物理学的历史和哲学

序

吴大猷先生 1907 年出生于广州一个书香门第。早年就读于天津南开大学物理系，得我国物理学界前辈饶毓泰的赏识。1931 年经饶毓泰和叶企孙推荐，赴美国密歇根大学深造。1932 年完成了《最重元素低能态》一文，预言铀后元素的存在，为后来超铀元素的发现做出了开创性的理论贡献。1933 年获博士学位。

1934 年，吴先生回国任北京大学物理系教授，时年仅 27 岁。抗日战争期间，吴先生转往昆明西南联合大学任教。尽管当时学校物质条件极差，但西南联大却以国际级学术水平见称。吴先生在艰苦条件下写成了《多原子分子的振动光谱和结构》一书。这是红外光谱、喇曼光谱和分子结构方面的首部研究专著，也是该研究领域的经典著作之一。吴先生因此书而于 1939 年获中央研究院丁文江奖。吴先生教书育人，奖掖后进，当年在他门下受业的学生，后来许多成为世界一流的科学家，其中包括两位诺贝尔物理学奖获得者杨振宁和李政道。

抗日战争胜利以后，吴先生去美国，先后在美国几所大学执教并从事物理学研究，其间有 14 年在加拿大国家研究院负责理论物理研究工作。

1956 年后，吴先生常到台湾讲课并兼任职务。1978 年从美国纽约州立大学退休后定居台湾。几十年来，他为台湾的科学和教育的发展绘制了蓝图，制订了发展规划和一系列政策，大大促进了台湾的科学和教育事业的进步，从而也推动了台湾的工业和经济的快速发展。

吴先生在从事科学研究和教学的同时，出版了 13 部专著、近 120 篇研究论文，涉及原子物理、分子物理、统计物理、天体物理与大气物理、原子核结构、散射理论、等离子体、气体和相对论等。吴先生在这些领域做出了不少具有独创性的贡献。

吴先生除了从事物理学研究，关心科技发展和教育改革之外，还涉足人文、社会等领域，对时代的弊病常常给以有力的针砭。吴先生自己说：“作为一个曾与任何组织和党派无瓜葛的‘局外人’，使我在决策时更独立也更客观。经年的严格学术生涯培养了我独立思考的习惯和思想的诚实性。对于无论个人或小集体利益的淡漠使我弥补我的坦率——对于过失直言不讳——可能带来的损失。”吴先生的坦诚率直赢得了公众的敬重。

70 年代以后，吴先生开始更多地思考有关科学的理论性问题，对科学，尤其是物理学的本质等进行深入的探讨。1974 年他出版了《现代物理学基础的物理本质和哲学本质》（The Physical and Philosophical Nature of theFoundation of Modern Physics）一书；1988 年 1 月，他因病住院，在医院写了《物理学的性质、简史和哲学》，此文发表在台湾中央研究院物理研究所集刊第 17 卷上；他在台湾大学物理系所作的 14 次系列讲演的汇集也以《物理学：它的发展和哲学》（Physics：Its

Development and Phi-losophy）为题刊登于国际现代物理杂志 1989 年第 4 卷上 ;1992 年，吴先生又将此文作了大幅度修改和扩充，正式成书出版，定名为《物理学：它的发展》（Physics：Its Develop-ment）。

吴先生谈物理学理论的这两本书——《现代物理学基础的物理本质和哲学本质》和《物理学：它的发展》，都是用英文写作和发表的，国内还没有译本，现在由中国社会科学院哲学研究所金吾伦、胡新和两位把两书结合在一起译成中文，介绍给更广大的读者，我认为，这是一件很有意义的工作。译本定会给国内广大物理学、哲学等学科的读者带来深刻的启示。

钱临照

吴大猷，中国物理学家和教育家。广东番禺县人，1907 年 9 月 29 日生。天津南开大学物理系毕业，美国密歇根大学博士，北京大学、密歇根大学、纽约州立大学教授。曾任加拿大国家研究院理论物理组主任，纽约州立大学物理系主任等职。1967 年任台湾科学发展指导委员会主任和科学委员会主任。1983 年至 1994 年任台湾中央研究院院长。

吴大猷科学研究领域广及原子分子理论、相对论、经典力学和统计力学等。他预言了后铀元素的存在，被誉为"后铀元素之父"，原子光谱中所作的发现被人以其名字命名，称为"吴态"。他是中国物理学界的泰斗，学生中有李政道、杨振宁这样的诺贝尔奖得主。

第一章　引论

物理学有若干个层面。首先，我们最关心的是那些我们看做物理学本身主题的东西，即自然界和实验室内发生的物理现象、定律和理论，以及它们的应用。接着，关心的是物理学发展的历史——观察中的重要发现，由实验所得出的理论，以及物理学各分支的形成。进而，有一个更深的层面，那就是基本概念本质的研究和理论结构的研究。这，我们将不严格地称之为物理学哲学。大多数学习物理学的学生都满足于物理学的“作业知识”（working knowledge），独独把物理学哲学搁置一旁。毕竟没有它，物理学家们仍能做好他们的工作。事实上，大量的贡献，尤其是应用物理学于实际问题上的贡献,都是那些对有关物理理论基础可能无太多兴趣的物理学家做出的。但是，一位物理学家从对物理学演变过程的一个更深刻更具批判性的理解中，能获得精神上的更大满足与享受。

让我们先看看“定律”与“理论”。物理“定律”乃是一种连接某些物理概念的关系。经验定律便是从观察或实验资料中发现的这样的关系。前者的一个例子是开普勒行星运动定律，这些定律是从丹麦天文学家第谷·布拉赫（1546—1601）的观测资料中得出的。后者的一个例子是玻意耳（1627—1691）的气体定律。

对观测资料或实验资料进行分析,为我们思考现象和表述经验定律而引进概念，这些构成了作为一门科学的物理学成长的最初阶段。第二阶段则是经过思考构建现象的“理论”。

理论的目的是使物理学家运用较少的或较简单的连接物理概念的假定关系，去力图“理解”（以定律形式表达的）经验事实。这些假定关系可能要求引进新概念。牛顿运动定律中的动量概念，牛顿万有引力概念，热力学中的熵概念，化学中和气体理论中的分子概念等等，便是例子。这些假定依据它们所扮演的角色而被叫做“假定”(具有一种尝试的性质）、“假设”、“公设”（具有一种较强的性质）和具有欧几里得几何学中公理性质的“公理”。

理论的第一个明显要求是，它要与所有已知的观测事实相一致。第二个更加严苛的要求是，该理论的一切逻辑推论，都必须与为检验它们而设计的所有进一步的实验结果完全相符。因此，在这个意义上说，理论总是“处在试验之中”的；一旦发现其逻辑结论与某种经验事实相矛盾，该理论就必须加以修正或被抛弃。

按照爱因斯坦的论述（1919年），理论有两大类。一类他称为“构建性理论”，气体动理论就是这类理论的一个范例。另一类他称为“原理性理论”，经典热力学和相对论是这类理论的范例。

牛顿运动三定律，严格说来，应称为理论。“三条定律”都是强的“原

理”，而非经验定律。但是，动力学的整个结构及其应用的成功（直到它们应用于原子和亚原子领域才出现例外），已把牛顿“理论”提升到“定律”的地位。

在本书中，我们将循着从 17 世纪经典动力学到目前的物理学的发展路子，强调基本概念和理论真正本质的基础性变化。因此，牛顿力学具有绝对时间的基本概念和决定论的运动方程 光学理论和电磁理论充分发展了场概念并为狭义相对论铺设了道路；狭义相对论抛弃了绝对时间和以太，并强调变换和协变的概念，推广到广义相对论，引进了一种时空引力的全新观点，开创了宇宙学的领域。在讨论巨量物质的性质时，发展出了三种方法：气体动理论、热力学和统计力学。当应用到黑体辐射问题时，热力学和统计力学导致了普朗克和爱因斯坦的量子论。原子物理学沿着玻尔理论的发展，导致了量子力学的发展，其中牛顿体系的基本概念和物理理论的本质，由数学形式体系和一种物理理论的新哲学所取代。

在物理学发展过程中，重要目标之一是已知各种相互作用的统一。电磁场的统一是最早的统一；接着于本世纪 10 年代后期开始致力于引力场与电磁场的统一。这个目标迄今尚未达到，但弱相互作用与电磁相互作用的统一在 60 年代首次达到，并于 1983 年得到实验的支持。这两种相互作用与强相互作用和引力相互作用的最终统一是目前研究的任务之一。这些话题，将在书中的有关章节中论及。

*　　　　　*　　　　　*　　　　　*

物理学家，或许还有科学哲学家都普遍赞同，相对论和量子力学是现代物理学的两块基石。每所大学的大学生都学这两门课程。数学方面，诸如洛伦兹变换及其推论、薛定谔方程为许多问题提供的解，所有物理系的学生都是相当熟悉的。然而，对于物理诠释和哲学诠释，却并不熟悉。教科书或讲演中通常也不对之作充分的强调。作者相信，为了对这两个课题有一个较深刻的理解和恰当的把握,某些关于历史透视和哲学涵义的知识是必要的。下面就这两个课题的几个方面作一简要的阐明。

1．观察及由此形成的物理概念

物理学，作为科学的一个学科，开始于对某些种类现象的观察。为了对这些观察作思考，以及将这些思考与别人进行交流，有必要形成某些概念。其中有些概念是基本的，它们是从我们的经验中形成起来的，例如空间和时间的概念。与空间概念相伴随，出现了几何学的概念，或许来自于耕种土地的实践经验，欧几里得几何学便诞生了。这种几何学是以其隐含在空间中能将一个图形变换成另一个全等图形的可能性为特征的。这种几何学,由于与我们物理世界的日常经验相近而如此深深地扎根于我们的思维之中，以致直到爱因斯坦于 1910—1915 年提出他的引力理论，在物理

学家和哲学家的头脑中，欧几里得空间依然是我们的物理空间。但是，其他的空间也是可能的，对于这些空间，上面所说的“全等”是不可能的，例如鸡蛋的二维表面。广义相对论将弯曲空间引入了物理学。

试考虑时间概念。我们将不涉及哲学家的时间概念。在物理学中，我们涉及的是，测量在同一空间彼此作相对运动的诸点上的事件之间的时间间隔。物理学中有绝对的时间（或宇宙时间）概念，这个概念一直沿用到爱因斯坦在他 1905 年的相对论对之作了批判性的重新考察，并且表明它没有物理意义为止。

由空间和时间概念，人们得到速度、加速度等导出概念。于是，在处理动力学现象时，便形成了诸如粒子、质量、动量、力、功、能量等附加概念。与此相似，也形成了波、电荷、温度等等概念。

2．物理定律

根据观察（和测量）的经验数据，应用已形成的概念，通过内插和外推，在概括的基础上运用归纳过程而得到种种关系，这些关系就成了物理定律。范例是：气体定律 $pV=RT$、开普勒行星运动定律、傅里叶热传导定律、安培定律、法拉第定律等等。物理定律在起源上是经验性的，尽管概括和抽象过程有时掩盖了这种经验本质。这后一种情况的实例是热力学定律和爱因斯坦—德布罗意关系（见第九章量子力学）。热力学定律是第一类和第二类永动机不可能这一经验事实的概括陈述；爱因斯坦—德布罗意关系则是由光电效应和康普顿效应等实验所揭示的波粒二象性的概括陈述。以这种概括形式，这些陈述在量子力学中可以被看做“原理”。

3．物理理论

单独的“物理定律”集合只不过是一种组织起来的、系统的经验资料的集合而已。物理理论的目的和功能是要将许多分离的定律通过一些观念集合成为一个统一的整体。这些观念被叫做“原理”或“公设”，它们由包含物理概念的关系（方程）组成。一个理论必须与所有已知事实相一致，并且必须能借助逻辑演绎导出能提示新的观察或实验的新关系。如果一个理论，它能导出一个丰富的新研究领域并且预言实验所提供的所有事实，那么，它就是一个“好的”理论。一个理论不是一个终极真理，因为即使一个理论的大量推论都被证实了，只要有一个预言与观察不符就将迫使我们修改、或者抛弃这个理论。通过修改我们的理论，我们力图接近“真理”。这便是理论的地位！

在经典物理学（见第二、三章）中，有形成物理学基础的重要理论，这就是牛顿运动定律、牛顿万有引力理论、麦克斯韦电磁理论以及气体动

理论。人们也可以认为热力学具有理论的地位。在现代物理学中，两块基石是（狭义、广义）相对论和量子力学。

由上可见，物理定律的表述基本上是一个归纳过程，而物理理论则不是。物理理论的创立依赖于物理学家的想像、直觉和创造力，尽管一些经验物理事实的知识是必要的。正是由于这些个人的因素——想像、直觉和创造力，“哲学”才进入了物理学。在相对论和量子力学中，我们不仅看到了对经典物理学的基本概念所作的某些基础性的重新考察和重新表述，而且也看到了对于物理理论的本质所取的批判的哲学态度。

在这里我愿顺便谈谈那个通常的陈述，即科学的目的是寻求“真理”。这个问题不存在争议，争议在于物理学中“真理”概念的含义。就拿苹果落地这个简单的、基本的现象来说吧。我不知道在牛顿理论之前这种现象是如何解释的，很可能其解释和“真理”在不同文化中是不同的。人们普遍同意在牛顿之后的两个世纪之中，物理学家中确定无疑地认为，“真理”是地球的重力吸引造成了物体落地。但是，随着爱因斯坦“引力理论”的提出（其预言与牛顿理论略有偏离，而实验所得的结果支持了爱因斯坦引力理论）似乎显示出“真理”已经改变了。再拿宇称守恒定律来说，以前它一直被视为是一个自然定律，直到李政道、杨振宁博士的工作和吴健雄及其他人的实验证明它在β衰变的弱相互作用中并不正确。其他例子也能说明这一点。由此足以说明，在物理学中和在科学中，“绝对真理”之说并不总是有意义的。

第二章　经典动力学

1．经典物理学的基本特性

本章和第三章都属于经典物理学。这里我们就经典物理学的性质，作一必要的说明。

经典物理学，即力学和电磁学的最重要特征，就是决定论的本性，其意是在时空内用微分方程描述现象，只要在任何一个时空内给定了条件，那么，微分方程就完备地和唯一地决定了在任何时空内的一个系统的态。（请注意：不必把这个特征与因果性概念等同起来，因果性概念可能还包含着某种与物理学无关的思想要素。）

经典物理学的这种决定论特征在人的天然思维中有它的形而上学起源，而在力学中有它的科学起源。现在经典动力学可以说在天体力学中有了它的基础,太阳系的行星运动能够经受重复的观察并且已经发现可以用运动方程高精度地加以描述。牛顿方程和以拉格朗日与哈密顿形式表述的牛顿方程，代表了最明确形式的经典决定论。

在经典力学中，“粒子”概念当然有它在我们日常经验中的起源，不过也可以被说成是在天体力学中有坚实的基础,这是指从天体观察中抽象出来而言，天体的运动可以用动力学方程描述。然而，还有另一些现象，

诸如声音、光和弹性物体，对于它们的描述则创造出另一个概念，即“波动”概念，发现用来描述它们最为合适——事实上合适到像“粒子”概念在物理学中一样的“自然”。现在，一个“波”的基本属性就是空间和时间中的周期性，这些属性可以用“波长”和“频率”概念来表达。于是这些属性就把波动与其基本属性是“动量”和“能量”的粒子区分开来。因此，在经典物理学中，我们就有了两种概念，“粒子”和“波”。

粒子	波
动量	波长
能量	频率

由它们的基本属性所表征的“粒子”和“波”是相互排斥的。它们的基本差别由于与波动有关的干涉和衍射现象而进一步加深。不以一种特设性的和强迫的方式给“粒子”增加非常复杂的性质，那么，粒子的“衍射”是难以设想的。

在经典物理学中，基本学科除力学外还有电磁学。而在电磁学中，电荷是在力学意义上的“粒子”，但对电荷之间相互作用的描述需要某些其他概念。法拉第引进“场”的概念，对物理学作出了最重要的贡献。“场”概念在麦克斯韦那里得到了充分的发展,他以场方程的形式给出了一种精确的数学表述。这些方程提示或预言了电磁场的波动性质，“电磁波”后来为赫兹所发现。伴随着麦克斯韦的理论和赫兹的发现，经典物理学可以说已臻“完成”了。“粒子”概念，为力学所庇护；“波动”概念，为电磁理论所涵盖。

经典力学中还有另一群现象，其中包含着大量的粒子。例如，考虑任何宏观量的任何气体。我们要处理的分子数目将达 10^{22} 数量级。在这种情况下，实际上不可能知道和鉴别系统的初始条件（动力学状态，即单个分子的坐标和动量）；我们也不着眼于这样一种细节的知识。我们只对系统在宏观尺度上的性质，即在我们观察和测量的尺度上的性质感兴趣。为此目的，我们引入宏观变量（作为与原子尺度上的‘微观’变量相对比），并且通过引入概率概念与统计方法来处理大量分子的平均性质。然而，在原子层次上理论仍然在本质上是决定论的,因为单个分子仍然受决定论定律支配。

以下我们将从经典动力学开始讨论，理由是：（1）经典动力学是物理学最早的部分，从 17 世纪初叶就由伽利略和牛顿发展成了一门定量的科学。（2）它是最简单的，只包含三个基本概念，即空间（长度的量度）、时间（持续的量度）和质量（一物体阻止其速度变化的量度）。

科学的发展可以追溯到亚里士多德（公元前 384—前 322 年）及其学派；但他们没有把他们的研究建立在实验之上。强调实验和观察或许始于弗兰西斯·培根（1561—1626）；不幸，这种强调是定性的，而不是定量的和测量的。开普勒（1571—1630）和伽利略（1564—1642）是同时代人；

他们同是“近代科学”的先驱者——开普勒强调数学，伽利略强调实验和定量测量。这里还必须谈到较他们年轻的同代人 R．笛卡儿（1596—1650）。笛卡儿起初是一位亚里士多德主义者；尽管他受了开普勒工作的影响，但他相信物理学是从先验原理中推演出来的（像欧几里得几何学那样），而不相信观察和实验。物理学作为一门科学，在近代意义上的确可以说是自伽利略和牛顿的经典动力学开始的。

现在，距离和时间持续的概念在以下意义上可以看做是“基本的”，即它们不但为人类所“理解”，而且也同样为动物所“理解”。从昼夜、满月和季节的更替等现象中，人形成了周期性的概念，并很快发现了大自然的经验规律：一年有大约 365 又 1／4 天。

距离或长度的原始概念，立即被扩展到二维和三维，人们从事土地测量的经验又导致了欧几里得几何学的形成，它也许是第一个逻辑体系，并且在两千余年内一直独领风骚，直至 19 世纪非欧几何学可能性的发现。几何学的这些发展，在物理学中扮演了重要的角色，因为“空间”乃是物理学真实基础的一个基本概念。

让我们回过头去看一看伽利略以前的时期。人们从空间和时间的基本概念中推导出像速度和加速度这样一些概念。伽利略用一个圆球在斜面上的运动做实验。他通过改变倾斜角所作的观察和结论，奠定了运动“定律”的基础，这后来又被牛顿（1642—1727）作为他的运动第一定律表述出来。这就是众所周知的“惯性定律”，该定律包含了空间和时间概念之外的另一个概念，即惯性或“质量”。一个物体的质量是该物体的一种属性，即“抵抗”物体运动“状态”（诸如速度）的变化的性质，因此名之曰“惯性”。

牛顿为了完善他关于一粒子（一物体抽象为一个质点）在力作用下的运动“理论”，表述了第二定律，即运动方程。关于力（作用）本身，牛顿又陈述了第三定律，即作用和反作用定律。

我们将从牛顿三个运动定律的陈述开始：

2．惯性定律

“一个质点，当它与所有其他质点相距足够远时，该质点之加速度消失。”

这一表达相当于那个若非更精确，也是更为人熟悉的表达形式：一物体，当无外力作用其上时，保持静止或匀速直线运动状态。请注意，作为一个基本理论，质点的概念是基本的，而一“刚性”物体的概念必须基于“大量的质点”（massive points）概念连同由质点组成的“物体”之结构（最终包含物质的原子本质）的辅助假设。

3．运动定律

“在外力作用下，一个质点的运动由以下方程描述

动量变化率＝力

或者

质量×加速度＝力。”

这里，一个质点的运动必须被理解为“相对于其他客体的运动”，而这个概念的数学表达是用笛卡儿坐标 x，y，z 作为“时间”t 的函数在一个适当选择的参考系内去描述这个点。很清楚，前面两个定律并非适用于任意参考系（牛顿的旋转桶，或一种旋转游戏装置的旋转台），而只适用于某些参考系。这些参考系被叫做“惯性系”。存在着无穷多个这样的参考系，即对一个惯性系作匀速（即无加速度）相对运动的所有参考系也都是惯性系。这是一种限制。广义相对论重新表述了动力学理论，以避免这种对惯性系的限制。

第二定律（ma＝f）是一个混乱的源泉。某些人争论说，如果将“力”视作原始概念，那么它是一个定义“质量”的方程，或者如果将“质量”视作原始概念，则它是一个定义“力”的方程。从逻辑的观点看，人们只要首先将质量作如下操作定义，便可避免上述混乱：不同物体（m_1，m_2，……）受相同“力”的连续作用（比方说，一个受压缩的弹簧），且在实验上它们的加速度的倒数比

$$\frac{1}{a_1}:\frac{1}{a_2}:\cdots\cdots$$

被用来定义它们的惯性（或质量）的比率

$$m_1:m_2:\cdots\cdots$$

人们可以把其中的一个视作标准单位。于是关系式 f＝ma 可以用来给出力的定量量度。

重要的是，不要把此混作循环定义。第二定律是一个理论，给出关于力（诸如万有引力，或在一个荷电粒子上的洛伦兹力）的定律意义上的理论，第二定律给出了运动方程，即对粒子在时空中的描述。

牛顿动力学在太阳系行星运动及所有其他方面应用的成功是如此之巨大，以致运动方程和万有引力的假设均被提升到“定律”的地位。

4．力的定律

牛顿第三定律是说，对于每一个力，都存在着一个大小相等方向相反的反作用力。

这个定律明显地应用在牛顿的万有引力理论中。

我们现在将从现代发展的观点，即从狭义相对论和广义相对论及量子

力学的观点，来考察一下牛顿体系的基本内容。

为清楚起见，我们先概述一下牛顿理论及其推论的基本内容：

（1）牛顿强调时间这一基本概念的意义，说：时间均匀和连续地流动，与任何其他事物无关。这样一种时间被称为“绝对的”或“普遍的”时间；因此所有观察者都有相同的时间，不论他们的相对运动如何。

在本书第四章中我们将看到，在物理学中，与哲学有别，这样一种绝对时间并不具有“操作”意义。

（2）在牛顿动力学中，首先，空间距离的测量与时间无关，这是根据绝对时间的定义规定的。第二，三维空间有欧氏几何学的几何性质。在牛顿时代，除了欧氏几何外，还不知有任何其他几何学，空间的几何学问题尚未提出。事实上，直到爱因斯坦于 1915—1916 年提出引力理论之前，并不存在空间几何学问题。

我们将在第四章中看到，空间和时间这两个概念并不是独立的，而且四维时空也并不一定与物质（或能量）的存在无关，因此事实上按照爱因斯坦引力理论它是黎曼时空。

（3）在牛顿动力学中，暗含着将以下一点视为当然的事，即同时测量（即知道）一个粒子（一个质点）的位置和动量在原则上是可能的。这种可能性隐含在运动定律本身中：运动的二阶微分方程的解要求知道 x 和 p_x 的某个同一时刻的初始值。

但是这种可能性在量子力学中从根本上被否定。

（4）牛顿动力学中运动方程是决定论的和因果律的，即从一个由系统的粒子之坐标和动量所规定的已知初态出发，运动方程以一种决定论的方式导致一切其后时刻的确定状态。这导致拉普拉斯（1749—1827）宣称：一旦给出了某一瞬间宇宙中所有星星的位置和动量，那么，宇宙过去和未来的状态都将完全被决定。

但这种决定论和因果律在量子力学中基本上被否定。

（5）让我们再次返回到牛顿动力学。我们已经强调过，牛顿的时间是“绝对的”。他的空间在下述意义上也是“绝对的”。

在惯性定律和运动定律的表述中，首先必须有参考系。于是很清楚，这些定律只对“非加速的”参考系才有效。但这隐含了一种“绝对的”空间；否则，“加速”一词便是空的。牛顿运动定律因而只对“惯性”（或伽利略）系有效。例如，在旋转系中，牛顿运动方程必须通过引入像离心力和科里奥利力等额外的“力”加以修正。这些“力”是由于参考系的加速才出现的；它们被认为是“虚拟的”，常称为“惯性力”。

虽则牛顿运动方程只对惯性系才有效，但它仍具有很大的普遍性——即它对于彼此作匀速相对运动的所有惯性系都有效。事实上，运动方程的形式在一切惯性系中都是相同的。我们将在第四章中回到这个“伽利略的相对性”。

5．哈密顿动力学：可积系统

牛顿的运动方程由拉格朗日（1736—1813）、泊松（1781—1840）、哈密顿（1805—1865）和其他人用各种形式加以表达。最早的形式是莫培督的最小作用原理（Maupertuis’ principle ofleast action，1744）

$$\triangle\int_{t_2}^{t_1} 2T\mathrm{d}t = 0$$

〔这里的变分号△意指，运动的变化路径和实际路径满足同样的能量条件，以使迁移时间 $t_1—t_0$ 对变化路径和实际路径来说是不同的，即 t 不是一个人们可以令其变分等于零的独立变量。〕

一个更普遍的形式则是哈密顿原理（1834 年）

$$\delta\int_{t_2}^{t_1} L\mathrm{d}t = 0$$

$$L = \sum p_k \dot{q}_k - H = 2T - H$$

而 H（p，q），哈密顿量是具有 n 个自由度的保守系统中的总能量：这里的变分号 δ 表示，t 是一个独立变量，并且所有的变化路径像实际路径一样具有相同的迁移时间 $t_1—t_2$。

从哈密顿原理，人们可以得到 n 个二阶拉格朗日方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0，k = 1，\cdots，n，$$

且哈密顿正则方程

$$\dot{p}_k = \frac{\partial H}{\partial q_k}，\dot{q}_k = \frac{\partial H}{\partial p_k}，k = 1，\cdots，n，$$

它们是 2n 个一阶微分方程。

虽则所有这些形式彼此等价并且在物理内容上也等价于牛顿的运动方程，但它们在进一步的数学发展中，例如在正则变换理论中就不同了。

让我们探索一下 2n 个 q_k，p_k 换成一个新的集合 Q_k，P_k，k＝1，…，n 的情况，这种在 Q_k，P_k 中的正则方程像在 q_k，p_k 中一样，具有相同的形式

$$\dot{P}_k = -\frac{\partial \overline{H}}{\partial Q_k}，\dot{Q}_K = \frac{\partial \overline{\overline{H}}}{\partial P_k}，k = 1，\cdots，n$$

$$\overline{H} = \overline{H}(Q_k，P_k)$$

〔这里，为简单起见，我们将考虑保守系统，即 H（q，p）并不明显依赖于 t。〕这种不变性的一个（充分）条件是

$$\sum p_k \mathrm{d}q_k - \sum P_k \mathrm{d}Q_k = \delta S$$

这里 S＝S（q，Q），从这一点出发，得出变换方程①

$$p_k = \frac{\partial S}{\partial q_k}，P_k = -\frac{\partial S}{\partial Q_k}，k = 1，\cdots，n$$

上述正则变换条件可以用许多等价形式表示，诸如 $(Q_i，Q_j)_{q,p}=0$，

$(p_i, p_j)_{q,p}=0$，$(Q_i, p_j)_{q,p}=\delta_{ij}$

其中

$$(u_i, v_j)_{qp} \equiv \sum\left(\frac{\partial u}{\partial q_k}\frac{\partial v}{\partial p_k}-\frac{\partial u}{\partial p_k}\frac{\partial v}{\partial q_k}\right)$$

是泊松括号表示。

2n 个一阶正则方程有 2n 个一次积分，其中之一是能量积分

$$H(q, p)=h$$

按照庞加莱的论证，一般地在 q_k，p_k 中不存在单值的和解析的其他积分（1892 年）。〔在特殊情况下，对应于动量和角动量守恒的一次积分存在。〕由此我们就能理解，为什么三体问题一般不可能用经典动力学“解决”。

最具有意义的是对一个已知其 n 个独立的单值的一次积分的系统，

$$f_j(q, p)=c_j,\ c_j=\text{常数},\ j=1, \cdots, n$$

如果任意对 f_j，f_k 的泊松括号消失，

$$(f_j, f_k)=0,\ j, k=l, \cdots, n$$

则正则方程的系统便是“可积的”。这一陈述可作如下证明：

让我们选择 $f_k(q, p)$ 是作用变量 J_k（具有相应的角变量 W_k），并且通过一个函数 $S^*(q, J)$ 按照

$$P_k=\frac{\partial S^*}{\partial q_k},\ W_k=\frac{\partial S^*}{\partial J_k}$$

作一个正则变换从 q，p 到 W，J，哈密顿量 H（q，p）转变为 $\overline{H}$（W，J），且正则方程

$$\dot{W}_k=\frac{\partial \overline{H}}{\partial J_k},\ \dot{J}_k=\frac{\partial \overline{H}}{\partial W_k}$$

但是因为 $J_k=f_k=$常数，可以知道 $\overline{H}$ 仅仅是 J 的函数，所以

$$J_k=\text{常数}$$

$$W_k=v_k t+\delta_k \qquad \left(\text{此处} v_k=\frac{\partial \overline{H}}{\partial J_k}=\text{常数}\right)$$

为了找出 $S^*(q, J)$，因 $J_k=$常数，所以

$$dS^*=\sum\frac{\partial S^*}{\partial q_k}dq_k=\sum p_k(q, J)\,dq_k$$

$$S^*(q, J)=\sum\int_0^{q_k} p_k(q, J)\,dq_k$$

从 $f_j(q, p)=J_j$，$j=1, \cdots, n$，人们就能得到作为 q，J 函数的 p_k，积分值也可求得；所以 $S^*(q, J)$ 就能找出。从 $W_k=\frac{\partial S^*}{\partial J_k}$，人们就能到 $q_k=q_k(W, J)$。最后得到

$$p_k=p_k(q, J)=p_k(W, J)$$

证毕。

阐明上述方法的最简单例子也许就是一个电子在核电荷为 Ze 的库仑场中的问题。当然无论是这种方法还是分离变量法均导致（类氢原子）系统能量的相同（索末菲）结果

$$E = -\frac{2\pi^2 mZ^2 e^4}{(J_v + J_\theta + J_\psi)^2}$$

这里 J_v，J_θ，J_ψ，为三个作用变量（也是一次积分）。

6．K－A－M 理论①

在一个反平方律引力场中,一个物体的运动方程因其高度对称性是完全可积的。运动是周期的，且因此是稳定的。一般说来，N 个物体的系统在无限长时间内运动的问题，由 2n 维相（或 r）空间内一个点 Q（q_1，…，q_n，P_1，…，p_n）的运动来描述。问题是点 Q 是否在时间进程中通过在 2n—1 维能量表面上的每个点（或者，为人们所要求的那样近乎每个点）。所谓各态历经假设（ErgodicHypoth-esis），即由 L．玻耳兹曼和 J．C．麦克斯韦于 19 世纪 80 年代提出的 Q 通过每个点的假设,已被认为是站不住脚的。庞加莱（1892 年）给出了所谓各态历经定理，即在长时间进程中，相点 Q 通过人们所希望的任何点附近，以致运动是准周期的，因此是准各态历经的。

然而，存在已知的行星运动，它在 10^9 年内是稳定的，即明显不是各态历经的。G．D．伯克荷夫已揭示了运动“形式”积分的存在（1927 年）；在稳定运动的“近邻”，系统不是各态历经的。

在该问题上的第二个重要进展来自 A．N．科耳莫戈洛夫（1954 年）、V．I．阿诺德（1963 年）和 J．莫塞（1962 年）的工作。他们研究了如下的问题：选取一个系统，它的哈密顿量 H（q，p）可以用一种正则变换而变换成前节所述的角作用变量（W，J），以使 H 可以表达成一个级数（用 ε 解析）

$$H(q, p) = H_0(J) + \varepsilon H_1(W, J) + \cdots$$

H_0（J）是其中 n 个一次积分为 J_k 的完全可积系统，ε 是一个“小的”参数，而 εH_1（W，J）是一个在系统中代表“相互作用”的“微扰”。

科耳莫戈洛夫、阿诺德和莫塞的重要结果是，对非常小的 ε，系统 H 具有稳定的行为（即非各态历经行为）。然而，情况十分复杂，且我们不可能对 K-A-M 理论作更详细的讨论。

7．万有引力理论

牛顿的万有引力理论，虽然不是普遍动力学原理的一部分，但在物理学发展中具有基本的重要性。我们将提出一系列的评论：（1）万有引力

理论是重要的，因为它把许多不同的现象统一成一个“定律”，即开普勒行星运动的三个经验定律（1609 年、1619 年）、月亮绕地球的运动、“苹果落地”及潮汐等。力图更简单地描述自然现象，一直是物理学追求的重要目标之一。我们将在下一章中看到另一个“统一”，即明显不同的电现象和磁现象统一成一个电磁场——相对论中的四维势；第十三章及附录 13 中将讨论电弱相互作用的统一和电弱强相互作用的统一。时下的目标是上述电、弱、强相互作用与引力相互作用的大统一。

（2）牛顿万有引力理论最初假定的是两大团粒子“超距”的相互作用，即中间无任何媒介。这种“超距作用”很难理解。牛顿事实上表达了这样的信念，即：经由真空而产生超距作用的思想是荒谬的。然而，牛顿在他的《原理》一书中说得很清楚，他的万有引力理论只是为（开普勒定律的）数学预言提供一个必要的工具，而不涉及引力的机制，即并不触及两个物体为什么互相吸引的问题。我们必须注意到，这一直是物理学家对待物理理论的主流哲学。一个引力定律只描述两个物体怎样相互吸引，而并不解释这两个物体为什么会相互吸引 对后一个问题的考虑已属物理学之外了①。

（3）在 r 点的一个物体 m 受到的来自 R 点的另一个物体 M 的万有引力，可以用一个（标量）势（能）加以表达

$$F=-\frac{GmM}{|r-R|^{3}}(r-R)=-m\nabla_{r}v \quad v=-\frac{GmM}{|r-R|}$$

这就引入了 R 点的一个质量 M 在 r 点产生的引力场概念。这样一个静态场满足拉普拉斯方程或泊松（1781—1840）方程。但是一个以有限速度在空间传播的场的完整概念还有待于 19 世纪法拉第和麦克斯韦的工作。

（4）随着牛顿万有引力理论说明开普勒行星运动三大定律的巨大成功，爱德蒙·哈雷（1656—1742）将此理论应用于彗星轨道。经数学家拉格朗日、拉普拉斯和高斯（1777—1855）之手，天体力学学科诞生了。这个理论也被提升到定律的地位。

在开普勒定律的时代，只知道有水星、金星、地球、火星、木星和土星等行星。当然引力定律并没有预言任何轨道上行星的存在，但在 1781 年天王星发现之后，人们发现它的运动表明有某些奇异的特性，按照万有引力定律，似乎指示出有来自一未知行星的扰动，即来自不是已知的从水星到土星的行星扰动。乌本·J．勒威耶和约翰·C．亚当斯通过计算，预言了未知行星的位置，1846 年，第八颗行星海王星被发现。又一个类似的故事重复出现，第九颗行星冥王星于 1930 年被发现。万有引力定律取得的成功确实是巨大的。

（5）物理学的进步和演化是一个无穷尽的过程，而万有引力定律也绝非已终结真理。爱因斯坦于 1915—1916 年在他的广义相对论基础上提出了一个新的“引力理论”，这个理论我们将在第四章和本书结尾的附录

3 中再作较详细的描述。

参考文献

吴大猷：古典动力学（理论物理第一册），科学出版社，北京，1983 年

（拉格朗日和哈密顿力学）

Articles by J .Moser,Y.M.Treve in Topics in NonlinearDynamics, A Tribute to Sir Edward Bullard,ed.bySiebe Jorne,American Iust. of Phys. ，New york，1978.

