



目 录

目录	2
现代物理学基础	4
(作者与译者题献)	5
作者前言	6
译者翻译规范	8
译者说明与阅读建议	9
1 早期原子理论	10
1.1 气体性质.....	12
实验关系	12
温标	13
理论解释	15
气体压强的粒子模型*	19
1.2 化学	23
元素	23
结合质量定律	24
体积结合定律	26
气体常数.....	28
Avogadro常数	29
1.3 电解	30
早期电学.....	30
早期磁学	31
电磁学	32
Ohm定律*	35
位移电流*	37
电解的发现	41

1.4 电子43

现代物理学基础

除了其开创性的研究之外，诺贝尔奖得主Steven Weinberg（史蒂文·温伯格）还因一系列备受赞誉的物理学著作而闻名——这些著作将他卓越的物理洞察力与清晰的讲解天赋完美结合。本书在历史背景下描绘现代物理学的基础，并提供了一些新的推导。Weinberg介绍了从原子理论的早期应用到热力学、统计力学、输运理论、狭义相对论、量子力学、核物理以及量子场论。本书既可作为高阶本科生和研究生物理课程的教材基础，也可供职业科学工作者作为现代物理学相关领域的便捷入门读物。

Steven Weinberg是德克萨斯大学奥斯汀分校物理与天文学系的教授。他曾荣获众多奖项，包括诺贝尔物理学奖、国家科学奖章、海涅曼（Heinemann）数学物理奖，以及最近的基础物理学特别突破奖。他是美国国家科学院、英国皇家学会以及美国和国际多个科学院的院士。美国哲学学会曾授予他本杰明·富兰克林（Benjamin Franklin）奖章，并引述他为“很多人心目中在世的最杰出的理论物理学家”¹。他著有多部广受好评的著作，包括《引力与宇宙学》

（*Gravitation and Cosmology*）、三卷本《量子场论》（*The Quantum Theory of Fields*）、《宇宙学》（*Cosmology*）、《量子力学讲义》（*Lectures on Quantum Mechanics*）以及《天体物理学讲义》（*Lectures on Astrophysics*）。

¹ 译者注：本书英文原著出版于2021年，Steven Weinberg（1933 – 2021）于同一年离世。

（作者与译者题献）

For Louise, Elizabeth, and Gabrielle

献给Louise，Elizabeth，和Gabrielle（作者）

For each other, and the minds that lit the way

献给彼此，以及那些为我们照亮道路的启蒙者（译者）

作者前言

本书源自我在德克萨斯大学一门本科物理课程的讲义。我认为自己超出了常见本科课程的内容，让读者初步品味了核物理与量子场论。我还追溯了一些不常被涉及的更早历史时期：回顾了十九世纪时为确立原子的存在及其性质所作出的努力，其中包括热力学的发展——它为这一确立过程造成了些阻碍，但也提供了不同的思路。

我担心一些读者可能会跳过这些更早期的内容，急于阅读他们心中的“主菜”——量子力学和相对论。真这样的话就有点儿遗憾了。我的经验是，许多希望从事原子、核、或粒子物理学的学生，经常是在不怎么熟悉熵（entropy）、能均分（equipartition）、黏滞（viscosity）、扩散（diffusion）等概念的情况下，就顺利通过了他们的正式学业——我自己便是如此。本书，或是基于本书的课程，或许能为一些学生提供最后的机会，了解上述及其他一些理解宏观世界所需的知识。

读者可能会发现本书对历史有特别的侧重。我专门提到一些五花八门的理论构想和没有得到充分理解的实验事实——而正是从他们当中，萌生出二十世纪的现代物理学。至少对我来说，在不了解这些背景的情况下，相对论和量子力学的一套假设看起来简直就是随意、凭空的。这也是件个人品味的事儿：我总觉得如果我们不把物理学研究看作伟大历史进程的一部分，它就会丧失不少精彩。我在本书索引中列出了一些重要的历史著作，还有一些我认为有益的原始文献。

但本书并非历史作品。历史学家的目的在于揭示过去的科学家如何思考他们所面临的问题——比如，1905年，爱因斯坦在建立狭义相对论时，他如何思考空间和时间间隔的测量问题。为了这些目的，历史作品必须深入到物理学家的个人经历、当时的制度发展，和历史上一些错误的理论起点，而要把我们的后见之明放到一边。我虽然也尽量精准还原了物理学在过去的面貌，可本书讨论物理历史问题的目的是不同的：我是为了让读者明确这些问题今天是如何被看待的。

本书主要面向已经有一定基础的物理本科生，或者那些希望简明了解现代物理学某一领域的科研人员。为此，我毫不犹豫地使用了微积分和矩阵代数，虽然并非其高阶版本。书中的数学水平随着内容的推进逐步升高。我尽可能选择具体而非抽象的物理理论表述形式。例如，在第五章的量子力学部分，我主要以波函数的形式来表示物理状态，直到章节末尾才介绍希尔伯特空间中的向量表示。在某些章节中，更详细的材料被放入附录中，可以跳过而不会影响理解理论主线。其中二个附录在我看来是对重要结果推导的改进：附录2.6修订了爱因斯坦关于布朗运动的公式，附录6.4则修正了费米关于 α 衰变的计算。

根据我的经验，经过一些明智的删减，截至本书第五章中间的内容可以在一个学期的本科课程中讲完。但我认为，要讲完整本书，需要整整二个学期的时间。

这本书涉及的主题范围非常广泛，以至于无法完全深入任一主题。当然，它对量子力学、统计力学、输运理论、核物理学和量子场论的介绍都不能替代研究生级别的课程，这些主题中的任何一个都至少需要一年的学习。尽管如此，这本书讲述了我认为在理想情况下，有志向的物理学学生在进入研究生院时就应该已经有的知识——至少，这是我希望自己当年在进入研究生院时就知道的东西。

总的来说，我希望学生或读者能够对书中的内容产生足够兴趣，以至他们愿意深入某一领域并学习更多的专业书籍或课程，我也希望这本书能够为这些进一步的学习提供良好的基础。

我非常感谢许多学生和同事指出了本书所基于的讲义中的错误，也感谢剑桥大学出版社的编辑 Simon Capelin 和 Vince Higgs 为我提供的专业和友好的帮助，他们指导了本书的出版。

STEVEN WEINBERG

译者翻译规范

人名一般不翻译，首次出现时加汉译名于其后括号中；

特定地名、机构名等（如历史地名或翻译后难以辨认的专用名），翻译后加原英文于其后括号中；

部分专有名称（如书名）、专业名称（如规律、公式名），翻译后加原英文于其后括号中，斜体；专有和专业名称中出现的人名一般不翻译，规范如前所述；

原文译文与译者释义扩充解说等以不同字体字号区分：正文译文 Iowan Old Style 16 pt（着重处 **Iowan Old Style Black 16 pt**），释义 Kaiti Sc 16 pt；

脚注区分作者注与译者注，脚注 Iowan Old Style 14 pt；脚注标号数字加在最直接相关的短语后，标点符号前；

图示区分作者图与译者图，作者图来自原书，译者图来自译者手绘或网络公共图片；

出现单位时：若原文为全拼英文，则译为中文；若原文用符号，则保留符号；

出现数字时，同样尊重原文，一般全拼则译；

文中插入公式前后各空一格，数字与单位符号之间空一格；必要时删除紧贴文中数学表达式之后的标点（尤其是句号“。”），以免符号混淆；

单独自然段插入公式时，一般公式居中，公式标号居右对齐；去除紧贴独立行数学表达式之后的标点；

公式符号凡例：黑体为向量，如 $\mathbf{v}$ ， $\mathbf{A}$ ；加“hat”为单位向量，如 $\hat{\mathbf{n}}$ ， $\hat{\mathbf{e}}$ ；译文花体显示与原文有差异（原文花体大概使用了LaTeX环境amsmath，有别于译文环境），如 $\mathcal{N}$ ， $\mathcal{T}$ ， $\mathcal{E}$ ；积分符号下标指明积分区域，如 $\int_{\omega}$ ， $\oint_{\partial\omega}$ ， $\int_{\infty}$ ，分别表示对区域 ω ，其边界 $\partial\omega$ ，及无穷大区域积分；不引起混淆的情况下，不区分多维积分元写法，如 $d^3\mathbf{x}$ 与 d^3x 等同（具体表达形式尽量尊重原文习惯）；平均值用bracket或bar，如 $\langle x \rangle$ 与 $\bar{x}$ ；

尽力信达雅：首先“信”不扭曲原意，尽可能“逐字译出”不添不减；为了“达”而不拘泥于原英文词序语序，必要时依汉语习惯添加辅助词汇；本书预设为学术教材，主要以事实、逻辑、数学模型等高度理性方式表达，翻译中慎用诗意的朦胧的情感渲染的“雅”。

译者说明与阅读建议

1. 译者认为作者呈现了十分难得的资料，既非常规教材的常规内容，亦非寻常科普的大众故事。前者难免陷于中规中矩的展现方式和千篇一律的取材，后者往往回避数学避重就轻甚至流于惊悚的标题与夸张的轶闻轶事。作者不仅取材逻辑自成一体他处罕见，讲解更是独树一帜不落窠臼。译者强烈建议读者既不要跳过数学与逻辑体系而只看故事，也不要断章只取技术细节而忽略历史与叙事之逻辑结构。
2. 原文数学部分程度变化剧烈，作者并不会刻意避免困难的数学。译者对部分演算步骤和推导细节添加了注释，以及另辟额外段落进行详细阐释。有些可能有助于读者跟上作者逻辑，有些添加的内容则可能因其广度与深度失当，给读者制造了更多困难。译者建议读者依数学程度渐进，大可不必要求一次透彻，完全可以折返。遇到译者添加的大段数学内容，以及单独以星号*标题的小节，首次阅读完全可以跳过。
3. 几乎任何一个具体话题都牵连着一系列其他话题，译者释义和扩充所预期的功能之一，是希望每个单独话题的呈现至少初看起来大致“完整”，或者叫“自给自足”（self-contained），略微减少读者在构建知识网络过程中四处搜索的麻烦。译者也尝试效法作者，尽量少一点人云亦云，不使读者感觉“重复的书又多了一本，雷同的话又说了一遍”。
4. 今日之人工智能技术刹那可将整部著作转换语言，文字翻译本身几乎贬值殆尽，甚至多此一举。译者私心在于迫使自己学习一番，籍此工作将自身如海绵一般漏洞百出的知识体系填实几分，不少笔墨都来自译者的学习笔记。若能使读者感觉稍许有益或有趣，译者必然大受激励，敬请读者纠错赐教。
5. 译者之一20多年前初到美国学习理论物理，恰好有一周时间没有电脑上网不便，于是静下心来阅读Steven Weinberg所著《量子场论》（*The Quantum Theory of Fields*）第一卷开头，作者第二章先概括了量子力学基本框架体系。彼时译者已经学过本科程度的量子力学，读来却如醍醐灌顶，头脑中一团杂乱的知识立现结构森严。译者才疏学浅，但自认为对学问的审美并不粗陋，Weinberg之学仰之弥高，钻之弥坚，Weinberg之作值得学子切磋琢磨。
6. 开卷请务必阅读作者前言。

1 早期原子理论

“物质由原子组成”是一个古老的观念：“原子”最初即指微小的、不可分割的颗粒，在空无一物的空间中运动。这一理论可以追溯到Democritus（德谟克里特）。他居于希腊阿卑德拉城（Abdera），位于爱琴海北岸。在公元前五世纪末期Democritus宣称：“只有原子和虚空是真实存在的。”他既没有为这一假说提供证据，也没有进行任何能够验证其预言的计算。尽管如此，这一观点依然影响力巨大，即使只是作为一个例子来说明在不援引神明的情况下依然可能解释自然现象。略晚于公元前300年，Epicurus of Samos（伊壁鸠鲁）创立了雅典四大学派之一的花园学派（The Garden），而原子被引入了Epicurus的物质主义（Materialism）哲学中。继之，罗马的Lucretius（卢克莱修）在其诗作《物性论》（*On the Nature of Things*）中又引用了原子学说和Epicurus的哲学。这首诗在1417年被重新发现后²，影响了Machiavelli（马基雅维利）、More（莫尔）、Shakespeare（莎士比亚）、Montaigne（蒙田）和Newton（牛顿）等人。Newton在他的《光学》（*Opticks*）中推测，物质表现出的性质源起于原子聚集成大颗粒，这些大颗粒又聚集成更大的颗粒，依此类推。正如我们将看到的，Newton曾尝试提出一种关于空气压强的原子理论，但并未取得显著成功。

Democritus（约460 – 约370 BC）认为只有原子这种颗粒和纯粹的空间（或称真空）是现实存在的，原子与原子之间就是真空。其后的希腊哲学家Epicurus of Samos（伊壁鸠鲁）和罗马哲学家Lucretius（卢克莱修）都引用了这种观点来论证自然现象是自然法则的结果，而无需引述神来解释。

Epicurus（341 – 270 BC）认为，人最大的痛苦来源包括恐惧死亡，以及害怕神明的惩罚，尤其是死后的审判。Epicurus引用Democritus的原子论来说明物质宇宙不是由某个至高的神（所谓“超级存在”）所创造的，而是由最基本的元素粒子混合所形成，这些粒子永恒存在并受某些简单规律的支配。他借助Democritus的原子理论，来论证人对死亡和神明的恐惧是非理性的：

1. 为何死亡不值得害怕：世间万物，包括人的灵魂，都是由原子构成的。死亡是原子的解体：当人死亡时，灵魂的原子分散，意识随之消失，因此死后并没有感知，也没有痛苦。

² 译者注：公元九至十五世纪的欧洲文献虽有少数提及Lucretius的记录，但这段时间内并无已知的《物性论》手稿幸存。《物性论》于1417年被文艺复兴意大利学者Poggio Bracciolini（波吉奥·布拉奇奥里尼）重新发现，很可能是在富尔达（Fulda）的本笃会图书馆中找到的。Poggio发现的手稿没有保存下来，但他的朋友Niccolò de' Niccoli（尼科洛·德·尼科利）为其抄写的副本得以保存，现保存在佛罗伦萨的劳伦蒂安（Laurentian）图书馆。

“当我们存在时，死亡并不存在；当死亡来临时，我们已不复存在。”（When we are, death is not come, and, when death is come, we are not.）既然死亡之后不再有意识，就没有任何痛苦可言。因此，害怕死亡是毫无意义的。

2. 为何神明不值得恐惧：Epicurus并不否认神明的存在，但他认为神明是由最精细、最稳定的原子构成的，神明永远生活在绝对的幸福与安宁中——与所有影响人类的热情、欲望和恐惧无关——并且对世界及其居民完全漠不关心，对人类的美德和罪行都彻底无动于衷。这意味着人类无需惧怕神明。自然现象有物理原因，而非神明意志，世间的风暴、瘟疫、地震等现象，都是原子运动的必然结果，而不是神明的愤怒或惩罚。

Epicurus将Democritus的原子论作为一种心理解放的工具，来消解人类对死亡和神明的恐惧，以帮助人达到所谓ataraxia（即内心的平和宁静与无忧无虑，有时译为“毫无纷扰”）之境界。

Lucretius（约99 – 约55 BC）选用了一种吸引人的形式继续阐明发展这些观点——他在长篇诗歌《物性论》（On the Nature of Things）中表达自然界中的一切都可以通过自然法则来解释，而不需要神明的干预。Lucretius通过论证³表明世界的运作可以归结为自然现象，而这些现象是微小的原子在空无的空间中有规则但无目的的运动和相互作用的结果，一切并无神明干预，从而反对对神明的恐惧。Lucretius写《物性论》这首史诗的目的之一是献给“Memmius（梅米乌斯）”，可能是指公元前58年担任裁判官（praetor，所谓裁判官是裁决公民与政府之间争议的司法官员）的Gaius Memmius（盖乌斯·梅米乌斯）。整部长诗中有十几处提到“Memmius”的地方，以各种语境出现，可翻译成如“我的梅米乌斯”，“梅米乌斯我亲爱的”，“杰出的梅米乌斯”等。根据Lucretius在诗中的频繁表述，这部作品的主要目的是要从对超自然的迷信和对死亡的恐惧中解放Memmius的心灵——并通过阐述Epicurus的哲学体系，将其引入宁静无忧之心境（即前文所述的ataraxia）。

原子理论真正严谨的科学应用开始于十八世纪，首先是在对气体性质的计算上——此前一个世纪以来，气体就已通过实验被研究。本章我们就以此为主题开头。原子理论随之在十九世纪化学和电解中的应用将在后续各节进行探讨。本章最后一节将讲述，随着电子的发现，原子的本质如何开始得到澄清。在下一章中，我们将看到估算原子的质量和大小是如何成为可能的⁴。

³ 译者注：此处的“论证”并非符合现代科学标准的（基于假说、做出预言、实验检验的）论证，而更像是对一种解释现实的可能性的“自我一致而非严谨的故事”，比如“无中不能生有”（nothing can be produced from nothing）。更多讨论可以阅读作者Weinberg的另一本书：《解释世界：现代科学的发现》（To Explain The World: The Discovery of Modern Science）。

⁴ 作者注：有关这些发展的更多历史细节，可以参考Weinberg所著《亚原子粒子的发现》（Discovery of Subatomic Particles），已列在参考书目中。

1.1 气体性质

实验关系

十七世纪实验研究热情的高涨主要集中在空气性质的问题上。这些实验的执行和结论报告并不依赖关于原子的假说，但此处我们需要回顾一下这些实验，因为它们的结果为后来的气体性质理论提供了背景，而后来这些理论则确实依赖于关于原子的假设。

