。只有展示具体差异，才能让“创造世界”成为具有解释内容的命题。

牛顿物体的力量正在这里显现：它允许多种日常之物在共同动力学结构中相遇，使相互作用、运动改变和可比较量共同参与物体的成立；它也明确地留下入口，让材料、转动、流动与其他未被充分展开的差异重新进入。质量不再只是沉重的别名，惯性不再只是静止的顽固，运动量也不只是速度的另一种称呼。新的物由这些相互约束的规定稳定成位。

下一步要问的是，这种已经成位的物如何在空间和时间中获得普遍可计算的运动秩序。局部作用怎样延伸为整体轨迹，近地经验怎样与天体运动发生连接，数学上的统一又需要付出哪些理想化代价？只有继续追踪这些问题，才能看清牛顿世界的创造既不止于三个定律，也不止于一个公式。

本章参考资料

[16-1] Isaac Newton. The Mathematical Principles of Natural Philosophy. Andrew Motte 译, 1729. “Definitions”，定义 I—VIII。本章依据公开原典转录核验术语，现代公式另作明确标识。https://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_(1729)/Definitions

[16-2] Isaac Newton. 同上，“Axioms, or Laws of Motion”，定律 I—III、推论 I—V。https://en.wikisource.org/wiki/The_Mathematical_Principles_of_Natural_Philosophy_(1729)/Axioms,_or_Laws_of_Motion

第十七章 万有引力
天空与地上为什么能成为同一物理

17.1 月亮为什么没有落下来

看见石块下落，人会指向脚下；看见月亮越过夜空，人会指向天上。这两个动作似乎天然划出了两个世界。一个世界里有支撑、碰撞与沉重，另一个世界里有遥远、周期与不可触及的光。倘若只凭日常形态分类，月亮显然不像石块；但物理创造的关键，往往正是改变哪些差异应该先被比较。牛顿式统一没有要求月亮与石块长得一样，而是重新规定它们能够怎样发生共同的运动关系。

“月亮也在下落”因此不是一句故意违背常识的修辞。这里的下落不是说月亮的地心距离每时每刻都在缩小，而是说，相对于没有相应作用时的切向前进，月球运动持续发生朝向地球的偏转。如果把“下落”限定成离地面越来越近，这句话当然令人困惑；若把它改写为运动变化的方向和大小，问题就进入另一套可测量关系。一个日常动词在新系统里取得了新的物理规定。

这一规定需要两种运动同时成立：切向运动使月球不断前进，径向作用使其轨迹偏离切线。只剩前者，就不再得到绕地运行；只剩后者而没有适当初始条件，也不会获得所讨论的轨道。轨道不是一个天体单独携带的形状，而是质量关系、空间分布、初始运动和时间比较共同构成的结果。本章借这个问题追踪的，正是天空之物与地上之物怎样进入同一个发生系统。

17.2 一次跨越距离的数量桥接

《原理》第三卷命题四把维持月球运行的作用，与地表重物下落的作用进行定量比较；原文同时注明，先把地球视为静止是一项计算假设，考虑共同质心运动需要修正。这一细节说明，天上地上的统一并非只凭相似故事完成，而是经过尺度转换与近似条件的检查。[17-1]

可以用一个刻意简化的现代算例重做这座桥。取地球半径为六百四十万米，月地距离为六十个这样的半径，即三亿八千四百万米；取月球运行周期约二百三十六万秒，并暂时按匀速圆周运动计算。所有数值都只是用于展示数量级的近似输入，不能当作一套精密月球历表。圆周加速度由半径与周期给出：

$a_{月} = 4\pi^2 r / T^2 \approx 0.00272$ 米每二次方秒。

如果这种作用随距离平方的倒数变化，那么把距离由六十个地球半径缩小到一个地球半径，加速度应放大三千六百倍，得到约每二次方秒九点八米。这与近地自由落体的熟悉数量级相接。反过来，从近地重力加速度出发，按距离平方衰减，也能够给出月球轨道所需要的向心加速度数量级。

有意义的不是最后两个数看起来接近，而是它们的来源具有不同操作结构。月球轨道的半径与周期来自空间定位及长时间的天文比较，近地落体来自另一套实验安排。它们不是把同一份读数换一个名称后再互相证明。当不同操作链条在明确转换规则下相容，统一便具有超过比喻的经验内容。

当然，这个算例也把自己的局限带到了台面上。月球轨道不是精确的圆，地球不是绝对静止的球，太阳等作用也没有消失。近似结果的一致只支持相应精度范围内的桥接，不能证明所有细节均已处理。越是要强调统一的创造力，越需要保留这些条件，否则统一会退化为一种无法失败的宣言。

17.3 从圆周规律到万有引力并非一步跳跃

用现代符号表示，两个质点之间的牛顿引力大小为：

$F = Gm_1m_2/r^2$。

方向沿连接两质点的直线，属于吸引作用。这个简式不能无条件套给任意延展物体，更不能把两个物体最近表面的距离代入其中。对于球对称质量分布，在相应外部区域可以用中心集中质量的等效表达；对于不规则分布，则需要把不同部分的作用按方向合成。公式短，不等于对象构成简单。

为什么距离平方的倒数能够与行星周期建立联系？在中心体远比绕行物体质量大的圆轨道近似中，向心加速度为四倍圆周率平方乘半径再除周期平方，引力加速度为中心体引力参数除半径平方。令二者相等，就得到：

$T^2 = 4\pi^2r^3/\mu$，$\mu = GM$。

于是，若围绕同一主导中心体的两个圆轨道半径之比为四，周期之比应为八。这里从运动记录到作用形式，有一个可检查的数学桥梁；但桥梁同时带着圆轨道、主导中心体和其他作用可忽略的条件。在一般椭圆两体问题中，对应量是半长轴，并且引力参数包含两体质量之和。不能把圆周推导换几个名词，就假装已经证明所有实际行星运动。

由此也能看出，观测轨道主要约束的是引力参数，而不是自动分别测出 G 和中心天体质量。如果一组观测只给出它们的乘积，单靠这组观测就无法把两个因子独立分开。

这一点对于本书很重要：新物理对象并不是由一个漂亮公式一次性全盘显露，量之间仍有辨识限制。要把参数进一步拆开，需要另一些相互独立的实验关系。

万有引力的“万有”也不是从一个轨道的拟合中直接跳出来的。它包含把同类作用扩展到各种物体及其部分的主张。《原理》第三卷有关重力的论证把实验所及物体的性质与更广范围的推广相联系；这是一种有经验支持的推论安排，不等于人类已经实际检查宇宙里每一份物质。[17-2]

17.4 整体不是把一点放大

把地球看作一个中心点，是一种非常有力的封装，但这项封装自身也需要理由。设想把一个大球切分成许多小块：离试验物体较近的小块作用较强，较远的小块作用较弱，不同方向的小块还会产生部分抵消。整体作用既不能只取最近部分，也不能把所有作用的大小不分方向地相加。

从局部作用得到整体规律，需要同时处理质量分布和方向结构。球对称提供的便利，在于不同方向上的成对贡献能够按严格关系合成，使外部作用具有简洁表达。这不是“很大的物体自然等于一个点”，而是特定空间组织让复杂局部关系形成可封装的整体结果。改变分布，封装是否仍然有效，就要重新检查。

如果把同样总质量从近似球体改造成细长形状，并在其附近进行测量，不同位置的作用就未必还能只由到某个中心的距离决定。总质量相同不能保证近场完全相同。距离足够远时，一些形状细节对结果的影响可能变得较小，但这个“足够”必须相对于所需精度和物体尺寸来说明。忽略细节不是宣布细节不存在，而是规定它们在当前尺度上的影响预算。

这给微积分提供了明确任务：在空间分布已经被规定、局部作用关系已经获得含义的条件下，对各部分贡献进行合成。数学积分帮助我们从局部组织整体，积分之前的边界和密度模型也同时参与结果。若密度分布是假定的，计算结果就必须保留这一假定；若从结果反推分布，问题又会面临不唯一性。相同的远处观测可能对应多种内部结构，不能把可计算误认成全知。

天空与地上因此不是被一根抽象的线简单绑起来。地球作为重力源、月球作为绕行物体、石块作为近地试验对象，获得了一套可以相互转换的规定，但各自内部结构和观测尺度仍然不同。统一建立的是相容关系，不是抹掉所有对象差异。

17.5 轨道是初始条件与作用关系共同写出的

同一种引力作用为什么既能对应坠落，也能对应环绕，还能对应离去？答案不能只在力的公式里找。先在中心体固定、试验质点质量可忽略的近似下讨论：给定一个位置，如果初始速度为零，运动与给予适当横向速度时显然不同。如果初始速度方向变了，即使速率不变，轨道也会不同。因此，“给出力”并没有完成全部运动问题。

在距中心 r 的位置，圆轨道速率满足 $v_{圆}^2 = \mu/r$。与此同时，把无穷远处势能规定为零，单位质量的机械能可写为：

$\varepsilon = v^2/2 - \mu/r$。

在两体近似、无耗散且仅有引力的条件下，能量为负可以对应束缚运动；但它并不保证质点永不撞到中心体的实体表面。径向运动与有角动量的运动仍须区分。所谓轨道类别，依赖速度大小、方向以及研究对象是否被视为数学点等条件的共同规定。

令离去时无穷远速度趋于零，可以得到逃逸速率 $v_{逃} = \sqrt{(2\mu/r)}$，它是同位置圆轨道速率的平方根二倍。这个结果并不意味着物体达到该速率后便不再受引力，而是意味着在所给模型内，它具有足够的能量不必返回有限距离。词语“逃逸”若被理解为走出引力存在的范围，就会把能量条件误写为作用突然消失。

上述关系构成另一种原创检验：如果半径扩大四倍，圆轨道速率减半，周期扩大八倍；如果把中心体质量增为四倍而半径不变，圆轨道速率扩大两倍，周期减半。参数改变给出的不是孤立数字，而是一组相互约束的响应。实验或观测可以围绕这些共同变化来检验模型，比只展示一个拟合成功的轨道更有区分力。

真正的轨道计算还必须面对初始数据的不确定性。位置差一点、速度方向差一点，后续轨迹就可能逐渐分开。计算机把方程积分到更多位数，并不会自动消除输入的不确定性。理论的确定性与具体预测的有限精度可以同时成立；不能把后者都当成理论失败，也不能用前者掩盖后者。

17.6 潮汐不是整片海水被拉向一边

海水涨落把万有引力的统一推到一个更困难的位置。这里不再只是追踪一个中心的位置，而要解释一个有空间范围的系统内部为什么产生相对运动。《原理》第三卷命题二十四将海潮与日月作用及其相对位置联系起来，讨论了作用的合成与涨落变化。这是原典层面的历史事实；下面使用的梯度表达是现代分析，不冒充牛顿原文。[17-3]

关键不在于月球是否吸引海水，而在于它对地球不同位置的吸引不完全相同。地球整体也受到月球作用，所以只计算海水的绝对加速度会漏掉参照系统的运动。要研究海水相对地球的变化，需要比较局部与整体的加速度之差。这样，问题从作用本身转向作用随空间位置怎样改变。

设地月距离为 r，沿地月连线取一个相对地心的小位移 x，且 x 远小于 r。把月球引力加速度在地心附近作一阶展开，沿该方向的相对加速度大小具有约二倍月球引力参数乘位移、再除距离立方的尺度：

$|\Delta a| \approx 2\mu_{月}|x|/r^3$。

这个式子只描述特定方向的小范围线性近似，不能直接当作全球海面高度公式。它揭示一个清楚差别：整体吸引的主尺度随距离平方倒数变化，而潮汐差异的主尺度随距离立方倒数变化。讨论谁对潮汐影响更大，不能仅比较谁对地球的总吸引更大。

近地一侧海水比地心更强地朝月球加速，远地一侧则比地心更弱地朝月球加速。相对于共同运动的地球，二者可以产生方向相反的拉伸趋势。于是，两侧涨起的基本图景并不需要给远侧再发明一种把海水推离月球的神秘作用。需要的是把参照运动与空间差异一起纳入，而不是只画一支指向月球的箭头。

但从这个趋势到某个港口几点涨潮，还隔着海盆形状、水深、地球自转、摩擦和流动响应。相同的天体强迫，进入不同海岸系统，可以形成不同振幅和相位。它们不是对引力解释的自动否定，也不是可以无限追加的免责理由。每引入一个因素，都应尽量说明它怎样改变可观测量，并以独立资料或新预测加以约束。

17.7 能够统一不等于能够一次算尽

日月、地球和其他天体一旦共同进入方程，理想两体问题就变成多体关系。一个天体的路径影响它与其他天体的距离，这些距离又改变下一时刻的作用。把各作用相加只是建立微分方程，通常不意味着已经得到一个简单的封闭轨迹表达。数值积分可以向前推进，却仍须控制步长、舍入和模型遗漏。

为了理解这一点，可以设想一个最小的预测任务：先用一段观测确定初始位置和速度，再预测随后未参与调整的数据。如果后段记录持续偏离，需要区分三类来源：输入估计不准、计算方法误差、作用模型不充分。三类问题可能同时存在，但修正它们的检验方式不同。把所有偏差都归于观测，会保护理论；把所有偏差都归于理论，则会忽略测量责任。

更有力的研究不是无限增加可调参数，而是让修正承担额外承诺。若认为遗漏了一项外部作用，就应计算它对其他方向、其他时段或其他对象有什么连带后果；若认为某种定位程序有偏差，就应通过改变观测安排检查偏差是否随之变化。独立后果越具体，解释越可能被否定，也越可能获得可信支持。

这让“统一”具有了工作性的含义。统一不是不允许例外记录出现，而是面对差异时，仍能提出相互制约的修改路线。当天空运动与地面运动在同一规则下进行比较，两个领域开始互相提出问题：地面测量约束天体参数，天体记录又检验地面规律的远距离延伸。物理世界从并排的现象集合，发展成能够相互检验的发生系统。

牛顿引力也有它的适用域。高速、强引力及要求更高精度的情况不能只靠经典模型概括。本章不在这里提前展开后面的相对论论证，但必须保留逻辑界限：广泛有效与绝对无界普适不是同一命题。一个理论能够被后继理论重新定位，并不会倒过来使此前真实实现的统一变成毫无价值的幻觉。

17.8 三维回路中的天空与地上

现在把本章的统一作为一个具体 SDE 问题。S 不是单独一个算出的引力数字，而是月球偏转、近地落体和潮汐差异之间可重现的相容显露；D 是长期定位、时间比较、近地实验、模型计算与预测回检的差异序列；E 则是距离、周期、质量分布、参照运动及作用规律之间相互规定的特征纠缠。若只把太阳、地球、海洋列成一张背景表，还没有解释它们如何参与当前显露。

S＝F(D,E)追问这些相容结果怎样产生；D＝G(S,E)追问已有相容与残差怎样推动新的观测顺序；E＝H(S,D)则要求结果和实际操作共同检视模型中的关系组合。例如，月球定位残差不能仅凭一句“还有别的因素”处理，它应推动更明确的差异测量，反过来决定哪些作用项需要进入。三式共同工作，统一才保持可修正性。

本章最值得走通的是 S—E—D 的进入路径。先发现月球实际位置与理想两体轨道之间存在系统残差，这是进入点；随后检查残差是否与某种空间配置相联系，把原先合并为一个“误差”的现象分解为可竞争的构成关系；最后选择那些能够让不同解释产生不同预测的观测时段与方位。回到新记录时，若残差仍不随所假设的作用变化，原解释就需要撤回。路径的终点不能只是一套更复杂的方程，而应是能够裁决解释的差异序列。

另一条具有实际意义的是 E—S—D 的锚定路径。先在既定作用结构中推导某项尚未用于调整的轨道特征，规定应该出现何种显露，再安排观测获取它。这与看到结果以后寻找一套能够包含结果的关系不同。两种工作都可能有价值，但证据地位不能混同：前

者让模型提前承担失败风险，后者还需在新的比较中检验。完整三维顺序把这种区别保留下来，避免把所有成功拟合都称为预测。

三原理关注的是两维张力推动第三维改变并返回检查。S 与 D 冲突时，检查 E 中是否把实际偏心轨道过度简化为圆；S 与 E 冲突时，重组 D 以区分周期测量偏差与模型问题；D 与 E 无法共同支持现有对象规定时，修订 S 的表达，例如由单一轨迹转向带有不确定范围的预测。每一次调整都要回到引发张力的记录，而不能只在概念层面宣布问题已被提升。

17.9 新的物理世界与未被关闭的问题

从落体到月球，再到潮汐，发生的并不是一种对象吞并了其他对象。触觉中的沉重与视觉中的天体位移，各自在完整 SDE 中成立，通过测量、数学和实验形成新的纠缠。人并没有亲手把月球放到天平上，却能够用不同过程之间的相容关系建立可检验的质量和引力参数。这种间接性不是知识的缺陷，而是新物理世界得以扩展的构成条件。

仍然可以提出一种竞争解释：这些工作仅仅发现了早已存在的统一规律。本书不以一次轨道计算反驳这种哲学立场，而是指出它容易略过的问题：当“同一种力”“同一种质量”“同一条可计算轨道”成立时，哪些操作与关系已经参与规定它们？若把这些条件全部当作外在通道，就可能看不到科学创造重组了物的同一性、量的可比性和检验的公共性。

因此，本书所说创造物理世界，保留着严格的论证责任。它必须显示，新发生系统能够产生以前不能如此关联、计算和回检的物理规定；同时承认，这一解释不授权研究者任意支配观测结果。月球不会因为我们偏爱一个公式就沿它运行，海潮也不会服从为了理论整齐而删去的边界。创造与受约束并不互相取消，它们共同决定科学创造的性质。

天空与地上成为同一物理，最终表现为不同操作可以在明确规则下相互转换并彼此追问。但要让这样的物理超出某个人、某架仪器和某座观测台，仍须解决一个尚未完成的问题：不同地方的钟、尺和天平怎样进入同一套量度关系？下一章转向这项经常被放在公式后面、实际上参与公共物理世界发生的工作。

本章参考资料

[17-1] Isaac Newton，The Mathematical Principles of Natural Philosophy，Motte 英译，1846 版，第三卷命题四，原印页 392。https://en.wikisource.org/wiki/Page:Newton%27s_Principia_(1846).djvu/398

[17-2] 同上，第三卷命题六相关推论，原印页
396。https://en.wikisource.org/wiki/Page:Newton%27s_Principia_(1846).djvu/402
[17-3] 同上，第三卷命题二十四，原印页
415。https://en.wikisource.org/wiki/Page:Newton%27s_Principia_(1846).djvu/421

第十八章 测量与单位
经典物理如何创造公共“物理”

18.1 两位诚实的人为什么会量出不同结果

两间实验室收到同一根金属杆。甲报告长度略大，乙报告长度略小。我们不能因为二者都诚实，就宣布杆子同时拥有两个互相冲突的长度；也不能立刻认定其中一人操作粗心。甲在较暖的房间测量，乙先把杆子放在较冷的恒温箱里；两人使用不同的支撑位置，也可能把端面上的不同接触点作为边界。数字之间的差异，先要求我们询问他们究竟在实现怎样一个“长度”。

这不是测量结束后才补写的注意事项。若测量对象的状态、边界和比较方式没有被规定，所谓相同的量尚未充分成立。金属杆在日常辨认中可以保持同一，却在温度、受力和放置方式改变时产生不同尺寸。公共物理必须提供一种程序，使这些差异能够被记录、转换和检验，而不把任何一次读数直接封为对象的全部真相。

“谁来测都一样”因此应当被改写为：不同合格测量者在规定条件及明确变换规则下，可以得到彼此相容的结果。相容不要求最后一位数字完全相同，更不意味着任何时间、任何地点、任何仪器都必须显示同一个数。它要求差异具有可检查的来由，并且在说明不确定性的条件下进入共同比较。

在本书的主轴下，这种公共性参与物理世界的创造。没有它，公式即使在一间实验室内有效，也还不能稳定进入其他人的实验。钟、尺、天平以及围绕它们形成的校准、记录和复查，构成了不同 SDE 能够相互接续的条件。它们并非把一个早已公共完备的物理世界搬运到各地，而是在实际工作中不断建立公共“同一”的资格。

18.2 先问量什么，再问量得多准

金属杆的长度可以指自由放置时的端面间距，也可以指一定张力下的标线距离。它还可以指沿弯曲中心线的长度，或者两个端点之间的直线距离。这些问题都可以合理，但不能把它们的结果未经说明地放在同一列中平均。若量的定义发生改变，更多小数位只会把概念差异包装得更加精密。

设一根略弯的细杆，沿表面的路径长于两端直线间距。一个人用软尺贴着表面量，另一个人用端点定位装置量，两者差异可能来自对象定义，而不是仪器失准。解决办法不是先买更贵的仪器，而是明确实际任务：制造插入某个槽口的部件时，端点位置或外

包尺寸可能关键；计算材料用量时，沿杆长度又可能更重要。测量问题的提出参与选择何种物理规定需要稳定显露。

这里也存在模型尺度。把一个表面视为平面，就暂时压缩了微小凹凸；提高分辨率以后，接触点的选择可能改变结果。若规定测头直径和接触力，测量便获得一个明确的作用尺度。这个尺度不必被视为缺陷，它可以是任务所需的合理封装。但若研究问题改变，就需要检查此前封装是否仍然保留了应当保留的差异。

对于时间也是如此。测一只摆的周期，要先规定“回到同一状态”是什么意思：从左端到右端，还是从左端经过右端再回到左端？若数过中点的次数而不记录方向，很容易把半周期当作全周期。钟并没有替人自动解决所有事件识别问题。读数之前，事件边界已经通过观察、触发和记录规则获得规定。

因此，精确测量并非先有一个孤立数字，再努力减少周围噪声。它是对象状态、操作事件、尺度与比较关系共同形成的显露。对定义的检查，不是反科学的怀疑，而是避免把不同事情误当作同一事情的第一道工序。

18.3 尺子也会变化

为了让开篇的金属杆问题可以计算，设定一个简化模型：在所考察的小温度范围内，杆的线膨胀系数近似恒定，且杆温度均匀、没有额外应力。用参考温度下的长度 L_0 表示，温度升高 ΔT 以后有：

$L \approx L_0(1 + \alpha \Delta T)$。

若 L_0 为一米，假设 α 为每开尔文百万分之十，温度高出十开尔文，则长度约增加零点一毫米。这里膨胀系数是本章自设的示例参数，不指认某种具体材料的标准值。对粗略木工来说，这个变化也许可以忽略；对精密配合来说，它可能足以影响使用。

但测量尺本身也可能膨胀。如果杆和尺具有相近的膨胀系数，又处于相同温度，它们的长度比可能变化很小。于是，稳定读数不等于对象没有变化；对象和参照共同变化，可以使比值保持近似不变。若换用另一种材料的尺子，又可能显出此前不明显的差异。测量从来不是对象单方面向一只完全不变的眼睛交出属性。

这类问题要求把温度计、尺、杆以及它们各自的响应关系放入一个可检查模型。E 在这里不是“房间温度、空气、人员”等背景名称的堆积，而是温度如何改变杆和尺的长度，温度计读数如何参与修正，接触与热平衡怎样影响读数的成立。只有明确这些相互规定，环境因素才成为当前测量的构成解释。

我们还要检查修正是否超出证据范围。如果只在很窄的温区估计了膨胀系数，就不能毫无检验地用同一线性式外推到极端温度。校正公式也会拥有适用域。把一项不可靠修正精确计算到十位小数，不会使结果更可靠；它只会使错误承诺获得更漂亮的外表。

18.4 钟如何把连续变化变成公共时间

测量摆动时，一个直接办法是手动启动和停止计时。若人的反应迟延在起止两个时刻都相同，部分影响可能抵消；若起点依靠视觉、终点依靠听觉，迟延结构可能不同。结果的偏差不是“人参与了”这一事实本身，而是参与方式中哪些差异没有被识别和控制。把人从叙述中删去，不会让这些差异自动消失。

常用的改善是计时多个周期，再除以周期数。假设起止总计时的不确定性规模约为零点二秒，测二十个周期时，平均每周期对应的这一分量约为零点零一秒；若只测一个周期，就无法获得同样的缩减。这只是一个说明操作设计的简化估计，未把所有误差都收入其中，也没有假设重复次数能无限消除系统偏差。

例如，摆的振幅若持续减小，而周期在研究精度下受振幅影响，二十个周期的平均值就不能无条件当作某个固定振幅下的周期。延长观察序列降低了一类计时影响，却可能引入状态变化的影响。操作改进必须同时检查它改变了什么，不能把“多测一些”当成普遍有效的按钮。

不同实验室要比较周期，还需要知道各自时钟的速率是否相容。零点相差一小时不会必然损害一秒间隔的比较，速率偏差则会持续改变间隔结果。把两个时钟调到同时显示十二点，只解决了一个时刻的对齐，不等于已经验证它们随后走得同样快。这一区分使同步、频率比较与长时间稳定性成为不同任务。

经典测量中的时间因而不是一串从天空直接落下的刻度。周期事件、计数方式、起止判定和时钟比较共同构成可使用的时间尺度。后来的物理还会进一步追问不同运动状态和引力条件下的时钟关系；在本章阶段，至少已经可以看到，同样的数字格式并不保证同样的测量含义。

18.5 天平为什么不是一只放大的手

手掂物体，感受到的是身体、支撑方式和物体共同产生的响应。天平把比较重组为另一种互动：在近似相同的重力条件下，比较两侧的作用及力臂关系。对于理想等臂天平，在忽略浮力等影响时，两边重力中的共同因子可以约去，于是能够比较质量。这里起作用的是一种对共同条件不敏感的结构，而不是人的感觉被简单放大。

如果两边力臂稍有差异，直接把平衡等同于质量相等就会出错。可以交换物体与砝码的位置，利用两次比较识别或抵消某些影响。这个动作很普通，却体现了重要的知识生成方式：不是增加一条抽象假设，而是改变差异序列，让原先隐藏在一次读数里的装置属性显露出来。

空气浮力也会参与精密称量。两个质量相同、体积不同的物体，在同一空气中受到的浮力不同；因此天平实际响应的作用差，不能在任何精度下都直接等同于真空中的质量差。修正需要密度、体积和空气状况的信息。如果这些输入不够可靠，就应把残余影响计入结果的不确定性，而不是因为已经写出公式就声称全部消除。

用弹簧装置称量还要区分力与质量。装置在一个地点按质量标度刻好，并不意味着移到不同重力条件下仍然无需调整就能显示同一质量。这里不是质量随意变化，而是仪器把作用量转换成显示数字的关系依赖条件。公共单位的名字印在面板上，不足以保证整个测量过程已经公共化。

从 SDE 分析看，天平创造的是一种新的可比较显露：物体、砝码、支点、交换操作和读数规则共同支持稳定的质量关系。客体性的质量规定不等于 SDE 中的 E；E 指的是这些特征怎样互相成立。主体选择何种精度，互动决定装置怎样响应，客体显露在这样的组织中取得可迁移性，三号位与三维必须始终分清。

18.6 单位的公共性需要制度，也需要实验

公共单位不是靠每个人心里觉得“大概一样”维持的。一八七五年签署的《米制公约》建立了国际计量合作的组织框架；一八八九年的第一届国际计量大会认可米与千克的国际原器，米原器的长度还连同规定温度一起获得定义。这里的物质实物、状态条件、比较程序与国际协作共同进入公共尺度。[18-1][18-2]

这段历史不应被简化为政府下令，所以大家采用同样数字。行政承认可以规定使用哪个单位，却不能凭命令使一根金属杆不再膨胀，也不能保证副本与原器完全无差别。公共性要求把实际差异测出来并记录，使各地使用者知道自己的参照与共同参照之间有什么关系。恰恰是允许记录不完全相同，标准体系才有可能支持更可靠的一致。

同时，实验能力也需要组织条件。标准器怎样保存、谁有资格使用、比较结果怎样公开、损坏后怎样处理，都会影响量度关系能否持续。把这些全称为科学外部背景，会看不到知识传播为何可能在一个环节中断。但把它们全称为权力决定，也会遗漏仪器和比较结果对制度安排提出的抵抗。

设一个地区宣布所有工作尺都等于标准尺，而实际比较显示其中某批尺存在稳定偏差。若制度只维护名义上的相同，公共测量会失效；若制度允许记录偏差、修正转换并追查来源，统一反而得到加强。可纠正性是公共标准的组成部分，不是标准权威受到威胁时才临时打开的例外。

这也说明公共物理的成立不等于知识从某个中心向外原样复制。副本进入新的温度、运输与使用条件后，需要当地的比较和维护。中心与地方通过回报差异不断修正关系网络。传播本身参与知识继续发生，不能把传播者想成只搬运不变数字的容器。

18.7 校准不是把指针拨到满意的位置

现代计量词汇把校准描述为，在规定条件下建立标准量值与仪器示值之间的关系，并利用这一关系得到测量结果；它明确区分校准与调整。这里借用现代术语说明经典测量问题，不把二十一世纪的定义逐字安到早期实验者头上。[18-3]

做一个自拟算例。假设某温度显示装置在窄范围内近似满足“示值等于参考温度值乘一点零二，再加一度”。那么直接读到五十二度时，按该模型修正得到五十度。把装置的零点调好，只能处理加一度这一部分，并不能同时保证一点零二的比例偏差消失。一个校准点也不足以决定一个带斜率与截距的线性关系。

即使两个校准点恰好落在线上，中间区间仍可能存在非线性；若要判断它，应增加能够检验线性假设的点。校准不是为了让仪器看起来与标准完全相同，而是为了把转换关系及其限度说明白。调整可以改善显示便利性，完整的校准则需要在调整后确认关系有没有改变。

计量溯源进一步要求，测量结果能够通过有记录且不中断的校准链联系到参照，各环节都贡献不确定性。现代定义还特别提醒，具备溯源性不自动保证结果适用于某项任务，也不自动排除错误。[18-4]

假设一项加工要求长度差控制在百分之一毫米内，而测量链的不确定性远大于这个尺度，那么一张有效的溯源说明并不能使结果突然足够精细。它说明结果与参照如何联系，却不能替代用途判断。这种区别阻止我们把“有标准”“有证书”“有数字”当作任务已经完成的三个同义句。

18.8 不确定性让差异进入结果

本章使用的不确定性不是“也许一切都错了”的空泛怀疑。现代计量学用它表达：依据已经使用的信息，对被测量赋值时还存在怎样的分散范围或规模；其中可以包含统

计重复、修正量和标准量值等来源。[18-5] 它不等于一个已知但故意不改正的错误，也不等于仪器显示的最小分度。

例如，一次长度测量显示一百点零零毫米，不能只凭末尾两位小数断言结果准确到百分之一毫米。显示格式可能很细，而温度修正、接触变形和校准输入仍较不确定。反过来，一项测量最后只报告到十分之一毫米，也可能已经完全满足所需任务。数字多寡与知识质量之间不存在简单的单调关系。

设两个互不相关的标准不确定性分量，分别为零点零三毫米和零点零四毫米；在适合使用方差合成的模型下，合成标准不确定性为零点零五毫米，而不是零点零七毫米。若两个分量具有相关性，计算就必须加入相应项。这里的数字是示例，不是一项真实实验报告。它说明不确定性的组织需要模型，不能把所有可能影响的最大值随意相加，也不能不看依赖关系就一律开平方。

同一仪器重复一百次所得的平均值，可能降低一部分随机分散，却保留共同的标度偏差。若一百次都使用同一支偏长的尺，读数彼此非常接近也不能证明尺度正确。高度重复与外部相容是两种不同检验：前者在相对固定条件下看稳定，后者通过改变条件和参照看可迁移性。

认识到不确定性，并不是让结果从公共世界退回私人意见。相反，它把原先含混的“差不多”转化为可以讨论的结构，使研究者知道哪些差异值得追查、哪些差异暂不足以判定冲突，以及下一次改善应该作用于哪一个环节。误差与不确定性管理因此参与物理对象的稳定，而不只是给已经稳定的对象贴上尾注。

18.9 跨实验室的相容是怎样检验的

设两个实验室测量已经约定状态的一根杆，结果分别为一百点零零毫米和一百点零六毫米，二者各有零点零五毫米的标准不确定性。如果两项结果不相关，差值的标准不确定性约为零点零七一毫米；若预先采用两倍标准不确定性作为比较阈值，这个差值没有越过该阈值。我们可以说它们在该判据下相容，不能说已经证明两个实验绝无差错。

现代计量相容性概念正是把结果差与差值不确定性联系起来，并要求考虑结果之间的相关性。[18-6] 两间实验室如果都使用同一来源的校准，可能分享某项共同偏差。它在两者相减时抵消，因而彼此接近不必然表示两者都接近一个独立参照。只看不同机构的招牌，不能自动认定证据独立。

更有区分力的比较，可以安排不同方法、不同参照链或不同条件下的交叉检验。还可以让同一件传递标准按一定次序经过各实验室，并在返回时检查它是否改变。如果标

准在运输途中变化，实验室之间的差异就不能全算到实验能力头上。公共比较不仅检查参与者，也检查承载比较关系的物。

为避免事后迎合结果，比较规则应尽可能预先明确：测量什么状态，报告哪些修正，采用什么不确定性说明，遇到离群结果怎样调查。规则可以在经验中改善，但不应在看见不喜欢的数据后悄悄重写。否则，表面上的一致可能只是筛选机制制造的，而不是测量系统实现的。

“谁来都一样”由此获得一种更强也更具体的含义：换人、换装置、换地点以后，结果能够在公开的变换与不确定性规则下维持相容；不能相容时，又有办法定位是哪种条件改变了。公共性不是差异的消失，而是差异能够参与共同检验。

18.10 从一次偏差到一个新测量系统

设想一组跨实验室比较反复出现同一方向的长度差。按 SDE 分析，S 是已经稳定显露的偏差模式，D 是运输、放置、温度平衡、读数和交换操作的序列，E 是材料响应、参照尺度与修正模型的纠缠结构。三方程在这里不只是再写三个符号：S＝F(D,E)要求说明偏差如何发生；D＝G(S,E)要求用偏差设计新的比较；E＝H(S,D)要求让比较结果有资格修改既有模型。

最值得实行的是 D—S—E 的原型路径。先安排同一根杆在两个温度下重复比较，观察差异是否随温度改变；由这个初步显露检视原先的温度修正是否错用了尺的膨胀系数。若只是重复原操作而从不改变温度，就可能长期得到稳定偏差，却无法区分参照问题与对象问题。这条路径的价值来自具体差异安排，而不是三个字母的出现。

随后可以转入 E—D—S 的沉淀路径：把修订后的温度、材料和读数关系纳入程序，在新的运输和测量批次中执行，再观察原偏差是否消失或缩小到可解释范围。若修订只在提出它的那一批资料中有效，就尚未完成沉淀。新数据迫使模型承担额外风险，公共性才不是自我认证。

三原理的张力检查贯穿其中。S 与 D 的张力要求审查 E 中的共同标度假设；S 与 E 的张力要求改变 D 以检验温度解释；D 与 E 的张力则可能要求重新规定 S，把含混的“杆长”改为某个规定温度与支撑状态下的长度。第三维的改变必须返回原问题接受检查，不能靠改名把差异从记录中抹去。

18.11 标准能够更新，公共物理才能继续发生

度量衡并未永远停在金属原器。二〇一八年国际计量大会通过修订，二〇一九年五月二十日起，国际单位制以七个定义常数的固定数值为基础，其中包括真空光速与普朗克常数。[18-7] 本章提到这一后续变化，是为了标明经典标准化工作的延伸，不是提前把现代量子实现归入牛顿时代。

从实物原器转向常数定义，并不意味着人们从此不需要仪器、校准与不确定性。定义规定单位，实际实现仍需要具体装置和受控过程。一个定义常数具有精确规定的数值，不等于每一间实验室实现相应单位时都没有误差。若把定义的精确性与实施的精确性混同，新的标准同样会被误用为不必检查的权威。

变化之中保持连续，是标准更新的重要任务。如果每次理论进步都让日常一米、一秒、一千克任意跳变，公共比较将无法累积。因此，新的定义与实现需要处理旧尺度的接续，使既有知识能够在可追查的关系中继续使用。这里再一次显示，稳定不是拒绝变化，而是能够组织变化而不丢失比较资格。

经典物理创造公共“物理”的成就，正体现在这些看似细碎的工作里。方程给出关系，测量让关系获得可共同实行的入口，标准让不同入口之间能够接续，误差分析又让接续不靠掩盖差异维持。本书的发生论解释并不否认物质对测量的约束，而是把这种约束与人的操作、制度和数学组织放回同一个复合发生过程。

由此通向后续章节的，不是一个终于排除所有人的纯净世界，而是一个能够通过训练、仪器与公开规则，让不同主体参与并相互纠正的物理世界。它之所以公共，不是因为没有任何人参与，而是因为参与方式能够被说明、比较、替换与修复。

本章参考资料

[18-1] BIPM，The Metre Convention。https://www.bipm.org/en/metre-convention

[18-2] CGPM，Declaration of the 1st CGPM (1889)，认可米和千克国际原器。https://www.bipm.org/en/committees/cg/cgpm/1-1889/resolution-

[18-3] JCGM，VIM3，2.39 Calibration。https://jcgm.bipm.org/vim/en/2.39.html

[18-4] JCGM，VIM3，2.41 Metrological traceability。https://jcgm.bipm.org/vim/en/2.41.html

[18-5] JCGM，VIM3，2.26 Measurement uncertainty。https://jcgm.bipm.org/vim/en/2.26.html

[18-6] JCGM，VIM3，2.47 Metrological compatibility of measurement results。https://jcgm.bipm.org/vim/en/2.47.html

[18-7] CGPM，Resolution 1 of the 26th CGPM (2018)，On the revision of the International System of Units。https://www.bipm.org/en/committees/cg/cgpm/26-2018/resolution-1

第十九章 拉格朗日、哈密顿 旧力学为什么能再生新形式

19.1 同一只摆，为何值得换一种语言

上一章把公共物理的成立放回钟、尺、天平及其校准关系之中。现在，测量程序已经比较稳定，研究者能够反复记录一只摆的位置，问题似乎只剩把数据代入力学方程。然而，当摆连着另一只摆，支点又可以移动，或者摆球被限制在一条曲线上时，最熟悉的空间坐标未必还是最合适的组织方式。测得准确，并不保证问题已经被有效地提出。

设一颗小球由不可伸长的轻杆连接到固定支点，只能在竖直平面内运动。用横、纵坐标描述它，必须时刻保证两坐标满足同一个圆的条件，还要引入杆对小球的约束作用。若改用杆与竖直方向的夹角，允许的位置则由一个数给出。少掉的不是某块真实空间，而是一个不必独立指定的变量。物理约束被吸收到变量的选择之中。

这样的改变有时只是计算方便，有时却会改写研究者看得见的问题。角度让振动的往复结构更清楚，两个角度可以把双摆组织为一个整体状态，进一步的坐标变换又可能显出共同振动与相对振动。新语言的价值不由符号是否新颖决定，而由它是否使原先被纠缠在一起的任务获得可操作的分解与重新连接来决定。

本章因此既要说明分析力学的创造，也要限制一种过快的本体论宣布：把牛顿方程改写成拉格朗日或哈密顿形式，并不自动产生新的经验预测，更不自动证明一个新世界已经从无到有。我们需要追踪，形式改变如何参与新的对象组织、计算操作和检验安排；只有这些环节实际发生变化，形式再生才超出对既有结果的重新排版。

19.2 广义坐标怎样让约束进入对象

广义坐标并不神秘。它们是一组足以在研究范围内指定系统构形的独立参数，可以是距离，也可以是角度或其他适当量。对于平面单摆，一个角度即可；对于两颗可独立运动的自由质点，通常需要更多参数。坐标数量所表达的是模型保留的自由度，不是器物表面看上去包含多少零件。

必须补上“在研究范围内”这个条件。一套坐标可能在某些位置失效，或者不能用一个数值唯一识别所有构形。一个角度增加一整圈，杆的空间朝向可以相同；若模型还要记录累计转过多少圈，就需说明是在圆上取位置，还是在展开的角度轴上记录运动历史。相同符号的数值差，未必代表相同类型的物理差异。

约束的性质也很重要。轻杆长度固定，可以写成位置之间的关系，适合通过减少坐标处理；某些滚动约束涉及速度，未必能积分为单纯的位置限制。若把所有约束都当作“少写一个变量”，可能改变允许的运动集合。分析力学的整齐依赖对约束的准确分类，而不是把不方便的作用藏进一个新坐标后便不再负责。

以摆为例，沿允许位移方向，理想杆的约束作用不做虚功，这使我们能够暂时不显式求出杆的拉力而得到角运动方程。但如果研究问题恰好是杆是否断裂，张力就必须重新求出。把约束作用从某组运动方程中消去，不等于作用在现实中消失，也不等于模型已经回答关于装置的一切问题。

这种对象重组对 SDE 具有直接意义：变量不是任意标签，它决定哪些差异能够被独立改变，哪些差异通过约束一起变化。把一组笛卡尔坐标换成广义坐标，可能重新组织 D 中的可实施变化以及 E 中的约束关系。若只把坐标变换看作纸面换名，就容易遗漏其中被重新封装的操作条件；若把每次换名都称为存在革命，又会把真正的创造门槛降得过低。

19.3 从受力列表到一个函数

对本章讨论的保守系统，可把拉格朗日函数取为动能减势能，即 L＝K−V。用广义坐标 q 和广义速度 q̇ 表达，运动方程写作：

d/dt(∂L/∂q̇) − ∂L/∂q = 0。

多自由度时，对每个独立坐标分别写一个方程。现代教材用这套形式系统处理坐标和约束，明确说明它与满足相应条件的牛顿运动方程相容；这里采用现代记号，不把教材排式冒称拉格朗日原稿的逐字复制。[19-1]

为什么一个函数能够替代许多分量关系？因为它不是一个随便挑选的曲线，而是已经把系统的动能、势能与坐标结构收进同一个表达。求导操作再从整体表达中取出各方向的变化关系。如果漏掉某一部分动能，或者把势能的方向写反，统一形式不会替我们修复错误。函数的紧凑恰恰要求对它的物理构成更认真。

对于单摆，取摆球质量 m、杆长 l、重力加速度 g，以最低点势能为零，则：

L = ½ml²θ̇² − mgl(1−cosθ)。

对 $\dot{\theta}$ 求偏导再对时间求导，得到 $ml^2\ddot{\theta}$；对 θ 求偏导，得到 −mgl sinθ。代回运动方程，有 $\ddot{\theta}+(g/l)\sin\theta=0$。这个结果没有先求出每时每刻的杆张力，却保留了沿允许运动方向的动力学。推导中每个系数都来自一个明确的构成条件：质量集中于摆球，杆质量忽略，长度不变，支点固定，空气阻力和摩擦暂时忽略。

若支点开始运动，或者杆可以伸长，不能只把旧方程中的某个常数随意换成时间函数便认为处理完成。动能的表达可能出现新的交叉项，自由度也可能增加。实际装置改变时，首先应回到函数的构成检查，再求运动方程。分析力学的优势，是提供一种可以重建的组织程序，而不是一条永远不用重新理解装置的公式。

19.4 一个能够复算的摆动检验

为了把上述推导与数字联系起来，取杆长一米，g 取每二次方秒九点八米。小角度时用 sinθ 近似 θ，得到$\ddot{\theta}$＋九点八 θ＝0，角频率约为每秒三点一三，周期约为二点零零七秒。设初始角为零点一弧度、初始角速度为零，则在线性近似中：

θ(t) = 0.1 cos(√9.8 t)。

从释放到第一次经过最低点的时间约为零点五零二秒。此时角速度大小约为每秒零点三一三弧度，摆球速率也约为每秒零点三一三米。用小角度势能近似计算，初始每单位质量能量为二分之一乘九点八再乘零点一的平方，即零点零四九焦耳每千克；最低点的动能给出同一结果。

还可以用未线性化的势能作一次独立检查。释放角仍为零点一弧度，由能量关系得到最低点速率 √[2gl(1−cos0.1)]，约为每秒零点三一二九二米。它与小角度结果接近，却不完全相同。这个微小差别不是两套力学互相矛盾，而是近似所删去的高阶影响开始显露。

如果把释放角增大，却仍要求周期严格保持二点零零七秒，问题就不是实验者不够熟练，而是线性近似被过度外推。一个良好的检验应当改变振幅，检查偏差是否随之系统增长，同时控制阻力和支点运动。若偏差只与计时方式有关，便不支持振幅解释；若在独立计时与不同装置中都跟随振幅变化，解释才获得更强约束。

这个算例同时划定创造的边界。牛顿方法、拉格朗日方法和能量方法在各自适用条件内给出相容结果，这不是三次独立实验支持。它们共享同一个模型假设。不过，三种表达能互相揭示计算错误，帮助我们辨别某个失败来自代数、近似还是物理构成。形式之间的相容属于理论自检，实验之间的相容属于经验检验，二者都重要，但证据地位不同。

19.5 驻作用量并不是自然预先知道终点

拉格朗日方程还可以从一个整体路径条件得到。为避免与 SDE 的显露 S 混淆，本章把作用量记为 A：

$A[q] = \int L(q,\dot{q},t)dt$。

在起止时刻及端点构形固定的变分问题中，真实运动路径满足一阶变分 $\delta A=0$。这通常称为驻作用量条件。“最小作用量”是广泛使用的名称，但一般不能把驻值直接等同于所有候选路径中的绝对最小值。边界条件、允许的变分范围和函数本身的性质，都参与这一判断。[19-1]

一个常见误解由此产生：物体似乎需要先知道终点，再比较所有路线，选择最好的一条。实际上，变分形式是对运动关系的另一种数学组织，不要求自然物体实施一次有意识的规划。在适当正则条件下，它导出局部运动方程；给定初始位置与速度，又可以按局部方程求后续运动。端点变分问题与初值预测问题有联系，却不是把物体的行动解释为预知未来。

可以用最简单的自由质点直接验证。令质量为 m，在时长 τ 内从位置零走到 a。候选路径取 $x(t)=at/\tau+b\sin(\pi t/\tau)$，其中第二项在起止时刻都为零。将速度平方积分，交叉项的积分为零，得到：

$A = ma^2/(2\tau) + mb^2\pi^2/(4\tau)$。

在这组候选中，$b=0$ 给出较小作用量，即匀速直线路径。这里没有神秘选择，只有一个可复算的比较：加入额外的弯折速度，会增加该模型中的积分值。这个特例确有最小值，但不能从它推出所有系统的驻作用量都必定是全局最小。

如果路径中含有摩擦耗散，单纯使用上述 $L=K-V$ 及封闭系统条件就未必够用，需要加入适当广义力、扩展系统或采用其他正确处理。驻作用量的威力不在于把任何过程硬塞进一个看似美丽的积分，而在于用明确条件连接局部方程与整体路径，允许我们发现哪些结构能够共存。

19.6 从构形走向相空间

只知道摆球在最低点，还不能知道下一刻往哪边走。它可能从左向右经过，也可能从右向左经过，甚至正停在那里。构形给出位置安排，尚未完整给出经典初值问题所需要的状态。把广义坐标与相应动量同时纳入，就得到相空间中的状态描述。

定义共轭动量 $p=\partial L/\partial \dot{q}$。若速度到动量的关系能够局部反解，便可以进行勒让德变换，写出哈密顿函数：

$H(q,p,t) = p\dot{q} - L(q,\dot{q},t)$。

右侧的速度最终须用 q、p 和 t 表示。多自由度时，对各 $p\dot{q}$ 项求和。哈密顿方程为 $\dot{q}=\partial H/\partial p$，$\dot{p}=-\partial H/\partial q$。这里的变换需要正则性；遇到退化拉格朗日函数或约束系统，不能不加处理地照用反解步骤。[19-2]

动量也不总是我们最初熟悉的“质量乘直线速度”。在单摆角坐标下，$p\theta=ml^2\dot{\theta}$，它具有角动量的量纲。坐标变化会改变共轭变量的形式，存在速度相关相互作用时更须分清正则动量与机械动量。若见到字母 p 就自动代入 mv，反而会被新语言误导。

哈密顿的研究并非只是把微分方程换行排列。他在一八三四年《动力学的一般方法》的引言中，把复杂运动问题转向特征函数的研究，并明确承认困难可能只是由一类积分转移到另一类问题。原文还指出其方法与此前光学研究的联系。[19-3] 这种自我限制值得保留：中心函数的出现，没有使所有多体问题自动得到简单答案。

相空间的意义在于把位置与继续运动的条件放进同一个状态点，把运动组织为状态空间中的流。这个空间不是日常三维空间的另一名称。一只一维振子的相空间已经有位置和动量两个方向；许多自由度组成更高维的状态组织。维数增加不等于宇宙多出若干肉眼可见方向，而是模型需要同时指定更多独立差异。

19.7 同一振子在两种空间里的不同显露

取一个一维弹簧振子，质量为一千克，劲度系数为每米四牛顿。忽略阻尼，平衡点取为位置零，释放时位移为零点一米、速度为零。哈密顿函数为 $H=p^2/2+2q^2$，以国际单位制数值计算。哈密顿方程给出 $\dot{q}=p$，$\dot{p}=-4q$，因而：

$q(t)=0.1\cos(2t)$，$p(t)=-0.2\sin(2t)$。

总能量为零点零二焦耳。物理空间里的运动是在两端点之间往复；相空间里的轨迹则是一条椭圆，满足 $(q/0.1)^2+(p/0.2)^2=1$。经过四分之一周期，即 $\pi/4$ 秒，位置为零、动量为负零点二千克米每秒。过最低势能点并不表示状态没有差异，动量方向记录了它将怎样继续。

这条椭圆把不同信息同时呈现出来。它让转折点、最大速率与能量约束互相定位，也使不同初始条件对应的轨道族容易比较。但不能因此把椭圆当成振子在房间里真正绕出的路线。视觉图上的几何是状态关系的几何，其物理意义由坐标轴和测量程序共同规定。图画得更直观，不等于图中的所有方向都是空间方向。

给模型加入阻尼后，能量会转移到未显式追踪的自由度，简单振子的相轨迹可以向平衡点收缩。若仍用无阻尼的封闭二变量哈密顿描述，却把收缩说成标准哈密顿流的性

质，就混淆了模型层级。为了在更完整层面讨论能量守恒，需要把相关环境自由度纳入，或者明确承认当前使用的是有效开放系统描述。

这个比较带来一项可败条件：如果宣称装置属于无阻尼模型，就应在适当误差范围内看到能量及相轨迹的稳定；若连续记录显示系统性收缩，必须检查阻尼、测量反馈与参数漂移。不能一边保持原模型的全部前提，一边把所有不相容结果都归入“复杂性”。相空间的价值，是把这种不相容变得更容易辨认。

19.8 守恒结构为什么比一条轨迹更有迁移力

对于一个没有显式时间依赖的正则哈密顿系统，沿运动有 $dH/dt=0$。直接代入两条哈密顿方程，就能看到坐标变化与动量变化的贡献相消。如果 H 显式依赖时间，结果则变为 $dH/dt=\partial H/\partial t$，不能再因它名叫哈密顿函数就宣布它永远守恒。[19-2]

同样，如果某个广义坐标没有显式出现在拉格朗日函数中，相应共轭动量可以守恒。这个结构使我们不必每次先求出完整轨迹，再从结果中猜测有没有不变量。我们可以在求解之前检查函数怎样依赖变量，由此发现降维和比较的可能。这是形式再生的重要收益：变化从哪些地方进入，哪些量能够在变化中保持，获得了系统的提问方式。

但是“不出现”需要相对于一套明确表示来理解。一个坐标变换可以把原先显然的对称藏起来，也可以把原先不明显的对称显露出来。若变换显式依赖时间，还要处理附加项和能量解释。数学形式的简洁并不是对象本身突然变得简单，而是研究者找到了与某种结构更相容的组织方式。

在自拟振子算例中，守恒椭圆可以帮助检查数值计算。假如使用一种粗糙算法，计算轨迹逐圈向外漂移，研究者应先检查离散化是否人为增加能量，而不是立即宣布物理装置能够自行增能。相反，算法恰好保存某个量也不保证所有轨迹细节都准确；相位误差仍可能积累。守恒检查是必要的结构检验之一，不是取代全部误差分析的通行证。

由一条具体轨迹转向轨道族、守恒量和变换结构，意味着研究对象的组织层级提高了。它让不同机械系统之间出现新的可比关系，为后来跨领域的理论迁移提供了条件。但这种可能性需要后续真实研究来兑现，不能从形式优美直接跳到所有未来成果都已包含其中的历史叙事。

19.9 数学等价为什么不是创造的自动证明

设有两套形式，在相同对象定义、相同初始条件和同样适用范围内，能够一一映射并给出相同可观测结果。就这一范围而言，它们没有通过互相改写增加一条新的经验事实。如果只在新形式中重新算出旧轨道，再宣布物理世界已经改变，论证是不够的。

然而，经验等价也不意味着认知与研究操作毫无变化。一套形式可能使约束系统更易处理，使守恒关系先于解显露，使模型能够与另一领域发生结构性接续。这些变化改变了人—仪器—数学复合体能够稳定完成的工作。发生论应在这里提出具体证据：哪些问题以前难以组织，现在可以系统提出；哪些预测或装置设计因此能够实现；哪些失败可以更早被识别。

一个竞争解释会把这些都称为工具改进，认为对象及世界完全不变。本书不靠贬低工具来反驳它，而是追问工具改进何时进入对象规定本身。若相空间只是画出已知轨迹的辅助纸面，它的本体论分量有限；若它参与规定统计状态、可达区域或后续理论中的基本关系，研究对象就可能不再由原先的孤立位置列表充分表达。这里需要逐案论证，而非一次性宣布所有形式都是同一个层级的创造。

判断也须接受反向可能。如果新形式只带来更多符号，既不改善推导，也不增加可检验结构，所谓再生可能只是修辞。如果形式变换不满足可逆条件，却被错误宣称与旧模型完全等价，则它可能悄悄增删了物理条件。创新评价不能只看新语言的气势，要看它改变了哪些实际发生能力，并且是否如实承担了改变后的检验责任。

19.10 两条路径怎样把形式改写变成研究

可以从约束摆的问题具体实行 D—E—S 的倒逼路径。原有逐分量计算 D 不断被张力和几何限制缠住，研究者于是重组 E，把允许构形、动能和势能组织为广义坐标中的关系；再由新结构产生可与原实验比较的角运动 S。如果推导无法回到同一装置的记录，程序便没有完成。少写了张力并不算成功，正确保留它对允许运动的限制才算。

进一步实行 E—S—D 的锚定路径：从已建立的哈密顿结构 E 先推定能量轨道的某种显露 S，再安排同时记录位置与速度的 D，检查轨迹是闭合、漂移还是收缩。这与事后在数据上画一条椭圆不同。新的测量安排需要处理速度估计、采样误差和模型范围，形式预测才真正进入可失败的操作。

三方程在这两条路径中共同运转。S＝F(D,E)要求轨道显露由实际序列与结构支撑；D＝G(S,E)要求用偏差修改测量或计算；E＝H(S,D)则允许稳定的不相容结果反过

来改变原先的模型。不是每一次偏差都要抛弃理论，也不是每一条理论都可以靠调整数据永远保留。

三原理进一步要求，两维间的张力推动第三维改变并返回检查。S 与 D 不相容，可推动审查 E 中的无阻尼假设；S 与 E 不相容，可推动改变 D 的采样与初始准备；D 与 E 不能共同支持旧状态描述，可推动 S 由单位置曲线扩充为相轨迹。每次扩充都须说明新增信息来自什么操作，而不是让抽象名词代替观察。

19.11 旧力学的再生留下了什么边界

拉格朗日与哈密顿形式的重要性，不在于它们让牛顿力学从此没有困难，而在于它们改变了困难能够怎样被组织。约束不再只能以额外作用逐个处理，状态不再只能由空间位置图表达，守恒关系不再只能在解出轨迹以后偶然发现。新的数学组织为物理世界的继续发生提供了可以复用的条件。

与此同时，本章一直保留几条边界：驻值不等于全局最小，广义坐标不等于任意删去自由度，共轭动量不总等于质量乘速度，哈密顿函数不在所有表示和条件下都等于守恒机械能，形式等价不等于经验新增。这些边界不会削弱理论的美，反而使美能够与可靠的推导和失败检验共同成立。

还剩一个更深的问题。无论使用哪种形式，经典模型仍需规定参考系、惯性响应与允许的运动结构。我们可以优雅地描述一个系统怎样运动，但惯性为什么具有这种地位？它只是物体独自带着的性质，还是需要与其他物体以及整体世界的组织相联系？分析形式的成功没有自动消除这个问题；下一章将看到，马赫正是在经典力学最稳定的根基处重新提出了发生性的追问。

本章参考资料

[19-1] David Tong，Classical Dynamics，Chapter 2，The Lagrangian Formalism，§§2.1—2.3。https://www.damtp.cam.ac.uk/user/tong/dynamics/dynhtml/S2.html。本章仅简述一般形式与条件；单摆、自由质点变分与数值算例为自拟推导。

[19-2] David Tong，同上，Chapter 4，The Hamiltonian Formalism，§4.1。https://www.damtp.cam.ac.uk/user/tong/dynamics/dynhtml/S4.html。

[19-3] William Rowan Hamilton，On a General Method in Dynamics，1834，Philosophical Transactions，247—308；David R. Wilkins 校订转录，2000，重点引言与开头。https://www.maths.tcd.ie/pub/HistMath/People/Hamilton/Dynamics/GenMeth.pdf。未声称逐页精读全文。

第二十章 马赫与经典物理的界限
惯性真是物体自带的吗

20.1 把一切算清以后，仍有一个问题

分析力学能够把受约束的运动组织成广义坐标，也能把位置与动量变成相空间中的一条轨迹。可是，公式使用得越熟练，某些前提越可能退到视野之外。我们说一个物体保持匀速直线运动，是相对什么保持？我们说它受到相同作用却较难加速，这种响应关系又怎样从比较中取得确定含义？计算完成并不意味着这些问题已经被取消。

想象一辆平稳前进的车，车厢里有人把小球轻轻放在足够光滑的水平面上。若车辆突然加速，小球相对车厢向后移动。在地面近似惯性系里，可以解释为车厢向前加速，而小球的原运动状态暂时没有相同改变；在随车描述里，则需要引入相应的惯性力。两种说法可以相容，但它们已经说明，“是否加速”不能离开参考描述直接宣布。

这并不意味着一切受力现象都只是观察者的幻觉。乘客会感到座椅的推压，加速度计会给出响应，物体间的接触也会改变。问题是：哪些量取决于坐标表达，哪些可观察关系在变换中保持，哪些作用需要由更完整的物理模型说明？如果把这些层次混在一起，就容易在“全是物体固有属性”与“全是观察者任意选择”之间摇摆。

马赫对经典力学根基的追问，可以放在这个裂口处理解。本章不会把他写成已经解开惯性之谜的人，而是追踪一个更准确的问题：惯性能够在多大程度上被当作独立物体的自带性质，又在多大程度上依赖与其他物体、参考框架和整体物理结构的联系？一个有创造力的问题，未必已经拥有经验证实的唯一答案。

20.2 参考系换了，哪些东西没有换

先做一个纯运动学检验。在一维中，设新坐标原点相对旧坐标的位置为 $X(t)$，则同一物体的坐标满足 $x' = x - X(t)$。对时间求两次导数，得到：

$$a' = a - \ddot{X}。$$

如果新原点匀速运动，$\ddot{X}$为零，两边加速度相同；如果新原点加速，加速度描述就相差原点的加速度。这是明确的变换关系，不是对力学本体的最终解释。它告诉我们，不同坐标里的不同数值可以对应同一个过程，但也不能把两者未经转换就拿来比较。

在旧参考系中，若合力为 F 并有 $F = ma$，那么新参考系中可以写成 $ma' = F - m\ddot{X}$。后一项通常作为平动惯性力处理。它与真实接触力不必具有相同来源：车厢里的小球可

能没有受到向后的接触推力，但在加速车厢坐标里仍表现出向后的加速度。区分来源不妨碍计算，只是要求说明方程中每一项的资格。

给出一个自拟数值例。车厢向前加速度为每二次方秒二米，质量为一千克的小球在地面系中暂时不受水平合力。随车系中的加速度便为向后每二次方秒二米，相应惯性力项为向后两牛顿。若小球被座椅或挡板限制而随车一起加速，地面系中就必须有向前两牛顿的实际合力；随车系里，这个力与惯性力项相抵，小球保持相对静止。

两个情形的相对位置都可以被安排成某种简单样子，接触与传感器响应却不同。不能只因为一个观察者看见小球静止，就推断它没有受到实际作用；也不能只因另一个观察者看见小球运动，就认定它必有一股沿运动方向的推力。经典理论本身已经包含比直观的“动不动”更精细的关系区分。

20.3 一面弯曲水面能证明什么

旋转提供了更尖锐的例子。让一只圆筒中的水逐渐随容器稳定旋转，水面会形成中心低、边缘高的形状。如果水已经与筒壁共同转动，二者彼此可以近似静止，而水面仍然弯曲。由此能够排除一个过于简单的说法：所有相关效应只由水与紧邻筒壁的相对速度决定。

但从这个排除到“已经唯一证明绝对空间存在”，还有一个逻辑距离。实验给出水面、容器、地球以及更广参考关系之间的具体响应，哲学解释则要说明为何这些响应只能由某一种存在论承担。排除一个狭窄的相对运动模型，不等于排除所有可能的关系性理论。反过来，指出解释不唯一，也不等于已经建立了一个可以同样准确计算的新理论。

马赫在《力学》中批评把实验结论无条件推广到脱离其他物体的情境。他借旋转容器追问：若容器壁变得极厚、极重，是否仍能把通常实验所得关系原样外推？同时，他没有声称已经用实验知道任意改变整个宇宙后的结果。[20-1] 这一原典上的谨慎，比后来把他概括成一句“恒星制造惯性”更值得保留。

以下讨论是本书沿这个问题所做的分析。容器变重以后，可能改变的因素很多：支撑受力、局部引力、机械振动以及测量基准。若出现新的水面变化，不能立刻把它归因于一种宇宙性的惯性生成机制。新的解释必须说明如何区分这些竞争因素。一个反事实问题可以打开研究，但它需要经过模型和实验设计，才取得可判断的物理内容。

这正是哲学批判与物理重建之间的门槛。批判能够指出原结论超出了证据，重建则需要进一步给出在什么改变下出现什么新结果。只完成前一步，不能冒称已经完成后一步；只因后一步尚未完成，也不应否认前一步揭示了真实的推论缺口。

20.4 把旋转水面算出来

为了避免讨论只停在立场上，我们建立一个现代经典流体算例。假设水已经达到绕竖直轴的稳定刚体旋转，角速度为 ω，流体密度近似均匀，重力加速度 g 近似恒定；忽略表面张力造成的边缘修正以及显著波动。令 r 表示到转轴的距离，z 表示高度。在惯性描述中，径向压力梯度提供向心加速度，竖直压力梯度平衡重力。

沿自由表面，压力近似相同，因此径向与竖直的压力变化须相互配合，得到 g dz/dr ＝ω²r。积分后，水面相对于轴心的高度差为：

z(r) − z(0) = ω²r²/(2g)。

这个结果是明确的抛物面关系。它并未回答惯性最终从何而来，却允许我们检查经典模型在给定装置中的预测。把解释层次分开，讨论便不会因为终极问题尚未解决而失去已有计算能力。

设圆筒半径为零点一米，角速度为每秒两弧度，g 取每二次方秒九点八米，则筒边相对轴心的水面高度差约为零点零零二零四米，即二点零四毫米。若角速度增为每秒四弧度，高度差增为约八点一六毫米，是原来的四倍。这些数值均由本章假设直接计算，不是马赫做过的历史实验数据。

在随水旋转的参考描述里，可以把径向惯性项与重力共同组成有效作用，再得到同一个表面条件。两种计算的相容表明坐标变换受到规则约束，却不构成两份独立实验支持。它们共享相同的经典物理结构。只有真正改变角速度、半径或其他条件并取得新记录，才会增加对该模型的经验检验。

这也给出可败条件。若在排除显著非稳态、毛细效应和装置倾斜后，高度差不随角速度平方变化，就应调查模型的适用性，而不是只重复“离心力存在”。若某种关系性新理论要替代或扩展经典解释，至少应在通常条件下恢复这个成功结果，并在特殊条件下给出可区分的额外预测。

20.5 把惯性写成关系，尚不等于解释来源

考虑两个相互作用的小车。在外界作用可忽略的条件下，它们的动量变化大小相等、方向相反。若质量在过程中恒定，可以由加速度比比较惯性质量之比。例如，甲的

加速度大小为乙的两倍，就得到乙的质量为甲的两倍。这里的质量比较由相互作用中的响应关系建立，并不需要先假定每个质量都被某种直接感觉独立读出。

但这种操作性关系尚未说明惯性为什么具有这个来源。用加速度比定义或测定质量比，与提出全部惯性由遥远物质产生，是两类不同命题。前者给出比较程序，后者给出因果或构成解释。若把“测量离不开其他物体”直接推成“惯性完全由其他物体制造”，中间就省略了必须论证的桥梁。

也可以从另一方向犯错。一个局部物体在许多实验中表现出稳定质量，并不单独证明质量在任何可能世界中都独立于其他存在。经验稳定可以支持有效封装，但它不自动回答把全部周围物质移走后的反事实。前提未被实际改变，便不能把没有观察到改变当作完全独立性的证明。

因此，本书把问题拆成三个可区分层次：怎样规定惯性运动，怎样比较惯性响应，怎样解释这种响应的成立。第一层涉及参考结构，第二层涉及具体量度，第三层涉及理论的生成机制。它们互相联系，却没有一条可以靠改写术语直接吞并另外两条。马赫问题的价值，恰在于迫使经典封装重新面对第三层；它的困难，则在于第三层需要比批判更具体的证据。

如果有人主张惯性质量与某个宇宙整体量成正比，就应说明那个整体量是什么、如何定义边界、怎样处理不同运动状态、如何预测局部质量变化。只写“所有质量的作用总和”，仍然没有给出可计算的量。全体一词可以表达研究方向，却不能代替一个收敛、可辨识而且可检验的模型。

20.6 爱因斯坦怎样接过问题，又留下限制

爱因斯坦在一九一八年的文章中使用“马赫原理”这一名称，把它解释为对惯性应追溯到物体相互作用这一要求的推广。[20-2] 这是一位后继研究者对问题的重新表述，不应把它与马赫原书的全部论证视为同一段话，更不能因为名称相同就假定所有后来的版本都提出完全相同的命题。

《相对论的意义》一九二二年版相关段落又显示，爱因斯坦在当时把惯性部分依赖质量之间作用看作对马赫思想的支持，同时将更强的整体要求与特定宇宙构想联系起来。[20-3] 读这类材料，必须保留它的历史条件：它展示思想如何进入研究，而不是向今天提供一条不受版本和模型限制的最终定理。

本书在这里区分三个容易被混同的陈述。第一，马赫的批判启发了爱因斯坦。第二，某种引力理论能够预测旋转质量对局部参考结构的影响。第三，宇宙全部物质唯一

决定一切惯性性质。第一项是思想史命题，第二项是需要给出模型和量值的物理命题，第三项是更强的总体要求。即使前两项成立，也不能不经进一步论证就断言第三项已经证实。

这与上一章形式等价的纪律相同。把一个方程写成更一般的坐标形式，不等于已经解释所有参考结构的物理来源；赋予背景以动力学内容，也不意味着所有边界和初始数据都从此不再需要。方程、解的条件和可观测预测必须分别说明。若把一般协变性直接翻译为“一切都由物质关系唯一决定”，就把数学表达性质与物理决定关系混在了一起。

我们可以用一个更简单的数学类比检查这种推理。在一个没有质量源的有限区域内，牛顿势可以满足拉普拉斯方程；但不同边界条件允许不同的调和势。区域内部没有源，不等于区域中的场必为零，也不等于仅凭局部源信息就唯一确定场。这不是广义相对论的证明，只是提醒我们，微分方程里出现源项，并不足以宣布源项独自给出全部解。

20.7 可测的拖曳不等于整个马赫命题获证

把讨论推进到可测命题，必须给出具体预测。引力探测器 B 的研究团队在二〇一一年报告了地球轨道陀螺仪的测量结果，其中参考系拖曳漂移率为每年负三十七点二角毫秒，不确定度为每年七点二角毫秒；对应广义相对论预测为每年负三十九点二角毫秒。[20-4] 这里引用一项有明确实验对象、量值和不确定性的原始研究，不把它泛化为终极来源问题的完成。

从这组数值可以作一个简单核对：测得中心值与理论值相差每年两角毫秒，约为所报告不确定度的零点二八倍。这个比较说明二者在该误差尺度下相容；它不独自证明理论的每一种可能扩展，也不能替代论文对系统误差和数据处理的完整论证。数值相容是一项有限而有力的结果，力量恰恰来自它的有限规定。

为什么这不等于“恒星创造了全部惯性”已经实验获证？因为实验检验的是某个具体引力理论在地球周围对陀螺仪漂移的预测，并没有把宇宙中的其他物质全部移走再观察惯性是否消失。若不同总体解释在该实验条件下给出同样结果，这项实验就不能区分它们。支持一个具体效应，不等于支持所有借它讲述的哲学故事。

项目研究者的官方说明也明确提醒，不应把这类实验直接当作笼统马赫原理的检验，而应看作对具体引力理论及其明确特征的检验。[20-5] 这一点并不使马赫追问失去

价值。它把追问从口号转入研究的必要条件说清楚了：必须把“关系决定惯性”变成一组能够失败的量化承诺。

一个新的模型若预测额外漂移，应给出方向、大小和随装置变化的尺度关系，并与已知机械、电磁及引力效应区分。若每次没有看到预期结果，都把作用推到尚不可达的宇宙尺度，却不再提供可区分后果，它就退回研究愿景。愿景可以保留，但不能与已通过检验的命题使用同一种证据语气。

20.8 整体质量分布不是可以随便补列的环境

在 SDE 语言里，马赫问题很容易被误说成“要考虑环境”。这种说法太弱，也太宽。把地球、恒星、星系写在模型外面，并不会自动解释惯性；把 E 当作容纳一切遗漏因素的袋子，只会让理论失去边界。需要追问的是，参考结构、响应量度和质量分布之间怎样相互规定，哪些关系能够进入具体实验。

例如，在普通旋转水面实验中，地面参考、重力方向与装置轴线之间的关系参与水面形状的测量；筒壁通过接触把角动量传给水，也参与稳定状态的形成。这些都可以逐项改变并检验。至于远处质量是否对某项局部惯性参数具有额外作用，则需要新的动力学关系，不能因远处物质同样属于环境而自动归入已经验证的解释。

本书的本体论立场要求把物体的稳定属性放回复合发生系统，但不因此预先替任何特定马赫式物理模型宣布胜利。属性不是脱离互动便能完整定义的孤立成品，与某个具体宇宙积分式就是正确引力定律，是两个不同层次的主张。前者提供研究取向，后者必须接受数学和实验的独立检验。

这一区分尤其重要，因为整体性语言很容易获得一种不应有的免检地位。一个模型声称考虑了全部宇宙，不会因此比局部模型自动更深刻。若它不能说明如何恢复已验证的局部极限，或不能定义所谓全部的计算方式，它反而可能比局部模型更模糊。整体性需要建立层级之间的可追踪关系，而不是把所有层级都写进同一个句子。

20.9 让批判经过两条完整路径

本章可以先走 S—E—D 的进入路径。从旋转水面、车厢响应等稳定显露 S 进入，审查其中默认的参考结构与惯性关系 E；再设计能够区分解释的差异序列 D，例如改变角速度、交换观测框架、控制容器质量和排除附加机械影响。若新的操作只重复旧条件，便没有获得判断强反事实命题所需的新证据。

另一条是 D—S—E 的原型路径。先建立一个可控模型或实验 D，观察其与已知经典预测的差别 S，再判断这种差别是否足以支持新的关系结构 E。它要求研究者先把可测后果做出来，避免从一个令人兴奋的整体解释直接跳到结论。原型可以失败；失败之后保存哪些条件已经受检，比把愿景改写得永不失败更有知识价值。

三方程在这个过程里承担不同约束。S＝F(D,E)要求现象有实际发生条件；D＝G(S,E)要求已有结果决定下一步怎样区分；E＝H(S,D)则要求新关系回应已取得的记录。三原理的两维张力推动第三维改造：S 与 D 的失配可推动重审参考假设 E；S 与 E 的失配可推动新实验 D；D 与 E 无法支持旧描述时，可推动重新规定需要测量的 S。改变以后都要回到原先现象，不得只在抽象层面宣布已经超越。

可败性在这里不是附加的审稿手续。若一个惯性来源模型预测局部响应随某项可控质量配置变化，那么在约定灵敏度内没有出现相应变化，就对这个版本构成限制。若模型只解释既有结果而不给额外可观察区别，它仍可作为解释候选，但不能冒称比竞争解释获得更多实验证实。证据的差别必须写进结论，而不是藏在脚注里。

20.10 惯性问题怎样打开新的发生空间

马赫没有通过一句批评废除经典力学。经典方程在明确条件下仍然成功，旋转水面依然可以计算，车厢中的相对加速度也依然受变换规则约束。他所打开的是另一种问题空间：最稳定的基本属性，是否也包含尚未被充分展开的构成关系？当研究者把某种稳定封装视为不需解释的最后实体时，怎样重新追问它的成立条件？

但这种追问同样要防止自己的过度。不能因为惯性需要关系性量度，就断言全部来源问题已经解决；不能因为爱因斯坦受过启发，就把启发升级成物理定理；不能因为发现参考系拖曳，就宣布所有马赫原理版本全部获证。思想史、方法论、数学模型和实验结果，各自有自己的证据责任。

对于本书的主轴，最有价值的收获是保持两种能力同时存在。一种能力使成功的物理封装稳定运行，让人们能够继续建造、测量和预测；另一种能力把封装重新打开，追问其中哪些性质由更大系统的特征纠缠维持。前者没有后者，容易把条件性稳定误作无条件自足；后者没有前者，则容易只剩没有计算和检验能力的宏大语言。

到这里，经典物理已经显出它的创造力与边界：它重建了物体、运动、公共量度和数学形式，也留下了整体关系与局部属性之间尚待展开的张力。下一编转向热与统计，不是简单添加一种新力，而是观察另一种物理世界怎样发生：当无数微观差异持续变化，宏观稳定、转换方向和分布类型又怎样取得自己的存在规定。

本章参考资料

[20-1] Ernst Mach，The Science of Mechanics，T. J. McCormack 英译，1893，第二章第六节，重点原印页 229—238。https://pages.jh.edu/rrynasi1/spacetime/eprints/Mach1893NewtonsViewsOfTimeSpace%2BMotion.pdf。

[20-2] Albert Einstein，Prinzipielles zur allgemeinen Relativitätstheorie，Annalen der Physik 55（1918），241—244。doi:10.1002/andp.19183600402。https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19183600402。本章仅据命名说明讨论“马赫原理”与原作者思想的区别。

[20-3] Albert Einstein，The Meaning of Relativity，1922，原印页 119。https://en.wikisource.org/wiki/Page:The_Meaning_of_Relativity_-_Albert_Einstein_(1922).djvu/131。

[20-4] C. W. F. Everitt 等，Gravity Probe B: Final Results of a Space Experiment to Test General Relativity，Physical Review Letters 106（2011），221101。doi:10.1103/PhysRevLett.106.221101。https://arxiv.org/abs/1105.3456。

[20-5] James Overduin，Spacetime Before Einstein，Gravity Probe B 项目说明，2007，Mach’s Principle 节。https://einstein.stanford.edu/SPACETIME/spacetime1.html。

第五编 热、时间与概率：统计物理的新物质规定

第二十一章 蒸汽机与卡诺
热为什么成为一门新物理

21.1 炉火没有单独推动一座矿山

设想一台正在抽水的蒸汽机。炉膛发亮，锅炉里的水沸腾，活塞往复，泵杆把水从深处送上来。若问动力从哪里来，指着炉火回答似乎已经足够。但把锅炉、汽缸和排汽端全部置于同样高的温度，让原有冷却过程停止，机器还能按照原来的方式持续工作吗？这个反问把我们从一种实体直觉中拉开：热的作用不能只靠某个地方有多热来规定，还必须考察热如何进入、通过什么过程发生变化，以及装置怎样重新获得进行下一轮操作的条件。

矿工需要的不是一次壮观的活塞冲程，而是能够重复的抽水；经营者关心的不是火焰颜色，而是一定燃料支出可以支持多少生产；制造者面对的则是密封、摩擦、材料、压力和冷却之间互相牵制的要求。在这里，理论尚未建立，问题却已经由连续工作的技术系统提出。一次推动能够被观察，一轮循环是否成立却需要更长的差异序列。机器若不能回到可再工作的状态，前一个冲程的成功便不能直接外推为持续动力。

这些要求不只是科学研究的社会背景。它们参与规定被研究的物理量：必须先规定一次完整操作在哪里开始、在哪里结束，才能谈每轮输出；必须规定输入与输出穿越哪一道边界，才能比较机器；必须把抬升水的高度与水量联系起来，才有可能把不同装置的效果放到同一尺度上。热学的新对象，正是在这些互相校正的操作中获得稳定轮廓。

本书所谓创造新的物理世界，在此有很具体的含义。火仍然燃烧，金属仍然膨胀，但炉火、温度指示、压力变化、活塞位移和负载提升被组织成新的特征纠缠聚合体。原来分散的冷热感觉、工匠经验和运动测量，开始共同规定一个能够计算、比较和反驳的热机世界。这不是说理论凭空制造燃料和金属，而是说“热机效率”这种可检验的物理存在，不能脱离其发生系统，预先放在某一块材料内部等待取出。

21.2 卡诺提出的不是另一张改装图纸

卡诺于 1824 年出版《论火的动力及发展这种动力的机器》。原文将动力问题放在热由较热物体向较冷物体转移的过程中讨论，并试图寻找不限于某一种工质的原则。当时

的论证仍使用热质观念，把热质的转移而非净消耗作为基本图景。这一点必须保留：不能把后来热力学的能量守恒写回 1824 年，仿佛卡诺已用现代第一定律完成推导。[21-1]

他的提问之所以重要，不是因为它立刻解决了所有热机设计，而是因为它改变了问题的尺度。改进一台机器，可以减少某个零件的损失；追问任何机器能够达到的界限，则必须找到跨越零件和工质差异仍能成立的比较办法。后一问题不能由搜集更多成功机器自动回答。即使所有已造机器都很低效，也不能由此证明将来的机器无法超过它们；即使某种工质表现优异，也不能直接断定优异来自它独有的物质本性。

因此需要一种理想操作。理想化的任务并非把现实中不喜欢的细节全部删除，而是控制删除的后果：哪些影响可以作为损失暂时排除，哪些条件一旦删除，研究对象便不再是同一种热机？摩擦可以在极限模型中趋于零，冷热两端的差异却不能随便抹去。否则得到的不是更完善的同一问题，而是失去原有发生条件的另一问题。

这里出现了属性纠缠的一个关键区分。工质的某些具体属性影响装置能否制造、怎样运行和达到多大的功率；在严格限定的可逆循环比较中，最高效率却具有跨工质的一般性。一般性并不证明材料无关紧要，而是说明我们经过限定和比较，构造出了对某些材料差异不敏感的关系。所谓普遍规律，因而不是把所有条件擦掉后的空白，而是在明确条件内保持稳定的结构。

21.3 循环使“同样的结果”变得可判定

让气体膨胀推动活塞，很容易产生一种印象：只要不断让它膨胀，便能不断得到功。但有限容器不能无限扩大。若要重复操作，气体必须被压回去；压回去本身又需要功。热机问题由此变成一个成对的问题：怎样让输出阶段得到的功，多于恢复阶段必须付出的功？仅描述膨胀的前半段，会把恢复条件偷偷交给没有进入计算的外界。

循环的意义正是把这种偷渡堵住。工作物质一轮以后回到相同热力学状态，使“这台机器靠消耗自身储备工作”的可能性与“它靠持续交换工作”的可能性能够区分。需要注意，回到相同状态并不等于世界没有发生变化：热源、冷源和负载可能已经改变。对循环的定义必须同时说明，哪些部分被要求复原，哪些部分保留作为净效果。

用现代符号表示，在只交换热与功、忽略整体动能和势能变化的封闭工作系统中，取吸热为正、系统对外做功为正，有

$$dU = \delta Q - \delta W。$$

对完整循环积分，内能回到原值，于是净功等于净吸热。若每轮从高温热库吸热量为正值 $Q_{热}$，向低温热库排热量的正值为 $Q_{冷}$，则

W = Q 热 − Q 冷，η = W/Q 热 = 1 − Q 冷/Q 热。

这是现代热力学表达，不是卡诺原文的公式。符号上的小区别尤其不能忽略：dU 表示状态函数的微分，δQ 与 δW 提醒我们，热和功由过程规定。回到原状态可以使内能变化为零，却不能使一圈交换的热与功各自为零。[21-2]

从发生学看，循环不是给既有实体加一个方便的计算括号。它建立了物理同一性的操作要求：相同状态如何重建，输入如何比较，净效果如何保留。微积分在这里首先要回答变量与路径的资格问题，而不是急着把一个函数积分。只有这些资格明确，闭合积分才具有物理含义。

21.4 可逆性检验的是整个过程能否撤回

可逆过程常被误读为“让机器反着转”。然而，把电动机接上轴，使活塞倒转，并不能自动消除先前的摩擦发热；把气体重新压缩，也不能保证原来热库的变化已经恢复。可逆性要求的严格性超出了运动方向：系统以及相关外界必须能够沿逆过程恢复，而不能留下另一项未偿还的变化。它是一种全局比较要求，不是齿轮形状的属性。

现代卡诺循环由两段可逆等温过程和两段可逆绝热过程组成。等温段让工作物质分别与两个热库交换热，绝热段则在不交换热的条件下改变其温度。对于这两个恒温热库之间的可逆热机，热量比由绝对温度比规定，从而有

Q 冷/Q 热 = T 冷/T 热，ηC = 1 − T 冷/T 热，T 热 > T 冷 > 0。

这里的温度必须是热力学绝对温度。式子适用于上述两热库可逆循环，不能把任意燃烧温度与室温代入，就宣称已经得到真实发电系统的效率。[21-3]

我们可以据此做一个独立的数值检验。设高温热库为六百开尔文，低温热库为三百开尔文，则理想上限为二分之一。若把两个热库温度都加上三百开尔文，温差仍是三百开尔文，上限却变成三分之一。这个简单结果揭示了“温差越大，效率越高”的口语表达为何不够严格：若不说明哪一个温度保持固定，仅凭温差无法唯一确定上限。

再设一轮吸热一千焦耳，在前一种温度配置下，可逆热机输出五百焦耳，并排出五百焦耳。实际机器若输出四百焦耳，并不能单从这两个数字判定是哪一个零件有问题。我们还要查测量边界、散热、摩擦、压降和交换过程。上限为判断提供方向，但不会代替诊断发生机制。由总结果退回过程，仍然需要新的实验与模型。

可逆极限还有另一种容易被忽略的代价。在常见有限传热系数的装置中，为使传热温差趋于零，完成有限热量交换所需时间会增长。极高效率并不自动等于很高功率。若

使用极限模型评价真实机器，就必须把每轮收益与单位时间收益区分开。所谓最好的机器，取决于需要实现的任务及约束；热效率上限不能独占全部工程价值。

21.5 压力—体积图中的路径不是装饰

在机械准静态、活塞边界无摩擦耗散且只考虑体积功的条件下，系统对外微元功可写为 $\delta W = p\,dV$，循环功为 $\oint p\,dV$。在压力—体积图上，闭合路径的有向面积对应净体积功。若存在有限压差，边界功则应根据外压和实际边界运动处理，不能不加区分地使用工作物质的平衡压强。微元表达的物理意义来自界面上力与位移的规定。[21-2]

这个关系允许我们自建一个简化比较模型：在较高外压下膨胀，在较低外压下压缩，若两段体积变化幅度相同，得到的净功为压强差乘以体积变化量。如果以定容段连接这两段操作，定容段不做体积功。它不是卡诺循环，也不足以确定吸热与效率；它仅显示为什么路径安排能够决定净机械效果。如果交换高低压段的方向，净功符号便改变。相同的两个体积端点，不能决定整个过程的功。

由此再看“热是物体具有多少东西”的说法，就会发现其解释能力不足。即使工作物质开始和结束的温度、体积相同，不同过程安排仍可能交换不同数量的热。需要区分物体的状态与穿越系统边界的过程量。若把这两层混为一谈，便会错误地寻找一种“物体内部的热量”，期望仅凭它解释所有热机表现。

积分也没有独自创造这一认识。曲线是否能够被视为一串近似平衡态，压力如何测得，体积边界是否明确，摩擦如何单独处理，都是积分成立以前和以后必须检查的内容。形式计算把已经建立的局部关系组织成整体结果；当测量过程破坏局部关系时，正确的积分步骤也可能导出错误的物理结论。

在视觉 SDE 中，坐标曲线使各阶段差异可以同时比较；在触觉 SDE 中，阻力、推压与热接触参与操作判断；在听觉 SDE 中，运行声响可能成为异常监测的线索。三者各有完整的发生结构，不能分别贴成 S、D、E 三个零件。它们通过仪器、训练和校准形成相互约束，而不是凭感觉相似便获得物理等价。

21.6 三方程怎样进入热机研究而不成为套话

把输出功、测得效率等稳定显露记入 S，不能因此把 S 等同于仪表读数；显露还包括这些读数在重复操作中的可辨认性。D 涉及升温、膨胀、换热、压缩等差异序列的形成与运行。E 则不是“社会、燃料、气候”等因素的随意罗列，而是温度、压强、体积、边界交换及测量尺度怎样互相规定的特征纠缠。一个量改变时，哪些关系必须随之调整，才是识别 E 的要点。

在这一限定下，S = F(D,E) 提醒研究者，效率显露由操作序列及其特征纠缠共同发生；D = G(S,E) 要求根据输出偏差与可成立的关系重新设计操作；E = H(S,D) 则把结果与过程用于检验原有关系组合是否需要改写。这些方程在本章是方法论约束，不是已经拟合完成的热机数值模型。不能把未知函数写在纸上，就声称得到了超越热力学的新预测。

六路径把不同起点组织为完整的研究遍历。E—D—S 的沉淀路径，从已有特征纠缠设计操作并检验输出；E—S—D 的锚定路径，从关系与目标显露重建过程；D—E—S 的倒逼路径，让运行异常推动关系调整，再检查结果；D—S—E 的原型路径，先形成可运行装置和结果，再辨认其纠缠结构；S—E—D 的进入路径，从已显露现象进入关系并设计后续操作；S—D—E 的再创路径，从结果改变操作，再重建关系。这六条路径不能只列名称：同一热机问题应允许不同路径互相校验，防止只沿一种方向保护既定解释。

设一项实验报告效率高于所给两热库温度的卡诺上限。第一反应不应是立刻宣布第二定律失效，也不能靠一句“不可能”关闭研究。我们需要询问：循环真的闭合了吗？是否消耗了未计入的化学储备？冷端和热端是否真是所报告的两个恒温热库？吸热测量有没有漏掉通道？这些检查不是为旧理论辩护的任意补丁，而是把公式适用条件逐项转化为可观察的反驳点。

三原理在此落实为两维张力推动第三维改变，再由改变后的第三维回检原来两维。D 与 E 不能支持预期输出时，需要重审 S 的显露条件；S 与 E 持续不相容时，需要改变 D 的操作组织；S 与已经核实的 D 仍不能协调时，则须重建 E 的特征纠缠。每次改变都要返回原来问题，检验它是否真正消除了张力，是否造成新的冲突。三方程提出互相生成的约束，六路径安排不同起点的完整遍历，三原理说明张力如何推动研究发生。缺少回检，新方案就可能只是把一个困难搬到另一个位置。

21.7 新物理从旧理论中保留了什么

卡诺热质图景的局限，使这一历史不能被写成某位天才一次说对全部答案。关于类比在热力学形成中的作用，科学史研究也提示：旧有图景可能帮助组织有效论证，同时把错误承诺带入其中。因此应分别考察保留下来的结构、被放弃的物质解释和后来重建的证明。[21-4]

从本章的 SDE 解释出发，理论转变值得追问的不是一句抽象的“旧理论错了”，而是原来哪些差异比较继续有效，哪些特征纠缠必须拆开重建。热质不被消耗的假设不能进入现代能量平衡；循环复原、冷热差异和可逆比较所打开的研究方向，却可以在新的理论结构中重新获得位置。保留并不是原封不动，重建也不是把此前工作全部清空。

竞争解释仍然需要认真对待。较弱的认识论解释可以说，卡诺只是给先在热世界提供更好的描述；本书则强调，热机、效率和可逆极限作为可操作、可计算、可公共检验的物理存在，依赖新发生系统的建立。双方都能承认实验事实，真正分歧在于这些物理对象的规定性被放在哪里。单个效率实验并不能独自裁决这一本体论分歧；本书的论证责任，是展示新对象的成立条件，并说明整体解释比孤立实体解释多解决了什么问题。

卡诺留给后续章节的，因而不是一台永远最高效的现实机器，而是一种更严格的问题形式：当工作物质复原时，世界其余部分究竟改变了什么？怎样比较不同过程的净效果？什么限制来自可修复的装置损失，什么限制属于热过程的组织本身？下一步，热与机械功的定量联系将使这些问题进入新的结构。热不再只是冷热程度的名称，而成为必须通过交换、转换和方向性共同规定的物理领域。

本章参考资料

[21-1] Sadi Carnot，Réflexions sur la puissance motrice du feu，1824；本章核对 1897 年英译本，正文原印页 44—49 附近。https://www.gutenberg.org/files/78610/78610-h/78610-h.htm。须与同页编者前言、汤姆孙论文及遗稿区分。

[21-2] MIT Unified Engineering，The First Law Applied to Engineering Cycles 及相关热力学课程。https://web.mit.edu/16.unified/www/FALL/thermodynamics/notes/node21.html。本章方程采用现代约定，自拟算例不作为历史事实。

[21-3] MIT Unified Engineering，3.3 The Carnot Cycle。https://web.mit.edu/16.unified/www/FALL/thermodynamics/notes/node24.html。

[21-4] John D. Norton，How analogy helped create the new science of thermodynamics，作者机构存档。https://sites.pitt.edu/~jdnorton/papers/Analogy_Carnot_online.pdf。本章仅用于提示类比与理论替换问题，不据此复述详细历史。

第二十二章 能量与熵：守恒之外为何仍有方向

22.1 一笔相等的能量，为什么不能把事情倒过来

把一杯温水放在较冷的房间里，过一段时间，杯中的温度降低，房间从中得到能量。若把涉及的能量全部核对，没有任何东西凭空消失。然而，仅凭这个核对结果，我们还不能说明为什么同一杯水不会自动吸取周围较冷空气的能量，逐渐变得更热。两种设想都可以写成一方失去多少、另一方得到多少，方向上的不同却没有被等量关系识别。

这正是热学必须越过的一道门槛。守恒规定变化不能任意多出或少掉，却不保证满足守恒的变化都可以自行发生。好比地图上两点之间存在连线，并不保证每种车辆都能沿线通行；但热力学的限制比这一比喻更严格，因为它要求给出可测量的判据，而不是泛泛诉诸某种阻力。理论必须说明同一能量在不同配置中为何具有不同的转换可能。

上一章研究了循环如何建立热机的比较尺度。本章把问题推进一步：如果热与功具有定量联系，是否意味着它们在所有过程中都完全对称？如果不是，那么它们的不对称在哪里取得物理规定？新概念必须既保留等量转换，又解释方向限制，不能用其中一条原则取消另一条。

从本书的主轴看，这不是为现成热世界再添加一条口头规则。内能、热、功、熵以及边界交换共同形成的新结构，改变了什么过程可以被认为相同、什么装置被判为可能、什么变化需要额外补偿。热的物理世界因此从单纯冷热与运动的联系，发展为具有状态、路径和方向三重约束的发生系统。

22.2 焦耳的实验怎样建立跨过程的尺度

焦耳关于热功当量的工作并不是看到摩擦发热便完成了。摩擦会发热是一种经验，给机械效果和热效果建立可重复的数值联系则是另一种工作。他的《论热的机械当量》发表于 1850 年的《皇家学会哲学汇刊》；相关手稿在 1849 年提交。实验涉及以机械作用使流体升温，比较机械作用量与热效果。[22-1]

我们不妨设计一个说明性的测量方案：由砝码下降带动搅拌装置，记录质量、下降高度与温升。若直接以砝码减少的重力势能除以水的温升，就宣布获得热功关系，仍可能遗漏转轴摩擦、容器热容量、散热和开始结束时残留的转动。每一种遗漏都改变比较中的等量对象。跨领域的统一因此必须通过拆分和回收影响才能建立，不能由两个数看起来接近便成立。

更深的难题是测量的独立性。若我们已经用某个预期热功比例校准温度装置，再用该装置验证同一比例，结果就可能包含循环支持。可信实验应让机械量的标定、温度变化的测定和热容量的评估有可追踪的独立依据，并让不同实现方式相互比较。量值之间的联系越强，对这些来源的要求就越严格。

由此可以看见能量统一性的发生方式。不同操作不必共享同一种可见运动；它们可以通过对状态变化的稳定贡献而被纳入共同量度。这个共同尺度并不把温度计、砝码和搅拌器变成同一种器物，而是让其效果在特定条件下可以相互转换与核验。统一性的力量来自差异被连接后的可检验结果。

但必须限制结论：一种搅拌实验不能单独证明所有物理过程的能量守恒，更不能证明全部热能都能在循环中化为机械功。它建立的是更广泛理论所必须解释的一类定量证据。科学史若把漫长的尺度形成写成一次灵光，就会遗漏真正艰难的工作：如何使许多不同操作持续支持同一个跨过程约束。

22.3 克劳修斯和开尔文为什么需要第二个原则

开尔文 1851 年的《论热的动力理论》明确处理热功当量与卡诺结论之间的冲突：原有热质守恒假设不能继续保留，但可逆机器的比较原则需要重建。文中同时讨论了他所采用的限制与克劳修斯关于热不能无外界作用而由冷体传向热体的限制，指出二者可以互相联系。[22-2]

现代教材常把相应限制表述为：不存在只从单一热库吸热并全部转化为功、而不产生其他净效果的循环机器；也不存在唯一净效果是把热从低温物体传向高温物体的过程。这里最容易漏读的不是“不能”，而是“循环”和“唯一净效果”。气体一次等温膨胀可以吸热并做等量的功，但它的体积已经改变；冰箱可以把热从冷处送往热处，但它需要输入功。

这两个例子并没有反驳第二定律，因为它们恰好展示了被口语删掉的条件。若把一次膨胀接成循环，必须恢复工作物质；若让冰箱连续工作，就必须维持电力供应并处理排热。局部过程的自由常常依赖更大系统承担变化。正确的边界不是为了躲开反例而临时扩张，而是应当在实验前就清楚规定输入、输出以及复原要求。

可以构造一个思想检验：假设一台机器能够从单一热库吸热并在循环后全部输出为功，用这份功驱动普通制冷机。把两台装置按适当规模组合，便可能消去外界功输入，让净效果只剩冷向热的传热。反过来，若冷向热传热无需补偿，也能与热机组合出违反

前一限制的装置。这里的组合不是制造建议，而是显示两个禁止命题怎样约束同一整体。[22-2]

其方法论意义在于，科学不能只检查零件各自是否看似合理，还必须检查连接之后的净效果。一个局部承诺如果与另一套已经成立的过程组合后导致矛盾，就受到跨装置的约束。热力学的普遍性因此不只是普遍词语，而是过程之间能够共同存在的组织条件。

22.4 熵首先是一种严格的状态规定

克劳修斯在 1865 年的论文中提出熵这一名称，并把它与能量并置，讨论转换的方向性。原文的命名和论述属于当时理论的发展阶段，不能把所有现代统计含义直接塞入其中。[22-3] 本章采用现代符号 Σ 表示热力学熵，以免与 SDE 中的显露 S 混淆；常见教科书通常写作 S，符号改变不改变物理内容。

对于相邻平衡态，沿可逆路径的热交换满足 dΣ = δQ 可逆/T，其中 T 为绝对温度。对于两个平衡态 A 和 B，熵差由任一可逆连接路径的积分给出。式中的“可逆”不是排版修饰，而是计算资格。实际过程若不可逆，不能把它的热交换直接除以温度，就当成全部熵变。熵是状态量，热是路径量，二者之间由特定过程条件建立联系。[22-4]

一个看似矛盾的问题随之出现：真实过程不可逆，为什么可以用不存在于实际历史中的可逆过程求结果？理由是我们求的是端点状态差，而不是伪造历史。只要状态函数已经建立，选择便于计算的连接路径不会改变其差值。就像测量高度差不必沿人实际走过的所有道路，替代路径用于计算状态差，并不宣称实际道路与它相同。

这个区分同样规定了替代的边界。真实过程的耗时、局部发热、瞬时温差和机械损失不能由这条替代路径还原。求出相同熵差，不等于获得相同过程。因此必须分别回答三个问题：端点改变了什么，过程中交换了什么，内部产生了多少熵。混在一起，任何一项都可能被解释成其他项的替身。

对只有热交换、没有物质交换的系统，可把熵平衡写为：系统熵变等于各边界热交换所携带的熵之和，加上内部熵产生量。边界温度变化时需要逐段积分；内部熵产生量非负。绝热系统没有热交换带来的熵输运，却仍可能有内部熵产生。故“绝热”与“等熵”不是同义词，可逆绝热过程才有熵不变这一结果。[22-4]

22.5 两个等量热交换，为何产生不同的熵变化

考虑两个足够大的恒温热库，一个为四百开尔文，另一个为三百开尔文。让一百二十焦耳的热量从高温端直接传到低温端，并忽略其他变化。高温端熵变为负零点三焦耳每开尔文，低温端熵变为正零点四焦耳每开尔文，合计增加零点一焦耳每开尔文。能量的代数和为零，熵变的和却为正。

这个自拟算例把守恒与方向分开显示。高温端失去的能量与低温端得到的能量相同，但同一热量除以不同绝对温度，得到的熵贡献不同。若设想没有任何其他效果而把同样的热量反向传递，系统总熵就会减少。第二定律所禁止的是这种无补偿的净效果，并不是禁止局部熵降低。

再考虑可逆制冷。冷端可以失去热量并降低熵，但输入功支持机器把更多的热排到高温端。冷端熵减与热端熵增在可逆极限中相抵；真实制冷还有额外熵产生。日常所谓“制造冷”，在完整过程里并不是消灭能量，也不是无代价地减少总熵，而是组织受约束的热功交换。

用 SDE 的语言说，冷热属性不能只归于两个孤立对象。传热方向的规定来自温度尺度、热量交换、边界划分和过程补偿之间的特征纠缠。换一个边界，某部分可以表现为熵下降；把供能装置和排热部分纳入相应边界，净结果又不同。边界选择参与对象的构成，但不能随意选择到使相互冲突的结果都成立，因为跨边界流量必须连续衔接。

这也解释了为什么“宇宙越来越无序”不足以替代计算。纸牌散乱、房间凌乱、图案复杂与热力学熵，没有一种无需说明模型便可通用的直接对应。房间看起来更整洁时，清洁设备可能耗电发热；一幅图看起来更复杂时，介质的热力学状态可能几乎未变。必须说明系统、状态变量与过程，才能谈相应熵变。

22.6 自由膨胀如何迫使我们放弃热量袋子的想象

设一个绝热刚性容器被隔板分成两部分，一边装有理想气体，另一边是真空。撤去隔板后，气体最终占据整个容器。对包括原来真空区域在内的整体边界，既没有热交换，也没有外边界体积功；若起止时没有宏观流动，内能保持不变。对固定物质量的理想气体，内能仅由温度决定，因此初末温度相同。

但初末状态并不相同：气体可占据的体积增加了。沿连接相同端点的可逆等温路径计算，得到 $\Delta\Sigma = nR\ln(V_2/V_1)$，其中 n 为物质的量，R 为摩尔气体常量。若体积加倍，熵变为 $nR\ln 2$。这个计算来自理想气体状态关系与熵的状态性质，不是说自由膨胀过程中气体实际吸收了 $nRT\ln 2$ 的热。[22-5]

反例的锋利之处在于：实际 Q 为零，而熵变大于零。若把熵定义成“吸收热量除以温度”，不注明可逆路径，便会立即给出错误答案。也不能把温度不变解释为全部物理属性不变；可用体积的改变已经重组了压强及其他可检验关系。一个显露保持，不等于整个发生系统保持。

可以进一步比较另一种操作：让同样的气体通过可逆等温膨胀到相同体积。在这一过程中，它从热库吸热并输出功，内能仍不变，熵变也相同。两条路径有相同端点，却有不同交换和不同熵产生。自由膨胀没有利用那部分可获得的功，相关差别无法由仅看端点的温度计独自揭示。

这里出现了物理学的新对象：过程的可利用性差别。它不是气体内部一块能够取出的特殊东西，而由初末状态、环境温度、允许的交换方式以及复原条件联合规定。
谈“能量品质”时必须给出这些条件，否则一个有用的工程概念很容易退化成对能量本身的价值拟人化。