（K—A—M 理论）第三章　光学和电磁学

1．光学

有关光现象的研究领先于电和磁现象的研究一个世纪。最早知道的光传播的特征是直线传播以及由 W．斯涅耳（1591—1676）于 1621 年实验发现的反射和折射定律。但是，光理论发展的历史（在 17～18 世纪期间）是一段复杂的历史，其中包含了许多伟大人物的名字，像笛卡儿、惠更斯（1629—1695）、胡克（1635—1703）、牛顿、杨（1773—1829）、菲涅耳（1788—1827）等等，都卷入了微粒说和波动说之间的争论。下面我们只能提供一个非常简要的梗概。

笛卡儿有一个关于事物的总纲要：宇宙的机械论观点。他把以太概念作为具有机械性质的介质引入物理学。他关于（微粒）光的折射定律的演绎暗示，它在（比方说）水中的速度大于在空气中的速度。顺便说一句，牛顿的光微粒说也导致了同样的错误结论。

费马（1601—1665）于 1657 年提出了光传播的最小时间原理（Principle of Least Time）。这个变分原理形式的数学定律具有普遍性和重要性，他是在物理学中以这种形式表达定律（或原理）的先驱；这个定律也优于笛卡儿的理论，因为它作了正确的假定，即光速在（比方说）水中比在空气中要小。然而，该原理的推导是基于形而上学考虑而非物理学考虑的。

罗伯特 胡克是一位比牛顿稍年长的同代人。他赞成光的波动说(1667 年），而牛顿坚持微粒说。牛顿于 1666 年发现棱镜分离太阳光成光谱，并于 1671—1672 年批评胡克的理论，由此在他们两人之间展开了一场争论，弄得关系紧张。人们认为，这可能是造成牛顿后来不愿出版他的著作的因素之一。牛顿拒绝接受波动说，是由于波动说不能说明光的直线传播，因为在托马斯·杨 1801—1803 年的实验之前，衍射现象尚不知道。另一方面，偏振和双折射现象已由牛顿于 1717 年在光的两侧性（two sidedness）基础上得到“解释”，而且值得注意的是，他事实上利用偏

振现象作为反对“波动说”的强有力证据，因为那时理解的波动说考虑的是纵向声波的那种波。杨和菲涅耳的波为人所知要在一个世纪之后。

惠更斯比牛顿大 13 岁，但他的重要贡献，即惠更斯原理，在 1678 年才作出，1690 年发表，晚于牛顿 1666—1672 年的工作。通过这一原理，惠更斯就能“理解”反射和折射定律以及在一块冰洲石晶体双折射中两射线的偏振。

对波动说最严重的质疑，即它不能说明光的直线传播，由于杨在 1801—1803 年关于直棱的光衍射实验而最终得以消除。利用波的干涉和惠更斯原理“解释”了这个现象。杨的理论由菲涅耳给出了一个数学处理，菲涅耳建议杨在圆物体的阴影中心寻找光亮点，而这个光亮点被找到了。菲涅耳为法国科学院的有奖征文提交了一篇论文，该 论文使委员会（由拉普拉斯、泊松和毕奥组成）信服，菲涅耳于 1818 年赢得了这笔奖金。

直线传播的理由现在已经清楚了：在普通的条件下，窗口或开口的尺寸与光的波长比较非常之大，以致衍射现象观察不到。事实上，人们现在知道，一个波，在极短波长的极限下，像“射线”一样沿波前的法线传播（布伦程函，1911 年）。

与有关光本性的波—粒争论连在一起的是光传播的介质（后来被名之为以太）是否稳定或是否随地球绕太阳轨道运动而运动的问题。早在 1728 年，一位英国天文学家 J．布喇德雷就观察到恒星光的光行差。按照介质不参与地球运动的理论，观察到的恒星光的方向（由望远镜的方向 θ 显示）将与光行差角的真正方向 θ' 有一角 α 的差别，$\alpha=\theta'-\theta$，

$$\cos\theta=\frac{\cos\theta'+v/c}{1+(\frac{v}{c})\cos\theta}$$

或者，如果我们令 $\theta'=\pi/2$，那么对于小的 α 来说就有 $\sin\alpha\approx v/c$。这个关系正好与布喇德雷的观察相符。

1818 年，菲涅耳得到一个在速度为 v 的流水（折射率 n）中光速 u 的公式

$$u=\frac{c}{n}\pm(1-\frac{1}{n^2})v$$

v 平行与逆平行于 u，假定以太只是部分地被流水“拖曳”。斯托克斯（1846 年）得到了相同的公式，他作了有点类似的关于以太受流水压缩和稀疏化的假定。令人惊异之处在于一个依赖于如此详细的以太模型的公式竟然被菲佐的实验（1851 年）所证实！更有甚者，菲涅耳的预言，即如果望远镜由水充满，布喇德雷的光行差将不会改变，也竟于 1871 年由爱里所证实。我们将看到，菲涅耳的公式为实验所证实，因为狭义相对论（洛伦兹变换）也精确地给出菲涅耳公式！

用光所做的另一个实验是 1881 年的迈克耳孙实验和 1887 年的迈克耳孙-莫雷实验。实验被设计来检验地球围绕太阳作轨道运动，光的传播在

实验室干涉仪上是否有任何效应，结果是否定的。一个可能的解释是以太被地球携带着一起运动；另一种推测是，所有物质在运动方向（即迈克耳孙实验中地球的轨道速度ϖ的方向）上以一个因子$\sqrt{1-(v/c)^2}$缩。

迈克耳孙实验事实上是由麦克斯韦（在他 1879 年）逝世前提出来的。我们将简要评述导致 1905 年相对论诞生的电磁场理论的发展。

2．电磁学

在 18 世纪的最后十年或其之前，多数学者都致力于静电学与静磁学的研究。因此，J．普里斯特利（1733—1809）于 1766—1767 年发现电荷的平方反比定律；H．卡文迪什（1731—1810）1771 年的发现在他死后由开尔文勋爵于 1879 年发表；J．米切耳（1724—1793）和 C．A．库仑（173—1806）于 1785 年发现平方反比定律；拉格朗日以$\frac{\partial V}{\partial r}$的形式表达了作用在一个电荷上的力（1771年）；拉普拉斯给出方程$\nabla^2 V=0$（1777年）；泊松给出了方程$\nabla^2 V=-4\pi\rho$（1831年）；G．格林（1793—1841）将“势”（Potential）这个名词引入函数 V（1828 年）。

电流可以说已由 L．伽伐尼（1737—1798）于 1780 年发现；电池则由 A．伏打（1745—1827）于 1800 年发现。

但是，电流发现之后的最重要发现是 H．C．奥斯特（1777—1851）关于电流在磁针上的效应之发现（1820 年 9 月 11 日）。不久，J．B．毕奥（1744—1860）和 F．萨伐尔（1791—1841）发现了以他们的名字著称的“定律”（1820 年 10 月 30 日）；A．M．安培（1775—1836）于 1820 年 9 月 18 日发现了两个平行和逆平行电流之间相互作用的定律。1825 年安培于呕心沥血的系列研究之后出版了他的伟大文集。

安培给出一个电流元 I_1ds_1 在 I_2ds_2 上的力 F

$$I_1ds_1 \nwarrow \underset{r}{\overline{\qquad\qquad}} \nearrow I_2ds_2, \qquad r=r_2-r_1$$

$$F=\frac{1}{r^3}I_1I_2[(ds_1\cdot r)ds_2+(ds_2\cdot r)ds_1-(ds_1\cdot ds_2)r]$$

$$=\frac{1}{r^3}I_1I_2[(ds_1\cdot r)ds_2+\{ds_1\times(ds_2\times r)\}]$$

把 1 和 2 相互变换导致—F，因此这个 F 服从牛顿第三定律。让我们将 F 在电路 s_1 上积分，且令

$$B=I_1\int_{s_1}\frac{1}{r^3}\{ds_1\times r\}$$

B 是在 r_2 处由电流 I_1 在其电路上产生的（感应）磁场，正如毕奥—萨伐尔定律所给出的。则

$$F = I_2 \ (ds_2 \times B)$$

因此一般地说，由磁场 B 作用在 r 处的一电流元 ids 上的力可以看成是

$$F = i \ (ds \times B)$$

这就是众所周知的安培定律。

1831 年 11 月，M．法拉第发现了（电磁）感应定律①。他未经过数学训练，但他的直觉能力极强；他的场概念（力的电磁线概念）在他的思想中和他研究的理解中起了基础的和重要的作用。

法拉第定律由楞次定律（1834 年）得以“完备”。

电和磁现象的理论从早年起就包含着是一种流体还是两种流体的理论之争。在法拉第的发现之后，该问题受到了众多数学物理学家的集中注意。此处略举这些数学物理学家的名字如下：F.E.诺埃曼（1798—1895），W.韦伯（1804—1890）；H．V．亥姆霍兹（1821—1894），G．基尔霍夫（1824—1887）；开尔文勋爵；J.C．麦克斯韦（1831—1879）；H．A．洛伦兹（1853—1928）。

在其他的贡献中，“矢势”的概念已为诺埃曼（1845 年）、韦伯（1846 年）、开尔文勋爵（1846—1847）、基尔霍夫（1857 年）以及麦克斯韦（1864—1865）使用。

所有基本电磁定律的大综合则由麦克斯韦于 1864—1865 年实现。四个基本定律是（在 mksa 制即米·千克·秒·安制中）

（1）对静电学，库仑定律（1785 年）

$$\mathrm{div}D = \rho, \ D = \varepsilon E$$

（2）对静磁学，米切耳定律（1750 年）

$$\mathrm{div}B = 0, \ B = \mu H$$

（3）安培定律（1825 年）

$$\mathrm{curl}H = j + \frac{\partial D}{\partial t}, \ j = \rho v$$

（4）法拉第定律（1831 年）

$$\mathrm{curl}E = -\frac{\partial B}{\partial t}$$

这些就是为人所知的麦克斯韦场方程。麦克斯韦理论中真正新的特点是引正新的特点是引进了位移电流 $\frac{\partial D}{\partial t}$，这是出于四个方主它们与表达电路中电荷守恒的连续性方程相一致的考虑而引进的。这是具有远见卓识的——如果麦克斯韦在狭义相对论之前 40 年没有引进它，那么狭义相对论也将召唤它出现。

<u>麦克斯韦电磁波理论</u>

$$B = \mathrm{curl}A, \ E = -\mathrm{grad}\varphi - \frac{\partial A}{\partial t} \qquad (\mathrm{mksa})$$

加上洛伦兹条件

$$\mathrm{div}A+\mu\varepsilon\frac{\partial\varphi}{\partial t}=0$$

则

$$\square A=-\mu j，\square\varphi=-\frac{\rho}{\varepsilon}，\square=\nabla^2-\mu\varepsilon\frac{\partial^2}{\partial t^2}$$

在真空中，ρ＝j＝0，且

$$\square A=0，\square\varphi=0$$

这些都是简单的波方程，传播速度为

$$v=\frac{1}{\sqrt{\mu\varepsilon}}$$

μ ε 的实验值（从电磁测量所得）表明$\mu_0\varepsilon_0=\frac{1}{c^2}$，c为真空中的光速。

对于介电常数为 ε 的晶体而言，预期的光速是$c/\sqrt{\varepsilon}$，与（观察到的）折射定律相符。

光和电磁这两类现象的统一是麦克斯韦电磁场理论的最大成功之一。

然而，麦克斯韦理论不能说明光的色散现象和旋光偏振现象。该理论不是完备的，因为它并没有说到关于介质中电荷行为（运动）方面的任何东西。这只有到洛伦兹电子理论才能做到。洛伦兹假定了作用在电荷密度 ρ、质量密度 σ 和速度 v 上的（洛伦兹）力密度

f＝ρE＋〔ρv×B〕，①和运动方程

$$\frac{d'}{dt}(\sigma v)=\rho(E+〔v\times B〕)$$

洛伦兹理论和麦克斯韦理论在一起形成了被称为电动力学的理论。色散现象和旋光偏振现象只有在这个理论中才能得以理解。

电动力学的重要结果之一就是，一个被加速的电偶极矩 p＝ez 以电磁波的形式辐射能量，其辐射率

$$Q=\frac{1}{6\pi\varepsilon_0c^3}(\ddot{p})^2$$

$$=\frac{e^2}{6\pi\varepsilon_0c^3}(\ddot{z})^2$$ 在米·千克·秒·安单位制中

$$=\frac{2e^2}{3c^3}(\ddot{z})^3$$ 在高斯单位制中

因此一个振子 ez＝eacos（ωt＋b）以如下的平均速率辐射能量

$$<Q>=\frac{(ea)^2}{3c^3}\omega^4$$ 在高斯制单位中

拉莫的这一结果（1897 年）在我们接触到卢瑟福的原子结构模型时，就会看出它的重要性。

3．对麦克斯韦理论的评论

（1）麦克斯韦用数学形式表达了法拉第场的思想，并且以在时空中场 E 和 B 的四个偏微分方程概括了电磁学基本定律。电荷和电流是场的来源，并通过它们的场而相互作用，场以光速传播。H．赫兹于 1884 年在实验室中关于电磁波产生和探测的实验，证实麦克斯韦理论是确定无疑的。

（2）诺埃曼、韦伯、开尔文勋爵、基尔霍夫和麦克斯韦引入的矢量势绝不能看做一种单纯的数学技巧。借助于四维矢量（A、φ），两个三维矢量 E 和 B 被统一为一个场,这种统一的意义在相对论中和现代场论中就显得很清楚了。

E，B 的表达借助于

$$B = \mathrm{curl}A，E = -\mathrm{grad}\varphi - \frac{A}{\partial t}$$

并不唯一地定义 A 和。E、B 场在“规范变换”A，φ→A′，φ′中保持不变

$$A' = A + \mathrm{grad}\chi，\varphi' = \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}$$

这里χ（x，y，z，t）是满足下述方程的标量函数

$$\Box\chi = ，\Box = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$$

（韦耳，1918 年，1929 年）这一变换也使洛伦兹关系保持不变

$$\mathrm{div}A + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} = 0$$

“规范变换”下的“不变性”这一概念是重要的。

（3）但是麦克斯韦理论仍未解答介质以太的作用问题，即以太是“静止的”还是由“运动”物体带动的。由于在麦克斯韦理论中光是电磁波，因此，问题便是：是否有可能通过光学的或电磁学的观察来探测“相对于以太的运动”。为此目的人们做了许多尝试。

在光的情况下，我们已经提到过恒星光的光行差，菲佐实验和迈克耳孙及莫雷的实验，结果导出了不同的推论。

在电磁实验的情况下，我们有了特劳顿（斐兹杰惹的一位学生）于 1902 年所作的实验，以及特劳顿与诺布尔于 1903 年的实验。这个实验被设计来探测地球运动对一个悬空的荷电平行板电容器的效应。

结果是否定的，即观察不到由地球绕太阳作轨道运动所引起的“相对于以太的运动”。这个结果与迈克耳孙—莫雷实验的结果是一致的。

参考文献

E. T. Whittaker, A History of Theories of Aether and Electricity, Vol. 1, The Classical Theory, Nelson, London,

1951

（有关光学和电磁学原始论文的历史性说明及引证）第四章　相对论

1．狭义相对论：背景

牛顿的物理学体系明显地使用了绝对时间的概念,并且也隐含了绝对空间的概念。因此，在不同两地的两个事件的同时性是绝对的，且空间两点之间的距离 r_{12} 对所有相对运动的观察者都是相同的。对处在匀速相对运动的两个参考系（例如，沿着它们的 x 轴恒定的相对速度 v）来说，它们的时空坐标变换是

$$x'=x-vt，t'=t$$

且牛顿的运动方程（第二定律）对于两个参考系都有同样的形式。运动方程在这种变换下的不变性被称为伽利略相对性。从这一点出发得出，借助于在一个参考系内所做的动力学性质的实验,不可能判定自己这个参考系（或自己的实验室）以什么样的恒速在运动。这个相对性原理是从我们的经验中产生的。

其次，我们把问题推广到电磁（包括光）现象，并且问，在一个实验室内做实验，是否有可能判定他的实验室是否在相对于以太运动。这个问题可以用不同的方式表达 地球的轨道运动影响光的传播吗？或者水的流动影响光在其中传播吗？我们已涉及过光的光行差（布喇德雷）、菲涅耳公式和菲佐实验，迈克耳孙—莫雷的和特劳顿—诺布尔实验。我们已经看到，在这些情况中，似乎相互冲突的答案事实上都不能令人满意。

2．洛伦兹变换和爱因斯坦的贡献

洛伦兹（1892，1895）和拉莫（1900 年）探索了一种形式变换（普适化的伽利略变换），但这种变换不适用于电磁定律。最后洛伦兹在 1903 年得到一种变换

$$x' = \frac{1}{\sqrt{1-\beta^2}}(x-vt),$$

$$y' = y，z' = z，ct' = \frac{1}{\sqrt{1-\beta^2}}(ct-\beta x)，\beta = v/c$$

这适用于以恒速 v 沿着它们的 x′，x 轴相对于 S 运动的参考系 S′。如果（A_x，A_y，A_z，$\frac{i}{c}\varphi$）被定义为一个四维矢量并且像(x, y, z, ict)那样按上述同样的（洛伦兹）变换成（A_x，A_y，A_z，$\frac{i}{c}\varphi$）

①，那么麦克斯韦方程事实上是不变的，即在两个参考系中都具有相同的

形式。

因此用下面的四维矢量和张量

$$A\ (A_x,\ A_y,\ A_z,\ \frac{i}{c}\varphi),\ s\ (\rho v_x,\ \rho v_y,\ \rho v_z,\ ic\rho)$$

$$F_{\mu\nu}=\frac{\partial A_\nu}{\partial X_\mu}-\frac{\partial A_\mu}{\partial X_\nu}=-F_{\nu\mu},\ G_{\mu\nu}\sqrt{\frac{\varepsilon_0}{\mu_0 f_{\mu\nu}}}$$

$$\mathrm{curl}H-\frac{\partial D}{\partial t}=\rho\nu,\ \mathrm{div}D=\rho$$

麦克斯韦方程组

$$\mathrm{curl}E+\frac{\partial B}{\partial t}=0$$

$$\frac{\partial\rho}{\partial t}+\mathrm{div}\rho v=0$$

$$\Box A=-\mu_0\rho\nu,\ \Box\varphi=-\frac{1}{\varepsilon_0\rho}$$

分别变为

$$\frac{\partial G_{\mu\nu}}{\partial x_\nu}=s_{\mu\nu}$$

$$\frac{\partial F_{\mu\nu}}{\partial x_\lambda}+\frac{\partial F_{\nu\lambda}}{\partial x_\nu}=0$$

$$\sum\frac{\partial s_\mu}{\partial x_\mu}=0$$

$$\Box A=-\mu_0 s$$

张量成为它们的协变性的保证（详细论证请见书末附录 1）。

但确实重要的是洛伦兹的成就是相当“形式的”；变量 t′被洛伦兹称为“当地”时间，而据牛顿的绝对时间观点，t′的物理意义还完全不清楚。

1905 年，爱因斯坦（1879—1955）通过设定的下面两条原理提出了相对论的核心：

（1）所有物理定律对一切处于匀速相对运动的参考系都有相同的形式。

（2）光速对一切处于匀速相对运动的参考系都有相同的值 c。

第一条原理在那时已不是新的；庞加莱 1899—1904 年在各种场合已表达了他的信念。第二条原理则并不像它所显示的那样无关宏旨。爱因斯坦分析了在两个地点的两个事件的“时间”和“同时性”的“物理”意义，以及运动物体长度测量的“物理”意义。两只相同结构的时钟放在 A 和 B，光信号在时间 t_A 从 A 发送，在时间 t_B 到达 B，又立即反射回 A，在时间 t'_A 到达 A。t_A，t'_A 和 t_B 均按分别在 A 和 B 的时钟所示。校正 B 点的钟让两只钟同步，以致有

$$t_B=\frac{1}{2}(t_A+t'_A)$$

注意这里隐含着第二条原理！每个参考系在其系内的所有点都有同步的钟。

在两个地点 A 和 B 发生的两个事件的同时性通过下式加以定义，即

$$t_A=t_B$$

这里 t_A，t_B 是 A 和 B 处同步的钟所示的时间。

测量一个以恒速 v 运动的物体的长度，是同时采取物体的两个端点在量杆上的读数。因此，运动物体长度的测量包含着时间概念，即空间和时间的测量不是彼此无关的。

通过运用相对论的两条原理，爱因斯坦以一种非常简单的方式推出了洛伦兹变换。

从洛伦兹变换出发，很快就得出下面的结果$(\beta=\frac{v}{c})$：

（1）斐兹杰惹—洛伦兹收缩

$L=L_0\sqrt{1-\beta^2}$，L_0为本征长度

（2）时间膨胀

$\triangle t=\frac{\triangle\tau}{\sqrt{1-\beta^2}}$，$\tau$为本征时间

（3）多普勒效应

两个参考系相互分离和接近，所接收的频率分别为

$\nu=\sqrt{\frac{1-\beta}{1+\beta}}\nu_0$，$\sqrt{\frac{1+\beta}{1-\beta}}\nu_0$，$\nu_0$为本征频率

（4）恒星光的光行差

$\cos\theta=\frac{\cos\theta'+v/c}{1+(v/c)\cos\theta'}$，像前面一样

（5）速度相加

$u_x=\frac{u'_x+v}{1+\boldsymbol{\beta}u'_x/c}$，$u_y=\frac{u'_y+v}{1+\boldsymbol{\beta}u'_x/c}$，

$u_z=\frac{u'_z+v}{1+\boldsymbol{\beta}u'_x/c}$，

（6）菲涅耳公式

$$u=\frac{\frac{c}{n}\pm v}{1\pm(v/cn)}=\frac{c}{n}\pm(1-\frac{1}{n^2})v$$

3．相对论性动力学

让我们再回到动力学。在这一点，情况如下所示：

变换	牛顿动力学	麦克斯韦方程组
伽利略	不变	非协变
洛伦兹	非协变	不变

重新表述牛顿动力学已为必需。令 τ，V_0，ρ_0 是本征时间、体积、密度，且 τ，V，ρ 是它们在以速度 v 运动的一物体的参考系内的值，则

$$dt=\frac{1}{\gamma}dz,\ dV=\gamma dV_0,\ \rho=\frac{1}{\gamma}\rho_0,\ \gamma=\sqrt{-(v/c)^2}$$

令

$$u(\frac{v_x}{\gamma},\frac{v_y}{\gamma},\frac{v_z}{\gamma},\frac{ic}{\gamma})\text{是一个四维速度，}$$

$$(u\cdot u)=-c^2$$

且 f_μ 是四维矢量 ρ

$$\rho_0\frac{du_\mu}{d\tau}=F_\mu,\ \mu=1,2,3,4$$

则

$$m_0\frac{du_\mu}{d\tau}=F_\mu,\ F_\mu=\frac{1}{\lambda}\int f_\mu dV$$

且

$$(F\cdot u)=\frac{m_0}{2}\frac{d}{d\tau}(u\cdot u)=\frac{m_0}{2}\frac{d}{d\tau}(-c^2)=0$$

定义一个三维力（K_x，K_y，K_z）

$$(K_x,K_y,K_z)=r(F_x,F_y,F_z)$$

此时（F·u）＝0 变成

$$\frac{1}{\lambda^2}(k\cdot v)+F_4u_4=0$$

或

$$\frac{d}{d\tau}(\frac{m_0c^2}{\sqrt{1-\beta^2}})=(k\cdot v)$$

即，动能 T 的变化率等于作用在物体上力 K 所做的功率

$$T=m_0c^2(\frac{1}{\sqrt{1-\beta^2}}-1)=(m-m_0)c^2$$

此处 $m=\frac{m_0}{\sqrt{1-\beta^2}}$

这些都是为人所熟知的结果（详见书末附录 2）。

运动方程$m_0\frac{d}{d\tau}u_\mu=F_\mu$**洛伦兹变换下是不变的。**

爱因斯坦的巨大贡献，并不在于相对性原理，而主要在于他对空间和时间基本概念本身的基础性分析。他强调用概念的测量程序定义概念——所谓的“操作观点”——在当时的物理学中是全新的观念。

牛顿的绝对空间和绝对时间被四维时空所取代；dr^2 和 dt 的不变性现在由 $ds^2=(cdt)^2-dr^2$ 的不变性所取代。闵可夫斯基的四维空间仍然是平直空间。机械力和电磁力并不像爱因斯坦“引力理论”中的引力那样影响四维空间的几何性质。

4．广义相对论

狭义相对论要求，所有物理定律对一切处于匀速相对运动的参考系都具有相同的形式。这一点在数学上是用物理定律在洛伦兹变换下的不变性要求表示的。

在广义相对论中，这个原理被推广到处于任意相对运动的所有参考系。这一点的数学表达是一切物理定律在四维时空内任意坐标变换下的不变性要求。

$$x'_\mu = x'_\mu(x_1, x_2, x_3, x_4), \quad \mu = 1, 2, 3, 4$$

广义相对论就是以这种形式按人们的哲学信念或口味被接受或被拒绝的。

5．爱因斯坦的“引力理论”

广义相对论的一个非常有趣和意义深远的应用是在 1905 年与 1915—1916 年间发展起来的爱因斯坦的“引力理论”。

爱因斯坦从牛顿动力学和他的万有引力定律开始。为了清楚起见，让我们将其思路表示如下。

（1）牛顿使用了一个三维空间和一个独立的时间，而三维空间有欧氏几何的性质。按照伽利略—牛顿运动定律，一物体不受力作用，要么保持静止，要么以恒速保持（在三维空间内）直线运动。

但当行星作椭圆运动时，就能得出，必定有力作用在行星上，并且这些力被设定为来自行星和太阳之间的万有引力。

现在，虽然我们的日常经验强烈地提示出欧氏几何，但它毕竟只是一种抽象的几何学。在物理学中，是否有可能空间的几何性质不同于欧氏几何？在物理学中，是否有可能空间（或几何学）的性质会依赖于物质的存在？太阳能否给空间几何学造成影响，以致它不再是欧氏几何的？一个检验物体（行星）能否在此空间内简单地按最短路线（“直”线）运动，且根本就没有任何“引力”存在？（注意，这里涉及的“空间”一词是指四维空间！）

（2）用“无引力但有非欧四维空间”的新选择代替牛顿“有引力的欧氏空间”观点的上述思想可以描述为“引力场的几何化”；但这个思想取决于一个重要的要素，即对所有物体来说，一物体的“引力质量 m_g”必

须具有与它的“惯性质量 m_1”相同的关系。这个要素幸好由经验结果所满足，爱因斯坦称之为“等价原理”。

一个加速度的效应等价于一重力在反方向上的效应,这种在电梯内的感受就是等价原理的一个基本范例。这个等价原理意味着，比率 m_g/m_1 对所有不同种类的物质都是相同的。这一点已为伽利略和牛顿的早期观察指出得十分清楚，以致一个单摆周期只取决于其长度，而与形成摆锤的物质本性无关。但是，深层的意义在爱因斯坦之前并没有为人们所充分了解。m_g/m_1 对所有不同种类物质是相同的这一点，已为厄缶（1891 年）和迪克（1961 年）所证实。后者所作实验的精确度为 $2/10^{12}$。

（3）根据这个等价原理，一个均匀引力场在负 Z 方向上的效应，能够用在自由空间中正 Z 方向上的匀加速参考系来“变换”掉。但是，一个引力场，例如太阳的引力场每一点都不同，不可能如此简单地被变换掉。引力场的“几何化”一般需要微分几何学。

在这种“几何化”理论之前，我们将以一种基本的方式来看一看一个加速度是怎样影响空间和时间测量的，再看看通过等价原理，一个引力场又是怎样影响空间和时间测量的。

试考虑一个惯性系 $_0$，以及一个以角速度ω旋转的旋转游戏装置（S），它的边缘在地面上画出一个半径为 R 的圆。S_0 内的一位观察者 A，用一根米尺，测量圆的圆周，得到 2πR。令一观察者 B，用一根米尺在 S 上测量同样的圆周。对 A 来说，观察者 B 以一个（横向）

速度$v = \omega R$运动，并且他的米尺显示出以$\sqrt{1-(\omega R/c)^2}$因子发生收缩。于 A，B 的测量结果将是

$$\frac{\text{一个圆的圆周}}{R} = \frac{2\pi R}{R\sqrt{1-(\omega R/c)^2}} > 2\pi$$

以致 B 的几何学必须是非欧几里得的，偏离 2π的比率随 R 连续地改变。

考虑一个惯性系 S_0 和一个以加速度 a 向上加速的参考系 S，

当 t=0，O ' 从静止的 O 出发，并向上发送一频率为 v■的光束。

在 t，光束在 S_0 内到达一高度 h=ct，在 S 到达 P 点，其（上行）速度相对于 S_0 是 v=at=ah/c。

对 S 内 P 点的观察者，因为多普勒效应，光的频率是

$$\nu_p = \sqrt{\frac{1-(ah/c^2)}{1+(ah/c^2)}}\,\nu_0$$

按照等价原理，P 点比 O 处在一较高的引力势。因此在 O（低引力势）的一只钟比在 P（高引力势）的一只全同的钟要慢。

上述近似的考虑在爱因斯坦 1916 年论述完整理论的论文之前的早期论文（1905～1911 年）中就已有了。

（4）为了使引力“几何化”，爱因斯坦求助于黎曼空间，它的度规

是 $ds^2 = g_{\mu\nu} x^{\mu} dx^{\mu}$

$g_{\mu\nu}$完全决定了空间的几何性质。例如，黎曼—克里斯托费耳张量 $R_{\lambda\mu\nu\rho}$是 $g_{\mu\nu}$的二次导数的函数，给出了空间的“曲率”。“平直空间”的条件是在所有各点 $R_{\lambda\mu\nu\rho}=0$。

在 1916 年的理论中，爱因斯坦设定黎曼空间中的短程线是在那个空间中一个检验物体的运动方程。在后来的一个理论中（英费耳德和霍夫曼提出，1938～1940 年），发现这个假定是不需要的，运动方程已经包含在“场方程”中了。

场方程是理论中把物质在时空内的分布与四维黎曼空间的度规性质联系起来的基本假定。爱因斯坦假定了如下关系

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -kT_{\mu\nu}$$

此处 $R_{\mu\nu}=g^{\alpha\beta}R_{\mu\nu\alpha\beta}$，且，而 $T_{\mu\nu}$是代表物质分布的能量—动量张量。这个方程起着泊松（和拉普拉斯）方程在牛顿引力理论中的作用。这是一个 $g_{\mu\nu}$的非线性偏微分方程组。

下页表给出了牛顿理论与爱因斯坦理论之间的比较。

爱因斯坦理论中的场方程在一般情况下是无解的。一个由于一重质点产生的静态的、球对称场的特殊情况——史瓦西问题——已有精确的解（请见书末附录 3）。

（5）1917 年，爱因斯坦把他的理论应用到宇宙学问题。他试图得到他的场方程

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -kT_{\mu\nu}$$

牛顿理论	爱因斯坦理论
绝对空间和绝对时间	
（3+1）维空间	四维时空
平直的	弯曲的
三维空间中的直线	弯曲空间中的短程线
引力势 V	度规张量 $g_{\mu\nu}$
运动方程	短程线方程
时空中行星轨道（椭圆）	四维空间中的短程线
引力定律	场方程
$\nabla^2 V = -4\pi\rho$	$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -kT_{\mu\nu}$
或	
$V = -\frac{kM}{r}$	

的一个解，它在均匀的、各向同性的和静态的宇宙中的一个解，但却找不到。对每个地方 $T_{\mu\nu}=0$ 的情况下存在一特殊解，即

$$g_{\mu\nu}=常数$$

这隐含着即使在宇宙中不存在任何物质（能量和动量），“引力”场可能依然存在。这时爱因斯坦采用马赫原理：如果宇宙中没有物质存在，也就根本不可能有任何时空或惯性。爱因斯坦（1917 年）给他的场方程增加了一个宇宙项① $R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R_{\mu\nu}-\lambda g_{\mu\nu}=-kT_{\mu\nu}$

这导致了单一重质点在 r=0 的情况下的度规(见书末附录 3 中谈及史瓦西处）：

$$ds^2=(1-\frac{2M}{r}-\frac{\lambda}{3}r^2)(cdt)^2-\left\{\frac{(dr)^2}{1-\frac{2M}{r}-\frac{\lambda}{3}r^2}+r^2d\theta^2+(r\sin\theta d\varphi)^2\right\}$$

这被称为爱因斯坦宇宙。λ项的结果是引入了大距离上的斥力。对于空的空间（即上述 M=0），度规变为

$$ds^2=(1-\frac{\lambda}{3}r^2)(cdt)^2-\left\{\frac{(dr)^2}{1-\frac{2M}{r}-\frac{\lambda}{3}r^2}+r^2d\theta^2+(r\sin\theta d\varphi)^2\right\}$$

这由德西特于 1917 年找到。它是一个具有恒定曲率的赝球形世界。

勒梅特（1927 年）指出，爱因斯坦宇宙（带有λ项）是一个不稳定的宇宙，当从静态扰动时，就连续过渡到德西特宇宙。

A．弗里德曼（1922 年）取得了一个重要进步，他指出，爱因斯坦最初的理论（不带λ项）具有一个解，该解是均匀的，并具有与时间有关的曲率。这项工作最初未引起更多的注意；但到 1929—1930 年，哈勃膨胀宇宙的发现为人所知了（从退行星云的钙吸收线的多普勒红移中获知的）。勒梅特的老师爱丁顿在勒梅特的工作中认识到哈勃的观察与爱因斯坦最初理论间的关系。事实上这是现代宇宙学的一个重要时代。1931 年，爱因斯坦抛弃了λ项,并且说λ项的引进是他一生所做的工作中最坏的错误。

6.对爱因斯坦引力理论的评论

（1）我们已经看到，闵可夫斯基四维空间是动力学定律和电磁学定律的基本框架结构。因此电磁力不要求对无电磁场时有效的平直四维空间的修正。在爱因斯坦引力理论中，“引力场”以四维空间的度规性质显示出来，即以 $g_{\mu\nu}$ 的度规性质显示出来，以致四维空间是一个弯曲的（黎曼）空间。因此物质的存在及其分布决定四维空间的度规性质，引力相互作用是通过度规性质而起作用的。这与电磁场不同，电磁场并不通过四维空间的度规性质作用在电荷上。

（2）当一给定的 $T_{\mu\nu}$ 分布决定度规 $g_{\mu\nu}$ 时，一切变换都在同一空间内。某些量，例如 ds_2 和标量曲率 R 是不变的。不存在从一弯曲空间到一平直空间的变换，反之亦然。

虽然引力场要求弯曲空间,但由此并不能推得加速度必然要求弯曲空间。一个例子是由莫勒（1943 年）、吴和李（1972 年）所给出的被加速参考系，它由一平直空间描述，并且从平直空间出发，一族被加速的运动能够通过变换生成。

参考文献吴大猷：相对论（理论物理第四册），科学出版社，北京，1983 年

（狭义相对论，微分几何，爱因斯坦引力理论）W.Pauli,Theory of Relativity,Pergamon press,London,1958
A. Pais，“Subtle is the Lord…”,Oxford University Press，1982

（许多历史笔记）

S. Weinberg,The First Three Minutes，Basic Books，Inc. Publ，N. Y,1977

（宇宙学）C. Mφller，Danske，Vid. Sel. Mat—Fys. Med. XX，No. 19，1943

T.Y. Wu，and Y. C. Lee，International Journal of Theoritical Physics,5，307，1972

（时钟或双生子佯谬）第五章　相对论对现代物理学哲学的影响

1900 年，普朗克建立了量子理论，由此导致了于 1924—1926 年由一批物理学家创立的量子力学。1905 年，爱因斯坦建立了狭义相对论，1915 年建立了广义相对论。这两门学科——相对论和量子力学——形成了现代物理学整个结构的基石。

我们在本章将力图为相对论的物理本质和哲学本质提出一个简要的说明。

1．变换概念和不变性概念

每一条物理定律，诸如开普勒行星运动定律或牛顿万有引力定律，都是物理概念之间的一种关系。在上面提到的诸定律中，包含的概念是空间、时间和质量等概念。一个定律不应该依赖于空间坐标系的特殊选择，或时间坐标的特定初始时刻的选择。现在人们能直接思考下面这些空间和时间坐标的变换：

（1）空间坐标系的平移，

$$r \to r', \quad r' = r + d$$

这里的 d 是一种恒常位移（constantdisplacement）

（2）空间坐标系的转动，

$$r \to r', \quad r' = Rr$$

这里的 R 是一个“算符”，它可以表达成一个矩阵，即：

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

并有

$$\sum_k a_{ik} a_{jk} = \delta_{ij}$$

（3）空间坐标的反演，

$$r \to r', \quad r' = P, \quad r = -r$$

这里的 P 被叫做宇称算符。

（4）时间坐标的平移，

$$t \to t', \quad t' = t + t_0, \quad t_0 \text{是常数}$$

（5）时间坐标的反演，

$$t \to t', \quad t' = -t$$

物理定律应独立于像这些坐标系选择的情况是用“变换下的不变性”概念来表达的。与物理定律在某种变换下的不变性相联系，有一个守恒定律，它是不变性性质的一个推论。下表列出其中的一些关联。

不变性	守恒定律
哈密顿算符在下述变换下的	
空间平移	动量的
空间转动	动量矩的
空间反演	宇称的
时间平移	能量的
在规范变换和能量守恒下的	
电磁定律的	电荷的

变换和不变性概念并不是新概念,但它的重要性也许只有在相对论发展后才充分体现出来，这一点我们将在下面看到。

2．洛伦兹变换和相对性概念

在经典动力学中，很容易看出，当我们相对于另一个参考系变换运动方程，这个参考系相对于原来的参考系作恒速运动时，运动方程（牛顿第二定律）保持不变，即：

$$x' = x - vt, \quad v = \text{常数}$$

在这种所谓伽利略变换下运动方程的不变性构成了一个相对性原理。它

意味着经典动力学在彼此作匀速相对运动的参考系中并无区别。

试图将这个相对性原理扩展到电磁现象并不成功,即麦克斯韦电磁场方程在空间、时间坐标的伽利略变换下并不是不变的，而牛顿方程在这种变换下则是不变的。这种形式的、数学的问题的解就是洛伦兹变换

$$x'=\frac{1}{\sqrt{1-\beta^2}}(x-vt)，t'=\frac{1}{\sqrt{1-\beta^2}}(t-\frac{\beta}{c}x)$$

这里的β=v/c,c 是光速，它是一个常数,v 是最初的参考系相对于其后的系的恒定速度。在爱因斯坦 1905 年相对论论文发表之前的 1902—1903 年，这种变换已经发现。然而，这些变换方程的意义并不清楚，虽则在极限情况β→0 时，这些变换的确可以简化为伽利略变换。奥秘就在 t'；因为在牛顿力学时代，事实上在自然哲学发展的所有时代，只有一种时间，那就是绝对时间。按照牛顿的理论，在宇宙中，时间是均匀地并与其他任何东西无关地流逝着。不难发现，这样的一种时间概念确有不能令人满意之处，因为没有参考物，均匀流逝的概念是不清楚的。牛顿和其他哲学家已经意识到了这一点，但事实上依然是始终找不出关于物理学中时间概念的更好表述，直到爱因斯坦！

洛伦兹的形式变换，当它满足电磁的不变性要求时，就不再使牛顿的动力学方程保持不变了。因此，就有了以下的情形：

	牛顿运动定律	麦克斯韦场方程
伽利略变换	不变	不再不变
洛伦兹变换	不再不变	不变

这种局面确实是令人非常不满意的。

3．爱因斯坦的空间、时间“操作”定义和相对性原理

爱因斯坦通过重新考察物理学中空间和时间的基本概念本身的性质来处理前面提出的问题。他强调了早先由恩斯特·马赫所表达的这样一个观点，即：在物理学中概念必须基于它的测量才有意义。爱因斯坦由此作了如下分析：

（1）为了测量相对于观察者的一个静止物体的长度，读出一根米尺上的终端标记这种通常程序是清晰明了的。为了测量对一个静止的观察者来说发生在同一地点的两个事件之间的时间间隔,读出在那同一地点上一只钟的时间差（time off）这种通常程序也是清晰明了的。

（2）测量在两个不同地点两个事件发生的时间要求有一个确定的约定以安置或校正两地的时钟。校正时钟的一个方便的程序如下。从 A 点在时间 t_A（A 点时钟上的时间）送出一束光给 B 点，在时间 t_B（B 点时钟上的时间）到达 B 点，而在 B 点通过一面镜将光束返回 A 点，到达 A 点的时间是 t'_A。A 点和 B 点的时钟都加以“校正”，使其读数有这样的关系

$$t_B = \frac{1}{2}(t_A + t'_A)$$

因此，在 A 和 B 两个事件的“同时性”就由在 A 和 B 的时钟读数使 $t_A=t_B$ 来加以定义。

（3）现在，为了测量相对于观察者以速度 v 运动的物体的长度，我们制定这样的标准，即在米尺上的客体（客体是静止的）的两个终端标记的读数必须按照相对于观察者校正的 A 和 B 两点上的时钟同时读出，即 $t_A=t_B$。这个同时性只是对那个观察者而言的；对于身处物体之中的观察者来说，按照他自己使用相同的标准校正过的时钟，上面所述的两个终端标记的读数并不是同时的。因此，同时性概念是“相对的”概念，它不具有绝对的意义。

（4）从上面的分析可以看出：①经典的绝对时间在物理学中是没有意义的；②空间和时间概念不是独立的，它们通过它们的测量定义彼此联系在一起。

与空间和时间这些概念一起，爱因斯坦引入了相对性以下的问题，如：

（Ⅰ）斐兹杰惹“收缩”和迈克耳孙—莫雷实验；

（Ⅱ）时间膨胀，多普勒效应；

（Ⅲ）速度相加；

（Ⅳ）麦克斯韦方程在洛伦兹变换下的不变性；

（Ⅴ）相对论性动力学，等等。

以上都是洛伦兹变换的一些简单而直接的推论。它们现在对每个物理系的学生都是熟悉的。对这些内容我们就不再在这里作深入论述了。

相对性原理，即物理定律对所有作匀速运动的参考系都具有相同形式，确切地说，并不是新的，它在此前已为 H．庞加莱猜测到了。洛伦兹变换的数学公式当然也在不久前为洛伦兹建立了。但是，对空间和时间概念的分析以及它们的操作定义乃是关键性的步骤,它们给了形式的洛伦兹变换以物理意义，在这些方面，爱因斯坦的贡献是巨大的。不幸的是，这一点并没有被广泛地理解。例如，著名的应用数学家 E．T．惠塔克，在他的《以太和电的理论的历史》一书中，写到了洛伦兹和庞加莱的相对性理论。在差不多 50 页的篇幅中，只有无关紧要的三处提到爱因斯坦的名字。它只采纳了洛伦兹本人所承认的爱因斯坦对这个问题的关键性贡献。

相对论的哲学方面，即强调对物理学中基本概念的批判分析，以及这些概念定义的操作观点，其影响是非常巨大的。正像我们在下面将要看到的，这种态度正是海森伯新量子力学出发点的指导精神，而且事实上在量子力学的哥本哈根哲学中起了重要的作用。但是，出乎意料的是，正是这种量子力学中的哲学后来成了爱因斯坦本人所不能接受的了。

4．广义相对论

概而言之，广义相对论就是使相对性原理从惯性系（即匀速相对运动的系，其中力学定律均适用）推广到任意运动的系。这意味着在任意的时空坐标变换下物理定律保持不变。以数学形式表达，用洛伦兹变换描述的狭义相对论使四维空间间隔 ds_2 保持不变，即，

$$ds^2 = dx_1^2 + dx_2^2 + dx_2^3 + dx_2^4 = dx'^2_1 + dx'^2_2 + dx'^2_3 + dx'^2_4$$

其中 $x_1=x$，$x_2=y$，$x_3=z$，$x_4=ict$。这是对普通三维空间的一种普遍化，三维空间间隔为

$$ds^2+dx^2+dy^2+dz^2$$

现在在微分几何学中，我们已有了一种普遍的非欧（黎曼）空间，对于这种空间，我们已讨论过全等性（congruence）概念是不适用的。ds^2 一般地由以下公式给出：

$$ds^2 = \sum g_{\mu\nu} dx^{\mu} dx^{\nu}$$

这里的 $g_{\mu\nu}$ 是 x ' s 的函数。张量 $g_{\mu\nu}$ 被称做度规张量，由它定义空间几何学。一个 ds_2 能够通过坐标变换而变换成对所有的点 x_{μ} 有形式

$$ds^2 = \sum dx^{\mu 2}$$

的空间被说成是平直空间。欧氏空间就是一种平直空间。如果 ds_2 不能被变换成这种形式，那么这种空间就可以被说成是弯曲空间。如果从这些定义中得出空间的平直或弯曲性质是空间本身的一种固有属性,那么也就能得出,不存在能够把一个弯曲空间变换成一个平直空间或者反过来也一样的坐标变换。用数学语言来说，一空间的曲率是由一张量描述的。如果曲率张量在一个坐标系中是恒等地非零的,那么它在坐标变换中依然是恒等地非零的。

在物理学中,用张量方程的形式表达的定律是按定义张量的坐标变换从一个参考系变换到另一个参考系的。这些变换把相同空间中的不同参考系联系起来。洛伦兹变换把赝欧氏四维空间中的不同惯性系联系了起来。但从一个惯性系出发，对于平直空间中的被加速的系，至少存在一个参数变换群（这是由 C. 莫勒、吴大猷和李荣章发现的）。因此，在平直空间有“ 广义相对论 ”是可能的，虽然对于任意被加速的系来说涉及弯曲空间。

爱因斯坦广义相对论认为物理定律在所有任意运动系中都是不变的（即具有相同的数学形式）。这个陈述的分量在于“ 相同数学形式 ”一词。为此目的所用的合适数学工具是张量微分。

一个张量的变换性质（其中矢量是一特例）是由坐标的变换定律定义的。一个物理量必须具有在坐标变换下的变换性质，并且总能用张量形式加以表达。一条物理定律是物理量之间的一种关系，它能被铸成一个张量方程的形式。一旦这样做了，它在坐标变换下的不变性（为此需首先定义张量的量）就自动地得到保证（一个基本的例子是麦克斯韦场方程的洛伦