Aristotle（亚里士多德）及其追随者曾认为，人们观察到的泵和风箱的吸力来自于自然界对真空的排斥。这一观点在1640年代遭到挑战，来源是意大利佛罗伦萨的博学家Evangelista Torricelli（埃万杰利斯塔·托里切利，1608–1647）发明的气压计。如果自然界排斥真空，那么当一端封闭的长玻璃管装满水银，保持封闭端在上方并竖立起来时，为什么水银会从下端流出，直到水银柱高度只剩为760 mm高，而在水银上方出现空隙呢？自然界对真空的排斥难道存在一个限度？Torricelli则提出，水银其实是被作用在玻璃管的开口端的空气压力撑住的（或者，如果是把玻璃管开口浸入水银池中，气压作用所在的水银池的液面），这个对应压强正好足以支撑760 mm高的水银柱。如果是这样，那么就应该可以使用竖直玻璃管中的水银柱来测量空气压强的变化——这种装置我们今天称之为气压计。Blaise Pascal（布莱兹·帕斯卡，1623–1662）从1648年到1651年进行了一系列这样的测量，他发现，当气压计移至山顶时，水银柱的高度会下降——此处（山顶）气压计上方的空气较少。

拉丁文Horror vacui，或者英文nature abhors a vacuum，“自然界厌恶（排斥）真空”，是一种被归于Aristotle的假说。他认为自然界不可能包含真空，因为一个区域一旦密度降低，周围更稠密的物质连续体（continuum）会立即填充即将形成的空洞。（该假说后来被Epicurus和Lucretius的原子论所批评。）Aristotle还在更抽象的概念层面上反对真空概念：因为真空（void）仅仅是虚无（nothingness），遵循他老师Plato（柏拉图）的观点，虚无（nothingness）不能被说成是存在的（即，根本逻辑上概念上，说“无”等于“有”就是不正确的）。此外，鉴于真空没有特征，它无法被感官感知；真空概念这一理论假设也并不产生对事物额外的解释能力。与真空（vacuum）对立的观念是“充实体”（plenum），即空间中总是充满了物质，不存在“无”物质的情况——即使并不是所有物质都对肉眼可见。

Torricelli对（基于他的水银柱实验装置制成的）气压计原理的基本解释是：空气有重量，压到地表时足够强以至于能托起水银柱；真空并不本质上被自然界排斥，一旦空气的压力不足以让水银柱填满玻璃管，真空就可以出现可以存在。而Aristotle学说的学者看法几乎正相反：空气无重量，不存在空气压力，真空本质上就不能存在，玻璃管中水银上方的空隙是由液体蒸气填充的。

1646年，Blaise Pascal（布莱兹·帕斯卡）在听闻Torricelli实验后，设计了一个实验来检验Aristotle学者提出的假设。他的实验比较了水和酒：如果玻璃管空隙里的气体是挥发来的，而己知酒比水更容易挥发成气体，应该可以观察到装酒的玻璃管在相同条件下比水挥发出更多气体（有更大的空隙）。Pascal进行了公开实验，并邀请Aristotle学者提前预测结果。Aristotle学者预测酒柱会更低，但结果并非如此。

Pascal的逻辑推演再进一步：如果空气有重量（像Torricelli设想的那样），那么在海拔更高的地方，能够施加压力的空气总重量小于更低的位置——气压在更高的海拔处会更低。Pascal写信给住在普伊·德·多姆山（Puy de Dôme）附近的姐夫Florin Perier（弗洛林·佩里耶），请他进行一个关键实验。Perier带着气压计爬上普伊·德·多姆山，在途中测量水银柱的高度，将测量结果与山脚下的结果对比。1648年，Perier严谨地完成了实验，发现Pascal的预测是正确的。当气压计被带到更高的海拔时，水银柱的高度确实较低。

Pascal在回应充实体（plenum）学说的Étienne Noël（埃蒂安·诺埃尔）⁵的信件中写道：“为了证明一个假说，所有现象都随它说得通并不够；正相反，只要假说导出的东西与哪怕一个现象相悖，就足以证明它是错的。”这当于后世科学标准中的可证伪性（falsifiability）观念。

早于任何对气体性质的正确理论理解，空气压强的定量实验研究就已经开始了。1662年，安格鲁-爱尔兰贵族Robert Boyle（罗伯特·波义耳，1627–1691）在其著作《关于空气弹性及其效应的新物理力学实验》（*New Experiments Physico-Mechanicall Concerning the Spring of the Air and its Effects*）第二版中，描述了对一份固定质量的空气，其压强（“空气的弹性”）与体积之间关系的实验。利用被水银柱封闭在玻璃管末端的空气研究，他发现，在恒定温度下，推动水银表面，空气会被压缩。这就是后来所谓的波义耳定律（*Boyle's law*）⁶：在恒温下，对一份固定质量和成分的气体，其体积与压强成反比（压强的概念就是被Boyle定义为了作用在气体上每单位面积的力）。

温标

关于“恒定温度”这一说法，必须做一点说明。Boyle生活在现代华氏温标和摄氏温标建立之前，后二者的先驱分别要等到1724年和1742年。然而，尽管Boyle的时代无法为任何给定物体的温度确定一个有意义的数值，却仍然可以精确讨论二个物体处于**相同**温度的情况：如果二个

⁵ 译者注：Étienne Noël（埃蒂安·诺埃尔）是一位耶稣会神父和自然哲学家，他曾是René Descartes（勒内·笛卡尔）的老师，Descartes也支持充实体（plenum）学说。

⁶ 译者注：这一关系最早由17世纪的Richard Towneley（理查德·汤恩利）和Henry Power（亨利·鲍尔）观察提出。然而，系统地进行实验并于1662年发表研究成果的是Robert Boyle。有趣的是，当时Boyle的助手，鼎鼎大名的Robert Hooke（罗伯特·胡克），被认为设计并制造了实验装置，使Boyle得以进行精确测量。

物体接触而没有任何一个显著变热或变冷，那么二者就处于相同温度。Boyle让玻璃管保持恒定温度的方法是将其浸入一个大浴槽中，例如，正在融化的冰形成的水浴。后来，华氏温标的建立是定义了冰融化的温度为 32°F ，以及在平均大气压下水沸腾的温度为 212°F ，然后在水银温度计的玻璃管上 32 和 212 之间蚀刻 $212 - 32$ 等分区间，从而定义 1°F 的单位温度。同样，在摄氏温标中，融化冰和沸水的温度分别被定义为 0°C 和 100°C ， 1°C 是从融化冰到沸水的加热过程中温度计中的水银体积增加 1 % 所需的温度差。我们将在下一章看到，存在一个更为深刻而普适的温度定义，相对其而言，基于水银温度计的温标仅是提供了一个不错的近似。

读者至此可能意识到，早期初创的温标恐怕比长度、重量等标准更模糊：

1. “等温”的判断不仅依赖于感觉，甚至冰融化或水沸腾是否总在同一“温度”都并不直观。经验感受上我们可能会接受常见物质中有某种冷热的流动（早期同样不清楚但颇为类似的另一个现象是物质中电的流动），这个冷热流动似乎和温度的概念纠缠在一起，而且二者都伴随物质样貌的变化，如冰融化，水凝结，颜色变化，热胀冷缩等。冰水接触时，如果“热”流出水则水“变冷”（温度降低），“热”流向冰则冰“变热”（温度升高），经常伴随冰与水比例变化，这个过程二者是否温度相同呢？直观感受上并非总如此，比如将冰置于热水中迅速融化殆尽。如同正文中所述，可以假设稳定共存很久的冰水混合物相互“冷热平衡”（温度相等），但难以确定这个相等的“温度”是否取决于其他环境条件。
2. 即使假设冰水共存和水沸腾总是二个固定的温度点，比如 0° 和 100° ，将一个玻璃管中热胀冷缩的液柱对应体积蚀刻后100等分（这种温度计还可以向 0° 和 100° 外侧继续刻度），但恐怕难以保证二个不同的温度计其 1° 是相同的：比如水银管和酒精管，它们刻度中间的 50° 难道是一样的吗？换句话说，用水银温度计测出来的 50° 的水，与酒精温度计测出的 50° 的水，能够冷热平衡吗？
3. 除了将在下一章出现的普适的不依赖于具体物质的温度定义，幸好还有下文Gay-Lussac和Charles的实验结论，既与普适温度等价，又方便在现实中进行测量。如同水银和酒精，稀薄气体的体积也是热胀冷缩的，二者可以在每一个温度相比较：如温度升高时，每当水银体积增大 1 %，气体体积对应增大多少；二者未必成线性关系，但谁是正比于温度变化的并不知道（或许都不是）。重要区别之处在于，如果仿照水银或酒精这类温度计，以一种等压气体的体积来为温度定标，即 0° 和 100° 之间将气体容器管等分，做成气体温度计，那么任何其他等压气体制作的温度计，都具有相同的 1° 。也就是说，虽然还不清楚温度到底是什么，但氢气，氧气，氮气，空气等各种气体的体积变化之间相互成线性关系，用他们制作的温度计是等价的，这些任何温度数值都是“同一温度”，他们测出的 50° 水的确能够冷热平衡。甚至可以这样说，在普适的温度概念与微观理论建立之前，温标实际上是用气体来定义的。

温标建立之后，就可以进行体积与温度关系的定量研究了。为了同时保持气体压强与质量恒定，可以将空气封闭在一个具有可动器壁的容器中——器壁膨胀或收缩以保持容器内部压强与外界空气压强相等。1802年，Joseph Louis Gay-Lussac（约瑟夫·路易·盖-吕萨克，1775–1850）在一次讲座中宣布了这一定量关系，并将其归功于Jacques Charles（雅克·查尔斯，1746–1823）在1780年代未发表的工作。此关系其后被称为查尔斯定律（*Charles Law*）：在恒定压强和质量下，气体的体积与 $\mathcal{T} - \mathcal{T}_0$ 成正比——其中 $\mathcal{T}$ 是通过如水银温度计等方式测得的温度， $\mathcal{T}_0$ 是一个常数，其数值依赖于所使用的温度单位（ $\mathcal{T}_0 = -459.67\text{ }^\circ\text{F} = -273.15\text{ }^\circ\text{C}$ ）。因此 $\mathcal{T}_0$ 是绝对零度，即可能的最低温度，在此温度下气体体积消失为零。使用摄氏单位表达温度差的话，绝对温度 $T \equiv \mathcal{T} - \mathcal{T}_0$ 今天被称为开尔文（Kelvin）温度，记作 K。

理论解释

在其伟大著作《自然哲学的数学原理》（*Principia*）之命题23中，Isaac Newton（艾萨克·牛顿，1643–1727）将空气考虑成由相隔一定距离、彼此相斥的粒子组成⁷——以此尝试解释Boyle定律。仅仅使用量纲分析，他证明如果间距 r 的粒子之间的排斥力按 $1/r$ 衰减，那么对于一份固定质量的空气，其压强 p 与体积 V 成反比；但正如他所指出的，如果排斥力按 $1/r^2$ 衰减，则 $p \propto V^{-4/3}$ 。不过他并未提供任何理由来说明为什么排斥力应该按 $1/r$ 变化。我们之后将看到：并不是按 $1/r$ 变化的力，而是作用范围极小的、只在碰撞时发生的力最能解释说明气体的性质。

Newton《原理》中量纲分析的大致思路：

考虑一个立方体的气体（流体），假设其由互斥的粒子组成。我们想要双向推出关系：如果气体质量不变时，其压强与体积成反比（与密度成正比），则粒子间排斥力与间距成反比；反之亦然。首先假设已知前者，在这样的关系下，保持系统位形（保持其中每个粒子的相对位置），将立方体边长压缩为1/2，其密度会变成8倍——压强也就变成8倍。我们考虑一个把立方体分隔成二部分的截面（上下，前后，左右，斜着等任何方向分隔都行），这个面在压缩中保持几何相似，但面积变为1/4，于是气体作用在这个面上的压力就变成2倍（8倍压强乘以1/4面积）。假如这个压力完全就是粒子们相互排斥力的总和（比如，选定一个横截面的话，压力来源于上部空间的每个粒子排斥下部空间的每个粒子），那就意味着平摊到每个（或每对）粒子上的排斥力变为了2倍。注意到粒子之间相对位置都没有变化，只是所有相互间距都缩为1/

⁷ 译者注：原文particles repelling each other at a distance，指的是action at a distance，即隔空（相隔一定距离）相互作用，称为“超距作用”；区别于action at contact，接触相互作用或“近距作用”，与后文提到的“碰撞时发生的力”相对应。

2, 于是得到结论: 粒子间的排斥力与间距成反比。反之, 如果已知粒子间的排斥力与间距成反比, 那么还是保持位形压缩正方体到 $1/2$ 边长 ($1/8$ 体积)。此时所有粒子间距都缩为 $1/2$, 所有排斥力都变为2倍; 还是考虑立方体内任意一个保持几何相似的分隔面, 粒子们作用在这个分隔面的压力如果就是相互排斥力之和的话, 也应该成同一比例变化为2倍, 对应压强就是8倍 (2倍压力除以 $1/4$ 面积)。于是证明了: 当粒子间的排斥力与间距成反比时, 气体满足Boyle定律。

按类似方法分析, 如果排斥力是距离平方反比的, 那么边长压缩到 $1/x$, 排斥力变成 x^2 倍, 任意一块面积变成 $1/x^2$, 于是压强 p 变成 x^4 倍, 而同时体积 V 变成 $1/x^3$, 这样就有 $p \propto V^{-4/3}$ 。

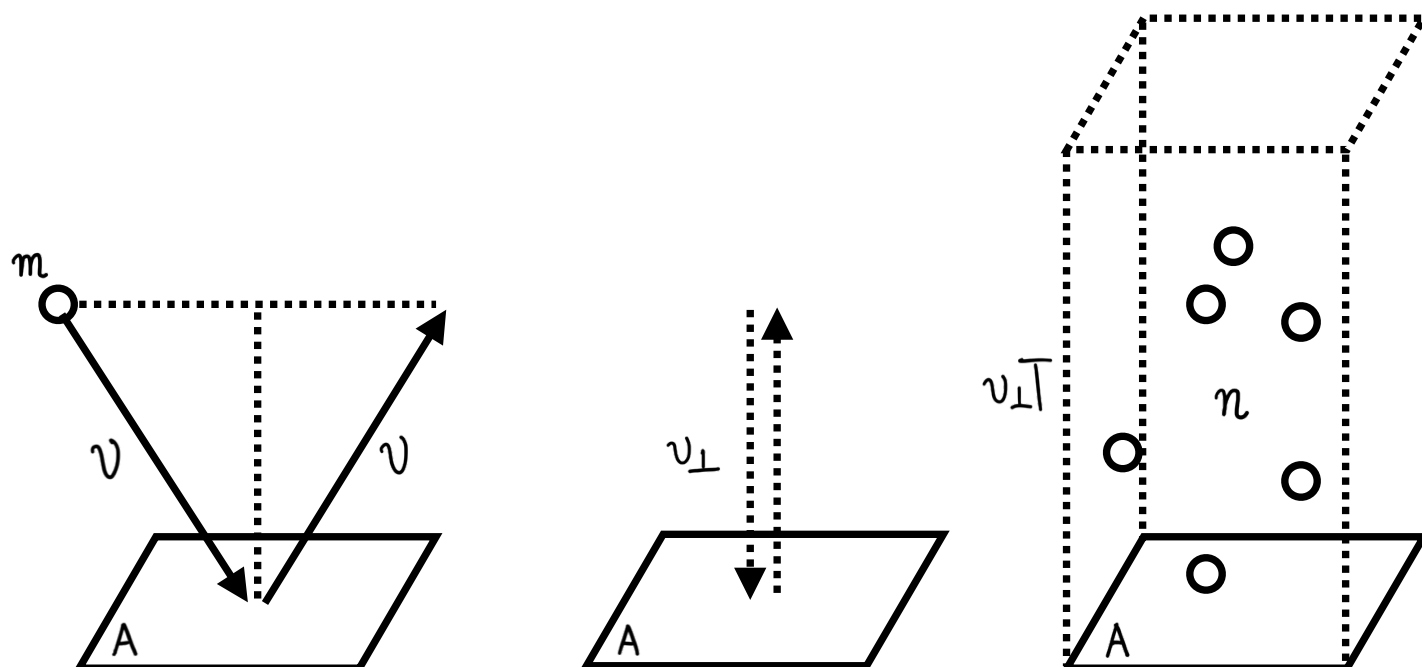
尽管Newton在《原理》命题23下方明确说粒子是运动的 (而且相互之间的离心力恰好有 $1/r$ 的变化关系): “如果一种流体由相互绕转飞行的粒子组成, 并且其密度与压缩程度成正比, 那么粒子的离心力将与它们中心之间的距离成反比”; 但其论证却是基于流体中每个粒子保持位置静止而存在“长程” (超距) 相互排斥而作出的。这和下文将看到的后世的模型几乎正相反: 粒子是运动的但只有“短程” (接触) 相互作用。

瑞士数学物理学家Daniel Bernoulli (丹尼尔·伯努利, 1700–1782) 最先作出如下理论假设来理解气体的性质: 气体由许多微小粒子组成, 除极短暂的碰撞以外, 这些粒子自由运动。1738年, 在其著作《流体力学》(Hydrodynamics) 之“关于弹性流体的性质和运动, 特别是空气”一章中, 他认为: 对于一份气体 (当时称为“弹性流体”), 如果其中每单位体积含有 n 个粒子以相同的速度 v 运动 (由于碰撞⁸), 所有方向都相同, 则压强与 n 和 v^2 成正比。这是因为在给定的时间内, 撞击容器壁上任意给定区域的粒子数, 与任意给定体积中的粒子数成正比; 又与粒子撞击容器壁的频率成正比, 而频率与 v 成正比; 再与每个粒子施加在容器壁上的力成正比, 这个力也与 v 成正比。对固定质量的气体, n 与体积 V 成反比, 因此 pV 与 v^2 成正比⁹。如果 (如Bernoulli所认为的) v^2 仅依赖于温度, 这就解释了Boyle定律。如果 v^2 与绝对温度成正比, 则亦可得出Charles定律¹⁰。

⁸ 译者注: 最简单的模型: 假设每个粒子都是质量相同的球体, 且只发生完全弹性碰撞; 速度不一样的两个粒子相撞, 如果只在法向 (球心连线) 方向上相互作用, 则法向速度改变 (在两个粒子的质心参照系中两者速度交换), 而切向维持原本的速度。但即使在如此理想化的情景下, 也并不能证明得到“碰撞导致速度趋同”, 而且速度趋同也并不是观察事实——但在下一章引入微观态和熵的概念后, 我们会看到这个假设为何起码是部分合理的。同时, 我们很快就会在本章节看到Clausius的推导并没有依赖于这个假设, 同时译者还加入了另一个推导方法, 可资比较。

