学校的理想装备

电子图书·学校专集

校园网上的最佳资源

物理学基本概念和基本定律溯源

BOOK
内部资料 非卖品

序

在教学中引入历史，和专门研究历史是不同的。专门研究历史，要求对史料作详细的考证工作；物理课程中引用史料，则是为教学目的服务的。在课程中讲述一些物理思想的发展过程，穿插一些掌故轶事，对提高教学质量，并结合课程内容进行一些有关的思想教育，都是大有好处的。但做的时候不能喧宾夺主，而应烘云托月。这个“月”便是课程中要讲授的物理原理和物理概念，特别是本课中的那些重点和难点。个体认识活动的逻辑过程与认识发展的历史过程，就其总体和梗概而言，是一致的。教学中学生的难点，往往也是物理学发展史上长期未能克服的困难。历史上物理大师们与之辩论和斗争的错误观点，往往仍保留在我们今天学生的概念之中。认识上的曲折和反复，正可反衬出正确理解物理概念的重要。物理学史中关键性的突破和前辈物理学家伟大贡献的精髓，也正是物理教学的重点。围绕这些问题倒不妨着意点染，利用历史资料把中心问题衬托出来，有时会收到意想不到的良好效果。

向义和先生的这本《物理学基本概念与基本定律溯源》由 14 篇论文组成，其中属于经典物理部分的 9 篇，属于近代物理部分的 5 篇，基本上包括了普通物理学中的重要课题。每篇论文构成一个独立的主题，回答一个物理定律是怎样得到的，或一个物理概念的思路是怎样发展而来的。与编年史或分期史相比，这种编排更便于读者了解各个物理原理发展的来龙去脉，特别有助于教师将物理学史的材料运用到物理教学中去，有利于学生更生动、更深刻地理解物理概念的内容实质，受到科学方法论的启迪和世界观的教育。愿这本书的出版将在我们的物理教学方面起到积极的推进作用。

赵凯华

1992 年 8 月

前　言

《物理学基本概念与基本定律溯源》是部大学物理教学参考书。

本书以一系列论文式的专题组成其独特的体系，不同于单纯的概述性体系。每篇论文构成一个独立的主题。陈述某一物理学概念是怎样形成的，某一物理学定律是怎样得来的，或者某一物理学理论是怎样建立的。全书所选的 14 个主题均是高等院校普通物理学课程中的重点内容，包括力、热、电、光与近代物理等各个部分。作者从物理学史的角度，对这些概念的起源和定律的建立作了论述。

物理学中一些基本概念都有一个形成过程。例如能量是个守恒量这一概念就是人们对自然界各种运动形式相互转化长期认识的结果，从机械能守恒到得出能量守恒与转化定律其间经历了大约 150 年的时间。在这些概念的形成过程中，包含了大量丰富多彩的内容。但是现有的物理学教材基本上是按演绎法编写的，能量概念是从推导动能定理中引进的，舍弃了历史发展过程中那些生动具体的细节，内容较为枯燥。本书从掌握概念的需要出发，对教学中一些重要而又难于理解的概念，如“熵”、“感生电场”、“位移电流”、“同时性的相对性”、“波粒二象性”等等，着力揭示它们的孕育和发展的历史脉络，使学生对其物理实质有较深入的理解。

在物理学概念形成的过程中曾经有过曲折与反复、分歧与斗争、停滞与突破。把这些过程介绍给学生，可使学生有身临其境的参与感，而且从正反两方面的对比中更能加深对概念的理解。例如力学和电学中超距作用和近距作用之争；热学中热的运动说和热质论之争：光学中波动说和粒子说之争；近代物理中在对波函数的物理解释上爱因斯坦与哥本哈根学派之争等。在现有的物理教材中往往缺乏这些两种物理学说论战史的介绍，常会造成一种物理学中无矛盾的假象，使学生难于从理性思维的高度把握本门学科的发展，也难于启发学生创造性的思维能力。本书力图弥补教材的这种不足。

本书以物理学家的原始论文为依据，在介绍物理定律建立的过程中，努力揭示物理学家的研究思路、创造性工作的特点以及所运用的研究方法，使学生可以从中受到丰富的科学方法论的教益与启迪。例如在发现万有引力定律的过程中，牛顿所采用的抽象简化、建立理想模型的方法就起了非常重要的作用。我们知道宇宙系统是一个多元的复杂系统，每个星体都是一个引力中心，都有一定的形状和大小，每个行星既不完全在椭圆上运动，也不在同一轨道上旋转两次，正如牛顿所说：“同时考虑所有这些运动之起因，是整个人类智力所不能胜任的。”牛顿采用简化模型的方法：从圆运动到椭圆运动；从质点到球体；从单体问题到两体问题。每次将他的理想模型与实际比较，再适当加以修正，最后使物理模型与物理世界基本符合。现有的物理教材所强调的是逻辑推理方法，对于在实际的科学发现中物理学家所使用的创造性思维方法，如物理类比、理想模型、理想实验、科学想像、科学直觉、试探猜测等等，都很少介绍，这就不利于发展学生的能力和培养开创性的人材。

全部物理学史告诉我们，新的物理概念和物理观念的确立是人类认识史上的一个飞跃，只有冲破旧的传统观念的束缚才能得以问世。例如，普朗克的能量子假设是在突破了“能量连续变化”的传统观念的基础上建立

的。同样，狭义相对论也是爱因斯坦在突破了传统的时空观念束缚的基础上建立的。这个思想发展的历程并不是一帆风顺的，而是经过不少思想上的曲折、困惑、疑虑、矛盾斗争，有时甚至动摇退却。在现有的物理教材中一般只把这一认识的结果，又是经过后人多次消化了的材料介绍给学生，使学生较难体会到科学工作者正确的科学观与世界观、科学素质与革命气质对科学发展的重要作用。本书努力于揭示物理学家这一心理发展过程，使读者从他们的成功与失败的经验教训中获得启示与鉴戒。

在编写本书的过程中得到了我校物理系教授张三慧、秦明华、徐亦庄、张泽瑜和副教授王以炳的热情帮助和具体指导，他们分别审阅了部分论文初稿，并提出了中肯的意见。国内物理史学界邹延肃、申先甲、杨福征等先生审阅了本书的初稿并提出了宝贵的意见。复旦大学物理系 85 年举办的量子物理史讲习班以及他们提供的参考资料对本书的编写也有不少帮助，作者在此一并表示衷心的感谢。

向义和

1992.7

一、万有引力定律的建立

万有引力定律的建立是牛顿“从运动现象研究自然力”的一个最辉煌的范例。本文将依据牛顿在各个时期写的手稿与论著，探讨牛顿论证的特色以及牛顿引力思想的发展过程，以期回答下述问题：牛顿是怎样从开普勒的行星运动规律和他的“离心力”公式推导出引力的平方反比律的？是如何解决椭圆轨道上运动物体的引力以及球体引力的问题？是怎样依据天文观测结果对引力定律进行实验验证的？

(一)开普勒定律的建立及引力思想的萌芽

开普勒定律描述了行星绕太阳运动的规律。它不仅使得人们有可能比较详细地进一步研究行星运动的“运动学”问题，而且还有利于研究行星运动的“动力学”问题。它为万有引力定律的建立奠定了基础。

约翰内斯·开普勒(Johannes Kepler， 1571—1630)是德国天文学家。1587 年，他进入杜宾根大学。在大学期间，他受到热心宣传哥白尼学说的天文学教授麦斯特林(Mästlin)的影响，成为日心说的忠实维护者。1591 年获文学硕士学位，后曾当路德教派牧师而学神学。1594 年，他得到大学的有力推荐，中止了神学课程，去奥地利格拉茨的路德派高级中学任数学教师并开始研究天文学。[1]

开普勒是一个深受毕达哥拉斯影响的数学家，他深信上帝是依照完美的数的原则创造世界的，他以数学的和谐性来解释哥白尼算出的行星的配置。在 1596 年发表的《宇宙的秘密》 (Myste-rium Cosmographicum)中，他把当时已知的六颗行星和从希腊时代就知道的仅有的五种正多面体联系了起来。他设想了一个模型：一个半径等于土星轨道的球内，内接一个正六面体，木星的轨道便在这个正六面体的内切球上；在木星轨道的球内，内接一个正四面体，火星的轨道便在这个正四面体的内切球上；其下依次是正十二面体，地球的轨道，正二十面体，金星的轨道，正八面体，水星的轨道。因为这种正多面体只有五种，所以开普勒相信行星只有六颗。而且各个球壳的大小和哥白尼算出的行星距离相差在 5%的范围内，这个安排全然是偶然性的和带有数学神秘性的。开普勒怀着敬意把新书寄给了布拉格的第谷，第谷看完这本书后，尽管他对书中的种种解释不太满意，但对开普勒的想象力和数学才能却很赏识，于是第谷写信给开普勒，请他到布拉格来研究天文学。[2]

第谷·布拉赫(Tycho Brahe，1546—1601)是丹麦天文学家，出生于一个贵族家庭，自幼喜欢观察星辰。1559 年进入哥本哈根大学学习法律，他的伯父希望他成为一个律师，但第谷并不热心于此。1560 年，通过一次日偏食的观察，他的注意力转向了天文学。他通过对行星在星空方位的观察和计算，发现当时的行星位置图表有严重的错误。他想要建立一个满意的行星理论，就必须有高度精确的星表，而这就需要长期进行新的准确的天文观测。1576 年，在丹麦国王腓特烈二世的资助下，他在哥本哈根海峡的汶岛上建立了一所宏大的天文台，他称之为天文堡。第谷对观测仪器进行了改进，增大了仪器的尺寸并安装在坚固的基础上，这就加强了仪器的稳定性，给仪器进行了精密的刻度，从而提高了仪器的精密度。第谷还按照大气对光线的折射效应对观察结果进行修正。他年复一年的观测，取得了

大量关于行星位置的准确记录资料。在 21 年的观测中，各行星的角位置误差仅有 2′，即 0.033°。1597 年，他离开汶岛，1599 年到布拉格任鲁道夫二世的御前天文学家。[3]

1600 年，开普勒接受第谷的邀请来到布拉格。在这里两位天文学上的巨人相会了。在第谷的安排下，开普勒觐见了国王，接受了“皇家数学家”的头衔。开普勒和他的老师有不同的特色和兴趣：第谷着重并善于实际观察，而开普勒则更醉心于数学和理论思考。第谷的精确观察与开普勒的深刻研究相结合得到了丰硕的成果，第谷多年精心观测得到的宝贵资料，为开普勒发现行星运动三定律奠定了基础。1601 年，重病的第谷把开普勒请到床边，作了临终的嘱托。第谷说：“我一生之中，都是以观察星辰为工作，我要得到一种准确的星表……现在我希望你能继续我的工作。我把底稿都交给你，你把我的观察的结果出版出来，题名为《鲁道夫天文表》，我们至少要有一点报答鲁道夫国王。”在开普勒允承以后，第谷就安然与世长辞了。开普勒精心整理并千方百计筹集资金，经过多年的努力，直到 1627 年才正式出版了这本有史以来最精确的天文表。[3]

第谷逝世以后，开普勒把大部分时间用于对火星的研究上。他试图使观察得到的准确的火星轨道新数据符合作匀速圆周运动的哥白尼体系。尽管他在计算中用了偏心等距点，经过二十多次不同方案的试验，历时四年的计算结果是：用哥白尼体系计算出的轨道比根据新数据计算出的轨道小八分。开普勒知道，第谷明察秋毫的慧眼和颇为精密的仪器记录的行星位置，其误差是远小于八分的。于是他敏锐地觉察到火星可能不是作匀速圆周运动。[3]

怎样确定火星的轨道呢？由于第谷对火星的观察资料是从运动着的地球上观察得出的，所以必须先弄清楚地球轨道的真实形状及它们的运行方式，以便确定在观察火星的日子里地球在什么地方，然后才可能利用这些数据来确定火星的运动。为此他充分利用每组火星年的观测数据，并用几何作图法确定地球轨道的形状，然后，又确定火星轨道的形状，他终于发现火星轨道是一个椭圆，进而又发现每个行星都沿着椭圆轨道运行，太阳就在这些椭圆轨道的一个焦点上，这就是开普勒的第一定律即著名的轨道定律。这个发现把哥白尼学说向前推进了一大步。用开普勒本人的话说：“就凭八分的差异，引起了天文学的全部革新。”[2]

在确定了行星轨道的形状后，开普勒又去寻求行星在轨道上的速率与位置之间的关系。他从观察火星的资料中发现火星距太阳近时运动得快，而在距太阳远时运动得慢。他试图从物理学上解释这一现象，设想太阳可能以某种力驱使行星沿轨道运行，这种力只作用在行星的轨道平面上，因而轨道上各点受太阳作用力的大小与离太阳的距离成反比；所以，按当时流行的动力学原则，行星运动的速率与这个距离成反比。这样，一颗行星沿它的轨道走过一小段距离所用的时间，应该同它到太阳的距离成正比。这是近似正确的。于是开普勒提出可把行星轨道上较长一段弓形分成一段段小圆孤，然后把各小圆弧到太阳的距离相加，用以计算行星沿其轨道走过较长一段弓形所用的时间。他假定每小段圆弧长为 2 个单位长，则这些距离的总和在数值上就等于太阳和行星联线所扫过的面积，从而得到行星到太阳联线扫过的面积与所经历的时间成正比的结论。[4]

开普勒实际上只计算出地球和火星这两颗行星在近日点和远日点时，

在相同的时间内其径矢扫过的面积相等。然而由于这种关系如此美妙和简单，致使他坚信这个关系无论对于哪个行星和在轨道的哪一部分都是真实的，这就是所谓的面积定律，即行星的径矢在相等的时间内扫过相等的面积。虽然开普勒在推导这个定律时所使用的假定是错误的，但这个定律却是正确的。它不仅准确地描述了围绕太阳的任何行星的运动，也适用于围绕任何行星的卫星的运动[5]。

图 1-1 开普勒假定，当行星沿其轨道运动时，所有从太阳到行星的联线的长度之和 SP_1+SP_2+…SP，在数值上近似等于 SP_1P_N 的面积之和。

图 1-2 开普勒面积定律

1609 年，开普勒在《新天文学》(NewAstron-omy)一书中发表了上述两个定律。但是他对自己获得的成就并不满意。他认为各个行星都沿椭圆轨道，以匀面积速度运行不是偶然的，必有某种更普遍的规律联系着太阳系的所有行星的运动；只有发现各个行星运动之间存在着统一的关系，才可以建立一个太阳系的整体模型，从而揭示出宇宙的谐和与一致。开普勒怀着这种信念，长年累月地考察了许多因素的各种可能的组合，终于在 10 年之后发现了这条规律。1619 年，开普勒在《世界的谐和》(TheHarmonicesoftheworld)一书中公布了这一定律 行星公转周期的平方同它们到太阳的平均距离的立方成正比。如果以 T 代表行星运行周期，以 R 代表行星到太阳的平均距离，则这个定律可以表示为

$$\frac{T^2}{R^3} = K$$

式中 K 是一常数。这个定律被称为谐和定律，它表明各个行星的运动速度和轨道大小之间很有节奏的比例关系，就象音乐中的和声一样。事实上，开普勒在《世界的谐和》中正是用乐谱的形式把六颗行星在远日点和近日点之间角速度的变化情况表征为一首“行星协奏曲”。[2]

开普勒坚信自然界存在着一条简单的法则，这是来自他早年对数学的迷恋。他承袭了古希腊毕达哥拉斯学派“数是实在世界的基础”的思想，深信上帝是依照完美的数的原则创造世界的，并认为根本性的数学谐和即所谓天体的音乐，乃是行星运动的真实的可以发现的原因。正是这种追求数学谐和的理想，对自然的单纯性和一致性的信念，使他克服了前进道路上的各种障碍，忍受着生活的不幸，在艰苦漫长而又毫无结果的工作期间能从这一泉源中得到精神鼓舞，以致在他最后发现这一定律时欣喜地写道：

“经过长时期不断的艰苦工作后，利用布拉赫的观测结果我发现了轨道的真正距离，最后终于找到了真实的关系……一下子消除了我心中的疑团，17 年来我对布拉赫观测结果的刻苦研究同我现在的这个研究是如此相符，以致我起初还以为是在做梦……。”[4]

在开普勒全面解决了行星运动的运动学问题之后，关于行星运动的动力学方面的问题就自然提出来了。“为什么行星会保持在轨道上？”“为什么行星这样运动呢？”开普勒在研究行星运动的规律时，就已经注意到

这个问题。他认为这种运动一定是由于力的作用而产生的。他在《宇宙的秘密》第二版中说：

“我一度坚信驱动一颗行星的力是一个精灵……然而当我想到这个动力随距离的增大而不断减小，正如太阳光随着与太阳的距离增大而不断减弱的时候，我得出了下面的结论：这种力必定是实在的——我说实在的并不是按字面的意义，而是……象我们说光是实在的某种东西一样，意思是说：那是从一实体发出的一种非实在的存在(unsubstantial entity)。”[4]

在英国人吉尔伯特(W. Gilbert，1540—1603)关于磁力的论文发表不久，开普勒在1605年给一个朋友的信中写道：“我的目的在于证明：天上的机械不是一种神圣的、有生命的东西，而是一种象钟表那样的机械，正如一座钟的所有运动都是由一个简单的摆锤造成的那样，几乎所有的多重运动都是由一个最简单的，磁力的和物质的动力造成的。我也要证明，何以应当用数字和几何来表达这些物理原因。”[4]开普勒设想发自太阳的磁力驱使行星沿其轨道运动。这一设想虽然是错误的，但是开普勒把可观察的实验现象作为出发点，从事实本身去寻求运动原因，这标志着近代物理学的主要特征之一的开端。

法国杰出的数学家和哲学家笛卡尔(Rene Descartes，1596—1650)通过对于“行星保持在轨道上运动的原因”的探索，发表了他的“旋涡说”，被当时很多人所接受，牛顿也是在这种理论的影响下成长起来的，因为当时在英国的大学里都讲授这个理论。旋涡说的理论是：宇宙空间充满一种稀薄的不可见的流质“以太”，各个聚集体周围的以太围绕聚集体形成大小、速度和密度不同的旋涡式运动，它产生的旋涡压力卷吸着周围的物体趋向中心物体，这就表现为引力作用。行星以其旋涡带着它周围的附属物沿着更巨大的旋涡围绕太阳旋转。笛卡尔的学说由于是从接触作用来说明引力的本质，因此比超距作用更易被理解和接受。[2]

1645年，法国天文学家布里阿德(I. Bulliadus)提出一个假设：从太阳发出的力，应和离太阳距离的平方成反比而减小。这是第一次提出平方反比关系的思想。牛顿正是在布里阿德思想的启示下产生了论证平方反比力的想法。

(二)引力平方反比律的发现

牛顿(Isaac Newton，1642—1727)，1642年12月25日的圣诞节出生于英格兰林肯郡的沃尔斯索普(Woolshorpe)村一个农户家里。12岁那年他进入了格兰姆中学。在毕业前他获得优秀学生的荣誉。1661年6月，牛顿以“减费生”身份考上著名的剑桥大学三一学院。在这里他受到“卢卡斯数学讲座第一任教授巴洛(Isaac Barrow)的引导而走向自然科学，特别是数学和光学的研究。这时他读了开普勒的《光学》和笛卡尔的《几何学》等著作。1664年经巴洛考试被选为助手。1665年，他获三一学院文学士学位。

就在牛顿毕业的这一年，英国发生了瘟疫。1665年到1667年，牛顿在故乡躲避瘟疫的大约十八个月的时间里，进行了在力学、天文学、数学和光学方面伟大的基础性研究工作。引力的平方反比律就是在这个时候发现的。后来牛顿在谈到他在1666年间一系列重要发现时写道：“这一年里，

我开始想到把重力推广到月球的运行轨道上去，在求出了在内球面上一个旋转的小球对球面的压力后，我就从行星运转周期的平方同它们到太阳的平均距离的立方成正比的开普勒定律推导出：使行星保持在它们的轨道上的力必定与它们到旋转中心的距离平方成反比。而后把使月球保持在它轨道上所需要的力和地球表面的重力作了比较，发现它们近似相等。所有这一切都是在 1665 和 1666 年瘟疫流行的年代里发现的。那时我正处于发明创造的青春年代，并且比任何时候都更关心数学和哲学。”[6]

牛顿是否在 1666 年发现了引力的平方反比律，物理学史家们众说不一。有人依据牛顿本人及其亲属朋友的回忆，认为牛顿在 1666 年发现了引力的平方反比律；有人认为这种说法依据不足。认为牛顿在这时只有验证平方反比律的“想法”，并未作任何验证。的确，这个时期牛顿没有发表任何著述，但是笔者从牛顿未发表的早期手稿中看到牛顿不仅验证了平方反比律，而且作了绝妙的论证。

牛顿在 1664—1665 年写的《未发表的记事手稿》(WasteBook)的手稿中，讨论了圆周运动的动力学，引进了“离心力”的概念，讨论了决定“离心力”大小的因素，并结合开普勒第三定律得出了离心力与半径平方成反比的定律。在公理及命题 20 中，他首次讨论了圆周运动的问题。他假设一个小球沿圆筒 def 的内表面运动，它必须持续地压在 def 上(图 1-3)。因为当小球在 c 时，它正向着 g 运动，如果 def 不再控制其运动，小球就会沿 cg 偏离中心：

图 1-3

“但是如果圆筒 def 本身对抗这种努力，使小球与 m 保持等距离，那么 def 就会从圆周 chb 每一点的切线上连续地制止或反射(Checking or reflection)小球。但是只有它们连续地相互紧压，圆筒 def 才能控制小球的运动方向。同样也可理解某物体被一根绳子束缚作圆周运动的情形。由此看来，所有作圆周运动的物体都力图脱离它们的运动中心(endeavour from the center)，否则小球就不会持续紧压 def”[6]

牛顿首次定量分析这个力的大小出现在《未发表的记事手稿》公理及命题 22 中，他推理说，在绕转半周中小球力图脱离中心的“全部力”，是大于能够产生或者破坏其运动的力的两倍，即两倍于使其运动的力：

“假设小球由 c 经 h 运动至 b。那么圆筒 def 的抗力(等于小球对 def 的压力)能够破坏小球从 c 向 g 运动的力，并使小球产生同样大小的，在相反方向上从 b 向 k 运动的力。”[7]

牛顿在《未发表的记事手稿》中用下述方法定量地讨论了离心力的大小。他从圆内接正方形路径开始讨论，这时一物体完成一周运动将被弹回四次，牛顿把四次碰撞的力与物体运动的力相比较，再把结果推广到了无穷多边的多边形，从而得出结论“全部的反射力与物体运动的力

之比如同全部边长(即周长)与半径之比。”其论证如下：如图 1-4 所示，球 b 绕中心 n 旋转，为了得到离心力的大小，牛顿先以正方形 abcda 路径的运动代替圆运动。在分析一次碰撞中小球在 b 处对器壁的作用力时，他写道：

“如果 ef＝fg＝gh＝he

图 1-4

=2fa=2fb=2gc=2ed，在小球从 a 运动到 b 时，2fa：ab=ab：fa=小球反射时对 fg 的压力：小球运动的力。”

“4ab=ab+bc+cd+da：fa=在一周中反射的力(即在 b、c、d、a 四次反射中的力)：小球运动的力。”

“照此推论下去，如果小球被边数为 6、8、12、100、1000 等等外切多边形的各个边反射，则全部反射的力对物体运动的力之比如同所有边的总和对这个圆的半径之比。所以如果物体被无限多边数的等边的外切多边形所反射，则全部反射的力与物体运动的力的比值如同全部边长(即周长)对半径之比。”

对这一最后结论我们可作如下一解说。“全部反射力”指小球在旋转一周的时间 T 内小球的离心力的“冲量”，即 fT。“小球运动的力”是指小球的动量 mv。按照上述结论可得

$$\frac{fT}{mv}=\frac{2\pi R}{R}\text{因为}T=\frac{2\pi R}{v}$$

由此得到$f=\frac{mv^2}{R}$。这就是我们现在都熟悉的向心力公式，但牛顿在当时并没有给出这个公式。

在 1665 年或 1666 年写的“仿羊皮纸手稿”(VellumManu-script)中，牛顿含蓄地表示了他的离心力定律，他写道：“当物体以速度 v 在半径为 R 的圆周上运动时的离心力(centrifugalforce)等于作用在一个从静止开始沿一直线运动的相同的物体上的力时，在圆周上运动的物体通过距离为 R的时间内，在直线上运动的物体将通过$\frac{1}{2}$R的距离。”[8]

对此我们可以作下面的演算：设离心力为 F，从静止开始在直线上运动的物体，在力F的作用下，加速度为F/m，在$t=\frac{R}{v}$的时间内所经过的距离为

$$\frac{1}{2}\frac{F}{m}\left(\frac{R}{v}\right)^2=\frac{1}{2}R$$

图 11

由此又可得

$$F=\frac{mv^2}{R}$$

牛顿在《未发表的记事手稿》中还把离心力推广到物体作椭圆运动的情形。他说：“如果物体在椭圆上运动，它在每一点(设它在该点的运动已知)上的离心力可以用一个在该点与椭圆相切的，具有同样曲率(Equalcrookedness)的圆求出。”

在 1669 年前的一份手稿《论圆运动》[9](OnCircularmotion)中，牛顿也讨论了离心力。他指出：“在物体由 A 到 D 作圆周运动的过程中，它

力图脱离中心的力的大小是这样的：在物体通过 AD(假定它很小)的时间内，它将使物体脱离圆周一段距离 DB。”(图 1-6)

图 1-6

因为 $$\frac{BE}{BA}=\frac{BA}{BD}$$

在无限短的时间内，BE 和 DE 之差以及 BA 和 DA 之差无限小，用 DE 代替 BE，DA 代替 BA，由此得出

$$\frac{DE}{DA}=\frac{DA}{DB}$$

如果这个力象重力一样在一条直线上作用，则使物体通过的距离将与时间的平方成正比。为了求出在绕转一周 ADEA 的时间内迫使物体通过的距离，就要求出 DB 与周长 ADEA 的关系。将上式中的 DA 用 ADEA 代替，即可得下式

$$DB=\frac{ADEA^2}{DE}$$

因为周长 ADEA＝2πR，直径 DE＝2R，所以

$$DB=2\pi^2R=19.7393R$$

DB 表示在周期 T 的时间内迫使物体在直线上通过的距离。接着牛顿提出一个推论：“在不同半径的圆上，运动物体的离心力正比于直径除以周期的平方。”对于这些结论我们可以作如下的计算，以 f 表示离心力的大小，m 表示物体质量，则

$$2\pi^2R=\frac{1}{2}\frac{f}{m}T^2$$

由此得到 $$f=\frac{4\pi^2mR}{T^2}$$

亦即 $$f\propto\frac{2R}{T^2}$$

若将$T=\frac{2\pi R}{v}$代入，我们又可得向心力与速率及半径的关系式

$$f\propto\frac{2R}{\left(\frac{2\pi R}{v}\right)^2}=\frac{v^2}{2\pi^2R}\propto\frac{v^2}{R}$$

牛顿在这篇文章中接着指出：由于行星运行周期的平方同它们到太阳的距离的立方成正比，所以行星退离太阳的离心力(Theendeavours of receding from the sun)与行星到太阳的距离的平方成反比。

牛顿在这里没有叙述他的推导过程。但是有了他的离心力公式和开普勒第三定律，我们很容易从圆运动得出上述结论

$$f\propto \frac{R}{T^2}=\frac{R^3}{T^2R^2}=\frac{1}{R^2}\times 常数$$

即$f\propto \frac{1}{R^2}$

值得注意的是，在牛顿早期的著作中没有使用向心力这一术语。他在《论圆运动》中所讲的物体脱离运动中心的努力，实际上指的是物体的惯性，即任何物体在没有外力作用下具有保持匀速直线运动的趋势。正如他所说的在圆筒中作旋转运动的小球，如果不受圆筒的控制，则小球将沿一条切线飞出。在 1684 年他写的《论运动》的手稿中，首次提出了向心力的定义：“我称之为的向心力是物体被吸引或迫使向某个中心点的力”。在 1687 年的《自然哲学的数学原理》一书中又进行了非常明晰的讨论，并把在圆运动中物体受到的向心力和它所具有的惯性离心趋势加以区别。

(三)平方反比律的地月验证

牛顿晚年的挚友彭伯顿(Pemberton)博士，在他的《哲学解释》的序言中讲述了牛顿对他讲的关于引力思索的一个生动的故事：

“1666 年，他在沃尔斯索普村家里躲避瘟疫时，有一次他独自坐在花园里，忽然看到一个苹果从树上掉了下来，他吃了一惊，同时便沉浸在引力的思索中，沉思引力的巨大威力。即便是在我们所能够达到的，离地心很远很远的地方这种引力也丝毫不见减小。不管是在最高的建筑的最上层，还是在最高的山顶上都是一样。因而他就想，是否可以设想，这种力的作用范围可能要比通常设想的还要大得多。比如说一直延续到月亮。如果是这样的话，月亮的运动必定受到引力的影响，甚至很可能这个力就是使月亮维持在它的轨道上的原因。”[6]

这个关于引力思索的故事，说明牛顿开始产生了把地面上物体的重力与保持月亮在它的轨道上的力看成同一本质的力的想法。苹果落地激起了一个伟大科学家的有准备的头脑的灵感，牛顿以他非凡的洞察力在落体运动与月亮运动之间搭起了一座智慧之桥，觉察到了天体运动与地球上的运动的统一性。这个故事是牛顿在发现万有引力定律的艰苦历程中的一个有趣的小插曲。当然不能把它讹传为牛顿看到苹果落地就发现了万有引力定律。

牛顿在 1669 年前写的《论圆运动》中，开始对引力的平方反比律进行地月验证。他把维持月亮在它的轨道上的力与地球表面上的重力作了比较。他利用了天体测量的结果，月亮与地球之间的距离是地球半径的 60 倍，月亮绕地球一周的时间为 27 天 7 小时 43 分钟或 27.3216 天。他采用了当时 1 纬度是 60 哩的公共量度，得到地球半径为 3500 哩。他根据在这篇文章中导出的公式，在向心力的作用下，月亮在绕转一周的时间内下落的距离为

$$2\pi^2R=19.7393R$$

因为在常力作用下，物体下落的距离与时间的平方成正比，且 1 哩等于 5000呎，所以月亮在一分钟内下落的距离为

$$h=\frac{2\pi^2R}{T^2}=\frac{19.7393\times 60\times 3500\times 5000}{(27.3216)^2\times (24)^2\times (60)^2}$$

$=13.6$呎

他把这一数据和伽利略测量的地球表面上的重物在一秒钟内下落16呎的数据加以比较

$$\frac{13.6}{16\times(60)^2}\approx\frac{1}{4300}$$

由此得出结论：“地球表面上的重力比月亮对地球中心的离心力大4000 多倍。”按照引力与距离平方成反比的关系应约为 3600 倍，二者相差 19%，这说明在开始进行地月检验时，牛顿的计算与预期的结果并不是“相当接近”，而是“不太符合”。

二十年后，牛顿在他的名著《自然哲学的数学原理》一书中，进一步阐述了维持月球在它轨道上的力与地面上物体所受的重力是同一本质的力的思想。他在该书第一篇关于向心力的定义 5 的说明中，设想了一个理想实验，论述了月亮的绕转运动与地面上平抛运动的联系。他写道：

“如果从山顶用弹药以一定的速度把一个铅球平射出去，那末它将沿一条曲线射到两英里以外落到地面；如果能消除掉空气阻力，而且发射速度增加到两倍或十倍，那么铅球的射程也会增加到两倍或十倍。而且用增加发射速度的办法，我们可以随意增加其射程，并同时减少它所画的曲线的曲率，使它终于在十倍、三十倍或九十倍远的距离处落到地面，或者甚至可以使它在落地以前绕地球一转；或者最后，也可以把它发射到空中去，在那里继续运动以至无穷远而永远不落到地面。”[10]

牛顿在该书第三篇命题 4 的注解中，继续用理想实验的科学方法，论证使月球保持在它轨道上的力与地面上物体所受的重力是同一本质的力。他设想有几个月亮围绕地球转动。“如果有一个小月亮很靠近地球，以至触及到地球上最高的山顶，那么使它保持轨道运动的向心力当然就等于它在山顶处所受的重力。这时如果小月亮突然失去了运动，它就如同山顶处的物体一样以相同的速度下落。如果它所受的向心力并不是重力，那么它就将在这两种力的共同作用下以更大的速度下落，这是与我们的经验不符的。可见重物的重力，月亮的向心力，必然是出于同一个原因。因此使月亮保持在它轨道上的力就是我们通常称为‘重力’的那个力。”[10]

牛顿在这个命题的正文中，根据天文观测和大地测量的结果进一步作了平方反比律的地月验证。月地之间的平均距离为地球半径的 60 倍。大地测量得出地球的周长是 123249600 巴黎尺，它的半径是 19615800 巴黎尺。月亮绕地球一周的时间为 27 天 7 小时 43 分，即 39343 分。我们根据牛顿在《论圆运动》中所导出的公式，可以算出月亮在一分钟下落的距离为

$$h=\frac{2\pi^2R}{T^2}=\frac{2\pi^2\times60\times19615800}{(39343)^2}$$

$$=15.009\text{ 巴黎尺}$$

牛顿当时的计算值是$15\frac{1}{12}$巴黎尺。因为引力按距离的平方反比而减小，在地面处引力应比月球处强约 60×60 倍，所以地面上物体在 1 秒钟下落的距离应为$15\frac{1}{12}$巴黎尺。这同惠更斯用单摆在巴黎测出一个物体从静止落下时第 1 秒通过的距离相等。因此平方反比律被证明是正确的。

(四)椭圆运动的平方反比律与引力普适性概念的确立

1666 年，牛顿在沃尔斯索普村家里，对引力定律进行了第一次检验失败后，第二年春天牛顿回到剑桥大学，结束了他的动力学创造性工作的第一个时期。从 1667 年到 1679 年。这方面的研究工作几乎完全中断。关于中断的原因有各种各样的说法：一种说法是当时他所知道的地球半径之值不准确，计算误差较大；另一种说法是，这时牛顿的兴趣转向光学研究并卷入了在光学方面与胡克的论战。还有人认为牛顿当时遇到了在动力学方面难以克服的困难。其困难之一是：牛顿的计算是以行星轨道是圆形为前提，事实上行星的轨道是椭圆形，这就给论证平方反比定律带来很大的困难。其困难之二是：在计算天体之间的引力时，牛顿未能证明可以把每个天体看作全部质量集中于中心的质点来看待。牛顿似乎估计到这些困难，意识到要解决这些动力学上的难题需要一个长期的准备时间。在这个时期中，他能够从容不迫地吸收并消化早期研究所积累起来的经验，寻求解决这些难题的工具。

1673 年惠更斯在他的著作《摆钟》(Horologium Oscillator-ium)中，导出了单摆运动的公式为

$$t=\pi\sqrt{\frac{L}{g}}$$

式中 L 为摆长，g 为重力加速度，t 为单摆往或返的时间，即半个周期。根据这一公式，他在巴黎用一个周期为 2 秒的单摆，精确测出摆长为 3.0565呎，从而计算出重力加速度为 30.1666呎/秒 ^{2}(换算为公制约等于 9.8 公尺/秒 ^{2})。

精密摆钟的发明和广泛使用，为发现重量与质量两个概念的差异提供了一个条件。1671 年法国的里切尔(JeanRicher，1630—1696)到南美赤道附近的卡因岛作天文观测时，发现他从巴黎带去的摆钟变慢了，当他把摆长适当缩短，走时就准确了。惠更斯对这一现象作了研究，认为这是因为赤道附近物体受到更大的离心力，从而抵消了物体的部分重力。这就得出结论：同一物体在地球表面的不同地点其重力也不同，使重量概念与质量概念的差异变得明显了。[2]

惠更斯在这本书中，还提出了他称之为的“离心力定理”。他提出：“二个等重量的运动物体在大小不同的圆周上作等速运动，则该二物体的离心力之比等于二圆周直径的反比。”“二个等重的物体在大小相同的圆周上以不同速度各作匀速运动，则运动较快的物体与运动较慢的物体的离心力之比等于其速度的平方比。”[11]

即 $$F\propto\frac{v^2}{R}$$

图 1-7

惠更斯是这样来确定向心加速度公式的。如图 1-7 所示。他以切线 AD 到圆周的距离来确定向心力的作用。图中设 AD 是物体在某个时间间隔 τ 内以 v 作匀速直线运动时通过的距离 s，(即图 1－7s＝AD)。而 δ ＝AC＝DB，认为是由于物体在时间 τ 内做加速运动而造成的与 AD 的偏离，则

$$s=v\tau$$

$$\delta = \frac{1}{2}a\tau^2$$

又有 $s^2=(2R-\delta)\cdot\delta$

当 δ 很小时 $s^2=2R\delta$

所以 $$\alpha = \frac{v^2}{R}$$

在此以前，胡克在引力的研究上就作出了贡献。胡克已觉察到引力和地球上物体的重力有同样的本质。1662 年和 1666 年，他曾在山顶上和矿井下用测定摆的周期的方法做实验，试图找出物体的重量随与地心距离而变化的关系，但没有得到结果。他在 1674 年的一次演讲“证明地球周年运动的尝试”中，提出要在一致的力学原则的基础上建立一个宇宙学说，为此他提出了三条假设：

“第一，据我们在地球上的观察可知，一切天体都具有倾向其中心的引力，它不仅吸引其本身各部分，并且还吸引其作用范围内的其它天体。”

“第二，凡是正在作简单直线运动的任何天体，在没有受到其它作用力之前，它将继续保持直线运动不变。”

“第三，受到引力作用的物体，越靠近吸引中心，其引力也越大，至于此力在什么程度上依赖于距离问题，在实验中我还未解决，一旦知道了这一关系，天文学家就很容易解决天体运动的规律了。”[2]

1679 年，哈雷与伦恩也按照圆形轨道由开普勒第三定律和惠更斯在 1673 年发表的向心力的公式，证明了作用于行星的引力与它们到太阳的距离的平方成反比。但是他们不能证明行星在椭圆轨道上也是如此。这年 10 月 24 日，胡克给牛顿的一封信中，提出了引力反比于距离的平方的猜测，并问道：如果是这样，行星的轨道将是什么形状？胡克给牛顿的信重新激起了牛顿对动力学的兴趣，使牛顿把他的注意力转到椭圆运动问题。1679 年底，胡克给牛顿写信介绍一种分析曲线运动的新方法，即物体沿曲线运动具有两个分量，一个是切向分量，一个是向心分量。1686 年 7 月 14 日，牛顿给哈雷的信中谈到这封信时说：“这是真的，他的信使我偶然发现应该用图形法，在椭圆上的验算是靠图形法的研究进行的，就这样搁置达五年之久，直到你要寻找那篇论文时。”牛顿表示他在 1679 年作了行星在椭圆轨道上时引力平方反比律的证明。

1679 年(或 1684 年)牛顿写了一份手稿，标题是：“引力指向太阳的行星可能在椭圆上运动的证明。”[14]这份手稿一开始就提出三条假设：

假设一如果物体不受到阻力或其它外力，将一直做匀速直线运动。

假设二运动的变化永远正比于使运动发生变化的力。假设三提出运动合成的平行四边形法则。

根据开普勒定律行星绕太阳作椭圆运动，则各个行星必然受到太阳的引力，这些引力方向指向太阳，太阳就是这些引力的中心。牛顿首先在命题一中证明了这个问题。

命题一：如果一个物体在真空中运动而且被一个固定中心所吸引，它将在一个固定平面内运动，而且在相等的时间内，物体与力心的连线，扫过相等的面积。

图 1-8

如图 1-8 所示令 A 为力心，假设把时间分成许多相等的小段，在第一段时间内，物体从 B 匀速运动到 C，如果没有受力，它将沿直线 BC 运动到 I，而且 CI＝BC。在 C 点它受到一个指向 A 的冲力，在冲力作用下它沿 CD 线运动。从 I 点作平行于 CA 的直线得 ID，D 就是第二个瞬时末物体所在的位置。由于同底等高，△ACD 和△ACI 的面积相等，也等于△ABC 的面积。同理，如果在第二、第三、第四、第五各瞬时末，物体在 D、E、F、G 各点受到指向 A 的冲力的作用，将使物体在每一段相等的时间内经过直线段 DE、EF 和 FG 等等。而且△AED，△AFE 和△AGF 的面积都等于△ADC，△ABC。因此在相等的时间内扫过相等的面积。

图 1-9　1684 年牛顿写的短文《论运动》的第一页

现在假定时间间隔无限地减小，而三角形的数目无限地增大，那末，引力的冲力就可能变成连续的，而由无限多小线段组成的折线就可能变成一条曲线。在连续的引力作用下，物体与力心的连线所扫过的面积和物体运行的时间成正比。

牛顿的这一论证表明只要给出了物体的初始位置和初速度和力作为位置的函数，根据这个作图法就可以确定物体的近似路径，这是牛顿动力学决定论的一个生动的例子。

把这一定律应用于开普勒行星运动图象。从开普勒的面积定律和命题一的逆命题，就可以得出作用在行星上的力都是指向太阳的，太阳是吸引行星的力心。解决了这一问题后，牛顿就可以讨论在指向焦点的力的作用下，物体在椭圆上运动的问题，进一步探讨物体受力的大小与物体到力心的距离的关系。

命题二：如果在椭圆上运动的物体受到一个指向椭圆焦点的引力，在椭圆的两个顶点处物体受到的引力与物体到焦点的距离的平方成反比。

如图 1-10 所示，A、C 是椭圆的顶点，F 是焦点。AFE 和 CFD 是在相等的时间内径矢扫过的相等的面积。假设 AE 和 CD 是如此之短，以致可以把它们视为直线。这样扇形的面积就变成半个矩形的面积。由此得

$$\frac{1}{2}\mathrm{AF}\times\mathrm{AE}=\frac{1}{2}\mathrm{FC}\times\mathrm{DC}$$

$$\frac{\mathrm{AE}}{\mathrm{CD}}=\frac{\mathrm{FC}}{\mathrm{FA}}$$

图 1-10

设 AM 和 CN 是通过顶点 A 和 C 的切线，分别作 AM 与 CN 的垂线 ME 和 DN。因为椭圆在 C、A 两点是同样弯曲的，所以 EM 与 DN 之比等于 AE 的平方和 CD 的平方之比。

$$\frac{\mathrm{EM}}{\mathrm{DN}}=\frac{(\mathrm{AE})^2}{(\mathrm{CD})^2}$$

所以

$$\frac{\mathrm{EM}}{\mathrm{DN}}=\frac{(\mathrm{FC})^2}{(\mathrm{FA})^2}$$

当物体在顶点不受力时，将沿切线 AM、CN 运动。在这段时间内，引力的作用相当于使物体从 M 运动到 E，从 N 运动到 D，所以物体在椭圆顶点 A 受的引力与在顶点 C 受的引力之比等于 ME 与 ND 之比，结果得到与距离的平方成反比，即

$$\frac{F_A}{F_C}=\frac{(FC)^2}{(FA)^2}$$

命题三：如果在椭圆上运动的物体受到一个指向椭圆焦点的引力，这个引力与物体到焦点距离的平方成反比。

图 1-11

如图 1-11，假设把时间分成许多相等的小段。当物体运动到椭圆上的 P 点时，如果不受力，在△t 的时间内它将沿切线匀速运动到 X 点。由于物体在 P 点受到指向焦点的引力的冲量，使物体从沿着切线运动变到沿着弦 PY 运动，Y 是△t 时间末物体在椭圆上的位置，这个距离 XY 是由于引力冲量作用的结果，因此 XY 必定与引力 F_P 成正比，并且平行于引力的方向，即平行于 PF。用同样的方法，可得在 p 点的引力 f_p 与 xy 成正比，xy 平行于 pF。由此得出

$$Fp：fp＝XY：xy$$

根据椭圆的几何关系和开普勒面积定律，最后得出物体在 P 点受的引力 F_p 与在 p 点受的引力 f_p 之比与该点到焦点距离的平方成反比

$$\frac{F_P}{f_P}=\frac{(pF)^2}{(PF)^2}$$

1684 年 1 月，雷恩、哈雷和胡克三位当时英国科学界著名人士在伦敦相叙，讨论行星运动的轨道问题。胡克说他已通晓，但拿不出计算结果。于是牛顿的好友哈雷专程去剑桥请教牛顿。牛顿告诉哈雷他已经计算过了，断然地说：行星绕日轨道是个椭圆，但手稿压置多年一时找不到。应允重新计算，约期三个月后交稿。哈雷按约期再度访剑桥，牛顿又提出一份手稿《论运动》(DE Motu)。牛顿在此基础上，在 1684 年 8—10 月间写了一本小册子《论球体在液体中的运动》[15](On the motion of Spherical Bodies in fluids)在这篇文章中，他首先讨论了物体在阻力介质中的运动，他给向心力下了明确的定义，并再一次证明了向心力定律：向心力正比于速度的平方被半径除。他还根据开普勒的面积定律和几何关系，证明了在椭圆上运动的物体指向椭圆焦点的向心力与物体到焦点的距离的平方成反比。更为重要的是，他在这篇文章中首次提出了引力的“万有性”或“普适性”。以前他只考虑太阳对行星的引力，没有考虑行星之间的引力。依照这种理论，行星绕太阳的运动是严格的椭圆。他根据太阳相对于太阳系重心(The centre of gravity 应理解为质心)的位移，发现向心力并不总是指向系统的重心。因此他认为“这些行星既不是严格地作椭圆运动，也不会在同一轨道上绕行两次”。“每个行星的轨道依赖于所有行星的联合运动以及它们之间的相互作用。”对这个结果如果只认为太阳才有引力是无法解释的，所以牛顿说只有计及行星彼此之间的作用，才能说明行星的运动。这意味着引力不仅是太阳的本性，同样也是行星的属性。到