兹不变性，方程都能以张量形式表达，所有场量都是按照洛伦兹变换的张量变换）。在某种意义上，狭义相对论普遍化达到任意运动系是一个大胆的，但却是十分自然的步骤。

向量代数是吉布斯在相对论之前发明的，张量在电磁学理论和连续介质物理学中已被应用。甚至张量演算——所谓绝对微积分——也已在广义相对论之前的微分几何中得到发展了。但是张量在物理学中使用的充分重要性只有到了相对论才被充分认识。在相对论中，它不是单纯的优美形式体系的问题；它是处理与广义坐标变换有关的协变性概念的自然数学语言。

5．爱因斯坦的引力理论

爱因斯坦引力理论的根源可以说基于他这样的认识：对于观察到的物理现象来说，参考系的加速度效应是与引力场的效应相同的。伽利略对教堂里的单摆和比萨斜塔上的自由落体的观察已经给出一个暗示，即一个物体的“惯性质量”与“引力质量”是等价的；而厄缶 1890 年和 1909 年的精确测量确立这种等价值达 $1/10^8$ 的精度。R．迪克 1961 年的更近期的测量确立这种等价值达 $5/10^{12}$。1905 年，爱因斯坦提出了“等价原理”，即对在一个被加速的参考系中的物理现象的描述与对在引力场中的一个惯性系内的物理现象的描述是等价的。从这一点出发，就产生了这样的思想，即按照牛顿理论在一个引力场中的运动可以被看做是在一个适当的加速系中的“自由运动”（即无引力场）。第二步就是用一个四维弯曲空间来描述这个加速系，四维弯曲空间的度规

$$ds^2 = \sum g_{\mu\nu} dx^{\mu} dx^{\nu}$$

代表任意时空变换（即洛伦兹变换不再限制在平直空间）。

在深入讨论爱因斯坦引力理论以前，让我们先来回忆一下经典物理学基础中的空间概念。在那里，三维空间直观上被当作是欧氏空间（即“平直的”），而时间是绝对的。甚至当时间与空间合并成狭义相对论中的四维空间时，这四维空间依然是平直的（或是赝欧几里得的，因为在 $ds_2=dx_2+dy_2+dz_2-c_2dt_2$ 中的符号不同）。在经典动力学中，为了描述平直时空中的行星（和其他）运动，设定一个引力场是必需的；因为若无各种引力，运动就将直线进行。这就是牛顿理论。

爱因斯坦提出了一个新观点。为了代替欧氏空间是物理空间这一先验假定，爱因斯坦提出，物理空间不是一种抽象空间，而是受物质（能量）所制约的，即物理空间有一种为度规张量 $g_{\mu\nu}$ 所规定的几何，它本身受宇宙中物质（能量）的分布所支配。这种几何学可以说已经包含了物质分布的性质，而且空间在微分（或仿射）几何学的意义上是被弯曲了的。在这种空间中的自由运动就取代了“在欧氏空间中的引力场中的运

动”。

我们将对爱因斯坦引力理论和牛顿理论中的相应思想作一个最简单和最清晰的对比。

牛顿理论	爱因斯坦理论
无引力场的情况	
平直三维空间，欧氏几何学引力势φ	平直四维空间，赝欧氏几何学，度规张量 $g_{\mu\nu}$
φ=常数	$g_{\mu\nu}$=常数
有引力场的情况	

牛顿理论	爱因斯坦理论
平直三维空间，欧氏几何学 引力 $= -\nabla\varphi$	弯曲四维空间，黎曼几何学克里斯托费耳3指标符号$\begin{Bmatrix} k \\ lm \end{Bmatrix}$，其中的40，包含了$g_{\mu\nu}$的一阶导数
在自由空间中	
引力定律 $\nabla^2\varphi = 0$	“引力”定律 $G_{\mu}^{\nu} = 0$ G_{μ}^{ν}是一个张量，包括3指标符号 方程是非线性的
在有物质的空间中	
引力定律 $\nabla^2\varphi = 4\pi k\rho$ 这里的ρ为物质密度。方程是线性的 一已知分布ρ按上面的泊松方程决定引力场φ	“引力”定律 $G_{\mu}^{\nu} = -kT_{\mu}^{\nu}$ 这里的T_{μ}^{ν}是物质（能量—动量）张量。方程是非线性的 宇宙中物质的已知分布 T 决定空间的几何学，即度规张量$g_{\mu\nu}$

牛顿理论	爱因斯坦理论
物质守恒定律：$\frac{\partial\rho}{\partial t} + \mathrm{div}(\rho v) = 0$	能量—动量守恒定律：$T_{\mu;\nu}^{\nu} = 0$ 这里的 T 是物质张量密度。分号“；”代表协变导数
粒子的运动方程：$m\frac{d^2 r}{dt^2} = -\nabla\varphi$	短程线方程：$\frac{du^4}{ds} + \begin{Bmatrix} k \\ lm \end{Bmatrix} u' u^m = 0$

有时，人们把爱因斯坦的引力理论说成是几何化的理论，即用一种几何学理论（在弯曲时空中的自由运动）代替牛顿的动力学理论（在平直欧氏时空中引力场下的运动）。人们还认为，爱因斯坦理论始终是一种对描述的纯形式改变，理论是有意义的，但并不真正深奥。然而，爱因斯坦理

论并非对描述的纯形式的改变；它导致了三个著名的问题，并对其作出了与牛顿理论不同的预言，这就是引力红移，太阳（或任何大质量的物体）引起光的弯曲，以及水星近日点的进动。所有观察似乎都与爱因斯坦理论的预言相一致。

6．评论

（1）对太阳光谱的光谱线红移的早期测量与实验室光谱线($\frac{\triangle\lambda}{\lambda}=2\times10^{-6}$相比较，在定性上与理论相符，但还不是结论性的，因为要把其他扰动因素，诸如斯塔克效应和湍流分离开来尚有困难。通过极其灵敏的穆斯堡尔效应对引力红移（与狭义相对论效应相区别的效应）所作的更新近的测量表明几乎与理论完全相符（邦迪和里布卡，《物理评论》1960 年第 4 期第 337 页）。

众所周知的如时钟佯谬或双生子佯谬这一效应，是与等价原理连在一起的。这个在平直时空中的广义相对性问题已由 C．莫勒（Danske Vid．Sel．Mat-Fys．Med．XX，No.19，1943）和吴大猷、李荣章（国际理论物理学杂志，1972 年第 5 期第 307 页）给出了精确的论述（请见书末附录 4）。

（2）对通过太阳边缘附近的（遥远星体的）光的弯曲的早期测量是在日全食期间用感光板做的，这些测量做起来非常困难，但整个结果表明与理论符合得极好。新近的测量是用雷达波通过太阳的边缘并从太阳以远的一个物体反射回来。结果看来与理论相符。

（3）水星近日点进动的观察数据为每百年 43．11″±0．45″，与理论的预言 43．03″相符。然而，这个观察值是在对许多扰动效应作了校正以后从一个大数值 5160″那里得到的小剩余数。近几年来，对这个小进动已提出了其他的解释（迪克等人）。

（4）对称度规张量 $g_{\mu\nu}=g_{\nu\mu}$只有十个独立的组元。自 1928 年起的许多年中，爱因斯坦，后来还有薛定谔，做了许多尝试，力图将电磁现象合并到爱因斯坦的引力理论中去，但没有成功，因为那需要另外六个独立的参数，才能与电磁场相对应。M．萨克斯 1968 年在台湾夏季科学讨论班的讲演中提出了一个理论，他在该理论中保留了爱因斯坦所有的基本思想，但用两个四元数的乘积代替每个 $g_{\mu\nu}$，使得度规张量具有 16 个独立的数。这一理论能给出爱因斯坦最初的理论所能给出的一切结果，并且还更多一些。我相信这个理论应比已被接受的理论更值得重视。

第六章　对称、变换和不变性

前几章，我们简略评述了经典动力学、电磁学和相对论，我们已经知

道（虽然没有明确地强调）物理定律在某些变换下的不变性概念，例如整个电磁场方程组在洛伦兹变换下的不变性。这些概念与对称概念有关，也与被称做“守恒定律”的定律有关，虽然其中的一些概念并不全是“新的”，但它们在现代物理学中比在经典物理学中有更加重要的地位。

关于对称、变换、不变性和守恒定律以及它们之间的关系的讨论，我们可以从一些基本范例开始：

（1）对于具有球对称的物理系统，诸如太阳的牛顿引力场，物理定律在坐标系对于原点的旋转下是不变的。这一不变性导出了角动量守恒定律（诸如开普勒的第二定律）。

（2）对于在空间一确定方向有平移对称性的动力学系统，动力学定律在此方向上坐标系的线性平移下是不变的，由此导出在此方向上的动量守恒定律。

（3）对于时间上具有平移对称的系统，即在时间平移下不变的系统，存在能量守恒定律。

（4）在经典动力学中，有坐标和共轭动量的正则变换，在此变换下，运动的（正则）方程形式不变。即其积分不变量为守恒量。

（5）对于匀速相对运动的系统，物理定律在洛伦兹变换下是不变的。这就给出了（或者说表达了）相对性原理。一个不变量是 ds^2，它是两点

$$(x_2, y_2, z_2, ct_2) 和 (x_1, y_1, z_1, ct_1)$$

之间距离的平方。

（6）宇称（P）对称（反演对称）

反演运算（In）是一种其中矢量 r（x，y，z）变为－r（－x，－y，－z）的运算。如果一个系统具有球对称，则它们的物理定律是反演不变的。宇称操作是一个平面上的反射。因此在 X—Y 平面上的一个反射（一个镜像）等价于一个反演加上绕 Z 轴旋转一个角 π。如果旋转对称被设定，那么反演对称也隐含宇称对称，反之亦然。

让我们考虑电磁定律。麦克斯韦方程对反演是不变的。这是因为电场矢量（极矢量）E 改变符号，而磁感应强度 B（一个轴向矢量或赝矢量）的符号保持不变①。电磁定律在反演（和宇称）操作中实际上不变这一点已在实验上以拉波特定律（1924 年）的形式确立，拉波特定律指出，电偶极辐射只在偶宇称态和奇宇称态之间跃迁时被一个原子所发射或吸收。这个定律也许更强有力地加强了一个直观感觉，即自然定律必须具有这种左右对称性。

因此，当李和杨对弱相互作用（在 β 衰变和 π－μ 衰变中）中这种宇称不变性可能不成立的设想（1956 年）几乎立即由吴健雄等人（1956—1957 年）在实验上得到证实时，人们才认识到宇称守恒并非一个普遍定律。这个发现确实对于差不多所有物理学家来说都是出乎意料的①。

（7）电荷共轭（C）对称

随着 1932 年正电子的实验发现和 1955 年负质子的实验发现，人们知道了有反粒子存在。由此提出了问题，物理理论在“电荷共轭”变换下是否不变，即粒子变为它们的反粒子（或反粒子变为粒子）时物理理论是否是不变的。

吴健雄等人关于宇称不守恒的实验表明，当 P 守恒遭破坏时，组合的 CP 对称在他们的实验中仍然是守恒的。

（8）时间反演（T）对称性

动力学定律和电磁学定律在时间 t 变为—t 时都被看成是不变的。在量子力学中，薛定谔方程也具有这种时间反演对称性（如果连同时间反演，人们做复共轭运算）。但是随着弱相互作用宇称不守恒的发现，提出时间反演不变性是否也可能会破坏的问题，就变成得当的了。

现在从实验上看，使时间反演（即使可能）也是不易的。但是存在一个定理，即组合对称性（CPT）将在某些一般条件下保持。按此定理，在时间反演下守恒或不守恒能通过对组合 CP 对称性的实验加以检验。

事实上，正是用了这种办法，克里斯坦森、克隆宁、菲奇和特雷（1964年）关于$K_2^0 \to \pi^+ + \pi^-$衰变的实验表明T反演对称也不是守恒的②。

（9）规范对称

第三章中，我们已经看到，电磁定律在四维势规范变换下是不变的（H．韦耳，1918，1929 年）

$$A' = A + \mathrm{grad}\,\chi\text{，}\ \varphi' = \varphi - \frac{1}{c}\frac{\partial x}{\partial t}$$

具有

$$\Box\chi = 0$$

我们将在第十三章中看到电荷为 e 的一个电子的波动方程在上述规范变换和下述波函数变换连同一起的情况下是不变的。波函数变换式是

$$\Phi' = \Phi \exp(-\frac{ie}{\hbar c}x)$$

这种不变性导致电荷 e 守恒。

一种广义的规范场理论是杨-密耳斯理论（1954 年）。

在基本粒子物理学的新近发展中，对称概念已被认为具有相当基本的作用。我们将在第十三章和书末附录 13 中再作论述。

第七章　气体动理论、热力学和统计力学

除了动力学的和电磁学的（包括光学的）现象之外，还有一类涉及大量物质性质的现象。在讨论后一类现象时，需要若干附加的概念（即不同于空间、时间、质量、电荷、电磁场等的概念）。一个是温度概念，这个概念与热力学平衡有密切的关系——而在动力学问题中从未遇到过。还有另一些概念为处理宏观现象所需要，并且在处理有关大量物质性质的问题上也还有不同的观点。

1．气体动理论

最早的理论就是气体动理论。1658 年，伽桑狄把物质看成是由恒定运动的原子（弹性球）所组成的，这些原子能形成气、液、固体状态。1662 年，玻意耳发现在常温下的经验定律 PV=常数。1678 年，胡克试图解释这个定律。1738 年，丹尼耳 · 伯努利从一个假定出发推导出玻意耳定律，他的假定是：气体的压力产生于气体分子对容器壁的撞击。我们在这里使用“ 假定 ”一词，是因为在那时原子或分子的存在还是一种假设。这些概念得到了化学中定比定律和倍比定律之发现和阿伏伽德罗于 1811 年原子和分子间有区别之发现的有力支持。气体压力等同于分子的运动能，由焦耳于 1848 年，克朗尼格于 1856 年和克劳修斯于 1857 年加以确立；他们得到了一种理想气体的状态方程 PV=RT。平均自由程的概念是由克劳修斯引进的。一项重要的工作便是 1859 年麦克斯韦的速度分布定律。

气体动理论在 1868 年以后的二十年内由于玻耳兹曼的工作而充分成熟。他构建了玻耳兹曼方程，该方程是一个基本方程，它原则上描述了一种气体从任意态到平衡态的不可逆过程，通过他的 H 函数渐进地减小到极小值加以定义。玻耳兹曼方程原则上可描述粘滞运动、热传导和气体扩散等输运现象。应用于这些输运现象的玻耳兹曼方程的一种解是由查普曼和恩斯考格于 1911—1912 年彼此独立地提供的。玻耳兹曼理论在实际问题上的成功是巨大的；从玻耳兹曼方程进行演绎，就有可能用温度、分子质量和分子间相互作用定律表示传导系数和粘滞系数。

在谈论气体动理论在第二次世界大战后的进一步发展之前，让我们先回顾一下这个理论的基本假定。一个假定是从分子组成气体的微观观点出发的，分子服从经典动力学。为了代替处理大量分子（比方说，10^{22} 数量级分子），人们引进概率概念、分布函数和宏观图像中的平均值。麦克斯韦最初的速度分布理论是基于分布函数的一种概率论假定.玻耳兹曼的分布函数方程是基于（分子碰撞）动力学与一种概率论本质的某些启发性假定的组合（即所谓分子混沌拟设［Stosszahlansatz］和用单体分布函数描述 N 体系统的基本假定）。把 N 体问题归化为一体问题的假定可能性之

证明直到1946年波哥留波夫的工作之后才得以清楚地理解。

1946年，波哥留波夫在苏联，玻恩和格林在英格兰，柯克伍德在美国，以及（稍早）伊翁在法国，独立地开始他们的理论研究工作，他们都是从N分子的相空间内（N分子）的一个气体系综的分布函数D（q，p，t）之刘维方程出发，q，p是坐标和它们的共轭动量。通过对dq_1dp_1，$dq_1dq_2dp_1dp_2$，…，$dq_1\cdots dq_{N-1}dp_1\cdots dp_{N-1}$连续积分（取平均），人们得到N－1，N－2，…2，1粒子内的分布f_{N-1}，f_{N-2}，…f_2，f_1，它们是由来自刘维方程的一个（嵌入）方程的系列给出的。这个方程谱系（B－B－G－K－Y方程谱系）在内容上是与刘维方程等价的，它们可从刘维方程中推导出来，但这种形式却适于近似处理，从而有助于更好地理解玻耳兹曼理论。

理解玻耳兹曼理论合理性这一过程的关键来自波哥留波夫的洞见。波氏指出，从事物的本质上说，气体中存在着三类时间标度，即最短时间t_0，对应于碰撞时间，在此时间内两个分子间发生碰撞，即$t_0\approx r_0/u$，这里r_0是分子间相互作用的距离（10^{-8}㎝数量级），u是分子的平均速度；两次连续碰撞之间的时间t_1，$t_1\approx\lambda/u$，这里的λ是分子的平均自由程；而时间$t_2\approx L/u_s$，这里L是一个宏观长度（例如容器的大小的长度）而u_s是声波的速度。因此t_2是气体对平衡之弛豫时间的量度。在通常的密度和温度条件下，这些时间之间的关系是

$$t_0<<t_1<<t_2$$

这里不可能详谈波哥留波夫理论的细节，只略述其重要的结果。在一种气体内如上所述大为不同的时间尺度的固有存在，提供了用单体分布函数（玻耳兹曼方程）去处理N体气体（刘维方程）的合理性基础，也为用流体动力学方程（这些方程能从B－B－G－K－Y方程谱系中得到）解玻耳兹曼方程的查普曼—恩斯考格方法提供了合理性的基础。

随着这一发展，分子动理论有了非常清晰的基础。原则上，它能描述气体不可逆地趋向于平衡的过程；然而，数学解法却是另外一回事。玻耳兹曼方程和波哥留波夫理论的更多细节将在书末附录5中给出。

等离子体（高温）动理论随着刘维方程作为一个出发点而获得类似的发展。这里的问题因为带电粒子之间存在长程的相互作用，而不同于气体的问题，并且在这种情况下数学上也更加复杂。

2．经典热力学

大约在焦耳、克朗尼格和克劳修斯得到理想气体的状态方程pV=RT的同时，处理物质在热力学平衡态中性质问题的另一种方法也得到了发展。热力学的课题最初起于与包含热现象有关的经验，并以两个原理（或后来称为定律）的形式演变。热力学第一定律就是能量守恒定律，它似乎

使许多物理学家化费了大量时间才建立起来，这些物理学家有 R．克劳修斯（1822—1888）、J．P．焦耳（1818—1889）、 H．V．亥姆霍兹（1821—1894）、J．R．迈尔（1814—1878）以及开尔文勋爵（1824—1907）。热力学第二定律是由克劳修斯（1850 年）和开尔文勋爵（1851 年）宣布的。它创始于萨迪·卡诺（1796—1832）的一个定律。克劳修斯对定律的表述为 :不可能把热从一热源转移到较高温度的热源而不使环境发生某些变化。开尔文形式则为，不可能从单一热源取热，并使之完全转变为功。能够表明，这两种形式在以下的意义上是等价的，即对哪一个的否定都将与另一个相矛盾。也还有别的等价形式，诸如 M．普朗克（1858—1947）的说法，不可能通过任何一个装置的单独效应使一热源冷却，并使一重物的温度升高，而奥斯特瓦尔德的形式是，不可能有“第二类永动机”（第一定律否定了“第一类永动机”的可能性）。

所有各种等价形式都能用克劳修斯于 1865 年引进的熵的概念予以更精确的表达。对于热力学平衡的任何系统，熵 S 都是热力学变量（或它们的其他函数）的函数，这些热力学变量用以定义系统的热力学态（例如，对气体来说是 p、V、T 三个变量中的两个变量）。对于“可逆”变化，如果系统吸收热量△Q，则系统熵的变化就是

$$dS = \frac{\triangle Q}{T}$$

这定义了S，辅之以一个附加常数。（对于不可逆变化，$\frac{\triangle Q}{T}$根本不涉及到熵，且变化 dS 必定由一可逆过程决定，这个可逆过程与考虑中的不可逆过程具有相同的初态和终态。）

第二定律陈述的是，对于任何封闭系统（即系统孤立于整个环境，或者系统包含了它的整个环境作为一个单一系统），熵总是增加（或保持不变）。

$$dS \geq 0$$

对于平衡态，熵必须是它的极大值。

这两条定律从未与我们的经验发生过矛盾。热力学作为一个从这两条定律出发的演绎理论，爱因斯坦称之为“原理性理论”，与之相对照，爱因斯坦把气体动理论称之为“构建性理论”。从这两条定律出发，能为各种热力学函数和变量推导出大量关系（不同的方程），应用并受实验和观察的检验。

现在热力学的这两条定律，在它们不需要使用有关系统的任何详细而专门知识的意义上，是非常普遍的。事实上，当它们应用于化学热力学时，无需涉及物质的原子本性。因此，热力学是一个强有力的理论。

但是正由于两条定律的普遍性使得热力学方法“强有力”，同样的普遍性也给它带来了局限性。例如，从第二定律演绎出的相律表明，对处于热力学平衡的任何气体，一种函数关系

$$f(p, V, T)=0$$

存在，它是一种气体的状态方程。但热力学并不告诉我们关于函数 f（p，V，T）形式的任何东西。甚至对理想气体也如此。另一方面，最简单的气体动理论对理想气体给出方程 PV=RT，对于真实气体，则在一些假定基础上，给出了范德瓦耳斯方程，我们通过引进关于这些气体的某些假定而获得这些附加的知识。

热力学第二定律用克劳修斯、开尔文、普朗克和奥斯特瓦尔德否定某种情况之可能性的负面态度加以表达，也许其中有某种可期待的哲学意义。因此，出于哲学意义上的兴趣，卡拉西奥多里给出了另一种表述（1909年），这种第二定律表述是以“正面”陈述出现的：“在一个系统的任意态 P 的附近，至少存在一个态 Q，从 P 出发通过任何（可逆的或不可逆的）绝热过程永不能达到态 Q。”这个原理担保熵函数 S 具有如下性质，即

$$dS \geq 0$$

这个原理的第二部分的证明本质上（包含帕夫形式的性质）是数学的，而第一部分的证明取决于经验，因此这种形式的第二定律所依据的是经验，而不是通常表述中所有假定的不可能性。

热力学和气体动理论在某种意义上是互补的：热力学从宏观概念出发，从热力学平衡态系统和两条定律出发，无需任何关于系统结构的细节知识。因而它的应用范围广，但它所产生的信息有局限性。另一方面，气体动理论从服从经典动力学定律的相互作用的分子和从运用概率观念出发，目的在描述系统从任意初始宏观态趋向于热力学平衡态的过程。这当然是一个比热力学更为雄心勃勃的理论，因为热力学只涉及平衡态！

3．玻耳兹曼的统计理论

研究平衡态物质性质的第三种处理方法是玻耳兹曼的统计方法（19世纪 70 年代）。首先，对熵的热力学概念给出了一种诠释，它不只是易于掌握，而且还导致了统计力学的发展。

熵概念在热力学中最初引进时，据说是学生最难以掌握的概念之一。通过在玻耳兹曼统计考虑基础上的一种诠释，就容易理解多了。

试考虑一种 N 个分子（视为质点）的气体。令体积为Ω的六维空间分成 s 个（大数目的）相格，每个体积为ω，使 sω=Ω。令

$$(n) \equiv (n_1, n_2, \cdots, n_3)$$

$$\sum_{i=1}^{s} n_i = N$$

n_i 为 N 个分子在第 i 相格中的分布数，i＝1，…，s，$\Omega=\Omega_x \cdot \Omega_v$。这里 Ω_x 为坐标空间的三维体积，而 Ω_x 是速度空间的三维体积。〔注意：分子速度原则上有有限范围，所有分子的总能量 E 是一个常数

$$\sum_{i=s} n_i \varepsilon_i = E \quad \}$$

N个不可区分的分子按照（n）的分布方式数是$\frac{N!}{\Pi n_i!}$，而分布（n）的概率是

$$W_{(n)} = \frac{N!}{\Pi n_i!}\left(\frac{\omega}{\Omega}\right)^{n_1+n_2+\cdots n_3}$$

（注意：这也是气体6N维Γ空间的总体积Ω^N的分数。）

令 $$NK_{(n)} \equiv \sum_i \left(n_i + \frac{1}{2}\right)\ln\frac{sn_i}{N} \approx \sum_i n_i \ln\frac{sn_i}{N}$$

运用斯特林近似，$W_{(n)}$可以表示为

$$W_{(n)} = \sqrt{2\pi N}\left(\frac{S}{2\pi N}\right)^{s/2}\exp(-NK_{(n)})$$

为了找出$W_{(n)}$是极大值（即$K_{(n)}$是极小值）的分布（n）=（n_1，n_2，…，n_s），我们得到变分

$$N\delta K_{(n)} = \sum_i \left(1+\ln\frac{sn_i}{N}\right)\delta n_i$$

解$\delta K_{(n)}=0$服从

$$\Sigma\,\delta n_i = \Sigma\,\varepsilon_i\,\delta n_i = 0$$

在引入拉普拉斯倍增数λ和μ后，为

$$\ln\frac{sn_i}{N} + 1 + \lambda + \mu\varepsilon_i = 0$$

或者

$$n_i = Ae^{-\mu\varepsilon_i}\text{，}A\text{为常数}$$

我们把它称为麦克斯韦速度分布定律,如果不存在势场且ε是一个分子的动能。

$$\varepsilon = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

详细的研究表明，上面所确定的稳定值$K_{(n)}$事实上是一个极小值，而$W_{(n)}$（当$n_i = Ae^{-\mu\varepsilon_i}$时）事实上是一个强极大值。这就强有力地提示出，分布n_i（是一个“最概然”分布）对应于一个最经常出现的分布，因而是与平衡分布等同的。

从前述关系可知

$$\ln W_{(n)} = -NK_{(n)} + \text{常数}$$

让我们来找一找前面所述函数$K_{(n)}$，玻耳兹曼H函数（见书末附录5），以及热力学熵之间的关系。

在玻耳兹曼理论中，使用了单粒子的分布f（r，v，t）。fd^3rd^3v是N个分子坐标r和速度v处于体积元d^3rd^3v中分子的分数，即

$$\iint fd^3rd^3v = 1$$

在符号ω，$s\omega=\Omega$，$\frac{n_i}{N}$，$\frac{sn_i}{N}$分别等同于d^3rd^3v、$\iint d^3rd^3v$，fd^3rd^3v，$sf\omega=f\Omega$的情况下，有

$$K_{(n)}=\frac{1}{N}\sum_i n_i \ln\frac{sn_i}{N}\leftrightarrow\iint f\ln(f\Omega)d^3rd^3v$$

因此

$$K_{(n)}=H+数$$

玻耳兹曼H定理$\frac{dH}{dt}\leq 0$连同$\frac{dH}{dt}=0$态的性质提示出H是与负熵成正比的，即

$$-H=S+常数$$

最后，我们可以用一个图解形式概括 H、W 和 S 之间的关系如下

这里下标“0”表示适合于一个分子的量。附加常数没有确定①。

4．吉布斯理论：Γ空间

玻耳兹曼理论把在六维μ空间 s 相格中分子的最概然分布（n_1，n_2，…，n_3）与平衡态分布等同起来，这可以用吉布斯的Γ空间理论表达得更清楚。

一种 N 个分子的气体的Γ空间是 N 个分子的坐标和动量的 6N 维空间。体积是Ω_N，Ω是一个分子的六维相空间（见第 3 节）。分子的运动方程由下式给出

$$p_k=-\frac{\partial H}{\partial q_k}，\dot{q}_k=\frac{\partial H}{\partial p_k}，k=1，\cdots，3N$$

气体由在此Γ空间中的一点 P 描述，而分子运动则由在Γ中（6N—1）维能量表面的 P 的轨迹描述。极大值$W_{(n)}$被看成对应于占Γ空间中可达到体积的绝大部分——可达到是借助于气体分子给定的总能量表达的。

按庞加莱的各态历经定理（Poincaré’sErgodic Theorem，1890 年），点 P 在一定时间进程（足够长的时间）中，能足够接近地通过Γ空间中的任意点。因此在一个长时间内，可发现 P 在大多数时间内处于对应于六维μ空间的极大值 W■的Γ中可达到体积的极大部分。这就是玻耳兹曼把最概然分布等同于平衡分布观点的另一种表达。

〔在上面讨论中，我们并未讨论各态历经性问题的严密性，或Γ空间的“度量传递性”。这种简化的论证只是提示出“似然性”。〕

5．统计力学：系综理论

（1）玻耳兹曼和麦克斯韦的早期理论

玻耳兹曼关于在μ空间（一个分子的六维相空间）中最概然分布等同

于热力学平衡分布的理论，纯粹是统计的；它不包含动力学，各分子被认为是独立的。

真正的"统计力学"开始于麦克斯韦和玻耳兹曼。在前一节中，N 个分子组成的一种气体的态是由在 Γ 空间的一个点 P 描述的。玻耳兹曼于 1868 年引进了各态历经假设（麦克斯韦称之为路径的连续性，1878 年），即：在时间进程中，P 通过 Γ 空间按照气体的总能量能为 P 所达到部分的每个点。根据人们所谓的热力学第零定律（ ZerothLaw），任何系统终将趋近于热力学平衡态。因此，一个系统的平衡性质 f（q，p）由（长）时间的平均给出

$$\bar{f}=\lim_{T\to\infty}\frac{1}{T}\int_{t}^{t+T}f(p_k，q_k)dt$$

P_k，q_k 由运动的正则方程给出。这样问题就成了计算 $\bar{f}$。

玻耳兹曼（1869 年）引进了系综概念，系综由（无限）多个类似系统（气体）组成，所有系统全都处于具有同样热池的热力学平衡中。因此一个系综由 Γ 空间中的一群点描述，一般情况下具有密度 ρ（p，q，t）。对于热力学平衡，密度 ρ 必须是固定的，即 $\frac{\partial\rho}{\partial t}=0$，ρ $=\rho_0$（P，q）。在经典动力学中，密度满足刘维方程

$$\frac{\partial\rho}{\partial t}+(\rho，H)=0$$

这里（ρ，H）是 ρ 和哈密顿函数的泊松括号。因此对稳定的密度 ρ_0（ρ_0，H）=0

另外，在经典动力学中，正则方程一次积分的任一函数 F（q，p）满足（F，H）＝0。按各态历经假说，点 P 留在 Γ 空间的（6N—1）维能量表面（或一个能量壳层上）。因此，对稳定的密度，ρ_0 只能是能量积分的任意函数，即

$$\rho_0=\rho_0〔E（q，p）〕$$

具有形式 f（q_k，p_k）的气体的任何物理性质，当对稳定系综取平均时，都是

$$<f>=\int\cdots\int f(q_k，p_k)\rho_0(q_k，p_k)dq_1\cdots dp_N$$

麦克斯韦和玻耳兹曼理论证明了

$$\bar{f}=<f>$$

即 f 的热力学平衡值能通过求系综平均值而获得，为此目的，可使用能量积分的任意函数。

不幸的是，数学研究已经表明，各态历经假说并不正确（罗森塔耳，普朗切利耳，1913 年）。

（2）吉布斯的系综理论

该理论抛弃了各态历经假设，而以下式开始

$$\rho_0(q，p)=\rho_0(E)$$

以此作为一种稳定密度的基本假设，$\rho_0(E)$ 是能量积分的任意函数。依据人们对 $\rho_0(E)$ 所作的假定形式，有各种不同的系综理论。一种是"微正则系综"，定义如下：系综中的所有系统都有相同的分子数 N，相同的总能量，处在 E 和 E＋△E 之间，△E 很小。令能量壳层中可达到的体积为Ω。令 ρ（q，p）符合如下条件：

$$\int_{\Omega}\rho d\Omega=1 \qquad \text{（归一化）}$$

且

$$\sigma\equiv\langle\ln\rho\rangle=\rho\ln\rho d\boldsymbol{\Omega} \qquad \text{是极小值}$$

（这最后的条件是通过与玻耳兹曼理论中 $K_{(N)}$ 的类比提示的，见上面第 3 节。）

变分问题

δ■ρdΩ=0，δ■ν ρ lnρdΩ=0

导出

$$\rho=\text{常数}$$

σ是极小值，是由二次变分表示的

$$\delta^2\int_{\Omega}\rho\ln\rho d\Omega>0$$

微正则系综则可由下式描述

$$\rho_0=\begin{cases}\text{常数，对}E_0\leq E\leq E_0+\triangle E\\0\text{，在能量壳层之外}\end{cases}$$

该理论的公设是，任何物理量的平衡值 f（P，q）皆是系综平均

$$<f>=\int_{\Omega}f(p,q)\rho_0 d\Omega$$

另一种系综是正则系综。正则系综是指在系综中只有一个系统的总能量 E 的系综平均是已知的（被指定的）。

$$<E>=\int_{\Omega}E\rho d\Omega$$

同样的要求

$$\langle\ln\rho\rangle\text{为极小值}$$

导出

$$\rho\Rightarrow\begin{cases}Ae^{-\beta E}\text{，}\beta\text{是常数}\\0\qquad\text{在能量壳层之外}\end{cases}$$

还有另一种系综被称为巨正则系综，我们就不在此表述了。

从一个系综出发，有可能推导出若干表达式，与某些热力学关系式比较，这些表达式可以与温度、自由能、熵等热力学函数等同。

</PARA/>例如，通过对能量E的表达，$E=\frac{1}{2}m(v_x^2+v_y^2+v_z^2)$，</PGN0081/PGN>"力学"就进入到该理论中。统计力学发端于分子的微观概念，分子间的相互作用也包含在内。它不依赖于各态历经假设。它是一种演绎的、数学的方案，可以从这个方案中获得结果。它等同于热力学函

数和关系式。在相同的条件下，玻耳兹曼理论、达尔文—富勒理论和吉布斯理论都导致相同的结果。

6．能量均分定理

玻耳兹曼和吉布斯的经典统计力学的一个重要结果是能量均分定理。这个定理是说，在热力学平衡时，每个自由度都有一个平均能量$\frac{1}{2}kT$。早在1845年，当热力学平衡时一种气体混合物中各类分子之间能量均分的思想，就由 J．J．沃特斯顿在给伦敦皇家学会的一篇论文中提出来了，但这篇论文没有发表。1859 年麦克斯韦重新发现了这一思想。金斯的论证，对麦克斯韦的论证作了改进（1879 年），基本上等价于微正则系综的方法。该定理从单原子气体和固体的比热（杜隆—珀替定律）中获得了大量的实验证据。但它在双原子气体情况下的失效和在非常低的温度下对杜隆—珀替定律的偏离是如此严重，以致开尔文勋爵在 1900 年说这种失效是密布在热的动力学理论之优美和清晰性上的“两朵乌云”之一。

量子理论的起源将是下一章的主题。经典物理学（动力学、电磁理论、热力学和统计力学）在说明黑体辐射观察谱线分布上的失败，就是由包括能量均分定理在内的瑞利—金斯定律的失败而积聚起来的。

7．量子统计

1924 年，S．N．玻色指出，把光子作为辐射处理并修正玻耳兹曼发现的它们在相格中分布的方式，有可能得出普朗克的辐射公式，而不是每一自由度为经典的$\frac{1}{2}kT$。玻色把他的（英文）手稿送给爱因斯坦，爱因斯坦发现这非常重要，于是将它译成德文，并送交《物理年鉴》（Zeitschrift für Physik）发表。爱因斯坦把玻色的思想推广到分子，并由此而创造了与玻耳兹曼统计相对的玻色—爱因斯坦统计。

1926 年，E．费米和 P．A．M．狄拉克提出了另一种统计法，这种方法适合于像电子、质子和中子这样的粒子。

本书的目的不在于提供任何理论的细节。但以下由于布里渊的贡献而形成的对各种分布的优雅（统一）处理是有趣的。

令类 j 的 N_j 个粒子分布在 G_j 格（或能级 j）内。再让我们假定，第一个粒子能以 $G_j-\alpha$ 种方式分布，第二个粒子以 $G_j-2\alpha$ 种方式分布，等等，第 N_j 个粒子是以 $G_j-(N_j-1)\alpha$ 种方式分布。N_j 个粒子就能以

$$G_j(G_j-\alpha)(G_j-2\alpha)\cdots(G_j-[N-1]\alpha)$$

种方式分布。由于 N_j 个粒子均是不可区分的，所以不同方式数是由

上面的数除以 N_j！。在 G_1 格内分布 N_1，在 G_2 格内分布 N_2 等等，分布方式总数便是乘积。

$$W=\prod_j \frac{G_j(G_j-\alpha)\cdots(G_j-[N_j-1]\alpha)}{N_j!}$$

分布 N_1，N_2，…N_j，对极大值 W 服从条件① $\sum N_j=N=$ 常数，总数目 $\sum N_j\varepsilon_j=E=$ 常数，总能量

由下式给定，

$$N_j=\begin{cases}\dfrac{G_j}{e^{-\gamma+\beta\varepsilon_j}-1}\text{，玻色—爱因斯坦}(\alpha=-1)\\ G_j\cdot e^{-\beta\varepsilon_j}\text{，玻耳兹曼}(\alpha=0)\\ \dfrac{G_j}{e^{-\gamma+\beta\varepsilon_j}+1}\text{，费米—狄拉克}(\alpha=1)\end{cases}$$

玻色—爱因斯坦统计，α=－1，意味着粒子有一种聚集在同一格中的趋势；玻耳兹曼统计，α =0，意味着格中有一粒子存在并不影响第二个粒子进入此格中的概率；费米—狄拉克统计，α =1，意味着粒子有一种不进入同一相格的趋势。这最后的情况类似于泡利不相容原理。

玻色—爱因斯坦和费米—狄拉克统计在 $G_j>>N_j$ 的极限情况下接近于玻耳兹曼统计。它们与玻耳兹曼统计的差别在高温与低密度下是很小的。在低温下，^4He 显示出很

强的量子效应（爱因斯坦凝聚），而且由于它们的质量小，因而原子或金属中的电子显示出很强的（简并）效应。我们不准备在这里进一步谈这些方面的问题。

再回到能量均分定理，这是我们第 6 节的出发点，经典理

论中每个自由度的平均能量 $\frac{1}{2}kT$，按玻色—爱因斯坦统计，被下式替代

$$kT\rightarrow\frac{x}{e^x-1}kT\text{，}x=\frac{hv}{kT}$$

可以看到，对于 x<<1，即高温 T 或 h→0，量子理论表达式接近于经典值 kT。

8．评论

（1）经典热力学使用宏观变量和“状态函数”来描述物质在热力学平衡时的热力学性质。经典统计力学（麦克斯韦、玻耳兹曼、达尔文一富勒、吉布斯）发端于微观概念（分子及分子间的相互作用），但借助系综和配分函数来定义宏观函数以描述热力学平衡时物质的性质。量子统计也是系统在热力学平衡时的理论。在所有这些理论中，都不包含“随时间变化”的概念。

还有非平衡热力学的较近期的发展（昂萨格，1931 年；普里戈金，20 世纪 40 年代；德·格罗特，20 世纪 40 年代）它涉及不可逆过程。一个简略的讨论将在书末附录 6 中给出。

（2）动理学理论旨在处理描述系统从一任意态不可逆地接近热力学平衡态的问题。近期理论（波哥留波夫，1946 年，及其他人）从刘维方程出发，并且原则上导出了玻耳兹曼方程为其一级近似的逐步求近的程序。这个理论的基础似乎是清楚的，但必须在不同阶段引进很多物理近似和数学近似，且方程的求解很困难。

书末附录 5 将给出玻耳兹曼和波哥留波夫理论的简略说明。

（3）总是存在着这样的问题，即怎样以及为什么有可能表述一个理论，它可以在动力学定律的基础上描述不可逆地趋近于热力学平衡的过程，而这些动力学定律本身又是时间上可逆的。热传导、扩散、布朗运动、欧姆热等等方程都“构成”不可逆的，也不像牛顿的运动方程或麦克斯韦方程那样具有基础的意义。玻耳兹曼方程确实有一个确定的时间箭头，这已由 H 定理表明了，但我们知道这是由于在“碰撞积分”中所谓分子混沌拟设所引入的假定，不是来自动力学而是具有概率的本性，并且我们也已知道时间方向是怎样通过概率考虑而得到的。如果人们从刘维方程出发，而刘维方程来自经典动力学并在时间上可逆，那么现在的问题就成了：人们怎能以一个描述不可逆地趋近平衡的理论告终？

在波哥留波夫理论中,基本的刘维方程的时间可逆性由于引入了某些“初始条件”（柯亨和伯林，1960 年）而遭破坏，引入这些初始条件定义了一个时间方向，仅在此方向上理论才有效。一个时间可逆的方程也还要通过诸如拉普拉斯变换（它自动地去掉了一个时间方向不作考虑）等其他手段而给出一个时间方向。

近期还有不少尝试以构建一个时间不可逆的基本理论(在牛顿动力学和量子力学层次上）。这里，包含着一个哲学态度问题：即我们是否满足于目前的观点，把时间可逆作为基本理论接受，并在统计基础上理解宏观不可逆性，抑或人们应该在基本理论本身中寻求宏观的时间不可逆性？

（4）“熵”的概念被推广到作为一个整体的宇宙，并且有人把热力学第二定律$\triangle S>0$ 当做定义时间的“方向”。按此观点，在宇宙尺度上“温度”的普遍“齐一性”（uniformization）是与“熵增”连在一起的，而诸如地球上的生命现象中的“熵减”则只是片刻的、局部的涨落。然而，按照宇宙演化的大爆炸理论，根据目前的知识，人们还不能肯定宇宙是否将不确定地一直膨胀下去，或者将停止膨胀并开始收缩，到条件成熟又开始另一次大爆炸。在后一种情况下，宇宙是“周期性的”，在宇宙学意义上定义时间方向变得不很清楚，尽管在一个缓慢收缩的宇宙中，包含“熵减”的生命过程和其他过程也许依然作为涨落而存在。