⁹ 译者注: 压强 $p \propto$ (给定时间内撞击数 $\times$ 单次撞击的力) $\propto$ (给定体积中的粒子数 $\times$ 撞击频率 $\times$ 单次撞击的力) $\propto n \cdot v \cdot v \propto \frac{1}{V} \cdot v^2$, 于是有 $pV \propto v^2$ 。

¹⁰ 译者注: 如果读者此时也和译者一样又一次忘记了理论的具体内容的话: Boyle定律说气体的体积与压强在在恒温下成反比; Charles定律说气体的体积与 $T \equiv \mathcal{T} - \mathcal{T}_0$ 在恒定压强下成正比。



译者图 1.1：单个粒子单次撞击容器壁，完全弹性反弹，其动量法向分量变化 $-2mv_{\perp}$ ，这些动量在相接触的时间 t 内转移给容器壁，贡献法向力 $F = -2mv_{\perp}/t$ 。而后考虑时间段 T 内柱体中所有粒子撞击（并非同时）的贡献，假设每个粒子只撞击一次（容器足够大，气体足够稀薄），撞击时刻“随机”，即每个 t 在 T 内均匀分布。

Bernoulli没有给出太多数学细节，也没有试图说明压强除了与 nv^2 成正比之外还可能与什么成比例——而弄清楚此事在化学史上至关重要。这些细节由Rudolf Clausius（鲁道夫·克劳修斯，1822–1888）于1857年在文章“我们称为热的运动之本质”（*The Nature of the Motion which We Call Heat*）中给出。以下是一个与Clausius本人的推导比较一致的呈现，使用了略微不同的符号。

设想一个粒子撞击容器壁（参见译者图 1.1），并在一个很短的时间 t 内与之保持接触，在此期间它施加于器壁的力沿器壁内法向的分量为 F 。于是粒子沿器壁内法向方向的动量将减少 Ft 。如果粒子撞击器壁之前的这一速度分量为 $v_{\perp} > 0$ ，并且它以 $-v_{\perp}$ 的法向速度分量被完全弹性（elastic）反弹，则动量的法向分量变化为 $-2mv_{\perp}$ ，其中 m 是粒子质量，因此

$$F = -2mv_{\perp}/t$$

现在，设想在时间间隔 $T \gg t$ 内，大量粒子持续不断撞击器壁，所有粒子都具有相同的速度向量 $\mathbf{v}$ 。在这段时间间隔内，将会撞击器壁上面积 A 的粒子数 $\mathcal{N}$ ，正好是以 A 为底，以 $v_{\perp}T$ 为高的柱体中粒子的数量，即

$$\mathcal{N} = nAv_{\perp}T$$

其中 n 是粒子数密度——每单位体积的粒子数。每个粒子与墙面接触的时长与 T 的比例为 t/T ，所以施加在墙面上的总和的力为¹¹

$$F\mathcal{N}(t/T) = 2mv_{\perp}/t \times nAv_{\perp}T \times (t/T) = 2nmv_{\perp}^2A$$

我们看到，所有对时间 t 和 T 的依赖关系都抵消了。压强 p 被定义为单位面积上的力，所以上述结果给出了关系

$$p = 2nmv_{\perp}^2 \quad (1.1.1)$$

上述情形并非物理现实：每个粒子都有相同的 $v_{\perp}$ 值，且都是正的，意思是粒子被假设都是朝向器壁运动的。在现实世界中，不同的粒子将以不同速率在不同方向运动，因此方程 (1.1.1) 应该被替换为

$$p = 2nm \times \frac{1}{2} \langle v_{\perp}^2 \rangle = nm \langle v_{\perp}^2 \rangle \quad (1.1.2)$$

其中尖括号表示对所有气体粒子的平均值，第一个表达式中插入了因子 $1/2$ 的原因是只有 50 % 的粒子会朝向任意给定的器壁区域运动。

为了将 $\langle v_{\perp}^2 \rangle$ 用方均根速度表示，Clausius没有证明而直接假设“平均而言，[粒子速度在]每个方向都等可能地出现¹²。”在这种情况下，速度每个分量的平方平均值都等于 $\langle v_{\perp}^2 \rangle$ ，因此速度向量的平方平均值为

$$\langle \mathbf{v}^2 \rangle = \langle v_1^2 \rangle + \langle v_2^2 \rangle + \langle v_3^2 \rangle = 3\langle v_{\perp}^2 \rangle$$

因此方程 (1.1.2) 可写作

$$p = nm \langle \mathbf{v}^2 \rangle / 3 \quad (1.1.3)$$

这实际上就是Bernoulli算的结果 $p \propto n \langle \mathbf{v}^2 \rangle$ ，多加了因子 $m/3$ ——方程 (1.1.3) 现在成为了等式，而不只是陈述正比关系。对于占据体积 V 的质量 M 的气体，其粒子数密度为 $n = M/mV$ ，因此Clausius可以利用Boyle定律（他称之为马里奥特定律（Mariotte's law））

¹¹ 译者注：并不是所有粒子都在同一时间撞击容器壁，每一个时刻容器壁上受到的力只由那一个时刻撞击其上的粒子决定。由于柱体里的每一个粒子都只有 t/T 比例的总时间在和容器壁接触，如果粒子们接触容器壁的时间分布得足够均匀，就可以认为平均来讲每一个时刻接触容器壁的粒子数只是总粒子数比例的 t/T 。

¹² 译者注：“on the average each direction [of the particle velocities] is equally represented.”可直译为：平均而言，[粒子速度的]每个方向都被均等体现（或展现、表达）。意思是没有任何方向比其他方向更特殊，所有方向是平权的等概率的。

——即恒定温度下 pV 是常数——推断出：对一份给定气体， $\langle \mathbf{v}^2 \rangle$ 仅依赖于温度。此外，正如 Clausius 所指出的，方程 (1.1.3) 结合 Charles 定律（Clausius 称之为盖-吕萨克定律（law of Gay-Lussac））表明： $\langle \mathbf{v}^2 \rangle$ 与绝对温度 T 成正比。如果我们愿意，可以采用现代符号并将这个比例常数写为 $3k/m$ ，这样

$$m\langle \mathbf{v}^2 \rangle / 3 = kT \quad (1.1.4)$$

因此方程 (1.1.3) 可以写成

$$p = nkT \quad (1.1.5)$$

其中 k 是一个常数，意思是独立于 p ， n 和 T 。但是这个符号的选择并不能告诉我们 k 是否因气体种类不同而变化，或者其是否依赖于分子质量 m 。Clausius 无法回答这个问题，也没有为 Boyle 定律或 Charles 定律提供任何理论上的论证。Clausius 称得上热力学的奠基人（这将在 2.2 节和 2.3 节中被讨论），但上述这些并非是仅靠热力学就能回答的问题。正如我们将在后续章节中看到的，关于气体化学的实验表明， k 对所有气体都相同——它是一个普适常数，现在被称为玻尔兹曼常数（Boltzmann's constant）。但对其以及对 Boyle 定律和 Charles 定律的理论解释必须等到分子运动论和统计力学的发展，这是 2.4 节的主题。

正如其文章标题“我们称为热的运动之本质”所示，Clausius 一心想要表明，至少在气体中，热现象应由组成气体的粒子的运动来解释。他通过使用自己的理论计算出气体的比热来支撑其可靠性，我们将在下一章中关注这一话题。

气体压强的粒子模型*

我们用更接近现代表述的方式推导 Clausius 对气体压强的结论以资比较：所谓粒子构成的气体，其模型是在体积 V 中 $\mathcal{N}$ 个满足牛顿定律的质点构成的系统，每个质点质量都是 m ，除彼此接触碰撞（近距作用力）外无其他相互作用（无超距作用力）。具体的计算需要不少假设：

- 假设 $\mathcal{N}$ 足够大，以至于一些变量可以处理成连续的，如密度、压强及粒子其他的分布函数：数密度就是空间分布函数 $n(\mathbf{x})$ ，位置处于 $\mathbf{x}$ 到 $\mathbf{x} + d\mathbf{x}$ （微分小方块）的粒子数为 $n(\mathbf{x}) d^3x = n(\mathbf{x}) dx_1 dx_2 dx_3$ ， $\int_V n(\mathbf{x}) d^3x = \mathcal{N}$ ，后文更经常使用一个粒子处于 $\mathbf{x}$ 附近的概率密度 $n(\mathbf{x})/\mathcal{N}$ ；我们还需要考虑速度分布函数 $f(\mathbf{v})$ ，也就是粒子在“速度空间的数密度”，即每单位速度区间有多少粒子，速度位于 $\mathbf{v}$ 到 $\mathbf{v} + d\mathbf{v}$ 区间的粒子数为 $f(\mathbf{v}) d^3v = f(\mathbf{v}) dv_1 dv_2 dv_3$ ， $\int_{\infty} f(\mathbf{v}) d^3v = \mathcal{N}$ ，同样，后文更经常使用一个粒子处于速度 $\mathbf{v}$ 附近的概率密度 $f(\mathbf{v})/\mathcal{N}$ 。

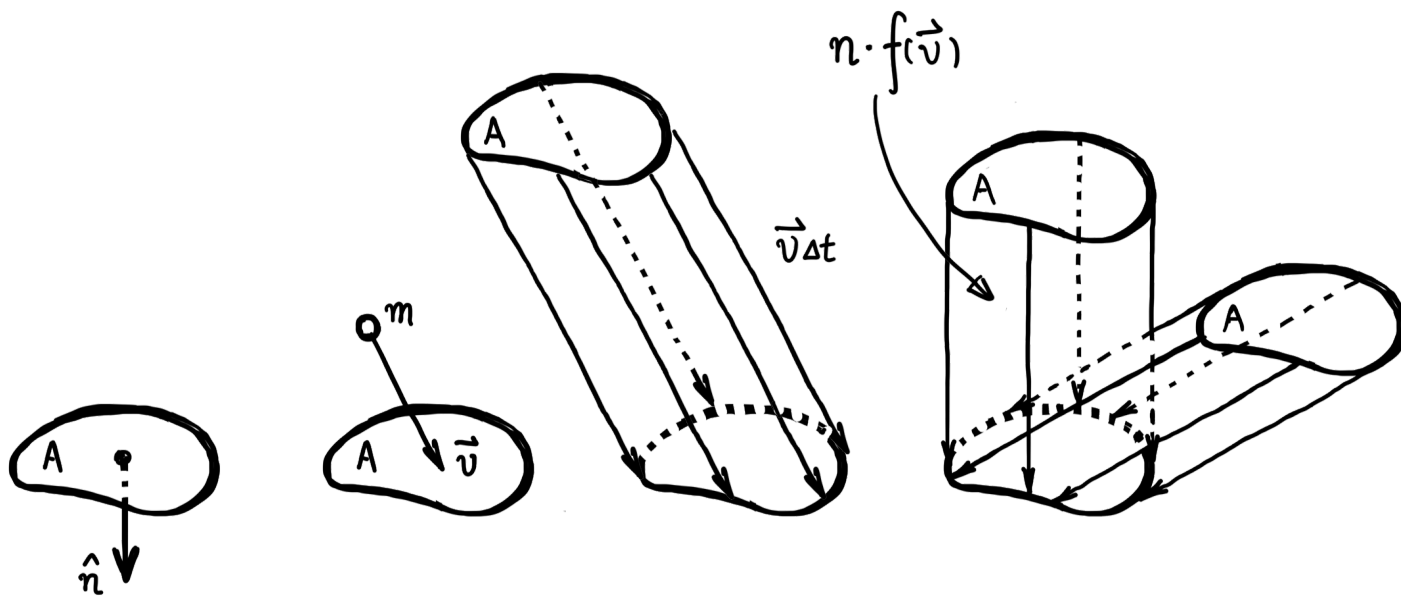
- 假设 V 足够大（相对于实际粒子的尺寸），以至于可以忽略边界的影响（若是，比如说，容器边界不对称，可能影响粒子的分布函数）；如果 V 很大，对一个粒子来说近乎无边无际，处于气体内部还是靠近边界并无区别，则可以认为粒子在系统中不同位置并无区别—— $n(\mathbf{x})$ 就是个常数，即均匀密度 $n = \mathcal{N}/V$ ；同时假设粒子向不同方向运动亦无区别，没有任何特殊方向，于是 $f(\mathbf{v})$ 只依赖于 $|\mathbf{v}|$ 或 $\mathbf{v}^2$ 。
- 假设空间分布与速度分布是独立的，换句话说，其联合分布（一个粒子既在位置 $\mathbf{x}$ 又具有速度 $\mathbf{v}$ 的概率密度）就是 $(n(\mathbf{x})/\mathcal{N}) \cdot (f(\mathbf{v})/\mathcal{N})$ 乘积¹³；也可以说成，处在任何位置的气体粒子其速度分布都一样；空间均匀分布的粒子系统，单位空间体积单位速度区间的粒子数（密度）为 $n f(\mathbf{v})/\mathcal{N} = (1/V)f(\mathbf{v})$ 。
- 假设气体足够稀薄，即 $n = \mathcal{N}/V$ 不大，以至于质点之间除了偶尔碰撞几乎总是相隔很远（相对于粒子尺寸），这样由运动碰撞产生压强的基本模型图景才合理（而不需要像Newton量纲分析中那样考虑其他可能对压强造成影响的粒子间相互作用力）。
- 假设一切都是稳定的，也就是系统各种量都不随时间变化（分布函数都不含时间），热力学中称之为热平衡状态。

现在可以开始计算了（参见译者图 1.2）。在气体中任意指定一小块平面 $\hat{\mathbf{n}}A$ ，其中 A 为面积， $\hat{\mathbf{n}}$ 为其单位法向量，我们将计算粒子们穿过这个面造成的压强：一小段时间 Δt 内，粒子们从各个方向以各种速率穿过 $\hat{\mathbf{n}}A$ ，形成动量流；我们要计算的是 Δt 时间内通过 $\hat{\mathbf{n}}A$ 的总动量 $\Delta \mathbf{P}_{\text{net}}$ （即从 $\hat{\mathbf{n}}A$ 一侧转移到另一侧的总动量）¹⁴。然后就简单了，单位时间转移的动量其法向分量就是压力，而单位面积的压力就是压强。

考虑一个速度为 $\mathbf{v}$ 的粒子，单独它穿过 $\hat{\mathbf{n}}A$ 而造成通过 $\hat{\mathbf{n}}A$ 的动量转移为 $m\mathbf{v}$ ，其法向分量为 $m\mathbf{v} \cdot \hat{\mathbf{n}}$ 。而 Δt 时间内，所有以 $\hat{\mathbf{n}}A$ 为底（底面边界称为Directrix）， $\mathbf{v}\Delta t$ 为母线（Generatrix）的“斜柱体”（ $\mathbf{v}\Delta t$ 围绕底面边界平行移动一圈形成的立体，其体积为 $\mathbf{v} \cdot \hat{\mathbf{n}}A\Delta t$ ）中速度为 $\mathbf{v}$ 的

¹³ 译者注：请注意符号习惯。粒子分布的速度部分，我们只在本章临时用 $f(\mathbf{v})$ 表示粒子数的速度分布（符号上对标位置分布 $n(\mathbf{x})$ ），并且不需要（暂时也得不到）其确切数学形式。下一章及之后，符号 $f(\mathbf{v})$ 一般都用来直接表示粒子速度的概率分布，并且会详细讨论和使用其具体数学形式。而空间部分， $n(\mathbf{x})$ 则始终表示粒子数的位置分布，如果要表示粒子位置的概率分布，则另寻其他符号。数分布还是概率分布在我们当前处理的粒子模型中区别只在归一化条件：数分布的积分（“总数”）等于总粒子数，而概率分布的积分（“总概率”）则等于1，数分布除以总数就是概率分布。

¹⁴ 译者注：英文字母只有26个，但是物理量远多于26个，同一个字母难免要被重复用。我们这里用小写 p 表示压强（与原书一致），大写 $\mathbf{P}$ 表示动量向量。



译者图 1.2: $\hat{n}A$ 处于气体内部任意位置, 粒子穿过则转移动量, 柱体体积 $\mathbf{v}\Delta t \cdot \hat{n}A$ 内速度为 $\mathbf{v}$ 的粒子数同时正比于空间分布密度 $n(\mathbf{x}) = n$ (均匀分布) 和速度分布密度 $f(\mathbf{v}) = g(|\mathbf{v}^2|)$ (各向同性); 倘若二个分布不是这样简单, 可以想像, 一般来说压强将依赖于小平面对气体内部的位置 $\mathbf{x}$ 和其单位法向量 $\hat{n}$ 的具体空间指向。

粒子都穿过了 $\hat{n}A$, 每个粒子对法向动量转移的贡献 $m\mathbf{v} \cdot \hat{n}$ 都一样¹⁵。这块体积中的粒子总数为 $n\mathbf{v} \cdot \hat{n}A\Delta t$, 但其中只有一部分是速度为 $\mathbf{v}$ 的粒子。借助前边提到的速度分布函数 $f(\mathbf{v})$, 整个气体系统所有粒子中速度为 $\mathbf{v}$ 的粒子 (连续分布对应的是速度位于 $\mathbf{v}$ 到 $\mathbf{v} + d\mathbf{v}$ 的粒子数) 占比为 $f(\mathbf{v})d^3v/\mathcal{N}$; 在这个斜柱体筒中也是同一比例, 于是穿过 $\hat{n}A$ 速度为 $\mathbf{v}$ 的粒子数为斜柱体中粒子数乘以该比例 $n\mathbf{v} \cdot \hat{n}A\Delta t f(\mathbf{v})d^3v/\mathcal{N}$, 再乘上每个粒子带来的动量贡献, 就得到全部速度为 $\mathbf{v}$ 的粒子们造成的法向动量转移: $nm(\mathbf{v} \cdot \hat{n})^2 A\Delta t (f(\mathbf{v})/\mathcal{N})d^3v$ 。