1685 年牛顿进一步在《自然哲学的数学原理》中写道：“依此定律一切物体必定互相吸引”。这就是万有引力了。

在这篇文章的后一部分，讨论物体在有阻力的介质中运动

SIR ISAAC NEWTON’S
MATHEMATICAL PRINCIPLES
OF
NATURAL PHILOSOPHY AND HIS
SYSTEM OF THE WORLD
Translated into English by Andrcw Motte
The translations revised,and suppiied with an historical
and explanatory appendix
by
FLORIAN CANJORI
LATE PROFESSOR OF
THE UNIVERSITY OF CALIFURNIA
CAMBRIDGE
AT THE UNIVERSITY PRESS
1934

图 1—12　1686 年牛顿写的《自然哲学的数学原理》

时，牛顿首次提出了质量概念，还把质量和重量加以区别。他指出介质的密度正比于它们的重量，取决于它们固有的物质的量。为了研究引力与质量的关系，他通过分析外力作用于物体产生的运动得出“加速力等于质量乘加速度”的规律。

(五)球体引力问题的解决与万有引力定律的确立

在《论运动》和《论球体在液体中运动》的基础上，牛顿于 1684 年底开始了《自然哲学的数学原理》的写作。1685 年春天完成了球体引力定理，以卓越的彻底精神和严谨的科学态度解决了求解球体引力的问题。[10]

首先，牛顿在《自然哲学的数学原理》第一篇命题 70 中，论证了均匀球面内任一粒子 p，球面对它的引力的合力为零(见图 1-13)。他通过 P 作两个圆锥面，在球面上截取两个面积元 HI 与 KL。面积元对 p 的引力与距离平方成反比，与面积大小成正比。而面积大小又与到 p 点的距离平方成正比。把这两个比相结合得一等比，即 1∶1。因此，这两部分的引力，大小相等，方向相反，互相抵消。据此可知，整个球面上的一切引力均有相反相等者与之抵消，故粒子 p 将不会被这些引力吸引到任何一个方面。

图 1-13

接着他在命题 71 中,论证了均匀球面外的一个质点被吸向球心的力与该质点到球心的距离平方成反比。

设球面外与球不同距离处的两个质点 P 和 p，通过 P 和 p 作两个小圆锥面，它们在球面上分别截取两段相等圆弧 HK、hk 和 IL、il，作垂线 SD，sd，SE，se，IR，ir，IQ，iq，如图 1-14 所示。由几何关系可得

$$PI : PF = RI : DF$$

和　　　　　　　　　pf∶pi=df∶ri

等式两边相乘，因为当圆锥角趋于零时，df=DF，所以

图 1-14

PI·pf∶PF·pi=RI∶ri=IH∶ih(1.5.1)

又　　　　　　　　　PI∶PS=IQ∶SE

ps∶pi=se∶iq

同理可得　　PI·ps∶PS·pi=IQ∶iq(1.5.2)

式 1.5.1 和式 1.5.2 中相乘得

$$(PI)^2\cdot pf\cdot ps:(pi)^2\cdot PF\cdot PS=HI\cdot IQ:ih\cdot iq$$

这就是当半圆 AKB 绕直径 AB 转动时弧 IH 所画出的圆环面，与当半圆 akb 绕 ab 转动时弧 ih 所画出的圆环面之比。吸引着粒子 P 和 p 的力沿连线的方向并指向这些面。力的大小与面积大小成正比，与面积到粒子的距离平方成反比，从而得 HI 和 hi 对粒子 P 和 p 的引力之比为

pf·ps∶PF·PS

因这些斜交的力分别是它们沿PS方向指向中心S的分力的 $\frac{PI}{PQ}$ 倍(或 $\frac{PS}{PF}$ 倍）以及 $\frac{pi}{pq}$ 倍(或 $\frac{ps}{pf}$ 倍），

所以把粒子 P 吸向中心 S 的引力与把粒子 p 吸向中心 s 的引力之比值为

$$\frac{F_P}{f_P}=\frac{PF\cdot pf\cdot ps}{PS}:\frac{pf\cdot PF\cdot PS}{ps}=\frac{(ps)^2}{(PS)^2}$$

据此理由，弧 KL 和 kl 绕转所画出的圆环面对粒子 P 和 p 的引力之比，同样是 $(PS)^2$ 与 $(PS)^2$ 之比。所有圆环面的力之比具有同样的比率。因此整个球面施加在质点上的力与质点到球心距离的平方成反比。

牛顿根据前两个命题很容易地得出下面一系列推论。在命题 73 中，他证明了在密度均匀的球中，粒子受到该球的引力正比于该粒子到中心的距离。在命题 74 中牛顿证明了位于球外的一个粒子受到球的吸引力与该粒子到球心的距离平方成反比。

在命题 75 中，牛顿指出一个球受到另一个球的引力与两个球心之间的距离平方成反比。他根据命题 74 得出均匀球对球外一个粒子全部引力，仿佛是由放在球心的一个单个粒子给于的。另方面，如果这个粒子被球上各个粒子以同样力吸引，则单个粒子对球的引力必定是与球对单个粒子的引力一样大。但是对粒子的引力是与它到球心的距离的平方成反比，因此对球的引力与两个球心之间的距离平方成反比。

在命题 76 中，牛顿进一步把质量概念引进球体引力定理。在推论 3 中他提出：在球心距离相同的条件下，两个球之间的引力正比于两个球的乘积。在推论 4 中，他提出：在球心距离不同的条件下，引力与两个球的乘积成正比而与球心之间的距离的平方成反比。牛顿在这里所说的物体的乘积就是指的物质的量或质量的乘积。牛顿在《自然哲学的数学原理》第

三篇《论宇宙系统》的命题 7 中，精辟地表述了万有引力定律：“一切物体所具有的引力正比于它们各自所包含的物质的量，与距离的平方成反比。”

如果我们用 m_1 及 m_2 分别表示二物体的质量，R 表示二者之间的距离，则引力 F 为

$$F \propto \frac{m_1 m_2}{R^2} \text{或} F = G\frac{m_1 m_2}{R^2}$$

这就是万有引力定律的数学表示式，式中 G 是引力常数。

万有引力定律的发现揭示了引力的普适性。牛顿在《自然哲学的数学原理》第三篇“哲学中的推理法则”第三条中，提到引力是物体的普遍属性时写道：“如果依靠实验和天文观察，普遍发现地球周围的所有物体都被吸向地球，而且这种吸引正比于这些物体各自所含的物质之量；月球同样也按其物质之量而被地球所吸引；另一方面，我们的海洋又被月球所吸引；所有行星都相互吸引；而且彗星也以同样方式被太阳所吸引；那末，根据这条法则，我们必须普遍承认，所有物体都天然具有相互吸引的本性。”[6]

在牛顿以前，无论是东方还是西方，天与地的区分是根深蒂固的。没有任何一项成果能说明天上运动与地上运动服从相同的规律。牛顿的引力定律体现了天上运动与地上运动的统一性，它把开普勒的行星运动和伽利略的落体与抛体运动统一了，从而把天体运动纳入到根据地面上的实验得出的力学原理之中。这是物理学史上第一次伟大的综合，也是人类认识上一次巨大的飞跃。

关于引力的机制，牛顿对此没有作过任何假设。他在《自然哲学的数学原理》最后一节中提到引力的起源时写道：“一直到现在，我已将天体现象及海洋运动用重力来说明了，但重力之来源如何，却没有说过，此项力必有一原因，贯彻至太阳及行星之中心，完全不受丝毫损失。……我还没有方法由此项现象以推及重力之根源，我亦不想设立一假设。”牛顿认为在没有从观察和实验中发现重力之原因时，决不杜撰假设。他在《自然哲学的数学原理》第三篇哲学中的推理法则第一条中写道：“除了那些真实而已足够说明其现象者外，不必去寻求自然界事物的其它原因。”

有种说法把引力有超距作用归之于牛顿，这是没有根据的，他在《自然哲学的数学原理》书中未提此说。在缺乏实验线索的情况下，他拒绝提出一种使引力效应从一物体传到另一物体的机制，并不意味着他同意超距作用。1692 年 2 月 25 日他在写给神学家本特利(Bentley)的信中，这样写道：“在我看来，说引力是物质本身固有的，内在的和根本的，因而一物体可穿过真空距离作用于另一个物体，毋需有其它的东西作为媒介把它们的作用和力传达到另一物体上，这是甚为荒谬的，因为我相信，凡是在哲学方面有思考能力的人决不会陷入这种谬论之中。引力必然是持续不断地按一定规律施加作用的动因造成的；至于这个动因究竟是物质的抑或非物质的，我则留给读者自已去考虑。”[4]

(六)引力定律的实验验证

在牛顿时代由于缺少精密的仪器，很难测量两个小物体之间微弱的引力。但是牛顿想出了一个巧妙的方法来证明 G 的常数性。他考虑地球表面

上一个质量为 m_1 的物体，它的重量为 m_1g，设地球质量为 M_e，因此根据万有引力定律

$$m_1 g = G\frac{m_1 m_{地}}{r^2} \text{或} G = \frac{r^2}{m_{地}}g$$

在一给定地点，$r^2/m_{地}$当然是一个常数。如果在这个地方，所有物体都恰好具有相同的 g 值，那么我们就证实了 G 也是个常数，而不管这些物质的化学组成，结构及形状如何。这正是牛顿用实验证明的。他不是简单地用大物体和小物体的下落来测量 g，而是使用长度相等但材料不同的计时摆这种精确得多的方法。当时已知，具有给定长度的一个单摆的周期 T 仅同 $1/\sqrt{g}$成比例。经过详尽的实验，所有结果都指出G为常数，因此他写道：“这就是在我们实验所及的范围内一切物体的性质；因此(根据法则 3)可以肯定这是无论哪种物体都具有的性质。”[4]

1798 年，即又过了 100 多年，英国物理学家卡文迪许(Cavendish)做了测量 G 的经典实验。如图 1-15 所示，两个小铅球被固定在轻杆的两端，用一根系在杆的中点的极细金属丝把杆沿水平方向悬挂起来，细丝上固定着一面小镜子。小铅球的附近对称地安放着两个较大的铅球，这两对 大质量和小质量之间的引力使杆在水图 1－15 平面上转动。当金属丝的扭转所产生的回复效应恰好与相吸引的力平衡时，杆就停在一个平衡方向上，反射光把微小的角偏转放大为光点相当大的位移。根据金属丝扭转的角度可以测出力的强度，证实了万有引力定律的正确性，并测定了 G 的数值为 6.670×10^{-11}牛顿·米^2/千克^2。

图 1-15

在《自然哲学的数学原理》第三篇宇宙体系中，牛顿就运用引力定律讨论了太阳系的行星、行星的卫星、彗星的运行，以及海洋潮汐的产生，行星之间运动相互受到引力干扰，即所谓摄动。1680 年 11 月与 1681 年 3 月有个大彗星两度出现，牛顿通过计算得出它是以太阳为焦点作抛物线运动的，它受太阳的向心力也服从距离平方反比定律。并指出：“彗星是行星之一种，它绕太阳运行具有极大偏心率。”[16]

预见并发现新的行星是显示引力理论威力的最生动的例证。1781 年的一个夜晚，英国人威廉·赫歇尔(WilliamHerschel)用他自己制作的 10 英尺望远镜观察天空，发现了现在称为天王星的行星。发现天王星后，它的运动就成了不断研究的主题。积累的资料表明它的运动有某些极小的不规则性，不能归因于任何已知天体的摄动效应，人们猜测在天王星之外可能有一个未曾预料到的行星，它对天王星的轨道起了附加的摄动作用。这个想法引起了剑桥大学一位青年学生亚当斯(J.C.C.Ada-ms)的兴趣。他使用万有引力定律，从观察到的天王星的运动，来计算这颗未知行星的位置，这是一项极其艰巨困难的数学工作。他毕业两年后得出了该问题的数学结果，推断出新行星在特定时刻出现在轨道上的位置。1845 年 10 月亚当斯写信给格林威治皇家天文台请求他们用强大的望远镜在预言的位置上寻觅这颗假设的新行星。然而，由于亚当斯是一个不出名的青年数学家，所以没有受到足够的重视。

1846 年 6 月，另一位法国青年勒维耶(U.J.J.Leverrier)发表了类似

的独立计算结果，他给出的这颗被猜测的行星位置很接近于亚当斯的预言位置。后来英国终于决定对亚当斯的理论计算予以验证，当英国人正在迟缓地进行某些观测时，勒维耶把他的预言送交柏林天文台台长那里，幸好正是在收到这封信的晚上，这位台长手边有一幅有助于寻觅该行星的新星图，于是他亲自寻觅，并在非常靠近预言位置的天区辨认出了这颗行星。这样在 1846 年太阳系里又增添了一颗海王星 这确实是万有引力定律的一次辉煌胜利！[4]

从牛顿建立万有引力定律的科学历程中，我们能够吸取哪些科学思想和科学方法论的教益呢？

首先，牛顿在科学研究中坚持以经验为基础，他认为在没有从观察和实验中发现引力之原因时，决不杜撰假设。牛顿的“ 不杜撰假设 ”具有方法论的意义，这种方法论与他同时代的大多数人所遵循的方法迥然不同。牛顿的同时代人都追随笛卡儿，探索自然现象的起因，构筑引力的机制。而牛顿则不然，他所关心的不是引力“ 为什么 ”会起作用，而是“ 如何 ”在起作用。他的目的是寻求引力所遵从的规律，提出准确的数学描述，证明行星系统如何依赖于引力定律。

但是，牛顿的认识路线也不同于受经验主义影响很深的胡克的认识路线。胡克强调从实验上去探求引力定律，忽视数学推理的必要性。他的表述停留在定性认识上，缺乏定量的成份。他没有认识到当时更需要的是数学推理，而不是实验，因为所有行星运动的实验资料都已总结在开普勒定律之中，而胡克面对实验事实，迟迟不能提出物理模型，进行数学推导，从而确立力的定律。这是他在方法论上不如牛顿的地方。

牛顿所遵循的认识途径是从实验观察到的运动现象去探讨力的规律，然后用这些规律去解释自然现象。正如他在《自然哲学的数学原理》一书的前言中写的：“ 我奉献这一作品，作为哲学的数学原理，因为哲学的全部责任似乎在于——从运动的现象去研究自然界中的力，然后从这些力去说明其它自然现象。 ”[10]爱因斯坦对牛顿的科学认识道路给于了高度的评价。他在《自述》一文中写道：“ 你(指牛顿)所发现的道路，在你那个时代，是一位具有最高思维能力和创造力的人所能发现的唯一的道路。 ”[17]

牛顿的科学认识道路对以后物理学的发展产生了深刻的影响，许多物理学家都沿着牛顿的道路进行工作。1827 年，安培在《电动力学理论》一书中，阐述了他处理电磁理象的方法：从观察事实出发，撇开力的性质的假说，推导出这些力的表示式，确立一般规律，最后他明确指出：“ 这就是牛顿所走过的道路，也是对物理学作出重大贡献的法兰西知识界近来普遍遵循的途径。 ”

牛顿研究方法的一大特点是对错综复杂的自然现象敢于简化，善于简化，从而建立起理想的物理模型。宇宙间星体的相互影响是无限复杂的，每个星体都是一个引力中心，所以它是一个相互作用的多元的复杂系统；而且每个星体都有一定的形状和大小；每个“ 行星既不完全在椭圆上运动，也不在同一轨道上旋转两次 ”。面对这一情况，不采用简化模型予以分别处理是极为困难的。1684 年牛顿在《论微粒》一书中指出：“ 同时考虑所有这些运动之起因，是整个人类智力所不能胜任的。 ”[18]牛顿是怎样对这一复杂系统进行简化的呢？他采用的简化模型的步骤是：从圆运动到椭

圆运动；从质点到球体；从单体问题到两体问题。

他一次又一次地将他的理想模型与实际比较，再适当加以修正，最后使物理模型与物理世界基本符合。所以牛顿的万有引力定律既解释了为什么行星的运动近似地遵守开普勒定律，又说明了为什么它们又是那样或多或少偏离开普勒定律。

牛顿把一切物体间的引力归结为粒子间的引力的思想，对以后的物理学家影响很大，19 世纪 20 年代，毕奥、萨伐尔、拉普拉斯和安培在研究电流之间的作用力时，总是把它归结为电流元之间的作用力。

牛顿研究方法的另一特色是运用形象思维的方法，进行创造性的思维活动，他构思了一些神奇的理想实验，创造了新的物理图象，来揭示天体运动与地面上物体运动的统一性。牛顿的科学思想和科学方法不仅使他少走弯路，发现了万有引力定律，而且深刻地影响着以后物理学家的思想、研究和实践的方向。这说明科学思维方法的极端重要性。从物理学的重大发现中去吸取科学思想、科学方法的营养，对提高教学水平，提高学生分析问题、解决问题的能力都是大有裨益的。

参考文献

[1]《中国大百科全书》，第 1 版，上海，中国大百科全书出版社，1980 年 12 月，61，189

[2]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，264—269，319—325

[3]杜正国，“ 开普勒 ”，《物理教学》，1983 年 3 月

[4][美]G.Holton 著，张大卫等译，《物理科学的概念和理论导论》，上册，第 1 版，北京，高等教育出版社，1983 年 2 月，62—68，221，246—248，232—234

[5]张三慧，《从伽利略到牛顿》，第 1 版，北京，北京出版社，1982 年 1 月，120，126—127

[6]HerivelJohn,TheBackgroundtoNewton ’ sPrinciple,Oxford Clarendon Press,1965，65—67

[7]Herivel John,ibid,129—130,147

[8] Herivel John,ibid,110

[9] Herivel John,ibid,195—197

[10] Newton,MathematicalPrinciplesofNaturalPhilosophyand his System of the World, Berkeley, University of Californiapress, 1962,3,193-199,407-409

[11][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》，第 1 版，北京，商务印书馆，1986 年 5 月，32—33

[12]何圣静主编，《物理定律的形成与发展》，第 1 版，北京，测绘出版社，1988 年 3 月，63—64

[13]阎康年，“ 关于万有引力定律的发现年代问题 ”，《物理》，1981 年，3 月

[14] Herivel John,ibid,246—254

[15] Newton,on the Motion of Spherical Bodies in Fluids,

unpublished Scientific papers of Isaac Newton,CambridgeEng,the Universitypress,1962,280—281285—287

[16]钱临照，“牛顿及其《原理》——纪念牛顿《原理》出版三百周年”，《物理》，1987 年 12 月

[17]许良英、范岱年编译，《爱因斯坦文集》第 1 卷，第 1 版，北京，商务印书馆，1976 年 1 月，14—15

[18][美]科恩，“牛顿革命”，《自然科学哲学问题》，1989 年 1 月

二、能量概念的发展及能量守恒定律的发现

能量是物质运动的一种量度，是人们认识客观世界的主要对象之一。19 世纪中期发现的能量守恒定律表明能量是个守恒量，它可以由一种形式转化为另一种形式。能量守恒定律深刻地揭示了各种形式能量的相互联系和自然界的统一性，被恩格斯称为伟大的运动基本定律，19 世纪自然科学三大发现之一。[1]能量守恒定律的发现以及能量概念的形成经历了漫长的历史过程，它是人类在生产实践和科学实验的基础上对自然界的运动转化长期认识的结果。从研究机械能守恒到得出广义的能量守恒定律其间经历了大约一百五十年的孕育时期。

(一)“活力”守恒的发现

从历史上考察，能量原理是从力学留传下来的。意大利物理学家伽利略(Galileo，1564—1642)在 1638 年出版的《关于力学和

图 2-1

局部运动两门新科学的谈话和数学证明》(简称《两门新科学》)中，讨论了自由落体运动和物体沿斜面的运动，提出了这样的假设：静止的物体不论是沿竖直方向自由下落还是沿不同倾斜度的斜面从同一高度下落，它们到达末端时具有相同的速度，这就是“等末速度假设”。[2]伽利略利用一个简单的实验检验了这个假设。摆球沿圆弧运动可看作是沿着一系列不同倾斜度的斜面的下落和上升运动。实验表明：使单摆由一侧开始摆动，当它经过最低点而到另一侧时，会升到几乎相同的高度，如果摆线中途为钉子 E 或 F 等所阻，则摆球将沿新的弧线上升，但仍达到相同的高度。这说明沿不同倾斜度的斜面对于下落速度没有任何影响。[2]物体下降时所得的速度正好等于能够把它送到原来高度的那个速度，一个物体下降的速度只决定于下降的竖直高度而与下降时实际经过的路程的形状无关。伽利略的这个假设为后来揭示重力场的保守性，即在重力场作用下物体的机械能守恒开了先河。

图 2-2

德国数学家、哲学家莱布尼兹(G.W.Leibniz,1646—1716)提出了“活力”概念及“活力”守恒原理。1686 年，莱布尼兹在他的论文《关于笛卡尔和其他人在确定物体的运动力中的错误的简要论证》中提出 mv 不宜作运动的原动力的量度，应把 mv^2 作为原动力的量度。他认为：“力必须由它所产生的效果来衡量，例如用它能将一个重物举起的高度来衡量……而不是用它传给另一物体的速度来衡量。”他把物体的重量和上升的高度的乘积作为运动的力的量度。他说：“我假定将 1 磅重的物体 A 从 D 提升到 4 爱尔(注 1 爱尔(ell)=45 吋)高的 C，其所需的力等于把 4 磅重的物体 B 从 F 提升到 1 爱尔高的 E。这个假定是笛卡儿派所承认的，也是我们时代的其它哲学家与数学家们所承认的。”[3]他根据伽利略的落体定律计算出物体自由下落的高度和它下落此高度所获得的速度的平方成正比，即 $v^2 \propto h$，而物体下落所得的速度正好等于把它送到原来高度的那个速度，所以物体能上升的高度就和这速度的平方成正比。以物体的重量和提升的高度的乘

积作为运动的力的量度时，应以 mv^2 作为运动的力的真正量度。莱布尼兹把 mv^2 叫做活力(visviva)

和笛卡尔一样，莱布尼兹也相信宇宙中运动的总量必须保持不变，不过和笛卡尔不同他认为应该用 mv^2 表示这个量，而不是 mv。莱布尼兹的活力守恒概念在当时的力学现象中也得到了验证。荷兰物理学家惠更斯(ChristianHuygens,1629—1695)在 1703 年作为遗稿发表的论文《论碰撞作用下物体的运动》中对完全弹性碰撞作了详尽的研究。惠更斯写道："在两个物体的碰撞中，它们的质量和速度平方乘积的总和，在碰撞前后保持不变。"[4]这就是完全弹性碰撞中"活力"守恒原理的具体表述。

对非弹性碰撞，动量是守恒的(这对笛卡尔派有利)，但是活力是减少的。莱布尼兹仍坚信活力是守恒的。为了说明在非弹性碰撞中活力并没有减少，他提出了一个巧妙的解释，即认为碰撞物体在整体上所减少的活力并未消失，而只是被物体内部的微小粒子吸收了，微粒的活力增加了。这个思想是深刻的。莱布尼兹当时当然没有现代的分子原子概念，他的这种解释纯属设想，但是却符合了近代分子运动论的观点——碰撞物体整体的动能变成了热能，即内部分子运动的动能。[5]

莱布尼兹还把他的活力守恒思想推广到碰撞以外的问题中。他看到当石头以一定的初速竖直抛上时，它的速度随高度而减少，到达最高点时活力变为零；然后回落，活力又逐渐增大，最后又可恢复原来的活力数值。对上抛过程中活力的显然减少，莱布尼兹仍坚持活力并未消失而是以某种形式被储存起来了，当物体回落时这储存的活力又被释放了出来。由此可以看出，莱布尼兹的活力守恒的信念何等牢固！他的这种活力被储存起来的想法是后来势能概念的先声。[5]

瑞士的数学家约翰·伯努利(JohannBernoulli,1667—1748)对活力概念作出了深刻的说明。他认为非弹性物体类似于在压缩后其膨胀受到限制的弹簧，在物体的压缩中活力减少了。但是在它们的变形中没有消失。他没有把非弹性碰撞中活力的损失归因于热，他直接用机械因素设想这种损失。[6]1738 年丹尼尔·伯努利(DanielBernoulli,1700—1782，约翰的次子)在他的《流体动力学》中，提出了活力的下降和位势的升高等同的原理。他说，用位势提高来代替"活力"的说法对某些科学家更容易接受。他把这一思想用于理想流体运动，得出了著名的伯努利方程，引出了势函数的概念，并认识到可以从势函数引出活力。[4]伯努利指出：一组质点在相互的引力作用下运动时，它们的总的活力的增量只决定于这些质点在开始和终了时的相对位置，而与各质点实际运动的路径无关。当质点组的相对位置复原时，它们的活力也就复原到原来的数值了。[5]约翰的学生，瑞士数学家欧勒(Euler,1707—1783)进一步发展了势函数的概念。他指出一个质点在有心力(质点受的力的方向总指向某一点时，此力叫有心力，此点叫做力点)作用下运动时，当该质点和力心的距离不变时，其活力在任何时候都是相同的。活力的增量只决定于质点在开始和终了时和力心的距离，而与实际的路径形状无关。[5]到 1800 年前后，物理学界已经认识到在一个彼此有相互作用力的系统内，活力仅仅取决于系统的位形和依赖于位形的力函数。

1801 年，英国物理学家托马斯·扬(ThomasYoung,1773—1829)在英国皇家学院的一次演讲中，提出用"能"这个词代替活力。他说："用'能'

这个词来表示物体的质量与速度平方之积是很妥当的。”[7]但“能”这个词并没有被当时科学界所接受。1829 年，科里奥利(Coriolis)虽然用 $1/2mv^2$ 来代替 mv^2，但仍用了活力这个词，为了避免有无 1/2 的混乱，柏兰吉尔建议把 mv^2 叫做活力，而把 $1/2mv^2$ 叫做“活动力”，这种不准确的、具有双重意义的“力”的词一直继续到 19 世纪中叶，直到能量守恒定律确立后，才明确建立了能量概念以及力学中的动能和势能概念。[5]

功的概念起源于早期工业革命中工程师们的需要。他们需要一个用来比较蒸汽机的效率的办法。在实践中大家逐渐同意用机器举起的物体的重量与行程之积来量度机器的输出，并称之为功。在 19 世纪初期用机械功测量活力已引入动力技术著作中。1820 年后，力学论文开始强调功的概念。1829 年，法国工程师彭塞利(J.Poneclet,1788—1867)在一本力学著作中引进“功”这一名词。之后，科里奥利在《论刚体力学及机器作用的计算》一文中，明确地把作用力和受力点沿力的方向的可能位移的乘积叫做“运动的功”。[4]力在一段距离上的总功用力对于距离的积分来测量。当时已经把被称为功的量与原来叫做活力的量联系这一起来，可以明确地写出

$$FdS = \frac{1}{2}mv^2$$

功作为机械能的测量的概念具有重要的意义。功与以后建立的能量概念具有相同的量度，功作为能量变化的量度为研究能量转化过程奠定了一个定量分析的基础。

(二)热的运动说的发展

在通向能量理论结构的途程上，热的运动说的实验证明，对能量守恒定律的建立有密切的关系。热是组成物体的粒子的运动这一学说，使得热和机械功的等效性在概念上是可以理解的，并为机械功和热的相互转化提供了一个解释的基础。在对热的本质的认识上，历史上有热质说与热的运动说之争。古希腊哲学家亚里士多德(Aristotle,公元前 384—322 年)提出宇宙是由土、水、气、火四种元素组成，把火看作四元素之一。随着古希腊原子论思想的复兴，热是某种特殊的物质实体的观点也得到流传。法国科学家和哲学家伽桑狄(Pierre Gassendi,1592—1655)认为，运动着的原子是构成万物的原始的、不可再分的世界要素。同样，热和冷也都是由特殊的“热原子”和“冷原子”引起的。它们非常细致，有球的形状，非常活泼，因而能渗透到一切物体之中。这个观念，把人们引向“热质说”。[4]

热质说的倡导者们称热是由无重量的某种特殊物质组成的。荷兰物理学家波尔哈夫(H.Boerhaave,1668—1738)认为，热的本原是钻在物体细孔中的、具有高度可塑性和贯穿性的物质粒子。它们没有重量，彼此间有排斥性，而且弥漫于宇宙。1789 年，化学家拉瓦锡还将“热质”和“光”列入无机界二十三种“元素”之中。[4]

在热质说观点指引下，热学研究取得了一定的进展。波尔哈夫在作混合物的实验时断言“热不能创造也不能消灭”，提出了混合时热量守恒的思想。1757 年前后，热质论的主要倡导者，英国化学家布莱克(JosephBlack1728—1799)主张把热和温度两个概念区分开来。他在研究热传导时发现，同重量而不同温度的两种物质混合在一起时，它们的温度变

化是不相同的。他把物质在改变相同温度时的热量变化叫做这些物质的‘对热的亲和性”、“接受热的能力”，并由此提出了“比热”概念，引进了“热容量”概念，得出了量热学的基本公式△Q＝Cm△T。布莱克在研究冰和水的混合温度时发现，在冰的溶解中需要一些为温度计觉察不出的热量，进而发现各种物质在发生物态变化时都存这种效应，他由此引进了“潜热”的概念。布莱克在他的化学原理讲义中写道：“大量的热或热物质进入溶冰以后，除了给它以流动性之外，并没有产生其它效应，也没有增加它的温度；这种热好象被溶冰吸收或潜伏在冰水之中，因此，用温度计去量也无法予以发现。”他指出使冰溶解的过程是潜热发生的过程，使水凝固的过程是潜热移出的过程。[8]

热质说对热学的发展又起着严重的阻碍作用。既然把热看成一种物质，而不是物质的一种运动形态，那就不可能有各种物质运动形态的转化。在热质论者看来，摩擦所以生热，只是由于摩擦把热逼出来而已，使摩擦后物体的比热比摩擦前小，所以温度计升高，而热质的量并没有增加。因此，在热质说占统治地位的18世纪，人们就不可能正确理解由蒸汽机的发明所揭示的热和机械运动之间的关系。

关于热的本质的另一种解释则认为热是一种运动。这种热的运动说认为热是组成物体的细小微粒的运动或分子的运动。18世纪前英国的哲学家和学者培根、法国的笛卡尔都持有这种观点。18世纪40年代，俄国学者罗蒙诺索夫在他的“论热与冷的原因”里，认为热现象是分子的转动引起的，对热质说提出了异议。但是，在那时热是运动的观点尚缺乏足够的实验依据，所以还不能形成为科学理论。

18世纪末，由于实验材料的增多，越来越表明热质说不能说明物体因摩擦力作功而生热的现象。1798年英国物理学家本·汤普孙(BenjaminThompson,即伦福德伯爵,1753—1814)向英国皇家学会提出了一个报告《论摩擦激起的热源》，叙述了他的著名实验，这一实验导致了后来具有深远意义的发现。他写道：

“我确信，如果一个人对于生活中的一切日常事务，随时加以注意，这种习惯就往往象人们在最普通事情上突动灵机或戏运遐思那样，可以导致种种有益的疑问和合理的研究与改进的方法，而且比哲学家们在专门研究时所作的一切深思更为有效。”

“出于偶然，我进行了下述实验……我最近在慕尼黑兵工厂车间监督大炮镗孔工作时，对于铜炮在钻腔时能在短时间内获得大量的热，和钻腔机从炮上切下的金属片所获得的更大的热感到十分惊奇(我由实验发现,这温度超过水的温度)。”[9]

这是一个极平常的观察，但由于产生的温度异乎寻常之高和伦福德的特殊的观察天赋而变得引人注目。他敏锐地感觉到彻底研究这一课题，对于热的本质可望获得进一步认识，从而对于热质存在与否这个自古以来哲学家们众说不一的问题作出合理的推测。

接着，他提出了大量热是从哪里来的问题。他说：“热是否来自钻腔机所切开的金属片？如果情形确是这样的话，那么，根据现代的潜热和热质学说，则金属片的热容量不仅应该变化，而且此变化还应该大到足以成为产生所有热的源泉。”但是，通过实验他比较了钻孔前后金属和碎屑的比热，发现钻削不会改变金属的比热。他还用很钝的钻头钻炮筒，半小时

后炮筒升高 70°F，金属碎屑只有 54 克，相当于炮筒质量的 1/948，这一小部分碎屑能够放出这么大的“潜热”吗？于是，他作出结论：“这些实验所产生的热，或者宁可说所激发的热，不是来自金属的潜热或综合热质。”他在论文的末尾写道：

“看来在这些实验中，由摩擦产生热的源泉是不可穷尽的。不待说，任何与外界隔绝的物体或物体系，能够无限制地提供出来的东西，决不可能是具体的物质实体：在我看来，在这些实验中被激出来的热，除了把它看作是‘运动’以外，似乎很难把它看作为其它任何东西。”[9]

1799 年，英国化学家戴维(HumphreyDavy，1778—1829)在《论热、光和光的复合》的论文里描述了这样的实验：在一个同周围环境隔离开来的真空容器里，利用钟表机件使里面的 29°F 的两块冰互相摩擦而溶解为水。他在论文中写道：“如果热是一种物质的话，它一定是从这几种方式之一产生的。或者是由于冰的热容量减少；或者是两物体的氧化；或者是从周围的物体吸引了热质。”可是明显的事实是，水的热容量比冰的热容量大得多，而冰一定要加上一个绝对量的热才能变成水，所以摩擦并没有减少冰的热容量。”“也不是由于热体的氧化引起的，因为冰根本不能吸引氧气”。最后，他得出结论：“既然这些实验表明，这几种方式不能产生热质，那么，热就不能当作物质。所以，热质是不存在的。”他明确指出热是物体微粒的运动。他说：物体既因摩擦而膨，则很明显，它们的微粒一定会运动或相互分离。既然物体微粒的运动或振动是摩擦和撞击必然产生的结果，那么，我们就可以作出合理的结论：热是物体微粒的运动或振动。”[10]

伦福德和戴维的实验与论证是令人信服的，可以说为以后热质说的最终崩溃和热的运动说的确立提供了最早的论据。但是，尽管如此，热质说在当时并没有被推翻，他们的实验没有被人们所重视，大多数学者并没有因此而改变自己关于热的本性的观点。这个问题一直到 19 世纪热力学第一定律问世时，才真正得到了解决。

(三)各种运动形式相互转化的发现

从 18 世纪末到 19 世纪前半期，包括物理学在内的自然科学进入到一个蓬勃发展的新时期。自然科学方面完成的一系列重大发现，日益揭示出各种运动形式之间的普遍联系和转化，这成为这一时期自然科学发展的一个显著特点。

各种自然现象之间的联系和转化的发现，使人们逐渐形成能量概念，认识到各种运动形式是同一种能量的不同表现形式。能量概念给物理科学提供了一种新的统一的框架。能量守恒定律就是在对力、热、光、电、化学各种运动形式相互联系认识的基础上建立起来的。

1780 年意大利人伽代尼(A.Galvani)发现了动物电和电流，接着 1800 年，意大利人伏打(Volta)发明了电池，这是化学运动向电运动的转化。不久之后，人们就利用电流进行电解，实现了电运动向化学运动的转化。

1821 年，德国的塞贝克(Seebeck)发现温差电现象，这是热向电的转化。13 年后，法国的帕尔帖(Peltier)发现了它的逆效应，实现了电流向热的转化。

1820 年奥斯特(Oersted)发现了电流的磁效应。1831 年法拉第

(Faraday)发现了电磁感应现象。他们的发现揭示了电和磁之间的相互联系。法拉第还发明了第一台直流发电机，实现了机械运动向电磁运动的转化。1845 年法拉第又发现了磁致旋光现象，揭示了电、磁、光三者之间的联系。

法拉第对自然界“力”的守恒性和统一性具有坚强的信念。他始终不渝地为证实各种现象之间的普遍联系而努力。他在关于偏振面的磁致偏转一文的开始就明确提出：“我早已持有一种见解，它几乎到达深信不疑的程度，而且我想这也是其他许多自然科学的追求者所持有的见解，即物质之力所表现出来的各种形式具有共同的起源，换言之，它们彼此是如此之互相依赖，以致于它们能够相互转化并具有力的当量。”他举出了电磁感应、电流的磁效应，电流的化学效应、电池的化学论思想，电化当量等为自己的观点提供例证。此外，他还考察过由不同的来源产生的各种形式的电的同一性。法拉第所作的每一个新发现，都使他坚信自然界不同运动形式之间有着无限多样的联系的观点的正确性。法拉第指出，在当时的认识水平上，要想弄清电磁效应、引力效应、光学效应等这些过程的具体机制，是不可能的，但是无论如何他都坚信这种多种现象的统一性，坚信世界的物理图景的统一性。他还考虑过在引力和电力之间的可能联系。他写道：“有一个古老而不可改变的信念，即自然界的一切力都彼此有关，有共同的起源，或者是同一基本力的不同表现形式。这种信念常常使我想到在实验上证明重力和电力之间的联系的可能性。”[4]

各种运动形式相互联系和相互转化的发现使科学家们预感到存在着一种“能”(或“力”)，这种“能”按照各种情况以机械能、化学能、电能、光能、热能或者磁能的形式出现，所有这些运动形式的任何一种都可以转化为其它一种。人们期望找到这种“能”的一个共同量度，使这种“能”成为概念上能够把握的东西。这就要求确定不同形式能量的数值当量或换算因子。

(四)能量守恒定律的发现

长期的科学实验和理论研究为能量守恒定律的最后确立奠定了基础。到了 19 世纪 40 年代，欧洲科学思想中已经普遍蕴含着一种气氛，以一种联系的观点去观察自然。正是在这种情况下，不同国家的数位科学家，同时从几个不同的途径，各自独立地发现了能量守恒定律，并算出不同形式能量的换算因子。其中迈耶、焦耳和亥姆霍兹作出了最杰出的贡献。

1．迈耶的贡献

罗伯特·迈耶(Robert,Mayer,1814—1878)是一位德国医生。他从事能量守恒和转化问题的研究是从对生理现象的分析开始的。1840 年他作为一名医生，在一次乘海船赴印度尼西亚爪哇的航行中，在赤道附近给船员作放血治疗时，他发现水手静脉流出的血好象动脉血一样鲜红，以致误认为切错了血管。随后，证明流出的血确实是静脉血。但回到欧洲后发现静脉血又变成暗红色。由此迈耶根据燃烧学说为基础的动物热理论，想到这些现象可能是由于人体与环境之间的温度差不同的原故。热带环境温度高，维持体温所消耗的动脉血中的氧较少，静脉血管中红色球还带有大量剩余的氧成分所以静脉血要红些。他把有机体的这种化学过程和无机的物理现象联系起来，从而产生了热和机械运动有一定当量关系的思想。[12]

1842 年迈耶发表了《论无机界的力》一文，第一次提出了力(即现在所说的能量)的不灭性和可转化性的原理,并初步计算了热功当量。他从“无不生有，有不变无”和“原因等于结果”等哲学观念出发，表达了他对物理、化学过程中的“力”的守恒和转化的思想。他这样写道：

“力”是原因：因此，我们完全可应用原因等于结果这条原理来表明它们的关系。如果原因 c 有结果 e，那么 e=c，……这个方程的性质清楚表明，在一连串的原因和结果中，某一项或某一项的一部分决不会化为零。这是所有原因的第一个性质，我们称之为不可灭性……如果在结果 e 产生之后，原因 c 仍完整或部分地保存，那么必有另一些结果[f,g…]对应于剩余的原因。因此，由 c 变为 e，而 e 变为 f，如此等等，我们必须把这些不同的量值视为同一实体出现时采用的不同形式。采取不同形式的能力是所有原因的第二个基本性质。同时采用这两个性质，我们可以说，原因在数量上是不可灭的，在性质上是一些可转化的实体……所以，“力”是不可灭的可转化的实体。[13]

迈耶分析了一系列实验事实，如两块金属相互摩擦产生了热；用机械“力”使一定体积的空气压缩，运动的结果使其温度升高；一个重物从高处落到地面，碰撞过程中产生热；反过来，在蒸汽机中蒸汽失去部分热结果产生了机械运动。迈耶把他的“力”不灭和可转化的学说应用于以上例子。他继续写道：

“在无数的例子中，我们看到运动在没有引起另一种运动或没有举起重物时便停止了；但是，‘力’一旦存在就不能消灭，它只能变更其形式，于是就产生了这样的问题：力——我们所熟悉的‘落体力’和‘运动力’(即现在所说的势能和动能)——会采取的另一些形式是什么形式呢？只有经验能引导我们得出结论。”[13]

迈耶作出如下推论：如果热是由两个表面相互摩擦这样的运动产生的，则热必定与运动等价。迈耶指出：“如果落体力和运动力等价于热，那么热自然要等价于落体力和运动力。正如热作为体积减小和运动停止的一种效应而产生那样，热作为它的效应——运动、体积膨胀、重物升高——的一种原因而消失。”[14]迈耶根据“原因等于结果”的原理推得以上的命题。他又提出“相当于动能或势能的热量究竟是多大”的问题。第一次提出了热功当量的概念。他并根据落体的势能全部转化为热作了初步估计：“重物从大约 365 米的高度下落(所产生的效应)对应给相同质量的水从 0℃到 1℃所加的热。”但是在这篇文章中他并未讲到如何获得这个值。[14]

1845 年，迈耶在《论有机运动与新陈代谢》一文中，具体地进行了热功当量的计算，并且由空气的定压比热与定体比热之差算出热功当量。当时测得在定体下 1 克空气温度升高 1℃所需热量为 0.188 卡，即定体比热 C_v=0.188 卡/g·℃。在 1 个大气压下气体膨胀时温度升高 1℃所需的热量为 0.267 卡，即定压比热 C_p=0.267 卡/g·℃。1 克气体在温度升高 1℃时，定压过程中吸收的热量与定体过程中吸收的热量之差为

$$\triangle Q=(C_p-C_v)1g1℃=0.079 \text{ 卡}$$

迈耶假定在定压下温度升高 1℃所需的热量也是 0.188 卡，而另外的 0.079 卡，则是气体膨胀推动活塞而做的机械功。气体在定压膨胀时温度

增加1℃体积约增大$\frac{1}{273}$，1克空气的体积约增大2.84厘米^3，所以在这个过程中气体对外作的功为

$$A=p\triangle V=1.0130\times10^5N/m^2\times2.84\times10^{-6}m^3$$
$$=0.288J$$

于是可得出热功当量为

$$J=\frac{A}{\Delta Q}=3.65J/\text{卡}$$

这个计算公式就是今天我们称为的迈耶公式

$$C_p-C_v=R[4]$$

普朗克在谈到迈耶工作的特点时指出：“按照他的整个思想倾向，他更爱好做哲学上的概括，而不太愿意一步一步从经验上建立起他的理论。”[15]