（5）关于生命现象，有人提出这样一种观点，即人们应看到太阳送

到地球上的不是能量而是负熵。人们还不清楚这种观点究竟有何助益。因为太阳送到月亮或火星上的负熵似乎并未减低任何熵。如果人们非要坚持使用“熵”的概念的话，那么，简单的观点就是，地球上的生命过程是特别复杂的低效率的机制，它以吸收能量为代价使“熵”减低。

参考文献

J.H.Jeans,The Dynamical Theory of Gases,Cambridge University Press.4th ed., Dovered., New York,1954（玻耳兹曼方程，玻耳兹曼统计，能量均分定理）

A.Pais，见第四章后参考文献

（玻色—爱因斯坦统计的历史陈述）

吴大猷，热力学、分子运动论和统计力学（理论物理第五册），科学出版社，北京，1983 年

（麦克斯韦—玻耳兹曼统计力学，达尔文—富勒理论，吉布斯系综理论，量子统计力学）R. B. Lindsay, Physical Statistics. John Wiley＆Sons, New York, 1941

（布里渊的量子统计方法）

第八章　量子论

1.量子论的起源

到 19 世纪末期，经典力学、包括光学在内的电磁学、热力学及其统计理论皆已完成。物理学的状况促使一位重要的物理学家宣称，物理学的进一步发展，将只不过是改进其精确度至下一位小数点而已。但事实上，此时物理学正处于一系列惊人进展接踵而至的新纪元的前夜。

通过对气体中导电性的实验研究，伦琴于 1895 年发现了 X 射线；J.J.汤姆孙于 1895 年至 1896 年间“发现”了电子；A.H.贝克勒尔于 1896 年发现了天然放射性。这些发现预示了“现代物理学”的曙光。

甚至在这些发现之前，实验物理学家就研究了作为温度函数的“黑体”辐射的光谱分布。理论物理学家则寻求给观察到的分布定律以一个理论公式。经过许多努力，终于得出瑞利—金斯定律（1900，1905 年），给出了物体在平衡温度 T 时辐射场中频率范围 dν 中的能量密度 $d\nu\psi_\nu d\nu$，这公式为

$$\psi_\nu d\nu = \frac{8\pi\nu^2 d\nu}{c^3}\times kT$$

其中第一个因子给出单位体积中（横向电磁振动）自由度的数目，第二个因子 kT 即每一（横向运动）自由度的平均能量，这是经典统计力学中均分定理的必然结果。这一公式在长波长一端与观察到的光谱相符，但由于其对所有频率的积分值发散而显然没有意义。

1900 年 10 月，马克斯·普朗克基于熵的观点，并通过在上述瑞利—金斯定律和由维恩得到的仅适用于光谱的短波端的定律之间尝试经验插入法，“猜出”了一个公式

$$\psi_\nu d\nu = \frac{8\pi\nu^2 d\nu}{c^3}\frac{a\nu}{\exp(b\nu/T)}-1$$

其中 a、b 是常数。这一公式与经验光谱相比，（在 ν 的大小两端）都有很好的渐近表现。经过一段紧张的工作，普朗克于 1900 年 12 月得出公式

$$\psi_\nu d\nu = \frac{8\pi\nu^2 d\nu}{c^3}\frac{x}{e^x-1}，\ x=\frac{h\nu}{kT}$$

当时他通过与“猜得的”公式比较，假设“振子”（辐射与它们处于热平衡）以比例于频率 ν 的单元 ε 吸收和发射能量。尽管这一假设按经典物理学（电磁学、热力学和统计力学）基本上是无法理解的，但普朗克公式却与对 ψ_ν 的最精确的测量完全相符。这一情形在物理学史上是异乎寻常的：物理学的实验事实，迫使人们对已牢固确立的概念和知识作出革命性的否定！普朗克本人，也只是这一新理论的勉强的创立者，正是这一新理论，在二十年后导致了全新的物理学，即量子力学。

2．爱因斯坦理论

普朗克的量子论是革命性的，爱因斯坦进一步于 1905 年发展了它。同样基于（空腔中辐射的）熵的观点，并与气体统计理论中的概率相比较，他提出不仅在辐射的吸收和发射过程中,而且在空间的传播中也存在分立的量子 hν（h 是普朗克在他的理论中引入的常数）。按这种形式，爱因斯坦的理论甚至比普朗克的更为革命,因为它把电磁波在空间中的传播图景描述为一连串的能量量子 hν 流。爱因斯坦提议他的理论可由光电效应加以检验，而密立根等人于 1916 年在实验中证实了这一量子效应。G.N.刘易斯于 1926 年将这一量子命名为“光子”。爱因斯坦于1917年完善了他的光子理论，给光子以动量$p=\frac{h\nu}{c}=\frac{h}{\lambda}$对频率为 ν 和波长为 λ 的电磁波,光子具有由下式规定的粒子性质——能量 E 和动量 p

$$E=h\nu，p=\frac{h}{\lambda}$$

光子理论还从康普顿效应（1923 年），特别是波特和盖革于 1924—1925 年间对反冲电子和散射光子的同时测量实验中，得到强有力的支持（参见书末附录 7）。

量子论早期的另一成功是爱因斯坦的固体比热理论（1907 年），它显示出在很低温度时对杜隆—珀替定律 $C_v=3R$ 的偏离。应用普朗克理论于固体中原子的振动，得出原子比热的表达式

$$C_v=3R\frac{x^2e^x}{(e^x-1)^2}，x=\frac{h\nu}{kT}$$

它当 x→0 即高温时趋于 3R。尽管这个理论后来为玻恩、V.卡尔曼和德拜等人所修改，但爱因斯坦的主要思想仍是现代固体理论的发端。

3.玻尔理论

电子的发现表明了原子不是不可分的。由此卢瑟福及其助手于 1911 年进行了 a 粒子被（薄金属箔中的）原子散射的实验。实验的想法是非常简洁的，但其结果却出乎意料的重要，它们导致了原子的“有核模型”。太阳系模型似乎应是最自然的,但由于按电磁学定律它是不稳定的而必须被抛弃。

任何由电子和带正电的核组成的原子模型,基于经典电磁理论似乎都难以克服其不稳定性困难，因为按照经典电动力学，一个被加速的电荷必然发射出辐射。尼尔斯·玻尔在 1913 年访问曼彻斯特的卢瑟福实验室时，通过提出两个假设而一举解决了这一难题：即由“量子化条件”

$$\oint p_i dq_i = n_i h，n_i = 1，2，3，\cdots i = r，\theta，\varphi$$

（以索末菲 1915 年得出的普遍化形式）确定的定态的新概念，和用于计算原子在两个能量分别为 E_M 和 E_N 的定态间发生跃迁时发射辐射的频率公式

$$h\nu = E_M - E_N$$

玻尔的氢原子光谱理论获得了圆满成功，包括它的应用于氦离子光谱，索末菲的由于电子质量的相对论性变化得出的精细结构理论（1916 年），正常塞曼效应（1916 年），空间量子化及其由施特恩—革拉赫实验（1921—1922 年）所作的证实。对（非氢原子中）定态的一个直接证据，来自关于在电子和汞原子间非弹性碰撞的弗兰克—赫兹实验（1914 年），及随后由他人所作的关于其他原子的实验。

玻尔的氢原子光谱理论的巨大成功,开辟了对原子和分子光谱进行广泛而深入细致研究的时期。

在对原子的复杂光谱的分析中，人们发现有必要引入多个新的“量子数”，例如多重性，L、J 与态的奇数性和偶数性等。起初（在 1925 年引入电子自旋之前），这些量子数的意义并不清楚，因此尽管朗德发现反常塞曼效应的 g 因子是一项惊人的功绩,但也仅仅是增添了累积起来的神秘性而已。乌伦贝克和古兹密特于 1925 年提出的电子自旋理论部分地“解决”了这些问题，同时也为稍早些由泡利（1925 年）提出的泡利不相容原理提供了所缺的一个要素，事实上泡利使用字母“S”来标示这个尚未知的“量子数”。泡利原理为玻尔对周期表中的化学元素的组建原理提供了解释。自旋和泡利原理这两个概念注定了要在基本粒子物理学中发挥重要作用。

4.跃迁概率

让我们回到普朗克辐射公式，它于 1917 年又得到一个重要的、基本的发展。爱因斯坦考虑处于热平衡中的辐射,引入了自发发射、吸收和“负吸收”的跃迁概率概念。“负吸收”意即由（适当频率的）电磁场诱发的发射。通过求助于瑞利—金斯定律作为长波或高温情况下的极限形式，爱因斯坦得出了普朗克黑体辐射公式。诱发发射的概念导致了汤斯（1954 年）微波的发展和 1960 年激光的发展。

5.玻色统计

与普朗克公式相关的另一项基本进展是 1924 年取得的。S.N.玻色把辐射处理为一种光子气体，它服从一种与 μ 空间（分子的六维相空间）相格中的玻耳兹曼分布不同的分布定律，从而导出了普朗克公式。这鼓舞了

爱因斯坦把这种新的分布扩展到气体中，玻色—爱因斯坦统计由此诞生（第七章第 7 节）。

6.量子论的困难

尽管量子论取得了如此巨大的成功，但日益明显的是，它在概念上和技术上，都面临许多不容忽视的困难。为清晰计，让我们以一系列评论方式讨论其中的几项。

（1）首先，辐射以量子形式的不连续发射和吸收，是与电磁波理论中的连续性这一基本概念相矛盾的。

其次，为光电效应和康普顿效应所充分支持的爱因斯坦关系式

$$E = h\nu,\ p = \frac{h}{\lambda}$$

使粒子的属性与波的属性相等同，也为经典物理学所无法理解。

再者，在玻尔的氢原子理论中应用了经典动力学概念，但却为保证稳定性而“撤销”了电磁辐射定律。在爱因斯坦关系和玻尔定态条件中都有一种不一致性：假设了概念之间的关系，却违背或否定了概念本身的含意。

（2）有一些很有趣的尝试，通过借助于庞加莱的积分不变式和（埃伦菲斯特、伯格斯于 1916 年提出的）绝热不变式来“理解”为什么定态会发生量子跃迁和不连续变化。但由于定态条件包含普朗克常数 h，任何由经典物理学作出的解释都是注定要失败的。

（3）为了尝试沟通量子论和经典理论，玻尔于 1918 年引入下述概念，即在大量子数的极限下两个理论间存在某些渐近联系。这一“对应原理”事实上在新量子力学的发展中也成为一种指导思想。

（4）玻尔氢原子理论的惊人成功很自然地使人们尝试把它推广到下一个最简单的原子——氦。但人们很快就意识到存在着基本的困难，这主要还不在于计算出的和经验的原子能量间的不一致,而在于无法制定核场中双电子问题的量子化条件。

（5）另外还有“较小的”技术性问题，即在双原子分子振动光谱和零点能中出现的半奇整数的量子数。半奇整数量子数本身并不比整数的更困难，难处出在对于给定情形应采用哪一种量子数并无指导规则可循。

参考文献

M. Jammer, The Conceptual Development of Quantum Mechanics, McGraw Hill, New York, 1966

（量子论和量子力学的历史叙述，有关原始论文的索引）吴大猷：量子论与原子结构（理论物理第二册），科学出版社，北京，1983 年

Quantum Mechanics,World Scientific Publ, Co., Singapore,1986

（普朗克、爱因斯坦、玻尔、索末菲、泡利、乌伦贝克和古兹密特等人的理论）

第九章　量子力学

1924—1926 年间，从不同的观点出发，独立地萌发和成长起一些新的思想，却又以互补的方式，汇聚成一个从很基本的概念上就不同于经典体系的统一的理论结构。路易斯·德布罗意（1892—1987）于 1923—1924 年提出了新的“物质波”概念，这导致 E.薛定谔（1887—1961）于 1926 年的最初 6 个月内系统地提出和发展了他的波动力学。W.海森伯（1901—1976）于 1925 年以一条完全不同的途径开始发展起他的矩阵力学，他的非对易算符的概念激发 P.A.M.狄拉克（1902—1984）很快发展起一种普遍的量子力学。矩阵力学与波动力学在数学上的等价由薛定谔本人于 1926 年加以证明。同年，M.玻恩通过他的概率诠释，补充了波动力学所失落的一个物理要素。1927 年，海森伯提出了不确定性原理，玻尔制定出他的互补性概念。量子力学的整体结构至此宣告完成。在处理原子和亚原子领域的现象时，量子力学取代了经典物理学。

在本章和随后几章中,我们将试图简要回顾这些进展以及由此而来的关于物理理论本质的哲学问题。

*　　*　　*　　*　　*　　*

德布罗意是从 W.R.哈密顿（1805—1865）于 1828—1829 年间注意到的光学上的费马最小时间原理（1657 年）

$$\delta\int_A^B \frac{ds}{u} = 0,\quad (u = c/n，n\text{为绝对折射率})$$

和质点动力学中的莫培督最小作用原理

$$\triangle\int_{t_1}^{t_2} 2T\,dt = 0 \qquad (T = \frac{1}{2}mv^2)$$

之间存在密切的形式相似性开始其研究的。通过引入 $\lambda = \frac{u}{\nu}$ 和 $2T\,dt = p\,ds$，这两条定律的形式变为

$$\delta\int_A^B \frac{ds}{\lambda},\quad \delta\int_A^B p\,ds = 0$$

如果

$$\frac{1}{\lambda} \propto p$$

其形式将是同一的，即粒子在势场 V（x，y，z），$p^2=2m(E-V)$ 中的轨迹与光在折射率为 n（x，y，z），$u = \frac{c}{n}$ 的介质中的传播，将由相同的数学方程给出。这一见解被埋没了近一个世纪，直到德布罗意提出粒子和波之间的关系可能不仅仅是形式上的,粒子可能具有某些与之相关联的波动性质。他的思路可能如下：

在一个惯性系中，能量—动量是一个四维矢量（p，$\frac{iE}{c}$）。在同一系统中，平面波由 $\exp(ik\cdot r-wt)$ 描述，（k，$\frac{iw}{c}$）和（r,ict）是两个四维矢量。德布罗意订出了比例关系

$$k\propto p，\omega\propto E$$

通过选择普朗克常数h为比例系数，并由于 $|k|=\frac{2\pi}{\lambda}$，他得出

$$\lambda=\frac{h}{p}，\nu=\frac{E}{h}$$

认为对一个能量为 E 和动量为 p 的粒子，即有一个频率 ν 和波长 λ 由上式确定相关的波，他称此波为“物质波”，尽管其本质尚不清楚。

德布罗意关系在形式上与爱因斯坦关于光子的关系式是完全相同的，但它们的意义却大不一样！

德布罗意的波长 $\lambda=\frac{h}{p}$ 的物质波概念立刻取得的第一个成功，是应用于氢原子中的电子时，从圆轨道中的驻波条件非常简单地给出玻尔的量子化条件 $2\pi pa=nh$。

关系 $\lambda=\frac{h}{p}$ 很快为戴维孙和革末的实验（1927年）、G．P．汤姆孙（1927 年）、拉普（1928 年）和其他人关于电子“波”为晶体和光栅衍射的实验所证实。

这样，德布罗意关系和爱因斯坦关于光子的关系式一起，构成了被称为“波粒二象性”的二难推论。后面我们将会回到对这些经典物理学所不能理解的关系的哲学讨论上来。

1．波动力学

这里我们将打破量子力学发展的年代顺序，先来谈薛定谔关于波动力学的工作。德布罗意留下了诸如他的“物质波”的本性和“支配这些波的规律”这类问题。1926 年，薛定谔从哈密顿提出的光学上的费马原理和粒子的莫培督原理之间的密切相似性出发，寻求德布罗意波的规律或方程。基于下述思考，他得出了几何光学和波动光学间的关系。对几何光学，给定光线方向，波前 S 的梯度∇S 为常数时，我们有程函方程

$$(\nabla S)^2-n^2(r)=0$$

利用变分方程，对函数

$$\psi(r)=\exp(kS)$$

人们可由 S（r）得出波动方程

$$(\nabla^2+k^2)\psi=0$$

薛定谔注意到经典动力学可以纳入哈密顿—雅可比方程形式：

$$\frac{1}{2\pi}(\nabla S)^2-(E-V)=0$$

该方程相似于、并可用以“对应”于光学的程函方程。借助于光学情形相同的变分方程，薛定谔对函数

$$\psi(r)=\exp(S/h)$$

得出波动方程

$$\left\{\frac{\hbar^2}{2m}\nabla^2+(E-V)\right\}\psi=0$$

这就是薛定谔方程。

至关重要的是,应记住上述程序并非从某种已知原理中推导出薛定谔方程。几何光学和质点力学间的相似以及它们在波动光学和新的波动方程之间的对应关系，并非逻辑的推断。上述说明的意义仅在于表述薛定谔的思路，薛定谔方程实际上是一种新理论的基本假设，甚至在量子力学已发展成为一个完整的结构之后也依然如此。

在这一理论中,玻尔的量子化条件现在已为薛定谔方程的本征值条件——一种纯数学问题所简单取代（参见书末附录 9）。

薛定谔方程在实际问题中的成功我们已知之甚详，这里不再赘述。

在薛定谔得出德布罗意波的（上述）规律后，还有另一个未解决的问题，即波的意义。1926 年，M．玻恩（1882—1970）提出，积分

$$\iiint\psi^*Q\psi dxdz=\langle Q\rangle$$

应解释为当系统处于由ψ描述的态时，测量一个物理量 Q 时的（期望）值。〔当 Q 为一坐标值（例如 x）时我们会更为熟悉。此时$\psi^*\psi$dxdydz 为在（x,y,z）处的体积元 dxdydz 中发现坐标 Q 的概率。〕ψ的这种诠释导致了（从经典物理学的观点看）非常奇怪的结论。它是量子力学的基本假设之一。我们将在下一章中更详尽地讨论它。

按历史的顺序，1927 年，海森伯从爱因斯坦—德布罗意关系

$$E=h\nu，p=\frac{h}{\lambda}$$

出发，推断出了坐标与其共轭动量的同时值之间的不确定性关系

$$\triangle x\triangle p_x\geq h$$

一个通过（γ射线）显微镜来测量电子的位置 x 及其相伴的康普顿散射中的反冲的例证已众所周知。这个关系是普遍的，只要假设了爱因斯坦—德布罗意关系式，它就是必然的。它已被通称为不确定性原理，是量子力学的基本原理之一。

从这一关系，可立即推断出它与经典动力学的下述基本假设相冲突，即原则上可以以任意的精确度同时知道 x 和 p_x。由此，在不确定性关系的基础上，经典动力学失去了它的基础作用；它仅在我们的宏观经验中，即当所涉及的动量如此巨大,以致不确定性$\triangle x=h/\triangle p_x$比之于相关问题微

不足道时才有效。

这里我们要注意，接受了$\psi^*\psi$的概率诠释之后，薛定谔理论不再是牛顿运动方程意义上的“因果”理论。薛定谔方程作为一个微分方程是“因果”的，但这是一种只适用于概率的因果性。

在概要介绍德布罗意和薛定谔理论的起源时,我们试图追寻他们思想的发展轨迹，并强调他们的创造力和独创性。我们也致力于说明，创造力和独创性对重大的突破是必要的因素，同时，掌握一定的基本知识也是很有帮助的。我们应该赞赏德布罗意对哈密顿的工作和相对论理论的知识，赞赏薛定谔在几何光学和波动光学、哈密顿力学及微分方程的本征值问题等方面的造诣。普朗克和爱因斯坦引入了量子论中的革命性概念，但他们都精通经典热力学和统计力学，都基于熵的观点开始自己的新理论。在电磁学的例子中，法拉第的直觉洞察力奠定了电磁场理论的基础，但把它发展成为一个完备的电磁学体系，则有赖于麦克斯韦的独创性和数学知识。我们还可以举出其他例证。爱因斯坦的广义相对论思想在深入研究张量微积分后得以充分发展 狄拉克的灵感来自他早年接触过经典力学中的泊松括号 海森伯的奇异想法借助于玻恩对于矩阵代数——一种当时还不为物理系学生所熟悉的数学——的知识而形成一种理论。这就是人们曾说过的，重要发现的机遇只青睐那些有充分准备的人们。

2.矩阵力学

现在让我们回到恰好早于薛定谔 1926 年初工作的 1925 年。1925 年夏，W.海森伯作为格丁根的玻恩的助手，访问了玻尔的理论物理研究所，与克喇末一起研究量子色散理论。他估量了量子理论的困难并寻求一种新的途径。他争辩说玻尔理论应用了不可直接观察的概念，例如电子轨道的半径和频率等,并表示要构造一个其中只有直接可观察量出现的理论——一种数学模式。这点他是受了爱因斯坦所坚持的物理理论中全部概念的“操作”意义的观点的影响（第四章）①。因此，他引入了可观察的光谱频率和强度，前者服从里兹组合原则

$$\nu_{mn}=\nu_{mk}+\nu_{kn}$$

如果频率能表示为（经验）项值之差，

$$\nu_{mn}=T_m-T_n$$

上述原则即可满足。对应于频率的“二维”本性，海森伯用量

$$q^0_{mn}\exp(2\pi i\nu_{mn}t)$$

的二维排列取代经典理论中的傅里叶分量$q_n e^{in\omega t}$，而傅里叶级数的现实性由要求

$$q^0_{mn}=q^{0*}_{mn}$$

所表示。里兹原则，加上基于玻尔对应原理精神的考虑，即可导出对二维

排列的乘法规则

$$(xx)_{mn}=\sum_j x_{mj}x_{jn}$$

即使仅凭这些初步的想法，海森伯已能得出许多成果（1925 年 7 月），例如，求一个谐振子的能量，与普朗克的表式 $E_n=nh\nu$ 相对照，可以得出

$$E_n=(n+\frac{1}{2})h\nu$$

海森伯把论文手稿交给了格丁根的玻恩，玻恩认识到海森伯的乘法规则正是矩阵的乘法规则，他与一个年轻的数学家 P．约旦一起，进一步发展了海森伯的思想（1925 年 9 月）。现在，由于两个矩阵 p、q 一般是不对易的，即 $pq-qp\neq 0$，问题首先就在于，如果 p 和 q 是共轭对动量和坐标的矩阵，那么 pq.—qp 的差应当是什么？从对应原理的观点和经学中的作用量 $=\oint pdq$ 出发，玻恩和约旦作出了下述基本假定

$$pq-qp=\frac{h}{2\pi i}I$$

位矩阵。（这里重要的是要注意 p、q 是连续的和无限的矩阵。）p 和 q 的不可对易与普朗克常数 h 的出现立即表明，这一关系在本质上是非经典的。这种“对易关系”已成为量子力学的一个基本假设（见第十章）。

玻恩和约旦（1925 年 9 月）及稍后玻恩、海森伯和约旦（1925 年 11 月），通过假设运动方程具有经典力学中的正则形式

$$\dot{q}=\frac{\partial H}{\partial P}\text{，}\dot{p}=-\frac{\partial H}{\partial q}$$

共同完成了矩阵力学，其中 H 现在是从经典哈密顿函数中用矩阵取代 q、p 并使其对称，以保证 H 的厄密性质而得出的（参见书末附录 8）。

在这些论文中，还借助于幺正矩阵 U，由变换理论保证了 $W=U^{-1}HU$ 的厄密性，微扰理论和各种应用也都发展起来。对氢原子的矩阵力学探讨则由泡利（1925 年）作出。

海森伯、玻恩和约旦的矩阵力学，完成于 1925 年 7 月、9 月、11 月的一系列论文中。薛定谔的系列论文发表于 1926 年 1 月、2 月、5 月和 6 月。薛定谔还于 1926 年 3 月指出矩阵理论和波动理论在数学上是等价的。这一等价性的证明实在是很简单的。

薛定谔方程可由对哈密顿—雅可比方程作形式代换，变 p_x，p_y，p_z 为 $\frac{h}{i}\frac{\partial}{\partial x}$，$\frac{h}{i}\frac{\partial}{\partial y}$，$\frac{h}{i}\frac{\partial}{\partial z}$ 所得出，而这时矩阵力学的对易关系

$$pq-qp=\frac{\hbar}{i}E$$

被看作算符方程

$$\frac{\partial}{\partial x}x-x\frac{\partial}{\partial x}=\frac{h}{j}$$

其后，证明即由表明积分

$$F_{mn} \equiv \int \psi_m^* F \psi_n dq$$

具有矩阵元的性质

$$(F+G)_{mn} = F_{mn} + G_{mn}$$
$$(FG)_{mn} = \sum_i F_{mi} G_{in}$$

得以完成。

狄拉克在发表于 1925 年 11 月、1926 年 1 月和 7 月的一系

列论文中，提出另一种处理非对易数（q 数）的更一般的理论。由经典的泊松括号表式和对应原理，得出了玻恩—海森伯的对易关系。

还应谈到狄拉克的变换理论（1926 年），它隐含着变换矩阵元的概率幅诠释，也是朝向不确定性原理的重要一步。

从历史的意义上，下述故事可能是有趣的。1925 年夏，海森伯应邀到剑桥大学作报告。他的演讲不包括他正在研究的新理论。P.A.M.狄拉克出席了这次演讲。海森伯 7 月论文的一个副本（以长条清样的形式）被送给 R．H．富勒教授，教授把它交给狄拉克看。在这篇论文中，“矩阵”的名字尚未出现，但论文却使狄拉克回想起经典力学中的泊松括号表示，立即去图书馆阅读有关泊松括号的文献，并激发他写下了关于（服从一种非对易代数的）q 数的理论的一系列论文（1925 年 11 月，1926 年 1 月和 6 月），由此发展起一种普遍的量子力学，波动力学和矩阵力学都是这种量子力学的特殊形式。狄拉克还把 δ 函数引入了量子力学，尽管它曾为基尔霍夫（1882 年）于光波理论中、为亥维赛（1893 年）于电磁学理论和在一点上有单一负载的振动弦理论中应用过。单一的 δ 函数是非常有用的，特别是在狄拉克（1926 年 12 月）和约旦各自独立发展起来的变换理论中。

参考文献

M．Jammer，见第八章后参考文献

（历史叙述和有关原始论文索引）

吴大猷，Quantum Mechanics，（见第八章后参考文献）

（矩阵力学，波动力学，原子结构）

P．A．M．Dirac，ThePrincipleofQuantumMechanics，OxfordUniversity Press，3rd ed．1947

W．Heisenberg，ThePhysicalPrinciplesoftheQuantum Theo-ry，Chicago University Press， 1930； Dover ed． ，1949

第十章　量子力学的公设形式及其物理诠释

1．历史进展

为清晰计，我们按时间顺序，列出迄 1927 年 3 月止量子力学的重要进展。

（1）爱因斯坦（1905 年）和德布罗意（1924 年）关系

$$E = h\nu，\ p = \frac{h}{\lambda}$$

与它们的实验证据：光电效应（密立根，1916 年），康普顿效应（康普顿，1923 年；戴维孙—革末，1927 年；G．P．汤姆孙，1927 年）。

（2）海森伯的矩阵力学和对易关系（1925 年 6 月、9 月、11 月）。

$$pq - qp = \frac{h}{2\pi i} I$$

（3）薛定谔的波动力学（1926 年 1—6 月）及其与矩阵力学的数学等价性（1926 年 3 月）。

（4）玻恩关于薛定谔波函数的概率诠释（1926 年 6 月）。

（5）海森伯的不确定性关系（1927 年 3 月）。

$$\triangle x \triangle p_x \geq h$$

它是爱因斯坦—德布罗意关系的一个推论。后来（1928 年）H．韦耳表明，由（2）中的对易关系和玻恩的概率公设（见下文），不确定性关系

$$\triangle x \triangle p_x \geq \frac{h}{4\pi}$$

可作为一个结论演绎得出。

所有上述结果（或公设）无论是在概念上，还是在形式上（由于普朗克常数 h 的关系），按照经典物理学都是不可理解的。爱因斯坦—德布罗意关系把根本不同的波和粒子的概念连接起来，已引出了很严重的困难，而它们的结论，即不确定性关系，则真正动摇了牛顿力学的概念基础。

1927 年 9 月，在意大利科莫举行的国际物理学会议上，玻尔首次提出了一种后来被称之为互补性原理的观点。按照这种观点，爱因斯坦—德布罗意关系应作如下理解。粒子和波的概念是我们从宏观经验中形成的。我们宏观概念的这种起源使得这些概念本身带有固有的局限性。当它们被应用于原子（和亚原子）领域中时，这种局限性就显现出来，而爱因斯坦—德布罗意关系正表明了在原子领域中这些宏观概念的不适当性。在这个意义上，波和粒子概念不应被看做“矛盾的”，而应看做彼此“互补的”。

在同一意义上，坐标 x 与其共轭动量 p_x 是互补的概念；它们在经典物理学中的意义在原子领域中是不适当的,这种在原子领域的不适当性由不确定性关系所展现。

人们或许会问：是否有可能，并且最好通过运用某些适当的概念，

以避免诸如波与粒子、坐标与共轭动量这些概念的不适当性？被称为哥本哈根学派的物理学家（其中最著名的倡议者有玻尔、海森伯和罗森菲耳德）对此的回答是：我们已从日常经验中形成的“经典”概念，是我们仅有的能借以思考并与他人交谈的概念；没有这些概念甚至就没有物理学。这些宏观概念在原子领域中的不适当性（对波和粒子的互补对）表现为爱因斯坦—德布罗意关系，并以更精确的数学形式（对共轭的 x 和 p_x）表现为对易关系。

玻尔关于互补性的科莫演讲似乎没有招致什么反对意见;一方面是因为极少物理学家完全理解玻尔的意旨,另一方面是由于爱因斯坦没有出席那次会议。同年 10 月, 第 15 届索耳维会议（首届索耳维会议举行于 1911 年）在布鲁塞尔举行，会上爱因斯坦既不承认概率诠释，也不接受不确定性关系，从而开始了与玻尔的长期论战，直至 1955 年爱因斯坦去世。我们将在下两章再讨论爱因斯坦—玻尔论战。

2．基本公设

在紧接着量子力学开端的那些岁月里，其数学基础也为诸如希尔伯特、冯·诺埃曼、P．约旦和其他数学家所奠定。海森伯、薛定谔、狄拉克、玻恩、玻尔和其他人的工作被统一进一个逻辑一致的体系内。下面，我们将给出一个量子力学的公设形式的基本表述。为这一考察的明晰起见，我们在每一公设后仅作少量评注，不加详细说明。

这些基本公设可被粗略地分为两类,即有关互补性原理的和有关概率的。

公设 1．物理系统的状态被一个无限多维抽象空间中的矢量所确定。

注释　如果｜a〉（狄拉克术语中的“右矢”）是描述某一状态的矢量，则±C｜a)（其中 C 为一常数）描述同一状态。因此｜a〉可通过〈a｜a〉＝1 归一化。态的变化由｜a〉的变化所表示。

公设 2．所有物理量由线性的厄密算符所表示。作用于态矢｜a〉上的算符改变态｜a〉。

注释（1）　如果 Q 是一个厄密算符，｜q_k〉是一右矢，使

$$Q|q\rangle = q_k|q_k\rangle$$

式中 q_k 为一个数，则｜q_k〉被称为 Q 的一个本征态；并且｜q_k〉是与本征值 q_k 相联系的本征态。

注释（2）　厄密算符的本征值是实数。

注释（3）　属于一个厄密算符的两个不同本征值的本征态矢是彼此

正交的，

$$\langle q_j | q_k \rangle = \delta_{jk}$$

其中$\langle q_j|$是处于与$|q_k\rangle$的空间共轭的空间中的一个矢量（左矢），并有

$$\langle q_j | Q = q_j \langle q_j |$$

即$\langle q_j|$是 Q 的属于本征值q_j的本征左矢。

注释（4） 厄密算符的线性保证了态迭加原理。

注释（5） 一个厄密算符 Q 的本征矢$|q_j\rangle$构成一完全系。一个任意矢$|a\rangle$能被展开为一组完全系

$$<a| = \sum_k \int |q_k><q_k|a>$$

和

$$<a| = \sum_k \int <a|q_k><q_k|$$

其中系数$\langle q_k|a\rangle = \langle a|q_k\rangle$*满足

$$\sum_k \int <a|q_k><q_k|a> = \sum_k \int |<a|q_k>|^2 = 1$$

注释（6）任意厄密算符的本征矢的完全系能被选为基矢。从基矢$|q_k\rangle$到另一基矢$|r_j\rangle$的变换导出量子力学的变换理论。

注释（7） 幺正变换$U^{-1}QU = R$保持了 R 的厄密性质。

注释（8） 变换理论与下文的概率公设一起，使量子力学成为一个完备体系。

公设 3．坐标 q 与其共轭动量 p 的厄密算符遵循关系

$$pq - qp = \frac{h}{2\pi i}$$

注释（1） 这个关系在某种意义上，是玻尔的互补性思想的数学表示。见下述公设 4。

注释（2） 这个关系和公设 4 是与爱因斯坦—德布罗意关系相关联的。

公设 1 和公设 2 是属于“定义”性质的。公设 3 的性质则为基本原理，其意义类似于热力学定律。

所有前三条公设与“物理学”尚无关联。这种联系是由概率公设所提供的：

公设 4．当一个物理系统在态$|a\rangle$时测量其一物理量 Q，则期望值为

$$\langle Q \rangle = \langle a | Q | a \rangle$$

注释（1） 这个公设把前三条公设中引入的数学概念与测量结果联

系起来，是量子力学体系的主要物理内容。

注释（2）　让一个系统在态｜a〉时测量其一物理量 Q。用 Q 的本征矢的完全系〔公设 2，注释（5）〕来表示｜a〉，我们得到

$$\langle Q\rangle = \langle a | Q | a\rangle = \sum_k \int | <a | q_k> |^2 dq_k$$

作为测量结果，任何一个本征值 q_k 都有可能出现，其出现概率为｜〈a｜q_k〉｜^{2}。仅当一个系统处于 Q 的一个本征态｜q_k〉时，才能肯定测量 Q 的结果精确地为 q_k。

这是一个非常重要的结果，它根本违背经典物理学的基本原理。一般而言，它否认了预言测量结果的准确值的可能性，只给出了所有可能出现值的相关概率；而这种精确知识的缺乏并非是由于人们必须处理大数量原子（如同在经典气体动理论中，人们为实用方便引入概率概念）；在这里，即使对一个原子也必须应用概率概念，即概率概念在量子力学中是内禀的。这一结论按公设 4 是不可避免的（参见书末附录 10）。

注释（3）　由公设 3 的对易关系和概率公设 4，韦耳（1928）推演出不确定性关系

$$\triangle x \triangle p_x \geq \frac{h}{2\pi}$$

注释（4）　人们可能要问：概率公设是否为量子力学作为一自洽体系所必需？

如先于公设 4 的段落和上述注释（1）所述，前三项公设 1、2、3 并不构成一个完备的理论，因为“算符”、“态矢”和“对易关系”还必须与对物理测量结果的预言和描述相联系。这些联系是由概率公设所提供的。没有这一公设，这个理论确实是不完备的。

概率公设的一个最重要推论，是它与对易关系一起，导出了不确定性关系，即概率公设是与互补思想一致的，而后者是哥本哈根学派量子力学的基本原理。从量子力学的内禀一致性观点看，这当无问题，但当排斥公设 4 而把前三项公设 1、2、3 称为“互补的公设”时，却引出了一个逻辑问题；因为如果公设 4 是互补思想的一个独立公设，那么就不应出现一致性问题。因此我们必须把概率公设与前三项公设一起，看做一个自洽的体系；把公设 1、2、3 称做“互补的公设”在逻辑上是不正确的，而仅仅是为了用以强调它们在表达互补思想方面的作用（参见书末附录 11）。

关于公设 4 形式的概率公设是否必需的问题，确实是一个涉及量子力学自身基础的深刻问题。我们将在第十二章再回到这个问题的讨论上来。

公设 5．一个物理系统的状态变化，由薛定谔方程

$$-\frac{h}{2\pi i\partial t} | a,t >= H | a,t >$$

决定，或者以波函数 ψ（q，t）来表示：

$$-\frac{h}{2\pi i\partial t}\frac{\partial\psi}{\partial t}=H\psi$$

注释（1） 薛定谔方程本身是一个基本公设，显然由 于普朗克常数 h 的出现，它不能由经典物理学中推导出。注意这个薛定谔方程是非相对论性的。

注释（2） 注意ψ（q，t）〔或通过傅里叶变换为φ（p，t）〕可被用于方程中，但由于公设 3，ψ（q，p，t）则不行。

注释（3） 薛定谔方程的作用类似于牛顿运动方程的作用。

注释（4） 在哈密顿算符 H 中，出现诸如矢势 A 和标势φ这样的势，而在经典物理学中，只有它们的微商，如 curlA（A 的旋度）和 gradφ（φ的梯度）出现。这是一个重要的区别，并导致了玻姆－阿哈罗诺夫效应（1959，1960 年）。（见第十三章，1（3）部分。）

注释（5） 薛定谔方程是一个微分方程，对ψ而言，它是决定论的和因果性的。因为$\psi^*\psi$只有概率性意义，按公设 4 它给出期望值，我们可说薛定谔方程仅对预言概率而言是“决定论的和因果性的”，并非对实际结果如此。

注释（6） 薛定谔方程内部包含某些守恒定理。例如，取一不可分辨粒子的系统（如在一原子中的一些电子）。如果哈密顿算符对两粒子的互易是不变的,则薛定谔方程的解对于两粒子的互易可能是对称的或反对称的，即

$$\psi_s(1,2)=\psi_s(2,1)\text{或}\psi_s(1,2)=-\psi_s(2,1)$$

现在的一个定理是：当ψ按薛定谔方程变化时，其对称性（或反对称性）不随时间改变。

另一个原理的性质则不同于上述公设。它使态函数对两个不可分辨粒子的互易的对称性或反对称性与它们的自旋相关。

公设 6．对自旋为半奇整数的粒子，态函数对于两个不可分辨粒子的互易是反对称的，对于自旋为整数或零的粒子则是对称的。

从这个公设可得出，半奇整数自旋的粒子服从泡利不相容原理（起初用以表示电子）和费米—狄拉克统计，而整数和零自旋的粒子则服从玻色—爱因斯坦统计。

以这种形式表示的量子力学已经成功地应用于原子物理学、分子物理学、固体物理学、核物理学和粒子物理学。然而，仍有若干有关量子力学基础的基本原理的问题，自从理论的诞生起，在物理学家中就没有达成共识。

3．小结

从量子力学公设形式及其物理诠释中可以窥见量子力学的基本思想。为清晰起见，我们扼要地强调以下几点：

（1）所有包含能量或动量转换的基元过程，因 h 的大小有限，所以是不连续的〔在经典物理学中，h 是零，而所有出自大小有限（不为零的）h 的后果给消除了〕。

（2）由此，所有测量，包括被测量的系统和测量仪器之间的相互作用，总包含一种相互扰动，以致不可能做到"人们要想多小就有多小"。

（3）物理学中的一切概念，只有在由它们的实验测量（实际上和想像中）所定义的范围内才有意义，而且测量程序及所得结果必须是能以经典物理学的概念表达的。

（4）经典概念已被发现在涉及原子现象时是不合适的。这种不合适性已由发现 h 的有限大小所揭示,它导致了为爱因斯坦—德布罗意关系所表达的波粒二象性。哥本哈根学派的哲学正是把这种波粒二象性看做粒子和波这些经典概念的根本不合适性的结果,当这两个概念应用于原子领域时就把它们看做彼此"互补"的。爱因斯坦—德布罗意关系就被当做对这种互补性的一种表达。

（5）从爱因斯坦—德布罗意关系出发，海森伯不确定性关系可作为一个推论得出。这必须被认为，不仅意味着在测量两个互补的性质时要求它们同时达到人们所希望的准确性在实践上的不可能性,而且还意味着要超越由不确定性关系所确立的准确性的限度,从概念上定义这样两种性质也有内禀的不可能性。

（6）由于从互补性引发的基本的不确定性，所以，经典物理学的严格的决定论特征在量子力学中是不存在的。为了形成一种新理论以适应这种新形势，有必要对"态"给出一个新的定义。人们设定：一个系统的态，要么用坐标和时间的函数ψ来完备描述，要么用动量和时间的函数φ来完备描述，但不能用坐标和动量两者来完备描述。物理量不能用普通的数描述，只能用算符描述，后者是使理论能够容纳在经典物理学中不存在的不确定性关系的一种方式。ψ随时间演变被设定为受薛定谔方程支配。

（7）为了能与不确定性关系相一致，对ψ的概率诠释被设定为甚至可以应用于单个的粒子。这是与经典理论中适用于很大数量的粒子的统计概念相矛盾的。

（8）在这些基本公设的基础上，有可能给出一个与量子力学中测量相一致的理论。因此，就能表明，互补性质的测量（借助实际的实验程序）在不确定性关系的意义上是互相排斥的。