当然粒子速度并不都一样, 不过对每个 $\mathbf{v}$, 表达式 $nm(\mathbf{v} \cdot \hat{n})^2 A\Delta t (f(\mathbf{v})/\mathcal{N})d^3v$ 都是通用的: 都是 $n\mathbf{v} \cdot \hat{n}A\Delta t (f(\mathbf{v})/\mathcal{N})d^3v$ 这么多粒子以相同速度 $\mathbf{v}$ 穿过 $\hat{n}A$, 每个粒子带来相同的 $m\mathbf{v} \cdot \hat{n}$ 这么多法向动量转移。现在我们需要把 Δt 时间内所有能通过 $\hat{n}A$ 的粒子都算上, 也就是把各种速度 $\mathbf{v}$ 对应的贡献都加起来, 流过 $\hat{n}A$ 的总动量转移其法向分量为

$$\hat{n} \cdot \Delta \mathbf{P}_{\text{net}} = \int_{\infty} nm(\mathbf{v} \cdot \hat{n})^2 A\Delta t \frac{f(\mathbf{v})}{\mathcal{N}} d^3v = nmA\Delta t \frac{1}{\mathcal{N}} \int_{\infty} (\hat{n} \cdot \mathbf{v})^2 f(\mathbf{v}) d^3v \quad (1.1.a)$$

虽然我们不知道 $f(\mathbf{v})$ 到底长什么样, 但总可以借助它将各种依赖速度的量写成该量的平均值, 比如速度 $\mathbf{v}$ 的平均值 $\langle \mathbf{v} \rangle = \frac{1}{\mathcal{N}} \int_{\infty} \mathbf{v} f(\mathbf{v}) d^3v$ 。类似的, $\int_{\infty} (\hat{n} \cdot \mathbf{v})^2 f(\mathbf{v}) d^3v = \mathcal{N} \langle (\hat{n} \cdot \mathbf{v})^2 \rangle$, 因而 (1.1.a) 可以写作

$$\hat{n} \cdot \Delta \mathbf{P}_{\text{net}} = nmA\Delta t \langle (\hat{n} \cdot \mathbf{v})^2 \rangle \quad (1.1.b)$$

¹⁵ 译者注: 注意体积的表达式 $\mathbf{v} \cdot \hat{n}A\Delta t$ 是有正负的, 而法向动量转移也有 $m\mathbf{v} \cdot \hat{n}$ 与之相一致的正负号, 出现在动量转移表达式中的二者乘积总是非负的, 因为因子 $(\mathbf{v} \cdot \hat{n})^2 \geq 0$ 。

接下来就是最关键的， $\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle$ 要改为用 $\langle\mathbf{v}^2\rangle$ 表示，前者是粒子速度法向分量的平方平均，而后者直接是粒子速度的平方平均。要去掉 $\hat{\mathbf{n}}$ 这个方向，就得回到我们的假设：所有的方向都相同（等价），无论 $\hat{\mathbf{n}}$ 指向何方，或者说 $\hat{\mathbf{n}}A$ 在空间中姿态如何， $\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle$ 都是同一个数值。具体而言常用如下二种做法：

- 如Clausius的方法，分别考虑 $\hat{\mathbf{n}} = \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ ，（三维空间）三个正交基底方向。既然所有方向速度分量的平方平均值都一样，就有 $\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle = \langle(\hat{\mathbf{e}}_1 \cdot \mathbf{v})^2\rangle = \langle(\hat{\mathbf{e}}_2 \cdot \mathbf{v})^2\rangle = \langle(\hat{\mathbf{e}}_3 \cdot \mathbf{v})^2\rangle$ 。而三个方向加起来就得到

$$3\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle = \langle(\hat{\mathbf{e}}_1 \cdot \mathbf{v})^2\rangle + \langle(\hat{\mathbf{e}}_2 \cdot \mathbf{v})^2\rangle + \langle(\hat{\mathbf{e}}_3 \cdot \mathbf{v})^2\rangle = \langle(\hat{\mathbf{e}}_1 \cdot \mathbf{v})^2 + (\hat{\mathbf{e}}_2 \cdot \mathbf{v})^2 + (\hat{\mathbf{e}}_3 \cdot \mathbf{v})^2\rangle = \langle\mathbf{v}^2\rangle$$

- 另一种做法是注意到，同一个假设也表现在速度分布函数“各向同性”，即 $f(\mathbf{v})$ 只依赖于 $|\mathbf{v}|^2$ 。不妨记 $f(\mathbf{v}) = g(|\mathbf{v}|^2)$ ，速度 $\mathbf{v} = |\mathbf{v}|\hat{\mathbf{e}}$ ，其中 $\hat{\mathbf{e}}$ 是速度方向的单位向量；于是 $\hat{\mathbf{n}} \cdot \mathbf{v} = |\mathbf{v}|\hat{\mathbf{n}} \cdot \hat{\mathbf{e}}$ ， $d^3v = |\mathbf{v}|^2 d|\mathbf{v}| d^2\hat{\mathbf{e}}$ ，其中 $d^2\hat{\mathbf{e}}$ 是围绕 $\hat{\mathbf{e}}$ 方向的微分立体角¹⁶。借助这套符号，粒子速度法向分量的平方平均就可以表达成

$$\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle = \frac{1}{\mathcal{N}} \int_{\infty} (\hat{\mathbf{n}} \cdot \mathbf{v})^2 f(\mathbf{v}) d^3v = \frac{1}{\mathcal{N}} \int_{\infty} |\mathbf{v}|^4 g(|\mathbf{v}|^2) d|\mathbf{v}| \int_{\mathbb{S}^2} (\hat{\mathbf{n}} \cdot \hat{\mathbf{e}})^2 d^2\hat{\mathbf{e}}$$

其中单位球面上的积分 $\int_{\mathbb{S}^2} (\hat{\mathbf{n}} \cdot \hat{\mathbf{e}})^2 d^2\hat{\mathbf{e}} = \int_0^\pi \cos^2 \theta \sin \theta d\theta \int_0^{2\pi} d\phi = 4\pi/3$ ，球面坐标以 $\hat{\mathbf{n}}$ 为极轴， $\hat{\mathbf{n}} \cdot \hat{\mathbf{e}} = \cos \theta$ ，习惯称 θ 为极角（polar angle）， ϕ 为围绕极轴 $\hat{\mathbf{n}}$ 一圈的方位角（azimuthal angle）。比较而言，粒子速度的平方平均

$$\langle\mathbf{v}^2\rangle = \frac{1}{\mathcal{N}} \int_{\infty} \mathbf{v}^2 f(\mathbf{v}) d^3v = \frac{1}{\mathcal{N}} \int_{\infty} |\mathbf{v}|^4 g(|\mathbf{v}|^2) d|\mathbf{v}| \int_{\mathbb{S}^2} d^2\hat{\mathbf{e}}$$

其中单位球面上的积分 $\int_{\mathbb{S}^2} d^2\hat{\mathbf{e}} = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi = 4\pi$ ，其余因子都一样。二者对照得到 $\langle\mathbf{v}^2\rangle = 3\langle(\hat{\mathbf{n}} \cdot \mathbf{v})^2\rangle$ ， $\forall \hat{\mathbf{n}}$ 。读者可能已经注意到，（1.1.a）中的积分和此处单位球面上的积分都是恒正的，也就是说我们计算的动量转移法向分量总是正的（前一个脚注处有说明），无论粒子从哪个方向穿过 $\hat{\mathbf{n}}A$ ；初看起来这不正确，因为相对于法向量 $\hat{\mathbf{n}}$ ，动量流总是一半正一半负，才能保证不打破平衡态（粒子以各种速度穿过 $\hat{\mathbf{n}}A$ 的净动量转移一定是零）。其实我们偷懒了，想象真的有个现实的物理的小平面 $\hat{\mathbf{n}}A$ 放在气体中阻挡粒子时，其二侧粒子都会（完全弹性地）弹回去，反弹的法向动量变化是加倍的（单个粒子穿过 $\hat{\mathbf{n}}A$ ，动量转移算1

¹⁶ 译者注：3维空间的向量微分，我们默认使用如下记号习惯： $d\mathbf{x}$ 一般表示全微分，比如单参数曲线 $\mathbf{x}(t)$ ， $d\mathbf{x} = \mathbf{x}'(t) dt$ ，双参数曲面 $\mathbf{x}(u, v)$ ， $d\mathbf{x} = \partial_u \mathbf{x} du + \partial_v \mathbf{x} dv$ 等。 $d^2\mathbf{x}$ 表示在曲面上积分的面积元，如双参数曲面 $\mathbf{x}(u, v)$ ， $d^2\mathbf{x} = |\partial_u \mathbf{x} \times \partial_v \mathbf{x}| du dv$ 。单位向量 $\hat{\mathbf{e}}$ 长度固定（起点在原点的话，端点处在单位球面 $\mathbb{S}^2$ 上），只有2个独立分量（比如球面经纬度），于是 $d^2\hat{\mathbf{e}}$ 表示单位球面上的面积元，也就是微分立体角，建立球面坐标系可以写成 $d^2\hat{\mathbf{e}} = \sin \theta d\theta d\phi$ 。 $d^3\mathbf{x}$ 表示区域体积元，此时理应有3个独立参数，比如Descartes坐标 (x_1, x_2, x_3) 于是 $d^3\mathbf{x} = dx_1 dx_2 dx_3$ 。

份，其反弹的动量转移就是2份）。粒子在一侧的反弹就等价于从另一侧穿过，于是我们本来要计算粒子从一侧反弹的法向动量变化，就等价于计算从二侧穿过的法向动量变化之和（都算正的），这就是我们偷懒的计算。不满意的读者不妨先计算半个球面积分（单侧穿过或单侧入射撞击）再乘2倍（反弹），完全一样。

无论何种方法最终都给出 $\langle (\hat{\mathbf{n}} \cdot \mathbf{v})^2 \rangle = \langle \mathbf{v}^2 \rangle / 3$ ，即在各向同性前提下，粒子速度任意一个分量的方均值都是整个速度方均值的1/3，代入 (1.1.b) 也就推出了压强与密度的关系，即 (1.1.3) 式

$$p = \frac{1}{A} \frac{\Delta \mathbf{P}_{\text{net}}}{\Delta t} = nm \langle (\hat{\mathbf{n}} \cdot \mathbf{v})^2 \rangle = \frac{1}{3} nm \langle \mathbf{v}^2 \rangle$$

1.2 化学

元素

所有物质都由有限数量的元素组成的这一观点可以追溯至关于物质本质的最早推测。起初，在 Socrates（苏格拉底）之前的一个世纪里¹⁷，有设想认为只有一种元素：水（Thales（泰利斯）），气（Anaximenes（阿那克西门尼）），火（Heraclitus（赫拉克利特）），或土（也许是 Xenophanes（色诺芬））。公元前450年左右，Empedocles of Acragas（恩培多克勒斯，Acragas是现代的Agrigento（阿格里真托），意大利市镇）提出了四元素说。在《自然论》（*On Nature*）中，他将元素辨别为“火、水、土和无尽高度之气¹⁸”。中国典籍列举了五种元素：水、火、土、木、金。

像原子理论一样，这些早期关于元素的假说并未伴随任何证据表明它们真的就是元素，也未提出关于如何获取这类证据的任何建议。Plato（柏拉图）在《蒂迈欧篇》（*Timaeus*）中甚至更加坚定地提出，元素之间的差异源于组成这些元素的原子的形状：土的原子是微小的立方体，而火、气和水的原子则是其他正多面体——分别是由4、8或20个全等的正多边形构成的立体，每个立体的每条边和每个顶点都与该立体其他边或顶点相同。

¹⁷ 译者注：Socrates (约 470–399 BC) 之前一个世纪，大约指公元前550–450年。哲学在公元前6世纪于古希腊出现，哲学史有所谓“前苏格拉底时期”，持续了大约二个世纪。

¹⁸ 译者注：英文中，四元素名中的土即earth（大地），气即air（空气）。

历史上很长时期，人们并不很区分作为纯粹数学（思维构造）的几何学和作为物理空间模型的几何学，我们生活的空间就被认为是三维欧几里得空间（3D Euclidean Space），其中只存在五种凸正多面体。Plato在约公元前360年的对话《蒂迈欧篇》（*Timaeus*）将四元素与其中四种正多面体关联：土与立方体，气与八面体，水与二十面体，火与四面体。对第五种正多面体——十二面体——柏拉图模糊地提到：“……神用其来安排整个天球上的星座。”Plato有一种从“审美”出发，基于理想（ideal）世界来建立关于现实的理论的倾向。Plato所谓的理型（Ideal Forms）被认为是独立于现实而完美存在的，满足完美的规律，理型的世界包含数学。我们此处不深入探讨Plato的哲学体系，而只是说明一下，Plato更倾向于认为（不完美的）现实要靠向（完美的）理论，而不是反过来；后世的科学标准则要求以可观测的现实为标准，理论要靠向现实。这里Plato想要通过一个漂亮的假设（每一种正多面体对应一种元素）来理解组成现实的元素——即使他并没有能像识别其他四种元素一样有把握地识别出第五种元素，但是基于第五种正多面体的存在，他也认为第五种元素应该存在。其弟子Aristotle（亚里士多德）增加了第五种元素——以太（拉丁文aether，英文ether），并假定天体由这种元素构成，但他并不关心将其与Plato的第五种正多面体对应起来。

到中世纪末期，这一元素列表已经显得难以说通。很难将任何特定的泥土样本认定为土元素，而火更像一个过程而非一种物质。炼金术士将元素数量缩减到仅有三种：汞、硫磺和盐。

现代化学大约始于十八世纪末，伴随着Joseph Priestley（约瑟夫·普里斯特利，1733–1804）、Henry Cavendish（亨利·卡文迪什，1743–1810）、Antoine Lavoisier（安托万·拉瓦锡，1743–1794）等人的严谨实验。至1787年，Lavoisier已经列出了55种元素。气的地位被多种气体取代：氢气、氧气和氮气；空气被辨别为氮气和氧气的混合物。元素列表中还有其他非金属：硫、碳和磷，以及一些常见金属：铁、铜、锡、铅、银、金、汞。Lavoisier还将一些我们现在知道是紧密结合的化合物列为元素：石灰、苏打和草木灰¹⁹。列表中还包括热和光，它们当然根本不是物质²⁰。

结合质量定律

第一个使用化学方法提供原子定量信息的人是John Dalton（约翰·道尔顿，1766–1844）——一个贫困织布工的儿子。他从1802到1804年的实验室笔记描述了对结合成化合物的元素质量的细致测量。他发现这些质量总是呈固定的比例。例如，他发现当氢气在氧气中燃烧时，1克

¹⁹ 译者注：原文中lime, soda, and potash，是其非化学术语的通俗名称，主要成分分别是氧化钙（石灰）、碳酸钠（苏打），和钾碱（草木灰）。

²⁰ 译者注：原文用词是substance，常指“组成东西的物理实体或化学物质”。热与光是否属“物质”当然取决于物质的定义，以今天的语言，不妨理解为：热与光并非原子分子类物质（atomic matter）。

氢气与5.5克氧气结合，产生6.5克水，没有任何剩余。在每个水粒子由一个氢原子和一个氧原子组成的假设下，一个氧原子质量必须是一个氢原子质量的5.5倍。

我们会看到，水很快被发现是H₂O：二个氢原子比一个氧原子。倘若Dalton知道这一点，他就会作出结论，一个氧原子质量是二个氢原子质量的5.5倍，即一个氢原子质量的11倍。当然，后来更准确的测量发现1克氢气与大约8克氧气结合，因此一个氧原子质量是二个氢原子质量的8倍，或者说一个氢原子质量的16倍。原子量很快被定义为与一个氢原子质量相比的原子质量²¹，因此氧的原子量是16。（这只是近似值。碳原子最常见的同位素的原子量被定义为精确的12；根据这一定义，氢和氧最常见同位素的原子量分别被测量为1.007825和15.99491。）

下表比较了Dalton假设的几种常见化合物的分子式与正确分子式：

化合物	Dalton的分子式	正确分子式
水	HO	H ₂ O
二氧化碳	CO ₂	CO ₂
氨	NH	NH ₃
硫酸	SO ₂	H ₂ SO ₄

以下是几种元素真实原子量的近似值，Dalton推测的值，以及Dalton如果知道真实化学分子式将会计算出的值（以星号标注的列）：

元素	真实值	Dalton	Dalton*
H	1	1	1
C	12	4.3	8.6
N	14	4.2	12.6
O	16	5.5	11
S	32	14.4	57.6

要在测量原子量方面取得进展，显然需要找到一种方法来推算各种化学化合物的正确分子式。气体化学反应的研究提供了一种方法。

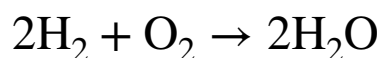
²¹ 译者注：原文用了atomic weight原子量，即relative atomic mass相对原子质量。

体积结合定律

1808年12月31日，Gay-Lussac（盖-吕萨克）在巴黎的哲学学会（Societe Philomathique）²²宣读了一篇论文，宣布他观察发现气体在相同温度和压强下，总是以特定的**体积**比例相结合。例如，二升氢气与一升氧气结合生成水蒸气，没有剩余的氢气或氧气。类似的，一升氮气与三升氢气结合生成氨气，也没有任何气体剩余，依此类推。

对这一实验结果的正确解释是由都灵（Turin，意大利城市）的Count Amadeo Avogadro（阿梅迪奥·阿伏伽德罗伯爵，1776–1856）在1811年给出的。阿伏伽德罗原理（*Avogadro's principle*）²³说，在相同温度和压强下，相同体积的气体总是含有相同数量的气体粒子。这些气体粒子被Avogadro称为“分子”，可以就是单个原子，也可以由相同或不同元素的多个原子连接在一起而成。水蒸气是由一倍体积的氧气与二倍体积的氢气形成这一观察表明，根据Avogadro原理，水分子是由氧分子及其二倍数量的氢分子形成的——这与Dalton曾经假设得不同。

数据中还有更令人惊讶的发现：二升氢气与一升氧气结合生成的不是一升，而是**二升**水蒸气。如果氧气和氢气分子由单个原子组成，而水分子由二个氢原子和一个氧原子组成，那么这就不是预期结果。预期情况下二升氢气加上一升氧气应该产生**一升**水蒸气。Avogadro可以得出结论，假如水分子的确包含二个氢原子和一个氧原子（这在整套的解释中说得通），那么氧气和氢气分子本身必须各包含二个原子。也就是说，将水分子视为H₂O，则生成水分子的反应是：