22.7 微分形式为什么不能抹去实验条件

熵的引入使数学结构发生关键变化。热量的微元 δQ 一般不是状态函数的全微分；在可逆条件下，除以绝对温度得到 $d\Sigma$，才可以在平衡态空间中建立路径无关的差。这个积分因子的物理意义，来自温标、可逆过程和第二定律的共同支撑，而不只是代数技巧。

我们可以独立检查其一致性。对固定物质量的简单可压缩理想气体，若定容摩尔热容近似常数，有 $d\Sigma = nC_v\,dT/T + nR\,dV/V$。由此两状态间熵差为 $nC_v \ln(T_2/T_1) + nR \ln(V_2/V_1)$。温度和体积的比例都没有量纲；对数里不能随意放进一个带单位的体积数。[22-5]

假如两段实验先升温后膨胀，与先膨胀后升温到达同样的平衡态，按这个状态关系算出的熵差必须一致。若热量记录不一致，并不自动构成矛盾，因为两条实际路径交换热和功可以不同；若据全部可逆校正仍得到不同熵差，则要检查理想气体近似、热容变化、样品量或测量过程。数学不负责替实验选择一个舒服的结果，它负责明确哪些差异相容，哪些差异构成问题。

这一点与本书关于微积分的主张直接相接。局部形式并不自动拥有跨过程的意义。变量怎样获得定义，局部关系适用于哪些状态，整体路径是否可以连接端点，结果如何独立核验，必须同时回答。把微分理解为 D、把积分理解为 E，会跳过最需要解释的发生工作。

22.8 三方程如何处理守恒成立而过程失败

设研究者设计了一个装置，能量平衡完全正确，但预期的循环输出始终未出现。若坚持只检查能量是否漏算，就可能永远找不到问题：装置也许要求一个无补偿的冷向热交换。S 与 D、E 的张力在这里迫使研究从数量平衡进入方向约束。新的 E 不是补进“环境因素”，而是重建过程之间可以兼容的关系。

S = F(D,E)要求输出和方向显露由具体过程与纠缠共同支持；D = G(S,E)要求针对未实现的输出和方向限制重新组织操作；E = H(S,D)则把已核实的结果与过程带回理论，检查能量关系之外还缺少什么。热力学熵 Σ 在这些分析中是一种可测算的物理量，不能因为它在其他书中写作 S，就把它与显露 S 合并。

在这个研究里，更有效的是让两条完整路径互相约束。D—E—S 的倒逼路径从已经运行失败的装置出发：先明确每轮的热功交换，再发现所设想的冷向热传递缺少补偿，由此引入方向约束，最后检验新结构能否预言装置会停在哪里。S—E—D 的进入路径则从已经测得的冷热变化出发，进入状态与熵平衡，随后设计另一个有外界功输入的对照过程。前一路解释为什么旧方案失败，后一路检验补上条件后是否成功。若只解释失败而不能产生新的可比较结果，所谓重建仍可能只是事后修辞。

三原理约束推进是否发生：D 与 E 之间的张力推动重新判断 S；S 与 E 之间的张力推动改变 D；S 与 D 之间的张力推动重建 E。每次第三维改变以后，必须回检原来两维。加入熵后，不能只宣称解释更完整，还要重新检验已被第一定律支持的实验是否仍成立，新判据是否真正区分可行与不可行。

这一重建并不能凭本章例子独自证明 SDE 本体论。传统热力学也能在操作性语言中严密描述相同关系。本书提出的较强解释是：能量与熵的意义由完整发生结构取得，不能还原为独立物体早已装有的两种实体。该解释的论证责任，是持续展示它如何防止错误拆分，并帮助研究发现此前被隐藏的问题。

22.9 方向性不是宇宙命运的廉价答案

克劳修斯曾把能量与熵的讨论推向宇宙，并指出应用需要考虑全部相关情况。[22-3]今天引用这一历史表述时，不能把它当作已经解决引力系统、宇宙膨胀、视界与非平衡过程的完整定理。有限实验室系统的边界、平衡态和热库可以明确控制；宇宙整体能否沿用同一套简单条件，需要另外的物理研究。

同样，熵增不等于所有局部结构必须立刻瓦解。局部系统可以通过与外界交换而形成稳定结构，其熵变化须与外界及内部产生共同分析。把第二定律直接搬去证明生命、

组织、文化必然衰败，会混淆物理量与价值判断。方向限制是真实的，但它不替代每一种复杂系统的具体发生机制。

本章的新进展，因而是把守恒与方向放进一个可检验的物理世界：内能使不同交换获得共同尺度，熵使过程的兼容与补偿获得另一种尺度。接下来还剩一个更尖锐的问题：如果组成气体的微观运动在理想模型中可以反演，那么宏观方向性为什么如此稳定？这个问题将把物理学从单次轨迹推进到分布、概率和大量自由度的组织之中。

本章参考资料

[22-1] James Prescott Joule，On the Mechanical Equivalent of Heat，Philosophical Transactions，140（1850），61—82。皇家学会手稿目录，PT/37/3：https://makingscience.royalsociety.org/items/pt_37_3/paper-on-the-mechanical-equivalent-of-heat-by-james-p-prescott-joule。核验范围为手稿目录元数据；实验类型另核对 William Thomson 1851 年论文 § 4。

[22-2] William Thomson，On the Dynamical Theory of Heat，1851，原论文节选转录，重点 § § 6—14：https://zapatopi.net/kelvin/papers/on_the_dynamical_theory_of_heat.html。正文对第二定律的现代表述与独立组合说明，不作为原文引语。

[22-3] Rudolf Clausius，Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie，1865；Magie 编 Source Book in Physics 英译节选：https://webserver.lemoyne.edu/giunta/Clausius.html。

[22-4] MIT Unified Engineering，Concept and Statements of the Second Law 及熵课程：https://web.mit.edu/16.unified/OldFiles/www/FALL/thermodynamics/notes/node37.html；https://www.ocw.mit.edu/ans7870/16/16.unified/thermoF03/chapter_7.htm。

[22-5] MIT Unified Engineering，Entropy Changes in an Ideal Gas：https://web.mit.edu/16.unified/www/SPRING/propulsion/notes/node39.html。本章冷热交换和理想气体比较均为自拟说明，不冒充历史实验数值。

第二十三章 玻尔兹曼：微观运动怎样生成宏观物理

23.1 温度计没有报告任何一个分子的速度

温度计在容器中的气体里给出稳定读数，并不意味着每个分子都以相同速度运动。压力表保持在某个数值，也不意味着容器壁每一瞬间受到完全相同的冲击。宏观稳定与微观变化可以同时成立，这种共存要求物理学改变提问方式：我们需要知道什么在变化中保持，怎样保持，又在多大的尺度上保持。

若执意追踪所有分子的每次碰撞，问题的复杂程度会迅速超出计算能力。但统计物理的意义不能只用“算不动所以取平均”解释。即使能够获得全部轨迹，温度、压强等宏观量仍然需要另外的构成规则。轨迹列表本身不会自动告诉我们，哪些差异可以忽略，哪些差异对于某个实验至关重要，什么条件下不同轨迹代表同一种宏观状态。

玻尔兹曼的研究使这个问题获得新的形式。在 1877 年的论文中，他通过分配与计数研究热平衡的概率基础，把具体分配方式和按占有数归类的分布区别开来。本文核对的是 Sharp 与 Matschinsky 所译原典；原文的记号和现代教材并不完全相同，以下自拟模型使用现代记号。[23-1]

从本书的 SDE 视角看，宏观物理世界并不是把微观小球放大后的世界。一个分子没有整份气体的宏观压强；一个分子的瞬时运动也不独自具有整个样品的热平衡温度。宏观类型由分布、约束、交换与测量共同成立。这种发生不否认微观运动，而是要求说明微观差异怎样被组织为另一个层级的稳定显露。

23.2 四个粒子足以揭示计数单位的陷阱

先构造一个刻意简化的模型。四个带编号的经典粒子可以分别处在盒子的左半部或右半部，两半体积相等。暂时只记录每个粒子的左右位置，并假定十六种位置配置等概率。四个都在左边只有一种配置；三个在左边有四种；两个在左边有六种；一个在左边有四种；一个都不在左边也只有一种。

这里同时出现三个层次。具体指定甲乙在左、丙丁在右，是一个微观配置；只说左边两个、右边两个，是一个占有数类别；说左右密度接近相等，则是更粗的宏观描述。把这些层次混成同一个“状态”，就会产生错误概率。如果五个占有数类别被错误地当成等概率，均匀状态的优势便凭空消失。

一般地，N 个独立粒子中左边有 n 个的配置数为 $C(N,n) = N!/[n!(N-n)!]$；在上述假设下，该事件概率为 $C(N,n)/2^N$。这个结果来自明确的样本空间与等概率假设，不能

用“自然喜欢平均”替代。若左右体积不等，单个粒子落入左侧的概率就不再是二分之一，最常见的占有数也随之改变。

另一个限制同样重要：这个位置计数模型没有给出随时间运动的规律。它说明何种宏观类别包含更多配置，却没有证明任意初始状态将在多久以后进入那个类别。若中间有不可穿越的隔板，最初全在左侧的粒子就不能到右侧，样本空间中的许多配置根本不可达。统计数量要与动力学可达性相接，才成为物理解释。

由此我们可以精确定位“分布”的作用。它既不是人的随意命名，也不是每个粒子天生携带的宏观属性。体积划分、粒子运动、可穿越边界及采样方式共同规定计数类别。改变这些纠缠关系，原来的“均匀”便可能不再是同一种状态。所谓类型，从一开始就受到发生条件的约束。

23.3 为什么熵要与配置数的对数联系

采用现代玻尔兹曼形式，可以写 $\Sigma = k_B \ln\Omega$，其中 Σ 为熵，Ω 表示给定宏观约束下相应的微观状态数或适当无量纲化的相空间测度，k_B 为玻尔兹曼常量。连续经典相空间不能直接把一个带量纲体积放进对数，微观计数也必须处理不可分辨性等条件。这个简写不是对一切系统不加说明即可套用的公式。[23-2]

为什么不是直接用 Ω？可以从独立组合的要求进行一个说明性推导。两个相互作用可忽略、约束可以分别处理的子系统，若各有 $\Omega_{甲}$ 和 $\Omega_{乙}$ 种允许配置，则联合配置数为二者乘积。宏观熵在相应条件下要求可加，而对数把乘法变成加法：$\ln(\Omega_{甲}\,\Omega_{乙}) = \ln\Omega_{甲} + \ln\Omega_{乙}$。对数形式因此把组合结构与宏观可加性连接起来。

这个论证也显示它的限制。若两个子系统存在必须保留的关联，联合允许配置不一定是全部配对；若有长程相互作用，分开的能量和状态约束也可能不再简单相加。此时不能把独立组合的结论强行移植。统计解释要求对关联负责，不能把所有不方便的关联都叫作随机性而删除。

回到左右盒子模型，可以比较全部粒子在左与左右接近均分的类别。对偶数 N，前者配置数为一，后者为 $C(N,N/2)$。熵差由 $k_B \ln C(N,N/2)$给出；当 N 很大，组合数的增长使后者拥有压倒性的配置数量。然而“压倒性”是相对于所选择的宏观类别和测度，不是说每一个均分配置比每一个全左配置更有资格发生。

这一点对理解概率尤其重要。若微观配置等概率，某一个具体均分配置与那个具体全左配置的概率相同。均分宏观状态更常见，是因为它由更多微观实现方式组成。宏观

规律的稳定来自许多不同实现共同支持一个类别，而非每个微观实现都朝向同一个预定结果。

23.4 概率的稳定并不要求每一个成员稳定

仍用 N 个独立左右选择的模型，左侧粒子数的平均为 N/2，方差为 N/4，标准差为 √N/2。标准差相对于平均值的比例是 1/√N。粒子数增加时，绝对涨落可以增长，相对涨落却缩小。这两个变化不能混为一谈：更大的系统未必每一瞬间少变动几个粒子，却可以在密度比例上显得更加稳定。

设 N 为一百，相对标准差为十分之一；设 N 为一万，则为百分之一。这是自拟数字，用来展示尺度关系，而不是宣称现实气体所有涨落都服从独立二项分布。真实体系存在关联；临界点附近的长程关联尤其会改变朴素独立估计。模型的价值，在于把“规模增大为什么常伴随稳定”拆成可检验条件，而不是把大数当成万能魔法。

从测量方面看，还存在空间和时间的采样尺度。一个足够小、响应足够快的传感器可能看见明显脉动；一个覆盖更大区域、积分更长时间的装置则给出平滑曲线。二者未必互相矛盾。问题在于它们测到的是不是相同的可观察量，以及平均过程有没有越过我们真正关心的变化时间。

如果研究的是短脉冲中的局部状态，长时间平均可能掩盖关键过程；如果研究的是容器整体平衡性质，追踪每个冲击又会把稳定关系淹没在无关细节中。宏观层级不是任意粗糙化，而是需要找到足以保留目标现象的分辨率。对不同任务，合适的层级可以不同，但其适用性应由重复比较来判断。

本书把这种形成称为分布稳定组织。稳定的是某种统计关系以及它与测量的连接，组成者仍在交换位置和速度。这既不同于静止实体的同一，也不同于一切都不可把握的流变。它让物理学获得一类新对象：能够在成员不断变化时维持自身规定的宏观类型。

23.5 从约束计数到速度分布

为了理解能量分布，可以再建立一个现代化的简化计数模型。把单粒子可用微观状态按能量 ε_i 分类，相应的简并数为 g_i，令其中有 n_i 个粒子。固定总粒子数与总能量，则有各 n_i 之和为 N、各 $n_i\varepsilon_i$ 之和为 U。在经典稀薄近似下，相应计数含有 N!乘以各 $g_i^{n_i}/n_i!$的乘积。

取对数，并对较大的占有数使用斯特林近似，最大化计数时必须同时保持粒子数和能量约束。引入两个拉格朗日乘子后，对每个 n_i 求导，得到 n_i 正比于 $g_i \exp(-\beta\varepsilon_i)$。这

里的关键并非把“指数”记成经验曲线，而是看见它由计数结构与两个约束共同产生；β再通过热平衡关系与绝对温度联系。[23-3]

如果忘记 g_i，就会把能量区间中可容纳多少微观状态这一条件删掉。等宽能量区间未必有相同数量的速度状态，因此不能把粒子速度大小的分布直接写成单纯指数。在三维速度空间中，同一速度大小对应一个球壳，球壳测度随速度平方增长。这一几何因素与能量指数共同决定常见的速度分布形状。

从这一结构可以自行检查两个极端。速度为零附近的单位速度区间，球壳测度趋于零，因此速度大小分布不在零点达到最大；速度很大时，动能随速度平方增加，指数衰减压过球壳增长。分布由两种趋势共同形成一个中间峰值。不是某个分子“追求最概然速度”，而是大量可用状态与能量约束共同规定了整体比例。

这个推导仍然不是对气体每一步弛豫的证明。最可能分布告诉我们平衡候选的结构；碰撞是否使样品接近平衡、多久接近、有没有其他守恒量阻碍，都属于动力学问题。把最大计数直接说成所有初始条件的必然终点，会省去最困难的桥梁。静态分类与实际发生必须保持区别，随后才能真正连接。

23.6 压强怎样从碰撞成为宏观特征

考虑边长为 L 的立方容器，取一粒质量为 m、垂直于某壁的速度分量为 v_x 的经典粒子。理想弹性反射中，该粒子向壁传递的动量大小为 $2m|v_x|$；在不改变该分量的简单往返图景中，两次撞击同一面壁的时间间隔为 $2L/|v_x|$。二者相除，平均力贡献为 mv_x^2/L，再除以面积 L^2，得到压强贡献 mv_x^2/V。

把粒子贡献求和，在各向同性的平衡分布中，三个速度分量平方的平均相同，因此得到 $pV = Nm\langle v^2\rangle/3$。这是一个可逐项复算的理想模型说明：压强与平均速度本身无直接比例，真正进入的是速度平方的平均。往返方向的抵消可以使平均速度为零，却不会使壁面受压消失。

进一步在经典单原子理想气体的热平衡条件下，平均平动动能与温度联系，便得到熟悉的 $pV = Nk_BT$。[23-3] 这里没有哪个单独分子携带宏观压强。压强由粒子运动、边界冲量、面积以及时间平均共同规定。若边界会吸附粒子、发生非弹性过程或持续向体系输入动量，相应模型就必须重新处理。

这段计算也说明微观理论为何能够超越宏观拟合。仅靠压力与体积的经验关系，不容易知道一个新装置改变了哪些作用；碰撞模型则明确指出应检查分布、壁面和时间尺

度。反过来，若这个模型不断误判实际压强，宏观实验就会逼迫微观假设接受修改。两层不是上下级服从关系，而是通过可计算的桥梁相互检验。

对于 SDE 研究，这种桥梁比“宏观由微观构成”一句话更有力量。前者把构成写成可失败的具体过程，后者可能只是把问题换个尺度再说一遍。创造一个统计物理世界，就是让分布的类型、边界上的作用与仪器的显露组成新的可追踪结构。

23.7 时间反演为什么使 H 定理不能被当成无条件证明

玻尔兹曼动力学研究中的 H 定理，给出分布泛函的单调趋势，但它的适用条件必须保留。在现代稀薄气体推导中，碰撞前相关性的处理具有关键地位；分子混沌近似把适当的两粒子分布近似为单粒子分布乘积。洛施密特式的反演问题正迫使人们检查这些条件，不能说只从可逆力学便无条件证明了任意微观初态都熵增。[23-4]

为了看清逻辑，不妨做一个理想思想操作：记录一个体系从特殊初态向平衡演化的完整微观轨迹，然后在某一时刻将所有速度严格反向。若动力学具有相应反演对称性，新的轨迹可以沿原路返回。这并不意味着日常气体经常自动复原；它意味着“任何初态都严格单调趋向平衡”的命题过强。反演初态包含经过精确安排的关联，不能自动满足常用统计闭合条件。

可以把关键不等式单独拿出来看。对正数 a、b，因为对数单调递增，所以(a−b)(ln a−ln b)不小于零。将这样的结构按非负碰撞权重积分，可以得到相应 H 泛函变化的符号。代数没有错误，但 a 和 b 怎样由碰撞前后的分布组成、积分权重怎样定义，已使用动力学及统计条件。符号结论的严格，不会使此前假设自动变成无条件真理。

这为方法论提供了一个重要反例。若实验中出现预期外的短时逆向变化，不能立即把它当作违反微观力学，也不能只用“统计罕见”排除检查。需要先问测量的量是不是相应熵，体系是否足够大，初态是否带有特殊关联，所用近似是否仍然适用。异常究竟来自理论边界还是测量错误，只能通过新的差异运行来区分。

不可逆性因此需要分层理解：基本运动方程的反演性质、初始条件的特殊程度、统计描述保留的关联以及宏观观察尺度，各自承担不同责任。把它们压成“分子乱动，所以时间只能前进”，既没有解释为什么稳定，也没有解释哪些情况下稳定会失效。

23.8 热力学极限不是把有限世界删除

在适合的短程相互作用体系中，常考虑粒子数和体积共同增大、密度保持一定的热力学极限。这个极限有助于区分体积效应与边界效应，建立宏观量的稳定形式；它不等

于现实容器真的有无限粒子，也不保证所有长程相互作用体系都具有同样简单的极限。[23-2]

若仅有有限粒子，涨落便不能从定义中删除。观察时间延长，某些小概率事件可能被看到；但一个小概率并不能仅凭取倒数就变成可靠的等待时间，因为连续观测之间可能高度相关，动力学也决定采样速度。计数告诉我们某类状态在适当测度下有多稀少，发生频率还需要时间结构。

同样，在相关可达相空间具有有限不变测度、演化保持该测度等条件的理想系统中，复现问题限制了永久单调性的强断言。[23-5] 但接近某个过去微观状态的可能性，并不能直接否认在可观察时间内稳定成立的宏观热力学。需要比较的是尺度和条件，而不是让“可能复现”与“趋向平衡”作为两个不带范围的口号互相消灭。

一个有价值的研究设计应当同时报告平均曲线与涨落。若只给平均值，我们可能错过有限尺度效应；若只展示某一次涨落，又可能把稀有事件夸大为宏观规律已经失效。稳定性不是没有偏离，而是偏离的尺度、分布和可重复结构也可以被规定。这使统计物理比单纯确定曲线更丰富，而不是更不可靠。

从分布稳定组织的角度看，热力学极限是构造宏观类型的一种方法，有限系统实验则负责检查类型在何处近似成立。二者必须往返：理想极限提供清晰关系，有限尺度回检关系的边界。不能为了保护理想模型，把所有有限偏差都排除出物理学。

23.9 两条研究路径怎样把宏观稳定重新带回微观差异

本章最有效的一条路径是 D—S—E 的原型路径。先进行粒子运动与边界碰撞的模型操作 D，产生可比较的压力和密度显露 S，再从这些结果中辨认分布、平均尺度与约束组成的 E。若换一个时间窗便得到完全不同的宏观类型，就需要重新检查这种类型是否具有预期的稳定范围，而不能只选最漂亮的一条曲线。

另一条是 E—D—S 的沉淀路径。从已经构成的统计纠缠出发，明确独立近似、守恒条件与边界性质，推导碰撞或采样的差异序列，然后检验平均值及涨落。如果只检验平均压强，许多不同模型可能同样通过；加入分布宽度和时间关联后，它们的差异便可能显露。路径的完整性体现在新增可失败的预测，而非写完三个字母。

三方程在这里承担相互追问：S = F(D,E)防止把宏观稳定视为不需解释的事实；D = G(S,E)要求根据稳定与偏差调整操作和采样；E = H(S,D)则让可复现的偏差进入关系重建。三原理的张力机制可以落实得很具体：S 与既定 D 不协调时，推动重审独立性等 E

条件；S 与 E 不协调时，改变 D 的时间窗或实验设计；D 与 E 不协调时，则重新界定什么 S 有资格称为平衡显露。改变之后都必须返回原来的平均与涨落检验。

本书的解释并不是说观察者任意选择粒度就能创造任何气体性质。分类会受到碰撞动力学、守恒关系和多种测量的一致性约束；不能稳定复现的分类，没有同样的物理资格。物理世界的创造发生在这样的抵抗中：研究者建立一种新组织方式，而后让它接受超出自身设计目的的实验检验。

玻尔兹曼所打开的方向，使物理学能够认识一种新的存在方式：微观差异持续发生，宏观类型仍有稳定结构；方向性并不要求每条轨迹都被禁止反向，概率也不只是对无知的歉意。它们共同提出了更深的问题：规律究竟在哪一个层级、以怎样的约束和时间尺度成立？此后材料、辐射与量子理论将不断重写答案，但再也不能轻易退回只把世界看成少量孤立物体轨迹的图景。

本章参考资料

[23-1] Ludwig Boltzmann，1877 年论文，英译 Kim Sharp、Franz Matschinsky，Entropy 17（2015），1971—2009，doi:10.3390/e17041971。作者上传全文转录：https://www.researchgate.net/publication/275220813_Translation_of_Ludwig_Boltzmann%27s_Paper_On_the_Relationship_between_the_Second_Fundamental_Theorem_of_the_Mechanical_Theory_of_Heat_and_Probability_Calculations_Regarding_the_Conditions_for_Thermal_Eq。本章核对分配、分布与排列计数段落，非全文精读。

[23-2] David Tong，Statistical Physics，Chapter 1，特别是微正则系综、熵与涨落：https://www.damtp.cam.ac.uk/user/tong/statphys/statmechhtml/S1.html。

[23-3] David Tong，Statistical Physics，Chapter 2，Classical Gases：https://www.damtp.cam.ac.uk/user/tong/statphys/statmechhtml/S2.html。正文四粒子、二项涨落及容器碰撞推导为本书自拟说明。

[23-4] David Tong，Kinetic Theory，Chapter 2，§2.2—2.3，尤其原印页 22—23、34—35 的分子混沌与 H 定理：https://www.damtp.cam.ac.uk/user/tong/kintheory/two.pdf。

[23-5] Sheldon Goldstein，Boltzmann’s Approach to Statistical Mechanics，2001：https://arxiv.org/abs/cond-mat/0105242。作为关于典型性与反演/复现问题的研究对话入口，不将作者的全部争论立场视为定论。

第二十四章 吉布斯与系综：概率是否只是无知

24.1 一只盒子，为什么要想象许多个系统

实验室里只有一个密封容器，理论却要求想象许多同类容器：它们具有相同的组成规则和控制条件，内部的具体位置与速度可以不同。这是否意味着物理学不再研究现实，只在头脑里制造一批替身？系综方法提出的挑战正在这里。理论对象与手上的样品不是简单的一一对应，而需要经过制备、统计和测量的连接。

吉布斯在 1902 年的《统计力学基本原理》中，将关注点从追踪一个系统的状态历程转向研究同类系统在相空间中的分布，并系统讨论正则分布及其与热力学量的联系。[24-1] 这一转移没有取消运动方程，也没有假定实际房间里存在无限多个复制品。它建立了一种比较结构：哪些微观实现归入同一制备条件，给它们怎样的权重，依照这些权重能够预言什么。

若将一枚硬币遮住，我们不知道正反，这种概率可以反映信息不足。但温度固定时小系统能量的起伏，并不只是因为有人没有看清一项永远不变的数值。它可能与热库真实交换能量，因而在不同测量时刻取得不同结果。把两种情况都简化成“不知道”，便会抹去知识缺口与物理运行之间的区别。

另一方面，也不能反过来说所有概率分布都是客观地漂浮在物体内部的实体。选择变量、划定系统、指定约束以及决定采样方式，都参与分布的定义。物理问题不在于给概率贴上纯主观或纯客观的标签，而在于说明这个分布由哪些条件支持，它怎样与可重复的测量接通，何种结果会使它失去资格。

本章的核心命题是：系综创造了一类新的物理对象，却不能仅因它表现为整体分布，就把它等同于 SDE 的特征纠缠场。概率密度是经过变量选择和模型规定形成的数学对象；E 则要求解释相关特征怎样互相规定。两者可以发生联系，但建立联系仍然需要具体工作，不能借“场”这个共同字眼直接合并。

24.2 选哪一种系综，其实在问允许什么交换

一个系统总能量近似固定，与一个系统持续接触恒温热库，是两种不同的制备。前者可以使用适当的微正则描述，后者在满足弱耦合、平衡及热库足够大等条件时，可以使用正则描述。若粒子也可以交换，则需要另行处理粒子数及化学势。三个名称所标识的，首先是不同的约束结构，而不是研究者喜爱哪一种求和公式。[24-2]

以一小块样品接触恒温平台为例，平台并非只提供一个写在旁边的温度数值。它允许样品能量改变，并在适当尺度下维持自身温度近似不变。如果接触太弱、等待时间不足，样品未必达到预期平衡；如果接触强到界面能量无法忽略，简单划分样品与热库的方式也可能需要修正。热库进入概率权重的成立条件，不能退回一张背景清单。

这使系综选择可以被实验反驳。研究者假定粒子数固定，但装置存在吸附和脱附，则该假定可能漏掉一类涨落；假定能量固定，却持续受到外界驱动，则所得分布可能根本不是微正则平衡分布。即使一段平均曲线碰巧拟合得很好，也不能据此证明选择正确。应当寻找会把这些约束差异放大的观测量。

一个很小的模型已经足以说明。设两个可区分单元，各有零和 ε 两个能级，且总能量严格为 ε。允许的联合状态只有一低一高与一高一低，两个单元的能量完全反相关。如果分别只看任意一个单元，高能级占有概率都是二分之一；若误把它们当成两个互相独立、各自高低等概率的单元，则会额外允许总能量零和二 ε。

两种模型给每个单元相同的边缘概率，却给总能量不同的分布。这不是少收集几次单元读数就能自动解决的问题，而是必须改变观测对象，测量联合占有或总能量。概率分布怎样聚合，因而与哪些关系被保留直接有关。若把相关性全部压掉后再谈整体，就可能连最基本的守恒都丢失。

24.3 配分函数为什么不是计算中的垃圾桶

在离散能级、正则平衡且能谱不显含温度的模型中，用 W_i 表示物理能级以区别 SDE 中的 E 维。令 $\beta = 1/(kBT)$，微观状态 i 的概率写为 $p_i = \exp(-\beta W_i)/Z$，其中 $Z = \Sigma_i \exp(-\beta W_i)$。求和中的每一项代表一个被计入的微观状态，而不一定只是一个不同的能量数值；有简并时必须计入相应多重性。Z 保证概率和为一，同时把所有允许状态及其权重联系起来。[24-3]

把 Z 叫作归一化常数，没有错，却容易让人低估它承载的内容。若删掉一个可以被占据的状态，改变的并不只是那一项概率；分母改变以后，其他状态的概率也改变。某种局部性质进入完整分布，往往必须经过整个允许状态集合的重新协调。这种联系是数学上明确的，但它是否对应实际样品，仍由模型中的状态与制备决定。

对 Z 求导，可以直接得到平均能量 $U = -\partial \ln Z/\partial \beta$。理由很简单：每一项 $\exp(-\beta W_i)$求导时带出负 W_i，再除以 Z，就成为能量按概率的加权平均。第二次求导得到 $\partial^2 \ln Z/\partial \beta^2 = \langle W^2\rangle - \langle W\rangle^2$，即能量方差。平均与涨落来自同一个函数的不同变化，因而不能在同一模型里随意独立指定。

这个推导还显示了条件的重要性。若能级随温度变化，直接求导会出现额外项；若以某种有效模型替代完整系统，应当检查这些参数依赖是否已经包含被消去自由度的贡献。不能把一个经验拟合能谱代入，再毫无说明地使用最简单的导数公式。计算很快不意味着它知道自己正在对什么保持不变。

由温度导数还可得定容热容与能量方差的联系：$C_v = \mathrm{Var}(W)/(kBT^2)$，其中体积、粒子数和哈密顿量参数按相应要求固定。于是，正则平衡系统的这种热容不能随意取负。若实际研究出现负热容，应先核对系综、长程作用或非平衡等条件，而不是用一个公式直接裁决所有可能物理系统。[24-3]

这里微积分的任务，是把概率组织中的变化转化为可测响应。它并没有凭符号求导创造热容实验；相反，实验要求导数所保持的条件与实际控制相对应。若实验是在定压而非定容条件下进行，就不能把测得数值不加转换地与这个 C_v 比较。变分对象和实验对象必须同一，公式才能进入检验。

24.4 一个二能级系统怎样检验整套关系

现在自行计算一个最小例子。一个单元只有两个无简并能级零与 ε，ε 为正，且接触正温度热库。令 $x = \beta\varepsilon$，则 $Z = 1 + \exp(-x)$，高能级概率 p 高 $= 1/[1 + \exp(x)]$，平均能量 $U = \varepsilon p$ 高。这里没有引入神秘的集体效应；全部结果都可由两项求和直接检查。

取 $x = \ln 3$，高能级概率为四分之一，低能级概率为四分之三，平均能量为 $\varepsilon/4$。能量方差为 $3\varepsilon^2/16$。由热容关系，$C_v/kB = 3(\ln 3)^2/16$，约为 0.2263。假如实验在相应温度附近得到平均能量约 $\varepsilon/4$，但能量方差远离 $3\varepsilon^2/16$，那么模型并没有被完整验证；可能遗漏简并、额外能级、相互作用或测量滤波。

再看极限。低温时 x 很大，高能级几乎不被占据，平均能量接近零；高温时 x 趋近零，两能级接近等概率，平均能量接近 $\varepsilon/2$。两端的热容都趋于零，中间存在峰值。高温热容趋零并不是系统能量不存在，而是这个只有两个能级的模型接近占有饱和，继续升温难以再改变平均能量。

如果把 N 个这样的单元独立组合，配分函数变为单元配分函数的 N 次方，总平均能量和方差各乘以 N。这一结论严格依赖独立与可加能量。若实验热容峰的位置或宽度随着单元密度显著变化，便应检查相互作用，而不能继续用独立组合公式把所有差异归为样品量误差。

这个算例最适合作为模型之间的竞争平台。两个模型也许都能拟合某一温度下的平均能量，却对温度扫描、方差和密度依赖给出不同预期。选择哪些观测留作预测，决定

一次拟合是否产生新知识。若用所有观测确定所有参数，再用同一批观测证明理论正确，就没有跨出自我复述。

还必须说明观测时间。若装置在能级跃迁发生以前便重复采样，多次读数高度相关，不能把样本数量直接当作独立实验次数。另一方面，测量过程若改变能级占据，原来的平衡分布又需要重新制备。概率的计算对象并没有脱离操作；它要求操作具备足以支持统计比较的稳定形式。

还可以用两个独立单元检验组合预测。每个单元的高能级占有率为四分之一时，联合能量取零、ε、二 ε 的概率分别是十六分之九、十六分之六和十六分之一，总平均能量为 ε/2。若观测完全没有二 ε 事件，而中间能量出现率异常增大，就需要区分样本不足与联合约束。按照独立模型，连续十六次独立测量不出现二 ε 事件的概率仍为$(15/16)^{16}$，约为 0.356，因此仅一次这样的短记录不能判模型失败。统计反驳也需要误差与样本规模，不能把“预期应该有一次”误写成“必定有一次”。

24.5 自由能把能量与状态数量放在同一张约束表上

对于相应正则平衡系统，亥姆霍兹自由能为 $F=-kBT\ln Z$。它同时满足 $F=U-T\Sigma$，其中 Σ 表示热力学熵。[24-4] 这不是说能量与熵是可以互相抵消的两种物质，而是说，在给定温度及相关约束下，状态的能量代价与实现方式的数量共同进入平衡选择。

仍用刚才的二能级例子，在 $x=\ln 3$ 时，$F=-\varepsilon\ln(4/3)/\ln 3$，约为 -0.2619ε。结合 $U=\varepsilon/4$，可以求出 $\Sigma/kB=\ln 4-(3/4)\ln 3$，约为 0.5623。负的自由能不意味着系统拥有负的全部能量；自由能的数值涉及能量零点，单凭一个正负号无法判断物体是否稳定或是否能凭空输出功。

将所有能级统一增加一个常数 C，可以直接检验这一点。Z 乘以 $\exp(-\beta C)$，平均能量和自由能都增加 C，概率分布保持不变。可见，有些绝对数值变化只是参考零点改变，而占有与可观察差异未必改变。比较两个过程时必须保持相容的零点约定，不能从两个不相容的定义里提取可用能量。

自由能最小原则也有边界。固定温度、体积和粒子数的条件，与固定温度、压力和粒子数的条件不同；后者常使用吉布斯自由能。若约束选错，极小化可以算得十分精确，却回答了另一个问题。数学工具的名称不能代替系统边界与控制方式的交代。

一个有用的竞争检验是比较两种可实现状态。若甲的能量更低，但乙拥有更多相容微观实现，升温可能改变自由能差的符号。这里不能只凭“系统总爱能量最低”解释全

部现象。能够使不同倾向共同进入可计算判据，是自由能结构的重要价值；这种结构也为下一章的相共存与新物态提供了入口。

24.6 时间平均什么时候能够代表系综平均

实验者无法真的同时测量无限多个复制系统，往往选择重复制备或在同一样品上连续采样。这两种办法与理论系综之间需要桥梁。重复制备若受到缓慢漂移影响，所得分布可能混合不同条件；时间采样若没有覆盖相关状态，也可能只看到一个长期滞留的局部区域。不能因为数据很多就自动宣布代表整个系综。

用一个两状态跳变模型作说明。设状态甲到乙的跃迁率为 a，乙到甲为 b，二者均为正且保持恒定。稳态乙占有概率为 a/(a+b)。当 a 和 b 都乘以十倍，稳态概率不变，但接近稳态和相关衰减的时间缩短。这说明，仅知道平衡概率不足以决定实际采样需要多久；分布相同的系统，可以有不同的动力学时间结构。

如果 a 极小而 b 也极小，观测窗口内可能一次跳变都没有。此时单条轨迹看起来完全不符合稳态比例，却不一定说明稳态计算错误；它可能意味着测量尚未进入适合的时间尺度。但这一说明不能成为无限延期的豁免。研究者应当独立测定跃迁率或延长窗口作预测，检验遗漏状态是否确实按照所给时间尺度出现。

遍历性在这里是需要条件支持的数学与物理命题，不能由一条平滑平均曲线证明。系统也可能存在额外守恒量、互不连通区域或极长弛豫时间。某些宏观量仍可在较弱条件下稳定，而不必要求每一种微观细节都被均匀遍历。因此，解释测量成功时，应说明成功需要哪一种代表性，避免把最强的假设不加区分地全部装进去。

对于知识发生，这段讨论尤为重要。统计量的稳定可能来自样品运行，也可能来自仪器积分，还可能来自数据处理人为平滑。只有改变采样窗口、制备次序和观测变量，比较这些改变对结果的影响，才能分辨稳定发生在哪个层级。一个平均数并不会自动带上它自己的来源说明。

24.7 杰恩斯的挑战：最好的推断是否已经足够

杰恩斯 1957 年的《信息论与统计力学》提出以信息约束与最大熵推断理解统计力学的路径。这使概率解释不必只诉诸实际轨道遍历，同时也把已知信息、约束选择和推断规则放到中心。[24-5] 这是本章必须认真面对的强竞争解释，不能把它贬成“因为无知，所以随便猜”。受约束的推断可以非常精确，并能够产生可检验结果。

如果只知道归一化和平均能量，最大化离散概率的信息熵，会得到指数形式的分布。它与正则分布的数学同形，说明两种建构可以在某些条件下汇合；但数学同形不自动证明二者回答同一个问题。推断方法说明在所给信息下怎样避免额外偏见，物理制备说明实际系统为什么可能实现这些约束和频率。

设两位研究者拥有不同信息。甲知道样品存在两个长期不互通的区域，乙不知道。乙用单一平衡分布作出最佳推断，仍可能系统性地预言错误；甲的条件化分布则可能更好。这里并不需要宣布最大熵原则被推翻，真正需要问的是哪些信息应当进入物理模型，以及如何通过实验获得它们。

对 SDE 而言，增量不能停在“概率也是互动生成”这一重复表态。我们必须追问：为何某个约束被纳入，另一个被忽略？样品制备与测量怎样使约束持续有效？新的显露是否促使原来允许状态的划分改变？如果这些工作都已由统计模型和实验设计清楚完成，SDE 就应承认其成果，而不能凭换名再收取一次原创性。

本书较强的解释在于，知识不只是在固定状态空间内更新权重，还可能通过新的操作与关系重建状态空间本身。发现原先合并的两类状态具有不同响应时，变化不只是概率由零点四改成零点六，而是被计数对象的身份发生调整。这种调整需要额外证据；若没有对象重构的实际必要，较弱的推断解释已经足够，不应强行升级本体论主张。

24.8 概率分布不是特征纠缠场的自动译名

现在可以更严格地回答章题。概率有时反映不完整信息，有时描述重复制备的频率，有时作为动力学模型的状态变量。它们之间存在联系，但不能由一张概率曲面同时证明所有解释。系综的成功也不等于“真实世界就是概率表”；它首先说明该表在给定构成与采样条件下具有稳定预测能力。

同样，相空间中连续的概率密度不因为被画成一张场图，就直接等于 SDE 的 E。一个概率密度必须先指定相空间、测度、状态变量和演化规律；特征纠缠分析则追问这些规定为什么以这种方式相互成立。如果把二者合并，原来需要研究的对象构成问题反而被一个方便的数学符号遮住。

可以选择 S—E—D 的进入路径处理一个具体失败：平均能量符合模型，方差却明显偏小。先从显露进入模型的约束关系，区分真实守恒限制、测量滤波和非独立采样三种解释；再设计不同带宽、不同间隔与独立重复制备的观测。若延长记录并适当调整采样间隔后，有限记录的方差估计趋近预言，操作 D 需要修正。相关采样改变有效样本数，

并不改变平稳过程的真实边缘方差；仪器滤波则可能改变测得的方差。若排除这些因素后不同装置仍显示偏小，便需要重审 E 中的系综选择。

另一条 E—D—S 沉淀路径从候选约束结构出发。例如认为样品与有限热库共同守恒，则推导允许的联合分布，设计改变热库大小的操作，再预言样品方差怎样随之变化。这样的研究使 S = F(D,E)、D = G(S,E)、E = H(S,D)各有具体任务。S 与 E 的张力推动 D 重排，S 与已核实 D 的张力推动 E 重建，新 E 和新 D 还必须返回原显露接受回检；不能只解释旧误差而不承担新预测。

吉布斯的思想创造因而远大于一个求和技巧。他使“同一种物理系统可以有哪些实现、这些实现怎样被共同规定”成为系统研究对象。宏观规律不再只由一条理想轨迹承载，也由约束下的分布、平均与涨落共同承载。在这个新世界里，概率不是认识失败的羞耻标记，也不是免于物理检验的哲学特权，而是一套必须说明如何成立的精确组织方式。

本章参考资料

[24-1] J. Willard Gibbs，Elementary Principles in Statistical Mechanics，1902，序言及第 I、IV 章。原典转录：https://en.wikisource.org/wiki/Elementary_Principles_in_Statistical_Mechanics；https://en.wikisource.org/wiki/Elementary_Principles_in_Statistical_Mechanics/Chapter_I。本章只短述研究对象与正则结构，未复述全书。

[24-2] David Tong，Statistical Physics，第 1 章微正则、正则及巨正则系综：https://www.damtp.cam.ac.uk/user/tong/statphys/statmechhtml/S1.html。

[24-3] MIT，Thermodynamics of Materials，Lecture 29：https://ocw.mit.edu/courses/3-020-thermodynamics-of-materials-spring-2021/mit3_020s21_l29.pdf。公式采用现代离散能级约定，二能级与导数演算为本书展开。

[24-4] David Tong，Classical Thermodynamics，第 4 章：https://www.damtp.cam.ac.uk/user/tong/statphys/statmechhtml/S4.html。

[24-5] E. T. Jaynes，Information Theory and Statistical Mechanics，Physical Review 106，620—630（1957），https://doi.org/10.1103/PhysRev.106.620；原论文存档：https://wwwusers.ts.infn.it/~milotti/Didattica/Bayes/pdfs/Jaynes_1957a.pdf。

第二十五章 相变与临界
新物态为什么从整体中出现

25.1 同一物质为何不能只用一套性质描述

水结冰以后，组成它的化学种类并未因此改名，承受剪切的方式却改变了。讨论相变，不能满足于说温度下降使分子运动变慢。缓慢与快速是连续差别，而固体能在相应时间尺度上保持形状，涉及组织方式和响应关系。物态名称之所以有物理内容，在于它联结了一组能够相互支持的测量：密度、热容、结构、形变和传输。新物态不是在每个分子上贴上一个新的标签，而是允许对整体提出新的、可以检验的判断。

这里的“整体”也不能被理解成无需分析的神秘力量。同样的分子可以在不同温压下形成不同相；同样的平均密度可能来自均匀相，也可能来自两相混合。因此，问题不是部分与整体谁更真实，而是哪些相互作用、约束及可达过程，使某组宏观关系获得稳定性。相变理论的困难恰在于：微观规律可以连续变化，宏观上最稳定的组织却可以突然更换。

第 23 章讨论微观运动如何支持宏观统计，第 24 章说明系综和自由能怎样连接约束与概率。本章把问题进一步收紧：当两个组织方案竞争同一份物质，什么决定它们共存、互相替代或长期僵持？第 53 章将讨论临界尺度与普适性；这里先以自由能相平衡为中心，区分平衡判据、形成路径和有限样品中的显露。三者相关，却不能互相代替。

25.2 平衡不是所有变化都停止

吉布斯在十九世纪七十年代关于非均匀物质平衡的研究中，把允许变化与极值判据联系起来。他在 1878 年所作作者摘要中，将固定能量条件下的熵极大以及相应的能量变分表述列为平衡理论的基础，并引入能量减去温度乘熵的函数。[25-1] 重要的不只是找到一个需要最大或最小的量，而是先交代哪些量允许改变、哪些条件保持不变。

设封闭的单组分样品与恒温恒压环境交换热与体积功，忽略额外场、化学反应以及宏观表面项。在这一限定中，吉布斯自由能 $G = U + pV - T\Sigma$ 给出平衡选择；Σ 表示熵。对纯物质的均匀体相，摩尔吉布斯自由能等于化学势 μ。这个等式不是说化学势是一种流体压力，而是给物质转移规定了一个可比较的代价。

假设两相甲乙已具有相同温度和压力，从甲相向乙相转移微小物质的量 $\mathrm{d}n$，系统总摩尔数不变。自由能的变化为 $\mathrm{d}G = (\mu_{乙} - \mu_{甲})\mathrm{d}n$。若两相都实际存在且允许双向转

移，平衡必须满足 μ 甲 = μ 乙。否则，总有一个转移方向降低 G。热平衡和力学平衡还不足以保证相平衡；物质能否转移以及转移后的自由能变化，是另一项条件。

这里已经能够看出一种常见误读。水和水蒸气共存时，分子并非停止穿过界面；宏观比例可以保持不变，而交换持续发生。宏观不变是相反过程在统计意义上的平衡，不是每个组成部分的静止。反过来，若隔板阻断物质交换，化学势不相等也可能长期保持。那是受约束的状态，不能用它否认相平衡判据。

这一推导还说明，为何同一种自由能不能无条件用于一切实验。固定体积而非固定压力时，允许变化的集合不同，适合的势也随之改变。把恒压条件所得的结论搬到刚性密闭容器，需要重新核对体积约束。自由能不是物质携带的一张永久排名表；排名依赖比较发生的条件，但这并不使排名变成任意的主观选择。

25.3 从相等的化学势推到共存线

沿两相共存线移动一个微小步长，化学势必须继续相等。对单组分体相，写成摩尔量的微分关系是 dμ = −s dT+v dp，其中 s 为摩尔熵，v 为摩尔体积。两相分别应用后相减，得到 Δv dp = Δs dT。因此，在 Δv 不为零且所用导数存在的共存段，dp/dT = Δs/Δv = L/(TΔv)。这里 L 是按甲转乙方向取符号的摩尔潜热。[25-2]

这个式子把三个原本可以分开测量的量绑在一起：相变吸放热、相变体积差、共存压力随温度的斜率。它不是从一张图上读取趋势后补上的解释。若其中两个量已独立测定，第三个量就不再自由。理论因此获得风险：斜率若持续不符，必须检查平衡、组分、场条件以及实验标定，而不能只把曲线重新画顺。

作一个不对应特定材料的算例。取共存温度 300 K，正向潜热 6000 J/mol，摩尔体积增加 0.002 m³/mol，则局部斜率为 10000 Pa/K。这只是一点附近的切线，不表示升温一百开尔文后仍可把压力差线性外推。若温度、潜热或体积差改变，必须重新计算或积分。一个局部微分关系，不能被偷换成全温区的经验直线。

若相变吸热而体积反而减小，斜率便为负。这种符号分析揭示了关系的力量：相变方向不能仅凭“加热使物体膨胀”一类常识决定。温度与压力共同选择相，体积与熵共同约束选择方向。对模型而言，符号也是证据；一个错误符号不能靠数量级相近来弥补。

接近两相差别消失的端点时，Δs 与 Δv 都可能趋于零，不能直接把公式读成无穷大或零。需要分析它们的比值以及自由能分支如何合并。这里已触及临界问题，但本章不

以“所有相变都有无限关联长度”概括它。普通一阶相变中的相共存与临界端点，是不同的物理组织情形。

25.4 相共存不是半个分子变成另一相

在给定共存温压下，设两相的摩尔体积分别为 v 甲和 v 乙，样品总平均摩尔体积处于二者之间。忽略界面贡献及其他组分，令乙相占总物质的量的比例为 x，就有平均 v = (1−x)v 甲+xv 乙。因而 x = (平均 v−v 甲)/(v 乙−v 甲)。这个杠杆关系来自守恒与加和，不必假设每个分子都处于介于两相之间的第三种状态。

取相同体积单位下 v 甲为一、v 乙为五、平均 v 为二，可得 x 为四分之一。这个四分之一是乙相的摩尔分数，却不是乙相占据的空间体积分数。后者为 xv 乙/平均 v，即八分之五。若一幅图像显示约百分之六十二点五的区域属于乙相，它可以与仅四分之一的物质转入乙相完全一致。把图像面积、三维体积与物质的量混为一谈，会制造虚假的理论矛盾。

这一点也提示我们，平均值常常隐藏组织。平均摩尔体积为二，可以是两种密度的空间组合；一个假设的均匀中间密度状态即使能短暂制备，也未必具有最低自由能。对足够大的短程作用系统，若分成两相降低体自由能，而增加的界面代价相对较小，分相就可能更有利。这里必须同时比较体积量级与面积量级，不能只看局部函数的一点高低。

在受限几何、小颗粒、强长程相互作用或明显弹性约束下，界面与非局域代价可能不再可以忽略。此时简单的直线混合关系即使仍满足总量守恒，也未必给出平衡相的组成与形态。条纹、液滴或其他微结构的选择需要额外计算。整体组织并不只有“完全均匀”与“无限大两块”两个选项。

所以，相的身份既不是平均数本身，也不是肉眼看到的边界本身。边界可以受分辨率影响，平均数可以由多种组织实现；相判据需要结构、响应与热力学关系相互印证。在 SDE 语言中，显露 S 提供区分的入口，特征纠缠 E 使多个指标形成可重复的联系，差异序列 D 则检验这种联系是否随过程保持。仅凭一张照片不能完成三者之间的回检。

25.5 一个能够突然换解的自由能模型

为把“组织竞争”变成可算的问题，考虑无外场、具有 m 与−m 对称的均匀标量模型。m 是无量纲序参量；把自由能密度及系数作适当单位化，写 f(m) = a m²/2−b m⁴/4+c m⁶/6，取 b、c 为正。六次项保证大幅度方向有界，负四次项允许零解与非零解在一段参数范围内并存。这个模型是本章的演算工具，不作为所有材料的通用拟合式。

驻点满足 $m(a-bm^2+cm^4)=0$。令 $y=m^2$，非零驻点由 $a-by+cy^2=0$ 给出。要让一个非零最低点与 $m=0$ 具有相同自由能，还要满足 $ay/2-by^2/4+cy^3/6=0$。用第一式消去 a，得到 $y=3b/(4c)$，于是共存参数 a 星 $= 3b^2/(16c)$。序参量从零变到非零，并不需要参数 a 发生跳跃；跳跃的是最低自由能所选中的解。

取 $b=c=1$，模型在 $a=3/16$ 发生这种竞争交换，非零最低点位于 $m=$正负$\sqrt{(3/4)}$。此时还存在 $y=1/4$ 的中间驻点，将它代回 f 可得屏障高度 1/96。用二阶导数检查，零点和外侧两个驻点都是局部最低点，中间两点则是局部最高点。仅仅解出驻点方程不够；若不比较函数值与稳定性，就会把能垒顶部误当成另一种平衡相。

非零驻点首先出现在 $a=b^2/(4c)$处，而零点在 a 降到零时才失去局部稳定性。于是两种局部稳定组织可以在一个区间内竞争，但真正的平衡交换只发生在 a 星。三个参数位置各有含义：出现局部解、自由能相等、局部稳定性丧失，不能合并为一个“相变温度”。若把 a 视作温度的光滑函数，还必须给出该函数及适用区间，才能换算成实际温度。

模型的贡献不是证明自然界一定服从六次多项式，而是将突然变化拆开为连续函数、多个极小值与稳定选择三个环节。它也有明显缺口：尚未描述空间界面、热涨落、序参量守恒与具体动力学。因而这里的局部稳定性边界属于均匀平均场模型，不能直接宣布真实样品必定在该处转变。任何实验检验都必须保留这层限定。

25.6 为什么自由能更低的相还没有出现

假设新相已具有更低的体自由能，但在旧相内部生成一个小液滴需要形成界面。经典均匀成核近似把球形液滴的代价写为 $\Delta G(r)=4\pi r^2\gamma-(4\pi/3)r^3\Delta g$，其中 γ 为正的界面自由能密度，Δg 为新旧两相的正体自由能差，忽略曲率依赖与复杂形变。对半径求导得到临界半径 r 星$=2\gamma/\Delta g$，能垒为 $16\pi\gamma^3/(3\Delta g^2)$。[25-3]

这里的“临界半径”不是临界点的关联长度。它表示长大与缩小之间的分界：小于该半径时界面代价占优，大于它时体自由能收益占优。两种“临界”共享边界之意，却属于不同模型中的不同量。词语相同不能替代量纲、定义和推导。

在 γ 近似不变的比较中，若 Δg 增至两倍，临界半径减半，能垒降为原来的四分之一。由于形成速率常对能垒呈指数敏感，这种变化可以使原本难以等待的转变迅速发生。因此，相同的平衡共存线并不保证相同的实际转变时刻。容器表面、杂质和既有界面提供异质成核通道时，上述球形均匀模型还需要改写，不能拿偏离结果简单指责样品“不纯所以不算”。

过冷与滞后因而提供了一种重要反例：观测到状态持续存在，不等于它具有最低自由能。样品可以被动力学屏障困住。反过来，观测到突跳也不自动证明平衡一阶相变；机械失稳、化学反应或反馈控制切换都可能产生突变。区分这些解释，需要改变扫描速率、等待时间、成核条件与热交换，并检查潜热、结构和可逆共存是否共同支持同一判定。

本章的核心张力由此显现：热力学比较给出一种组织更有利，实际显露却显示旧组织仍在。E 与 S 之间的不一致，推动 D 从单次升降温扩展为不同速率、等待时间及边界条件的实验序列，再用结果回检能垒解释。如果新路径改变转变时刻而平衡共存关系保持，动力学解释得到支持；如果连共存关系都系统改变，则需重查组分、应力或自由能模型。

25.7 有限样品中的圆滑转变与真实物态

对只有有限个状态、能量随参数平滑变化的模型，在正温度下，配分函数是有限个正指数项的和。它不会因为某两个权重恰好相等就产生真正的不连续。连续自由度系统还须另行保证积分收敛及参数正则性，不能仅凭“有限”二字作无限制断言。宏观奇异性与有限实验曲线之间，需要说明极限如何取得。[25-2]

一个简化的两分支模型足以显示这种区别。设每粒子自由能为 f 甲与 f 乙，总粒子数为 N，忽略界面态时 Z 约等于 exp(−βNf 甲)+exp(−βNf 乙)。乙分支的概率于是为 1/[1+exp(βN(f 乙−f 甲))]。对任何有限 N，这都是光滑的。两分支明显混合的自由能差范围约为 kBT/N；N 增加，转换窗口收窄，却不要求每个微观变量都作同样的跳变。

这不是实际相共存的完整模型，因为它只保留整块样品的两个分支，没有保留液滴和界面等中间组织。其价值是说明，大样品中的陡峭变化可以来自极大数量的统计权重竞争。实际系统还要处理界面、边界条件与观察时间，不能从这一个式子直接预测所有跃迁频率。

有限性的存在也不使物态概念失效。我们并不要求一块有限晶体拥有数学上无限长的晶格，才允许讨论它的弹性和衍射。合理的问题是：在给定尺度与时间范围内，这种组织是否形成稳定的响应体系，有限修正是否可以计算，扩大样品后结论是否趋于一致。热力学极限提供清晰的分类工具，有限尺度实验则检验它是否抓住了现实中已经显著的组织差别。

涨落在此并不是宏观定律之外的噪声残余。它能让系统跨越局部屏障，也能使接近临界点的简单均匀模型失去定量可靠性。第 53 章将详细讨论关联长度与粗粒化；本章只

保留一个必要区别：相共存的自由能选择、成核的能垒与临界关联，承担不同解释任务，不能用“整体涌现”一词把三种计算压成同一件事。

25.8 新物态的创造命题怎样接受检验

一个强有力的竞争解释是：相变并未创造新的基本自然律，它只是同一微观哈密顿量在不同参数下的集体现象。只要确定粒子、作用与边界，原则上便可得到各相。这一解释足以说明许多新性质，并且提醒我们，不应把计算困难夸大为自然界断裂。本文认可这一约束，也不把“新”定义为必须违反微观规律。

知识创造所要进一步说明的是，何种稳定组织使新的物理判断成为可能。液体与固体不仅占据微观状态集合中的不同区域，它们支持的有效变量、响应方式与实验操作也不同。把一块材料当作弹性体，意味着可以建立应力、应变与边界条件之间的一组关系；若这种关系持续失败，它就不能仅靠命名保住身份。新物态在这个层级上是新的可操作物理世界，其成立仍须接受既有微观理论与实验的双重约束。

具体可以走 S—D—E 的再创路径：先遇到密度突跳及热流峰值，再组织慢速扫描、定点等待与相比例测量，最后将显露重构为自由能分支、相共存和成核的联合模型。D 与 S 的张力推动 E 改变，而改变后的 E 必须同时预言共存线斜率与杠杆关系，不能只拟合最醒目的峰。

这一路径也有明确可败条件。若峰值随传感器带宽移动而结构与平衡量均无变化，优先修正测量模型；若产物组成改变，应转向化学反应解释；若只有有限时间的滞后却没有相应平衡证据，就不宜贸然宣布新平衡相。反之，若不同探测方法得到可协调的相边界、潜热、体积差与响应变化，组织重构便具有超出单一读数的实在根据。

因此，整体性并不是把证据门槛降低，而是把证据要求提高。一个局部信号必须进入多个关系的交叉检验，才能支撑相的身份。相变思想给物理学留下的深刻改变，正在于它使“同一种物质”与“同一种物理组织”不再等同：组成保持，可以伴随可检验的新关系出现；新关系成立，又不要求抛弃对组成及作用规律的追问。

本章参考资料

[25-1] J. W. Gibbs, “On the Equilibrium of Heterogeneous Substances”, 1870 年代论文及作者摘要，American Journal of Science, 3rd series, 16, 441–458, 1878。阅读作者摘要开篇的平衡变分原则与自由能定义，收入 Scientific Papers 扫描本第 390 页。https://en.wikisource.org/wiki/Page:Scientific_Papers_of_Josiah_Willard_Gibbs.djvu/390

[25-2] David Tong, Statistical Physics, Chapter 5, “Phase Transitions”。阅读相共存、Clausius–Clapeyron 关系与有限系统配分函数的有关段落；正文算例与两分支演算为独立展开。https://www.damtp.cam.ac.uk/user/tong/statphys/statmechhtml/S5.html

[25-3] James P. Sethna, “Critical Droplets and Nucleation”。阅读球形液滴的界面与体自由能竞争，以及均匀和异质成核的区别。https://sethna.lassp.cornell.edu/Nucleation/

第六编 场之物：法拉第、麦克斯韦与电磁世界

第二十六章 电与磁的相遇
实验怎样把分离的世界联结起来

26.1 相同的数学形式为什么还不是统一

电荷能够吸引或排斥，磁体也能够吸引或排斥。若只看这些动词，很容易认为电与磁原本就是同一种作用。但熟悉的相似性并不足以证明物理同一：引力也表现为吸引，不能因此把石块、带电小球与磁针看成同一种对象。一个严格的统一，必须说明原来独立规定的现象，怎样在同一操作中彼此转换或相互约束。

此前热学诸章已经展示，能量守恒没有抹去热与功的过程区别，统计力学也没有把温度直接等同于一个分子的速度。电磁学的统一同样需要保存差别。电荷分布、持续电流、磁针方向和导线受力并非同一读数；它们能够统一，是因为存在一组共同关系，使一种读数的改变限制另一种读数，而不再允许各自随意变化。

本章所说的实验统一，因而有三个层次。首先是可重现的交叉作用：原来归在电学中的安排使磁针发生规律变化。其次是可区分的关系：效应依赖方向、位置和通断方式，不只是偶然发生。最后是共同模型：不同装置的变化能够由相同参数解释，而且模型对尚未用于调整的安排给出预言。这三个层次若被压缩成“发现电能生磁”，便会失去知识生长的主要内容。

还应区分历史研究与现代演算。下面将使用电流 I、磁感应强度 B 和现代磁静力学公式，以便清楚计算；这些记号与理论结构不能逐项倒写到 1820 年的实验者头脑中。思想史的任务不是装作不知道后来的理论，而是明确哪些是原始报告，哪些是后来的重述，哪些是本章用来检验关系的独立模型。

26.2 一根导线怎样改变磁针的证据身份

奥斯特 1820 年发表的报告描述了闭合电路附近磁针的偏转，并系统改变导线相对磁针的位置。导线在针上方和下方时，偏转关系发生反转；报告还比较了不同导线材料和插入物的情况。[26-1] 这不是仅仅记下针动了一次，而是逐步建立一个操作集合，说明什么变化与电路安排共同发生。

这一细节尤其重要。磁针本来用于辨认磁方向，现在又成为探测电路效应的工具。它的材料不必改变，但证据身份改变了：指针不再只报告某个既有磁体的位置，而开始参与测定电流周围的空间组织。仪器的新用途并非研究者单方面宣布；只有通断、换向和移位产生稳定对应，它才获得这种资格。

设一根近似南北向的直导线放在水平磁针上方。不能仅凭磁针偏离南北便断言出现了新的基础作用。导线通电可能发热，支架可能轻微移动，附近铁件可能重新磁化。真正有区分力的设计，是让这些解释在同一组操作中作出不同预期。例如，在接近相同发热条件下反转电流，磁效应应改变方向，而通常的焦耳发热主要依赖电流平方，不会随电流符号一同反转。

这个比较并不意味着一次换向能够排除所有非磁干扰。接触不对称、热电过程或机械耦合仍需分别检查。它说明的是，证据的力量来自候选解释之间的差异，而不是操作次数本身。重复一百次同一种安排，可以提高稳定性估计，却未必区分两种都能解释它的模型。

在导线与磁针之间插入某种物体，也必须说明插入改变了什么。绝缘板不阻断通电导线本身，却可以改变热对流与机械接触；磁性材料还会重组附近磁场。不能笼统说“穿过障碍物，所以作用没有媒介”。穿透某种材料、是否需要接触、传播是否有延迟，是三个不同问题。初期实验已经强有力地联结电与磁，却没有单独决定后来的场本体论。

26.3 方向进入定律，量的大小也必须重新解释

用现代语言描述长直导线的稳恒电流，在真空或近似非磁性环境中，远离端部且导线半径远小于观测距离时，磁场沿环绕导线的方向，大小 $B = \mu_0 I/(2\pi r)$。这里 r 为垂直距离，μ_0 为真空磁导率。本章数值仅取 μ_0 约等于 $4\pi \times 10^{-7}\ \mathrm{N/A^2}$，不把这个近似值当作不需限定的精确常数。[26-3]

值得注意的是，方向结构不是对大小公式的装饰。磁针既不简单指向导线，也不简单背离导线；它对当地合成磁场作出定向响应。若把所有远距离作用都预先理解为沿两物体连线的吸引与排斥，就会遗漏这种环绕关系。新问题要求研究者把空间方位纳入对象规定，而不再只问距离越近作用是否越强。

设导线产生的水平磁场与地磁水平分量 B 地正交，磁针足够小，支点摩擦和附加力矩可忽略。设偏角为 θ，平衡条件给出 $\tan\theta = B_{线}/B_{地}$，而不是 θ 直接正比于电流。这

个关系可由力矩相消得到：导线场产生的转动力矩与 B 线 cosθ 成正比，地磁回复力矩与 B 地 sinθ 成正比，两者相等便得到正切关系。磁针磁矩因出现在两边而消去。

作一个假设算例，令 B 地为 20 微特斯拉，导线距针 0.02 米。I 为 1 安培时，B 线约 10 微特斯拉，偏角约 26.6 度；电流增至 2 安培，B 线约 20 微特斯拉，偏角变成 45 度，并不是 53.2 度。若研究者直接用角度比估计电流比，就会把仪器响应的非线性误写成电磁规律的变化。

测量误差也可以定量追踪。若角度的小误差为 δθ，并用弧度表示，则由正切关系有 δB 线约为 B 地 $\sec^2\theta\,\delta\theta$。在相同角度分辨率下，接近直角的读数对应更大的场误差。把指针偏得更明显，并不一定意味着测量更精确；应同时比较信号大小和响应斜率。

这个算例还能产生新的检查。保持电流不变，把距离加倍，长直导线近似预言 B 线减半，因此正切偏角也减半。若偏角本身近似减半，只能说明处于小角度范围，不能当作普遍公式。角度接近直角时，微小读数误差会被正切显著放大，测量程序便需要选择合适范围。一个现象成为量，既需要外部规律，也需要仪器的响应模型。

26.4 有限导线提醒我们整个回路没有消失

真实导线没有无限长。对于以观测点垂足为中心、总长 2a 的直线段，若观测点离线段为 r，现代毕奥—萨伐尔积分给出的这一线段的贡献为 B 段 = $\mu_0 Ia/[2\pi r\sqrt{(r^2+a^2)}]$。它相对无限长导线结果多了 $a/\sqrt{(r^2+a^2)}$这个因子。[26-3] 本章借此做一次模型误差的独立估计，而不把真实实验的所有导线都缩成一个线段。

取 a = 0.1 米、r = 0.02 米，该因子约为 0.9806；使用无限长公式会使这部分场的估计高约百分之二。如果 a 与 r 相同，因子降为 $1/\sqrt{2}$，差别已经不能忽略。这里没有出现新物理，只是原来隐藏在“足够长”之中的比例条件开始显露。一个近似好不好，取决于无量纲比例，不取决于导线在肉眼看来是否很长。

更深一层的问题在于，稳恒电流不能在孤立线段的一端不断消失、另一端不断产生。它需要完整回路，返回导线也会贡献磁场。将返回段放远，可以使它在研究区域的影响变小；将它贴近正向导线，两段磁场可能部分抵消。若两个实验使用相同的直线段、电流与测量距离，却有不同的回路线形，结果并不必然相同。

因此，“同一个装置条件”并不是列出三项数字就完成。研究者必须说明哪些关系可以忽略、为什么可以忽略，以及忽略后误差如何控制。回路形状原先被放在边缘，现在进入待解释结构；这正是特征纠缠 E 的具体内容。它不是把所有实验室物件都收入背景列表，而是辨认哪些连接实际改变被测量的场。

还可以借此设计一个竞争检验。若偏差随返回导线的位置系统改变，并符合两段场的矢量相加，应优先修正几何模型；若只随电源工作时间漂移，应检查电流与热变化；若在这些修正之后仍出现可重复的额外方向依赖，才有理由扩大理论挑战。把所有偏差直接称为发现，或一律称为误差，都绕开了研究所需的分层判断。

26.5 从磁针转向另一条电流

安培在 1820 年的研究中把问题推进到两条载流导线之间的作用，随后发展电动力学方案。他以平行同向电流相吸、反向电流相斥等关系，探索将磁现象与电流统一起来的可能性。[26-2] 这种改变并非只是更换一种探针：实验中的两端现在可以由相同类型的电路操作控制，从而得到对称且丰富的比较。

在现代磁静力学近似下，两条相距 d、足够长的平行细导线，其单位长度受力大小为 $f = \mu_0 I_1 I_2/(2\pi d)$。这里假设稳恒电流、端部和返回段影响受到控制，并使用导线受力关系与另一导线产生的磁场相结合。这个推导属于现代场表述，不能说安培已经以相同方式先算 B 再算受力。

取 $I_1 = 2$ 安培、$I_2 = 3$ 安培、$d = 0.01$ 米，可得每米受力约 1.2×10^{-4} 牛。若只反转一条导线中的电流，方向由吸引改为排斥或相反；同时反转两条，乘积不变，作用方向保持。若同时把两电流加倍，受力增至四倍。这一组符号与尺度关系，比“通电就有作用”严格得多。

它还允许把磁针测量链与机械受力测量链联结起来。先用已标定的磁针响应推测导线场，再预言另一导线的受力；或者由电流间受力反推磁场，并与磁针比较。两条链不完全独立，因为可能共享电流标定和几何测量，但探测响应不同，可以对各自的系统误差形成约束。统一的强度来自这种交叉限制，而不是把不同实验都写上同一个名称。

一个反例可以说明限制所在。实际导线若带有额外净电荷，就可能同时出现静电力；若导线悬挂结构温度改变，也会出现机械位移。单凭两线靠近不足以确认电动力学解释。反转一条电流、反转两条、改变电流大小及导线间距，并监测电位与支撑，才能让竞争解释承担各自的可败条件。

这里也不应把现代电子理论提前赋予安培。把磁性归于内部电流，为统一提供了重要研究方案；后来关于材料磁性的解释还涉及微观运动及量子性质。一个宏观环流类比取得成功，并不等于物质内部所有机制已经被直接观察。思想史中的创造可以真实而有边界，无须靠后见之明补成无所不知。

26.6 统一是减少任意性，而不是删除差别

假设有人为每一种导线形状都指定一个独立的偏转系数，为每一种材料再加一个修正常数。只要参数足够多，这个方案也许能够覆盖已知数据。但它尚未实现有力统一，因为改变实验以后仍需重新调整，过去经验无法严格限制未来。更强的模型应让相同电流、几何规则与叠加关系跨装置工作，并暴露哪些新结果会使它失败。

反过来，也不能仅凭公式数量少就判定理论优越。一条短公式若把所有复杂性藏入一个任意函数，同样可能没有预测力。需要比较的是自由参数、适用条件和预言的独立性。例如，固定一套电流标定以后，改变距离所得曲线应受到约束；不能每换一个距离就偷偷重新定义电流，使理论永远正确。

统一还必须保存尚未统一的差异。稳恒电流附近的磁场，不等于电荷静止时的电场；磁针转动，不等于导线内部载流子的微观运动。磁静力学中的即时积分表达，也不能直接推广到快速变化与远距离传播。若观察区域很大、变化很快，电流源当前的数值未必足以描述远处此刻的响应。下一阶段的问题，正是在这些条件不再可忽略时产生。

因此，本章既不把 1820 年写成麦克斯韦理论的一次缩小预演，也不把现代统一全部归结为最初的磁针偏转。奥斯特打开了交叉作用的稳定入口，安培扩大了电流间的实验与理论关系；随后还需要研究变化的磁状态能否生成电效应，传播及能量怎样进入空隙。每一步都增加新的可问对象，也增加新的检验责任。

这让“电磁世界”的出现具有具体含义：电流开始能够通过磁响应被辨认，磁作用开始能够通过电路条件被调控，几何方向成为共同定律的一部分。这些变化超出词汇合并，因为它们扩展了允许的操作与推断。但如果较弱的共同测量模型已经解释全部结果，就不能另行宣称每一次偏针都证明了某种额外的本体断裂。

26.7 一次反常偏转怎样走完研究路径

设在磁针实验中，正负电流所得偏角大小不对称。最直接的反应可能是给理论增加一个修正因子；更有用的起点是保留完整记录，询问不对称来自场、仪器还是操作。这个问题可以沿 S—E—D 的进入路径展开：由显露 S 进入合成场与探针响应的结构 E，再安排能够区分候选的差异序列 D。

例如，若附近存在一个与导线场同方向的固定背景分量 B 偏，正向时测得 B 偏+B 线，反向时测得 B 偏−B 线。经磁针模型换算成有符号场值后，二者半差给出 B 线，半和给出 B 偏。这一分解要求测量关系已正确线性化；不能直接对大偏角取半差而忽略正切变换。操作、模型与结果必须联立。

若所得 B 偏随装置整体转动而改变，可进一步检查地磁分量和附近磁性物件；若它随电流方向改变而不符合固定背景假设，则要检查换向电路或探针滞后。两维张力推动第三维的改变在这里有明确内容：S 与 E 不符，推动 D 加入换向与旋转；新的 D 与 S 仍不相容，推动 E 调整探针或背景模型；修订后的 D 与 E 又必须预言下一组 S。

三方程因而不是附加的装饰。S = F(D,E)落实为通电、方向与场响应怎样生成偏角；D = G(S,E)落实为根据不对称选择哪次反转与移位；E = H(S,D)落实为根据操作后的记录重建哪些场分量与仪器条件。这里的 F、G、H 是发生学关系标记，不代替上文可计算的磁场与力矩方程，更不把实验者、导线和磁针分别塞进三个字母。

这个路径也允许失败。若各种探针与布线修改都不能消除异常，且异常能够跨装置复现，就应承认现有模型的边界；若异常跟随某一支磁针移动，则不宜先重写电磁学。发生学的开放性不意味着所有解释平等，而意味着每次结构改造都须回到有区分力的观察，承担比原来更多而不是更少的约束。

26.8 一个联系还会产生相反方向的问题

电流改变了磁针，那么磁的变化能否产生电的变化？这不是逻辑上已经被保证的对称命题。自然界中的因果关系不能因为句子可反读便自动可逆，静止磁体也不会无条件在静止导线中维持电流。正因为这种逆向期待可能失败，它才值得转化为实验，而不是停留在自然统一的信念中。

下一章的法拉第铁环实验，将显示一个比“磁能生电”更精细的答案：关键并非仅有磁状态，而是相关状态怎样改变，以及回路如何把改变转成可观测电效应。由稳恒关系走向过程关系，意味着对象身份再次发生调整。电流不只是已经存在的原因，也可能是另一套过程的短暂结果。

电磁统一的第一阶段因此留下两项并行任务。需要把既有实验放进更稳定的量化结构，也需要让这种结构生成它尚不能回答的问题。前者防止统一沦为漂亮比喻，后者防止成功模型变成禁止新观察的边界。一个共同物理世界的创造，正发生在这种稳定与开放的相互要求之中。

本章参考资料

[26-1] H. C. Ørsted，“Experiments on the Effect of a Current of Electricity on the Magnetic Needle”, Annals of Philosophy 16 (1820):273–276。核读 Samuel Devons 编 Oersted: The Beginnings of Electromagnetism 附录所收同时代英译首页，重点闭合电路、导线上下位置与插入物比

较。https://stwww1.weizmann.ac.il/wp-content/uploads/2016/08/Oersted-%E2%80%93-The-Beginning-of-Electromagnetism.pdf

[26-2] Christine Blondel and Bertrand Wolff, “Ampère lays the foundations of electrodynamics (September 1820–January 1821)”及“In Search of a Newtonian Law of Electrodynamics (1820–1826)”, Ampère Archives。阅读电流间作用、安培方案与磁流体竞争解释的历史讨论。https://ampere-archives.fr/histoire-electricite/parcours-historique/lois-courants/ampere-electrodynamique/eng

[26-3] Neville Harnew, Electromagnetism, Lecture 11: Magnetostatics and the Biot–Savart Law, University of Oxford, 2022。阅读直导线积分与两平行导线受力；本章磁针响应、数值与误差分析为独立演算。https://users.physics.ox.ac.uk/~harnew/lectures/EM-lecture11-handout.pdf

第二十七章 法拉第的铁环
短暂变化怎样成为新的物理对象

27.1 期待一个持续电流，却遇到两次短暂偏转

一八三一年八月二十九日，法拉第用铁环和相互绝缘的绕组获得了电磁感应的明确结果。现存装置由铁、铜线及绝缘材料组成；英国皇家研究院的馆藏说明把这一日期与铁环实验联系起来。[27-1] 它后来可以被理解为变压器的先驱，但把后来的设备名称放回这一刻，并不能解释实验究竟改变了什么。

《电学实验研究》第一组论文记载，铁环一侧绕组接电源，另一侧接检流计。接通时指针偏转，保持通电后指针回到原位；断开时又向相反方向偏转。法拉第也强调，报告的编排并不完全遵循发现次序。[27-2] 因此，思想史既不能把印刷段落顺序当成实验日程，也不能只保留成功接通的一瞬间，删去保持通电时的无持续效应。

这里最值得追问的，恰是“没有持续效应”为何成为发现的一部分。若期待磁状态只要存在就持续生成电流，保持通电后的零读数会被视为失败；若把通断与变化纳入比较，同一个零读数却开始划定规律的适用对象。它与两次相反的偏转共同表明，不能只用磁作用的大小描述结果，必须研究它怎样建立和撤回。

铁环没有单独提供一条完整方程，却迫使研究对象从静态配置移向过程关系。电流不再只作为已存在的磁作用来源，也成为另一段改变的结果。短暂效应原本容易被当作接触扰动或仪器颤动，现在必须取得稳定身份。这种身份不是因为现象足够壮观，而是因为它能够在时间次序、方向和回路条件的改变中被反复追踪。

27.2 瞬间并不等于没有结构

把开关接通、等待、断开三个动作写在纸上，看似只是一条操作顺序。但对感应研究而言，每一段都具有不同意义。接通使电流从某个初值变到另一个值，等待使它接近稳态，断开又改变电流及相关磁状态。如果把全部读数平均成一个数，正负效应可能抵消，最重要的证据反而消失。

法拉第当日的日记保留了绕组连接、远置磁针以及接通后的摆动记录；这使铁环不只是后来概念的图解，而有可核对的实验构成。[27-3] 不过，本章不把原始指针运动直接等同于瞬时电流曲线。机械指针有惯性、回复力与阻尼；一个短脉冲已经结束，针仍可能继续摆动。仪器运动持续多久，与电流持续多久并不是同一个时间。

考虑现代的简化检流计模型：指针角度 θ 满足 J 转 θ¨+bθ˙+κθ = K i(t)。J 转是转动惯量，b 是阻尼系数，κ 是回复系数，K 把电流转成力矩。这个线性近似要求小偏角以及稳定磁矩等条件。它说明可见角度并非电流的无延迟复印，而是输入信号经过仪器动力学后的响应。

若脉冲短到在它持续期间 θ 几乎未改变，积分方程可得 J 转 Δθ˙约等于 K ∫ i dt。初始角速度由通过的电荷量决定；后来针摆得多远，还由阻尼与回复过程决定。因此，同样高但不同宽的电流脉冲，可能产生不同摆幅；同样电荷量但不同峰值的短脉冲，也可能给出近似相同的最初冲量。只读最大偏角，未必能唯一恢复峰值。

这不是以现代仪器学贬低早期实验，而是说明瞬态对象需要什么样的证据桥梁。直接把指针曲线叫作电流曲线，会把探测器的时间结构偷偷转交给被探测过程。相反，若更换响应时间不同的探针仍得到一致的积分量，短暂效应便获得更强的独立地位。实验对象由仪器显露，却不能与某一种仪器运动混同。

27.3 磁通把局部变化与整个回路联系起来

现代感应定律首先需要定义磁通。对一个有方向的曲面，Φ = ∫ B·dA；B 为磁感应强度，dA 含有曲面法向。均匀场穿过平面线圈时，磁通为 BA cosα，α 是场与法向的夹角，而不是场与线圈平面的夹角。这个定义使面积、方向和场强进入同一个量，防止把“磁体更近”当作一切情形中的唯一变量。

对于固定闭合回路，感应电动势为 ε = −dΦ/dt；N 匝在各匝链接相同磁通时为 ε = −N dΦ/dt。若各匝磁通不同，应先把磁通相加，形成磁链，再求导，不能只凭匝数直接相乘。负号依赖回路正方向与曲面法向的一致约定，表示感应响应的方向关系，不是任意附加的减号。[27-4]

这一定律首先给出的是电动势，不是电流。电流还依赖回路电阻、自感以及其他元件。开路时没有同样的闭合传导电流，却不能因此断言所有感应电效应消失。相反，一个连接检流计的闭合回路能够把电动势转成可测电流，但它也改变系统响应。测量不是毫无代价的窥视，它属于模型的一部分。

磁通的整体性同样容易被误解。它并不是一条神秘绳索穿过线圈，也不是只有位于导线材料上的 B 才重要。曲面上场分布的积分受到回路边界约束；在无磁单极的经典描述中，不同同边界曲面的磁通相容。若更换曲面就得到相矛盾的答案，应检查场模型和方向，而不是让同一回路同时拥有多种任意电动势。

对固定回路，这还意味着诱导电场一般不能被全球单值的静电势差完全描述。沿闭合路径的电场积分可以不为零，而纯静电场的闭路积分为零。这里用电场记号 E 电，以区别 SDE 中的特征纠缠 E。电的世界不再只有两端之间的静态电势高低；围绕整个闭路的电作用也开始成为基本描述对象。

27.4 两个相同终点，可以产生不同波形

设在线性、固定几何且弱反馈的近似下，初级电流 I_1 在次级产生磁链 $\lambda_2 = M\,I_1$，M 为常量互感。于是次级感应电动势 $\varepsilon_2 = -M\,dI_1/dt$。这不是说互感是两只线圈各自的孤立属性；它取决于它们的共同几何与材料关系。转动或移动其中一只，M 便可能改变，公式须相应扩展。

作一个自拟算例：$M = 0.02$ 亨，次级电阻 $R_2 = 10$ 欧，暂忽略次级自感和反作用。令初级电流在 0.1 秒内从零线性增至 2 安培，则 ε_2 为−0.4 伏，次级电流为−0.04 安培；保持初级不变时，该近似下次级电流回到零。随后若在 0.2 秒内降回零，电动势为正 0.2 伏，电流为正 0.02 安培。

两个脉冲峰值不同，持续时间不同，但通过次级的有符号电荷量分别是−0.004 库和正 0.004 库。原因可以直接从积分得到：$Q_2 = \int i_2 dt = -M\Delta I_1/R_2$。只要上述条件成立，积分量只依赖初级电流的端点，而峰值依赖路径如何走。状态变化与过程形状在同一实验中承担不同任务，不能拿其中一个替代另一个。

这提供一个比“变化越快越强”更严格的检验。用两种斜坡改变相同的初级电流差，积分电荷应一致；峰值则按斜率改变。若只看指针最大摆幅，可能把电荷响应误读成峰值不随速度变化，从而错误否定感应规律。需要先确定探针测的是哪一类量，再决定什么结果构成反例。

还可以设计零净变化的往返过程。初级电流先上升再下降，整个过程 ΔI_1 为零，于是有符号总电荷可为零；但电阻发热由 $\int R_2 i_2^2 dt$ 决定，两个脉冲都作正贡献。这个区别非常关键：总积分为零不表示没有发生能量转移，正负电流也不会以同样方式抵消热。单个汇总指标无法代替过程的完整结构。

对刚才两个脉冲，忽略自感时次级总焦耳热为 $10\times0.04^2\times0.1$ 加 $10\times0.02^2\times0.2$，等于 0.0024 焦。这个数字只属于假设模型，不是 1831 年铁环的实测热量。它显示可以从同一组波形得到不同的物理问题：电荷转移、峰值电流、持续时间和能量损耗，彼此相联而不等同。

27.5 自感怎样使“立即归零”失效

若次级自感 L_2 不可忽略，电路方程变为 $L_2\,di_2/dt+R_2i_2=-M\,dI_1/dt$。初级停止变化以后，右侧变零，次级电流却一般不会立刻消失，而是按自身时间尺度 L_2/R_2 衰减。这不是感应定律失效，而是原先忽略的响应自由度参与了过程。

在初级电流以常斜率 a 上升、次级初始电流为零时，解为 $i_2(t)=-(Ma/R_2)[1-\exp(-t/\tau)]$，其中 $\tau=L_2/R_2$。若 τ 远小于斜坡持续时间，除起止附近外，可近似使用上一节的电阻模型；若二者相当，波形明显变钝，峰值也不能再用 ε_2/R_2 瞬时计算。

取 $L_2=0.2$ 亨、$R_2=10$ 欧，则 $\tau=0.02$ 秒。若一个斜坡只持续 0.01 秒，终点电流仅达到相应稳态幅值的 $1-\exp(-0.5)$，约百分之三十九点三。忽略自感会高估这一时刻的电流。仪器若同时存在机械迟滞，实际角度还要再经过另一层响应，不能把两种时间常数合并成一个原因不明的“延迟”。

尽管波形改变，积分关系可以在更宽条件下保留。对整个过程积分，有 $L_2[i_{2\,末}-i_{2\,初}]+R_2Q_2=-M\Delta I_1$。若起点和终点均等到次级电流回到零，便仍得到 $Q_2=-M\Delta I_1/R_2$。若过早结束记录，残余自感项不为零，所谓电荷不守恒的表象可能只是漏掉了尾部。这是一项可以通过延长观测窗口直接检验的解释。

不过，常量 M、L_2 也不是永久正确。初级电流受到次级反作用时，需要联立两个回路；高频时还可能出现电容、辐射与分布效应。上述常微分方程属于集中参数近似，要求器件尺度上的传播时间相对过程尺度足够小。本章用它说明瞬态结构，不能以它替代下一章将讨论的空间传播理论。

在线性、互易且磁能为非负二次型的双回路模型中，还存在一项不能任意越过的约束：$W_{磁}=L_1I_1^2/2+MI_1I_2+L_2I_2^2/2$。将它对 I_1 配方，可得 $L_1[I_1+(M/L_1)I_2]^2/2$，再加 $[L_2-M^2/L_1]I_2^2/2$。若 L_1 为正，要使任意电流组合都不产生无下界的负磁能，就需 M^2 不大于 L_1L_2。互感不能为了拟合大信号随意增大；自感与互感共同受储能结构限制。

这一限制也说明“把更多能量从一个绕组送到另一个绕组”需要改变哪些条件。次级接上负载以后，其电流产生的磁作用会返回初级，电源为了维持指定输入波形，通常必须提供相应额外能量。把初级波形固定作为演算条件是可以的，但不能同时把保持波形的电源排除，进而宣称次级热量没有来源。理想电压比可以由共同磁通推得，完整功率收支仍要求说明负载及损耗。

27.6 铁环既放大效应，也带来材料历史

铁环的作用不能仅用“它是一个媒介”概括。在线性近似下，铁芯可以增加并引导链接绕组的磁通，使同样初级变化产生更明显的次级响应。但真实铁磁材料还可能具有非线性、剩磁及损耗，磁状态不一定只由当前电流唯一决定。把铁环当作一个永远不变的倍增器，会遗漏实验最容易产生路径差异的部分。[27-5]

因此，互感常数往往只是某一区间或工作点附近的有效参数。若磁链 λ_2 是 I_1 的非线性函数，而暂不考虑额外历史变量，电动势应写为 $-(d\lambda_2/dI_1)(dI_1/dt)$。同样的电流斜率，在不同工作点可以得到不同电压。变化的是磁链对电流的响应，不必立刻宣布感应定律本身改变。

更进一步，若存在磁滞，λ_2 还依赖此前怎样磁化与退磁，不能用一条单值曲线覆盖全部记录。实验序列就必须包含制备历史：从何种状态开始、最大电流曾达到多少、循环多少次、等待多久。这些不是为了迁就结果而补加的背景说明，而是候选材料模型明确要求记录的状态信息。它们只有能改善独立预测，才值得进入结构。

一个强检验是把两组初级电流波形做得相同，只改变此前磁化过程。如果次级输出随历史系统变化，线性常互感模型就受到挑战；若采用无磁芯装置后这种差异明显减少，材料解释得到支持。还需要保持几何、电阻和测量带宽可比，不能把所有变化都归给取走铁芯这一动作。

这也给“实验创造对象”一个清楚界限。实验者可以通过退磁和偏置准备不同的可重复工作状态，但不能随意宣布铁芯忘记历史。新对象是带有制备、响应和适用范围的有效电感元件；它并不因为被画成电路图上的一个符号，就与材料内部过程断绝联系。封装成功意味着这些过程在指定范围内可被可靠概括，而不是它们从此不存在。

27.7 运动与变化为什么可以在同一规律中会合

改变初级电流是一条路径，移动磁体或回路是另一条路径。对运动的导体，相关电动势需要同时考虑电场与导体运动项：$\varepsilon = \oint (E_{\text{电}} + v \times B)\cdot dl$，其中 v 是回路元的速度。在适当的经典条件下，这与运动回路所链接磁通的总变化相联系。不能一方面固定积分曲面，另一方面把回路移动造成的面积变化悄悄漏掉。

一个简单例子是长度 ℓ 的导体棒沿两条平行导轨运动，棒、导轨和固定电阻构成闭路，匀强磁场垂直回路平面。忽略导轨电阻、自感及边缘效应，棒以速度 v 扩大回路面积，磁通变化率为 $B\ell v$，电动势大小为 $B\ell v$，电流大小为 $B\ell v/R$。这里即使 B 不随时间改变，回路本身的几何变化仍能产生感应。

取 B = 0.5 特斯拉、ℓ = 0.2 米、v = 1 米每秒、R = 2 欧，可得电动势 0.1 伏，电流 0.05 安培。导体棒受到阻碍运动的磁力，大小 IℓB 为 0.005 牛；维持匀速所需机械功率为 0.005 瓦，恰好等于电阻的 I^2R。这个能量检查约束了方向：若感应电流反而无条件帮助原运动，就会破坏这个被动装置的能量平衡。

但不能反过来认为能量守恒单独决定全部感应公式。它约束符号和功率对应，却不能代替磁通规律对任意几何给出的量值。只有把运动、磁作用、电路响应和机械功共同写出，才能完成此处的检查。统一并不是一句最一般原则包办所有具体关系，而是不同层次的规律彼此限制。

在铁环中，两绕组可以静止；在导轨中，部分回路运动。它们都能生成电效应，提示“磁生电”需要更精确的过程规定，却不说明两种装置的每一个局部机制都完全相同。以哪个参考系描述、导体怎样运动、电场如何分布，仍然需要具体交代。到第 31 章，磁体与导体相对运动的解释关系还会成为更深问题的入口。

27.8 如何击败一个过于宽泛的感应解释

对铁环中的短暂偏转，至少有三类竞争解释值得认真对待。第一是两绕组绝缘不佳，初级电源直接漏到次级；第二是寄生电容在通断时耦合脉冲；第三才是与磁链变化相联系的感应。三者都可能在某些安排中产生短暂信号，因此“只在开关时发生”本身不够。

可检验的区别在于它们怎样响应几何、材料、频率和连接变化。漏电通常应在相应绝缘路径与电位差下留下持续或可另行测量的传导；电容耦合与电位变化及电场结构相关；磁耦合则对链接磁通与绕组方向有相应依赖。实际装置可能同时具有三种贡献，研究任务不必预设只能选一个原因，而要估计何者在什么区间占优。

屏蔽试验也不是一招定胜负。导电屏蔽如果形成闭合回路，可能通过涡流改变磁响应；屏蔽改变布线，又会改变电容和互感。一个“加了屏蔽所以排除某机制”的判断，必须说明屏蔽实际抑制和引入了哪些作用。更可靠的策略，是让多种独立改变的结果同时符合一个参数受约束的模型。

还应区分基本规律与材料模型的失败。如果铁芯历史导致输出差异，而按实测磁通求导仍能解释电动势，受挑战的是常互感近似；如果测得的电动势与磁通关系在带宽、寄生耦合和回路定义都受控后仍系统不符，挑战才更直接指向所用感应描述。把两种失败都写成“法拉第定律被推翻”，会丢失实验真正提供的信息。

这种竞争检验也使本书的哲学主张受到限制。较弱解释可以说，实验只是发现既有感应规律，新的对象来自更好的测量分类；这种说法有相当力量。发生学分析的增量，必须具体落在分类怎样通过操作改变、何种短暂过程取得独立身份、哪些关系从可忽略因素进入构成条件，而不能仅给成功实验再加一个“创造”称号。

27.9 把零读数与非零读数共同纳入一个世界

铁环案例可以沿 D—S—E 的原型路径重建。先组织通电、保持和断电的差异运行 D，取得短暂偏转及稳态归零的显露 S，再把对象 E 从单纯“有磁作用便有电流”改造成磁链变化、回路响应与探针时间共同规定的过程。新的结构必须返回另一种斜坡、另一种负载及另一段观察窗口，预言新的显露。

这里三方程各有清楚的实质。S = F(D,E)要求由输入变化及电磁—仪器响应得到记录；D = G(S,E)要求依据波形异常选择延长尾部记录或改变磁化历史；E = H(S,D)要求通过这些改变判断应保留常互感、加入自感，还是引入材料记忆。三者联合，不容许把 E 当成事后塞入所有遗漏原因的袋子。

两维张力推动第三维的改变也可以逐项辨认。短暂 S 与持续输入 D 不相称，推动 E 引入变化率；已修订 E 与积分读数 S 不相称，推动 D 延长记录；短时操作 D 与含自感及仪器迟滞的响应 E 不能由单一峰值协调时，推动 S 从最大偏角改造成含完整尾部的响应曲线与积分电荷，再检验延长窗口后能否恢复预测。这个循环既让研究能够改变前提，也要求每次改变都提高区分和预测能力。

法拉第的铁环由此打开一个新的物理层次：状态变化本身开始具有可测量、可组合和可转换的规律，静止的空隙不再只是导体之间无事发生的地方。但它没有单独解答变化怎样穿过空间，也没有证明传播速度。下一章将研究麦克斯韦怎样把这种关系写进场的共同方程，使局部变化不仅形成另一回路的脉冲，而且能够成为在空间中演化的物理过程。

本章参考资料

[27-1] Royal Institution，“Michael Faraday's ring-coil apparatus”。馆藏日期、材料与 1831 年 8 月 29 日实验记录。https://www.rigb.org/explore-science/explore/collection/michael-faradays-ring-coil-apparatus

[27-2] Michael Faraday, Experimental Researches in Electricity, Vol. I, First Series, read 24 November 1831；1849 第二版转录。核读 § 5、10—18、27—43，正文只简述报告顺序与通断响应，现代方程另作说明。https://www.gutenberg.org/files/14986/14986-h/14986-h.htm

[27-3] Michael Faraday, Faraday's Diary, Thomas Martin 编，2008 重印预览版，1831 年 8 月 29 日条目。核读预览 PDF 第 37 页附近连接与磁针记录。https://faradaysdiary.com/ws3/faraday.pdf

[27-4] MIT OpenCourseWare, Physics II, Summary of Class 22: Mutual Inductance & Transformers; Inductors, 2007。互感、自感与感应公式；本章脉冲、积分量和电路解为独立演算。https://ocw.mit.edu/courses/8-02-physics-ii-electricity-and-magnetism-spring-2007/1762da3d55d1798e584c08bddb29235c_summary_w09d3.pdf

[27-5] MIT OpenCourseWare, 6.007 Lecture 9: Magnetic Materials, 2011。阅读非线性磁响应与磁滞有关部分。https://ocw.mit.edu/courses/6-007-electromagnetic-energy-from-motors-to-lasers-spring-2011/ac90133277835c7ae77d20793708fbc1_MIT6_007S11_lec09.pdf

第二十八章 麦克斯韦
位移电流怎样使场成为共同演化的系统

28.1 把一条实验关系推进为空间中的理论

法拉第的铁环说明，磁状态的改变可以在另一回路产生短暂电效应。但回路之间如何联系，变化怎样经过空间，导体以外是否也有需要描述的物理过程，仍不是一个互感常数能够完整回答的问题。互感把装置整体压缩成参数，场论则要求追问这个整体内部的分布和演化。

麦克斯韦 1865 年发表的《电磁场的动力学理论》，把周围空间纳入电磁过程的解释，并将电位移的变化与传导电流共同列入总电流。他又把电磁扰动的传播与光联系起来。[28-1] 这篇论文采用的变量、方程编组与机械解释并非今天教科书中的四行向量式。本章使用现代真空形式作推导，不能把这种重述冒充原文的全部面貌。

这一历史区分关系到我们怎样评价创造。若把过程写成“已有三条定律，补上一项就结束”，便会省略对象选择、媒质观念和电学量之间的重新组织；若只强调以太机械模型又会看不到，其中某些关系后来获得了超出原模型的生命。一个构想可以在解释上受时代限制，在数学结构上却打开更长远的问题。

场论的新意不只是给空间每一点添上几个数字。温度分布也有空间数字，地图也有坐标；场若要成为动力学对象，必须说明数值如何变化、哪些初始配置允许、变化如何影响其他对象，以及局部描述怎样汇成整体。位移电流的意义，正可以从这些相互要求进入，而不是从一个听起来神秘的名称进入。

28.2 稳恒电流公式为什么不能原封不动地处理充电

先从现代磁静力学的安培关系开始：$\nabla \times B = \mu_0 J$，其中 J 是电流密度。对两边取散度，由矢量微积分恒等式 $\nabla \cdot (\nabla \times B)=0$，得到 $\nabla \cdot J=0$。对于稳恒电流，这并不奇怪；局部没有电荷持续积累，流入与流出相平衡。问题发生在把同一表达直接用于随时间变化的电荷分布时。

电荷连续性方程是 $\partial \rho/\partial t+\nabla \cdot J=0$，$\rho$ 为电荷密度。如果 $\nabla \cdot J$ 被强制为零，就同时要求 $\partial \rho/\partial t$ 为零。电容器充电时，电荷在极板上增加，恰好不满足这个限制。这里不是电荷守恒与实验冲突，而是把稳恒关系扩展到不稳恒过程时，忽略了原公式的条件。

电场的高斯关系 $\nabla \cdot E_{电} = \rho/\varepsilon_0$ 提供一个可用连接。对它取时间导数，可得 $\varepsilon_0 \nabla \cdot (\partial E_{电}/\partial t) = \partial \rho/\partial t$。于是把安培关系写成 $\nabla \times B = \mu_0[J + \varepsilon_0 \partial E_{电}/\partial t]$，右侧散度成为 $\mu_0[\nabla \cdot J + \partial \rho/\partial t]$，与连续性相容。$E_{电}$ 表示物理电场，不能与本书 SDE 中的特征纠缠 E 混淆。

这是一个非常清楚的结构修复，却不能被夸成由逻辑唯一推出整个电磁学。连续性只限制新增项的散度；理论上还可以添加某些散度为零的项而不破坏这项相容性。若要选定具体规律，还需要局域形式、对称性、其他经验关系与新的预测共同约束。数学一致性排除一些错误方案，不自动证明剩下的方案就是自然界。

因此，位移电流既不是任意补丁，也不是只靠一条恒等式便无条件获得的结论。它的力量来自连接：与已知电荷关系相容，能用于充电过程，又能与感应定律共同产生传播预言。如果一个修复只让纸上的散度相消，却在其他实验中不断要求额外豁免，就还没有获得同样的理论地位。

28.3 同一条边界为什么不能有两个磁环量

电容器提供了一种几何上直观的检验。沿围住引线的闭合路径计算磁场环量，若选择穿过引线的跨越曲面，传导电流通过曲面；若将曲面鼓起，绕开导线并穿过极板间隙，理想绝缘间隙中没有相同的传导电流。两曲面有同一边界，左侧磁环量不能取两个互相冲突的值。

加入位移项后，环量关系成为 $\oint B \cdot dl = \mu_0 \int [J + \varepsilon_0 \partial E_{电}/\partial t] \cdot dA$。穿过导线的曲面以传导项为主，穿过电容间隙的曲面则由变化的电场提供相应贡献。在理想平行板近似下，$E_{电} = Q/(\varepsilon_0 A)$，故 $\varepsilon_0 A \, dE_{电}/dt = dQ/dt = I$。曲面的改变不再任意改变同一个回路的答案。[28-2]

这项等式不表示有同一批电子穿过真空间隙。传导电流描述电荷的运输，真空位移项描述电场的时间变化；它们在磁场方程中共同出现，不等于微观过程完全相同。若间隙中确实发生漏电或击穿，则传导项也应计入，不能仍按理想绝缘情形计算。

曲面独立性也不是只对画得漂亮的两张曲面成立。两张同边界曲面的差构成一个闭合面，对总电流密度取通量，其值由相应散度决定。连续性与高斯关系使这个闭合通量为零，才保证所有允许曲面得到相容结果。一个局部微分约束因此支持了全局积分的一致性。

从物理思想看，这比“空隙里也有一种电流”更精确。原来我们追踪的是导体中的电荷流，现在必须同时追踪充电所改变的空间电场；导体与间隙不能再用互不相干的两套规则处理。新对象是在共同约束中成立的，而不是仅因有人给空隙命名便获得存在。

28.4 一个充电模型中的空间预言

设理想圆形平行板半径为 a，板间距远小于 a，电场在极板投影内部近似均匀，充电安排具有足够的轴对称性。忽略边缘区并限于准静态阶段，在板间取半径 r 小于 a 的圆形安培回路。它所包围的位移电流为 $I\,r^2/a^2$，于是 $B(r)=\mu_0 I\,r/(2\pi a^2)$。场在中心趋零，随 r 线性增加，而不是在整个间隙中处处相同。

在相应理想化的外侧区域，若所包围的总位移电流为 I，公式转为 $B(r)=\mu_0 I/(2\pi r)$。内外表达在 r=a 处相接，却有不同的空间趋势。这个分段结果来自分布和对称条件，不能推广到任意接线、有限极板边缘或任意快速充电。实际引线破坏对称时，还须加上相应磁场贡献。

取 a=0.05 米、I=0.01 安培，在 r=0.025 米处，B 约为 2×10^{-8} 特斯拉，即 20 纳特斯拉。若把测量位置移到 r=0.05 米，理想式给出 40 纳特斯拉。这里电流完全相同，读数因几何位置改变；只用总电流而不说明场分布，便无法把两个读数看成同一理论的结果。

板面积约为 0.00785 平方米，取 ε_0 约为 8.85×10^{-12} 法每米，电场增长率 $I/(\varepsilon_0 A)$ 约为 1.44×10^{11} 伏每米每秒。若板间距为 0.001 米，100 纳秒内相应电压增量约为 14.4 伏。这个数值只是理想、短时、恒流充电的演算，不是对历史装置的重建，也不包含实际供电回路的全部建立过程。