由于迈耶的工作带有哲学思辨的色彩，开始时未能得到学术界的支持。他在推理时所用的那种哲学风格，对那些习惯于用牛顿力学计算和定量实验来定义科学的人们来说，是无法使人信服的。况且迈耶的新思想所解决的那个概念危机并不是广泛承认的：他正在阐述的是一种可以把过去物理科学几个独立分支统一起来的新观点，可是这种观点的成效在当时还不明显。大约有 10 年之久，迈耶的工作几乎没有受到注意。他受到了激烈的反对，粗暴的中伤，一直在逆境中奋斗。

对于迈耶的研究思想，德国物理学家劳厄作了全面的评价。他在《物理学史》中指出：“不管人们对迈耶的推论采取怎样的态度，无论如何都应当承认：因为物理学的任务是要发现普遍的自然规律，而且又因为这样的规律性的最简单的形式之一是它表示了某种物理量的不变性，所以对于守恒量的寻求不仅是合理的，而且也是极为重要的研究方向，在物理学中也要经常指出这种方向，基本上，人们把能量守恒的早期信念归功于他。当然，只有经验才能真正决定一个被当作守恒的量是否实际上是不变的。就象电的守恒定律一样，能量定律也是一个经验定律。不管怎样，当迈耶计算热的机械当量时，他实际上走的是经验的道路。”[15]

恩格斯对迈耶的工作给以极高的评价。他说：“运动的量的不变性已经被笛卡儿指出了。……但是，运动形态的转化直到 1842 年才发现出来，而且新的东西正是这一点，而不是量方面不变的定律。”[1]这里所指的就是 1842 年迈耶的论文。

2．焦耳的贡献

焦耳(JamesPrescottJoule,1818—1889)关于热功当量的测定是确定能量守恒定律的实验基础。1840 年至 1841 年间，他测量了电流通过电阻放出的热量，总结出了《论伏打电所生的热》等两篇论文。他发现导体在一定时间内放出的热量同电路的电阻以及电流强度平方之积成正比——这就是焦耳定律。但是焦耳明白，这个实验不能对热的来源作出判断。

1843 年，焦耳想到由磁电式机器(即发电机)通过感应而产生的电流应该与来自其它电源的电流一样产生热效应。在《论磁电的热效应和热的机械值》一文中，他写道：

“当我们不把热看作一种实物，而是看作一种振动状态时，没有理由认为它为什么不能由一种单纯的机械性质的作用所引起，例如象一个线圈

在一个永磁体的磁极前转动的种种作用。”[4]

于是他设计了这样一个实验，使一个绕在铁芯上的小线圈在一电磁体的两极间转动，用一个电流计测量线圈中的感应电流，把线圈放进一个盛水的量热器里以测定水温升高所获得的热量。实验得出这样的结果：“磁电机的线圈所生成的热量正比于电流的平方。”

为了确定转动线圈所消耗的机械功与生成的热量之间的数量关系，就需要同时确定转动该装置所需要的机械能。焦耳采用的办法就是在带动线圈的轴上绕两条细线。跨过两个滑轮挂以适当重量的砝码，由砝码的重量及下落的距离计算出所作的功。他总共作了 13 次实验测得热功当量大约为 460kg·m/千卡。在文章的最后，他得出这样的结论："由于创世主的意旨，自然界的全部动因是不灭的；因此有多少机械能被消耗掉，就有完全等量的热被得到。”[16]

1850 年，焦耳在他写的《论热功当量》一文中，总结分析了以往工作的经验，并介绍了现在物理教科书上所介绍的那个经典实验：在金属量热器内装上带有桨叶的轴，叶片分布在彼此成 45°角的垂直平面内(共八列)，在侧壁上装有呈放射状的四列平板，以阻止水的运动。在轴的外端的木圆柱上绕上绳子，绳子通过定滑轮吊起重物，重物下落的高度由标尺测量。焦耳测得热功当量值为 425kg·m/千卡或 4.16J/卡，与现在采用的热功当量值 427kg·m/千卡或 4.18J/卡，差 0.5%。在得到了上述热功当量数值后，焦耳作了如下总结。他说：“由物体(不论固体或液体)的摩擦所产生的热量与消耗的功之量成正比。”焦耳的研究是以测量、严格的实验与观察和对自然界能量不灭的执着信念为基础的。他还力求从理论上说明功热互相转化的可能性，用他的话说：“热和机械功的联系体现了热作为物体粒子运动模型的理论。”他在 1848 年写的《关于热与弹性流体之组成的意见》一文中指出：

“由这些实验所得出的明显结论是，热与机械力是可以互相转化的，因此很显然，热或是可称衡的粒子的活力，或是能够产生活力的一种吸引或排斥的状态。关于空气的胀缩引起温度改变的一些实验同样揭示出弹性流体的组成内幕，因为这些实验表明弹性流体的热是它们所具有的机械力。又由于已知气体的温度决定着它的弹力，所以弹力(或压强)必定是组成任何气体的那些粒子运动的效果。”[4]

焦耳的工作一开始也没有得到人们的重视。1844 年英国皇家学会拒绝他宣读论文。1847 年英国科学促进会只让他报告实验提要，本来对他的发言不准备讨论的，但汤姆孙(W.Thomson)却提出了问题，指出焦耳的研究结果是同法国工程师们所建立的热机理论相矛盾的，因为后者是以热质说为出发点的。这个提问引起了听众对新理论的强烈兴趣，从而使焦耳的论文受到了科学界的注意。

焦耳一生专心致志于热功当量实验几乎到了着迷的程度，给当时的科学家留下了深刻的印象。我们从汤姆孙在 1893 年讲的一段话可以看到这一点：

“我永远不会忘记，英国科学促进会 1847 年在牛津举行的会议。在一个小组会上，我聆听了一个非常谦虚的青年人宣读了一篇论文。他并没有流露出他有一个伟大理论要发表的神气。他的论文深深地打动了我。最初，我想那可能不正确，因为它与卡诺理论不同，当他宣读完毕，我立即同作

者焦耳谈了几句话，这就是我们四十年相识与友谊的开始。就在当天晚上，在英国科学促进会这个可贵的学会上，我们有幸谈到和讨论他和我所知道的有关热力学的一切。我听到自己以前从未想过的一些意见，我想我也提出了值得焦耳考虑的东西，例如我告诉他卡诺理论。我们那时在牛津拉德克利夫图书馆(RadeliffleLibraryOxford)那里分手，我可以肯定，我们俩人都感到彼此都有很多话要说，这天晚上的讨论有很多事要回顾。然而使我惊奇的是，在两星期以后，当我走下契忙里克斯(Chamounlx)山谷时，我远远看见一个青年人沿山路迎面向我走来，他手里拿着一条象手杖一样的东西，但是他既不把它用作铁头登山杖，也不用作行走手杖。他就是焦耳，手里拿着一支长温度计，他不放心把它放在跟在他后面慢慢往上爬的游览车上，怕把它碰断，尽管他的新娘——以后他工作的同情伴侣及参与者——安详又舒适地坐在游览车中。在那次 A 小组会上或拉德克利夫图书馆里，他都没有吐露三天后他要结婚，但是现在，在契忙里克斯山谷，他向我介绍了他年轻的妻子。我们约好两星期之后在马秩尼(Martigny)相会，共同用此温度计测量色朗契斯(Sallanches)瀑布的温度。以后我们便经常会面，并从此成为密友，一直到他去世。许多年来我兴高彩烈，心满意足，从四十年前开始，我与焦耳共同实验，取得了关于热力学理论的一些重要的成果。这是我一生最值得回忆的之一，这对一个有志于科学的人同样是我能想起的最有价值的回忆。”[19]

热功当量的测量，即使在焦耳的蜜月时间也吸引着他，这种对验证科学原理的近似着迷及坚定不移就是焦耳的性格特征。

3．亥姆霍兹的贡献

海尔曼·亥姆霍兹(Hermann Helmholtz,1821—1894)被劳厄称为是一位“能够充分发展能量守恒定律的普遍意义的具有全面才能的人。”[15]他是医生出身，后来成为生理学家、物理学家和数学家。他是从生理学问题开始对能量守恒定律进行研究的，他在大学学医时就对当时流行的在活的生物体中存在着生命力的看法表示怀疑。在他七十岁寿辰时发表的自传式谈话中，他回忆道：

“我怀疑在这种解释中有某些违背自然之处，我花费了大量的努力，以把我的怀疑以准确的问题的形式表述出来。最后，在我大学生生活的最后一年，我认识到施培尔的学说对每一个生物体都赋予了永动机的性质。我对于这一问题的争论是相当熟悉的，在我中学时代就曾听到我父亲和我们的数学老师讨论这个问题。并且当我还是威尔亥姆学院的学生时，我就在图书馆里阅读过姆·伯努利、达朗贝尔以及前一世纪其它一些数学家的著作。于是，我提出了这样的问题：如果永动机是不可能的话，那么在自然界的不同力之间应该存在什么样的关系呢？而且，这些关系是否真正存在呢？”[4]

亥姆霍兹反对“生命力”的说法，主张生物的一切活力都来源于化学力或食物的燃烧。永动机之不可能实现的事实，成为他探索支配自然界的各种运动的普遍规律的出发点。1847 年 7 月 23 日亥姆霍兹在柏林物理学会上宣读了论文《论力的守恒》。他论述了他的能量守恒与转化方面的基本思想。他提出自然科学的基本任务“就在于把一切自然现象归结为其大小依赖于距离的、不变的引力和斥力。”他把力的方向沿着质点连线而力的大小只决定于距离的力称为中心力，把中心力看作机械运动的最后的原

因，并证明了这种力的保守性。[17]他说：

“在质点之间的引力和斥力作用下的质点的所有运动中，如果引力和斥力的强度只与距离有关，那末张力(即势能)在量上的损失始终等于活力$\left(\text{即动能}\frac{1}{2}mv^2\right)$的增加，反之张力的增量始终等于活力的损失。因而，所有活力和张力之和始终是一个常数。这条具有最普遍形式的定律，可以称为力(即能力)的守恒定律。”[4]

为了与功的量等同，亥姆霍兹把$\frac{1}{2}mv^2$称为“活力”。设质量为m的质点，在中心力 F 作用下从距离 r 运动到距离 R 的位置，速度由 v_0 变到v，则“活力”的增量等于“张力”$\left(\text{即}\int_r^R Fdr\right)$的损失，即

$$\frac{1}{2}mv^2-\frac{1}{2}mv^2=-\int_r^R Fdr$$

亥姆霍兹在这里表述了在中心力作用下的机械能守恒定律。接着他把这一定律应用于对其它物理过程的分析，充分地阐明了能量守恒定律的普遍意义。他指出：无摩擦下的力学过程，引力作用下的运动，不可压缩的液体、气体和理想弹性体的运动都服从上述规律。在分析光的干涉时指出，干涉花样中的明暗现象，并不表示能量的消失，而只是能量的重新分布。他证明了热质说的毫无根据，推演出对热的本性的解释。他写道：

“热的数量可以通过机械力来使之增加，因此，热现象不可能用一种物质推论出来，不可能以这种物质的存在为条件，而只能从某些熟知的有质体的变化和运动中，或者是从无质体如电或以太的变化或运动中推论出来。根据这个说法，到现在为止被称为热的量的那种东西，一部分是指热运动活力的量，另一部分是指原子之间张力的量，这些张力在原子的排列发生变化时能够引起热运动。第一部分相当于称之为自由热的部分，第二部分相当于称之为潜热的部分”。[4]

亥姆霍兹还研究了能量守恒定律在电磁现象方面的应用，给出了静电力作功时电荷系动能改变的关系式，确定了静电势的概念。他还分析了电池中的化学-电的作用。他用能量守恒定律来反对所谓“接触说”。他指出接触电势差不与接触表面的大小和形状有关，而只与两种相接触的金属的性质有关。

在论文的最后，他说明了能量守恒定律的意义。他说：“从上述内容可以证明，这一定律与自然科学中任何一个已知现象都不矛盾，而大量的现象倒很明显地证实了它。我力求更完全地叙述出这个定律同至今已知的自然规律相结合得出的结论。它们还需要加以实验的验证。我希望，目前研究的假设部分将使我冷静下来，而这种研究的目的在于以尽可能充分的理由向物理学家们表明这个定律在理论上、实践上重要的启发意义，而这个定律的完全证实将是不远的未来物理学家的基本任务之一。”[4]

亥姆霍兹 1847 年的论述没有立即得到普遍的同意,他的年长的同事们担心这些论述中包含有复活黑格尔自然哲学中的空想的成份而拒绝发表他的文章。后来，他不得不以小册子的形式在柏林单独出版了这篇论文，只有数学家雅可比对他的论文感到兴趣，直到 1860 年左右，能量守恒定律才得到普遍承认和更广泛的应用。

值得一提的是，亥姆霍兹的这一工作无疑是从理论上对能量守恒定律所作出的重要概括，他基本上是独立地作出这一发现的。但是，亥姆霍兹决没有去要求这一发现的优先权。后来，当他了解到迈耶在他之前已发表的两篇论述之后，公正地说：“我们必须承认，迈耶不依赖于别人而独立地发现了这个想法，而这个思想使自然科学获得了长足的新进展。”他还谦逊地说：“和焦耳的工作相比，在那时就已经谈不上要为我提出什么优先权的要求了。”

(五)能量守恒定律的确立和正确表述

在能量守恒定律已经建立起来后，人们还是用“力的守恒”定律来表述它，用具有双重意义的词“力”来表示能量。1851 年，汤姆孙等人大力提倡用清晰的能量概念来代替含混不清的“力”的概念，这才逐渐把能量和“力”两个概念区别开来。1853 年汤姆孙给能量概念作了一个精确的定义：

“我们把给定状态中的物质系统的能量表示为：当它从这个给定状态无论以什么方式过渡到任意一个固定的零态时，在系统外所产生的用机械功单位来量度的各种作用的总和。”[4]

1853 年格拉斯哥的力学教授兰金(Rankine,1820—1872)先把“力”的守恒表述为能量守恒。他指出能量概念是从确立“在宇宙中所有不同种类的物理能是相互转换的”这个定律中得出的。能量是一个守恒量。而力的概念是由牛顿运动定律所定义的，力不是守恒量，对于万有引力来说，力是与距离有关的一个物理量。他强调能量概念的普遍性和统一作用，他提出了用机械能、热能、光能、电能、磁能等名称来区别能量的各种形式。他还提出了现实的能和潜在的能的概念，后者就是我们称为的势能。汤姆孙又把“现实的能”称为动能。

但是，这一定律的发现者们只是着重从量上去表述能量守恒，而没有从质上去强调运动的不灭性。恩格斯指出了这种表述的不完善性，他认为运动的不灭不能仅仅从数量上去把握，而且还须从质的转化上去理解。他说：

“如果说，新发现的，伟大的运动基本定律，10 年前还仅仅概括为能量守恒定律，仅仅概括为运动不生不灭这种表述，就是说：仅仅从量的方面概括它，那么，这种狭隘的、消极的表述日益被那种关于能的转化的积极的表述所代替，在这里过程的质的内容第一次获得了自己的权利……”[1]

能量守恒定律是 18 世纪中叶物理学上最重要的发现。它揭示了自然界各种运动状态的普遍性联系和统一性，找到了各种现象的一种公共量度——能量。它是自伽利略和牛顿之后朝向科学的统一迈进的很重要的一步，并成为全部自然科学的基石，成为检验一个理论是否正确的基本准绳之一，成为辩证自然观重要的自然科学基础。

参考文献

[1]《马克思恩格斯选集》，第 1 版，第四卷，人民出版社，1972 年 5 月，241

[2][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》北京，第 1 版，商务印书馆，1986 年 5 月，4—10
[3]同[2]，58—62
[4]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，318，348，351，411，453—470，570
[5]张三慧，《从伽利略到牛顿》，第 1 版，北京出版社，1988 年 1 月，89，92，96
[6]P.M.Harman,Energy,Force,and Matter:The ConceptualDevelopment of Nineteenth-Century Physics,New York,CambridgeU.P.1982 年，35—41
[7]同[2]，66
[8]同[2]，157
[9]同[2]，166—174
[10]同[2]，175—179
[11]同[2]，236—243
[12]朱尔恭、宋德生，“能量守恒和转化定律的发现”，钱时惕主编，《重大科学发现个例研究》，北京，科学出版社，1988 年
[13][美]G·Holton 著，张大卫等译，《物理科学的概念和理论导论》，上册，第一版，北京，高等教育出版社，1983 年 2 月，405
[14]同[2]，217，218
[15][德]M·V·劳厄著，范岱年、戴念祖译，《物理学史》，商务印书馆，1978 年，80—82
[16]同[2]，220
[17]同[2]230，233
[18]莱尔曼，“能量并非是作功的本领”，《物理教学》，1980 年 2 月
[19]胡南奇，“焦耳的故事”，《物理教学》，1984 年 6 月

三、热力学第二定律的发现

热力学第一定律是热学中的能量守恒定律。热力学第二定律是独立于热力学第一定律的另一条基本定律，它表明自然过程进行的方向性。热力学第二定律是在研究如何提高热机效率的推动下逐步被发现的，是汤姆孙和克劳修斯对卡诺热机理论加似改造，使之建立在焦耳的热功等效和热功转化学说的基础上提出来的。

(一)卡诺的热机理论

热力学第二定律的起源要追溯到卡诺的热机理论。从 19 世纪起，蒸汽机在工业、交通运输中起到愈来愈重要的作用。但是，蒸汽机的效率是很低的，还不到 5%，有 95%以上的热量都没有得到利用。在生产需要的推动下，一大批科学家和工程师开始由理论上来研究热机的效率。

萨迪·卡诺(Sadi Carhot,1796—1832)是法国工程师。他目光敏锐，富有远见卓识。他在 1824 年发表的《关于热的动力的思考》的论文中这样写道：“研究这些蒸汽机是很有意义的，蒸汽机是极为重要的，其用途将不断扩大，而且看来注定要给文明世界带来一场伟大的革命”。[1]在这篇文章的一开始他就指出，与机器所产生的功相比，燃料的消耗太高了；阐明从热中获得动力的条件就能够确立利用热的原理，改进热机的效率。他提出了两个重要的问题：“一个是热机的效率与工作物质有无关系，一个是热机效率是不是有个限度。”

1．卡诺循环的提出

卡诺给自己提出的任务是“从足够普遍的观点”去研究，“由热得到运动的原理。”他说：“为了以最普遍的形式去研究由热得到运动的原理，必须不依赖于任何机械和任何特殊的工作物质，必须使所进行的讨论不仅限于只能应用于蒸汽机，而且还要建立起能应用于一切可以想象的热机的原理，不管他们用的是什么物质，也不管它们如何运转。”[3]卡诺在《关于热的动力的思考》中分析了蒸汽机的基本结构和工作过程，抓住了问题的本质，撇开了各种热机的具体结构及一切次要因素，提出了理想循环过程和卡诺定理。

当时热的运动说还未为人们广泛接受，卡诺也信奉热质说，他认为蒸汽机的工作过程总要伴随着热质的流动和重新分布，还认为蒸汽机和另一种水车是相似的。他说：

“我们可以恰当地把热的动力和一个瀑布的动力相比。……瀑布的动力依赖于它的高度和水量；热的动力依赖于所用的热质的量和我们可以称之为热质的下落高度，即交换热质的物体之间的温度差。”[3]

我们可以把这个类比看作卡诺在建立自己理论时的一个形象依据。这个类比使他得出如下的结论：正如水通过落差而带动水车作功并不改变水的总量一样，在蒸汽机的工作中，热质总量并没有损失。从高温加热器放出的热量和低温冷凝器所接收的热量是相等的。正如他在文中所说的：“在蒸汽机内，动力的产生不是由于热素的实际消耗，而是由热体传到冷体，也就是说重建了平衡。”[2]在这里，卡诺关于热只在机器中重新分配，热量并不消耗的观点是不正确的，他没有认识到热和功转化的内在的本质联系。

但是这个类比也使他得到一个有益的见解：蒸汽机必须工作于高温热源与低温热源之间。他说：“凡是有温度差的地方，就能够产生动力；反之，凡是能够消耗这个动力的地方就能够形成温度差，就可能破坏热质的平衡。”[2]卡诺认为热的动力是由热质在两个热源之间的下落而产生的，两个热源之间具有一定的温度差是热机产生机械功的关键因素。

图 3—1

在建立热机的一般原理时，卡诺提出了一个理想热机模型和一个理想循环过程。卡诺热机被设想为由汽缸、活塞、空气工作物质以及温度较高的热源 A 和温度较低的热源 B 所组成。卡诺设想的理想循环由四个阶段所组成：第一阶段，将热源 A 接触汽缸，汽缸内的空气吸收热源 A 提供的热质，在恒定温度下缓慢膨涨。第二阶段将热源 A 撇开，空气在没有接受热质的情况下继续膨胀，空气的温度降到与热源 B 的温度相等。第三阶段，将热源 B 接触汽缸，空气在恒温下被压缩，把热质输送给热源 B。第四阶段，撇去热源 B，空气继续被压缩，直到回到原来的状态。卡诺指出：“由于空气膨涨时的温度比它被压缩时的温度为高，所以空气膨胀时的弹力较大，因而膨胀时所产生的动力量比压缩时所耗用的动力量大。这样就有动力剩余下来，可以供作其它方面的用途。”[2]这就是我们通常称为的卡诺循环。用现代的术语来表达，这种新的循环包括等温膨胀、绝热膨胀、等温压缩和绝热压缩四个过程。

1834 年，法国工程师克拉珀龙(Clapeyron，1799—1864)提出用压强体积图来表示卡诺循环，他指出等温膨胀中的吸热与等温压缩中的放热是相等的。在一次循环中热是守恒的。所作的功由曲边四边形 ABCD 围出的面积来表示。他把热机的效率定义为循环所作的功与吸收热量之比。

图 3—2

卡诺在《关于热的动力的思考》中还引进了可逆循环的概念。他明确指出：“上述全部操作，可以顺序也可以逆序进行。”如果逆序操作，则逆序操作所耗用的动力等于顺序操作中所产生的动力，从热源 B 还原到热源 A 的热质等于从热源 A 移到热源 B 的热质,最后使动力和热质均无损失。所以，两系列的操作具有相互抵消或相互平衡的意义，因此这一循环就是可逆循环。恩格斯对卡诺抓住主要矛盾进行分析的研究方法给予了很高的评价。他写道：“萨迪·卡诺是第一个认真研究这个问题的人。”“他研究了蒸汽机，分析了它，发现蒸汽机中的基本过程并不是以纯粹的形式出现，而是被各种各样的次要过程掩盖了；于是他撇开了这些对主要过程无关重要的次要情况而设计了一部理想的蒸汽机(或煤油机)。的确，这样一部机器就象几何学上的线或面一样是决不可能制造出来的，但是它按照自己的方式起了象这些数学抽象所起的同样的作用：它表现为纯粹的、独立的、真正的过程。”[1]

2．卡诺定理及其论证

卡诺在讨论了理想循环的基础上，提出了对热机效率所作的理论探讨的核心论点：在工作于两个给定温度之间的所有热机中，以这种理想可逆热机所产生的动力为最大；热动力的产生与所用的工作物质无关，它的量完全决定于两个热源的温度。”[2]因此，一切可逆热机如果在同样温度的

两热源之间工作，就会具有同样的效率，与所用的工作物质无关。这就是现在所称的“卡诺定理”。

卡诺根据热质守恒思想和永动机不可能实现的原理对有关热机效率的卡诺定理作出了证明。他说，如果有一种比他的理想循环在热的利用方面更加有效的方法，“则只需要这个动力的一部分就可以把热质由物体 B 送到物体 A 去，即从冷源送回到热源，于是，起始的状态就得以复原。这样，又可以重新开始类似的操作并如此继续下去，这就不仅是一种永恒的运动，而且将不消耗热质或其它工作物质而无限地创造出动力来。这是跟公认的观念以及力学规律、健全的物理学相矛盾的，这是不允许的。”[3]

卡诺是用文字表述了他的证明，如果用现在的符号和公式则

图 3—3

可说明如下。设有一部任意的热机 C 和一部理想的可逆卡诺热机 D，它们在相同的高温热源 T_1 和低温热源 T_2 之间工作，它们从 T_1 处吸收的热量分别为 Q 和 Q'，卡诺按照热质论认为放给 T_2 处的热量也是分别为 Q 和 Q'。设在一个循环中做的功分别为 A 和 A'，则它们的效率各为

$$\eta=\frac{A}{Q}\ ;\ \eta'=\frac{A'}{Q'}$$

假设这部任意热机 C 的效率大于卡诺热机 D 的功率，即 $\eta>\eta'$，并且使 Q=Q'，则有 A＞A'。这样就可以用 C 来推动 D 使它作逆循环，从而使热量 Q'放回给热源 T_1 处，使各个热源和工作物质都恢复原状，但是 A＞A'，即得到净功 A-A'＞0。

由此可以看到只要在同样的温度之间工作的任意热机的效率大于卡诺热机的效率，就可以在热状态保持不变的情况下源源不断地产生出多余的净功来，这就是所谓的永动机。卡诺认为永动机是不可能实现的，所以任意热机的效率不可能大于可逆卡诺机的效率，可逆卡诺机的效率为最大。从卡诺证明还可以看到在同样温度之间操作的两部工作物质不同的可逆热机具有相同的效率，即热机的效率与工作物质无关。正如卡诺在他的笔记中所作的论断：“热动力不依赖于提供它所用的工作物质，动力的大小唯一地由热质在其间转移的一些物体的温度决定。”[3]

卡诺的论证是建立在错误的热质论观念的基础之上的。他对卡诺定理所给出的那个证明是不能成立的，实际上联合热机并不违反能量守恒定律。它违反的是热力学第二定律。

尽管如此，卡诺定理的两点结论是正确的。他的研究为热机理论的形成和发展做出了开拓性的贡献，为提高热机效率指明了方向。所以恩格斯说：“他差不多已经探究到问题的底蕴，阻碍他完全解决这个问题的，并不是事实材料的不足，而只是一个先入为主的错误理论。”[4]由于热质论的限制，使他在当时不可能认识到热和功转化的内在的本质联系，因而忽视了实际热机中动力可以全部转化为热，而热却不可能完全转化为动力这个普遍存在的实际问题的重大意义。不过，卡诺当时也已经觉察到热质的观点不是完全令人满意的。他说：“我们所提出的基本定律，看来还需要

作进一步的研究；这个定律是根据我们现在的热的理论建立的，应该说，这个基础看来还不是已经没有什么问题了，还需要新的实践的检验。”1830年，卡诺转向了热的运动说，并得到了热功当量值，他在一份遗留的笔记中写道：

“热不是别的什么东西，而是动力，或者可以说，它是改变了形式的运动，它是物体中粒子的一种运动形式。当物体的粒子的动力消失时，必定同时有热产生，其量与粒子消失的动力精确地成正比。相反地，如果热损失了，必定有动力产生。”

“因此，人们可以得出一个普遍命题：在自然界中存在的动力，在量上是不变的。准确地说，它既不会创生也不会消灭；实际上，它只改变了它的形式。”[3]

卡诺未推导而基本上正确地给出了热功当量的数值：370kg · m/千卡。但是他仍然没有触及热转化为功的过程中的热耗问题及热由冷体向热体传递的条件和规律。卡诺过早地死去，他的弟弟看过他的遗稿却不理解这一发现的重要意义，直到 1878 年才把这份遗稿交给法国科学院，这时能量原理早已确立了。

(二)热力学第一定律的提出与热力学第二定律的克劳修斯表述

热力学第二定律的两种表述是克劳修斯和汤姆孙接受了焦耳的热功等效和热功转化的思想后，对卡诺热机理论加以改造，对卡诺思想批判地继承，在论证卡诺定理中提出来的。

对卡诺理论的修正工作最早是由德国物理学家克劳修斯(Rudolf Clausius， 1822—1888)进行的。1850 年，克劳修斯在论文《论热的动力和由此得出的热学定律》中，对卡诺的热机理论作了全面的分析。对于卡诺得出的：“作为产生当量功的热，在由热体向冷体传递时，没有损失热量”[1]的结论，他指出卡诺论点的前一部分是主要的、正确的，在一个循环过程中每当作了功就有热从热物体传到冷物体，但是后一部分，即“没有损失热量”是不对的。他指出：焦耳的实验“明确无误地证明热在任何情况下不仅有增加的可能，而且还证明了一条定律：热量的增加与操作中所耗的功成比例。”[7]克劳修斯又叙述了热是物体粒子的运动组成的观点，热由粒子的“活力”(即动能)来测量，这些运动粒子的“活力”能够转化成机械功。这一观点使得热和机械功的等效性在概念上是可以理解的，为热和功之间的关系提供了一个力学基础。[6]这些事实迫使克劳修斯把卡诺的结论修正如下：“功的产生不仅要求热的分布有所变化，而且确实耗用了热，反过来说，热也可以因功的耗损而再生。”[7]克劳修斯对热机的工作过程作了新的分析，他写道：

“和新的观点相抵触的并非卡诺基本原理本身，而只是那个热并无损失的附加部分。因为功的产生很可能伴随着两种过程，即一些热量被消耗了，另一些热量从热物体传到了冷物体，而这两部分热量和产生的功之间可能存在某一确定的关系。”[3]

克劳修斯的叙述与焦耳的热功等效及热功转化的思想是完全一致的。在克劳修斯看来，提供给系统的热量，一部分转化为功，另一部分变为系统内部的热量。他把这个关系称为热力学的基本原理，并用微分方程的形式表示为

$$dQ=dU+dW$$

这就是我们通常所说的热力学第一定律的数学表示式。式中 dQ 表示传递给物体的热量，dW 表示物体所作的功；U 是克劳修斯首先引进热力学的一个新函数，“它包括增加的自由能和变成内功所耗去的热。”他还指出：“U 的性质有如人们通常假定它为总热量那样,是体积 V 和温度 t 的一个函数，由变化过程的初态和终态完全确定。”后来汤姆孙把这个函数称为物体的能量，即热力学系统的内能。

关于卡诺定理，克劳修斯指出卡诺的结论的正确的，但是卡诺在证明卡诺定理时所依据的热质论是错误的。克劳修斯吸取了卡诺理论中基本的合理的方面，即热总是倾向于从热物体传到冷物体，通常显示出一种使温度差减小的倾向，他把卡诺理论中的这个真正的基本部分与他提出的热力学基本定理相并列，发展成为热力学的第二个基本定理。并用它来论证卡诺定理的正确性。克劳修斯写道：

“如果有两种物质，在一定的热量转移下一种比另一种能够产生较多的功；或者当产生一定的功时，一种比另一种需要从 A 转移到 B 的热量较少，那么我们就可以利用前一种物质通过上述过程来产生功，再将这个功作用于后一种物质使它实现相反的过程。于是，这两部分物质都回到原来的状态时，所作的正功和负功正好抵消，因而根据前一个基本定理(指克劳修斯提出的热力学基本原理)则热量没有增损,不过热量的分布则发生了变化，从 B 移到 A 的热量将大于从 A 移到 B 的热量。因此人们就可能通过交替利用这两个过程，在没有任何其他变化的情况下，把任意多的热量从一个冷物体转移到一个热物体，而这是同热的一般行为相矛盾的，即热总是表现出这样的趋势，它总要从较热的物体转移到较冷的物体而使温度差趋于消失。”[3]

克劳修斯只是用文字表述了他的证明，如果用现在的符号和公式表示，那就是今天物理教科书中所给出的证明。设热源 A 和热源 B 的温度分别为 T_1 和 T_2，任意热机 C 和可逆卡诺机 D 从 T_2 处吸收的热量分别为 Q_1 和 Q_1'，放给 T_2 处的热量分别为 Q_2 和 Q_2'，在一循环中两热机作的功相等 $A=A'$，则它们的效率各为

$$\eta=\frac{A}{Q_1}\qquad\qquad \eta=\frac{A'}{Q'_1}$$

如果任意热机的效率大于卡诺热机的效率,即 $\eta>\eta'$,则 $\frac{A}{Q_1}>\frac{A'}{Q'_1}$,

所以 $Q_1<Q_1'$，按照克劳修斯提出的热力学第一基本原理，因为在一个循环过程中 $\triangle U=0$，以及题设 $A=A'$ 可得

$$Q_1-Q_2=Q_1'-Q_2'$$

则

$$Q_2'-Q_2=Q_1'-Q_1>0$$

把两机看成一复合机，使 C 机推动 D 机并使它逆运转，外界没有对复合机作功，整个系统的热量没有增损，并不违反能量守恒定律，但是却违背克劳修斯提出的热力学第二个基本原理，总有热从较冷的物体传到较热的物体，具有加大温差的倾向。

图 3-4

1854 年，克劳修斯在他的《论机械热理论第二定律的一个改变形式》中提出了更明确的科学表述：“如果不同时引起其它关系的变化，热不可能由冷体传到热体。”1875 年，克劳修斯重新把它表述为“热不可能自动地从冷的物体传到热的物体”；“热从冷的物体传向热的物体不可能无补偿地发生。”这就是热力学第二定律的一种表述。

(三)热力学第二定律的开尔文表述与能量耗散概念的提出

威廉·汤姆孙(William Thomson，即开尔文勋爵，LordKelvin，1824—1907)是英国著名的物理学家，在 19 世纪 40 年代末发表的论文中，他仍然接受在热机产生机械功时热质守恒的观点，对于热的动力研究，他认为卡诺定理是最合适的基础。他写道：“没有进一步的实验研究，这是不可逾越的，而且在其基础上改革热理论有着无数的困难。”[5]1847 年 6 月在英国科学促进会的牛津会议上，汤姆孙见到了焦耳。焦耳宣读了他们测量热功当量实验的论文，汤姆孙对焦耳的结论感到惊讶。他感到卡诺的在热机产生机械功时热量守恒的观点与焦耳的必须消耗与功等量的热的见解存在着明显的分歧，他被这一尖锐的矛盾所困扰。这次会后，他开始研究焦耳关于热和机械功相互转换的论文。

汤姆孙从焦耳实验中，一方面看到热质论及热的守恒假设已经成为问题；另方面也看到焦耳的热转化为功的理论面临一个基本困难，这一困难是由于不可逆的热传导现象引起的。汤姆孙指出当通过固体传热时，没有观察到有机械效应发生。他问道：在通过固体传热有热量消耗时，这个热量变成了什么？

1849 年，汤姆孙已经认识到需要重新修正卡诺理论，把它建立在焦耳实验的基础上。他在《关于卡诺学说的说明》中指出：卡诺关于热在机器中的重新分配、热量并不消耗的观点是不正确的；但是如果抛弃了卡诺关于热产生功的条件的结论，那就会碰到不可克服的困难。汤姆孙的结论是：热的理论需要从根本上进行完全的改革和进行进一步的实验研究。[3]

1850 年，格拉斯哥大学教授兰金观察到来自高压锅炉的蒸汽通过孔口而未液化的现象，他指出必须由外界提供热量来防止蒸汽的液化。汤姆孙认为热是蒸汽通过孔口时由于摩擦而获得的。这个发现支持了焦耳的热是由运动中液体的摩擦产生的理论。因此在 1850 年后期，汤姆孙在陈述热力学理论时观念已经改变，他接受焦耳的热和功相互转化的理论以及热机产生功时有热的消耗的观点。[6]

1851 年汤姆孙发表了《论热的动力理论》等三篇论文。在引言中他说，这篇论文要解决的任务之一，就是要表明“当我们采取同卡诺的假设相反的动力理论的假设时，卡诺以及追随他关于热产生动力的特殊推理方式的人们所达到的那些结论应该如何修改。”他的论文系统地陈述了热的动力理论。他明确提出的以下两个命题构成了全部热之动力理论的基础：

“命题Ⅰ(焦耳)——当等量的机械效应以不论什么方式从纯粹热源产生或在纯粹热效应中丧失时，则有等量的热因之消耗或据此产生。”

“命题Ⅱ(卡诺与克劳修斯)：如果有这样一部机器，当它反过来运转时，它的每一部分的物理的和力学的动作全部逆转过来；那么，它将象具

有相同温度的热源和冷凝器的任何热机一样，由一定量的热产生同样多的机械效应。”[3]

命题Ⅱ就是前面说的卡诺定理。为了证明卡诺定理，他提出了下述公理：

“我们不能从物质的任何部分，用冷却到低于其四周物体最冷温度的方法，借助非生物的媒质来产生机械效应。”[8]

这说明，在某些条件下热转化为功是不可能的。这个公理后来常被叙述为：“从单一热源吸取热量使之完全变为有用的功而不产生其他影响是不可能的。”这个说法强调了两个热源的必要性。王竹溪在《热力学简程》中说：“开尔文的说法相当于说摩擦生热的过程是不可逆的。从开尔文的说法我们马上可以看到，卡诺发现的热机必须工作于两个热源之间的结论具有原则性的意义。”[9]这种原则性的意义也表明了与热力学第一、二定律相对应的第一、二类永动机的不可能性。克劳修斯指出：如果这个公理不成立，就必须承认可以有一种永动机，它借助于使海水或土壤冷却而无限制地得到机械功，这就是所谓的第二种永动机。所以汤姆孙的表述后来也被奥斯特瓦尔德说成“第二类永动机是不可能制成的。”[8]

汤姆孙是这样论证卡诺定理的。他设 A、B 两个热机工作在同样的高温热源和低温热源之间。A 机是任意机，B 机是可逆机，A 机的效率大于 B 机的效率，让 A 机输出的功去推动逆向运转的 B 机，并使 B 机送回到高温热源中的热量等于 A 机从高温热源中吸收的热量。联合热机的总效果是从单一的低温热源中吸热，而不断对外输出功，这是与上述公理相违背的。[8]

汤姆孙指出他们论证所依据的公理与克劳修斯论证所依据的公理，“虽然形式不同，却是互为因果的。至于两个论证的推理方法，则都与卡诺原来的推理类似。”[8]他还把克劳修斯的公理表达为：

“一个没有任何外力为助的自行作用机械不可能将热从一物传递到温度较高的另一物。”

汤姆孙强调这个不可逆现象表明热流的方向性，即热总是倾向于从热物体传到冷物体。在解决他自己所提出的在通过固体传热时，没有观察到机械效应发生这一难题时，他引入了能量耗散(Dissipation of Energy)的概念。在他接受了焦耳的热功等效和热功转化的学说后，汤姆孙断言由物体粒子运动构成的热的力学或动力学理论是焦耳理论的物理学基础，焦耳的热和功的相互转化意味着“热不是物质，而是运动的状态。”于是热的动力学理论断言，在固体的热传导中消耗的热没有失去，这个热转化成物体粒子运动的能量，虽然这个热不能自动复原，它被转换、被耗散，但没有消失。[6]

因热流从高温物体传到低温物体而引起的温度均衡过程，不仅在热机中而且在自然界的所有系统中都总是要发生的；而温度的均衡过程表示做机械功的可能性的丧失。这正是能量“耗散”的含意，一闭合系统中的总能量总是保持不变，但能量趋向于转变成较为无用的形式。[10]汤姆孙认为热力学的两个定律表明了能量的不灭性和耗散性。这两个定律是一致的。因为耗散的能量并未消失，只是未复原。他在论文中指出“物质世界中每个事物是前进的，能量的耗散性表示物理宇宙方向性发展的特征。”对能量不灭，汤姆孙未加以科学的论证，而是求助于神灵的帮助。他在论文的草稿中说：“能量是不变的自然力，除了凭借神力的作用，能量不能

被创造，也不会被消灭。”[6]

汤姆孙提出了能量耗散原理，指出了自然过程发展的方向性特征，进一步发展了能量概念，扩展并深化了对物质世界的认识。汤姆孙表示他的工作是独立进行的。他说：“我的论证是在我知道克劳修斯宣布或论证该命题之前完全依照我当时的想法而提出的。”[8]但是在对待发现的优先权问题上，他表现出了科学的诚实和高尚的风格。汤姆孙说：“首先建立定理的功劳，应属于克劳修斯，因为第一次依据种种正确原理建立该命题的是克劳修斯，他在去年五月已将他的论证在他的热动力论文第二部分中发表了。”[8]他在其他著作中又说：“优先权的问题，对于对科学有兴趣的人来说，在获得深入洞察自然奥秘的任何方面，都是没有意义的。”[5]

参考文献

[1]袁运开、戚越然，“萨迪·卡诺——热力学的奠基者”，《自然杂志》，1983 年 7 月，545—549

[2][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》，237—243

[3]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，469—481

[4]恩格斯著，曹葆华、于光远等译，《自然辩证法》，北京，第 10 版，人民出版社，1960 年 9 月，85

[5]阎康年，“热力学第二定律和热寂说的起源与发展”，《物理》，1986 年 2 月

[6]P.M.Harman,Energy,Force,andMatter;TheConceptualDevelopment of Nineteenth—CenturyPhysics,NewYork,Cam-bridge U.P,1982,45—58

[7]同[2]，244—249

[8]同[2]，258—261

[9]王竹溪，《热力学简程》，第 3 版，北京，人民教育出版社，1978 年 5 月，9

[10][美]G.Holton 著，张大卫译，《物理科学的概念和理论导论》，上册，第一版，北京，高等教育出版社，1983 年 2 月，435

四、统计概念的发展以及麦克斯韦分布的建立

在 19 世纪 40 年代能量守恒定律被发现后，热是分子运动的观点已经得到普遍承认。从 50 年代开始，克劳修斯、麦克斯韦、玻耳兹曼对分子运动论进行了全面的、系统的研究。这些研究深入到物质内部、把唯象热力学和分子运动论结合起来，由系统的微观状态预言系统热运动的宏观性质。他们引进了概率概念、运用统计方法，把系统的宏观状态参量描述为系统相应的微观量的统计平均值。概率统计的研究方法不仅在揭示热现象的本质上起着重要作用，而且在近代科学技术中获得了广泛的应用。

(一)克劳修斯的气体压强公式以及气体分子平均自由程公式的建立

1850 年当克劳修斯建立热力学基本定律时，对于热是物质分子的运动，他已形成了明晰的概念，但是在以后的论文中，他回避了考虑分子运动的性质。这是因为他希望把从建立在实验基础上的热力学原理所导出的结论与预先假定某种分子运动来研究问题加以区别。他希望单独写篇文章来阐述分子运动的性质，直到 1857 年，克劳修斯发表了一篇非常重要的论文《论我们称之为热的运动》才实现了他的这一夙愿。

这篇文章虽然不是第一篇气体运动理论，但是克劳修斯的论文是首次系统地研究了这一课题，现代气体动理论的发展就是从这一论文开始的。他不仅分析了气体分子的平移运动，还指出了旋转运动和振动运动的可能性，这些运动体现了“气体所含的热”，并且“对化学成分复杂(一个气体分子含有大量原子)的气体有特别重要的作用”。在这篇文章中克劳修斯用分子运动论的观点，生动具体地描述了固体、液体的压力，提出了热动平衡的概念。[1]

克劳修斯还提出了理想气体分子运动的微观模型，为了严格遵守由实验得出的玻意耳定律和盖-吕萨克定律，他提出了理想气体必须满足的条件：

(1)分子本身占有的空间相对于气体所充满的整个空间来说是无限小的，即把分子作为数学点来处理。

(2)分子间一次碰撞过程所经历的时间比起相继两次碰撞之间的时间间隔来说是无限小的。

(3)分子力的作用是无限小的。[1]

克劳修斯在建立了理想气体的模型后，着手推导气体的压强公式，他第一次明确地提出了物理学中的统计概念，他指出气体对容器壁的压强是大量分子对容器壁碰撞的平均效果，他写道：

“由于分子的质量很小，每一次单独的碰撞的作用是微不足道的，但是在单位时间内即使在我们目力所能达到的最小的表面元上的碰撞数目也是非常大的，于是在我们的感觉上就产生了一个虚假的印象，似乎器壁所得到的动量不是由于单独的碰撞产生的，而是由于从中心指向四周的不变的力的作用产生的，这就是我们称之为压力的那个力。”[2]