Avogadro原理的使用迅速为诸如CO₂、NH₃、NO等气体提供了正确的化学式。知道了这些化学式并测量参与各种反应的气体**质量**，就可以修正Dalton的原子量，并计算出气体分子中各原子的更可靠的原子量数值——可相对于其中任何一种原子而计量。以氢的原子量作为1个单位基准，这些计算得出的原子量为：碳约12，氮约14，氧约16，硫约32，等等。随后，凭借这些已知原子量，测量多种化学反应中元素结合的质量，可以得出许多其他元素的原子量，不仅限于通常存在于气体中的元素。

我们不妨梳理一下写出氢气分子、氧气分子和水分子的化学式以及氢原子和氧原子的相对质量所动用的假设和逻辑：

1. 只用元素质量守恒的假设（以及假设氢气只包含氢元素，氧气只包含氧元素，生成物水只包含这两者），如果有精准的质量测量（生成的水的质量和参与了反应的氢气、氧气的质

²² 译者注：“Science（科学）”一词直到十九世纪中叶才被广泛应用作今天其常见的含义：“自然及物理科学”。更早一段时期，几乎任何系统性、理解世界的理论一般都被归在“Philosophy（哲学）”一词下。

²³ 译者注：或许“阿伏伽德罗猜想”是更恰当的说法，这一猜想的理论依据要到气体分子运动论（Kinetic Theory of Gas）出现才被解释清楚，我们后文会看到。

量)，我们就可以确定生成物水中氢元素和氧元素的质量比例（也就是总氢原子和总氧原子质量比例）为 1:8。

2. 在此基础上，我们再假设物质以分子/原子的形式存在，氧气、氢气和水中各只包含一种分子/原子。如果每一个分子都由整数个固定质量的原子组成，结合元素质量守恒的假设，这相当于说每一个元素的原子数是守恒的。因此我们要求 $xH_a + yO_b = zH_cO_d$ ，其中 x, y, z, a, b, c, d 均为正整数。
3. 动用Avogadro的假设我们至少可以得到第一组系数 x, y, z ：固定气压和温度下，同体积的气体中分子数是一样的——不管每种分子是由什么样的原子构成的。于是只要有了参与反应的氢气和氧气以及生成的水蒸气的体积之比，也就有了三者的分子个数之比：每二个氢气分子和每一个氧气分子结合，生成二个水分子。于是立刻有： $2H_a + O_b = 2H_cO_d$ 。
4. 再通过元素原子数守恒，我们可以得到 $2a = 2c, b = 2d$ 。由于 a, b, c, d 都是正整数，则 $c = a \geq 1, b = 2d \geq 2$ ——对氢气分子和水分子的组成成分我们没有得到什么有效信息，但是足以判断氧气分子应该含有偶数个原子，至少是双原子。如取最小整数解： $a = c = d = 1, b = 2$ ；次小整数解： $a = c = 2, d = 1, b = 2$ 。
5. 要确定氢气分子中的原子数（同时也就确定了水分子中的氢原子数），还需要氢气在不同环境中的表现。Sir Humphry Davy（亨弗里·戴维爵士，下文讲到电解会出场）和Joseph Louis Gay-Lussac（约瑟夫·路易·盖-吕萨克）在19世纪初进行了一系列实验，表明氢气也至少是双原子。例如，1升氢气 + 1升氯气 = 2升氯化氢，按Avogadro原理转为分子数： $H_a + Cl_b = 2H_cCl_d$ ，得到 $a = 2c \geq 2$ ，同样表明氢气至少是双原子的。
6. 也就是说，现在我们能够确定到氢和氧都至少是双原子（一定是偶数个原子的），但具体数字并不确定，而且二个值仿佛是彼此独立的（认为氧气双原子、氢气四原子并不违背已知信息）。要完全确定下来两种气体都是双原子，还需要等几十年汇集更多信息，如分子运动论将给出微观的第一性原理解释——下一章会涉及到一些。其中一个关键的预测结果来自下一章2.4节会详加论证的能均分定理（Equipartition Theorem）：气体的总能量与其总自由度成正比（单原子气体仅具有平动自由度，而双原子气体还包含转动和振动自由度……），而这会导致平均每分子的热容量不同。氢气和氧气分子究竟是双原子还是包含更多偶数个原子，原则上应该能由测量其热容量判断。然而，下一章会看到，实验测量结果却又受到量子效应的影响——这就得等到二十世纪了。因此，我们在这里只好先接受氢分子和氧分子确实都是双原子的事实。
7. 我们此时还没有用到化学反应中元素质量比的信息：1克氢气 + 8克氧气 = 9克水蒸气。也就是说：（质量 m 以下标 m_H, m_O 等标记各种原子质量）

$$\frac{2am_H}{1} = \frac{bm_O}{8} = \frac{2cm_H + 2dm_O}{9}$$

于是

$$2a = \frac{b}{8}\mu = \frac{2c}{9} + \frac{2d}{9}\mu$$

其中 μ 为氧原子相对于氢原子的质量（即氧原子量）。将假设氢气分子和氧气分子都是双原子的正确解代入

$$4 = \frac{2}{8}\mu = \frac{4}{9} + \frac{2}{9}\mu$$

得到 $\mu = 16$ 。类似方法可以计算各种气体分子中各类原子的相对原子质量。

气体常数

如上一节所述，1857年Clausius（克劳修斯）证明了：对于每单位体积由 n 个质量为 m 、速度平方平均值为 $\langle v^2 \rangle$ 的粒子组成的气体，其压强为 $p = nm\langle v^2 \rangle/3$ 。利用Charles定律，他得出结论： $\langle v^2 \rangle$ 与绝对温度成正比。将这一关系写为 $m\langle v^2 \rangle/3 = kT$ ，其中 k 为某个常数，便可得出方程（1.1.5）， $p = nkT$ 。但这一结论本身并没告诉我们 k 如何随不同气体而变化。Avogadro原理回答了这个问题。对于体积 V 中的 N 个粒子，其粒子数密度为 $n = N/V$ ，因此公式（1.1.5）可改写为：

$$pV = NkT \quad (1.2.1)$$

如果按Avogadro所述，在给定压强、体积和温度下，任何气体的分子数均相同，那么 $k = pV/NT$ 一定对任何气体都相同。然而，Clausius并未下这一结论，可能因为当时尚不知道Avogadro原理有何理论基础。常数 k 的普适性——以及Avogadro原理——后来通过分子运动论（Kinetic Theory）获得了解释，这将在下一章中讨论。常数 k 后被称为玻尔兹曼常数（*Boltzmann's constant*），以纪念Ludwig Boltzmann（路德维希·玻尔兹曼），我们将看到，他是分子运动论的主要创立者之一。

化合物的分子量 μ 被定义为其单个分子中各原子的原子量之和。分子的实际质量 m 等于其分子量乘以具有单位原子量的一个假想原子的质量 m_1 ：

$$m = \mu m_1 \quad (1.2.2)$$

在现代原子量体系中，碳最常见的同位素的原子量被精确定义为 12， $m_1 = 1.660539 \times 10^{-24}$ g，这在Avogadro的时代当然尚未知晓。一份质量 M 包含 $N = M/m = M/m_1\mu$ 个分子，因此理想气体定律 (1.2.1) 可写为：

$$pV = MkT/m_1\mu = (M/\mu)RT \quad (1.2.3)$$

其中 R 为气体常数：

$$R = k/m_1 \quad (1.2.4)$$

19世纪早期的物理学家可以利用公式 (1.2.3) 测量 R ，他们得到的数值与现代值 $R = 8.314$ J/K 非常接近。如果 k 已知，凭借 R 便可确定 m_1 ，进而计算所有已知原子量的原子质量。然而，要等到2.6节所描述的发展， k 才得以确知。

AVOGADRO 常数

顺便一提，一摩尔 (mole) 分子量为 μ 的任何元素或化合物被定义为 μ 克，因此公式

(1.2.3) 中的比值 M/μ 以克为单位正好等于气体的摩尔数。由于 $N = M/m_1\mu$ ，一摩尔包含的分子数等于 $1/m_1$ ，其中 m_1 以克为单位。这被称为阿伏伽德罗常数 (Avogadro's number)。当然Avogadro本人并不知道Avogadro常数。现在我们知道每摩尔是 6.02214×10^{23} 个分子，对应于单位分子量 $m_1 = 1.66054 \times 10^{-24}$ 克。Avogadro常数的测量在19世纪末被广泛认为是物理学面临的重大挑战之一。

1.3 电解

早期电学

电在古代世界就已被认识，尤其是我们现在所说的静电。琥珀用毛皮摩擦后，会看到其吸引或排斥轻质材料的小块小片。Plato在《蒂迈欧篇》（*Timaeus*）中提到“关于琥珀吸引的奇迹。”（这就是“electricity（电）”一词的来源；希腊语琥珀一词为“elektron”。）

电在十八世纪开始被科学研究。二种类型的电被区分开来：琥珀棒被毛皮摩擦，“树脂电”（resinous electricity）留于其上；而玻璃棒以丝绸摩擦，“玻璃电”（vitreous electricity）留于其上。不同电荷被发现互相吸引，而相同电荷则互相排斥。Benjamin Franklin（本杰明·富兰克林，1706–1790）给予“玻璃电”和“树脂电”相应的现代称谓：正电和负电。

在1785年，Charles-Augustin de Coulomb（查尔斯-奥古斯丁·德·库仑，1736–1806）研究报告称携带电荷 q_1 和 q_2 相隔距离 r 的二个物体之间作用力 F 为：

$$F = \frac{k_e q_1 q_2}{r^2} \quad (1.3.1)$$

其中， k_e 是一个普适常数。对于相同或不同电荷，乘积 $q_1 q_2$ 分别为正或负，表明是排斥力或吸引力。Coulomb无法实际测量这些电荷，但他通过将带电体接触到一个相同材质和大小的无电物体上，可以使电荷减少一半，并观察到这会使它与任何其他带电物体之间的力也减少一半。现代电荷单位的引入要等到对磁性的定量研究。

Coulomb的做法也可看作一种对称性考虑：对于两个半径相同、材料相同的导体金属小球，假设电荷可以在其上流动，相互接触后没有道理认为电荷会厚此薄彼，而是该平均分配。

截止到现在，关于对称性的考虑我们已经见到了三次：一是在研究气体粒子的运动时，我们认为在没有特殊的边界条件时，粒子在（无穷大）箱子里的位置分布是均匀的；二是同一情景中，我们假设了粒子在各个方向的速度分布不该有什么区别；第三次就是以上电荷的均分。

如果我们观察到一组物体彼此完全相同，或者是他们所在的环境/空间符合某种对称性（如几何对称或现实条件并无差别），则我们认为物理定律不应该对这些物体或者对称的“维度”有所区分。事实上，历史上我们一次又一次地发现关于对称性的假设在绝大多数情况下都符合物理现实，并且对物理理论的基础建构和扩建具有至关重要的作用，甚至经常成为理论推演的起点之一。比如，此处并不比彼处特殊（空间平移对称），此时并不比彼时特殊（时间平移对

称)，此方向并不比彼方向特殊（空间旋转对称），此参照系并不比彼参照系特殊（Galileo或Lorentz对称），等等。

更根本地说，当我们试图理解世界的规律时，必须假设物理定律遵循某种对称性。物理学的任务之一可以说是寻找规律、压缩我们对自然的描述：当A出现时，B必然发生（B可以是概率叙事）。如果在相同的条件A下，B的结果却可以是像有主观意愿一样“任性”/任意的，那么我们便几乎无法建立任何可预测的规律。当我们观察到两个情境在可观测量上完全相同（对称）时，只有不区分这两种情境的物理定律（物理定律对两者一样），我们才能管其叫做“规律”。换句话说，对称性假设是我们构建任何物理知识体系的前提。也因此，如果看到一组看似相同的物体在看似相同的情境下呈现出不同的物理行为，与其认为物理定律本身是无常或不存在的，我们反而会想：这些物体或情境或许并不完全相同，而是存在我们尚未发现的差异。这是对对称性假设的逆否命题：如果物理定律区分两种情境，那么二者就不该是相同/对称的。

或许可以说，“对称性”这一概念本身就与人类对观测的主观认知密切相关：人类所能区分者为异，不能区分者为同。自然/绝对客观上何为同何为异我们不得而知，常有所谓“独立于主观的客观”这个概念提法恐怕自动就将其自身（如 thing-in-itself 物自身或“自在之物”）归属成为了“不可知”；但作为观测者与理解者，我们所总结出的可称之为“规律”的东西，都要遵从我们主观对观测（现象）所提出的对称性要求（“对称性”至少必须在认知的框架之中，似乎是人脑构建世界模型所倚赖的基本框架的一部分，脱离它认识世界似乎就没办法了）——我们判断为一样的情形，就要具有一样的物理。

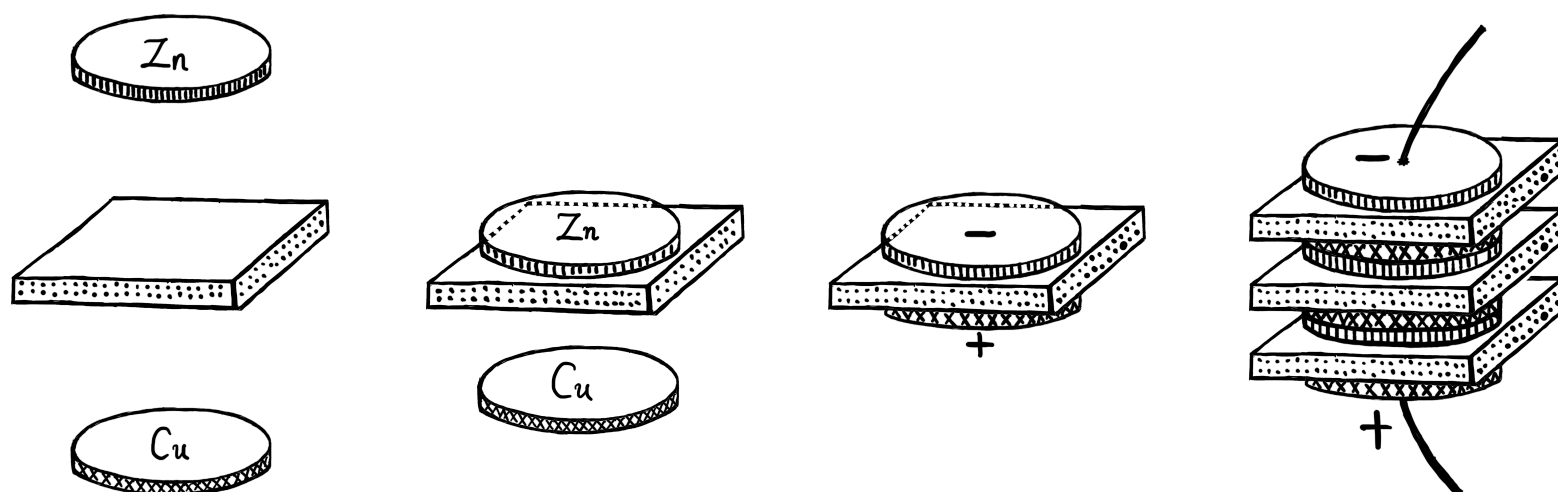
早期磁学

磁在古代世界也为人所知，当时被观察讨论的磁属于我们现在所称的永久磁（Permanent Magnetism）。希腊人知道天然产生的磁石能够吸引或排斥小铁块。Plato在《蒂迈欧篇》（*Timaeus*）一书中提到磁石称之为“赫拉克勒斯之石²⁴”。（今天英文“magnet（磁铁）”一词来自于亚洲小亚细亚的马格尼西亚城（Magnesia），那附近经常有磁石被发现。）

很早期中国人也发现了磁石并将其用作指南针（“寻南石”），用于风水和航海等目的。每块磁石在其一端有一个寻南极²⁵，被吸引向地球南极附近的某个点；另一端则有一个寻北极，吸引向地球北极附近的某个点。磁性最早由William Gilbert（威廉·吉尔伯特，1544–1603）——伊丽莎白一世（Elizabeth I）的宫廷医生——进行科学研究。他观察到，不同磁石的寻南极彼此排斥，寻北极也是如此；而一块磁石的寻南极吸引另一块磁石的寻北极。Gilbert得出结论，磁

²⁴ 译者注：Heracleian stones，赫拉克勒斯是希腊神话中的半神英雄，能勇善战的大力士。

²⁵ 译者注：原文语言并未使用南/北磁极，而是south/north-seeking pole



译者图 1.3：伏打电池一个单元的构成，例如：金属锌 Zn 板，金属铜 Cu 板，夹在二者中的盐水纸板；单元可以多层堆叠构成整个电池，最下端的铜板是电池正极，最上端的锌板是电池负极。