（9）按照哥本哈根学派的意见，量子力学目前的体系和它们的诠释的逻辑一致性已经确定性地建立起来了。

参考文献

吴大猷，量子力学（见第八章后参考文献）
（公设形式，互补性）
P．A．M．Dirac，见第九章后参考文献
M．Jammer，见第八章后参考文献

第十一章　量子力学本质的哲学问题

在前几章中，我们已经论述了量子力学发展的简要历史。人们看到，各种完全革新性的思想一开始几乎同时由不同的物理学家提出,但具有不同的形式，即德布罗意和薛定谔的波动理论、海森伯的矩阵理论和狄拉克的更一般的理论。一个重要的事实是，在每一种情况下，都是首先充分发展数学部分，然后才是多种理论的统一，最后是它们的物理诠释。这些发展以由公设的形式表述量子力学而达到顶点,我们已在第十章中作了基本的概述（“基本的”是以损失“数学的严密性”为代价的）。这个体系在逻辑上是自洽的，并且也是与哥本哈根学派的哲学诠释一致的，关于哥本哈根学派的哲学诠释已在前面作过简要说明。但是出人意料的是，许多量子理论与量子力学的创始人程度不同地不能接受哥本哈根的观点。爱因斯坦就是一位主要的反对者，普朗克、薛定谔和德布罗意也不满于哥本哈根的态度。我们将尝试对爱因斯坦与哥本哈根学派的争论作一个简略的评论（参见书末附录 12）。

但在做这一点以前,让我们再次不揣重复地概述一下物理学中的重要发展，正是这些发展带来了基本概念和理论的基础性变化。为了避免太多的重复，我们略去了实验支持方面。

（1）17 世纪，牛顿奠定了（经典）物理学的基础，它具有独立的（欧几里得）空间和（绝对的）时间概念，以及隐含决定论与因果性的运动定律，这个体系在两个世纪内未受到物理学家或哲学家的挑战。

（2）1900 年，普朗克多少受实验结果所迫引进了辐射的量子理论。1905 年，爱因斯坦由此进而提出了光子理论，具有如下关系：

$$E = h\nu，p = \frac{h}{\lambda}$$

这就使波和粒子的不同概念等同起来，造成了它与经典物理学（事实上，与“常识”）的基本区别，引人瞩目。

（3）在狭义相对论（1905 年）中，爱因斯坦指出，独立时空、绝对时间和绝对同时性的概念必须抛弃。在广义相对论（1915—1916 年）中，他又引进了非欧时空概念以代替牛顿体系中的欧氏空间。

（4）1913 年，玻尔理论引入了革命性的由量子化条件所规定的定态概念。对原子光谱的广泛而精细的研究，表明氢原子理论是成功的。每一个理论的或实验的结果（原子光谱、光电效应、康普顿效应、施特恩—革拉赫实验中的空间量子化，等等）都显示出普朗克常数 h 的存在，这增加了经典物理学不合适性的证据。

（5）但到了 20 年代初，对玻尔理论的最初的高度热情开始让位给深刻的思考。该理论不仅没有提供使理论推广到多电子原子问题的线索；物理学家们也因为理论的不一致而变得不安,理论中经典概念和原理的使用

只是被忽视，并与外来的公设相矛盾。有些物理学家开始相信，进步不能只通过对经典理论与量子理论拼拼凑凑去寻求,这样的量子理论仍然只具有修正和改进的性质，它明显地还不足以形成一个一致的和完备的理论；一个一致的和完备的理论可能需要一个完全新的基础。

（6）物理学最值得庆幸的时期是在 20 年代早期，当时有一批才华横溢的年轻理论物理学家〔泡利（1900—1958）、费米（1901—1954）、海森伯（1901—1976）、狄拉克（1902—1984）〕，有一批稍为年长一点的理论物理学家〔薛定谔（1887—1961）、德布罗意（1892—1987）〕，还有一大批资深的大物理学家〔普朗克（1858—1947）、索末菲（1868—1951）、爱因斯坦（1879—1955）、玻恩（1882—1970）、玻尔（1885—1962）〕。他们是量子力学和量子理论的创始人。

德布罗意点燃了思想的火花（1924 年），薛定谔发展了整个数学结构（1926 年），而玻恩提供了波动力学的概率诠释（1926 年）。海森伯从一个不同的方向出发提出了矩阵力学的思想并和玻恩、约旦一起完成其数学结构(1925 年)。狄拉克发展出量子力学的普遍形式和变换理论(1926—1927 年）。在量子力学的物理意义上，海森伯从爱因斯坦—德布罗意关系出发发现了不确定性原理（1927 年），此后不久，玻尔提出了他的互补性思想（1927 年）。到 1927 年底，量子力学的数学结构和物理诠释就完备了。在前一章中，我们已经看到了量子力学的公设形式结构。

（7）公设间的内在关系（爱因斯坦—德布罗意关系和不确定性关系并没有明显地出现在公设中）可以方便地以下面的图示表达出来。

薛定谔		海森伯
波动力学（1926 年）		矩阵力学（1925 年）
$q，p=\frac{h}{2\pi i}\frac{\partial}{\partial q}$	$\underset{狄拉克（1927年）}{\overset{薛定锷（1926年）}{\rightleftharpoons}}$	$pq-qp=\frac{h}{2\pi i}$
＋		＋
概率公设		概率公设
玻恩（1926 年）		玻恩（1926 年）
｜		｜
傅里叶变换		韦耳（1928 年）
↓		↓
$E=h\nu，p=\frac{h}{\lambda}$	$\xrightarrow[海伯森（1927年）]{}$	$\triangle p\triangle q=\frac{h}{4\pi}$ （1927年）
爱因斯坦（1905 年）		
德布罗意（1924 年）		

这个图示中，箭头所指的方向意为“导致”。例如，算符（q，p）＝

$(q, \frac{h}{2\pi i}\frac{\partial}{\partial q}$ 满足对易关系，就像薛定谔在他1926年3月的论文中发现的那样，从爱因斯坦—德布罗意关系，海森伯于 1927 年 3 月证明了不确定性关系；韦耳在 1928 年指出，从对易关系和概率公设，不确定性关系就能作为一个推论得出。从对易关系出发，得出的算符形式$p=\frac{h}{2\pi i}\frac{\partial}{\partial q}$，在狄拉克的《量子力学原理》中得以证明。$p=p=\frac{h}{2\pi i}\frac{\partial}{\partial q}$和概率公设使得$p=\frac{h}{2\pi}k=\frac{h}{\lambda}$可以借助傅里叶变换得到证明①。

（8）我们现在可以简述要旨。

在前节中，已经表明经典物理学的失败（至少是不合适）可以概括在爱因斯坦—德布罗意关系中，这个关系展示出基本经典概念的明显矛盾，而这个关系又是得到实验的强有力支持的。它同时也表明，普朗克—爱因斯坦—玻尔的量子理论除氢原子以外未能成功,并且在把经典概念和原理与完全外来的公设结合中常有内在的不相容性。

在前面第九章中所描述的德布罗意和薛定谔的工作,以及海森伯的工作，都力图构建数学框架以处理原子现象。到 1927 年，量子力学的数学结构已经清晰了。这个理论的物理意义 1927 年由玻尔以他的互补性思想提示，并已在本书第十章中作过简略描述。量子力学作为一个“自洽的”和“完备的”体系以公设形式被表达，在本书第十章中也提到了。

从本章前述（7）的图示中可以看出，互补公设（算符 p、q 和它们的对易关系）连同概率公设（对于期望值）一起，导致德布罗意关系和不确定性关系！这就很清楚，引进非经典概念（算符和对易关系）的代价是由演绎出爱因斯坦—德布罗意关系和不确定性关系为补偿的。换言之，使用算符 p、q 和它们之间的对易关系，便放弃了从宏观经验中形成的 p 和 q 的经典概念。（正是由于这个理由，第十章中的公设 1、2、3 被称作互补公设。）但是，这些公设本身并不足以形成一个完备的理论；概率公设是必需的！

（9）现在，根据概率公设，对于一个处于任意态｜a〉的体系中的量 Q 作一次测量，理论并不能预言一个确定的值，而只能以概率$|\langle a|q_k\rangle|^2$预言所有的值 q_k。有些物理学家已经发现量子力学的这个方面是难以接受的。薛定谔通过下面的例子说明了这个推论的本质。试考虑一只猫关在一个封闭的盒子内，盒子上有一个用极化晶体遮住的孔，当一束由两个垂直极化分量构成的光射来时，一个分量将被吸收，而另一个分量将通过这个孔，扣动光子枪的扳机，使枪击发，射向盒内的猫。现在有单个光子通过孔进入盒内。按照量子力学，击发枪机，光子（只作为一个整体）有 50%的机会通过晶体，进入盒内，把猫射死，即量子力学不允许人们知道

这只猫是已经被杀死了还是没有杀死,除非人们打开盒子才能发现猫的死活。

这个结论是量子力学的一个严格推论。哥本哈根学派把这个结论视作我们所能期望得到的答案,并且相信量子力学告诉了我们所有我们将能知道的东西。另外，这种态度还是一种哲学，亦即是我们关于物理理论应该是什么的信念，这也是爱因斯坦和哥本哈根学派之间争论的根源。

（10）哥本哈根哲学的主要信奉者是玻尔和海森伯，泡利、狄拉克、维格纳和其他人也接受目前的量子力学体系,它的数学结构和内在的物理意义，即互补性思想。但就“ 哥本哈根哲学 ”而言，其意蕴不仅是互补性思想，而且也是一种态度，尤其如玻尔所强调的，认为现行量子力学体系是一种“ 完备的 ”理论（这是在这样的意义上说的，即它回答了所有有意义的问题，量子力学不能回答的其他问题必须被认为是“ 不允许的 ”）。因此，这里的概率概念（连同上面涉及的以薛定谔猫为例证的推论），按照玻尔的意见，具有内在的本性，而要问理论所预言的“ 概然的 ”结果背后的问题则是“ 无意义的 ”和“ 不允许的 ”。

但是玻尔的哲学甚至比这一点走得更远。按照他的意见，量子力学现行体系不仅是一个“ 完备的 ”理论，而且也是唯一的理论，即：某些问题不仅是量子力学所“ 不允许的 ”,而且也是我们的知识本性所“ 不允许的 ”。换句话说，玻尔相信，量子力学的现行体系已经规定了我们知识的限度，也规定了我们探求理解的限度。

薛定谔（1955 年）不能接受这种态度，他认为这是教条。尽管对量子力学作出新诠释或新表述的许多尝试都屡屡失败（德布罗意、玻姆及其他人），但玻尔（还有海森伯、罗森菲耳德）的态度似乎是太强了。我曾一度（1956 年）把这些情势与另一种情势作过比较，在后一种情势中人们坚持欧几里得几何学的内在一致性，并否定其他几何学的可能性，当然这种比较或许不甚准确，因为量子力学不仅包含逻辑一致性，而且也与物理现象对应。

参考文献

A.Pais，见第四章后参考文献

N. Bohrandothers，articlesinP. Schilpp ’sAlbertEinstein：Philosopher-Scientist，Tudor，NewYork，1949

第十二章　爱因斯坦—玻尔关于量子力学诠释之争

爱因斯坦—玻尔的讨论与争论开始于 1927 年 10 月的索耳维（Solvav）会议。在那次会议期间，爱因斯坦不接受不确定性原理（这年的 3 月发现该原理）并试图构思出一些思想实验，使之能超越不确定性关系的局限。玻尔则能通过爱因斯坦—德布罗意关系而扭转每一种情况。爱因斯坦在他的关于一个放射性原子的能量与它衰变时间的知识的思想实验被玻尔诉诸爱因斯坦自己的广义相对论（引力论）而作出成功回答之后，最终放弃了要找出“违背”不确定性关系的事例的任何进一步的努力。爱因斯坦似乎已承认了量子力学的内禀一致性，但他依然反对内禀的概率，反对关于物理理论本质的哥本哈根哲学。

爱因斯坦相信，自然界是决定论的（在一封给玻恩的信中，他说：“上帝是不掷骰子的”，这是在文献中屡屡被引用的一句话）。他接受量子力学（概率）分布的结果，但他希望把概率诠释为是在我们讨论许多原子时出现的，如同我们在讨论气体动理论时出现的那样，而不是单个原子的内禀概率。现在，在动理论中，使用概率方法是为了方便，但基础则是决定论的动力学定律。因此爱因斯坦的观点已导致了许多关于量子力学概率是否也具有一种内在的决定论基础的理论探索。在动理论中，最终只有平均值（宏观概念）出现，分子坐标和动量不再出现。因此，问题就成了量子力学中是否有“隐变量”存在。研究似乎表明，在量子力学目前的形式中，回答是否定的。

然而，这个结果并没有改变爱因斯坦的基本观点，因为这种情况并非是量子力学还是爱因斯坦对或不对的一种“证明”，而正体现了爱因斯坦关于物理理论本质的观点。爱因斯坦坚持，一个“完备的”物理理论必须包含物理实在的一切元素，而目前的量子力学由于不能同时给出例如 p 和 q 的精确知识，所以它是“不完备的”。换言之，爱因斯坦不满意目前的量子力学，因为它从根本上否定确定性知识。

从前两章我们关于量子力学现行体系连同哥本哈根诠释的讨论可以看出，通过实验测量以定义坐标和动量，这一观点本是爱因斯坦在相对论中的观点。波粒二象性就被简单地理解为经典概念应用于原子领域时的局限性的一种表现。不确定性关系也是在这种相同的意义上被理解的。态函数ψ被认为是对一个系统的完备描述，即ψ包含了具有物理意义的所有信息。因为互补性和概率公设，理论只能给出测量结果的概率预言，而 $|\psi|^2$ 正是对单个系统给出了这种内禀概率。

爱因斯坦这位第一个在相对论中用测量定义空间和时间概念的人，现在在量子力学中却从这种观点后退了。他坚持认为，即使 p 和 q 在它们的同时测量中受到不确定性关系的限制，但它们的精确值在我们不测量它们时是存在的，因此，p 和 q 应在ψ中都出现。由此他相信，量子力学用态

函数ψ所作的描述是不完备的。

爱因斯坦强烈反对量子力学中的内禀概率诠释。他相信大自然是决定论的。他用“上帝不掷骰子”来表达这点。他相信，$|\psi|^2$应被诠释成给出概率，但这一概率不是对单个系统的，而是对系综的，即类似于经典统计理论的那种系综。然而，这就暗示了量子力学中迄未所知的所谓“隐变量”的存在。（在热力学中，使用宏观变量，诸如温度、体积和压力等，就暗含着隐匿了单个分子的坐标和动量——“隐变量”。）

爱因斯坦对目前量子力学的批评表达在爱因斯坦、波多耳斯基和罗森1935年5月的一篇论文中（《物理评论》，第47期，第777页），其基本哲学是，所有能被测量的量都是物理实在的元素，一个物理理论只有在包含了所有物理实在的元素时，它才是完备的。

令两个系统A和B在一段时间内彼此有相互作用，在这段时间内，两个对易量p_A+p_B和q_A-q_B被测出且精确已知。再令这两个系统在空间分离到很大的距离，使得从一切实际效果上看，它们彼此间不再有相互作用。现在如果测量p_A，就能从p_A+p_B的总和中确定地推演出p_B，因为在A上所作的任何观察都不可能影响到B。另一方面，人们能够测量q_B。因此就得出，p_B和q_B原则上都是物理实在的元素，并且它们能达到任何所需要的精确度。

为了避免与不确定性关系的矛盾，这些作者相信，要么①人们必须假定,在一个系统上所作的测量不可能对作为一个整体的结合系统A+B没有干扰,即使A和B在空间分离得足够远,要么②人们必须抛弃这样的假定，即ψ对一个系统的描述是完备的。

这几位作者相信，我们必须保留这样一个“明智的”假定，即有可能对一个系统进行观察而不在可觉察的程度上影响在空间上与之分离得很远的另一个系统。因此他们只能得出结论，用ψ作出的量子力学描述是不完备的。这种不完备性是指，量子力学只容许要么p要么q而不能两者同时有精确的了解，而上述假想性实验表明，p和q两者都有精确的意义。

玻尔对这个论证的回答是，一旦在系统B上测q_B，那么由于h的有限大小和不可控制，系统和仪器之间有限的相互作用就将破坏对p_B任何精确的了解，以致p_B和q_B仍然不可能同时知道，它们只能服从不确定性关系给出的精确度。因此，在现行量子力学体系中的ψ描述在这样的意义上是完备的，即它包含了人们有权索要的所有信息。

在另一个例子中，爱因斯坦考虑一个放射性核。量子力学只给出衰变的概率，而不能给出衰变的确切时间瞬间。爱因斯坦相信，尽管衰变的能量和衰变的瞬时无法同时精确地确定,但仍然必定有一个与衰变恰好同时的瞬时，而由于现行的量子力学并不包含这个瞬时，因此，它不可能是一个完备的理论。

玻尔的论证是，由于“时间”只有当它被测量时才有意义，又由于对

衰变时间的精确测量将使系统的能量完全不确定，反之亦然，故而该理论已经是完备的，这就是说，它回答了人们有权提问的所有问题。

从前述讨论，人们必定注意到，爱因斯坦和玻尔的态度之不同，在于对待这样一个问题的基本态度，即：具体体现在一个理论中的概念应该是什么类型的，或者说，我们所要的是哪种类型的量子力学理论。玻尔的态度是，现行体系不仅在逻辑上自洽，而且包含了人们有权追问的一切信息。另一方面，爱因斯坦则相信，现行体系虽然完全自洽，但仍然是不完备的，它对于一个完备性理论的地位，多少有点像经典理论中统计力学相对于力学的地位那样。

关于爱因斯坦和哥本哈根学派之间的这种争论,在许多文献中还存在着某些混乱。似乎很少有讨论明确地集中在爱因斯坦观点与量子力学之间的真正不同点上。一个突出的例子是在希耳普编的《阿尔伯特 ·爱因斯坦：哲学家—科学家》（1949 年）一书中，书中除了爱因斯坦自述之外，还有许多其他文章。在玻尔的文章中，重复的论证都建立在量子力学自洽性基础上。这种论证如果在早期，爱因斯坦还拒绝不确定性原理时（例如在 1927 年 10 月的科莫会议上），那会是恰当的，但到了后期，在爱因斯坦—波多耳斯基—罗森的 1935 年著作（爱因斯坦在其中已经清楚地表达了他关于物理理论“本质”的“哲学”）以后，玻尔和别人的许多论文就似乎是语不中的的了。因为这些论文只是有助于说明体系的逻辑上的一致性，而没有面对爱因斯坦针对现行体系基础本身的质疑。爱因斯坦与玻尔态度之间分野的要害在于他们关于物理理论本性的质疑。我记得在 50 年代中期（？）读到一篇罗森菲耳德对 BBC 谈话的稿子，其中的语言带有不必要的讥讽。这激起我写了一篇论量子力学物理基础和哲学基础的文章。1956 年 3 月，我撰文就福克、薛定谔、朗德和其他人论量子力学和论爱因斯坦哲学的工作进行了讨论。

这里再次需要强调指出的是，对于像这样的一种基本态度的不同，所有论证包括现行体系的自洽问题在内都是完全不相干的。单独以现行体系的内在一致性为基础来反对爱因斯坦的态度,正好像以欧氏几何的内在一致性为基础去反对非欧几何一样。即使构建另一种量子力学体系的努力屡遭失败，但一种真正的科学态度要求一种开放的精神。科学只有通过对现存状况作连续不断的重新考察,甚至对那些似乎已确立的理论作连续不断的重新考察，并且寻求新思想，才能获得进步。

参考文献

A. Einstein，B. PodolskyandN. Rosen，Phys. Rev. 47，777 （1935）第十三章　经典场论

和量子场论

1.量子力学中的场

经典物理学中，我们所熟悉的两种场是电磁场和引力场。

在经典理论中，直接可测量的量是力，例如作用于一个粒子或物体或电荷上的力。因此即使当“势”的概念被引入引力场和电磁场时，出现于运动方程或麦克斯韦场方程中的也只是它们的微商。一个场的“势”，在引力场中仅被定义为一个常数，在电磁场 A、中则为一规范变换。

在量子力学中，直接出现于波动方程中的是“势”而不是“力”。因此一个粒子的薛定谔方程是

$$(-\frac{\hbar^2}{2m}\nabla^2+V)\Psi=-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t}$$

而电磁场中一个电子的狄拉克方程是

$$\gamma_\mu(\frac{\hbar}{i}\frac{\partial}{\partial x^\mu}-\frac{e}{c}A_\mu)\Psi-imc\Psi=0 \qquad (1)$$

式中出现的是势 V、A、Ψ而不是力。

重要的问题是：势是否有可观察的效应？

（1）不可积的相

容易看出，一个自由电子的波函数Ψ^0满足

$$(\frac{\hbar}{i}\frac{\partial}{\partial x^\mu}-imc)\Psi^0=0 \qquad (2)$$

它与（1）式中波函数Ψ之不同在于位相

$$\Psi=\Psi^0\exp(\frac{ie}{\hbar c}\int A_\mu dx^\mu) \qquad (3)$$

这个位相$\beta\equiv\frac{ie}{\hbar c}\int A_\mu dx^\mu$在两点$x^\mu_{(2)}$和$x^\mu_{(1)}$之间的差别，依赖于从$x^\mu_{(1)}$到$x^\mu_{(2)}$的路径。（我们称这位相为不可积的。）

（2）可积的相

如果方程（1）中的（A，Ψ）经过规范变换

$$A'=A+\nabla\chi，\Psi'=\Psi-\frac{\partial\chi}{\partial t}$$

其中

$$\Box\chi=0，\chi\text{为一标量函数}$$

容易得出方程（1）变为

$$\gamma_\mu=(\frac{\hbar}{i}\frac{\partial}{\partial x^\mu}-\frac{e}{c}A'_\mu)\Psi=imc\Psi' \qquad (3)$$

其中

$$\Psi'=\Psi\exp(-\frac{ie}{\hbar c}\chi)$$

由此，两点$x^\mu_{(2)}$和$x^\mu_{(1)}$之间的相差独立于路径$x^\mu_{(2)}$—$x^\mu_{(1)}$。（我们称此位相

为可积的。）

（3）阿哈罗诺夫—玻姆实验阿

哈罗诺夫—玻姆（1959 年）提出由下述实验来探测位相（3）的效应。取一细长的螺线管（垂直于纸平面），让一束电子在点 A 处被分为纸平面中环绕螺线管的两个半圆 C_1、C_2，并于 B 点重新汇合起来。除了螺线管内部以外，磁场 H 处处为零。因为 $A_\mu dx^\mu$是洛伦兹不变量，

我们可取一坐标系，其中$dx_4=0$，这样$\int A_\mu dx^\mu \to \int (A\cdot dx)$（A・dx）

。B 点处两束电子的相差为

$$\int_{C_1} A\cdot ds-\int_{C_2} A\cdot ds-\oint (A\cdot dx)=\int_\sigma H\cdot d\boldsymbol{\sigma}\neq 0$$

其中 σ 为被两束电子圈起的面。由这种相差引起的干涉效应已为钱伯斯（1960 年）所观察到。

这个实验非常重要：①在量子力学中，势 A 产生了一个可观察效应(尽管与规范变换相联系的势将不具有这一阿哈罗诺夫—玻姆效应）；②螺线管场的存在好像改变了空间的（拓扑）性质——它不再简单地像在没有 A 场时那样连接了。

这里可以提一下，另一个磁场影响波长的位相的例子，是狄拉克的磁单极理论（1931 年）。

2．统一场：引力场和电磁场

正如我们所已经知道的，最简单的两种场统一的例子是电场和磁场的一。这种统一的取得是通过引入一个四维矢量（A_1，A_2，A_3，A_4=（A，$\frac{i}{c}\psi$）和张量

$$F_{\mu\nu}=\frac{\partial A_\nu}{\partial x_\mu}-\frac{\partial A_\mu}{\partial x_\nu}$$

其分量给出了两个三维矢量 E，B 的 6 个分量。在这个例子中，闵可夫斯基四维空间是平直的，即电磁场不影响自由场空间的度规。这一统一不仅“令人愉悦”，而且也是狭义相对论理论的必要成分。

在爱因斯坦的引力理论中，牛顿的万有引力“作用”被看作是四维空间的内禀性质，即“引力”是植根于时空结构的度规性质中的。这可以与电磁相互作用作一对照。在他完成引力理论（1916 年）后不久，爱因斯坦就开始考虑电磁场和引力场的统一。从上所述，我们可以看出这一问题的困难性质。爱因斯坦后来在这个问题上花费了大量精力。

关于这种统一的第一个理论，是由数学家 T．卡鲁查发表的；他的手稿于 1919 年（？）交给爱因斯坦，并由爱因斯坦于 1921 年送去发表。卡鲁查提出了一个五维世界，具有度规

$$d\sigma^2=g_{\alpha\beta}dx_\alpha dx_\beta,\quad \alpha,\ \beta=1,2,3,4,5$$

其中 $g_{\alpha\beta}$ 被假设具有简单的形式

$$
g_{\alpha\beta}=\begin{array}{ccccc}
5 & 1 & 2 & 3 & 4 \\
\hline
1 & A_1 & A_2 & A_3 & A_4 \\
A_1 & g_{11} & g_{12} & g_{13} & g_{14} \\
A_2 & g_{21} & g_{22} & g_{23} & g_{24} \\
A_3 & g_{31} & g_{32} & g_{33} & g_{34} \\
A_4 & g_{41} & g_{42} & g_{43} & g_{44}
\end{array}
$$

其中（A_1，A_2，A_3，A_4）是电磁场的四维矢势，而 $g_{\mu\nu}$（μ，ν=1，2，3，4）是（爱因斯坦理论中的）四维（黎曼）空间的度规张量分量。所有 $g_{\alpha\beta}$ 都假定分别为 x_1，x_2，x_3，x_4 的函数。场方程（通过扩展爱因斯坦理论）假定为

$$R_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}R=-kT_{\alpha\beta}$$

其中 $R_{\alpha\beta}$ 是（收缩后的）黎曼张量，$T_{\alpha\beta}$ 是排除了纯电磁学部分的能量—动量张量。通过对一质量为 m、电荷为 e 的单个粒子的情况作特殊简化，

$$T_{\alpha\beta}=mu_{\alpha}u_{\beta}，\ u_{\alpha}=\frac{dx_{\alpha}}{d\boldsymbol{\sigma}}$$

于是对 α，β=1，2，3，4，人们得到引力场方程；对 α，β=1，2，3，4，5，人们得到麦克斯韦方程；此时人们须确定 $\mu_5=e/m$。$m\mu_{\alpha}$ 是五维动量—能量—电荷矢量。

五维理论是有趣的，但结果多少来自对 $g_{\alpha\beta}$ 构造的不同假定。爱因斯坦于 1923 年致力于继续研究这一理论，但在 1926 年，O．克莱因发展了卡鲁查的理论，并把新发展起来的量子力学结合进来。爱因斯坦于 1929 年与爱丁顿一起致力于五维理论；于 1930 年至 1932 年间与 W．迈耶尔致力于它；于 1938—1941 年间又与 P.柏格曼和 V.柏格曼致力于它。此后不久，他放弃了这条研究途径。

由克莱因的工作和以后的发展，人们认识到把量子力学原理与五维理论联系起来，只会使电磁场和引力场的统一更加复杂化。看起来广义相对论与量子力学的结合是没有希望的（泡利，1955 年）。场理论在爱因斯坦的意义上是决定论的，因而不能以一种根本的方式与其基础为不确定性原理的量子力学相混合。近年来，统一引力场和其他已知场（电弱的和强的）的尝试采取了新的方向（见下面第 7 节）。

3.经典场

在经典物理学中，“粒子”和“场”是两个不同的概念；它们涉及不同的现象，并分别用于描述分立性和连续性。一个系统可能包含巨大数量的粒子，例如气体中的分子，但只要它们是可数的，基本理论就是经典力学。但如果一个系统的变量是不可数的，如电场强度，则基本理论就是所

谓的场论①。

在经典场论，例如电磁场中，定律通常以偏微分方程表示，这些方程为从拉格朗日密度得出的哈密顿方程，且皆为相对论性协变形式。

但自爱因斯坦 1905 年的辐射量子论始，诸如光电效应和康普顿效应等直接实验对电磁辐射的粒子性质给出了有力的证据。这引导我们去寻求一种数学理论以描述连续场的量子性质——类似于量子力学中原子和分子现象的量子化。把量子力学拓展到场的量子化性质，这个理论就称作量子场论。

粒子动力学的“量子化”数学方法由为索末菲和威尔逊所普遍化的玻尔量子条件开始，

$$\oint p_k dq_k = n_k h$$

这个条件是一开始引导玻恩和海森伯在他们的矩阵力学中假定对易关系

$$pq - qp = \frac{h}{2\pi i} I$$

的线索。

我们已经看到，这个关系如何与量子力学中的概率公设一起，成功地处理了经典物理学中所描述的“粒子”和“波”，并导出爱因斯坦—德布罗意关系式

$$E = h\nu，p = \frac{h}{\lambda}$$

现在，我们希望从粒子回到老问题——电磁场的量子理论，以把粒子动力学（坐标和动量、角动量和能量）的量子化方法扩展到连续场。

这种方法由下述程序组成：

（1）以经典动力学的哈密顿方程形式中的四维势，来表示电磁场规律，即麦克斯韦的场方程组。

（2）定义场变量（“坐标”，例如四维势及其正则共轭）。

（3）通过类似于上述对易关系的关系使场变量量子化。

这种程序当用之于所谓“自由场”（例如不受电荷密度 ρ 和电流密度 J 影响的电磁场）时，能够很容易地实行，并且事实上量子化场由光子表述。这一理论首先由狄拉克于 1927 年提出。

具有四维势的电磁场并非最简单的经典场。如果我们把克莱因—高登方程

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi，\lambda_c = \frac{h}{mc}$$

看作一经典标量 ψ 场的方程，则量子化过程甚至更为简单。量子化场也适用于描述 π 介子（π^0，π^1，π^2）。

4.量子化场

我们已经看到，由法拉第引入、麦克斯韦发展的电磁场的概念——尤其是连续统的概念——首先被电磁现象的整个波谱所确立,尔后又发现了其他一些现象，这就迫使物理学家引进量子概念，这是一个关于离散的概念。这个二象性的两难推论，在概念方面，在哥本哈根学派的实证哲学中是用“互补原理”来“解决”的。在数学方面，这个二象性是用“量子化场”的形式体系描述的，在“量子化场”中，具有能量和动量这些粒子属性的光子乃是使电磁场“量子化”的结果，而荷电的粒子（电子）则是使由粒子的波函数所描述的“场”量子化的结果。电荷之间的相互作用，是通过它们与电磁场的相互作用进行的（电荷本身是电磁场之“源”）。法拉第最初的这种观念,现在被翻译成为相互作用场——电子场和电磁场——的量子理论语言了。相互作用场的数学理论成为量子电动力学的主题。这个理论的发展自 1930 年代以来使许多大物理学家为之忙碌。这些大物理学家中，30 年代有海森伯、泡利、狄拉克、费米；40 年代后期有朝永振一郎、施温格和费因曼。这种相互作用场理论的数学结构有着许多非常基本的困难——在理论中出现无穷大问题难以消除。这个理论决非一个单纯的形式结构；它导致了电子的所谓兰姆移位和 g 值（磁矩与角动量之比），它是由兰姆、卢瑟福和库什在实验中发现的。

量子化场的概念，首先是从电磁场和电子场（狄拉克场）发展起来的，接着由汤川秀树卓有成效地于 1935 年应用于一种当时还未知的场，如今我们称它为介子场，它与核子（质子与中子）的耦合可以说明核子间的强相互作用。介子场的量子化“粒子”后来在宇宙射线中找到，接着又在实验室的高能加速器中发现，例如π介子。

法拉第的同一个基本思想,即粒子通过与一个场的相互作用而形成粒子间的相互作用，也被费米于 1933～1934 年应用到电子—中微子场，以说明另一种相互作用——所谓的弱相互作用——它们支配诸如原子核的β衰变和介子的π-μ衰变。

下表列出各种已知场及其所起作用的相似性。

场	量子化“粒子”	强度	耦合
介子	π介子	“强”	核子与核子
电磁	光子	$\alpha=\frac{e^2}{\hbar c}$	电荷与电荷
中微子	中微子	“弱”	核子—电子—介子
引力	引力子	“极弱”	（质量）粒子与粒子

在基础物理学发展的目前阶段,对于自然界中所有各种相互作用的统一性还远未得到完备的理解，不过，它们之间在定性的相似性上已有明显的统一。物理学中真正的基本问题之一，就是这些相互作用的本质问题。

1928 年，狄拉克建立了电子的相对论性波动方程。它取得了下述成

功：它的形式是洛伦兹不变的；它以回磁比 g=2 包含电子自旋；它预言了反粒子正电子的存在，并于 1932 年为安德森在实验中发现；它给出了氢原子能级的正确的精细结构。狄拉克方程也有一些“困难”。其一是把这一为单一电子建立的理论推广到一个多电子系统。另一困难则更为基本，即着手于描述单一电子的理论结果却由于为说明处于负能态的无限电子海而不可避免地具有多体效应。在一个强场中，一个电子不再孤立存在，而是一个包含无穷多（并且其数目非不变的）电子和正

电子的系统。关于这一点，在狄拉克理论应用于Z趋近于 $\frac{1}{\alpha}=137$ 时氢原子中的电子的例子中已有所提示，当 Z=137 时，通常的能级理论遭破坏。因此，严格说来，单一电子理论仅在自由电子的极限情形中有效，并且必然为“多电子”理论所取代。

眼下，一个数目无限大并且可变的“粒子”系统的适当表述是“场”——正像“光子”表述的是电磁场。由此我们可以说，从粒子的观点出发，人们寻求一种满足相对性原理的粒子理论,结果发现合适的理论为一多体理论，而多体系统的适当表述是“场”。

这样，问题就在于构造一多电子场论。人们可以从一自由电子的狄拉克方程出发

$$\left(\gamma_{\mu}\frac{h}{2\pi i}\frac{\partial}{\partial x_{\mu}}+im_0c\right)\Psi=0$$

并把它看做一经典场Ψ（在此情形中为一四分量场）的方程。然后把 Ψ_{μ} 看作场变量（类似于电磁场中的四维势 AΨ）及其正则共轭。再以对易关系形式引入量子化条件，如所期望的，人们可得到电子作为量子化场的粒子。这样一种量子化程序被称为“二次量子化”。量子化的ψ拥有波动和粒子两种性质——正如量子化的电磁场拥有光子和波动性质一样。

纯电子场的量子理论不存在困难，但对电子加上电磁场的耦合系统，一种相对论性量子理论（即量子电动力学）则存在严重的根深蒂固的困难，即物理量计算结果中顽固的无穷大问题。

紧随狄拉克 1927 年的量子化电磁场论而来的，有约旦和维格纳、海森伯和泡利以及费米（1930 年）等人的工作。但直到 40 年代中叶，随着日本的朝永振一郎和美国的施温格、费因曼和戴森等人的工作，才取得突破性进展。尽管无穷大依然存在，但这种理论以一种确定的、协变的方式成功地“消减”它们，从而可得出有限的结果，并与观察到的兰姆移位和“反常磁矩 g”〔g=2（1+α／2π+…），α=1／137〕精确地一致。从 40 年代中叶到 50 年代中叶的 10 年，是热衷于量子电动力学研究的时期，既进一步分析这一“重整化”理论，也致力于使理论摆脱无穷大（即不仅仅分离和掩盖它们）。但是，迄今仍未找到一种令人满意的解决问题的办法①，而同时，物理学家已把他们的兴趣转向其他领域——基本粒子及其相

互作用和它们的统一。

5.粒子和相互作用

1935 年，汤川秀树提出了关于两个核子间的“强”相互作用的介子理论。介子（π介子）的波动方程（取为克莱因—高登方程）给出了介子场，经量子化导出了π介子。

30 年代早期，泡利提出了中微子假设，费米则对作为“弱相互作用”典型的β衰变提出了中微子理论。 1932 年，查德威克发现了中子，而在 30 年代后期，实验研究证明了核子—核子相互作用与电荷无关。1946—1947 年间π介子的发现证实了汤川秀树的“强相互作用”。到 50 年代前期，在自然界中已知有四种相互作用，按强度顺序依次排列，即

（1）核子间的强相互作用，强度级为 1，

（2）电磁相互作用，强度级为 10^{-2}，

（3）β衰变、π～μ衰变中的弱相互作用，强度级为 10^{-12}，

（4）引力相互作用，强度级为 10^{-29}。

到 60 年代，随着高能加速器的发展而开展的实验研究和理论工作，已导致了许多粒子、包括中微子（电子中微子和μ子中微子）、介子和超子的发现。到 80 年代，已知上百种粒子：

轻子：电子，μ子，τ轻子，与之相应的中微子；

重子：核子，Λ，Σ^+，Σ^0，Σ^-，Ξ^-，Ξ^0，Ω^-，$\triangle^-$，$\triangle^0$，$\triangle^+$，$\triangle^{++}$，Λ_c，Λ_b；

介子：$\pi^\pm$，π^0，$K^\pm$，K^0，η，η'，$\rho^\pm$，ρ^c，J／ψ，Y 族；

夸克：u，d，c，s；（t），b，每种味有三种可能的色（红、黄、蓝）。

尚未发现孤立形式中的这些夸克,但有很强的证据说它们是构成重子和介子的“基本”粒子。例如，质子由（u，u，d）组成，中子由（u，d，d）组成，而π^+由（u，$\bar{d}$）组成，其中$\bar{d}$是d夸克的反粒子。

这些粒子及其“量子数”（例如重子或轻子数，电荷，自旋 J，同位旋 I、I_3，超荷 Y，奇异数 S，粲数 C 等等）是如此数目众多，我们不想在这里解释它们的意义。我们的所有兴趣在于理解不同的相互作用，以及统一这些相互作用的可能性。这里的统一，与爱因斯坦试图统一电磁与引力相互作用中的意义相同。

6.统一场：电磁与弱相互作用

在上一节里，我们已经认识了四种相互作用。我们现在的兴趣是探讨它们是否能统一起来。为此目的，相互作用（场）的规范变换的对称群将作为指导性原理。

在第六章第 7 节中，我们已知电磁场和狄拉克方程在规范变换（现在是四维表示法）

$$A'_\mu = A_\mu + \frac{\partial \chi}{\partial x_\mu}$$

$$\psi' = \psi \exp(-\frac{ie}{hc}\chi)$$

下是不变的。这些变换形成局部规范群 U（1），（由ψ到ψ'的变换矩阵为 1×1 幺正矩阵）。

由β衰变

$$n \rightarrow p^+ + e^- + \bar{v}$$

所代表的弱相互作用涉及一个由 2×2 幺正矩阵表示的同位旋空间的变换，这一变换属于规范群 SU（2）（二维幺正单模群）。

电磁相互作用与弱相互作用的统一，由温伯格、萨拉姆和格拉肖于 60 年代的研究所取得。两种相互作用统一为一种现在称之为"电弱"的相互作用，它属于对称性 SU（2）×U（1）。电弱相互作用的传递者为中间玻色子 W^+、W^-、Z^0 和γ（γ是仅有电磁相互作用时的光子）。$W^\pm$ 和 Z^0 的质量分别约为 82GeV 和 94GeV。

这一理论的成功确实是令人惊异的①。

7.强相互作用

强相互作用理论的历史很长，也很复杂。它始自 50 年代早期费米和他的同事们的工作。很快就发现了与不同的"基本"粒子及其激发态相一致的大量"共振态"。对这些粒子的衰变过程，用非常类似于原子光谱分析的方法进行了分析。为了使衰变结构系统化，引入了一些不同于熟悉的质量、自旋、电荷等的概念，以便在目前即使其中某些物理意义尚未搞清的情况下，可通过经验选择定则来表达某些规则性。这使人联想起 20 年代早期原子光谱学中的情形，当时在 1925 年电子自旋引入之前，许多诸如 J、S 这样的量子数的意义也不清楚。

强相互作用中（多少经验地）引入的"量子数"是：

宇称	电荷	自旋	同位旋	奇异数	超荷	重子数
±	Q	J	I	S	Y	B

其中

Y= B+S

$$Q = I_3 + \frac{1}{2}Y，I_3是I的z分量$$

经验的"选择定则"是：

1）在变换（衰变）中，重子数不变，△B=0

表示"重粒子守恒"；

2）总电荷数不变；

3）奇异数 S 不变，即

$$\triangle S=0$$

这一定则有时被破坏，例如在下述情况下：

$$\Lambda \rightarrow p^{+}+\pi^{-}，\triangle S=（-1）-（0）=-1$$

$$\Sigma^{+}\rightarrow p^{+}+\pi^{0}\ \triangle S=-1$$

但这些（违背定则的）过程非常慢，这些转变中的Λ、Σ^{+}、Σ^{0}的寿命约为 10^{-10}秒，相比之下，那些遵从△S=0 的寿命约为 10^{-23}秒。

4）从未发现违背 B 和 L 的守恒（其中 L 是轻子数）的情形。

这些定则与大量关于许多强子和介子的经验结果(其中有些见本章第 4 节）一起，导致盖耳—曼和奈曼分别独立建议这些粒子（或“层”）由 SU（ 3）群来表示。这些思想引导盖耳—曼和兹韦克于 1964 年分别独立地提出夸克理论,认为每个核子和介于分别由 3 个和 2 个夸克和反夸克构成。现行的理论中有六种“味”的夸克，每味又有三种“色”，及其反夸克，所有这些都具有自旋 1/2 和重子数 1/3。下页表概括了它们的性质。

所有的反夸克都有相同的 I，但对重子和所有其他量子数都带有相反的符号。

由此，对质子、中子、π介子，我们有

p（u，u，d），n（u，d，d），π^{+}（u，$\bar{d}$），π^{0}（u，$\bar{u}$），（d，$\bar{d}$），π^{-}（d，$\bar{u}$）。