在推导压强公式时，克劳修斯遇到了不能用力学定律解释大量微观粒子运动的问题。对于大量分子组成的系统要确定每个分子的碰撞过程和行为细节是不可能的，也是没有意义的。为了解决已经出现的困难他引进了

统计平均值的概念，以代替对单个分子运动的描述。他说：

“当然，个别分子的速度实际上是千差万别的。但在计算时可以赋予所有分子一定的平均速度。由以下公式可知，为了得到相同的压力，必须这样选择平均速度，使得所有分子的活力在平均速度下就象实际速度时一样。”[3]

克劳修斯在论文中推导了压强公式。他设容器(如图 4－1)的两平行壁面间的距离为 h，分子的平均速度为 u，与器壁法线间的夹角为 θ，N 为总分子数，m 为每个分子的质量。于是，一个分子在单位时间内与某一器壁的碰撞次数为

$$\frac{u\cos\theta}{2h}$$

假设各个方向都是等概率的，则在与某一器壁可能发生碰撞的所有分子中，处于 θ 与 θ+dθ 之间的分子数占总分子数的比例，将是相应的圆环面积与半个球面面积之比(如图 4—2)

$$\frac{2\pi r\sin\theta r d\theta}{2hr^2}=\sin\theta d\theta$$

图 4—1　图 4—2

所以，在 θ 与 θ+dθ 之间碰撞壁面的分子总数为

$$\frac{Nu}{2h}\cos\theta\sin\theta d\theta$$

每次碰撞在垂直于器壁方向的动量变化为 2mucosθ。所以全体分子施乐器壁的力为

$$\begin{aligned}F&=\int_0^{\pi/2}(2mu\cos\theta)\left(\frac{Nu}{2h}\cos\theta\sin\theta d\theta\right)\\&=\frac{Nmu^2}{3h}\end{aligned}$$

若以 a 表示器壁的面积，则容器的容积 V=ah，于是，压强

$$p=\frac{F}{a}=\frac{Nmu^2}{3V}$$

或

$$p=\frac{1}{3}nmu^2$$

n 为单位体积的分子数。上式又可写成

$$\frac{3p}{2}=\frac{1}{2}nmu^2$$

等式右端是分子平移运动的“活力”(即动能)。

按照玻意耳定律和盖-吕萨克定律

$$pV=T\times 常数$$

在这里 T 是绝对温度，V 是单位体积，因此

$$\frac{1}{2}nmu^2=T\times 常数$$

可见，平移运动的“活力”正比于绝对温度。

克劳修斯又根据上述压强公式得出平均速度的公式。设单位体积中气体的重量即比重为ρ，因为 $\rho=nmg$ 所以 $u^2=\frac{3p}{nm}=\frac{3pg}{\rho}$ 由此他得出在标准状态下，几种气体的平均速度(严格地说是方均根速度)。对于氧、氮、氢平均速度分别为：461m/s，492m/s，1844m/s。

克劳修斯的论文引起了一连串的反对意见。特别是拜斯-巴洛特(Buys-Ballot)，他在 1853 年 2 月发表的一篇文章中指出，分子的高速度并不符合所观察到的现象，诸如气体的缓慢扩展，烟雾的缓慢散开。他说：“如果在房间的一个角落里出现了硫化氢或氯气，当在另一角落的人闻到这个气味时，好几分钟过去了，然而气体分子在一秒钟必定飞过这个房间几百次，这个现象如何解释呢？”对于这些反对意见，克劳修斯采取了积极的态度。他在 1858 年发表的一篇题为《关于气体分子运动的平均自由程》的文章中说：“咋一看来，这些异议显现出具有很大的分量，因此，我考虑对它必须加以证明。……的确，我对巴洛特提出的这个问题而感到高兴，因为它给我提供了一个解释我的理论的机会。”[4]

克劳修斯在回答巴洛特提出的问题时，进一步发展了自己的理论。他为了解释气体分子的运动和相互作用引入了更加复杂的气体分子模型，在理论中引进了分子作用范围这一概念。他指出，在不存在化学亲和力的情况下，必须区分两种力。当两个分子接近时最初起作用的是吸引力，这种力在某一距离上很明显，并随距离的缩小而增大，而当这些分子非常接近时，又会出现另一种使它们彼此离开的力。在斥力与引力相互平衡的位置，克劳修斯把这一距离 ρ 定义为分子作用球的半径。当瞄准距离(用现代术语说)大于 ρ 的情况下，就会发生分子路程的弯曲，而在比 ρ 小的瞄准距离下，分子就会相互排斥，克劳修斯把后一情况看作碰撞。克劳修斯引入了一个新的重要的概念，关于平均自由程的概念，它指的是分子的重心运动到另一个分子作用球的平均距离。[4]

在计算平均自由程的时候，克劳修斯指出了平均自由程对分子运动的平均速度的依赖性。他写道：对于我们的研究特别有意义的是所有分子具有一样的平均速度的情况，在这种情况下，只考察平均速度就能够较简单地接受所有分子都以相同速度运动的设想，这样我们就得出以下结果：在两种情况下，当其余分子或者和所考察分子一样以相同速度运动或者静止不动时，平均自由程的比例为(3/4)：1。[4]

在计算平均自由程时，克劳修斯使用了概率思想。他首先假定除了一个运动的分子外，其余分子都静止不动，而这一分子则以平均速度 u 运动。

他假定在包含大量分子的空间中，分子的密度是相等的，分子的排列是均匀的。求运动分子自由通过 x 距离时，没有与其它分子作用球相碰撞的概率是多大？设对应于单位厚度的概率为 a，则对应于厚度为 x 的概率为 a^x。因为 a<1，让我们变换这个表示式，令 $e^{-a}=a$，则 -a=1na，设自由通过 x 距离的概率为 W，则 $W=e^{-ax}$ 对于厚度 x=δ 的薄层，这个概率可写为

$$W_\delta=e^{-a\delta}=1-\alpha\delta\ (1)$$

为了用分子作用球的半径ρ表示上式中的系数，克劳修斯设想分子排列在与运动分子速度垂直的一个个正方形平面上。设不动分子中心之间的平均距离为λ，n 个分子排成的一层的面积为 $n\lambda^2$，而 n 个分子作用球的截面积为 $n\pi\rho^2$，因此，分子作用球所遮盖的面积占总面积的比为 $\pi\rho^2/\lambda^2$。对于厚度为δ的气体层来说，被分子作用球遮盖的面积占总面积的比将为

$$\frac{\pi\rho^2}{\lambda^2}\cdot\frac{\delta}{\lambda}$$

一个分子无碰撞地通过这一层的概率正等于不被其它分子作用球所遮盖的面积占总面积的比值

$$W_\delta=1-\frac{\pi\rho^2}{\lambda^3}\delta \qquad (2)$$

式(1)与式(2)联立得

$$\alpha=\frac{\pi\rho}{\lambda^3}$$

这样，一个分子通过厚度为 x 而不发生碰撞的概率为

$$W=e^{-(\pi\rho^2/\lambda^3)x}$$

而在 x 与 x+dx 之间不发生碰撞的概率为

$$e^{-(\pi\rho^2/\lambda^3)}\left(\frac{\pi\rho^2}{\lambda^3}dx\right)$$

所以就 N 个分子来说，平均自由程为

$$L'=\frac{1}{N}\int_0^\infty Ne^{-(\pi\rho^2/\lambda^3)x}\cdot\frac{\pi\rho^3}{\lambda^3}\cdot xdx$$

$$=\frac{\lambda^3}{\pi\rho^2}$$

如果N个气体分子所充满的空间体积为V，则 $\lambda^3=\frac{V}{N}$，所以

$$L'=\frac{V}{n\pi\rho^2}$$

如果 $n=\frac{N}{V}$ 表示单位体积中的分子数，则

$$L'=\frac{1}{n\pi\rho^2}$$

克劳修斯指出如果考虑到所有分子都以速度 u 运动，那么碰撞的概率将会增大。可以计算出，当分子间的相对速度为 $\frac{3}{4}$u时，平均自由程将减小到原来的 3/4，即为

$$L=\frac{3}{4}\frac{1}{n\pi\rho^2}$$

这样，克劳修斯就初步回答了对分子运动的责难：扩散过程与其说是取决于分子的速度，不如说是取决于两次碰撞之间的平均距离，而这个距离比起分子的速度来是小得多的。

由此看来，克劳修斯在论文中已经使用了概率思想，运用了统计方法，引进了统计平均值。但是这些思想在克劳修斯那里没有得到充分发展，他对气体分子速度的研究还是停留在平均速度的水平上。揭示气体分子速度分布的规律是由麦克斯韦完成的。

(二)概率概念的引进以及麦克斯韦分布的建立

1850 年 6 月在《爱丁堡评论》上发表了英国物理学家和天文学家赫谢尔(Herchel)的长篇论文《论魁泰勒特(Quetlet)关于统计理论的著作》。在这篇论文中赫谢尔把欧洲大陆的统计理论介绍给了英国的科学界。它给麦克斯韦一个强烈的印象，那时他是爱丁堡大学的一个 19 岁的学生，不久，他就进入剑桥大学，在麦克斯韦给他的朋友卡卜贝尔(Campbell)的信中反映了赫谢尔的一些思想。

“这个世界的真正逻辑是或然率的计算……因为人类知识是以这种方式感觉到的，即从不同感觉、理解和根据正确规律的行动的和谐一致(不是类似)的证据所推出的结论,表明外部事物将把不同程度的概率分配给不同的情况(或事实，或证据，或我们所称为的那些东西。)”[5]这段话清楚地说明了应用概率理论的必要性。因为概率分布是客观存在的，所以物理学的任务就是要找出这一规律性。

除此之外，克劳修斯的研究对麦克斯韦形成对气体动理论的兴趣产生了巨大影响，麦克斯韦正是在读了 1859 年 2 月《哲学杂志》上发表的克劳修斯论文的英译文后研究这一理论的。他在 1866 年发表的《论气体的动力理论》中说：“我们把完整的气体动力理论归功于克劳修斯教授，他关于热的动力理论的研究是众所周知的。他的《论我们称之为热的运动》的论文是对分子理论的完满的解释”。接着他讲了在读了克劳修斯著作后所进行的一系列理论研究工作。[6]

麦克斯韦对气体动理论的研究可分为两个阶段。第一阶段最著名的论文是 1859 年 9 月 21 日他在英国科学促进会上宣读的报告《气体动力论的说明》，这篇文章 1860 年发表于哲学杂志上。他在文章一开始就谈到了他的探索的动机。他说：“物质的特别是气体物质的许多性质，都可以从物质的微小部分在迅速运动中其速度随着温度的增加而增加这个假定推导出来，所以这种运动的真正性质就成为理性的好奇心(Rational Curiosity)的课题。”[7]接着他提到了丹·伯努利、克劳修斯等人的工作。针对当时一些人不相信原子论的物理解释的可能性，特别是不相信测定分子大小的

可能性的观点，麦克斯韦卓有远见地指出，虽然现在还无法测定克劳修斯所引入的平均自由程以及分子的有效直径，但是已知的一些现象，如气体的内摩擦、热传导和扩散等，却提出了精确测定平均自由程的可能性。”为了把这一研究置于严格的力学基础上，他将论证“数量很大的、非常微小的、完全弹性的，只在碰撞时才有相互作用的坚硬小球所组成的系统的运动规律。”[7]

麦克斯韦通过讨论完全弹性球的碰撞与运动，得出气体分子经过碰撞后沿各方向运动的概率相等的结论。他认为应当考虑这样一个事实，即任何速度都不可能是被禁止的，分子的速度可以从 0 到∞。他指出气体分子间的频繁碰撞并不使它们的速度趋于一致，而是出现一个不同范围内的某种分布，在平衡态下这种分布不变。如果知道了气体分子按速度的分布，则气体大部分可观察的性质都可以由此计算出来。他在这篇文章中写道：“如果有大量相同的球形粒子在完全弹性的容器中运动，则粒子之间将发生碰撞，每次碰撞都会使速度变化，所以在一定时间后，活力将按某一有规则的定律在粒子中分配，尽管每个粒子的速度在每次碰撞时都要改变，但速度在某些极限值内的粒子的平均数是可以确定的”。

在建立这一分布定律时，麦克斯韦以下述三个假设为出发点：两个弹性球相碰撞时，在一切方向上的反冲都有同等概率；速度的各分量 X、Y、Z 的分布彼此独立；速度分布不受外界影响。

接着他开始确定在大量同类粒子之间很多次碰撞之后，速度在给定范围内的粒子平均数，即速率分布律。

他令 N 为粒子总数，X，Y 和 Z 为每个粒子的速度在三个垂直方向上的速度分量。速度分量 X 值在 X 与 X+dX 之间的粒子数为 Nf(X)dX，f(X)是 X 的待定函数；速度分量 Y 在 Y 与 Y+dY 之间的粒子数为 Nf(Y)dY；速度分量 Z 在 Z 与 Z+dZ 之间的粒子数为 Nf(Z)dZ，这里 f 始终代表同一个函数。

由于三个速度分量彼此垂直并且互相独立，所以速度 X 的存在毫不影响速度 Y 和 Z。于是速度值在 X 与 X+dX，Y 与 Y+dY 以及 Z 与 Z+dZ 之间的粒子数为

$$Nf(X)f(Y)f(Z)dXdYdZ$$ ”

如果假设 N 个粒子在同一时刻由原点出发，则此数将为经过单位时间以后在体积元(dXdYdZ)[麦克斯韦在此引进了速度空间的概念，(dX dY dZ)指的是速度空间内的体积元。]内的粒子数，因此单位体积内的粒子数应是

$$Nf(X)f(Y)f(Z)$$

由于坐标轴方向的选择是完全任意的，所以这个数目必然只依赖于到原点的距离，即

$$f(X)f(Y)f(Z)=\phi(X^2+Y^2+Z^2)$$

解此函数方程，可得

$$f(X)=Ce^{Ax^2}\quad \phi(r^2)=C^3e^{Ar^2}$$
$$(r^2=X^2+Y^2+Z^2)$$

如果取 A 为正数，则当速度增大时，粒子数随之增大，这样，粒子总数将为无穷大。所以取A为负数，并令其等于 $-\frac{1}{a^2}$，则X与X + dX之间的粒子数为

$$NCe^{-(x^2/a^2)}dX$$

从 X=-∞到 X=+∞积分，我们得到粒子总数为

$$\int_{-\infty}^{+\infty} NCe-(x^2/a^2)dX = N$$

求解得 $$NC\sqrt{\pi}a = N$$

即 $$C = \frac{1}{a\sqrt{\pi}}$$

所以 f(x)为

$$\left(\frac{1}{a\sqrt{\pi}}\right)e^{-(x^2/a^2)}$$

这是分速度 X 的分布函数，Y 和 Z 的分布函数与此类似，麦克斯韦进一步得到如下结论：

“第一,速度分解在某一方向上的分量 X 在 X 与 X+dX 之间的粒子数为

$$N\left(\frac{1}{a\sqrt{\pi}}\right)e^{-(x^2/a^2)}dX$$

第二，速率在 v 与 v+dv 之间的粒子数为

$$N\left(\frac{1}{a^3\sqrt{\pi}}\right)u^2e^{-(v^2/a^2)}dv$$

第三，求 v 的平均值，把所有粒子的速率加在一起，除以粒子总数，其结果是

$$\text{平均速率} = \frac{2a}{\sqrt{\pi}}$$

第四，求 v^2 的平均值，把所有粒子的速率的平方相加再除以 N

$$v^2\text{的平均值} = \frac{3}{2}a^2$$

这比平均速率的平方大，正应如此。”[7]

第二个结论所给出的公式即著名的麦克斯韦速率分布律，它与高斯由概率导出的误差分布律很相似。速率分布曲线是不对称的，误差分布曲线是对称的。不难看出 a 就是最概然速率，粒子在这个速率值附近出现的概率最大。

在作了以上推导以后，麦克斯韦作出结论：“由此可见，粒子的速度按照‘最小二乘法’理论中观测值误差的分布。速度范围从 0 到∞，但是具有很大速度的粒子数相当少。”[7]

麦克斯韦论文选集主编尼文(W.D.Niven)对麦克斯韦的这一著作给予了充分的评价。他在序言中说：“虽然这篇论文推导的方式在以后其它文章中被抛弃了，但是这篇论文本身是极其有意思的，它清楚地表明由麦克斯韦提出并由麦克斯韦所解决的理论问题，迄今仍然包含着在他以后的论文中所要处理的大量问题的萌芽。”[8]

麦克斯韦的这一推导起初并未引起人们多少兴趣。当时概率论的方法

基本上还只用于描述社会过程，在物理学中的应用极其有限，多数物理学家把力学方法作为研究一切问题的普适方法，他们对建立在概率论基础上的速度分布律漠然置之。麦克斯韦的这一推导还受到了克劳修斯的批评，也引起了其他物理学家的怀疑，这是因为他假设它们互相独立地分布，麦克斯韦自己也承认“这一假设似乎不大可靠”，难以令人信服，在以后的几年里他继续研究，直到 1866 年麦克斯韦向英国皇家学会提出了一部新的、最重要的和篇幅巨大的分子运动论著作《气体的动力理论》。这篇文章讨论气体输运过程等若干热力学问题。其中有一段是关于在热平衡状态下气体速度分布定律的推导，这一推导不再有“速度的三个分量的分布互相独立”的假设，也得出了上述速度分布律。尼文在他写的序言中说：“在这篇文章中他给予了速度分布律一个新鲜的证明，但是这个方法具有永恒的价值，它在没有在一个点的领域内各个方向上的速度分布相同的假设下，详细完成了论述”。[8]麦克斯韦这一新的推导过程如下：

“从给定点 O 作线段，代表单位体积中任一种分子速度的大小和方向，这些线的终点分布在空间。(指速度空间)任选一体积元 dV(指速度空间内的体积元 $dV=dV_xdV_ydV_z$)，终点在 dV 内的这样的线数为 $f(r)dV$，其中 r 是 dV 到 O 点的距离。

令 OA=a，是互相碰撞前第一类一个分子的速度；OB=b，是第二类中一个分子的速度，如果与分子质量 M_1、M_2 成反比地分割 AB 于 G 点，则 OG 是两分子重心的速度。如图 4-3 所示。

图 4-3

令 OA'=a'及 OB'=b'，为两分子碰撞后的速度。使 GA=GA'及 GB=GB'。A'GB'是一直线，但不一定在 OAB 平面内。∠AGA'=2θ是在碰撞中相对速度转过的角度。

如果我们知道碰撞前的相对速度 BA，在碰撞中 BA 转过的角度为 2θ，以及决定 AB 和 A'B'所在平面的方向的夹角φ，则分子的相对运动就完全确定了。在所有碰撞中，如果使 BA 的大小和方向，θ，φ角都在某一几乎邻近的限度内，则在单位时间内这一类碰撞的数目应是

$$n_1n_2Fde$$

其中 n_1 和 n_2 是所讨论的每种分子的数目，F 是相对速度和角度θ的函数。de 依赖于变动的限度，我们把这个限度以内的碰撞作为同一类。

第一类分子的数目(即代表速度末端止于 A 的体积元 dV 中的线数)应为

$$n_1=f_1(a)dV$$

具有相当于 OB 的速度之第二类分子数为

$$n_2=f_2(b)dV$$

两组分子之间的碰撞数为

$$f_1(a)f_2(b)(dV)^2Fde$$

代表这些分子在碰撞后的速度的线，将终止于等于 dV 的体积元内的 A'和 B'。

与此类似，可以求出原来速度相当于 A'和 B'所描绘的，后来速度相当于 A 和 B 所描绘的、等于 dV 体积元中的分子之间的碰撞数为

$$f_1(a')f_2(b')(dV)^2F'de$$

其中 F'是 B'A'和∠A'GA 的函数，它与 BA 和∠AGA'的函数 F 相同，所以 F 等于 F'。

当速度从 OA 和 OB 变到 OA'和 OB'的分子对数目等于从 OA'和 OB'变到 OA 和 OB 的数目时，就可以得到速度的最终分布，不再因以后的碰撞交换而变动。这就是下列情形

$$f_1(a)f_2(b)=f_1(a')f_2(b')=\text{定值}$$

在 a 和 b 及 a'和 b'之间应满足能量守恒关系。

$$M_1a^2+M_2b^2=M_1a'^2+M_2b'^2$$

这两个方程的解为

$$f_1(a)-C_1e^{-(a2/a2)}$$

和 $$f_2(a)-C_2e^{-(b2/\beta 2)}$$

其中 $$M_1a^2=M_2\beta^2$$

为了定出常数 C_1 和 C_2，可以对 $f_1(a)$和 $f_2(b)$进行积分

$$\iiint C_1 e^{-(\xi^2+\eta^2+\zeta^2)/a2}d\xi d\eta d\zeta$$

令其结果等于 N_1，就可得 C_1 值。所以，如果 N_1 个分子的速度分布是这样，其速度分量在 ξ 到 ξ+dξ，η 到 η+dη 和 ζ 到 ζ+dζ 之间的分子数为

$$dN_1=\left(\frac{N}{\alpha^3\pi^{3/2}}\right)e^{-(\xi^2+\eta^2+\zeta^2)/a^2}d\xi d\eta d\zeta$$

则这一速度分布将不因分子间相互作用，交换速度而改变。”[6]

这个推导以分子的弹性碰撞为出发点，推导的基础是每一种碰撞中往返过程数量相等这一条件(后来称作细致平衡原理)，并无其它任何假设，因而结论是普遍的。

麦克斯韦的推导继续受到克劳修斯的批评。克劳修斯在 1881 年麦克斯韦去世后出版的《热的力学理论》中，用了“气体动理论”一章专门分析麦克斯韦的分布定律，指出“不应把它看作是任何条件下都准确地符合实际的定律”。克劳修斯的理由是：“推导这一定律所依据的假定是：在三个相互垂直的坐标方向的运动分量是相互独立的，而推导的出发点则是考察了固体弹性球的行为，它们满足上述假定。”[3]对推导分布定律的这一批评似乎表明了克劳修斯并不知道麦克斯韦后来的著作，因为麦克斯韦在 1866 年的著作中没有“速度三个分量的分布互相独立的假设”。

麦克斯韦在回答克劳修斯对他的批评时，特别强调方法的实质，而不是它的附加部分。1877 年麦克斯韦在沃森写的《论气体动理论》一书的书评中，对他采用的方法作了详尽描述。他指出“有两种测定复杂物质系统状况的方法，一种是以力学定律为基础的严格的动力学方法；另一种是可以称为统计法的方法，它的基础是类似于用于观察人口涨落的方法”。他自己偏向于统计方法。他对这一方法的特点描述如下：“我们把物体系统按它们的位置、速度或其它特性分组。我们的注意力不是在物体本身，而是任一时刻内属于某一组物体的数目，这个数目当然会由于物体进入或离开这个组而发生变化。我们应当研究发生这种变化的条件，并按照动力学方法跟踪这些物体。但是，过程一旦结束，即物体一旦进入或离开了该组，我们就停止跟踪它。如果它重新出现，我们就把它算作一个新的物体，这

就象博览会的旋转门计算入场观众那样，不管他们做过什么和将做什么，也不管他们先前是否曾通过这个旋转门。”[3]这一段话生动具体地说明了统计方法研究的对象是研究物体在每一组内的概率。

在这篇论文中麦克斯韦还把它的速度分布律用于解释输运过程，得出常压下粘滞系数μ与压强p(或分子数密度n)无关的结论。这个预测与当时人们的认识不符合。人们认为温度不变时，若是压强减小，则分子数密度就成比例减小，双方交换的分子数也减少，因此粘滞系数也应减少，从而对气体动理论加以指责。麦克斯韦为了验证从理论上得出的这个结论，特地做了一系列实验来加以验证。后来，英国物理学家瑞利(Raylelgh，1842—1919)在1890年评论刚出版的麦克斯韦两卷本科学论文集时写道：“在整个科学领域里，没有任何发现能比发现气体粘滞性在任何密度下均不改变更加美好和更加有意义的了。麦克斯韦从理论上预见到，后来又从实验上证明了：在有限空间内振动的物体和所受到的阻滞作用，和在大气压力下是一样的”。迈耶在德国完成了空气的内摩擦系数的实验测定，他在《论气体的内摩擦》中得出结论说：“在密度不断降低时，摩擦系数的改变比密度的改变小得多。因此，麦克斯韦定律在任何情况下都近似正确。”[3]麦克斯韦速度分布律在当时的条件下无法进行直接的实验验证，直到1920年施特恩(O·Stern)发展了分子束方法，才第一次直接得到速度分布律的证据。

麦克斯韦在科学工作中最突出的特点是，他能够把数学思维和物理实验密切结合，他善于大量地应用数学去解释物理现象。他在处理科学问题时是实事求是地从物理概念出发，而不是抽象地从数学符号出发，从方程式到方程式，忽略物理过程的分析；他既充分利用数学工具，又紧紧抓住问题的物理实质；既重视直观，又有严密的逻辑推理；他既进行理论研究，又亲自参加实验，他历时约10年主持了卡文迪许实验室工作。尼文对麦克斯韦曾这样评论过：“创造和发明的才能，对物理科学的热爱和数学处理的本领，都同时地存在于一个人的心灵里，这是罕见的。”

参考文献

[1]Rudolf Clausius, The Nature of the motion which we call Heat, Selected Readings in Physics Kinetic Theory, Vol.1,112—131

[2]申先甲等编著，《物理学史简编》，502—514

[3][苏]M.A.叶里雅舍维奇，T·C·普罗纪柯，“气体分子运动论中的克劳修斯纲领和麦克斯韦纲领”，《科学史译丛》，1987年1月，38—42

[4]Rudolf Clausius, ibid,136—141

[5]Stephen, G.Brush, Statistical Physics and the Atomi Theory of Matter, Princeton, N.T.PrincetonUniv,1983,59

[6]James Clerk Maxwell, On the Dynamical Theory of Ga-ecs, Selected Reading in Physics Kinetic Theory, Vol.2, 27,45—48

[7]James Clerk Maxwell, Illustrations of the Dynamical Theory of Gases,ibid[1],149—155

[8]W.D.Niven, The Scientific Papers of James Clerk Ma-xwell, the first edition, Vol,1. Cambridge Vniversity Press, 1890年, Preface.

13,24
[9]沈慧君，“ 麦克斯韦是怎样推导速度分布律的 ”，《物理》，1986 年 5 月

五、熵的概念的建立和热寂说的起源

热力学第二定律是有关过程进行方向的规律，它指出自然过程的方向性。1850 年克劳修斯开始系统阐述热力学定律的时候，他依据热传导的方向性，即热倾向于从温度较高的物体传到较低的物体，表述了热力学第二定律。他把这一定律认为是经验的概括，不涉及不可逆性概念的应用。然而，在他以后的论文中，他提出变换等效值的概念，熵是作为变换的等效值而提出来的，熵表示了物理过程的方向性的特征，物理过程的方向性用熵增加来表示。[1]

(一)熵的概念的建立

1854 年克劳修斯发表了一篇论文，题为《论热的动力理论的第二原理的另一形式》。他从分析卡诺热机开始，假设热经历两种变换：一种是热从高温物体传到低温物体的传送变换(trans-mission transformation)；另一种是热转化为功的转化变换(conversion transformation)。克劳修斯指出每种变换有两种可能的方向：一种是自然的方向(natural direction)；这种变换能够自发地独立地进行；另一种在非自然的方向上，在没有外界影响的迫使下不可能进行。对于转化变换来说由功转变到热的方向是它的自然方向；而由热转化到功的方向是它的非自然方向。对于传送变换来说，热由高温物体传到低温物体是它的自然方向；而由低温物体传到高温物体是它的非自然方向。克劳修斯看到在热机运转中两种变换同时发生了，传送变换在它的自然方向上发生了，而转化变换在它的非自然方向上发生了。这好象是传送变换推动了在非自然方向上的转化交换，传送变换起支配作用，给在非自然方向上的转化变换提供推动力。如果使热机逆向运转，则转化变换就处在它的自然方向上，推动着在非自然方向上的传送变换，这时转化变换起支配作用。[2]

这幅热力学的图画使克劳修斯受到启发，在他的两种变换中的任一个支配另一个变换的可逆循环中，两种变换几乎是均衡的。在某种意义上它们是彼此等效的，克劳修斯着手按照这条途径建立一个定量的变换理论。他的目的是确定两种变换的等效值。他希望这个等效值能够以新的自然规律表示这个均衡条件或者象他称为的补偿(compensation)。[2]

克劳修斯假设对于任意变换的等效值正比于热量 Q 和某个或某些温度函数 f(t)或 $f(t_1, t_2)$的乘积 Qf(t)或 $Qf(t_1, t_2)$。并假定同一变换在自然方向上和非自然方向的等效值大小相等、符号相反。规定在自然方向上的等效值为正，在非自然方向上的等效值为负。并令在一个可逆循环中，两个等效值的和为零。在这些条件下，又依据热力学第一定律和理想气体状态方程得出如下结论:在循环中发生的所有变换的等效值是积分$\oint \frac{dQ}{T}$。dQ 是热量的微小变化，T 是绝对温度。[2]

热的变换的等效值的概念对可逆与不可逆过程之间的区别提供了一种说明。对于可逆循环过程这个值趋于零；对于非可逆循环过程这个积分总是具有负的值

$$\oint \frac{dQ}{T} = 0$$ （可逆循环过程）

$$\oint \frac{dQ}{T} < 0$$ （非可逆循环过程）

上式表明如果这个循环是可逆的，循环中的所有变换是相互抵消的或互为补偿的。如果这个循环是不可逆的，就有一些变换未被补偿，例如从热到冷的变换没有补偿从热到功的转化变换。[2]

1865 年，克劳修斯在《论热的动力理论的主要方程的各种应用形式》的论文中，提出了态函数 S 的概念。关于可逆循环，克劳修斯指出：

“如果物体从任意一个初态开始，连续地经过任意的一系列状态又回到初态时，∫dQ/T 总等于零，那么积分号里的表示式 dQ/T 必定是一个量的全微分，这个量只与物体当时所处的状态有关，而与物体到达这个状态所经过的途径无关。如果用 S 表示这个量，则我们就可以规定

$$dS = \frac{dQ}{T}$$[3]

因为 S 是一个态函数，dS 沿任意可逆过程的积分等于 S 的末态值 S_2 与它的始态值 S_1 之差，即

$$\int_{(1)}^{(2)} dS = S_2 - S_1$$

关于 S 的概念，克劳修斯称它为物体的变换容度(transformationconcent），即物体的转变含量，因为这个量通常是用 $\frac{dQ}{T}$ 进行积分求变换的等效值。他建议称量 S 为熵(entropy)，它是来自意思为变换的希腊字“trope”，加了一个前缀 en，以便与能量(energy)这个词相对应。克劳修斯说：“我有意把这字拼为 entropy，以便与 energy(能)尽可能地相似，因为这两个字所表示的量，在物理上都具有重要意义，而且关系密切，所以名称上的相似，我认为是有好处的。”[4]可见，在克劳修斯看来，熵和能这两个概念是有某种相似性的。事实上，能这一概念，从正面量度着运动转化的能力，能越大，运动转化的能力越大，熵却从反面，即运动不能转化的一面量度运动转化的能力，表示着转化已经完成的程度，亦即运动丧失转化能力的程度。熵这个词的中文译名是我国物理学家胡刚复教授确定的。1923 年 5 月 25 日，德国物理学家 R·普朗克在南京东南大学作《热力学第二定律及熵之观念》的报告,他为普朗克翻译时,译成为“熵”。据钱临照教授的回忆，胡刚复曾亲自对他谈过，熵一词是胡刚复首先译名的，因为熵这个概念太复杂，所以他从它是温度去除热量变化即求商数出发，把“商”字加“火”字旁，译成了“熵”。据王竹溪教授说，克劳修斯曾造了很多词，只有德文的熵(entropie)这个词流传了下来。[5]

(二)熵增加原理的提出

在 1865 年的论文中,克劳修斯力图把热力学第二定律所揭示的自然过程的方向性定量地表示出来。他说：“按照我的想法，第二定律所说明的事实是，自然界中出现的一切交换，都是在我称之为的‘正’的意义之下自行出现的，能够无补偿地发生；而沿着相反的方向，即负方向出现的变

换，只能在同时出现的正变换对之进行补偿时发生。”克劳修斯提出这种自然过程的方向性及其限度是否能用简单而明确的方式表征出来的问题。他回答说：“可以照我所做的那样把这些变换作为数学量来看待，计算这些变换的等效值，并通过代数加法求和。”[4]这说明克劳修斯是通过引入熵的概念来定量地表示热力学第二定律的。

在这篇论文中，克劳修斯还严格地证明了任何孤立系统的熵永远不会减少；或者说，自然界里的一切自发过程，总是沿着熵不减小的方向进行的。这就是“熵增加原理”。

他设想了一个由状态(1)到状态(2)的不可逆过程和从状态(2)返回到状态(1)的可逆过程构成的不可逆循环过程，对这个循环过程可得

$$\int_{(1)}^{(2)}\frac{dQ}{T}+\int_{r(2)}^{(1)}\frac{dQ}{T}<0$$

$\int$和$\int_r$ 分别表示沿着不可逆与可逆路径的积分。对于可逆过程有

$$\int_{r(2)}^{(1)}\frac{dQ}{T}=S_1-S_2$$

代入上式可得 $$\int_{(1)}^{(2)}\frac{dQ}{T}<S_2-S_1$$

对于与外界没有任何热交换的孤立系统：因为 $dQ=0$，所以 $S_2>S_1$[2]这就是熵增加原理，它是利用熵的概念所表述的热力学第二定律。

熵增加原理表明，在没有外界的作用下，一个系统的熵沿着熵增加的正方向进行，一个系统的熵越大，就愈接近于平衡状态。孤立系统里的每一种平衡必定对应于熵的极大值。德国物理学家劳厄指出：“这个结论已被证明是更为重要的。一旦人们能够表示出各种物质的熵函数，那么人们就能够因此对它们之间的平衡作出论述。所以，克劳修斯能够作出关于同一物质的不同聚集态之间的平衡理论。”[6]他还指出：“熵对于热力学统计方法是必不可少的，在普朗克辐射定律的发现过程中，熵起着重要作用，人们或许应当说，它起了决定性的作用。”[6]普朗克就是从空腔辐射场中的振子在平衡态时的熵来研究热辐射的。

熵增加原理揭示出自然过程的不可逆性，或自然过程对于时间方向的不对称性。不平衡状态可以自动地趋向平衡态，而平衡态却不能自动地转化为非平衡态。它也表明不同运动形式的转化在一个方向上存在着限制，如机械能可以完全转化为热，而热却不可能自动地完全转化为机械能。

热力学第二定律由于表明了与热运动形式联系着的能量转化的新的特点，即能量转化的方向性和限度，从而成为独立于热力学第一定律之外的另一重要定律，它使自然过程中能量转化的表征更加全面了，这在物理学理论的发展中无疑是一个重要的进步。

(三)热力学第二定律统计解释的提出

应用气体分子统计理论对热力学第二定律进行统计解释的有麦克斯韦和玻耳兹曼。

英国物理学家麦克斯韦(J.C.Maxwell,1831—1879)第一次讨论这个问题是在 1867 年他写给泰特(Tait)的一封信中，他设想了一种方式，在外界没有给系统输入功的情形下，热物体能够从冷物体获得热。麦克斯韦设想用一个膜片把容器分成 A 和 B 两部分。假设 A 中气体的温度比 B 中气体的

温度高。然后，他又设想了一个能够见到单个分子的极小的生物(finite being)，后来威廉·汤姆孙用“精灵”(demon)这个词来表示，后人又把它称为“麦克斯韦妖”。这个精灵能够打开和关闭在膜片上的小孔，可以任意地允许分子从 A 和从 B 通过这个小孔，而且有选择地只让 B 中速度快的分子进入 A，而慢分子从 A 进入 B。其结果是 A 中的能量增加，B 中的能量减少；热物体变得更热，冷物体变得更冷。这样，它将在不消耗功的情形下，只用一个观察力极其敏锐的，且能熟练拨开小孔的极其灵敏的精灵，就能实现把热从冷物体送到热物体。[1]

麦克斯韦提出这个机智的论据的用意是什么呢？他既不是用这种推测方法来论证单个地操作分子在物理上的可能性；也不是用这样一个能够操作分子的生物来否定热力学第二定律。他反对汤姆孙使用“精灵”这个词，强烈要求泰特告诉汤姆孙，不再是一个“精灵”，而是一个“活门”(valve)。克劳修斯举出这个例子的用意是要表明热力学第二定律是描述大量分子系统性质的统计性规律，而不是描述系统内单个分子的行为。上述单个快分子从冷物体流向热物体的过程是在分子级别上自发出现的。在不断出现的单个分子的自发涨落(spontaneous fluctuation)中，通过分子的无规则的运动(random motion)，快分子从冷物体运动到热物体，这种随机的涨落没有造成对热力学第二定律的违反，因为热力学第二定律描述的是明显的热流，而不是分子的随机涨落。正象麦克斯韦对泰特所说的那样：他的目的是“表明热力学第二定律只具有统计的确定性”。这个“精灵”佯谬是要说明气体分子的速度是按照统计规律分布的，特别要强调在由大量分子组成的系统中有自发涨落的存在。因此，它意味着热力学第二定律的统计实质。[1]

1871 年，麦克斯韦在《热的理论》一书的末尾，在“热力学第二定律的限制”的标题下强调热力学第二定律必须建立在大量分子运动的统计分析的基础上，它不是描述单个分子运动的动力学定律，因此物理学家必须接受统计的计算方法。

这时，奥地利物理学家玻耳兹曼(Ludwig Boltzmann，1844—1906)接受了热力学第二定律是一个统计规律的观点，他用分子运动的统计平均规律确立了熵增加的概念。

当时，对于熵增加原理所表征的自然过程的不可逆性，不少科学家是难于接受的。他们提出了所谓“可逆性佯谬”。这个佯谬指出单个粒子的运动是服从牛顿力学原理的，所以单个分子的运动以及分子之间的相互作用是完全可逆的，就是说，微观运动过程是完全可逆的，然而由大量分子在相互作用中所表现出的宏观热力学过程 S 函数单调增加的规律，即表现出不可逆性，这两者是相互矛盾的。这就是佯谬之所在，由单个粒子运动的可逆性如何会得出宏观过程的不可逆性这一结论？

为解决这个矛盾所作的努力，把玻耳兹曼的研究工作推向了高峰。玻耳兹曼引入了概率概念，找到了熵增加原理的统计解释。在他 1877 年出版的论文《关于热动力学第二定律与概率的关系、或热平衡定律》中，极其深刻地阐明了熵与系统内分子位置分布之间的关系，他指出系统的熵是它的概然性的测量，自然过程中的熵的增加对应于系统达到最概然的分子分布。于是，热力学第二定律表述的自然过程的不可逆性是系统倾向于达到最概然的热力学状态，即热平衡状态的结果。他在这篇文章中写道：

“很清楚，从某种初始状态开始，经过一定时间以后，发生的任何个别均匀状态是与发生特定的非均匀状态一样，几乎是不可能的，这正如接龙游戏一样，出现几个相同的号码牌是和刚好出现 12345 号码牌一样，几乎是不可能的，只是因为均匀状态比非均匀状态多得很多，所以概率较大，从而在时间的进程中变得均匀了。我们甚至可能从不同状态数目的关系中计算出它们的概率，从而可能导致出一种计算热平衡的有趣的方法。因此，我们深信，我们能从研究系统中各种可能状态的概率去计算热平衡状态。在大部分的情形下，初始状态是出现概率很小的状态，但是从初始状态开始，这体系将逐渐走向出现概率较大的状态，直到最后进入最概然的状态，那就是热平衡。如果我们把这种计算应用于第二定律，我们就能将普通所谓熵的那种量等同于实际状态的概率。”[7]

1877 年，玻耳兹曼把熵和系统相应的热力学状态的概率 W 联系起来，得出具有重要意义的关系式

SalnW

1900 年，普朗克引进了比例系数，写出了玻耳兹曼-普朗克公式

$$S=k\ln W$$

k 为玻耳兹曼常量。根据这一关系，玻耳兹曼把力学过程的可逆性与热力学过程的不可逆性辩证地统一起来。它揭示出热力学规律性是物质结构的原子性的表现，其统计规律性植根于体系中巨大数目粒子的随机运动。一个热力学状态的概率 W 就是这个宏观状态所对应的微观状态数。熵增加原理所揭示的孤立系统中自发过程的方向性，对应于系统从热力学概率小的状态向热力学概率大的状态过渡，由包含微观状态数目少的宏观状态向包含微观状态数目多的宏观状态过渡。平衡状态是热力学概率最大的状态，亦即熵取最大值的状态。

玻耳兹曼揭示了热力学第二定律的统计本质，指出这个定律是一个统计规律。他所揭示的熵和概率之间的联系是物理学的最深刻的思想之一。玻耳兹曼的工作有力地推动了热学理论的进展。

(四)热寂说的起源

最早提出热寂说的物理学家是汤姆孙。他在 1852 年关于自然界中机械能耗散的一篇论文中，从他所提出的公理导出结论，在自然界中占统治地位的趋向是能量转变为热而使温度趋于相同，最终导致所有物体的工作能力减小到零，达到热死状态。他在 1862 年发表了《关于太阳热的可能寿命的物理考察》论文，明确提出“热寂说”。他写道：

“热力学第二个伟大定律孕含着自然的某种不可逆作用原理，这个原理表明虽然机械能不可灭，却会有一种普遍的耗散趋向，这种耗散在物质的宇宙中会造成热量逐渐增加和扩散，以及势的枯竭，如果宇宙有限并服从现有的定律，那么结果将不可避免地出现宇宙静止和死亡状态。但是，对宇宙中物质的广延设想出一个界限是不可能的。因此，科学宁可认为它通过一个无限空间的无限进程，这种进程包括势能转变成可感知的运动，然后又转变成热，而不认为它是单一的有限机构，象一台时钟那样停下来，并永远停下去。”

从汤姆孙的这段话可以看出，他从机械能转化为热而耗散和热力学第二定律，得出宇宙热寂的观点。克劳修斯在 1865 年指出：“这个定律在宇

宙中的应用，已得出一个结论，那是汤姆孙首先注意得出的，因此我才发表我所说的论文。”可见克劳修斯承认汤姆孙先于他提出的热寂说，并启发他做进一步的尝试。

克劳修斯在前述的 1865 年的论文中把宇宙看作一个孤立的绝热系统，在这系统中热的正向变化总是大于负向变化，因此宇宙热量的总和向一个方向变化而趋于最终状态。另外，他指出他的熵，只包含了“热含量”和热离散度(disgregation) [5]，而未考虑当时已知的热辐射和由“以太”传播的热量等。他写道：“由此熵尚未用尽，还必须考虑辐射热，或以以太振动方式通过宇宙空间弥散热的其他形式，以及不包括在热名义下的那些扩展更远的某种运动。”[5]正是在上述前提下他得出宇宙的基本定律：1．宇宙的能是恒定的；2．宇宙的熵趋于极大。[4]克劳修斯在 1867 年《关于机械热理论的第二定律》的讲演中，又进一步提出：“宇宙越是接近于其熵为一最大值的极限状态，它继续发生变化的可能就越小；当它最后完全达到这个状态时，就不会再出现进一步的变化了，宇宙将永远处于一种惰性的死寂状态。”[4]克劳修斯提出上述结论，忽视或回避了他在 1865 年论文中提出的前提条件，因此引起百多年的激烈争论。一些物理学家认为把以地球上的实验为根据建立的原理推广到整个宇宙，这是很难置信的。

热寂说象物理学中许多其它观念一样，在社会上引起了巨大反响。美国历史学家亨利·亚当斯把它解释为 19 世纪所特有的低落情绪的原因；还把它与对社会进步的失望情绪相联系。正是这一观念给一些作家带来了一种宇宙热死亡的忧郁心态。具有资产阶级自由思想的英国诗人史文明这样描述了热寂：

不论是星星还是太阳将不再升起，
到处是一片黑暗，
没有溪流的潺潺声，
没有声音，没有景色，
既没有冬天的落叶，
也没有春天的嫩芽，
没有白天，也没有劳动的欢乐，
在那永恒的黑夜里，
只有没有尽头的梦境。[8]