石的一极被拽向北方，另一极被拽向南方，这是因为地球本身就是一个磁体，而磁石的寻南极和寻北极分别指向（更接近于）地球自己作为磁铁的的北极和南极附近。

电磁学

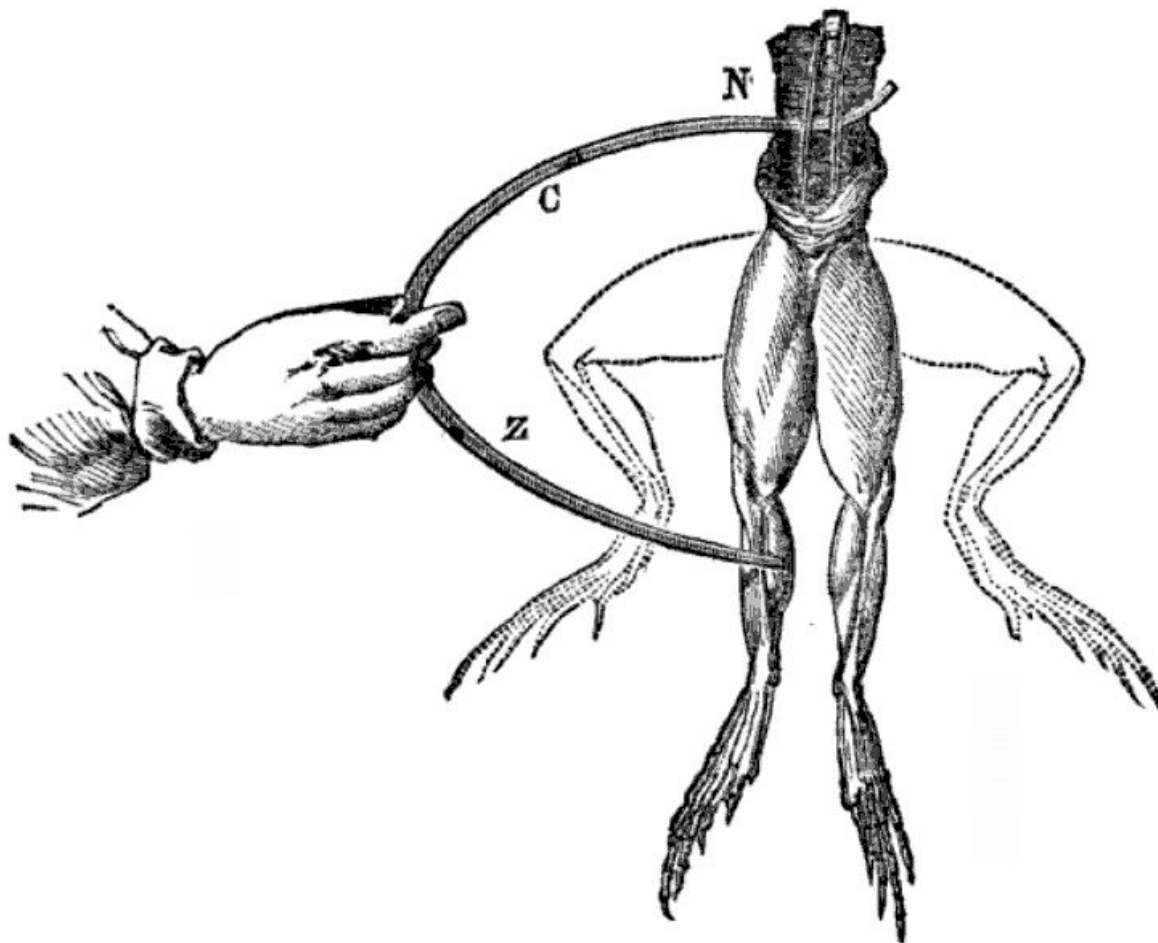
随着1809年Count Alessandro Volta（亚历山德罗·伏打伯爵，1745–1827）发明电池，探索电与磁之间定量关系成为可能。伏打电池由两种不同金属的圆盘堆叠而成，圆盘之间用浸透盐水的硬纸板隔开（参见译者图 1.3）。这种电池驱动稳恒电流通过连接在堆叠两端的导线，其正负极分别对应电流从堆叠流出端和向其流入端。

伏打电池（Voltaic Cell）其实也叫做伽伏尼电池（Galvanic Cell）——Volta是在研究如何反驳Luigi Galvani（路易吉·伽伐尼）提出的生物体发电理论的过程中才发明了电池。

Volta对电学产生兴趣的时代也是神经科学起步的时代：Galvani是意大利物理学家、生物学家。据说他在铺满各种器具、材料的实验室里意外发现（看起来有时候实验室乱一些也有意想不到的好处）：当青蛙腿部肌肉的两端分别和两块不同的金属相连（他用了铜和锌），且保持两块金属相接触时，青蛙腿会收缩（参见译者图 1.4）。Galvani因此得出结论：动物体可以产生“电（Animal Electricity）”（金属刚好可以激发动物的这一功能，但“电”是动物自己产生的），从而控制肌肉的收缩，产生运动。

不过，虽然用了“电”来形容他的发现，Galvani并不认为这种“电”和当时人们所熟悉的、能够被储存的静电荷是根本上相同的物质。Galvani认为生命所产生的“电”是一种直接由神力引导的电，并不能被人为操控。

Volta则认为：哪来什么动物电，所有的电都是同一种。他与Galvani的敌对也部分源自于更根本的立场不同：Volta是启蒙运动积极分子，而这场运动的一个根本观点就是我们不需要神力



译者图 1.4：这是Galvani在青蛙腿部外接的电路，上面接着坐骨神经，下面接着小腿肌肉神经。其中C是铜 Cu，Z是锌 Zn，青蛙腿里的液体相当于电解质，被铜和锌夹在中间，形成一个跟伏打电池异曲同工的电池结构；而金属既是电池的能量来源（通过化学反应），也在这里充当了导线——电子可以通过其形成电路。

来解释现实，自然的规律可以被人类发现和利用。Volta觉得真正产生电的很可能是金属，而不是青蛙腿。青蛙腿抽动仅仅可以证明生物体能对电流做出反应，不能证明生物可以自己产生电，更别提什么特殊的“动物电”。

于是，他做了个实验，用盐水浸过的纸代替青蛙腿连接起来金属，发现同样可以检测到电流流通。他甚至还用自己的舌头代替青蛙腿，让两枚不同材料的硬币一侧相接，另一侧分别夹在舌头上下，发现会产生一种苦涩的味道。

Volta是对的：电的确不来自于青蛙腿，而是来自金属。但是，青蛙腿之所以会动，是因为外界来的电流引发了神经中的动作电位（Action Potential），从而控制了肌肉收缩。生物体的运动行为的确依赖于体内自发产生的电信号，只是在这一个情景当中，电信号来自于外界而不是已经死了的青蛙。有人认为，声誉极高的Volta对Galvani的公开反驳导致许多同时代的科学家彻底排除了动物可以自发产生电信号的可能性，从而或许搁置了对神经电路的正确认识。

（Volta倒是并不否定Galvani的发现，他甚至以Galvani的名字命名了用二个金属产生的电流——Galvanism。）

不过，Volta对电池的理解也是不完全准确的：他以为电的产生仅仅来自于二种不同金属材料本身，但Michael Faraday（迈克尔·法拉第）会最终发现，电流产生的真正原因在于金属和电解质（Electrolyte）交界处发生的化学反应（在Galvani的实验里充当电解质的是青蛙腿上的组

织，Volta的电池中则是夹在中层的盐水纸）——Volta误认为电解质只起到连接作用。很快我们就会在后文中看到更正确的解释。

Volta的电池有史以来第一次提供了稳衡电流。这不仅开启了电被作为能源在工业日常中应用的可能性，也是给了当时的科学家们一个很大的好消息：他们发现，当分别把与电池正负极相连的两个电极（Electrode，经常是金属棒）同时放到水里时，水中会产生氢气和氧气气泡——这极大便利了分解研究化合物的成分。我们后文就会继续讲。

1820年7月，在哥本哈根的Hans Christian Ørsted（汉斯·克里斯蒂安·奥斯特，1777–1851）注意到当电流接通时，附近的指南针发生偏转，从而得出结论，电流对磁铁产生力。反过来，他还发现磁铁对通电流的导线也有作用力。

这些发现几个月后由André-Marie Ampère（安德烈-马里·安培，1775–1836）在巴黎进一步发展，他发现携带电流的导线会相互施加作用力。对于长度为 L 、携带电流（每秒电荷量）分别为 I_1 和 I_2 、相距 $r \ll L$ 的两根平行导线，它们之间的力为：

$$F = \frac{k_m I_1 I_2 L}{r} \quad (1.3.2)$$

其中， k_m 是另一个普适常数。如果电流方向相同，力为排斥力；如果电流方向相反，力为吸引力。一安培（ampere）被如此定义：当 $I_1 = I_2 = 1$ 安培时， $F = 10^{-7} \times L/r$ 牛顿。（即， $k_m \equiv 10^{-7} \text{ N/ampere}^2$ 。）电荷的电磁单位“库伦（coulomb）”被定义为1安培电流在1秒钟内携带传输的电荷量。现代电流表通过观察电流流过线圈产生的磁力来测量电流。

电与磁的联系在1831年由Michael Faraday（迈克尔·法拉第，1791–1867）在伦敦的皇家研究院²⁶进一步加强。他发现变化的磁场会产生电力，可以在导线中驱动电流。这是当今发电产生电流背后的原理。电力其后很快有了重要的实际应用，1831年美国画家Samuel F. B. Morse（塞缪尔·F·B·莫尔斯，1791–1872）发明了电报。

最后，在1870年代，伟大的苏格兰物理学家James Clerk Maxwell（詹姆斯·克拉克·麦克斯韦，1831–1879）证明了电流产生磁场这一方程的一致性²⁷要求变化的电场也产生磁场。特别是，当振荡的磁场产生振荡的电场时，振荡的电场也产生振荡的磁场，因此电场和磁场这一自我维持的振荡可以在看似空无一物的空间中传播。Maxwell计算了这一传播的速率，并发现其等于

²⁶ 译者注：1660年建立的皇家学会（Royal Society）和1799年建立的皇家研究院（Royal Institution）并非同一机构。

²⁷ 译者注：方程的一致性（consistency）指的是在一定前提条件下（如能量守恒，对称性，已经接受的理论等）数学和逻辑的一致性自动要求必然得到特定结论。如果电生磁，而磁不生电，数学上会与已有的结论相矛盾；我们稍后单辟一小节“位移电流”讲述此事。

$\sqrt{2k_e/k_m}$ ²⁸，数值上大致与测得的光速相等，强烈表明光就是电场与磁场中的这种自我维持的振荡。我们在后续章节中，特别是在第四章和第五章中，会看到Maxwell方程的更多内容。

历史上 k_e 和 k_m 这两个常量的值分别是在其对应静电学、静磁学的情景当中测量得到的。如正文所说，把这两个常量放在一起我们可以得到光速 c ——这不仅暗示了光是电磁波，也为后来狭义相对论的理论基础假说——光速不变——做出了一致的预言。

今天，我们已经有了相对论的理论体系，也可以从另一个相反的思路方向出发来理解这三个量之间的关系：结合静电场Gauss定律、电荷的相对论不变性和狭义相对论，可以推出电流一定会产生一种垂直于电流方向的力——这也就是磁力（粗略地说，在一个参照系下的静电力在另一个参照系下被叫做磁力——电和磁更像是同一现象在不同观察者看来人为划分成的不同部分）。从这个情景出发推导无穷长电流的磁场，光速 c 就已经自动在公式当中了。于是从这一角度思考，我们也可以通过 k_e 和 c 这两个量得到 k_m 。感兴趣的读者可以参考Edward Purcell（爱德华·珀塞尔，1912 – 1997）的经典电磁教材《电磁学》（Electricity and Magnetism）第5章The Fields of Moving Charges和第6章The Magnetic Field。

O H M 定 律 *

初级教材经常将1827年发表的欧姆定律（Ohm's Law）放在很靠前的位置，而电流的磁力（1.3.2）呈现次序则靠后。这难免给人造成一种误解：仿佛有了伏打/伽伏尼（Voltaic/Galvanic）电池，不仅可以维系稳恒电流，电解熔融盐（见后文），而且立刻可以建立Ohm定律，似乎测量电流与电势差是唾手可得的事。事实上，电流的磁力（1.3.2）恰恰是可靠的电流测量之基础，有如理想气体温度计对温度的定标。

如同Coulomb定律（1.3.1）中的电量 q ，在理论建立之初几乎没有办法测量（作者在文中也说到，电量的可靠标准同样依赖于能够测量电流，又一次归结到确立磁力与电流的依赖关系（1.3.2））；但凭对称性考虑，原则上可以将带电量对半分而观察力的变化。十分类似，（1.3.2）中电流的磁力正比于电流大小，而如何测量电流 I 变成了困难的焦点。此时对称性又可以派上一些用场：比如完全一样的（实验上差异足够小的）电池与长导线应该产生一样的电流，将这样“全同”的携带电流的直线导线平行排列，彼此足够靠近时相当于若干倍的电流；亦或将一条长导线绕一个大圈折返回旋而形成平行直线，也相当于几倍电流；还可以将完全相同的多条导线连接同一个电池，每条导线似乎应该分配相同的电流，这样也可以彼此靠近而相当于多倍电流。如此就可以测量多倍电流施加于其他电流的磁力如何变化（是否呈现简单倍数关

²⁸ 作者注：该量与电荷单位的选择无关——只要方程（1.3.2）中出现的电流被定义为方程（1.3.1）中电荷的流动变化率，而电荷以相同单位计。该量显然也与所采用的力的单位无关，只要在方程（1.3.1）和（1.3.2）中使用相同的力的单位。

系)，这些方案至少在原则上可行。历史上André-Marie Ampère所设计进行的一系列实验比上述构想精巧得多。一旦确定了磁力正比于电流（1.3.2），立刻就可以用其测量电流，Ohm定律的实验思路正是使用磁力扭秤测量电流，而用正比于电流的磁力测量电流的基本方法一直沿用至今。

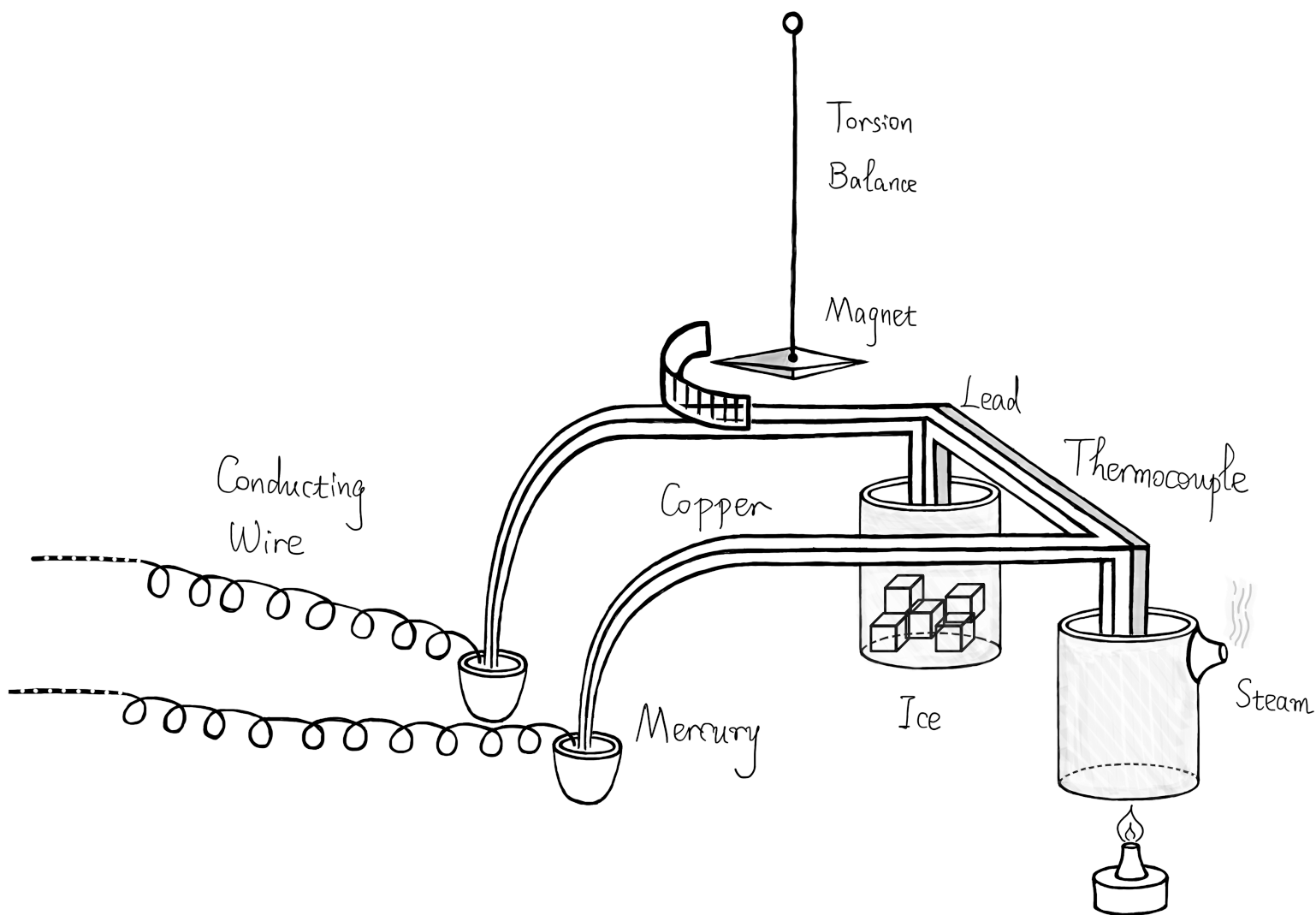
Ohm定律实验还需要一个要素：电源。Ohm在1825 – 1826年间的研究起初采用伏打/伽伏尼

(Voltaic/Galvanic) 电池，而后更换成了此前不久德国的托马斯·约翰·塞贝克 (Thomas Johann Seebeck) 发现的更稳定的热电偶。从1821到1823，Seebeck进行了一系列实验，试图理解Ørsted在1820年的发现“电生磁”。在此期间，1822年Seebeck发现：如果用二种不同金属构成一个闭合电路，并让它们的二个接触点处于不同的温度，这个电路会使磁针发生偏转。用现代语言说，二个接触端点之间的电势差/电压 ΔV 与温度差 ΔT 成正比 $\Delta V = a\Delta T$ ，其中系数 a 依赖于金属类型等参数。从后见之明看这个现象，其实单独一种金属导体跨接在二个不同温度，原则上就会产生电动势。金属中的自由电子如同气体分子，温度高处平均速率更大，于是平均而言电子会向温度低处扩散（逆着温度梯度 ∇T 方向扩散），很快低温处就聚集了更多电子，形成更高的负电荷密度（下一章中我们会详细讨论类似的输运过程，包括扩散）。这样的电荷移动与不均匀的空间分布构建起 $-\nabla T$ 方向的电场 $\mathbf{E} = -\nabla V$ ，使得持续扩散的电子受到反向静电力，相当于扩散的阻碍力量。稳定时（此时金属相当于开路电池）正比于 $-\nabla T$ 的扩散的“动力”与反向的电场力相平衡；平衡也可以说成是，一方面电子由高温向低温扩散，另一方面电子在反向的电场力作用下由高电势能向低电势能流动，二个流动效果抵消。最终 $\nabla V \propto \nabla T$ ，意味着电动势密度正比于温度梯度，因而最终二个温度端点间电动势正比于温度差。但这时用同一种金属连接成回路显然无法产生电流，因为这相当于二个电动势一样的电池平行连接，效果互相抵消。如Seebeck所发现的，二种不同金属每一个都满足电动势 $\mathcal{E} = k\Delta T$ ，但由于金属电子密度差异，电子扩散程度不同等区别，造成最终比例系数 k 不同，相当于二个大小不同的电动势平行连接抵消不掉，于是净余 $\Delta V = \mathcal{E}_1 - \mathcal{E}_2 = (k_1 - k_2)\Delta T = a\Delta T$ 。