准静态要求同样可以估计。极板半径对应的真空传播时间约为 a/c，即 0.17 纳秒；若驱动变化时间远长于这一量级，忽略许多空间延迟可能合理。若试图在比它更短的时间内建立整个板面均匀电场，上述模型本身便不再自洽。集中参数电容不是说所有部分真的同时知道电源做了什么，而是把足够小的延迟有效省略。

28.5 四组方程如何共同承担演化

在真空、经典连续场描述中，现代方程可写为：$\nabla\cdot E_{电}=\rho/\varepsilon_0$，$\nabla\cdot B=0$，$\nabla\times E_{电}=-\partial B/\partial t$，$\nabla\times B=\mu_0 J+\mu_0\varepsilon_0\,\partial E_{电}/\partial t$。前两组规定散度约束，后两组联系空间旋度与时间变化。电与磁在其中仍保留源结构的差别：电荷进入电场散度，经典无磁单极模型则把磁场散度置零。所谓统一并不要求两种场的每个位置都填入形式对称的项；如果仅为了整齐而虚构一类未获支持的源，数学美感就越过了经验责任。

这不是四条互不相干的口号，任意指定电荷、电流和场而不检查彼此相容，通常不会得到一个物理解。

例如，对法拉第关系取散度，得到 $\partial(\nabla\cdot B)/\partial t=0$。若初始时没有磁散度，之后的演化保持这个条件。对修正后的安培关系取散度，并使用电荷连续性，可得 $\partial[\nabla\cdot E_{电}-\rho/\varepsilon_0]/\partial t=0$。高斯约束若一开始成立，也会由相容的演化保持。初始条件与动力学因此不能被分成两套互不检查的规定。

这项结果还说明，仅有演化方程并不足以挽救错误初值。若一开始就塞入违反约束的场，演化不会自动保证它变成正确物理状态；某些不相容残差可能继续存在。数值计算中出现看似平滑的波形，也不意味着所算配置具有合格的电荷与场关系。稳定运行和物理一致性是两项需要分别验证的要求。

还必须给定边界与源的历史。金属表面、开口空间、导波结构会限制允许解；相同微分方程在不同边界下，可以产生传播、驻波或局域响应。方程不是一张不需要情境的世界照片，而是组织一族可能过程的规则。规则的普遍性，恰表现为它能够处理明确不同的边界，而不是假装边界从来不重要。

在这里，场的身份已经超过探针受力表。即使某处没有放置测试电荷，该处场也须与邻近区域及初始条件协调，并影响后来的可测过程。探针可以揭示它的某些作用，却同时可能扰动源和边界；把场定义成“此刻放入一只探针所得读数”只提供操作入口，不能替代全部演化内容。

28.6 从耦合变化推导波动，而不是让两个场互相叫醒

在没有电荷与电流的真空区域，取感应方程的旋度，有 $\nabla\times(\nabla\times E_{电})=-\partial(\nabla\times B)/\partial t$。利用恒等式 $\nabla\times(\nabla\times E_{电})=\nabla(\nabla\cdot E_{电})-\nabla^2E_{电}$，以及该区域 $\nabla\cdot E_{电}=0$ 和修正后的安培关系，得到 $\nabla^2E_{电}=\mu_0\varepsilon_0\partial^2E_{电}/\partial t^2$。磁场以同样方式满足波动方程。[28-2]

这是联合关系产生的新结论。一条方程约束电场的旋转结构与磁场变化，另一条约束磁场的旋转结构与电场变化；合起来才形成空间变化与时间二阶变化之间的传播关系。波速为 $c=1/\sqrt{(\mu_0\varepsilon_0)}$。如果删去真空位移项，真空中的磁旋度便失去对应的时间变化项，上述传播推导也就无法按原样成立。

常见比喻说电场生磁场、磁场再生电场，像两个人轮流推动。这可以帮助记住耦合，却容易制造错误时间图像。对于单向真空平面波，电场和磁场可以同相变化，并非

一个达到最大后另一个才开始。它们是同一解中相互约束的分量，不是各自独立的物体在相邻时刻互相创造。

更不能从耦合关系推断空无一物会自动产生波。若初始场为零、无源且边界没有输入，在适当唯一性条件下零解持续成立。要有非零辐射，必须有非零初始条件、过去源的活动或边界激发。传播区域无局部电荷，不等于波没有起源；局部无源与全部历史无源，是两种不同判断。

这种区分对本书的创造命题同样有约束力。一个新的理论对象并不因为方程可以写出便在任何地方自动发生。物理创造需要实际制备、演化与边界条件；知识创造则需要把这些条件联结成可追踪的对象身份。无论强调哪一个层级，都不能用“关系相互生成”删除初值和能量来源。

28.7 横向关系给统一增加了哪些负担

考虑沿 z 正方向传播的简谐平面波，令电场沿 x 方向，大小为 $E_0\cos(kz-\omega t)$。散度约束要求这种均匀横截面平面波的振荡电场没有沿传播方向的独立分量；磁场也须横向。由旋度方程可得 B 沿 y 方向，幅值 $B_0 = E_0/c$，且 $\omega=ck$。传播方向、电场方向和磁场方向按相应右手关系协调。

如果取频率 100 兆赫，波长约 3 米；取电场幅值 3 伏每米，则磁场幅值约 10 纳特斯拉。这个例子同时预言频率—波长关系及两场幅值比。只测到某个周期性电信号，尚不足以确认这种自由空间平面波；近场、导线信号或探针自身振荡，都可能具有相同频率，却不满足完整空间关系。

麦克斯韦在原论文中比较电学测量给出的速度尺度与独立光学测量，并将横向扰动与光联系起来。[28-1] 这里的关键并非今天把两个常数相乘后得到一个熟悉数值，而是来自不同实验链的量值开始共同限制一种传播对象。若先用光速定义某些单位，再把相应数值代回公式，就不能把同一个关系重复计算成第二份独立证据。

速度相近仍不是全部统一。光还有偏振、反射、折射和材料响应，电磁理论必须在适当条件下解释这些关系。一个候选波速巧合地接近光速，并不能单独排除所有竞争机制；与横向性质、边界效应及更广实验共同符合，才使联系逐步坚固。下一章的赫兹实验将把这类要求转成电振荡装置能够直接检验的问题。

横波结论也有范围。在导体附近、带电区域、波导或某些介质中，局部电磁场不必呈现无源真空平面波的简单横向形式。若看到纵向电分量，就宣布麦克斯韦方程被否

定，是把特殊解的性质冒充整套方程的普遍约束。理论越统一，越需要清楚区分一般规律与选定几何的解。

28.8 介质与传导使“变化项”不能随意忽略

在物质中，电位移通常写为 D 电 = ε₀E 电+P，P 为极化。在宏观描述下，位移电流项为 ∂D 电/∂t；它包含真空项和极化变化，但不意味着所有物质响应都与 E 电瞬时成比例。若使用 D 电=εE 电，便已作出线性、相应频率范围内适用等假设。这里的 D 电同样不等于 SDE 的差异序列 D。

设一个均匀线性介质在所考察频率附近可用实常量 ε 与电导率 σ 近似，简谐场角频率为 ω。传导电流幅值为 σE₀，位移项幅值为 ωεE₀，二者比为 σ/(ωε)。因此，是否可以忽略位移项，不由材料叫不叫绝缘体一句话决定，而由时间尺度与材料响应共同决定。

作一个纯假想介质例子，取 ε=4ε₀、σ=10⁻⁸ 西门子每米。在 1000 赫附近，这个比值约 0.045；若同样参数还能用于 1 赫，则比值约 45。前一情形变化项较大，后一情形传导项较大。但实际材料的 ε 与 σ 可能随频率改变，后一计算是条件推演，不能不经测量就当作某种具体物质的频谱。

这也解释了为何“微小修正”可能改变问题类型。某一慢过程里很小的项，在另一时间尺度上成为主要项；保留与删除它，不只是数值精度差别，还可能决定模型是否支持某类传播。正确的近似应说明比值及预期误差，而不是把曾经有效的省略永久化。

媒质模型的失败必须与基本场方程的失败分开。如果折射和衰减不符合常 ε 模型，可以先检查色散、吸收、各向异性及空间非局域响应。只有把这些可独立测试的条件纳入后，才知道偏差指向哪里。统一理论并不取消材料研究；它反而提供共同框架，使材料的不同响应能够被精确表达。

28.9 边界不是方程外面的一圈空白

还可以把连续性要求推进到材料表面。设一块薄片存有表面电荷密度 σ 面，两侧法向电流不相等，薄片上的电荷就必须改变。若另有沿表面流动的电流，也须计入它的表面散度。一个边界并不会因为在图上没有厚度就免除守恒；它只是把某些体分布压缩成面分布，仍需独立说明允许哪些积累。

最简单的例子是一束电流进入电容极板，却没有传导电流穿出绝缘间隙。电荷留在表面，改变两侧电场法向分量；高斯关系使电场的跳变与表面电荷对应。若一边计算极

板电荷增加，一边强行令两侧法向电场始终相同，就会人为制造不相容。源的积累、场的边界跳变及电流连续性必须一起成立。

另一个独立演算可以检验传导的时间尺度。考虑均匀线性欧姆介质，取 J=σE 电、∇·E 电=ρ/ε，并假设 σ 与 ε 为常数。连续性给出 ∂ρ/∂t=−(σ/ε)ρ，于是局部体电荷按 exp(−t/τ 电)衰减，τ 电=ε/σ。这不表示电荷在介质内部凭空消失，而表示它经电流迁移，常会重新分布到边界。

沿用上一节假想参数，τ 电约为 0.00354 秒。若只观察远短于这个时间的过程，介质来不及按该模型消除体电荷；若等待许多个时间常数，体内部才可能接近相应无自由体电荷状态。以静电平衡图像描述所有快速过程，和以高频绝缘近似描述任意漫长等待，都可能错误。这个时间常数与前述 σ/(ωε)互为尺度表达：它们共同问的是驱动给材料留下多少重排时间。

当然，真实介质若有扩散、陷阱、非线性或电极反应，单指数松弛便可能失败。不能先用未检验的欧姆关系推出指数，再把任何偏离都称为违反电荷守恒。相同的守恒方程可以与不同材料动力学组合，产生不同松弛曲线。基本约束与构成模型的区别，正是统一理论容纳多种物理对象而不变得任意的条件。

28.10 场的实在性不能由一幅力线图裁决

一个强竞争解释认为，场只是整理源之间作用的方便记号；只要某种直接作用理论能够给出相同可观测结果，没有必要把每一点场值都实体化。麦克斯韦本人面对的也不是毫无成绩的旧理论，他在原论文开篇承认韦伯等方案组织多种现象的能力。[28-1] 因此，不应把历史竞争写成有方程的真理对没有计算的迷信。

本书的判断是，局部场描述取得了新的解释能力：它把过去源的活动、空隙中的传播、后来的探测与边界响应组织成连续可计算的过程。但若两种数学表述在明确条件下完全等价，单凭语言偏好无法宣布实验已经裁决全部本体论。需要寻找不同预测、不同适用边界，或说明一种表述怎样增加可计算与可控制的过程。

这一判断也限制 SDE 的使用。物理电磁场不是因为具有空间分布便自动等于特征纠缠场 E；前者有具体量纲、方程和测量定义，后者分析属性如何共同规定。二者的联系，应落实为电流、充电、磁环量、传播与探测怎样形成一个不能任意拆开的规定系统，而不是仅凭同一个“场”字发生同一化。

可选择 E—D—S 的沉淀路径：从场方程与源条件出发，设计充电及不同半径的磁探测，事前计算幅值和时间关系，再检查显露。若 S 与 E 不符，先推动 D 改变探针位置、

带宽或接线；若不同 D 都保留同一偏差，再推动 E 中的对称近似或材料响应改造。改造以后必须返回未参与拟合的测量，不能只解释旧图。

三方程分别要求结果能由构成与操作产生、操作能由已有结果与候选构成选择、构成能由跨操作结果重建。两维张力推动第三维并回检，在本章不是审美和谐，而是对曲面一致性、时间尺度与幅值关系的严格要求。数学上的相容性打开可能世界，独立实验则决定哪一种可能性获得物理资格。麦克斯韦的新世界正是在这两种要求的交会处形成，而不是在四行符号写完的瞬间自动完成。

本章参考资料

[28-1] James Clerk Maxwell, “A Dynamical Theory of the Electromagnetic Field”, Philosophical Transactions 155 (1865):459–512。核读 Part I §1—3、11、18—20；Part III §55 的总电流；Part VI §95—97 的横向传播与速度比较。现代向量推导不作原文逐字还原。https://en.wikisource.org/wiki/A_Dynamical_Theory_of_the_Electromagnetic_Field/Part_I

https://en.wikisource.org/wiki/A_Dynamical_Theory_of_the_Electromagnetic_Field/Part_III

https://en.wikisource.org/wiki/A_Dynamical_Theory_of_the_Electromagnetic_Field/Part_VI

[28-2] David Tong, Electromagnetism, Chapter 4, Electrodynamics。阅读位移电流、Maxwell 方程及波动关系；本章充电数值、约束保持、媒质尺度比和方法分析独立展开。https://www.damtp.cam.ac.uk/user/tong/em/em3.pdf

第二十九章 赫兹
怎样使看不见的电磁波取得实验身份

29.1 一个波动解还不是一项波动证据

麦克斯韦方程容许电磁扰动以有限速度传播，但数学上能够写出一种解，并不保证某个实验室已经生成它，更不保证手边的探测器能够认出它。源怎样激发、哪些频率足够强、接收器对什么响应、房间边界怎样改变场，都是从理论可能性进入经验对象必须完成的工作。

赫兹在十九世纪八十年代后期的研究，将快速电振荡、接收回路和空间比较组织起来。《电波》所收论文分别讨论传播速度、空气中的波及其反射，后来又讨论具有光学性质的电辐射。[29-1] 这些研究不宜压缩成某一天偶然看见第二个火花的传奇。火花是证据的入口，电磁波则是经过一组差异检验才成立的对象。

尤其需要区分三个命题。远处接收器受到电作用，说明源与接收器之间存在耦合；作用具有波动空间结构，说明不能只用某种随距离减弱的静态影响概括；传播与麦克斯韦理论给出的速度、方向及边界关系相容，才进一步支持电磁波解释。第一个命题并不自动包含后两个。

本章将使用简化振荡电路和波模型，重建这些命题之间的证据阶梯。所有具体数值均为本书假设算例，不充当赫兹装置参数的复原。原典的意义在于约束历史主张，现代计算的意义则在于把当时研究打开的问题展开为读者能够独立检查的关系。

29.2 发射器的火花并不是波的每一次振动

一个火花间隙可以在适当条件下使原来积累的电荷迅速放电，从而激发电路的快速振荡。但可见火花的重复频率，与一次放电内部电荷往复变化的频率，可以相差很多。将耳朵听到的噼啪次数直接当成电磁波频率，会混淆激发事件与被激发的振荡。

用集中参数近似，设电路电荷 q 满足 $Lq''+Rq'+q/C=0$。这里 L、C 是有效电感和电容，R 代表适用频带内可近似处理的损耗。欠阻尼条件 $R^2<4L/C$ 成立时，解含有 $\exp(-Rt/2L)$的衰减包络和角频率 $\omega_{阻}=\sqrt{[1/(LC)-R^2/(4L^2)]}$的振荡。只有损耗足够小时，频率才近似为 $1/(2\pi\sqrt{LC})$。

取 L=1 微亨、C=100 皮法、R=10 欧，理想无阻尼频率约 15.9 兆赫；衰减时间 2L/R 为 0.2 微秒，约对应三个振荡周期。不同放电的起始时刻若有抖动，不经同步便直接平

均，会把快速振荡洗掉，却未必消除包络。平均后没有载波，不等于每次激发都没有振荡；还须说明记录如何对齐。这个波包不会包含无限多次完全相同的振动。若用单频正弦波来分析它，只能作为描述中心频率附近行为的近似，不能忘记有限持续时间带来的频谱宽度。

这也表明，发射器不是一只把已有波取出来的盒子。它通过电荷、几何、损耗和激发方式共同生成可传播的扰动。改变某一尺寸可能同时改变电容、电感和辐射损耗，不能只挑其中一项解释。等效电路的价值是把这种联合影响压缩成可测参数；它的边界则是，当导体尺度与波长可比时，电压、电流不再处处同相，简单集中参数描述需要检查。

若把电容增大四倍而保持其他有效参数可比，低损耗近似预言频率减半；若增加阻尼，波包更快衰减，过大时甚至不再振荡。不同改变对频率与包络的作用不同，因此可以形成独立诊断。一个只看到“火花更亮”的观察，无法告诉我们频率变了、能量变了，还是放电阈值变了。

29.3 接收器选择什么，实验就容易看见什么

接收回路也有自己的动态结构。若在弱耦合近似下，它受到角频率 ω 的外部电动势驱动，电流幅值可写为 $I_{接}=\varepsilon_{驱}/\sqrt{[R_{接}^2+(\omega L_{接}-1/(\omega C_{接}))^2]}$。在串联理想模型中，感抗与容抗抵消时，电流响应最大。这说明同一个空间场可以因接收器参数不同，变成明显或几乎不可见的信号。

设接收器在 ω_0 处谐振，定义品质因数 $Q_{品}=\omega_0 L_{接}/R_{接}$。对于高品质因数和小失谐，若 $\delta=(\omega-\omega_0)/\omega_0$，归一化幅值近似为 $1/\sqrt{[1+(2Q_{品}\delta)^2]}$。当 $Q_{品}=10$、失谐百分之五时，幅值约为谐振值的 0.707；若某火花间隙只有在更高电压下才击穿，这个下降足以使火花突然“消失”。

火花消失不等于场严格为零。击穿是有阈值、受电极形状及条件影响的响应；它能够把连续的场强变化转成近似二值的可见结果。若每次移动接收器都重新调间隙直到看见火花，却没有记录调整量，就会悄悄改变探测标准，使空间强弱图失去共同尺度。

现代研究可以记录电压波形或频谱，历史装置则使用当时可获得的指标。两者的证据结构有共同问题：接收器响应必须与所探测对象区分。仪器越敏感，并不自动意味着它越忠实反映全部频率；尖锐共振可以放大弱信号，也可以遮蔽宽频输入，只留下接收器愿意接受的一小段。

因此，接收器既参与对象的显露，也受到对象约束。改变它的尺寸或电容，若所见主要频率随之改变，就必须区分源频谱、空间模式与接收选择。若不同接收器在校正响应后恢复相同传播关系，波的身份才获得超出单一仪器的稳定性。把实验者、接收器和电磁场分别安排到 S、D、E 三个字母，反而会错过这种完整的共同规定。

29.4 近处点亮接收器还不能证明远场传播

一个重要竞争解释是普通感应耦合。上一章之前，我们已经知道两个邻近线圈可以交换能量；因此，接收端出现电流甚至点亮灯泡，并不单独证明它处在自由传播的远场。源附近的电磁结构可以含有随距离较快衰减的分量，不能一概套用平面波关系。

一个有用的尺度是 kr，其中 k=2π/λ。在简单小源的讨论中，kr 远小于一往往属于近场需要认真处理的范围，kr 远大于一才可能进入辐射项占优的区域；实际天线尺寸和方向还会加入限制。对于频率 3 兆赫、波长约 100 米的过程，距离 1 米时 kr 约 0.063。接收器即使已经明显响应，也不能仅以“隔了一米空气”便宣布完成远场测量。

Faccio 等人在 2006 年的赫兹实验重访中，明确把一套现代装置用于近场衰减和耦合共振的研究，并说明该装置不便直接重现波长测量。[29-2] 这提供了一项重要提醒：演示成功的内容应由实际尺度与测量方式决定，不能因为演示名叫赫兹实验，就把历史研究的全部结论一次搬来。

在本章的独立模型中，若磁偶极近区的主要磁场随 $1/r^3$ 变化，将距离加倍会使这一幅值约降到八分之一；而固定方向、足够远处的辐射场幅值通常以 1/r 为主要衰减，距离加倍约减半。实际测量还受有限源、反射、方向图和探针大小影响，所以两个距离点不能自动判定幂律。需要在可说明的区间内比较一组距离，并检验模型残差。

这里的区别不是“近场不真实、远场才真实”。两者都是电磁场解的不同部分，只是承担不同传播与交换结构。对电磁波实验的判断，关键是不能拿一种区域的证据替代另一种区域的结论。近场实验可以准确检验耦合，远场实验可以检验辐射传播；统一理论允许它们相联，却不抹去测量任务的不同。

29.5 让反射产生一把空间标尺

赫兹关于空气中电磁波及其反射的原文，利用直接与反射作用的比较研究空间分布；后来的电辐射研究又通过反射器和方向变化建立与光学的联系。[29-1] 这条路线的力量在于，波不再只作为某个探针中一闪而过的事件，而开始在空间中留下有组织的相位关系。

取一个理想化例子：平面波垂直入射到位于 z=0 的理想导体表面，电场平行于表面。为使表面切向电场为零，反射电场与入射电场在边界相消。两个反向传播波相加后，电场可写为一个与 sin(kz)成正比的空间因子，乘以随时间振荡的因子。因此，相邻电场节点之间的距离为 λ/2。

假设一次模型实验中相邻节点相距 0.75 米，可推得波长 1.5 米。若另有独立频率测量给出 200 兆赫，则相速度 fλ 约为 3×10^{8} 米每秒。这里每一步都有证据责任：节点究竟对应电场还是磁场，距离是相邻同类节点还是峰谷间距，频率来自独立标定还是已用光速倒算，都影响结论。

若将相邻节点与波腹的距离当作半波长，结果便会错一倍。理想站波中电场节点与磁场节点也错开，位置比较需要说明探针耦合什么。一个小环主要对链接磁通敏感，一根适当取向的短导体则可以主要耦合电场；实际装置可能混合响应。探针名称不能代替响应分析。

真实反射并不总是全反射。设入射与反射幅值之比中的反射幅值系数大小为 η，小于一。两波叠加的最大幅值与最小幅值之比为(1+η)/(1−η)。若 η=0.6，比值为四；最小点并不为零。如果接收阈值恰好高于最小幅值，就会看见一段没有火花的“暗区”，但这不是严格节点的直接图像。

反过来，阈值太高也可能使两个相邻波腹之间的大段区域都变暗，节点位置因观察者判定而漂移。较好的程序是多次改变阈值或使用连续响应，检查极小位置是否稳定。几何和模型应帮助提取相位结构，而不是把一个模糊的暗区中心宣布为无误差的数值。

探针的有限尺寸还会把空间图样作平均。设一个理想线性探针沿扫描方向均匀覆盖长度 d，读数与这段区域内 sin(kz)的平均成正比。以探针中心 z_0 为坐标，直接积分得到 $\sin(kz_0)$乘以 sin(kd/2)/(kd/2)。当 d 远小于波长，后一个因子接近一；当 d 为半个波长，它变成 2/π，约 0.637。幅值被压低，即使中心位置和真实波形都没有变化。

在这一个完全对称的平均模型里，节点位置保持，但非均匀灵敏度、探针倾斜或电磁混合响应可以进一步改变极小位置。因而，缩小探针后若波腹变高，未必是发射器更强；若极小点移动，也需先检验取样核是否改变。把分辨率当成理论外面的纯技术细节，会使对象的空间结构与读取它的尺度发生混淆。

29.6 偏振把方向选择变成另一项检验

波长和传播速度之外，横向方向结构提供独立约束。设局部近似线偏振电场的方向固定，一根足够短的线性接收天线与它夹角为 θ，其有效驱动幅值在理想近似下与 cosθ

成正比。若负载响应线性且接收功率与驱动平方相联，则功率随 $\cos^2\theta$ 变化。转过直角后信号大幅削弱，因而不只报告“哪里有作用”，还报告作用的方向。

这里仍有必要限定。实际接收器在正交方向可能保留有限响应，反射可能改变偏振，天线转动也可能改变周围耦合。因此，未见绝对零点并不必然否定横波；见到一次零点也不够。应比较完整角度关系、反转后的重复性及不同位置的结果，检查哪些偏离可以由已标定的仪器解释。

对于假设的理想电场天线，转到 60 度时幅值减半、功率降到四分之一。若火花长度不是线性电压指标，就不能期待它也恰好减半；若只记录有无火花，更不能直接把出现概率当作 $\cos^2\theta$。理论预言的是某个被定义的量，实验必须建立该量与实际可见指标之间的连接。

偏振也帮助击败一些过于宽泛的竞争解释。例如，若所有接收效应仅来自房间中无方向性的声振动，就难以同时说明电天线的可重复角度响应、反射关系和频率选择。这个论证不是说声学没有偏振或方向现象，而是说某个具体声学干扰模型必须对同一组定向测量作出定量预期。真正的竞争发生在明确模型之间，而非两个抽象名词之间。

赫兹原文关于电辐射的讨论，将接收安排的转动与偏振作比较，并注意实验直接识别的电作用与理论中的磁分量有区别。[29-1] 这种克制值得保留。不能因为理论说两个分量都存在，就宣称一种探针已经把两者分别测量；联合理论可以支持推断，独立探测仍是另外一种证据。

29.7 两个共振器一起工作，会改变彼此的频率

如果发射器与接收器耦合较强，把接收器当作完全被动的旁观者就会出错。设两只相同的无损 LC 回路，自感为 L、互感为 M、电容为 C，电荷分别为 q_1、q_2。选择一致方向后，方程为 $Lq_1''+Mq_2''+q_1/C=0$，以及 $Mq_1''+Lq_2''+q_2/C=0$。这是一个适用范围明确的双振子模型，不包括全部辐射与火花非线性。

取同相组合 $q_1=q_2$，得到角频率 $\omega_{同}=1/\sqrt{[(L+M)C]}$；取反相组合 $q_1=-q_2$，得到 $\omega_{反}=1/\sqrt{[(L-M)C]}$。对于正 M 且 $M<L$，原来单回路的一个频率分裂成两个。
若 $M/L=0.1$，同相频率约为单回路频率的 0.9535 倍，反相约为 1.0541 倍。

这个结果直接反对一种简单判断：接收频谱出现两个峰，未必意味着空间中原来就有两种互不相关的辐射源。它可能是源与探针组成的联合系统具有两个模式。随着距离增加或取向改变，M 减小，两频率趋近；这个趋近关系提供可检验的区分，不能仅靠事后给两个峰分别起名字。

在有阻尼的实际记录中，两个近频模式还会产生拍振，包络起伏看起来像周期性增强和减弱。若仪器带宽不足，只看包络，可能把拍频误当作电磁振荡频率。若火花只在包络最高时出现，可见事件序列又会进一步选择信号。必须把激发重复、载波振荡、模式拍频与阈值越过区分，才知道哪种周期正在被测量。

这并不说明仪器参与便使客观测量不可能。我们可以降低耦合，独立测定单回路参数，改变接收器位置，再检查模式变化是否符合联合模型。被研究对象与探针之间的联系越明确，越能够计算并控制测量反作用。客观性在此来自可公开改变的关系，而不是假设存在一个永远不改变对象的理想眼睛。

29.8 真实历史中的速度难题不能删去

赫兹的早期比较中，空气与导线附近传播所对应的速度关系曾出现困难。O’Hara依据往来信件及后续实验，讨论了赫兹如何承认不一致、要求复验，并追查装置周围条件。[29-3] 本章不把这段过程写成所有偏差已经由一个简单原因当场消除，也不把早期数据等同于现代高精度光速测量。

这项困难对思想史的重要性，不在于给成功故事加一点曲折。它表明建立波的存在，与精确测定传播速度，是可以具有不同完成程度的任务。反射和驻波可能已经强烈支持有限传播的波动结构，而频率估计、接收器选择或复杂场分布仍影响数值。一个理论获得实质支持，不意味着每个测量值已毫无争议。

若现代重建一个类似研究，应当把误差链拆开。节点位置的不确定度进入波长，源频率的不确定度进入速度，二者的相关性则决定能否简单合成。设 λ 由独立距离测量得到，相对不确定度百分之二，频率独立测量相对不确定度百分之一，在小误差近似下，速度相对不确定度约为 $\sqrt{(0.02^2+0.01^2)}$，即百分之二点二四。

若频率其实是用同一波长和预设光速反算，上述独立合成便无效；随后再乘 $f\lambda$ 只会找回预设值。这种循环在叙述中很容易被遮蔽，因为每一步看起来都有公式。必须追溯每个量怎样获得，才能判断一致性究竟是新证据，还是相同假设的另一种写法。

还需区分理论与辅助假设。若有限房间造成多重反射，修正空间模型后可能改善节点估计；若接收器强共振偏选某一频率，需要重建响应函数。这些修正有独立后果，例如移动反射面或改变接收参数后图样应怎样变化。若修正只能事后解释旧结果而不能预言新变化，它就不能仅凭“实验复杂”得到无限宽容。

29.9 波的证据怎样形成一张不能随意松开的网

设有两种候选。甲是传播波模型，乙是某种近似瞬时、只随距离单调减弱的耦合模型。在一个固定位置，两者都可以调整幅值以产生接收火花。于是单点火花几乎没有区分力；但引入反射面、扫描空间极小点、独立改变频率，再比较节点间距是否按相应波长改变，便会让乙承担更困难的预测任务。

强竞争者当然可以更复杂。它可能容许延迟、共振或与场理论等价的直接作用关系。对这种方案，简单宣称“观察到波，所以所有直接作用理论都错了”并不严谨。实验首先限制的是具体传播关系，不是所有哲学词汇。若不同表述得到完全相同可观测量，需承认本体解释尚有余地；若它们对反射或偏振给出不同结果，才可安排裁决。

从本书的角度，赫兹实验创造的不是人眼原本看不见、现在突然能直接看见的一串波峰，而是一套可运行的波身份。源可调，接收有响应，空间比较能复做，反射和方向关系能计算，频率与波长能相互约束。这套身份把多个原来分散的操作连成共同对象，也使研究者能够据此设计此前没有的传播与探测方式。

这种创造命题有明确可败条件。如果所谓节点全跟着某只接收器内部参数走，却与反射面位置和独立源频率无关，就不能把它们直接当作源波长；若角度效应随电极缺陷转动，而不是随所测场方向变化，就应修正探针模型；若重复装置不保留任何共同关系，则必须承认对象规定尚未稳定。不能因为理论预言了波，就保证每一个模糊信号都属于波。

29.10 用异常节点走通一次再创

设想空间扫描得到的极小间距随接收器大小改变。可以沿 S—D—E 的再创路径推进：从异常间距这一显露 S 出发，组织不同接收器、不同源频率以及不同反射距离的运行 D，再区分“接收器选择了谱中另一段”“耦合改变了模式”和“传播介质改变了波长”等候选结构 E。这三种解释各有不同预言，不能只说整体条件变了。

三方程在此共同工作。S = F(D,E)要求记录由传播与接收共同生成；D = G(S,E)要求下一次改变能区分候选，而不只是增加照片；E = H(S,D)要求模型从跨操作的稳定关系重建。若已有 S 与传播 E 不相容，推动 D 加入独立频谱测量；若 D 与 S 的对应排除单一模式，推动 E 加入耦合或多路径；若快速扫描 D 与慢响应 E 不相容，则推动 S 改用经标定的稳态幅值或完整时间记录，并回检原节点判断。

这里没有哪一维可以永远免检。理论可以指导我们寻找节点，却不能指定节点必然在那里；仪器可以形成显露，却不能把自身共振偷偷改称外界规律；操作可以改变对

象，却必须记录改变的实际后果。三原理的内容是两维张力对第三维的具体推动及返回，而不是让实验、理论和语言最后显得彼此协调。

电磁波由此成为一个新的可制备、可比较的物理世界。它的存在不依赖人是否在接收端看到火花，但我们能够确认什么波、以何种方式传播，离不开源、空间与探测之间经过检验的关系。下一章将继续追问：当一个扰动已经离开发射器，却尚未到达吸收体，能量和动量在哪里？要回答这个问题，场不仅需要传播方程，还需要承担局部收支的责任。

本章参考资料

[29-1] Heinrich Hertz, Electric Waves, D. E. Jones 授权英译，Macmillan, 1893；核读 Dover 1962 原样重印本，论文 VII pp.107—123、VIII pp.124—136 相关实验段及 XI pp.172—185 的反射与偏振讨论。https://commons.princeton.edu/josephhenry/wp-content/uploads/sites/71/2020/02/Electric_Waves.pdf

[29-2] D. Faccio, M. Clerici and D. Tambuchi, “Revisiting the 1888 Hertz experiment”, American Journal of Physics 74 (2006):992–994, doi:10.1119/1.2238886。阅读现代装置的近场、耦合频率及测量范围；本章数值和双回路推导另行构造。https://www.physics.gla.ac.uk/xtremeLight/publications/AJP-Hertz-Facio(2006).pdf

[29-3] James G. O’Hara, “A ‘Horrible Conflict with Theory’ in Heinrich Hertz's Experiments on Electromagnetic Waves”, European Review 15 (2007):545–559, doi:10.1017/S106279870700052X。阅读开篇历史问题及赫兹与菲茨杰拉德通信的讨论，正文限于速度争议、复验及装置条件。https://oharas.com/Hertz/OHARA.pdf

第三十章 从作用到场
能量与动量怎样取得局域的历史

30.1 能量离开以后，到达以前

赫兹的实验使传播成为可操作的问题，接下来的问题却并不只属于波速测量。设一个发射器已经停止供能，而远处的吸收器还没有收到脉冲。在这段时间内，能量在哪里？如果只在两个物体的账本上登记，发射器的账上已经减少，吸收器的账上尚未增加，两个账本相加便出现暂时的缺项。坚持总能量守恒，可以说还有一个没有登记的储库；但要进一步说明储库怎样变化、能量通过哪个边界、什么时候抵达，就需要比一个总数更细的物理结构。

场论的推进正在这里。场不再只是计算远处物体受力的方便表格，而取得了自身的状态、储存和传输能力。这里的“自身”并不意味着场能脱离源、边界和测量而独立自足。它意味着在给定初始条件之后，即使源的当前状态已经改变，空间中的场仍可保留先前作用的结果，并继续对后来遇到的物体发生作用。一个脉冲因而拥有沿空间展开的历史。

坡印廷在一八八四年发表的论文，把研究重心明确移到能量从电源经周围空间传往耗散部位的路径，并从麦克斯韦理论提出局域能流的表达。[30-1]这个问题不应被压缩成又发现了一个向量。向量的意义在于，它把“总量没有少”改造成可以对任意局部区域逐项核查的收支。以下采用现代真空形式与国际单位制推导，不把现代符号、后来的动量表述或本章算例倒写成原论文原貌。

30.2 把总账拆成每一个小体积的账

设电场写作 E 电，磁感应强度写作 B，电流密度写作 J。E 电与本书方法论中的 E 不是同一符号对象。在真空形式的麦克斯韦方程中，将安培—麦克斯韦方程除以 μ_0 后与 E 电作点乘，再减去法拉第方程与 B／μ_0 的点乘。电场变化项给出 ε_0E 电·∂E 电／∂t，磁场变化项给出 B·∂B／∂t 除以 μ_0；它们恰好分别是两个平方项对时间的导数。

定义场能量密度和能流密度为

$$u=\varepsilon_0 E_{电}^2/2+B^2/(2\mu_0),\quad J_{能}=E_{电}\times B/\mu_0。$$

利用散度恒等式 $\nabla\cdot(E_{电}\times B)=B\cdot(\nabla\times E_{电})-E_{电}\cdot(\nabla\times B)$，得到

$$\partial u/\partial t+\nabla\cdot J_{能}=-J\cdot E_{电}。$$

这个等式的每一项都有清楚的去向。第一项是单位体积内场能储存的增减，第二项是净流出，右边是场向带电物质交付功率的负值。这里 J 包括所讨论微观描述中的电荷运动；若改用介质的宏观平均量，物质内部的储能与耗散必须重新分配，不能不加条件地把 ε_0 换成任意频率相关介电常数。[30-2]

对于一个固定体积，积分形式更加直观：场能增加率，加上穿出闭合边界的总能流，等于负的内部物质受功率。边界法线向外，因此流入给负的表面积分。正负号不是排版细节：把流入误读为流出，就会把正在充电的电容器误写成能量源。

这个推导使用场方程，而不只是能量守恒的愿望。单凭一句能量不能凭空产生，并不能告诉我们储能为何是场强平方，更不能给出能流方向为何涉及叉乘。与此同时，推导本身也不证明所有真实材料都能由这个最简模型覆盖。它提供的是一套可以被局域实验、边界功率和热量测量共同约束的关系。只要对同一个区域使用同一套划分，源端减少、区域储存、接收端增加就必须接得起来。

磁力在这里有一个容易混淆的位置。点电荷的磁力为 $q\mathbf{v}\times\mathbf{B}$，与速度垂直，所以这一项不直接改变粒子的动能；功率来自 $q\mathbf{v}\cdot\mathbf{E}$ 电。这不等于磁场对机械过程毫无贡献。磁场可以改变运动方向、传递动量，并参与导体、约束和电场共同组成的能量转换。把单粒子磁力不做功，推成整台含磁场机器不能输出机械功，是把不同层次的系统混成了一个。

30.3 电阻发热，能量从哪里进入

考虑一个长直圆柱电阻，半径为 a，有效长度为 ℓ，端电压为 V，电流为 I。为突出局部关系，忽略端部的场畸变，取内部轴向电场近似均匀，$E_{电}=V/\ell$；圆柱外表面的磁场绕轴环行，大小 $B=\mu_0 I/(2\pi a)$。两者叉乘指向圆柱内部。因此能流并非沿着电子漂移方向画一根箭头就能说明，而是在侧面有明确的径向流入。

侧面总输入功率为

$$P_{入}=2\pi a\ell\times E_{电}B/\mu_0=VI。$$

这个结果正好等于电路计算的焦耳功率。这里没有发现另一份热，也没有推翻电流与电阻的关系；同一个功率获得了空间分辨的说明。若取 $V=12$ 伏、$I=0.2$ 安、$\ell=0.1$ 米、$a=0.001$ 米，则表面电场约为 120 伏每米，磁场约为 4×10^{-5} 特斯拉，向内能流密度约为 3820 瓦每平方米。乘以约 0.000628 平方米的侧面积，得到 2.4 瓦。

这是本章设置的理想算例，不能推广成所有真实电阻都严格只从侧面进能。接线、端部、附近导体和非均匀材料会改变局部场形。建立轴向电场的表面电荷分布也不可省

去；它由整个电路及其边界共同决定。若只保留一条孤立的载流线，再任意补画电场，就可能给出错误的能流路径。

这一算例的思想力量在于，同一个 VI 不再只能解释为“电流把能量带进电阻”。电流规定磁场的一部分，电压和表面电荷规定电场的一部分，二者的空间关系共同规定能量怎样穿过边界。导体参与组织传输，但传输所占据的空间不能预先缩减为导体内部。近似电路模型把这些细节消去，仍能精确计算许多端口行为；场论则说明了这种消去在哪些条件下成立。

30.4 同轴线把几何变成可检验的功率

再看一条理想同轴传输线。内导体半径为 a，外导体内半径为 b，中间为真空，忽略电阻和端部，并考虑沿线传播的横电磁模式。两导体之间的径向电场与环向磁场分别为

$E_r = V / [r \ln(b/a)]$，　$B\phi = \mu_0 I / (2\pi r)$。

两场的叉乘沿线轴向，因此 J 能的轴向分量为 $VI / [2\pi r^2 \ln(b/a)]$。把它对环形横截面积分，面积元为 $2\pi r\, dr$，便得到

$P = \int J_{能} \cdot dA = VI$。

这里的积分展示了一个容易丢掉的因素：靠近内导体处的能流密度较大，但每一小层的面积也不同。不能把某个半径处的数值简单乘以整个截面积。几何既改变局部密度，也规定应怎样汇总成端口功率。

从同一电场分布可求单位长度电容 $C' = 2\pi\varepsilon_0 / \ln(b/a)$；从磁场能可求单位长度电感 $L' = \mu_0 \ln(b/a) / (2\pi)$。二者相乘为 $\mu_0\varepsilon_0$。对于单向传播的波，电压电流之比为特性阻抗 $Z_0 = \sqrt{L'/C'}$，即约为 $60 \ln(b/a)$ 欧姆。取 $b/a = e$，便得到约 60 欧姆；12 伏的前行波对应 0.2 安，传送 2.4 瓦。

这不是说给任何负载加 12 伏就一定出现 0.2 安。它是单向行波的电压电流关系。若末端不是匹配负载，反射波会叠加，局部总电压与总电流的比值不再处处等于特性阻抗。理想开路端总电流为零，但入射波和反射波各自并不消失；在端点处把总电流为零误读成整条线从未传过能量，会抹去建立状态的过程。

设这条线长 30 米，真空模式的单程传播时间约为 100 纳秒。在线路开始受激的最初几十纳秒，远端尚不能以当前负载响应反馈源端。等到反射返回，源端条件才受到那一部分信息影响。这是静态端口图所省略的时间层次。集中电路近似之所以常常有用，是操作时间远大于这些传播时间，且分布效应可控；它不能因此被提升为瞬时作用的普遍证明。

更重要的是，功率测量和传播时延现在能够相互检验。若由端口算出的输入输出差，不能与线内储能变化相接，就需追查导体损耗、辐射、探针负载或波形标定。局域理论增加的不仅是一个图景，也增加了多种测量之间不能随意独立调整的约束。

还可以让同轴线的账本完整地运行一次。仍取匹配线中的功率为 2.4 瓦，输入一个持续 200 纳秒的近似矩形脉冲，忽略上升沿所占时间。总输入能量为 0.48 微焦。脉冲前端抵达负载以前，输入能量都存在线路的电磁场中；在 100 纳秒时，线内能量达到 0.24 微焦。此后到 200 纳秒，源继续输入而负载已经吸收，若两端功率相等，线内储能保持这一数值。源停止以后，后沿仍需 100 纳秒到达末端，所以负载还能继续收到能量，直到 300 纳秒才结束。

这个分段过程没有把能量当作不可见小球追踪，而是逐段使用边界收支。第一段输入有、输出无；第二段输入输出相等；第三段输入无、输出有。若只在源停止那一刻拍摄仪器读数，可能看到源功率为零而负载仍在发热。它不是永动现象，也不是电源停止的消息尚未被人知道，而是有具体量值的场储能正在减少。把输入输出功率分别对时间积分，最终得到相同的 0.48 微焦。

在第二段，电场能与磁场能各占线内储能的一半；这是理想单向横电磁波的关系，并不适用于任意反射叠加的瞬时状态。若人为把线内全部储能都归为电容储能，就会少记一半；若又在输入能量之外重复加上场能，则会多记一份。选择账本的边界与时刻，是验证守恒的组成部分。这个例子还给仪器提出具体要求：记录的时间窗若只延续到源脉冲结束，输出积分会被截短，看起来像出现损耗；记录必须覆盖传播尾部，才能区分暂存与耗散。

30.5 缺失的反作用可能在场中

能量之外还必须处理动量。两个带电物体相隔较远，其中一个突然改变运动，另一个尚未收到变化；若要求同一时刻两个物体之间的机械力必定大小相等、方向相反，就会遇到困难。场论的回答不是取消动量守恒，而是把系统边界从两个物体扩展到物体与场。

在真空电磁描述中，场动量密度为 $g=\varepsilon_0 E$ 电 $\times B=J$ 能 $/c^2$。能量流和动量密度因此不是彼此无关的附加定义。为了知道动量如何穿过边界，还需要应力张量。其分量写作

$$T_{ij}=\varepsilon_0\left(E_iE_j-\delta_{ij}E\text{ 电}^2/2\right)+\left(B_iB_j-\delta_{ij}B^2/2\right)/\mu_0。$$

这里 E_i 是 E 电的笛卡尔分量，δ_{ij} 在两个指标相同时为一，否则为零。张量的必要性并不来自数学炫技：动量本身有方向，它又能通过不同朝向的面流动，单一向量不足以同时记下这两种方向。按照这一应力符号约定，力密度满足

$f_i = \Sigma_j \partial T_{ij} / \partial x_j - \partial g_i / \partial t$。

对体积积分，物质所受总力等于边界应力的合力，减去内部场动量的增加率。若把动量通量张量定义为 Π，则这里 Π＝－T，不能在公式中任意交换符号。这些关系可以直接从洛伦兹力密度 ρE 电＋J×B 与场方程推得；它们也要求物质和场在同一边界定义下共同记账。[30-3]

因而，机械物体的动量没有闭合，并不立刻表示守恒失效；但也不能用“动量在场里”当作不可核查的借口。场强、边界应力和变化率必须给出缺项的具体量值。若所需缺项与独立场测量不符，就要修改模型或承认遗漏了支撑物、接地回路、辐射等系统成分。扩展系统是有代价的解释，不是无限增加不可见因素的许可证。

应力的符号还可以用平面波自行检查。若波沿 z 方向走，电场和磁场都横向，便有 Tzz＝－u。体积左端面的外法线指向负 z，于是应力对该面给出的作用沿正 z；这与入射波把物质向前推的方向一致。只记住压力等于能量密度而不检查法线，容易在吸收端算出相反的力。这个简单检查也说明，局域量的物理意义必须连同积分方向使用，不能从孤立的一个负号判断能量或动量“不真实”。

30.6 一束光怎样推动物体

对于真空中近似平面的单向波，E 电＝cB，两种场能密度相等，J 能＝cu 沿传播方向。把 g＝J 能／c^2 在一个有限脉冲中积分，得到脉冲动量大小 p＝U／c。这里要求波束近似沿同一方向，且选取的区域包含所讨论脉冲；强聚焦或多方向叠加时，应把向量动量密度直接积分，不能把所有能量机械地除以 c 就当作某一方向的净动量。

设一束能量 10 毫焦的脉冲被静止靶吸收，忽略反射和其他出射辐射，在靶反冲速度很小的近似下，它交付的动量约为 3.33×10^{-11} 千克米每秒。理想镜面把脉冲沿原路反射，入射与出射动量方向相反，交付量约加倍。对 100 瓦连续入射功率，相应平均力约为 3.33×10^{-7} 牛顿和 6.67×10^{-7} 牛顿。

这些数值是设计量级，不是某次历史实验的数据。它们说明为何验证不能只看“照光后物体动了”。吸收引起的温升、残余气体的运动、悬挂系统漂移，都可能造成力学响应。一个强的竞争解释是热效应驱动，而非光场直接的动量交付。区分二者需要改变

反射率、压力、调制频率与照射几何，并分别测量吸收功率和机械响应。改变反射率时，吸热与动量转移未必同比变化，这正是可利用的差异。

理想反射的二倍关系也有条件。镜子若显著运动，反射光会改变频率，能量与动量的交换需一起计算；若只是漫反射，出射方向分布就决定反冲大小；若有透射，则透射部分也必须入账。用最简关系提出清楚的失败条件，比把一切偏离都归咎于实验不够理想更有科学价值。

30.7 局域账本究竟证明了多少

局域能流可以存在于时间不变的场中。电阻稳定发热时，场能密度不随时间变，能量却持续流入物质；稳态意味着储存不再积累，并不意味着传输停止。其他静态配置也可能出现非零的局域叉乘，而闭合边界上的净通量为零。此时必须检查整个空间的流线、源和机械支撑，不能从某处非零箭头推断出无源的持续输出。

还有一个更深的限定。连续性方程规定的是能流的散度。若给能流加上某个足够光滑向量场的旋度，其散度为零，闭合边界的净通量也不改变。因此，只凭连续性方程，不能声称已经唯一确定所有局部流线。通常采用的坡印廷表达还连同场动力学、局域做功关系和空间变换性质使用；这些额外结构大幅限制允许的解释，却不能被一句“守恒必然如此”省略。

这些限制并未取消局域描述的价值。它要求同一组场既解释热量，又解释力和延迟；每增加一种独立测量，任意补账的余地就更小。解释的可信程度来自这种交叉约束，而不是来自能流图是否符合直觉。

这一区别给直接作用论留下的最强位置，也不是否认已测得的传播延迟，而是尝试用不同变量重新表达相同可观测关系。若某种包含延迟和完整边界条件的表示，在所有可观测量上与场论等价，仅凭某一幅能流图未必能将它排除。另一方面，只计算物体之间瞬时作用、不提供线中储能和反射延迟的简化模型，会在前述快速开关实验中遭遇明确限制。需要比较的是模型实际承诺的结果，而不是“作用”与“场”两个词的气势。

介质还提示另一条边界。材料的极化、磁化、运动和色散会参与能量与动量存储。把真空场的账本和物质的账本怎样划分，必须与所选宏观变量一致。只检验总动量并不足以唯一判定某种分账，但任一分账若遗漏可测的机械部分，便不能靠命名自救。本章不把真空公式冒充所有介质中的完整答案。

30.8 从一个缺项重新规定一个对象

以 SDE 语言回看脉冲传输，S 是经过时钟和功率标定的输入输出差与机械响应；D 是切换、延迟、负载变更和边界移动的操作序列；E 则是储能、传播、端口和应力之间彼此规定的关系。E 不是电源、导线、仪器等物件名单。没有场能储存与边界功率之间的约束，同一条暂时的功率差曲线既可能被解释为损耗，也可能被解释为尚未到达的能量，证据的意义会改变。

一条具体的 S—D—E 再创路径从源端与接收端暂时不平衡的显露开始；研究者改变线长、脉冲宽度和负载，构成有辨别力的差异；再依据时延和反射重建含储能与局域通量的关系结构，最后用另一组线长和边界回检。S＝F（D，E）在这里要求同一模型给出各组波形；D＝G（S，E）意味着用已见缺项和现有传播关系选择下一步操作；E＝H（S，D）则意味着结构修订必须受重复差异约束，不能只改一个解释词。

三种张力也有各自实处。显露 S 中的功率缺项与差异 D 中的切换时序不能在瞬时电路内相容，推动 E 增加分布储能并用长度比例回检。显露 S 中的受力与既有 E 的纯热解释相抵触，推动 D 改为吸收率和反射率分开的比较，再检验力与功率的关系。操作 D 的时间分辨率与 E 所要求的传播时间相冲突，则推动 S 由单一平均功率改成同步波形和时间积分，随后检查新读数能否闭合收支。第三种变化不是要求“再测一次”，而是改变何种显露才足以承载问题。

本书所谓物理世界的创造，在此首先指可操作对象的重组：场能不再是只为总账平衡而写下的符号，它取得空间区域、边界通量、延迟和机械回检。这样的创造受到世界的抵抗，并不表示研究者制造了本来不存在的全部物理实在。局域化也把新的困难推到前台：不同运动状态下，电场和磁场怎样相互转化？能量和动量的分账如何保持一致？既然传播具有特征速度，长度和时间的测量又该如何共同规定？下一章的相对论危机正从这些仍未闭合的接口进入。

本章参考资料

[30-1] J. H. Poynting, “On the Transfer of Energy in the Electromagnetic Field,” Philosophical Transactions of the Royal Society of London, 175 (1884), 343–361。阅读原论文公开转录的开篇问题、能流推导及应用结构；历史概述仅据开篇。https://en.wikisource.org/wiki/On_the_Transfer_of_Energy_in_the_Electromagnetic_Field

[30-2] David Tong, Electromagnetism, Chapter 4, § 4.4 “Transport of Energy: The Poynting Vector”。阅读局域能量收支与真空波能流；本章另行推导并设置电阻与同轴线数值。https://www.damtp.cam.ac.uk/user/tong/em/em3.pdf

[30-3] Richard P. Feynman, Robert B. Leighton and Matthew Sands, The Feynman Lectures on Physics, Vol. II, Chapter 27, “Field Energy and Field Momentum”。阅读场能、动量与局域解释的限定；本章应力符号与算例统一采用文内约定。https://www.feynmanlectures.caltech.edu/II_27.html

第七编 时空之物
洛伦兹、爱因斯坦、闵可夫斯基与诺特

第三十一章 迈克耳孙、洛伦兹、庞加莱
旧变换遇到新困难

31.1 一套成功的物理学，为什么会遇到另一套成功的物理学

最难处理的科学危机，往往不是正确战胜错误，而是两套各有成效的系统进入同一场合，却不能同时维持原来的规定。在运动物体的电动力学问题上，困难正来自这里。机械运动的测量有成熟的坐标规则；电磁传播有能够连接实验的方程；光学装置又把长度、方向、传播与时间放在同一台仪器内。它们原先可以分别运行，合起来却要求回答：换一个运动状态，哪些量可以不变，哪些量必须连同测量方式一起改变？

设一个参考系相对另一个参考系沿同一轴匀速运动。通常写出的伽利略变换是 $x'=x-vt$、$t'=t$。这个式子不仅调整一个位置数值，也保留了一个强条件：不同地点、不同运动状态所使用的时间，可以不经结构性改变而相互通用。因此，如果在一个坐标系中某个信号的位置满足 $x=ct$，那么直接代入以后便有 $x'=(c-v)t'$。在这套变换里，速度相减不是额外意见，而是坐标与时间条件的结果。

困难在于：如果又要求各惯性系采用同一真空电磁传播规律，且其中的特征速度仍为 c，那么刚才的变换就不能原样保留。这里的不兼容是有条件的。单独写下麦克斯韦方程，并不会自动告诉十九世纪物理学家所有惯性系都必须拥有同样形式；人们也可以设想只有以太静止系具有某种优先地位。真正需要检验的是整套条件能否共同成立，不能把后来确定的条件偷偷提前塞回历史。

本书据此把这场危机解释为两个物理发生系统之间的相遇。一个系统把共同时间、刚性长度和速度叠加连接起来，另一个把电场、磁场和有限传播连接起来。它们的相遇不是两个名词之间的争执，而是要求同一束光、同一把尺和同一组读数同时服从两个构成方案。新物理世界的创造，将发生在这些构成方案无法继续各行其是的地方。

31.2 干涉条纹怎样把一个宇宙问题变成可比较的差异

迈克耳孙与莫雷在一八八七年的实验，把沿不同方向往返的光带回同一处比较。研究者转动装置，寻找与设想中的以太相对运动有关的条纹变化。报告没有得到预期规模

的位移。这里的“零结果”是相对理论预期与实验分辨能力而言，不表示每一次读数都严格等于零。[31-1]

仪器的重要性在于，它把一个极为遥远的问题压缩成现场可以辨认的变化。人不必先看见以太，再测量自己穿过它的速度；只要一个以太模型要求两条光路出现可计算的时间差，时间差又能够产生相位差，相位差便可进入条纹比较。原来不能直接看见的假设，由此获得能够失败的实验出口。

这不是单向的“不可见对象导致可见现象”。从研究过程看，光源稳定性、反射镜位置、光程构造、旋转方式与条纹辨认共同规定了什么算作有效现象。随温度漂移而缓慢移动的条纹，与随装置方向系统变化的条纹，不具有同一种证据资格。必须构造有区分力的差异序列，才能把许多变化中的某一种稳定提取出来。

以本书的 SDE 语言说，S 不是未经加工的一幅亮暗图，而是能够在操作重复中被认作同类证据的显露；D 不只是装置转了一圈，而是包含转动次序、比较方式和读数选择的差异序列；E 也不是实验室周围的普通背景，它是光程、相位、方向、预期量级与判定规则彼此取得意义的特征纠缠。少掉其中一个关键条件，“没有出现预期条纹位移”就可能变成另一种完全不同的陈述。

然而，这种分析不能把证据变成任意解释。正因为条纹的证据意义有构成条件，那些条件才必须公开、稳定、可复做。研究者不能在结果不合意时随意更换有效读数的标准。一个发生系统能创造新的物理对象，并不意味着它可以按意愿创造任何结果。世界的抵抗，正通过重复操作之后仍然不能消除的差异进入理论。

31.3 计算先暴露了什么，实验又没有单独决定什么

为了看清问题，考虑一个理想化模型：装置在静止以太中以速度 v 运动；两条垂直臂在模型所用坐标下都长 L；光相对以太以速度 c 传播；暂且不允许臂长随运动改变。沿运动方向往返所需时间为 $T_{\parallel}=L/(c-v)+L/(c+v)=2Lc/(c^2-v^2)$。这里追上移动反射镜与返回分光点需要分别计算，不能把两个单程时间当成相等。

垂直臂的光也不能沿以太坐标中的固定竖线往返，因为镜子随装置平移。满足相遇条件的光程构成斜线，故 $T_{\perp}=2L/\sqrt{(c^2-v^2)}$。令 $\beta=v/c$，在 β 很小时展开可得 $T_{\parallel}$ 约为 $(2L/c)(1+\beta^2)$，$T_{\perp}$ 约为 $(2L/c)(1+\beta^2/2)$，两者之差约为 Lv^2/c^3。将装置转过直角，两臂的角色交换，理想差值的变化量因此多出一个二倍因子。以上是为说明假设链而作的标准理想推导，不是对真实装置全部光路和误差的替代。

这段计算值得认真读，因为它让一项常被遮蔽的工作显现出来：理论预期不是从“存在以太”四个字直接跳出来的。它同时使用传播速度规则、镜面运动规则、长度条件与几何相遇条件。若实验不支持预期，则这些条件的合取出了问题；但究竟应修改哪一项，还需其他证据。

例如，让沿运动方向的臂按 $L_{/\!/}=L_0\sqrt{(1-\beta^2)}$ 变化，而垂直臂仍为 L_0，就会使上述两个往返时间相等。这一代入只能证明某种收缩规则具有消除理想时间差的能力，不能独立证明材料实际上依照此规则变化，更不能说明这条规则为何适用于不同物质。数学上的可行修补，需要接受物理上的跨对象约束。

同样，只凭一项零结果不能直接推出光速在所有惯性系都相同，也不能一次排除一切可能的以太模型。一个实验有力量，是因为它排除了一类具体预测，而不是因为它可以代替所有后续推理。物理思想史若把一次实验写成某种新理论的完整证明，便省略了最值得研究的创造过程：人在多个仍然可能的改造方向之间，如何形成更强的连接与更严的回检。

31.4 洛伦兹：把临时补救推向共同变换

洛伦兹一九〇四年论文开篇讨论二阶效应与若干零结果，并承认每遇到一种新事实就添加特殊假设的做法不够令人满意。他尝试从共同假设出发，处理一类运动系统；文中引入包含空间项的局域时间，并把场量与坐标一起变换。[31-2]

这项努力的重要性，不宜简化成“给失败的以太打补丁”。如果一种修正只对一台仪器有效，它很难成为理论；如果同一变换能够约束许多装置、许多材料和许多过程，那么原来分散的补救便开始转化为共同结构。评价一个研究纲领时，关键问题不是它有没有保留旧名词，而是它的修改究竟减少了独立假设，还是增加了任意豁免。

局域时间尤其值得注意。它意味着，在建立可比较的运动过程时，时间变量不能保证永远只依赖旧时间。以现代记号表示洛伦兹变换的一维标准形式，是 $x'=\gamma(x-vt)$、$t'=\gamma(t-vx/c^2)$，其中 $\gamma=1/\sqrt{(1-v^2/c^2)}$。这里展示的是后来通行的标准形式，并非声称洛伦兹在所有历史阶段都已赋予它相同解释，也不把一九〇四年的全部物理论证等同于现代狭义相对论。

空间项进入时间，最直接的含义是：两件事若在某系满足相同 t，却发生在不同 x，那么变换后的时间差一般不为零。代入可得 $\Delta t'=-\gamma v\Delta x/c^2$。这是一条代数结论。它在物理上究竟表示“真实时间之外的辅助变量”，还是另一套实际测量程序所建立的时间，则涉及解释地位的改变。

因此，思想创造并不总发生在公式从无到有的那一刻。有时公式已经能够工作，但它仍被放在旧世界的从属位置。将辅助变量提升为可操作的基本规定，并据此重新安排什么具有优先地位，可能产生另一种深刻转变。反过来，也不能因为后来解释更统一，就抹去此前形成公式、辨识约束、建立计算方法的创造。

这里可以提出一个发生学判断标准：新变量有没有改变旧变量之间可以进行的操作？如果局域时间只是抄写方便，删掉它不会改变问题的结构；如果它使过去互不协调的传播过程能够在同一变换下比较，那么它已经承担新的组织作用。问题随后才推进到：这种组织作用需要一个不可观测的绝对时间作为底座吗？

31.5 庞加莱：一次变换成功，还不够

庞加莱一九〇五年六月的短文明确讨论洛伦兹变换的群性质，并把相对运动问题与共同变换要求联系起来；他也讨论电子稳定所需的非电磁作用。稍后的长文进一步展开这一研究。[31-3][31-4]本章关注其中一个方法论跃迁：对某一次转换有效的规则，能否在连续转换中保持一致？

设甲系与乙系之间有一套变换，乙系与丙系之间也有一套变换。如果由甲到丙直接计算，与经过乙系再到丙系的结果无法相容，那么这些变换尚未组成一个可共同使用的世界。单对关系的成功，不能代替关系集合的封闭性。所谓群性质，在这里不只是数学术语，它要求变换可以复合、有对应的逆操作，并具有共同的结构规则。

以一维标准洛伦兹变换作本书的说明性计算，两个同向变换的速度参数 u 和 v 组合以后，所得参数为 $w=(u+v)/(1+uv/c^2)$。这个表达式说明，速度组合方式与空间、时间的组合方式不能彼此独立选择。若保留旧的时间通用性，却单独采用新的速度组合式，便会在其他环节出现矛盾。数学一致性要求局部修改能够传到整体。

这个要求同样适用于实验。一个实验室可能在特定条件下获得稳定结果，但若不同实验室的测量不能通过明确规则相互转换，稳定性仍然局限于局部。通过转换结构建立的可比性，使“同一个物理过程”获得跨状态延续的规定。这种同一性不是忽略变化后的抽象口号，而是变化被允许、结果仍可协调的构造成果。

本书由此把庞加莱的作用理解为：他推动人们追问一个变换世界是否能够自洽地聚合成体。这里的“创造世界”具有具体含义。对象能够带着哪些性质跨参考系被追踪，哪些性质只在特定坐标表达中成立，哪些组合能够保持，都开始受新的结构制约。世界的共同性，不再靠事先宣布所有时间相同来担保，而必须通过变换的组织能力获得。

但群结构不能单独替代全部物理。许多数学变换可以组成群，只有得到测量与动力学支持的变换才构成合适的物理模型。电子模型中的稳定性问题也提醒我们：写出漂亮的变换，还需要说明被变换的对象如何维持。因此数学一致、物理实现与经验支持必须相互牵制，任何一项独自成功，都不足以终结研究。

31.6 三方程不是贴标签，六路径必须承担改造后果

现在可以把这一章形成的张力放回 SDE 三方程。S＝F（D，E）要求解释可观察的条纹、计时与运动关系如何成立；D＝G（S，E）要求根据已经显露的结果和既有纠缠，重新组织比较、转动、变换及测量序列；E＝H（S，D）则追问：在结果与操作不能继续按旧方式协调时，哪些特征之间的构成关系必须重建。这些式子在本章是研究框架，不能冒充已经量化拟合的物理定律。

以干涉实验为例，旧模型下的预期位移与实测显露形成张力。如果只说“改一改 E”，便没有完成分析。必须指出究竟改了什么：是长度与运动的连接，是时间与位置的连接，还是传播规律与参考系的连接。随后还要让改变后的纠缠重新生成可计算的 D，并回到可独立比较的 S。否则 E 就会退化成能够容纳任何结果的口袋。

六路径在这里承担完整的次序检验。沿 E—D—S 的沉淀路径，要检查既有纠缠如何经测量运行成为稳定显露；沿 E—S—D 的锚定路径，要由结构与已知显露重新规定测量序列。D—E—S 的倒逼路径要求操作困难推动构成关系改变，再生成可检验结果；D—S—E 的原型路径则从具体运行及其显露，重建哪些特征可以共同成立。S—E—D 的进入路径，由零结果进入旧结构并调整比较程序；S—D—E 的再创路径，则从已显露问题与差异运行重建构成系统。这些路径不是六种随意替换的叙述顺序，每一路都必须说明改了哪项条件、造成什么后果。

三原理在这里要求把两维之间的张力落实为第三维的改变，并在改变后回检。S 与 D 的张力推动 E 改变，不能只为失败结果另添名词；S 与 E 的张力推动 D 改变，不能不经重新测量便宣布结构获救；D 与 E 的张力推动 S 的重新生成，也不能预先指定结果一定符合理论。以收缩假设为例，改动长度与运动的构成关系之后，需要重新算出光程差，再到实验显露接受检查。以局域时间为例，改动时空变量关系之后，需要检查整个变换程序是否产生一致结果。三方程规定相互依存，六路径展开改造次序，三原理要求张力转化与返回检验，三者联合才能避免方法空转。

这里尤其要防止三号位与三维混淆。观察者采用的读数规范可以涉及主体性属性，仪器与光路的作用可以涉及互动性属性，稳定测得的量可以涉及客体性属性；但每一号

位属性的成立都涉及完整的 S、D、E。不能将人放进 S、实验放进 D、以太放进 E，然后说完成了解构。那样只是把旧的人、过程、外界分类换了一组字母。

31.7 不能把三个人写成通向一个英雄的台阶

迈克耳孙与莫雷提供的并不是一段等待天才解释的空白。他们构造了新的证据能力：一种原来难以进入经验比较的相对运动假设，现在必须面对细微而可重复的光学差异。洛伦兹构造的也不仅是一道防线；他让多个运动效应开始接受共同变换约束。庞加莱进一步把单项兼容问题推向变换集合的一致性。这些创造具有各自的完成程度与问题边界。

因此，本章不把他们排成“差一点就成为爱因斯坦”的队伍。这样的叙事把后来成果变成唯一尺度，反而看不见每一种研究对象怎样形成。严格的思想史可以比较不同方案解释的代价与生产能力，却不应把历史人物的全部工作浓缩成他们距离后来某篇论文还有几步。

爱因斯坦一九〇五年论文开篇讨论磁体与导体相对运动的解释不对称，也提到探测地球相对光媒质运动的失败尝试，并把运动学基础带回问题中心。[31-5]这一原文足以反对“只有迈克耳孙—莫雷实验触发相对论”的单因叙事，却不足以还原全部个人思想经历。没有进一步的一手材料，本书不替作者补写某次阅读或某个瞬间的心理路线。

本章所建立的接口因此十分明确：下一步需要研究的，不只是怎样得到更合适的变换，而是测量时间、比较长度、辨认同时性这些基本行动，如何在同一套新原则下取得意义。物理世界的重建已经开始，但尚不能用一个成功公式宣布完成。它必须让计算、仪器与对象构成在更多场合共同运行，并允许新结果再次改写它。

本章参考资料

[31-1] Michelson, A. A.; Morley, E. W. On the Relative Motion of the Earth and the Luminiferous Ether. American Journal of Science, 1887, s3-34:333–345. 原刊扫描：https://history.aip.org/exhibits/gap/PDF/michelson.pdf

[31-2] Lorentz, H. A. Electromagnetic Phenomena in a System Moving with any Velocity Smaller than that of Light. 1904. 原文转录，重点 §1—4：https://en.wikisource.org/wiki/Electromagnetic_phenomena

[31-3] Poincaré, H. Sur la dynamique de l’électron. Comptes rendus, 1905, 140:1504–1508. 六月短文英译：https://en.wikisource.org/wiki/Translation:On_the_Dynamics_of_the_Electron_(June)

[31-4] Poincaré, H. On the Dynamics of the Electron: Introduction, §§1, 9. 庞加莱文献项目所收英译，长文写于 1905、刊于 1906：https://henripoincarepapers.univ-nantes.fr/chp/text/hp2007gg.html

[31-5] Einstein, A. Zur Elektrodynamik bewegter Körper. Annalen der Physik, 1905, 17:891–921. 引言与前两节英译：https://sites.pitt.edu/~jdnorton/teaching/HPS_0410/chapters/origins_pathway/On-the_electrodynamics/index.html

第三十二章 爱因斯坦 1905
同时性使“长度”成为另一种物

32.1 先问两端在什么时候，再问一根尺有多长

一根尺放在桌上，人们把标准尺与它并排，两端一比，长度便出来了。这种简单经验很容易使人相信，长度是一项已经完整装在物体里面的属性，测量只负责把它取出来。运动物体的问题揭示了被日常经验遮住的一层：若物体的前端和后端都在移动，测量者必须说明，他把前端在哪个时刻的位置，与后端在哪个时刻的位置相减。

假设一辆匀速行驶的车，研究者在某一时刻记录后端位置，一秒后才记录前端位置；这里说的是两次读数的先后，并不是两端通过同一标记的先后。把这两个位置直接相减，不一定得到车长，因为其中混入了车在一秒内移动的距离。这不是精密物理才有的问题，而是每一次运动长度测量的基本条件。只是日常速度较小、计时精度较低时，人们往往已经默认了一套共同的同时规则，不再把它当成需要解释的构成部分。

爱因斯坦一九〇五年论文第一、二节把时钟同步与运动尺测量放在相邻位置，区分随尺共同运动的测量和另一参考系对两端同时取位置的测量。[32-1]本章将以现代记号展开这种区分的后果，而不把后来的全部几何解释说成原论文已经一一完成。问题的重心从“尺究竟被什么挤短”移向“在不同测量系统中，怎样的两端事件共同构成一次长度”。

章题所谓长度成为“另一种物”，并不是把抽象数值说成新材料，也不是宣布同一根尺被人的观看切成几件物品。这里的“物”指能够在物理理论和测量活动中稳定成立的规定。旧图像倾向于让长度脱离时间组织而成立；新的运动学则要求明确端点事件、同时性条件和相对运动状态。看起来仍是一个带单位的数，组成这个数的条件却已改变。

从本书主轴看，这种改变具有创造物理世界的意义。主体不是站在已经定型的长度之外任意挑选意见；标尺、时钟、信号、物体运动与比较规则共同进入长度的发生。它们形成的约束可以被别的研究者复做，可以产生一致的转换关系，也会抵抗不合适的解释。正因为测量结果不能按人的愿望任意更改，新的规定系统才具有物理力量。

32.2 让远处时钟进入同一套时间

一只钟可以给发生在旁边的事件标时。例如指针到达某处与一次碰撞在局部相遇，这类对应相对直接。困难来自远处：如果另一只钟与第一只钟相隔很远，仅仅说它们显

示一样的数字，还不能保证它们属于同一套时间。两只钟可以保持相同走时速率，却从不同的起点开始；也可能在某次对表后因运行条件不同逐渐分开。

考虑一种理想同步方案。甲处在时刻 t_1 发出信号，乙处收到后立即反射，甲处在 t_2 收到返回信号。把乙处反射事件的时间规定为（t_1+t_2）／2，是一种对往返传播作对称分配的规则。在理想惯性系中，配合真空光传播与一致时钟的假定，它建立了一套可用于远处事件比较的时间坐标。这里采用的是明确的同步约定，并不假装一个局部读数能够独自测出已经同步之后才有定义的所有单程传播时间。

这种规则不是人们喜欢哪个数字就填哪个数字。它必须与多地点时钟网络一致：甲与乙、乙与丙建立的同步关系，应能够在适用条件下与甲丙关系相容；重复操作不应产生无法解释的累积冲突；换一种等效装置也不能改变物理预测。约定为坐标赋予意义，经验则约束以这些坐标表达的规律能否共同成立。两者合作，不等于两者可以相互取代。

还必须把“同时收到光”和“远处同时发生”分开。一位观察者同时收到两束光，可能只是因为较远事件较早发光，传播时间刚好补偿。反过来，两个在所选参考系中同时发生的事件，如果到观察者距离不同，也不会同时进入他的眼睛。若不处理传播延迟，讨论的只是到达事件；处理后讨论的才可能是被定位的远处事件。相对论的同时性不能归结为眼睛看得快慢不同。

现在可以看见时间变量的一项深层变化：它不再只是一条在所有地方自动流过的公共刻度。对一个参考系而言，远处时间由一套能够协调使用的时钟与信号规则建立。这套结构可以具有很高的稳定性，却仍然需要说明成立条件。把条件纳入时间，不是降低时间的客观性，而是让客观比较拥有可追查的程序。

32.3 必须区分三种“改变”

讨论相对论时，许多争论来自把不同操作混成同一个动作。第一种是同一个事件换坐标。譬如某次闪光在一套坐标中记为（t，x），在另一套坐标中记为（t′，x′）。事件没有因为改用另一套数值而发生第二遍。两种记录之间的转换，描述的是同一事件如何进入不同但相互协调的定位系统。

第二种是为了测量某个物理量，选择了不同的事件对。测车长需要车头和车尾各取一个事件；“在甲系同时”与“在乙系同时”一般从同一辆车的历史中切出不同的一对事件。这时长度数值之差不能单纯解释为同一事件对的空间坐标差换了一种单位，因为

选择规则本身也不同。要得到甲系长度，应按甲系同时性挑选端点；得到乙系长度，也必须按乙系同时性重新挑选。

第三种是真实干预：给尺加速、拉伸、加热，或者改变它内部的应力状态。这些操作改变物体的物理历史，不是被动更换坐标。加速后尺能否保持原来的静止长度，取决于具体加速方案、材料响应和是否回到相同的内部状态。狭义相对论的长度变换不能单独回答一块实际材料在任何驱动下最终变成什么形状。

这三者可以共同出现在一个实验里，却不能相互冒名。例如实验者先把一根尺加速到稳定运动，再在实验室测长，既涉及真实准备，也涉及端点事件的选择。若发现变化，应先分清哪些来自物体状态改变，哪些来自运动学测量条件。否则把所有差异都叫“观察角度不同”，会把材料的真实变形抹去；把所有差异都叫“物体真的被压缩”，又会误把坐标组织当成一种额外的挤压力。

本书使用 SDE 语言，也必须保持这道区分。不同观察位置不等于发生了不同物理干预；同一结果的两种表达不等于创造了两次实体；实际差异序列的改变，也不能被降为纯粹表达替换。发生学分析应比普通叙述更清楚地说明发生了哪种改变，而不应借“都在互动中发生”消去各种操作之间的差别。

32.4 同一把尺的两个长度如何同时成立

设一根尺静止在 K′系中，两端坐标为 $x'=0$ 与 $x'=L_0$。L_0 是这根尺在指定内部状态下的静止长度。K′相对 K 沿正方向以速度 v 匀速运动，且 $|v|<c$。取 $\gamma=1/\sqrt{(1-v^2/c^2)}$，标准变换为 $x'=\gamma(x-vt)$，$t'=\gamma(t-vx/c^2)$。本节只讨论沿运动方向的尺，且各系采用上述标准同步方式。

K 系要测运动尺长，需要在相同 t 记录两端。将两端事件代入空间变换相减，得到 $\Delta x'=\gamma(\Delta x-v\Delta t)$。因为 K 系选择了 $\Delta t=0$，且两端在 K′中固定相隔 L_0，所以 $L_0=\gamma\Delta x$，即 $L=L_0/\gamma$。代数上没有加入一个把尺压短的额外作用；所得长度来自端点轨迹和 K 系同时截取条件的联合。

但必须立即检查同一对事件在 K′中的时间差。由时间变换可得 $\Delta t'=-\gamma vL/c^2=-vL_0/c^2$。它们在尺的静止系中不是同时发生的。这并不妨碍其空间坐标差恰为 L_0，因为两端在该系中始终固定；然而它提醒我们，不能把两套测量各自使用的事件对看成自动相同。相对论中的“同时”不是写在事件本身上、所有坐标系都照抄的标签。

举一个便于复算的数值。若 $v=0.6c$，则 $\gamma=1.25$；一根静止长度十米的理想尺，在 K 系同时测量所得为八米。这个结果既不说尺的每一段材料在自身局部状态中都缩掉了

两成，也不说八米只是光学幻觉。它说，在已经约定的测量程序下，实验室同时定位的两端相距八米。拍照结果还包含两端光信号到达相机的不同延迟，必须另行处理，不能把照片上的外形与同时测量的长度直接等同。

尺的同一性因此不能由“所有人必须得到同一个长度数”维持。它可以由端点历史、静止长度、运动状态以及跨参考系转换关系共同维持。如果不同研究者按照清楚的条件得到不同数字，而这些数字满足统一的可检验关系，那么这里出现的是有组织的差异。反之，若同一条件重复测量却任意变化，才是另一种需要解释的不稳定。

32.5 一只钟走过的时间与一排钟给出的时间

长度问题涉及两端，计时问题也要先问取的是哪些事件。考虑一只静止在 K′中的理想钟，先后两次滴答发生在相同 x'。两次滴答之间的本地读数差记为 $\Delta\tau$。对 K′而言，这就是该钟经历的时间间隔。对 K 而言，这只钟在移动，两次滴答发生在不同位置；用 K 系已同步的时钟网络标时，得到间隔 Δt。

从逆变换 $t=\gamma\left(t'+vx'/c^2\right)$ 取两事件之差，因为 $\Delta x'=0$，就有 $\Delta t=\gamma\Delta\tau$。因此在 K 系比较中，运动钟两次滴答之间占用更长的坐标时间，也就是这只钟相对 K 系时钟网络走得较慢。这一推导比较了同一对滴答事件，却使用不同计时结构。它既不是研究者精神感受到时间变慢，也不是假设机械钟出了故障。

仍取 $v=0.6c$。如果运动钟两次滴答间隔为四秒，K 系给出的相应间隔为五秒。另一参考系也可以说 K 系的运动钟走得慢，但这不构成逻辑矛盾，因为双方各自选择的一只钟与对方一排同步钟，形成了不同的比较配置。要把两句话合成一项矛盾，必须先证明它们对同一对事件、同一比较程序作出了互相否定的判断。

“既然双方都觉得对方慢，最后谁真的年轻”是另一个问题。若两只钟离开后再相遇，应比较各自沿实际路径累积的时间。重新相遇的方案通常包含转向或其他使两条历史不对称的安排，不能把两个始终匀速分离的参考系对称性直接套在整个过程上。加速度参与改变路径和比较条件，但不能简单说年龄差只发生在加速瞬间。下一章的世界线与固有时，能够把这个区别表达得更清楚。

在实验层面，真实时钟还需要校验。温度、振动、电磁干扰或结构变化可能影响某类装置，理想钟是理论模型，不是对所有钟具的免检证书。正因如此，若不同原理的时钟在适当控制后都支持同一运动学关系，这种一致性才具有意义。物理时间的共同性，不是由一种钟垄断出来的，而是在可以相互比较的多种运行中被稳定建立。

32.6 相对性没有取消精确性，反而增加了条件

人们有时把相对论理解成“每个人都有自己的真理”。这恰好错过了理论最严格的部分。给定同一物理过程、参考系相对速度、同步方法和事件选择，转换结果受到明确约束，不能按个人感受更改。相对性撤回的是某些量无条件共享同一数值的要求，建立的是条件明确时的共同变换关系。

一项好的测量报告因此不能只写“尺长八米”，而应至少让读者知道尺的状态、相对运动以及怎样取得两端事件。不是每次日常测量都要列出整页理论前提，而是在这些前提影响比较时，必须能够恢复它们。日常语言的简化之所以安全，是因为某些差异小到可以忽略；当精度或速度使这些差异进入结果，就需要打开原先封装的条件。

这也使属性纠缠论获得一项具体研究任务：不是给长度添加许多相关因素，而是指出哪些关系进入长度的定义与可测构成。空气颜色通常不参与理想尺长的定义；同时性却直接规定取哪两个端点事件。材料温度可能通过改变静止长度进入另一个问题；观察者的主观喜好则不能改变标准变换。在构成性、动力学影响和无关条件之间作区分，才使“纠缠”具有分析力量。

与此相关，单位改变和参考系改变也应分开。把米换成厘米，会把所有相应长度数值乘以相同换算因子，并不需要重新选择同时事件。参考系变换会混合时间与位置，测量长度时还涉及不同的同时截面。两者都能导致数字变化，但改变发生在不同层级。若只因数字不同就把它们等同，便看不见狭义相对论真正重组了什么。

本章的发生学命题因而不是“任何属性都随条件随意改变”。更精确的说法是：属性必须通过明确的构成系统取得身份，不同系统之间既可能存在稳定的转换，也可能存在真实状态变化，还可能根本不构成有效比较。研究工作的任务，就是辨明这些情况，并建立能够排除错误比较的方法。

32.7 一项测长报告应该怎样被复核

设想一项待复核报告：研究者说某根高速尺比预测值更短，而且把差额直接归结为相对论失败。本书的第一步不会先选立场，而是恢复这项显露 S 的生成条件。前后端有没有在同一套时间中定位？钟的偏差有没有修正？尺是否仍保持原来的内部状态？这里的 D 是可追踪的操作序列，E 则是端点身份、同时性、传播规则与物体状态的构成关系。它们不是实验之外临时补列的背景。

先沿 S—E—D 这条完整路径进入。异常长度只有同一个理论预言量相比较，才有资格称为异常。如果报告实际使用相机照片的投影长度，理论却给出同时截取长度，那么

首先缺少的不是新定律，而是同一被测量的构成条件。恢复 E 之后，应改造 D：记录来自两端的光的传播时间，或使用已同步的端点探测器，重新取得能够与预测相比较的数据。这一路径是否有效，要看新操作是否使原来的异常以可重复方式消失或保留。

再沿 D—S—E 的路径作反向压力检验。假设不同实验室使用不同但等效的端点探测方案，都在校准后得到同方向的偏差，就不能永久用“操作大概有错”保护理论。这些 D 共同生成稳定 S，便要求重新检查 E 中的模型规定。实际尺在加速中是否受热，静止长度是否改变，所假定的惯性运动是否真的保持，都应成为可以独立测试的问题。若这些解释也被排除，才进一步考虑基本变换的适用边界，而不是用一句物体复杂就结束研究。

两个案例把三方程落实为不同的研究负担：预测 S 时必须交代 D 与 E；修订 D 时必须依据 S 与 E；重建 E 时必须接受 S 与 D 的双重限制。它们也使三原理获得具体内容：显露与构成之间的张力可以推动程序改变，程序与显露之间的张力可以推动构成改变，改变后的程序和构成又必须生成新的显露接受回检。这里没有哪一个维度拥有永久免检权。

复核还需要安排停止条件。若偏差可由独立测得的温度变化定量解释，且冷却恢复后按同一公式消失，那么继续无限追问新时空理论没有根据；若数据持续冲突，重复校准也无法解决，就不应为了维护已有解释停止调查。好的方法既能打开被封装的前提，也能在问题已被具体解释时收束，不让怀疑本身取代证据。

这种分层判断不是退回孤立客体论。恰恰因为属性由多种构成条件共同成立，研究者才需要分辨哪一种条件真正参与当前异常，哪一种只是被随手列出的可能性。把所有条件统称互动，容易使每次失败都能被无限解释；让具体条件产生独立预测，才使属性纠缠成为有区别力的分析。主体性、互动性、客体性属性都可以进入这项工作，但每一类属性仍须追查其完整的 S、D、E，而不是分别把三号位塞进三只抽屉。

32.8 另一种长度的诞生，通向另一种整体

现在回看桌上的那根尺，它并没有因为相对论出现就失去日常用途。低速、近似共同时刻与足够稳定的材料，使传统测长成为可靠的局部封装。新理论所做的，是把这种可靠性放回成立条件中：当相对速度趋于零时，γ 趋于一，两种长度差异随之缩小；这是一条可计算的近似关系，不是对旧经验的礼貌保留。

更重要的是，新的规定系统允许旧系统难以共同处理的测量相互进入。时间不再只是给位置变化贴标签，长度也不再只是某个时刻一眼可见的轮廓。二者通过事件、时钟

网络与变换关系共同成立。既不能把时间完全溶解为人的意识，也不能把长度重新锁回一件与全部操作无关的孤立物体。

一个谨慎的实在论者仍可主张，同一根尺的物理历史不依赖某位观察者是否知道它。本书无须否认这种独立抵抗。需要推进的是：关于这段历史的“长度”规定，怎样才能获得跨过程可比较的具体意义？如果回答必须调用同步、取点与变换，那么这些关系就不能永远被当作物体之外的附带认识手段。它们参与构成了物理理论中这项属性的身份。

也不能走向另一端，把测量定义当成全部物理。尺由什么材料组成、为什么保持稳定、加速中会不会断裂，都需要动力学与材料理论。操作定义让我们知道在量什么，并不替我们解释所有对象怎样维持。新的物理世界因此具有多层结构：测量规定、变换规则、物体动力学与实验实现彼此纠缠，但各自承担不同问题。

这一章把问题推进到一个新的门槛。如果同一物体可以由不同同时截面获得不同长度，而稳定性又不能只靠数值相同来表达，那么有没有一种整体结构，能够把这些截面、事件与转换放到同一框架内？闵可夫斯基将把运动学的操作问题，进一步转化为世界线与时空间隔的几何问题。下一章研究的，正是这种整体如何被创造，以及它为什么不等于把世界画成一张四维照片。

本章参考资料

[32-1] Einstein, A. Zur Elektrodynamik bewegter Körper. Annalen der Physik, 1905, 17:891–921, 重点§§1—4。原文英译：https://fourmilab.ch/etexts/einstein/specrel/www/。本文仅短述原论文的同步与两种测长结构；其余推导与分析采用明确标示的现代记号。

[32-2] Einstein, A. 同文§§1—2，匹兹堡大学所收英译及对同步句译法的校注：https://sites.pitt.edu/~jdnorton/teaching/HPS_0410/chapters/origins_pathway/On-the_electrodynamics/index.html。

第三十三章 闵可夫斯基
四维时空怎样重组物理对象

33.1 把一生画成线，会失去什么，又获得什么

我们在纸上画一辆车的路线，通常只保留它经过哪些地方，不保留它什么时候经过。两辆车可以走过完全相同的路线，却从未相遇；也可以在平面路线图上交叉，却因到达时间不同而安全通过。空间轨迹不能独立回答相遇问题。若把时间一并记录，车辆的历史就不再只是平面上的线，而成为同时包含位置与时间的事件序列。

闵可夫斯基一九〇八年在科隆的演讲，把世界点、世界线、光锥与固有时组织在共同的时空表达中，强调物理规律与这些结构之间的关系。[33-1]这项工作不是凭空给宇宙再添一个看不见的方向，而是让此前分开处理的时间、位置和变换获得统一的数学组织。下文采用现代实坐标与符号约定作分析，不逐句重述演讲，也不把后来成熟理论的全部内容倒放到一九〇八年。

一条世界线表示一个理想点状对象的一段历史。它上面的一个点对应一次事件，不是某个空间点永远不变的身份。若对象在所选坐标系中静止，空间坐标不变，世界线沿时间方向延伸；若匀速运动，空间坐标随时间线性变化；若速度改变，线的方向也改变。这些图形使不同过程能够放在同一个构造里比较。

但画成一条线，也意味着作了抽象。真实物体有大小、内部结构和复杂运动，不能总由一个点代表。若研究整根尺，应考虑尺中不同部分的历史形成的世界管，而不是只追踪质心就以为获得了全部信息。物体的温度变化、破裂和形变可能需要额外变量。四维表示并不会自动保留全部物理，只是把事件的时空关系组织成一种更强的共同框架。

更不能把世界线当成一条真的铺在宇宙里的发光轨道。画线是理念活动，追踪事件依赖仪器与操作，实际对象沿其历史参与相互作用。三者可以紧密接合，却不是同一种存在。发生学的任务是解释它们如何取得可检验的联系，而不是因为图形直观，就把数学图上的线直接当成感官经验中现成可摸的物。

33.2 一个数值会变，为什么仍能描述同一段关系

上一章中，时间差和空间距离都可能随参考系改变。若把稳定性理解成“任何人写出的每一个数都一样”，相对论似乎破坏了稳定性。闵可夫斯基式的组织让另一个问题进入中心：哪些组合在允许的变换下保持，哪些关系虽然分量改变却仍能严格对应？物理同一性开始由结构承担，而不是由每个分量单独承担。

采用一时间三空间的现代记号，两个事件的间隔写为 $\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$。本章使用时间项取正的符号约定。也有教材把所有符号反过来，这不构成物理分歧；重要的是整套判定和公式始终一致。这个间隔不是四维欧氏距离，因为空间项带负号。若误把它写成四项平方全加，便会消除相对论特有的几何结构。

只看一维运动，可以直接复算不变性。令 $\Delta x' = \gamma(\Delta x - v\Delta t)$，$\Delta t' = \gamma(\Delta t - v\Delta x/c^2)$，代入 $c^2\Delta t'^2 - \Delta x'^2$，展开后两项含 $\Delta t\Delta x$ 的交叉项抵消，剩下 $\gamma^2(1 - v^2/c^2)(c^2\Delta t^2 - \Delta x^2)$，因此回到原值。稳定组合不是凭一句“几何告诉我们”指定出来的，它与变换规则具有可核查的代数联系。

作一个具体例子，以光秒表示长度、秒表示时间，使 c 的数值等于一。两个事件在 K 系中相隔五秒、三光秒。间隔为二十五减九，等于十六平方光秒。转到相对速度 0.6c 的 K′系，$\gamma = 1.25$，得到时间差四秒、空间差零。两套数值很不同，但间隔仍是十六平方光秒。K′中两事件发生在同一地点，这说明这一类事件对允许某个随行系把它们表达为同地而不同时。

这个例子也说明，不变量不等于一切都不变。它保持的是经明确规则组成的关系量；各分量变化正是这项关系在不同分解中的体现。若只寻找一个永远不变的空间距离，就会错过新系统的稳定性所在。反过来，若只强调分量都变了，也会把严格可比的物理世界误读为没有共同标准的碎片。

33.3 光锥划出的不是心理视野，而是可能连接的结构

间隔的正、零、负，把事件对分为三种不同类型。按本章约定，正间隔称为类时，零间隔称为类光，负间隔称为类空。在平直时空的标准相对论结构中，这种分类不因洛伦兹变换而改变。它并不由观察者喜欢哪种描述决定，因为间隔本身在这些变换下保持。

以一个事件为原点，满足 $c^2t^2 = x^2 + y^2 + z^2$ 的事件组成光锥。未来一侧对应从原点出发的光信号可以到达的边界，过去一侧对应光信号能够到达原点的边界。锥内的类时事件可以由低于光速的路径连接，锥外的类空事件不能由遵守这一速度限制的信号直接连接。这里说的是理论允许的连接条件，不是断言每一对类时事件之间都真的存在因果作用。

因果可能性与已经发生的因果关系必须区分。某次敲击能够在足够长时间后影响远处探测器，不意味着探测器的所有变化都由那次敲击造成。传播还需要相互作用渠道、

初始状态和具体动力学。光锥排除一部分不可能的解释，却不独立选择唯一真实原因。把几何允许误当成事实发生，会让物理因果分析失去最重要的证据步骤。

再给一个反例。两个事件相隔三秒和五光秒，间隔为负十六平方光秒。选择 0.6c 的变换，可以使它们在新系同时，空间差为四光秒。若再选择合适的更大相对速度，其中原本较晚的事件可以在新坐标中较早。但这并不意味着谁能把消息发回自己的过去，因为这一对事件属于类空分离，标准结构本来就不允许用普通低于或等于光速的信号直接把它们连接起来。

类时事件则不同。对已经选定未来方向的物理变换，一对类时事件的先后不会被倒置。相对论不是宣告所有先后都相对，而是精确说明哪类先后具有跨系稳定性，哪类排序依赖时间切分。把这两者混在一起，会把“远处分离事件的同时性无普遍固定值”夸大成“原因和结果可以任意交换”。新几何增加的正是这种分类能力。

33.4 同一物体为何拥有不同的空间截面

把一根理想尺的两个端点各自画成世界线，就得到两条平行的历史边界。K 系取某个共同时间，横截这两条线，得到它在 K 系中的两端位置；K′系按自己的共同时间截取，则选出另一对事件。在时空图里，不同同时条件对应不同切法。尺长变化不再是两句互不相容的报告，而是同一历史结构在不同测量切面上的结果。

这里的切面并不是真有刀把物体切成片，而是测量如何选取事件的数学表达。若一根尺始终匀速并保持相同内部状态，不同切面仍然能够按照统一规则相互联系。若尺正在弯曲、升温或断裂，不同部分的历史就更复杂，原来简单的平行边界模型失效。这时需要补入动力学，不能要求一幅匀速世界管图解释所有材料过程。

把对象理解为历史结构，还能纠正一种容易出现的误会：物体的同一性不需要每个时刻的全部属性都相同。一个人的位置、体温和所持物品持续改变，研究者仍可能通过连续历史与因果联系追踪同一人；物理对象也可以通过规定清楚的历史连接保持身份。世界线为这种追踪提供结构，但“跟踪的是同一对象”仍需要模型与实验依据，不能只凭画线就解决粒子产生、湮灭或对象分裂的问题。

本书所谓新物理世界创造，发生在这种构成单位的改变中。孤立的空间形状退到特定截面的地位，事件关联与历史结构成为更有组织力的对象。不是说三维物体从此不存在，而是说它的形状、长度与同时状态必须交代所依赖的切分。不同层级能够共存，但不能不加条件地把某一层级的规定提升为全部世界的本体。

这一解释也有边界。若主张世界管表明全部未来已经以与过去完全相同的方式存在，那就进入更强的时间哲学立场。数学可以把一段完整解表示在同一张图上，却不能仅凭“图上都有”推导一种唯一的存在论。初始条件、动力学、概率解释与现实中的知识开放性仍是不同问题。四维形式提供重要约束，但没有凭一个图形替哲学完成所有判决。

33.5 固有时：两次相遇之间，路径怎样进入时间

上一章留下了一个问题：两只钟离开后再次相遇，究竟比较什么？在平直时空中，一只理想钟沿类时路径经历的固有时，可写为 $\tau=\int \sqrt{\left(1-v(t)^2/c^2\right)}\,dt$。这里的 $v(t)$ 是相对所选惯性系的瞬时速度。积分将局部时间间隔累积为整段路径上的钟读数；不同坐标计算同一条路径，应给出相同结果。

这个公式使微积分的作用格外清楚。局部关系不是随意把若干钟慢公式拼起来，而是沿实际路径逐段计算；整体结果也不只由两个端点的坐标差决定。对弯曲路径，端点间的时空间隔与路径累积的固有时一般不同。只有在相应的惯性直线路径等特定条件下，两者才按简单关系一致。把端点关系与路径积分混用，会制造假的悖论。

考虑一项理想化往返。甲留在某惯性系原点，乙以 $0.6c$ 离开，在该系经过五年后立即改变方向，再以同样速率返回，五年后与甲相遇。忽略极短转向阶段的持续时间，甲经历十年；乙每一段的时间因子都是 0.8，两段共经历八年。这个例子假定理想钟并把瞬时转向视为有限平滑转向的近似，不能把它直接当成可实施的无限加速度装置。

年龄差并不是因为一个人“觉得”另一个人慢，也不是在某个转向瞬间凭空扣掉两年。它来自两条完整历史累积的固有时不同。转向使乙不能从出发到返回一直使用同一个与自身共动的惯性系，因而破坏了把全过程解释为两个对称匀速系的企图。至于具体加速度造成的时钟机械误差，则属于另外需要控制的装置问题。

还可以看到一个与欧氏直觉相反的性质。在给定的两次类时分离相遇之间，平直时空中的惯性直线路径具有最大的固有时；绕路的类时路径可以累积更少。这里不能用普通几何“直线最短”的口号判断，因为度量的符号不同。数学对象的规则不是日常空间感觉的自动延长，正是这种不自动延长，才使新模型能够解释旧直觉难以协调的现象。

33.6 四维不是四个同性质的方向

把时间乘以 c 之后，它与空间坐标具有相同长度量纲，这有助于统一表示，却不意味着时间与三个空间方向在所有物理意义上相同。度量的符号、因果方向与可实现路径

都区分了时间性和空间性。人可以在房间里向前后左右移动，但不能据此把自己的世界线任意折回过去；这一差别不是单位换算能够抹去的。

同样，时空图上的斜率也需要按照坐标解释。横轴画 x、纵轴画 ct 时，世界线的倾斜程度与速度相关，但图纸比例一改，视觉角度就会改变。不能用普通量角器测出的角度直接当成所有图中共同的物理量。标准图形把光线画成四十五度，是选择了相同尺度后的便利，而不是光在宇宙中拥有一个绝对的纸面倾角。

这正是本书需要区分数学对象与感官经验的地方。人眼看见图上的距离、弯曲和交叉，属于一项具体视觉 SDE 活动；图所表示的时空间隔与因果类型，则通过符号规定和推理规则进入理念结构；探测器的实际记录又以另一套完整 SDE 参与检验。三者可以纠缠，但不能把视觉长度直接等同于不定度量的间隔，也不能把理念中的四维结构宣称为眼睛直接经验到的四维实体。

一个图画得漂亮，并不能证明模型正确；一个模型难以直接想象，也不构成反证。判断应追踪图形背后的运算能否与实验操作连接。若图只是帮助研究者记住已核实的关系，它是有效工具；若图诱使人把普通空间直觉带入不同度量，则需要主动限制图的解释权。可视化的力量与危险，来自同一个事实：它能让抽象关系迅速获得熟悉感。

从 SDE 立场说，数学并非站在发生之外的透明镜子。选择哪些事件、怎样建立等价、哪些变换保持结构，都参与理念对象的构造。但数学构造也不是任意想象：内部一致性与经验接口共同约束其使用。四维时空的创造，是一种受严格限制的结构创造；它既不能被缩小为画图方便，也不能被扩大为数学自动生产全部现实。

33.7 一个不变量还不足以保证找对了物理

假如有人提出一种新变量，使某组实验数字在重新表达后完全不变，这是否已经发现自然规律？未必。通过适当定义，很多数据都可以被压缩成不随某种人为操作改变的量。关键在于：变换是否具有物理意义，不变量是否约束未参与定义的新情况，以及这个结构是否能够连接相互独立的测量。形式稳定必须承担经验责任。

本章可以沿 D—E—S 的倒逼路径检验这种责任。已有运动比较 D 要求长度和时间联合变化，旧的独立属性组织不能保持协调，于是 E 被改造成时空间隔与变换结构；随后必须由新 E 生成新的可比较显露 S，例如对同一对事件在不同惯性系中的分类一致、同一路径的固有时一致。最后再用独立定位或时钟比较检查这些显露，不允许只重画同一组数据就宣布成功。

反方向的 S—D—E 再创路径则用于异常。如果两套事件记录算出的间隔不一致，首先追踪 D 中有没有选错事件、混淆单位或使用未同步时钟；若这些都得到独立排除，再检查 E 中的平直时空假设、惯性坐标条件或传播模型。异常可能表明实验误差，也可能表明简单模型不够。研究者不能在证据出现之前替所有异常指定同一个原因。

三方程在这里变成具体的检验纪律：一个声称能生成 S 的结构，必须说明其 D；一个重新组织 D 的方案，必须保留已有 S 所施加的限制；一个重建 E 的主张，则必须同时回应 S 与 D 的张力。三原理要求两维冲突推动第三维改变，并让新状态回到原问题接受检验。不能把不一致的数据删掉后称为统一，也不能因为所有环节相互依存，就说永远无法定位哪一环出了问题。

还要区分坐标变换与模型扩展。任意改用弯曲坐标，并不会自动使原本平直的时空变成有曲率的时空；坐标表达复杂，也不等于出现了引力。另一方面，若研究范围需要引入真正的几何曲率，平直时空的全局公式就不能原样覆盖。这里为后续引力研究留下问题，但不提前用一个“几何化”词语把狭义相对论与广义相对论之间的创造距离抹平。

33.8 从恒常属性转向可保持的关系

经过这一章，“物理对象是什么”又被推进了一步。一个点状对象的运动不再只是某一空间内位置不断变化，它的历史可以作为世界线进入整体结构；一根尺的长度不再承担跨所有参考系的数值恒常性，世界管与各同时切面的关系承担了更广的可比性；一只钟的经历不再由全宇宙共享的单一时间自动规定，而要沿它的路径累积。

这些改变没有取消对象的稳定。相反，稳定性被安排在更有约束力的地方：事件重合可以保持，间隔类型可以保持，同一路径固有时可以保持，规律在适当变换下的形式可以保持。某些分量改变，恰好是为了让更整体的关系保持。物理学因此不必在“所有数字一样”与“完全没有共同世界”之间二选一。

属性纠缠论可以从这里提出一个值得继续检验的论题：在重大理论转变中，原来由单项属性承担的同一性，可能转交给多个属性的共同转换结构。长度与时间的例子已经显示这种可能性；但不能据此断言所有科学革命都遵循完全相同的机制。其他领域可能出现对象层级改变、概率结构改变或新的测量不可兼容性，需要分别研究。

对本书的主轴而言，闵可夫斯基的创造不是仅给一个先在图景换上更优雅的数学外衣。世界点、世界线、间隔和光锥改变了哪些关系能够成为基本对象，改变了怎样把局部经验聚合成可工作的整体。这一理念系统进入时钟、信号、轨迹与实验判断之后，才

形成新的物理世界的现实运行能力。数学表达和现实操作之间的相互回写，是这项创造继续发生的条件。

与此同时，四维时空不是最后答案。它给出了强大的共同结构，却仍需要回答具体物体怎样受力、能量动量如何协调、哪些系统适合点状近似，以及更复杂情况下几何本身是否参与动力学。一个成功整体会创造新的问题，而不是让后续研究失去对象。下一步的工作，是让这种新结构进入具体的物理守恒与相互作用，检验它能否从测量协调走向更完整的动力学创造。

本章参考资料

[33-1] Minkowski, H. Raum und Zeit. 科隆演讲，1908 年 9 月 21 日；发表于 1909 年。核读英语转录：https://en.wikisource.org/wiki/Translation:Space_and_Time。重点第一至三部分；该版说明沿用并校订 Saha 译本、恢复原记号。本文只作简短历史定位，数学示例均采用现代实坐标和时间正号约定。