夸克 / 性质	u 上	d 下	s 奇异	c 粲	b 底	(t)（顶）
同位旋 I	/1/2	1/2	0	0	0	0
I	1/2	－1/2	0	0	0	0
电荷 Q	2/3	－1/3	－1/3	2/3	－1/3	2/3
奇异数 S	0	0	－1	0	0	0
粲数 C	0	0	0	1	0	0
底数	0	0	0	0	－1	0
顶数	0	0	0	80	0	1

关于夸克的量子理论的简略概要，请见书末附录 13。强相互作用与电弱相互作用的统一，将意味着一个立足于规范群 SU（3）×SU（2）×U（1）的理论。这种相互作用的传递者是胶子，而以电弱理论作为一个部分。这一理论尚未完成。

8.引力场

关于引力波存在的证据已由韦伯报告过，但迄今还是非结论性的。所以，“引力子”目前还只是一个理论上的概念。自从爱因斯坦和韦耳时期以来,理论物理学的最终目标就是把所有已知的相互作用统一进一个单一

的相互作用。近年来，有许多尝试性的理论想把引力场和其他的（电弱和强）场统一起来，“弦”和“超弦”理论在数学上是很难的（并且是猜测性的，像所有全新理论一样）。评述这些尝试已非本书的范围所能及。关于对称性和场的进一步论述，请见书末的附录 13。

参考文献

T.Y. WuandPauchyW.Y.Huang,RelativisticQuantumMechanics and Quantum Fields，WorldScientificPubl. Co.，Singapore，1991

（经典场论，量子场论，量子电动力学，标准模型,阿哈罗诺夫-玻姆理论）

W. Pauli，见第四章后参考文献

A.Pais，见第四章后参考文献

（卡鲁查和克莱因的统一场论）第十四章　结束语：物理学中实验与理论的相互作用

（1）我们已经考察了物理学的重要发展——它的各个分支，尤其是它的基础的重要发展。在本世纪中，经典物理学经受了三次革命。狭义相对论揭示了光速的有限性,抛弃了绝对时间的概念并在诸如惯性系之间的对称性这样的基本原理基础上建立了一个理论。这一理论由此而使麦克斯韦电磁理论摆脱了不可观察的和不必要的以太。广义相对论（和引力理论）由于摆脱了狭义相对论所受惯性系的限制而进一步向前发展，并改变了物理学家承袭下来的不加怀疑地接受的欧几里得空间和牛顿的引力理论。量子力学，至少按互补性观点，由像波—粒概念、不确定性原理这样的二难推理，进而认识到我们基本概念的本质（凭借它们的宏观起源）并用数学量代替了不可观察的轨道，等等。在从早期的粒子概念发展到场论的过程中，在处理“场”和“相互作用”时，理论借助于对称性原理和数学（群论）形式取代了素朴的形象化方法。这就是现代物理学的进展。

（2）相对论和量子力学已被普遍地承认是物理学的两块基石。然而却一直未能成功地使两者统一起来。泡利认为，广义相对论和量子力学之间的清晰联系迄今亦非近在眼前（1956 年）。甚至狭义相对论与量子力学之间，也依然存在着隔阂。这可以从以下的基本情况中看到。

在狭义相对论中，空间和时间形成一个四维的世界，以致在洛伦兹变换中，x，y，z，t 在它们各自的意义上是“同步的”。但在量子力学中，虽然 p 和 q 被赋予了非经典的作用（算符和对易关系），而时间 t 则仍保留它的经典意义,而且在经典动力学中是规范共轭偶的 t 和 E 之间不存在对易关系。这就是说，对于 t 没有满足下列关系的算符

$$\mathrm{Et}-\mathrm{tE}=-\frac{h}{2\pi i}$$

这是狭义相对论提示出来的。这个情况为泡利于 1933 年所指出。

（3）在前面的章节中我们已经概述了爱因斯坦对待量子力学和对待物理理论的哲学。我们必定已经注意到，他的哲学在 20 年代晚期以后与他年轻时代相比已有了改变，他在年轻时代，“创造”并“发展”了相对论和量子论。然而在一个层面上，他依然没有改变，这就是，他经常十分强调不可能存在一种“归纳的理论”;一切物理理论必定建立在某些原理，因而是演绎的基础上。就此而论，他并不是说基本原理就不可能由经验体验所提示，因为无疑，最初表述的热力学定律不可能都是不涉及经验体验的纯思维的结果。

事实上，量子力学是一种演绎的理论；它体现了很多爱因斯坦早期的精神，可是它却变得不被爱因斯坦所采纳了。

（4）在物理学发展史上，人们看到，实验发现和研究与理论探索是互补的，它们共同或交替促成了物理学的进步。有的时候，是某些实验发现导致了理论工作的巨大发展，而有的时候，则是某项理论工作导致了重要的实验结果。李政道（1987 年）表述了物理学家的两个定律：

第一定律：没有实验家，理论家趋于浮泛。

第二定律：没有理论家，实验家趋于摇摆。

在下面的表中，箭头表示“导致”。此表显示物理学中实验发展与理论发展之间的双向因果关系。

现象、实验		理论
地球上物体的运动	→	伽利略—牛顿动力学
开普勒行星运动定律 （17 世纪早期）	→	牛顿万有引力定律 （17 世纪中期）
海王星的发现（1846 年） 冥王星的发现（1930 年）	←	牛顿万有引力定律 （17 世纪中期）
光的衍射 （杨，19 世纪早期）	↔	光的波动理论 杨，菲涅耳（19 世纪早期）
蒸汽机的效率 （卡诺，19 世纪早期）	→	热力学第二定律 （19 世纪中期）
安培定律（1825 年） 法拉第定律（1831 年）	→	电磁场理论 （麦克斯韦，19 世纪中期）
电磁波的发现 （H．赫兹，1884 年）	←	电磁场理论 （麦克斯韦，19 世纪中期）
迈克耳孙-莫雷实验 （1881 年，1887 年）	←	麦克斯韦理论

孤立波的观察 （罗素，1842 年）	→	**孤立子理论 （19 世纪 90 年代；1960 年-）**
气体电导；电子的发现（19 世纪 90 年代；1896 年）	→	**原子结构理论**
黑体辐射的光谱分布（19 世纪 90 年代）	→	**普朗克量子论（1900 年）**
黑体辐射的光谱分布 （19 世纪 90 年代）	→	**爱因斯坦光子理论 （1905 年）**
黑体辐射的光谱分布 （19 世纪 90 年代）	→	**玻色统计（1924 年）**
迈克尔孙-莫雷实验； 菲佐实验（1851 年）； 特劳顿实验（1902 年）； 诺布尔实验（1903 年）； 布喇德雷实验（1729 年）；爱里实验（1871 年）	→	**洛伦兹变换（1902-1903） ↓ 爱因斯坦狭义相对论 （1905 年）**
光电效应 （密立根，1916 年）	←	**爱因斯坦光子理论 （1905 年）**
康普顿效应（1923 年）	↔	**爱因斯坦光子理论（1905 年）**
低温下固体的比热	→	**爱因斯坦理论（1907 年） ——现代固体理论的开端**

卢瑟福α粒子散射实验 （1911 **年**）	→	**玻尔的氢原子理论** （1913 **年**）
氢光谱的巴耳末系 （1885 **年**）		
弗兰克-赫兹实验 （1914 **年**）	←	**玻尔的氢原子理论** （1913 **年**）
施特恩-革拉赫实验 （1921 **年**）	←	**玻尔理论与索末菲的修正**
H **α线的精细结构** （1920 **年**）	←	**玻尔理论与索末菲的修正**
↓		
兰姆-雷瑟福移位 （1947 **年**）	↔	**量子电动力学** （1946 **年—**）
放射性元素 （20 **世纪** 10 **年代**）； **莫塞莱实验**（1913 **年**）； J.J.**汤姆孙极隧射线实验** （1914 **年**）	→	**同位素概念**
佩兰实验 （20 **世纪** 10 **年代**）； **确立物质由原子构成**	←	**爱因斯坦的布朗运动理论** （1905 **年**）
超导电性（1914 **年**）	→	B-C-S **理论**（1957 **年**）

惯性质量和引力质量 〔牛顿；厄缶（1890 年）； 迪克（1961 年）〕	→	爱因斯坦等效原理 （1907 年）
恒星光的偏转（日食观察）， 1919 年和随后的日食	←	爱因斯坦广义相对论和 引力理论（1915-1916 年）
水星近日点的进动 （19 世纪资料）	↔	爱因斯坦广义相对论和 引力理论（1915-1916 年）
由穆斯堡尔效应测 出的光谱线的 引力红移（1960 年）	←	爱因斯坦广义相对论和 引力理论（1915-1916 年）
哈勃红移（1930 年）	↔	爱因斯坦广义相对论和 理论（1915-1916 年）
尝试观察引力波 （J.韦伯，20 世 （纪 60 年代初）	←	爱因斯坦广义相对论和 引力理论（1915-1916 年）
		↓
自 20 世纪 60 年 代开始的观测	←	黑洞理论（1967 年）
微波激射（1954 年） 激光（1960 年）	←	爱因斯坦受激辐射理论 （1917 年）
恒星光谱（20 世纪初）	→	萨哈星体表面条件理论 （1922 — 1923 年）
元素周期表	→	泡利不相容原理 （1925 年）

原子光谱，反常塞曼效应（20 世纪 20 年代）	→	电子自旋（1925 年）
原子光谱	→	海森伯矩阵力学（1925 年）
电子衍射实验（戴维孙-革末，1927 年，G.P.汤姆孙，1928 年）	←	德布罗意物质波理论（1924 年）
		↓
		薛定谔波动力学（1926 年）
		↓
		概率诠释（1926 年）
		↓
		海森伯不确定性关系（1927 年）
		↓
		玻尔互补原理（1927 年）
		↓
原子光谱和结构，分子光谱和结构（20 世纪 20 年代-40 年代）	↔	作为完备体系的量子力学（1927 年）

拉曼效应（1928 年）	→	分子对称与结构（20 世纪 30 年代-）
β 衰变（20 世纪初）	→	泡利中微子（1931 年）
人工核衰变（1919，1932，1934 年） 中子的发现（查德威克，1932 年） 回旋加速器的发明	→	核结构理论（20 世纪 30 年代初）
正电子的发现（安德森，1932 年）	←	狄拉克相对论性电子方程（1928 年） 空穴理论（1930 年）
反质子的发现（塞格雷，1959 年）	←	狄拉克相对论性电子方程（1928 年） 空穴理论（1930 年）许
多未获成功的实验探索	←	狄拉克磁单极理论（1931 年）
质子-质子，质子-中子散射实验（20 世纪 30 年代）	→	核子-核子相互作用理论（20 世纪 30 年代，40 年代）
π 介子的发现（在宇宙线中，1946 年；在加速器实验中，1947 年）	←	汤川介子理论（1935 年）

		↓
核子—核子散射实验 （100 兆电子伏能量， 20 世纪 40 — 50 年代）	↔	介子理论和核子—核 子相互作用理论 （20 世纪 40 年代）
核相互作用实验 （20 世纪 30 年代）	→	恒星内部的能量 生成理论（1937 年）
核裂变（1938 年）	→	裂变理论（1939 年）
裂变链式实验 （1941 — 1944 年），核能 （自 40 年代晚期起）	←	裂变理论（1939 年）
原子束实验（拉 比等，40 年代早期）	→	氘结构理论 （中子—质子相互作用）
兰姆—雷瑟福移位 （1947 年） 回磁比 g（库什，1947 年）	↔	量子电动力学重正化理论 （40 年代，50 年代）
半导体（50 年代）	→	固体量子力学
核内的幻数	→	核壳层模型（50 年代）
核光谱学（50 年代后期）	↔	玻尔—莫特森 核模型（50 年代）

K介子衰变（50年代）	→	宇称不守恒（李政道和杨振宁，1956年）
β衰变；π-μ衰变；μ-e衰变（1956年）	←	宇称不守恒（李政道和杨振宁，1956年）
PC和T不守恒（克里斯坦森等，1964年）	←	时间反演守恒（1957年）
电子中微子和μ子中微子（莱德曼等，1962年）	←	弱相互作用理论（50年代开始）
R.G.钱伯斯实验（1960年）	←	Y.阿哈罗诺夫—D.玻姆理论（1959年）
宇宙3K辐射（1964—1965年）	→	宇宙学
高能实验中超子、介子的发现（50年代）	→	基本粒子理论〔SU（3）对称，60年代〕
重子Ω^-的观测（1964年）	←	基本粒子理论〔SU（3）对称，60年代〕 ↓

高能粒子物理学（60年代）	→	夸克理论（盖耳—曼，兹韦克，1964年）
J/ψ的发现（1974年）	←	夸克理论（盖耳-曼，兹韦克，1964年）
		规范场（杨—密耳斯，1954年） ↓
规范玻色子$W^{\pm}$，Z^0的发现（1983，1984年）	←	电磁相互作用和弱相互作用的统一〔SU（2）×U（1）理论，1967年〕

附　　录

1．电动力学方程在洛伦兹变换下的不变性

电磁学的基本定律是法拉第的电磁感应定律、库仑的静电学定律、具有麦克斯韦位移的安培定律，以及库仑的静磁学定律。它们是

$$\mathrm{curl}E=-\frac{1}{c}\frac{\partial H}{\partial t} \qquad (1)$$

$$\mathrm{div}E=4\pi\rho \qquad (2)$$

$$\mathrm{curl}H=\frac{4\pi}{c}\rho v+\frac{1}{c}\frac{\partial E}{\partial t} \qquad (3)$$

$$\mathrm{div}H=0 \qquad (4)$$

麦克斯韦引入位移电流 $\frac{1}{c}\frac{\partial E}{\partial t}$，保证了方程的连续性。

$$\mathrm{div}(\rho\nu) + \frac{\partial E}{\partial t} \qquad (5)$$

通过以下关系式引入矢量势 A 和标量势φ。

$$H = \mathrm{curl}A \qquad (6)$$

$$E = -\frac{1}{c}\frac{\partial A}{\partial t} - \mathrm{grad}\varphi \qquad (7)$$

方程组（1），（2），（3），（4）可以由以下方程组替代：

$$E + \frac{1}{c}\frac{\partial A}{\partial t} = -\mathrm{grad}\varphi \qquad (a)$$

$$\square\varphi = -4\pi\rho \qquad (b)$$

$$\square A = -\frac{4\pi}{c}\rho\nu \qquad (a)$$

$$\mathrm{curl}A = H (d)$$

这里 $\square = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$

方程（6）和（7）不能由给定的 E 和 H 来完全定义 A 和φ，为此洛伦兹引进方程

$$\mathrm{div}A + \frac{1}{c}\frac{\partial\varphi}{\partial t} = 0$$

方程组（a）…（e）用于电磁场。对于电荷（其密度为ρ）的运动，洛伦兹引进了（作用于电荷单位体积上的）力的表达式

$$f = \rho E + \frac{1}{c}\rho\,[\nu \times H] \qquad (f)$$

方程组 (a)、(b)…(f)在洛伦兹变换下的不变性

$$x'_\mu = \sum a_{\mu\nu} x_\nu,\quad x_1, x_2, x_3, x_4 = x, y, z, ict \qquad (8)$$

$$a_\mu = \gamma \begin{bmatrix} 1 & 0 & 0 & i\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta & 0 & 0 & 1 \end{bmatrix}$$

$$\beta = \frac{v}{c},\quad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

这可以从以下考虑中建立起来：

1）前文定义的□算符在变换式（8）中是一个不变算符。

2）$dx_1dx_2dx_3dx_4 = dx'_1dx'_2dx'_3dx'_4$=不变量

3）由于 $\rho\, dx1dx2dx3 = \rho'\, dx'_1dx'_2dx'_3$=不变量(电荷），则ρ必须按 dx_4 变换，即

$$\rho = \rho_0\frac{1}{\sqrt{1-\beta^2}}$$

4）由 3）可得 I(ρv_x，ρv_y，ρv_z，$ic\rho$）如 ds(dx_1，dx_2，dx_3，dx_4）变换，即 I 是一个四维矢量。

5）如果我们建立 A=A(A_x，A_y，A_z，$i\varphi$），那么，方程(b) 和(c) 合起来就可写成

$$\Box A = -\frac{4\pi}{c} I$$

但 I 是一个四维矢量。因此按照一个普适定理，A 也必定是一个四维矢量，四维势。

6）方程(e) 现在变为

$$\frac{\partial A_\lambda}{\partial x_\lambda} = 0，在 \lambda = 1，2，3，4 范围内求和 (e)')$$

7）如果我们通过以下方程定义张量 F

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} = -F_{\nu\mu} \qquad (9)$$

即

$$F_{\mu\nu} = \begin{bmatrix} 0 & H_z & -H_y & -iE_x \\ -H_z & 0 & H_x & -iE_y \\ H_y & -H_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{bmatrix}$$

那么方程(a)，(d) 合起来可用张量方程形式表达

$$\frac{\partial F_{\lambda\mu}}{\partial x_\nu} + \frac{\partial F}{\partial x_\lambda} + \frac{\partial F_{\mu\nu}}{\partial x_\mu}，\lambda \neq \mu \neq \nu \quad (a+b)$$

并且上面(b+C)' 可用矢量方程形式表达成

$$\frac{\partial F_{\lambda\mu}}{\partial x_\mu} = \frac{4\pi}{c} I_\lambda \qquad (b+c)$$

这里的 $\frac{\partial F_{\lambda\mu}}{\partial x_\mu}$ 隐含着一个 $\mu = 1，2，3，4$ 范围内的求和

8）方程(f) 可写成

$$cf_j = F_{j\nu} I_\nu，j=1，2，3\ (x,y,z)$$

这里隐含在 ν=1，2，3，4 范围内的和。如果我们定义

$$cf_4 = F_{4\nu} I_\nu = i\rho(Ev) \qquad (10)$$

则(f+10) 连同(b+c) 合起来能写成

$$f_\lambda = \frac{1}{c} F_{\lambda\nu} I_\nu，\lambda = 1，2，3，4，$$

$$= \frac{1}{4\pi} F_{\lambda\nu} \frac{\partial F_{\lambda\nu}}{\partial x_\mu}，对 \nu 和 \mu 求和 \qquad (f+10)')$$

由于右边是一个四维矢量，通过一个普适定理，左边 f_1，f_2，f_3，f_4 就必定是一个四维矢量。

这样，方程组(a)-(f) 全都采用张量方程的形式，并且它们在洛伦兹变换下的不变性得到了保证。

9）定义广义场张量 $T_{\lambda\mu}$

$$T_{\lambda\mu}=\frac{1}{4\pi}(F_{\lambda\mu}F_{\nu\mu}+\frac{1}{4}\delta_{\lambda\mu}F_{\alpha\beta}^2$$

由此

$$T_{xx}=\frac{1}{8\pi}(E_x^2-E_y^2-E_z^2+H_x^2-H_y^2-H_z^2)\quad \text{等等}$$

$$T_{xy}=\frac{1}{4\pi}(E_xE_y+H_xH_y)\quad \text{等等}$$

$$T_{x4}=-\frac{i}{4\pi}(E\times H)_x=-\frac{i}{c}S_x\quad \text{等等}$$

$$T_{44}=\frac{1}{8\pi}(E^2+H^2)\equiv u$$

于是，方程(f+10）可写成

$$f_\lambda=\frac{\partial T_{\lambda\mu}}{\partial x_\mu}\quad（对\mu求和）\qquad (f+10)'$$

这里的 S 是坡印廷矢量而 u 是场的能量密度。从方程(f），我们看到 f(f_1，f2，f3）是由场作用在(每单位体积的）电荷密度上的力，而从方程(10），$\frac{c}{i}f_4$是场在 ρ 上所作功的功率。如果 $\rho=0$，则场的能量守恒由下式表达

$$\frac{\partial T_{4\mu}}{\partial x_\mu}=，\text{或}\operatorname{div}S+\frac{\partial u}{\partial t}=0$$

2．相对论形式的经典动力学

在第四章第 3 节中我们已指出,牛顿运动方程在洛伦兹变换下不是不变的。不过，我们可以重新表述运动方程，以使其成为不变的。

我们注意到，在洛伦兹变换下，$ds^2=c^2dt^2-(dx^2+dy^2+dz^2)$ 是不变的。

在静止系中(粒子在其中处于静止状态），dx=dy=dz=0，且称为本征时间的时间 τ 由以下方程给出

$$ds^2=c^2d\tau^2$$

在粒子在其中速度为 v 的系统中，

$$ds^2=c^2dt^2(1-\beta)^2$$

$$\text{即}\quad dt=rd\tau，r=\frac{1}{\sqrt{1-\beta^2}}，\beta=\frac{v}{c}$$

$$\text{且}\quad \frac{dx_4}{ds}=\frac{icdt}{ds}=ir$$

质量密度 ρ 与静止质量密度有下述关系〔参见前述附录 1 中 3），电荷密度〕

$$\rho=r\rho_0$$

$$\text{且}\quad \rho d^3x=r\rho_0d^3x=\rho_0d^3x_0$$

这里的 d_3x 是三维体积元。

令 $$u_\mu=\frac{dx_\mu}{ds}$$

$f(f_1, f_2, f_3)$ = 作用在单位体积物质上的力，$\frac{c}{i}f_4$ = 对单位体积物质所作功的功率，

$$R_{\mu\nu}\equiv\rho_0c^2u_\mu u_\nu$$
$$T_{\mu\nu}\equiv R_{\mu\nu}+S_{\mu\nu}$$
$$\frac{\partial S_{\mu\nu}}{\partial x_\nu}=-f_\mu$$

现在

$$\frac{\partial R_{\mu\nu}}{\partial x_\nu}=c^2u_\mu\frac{\partial}{\partial x_\nu}(\rho_0u_\mu)+\rho_0c^2u_\mu\frac{\partial u_\mu}{\partial x_\nu}$$

且

$$\frac{\partial R}{\partial x_\nu}(\rho_0u_\nu)=\sum_{j=1}^{3}\frac{\partial}{\partial x_j}(\rho_0\frac{\partial x_j}{\partial x_4}\frac{\partial x_4}{\partial s})+\frac{\partial}{\partial x_4}(\rho_0\frac{\partial x_4}{\partial s})$$
$$=\frac{1}{c}\left[\mathrm{div}(\rho v)+\frac{\partial\rho}{\partial t}\right]$$

因此连续性方程导出 $\frac{\partial T_{\mu\nu}}{\partial x_\nu}=0$

方程

$$\frac{\partial T_{\mu\nu}}{\partial x_\mu}=0$$

现在导出

$$\rho_0c^2u_\nu\frac{\partial u_\mu}{\partial x_\nu}-f_\mu=0$$

或 $\rho_0c^2\frac{du_\mu}{ds}=f_\mu$，$\mu=1，2，3，4$

这是矢量形式的运动方程。

在该方程两边乘以 $d^3x_0=\frac{1}{\sqrt{1-\beta^2}}d^3x$，并且对物质（$\rho_0V_0=m_0$）的体积 $V_0=\frac{1}{\sqrt{1-\beta^2}}V$ 积分，我们就得到

$$m_0c^2\frac{du_\mu}{ds}=r\int f_\mu d^3x\equiv F_\mu$$

这里 F_μ 是一个四维力矢量。

这个运动方程在相对论形式中可以写成另一种形式。

让我们定义一个四维速度

$$V\left(\frac{dx_1}{dt},\ \frac{dx_2}{dt},\ \frac{dx_3}{dt},\ \frac{dx_4}{dt}\right)$$

从 $\frac{dt}{d\tau}=r$，我们有

$$V=V(rv_x, rv_y, rv_z, icr)$$

且 $\frac{du_s}{ds}=\frac{1}{c^2}\frac{d}{d\tau}\left(\frac{dx_\mu}{dt}\frac{dt}{d\tau}=\frac{1}{c^2}\frac{dV_\mu}{d\tau}\right)$

此时四维矢量形式的运动方程便是

$$\frac{d}{d\tau}(m_0V_\mu)=F_\mu$$

从这个关系式，我们得出一个意义深远的推论结果。三维力（即在三维空间内）由下式与 F_1，F_2，F_3 相关

$$f_k=F_k r,\ k=x,\ y,\ z$$

而通过下式使三维速度 v 与 V_x，V_y，V_z 相关

$$v_k=V_k r,\ k=x,\ y,\ z$$

因此（F・V）=0 给出

$$r^2(f\cdot v)+\frac{d}{d\tau}(m_0icr)icr=0$$

或 $\frac{d}{dt}\frac{m_0c}{\sqrt{1-\beta^2}}=(f\cdot v)$

该方程表示，量 $\frac{m_0c^2}{\sqrt{1-\beta^2}}$ 增加率等于力 f（在质量 m_0 上）所作功的功率。

这导致把 $\frac{m_0c^2}{\sqrt{1-\beta^2}}$ 解释作动能 T，出现一个附加常项，例如

$$T=m_0c^2\left[\frac{1}{\sqrt{1-\beta^2}}-1\right]\equiv(m-m_0)c^2$$

3.爱因斯坦理论中的静态球对称场

在爱因斯坦引力理论中，由能量-动量分布张量■给出四维空间的度规性质的“定律”是“场方程”

$$R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R=-kT_{\mu\nu} \qquad (1)$$

这些方程一般是 $g_{\mu\nu}$ 的非线性偏微分方程，并且是尚未解出的。不过，在静态球对称场的特定情况则有解。而即使是这种特定情况，也已导致了很重要的发展。

对于静态球对称场，度规能用以下形式表达

$$ds^2=g_{11}dr^2+g_{22}d\theta^2+g_{33}d\psi^2+g_{44}dt^2 \qquad (2)$$

这里

$$g_{11}=-e^{\lambda},\ g_{44}=e^{\mu}$$
$$g_{22}=-r^{2},\ g_{33}=-r^{2}\sin^{2}\theta \qquad (3)$$
$$g_{\mu\nu}=0,\text{在}\mu\neq\nu\text{时}$$

而λ（Y），μ（Y）是两个待定的Y的函数。

未消去的第二类克里斯托费耳3指标符号｛αβ;σ｝=｛βα;σ｝

$$\{\alpha\beta;\sigma\}=\frac{1}{2}g^{\sigma\rho}\left(\frac{\partial g_{\rho\beta}}{\partial x^{\beta}}+\frac{\partial g_{\rho\beta}}{\partial x^{\alpha}}-\frac{\partial g_{\alpha\beta}}{\partial x^{\beta}}\right)$$

是

$$\{11;1\}=\frac{1}{2}\frac{d\lambda}{dr},\ \{22;1\}=-re^{-\lambda},\ \{33;1\}=-r\sin^{2}\theta\, e^{-\lambda}$$
$$\{44;1\}=\frac{1}{2}e^{-\lambda}\frac{d\mu}{dr},\ \{12;2\}=\frac{1}{r},\ \{33;2\}=-\sin\theta\cos\theta$$
$$\{13;3\}=\frac{1}{r},\ \{23;3\}=\cot\theta,\ \{14;4\}=\frac{1}{2}\frac{d\mu}{dr} \qquad (4)$$

缩并的黎曼—克里斯托费耳张量$R_{\mu\nu}$是

$$R_{\mu\nu}=\frac{\partial}{\partial\chi^{\nu}}\{\mu\rho;\rho\}-\frac{\partial}{\partial\chi^{\rho}}\{\mu\nu;\rho\}+\{\mu\rho;\alpha\}\{\alpha\nu;\rho\}-\{\mu\nu;\alpha\}\{\alpha\rho;\rho\} \qquad (5)$$

张量T_{μ}^{ν}满足守恒定律

$$T_{\mu;\nu}^{\nu}=0 \qquad (6)$$

且$R=g^{\mu\nu}R_{\mu\nu}$，从$R_{\mu\nu}$，$g_{\mu\nu}$和场方程，在μ≠ν时可得出

$$T_{\mu}^{\nu}=0 \qquad (7)$$

守恒方程（6）又导出

$$\frac{dT_{1}^{1}}{dr}+\left(\frac{2}{r}+\frac{1}{2}\frac{d\mu}{dr}\right)T_{1}^{1}-\frac{1}{r}(T_{2}^{2}+T_{3}^{3})-\frac{1}{2}\frac{d\mu}{dr}T_{4}^{4}=0$$
$$\frac{dT_{2}^{2}}{d\theta}+\cot\theta(T_{2}^{2}-T_{3}^{3})=0 \qquad (8)$$
$$\frac{dT_{3}^{3}}{d\psi}=\frac{T_{4}^{4}}{dt}=0$$

从（5）和（4）式得出μ≠ν时

$$R_{\mu\nu}=0 \qquad (9)$$

且仅有的非消去组分是R_{11}，R_{22}，R_{33}，R_{44}。

爱因斯坦场方程（1）现在给出四个方程

$$R_{ii}-\frac{1}{2}g_{ii}R=-kT_{ii}=-kg_{ii}T_{i}^{i},\ i=1,2,3,4 \qquad (10)$$

这里没有因袭累加的方式，且R是标量，$R=g_{\mu\nu}R_{\mu\nu}$。利用（4）式，就有下式

$$\frac{1}{2}R=-e^{-\lambda}\left[\frac{1}{2}\frac{d^2\mu}{dr^2}-\frac{1}{4}\frac{d\lambda}{dr}\frac{d\mu}{dr}+\frac{1}{4}(\frac{d\mu}{dr})^2\right]+$$
$$\frac{1}{r}e^{-\lambda}(\frac{d\lambda}{dr}-\frac{d\mu}{dr})+\frac{1}{r^2}(1-e^{-\lambda}) \qquad (11)$$

于是方程（10）就成为下列各式

$$-\frac{1}{r}\frac{d\mu}{dr}+\frac{e^{\lambda}-1}{r^2}=+ke^{\lambda}T_1^1 \qquad (12a)$$

$$-e^{-\lambda}\left[\frac{1}{2}\frac{d^2\mu}{dr^2}+(\frac{1}{4}\frac{d\mu}{dr}+\frac{1}{2r})(\frac{d\mu}{dr}-\frac{d\lambda}{dr})\right]=kT_2^2 \qquad (12b)$$

$$-e^{-\lambda}\left[\frac{1}{2}\frac{d^2\mu}{dr^2}+(\frac{1}{4}\frac{d\mu}{dr}+\frac{1}{2r})(\frac{d\mu}{dr}-\frac{d\lambda}{dr})\right]=kT_3^3 \qquad (12c)$$

$$e^{\mu-\lambda}\left[\frac{1}{r}\frac{d\lambda}{dr}+\frac{e^{\lambda}-1}{r^2}\right]=ke^{\mu}T_4^4 \qquad (12d)$$

从球对你出发，按式（12b），（12c），就有

$$T_2^2=T_3^3$$

这就把四个场方程约简为三个不同的方程。

因为 $e^{\lambda}\neq 0$（即，λ不是—∞），三个方程就是

$$\frac{d\mu}{dr}-\frac{1}{r}(e^{\lambda}-1)=-ke^{\lambda}rT_1^1 \qquad (14)$$

$$\frac{d\lambda}{dr}+\frac{1}{r}(e^{\lambda}-)=ke^{\lambda}rT_4^4 \qquad (15)$$

$$\frac{d^2\mu}{dr^2}+(\frac{1}{2}\frac{d\mu}{dr}+\frac{1}{r})(\frac{d\mu}{dr}-\frac{d\lambda}{dr})=-2ke^{\lambda}T_2^2$$

把（14），（15）代入（16）并利用（8）式的前一个方程，就能得到恒等关系。因此（14）和（15）的解自动满足（16），而且不必再解（非线性的）方程（16）**。（当然，这一点也适用于**$T_1^1=T_2^2=T_3^3=T_4^4$**的史瓦西问题。）**

（1）史瓦西的情形：在 r=0 有一个质点，便有

$T_1^1=T_2^2=T_{34}^3=T_4^4$**，对**$r>0$**的情况**（17）

（14）和（15）的解是

μ（r）**＋λ**（r） = **常数**$\equiv q$（18）

而对 $e^{\lambda}\neq 0$ 的情况

$$e^{-\lambda}=1-\frac{2M}{r} \qquad (19)$$

积分常数 q 可从渐近条件→∞得以确定

$$\lim_{r\to\infty}e^{\lambda}=\lim_{r\to\infty}e^{\lambda}=1$$

于是 q=0，且

μ（r）=－λ（r）

因此从（3）得

$$g_{11}=-e^{\lambda}=-\frac{1}{1-\frac{2M}{r}}，g_{44}=e^{\mu}=1-\frac{2M}{r} \qquad (20)$$

积分常数 M 可以看做 r=0 处质点的质量，以长度单位 $C^2/2k$ 表示，k 是引力的宇宙常数，$k=6.67\times10^{-8}cm^3/g\cdot s^2$，c 为光速，$3\times10^{10}$cm／s。〔太阳的点质量 M=1.5km，地球的点质量 M=0.5cm〕

史瓦西情形的线元便是①

$$ds^2=-\frac{1}{1-\frac{2M}{r}}dr^2-r^2d\theta^2-r^2\sin^2\theta d\psi^2+(1-\frac{2M}{r})dt^2 \qquad (21)$$

上述度规已被应用于三个问题,它们在史瓦西求解之前就由爱因斯坦用近似方程求解了。这三个问题是：

（1）水星近日点的进动。计算所得的角Δ事实上是极小的，每一个世纪（100 年）为 43"；

（2）当恒星光从太阳边缘通过时有偏转角δ，计算值是 1.75"，一个非常小的效应；

（3）来自恒星和太阳的光谱线，相对于来自实验室的光谱线有引力红移。对于太阳来说，红移（以波长λ计）由下式给出

$$\triangle\lambda=2\times10^{-6}\lambda$$

这也是非常小的。

对Δ的观察值是 43.11"±0.45"。从日全食期间所拍摄的照片中得到δ极不容易，因为δ太小，并且与底板上所做的测量有许多关联。但就整个来说，观察值一般地与理论值 1.75″相符。

对于从恒星和太阳来的光谱线红移Δλ的测量是困难的,因为在发射原子时有许多扰动效应，诸如斯塔克效应和多普勒效应。但是庞德和里布卡于 1960 年通过利用穆斯堡尔效应的非常灵敏的方法已对引力红移作了准确的检验。爱因斯坦理论是与观察结果相一致的。

因此，在爱因斯坦理论的预言与牛顿理论有非常小差异的三个问题中，观察数据看来都支持了爱因斯坦的预言。从实践的意义上说，两个理论之间的小差别是不重要的，但两者存在差别，尽管差别极小，在原则上却很重要,因为它们意味着爱因斯坦理论决不仅仅是牛顿理论的一种翻版（一种单纯的几何化形式）。

半径 r＝2M 的球表面是一种 g_{11} 为奇异的球表面。它具有下面的性质：它把空间分为两个区域。对 r＜2M，有

$$g_{11}>0，g_{44}<0 \qquad (22)$$

以致空间坐标和时间 t 具有它们通常的相互变换的作用。

在球 r=2M 的外面，试考虑从一点 $r_0>2M$（$=r_s$）出发沿径向方向的光线。对光而言，零线便是

$$-\frac{1}{1-\frac{2M}{r}}dr^2+(1-\frac{2M}{r})dt^2=0$$

对它积分，就得出到达点 r 所花的时间

$$t=r-r_0+r_s\ln\frac{r-r_s}{r_0-r_s} \qquad (23)$$

随着 r_0 接近 r_s，t 增加，且光离开表面 $r=r_s$ 所花的时间变成无限。因此史瓦西球从外面是“不可见的”。J．A．惠勒（1967 年）提议将它命名为“黑洞”。

（对太阳而言，假如它是一个点质量，半径 r=2M=3km。太阳半径是 7×10^5km。史瓦西模型在太阳内部是不适用的。）

（2）推广到恒星

场方程是（14）和（15）。对之积分，便得

$$e^{-\lambda}=-\frac{k}{r}\int_0^r T_4^4(r)r^2dr \qquad (24)$$

$$\lambda+\mu=k\int_0^r(T_4^4-T_1^1)re^{\lambda}dr \qquad (25)$$

就T_1^1、T_4^4的有限范围而言，在极限$r\to\infty$下，解（24）与（17）相一致。对于质量不大于太阳质量的密度很大的恒星（例如密度约为 10^{13}g／cm³的中子星），其半径 r_B 由下式确定

$$1-\frac{k}{r_B}\int_0^r T_4^4(r)r^2dr=0$$

此半径 r_B 可以大于恒星的“半径”，球面 $r=r_B$ 具有“黑洞”的各种性质。

以上概述适用于静态球对称系统；非静态的情况，诸如“旋转”系统，便非常复杂。

从 20 世纪 20 年代到 40 年代的 20 年间,物理学的主要兴趣是在量子力学，以及它的应用于原子、分子和核物理学上。接着发生了第二次世界大战。在此期间，我们可以提及爱因斯坦、英费耳德和霍夫曼于 1938—1940 年在 N 体运动问题上的工作（见本书第四章第 5 节）和奥本海默及其学生赛培尔、斯奈德和伏耳科夫在 1938—1939 年所做的工作。

奥本海默和伏耳科夫关于静态球中子星稳定性的工作设立了质量大约 0.7 个太阳质量的极限。这项工作以后斯奈德一起参与，表明了当一个足够重恒星的热核能量耗尽时，该恒星将无限地收缩。这些在战时所做的工作未引起太多注意，但 20 年之后，年轻一代的物理学家复活了对广义相对论的兴趣，并展开了大规模的研究。

上面提及的奥本海默、伏耳科夫、斯奈德等人的工作已成了“黑洞”思想的出发点。如前所述，“黑洞”这个名字是由惠勒于 1967 年首倡的。

参考文献

W．Pauli，见第四章的参考文献
A．Pais，见第四章的参考文献
吴大猷，见第四章的参考文献

（微分几何学；爱因斯坦引力理论；史瓦西问题；三个检验）

4．在平直时空中的加速运动——时钟佯谬

令（X，T）是在一个惯性系 S 中的时空坐标，且（x，t）是在系 S′中的时空坐标，S′相对于 S 被加速，其加速方式是 S′的原点 x=0 按照方程

$$(1+\frac{aX}{c^2})^2=1+(\frac{aT}{c})^2$$

运动。式中 a 是一个常数。人们发现，变换

$$(1+\frac{aX}{c^2})^2=(1+\frac{aX}{c^2})^2+(\frac{aT}{c})^2$$

$$\frac{2at}{c}=\ln(1+\frac{aX}{c^2}+\frac{aT}{c})-\ln(1+\frac{aX}{c^2}-\frac{aT}{c})$$

在惯性系 S 和加速系 S′之间有如下性质：

（1）它代表加速系的一个单参数（a）族，

$$ds^2=-dX^2+c^2dT^2=-dx^2+c^2(1+\frac{ax}{c^2})^2dt^2$$

（3）该变换是定域洛伦兹变换，且

（4）对α的一个已知值，x 被限定在 x＞－c^2/a。在 x=-c^2/a 时，速度 v＝c。v 将不会超越此限度。这个要求是通过上述变换来满足的。由于这一理由，其他任意加速系可能必须在弯曲空间内才能加以描述。

（5）它允许对时钟佯谬问题进行非常精确的数学计算，并且表明一只被加速的时钟相对于在一个惯性系中的时钟会“丢失时间”。试考虑下面的一次火箭旅行，作为出发点的地球取作惯性系（S）。

令被加速部分 AB、CD、DE、FA 由前面提到的加速运动给出。令在匀速运动部分 BC 和 EF 的相对速度为 V_0 和－V_0。令Δτ$_2$ 为本征时间间隔，它由在 BC（和 EF）的火箭（S′）内的一只钟记录下来。计算表明，从 S 的立场看，他的钟将记下的是整个事件系列 ABCDEFA 的总时间，精确结果是

$$\frac{\triangle T_s'}{\triangle T_s}\approx\sqrt{1-(\frac{v_0}{c})^2}$$

而火箭 S′内的钟将记下整个旅行的本征时间

$$\triangle T_s'=\frac{4c}{a}\tanh^{-1}(\frac{v_0}{c})+2\triangle\tau_2$$

人们可以看出，总是有

$$\frac{\triangle T_s'}{\triangle T_s} \leq 1$$

在 $v_0/c=0$ 时，得到等号。

特别是，如果常数 α 非常大，我们就得到

$$\frac{\triangle T_s'}{\triangle T_s} \approx \frac{c}{v_0}\sqrt{1-(\frac{v_0}{c})^2}$$

这就是在不严格的论证基础上所得到的通常结果。为了表明得出返回火箭内的钟已经丢失了时间这一结果的通常论证（即让滑行部分 BC 和 EF 变得非常长）是不相关的，我们可以代之以完全废除滑行部分，以得到

$$\frac{\triangle T_s'}{\triangle T_s} = \frac{c}{v_0}\sqrt{1-(\frac{v_0}{c})^2}\ \tanh^{-1}(\frac{v_0}{c}) \approx 1-\frac{1}{12}(\frac{v_0}{c})^2$$

因为 $(\frac{v_0}{c})^2 << 1$

它总是小于 1 的，即在 S′内被加速的钟相对于 S 丢失了时间。

为了使爱因斯坦的论证完备，有必要指出，上述结果也还是从 S 的立场得到的。从 S′看来，参考系 S 经历了如下一系列的运动 A′—B′—C′—D′—E′—F′—A′，在 A′B′期间，S 在一个宇宙引力场 g（=-a）中朝负 x 的方向自由下落。在 B′，S 已达到一个相对于 S′的速度 $-v_0$。这个场消失，S 滑行一段时间 $\triangle\tau_2$。（在 S 内的一只钟所记下的本征时间）直到它到达 C′，这时一个宇宙引力场 g（=a）在正 x 方向起作用。S 在这个场中运动，从 C′到 D′再到 E′，到 E′后这个场又消失。S 以匀速 v_0 滑行到 F′，此时宇宙引力场 g（=-a）又起作用并把 S 送回到相对于 S′而言是静止的 A′。