有了产生电动势的热电偶和测量电流的磁力扭秤，Ohm的实验就可以成形。参见译者图1.5的示例：如果使用铜和铅二种金属，一端都置于冰水中，另一端都置于热蒸汽中，这个热电偶作为电池可以调节温差改变电动势。将测试导线接到热电偶上形成回路，磁针则置于直线导体附近，有电流时会产生偏角 x 。改变测试导线的长度 l ，其他因素不变，测量磁针偏角 x 如何依赖于导线长度 l 。Georg Simon Ohm（格奥尔格·西蒙·欧姆，1789 – 1854）于1827年发表了如下结果：

$$x = \frac{a}{b + l}$$

其中， a 与热电偶接点的温度有关， b 是整个实验装置的一个常数。用现代的记号，磁针偏角 x 正比于磁力，而后者正比于电流 I ，于是 $x \propto I$ 。测试导线的长度 l 正比于其电阻 $R = \mathcal{R}l$ ，记



译者图 1.5：欧姆定律实验装置示意图，用现代术语描述：二种金属（如铜和铅）相接触，二端放置在不同温度（如一端沸腾蒸汽，另一端冰水）构成热偶，其电动势正比于温度差。扭秤悬挂磁针受到电流磁力而偏转，改变接入电路的导线长度，电流随之变化，偏转角亦随之变化。磁力导致偏转角正比于电流，测量磁针偏转角与导线长度的关系，即电流与电阻的关系；测量磁针偏转角与温度差的关系，即电流与电动势的关系。

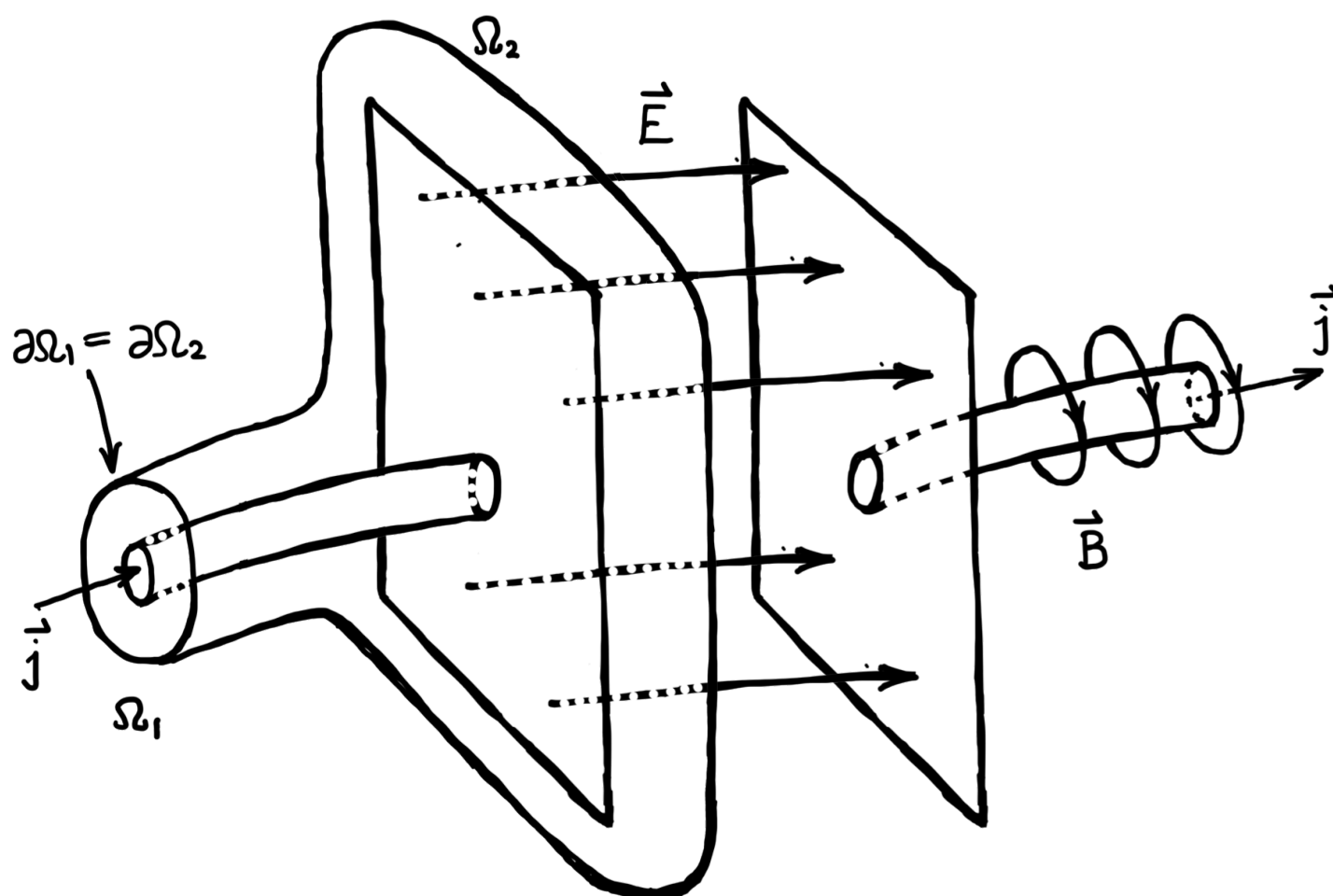
$\mathcal{R}$ 为导线单位长度的电阻（默认导线材料均匀截面不变）。 b 对应于测试导线之外电路其他部分的电阻 r ，而 a 对应于电动势 $\mathcal{E}$ ，Ohm得到的结果相当于

$$I = \frac{\mathcal{E}}{r + \mathcal{R}l}$$

这其实已经是一个更完整的对整个电路的结论： $\Delta V_r + \Delta V_R = \mathcal{E}$ ，其中测试电阻上的电势差 $\Delta V_R = IR$ 和其他电阻部分的电势差 $\Delta V_r = Ir$ 都正比于电流 I 。

位移电流*

我们来稍微详细叙述一下原文电磁学一节末尾所谈及的“电流产生磁场这个方程其一致性要求”，顺便也为后文做准备；Maxwell理论在现代物理基础的构建过程中登场数次，我们不妨权当熟悉/复习一下。



译者图 1.6: 假设电容器板很大, 电场 $\mathbf{E}$ 只分布于二块板之间, 而电流 $\mathbf{j}$ 只在导线之内。电流穿过 Ω_1 , 进入 Ω_1 和 Ω_2 围住的区域, 增加区域内的电荷 (电容器板上), 而电荷增加又伴随 Ω_2 上的电场通量增加。一边的“电荷流 $\mathbf{j} = \rho \mathbf{v}$ ”通量和另一边的“电场变化率 $\partial \mathbf{E} / \partial t$ ”通量此消彼长。虽然这是随时间变化的动态情况, 但总可以考虑变化得足够慢, 乃至静电场方程 (Gauss 定理) 依然成立, 于是位移电流这个逻辑结论几乎一定可靠 (当然也得假设场方程数学形式本身也是从静态到动态“连续变化”的)。注意即使考虑到电流可能无法保持稳恒, 从而可能有变化的磁场 $\partial \mathbf{B} / \partial t \neq 0$, 法拉第定律依然保证了伴随而来的“磁生电”那部分电场不产生额外的电场通量。Gauss 定理“兼容”Faraday 定律, 最终只有电荷变化才对应闭合曲面净电场通量的变化。

入门教材常以电容器充放电过程中 (此时电荷分布与电流都是随时间变化的) 要求磁场的 Ampère 环路定理 (定理原本是只对稳恒电流证明了的) $\oint_{\partial \Omega} \mathbf{B} \cdot d\mathbf{x} = 2\pi k_m \iint_{\Omega} \mathbf{j} \cdot d\mathbf{A}$ 对任意曲面 Ω 都成立来引入“位移电流”的概念 (参见译者图 1.6): 对同一个曲面边界 (闭合曲线) $\partial \Omega$, 曲面 Ω 可以切过导线从而有电流穿过 (不妨称其为 Ω_1), 也可以绕道从电容器中空部分经过从而没有电流穿过 (称为 Ω_2)。没有电流穿过的情形, $\mathbf{j} \cdot d\mathbf{A}$ 就无从谈起了, 但此处电场 $\mathbf{E}$ 一定是随时间变化的, 注意到只有 Ω_1 上有电流穿过, 而只有 Ω_2 上有电场穿过, 于是 (注意 Ω_1 的法向量 $\hat{\mathbf{n}}_1$ “向右”, 对于 Ω_1 和 Ω_2 构成的封闭曲面 $\partial \omega$ 而言是内法向量; 所有封闭曲面积分中 $d\mathbf{A}$ 默认外法向量方向, 在 Ω_1 上与 $\hat{\mathbf{n}}_1$ 方向相反, 相差了负号; 而 Ω_2 的法向量 $\hat{\mathbf{n}}_2$ 还是“向右”, 对 $\partial \omega$ 而言是外法向量):

$$\iint_{\Omega_1} \mathbf{j} \cdot \hat{\mathbf{n}}_1 dA = \oint_{\Omega_1 + \Omega_2 = \partial \omega} -\mathbf{j} \cdot d\mathbf{A} = \frac{\partial}{\partial t} \iiint_{\omega} \rho d^3x = \frac{1}{4\pi k_e} \frac{\partial}{\partial t} \oint_{\partial \omega = \Omega_1 + \Omega_2} \mathbf{E} \cdot d\mathbf{A} = \frac{1}{4\pi k_e} \frac{\partial}{\partial t} \iint_{\Omega_2} \mathbf{E} \cdot \hat{\mathbf{n}}_2 dA$$

其中第二个和第三个等号分别用到了电荷守恒和Gauss定理（下文详述）。关键是， $(1/4\pi k_e) \partial_t \mathbf{E}$ 替代了电流 $\mathbf{j}$ ，是为“位移电流”（displacement current），不再赘述。

我们打算换个方式，不拘泥于电容器等特殊情形，直接从电磁场的微分方程来分析。在引入位移电流之前²⁹，针对静电场和静磁场，以及唯一已知的“动电生磁”（电磁感应）情形，已经建立了如下方程：

- 静电场的高斯定理（Gauss' law）：

$$\nabla \cdot \mathbf{E} = 4\pi k_e \rho \quad (1.3.a)$$

假设其对不随时间变化的电荷分布 ρ 是正确的，这种情况下 (1.3.a) 等价于 (1.3.1) 库仑定律（Coulomb's law）。

- 静磁场的安培环路定理（Ampère's circuital law）：

$$\nabla \times \mathbf{B} = 2\pi k_m \mathbf{j} \quad (1.3.b)$$

假设其对不随时间变化的稳恒电流 $\mathbf{j}$ 是正确的，这种情况下 (1.3.b) 与 (1.3.2) 一致。

- 静磁场的高斯定理（Gauss' law for magnetism）：

$$\nabla \cdot \mathbf{B} = 0 \quad (1.3.c)$$

假设不存在磁荷，稳恒电流的磁场和其他如永磁体的静磁场总是是没头没尾的。

前三个方程都是对静电分布和稳恒电流，即不随时间变化的场得到的，尚不知其对随时间变化的情形变成什么样子；

- 法拉第定律（Faraday's law）：

$$\frac{\partial}{\partial t} \mathbf{B} + \nabla \times \mathbf{E} = 0 \quad (1.3.d)$$

磁场变化会产生（伴随）无头无尾的电场，有导体（如导线）在的话就能驱动电流。注意这个表达式中的 $\mathbf{E}$ 和 $\mathbf{B}$ 自动与第三个方程 (1.3.c) 一致，因为 $\nabla \cdot \nabla \times \mathbf{E} = 0$ 是恒等式，自然

²⁹ 译者注：位移电流（displacement current）这个概念构想是James Clerk Maxwell（詹姆斯·克拉克·麦克斯韦，1831–1879）在其1861年论文《论物理力线，第三部分》（*On Physical Lines of Force, Part III*）中提出的，联系着电介质中电荷的位移，Maxwell在Ampère环路定理的电流项中加入了位移电流。在其1865年论文《电磁场的动力学理论》（*A Dynamical Theory of the Electromagnetic Field*）中，Maxwell使用这一修正过的Ampère环路定理推导出电磁场的波动方程，这一推导被认为是物理学史的里程碑之一。

就有 $\partial_t(\nabla \cdot \mathbf{B}) = 0$ ，即 $\nabla \cdot \mathbf{B} = \text{const.}$ 不随时间变化；只要初始 $\nabla \cdot \mathbf{B} = 0$ ，就永远如此。换言之，随时间变化的磁场依然没头没尾，但前二个方程对变化的场尚需讨论。

• 电荷守恒定律：

$$\frac{\partial}{\partial t}\rho + \nabla \cdot \mathbf{j} = 0 \quad (1.3.e)$$

先看不随时间变化的情形，静止电荷与稳恒电流对应的电荷守恒 $\partial_t \rho = -\nabla \cdot \mathbf{j} = 0$ 当然是成立的，因为电荷分布 ρ 不随时间变化，而如果 $\nabla \cdot \mathbf{j} \neq 0$ 则会导致电荷聚集或分散，就矛盾了。假如我们认为即使电荷运动，电荷在我们的参照系中还是守恒的，即使电荷分布随时间变化，也只能是“流进流出”，分离中和，电荷净代数值不能“凭空”创生或消灭，那么这个动态关系就成了线索。

我们的思路是：如果电荷与电流分别联系着电场与磁场，那么将前二者约束在一起的守恒，大概就把后二者（场）也约束在一起了。也就是说，即使没有Faraday定律，动态的 $\mathbf{E}$ 和 $\mathbf{B}$ 也不是完全独立的，至少电荷守恒对二者施加了约束。

下一步假设是，即使超出静电荷分布的情形， $\mathbf{E}$ 还是满足 (1.3.a)。就是说， $\mathbf{E}$ 可以不全是静电荷的电场，还可以有别的不知道的成份；但即便如此，也只有电荷可以成为 $\mathbf{E}$ 的头和尾，电场的其他成份都是没头没尾的，Faraday定律 (1.3.d) 中“磁生电的电场”与此假设相符。接下来，在电荷守恒 (1.3.e) 中用 $\mathbf{E}$ 替代 ρ ，可以得到

$$\frac{\partial}{\partial t}\rho + \nabla \cdot \mathbf{j} = \frac{1}{4\pi k_e} \nabla \cdot \frac{\partial}{\partial t} \mathbf{E} + \nabla \cdot \mathbf{j} = \nabla \cdot \left(\frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} + \mathbf{j} \right) = 0$$

上式括号中已经可以辨认出位移电流的影子了。对于任意的封闭曲面 $\partial\omega$ 计算“组合流”

$$\oiint_{\partial\omega} \left(\frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} + \mathbf{j} \right) \cdot d\mathbf{A} = \iiint_{\omega} \nabla \cdot \left(\frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} + \mathbf{j} \right) d^3x = 0$$

这意味着通过任意封闭曲面的“组合流” $(1/4\pi k_e) \partial_t \mathbf{E} + \mathbf{j}$ 是个此消彼长的东西，电流 $\mathbf{j}$ 的变化由 $(1/4\pi k_e) \partial_t \mathbf{E}$ 抵偿。完全没有电流的曲面上全是电场变化 $(1/4\pi k_e) \partial_t \mathbf{E}$ 的通量，反之亦然。也可以说成是，借助流体的形象语言，沿着组合流场 $(1/4\pi k_e) \partial_t \mathbf{E} + \mathbf{j}$ 构成的流线走，一簇流线围成流管，流管中的组合流量完全不变。总之， $(1/4\pi k_e) \partial_t \mathbf{E}$ 被定名为“位移电流”完全说得通。

一旦超出稳恒电流的情形，我们就不知道 $\nabla \times \mathbf{B}$ 长什么样子了。如同 $\mathbf{E}$ 可以有非静电的不知道的成份（如“磁生电”的成份）， $\mathbf{B}$ 也可以不全是稳恒电流的磁场，而有其他未知成份。假设 $\nabla \times \mathbf{B} = 2\pi k_m (\mathbf{j} + \mathbf{f})$ ，稳恒（不随时间变化的）情形中应该有 $\mathbf{f} = 0$ 。无论 $\mathbf{B}$ 是具体什么样子，数学上 $\nabla \cdot \nabla \times \mathbf{B} = 0$ 对任意向量场恒成立，于是就有 $\nabla \cdot (\mathbf{j} + \mathbf{f}) = 0$ ，与组合流的方程

$$\nabla \cdot \left(\mathbf{j} + \frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} \right) = 0$$

如出一辙。当然也可以在电荷守恒 (1.3.e) 中用 $\mathbf{E}$ 替代 ρ ，用 $\mathbf{B}$ 和 $\mathbf{f}$ 替代 $\mathbf{j}$ ，得到

$$\frac{1}{4\pi k_e} \nabla \cdot \frac{\partial}{\partial t} \mathbf{E} + \nabla \cdot \left(\frac{1}{2\pi k_m} \nabla \times \mathbf{B} - \mathbf{f} \right) = 0$$

也就是

$$\nabla \cdot \left(\frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} - \mathbf{f} \right) = 0 \quad (1.3.f)$$