第三十四章 广义相对论
时空几何与物质怎样互为条件

34.1 如果舞台也需要解释

上一章把事件、世界线和时空间隔组织成一个共同结构。但在那里，平直时空的几何仍然先被给定，物体在其中运动，理想钟沿路径累积时间。这套框架具有强大的解释能力，却留下一个新问题：引力应当怎样进入？如果引力只是向已有时空中的物体添加一种力，为什么不同材料的测试体会呈现共同的自由落体规律？这种普遍性为何提示我们重新研究运动结构本身？

自由落体的共同性构成了一条特殊线索。忽略空气阻力和其他非引力作用，在相同初始条件下，不同测试物体具有相同的自由落体运动。这并不是说所有现实物体在任何条件下都沿同一路径，也不是说带电物体可以忽略电磁作用；它指出，在测试体近似下，引力响应具有值得从运动结构本身研究的普遍内容。

爱因斯坦一九一六年的系统论文把等效原理、一般坐标描述、度量与引力场方程组织起来，并讨论了物理检验。[34-1]本章采用现代记号说明这些问题，不把后来的全部概念和实验成就归到这篇论文中。与前章相比，关键变化在于：确定钟尺与传播关系的几何结构，开始作为需要通过方程求解的物理内容，而不是永远置于研究之外的固定容器。

把这个变化称作舞台也进入演出，只能作为入口。舞台比喻容易误导，因为时空几何不是一块被物质压弯的舞台布，也不需要更高维的外部空间才能定义曲率。真正需要理解的是，局部的计时测距关系、邻近自由落体的相对运动与能量动量分布，怎样共同进入一个动力学系统。

本书据此提出的发生学判断是：广义相对论重建了何种物能够成为物体、何种运动能够被认作自由运动、何种距离和时间能够被共同测定的条件。它的创造性不只是把引力的解释换了一套词，而是让先前分别处理的物体、测量与几何发生了新的结构性纠缠。这个判断需要下面的具体运算和反例支撑，不能靠一个宏大的“整体”概念自行成立。

34.2 自由下落的房间与被拉动的房间

想象一个足够小、观察时间足够短的封闭房间。如果房间及其中物体一起自由下落，实验者松开手中的小球，会发现它在自己的附近保持漂浮。不是因为地球突然消

失，而是房间和小球具有近似相同的下落运动。支撑身体的压力解除以后，某些原来被视为引力的局部效应可以消失。

再想象远离其他影响、被火箭持续拉动的房间。小球不受地板支撑时保持自身惯性运动，地板却向它加速，实验者因此看见小球落向地板。用房间内的坐标描述，需要引入类似重力的效应。这两个例子把一类局部引力现象与加速描述联系起来，但它们不能证明所有引力场在任意大小的区域内都与一个加速房间完全相同。

限制来自潮汐。在地球附近同时释放两个相隔一定距离的小球，它们指向地心的自由落体方向并不严格平行；若一个略高一个略低，它们的加速度也可能不同。把房间缩得足够小、观察时间缩得足够短，这种相对效应可以近似忽略；扩大范围后，它会显露。等效原理的局部性不是无关紧要的修辞，而是辨认几何曲率的入口。

这也让我们重新理解一个人站在地板上的感受。身体受到地板的支撑，随身加速度计可以记录非零的固有加速度；理想自由落体者的加速度计则读零。某个坐标系中位置的二阶变化，与加速度计的读数并不总是同一件事。研究者必须说明使用的是坐标加速度还是固有加速度，不能只因都叫加速度就直接相互替代。

在 SDE 框架中，这是一项关于属性身份的具体教训：身体感受到的支撑、坐标中的下降与两个自由物体之间的接近，是三种不同的显露组织。它们可以共同参与引力研究，但每一项都需要自己的差异序列与特征纠缠。把它们简单合称“主体感觉到重力”，既遮蔽仪器判定，也遮蔽跨过程稳定的关系。

34.3 一个坐标弯了，不等于时空弯了

先从没有引力的平面做一个简单检查。直角坐标中的线元为 $\mathrm{d}l^2=\mathrm{d}x^2+\mathrm{d}y^2$，改用极坐标后变成 $\mathrm{d}l^2=\mathrm{d}r^2+r^2\mathrm{d}\varphi^2$。表达式中出现随位置变化的系数，圆周方向的坐标网格也弯了，但平面的内禀曲率仍为零。因此，“公式系数变了”“坐标线弯了”和“空间有曲率”是三个不能自动等同的判断。

时空中同样如此。在平直时空中使用加速坐标，可以产生非零的联络系数，使自由粒子的坐标运动看起来弯曲；这并不意味着已经生成真正的时空曲率。联络帮助我们比较邻近位置上的方向与变化，曲率则检验这种比较在局部闭合回路或邻近自由运动之间出现的不可消除内容。二者有联系，却不是同一个数学量。

在任一合适事件附近，可以选择局部惯性坐标，使度量在该点具有平直形式，并使联络在该点消失。但一般不能靠同一次坐标选择，把一个有曲率区域的所有曲率效应消

掉。类比地球表面：在很小的范围内，地图可以近似平坦；扩大后，路径和角度之间的差别不能通过给街道重新编号而全部消除。

这里还必须防止一种本体论偷换。坐标系不是心理主体的同义词。两位研究者可以使用同一套坐标，一位研究者也可以切换多套坐标；改变记号不等于物理干预，更不等于个人意识创造了一次新的引力场。坐标表达参与知识发生，但其可变性受到表达同一物理关系的严格要求约束。

真正有力量的发生学分析，应当追查哪些变动只改写了表示，哪些改变了实验安排，哪些改变了被研究的几何或物质状态。若三者都被称作“世界因主体而改变”，理论便失去区别力。广义相对论的数学精度恰好提供一项反约束：某些在坐标中很显眼的变化没有独立物理意义，而某些在一点看不出的差异，却会在邻近过程的比较中稳定显露。

34.4 潮汐让曲率获得一个实验出口

把两粒彼此邻近、互不施力的测试体同时释放，跟踪它们的相对位移。即使每一粒都沿自由落体路径运行，它们之间的距离仍可能改变。这种相对加速能够揭示曲率，因为单独把某一粒的运动描述为局部惯性运动，并不能同时删除所有邻近路径之间的相对变化。

现代语言用测地线偏离方程表达这种联系。设 u 为参考自由落体路径的切向量，ξ 为指向邻近路径的分离向量，方程把 ξ 沿路径的二阶协变变化，与曲率张量在 u、ξ 和 u 这三个方向上的作用联系起来。不同曲率符号约定会改变公式的整体符号，但可观测的相对加速预测必须一致。这里的协变微分要求同时处理向量变化与几何比较规则，不是对每个坐标数字机械地连求两次普通导数。

不熟悉张量也可以理解这一关系提出的任务：不是只测一粒物体掉得多快，而是测邻近自由路径之间怎样分开或靠近，并把这种变化与几何关系对应。不同方向上的分离可以产生不同变化，故单个数常常不足以表达全部潮汐结构。这正是使用具有多个分量的曲率对象的理由之一。

在牛顿弱场近似中也能看到相同入口。若引力加速度为 $g(x)$，两物体相距很小的 ξ，则加速度差近似为 g 在该处的空间梯度作用于 ξ。均匀加速可以被共同运动消去，梯度效应却留下相对变化。广义相对论把这种局部相对运动组织进时空几何，而不是仅仅在旧力学中多加一个梯度项。

实验出口的存在对哲学解释很重要。曲率不是一幅橡皮膜图带来的审美感，也不是研究者在某套坐标中看到轨迹弯曲后的直觉结论。它必须能够约束钟、光束或自由物体之间的比较。即使某个曲率标量恰好为零，也不能立即断言整个曲率张量为零；某些结构需要更完整的关系检验。用一个漂亮数字代表全部几何，仍可能重演旧的属性简化。

34.5 场方程把哪两种东西联系起来

用现代通行记号，包含宇宙学常数的爱因斯坦场方程可写为 $G\mu\nu+\Lambda g\mu\nu=(8\pi G/c^4)T\mu\nu$，其中 $G\mu\nu$ 是爱因斯坦张量，$g\mu\nu$ 是度量，$T\mu\nu$ 是能量动量张量，μ 与 ν 标识时空分量。常数 G 是万有引力常数，不是 SDE 三方程中的函数名称。这个包含 Λ 的现代形式不应被说成一九一六年论文已采用的全部原样表达。

能量动量张量不仅记录静止质量密度，还包括能量流、动量流与应力等内容。因此“物质告诉空间怎样弯”这句流行口号，如果被理解成只有一堆静止质量决定一个三维弯曲形状，就遗漏了主要结构。度量同样规定钟尺关系与运动方程中的几何，而物质的演化、应力和流动又参与决定几何。二者必须共同求解。

不过，“互为条件”不等于一个位置上的 T 就唯一决定该处及全宇宙全部几何。场方程还需要初始条件、边界条件、约束条件和适当的物质方程。不同整体解可以具有不同的几何结构；在没有局部物质的真空区域，也可以存在非零曲率，例如外部引力场与引力波。因此不能把关系简化成每处放多少物质、该处就有多少曲率的一张一一对应表。

真空中 $T\mu\nu$ 为零，不意味着所有几何自由度自动消失。场方程约束的是曲率的特定组合，并非把曲率张量的每个分量都逐一置零。这个区别对于理解远处天体如何影响一段没有物质的传播区域尤其重要。若把“空”当成“任何结构都没有”，就无法理解真空几何的物理作用。

同样，几何与物质联合成立也不证明宇宙中的一切性质都已经被这一个方程解释。具体物质模型、量子效应和测量实现仍有各自问题。本书借场方程论证的是一种更严密的共同规定方式，不能把它当作 SDE 全部本体命题的数学证明。物理方程与发生学解释之间可以形成研究接口，却必须保留不同层级的证据责任。

34.6 从新方程返回旧引力，不能只靠说相容

一个新理论若声称能够处理普通引力现象，就必须说明旧近似怎样从中出现。在弱场、缓慢变化与低速条件下，可以把时间分量写成与牛顿势 Φ 有关的近似形式。采用时

间正号约定，静止钟满足 $d\tau$ 约为 $(1+\Phi/c^2)\,dt$，其中还假定势的绝对值相对 c^2 足够小。这个关系已经显示：不同高度的钟读数比较，与引力势差有关。

在同样适用的近似中，空间运动回到 d^2x/dt^2 约为 $-\nabla\Phi$，场方程则回到泊松方程 $\nabla^2\Phi=4\pi G\rho$。这里的回归需要具体展开与物质条件，不是把所有张量符号删去就能得到。压力不可忽略、速度很大或场很强时，简单的牛顿表达不再提供全部内容。

可以用一项量纲检查建立直觉。若两只静止钟之间的势差为 $\Delta\Phi$，领先阶的相对走时差约为 $\Delta\Phi/c^2$。势差的量纲与速度平方相同，除以 c^2 得到无量纲比值。若有人把势差直接写成时间差而没有经历时长或其他尺度，表达便缺少必要条件。量纲检查不能证明理论正确，却能迅速排除不少不成立的改写。

取一个自拟的地面近似例子：两只静止理想钟相差一百米，暂取重力加速度 $g=9.8$ 米每二次方秒，c 约为 3×10^8 米每秒，并忽略地球自转及其他校正。势差约为 $gh=980$ 平方米每二次方秒，较高钟的相对走时优势约为 1.09×10^{-14}。经过同一近似坐标时间的一天，累积差约为 0.94 纳秒。这个数字不是本章引用的一项实验结果，而是显示尺度的模型计算；小到日常感官无法区分，并不等于在精密时钟比较中可以永远忽略。

这项检验也说明，所谓时空变化不等于计时装置随便失灵。具体钟可能受温度等非引力因素影响，但理论要预言的是在控制这些因素后，不同运行条件下理想时钟之间具有共同的比较关系。只有多种装置和操作能够进入同一关系，几何计时才具有超出某一仪器的物理意义。

从发生学看，旧引力世界没有因新方程出现而被全盘删除。它作为特定条件下的稳定近似重新获得位置。新旧之间的关系不能只用“后者更正确”结束，而应展示哪些变量保持、哪些解释角色改变、哪些条件原来被封装。这样的回检既保护日常与工程经验，也给强场和高精度的新问题留下明确入口。

34.7 光线弯曲与引力透镜如何形成共同证据

光在弯曲时空中沿适当的零测地线传播。在静态、球对称、弱场且光线远离强场区的近似下，点质量造成的总偏折角领先阶为 α 约等于 $4GM/(bc^2)$，b 为冲击参数。这个表达与简单把光当作受牛顿力作用的高速小粒子所得结果不同；完整系数来自相对论几何。它不是仅凭等效原理一句话就可以无条件推出的全部光传播定律。

一九一九年日食远征的星位观测，由戴森、爱丁顿与戴维森在一九二〇年的报告中整理，用来比较太阳附近光线的偏折。[34-2]本章不重演一张照片当场证明整个理论的

传奇，而关注证据结构：必须把日食时星像位置与对照位置连接起来，处理仪器、尺度与误差，再与具体预期比较。可见图像成为引力证据，需要一条完整的差异运行。

引力透镜把这个结构推广成更丰富的形态。光源、透镜质量分布与观察位置共同决定可能出现的像的位置、放大和形状。多个像不表示原天体被复制成多个独立天体，而是不同传播路径把同一源的辐射带到观察系统。图像在视觉经验中是多个，在源模型中可以对应同一个对象；二者之间需要传播几何来建立身份联系。

反演也不是天然唯一。不同的质量分布、源结构或传播条件可能在有限数据上形成相近图像，因此需要额外信息来限制模型。若只因某幅弧状图像看起来符合理论，就宣告内部所有参数已经被准确测定，便高估了证据。几何理论创造新的识别能力，同时也创造新的反问题；能力增长与不可辨识性需要一起研究。

SDE 解释在这里应当追踪：视觉显露怎样经标定、对照与模型约束取得天体身份；实际传播关系怎样限制解释；新解释又怎样决定下一次观测。它不能把“图像由系统生成”偷换成“远处天体由人的视觉愿望制造”。创造的对象是一个能够被持续检验和应用的物理规定系统，实验之外的抵抗仍然是这一系统不能取消的条件。

34.8 引力波让几何不再只是静态形状

当几何具有动力学，研究者便要问扰动是否能够传播。在适当的弱扰动近似中，可以从场方程研究相对平直背景的小扰动，区分其中可以通过坐标或规范选择消去的部分，与对测试体相对运动产生物理效应的部分。并不是度量表达式里任何摆动都自动是一列可探测引力波。

引力波的一个物理出口，是使自由测试体之间的相对距离或光传播比较出现特定的时间变化。用小尺度探测器近似描述时，这与潮汐结构相连。仪器并不直接伸手抓住一个抽象的度量分量，而是比较光路、相位和测试质量的相对行为，再将这些记录与传播模型联系起来。

LIGO 与 Virgo 合作组发表的首次直接探测论文报告，两台 LIGO 探测器于二〇一五年九月十四日记录到信号，其波形与双黑洞并合的广义相对论预测相符，结果于二〇一六年公布。[34-3]这是一项具体历史事实，而不是所有引力模型已经永久退出研究的声明。本章不列未经本次逐项复核的参数，也不把模型拟合与原始仪器读数混成同一种资料。

一条可信证据链需要多项约束：两个地点的记录能否按传播时差协调，仪器与环境噪声能否解释信号，校准过程是否可靠，采用不同分析方法时是否保留主要结构，候选

波形对未用于调整的特征有什么预言。理论可以帮助从噪声中识别结构，但这种帮助必须受独立检验限制，不能靠只保留“像理论”的片段完成自我证明。

把引力波转换成声音，有利于人类感知频率随时间变化的模式，但这种听觉展示不是声波穿过宇宙空间到达耳朵。实际记录、数据变换和听觉经验是不同而相联的完整SDE。只有清楚交代转换，视觉与听觉才会扩展物理经验，而不是制造关于感官来源的误解。

34.9 用异常计时走通几何与物质的互检

设想一组跨高度的时钟比较出现稳定偏差。直接把它称为新引力，比直接把它称为仪器错误都过早。先沿 S—E—D 进入：从走时显露 S 检查解释所需的纠缠 E，其中包括高度、速度、信号传输、钟的内部响应以及所用引力近似；再组织有区分力的运行 D，例如交换时钟位置、改用另一种钟、独立测定运动和传输延迟。不是把可能因素列得越多越好，而是让每种候选产生不同的可检验后果。

若交换位置后偏差随钟具走，优先检查装置；若随地点走，再比较环境条件与几何预期；若不同原理的钟和不同传输方案都保留同一偏差，就需要进一步检查物质分布、近似条件或理论结构。一次改造不必直接上升到基础革命，却必须允许更高层前提在证据要求时进入检验。

随后可以沿 D—E—S 倒逼，主动改变已掌握的运行条件，修改候选结构，事前预言新的走时关系。改变后只有回到 S 接受独立比较，才完成循环。S 与 E 之间的张力推动 D 重组，D 与 S 之间的张力推动 E 改造，新的 D 与 E 联合再生成 S；这使三原理成为具体的判断纪律，而不是关于协调和美感的空泛赞词。

三方程在此分别要求说明结果如何由操作与构成产生、操作如何由结果与构成选择、构成如何在结果与操作的约束中重建。没有哪个环节能够永远免责。若只把几何称为客体、仪器称为互动、科学家称为主体，再分别塞进 S、D、E，便完全错过了三维相互规定的分析；每种号位属性都需要追踪其完整发生过程。

广义相对论因此给本书留下的不是“关系决定一切”的便捷口号，而是一项更艰难的研究范式：让共同规定进入方程，让方程返回可区分的经验，让新的经验有能力改写共同规定。物质与几何互为条件，并不使检验变得不可能；它要求我们更精确地组织检验。下一章的诺特定理，将进一步展示变换结构怎样与守恒量共同成立，以及为什么不能离开条件把守恒当作永远不变的口令。

本章参考资料

[34-1] Einstein, A. Die Grundlage der allgemeinen Relativitätstheorie. Annalen der Physik 49 (1916):769–822. 英译转录：https://en.wikisource.org/wiki/The_Foundation_of_the_Generalised_Theory_of_Relativity。仅据其结构作简短历史定位，公式采用现代记号。

[34-2] Dyson, F. W.; Eddington, A. S.; Davidson, C. A Determination of the Deflection of Light by the Sun’s Gravitational Field. Philosophical Transactions A 220 (1920):291–333. https://royalsocietypublishing.org/rsta/article/220/571-581/291/40498/IX-A-determination-of-the-deflection-of-light-by

[34-3] Abbott, B. P. et al. Observation of Gravitational Waves from a Binary Black Hole Merger. Physical Review Letters 116 (2016):061102. https://arxiv.org/abs/1602.03837

第三十五章 诺特定理
为什么变换与守恒可以共同规定物理

35.1 守恒究竟是不变的什么

一辆小车沿无摩擦直轨匀速运动，它的位置不断变化，动量却可以保持。弹簧上的小球往复振动，速度与位置轮流改变，适当条件下总能量保持。所谓守恒不是一切停止变化，而是在持续变化中有某个结构能够保持。问题在于：为什么有些组合可以保持，另一些不能？我们是在每种现象里分别碰巧找到一个常数，还是能够从更共同的条件解释它们？

前面章节已经反复出现变换：更换参考系、改变坐标、比较相同规律在不同表达中的形式。诺特的贡献把这一线索与变分结构连接起来。她一九一八年的《不变变分问题》讨论连续变换与变分方程之间的关系，区分有限参数群和依赖任意函数的变换所产生的不同结果。[35-1]本章从一个粒子模型出发，理解这种连接为什么能够改变守恒问题的身份。

先作一个重要限定：不是凡是看起来对称的图形，都自动对应一个守恒量。诺特定理需要关于作用量、变换、边界和运动方程的明确条件。离散对称性与连续对称性的结果也不能直接混同。把一幅左右对称的照片指给读者看，然后宣布能量必定守恒，缺少了必要的理论链条。

同样，定理不是给物理世界添加一条来自数学外部的命令。研究者建立某个作用量模型，检查它在何种变换下保持相应结构，再由定理推出沿运动方程成立的关系。如果模型本身遗漏了外界驱动，推出来的守恒可能不适用于实际系统；如果把实际变化错误地记成封闭系统内部变化，数学不会替研究者修正对象选择。

本书关心的创造发生在这里：守恒量不再只是物体身上孤立带着的一项不变量，变换也不只是把同一现象换个写法的工具。它们通过作用量与实际运动共同规定彼此。这个共同规定有严格的数学条件，能够计算，也能够失败。因此它为属性纠缠论提供具体研究案例，却不能被直接说成整个 SDE 本体论已经得到证明。

35.2 从一个作用量开始，而不是从口号开始

设一个粒子的广义坐标为 q，速度为 $\dot{q}$，模型由拉格朗日量 $L(q, \dot{q}, t)$给出。作用量记为 $A=\int L\,dt$。本书不用 S 表示作用量，以免与 SDE 中的显露 S 混淆。变分问题研

究一族相邻的候选路径，在端点条件规定后，实际路径满足作用量的一阶变化为零。这不表示粒子有意识地试走所有道路，而是一种组织动力学的数学条件。

把路径改变写为 δq，且固定端点，则 δA＝∫［（∂L／∂q）δq＋（∂L／∂q̇）δq̇］dt。对第二项分部积分，端点项因 δq 在端点为零而消失，剩余积分要求 d/dt（∂L／∂q̇）－∂L／∂q＝0。这就是欧拉—拉格朗日方程。推导的力量在于，它把整段路径的变分条件转化成局部微分关系。

然而，端点项不能因为常见就被永久忘记。若研究的变换移动端点、改变时间参数，或者作用量只相差一个边界贡献，推导就需要相应调整。物理学中很多细节正藏在这里：同一个局部微分方程可以对应不同边界问题，守恒量的表达也可能受到边界项影响。把分部积分的边界一律划掉，会把问题最有意义的部分一并删去。

定义广义动量 p＝∂L／∂q̇。沿满足运动方程的路径，有 ṗ＝∂L／∂q。这个关系立即给出一个入口：若 L 不显含 q，便有 ṗ＝0。但这还只是某种坐标选择下的方便表达。接下来要看，为什么“不显含 q”与一个连续变换有关，又怎样在更一般情况下表达同一结构。

在发生学层面，q 的身份也不能被跳过。它可能是位置、角度或其他合适变量；不同选择可以表达同一个系统，也可能对应不同的模型简化。L 不是与对象选择无关的一行万能公式。变量、端点、约束和相互作用共同确定研究的系统，定理才有明确作用对象。

35.3 平移对称怎样让动量出现

考虑固定时间下的微小连续变换 δq＝εξ(q，t)，其中 ε 是常数小参数。假设变换使拉格朗日量改变为一个全时间导数：δL＝εdB／dt。沿运动方程，可以把 δL 改写成 εd(pξ)／dt，因此 d(pξ－B)／dt＝0。于是 J＝pξ－B 为守恒量。这是本章所需的一维固定时间版本；它没有覆盖所有时空变换，却足以展示对称与守恒之间怎样逐步相接。

取自由粒子 L＝mq̇²／2。把整条路径沿直线平移 ε，速度不变，拉格朗日量也不变，所以 ξ＝1、B＝0，得到 J＝mq̇。无论把实验放在直轨哪一段，在这个理想模型中动力学结构相同；与这种平移相联系的动量沿实际运动保持。这不是“位置不变所以动量不变”，恰恰是位置可以整体改变而规律保持。

若加入势能 V(q)，L＝mq̇²／2－V(q)，平移引起的 δL 一般为－εV′(q)，不再具有自由粒子那种对称。运动方程给出 ṗ＝－V′(q)，动量因力改变。对于 V＝kq²／2 的弹簧

模型，选定的平衡位置进入了结构；把小球位置平移而保持弹簧固定，并不是同一个物理条件。

这个例子提醒我们，所谓“空间均匀”必须说明系统包括什么。如果把小球和整个装置连同支撑一起整体平移，可能讨论另一种更大的系统对称；如果只移动小球而把外部势场固定，讨论的就是子系统受到的位置依赖作用。两种变换不是同一个实验。系统边界选择直接影响什么量有理由被期待守恒。

一个更有意思的检查是自由粒子的匀速助推：$\delta q=\varepsilon t$，$\delta\dot q=\varepsilon$。此时 $\delta L=\varepsilon m\dot q=\varepsilon d(mq)/dt$，故 $B=mq$，$\xi=t$，得到 $J=pt-mq$。这个量守恒，表达匀速运动的轨迹结构。拉格朗日量本身并非严格不变，而是差一个边界导数；若只接受 δL 必须逐项为零，就会漏掉这一类重要对称。

再把问题扩到平面中的粒子，令 $L=m(\dot x^2+\dot y^2)/2-V(r)$，其中 $r^2=x^2+y^2$。绕原点的微小转动给出 $\delta x=-\varepsilon y$、$\delta y=\varepsilon x$，速度也按同一方式改变。因为速度平方与 r 都在这种转动下保持，$\delta L=0$，相应守恒量为 $x\,p_y-y\,p_x$，即垂直于平面的角动量。这里不需要先猜一个角动量再碰运气求导，而可以由变换及其对变量的作用构造它。

若把势改成 $V=(k_x x^2+k_y y^2)/2$，且两个劲度不同，原来的连续转动对称消失。直接求导得到角动量变化率为 $(k_x-k_y)xy$，通常不为零。然而只要劲度不随时间改变，能量仍可守恒。这个反例很重要：破坏一种对称，不表示所有守恒量一并消失。研究者必须逐项辨认哪一种变换仍满足条件，不能把“系统对称性降低”当作不加区分的结论。把两个劲度调回相等时，上述角动量变化率重新归零；这一可控恢复还提供了检查模型的直接办法，而不是只在概念上宣布对称已经恢复。

35.4 时间平移与能量：一条可以自己复算的式子

定义能量函数 $H=p\dot q-L$。沿运动方程求导，得到 $dH/dt=\dot p\dot q+p\ddot q-(\partial L/\partial q)\dot q-(\partial L/\partial\dot q)\ddot q-\partial L/\partial t$。前两组项分别抵消，因此 $dH/dt=-\partial L/\partial t$。对多个坐标，只需把相应动量速度项求和，结论同样成立。

如果拉格朗日量不显含时间，便有 H 守恒。在通常的标准粒子模型 $L=m\dot q^2/2-V(q)$中，H 就是动能加势能。但不应把任何模型里的 H 都未经检查叫作日常意义的机械能；速度相关耦合、约束及时间依赖变量变换可能改变表达。定理给出的是模型中精确定义的量，而不是保证每个叫能量的符号自动具有相同解释。

时间平移对称也不能理解为系统不发生任何变化。弹簧小球不断振动，但若模型中没有随日历时刻改变的参数，把同一初始条件安排在明天，运动规律仍相同。这种“开

始时间可平移”的结构与能量守恒联系起来。变化发生在路径内部，对称涉及整个动力学条件在某种转换下的保持，二者处于不同层次。

现在加入一个明确反例：令弹簧劲度 k 随时间改变，$L=m\dot{q}^2/2-k(t)q^2/2$。此时 $H=m\dot{q}^2/2+k(t)q^2/2$，但 $dH/dt=\dot{k}(t)q^2/2$。只要 $\dot{k}$ 不为零且小球离开平衡位置，这个子系统的机械能就可能改变。它并未违反推导；相反，正是推导告诉我们原来的守恒前提已经撤回。

例如在某一瞬间 $q=0.1$ 米，若 $\dot{k}=20$ 牛顿每米每秒，则当时 H 的变化率为 0.1 焦耳每秒。这个自拟数值只使用上述理想模型，不是实验测量。调节弹簧的外部装置做了什么，必须在更大的系统中研究。如果把装置和能源包括进来，可能建立新的总能量平衡；不能一面把驱动排除在系统外，一面要求子系统能量永远封闭守恒。

35.5 从一个常数到一条守恒流

粒子力学常用一个随时间不变的数表达守恒，场论则需要处理量在空间中的流动。现代写法中，连续对称可以在适当条件下给出守恒流 j，其局部关系为 $\partial_\mu j^\mu=0$。用时间和空间拆开，便是 $\partial\rho/\partial t+\nabla\cdot j=0$：某处密度的变化，与流入流出的差额相协调。

把这个关系在一个固定空间区域积分，利用散度定理可得 $dQ/dt=-\oint j\cdot n\,dA$，其中 Q 是区域内的总量，n 是边界外法向。这个式子立即纠正一种常见误读：局部守恒不保证任意有限区域中的总量不变。只要有净通量穿过边界，区域内部的量就会变化；若边界通量为零或无穷远衰减条件适当，才得到相应总量守恒。

水箱的类比可以帮助理解流量平衡，但不能据此把所有守恒流认作实际液体。电荷流、能量流和其他守恒量各有不同物理定义。共同的连续性形式，使我们能调用相似的积分工具；守恒的究竟是什么，仍由作用量、对称和实验对应决定。数学形式的重复使用，不能代替对象身份的说明。

边界的角色在这里十分清楚。一个实验盒中的能量减少，可能是辐射穿过边界，并不自动说明总能量不守恒。反过来，若研究者没有检查边界，只把缺少的部分命名为“某处流走了”，也不是完成解释。通量应当具有独立的模型或测量约束，否则守恒就变成可以无限补账的信念。

与前章广义相对论相接，能量动量的协变散度条件不应被草率写成任意坐标下都有一个简单、全局唯一的守恒能量。几何、边界与是否存在相应时间对称会影响全局量的定义。局部平衡、对称生成的守恒荷与特定整体条件下的总能量，是相关而不同的层

级。诺特思想的价值，正是迫使我们给出这些条件，而不是让“守恒”两个字替我们省略它们。

35.6 物理对称与规范冗余不是一回事

把整套孤立实验装置平移，可能得到物理位置不同而动力学等价的安排；把描述同一电磁场的一组势换成另一组势，则可能只是改变了表示。两者都使用变换语言，但变换与物理状态的关系不同。若把所有变换都认作创造了新物理状态，便会把数学描述中的冗余实体化；若把所有变换都认作只换记号，又会抹去真实实验之间的等价关系。

以电磁势为例，$B=\nabla\times A$，电场为 $E_{电}=-\nabla\phi-\partial A/\partial t$。取任意适当光滑函数 χ，令 $A'=A+\nabla\chi$、$\phi'=\phi-\partial\chi/\partial t$。由于旋度梯度为零以及混合导数相容，得到的电场和磁场不变。这里 $E_{电}$ 特意注明为电场，不是 SDE 特征纠缠 E。规范变换的这一简单例子表明，势的数值差异不能自动被认作场的物理差异。

诺特对依赖任意函数的变换所作分析，涉及运动方程之间的恒等联系，不能直接套成“每个位置有一个自由函数，所以每个位置都有一个独立新守恒荷”。第一定理与第二定理处理的问题相互联系，但不能以同一句通俗口号替代。对于规范理论，约束、边界和允许的变换范围尤其重要。

同时，不能把所有规范相关问题都用“只是冗余”一句话彻底结束。某些边界条件下的变换可能与非平凡的边界荷或渐近对称有关。究竟哪些变换被视为同一物理状态的不同表示，哪些在所研究边界下具有物理作用，需要具体分析。本章不把复杂问题压成普遍判词，而以电磁场不变的基本例子建立最低限度的区分。

这也给本书的“创造物理世界”命题增加了一项约束。不是每次重写势、换坐标或改变基底都创造一个新世界；如果全部相关物理预测保持，可能只是同一结构的不同表达。只有当允许对象、操作范围或共同规定发生实质重建，才应进一步讨论新的物理世界如何成立。创造必须有区分标准，不能因词语宏大而免除标准。

35.7 守恒量不见了，研究应当先查哪一层

设一项振动实验发现机械能缓慢增加。若只说“能量本来守恒，所以测量错了”，理论拒绝了证据；若立刻说“发现能量不守恒”，又可能忽略系统边界。一个更有效的入口，是把异常显露 S 与原来期待守恒的结构 E 放在一起：模型有没有显式时间依赖，支撑或外场有没有驱动，所计算的能量是否覆盖了实验对象？

沿 S—E—D 进入路径，可以据此安排有区分力的差异运行。测定 k 是否随时间变化，改变控制程序，监测支撑运动，比较关掉驱动后的结果。如果能量增长率与 $\dot{k}q^2/2$ 的预言相符，便获得一个具体解释；如果不符，不能用“总有外界作用”把缺口永久遮住，应继续检查阻尼、非线性与测量模型，或重新界定实验系统。

随后沿 E—D—S 沉淀路径，用已修订的时变作用量组织新的驱动序列，提前计算能量增长的符号和大小，再用未参与拟合的过程验证。例如在其他条件相同的近似下，把 $\dot{k}$ 改号，会改变对应能量交换项的符号。这种预言比事后给每段曲线补一个外力更有约束力，因为下一次结果能够否定它。

三方程在这里分别承担明确任务：S＝F（D，E）要求显露来自可说明的驱动和模型，D＝G（S，E）要求下一次操作针对候选解释的差别，E＝H（S，D）要求结构在数据与操作之间接受重建。三原理则要求两维的张力推动第三维改变，并在改变后回到原问题。不会变化的模型无法接受学习，随结果任意变化的模型也无法形成知识。

真正有研究价值的是区分两种失败：一个合法推导所用前提未被实验满足，和满足前提后预测仍不成立。前一种失败推动重新组织对象与条件，后一种才更直接挑战理论结构。两者都可能产出新知识，但不能用同一个“推翻”词语抹平。发生学不是降低逻辑要求，而是把逻辑推导成立所需的条件重新纳入研究。

35.8 一个守恒世界怎样在变换中成立

诺特定理给思想史带来的变化，可以从自由粒子的例子重新概括。空间平移没有改变模型的作用量结构，因而动量沿实际路径保持；时间平移没有改变模型，能量函数因而保持；作用量允许差一个边界项时，相应守恒量需要包含这个贡献。对称、动力学与守恒共同成立，每一环都可以写出明确条件。

这种共同成立不同于静止属性的简单相加。小车的位置不断变化，平移对称讨论的是整套规律怎样应对位置变换，动量守恒讨论的是沿运动路径什么保持。三个层次相互联系，却不能互相替代。把变换、轨迹与不变量分别安排在正确位置，才真正得到一种能够工作的物理整体。

从属性纠缠论看，守恒量的身份由一组特征共同构成：采用何种变量、作用量怎样写、变换如何作用、边界允许什么、实际运动满足哪些方程。它不是一个脱离这些条件而先行存在、等候被抄下的数字。但这种条件性并不使守恒任意化，反而让我们知道如何严密推出、测量和反驳特定的守恒主张。

同时，数学定理有自己的证据地位。只要前提满足，其结论不依靠一次次实验投票决定；真正需要经验检查的是所用模型与现实过程之间的接合，以及定理前提是否被正确实现。本书不能因为强调知识发生，就把已经证明的数学关系说成每次互动都能随意改写；也不能因为定理严格，就把模型选择免于经验检验。

这章因此为全书的物理创造主线加入一项新尺度：伟大的理论不仅给出更多公式，也可能让原本分散的规律在共同条件下相互生成。它使研究者知道何时应期待守恒、何时应寻找流出边界的量、何时应检查外部驱动，以及何时需要改变整体结构。守恒世界的力量，不在于禁止一切变化，而在于让变化有可追查的组织，并让组织始终能够回到具体问题接受检验。

本章参考资料

[35-1] Noether, E. Invariante Variationsprobleme. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (1918):235–257. 英译：M. A. Tavel, Invariant Variation Problems, Transport Theory and Statistical Physics 1 (1971):183–207。核读版本：https://web.stanford.edu/~cantwell/AA218_Course_Material/Resources/Invariant_Variation_Problems_Emmy_Noether.pdf。本章使用现代一维粒子模型作独立推导，历史定位限于有限参数与任意函数变换的区分。

第八编 从辐射到原子：量子物理的早期创造

第三十六章 黑体与普朗克
能量怎样成为可以计数的份额

36.1 从守恒追问到分配

前章讨论对称为什么能够规定守恒，但守恒并不告诉我们能量怎样分布。同样一焦耳，可以主要集中于低频振动，也可以分散在许多不同频率之中。总量相同，不代表颜色、温度响应和物质交换方式相同。当研究者把这些差别放进一只加热的空腔，问题便从“能量有没有丢失”推进到“什么分配能够稳定存在”。这条路线通向量子论，却不是以否定守恒为起点。

设想一只内部不透明的小炉，在壁上开一个很小的孔。进入小孔的光经过多次反射后很难原路逃出；从孔中出来的辐射，在适当平衡条件下能够接近理想黑体辐射。这里的“黑”不是指定常温外观，也不意味着不发光。对象的身份来自吸收、发射与热平衡的关系。加热后小孔可以明亮，而这一点与其近似黑体性质并不矛盾。

要把这种安排变成可靠研究对象，必须分清温度、频率区间、观察方向和探测响应。若仪器对红外迟钝，原始读数的低值不能立即解释成炉内没有红外能量。若不同波段使用不同探测方式，交接区的校准就可能决定一条谱线是否真的连续。所谓一条黑体曲线，是经过测温、分光、定标与误差控制后成立的共同对象，并不是肉眼看到的颜色被原封不动抄到纸上。

普朗克一九〇一年的论文开篇指出，新的光谱测量使维恩分布的普遍有效性受到挑战；他的任务是修正有关谐振子熵与能量的理论联系。[36-1]这里只保留这个足够重要的历史入口。不能把整个事件压成“经典公式先预言紫外无穷，普朗克立刻发明光子”的顺畅故事：现代教材便于比较的逻辑排列，与当年问题逐步形成的时间顺序并不相同。

36.2 频率区间本身参与规定曲线

用 u_ν 表示单位体积、单位频率区间内的辐射能量密度，则一个窄区间的能量为 $u_\nu d\nu$。也可以用单位波长区间的密度 u_λ 描述同一辐射，但不能只把公式里的 ν 换成 c/λ 就算完成。由于 $\nu=c/\lambda$，有 $|d\nu|=c\,d\lambda/\lambda^2$，所以 $u_\lambda=u_\nu c/\lambda^2$。两种密度针对的区间宽度不同，纵轴也随之改变。

这不是印刷格式的小事。每赫兹一个格子与每纳米一个格子，对同一束辐射划分了不同的份额。一张频率谱的峰值位置，不能简单取倒数便称作波长谱的峰值。若忽略这项转换，可以在没有任何新物理的情况下制造两种仪器互相矛盾的假象。对象是否相同，必须连同积分区间与转换规则一起判断。

再看空腔中的波。对于足够大的三维空腔，考虑两种偏振，单位体积、频率 ν 到 ν＋dν 之间的模式数近似为 $8\pi\nu^2 d\nu/c^3$。它来自允许波矢的空间计数：球壳体积随半径平方增大，而频率与波矢长度成正比。这一步没有使用能量量子。它只说明高频处有更多可能的振动方式，尚未说明每一种方式实际分到多少能量。

把平均每模式能量记为 ε 均，便有 $u\nu=(8\pi\nu^2/c^3)\,\varepsilon$ 均。整个问题因而可以暂时拆成两项：允许多少种振动方式，每种方式在平衡时获得多少平均能量。两项相乘形成可检验的谱密度。模式计数与能量分配承担不同职责；若测量不符，不能未经分析便把其中任何一项宣布为全部错误。

在连续能量的经典平衡模型中，谐振模式的平均能量为 kT，因此 $u\nu=8\pi\nu^2 kT/c^3$。频率越高，密度反而不断增大；对全部频率积分不能得到有限的总能量。有限炉子却不能在有限温度容纳这个无穷结果。矛盾并非一台探测器显示了“无穷大”，而是理论的一组前提联合产生不可接受的预测。此处的推导是现代回顾，不把完整论证冒充普朗克当时唯一的思路。

36.3 先数清楚我们究竟在数什么

普朗克的关键构造涉及把总能量划分为有限大小的能量元，再统计这些份额在谐振子之间的分配。[36-1]为了看清这种步骤改变了什么，下面使用独立的小数目例子。它不模拟原论文每一步措辞，只把组合规则展开到能够手工检查。

设三个可区分的谐振子分享四份相同能量元，每份大小为 ε。一个安排可以写为（4，0，0），另一个为（2，1，1）。三个位置标记三个不同谐振子，所以（4，0，0）与（0，4，0）不同；但同一谐振子内的四份能量元不附加姓名。问题是非负整数 $n_1+n_2+n_3=4$ 有多少个解。

可以用两条分隔线把四个相同记号分为三组，全部排列位置共有六个，从中选择两个放分隔线，因此共有十五种安排。逐类核对也得到同一结果：四份全在一个位置有三种；三份与一份分开放有六种；两份与两份有三种；两份、一份、一份有三种。总数确实为十五。这个核对比直接背诵阶乘更能说明“可区分”到底落在哪里。

推广到 N 个可区分谐振子、P 份相同能量元，分配数为 W＝（P＋N－1）!／［P!（N－1）!］。它与把每一份能量都贴上独立标签的计数不同，也与允许能量无限细分后在连续空间求体积不同。计数不是在一个完全固定的世界外面加一个数学按钮；哪些安排算同一种、哪些变化算新状态，已经参与规定了物理模型的基本结构。

还需要澄清两种“相同”。三个谐振子可以具有相同频率，但仍因位置等标识在分配中可区分。能量元不编号，也不等于已建立后来关于不可区分粒子的完整理论。若把这个有限分配公式直接翻译成“普朗克已经完成现代光子统计”，就把历史上尚待发展的对象身份提前塞入了早期计算。

对于宏观数量，利用阶乘的对数近似，并令每个谐振子的平均能量为 U，P/N＝U/ε，可以把每个谐振子的熵写成 k［（1＋U/ε）ln（1＋U/ε）－（U/ε）ln（U/ε）］。这里用熵的汉字名称，避免与显露 S 混同。该式省略大数极限下不影响主项的修正；对于刚才三个谐振子的例子，应使用精确整数计数，不能要求宏观近似逐项吻合。

36.4 一项离散规则怎样改变整条曲线

对上述熵关于 U 求导，得到 1/T＝（k/ε）ln（1＋ε/U）。移项后有 exp（ε/kT）＝1＋ε/U，因而 U＝ε／［exp（ε/kT）－1］。这段代数说明，分配方式一旦改变，热平衡时平均能量随温度的关系便随之改变。不是先任意画出理想曲线，再给它贴一个量子名称。

令能量元 ε＝hν，就得到现代写法的普朗克谱：uν＝（$8\pi h\nu^3/c^3$）／［exp（hν/kT）－1］。h 的量纲是能量乘时间，hν 才是能量。量纲检查能立刻排除把能量元写成 h/ν 的误植，却不能单独选出正确的分配规则。量纲提供必要约束，不代替物理证据。

引入无量纲比值 x＝hν/kT，平均能量与经典值之比为 x／（e^x－1）。当 x 很小时，e^x－1 约等于 x，于是 U 接近 kT，旧结果恢复。当 x 很大时，U 约为 hνe^{-x}，高频模式虽然数量增多，却很难获得一整份所需能量。多种模式与每种模式的热占有共同作用，形成高频端的衰减。

取一个自拟对比：若 x＝0.1，这个比值约为 0.951；若 x＝1，则约为 0.582；若 x＝5，则约为 0.0339。三个数说明新规则并非对所有频率平均打折。它在低频几乎保留经典分配，在高频强烈抑制平均能量。测量若只覆盖第一个范围，便可能难以区分两种结构；扩大可比较的频段因而具有实质意义。

还可以在同一频率把温度从 T 提高到 2T。若原来 x＝2，平均能量之比为（$e^{2}-1$）／（$e^{1}-1$），约为 3.72，并不等于二。经典模型预言平均能量随温度成正比，而离散模型在这个范围给出更陡的增长。这个算例提供另一条检查路线：不必只横向比较不同频率，也可以纵向改变温度而固定探测波段。

把普朗克谱对频率积分，再换元 ν＝（kT/h）x，可得总能量密度等于一个不随温度改变的系数乘 T^{4}。这里的四次方来自 ν^{3} 与 dν 共同贡献，不是对实验幂律另加的命令。既能回到低频近似，又能给出有限总能量和正确温度尺度，意味着新结构需要同时满足几项独立限制。

36.5 成功的公式还没有完成全部解释

科学史家吉尔哈特梳理过一项长期争论：普朗克在一九〇〇年前后究竟怎样理解能量元，不能仅由后来熟悉的量子能级倒推。他既讨论库恩、达里戈尔的连续性解释，也指出原始文字及后续发展的复杂性。[36-2]本章因此把可明确核验的计数步骤与尚有争议的思想承诺分开，不以一句“纯数学技巧”或“完整量子革命”替代历史判断。

谱公式的成功不能让早期理论中尚未解决的问题消失。能量元究竟只是统计计算的分割单位，还是物质能量交换必须遵守的物理间隔？这种规则约束谐振子、辐射本身，还是两者交换的方式？不同答案可以在某个平衡谱上相遇，却对光的传播和吸收提出不同承诺。仅靠一条平均分布曲线，未必能够把它们全部分开。

尤其不能把 hν 立即画成一颗具有微小半径的硬球。能量与频率的关系告诉我们一份交换的大小，并没有给出小球表面在哪里、内部由什么组成、如何穿过狭缝。把一个有明确计算作用的规则过快转化成机械图像，会让图像带入许多并未得到支持的额外假设。

也不能反过来说，这不过是拟合技巧，所以与物理世界毫无关系。一项计数规定若能够在温度改变、频段扩展和独立定标中保持，就已改变允许哪些分配共同成立。它为研究者提供新的比较对象和推导关系。问题不是数学还是物理的简单二选一，而是数学构造通过哪些接口获得物理约束。

一个竞争方案可以保持能量连续，却修改高频模式数；另一个可以保持模式数，却改变平衡分配。两者也许都能压低高频尾部。要判断何者更好，应检查空腔几何改变时的模式预测、低频极限和不同温度下能否使用同一组常数。若每换一种条件就要重新调整整条函数，它虽然可能拟合，却缺乏相同结构跨条件工作的力量。

任何一次拟合都还有仪器竞争解释。探测器高频效率降低，可能伪造一条衰减尾部；漏光和温度不均可能改变谱形。理论上的高频抑制与装置造成的信号抑制，在图上都可能表现为下降，只有改变仪器条件、独立测温和重复比较，才能使它们不再只是看起来不同的故事。

36.6 黑体研究中的两条完整路径

从 S—E—D 进入路径看，入口是不同温度和频段下稳定出现的谱形。研究者把它与模式数、热平衡及能量分配的特征纠缠放在一起，找出连续分配在何处产生压力；随后安排扩大频段、比较温度响应与核查探测效率的差异运行。路径走完三维，才把“谱线不对”变成有方向的研究，而不是停在异常命名上。

反向的 E—D—S 沉淀路径，则从新的分配计数出发，明确频率、温度、空腔和定标条件，计算预期曲线，再观察未被用来选取参数的结果。比如前面的温度加倍算例，在确定 h 与 k 等常量后便不再有任意调节余地。如果实际增长始终只有二倍，就应查明所处频段是否真的满足 x＝2，或模型及测量链条哪里有误。

三方程在这里不是三种诗意说法。S＝F（D，E）要求读数能够从具体操作与分配结构推出；D＝G（S，E）要求新操作针对连续与离散模型最容易分叉的条件；E＝H（S，D）要求计数规则在显露和运行之间接受修订，而非只由理论偏好宣布。E 既含允许模式和能量份额怎样互相限定，也含研究尺度下不可省去的平衡条件，不是把周围环境统统装进一个背景框。

两维张力推动第三维改变，在此有明确对象。谱形与已有探测操作彼此稳定，却与连续分配不相容，推动 E 改变；新结构与现有操作尚不能形成可比结果，推动 S 重新定标并规定可观察量；预测与实测相冲突，则推动 D 改进温度控制或选择更有区分力的波段。每次修改必须回检原矛盾，否则所谓修正可能只是把失败挪到另一个位置。

这套解释不声称普朗克写作时使用本书术语。它重构的是问题怎样具有发生条件，也不把一种物理成功当成 SDE 哲学的演绎证明。其价值取决于能否帮助我们区分计数改变、仪器改变和对象改变，避免把三者全说成同一种“观念革命”。

36.7 平均值不能替代一次事件

还可以用同一个三谐振子模型检查平均的含义。四份能量由三个位置分享，若十五种安排等权，则每个位置的平均占有数为 4/3。这个非整数平均不与每次占有整数相矛盾，因为它不是某一次安排的占有数，而是许多安排中同一位置的统计平均。

为免语言绕行，只需列出第一位置得到四份、三份、两份、一份、零份时分别有一、二、三、四、五种安排，加权总数为二十，除以十五就是三分之四。平均值不必属于单次允许结果的集合。这一点在量子史中会反复出现：若研究者只测平均能量，便不能根据一个连续变化的平均数否定离散的单次规则。

不过，等权本身是一项模型规定，需要统计条件支持。一个持续受外部驱动、来不及平衡的谐振子集合，不必遵循这里的分配概率。若把同一表达直接搬到任意闪光或激光过程，就越过了本章的适用边界。平衡黑体不是所有发光物的同义词，温度也不是任意光谱都能被赋予的唯一标签。

由此还可以提出一个新实验问题：两种模型给出相同均值时，能否比较涨落？同样的平均占有数，可能来自经常出现的中等交换，也可能来自大多数时候没有交换、偶尔出现较大交换。方差与时间关联能够提供均值之外的限制。这里仅提出问题，不预先搬用后来的完整光子计数理论；思想上的推进在于认识到，一条平均谱成功之后，事件层面的证据责任仍未结束。

因此，计数规则的检验应按层次推进：先检验均值，再检验不同条件间的同一参数，最后探索足以区分候选机制的事件统计。任何一层成功都值得保留，却没有理由替下一层预签通行证。

36.8 从份额到光的身份仍有一步

本章的转折不是连续曲线突然变成一排断点。黑体的宏观谱仍可以是一条光滑函数；离散规则进入的是允许的能量分配及其统计，而大量事件平均后可以产生连续的外观。因此，肉眼或仪器见到平滑曲线，不能立即判决底层过程必然连续；反过来，一次读数跳变也不能自动证明存在普适的能量量子。

前章的守恒由此取得新的任务。现在除了问总量是否守恒，还要问守恒的总量由怎样的允许份额组成，哪些交换能够发生。相同的守恒式可以容纳不同的状态结构。物理创造并不总是推翻一个旧方程，也可能保留总量约束而重建总量的组成方式，使旧方程在新的对象世界中继续工作。

黑体之所以值得作为思想史节点，正因为它迫使研究者把计数、熵、频率与仪器曲线放进共同限制。炉子的颜色没有因一个公式而被任意创造；新发生的是一套能够跨温度、跨频段重新组织辐射的规定。它既接受经验抵抗，也改变哪些问题可以被精确提出。

下一章要走的那一步并不自动包含在这里。若把能量份额的结构从谐振子分配推进到光的产生、传播与吸收，就必须面对更多证据，也要承担更多解释义务。一个平衡谱中成功的计数，能否成为关于光本身的假说？让这个问题保持开放，比把后来成熟的答案提前写回一九〇〇年，更能显示量子思想真正的创造难度。

本章参考资料

[36-1] Planck, M. Ueber das Gesetz der Energieverteilung im Normalspectrum. Annalen der Physik 4 (1901):553–563. 英译 On the Law of Distribution of Energy in the Normal Spectrum, https://strangepaths.com/files/planck1901.pdf。历史定位依据开篇及分配计数部分；本章小整数计数、极限比较、温度比值和 SDE 分析均为独立教学重构。

[36-2] Gearhart, C. A. Planck, the Quantum, and the Historians. Physics in Perspective 4 (2002):170–215. https://employees.csbsju.edu/cgearhart/planck/pqh.pdf。使用引言及“Arguments for Continuous Resonator Energies”部分对史学争论的梳理。

第三十七章 爱因斯坦光量子 光怎样在一次交换中取得身份

37.1 同一束光的两个问题

一束光经过透镜，在屏上形成明暗条纹；把屏换成适当的金属表面，又可以使电荷逸出。前一安排追问空间中的强度怎样分布，后一安排追问能量怎样进入物质。两者都在研究光，却不能因名称相同，就假定前一个问题中成功的所有构成条件能够不加修改地回答后一个问题。

前章的普朗克计数规定了平衡分配的一种结构。这一章进一步追问：能量的份额性是否只在谐振子的计算中成立，还是光的产生和吸收也应以不可任意细分的交换单位组织？爱因斯坦一九〇五年的光量子论文并不只讨论光电效应。它从辐射与物质描述方式之间的差别出发，利用熵的比较提出启发式假说，再讨论发光、光致电子发射和电离等问题。[37-1]

“启发式”并不是免于证据的同义词。它意味着作者提出一条尚未完成全部动力学解释的研究方向，并让这条方向接受具体结果限制。我们不应把后来的光子概念、场量子化和探测理论全部装进一九〇五年的术语，更不应把这篇论文写成爱因斯坦已经同时解决波粒关系的终章。

本章关注一种新的物理身份如何出现：频率不再只是波形每秒重复多少次，也进入一次能量交换的大小；强度不再独自决定单次能量，而与交换次数发生联系；材料不再只是被动承受光的表面，而通过逸出功与电子损失过程参与规定可见结果。这几项属性共同改变，才构成光量子问题的实质。

37.2 从体积变化得到的提示有多强

先用一个简化模型理解熵比较的思路。设 N 个相互独立的粒子均匀分布于体积 V_0，现问它们恰好全在其中一块体积 V 内的概率。单个粒子落入该区域的概率为 V/V_0，独立性使联合概率为 $(V/V_0)^N$。取对数后，相关熵差与 $Nk\ln(V/V_0)$ 成正比。这一结构把体积比的对数与可计数对象的数目联系起来。

若某种单色辐射在适用条件下的熵变化，也呈现一个能量除以 $h\nu$ 的系数乘 $k\ln(V/V_0)$，就会出现一个有力提示：它的体积依赖好像由 N＝能量总量/$h\nu$ 个独立份额构成。爱因斯坦的原文把这一比较限定在维恩辐射规律适用的区域，然后把它作为进一步研究光的产生与转化的理由。[37-1]

诺顿的科学史研究特别反对把这项提出过程讲成任意破坏旧规则，或仅由光电效应触发的灵感。他把论文的热力学论证放回中心。[37-3]这一研究提醒我们：后来最著名的验证场景，并不必然就是最初论证的全部来源；发现的理由与传播时使用的代表性例子，需要分别追踪。

这种论证的力量与限度应一起保留。相同的熵形式并不自动保证所有微观机制完全相同，正如相同压强规律不能证明两种气体的分子构造相同。它建立的是一项可继续检验的结构相似，尚未给出光量子怎样形成干涉、如何在传播中相互关联的完整答案。

为了看出限定的重要性，可以把单色辐射换成宽频混合光。若每个频率对应不同的 hν，单一总能量除以某个任意平均频率，不再自动给出唯一的对象数。还需要说明频率分布怎样计权，以及每一份能量属于哪个区间。一个在窄频、低密度条件下清楚的比较，不能不加条件地扩展为所有光场的普遍计数公式。

这里还藏着一个方法论转折。研究者没有先用显微镜找到光的小球，再根据小球性质写熵；他先从已经可约束的宏观关系中提出关于交换份额的新规定，然后寻找能够扩大或否定这种规定的实验。对象身份通过多种关系逐步取得，而不是由一幅想象图像一次性保证。

37.3 先区分电流和截止电压

考虑一束近似单色光照射金属，逸出的电子在电场作用下到达收集极。测得电流，说明每秒有一定净电荷通过；改变反向电压，可以阻止动能不足的电子到达。使光电流在理想条件下刚好消失的截止电压 V 停，与可逸出电子的最大动能满足 K 最大＝eV 停，其中 e 表示电子电荷的正大小。

电流与截止电压因而不是同一性质的两个读数。较大电流可能意味着更多电子到达，却不一定意味着每个电子都更快；较高截止电压限制的是能量上端，却不能单独告诉我们每秒有多少电子。若把二者混同，任何关于增强光照会发生什么的讨论都会失去明确对象。

在最简单的单量子吸收模型中，光的一份能量 hν 转移给电子，电子离开材料需要付出逸出功 Φ，剩余能量上限为 K 最大＝hν－Φ。因此 eV 停＝hν－Φ。这里的最大值十分关键。实际电子可因起始状态不同、在内部散射或表面条件不同而损失能量，不应期待所有逸出电子都具有完全相同的动能。

这条式子同时给出三项可区分预言。固定材料而改变频率，截止电压随频率线性变化；在理想单光子范围内，固定频率而增加强度，主要改变电子数而非能量上限；当 hν

低于逸出功时，简单模型不允许由一次吸收产生逸出电子。三者应联合检查，不能只挑最容易画成直线的一项宣布完成验证。

还要区分截止电压与外加加速电压。若收集极吸引电子，电子在飞行中获得额外动能；那不应算回光在表面交给电子的能量。仪器测的是经过电势差和探测过程之后的效果，理论需要把这些操作明确折算。所谓“直接测得光子能量”也往往包含这样的推断链。

37.4 一个可复算的光电算例

设想一种理想表面的逸出功为 2.0 电子伏特。这里的数值仅用于演示，不对应某一种真实样品的保证值。用 hc 约等于 1240 电子伏特纳米，则 400 纳米光的一份能量约为 3.10 电子伏特，因此最大电子动能为 1.10 电子伏特，理想截止电压为 1.10 伏。

若换成 500 纳米光，单份能量约为 2.48 电子伏特，最大动能降为 0.48 电子伏特。两种光的功率可以完全相同，但能量上限不同。这不是能量守恒失效，因为相同功率意味着每秒送入的总能量相同，不意味着每次送入的份额相同。较长波长的份额较小，相同总功率对应的份额数反而较多。

临界波长为 hc/Φ，约等于 620 纳米。用 700 纳米光时，一份能量约 1.77 电子伏特，低于上述门槛。不能把公式结果负 0.23 电子伏特理解成电子以负动能飞出；正确解释是在这些假设下，单次吸收没有足够能量产生这种逸出事件。公式右边为负，标记的是过程不允许，不是一个新的可测负动能。

再固定 400 纳米，将入射功率取为一微瓦，理想份额数率约为一微瓦除以 4.97×10^{-19} 焦耳，即每秒 2.01×10^{12} 份。若每一份都导致一个被完整收集的电子，电流约为 0.322 微安。若总有效效率只有百分之十，则约为 0.0322 微安。这个效率同时概括吸收、逸出和收集，在真实研究中通常需要进一步拆分。

现在把功率增加十倍而保持单光子、低密度和表面稳定等条件，模型预言电流增加十倍，截止电压仍约为 1.10 伏。如果实测截止电压显著变化，不应先宣布假说失败，也不应自动归于误差；应检查空间电荷、样品升温、光谱变化和多光子过程是否已经进入。明确适用条件，使反例能够指向具体前提。

如果把光束面积减半而保持总功率，单位面积的份额率加倍，总份额率不变。理想整体电流可以保持，局部表面负荷却改变。这个区别提醒我们，“强度增加”究竟指总功率还是单位面积功率必须说清楚。一个没有注明面积和收集范围的比较，可能把空间分布变化误写成单次能量变化。

37.5 材料变了，哪一项应该保持

换另一种表面，其逸出功假设为 2.5 电子伏特。在 400 纳米照射下，最大动能变成 0.60 电子伏特，截止电压比第一种表面低 0.50 伏。但把截止电压对频率作图，在理想模型中两条直线斜率同为 h/e，只是截距不同。跨材料保持的斜率与材料相关的截距，共同构成检验的关键。

若所有差别都可归入任意变化的逸出功，模型是否会变得无法反驳？答案取决于逸出功是否在一组实验中保持约束。对同一稳定表面用一个固定 Φ 解释多个频率，比每个频率单独设置一个 Φ 严格得多。后者可以把任意曲线强行改写成直线关系，因而失去检验意义。

实际表面还会受到吸附、氧化、晶面和接触电势的影响。此处不把它们展开成材料学大全，只强调实验逻辑：两块名义上相同的金属，不一定实现相同的表面条件；一个测量电压的常数偏移，也可能改变截距而不改变斜率。应先检查哪类误差改变哪类可观测量，再决定它是否能够模仿理论预言。

密立根一九一六年的研究以光电测量确定普朗克常量为题，显示检验已推进到精密的频率—电压关系。[37-2]它在本章承担的历史作用不是给某个人物编排“先顽固反对、后彻底投降”的心理戏剧，而是说明一条新假说的力量必须进入精密操作。测量关系获得支持，与其全部本体解释立即被所有人接受，是不同的历史判断。

独立测定 h 还使两类原先不同的实验发生连接：黑体分配中的常量与光电关系的斜率使用同一数值尺度。如果每种现象都需要专属的“普朗克常量”，理论统一就会丧失实质。跨实验的同一参数不是装饰，它把不同操作体系绑在一起，使一边的结果能够限制另一边。

37.6 “光越弱越要等”为什么不是一个充分理论

常见的连续积累图像认为，光的能量均匀流入一个微小电子，只有积累到门槛后才能逸出。若把电子截取光的面积固定为 a，入射强度为 I，积累所需时间就大致为 Φ/(Ia)。这样减弱光照会延长等待，增加频率本身却未必改变每次获得能量的方式。这个粗略模型与份额交换给出不同的因果结构。

但不能把它当作所有经典或半经典理论的唯一代表。电子与材料可具有复杂响应，量子化的物质与连续经典场结合，也能解释许多平均光电关系。因此，单凭一个漂亮的截止电压直线，不能逻辑上排除所有连续场描述。要裁决光场本身的量子性质，还需要更精细的相关和统计证据，后续章节才有条件展开。

同样，份额假说不保证任意弱光一开灯就必然在下一瞬间看到电子。若份额到达率很小，等待首个事件的时间可以很长；这与每个电子需要逐渐蓄足许多小份能量不是一回事。前者是事件出现的统计等待，后者是同一个响应单元的连续积累。把这两种时间混同，会使“无延迟”的口号产生错误预言。

用一个自拟统计模型说明：若有效事件平均每秒发生一次，假设独立泊松到达，则一秒内没有事件的概率约为 e^{-1}，即百分之三十七；十秒内没有事件的概率约为 e^{-10}。这里泊松过程只是演示随机等待与能量积累的区别，并不声称所有光源必然具有这种统计，更不以它替代量子光学。

在单份能量足够的条件下，弱光使事件更少，并不使一个已发生事件的能量上限按强度同比缩小。若长时间没有读数，需要问入射份额率、吸收效率、探测阈值与暗噪声，而不是立即判断光子假说被否定。相反，如果研究者不断把所有不符结果归于看不见的效率变化，也会使理论免于检验。效率应通过独立操作约束。

37.7 反例应当落在它真正挑战的地方

强光下，多个光量子可以共同参与一次过程；材料本身也可以因升温而提供能量。因此，“低于临界频率，无论多强都不可能逸出”只在规定的单光子、无其他能量来源模型中成立，不能作为不带条件的自然禁令。若观测到阈值以下的发射，首先要检查它是否随强度呈不同次数依赖，以及是否存在热效应。

例如一种假设性的双光子过程需要 $2h\nu$ 超过 Φ，在足够弱且其他条件固定的近似下，发生率可能随强度平方变化。若强度加倍后事件率约增四倍，它与单光子率的线性响应形成区别。不过仅看到四倍仍不是唯一机制的身份证，因为探测器非线性也可能模仿。必须交叉检查频率、脉冲形状与探测线性范围。

另一种反例来自能量上限以下的宽分布。有些电子只带走一部分入射能量，并不违反最大值关系；但若在排除额外能量来源后可靠发现显著高于 $h\nu-\Phi$ 的单事件，就会触及更直接的模型问题。哪一部分是预测的上界，哪一部分是分布形状的留白，应在检验之前写清楚。

相同原则也适用于波的证据。干涉条纹不能被一句“光其实是粒子”取消。新的份额结构如果要扩展到传播，必须解释为何相位关系依然可检验。本章不以“光一会儿是波一会儿是粒子”填补机制空白，因为这句话若没有说明何种条件保留什么关系，便只是把两套成功语言并排摆放。

37.8 交换的共同规定怎样进入 SDE

在光电实验中，显露不只是眼睛看到的一点亮光，而包括经过仪器规定的电流、截止曲线和统计记录；差异运行包括改变频率、功率、面积、表面与电压；特征纠缠包括 hν、逸出功、电子损失和收集条件如何共同限制一次记录。这不是把视觉、电学和材料分别配给三个字母，每一种感知和操作过程仍有自己的完整发生结构。

可以沿 D—S—E 原型路径，先建立一组固定表面、扫描频率的操作，得到初步截止曲线，再检验共同斜率与不同截距能否成立。如果能，就把强度扫描加入同一实验，检查能量上限与计数率是否具有不同响应。E 的形成来自一组被交叉约束的差别，不能从单个读数直接跳到“光子已经被看见”。

再沿 E—D—S 沉淀路径，用确定的份额模型预测另一波长或另一表面的结果，预先给出截止电压和计数改变方向，最后回到观测。前面的 400 纳米与 500 纳米算例就是这种路径的最小版本。若只报告模型解释了已参与拟合的点，没有新的预测回检，所谓沉淀便仍然停留在重述。

S＝F（D，E）在此要求结果从照射、表面和收集结构中生成；D＝G（S，E）要求实验优先选择能区分“更多电子”与“更快电子”的条件；E＝H（S，D）要求理论解释接受共同斜率、阈值和损失过程的联合限制。三方程不提供 h 的数值，而规定论证不能在哪一环悄悄省略条件。

三种张力也各有落点。稳定的最大能量与连续积累操作预期不合，推动交换结构 E 改变；新结构 E 要求比较表面逸出能量，而现有电压扫描操作 D 尚未扣除接触电势，二者张力推动显露 S 重新定标；曲线与理论不能协调，则推动 D 改变频率区间、表面制备或强度范围。任何修改都要返回原来两维，检查是否真正解决冲突，而不是靠更换术语避开。

37.9 光量子不是缩小版弹丸

到这里，光量子的身份已与“一个很小的发光球”明显不同。它的可检验规定在于频率与能量交换的联系、计数与总功率的联系，以及对材料电子响应的限制。半径、硬壳和经典轨迹都不是从这些关系自动推出的属性。保留物理对象的实在约束，不要求把每个对象强行画成日常触觉中的小物件。

工具主义可以强调这条公式只是有效算法；朴素粒子实在论可以强调光终于恢复了由小颗粒组成的真相。本书需要同时面对两种提醒：没有预测约束的实体图像过于自

由，没有对象重组意识的算法叙述则可能忽略频率、强度和能量身份已被改写。发生学的任务是把这次改写落实到可操作的关系，而不是凭哲学口号宣布胜负。

新的世界在这里并不等于观察者用意识制造照射能量。它意味着一套新的共同规定进入仪器、实验和计算，使原先不能协调的结果成为可比较对象。自然通过阈值、斜率与噪声持续限制这种创造；人们则通过选择问题与重组模型改变何种物能够在研究中成立。两者缺一，都会失去这段历史的张力。

光量子的启发也不会独自解决原子内部怎样安排的问题。光从材料带走电子，使电子身份更加迫切；散射和阴极射线又将推动原子结构重建。接下来的章节要把目光转向这些物质对象，再看能量份额怎样进入原子稳定性与谱线。只有经过这些不同入口，量子论才会逐步获得比一条光电方程更广的组织力量。

本章参考资料

[37-1] Einstein, A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. Annalen der Physik 17 (1905):132–148. 英译 Concerning an Heuristic Point of View Toward the Emission and Transformation of Light，American Journal of Physics 33 (1965):367–374。https://icg.isy.liu.se/en/phd-courses/quantum_foundations/references/Einstein1905b.pdf。历史定位取引言、熵比较及第八节；算例使用现代记号独立展开。

[37-2] Millikan, R. A. A Direct Photoelectric Determination of Planck's “h”. Physical Review 7 (1916):355–388. DOI:10.1103/PhysRev.7.355；原件：https://users.physics.ox.ac.uk/~palmerc/Datafiles/Millikan1916b.pdf。本章仅用其测量目标与光电关系检验的历史定位。

[37-3] Norton, J. D. How Einstein Did Not Discover. Physics in Perspective 18 (2016). https://sites.pitt.edu/~jdnorton/papers/Einstein_Discover_final.pdf。使用“How Einstein Discovered the Light Quantum”部分，讨论热力学论证与光电效应在发现叙述中的不同地位。

第三十八章 汤姆孙与卢瑟福 原子怎样从不可分之物成为可追问的结构

38.1 比看见更困难的是认定同一个东西

假如玻璃管中出现一道亮痕，磁铁靠近时亮痕移动，我们已经看见什么？可以确定装置出现了受磁场影响的效应，却还不能单凭这一事实断定，亮痕就是某种小粒子的完整轨迹。发光介质、带电运动、气体导电和玻璃响应都可能参与结果。要使“电子”成为一个能够跨装置保持身份的物理对象，必须让多种效应受到同一套数量关系限制。

原子结构的改变也不是把想象中的实心球敲开，再直接看见里面的核。阴极射线实验从电荷与偏转关系出发，散射实验从出射角分布出发，两者都通过可控制作用追问无法直接看见的结构。这里的推断并不比普通观看天然低一等；可靠性来自条件明确、不同证据互相限制，以及竞争解释能够被安排在同一张比较桌上。

汤姆孙一八九七年的论文把阴极射线性质作为问题，研究电荷携带、电场偏转、磁场偏转及质量电荷比。[38-1]卢瑟福一九一一年的论文则把大角度散射与集中电荷的原子构造联系起来，并明确暂缓处理原子稳定性。[38-2]这两项工作回答的不是同一道题，也没有在一次实验中完成后来所有关于电子与原子核的知识。

本章先分析如何由偏转得到一个比值，再分析如何由散射分布约束内部结构。两次分析之间的共同线索是：物的身份由一组跨操作保持的关系取得。但不能因此说两种实验只是同一个抽象模板。电荷质量比和散射截面各有独特的推导、误差与不可识别性，正是这些差别使发生学讨论有具体内容。

38.2 电与磁给出的不是同一种偏转

用现代记号重构一个理想模型。设带负电粒子以速度 v 进入匀强电场，电场垂直于原运动方向，粒子电荷大小为 e，质量为 m。电力大小为 eE 电，加速度大小为 eE 电/m。若场区长 l，纵向速度近似不变，停留时间为 l/v，场区出口的横向偏移为 eE 电 $l^2/(2mv^2)$。

电子离开场区后仍带着横向速度，所以若屏幕位于出口之后距离 L 处，还会继续偏移 eE 电 $lL/(mv^2)$。屏上总偏移为 eE 电 $l\,(L+l/2)\,/(mv^2)$。这个简单补项很重要：屏的位置改变，并不一定表示粒子受到的力改变；它也可能只是同一偏转角经过更长飞行距离放大。测量几何属于推断结构的一部分。

若换成垂直于速度的匀强磁场，磁力大小为 evB，且在此理想情形下不做功。轨迹半径 r 满足 $mv^2/r = evB$，故 $e/m = v/(Br)$。电偏转依赖速度平方的倒数，磁曲率依赖速度的一次倒数。利用这种不同依赖，便有可能把未知速度与电荷质量比从多组结果中分别限制出来。

一个便于教学的组合方式是交叉电场与磁场，让两种力方向相反。当束流不偏转时，$E_{电} = vB$，因此 $v = E_{电}/B$。再用单独磁场测 r，得到 $e/m = E_{电}/(B^2 r)$。这是现代简化的速度选择与偏转方案，不把它说成汤姆孙原文全部实验装置的逐件复原。它的作用是让读者自己看清多种作用为什么能够解除未知量之间的纠缠。

取自拟数据 $E_{电} = 2.0 \times 10^{4}$ 伏每米、$B = 1.0 \times 10^{-3}$ 特斯拉，则 $v = 2.0 \times 10^{7}$ 米每秒。若相应磁曲率半径约 0.114 米，则 e/m 约为 1.75×10^{11} 库仑每千克。这里速度约为光速的十五分之一，非相对论近似足以展示方法；精密测量应评估相对论修正、边缘场和速度分布，而不能把演示值当作原始测量表。

38.3 比值相同还不是所有属性都已确定

测出 e/m 不等于分别测出 e 与 m。假设电荷和质量同时乘二，比值不变，在上述理想偏转中可能得到同样结果。要分别确定它们，需要额外信息或独立实验。这个简单的不可识别性限制，能够防止一种常见历史缩写：把后来已知的电子质量和电荷，全部归入一八九七年一项测量。

同样，比值很大可能来自质量小，也可能来自电荷大，或者两者共同改变。若有其他证据限制电荷大小，关于质量很小的推断才更有力量。一个模型中可由数据直接识别的参数组合，与借助附加假设得到的单个参数，应在叙述中分开。科学的可靠性并不要求每一步都“直接”，而要求推断链能够被公开检查。

跨材料比较提供另一层约束。若来自不同气体或电极的射线在适当条件下具有相同偏转特征，就会削弱“它们只是每种材料特有的普通离子”的解释，并支持某种普遍成分。这里的普遍性不是凭名称宣布；它由改变来源后仍保持的数量联系逐渐成立。

不过，比较不同气体时不能只改变标签而忽略放电条件。气压、电压与速度分布可能共同改变，若两条亮痕不同，未必说明粒子种类不同；若相同，也要确认探测器分辨率足以发现候选差异。对照实验需要保持哪些条件、允许哪些条件变化，是“相同粒子”判断的组成部分。

汤姆孙原文还讨论高真空条件下电偏转的出现，以及残余气体导电可能使预期场效应失真的问题。[38-1]这一点具有方法上的重量：没有观察到偏转，可能是对象不带电，

也可能是预期电场没有真正作用于它。负结果不能脱离装置运行条件独自裁决对象性质。

若在真空改善后出现偏转，也不能立刻把每一种变化都归于“终于看见电子”。真空变化还会影响束流能量与碰撞损失，必须用多种测量确认解释一致。对象不是在一次开关动作后突然获得全部属性，而是在电荷、偏转、速度、来源等关系中逐步稳定。

38.4 从电子成分到原子内部还有距离

接受原子含有带负电成分后，电中性仍要求有相应正电荷。但正电荷怎样分布，是新的问题。把它均匀铺在一个原子大小的球内，是一种具体模型；把它集中在很小区域，则是另一种。二者都可以使总电荷为零，所以电中性本身不能选择内部结构。

假设均匀带正电球半径为 R，总电荷为 Q。球内距离中心 r 处的电场大小为 $Qr/(4\pi\varepsilon_0 R^3)$，随 r 线性增加，在中心为零；球外则与集中点电荷相同，按距离平方反比下降。因此，只测足够远处的场，难以区分均匀球与集中核。要探问内部，入射探针必须有机会进入能够分辨两种场结构的距离范围。

这个模型比较显示，空间分辨并非只由显微镜放大倍数决定。探针能量、作用方式与可达到距离也会规定它对内部结构的敏感性。一个只能经过外部区域的粒子，即使出射角测得十分精确，也可能对中心分布缺乏区分力。

同时，含电子的均匀正电模型不能只用一个静电球图画概括全部研究。电子怎样运动、如何维持平衡、怎样解释化学与辐射，都需要进一步安排。卢瑟福的集中核方案也面临相同压力。不能只列旧模型缺陷而把新模型的未决问题藏起来；思想史中的竞争往往是不同解释各自擅长一部分证据，而不是一方全无理由地抵抗明显真相。

科学史家克拉格把核原子的提出放在卢瑟福长期的放射性研究与研究团队工作中讨论。[38-3]这提示我们，不应让一个后来广为传播的反弹比喻吞掉实验来源、闪烁计数和模型分析。惊讶可以推动研究，只有分布与推导才能决定惊讶最后支持什么。

38.5 为什么尾部比平均角更有力量

设一束 α 粒子穿过薄箔，绝大多数只略微偏离原方向，少数却发生大角度偏转。若只报告平均偏转很小，关键证据可能被多数事件淹没。统计分布的尾部有时直接区分机制：许多小作用的累计，与一次强作用，可能产生相似的中心宽度，却产生极不同的大角度概率。

一个直观但有边界的例子是随机小偏转。若许多次独立横向偏转各自平均为零、方差为 σ^2，N 次累计的方差约为 $N\sigma^2$，典型宽度按 $\sqrt{N}$ 增长。但典型宽度并不自动决定远端尾部，独立性、单次分布与大偏转近似都需检查。不能只说“次数很多，所以任何大角度都可能”便完成解释。

反过来，发现一次大角度事件也不是立即确立原子核。它可能来自杂质、边缘碰撞、背景计数或装置散射。必须考察事件随薄箔厚度、入射能量、材料和观察角度如何变化。若同一集中电荷模型同时给出这些变化的比例，证据强度才明显增加。

卢瑟福论文的核心区分正是单次大偏转与多次小偏转的组合；它利用集中电场建立可比较的角分布，并未声称已经解决完整原子动力学。[38-2]本章下面以现代符号把这种推理展开，重点放在可复算关系，避免把某一历史数值当作所有金箔实验通用的反射比例。

38.6 把入射位置转成散射角

设质量较轻的入射粒子以动能 K 接近足够重、近似静止的散射中心，二者排斥势为 κ/r，其中 $\kappa = zZe^2/(4\pi\varepsilon_0)$。z 与 Z 分别是入射粒子和中心的电荷数。冲击参数 b 表示远处入射直线与中心的垂直距离，不是粒子真实轨迹的最近距离。

在经典库仑散射模型中，能量与角动量守恒给出 $b = [\kappa/(2K)]\cot(\theta/2)$，θ 为出射方向相对于入射方向的偏转角。b 很大时偏转小；b 趋近零时粒子接近迎头碰撞，偏转趋近一百八十度。这个连续关系不需要把中心当成有硬表面的碰撞球，强排斥场本身就能产生反向出射。

由入射面积环 $2\pi b\,db$ 与出射立体角环 $2\pi\sin\theta\,d\theta$ 相对应，微分截面为 $d\sigma/d\Omega = [\kappa/(4K)]^2/\sin^4(\theta/2)$。推导中必须对 $db/d\theta$ 取绝对值，因为 b 随 θ 增大而减小，截面却不能为负。这一小细节显示，角分布是几何映射与动力学关系共同给出的结果。

例如比较六十度和一百二十度方向，每单位立体角的散射强度比为 $[\sin60^\circ/\sin30^\circ]^4 = 9$。若直接比较两个相同角宽、但不相同立体角的环形探测区，则还要计入 $\sin\theta$ 因子；若是固定面积、不同距离的探测屏，也要折算立体角。裸计数比不是自动等于截面比。

在其他条件固定时，截面随 K 的平方倒数变化。把入射动能加倍，模型预言同角度、同立体角的计数贡献降到四分之一。电荷数 Z 加倍则增加四倍，但真实材料更换还会改变原子数密度、电子屏蔽和能量损失；这些因素需独立校正，才能把材料比较变成集中电荷的检验。

对于薄靶，若单位面积原子数为 n 面、入射粒子数为 N 入，探测效率为 η，小立体角为 ΔΩ，则某角度计数近似为 N 入 n 面（dσ/dΩ）ΔΩη。这个表达把核模型与记录系统接起来。只要某一项随角度未知变化，就可能模仿散射规律；因此角扫描应同时检查几何接受度和效率。

38.7 最接近中心不等于摸到核表面

迎头碰撞时，入射动能在最近点转成排斥势能，故最近距离 r 近＝κ/K。以现代金核电荷数 Z＝79、α 粒子 z＝2、动能 5 兆电子伏特作示范，取 $e^2/(4\pi\varepsilon_0)$约 1.44 兆电子伏特飞米，得到 κ 约 227.5 兆电子伏特飞米，最近距离约 45.5 飞米。这个演示使用今天的电荷数，不冒充卢瑟福一九一一年的原始参数。

四十五飞米远小于通常原子尺寸约十万飞米，却不能据此说核半径就是四十五飞米。粒子可能在核表面之外很远处已被电场折返；若一路符合点电荷散射，只能说明在探测到的尺度内，该分布尚未被分辨。测得的是可到达距离及对结构的约束，不是拿尺子直接量核的边缘。

非迎头情况下还保留角向运动。最近距离为 κ/(2K)加上［（κ/(2K)）²＋b²］的平方根，大于迎头情况的相应极限。它再次说明冲击参数、最近距离和核半径是三个不同量。把一张轨迹图上的任意短线都叫“原子核大小”，会抹去实验究竟测到什么。

使用更高能量能够接近更小距离，却也可能让原先可忽略的相互作用进入。若出现偏离，可能意味着有限核尺寸，也可能涉及核力、非弹性过程或靶反冲，不能只选自己期待的一项。理论外推会改变探针的敏感范围，同时改变解释所需条件，这种双重变化是散射方法必须面对的事实。

重靶近似同样有限。若散射中心质量不够大，它会带走动量和能量，实验室角与质心系角不再简单相同。中心能动与中心电荷分布是两种不同修正，不能把所有偏差一概归为“核并非点状”。此处保留这些限制，是为了让漂亮的 $\sin^{-4}$ 关系不被误用成任何散射的万能指纹。

还可以用厚度变化区分“没有事件”与“事件太少”。假设某角区在一块薄箔上预计十万次入射产生十个有效计数，平均率为万分之一。若只发射一百个粒子，没有计数十分正常，不能据此否定该角区散射。若发射一千万个仍然没有计数，则在效率可靠且条件相同的前提下，就会对原预测形成强压力。实验规模属于判断负结果的必要条件。

在单次散射占主导且靶很薄时，厚度加倍应使相应计数近似加倍。如果增厚后计数不再线性增长，可能因多次散射、吸收与能量降低进入，也可能因探测器计数饱和。此

时最有信息的操作通常是重新减薄或改变入射率，观察是否恢复线性。一个能够通过反向操作恢复的效应，比只凭一次异常猜测更容易定位。

统计误差也不应与系统误差混为一谈。十个计数的随机相对起伏很大，积累更多事件能够改善；角度刻度整体偏移却不会靠长时间计数自行消失。反复测同一设置可以获得很精密但仍偏离真实角度的分布。模型检验需要两种工作同时进行：扩大样本以稳定随机量，改变校准方式以暴露共同偏差。散射图像的说服力正来自这种多层约束，而非少数闪光的戏剧性。

38.8 稳定性留下的不是小漏洞

集中核方案对大角度散射有力量，却不能仅靠它解释电子为什么长期留在原子中。若电子按经典电磁学绕核加速运动，它会辐射能量；能量不断损失将改变轨道。若把电子静止放在核外，仅有库仑吸引又无法提供稳定的静态平衡。一个能够解释探针轨迹的模型，仍可能没有完成被探测对象的内部动力学。

这不是要求每个新理论诞生当天包办一切，而是区分成功的范围。核的集中电荷、电子的存在与稳定原子的完整结构，可以由不同证据逐步建立。保留尚未解释的部分，允许下一项理论创造有真实任务；若提前把原子画成稳定太阳系，就会把玻尔随后要处理的矛盾从历史中删掉。

散射数据对中心电荷符号的限制也应谨慎。某些角分布主要依赖耦合强度的平方，单靠这种分布不能任意宣称已辨认所有符号信息；电子与物质电中性等独立知识共同参与判断。卢瑟福原文明确指出主要散射推论对中心电荷正负具有一定独立性。[38-2] 把“正核”作为整个证据网络中的结论，比让一条曲线独自承受全部证明责任更准确。

在这里，原子不是从“没有内部结构”直接跃迁成现代教科书插图。它先成为具有可分离电荷成分的对象，再成为具有高度集中散射中心的对象，而稳定状态和辐射规律仍待重建。不同属性可以跨理论保留，也可以在新的关系中改变意义；全有或全无的发现叙述不能反映这种分层发生。

38.9 让对象经受多种操作的共同限制

阴极射线的一条有效路径是 S—E—D 进入：从偏转与电荷记录的联系出发，检查负电荷、速度、质量和气体导电怎样互相限制，再改变真空、场向与收集位置。若改善真空后电偏转出现，第三维的变化必须回检磁偏转与收集电荷是否仍指向相容对象。不能只保留最喜欢的一项读数。

散射中则可走 D—S—E 原型路径：改变箔厚与观察角，取得角分布及厚度依赖，再比较集中核单次散射和复合小散射的结构。结构形成后用新的能量范围检验，才完成由原型到可迁移规律的推进。尾部计数虽少，却可能比中心区域更加决定竞争解释的差别。

S＝F（D，E）要求角分布由入射制备、靶结构和计数条件共同生成；D＝G（S，E）要求下一轮操作选择能放大模型差异的角度与能量；E＝H（S，D）要求“电子”“集中核”这些对象规定在实验联系中接受重建。三方程并不把核变成心理产物，也不把所有实验条件都泛称为背景。

两维张力推动第三维的回检在这里可以具体执行。大角尾部与可重复的薄靶操作冲突于弥散场模型，推动 E 改写；相同角宽采集的操作 D 与按单位立体角定义的截面结构 E 不相容，推动 S 从裸计数重组为经立体角和效率校正的可比较量；同一结构与不同材料数据不合，则推动 D 测定面密度、调整能量并检查多次散射。修改只有经原证据重新检验才有资格保留。

一种实在论会说核早已存在，实验只是把它找出；发生学并不靠否认靶箔对探针的抵抗来回应。它追问“核”的哪些规定由何种作用稳定建立，哪些性质仍不能从当前证据推出。集中电荷的实在约束与对象身份的历史形成可以同时被讨论，不需要把二者强迫变成互相排斥的口号。

下一章从这项未完成处继续：若原子核与电子的关系已经受到散射和电学约束，怎样使原子既稳定存在又发出离散谱线？仅在太阳系图像中缩小长度，无法解决这个问题。必须改变允许状态与辐射之间的关系，物质结构与光量子两条路线才会真正交会。

本章参考资料

[38-1] Thomson, J. J. Cathode Rays. Philosophical Magazine 44 (1897):293–316. DOI:10.1080/14786449708621070. https://staff.emu.edu.tr/huriyegursel/en/Documents/J.%20J.%20Thomson%201897.pdf。使用电荷收集、电偏转与真空条件部分；教学公式采用现代理想装置重构。

[38-2] Rutherford, E. The Scattering of α and β Particles by Matter and the Structure of the Atom. Philosophical Magazine 21 (1911):669–688. DOI:10.1080/14786440508637080. https://uni-tuebingen.de/fileadmin/Uni_Tuebingen/Fakultaeten/MathePhysik/Institute/IAP/Forschung/MOettel/Geburt_QM/rutherford_PhilMag_21_669_1911.pdf。使用开篇、集中电荷假设、单次散射及稳定性限定。

[38-3] Kragh, H. Rutherford, Radioactivity, and the Atomic Nucleus. 2012. https://arxiv.org/abs/1202.0954，全文 https://arxiv.org/pdf/1202.0954。使用其放射性研究与核原子提出的历史联系；散射数字和参数识别讨论为本章独立分析。

第三十九章 玻尔原子
稳定状态与谱线怎样共同规定一种物

39.1 原子不能靠一幅太阳系图画稳定下来

把电子画在圆圈上，把正电核画在中央，图像似乎已经给出一个原子。但它还没有回答为什么圆圈具有这个大小，为什么电子不沿其他半径运动，为什么原子能够长久存在，又为什么同一种气体发出具有特定频率的光。图中的圆圈容易把这些问题遮住，使一个有待解释的状态被误认为已经由图画保证。

前章的散射支持集中电荷结构，却没有完成电子系统的动力学。若按照经典电磁学，电子在圆周上持续加速就会辐射；能量减少会改变轨道，原子难以保持稳定。若停止电子的运动，又没有其他机制平衡吸引。这个困境涉及结构与运动之间的不相容，不只是把某个测量常数调准即可解决。

玻尔一九一三年论文以核原子为出发点，提出经典理论不足以处理其稳定性，并把定态和辐射的量子条件联系起来。[39-1]本章采用现代简化的单电子圆轨道推导来理解其组织力量，不把教材中熟悉的演算顺序当作原论文全部论证的逐句重演，也不把随后发展的波函数和算符理论提前放入他的模型。

克拉格研究早期接受史时提醒读者，玻尔的三篇论文涉及的范围远不止后来最常见的氢原子图像，还包含分子与物质构造的广泛设想；当时研究者注意的内容并不完全相同。[39-2]因而，本章选择单电子模型，是为抓住最可复算的一条线索，不意味着历史上只有这一条线索，也不意味着它一公布就获得整齐一致的赞同。

39.2 给经典圆轨道增加一道选择规则

先设核无限重，电荷为 $+Ze$，电子质量为 m，电荷为 $-e$。把 κ 写为 $Ze^2/(4\pi\varepsilon_0)$。圆轨道的向心条件为 $mv^2/r=\kappa/r^2$，因此动能为 $\kappa/(2r)$，势能为 $-\kappa/r$，总能量为 $-\kappa/(2r)$。负号意味着以电子和核相距无限远且静止为零点时，束缚系统能量低于分离状态，不表示它有一种神秘的“负物质”。

到这一步仍是经典模型，允许半径连续取值。若再施加角动量条件 $mvr=n\hbar$，其中 $\hbar=h/(2\pi)$、n 为正整数，就得到 $r_n=n^2\hbar^2/(m\kappa)$，$v_n=\kappa/(n\hbar)$，以及 $U_n=-m\kappa^2/(2n^2\hbar^2)$。这里以 U 表示原子能量，避免与本书特征纠缠 E 混同。

推导非常直接：从 $mvr=n\hbar$ 得 $v=n\hbar/(mr)$，代入向心条件，便确定 r；再代回总能量式得到 U。真正的新规定不在代数步骤，而在只允许整数标记的状态。牛顿方程本身没有禁止半径介于 r_1 与 r_2 之间的轨道，离散选择不是从向心力式偷偷长出来的结论。

因此不能说玻尔仅仅发现电子在经典轨道上“恰好喜欢整数”。整数规则改变了什么状态被允许存在；定态不辐射又改变了运动与辐射之间的经典联系。若继续把允许状态里的电子当作普通天线并计算连续能量损失，就会把已被模型撤回的规则重新放回，制造一套自相矛盾的混合推导。

另一方面，模型保留经典力学来计算定态半径和能量。这种保留与撤回并存，是旧量子论的张力所在。它可以产生惊人的数量关系，却还没有一种统一动力学说明为什么某些环节按经典规律处理、某些环节必须使用量子条件。承认这种不完整，并不抹去模型的成绩，反而指出成绩为什么需要下一轮创造。

39.3 尺寸不是任意画出来的圆圈

对氢的固定核近似，第一轨道半径约为 0.0529 纳米，第一能级约为 -13.6 电子伏特。第二轨道半径是第一轨道的四倍，第二能级为 -3.40 电子伏特；第三轨道半径是九倍，能级约为 -1.51 电子伏特。半径扩大并不表示能量变得更负，相反，能量向零接近，电子更容易被移到分离状态。

这个次序可以用电离检验。若电子原在最低态，需要约 13.6 电子伏特把它带到无限远静止的极限；若原在第二态，只需约 3.40 电子伏特。在实际实验中，外场、碰撞与有限线宽会影响过程，但模型已经给出状态依赖的能量门槛。原子不再只有一个与过程无关的“坚固程度”，而具有被状态规定的响应。

对只有一个电子的 He^{+}，$Z=2$，固定核模型给出半径缩为相应氢轨道的一半，能量绝对值增为四倍。最低态约为 -54.4 电子伏特。这个 Z^2 关系只适用于这里的单电子库仑问题；不能把中性氦的两个电子简单都塞入同一公式，再宣布解释已经完成，因为电子相互作用改变了共同条件。

原子尺寸由 m、e、$\hbar$ 和核电荷共同给出，意味着不能随意把圆圈缩放而保持全部性质。若把半径改为原来一半，却仍维持相同整数和相同常量，向心条件与角动量条件便不能同时满足。图画之所以取得物理内容，正在于它被这种数量联锁约束。

但半径的模型数值也不应未经说明地等同于现代意义的电子云半径、碰撞半径或化学键长度。不同概念通过不同操作定义，可能在某些尺度同量级，却不是一把尺上同一

刻度的不同名字。本章讨论的 r_n 属于圆轨道模型，在后面成熟量子理论中，其几何解释将重新安排。

39.4 光的频率不等于绕圈的频率

定态的能量已经规定，辐射还需要另一条关系。若原子从较高能量 U 上转至较低能量 U 下，发出光的频率满足 hν＝U 上－U 下。吸收则要求光向原子提供相应能量差。这里出现的是两个状态之间的关系，不能仅由其中一条轨道的绕行频率决定。

以氢从 n＝3 到 n＝2 为例，发射能量约为 13.6（1/4－1/9）电子伏特，即 1.889 电子伏特，对应波长约为 656 纳米。这条可见红线的数值来自状态差，不是把电子沿第三轨道转一圈的周期倒数直接当成发光频率。若忽略两者的区别，就会失去玻尔理论最关键的改写。

再从 n＝4 降到 n＝2，能量为 13.6（1/4－1/16），约 2.55 电子伏特，对应波长约 486 纳米。改变上态而保持下态，便形成一个系列。不同谱线由同一组能级差联系，不需要为每条线分别设一个独立振子频率。允许状态因此获得一种跨谱线的共同身份。

把一般公式写成 $1/\lambda＝RZ^2（1/n_{下}^2－1/n_{上}^2）$，R 在固定核近似下由 m、e、h、c 等常量组成。这不是只有外形像实验规律才算成功，系数还必须在独立测量的常量范围内符合。若任意选一个 R 拟合所有氢线，却不问它与其他量的关系，就减少了模型的解释任务。

这里需要慎用“跃迁”图像。一条从外圆画到内圆的箭头标记两个状态的变化，并不等于模型已经给出电子沿箭头飞行的连续时程。若要求它在中途每个半径都有普通圆轨道状态，就与离散允许条件冲突。箭头表达关系与能量收支，未提供完整过程机制；这项缺口将成为后来理论的动力，而不应被插图悄悄填平。

39.5 串联路径检验能级是否共同成立

假设氢从第三态先降第二态，再降第一态，两次发光能量分别约为 1.889 与 10.20 电子伏特，总计 12.089 电子伏特。若直接从第三态降第一态，按能量差计算也是约 12.089 电子伏特。这种加法来自同一组状态能量的相减，独立于我们是否已经知道每种跃迁在现实中有多大概率。

相加的是光量子能量，也就是频率乘 h，不是波长。两次光的波长不能直接相加来得到单次光波长。把谱线列表按频率组织，会显露出差值网络；按波长看则需要先取倒

数。这与黑体章中密度变量转换的教训相呼应：表格里使用什么量，会影响哪些关系显得简单。

若一组假设的谱线给出 A 到 B 为两单位能量、B 到 C 为三单位，A 到 C 却为六单位，就无法由三个固定能级同时产生。可能的解释包括谱线归属错误、来自不同物种、外场改变状态，或候选模型不足。这种闭合检查比单独匹配某一条线更严格，因为任何单线都容易被一个自由参数追上。

同时，能量差允许不等于跃迁必然发生。最简圆轨道模型没有充分给出所有选择规则、强度和寿命。若某条差值对应的线没有出现，不能不加分析便说能级不存在；需要研究制备是否占据上态、过程是否允许及探测是否敏感。另一方面，也不能用“跃迁不允许”给每条缺失线事后补借口，而应寻求统一的选择规定。

这个区别使原子世界具有双重层次：状态决定能量如何比较，跃迁动力学决定哪些比较在实验中成为事件。玻尔模型首先强力重建前者，并为后者提供频率条件；成熟量子理论后来才把更多任务放入统一形式。本章停在这个不完整而富有生产力的位置，才能与后面的矩阵力学保持真正接口。

39.6 让核也运动：一个小修正怎样约束大解释

固定核近似方便，却不是精确的两体问题。若核质量为 M，电子与核围绕共同质心运动，相对坐标中的惯性由约化质量 μ＝mM/(m＋M)承担。于是能级表达中的 m 要换为 μ，相关谱线系数也随 μ 变化。核很重时 μ 接近 m，所以固定核近似能够成功；它并不是对所有原子都毫无误差。

设质子质量约为电子的 1836 倍，则氢的 μ/m 约为 1836/1837，约 0.999456。若另一个核质量约为 3670 个电子质量，相应比值约 0.999728。在相同 Z 与同一跃迁下，第二种系统的频率略高，两者相对差约为万分之二点七二。这种微小差别来自核运动，不需要更换电子电荷或为每种同位素设一套 h。

用 656 纳米附近的一条线演示，相对频率增大约 0.000272，对应波长缩短约 0.18 纳米。这个值只作最低阶模型估算，未包含精细结构、实际核质量与其他修正。它说明“核很重所以不用管”与“核可近似固定但仍有可测后果”是两种不同判断。近似应当能说出误差方向和尺度。

约化质量也帮助区分竞争解释。若误差来自波长标尺的乘性偏差，多种物种可能出现共同比例偏移；若来自核质量，偏移依赖 M 且具有特定极限。选择不同核质量而相同

电荷的系统，可以让两种解释分叉。理论修正只有给出跨对象约束，才比为每条谱线补一个修正数字更有内容。

还应检查坐标的角色。约化质量不是说电子真的变轻，而是把两体运动改用相对坐标表达后，动能获得一个有效惯性系数。真实质量、有效参数与坐标选择不能混同。这一点与相对论章节的经验相接：数学表达变化可能重新组织计算对象，却不自动等于对材料实施了真实干预。

39.7 旧运动频率在什么极限回来

对圆轨道模型，轨道频率为 $f_{轨}=v/(2\pi r)$。代入前面的表达，有 $f_{轨}=m\kappa^2/(2\pi n^3\hbar^3)$。而相邻态 $n+1$ 到 n 的发光频率，由能量差给出 $\nu=[m\kappa^2/(2\hbar^2 h)]\ [1/n^2-1/(n+1)^2]$。两者在低 n 一般不同。

计算它们的比值，可化为 $\nu/f_{轨}=n(2n+1)/[2(n+1)^2]$。$n=1$ 时比值为 3/8，不能说两个频率近似相同；$n=10$ 时约为 0.868；$n=100$ 时约为 0.985。随着 n 增大，比值趋近一，相邻状态能量差的频率逐渐接近下态的经典轨道频率。

这给出一个具体的对应关系。新规则不必在每个状态都复制经典图像，但在适当极限应能恢复经典成功的联系。这里“适当”包含高量子数、相邻跃迁和当前圆轨道模型；若选很远的两态，不能直接使用同一近似。对应不是一句要求新理论尊重旧理论的礼貌话，而是一项能够检查的极限计算。

这项极限也揭示经典连续性如何可能出现。高 n 处相邻能级更密，能量变化的细分在某些观察尺度下可以近似连续。但这不表示整数条件在微观层面被取消，而是离散差别相对于测量尺度变小。连续经验与离散规则能够在限定范围共同成立，不能靠“一个真一个假”的简单对立处理。

同时，高 n 态的半径扩大、束缚变弱，外界电场、碰撞等因素可能更容易扰动它。数学上取 n 很大，不代表实验中可以忽略所有其他条件。要在现实里观察对应极限，还需要控制环境耦合和分辨尺度。E 在这里涉及特征怎样共同限制，而不只是被动摆在原子周围的外部场景。

39.8 成功的限制与失败的分类

单电子谱的成功容易使人误以为，只要继续画更多圆圈就能得到所有元素。实际多电子问题存在电子之间的相互作用，简简单单给每个电子一个独立氢式轨道，并不能使

所有能量、稳定性和化学规律同时成立。对某些谱线位置的改进，也不保证强度、精细分裂及磁场响应都已解决。

这里至少有三类失败需要区分。第一是算术或近似失败，例如忘记核运动、把波长相加；第二是模型适用范围失败，例如把单电子式直接用于中性多电子原子；第三是结构性失败，例如需要一套统一规则给出状态转换及强度，却只有定态与频率两项拼接条件。三类失败要求不同的研究行动。

若把所有问题都称为“玻尔理论不够精确”，就可能暗示只需把小数多算几位。结构性失败则不同：计算得越准确，也未必补得上缺少的构成规则。反过来，若因为后来理论更广便把早期模型的全部成功说成偶然，又无法说明它为什么能够组织那么多数量关系。继承需要逐项辨认，不能靠整体判决省略。

另一种竞争解释会把每条谱线归于一个独立经典振子。这可以描述有限数据，但如果振子频率互不相关，就不自动给出能级差的组合规律、Z^2 缩放和约化质量修正。玻尔方案的优势在于把这些关系绑定在一组状态参数中。它承担更多风险，也因而具有更强的解释统一性。

不过，统一性不是逻辑唯一性。不同理论可能产生相同的谱线位置；仅靠这些位置不能裁决所有关于电子轨道的实体主张。后来可以在没有经典圆轨道的框架中保留相同的某些能量结果，恰好说明数值成功与图像真实性并不等价。本章不提前推导那一框架，只把需要区分的证据层级明确留下。

还需辨认“定态”与“永不改变”。定态规定原子在相应状态中不按经典轨道方式持续辐射，不表示任何外界作用都不能使它变化。吸收、碰撞和电离仍可以把系统带离原状态。把定态说成一个完全不受作用的封闭实体，会使理论失去与实验接触的入口，也无法解释激发后为何出现不同谱线。