美国物理学史家霍尔顿把这种没落情绪正确地归之于社会原因。他在《物理科学的概念和理论导论》一书中指出：“热寂说对于一些流行作家有一种不健康的吸引力，这些作家沉湎于席卷欧美社会某些部分关于世界末日的悲观情绪。由于熵的增加意味着更大的无秩序和混乱，这也许就是对社会崩溃和环境衰退的一种解释！”[9]

恩格斯于 1869 年 3 月 21 日给马克思的信中，曾严厉批判过宇宙热寂说。他说：“既然这种理论认为现在世界上转化为其它各种能的热能的数量日益超过可以转化为热能的其它各种能的数量，那末，作为冷却的起点的最初的炽热状态自然就绝对无法解释，甚至无法理解，因此，就必须设想有上帝存在了。牛顿的第一推动力变成了第一炽热。”[10]在涉及到宇宙的起源与终结的问题上，在英国物理学家中求助于神学是很普遍的，这是“自然神学”传统继续影响的表现。[1]

恩格斯在《自然辩证法》中用能量守恒与转化的观点对热寂说作了精辟地分析。他在《导言》中说：“散射到太空中去的热必须有可能以某种方法——阐明这种方法将是以后自然科学的课题——转变为另一种运动形态，在这种运动形态中它能够重新集结和活动起来。”[10]恩格斯依据天文观测资料“新星之突然地闪现以及熟知的旧星的突然增加光亮”指出散射到太空中的热能有重新集结的可能。他坚信辩证自然观的正确性，在《导言》的最后他写道：“我们确信，物质在它的一切变化中永远是同一的，它的任何一个属性都决不会丧失，因此它在某个时候以铁的必然性毁灭自己在地球上的最高的花朵——思维着的精神，而在另外的某个地方和某个时候又一定以同一种铁的必然性把它重新产生出来。”[11]

根据近代天文观察，发现了星球创生、能量重新聚积的现象。在万有引力作用下，当恒星物质向中心激烈坍缩时，释放出的引力势能可以达到使整个恒星爆炸开来，这种现象称为超新星爆炸。当超新星爆炸时，恒星的亮度可以瞬间增大千万倍。我国自殷代到公元 1700 年共记录了 90 颗新星和超新星。其中最著名的是宋至和元年(公元 1054 年)出现在金牛座 ξ 星(天关星)附近的超新星。史书《宋会要》中这样写道：“初，至和元年五月，晨出东方，守天关，昼见如太白，芒角四出，色赤白，凡见二十三日。”18 世纪末，有人通过望远镜观测，在天关星附近，发现一块外形象螃蟹的星云，取名叫蟹状星云。1921 年发现这星云在不断向外膨胀，根据膨胀速度可以反回推算出这星云物质大约是在 900 年前形成的，是超新星爆发的产物。距今最近的一次是 1987 年在南天区大麦哲仑星云处观测到的一次超新星爆炸。2 月 23—24 日的 24 小时内，超新星增亮了 2000 倍，一跃成为大麦哲仑星云中最亮的天体。由此看来，在宇宙中不仅有能量的分散过程，也有能量的重新集结过程。

热寂说是以宇宙整体正在从非平衡趋于平衡的结论为前提的。然而近代宇宙论的观测和研究表明，宇宙正在膨胀，它不是趋于平衡，而是越来越趋于不平衡。热力学第二定律在此条件下不成立，当然由此导出的热寂说也不成立。

参考文献

[1]P.M.Harman, Energy, Force, and Matter: the Concep-tual Development of Nineteenth—Century Physics, NewYork, Cam-bridge U.P.1982 年，66—69，139—143

[2]Williamh. Cropper, Rudolr Clausius and the road tocntropy, American Journal of Physics, 1986.12

[3]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，483—484

[4][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》，第 1 版，北京，商务印书馆，1986 年 5 月，249—252

[5]阎康年，“热力学第二定律和热寂说的起源与发展”，《物理》，1986 年 2 月

[6][德]劳厄著，范岱年、戴念祖译，《物理学史》，商务印书馆，1978 年，83

[7]同[4]，279，280
[8][美]库珀著，杨基方、汲长松译，《物理世界》，上卷，海洋出版社，1981 年，429
[9][美]G.Holton 著，张大卫等译，《物理科学的概念和理论导论》，上册，439 页
[10]马克思恩格斯全集，第 32 卷，267—268
[11]恩格斯著，曹葆华、于光远、谢宁等译，《自然辩证法》，人民出版社，1960，20

六、电流磁效应的发现以及电流元作用定律的建立

电流磁效应的发现揭示了电与磁的内在联系，拉开了电磁统一的序幕。电流元磁场公式的提出以及电流元作用定律的建立，奠定了电磁理论的基础。本文探讨的问题是：奥斯特的发现是在什么背景下产生的？毕奥、萨伐尔、拉普拉斯定律是怎样得到的？安培的电流元理论又是如何建立的？

(一)奥斯特的发现

在历史上很长的一段时间里，磁学和电学的研究一直彼此独立地进行着。对电和磁现象进行系统的研究是从英国的吉尔伯特(William Gilbert,1544—1603)开始的。他是英国女王的御医。在1600年出版的《论磁、磁体和作为一个巨大的磁体的地球》一书中，他指出了电现象和磁现象之间深刻的差异，诸如磁性质是几种少数磁体具有的性质，而电性质是物体通过摩擦而具有的普遍性质等等，从而认为电和磁是两种截然无关的现象。这对后来电磁学的发展产生了深刻的影响。18世纪80年代库仑论证电和磁是两个完全不同的方面，它们的作用定律在数学上很相似，但它们的性质却很不相同，认为电与磁相互转换简直是不可思议的。直到19世纪初叶，科学界还是普遍把这两种现象看作是相互独立的。[1]

但是实际上电和磁之间相互联系的现象早已引起了一些人的注意。早在18世纪30年代，就有人描述过闪电能使箱中刀、叉、钢针磁化的现象。1751年富兰克林(Benjamin Franklin，1706—1790)已发现用莱顿瓶放电的方法可以使焊条、钢针磁化或退磁。当时关于闪电改变钢铁物件磁性的事情屡见报道，但既未作过系统的研究，更想不到这与电流之间有何联系。1774年巴伐利亚电学研究院还特地出了一个有奖征文题目：“电力和磁力是否存在着实际的和物理的相似性？”不少人用实验来加以研究。1805年，德国的哈切特(J.N.P.Hachetle)和笛索米斯(C.B.Desormes)用一根绝缘绳将伏打电堆悬挂起来，企图观察它在地球作用下是如何取向的，但实验没有得出结果。19世纪初，戴维在研究电极的炭棒之间的弧光时，曾观察到磁铁能够吸引或排斥弧光，并使弧光偏转，这是关于电磁之间相互作用的早期发现之一。[1]

丹麦物理学家奥斯特(Hans Christian Oersted，1777—1851)出生于一个贫穷的药剂师家里。他17岁时便考取了哥本哈根大学的免费生。他一边当家庭教师一边在学校学习药物学，对物理学、天文学、哲学和文学都很有兴趣。1797年，他在哥本哈根大学毕业取得药剂师称号，并由于他写的美学和医学方面的论文而获得金质奖章。1799年，他由于一篇关于康德哲学的论文被授予哲学博士学位。1806年，任哥本哈根大学物理学教授。1821年，被选为英国皇家学会会员。两年后，又被选为法国科学院院士。后来出任丹麦皇家科学协会会长。[1]

奥斯特是一个深受德国古典哲学影响的物理学家，他受到德国“自然哲学”，特别是谢林(Friedrich Schelling，1775—1854)关于各种现象相互联系和相互转化思想的影响。他说：“谢林想给我们一个物理学的完整的哲学体系，但他除教科书外没有任何自然界的知识。”自然哲学强调知识的先验性，反对知识的经验性，带有思辨的色彩，它把虚构的联系强加

予自然界，而不是从自然界中去寻求真实的联系。但是它关于世界的统一性和运动形式相互转化的思想，对自然科学的发展起着积极的作用。奥斯特从中吸取了思想营养。1803 年，奥斯特就说过：“我们的物理学将不再是关于运动、热、空气、光、电、磁以及我们所知道的任何其它现象的零散的罗列，而我们将把整个宇宙容纳在一个体系中。”奥斯特通过对电的研究增强了他对自然力统一的信念。他说：“既然长期以来，我认为电力是自然界最一般的力，我也必须从它们得到磁的效应。”[2]1813 年出版的《关于化学力和电力的统一性的研究》中，奥斯特根据电流流经直径较小的导线时会发热的现象推测，如果通电导线的直径变得更小，小到一定程度时，电流就会发生磁效应。他指出：“必须检验电是否以其最隐蔽的方式对磁体有所影响。”寻找这两个自然力之间联系的思想，经常盘绕在他的头脑中。

1819 年冬天，奥斯特在哥本哈根大学讲授电学、伽伐尼电流和磁学的课程时，进行了电流对磁针效应的第一批实验。他曾经说过：“这些实验似乎表明磁针因伽伐尼仪器的影响而移动了位置，而且这种情况发生时，伽伐尼电路必须是闭合的，而不是开路的；后一种方法几位著名的哲学家几年前已经徒劳地尝试过。由于这些实验是用有缺陷的仪器进行的，所以对这样重要的问题不足以下结论。”[3]这是他在 1820 年 7 月 21 日的论文开首的一段话。在以后的文章中他一直明确指出电流磁效应发现的时间在 1820 年春天。《早期电动力学》一书的编者 R.Tricker 认为奥斯特在 1819 年冬天的实验是失败的。他在注解中写道：“直到 1819 年末奥斯特仍然证明这一实验是否定的实验。遗憾地是当他试图在闭合回路和悬挂的磁铁之间发现某些作用力时，他把导线放在与磁针垂直的方向上，所以未获得效应。”[2]

奥斯特在谈到自己的发现时写道：“1820 年春天，我在哥本哈根开办了一个电和磁的讲座，听众多数是在科学上已有相当成就的人，所以这些讲座的备课导致我更深入地研究通常讲课中的一些看法，于是我以前对电力和磁力一致性的确信以新的清晰性出现了。我决心通过实验检验我的看法。”[2]

1857 年，在 Hansteen 教授致法拉第的一封著名的信中，叙述了他曾经作为奥斯特的实验助手亲眼目睹奥斯特发现电流磁效

图 6-1 奥期特的实验装置

应的具体过程。“在 1820 年 4 月的某一天晚上，奥斯特在课堂上作电流磁效应的实验，他把联接伽伐尼电池两极的导线垂直地放在磁针之上，但是没有显示出明显的运动。当他讲课一结束，他说：‘让我们现在立即把导线放在与磁针平行的位置来试试看。’当他这样作后，突然发现电流附近的小磁针向垂直于导线的方向大幅度地转过去，振荡起来，并在几乎垂直于导线的方向停下来。奥斯特激动万分。于是他说：‘现在让我们使电流反向流动。’他把电池转了 180 度，再接上电源，结果磁针就向相反方向转去。伟大的发现就这样得到了。”[2]

奥斯特紧紧抓住这一发现，先后作了一系列实验。1820 年 7 月 21 日，在《关于磁体周围电冲突的实验》的报告中，公布了他的实验结果。他把电流在周围空间产生的磁效应称为“电冲突”(electricconflict)，并总

结了所谓电冲突的某些特性。他在导线和磁针之间放上玻璃片、金属片、木片等等物质，都不妨碍电流对磁针的偏转作用。而把悬吊的磁针换为钢针、玻璃针、树胶针，则都静止不动。根据这些现象，他得出结论如下：“电冲突只能对物质的磁质点起作用。一切非磁体好象能让电冲突通过。但是磁体，或者宁可说物体的磁质点好象是抵抗电冲突通过，所以它们能由于这种对抗的力产生的冲量而运动。”[4]

他又写道：“从上述事实，我们还可以推出电冲突呈现为圆形。否则就不可能发生这样的情形，将联接的导线放在磁极下面时，磁极被推向东方，而放在磁极上面时，就被推向西方。其原因是只有圆才具有这种性质，其相对部分的运动具有相反的方向。”[4]

奥斯特发现电流对磁极的作用力是一种旋转力，这是对中心力观念的有力冲击。当时人们知道的重力和静电力都是沿着物质之间或电荷之间的连线发生作用的，即沿着纵向发生的。人们习惯于这种预测。当奥斯特开始作实验时也是这样假设：电流对磁极的作用力是纵向力，所以他把电流垂直磁针放置。当他从多次的失败中意识到电流对磁针的作用力可能是横向力的时候，沿着这一新的方向探索，他终于取得重大突破。

有些人过分地宣染奥斯特发现的偶然性。其实他的发现具有深刻的思想文化背景和科学技术发展前提，只有在 1800 年伏打发明电池之后，才有可能为实验提供稳定电流；此外这也与奥斯特本人多年的探索分不开。正如法国生物学家巴斯德说的一句话：“在观察的领域中，机遇只偏爱那种有准备的头脑”。奥斯特发现的电流的磁效应揭开了把电学和磁学联系起来的电磁学的新篇章，为近代电磁理论的发展奠定了实验基础。

(二)毕奥-萨伐尔定律的确立

奥斯特的发现使整个科学界大为震动，人们长期以来所信奉的电和磁没有内在联系的信条崩溃了。这个发现开启了一扇通向新的研究领域的大门，导致了进一步探索的激流。正如法拉第所说：“猛然打开了科学中一个黑暗领域的大门。”

1820 年 8 月间，法国物理学家阿拉果在瑞士听到了奥斯特发现电流磁效应的消息，立即敏锐地感到这一成果的重要性。回到巴黎后，他在 9 月 11 日法国科学院的一次会议上报告了奥斯特的新发现，并详细地描述了电流磁效应的实验。阿拉果的报告立即在法国引起了巨大的反响。

对这一效应首次进行精确分析的两位研究者是毕奥(Jean Baptiste Biot，1774—1862)和萨伐尔(Felix Savart 1791—1841)。毕奥曾任法兰西学院物理学教授，他的科学兴趣是多方面的，对光学尤有研究，还写了许多数学著作。萨伐尔早年曾行医，1819 年他给毕奥呈送一篇论文，毕奥对这个年轻人发生了兴趣，鼓励他继续研究，1828 年他任法兰西学院实验物理教授。[2]

1820 年 10 月 20 日，在法国科学院的一次会议上，毕奥和萨伐尔宣读了题为《运动的电传递给金属的磁化力》的论文。报告了他们发现直线电流对磁针作用的规律：直线电流对磁极的作用正比于电流的强度，反比于它们之间的距离，作用的方向则垂直于磁极到导线的垂线。

他们的实验装置如图 6－2 所示，一条垂直的金属丝 CZ 两端

图 6-2 毕奥-萨伐尔实验装置

联接到伏打仪器的两极，一根磁化的钢针用一根丝线悬挂在水平位置上，通过调节装置可以改变磁针与金属丝的距离，在一定位置和方向上放一条形磁铁以便消除地磁力的影响。

图 6-3

当电流通过导线时，磁针就转到与其中心到导线的距离相垂直的水平方向，和奥斯特所表示的旋转方向一致。我们通过磁针取一个垂直于导线的水平面，磁针的平衡位置如图 6-3 所示，A、B 是磁针的两极，C 是它的中心，F 是导线与水平面的交点。因为导线是无限长的，磁针的每个极受到的合力必定在水平面内。假定北极 N 的北磁分子受到的力沿 BD 线方向，南极 S 的南磁分子受到的力沿 AE 线方向，很显然在与导线距离相同处，导线对磁分子的作用是相同的，所以角∠EAF=∠GBF。又因为磁针处于平衡状态，由实验观测得 BCA⊥CF，所以 BD 与 AE 必须与磁针具有相同的倾角，即∠DEC=∠EAC，这就要求∠DBF=∠EAF。因此∠DBF=∠GBF。又因为∠DBF+∠GBF=180°，所以∠DBF=∠GBF=90°。

毕奥-萨伐尔由此得出结论：“一条无限长的载流导线作用在南磁分子或北磁分子上的力垂直于该分子到导线的距离。”[5]

如何确定力的大小呢？毕奥-萨伐尔是通过测量磁针的振荡周期来确定的。根据单摆原理在小振幅的情形下，力的大小与振荡周期的平方成反比。因此，如果我们调节磁针与导线处于不同距离处，测量磁针完成一定振荡次数所需要的时间，并把它们加以比较，就可以确定导线产生的磁力与距离的关系。

如果设 F 是磁针在某一距离 d 处所受的力，t 是完成一定振荡数所需的时间。F′是在 d′处所受的力，t′是完成同样振荡数所需的时间，则在 d′处磁针受的力 F′为

$$F'=\frac{t^2}{t'_2}F$$

实验表明观测到的力的比率几乎严格地与磁针到导线的距离成反比，即

$$\frac{F'}{F}=\frac{d}{d'}$$

于是，毕奥-萨伐尔得出结论：“载流导线对南磁分子或北磁分子的作用力与磁分子到导线之间的距离成反比”。他们又通过实验证明了载流导线对磁分子的作用力与电流强度的大小成正比。

用现代形式表示，设长直导线通过电流强度为 I 的电流，在距它为 r 处的磁感应强度为 B，则 B 与 I、r 具有如下的关系

$$B=K\frac{I}{r}$$

K 是比例常数。

法国数学家、物理学家拉普拉斯(PierreSimonLaplace，1749—1827)遵循将一切物理现象简化为粒子间的引力或斥力现象的途径，根据毕奥-萨伐尔由实验得出的长直导线公式，从数学上推导出每个线段元施加在磁极上的作用力的规律。象合力一样，线段元对磁极的作用力垂直于线段元

和它到磁极的距离所构成的平面，这个力的大小反比于距离的平方。用现代形式表示，设电流元为 Ids，Ids 与距离 r 的夹角为θ，如图 6-4 所示，则电流元 Ids 在 P 点产生的磁场 dB 为

$$dB = K\frac{Ids\sin\theta}{r^2}$$

写成矢量形式 $$dB = K\frac{Ids \times r}{r^2}$$

这就是毕奥-萨伐尔定律。

毕奥和萨伐尔为了验证拉普拉斯导出的这个线段元磁场公

图 6-4　图 6-5

式，作了第二组实验。他们把长直导线 ZMC 弯折成与水平面的夹角都为 a 的 ZM 和 MC，如图 6-5 所示，这条折线与竖直导线 Z′M′C′和磁针中心都在同一个竖直平面内，拆线与竖直导线在 MM′处用一纸片隔开。

据拉普拉斯公式可以导出载流斜线对单位磁极的作用力，从而确定磁物 B 与斜线倾角 a 的关系

$$B = 2K\int_{\theta=0}^{\theta=a}\frac{Ids\sin\theta}{r^2} = 2K\int_0^a\frac{Id\theta}{r}$$

$$= 2K\int_0^a \frac{I\sin\theta}{d\sin\alpha}d\theta = \frac{2KI}{d\sin\alpha}(-\cos\theta)\Big|_0^a$$

$$= \frac{2KI}{d\sin\alpha}(1-\cos\alpha) = \frac{2KI}{d}\mathrm{tg}\frac{\alpha}{2}$$

由此看出这条斜线的作用比直线的作用多了一个$\mathrm{tg}\frac{\alpha}{2}$的因子。如果长直导线对磁极的作用力为 F，则通有同样大小电流的倾斜导线对磁极的作用力应为$F\mathrm{tg}\frac{\alpha}{2}$。毕奥-萨伐尔讨论了$a = 45°$的情形。

图 6-6

$$F\mathrm{tg}\frac{\alpha}{2} = F\mathrm{tg}22^\circ 30' = 0.414214F$$

在倾斜导线中通过两倍电流，则将施加两倍的力在磁针上，它与竖直导线对磁针作用力的比值将变为 0.828427。

设 t 是没有电流的情形下，磁针完成一定振荡次数所需的时间，t′是导线通有电流时，磁针完成同样振荡次数所需的时间，则电流本身对磁针的作用力为

$$\frac{K}{t'^2} - \frac{K}{t^2} = K\frac{(t-t')(t+t')}{t^2t'^2}$$

K是取决于磁针线度的常数。设倾斜导线与竖直导线作用的比率为$\frac{O}{V}$,则

$$\frac{O}{V}=\frac{(t-t_0')(t+t'_0)}{(t-t_V')(t+t'_V)}\quad\frac{t'^2_V}{t'^2_0}$$

式中 t'$_V$ 直线通有电流时，磁针完成一定振荡次数所需的时间。t'$_0$ 斜线通有两倍电流时，磁针完成同样振荡次数所需的时间。

当考虑到两条倾斜导线，每边偏离竖直导线右边或左边 3㎜，则它们到磁针中心的距离不是d，而是$\left(d^2+9^{\frac{1}{2}}\right)$或$d\left(1+\frac{9}{d^2}\right)^{\frac{1}{2}}$，因$\frac{9}{d^2}$是个小数，可以导出

$$d\left(1+\frac{9}{d^2}\right)=d\left[1+\left(\frac{9}{2d^2}\right)\right]$$

为了推导出倾斜导线与竖直导线在同一距离上对磁针作用力的比$\frac{F_0}{F_v}$，必须用（$1+\frac{9}{d_2}$）的倒数乘上比率$\frac{O}{V}$，这样就可以得到

$$\frac{F_0}{F_v}=\frac{O}{V}\quad\frac{1}{\left(1+\frac{9}{2d^2}\right)}$$

毕奥-萨伐尔根据实验观测结果得到的倾斜导线与竖直导线对磁针作用力的比率为0.827545几乎严格的与0.828427一致。这就说明$Ftg\frac{\alpha}{2}$普遍表示了一条折成彼此夹角为 2a 的两条臂的斜线对磁针的总的作用力。这就从实验上再一次证明了毕奥-萨伐尔定律的正确性。

(三)安培的基本实验和电流元作用理论

法国物理学家安培(Andre Marie Ampe re，1775—1836)出生于里昂一个富商家庭。少年时代就表现出对数学的浓厚兴趣，12 岁时就学习微积分，随后又学习了拉格朗日的《分析力学》，他是在刻苦自学中成长的。1793 年，他的父亲在法国革命中被处死，使年轻的安培在精神上受到很大的打击。1803 年，他的夫人又病故，在坎坷的生活道路上他坚持科学研究。1809 年，他任巴黎工业大学数学教授，1814 年，被选为法国科学院院士，1824 年担任法兰西学院实验物理学教授，1827 年被选为英国皇家学会会员。

安培在科学上极其敏锐，很善于接受新的成果。作为一个数学家他对奥斯特的新发现作出了异乎寻常的反应，他立即转向了电学研究。1820 年 9 月 25 日，在阿拉果报告后的第二周，他就向法国科学院报告了关于两条平行载流导线之间相互作用的发现；证明了同向电流相互吸引，反向电流相互排斥。

安培的目的是要通过实验确立支配电磁现象的普遍规律。他所遵循的途径，象牛顿把质量分解成质量元那样，把电流分成无限多电流元，进而推导出类似质点引力公式的电流元相互作用公式。如果找到了这个关系

式，就可以通过积分得到所有电磁现象的定量结果。为了这个目的，他从1820 年 10 月开始，集中精力研究电流之间的相互作用。他设计了四个精巧的实验来探讨电流之间相互作用的性质。

实验一：安培用硬导线做成如图 6-7 所示形状的线圈，这线

图 6-7

圈由两个形状和大小相同，电流方向相反的平面回路固定连在一起，整个犹如一个刚体。线圈的端点 A、B 通过水银槽和固定支架相连。这样，线圈既可通入电流，又可自由转动。这种装置叫做无定向秤(astaticbalance)，它在均匀磁场(如地磁场)中不受力和力矩，随时可以平衡，但对于非均匀磁场将会作出反应。

安培用上述装置来检验图 6-8(a)所示的通有电流的对折导线对无定向秤是否有作用力，结果是否定的。从而证明当电流反向时，它产生的作用力也反向。

实验二：把对折导线的另一臂绕成螺旋线(图 6-9a)，结果也是否定的。从而证明电流元的作用具有矢量性质，即许多电流元的合作用是单个电流产生作用的矢量叠加(图 6-9b 和图 6-8b)。

图 6-8　　图 6-9

实验三：他将一圆弧形导体架在水银槽上(图 6-10)，导体与一绝缘柄固连，柄架在圆心 C 处的支点上，这样既可给弧形导体通电，弧形导体又可绕圆心转动，从而构成一个只能沿自身长度方向移动，但不能作横向位移的电流元。然后用各种载流线圈对它作用，结果都不能使弧形导体沿其电流方向运动，这就证明了作用在电流元上的力是与它垂直的。

实验四：他用三个几何形状相似的线圈(图 6-11)，1 和 3 两线圈固定并串联在一起，通入电流 I_1，线圈 2 可以活动，通入电流 I_2。当三个线圈间距之比与三个线圈的线度之比一致时，1、3 线圈对 2 线圈的合作用为零。从而证明所有几何线度(电流元长度、相互距离)增加同一倍数时，作用力不变。[6]

1．弧形导体 2．绝缘柄 3、4．水银槽

图 6—10

图 6—11

安培的四个实验独具匠心。他不去直接测量环流之间的作用力，而是把实验建立在精确的平衡基础上。采用示零实验(“null”experiment)的方法，标志着物理实验一个明显的进步。[6]

安培设计的四个实验，每个实验都可以用来检验自己的物理思想，确定一个方面的问题。然后他把这四个实验的结果综合起来进行分析，再加上一个假设：两个电流元的相互作用力沿着它们的连线。于是便导出了电流元之间相互作用力的公式,从设想到完成仅用了不到两个月的时间。1820 年 12 月 4 日，他在法国科学院的会议上报告了电流元相互作用定律，即安培定律。这份报告收集在 1823 年的回忆录中，下面叙述的就是安培的原始

推导。[7]

首先论证两个电流元的相互作用力正比于它们的长度。他假设把电流元分成无限多相等的部分，各个微小部分之间的相互作用几乎沿同一直线，所以这些作用力可用代数相加。因为电流强度是用作用力来定义的，当然作用力正比于电流强度。令 i 和 i′是两个电流元的电流强度，ds 和 ds′是相应的长度，当它们彼此平行且垂直于二者中点的连线并相距单位长度时，作用力正比于 ii′dsds′，两电流方向相同取正号，方向相反取负号。

图 6-12

任意放置的两个电流元的相互作用还取决于它们的相对位置。设两个线元与其中点连线的夹角为θ和θ′，线元与中点连线形成的二平面之间的夹角用ω表示，令ρ为力与角θ、θ′，ω的未知函数，假设力与两个线元距离的 n 次方成反比，n 是待定常数，于是两个电流元作用力的一般表示式为ρii′dsds′/r^n 留待解决的问题是如何确定ρ与 n。

考虑两个彼此平行垂直于中点连线距离为 r 的电流元 ad 和 a′d′，它们的作用力为 ii′dsds′/r^n。假设 ad 固定不变，a′d′作平移运动且二中点距离保持不变，ω总是为零，则它们的作用力只取决于θ、θ′。按照电流方向θ与θ′或是相等，或互为补角，因此它们的作用力为 ii′ds ds′ω(θ，θ′)/r^n。当 a′d′处在 ad 延长线上 a″′d″′的时候，令此处的ω(θ、θ′)为 K，则 ad 对 a″′d″′的作用力为 Kii′dsds′/r^n。常数 K 表示了 ad 对 a″′d″′的作用与 ad 对 a′d′作用的比率，这个比率不依赖于距离 r、电流强度 i、i′和两个线元的长度 ds、ds′。

令两个电流元 Mm、 M′m′的长度分别为 ds、 ds′，其中点分别为 A、A′，通过线元 Mm 和二中点的连线作平面 MA′m。我们用 ds 在直线 AA′上的投影 Nn=dscosθ和在 MA′m 平面内垂直于 AA′的投影 Pp=dssinθ来代替 ds。再用 ds′在直线 AA′上的投影 N′n′=ds′cosθ′和在 M′Am′平面内垂直 AA′的投影 P′p′=ds′sinθ′，来代替 ds′。最后我们用 P′p′在 MA′m 平面内的投影 T′t′=ds′sinθ′cosω和 Pp′在垂直于 MA′m 平面的投影 U′u′=ds′sinθ′sinω来代替 P′p′。

图 6-13

实验表明两个端点重合的直线电流元与曲线电流元对外界的作用是相同的。所以电流元 ids 对 i′ds′的作用与 idscosθ和 idssinθ对 i′ds′cosθ′、i′ds′sinθ′cosω、i′ds′sinθ′sinω的作用一样。

因为 i′ds′sinθ′sinω的中点在 MA′m 平面内而且它又垂直于 MA′m 平面，所以它和这个平面内 idscosθ和 idssinθ没有作用力。同理，i′dscosθ和 i′ds′sinθ′之间以及 idssinθ和 i′ds′cosθ′之间没有作用力，因为设想通过直线 AA′且垂直于 MA′m 作一平面，dscosθ和 ds′cosθ′在这个平面内而 ds′sinθ′cosω和 dssinθ在垂直于它的平面内。因此两个线元 ids 和 i′ds′之间的作用最终减少到两个联合作用，即 idssinθ和 i′ds′sinθ′cosω以及 idscosθ和 i′ds′cosθ′之间的作用。这两个作用都是沿着中点的连线。它们之和就是 ids 和 i′ds′之间的相互作用。因为 idssinθ和 i′ds′sinθ′cosω是在同一

个平面且都垂直于直线 AA′，所以沿 AA′方向的作用力为

$$\frac{ii'dsds'\sin\theta\sin\theta'\cos\varpi}{r^n}$$

而 idscosθ 和 i′ds′cosθ′沿 AA′方向的作用力为

$$\frac{Kii'dsds'\cos\theta\cos\theta'}{r^n}$$

因此两个电流元 ids 和 i′ds′的相互作用力为

$$\frac{ii'dsds'}{r^n}(\sin\theta\sin\theta'\cos\omega + K\cos\theta\cos\theta')$$

这就是 1820 年 10 月 4 日安培在法国科学院会议上公布的电流元相互作用力的公式。1833 年，他在回忆录中推导出系数 n 和 K 的关系式为：1-n-2K=0。在考虑到拉普拉斯从毕奥-萨伐尔的实验中得到的电流元磁场公式后，他得到$n = 2$，$K = -\frac{1}{2}$。Tricker根据量纲分析，式中分子具有长度的二次方，同样得到上述结果。这样安培公式的最后表现形式为

$$F = \frac{ii'dsds'}{r^2}\left[\sin\theta\sin\theta'\cos\varpi - \frac{1}{2}\cos\theta\cos\theta'\right]$$

Tricker 在他编写的《早期电动力学》一书的注释中证明了应用安培公式求解长直导线对电流元的作用力与由毕奥-萨伐尔定律得出的结果是一致的。[2]

求载流长直导线对平行于它的电流元的作用力。考虑任一电流元 ids 施加在 i′ds′上的力，设ω=0，力沿 PQ 线，θ=θ′，则

$$dF = \frac{ii'ds'}{r^2}\left(\sin^2\theta - \frac{\cos^2\theta}{2}\right)ds$$

在垂直 AB 方向上的分力为

$$dF_\perp = \frac{ii'ds'}{r^2}\left(\sin^2\theta - \frac{\cos^2\theta}{2}\right)ds$$

因为$r = \frac{d}{\sin\theta}$和$r\frac{d\theta}{ds} = -\sin\theta$

所以 $$F_\perp = \frac{ii'ds'}{d}\int_0^\pi\left(1 - \frac{3}{2}\cos^2\theta\right)\sin\theta d\theta$$

$$= \frac{ii'ds'}{d}\left(\cos\theta - \frac{\cos^3\theta}{2}\right)\Bigg|_0^\pi$$

$$= \frac{iids'}{d}$$

图 6—14　 图 6—15

在平行 AB 方向上的分力为

$$
\begin{aligned}
F_{ll} &= \frac{ii'ds'}{d}\int_0^{\pi}\left(\frac{3}{2}\sin^2\theta-\frac{1}{2}\right)\cos\theta d\theta \\
&= \frac{ii'ds'}{d}\left(\frac{\sin^3\theta-\sin\theta}{2}\right)\Bigg|_0^{\pi}=0
\end{aligned}
$$

求载流长直导线对垂直于直导线且与直导线在同一平面内的电流元 i ′ds′的作用力。在平行于 AB 方向的分力为

$$
\begin{aligned}
F_{ll} &= \frac{ii'ds'}{r^2}\int_0^{\pi}\left(\sin\theta\cos\theta+\frac{1}{2}\sin\theta\cos\theta\right)\cos\theta ds \\
&= \frac{ii'ds'}{d}\int_{\pi}^{0}\frac{3}{2}\cos^2\theta\sin\theta d\theta \\
&= \frac{ii'ds'}{d}\left(\frac{\cos^3\theta}{2}\right)\Bigg|_0^{\pi}=\frac{ii'ds'}{d}
\end{aligned}
$$

可以很容易地证明，在垂直于 AB 方向的分力 $F_{\perp}$=0。

当电流元 i ′ds′垂直于它的中点与直导线组成的平面时，有

$$\omega=\frac{\pi}{\ },\ \theta'=\frac{\pi}{\ },$$

因此

$$\sin\theta\sin\theta'\cos\omega-\cos\theta\cos\theta'=0$$

所以对直导线上所有电流元对 i ′ds′的全部力也是为零。以上讨论是与依据毕奥-萨伐尔定律计算的结果一致的。

1827 年，安培在《电动力学理论》一书中阐述了他处理电磁现象的方法。他说：“首先观察事实，尽可能地变换条件以精确的测量来实现这一步。目的在于推导建立在经验基础上的一般定律，在于独立于所有关于产生现象的力的性质的假说推导这些力的数学值，就是说目的在于获得表示它们的方式。这就是牛顿所走过的道路，也是对物理学作出重大贡献的法兰西知识界近来普遍遵循的途径。同样，它也在引导我把所有研究都投入到电动力学现象中去。”[8]

这里所讲的牛顿所走过的道路指的是从运动现象研究自然力的规律，然后又以这些力的规律来解释运动现象的道路。万有引力定律的发现就是这样。关于引力的机制，牛顿认为在没有从观察和实验中发现重力之原因时，决不杜撰假设。安培遵循这条途径，从实验出发，推导出电流元相互作用的规律。

安培定律的建立奠定了电磁理论的基础。麦克斯韦对安培的工作给以高度的评价：“安培为建立电流间的作用定律而进行的实验研究在科学上是最光辉的成就，整个理论和实验似乎是从‘电学中的牛顿’的头脑中跳跃出来的，曾在那里得到全盘思考和全副武装，它们是完善的无懈可击的。

它们总结在一个公式中，所有电磁现象都可以从这个公式推导出来，这个公式将永远是电动力学的基本公式。”[8]

参考文献

[1]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，422，548—550

[2]R.A.R.Tricker，Early Electrodynamics， First Editicn，New york，Pergamon Press，1965 年，9—15，21—25，48—53

[3]R.A.R.Tricker，ibid，113—117

[4][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》，第 1 版，北京，商务印书馆，1986 年 5 月，456—461

[5]R.A.R.Tricker，ibid，118—139

[6]赵凯华，“安培定律是如何建立起来的？”《物理教学》，1980 年 1 月

[7]R.A.R.Tricker，ibid，172—181

[8]宋德生，“安培和他在科学上的贡献”，《自然杂志》，1984 年 4 月

七、电磁感应现象的发现和电磁感应定律的确立

电磁感应现象是电磁学中最重大的发现之一，是法拉第所获得的最伟大的实验成果，这一发现进一步揭示了电与磁的相互联系和转化。电磁感应定律是电磁理论和电磁测量的基石，是发电机的理论基础。它的确立开创了人类利用电能的新时代。

本文主要依据《法拉第日记》和法拉第的《电学实验研究》以及有关著作，探索以下几个问题：电磁感应现象是怎样发现的？力线概念是怎样形成的？电磁感应定律的定量表达式是怎样从实验上和理论上得来的？

法拉第(MichaelFaraday，1791—1867)是英国伟大的物理学家。他生于伦敦近郊一个铁匠家庭，13 岁时就到一家书报店里当报童，不久又成为装订书的学徒。1813 年经化学家戴维推荐他作了皇家研究院的助理实验员。1821 年担任了皇家研究院实验室主任,同年发现了电磁旋转现象。1824 年当选为皇家学会会员，1825 年成为皇家研究院教授，1831 年发现了电磁感应现象。20 年后他从实验上确立了电磁感应定律。1852 年他引进了力线概念，强调电磁场是物理存在，在理论上作出了重大贡献。

(一)电磁旋转的发现

1821 年，《英国哲学年鉴》的编辑邀请法拉第撰写一篇文章，综合评述奥斯特发现电流磁效应一年来电磁学实验和理论的发展状况。他以极大的热情投入这项工作，力图给这一科学新领域的发展状况以精确的说明。他收集了各种杂志中许多物理学家的论文，

FARADAY'S DIARY
Being the Various Philosophical Notes
of Experimental Investigation
made by
MICHAEL FARADAY
D.C.L., F.R.S.
during the years 1820—1862
and bequeathed by him to the
ROYAL INSTITUTION OF GREAT BRITAIN
Now, by order of the Managers,
printed and published for the first time,
under the editorial supervision of
THOMAS MARTIN. M.Sc.
with a Foreword by
SIR WILLIAM H. BRAGG, O. M., K.B.E., F.R.S.
Director of the Laboratory of the
Royal Institution

Vol.I
SEPT, 1820—JUNE II, 1832

LONDON
G.BELLANDSONS，LTD
1932
图 7—11820—1862 年间的《法拉第日记》

并重复作了这些论文中所描述的大多数实验。同年 10 月 1 日，这篇论文以《电磁学发展历史概况》为题发表在《哲学年鉴》上。这是法拉第发表的第一篇电磁学论文，是他转向电磁学研究的开始。[1]

在奥斯特已经作出了电流磁效应的伟大发现后，法拉第曾经想过对于这一课题还能作些什么新的、有趣的事情呢？他在仔细重复奥斯特实验时，发现磁针的磁极不是在磁针端点，而是在磁轴上距离磁针端点很小一段距离处。进而他又发现一种新的电磁运动形式，即电磁旋转现象。这就大大丰富并发展了奥斯特实验的内容。为了解释这一现象，他提出了载流导线或磁铁周围存在环形的“电磁流”的设想，开始萌发了“力线”概念。

早在 1820 年 8 月，英国物理学家沃拉斯顿(W. H. Wolla-ston，1766—1828)首先想到了在电流回路中使一段导线绕它自己的轴旋转的可能性。1821 年初，他为此目的作了一套装置来到皇家学会试图表演这项实验，但没有成功。

1821 年 9 月，法拉第开始进行电磁旋转的研究。他是从确定磁针相对于载流直导线的位置中发现电磁旋转现象的。在 9 月 3 日的法拉第日记上记述了他测定磁针的每个极相对于导线的吸引位置和排斥位置的实验。实验装置如图 7－2 所示，把一根垂直导线的两端联接到电池的两极，在导线中电流由 Cu 极到 Zn 极。[2]在导线近旁放一水平磁针使之处于“自然位置”，即垂直于导线的位置，然后沿磁针的轴线方向移动磁针，例如 N 极趋近导线，如图 7－3，他发现先是 N 极被导线吸引，继续而又出现被排斥的情况。他提出这和奥斯特在同样情况下只观察出吸引一种情况不同。并由此得出结论磁针两臂的作用中心或真实磁极位置并不在磁针的端点，而是在磁轴上离磁针端点一段距离处。[3]这说明法拉第的实验观察是非常仔细的。他不但观察了磁极受电流吸引或排斥的情况，而且观察了磁极绕电流转动以及电流绕磁极转动的情况，即所谓电磁旋转现象。

图 7—2 图 7—3

法拉第说明电磁旋转现象的实验是很精巧的，如图 7－4 所示，左边为磁极绕电流转动的装置，右为载流导线绕磁铁转动的装置。在一个装着水银的容器的上下两端分别装上金属棒 A 和铂条 C，将一根条形磁铁的南极固定，用导线连接在铂条上，北极露出水银面，当电流由 C 经水银到 A 通过时，北极 N 就绕着容器沿反时针方向旋转。如果北极在下，南极在上，则旋转方向相反。如果把磁铁固定，使容器上部的金属棒 A 能绕它的上端自由转动，当通电时，金属棒 A 就绕着条形磁铁转动。

1921 年 9 月 11 日他写了一篇总结这一实验成果的论文，题为《论某些新的电磁运动和磁的理论》，同年 12 月底发表在《英国哲学年鉴》上。他在这篇文章中探索了电磁旋转现象的本质，提出了“电磁流”的概念。他认为导线绕磁极旋转是因为磁极周围存在环形的“电磁流”；磁极绕导线旋转是因为导线周围存在环形的“电磁流”。联系到撒在磁体周围的铁

屑所排列的曲线，他萌发了磁力线的想法。他写道：“我相信每个磁学实验对于磁极或金属屑的全部影响是与环流的概念一致的。”[3]

图 7—4

法拉第关于电磁旋转的发现是奥斯特发现之后的一个重大发现。它不但加深了人们对电磁相互作用的认识，而且这也是由磁运动产生连续的机械运动，由电能转换为机械能的第一个装置。实际上这就是以后被广泛应用的电动机的雏形。

(二)电磁感应现象的发现

1822 年，安培和德莱里弗做了一个电流的感应效应的实验。如图 7—5 所示，他们将强磁体移向铜环，当线圈与电池接通时，发现铜环发生了偏转。这本来是由于铜环中产生了感应电流的缘

图 7—5 安培在 1821 年和 1822 年所做实验的装置

故。而当时他们两人却局限在安培的分子电流框架内作出解释，认为当线圈与电池接通时，铜环暂时地被磁化了，在铜环中有分子电流产生，铜环的转动是由于强磁极对分子电流作用的结果。不久后，法拉第便得知这一重要实验的信息，但他在重复这一实验时，没有使用铜圆环，而是用了一个铜盘，由于铜盘的转动惯量太大，所以没有观察到安培所看到的效应。

1822 年阿拉果(D. F. Arago，1786—1853)和洪堡(A. V. Humboldt，1769—1859)在测定格林威治山的磁强度时,偶然发现金属可以阻尼磁针的振荡。阿拉果想，既然一个运动着的磁针可以被金属片所吸引，那末，一个静止的磁针也一定可以被一个运动着的金属片带走。1825 年，他作了一个圆盘实验，在一个可以绕垂直轴旋转的铜盘正上方悬挂一根磁针，当铜盘旋转时，磁针跟着旋转，但有所滞后。这本来是由于铜盘中产生了感应电流的缘故。而他们当时却作了机械论的解释，认为是由于金属盘的旋转离心力使盘中两种电流体分离而形成电流，这种电流作用于磁针中的分子电流，便扰动了磁针。阿拉果的实验给法拉第留下了深刻的印象，引起了他不断的思索，他力图揭开这一令人不解之谜。

1823 年，法拉第继续作一系列电磁旋转的实验，通过实验使他想到既然电对磁有作用，一定有磁对电的反应用；既然电流能产生磁，则磁也一定能产生电流。1825 年，他在《科学季刊》上发表了一篇文章《在磁影响下的电流》，他写道：“当磁极绕着载流导线运动时，可能有一种反作用施加在电流上，会产生明显的效应。据各种理由预测，一个强磁铁的磁极接近导线时会减少导线中的电流。”[5]

1831 年 10 月 24 日，在法拉第向皇家学会提交的论文中谈到了探索电磁感应的动机。“……一方面，各种电流都伴随有相应强度的磁作用，它的方向与电流的方向呈直角；而另一方面，若将电的良导体放入有磁作用的环境中，在导体内竟然完全不会引起感应电流，也不产生任何可觉察到的等效于这种电流的作用，这是很不平常的。”“对这些问题及其后果的考虑，再加上想从普通磁中获得电的希望，时时激励着我从实验上去探求电流的感应效应。”[6]

为了实现磁产生电，法拉第从 1824 年起作了一系列实验。1824 年 12 月 28 日，他在日记上写道：“期望一个强磁极接近导线时通过导线的电流会受到影响，以便显示在导线其它部分中的某些反作用效应，但是未觉察到任何这类效应。”他把不同长短不同粗细的铜线或银线作成螺旋线与电流计和电池相联接构成一个回路，把磁极放入螺旋线，并未发现电流计指针的偏转。[2]

1825 年 1 月 28 日，他用载流导线进行感应实验。实验 1：用一根直导线与伏打电池的两极联接，使另一根导线与之平行且相距两张纸的厚度，后者两端与电流计相连，未显示出任何作用。实验 2：把一根螺旋线与电池两极联接，直导线穿过它，后者两端连接电流计，没有影响。实验 3：用一根直导线与电池联接，在直导线上放一根螺旋线，后者两端联接电流计，没有影响。由此他得出结论：“不能以任何方式从这种联线中提供任何明显的效应。”[2]

图 7—6

1828 年 4 月 22 日，法拉第用一根线把一个铜线环固定在导线上并把它悬挂起来，象一个扭秤一样。把强磁棒的极穿入环内，假定铜线环内产生了感应电流，但是把其它磁铁靠近导线，无论放在什么位置，都没有任何效应产生。[2]

年复一年的实验，失败一个接着一个，给他带来了不少的苦脑，但是他并不灰心，他坚信电与磁是相互联系的，磁一定可以转化为电。正在此时，英国物理学家斯特金发明了电磁铁。他在一块原来没有磁性的软铁上绕以导线，通电以后，软铁就变成具有了强磁性的磁铁。后来，美国物理学家亨利改进了斯特金的电磁铁，用彼此绝缘的丝包铜线代替裸铜线，制成能吸引三百公斤铁的电磁铁。这些对法拉第的进一步研究有一定的启发和帮助。

图 7—7

1831 年 8 月 29 日，法拉第在日记中记述了他第一次成功的实验。他在软铁环的 A 边绕了三个线圈，可以串联起来使用，也可以分开使用。在 B 边以同样的方向绕了二个线圈。他把 B 边的线圈接到检流计上，把 A 边的线圈接到电池组上。当电路接通时，法拉第看到检流计的指针立即发生明显的偏转、振荡，然后停止在原来的位置上。这表明线圈 B 中出了感应电流。当电路 A 断开时，他又看到指针向相反方向偏转。把 A 边的三个线圈串联成一个线圈重作以上实验，对磁针产生的效应比以前更加强烈。他看到 B 边的感应电流是明显的，又是瞬时的，只在 A 边断开和接上电源时的瞬间产生。[2]

在第一次发现之后，法拉第继续进行了大量的实验，探讨电磁感应产生的条件。他提出这样的问题：是否可以用其它方法产生同样的效应？铁环是必需的吗？线圈 A 是必需的吗？9 月 24 日，法拉第在两条磁棒的 N、S 极之间放上一条带有线圈的圆铁

图 7—81831 年 8 月 29 日法拉第记录的电磁感应的首次成功实验

图 7—9

棒，线圈与一检流计连接(图 7－9)。他发现当圆铁棒接触 N、S 极和脱离 N、S 极时，检流计的指针就会偏转，他指出这一效应不是永恒的而是瞬时的，“因此，在这里磁转化为电是清楚的”。[2]

10 月 1 日，他把两条长 203呎的丝包铜线绕在木筒上。其中一个线圈和检流计相联接，另一个线圈和电池相联接。他发现当电池接通和断开的瞬间，“对电流计的指针有影响，但是如此之小，以致于很难感觉到。因此在没有铁芯的情形下也有感应效应。”[2]

10 月 17 日,法拉第用另一种方式得到了感应效应.他在直径为 3/4吋，长为$8\frac{1}{2}$时的空心纸筒上绕了八层螺旋线，把八层线圈并联后再接到电流计上，然后把磁铁棒迅速地插入螺线管时，电流计的指针就偏转了，然后又迅速地拉出来，指针在相反的方向上发生了偏转。“每次把磁棒插进或拉出时，这效应都会重复，因此电的波动只是从磁铁的接近而不是磁铁停止在那里产生的。”[2]

10 月 28 日，他把一个空心螺线管迅速送入一对大的磁极之间(图 7－10)，检流计的磁针受到强烈的影响。然后又迅速取出，磁针同样受到强烈影响。这是在磁铁与线圈的相对运动中所产生的一种效应。[2]

图 7-10　图 7-11

从 10 月底到 11 月初，法拉第进行了著名的圆盘实验。他在一个铜轴上安装了扁平的铜盘，把它放在磁铁的两极间，用一根导线从铜轴上引出，另一根导线与铜盘边缘接触，然后把这两根导线与电流计相联接。使铜盘转动时，指针就发生了偏转。当反方向转动时，指针的偏转方向相反。在铜盘继续转动时，指针持续地偏转。[2]这就是一台原始的发电机，通过铜盘的机械转动而产生了电流。这一年圣诞节前夕，法拉第在朋友面前表演了这一实验。当时在场的一位贵夫人取笑地问：“先生，你发明这个玩意儿有什么用呢？”法拉第风趣地回答：“夫人，新生的婴儿又有什么

EXPERIMENTAL RESEARCHES

IN

ELECTRICITY.