无论哪种写法，都暗示 $\mathbf{f} = (1/4\pi k_e) \partial_t \mathbf{E}$ ，也就是位移电流，而且已经自动满足不含时间的情况下 $\mathbf{f} = 0$ 。如此一来，第二个方程 (1.3.b) 就改造成

$$\nabla \times \mathbf{B} = 2\pi k_m \left(\mathbf{j} + \frac{1}{4\pi k_e} \frac{\partial}{\partial t} \mathbf{E} \right) = 2\pi k_m \mathbf{j} + \frac{k_m}{2k_e} \frac{\partial}{\partial t} \mathbf{E}$$

即Maxwell方程组其中一个方程的形式，此处已经出现了对应于光速平方的 $2k_e/k_m$ 这个因子。严格地说（向量场的亥姆霍兹分解Helmholtz decomposition定理），(1.3.f) 意味着 $(1/4\pi k_e) \partial_t \mathbf{E} - \mathbf{f}$ 是一个无散场，可以写作 $\mathbf{f} = (1/4\pi k_e) \partial_t \mathbf{E} + \nabla \times \mathbf{g}$ ，其中 $\mathbf{g}$ 是某个向量场，而且不含时情况须要 $\mathbf{g} = 0$ 。保留 $\mathbf{g}$ 是冗余的，因为最终 $\nabla \times \mathbf{B} = 2\pi k_m \mathbf{j} + (k_m/2k_e) \partial_t \mathbf{E} + 2\pi k_m \nabla \times \mathbf{g}$ 完全可以将 $\mathbf{g}$ 吸纳入磁场 $\mathbf{B}$ ，只要定义 $\mathbf{B} - 2\pi k_m \mathbf{g}$ 为最终的磁场即可，也就是说， $\mathbf{g}$ 与磁场 $\mathbf{B}$ 的一部分不可区分（整组微分方程形式完全确定之前，场尚未完整定义）。截至此，Maxwell方程之间已经完全满足彼此一致性的要求：

$$\text{非齐次方程:} \quad \nabla \cdot \mathbf{E} = 4\pi k_e \rho \quad \nabla \times \mathbf{B} - \frac{k_m}{2k_e} \frac{\partial}{\partial t} \mathbf{E} = 2\pi k_m \mathbf{j} \quad (1.3.g)$$

$$\text{齐次方程:} \quad \nabla \cdot \mathbf{B} = 0 \quad \nabla \times \mathbf{E} + \frac{\partial}{\partial t} \mathbf{B} = 0 \quad (1.3.h)$$

电解的发现

电解是由化学家William Nicholson（威廉·尼科尔森, 1753–1815）和外科医生Anthony Carlisle（安东尼·卡莱尔, 1768–1840）于1800年发现的。他们发现，当伏打型的电池的负极和正极电线分别插入水中时，会产生氢气和氧气的气泡。Sir Humphrey Davy（亨弗雷·戴维爵

士, 1778–1829) ——Faraday在皇家研究院的上司——进行了大量关于熔融盐电解的实验。他发现, 例如在对熔融食盐进行电解时, 一种前所未知的金属, 钠, 会在与电池负极连接的电线上析出, 而一种绿色的气体, 氯气, 则会在与另一端与正极连接的电线上析出。Davy的电解实验将多个金属加入到Lavoisier (拉瓦锡) 的元素表中, 除了钠, 还有铝、钾、钙和镁。

Faraday提出了电解的理论。用现代的术语来说, 水分子中有一小比例部分 (在室温下大约为 1.8×10^{-9}) ³⁰ 会常规离解形成正的氢离子 (H^+), 它们被吸引到连接电池负极的电线上; 同时形成负的氢氧根离子 (OH^-), 它们被吸引到连接电池正极的电线上。在连接负极的电线上, 二个 H^+ 离子与电池提供的二个单位的负电荷结合, 形成一个中性的 H_2 分子; 在连接正极的电线上, 四个 OH^- 离子会生成一个 O_2 分子、二个 H_2O 分子和四个单位的负电荷——这些负电荷流过电池到负极³¹。

同样地, 熔融的食盐 ($NaCl$) 中也有一小部分分子常规离解成 Na^+ 离子和 Cl^- 离子。在连接电池负极的电线上, 一个 Na^+ 离子加一个负电荷, 给出一个金属钠原子 (Na); 在连接正极的电线上, 二个 Cl^- 离子给出一个氯气分子 (Cl_2) 和二个单位的负电荷——这些负电荷流经电池到负极。

在Faraday的理论中, 需要一个单位的电荷才能将一个单电荷离子——如 H^+ 或 Cl^- ——转化为中性原子或分子。因此, 由于分子量为 μ 的分子质量为 μm_1 , 将质量为 M 的单电荷离子转化为质量为 M 、分子量为 μ 的中性原子或分子需要 $M/\mu m_1$ 单位的电荷。实验表明, 大约需要 96500 库伦 (例如, 一个安培的电流持续 96500 秒) 才能将 μ 克 (即一摩尔) 的单电荷离子转化为中性原子或分子 (这被叫做一法拉第常数 (faraday); 其现代值为 96486.3 库伦/摩

³⁰ 译者注: 意思是离解的分子数 (H^+ 或 OH^-) 占全部水分子数的比例为 1.8×10^{-9} , 差不多每十亿个水分子有 1.8 个离解了。

³¹ 作者注: 我们现在知道, 流过电池的是负电荷, 即电子。就Faraday当时所知, 正电荷流过电池也是同样可能的。在这种情况下, 在连接负极的电线上, 两个 H^+ 离子将形成一个 H_2 分子加二个单位的正电荷, 这些正电荷将通过电池流向连接正极的电线, 在那里, 四个 OH^- 离子加上四个单位的正电荷将生成一个 O_2 分子和二個 H_2O 分子。

译者注: 今天, 我们认为负电荷移动、正电荷不移动是区别元素/原子种类的一个关键特征: 质子数决定了物质的性质, 质子全都固定在原子核, 而对应个数的电子中的一部分可以更轻易被移动。但是要知道, 原子模型在当时还不存在, 电被看成一种流进流出的无质量“流体” (后文有长注解), 电荷量和元素之间并没有被认为有什么关系。

尔)。由此，Faraday就知道了 $e/m_1 \simeq 96500$ 库伦/克³²，其中 e 是电荷单位——1874年爱尔兰物理学家George Johnstone Stoney（乔治·约翰斯通·斯托尼, 1826–1911）将其称为“电质”（electrine）。已经测量了法拉第常数，如果物理学家还知道 e 的值，就能知道 m_1 ，但这一信息直到后来才获得。此外，当时没有人知道 e 其实是某个粒子的电荷。

1.4 电子

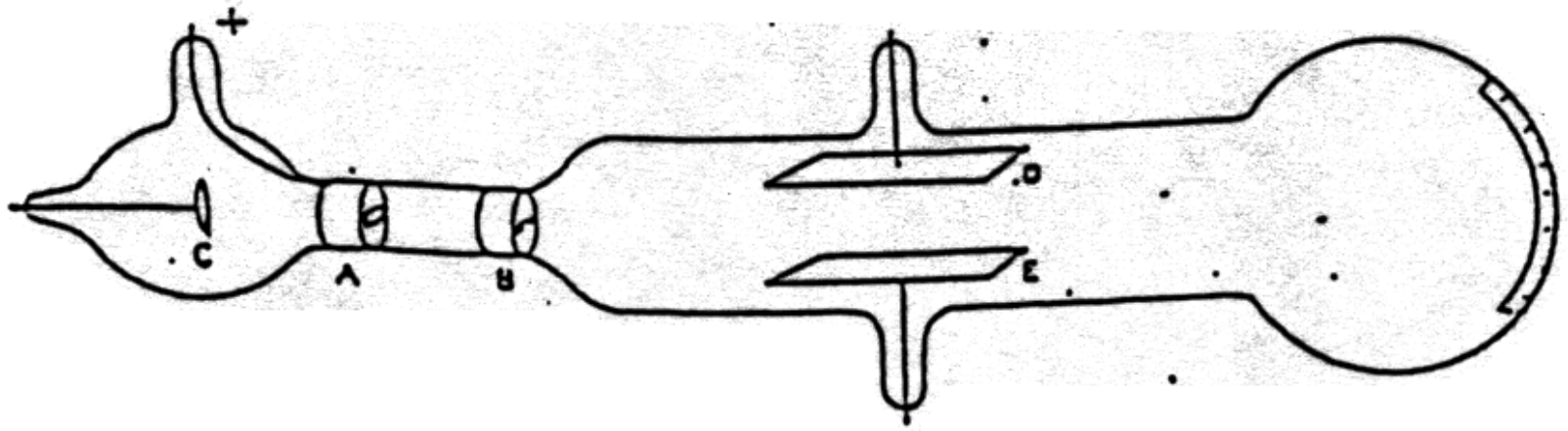
正如有时发生的那样，1858年，基础物理学的一条新路径随着一种实用设备的发明而开启，这回是改进版的空气泵。在他的泵中，波恩工匠Heinrich Johann Geissler（海因里希·约翰·盖斯勒, 1814–1879）使用汞柱作为活塞，大大减少了困扰以往所有空气泵的活塞漏气的情况。凭借这个泵，Geissler能够将密闭玻璃管中的压强降到地球表面典型空气压强的万分之一。

在玻璃管中这样的近真空环境下，电流可以在没有导线的情况下在管中传输。人们发现，电流会从阴极流出（连接到强电池负极的一块金属板），然后飞过阳极上的一个孔（连接到电池正极的另一块金属板），最后在管壁的远端点亮一个光点。向管内添加少量各种气体会使阴极射线亮起：当添加氖气、氦气或汞蒸气时，沿射线路径会发出橙色、粉红色或蓝绿色的光。利用Geissler的泵，Julius Plücker（尤利乌斯·普吕克, 1801–1868）在1858至1859年间发现，阴极射线会被磁场的偏转，从而移动射线打到管端玻璃上的光点位置（参见译者图 1.7）。

1897年，Joseph John Thomson（约瑟夫·约翰·汤姆森, 1856–1940），Maxwell之后的剑桥大学卡文迪许（Cavendish）教席教授³³，开始了一系列对阴极射线偏转的测量。在他的实验

³² 译者注： $M/\mu m_1$ 就是质量 M 所含的原子或分子数（包括离子），如果所谓单电荷离子就是每原子/分子增减了电荷 e ，那么将这个数量的单电荷离子恢复中性，或者将这个数量的中性原子/分子电离成单电荷离子，就需要给予或移走 $Q = e(M/\mu m_1)$ 这么多电荷。当 $M = \mu$ 克的时候（也就是一摩尔物质），这时候电量 Q 被称为1个faraday，消去 $M = \mu$ 得到 $1 \text{ faraday} = e/m_1$ ，即一摩尔 e 的电量是多少。正如文中所言，从逻辑判断，单个 e 并不必须和某个粒子绑定，也就是说电离完全可能只是一定量的“电流体”流进流出原子/分子（类比于宏观物体的热流进出），而不必要伴随质量或其他“结构”变化。

³³ 译者注：卡文迪许教授职位（Cavendish Professorship）是剑桥大学物理系的高级教职之一。该教席职位于1871年成立，卡文迪许实验室也随之在三年后建成。第七代德文郡（Devonshire）公爵William Cavendish（威廉·卡文迪许）为了纪念他的家族先辈，18世纪物理学家与化学家Henry Cavendish（亨利·卡文迪许），而捐赠了该教席和实验室。



译者图 1.7: Thomson对克鲁克斯管(Crookes tube)的示意图, 借此他观察了阴极射线在电场中的偏转(并后来测量了其荷质比)。阴极射线从阴极 C 发出, 穿过狭缝 A (阳极) 和 B (接地), 然后通过 D 和 E 两块电板之间产生的电场(当然也可以用电流线圈施加磁场), 最终撞击远端的表面。

中, 射线粒子穿过阳极后会受到电或磁力 F , 作用方向与射线运动方向呈直角, 在沿着射线的一段距离 d 内持续作用。随后, 它们会在没有力的区域漂移一段距离 $D \gg d$, 直到它们撞上管的末端。如果一个射线粒子沿射线方向的速度为 v , 它在时间段 d/v 内感受到电或磁力, 之后漂移一段更长的时间 D/v 。垂直于射线的力 F 给予质量为 m 的射线粒子一个垂直于射线的速度分量, 等于加速度 F/m 乘以时间 d/v 。因此当射线粒子撞击管子末端时, 它们已经被偏移了一定程度:

$$\text{位移} = (F/m) \times (d/v) \times (D/v) = \frac{FdD}{mv^2}$$

与射线呈直角的电场 E 或磁场 B 作用在电荷 e 上的力分别为:

$$F_{\text{elec}} = eE, \quad F_{\text{mag}} = evB$$

因此:

$$\text{电位移} = \frac{eEdD}{mv^2}$$

$$\text{磁位移} = \frac{eBdD}{mv}$$

Thomson希望测量 e/m 。他知道 D 、 d 、 E 和 B , 但不知道 v 。如果他能测量电位移和磁位移, 就可以从方程中消去 v , 但是电位移的测量非常困难。强电场易于电离管内的残留空气, 而正负离子会分别被吸引到产生电场的带负电和带正电的极板上, 从而中和它们的电荷。最终, Thomson通过使用极低气压的阴极射线管, 成功测量了电和磁偏移。(电和磁位移二者都只有几英寸。) 由此给出的电荷与质量的比值结果范围从 6×10^7 到 10^8 库伦每克。

Thomson将这一结果与Faraday在电解实验中测量所得的结果进行了比较，后者是 $e/m_1 \approx 10^5$ 库伦每克，其中 e 是单电荷电离的原子或分子（如在NaCl电解中的钠离子）所带的电荷， m_1 是假想的带有单位原子量的原子质量，接近氢原子的质量。他推测，如果他阴极射线中的粒子与电解中转移的粒子相同，那么它们的电荷一定和 e 相同，因此它们的质量必然是大约 $10^{-3} m_1$ 。Thomson得出结论：由于阴极射线粒子比离子或原子轻这么多，它们肯定是离子和原子的基本构成成分³⁴。

Thomson仍未测量出 e 或 m 。他甚至没有证明阴极射线是粒子流；它们可能是带电流体的流动，而任意体积流体具有的荷质比都与他所测得的 e/m 相等。尽管如此，在接下来的十年里，人们普遍接受了Thomson确实发现了原子中存在的粒子，这个粒子后来被称为电子。

我们稍微整理一下“电子”概念的逻辑，从已知的事实与一整套环环相扣的假设来看：

- 每个原子/分子 $\rightleftharpoons$ 单电荷离子进出一份电荷 e ，当然可以折算成每 μ 克原子/分子/离子（也就是每摩尔）对应的进出 1 faraday $\approx 10^5$ 库伦。比如每1克氢原子 $\rightleftharpoons$ 氢离子进出/携带电量 10^5 库伦，折算成每个氢原子进出 $e \approx 10^5 m_1$ 库伦。
- 进出原子分子的电荷与导线中宏观电流的电荷似乎没什么不一样，尤其是电解实验所观察到的：比如，伏打电池金属Cu和金属Zn化学反应（原子 $\rightleftharpoons$ 离子）转移的电荷不断进入或来自导线的宏观电流，而同一导线中流动的电荷又可以电解水，从而进出分子（ $2H^+ + 2e^- \rightleftharpoons H_2$, $4OH^- \rightleftharpoons 2H_2O + 2O_2 + 4e^-$ ）。电解的事实甚至意味着：如果电荷（无所谓“哪里来的，谁的”电荷）和电的力量能够将物质（原子分子）拆解重组，那么日常将物质绑定在一起（相对稳定的原子分子）的可能也是同样的电荷和电的力量。
- 从阴极金属板飞出到阳极金属板的“阴极射线”好像也是一样的电流，只是此时电荷不在导线内流动，而是在真空管里面飞。Thomson测量了这种飞行电流还带质量，每克阴极射线 $\approx 10^8$ 库伦。显然，同样电量的话，阴极射线比氢原子质量还轻很多。
- 如果其实是一样的电荷飞出原子分子飞出导线（当然也可以飞回去），那么这种带质量的电荷就应该是原子分子的一部分（包括其质量）。换句话说，原子/分子 $\rightleftharpoons$ 离子这样的过程其实是转移了一点点阴极射线。
- 进一步假设如果这种带质量的电荷如同原子分子一样（毕竟它们好像也是按份按比例分配给原子分子的，至少电荷如此），是分成一份一份的颗粒/粒子；而且成分单一，就只是一种质量为 m 的颗粒，每个的电量就是单电荷离子的电量 e ，那么（Thomson测量的结果）每克

³⁴ 译者注：也就是说，原本构想的不可拆分的原子和离子还不是最基础的构成物质的单位——因为如果一种质量更小但不为零的粒子被证明存在，那么原子和离子很可能还能再被进一步拆解成这一粒子。

阴极射线的电量折算成每个颗粒的电量就是 $e \approx 10^8 m$ 库伦（而单个氢原子 $\rightleftharpoons$ 氢离子进出的是 $e \approx 10^5 m_1$ 库伦，假设就是同一种“阴极射线颗粒”）；

- 最后，相较氢原子的质量， $m \approx 10^{-3} m_1$ ，这种假想粒子的质量是其千分之一的量级，也符合此前在化学反应、电解、宏观电流、静电荷等情形，包括宏观偏离电中性的情况都没有观察或测量到伴随电荷进出的质量变化。比如电容器的二块板，似乎并未因其增减电荷而变轻变重，做个简单的量级估计：假如一块铜板正好是1摩尔Cu，63.55克。即使增加巨大的1库伦负电荷（典型的数量级估计下，铜板附近空气在强电场下早已电离击穿），分配到每个原子上只有 10^{-5} 比例（十万分之一摩尔）的铜原子多了电子，而增加的质量只有 10^{-9} 克（1摩尔氢原子质量是1克， 10^{-5} 摩尔氢原子是 10^{-5} 克，而电子质量是其 10^{-3} ，接近二千分之一），相当于铜板质量的 10^{-10} ，百亿分之一。长期以来，人们没有认为“电流体”同时还是质量流（如水流），而更像热流（伴随着热流也几乎观察不到质量变化；我们在后续章节将看到的，在19世纪的发展中，热流被看作能量流）——电子质量相对原子的其他组成部分太小了，以至于宏观测量不会受到什么影响。