用前面的变换方程计算（适当地采用分段进行，分成 C′—D′—E′和 A′—B′，F′—A′各段）表明，S 中所记录的时间 ΔTs 和 S′中所记录的时间 ΔTs′与前面从 S 的立场所给出的值精确地一致。

具体细节以及一般的（即弯曲时空内）加速运动之讨论，请参见吴大猷等发表于《国际理论物理学》杂志（第 5 卷，第 307 页，1972 年）上的文章。

5．气体动理论

5．1 玻耳兹曼方程

玻耳兹曼关于单粒子分布函数 f（r，v，t）的方程是

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\nabla\right) f(r, \mathbf{v}, t) = \left(\frac{\partial f}{\partial t}\right)。$$

$$\left(\frac{\partial f}{\partial t}\right)_c = \iint dw \mid \nu - \nu_2 \mid \boldsymbol{\sigma} \, [f(r, \nu', t) f(r, \nu_2', t)$$

$$- f(r, \nu, t) f(r, \nu_2, t)$$

这里的 f 被按下式归一化

$\int f(r, v, t) dv = n(r, t)$，即数密度，

而｜v-v_2｜是该分子在（r，t）处与另一个在 r 和 t 相碰撞的分子“2”之间相对速度的绝对值；dw 是碰撞截面；dw 是碰撞后分子被散射之立体角元。在碰撞积分 $\left(\frac{\partial f}{\partial t}\right)_c$ 中，$f(r, v, t) f(r, v_2, t)$ 项代表由于于与另一个分子碰撞 f（r，v，t）的减少率；f（r，v′，t）f（r，v_2′，t）项代表因“恢复碰撞”f（r，v，t）的增加率，“恢复碰撞”是指使碰撞粒子之一的速度进入 v 和 v+dr 之间范围的碰撞。

我们将在考察这个理论的基础之前简略地概括一下该理论的成功之处。

5．1．1H 定理

玻耳兹曼定义 H（t）函数为

$$H(t) = \iint f(r, \nu, t) \operatorname{In} f(r, \nu, t) \, drdv$$

这样从玻耳兹曼方程可得出

$$\frac{dH}{dt} < 0,$$

H 单调地减少到一个最低值，在此值假定 f（r，v，t）处于均匀空间密度和麦克斯韦速度分布的“平衡”值。

5．1．2 查普曼—恩斯考格解法

1911—1912 年，查普曼和恩斯考格独立地发展了解玻耳兹曼方程的一种方法。该法在于从玻氏方程获得方程组

$$\frac{\partial n}{\partial t} + \operatorname{div}(nu) = 0$$

$$\rho\left(\frac{\partial}{\partial t} + u \cdot \nabla\right) u_x + \sum \frac{\partial P_{xj}}{\partial r_j} = 0, \quad \rho = mn(r, t)$$

$$\frac{3}{2} nk \left(\frac{\partial n}{\partial t} + u \cdot \nabla\right) T + \operatorname{div} J + \sum_{ij} P_{ij} D_{ij} = 0$$

这里

$$2D_{xy}=\frac{\partial u_x}{\partial y}+\frac{\partial u_y}{\partial x}，D_{ij}=\mathrm{div}u$$

$$n（r，t）=\int f\mathrm{d}v$$

$$u（r，t）=\int vf\mathrm{d}v$$

$$p_{xy}（r，t）=\iint mU_xU_yf\mathrm{d}v$$

$$T（r，t）=\frac{m}{3nk}\int v^2f（r，v，t）\mathrm{d}v$$

$$J（r，t）=\int\frac{1}{2}mU^2U_xf（r，v，t）\mathrm{d}v$$

$$U（r，t）=v-u（r，t）$$

三个方程分别为连续性方程、运动方程和热通量 J 方程。

它们并不形成一封闭系统，因为胁强张量 P_{xy} 和通量 J 由更高的矩方程给出。

在不远离局部平衡的条件下，查普曼和恩斯考格方法将 f（r，v，t）展开为一个测量偏离局部平衡分布 f^0（v，t）的小参数 ξ 的幂级数

$$f^0（r，v，t）=f^0（v，t）〔1＋\xi\phi+\xi^2\psi\cdots〕$$

于是一个新思想被引入了。由于我们最感兴趣的是趋向平衡态的较晚阶段（其中流体动力学量 n，u，T 都随时间“缓慢”变化），我们将只在接近平衡态时 n，u，T 变化的时间尺度上寻求 f（r，v，t）的时间演变。数学上，这是由把 f（r，v，t）视为 n（t），u（t），T（t）的一个函数来表达的，

$$f〔r，v\,|\,n（t），u（t），T（t）〕$$

解玻耳兹曼方程的程序如下：上述 f（r，v，t）的推广被用于计算 n，u，T，且流体动力学方程是以连续近似方式解的。对于零次 ξ，流体动力学方程给出连续性方程、非粘滞流体的欧拉方程以及绝热过程的能量方程。这些方程都是时间可逆的〔由于它们并不包含玻耳兹曼方程的 $\left(\frac{\partial f}{\partial t}\right)_c$ 项。

对于一次 ξ，我们得到粘滞液体的纳维—斯托克斯方程与热流方程，两者都是时间不可逆的。从这些方程出发，我们可以导出粘滞系数 μ 和热导率 k，

$$\mu=-\frac{m}{2\sigma_0\zeta}\left(\frac{2kT}{m}\right)^{\frac{1}{2}}$$

$$k=\frac{15k}{4m}\mu$$

此处 σ_0 是量纲为面积的常数，ζ 是由两分子之间的相互作用力决定的一个（负的）常数。这个结果乃是玻耳兹曼理论的一个重要成就。

查普曼—恩斯考格理论的计算已经做到 ξ 的二次幂了（伯奈特，

1935）。

5．2 玻耳兹曼理论的基础

我们将回过头来考虑玻耳兹曼理论的基础。

起初 H 定理促使玻耳兹曼认为他的理论已构成了热力学第二定律的一个动力学论证（—H 与熵函数等同）。不过，这一观点立即遇到了洛喜密脱的重归佯谬和策梅洛的回归佯谬的批评，也就是说，如果 H 定理是动力学定律的一个推论，那么，把所有分子速度的方向反转就将导致 H 函数的增加；而按庞加莱的各态历经定理，一个系统的动力学态，在一个准周期之后，将返回到一个能如我们所愿地接近初态的态，故而 H 函数不可能在一切时间内都减少。人们已经知道，玻耳兹曼方程与一些动力学定律不同，它不是时间反演可逆的，因为碰撞积分（$\frac{\partial f}{\partial t}$）$_c$不是动力学定律的一个推论，而是一个概率本质的假定！玻耳兹曼因此而抛弃了他一度热衷过的这个观点。

〔事实上，从任何完全基于动力学的且因此而随时间可逆的“运动”方程出发，H 函数必须不随时间变化，而且在这种情况下，不会再发生洛喜密脱和策梅洛的质疑，H 不随时间改变。这一点本书作者已于 1975 年指出过了。〕

甚至在认识到碰撞积分（$\frac{\partial f}{\partial t}$）$_c$的概率本质之后，仍然有许多人力图去“理解”H 函数无涨落的单调递减（例如在埃伦菲斯特的文章中可见到的）。然而，涨落问题已经（由李保罗和我于 1973 年）澄清了；可以表明，从更基础的理论（下面 5.4 小节内要谈到的波哥留波夫理论）出发，再考虑到粒子—粒子关联效应，人们的确可以得到代表涨落并且导致 H 函数随涨落而递减的玻耳兹曼碰撞积分的附加项。

5．3 主方程和时间可逆性

我们已经看到，玻耳兹曼方程在时间上是不可逆的，且作为一个推论，H 函数具有时间不可逆性。我们曾强调，这个时间不可逆性来自碰撞积分中所谓的分子混沌拟设的概率本质。我们将通过泡利主方程（1928 年）以弄清概率考虑和时间可逆性之间的关系。

设 W_k（t）为时间 t 时一个系统处在态 k 的概率，W_i（t+Δt）为在时间 t＋Δt（Δt＞0）时处在态 i 的概率。令 $A_{ik}＝a_{ik}$Δt 是从 t 时态 k 到 t+Δt 时态 i 的跃迁概率。从上述“定义”，我们就有

$$\sum_k W_k(t)=1,\ \sum_i W_i(t+\nabla t)=1$$

$$\sum_k A_{jk}=1,\ \sum_k A_{ik}=1$$

且

$$0\le A_{ik}\le 1$$

我们假定

$$W_i(t+\Delta t)=\sum_k A_{ik}W_k(t)=\sum_k a_{ik}W_k(t)\triangle t$$

从这些，我们得到

$$W_i(t+\Delta t)-W_i(t)=\sum_k A_{ik}W_k(t)-(\sum_k A_{ik})A_i(t)$$

和

$$\frac{dW_i(t)}{dt}=\sum_k(a_{ik}W_k-a_{ki}W_i)$$

这就是所谓主方程，可以看出它具有以下性质：

①平衡的充分条件是

$$\sum_k(a_{ik}W_k-a_{ki}W_i)=0$$

因此“细致平衡”是不必要的。

②方程在时间上是不可逆的。

这最后的陈述可证明如下：

借助于不等式

$$x\int_1^{y/x}\ln z\quad dz=y(\ln y-\ln x)+x-y\geq 0$$

依据吉布斯方程，人们已得到

$$-\sum_i W_i(t+\triangle t)\ln W_i(t+\triangle t)\geq-\sum_i W_i(t)\ln W_i(t)$$

这类似于玻耳兹曼的 H 定理

$$-H(t+\triangle t)\geq-H(t)$$

（泡利，1952 年，是由斯蒂克尔贝格报告的）

首先，似乎从反向关系

$$W_k(t)=\sum_i A_{ki}^{-1}(t+\triangle t)\ln W_i(t+\triangle t)$$

出发，人们将会以一种类似方式得到关系式

$$-\sum_i W_i(t)\ln W_i(t)\geq-\sum_i W_i(t+\triangle t)\ln W_i(t+\triangle t)$$

如果这是正确的，那么将导出

$$-\sum_i W_i(t)\ln W_i(t)=-\sum W_i(t+\triangle t)\ln W_i(t+\triangle t)$$

这个结论的谬误起因于这样的事实，即A_{ki}^{-1}对反向跃迁并不满足

$$0\leq A_{ki}^{-1}\leq 1$$

这可从$A^{-1}A=1$，即$\sum_i A_{ki}^{-1}A_{ij}=\delta_{kj}$看出。由此可知，首先吉布斯不等式不能用于$A_{ki}^{-1}$，其次，$A_{ki}^{-1}$不具有跃迁概率的

PGN

意义（吴大猷和里维尔，1961年），因此，关系$W_i(t+\triangle t)=\sum_k A_{ik}W_k(t)$在时间上是不可逆的，所以它是主方程。

5．4 波哥留波夫理论（1946 年）

有关玻耳兹曼理论真正深刻的问题如下所述。在讨论有 N 个粒子的一

个系统时，我们可考虑从 6N 维空间的分布函数着手。那么人们怎样证明以单粒子在六维空间中的分布函数 f（r，v，t）来描述它是有道理的呢？

在解玻耳兹曼方程的查普曼—恩斯考格方法中,六维空间中的问题进一步由一个三维空间中的问题取代（在流体动力学方程中）。

这个降低维度的程序（从 6N 到六，又从六到三，乌伦贝克称作为“收缩”）并不仅仅是个数学手段；为其成功，还必须有个物理基础。事实上，它已隐含在查普曼—恩斯考格的方法中。然而，要清楚地理解只有到波哥留波夫的工作中。

5.4．1B—B—G—K—Y 方程链

波哥留波夫、玻恩和格林、柯克伍德、伊翁开始都以刘维方程求一个 N 粒子的分布函数 F_N

这里 H_N 是 N 粒子的哈密顿量

$$\frac{\partial F_N}{\partial t}+\sum_i\left(\frac{\partial F_N}{\partial r_i}\frac{\partial H_N}{\partial p_i}-\frac{\partial F_N}{\partial p_i}\frac{\partial H_N}{\partial r_i}\right)=0$$

令x为以r_0为单位表达的距离r，τ是以t_0为单位表达的时间t，v是以$u_0=\left(\frac{kT}{m}\right)^{\frac{1}{2}}$为单位表达的速度u，便有

$$x=\frac{r}{r_0}，\tau=\frac{t}{t_0}，v=\frac{u}{u_0}\text{和}t_0=\frac{r_0}{u}$$

令相互作用 V_{ij} 用以下形式表达

$$V_{ij}=\varphi_0 W_{ij}(|x_i-x_j|)$$

这里是一个恒定能量参数（范德瓦耳斯相互作用的量级），而 W_{ij} 是一个 $|x_i-x_j|$ 的无维度函数。令 F，F_{N-1}，…，F_1 为（无维度的）分布函数，由下列各式予以定义

$$\left(\frac{r_0^3}{V}\right)^N\int\cdots\int F_N(dxdv)^N=1$$

$$F_s\left(\frac{r_0^3}{V}\right)^{N-S}\int\cdots\int F_N(dxdv)^{N-S}，S=1,2,\cdots,N-1$$

$$F_1(x_1，V，\tau)=\left(\frac{r_0^3}{V}\right)^N\iint F_2(x_1,x_2,V_1,V_2,\tau)dx_2dV_2$$

这里 V 是气体之体积。

令

$$\Theta_S=\sum_{1<i<j}\nabla_i W_{ij}\cdot(\nabla_{v_i}-\nabla_{v_j})，\Theta_1=0$$

这里∇_{V_i}是v_i中的梯度算符。

可以看出刘维方程能采用以下形式

$$\frac{\partial F_N}{\partial\tau}+\left(\sum_i V_i\cdot\nabla_i-\frac{\varphi_0}{kT}\Theta_N\right)F_N=0$$

而从它出发，人们可获得方程的谱系

$$\frac{\partial F_S}{\partial\tau}+\left(\sum_i V_i\cdot\nabla_i-\frac{\varphi_0}{kT}\Theta_S\right)F_S=\frac{N-s}{N}\frac{\varphi_0}{kT}(nr_0^3)L_sF_s+1$$

s=1，2，…

这里$n=\frac{N}{V}$是每单位体积中之粒子数

①，且

$$Ls=\sum_{i=1}^{S}\iint dx_j dv_j(\nabla W_{ij}\cdot\frac{\partial}{\partial V_i}，j=s+1$$

对于单粒子函数F_1，因$\frac{N-1}{N}\approx 1$，则有

$$(\frac{\partial}{\partial\tau}+V_1\cdot\nabla_1)F_1(x_1，V_1，t)=\frac{\varphi_0}{kT}(nr_0^3)\iint dx_2 dV_2\frac{\partial}{\partial x_1}W_{12}\cdot\frac{\partial}{\partial V_1}F_2(x_1，V_1，x_2，V_2，\tau)$$

可以看出，F_1的方程包含F_2，而F_2又由含F_3的方程给出，等等。这个所谓的B－B－G－K－Y方程链在内容上当然与刘维方程是同一的,因为这些方程可从刘维方程中推导出来,但它们是一种更适于做近似处理的形式，而且也是理解玻耳兹曼方程的较适合方式。

5．4．2 时间尺度

波哥留波夫指出，气体中存在三种固有的时间尺度，即，两分子间一次碰撞的时间t_0，一个分子与另一个分子两次连续碰撞间的平均自由时间t_1和弛豫时间t_2，即

$$t_0=\frac{r_0}{u_0}，t_1=\frac{\lambda}{u_0}，t_2=\frac{L}{u_s}\approx\frac{L}{u_0}$$

这里r_0是分子间相互作用的范围（10^{-8}㎝量级），λ是平均自由路程$=\frac{1}{nr_0^2}$（10^{-4}cm量级），L是一个宏观长度，u_s是 声速，$u_s\approx u_0\approx 10^4$㎝/s，

$$t_0<<t_1<<t_2$$

分布函数F_s在所有这些时间尺度内都有变异。

若在上述F_1的方程中，我们取r_0为分子间相互作用的范围，则

$$\in\equiv nr_0^3=\frac{r_0}{\lambda}=\frac{t_0}{t_1}\approx 10^4$$

对于范德瓦耳斯力，$\frac{\varphi}{kT}$在常温下量级为1。因此对低密度n来说，右边的项是“小的”。这一项包含F_2，即它代表F_1由于分子与另一分子相互作用（碰撞）的变化率。这个变化率用时间$t_1=\lambda/t_0$。来表征，尽管F_2也包含较快的时间t_0。如果我们对t_0量级的时间间隔的快速变化不感兴趣，而只对时间尺度t_1感兴趣，那么，就有两种处理这种情况的方法。

一种方法是将F_2（x_1,x_2,v_1,v_2,t）表达为F_1的函数

$$F_2(x_1,x_2,v_1,v_2,t)\to F_2〔x_1,x_2,v_1,v_2\mid F_1(t)〕$$

另一种方法是利用所有的F_1，F_2，…明显的多倍时间尺度。

这两种方法都是属于波哥留波夫的。应用于目前问题的第一种方法已

由波哥留波夫给出，第二种方法由弗里曼和桑德里给出。事实已经表明，用这两种方法，上述 F_1 方程的某些近似确实导致玻耳兹曼方程！

因此，波哥留波夫的理论已为玻耳兹曼方程提供了一个基础和辩护，同样也为解玻耳兹曼方程的查普曼和恩斯考格方法提供基础和辩护 换言之，认识到有相当不同的特征时间尺度的存在，已为从 6N 维问题到六维，最后到三维问题的“收缩”程序提供了物理基础。

5.4.3 涨落

但波哥留波夫理论远超过一级近似，这一级近似就是玻耳兹曼方程，它导致 H 函数的单调递减。从 B－B－G－K－Y 方程链的 F_1 方程出发，如果人们不忽略双粒子函数 F_2 中分子—分子的关联效应，则已由李保罗所指出，在 t_1 的时间尺度（平均自由时间）上，有附加“随机”项加到 dH/dt 上，它们代表涨落，以致 H 并不单调递减，但随着这些涨落，H 减小到一极小值（李保罗与吴大猷，1973 年）。

参考文献

T. Y. Wu，Kinetic Equations of Gases and Plasmas，Addison－wesley Publ. Co. ，1966

（玻耳兹曼方程，B－B－G－K－Y 方程链，波哥留波夫理论）

W. Paull，article in Probleme der Moderne Physik，SommerfeldFestschrift，1928

（主方程）

T. Y. Wu＆D. Rivier，Hely. Phys. Acta 34，661，（1961）PaulS. Lee ＆T. Y. Wu，Intern. J. Theor. Phys. 7，267

（1973）

（H 定理中的涨落）

T. Y. Wu，Intern. J. Theor. Phys. 14，289（1976）（洛喜密脱和策梅洛佯谬）

6．非平衡热力学

6．1 不可逆过程

经典热力学处理由确定热力学平衡态的宏观概念（变量、函数）所描述的系统。因此严格说来，像温度、熵这样的概念仅对平衡态有意义。当一系统由平衡态 A 经一可逆过程 R（即经过一系列连续的平衡态）到 B 时，其熵的变化△S 为

$$S_B - S_A = \int_A^B \frac{dQ}{T}$$

但当这系统由态A经一不可逆过程IR到另一平衡态C时，上述积分$\int \frac{dQ}{T}$则

与熵的变化 S_C-S_A 完全无关；在这种情况下，S_C-S_A 只能用于计算导致从 A 到 C 的一个可逆过程 r （或一系列可逆过程）。

这样就产生了问题：“可逆的”或“不可逆的”过程的确切含义是什么？我们必须注意，这个问题并不简单。一个从态 A 到态 B 的可逆过程 R，是一个可被逆转的过程（在一连续的无限小过程系列的意义上），或是允许存在另一过程(或一系列过程)使系统由态 B 返回态 A 而不在环境中(在宇宙中）留下任何变化。一个不可逆过程 IR 则是这样的过程，对它来说，没有任何过程能使系统从 C 回到 A 而不在环境中（在宇宙中）留下任何变化。

强调“不在环境中留下任何变化”是十分重要的。这种发现一个把系统从 C 带回 A 而在环境中不留任何变化的过程的不可能性，是从第二定律（用克劳修斯、开尔文、或奥斯特瓦尔德的形式）的意义上理解的。因此，热力学中不可逆性的定义与第二定律本身密切相关。有许多涉及非平衡态的不可逆过程。例如，这里可提及：

（1）气体的自由膨胀

其中间状态为不能由变量 p，V，T 限定的非平衡态。

（2）粘滞流动（牛顿定律）

$$F_x=-\mu\frac{\partial u_x}{\partial z}\text{，}\mu\text{为粘滞系数}$$

流体的纳维—斯托克斯运动方程

$$\left(\frac{\partial}{\partial t}+u\cdot\nabla\right)u_x=\rho X-\frac{\partial\rho}{\partial x}+\mu\nabla^2u_x$$

其中 X 是单位质量上外力的 x 分量。

（3）热传导（傅里叶定律）

$$\text{热通量}J_x=-k\frac{\partial T}{\partial t}\text{，}k\text{为热传导率}$$

$$\rho C_v-\frac{\partial T}{\partial t}=k\nabla^2T$$

（4）扩散（斐克定律）

$$\text{质量流}=-D\nabla n\text{，}D\text{为扩散系数}$$

$$\frac{\partial n}{\partial t}=D\nabla^2n\text{，}n\text{为数密度}$$

（5）电流（欧姆定律）

$$\text{电流密度}j=-\sigma\nabla V\text{，}\sigma\text{为电导率}$$

$$\frac{\partial\rho}{\partial t}=\sigma\nabla^2V\text{，}\rho\text{为电荷密度，}V\text{为电势}$$

在粘滞流动中，定向运动的动能转换为分子热运动的能量。热传导涉及热能从高温到低温的传输；扩散涉及密度的均匀化；电流的流动产生出欧姆热。所有这些过程都是不可逆的，它们的半理论、半唯象的定律（方程）在时间 t 中也是不可逆的，即不能作为基本热力学定律的推论。

此外，还有更复杂的交叉效应，如：

（6）迪富尔的热浮流效应（1872 年）

$$\text{热通量} J = \beta \nabla n$$

而当温度梯度和浓度梯度俱在时

$$J = -k\nabla T + \beta \nabla n$$

（7）索里特热扩散效应（1893 年）

$$\text{质量扩散} N = \alpha \nabla T$$

而当温度梯度和浓度梯度俱在时

$$N = -D\nabla n + \alpha \nabla T$$

迪富尔效应和索里特效应是“倒易的”效应。上述系数α和β后来在昂萨格的非平衡热力学（1931 年）中被发现是彼此相关的。

6．2 非平衡热力学

在经典热力学主要讨论处于热力学平衡中的物质的性质时，第二定律本身也已涉及非平衡态的例子，因为它表明系统在（从非平衡）趋向平衡态时总是增加自己的熵。因而想把热力学扩展到非平衡情形是很自然的事。

为了做到这一点，我们首先要把某些严格说来起先限定于平衡态的概念，推广到非平衡情形。因此，我们将扩展熵的意义到非平衡情形中，然后再来评判这种推广。

6．2．1 熵产生率

设ρ为质量密度，u 为单位质量的能量，J 为热通量。这样能量方程为 $\rho \dfrac{\partial u}{\partial t} + \mathrm{div} J = 0$

第二定律为

$$Tds=du+pdv$$

其中 s 是单位质量的熵。对固体，膨胀可忽略，则

$$\rho T \frac{\partial s}{\partial t} + \mathrm{div} J = 0$$

上式可用下述形式表示

$$\rho \frac{\partial s}{\partial t} + \mathrm{div}\left(\frac{J}{T}\right) = -\frac{1}{T^2}\left(J \cdot \nabla T\right)$$

我们定义 $\dfrac{J}{T}$ 〔量纲为能量 /（面积·秒·摄氏度）〕为熵通量

$$S = \frac{J}{T} \text{（矢量）}$$

并运用傅里叶定律 $J = -k\nabla T$ 得出“熵平衡方程”

$$\rho \frac{\partial s}{\partial t} + \mathrm{div} S = \frac{k}{T^2}(\nabla T)^2 \equiv \Theta$$

其中

$$\Theta=\frac{1}{T}(J\cdot X)\geq 0,\ X=-\frac{\nabla T}{T}$$

Θ被称作熵产生率。这个必定总是正的Θ，也就是（不同形式的）第二定律。

6.2.2 宏观定律

对各向异性媒质中的热传导

$$J=-K\nabla T\text{，}k\text{为张量}$$

$$J_i=-\sum_j k_{ij}\frac{\partial T}{\partial x_j}\text{，}i\text{，}j=x\text{，}y\text{，}z$$

而$\Theta=-\frac{1}{T^2}\sum_i J_i\frac{\partial T}{\partial x_i}=\frac{1}{T^2}\sum_{i,j}k_{ij}\frac{\partial T}{\partial x_i}\frac{\partial T}{\partial x_j}$

为保证Θ≥0，二次形式$\sum k_{ij}x_i x_j$必须为非负定的。

对温度梯度和浓度梯度同时存在的热传导和热扩散的情形，如前所示，我们有

$$J=-k\nabla T+\beta\nabla n$$
$$N=\alpha\nabla T-D\nabla n$$

我们可以用下述形式重写这些方程

$$J(\alpha)=L_{\alpha\alpha}X_\alpha+L_{\alpha\beta}X_\beta$$
$$J(\beta)=L_{\beta\alpha}X_\alpha+L_{\beta\beta}X_\beta$$

并且

$$\Theta=\frac{1}{T}\sum_\alpha J(\alpha)X_\alpha=\frac{1}{T}\sum_{\alpha,\beta}L_{\alpha\beta}X_\alpha X_\beta$$

第二定律要求的Θ≥0 现在为

$$\sum_{\alpha,\beta}L_{\alpha\beta}X_\alpha X_\beta\text{是正定的，}$$

即

$$L_{\alpha\beta}>0,\ \begin{vmatrix}L_{\alpha\alpha} & L_{\alpha\beta}\\ L_{\beta\alpha} & L_{\beta\beta}\end{vmatrix}>0$$

因而仅当 $x_\alpha=x_\beta=0$ 时Θ=0，在这种情形中热流通量

$$J(\alpha)=J(\beta)=0$$

6．2．3 倒易关系 1931 年，L．昂萨格证明基于微观可逆性原理的考虑，一般有

$$L_{\alpha\beta}=L_{\beta\alpha}$$

这被称作昂萨格倒易关系。这个关系的证明并非全然明白，因为系数 $L_{\alpha\beta}$是源于宏观（来自傅里叶、斐克、迪富尔和索里特等定律）的。而卡西米尔给出的证明或许更易于理解。它立足于通过热力学平衡系统中涨落的统计力学来连接宏观性质和微观性质。以下我们将简略介绍这一理论。

6．3 昂萨格理论

6．3．1 涨落与回归

考虑一个热力学平衡中的绝热孤立系统。设 a_i 为偏离平衡态的涨落

的（小的）宏观变量。对小的涨落，（偏离平衡态的）熵变为

$$\triangle S=-\frac{1}{2}\sum_{i,j}S_{ij}a_i a_j<0$$

即二次形式$\sum_{ij}S_{ij}a_i a_j$为正定形式。

如果我们尝试性地以

$$X_i=\frac{\partial(\triangle S)}{\partial a_i}=-\sum_j S_{ij}a_j$$

$$J_i=\frac{da_i}{dt}=\dot{a}_i$$

来定义“力”X_i和“通量”J_i，则在偏离平衡态的涨落中的熵产生率Θ为

$$\Theta=\frac{d}{dt}\triangle S=-\sum_{i,j}S_{ij}\dot{a}_i a_j=\sum_i J_i X_i$$

它是负的，因为系统偏离平衡态时熵减小了。

在表达式 $J_i=\dot{a}_i$ 中，我们已过分简化了问题。为了说明问题的动理学—统计学本性，人们在一时间尺度τ上处理 a_i 的时间变化率，这一尺度较之确立一稳态通量 J 所需的时间 τ_1 为长，而较之涨落的衰退（或弛豫）时间 τ_2 为短，即

$$\tau_1<<\tau<<\tau_2$$

$\tau_1<<\tau_2$ 这一假设暗含着关于系统的力学的某些条件。然而，我们看到 $\tau_1<<\tau_2$ 这个条件在气体情形中确实得到满足。在气体情形中，τ_1 是玻耳兹曼方程中的平均自由时间,而 τ_1 是流体动力学方程的弛豫时间(见附录 5 中的 5．4．2 节）。

考虑到问题的统计本性，微商$\dot{a}_i$现在由比率

$$\left[\frac{a_i(t+\tau)-a_i(t)}{\tau}\right]$$

取代，其中〔…〕是对微观正则系统中对应于特定值 α_i（t）（i=1，2，…，n）的那部分取平均。

我们已在上面 6．2．2 节中看到，自然的、不可逆过程的唯象的、宏观的定律（傅里叶、斐克、迪富尔、索里特等）具有形式

$$J_i=\sum_k L_{ik}X_k$$

其特征为通量和广义力之间的线性关系。在昂萨格理论中，假设涨落的衰退（即系统的回归平衡态）类似于自然的不可逆地趋向平衡，并遵循同样的上述宏观定律。因此

$$[\alpha_i(t+\tau)-\alpha_i(t)]=\tau\sum_j L_{ij}X_j(t)$$

6．3．2 微观可逆性

微观可逆性原理以经典动力学中运动方程关于时间方向的反演

$$t \to -t$$

的不变性为基础。昂萨格理论中倒易关系的证明依赖于这一原理。不可逆宏观过程的系数间的这种倒易关系 $L_{ij}=L_{ji}$ 在基础动力学定律的可逆性中有其起源，这似乎并不明显。然而，这种时间反演不变性和不可逆过程之间的联系，N．玻尔在他关于金属的电子理论研究（就开尔文勋爵的热电现象关系而论）的博士论文（1913 年）中就已注意到了。

对宏观变量 a_i，微观可逆性（M．R．）能以下述形式表示：如果 a_i 是单个分子的速度的偶函数，那么微观可逆性可由 a_i（t）对时间反演的不变性表示，

$$a_i(t)=a_i(-t)$$

由此导出关系

$$〔a_i(t+\tau)〕=〔a_i(t-\tau)〕$$

其中〔…〕具有前已给出的含义。这个关系表明，平均而言，为所有 a_i（t）所限定的系统，其变化在未来 t+τ 与其过去 t－τ，是相同的。

上述关系也可以由在 t 时的 α_i（t）和在 t＋τ 时的 α_j（t+τ）之间的伴随关系的形式来表示：

$$\langle \alpha_i(t)\ \alpha_j(t+\tau)\rangle=\langle \alpha_i(t)\ \alpha_j(t=\tau)\rangle$$

其中〈…〉是对整个微观系综、即对 a_i（t）（i=1，2，…，n）的所有可能值取平均。这一关系也能表示为

$$\langle a_i(t)\ a_j(t+\tau)\rangle=\langle a_i(t+\tau)\ a_j(t)\rangle$$

也可以以另一种方式取平均：

$$\langle a_i(t)\ 〔a_j(t+\tau)〕\rangle=\langle a_j(t)\ 〔a_i(t+\tau)〕\rangle$$

对 τ=0，即还原为

$$\langle a_i(t)\ 〔a_j(t)〕\rangle=\langle a_j(t)\ 〔a_i(t)〕\rangle$$

把最后两个关系结合起来，我们得出

$$\langle a_i(t)\ \{〔a_j(t+\tau)-a_j(t)〕\}\rangle=\langle a_j(t)\ \{〔a_i(t+\tau)-a_i(t)〕\}\rangle$$ 6．3．3 倒易关系

在上面的 6．3．1 节中，作为昂萨格理论中的一个假设，我们有关系

$$J_j=\frac{1}{\tau}〔a_j(t+\gamma)-a_j(t)〕=\sum_m L_{jm}X_m(t)$$

由这一关系，我们得出系综平均值

$$\begin{aligned}\langle a_i(t)\ J_j(t)\rangle&=\frac{1}{\tau}\langle a_i(t)\left\{〔a_j(t+\tau)-a_j(t)〕\right\}\rangle\\&=\sum_m L_{jm}\langle a_i(t)\ X_m(t)\rangle\\&=-k\sum_m L_{jm}\,\delta_{im}\end{aligned}$$

①k 为玻耳兹曼常数=－kL_{ij} 但是类似地，我们可得出

$$\begin{aligned}\langle a_j(t)\ J_i(t)\rangle&=\frac{1}{\tau}\langle a_j(t)\left\{〔a_j(t+\tau)-a_i(t)〕\right\}\rangle\\&=-kL_{ij}\end{aligned}$$

由前一小节的最后一个方程，可得

$L_{ji}=L_{ij}$ 这就是昂萨格的倒易关系。

在上述证明中，我们假设了 a_i（t）是分子速度的偶函数。然而可以证明，如果 a_i（t）是奇函数，或者其中有些是偶函数、有些是奇函数，也有同样的结果。因此倒易关系是普遍的。

仅有的规定是，如果磁场 B 出现并且相关，则

$$-L_{ij}(B)=L_{ji}(-B)$$

这是由于 B 与电流方向相关，所以随时间反演改变符号。

6．3．4 热电现象的不可逆热力学

昂萨格理论可应用于许多非平衡现象,但其中最有趣的应用是用于解决热电效应理论中长期存在的“谜题”。这涉及珀耳帖系数Π和塞贝克绝对温差电势率ε之间的汤姆孙关系

$$\Pi-T\varepsilon=0$$

让两种不同金属 A 和 B 的两极导线构成一个闭合电路,并使两处接合点处于不同温度。当珀耳帖效应和汤姆孙效应为可逆过程时，电路中的电流必然为（温度梯度下的）热传导和欧姆热产生等不可逆过程所伴随。汤姆孙在他的理论中，把系统看作处于“卡诺循环”中，并通过忽略所有不可逆部分以运用第二定律于系统，由此得出关系式 $\Pi-T\varepsilon=0$。

这一理论显然是错误的，因为没有理由为一时之需而“忘记”确实存在的不可逆部分。谜题在于汤姆孙理论的结果被发现与实验测量完全相符。许多证明汤姆孙理论的尝试都失败了。1928 年，索末菲运用新发展起来的量子统计法，计算了金属电子理论的系数Π，ε，证实了关系 $\Pi-T\varepsilon=0$。但关于这一关系的最终解决却是在昂萨格 1931 年的理论之后，此时发现这一关系 $\Pi-T\varepsilon=0$ 是倒易关系 $L_{ij}=L_{ji}$ 应用于关于能量通量 J 和电场 E 的下述唯象关系时的一个推论。

对一种分子量为 M、密度为ρ、单位质量的能量为 u、单位质量的熵为 S、能通量为 J、电流密度为 I、电势为Φ、化学势为μ的金属，其能量方程为

$$\rho\frac{\partial u}{\partial t}+\mathrm{div}(J+\Phi I)=0$$

而（忽略了体积变化的）第二定律为

$$MTdS=Mdu(\mu-F\Phi)dn$$

其中 F 是法拉第常数 Ne，而 n 是金属中每个原子所拥有的电子数。唯象关系为

$$J=-k\nabla T-\left(\Pi+\frac{\mu_0}{e}\right)I$$

（傅里叶）（珀耳帖）

$$E=\frac{1}{\sigma}I-\varepsilon\nabla T-\nabla\left(\frac{\mu_0}{e}\right)$$

（欧姆）（塞贝克）（接触电势）

其中 $E=-\nabla\Phi$ 是电场，$\mu_0=\mu/N$（N是阿伏伽德罗数）是每个电子的化学势。熵平衡方程

$$T\rho\frac{\partial S}{\partial t}=-\mathrm{div}J+（E\cdot I）-\frac{\mu_0}{e}\mathrm{div}I$$

可以写成

$$T\frac{\partial D}{\partial t}+\mathrm{div}S=-\frac{1}{T}（J\cdot\frac{\nabla T}{T}）+\frac{1}{T}I\cdot（E+T\nabla（-\frac{\mu_0}{eT}））$$

其中

$$S=\frac{1}{T}（J+\frac{\mu_0}{e}I）$$

上述唯象关系可以改写为下列形式：

$$J=L_{11}X_1+L_{12}X_2$$
$$I=L_{21}X_1+L_{22}X_2$$

其中

$$X_1=-\frac{1}{T}\nabla T$$
$$X_2=E+T\nabla(\frac{\mu_0}{eT})=-\nabla\Phi+T\nabla（\frac{\mu_0}{eT}）$$
$$L_{11}=kT+\sigma（\Pi+\frac{\mu_0}{e}）（T\varepsilon+\frac{\mu_0}{e}）$$
$$L_{21}=-\sigma（T\varepsilon+\frac{\mu_0}{e}）$$
$$L_{12}=-\sigma（\Pi+\frac{\mu_0}{e}），L_{22}=0$$

前述熵产生方程现在取形式

$$\rho\frac{\partial s}{\partial t}+\mathrm{div}S=\frac{1}{T^2}\left\{k(\nabla T)^2+(\Pi-T\varepsilon)(I\cdot\nabla T)+\frac{1}{\sigma}TI^2\right\}$$

所有这些都严格地是唯象关系和第二定律的结果。

现在，由昂萨格倒易关系 $L_{12}=L_{21}$ 导出

$$\Pi-T\varepsilon=0$$

这就是汤姆孙的“第二关系”。借助于 $\Pi-T\varepsilon=0$，则熵产生率为

$$\Theta=\frac{1}{T}\sum_i L_{ij}X_iX_j$$
$$=\frac{1}{T^2}\left\{k(\nabla T)^2+\frac{1}{\sigma}TI^2\right\}>0$$

珀耳帖系数 Π 在 Θ 中不出现，现在可从多个方面来理解：①珀耳帖效应本身为一可逆现象，只有与之相伴的焦耳热效应才对熵的产生有贡献；②在汤姆孙理论（包括珀耳帖、塞贝克和汤姆孙等效应）中，所有起因于热传导和焦耳热的不可逆部分都被忽略，而只考虑可逆部分，由此得出关系 $\Pi-T\varepsilon=0$ 并导致（$\Pi-T\varepsilon$）（$I\cdot\nabla T$）这一项在 Θ 的表达式中消失；③最后，昂萨格理论证明了普遍的倒易关系，这一关系在珀耳帖—塞贝克效

应情形中精确导出关系Π－Tε=0。

汤姆孙关系Π－Tε=0 尽管起初是从一个错误的理论得出的，但它毕竟是正确的，这是物理学史上最独特的事例之一。另一例子是索末菲关于氢原子能态的电子质量的相对论性变化理论。尽管这一理论后来被发现并不正确，但能量表达形式却与 14 年后由狄拉克相对论性波动方程得出的表达形式上相同。

6．3．5 对非平衡热力学理论的评论

在这一附录的第 6．2 节的开始曾提到，在严格的平衡概念如熵、温度等被推广和扩展以发展出非平衡热力学理论后,必须理解制约这种推广和扩展的合理性的条件。我们已经简要叙述了昂萨格理论，并且提到了理论中的若干假设（见上述 6．3．1 和 6．3．2 节）。这里，我们要对非平衡热力学理论再作一些评论。

（1）我们记得在求解玻耳兹曼方程的查普曼—恩斯考格方法（见附录 5 的 5．1．2 节）中，分布函数 f（r，v，t）按参量ξ的幂次展开，这里ξ是对偏离“局域平衡”状态的量度。局域平衡指的是这样一种状态，其中系统的“小的”体积可由诸如温度这样的宏观“平衡态概念”来“局域”地描述。这种展开法仅当系统对平衡态的偏离很小时才可运用；显然当系统处于一很大的湍动状态，如冲击波中时，它是没有意义的。

在非平衡热力学理论（例如昂萨格理论）中，我们也同样要作从平衡态的偏离很小的限制（见本附录的 6．3．1 节），意即系统的“很小”体积可被看作处于“熵”有意义的平衡态中。这一观点不仅仅是“合理的”。普里戈金（1946—1949）指出，当运用查普曼—恩斯考格方法（在ξ一次幂上），以动理学理论来计算非平衡气体的“熵”（通过玻耳兹曼的 H 函数和 S=-H+常数）时，“熵”显现出不再明显依赖于宏观变量的梯度，因此熵概念对非平衡气体的推广是“无可非议的”。但当计算拓展到ξ^2级（伯奈特近似）时，熵开始包含梯度。这被认为意味着“熵”的推广或扩展是成问题的。因此在非平衡热力学理论中，人们可以“证明”熵概念推广到从平衡态“微小”偏离状态的正当性，与在动理论中的情形含义相同。

（2）在 6.3.1 中已提及的“时间尺度”问题，即理论适用的时间τ由下式限定

$$\tau_1 \ll \tau \ll \tau_2$$

对非平衡系统有足够的时间，使唯象关系$J_i = \sum_j a_{ij} x_j$具有意义。对短时短时间中处于剧烈湍动的系统，这个理论显然是不适用的。

参考文献

L.Onsager，Phys.Rev.37，405；38，2265（1931）H. B. G. Casimir，Rev. Mod. Phys. 17，343（1945）