By

MICHAEL FARADAY, D.C.L.,E.R.S.

FULLERIAN PROFESSOR OF CHEMISTRY INTHE ROYAL INSTITUTIOS.

CORRESPONDING MEMBER,ETC.OF THE ROYAL AND IMPERIAL ACADEBLLES OF

SCIENCE OF PARIS,PETERSBURGH,FLORENCE,COPENHAGEN,BERLIN,

GOTTINGEN,MODENA,STOCKHOLM,PALERMO,ETC.ETC.

Reprinted from the PHILOSOPHICAL TRANSACTIONS of 1838-1843.

With other Delectrical i’apers

From the QUABTERLY JOURNAL OF SCIENCE and PHILOSOPHICAl MAGAZINK.

VOL.H.

Facsimile-reprint.

LONDONn:

RICHARD AND JOHN EDQARD TAYLOR

PRINTERS AND PUBLISHERS TO THE UXIVERSITY OF LOSDOS.

RED LOS COURT<FLEET STREET.

1844

图 7－12 法拉第写自《电学实验研究》一书的封面

用呢？”[7]

1831 年 10 月 24 日，法拉第在提交给皇家学会的一篇论文中，把产生感应电流的情况概括成五类：变化着的电流；变化着的磁场；运动的稳恒电流；运动的磁；在磁场中运动的导体。在《电学实验研究》19 节中，他还讲到感应电流的方向：“当一条载流导线与另一条与之平行的导线相互接近时，感应电流的方向与施感电流的方向相反。”“它们彼此排斥，反对相互接近。”“当两线离开时，感应电流的方向与施感电流的方向相同。”“它们彼此吸引，反对相互分离。”[6]但这只是确定感应电流方向的一个特例，还没有提出确定感应电流方向的普遍法则。他在《电学实验研究》119 节中指出：当一块金属通过磁极前面或两极之间时，所产生的电流与运动方向成直角。据此理由他解释了阿拉果实验，当圆盘在磁场中旋转时，感应电流的方向近似沿半径方向，在盘内形成闭合的感应电流，即涡电流，这个电流趋向于阻止磁针和圆盘的相对运动，因此磁针就随着圆盘转动起来。

1832 年，法拉第发现在相同条件下不同金属导体中产生的感应电流与导体的导电能力成正比(欧姆定律已在 1827 年得出)，他由此意识到在电磁感应中产生了感应“电动力”(Electrometiveforce，用现在的术语应为电动势)。这个“电动力”与导体的性质无关，只取决于导线和磁力的相互作用。在闭合回路中感应“电动力”产生了感应电流，在开路中没有感应电流，但感应“电动力”还存在。

在法拉第工作的前后，还有不少物理学家从事电磁感应的研究。1829 年美国物理学家亨利(Josoph Honry，1799—1878)在实验不同长度导线电磁铁的提举力时，意外地发现了通电线圈在断开时所产生的强烈的电火花。1832 年，他发表了题为“关于磁生电流与电火花”的论文，宣告了自感现象的发现。他把电池、铜线与水银杯连成一个回路，当铜线不到 1呎时，断开电路觉察不出火花。如果用长达 30呎或 40呎的铜线，把铜线的一端从水银杯中抽出，火花立即产生。如果把导线扭成螺旋线，则断开时，火花更为强烈。这种效应的强弱依据导线的长度、粗细、形状而定。[7]

1832 年，俄国物理学家楞次(H. F. E. Lenz，1804—1865)获悉法拉第发现电磁感应现象后作了许多电磁实验。11 月 29 日发表了题为“论电磁感应引起的电流方向的决定”一文，提出了确定感应电流方向的一般原则。他提出：“如果一个金属导体在一电流或一磁体附近运动，其所产生

的感应电流的方向是这样决定的，感应电流在磁场中受到的作用力与导线的运动方向相反。”[8]

(三)力线概念的发展

在发现电磁感应现象后，为了解释这一现象，法拉第提出了电紧张态(electro-tonicstate)概念。他在 1831 年 10 月 24 日提交给皇家学会的论文中阐述了这一概念的含意。在《电学实验研究》60 节中，他写道：“导线在经受电流的或磁电的感应时似乎处于一种特殊的状态，如果在通常情况下要产生电流的话，这种状态抵抗导线内部电流的产生；当导线不受影响时，这种状态有引起电流的能力，而在通常环境中，这导线不具有这一能力。我把物质的这种状态称为电紧张状态。”[6]

法拉第依据物质的电紧张态假设提出了两个预测。其中之一是：一个环路中的电流在另一个环路的物质中感应出的电紧张态对原电流应该有一个反作用。他写道：“无论这个状态的张力(Tension)是大还是小，不对原电流施加一个反作用并且产生某种平衡是很难设想的，可以预期必定对原电流变化起阻碍作用。”他的第二个预测是：在邻近导线中感应出电紧张态的电流，在它自身的导线中也感应出这种状态。”从而预言了自感现象的存在。[6]

与此同时，他还用磁力线概念来解释电磁感应现象。1831 年 11 月，他在《电学实验研究》中指出：“当导线与电源接通时，磁力线向四周扩张，在它横穿过的导线中产生感应电流；在断开电源时，磁力线向着减弱的电流收缩和返回，因此在相反方向上横穿过导线运动，引起了与第一种情况相反的感应电流。”[6]

他认为用磁力线这一明确概念解释电磁感应现象比用电紧张态这一模糊概念为好。他在《电学实验研究》231 节中指出：“相对于磁铁运动的金属中存在的感应电流取决于金属横切的磁力线，这一定律为我们提供了更精确、更明确的表述。为了对产生的效应提供一个更完善的理由，我放弃我所称为的‘电紧张态’的假定。”[6]

1832 年 3 月 12 日，法拉第写了一封密封的信给英国皇家学会，信封上写着“现在应当收藏在皇家学会档案馆里的一些新观点”，这封信直到 1938 年才被找出来启封公布。[7]法拉第在信中写道：

“……磁作用的传播需要时间，即当一个磁铁作用于另一个远处的磁铁或者一块铁时，产生作用的原因(我以为可以称之为磁)是逐渐地从磁体传播开去的；这种传播需要一定的时间，而这个时间显然是非常短的。

我还认为：电感应也是这样传播的。我以为，磁力从磁极出发的传播类似于起波纹的水面的振动或者空气粒子的声振动，也就是说，我打算把振动理论应用于磁现象，就象对声作的那样，而且这也是光现象最可能的解释。

类比之下，我认为也可以把振动理论运用于电感应。我想用实验来证实这些观点。……”

这表明，在法拉第那里已经孕育着电磁作用传播的波动性质以及它们传播的非瞬时性思想。

20 年后，法拉第进一步发展了他的力线思想。他在 1852 年 1 月 11 日发表的《关于磁力的物理线》一文中，强调力线是一种物理存在。他写道：

“兹举太阳施加给地球的照明或热力为例。在这情况中，射线(即力线)通过中间的空间；但是我们也可以在它们的路径中间用不同介质来影响它们。我们可以用反射或折射变更它们的方向；我们在光源处切断它们，在它们到达目的物之前去寻找和发现它们。它们与时间有关，来自太阳的射线需要 8 分钟才能到达地球，所以它们可在离开它们的来源或老家之后而独立存在，事实上有个明显的物理存在。”[4]

在论述了静电的力线后，他又转向动电的力线。他写道：“至于动电，则物理力线的证据更为确凿得多。与伏打电池相连接的导线，具有人们所讲的环绕电路的力流，但是这种力流具有一对大小相等方向相反的力轴，它所含的力线能根据导体的横向作用而收缩或扩张，并能随着导体的形状而改变方向；它存在于导体的各部分，并能经由适当的途径依我们的目的而从任何地方取出。毫无疑义，它们是物理力线。”[4]

在谈到磁力线的物质承担者时，他指出：“它可能象光线一样靠以太而存在。光与磁已经联在一起了。它的存在可能取决于与磁力密切有关的某种张力状态，振动状态或与电流有关的其它状态。”[4]他认为磁力线依赖于物质才能存在，但不能把物质简单地理解为有质的或有重力作用的物质，他反对超距作用的观点，强调物质之间的电磁力是通过媒介传递的近距作用力。他说：“如果我们假定它须靠以太才能存在而承认以太是属于物质种类的话，那末这种力线可能要靠物质的某些作用才能存在。”[4]

为了定量地具体地描述空间任何区域磁力的本性、情况、方向和大小，他提出了力管的思想。“如果在空间取一个任意小的闭合曲线，与这个闭合曲线相切的力线形成一个返回到自身的管状的表面，这个表面称之为力管。”力管的设想，不只考虑到磁场强度的方向，也考虑到磁场强度的大小，沿着整个管的长度磁场强度和管的切面积的乘积是一个常数，即力线数的总和不变。在这个基础上法拉第设想把整个空间用力管分成许多部分，并使每个力管具有同样确定的值。为了简单起见，每个力管可以称为一个“力线单位”，于是磁场强度就由单位力线的分散和集中来表示，在任何点通过垂直于力管方向上单位面积的力管数就表示了磁场强度。”[9]

法拉第关于力线和场的概念对于电磁学的发展以及整个物理学的发展都是很有影响的。几十年后，约·汤姆孙(J. J. Tho-maon)评论说：“在法拉第的许多伟大的贡献之中，最伟大的一个就是力线概念了，我想电场和磁场的许多性质借助于它就可以简明而富有启发性地表示出来。”[7]场的概念和力线的模型，对当时的传统观念是一个重大的突破，从此超距作用的观念逐渐衰败，新型的近距作用观念日益强化和完善。

(四)电磁感应定律的确立

在法拉第着手通过实验确立电磁感应定律之前，他的同时代人德国物理学家纽曼(F. E. Neumann，1798—1895)在 1845 年首先从理论上导出了电磁感应定律的定量表达式。

纽曼从楞次定律，即在磁场中移动一个导体环路要克服磁力作功这个事实出发，认为如果要把一个导体环路从无限远移到磁场中某一位置，就必须作出一定的功，他把这个功称为势。感应电流仅仅取决于由于运动引起的这个势函数的变化。[9]

设一个磁分子的磁矩 $dM=idS$，磁分子在磁场 B 中具有的势能为

du=B·dM=iB·dS。因此对一个通有电流 i 的环路 L 在磁场 B 中具有的势能为

$$U_i = \iint_s B \cdot dS \qquad (7.4.1)$$

在这里 S 是环路 L 所包围的面积。如果磁场 B 是由通有电流 i' 的环路 L' 产生的，根据毕奥-萨伐尔定律我们有

$$B = i'\int_{L'} \frac{dL' \times r}{r^3} \qquad (7.4.2)$$

利用 $\frac{r}{r^3} = -\nabla\left(\frac{1}{r}\right)$ 及Stokes公式得

$$\begin{aligned} U_i &= ii'\iint_S \int_{L'} \left(\frac{dL' \times r}{r^3}\right) \cdot dS \\ &= ii'\iint_S \int_{L'} \left(\nabla \times \frac{dL'}{r}\right) \cdot dS \\ &= ii'\int_L \int_{L'} \frac{dL \cdot dL'}{r} \end{aligned}$$

我们从式(1)看到势函数等于电流 i 与由 i' 产生的磁场穿过环路 L 的力线数 $\iint_S B \cdot dS$ 的乘积，因为根据法拉第定律，通过环路L的感应电流只取决于穿过环路的力线数的改变。所以纽曼假定，当载有电流 i' 的环路 L' 运动时，在环路 L 中产生的感应电动势正比于该环路的势函数对时间的变化率，即

$$\varepsilon \alpha \frac{dU_i}{dt}$$

$$\varepsilon \alpha i \frac{\partial}{\partial t} \iint_s B \cdot dS \qquad (7.4.4)$$

纽曼引入矢量函数α

$$\alpha = i'\int_{L'} \frac{dL'}{r} \qquad (7.4.5)$$

α是众所周知的矢势，是电流元 i'dL' 到场点距离 r 的函数。将式(3)式(5)代入式(4)，得由电流 i' 的改变在环流 L 中产生的感应电动势

$$\varepsilon \alpha i \int_L \frac{\partial a}{\partial t} \cdot dL \qquad (7.4.6)$$

式(4)、式(6)联立可得

$$\int_L \frac{\partial a}{\partial t} \cdot dL = \frac{\partial}{\partial t} \iint_S B \cdot dS \qquad (7.4.7)$$

上式表明穿过面积S的磁通量 $\iint_S B \cdot dS$ 可以用矢量α对包围该面积的环路L的积分 $\int \alpha \cdot dL$ 来表示。所以麦克斯韦认为矢势α是描述法拉第电紧张态的物理量。可以把α的环路积分看成是导线中产生的电紧张态的量度。

纽曼据楞次定律考虑到感应电流使线圈受到的作用力与线圈在磁场中运动的方向相反，在式(4)式(6)中加上一个负号。又根据法拉第关于感应

电动势与导体性质无关的论断,同时也为了把式(6)写成等式时两边量纲相等，除去式(6)右边的 i，就给出了电磁感应定律定量表示式[10]

$$\varepsilon=-\int\frac{\partial\alpha}{\partial t}\cdot dl$$
$$=-\frac{\partial}{\partial t}\iint_S B\cdot dS$$
$$=-\frac{dN}{dt} \qquad (7.4.8)$$

$N=\iint_S B\cdot dS$ 表示通过S面的磁力线总数，即磁通量。

1846 年，德国物理学家韦伯(W.E.Weber，1804—1890)从另一角度得出了电磁感应定律定量表达式。他从安培的电流元作用定律出发导出了两个运动电荷的相互作用力公式

$$F_{ee}'=\frac{ee'}{r^2}\left[1-\frac{1}{2C^2}\left(\frac{dr}{dt}\right)^2+\frac{r}{C^2}\frac{d^2r}{dt^2}\right] \qquad (7.4.9)$$

式中 $\frac{ee'}{r^2}$ 是库仑力，其它两项是由于e、e'的相对速度 $\frac{dr}{dt}$ 以及相对加速度 $\frac{d^2r}{dt^2}$ 引起的作用力。

他设电流元 i ' dL ' 静止，idL 切割磁力线运动，电荷除了沿导线运动外还有随导线的运动，从而使导线受到附加的作用力。韦伯指出这个附加的作用力正是产生感应电动势的根源。经过详细计算他得出

$$\varepsilon=-\frac{d}{dt}\left(I'\int_L\int_{L'}\frac{dL'\cdot dL}{r}\right)$$
$$=-\frac{d}{dt}\int_L a\cdot dl \qquad (7.4.10)$$

与纽曼的结果相同。但韦伯的讨论只限于动生电动势。韦伯的推导过程表明动生电动势是安培公式的结果，而不是独立的新现象。这为后来动生电动势和感生电动势的区分，为揭示电磁感应现象的本质创造了条件。[10]

五年后，法拉第从实验上验证了从理论上推导出的电磁感应定律的定量表达式，证明了在磁场中运动导线产生的感应电动势只与导线横切的磁力线有关。1851 年 11 月，他在“论磁力线”，即《电学实验研究》3082 节中指出：“无论导线是垂直地还是倾斜地横切磁力线，也无论是沿某一方向或另一方向，该导线大体上以同样的精确性，把它汇总起来的这些力的量用它所横切的力线数来表示。”[11]法拉第所说的总的力的量就是导线中的感应电动势。

1852 年 3 月，他在“关于感应电流在测量磁力的应用”，即《电学实验研究》3177 节中明确提出：“我应当证明感应电流的量精确地正比于运动导线横切的磁力线数。”法拉第是在地磁场中对这一定律进行实验验证的，因为在一个小范围内，地磁场对磁针的作用力具有同样的大小和相同的方向，可以把地磁场看成是一个均匀场。[11]

因为导线在地磁场中切割磁力线时，产生的感应电流很微弱，在《电学实验研究》3123 节中，法拉第重新设计了一个灵敏度较高的电流计。如

图 7－13 所示，他用一条很粗的卷绕

图 7—13

图 7—14

的铜线来代替细的线圈，铜线的直径为 0.2吋，铜线水平地通过下面的磁针，然后在上下磁针之间通过，越过上面的磁针后，再从上下磁针之间通过，导线长达 19呎，这根导线具有很好的导电性，电流计具有很高的灵敏度。对同一感应电流磁针的偏转值要高出许多倍。[11]

从《电学实验研究》第 3192 节到 3199 节，法拉第详细地叙述了电磁感应定律的实验验证。[11]他用的运动导线是一个矩形线圈，先让线圈的平面法线与磁力线平行。然后让它绕 ab 轴转动一周，cd 和 ef 两部分将两次切割面积 cedf 内的磁力线。在转动 180 ° 后，cd 和 ef 两部分切割磁力线的方向相反，它们将产生一个反方向的电流。所以如果头一个电流是从 d 经 ce 和 f 到 d，则第二个电流将从 d 经 fe 和 c 到 d。如果这个矩形不是闭合回路，而是在 b 点断开的开路，那么把 b 点与整流器连接，就能够输出连续的电流到电流计测量。矩形的 ce 和 df 部分在运动中不切割磁力线，所以它们不产生任何电流。

图 7-14

在测量过程中要求线圈转动的时间在磁针(即灵敏电流计的指针)从零位置摆动到最大振幅所需的时间内。在《电学实验研究》 3104 节中说明了这一条件的必要性。如果线圈慢慢地转动，在线圈内部产生微弱的电流；如果它较快地转动，在较短的时间内产生较强的电流。地磁对电流计磁针偏转的作用与感应电流的作用相反，它趋向于使磁针回到零位置。除了磁针摆动到最大位移的时间明显大于产生感应电流的时间外，微弱的感应电流不能使磁针的偏转值达到强电流引起的偏转值。如果磁针摆动到最大偏转值的时间是 10 秒，线圈转动 10 次的时间是 6 秒，所有感应电流的效应都会施加到摆动的磁针上。但是如果转动 10 次的时间是 12 秒或 15 秒, 则有部分感应电流对磁针摆到最大偏转角无贡献。所以要求线圈的转动时间在磁针摆到最大偏转值所需的时间内。这样在转动停止后，能够看到全部感应电流所引起的磁针的最大偏转值。一般选取线圈的转动时间为磁针摆动到最大值所需时间的$\frac{1}{3}$到$\frac{1}{4}$。

在《电学实验研究》3197 节中，他测量了在不同转数不同转速下，每转一周电流计磁针的平均偏转值，实验结果表明它们是比较一致的，精确度接近千分之一。

在 3198 节中，他用了一个边长分别为 8吋和 16吋的矩形圈作实验。与磁力线相切的部分一个是短边，另一个是长边。实验结果表明，只要两个矩形的面积不变，改变矩形的位置和它与导线相切割的部分，矩形每转一周，电流计磁针的偏转值是一致的。

图 7-15

接着他又用了二个周长相同面积不同的矩形线圈与正方形线圈进行比

较。矩形线圈与上述的相同，面积为 128 平方吋，线圈每转一周产生的感应电流引起的磁针的平均偏转角为 2.312°。对于面积为 144 平方吋的正方形线圈，每转一周磁针的平均偏转角为 2.61°。现在面积之比与角度之比非常接近，这就证明感应电流正比于运动导线在单位时间内切割的磁力线数。据前面讲的感应电动势与感应电流的关系，我们可得到如下结论：感应电动势 ε 正比于穿过回路的磁力线通量的变化率，即

$$\varepsilon = -\frac{dN}{dt}$$

这就从实验上证明了电磁感应定律的正确性。

法拉第是位伟大的实验物理学家，他始终把实验作为检验理论与概念的试金石。他不仅在实验上作出了一系列重大发现。在理论上特别是场的概念上也作出了重大贡献。他对自然力的统一性怀有坚定的信念，为揭示自然力间的相互联系和相互转化作出了不懈的努力。法拉第把他的全身心献给了科学事业。他谢绝了重金聘请，全力投入实验研究工作。他的学生和朋友丁铎尔(J. Tyndall，1820—1893)在《作为一个发现者的法拉第》一书中写道：“这位铁匠的儿子，订书商的学徒，他的一生一方面是可以得到十五万镑的财富，一方面是完全没有报酬的学问，要在这两者之间作出选择，结果他选择了后者，终生过着穷困的日子。然而这却使英国的科学声誉比各国都高，获得接近四十年的光荣”。1851 年，法拉第被一致选举为英国皇家学会会长，但他坚决辞掉了这个职务，并说：“我希望我一直保持只有一个称号，这就是法拉第。”

参考文献

[1]Faraday，Experimentai Researches in Electricity，VoL. II，London，I．M. Dent and sons，LTD，159—162

[2] Faraday，Faraday'sDiary，FirstEidtion，London，G·Bell and sons，LTD，1932，48，178，279，310，367—369，372—373，375，381，385

[3]同[1]，127—147

[4][美]威·弗·马吉编，蔡宾牟译，《物理学原著选读》，第 1 版，北京，商务印书馆，1986 年 5 月，463—464，526—531

[5]同[1]，162—163

[6]R. A. R. Tricker，The Contributions of Faraday andMaxwell to Electrical Science，Addison-Wesley，1981，135－136，139，150，28，29，27

[7]申先甲等编著：《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，565，567，576，577

[8]W.F.MAGIE,ASourceBookinPhysics,First Edi-tion,New York，Mcgraw，Hill book Company，1935，513

[9]Whittaker，A History of the theories of Aether and Electricity，London，Lorgmans，Vol.1，1951 年，171—172，198—200

[10]陈秉乾，王稼军，“电磁感应定律的定量表达式是怎样得出来的？”《大学物理》，1987 年 7 月

[11]同[1]，VoL.IV，1332，371，349—350，378—383

八、麦克斯韦电磁场理论的建立

麦克斯韦是继法拉第之后，又一位集电磁学大成于一身的伟大科学家。他全面地总结了电磁学研究的全部成果，并在此基础上提出了“感生电场”和“位移电流”的假说，建立了完整的电磁场理论体系，不仅科学地预言了电磁波的存在，而且揭示了光、电、磁现象的内在联系及统一性，完成了物理学的又一次大综合。他的理论成果为现代无线电电子工业奠定了理论基础。

本文以麦克斯韦三篇主要电磁学论文为依据，试图探讨他的电磁理论建立过程以及他的科学方法的特色，期望能回答以下几个问题：麦克斯韦是怎样用类比研究方法建立他的电磁场理论的？“感生电场”和“位移电流”的假设是怎样提出来的？场的概念和麦克斯韦方程是怎样建立的？

麦克斯韦(James Clark Maxwell，1831—1879)是英国物理学家，诞生于苏格兰的古都爱丁堡。他少年时就受到很好的教育，有良好的思维习惯。15 岁时写了一篇题为“论二次曲线的几何作图”的论文，发表在《爱丁堡皇家学会学报》上。16 岁时考进了爱丁堡大学，专攻数学和物理。19 岁时又进入剑桥大学。1854 年，他在剑桥大学数学竞赛中名列第二。1856 年任苏格兰 Marishal 学院的自然哲学讲座教授。 1860 年至 1868 年任伦敦皇家学院和剑桥大学物理学教授。1870 年创设并主持卡文迪什物理实验室，并担任剑桥大学首任实验物理学教授。他是英国皇家学会会员。他的主要贡献是建立了电磁场理论和气体分子速度分布律。

(一)《论法拉第力线》——电紧张态的数学描述感生电场的提出

在大学期间，麦克斯韦就阅读了法拉第的《电学实验研究》，深深地被法拉第的电磁思想所吸引，他认识到“力线”概念的重要性，也看到法拉第定性表述方面的弱点，决心以数学手段弥补法拉第的不足，把法拉第的天才观念用清晰准确的数学形式表示出来。

1856 年 2 月，麦克斯韦的第一篇电磁学论文《论法拉第力线》(On Farday'sLinesofForce)在剑桥哲学学报上发表了。这篇文章不仅用数学形式解释了法拉第的力线图象，而且包藏着他后来一切新思想乃至麦克斯韦方程的胚胎。因此，法拉第从一开始就对它加以赞扬。在他给麦克斯韦的信中有这样一段话：“我亲爱的先生，接到你的论文，深表谢意。并不是说，我感谢你是因为你谈论了力线，而是因为我知道你已经在哲学真理的意义上处理了它。你的工作使我感到愉快，并鼓励我去作进一步的思考。但当我知道你要构造一种数学形式来针对这样的主题时，起初我几乎是吓坏了。然后我才惊讶地看到这个主题居然处理得如此之好！”[1]法拉第认为这位年青人是真正理解他的物理思想的人，并鼓励他要继续探索，有所突破。

论文一开始，麦克斯韦就环顾了电磁学研究的现状，指出已经建立了很多实验定律和数学理论，但未能揭示各种电磁现象之间的联系。他写道：“电科学的现状看来特别不便于思索。”他认为“有效的科学研究的第一步必须把已有的研究成果简化和归纳成一种思维易于领会的形式”，“寻找一些在思维发展的每一步保持清晰物理概念的研究方法”。[2]他认为“物理类比”就是这一研究方法。

麦克斯韦指出：“物理类比(physical analogy)的意思是利用

THE
SCIENTIFIC PAPERS
OF
JAMES CLERK MAXWELL
M.A,LL.D. EDIN, D.C.L, F.R.SS. LONDON AND EDINBCRGB,
HONORARY FELLW OF TRINITY COLLEGE,
CAVENDISH PROFESSOR OF EXPERMENTAL PHYSICS IN THE UNIVERSITY
OF CAMBRIDCE

EDITED BY
W.D. NIVEN, M.A,, F.R.S.,
DIRECTOR OF STUDIES AT THE ROYAL NAVAL COLLEGE, GREENWICH;
FORMERLY FELLOW OF TRINITY COLLEGE.

VOL.I.

PARIS
LIBBAIBIE SCIENTIFOQCE J. HEBMANN

图 8—1

一种科学定律和另一种科学定律之间的部分类似性，用它们中的一个去说明另一个。”“类比是建立在两类定律在数学形式上相似的基础上。”[2]类比可以构通不同领域的研究方法，可以在解析的抽象形式和假设之间提供媒介，还可以启发新的物理思想，帮助人们去认识和发现一些尚待研究的物理过程和规律。

在这方面，W·汤姆孙的研究给麦克斯韦以很大的启示。1842 年，当汤姆孙还是剑桥大学的学生时，就把包含带电导体的区域内的静电力分布与无限固体中的热流相比较，指出前一种情形下的等势面与后一种情形下的等温面相对应，前者的电荷与后者的热源相对应；与距离成平方反比的吸引力方程与均匀媒质中的均匀热流方程相对应。[3]由于热传导是通过连续媒质中相邻粒子之间的相互作用进行的，这就启发人们思索是否可以把电作用看作是经某种连续媒质传递的呢？

1846 年，汤姆孙获得了剑桥大学学位后，研究了电现象和弹性现象的类似性。他考察了处于应力状态的不可压缩弹性固体的平衡方程，指出表示弹性位移的矢量分布可与静电系统的电力分布相比拟。而且，弹性位移还可以同样好地与一个通过 B 由 Curl α=B 定义的矢量α相一致。这里α是与纽曼，韦伯等人在电流感应论文中引用的矢势(vector potential)等价。但汤姆孙是在不知道这一等同性时，经过不同途径，独立得到的。[3]

麦克斯韦的《论法拉第力线》采用一种几何学的观点，为法拉第的力线概念作出了数学描述。他在描述这些力线的构成时说：“如果我们从任意一点开始画一条线，并且当我们沿这条线走时，线上任意一点的方向，总是和该点合力的方向重合，那么这条曲线就表示它所通过的各点的合力

的方向，在这个意义上才称为力线。用同样的方式我们可以画出其它的力线，直到曲线充满整个空间以表示任一指定点力的方向。”[2]

这篇论文可分为两部分。第一部分阐述了力线和不可压缩流体之间的类比。为了对法拉第的观念作出精密的数学处理，把这一个物理图象表示为清晰的几何图象，麦克斯韦采用了类比的思考方法。他把力线和不可压缩流体的流线加以比较，因为流体的速度方向与流线的切线方向相同且反比于流管的截面积，所以力的强度也反比于力管的截面积，这就获得了一种既表示力的强度又表示力的方向的几何模型。

麦克斯韦着重分析了不可压缩流体通过有阻力介质的运动。这个阻力取决于介质的性质，与运动方向相反，与速度大小成正比。如果用 v 表示速度大小，K 表示阻力系数，则作用在单位流体体积上的阻力为 Kv。因此，为了保持这一速度，在流体的后一部分必须比它的前一部分保持较大的压强，这一压强差可以抵消阻力的效应。显然，压强相等的面必定垂直于流线。

考虑一个单位立方体，让流体在垂直于它的两个面的方向上流动。在这个运动方向上单位长度上的压强差为 Kv，如果长度为 h，则压强差为 Kvh。在单位流管上取两个压强差为 1 的等压面，如果这两个等压面间沿流线的长度为 h，则 Kvh=1。据单位流管的定义 vS=1，由此我们得出 S=Kh。他把截面积为 S 长为 h 的单位流管称为单位元(unit cell)。在每个单位元中，在单位时间内通过单位流体体积时压强从 p 降到 p-1，因此克服阻力所消耗的功在每个单位时间，每个单位元中是一个单位。

麦克斯韦把流线的开始处称为源(Source)，流线的终止处称为穴(Sink)。他讨论了一块中间嵌有一个单位流体元的各向同性的无限大的均匀介质，以此流体源为球心，在单位时间内从任意一个球面流出一个单位体积的流量，与源距离r处的速度为$v=\frac{1}{4\pi r^2}$，因此压强的减少率$Kv=\frac{K}{4\pi r^2}$。因为在无限远处压强$p=0$，所以在距离流体源为r处的压强$P=\frac{K}{4\pi r}$，显然对于一个单位穴产生的压强为$-\frac{K}{4\pi r}$，对于N个单位源产生的压强为$\frac{KN}{4\pi r}$。

麦克斯韦把流体源产生的流线与点电荷产生的力线加以类比。因为点电荷之间的作用力与距离的平方成反比，所以点电荷产生的电场强度与流体源在流体中产生的速度相对应，从而得出：流体中的压强与静电电势相对应；流体中的压力梯度与电势梯度相对应。

麦克斯韦把流体运动的理论应用到电流的情形。设在导体环路中有一个均匀流动的电流，由于导体中电阻的存在，为了克服阻力，导体环路中必须有电动力存在。流体中的压强相应于导体环路中电的张力(electrical tension)，物理上等同于静电电势。流体中某方向的压力减少率，相应于沿该方向该点的电动力。他明确指出为了在一闭合回路中产生稳恒电流，除了静电力外还必须有其它作用力存在，他把这个力称为环路的电动力(即现代所说的电动势)。

当不可压缩流体由一种介质进入另一种不同阻力的介质时，流动是连

续的，从一种介质中流出的流量等于在另一种介质中流入的流量。但是通过介质的界面后，两旁存在的压强是不同的，而在两种介质的边界面上法向速度是相同的，因此压强的减少率正比于阻力。如果在介质的边界面上在一个单位流管进入介质处用一个单位源代替，在一个单位流管流出处用一个单位穴代替，则介质中流体的运动将和以前一样。他用这种在边界面上引入源和穴的形式，获得了在两种介质中压强的分布。

接着，他把流体通过边界面的理论应用到电介质中，对于不同的介质，阻力是不同的，如果这个阻力较小，就使我们获得一种类似于更容易传导力线的性质。显然，在这种情形下介质的表面上总是有一个明显的电的分布。在力线进入的地方有负电荷分布，在力线出来的地方有正电荷分布。在流体的情形下，在界面上没有真实的源，我们使用它们只是为了计算的目的。在介质内可能没有电荷存在，但是由于表面极化电荷的存在，有明显的电作用存在。

麦克斯韦在描述流线与力线时，涉及到“量”(quantity)和“强度”(intensity)的术语。他在把流体运动和电流加以类比时，着重指出区别这两个术语是很有必要的。对于流体的情形，前者表示“流量”(fluxes)，后者表示克服阻力的“力”(forces)。对于电流的情形，电流的量指的是在单位时间内通过一个截面的电量。电流的强度指的是电动力，即克服阻力的能力(Power)。对于磁的情形，在一个磁体内任何截面上的磁化量用穿过该面的磁力线条数来表示，磁化强度取决于穿过该面的力线数和阻力。在把力学量与电学量类比的基础上，麦克斯韦对矢量场的量和强度进一步从数学上予以区分，场中的“量”，即力线的通量，可以通过面积分求出，场中的“强度”可以通过线积分求出。

论文的第二部分主要是讨论了电磁感应现象。在第一部分的基础上，他对法拉第的力线观念作出了精密的数学处理，确立了各电磁量之间的相互联系，建立了对于电紧张态的数学描述，提出了感生电场的概念。

首先，麦克斯韦全面地阐述了法拉第的电磁感应定律。他写道：“当导体在电流或磁铁附近运动时，或者当电流或磁铁在导体附近运动时，或者在电流强度变化或磁场强度改变时，在导体中就产生了一个电动力(electromotive forces，即现在所说的电动势)。”麦克斯韦注释：“必须把电动力与电磁力(electromagneticforces)加以区别,电磁力趋向于产生导体的运动，电动力趋向于产生电流。如果导体环路是闭合的，就有连续的电流产生；如果导体环路是开放的，在它的端部就产生电的张力(tension，即现在所说的电压)。如果在闭合导体横切磁感应线的运动中，通过环路的力线数不变，整个环路的电动力就为零。”“无论用什么方法，只要使得通过环路的力线数改变，环路中就有电动力产生，而感应电流所产生的磁力线总是反抗磁力线的变化。”[2]

接着麦克斯韦探讨了法拉第电紧张态的起源，阐述了法拉第把电磁感应现象与状态概念联系起来的思索过程。他写道：“导体环路内的感应电流只是由于它周围的电的或磁的现象的改变而产生的，只要这些不变，在导体中就没有任何效应产生。当导体接近电流或磁铁，以及远离它的影响时是处在不同的状态中。”他指出，法拉第关于取决于力线数改变的电动力是由于状态的改变产生的，该状态由力线数目来确定。在磁场中，一个闭合导体处在由磁作用引起的一定的状态中，只要这种状态不变，就没有

效应发生，但是当这个状态改变时，电动力就产生了，它的强度与方向取决于状态的改变。这一状态就是法拉第所说的“电紧张态”。但是这些概念和想法还没有作为数学研究的课题。麦克斯韦强调：“电紧张态是电磁场的运动性质，它具有确定的量，数学家应当把它作为一个物理真理接受下来，从它出发得出可通过实验检验的定律。”〔2〕

接着麦克斯韦讨论了电流的量与强度的关系即电流密度与电动力的关系，在定义了电流密度概念后，他引进了某一点的电动力的概念。他写道：“导体中任何一点的电流是由于作用在那个点的电动力产生的。这些电动力可能是外部的，也可能是内部的。外部的电动力或者是由于电流和磁铁的相对运动引起，或者是来自它们的强度的改变，或者是来自于其它的原因。内部的电动力是由于导体中相邻点之间的电张力之差引起的。

令 P_2 表示任意点的电张力；X_2、Y_2、Z_2 表示该点电动力在坐标轴 X、Y、Z 上的分量；α_2、β_2、γ_2 表示该点有效电动力的分量。它们有如下的关系

$$\begin{aligned}\alpha_2 &= X_2 - \frac{dP_2}{dX}\\ \beta_2 &= Y_2 - \frac{dP_2}{dY}\\ \gamma_2 &= Z_2 - \frac{dP_2}{dZ}\end{aligned} \qquad (8.1.1)$$

如果用 a_2、b_2、c_2 分别表示沿坐标轴 X、Y、Z 方向上的电流密度，K_2 表示均匀各向同性介质的阻力，则有

$$\alpha_2=K_2a_2，\beta_2=K\ b_2，\gamma_2=K_2c_2 \qquad (8.1.2)$$

α_2、β_2、γ_2 可以认为是坐标轴 X、Y、Z 方向上的电作用强度。沿任一线元 $d\sigma$ 测量出的强度由下式给出

$$E=l\alpha+m\beta+n\gamma \qquad (8.1.3)$$

式中 l、m、n 是这一线元的方向余弦。

对于一给定曲线积分 $\int Ed\sigma$ 表示沿该线的总的强度。如果这曲线是闭合曲线,它就表示在这条闭合曲线中电动力的总强度°将方程(8.1.1)中的 α、β、γ 值代入式(8.1.3)，等式两端取积分可得

$$\int Ed\sigma = \int (Xdx + Ydy + Zdz) - P + C$$

对于闭合曲线

$$\int Ed\sigma = \int (Xdx + Ydy + Zdz)$$

由此得出在一条闭合曲线内有效电动力的总强度等于从外部施加的电动力的总强度。

通过任何面电流的总量用 $\int edS$ 表示，在这里 $e = la + mb + nc$，所以

$$\int edS = \iint adydz + \iint bdzdx + \iint cdxdy$$

这个积分在给定的面上是有效的。当这个面是一个闭合面的时候，我们可以通过各部分的积分求出

$$\int edS = \iiint\left(\frac{da}{dx} + \frac{db}{dy} + \frac{dc}{dz}\right)dxdydz$$

如果我们令

$$\frac{da}{dx} + \frac{db}{dy} + \frac{dc}{dz} = 4\pi\rho \qquad (8.1.4)$$

则 $$\int edS = 4\pi\iiint \rho dxdydz$$

在这里方程右边对 e 的面积分等于在这个面包围的空间内对 4π ρ 的体积分。在包括均匀电流情形的一大类现象中，量 ρ 为零。

接着麦克斯韦讨论了磁场强度和电流的量的关系。他在电张力状态理论总结定律三中写道：“绕着任一曲面边界的总磁场强度

图 8—2

量度了通过该表面的电流量。”即现在所讲的安培环路定律

$$\oint H\cdot dl = \sum I$$

麦克斯韦认为据此定律可以求出穿过一个环路的电流量。

把上述定律应用于面积元 dy dz，设面元中心的坐标为 x、y、z，该点的磁化力强度(inten-sity of magnetizing force 即磁场强度)沿坐标轴 X、Y、Z 方向的分量为 α_1、β_1、γ_1，于是围绕这四个边测量的总的磁化力强度分别为

$$+\left(\beta_1 + \frac{d\beta_1}{dz}\frac{dz}{2}\right)dy \qquad -\left(\gamma_1 + \frac{d\gamma_1}{dy}\frac{dy}{2}\right)dz$$

$$-\left(\beta_1 - \frac{d\beta_1}{dz}\frac{dz}{2}\right)dy \qquad +\left(\gamma_1 - \frac{d\gamma_1}{dy}\frac{dy}{2}\right)dz$$

$$\text{总的强度} = \left(\frac{d\beta_1}{dz} - \frac{d\gamma_1}{dy}\right)dydz$$

通过面元 dydz 传导电流的量为 a_2dydz。因此如果我们用一个围绕着闭合曲线的总磁化力强度来量度电流的量，我们就有

$$a = \frac{d\beta_1}{dz} - \frac{d\gamma_1}{dy}$$

$$b = \frac{d\gamma_1}{dx} - \frac{d\alpha_1}{dz}$$

$$c = \frac{d\alpha_1}{dy} - \frac{d\beta_1}{dx} \qquad (8.1.5)$$

因此只要我们知道磁场强度(magnetic intensities) α、β、γ 的值，我们就能求出电流的分布。我们看到通过对上述方程的微分可以得到

$$\frac{da_2}{dx}+\frac{db_2}{dy}+\frac{dc_2}{dz}=0$$

这就是闭合电路的连续方程。

接着麦克斯韦作了一系列论证。他证明了如果 a_1、b_1、c_1 是磁感应量的三个分量，ρ_1 是磁密度，满足

$$\frac{da_1}{dx}+\frac{db_1}{dy}+\frac{dc_1}{dz}+4\pi\rho_1=0$$

可以找到三个函数 α_0、β_0、γ_0 和第四个函数 V 使其满足

$$\begin{aligned} a_1 &= \frac{d\beta_0}{dz}-\frac{d\gamma_0}{dy}+\frac{dV}{dx} \\ b_1 &= \frac{d\gamma_0}{dx}-\frac{d\alpha_0}{dz}+\frac{dV}{dy} \qquad (8.1.6) \\ c_1 &= \frac{d\alpha_0}{dy}-\frac{d\beta_0}{dx}+\frac{dV}{dz} \end{aligned}$$

$$\frac{d^2V}{dx^2}+\frac{d^2V}{dy^2}+\frac{d^2V}{Dz^2}+4\pi\rho_1=0 \qquad (8.1.7)$$