最低态与激发态也不应因同被整数标记，就被认为具有相同的长期命运。模型允许列出若干能量，但列出能量并未给出各状态在具体环境中的占有比例和寿命。若实验主要看到某几条线，可能是制备选择了特定状态，并不表示剩余能级被自然永久取消。状态空间与实际样本分布，是需要通过制备过程连接的两个层次。

一个有用的思想实验是用两种不同办法激发同一种稀薄气体：一种输入能量不足以到达某高态，另一种足以到达。高态相关谱线只在第二种条件出现，可以支持制备差异的解释；若所有线都随装置移动相同幅度，更应怀疑标尺或整体运动影响。实验改变的是入态、测量条件还是物种身份，应由数据关系决定，不能把所有光谱变化都称为原子结构发生了根本改变。

39.9 从稳定性冲突走向新的对象规定

玻尔案例可以沿 S—E—D 进入路径重构。稳定物质与离散谱线是入口；把它们放入集中核、库仑作用和经典辐射的关系中，出现不能共同满足的张力；据此改变差异运行，把研究从任意轨道演化转向允许状态、能量差与频率组合的计算和比较。完整路径不是从看到一条线直接宣布世界量子化，而是经过共同条件的重建。

随后沿 E—D—S 沉淀路径，用同一状态规则推算其他跃迁、改变核电荷与核质量，再把未用于定参的数据带回检查。前面的 656 纳米、486 纳米、He^+缩放和核质量修正，都属于能够把一个结构推进到新条件的操作。每一次外推都应注明适用边界，否则跨域成功可能只是人为选择有利例子。

三方程各自约束一个容易跳过的环节。S＝F（D，E）要求谱线由制备、跃迁和状态规则共同形成，不能把它当作直接拍到电子轨道的照片；D＝G（S，E）要求选择能区分独立振子与共同能级网络的比较；E＝H（S，D）要求允许状态在稳定性与谱学约束下重建，而不是凭几何美感选圆圈。

两维张力推动第三维改变，在本章首先表现为稳定性显露与经典运动推演之间的冲突推动 E 重写。新定态结构 E 要求比较能量差，而已有分光操作 D 仅按波长位置直接相加时，S 需要重新形成按频率或能量表达的可比谱线关系；确定的谱线与状态结构冲突，则推动 D 改变激发条件、核质量比较或外场控制。修改之后仍要回到原来的稳定性和谱线问题，不能只庆祝新术语出现。

这种发生学解释也面临自己的检验义务。如果它只在每个成功理论后面贴上三字母，就没有增加历史理解。本章尝试使它承担一项更具体的责任：指出圆轨道几何、定态能量和跃迁频率并非先已独立完成的属性，而是在互相限制中取得新的身份；同时指出哪些联系依然缺少统一机制。成功与空缺应当一起进入解释。

39.10 保留模型的桥梁，同时看见桥下的空处

玻尔原子具有一种特殊的思想史力量：它没有彻底放弃经典轨道，却让轨道不再能单独决定原子的全部物理身份。允许哪些状态、状态何时辐射、辐射频率取决于什么，都由新关系重新安排。轨道图像仍在，但它已不再拥有原来的全部权利。

守恒在这次改变中没有被抛弃。状态差与光量子能量必须相等，多步转换的总收支仍应闭合。改变的是守恒参与的对象：它不再只追踪一个电子沿任意连续轨道平滑失去能量，而是在离散状态和辐射事件之间建立约束。同一守恒思想能够进入不同的物理世界，却不保证这些世界允许相同的过程。

这也说明创造物理对象不等于随意创造参数。状态规则越紧密，允许的成功组合越少，模型越容易在新谱线和新物种上受挫。创造的力量与可失败性同时增强。把物理世界说成关系中成立，若不能保留这种抵抗，就会从发生学滑向任意化；本书所要求的回检正为防止这一步。

最后仍有一个开放问题：为什么允许状态要服从整数条件？如果整数仅是一项成功规则，理论便还欠一个结构来源。把电子的运动与某种波的相位条件联系起来，可能为整数提供新的解释入口；与此同时，光的散射也将要求能量份额携带明确动量。下一章沿这两条路线推进，让旧有的波与粒子分类面临更深的重组。

本章参考资料

[39-1] Bohr, N. On the Constitution of Atoms and Molecules, Part I. Philosophical Magazine 26 (1913):1–25. DOI:10.1080/14786441308634955. https://www.informationphilosopher.com/solutions/scientists/bohr/Constitution_of_atoms_I.pdf。使用引言、定态与频率关系部分；本文圆轨道、约化质量与极限算例均为现代记号重构。

[39-2] Kragh, H. The Early Reception of Bohr's Atomic Theory (1913–1915). RePoSS No.9, Centre for Science Studies, Aarhus University, 2010. https://css.au.dk/fileadmin/reposs/reposs-009.pdf。使用引言与“Content and character of Bohr's trilogy”，限定单电子教材图像与历史理论范围的关系。

第四十章 德布罗意与康普顿 波和粒子的边界怎样被重新组织

40.1 分类本身开始成为问题

如果把波规定为铺展于空间、可以干涉的连续过程，把粒子规定为携带一定能量与动量、在局部事件中交换的对象，那么光与电子似乎各有自己的位置。但前面几章已经使这张分类表不再安稳：光的传播保留波的规律，光与物质的交换却表现出份额约束；电子能够偏转并携带电荷，原子的允许状态却不能由任意经典轨道构成。

新的问题不只是某一种东西“究竟属于哪一栏”，而是两栏中哪些属性必须一起保留，哪些原来的绑定可以拆开。局部交换不必要求一颗有硬表面的微型弹丸，空间相位联系也不必要求把电子当成一团普通液体。若不重新检查这些默认绑定，波粒二象性便很容易变成一个不解释任何事的名词。

德布罗意在一九二四年的论文中尝试把物质运动与相位波联系起来，并讨论相位速度、群速度和周期运动条件。[40-1]康普顿一九二三年的散射论文则把光量子的能量动量交换与散射后波长联系起来。[40-2]本章标题把两位作者并置，不表示康普顿工作发生在德布罗意之后，也不把两条研究路线压成单向灵感链。

科学史家斯图韦尔研究康普顿效应时，强调他在一九一六至一九二二年间持续推进的实验与理论计划，并指出不能把最终结果简单解释为直接照搬爱因斯坦光量子假说。[40-3]这个提醒使我们将注意力放到问题怎样在自身的反复修正中成熟，而不是用后来清楚的公式替代当时艰难的探索。

40.2 给电子关联一个波长究竟增加了什么

德布罗意关系的现代写法是 $\lambda = h/p$。这里 p 是粒子动量的大小，λ 是关联的空间波长。它不是由“所有东西都像水波”推出来的，也没有单独说明波的振幅是什么、通过什么方程演化、如何与探测概率连接。它首先提出一条新的数量桥梁，使机械动量与相位变化的空间尺度能够共同限制。

若电子以非相对论速度运动，p 约为 mv。通过电势差 V 从近似静止加速，动能约为 eV，于是 $p = \sqrt{(2meV)}$，$\lambda = h/\sqrt{(2meV)}$。这说明电压增大四倍，波长缩短一半；它不是随电子束亮度或粒子数直接改变的量。束流强度与动量制备应分开比较，正如第三十七章讨论的光强与单次能量不能混同。

以电子经过 150 伏加速为例，波长约为 0.100 纳米，已与常见晶体中的原子间距处于相近尺度。若取 600 伏，则约为 0.050 纳米。数值说明关联波长不是只能停留在哲学比喻里的概念，它可以把电压控制与衍射角联系起来，给出有区分力的实验目标。

假设一个理想晶面间距为 0.20 纳米，用 $2d\sin\theta=\lambda$ 的一级条件作几何演示，0.100 纳米对应 θ 约 14.5 度，0.050 纳米对应约 7.2 度。这里 θ 是相对晶面的掠射角，实际入射与出射方向之间的角度定义还需按装置说明；不能把一台仪器上标注的散射角不加转换地代入另一种约定。

这个算例不等于已经给出真实晶体衍射的完整理论。晶面方向、多重散射、表面势和晶格结构都会影响强度与峰形。但如果电压变化后峰位按预测移动，就能把“存在某种周期性散射”推进为“周期尺度与动量通过同一 h 相接”。若只看到条纹便宣称任意物质都已被证明为波，仍然跨过了必要步骤。

40.3 相位走得很快，不代表电子超光速

从相对论能量动量关系出发，记总能量为$\mathcal{E}$，有$\mathcal{E}^2=p^2c^2+m^2c^4$。若采用$\mathcal{E}=\hbar\,\omega$、$p=\hbar\,k$，则相位速度 $\omega/k=\mathcal{E}/p$。对有静质量的自由粒子，$\mathcal{E}/p=c^2/v$，因此当 $0<v<c$ 时，相位速度大于 c。这个结果并不要求电子本身以该速度运动。

相位速度跟踪的是一个等相位面的移动。例如在形如 $\cos(kx-\omega t)$ 的表达中，令 $kx-\omega t$ 为常数，就得到 $dx/dt=\omega/k$。这个数学面不是一块携带固定能量的材料薄片；它可以在空间不同点取得相同相位，却不因此建立可用于超光速通信的信号通道。

若把一组相近波数的自由波叠加，包络的群速度为 $d\omega/dk$。由上述能量动量关系，$d\omega/dk=d\mathcal{E}/dp=pc^2/\mathcal{E}=v$。相位面与包络因而具有不同速度。此处结论依赖当前自由粒子的色散关系，不应推广为任何介质、任何脉冲下群速度都等于信息速度的普遍命题。

用 $v=0.6c$ 演示，$\gamma=1.25$，相位速度为 $5c/3$，而群速度为 $0.6c$。两种速度不同并不矛盾，因为它们追踪的对象不同。若把相位面当作粒子世界线，就会像在相对论中把同一事件的坐标变化误认为真实干预一样，把描述关系混成物质运动。

还有一种更隐蔽的混淆。与静止粒子关联的固有周期频率可写为 $\nu_0=mc^2/h$；随粒子运动的钟样周期在实验室每单位时间只经历 ν_0/γ 次，而波在固定空间点的频率为 $\gamma\nu_0$。二者看似一增一减，是因为比较不同的事件序列。沿粒子路径 $x=vt$ 代入相位 $2\pi\nu(t-vx/c^2)$，得到 $2\pi\nu_0 t/\gamma$，正好与随动周期一致。这是相位协调的数学要点，而非两个互相矛盾的时间膨胀公式。

40.4 整数条件可以从闭合相位重新提问

前章留下的问题是为什么角动量条件出现整数。若把某条圆周上关联的波要求为绕行一周后相位复原，则周长应满足 $2\pi r = n\lambda$。代入 $\lambda = h/p$，便得 $pr = n\hbar$；对非相对论圆轨道 $p = mv$，又回到 $mvr = n\hbar$。整数由闭合条件取得一种新的结构解释。

这个推导的力量在于，把看似外加的允许轨道规则，与相位一致性相接。一个绕行后相位不能接合的候选方式，不再只是“自然不喜欢的半整数”，而是与规定的闭合边界不相容。量子条件因而有机会从额外禁令转化为波结构的允许条件。

它的限度同样明确。我们事先指定了圆轨道，并在这条路径上要求相位闭合，尚未由更普遍方程推出全部原子状态。一般势场中的允许态、三维角向结构、多电子问题和跃迁强度，都不能从一条周长等式自动得到。如果把它说成已经证明现代氢原子的电子就是一条真实圆形驻波，就超出了论证。

还可以做一个反向检查：将 $n = 2$ 的允许圆周固定，而把动量改为原来的一点一倍，则波长缩短，原周长包含二点二个波长，不满足当前闭合条件。要恢复允许关系，不能任意只改一个量而保持所有其他条件；半径、动量及力学约束必须一起协调。这个小例子说明波不是在既有轨道外面添加一层装饰，它会重新限制轨道能否成立。

但闭合相位也不是所有量子化的唯一来源。开放边界、散射态和自由粒子可以具有连续动量范围。若把“量子”一概解释成所有变量都只能取整数值，就会在最简单的自由运动问题上出错。离散与连续要由具体边界和动力学共同判断，而不能由一个总名称替代。

40.5 光的动量在一次碰撞中接受检验

现在转向康普顿问题。设入射光量子能量为 ε、动量大小为 ε/c；一个电子初始静止。散射后光量子以角 θ 离开，能量变为 ε'，电子反冲。若每次事件都满足能量与动量守恒，就有 $\varepsilon + mc^2 = \varepsilon' + \mathscr{E}_{电}$，其中 $\mathscr{E}_{电}$ 是末态电子的总能量。

动量守恒还给出末态电子动量的平方：$p_{电}^2 c^2 = \varepsilon^2 + \varepsilon'^2 - 2\varepsilon\varepsilon'\cos\theta$。这里不能只保存能量而随意决定反冲方向，也不能只保存动量而仍把电子能量写成未经检查的经典小速度表达。两项守恒必须与电子相对论能量关系共同使用。

由 $\mathscr{E}_{电}^2 = m^2c^4 + p_{电}^2 c^2$，把 $\mathscr{E}_{电} = mc^2 + \varepsilon - \varepsilon'$ 代入，展开并抵消相同项，得到 $mc^2(\varepsilon - \varepsilon') = \varepsilon\varepsilon'(1 - \cos\theta)$。于是 $1/\varepsilon' - 1/\varepsilon = (1 - \cos\theta)/(mc^2)$。用 $\varepsilon = hc/\lambda$，最终得到 $\lambda' - \lambda = [h/(mc)](1 - \cos\theta)$。

推导只用了几行，但每一项都承担物理任务：光量子能量与频率关系给出谱学接口，动量给出方向接口，电子质量给出反冲尺度，逐事件守恒把它们锁在同一过程。若删掉任何一环，波长随角度变化的特定形式就不再由这组条件共同规定。

h/(mc)称为电子康普顿波长，约 2.426 皮米。它不是电子半径，也不是电子在任何状态下的德布罗意波长。前者由静质量形成一个固定尺度，后者随动量改变。两者都含 h，不能因此被当成同一种几何长度。

40.6 用数字区分不同尺度

设入射波长为 0.071 纳米。若散射角为九十度，波长增加 0.002426 纳米，变成约 0.073426 纳米；初始能量约 17.46 千电子伏特，末态光能量约 16.89 千电子伏特，电子获得约 0.58 千电子伏特动能。数值四舍五入造成的末位差异应与实验误差区分。

若散射角为一百八十度，位移增为两倍，即 0.004852 纳米；若为零度，在初始静止自由电子的理想模型中位移为零。它不是“光只要碰到电子就一定变红”的规则，角度参与决定能量交换。若入射电子原本运动，还可能出现不同的能量变化方向，本章的静止靶结论不可直接套用。

把入射波长改为 0.142 纳米，在相同九十度下位移仍为 0.002426 纳米，但相对变化比例减半。因而实验应比较绝对波长差而不只是百分比变化。相同位移而不同相对比例，是这一模型区别于某些简单缩放假说的重要特征。

对于 500 纳米可见光，同一位移相对于原波长仅约百万分之四点九。即使过程符合相同公式，直接分辨这样的变化也比 X 射线例子困难得多。效果未被普通棱镜明显分开，不表示能量动量关系在那里自动失效；可分辨性还取决于尺度与装置。

如果散射由质量很大的整体近似承受反冲，公式中的相关反冲质量尺度会大得多，位移可以小到近似不变。这有助于理解为什么真实散射光谱不一定只有一条完整移位峰。束缚、相干作用和初始电子运动都可能改变结构，不能把真实材料中的每个电子都当成独立静止靶而不检查条件。

40.7 与经典散射及统计守恒的竞争

最简单的经典线性散射图像中，入射电场驱动电子在相同频率振动，再辐射同频率的光。它可以解释某些角分布和偏振关系，却不会在这套最简前提下产生上述特定角度相关的波长位移。这里受到挑战的是一个明确模型，不是所有含连续场的理论已经被一条式子从逻辑上全部排除。

别的候选解释包括材料特有荧光、入射光谱混合、多重散射和多普勒变化。若是材料特征线，其位置未必按同一个（1−cosθ）规律连续移动；若是固定的仪器标尺误差，也不应随观察角按理论形状变化。独立改变入射波长和散射角，能让这些解释承担不同预言，而不只是比较谁的故事更直观。

平均谱符合仍不等于逐事件守恒已被全部确认。设想一种理论仅要求大量事件的能量动量平均平衡，而不要求每个电子与每份散射光在同一事件中配对，它可能在平均量上隐藏差异。若能同时记录反冲电子和散射光的时间、方向及能量关联，就能对这种区别施加更强限制。

这一问题在历史上并非多余：斯图韦尔将玻尔—克拉默斯—斯莱特方案及其经验命运列为康普顿效应之后的重要回应。[40-3]这里不详细重演后续实验，只指出思想上的证据层级。守恒既可能被理解为单事件关系，也可能被候选理论改写为统计关系；要判断哪一层成立，必须选用能够区分它们的观测。

第三十五章讨论的守恒与对称、边界条件关系，在这里又与碰撞的对象规定相接。守恒本身并没有告诉我们光究竟是什么，却在与能量频率、动量方向及电子质量联合后，强烈限制了光量子的身份。它既是跨理论延续的资源，也会因对象与过程的重组获得新的检验方式。

40.8 一次交换与一组相位不能互相取消

德布罗意路线把波长带给有质量粒子，康普顿路线把定量动量交换带给辐射。两条路线的相遇不是让世界中出现两类互相替换的照片，而是迫使我们放弃一种过强的分类：只有铺展的东西才可干涉，只有硬球式的东西才可携带局部交换。

因此，“电子变成波”不意味着电子电荷在实验记录中消失，“光变成粒子”也不意味着衍射规律被撤销。新的理论必须同时保存那些已得到不同操作支持的关系，并说明它们如何共同存在。把互相困难的关系简单称为互补、二象或奇异，都不能替代这项构造任务。

这里也不能直接把波的空间分布认作观察者的意识状态。相位关系通过可控制的路径、晶体和动量制备接受检验；记录由真实相互作用产生。数学对象与视觉形状不同，但不因此失去客观约束。相反，越不能依赖日常形状，我们越需要清楚说明哪些操作能够支持它。

从 SDE 看，电子的“粒子性”与“波性”不是两个先已完成的属性标签，再被外力粘在一起。动量、波长、边界相位、探测事件和制备方式在新的共同规定中重建彼此的

作用。不过，这仍是一种发生学解释，不是从 λ＝h/p 演绎出 SDE 哲学已被证明；物理公式的检验范围必须与哲学概括的责任分开。

40.9 把两条路线放进可回检的推演

物质波的有效路径可从 E—D—S 沉淀开始：先给出动量与波长的关系及晶格周期条件，再安排电压扫描和角扫描，最后比较峰位是否按共同参数移动。S＝F（D，E）在这里要求图样来自制备与相位结构，而不是只因“像波”就宣布波的身份；D＝G（S，E）要求下一次电压选择使候选缩放规律明显分叉。

若相同电压操作产生的图样与预期不合，E＝H（S，D）要求检查动量制备、表面势及散射结构。S 与 D 之间的稳定差别可推动 E 修订；D 采用宽能量束而 E 假定单一动量时，二者张力推动 S 改为带有展宽和分辨限制的预测图样；S 与 E 仍冲突则推动 D 收窄能量分布并重新校准。改变必须返回原始峰位关系，而非把所有差别重新命名为新波态。

康普顿研究则可沿 S—E—D 进入：从角度相关的谱位移进入能量、动量和反冲的共同限制，再组织独立改变入射波长、散射角及材料的操作。九十度下绝对位移保持而相对位移改变，正是一个明确回检点。理论若只能够事后解释某一张谱图，还没有完成这条路径。

两条路径需要不同的显露定义。物质波比较相位条件下的峰位，散射交换比较光与电子的关联；它们不能由同一张“波粒图”包办。也不能把波统一配给 E、运动统一配给 D、亮点统一配给 S，因为每一个过程内部仍有其差异运行、特征关系与稳定显露。三维划分的价值在于追问条件，不能成为新的实体分类表。

40.10 走向一个不必先给出轨道的理论

本章完成了两项桥接，却留下一个共同缺口。λ＝h/p 规定动量与波长怎样相接，但尚未完整决定任意相互作用中的状态演化；康普顿关系规定一次交换怎样满足守恒，但尚未给出所有过程的发生概率与振幅。早期量子理论获得更多紧密关系，也越来越清楚地暴露统一语言的需要。

更深的困难仍来自原子。闭合相位能够解释某些整数规则，却仍借用轨道；谱线频率由两个状态的差决定，轨道频率却属于一个状态。若继续要求完整轨道先行存在，再把量子条件附加其上，研究将不断面对同一对象在不同环节被不同规则规定的矛盾。

下一章的矩阵力学将从这处张力出发，追问是否必须先提供电子的完整轨道，才能组织频率、跃迁及其组合。本章不提前给出答案，只说明问题为何已经成熟：波与粒子的旧边界被松动以后，理论不能永久停留在两套图像轮流借用的阶段。

物理思想的创造在这里表现为一种严格的重新安排：同样的 h 把热分配、光能量、原子状态和物质动量连接起来，但每一处连接都有不同边界。统一常量并不使对象差异消失，而让不同领域能够互相施加压力。新的世界由这些压力推动形成，也必须在新的实验中继续承担失败的可能。

因此，所谓统一不能只看公式里是否重复出现同一个符号，还要看常量能否跨不同制备、仪器与对象保持数值联系。若相同记号背后每次都重新调整含义，统一只是语言上的整齐；若不同实验能够彼此限制，它才成为物理上可工作的共同结构。

本章参考资料

[40-1] de Broglie, L. Recherches sur la théorie des quanta. Thèse, Paris, 1924；Annales de Physique 3 (1925):22–128. A. F. Kracklauer 英译 On the Theory of Quanta，https://fondationlouisdebroglie.org/LDB-oeuvres/De_Broglie_Kracklauer.pdf。使用第一章相位与群速度及周期条件概述，现代演算为本章重构。

[40-2] Compton, A. H. A Quantum Theory of the Scattering of X-Rays by Light Elements. Physical Review 21 (1923):483–502. DOI:10.1103/PhysRev.21.483. 原件：https://history.aip.org/exhibits/gap/PDF/compton.pdf。本章使用散射能量动量关系的现代独立推导。

[40-3] Stuewer, R. H. The Compton Effect: Transition to Quantum Mechanics. Annalen der Physik 9 (2000):975–989. DOI:10.1002/andp.200051211-1216. https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.200051211-1216。使用论文摘要对研究计划及后续竞争理论的概述。

第九编 量子态、测量与属性纠缠的最艰难检验

第四十一章 海森堡矩阵力学 经典轨道之外还有什么物

41.1 当轨道不能再承担原子的身份

如果把原子想成一个极小的太阳系，那么理解原子的第一任务似乎十分清楚：找到电子在哪里、怎样运动，再从运动推出它发出的光。这里隐藏着一项比计算公式更深的承诺：在这种图像里，电子之所以是一个物，首先因为它能够在每个时刻占据一个确定位置；位置连成轨道，轨道赋予运动连续性，辐射则是这个已经成形的物向外发出的表现。即使暂时看不见轨道，人们也容易相信，只要仪器足够精密，终会补齐那些空白。

前章所讨论的量子条件已经动摇了这幅图景，却没有完全撤走它。若仍由经典运动规定轨道，再由额外规则筛选允许的状态，原子的同一性就同时受两种不完全相容的组织方式支配。轨道在机械方程中可以连续变化，允许状态却不能照此任意变化；辐射关联于能级差，未必等于电子沿轨道运行的频率。问题不仅是算得不够准，而是不同计算环节对同一个原子提出了彼此冲突的身份要求。

海森堡于一九二五年提交的重释论文，把原则上可观察的量之间的关系置于理论出发点。他不再以完整电子轨道作为必要中介，而以频率、跃迁振幅及其组合来重组运动学，并在一自由度模型中检验这种方案。[1] 这几项史实已经足够说明转折的方向。至于他当时每一步的心理动机，本章不依靠后来流传的灵感故事加以补齐。一个理论的发生可以有偶然场景，但它取得物理效力，终究要经过能够被他人重复的推导与检验。

本章的主要问题因此是：拿掉经典轨道之后，剩下的是否只是仪器读数的清单？如果不是，新物理中的电子靠什么获得稳定规定？我们的回答将沿着一个具体环节展开：乘法。乍看之下，乘法只是已经知道两个量以后才使用的工具；在这一转折中，怎样把量相乘，却直接参与决定什么样的量能够成立。物的规定不再先于所有关系而完整地摆在那里，关系的可组合性开始进入物的身份。

从本书的 SDE 立场看，这是物理世界的创造问题。这里的创造指新的差异序列与特征纠缠，使一种过去无法稳定成立的物理规定系统成为可能；它不表示人的意识凭空制造电子，也不表示电子是否存在由个人相信与否决定。真正需要说明的是：实验中的稳

定抵抗、数学中的一致要求与操作中的可重复效果，怎样共同约束一种新物理世界的形成。

41.2 从一个位置到两个状态之间的量

为理解这个变化，可以先做一个不依赖历史心理复原的数学比较。经典周期运动常可写成傅里叶展开。一个函数的位置值随着时间改变，而不同频率分量共同组成这条运动曲线。用示意记号写，便是各个振幅乘以相应的周期因子，再对谐波指标求和。这里每一项仍然服务于同一条轨道；它们好比同一句话里的不同音调，最终能够合回一个完整的时间函数。

转向跃迁描述之后，关键量带有两个状态指标。我们采用现代化的记号，以 W_n 表示第 n 个定态的能量，以 ω_{nm} 表示两个状态之间带符号的跃迁角频率，那么频率条件写成 $\omega_{nm}=(W_n-W_m)/\hbar$。字母 W 用于能量，是为了避免与本书表示特征纠缠的 E 相混。这一表达不是本书提出的新物理定律，而是本章说明计算结构的共同起点。

双指标意味着什么？在轨道图像中，问题是某一时刻的电子位置；在这里，指标 n 和 m 共同限定一个量。单独留下其中一个，通常不能恢复原来的规定。这种改变并不等于把电子换成两颗电子，也不等于说每一次辐射都能被看作一段已经拍摄下来的空间路径。它意味着描述所需的最小关系单位发生了变化：从沿一条曲线取值，转到一套状态之间的量如何组织。

因此，“只研究可观察量”不能被理解成只准使用肉眼直接看见的东西。谱线可以在仪器上显露，能级差、振幅与相位却需要通过理论和实验程序得到限定。某个矩阵元也不是一枚可以单独装进试管的物体。如果要求所有数学量都必须逐个变成直接可见的实体，新的理论反而无法建立。问题的核心是数学量与可检查结果之间有没有受约束的通路，而不是它是否长得像日常物件。

这对 SDE 分析尤其重要。视觉仪器记录下来的谱线不是一份脱离生成条件的纯客体画像。光源制备、分辨程序、频率比较和重复辨认使差异获得稳定形式；但这些环节也不是可以任意增删的背景列表。某项条件只有在改变了谱线的身份、关联或可复现性时，才进入当前特征纠缠的构成分析。我们要追查的是条件怎样进入规定，而不是给历史事件配一张无限扩展的环境清单。

41.3 乘法为什么必须经过中间指标

现在考虑两个量 A 和 B。若它们都是普通数，把它们相乘似乎不需要先问来自哪两个状态。然而对带有双指标的量而言，必须明确什么叫组合。设一个分量带有时间因子

$\exp(i\omega_{nk}t)$，另一个带有 $\exp(i\omega_{km}t)$。两者相乘，指数里的频率相加。由于能量差相消，便有 $\omega_{nk}+\omega_{km}=\omega_{nm}$。中间指标 k 因此并非装饰，它保证组合后的频率仍然属于起点与终点所限定的同一类量。

若要把所有允许的中间指标纳入组合，便得到熟悉的形式：$(AB)_{nm}=\Sigma_k A_{nk}B_{km}$。这一式子可以直接复算：每项的中间指标前后相接，所有项具有一致的起止指标，时间因子也具有同一总频率。矩阵乘法在这里展现的价值，是让组合之后的量仍留在同一套可使用的规定系统中。这个闭合要求，比“写成一个方阵比较方便”深得多。

这里需要一个重要区分。式中的中间指标求和，首先是数学组合规则；它不自动表示电子在真实时间里先走到每个中间状态，再走到终态。某个计算含有中间状态，不等于已经观察到一串实际跃迁事件。若把代数中的索引连接直接画成微小弹珠旅行图，经典轨道会以另一种未经证明的方式被偷偷放回。我们的讨论关注量的组合规定，实际过程的解释还必须结合制备、耦合和测量条件。

可以用一个纯代数例子检查顺序的意义。取 A 为只有第一行第二列等于一的二阶矩阵，B 为只有第二行第一列等于一的二阶矩阵，其余元素都为零。按上式相乘，AB 的第一行第一列等于一，其余为零；BA 的第二行第二列等于一，其余为零。所以 AB 不等于 BA。这里没有任何仪器误差，也没有人的意志干预；不同结果来自组合规则本身。

这个例子当然不是电子位置与动量的完整模型。它只证明：双指标对象的乘法一般可以不交换，不能继续把普通数的所有习惯未经检查地带过来。要把某一组矩阵认作物理量，还必须规定状态空间、动力学及与实验结果的对应关系。不对易是一个结构特征，物理意义则要在更完整的理论中成立。代数演示和物理认定是两项相关而不同的工作。

41.4 不对易改变了什么，尚未说明什么

玻恩与约尔当随后把海森堡的方案发展成系统的矩阵力学，明确使用矩阵演算，并将量子条件、运动方程和能量守恒联系起来。[2] 用现代常见约定表示，位置与动量满足 $[q,p]=qp-pq=i\hbar I$。这个公式不能被随意简化为“先量位置和后量位置总是不一样”；它首先表达两个算符的代数关系，具体测量次序产生什么统计结果，还需要相应的测量模型。

位置和动量不再只是两张可以无限精确地同时填满的普通数值表，但这也不是宣布计算从此失去规则。恰恰相反，原先被默认为相同的两个组合，现在必须清楚地区分。

理论的纪律加强了：每次交换次序都需要理由，每次把复合表达式改写都要检查代数条件。物理创造并非摆脱约束，而是让过去不曾显露的约束取得可运算的形式。

读者也许会问，既然刚才可以用二阶矩阵展示不对易，为什么不直接用二阶矩阵表示所有位置和动量？答案可以通过一个简短推导看出。有限阶矩阵满足 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$，因此任何对易子的迹都为零。如果两个 N 阶矩阵精确满足 $[q,p]=i\hbar I$，对两边取迹，左边为零，右边则是 $i\hbar N$。对非零 $\hbar$ 和正整数 N，这是矛盾。

这个结果并不否定有限维量子系统，也不否定有限维近似在计算中的用途。它说明，有限维矩阵不能精确满足这一特定的正则对易关系。无限维算符的迹与定义域还需要专门处理，不能机械搬用有限矩阵的论证。数学上“看起来都是矩阵”的对象，可能有重要的成立条件差别。教学若省去这些条件，就很容易把漂亮的小例子误当成完整自然图像。

从发生学角度看，这个细节提供了一项有效约束：不能只凭某种数学形式与哲学词语相似，就宣布它证明了整个本体论。双指标、不对易和中间状态求和确实支持我们重新追问物的规定方式；它们并不独立推出 SDE 理论的全部命题。SDE 解释要有用，就必须准确辨认这些结构在哪些条件下成立，并说明由此能够提出什么新问题，而不是把所有公式都译成同一句口号。

41.5 轨道退场以后，微积分并未退场

矩阵力学还纠正了另一种过于简单的历史叙事：经典理论使用微积分，量子理论换成矩阵，所以两者之间只剩断裂。事实上，即使不再用一条经典轨道组织原子，时间变化仍需要表达。矩阵元可以随时间变化，对它们求导便会带出相应频率。连续时间与非交换代数可以在同一理论中共存；离散状态指标也不等于时间必须由一格格不可再分的跳步组成。

采用现代算符语言，对没有显式时间依赖的物理量 A，海森堡运动方程写为 $dA/dt=(i/\hbar)[H,A]$。如果 A 本身还有显式时间依赖，右端应再加上偏时间导数。这里的 H 是哈密顿算符。本章使用这一成熟形式来展示数学关系，不把它的所有现代解释倒置为海森堡最初论文已经完整具备的内容。

以 $H=p^2/(2m)+V(q)$ 为例，在满足正则对易关系且定义域允许计算的情况下，可以推出 $dq/dt=p/m$ 以及 $dp/dt=-V'(q)$。表面上它们很像经典方程，但相同的字形承载着不同的对象类型。q 与 p 不再被理解为沿某条预先存在的轨道同时取确定值的普通函

数；它们是处于特定代数结构中的算符。形式相似有助于建立联系，却不足以恢复原来的对象身份。

这里的推导也让本书关于微积分的主线变得更精确。微分可以保留，微分所作用的变量却已改变；局部变化关系可以沿用部分形态，乘积法则中允许的交换步骤却必须重新检查。整体结构也不是把许多瞬时位置简单拼成一条路径，而是需要状态、算符和演化之间保持相容。初始条件究竟应指定什么，因而成为理论构成的一部分，而不是末尾随意补上的数字。

这是一种重要的创造方式：不是废掉所有旧工具，而是改变工具的合法作用对象与组合边界。数学史与物理史在这里相互牵动。若把微积分直接等同于 D，再把矩阵直接等同于 E，就会错失这种变化；同一种数学操作可以在不同发生系统中承担不同任务，同一个发生维度也可能调用多种数学资源。对应关系必须由具体工作建立，不能靠名称一次性分配。

41.6 三方程中的双重约束与第三维改写

把前面的过程放入本书的方法论，首先应区分三个层次。谱线、选择规则与可检验响应构成一定尺度上的显露 S；制备、激发、比较、组合与回检中的差异运行构成 D 的研究对象；能量差、跃迁振幅、相位关联、仪器分辨与可组合条件怎样互相限定，则进入 E 的分析。这个划分只是明确本章问题的入口，不意味着每一项材料从此只能被固定放在一个抽屉里。

三方程 S＝F(D,E)、D＝G(S,E)、E＝H(S,D)在这里是研究约束，不是可以替代薛定谔方程或对易关系的数值算法。第一个问题是：给定怎样的差异运行与特征纠缠，可以生成可重复的谱学显露？第二个问题是：既有显露与纠缠结构要求怎样改变实验和计算的运行程序？第三个问题则是：面对稳定显露与现有程序之间的冲突，必须改写哪些特征共同成立的条件？

六路径则检查三维进入研究的不同完整次序。E—D—S 的沉淀路径，从已有组合约束组织计算与实验序列，沉淀可重复的显露；E—S—D 的锚定路径，先由纠缠规定辨认可保留的结果，再据此安排后续操作；D—E—S 的倒逼路径，从旧计算遇阻出发，迫使组合结构改变，进而形成新的预测。这些次序都经过三维，不能缩成两个字母之间的一支箭头。

其余三路也有不同入口。D—S—E 的原型路径，可以从一组实际计算出发，用得到的初步结果检验怎样的组合结构能够成立；S—E—D 的进入路径，从谱线的既有显露进

入其能量差与振幅约束，再组织可检验的操作；S—D—E 的再创路径，则从异常结果改变差异运行，并据新运行重建相互规定的结构。这里所列是本书对研究任务的组织方式，不声称海森堡曾按这些名称逐条执行。

三原理进一步规定改写的动力：S 与 D 之间的张力推动 E 改变，D 与 E 之间的张力推动 S 改变，S 与 E 之间的张力推动 D 改变。例如，稳定谱线与旧轨道计算不能协调，就要求重建特征组合；新代数与实际计算程序接合不良，就要重新推求能够成立的显露；预测显露与结构约束不一致，则应改变计算近似或实验序列。第三维改变后，必须返回原来两维构成的张力场，检验矛盾是否真正得到处理。

三方程由此提出互相依赖的约束，六路径组织研究展开的完整次序，三原理则辨认两维张力怎样推动第三维改写。它们须联合工作。若只把失败结果重新命名为“新的显露”，却不愿修改任何条件，也不回到原证据检查，方法便失去辨别力。理论创造需要说明改变发生在哪里、为什么发生，以及哪些结果能够证明这次改变仍然不够。

41.7 什么意义上的物理世界被创造了

现在可以正面回答章题。经典轨道之外，仍有能够制备、相互作用、产生稳定效应并接受共同检验的物理对象；但这种对象不能一概依靠一条经典位置曲线取得全部规定。其身份通过允许状态、算符关系、跃迁结构和实验效应的相互限制而成立。“有什么物”与“哪些操作、关系和显露能够共同稳定”因此被更紧密地联系起来。

这不只是把旧电子换一个名称。假如一种理论允许交换两个量而不改变结果，另一种理论却禁止这个步骤，那么两者所许可的计算世界已经不同。假如一种理论要求先给轨道才能谈辐射，另一种理论能在不指定这种轨道的情况下建立严格预测，那么何者是必要构成、何者只是某种表示方式，也发生了变化。创造的实质在这些可工作的差别中，而不在叙述语气有多宏大。

不过，不能由此推断经典世界从历史中消失了。宏观轨道描述在其适用范围内仍可以高度可靠。我们需要解释的是它怎样作为特定尺度、制备条件和近似关系中的稳定形式继续成立。新理论创造更广的规定系统，并不要求所有旧语言立即作废；它要求旧语言说明自己的有效边界。所谓继承，不是把新理论重新塞回旧本体，而是在新结构中重新限定旧结构何以可用。

还应区分 SDE 特征纠缠与量子理论中严格定义的纠缠态。前者在本书中讨论特征如何互相构成和限制，后者涉及复合量子系统状态能否按特定张量分解写成乘积等专门问题。一个位置算符和动量算符不对易，并不自动等于已经展示了两个子系统之间的量子

纠缠。两者可以在进一步研究中相遇，但不能凭同一个汉语词就把不同层次的概念合并。

因此，本章所说的新物理世界，既包括新的理念规定，也包括这些规定进入仪器、预测、操作与重复经验后形成的现实结构。视觉谱线并非被宣布为幻觉，触觉操作也并非被降为无关附属；每种感觉参与的活动都是完整的 SDE 发生，只是它们在当前研究中以不同方式接合。我们不能把视觉配给 S、听觉配给 D、触觉配给 E，借一个整齐的三分表取代发生分析。

41.8 与竞争解释的对话及下一步问题

一种谨慎的实在论解释会说：电子一直存在，海森堡只是找到更好的描述方式。这种解释保留了自然对理论的约束，也防止把科学说成随意建构。SDE 研究若要推进讨论，就不能靠否认实验抵抗取胜，而应追问“电子的哪些规定一直存在”这一说法需要哪些可说明条件。电荷等跨理论延续的规定，与经典轨道这种被撤回的构成要求，不应被混作一个不可分辨的整体。

另一种工具主义解释会说：理论只须预测仪器读数，不必回答物是什么。它提醒我们不要把数学图像过早当成实体，却可能低估预测工具本身如何重组对象身份。若允许什么量、怎样组合量、哪些制备可以比较都改变了，工具已经参与建构一个可工作的物理领域。把这些变化全部归入“仅仅计算”，反而会遮蔽计算能够成立的条件。

本书提出的发生学解释尝试保留两者各自的压力：既要求物理规定进入可重复的抵抗与检验，又要求说明这些规定不是脱离互动而一次性给定。它的弱点也必须公开：如果不能进一步展示哪些历史选择、实验条件和数学限制决定了某一结构的形成，那么“世界被创造”仍可能只是一种概括性表述。只有把概括落到可比较的案例，才能显示它比“观念改变”多解释了什么。

矩阵力学还没有在这一章中解决测量问题，没有独自裁决所有量子解释，也没有证明任何观察者意识引起物质存在的结论。对于谱线之外更复杂的现象，新的结构还必须继续扩展并接受检验。我们停在这里，是为了保留一个真实的历史开放点：轨道不再是唯一入口，但如何让这种抽象组织获得更广泛、可计算、可比较的表达，还需要新的创造。

下一章进入波动力学时，应带着这个问题前行：如果相同的物理约束可以在另一种数学表达中成立，那么究竟什么属于物理世界的稳定规定，什么属于可替换的表达形式？海森堡的贡献不是交给我们一幅更小、更精致的原子画像，而是迫使我们检查画像

之所以被当成“物本身”的条件。由此，物理创造的任务从寻找轨道，推进到重建可检验的规定系统。

本章参考资料

[1] Heisenberg, W. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. Zeitschrift für Physik 33 (1925), 879–893. DOI: 10.1007/BF01328377. 本章核读 D. H. Delphenich 英译，前言及第一节；网页：https://neo-classical-physics.info/uploads/3/4/3/6/34363841/heisenberg_-_qm_interp_of_kin_and_mech.pdf。仅将原文作为简短历史事实依据；余下结构讨论为本书的现代记号分析及发生学解释。

[2] Born, M.; Jordan, P. Zur Quantenmechanik. Zeitschrift für Physik 34 (1925), 858–888. D. H. Delphenich 英译：https://www.neo-classical-physics.info/uploads/3/4/3/6/34363841/born_and_jordan_-_qm_1.pdf。用于矩阵演算系统化、量子条件与能量守恒的历史定位。

第四十二章 薛定谔波动力学
波函数是怎样的“物”

42.1 从看不见的轨道转向不能照搬的波

上一章的矩阵力学已经表明，电子的物理规定不必全部寄托于一条经典轨道。但是，把轨道撤走之后，读者仍可能感到空落：一张张数组怎样成为运动的描述？当薛定谔把波动方程带入原子研究，熟悉的波形似乎又把直观世界带了回来。绳子的驻波、乐器的泛音、水面的干涉，都可以帮助人们想象新的理论。然而，最有帮助的类比，也可能成为最顽固的误解来源。

波函数写成空间和时间的函数，绝不等于它就是某种肉眼无法看见的水波。水波涉及水面位移、流速和密度等物理量，电磁波有电磁场的规定；量子态的波函数则必须由它在整个理论中的职能来辨认。我们不能先把“波”认作一种现成物质，再把量子对象放进这类物质的口袋。要问的是：波函数允许哪些运算，怎样演化，怎样与实验结果连接，以及哪些不同写法实际上代表同一状态。

薛定谔一九二六年的《量子化作为本征值问题》系列工作，把原子允许能量与微分方程的本征值联系起来。[1]这一历史事实已经提示一个根本变化：允许能量可以由整体解的条件产生，不必先给出经典轨道，再从外面附加一套筛选规则。本章不把现代教材的全部解释倒写进最初论文，也不把一套后来成熟的语言当作某个清晨同时发生的成品。

我们将从一个具体疑问进入：同样是求解微分方程，为什么有些解能够成为物理状态，另一些不能？这个疑问把方程、边界、归一化、相位和演化连接起来，也把波函数的身份从一幅静态图像推进为一套可共同工作的规定。按照本书的主轴，新物理世界正是在这种规定系统的形成中被创造，而不是在某个词语被换成“波”时已经完成。

42.2 本征值不是给任意答案贴上的号码

先暂时不讨论波函数究竟代表什么，把一个简单计算完整做出来。设粒子被限制在长度为 L 的一维理想无限深势阱中，阱内势能取零，边界要求波函数在两端为零。在阱内，定态方程写为 $-\hbar^2\psi''(x)/(2m)=W\psi(x)$。这里 W 代表能量，E 仍保留为本书的特征纠缠记号。我们研究的是这一明确理想化模型，而不是声称自然界存在绝对无限高的真实墙壁。

若 W 为正，解可以写成正弦和余弦的组合。左端 ψ(0)＝0 排除余弦项，右端 ψ(L)＝0 进一步要求波数 k 满足 kL＝nπ。非零解由此只在一系列能量上成立：$W_n = n^2\pi^2\hbar^2/(2mL^2)$，n 为正整数。若把 n 取零，正弦解整体变成零，不能表示归一化状态；若取负整数，只是把同一空间形状乘以负号，并没有增加新的能量。

这里的量子化不是在所有连续能量中人为挑选喜欢的数值。微分方程要求局部曲率与能量协调，两端边界又要求同一个解整体接合。两种要求联合施加，非零解才只能在特定能量上成立。这个推导给出了一个清晰的发生结构：局部规律与整体条件共同限定允许状态。只背诵能量公式，会把最重要的构成过程隐藏起来。

同样重要的是，不应由这个例子推出所有量子能量都离散。若把粒子放在整个空间中，改变边界与束缚条件，能谱可以包含连续部分；某些系统则同时具有离散谱和连续谱。离散性不是“量子”一词随身携带的一枚徽章，它必须在给定算符、定义域和条件下被求出来。一个理论有价值，恰在于能够说明何时离散、何时连续，以及两者如何联系。

这也改变了“物”的边界含义。在日常图像里，墙壁只是围住一个已经有完整运动规律的物；在此模型中，边界参与规定允许状态本身。改变 L，能级整体变化，尽管质量 m 和方程内部的形式没有改变。边界并未凭空创造粒子，却进入了该粒子在此系统中能够具有怎样能量的构成条件。这是关系进入物理身份的一个可以复算的例子。

42.3 归一化如何把数学解变成可用的状态表示

刚才得到的正弦解还带有一个任意常数 C。如果 ψ 满足线性齐次方程，Cψ 也满足。因此，单凭微分方程并不能决定这个整体尺度。采用后来成熟的概率解释，必须要求整个允许区域内的概率为一，即 $\int |\psi(x)|^2 dx = 1$。在阱内，这使归一化定态成为 $\psi_n(x) = \sqrt{(2/L)}\sin(n\pi x/L)$。积分区间是零到 L，平方根系数来自正弦平方积分等于 L/2。

归一化并不是把自然界调整成总数为一，而是把状态的表达与概率规则接合起来。它说明一个数学解要参与预测，必须在表达尺度与实验含义之间取得约定明确的对应。若只保留曲线形状，忘记其量纲与归一条件，便无法从一段区域的积分得到正确概率。对一维位置波函数，归一化还要求它具有长度负二分之一次方的量纲；这与水面位移的长度量纲根本不同。

这时也会出现一个看似悖论的问题：自由粒子的平面波在无限空间里不能按普通积分归一到一，难道自由粒子不存在？正确处理并不是丢掉积分条件，而是区分理想化的广义本征态与实际可归一化的波包。平面波可以作为展开工具，有限宽度的动量分布可

以叠加成波包。不同数学对象的使用条件必须说清，不能让一个理想极限冒充已经制备出的正常状态。

归一化还让我们看到，波函数的某些差异并非物理差异。把 ψ 整体乘以一个模为一的常数相位，所有只依赖同一纯态的概率预测保持不变。在通常的纯态表述中，这些向量代表同一条射线。这里并不是说所有相位都无意义；整体相位与不同分量之间的相对相位具有不同地位。先分清什么差异可以被消去，才能知道什么差异必须保留。

从发生学角度看，状态身份因此包含一种筛选：不是数学表达式每出现一个不同符号，就产生一个新的物理世界；也不是所有看起来相似的曲线都可以当作相同状态。我们需要由可执行操作和预测结果来检验哪些表达差异保留下来。这种筛选具有客观约束，却必须通过计算与实验的互动才能具体成立。数学形式与物理身份之间，需要一条可追问的通道。

42.4 同样的密度，不一定是同样的状态

考虑两个归一化波函数，它们在每个位置具有相同的模，但相位随位置变化的方式不同。它们在同一时刻的位置概率密度相同，却可能具有不同的动量分布和后续演化。为避免只有口头描述，可以设 $\psi_2(x)=\exp(ik_0x)\psi_1(x)$。在适当的光滑性和衰减条件下，用 $p=-i\hbar\,\partial/\partial x$ 计算，第二个态的平均动量比第一个增加 $\hbar\,k_0$。

这个差别很容易验证：对乘积求导，导数既作用在原波函数上，也作用在额外相位因子上。后一个贡献正好给出 $\hbar\,k_0$。只画 $|\psi|^2$ 的两张图，看不出这一变化；但让两态演化或测量其动量，结果就能够不同。因而波函数不能被简化为一团“电子有多少”的空间浓淡图。概率密度保留了重要内容，却没有保留全部状态规定。

再考虑两个正交态 ϕ_1 与 ϕ_2 的等权叠加。$\psi_+=(\phi_1+\phi_2)/\sqrt{2}$ 与 $\psi_-=(\phi_1-\phi_2)/\sqrt{2}$，在测量这两个基态对应的投影时都各有一半概率；若改用包含两者叠加的测量基，则可以把它们区分。这个小模型说明，“得到同一组频数”永远需要补问：在哪一种测量安排下？一个测量的相同，并不能保证所有测量的相同。

这也是本书研究“显露”的一条纪律。显露不是没有条件的总画像，而是在具体差异运行与特征纠缠中取得规定的结果。如果两个态在一次观察中相同，不能立即把它们在所有可能操作中的差别删除。反过来，如果一个相位变化在给定理论允许的全部操作中都没有可辨结果，也不应靠想象为它制造物理差别。可辨与不可辨都需要论证。

这里尚不必把不同量子解释提前裁决。把波函数理解为实在的状态、预测结构或有关制备的信息，各有额外承诺；但它们都不能忽略相对相位带来的可检验差别。本章要

保住的是这一共同约束。发生学分析的力量不来自宣布某个词已经终结争论，而来自把争论者都必须面对的具体关系列清，让不同解释在相同压力下展示自己的解释能力。

42.5 演化是怎样保留整体规范的

含时薛定谔方程采用现代记号写为 $i\hbar\,\partial\psi/\partial t=H\psi$。对适当的自伴哈密顿算符，它把初态与后续状态关联起来，并在通常封闭系统的描述中保持态的范数。这里的“确定演化”指给定哈密顿算符、初态和相应数学条件后，态演化被确定；它不等于每次测量都只有一个预先给定的结果，也不意味着实际实验永远与环境隔绝。

范数保持可以通过求导看出。对 $\langle\psi|\psi\rangle$ 求时间导数，会出现一项由态矢导数贡献，另一项由其共轭导数贡献。当 H 自伴且有关运算成立时，这两项相消。因此归一化不是每隔一会儿由研究者手动重新涂改的数值，它与演化结构相容。若数值程序使封闭系统的总范数不断漂移，首先应检查离散算法和精度，而不是马上宣布概率守恒已经被自然推翻。

对无显式时间依赖的 H，若 ϕ_n 是能量本征态，时间因子为 $\exp(-iW_nt/\hbar)$。于是 $\psi_n(x,t)=\phi_n(x)\exp(-iW_nt/\hbar)$，其位置密度不随时间改变。所谓定态，不表示波函数每个复数值都不变，而表示这类态在演化中只获得整体相位，相关静态概率分布保持。把“定态”误写成“电子完全静止”，会重新把波函数当作经典轨道照片。

若初态包含多个不同能量成分，各成分获得不同时间相位，相对相位便随时间改变；某些观测量的期望和分布也会随之变化。定态与叠加态因此提供两种不同的时间结构。变化不必来自某个小球沿着屏幕上的浓淡区域跑动，它可以由相对相位的演化产生。我们看见的是可预测的显露变化，解释必须忠于产生这些变化的数学关系。

对一维实势能模型，还可从方程推出连续性方程 $\partial\rho/\partial t+\partial j/\partial x=0$，其中 $\rho=|\psi|^2$，$j=(\hbar/m)\mathrm{Im}(\psi^*\partial\psi/\partial x)$。它使密度变化与概率流相互协调。这一关系不是液体模型的充分证明；同一种连续性形式可以在不同理论里表示不同守恒内容。数学相似告诉我们可以借用某些分析工具，不替我们决定流动的究竟是哪一种物。

42.6 一个粒子的三维空间与许多粒子的配置空间

只看单粒子问题，人们很容易把波函数想成普通三维空间里的一种场。对于许多粒子，事情立即复杂起来。两个无自旋粒子的波函数一般依赖两个位置向量，即 $\psi(r_1,r_2,t)$；更多粒子则需要更多坐标。这里的配置空间不是说日常空间突然增加了人可以走进去的方向，而是完整配置需要由多组坐标共同标识。

如果两个粒子的联合态可写为 $\phi(r_1)\chi(r_2)$，它们具有一种特殊的可分结构；一般联合态却不能作这样的乘积分解。此时，仅给两个粒子各自一张普通三维浓淡图，并不能重建全部联合预测。关联结构丢失在哪里，可以通过联合分布与不同测量安排的比较来追查。不能因为难画，就把高维表达当作纯粹计算多余物；也不能因为它难画，就宣布高维空间必定是唯一真实空间。

一个有限维例子更加直观。把两个两态系统的联合态取为$(|00\rangle+|11\rangle)/\sqrt{2}$，在零一基上，各系统单独看都是一半零、一半一；联合看却只得到零零或一一。单独概率没有告诉我们这种对应关系。若考虑另一种混合制备，也可能在这一基上得到相同联合频数，但换测量基后显示不同结果。因此，连一张联合概率表也未必包含完整量子态结构。

这个例子为本书的特征纠缠讨论提供启发，但仍须保持术语边界。量子纠缠有依赖张量积结构的精确定义；SDE 的特征纠缠研究更广的规定怎样互相构成。我们可以研究量子态的特征怎样进入实验身份，却不能宣称两者仅凭名称便相互证明。跨层次解释需要指出具体映射，也需要允许映射失败。

波函数究竟是怎样的“物”，到这里已经不能靠一幅三维波峰图回答。它至少是一种受到归一化、相位等价、组合规则与演化要求共同约束的状态表示，并通过概率规则进入实验预测。如果提出更强的本体论主张，就必须说明它如何处理多体结构、表示变换和测量结果。形象的亲切程度不能代替这些论证。

42.7 矩阵与波函数为什么能够相遇

设某一物理量由算符 A 表示，在一组正交归一基$\{\phi_n\}$中，它的矩阵元为 $A_{nm}=\langle\phi_n|A|\phi_m\rangle$。同一个态也可以写成这组基上的系数展开。这样，波函数中的算符作用便可转为系数上的矩阵作用。这个关系说明上一章与本章并不只是两种互相排斥的直观图像：在适当数学条件下，它们可以表达相同的状态与观测量结构。

需要注意这里的限定。基是否完备，算符定义在哪个空间，连续谱如何处理，都是实际工作的一部分。若只取有限多个基函数，得到的是截断近似，不能因此声称原来无限维问题的一切关系均被精确保留。特别是上一章指出的正则对易关系有限维障碍，会在这种处理中再次提醒我们：近似可用与数学等同不是同一判断。

为了看清表示变换，可以设换到另一组正交基，态系数和算符矩阵一起作相应变换。如果只变换态而把测量算符留在原来的坐标说明中，预测可能被错误地改变；若两

者协调变换，物理概率保持。这就像变换坐标时必须同时调整对象分量与计算规则，但量子状态空间的结构又不能完全被普通空间旋转比喻替代。

从思想史角度看，这种相遇具有双重意义。它使波动力学提供的微分方程工具与矩阵力学的代数结构互相支持，也防止我们把某一种写法的直观外观直接等同于全部物理实在。同一规定可以拥有不同表达，并不意味着规定是空洞的；恰恰是那些跨表达保持的关系，让理论能够从一张纸上的计算变成共同体可以接续的工作对象。

本书所谓创造物理世界，并不要求每换一次表示便宣布世界重新诞生。表示变换若保留全部相关预测，属于同一结构的不同展开；若改变了允许状态、操作范围或可检验关系，则需要进一步判断是否形成新的规定系统。把这两个层级分开，才能避免让“创造”一词失去区分能力。

42.8 被经典禁区排除的解为什么还能延伸

有限势垒还提供一个具体机会，让读者看到波函数怎样改变物理可行性的边界。设某一区域的势能为 V_0，而定态能量 W 小于 V_0。在经典粒子图像中，如果只含这种势能和动能，动能 $W-V_0$ 为负，因而该区域不允许经典轨道进入。波动方程却并不因此要求该区域的解恒为零，而给出指数增长与指数衰减的线性组合，特征尺度由 $\kappa=\sqrt{[2m(V_0-W)]}/\hbar$ 确定。

若势垒具有有限厚度，必须把左侧、垒内和右侧的解接续。对于有限且非奇异的势能跃变以及常质量模型，波函数及其一阶导数保持相应连续。满足这些条件的散射解可以在右侧留下非零透射流。对足够厚的势垒，透射率的主要厚度依赖含有 $\exp(-2\kappa a)$ 这样的衰减因子；完整系数还取决于两侧与垒内的匹配，不能把这个近似因子当成任意势垒的精确答案。

这里没有必要说粒子临时借来能量，穿过去再还回去。对静态势垒的定态散射，总能量在方程里一直是同一个 W。真正改变的是可允许状态的结构：经典轨道判定禁入的区域，并不等同于波函数必须为零的区域。把隧穿说成暂时违反能量守恒，会把新理论的规则误写成旧理论之外的一次作弊，反而掩盖创造发生在哪里。

这个例子也说明，波函数有多大与粒子被切成多少块不是同一个问题。透射概率和反射概率分别联系不同实验结果，概率流的守恒约束两者。在相同两侧势能和适当归一条件下，两者之和为一；若两侧速度不同，透射率还要包含相应的流速因子。正确比较需要概率流，而不能只看透射波振幅的平方就草率下结论。

42.9 微分方程的力量与条件的回归

波动力学使微积分再次处于物理创造的核心，但它同时拆掉了“有一个方程就有完整物理”的幻觉。方程需要变量的身份，解需要定义域与边界，状态需要归一化与相位规则，预测需要观测量与实验对应。这些不是在主要工作结束后随便补上的说明，而是让数学表达拥有物理效力的条件。

以无限深势阱为例，同一个阱内微分表达式，配不同的边界条件，可能对应不同的谱和不同的物理设置。研究者不能为了算出喜欢的答案，任意选择边界；必须说明边界怎样由实际约束的理想化取得。对有限势垒，隧穿和波函数延伸使简单的“墙边为零”近似发生变化。这时修正不是在原图外添加一个神奇例外，而是恢复先前被理想化省略的结构。

再考虑数值求解。网格间距过大可能抹去快速变化，区域截断过小可能把波包错误反射回来，时间步长不合适可能破坏稳定性。这些失败首先属于计算过程与目标问题不匹配。它们也有发生学意义，因为实际知识依赖这些过程；但不能把数值发散直接改名为自然界处于混沌，更不能据此推断普朗克常数本身在变化。

严格的研究必须做收敛检查：细化网格、扩大区域、改变算法以后，所关心的结果是否趋于稳定；同时还要做独立检验，确认收敛到的并不是一个错误模型的答案。数值稳定不等于物理正确，物理模型合理也不保证每一种程序都正确。这些差别构成一种分层判断能力，使波函数研究能够在复杂计算中继续保持可反驳性。

因而，新的方程并没有终止发生过程。它为新的制备、谱学、化学与材料研究开放空间，而每一种应用又会把新的边界与相互作用带进来。物理世界的扩大，不是把所有对象塞进同一条公式就告结束，而是让公式与具体对象之间建立越来越明确、越来越可以回检的连接。

42.10 改变边界之后，三维怎样互相改写

让我们把方法论放回一个尚未解决的实验问题：同样的阱宽扫描，实测能级没有完全遵循无限深势阱的平方反比关系。首先不能因不符合公式就宣布波动力学失败，也不能因为公式著名就把偏差一律归给仪器。需要比较至少三个候选：有效阱宽的标定有误，势垒并非无限高，或者单粒子近似遗漏了相互作用。它们要求不同的追加测量，能够在不同条件下给出不同的偏差形式。

采用 S—D—E 再创路径，研究由异常能谱 S 进入操作 D 的改变：重新标定几何尺寸，改变势垒高度，并在不同制备条件下重复测量。若宽度标定修正后异常消失，就不

必为此重建全部理论；若偏差随势垒高度系统变化，则应把有限势垒纳入特征纠缠 E。这个新结构不只是多写一个参数，它改变了波函数在边界附近的接续，继而改变允许能量。

这正体现 S 与 D 的张力推动 E 改变：显露稳定，而原操作与原模型的连接不能解释它，第三维就必须受审。随后沿 E—D—S 沉淀路径，由有限势垒结构组织新的求解和实验程序，产生可事前检验的谱线预测，再与原异常比较。若只在已经知道的数据上调节参数而无新检验，循环尚未完成。回检也可能否定刚刚提出的新结构，使研究再次改道。

三方程在这一过程里各有实际问题：S＝F(D,E)问新结构如何经实际制备显露；D＝G(S,E)问已有能谱与候选结构如何决定下一轮扫描；E＝H(S,D)问哪些边界或耦合关系必须重建。另两种张力同样可能接管工作：操作与候选结构不能相容时，要重新推求可成立的显露；显露与候选结构仍冲突时，要改进区分它们的操作。这些步骤不由方法名称保证成功，而由最终预测是否经受新证据决定。

42.11 波函数问题留下的开放口

波动力学让我们看到，一种物理对象的身份可以通过解空间、允许操作与可检验结果取得稳定，而不必由日常形状先行保证。波函数既不是任意符号，也不能因为画成曲线就被认作普通流体。它的物理意义由一组相互约束的规则支撑，这些规则又必须在实际制备与测量中获得持续检验。

对实在论者而言，这可以意味着我们找到了更准确的实在结构；对工具主义者而言，这可以意味着我们拥有更强的预测工具。本书的发生学追问则是：什么新的特征组合和差异运行，使这一结构或工具能够成为实际物理世界的一部分？这一追问不以否认自然约束为代价，也不满足于把所有改变缩为观察者头脑中的印象。

本章从本征值走到归一化，从相位走到演化，再到配置空间与表示变换。每一步都让“波函数是什么”变得更加具体，同时限制了过早的本体断言。我们仍未回答某次测量为何出现这一结果而不是另一结果，也尚未解释概率规则怎样与单次事件相接。接下来，玻恩的统计解释与不确定性关系，将把这个尚未封闭的问题推向新的层面。

本章参考资料

[1] Schrödinger, E. Quantisierung als Eigenwertproblem. Annalen der Physik 384 (1926):361–376. DOI:10.1002/andp.19263840404。出版方记录：https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19263840404。核验范围为题名、年份、页码与原论文系列书目信息。

[2] Schrödinger, E. The Fundamental Idea of Wave Mechanics. Nobel Lecture, 12 December 1933. 官方入口：https://www.nobelprize.org/prizes/physics/1933/schrodinger/lecture/。仅据题名及历史定位列入；正文数学论证为本书自行展开的现代记号分析。

[3] University of Zurich, Erwin Schrödinger—Nobel Prize in Physics 1933. https://www.uzh.ch/en/researchinnovation/excellence/nobelprize/schroedinger.html。用于 1926 年首篇工作与能级计算的机构历史核验。

第四十三章 玻恩与不确定性
概率、分布与可观测量

43.1 波函数给出了答案，为什么还要问结果在哪里

波动力学可以算出能量、演化与干涉结构，却不会因此自动告诉我们每次实验将出现哪一个记录。屏幕上一次局域的点击，与整个空间中的波函数不是同一种东西。若把波函数画成浓淡连续的云，便要回答为什么探测记录往往表现为一个个事件；若把它当作尚未看清的小粒子轨迹，又要说明相位与干涉从何而来。这里的困难不只是缺少一个漂亮比喻，而是理论表示与实验事件之间尚需明确规则。

玻恩在一九二六年的散射研究中推进了波函数的统计解释；他在一九五四年的诺贝尔演讲中回顾了用振幅模平方联系散射方向概率的思路。[1]本章采用成熟的玻恩规则来讨论它的结构，不把后来的一般测量形式都归为最初论文已经完成的内容。历史上的突破与后来的数学完善，需要在叙事中彼此联系，也需要保留时间差别。

概率进入这里，并不表示物理学退回了随便猜测。恰恰相反，一组不能逐次预测的事件，可以呈现非常严格的分布限制；不同测量分布之间，还可能受到同一状态结构的共同约束。不确定性关系所规定的，正是某些分布宽度不能任意同时缩小。它的力量来自精确条件，而不是来自“世界什么都说不准”的感叹。

本章要分开四类常被混用的问题：状态本身给出的结果分布，仪器的误差与扰动，有限样本带来的统计误差，以及数学模型或计算程序尚未收敛。只有把它们分开，才谈得上研究它们怎样在实际实验中纠缠。把所有不稳定一概叫作量子不确定性，会同时损害物理学和发生学的解释能力。

43.2 概率规则究竟连接了什么

在最简单的离散情形中，设 A 的某个结果 a 对应投影 P_a，归一化态为 ψ，则得到该结果的概率为 $p(a)=\langle\psi|P_a|\psi\rangle$。如果本征值非简并，且对应归一化本征态为 a，这可写成 $|\langle a|\psi\rangle|^2$。位置表象中的区域概率则是对该区域积分 $|\psi(x)|^2$。公式连接的是给定制备状态、给定测量安排与结果类别，而不只是一个孤立符号与一串数字。

对混合态，规则推广为 $p(a)=\mathrm{tr}(\rho P_a)$，其中 ρ 是密度算符。更一般的测量可以用正算符组成的效应集合表示，概率写为 $\mathrm{tr}(\rho M_a)$，各效应之和为单位算符。这里引入这一现代形式，是为了防止读者把所有真实测量都当成理想投影；本章并不需要展开一般测量理论的全部技术。

归一化、非负性与总概率为一是最基本的检查。但它们还不足以决定量子概率的完整形式。许多普通概率模型也满足这些要求；量子理论特别之处在于，概率必须与态空间、叠加、可观测量和复合系统结构共同协调。若只说“加起来等于一，所以自然必定服从玻恩规则”，论证便跳过了最关键的结构条件。

考虑一个两态系统，$\psi = \sqrt{p}|0\rangle + \exp(i\phi)\sqrt{(1-p)}|1\rangle$。在零一基中测量，两个结果的概率为 p 和 $1-p$，相位 ϕ 不出现。但换到叠加基，ϕ 一般会影响结果。概率不是在波函数外面另贴的一张独立标签，它继承了状态与测量的相互规定。一个条件下不显露的差异，可能在另一个条件下变成能够辨认的差异。

这也说明本书为什么坚持把分布当作物理显露的一种组织方式。单次结果不是全部，频数曲线也不是脱离制备与测量就独立成立的对象。结果如何分箱、哪些试验算同一制备、漏检怎样处理，都参与概率与经验的接合。承认这些构成条件，并不是把概率交给研究者任意决定；相反，它使错误的接合方式可以被指出、修正和重复检查。

43.3 两条可能路径不能总像两只袋子一样相加

设到达某个探测结果的两个相干贡献分别为复振幅 a 和 b，总概率与 $|a+b|^2$ 成正比。展开后得到 $|a|^2+|b|^2+2\mathrm{Re}(a^*b)$。最后一项取决于相对相位，可以增加概率，也可以减少概率。若把两个贡献当作经典互斥事件的概率直接相加，就删掉了这一项。干涉不是在算完普通概率以后附赠的小修饰，而是组合规则带来的可检验差别。

可以取一个纯数学例子：$a=1/2$，$b=\exp(i\phi)/2$，则所考察输出的概率为 $(1+\cos\phi)/2$。ϕ 从零变化到 π，这个概率从一变成零。在一个合适归一化的两输出干涉模型中，另一个输出的概率作互补变化，总概率保持一。变化发生在结果如何分布，而不是总概率突然凭空增减。

若某种相互作用把两条贡献与可区分的辅助状态关联起来，干涉项会受到辅助态内积的影响。辅助状态完全正交时，忽略辅助系统后，相关干涉项可以消失。这里“可区分”是物理态之间的关系，不要求某个人必须坐在那里知道答案。一个无人查看的记录装置，也可能通过相互作用改变联合态的结构。

不过，干涉消失本身并不自动解决单次结果为何如此出现的全部问题。它解释某些相干项如何不再出现在子系统统计中，却不能在不加额外解释的情况下，直接等同于已经证明某一种量子本体论。我们的任务是保留它真正解释的部分，也指出它尚未完成的部分。把未解决的问题藏进“观察者知道了”这句话，反而会让理论失去可检验的边界。

这一例子把概率发生的结构说得十分具体：同样两个候选贡献，能否相加为振幅、是否留下干涉，依赖它们与其他系统的关系。概率分布由此不能被看成物体内部独立携带的一份名单。它通过制备、相干保持、耦合与探测之间的组织而成立，同时又以稳定的统计规律约束这些操作。

43.4 分布的宽度与平均值的不确定是两回事

假设在完全相同的制备与理想测量条件下，一个二值结果出现的概率为 p。独立重复 N 次，观察到该结果的次数 K，其期望为 Np，方差为 $Np(1-p)$。样本比例 K/N 的标准差则为 $\sqrt{[p(1-p)/N]}$。随着 N 增大，我们对概率的估计可以更精确，但单次结果的随机性并不因此按同样比例缩小。

这个区别对于理解量子统计十分重要。把重复次数扩大一万倍，某个平均量的标准误差通常能够显著减小；这并不意味着同一制备态的位置分布也被压缩一百倍。前者描述估计量在重复抽样中的波动，后者描述单次测量结果的分布。它们可以使用相似的平方根符号，却不是同一种物理问题。

例如，一个制备态的位置标准差为 σ。测量许多独立同样制备的系统，位置样本均值的标准误差可按 $\sigma/\sqrt{N}$ 缩小。没有任何位置—动量不确定关系禁止我们通过更多样本精确估计这个平均位置。真正不能任意缩小的是同一态中相应位置与动量结果分布的标准差乘积，而不是实验报告中所有写作“误差”的数字。

独立同分布假设也不能被忘掉。若源随时间漂移，连续事件有关联，探测效率变化，或者数据筛选依赖结果，简单的平方根规律可能失效。此时要研究新的统计模型和操作条件，而不能把异常一律归入自然的不可预测性。量子理论允许统计规律精确成立，并不保证研究者的每一次数据处理都已满足它的条件。

由此可以看到，统计稳定并不是没有过程的天赋事实。实验要把不同试次组织成可比较的集合，并持续检查哪些条件确实保持。SDE 分析在这里能够追问可重复制备怎样发生、样本同一性怎样维持，而不是把“统计”仅仅当作事后画图。只有这种同一性经受检验，频数才有资格进入理论比较。

还可以进一步区分同一量的两种增宽机制。设制备装置有时产生一种态，有时产生另一种态，我们却没有把两类试次分开。第一类的结果均值为 μ_1、方差为 v_1，第二类为 μ_2、方差为 v_2，出现权重为 w 与 $1-w$。混合后的总方差等于 $wv_1+(1-w)v_2+w(1-w)(\mu_1-\mu_2)^2$。前两项是各类内部的分布，最后一项来自两类均值差别。

因此，观察到宽分布，不能立刻说每一次制备都产生了同样宽的量子态。它可能部分来自制备标签被遗漏。若独立记录装置状态并按事先规定分类，就可能分离这两种贡献；若分类规则事后专门挑选数据，又可能制造虚假的窄分布。区分物理制备与统计筛选，是概率解释能够接受检验的必要条件。

这个全方差算例也提醒我们，“内禀统计”在此首先是给定态与给定测量规则所赋予的分布，并非用一个术语已经裁决了随机性的最终形而上学地位。纯态与混合态、条件分布与总体分布可以有不同解释任务，不能因为都画成直方图，就把它们当作同一种发生过程。

43.5 不确定性不等式怎样从状态结构推出

现在讨论罗伯逊形式的不确定关系。其一九二九年论文的书目与发表时间可由原刊记录核实。[2]下面给出本章自己的现代推导，以免把一个名称当作证明。设 A 和 B 为自伴可观测量，态已归一化，并满足所需定义域与有限方差条件。定义中心化算符 $\delta A=A-\langle A\rangle$，$\delta B=B-\langle B\rangle$，标准差分别为 ΔA 和 ΔB。

令 $u=\delta A\psi$，$v=\delta B\psi$。柯西—施瓦茨不等式给出 $\|u\|^2\|v\|^2\geq|\langle u,v\rangle|^2$。左边就是 $(\Delta A)^2(\Delta B)^2$；右边至少不小于其虚部的平方。由于 $\langle[A,B]\rangle$ 等于相应复数虚部的两倍乘以 i，便得到 $\Delta A\Delta B\geq|\langle[A,B]\rangle|/2$。每一步都可以检查，没有用到一台仪器先撞了粒子一下的假设。

对正则位置与动量，$[x,p]=i\hbar I$，于是得到 $\Delta x\Delta p\geq\hbar/2$。在这里，$\Delta x$ 指同一制备态的位置结果分布标准差，Δp 指同一制备态的动量结果分布标准差。可以分别用两组同样制备的系统测量它们，并不要求在同一个粒子上先后读取两个数。因此，这个推导本身是一项制备分布的限制，而不是完整的测量误差—扰动理论。

定义域条件不是挑剔的附注。有些态的某阶矩不存在，有些算符乘积不能对所选态合法作用，此时直接代入形式公式可能没有意义。不能因为无限大方差或未定义表达式出现，就宣布理论被一个代数例子推翻。首先应核对论证前提是否满足，再问是否需要换用其他衡量宽度的指标。

若保留柯西—施瓦茨推导中的实部，还能得到包含对称协方差项的更强关系。这说明通常写出的下界并不总囊括全部状态约束。一个简短公式可以揭示重要限制，却不应被当作所有不确定性的总目录。正因为有严格推导，我们才知道它说明什么，也知道它没有说明什么。

43.6 一个高斯波包把抽象下界变成可算结果

取一个一维归一化波包，其位置概率密度是中心在 x_0、方差为 σ^2 的高斯分布。可以把波函数写为归一化系数乘以 $\exp[-(x-x_0)^2/(4\sigma^2)]$，再乘以 $\exp(ip_0x/\hbar)$。前一个因子决定空间包络，后一个因子决定平均动量。对这个状态，$\Delta x=\sigma$；用动量算符求期望与方差，可得 $\Delta p=\hbar/(2\sigma)$。乘积恰好等于 $\hbar/2$。

这个计算给出了直观但严格的权衡。减小 σ，位置分布更窄，动量分布更宽；增加 σ，则相反。它不意味着仪器粗糙到无法看清，也不意味着每个粒子只是被我们的思想模糊了。它说明在所用态空间与算符关系下，这一族合法状态如何分配两种结果的分布宽度。

接着给同一空间包络加入二次相位 $\exp[i\chi(x-x_0)^2]$，其中 χ 为实数，量纲为长度的负二次方。位置概率密度完全不变，但相位梯度随位置变化，动量方差因此增加。直接计算给出 $(\Delta p)^2=\hbar^2/(4\sigma^2)+4\hbar^2\chi^2\sigma^2$。这里的位置—动量协方差也不再为零。于是，同样的空间浓淡并不保证同样的动量宽度，上一章关于相位的讨论在此获得定量延伸。

这个例子还防止一种误读：不确定关系给出下界，不是断言每个状态都恰好达到下界。许多态的乘积远大于 $\hbar/2$；若测得大于下界，并不意味着发现了新的量子原理。研究者要判断差额来自态本身的结构、额外制备混合、仪器分辨，还是其他条件，而不是把所有宽度一律称作最小量子噪声。

若让最初的自由高斯波包演化，在无初始位置—动量关联的条件下，它的空间宽度随时间增大：$\sigma_t^2=\sigma^2+(\hbar t/2m\sigma)^2$。动量分布在自由演化中保持，而位置分布展开。这个变化来自已给定的动力学，不能被说成常数 $\hbar$ 忽然变得不稳定。它展示的恰是同一常数和同一演化规律怎样生成随时间改变的分布。

43.7 非交换并不保证每个态都有同样的正下界

位置与动量的正则对易子是单位算符的常数倍，因而下界不依赖态。对于一般可观测量，情况不同。罗伯逊关系右侧含有对易子的期望值，这个期望可能随态变化，甚至为零。因此，“两个算符不对易”不能被直接翻译为“任何态中两者都有同一个严格正的标准差乘积下界”。

以两态系统的泡利算符 σ_x 和 σ_y 为例，它们的对易子为 $2i\sigma\_z$。如果系统处于 σ_z 的正本征态，$\Delta\sigma_x$ 与 $\Delta\sigma_y$ 都为一，下界也为一，关系取等号。若系统处于 σ_x 的本征态，$\Delta\sigma_x$ 为零，$\Delta\sigma_y$ 为一，而 $\langle\sigma\_z\rangle$ 为零，因此不等式右边也为零。算符仍然不对易，关系却没有给出严格正的乘积下界。

这不是漏洞，而是公式忠实保留了态依赖性。某一测量在某个态中可以完全确定，并不表示所有其他测量也同时确定；某个乘积下界退化为零，也不表示我们已经获得一套无条件的经典共同属性表。要判断还有哪些约束，需要查看更完整的状态结构或采用其他不确定关系。

这个例子在思想史写作中尤其有价值，因为它阻止我们从一个著名公式跳到过大的哲学结论。不确定性并非一张把所有对象都盖成模糊的印章。它具体依赖所问的量、准备的态与采用的衡量方法。科学创造的深度不在于把所有问题说得神秘，而在于创造能够分别说清这些差异的语言和操作。

同样，不能从位置—动量关系直接推出时间和能量在所有情境中拥有完全相同的算符地位。各种时间—能量关系涉及不同定义与条件，需另行论证。把 Δ 符号摆在任何两个名词前，再写一个普朗克常数，就不是物理推导。本章保留这一边界，是为了让后续讨论动力学时间尺度时仍有清楚的概念起点。

43.8 仪器误差、测量扰动和态的分布必须分层

一台理想测量装置对同样制备态得到分散结果，并不自动说明它不准确。准确性要比较它实际实现的测量与目标测量之间的偏差；分布宽度则可能已经由输入态给出。反过来，一台仪器每次都输出同一个数，也不因此准确。若它根本不响应输入，只是卡在某个刻度上，读数虽稳定，测量却可能完全失效。

测量扰动又是另一问题：某个测量相互作用如何改变系统后续的统计或状态。装置可以具有很好的某项分辨能力，同时显著改变另一项量的分布；也可能在特定状态与任务中减小某些扰动。讨论这些现象，需要具体测量模型和误差定义，不能仅用制备标准差替代所有装置量。

小泽正直关于一般测量的研究明确区分了这两类问题，并指出简单的误差—扰动乘积下界不是无条件适用于所有测量装置的关系。[3]这里只借此确立一个概念边界，不把后来的技术讨论当作海森堡一九二七年已经完整说出的结论。对某些常用的均方根误差与扰动定义，需要包含额外项的关系；其他操作性误差定义也有各自条件。

一个经典化的校准例子能帮助辨认层次。设目标位置结果 X 具有方差 σ^2，而仪器附加独立、零均值噪声 N，方差为 r^2，读数 $Y=X+N$。此时 $\mathrm{Var}(Y)=\sigma^2+r^2$。改进仪器能减小 r，却不必改变输入态的 σ。如果噪声与输入相关，还会多出协方差项，简单相加便不成立。这个模型不是全部量子测量理论，但足以证明“读数宽度等于态宽度”需要额外假设。

实际实验中还可能同时有分辨极限、效率损失、背景事件和选择偏差。纠正这些效应需要独立标定，不能用理论预期去强行把数据修成想要的形状。否则，模型与证据之间的回检就被改造成自我确认。统计解释赋予物理学严格的新能力，也要求对数据产生的过程承担更细的责任。

43.9 不收敛不等于常数不稳定

假如一个数值程序计算波包演化，网格改变后结果明显不同，我们首先遇到的是数值收敛问题。假如计算已经收敛，却不能描述实验，可能遇到的是模型误差。假如实验重复时分布中心缓慢移动，可能需要检查源漂移、仪器温度或外场变化。只有在相关普通解释受到充分检验后，才有理由提出更深层的物理修改。

这几种变化不等于物理常数发生变化。$\hbar$ 出现在不确定关系中，表明理论的量子尺度；态的宽度变化、样本频数起伏或算法不稳定，并不能独立证明 $\hbar$ 不再固定。若要研究常数是否变化，需要明确可比较的无量纲量、实验过程和替代模型。把一个估计值随样本变化的现象，写成常数本身漂移，是把估计过程与所估对象混为一谈。

可以设计一个简单检验链。先固定同一数值问题，比较不同网格与独立算法，判断计算是否可靠；再改变实验中的某项可控条件，检查预测变化的方向和规模；最后引入未用于拟合的数据，比较竞争模型。任何一步失败，都提供了进一步研究的入口，但失败发生的层级不同，后续行动也应不同。

这里的 SDE 意义不是把一切差错都称作新发生，而是追踪差错怎样进入显露。代码离散、探测响应与态制备可能共同影响一条曲线，但它们并不因此成为可随意替换的原因。只有改变某环节并观察相关结果，才能判断它在当前特征纠缠中承担什么作用。发生学需要因果辨别与模型比较，不能用一个包容一切的“整体”取消分辨工作。

这一纪律还保护了真正的创造。若任何异常都立即被宣布为基础常数失稳，研究会失去优先次序；若任何异常都被旧理论一票否决，又可能错过新的结构。分层检查既为已有理论保留合理的检验机会，也为新理论留下明确的进入条件。大胆与严格并不互相排斥，前者需要后者提供可前进的道路。

43.10 从一条异常分布走通发生路径

设一组研究者准备接近最小不确定性的波包，却发现动量分布总比预期更宽。这个显露 S 不能单凭一张直方图说明原因。候选解释至少有仪器分辨不足、制备过程引入随机动量偏移，以及态本身存在二次相位。三者都可能增宽动量分布，却对位置—动量关联和时间演化给出不同预言。

沿 S—E—D 进入路径，可以先根据宽度异常追查特征之间的约束 E：如果是二次相位，位置与动量应存在相应关联；如果是独立的仪器噪声，关联结构可能不同。随后组织差异运行 D，增加不同演化时间的测量、独立校准以及可控相位调节。路径在这里具有实质内容，因为它说明下一次实验为什么这样做，而不是只把三个字母排成装饰。

若新增操作显示宽度主要来自仪器，便应修正测量模型；若校准后仍保留与二次相位一致的关联，则可以重建态的规定，并作新的演化预测。S 与 E 的张力推动 D 改变，改变后的操作又回到原宽度异常检查；若 S 与 D 长期不能协调，则推动 E 重新规定噪声或相位的构成。第三维改写必须产生能够再次失败的预言。

三方程在这一案例中不是新的动力学公式。S＝F(D,E)要求把制备、演化与仪器共同连到分布；D＝G(S,E)要求由候选结构与异常决定区分程序；E＝H(S,D)要求根据实际分布与操作结果重建结构。另一条可用的 D—E—S 倒逼路径，则从主动施加相位的操作开始，改变相关结构，再预测分布如何响应。两条路径相互校验，比反复宣称量子世界神秘更有研究价值。

本书所谓物理世界创造，在这里指一种新的可分辨系统成立：过去被同一个“测不准”覆盖的现象，被拆成态分布、相关结构和装置过程，并能够重新组合成更可靠的预测。它不需要人的意识制造随机性，也不需要否认实验对象的抵抗。创造发生在新的规定与新的操作共同获得稳定的地方。

43.11 概率没有结束因果问题，而是改变了问题的形式

玻恩规则把态与结果概率连接起来，不确定关系则限制不同分布能够怎样共同成立。两者联合之后，物理学可以精确预言统计结构，而不必先承诺每一次结果都有一条可恢复的经典轨道。这是一种解释能力的增长，也是一种问题形式的改变：因果研究可以追问制备与相互作用怎样改变分布，而不只追问某个点下一刻必到哪里。

但这不意味着所有哲学争论都已经结束。概率究竟表示根本随机、关于系统的信息，还是更深理论中的某种统计结构，不同解释仍有不同回答。仅凭本章的不等式，不能证明所有隐变量方案都不可能；也不能证明观察者意识必不可少。更强的结论需要额外假设与独立结果，不能从一个著名名称中自动领取。

对本书而言，最有生产力的方向，是研究何种差异序列与特征纠缠使概率分布取得稳定身份，又在何种改写下产生新的分布关系。单次事件、总体统计、仪器响应与数学表示由此形成一组需要持续协调的层级。协调不是把层级混成一团，而是让它们的连接能够被检验、被反驳并被重建。

下一步进入测量与互补性的讨论时，我们应当带着这些已经分清的问题：什么量在何种安排下获得规定，测量怎样改变后续可能性，哪些统计可以共同进入同一解释。量子理论真正提出的困难，比“人一看世界就改变”更具体，也更深刻；只有保持这些具体困难，发生学才有机会提出有内容的新问题。

本章参考资料

[1] Born, M. The Statistical Interpretation of Quantum Mechanics. Nobel Lecture, 11 December 1954. https://www.nobelprize.org/uploads/2018/06/born-lecture.pdf。阅读范围为散射与概率回顾的官方节选；历史摘要限于短段，后续分析为本书现代模型推导。

[2] Robertson, H. P. The Uncertainty Principle. Physical Review 34 (1929):163–164. DOI:10.1103/PhysRev.34.163。https://journals.aps.org/pr/abstract/10.1103/PhysRev.34.163。原刊书目与 1929 年 7 月 1 日发表记录核实，正文需订阅；本文自主给出标准推导，不声称逐句转述原文。

[3] Ozawa, M. Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement. Physical Review A 67 (2003):042105. DOI:10.1103/PhysRevA.67.042105。https://journals.aps.org/pra/abstract/10.1103/PhysRevA.67.042105；作者预印本：https://arxiv.org/abs/quant-ph/0207121。仅采用原论文摘要所支持的概念区分与适用边界，不将所有误差定义视为等价。

第四十四章 玻尔、爱因斯坦与量子测量装置参与了什么

44.1 测量装置不是放在世界外面的眼睛

在经典实验的理想图像中，物体先具有位置、速度与能量，仪器随后来到，把这些性质读出来。仪器当然可能扰动对象，但这种影响似乎只是一层可以不断减小的麻烦。如果最终能够把扰动扣除，剩下的便应当是对象早已具有的完整状态。量子理论迫使人们追问：在所有测量问题中，这套图像是否都能保持？

上一章已经区分了制备分布、仪器误差和测量扰动。本章不再把这些区别重新说一遍，而把它们推进到一个更尖锐的问题：同一装置安排既产生读数，也规定读数所回答的问题，那么它参与的究竟只是数值的获取，还是物理量取得明确含义的过程？这里的“参与”不能被误解成意识可以按愿望改变电子。需要分析的是装置与对象之间真实的相互作用、记录结构以及可以共同成立的预测关系。

玻尔关于量子公设与互补性的讨论，在一九二八年的论文中得到系统表达；一九三五年，他又回应了爱因斯坦、波多尔斯基和罗森提出的完备性问题。[1][2]这些争论不宜被写成一个人坚持理性、另一个人拥抱神秘的戏剧。双方都要求物理学说明自己的概念与证据，只是对什么算作完整描述、何种条件允许赋予物理实在性，有不同判断。

本章将用三个承重环节展开：测量怎样把状态与记录关联起来，EPR 问题怎样逼出完备性与局域性的张力，贝尔类型不等式怎样把部分争论变成可检查的统计限制。我们既不把后来的实验当成对某个人全部哲学的最终判决，也不因解释仍有争议就抹去实验已经排除的具体结构。

44.2 一个两态测量模型揭开了哪一层困难

设被测系统有两个正交态$|0\rangle$与$|1\rangle$，装置准备在$|R\rangle$。一个理想相关过程把$|0\rangle|R\rangle$变成$|0\rangle|R_0\rangle$，把$|1\rangle|R\rangle$变成$|1\rangle|R_1\rangle$，其中两个记录态能够被区分。若输入为$a|0\rangle+b|1\rangle$，且整体演化线性，那么输出便是$a|0\rangle|R_0\rangle+b|1\rangle|R_1\rangle$。这不是一份实验报告，而是把测量关联问题写清的最小模型。

这个模型已经说明装置并非消极地悬在对象旁边。记录必须通过作用形成，系统态与记录态之间必须建立特定联系；若两种输入都使装置落在同一个记录态，仪器便不能辨别它们。测量准确性因此涉及输入与输出的对应结构，而不只是屏幕亮不亮。校准的任务，就是检验这种对应在什么条件下可靠。

然而，线性关联本身没有自动选出一个唯一记录。对叠加输入，整体模型给出的是联合叠加。只写下这条式子，再说“人一看便知道了”，并没有解释为什么一个结果出现，也没有说明所谓“看”在方程里增加了什么。测量问题在这里具有精确形态：线性演化、明确单次记录与状态解释之间怎样相容。

若再把环境加入，记录态可与不同环境态关联。环境态近似正交时，忽略环境的局部描述中，某些相干项受到强烈抑制。这样可以理解为什么许多宏观记录不容易呈现可见干涉，也能说明记录稳定与相互作用的关系。但局部密度矩阵的形式像经典混合，并不单独证明整体态在本体上已经变成某一确定分支。退相干解释了一项重要机制，不能被要求独自承担它没有证明的全部结论。

因此，装置参与有三个层次：建立输入与记录的物理关联，限定可以辨认的结果类别，使某些记录通过环境作用稳定下来。是否还需要额外坍缩机制、怎样理解整体叠加、单次结果的地位为何，是进一步的解释问题。把三个层次压成“观察改变世界”，会让一句话显得深刻，却使实际研究失去精度。

44.3 读数相同的两台仪器，是否完成了同一测量

为了看清装置参与的具体程度，考虑两台仪器都报告零或一，而且对任意输入态，报告零的概率都等于该态在$|0\rangle$方向的权重。第一台仪器在报告零后留下$|0\rangle$，报告一后留下$|1\rangle$；第二台先取得同样记录，再把系统的零一态翻转，所以报告零后留下$|1\rangle$，报告一后留下$|0\rangle$。两台仪器的当次结果分布完全相同，测后系统却不同。

用算符可以把区别写得更清楚。第一台的两个操作为$K_0=|0\rangle\langle 0|$、$K_1=|1\rangle\langle 1|$；第二台为$L_0=|1\rangle\langle 0|$、$L_1=|0\rangle\langle 1|$。分别计算$K_i^\dagger K_i$和$L_i^\dagger L_i$，得到相同的投影。因此两者具有相同的测量效应集合，却具有不同的状态更新。这里的翻转可以由后续物理操作实现，不是观察者在头脑中改写标签。

这项差别可以立刻检验：把第二次零一测量接到第一台后面，理想条件下会重复前次结果；接到第二台后面，则得到相反结果。若只保留第一次的直方图，两台仪器看起来等价；若把实验任务延伸到序列，等价就消失。测量描述必须说明它是只描述结果概率，还是同时描述系统怎样继续进入后续作用。

这个算例给“装置参与了什么”一个比泛论干扰更精确的答案：装置能够在相同即时统计下实现不同后续运行结构。于是，测量的身份不是由屏幕上的同一数字完全规定的。对于要把结果反馈到下一步的实验，记录与测后态的联系成为不可缺少的部分。不同装置在一个任务上等价，在另一个任务上可能不等价。

这也解释为什么测量链不能只按单次精度排名。如果任务是判断初始零一分布，两台理想装置同样合格；如果任务要求保留对应本征态供后续操控，第二台就不满足这一特定要求。所谓好仪器，需要相对于任务说明保留什么、改变什么，以及这些改变如何被检查。测量的价值与物理作用在此交织，但仍能用明确程序区分。

本书由此提出一项具体的发生学观察：同一显露 S 可以接续不同 D，因此不能仅凭当次显露识别完整 SDE。要重建装置与对象的特征纠缠，应把可接续的操作一起纳入。这个命题可被实际序列检验支持或修正；它不需要把意识当作额外物理力，也不以否认测量前系统的存在为前提。

44.4 互补不是两个互相矛盾的故事都可以随便讲

设想一个两路干涉安排。若保留两路的相干关系，输出概率随相对相位变化；若安排一个能够区分路径的辅助系统，两路便与不同记录态相关。干涉项包含辅助态之间的内积，因此路径记录的可区分性与干涉显露之间发生具体联系。装置配置改变的不是研究者的心情，而是联合状态与可实现测量的结构。

互补性可以从这种任务差异中得到一个谨慎入口：不同安排支持不同物理规定，不能把各安排下成立的叙述不加条件地叠成一幅经典照片。这不是说波与粒子的任何说法都正确，而是要求说明某个说法对应怎样的制备、作用与记录。一个词语若离开条件，可能从有用概念变成误导图像。

例如，探测器记录一次局域事件，不能单独证明系统在此前所有时刻都沿一条已经确定的经典路径运动；观察到干涉，也不能把每一次局域记录抹去，说它只是连续波的幻觉。真正困难是让两类结果在同一理论中取得协调。量子态、演化与概率规则提供一种协调方式，但对这一方式的本体解释仍有不同方案。

本书也不能把互补直接改名为 SDE 便宣告解释完成。若 SDE 分析要增加内容，应当说明哪些特征在某种运行中共同成立，改动哪种作用会破坏这一共同成立，以及怎样设计新的检验区分竞争解释。只有提出这种可工作的区别，主客互动的整体性才不会退化为“什么都有关”的宽泛陈述。

装置条件的作用也有边界。它不是任意生产想要的结果：若把相位调节器转过一定角度，输出应遵循可检查的变化；若记录器声称保存路径，却没有任何相应可辨信息，它的名称不会使干涉自动消失。物理规定的发生既依赖条件，也受到条件与结果之间的稳定关系约束。

44.5 EPR 把什么困难推到了远方

爱因斯坦、波多尔斯基和罗森一九三五年的论文提出：在不扰动一个系统的情况下，若能确定地预言某个量，就有理由讨论该量对应的实在性，并据相关体系考察量子描述是否完备。[3]本章只用这一短述定位原论证，不把原文所采用的位置—动量构造与后来的自旋例子混成同一历史文本。

为了便于计算，下面采用后来常用的两个自旋二分之一系统的单态。对同一方向分别测量，两端结果完全反相关，各端单独看仍各有一半概率。测量者若知道左端结果，就能推断右端同方向的结果；但在实验前，右端的局部统计没有给出固定的正或负标签。

困难在于，左端可以选择不同方向。若这种选择不改变远端系统的实在状态，而每一种选择都允许作相应确定推断，那么是否必须认为远端系统已经同时具有多个方向的确定结果？量子态通常不提供这样一张全部方向的共同数值表。完备性、局域性以及实在性的判据由此发生张力。

这不是简单地说“左边一测，右边立刻变”。条件概率的更新与可操纵物理信号不同；也不是说 EPR 已经用逻辑证明量子理论自相矛盾。原问题是在一组明确前提下，现有量子描述是否足够。要回答它，就必须指出接受或放弃了哪些前提，而不是仅凭某个哲学立场宣布对手不懂量子。

玻尔的回应强调，对预测条件与物理描述的理解不能离开实验安排。[2]但如果本书只停在双方分别强调的词语上，读者仍无法判断这场争论产生了什么研究增量。真正的下一步是问：某类满足局域条件的更完整模型，是否能够重现量子理论允许的全部相关？贝尔的工作使这个问题获得了新的数学形式。

44.6 把局域模型的前提逐项写出来

考虑两个相距较远的实验端。左端选择设置 a 或 a'，右端选择 b 或 b'，每次输出正一或负一。设 λ 表示该次试验共同原因或补充变量的描述。一个贝尔局域模型把联合概率写成给定 λ 后两端条件概率的乘积，再对 λ 平均。形式上是 $P(A,B|a,b)=\int \rho(\lambda)P(A|a,\lambda)P(B|b,\lambda)d\lambda$。

这项表达包含重要承诺。第一，给定足够的 λ 之后，左端结果概率不再需要远端设置与结果作为额外输入，右端亦然。第二，设置选择与 λ 之间满足所用测量独立性条件，因此同一个 $\rho(\lambda)$ 用于比较不同设置组合。第三，各组合中的试次必须按预先确定的规则进入统计，不能依赖结果偷偷改变样本。

为了展示不等式，可以先用确定性响应 A(a,λ)和 B(b,λ)，它们取正负一。这个做法不是额外宣布自然必定决定论；随机局域模型可以通过加入局部随机变量或直接对有界条件均值处理，得到同样的限制。我们选确定性形式，是让关键代数在纸上能够一眼复算。

这些假设不能被压缩成一个含义飘忽的“实在论”总标签。有人愿意保留隐变量却放弃贝尔局域性，有人修改测量独立性，有人重新理解结果与状态的关系。贝尔实验对不同方案的压力取决于它们是否落在上述结构内。只要前提没有写清，所谓“实验否定了实在”就比实验真正支持的结论走得更远。

还要说明，λ 可以是很多变量，甚至具有复杂分布；不等式并不要求它只是一个简单数字。若局域模型失败，不能靠把 λ 从一个变量增加成一百万个变量自动解救。限制来自因果分解与设置独立等组织方式，而不是变量数量不足。这正是它超越某个具体机械模型的重要之处。

44.7 CHSH 上界怎样在四项相关中出现

对每个 λ，定义组合 Cλ＝A(a,λ)[B(b,λ)＋B(b′,λ)]＋A(a′,λ)[B(b,λ)−B(b′,λ)]。由于 B 的两个值各为正负一，方括号中的和与差必有一个为零，另一个为正二或负二。因此 Cλ 只能等于正二或负二，其绝对值为二。

以同一个归一化分布 ρ(λ)平均，并使用平均绝对值不小于平均值绝对值，就得到|C|≤2，其中 C＝K(a,b)＋K(a,b′)＋K(a′,b)−K(a′,b′)，K 表示两端乘积的平均相关。这里特意不用 S 表示 CHSH 组合，以免与本书的显露 S 混淆。这个上界没有使用量子方程，完全来自前节的局域模型结构。

读者可能会问：同一次试验并没有同时选择四种设置，为何可以在同一个 Cλ 里写四个响应？答案不是仪器已经同时测到它们，而是所检验模型允许在同一 λ 下定义对各个局部设置的响应。借助设置与 λ 独立，不同试次获得的四类相关可以归入同一模型平均。若放弃这些组织条件，推导便不再直接适用；这正是需要公开的假设，而不是隐蔽技巧。

也可以用条件均值表述随机模型。令每个局部均值都落在负一到正一之间，固定右端两个均值 u 与 v，有|u＋v|＋|u−v|≤2。左端均值的绝对值不大于一，于是相同的平均上界仍成立。这证明限制不只针对预先决定每一个结果的特殊模型。一般概率性局域响应也不能任意突破它。

CHSH 论文于一九六九年把相应思路推进为实验检验方案。[4]本章的代数说明是现代记号下的自主推导，不把它伪装成原文逐句译述。它的思想力量在于：一个原先似乎只能围绕“真实究竟是什么”争论的问题，现在可以在明确条件下转化为四类可统计数据之间的数量限制。

44.8 量子预测为什么能够达到另一数值

对自旋单态，沿单位向量 a 与 b 测量并把结果归一成正负一，量子相关为 $K(a,b)=-a\cdot b$。取 a 沿 z 方向，a′沿 x 方向，b 沿 $(z+x)/\sqrt{2}$，b′沿 $(z-x)/\sqrt{2}$。四个点积给出三项负 $1/\sqrt{2}$ 和一项正 $1/\sqrt{2}$，代入前述符号约定得到 $C=-2\sqrt{2}$，故 $|C|=2\sqrt{2}$，大于二。

这个计算展示的不是一个粒子向远方发出可控指令，而是某种联合态与测量组合允许的统计结构。每一端单独的正负结果仍各占一半。只有把两端记录按设置和试次对应起来，才能形成这些相关。相关结构超出上述局域模型上界，与单端能够收到一个可随意编码的超光速消息，是两项不同命题。

可以直接检查无信号的局部概率。若右端进行一个完备测量而左端不知道右端结果，对右端全部结果求和，相当于在联合态中插入单位算符；左端自己的结果概率不因右端选择哪一组完备测量而改变。这个结论依赖标准量子操作规则，也不意味着所有关于非局域性的解释问题已经消失。它只明确说明，量子相关本身不能按上述方式被用作任意通信。

实验不需要完美单态才有讨论意义，但噪声会影响能否越过二。考虑一类简化混合态：以权重 v 取单态，以权重 1−v 取四维单位混合态。后一部分不贡献所用自旋相关，于是最佳上述组合的大小为 $2\sqrt{2}v$；当 v 大于 $1/\sqrt{2}$，才越过该 CHSH 上界。这个算例说明，观察不到违反，并不自动证明系统完全没有量子纠缠。

事实上，纠缠与某个特定贝尔不等式的可违反性不是同一个分类。测量选择、噪声结构和检验形式都会影响结果。若一次实验没有超过二，需要问状态质量、设置和统计检验是否足以区分候选，而不能把“未观察到”立即变成“原则上不存在”。反过来，若声称超过二，也必须核查数据规则与假设，而不能只看一个中心值漂亮。

44.9 实验怎样把纸上的上界变成有力度的证据

一次相关实验涉及试次的界定、设置选择、时间记录、探测效率和数据配对。若只保留双端都被探测到的少量事件，且保留机会依赖设置或潜在结果，样本可能不能代表推导所要求的总体。若设置与测量过程没有足够时空分离，局部传播的影响也可能成为替代解释。这些问题不是实验外部的手续，而是理论假设怎样进入实际证据的接合点。

因此，探测漏洞与局域性漏洞的处理有实质意义。前者要求避免用未经检验的公平抽样替代所需样本，后者要求实验时序与距离限制相关信息在测量期间以通常局部方式传播。设置随机性还需要说明，而不是只在装置上写“随机”两个字就结束。没有实验可以在完全不作任何假设的意义下裁决所有可想象世界；重要的是公开实际依赖并尽可能收紧替代说明。