（关于非平衡热力学和倒易关系）

A.Sommerfeld,ThermodynamicsandStatisticalMechanics,Acad.Press,NewYork，1964，及第七章参考文献中吴大猷的著作

（关于开尔文勋爵的热电理论）

S．R．deGrootandP．Masur，Non-EquilibriumTher-modynamics，North-Holland，Amsterdam，（1962）

7．爱因斯坦光子理论——波粒二象性的起源

普朗克的独创思想是原子（指构成与辐射平衡的空腔壁的物质原子）以离散的数量发射与吸收辐射。爱因斯坦（1905 年）进一步提出，辐射本身兼有连续与离散的性质。他的这些思考是建立在统计理论基础上的。

就一般的统计考虑而论，人们知道，具有许多自由度的一个系统之能量的均方偏离或涨落由下式给出

$$\overline{(\triangle\varepsilon)^2} \equiv \overline{(\varepsilon-\bar{\varepsilon})^2} = \frac{d\bar{\varepsilon}}{d\beta}, \quad \beta = 1/kT \qquad (1)$$

此时如果辐射是一个其所有频率、振幅、方向的谐波系统，那么经典理论给出的涨落和平均能有如下关系

$$\overline{(\triangle\bar{\varepsilon})^2} = \bar{\varepsilon}^2 \qquad (2)$$

从（1）式和（2）式，就能得到瑞利—金斯公式，该公式对于大的λT 值来说与观察到的谱线分布符合得极好。在解释λT 值较小时的维恩定律中，上两式对$\bar{\varepsilon}$不起作用。

爱因斯坦提出,辐射的能量分布也是离散的,并且频率为 v 的辐射‘量子”由下式给出

$$\varepsilon_0 = hv$$

如果 ndv 是每单位体积的量子数，其频率范围处于 v 和 v+dv 之间，且■是平均数密度，则

$$-\bar{\varepsilon} = \bar{n}\varepsilon_0$$

且$\overline{(\triangle\varepsilon)^2} = \varepsilon_0^2\overline{(\triangle\varepsilon)^2}$。此时，根据统计理论，对于很大数量的独立的“粒子”则有

$$\overline{(\triangle n)^2} = \bar{n} \qquad (3)$$

如果连续和离散的性质的起因是独立的，那么它们对涨落的影响可以相加，于是，当（2）和（3）结合起来时，我们就能从（1）式中得出方程

$$-\varepsilon_0\frac{d\bar{n}}{d\beta} = (\bar{n}^2 + \bar{n})\varepsilon_0^2$$

它的解是

$$\bar{n} = \frac{1}{e^{\alpha+\beta\varepsilon_0} - 1}$$

常数a必须调整到当T→∞时等于0的标准极限情况。从$\bar{\varepsilon} = \bar{n}\varepsilon_0 = \bar{n}h\nu$，

就得出普朗克黑体辐射公式。

作为这种量子理论的一个例子,爱因斯坦提出了已为我们所熟悉的光电效应理论。

8．海森伯新理论（1925 年）的指导思想

在经典理论中，作为时间函数的坐标 q 能表达为一个傅里叶级数

$$q（t）=\sum_{n=0}^{\infty}（a_n\cos 2\pi n\nu t+b_n\sin 2\pi n\nu t）=\sum_{-\infty}^{\infty}q_n e^{2\pi in\nu t}$$

此处$q_n=\frac{1}{2}（a_n-ib_n）$，$q_{-n}=q_n^*$，组分$n\nu$的强度和偏振由复振幅$q_n$给出。

傅里叶级数的一个重要性质是，所有运算，包括加、乘和微分，诸如$\dot{q}$，$q\dot{q}$，q^2，等等，都不引进不同于系列ν，2ν，3ν，……的新频率。

原子的光谱频率并不形成一个谐波系列 ν，2ν，3ν，……但满足一个通称为里兹（Ritz）组合原则的经验关系，即频率由下式给出

$$\nu_{mn}=T_m-T_n$$

此处 T_m，T_n 是凭经验确定的“项值”（确定趋向于一个附加常数）。因此，海森伯不得不用对应于 ν_{mn} 的二维体系代替频率级数 ν，2ν，3

ν，……即$q_n e^{2\pi in\nu t}$由$q_{mn}=q_{mn}^*$的$q_{mn}e^{2\pi i\nu_{mn}t}$代替。并且平方值$|q_{mn}|^2$和复 q_{mn} 的相代表谱线 ν_{mn} 的强度和偏振。为了像 q_2 和 qq 这样的乘法不引进已有的 ν_{mn} 中间所没有的新频率，M．玻恩注意到，海森伯思想的适当数学形

式体系可以用矩阵代数形式表达，因此就有

$$(qq)_{mn}=\sum_k q_{mk}e^{2\pi i\nu_{mn}t}q_{kn}e^{2\pi i\nu_{kn}t}=\sum_k q_{mk}q_{kn}e^{2\pi i(\nu_{mk}+\nu_{kn})t}$$

$$=(\sum_k q_{mk}q_{kn})e^{2\pi i\nu_{mn}t}$$

$$(\dot{q})_{mn}=2\pi i\nu_{mn}q_{mn}e^{2\pi i\nu_{mn}t}$$，等等

由于矩阵乘法是不对易的，即通常

$$AB-BA\neq 0$$

问题就是如何处理 pq－qp，例如当这里的 p，q 是经典意义上的正则共轭对时。

海森伯受了玻尔对应原理的提示和导引。在经典动力学中，对于周期系统，人们定义正则变量对 W，J，其作用变量是

$$J=\oint pdq=\int_0^{\frac{1}{\nu}}p\dot{q}dt$$

q 和 p 都使用傅里叶级数，我们就有

$$J=\int_0^{\frac{1}{v}}\sum_n p_n e^{2\pi i n v t}\sum_m 2\pi i m v q_m e^{2\pi i m v t}dt$$

$$=2\pi i v\sum_{n,k}\int_0^{\frac{1}{v}} p_n q_{n-k}(k-n)e^{2\pi i k v t}dt=-2\pi i\sum_{\tau=-\infty}^{\infty}\tau p_\tau q_{-\tau}$$

且 $1=\frac{\partial J}{\partial J}=-2\pi i\sum_\tau\frac{\partial}{\partial J}(p_\tau q_{-\tau})$。

在玻尔—索末菲的理论中，J 被量子化，

J=nh

且△J＝（△n）h=τh，τ≡△n

按照对应原理，经典动力学中任一函数F的导数$\frac{\partial F}{\partial J}$在量子理论中可用$\frac{\triangle F}{\triangle J}=\frac{\triangle F}{\tau h}$替代，$\frac{\partial}{\partial J}$（$q_\tau q_{-\tau}$）用$\frac{A(q_\tau q_{-\tau})}{\tau h}$替代，它们依次在新理论中又可用 p、q 在其中是矩阵的表达式替代，即

$$\frac{(\triangle p_\tau q_{-\tau})}{\tau h}\rightarrow\frac{\triangle}{\tau h}(p_{n,n-\tau}q_{n-\tau,n})$$
$$=\frac{1}{\tau h}(p_{n,n-\tau}q_{n-\tau,n}-q_{n,n+\tau}p_{n+\tau,n})$$

将此代入上面所述$1=\frac{\partial J}{\partial J}$的代表式，海森伯就得到关系式

$$1=\frac{2\pi i}{h}\sum_\tau(p_{n,n-\tau}q_{n-\tau,n}-q_{n,n+\tau}p_{n+\tau,n})$$

这里求和是对τ从-∞到∞范围内。这可以用矩阵方程的形式重新写成

$$(pq-pq)_{nn}=\frac{h}{2\pi i}$$

海森伯随即作了这样的假设，即两矩阵 p 和 q 满足关系

$$pq-pq=\frac{h}{2\pi i}E\text{，}E\text{为单位矩阵}$$

必须强调指出的是，上述对易关系的推导是不错的。事实上，它具有基本公设的性质，因它不可能从经典理论中“推导”出来。①一旦有公设

$$pq-pq=\frac{h}{2\pi i}E$$

许多结果便可随之得到。因此从这一点出发，对于涉及 p，q 的加或乘的任何矩阵函数 F（q，p），人们都可得出以下关系

$$Fq-qF=\frac{h}{2\pi i}\frac{\partial F}{\partial p}\text{，}Fp-pF=\frac{h}{2\pi i}\frac{\partial F}{\partial q}$$

如果哈密顿量 H=H（q，p）矩阵是对角的，并且人们规定了玻尔的频率条件

$$h\nu_{mn}=H_{mn}-H_{mn}$$

那么，人们可得出

$$(Hq-qH)_{mn} = (H_{mn}-H_{mn})(q)_{mn} = h\nu_{mn}(q)_{mn}$$
$$\dot{q}_{mn} = 2\pi i\nu_{mn}q_{mn}e^{2\pi i\nu_{mn}t} = 2\pi i\nu_{mn}(q)_{mn}$$
$$\dot{q} = \frac{\partial H}{\partial p} = \frac{2\pi i}{h}（Hq-qH）$$

类似地 $p = -\frac{\partial H}{\partial q} = \frac{2\pi i}{h}（Hq-qH）$

这些都是海森伯理论中的“运动方程”；它们除了是矩阵方程之外，与经典动力学中的正则方程具有同样的形式。

9．薛定谔波动力学的出发点

这个出发点是德布罗意与粒子相关联的“物质波”的建议。德布罗意的建议受了基于狭义相对论的考虑的引导。这一点对我们来说也许太熟悉了，因而不必再在此多费笔墨。薛定谔最早（1926 年）的考虑见第九章第 1 节。

粒子动力学与波动光学之间的一种形式关系可以追溯到哈密顿 19 世纪 30 年代初的工作。对于波动力学，有一个费马最小时间原理，它说的是，在任何介质中，光从 A 点行进到 B 点采取的方式是，它所走的实际路径花的时间为极小值，即

$$\delta\int_A^B \frac{ds}{u} = \delta\int_A^B \frac{ds}{\nu\lambda} = 0$$

此处 ds 是弧元，而 u 是相速度（费马，1601—1655）。在光的微粒论中，有一个莫培督最小作用原理（1732 年），按此原理，光微粒的实际轨道，其“作用”的路径是一极小值。在动力学的变分原理的更普遍的形式中，最小作用原理采取以下形式（对一个在势能为 V 的场中质量为 m 的粒子）

$$\delta\int_A^B 2T dt = \delta\int_A^B \sqrt{2m(E-V)}ds = \delta\int_A^B p ds = 0$$

此处 T 和 E 是粒子的动能和总能量。因此，当动量 p 正比于 1/λ 时粒子理论和波动理论就变成等价的了。德布罗意提出了如下关系：

$$\lambda = \frac{h}{p}$$

在薛定谔理论中的第二步是得到一个“波动”方程。指导思想又是从经典动力学中得到的一个提示——哈密顿—雅可比方程的形式

$$H（q，p=\frac{\partial S}{\partial q}，t） = \frac{\partial S}{\partial t}$$

H 是哈密顿函数，即用广义坐标 q_n 和共轭动量 p_n 表达的一个系统的总能量。上述哈密顿—雅可比方程是用 $\frac{\partial S}{\partial p_i}$ 代替 p_i 得到的，并因此是一个二次函数 S（$q_1 \cdots q_n$，t）的偏微分方程，这个二次函数被称做主函数。在 H

明显地不依赖于时间的情况下，H－J 方程对 S_0（q）=S（q，t）－Et 来说就简化为下面的方程

$$H\left(q, p=\frac{\partial S}{\partial q}, t\right)=E$$

此处 E 是系统的恒定能量。

在波动理论中，我们有波动方程

$$\nabla^2\psi=\frac{1}{u^2}\frac{\partial^2\psi}{\partial t^2}$$

此处ψ=ψ（r，t），相速度 u 是空间坐标 r 的函数。这时就德布罗意波来说，我们有德布罗意关系

$$E=h\nu,\ p=\frac{h}{\lambda}$$

由此

$$u=\lambda\nu=\frac{E}{p}$$

而在非相对论性理论中

$$u=\frac{E}{\sqrt{2m(E-V)}}$$

如果我们假定

$\psi(r,t)=\varphi(r)e^{-2\pi iEt/h}$

则上述波动方程引出

$$\left(-\frac{h^2}{8\pi^2m}\nabla^2+V\right)\varphi=E\varphi$$

且$\left(-\frac{h^2}{8\pi^2m}\nabla^2+V\right)\Psi=-\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}$

可以看出，这些方程与哈密顿—雅可比方程具有相同的形式，若在经典哈密顿函数中，动量 p 由算符取代，

$$p\rightarrow\frac{h}{2\pi i}\frac{\partial}{\partial q}$$

这就导出函数Ψ〔r，t）（或φ（r，t）〕的一个线性偏微分方程。上面Ψ的方程叫做薛定谔方程，而对φ的方程则是与时间无关的薛定谔方程。必须强调指出，薛定谔方程本身是量子力学中的一个基本公设，不可能从经典理论的已知原理中推导出来。决不可把上述“演算”看做一种推导。

在$H-J$方程中$\left(\frac{\partial}{\partial q_n}\right)^2$由$\left(\frac{h}{2\pi i}\frac{\partial}{\partial q_n}\right)$代替（以及在上述演算中我们从波动方程$\nabla^2\Psi=\frac{1}{u^2}\frac{\partial\Psi}{\partial t^2}$出发）乃是因为薛定谔希望假设一个线性方程，旨在使经典电磁理论中为人所熟知的叠加原理有效。从纯粹的数学观点看，在非线性理论还较鲜为人知的情况下，发展线性理论

是可取的。

10.量子理论中的概率概念

为了说明量子力学中“内禀概率”概念与经典物理学中“统计概率”概念之间的不同，让我们重温一下对经典统计力学情况的理解。

在经典统计力学（和气体动理论）中，人们处理的是宏观系统。一个宏观系统，例如一种气体，由大量分子，比方说，10^{23}个分子组成。分子的速度可能全部起作用。分布函数 N（v）dv/N，表明一个分子随机地取速度在 v 和 v+dv 之间的“概率”。这个概率具有以下的统计意义：想像一个大数系统（如上述气体），全都经过类似的制备，使之具有同样的宏观条件。这就是吉布斯“系综”。如果对分子速度所作的测量是用相同的程序对所有系统进行的，那么 N（v）dv 将代表分子的所有速度值的分布。因此，“概率”就不可能通过一次单一测量来证实；它必须通过对系综的类似制备的系统作大数测量来证实。

那么这一概率在下述意义上就不是“内禀的”。气体的所有分子的运动，在经典物理学中，是受运动的动力学方程（例如，哈密顿正则方程）支配的。如果知道了任一瞬间所有分子的坐标和动量，那么任何其他时刻——过去和未来的任意时刻——的坐标和动量，原则上都能借助运动方程，从分子间的力（和分子与容器壁的相互作用）而完全确定。因此，原则上，任何时刻的速度分布都能作确定性的预言，并不需要用概率函数 N（v）。概率概念被引进，是因为我们并不想知道 10^{23}个单个分子的坐标和动量的所有细节。在谈论一个分子的速度是 v 的概率为 N（v）/N 时，我们忽略了其他 N—1 个分子的坐标和动量，或是取了它们的平均。所以，经典物理学中使用概率概念是为了实践的方便而选择的，而不是原则上绝对必需的。

在量子力学的现行体系中，概率概念起着一种基本的作用。在下一个附录里，我们可以看到，对一个处于态 u_n（它不是 C 的一个本征态）的系统（比方说，一个单个原子）的物理量C作测量，就会以概率$|S_{nj}|^2 = \left|\int v_n^* v_j d\tau\right|^2$导致C的一个本征值$c_j$，$v_j$就是对$c_j$而言的C的本征态。对于有 j，概率$|S_{nj}|^2$同样不能由单一次测量，而只能用对一个系综的系统作大数测量来证实，所有这些系统都制备处于态 u_n。在这一点上，概率涵义是与经典物理学中的涵义相似的。但是有着这样一个基本的不同：在量子力学里，概率特征不是来自我们忽略原则上存在的其他详细的信息（这些信息可以使概率概念的引进不必要），而是内在于事物的本质之中。这种不同被表达为，在经典物理学中，概率概念伴随着或隐含着有“隐变量”存在；在量子力学中，按照盛行的哥本哈根学派的哲学，概率概念不伴有隐变量存在。另一种表达方式是说，经典物理学基本是决定论的，

而量子力学则基本是概率论的。

物理学中定律具有内禀概率本性的最早表述是卢瑟福和索迪关于原子核的放射性衰变 $-\frac{dN}{dt}=\frac{N}{t}$

或 $N=N_0e^{-t/\tau}$

这个经验定律暗示出，单个核的衰变是受随机性支配的。

1917 年，爱因斯坦引入了跃迁概率的概念，特别是感生发射的概念。令N_1，N_2为处在E_1，E_2能量状态的原子的数密度，$\bar{n}$为能量$E=E_2-E_1$的平均光子数。$\bar{n}$由在原子和辐射间平衡态下的普朗克公式给出，N_1，N_2由玻耳兹曼公式给出，

$$N_2=N_1e^{-\beta\varepsilon}，\beta=1/kT$$

令AN_2为每秒$2\to1$的自发跃迁数，$b_{21}N_2\bar{n}$是每秒$2\to1$的感生跃迁（辐射）而$b_{12}N_1\bar{n}$是相应的$1\to2$的吸收跃迁数。平衡时

$$AN_2+b_{21}N_2\bar{n}=b_{12}N_1\bar{n}$$

但从$N_2=N_1e^{-\beta\varepsilon}$和$\bar{n}$=〔$e^{-\beta\varepsilon}-1$〕$^{-1}$，能得到

$$N_2=\bar{n}N_2=\bar{n}N_1$$

因此

$$A=b_{21}=b_{12}$$

这个 $b_{21}=b_{12}$ 关系式明显地可用量子力学中矩阵元之间的关系表达。

不平衡时，动理学方程为

$$-\frac{dN_2}{dt}=AN_2+b_{21}\bar{n}N_2-b_{12}\bar{n}N_1=A〔\bar{n}(N_2-N_1)+N_2〕$$

在无任何辐射场时$\bar{n}$＝0，且

$-\frac{dN_2}{dt}=AN_2$从这里可以看出，这一关系与放射性衰变的关系完全相同。因此,纯随机性概念或内禀概率概念在其特征在量子力学中显露出来以前就已经隐含了。

11．概率公设与互补公设之间的一致性

互补公设的若干数学推论是：

（1）两个算符具有共同的本征态的必要和充分条件是：算符对易（特殊情况除外）。

（2）一个厄密算符的本征态形成一个完全集。

因此，如果 AB－BA=0，则 A，B 具有共同的本征态

$$Au_n=a_nu_n，Bu_n=b_nu_n$$

此处 a_n，b_n 是本征值，u_n 是本征函数。

如果对一个系统测量 A，而该系统处在 A 的一个本征态 u_n，则按概率公设，测量结果有一个确定性的值 a_n。

$$<A>=\int u_n^* A u_n d\tau = a_n$$

在状态 u_n 中对 B 的测量则确定性地是 b_n。

但如果我们测量 C，而对 C 有 AC－CA≠0，那么 C 的期望值可作如下计算。令 C 的本征函数是 v_n。

$$Cv_n=c_nv_n$$

u_n 能展开为 v_n 的完全集

$$u_n=\sum_j S_{nj}v_j，\sum_j |S_{nj}|^2=1$$

求和包括对谱 V_j 的连续部分的积分。这时我们得到非常重要的结果，甚至对一个单原子而言也有

$$<C>=\int u_n^* C u_n d\tau=\sum_{j,k} S_{nj}S_{nk}\int v_j^* C v_k d\tau=\sum_j |S_{nj}|^2 c_j$$

这就是说，C 不是确定性地知道，C 的每个本征值 C_j 都可以概率 $|S_{nj}|^2$ 获知。这个结果是与互补公设一致的，

即正则对p，q服从$pq-qp=\frac{h}{2\pi i}$和不确定性关系。

12．哥本哈根学派和量子力学其他观点

对量子力学的哥本哈根哲学不满的重要物理学家不只是爱因斯坦一人。最有趣的是，大多数量子理论和量子力学的奠基者都对它感到“不快”。为此，薛定谔在发展初期就强烈反对Ψ的概率诠释，当玻尔邀请他去哥本哈根并试图说服他时，他说他要早知如此，宁愿从未创造过波动力学。直到 1955 年，薛定谔（NuovoCimento，第十集，第二期）还表达了他的看法，可以概述如下：

（1）经典力学基本上涉及的是从对系统的观察资料中发现和表达有关定律（通过运动方程表达的因果定律），而且在方位天文学中有其基础。量子力学是对经典力学的模仿，无论在形式上（哈密顿—雅可比方程）还是在精神上（即由薛定谔方程所隐含的一种因果关系，虽然因果性现在指涉的是测量结果的概率，而不是经典意义上的状态）都是如此。

（2）但是经典力学与量子力学之间的情况有一个基本的差别。在经典力学中，它的起源及其应用的主要领域是方位天文学，存在三种可能：①对系统进行观察乃至跟踪观察，②从观察数据中找出系统的本质（用运动的哈密顿量和运动方程描述），以及③从已知本质（哈密顿函数和运动方程）出发，作出系统在任何其他时刻的因果预言。在量子力学中，当要把同样的观点保留在有关测量结果的公设中时，存在两种不可能：①对一个单一系统进行观察或跟踪观察（由于所有设想的实验，诸如γ射线显微镜等，都只是发明者的智力创造），②从观察结果（绝不是同一系统在时间进程中的观察结果）演绎出系统的本质（哈密顿量），因为理论容

许给出的只是测量的可能结果。

（3）这些差别和由此而在量子力学中引进的修正，其本身是无可置疑的。但量子力学的哲学态度，即那种主张只直接与所有实验观察打交道，而这些实验却只能想像而不能实际进行的测量理论，则是可置疑的。还有这样的涵义（当然，它是量子力学诸公设整体的一部分），即所有观察到的和可能观察到的东西与量子力学所选择的称为可观察量的东西精确相符。薛定谔认为，这确实是一个与哲学兴趣相关的问题，但它将迫使我们改造量子力学的概念体系。

在量子力学发展的极早期，德布罗意曾试图寻找一个“双重解”（法国《物理杂志》,1927 年 5 月）。他对泡利的批评一直保持沉默,直到 D.玻姆为量子力学现行的数学结构给出一种经典诠释,使他的主导哲学重新复活（《物理评论》85 卷第 166 页、第 180 页，1952 年；87 卷第 389 页，89 卷第 1458 页，1953 年）。

德布罗意和他同事们的一些著作有：

德布罗意，La Physique Quantique Restera-t-elleIndeterministe（1953 年）；

德布罗意，Nuovo Cimento，Ser. X，vol. I（1955 年）；

Une Tentative d’lnterpretationCausaleetNonLinéare de la Mécanique Ondulatoire（1956 年）；

J. P. Vigier，StructuredesMicro-Objectsdansl’Interpretation Causale de la Théorie des Quanta（1955 年）。

M. 普朗克也反对哥本哈根哲学（即认为一个物理量只有当它被观察和测量时才具有意义）。

在附录 10 中，我们曾按照哥本哈根学派现有的哲学强调了量子力学的内禀概率特征。按照这里的基本公设（见附录 11），对一个处于 u_n 态的系统测量物理量 C,其结果为 C 的一个本征值 c_j，c_j 出现的概率为 $|S_{nj}|^2=|Su_n^*u_j d\tau|^2$。由于任何单一测量的结果不能确定性地加以预言，因此,使倾向于经典决定论哲学的人们要问量子力学中是否也存在某种迄今未知的“隐变量”，以致概率概念不是内禀的，而是对隐变量的某个平均过程的结果。这个问题已由冯·诺埃曼考察过（Mathematishe Grundlagen Viriablesder Quantenmechanik，1932），答案为隐变量不可能存在。但对隐变量存在的这一否定取决于所使用的理论本身的条件，而不是广泛尝试了其他可能理论的结果。其他的研究也有人做了（D. 玻姆，N. 维纳和 A. 西格尔等人）。F. J. 贝林法蒂新近对各种隐变量理论作了评论，其书名是《隐变量理论概述》（ASurveyofHiddenViriablesTheories，Pergamon Press，N. Y.，1973）。

13.对称性

13.1　引言

在第六和第十三章中,我们已简要考察了一些熟悉的对称性和与系统的某些物理性质的不变性相关联的守恒律，即 1）对空间平移的对称性和动量守恒，

对空间转动的对称性和角动量守恒，

对时间平移的对称性和能量守恒；

2）对空间反射的对称性和“宇称”守恒，

对时间反演的对称性（基于对宇称 P、电荷共轭 C 和时间反演 T 三者混合的对称性的一般信念，这种对称性等价于对 CP 的对称性）；

3）四维电磁势和波函数在规范变换

$$A_\mu \rightarrow A_\mu + \frac{\partial}{\partial x_\mu}\chi$$

$$\psi \rightarrow \psi \quad \exp(\frac{ie}{\hbar c} \cdot \chi)$$

下的对称性导致电荷 e 的守恒。场量子为光子。

4）与中子β衰变相联系，我们提到为一 2×2 的■正矩阵所表示的同位旋空间的转换，并谈到了 SU（2）群。关于电弱统一理论，我们说到 SU（2）×U（1）规范场。

13.2 物理现象与对称群

在本附录中，我们将以一种基本的方式阐释这些术语的含义，以及物理现象与群的数学概念间的关系。

13．2．1U（1）

我们将从电磁场和波函数的规范变换入手，说明物理现象、对称性概念、规范场和对称群之间的联系。

我们先注意到麦克斯韦电磁场方程组和量子力学中的波动方程在规

$$A_\mu \rightarrow A_\mu + \frac{\partial}{\partial x_\mu}\chi$$

$$\psi \rightarrow \psi \quad \exp(\frac{ie}{\hbar c} \cdot \chi)$$

下是不变的。我们把这表述为由场 Aμ和ψ所描述的物理现象对于为相 $\exp(\frac{ie}{\hbar c}\chi)$所表征的变换具有对称性。代表这些变换的矩阵是一维的，即是一个数。变换$\exp(\frac{ie}{\hbar c}\chi)$显然是■正的。因此名词U（1）的含义是“■正的、一维的”。

场方程在 U（1）下的不变性的一个推论，为一守恒律，即在 exP $(\frac{ie}{\hbar c}\chi)$式中的电荷e的守恒。在量子场中，“量子化粒子”为（矢量量子）子，其静止质量为零，是电荷相互作用的传递者。

现在，我们以稍微不同的观点加以总结：我们起始于具有由 U（1）或其他群代表的对称性的变换。这种对称性决定了此场（即在这种变换下不变的场）的性质，也决定了量子化粒子的性质和与此不变性相联系的守恒律。

13．2．2SU（2）

下一步，我们要把这些概念扩展到更复杂的情形中。我们从在下述意义上质子和中子“非常相像”的物理唯象论开始。从“低”能散射实验中发现，p 和 p、p 和 n、n 和 n 之间的“强”相互作用都是同样的。如果我们忽略与强相互作用相比“很小”的库仑相互作用和它们之间的质量差，则可以把 p 和 n 表示为一个同位旋 I=1/2 在抽象同位旋空间中的两个分量态，$I_3 = \pm\frac{1}{2}$。这并不仅仅是个名称上的约定，而是有两个粒子间相互作用的“电荷无关性”为其物理基础的。这样，对一个复空间中的二分量矢量，它到另一矢量的变换由下式给出

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

由■正性条件和矩阵等于 1，导出

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$$

及$aa^* + bb^* = 1$。这一由三参量、二维■正矩阵$\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$确定的变换构成一个群，称作 SU（2）（意指特殊■正单模二维群）。

SU（2）具有矢量（u_1，u_2）“长度”的不变性

$$u_1'u_1'^* + u_2'u_2'^* = u_1u_1^* + u_2u_2^*$$

SU（2）群是非阿贝耳群（即非交换群）。

SU（2）群与三维空间中的旋转群R_3是同态的。对一个旋量$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$在旋转中转过角$\Theta$，变换为

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \exp\left[i(\sigma \cdot n)\frac{\Theta}{2}\right]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

其中σ_x，σ_y，σ_z为泡利矩阵（迹为 0）。

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}，\sigma_y = \begin{bmatrix} o & -i \\ i & o \end{bmatrix}，\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

因此对无限小转动

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} i(\sigma \cdot n)\frac{\Theta}{2}\right]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$= \left[1 + \frac{i}{2}\sum_x \sigma_x\Theta_x\right]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

现在，我们应用SU（2）群于同位旋$I = \frac{1}{2}$的变换。这种变换保持同

位旋不变。但如前所述，质子和中子的不同不仅仅在于它们的 I_3“量子数”；它们的质量和电荷也不同。因此，p 和 n 严格说来并不构成 SU（2）的基本粒子。

13．3SU（3）

13．3．1SU（3）——八重法

SU（3）群由特殊■正单模三维矩阵表示，其行列式等于 1。它是一个八参量群，即有 8 个独立 3×3 的矩阵，就像 SU（2）有 3 个 2×2 的 6 矩阵一样。它是一种非阿贝耳群。在群理论表示中，有许多不可约的表示 $D^{(1)}$，$D^{(3)}$，$D^{(6)}$，$D^{(8)}$，$D^{(10)}$，$D^{(15)}$……等，它们给出单态 1、三重态 3、六重态 6、八重态 8、十重态 10……。

在 20 世纪 50 年代，从高能加速器中发现了许多粒子。从这些粒子的性质（如自旋、质量、寿命、衰变方式等）中，人们发现若赋予这些粒子以一些新形态的量子数，如重子数 B、同位旋（I，I_3）、奇异数 S、超荷 Y 等将是很有助益的，这些量子数之间有下述关系

$$Y=B+S$$

$$Q=I_3+\frac{1}{2}Y$$

盖耳—曼和奈曼依循坂田稍早些时候（1955 年）提出的某些基本粒子能被归类于一个对应于 SU（3）的 8 维表示的八重态的想法，加以改进，于 1961 年编制出一个许多基本粒子的“周期表”。9 个赝介子（J=0）适应于一个八重态加单态；9 个矢量介子（J=1）适应

于一个八重态加单态；8个重子$(J=\frac{1}{2})$适应于一个八重态；10个重子（J=3/2）适合于一个十重态。8 个反重子（J=1/2）也构成了一个八重态 $D^{(8)}$。

这些都列于下表中，其中括号内的数字为以百万电子伏特计算的质量：

介子 B=0				
赝	矢	I_3	Y	S
K^+（493）	K^*（891）	$\frac{1}{2}$	1	1
K^0（498）	K^{*0}（891）	$-\frac{1}{2}$	1	1
π^+（140）	ρ^+（769）	1	0	0
π^0（135）	ρ^0（769）	0	0	0
η（548）	ω（769）	0	0	0
π^-（140）	ρ^-（769）	-1	0	0
$\overline{K}^0$（498）	$\overline{K}^{*0}$（891）	$\frac{1}{2}$	-1	-1
K^-（491）	$\overline{K}^{*-}$（891）	$-\frac{1}{2}$	-1	-1
η'（958）	φ（1019）	0	0	0
$D^{(8)}$，$D^{(1)}$，	$D^{(8)}$，$D^{(1)}$			

重子					B=1				
$J=\frac{1}{2}$		I_3	Y	S	$J=\frac{3}{2}$		I_3	Y	S
P	(938)	$\frac{1}{2}$	1	0	△	(1236)	$\frac{3}{2}$	1	0
n	(939)	$-\frac{1}{2}$	1	0	△	(1236)	$\frac{1}{2}$	1	0
Σ^+	(1189)	1	0	-1	△	(1236)	$-\frac{1}{2}$	1	0
Σ_0	(1192)	0	0	-1	△	(1236)	$-\frac{3}{2}$	1	0
ˆ	(1115)	0	0	-1	Σ^{*+}	(1382)	1	0	-1
Σ^-	(1197)	-1	0	-1	Σ^{*0}	(1382)	0	0	-1
Ξ_0	(1314)	$\frac{1}{2}$	-1	-1	Σ^{*-}	(1382)	-1	0	-1
Ξ^-	(1321)	$-\frac{1}{2}$	-1	-2	Ξ^{*-}	(1530)	$\frac{1}{2}$	-1	-2
					Ξ^*	(1530)	$-\frac{1}{2}$	-1	-2
					Ω^-	(1675)	$-\frac{3}{2}$	-2	-3
$D^{(8)}$					$D^{(10)}$				

把介子和重子归之于 SU（3）的单态 1、八重态 8 和十重态 10 的想法，得到随后对质量为 548 的η和质量为 1675 的Ω^-的实验发现的有力支持。但在迄今已知的粒子中，尚未发现三重态 $D^{(3)}$和六重态 $D^{(6)}$。

13．3．2SU（3）——夸克和量子色动力学

盖耳—曼成功的“八重法”方案表明 SU（3）对称性概念的方向是正确的。接下来一步的重要进展是一个全新的概念，分别由盖耳—曼和兹韦

克于 1964 年独立提出，即所有“基本”粒子都由三个“夸克”粒子和它们的反粒子，即u，d，s和ū，$\bar{d}$，$\bar{s}$构造而成，这些夸

克和反夸克自旋都为 1/2，并具有分数电荷（见第十三章中的表）。这些 u，d，s 是被称为“味”的“量子数”。SU（3）$_f$应用于味，u，d，s 就形成了三重态 $D^{(3)}$表示（这在介子和重子中都没表现出来）。

但很快人们就意识到，需要量子数为“粲”的第四种夸克。1974 年，丁肇中和里希特各自独立发现的 J／ψ粒子，以及随后发现的ψ′、ψ″、d、f、η^c、Λ^c等粒子确证了这种粲夸克。

1979 年，另一种新的粒子 Y 的发现要求引入第 5 种夸克 b（称“底”），而这得到了随后发现的其他 Y′、Y″粒子的支持。有人提出存在第 6 种夸克 t（称“顶”），但目前尚未“确定地证实”。

在由夸克来构造所有介子和重子时，对Ω^-遇到了必须把三个同一的 s 夸克置于最低的 S 态的困难，因为所有夸克都有自旋 J=1／2 违背泡利原理。韩和南部阳一郎于 1965 年提出，对夸克肯定还有另一个量子数。这个量子数是“色”，有三种“色”。每一个 u、d、s、c、b、（t）夸克都有三种“色”。三种色构成了 SU（3）$_c$群的三重态 $D^{(3)}$。

在与应用 U（1）于电磁场相同的意义上，应用 SU（3）$_c$。于夸克，得出“规范潮的量子化粒子，称作“胶子”。胶子传递夸克之间的相互作用，就像光子在电磁场中所起的作用。但与电性为中性的光子不同的是，胶子带有“色荷”，并且彼此有相互作用。再有，类似于对电磁场的 U（1）导致量子电动力学理论，SU（3）$_c$也导出称做“量子色动力学”的理论。

13．4SU（2）×U（1）——电弱尝规范场

在第十三章第 4 节，我们已很简略地提及电磁相互作用与弱相互作用的统一。这里，我们要进一步加以说明。

弱相互作用理论的历史是相当长的，它始自费米 1934 年的β衰变理论。到 50 年代，普遍认为弱相互作用由大质量的中间玻色子传递。那时规范场理论仅在电磁场情形中为人们所知，对称性为 U（1），它是阿贝耳群（见本附录 13．2．1 节）。普遍的规范场理论是由杨和密耳斯于 1954 年发展起来的，它把对称性推广到非阿贝耳群，例如 SU（2）、SU（3）。人们发现对“精确的”对称性，这个理论给予量子化矢量玻色子（b 场）（对应于 U（1）对称性情形中的光子）以零静止质量。这也是这个理论是否能重整化的问题。

1961 年，格拉肖提出为统一弱相互作用和电磁相互作用，使用 SU（2）×U（1）群于中间玻色子和光子。但是按杨—密耳斯理论，玻色子的零静止质量的问题仍未解决，直到 1964 年黑格斯引入自发对称破缺，使玻色子的静止质量不等于零。借助于黑格斯机制，温伯格和萨拉姆随后于 1967—1968 年系统提出了 SU（2）×U（1）理论。但玻色子场的重整化问题还未能解决，时至 1971 年才由霍夫特证明它确实是可重整的。这个统一的

场称为电弱场。

格拉肖-温伯格-萨拉姆理论终于形成了。在这个理论中，四个量子是玻色子 $W^{\pm}$，Z^0 和光子。$W^{\pm}$ 和 Z^0 粒子已于 1983 年在欧洲核子中心（CERN）发现。

13．5 更多的对称性

近来，人们在三族夸克 $\begin{pmatrix} u \\ d \end{pmatrix}$，$\begin{pmatrix} c \\ s \end{pmatrix}$，$\begin{pmatrix} t \\ b \end{pmatrix}$

和三族轻子

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$$

之间，从它们如下表所示的电荷和质量中，注意到某种明显的对称性：

夸克	$u(\frac{2}{3})$	$c(\frac{2}{3})$	$t(\frac{2}{3})$	(e)
	～10	～2000	?	(MeV)
	$d(-\frac{1}{3})$	s（—1/3）	b（—1/3）	(e)
	～5	～100	～5，000	(MeV)
轻子				
	e（—）	μ（—）	τ（+）	(e)
	0.5	105	1784	(MeV)
	$\nu_e(0)$	$\nu_\mu(0)$	$\nu_\tau(0)$	(e)
	$<2\times10^{-4}$	<4	<143	(MeV)

如果这种相似性不是偶然的，这种“对称性”是真实的，那么它将是某种意义非常深远的东西。

参考文献

A. Pais，InwardBound，（OxfordUniversityPress，1986）W. R. Frazer，ElementaryParticles，（Prentice—Hall，1966）

（关于“八重态”）

T. D. Lee，Symmetries，AsymmetriesandtheWorldofParticles，（University of Washington Press，Seattle，1988，）pp. 44—5 译者后记

吴大猷先生是我们敬仰已久的中国物理学界的老前辈。我们为能翻译他的物理学历史和哲学著作而深感荣幸。在这里，让我们对他这本著作的来龙去脉作一交待。

1974 年夏开始，吴先生曾应同仁之邀在台湾清华大学、台湾大学作过数次有关物理学概念发展历史和哲学的系列讲座。讲座的对象是具备一

定物理学背景知识的物理系学生,而这些讲座的内容又往往是他们以前所不熟悉、或不曾有过系统的了解的，同时这些内容对于他们理解物理学又极有帮助。

第一次是在台湾清华大学作讲演。那次系列讲演着重评述现代物理学的两门基础学科，即相对论和量子力学的物理意义和哲学意义。此讲演的英文版已由台湾联经出版社于 1975 年出版，题目是《现代物理学基础的物理本质和哲学本质》（The Physical and Philosophical Nature of the Foundation of Modern Physics）。吴先生在那本书中用中文写了一个自序，其中说：

“量子力学和相对论，乃物理学的两大基石。目前整个宇宙的现象定律和理论，小至原子，基本粒子，大至星云，无不建在这两基石上。惟量子力学和相对论的基本重要性，不仅是它们在各部门科学的‘应用’；它们对物理学的观念和物理学的理论的性质等问题，都引起第 19 世纪物理学家和科学哲学家所未想到的问题。爱因斯坦于 1905 年提出一观点，即在物理学中，一个观念，如‘时’、‘空’，必须有‘运作’的定义。这个观点，是他的相对论的枢纽。同是这个观点，也是 20 年后 W．海森伯氏创立他的新量子力学的出发点。后来量子力学的发展,更深地牵涉到‘物理学理论的基本性质’的哲学问题。半个世纪来，玻尔氏代表所谓‘哥本哈根派’的观点，爱因斯坦和些许大物理学家——包括量子论和量子力学的创始者，如 M．普朗克、L．德布罗意、E．薛定谔等——则不愿接受这观点。这是目前仍在研辩中的一个科学哲学上很基本的问题。”

此后，吴先生又应邀在台湾大学作了十四次每次两小时的系列讲座。讲座的内容是根据他发表于《国际现代物理学杂志》上的一篇题为《物理学：它的发展和哲学》（1989）文章加以修改和扩充的。这次讲座具有导论性质，预设了听众中的大部分人都已具备一定的物理学背景知识。在这个讲座基础上作了扩充，于 1992 年形成了一本讨论物理学的历史和哲学的专著，定名为《物理学：它的发展》（Physics：Its Development）。正文计 11 章，还有若干注释。此书还未公开出版过。

现在展开在读者面前的这本中文本《物理学的历史和哲学》是译者将原英文著作《现代物理学基础的物理本质和哲学本质》（正式出版物）及《物理学：它的发展》（内部参考资料）两书合并而成的。合并过程中，译者删去了若干重复的小节，尽力保持其原来论述的顺序。它能为读者提供一个物理学演变过程和哲学问题的纲要性框架。吴先生一直认为，一位物理学家，从对物理学的历史和哲学的更深刻更具批判性的理解中，会有助于更全面地了解物理学，同时也能获得精神上的更大满足与享受。

书末的附录和每章结尾列出的权威性参考文献是原有的,我们只在两书合并过程中作了一些顺序上的编排。

本书的第 1，2，3，4，5，6，7，11，12，14 章和相应的附录系金吾

伦译，8，9，10，13 章及相应附录为胡新和译，译后互校定稿。胡新和做了人名索引。

本书翻译和出版过程中，得到东方国际易学研究院院长朱伯■教授、副院长董光璧研究员、台湾朱高正先生和中国大百科全书出版社副总编王德有教授的鼎力相助，特致谢意。

译者

1996 年 12 月于北京