如果 a_1、b_1、c_1 表示磁感应强度 B 的三个分量，α_0、β_0、γ_0 就是法拉第电紧张态函数 α 的三个分量，据矢量分析中旋度和梯度的定义可得

B=curla+gradV (8.1.8)

当不存在磁极时，此式右边第二项可被变量的适当变换略去，可以得到

B=curla (8.1.9)

这是与麦克斯韦在这篇论文的电紧张态的理论总结中的叙述完全等效的。他在该总结的定律一中指出：“绕着一个面边界的整个电紧张强度等于通过那个面的磁感应量，或者换句话说，通过那个面的磁力线的数量。”用现代的形式表示就是

$$\oint_\Lambda \alpha\cdot\delta\lambda=\int_s\cdot B\cdot\delta\Sigma=\phi \qquad (8.1.10)$$

式中 φ 表示磁力线的数量。

麦克斯韦在定律五中写道：“一个闭合电流总的电磁势(即总能量)是由电流的量(即电流密度)与同方向绕该环路的电紧张强度的乘积所量度。”用现代的形式表示即为

$$W=\oint_L j\cdot adL$$

麦克斯韦在定律六中写道：“任何导体元上的电动力，无论是在大小还是在方向上由在那个元上电紧张强度的瞬时变化率来量度。这里所说的电动力就是变化磁场引起的感生电场。”用现代形式表示即为

$$E=-\frac{\partial a}{\partial t} \qquad (8.1.11)$$

对感生电场在一个闭合回路上进行积分就得到感生电动势。所以麦克斯韦写道：

“在一个闭合导体上的电动力(即电动势)是由绕着这个环路的整个电紧张强度对时间的变化率来量度。虽然产生的电流是随电阻而改变，但是电动力不依赖于导体的性质；而且，无论用什么方式使得电紧张强度改变，或者是由于导体的运动，或者是由于外部环境的改变，产生的电动力都是相同的。”[2]

(二)《论物理力线》——磁力及磁场能量感生电场与变化磁场的关系位移电流概念的提出

1862 年，麦克斯韦发表了《论物理力线》(OnPhysicallinesof force)这篇重要论文。他觉得要更好地体现法拉第的力线思想，仅仅从几何学方面讨论力线是不够的，还应该对已经确立的电学量和磁学量之间的关系给以物理解释，为此目的他设计了电磁作用的力学模型，试图用力学观点解释这些现象。他写道：“在这篇文章中，我的目的是研究介质中的张力和运动的某些状态的力学结果，并将它与观察到的磁和电的现象加以比较，来澄清这方面的思考。”[4]

在论文的第一部分，应用于磁现象的分子涡旋理论中，麦克斯韦通过他所提出的分子涡旋假设讨论了磁场作用在磁极上，作用在磁感应物质上以及作用在电流上的力。

在这以前，“以太旋涡”的思想早已有了。1856 年，汤姆孙从研究光的偏振面在磁场中的旋转效应得出磁具有旋转的特征。认为可以把磁致旋光效应归结为以太振动和分子旋转运动之间的耦合，这给麦克斯韦以很大的启发，使他认识到磁是一种旋转的效应。他写道：“对于由电流引起的电离质在一定方向上的传送和由磁力引起的偏振光在—定方向上的转动这一事实的思考，导致我把磁认为是一种旋转现象。”[3]

麦克斯韦把磁旋转这一概念与法拉第的力线思想相联系。按照法拉第的力线思想，力管倾向于纵向收缩和横向膨胀。他想，如果假设每个力管所包含的流体是处在绕它的管轴的转动中，这样一种倾向就可以归因于离心力。于是他设想了一个“分子涡旋”(molecularvortex)模型，假设涡旋绕磁力线旋转，即从 S 极到 N 极沿磁力线看去，涡旋在顺时针方向旋转，由于旋转引起的离心力使每个涡旋在横向扩张，从而纵向收缩，因而磁力线在纵向表现为张力，就象绳上的拉力一样。横向表现为压力。

图 8—3

麦克斯韦假设在场中任何一部分的所有涡旋是围绕几乎平行的轴在相同的方向上以相同的角速度转动。磁的影响是作为介质中的压力或张力形式而存在。这种压强不同于通常流体的压强，在介质中每一点在不同方向上的压强是不同的，在垂直于轴线方向上的压强是相等的，且具有最大值；最小的压强在平行于轴线的方向上。如果设轴上的压强是 p_0，在涡旋圆周上的压强将是$p_1 = p_0 + \frac{1}{2}\rho v^2$。$\rho$ 是涡旋物质的密度，v是每个涡旋周围的线速度。p_1 的方向垂直于轴线。在平行于轴线的方向上，平均压强为

$p_2 = p_0 + \frac{1}{4}\rho v^2$。因此最大压强与最小压强之差$p_1 - p_2 = \frac{1}{4}$ ρ v_2。这一压强差就是在轴线上的张力。他用$\frac{\mu}{\pi}$代替ρ得到到下式

$$p_1 - p_2 = \frac{1}{4\pi}\mu v_2 \qquad (8.2.1)$$

麦克斯韦认为实际应力可以分解成作用在所有方向上的简单的流体压力p_1和沿着应力轴作用的简单张力$p_1 - p_2$或$\frac{1}{4\pi}\mu v^2$。如果涡旋轴线相对于 X、Y、Z 轴的方向余弦是 l、m、n，则平行于三条轴的法向应力 p_{xx}，p_{yy}，p_{zz}和在三个坐标面上的切向应力 p_{yz}、p_{zx}、p_{xy}应为

$$px = \frac{1}{4\pi}\mu v^2 l^2 - p_1 \qquad p_{y2} = \frac{1}{4\pi}\mu v^2 mn$$

$$p_{yy} = \frac{1}{4\pi}\mu v^2 m^2 - p_1 \qquad p_{zx} = \frac{1}{4\pi}\mu v^2 nl$$

$$pzz = \frac{1}{4\pi}\mu v^2 n^2 - p_1 \qquad p_{xy} = \frac{1}{4\pi}\mu v^2 lm$$

如果我们令α=vl，β=vm，γ=vn，则有

$$p_{xx} = \frac{1}{4\pi}\mu\alpha^2 - p_1 \qquad p_{yz} = \frac{1}{4\pi}\mu\beta\gamma$$

$$p_{yy} = \frac{1}{4\pi}\mu\beta^2 - p_1 \qquad p_{zx} = \frac{1}{4\pi}\mu\gamma\alpha$$

$$p_{zz} = \frac{1}{4\pi}\mu\gamma^2 - p_1 \qquad p_{yz} = \frac{1}{4\pi}\mu\alpha\beta$$

由于内部应力的改变，在体积元内将引起一个合力，麦克斯韦根据应力的平衡定律得出每单位体积在 X 轴方向的力为

$$X = \frac{d}{dx}p_{xx} + \frac{d}{dy}p_{xy} + \frac{d}{dz}p_{xz}$$

代入以上数据，经计算得

$$X = \alpha\frac{1}{4\pi}\left(\frac{d\mu\alpha}{dx} + \frac{d\mu\beta}{dy} + \frac{d\mu\gamma}{dz}\right) + \frac{1}{8\pi}\mu\frac{d}{dx}\left(\alpha^2 + \beta^2 + \gamma^2\right)$$

$$-\mu\beta\frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right) + \alpha\gamma\frac{1}{4\pi}\left(\frac{d\alpha}{dz} - \frac{d\gamma}{dx}\right) - \frac{dp_1}{dx} \qquad (8.2.2)$$

接着麦克斯韦对上式中每一项进行物理解释。他假设α、β、γ是磁场强度，即磁场对单位北磁极的作用力在坐标轴 X、Y、Z 方向的分量，它们与涡旋周围的速度成正比。μ表示介质中任何一点的磁导率，与涡旋的密度成正比。μα、μβ、μγ表示通过垂直于 X、 Y、 Z 轴的单位面积上的磁感应强度。

通过围绕着一个磁极的闭合面的磁感应总量完全取决于闭合面内的磁物质的量。所以如果 dxdydz 是一个体积元，m 是磁极强度，即单位体积中北极磁物质的量，则有

$$\left(\frac{d}{dx}\mu\alpha+\frac{d}{dy}\mu\beta+\frac{d}{dz}\mu\gamma\right)dxdydz = 4\pi mdxdydz \qquad (8.2.3)$$

因此 X 值的第一项

$$\alpha\frac{1}{4\pi}\left(\frac{d}{dx}\mu\alpha+\frac{d}{dy}\mu\beta+\frac{d}{dx}\mu\gamma\right)$$

可以写为 $\alpha m \qquad (8.2.4)$

这一项的物理解释是，在 X 轴正方向推动磁北极的力是在 X 轴方向上的磁场强度和磁极强度的乘积。

然后麦克斯韦按照分子涡旋假设对磁场作用于磁极上的力进

图 8—4

行力学解释，给出了定性的物理图象。他假设有一组磁力线是从左向右的平行线。放置在磁场中的磁北极 N 向外发出力线，在 N 极的右边与原磁场方向一致，使磁场得到加强；在 N 极的左边与原磁场方向相反，使磁场受到削弱。因此这个场的涡旋的转速在 N 极的右边将加快，在 N 极的左边将减慢。从而使 N 极右边的张力将比左边的张力大，所似磁北极将沿着场的方向被拉向右边。在同样的意义上，南极将被拉向左边。

X 值的第二项是

$$\frac{1}{8\pi}\mu\frac{d}{dx}\left(\alpha^2+\beta^2+\gamma^2\right)$$

在这里 $\alpha^2+\beta^2+\gamma^2$ 是场中任意一点的磁场强度的平方，μ 是同一地方的磁导率。这一项的物理意义是在 X 轴方向上磁场作用在磁感应物质上的力。对于磁导率 μ 是正的物体将被推向强磁场的地方。

X 值的第三项是

$$-\mu\beta\frac{1}{4\pi}\left(\frac{d\beta}{dx}-\frac{d\alpha}{dy}\right)$$

在这里 $\mu\beta$ 是通过垂直于 Y 轴的单位面积上的磁感应量，即磁感应强度在Y轴上的分量。据第一篇论文中式(5)得到的结果，$\frac{1}{4\pi}\left(\frac{d\gamma}{dy}-\frac{d\beta}{dz}\right)$表示通过垂直于 X 轴的单位面积上的电流强度。如果我们令通过垂直于 X、Y、Z 轴的电流密度分别为 p、q、r，则有

$$\frac{1}{4\pi}\left(\frac{d\gamma}{dy}-\frac{d\beta}{dz}\right)=p$$

$$\frac{1}{4\pi}\left(\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\right)=q$$

$$\frac{1}{4\pi}\left(\frac{d\beta}{dx}-\frac{d\alpha}{dy}\right)=r \qquad (8.2.5)$$

第三项的物理解释是 Y 轴方向的磁感应强度 μ β 对 Z 轴方向的电流密度 r 的作用力沿 X 轴负方向。

图 8—5

麦克斯韦继续用分子涡旋假设对磁场作用在电流上的力给以力学解释。他假设一垂直于纸面的长直电流放在一垂直向上的均匀磁场中，由电流产生的磁力线使它右边的磁场加强，而使左边的磁场削弱。因此在导线右边的涡旋转速加快，给导线向左的压力加大；在导线左边，涡旋转速减慢，给导线向右的压力减小，其结果将把导线推向左边。

X 值的第四项是

$$+\mu\gamma\frac{1}{4\pi}\left(\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\right)\text{或}+\mu\gamma q$$

这一项的物理解释是 Z 轴方向的磁感应强度 μ γ 对 Y 轴方向的电流密度 q 的作用力沿 X 轴正方向。它与第三项一样都是指的磁场对电流的作用力。

X值的第五项 $-\frac{dp_1}{dx}$ 表示这个体积元将被推向流体压力减小的方向。

现在我们可以写出磁场作用在单位体积介质元上的合力的分量的表示式为

$$\left.\begin{aligned}X&=\alpha m+\frac{1}{8\pi}\mu\frac{d}{dx}v^2-\mu\beta r+\mu rq-\frac{dp_1}{dx}\\Y&=\beta m+\frac{1}{8\pi}\mu\frac{d}{dy}v^2-\mu\gamma p+\mu\alpha r-\frac{dp_1}{dy}\\Z&=\gamma m\frac{1}{8\pi}\mu\frac{d}{dz}v^2-\mu\alpha q+\mu\beta p-\frac{dp_1}{dz}\end{aligned}\right\}(8.2.6)$$

每个表示式中，第一项表示磁场作用在磁极上的力；第二项是作用在磁感应物质上的力；第三项和第四项是作用在电流上的力；第五项是简单的压力效应。

对于电流不存在，即 p、q、r 为零的情形，据式(5)有

$$\frac{d\gamma}{dy}-\frac{d\beta}{dz}=0\qquad\frac{d\alpha}{dz}-\frac{d\gamma}{dx}=0\qquad\frac{d\beta}{dx}-\frac{d\alpha}{dy}=0$$

因此有$\alpha dx+\beta dy+\gamma dz=d\varphi$，即$\alpha dx+\beta dy+\gamma dz$是$\varphi$的全微分，所以

$$\alpha=\frac{d\varphi}{dx}\qquad\beta=\frac{d\varphi}{dy}\qquad\gamma=\frac{d\varphi}{dz}\qquad(8.2.7)$$

据式(3)有

$$\begin{aligned}m&=\frac{1}{4\pi}\left(\frac{d}{dx}\mu\alpha+\frac{d}{dy}\mu\beta+\frac{d}{dz}\mu\gamma\right)\\&=\frac{1}{4\pi}\mu\left(\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}+\frac{d^2\varphi}{dz^2}\right)\end{aligned}\qquad(8.2.8)$$

如果在我们所考虑的空间无磁极物质，即 m=0，则

$$\frac{d^2\varphi}{dx^2}+\frac{d^2\varphi}{dy^2}+\frac{d^2\varphi}{dz^2}=0\qquad(8.2.9)$$

在 μ 是均匀的而且没有电流的空间中，磁力线必定是来自于磁物质。如果设想在与我们所考虑的空间点距离 r 处有一磁极 m，则上式唯一的解是

$$\varphi = -\frac{m}{\mu}\frac{1}{r}$$

单位北磁极在 r 处受到的排斥力是

$$\frac{d\varphi}{dr} = \frac{m}{\mu}\frac{1}{r^2} \qquad (8.2.10)$$

这就是磁库仑定律。

在论文的第二部分，应用于电流的分子涡旋理论中，讨论了电磁感应现象。这就要求对电流与分子涡旋的物理联系有所理解。而它又引起了另一个问题，由于相邻涡旋的表面在空间任意一点运动方向相反，这两个相邻的涡旋如何在相同的方向上自由转动呢？麦克斯韦想到在机械齿轮机构中有一种惰轮(idlewheel)，它们与两边的齿轮相互啮合，可以保证两边的齿轮在相同的方向上旋转。这个惰轮的中心是可以运动的，其中心的运动是两边齿轮周边运动之和的一半。麦克斯韦设想每个涡旋同它相邻的涡旋被一层细微的粒子隔开，这些细微的粒子起着齿轮系列中可动惰轮的作用，这些粒子远比涡旋的线度小，它们的质量与涡旋相比微不足道。由于磁体的磁力线可长期保持而不消耗，因此粒子作无滑动的滚动。粒子与涡旋的作用沿切线方向。麦克斯韦通过计算得到这些粒子的运动与电流相对应，穿过单位面积的粒子数等于流过单位面积的电量。因此，电流由这些穿插在相邻涡旋之间的粒子的移动来表示。粒子的滚动带动涡旋旋转，好象齿条带动齿轮。这就是电流产生磁力线的类比机制。

然后麦克斯韦计算了介质中涡旋运动引起的能量。他假设单位体积中涡旋的实际能量正比于密度和速度的平方，即

$$E = C\mu(\alpha^2 + \beta^2 + \gamma^2) \qquad (8.2.11)$$

式中 C 是待确定的常数。在磁场中没有电流的情形下，据式(7)有

$$\alpha = \frac{d\varphi}{dx} \qquad \beta = \frac{d\varphi}{dy} \qquad \gamma = \frac{d\varphi}{dz}$$

令$\varphi=\varphi_1+\varphi_2$据式(8)有

$$\frac{\mu}{4\pi}\left(\frac{d^2\varphi_1}{dx^2} + \frac{d^2\varphi_1}{dy^2} + \frac{d^2\varphi_1}{dz^2}\right) = m_1$$

$$\frac{\mu}{4\pi}\left(\frac{d^2\varphi_2}{dx^2} + \frac{d^2\varphi}{dy^2} + \frac{d^2\varphi_2}{dz^2}\right) = m_2$$

φ_1 表示磁极 m_1 在空间任何点产生的势，φ_2 是磁极 m_2 产生的势。所有涡旋的实际能量是 E 式对整个空间进行积分

$$E=\Sigma C\mu(\alpha^2+\beta^2+r^2)dV$$

通过分部积分可以证明

$$\varphi_1 E=-4\pi C\Sigma(\varphi_1 m_1+\varphi_2 m_2+\varphi_1 m_2+\varphi_2 m_1)dV$$

又因 $\Sigma\varphi_1 m_2 dV=\Sigma\varphi_2 m_1 dV$

则 $$E=-4\pi C\Sigma(\varphi_1 m_1+\varphi_2 m_2+2\varphi_1 m_2)dV \qquad (8.2.12)$$

现在令磁场 m_1 处于静止状态，m_2 通过空间 δx 的距离，因为φ_1只取决于 m_1，它仍然与以前一样，所以$\varphi_1 m_1$是常数。因为φ_2只取决于 m_2，在 m_2 的周围φ_2

的分布仍然是相同的，所以$\varphi_1 m_1$仍然与以前一样。由于m_2的位移

φ变成了$\varphi_1+\frac{d\varphi_1}{dx}\delta x$,所以式中的$2\varphi_1 m_2$发生了改变.因此,由于这个位移实际的能量变化为

$$\delta E=-4\pi C\sum\left(2\frac{d\varphi_1}{dx}m_2\right)dV\delta x$$

在这个运动中，磁场强度$\frac{d\varphi_1}{dx}$作用在磁物质m_2上作的功为

$$\delta W=\sum\left(\frac{d\varphi_1}{dx}m_2 dV\right)\delta x$$

根据能量守恒有$\delta E+\delta W=0$,即磁场对运动磁铁所作的功等于涡旋能量的减少。所以，

$$-4\pi C\sum\left(2\frac{d\varphi_1}{dx}m_2 dV\right)\delta x+\sum\left(\frac{d\varphi_1}{dx}m_2 dV\right)\delta x=0$$

由此得

$$C=\frac{1}{8\pi} \qquad (8.2.13)$$

所以单位体积中涡旋的能量为

$$\frac{1}{8\pi}(\alpha^2+\beta^2+\gamma^2) \qquad (8.2.14)$$

这就是磁场能量密度的公式$\frac{\mu H^2}{8\pi}$。涡旋能量的变化或者来自粒子层切向力所作的功，或者来自涡旋形状的改变。

接着，麦克斯韦根据粒子层的切向力对涡旋作的功应等于涡旋能量的改变，推导出了重要的磁场状态的改变和感生电场的关系式。

令 P、Q、R 是在坐标轴 X、Y、Z 方向上作用在单个粒子上的力。因为每个粒子的端部接触了两个涡旋，所以单个粒子对涡旋的反作用力分别为$-\frac{1}{2}P$，$-\frac{1}{2}Q$，$-\frac{1}{2}R$。又因为在确定电流与粒子运动之间的夫系时，他得到单位涡旋面上接触的粒子数为$\frac{1}{2\pi}$，所以作用在涡旋单位面积上力为

$$-\frac{1}{4\pi}P \qquad -\frac{1}{4\pi}Q \qquad -\frac{1}{4\pi}R \qquad (8.2.15)$$

现在令 dS 为涡旋的面积元，它的法线的方向余弦为 l、m、n，它的坐标是 x、y，z。它在坐标轴 X、Y、Z 方向的分速度为 u、v、w。于是粒子层作用在 dS 面上的力所作的功为

$$\frac{dE}{dt}=-\frac{1}{4\pi}(Pu+Qv+Rw)dS \qquad (8.2.16)$$

让我们从第一项开始，P 可以写为

$$P_0+\frac{dP}{dx}x+\frac{dP}{dy}y+\frac{dP}{dz}z$$

$$u=n\beta-m\gamma$$

因为涡旋表面是一个闭合面，所以

$$\Sigma\, l dS=\Sigma\, m dS=\Sigma\, n dS=0$$

$$\Sigma\, nx dS=\Sigma\, mx dS=\Sigma\, ny dS=\Sigma\, mz dS=0$$

且 $$\Sigma\, my dS=\Sigma\, nz dS=V$$

式中 V 为该闭合面所包围的体积。由此得到

$$\sum pudS=\left(\frac{dP}{dz}\beta-\frac{dP}{dy}\gamma\right)V$$

在单位时间内对涡旋作的总功为

$$\frac{dE}{dt}=-\frac{1}{4\pi}\sum(Pu+Qv+Rw)dS$$

$$=\frac{1}{4\pi}\left\{\alpha\left(\frac{dQ}{dz}-\frac{dR}{dy}\right)+\beta\left(\frac{dR}{dx}-\frac{dP}{dz}\right)+\gamma\left(\frac{dP}{dy}-\frac{dQ}{dx}\right)\right\}V$$

据式(14)，对体积为 V 的涡旋其能量为

$$E=\frac{1}{8\pi}\mu(\alpha^2+\beta^2+\gamma^2)V$$

而 $$\frac{dE}{dt}=\frac{1}{4\pi}\mu V\left(\alpha\frac{d\alpha}{dt}+\beta\frac{d\beta}{dt}+\gamma\frac{d\gamma}{dt}\right)$$

因为粒子对涡旋作的功应等于涡旋能量的改变，所以由此可得

$$\alpha\left(\frac{dQ}{dz}-\frac{dR}{dy}-\mu\frac{d\alpha}{dt}\right)+\beta\left(\frac{dR}{dx}-\frac{dP}{dz}-\mu\frac{d\beta}{dt}\right)+\gamma\left(\frac{dP}{dy}-\frac{dQ}{dx}-\mu\frac{d\gamma}{dt}\right)=0$$

这个方程对 α、β 和 γ 的任何值都是成立的。

首先令 β=γ=0，并用 α 除等式两端，然后分别令 α=γ=0、α=β=0，并相应除以 β、γ，就得到一组公式

$$\left.\begin{aligned}\frac{dQ}{dz}-\frac{dR}{dy}&=\mu\frac{d\alpha}{dt}\\ \frac{dR}{dx}-\frac{dP}{dz}&=\mu\frac{d\beta}{dt}\\ \frac{dP}{dy}-\frac{dQ}{dx}&=\mu\frac{d\gamma}{dt}\end{aligned}\right\}(8.2.17)$$

从这些方程可以确定涡旋速度的改变 $\frac{d\alpha}{dt}$ 等等和施加在粒子层上的力之间的关系。按照前面的假设，这就是磁场强度 H 的改变和感生电场 E 之间的关系。用现代的形式表示，即为

$$\mathrm{Curl}E=-\mu\frac{\partial H}{\partial t}\qquad(8.2.18)$$

麦克斯韦将这一公式与他在第一篇论文中得到的磁场强度和电紧张态之间的关系式 $\mathrm{Curl}a=\mu H$ 相比较,再一次得到 $E=-\frac{\partial\alpha}{\partial t}$ 这表明作用在粒子层上的力就是感应电动力，也就是描述“电紧张状态”变化率的物理量。

图 8—6

接着麦克斯韦按照他的分子涡旋假设对上述关系式进行力学解释。如

图(8－6)所示，六角形空白代表涡旋，分开涡旋的小圆圈代表带电粒子。当电流从左到右通过 AB 时，AB 上面的一排涡旋 gh 按逆时针方向旋转，如果 KL 排涡旋仍然处于静止状态,则这两排涡旋之间的粒子将在顺时针方向上转动，粒子层 pq 由右到左运动，即产生与 AB 电流方向相反的感应电流。当电流稳定时，空间涡旋运动都已达到平衡，它们之间相互没有影响。如果 AB 中电流突然停止，则 gh 排涡旋将受到阻碍，而 KL 排涡旋仍然在反时针方向上快速转动，这就使 pq 层中的粒子从左向右运动，即产生与 AB 电流方向相同的感应电流。因此感应电流的现象是涡旋的转速改变时发生的作用力的一种效应。

麦克斯韦在论文的这部分还讨论了电流或磁铁运动时产生的感应现象，以及导体运动时产生的感应现象，并用分子涡旋假设对动生电动势的产生作了力学解释。

在论文的第三部分，应用于静电的分子涡旋理论中，麦克斯韦引进了“位移电流”的概念。为了计算涡旋运动从场中的一部分传到另一部分，他假设涡旋物质具有弹性。当分子涡旋之间的粒子层受到电力而位移时，运动粒子给涡旋切向力使之变形，而形变涡旋给该粒子以大小相等、方向相反的弹性力，当激发粒子的电力撤销时，涡旋恢复原来的形状，粒子也回复原来的位置。

麦克斯韦根据物理图象的分析，对比了在电场作用下导体内和电解质内所发生的过程。他把导体比作多孔的膜，允许电流通过。把介质比作一个流体通不过的弹性膜，能把流体的压力从一边传到另一边。他指出：“虽然电流通不过介质，但是电的效应可以通过介质传播。”他肯定在变化电场的作用下，电解质中也会有一种能够产生磁效应的特殊的“电流”。他写道：

“只要有电动力作用在导体上，它就产生一个电流……。作用在电介质上的电动力使它的组成部分产生一种极化状态，犹如铁的颗粒在磁体的影响下的极性分布一样……在受到感应的电介质中，可以想象每个分子中的电是这样移动的，使得一端为正，另一端为负，但是这些电仍然完全同分子联系在一起，而不从一个分子跑到另一个分子上去。这种作用对于整个电介质的影响是引起电在一定方向上的一个总位移。这一位移并不构成电流，因为当它达到某一定值时就保持不变了。不过当电位移不断变化时，随着电位移的增大或减少，就会形成一种沿着正方向或负方向的电流。”[4]

麦克斯韦就是这样表述了位移电流的概念。所以位移电流出现在任何电场强度有变化的电介质中，传导电流与位移电流一起，保证了电流的连续性；位移电流和传导电流一样地在它的周围产生磁场。麦克斯韦按照他的分子涡旋模型，根据力的平衡条件，得出了电动力 R 与电位移 h(指的是穿过垂直于运动方向上的单位面积的电量)成正比的结论，即 Rαh。

微分得 $\frac{dh}{dt}\alpha\frac{dR}{dt}$，即位移电流与电动力随时间的变化率成正比。用现代的形式表示 $i_{位移}\alpha\frac{\partial}{\partial t}$。

最后，麦克斯韦讨论了在涡旋物质中波的传播速度。对于密度为 ρ 和

切变模量为m的弹性媒质，横波的传播速度$v=\sqrt{\frac{m}{\rho}}$。他认为这一公式也适用于具有弹性结构的涡旋物质。涡旋介质的密度ρ与介质的磁导率μ有关，$\mu=\pi\rho.m$与涡旋介质的电学量有关.经他推导得$\sqrt{\pi m}$是电量的电磁单位和静电单位的比值。1856年韦伯和柯尔劳斯(R.Kohlrausch,1809—1858)得出这一值为 3.1074×10^{3}m/s，这正是μ=1 的涡旋介质中波的传播速度。这个值与斐索(Fizeau)在 1849 年测定的光在空气中的速度 3.14858×10^{3}m/s 符合得很好。这一惊人的结果进一步揭示了电磁现象和光现象之间的联系。麦克斯韦在论文中着重指出：“我们不可避免地推论，光是介质中起源于电磁现象的横波。”[4]

（三）《电磁场的动力学理论》——场的概念的确立位移电流的定量表示 电磁场方程组的建立

1865 年，麦克斯韦发表了电磁场理论的第三篇重要论文《电磁场的动力学理论》(ADynamicalTheoryoftheElectromag-netic Field)。在这篇论文里，他不再采用以前论文中的力学模型，因为他认识到那里的分子涡旋运动只是一种力学运动，用它去解释复杂的电磁现象是不够的，从而放弃了这类假设。他写道：“在前一工作(指他的‘论物理力线’)中，我曾试图描述一种特殊的运动和一种特殊的应力，用以说明电磁现象。在本文中，我避免任何这类假设，而使用与电流感应和介质极化这些熟知现象有关的电动量和电弹性这样一些词汇，我仅希望指点读者想到一些力学现象，它们将帮助读者理解电现象。本文所有这些用语都应看作是说明性的，而不是解释性的。”[5]

在这篇论文的第一部分引言中，麦克斯韦最早明确地提出了电磁场的概念。他在评价韦伯和纽曼的超距作用的电磁理论时写道：“然而，在依赖于粒子速度的力超距作用在另一些粒子上的假设中包含着力学上的困难，阻止我认为这一理论是最终的理论。”“所以，我宁愿从另一方面去寻找对这一事实的解释，假设它们是被周围介质以及在激发物体中所发生的作用而产生，而不需要假定在相当距离上直接作用的力存在就可以解释远距离物体之间的作用。”“所以，我提出的理论可以称为电磁场理论，因为它与带电体或磁体附近的空间有关；它也可以称为动力学理论，因为它假定在该空间中有运动的物质，从而产生了我们所观察到的电磁现象。”[5]

麦克斯韦认为电磁场既可在物体内存在，也可在真空中存在。他写道：“电磁场是包括和环绕那些处于电或磁状态的物体的那一部分空间，它可以被任何种物质所充满，也可以抽成没有任何宏观物质的空间，就象在盖斯勒管或其它称为真空的情形下一样。”他假定以太物质是电磁场的物质承担者。他指出：“从光和热的现象看来，我们有理由相信，有一种以太介质充满空间和渗入物体；它能运动并将该运动从一部分传到另一部分；它能将该运动传到宏观物体，使其加热，并以各种方式影响它。”[5]

在物理学发展史上，麦克斯韦第一个表述了能量局域性的概念。这就是说对于带电物体或磁性物体，能量并非只存在于这些物体内部，而且也存留在该物体周围的空间中。他写道：“一切能量，不论它以运动的形式

存在，或以弹性的形式存在，还是以任何其它形式存在，都是和机械能一样的。……剩下的问题仅仅是：它存留在什么地方？……按照旧的理论，它存在于带电物体、导体环流以及磁铁的内部……按照我们的理论，它存在于电磁场中，存在于带电体的周围空间和这些物体内部。它以两种不同的形式表现出来，这两种形式就是电极化和磁极化；或者按照一种可能的假设，把它们看成是同一介质的运动和应力。”[5]麦克斯韦对电磁场的能量还作了定量计算，推导出了电磁场能量密度公式和总能量方程。

在论文的第三部分麦克斯韦建立了电磁场的普遍方程，它与我们今天所熟悉的麦克斯韦方程组已经非常接近，一共有八组方程，他把前六组矢量方程写成直角坐标分量式，所以这是一组包括 20 个变量的由 20 个方程构成的方程组。在这篇文章中，麦克斯韦直接根据电磁学的实验事实和普遍原理给出这些方程。下面我们用今天使用的术语和符号，把这些方程表示如下。

A.全电流方程

$$j' = j + \frac{\partial D}{\partial t}$$

式中j为真实的传导电流密度，j'为总电流密度，$\frac{\partial D}{\partial t}$为电位移对时间的变化率，即位移电流密度。这个方程是麦克斯韦电磁理论的核心，它把麦克斯韦关于位移电流的思想定量化了。

B.磁力方程

$$\mu H = \mathrm{Curl}\,\alpha$$

式中 H 为磁场强度，μ为磁导率，μH 即磁感应强度 B。α为电磁动量，麦克斯韦把它定义为：$\alpha = \int E dt$。他指出“α是与法拉第称为的电紧张态相同的量”。上式表明了在磁场中沿任意闭合回路的电磁动量的总和等于穿过回路的磁力线数。对此式两边取散度，即得现代形式的方程 divB=0。

C.电流方程

$$\mathrm{Curl}H = 4\pi j'$$

根据实验可知，当磁极在磁场中移动的闭合回路未绕过电流时不产生功，而磁极沿绕过电流的回路移动时所产生的功与绕过的电流有关，由此得出电流方程。将它与A式结合起来，即为$\mathrm{Curl}H = 4\pi(j + \frac{\partial D}{\partial t})$，写为现代形式（高斯制，以下同此）的方程为$\mathrm{Curl}H = \frac{4\pi}{c} j + \frac{\partial D}{\partial t}$。

1873 年，麦克斯韦在他的《电学和磁学通论》(TreatiseonElectricity and Magnetism)这部经典著作中，叙述了引入位移电流概念这一思想过程。他在该书的第 607 条中作出这一评述：“只有很少的实验证明介质中位移电流的改变与电流的电磁作用相联系。但是协调电磁定律与不闭合电流存在的极大困难使我们必须接受瞬变电流的存在是由于位移变化产生的。这是许多理由中的一个理由。”[6]

对安培环路定律的微分形式有

$$\mathrm{Curl}H=\frac{4\pi}{c}j$$

这就要求 divj=0

也就是说该定律只能在闭合回路中成立。在闭合环路中的任何点上没有电荷的累积。在开路的情形下，例如在电容器充电时，我们有

$$\mathrm{div}j=-\frac{d\rho}{dt}$$

ρ是单位体积中所包含的电荷量。这时在电容器的极板上有电荷累积，j的散度不为零，这就表明在不稳定电流的情形下，安培环路定律不成立。为了克服这一困难，麦克斯韦引进了全电流思想。

因为 $$4\pi\frac{d\rho}{dt}=\mathrm{div}\frac{dD}{dt}$$

所以 $$\mathrm{div}\left(j+\frac{1}{4\pi}\frac{dD}{dt}\right)=0$$

如果假设$\frac{1}{4\pi}\frac{dD}{dt}$这项象j一样，具有电流的性质，于是有

$$\mathrm{Curl}H=\frac{4\pi}{c}j+\frac{1}{c}\frac{\partial D}{\partial t}$$

D.电动力方程

$$E=\mu v\times H-\frac{\partial\alpha}{\partial t}-\mathrm{grad}\varphi$$

式中 E 为电动力，即电场强度。第一项表示导体本身运动产生的电动力。麦克斯韦指出这个电动力与运动方向和力线垂直。如果以速度 v 和磁感应强度 μH 为平行四边形的两个邻边，则这个力的大小等于平行四边形的面积，力的方向垂直平行四边形的平面。第二项表示在场中由于磁体或电流强度的改变，或位置的变化而引起的电动力。第三项表示电势φ引起的电动力。将此式两边取旋度并与B式结合起来，即得$\mathrm{Curl}E=-\frac{\partial B}{\partial t}$，写成现代形式的方程为$\mathrm{Curl}E=-\frac{1}{c}\frac{\partial B}{\partial t}$。

E.电弹性方程

E=KD

电动力作用于电介质，使它的每一部分极化，它的相对的面上出现相反的电荷，这个电量取决于电动力和电介质的性质。对各向同性的电介质电动力 E 与电位移 D 方向相同，其大小成正比，比例系数为 K。写成现代形式的方程为 D=εE。

F.电阻方程

E=-ρj

电动力作用在导体上产生通过导体的电流。式中ρ为单位体积导体中的电阻，写成现代形式为 j=σE，σ为电导率。

G.自由电荷方程

e+divD=0

式中 e 是单位体积内的正的自由电荷量。写成现代形式的方程 divD=4

$\pi\rho$。

麦克斯韦在《电学和磁学通论》一书的第 60 条中对上式说明如下："如果电荷 e 均匀地分布在一个球面上，在介质中与球心距离 r 处的任何点的电动力强度正比于 $\frac{e}{r^2}$。因此介质中的电位移将正比于 $\frac{e}{r^2}$，如果我们画出半径为 r 的同心球面，通过这个面的总电位移将正比于 e，而与 r 无关。如果 U_1 和 U_2 分别是内球面和外球面的电势，则增加电位移 dE 所作的功将为 $(U_1-U_2)dE$。如果取外球面在无限远，则 U_1 成为带电球的势，而 U_2 就变为零，于是这个功就为 UdE。但是这个功也是 Ude，在这里 de 是球的电荷的增加。如果我们承认电能存在于介质中，则 dE=de，即穿过任何同心球面的电位移等于球上的电荷。由此得出结论：位移电流给任何其它有限长的电流提供了一个连续的，等同于闭合环路中的电流。"[6]

H.连续性方程

$$\frac{\partial e}{\partial t} + \mathrm{div} j = 0$$

这就是电荷守恒定律。

以上八个方程就是麦克斯韦最早提出的电磁场方程组。这是一套完备的电磁场方程组，只要知道问题的条件，方程中的变量是完全可以确定的。这个方程组概括了各个电磁学的实验规律，是能够完整和充分地反映电磁场客观运动规律的理论。

方程组最简单的完美对称形式是 1890 年由赫兹写出的，它包括四个方程，其现代形式为

$$\mathrm{div}D=4\pi\rho_0$$

$$\mathrm{div}B=0$$

$$\mathrm{CirlH} = \frac{1}{c}\frac{\partial D}{\partial t} + \frac{4\pi}{c} j_0$$

$$\mathrm{curlE} = -\frac{1}{c}\frac{\partial B}{\partial t}$$

其中第一个方程即电学的高斯定理，D 为电位移矢量，ρ_0 是自由电荷体密度。第二个方程即磁学的高斯定理，说明磁场是涡旋场。第三个方程即麦克斯韦推广了的安培环路定理。H 为磁场强度，j_0 是传导电流密度，$\frac{\partial D}{\partial t}$ 是位移电流密度。说明传导电流和变化电场在其周围产生磁场.第四个方程即电场的环路定理，说明变化的磁场产生感应电场。

麦克斯韦的电磁场理论从超距作用过渡到以场作为基本变量，实现了科学认识的一个革命性变革。在《物理学的进化》一书中，爱因斯坦和英费尔德评论说："这些方程的提出是牛顿时代以来物理学上一个最重要的事件，这不仅是因为它的内容丰富，并且还因为它构成了一种新型定律的典范。"

参考文献

[1]沈葹，《从法拉第到麦克斯韦》，《物理教学》，1980 年 3 月

[2]Maxwell,onFarday'sLinesofForce,ed.W.D.Niven,TheScientific papersofJamesclerkMaxwell,CambridgeUniver-
sityPress,Vol.1,(1890)pp155—209
[3]E.T.Whittaker,AHistoryoftheTheoriesofAetherandElectricity, London,Longmans,Vol.1,1951 年,pp242—247
[4]Maxwell,ibid,451—514
[5]Maxwell,ibid,526—536,554—564
[6]R.A.R.Tricker,TheContributionsofFaradayandMax-
welltoElectricalScience,Oxford,PergamanPress,1960 年,126—129
[7]爱因斯坦、英费尔德著，周肇威译，《物理学的进化》，上海科技出版社，1962，91
[8]陈熙谋等，“Maxwell 电磁场理论的建立和它的启迪”，《大学物理》，1986 年 10 月，33—36，1986 年 11 月，27—33

九、光的本性认识的发展

光的本性问题是贯穿在光学发展中的一个根本问题。正是这种对光的本性的探讨有力地推动了光学以及整个物理学的发展。人们对光的本性的认识，从光是“物质的微粒流”，经历了光是“以太的振动”，光是电磁波到光是波粒二象性的统一等各个认识阶段。这一认识历程从牛顿和惠更斯之争算起到现在其间经历了三百多年。人们遵循实验——假设——理论——实验这条途径，逐步达到了对光的本性的认识，这一认识揭示了物质世界光和电磁的统一，光的波动性和微粒性的统一。德国物理学家劳厄在谈到这一认识的重大意义时指出：“在这以前还是完全互不相依的光的理论和电动力学理论的这种自然的结合发展是作为物理知识的真理性证明的一个最伟大的事件”。他在《物理学史》的导言中着重指出了研究两类不同的物理思想“它们不期而遇并且自然地相结合”的意义。他说：“凡是经历了这种令人极为惊奇的事件的人，即使是在很远的距离经历的，或者至少能在事后加以回顾的，都不会怀疑：这些相互结合的理论，即使不包含完全的真理，终究也包含了与人类的附加因素无关的客观真理的一种重要的内核。否则，它们的结合只能理解为奇迹。物理学史的理想必须是把这样的事件尽可能明晰地刻画出来”。[1]下面我们就来叙述人类对光的本性认识的发展过程。

(一)微粒说与波动说的思想渊源

关于对光的本性这一古老之谜的认识要追索到古希腊时代。古希腊杰出的原子论者德漠克利特（Democritus，公元前460—前370)最早提出光是物质微粒的观点。他认为视觉是由物体射出的微粒进入眼睛而引起的。古希腊的另一个原子论者伊壁鸠鲁(Epicurus,公元前341—前270)和古罗马的原子论者卢克来修(Lucretius，公元前99—前55)坚持这一学说。卢克来修说：“从任何我们看见的东西，必定永远有许多原初物体流出来，被发放出来，被散布到四周各处，这些物体撞击眼睛，引起了视觉。”[2]原子论者的这一观点是后来把光看作某种物质实体的粒子说的萌芽。古希腊杰出的思想家亚里士多德(Aristotle，公元前384—前322)认为，视觉是在眼睛和可见物体之间的中间介质运动的结果。他认为这种中间介质有让光通过的可能性(潜在能力)，即是透明的，光则把这种可能性变为现实。所以，没有中间介质就没有视觉。在这个理论中包含着后来的光的波动说的思想。[2]

科学发展到了 17 世纪，法国哲学家、物理学家、数学家笛卡儿(ReneDescartes，1596—1650)提出了对光的本性的看法。英籍德国物理学家玻恩(MaxBorn，1882—1970)和美国物理学家沃耳夫(Emil.Wolf，1922—)在他们的《光学原理》的历史引言中说：“在新哲学的创立者当中，笛卡儿可以提出来说一说，因为他根据他的形而上学观念系统地陈述了关于光的本性的见解。笛卡儿认为，光在本质上是一种压力，在一种完全弹性的，充满一切空间的介质(以太)之中传递，他并且把颜色的差异归因于这个介质中粒子的不同速度的旋转运动。”[3]笛卡儿对光的本性没有明确而统一的观点。他在他的著作《光的折射》中提出了一个比喻：光通过介质传入人眼，就象机械脉冲沿着手杖传入盲人的手和脑中一样，并没有某种

物质性的东西传入眼睛使我们看到光和色。笛卡尔在这里强调了介质的影响和接触作用，认为光是以太介质中某种压力的传播过程，所以可以把他算作波动论者。[2]

另一方面，笛卡儿又从光的微粒观念中推导出反射定律与折射定律。笛卡儿在《光的折射》中写道：“假设我们将球从 A 点击到 B 点，碰到地面 CBE 时,球因受阻而偏离原来的运动方向(图 9—1)。它将往哪个方向偏？为了简化所研究的问题，我们假定：地面是平滑坚硬的，并且球在下落和弹跳时速度保持不变。我们不考虑球在离开球拍以后能够继续运动的原因，也不考虑球的重量、大小以及形状对运动方向会有什么影响。