亨森等人在二〇一五年报告了相隔一点三公里电子自旋的贝尔检验，同时处理主要探测与局域性漏洞。[5]本章采用这一事件说明检验能力的提高，不以“无漏洞”作为不需条件的终极词语，也不把一篇论文当作关于宇宙一切因果结构的最终证明。其力量来自实验设计让特定局域替代解释难以继续承担结果。

统计判断还不能只靠计算四个样本均值。有限样本会波动，装置可能有记忆，停止试验的时机与选择规则也可能影响显著性。因此检验应预先说明零假设、统计方法和数据保留原则。一个略大于二的中心值，若不报告不确定范围或适当的假设检验，不能单独承担同样强的结论。

这个层次上的创造尤其容易被忽视。新的物理世界不只是提出纠缠态，还包括能够让远处分开的设备形成可信相关数据的制备、同步、校准与统计结构。它们不是附属于“真正理论”的杂务，而是让某类新关系获得共同可检验身份的发生条件。关系是否真实，不能脱离关系怎样经受独立证据的过程。

44.10 哪些解释受到压力，哪些问题仍然开放

一类保持贝尔局域分解、设置独立与适当抽样条件的隐变量模型，不能重现超过上界的相关；这已经是很强的排除结果。但它没有排除所有隐变量理论，更没有证明意识必须引起坍缩。非局域的隐变量方案在逻辑上不属于已被这个上界排除的同一类，代价是需要清楚说明其非局域结构以及与可观测无信号关系如何协调。

另一些方案保留量子态的普遍演化，对分支或观察结果采取不同解释；另一些方案把坍缩当成需要额外动力学的过程。若两个解释在全部可实施实验上给出完全相同预测，仅凭本章所列数据便不能区分它们。若某种修改给出额外噪声、干涉损失或尺度依赖，则应把这些差别转成独立检验，而不是只比较谁的故事更合日常直觉。

放松设置独立性的方案也必须承担具体义务。仅仅说“万事都有共同原因”，不足以形成一个可与实验竞争的物理理论；它还需要说明相关如何产生、哪些设置与源的联系被预测、如何在未用于拟合的新安排中接受检验。若允许每次结果出现后任意补一段不可检查的共同历史，模型便可适配任何数据，也因此失去解释区分力。

本书的 SDE 解释同样不能豁免于这一要求。说装置参与属性发生，并不自动推出贝尔违反，也不自动指定哪个量子解释正确。它应当提出更具体的问题：哪些局部属性表无法共同成立，怎样的联合结构才允许实验结果，哪些装置变化真正改变这种结构。若没有这些可辨内容，“主客互动”也可能成为包容一切却排除不了任何错误的语言。

对爱因斯坦的评价因此应当更精确。他对完备性与局域性的追问促进了后来的可检验问题，即使某些希望保留的结构不能承担实验相关，也不能因此抹去问题创造的价值。对玻尔亦然，强调实验条件不等于已经给出所有后来数学与实验结果。思想史的任务是辨认每一步怎样改变问题空间，而不是追授一个永久获胜者。

44.11 让一条异常相关真正推动结构改写

设研究者在某组设置下得到超出旧模型上界的样本相关。沿 S—E—D 进入路径，第一步不是立即宣布意识或新本体获胜，而是由显露 S 追查上界所依赖的 E：设置与源是否独立，样本是否完整，局部响应能否因果分解。随后改写 D，重新规定试次、提高探测完整性、改变时序与距离，并在新的条件下复验。

这里 S 与 E 的张力推动 D 改变。若新程序消除了异常，则原判断需要收回；若异常在更严格程序中保持，则旧结构受到更强压力。接着可沿 D—E—S 倒逼路径，从主动改变空间分离与设置时序出发，迫使候选模型明确其允许的关联，再提出新的统计预期。每一步必须在结果出现前说明什么会支持、什么会削弱候选解释。

三方程在此分别提出可执行问题。S＝F(D,E)问相关数据如何由制备、测量与联合结构共同形成；D＝G(S,E)问异常和候选结构怎样决定下一次实验；E＝H(S,D)问哪些局域或样本假设应当在新增结果下重建。若 S 与 D 长期保持而旧 E 不能协调，两者的张力便推动 E 改写；若 D 与 E 不相容，则要重新求出能够成立的显露，而不是沿用旧预测。

这种研究不把装置视为偶然外壳，也不让装置承担没有证据的神秘力量。它在每个环节区分实际作用、统计选择与理论条件，让整体性通过可操作差异取得内容。新的物理规定只有在这种回检中保持，才有资格进入共同知识；一套理论语言不能仅凭自称发生学就跳过此关。

44.12 物理对象的身份怎样跨过装置的边界

到这里，装置参与了什么已经可以作出分层回答。它规定测量任务，建立输入与记录的关联，参与某些相干关系的保持或消失，并把联合结构转换为可比较数据。它没有

因此成为随意创造物质的主体意志，也没有让所有实验结果失去约束。装置越是进入理论，越需要接受明确的模型与独立校准。

一种旧式物的图像要求，任何时候都可把对象从关系中取出，仍保留全部已成属性。量子测量与贝尔问题则逼迫我们检查：哪些属性能够这样保留，哪些只在特定联合结构和实验条件下具有可同时使用的规定。这里的重建不只是改变看法，而是改变哪些对象描述可以共同进入实际运算和实验组织。

本书把这种重建理解为物理世界创造的一部分，但保留一个必要限度：发生学解释必须接受标准量子预测与已核实实验的约束，不能以新的本体词语代替具体结果。它的增量应当体现在更准确地识别系统边界、属性共存条件和研究改道点，而不是声称一个总框架已经消除了所有困难。

测量争论留下的不是一片无规则的空白，而是一组被不断压实的问题。完备性与预测成功不同，相关与通信不同，退相干与单次结果不同，局域模型排除与所有实在论排除不同。守住这些差别，下一章才能进一步讨论：两套数学方法如何在预测层面会合，而会合以后为什么仍然容得下不同的物理解释。

本章参考资料

[1] Bohr, N. The Quantum Postulate and the Recent Development of Atomic Theory. Nature 121 (1928):580–590. DOI:10.1038/121580a0，https://www.nature.com/articles/121580a0。核验原刊题名与年份。

[2] Bohr, N. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? Physical Review 48 (1935):696–702. DOI:10.1103/PhysRev.48.696，https://journals.aps.org/pr/abstract/10.1103/PhysRev.48.696。原刊摘要定位回应问题，未声称逐页精读全文。

[3] Einstein, A.; Podolsky, B.; Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? Physical Review 47 (1935):777–780. DOI:10.1103/PhysRev.47.777。原文扫描：https://inters.org/files/einsteinetal1935.pdf。历史概括为短述。

[4] Bell, J. S. On the Einstein Podolsky Rosen Paradox. Physics Physique Fizika 1 (1964):195–200, 原文：https://quantum.country/assets4/Bell1964.pdf；Clauser, J. F.; Horne, M. A.; Shimony, A.; Holt, R. A. Proposed Experiment to Test Local Hidden-Variable Theories. Physical Review Letters 23 (1969):880–884. DOI:10.1103/PhysRevLett.23.880。https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.23.880。本章 CHSH 为现代记号自主推导。

[5] Hensen, B. et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. Nature 526 (2015):682–686.
DOI:10.1038/nature15759，https://www.nature.com/articles/nature15759；作者预印本：https://arxiv.org/abs/1508.05949。仅使用摘要支持的实验设计与主要漏洞处理事实。

第四十五章 海森堡与薛定谔：两套方法如何会合成新的物理

45.1 会合不能只靠一句“它们等价”

一套方法从跃迁量与不对易乘法出发，另一套方法用微分方程寻找允许状态。如果它们给出相同能级，是否已经证明两套理论完全相同？如果它们在所有已做实验中一致，是否已经证明背后的物理解释也必须相同？这两个问题常被一句“矩阵力学与波动力学等价”同时盖住，实际上却需要不同层次的论证。

薛定谔在一九二六年专门讨论了自己的理论与海森堡、玻恩、约尔当量子力学之间的关系；埃卡特同年的论文也建立了联系两者的算符演算。[1][2]这些工作构成重要会合节点。但是，历史上最初的论证范围、后来完善的数学结构与今天教材中的标准结论，不宜被压成同一个瞬间。本章将使用明确模型展示会合实际要求什么。

我们区分三个层次。形式等价要求给出态、算符与运算之间保结构的对应；预测等价要求相同制备与测量任务得到相同概率；解释一致则还涉及理论中的数学对象被赋予怎样的物理地位。前两层在适当条件下可以严格证明，第三层不会由一个矩阵乘法自动决定。这个区分不是削弱会合，而是让会合的成果与边界都可见。

上一章关于测量的讨论也提供了提醒：两台仪器可能拥有相同即时概率，却在后续序列中不同。类似地，两套计算若只比较一张能级表，检验范围仍然有限。要判断它们是否表达同一物理结构，必须把状态制备、演化、观测量和可接续的操作一起纳入，而不能仅凭几项漂亮数字相同就提前宣布完成。

45.2 一个双能级实验任务的完整设定

考虑一个理想封闭双能级系统，选取正交基$|0\rangle$、$|1\rangle$，在这一基上定义泡利算符。令哈密顿算符 $H=\hbar\,\Omega\sigma_x/2$，其中 Ω 为实常数，初态为$|0\rangle$。实验任务是在等待时间 t 之后测量零一基，求得到$|1\rangle$的概率。这里没有默认这两个基态就是 H 的能量本征态；它们是制备与读出采用的基。

这一点很重要。若把“零一基”与“能量基”混作一个概念，就会误以为初态必定只获得整体相位，无法理解后续概率振荡。在所给哈密顿算符下，能量本征态是 σ_x 的两个本征态，能量为正负 $\hbar\,\Omega/2$。制备的$|0\rangle$在这个能量基中具有两个分量，后续相对相位变化能够改变零一读出的概率。

由于 $\sigma_x^2=I$，指数演化算符可以直接分成偶次幂与奇次幂，得到 $U(t)=\cos(\Omega t/2)I-i\sin(\Omega t/2)\sigma_x$。这个式子无须对每个时间点重新求解微分方程，且可检查 $U^\dagger U=I$。以它作用初态，得到 $\psi(t)=\cos(\Omega t/2)|0\rangle-i\sin(\Omega t/2)|1\rangle$。

因此，得到一的概率为 $\sin^2(\Omega t/2)$，得到零的概率为 $\cos^2(\Omega t/2)$。当 $\Omega t=\pi$ 时，系统在理想模型中完全转到 $|1\rangle$；当 $\Omega t=2\pi$ 时，态向量为 $-|0\rangle$，所有与该纯态相关的概率回到原来值。这个负号体现整体相位等价，不需要再创造一种新的可区分制备。

这条计算路线把时间变化放在态中，观测投影 $P_1=|1\rangle\langle 1|$ 保持不变。我们已经得到一个完整的、可以与另一条路线逐步比较的结果。下面不是换一个名字把同一行式子重复一遍，而是把时间变化重新分配，检查读出算符如何承担演化。

45.3 让观测量演化，结果还能一样吗

在海森堡绘景中，把初态固定为 $|0\rangle$，令观测量随时间变成 $A_H(t)=U^\dagger(t)AU(t)$。对 σ_z 计算，可以得到 $\sigma_{z,H}(t)=\sigma_z\cos(\Omega t)+\sigma_y\sin(\Omega t)$。这个正号可以用初始导数检查：$i[H,\sigma_z]/\hbar=\Omega\sigma_y$。若在变换时把 U 与 $U^\dagger$ 次序写反，就会把这一项的符号改错。

零一读出中的 $P_1=(I-\sigma_z)/2$，因而 $P_{1,H}(t)=[I-\sigma_z\cos(\Omega t)-\sigma_y\sin(\Omega t)]/2$。对固定初态取期望，利用 $\langle 0|\sigma_z|0\rangle=1$、$\langle 0|\sigma_y|0\rangle=0$，得到 $p_1(t)=[1-\cos(\Omega t)]/2=\sin^2(\Omega t/2)$。这与上一节完全一致。

为了避免只检验一个特殊概率，再计算 σ_y。海森堡绘景给出 $\sigma_{y,H}(t)=\sigma_y\cos(\Omega t)-\sigma_z\sin(\Omega t)$，因此其期望为 $-\sin(\Omega t)$。直接用上一节的时间态求 $\langle\psi(t)|\sigma_y|\psi(t)\rangle$，也得到同样结果。两个不同观测量的检验，让符号、时间尺度与状态对应同时受到约束。

两条路线的相同并非巧合。对任意合适的 A，都有 $\langle\psi(t)|A|\psi(t)\rangle=\langle\psi(0)|U^\dagger AU|\psi(0)\rangle$。这项恒等式清楚说明，时间变化可以分配到态或算符中，而可观测预测不变。它并不证明“真正变化的只是仪器”或“真正变化的只是电子”，因为绘景选择首先是理论描述的重新组织，不是单凭写法决定的本体裁决。

这个算例也帮助我们辨认错误比较。若一边使用演化后的态，另一边却仍使用未变换的固定态与固定算符，两边当然可能不同；那不是两套量子力学在物理上冲突，而是变换只做了一半。等价有操作纪律，必须把对应对象一起带过去。所谓同一理论，不是允许随便混搭各表示中的公式。

45.4 改变哈密顿量与改变表示不是同一个动作

现在稍作修改，令 H＝ħ(Ωσₓ＋δσ_z)/2，其中 δ 表示另一项频率尺度。仍从|0⟩出发，令 R＝√(Ω²＋δ²)。利用(Ωσₓ＋δσ_z)²＝R²I，可得演化算符，并推出 p₁(t)＝(Ω²/R²)sin²(Rt/2)。若 δ 不为零，最大转移概率一般小于一，振荡频率也改变。

这次改变不是把同一个问题换一种表示，而是改变了动力学。若一组计算采用 δ＝0，另一组无意中加入 δ 项，再把不同曲线称为“海森堡与薛定谔的冲突”，便错误定位了差异来源。真正需要比较的是两组模型是否具有同一哈密顿算符、同一初态和同一读出，而不只是各自用了哪位物理学家的名称。

同样，改变基以后，H 的矩阵通常会改变，但若态与测量也按相同幺正变换对应，预测保持。矩阵数字发生变化，不必意味着物理动力学发生变化；矩阵外观相同，也不保证物理含义相同，因为基态和实验读出的对应可能已经改变。形式与物理之间的接口必须被写出来。

这为本书的世界创造主轴增加一条必要限制：不是每一次符号重排都创造新世界，也不是只有增加一种物质才算创造。若变换保留完整可操作结构，它是同一物理系统的另一展开；若改变了允许作用、状态空间或可辨预测，则需要分析新的规定是否成立。两者混淆，会使“新”失去可判断的尺度。

在实际研究中，双能级近似本身也有条件。若外部作用显著耦合到被排除的其他能级，二阶矩阵再精确求解也可能不再描述目标系统。误差此时来自模型裁剪，而不是矩阵运算不够准确。计算可以完全正确，却在回答一个与实验不同的问题；这正是后面截断讨论要继续追查的地方。

45.5 谐振子的代数谱如何成为空间中的解

为把两种方法的联系推进到微分方程，考虑一维谐振子 H＝p²/(2m)＋mω²q²/2。在适当定义域上，令 a＝√(mω/2ħ)q＋ip/√(2mħω)，其伴随为 a†。利用[q,p]＝iħI，可以推出[a,a†]＝I，哈密顿算符写成 H＝ħω(a†a＋I/2)。

算符 N＝a†a 为非负，因为任意合法态的⟨N⟩等于||aψ||²。若存在归一化基态|0⟩满足 a|0⟩＝0，对其反复作用 a† 并归一化，得到|n⟩，且 a|n⟩＝√n|n−1⟩，a†|n⟩＝√(n＋1)|n＋1⟩。相应能量为 Wₙ＝ħω(n＋1/2)。这是一条代数构造路线。

把同一条件转到位置表象，令 p＝−iħd/dq，基态条件便成为[d/dq＋mωq/ħ]φ₀(q)＝0。解出 φ₀(q)正比于 exp[−mωq²/(2ħ)]，再由归一化得到系数(mω/πħ)的四分之一次方。代数中的“被降算符消去”，在空间表示中变成一个一阶微分方程。

继续作用升算符，会产生高斯函数乘以不同阶多项式的解。对于第一激发态，$\phi_1(q) = \sqrt{(2m\omega/\hbar)}q\phi_0(q)$。把它代入谐振子微分方程，得到能量 $3\hbar\omega/2$，与代数结果一致。这个具体对应比“矩阵和函数可以互换”更有内容：它把基态条件、归一化、激发构造和能量一起连接起来。

反过来，如果从微分方程得到这些正交解，计算 q 和 p 在它们之间的矩阵元，又会恢复相邻能级的升降结构。例如 $q_{nk} = \sqrt{(\hbar/2m\omega)}[\sqrt{k}\,\delta_{n,k-1} + \sqrt{(k+1)}\delta_{n,k+1}]$，这里 m 仍表示质量，n 与 k 表示能级指标。位置算符在这个基中不是对角矩阵，其非零元连接相邻能级。这也说明，空间位置与能量标签各有自己的组织方式，不能把同一个指标直观化为固定轨道。

这项会合让一个系统的不同特征可以由不同工具进入。代数方法迅速揭示能级结构，空间解显示节点与局部形状；二者不是谁消灭谁，而是通过明确对应共同约束物理预测。它们的协作能够减少某一路线中的盲点，也能够暴露不一致究竟来自条件、近似还是实际模型差异。

45.6 有限矩阵怎样悄悄改变了代数

计算机不能直接保存无限维矩阵。一个常见办法是只保留|0⟩到|N−1⟩，把更高能级删去。定义截断降算符 a_N，使它在保留能级上仍按 $\sqrt{n}$ 降一级；其伴随在最高能级处不能再升上去。这个最后的边界看似只是存储限制，却会改变对易关系。

直接作用各基态可得[a_N,a_N†]=I_N−N|N−1⟩⟨N−1|。对非最高能级，结果像原来的单位算符；对最高能级，却给出 1−N。这一修正不是浮点误差，无论把数值精度提高多少，只要维度有限且采用这种截断，它就存在。其迹为零，也与第四十二章讨论的有限矩阵一般性质相容。

这意味着一个特别需要警惕的差别。可以先把完整 H 投影到有限空间，得到 H_proj= $\hbar\,\omega$(a_N†a_N+I_N/2)。也可以先分别截断 q、p，再套用 H_quad= $p_N^2/(2m)+m\omega^2 q_N^2/2$。第二条路线给出 H_quad= $\hbar\,\omega$[a_N†a_N+I_N/2−(N/2)|N−1⟩⟨N−1|]，与第一条在最高层相差一项。

以 N=2 为例，正确投影的能量为 $\hbar\,\omega/2$ 与 $3\hbar\omega/2$；若先截断 q、p 再平方，两个对角能量都变成 $\hbar\,\omega/2$。错误不是来自“波动力学和矩阵力学不等价”，而是投影与乘法不交换。中间经过被删空间的贡献，在第二条计算中丢失了。

一般地，令 P 为投影、Q=I−P，则 $PA^2P = PAPAP + PAQAP$。最后一项记录从保留空间经过被删空间再返回的贡献。若先把 A 压成 PAP 再平方，这项便不见了。这个恒

等式揭示一种跨许多模型都重要的近似问题：删掉看似不被占据的层级，仍可能改变复合运算与有效动力学。

因此，截断必须结合目标任务评价。若初态与演化始终远离最高保留能级，某些低能结果可以非常准确；若外场驱动把振幅推到边缘，误差可能迅速扩大。不能用一个低能本征值的吻合，证明所有长时间演化和所有观测量都可靠。近似的适用域由状态、作用强度、时间与所问量共同规定。

45.7 被删除的能级怎样仍然改变低能结果

一个更小的模型可以把被删层级的影响算出来。设保留态的未耦合能量为零，被删态能量为 Δ，二者实耦合为 g，哈密顿矩阵为[[0,g],[g,Δ]]。其两个本征值为[Δ ± √(Δ² + 4g²)]/2。若 Δ 为正且|g|远小于 Δ，较低本征值约为−g²/Δ。

如果直接删掉高能态，剩下一维哈密顿量只是零，就失去了这个低能位移。高能态的实际占据可以很小，却通过耦合改变低能谱。把“没有显著占据”当成“没有任何作用”，是近似研究中的常见错误。有效理论之所以需要修正项，正因为被消去的自由度不总能无痕消失。

这个例子同时说明，扩大矩阵维度不是唯一办法。也可以在有条件的微扰近似下，把高能影响压缩成低能有效项；但必须保留近似阶数与能量间隔条件。若 g 接近 Δ，低阶展开不再可靠，继续使用−g²/Δ 便可能误导。有效理论有自己的成功范围，而不是无限维原理论的任意缩写。

与位置网格方法比较时也有相同纪律。一边保留更多基函数，另一边细化网格，并不自动保证双方正在逼近同一个算符。边界条件、截断区域与势能处理方式必须相容。若一边把波包离开计算区视为吸收，另一边却把它反射回来，两条时间曲线的分离首先来自不同模型边界。

这一层的会合是研究实践，不只是抽象定理。它要求明确目标误差，选择对目标最敏感的比较量，并在改变近似后观察稳定性。能量、跃迁概率、长时间相位与局域密度可能以不同速度收敛；一个量已经稳定，不意味着其他量也自动可靠。研究者必须说明“等价”的检验到底覆盖了什么。

45.8 两个脉冲的次序为什么不能随意交换

会合还要经过时间序列的检验。前面的双能级模型使用不随时间改变的哈密顿算符；实际操控可能先沿一个方向作用，再沿另一个方向作用。设第一脉冲给出 $U_x=\exp(-i\theta\sigma_x/2)$，第二脉冲给出 $U_z=\exp(-i\phi\sigma_z/2)$。先施加 x 脉冲再施加 z 脉冲，整体算符为 U_zU_x；反过来则是 U_xU_z。矩阵的书写顺序与作用顺序必须清楚说明。

从 $|0\rangle$ 出发，令 $\theta=\phi=\pi/2$。先经过 x 脉冲，态成为 $(|0\rangle-i|1\rangle)/\sqrt{2}$；再经过 z 脉冲后，除整体相位之外，态成为 $(|0\rangle+|1\rangle)/\sqrt{2}$。因此测量 σ_x 必得正一。若先施加 z 脉冲，初态仅获得整体相位，再经 x 脉冲，最后对 σ_x 的正负结果各有一半概率。两种次序具有可检验的差别。

这不是两个绘景的分歧，而是两个不同实验序列。若在海森堡绘景计算第一种序列，观测量应变为 $U_x^\dagger U_z^\dagger A U_z U_x$；若误把两个脉冲互换，得到的正是另一任务。绘景等价保留次序结构，不允许把非交换操作重新排列后仍称作相同实验。数学会合的条件越明确，装置序列的物理意义就越清楚。

当哈密顿算符随时间变化，若不同时间的 H 彼此不交换，演化通常不能简单写成普通指数 $\exp[-i\int H dt/\hbar]$ 而忽略时间排序。把时间切成短段并按正确次序相乘，可以看出问题的来源：先后的短时作用一般不能合并成一个与顺序无关的总量。数值演化中交换步序，有时改变的不只是近似误差，而是所模拟的实验任务。

还可以采用随时间改变的表示 $R(t)$。若新态写为 $R^\dagger\psi$，新哈密顿算符须包含额外项 $-i\hbar R^\dagger dR/dt$，不能只写 $R^\dagger HR$。这个项来自表示自身的变化，不是新发现的一种外力。是否保留它，会直接影响两套计算能否给出相同结果。由此可见，等价并不意味着可以少做计算，而意味着必须把描述变化带来的全部项正确搬运。

45.9 相同期望值还差多少才算预测等价

即使两套方法对某个量给出相同平均值，也不能立刻宣布它们预测相同。设一组结果以相同概率取正一或负一，另一组以四分之一概率取正 $\sqrt{2}$，以四分之一概率取负 $\sqrt{2}$，以二分之一概率取零。两者平均值都为零，方差都为一；第四阶矩却分别为一和二。比较平均与方差，仍会漏掉可辨差别。

这个例子可以作为同一个具有五种可能结果的观测量的两种状态分布，因此不是把两种不同刻度硬凑在一起。两组态只在部分统计指标上相同，并没有给出相同的全部结果概率。若实验能够分辨零结果，第二种分布就可与第一种区分，不必等待高阶矩估计。判断预测等价，应优先说明具体测量分辨能够识别哪些结果类别。

对于有限维系统，若一组足够完整的测量确实区分全部密度算符，就可以用它检验状态对应；但实际有限数据仍带有统计与校准条件。有限组观察通过，只能支持相应精度与任务范围内的一致。抽象定理中的“所有观测量”与实验中“已经测过若干观测量”，具有不同的论证强度。

这也解释为什么多时刻与条件预测值得单独检验。两个模型可能拟合相同静态谱，却对受控脉冲序列或测后演化给出不同结果。前一节的次序实验提供了一种明确的区分方法。理论会合要保留的是整套允许任务之间的联系，而不是只保留几张静态表；否则，形式上的漂亮对应仍可能掩盖物理上没有接通的环节。

45.10 预测相同为什么没有替我们选定全部解释

在适当的同一数学结构中，薛定谔绘景与海森堡绘景可以给出完全相同预测。但这种相同不能单独决定波函数是实体、知识还是某种更深结构的表征；也不能决定测量记录的唯一性应怎样理解。数学变换改变描述位置，不会自动提供所有本体承诺。

相反，如果某种解释额外提出会改变演化的动力学，它就不再只是对相同数学内容的另一说法。它必须给出与标准模型何时不同、差异有多大、怎样检验。例如，某项额外过程若导致相干衰减，应说明衰减率与系统条件的关系，而不只说“宏观时会坍缩”。能产生不同预测的修改，与预测完全相同的解释，需要不同评价方式。

也应避免把表示选择误当成科学家私人性格的直接外化。一个人可能更习惯图像，另一个人更习惯代数，但理论的成立不能靠心理标签证明。真正可继承的创造是他们怎样让原本分离的计算对象建立联系、怎样发现不变量、怎样把困难转成别人能够继续工作的程序。思想史可以讨论风格，却不能让风格替代数学与实验条件。

本书所说“创造新的物理世界”，在此应理解为一个共享规定系统的形成。来自不同路径的状态、观测量和演化被连接起来，使研究者能够跨表达开展同一问题，发现彼此错误并扩展可计算范围。会合不是把差异全部磨平，而是让差异在共同约束下发挥生产力。

这也为跨学科研究提供一项限制：两种理论偶尔给出相同结论，远不如存在系统的对应关系强；两种理论词语相近，也不证明对象相同。要声称某个本体框架“生出”另一理论，至少应说明变量、运算、有效条件与检验如何对应。若这些环节缺失，宏大的发生叙述尚未成为可工作的继承关系。

45.11 从两套不一致计算走通一条研究路径

设一个团队用能量基截断计算受驱谐振子，另一个团队用空间网格计算同一任务，短时结果一致，长时激发概率逐渐分离。沿 S—D—E 再创路径，先从不一致显露 S 进入操作 D：扩大基维数、改变网格与区域、控制时间步长，并检查两边是否使用相同外场与读出定义。只有先把这些差别变成具体对照，才知道需要重建哪一层结构 E。

若扩大基维数后差异推迟出现，且最高保留能级的占据同时减小，就有证据指向截断边缘；若扩大空间区域改变网格结果，则应追查边界反射。这里 S 与 D 的张力推动 E 修正的，不一定是量子理论本身，而可能是近似模型对可用态与边界的规定。修正后必须返回原长时预测重新比较。

三方程让诊断不能只盯一条曲线。S＝F(D,E)要求说明程序与模型共同产生了什么结果；D＝G(S,E)要求选择能区分截断和边界解释的操作；E＝H(S,D)要求据对照结果重新规定有效空间。若新 E 与现有 D 仍不相容，就应重新确定可以可靠预言的 S，例如缩短时间窗口或改换对边缘较不敏感的观测量，而不是继续宣称所有结果精确。

还可以沿 E—D—S 沉淀路径，由已知的截断对易子修正出发，设计一个专门接近最高能级的测试，预先比较两种哈密顿构造给出的分裂。这个测试可能揭露实现错误，也可能确认某项有效补偿的价值。方法的作用在于使下一步检验有方向，不能代替具体代数证明与误差分析。

两套方法最终吻合，也不是研究结束的充分条件。如果它们共享同一错误势能或同一错误标定，便可能一起偏离实验。因此交叉计算提供的是内部一致性证据，仍需独立的物理检验。把多种程序意见相同等同于自然已经同意，会把方法协作误用为自我确认。

45.12 会合创造了一种可接续的物理共同体

矩阵力学与波动力学的会合，使物理学能够在代数结构、空间解和实验预测之间往返。某一路线难以处理的问题，可能在另一种表示中更容易计算；某种直观诱发的误解，也可能被另一种运算揭露。这种互相校正的能力，比保留两套彼此隔绝的成功故事更为强大。共同体可以据此把失败定位到可修正的环节：是实验任务理解不同，还是算符变换不完整，是截断丢掉了中间贡献，还是候选动力学本来不同。能够区分这些失败，比要求所有研究者采用同一套符号更接近真正的理论统一。

我们通过双能级系统看到了同一概率如何由态演化与算符演化分别得到，通过谐振子看到了升降结构如何成为空间微分解，又通过有限截断看到了近似怎样破坏原本精确的对应。这三步共同说明，等价不是口头宣布，而是一套需要条件与检验的工作关系。

因此，物理世界的创造不只发生在提出一条新方程的时候，也发生在不同研究路线形成可靠接续的时候。新的统一不能靠忽略边界实现，新的多样也不能只是并列互不相通的语言。共同结构必须能够容纳不同工具，允许它们揭露彼此的失效范围，并把争论推向更精确的问题。

这一会合没有抹去测量与诠释的开放性，却提供了继续前进的共同基础。我们不必先解决所有哲学问题，才能计算并检验具体系统；也不能因为计算成功，就宣布所有哲学问题都没有内容。两种责任可以同时保留：把已经成立的物理关系做扎实，把尚未成立的本体承诺留在明确的研究位置上。

本章参考资料

[1] Schrödinger, E. Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen. Annalen der Physik 384 (1926):734–756. DOI:10.1002/andp.19263840804，https://onlinelibrary.wiley.com/doi/10.1002/andp.19263840804。书目及讨论对象由原刊记录核验。

[2] Eckart, C. Operator Calculus and the Solution of the Equations of Quantum Dynamics. Physical Review 28 (1926):711–726. DOI:10.1103/PhysRev.28.711，https://journals.aps.org/pr/abstract/10.1103/PhysRev.28.711。原刊摘要说明其算符演算连接矩阵动力学与薛定谔方法。

[3] Skinner, D. Principles of Quantum Mechanics. University of Cambridge，作者课程讲义，https://www.damtp.cam.ac.uk/user/dbs26/PQM.html。供海森堡绘景与谐振子标准形式对照；本章双能级、截断与消去模型均独立展开计算。

第十编 粒子、对称性与量子场
新物身份的持续生成

第四十六章 狄拉克与反物质
方程怎样改变粒子的身份

46.1 相容性要求为什么会超出原来的对象

上一章的会合解决了怎样把同一量子系统放进不同表示。狄拉克面对的困难更进一步：如果把相对论的能量关系、量子的状态演化以及电子的磁性放在一起，原来的电子概念是否足以容纳它们？这不是给旧方程增加一个高速修正系数的问题。不同要求相互牵制，可能迫使状态的组成和粒子的分类一起改变。

一九二八年的《电子的量子理论》把相对论与量子变换理论的相容性作为问题，并以矩阵线性化组织电子方程。论文同时明确承认，有关正负能量分支的解释困难并未全部解决。[1]因此，不能把后来完整的反粒子理解倒放回第一篇论文，写成一个人看见一个负号，便立即知道整个反物质世界。真正重要的是，方程产生了超过最初解释能力的结构，而物理学不得不继续处理它。

本章从这一尚未封口的结构进入。我们将分别考察三个层次：为什么需要多分量状态，为什么负能量不能靠删除来解决，以及为什么今天谈反粒子要进入可变粒子数的场论。三层有关联，却不是同一句话的重复。数学允许的解、理论赋予解的身份与实验确认的粒子，需要通过不同步骤接通。

46.2 给平方根开出一个矩阵空间

自由相对论粒子的能量关系可写成 $W^2=c^2p^2+m^2c^4$。若直接取正平方根作为哈密顿算符，便得到一个包含动量平方根的算符。它可以在动量表象中使用，却不直接提供空间和时间同时为一阶导数的局域方程。狄拉克的构造要求寻找 $H=c\alpha\cdot p+\beta mc^2$，使 H 的平方恢复上述关系。

把这个式子平方，不难看出条件从哪里来。不同动量分量的交叉项必须消失，因此需要 $\alpha_i\alpha_j+\alpha_j\alpha_i=2\delta_{ij}I$；动量与质量之间的交叉项也须消失，因此 $\alpha_i\beta+\beta\alpha_i=0$；最后 $\beta^2=I$。普通实数或复数不能同时完成这些任务。例如，若 α_i 与 β 都是非零数，它们的反对易和不可能为零。

问题于是迫使我们扩展运算对象。这里不是研究者先爱上矩阵，再把自然改写为矩阵；是试图保持的一组关系排除了普通数作为系数的方案。在三维空间，用四阶矩阵可以实现这些条件。一种常用表示取 β 的上下对角块为 I 与 −I，α_i 的非对角块为泡利矩阵 σ_i。相应状态必须有四个分量。

为什么不是两个分量就够？任何厄米二阶矩阵都可展开为单位矩阵和三个泡利矩阵的组合。四个平方为单位且两两反对易的非平凡厄米矩阵，都必须无迹：若 $AB=-BA$ 且 B 可逆，则 $BAB^{-1}=-A$，取迹即得 $\operatorname{tr} A=0$；其泡利系数于是相当于三维实空间中的四个两两正交单位向量，这不可能。扩到四维不是随意增加一个藏身之所，而是代数约束的结果。

但是，四分量并不意味着电子是空间中的四块物质，也不意味着存在四种普通电子。分量要按洛伦兹变换相互组织，并在自由场量子化后与自旋及粒子、反粒子的模式联系。把分量数直接当作小球数，会在数学刚刚完成第一步时便误判对象。

46.3 同一个矩阵怎样保留两个能量分支

先把三维问题缩成一维演示，取 $H=cp\sigma_x+mc^2\sigma\_z$。这是帮助检查代数的简化模型，不用它代替完整三维电子的自旋结构。因为 σ_x 和 σ_z 反对易，$H^2=(c^2p^2+m^2c^4)I$，本征值立即是正负 ε，其中 $\varepsilon=\sqrt{(c^2p^2+m^2c^4)}$。线性化没有消灭负分支，反而使它在同一有限矩阵里格外清楚。

这个模型还允许写出投影 $P_+=(I+H/\varepsilon)/2$ 与 $P_-=(I-H/\varepsilon)/2$。它们满足 $P_+^2=P_+$、$P_-^2=P_-$、$P_+P_-=0$，且相加为单位。对于恒定动量的自由演化，可以分别给两支乘上相反方向的时间相位。正负分支于是都是完整分解的一部分，不能说其中一半只是平方时偶然制造出来的数学垃圾。

如果永远保持自由运动，只研究固定的正能量子空间，可以形成有效的单粒子描述。但一旦引入空间变化的势或强外场，原先选定的投影与相互作用不一定交换。先把负分支删掉，再作用外场，未必等同于先在完整空间演化，再取适当低能结果。这与上一章讨论投影和复合运算不交换的结构呼应，不过这里的后果涉及相对论粒子的身份，而不只是数值截断精度。

当动量远小于 mc 时，正支展开为 $\varepsilon=mc^2+p^2/(2m)-p^4/(8m^3c^2)+$ 更高阶项。减去共同的静能，熟悉的非相对论动能便重新出现。这个展开解释了旧理论为什么可以成功，也给出它何时需要修正。非相对论电子并非被宣判不存在，而是成为一套有明确尺度条件的描述。

展开的控制参数是 $p/(mc)$，不是“速度看起来不快”的印象。若某一试验要求很高精度，即使该比值不大，第四次幂项也可能不可忽略；若只需要粗略量级，较低阶近似便可能足够。理论替换应由目标精度和尺度共同判断，而不是按新旧年份排出一条单向淘汰链。

46.4 自旋怎样从运动结构中进入磁性

四分量方程的重要成果之一，是电子的磁耦合不必再完全作为外加的小磁针假设。为看清这一联系，在缓变外电磁场中写机械动量 $\pi = p - qA$，q 表示带符号电荷。把四分量状态分成上下两组二分量，并除去静能相位。在能量和动量远小于静能尺度时，下分量近似为 $\chi = (\sigma\cdot\pi)\phi/(2mc)$。

代回上分量方程，可得有效哈密顿量 $H_{低} = (\sigma\cdot\pi)^2/(2m) + q\Phi$。利用泡利恒等式 $\sigma_i\sigma_j = \delta_{ij}I + i\varepsilon_{ijk}\sigma_k$，并保留动量分量的不对易性，有 $(\sigma\cdot\pi)^2 = \pi^2 - q\,\hbar\,\sigma\cdot B$。于是 $H_{低} = \pi^2/(2m) + q\Phi - q\,\hbar\,\sigma\cdot B/(2m)$。最后一项正是自旋磁耦合。

这里有一个值得停下来的细节。如果把 π 的各分量当作普通数相乘，叉乘部分就会被错误地删掉，磁耦合也随之消失。磁场不是在最后被补到答案上，而是通过机械动量的对易关系进入演算。一个貌似纯数学的次序问题，承担了可测的能级劈裂。

写自旋角动量为 $s = \hbar\,\sigma/2$，磁矩为 $\mu = gqs/(2m)$，上述系数给出最低阶 $g = 2$。这个结论有严格范围：它属于点状狄拉克粒子的最低阶结果。辐射修正会使电子的实际磁矩略有改变，下一章的量子电动力学正要接住这个细小但重要的差别。若把 $g = 2$ 写成不许修正的终极预言，便会把狄拉克理论的成功与它需要扩展的接口混淆。

这一推导也不能用来把自旋还原成一个带电球表面的机械转动。模型中需要的是旋转变换下的状态组织与磁耦合关系，没有使用球半径和表面速度。自旋作为物理量的客观性，不依赖它必须符合某种微缩日常机械图像。

46.5 负能量海解决了什么，又留下什么

如果把所有解都解释成一个电子可以占据的能量状态，负能谱向下无界会造成稳定性困难。电子为什么不连续释放能量并落到更低状态？狄拉克后来用占满负能态的图景处理这一问题：泡利不相容使已满状态不可再占据，缺去一个负电荷则相对于填满的参照呈现正电荷。这个历史性的空穴解释打开了反粒子问题，却不应被当作今天一切反物质现象的唯一底图。[2]

现代场论将正频率和负频率模式分别与粒子湮灭、反粒子产生联系起来。用离散模式的简写，场算符包含 a u exp(−iεt/ ħ)和 b † v exp(+iεt/ ħ)两类项；a 作用为消去一种粒子，b † 作用为产生其反粒子。这里的场算符不再是单个电子的位置概率波，因而不能把第二项误读为实验中出现了负能量电子。

其代数含义可以在一个模式上看清。负支对哈密顿量先给出−εbb † ；费米反对易关系 bb † ＋b † b＝1 把它写成 εb † b−ε。扣除与占据无关的真空常数以后，反粒子的每次激发增加正能量 ε。电荷算符则对两类占据给出相反号。能量都为正与电荷相反可以同时成立，根本不需要让反物质具有负质量。

这个小演算只展示自由模式的结构，不能独自证明完整相互作用场论的存在或真空能问题已经解决。真空常数在没有引力的散射计算中可以重设零点，不代表它在所有问题中都不重要。更不应由形式重排推断“无穷大已经被彻底消灭”。本章需要的结论较具体：单粒子负能谱的解释困难，促使理论改用允许产生和湮灭的状态空间。

这一改变还说明，反粒子不专属于依赖泡利填满的费米子。带电玻色场也有反粒子，不能用同一片负能电子海替所有物种解释。历史模型能够引导发现，并不等于后来理论必须原封不动地保存它的每一项图像。

46.6 一条弯曲轨迹还不能独自证明新粒子

安德森关于正电子的论文发表于一九三三年，其题目就是《正电子》。[3]粒子发现所依靠的不是在照片上直接读出一个名词，而是把轨迹曲率、运动方向、能量损失与质量判断结合起来。下面把这类判别抽象成一个可复核的装置问题，并不逐项重演原始照片的全部分析。

在均匀磁场中，运动方向与磁场垂直时，动量大小满足 p＝|q|Br。仅凭轨道向某一边弯，若不知道粒子向上还是向下走，就还不能唯一确定电荷正负。把速度反向同时把电荷反号，可以产生同样的几何弯曲。照片给出的是一条曲线，物理身份需要补上曲线的方向。

如果轨迹穿过一层吸收材料，并可合理认为穿过后损失动量，那么穿出侧的曲率半径应较小。由此前后半径的差别，可以约束运动方向；方向确定后，磁场方向与曲率才共同约束电荷符号。这里吸收层不是遮挡观察的累赘，而是帮助消除身份歧义的组成部分。

但是，正电荷还不足以区分正电子与质子。相同电荷大小、相同动量的两种粒子在同一磁场中有相同曲率，质量并未直接进入 p＝|q|Br。必须再使用能量损失、速度或射

程等独立信息。若只凭弯曲方向就说“看到了电子的反粒子”，实际上遗漏了决定性分类步骤。

设两种候选都有给定动量 p。它们的速度满足 $v=pc^2/\sqrt{(p^2c^2+m^2c^4)}$，质量越大，在相同动量下速度越低；与介质作用的响应也可能显著不同。由多个可测量共同约束同一候选，远强于从一张照片的外观作直觉命名。这不是用理论压倒观察，而是把观察组织为能够区分竞争解释的程序。

46.7 为什么真空里一个光子不能随意变成一对粒子

把反物质理解为“负号对应的新物”还不够，产生过程必须满足守恒条件。设一个光子在空无其他参与者的真空中变成电子和正电子。初始总四动量平方为零，末态电子正电子总四动量却为类时量，其不变质量至少为 $2m$。两者不能相等，因此这一孤立过程不可能发生。

换成质心系更直观：一对有质量粒子的总动量可以为零，总能量仍至少为 $2mc^2$；单个实光子只要有非零能量，就有大小 W/c 的动量，不存在总动量为零的静止系。知道光子能量超过两份静能，仍不足以保证单光子成对产生。能量阈值与动量守恒必须一起检查。

若有一个质量为 M、初始静止的核参与并吸收反冲，情况就变了。初始不变量为 $M^2c^4+2Mc^2W\gamma$；阈值时末态总静质量为 $M+2m$，因而 $W\gamma$ 阈 $=2mc^2(1+m/M)$。重核极限接近 $2mc^2$，但有限反冲仍要求额外能量。核不是替能量守恒开后门，它承担了必须被结算的动量。

反过来，在电子正电子质心系，若二者近乎静止并湮灭为两个光子，动量守恒要求两光子反向，能量守恒给出各约 mc^2。这里没有“物质变成虚无”：初态和末态的粒子种类、数量改变，能量动量和电荷等相应守恒。所谓物的稳定身份，因而不能再要求一个小球从开始到结束始终存在。

这些简单阈值演算把粒子概念从永久成分清单转成允许过程中的激发身份。某种粒子可以稳定地被识别，却不必在所有作用中保持数目。稳定的是一套关于质量、电荷、变换与相互作用的约束，而不是每个组成者的永生。

46.8 局域化为何会碰到粒子数的边界

相对论粒子的身份重组还可以从定位任务理解。假设要把一个电子制备在宽度为 a 的区域内，动量展宽的量级至少为 $\hbar/a$。只要 a 远大于约化康普顿波长 $\hbar/(mc)$，典型动

量仍可能处于非相对论范围；当 a 缩小到这一尺度附近，典型动量便达到 mc，继续使用单纯的 $p^2/(2m)$动能就失去充分根据。

这个量级判断并不等于宣称存在一堵绝对不可穿越的定位墙。它说明要提高空间分辨，需要改变所用能量和作用过程，而这些改变可能打开粒子产生通道。探针不再只把原有电子位置照亮，它还可能改变末态中有哪些粒子。若仍坚持整个过程只含一个永不增减的电子，问题就在模型的状态空间，而不只在测量仪器不够精细。

也不能从这里得出“位置没有意义”。实验仍可以定义顶点、能量沉积和轨迹起点，并在一定分辨范围内重建空间信息。需要修订的是把所有空间信息都解释为同一个持久小球的既存坐标。空间辨识任务继续成立，但它与粒子数、相互作用和末态分类的关系更复杂。

这一尺度分析可以提出可败条件。若在本来预计只有弹性散射的条件下系统地出现额外带电轨迹，就应检查是否已跨过产生阈值，是否有背景过程或探测器转换；若提高探针能量仍完全符合单粒子有效模型，则该模型在所测任务上继续成立。不能仅凭一句“高能必定复杂”就预先取消它的成功，也不能把低能成功延伸成无条件的单粒子本体。

由此看，反粒子的进入不是给现有粒子表加一列镜像标签。它让“系统包含什么”从一次制备时给定的固定答案，变成必须随允许过程重新结算的问题。一次完整预测需要给出各可能末态及其概率，而不限于为同一对象计算一条连续轨道。这正是接下来相互作用场论承担的新责任。过程的分类必须能与探测记录接续，否则粒子表扩充得再漂亮，也还没有形成可以检验的物理说明。

46.9 从身份歧义走到可以失败的规定

用本书的方法看，反物质不是从纯思维里制造出来的物体。发生改变的是，哪些显露可以被共同规定成同一种粒子，哪些装置差异能够排除竞争身份，以及哪些数学关系可以在新过程里继续成立。沿 D—E—S 倒逼路径，吸收层、磁场与轨迹方向的操作差异 D 先逼迫重新组织电荷、动量和能损之间的关联 E，再产生比单纯曲线更严格的粒子身份显露 S。

三方程在这里不替代轨道方程。S＝F(D,E)要求身份判断依赖实际的辨识序列与共同约束；D＝G(S,E)要求现有轨迹和候选身份一起决定下一项有区分力的测量；E＝H(S,D)则要求无法协调的结果迫使我们修订质量、电荷与过程的配对规则。它们的价值在于避免把“发现”缩成仪器读数或数学预言中的任意一项。

具体的张力可以逐项检查。若同一操作反复产生不符合既定候选的曲率与能损组合，S 与 D 的张力推动 E 增补新的身份，并须到新样本回检。若已有身份结构与实际操作范围不相容，D 与 E 的张力推动寻找能显露差别的新信号 S，不能仅靠名称维持分类。若理论身份与当前轨迹显露相互冲突，S 与 E 的张力推动改变 D，例如添加方向判据；新操作必须消除歧义，而不是人为挑选想要的轨迹。

负能海、场量子化与候选正粒子的竞争也要按不同层次处理。有些历史解释因不能覆盖更广物种或过程而失去基础地位；有些候选模型在有限测量上暂时相同，需要新的区分任务；有些数学重写保留全部预测，则不能仅凭外观宣布已被实验裁决。清楚区分这些失败方式，比把思想史写成一串胜利口号更能保存创造的实际内容。

狄拉克方程把电子从附带自旋的小点，推进到与相对论变换、多分量状态和反粒子过程共同规定的对象。它的成就不在于第一天便说尽一切，而在于形成了一套能够迫使自身解释扩展的关系。下一章将看到，即使粒子身份已经重组，电磁相互作用中的自作用、发散和精密修正仍会继续要求新的工作。

本章参考资料

[1] P. A. M. Dirac, The Quantum Theory of the Electron, Proceedings of the Royal Society A 117 (1928), 610–624. DOI:10.1098/rspa.1928.0023。原文扫描：https://mathweb.ucsd.edu/~nwallach/Dirac1928.pdf。参见导论及第一、二节；本章采用现代记号重新展开线性化。

[2] David Tong, Quantum Field Theory, 第四、五章，作者课程讲义：https://davidtong.org/teaching/quantum-field-theory/。供旋量、场量子化与空穴解释的层次区别对照。

[3] C. D. Anderson, The Positive Electron, Physical Review 43 (1933), 491–494. DOI:10.1103/PhysRev.43.491。机构典藏：https://authors.library.caltech.edu/records/3cqmp-9q965。本文轨迹与阈值算例为独立简化推演。

第四十七章 费曼、施温格与朝永：相互作用怎样成为可计算的物理

47.1 小小的偏差为何能要求巨大的重组

狄拉克方程给出了电子自旋磁矩的最低阶结构，但电子并不是一个与电磁场永不相遇的孤立系统。若它能发射和吸收辐射，同一相互作用也应参与质量、磁矩和能级的规定。问题随即变得尖锐：计算这些修正时出现的无穷大，究竟表明电子不存在，理论没有意义，还是我们把输入参数与可测量混在了一起？

施温格一九四八年的短文明确把质量、电荷的重整化与有限的磁矩修正联系起来；朝永一九四六年的论文处理场论的相对论不变表述；费曼一九四九年的论文发展时空方法。[1][2][3]这些路线不是三位天才分别给同一公式换了名字，而是在相互作用怎样组织、相对论条件怎样保持、哪些量需要重新规定的问题上形成不同的工作入口。

戴森一九四九年的论文把这些路线联系起来，同时在摘要中说明，高阶过程的讨论没有给出一般的收敛证明。[4]这个限定值得保留：一个可以系统计算并得到极佳结果的微扰理论，不因此自动成为每种极限下都已严格完成的数学对象。其力量正应在明确适用条件内评价。

本章不把量子电动力学的建立缩成“消去无穷大”的奇术。我们将依次区分振幅与概率、内部线与探测轨迹、裸参数与测量参数、紫外困难与低能软辐射困难。这些区分接在一起，才形成从形式展开到实验数字的工作系统。

47.2 微扰展开不是把电子切成很多份

设哈密顿量由可解的 H_0 与相互作用 V 组成。在相互作用绘景中，演化算符满足 $i\hbar\, dU/dt = V_I(t)U$。把这个积分方程反复代回自身，第一阶是一次时间积分，第二阶是对两个按次序排列的时刻积分，依次出现更高阶。每一阶表示运算中相互作用因子的次数，并不是宣称电子在自然中被切成若干按阶数命名的碎片。

若相互作用带一个小参数 λ，某个跃迁振幅可写成 $A = \lambda A_1 + \lambda^2 A_2 + \text{更高阶项}$。概率并不是各阶概率的简单相加。展开模平方得到 $|A|^2 = \lambda^2|A_1|^2 + 2\lambda^3 \mathrm{Re}(A_1^{*} A_2) + \lambda^4|A_2|^2$，再加上同阶其他贡献。第二项是干涉，可以为正也可以为负。

取一个完全透明的数值例子，令 $\lambda = 0.1$，$A_1 = 1$，$A_2 = -2$。只保留两阶振幅，得到 $A = 0.08$，概率为 0.0064。若错误地把两阶分别平方再相加，得到 0.0104；若只保留概

率到三阶，得到 0.006。两者的差别说明，近似阶数必须在整个计算链上保持一致，不能把“我算了两个图”自动理解成“我完成了某阶精度”。

这个例子不是量子电动力学的具体截面，只是展示展开纪律。真实计算中的阶数还要按电荷耦合、圈数、外态与过程分类。某一过程若最低阶振幅恰好为零，其首个非零概率阶数也会改变。因此，所谓修正很小必须相对于实际主导项判断，不能只看公式前面有一个小常数。

此外，小耦合并不保证任何情况下有限阶都够用。如果每阶同时伴随很大的对数，真正控制展开的可能是 $\lambda \ln(Q/q_0)$；当尺度比特别大，这个组合不再小。研究者需要重新求和某些重复贡献，不能用“耦合仍小”忽略大对数。有效计算取决于全部无量纲组合。

47.3 图是一种运算语法，不是一段被偷拍的电影

费曼图把相互作用展开中的连接关系画出来，使复杂表达的生成与核对变得高效。外线关联所制备或探测的粒子状态，内部线关联传播函数，顶点关联相互作用规则。图的直观性来自它保存了连接与组合结构，不在于它已经拍下每个粒子的实际路线。

考虑两种不同带电费米子的弹性散射。在最低阶，电子流与另一粒子流通过光子传播函数连接，振幅的结构可以简写为电荷平方乘以两侧电流，再除以转移动量平方。这里转移动量 q 等于入射与出射四动量之差。内部光子通常不满足自由实光子的 $q^2=0$ 条件，因此不能把它当作某个探测器漏拍的普通光子。

这也不是说内部粒子“暂时违反能量守恒”。在动量空间规则中，每个顶点都有能量动量守恒；内部四动量可以不满足实粒子的质量壳关系，是另一件事。把两者混为一谈，再用能量时间不确定性解释“借来能量后马上归还”，会用一个形象说法遮盖真正的计算条件。

内部线也不是单独可观测的零件。改变规范选择，若干图或某些传播项的外观可以改变，而完整的物理振幅必须保持相应不变性。若把某一张图中的一条线当作独立发生的事件，就可能把随计算约定改变的对象误认为实验实在。图的有效性来自全部规则共同产生可检验结果。

可以用一个代数检查显示这种约束。电磁传播项中的规范依赖部分通常带有 $q_\mu q_\nu$。它与守恒电流连接后产生 $(q\cdot j_1)(q\cdot j_2)$，在合法外态与完整处理下为零。对于自由旋量电流，q_μ 乘以电流可用入出态的狄拉克方程化为两个相同质量项之差。某项消失依靠状态条件，不是因为我们觉得它“不物理”便任意擦除。

这个检查具有实践意义。若计算结果依赖本应任意的规范参数，可能漏了某类图、反项或条件。形式冗余由此成为诊断错误的工具。物理世界没有多出一个可测的规范旋钮，理论却多出了一种检查计算是否完整的能力。

47.4 一个对数模型怎样区分输入与预言

为看清重整化的逻辑，先研究一个不冒充完整 QED 的简化模型。某个无量纲响应在截断 Λ 下写为 $R(Q)=g_0+b\,g_0^2\ln(\Lambda/Q)+O(g_0^3)$，其中 Q 表示探测尺度，g_0 是表达式中原有参数，b 为已知系数。若固定 g_0 后让 Λ 无限增大，第二项发散。

实验并不直接测 g_0。选一个参考尺度 Q_0，把实测响应定义为 $g_R=R(Q_0)$。在同一阶数反解，$g_0=g_R-b\,g_R^2\ln(\Lambda/Q_0)+O(g_R^3)$。将它代回另一个尺度的响应，截断对数相消，得到 $R(Q)=g_R+b\,g_R^2\ln(Q_0/Q)+O(g_R^3)$。

这一步没有从无穷大减去无穷大后任意宣布答案。它在有限截断下使用一次测量替代不可直接观测的参数，并按相同近似阶数消去截断依赖。最后剩下的尺度差别不能再任意调整：g_R 由参考测量固定以后，另一尺度的响应成为预言。若预言失败，不能再为每个 Q 独立重设 g_R 而仍声称同一理论成功。

同样重要的是，这个例子只显示单个对数结构如何处理。真实量子电动力学必须证明需要吸收的发散具有相容形式，质量、场归一化和电荷重定义之间满足约束，并把不同过程接成一套共同方案。一个玩具式子可以澄清逻辑，却不能替代完整理论的证明。

把重整化说成“用实验数据填掉所有错误”因此不准确。它确实需要实验输入，但预测能力取决于输入数量有限、同一输入约束大量结果，以及在改变截断或方案时保留相同物理关系。任何经验模型都可以逐点拟合；重整化理论的要求远比逐点拟合严格。

47.5 为什么换方案不能改变实验结果

同一个模型还允许观察重整化方案的作用。假设另一套记号把耦合定义为 $g'=g_R+a\,g_R^2$。若把 g_R 反解并代入响应，常数项会重新分配，某些二阶系数随之改变。但若坚持把同一个测量条件转换过去，最终 $R(Q)$到所保留阶数不变。参数数字的变化本身不是自然相互作用突然改变。

可以把它类比为采用不同坐标描述一条曲线，但不能把这个类比用得过宽。重整化方案改变还与有限阶截断结合，因而实际近似结果可能保留更高阶大小的残余方案依赖。这个残余可以帮助估计尚未计算部分的影响，却不是严格的概率置信区间，也不能保证覆盖所有遗漏机制。

若某一结果对合理的方案或尺度选择异常敏感，应检查展开是否稳定、是否缺少大对数重求和，以及是否已接近强耦合区域。它不是自动发现了新粒子，也不是自动推翻了理论。先定位计算适用域，是把异常转成物理证据前不可省略的一步。

更根本的问题是，允许哪些反项。如果每提高一点精度都需要无穷多个同等重要的新输入，模型的预测能力会受限；但在有效场论中，按照低能与高能尺度的比值排列算符，给定精度往往只需有限项。传统的可重整化与现代有效理论视角由此既有区别，也有接续。下一章以后还会看到，不能把“可重整化”当作一切物理存在的入场证。

47.6 磁矩为何成为理论的精密接口

上一章得到最低阶 $g=2$。定义反常磁矩 $a_e=(g-2)/2$，施温格的首阶结果为 $a_e=\alpha/(2\pi)$，其中 α 是精细结构常数。[1]这一结果的重要性不是数字看上去小，而是原先由最低阶结构确定的磁耦合，被同一电磁相互作用系统地修正；磁矩成为连接理论展开与测量的精密接口。

取 α 约为一百三十七分之一，只作量级估算，得到 a_e 约为 0.00116，因而 g 约为 2.00232。此处不使用这些舍入数字作现代精密实验比较。若要在极高精度下检验，必须包括高阶项、其他相互作用贡献、输入常数的不确定度和实验系统误差。首阶成功不能被扩写成只凭一个简式就达到全部精度。

磁矩测量也不必直接把一只微小磁针摆在桌上称量。理想非相对论条件下，带电粒子在磁场中的回旋频率大小为 $\omega_c=|q|B/m$，自旋进动频率为 $\omega_s=g|q|B/(2m)$。相减得到 $\omega_a=a_e\omega_c$，因此频率比 ω_a/ω_c 可以给出 a_e。共同磁场因素在理想比值中消去，这说明测量设计如何减少某些标定依赖。

实际装置包含束缚、电场、相对论效应与场不均匀等条件，不能直接套这个最简关系便宣称全部误差消失。本章用它说明一个方法上的要点：精密物理不是单纯把仪器刻度画得更细，而是寻找能使不需要的变量相消、需要的参数留下的观测组合。

还有一个循环论证风险。如果用磁矩测量反推出 α，再把同一个 α 代回磁矩公式，不能把必然得到的一致当作独立检验。独立比较需要来自其他测量的 α，或明确说明正在进行的是参数提取。相同公式可以承担测量常数或检验理论两种任务，区别由输入输出关系决定。

47.7 能级位移不是给电子轨道抹上一层油

辐射修正不仅影响自由电子磁矩，也影响原子能级。为了理解不同状态何以响应不同，考虑一个极短程有效扰动 $\delta V = C\delta^3(r)$。一阶能量修正为 $\delta W = C|\psi(0)|^2$。这个模型不等于完整的兰姆位移计算，却把“局部相互作用如何选择状态”这件事变得透明。

氢样原子的低角动量状态在原点可以有非零振幅，而带非零轨道角动量的规则解通常在原点消失。因此，两个在某个较粗近似里简并的状态，可能受到不同修正。新的能级差并不是电子在经典轨道上受到一层无法说明的摩擦，而是波函数结构与新有效作用之间的矩阵元不同。

这里 C 是何物必须说清。若把它从一个能级差拟合出来，另一个能级差只有在同一有效作用和适用条件下才是预测；若不同状态需要任意不同的 C，原来的短程模型便失去了共同解释能力。一个最小模型的价值，在于把“同一种机制”转成跨状态的比例约束。

例如，两种具有相同局部作用、但原点密度比为八的理想状态，其一阶位移也应为八比一。若测量明显违反这一比例，应检查有限作用范围、其他辐射项、核尺寸或近似失效，而不是立即把每个状态独立命名为一种异常。简化模型使偏离有了方向，也使修订负担能够具体分配。

这说明精密偏差能够创造新的研究对象。原先被能级简并遮住的区别，经更细的测量与相互作用计算取得稳定位置。世界的扩大不总表现为发现一种新粒子，也可以表现为原先被视作相同的状态现在具有可重复、可计算的差别。

47.8 低能光子的困难为什么不能用紫外药方处理

还有一种发散来自极低能光子，与高动量积分的紫外困难不同。真实探测器有有限能量分辨率，不能把“完全没有发出任何光子”与“发出一个能量极小、未被分辨的光子”任意精细地区分。如果理论只计算前一种理想化末态，其红外行为可能不适合直接与实验比对。

在相关条件下，应把无法分辨的软辐射末态与虚修正合在一起，形成与实验分类一致的包容量。用一个示意式说明，虚修正含 $-A\ln(1/\lambda)$，实软辐射含 $A\ln(\Delta/\lambda)$，二者相加后，人工的红外调节量 λ 消去，留下与分辨阈值 Δ 有关的有限项。具体量纲比值应由过程中的能标补足，此处只展示相同调节量的抵消结构。

Δ 依赖不该被一并删掉。改变探测阈值，实际上改变了哪些末态被算入同一事件类别，因此可测截面本来就可能变化。正确目标是消除非物理调节量依赖，保留真实定义

依赖。若把二者都称作“仪器干扰”，就无法理解理论为什么必须知道实验怎样分类事件。

这个例子与紫外重整化承担不同工作。紫外处理组织参数与短距离贡献，红外包容处理组织不可分辨的末态；两者有时同时出现在一项计算中，却不能用同一套口头解释带过。无穷大的来源不同，要求重新规定的环节也不同。

47.9 逐阶有限与无穷级数收敛不是一回事

一个简单积分可以帮助辨认这两个命题。令 $f(\lambda)=\int_0^\infty \exp(-x)/(1+\lambda x)\,dx$，其中 λ 为正。这个积分本身有限。如果在分母处形式展开成几何级数，再逐项积分，会得到 $1-\lambda+2!\lambda^2-3!\lambda^3+$ 后续项。每一项都是有限数，但阶乘最终比任何固定 λ 的幂衰减更快地增长，所以该幂级数对非零 λ 并不收敛。

不收敛不意味着前几项没有用。当 λ 很小时，初始若干项先变小，在适当位置停止，仍可逼近原积分；一直加到极高阶反而变差。这里可以由有限几何展开的余项直接估计截断误差，不需要把发散说成神秘现象。积分提供的是一项独立数学反例，证明“各阶有限”不足以推出“无限求和成立”。

这个反例不证明 QED 的所有非微扰性质，也不能给它指定最佳截断阶。它要求我们在使用微扰成功时保持命题准确：逐阶重整化、有限阶实验吻合与无穷级数的数学完成，是三个不同判断。理论可以在前两方面非常有力，同时仍留下后一方面的问题。承认这种差别不会削弱精密计算，反而防止把已证成的成果扩写成尚未建立的结论。

47.10 多路线会合如何变成相互纠错

朝永与施温格的协变场论方法、费曼的时空展开，使研究者能够从不同位置处理同一相互作用任务。若只说它们最终等价，就会漏掉会合的实践成果：复杂项如何归组、哪些贡献容易漏算、相对论和规范约束如何被保存，以及怎样把共同结果交给下一位研究者复核。

可以设想两组计算对某个散射修正给出不同答案。一组使用算符推演，另一组使用图规则。如果差别只是中间参数定义，两边换到同一测量条件后应消失；如果差别来自缺少一类末态，统一参数仍无济于事；如果差别来自展开阶数混用，则需要统一截断。会合只有能够区分这些失败原因，才超过形式上的互相致敬。

这也是解释竞争的边界。若两种方法在同一状态、同一测量定义及相同精度下保留全部振幅，它们是计算表示的不同组织，不能通过一次散射决定哪种图像才是粒子偷偷

采用的路线。若某种候选理论真的修改电磁顶点或加入新相互作用，则会改变特定过程中的系数和能量依赖，应以这些差别接受检验。

因此，新的物理规定必须先说清自己改变了什么。仅把内部线重新命名为某种本体，不会自动增加预测；给出可测的偏离，却又在失败后退回“只是解释”，同样不成立。QED 的精密性要求解释承诺与可检验内容之间保持诚实的接口。

47.11 把一次精密偏差走成完整研究路径

取一项磁矩或能级差作为显露 S，取改变场强、测量频率组合和独立标定作为差异运行 D，取粒子状态、相互作用项、参数定义与末态分类之间的共同制约作为特征纠缠 E。沿 S—D—E 再创路径，研究不能从偏差直接跳到新粒子，而应先改变可控制条件，看偏差是否跟随仪器、截断精度或某种物理尺度。

S＝F(D,E)在这里要求理论预测与实验事件定义相容；D＝G(S,E)要求下一轮控制试验根据异常形状与候选机制选择；E＝H(S,D)要求能够通过控制试验保留下来的差别进入新模型。三式共同防止两种捷径：一是让任意小偏差自动证明新物理，二是让理论声望自动取消偏差。

若不同 D 稳定产生相同残差 S，二者张力推动 E 检查缺少的相互作用，再用新过程回检。若既有 E 预测某种尺度依赖，而现有 D 无法分辨，二者张力推动建立新的 S，例如频率比而非单一频率；新显露应减少实际歧义。若 S 与 E 不相容，则推动 D 改变输入常数的独立来源、观测阈值或测量次序，并检查冲突是否仍在。第三维的改变只有返回原问题才有解释价值。

这种联合分析没有把量子辐射变成一般方法论的例证材料。相反，质量与电荷如何由测量条件固定、哪些不可分辨末态须共同计算、有限阶结果怎样判断稳定，构成了方法论本身需要学习的具体纪律。特征纠缠在这里是这些关系的共同规定，不能被缩成“当时的社会背景”或“科学家之间的合作”。

量子电动力学让带电粒子不再只是先定义完毕、然后偶尔交换光子的孤立对象。它的可测质量、磁性和散射响应必须在相互作用体系中被共同规定。与此同时，这种关系性没有取消精确数字，反而要求更严格地说明哪些数字是输入、哪些是输出、哪些差异属于实验，哪些只是表达方案。下一章将从电磁的单一相位自由进一步走向不交换的内部变换，看看相互作用结构怎样继续扩大。届时，计算的纪律仍然保留：先界定变量与观测的关系，再判断新的形式究竟增加了什么物理内容。

本章参考资料

[1] J. Schwinger, On Quantum-Electrodynamics and the Magnetic Moment of the Electron, Physical Review 73 (1948), 416–417. DOI:10.1103/PhysRev.73.416。原文：https://physics.umd.edu/grt/taj/624c/Schwinger_AnomalousMoment.pdf。参见质量、电荷重整化及磁矩修正段。

[2] S. Tomonaga, On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields, Progress of Theoretical Physics 1 (1946), 27–42. DOI:10.1143/PTP.1.27。原刊记录：https://www.jstage.jst.go.jp/article/ptp1946/1/2/1_2_27/_article。

[3] R. P. Feynman, Space-Time Approach to Quantum Electrodynamics, Physical Review 76 (1949), 769–789. DOI:10.1103/PhysRev.76.769。原文：https://faculty.washington.edu/seattle/physics541/%202010-reading/spacetime-approach-qed.pdf。

[4] F. J. Dyson, The Radiation Theories of Tomonaga, Schwinger, and Feynman, Physical Review 75 (1949), 486–502. DOI:10.1103/PhysRev.75.486。摘要：https://journals.aps.org/pr/abstract/10.1103/PhysRev.75.486。

第四十八章 杨振宁与米尔斯 局部比较怎样成为相互作用

48.1 不同地方的内部方向怎样比较

电磁理论已经教会物理学：同一个物理状态可以有不同的势与相位描述。但当内部状态不再只有一个相位，而有几个能够相互转动的分量时，局部描述如何接续？在这里，比较本身成为问题。两个地点各自选用一套内部基，不能未经说明便把它们的分量直接相减。

杨振宁和米尔斯一九五四年的论文从同位旋的局部变换出发，构造新的规范场，并指出它的场方程具有非线性。[1]这项构造后来成为非阿贝尔规范理论的重要基础，但不能将当年以核子同位旋为入口的具体模型直接等同于今天的强相互作用理论。可继承的是结构方法，后来确定的群、物质表示与实验身份还需要另行建立。

本章集中解释这个结构方法改变了什么。我们用协变导数、两条路径的比较以及格点上的闭合乘积，把“局部内部变换”变成能够检查的关系。随后再问，这些数学关系怎样产生动力学约束，哪些仍只是描述冗余，哪些可以进入实验。

这里有一条贯穿全章的界线：选择内部基的自由，不等于自然多了一个可观测运动。规范势的引入允许我们组织相邻地点的比较；若要说引入了新的物理场，还须给它动力学、状态条件与可测后果。不能从“允许换写法”直接跳到“新相互作用已经被证明存在”。

48.2 普通导数在哪里失去共同意义

令 $\psi(x)$是一个内部多分量场。在不同地点采用新的基，相应分量变为 $\psi'(x)=U(x)\psi(x)$，其中 $U(x)$是幺正矩阵。如果只做全局变换，U 不随位置变化，那么普通导数也简单地乘上 U。若 U 因地而异，求导就多出一项：$(\partial_\mu U)\psi$。因此 $\partial_\mu\psi$ 不再与 ψ 按同样方式变换。

这个额外项不是粒子突然受到一种由书写者制造的力。它首先来自比较规则不完整：新坐标基在空间变化，普通导数同时记录了对象变化与基变化。若把两者都解释成物理场的变化，就会把描述任意性带进可观测预测。

为分离二者，定义 $D_\mu = \partial_\mu - igA_\mu$，其中 A_μ 是内部空间的厄米矩阵，g 为耦合参数。要求 $D'_\mu\psi' = U D_\mu\psi$，逐项代入便得到 $A'_\mu = UA_\mu U^{-1} - (i/g)(\partial_\mu U)U^{-1}$。这个非齐次变换项恰好抵消基变化产生的额外导数。

可以检查阿贝尔特例。令 $U = \exp(i\theta)$，矩阵都可交换，则 $A'_\mu = A_\mu + \partial_\mu\theta/g$。这个结果说明新构造保留了电磁规范结构作为简单情况。非阿贝尔并不是把电磁理论全盘推倒，而是允许内部变换不交换，从而改变局部比较的代数。

但“必须用协变导数”仍是带条件的命题：我们已经选择了局部描述冗余、场的变换表示以及希望保持的局域形式。它不证明所有可能世界都必须拥有同一规范群，也不单独决定 g 的数值。结构约束与经验输入共同组成理论，不能让前者冒充后者。

48.3 两次比较为何会留下曲率

现在比较沿两个不同方向的微小变化。把 $D_\mu D_\nu$ 与 $D_\nu D_\mu$ 相减，作用在任意合法 ψ 上，可得 $[D_\mu, D_\nu]\psi = -igF_{\mu\nu}\psi$，其中 $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$。前两项熟悉，最后的矩阵对易子是新的关键。

如果 A_μ 都是普通数，最后一项为零；若它们指向不同内部方向，就可能不交换。以 SU(2) 为例，令 $T_a = \sigma_a/2$，它们满足 $[T_a, T_b] = i\varepsilon_{abc}T_c$。取空间中恒定的 $A_x = aT_1$、$A_y = bT_2$，普通导数全为零，却仍有 $F_{xy} = gabT_3$。没有位置梯度不代表没有非阿贝尔曲率。

这一算例对直觉提出了明确要求。对于普通电磁势，我们习惯从旋度判断磁场；现在必须同时考虑内部方向的连接次序。两个在空间上都恒定的矩阵，可以因内部运算不交换而产生非零场强。新物理不是把“颜色”当作一种肉眼可见的油漆，而是让内部表示的组合规则参与相互作用。

也可以直接比较小步矩阵。令 $X = \exp(iga\ell T_1)$、$Y = \exp(igb\ell T_2)$，则 XY 与 YX 在二阶出现差别。闭合组合 $XYX^{-1}Y^{-1}$ 约为 $I - ig^2ab\ell^2T_3$；方向约定若反转，符号也反转。这个结果与上述曲率相符，说明微分表达与路径乘法保存了同一结构。

此处的“绕一圈”是内部平行运输的数学比较，不等于实验中任意粒子必定沿这四条边走了一遍。要把它变成测量，需要指定探针、状态以及如何形成规范不变的输出。几何语言能够指导构造，却不能省略实验接口。

48.4 场强为何也不总能直接当观测值

在电磁学中，电场与磁场对通常的势规范变换保持不变。非阿贝尔场强一般只按 $F'_{\mu\nu}=UF_{\mu\nu}U^{-1}$ 协变，并不逐个分量不变。某一内部颜色分量的数字可以随基选择改变，不能把这个数字本身当作独立的仪器读数。[2]

不过，矩阵的适当组合可以消除这类任意性。例如 $\mathrm{tr}(F_{\mu\nu}F_{\rho\sigma})$ 在共同共轭下不变，因为迹可以循环移动矩阵。对刚才的恒定场例子，$\mathrm{tr}(F_{xy}^2)=g^2a^2b^2/2$；若用同一个 U 旋转全部内部方向，这个数保持。它表明我们可以从冗余变量中构造不依赖内部基的关系。

“可构造规范不变量”仍不等于“已经说明怎样测量”。一个局部复合算符可能需要重整化，实验提取也可能依赖复杂过程。但它至少通过了第一道必要检验：不同描述不会为同一物理安排给出互相矛盾的数字。理论的可观察内容要继续与动力学、制备和探测连接。

因此，规范冗余不应被赞美为自然的无限自由，也不应被贬成毫无用途的多余符号。它允许用局域且相对论形式清晰的变量组织问题，同时要求我们识别哪些差异不代表新状态。忽略前一面，会错失结构方法；忽略后一面，会凭空增加不可观测的物理内容。

在有边界或非平凡整体条件时，还必须说明允许的规范变换范围。某些边界变换关联实际荷或边界数据，不能把所有形式上相似的变化不加区分地宣布为纯冗余。本章主要处理局部平滑变换及固定边界条件；这个范围限定防止一句“规范都是假的”代替具体分析。

48.5 从开链到闭环的一次离散推导

把空间暂时分成格点，是看清比较结构的另一方法。令 L_{xy} 表示从地点 y 向地点 x 运输内部状态的矩阵。若每个地点独立换基 G_x，必须有 $L'_{xy}=G_xL_{xy}G_y^{-1}$。端点的两个变换都在式中，才能使运输后的分量与新基相容。

考虑三点 x、y、z 组成的闭路。乘积 $W=L_{xy}L_{yz}L_{zx}$ 的变化为 $W'=G_xWG_x^{-1}$，因为中间地点的 G_y、G_z 与其逆矩阵依次抵消。因此 $\mathrm{tr}\,W$ 不变。这个抵消可以逐项核对，不依赖对“闭环天然真实”的直觉。开链矩阵本身保留端点基依赖，闭环的迹才去除了起点的共同共轭。

若所有连接都可以写成 $L_{xy}=G_xG_y^{-1}$，那么沿任意可收缩闭路相乘都得到单位矩阵。这是一种纯规范结构：各点记号可能变化很大，局部闭环却没有曲率。与之相

反，前面的不交换恒定矩阵在小方环上产生非平凡结果，不能通过同类局部换基让全部连接同时变成单位。

这个例子揭示“连接不同”与“物理不同”的区别。两组连接若只差格点换基，闭环迹相同；但不同闭环数据通常能揭示不同的场组织。不能先把所有连接当作独立物理量，再额外用规范变换制造许多副本。正确计数要在等价类或相应约束之下进行。

格点间距在这里是分析与计算工具，不代表自然已经被证明由固定网格组成。若用格点近似连续理论，必须研究间距、体积和时间方向等极限。一个有限格点模拟给出漂亮图形，只说明有限问题已被计算；要声称对连续物理成立，还须展示相应控制。

48.6 一个有限闭环怎样显示非交换的大小

小环展开给出的是最低阶结构，还可以用二阶矩阵把有限环精确算出。令 $X=\cos u\, I+i \sin u\, \sigma_x$，$Y=\cos v\, I+i \sin v\, \sigma\_y$，闭环矩阵为 $XYX^{-1}Y^{-1}$。反复使用 $\sigma_x\sigma\_y=i\sigma\_z$ 及循环关系，取迹后得到半迹 $w=1-2\sin^2u\, \sin^2v$。这里 u、v 都是无量纲运输参数，已经包含耦合与路径长度。

若 u 或 v 为零，$w=1$，因为其中一段内部作用消失，闭环乘积退回单位。如果 $u=v=\pi/2$，$X=i\sigma_x$、$Y=i\sigma\_y$，两者反对易，闭环得到 $-I$，$w=-1$。这一结果无法由把两种内部转动当作普通相位直接相加获得。非交换留下的差别不仅存在于抽象对易子，也可以在有限矩阵中完整计算。

但半迹不是闭环的全部信息。在 SU(2)中它决定这一单个矩阵的共轭类，不能据此恢复沿路每段连接，更不能独自重建整个空间的场配置。不同配置可能在某一个环上给出相同 w，需要改变环的形状、大小与位置取得更多约束。这与单个散射总截面不能唯一确定全部相互作用的道理相似。

还要注意小参数下的阶数。矩阵偏离单位的首项可以是 uv 阶，而迹中泡利矩阵为零，使 w 的首个偏离到 u^2v^2 阶才出现。一个可测不变量的首个敏感阶数，可能高于中间矩阵变化的阶数。因此不能只看内部变量已经变化明显，就断言某个选择的输出必然具有同样灵敏度。

48.7 局部为零为什么仍须检查整体条件

局部曲率说明小范围绕行的结构，却不能在任何空间上独自决定全部闭路信息。考虑一个具有周期坐标的圆环，先用阿贝尔模型说明。若势沿环恒定，局部场强可以为

零，但环绕一周的相位为 exp(ig ∮ A·dl)。只要这个相位不为一，保持单值与周期条件的规范变换就不能把它处处消去。

若企图用 θ 随坐标线性变化来抵消 A，绕回原点时 θ 已经改变。只有当总相位差是整周倍数，新旧场的周期条件才相容。局部微分运算没有错，遗漏的是允许变换的整体边界要求。由此可见，宣布某场“只是纯规范”以前，应说明讨论的区域和可允许的变换。

这个圆环例子不证明所有规范势都能直接观察，也不把非阿贝尔问题还原成电磁学。它展示的是一个共同纪律：局部关系与整体条件共同决定等价类。将场强为零误写成一切运输效应为零，会在非单连通区域失效；反过来，承认整体效应也不意味着每个势分量变成可观测自由度。

对于研究实践，这项区别要求先辨认候选模型是否在相同边界条件下比较。如果一个计算允许周期单值变换，另一个偷偷改变了边界扭转，两者的谱差别可能来自不同物理任务，而不是同一理论的两种规范不一致。局部对称原则越精巧，越需要把这些看似次要的条件写清楚。

48.8 自相互作用怎样从同一个结构出现

若采用由 tr(F_μνF^μν)组织的规范场动力学，非阿贝尔 F 中的对易子使作用量除了导数平方，还包含三次和四次的规范势项。这些项意味着规范场之间存在直接的自相互作用。它们不是在实验不吻合以后任意加上的每条独立规则，而受同一个场强结构和耦合约束。

用示意记号 F＝∂A＋gAA，平方便有(∂A)²、g(∂A)AA 以及 g²AAAA。这里省略内部指标和时空收缩，只为显示阶数关系。三点与四点相互作用的系数不能在保持原结构时任意独立改变。若人为把其中一类删掉，某些相容性和高能抵消就可能失效。

这与电磁场在最低阶没有同类纯光子自耦合形成区别。但应避免过度陈述：量子电动力学中可以通过带电粒子圈产生光子之间的有效散射。准确的比较是，非阿贝尔规范作用量本身包含相应的树级自耦合，而不是“光子永远不会彼此影响”。不同机制可能产生相似的过程名称，却有不同阶数和条件。

同一个几何结构也不唯一决定完整理论。我们仍须选择规范群、物质场所在表示、允许的质量与相互作用，以及可能的更高维项。局域性、幺正性和指定精度下的有效展开会进一步限制选择，但不能从“有规范对称性”直接读出全部自然常数。

因此杨—米尔斯的创造在于建立一个强大的生成规则：一旦给定若干结构输入，许多相互作用关系不再彼此独立。它减少任意性，增加跨过程检验的可能，却并未取消经验选择。物理统一的力量正在这些受约束的联系中，而不在把所有未知都藏进一个漂亮名称。

48.9 分量多了，物理自由度为何不能照数相加

四维时空中，每个内部生成元对应 A_μ 的四个分量。若规范群有三个生成元，形式上便有十二个势分量；这不表示十二种独立传播的偏振。无质量规范场受到约束且具有描述冗余，在线性化的传播分析中，每个生成元对应两种物理横向偏振。[2]

以平面波说明最简单的计数。设波矢沿 z 方向，极化四矢量满足相应的横向条件，并允许加上与波矢成比例的纯规范部分。两个关系共同去掉不独立的时间和纵向组合，剩下 x 与 y 方向的传播偏振。非阿贝尔理论在围绕平凡背景的自由极限重复这一计数，但相互作用时还须维持约束的一致性。

如果直接给矢量势加入一个普通质量平方项，其非齐次规范变换一般不会保持原形式。不能只说“实验需要重粒子，所以加上质量就结束”。建立自洽质量机制、追踪纵向自由度从何而来，是下一章电弱理论必须处理的核心问题之一。

这个计数也不能自动推广到强耦合区所有可观测粒子。某种理论的微扰场变量不一定对应能够自由飞到探测器的单粒子。例如强相互作用中的色自由度与色单态强子的关系，需要非微扰动力学。看到拉格朗日量中一个场名，并不意味着实验必然能单独装进一瓶这种粒子。

本书所说新的“物”，因此应当在层次之间保持区别：内部场变量、规范不变组合、散射外态与实验重建粒子，可能相关但并非同义。一个理论的成熟，表现为它能说明这些层次怎样接续，而不是把其中最直观的一层强加给全部公式。

48.10 渐近自由让强相互作用获得新的尺度结构

一九七三年，格罗斯与威尔切克、波利策分别研究了非阿贝尔规范理论的高能行为，建立渐近自由这一关键结果。[3][4]其意义不是强相互作用从此处处变弱，而是某些理论在高能、短距离条件下的有效耦合减小，使微扰工具获得可控入口。

把一圈结果的结构写为 $dg/d\ln\mu=-b_0 g^3$，其中 b_0 为正。两边积分，得到 $1/g^2(\mu)=1/g^2(\mu_0)+2b_0\ln(\mu/\mu_0)$。因此 μ 增大时 g 减小。这个推导只使用给定的微分方程，足以显示对数运行；b_0 的数值和符号则由群结构与物质内容的量子计算决定。

特别需要强调，加入足够多物质场可能改变系数符号。不能从“非阿贝尔”四字无条件推出任意物质配置都渐近自由。更不能由高能耦合趋小直接证明低能一定禁闭。高能可控展开与低能强动力学之间存在需要额外研究的距离。

若令 μ 向低处移动，公式可能在某个尺度给出很大的耦合，甚至形式发散。这是低阶微扰解失去可靠性的警报，不是实验已经发现一个无穷大的力。需要格点、有效理论或其他非微扰方法进入。将近似失效处的数学奇点直接当作自然奇点，是从计算到本体的一次越界。

这里 μ 是组织不同探测尺度的能标，不是宇宙时钟。今天用较高动量转移研究过程，可以得到比低动量过程更小的有效强耦合；这不表示同一实验放置若干年后强相互作用自行减弱。尺度运行与随宇宙时间变化，是不同命题，第五十章将专门追究它们的区别。

48.11 从微扰场到可检验的集体输出

若规范势的颜色分量不能直接拿来测，非阿贝尔理论怎样接受检验？答案不能只是“通过一切实验”。必须指定由理论组织的可观察量，例如特定散射中的角分布、多个喷注之间的关联、色单态谱或合适定义的闭环量，并说明从场变量到这些量使用了哪些近似。

以喷注为例，高能过程中的短距离展开描述最初的夸克与胶子活动，而探测器记录的是大量强子、能量沉积与轨迹。一个喷注算法把这些记录按明确规则归类。改变归类规则可能改变数值，因此理论比较必须采用相同定义。喷注不是把一个孤立胶子直接放大到肉眼尺度，而是通过集体末态保留其部分可计算特征。

这类比较能够区分只含独立矢量传播的候选模型与具有非阿贝尔自耦合的模型，但需要综合不同过程、角关联和背景。不能说看到“三股喷注”这幅外观就无需计算地看见了某个抽象群。实验图形与结构归因之间有一条可以检查、也可能出错的推断链。

对格点谱计算也一样。一组输出能级接近实验，仍须追查参数是否已由这些能级输入、有限体积和格距误差怎样控制、其他独立量是否被预测。规范理论的解释力不是由公式的复杂度保证，而由同一有限输入能否约束丰富结果来体现。

48.12 闭环比较怎样改变研究的入口

沿 E—D—S 沉淀路径，可以先从局部比较和不交换连接构成的特征纠缠 E 出发，安排能够检验路径、尺度或自耦合差别的差异运行 D，再形成可比较的显露 S。这个顺序

不是要求实验服从理论想象，而是要求理论将自己的结构承诺转成别人能够反驳的操作任务。

S＝F(D,E)要求输出与过程定义、约束和场结构共同相容。D＝G(S,E)要求已经显露的残差与候选连接规则一起决定下一种实验或计算。E＝H(S,D)则要求被多种操作稳定保留的差别反过来修订相互作用结构。这里 E 是比较、变换、传播与观测之间的制约，不是把群名、年代和实验室并列成背景表。

三种张力在具体工作中各有负担。S 与 D 在独立过程间不能协调，推动 E 检查是否缺了自耦合或错误物质表示，再用未参与拟合的过程回检。D 与 E 在当前分辨率下不能协调，推动构造新的 S，例如不同角关联而不是单一总事件数，并检查新输出是否确实区分候选。S 与 E 保持冲突，则推动 D 改变能标、末态分类或格点参数，检查差异来自物理还是近似。

这一研究结构也保护我们不把“局部任意性”误认为主观意识创造相互作用。局部基可以任意选择，实验结果不能任意选择；理论允许的变换越多，越要说明哪些关系在这些变换下保留。规范理论的创造力不是给自然增添无法约束的自由，而是在有冗余的表示中建立更严格的共同约束。

杨—米尔斯理论使相互作用的身份进一步改变：不再只是物体之间的一种外加作用，还涉及不同地点的内部状态怎样比较、这些比较如何绕路、以及连接自身怎样参与动力学。这个结构成为下一章的工具，但具体标准模型仍须说明为何选择那一套群、物质和破缺结构，并用新粒子的性质逐项接受检验。某套结构能够写得自洽，只取得了成为候选理论的资格；它是否对应自然，还要由独立数据决定。理论创造与经验约束在这里并行推进，任何一方都不能把另一方的工作预先宣布完成。

本章参考资料

[1] C. N. Yang and R. L. Mills, Conservation of Isotopic Spin and Isotopic Gauge Invariance, Physical Review 96 (1954), 191–195. DOI:10.1103/PhysRev.96.191。原文：https://eclass.upatras.gr/modules/document/file.php/PHY2129/PhysRev.96.191.pdf。参见局部变换、场强及非线性场方程。

[2] David Tong, Gauge Theory，第 2 章 Yang–Mills Theory，作者课程讲义：https://davidtong.org/pdfs/teaching/gauge-theory/gauge2.pdf。参见规范冗余、自由度与闭合运输。本章矩阵与格点算例独立展开。

[3] D. J. Gross and F. Wilczek, Ultraviolet Behavior of Non-Abelian Gauge Theories, Physical Review Letters 30 (1973), 1343–1346. DOI:10.1103/PhysRevLett.30.1343。原刊：https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.30.1343。

[4] H. D. Politzer, Reliable Perturbative Results for Strong Interactions?, Physical Review Letters 30 (1973), 1346–1349. DOI:10.1103/PhysRevLett.30.1346。原刊摘要：https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.30.1346。

第四十九章 标准模型：粒子怎样成为场的激发

49.1 粒子表背后还有一套生成规则

把标准模型画成若干方格，很容易获得一种印象：物理学已经找到宇宙的全部积木，剩下的只是用它们搭出复杂物体。这样的图方便记忆名称，却遮住了更深的改变。粒子不再仅按质量、电荷和寿命分别登记，它们被放进场的表示、相互作用以及真空附近激发的共同结构中。分类因此同时规定了允许过程。

标准模型通常以SU(3)颜色与SU(2)左手弱同位旋、U(1)超荷的局部规范结构表述。这里指常用的局部群结构；全局商群等更细问题不由这行符号自动确定。它也不是把强、电磁与弱作用简单化成一个唯一耦合的力。不同部分仍有不同耦合与物质表示，但可以放在一套相容的量子场论中工作。

一九六七年温伯格的《轻子模型》明确探索通过规范场与标量结构连接电磁和弱作用，并提出其可重整化的可能性。[1]这个原始模型不是今天全部标准模型的完整初稿。我们在此把它作为一个历史节点，然后用现代形式分析质量与激发怎样共同出现，不将后来全部经验成果倒填进一篇短文。

本章的问题是：如果场比粒子表更基础，那么实验中的粒子身份究竟由什么支撑？回答不能只靠一句“粒子是场的振动”。必须说明振动围绕什么状态、哪些组合能够传播、怎样取得质量，以及它们在短暂衰变后如何留下可辨记录。

49.2 先从两个耦合场看“哪个才是粒子”

考虑两个实标量场 ϕ_1、ϕ_2，它们的动能已经规范化，二次势能为二分之一乘以 ϕ 转置、质量平方矩阵 M^2 和 ϕ 的乘积。取 $M^2=[[a,b],[b,d]]$。这里 a、b、d 的量纲都是质量平方；暂用 $\hbar=c=1$，避免把量纲转换掺入对角化问题。

如果 b 不为零，原来命名的 ϕ_1 与 ϕ_2 并不是独立传播的质量模式。质量平方本征值为 $m^2_{\pm}=[a+d\pm\sqrt{(a-d)^2+4b^2}]/2$。对称矩阵可由正交变换对角化，新组合才具有确定的自由传播质量。可稳定展开的真空要求相应二次本征值非负；若出现负值，说明选定展开点沿某个方向不稳定。

以 $a=d=m_0^2$、$b=\kappa m_0^2$ 为例，新模式是 $(\phi_1+\phi_2)/\sqrt{2}$ 与 $(\phi_1-\phi_2)/\sqrt{2}$，质量平方为 $m_0^2(1+\kappa)$ 与 $m_0^2(1-\kappa)$。当 $|\kappa|$ 小于一，两种模式都稳定；当 κ 大于一，其中一个方向不稳定。矩阵并没有只改变两张标签，它决定了哪些线性组合能够作为稳定的小激发。

若探测器与 ϕ_1 耦合，便会同时耦合到两种质量模式，而不是只看到一粒叫作 ϕ_1 的物。反过来，质量谱中出现两个峰，不一定意味着最初写下的场变量恰好就是这两种粒子。相互作用基与质量基的区别，是理解混合问题的重要入口。

这个简单模型没有重讲量子态一般的基变换。此处对角化的对象是场论中的二次动力学，目的是识别给定真空附近的传播模式；同时还须把相互作用项变换过去。若只对角化质量而忘记耦合，得到的粒子表虽然整齐，过程预测仍会错。

49.3 最低点不是一锅黏住粒子的糖浆

为理解标量背景的作用，取复标量 ϕ 及势 $V=\lambda(|\phi|^2-v^2/2)^2$，其中 λ 为正。势的最低点满足 $|\phi|=v/\sqrt{2}$。选取一个适合局部展开的表示，写 $\phi=(v+h)\exp(i\theta/v)/\sqrt{2}$。$h$ 描述径向涨落，θ 描述角向变量；这种写法在相应局部范围内方便，不是关于所有配置的全球唯一坐标。

如果没有规范场，这个连续全局对称模型会有质量为零的角向模式。加入 $D_\mu=\partial_\mu-igA_\mu$ 以后，动能项展开为一半 $(\partial h)^2$，加上一半 $(v+h)^2(\partial\theta/v-gA)^2$。角向变量与矢量势不是各自独立地随意出现，而是以共同组合进入。

通过适当局部规范选择令 θ 消去，得到矢量质量项 $g^2v^2A^2/2$，因此 $m_A=gv$；标量势的二次项为 λv^2h^2，对应 $m_h^2=2\lambda v^2$。与此同时，动能展开还产生 hA^2 和 h^2A^2 相互作用。质量与若干耦合由同一结构联系，不能为每一项另设互不相干的解释。

这不是说一个原先无质量的粒子穿过某种摩擦介质，因被拖慢而长出质量。均匀标量背景不挑选某个运动方向，质量仍进入相对论的能量动量关系。真实摩擦则通常涉及能量流向介质和优先静止系，两者不是同一机制。

也不能把选定 ϕ 的某个方向理解为可观测的绝对内部指针。规范相关的真空方向是描述选择，物理内容应由谱、相关函数和散射等合适量表达。说“规范对称性自发破缺”是常用语言，但不能把局部描述冗余当作普通全局对称性，凭空增加一批可直接测量的真空方向。[2]

49.4 纵向偏振从哪里来

上述模型的自由度计数可以逐项检查。起初一个无质量矢量有两种传播偏振，一个复标量有两个实模式，总共四个。重组后，一个有质量矢量有三种偏振，一个实标量 h 仍保留一个，总数仍为四。角向变量进入了有质量矢量的纵向组织，没有凭空产生或消灭物理自由度。

常见的“规范玻色子吃掉戈德斯通玻色子”是对这种重组的形象概括，不宜当作实际发生的一场吞食过程。粒子概念是在特定背景附近定义的激发，前后两套记号承担不同展开。若把两套记号中的全部模式同时算入粒子表，就会重复计数。

这种机制也不要求实验先看见一个自由存在的角向粒子，再看它钻进矢量粒子。可以直接通过矢量粒子的质量、纵向作用和标量耦合检验整体结构。所谓发生，不总是一段逐帧可拍的时间故事；也可能是哪些状态和过程能够共同成立的结构推导。

但是，结构推导并不能保证自然使用最简单的一个复标量。可以考虑更多标量、更复杂表示或其他机制，它们可能在某些低能量上近似一致。竞争模型需要比较额外激发、耦合关系和高能行为，而不是只凭“都能写出质量”宣布等价。

49.5 电弱混合怎样留下一个无质量组合

现在把标量推广为 SU(2)双重态 H，并取超荷 Y＝1/2。常用局部表示写 H＝(0,v＋h)的列向量再除以√2。协变导数为 D＝∂−igTₐWᵃ−ig′YB，其中 Tₐ＝σₐ/2。对背景代入|DH|²，二次矢量项为 v²[g²((W¹)²＋(W²)²)＋(gW³−g′B)²]/8。

带电组合 W±＝(W¹∓iW²)/√2 具有 m_W＝gv/2。中性部分则对应质量平方矩阵 v²[[g²,−gg′],[−gg′,g′²]]/4。它的行列式为零，因此必有一个无质量组合；另一个本征值为 v²(g²＋g′²)/4。这个零不是把光子质量手动设为零后装饰在理论上的结果，而是给定表示与背景结构的后果。

定义 sinθ_W＝g′/√(g²＋g′²)，cosθ_W＝g/√(g²＋g′²)。有质量组合为 Z＝cosθ_W W³−sinθ_W B，无质量组合为 A＝sinθ_W W³＋cosθ_W B。对应 m_Z＝v√(g²＋g′²)/2，故最低阶 m_W＝m_Z cosθ_W。

再看它如何与物质耦合。将 W³、B 反解代回 gT₃W³＋g′YB，A 的系数是 e(T₃＋Y)，其中 e＝g sinθ_W＝g′cosθ_W。电荷生成元于是为 Q＝T₃＋Y。质量矩阵的零方向与电磁耦合的共同方向相接，使“光子是什么”不再只是一张单独的无质量粒子标签。

这些关系都是指定模型中的最低阶表达。精密质量关系包含辐射修正，并依赖参数定义方案；不能拿不同方案下的弱混合角和质量直接混算，再把差别称为模型失败。独立实验确实能够严厉检验关系，但需要把输入条件统一。

49.6 同一双重态怎样产生不同电荷

左手轻子双重态可记为(ν_L,e_L)，其 Y＝−1/2。上分量的 T_3 为 1/2，因而 Q＝0；下分量的 T_3 为−1/2，因而 Q＝−1。右手带电轻子是 SU(2)单态，T_3＝0，取 Y＝−1 也得到 Q＝−1。电荷相同不代表左右分量具有相同的弱作用表示。

这一手征结构说明，不能把所有粒子都按“有多少普通电荷”排列后认为相互作用已完全确定。一个电中性对象可能仍有弱作用，一个相同电荷的左右分量可能接受不同作用。粒子身份由多个表示与耦合共同规定，而不由单一数字决定。

夸克又提供另一组合。左手双重态 Y＝1/6，上下分量分别给出 2/3 与−1/3 的电荷；它们同时具有颜色表示。这里颜色不是可见光颜色，同位旋也不是粒子在普通空间的转动。不同内部标签各承担不同变换结构，必须避免把几个熟悉词的直观混合起来。

这些超荷赋值不是理论随意散发的编号。量子规范一致性对物质配置施加反常抵消条件。不过，一条抵消式子也不能单独解释全部粒子世代、质量和参数。标准模型把许多原本无关的数据纳入共同结构，同时仍留下需要经验输入的部分。

49.7 一代粒子的分数怎样通过一致性检验

可以具体检查混合的 SU(2)平方与超荷反常系数。在一代中，左手夸克双重态有三种颜色，每个超荷为 1/6；左手轻子双重态超荷为−1/2。忽略共同的群论因子，系数为 $3\times(1/6)-1/2=0$。这里夸克和轻子看似分属不同粒子格子，却共同参与一项量子一致性条件。

再检查纯超荷三次项，统一使用左手场。右手粒子转写为左手共轭后，超荷变号；一代的和为 $6(1/6)^3+3(-2/3)^3+3(1/3)^3+2(-1/2)^3+1^3$。通分后是 1/36−8/9＋1/9−1/4＋1，结果也为零。

这个算术并不意味着分数相加漂亮便证明世界真实。它检验的是一个指定候选谱能否通过必要的量子相容性关口。若擅自删除某种手征粒子而不补偿相应贡献，原来的一致性可能破坏；若增加新的粒子，必须重新结算。分类由此具有跨格子的约束。

同样，反常抵消也不独自选定三代，更不计算电子质量。可以有其他通过相容性要求的候选结构，仍须由实验筛选。理论限制与经验选择在这里相互加强，不能将一项必要条件宣称为唯一存在证明。

49.8 希格斯作用解释质量关系，却不替我们选定全部数值

带电费米子的质量可由汤川耦合产生。以电子为例，允许的作用连接左手双重态、标量双重态与右手单态。把标量背景代入，得到 m_e＝y_ev/√2，同时剩下 h 与电子的耦合强度 m_e/v。质量项与标量作用共享同一参数，因而可以跨过程比较。

但是，y_e 的大小并未由这一式子计算出来。若每个物种都有各自汤川矩阵，质量层级与混合仍包含经验输入。说“希格斯机制解释了电子为什么有质量”，不等于说“它已从零输入推导电子质量数值”。前者解释一种结构，后者要求更强的理论。

尤其不能把普通物体的全部质量都写成希格斯直接赠送的重量。核子质量主要涉及强相互作用的动力学尺度和束缚系统的总能量，不能简单等于内部夸克静质量之和。希格斯对基本费米子质量的作用，与复合物体质量的形成，是相关而不相同的层次。

这一点并未取消质量能量关系。恰恰因为系统静质量由质心系总能量决定，内部运动、场和相互作用都进入结算。最小组成者的质量表不足以直接给出复杂对象的质量，也不意味着复杂质量永远不能从理论计算。真正的难处在于动力学及其可控求解。

49.9 短程作用怎样从有质量传播产生

有质量媒介的另一个后果，可以从静态响应直接算出。设一个标量示范模型满足 $(-\nabla^2+M^2)G(r)=\delta^3(r)$，仍采用自然单位。三维空间的解为 $G(r)=\exp(-Mr)/(4\pi r)$。当 M 趋零，它回到熟悉的一比距离形式；当距离远大于一比 M，指数因子显著压低响应。

这个式子说明“作用短程”怎样与传播结构联系，但不把完整弱作用简化成普通中心势。实际弱作用具有手征、电荷交换和不同过程，不能仅凭指数势推出全部衰变规律。本例只展示质量尺度怎样控制空间传播的典型范围；恢复单位后，该长度是 $\hbar/(Mc)$。

同一结构在动量空间更容易连接低能理论。若某个交换振幅含有 $g^2/(M^2-q^2)$，且 $|q^2|$ 远小于 M^2，可以展开为 g^2/M^2 乘以 $[1+q^2/M^2+$后续项$]$。首项在低能任务中表现为接触作用，高阶项记录有限传播范围。于是低能观察者可以暂时不保留显式媒介场，却仍通过有效系数感受它。

这也产生一种识别上的退化：只知道首项 g^2/M^2，不能分别确定 g 和 M。一个更重而耦合更强的候选，可能与较轻而耦合较弱的候选给出相同最低阶低能响应。提高动量转移、测量下一阶依赖或直接产生媒介，才可能解除这种退化。不同能标的实验不是重复同一个故事，而是在辨识链中承担不同任务。

因此，标准模型的统一还包含尺度之间的接续。低能有效作用不因为没有显式写出全部粒子便是错误；高能理论也不能只凭包含更多场就宣称解释了低能。二者必须在共同适用范围内匹配同一响应，并说明截断误差怎样随尺度比增长。

49.10 负质量平方并不是一种超光速粒子通行证

前面质量矩阵的例子出现过负本征值。对此应区分稳定真空附近粒子的质量，与任意展开点势能曲率。若在势顶附近写出负二次曲率，线性化方程说明微小扰动会增长；它首先表明这个展开点不稳定，不是实验已经找到具有虚质量的自由粒子。

取最简单的实场势 $V(\phi)=-\mu^2\phi^2/2+\lambda\phi^4/4$，$\mu^2$ 和 λ 均为正。原点的二阶导数为 $-\mu^2$，因此原点不是稳定最低点。最低点位于 $\phi=$ 正负 $\mu/\sqrt{\lambda}$，在任一最低点附近写 $\phi=v+h$，二阶曲率成为 $2\mu^2$，局部小激发具有正质量平方。负号在寻找真空时有作用，稳定激发的质量由正确展开点决定。

这个全局实标量示例与规范模型有不同的对称性内容，不能直接把两个最低点的全部解释搬到电弱理论。它只澄清曲率的数学角色。对于规范模型，某些表面不同的方向属于描述等价类，物理真空与谱的讨论需要相应约束。

这一点也约束思想史的叙述方式。不能把理论中每一个不熟悉的符号立即命名为一种新物种；应先问它出现在哪一级近似、围绕什么状态、对应稳定激发还是不稳定方向。粒子的创造不是为形式可能性逐个登记，而是建立能够延续到可观察过程的稳定规定。

49.11 一个极短寿命粒子怎样留下质量峰

二〇一二年，ATLAS 与 CMS 的原始发现论文报告了约一百二十五吉电子伏附近的新玻色子信号，并比较了多个衰变通道。[3][4]发现不是让一个希格斯玻色子停在显微镜下拍照。它依靠末态的能量、动量与统计分布重建一个不直接持久留下轨迹的激发。

以两个光子为例，设各能量为 W_1、W_2，夹角为 θ，其共同不变质量满足 $M^2c^4=2W_1W_2(1-\cos\theta)$。这可以由两个无质量四动量相加后平方直接推出。如果母粒子在自身静止系衰变为两光子，二者反向且能量相同，便得到各为 $Mc^2/2$。

不变质量不要求母粒子在实验室静止。整体推动改变单个光子的能量与夹角，但上述组合保持，因此不同运动状态的母粒子仍可在重建分布上聚成相同质量位置。这使短寿命粒子的身份能够跨事件显露，而不依赖逐个看见它在两次相互作用之间旅行。

然而，每两个光子都能算出一个不变质量，并不是每一对都来自同一种新粒子。背景过程也产生光子对。证据需要比较候选信号形状与连续背景、校准能量响应、检查选择条件，并寻找其他通道的相容信号。质量峰是一组统计关系，不是图上一个突出的柱子就能自行完成命名。

峰的宽度也不能随意等同寿命倒数。理想的不稳定态宽度与寿命有 $\Gamma\tau$ 约为 $\hbar$ 的关系，但实验峰还受探测器分辨率和重建影响。若自然宽度远小于仪器分辨，观察到的宽峰主要反映仪器，不能直接读出自然寿命。理论量与重建量需要经过响应模型连接。

49.12 发现一种新玻色子之后还要确认什么

一个质量峰首先支持“存在额外过程或共振”的判断，随后才是更精细的身份研究。不同自旋、宇称和耦合结构可以产生不同角分布与通道比例。若候选质量相同，却对这些独立量预测不同，就有机会通过更多数据分辨。发现与身份确认是相连但不重合的任务。

把某个标量候选与最简标准模型相比较，可以定义信号强度 μ_{sig} 为测得的产生率乘衰变分支比，相对于参考预言的比值。若 μ_{sig} 接近一，它支持该任务上的相容性，却不自动证明所有耦合分别等于标准模型值。产生与衰变的不同改变可能在乘积中抵消。

举一个简化例子：某通道产生截面变为参考值的两倍，衰变分支比变为一半，则两者乘积保持不变。只测这一个通道的总事件数，无法区分。加入另一个产生机制或另一个末态，才可能解除退化。高精度不只是把同一数字测得更多位，还包括选择能够打破参数退化的新任务。

同理，若增加未被看见的衰变通道，总宽度改变，已知末态的分支比也会变化。不能把一个通道偏低直接解释为对应耦合变小。对新模型的检验，需要把总宽度、产生方式和多个末态共同放进一致分析。场的激发身份由这种关联结构逐步获得约束。

这使标准模型的成功保持开放性。数据可以强烈支持某组关系，同时仍允许更高能的新自由度或很小的额外作用。说它在已检验范围内成功，与说它是永恒且封闭的全部物理，是不同结论。模型名字中的“标准”不提供超出证据范围的豁免权。

49.13 中微子使粒子清单重新打开

若严格采用没有右手中微子、也没有高维质量算符的最简重整化版本，中微子质量为零。中微子振荡所要求的质量平方差意味着这个最简版本需要扩展。这里的修订可以有不同结构，不只是给一个原有方格填上新的小数字。

在两味理想模型中，味态由两个质量态按角 θ 混合。在真空、相对论近似及相干条件下，转化概率为 $\sin^2(2\theta)\sin^2(\Delta m^2 L/(4W))$，这里继续使用自然单位，W 为粒子能量。两个质量态累积不同相位，味读出的概率随距离和能量比改变。

这个公式说明了能知道什么，也说明了不能知道什么。振荡依赖质量平方差，不直接给出绝对质量尺度；两个质量平方同时加上同一常数，并保持非负，质量平方差仍不变。振荡也不独自决定中微子是否为自己的反粒子。需要不同实验任务补充这些身份信息。

因此，“场的激发”不是把所有粒子变成同一种简单波纹。不同场的表示、质量类型、混合、可观察过程各有具体内容。标准模型提供共同语法，未决问题则要求继续扩充或修订语法中的某些规定。共同结构与未完成性能够同时成立。

49.14 从一座峰到一种场激发的发生路径

本章可以沿 S—E—D 进入路径走完一次身份研究。初始显露 S 是一组末态中稳定的额外分布；先进入 E，列出能够同时解释质量、角分布和产生衰变关联的候选场结构；再转到 D，选择最能区分候选的通道、分辨条件与独立控制。路径终点不是一句漂亮解释，而是下一轮可败的试验安排。

S＝F(D,E)要求峰形既受物理过程也受选择与响应共同规定。D＝G(S,E)要求不能仅按容易积累事件的通道重复，而应按照候选之间的实际差别安排测量。E＝H(S,D)要求新的约束改变对耦合、混合与可能未见通道的组织。这里的三式是研究接口，不能替代质量矩阵和衰变振幅的定量推导。

当 S 与 D 在多套控制后仍保留冲突，推动 E 检验额外激发或错误耦合，并以独立通道回检；当 D 与 E 无法产生足够区分力，推动建立新的 S，例如角关联而非总事件数；当 S 与 E 互不相容，推动 D 改变产生模式或重建条件，检查残差是否来自共同校准。每一次推进第三维，都需要返回原有证据，确认它既减少旧矛盾，也没有制造新的任意参数。

在这个意义上，粒子成为场的激发，并没有使物理对象变得更虚幻。对象的规定更严格了：名称必须连接质量谱、表示、耦合与可观察过程。一个新峰不由意识宣布为新

粒子，一套新场也不因数学优美自动进入自然。理论与实验共同形成可以继承、比较和反驳的身份结构。

标准模型的思想贡献正在这里：它将大量粒子的差别与相互作用的共同规则组织起来，同时迫使我们承认参数输入、近似层级与未决边界。下一章将专门追问这些参数为何称为常数、这些数字怎样依赖测量定义，怎样与其他参数共同组成可检验的关系，它们又怎样跨尺度规定同一物理，以及“常数发生变化”到底需要什么样的证据。

本章参考资料

[1] S. Weinberg, A Model of Leptons, Physical Review Letters 19 (1967), 1264–1266. DOI:10.1103/PhysRevLett.19.1264。原文：https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.19.1264。

[2] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Physical Review Letters 13 (1964), 508–509. DOI:10.1103/PhysRevLett.13.508；David Tong, The Standard Model，第 5 章，https://davidtong.org/pdfs/teaching/standard-model/standardmodel5.pdf。本章采用现代记号独立展开简化模型与质量矩阵。

[3] ATLAS Collaboration, Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC, Physics Letters B 716 (2012), 1–29. https://arxiv.org/abs/1207.7214。本文只引用发现论文的信号性质与多个通道，不采用其数值作为后续最新测量。

[4] CMS Collaboration, Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC, Physics Letters B 716 (2012), 30–61. https://arxiv.org/abs/1207.7235。

第五十章 物理常数与尺度
什么在变化中保持，什么需要重新规定

50.1 常数为何不只是公式里的固定数字

上一章的标准模型用有限的结构和参数约束许多过程，但参数仍须通过测量确定。于是，一个古老问题重新出现：为什么自然选择这些数值？另一个更直接的问题却应先回答：当物理学说某个量是常数，它究竟承诺了哪一种不变？是换单位不变，换能标不变，换时间不变，还是在同一模型的一类过程里可共用同一个参数？

这些问题如果不分开，很容易出现表面深刻、实际混乱的结论。推荐数值更新，被说成自然本身漂移；耦合随能标运行，被说成宇宙年龄使力衰弱；单位定义固定某个数，被说成科学家投票让自然停止变化。思想史中的常数观，需要在这些不同任务之间建立明确边界。

狄拉克一九三七年的《宇宙学常数》把基本常数与宇宙尺度问题联系起来，是这一问题史中的一个节点。[1]但数值相近或量级巧合本身不是变化定律。若要从关系猜想走向物理理论，必须给出哪些无量纲组合随什么变量改变、变化怎样进入动力学，以及什么观察能够否定它。

本章从单位和比值开始，经过能标运行、参数估计与时空变化检验，最后把常数问题接到材料中的有效参数。重点不是列一张最全的数字表，而是说明同一个“常数”概念怎样在不同层级承担不同责任。

50.2 把米换成厘米，自然没有加速

任意有量纲物理量都可以写成数值乘以单位。若速度单位从米每秒换成厘米每秒，同一速度的数值增大一百倍；这不表示运动突然加快。单位变换改变的是表示中的数，物理比较关系保持。讨论常数数值的意义，必须先说明单位。

一般地，若基本长度、时间和质量单位分别变成原来的 a、b、d 倍，具有量纲 L^rT^sM^u 的同一物理量，其数值便乘以 a^(−r)b^(−s)d^(−u)。因此可以通过单位约定把某些有量纲常数写成一。自然单位中的 $c=1$、$\hbar=1$ 不是宣布光速或量子作用量消失，而是将它们作为不同量纲之间的换算关系。

这也说明，单独说“光速数值减小”可能还没有可检验内容。若尺与钟的实现随某种新动力学一起变化，需要说明实际比较的是什么。实验读出的最终是两个过程、长度或频率之间的比值；无量纲组合能够避开仅由单位重标造成的变化。

以精细结构常数 $\alpha = e^2/(4\pi\varepsilon_0 \hbar c)$为例，在国际单位制下它无量纲。另一种常用电磁单位写法可能省去 ε_0 或改变电荷归一化，但同一物理 α 不因此改变。质子与电子质量比 m_p/m_e 也是无量纲量，换公斤为克不会改变它。

这并不意味着有量纲常数毫无物理意义。它们在理论、量纲桥接和实验实现中十分重要。较准确的要求是：声称它们发生物理变化时，应给出完整理论与测量约定，并能翻译成不依赖任意单位选择的可观察关系。

50.3 定义常数与测量常数不能混为一类

国际单位制自二〇一九年五月二十日起，用七个定义常数的固定数值规定单位系统，其中包括真空光速、普朗克常数和基本电荷。[2]这种固定改变的是单位定义方式，不是用决议替自然定律作实验。单位的实际实现仍需要装置、校准和不确定度评估。

以长度为例，秒有相应实现，米则通过固定的 c 数值与光传播相接。于是，在这一单位体系中再报告“测量 c”的工作，常常实际检验的是长度、时间实现及传播条件的一致性。若不同实现发生冲突，需要研究具体物理和计量环节，而不是机械地从定义精确推出所有实验都精确。

定义值没有数值不确定度，实验实现却有。这两句话可以同时成立。一台装置以固定常数为基础，不表示它没有温度、几何、噪声和模型误差。把定义精确当成仪器无限精确，会消除科学测量最需要说明的部分。

定义改变还会转移哪些量需要实验确定。例如过去被定义约束的某些电磁常量，在修订单位制后可以通过其他常数关系与测量得到。[2]这不是物理作用忽然改变，而是约定与测量的分工重排。同一套自然关系允许多种一致的单位组织。

因此常数在科学共同体中也承担协调作用：不同实验室能够通过可复现过程连接自己的尺、钟和电学标准。协调需要制度约定，但约定的可用性持续接受自然过程约束。它既不是纯主观约定，也不是无需实践即可从天而降的数字。

奎因关于从实物基准转向基本常数的计量史研究，将这一转变放在长期稳定、全球可获得的标准与国际协调结构之间考察。[5]这个历史角度提醒我们，定义的变化不是单独一行数学等式的替换。对本书而言，更应继续追问：新的实现路线能否由不同实验室

重复，旧标准怎样与新标准比较，误差从哪里进入。只有这些接口接通，抽象的固定数值才成为共同使用的物理尺度。

50.4 一个能标变化怎样保留同一物理预测

量子场论中的耦合运行属于另一个层次。取一项无量纲观测量的简化有限阶表达 $R(Q)=g(\mu)+b\,g^2(\mu)\ln(Q/\mu)$。$Q$ 是实际过程的典型动量尺度，μ 是计算中选择的重整化尺度。要求 R 在所保留精度内不随任意 μ 改变，给出 $dg/d\ln\mu=b\,g^2+$更高阶项。

检查这个符号：对 R 求 $\ln\mu$ 导数，第一项给出 β，显式对数给出 $-bg^2$，二者相消到这一阶。因此，g 随着 μ 改变，并不意味着同一个物理过程随我们的计算选择改变。参数的隐含依赖与公式的显式依赖一起保持观测量。

将微分方程积分，可得 $1/g(\mu)=1/g(\mu_0)-b\,\ln(\mu/\mu_0)$。若选择 μ 接近 Q，对数变小，有限阶表达通常更方便。若选择相距很远的 μ，对数可能很大，需要更多阶或重求和。改变 μ 是一种组织计算的动作，不能把它画成宇宙在时间中推动常数的演化。

另一方面，真实改变 Q 可以改变可观察响应。不同动量转移的散射探查不同尺度的相互作用，相关有效耦合可能表现出能量依赖。这里依然是同一理论在不同过程尺度下的关系，不是理论参数表每过一天自动改写。两种“变化”必须用不同自变量标明。

若新物理真的使背景场随宇宙时间变化，可能需要写 $g(\mu,t)$，并分别分析对 μ 的运行与对 t 的动力学。前者不能代替后者的证据；后者也不能绕过前者的计算条件。把两个变量写在同一个式子里，反而能看出混淆为什么不成立。

50.5 阈值让不同尺度使用不同有效说明

运行也不是把同一个简单公式无限延伸。设某粒子质量为 M，在远低于 M 的过程里，它未必能够被直接产生，其影响常可压缩到低能有效系数；越过相应尺度后，需要显式计入它的贡献。有效理论的物质内容改变，运行方程中的系数也可能改变。

一个分段模型足以显示匹配的必要性。假设低能区 $d(1/g^2)/d\ln\mu=2b$ 低，高能区为 $2b$ 高。在阈值附近指定匹配尺度 M，先用低能方程把已知输入运行到 M，再按所用方案完成匹配，随后用高能方程继续。若忽略连接条件，两段曲线可以任意错开，便不再是一项完整预测。

最低近似有时可以直接令两边耦合在 M 相等，更高精度则需要有限阈值修正。这个修正不是让相互作用在某个任意计算刻度突然跳跃，而是不同有效描述分配自由度的方式不同。完整物理量应在相同精度内接续。

可以把它与地图尺度作有限类比：远处不需要画出每条街，近处必须加入；但物理匹配比地图比喻更严格，因为遗漏自由度的影响要以算符和系数保留下来。一个高能粒子没有出现在低能末态，不代表它对低能一切量都毫无作用。

这使“基本常数”和“有效常数”之间出现层级。某个低能系数可能由更高层的质量、耦合和对称结构共同决定。在低能实验中它表现为可共用的常数，在高能描述中却可能是若干参数的组合。常数的层级变化不等于自然关系变得任意。

50.6 数据表更新为什么不能证明常数漂移

CODATA 推荐值通过理论与实验数据的联合调整形成；二〇二二年度调整的正式说明明确采用最小二乘方法，并讨论输入之间的不一致。[3]因此，一张常数表不是每个数字都由一台仪器独立直接读出，也不是新的推荐值出现便证明旧时代的自然具有不同常数。

一个简单估计例子可以说明。若两次独立测量得到 x_1、x_2，不确定度为 σ_1、σ_2，在高斯误差模型下，共同参数的加权估计是$(x_1/\sigma_1^2+x_2/\sigma_2^2)/(1/\sigma_1^2+1/\sigma_2^2)$。加入精度更高的新测量，推荐值移动是正常的推断更新。它描述知识怎样改变，不直接描述对象怎样改变。

若两项测量之间有共同系统误差，上述独立加权便不充分，必须使用协方差矩阵。相关性可能使看似两份证据实际上共享同一个标定来源。增加实验数量不总能按平方根规律降低全部不确定度，尤其不能把共同偏差当成独立随机噪声平均掉。

检验真实时间漂移，需要保留原始测量时刻、方法、误差与共同系统项，建立类似 $x(t)=x_0+kt$ 的明确模型，再比较 k 是否获得独立支持。拿两个版本的推荐表直接相减，往往把数据集合、理论修正和统计方法的变化一起混进来，不能达到这个目的。

如果不同时期的独立实验确实给出稳定、可复现且排除已知系统效应的趋势，当然不能用“常数就该不变”把它取消。问题是证据必须针对物理漂移，而不是只针对估计值变化。名称不是论证，反常也不是免除论证的理由。

50.7 两种钟比较究竟约束了什么

原子钟比较提供一种清晰的无量纲检验。设两个跃迁频率之比为 $R=\nu_1/\nu_2$，在某个参考参数附近，小变化可以写成 $\delta\ln R=K_\alpha\,\delta\ln\alpha+K_\eta\,\delta\ln\eta+$其他项，其中 η 可表示某个质量比，K 是相应灵敏度系数。这些系数由具体原子结构计算决定，不能仅凭两个频率都很稳定便猜出。

如果两个跃迁对 α 具有完全相同的依赖，那么频率比中这项共同变化会抵消。高稳定性未必等于高灵敏度。选择灵敏度不同的钟，是让待检变化从共同因素中显露的关键。实验设计不只是寻找最好的钟，还要寻找对所问参数具有不同响应的组合。

再考虑两个待估变化 $x=\delta\ln\alpha$、$y=\delta\ln\eta$。若两项比值变化分别满足 $r_1=2x+y$、$r_2=-x+y$，那么 $x=(r_1-r_2)/3$，$y=(r_1+2r_2)/3$。两条独立响应方向才把两个未知量分开；若两项测量的系数都成比例，无论单项精度多高，仍留下退化方向。

这个例子也显示联合误差的重要性。若 r_1 与 r_2 共享同一个参考钟，它们的噪声可能相关。求 x 的误差时必须保留协方差，不能把两个标准差机械平方相加。更不能把某一组合的零结果解释成所有基本参数都绝对不变，它只约束灵敏度矩阵真正看得见的方向。

Uzan 对常数变化检验的系统研究强调，原子、天文与引力等不同体系对参数的依赖和系统条件各不相同。[4]这里的两钟线性模型只是把这种比较要求具体化，不使用某个旧综述中的数值上限冒充当前最强约束。

50.8 宇宙光谱为什么不能只比较颜色偏红

来自遥远天体的谱线为跨宇宙时间检验提供另一入口。但观测频率包含整体红移、运动和仪器响应，不能看到谱线偏移便直接宣布常数改变。必须寻找不同跃迁相对于共同红移的差异结构。

设同一个吸收系统中两条线的静止波长为 λ_1、λ_2，在最简共同红移模型里，观测值均乘以 $1+z$。其比值 λ_1 观/λ_2 观等于静止比值，整体红移消去。如果 α 的变化对两条跃迁产生不同相对位移，这个比值才可能提供额外信息。

实际问题比两线模型更困难。不同离子的速度成分可能不完全相同，同位素组成改变可以移动多条成分叠加后的谱线中心，波长校准也可能随频段产生系统趋势。这些效应若与候选变化模式相似，就需要更多跃迁、独立仪器或不同吸收系统共同约束。

因此，跨时间不等于无条件地看见过去的常数。观测把古老相互作用、光传播和今日仪器接在一起，推断必须分离这些环节。能够经受多种方法交叉检验的参数变化，才具有超出某一拟合的物理意义。

同时也不能反向声称，只要存在系统误差可能性，天文检验就没有价值。科学任务正是把竞争效应转成不同的可比预言：某项校准误差可能跟随波段，某种速度结构可能跟随离子，某个基本参数变化则受共同灵敏度关系约束。增加正确的差别，比笼统争论相信或不相信更有用。

50.9 如果常数真的变了，理论需要支付什么代价

一个可工作的变化模型不能只把原方程中的常数替换成任意时间函数便结束。例如在场论中引入标量 χ，让某个耦合成为 $g(\chi)$。如果 χ 随时空变化，耦合便可能变化；但 χ 本身必须有动力学、能量和与物质的作用，不能成为一个只在需要拟合时才出现的外部旋钮。

这样的模型通常还会带来其他后果：不同物质组成对 χ 的响应可能不同，产生额外作用或改变自由落体比较。具体是否出现、大小如何，依赖模型与屏蔽机制等条件，不能用一个抽象公式作普遍数值断言。重要的是，常数变化应与其他可检验后果组成共同理论。

如果候选模型用一个 χ 解释某项钟比变化，却对另一种钟、天体运动或组成依赖作出相反预言，它可能被排除。若为了每个失败再加入一项独立函数，预测能力就会快速下降。新的自由度可以合理增加，但必须说明新增解释力与新增输入之间的关系。

这也界定了大数巧合的地位。两个极大无量纲数相近，可以提示一个候选关系；要成为动力学，应说明相近为何保持、是否随宇宙演化改变、与其他参数如何相容。单靠数位排列无法区别必然关系与在许多可选组合中找到的巧合。

常数问题由此通向一种严格的创造观：创造新理论不是为固定数字编一个故事，而是建立能把多个独立变化一起约束的结构。一个理论若允许任何结果都被叫作常数的“发生”，便已经放弃常数概念原来承担的跨过程约束。

50.10 材料中的有效质量为什么不推翻电子质量

到这里，还应把“随条件变化的有效参数”与基本常数变化区别开。设晶体中某能带在极小值附近为 $\varepsilon(k)=\varepsilon_0+\hbar^2k^2/(2m^*)$。这里 m^* 由能带曲率定义，描述特定状态附近的运动响应；它可以随材料、能带甚至方向改变，不意味着真空电子的静质量在不同晶体中被任意重写。

以一维最近邻链为例，$\varepsilon(k)=\varepsilon_0-2t\cos(ka)$。在 $k=0$ 附近展开，得到 ε 约为 $\varepsilon_0-2t+ta^2k^2$，因此 $m^*=\hbar^2/(2ta^2)$，假设 t 为正。改变跃迁参数 t 或间距 a，便改变 m^*。这个有效质量来自电子与周期结构共同组织的允许状态，不能只由孤立电子质量表读出。

在能带顶部，曲率可以为负，相应电子有效质量也可为负。这不是产生负静质量反物质；它说明在该带的半经典近似下，速度对外力引起的波矢变化呈相反曲率响应。使用空穴描述时又可将相关响应重新组织。适用范围由能带、散射和外场条件共同决定。

材料的介电常数、弹性系数和热容也属于类似的条件化规定。叫作常数，常常表示在限定频率、温度、应变或弱场范围内可以近似共用某个数。跨过相变或明显非线性区域后，它可能变成函数或需要替换模型。不能因为它们并非宇宙普适常数，就否认它们具有可重复的物理意义。

这正是从基本粒子理论通往下一章材料世界的接口。不同层级的稳定性有不同来源：一项参数可以在低能理论中稳定，在材料条件改变时有规律地变化，又在更深理论中被解释为共同作用的结果。层级说明增加解释责任，并不许可随意命名。

50.11 以跨尺度一致性检验共同规定

沿 E—S—D 锚定路径，先从参数、单位实现、尺度运行与过程定义的共同制约 E 出发，锚定应该比较的无量纲显露 S，例如两种钟的比值或两能标下同一响应；再安排差异运行 D，改变能标、时间或材料条件中的一项，同时控制其余项。只有明确改变了哪个变量，才能判断观察到哪一种变化。

S＝F(D,E)要求比值的意义依赖实际比较任务与结构条件。D＝G(S,E)要求异常趋势与候选模型共同决定下一轮控制。E＝H(S,D)要求经操作检验保留的差异推动参数定义或动力学修订。三式在这里特别约束“常数变化”这个诱人的短语，避免让它覆盖彼此不同的工作。

若 S 与 D 显示稳定时间趋势而旧参数组织无法容纳，张力推动 E 引入并约束新的动力学，再到独立系统回检。若 D 与 E 只共同改变了单位或任意重整化尺度，就应推动建立真正不变的 S 来检查是否仅为表示差别。若 S 与 E 在不同测量中冲突，则推动 D 检查共同校准、阈值和模型输入，随后再判断是否需要改变基本理论。

常数的思想意义因而不是把世界冻结。它是在允许变化中找到可跨过程继承的关系，并说明这些关系何时需要重新规定。知道一个参数在什么条件下保持，和知道它在什么条件下变化，是同一研究责任的两面。

从狄拉克到量子电动力学，再到规范场和标准模型，物理学逐渐把粒子身份放进越来越丰富的关系系统；常数则让这些关系能够落到数字并跨实验接续。接下来进入晶体与材料，问题不会退回简单的组成清单。新的组织、占据和边界将继续产生新的有效身份，而我们仍须追问：哪些关系稳定地留下，哪些变化迫使“这一类物”获得新的规定？

本章参考资料

[1] P. A. M. Dirac, The Cosmological Constants, Nature 139 (1937), 323. DOI:10.1038/139323a0。原刊书目与摘要：https://www.nature.com/articles/139323a0。

[2] CGPM, Resolution 1 of the 26th CGPM (2018), On the Revision of the International System of Units. https://www.bipm.org/en/committees/cg/cgpm/26-2018/resolution-1。参见生效日期、定义常数及单位定义附录。

[3] P. Mohr, D. Newell, B. N. Taylor and E. Tiesinga, CODATA Recommended Values of the Fundamental Physical Constants: 2022, Reviews of Modern Physics 97 (2025), 025002. https://www.nist.gov/publications/codata-recommended-values-fundamental-physical-constants-2022。本文引用联合调整的方法性质，不列最新数值表。

[4] J.-P. Uzan, Varying Constants, Gravitation and Cosmology, Living Reviews in Relativity 14 (2011), 2. https://arxiv.org/abs/1009.5514。用于常数变化检验、模型与系统条件的分类；本文线性灵敏度、两线红移与运行模型均独立展开。

[5] T. Quinn, From Artefacts to Atoms—A New SI for 2018 to Be Based on Fundamental Constants, Studies in History and Philosophy of Science A 65–66 (2017), 8–20. DOI:10.1016/j.shpsa.2017.07.003。摘要与书目：https://pubmed.ncbi.nlm.nih.gov/29195652/。用于标准转变的历史问题定位。

第十一编 复杂物质与宇宙：整体何以产生新的物理

第五十一章 固体与材料
晶体、能带和“这一类”的新身份

51.1 把一块固体拆完，材料还在不在

握住一块铜，人会感到重量、凉意和坚硬；把它接进电路，又会得到可测的电流。这些特征看似挤在同一块物体里面，因而很容易被理解为铜原子各自携带的小份性质：小份相加，就成为大块材料的性质。但是，单个铜原子没有一根铜导线的电阻；一个孤立原子也没有试样在两电极之间表现的电导率。组成材料的原子身份当然重要，然而组成者的清单还不能替代材料身份的成立过程。

固体物理学把这一差别推进到可以计算、制造和检验的程度。它不只问材料里面有什么，还要问这些组成者以怎样的空间秩序、占据规则和相互作用共同成立，怎样与光、热、电极和应力接续。材料成为“这一类”，既涉及组成，也涉及组织和可重复的响应。这个问题在元素分类之后展开，却不能被元素分类收回。

本章用“新身份”指一种受到实验约束的物理分类能力：原来只能凭外观、手感或化学成分辨认的固体，如今可以凭衍射、能谱和输运的互相校验取得身份。这不是给同一物体多挂几张说明牌。新的制备与测量系统使原来不能稳定区分、不能控制生成的材料差别成为物理世界中的有效差别，并进一步让这些差别进入器件设计。

按照本书的 SDE 主轴，改变的也就不只是看材料的眼光。原子排列、探测过程、信号转换和理论推断被重新组织，新的视觉与触觉特征纠缠聚合体由此成立。屏幕上的峰不等于手中的硬度，但两者能够经有条件的推断、制备和复测相接。一种可研究、可制造、可失败的新材料世界正是在这些接续中形成的。

51.2 衍射使内部排列成为可以争论的对象

晶体外形的规则性早就令人注意。问题是，外形整齐是否意味着内部按某种方式排列；即使如此，又凭什么在数种内部模型之间选择？只画一幅小球整齐堆放的图，不能回答这个问题。内部秩序必须在某种与它发生作用的过程中留下区别，而且不同的候选排列需要给出不同的可比较结果。

劳厄发现晶体能够产生 X 射线衍射之后，布拉格父子的研究把衍射与晶体结构分析紧密连接起来。劳伦斯·布拉格的诺贝尔演讲以晶体的 X 射线衍射为题，直接讨论通过这种方法研究固体内部原子排列的进展。[1]这里真正具有发生学意义的动作，是把几何候选、波的相位差和记录装置结合成能够排除错误排列的过程。

在最简单的几何图景中，设相邻晶面间距为 d，入射线与晶面的夹角为 θ，波长为 λ，相邻晶面反射路径的光程差为 2d sinθ。相长条件写为 2d sinθ=nλ。这个表达不是一把可以离开仪器单独使用的尺子。角度零点、波长标定、样品取向以及哪些峰属于样品，都必须先获得可检验的规定。一个峰的位置帮助约束晶面间距，不能独自告诉我们全部原子怎样排列。

设想两个候选结构具有相同的某组晶面间距，却在晶胞内部放置了不同的原子。它们可能在若干角度上都满足上述条件，但不同散射贡献相加时会增强或抵消，造成强度与消光的差别。这就说明：周期尺度与内部组成不是同一项信息。若只保留峰位、删去强度与样品条件，再宣布“已经看见原子”，便把推断链中最重要的约束删掉了。

更值得注意的是，衍射图与结构图之间并非照相式的一一对应。有限数据、噪声和相位信息缺失都会约束结构重建。研究者必须问，当前模型还能解释哪些未用于拟合的信号；更换入射条件或采用另一种探针，结论是否保持。所谓结构的客观约束，正通过可被他人接续和反驳的互动过程获得力量，而不因承认这一过程就消失。

因此，晶格不是随意想象的后台脚手架，也不是看一眼便完整显露的实体。它是许多不同测量差别能够在同一组织方式下获得协调解释的结构规定。结构身份的稳定，需要跨过程的一致性；一幅漂亮图像只完成了其中很小的一部分。

51.3 从原子能级到能带：整体允许什么

在原子问题中，人们学习辨认离散能级。到了固体，若把每个原子仍当作彼此无关的孤岛，就无法解释电子在整个材料中的行为。关键变化不是原子数变得太大，算起来麻烦，而是相邻位置之间的耦合改变了允许状态本身。能带由此进入材料身份：它描述在一定结构和近似条件下，电子可以具有怎样的能量与运动状态。[2]

下面给出一个供本章分析使用的简化模型。取一条具有相同间距 a 的链，每个位置只有一个相关轨道，位置能量为 ε_0，最近邻之间的跃迁参数为 t。将第 j 个位置的振幅记作 c_j，模型方程为 $\varepsilon c_j=\varepsilon_0 c_j-t(c_{j+1}+c_{j-1})$。这里的 t 不是一个电子每秒跳跃多少次的直接记录，而是哈密顿量中描述相邻轨道耦合的参数。

若选取 c_j 与 $\exp(ikja)$成正比的状态，代回方程可得 $\varepsilon(k)=\varepsilon_0-2t\cos(ka)$。于是，同一组位置能量因为连接方式而展开为一段允许能量。令 t 为零，位置间不再通过此项耦合，能量范围收缩；改变 t，能带宽度随之变化。即使不增加新的原子种类，允许状态也会变化。这一演算直接揭示，材料的电子身份不能只由孤立原子的标签决定。

这个模型还允许读者看清“更多”何以成为“不同”：增加无关的原子，只是增加重复份数；把这些原子置入可延展的耦合组织，才使跨位置状态成为问题。由此不能推出所有宏观性质都不可从微观理论导出。较准确的结论是，导出宏观性质必须把相互作用、组织方式和适用条件带入；单个原子的性质表不是足够的输入。

周期结构中的布洛赫表述把这种思想一般化：波函数可以写成平面波因子与具有晶格周期的函数之积。它提供的是周期哈密顿量的状态组织原则，不是对任意实际材料的无条件保证。[2]真实样品有边界、缺陷和温度变化；强关联也可能使简单的独立电子能带解释不足。把能带理论的成功写成“晶体里所有问题都已经解决”，会遮住固体物理学继续发展的理由。

这一节的公式还提示一个记号纪律。物理能量用 ε 表示，本书 SDE 中的 E 表示特征纠缠；二者不能因为常见教材都用字母 E 就混为一谈。能带可以参与某个 SDE 的特征纠缠组织，但能带本身不是 E 维度的同义词。

51.4 金属和半导体不是原子的两种脾气

有了允许能量，还必须问哪些状态被电子占据，以及外场施加后哪些改变能够发生。在简单能带图景中，费米能附近是否存在可用状态，是理解金属与能带绝缘体差别的关键之一；半导体则使载流子生成、掺杂与控制成为可以组织的工作对象。[3]这里不能把导电简化为“电子松一点就跑，紧一点就不跑”。占据规则与可用状态共同约束响应。

为了看出这一点，可以沿用前节的链模型，比较两种不同填充情形。即使允许能量函数相同，低能状态被占满到哪里不同，可供小扰动重新分配的状态就可能不同。因此，同一个几何组织不能脱离电子数和占据条件独自决定电学身份。反过来，仅仅报告电子数量也不够，因为能量状态的组织已经由耦合与结构改变。

材料分类在这里从单一标签走向条件化判断。说某试样是半导体，需要说明所研究的组成和结构、温度范围及测量方式。把它加热、照光或改变掺杂之后，电流变化并不是它忽然“背叛本性”，而是相关的载流子过程发生了变化。一个可靠类别应该能够说明哪些变化仍属于该类别的正常响应，哪些变化意味着进入了另一种相或另一套机制。

这里还有一个容易混淆的地方：电阻是具体试样和测量配置中的量，电阻率是在一定均匀性和几何假定下用来比较材料的量。若把导线做得更长，测到的电阻变化，不能直接宣布材料种类改变；若接触不良，电压降的一部分发生在电极处，也不能全归到材料内部。这些区分使“材料自身”的含义逐步成立，它们本身就是实验实践的产物。

可以设计一项课堂思想实验：两组人拿到同样成分的样品，一组只测两端电阻，另一组还改变长度、接触方法和温度。第一组可能把接触变化误判为材料差别，第二组则有机会将不同贡献拆开。但第二组也不因测得更多便自动正确；它仍须证明所用模型在这些条件下适用。知识的进步在于建立能够指出自己何时失效的差别组织。

因此，“这一类”既不由命名者任意决定，也不是无需条件的天赋身份证。它由可保留的特征及其生成条件共同规定。材料科学的创造力，恰在于使这些条件部分地成为可设计的变量，而不只是接受自然样品呈现的偶然结果。

51.5 晶体管：把失效的边缘变成研究中心

一块半导体材料不会仅因具有合适的能带，就自动成为可靠的放大器。接触、表面、电场和载流子的产生与输运需要共同满足器件工作的条件。巴丁、布拉顿和肖克利因半导体研究及晶体管效应的发现获得一九五六年诺贝尔物理学奖；贝尔实验室的相关历史资料将最初器件的突破放在一九四七年。[4][5]这里可核实的史实已经足以建立一个重要接口：材料理论与器件实现彼此促进，却不能相互替代。

可以把这段历史的发生学问题抽象为一个实验场景。研究者希望外加控制作用改变半导体中的电流，模型给出了方向，器件却没有出现预期响应。此时至少存在几种可能：理论机制不适用，样品参数不对，表面或接触引入额外过程，测量线路遮蔽了效应。重复同一次操作只会重复困惑；必须把这些可能变成不同的可检验预言。

此处所谓“边缘”，是指最初模型常被简化处理的部分，不是物理上天然不重要的部分。若接触处支配电流，表面就不是可以最后补上的装饰。若器件性能受杂质分布控制，平均成分就不能替代空间分布。实际失败迫使研究的分辨层级改变：从一块材料的总体标签，进入内部与界面过程怎样共同构成器件响应的问题。

这也说明技术并非把已经完成的科学结论照单执行。器件要求稳定工作，会把原来可以在论文理想化中暂时搁置的问题变成核心问题。科学研究因技术任务而获得新的差异序列，技术也因新的物理规定而改变可行路线。把二者画成单向传送带，会漏掉失败怎样反过来创造研究对象。

不过，不能把每一次技术改良都称为物理世界的根本重建。若只是同一机制中的参数优化，新意的层级应如实说明；当新的可控制对象、新的可区分类别或新的响应机制取得稳定位置时，才需要论证发生系统改变到了哪里。创造世界不是给所有成功贴同一张宏大标签，而是细分什么新的规定性真正成立。

51.6 三方程怎样进入材料研究

现在可以将本章的问题放回联合方法，而不在史实结束后附上一段口号。取 S 为一次研究中已经取得规定的显露，例如某温度与电路条件下的电流响应；D 为使响应可比较的差异序列，例如电压扫描、样品处理和复测次序；E 则是这些特征能够相互规定的纠缠组织，包括载流子状态、结构、接触以及测量量之间的具体制约。E 不能被写成“实验室、年代、社会”等背景词的任意清单。

S=F(D,E)提醒我们，一条曲线怎样显露依赖操作序列与组织条件。D=G(S,E)提醒我们，异常曲线和已有结构约束会共同迫使下一轮实验改道。E=H(S,D)则要求研究者面对结果与操作之间无法协调的张力，修改对哪些过程共同构成系统的理解。三个式子在这里是研究结构的表达，不能冒充取代输运方程的数值公式。

六路径的作用，是防止研究只沿一条惯性方向前进。S→D→E 的再创路径与 S→E→D 的进入路径，从显露进入但分别先检查操作或结构；D→S→E 的原型路径与 D→E→S 的倒逼路径，从实验差别进入但先追踪结果或联系改变；E→S→D 的锚定路径与 E→D→S 的沉淀路径，从纠缠组织进入但分别先问允许何种显露或如何实施操作。这里的顺序是研究进路，不是单向历史因果。曲线与接触模型相冲突，可以推动实验序列改变；新序列仍产生异常，就要求回到结构解释。

三原理在这里落实为两维张力推动第三维改变并回检：显露与实验序列不能协调时，检查纠缠组织；实验序列与既有组织不能协调时，检验能否产生新的显露；显露与结构要求不能协调时，重排实验序列。这三种工作不能各自封闭。每一次第三维的修改，都必须返回原问题，检查矛盾是否缓解以及有没有引入新失败。不能靠删去异常数据取得统一，也不能把互相冲突的说法并列便宣布研究完成。

这里尤其不能把客体属性直接等同于 S、互动属性等同于 D、主体属性等同于 E。研究者的分类目的、器件的作用方式、晶体的结构约束，都可能在同一个完整 SDE 内部发生；三号位与三维度具有不同的分析职能。将二者机械对应，会使本来用来揭示复合发生的理论退化为旧分类表。

51.7 材料世界的扩大及其边界

从晶体衍射到能带，再到半导体器件，物理学连续改变了物体取得身份的方式。结构不再只是外观的延伸，性质也不再只是成分的附件。可允许状态、占据、界面和制备历史进入判断，使“相同物质”与“相同材料行为”之间出现可以研究的距离。

这一变化没有取消微观解释，反而提高了微观解释的要求。若声称从原子推出材料，就必须说明用了哪些相互作用和近似，怎样处理尺度与边界，哪里使用经验参数，哪些结果是独立预测。宏观性质不是孤立原子的特征，并不等于宏观性质永远不能被微观模型解释；前者是本章明确维护的判断，后者需要另外的论证。

本书的进一步推论是，材料的发生不仅创造了更多物品，还创造了新的可区分性与可行动性。某种差别原来只能偶然出现，后来可以稳定制造、检测并接入其他系统，它就在物理世界中取得了新的位置。这个位置既由实验抵抗任意想象，也向新的制造与理论工作开放。

所以，回答“一块固体是什么”，不能只退回最小组成者，也不能只停留在人的用途。应当追问组成、结构、操作与响应怎样共同使它成为这一种材料，又在何种改变下失去或重获身份。下一步面对更复杂的集体现象时，这个问题将继续加深：当许多组成者共同运行，新的整体规定究竟怎样出现，又怎样经受独立检验？

本章参考资料

[1] W. Lawrence Bragg, The Diffraction of X-Rays by Crystals, Nobel Lecture, 1915；演讲入口：https://www.nobelprize.org/prizes/physics/1915/wl-bragg/lecture/；事实页：https://www.nobelprize.org/prizes/physics/1915/wl-bragg/facts/。本章仅采用演讲题名及方法主题的官方摘要信息。

[2] David Tong, Solid State Physics，作者剑桥大学课程原稿，2017，能带部分：https://www.damtp.cam.ac.uk/user/tong/aqm/solid2.pdf。本章紧束缚算例为自行展开的简化推演。

[3] David Tong, Applications of Quantum Mechanics，作者课程入口与章节目录：https://www.damtp.cam.ac.uk/user/tong/justaqm.html，参见电子动力学、金属与绝缘体。

[4] Nobel Prize in Physics 1956，官方授奖记录：https://www.nobelprize.org/prizes/physics/1956/summary/。

[5] Bell Labs, 1956 Nobel Prize in Physics，研究机构历史资料：https://www.nokia.com/bell-labs/about/awards/1956-nobel-prize-physics/。

第五十二章 超导与超流
整体性究竟创造了什么新物态

52.1 原子没有换，世界为什么换了

上一章讨论结构与能带如何参与材料身份的成立。本章再向前一步：当材料组成没有发生化学意义上的改换，某些宏观响应却在特定条件下发生根本改变，我们究竟应当说旧物具有了新数值，还是说一种新物态已经出现？超导与超流把这个问题从哲学争论带到了低温装置、磁场、流动通道和电学测量之中。

想象研究者逐次降低一个金属样品的温度。仪表显示的电压越来越小，最后落入分辨率以内。这时最容易出现的判断是“电阻为零”。但仪器没有分辨出信号，与信号在理论意义上严格为零，并不是同一句话。新物态不能只靠一个难以读出的数字获得身份；它必须在另外的响应中留下能够彼此约束的规定。

同样，一种液体流过细小孔隙异常容易，并不立即证明它在一切运动情形中都没有耗散。通道、温度、速度和激发方式会影响我们测到什么。若在一种配置中看见流动几乎不受阻，在另一种配置中又测到阻尼，研究者不能先把其中一项删掉，来保全一个漂亮的名称。

本章要守住的判断是：整体性并非用“无法理解的集体奇迹”代替微观机制。它要求重新确定什么变量足以描述物态、哪些相关关系可以跨越许多组成者保持、哪些扰动会破坏这种保持。整体新的规定性与组成者已有的物理性质并不互相排斥，问题在于怎样经过组织与运行，从后者产生前者。

52.2 零电阻为什么还不够

超导的电学响应很醒目，但磁响应使物态问题变得更加尖锐。迈斯纳效应提示我们，超导不能简单等同于一个电阻率恰好取零的普通导体。在适当的低场条件和样品区域内，磁场的排斥与屏蔽属于需要解释的特征；它与系统怎样经过转变进入该状态有关。[1]第二类超导体混合态中的磁通穿入等情形，也提醒我们不要把“排磁”写成所有条件下内部磁场处处严格为零。

可以用一个概念比较来说明。假定甲模型只规定电阻消失，却让磁场分布继续保留进入理想导电状态前的历史。乙模型还规定一种新的平衡磁响应。两者都可能与某次极低电阻测量相容，却对改变磁场和冷却路径后的结果给出不同预期。因此，新物态的识别不能停留在共享的那项成功上，应当走到它们预言不同的地方。

这不是要求每个实验同时测尽一切，而是要求知道单项证据能排除什么。一个短路故障可能造成电压读数过低；不均匀样品中的连通路径可能影响电流测量；磁信号还可能混有载台或其他成分的贡献。如果电学、磁学和热学证据能够在同一组条件下协调，论证就比单条曲线坚实。反过来，如果这些证据不协调，不协调本身就构成研究对象。

在 SDE 语言中，新增一个测试不是在同一个 S 旁边多放一张照片。测试改变了 D：样品遭遇的差异序列及读数形成方式被重新安排；不同结果迫使 E 中各特征的相互规定受到检验。新的物态身份因此不是某个孤立显露，而是能够承受不同进入方式的显露组织。这里的创造，表现为一套此前没有稳定成立的辨认与行动条件获得了现实位置。

52.3 配对不是把电子绑成小球

巴丁、库珀和施里弗于一九五七年发表的《超导理论》，以有效吸引作用和成对相关构造超导态，其计算涉及能隙、热容和磁响应；这项共同发展的理论获得了一九七二年诺贝尔物理学奖。[2][3]这一历史节点的重要性不只是出现了一套更精细的拟合公式，而是常态电子图景与集体量子状态之间有了具体的构造途径。

“库珀对”这个名称容易诱发一种过分直观的想象：两个电子像两粒弹珠系在一起，然后很多小对排队运动。这个图像会把最重要的状态相关误画成固定成员、固定距离的机械捆绑。本章把配对理解为多体状态中的相关组织；具体空间尺度和配对机制由模型与材料条件决定，不能仅靠“二合一”的图像推出。

这里应当区分三个判断。第一，某种相互作用使配对相关在给定条件下有利。第二，这些相关能够形成具有集体相干性质的状态。第三，这个状态在实际样品与测量条件下表现出相应响应。三者并非可以随意合并：找到一种局部有利关联，不等于已经证明宏观样品处在稳定超导态；在有限区间看到异常信号，也不等于完整机制已经查明。

为了显示推理的缺口，设想一种玩具系统，每个小区域内部都有强相关，但区域之间的相位安排不断漂移。局部探针可能稳定看到相关的痕迹，跨整个样品的响应却未必具有同样的相干组织。这个例子不是某种具体材料的实验报告，而是说明局部形成条件与整体接续条件承担不同的论证任务。二者相互关联，不应以其中一个替代另一个。

宏观量子性的“宏观”，也不是把微观波函数放大到肉眼可见的尺寸。它指某些量子状态关系能够进入跨越许多粒子的可测响应，因而不能只用彼此独立的经典小物件来描述。屏幕上的电流曲线仍是宏观读数；其成立却依赖一个经过制备、模型和多种检验连接起来的状态组织。视觉记录与触觉操作都在各自完整的 SDE 中参与了这种物态的成立。

52.4 用一个最小模型辨认“新的整体”

下面采用一个供思想分析的自由能模型，展示新整体变量怎样进入计算。设复数量 ψ 描述某种有序状态的幅度与相位，令 $r=|\psi|$；在均匀近似下取自由能密度差 $f(r)=\alpha r^2+(\beta/2)r^4$，其中 β 为正。这个形式属于朗道—金兹堡一类现象学思想的简化表达，而不是本章新发现的超导微观方程。[4]

对 r 求导，得到 $df/dr=2r(\alpha+\beta r^2)$。若 α 为正，$r=0$ 为最低点；若 α 为负，则 $r^2=-\alpha/\beta$ 处可以取得较低自由能，代回得到 $f=-\alpha^2/(2\beta)$。于是，当控制条件改变 α 的符号，最低自由能状态的结构也改变。变的并不只是一个既有响应的大小，而是描述稳定状态所需的非零有序幅度是否成立。

这个演算非常简单，却能防止两种误读。它没有说明在没有任何组织条件的情况下，“整体”凭空多出一种力量；α 和 β 把材料及近似条件带入模型。它也没有允许我们把每个组成者原有的性质直接相加，便得到非零的 r。新的变量代表一种集体状态组织，必须有与实验量的联系，才不是数学纸面上的装饰。

若只保留均匀幅度，就会漏掉不同位置之间怎样接续。因此可以加入正的梯度代价 $K|\nabla\psi|^2$。考虑幅度近似不变而相位缓慢改变的情形，$\psi=r\exp(i\phi)$，则这一项变为 $Kr^2|\nabla\phi|^2$。相邻区域相位任意错开不再免费，空间组织获得了可计算的代价。本章将它作为理解相位刚性的一个入口，不把该简式当成带电超导体的完整表达；电磁情形还需保持规范不变性。

模型也有清楚的边界。它不能仅凭一次极小化就给出所有涨落效应，不能无条件穿越临界区，更不能决定每一种材料中配对吸引从何而来。若把模型系数全部由已知曲线反推，再把同一曲线当成独立预测，会夸大解释能力。可靠的工作要指出哪些量用于定参，哪些量保留作检验，以及改变条件后这种区分是否仍成立。

这里，微积分参与的是整体变量如何获得空间结构、稳定性怎样从局部变化条件中推出的问题。$df/dr=0$ 本身不保证找到了稳定物态，还要检查极值类型、允许变量域和模型是否适用；梯度项也不能自动解决边界条件。数学推演与对象成立各自有工作，二者必须往返衔接。

52.5 超流让同一种液体呈现不同的运动身份

一九三八年的《自然》分别刊登了卡皮察关于低于 λ 点液氦黏度的论文，以及艾伦和米塞纳关于液氦 II 流动的论文。[5][6]这些研究不是简单地把液体的黏度表往下添一

行。异常流动迫使人们重新辨认“这份液体在怎样的运动中表现出什么”，并检查不同装置是否确实询问了同一个问题。

同年关于液氦 II 黏度的另一篇实验报告明确指出，小孔流动与振荡圆柱方法所得结果之间存在不一致。[7]我们不必把全部历史细节压成一个一夜完成的发现故事；仅这一点就足以说明，出现矛盾的测量并不总意味着某一方粗心，也可能意味着旧的单一变量描述尚未覆盖不同的运动方式。

用现代两流体语言，人们区分正常组分与超流组分及其不同响应，但这不表示把液体按空间切成两层互不相干的物质，也不意味着能给每个原子永久贴上其中一种标签。[8]这里的“组分”属于集体运动的描述。它的价值在于使热输运、流动和边界作用能够在相容的结构中讨论，而不在于提供一幅简单的双色小球图。

这一变化与上一章的材料分类形成更深的联系。单个氦原子不能具有一整杯液体的黏性系数，而一整杯液体也不是只需报告一个与测量方式无关的词语。变量的成立有尺度、有过程、有可辨条件。面对异常，研究的任务是寻找能协调这些条件的状态描述，并让新描述承受独立测量的限制。

不能由此滑向“实验怎样做，物态就随便怎样变”的结论。测量配置既可能改变系统，也可能只对同一系统的不同响应敏感；两种情形需要分开。若一种新解释只在某个仪器上成立，换一个有效测量方式就失败，它的对象性仍然不足。真正的新物态需要从单台装置中的惊奇走向可以跨装置接续的规定性。

52.6 为什么氦三不是把氦四的故事照抄一次

超流研究还把“相同外观液体”内部更深的区别带入物理世界。氦三超流的发现获得一九九六年诺贝尔物理学奖，官方材料将李、奥谢罗夫和理查森列为获奖者；专业介绍指出，相关转变最初以低温实验中的细小异常进入研究。[9]本章不把发现、解释与发表压缩成同一天，也不将官方概述替代全部实验记录。

氦三作为费米子系统，其超流与成对相关的集体状态相联系；氦四则提供不同的量子统计入口。[10]这一差别的重要性在于，两个系统可以共享“超流”这一宏观类别，同时具有不同的微观实现。类别一致不等于机制完全一致；机制不同也不阻止某些有效变量在较高层级共同发挥作用。

可以把这种关系画成两个方向的提问。向下问：此处的组成者、相互作用和统计规则怎样支持这个状态？向上问：这个状态具有哪些跨具体实现仍能保持的响应关系？若

只向下，容易把宏观相似看作偶然；若只向上，又可能把一切超流都讲成同一个故事。物理思想的创造在于能够转换层级，并明确转换保留什么、舍弃什么。

由此还应区分超流与凝聚等概念，避免把所有宏观量子现象都替换成“大家进入同一状态”。对于具体系统，哪一种凝聚或相关量与超流响应相联系，需要借助相互作用、维度及边界条件分析。一个形象比喻能够帮助入门，不能充当跨系统的定理。本章所维护的整体性，是经条件约束的状态组织，而不是抹去各类物态差别的总名称。

52.7 涌现为何不等于原则上无法推导

超导与超流常被用来支持“整体大于部分之和”。这句话若指孤立组成者的性质列表不足以给出集体状态，具有明确意义；若进一步声称任何足够完整的微观数学都绝不可能解释整体，则论证已经换了问题。一个物态需要新有效变量，与它在原则上无法从微观理论导出，并不相同。

BCS 理论的存在本身就要求本书认真对待微观解释。它通过多体构造解释一类超导现象，而不是因为拥护整体性就停止计算。另一方面，具有微观方程也不意味着一切宏观问题已经自动解决：选择近似、发现合适变量、辨认稳定解、处理边界与缺陷、建立实验联系，都仍需创造性工作。不能把“方程写下了”当成“物态已经理解了”。

一个有用的区分是计算可行性与解释充分性。某个有限模型原则上可数值求解，实际计算却可能昂贵；这是能力与方法的问题。另一个模型算得很快，却选错了相关自由度，可能对真正的转变无能为力；这是对象组织的问题。再有一种情况，计算准确拟合实验，但人们不知道哪组参数可辨识，解释就还没有完成。把这三种困难统称不可还原，会使研究失去方向。

本书的原创推论是：整体物态的发生，使一些原来不具备独立操作意义的规定成为可以制备、检测、改变和接续的对象。这个“新”位于物理世界的发生层级，不需要借“微观理论永远失败”来保证。即便某种宏观关系能够成功推导，形成其变量、制备方法与检验体系的创造仍然真实存在；推导的成功恰能说明这一新世界如何与其他物理规定接通。

52.8 把联合方法放回低温实验台

在本章的实验问题中，S 可取为某组电磁、热学或流动响应；D 包括降温、升温、场扫描、改变通道和重复读数的差异序列；E 则是这些特征在特定样品、相干状态及测量作用中的相互规定。这里没有把仪器简单堆到“背景”里。哪些仪器特征确实改变结果的成立，必须通过对照说明。

考虑一个尚未判明的实验：电阻信号下降，但磁响应没有出现模型预计的变化。沿 S→E→D 的进入路径，先不宣布新物态成立，而是问电流通路与磁测量覆盖的体积是否一致，再改变电极布置、比较样品不同部位、重做场扫描。这里 S=F(D,E)成为对同一次异常的拆解义务，D=G(S,E)使下一次测量从异常与结构约束中产生，E=H(S,D)则等待新结果决定原有整块均匀样品假定能否保留。另一条 E→D→S 的沉淀路径，可以从候选非均匀结构出发，设计它区别于均匀超导模型的测量，再看显露是否符合，而非在结果出现以后任意补解释。

三原理进一步规定改动的理由。若两种降温序列产生相冲突的显露，S 与 D 的张力推动 E 重组，检查是否存在磁通历史、非均匀性或状态变化。若已有结构理解与可实施操作相冲突，D 与 E 的张力推动寻找新的 S，检验此前未被分辨的响应。若显露与结构模型不能协调，S 与 E 的张力推动 D 改变，例如加入另一种磁场或流动测量。每次改变第三维以后，都必须返回原有矛盾检查。

还可以在这项思想实验中引入一个保留样品。研究者先用第一块试样选择参数，再把冷却与磁场次序固定，用第二块试样作检验。如果第二块失败，应分别检查制备差异和模型失配，而不是把两块样品平均成成功。若改变测量电流就使低电阻现象消失，则需要进一步分辨临界响应、局部发热与接触变化。只有这些竞争说明给出不同预期，追加实验才真正增加知识；单纯重复最漂亮的曲线不会自动加强结论。

这套方法并不保证结果必然是发现新物态。有时查明的是传感器漂移，有时是样品连通性，有时才是旧状态分类不足。它的价值恰在于不预订答案，却坚持寻找哪一层规定需要重建。低温世界的深刻之处，不只在温度数字极小，而在看似同一物质能经可复验的发生过程，取得原先不存在的整体身份。

本章参考资料

[1] 剑桥大学课程讲义 D21L，第 5.2 节，超导磁响应与规范不变电流：https://www.damtp.cam.ac.uk/user/examples/D21L.pdf。阅读范围为第 5.2 节相关正文节选。

[2] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Theory of Superconductivity, Physical Review 108, 1175 (1957), https://doi.org/10.1103/PhysRev.108.1175。已读出版页摘要及参考文献。

[3] Nobel Prize in Physics 1972，官方授奖说明：https://www.nobelprize.org/prizes/physics/1972/press-release/。阅读范围为官方授奖说明的摘要节选。

[4] David Tong, Statistical Field Theory，§1.3，作者课程原稿：https://www.damtp.cam.ac.uk/user/tong/sft/sfthtml/S1.html。本章自由能极值及梯度代价为自行展开的演算。

[5] P. Kapitza, Viscosity of Liquid Helium below the λ-Point, Nature 141, 74 (1938), https://doi.org/10.1038/141074a0。核验范围为出版记录与摘要。

[6] J. F. Allen, A. D. Misener, Flow of Liquid Helium II, Nature 141, 75 (1938), https://doi.org/10.1038/141075a0。同上。

[7] Viscosity of Liquid Helium II, Nature 142, 72 (1938), https://doi.org/10.1038/142072a0。阅读摘要所列装置结果差异。

[8] David M. Lee, Nobel Lecture，官方存档：https://www.nobelprize.org/uploads/2018/06/lee-lecture-2.pdf。阅读范围为两流体描述节选。

[9] Nobel Prize in Physics 1996, Advanced information：https://www.nobelprize.org/prizes/physics/1996/advanced-information/；获奖页：https://www.nobelprize.org/prizes/physics/1996/summary/。阅读范围为专业说明摘要及获奖记录。

[10] D. Vollhardt, Pair Correlations in Superfluid Helium 3 (1997), https://arxiv.org/abs/cond-mat/9706100。读取作者摘要，未冒称全文审阅。

第五十三章 混沌、湍流与临界
不收敛一定是新本体吗

53.1 先问什么没有收敛

一条实验曲线反复摆动，一段数值计算越算越不同，一个理论面对新资料不断改动，都可以被日常语言称为“不收敛”。但是，它们可能指三件完全不同的事：系统轨道不趋于定值，数值近似没有获得稳定精度，或者现有解释不能协调证据。若不先分开这些含义，任何不顺利都可以被宣布为旧世界崩溃，任何复杂图像都可以被包装成新本体出现。

本章对这种捷径提出限制，同时维护复杂现象真正带来的思想创造。混沌告诉我们，确定性规则与长期轨道可预测性不能等同；湍流使尺度之间的传递和统计描述取得核心地位；临界现象则把相关范围、涨落与尺度变换带进物态身份。这些进展都产生新的有效对象和新的检验方式，但它们不是同一种现象，更不共同构成“只要混乱就超出数学”的证明。

上一章讨论了相干的集体状态，本章面对的是另一种挑战：当局部变化持续发生、单条轨道难以长期追踪，物理学还能建立什么样的稳定知识？如果稳定只被理解为每个数最终停在同一位置，答案似乎只能是否定的。但物理学能够稳定的，还包括分布、相关函数、尺度关系以及对扰动的响应。移动中的秩序需要相应的提问方式。

53.2 洛伦兹改变的是确定性的含义边界

洛伦兹一九六三年发表《确定性的非周期流》，研究有限维确定性非线性微分方程所产生的非周期行为及其对微小改变的敏感性。[1]本章采用这个可定位的原始成果作为入口，不以广为流传的偶然发现故事代替论文论证。历史意义在于，简单的确定性模型已经足以迫使人们区分“规则明确”与“远期状态能够精确预告”。

常用的洛伦兹三变量系统可写成 $dx/dt=\sigma(y-x)$，$dy/dt=x(\rho-z)-y$，$dz/dt=xy-\beta z$。这里 x、y、z 不是三维空间里一只蝴蝶的位置，而是简化模型中的状态变量；σ、ρ、β 是参数。只有在相应参数与初值条件下，系统才可能表现所讨论的复杂行为，不能看到非线性项 xy 就宣布所有解都混沌。

这一系统仍然给出确定的局部演化规则。若两个初态非常接近，后续距离在某个范围内迅速增长，测量精度便成为预测时限的关键。设小偏差在适用区间近似满足 $\delta(t)=\delta_0 \exp(\lambda t)$，$\lambda$ 为正，允许误差阈值为 Δ，则达到阈值的时间约为 $t^*=\ln(\Delta/\delta_0)/\lambda$。这个演算不是对所有系统的普遍定律，而是把指数敏感性对预测的影响明确写出。

假定初始精度提高一千倍，预测时限增加的是 $\ln(1000)/\lambda$，而不是自动提高一千倍。这一点足以纠正“只要仪器再精密一点，远期轨道总能无限延长”的直觉。但它没有否定短期预测，没有取消局部因果，也没有证明任何细节都不可知。误差增长模型本身就要求选择变量、范数和适用时间，不能把这一结论放大为普遍的不可认识性。

同样，长期轨道不同不直接意味着数值算法错误。研究者需要比较有限时间误差、守恒或耗散结构以及统计特征；若一味要求两个近似计算永远逐点重合，反而可能提出不合适的验收条件。这不是降低准确性要求，而是把准确性放在该系统允许的层级上，同时保留发现数值伪影的办法。

53.3 一条递推式如何产生分岔

罗伯特·梅一九七六年的论文强调，简单数学模型可以产生复杂动力学，其讨论使这类问题更广泛地进入研究与教学。[2]本章采用常见的一维映射自行演算：$x_{n+1}=r x_n(1-x_n)$，令 x 处于零与一之间，参数 r 在零到四的范围内。这不是给真实人口、流体或社会系统直接套上的普遍方程，而是一块能够逐项检查推论的数学试验台。

先寻找不动点。令 $x=rx(1-x)$，得到 $x=0$，以及 r 不为零时的 $x^*=1-1/r$。令一个小偏差为 η，则下一步偏差的一阶近似为 $\eta_{n+1}=f'(x^*)\eta_n$，而 $f'(x)=r(1-2x)$。在非零不动点处，导数变为 $2-r$，所以当 $1<r<3$ 时，其绝对值小于一，不动点局部稳定；到 $r=3$，线性稳定条件到达边界。

例如 $r=2.5$ 时，非零不动点为 0.6，局部误差乘子为 -0.5，意味着误差在交替方向的同时缩小。$r=3.2$ 时，不动点为 0.6875，误差乘子为 -1.2，微小偏离在这一近似内反而增长。读者可以直接复算两组数值，而不必凭分岔图的视觉密度判断稳定性。

这个小计算给“不收敛”一个十分具体的含义：原先稳定的不动点失去稳定性，不等于映射不再定义，也不等于所有统计量发散。它可能进入周期二等不同的吸引行为。对周期二的判别应检查两步映射，而不是继续拿一步不动点的标准强行裁决。这里新对象的出现首先是数学状态组织发生变化，不需要人为增加一个未被定义的外部力量。

取 $r=2$ 且 $x_0=0.2$，前几步为 0.32、0.4352、约 0.4916，趋近 0.5 的趋势可以直接看出。改变 r 以后，变化并非只是趋近速度略有不同；稳定结构可能改变。为此，参数扫

描必须固定初值与去除暂态的规则，不能把不同计算长度混在同一张图里，再把散点的宽度一概叫作混沌。

还可以构造一个更直接的反例：$x_{n+1}=1-x_n$。除初值为二分之一外，序列在两个值之间往复，永远不趋于单个数，却完全容易预测。可见“不趋于定值”连混沌都不能单独证明，更不用说新本体。若序列不断增大，也可能只是存在一个不稳定线性模式；发散与复杂性的身份仍然不同。

53.4 图像很乱，证据可能很弱

真实时间序列比递推试验台困难，因为研究者不知道全部变量，也不能无限精确地重置状态。一个传感器的随机噪声能够画出不规则曲线；两个频率不相称的周期信号叠加，也可能在短窗口内显得复杂；缓慢漂移的参数又会让相同机制在不同时间段呈现不同表现。视觉上的“乱”不能把这些机制区分开。

因此，从观测宣布确定性混沌，需要比画相图更多的证据。采样频率太低会把快速结构折叠进慢变化，数据过短会夸大拟合结果，强平滑又会制造看似低维的曲线。研究应当对照不同采样与滤波选择，检验结论是否依赖某个处理按钮。替代数据或噪声模型比较可以提供约束，但某项检验通过，也不自动等于机制唯一确定。

这里可以设计一个诊断练习。甲组提交一段受周期驱动的确定性信号，乙组提交带噪的线性随机过程，丙组提交非线性映射输出，三组都经过相同的测量分辨率截断。评审先隐藏来源，只比较时间曲线，很可能出现误判；若加入响应于可控扰动的结果、跨窗口稳定性和模型外预测，候选解释才逐步分开。实验干预由此比图像相似更有辨别力。

参数未辨识又是另一种常见困难。两个不同参数组合可能在已经观测的区间给出近乎相同的曲线，到了外推区间却完全分开。此时继续增加同一条件下的数据未必有效；需要改变实验条件，让原来互相抵消的参数作用分离。模型频繁换参，不必然意味着系统自发创造新规则，也可能说明我们一直没有真正识别旧规则。

在 SDE 分析中，第一步不是给“不收敛”加上哲学荣誉，而是确定差别在哪一层产生。若噪声来自测量，改动 D 可能足够；若变量漏掉了实际耦合，E 需要重构；若旧读数把两种行为混在一起，S 的分类也需要更新。三方程不能替代这项排查，反而把排查变成不可省略的研究义务。

53.5 湍流不是一幅放大的洛伦兹图

低维混沌提供了理解确定性复杂行为的有力入口，但真实湍流包含空间自由度、边界以及多尺度相互作用，不能被一张三变量吸引子图替代。剑桥高雷诺数流动研究组的工作说明，实验、数值与渐近分析在不同尺度的流动研究中联合使用。[3]它们面对的不是“没有规律”的同义词，而是怎样组织规律并获得可检验近似的问题。

以不可压、常密度的牛顿流体为例，运动方程包含局部加速度、对流项、压力梯度与黏性扩散。取特征速度 U 和长度 L，令运动黏度为 ν，尺度比较产生雷诺数 $Re=UL/\nu$。它把惯性与黏性项在所选尺度下的相对大小写成无量纲量。这个比值并不独自决定任意流动何时转捩，几何、扰动和边界条件同样重要。

为何尺度选择是实质问题？假定同一流场里既有宽大的循环，也有细小的速度变化层，用一个 L 描述所有局部过程会损失信息。对流可以把不同尺度的运动联系起来，黏性又在相应梯度上产生作用。模型要说明哪些尺度被直接分辨，哪些被平均，其效果以什么方式回到保留变量中。把未分辨部分统统叫“背景”，无法形成有效闭合。

对速度作平均分解 $u=\bar{u}+u'$，就能看到困难来自哪里。平均后的非线性乘积一般不等于平均量的乘积；例如平均 $u_i u_j$ 会包含平均 $u'_i u'_j$。这些涨落相关进入平均运动方程，使单纯平均不能自动封闭问题。新变量并非为了增加复杂词汇而加入，而是原运算确实产生了需要处理的项。

本章不把这条代数事实夸大为湍流全部理论。它只说明，从瞬时描述转到平均描述，有明确的得失：逐点轨道信息减少，但统计相关成为新的对象；若要预测平均阻力，就必须处理这些相关，不能假定它们因平均而消失。湍流使“稳定知识”从一条确定轨迹扩展到一定条件下的统计结构。

53.6 统计稳定不要求每个涡都回到原位

两次流动实验即使使用相近初态，瞬时涡结构也可能迅速分开。但在合适的定常驱动条件下，某些时间平均或分布仍可能具有可重复性。这里的“可能”由实际系统和观测时间决定，不能把有限样本的平稳外观当成严格遍历性的证明。研究者需要检验平均随窗口延长是否稳定，以及不同初态是否进入可比较的统计状态。

例如，若关注输送相同流量所需的平均压差，逐点预测每个小涡并非唯一有意义的成功标准。模型若能够在独立条件下预测平均压差及其波动范围，就取得了明确成绩；若只能追踪很短时间的漂亮涡图，却预测不了整体阻力，也不能凭图像复杂宣称解释更深。问题决定哪些稳定量值得建立。

与此相对，非定常过程不能被强行装入定常平均。衰减湍流的能量随时间变化，把很长一段数据混为一个分布，可能遮蔽过程。关于非定常湍流级联的原始数值研究也指出，某些谱形特征的出现不等于通常的定常级联平衡已经成立。[4]这为本章提供一个具体警告：相似的幂律图形，不能独自证明相同的发生机制。

因此，统计稳定也需要发生条件。输入、边界、耗散和观测窗口怎样共同支持某个分布，是应当明说的部分。分布不是撤退到较粗糙知识的借口；当它承载独立可验的约束时，它就是新的有效知识形态。但当它只是把不理解的差别平均掉时，所谓稳定也可能是分析过程制造的幻觉。

53.7 临界点为什么要求改变尺度

临界现象带来的困难与混沌不同。在某些连续相变附近，涨落与相关范围改变，使只考虑局部平均的简单近似不足。威尔逊因相变相关的临界现象理论获得一九八二年诺贝尔物理学奖，重整化群构成其工作的核心方向。[5]这里的创造并非把“不稳定”改名，而是通过尺度变换讨论哪些结构在改变观察尺度后仍然重要。

可以从相关函数入手。设局部变量为 q(x)，扣除平均后的相关可写为 C(r)=⟨q(x)q(x+r)⟩−⟨q⟩²。在适当情形中，相关随距离衰减的尺度用 ξ 表征。若 ξ 远小于样品尺寸 L，不同区域在某些问题上可近似独立；若 ξ 增长到与 L 可比，边界和整体尺寸就不能继续忽略。这不是说每个自由度都变成一样，而是局部独立近似的有效范围改变了。[6]

设某个理论在适用区间预言 ξ≈ξ₀|τ|^(−ν)，其中 τ=(T−T 临)/T 临，其中 T 临为相变温度，ν 为正。这个形式并不允许有限样品真的容纳无限相关长度。至少要检查 ξ 与 L 的比例，并考察有限尺寸怎样改变所测峰形及位置。将有限范围内拟合出来的陡增直接称为无限发散，会把理论极限与实际样品混为一谈。

上述关系还给出一个实验设计提示。不是只把温度测得更细，而应比较不同尺寸样品，在相应边界条件下观察特征怎样移动或钝化。如果只测一块样品，一个窄峰可能来自转变，也可能受到温度不均匀或测量分辨率限制。尺寸变化为解释提供新的限制维度，使候选机制不再只在一条曲线上竞争。

重整化的粗粒化也不是宣布小尺度从此不真实，而是系统地追问：改变尺度后，哪些参数影响减弱，哪些仍支配长距离行为，哪些新项必须保留。[7]不同微观系统在特定条件下可以共享某些临界规律，但不能由此说所有材料没有差别。普适性是有分类条件的共享，不是删除微观世界的总命令。

53.8 三种不稳定，三种研究改道

现在可将本章的三个问题真正分开。混沌中的敏感性主要挑战长期轨道预测；湍流中的尺度耦合要求处理大量自由度与统计闭合；临界附近的长相关则迫使人们重新组织尺度和涨落。三者可以在具体问题中交汇，却不能因都具有复杂图像，就被同一个“混沌态”名词取代。本书 SDE 理论使用的混沌或介生态概念，若要与动力系统混沌联系，必须另建定义和证据桥梁。

在一项临界实验中，可以把观测峰形作为 S，温度与尺寸扫描作为 D，相关结构及边界制约作为 E。若峰形随扫描速度改变，S 与 D 的张力首先推动检查 E 中是否漏掉非平衡过程；若既有相关结构要求更大的观测范围，而设备操作无法达到，D 与 E 的张力推动寻找可替代的 S，例如其他可测相关指标；若峰形与模型预测不符，S 与 E 的张力推动重排 D，以不同尺寸或路径区分原因。每次改动都需回到原矛盾复验。

这时，S=F(D,E)不再是一句抽象依赖声明：它要求说明扫描速度如何影响峰形。D=G(S,E)要求根据峰形与相关模型选择下一组扫描，而不是无限加密相同区间。E=H(S,D)要求根据不同操作所得显露，决定是否保留平衡与均匀假定。沿 D→E→S 的倒逼路径，主动改变尺度逼出结构差别；沿 S→D→E 的再创路径，从异常峰出发重排操作，再检验是否需要新的状态组织。

这个闭环允许两种同样有价值的结果。一种是纠正了假峰，保留原有理论；另一种是旧变量确实不能协调新增证据，需要新的有效对象。只有第二种情况才进入本体重建的进一步论证，而它仍须说明新规定怎样成立、如何超过旧框架、能在哪里失败。方法的严谨不是压低创造力，而是让创造力不被自己的惊奇轻易满足。

53.9 什么才配得上“新”的判断

不收敛本身不是勋章。一次计算爆掉、一个拟合换参、一幅曲线复杂，都只能成为问题入口。若研究者能够排除主要噪声解释，澄清尺度与参数可辨识性，建立跨操作保持的关系，并使新变量产生独立可验的预期，那么“不收敛”才可能成为发现新规定性的起点。

传统复杂系统研究已经拥有非线性动力学、统计物理、分岔、闭合和尺度分析等强大工具。本书不应把这些成果先说成无效，再由 SDE 重新赐予意义。SDE 若要推进研究，必须在具体工作中显示增量：是否更早识别错误对象边界，是否把遗漏的操作差别纳入，是否促成新的可验证模型。若做不到，宏大的本体论语言也只是另一层描述。

本章因而为“物理世界的创造”增加一项约束。新世界的成立并不以旧计算失败为充分条件，而以新的显露、差异序列与特征纠缠能够形成有效接续为正面证据。从轨道到分布，从局部到尺度，从单一读数到相关结构，变化必须落实为可研究的新规定。真正的思想创造既敢于改变什么算作物理对象，也能说明为什么这一次改变值得被保留。

本章参考资料

[1] E. N. Lorenz, Deterministic Nonperiodic Flow, Journal of the Atmospheric Sciences 20, 130–141 (1963), https://doi.org/10.1175/1520-0469(1963)020%3C0130:DNF%3E2.0.CO;2；原论文扫描存档：https://cdanfort.w3.uvm.edu/research/lorenz-1963.pdf。出版摘要及 PDF 入口已核验；本章不声称完成全文逐式重算。

[2] R. M. May, Simple mathematical models with very complicated dynamics, Nature 261, 459–467 (1976), https://doi.org/10.1038/261459a0；原文机构转载：https://ned.ipac.caltech.edu/level5/Sept01/May/May_contents.html。本章映射稳定性和数值为自行演算。

[3] University of Cambridge, High-Reynolds-Number Fluid Flow，研究组官方说明：https://www.damtp.cam.ac.uk/research/fluidshr/index。

[4] Unsteady turbulence cascades, Physical Review E 94, 053108 (2016), https://doi.org/10.1103/PhysRevE.94.053108。阅读原论文摘要，限用于非定常级联与谱形关系的提醒。

[5] Kenneth G. Wilson, The Renormalization Group and Critical Phenomena, Nobel Lecture, https://www.nobelprize.org/prizes/physics/1982/wilson/lecture/；授奖记录：https://www.nobelprize.org/prizes/physics/1982/summary/。仅据官方演讲题名、主题及授奖记录作历史定位。

[6] David Tong, Statistical Field Theory, Chapter 2，作者原稿：https://www.damtp.cam.ac.uk/user/tong/sft/sfthtml/S2.html。

[7] David Tong, Statistical Field Theory, Chapter 3，作者原稿：https://www.damtp.cam.ac.uk/user/tong/sft/sfthtml/S3.html。

第五十四章 宇宙学
整体宇宙怎样成为可研究的对象

54.1 一个不能放在面前的对象

研究一块晶体，可以移动样品，可以更换电极，也可以制造另一块晶体作比较。研究宇宙似乎没有这些便利：观察者在宇宙之内，不能走到外部拍摄全景，也不能重新启动一次宇宙演化。倘若可研究性必须等于把对象完整置于面前，宇宙学就只能是一种宏大想象。然而实际科学已经建立起另一种可研究性：把有限观测窗口接到明确的动力学和传播模型上，使关于整体的假设产生能够在窗口内相互核查的差别。

这里的“整体”首先不是一张包罗万象的照片，而是一组共同约束。某个模型给出随时间变化的尺度，便不能另给一套与它矛盾的红移关系；它说明物质平均密度怎样演化，也必须说明光怎样走到观察者这里。我们看见的星系、微波天空和元素丰度各有自己的形成过程，却不能被随意拼成互不承担责任的故事。整体宇宙作为物理对象，首先成立于这些不同历史痕迹能够接受同一组关系检验。

这不意味着模型涵盖的全部区域都已被观察。一个覆盖整个时空的数学解，可能包含永远不能向我们传来消息的事件。其计算上的完整与经验上的可达不是同一种完整。宇宙学最需要防止的跳跃，正是把“模型能写到那里”直接说成“我们已经测到那里”。但反过来，尚未观测的部分也不能一概视为无意义：只要关于它的假设会限制可观测区域中的统计关系，它仍可能间接承担检验责任。

54.2 从一条谱线到一部膨胀史

红移的直接定义很简单：同一原子跃迁的观测波长与实验室波长相比，有 $z=\lambda_{\text{观}}/\lambda_{\text{发}}-1$。困难在于这项比值不自己说明原因。发射体的局部运动、引力势差、宇宙尺度的演化，都可能进入频率变化。将一批星系红移解释成宇宙膨胀，需要同时建立距离尺度、校正局部运动，并规定光传播的几何。

哈勃一九二九年的论文把星系距离估计与径向速度资料相比较，使距离和退行趋势的关系成为可检验的问题。[1]这项历史贡献不应被改写为现代宇宙模型已经在一条散点拟合中全部完成。距离零点、样本数量、局部运动以及理论解释仍须分别处理。红移与距离进入共同图表，是新研究对象成立的重要一步，却不是它最后的形式。

在均匀膨胀模型中，令当前尺度因子 $a_0=1$，则 $1+z=1/a$ 发。相邻两次光波波峰沿同一传播关系到达时，它们的时间间隔也按这个比例伸长。因此，红移不只是“颜色变

红”的命名，它把频率差别与传播历史连在一起。若某种解释只让每个光子失去能量，却不给出相应的到达时间伸长，它与膨胀模型便不再是同一个经验命题。检验的对象由一个颜色读数扩展为频谱、光变时标和距离关系的联合。

在足够低的红移范围，可以把光度距离写成 $d_L \approx cz/H_0$。作一个纯示意计算：取 $H_0=70$ 千米每秒每兆秒差距，红移 $z=0.01$，则 d_L 约为 42.9 兆秒差距。若星系具有沿视线方向三百千米每秒的本动速度，它在这个红移上的影响约占一成。把同样本动速度放在 $z=0.001$ 的目标上，误差就与膨胀速度同量级。这说明“更近”未必总能使某个宇宙学参数更容易测准；距离太近会使局部动力学遮蔽所要抽取的整体规律。

54.3 均匀与各向同性不是肉眼印象

通常所谓 FLRW 模型，把大尺度空间的均匀性和各向同性作为几何条件。均匀性是各处没有特权位置，各向同性是各方向没有特权方向。这两句话既不要求每个小体积包含同样数量的恒星，也不否认星系团、空洞和丝状结构。它们要求在指定的粗粒化尺度上，统计结构可以由共同的背景几何描述，再把局部偏离作为扰动处理。

从一个观察位置看各方向相似，并不足以在逻辑上证明所有位置相同。一个以观察者为中心的球对称不均匀模型，也可以在此处表现为各向同性。由观测天空进入均匀背景，包含着非特权位置的假设和其他证据的支持。把这个假设公开并不削弱科学，反而使问题变得明确：哪些观测能够检验空间不均匀性，哪些只约束我们看到的方向差别？

采用无量纲尺度因子时，可将空间线元写作 $ds^2=-c^2dt^2+a^2(t)[d\chi^2+S_\kappa^2(\chi)d\Omega^2]$。这里 χ 是共动径向长度，S_κ 按曲率选择相应函数；平坦空间中它就是 χ。这个表达式的作用不在符号庄严，而在于把时间变化与空间形状分开，使一个光子的传播、两个共动点的间距和同一时刻的平均密度可以置于共同计算中。

要特别注意，“共动”也是一种指定。它随理想化宇宙物质的平均运动建立坐标，不宣称所有真实星系严格固定于网格。引力束缚系统中的恒星、行星和实验室尺子，不能简单套用 $a(t)$ 成比例增长。宇宙膨胀是背景时空与物质分布的关系，不是一种均匀拉伸每个局部物体的橡皮膜力量。只有把适用尺度写清，整体模型才不会侵吞已经成立的局部物理。

54.4 膨胀率不是宇宙年龄的别名

在广义相对论和均匀理想流体条件下，尺度因子的演化受 Friedmann 方程约束。用 ρ 表示质量当量密度，κ 表示具有倒长度平方量纲的曲率参数，可写 $H^2=8\pi G\rho/3+\Lambda c^2/3-\kappa c^2/a^2$，其中 $H=\dot{a}/a$。压力同时进入加速度关系：$\ddot{a}/a=-4\pi G(\rho+3p/c^2)/3+\Lambda c^2/3$。这里 ρ、p 不含已单列的宇宙学常数项，避免把它计算两次。

这些式子不是由红移定义直接推出的。它们加入了引力理论、时空对称性和物质内容。红移资料因而能够反过来检验这些条件的组合，但不能只凭一个拟合曲线判断是哪一项一定正确。相同的背景膨胀史，可能来自不同组分，甚至来自不同的动力学解释。

一个可完整算清的例子，是平坦、无宇宙学常数、由无压物质主导的模型。物质密度随体积稀释，$\rho\propto a^{-3}$，于是 $H\propto a^{(-3/2)}$。由 $\dot{a}=aH$ 积分可得 $a\propto t^{(2/3)}$，故 $H=2/(3t)$，模型年龄 $t_0=2/(3H_0)$。如果仍取前面的示意 H_0 值，其倒数约为一百四十亿年，而这个模型的年龄约为九十三亿年。相同的当前膨胀率，并没有给出相同数值的宇宙年龄。

再换成平坦的辐射主导模型，辐射能量除了随体积稀释，还因红移减少，$\rho\propto a^{-4}$，积分得到 $a\propto t^{(1/2)}$ 和 $t_0=1/(2H_0)$。两个结果的差别不是计算约定，而是不同压力和密度演化留下的历史差别。所谓“宇宙多少岁”实际上要求恢复一段膨胀历程，而不能把今天的速度尺度机械地倒置。这也解释了为什么恒星年龄等独立时钟能够成为宇宙模型的交叉约束。

54.5 距离必须说明怎样测量

在日常空间里，角大小、光强和尺长通常给出相容的同一种距离。宇宙学中，光从发射到接收经历了膨胀，这些测量规定不再重合。角直径距离 d_A 由物体横向实际尺度与角大小之比定义；光度距离 d_L 则通过流量 $F=L/(4\pi d_L^2)$ 定义。名称都含“距离”，不等于它们已经是同一物理量。

以下用平坦模型作独立推演。一个在红移 z 处发光的共动源，与我们之间的共动径向距离为 $\chi=c\int_0^z dz'/H(z')$。发光时的横向长度对应 a 发 χ，因此 $d_A=\chi/(1+z)$。另一方面，光子各自的能量降低一个 1+z 因子，单位接收时间内到达的光子数再降低一个同样因子；当前球面面积为 $4\pi\chi^2$。所以在透明传播、光子数守恒等条件下，$d_L=(1+z)\chi$，二者满足 $d_L=(1+z)^2 d_A$。[2]

这里两个红移因子承担不同作用，不能只凭符号合并忽略其物理来源。若红移 z=1，同一模型中 d_L 是 d_A 的四倍。把角直径距离误当光度距离，代入平方反比律，预

测流量会差十六倍。这个错误不能靠给某颗超新星微调亮度消除，因为它会随红移系统变化。关于宇宙几何的争论因此必须同时讨论测量规定，而不只是比较谁给出的距离更大。

对前节的物质主导模型，H(z)=H₀(1+z)^(3/2)，可以直接积分出 χ=2c[1−1/√(1+z)]/H₀。令 x=1+z，角直径距离正比于 x^{-1}−x^(−3/2)。对 x 求导，极值条件为 $-x^{-2}$+(3/2)x^(−5/2)=0，得 √x=3/2，即 z=1.25。在这个示意宇宙中，同样实际大小的物体越远并非始终显得越小。这个结论不表示真实星系大小不演化，只说明几何效应与源演化必须分别建模。

54.6 不同探针怎样共同承重

超新星研究需要把可校准的光变和光谱特征与亮度联系起来，再比较不同红移处的距离关系。里斯等人在一九九八年的论文中利用高红移与近邻样本约束宇宙学参数，提出支持加速膨胀的观测结果。[3]它所改变的不是“天上多出一种暗星”，而是距离随红移变化的形状与既有减速模型之间的关系。尘埃、源演化、选择效应和标定仍是这类推断必须面对的竞争解释。

声学尺度提供另一种关系。早期物质与辐射中的传播过程，可以在后来的星系空间相关中留下尺度特征。艾森斯坦等人二〇〇五年的研究在红星系样本的相关函数中报告了重子声学峰。[4]峰不是天空上放着的一根刚尺，它要经过早期动力学、后期结构形成和巡天选择函数的共同说明，才能用于距离测量。其价值恰恰在于主要误差来源与超新星亮度标定不同。

微波背景则把温度、偏振和引力透镜信息引入同一早期演化问题。Planck 二〇一八年参数论文明确在基础 ΛCDM 模型下给出晚期参数推断，而不是把所有参数当作仪器直接刻度。[5]早期物理到晚期数值的翻译，依赖组分、初始扰动和演化假设。一个模型能够同时解释这些不同资料，是实质成功；但成功的条件也必须连同数字一起保留。

可以用一个二维示意说明联合测量的力量。设两个未知参数为 u、v，第一类资料约束 u+v=1±0.1，第二类约束 u−v=0.2±0.1。若误差独立且近似高斯，解得 u=0.6、v=0.4，各自标准差约 0.071。第一项单独留下整条带状区域，第二项横穿它才使两个参数分别得到限制。若第二项其实只重复测量 u+v，即使精度提高，沿另一方向的退化仍然存在。多探针的关键是约束方向和物理误差的差异，并非图表数量。

54.7 同意也可能来自同一种偏差

联合分析不能把“不同仪器”自动等同“独立证据”。两套巡天可能共享同一恒星标定，两组数据可能使用同一前景模板，不同参数结果可能预先固定了相同早期声学尺度。若这些共同条件偏移，多项结论可以整齐一致地偏移。共同结果必须追溯到共同条件，而不是以相互引用替代检验。

令两项测量的独立噪声标准差都是 σ，共同标定偏差的方差为 τ^2。它们平均值的总方差是 $\sigma^2/2+\tau^2$。若把共同偏差误算成两项独立误差，就会得到 $(\sigma^2+\tau^2)/2$，把标定部分也虚假缩小了一半。当 σ 不断下降，真实不确定度仍有 τ 的下限。宇宙学中看似越来越精确的小数，只有在相关误差模型正确时才表达相应知识。

亮度零点还给出另一种不能靠增加样本自行消除的退化。若所有标准化光源的真实光度 L 都同时乘以 q，而所推距离乘以 $\sqrt{q}$，流量 F 保持不变。在低红移关系中，这又对应 H_0 除以 $\sqrt{q}$。因此，只有相对亮度与红移的资料，不能独立确定绝对光度和绝对膨胀率；必须加入几何距离等外部标定。更多同类光源可以精确描出曲线形状，却不会凭数量自动提供缺少的长度单位。把参数退化说清，才能判断新增观测究竟增加哪一种知识。

此外，一片天空只能提供有限数目的大尺度模式。将天空温度展开为球谐系数，在理想的全天观测条件下、统计各向同性、近似高斯而忽略仪器噪声的情形，第 ℓ 阶功率估计的相对标准差为 $\sqrt{2/(2\ell+1)}$。例如 $\ell=2$ 时约为百分之六十三。这不是望远镜造得不好，而是该阶可用的独立模式有限。增加仪器灵敏度可以改善很多事情，却不能复制另一片独立天空。

上述限制不等于“宇宙学不能实验”。它迫使实验性表现为更细致的窗口控制：更换频段以检验前景，改变样本切分以检查选择偏差，使用不同物理时钟和尺度以破除退化。在无法重复整个对象的条件下，重复的是推断链中可替换、可分离、可交叉测定的环节。整体对象的可信性从这些局部可纠错关系中建立，而不是来自整体一词本身。

54.8 可观测宇宙的边界怎样产生

光传播速度有限不直接意味着可观测宇宙的当前半径等于 ct_0，因为传播过程中尺度也在变化。粒子视界的当前适当距离为 $a_0\int_0^{t_0} c\,dt/a(t)$，前提是该积分收敛并且模型在积分区间适用。它计算自模型初始边界以来可能送来信号的最远共动区域，不是我们已经通过所有信使实际看见的边界。

在 $a=(t/t_0)^p$ 且 $0<p<1$ 的示意模型中，积分可直接算为 $ct_0/(1-p)$。物质主导的 $p=2/3$ 给出 $3ct_0$，辐射主导的 $p=1/2$ 给出 $2ct_0$。同一个年龄乘光速，既不是唯一距离，也不包含膨胀史的全部信息。这个简单反例足以说明，把宇宙年龄和当前可达尺度混写会制造什么误解。

粒子视界与事件视界也不同。后者问的是从现在出发的信号能否在无限未来到达某处，其有关积分是从 t_0 到无穷未来的 $\int c\,dt/a(t)$。若未来按指数方式膨胀，积分有限；若未来始终按上述小于一的幂次膨胀，它就发散。前一问题依赖过去，后一问题依赖未来模型。不能因为观测到过去某个阶段的加速，就把某种未来边界当作直接拍下的景物。

电磁观测还受到物质透明度的限制。某个早期阶段的光即使在因果上能够关联到我们，也可能在传播中反复散射，不能保留原先的方向图像。不同信使有不同作用截面与释放条件，因而拥有不同的有效观测窗口。可观测宇宙不是一个与实验方式无关的透明球壳，而是因果可达性与信号保存机制共同限定的领域。全部宇宙的拓扑、范围和超出窗口的历史，则仍须与这一区域明确区分。

54.9 三维关系怎样进入宇宙学而不冒充宇宙方程

本书的三方程在这里承担结构审查。S=F(D,E)要求解释何种显露成立：同一红移点在距离标定、传播和选择条件改变后，是否仍支持原来的膨胀差别。D=G(S,E)要求说明差异序列如何构造：把不同红移看作不同宇宙阶段，需要已显露的谱线身份和演化联系，不能只按图上横坐标排列。E=H(S,D)则要求特征纠缠受到反向约束：光源、几何、探测器和标定的结合不能无限吸收所有残差，必须被不同显露之间的差异重新限定。这些是本书的哲学表达，不是取代 Friedmann 方程的数值算法。

以距离测量为例，S-D 张力表现为某一红移段的亮度显露偏离预期序列，它首先推动 E 改变：检查消光、源群体和传播条件，并回检能否同时解释颜色与时标。若这些条件在独立资料中受到约束，D-E 张力又推动 S 改变：原先只比较峰值亮度的显露方式必须扩展为完整光变、光谱和偏振等判别。若新的联合显露稳定而旧距离序列持续失败，S-E 张力才推动 D 改变：重新估计距离关系或膨胀史，并回到其他探针验证。三种推动不是按固定顺序完成的仪式，而是让一次修订必须交出旁证。

这里可以走通 D-E-S 的倒逼路径。先发现不同距离尺度或演化阶段之间的系统差别，再问现有标定和动力学联系能否承受这些差别，最后提出新的显露任务，使竞争解释产生可区分的读数。若只从差异直接跳到一种新成分的名字，E 未受检验，S 也没有更

新，这条路径尚未走完。反过来，E-D-S 的沉淀路径可以从成熟的原子频率、时钟与传播理论出发，建立跨红移差异，最终沉淀出可复用的宇宙距离对象。

54.10 整体性应当增加责任

宇宙学改变了“对象”的尺度，却不应降低成立对象的要求。对于无法操纵的整体，必须更加明确哪些关系来自观察，哪些来自理论闭合，哪些属于跨越观测边界的延伸。一个整体理论若总能通过增加未检验区域的任意性质解释任何结果，它便失去了约束。整体性只有让原来分离的读数承担相互检验的责任，才具有物理意义。

在这一意义上，宇宙成为可研究对象，不是人类终于站到了宇宙之外。恰恰是身处其中的有限位置，被谱线、时标、相关尺度、微波结构和传播方程连接成了一套可以修订的发生系统。宇宙的物理身份不再只是“万物的总称”，而成为有尺度、有历史、有可检验演化关系的对象。

这个身份不会消除未知。关于早期阶段的外推、背景近似的边界、引力理论的适用范围以及暗组分的解释，仍然开放。下一章将进入这些开放处，考察黑洞、引力波、暗物质和暗能量。它们并不是同一种“神秘事物”：有的已拥有高度稳定的跨窗口指认，有的仍以一组受约束的解释角色存在。只有把这几种完成程度分开，宇宙学的成就与尚未完成的工作才都能被准确看见。

本章参考资料

[1] Hubble, E. “A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae.” Proceedings of the National Academy of Sciences 15 (1929): 168–173. doi:10.1073/pnas.15.3.168。用于历史论文主题及问题定位。

[2] Hogg, D. W. “Distance Measures in Cosmology.” arXiv:astro-ph/9905116，修订版 2000。距离定义与条件参照；正文平坦物质主导模型、极值与视界积分为本章独立推演。

[3] Riess, A. G., et al. “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant.” Astronomical Journal 116 (1998): 1009–1038. doi:10.1086/300499。采用摘要所报告的样本比较与参数约束范围。

[4] Eisenstein, D. J., et al. “Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies.” Astrophysical Journal 633 (2005): 560–574. doi:10.1086/466512。采用论文摘要的声学峰探测与标准尺度意义。

[5] Planck Collaboration. “Planck 2018 Results. VI. Cosmological Parameters.” Astronomy & Astrophysics 641 (2020): A6，及勘误。arXiv:1807.06209。采用摘要中的温度、偏振、透镜及基础 ΛCDM 条件说明；不将 2018 结果写为现今唯一或最终参数。

第五十五章 黑洞、引力波、暗物质与暗能量 未完成之物

55.1 “看不见”不是同一种情况

黑洞、引力波、暗物质和暗能量经常一起出现在关于宇宙奥秘的叙述中。但它们被放在一起，容易造成一种错觉：仿佛凡不以普通光学图像出现的东西，都只有相同程度的猜测。事实上，某个对象是否可以直接发光，与它是否具有充分的物理证据没有简单对应关系；某种解释尚缺微观身份，也不意味着支持它的现象从未成立。

本章将三个层次分开。第一层是读数：干涉仪的时间序列、背景星系的形变、射电阵列的相关信号、超新星的流量。第二层是理论对象：这些读数在校准与动力学模型下支持怎样的质量分布、传播模式或膨胀历史。第三层是存在解释：究竟是什么承担这些关系，它的结构、组分和规律能否在更多条件中继续成立。三个层次相互联系，却不能因共享一个名称就被压成一层。

所谓未完成之物，因而不是没有任何成立之处的东西。它可能已经有可靠的动态身份，而内部微观结构未知；可能具有多类引力证据，而粒子候选尚未确定；也可能在一种有效模型中承担稳定角色，而不同根本解释仍难区分。把未完成分解为具体的未完成，才能决定下一项实验到底需要改变什么。

55.2 黑洞阴影不是黑色表面的照片

在不旋转、不带电、孤立且静态的理想模型中，Schwarzschild 半径为 $r_s=2GM/c^2$。这个尺度并不是物质表面的颜色边界，而与因果结构相联系。把 M 取作十个太阳质量，用太阳质量约对应三千米 Schwarzschild 半径的换算，可得约三十千米。这说明同样质量可以有非常不同的紧致程度，但半径公式本身不证明某个天体已实现该状态。

用光学现象检验紧致对象时，还必须区分视界、光子轨道和远方观察者的阴影。在同一理想模型中，圆形光子轨道位于 $r=3GM/c^2$；临界冲击参数，即从远处描述光线相对中心的横向偏移尺度，其值为 $b=3\sqrt{3}GM/c^2$。阴影角尺度联系后者，不等于把视界直径除以距离。这三个尺度的差异，已经足以推翻“亮环就是物体表面轮廓”的简单想象。

临界参数可以由有效势直接看出。把零质量粒子的径向势写成与 $(1-r_s/r)/r^2$ 成比例的函数，对 r 求导，极值条件给出 $-2r+3r_s=0$，即 $r=3r_s/2$。再用光子在该处的能量与角

动量关系，得到 $b^2=r^2/(1-r_s/r)=27r_s^2/4$。这是理想几何的计算；真实图像还叠加发光物质的分布、吸收、自旋和观察方向。

事件视界望远镜合作组二〇一九年的 M87 结果报告了中心亮度下降的环状射电结构，并用不同标定和成像方案检查其稳定性；其形态与广义相对论黑洞阴影的预期相容。[1]这里观测的是周围等离子体发出的辐射，经过强引力传播后留下的结构。说它为黑洞提供强证据是有根据的；说它拍到了视界以内的物质或奇点，则超出了读数能够承担的结论。

55.3 微小位移怎样成为传播事件

引力波探测的核心不只是测一段长度，而是测量随时间变化的相对光程，并将其与仪器响应联系。用 h 表示应变数量级，在适当几何响应下可以估算 $\Delta L\approx hL$。假设臂长四千米、应变为十的负二十一次方，相应长度变化约为四乘十的负十八次方米。这个示意数值的意义不是仪器有一把刻满微小格子的机械尺，而是相位、光学反馈、隔振和统计处理组成了可分辨这样的变化的系统。

一个局部扰动也可能改变光程。地面振动、散射光、电学瞬变和控制系统异常，都可能生成某种时间结构。可信信号因而需要跨站点相容的到达时差、波形和响应关系，同时检查环境监测与仪器状态。不能因为某条曲线看起来像理论图，就直接把每一次起伏宣布为宇宙事件。

二〇一六年发表的 GW150914 论文报告，两台 LIGO 探测器记录到同一次瞬态信号，其频率上升并与双黑洞旋近、并合和最终振铃的广义相对论波形相符。[2]这个结果把一种可传播的时空扰动和一种高紧致度双体动力学接到同一证据链上。它的力量在于连续演化各阶段可以共同约束参数，不只是存在一个超过背景的峰。

仍应分清两件事：数据中是否存在共同的天体信号，和某个具体模型在多大精度内解释了信号。前者可以相当稳固，后者仍有可检验的细节。未来若发现某段波形偏离现有模板，未必意味着过去的全部信号都是噪声；它也可能要求改进源模型、传播规律或强引力理论。可靠对象与可修订解释能够同时成立。

55.4 频率加快怎样暴露质量组合

双体旋近给出一个可以独立推演的例子。设两质量为 m_1、m_2，总质量 $M=m_1+m_2$，约化质量 $\mu=m_1m_2/M$。在缓慢运动、近圆轨道的最低阶近似中，角频率 Ω 与间距 r 满足 $\Omega^2=GM/r^3$，轨道能量 $W=-G\mu M/(2r)$。主导引力波频率 $f=\Omega/\pi$，因为四极矩在一圈内重复两次。

以最低阶四极辐射功率 P=32G⁴μ²M³/(5c⁵r⁵)和能量平衡 dW/dt=−P 相结合，可消去 r。由于 W 正比于−f^(2/3)，而 P 正比于 f^(10/3)，得到 df/dt 正比于 f^(11/3)。系数整理为 df/dt=(96/5)π^(8/3)(G𝓜/c³)^(5/3)f^(11/3)，其中𝓜=μ^(3/5)M^(2/5)称为啁啾质量。这样一个波形时间序列首先约束的是质量组合，而不必立即分别给出两个质量。

对等质量系统 m₁=m₂=m，代入可得𝓜=m/2^(1/5)，约为 0.871m。若啁啾质量为二十六个太阳质量，等质量近似对应每个约三十个太阳质量。但不等质量系统也可具有同一啁啾质量，需要更高阶相位、自旋和其他波形特征进一步分辨。把一个组合参数说成直接称出了两个物体，会抹去推断真正增加知识的步骤。

这个公式还给出一个不依赖距离的形状检验：频率提高到两倍时，最低阶频率增长率应增加约十二点七倍。若时间序列只是一个恒定频率的机械振动，便不满足这个关系。靠近并合时缓慢运动近似失效，不能坚持用该幂律拟合全部信号，而要接入更高阶近似和数值相对论。理论在何处停止适用，本身就是对象身份的一部分。

对于宇宙学距离上的源，观测时间经过红移伸长，波形首先确定红移后的质量组合 (1+z)𝓜 源。没有额外红移信息，不能从这项相位尺度单独分开源质量和红移。引力波虽绕过电磁发光机制，却没有摆脱传播条件。新的窗口改变了哪些联系被观测，也引入了必须公开的新退化。

55.5 旋转曲线中的缺口是什么

暗物质问题可以从一个简单条件开始。对球对称质量分布中的近圆轨道，v²=GM(<r)/r。如果外围速度保持近似恒定，则 M(<r)必须随 r 近似线性增长，相应密度 ρ=(1/4πr²)dM/dr 近似按 r⁻² 下降。这是给定引力规律和几何近似下的质量推断，不是望远镜直接数出了不可见粒子。

以 v=200 千米每秒、r=20 千秒差距作数量级演算，利用 G 约为 4.30 乘十的负六次方千秒差距乘速度单位平方每太阳质量，得到 M(<r)约为一千八百六十亿个太阳质量。若已估计的恒星和气体质量不足以支持这个速度，就出现动力学质量与可见物质模型之间的缺口。它要求解释，但尚未单独指定是哪一种粒子承担缺口。

实际星系不是严格球形，倾角、距离、恒星质量光度比以及气体的非圆运动都会影响推断。以倾角为例，视线速度约为 v 观=v sin i；若真实倾角三十度而误用四十五度，推断速度将低约百分之二十九，在半径不变条件下推断质量会低一半。这说明误差检查不是形式性的退让，而有可以计算的影响规模。

同时，不能因为某些误差存在，就认为所有异常都可随意归于误差。必须问：一种误差模型能否同时解释不同半径、不同星系、透镜和结构形成资料？把每个目标的倾角或恒星质量任意调到需要的值，会丢掉这些参数原本受到的独立约束。解释缺口不能以破坏其他已经成立的关系为代价。

55.6 增加质量与改变规律怎样竞争

米尔格罗姆一九八三年提出修改低加速度动力学作为隐藏质量假说的替代路线。[3]其重要性在本章不在于宣布哪条路线已经终结，而在于使竞争问题具体化：同一组速度读数，可能要求改变物质分布，也可能要求改变从物质到加速度的关系。二者需要在更多观测中承担不同后果。

用一个低加速度形式 $a^2/a_0=GM/r^2$ 作演示，代入圆轨道 $a=v^2/r$，便得 $v^4=GMa_0$。半径从结果中消失，于是外围可以出现近似恒定速度，并把速度四次方与重子质量联系。这个短计算说明，平坦旋转曲线本身不具有唯一的解释出口。但这条非相对论关系尚未自动给出光怎样弯曲、宇宙怎样演化或星系团怎样碰撞；完整竞争理论必须继续补足这些部分。

反过来，给定暗晕密度函数并拟合速度，也并未自动解释暗物质的产生、分布与微观相互作用。若它的额外自由度只用于逐个吸收残差，就不能据拟合优度宣称已经完成物理说明。双方都应面对同一种要求：用先行规定的结构产生新的区分，并在未用于定参的资料中承担风险。

合并星系团把问题推进到不同物质成分的空间分离。Clowe 等人二〇〇六年的研究比较弱透镜重建与热气体、星系位置，发现引力分布不简单跟随主要发光气体。[4]这种关系为不可见质量提供强证据，也给只按局部可见物质修改一个力函数的方案提出严厉约束。然而透镜质量图仍是经几何与引力模型重建的量；它没有给某一种暗粒子拍摄显微照片，也不能把一切可能扩展理论仅凭名称预先排除。

55.7 没有候选事例怎样形成知识

若暗物质可以与探测器发生某种散射，可以把天体动力学的解释角色接到实验室反冲事件上。这是身份深化的重要路径：一个候选对象必须同时满足宇宙中的丰度和分布、实验室中的相互作用，以及其他已有实验约束。仅有“非常难看见”这一性质，不足以让任何新粒子成为合格暗物质。

一个最简单的无背景计数例子足以说明空结果的意义。若预期信号数为 s，观察到零个事件的概率为 e^{-s}。按这个特定模型，以百分之十作为单侧尾概率，解得 s 约小于

2.30。这并不意味着自然界不存在暗物质，而是排除那些在该曝光量、阈值、效率和散射假设下本来应产生更多事件的参数。若背景不为零、效率不确定或能谱结构重要，就不能直接照搬这个数值。

以 LZ 的四点二吨年曝光研究为例，其论文给出针对特定 WIMP 核反冲搜索的限制，而不是对一切暗物质实现方式的一次总判决。[5]没有把某种候选探测出来，会改变候选空间，也可能推动新靶材料、新能区或新耦合方式的搜索。它既不是徒劳，也不等于整个引力证据链被消除。

这里最容易混淆“目前某种证据尚不足”和“已经证明绝不存在”。前一句规定后续任务，后一句要求远强得多的论证。科学能够通过限制一步步排除结构，即使最终留下的对象不同于最早设想，也不因此使早期所有约束失去意义。未完成的身份可以有坚实的负面边界。

55.8 暗能量不是暗物质的稀薄版本

暗物质在通常宇宙学模型中承担聚集、引力势和结构增长的角色；暗能量则用来描述影响大尺度膨胀的另一类成分或有效项。二者都被冠以“暗”字，不意味着具有同一种物态，也不意味着前者数量多一点就可以得到后者的功能。

把均匀组分的压力写成 $p=w\rho c^2$，局部能量守恒关系为 $\dot{\rho}+3H\rho(1+w)=0$。若 w 为常数，积分得到 $\rho\propto a^{-3(1+w)}$。无压物质 $w=0$ 随体积稀释，辐射 $w=1/3$ 还包含红移损失，$w=-1$ 则使密度保持常数。这里 ρ 是质量当量密度，w 是状态参数，并不是由“暗”这个名称定义出来的性质。

在平坦、物质与该组分共同主导的简单模型里，若忽略其他成分且该组分取 $w=-1$，当前减速参数可写 $q_0=\Omega_m/2-\Omega_\Lambda$。取纯示意数值 $\Omega_m=0.3$、$\Omega_\Lambda=0.7$，得 $q_0=-0.55$，对应加速；若把两者都改为无压组分并保持总密度一，则 $q_0=0.5$。相同总质量当量密度不能决定相同膨胀加速度，压力与组分演化承担不可省略的作用。

负压力也不宜画成到处向外推物体的小箭头。均匀组分的引力效应进入整体时空演化方程，与局部压力梯度推动流体的力不是同一计算。宇宙学常数可以形式上移到物质一侧成为有效应力，也可以留在几何一侧。写在等号哪边并没有解决其根本物理来源。能在方程中表示，与已经解释为什么存在，是不同层次的完成。

55.9 拟合出变化还需要追问什么

允许 w 随时间变化，就能增加模型适应距离资料的能力，但也增加了自由度。常用示意参数式 w(a)=w₀+wₐ(1−a)可由守恒关系积分，得到 ρ(a)/ρ₀=a^[−3(1+w₀+wₐ)] exp[3wₐ(a−1)]。这个式子不是任何微观场的必然定律，它是在有限观测窗口中整理演化的一种规定。

若取 w₀=−1、wₐ=0，密度回到常数；若固定 w₀=−1 而令 wₐ=0.2，则在 a=0.5 时密度比为 2^0.6e^(−0.3)，约为 1.12。现在若只在狭窄红移窗观测，百分之十量级的组分变化还要通过总膨胀率和距离积分，才进入实际读数。它可能与物质密度、曲率、距离零点等因素混合，不能把参数变化率直接当作某个仪器测到的流量变化率。

DESI 第二次数据发布的宇宙学分析在二〇二五年报告，联合不同外部资料时，对暗能量演化的约束与偏好会随数据组合而变化。[6]本章将它看作进一步判别的具体任务，而不据此宣布一种微观暗能量已经被发现。需要检查参数化的适用范围、共享标定、样本系统误差，并问另一种动力学解释是否在相同资料中具有不同预测。

同样的膨胀史有时可以由不同有效描述重现。只测背景距离关系，可能尚不足以区分新组分与引力规律改变；结构增长、透镜和其他传播信息可以增加区分方向。这里的未完成不是因为没有更多术语，而是因为某些候选关系在现有窗口中还未充分分离。把窗口换得更有判别力，比给残差取更富想象力的名字更能推进对象的完成。

55.10 未完成的对象怎样继续发生

对本章各类问题，S=F(D,E)要求新的显露不是凭空宣布：亮环、啁啾、质量图和膨胀残差，分别依赖怎样的差别结构和发生联系。D=G(S,E)要求比较序列有物理支撑：由一段频率演化区分质量，由不同半径区分物质分布，都须说明哪些条件固定、哪些可以变化。E=H(S,D)则要求用已经成立的显露和差异收紧联系，禁止把一个“未知机制”当作可以解释任何结果的容器。三方程在这里组织审查义务，不替代波形、透镜或场方程。

以暗物质为例，S-D 张力是稳定旋转读数与可见质量所预测序列不合，它推动 E 改变，检查几何、重子分布和引力关系，再用其他读数回检。D-E 张力是某个候选联系在星系、星系团和早期宇宙间不能保持相容，它推动 S 改变，转向能区分候选的新透镜组合或实验室反冲窗口。S-E 张力则出现在新窗口提供稳定约束而旧参数排序失效时，推动 D 改变，重新组织允许的质量、截面或动力学区间，并回检此前成功的领域。未完成由此成为一组有方向的工作，而不是永久等待。

可以把这一过程走成 S-E-D 的进入路径：从新的可靠读数进入其传播与校准关系，再形成能够跨条件比较的对象差异。引力波的频率时间结构正是这种路径的清楚例子。另一种 D-E-S 倒逼路径则从跨尺度的质量缺口出发，迫使候选机制承担不同后果，再为这些后果设计新的显露窗口。只有走完三个环节，新增名称才可能成为可持续的物理身份。

本书的存在解释进一步主张：对象的物理规定在这些互相约束的发生系统中成立和发展。这个主张不能被偷换成 SDE 已经解决黑洞内部、暗物质粒子或暗能量来源。它应当接受自己的限制：如果不能指出哪些规定因何种新关系而改变，便没有比原来的物理解释多说明什么。哲学的作用是使成立条件与修订责任清楚，而不是提前占有尚未取得的实验成果。

黑洞和引力波已经展示了跨窗口指认如何获得高度稳定性；暗组分则展示了受约束的角色如何等待更具体的身份。这两面共同说明，物理世界的发生并不要求每个对象一次完成全部规定。能够稳定计算、能够跨条件追认、能够在失败时指明需要改变的关系，就是继续研究的现实起点。未完成之物不是科学边缘的一团黑暗，而是下一轮差异、联系与显露已经开始组织的地方。

本章参考资料

[1] Event Horizon Telescope Collaboration. “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole.” Astrophysical Journal Letters 875 (2019): L1. doi:10.3847/2041-8213/ab0ec7。采用摘要的观测、重建稳定性与模型相容范围；Schwarzschild 光子轨道计算为本章示意推导。

[2] LIGO Scientific Collaboration and Virgo Collaboration. “Observation of Gravitational Waves from a Binary Black Hole Merger.” Physical Review Letters 116 (2016): 061102. doi:10.1103/PhysRevLett.116.061102。采用摘要的事件与波形结论；最低阶啁啾关系为本章教学推导，不用其替代完整参数估计。

[3] Milgrom, M. “A Modification of the Newtonian Dynamics as a Possible Alternative to the Hidden Mass Hypothesis.” Astrophysical Journal 270 (1983): 365–370. doi:10.1086/161130。参照原文低加速度假设；正文四次方关系及与暗晕的比较为独立分析。

[4] Clowe, D., et al. “A Direct Empirical Proof of the Existence of Dark Matter.” Astrophysical Journal Letters 648 (2006): L109–L113. doi:10.1086/508162。采用摘要关于合并星系团、弱透镜与气体空间关系的报告，区分其质量证据与微观粒子身份。

[5] Aalbers, J., et al. “Dark Matter Search Results from 4.2 Tonne-Years of Exposure of the LUX-ZEPLIN (LZ) Experiment.” Physical Review Letters 135 (2025): 011802. arXiv:2410.17036。采用摘要所限定的 WIMP 核反冲搜索范围；零事件泊松例子是理想模型，不是对该实验分析的复现。

[6] DESI Collaboration. “DESI DR2 Results II: Measurements of Baryon Acoustic Oscillations and Cosmological Constraints.” arXiv:2503.14738，2025 修订版。采用摘要关于数据联合与演化暗能量约束的条件说明，不作为最终微观身份的证明。

第十二编 新物理世界的下一轮创造
科学典范与 SDE 研究纲领

第五十六章
视觉、听觉、触觉的 SDE 怎样在现代仪器里聚合

56.1 同一次振动为什么出现三种面貌

实验台上有一块正在振动的薄板。人可以看见它的边缘颤动，听见空气中的声音，也可以将手指轻放在边缘感到震颤。倘若把这些经验简单排成三栏，再说视觉提供显露、听觉提供差异、触觉提供实体，就已经破坏了本书的方法。每一种感觉都以自己的完整 SDE 成立：可辨的现象、分辨现象的差异序列以及使特征相互规定的纠缠关系，不能由三种器官分别承包。

看见颤动，不只是眼前出现一条边缘。照明、观察角度、背景对比和时间分辨共同决定什么能够成为边缘的移动；改变这些关系，显露也会改变。听到音高，不只是接收一个孤立频率；声源、传播路径、听觉辨别和比较程序使某种差别成为音高差别。触到震颤也不是越过一切媒介直接接触物自体；接触力、皮肤响应、手指位置与运动方式，都参与所感对象的规定。

这三种完整系统并不因此被封闭在各自的私人世界。薄板的某些运动关系可以使它们发生稳定协调：振动频率改变，视觉轨迹、声压记录与接触响应可能共同改变。但协调需要建立转换，不能因为发生在同一张实验台上就宣告完成。薄板振动很强而发声很弱，可能来自辐射效率低；听到很响的声音而接触点振动很小，可能因为手指恰在接近节点的位置。所谓同一物，必须经受这种不同显露之间的差异。

现代仪器把这项工作扩展到感觉难以直接分辨的尺度。它们并不只是把人眼放大或把人耳变灵敏，而是在光、力、电、声与数字记录之间组织可以校准的转换。对于本书，问题也随之改变：不是机器究竟代替了哪一种感觉，而是多套完整的发生系统，怎样形成能够回检的共同对象。

56.2 仪器不是把三种感觉装进一个盒子

设想用激光测振仪观察薄板，用加速度传感器贴在板上，再用麦克风记录附近声场。三个设备都输出电压，并不表示它们已经测量同一个量。光学通道可能给出视线方

向的速度，加速度传感器给出特定方向的惯性响应，麦克风则以自身频率响应把局部声压转成电信号。单位、位置、方向和传递规律都需明确。

NIST 研究者关于数字测振的原始报告说明，激光测振方法可用于加速度计校准，并与其传统干涉测量系统进行比较。[56-1] 这里值得注意的不是一种设备战胜另一种设备，而是光学运动记录能够成为机械响应校准的一部分。参与比较的每个系统，都必须先有足以承受比较的定义与误差说明。

不要把仪器拟人化。光学传感器不是另一个有视觉经验的人，接触探头也不具有手指的全部触觉。所谓视觉、听觉、触觉系统在仪器中聚合，是指原来在这些通道中展开的辨别与操作关系，经过新的物理转换被重新组织；不是说感觉内容被原封不动搬入电子线路。红外图像的颜色可以人为编码，超声记录可以被变成可听声音，这些显示方式并不使红外辐射自然具有那种颜色，也不使超声本来就是低频声音。

显示形式可以更换，物理约束仍须保留。若更换配色表就改变了对温度排序的判断，问题可能出在显示设计；若对原始数据进行相容转换后仍出现同样的异常，才需要追查测量链与样品。人的读图行为进入知识形成，却不能凭这一事实取消仪器中相对独立的物理检验。

56.3 用一个振动算例建立转换责任

自行规定一个理想振动：测量位置的位移为 $x(t)=A\cos(2\pi f t)$，取振幅 A 为十微米，频率 f 为一百赫兹。对时间求导，速度振幅为 $2\pi fA$，约每秒零点零零六二八三米；加速度振幅为 $(2\pi f)^2A$，约每二次方秒三点九四八米。这里三个量的数值不同，正是同一运动在不同量度中的相容结果。

如果光学通道报告十微米，而加速度通道报告三点九五米每二次方秒，不能因为数字相差巨大就说它们矛盾。反之，两个显示屏恰好都出现数字十，也完全不能证明相互支持。比较的单位不是屏幕上的字符串，而是由模型和校准联结的量值。位移的峰值与加速度的均方根值同样不能直接比较；对这个正弦例，后者应为加速度振幅除以根号二。

相位也带来约束。位移与加速度相差半个周期，速度与位移相差四分之一个周期。若软件为了让曲线看起来一致，任意移动各条曲线，可能把理论应有的相位关系一起擦掉。假设某通道额外延迟两毫秒，在一百赫兹处就产生七十二度的相位差。它不等于薄板上必然存在另一个机械模式；也不能未经检查就全部归于电子延迟。

可检验的区分办法是改变频率。固定时间延迟引起的相位差随频率成比例变化；某些机械共振附近的相位则以不同规律变化。先用独立已知信号测定仪器延迟，再把该修正用于薄板记录，才形成一次有内容的回检。若把每个频点的延迟都当作任意参数，任何曲线都能被对齐，比较便失去了风险。

麦克风不能仅凭上述导数链自动加入。声压与薄板位移之间还需要声辐射、距离、边界反射等具体关系。最简单的线性模型可以在频域写成 p(f)=H(f)x(f)，但 H 必须由理论条件或独立测量约束，不能定义成每次记录的商之后，再用它证明所有记录都被解释。若换到另一个位置，原来的 H 一般也需要重新核对。

于是，这个小例子同时显示聚合的能力和限度。位移、速度、加速度可以在规定运动模型内相互换算；声压需要额外结构才能与之连接。关系越多，共同对象的规定就越丰富；但额外关系也增加了可能失败的地点，不能把连接的数量直接当作证据的强度。

56.4 校准不是事后把三条曲线修成一样

国际计量学词汇将校准规定为在指定条件下建立标准量值与示值的关系，并据此从示值取得测量结果的操作；它还明确区别校准与调整。[56-2] 这一规定让我们看到，校准首先承担转换责任，并不是把仪器调到预期答案就算成功。

例如某电压通道满足 y=kx+b。应当用独立参考输入估计灵敏度 k 与零点 b，再对未知样品使用 x=(y−b)/k。若一边观察未知样品，一边任意改变 k 与 b 使它贴合理论，这叫参数配合，不能冒称已经建立独立校准。适当的研究可以包含联合估计，但必须说明哪些参数可被数据分别辨认，保留哪些观测用于检验。

接触式测量还可能改变样品。给轻薄结构贴上传感器，会增加局部质量和机械约束；激光虽不需要相同接触，也有视线方向、反射和散斑等自身条件。没有一种通道仅凭被叫作非接触，就获得绝对无扰动的地位。需要比较传感器安装前后、不同质量负载以及不同光学几何，判断所测对象是否仍属于同一模型。

设简化振子的刚度不变，原质量 m 增加一个小附加质量 δm，固有频率由根号(k/m)变为根号[k/(m+δm)]。若 δm/m 为百分之四，频率比为 1/√1.04，约零点九八零六，即下降约百分之一点九四。两个仪器因此报告不同频率，未必有一个失灵；它们可能已经通过不同操作造出了不同的被测系统。

这时聚合不能强行把两者平均成一个中间频率。应先回到对象边界：需要报告未负载结构，还是包含传感器的总成？模型若能够预测负载依赖，再用不同负载复验，差别

反而帮助我们确定结构参数。所谓恢复共同对象，不是删除操作的痕迹，而是说明哪些痕迹可计算、哪些尚未控制。

56.5 三路一致为什么仍可能一起出错

考虑三个通道都引用同一时间基准。若这只时钟的频率标定偏差，三条记录可能高度一致，却共同偏离外部参考。独立通道的数量与独立证据的数量并不相同。仪器外壳不同，不能证明误差来源彼此独立；软件共享同一个预处理模块，也可能把同样的偏差分发到全部输出。

用一个简单误差模型说明。两次估计分别写成 $z_1=z+c+e_1$、$z_2=z+c+e_2$，其中 c 是共同误差，e_1 与 e_2 是独立误差，均值均设为零。若 c 的标准差为二个单位，而两个独立误差各为一个单位，则每项估计的方差是五，平均值的方差却是四加二分之一，即四点五。平均后标准差约二点一二一，仅比原来的根号五略小。

若错误地把两项估计视为完全独立，就会把平均值方差算成二点五，低估了共同误差。差值 z_1-z_2 又恰好消去了 c，所以相互比较可以很好，却无法发现共同偏移。这个算例并非否定融合，而是明确融合必须处理哪些关系。差值检验擅长发现相对不一致，外部参考擅长检验共同标尺，二者任务不同。

计量溯源也不是一条线的机械传递。相关词汇指出，多输入测量模型的校准关系可能构成分支或网络，且具备溯源性并不自动保证不确定度适合用途。[56-3] 对多模态测量，这意味着应追问各通道在哪里共享时钟、标准和模型，而不能把一份证书当作整个聚合系统无误的证明。

56.6 从重物到电压：聚合改变了什么对象

基布尔天平提供了不同于薄板的例子。其理想关系在称重阶段写成 $mg=KI$，在运动阶段写成 $U=Kv$，K 表示相应电磁几何因子。若两个阶段的 K 能够相容，消去它便得 $mgv=UI$。NIST 的装置说明把力平衡、运动感应、电学量度与光学速度测量连接起来。[56-4] 这里比较的是两个工作阶段构成的虚拟功率关系，不应想成称重时重物正在按速度 v 持续输出机械功。

对本章而言，关键不是重新讲一遍千克史，而是注意“这块物体有多重”经过怎样的聚合才变成一个精确可传递的质量结果。接触承载、空间运动、光学记录、电磁关系与局部重力量度彼此制约。任何一个通道都不能独占物体的全部身份，也不能被机械安置到 S、D、E 的某一个固定角落。

这个例子还防止我们把聚合理解为单向翻译。不是先有一个无需说明的纯机械事实，再把它原样译成电压；称重与运动阶段能否共享 K，本身需要装置稳定性检验。反过来，电学关系的精密也不能替代几何对准与力学控制。一个新的公共对象通过这些不同结构互相限制而成立，其稳定性来自能够回到具体失败点的网络。

56.7 完整 SDE 怎样相接，而不被拆散

现在沿 S—E—D 进入路径处理一次三通道分歧。先取得相位不相容的显露 S，进入延迟、负载和声辐射的关系组织 E，再设计频率扫描、传感器拆装与空间移动的差异序列 D。若独立延迟校准消除了差异，就保留原薄板模型；若拆装使共振稳定移动，则需要将负载纳入对象规定。路径终点不是获得新名称，而是形成可以回检的实验安排。

另一路 E—D—S 的沉淀，从候选的耦合模型出发，决定在哪些位置、哪些频率测量，预言三通道应当怎样共同改变。两条路径均在每种感觉或仪器通道的完整 SDE 之间建立连接，不把视觉等同显露，不把触觉等同纠缠，也不把听觉等同差异序列。研究者、仪器作用与被测结构的号位分析，与这三维分析同样不能互相替换。

三方程在聚合中提出具体要求：S=F(D,E)规定记录不能脱离采样和耦合模型；D=G(S,E)使新增测量针对尚未区分的解释；E=H(S,D)要求结构修正回应实际跨通道证据。三原理则表现为相位记录与操作安排的张力推动延迟模型改变，记录与模型的张力推动校准序列改变，模型与既有操作无法相接时推动重新选择显露指标。第三维的改变必须返回最初分歧接受检验。

一种强竞争解释会说，这一切只是传统测量理论、系统辨识和传感器融合。本书不能靠重命名否认这种解释。SDE 只有在帮助识别对象边界、区分完整通道与显示形式、发现共享误差以及组织更有辨别力的回检时，才提供额外的研究价值。若所有工作已经由现成方法清楚完成，发生论应准确说明其成就，而不是宣称功劳从未存在。

现代仪器中的聚合因此不以所有人最终看见同一幅图为目标。相同的物理关系可以用图像、声音或触觉提示来表达，不同的人也可以通过不同方式进入它。公共性成立于转换有据、误差可追踪、差别能回检。物理世界由此扩展到任何单一感觉都不能独立维持的范围，同时又始终保留返回具体操作的道路。

本章参考资料

[56-1] B. F. Payne, R. A. Allen, C. E. Hood, Improvements in Accelerometer Calibration at NIST Using Digital Vibrometry, Proceedings of the 87th Shock and Vibration Symposium，2017。阅读

作者摘要所述校准与系统比较：https://www.nist.gov/publications/improvements-accelerometer-calibration-nist-using-digital-vibrometry。

[56-2] JCGM，International Vocabulary of Metrology，第三版，2.39 Calibration 及说明：https://jcgm.bipm.org/vim/en/2.39.html。

[56-3] JCGM，同书 2.41 Metrological traceability 及多输入模型注释：https://jcgm.bipm.org/vim/en/2.41.html。

[56-4] NIST，Kilogram: The Kibble Balance，阅读称重、运动阶段及速度量度原理：https://www.nist.gov/si-redefinition/kilogram/kilogram-kibble-balance。本文仅使用装置原理，不以页面历史规划语句描述当前制度状态。

第五十七章 微积分之后
非局域、随机、离散与计算物理

57.1 不是微积分结束，而是问题的组织扩大

一块材料中出现裂缝，经典弹性方程最常使用的位移导数在裂缝处不再按原来方式存在。一个微粒被液体持续碰撞，记录有清楚的统计规律，却没有实验者能够完整追踪的每次分子作用。一台计算机给出几百万个网格值，图像很像连续流动，但机器实际更新的是有限数量的变量。这些情况共同提出一个问题：物理的数学规定是否必须始终采取逐点光滑、单条确定轨道的形式？

章题中的“之后”不是废除微积分的年代宣告。偏微分方程、积分关系、随机过程和离散结构之间，既有不同对象规定，也有可以严格建立的转换。重要的是辨认转换保留什么、需要什么条件、改变了怎样的检验任务。如果把连续当作高级、离散当作低级，或者反过来把数字计算当作自然界本来就是代码的证据，都会过早关闭问题。

前面已经讨论过差分、尺度截断以及复杂系统的预测限度。本章不再用同一条网格稳定性条件讲述它们，而把重点放在三个新的接口：局部方程如何容纳跨距离作用，连续边值问题如何成为有限维变分问题，随机采样怎样成为有证据责任的求解操作。最后再追问，连接关系本身能否先于长度和角度进入离散物理。

这些接口使计算不只是替人完成更多算术。它参与规定哪些量被保留、哪些过程可被询问，以及哪些结果可以在不同实验和算法之间迁移。但参与规定不意味着任意创造答案。恰恰因为数学组织更灵活，研究者更需要说明哪一部分来自物理，哪一部分来自计算选择。

57.2 非局域关系并非一句“一切影响一切”

用一维简化模型表示跨距离耦合，令响应写成 $Lu(x)=\int K_\delta(r)[u(x+r)-u(x)]dr$，积分只取规定邻域。核 K_δ 说明距离为 r 的差别以多大权重进入响应，δ 规定作用范围。它不是把周围所有因素列成清单，而是把相隔位置之间的制约写进可以计算的关系。实际力学模型还必须指定密度、力的方向、边界和本构条件。

若核对称，奇次矩消失；若 u 在邻域内足够光滑，展开 $u(x+r)$，可以得到 Lu 约等于二阶导数乘以核的二阶矩的一半。这个局部极限并不来自“距离足够小”一句话，而需要核随 δ 如何缩放、相关矩是否有限以及解的正则性等条件。若这些条件缺失，直接把积分换成二阶导数可能删掉模型最重要的内容。

自行选一个易算的核：在绝对值 r 不超过 δ 时令 Kδ(r)=3/δ³，其余为零。它的二阶矩为二。代入 u(x)=x²，差为 2xr+r²，奇项积分消去，结果正好为二，与 u 的二阶导数相同。再代入 u(x)=x⁴，则结果为 12x²+6δ²/5；局部二阶导数只有 12x²。这个额外项明确记录了有限作用范围。

当 δ 减半，该多项式例中的附加项缩至四分之一。如果研究者在一系列已知形状上检验该核，就可以区分局部极限与有限范围模型；但仅在二次函数上测试，两者恰好一致，无法识别 δ。这说明选择检验对象本身具有数学责任。一个模型通过过于特殊的测试，不代表它已在需要的尺度上成立。

Silling 关于近场动力学的原研究提出用新的连续体关系处理裂缝等不连续，并讨论相关波传播与稳定性。[57-1] 本章的标量核只是用来说明结构，不冒充其完整材料理论。裂缝、界面和边界邻域会改变作用关系；增加非局域性不能自动解决所有本构问题，也不能凭一个作用范围参数就拟合全部断裂现象。

尤其不能把这里的空间非局域核与量子纠缠、超光速通信或遥远事物无条件互相影响等同。积分方程描述的是什么量、时间演化怎样规定、它属于何种有效尺度，都需要另行说明。不同理论共享一个词，并不共享全部物理结论。

57.3 有限元首先改变“解”的承载方式

偏微分方程通常要求在区域中满足某种关系，并在边界上满足另外的条件。复杂边界使显式解析式难以得到，然而没有简洁公式并不等于没有可检验答案。Courant 一九四三年的论文讨论变分近似，并指出可在三角剖分上采用分片线性函数，把积分问题化为节点值的方程。[57-2] 这一思路的重要性在于，近似解先被安置到一个规定的函数空间里，再接受物理条件的约束。

用最小的一维例子把它算出来。在无量纲区间零到一上求 −u″=1，边界 u(0)=u(1)=0。确切解是 u=x(1−x)/2，但先假装我们不使用这个结果。将方程乘以在两端为零的试验函数 v 并积分，分部积分后得到 ∫ u′v′dx= ∫ v dx。这里把对 u 的二阶导数要求改写成一阶导数的积分关系。

只把区间分成两个等长单元。取一个帽形函数 φ：左半段为 2x，右半段为 2(1−x)，两端为零、中点为一。令近似解 uₕ=aφ，同时取 v=φ。由于 ∫ (φ′)²dx=4，∫ φ dx=1/2，方程化为 4a=1/2，所以 a=1/8。仅一个未知数，就得到了整个分片线性的近似形状。

在中点，近似值恰好与确切解相同。但在 x=1/4 处，近似值为 1/16，确切值为 3/32，两者相差 1/32。不能因为唯一的内部节点完全正确，就宣布整条曲线已经精确。若需要最大挠度，这个网格表现很好；若需要局部斜率或应力，误差任务就不同。计算质量必须相对于所问物理量来判断。

这个例子还揭示了“弱形式”的准确含义。弱并不表示可以不遵守定律，而是把定律放进对一类试验函数的积分要求中。有限元近似只在所选择的有限空间里满足这些要求；扩大空间、改进网格和检查边界条件，构成后续验证工作。若材料参数或载荷设错，网格加密只会更准确地解出一个错误的物理问题。

在真正的多维问题中，三角形、四面体或其他单元帮助处理几何。它们不是材料真的被切成这些小块的证据。节点数增加代表计算表示更细，也不必然代表获得更多独立实验信息。有限元把无限维关系转成有限维方程，却没有取消模型、表示与样品之间的区别。

57.4 离散守恒与拓扑：先问谁与谁连接

还有一种组织顺序并不先从每一点的导数出发。将区域分成若干控制体，对每个体记录总量，并把相邻体之间的通量使用同一个带方向的数值。一个公共界面上，甲的流出就是乙的流入；将所有体的变化相加，内部通量成对抵消，只剩外边界与源项。这种整体守恒可以在有限网格上精确保持，不必等待网格趋于零以后才出现。

取三个容器，交换量分别记为从甲到乙的 j_1、乙到丙的 j_2、丙到甲的 j_3。各自总量变化率为 $-j_1+j_3$、j_1-j_2、j_2-j_3，和恒为零。即使三个交换量计算得并不精确，只要使用同一界面值并保持符号配对，内部交换也不会凭空创造总量。不过，守恒正确仍不保证局部分布正确；把全部 j 设成零同样守恒，却可能错过实际输运。

这段算式的第一层只依赖连接和方向，还没有用到容器之间的距离。距离、截面积和材料系数进入下一层，决定给定浓度差究竟产生多大通量。区分这两层，能防止把几何精度与结构一致性混为一谈：一个画得很精细的网格，仍可能因接口符号出错而漏掉守恒；一个粗网格，也可以正确保留某些结构恒等式。

用三角形边界再作检验。从顶点甲走到乙，再走到丙，最后返回甲，三个有向边的起终点相加为零。一般的离散边界算子满足“边界的边界为零”。这不是一条特殊材料定律，而是连接组织的恒等关系。把适当的物理量放在点、边、面或体上，可以使某些离散梯度、旋度与散度关系保持相容；具体物理系数仍需另行加入。

拓扑在这里没有神秘地替代动力学。它先限制哪些环路、边界与守恒关系允许存在，再由几何和本构赋予量值。如果删掉一条连接边，系统能够交换的路径改变；但仅知道图是连通的，还不能算出传输需要多久。离散结构创造了一种可计算的对象组织，同时保留了自己无法独立回答的问题。

57.5 随机不是计算失败时撒入的噪声

考虑一维无偏随机行走。每隔时间 τ，粒子以相同概率向左右移动长度 ℓ，且各步独立。走 n 步后的平均位移为零，位移平方的平均值为 $n\ell^2$，因为不同步长的交叉平均为零。令 $t=n\tau$，便得平均平方位移为 $2\kappa t$，其中 $\kappa=\ell^2/(2\tau)$。κ 在这里是扩散系数，避免与本书差异序列 D 混用。

取 ℓ 为一微米，τ 为一毫秒，经过一秒共走一千步，平均位移仍为零，均方根位移却约三十一点六二微米。只跟踪平均位置会看见静止，只看单条轨迹又可能看见偏向某侧；分布宽度把两种看似矛盾的记录协调起来。这里的随机模型有精确预言，并不是对任何不规则曲线都作解释。

若相邻步之间存在持续方向关联，交叉项就不再消失；若等待时间分布具有不同性质，扩散标度也可能改变。所以“走势像乱走”不够。应分别检验增量分布、相关性与时间标度，再判断简单模型适用到何处。把每个偏离都称为偶然，会使随机理论丧失与确定理论同样重要的可败性。

在适当的连续极限中，随机过程可以联系到确定的概率演化偏微分方程。方程确定地演化的是分布，不是预先指定每个实现的曲折路径。使用随机微分方程时，还必须说明噪声模型和积分约定；不能因为英文缩写偶尔同样写作 SDE，就把它与本书的显露—差异序列—特征纠缠三维体系当作同一理论。

更需区分物理随机过程与算法随机过程。模拟者可以随机抽样一个确定积分，而不主张被积分对象本身按这些抽样步骤运动。物理中的时间与计算机中的迭代次数只有经过专门论证才能对应；把算法快慢直接解释成材料弛豫快慢，是计算物理中特别容易产生的错位。

57.6 拒绝一次移动为什么也算一份数据

Metropolis 及合作者一九五三年的论文建立了按正则权重采样构型的计算方法，明确保留被拒绝提议后的原构型进入平均。[57-3] 这一看似技术性的规定，实际上决定计算的概率对象是否正确。算法不是产生一串新奇图像，而是要让不同状态在记录中获得符合目标分布的权重。

用本章的两状态例子说明。甲的能量为零，乙比甲高 Δ，令 βΔ=ln2。目标正则概率应为甲三分之二、乙三分之一。每一步都提出切换到另一状态；从甲到乙的接受概率为一半，从乙到甲则总接受。于是平衡流量为(2/3)×(1/2)=(1/3)×1，两向相等，目标分布与转移规则相容。

若只保存接受的移动，记录会变成甲乙甲乙交替，错误地给两者各一半。问题并不是随机数不够多，而是统计对象被悄悄更换：原来要统计每次尝试后的状态，现在却统计接受事件。一个极小模型就能揭示代码中的概念错误，远比让庞大模型多运行几天更有辨别力。

这个例子不意味着任何满足两向平衡的程序都能迅速达到目标。若状态空间被实际不可跨越的障碍分开，或者跨越时间远超预算，有限记录仍可能偏在局部。应比较不同初态、检查相关时间，并选择已知结果或可独立计算的量作对照。随机方法把高维求和转成采样任务，也把误差责任转成与访问频率和相关性有关的新形式。

57.7 计算结果何时成为物理证据

至少要区分两层问题。第一，程序是否正确求解了所写模型？这可以通过解析小例、守恒检查、网格研究、不同实现与统计核验来考察。第二，这个模型是否适合所研究的样品和操作？这需要实验、独立参数以及能够区分竞争解释的观测。第一层成功不能自动回答第二层，第二层的一次拟合也不能掩盖第一层的错误。

以非局域核为例，程序精确复现二次和四次多项式，是算法核验；在真实材料上预言不同长度尺度的响应，才进入物理检验。如果每改变样品尺寸就任意重调核，计算仍可能漂亮，却没有证明一种可迁移结构已经成立。类似地，有限元网格加密后结果稳定，也只能说明所选方程和边界被稳定地求解。

本书可沿 D—E—S 倒逼路径理解这项变化：原来的求解操作 D 无法处理不连续或复杂边界，促使重新组织函数空间、邻域核或连接结构 E，再形成可比较的响应 S。也可沿 S—D—E 再创，从算法间不一致的显露出发，组织小模型、尺度变化与实验对照，判断应修正代码还是更换物理对象。两者都不能跳过回检。

三方程在此分别约束输出、程序和模型的相互成立。S=F(D,E)要求报告结果如何由算法与物理假设产生；D=G(S,E)要求检验步骤针对尚存的歧义；E=H(S,D)要求新增结构回应独立差异。若输出与可靠操作冲突，推动模型 E 改变；若输出与模型冲突，推动操作 D 改变；若操作和模型只能稳定给出分布而非轨迹，推动显露 S 改以分布指标组织。三种张力的处理都须再回到原问题。

强竞争解释仍然值得保留：这些只是旧物理的新求解技术。对于许多计算，这已经是准确评价。只有当新组织使原来无法稳定规定的对象、操作和预测取得共同身份时，本书才进一步谈物理世界的扩展。非局域、随机和离散不是自动加分的标签；它们必须说明自己究竟保存了什么，增加了什么，又愿意在哪里被证明不适用。

本章参考资料

[57-1] S. A. Silling，Reformulation of Elasticity Theory for Discontinuities and Long-Range Forces，Journal of the Mechanics and Physics of Solids 48（2000），175—209。阅读作者机构所载摘要与研究范围：https://www.sandia.gov/research/publications/details/reformulation-of-elasticity-theory-for-discontinuities-and-long-range-force-2000-01-01/。相关实现分类参见 INL MOOSE 官方文档：https://mooseframework.inl.gov/docs/site/modules/peridynamics/PeridynamicModels.html。

[57-2] R. Courant，Variational Methods for the Solution of Problems of Equilibrium and Vibrations，Bulletin of the American Mathematical Society 49（1943），1—23。阅读变分近似及原页 15—17 的分片线性网格、积分核与随机行走联系。原论文扫描：https://home.agh.edu.pl/~milenin/programs/doc/Courant_1943.pdf。

[57-3] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller，Equation of State Calculations by Fast Computing Machines，Journal of Chemical Physics 21（1953），1087—1092。阅读原页 1088—1089 采样规则与拒绝步计数；doi:10.1063/1.1699114。原论文扫描：https://www.huy.dev/Metropolis-JChemPhys-1953.pdf。本文核、多项式、帽形函数、容器、随机行走及两状态算例均为独立展开。

第五十八章 科学典范怎样碰撞
库恩、拉卡托斯与 SDE 对话

58.1 改变答案，还是改变问题的资格

在一间研究室里，两组人可以对同一张图说出完全不同的话。一组追问曲线能否由更精确的参数拟合，另一组追问画这条曲线时遗漏了哪些状态。前者希望把答案做得更准，后者怀疑问题本身是否切分得恰当。两组人未必一组保守、一组先进，也未必一组只信数据、另一组只爱哲学。真正需要说明的是：他们保留什么，容许改变什么，用什么判断改变是否成功。

前面各章已经多次遇到这种局面。相空间改变了状态的组织方式，统计分布改变了宏观稳定的解释入口，测量理论迫使装置及记录进入问题规定。我们不能把这些变化都缩成“旧理论错了、新理论对了”，也不能因为知识有历史性，就宣布任何解释都一样。科学思想史需要一种能同时处理稳定、变革和检验责任的语言。

本章把库恩、拉卡托斯与 SDE 放在同一张讨论桌旁，不安排一场由最后出场者必胜的辩论。三个方向的关切并不完全一致：共同体如何维持研究，纲领怎样在时间中接受评价，新对象如何由差异序列与特征纠缠取得显露。它们可以互相追问，但不能仅靠更换术语便宣布已被统一。

58.2 先把库恩从口号中请回来

库恩在《科学革命的结构》后记中区分学科矩阵与示范性解题实例，讨论共同体共享的符号、模型、价值和训练。他也明确反对把不可通约理解为完全不能交流，或理论选择没有好理由；共同价值并不保证每个人按同一规则立即作出同一选择。[58-1]

这个限定对本书很重要。若把库恩写成“科学家全凭信仰换阵营”，SDE 即使提出操作、计算和检验，也只是打倒一个自己制造的对手。相反，若承认共同体成员拥有理由，却会在理由的权重、适用场合与对象分类上分歧，研究才真正开始。此时要追问的不是有没有理性，而是哪一种比较关系尚未建立。

例如，两组人都要求准确，却分别统计不同量的误差。一组计算逐点偏差，另一组优先检查整体守恒。前者可以在每个采样时刻拟合得漂亮，却在长期积分中不断创造能量；后者保住能量，也可能把相位预测错开。两者并非无法对话，但把一个分数称为“准确性”还不足以解决争论。需要拆开任务，列出各自承诺保留的关系，并安排能够让两种失败同时可见的比较。

从这个自拟例子看，典范冲突不必始于互不相识的语言。相同词语、相同仪器甚至相同公式，也可能承载不同任务。历史分析若只数理论中出现了多少新名词，便会漏掉这些变化；若只问大家是否仍说“能量”“状态”，又会把对象规定的改变误作语言连续。

58.3 拉卡托斯要求我们追踪一段路

拉卡托斯在《科学与伪科学》中主张，把研究纲领而非孤立命题作为评价单位，区分相对稳定的核心、可修订的辅助条件及解决问题的方法。他以能否产生并获得支持的新事实区别进步与退化，同时允许研究者暂时坚持处于困难中的纲领，前提是公开记录其处境。[58-2]

这里最值得带入本书的是时间尺度。一个新模型第一次失败，可能说明模型没有前途，也可能说明关键仪器尚未校准；一个老模型连续成功，可能仍有结构性盲区，也可能确实是目前最好的工作工具。评价不能只截取研究者最得意的一刻，更不能把多年修改后的成果倒填到最初提出时。

但拉长时间并不意味着无限延后判断。假设研究组先承诺半年内预测五个未测参数，半年后只解释已经知道的三个，再把“预测”改成“重新理解”；若不保留最初承诺，时间尺度就成了逃避失败的空间。相反，若明确记录哪些预测成功、哪些没有成功，以及为什么改变任务，继续投入便可以接受公开讨论。

我们因此把一个纲领的档案看成有日期的责任链。它至少应能回答：当时已知什么，哪些数据尚未开放，模型在看到结果之前承诺什么，修改以后增加了哪些可检验后果。历史的公平不在于要求前人知道后来的答案，而在于不让后来答案抹去他们真正面对的不确定性。

58.4 第三组合不能替代三个发生维度

《科学发生的结构》第三四二卷把典范辨识为 Function、D、E 的组合，用 S=Function(D,E)表达产出；正文强调 Function 涉及把操作与关联组织成何类结果，而不只是最后一道步骤。该书还提出多种典范模型共存、对话和形成新组织的研究方向。[58-3]

这里有一个必须守住的记号区别。典范的三组合写作(Function,D,E)，SDE 的三个发生维度仍是(S,D,E)。前者试图辨认一套研究方式如何稳定工作，后者分析具体显露、差异序列与特征纠缠怎样共同发生。Function 不等于 S，也不是在 SDE 外发现了第四种存在实体。若把两套三项表直接对齐，便会把产出及其构成方式混为一谈。

本章用一个自拟统计例子说明区别。给同一组四个读数一、二、三、十四，程序按同样顺序完成读取、单位核验和整理。若研究任务以平均值汇总，结果为五；若以中位数汇总，结果为二点五。相同读数可以支持不同结果类型，究竟采用哪一种，取决于要回答总体负荷还是典型水平等问题。这里的五与二点五是不同 S，汇总准则的区别帮助我们识别 Function。

不过，这个例子不能证明 Function 在所有定义下都独立于 D。若把 D 定义为包括每一步聚合运算的完整可执行程序，均值与中位数的差别已经写进程序，Function 便可以被编码在 D 中。因此三组合的价值不是数出三个永远不可互译的箱子，而是使研究者能够在所采用的分析尺度上，分别追踪程序次序、关联结构与结果组织准则。

这也给出一项严格的检验。若评审者在保持任务定义一致的前提下，只使用 D、E 便能同样稳定地预测模型的结果类别与失败方式，而增加 Function 标记不能改善辨识，那么三组合在这个研究范围内没有显示额外分析收益。不能把这种结果改称“Function 其实早已暗中存在”便宣布理论获胜；那只说明定义可以重写，尚未证明新增区分值得保留。

58.5 真正碰撞必须留下可检查的改变

两位专家在会议上轮流发言，不等于典范已经碰撞。十个模型重复同一个结论，也不等于发生了多元研究。碰撞至少需要一处可定位的不相容：输入怎样取得，变量怎样划分，哪些关系被保留，什么形式的结果才被认作答案。若这些都没有被触动，所谓对话可能只是增加说明文字。

用运动重建作一个小模型。模型甲以局部速度差估计阻力，模型乙以全程能量收支估计耗散。两者处理同一台装置，甲在低速段预测良好，乙能保持总能量账目，却不能定位耗散发生的时段。若新模型只是把两个分数平均，冲突被遮住而非解决。更有意义的改变，是引入可观测的内部温度或接触状态，使局部阻力与总体耗散通过新的变量关系相接。

但新增变量也可能只是多给了拟合自由度。假如内部温度从未独立测量，只按误差需要随意调节，那么它很容易解释任何结果。碰撞产生的新列，必须有测量资格、定义范围和跨任务约束。把“变量表长出一栏”当成成功标志，只能是研究提示，不能成为自动颁奖规则。

可以据此建立最低限度的前后比较。保留甲乙原有预测，再记录新模型用到了哪些额外信息。若新模型在相同信息预算下改善未见条件的预测，并且新变量能被独立测

定，其发生内容才较清楚。若它只是动用了更多事后数据，或在旧测试上更好、在新测试上更差，就不能把改进归给典范碰撞本身。

58.6 一次可复算的碰撞评估

设有十项预先固定的测试，每项只记预测是否落入约定误差内。甲成功七项，乙成功六项，其中两者同时成功五项。于是至少一者成功的任务数为七加六减五，等于八。这个八说明两者的成功范围不完全相同，却不意味着一个真实调度系统已经会在每次测试前挑对模型。

如果在知道答案以后总选成功者，八项只是事后可达的上界。要实现它，需要提前可用的选择线索，例如温度、尺度或边界状态。设调度器实际只成功七项，就不能宣传为八项；若新组合成功九项，也不能只凭这个小样本就断言已经开创典范。还需重复、说明样本难度，并与增加同等算力的单模型进行比较。

这个简单算例使“多元优于单一”的命题具有可争论内容。多元系统可能因覆盖互补而改善，也可能因路由错误、接口损失和额外成本而变差。一个大模型内部也可能保存多套方法，一个由多个模型构成的系统反而可能高度同质。因此，物理上的模型数量不能直接替代认识上的方法差异。

若要测量差异，可要求各模型对同一失败说明其假设、替代变量和可区分实验，再由独立评审检查是否真有不同。不能只统计措辞的多样性。十种写法可以表达同一条推理；一句相似的话也可能隐藏不同的约束。真正需要保存的是解决问题方式的差异，以及这种差异在新任务中的实际效力。

58.7 SDE 怎样进入，而不把一切都收归自己

在上述模型竞争中，可沿 S—D—E 再创路径展开：先发现两个模型对同一测试产生稳定分歧 S，安排改变条件的试验 D，再识别造成分歧的特征纠缠 E，例如接触状态与耗散率的联系。识别以后还须返回预测，检查新联系能否解释原失配并承担新的检验。

另一条 E—D—S 沉淀路径可以从明确的关联猜想开始：若猜测温度改变局部阻力，就先规定这种联系 E，再设计控温与独立测速序列 D，最后看是否出现事前承诺的差异 S。这两条路径的区别是研究从哪里进入，不是两套互不相容的科学逻辑。它们也不保证实验一定成功。

S=F(D,E)要求每个结果说明其生成条件；D=G(S,E)要求后续操作响应已知结果与关系，而不是机械重复；E=H(S,D)要求构成判断能够被操作与显露修订。三个式子提供组织研究的约束，不能直接推出任何具体阻力系数，更不能替代十项测试的真实记录。

三原理也应落到这张记录表上。若 S 与 D 的重复关系不稳定，推动检查 E 中被当作恒定的接触状态，再回检；若 S 与 E 的预测不符，推动改变 D 以区分温度、速度和仪器漂移，再回检；若 D 与 E 的组织不足以支持旧 S，推动把“总耗散”分成可区分的时段响应，再回检。每一次第三维变化都增加责任，而不是增添一层免受反驳的语言。

58.8 哪些反例会使历史方向论失败

本书若主张科学发展的方向总是让旧典范成为新典范的极限特例，就必须面对不被包含的情况。一个新理论可能保留旧计算在有限范围内的数值近似，却拒绝旧理论的实体身份与因果解释。数值可恢复、对象可互译和本体被包含，是三个不同命题，不能看到第一项便宣布第三项成立。

由此可以提出一项足以击败强方向论的反例标准：若严谨历史研究反复发现，最具持续预测力的新体系只能靠放弃旧体系的核心对象关系而成立，且无法在预先说明的映射下把旧关系恢复为受控极限，而这些体系仍保持公共检验能力，那么“必然包含而非替代”的普遍命题就失败。失败后应缩为某些领域常见的机制，不能把放弃重新命名为更深包含。

另一个反例针对多典范模型的制度预测。若未来在同等资源、任务难度和评估尺度下，广泛使用统一模型的研究群体持续产生更多相互独立的方法、对象与可验证成果，而专门保持模型分离的群体并无优势，那么“统一必然压低多样性”的强主张就不能成立。这与第三四二卷主动列出的方向性败局相呼应。[58-4]

还有一个更贴近 SDE 本身的失败条件。若盲评研究中，SDE 编码人员对同一案例的三维归属长期不能取得可复核的一致，分析又不能比更简单的任务—假设—证据描述多预测任何失败位置，那么 SDE 作为历史研究工具的操作价值便受到实质限制。不能以“发生过于复杂”解释所有不一致，又继续要求他人接受其方向判断。

这些条件不是为了让一本书显得谦虚。它们决定该书是否允许未来材料改变自己的组织。能够批判前人的理论，却不允许自身在相同层级失败，不是更深的发生学，而是把发生停在了自己这里。

58.9 历史方向与研究选择应当分别论证

即使多典范碰撞经常产生新方法，也不能由此直接推出每个实验室都应同时维持所有路线。资源有限、研究成熟度不同，某些任务需要长期稳定的校准，某些任务更需要探索新对象。多样性的价值必须与维持成本、交流成本及重复验证能力一起考察。

反过来，一项研究暂时只能使用统一平台，也不意味着其研究方向必然贫乏。研究者可以在同一平台内保存不同假设、隔离训练材料、使用不同变量划分，并请互不依赖的团队复核。是否形成有效差异，需要检查过程与产出，不能只看采购合同上有几个模型名字。

本章因此保留一个强竞争解释：许多所谓典范碰撞收益，也许来自一般的集成、分工与误差互补，无须 SDE 本体论。对它的回应必须具体。如果 SDE 能提前定位传统集成遗漏的对象不相容，指导建立新变量，并使新测试取得额外成功，它便有贡献；如果没有这些差别，应承认收益由更普通的机制解释已经足够。

这种对话并不妨碍本书继续坚持发生取向。它只是要求把哲学主张、研究方法与实验成绩分层。对象如何取得规定，是值得追问的问题；哪种编码能帮助研究者，是方法比较；某个装置预测更准确，是经验结果。三者可以相互支持，却不能用一个层次的成功替另两个层次提前盖章。

58.10 把最后的裁判权留给下一次工作

库恩提醒我们看见共同训练与评价条件，拉卡托斯要求在有时间次序的研究轨迹中比较成果，SDE 进一步要求追踪显露、操作序列与特征纠缠如何互相生成。把这三种关切放在一起，能够迫使一本物理思想史对自己的叙事方式负责。

科学史不只是胜者如何终于说出正确公式的故事，也不是每个共同体都困在自己的语言里。人们能够建立局部翻译、改变装置、保留失败，逐步创造原先尚不存在的共同判断条件。这样的过程可能形成包含，也可能伴随舍弃；可能通过碰撞，也可能通过一条路线内部长期、艰难的展开。

真正需要留下的不是一幅保证未来照着走的路线图，而是一套允许路线图被改写的研究纪律。若碰撞有创造力，让它在新的问题上显示；若它没有，让失败进入记录。只有这样，关于科学发生的理论才仍然属于它所描述的开放历史。

本章参考资料

[58-1] Thomas S. Kuhn, The Structure of Scientific Revolutions, 2nd ed., 1970，Postscript—1969，学科矩阵、示范例、不可通约及理论选择段，原印页 182—187、198—200。https://www.lri.fr/~mbl/Stanford/CS477/papers/Kuhn-SSR-2ndEd.pdf。

[58-2] Imre Lakatos，“Science and Pseudoscience”, 1973，演讲转录及后续文本增补，研究纲领、进步与退化、公开记录部分；转载原 LSE 转录。https://studylib.net/doc/27074072/bf3-lakatos-1973-science-and-pseudoscience。

[58-3] 王德生：《科学发生的结构——超越库恩》，德麦国际专著第 342 号，2026，第 18—20 章相关原文。https://sdeuniverses.com/books/m/342/text/。本章从其定义出发的数值例、分析尺度检验和竞争设计为本书展开。

[58-4] 同书主页“保底与死法”，方向性反例声明。https://sdeuniverses.com/books/m/342/。

第五十九章 十大物理学家的十种创造同一张比较图

59.1 把十个人放在同一张图上，不能先删掉他们的差别

全书已经沿不同问题进入物理学的创造现场。现在需要一次横向比较：牛顿、爱因斯坦、麦克斯韦、伽利略、法拉第、普朗克、玻尔、海森堡、薛定谔与狄拉克，究竟分别改变了物理世界的哪一项构成？这份名单是本书选择的比较范围，不是按名望排出的十个最高分，也不意味着名单以外的人物只提供背景。

比较有两个容易滑落的方向。一种把每个人压成一句口号：牛顿是力，爱因斯坦是相对，麦克斯韦是场。这样的表格容易记，却不能说明新对象怎样成立。另一种把前面各章缩短后依次复制，读者得到十篇小传，却仍不知道它们能够在哪一种尺度上互相比较。

本章采用同一组问题：什么对象的身份被重新规定，哪些原来可以分开的量被迫共同成立，新增关系如何进入检验，以及哪些成功不能被无限扩张。这组问题不预设所有创造都发生在同一层，也不要求每位作者独自完成发现、形式化、测量与验证。它把姓名当作进入研究网络的入口，而不是把整个网络压缩成一个人的头脑。

下表的每一行都指向前面已经展开的论证。这里的“创造类型”是本书的比较性判断，作者自己的历史用语与当时物理主张，仍以各章原典范围为准。它们不是十位作者曾经共同签署的一份方法宣言。

人物	重新规定的对象	关键转换与检验关系	必须保留的界限
牛顿	受共同运动规律约束的力学物体	质量、运动量与作用联合；地面及天体关系相接	定义不能替代外力、初值及适用条件
爱因斯坦	由时空比较规定的事件与物体	同步、测长和变换相容；另以光量子重问交换	参考系变化不等于人的意识改变自然
麦克斯韦	具有约束和演化的电磁场	电荷连续性、位移项、传播与光学关系联合	逻辑相容不能独自选出全部物理规律

人物	重新规定的对象	关键转换与检验关系	必须保留的界限
伽利略	可按时间关系比较的运动	距离、时间与理想化过程形成可复查比例	理想落体不等于任何介质中的实际运动
法拉第	由变化、几何与回路共同规定的感应	通断、反向、移动与零响应构成实验序列	后来的场方程不是原实验的逐字内容
普朗克	可按能量份额计数的分配	能量元素、计数与辐射谱共同受约束	份额计数不直接等于自由光子理论
玻尔	定态及其能量差构成的原子	稳定性、谱线组合与量子条件相接	圆轨道模型没有完整给出跃迁动力学
海森堡	以跃迁关系组织的动力学量	双指标、乘法次序与谱学约束共同改变	非交换不是对一切经典知识的否定
薛定谔	由波函数及算符关系规定的状态	本征问题、边界与演化形成统一求解	波函数不能无条件当作三维物质波
狄拉克	相对论量子结构中的电子与反粒子接口	线性化、旋量、磁耦合与新粒子检验相接	方程解、历史解释和实验发现不能合并

59.2 第一条比较轴：对象是被压缩，还是被打开

牛顿力学的突出力量之一，是把对象的大量日常特征放到研究边界之外，使质量、位置、运动量与作用关系获得稳定的计算身份。这样的压缩并不等于对象从此只剩三个数字。颜色、温度、形变与内部结构可以在另一问题中重新进入；当它们影响当前运动时，压缩就需要重审。创造发生在一种可迁移而有边界的封装中。

麦克斯韦的场则打开了另一种内部结构。一个回路间的整体作用参数，不再足以承担全部问题；空间分布、初始约束和传播必须进入对象。与牛顿比较，不能简单说一个把物拆散，另一个把物联结。力学的质量与作用同样必须联结，场论也同样要作理想化省略。真正差别在于：哪些自由度被保留，哪些关系成为后续变化不可绕开的条件。

普朗克与玻尔又使状态的允许范围成为中心。不是把已有运动多画几条线，而是改变哪些能量分配或定态被准许进入计算。海森堡与薛定谔进一步改变状态怎样承载变化：一条经典轨道不再自动是完整答案。到了狄拉克，状态分量与变换要求之间的相容性又迫使电子的规定扩展，原来的单粒子图像留下新的解释问题。

这些变化不能排成“对象越来越少”或“属性越来越多”的单调曲线。有时通过省略获得新的稳定性，有时通过恢复内部结构获得新解释，有时需要改变允许状态，有时需要改变表示和组合。共同标准不是复杂程度，而是新组织能否承担旧组织不能同时承担的约束，并继续在新的条件下受检验。

59.3 第二条比较轴：什么差别能够成为研究的发动点

伽利略与法拉第特别提醒我们，创造不能只从最终方程回溯。可比较的时间序列、可控制的装置改变以及对零结果的辨认，先决定哪些差异能够进入研究。没有稳定的操作差别，再漂亮的拟合也可能只是把偶然记录包装成定律。可是把两人分别称为运动实验家、电磁实验家，仍不足以解释他们怎样改变了对象。

对于伽利略式运动研究，操作的价值在于把快慢印象推进到距离与时间的组织；对于法拉第的感应研究，关键差别则在通断与保持、靠近与离开、同向与反向之间。前者使变化成为连续量度的对象，后者使短暂响应与共同几何获得新的联系。它们并不是“多做实验”这个抽象行动的两个例子，而是不同差异结构的创造。

法拉第第一系列报告中，通电持续期间与接通、断开瞬间呈现不同结果，两个瞬间的方向也不同。[59-1] 本章据此提出一项比较：零记录不总是等待被改良仪器消除的失败。它可能与非零记录共同限制一种规律。若只选取最显眼的偏转，便会把变化条件降格为故事背景，遗漏对象为何成为感应而非持续传导的问题。

爱因斯坦的同步分析则显示，差异运行并不只等于在实验台上调动设备。先澄清异地事件怎样被共同计时，能够揭示原来混在同一个“长度”词里的不同操作。[59-2] 这样的分析既不是凭空选择定义，也不是读数自己说话；它以物理信号、惯性系假设与变换要求共同限制比较程序。方法上的创造可以发生在实际操作，也可以发生在对操作资格的重构。

59.4 第三条比较轴：允许把什么同什么组合

最不显眼却最深刻的改变，往往发生在组合规则中。普朗克使总能量的分配计数受到有限份额的规定；玻尔使发光频率进入两个定态之间的能量差；海森堡则重新追问，

用来替代经典量的对象应当怎样相乘。三个问题都超出“再测一个数”，却不能用同一个量子跳跃故事吞并。[59-3—5]

普朗克的计数要求先确定什么算不同分配，再讨论概率与谱形；玻尔的能量差要求多个谱线能够归属于一组共同状态；海森堡的双指标乘法要求经过中间状态的组合遵守新的规律。如果只把它们概括为不连续，就会漏掉它们各自建立的联系。离散数值可以出现在许多经典系统中，量子创造的内容不能由离散二字独自承担。

麦克斯韦也改变组合，但不是同一种组合。他把源、场约束与演化放在一起，使局部连续性和空间传播互相制约。原论文将电磁扰动速度与光学量度相比较，给统一增加了跨实验领域的负担。[59-6] 这与把谱线纳入能级网络有可比之处：原来可被分别调整的结果，开始受到同一结构的联合限制。

这种可比性不能升级成“所有理论都是同一方程”。电磁场是带有空间和时间结构的物理量，跃迁矩阵是另一种运算组织，热辐射计数又有自己的样本空间。SDE 的比较任务是指出重新组合发生在哪一层，并说明组合如何得到回检，而不是把这些不同对象全部溶解成没有量纲、没有定义的关系。

59.5 一个二状态演算：形式改变的创造边界

为了让比较不只停在语言，我们设置一个两维状态空间的小演算。它不是十位人物中的任何一位做过的历史实验，而是本书检验“新表示是否已经等于新物理”的数学工具。取归一化状态 $\psi=(1,0)$，可观测量 Z 为对角矩阵，其两个对角元分别为一与负一。于是 ψ 的 Z 期望为一，测得负一的概率为零。

现在换一组正交基，变换矩阵 U 的第一行为 $(1,1)/\sqrt{2}$，第二行为 $(1,-1)/\sqrt{2}$。它满足 U 的转置乘 U 等于单位矩阵。新的状态坐标为 $\psi'=U\psi=(1,1)/\sqrt{2}$；同一个可观测量也必须变为 $Z'=U\,Z\,U^{\mathsf{T}}$，其中 U^{T} 为 U 的转置；直接相乘可得 Z' 的对角元为零、两个非对角元为一。

再计算期望，Z' 作用在 ψ' 上仍得到 ψ'，所以期望依旧为一，方差仍为零。坐标数值和算符矩阵都变了，可检验的判断保持。若只改变 ψ 而坚持保留旧 Z 矩阵，便会算出期望零；这不是同一过程因表示改变出现新结果，而是没有同时转换全部相关结构，实际询问了另一个可观测量。

这个小演算不能证明所有波动力学与矩阵力学之间的等价细节，却清楚标示了评价原则：在可逆、相容的表示变换中，数学外观变化本身不增加实验事实。海森堡与薛定

谔的历史创造，不能仅靠他们使用不同符号来证明；需要看各自怎样打开新的求解、对象组织与预测能力，并怎样被后来的统一结构承接。

也不能反过来说，既然有等价关系，所有形式创造就只是装饰。适合本征问题的表达、适合变换与复合运算的表达，可以让不同问题变得可计算。把实验、计算和状态规定稳定地接起来，可能产生新的研究能力。只是研究能力的增加，应与关于自然界新增独立事实的宣告分开，两者分别承担证据。

狄拉克提供了更进一步的比较。他要使相对论与量子演化共同成立，不仅是给旧状态换一组基。其一九二八年原文已经明确把相容性问题和未解决的正负能量解释问题分开。[59-7] 新的分量结构、磁耦合以及后来反粒子解释不能由上述两维换基例自动推出。比较有价值，正因为它能够同时识别同类任务与任务层级的改变。

59.6 第四条比较轴：旧成功怎样被新世界保留

一项创造若只能宣布旧理论全部错误，却无法解释它为什么长期有效，就会丢掉自己的重要约束。牛顿的力学世界进入相对论时，低速条件下的许多数量关系仍须恢复；经典电磁场进入量子理论时，在适当状态与尺度下的波传播仍有有效位置；旧量子模型的若干谱线结果可以被更统一的理论重新取得。

保留并不意味着原来的图像全部保留。相同的能级结果可以在不依赖经典圆轨道的描述中出现，相同的运动近似也可以被放到不同时间结构中解释。究竟保留了数值、关系、操作还是实体图像，需要逐项区分。若把它们打包，继承要么变成所有旧话都对，要么变成所有旧话都不曾真正成立。

同样，新理论的极限也不是一句“当参数足够小时当然回来”。需要指出控制参数、研究量和误差目标。低速并不保证任何高精度实验都可忽略相对论修正；大尺度也不保证一切量子性质均不可见。理论桥梁是一套有条件的转换，不是一张允许随意跨越领域的通行证。

这个尺度让十种创造能够共同比较而不互相压平：伽利略的理想化、牛顿的动力学封装、法拉第的实验差异、麦克斯韦的场组织、爱因斯坦的比较重构以及量子诸方案的状态改变，都要说明自己如何承接已经可靠的结果。创造不是越少继承越激进，而是既能保留必要联系，又敢于改变无法继续维持的规定。

59.7 三个扩展节点怎样使比较图免于英雄谱

马赫、诺特和玻尔兹曼在本书中不是可以随意替换十人名单的候选名字，而是三种必要的扩展节点。马赫使惯性与参考结构的前提重新成为问题，提醒我们不要把操作稳定性误作无条件的独立属性。他的思想批判并不等于已经建立一套获证的宇宙惯性定律，这一限制必须与启发同时保留。

诺特让变换与守恒之间的关系取得定理形式。她的作用不是在已有守恒量后面补一条哲学说明，而是使人们能够从作用量和对称条件追问哪些守恒关系应当存在。可是定理的适用条件、边界项以及局域规范情形仍需分辨。把一切不变都叫诺特守恒，会使这个扩展节点失去精确性。

玻尔兹曼则把宏观稳定与微观实现的组织推进到中心。统计规律不能仅靠无知两个字解释，也不能从可逆微观运动不加条件地推出一切宏观时间方向。这个节点迫使比较图加入分布、典型性、制备与时间尺度，使十个人的工作不再只是少数基本方程在真空中轮流登场。

三人还暴露了单一姓名叙述的局限。矩阵力学需要玻恩、约尔当等人的系统化，波函数的统计解释不能全归于薛定谔，电子及核的规定依赖汤姆孙、卢瑟福与实验共同体，反粒子身份也要经过实验判别。承认这些关系不是削弱十位人物，而是把他们真正能够改变研究的接口显示出来。

因此，比较图中的一行不是一个封闭贡献账户。它是一组特别有解释力的入口，向前连着问题来源，向侧面连着竞争方案，向后连着进一步的修正与应用。人物比较只有保持这些开放连接，才可能成为思想史分析，而不是用哲学术语重新装饰的名人名录。

59.8 十种创造怎样接受同一套发生学追问

现在可以回到 SDE。不能给伽利略和法拉第只分配 D，给牛顿和麦克斯韦只分配 E，给量子理论只分配 S。每一次创造都包含完整的显露、差异序列与特征纠缠，只是问题从不同入口显现。人物的研究风格不等于三维度的分工表，主体、互动和客体号位也不能与这些维度一一配对。

例如沿 D—E—S 倒逼路径比较：法拉第的操作变化迫使解释短暂感应，海森堡对经典量的重新计算迫使改变组合规则；两者都可从差异操作进入，但前者主要在装置、过程与响应之间展开，后者还在代数对象的构造中展开。路径相同不表示内容相同，更不表示历史发现实际按一个预先画好的顺序进行。

再沿 E—D—S 沉淀路径比较：麦克斯韦的联合场关系与玻尔的定态关系，都要求推出尚未被用于定参的后果，再安排相应比较。可是它们的证据范围不同，完成程度不同，不能因为都走过一次回检就获得相同理论地位。六路径帮助询问研究往返，不是给所有故事盖同一种成功印章。

三方程要求每一行都能说明 S 怎样由 D 与 E 成立，D 怎样因已有 S 与候选 E 重排，E 怎样在 S 与 D 的共同限制下改造。三原理则进一步追问，究竟哪两维不能协调，才使第三维必须改变。例如旧操作和稳定显露共同压迫原子结构重建；候选场关系与已知记录共同要求新的空间探测；可执行操作与形式结构共同促使改用谱线、分布或关联作为显露指标。改动之后必须回到原问题，不得只靠新词消除张力。

一种最强的竞争解释是：所有案例都可以由成熟的实验方法、数学重构与理论检验来说明，未必需要额外本体论。对此，本书必须承认既有说明的力量。属性纠缠论的增量应体现为它更准确指出哪些原以为独立的属性实际共同成立、哪些对象身份因关系改造而改变，以及为什么某些稳定不能由单个属性承担。若只能把成功案例翻译成三字母，它就没有完成这项任务。

59.9 同一张图所保留的不是同一种答案

这张比较图最终留下的，不是十位人物共同证明了一套后来的哲学，而是十种可以互相校正的创造经验。有的以可比较的运动打开局部规律，有的以跨装置联系重建作用，有的以共同演化使场取得身份，有的以状态、变换和组合改变物理判断的基本单位。它们的统一性在于承担同一类责任，差异则在于各自改变了哪些实际规定。

一个新世界的判断，因而不应只看新名词多少、方程是否陌生或图像是否违背直觉。应看新的规定能否使本来不能共同成立的证据形成更严密的联系，能否保留必要旧成功，能否产生未被事后调参吞掉的新预期，以及是否清楚承认仍然悬而未决的部分。

这样的比较也为下一章留下一个开放要求。今天出现的新计算、新材料或新实验，不能因为尚未拥有历史名人的声望就被排除在创造之外，也不能因为声称颠覆便自动获得同等地位。十人的价值在于提供可反复使用的研究尺度：重新打开对象，让关系接受改造，再让改造承担能够失败的回检。

本章参考资料

[59-1] Michael Faraday，Experimental Researches in Electricity，Vol. I，First Series，§5、10—17。https://www.gutenberg.org/files/14986/14986-h/14986-h.htm。

[59-2] Albert Einstein，On the Electrodynamics of Moving Bodies，1905，§§1—2 英译。https://fourmilab.ch/etexts/einstein/specrel/www/。

[59-3] Max Planck，On the Law of Distribution of Energy in the Normal Spectrum，1901，开篇与§§1—4 英译。https://strangepaths.com/files/planck1901.pdf。

[59-4] Niels Bohr，On the Constitution of Atoms and Molecules，Part I，1913，定态、辐射与频率关系。https://www.informationphilosopher.com/solutions/scientists/bohr/Constitution_of_atoms_I.pdf。

[59-5] Werner Heisenberg，On the Quantum-Theoretical Reinterpretation of Kinematical and Mechanical Relationships，1925，D. H. Delphenich 英译开篇与第一节。https://neo-classical-physics.info/uploads/3/4/3/6/34363841/heisenberg_-_qm_interp_of_kin_and_mech.pdf。

[59-6] James Clerk Maxwell，A Dynamical Theory of the Electromagnetic Field，1865，Part VI，传播与光学速度比较。https://en.wikisource.org/wiki/A_Dynamical_Theory_of_the_Electromagnetic_Field/Part_VI。

[59-7] P. A. M. Dirac，The Quantum Theory of the Electron，1928，导论与§§1—2。https://mathweb.ucsd.edu/~nwallach/Dirac1928.pdf。

[59-8] 本书第 10、16—20、23、27—28、32—37、39、41—46 章及相应参考资料：本章比较矩阵承接这些章节已经限定的历史与数学论证，不以一张表代替各章证据。矩阵行名为比较入口，二状态表示演算为本章自设例。

第六十章 第十一位是谁
AI、科研共同体与未发生的物理

60.1 一个留空的位置

如果一本物理思想史以若干位伟大人物作为路标，读到末尾时很容易追问：下一位是谁？这个问题既有力量，也有危险。它使读者不满足于记住过去，却可能把未来再次想成一位天才独自推开大门，其余人只需等待。所谓“第十一位”，在本章不是一张尚未公布姓名的排行榜，而是一个留空的位置：下一种足以改变物理对象、操作和检验方式的创造，将由什么样的主体完成？

这个主体可能包括提出问题的人、维护装置的人、编写算法的人、整理失败记录的人，也可能包括能够生成候选方案、执行计算与安排实验的 AI 系统。它不必有一张统一的脸，也不会因为成员多了便自动更聪明。真正的问题是：这些成员之间怎样形成能够发现错误、承受反例并持续改造自身的联系。

本书把物理世界理解为在具体复合系统中获得规定，并不意味着未来科学可以只靠重新说话发生。文字中的新对象必须取得操作资格，数学中的新结构必须说明与经验的联系，机器生成的猜想必须经过共同体能够检查的程序。未发生的物理尚未被完成，本章所能交出的，是可进入工作、可留下失败的开放纲领。

60.2 从找到公式到创造对象，还有多远

Udrescu 与 Tegmark 的 AI Feynman 研究将神经网络拟合与对称性、可分性等物理启发结合，用符号回归重建测试函数。论文报告在选定的一百条费曼讲义方程测试中恢复全部表达。[60-1] 这一成绩说明特定搜索问题可以获得强大的自动化帮助；它没有直接表明机器已经独立建立这些方程中的物理量和实验意义。

假设一个系统从时间与距离数据中找到二次关系。它还需要知道距离量自哪里，计时如何启动，装置是否存在固定延迟，数据是否经过筛选，以及二次关系是在什么范围内成立。若这些前提由人准备，贡献就不应全部归给最后输出公式的程序；但也不能因此否认程序在庞大候选空间中发现结构的真实作用。

在这里，功劳与认识论可以分开。谁先想到算法，是贡献分配；公式是否正确，是数学问题；它是否描述装置，是经验问题；它是否建立新物理对象，是构成问题。一项工作可以在某一层很突出，在另一层仍待完成。把所有层次压成一个“AI 发现定律”的标题，反而使我们无法判断它究竟推进了哪一步。

未来的系统也可能参与变量选择与仪器设计，而不限于接受现成数据。即便如此，仍须记录哪些选择来自模型，哪些来自既有数据库，哪些经过人类修改。自动化程度提高，不会使研究失去来源；它只会让来源关系更复杂，更需要明确保存。

60.3 让新颖性接受独立检查

自动实验室把候选生成、实际制作与测量连接起来，已经比纯文字生成走得更远。不过，成功需要经得起重新解释。A-Lab 相关论文在二〇二六年的作者更正中澄清，新颖性原意涉及预测平台所知范围，不能直接等同于科学史上从未出现；衍射重分析确认四十项原报告成功中的三十六项，四项仍不确定，并移除一项误入训练数据的材料。[60-2]

这个案例在本章的价值，不是用更正贬低整个方向，而是使“成功”这个词获得层次。算法认为某个目标可制备、装置产生了粉末、衍射支持某种结构、该结构对数据库而言陌生、它在文献史上确实新颖，是不同判断。各个环节可以相互支持，却不能由前一个自动保证后一个。

若共同体只保留首次新闻，而不把更正接回评价，自动化会放大认识误差；若因为出现更正就宣布自动实验没有价值，又会惩罚正常的纠错。可靠的研究记录应该容纳版本变化，使后来的人知道哪项结论仍成立，哪项被缩小，哪些样品还需独立测量。

在 SDE 语言中，数据库并不是装置外面的一摞参考书。它参与候选如何被认作新，参与哪些结构值得制备，也参与测量图样如何被匹配。因而数据来源与样品身份之间的特征纠缠，属于研究对象成立的条件。纠错不是把一个数字擦掉，而可能重新规定什么算作这项工作的显露。

60.4 一项小任务怎样承担真实的失败

宏大纲领需要从可完成的小任务开始。本章提出一项自拟任务：由 AI 辅助判断一个受控下降装置在给定时段内更符合恒速模型还是从静止开始的恒加速模型。任务不宣称发现新自然定律，而用来检验人机系统能否保持变量、条件与证据的完整性。

先规定训练阶段只开放一秒时的距离零点二米。模型甲取 s=零点二 t，模型乙取 s=零点二 t 平方，两者在这一点完全一致。若系统根据这一条记录便宣布乙得到唯一证明，它已经失败。它应当保留两个候选，并提出能区分它们的后续测量。

在两秒时，甲预测零点四米，乙预测零点八米。若距离误差界预先规定为零点零二米，两个预测区间分别是零点三八至零点四二、零点七八至零点八二米，互不重合。一

次合格的新测量便可能区分这个有限候选集。不过，结果若为零点六米，正确处理不是强行选择较近者，而是检查两种模型是否都不充分。

这项算例还必须防止一个偷换：排除甲不等于在所有可能函数中唯一证明乙。仍有无限多条曲线经过有限数据点。可以继续在三秒处测试，或改变初始条件，借助更广的机制约束减少候选；不能把候选列表很短，误作自然界本来只允许那几种答案。

这样的小任务看似不够壮观，却能检查未来科研系统是否具备关键能力：承认不可辨识，提出区分操作，保持测量单位，区分模型失败与仪器失败，并把新的结果写回假设。缺少这些能力，生成再多复杂方程也不能补足。

60.5 审计应当保存什么

可审计不等于把所有中间文字无限堆积。真正需要保存的是能影响结果的选择：数据版本、单位换算、训练与检验边界、模型代码、随机种子、参数范围、人工干预、失败条件及修改日期。对于实验，还应保留样品身份、仪器校准和处理链，使外部人员能够重建关键判断。

这里不要求研究者公开无法解释的内部心理活动，也不把机器生成的每段自述当作真实推理记录。一个语言模型可以说自己“先比较再推导”，但这句自述不等于可验证的计算轨迹。更可靠的记录是它实际调用了什么数据，提交了什么表达，哪些程序产生了什么结果，操作者在哪一步改变了方案。

应把发现用的数据与评价用的数据分开。若系统反复看到同一测试集并据此修改，那个测试集便逐渐成为训练的一部分。继续称其为独立检验，会夸大可靠性。对于重大主张，可以由另一团队保存最终测试条件，在冻结模型后才开放结果。

审计还要记录没有发生的事。某个候选因为设备能力不足未能试验，不能记作实验否定；某项测试因校准失效被取消，不能记作模型成功；没有发表的负结果，也不应自动从效率分母中消失。一个人机系统每提出十个候选才成功一个，与只公布成功样品的故事，可能看起来完全不同。

60.6 共同体不能只在最后签字

把人类安排成机器论文的最后签字者，不能构成充分的科研共同体。人的责任需要在问题选择、可接受风险、资源分配、证据裁决与错误处理处有实际入口。否则“人在回路中”只是一句形式保证，签字者并没有足够信息和能力改变结果。

同样，研究者也不能把 AI 当作永远不得反驳的廉价助手。机器提出的候选可能挑战团队习惯，暴露人为选择的数据偏差。共同体若只接受符合先验期待的输出，自动化便会把旧有封闭放大。保留人类裁定，不等于让人类直觉免于检验；它意味着由可追责的主体维护公开的判断程序。

对于合作成果，应说明贡献类型。提出问题、设计实验、维护硬件、开发软件、完成验证、解释结果和撰写文本，不是同一件事。记录这些工作有助于避免把整个复合过程归于一个系统名字，也有助于防止不可见的技术劳动被抹去。未来科研主体可以扩大，责任却不应因此稀释。

强竞争解释认为，AI 科研仍只是传统研究的工具升级，无须诉诸新的主体结构。这个解释完全值得保留。只有当系统持续参与任务重定、变量生成与实验选择，人的判断也因其反馈发生可追踪改变时，复合主体分析才显示额外内容。若机器只是批量执行预定计算，本书就应准确称其为自动化工具，而不强行提升概念等级。

60.7 防止理解感跑在证据前面

Messeri 与 Crockett 的研究分析了 AI 可能带来的理解幻觉及方法单一化风险，提醒人们生产更多结果不必然意味着理解更多。[60-3] 这是一项应被认真检验的风险判断，不是任何使用 AI 的实验室已经丧失创新的定论。

理解感之所以危险，在于它常由表达质量触发。一段文字把量子、几何和信息连得很顺，读者便觉得三者已经统一；一幅精美图展示一条新轨迹，人们便忽略它依赖哪些拟合参数。流畅表达有传播价值，却没有独立增加模型与世界之间的约束。

因此，未来科研系统应当被要求回答粗糙而具体的问题：哪些符号可测，怎样换单位，在哪个极限恢复已知结果，哪两个竞争模型会给出不同数字，什么观察会使当前版本退出？这些问题不能保证出现伟大理论，却能阻止尚未受检的组织冒充完成的物理。

对于量子测量与量子引力尤其如此。系统可以整理解释之间的差异，检查某个模型的数学一致性，寻找近似方法或提出可行的区分方案；但这些工作本身不等于已经解决测量问题，更不等于完成引力与量子的统一。本书的发生语言同样不能凭借把装置、主体和关系放进一句话，就宣布世纪难题已被消除。

60.8 让开放纲领具有明确期限

一个可执行的开放纲领可以包含三种不同尺度的任务。第一种是重建：在公开数据上复现已有关系，检查系统是否保持条件与来源。第二种是迁移：改变尺度、边界或噪

声后，看它能否说明原模型在哪些地方失效。第三种是生成：在明确空缺处提出新的变量或关系，并交给独立实验评价。三者应分别报告，不以重建成功冒充第三种成功。

每项任务需要期限和退出条件。假设团队承诺三个月内完成二十个冻结测试，并在其中至少十五个达到事前误差要求，那么结果只有十二个时，当前目标没有达到。研究者可以分析原因，提出新版本，但不能把门槛改成十二个后再把这一轮记为成功。这里的数值是制度设计示例，不是普遍适用的科研评分标准。

多模型碰撞也要接受成本比较。若系统使用十倍算力、两倍实验时间才多得到一个已知结果，其价值不能只按成功数判断。另一方面，一个揭示共同盲点的失败可能比十次重复成功更有意义。评估应同时记录资源、独立新信息和对后续操作的影响，避免用单一排行榜替代科学判断。

可以设置彼此隔离的预测组与反例组。前者提出最有前途的方案，后者专门寻找其条件漏洞；最后由独立实验裁定。反例组的作用不是为了制造争吵，而是把不相容在公开记录中保留下来。若两组实际使用同一数据与同一默认假设，它们的人员不同也不能保证认识差异。

60.9 SDE 在未来研究中应怎样工作

沿 D—E—S 倒逼路径，先让系统建立可追踪的操作序列 D，再明确数据、装置、变量与模型之间的特征纠缠 E，最后提出可重复的显露 S。这里的顺序强调先把任务做成可以执行的过程，而不是先给未来发现起一个宏大名字。

另一条 S—E—D 进入路径则从稳定异常进入：当某个结果持续不能被既有模型解释，先检查关联规定 E，再设计新的差异序列 D 去区分候选。AI 可能帮助搜索异常模式，人类可能提供装置知识，新的操作又可能否定双方最初的解释。复合主体的创造力正体现在这种可互相改变的关系中。

三方程在这里联合约束研究。S=F(D,E)要求结果能够追溯条件，D=G(S,E)要求下一个操作回应已知分歧，E=H(S,D)要求关联解释接受结果的修订。若把它们当作三句总能套用的结束语，便没有增加研究能力。只有在它们改变了变量表、测试安排或失败判断时，方法才实际工作。

三原理同样承担具体责任。S 与 D 的张力可以推动重建 E 中的数据泄漏判断；S 与 E 的张力可以推动 D 中新增独立测量；D 与 E 的张力可以推动把模糊的“发现成功”S 分成重建、预测和新颖性三种可区分结果。每次改造后都要回到原测试，说明哪里改变，哪里仍未解决。未来不是被三原理保证的胜利，而是它们也必须接受检验的场所。

60.10 留给下一位的不是王座

第十一位也许不会以一位独立人物的名字出现。它可能是一种新的合作方式，一套使不同方法保持差异而又能够共同检验的组织，或者一个让旧问题获得新操作入口的装置系统。也可能，下一次深刻创造仍由一位研究者提出关键想法，再由许多人用多年工作使其成立。历史没有承诺只采用其中一种形式。

本书愿意押注的，是人能够与新工具一起扩大提问和检验的能力；它不愿预先替未来宣布答案。若人机系统在透明比较中没有改善问题质量，或持续掩盖误差，就应改变组织方式。若新的共同体能够产生独立可复核的对象关系，即使其形式超出本书预想，也应让事实修订我们的概念。

物理思想简史因此不以“全部完成”结束。它把过去留下的稳定成果交给后来者，也把每一次稳定背后尚可重新打开的条件交给后来者。下一位不需要先接受本书的全部语言，才有资格继续创造；他、她，或他们，需要能把新的判断交给共同世界检验。

空着的位置不是王座，而是一张仍可增写、也可划去旧结论的工作桌。桌上有原典、仪器、算法、失败记录与尚未提出的问题。物理世界的下一次发生，要从这里继续。

本章参考资料

[60-1] Silviu-Marian Udrescu and Max Tegmark, “AI Feynman: A Physics-Inspired Method for Symbolic Regression”, Science Advances 6, eaay2631, 2020；论文摘要与研究范围。https://arxiv.org/abs/1905.11481。

[60-2] Nathan J. Szymanski et al., “Author Correction: An autonomous laboratory for the accelerated synthesis of inorganic materials”, Nature 650, E1, 2026，2026 年 1 月 19 日发布；更正正文关于新颖性、结构复核与训练数据。https://www.nature.com/articles/s41586-025-09992-y。

[60-3] Lisa Messeri and M. J. Crockett, “Artificial intelligence and illusions of understanding in scientific research”, Nature 627, 49–58, 2024；作者机构所载摘要。https://collaborate.princeton.edu/en/publications/artificial-intelligence-and-illusions-of-understanding-in-scienti/。

结语　物理世界为什么仍将不断发生

我们得到的不是一张补齐的名录

一本物理学思想史走到最后，最容易留下的印象，是人类逐项发现了原先不知道的东西：先有运动和力，后来有场和原子，再后来有粒子、黑洞与宇宙整体。这种排列便于记忆，也包含真实的知识增长，却不足以说明本书一直追问的变化。为什么有些现象曾经只能被描述，后来却能够被测量？为什么同一个名字下面的对象规定会改变？为什么一种新的数学关系，能够使原来各自成立的读数进入共同检验？

本书所追踪的，正是这些变化中的物理创造。它不只给已有名录补进条目，也改变条目怎样成立、怎样保持身份、怎样与其他条目发生关系。粒子不是一律以微小硬球的方式存在于理论，场不是一种遍布空间的颜色，温度不是无条件附着在每个单元上的标签。每一种稳定规定背后，都有辨别、比较、相互作用和再检验的具体组织。

所以，“物不是预先穷尽的清单”，不是说研究者能够任意宣布世界中出现什么。它首先意味着，不能用某一历史时刻已经掌握的分类，封闭后来一切可能成立的物理对象。曾经不具有共同量度的现象，可以通过新的关系取得比较资格；曾经被认为不可分开的属性，也可能在新的条件中显示不同结构。未来物理之所以仍可能发生，在于这种成立条件尚未被任何一份现成目录穷尽。

创造不要求抹去过去

说物理世界被创造，立即会遇到一个合理问题：在人类研究以前，恒星是否发光，物体是否相互作用？如果把创造写成意识宣布一件事存在，这本书就会与它依赖的全部历史证据发生冲突。研究能够解释久远过去的辐射、元素和地质痕迹，恰恰要求承认那些过程不以当时已经有人描述它们为条件。

本书的发生取向应在这里说得更精确。物理规定不是把一个词贴到毫无关系的东西上；它在实际作用与可保持的差异中成立，并通过研究取得可追认、可运算、可迁移的公共形式。过去的作用过程可以在没有人类的情况下发生，后来建立的科学对象则要说明，怎样凭现在保留下来的痕迹追认这些过程。不能把这两种历史次序混为一谈，也不能假定它们之间没有桥梁。

当本书谈论一种新的物理世界，强调的是对象、作用和规律的规定方式成为新的可持续系统。它比“换一种说法”要求更多，又不意味着研究者在实验当天制造了全部相关宇宙历史。新系统必须对已经存在的痕迹负责，对可以重复的作用负责，对不随我们

的愿望改变的结果负责。若一个创造论允许任意改写这些责任，它就没有解释物理创造，而是取消了物理约束。

这里仍保留着哲学上的争论空间。科学实在论、结构实在论和操作取向，可以对同一研究成果作出不同存在解释。SDE 的价值不能仅凭它使用了“发生”或“创造”两个词便获得。它必须让读者更准确地看见哪些规定共同成立，哪些条件被隐去，以及怎样的改变才真正构成新对象。这个要求也适用于本书自身。

物体的同一不是永远不变的外壳

我们在实验中不断说“同一个物体”“同一种粒子”“同一块材料”。这些说法并非无条件成立。升高温度可能改变相态，移除传感器可能改变振动，改换参照系会改变部分量值，交换测量程序会改变可比较的结果。物体的同一不应被理解为全部属性始终保持原样，否则任何变化都将让追踪失去对象。

同一性的实际工作，是说明哪些变化可以在一个规定系统内继续追认，哪些变化迫使我们改变身份标准。旋转一块晶体，不必宣布获得了另一种材料；改变其组成或相结构，却可能使原来的材料名称不再充分。换一组坐标不必制造一个新事件；改变复合系统的划分，有时会使原来讨论的关联失去相同意义。判断的依据不是变化有没有发生，而是关系能否按既定规则被保持和转换。

这使本书关于物体的理解具有双重要求。一方面，不能把装置、操作和比较条件全部排除在身份之外，只留下一个永远不可触及的核心。另一方面，也不能因为每次操作有所不同，就宣布没有任何可持续的对象。科学的力量恰在两者之间：它建立可以跨越某些变化而保持的规定，同时让保持的范围受到检验。

未来新对象的出现，未必总是突然增加一类陌生实体。它也可能发生在身份标准的重组处：原来的单一对象被分成不同类型，原来不同的类型通过新的变换进入共同结构，原来仅在局部成立的追认扩展到跨尺度过程。物理思想史因而不能只记录名词第一次出现的时间，还要追踪这些名词承担同一性责任的方式怎样改变。

物理的公共性不需要只有一种形式

从几何证明到微分方程，从统计分布到矩阵代数，从连续场到离散结构，物理的公共表达从未只有一种形式。要求所有真实规律必须长得像某种熟悉方程，会把一种成功历史误写成永恒许可证。反过来，形式多样也不意味着每种语言都能在任何地方同样有效。

公共性的核心，是不同研究者能够知道怎样进入问题，怎样得到结果，怎样检验转换。某些领域以解析公式提供这种入口，某些领域还要依靠数值算法、样品制备、统计规则和校准网络。只要关键条件能够被明确追踪，复杂并不自动意味着不公共；若关键判断只能凭作者权威接受，再短的公式也不能补足它。

本书没有把物理的公共性缩成人人具有相同感觉。图像、声音、触觉提示和数学符号都可以成为进入对象的通道。真正需要共享的是转换责任：同一个结构为何在不同通道中给出不同表现，什么关系允许把这些表现放在一起，哪些差别仍无法协调。以此理解公共性，也就不必把某一种感官经验当成所有物理的唯一原型。

这里有一项尚未完成的研究。视觉、听觉和触觉的 SDE 不能永远停留在直观类比。必须规定具体任务、操作者差异、刺激条件、仪器转换和判断标准，比较何时跨通道组织提高辨别力，何时只是改变显示方式。若这些区分不能被操作化，相关分析就仍主要是一种解释语言；不能直接把它写成已经得到普遍验证的感知科学。

继承不是把后来的一切塞回早先公式

这部简史中的人物创造并没有形成一条整齐的接力线。伽利略与牛顿、法拉第与麦克斯韦、普朗克与玻尔、海森堡与薛定谔、爱因斯坦与狄拉克之间，既有可以追踪的继承，也有对问题资格的改变。马赫、玻尔兹曼和诺特等扩展节点提醒我们，批判、统计组织和结构证明同样能够改变物理，不能只按一条英雄序列计算贡献。

一个后来理论在某些条件下恢复先前公式，是有力的联系，但必须指出恢复的是哪些量、哪些近似以及哪些操作。数值结果相近，不一定意味着全部对象观念相同。反过来，对某种实体解释的放弃，也不必抹去它曾经支持的测量技术和计算成果。科学继承可以同时包含保留、重新定位和舍弃。

因此，全书的共同图并不是要证明每个人早已在不自觉地使用本书的语言。把历史人物改装为一种哲学的代言人，会使思想史失去最重要的材料：他们真正使用的概念、他们面对的困难，以及他们与后来解释之间的距离。本书以 SDE 分析他们的创造，是作者的解释工作，不能倒写成前人自己作出的理论宣告。

更有价值的继承方式，是保留不同路线曾经解决问题的能力，同时公开其边界。后来者可以使用经典力学计算某些运动，而不因此拒绝相对论；可以使用热力学稳定关系，而不要求每次都从微观态重新推导。领域中的有效自治并非理论落后，它使不同尺度的可靠操作得到保存，也为跨尺度问题留下明确接口。

三方程必须继续交出具体工作

S=F(D,E)、D=G(S,E)、E=H(S,D)把显露、差异序列和特征纠缠放在相互规定的关系中。经过全书案例，它们所要求的工作应当比书首更清楚：不能脱离形成条件谈显露，不能脱离可辨结果与联系安排差异，也不能把纠缠组织成一个永远吸收失败的无限背景。第三个式子的责任尤其重要，因为它要求联系本身接受证据限制。

但写下三式，还没有得到任何一条新的物理方程。要与具体物理接合，必须明确变量是什么、输入和输出在哪个层级、哪些映射只是解释、哪些映射真的给出额外约束。如果把所有已有模型一概代入 F，又在每次失败时任意换掉 F，就不能说三式已经获得预测成功。它们作为研究纲领的成绩，应当由它们实际改变了什么来衡量。

同样，六路径不是给每个章节贴一种标签。完整顺序的价值，在于让研究起点经过另两维而形成新的回检任务。从异常出发，与从成熟联系出发，可以通向不同操作安排；但两种顺序都不能跳过证据责任。若研究者只是事后给一段成功过程配上路径名称，而无法据此指出下一项有区分力的工作，方法仍需深化。

三原理的要求也应留在这种具体性中。两维之间出现张力，推动第三维改变，并返回原关系检验。这里没有任何一维被保证永远正确，也没有任何一类修订被保证一定产生新物理。一次校准错误可以被校准修复，一个粗糙参数可以被更精确估计取代。只有当对象规定和发生联系确实改变时，才有理由作更强的创造判断。

尚未完成的数学接口

本书多次从局部差别走向整体组织，也多次讨论整体结果反过来限制局部描述。但差异序列不能简单等于微分，特征纠缠不能简单等于积分。微分与积分有严格的定义及适用条件；三维体系谈论的发生关系，则需要通过具体模型才能接到这些数学操作。两者的接口是研究任务，不是由相近意象完成的同一证明。

一个有内容的后续工作，应选择有限而明确的领域。比如在一种材料测量任务中，预先规定 S、D、E 各自包含哪些量，与传统系统辨识使用同样的数据和资源进行比较。若 SDE 分析能够提前发现未纳入的负载关系，并指导一项新的测试，它就取得了具体方法收益；若没有，就应记录它没有增加预测或诊断能力，而不是无限扩大概念以追回成功。

对于离散、拓扑和随机数学，也需保持同样纪律。它们可以帮助组织连接、不变量、访问频率和尺度转换，却不会因为比初等微积分陌生便自动更接近存在本身。下一

代数学对象的发生，需要同时展示内部一致性、可用的转换和它所打开的物理问题。抽象的增多与物理内容的增多不是同一计数。

因此，未来研究不只寻找“什么公式对应 SDE”，还要辨认对应的层级。某个式子可以帮助整理实验流程，而不描述样品动力学；某种图结构可以表达模型依赖，而不是空间中真实传递的力。区分这些层级，会使本体论主张更有承担，而不是更贫乏。没有清楚接口的宏大对应，反而无法知道自己在哪里错了。

开放不是保证一切问题终将解决

从本书走向未来，不能把开放性理解为一种必胜预言。新的窗口可能遇到不可消除的噪声，某些参数可能在现有条件下不可辨识，某些宇宙区域可能无法传来信息。某种数学问题也可能具有严格的不可能性结果。承认这些限制，不是放弃物理世界的发生，而是使它不再依赖无限能力的幻想。

有时，研究的进展正是发现不能同时得到哪些规定。一个不可兼得关系可以重新组织实验，一个无解证明可以停止徒劳搜索，一项排除范围可以使候选理论必须改变。物理世界的丰富并不只由正面发现组成，也由这些有根据的边界组成。边界把可能与不可能分开，使后续工作知道自己究竟还有哪些选择。

未完成也不总意味着同一种缺陷。黑洞内部的理论问题、暗物质的微观身份、量子测量的解释分歧、复杂材料的可计算性，各自缺少的东西并不相同。把它们统称为旧物理的失败，再用一个总词统一回答，只会抹去真实难度。新的研究需要把困难分解到能提出竞争方案、能安排核验或能给出数学限制的层级。

本书因而不承诺物理学必定沿某条方向永远增长。它提出的是一种有条件的判断：只要新的差异还能被建立，新的联系还能受到约束，新的显露还能被公共地检验，物理规定就仍有被扩展和重组的可能。这种可能性不为任何理论预留胜利，也不要求未来研究沿用本书的全部术语。

把书的边界也留在世界之内

写作一部思想史，同样需要选择窗口、安排比较并形成解释。作者不在所研究的历史之外，也不能因掌握了后来知识，就把历史中的岔路全部写成必经之路。本书提供的是一种可以接受纠正的整体组织，而不是从物理学之外颁布的最终判决。

它的主张也应被分开评价。史料引证是否准确，应由原文与研究资料核对；数学例子是否成立，应由推导和条件检验；哲学解释是否有价值，应看它是否澄清了原来混淆

的关系，是否提出了可以继续工作的限制。某一个层次的错误必须修正，某一个层次的成功也不自动赦免其他层次。

全书最后留下的，不是等待后来者赞同的封闭系统，而是一组可以被接续的区别：物与物体，量值与身份，结果与产生条件，计算表示与物理机制，方法收益与本体论主张。这些区别若能帮助读者更清楚地提出一个问题、检查一次测量、拒绝一项过强推断，它们就已经进入实际研究。若未来工作发现更好的区别，它们也应当被修改。

物理世界仍将不断发生，不是因为世界欠缺一个终极名字，而是因为稳定的关系从未替我们完成所有可能的提问。已有成果为新的操作提供支点，也使新的主张必须承担更严格的责任。后来者不必从头开始，却也不能只重复已经成立的答案。

当一个尚无共同量度的差别获得了可比较的形式，当一种无法迁移的现象找到了可以回检的联系，当一个熟悉对象在新的条件下要求改变其身份，新的物理工作就已经开始。它可能发生在庞大的观测网络，也可能发生在一次很小的校准与反例中。书写到这里可以结束；那些仍能被追问、被检验并被重新组织的关系，不会随着最后一页合上而停止。

参考与说明

本结语回收全书已论证的史例与方法边界，不另增独立史料断言。人物与物理案例的引证见相关各章；三方程、六路径与三原理沿用第四章及第十二编的统一定义。