因为我们的目的并不是研究这些问题，而是研究光；并且上述因素对光都没有影响。”[4]

图 9—1

接着笛卡儿开始了对平滑表面上光的反射的研究。他把球的速度分解为垂直分量及水平分量。当球碰到地面时，只是球速的垂直分量方向相反，大小不变。水平分量是不变的，由此很容易证明光的入射角等于它的反射角。

笛卡儿继续写道：“现在让我们来观察折射现象。首先假定球从 A 点被抛至 B 点，在 B 点碰到的不是地面，而是一块布 CBE，它非常稀疏和不结实，将减慢球的速度。”[4]

他仍然把速度分成垂直分量及水平分量，垂直方向速度减小而水平分量不变图 9—2。由此得

$$v_1 \sin i_1 = \mu_2 \sin i_2$$

即

$$\frac{\sin i_1}{\sin i_2} = \frac{v_2}{v_1}$$

图 9—2

对各向同性介质，v_1、v_2 与光的传播方向无关。上式右端是一个与入射角无关的常数，它表明入射角的正弦与折射角的正弦之比是一个常数，这样便解释了折射定律。

但是实验表明当光从光疏介质(例如空气)进入光密介质(例如水)时，光折向界面法线即 $i_2 < i_1$，按照上述公式则 $v_2 > v_1$，从微粒论的观点看来，光在光密介质中的光速应当大于光在光疏介质中的光速。后来，实验证明这一结论是错误的，但这是在笛卡儿以后 200 年的事了。当时他为自己的观点解释说：“……你也许会感到惊奇……但如果你还记得我所描述过的光的特性，你就不会为此而感到惊奇了。我曾说过，光只能设想为一种能填塞所有其它微孔的很稀薄的物质的某些运动或作用。……所以，事情是这样的，透明体的细小部分越是坚硬，光的通过越是容易，因为光并不要使透明体的任何部分离开它们的位置，它不象球那样必须排开部分水才能从中通过。”[5]

（二）波动说与微粒说之争

英国物理学家、天文学家胡克在他 1665 年发表的《放大镜下微小物体

的显微术或某些生理学的描述》(以下简称《显微术》一书中，明确主张光是一种振动。同时他还以金刚石的坚硬特性，提出这种振动必定是短促的。在讨论了光的直线传播和光速是有限的之后，他写道：“在一种均匀介质中这一运动在各个方向都以相等的速度传播。所以发光体的每一个脉动或振动都必将形成一个球面。这个球面将不断地增大，就如同把一石块投入水中后在水面一点周围的环状波膨胀为越来越大的圆圈一样(尽管肯定要快得多)。由此可知，在均匀介质中扰动起来的这些球面的一切部分与射线交成直角。”[2]可见，胡克实际上已接触到波前或波面的概念。

胡克还研究了薄膜的彩色。他在《显微术》中记载了这样的现象：厚云母片是无色的，但将它揭成薄片时却呈现出了象虹一样美丽的颜色；他还发现，在反射面上盖上一层具有不同折射作用的透明薄膜时，也会产生彩色现象。他对薄膜的这一光学现象提出了解释，他认为这是由于直接从前表面反射的光和经过折射从后表面反射的光相互作用而形成的。胡克提出：“一束最弱的成分领先而最强的成分随后的光脉冲的混合，在网膜上引起蓝色的印象，一束最强的成分领先而最弱的成分随后的光脉冲的混合，则引起红色的印象。其他的颜色印象都可以由两种成分的先后排列情况作出解释。”[2]胡克的这种解释是不正确的。他认识到了这一现象与两薄膜的厚度有关，但是他还没有建立相位差的概念，对干涉现象不能作出正确的解释。

经典物理学的奠基者牛顿基本上是主张微粒说。他根据光的直线传播性质，提出光是微粒流的理论。他认为这些微粒从光源飞出来，在真空或均匀物质内由于惯性而作匀速直线运动。早在 1672 年 2 月 6 日牛顿送交英国皇家学会的一封信“关于光和色的新理论”中，就提出了光的微粒说，认为光线可能是由球形的物体所组成，并用这种观念解释了光的直线传播和光的反射折射定律。[6]

胡克和牛顿争论时，提出不少问题，特别是微粒说所不能解释的一些事情。为了回答胡克提出的问题，牛顿又进一步研究，想办法如何来完善自己的假说和理论。他认为他的光的粒子结构理论是正确的，但他也没有绝对肯定这个结论，而只能用“可能”两个字来表示。1675 年 12 月 9 日，牛顿在送交皇家学会的一篇论文《涉及光和色的理论的假说》中，提出了一个把光的微粒和以太的振动相结合的新假说。他写道：“以太的振动在这一假说和那一假说中都是一样有用的和不可缺的。因为假定光线是从发光物质向各方面发射出去的小的微粒的话，那末当它们碰到任何一种折射或反射表面时，就必然要在以太中引起振动，正象石块被投到水中要引起振动一样。我还假定，这些振动将按照激发它们的上述颗粒性光线的大小和速度不同而有不同的深度和厚度。”[6]

除此之外，牛顿在 1675 年 12 月 21 日写给奥尔登堡(HenryOldenburg，当时皇家学会的秘书)的信中谈到他和胡克看法不同之处。牛顿认为，“除了假定以太是一种能振动的介质以外，我和他没有什么共同之点。然而我对这个假定有和他很不相同的看法：他认为能振动的以太就是光本身，而我则认为它不是，这是一个很大的差别，正如他和笛卡儿的差别很大一样。”[6]

当牛顿在皇家学会宣读新的论文、阐述新的假说时，胡克却提出了关于优先权的要求。于是牛顿在愤慨之下，决定不发表光学著作。而牛顿多

年来的光学研究成果，在 1704 年(胡克死后的一年)才发表在他的《光学》著作中。这一事情看来是影响了光学的发展。[6]

牛顿在《光学》中指出光的波动性不能很好地说明光的直线传播这个最基本的事实。他在《光学》的问题二十八中写道：

“在有些假设中，把光描述成在某种液体介质中传播的一种压力或运动，这些假设难道会是正确的？……如果光是瞬时传播或在时间中传播的压力或运动，它就应当朝影子内部弯曲。因为，在障碍物附近(它挡住了一部分运动)，压力或运动不可能在液体中沿直线传播——它们将发生弯曲，并在位于障碍物后面的静态介质中到处传播。”[4]

牛顿接着又写道：

“静水水面上的波沿较大的障碍物(它挡住了一部分波)边缘传播时，它会发生弯曲，并不断地向障碍物后面的静水水域扩展。空气波，空气的脉动或振动(它们构成声音)显然也会发生弯曲，但不会象水波那样强烈。小山虽然可以挡住我们的视线，使我们看不到钟或大炮，但在山后仍然可以听到它们的声音；声音很容易沿着弯弯曲曲的管道传播，如同在笔直的管子中传播一样。至于光，从来还没有听说过它可以沿着蜿蜒曲折的通道传播，或者朝阴影内弯曲，因为当一颗行星运行到地球与另一颗不动的恒星之间时，这颗恒星就看不见了。”[4]

在《光学》的问题二十九中牛顿明确地表述了光是微粒的观点。他指出：“光线是不是发光物质发射出来的很小的物体，因为这样一些物体能直线穿过均匀介质而不会弯到影子区城里去，这正是光线的本质。”[4]

“牛顿环”现象是牛顿的一项重要发现，从这一发现中他提出并确立了光的周期性。他在《光学》中详细地描述了这一实验现象。当他把一个平凸透镜放在一个双凸透镜上时，他观察到一系列明暗相间的同心圆环。压紧玻璃体，改变其间空气膜的厚度，又发现了条纹的移动。他还精确地测量了环的半径，发现环的半径的平方构成一个算术级数。这里最重要的是对光的周期性的发现。牛顿这样写着：“有时我一连数了三十多次周期性变化的序列，在每一个序列中都包括一明一暗的环，但是由于它们太窄无法数清楚。”牛顿的精确观察，本来是光的波动性的证明；他的测量也为确定各种色光的波长提供了根据，但是他并未由此走向波动说。

牛顿用他的光的微粒和以太振动相结合的新观点，解释了他发现的牛顿环现象。他设想光微粒在介质的界面上可以引起以太的各种大小的振动，即以太的压缩和扩散，并且按照其大小而激发起不同颜色的感觉，正象空气的振动按其大小而激起不同的声音感觉一样。他设想光微粒在介质界面处所激起的以太振动会在介质中传播开，而且是快于光速的，因而可以追上光线。由于这种追得上光线的以太振动的作用，使光微粒时而被加速时而被减速，从而使它一阵容易透射，一阵容易被反射。他在《光学》第二编第三部分中写道：

“每一条光线在通过任何折射面时，便处于某种为时短暂的过渡性结构和状态之中。在光线的前进过程中，这种状态每隔相等的间隔(等时或等距)内复发一次，并使光线在它每一次复发时容易透过下一个折射面，而在它相继两次复发之间容易被这个面所反射。”[2]这里所谓“等间隔”的复发，其实就是指光线的某种周期性。他还定义说：

“我将把任何一条光线返回到偏向于反射的状态，称为它‘容易反射

的猝发’，而把它返回到倾向于透射的状态称为它‘容易透射的猝发’，并且把每一次返回和下一次返回之间称为它猝发的间隔。”[2]

他还把这种“猝发的间隔”同光的颜色联系起来，认为红色光的间隔最大，紫色光的间隔最小。可以看出，牛顿所说的这种‘猝发间隔的距离’，在某些方面和后来波动说中所说的“波长”是很相似的。

坚持光的波动说，并想建立理论来解释它的是荷兰物理学家惠更斯(ChristianHuygens，1629—1695)。惠更斯于1678年向法国科学院提交了《光论》这本著作，该书于1690年出版。惠更斯以下列论据来驳斥牛顿的微粒说，他写道：

“假如注意到光线向各个方向以极高的速度传播，以及光线从不同的地点甚至是完全相反的地方发出时，光射线在传播中，一条光线穿过另一条光线而相互毫不影响，就能完全明白这一点：当我们看到发光的物体时，决不可能是由于从它所发出的物质，象穿过空气的子弹和箭一样，通过物质迁移引起的。”[2]

惠更斯在批评微粒说时，提出了他的波动说。他认为光是由发光体的微小粒子的振动在充满于宇宙空间的介质“以太”中的一种传播过程，光的传播方式象声音的传播方式一样。他写道：

“我们知道，声音是借助看不见摸不着的空气向声源周围的整个空间传播的，这是一个空气粒子向下一个空气粒子逐步推进的一种运动。而因为这一运动的传播在各个方向是以相同速度进行的，所以必定形成球面波，它们向外越传越远，最后到达我们的耳朵。现在，光无疑也是从发光体通过某种传给媒介物的运动而到达我们的，因为我们已经看到从发光体到达我们的光不可能是靠物体来传递的。正如我们即将研究的，如果光在其路径上传播需要时间，那么传给物质的这种运动就一定是逐渐的，象声音一样，它也一定是以球面或波的形式传播的；我们把它们称为波，是因为它们类似于我们把石头扔入水中时所看到的水波，我们能看到水波好象在一圈圈逐渐向外传播出去，虽然水波的形成是由于其它原因，并且只在平面上形成。”[8]

惠更斯认为光波是一纵波，这一波以非常大但又是有限的速度在以太中传播。这种以太由不均匀的、微小的、弹性的、压缩得非常紧密的颗粒组成。因此光不是一种实际上的物质的迁移，而更类似于是一种“运动的趋势”，类似于通过一列球的碰撞而产生的系统的位移。因为以太颗粒不是成排而是无规则地分布的，一个被撞击的颗粒将把这种“运动的趋势”传递给它运动方向上所有它所触及的颗粒。惠更斯由此断言，新的波前在被光所触及的每一个颗粒周围产生，并以半球形式散布开来。他认为产生于单一的点的单一波前是无限微弱的，不产生光，但无限多的这种波前重叠的地方就产生了光。这就是惠更斯原理。[6]

惠更斯原理的最有意义之处是用来确定波的传播方向。传播中波前上的每一点，都可以看成是一个新的波或子波的波源，新的波前位置就是这些小子波的包络线，这些子波是从原先波前上所有的点发出的。惠更斯应用这条原理成功地解释了反射和折射现象，并得出稠密介质中光的速度小于稀疏介质中光的速度的正确结论。

惠更斯在《光论》中强调他“解释了冰洲石奇异折射的原因”。1669年丹麦哲学家巴托林纳斯发现有一种称为冰洲石的透明矿物晶体具有一种

特殊的性质。当光线沿一定方向穿过它时，它能把光线分裂为独立的两束光。如果绕着入射光的方向转动这块晶体，它发出的两束光，有一种称为正规的折射将保持不变，另一种称为不正规的折射，要随着晶体的转动而转动。他还发现当通过冰洲石看物时所有物象都成双地出现。他提出用“半回转椭圆波”来解释非正规折射。[9]当光线垂直入射到冰洲石的晶体表面上时，就形成两列子波，一是球面波，另一列是椭球面波。球面子波构成的是一个沿原来的入射方向继续前进的波前，而椭球面子波产生的波前要不断地偏向一旁，因而形成了非常光。惠更斯所作的这个

图 9—3

解释虽然是完全正确的，但他无法解释为什么光波在晶体中有两种不同的传播方式，正如他自己说的“我还未能找出产生这现象的理由”。[9]这是因为惠更斯以为光波中振动也象声音的情况那样是沿着其传播方向发生的(纵向振动)，在这种情况下，如果绕着入射光束的方向转动晶体就根本不应当有什么不同。另一方面，牛顿并不相信惠更斯的波和子波，他企图用如下的假设来解释这个称为双折射的现象：构成正规折射和非正规折射光的粒子在垂直于光线的方向上有着不同的指向。在《光学》一书的第二版中，牛顿把这两束光线之间的区别比做两根长棒之间的区别，其中一根棒的截面是圆形的，另一根是矩形的，如果绕其轴线转动第一根棒，就看不出会有什么差别，第二根棒的情况就肯定不一样了。牛顿写道：“所以每条光线都有一对侧面，原来就具有与非常折射有关的性质，而另一对侧面则不具有这种性质。”牛顿显然认识到光线一定具有某些横向的(即垂直于传播方向的)性质，但他未能具体看出这可能是些什么性质。[8]

17 世纪，光的波动理论处于萌芽阶段。笛卡儿认为光是一种压力，胡克却说光是介质中的迅速颤动(脉冲)，这个波动说经过惠更斯的加工，得到了进一步的发挥，但仍然是很粗略的。他没有指出光现象的周期性，没有提到波长的概念，更没有确立相位的概念。他认为光波是纵波，不能解释偏振效应。虽然当时已发现了干涉现象、衍射现象，但是波动论在对这些现象的解释上软弱无力。由于牛顿的权威，使得微粒说在光学中占统治地位达一个世纪之久。

(三)波动说的胜利

19 世纪初叶，一系列决定性的发现导致人们普遍接受波动理论。托马斯·杨和菲涅耳等人的实验和理论工作把光的波动理论大大向前推进，解释了光的干涉、衍射现象，初步测定了光的波长，并根据光的偏振现象确定光是横波。到 19 世纪中叶，光的波动说战胜了微粒说，波动理论在比较坚实的基础上建立起来。

1．托马斯·杨的干涉理论和双缝实验

英国物理学家托马斯·杨(Thomas Young， 1773—1829)迈出了这胜利进军的第一步。1800 年他向皇家学会提出了《在声和光方面的实验和问题》的报告。他认为声和光都是波的传播，光是在充满整个空间的以太流体中传播的弹性振动，由于以太极稀薄，所以光是以纵波形式传播的，光的颜色和不同频率的声音是类似的。在声波叠加的情况下，会产生声音的加强和减弱，复合的声调和拍频。“干涉”这个术语就是他最先提出的。

在 1801 年发表的一篇报告中，托马斯·杨提出了光波的频率和波长的概念，提出干涉原理并解释了“牛顿环”现象。他写道：

“同一束光的两个不同部分，以不同的路径要末完全一样地，要末在方向上十分接近地进入眼睛，在光线路程差是某个长度的整数倍的地方，光就增强，而在干涉区域的中间部分，光将最强。对于不同颜色的光束来说，这个长度是不同的。”[2]

他指出牛顿环的明暗条纹，就是由不同界面反射出的光互相重合而产生“干涉”的结果。相位相反的振动叠加起来就互相抵消，相位相同的则互相加强。他用实验的方法确证了他所提出的这一假设。他用紫外光投射到薄层上，使紫外光从上下两个界面反射产生干涉。由于紫外光是人眼所看不见的，他就让反射光落在涂有氯化银溶液的纸上，结果出现了黑环，从而证明了他的干涉原理。他精确地确定了各种色光的波长值。他指出：“根据各种实验的比较，组成极端红光的波的宽度，在空气中，应假定约为三万六千分之一英寸，极端紫光的约为六万分之一英寸。”[2]

在 1803 年 11 月 24 日写的《关于物理光学的实验和计算》的论文中，托马斯·杨把干涉原理应用于解释衍射现象，并用一个非常简单、非常直观的实验证明这种衍射条纹是直接通过衍射缝的光和边界波的干涉产生的。杨氏在护窗板上钻一个小孔，用一张钻有针孔的厚纸片把它盖住，在通过孔后的阳光途中放一宽约 1/30 英寸的硬纸条,然后观察硬纸条投射到墙上或屏上的影子。他写道：

“除了在影子两侧出现了彩色的带外，我还看到影子本身也被分割成若干个这样的带，带的数目与硬纸到影子的距离有关，而且，影子的中央部分总是白色的。由于光线从硬纸条的两个侧边通过时它会发生曲折或者向影子内部衍射，所以光的这些不同部分互相共同作用的结果就产生了这些彩色带。我只要把一块不大的纸板放在硬纸条前面，使它只挡住硬纸条一侧的光线(纸板阴影的边缘正好落在硬纸条上)，或者将纸板放在硬纸条后面数英寸处，使它只挡住硬纸条投向墙上的阴影的一部分，此时，原先观察到的硬纸投射到墙上的影子中的全部彩带立刻就消失了，虽然这时候从硬纸条的另一侧边衍射出的光并没有被挡住……。”[4]

托马斯·杨在 1807 年出版的《自然哲学讲义》中，进一步阐述了干涉原理，描述了著名的衍射试验。他首先指出干涉现象是波动的普遍特征。他写道：“如果认为任何一定颜色的光都是由一定宽度即一定频率的振动所组成，那么，该波一定是会产生我们在水波和声脉冲中所考察过的那种效应。我们已经指出，由两邻近中心发出的两个相同的波系，可以在某些点上相互抵消其效应，而在其他一些点上倍增其效应。两音叉的拍频也可以从同样的干涉获得解释。对于光的加强与削弱，我们现在应用同一的原理。”[11]

接着，他提出了光的干涉现象产生的条件以及获得相干光的方法。他说：“要使光的两部分效应可以这样地叠加，它们必定要来自同一光源，而且必须从彼此偏离不大的方向经过不同的途径到达同一点。”[11]为了获得两束相干光必须使用同一个光源，然后再用衍射、反射、折射等方法，或者同时使用这几种方法把光束劈开。他指出“最简单的方法是使均匀的光束射到有两个小孔或两个狭缝的屏上，这些小孔或狭缝可以看作是散射中心，光由此向一切方向衍射。”[11]接着杨氏谈到干涉条纹的形状，并

指出产生亮纹、暗纹的条件。

他写道：“在这种情况下，当两个重新形成的光束投射到安置在它们行程上的屏面上时,光束好象被许多黑线条分割成大体上相等的亮带,……在图象中心总是出现亮带，而其他的亮带则分布在图象中央的两侧。亮带之间的间距是：光分别从两个狭缝到达各亮带的光程差正好等于半波长的偶数倍；光分别从两个狭缝到达各条暗带的光程差正好等于半波长的奇数倍。”[4]

图 9—4 杨氏的干涉实验

杨氏完满地解释了光的干涉现象，提出了干涉原理，并且测定了光的波长，对光的波动理论作出了重要的贡献。但是，他的见解大部分是定性的，而且由于他认为光是纵波，这给他的理论带来了很大的弱点，没有立即得到科学界的普遍承认。这个自牛顿以来在物理光学上最重要的成果就这样被埋没了将近 20 年。直到菲涅耳提出他的波动理论后，托马斯·杨才获得了应有的荣誉。

2．菲涅耳的衍射理论和泊松亮斑

最早发现光的衍射现象的人是意大利物理学家格里马弟(FrancescoGrimaldi，1618—1663)。在他逝世后于 1665 年出版的《光、色、虹的物理数学》中，描述了他所作的光的衍射实验。如图 9－5 所示，他在百叶窗上钻一条狭缝 AB，使阳光从晴朗天空通过该缝进入一个完全封闭的黑暗空间，该光的漫射状如锥形 ACDB，在光中放一不透明的物体 EF。他发现 EF 的影子

图 9—5 图 9—6

MN 比按几何光学计算的应有的大小 IL 要宽一些，在受到强烈照射的 CM、ND 部分，靠近影子边缘的区域，还有几层带颜色的带子，越往外越窄。如图 9—6 所示，第一层最宽的带是 NMO，N 是蓝色条纹，O 是红色条纹，M 无色是最宽、最亮的区域。第二层 QPR，第三层 TSV 仍然是靠近阴影的部分带蓝色，远离影子的部分带红色。他设想光是一种能够作波浪式流动的流体，不同的颜色是波动频率不同的结果。他发现光在物体的边缘发生了微小的弯折，他第一个称这种现象为衍射。他指出：“这种光带显然与孔 AB 的大小有关，因为如果孔很大，它们就不见了。”[12]牛顿重复了这些实验，他让光通过两个刀口之间的狭缝，使光线产生了更大的弯曲，牛顿用小孔边缘对微粒的引力来解释这种现象。[4]在菲涅耳发表他的论文以前，衍射效应一直没有得到正确的解释。[3]

1815 年，法国物理学家菲涅耳(A.J.Fresnel，1788—1827)向法国科学院提交了关于光的衍射的第一份研究报告，这时他还不知道托马斯·杨关于干涉和衍射的论文。根据惠更斯的子波假设，菲涅耳以子波相干叠加的思想补充了惠更斯原理，发展成为惠更斯-菲涅耳原理。他认为在各子波的包络面上，由于各子波的互相干涉而使合成波具有显著的强度，这给予惠更斯原理以明确的物理意义。[2]

1818 年，法国科学院提出了征文竞赛题目：一是利用精确的实验确定光线的衍射效应；二是根据实验，用数学归纳法推求出光通过物体附近时的运动情况。在法国物理学家阿拉果与安培的鼓励和支持下，菲涅耳向科

学院提交了应征论文。这篇论文的主体是由惠更斯的包络面作图法同杨氏干涉原理结合而组成，建立了他的作图形式的衍射理论。他用半波带法定量地计算了圆孔、圆板等形状的障碍物产生的衍射花纹。菲涅耳把自己的理论和对于实验的说明提交给评判委员会。参加这个委员会的有：波动理论的热心支持者阿拉果；微粒论的支持者拉普拉斯、泊松和比奥；持中立态度的盖·吕萨克。菲涅耳的波动理论遭到了光的粒子论者的反对。在委员会的会议上泊松指出，根据菲涅耳的理论，应当能看到一种非常奇怪的现象：如果在光束的传播路径上，放置一块不透明的圆板，由于光在圆板边缘的衍射，在离圆板一定距离的地方，圆板阴影的中央应当出现一个亮斑，在当时来说，这简直是不可思议的，所以泊松宣称，他已驳倒了波动理论。菲涅耳和阿拉果接受了这个挑战，立即用实验检验了这个理论预言，非常精彩地证实了这个理论的结论，影子中心的确出现了一个亮斑。这个亮斑后来称之为泊松亮斑。在托马斯·杨的双缝干涉和泊松亮斑的事实的确证下，光的粒子说开始崩溃了。[2]

3．光的横波理论和偏振光的干涉

早在牛顿时代，人们就知道通过冰洲石的光束会分裂成两束折射光的现象。

1808 年底的一个傍晚，在巴黎卢森堡宫殿外，法国工程师和物理学家马吕斯(Etienne Louis Malus， 1775—1812)用冰洲石晶体来看落日在玻璃上的反射现象时，他惊奇地发现只出现一个太阳的像，而不是一般双折射时的两个像。原来人们认为，光被反射或折射时，它的物理性质是不会改变的，这个偶然发现打破了这一见解，马吕斯想到这可能是反射造成的。夜间他在验证自己的发现时，观察了蜡烛在水面上的反射，发现当光束和表面成 36°反射时，在晶体中的一个像就消失了；在其他角度下，两个像的强度一般是不同的。在晶体转动时，较亮的像将会变暗，较暗的像将会变亮。[2]

利用其他反射表面，也会看到类似现象，只是发生一个像消失的角度不同(对玻璃来说，光束与表面成 35°)，1851 年，英国的布儒斯特(David Brewster,1781—1868)发现发生这一现象的入射角的正切等于折射光束所在介质的折射率与入射光束所在介质的折射率之比。

马吕斯对寻常光线和非寻常光线的反射作了进一步的研究后发现，如果一条光线反射了，另一条光线就会进入第二种介质。他由此引入了“光的偏振”这个术语。马吕斯证明了寻常光线和非寻常光线在互相垂直的平面内偏振。[2]

马吕斯进而研究了在简单折射现象中的偏振，他发现光在折射时是部分偏振的，折射光的偏振和反射光的偏振是成相反分布的。他对这一发现非常高兴，认为它击中了(纵波)波动论的要害，而有利于确证把光粒子看作有不同“侧面”的粒子说。并被微粒论的支持者认为是对光的粒子说的“真理性的数学证明”。[2]

1816 年，菲涅耳和阿拉果一起研究了偏振光线的干涉[3]。他们发现来自同一光源的通过冰洲石分裂成的两条折射光互不干涉。其中每一条折射光束可以用某种光学方法进行分光，分光后的光束都可以自己相互干涉。这一事实给光是纵波的观点造成很大的困难。一直在为波动说的困难寻求解决办法的托马斯·杨当知道菲涅耳和阿拉果的工作结果之后，于

1817 年觉察出，如果光的振动不是象声波那样沿运动方向作纵向振动，而是象水波或拉紧的琴弦那样垂直于运动方向作横向振动，问题或许可以得到解决。 1817 年初，他写信给阿拉果说：“……虽然波动说可以解释横向振动也在径向方向并以相等速度传播，但粒子的运动是在相对于径向的某个恒定方向上，而这就是偏振。”[2]阿拉果立即将托马斯·杨的这一新想法告诉了菲涅耳，菲涅耳当时已经独立地领悟到了这个思想。他立即以横向振动的假设解释了偏振光的干涉现象，而且还得出了一系列其它的重要结论。

菲涅耳同阿拉果在研究来自同一光源的两条折射光互不干涉的现象时，阿拉果提出了这样一个设想：光波系统可能是由两个相互垂直的振动所组成。菲涅耳发展了这个思想。他提出，光波可用垂直于其传播方向上的一种位移来描述。所以，由于自身结构上的特点，冰洲石晶体和电气石晶体能够把光束分解为相互垂直的两个光分量，这就出现了偏振现象。因为这两个分量是相互独立的，所以它们不能相互干涉，在水平方向上的位移不可能与在垂直方向上的位移相互抵消。菲涅耳写道：“当我们用这样的观点观察事物时，偏振的作用就不再是激起横向运动，而是把横向运动沿两个相互独立、相互垂直的方向分解，并把这两个分量单独分离出来。因为只有在这种情况下，每一个分量中的振动才处于同一平面内。”[4]

1819 年，阿拉果和菲涅耳在《化学与物理学年刊》上，联名发表了题为《关于偏振光线的相互作用》的论文。论文一开始，他们概括了杨氏干涉实验所得的“美妙的结果”，又引入了光程的概念。他们指出：“在一切干涉现象中，如果两个不同介质的厚度与其折射率的乘积相等，则该不同介质产生同样的效应。”接着，他们详细地描述了他们所作的一系列偏振光干涉的实验，最后得出五点结论：

(1)振动方向相互垂直的偏振光不发生干涉。

(2)振动方向相同的偏振光发生干涉，与自然光产生的干涉相同。

(3)原来振动方向相互垂直的两束偏振光，当用某种光学方法，使偏振面转动到同一平面后，仍然不发生干涉。

(4)原来振动方向相同的一束偏振光,当用某种光学方法分裂成相互垂直的两束偏振光，转动偏转面后又回到同样的偏振状态，两束光发生干涉，与自然光的干涉相同。

(5)双折射光线的干涉条纹的位置，不只是由路程差和速度差决定，在某些情况下，还必须把半波损失考虑在内。[13]

但是，新的光的波动说也引起了关于以太的性质的许多问题，这甚至使菲涅耳在接受光的横波假设之前，也经过一段犹豫。他提出：“这个假设与公认的弹性液体振动本质的概念如此矛盾，以致我长久以来不能决定采用它；甚至当全部事实和长久的思考使我相信这个假设对于说明光学现象是必要的……”[2]他指出：纵波可以通过气体介质传播，而横向振动只能在固体物质中产生。但很难设想一种能传播横波的固态以太，却能让天体自由通过。杨氏在谈到菲涅耳关于光波系统是由垂直于传播方向的两个相互垂直的振动所组成的假设时，写道：“菲涅尔先生的这个假设，至少应当被认为是非常聪明的。利用这个假设可以进行相当满意的计算。可是，这个假设又带来了一个新问题，它的后果确实是可怕的……到目前为止，人们都认为只有固体才具有横向弹性。所以，如果承认波动理论的支持者

们在自己的‘讲稿’中所描述的差别，那么就可以得出结论说：充满一切空间并能穿透几乎一切物质的光以太，不仅应当是弹性的，并且应当是绝对坚硬的！！！”[4]

菲涅耳的研究成果，标志着光学进入了一个新时期——弹性以太光学的时期。这个学说的成功，在牛顿物理学中打开了第一个缺口。由于菲涅耳的工作给波动理论奠定了如此牢固的基础，他被人们称为“物理光学的缔造者”。

在 1850 年进行了一项由阿拉果首先建议的，由傅科(L.Fou-cault)、斐索(A.H.L.Fizeau，1819—1896)和布雷格特(L.Bre-guet)作的关于光速的仲裁实验。当光从光疏介质进入光密介质时，按照微粒论者的解释光在光密介质中的光速比较大；相反，波动论者依据惠更斯作图，要求光密介质中的光速比较小。傅科等直接测量了空气和水中的光速，结果毫无疑问，判定波动理论获胜。相对于波动理论所取得的胜利来说，这个仲裁实验显得近乎多余了。[3]

(四)光是电磁波的发现

当光的弹性以太理论遇到许多困难的时候，电磁学的一系列发现，揭示了光与电磁的内在联系，证明了光是电磁波。

1845 年法拉第发现了光的振动面在强磁场中的旋转。这表示光学现象与磁学现象间存在内在的联系。当他用一束偏振光顺着磁力线方向透过置于强电磁铁的两个磁极之间的“重玻璃”时，利用尼科耳棱镜，他发现，光的偏振面发生了一定角度的偏转，磁力越强，偏转角越大。这就是法拉第的磁致旋光效应。这个发现载于他的《电学的实验研究》第十九部分。法拉第兴奋地说：“我确信，光与电和磁的关系是从这里开始被发现的”“这件事更有力地证明，一切自然力都是可以互相转化的，有着共同的起源。”[2]我们知道，这种效应实际上是磁场使位于其中的物质受到影响，间接地使光的偏振面发生旋转，并非磁场对光的直接作用。

在电磁学中，电量的单位有静电单位与电磁单位。电量的静电单位是根据库仑定律定义的：一个静电单位的电荷，对一个相距一厘米远的同样电荷的排斥力是一达因。在电磁单位中，电流强度的单位定义为：在两根相距一厘米的长平行导线上，当它们的每单位长度彼此以二达因的力相互作用时所流过的电流。由此就可以得到电量的电磁单位的定义：单位电流强度在单位时间内流过的电量。1856 年韦伯和柯尔劳斯在莱比锡做的电学实验结果，发现电荷的电磁单位和静电单位的比值等于光在真空中的传播速度，即 3×10^8m/s。[10]这一惊人的结果进一步揭示了电磁现象和光现象之间的联系，这是对于光的电磁理论具有根本性意义的一个重要发现。[2]

1865 年，麦克斯韦发表了一篇著名的论文《电磁场的动力理论》。在这篇文章中他提出了完整的电磁场方程组。从方程组推出了电场强度 E 和磁感应强度 B 的波动方程。方程表明电场和磁场以波动形式传播，两者相互垂直并都垂直于传播方向。变化的电场和磁场构成了统一的电磁场，它们以横波的形式在空间传播，形成电磁波，并求出电磁波的传播速度为

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

μ和ε分别为介质的磁导率和介电常数，于是在真空中的$\frac{1}{\sqrt{\varepsilon_0\mu_0}}$之值等于电量的电磁单位与静电单位之比，其值为 3×10^8m/s 恰好等于由实验测定的光速。这个奇妙的结果促使麦克斯韦在他的思想里实现了一个极具创造性的巨大飞跃：“两个结果的一致性表明，光和磁乃是同一实体的属性的表现，光是一种按照电磁定律在场内传播的电磁扰动。”[2]后来，他在给威廉·汤姆孙的信中写道：“我得出了自己的方程。我住在外省，我并不怀疑，我所得出的磁效应的传播速度与光速很接近，因此我想，我有一切根据可以认为，磁的介质和光传播的介质是同一个介质……。”[4]

1868 年，麦克斯韦发表了一篇短而重要的论文《关于光的电磁理论》，明确地把光概括到电磁理论中，这就是著名的麦克斯韦创立的电磁波学说。这样，麦克斯韦就把原来相互独立的电学、磁学和光学这三个重要的物理学领域结合起来。成为 18 世纪中叶物理学上实现的一个重大综合。[2]

1886 年 10 月，德国物理学家赫兹(Heinrich Hertz，1857—1894)在作放电实验时，偶然发现近旁一个线圈也发出火花，赫兹敏锐地想到这可能是电磁共振。于是他从 1886 年 10 月 25 日开始集中力量进行验证麦克斯韦的电磁波是否存在的实验。1886 年 12 月 2 日赫兹发现“两个电振荡之间成功地引起了共振现象”。他用火花隙振荡放电作电磁波发射体，通过改变感应线圈火花间隙处所附的金属电极的形状和大小来控制火花振荡的频率。用一个开口线圈作接收天线。开口线圈的两端装有两个小铜球。在这两个小铜球之间，每通过一次火花，就表示接收到了一个入射电磁波。当他把接收天线放在感应圈附近，他观察到当感应圈两极间有火花跳过时，接收天线的间隙处也有火花跳过，从而证实了电磁波的存在。[2]

1888 年，赫兹直接从频率和波长测量了电磁波的传播速率。[10]接着他又用这个简单的仪器证明了这种电磁辐射具有和光类似的特性，包括在固体表面上的反射，用金属凹面镜聚焦，通过小孔时的衍射，显示干涉效应，以及通过非导体材料时的折射等等，至此就确立了光的电磁理论的基础。[2]

1896 年洛伦兹(H.A.Lorentz)创立了电子论，他认为原子和分子内含有带负电的电子，在无外力时电子处于平衡位置；在外力作用下，电子作阻尼振动而产生光的辐射，当光通过介质时，介质中电子的自然频率与外场的频率相同时，则受缚电子成为吸收体。这样，利用洛伦兹电子论不仅解释了发光和物质吸收的现象,也解释了光在物质中传播的各种特点。[14]

光的电磁理论在整个物理学的发展中起着很重要的作用，它指出光和电磁现象的一致性，并且再一次证明了自然现象存在着相互联系这一辩证唯物论的基本原理，使人们在认识光的本性方面向前迈出了一大步。

(五)爱因斯坦的光量子假设光的波粒二象性的统一

19 世纪末到 20 世纪初，光学的研究深入到光的发生、光和物质相互作用的微观机构中，在解释光电效应现象时，近代物理学革命的先锋爱因

斯坦(Albert Einstein， 1879—1955)提出了光量子假设。

1887 年，赫兹在进行证明电磁波存在的实验时，发现了光电效应现象。他注意到当接收电磁波的电极之一受到紫外光照射时，两极之间就容易出现电火花。他在 1887 年发表的《论紫外光的放电效应》一文中，首先描述了这些现象。1889 年，霍尔瓦克斯(W.Hallwachs，1859—1922)指出如果用光照射锌、钠、钾等金属表面，就会有负电粒子释放出来。赫兹的助手勒纳德(P.Lenard,1862—1947)在 1902 年发表了对光电效应的第一批定量研究结果，他测量了在紫外光照射下，铝板发出的电子的荷质比。他确信赫兹看到的火花加强的现象是金属表面发射电子的结果。[15]他发现光电效应有三个为经典波动理论无法解释的主要性质：第一，出射光电子的动能只同入射光的频率有关。同入射光的强度无关；第二，每种金属表面都存在一个特征截止频率，频率小于 v_0 的入射光不管其强度有多大，都不能发生光电效应；第三，只要入射光的频率大于截止频率，则无论它多么微弱，都会立即引起光电子发射，不存在滞后时间。[2]

这三点对经典电磁理论提出了挑战。按照经典理论，光的能量是连续的，电子从光波的波阵面上连续获得能量，能量的大小应当与光的强度有关，而光强与光矢量(即电场强度 E)振幅的平方成正比，与频率无关。但实验结果表明，光电子的逸出只与入射光的频率有关。只要入射光的频率 ν 大于截止频率 v_0，不论光强多么微弱，就有光电子发生，如果 $v<v_0$，不论光强多强，也没有光电子产生。按照经典理论，光能量分布在波面上，金属中的电子吸收能量的范围是有限的，电子从金属中逸出需要一段积累能量的时间，但实验表明光电子的逸出是瞬时的，不存在滞后时间。光电效应使经典理论陷入了困境。

1905 年，年轻的爱因斯坦勇敢地抛弃了经典理论的传统偏见，接受了普朗克(Max Planck，1858—1947)提出的能量量子化概念，发表了题为《关于光的产生和转化的一个启发性观点》的论文。他写道：“用连续空间函数来运算的光的波动理论，在描述纯粹的光学现象时，已被证明是十分卓越的，似乎很难用任何别的理论来替换。”“但是可以设想，当人们把用连续空间函数进行运算的光的理论应用到光的产生和转化的现象上去时，这个理论会导致和经验相矛盾。”[16]爱因斯坦在指出了经典理论的缺点后，接着，提出了他的光量子假设。他写道：

“在我看来，关于黑体辐射，光致发光，紫外光产生阴极射线，以及其他一些有关光的产生和转化的现象的观察，如果用光的能量在空间不是连续分布的这种假设来解释，似乎就更好理解。按照这里所设想的假设，从点光源发射出来的光束的能量在传播中不是连续分布在越来越大的空间之中，而是由个数有限的、局限在空间各点的能量子所组成，这些能量子能够运动，但不能再分割，而只能整个地被吸收或产生出来。”[16]

爱因斯坦在这里进一步发展了普朗克的能量量子化思想。他指出光不仅在发射或吸收时具有粒子性，光在空间传播时也具有粒子性。也可看作光是一粒一粒以光速 c 运动的粒子流。这些光粒子称为光量子，也称为光子。每一个光子的能量 $\varepsilon=h\nu$，与频率 ν 成正比，h 是普朗克常量。

按照光子假说，光电效应解释如下：当光子照射到金属表面时，一次为金属中的电子全部吸收，而无需要积累时间。电子把光子能量的一部分变成它逸出金属表面所需的功 W。另一部分转化为光电子的动能

$$hv = \frac{1}{2}mv^2 + w_0$$

这就是爱因斯坦光电效应方程。由这一方程可以直接解释前述的光电效应的几个性质。

至此，人们一方面从光的干涉、衍射和偏振等光学现象证实了光的波动性；另一方面又从黑体辐射、光电效应、康普顿效应证实了光的粒子性。如何将光的本性的两个完全不同的概念统一起来，人们进行了大量的探索工作。1924 年德布罗意创立了物质波学说，他设想每一物质粒子的运动都和一定的波相联系。这一假设在 1927 年为戴维孙(C.J.Davisson，1881—1958)和革末(L.H.Germer，1896—1971)的电子束衍射实验所证实。事实上，不仅光具有波动性和微粒性，也就是所谓波粒二象性，而且一切实物粒子同样具有这种二重性。1925 年玻恩(M.Born1882—1970)提出波粒二象性的概率解释，建立了波动性和微粒性之间的联系。电子的双缝衍射实验表明：单个粒子在何处出现有一定的偶然性，但大量粒子的分布表现为具有波动性。这就是微观粒子波动性的统计解释。今天人们对光的本性的认识还远远没有达到最后境地，还要不断探索、不断前进。

参考文献

〔1〕〔德〕M·V·劳厄著，范岱年、戴念祖译，《物理学史》，商务印书馆，1978 年，11—12

〔2〕申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月， 229， 367—370， 377—378，385—391，572，579—580，594，589—591， 721—722

〔3〕M·玻恩、E.沃耳夫著，杨葭荪译校，《光学原理》，上册，第一版，科学出版社，1978 年 12 月，1—12，487

〔4〕〔美〕库珀著，杨基方、汲长松译，《物理世界》，上卷，海洋出版社，1981 年， 247—252，303—310，313—316

〔5〕〔美〕威·弗·马吉编，蔡英牟译，《物理学原著选读》，第一版，北京，商务印书馆， 1986 年 5 月， 287—289

〔6〕张瑞琨，《近代自然科学史概论》上海，华东师范大学出版社，1986 年 9 月，167—172

〔7〕同〔5〕，322—323

〔8〕〔美〕乔治·盖莫夫，高士圻译，《物理学发展史》，北京，商务印书馆，81，85—86

〔9〕同〔5〕，306－310

〔10〕姚启钧，《光学教程》，第三版，人民教育出版社，1983 年 3 月，5—8〔11〕同〔5〕，327

〔12〕同〔5〕，311－314

〔13〕同〔5〕，342－351

〔14〕母国光、战元令编，《光学》，第一版，人民教育出版社，1979 年 3 月，5—8

〔15〕沈珊雄，“爱因斯坦与光量子假设”，《物理教学》，1979 年 2 月

〔16〕许良英，范岱年编译，《爱因斯坦文集》，第二卷，北京，商务印书馆，37－38

十、黑体辐射定律的建立及普朗克常量的发现

普朗克常量是表示量子物理特征的基本常量，它的发现是物理学史上一次重大革命的开始。可以说物理科学的新时代是随着普朗克常量的发现而开始的。有了普朗克常量才有人们对微观世界进一步的认识，并且在此基础上建立起一个完整的量子理论体系，成为现代物理学的重要组成部分。普朗克常量是普朗克在建立黑体辐射定律中发现的。

本文的主要任务是探索在这一发现中普朗克的思想发展过程。他的科学发现是在什么背景下产生的？他受到了哪些事情的启发？他的研究具有什么特色？他又经历了哪些艰难曲折才走上量子论道路的？

(一)黑体辐射的经验定律

在 19 世纪开始的时候，天文学家赫谢耳(F.W.Herschel，1739—1822)发现了红外辐射的热效应。[1]他在实验中用灵敏温度计测试太阳光谱各部分的热效应，结果发现在红外光谱以外的区域温度升得最高，他认为在可见的红光之外还有不可见的辐射，这就是通常所指的热辐射。以后物理学家们对于热物体发射的辐射感到有兴趣，为了研究谱线的可见光部分，使用了照像的方法，对于红外区域即热辐射部分用热电偶测量。[2]

在实验发现的基础上，理论研究也活跃起来了，总结实验发现的经验规律也就相继地提出来了。1859 年德国物理学家基尔霍夫(G.R.Kirchhoff，1824—1887)得到如下结论：“在相同的温度下同一波长的辐射本领与吸收系数之比对于所有物体都是相同的，是一个取决于波长和温度的函数。”如果这一函数用φ(λ，T)表示，物体的辐射本领，即从物体表面单位面积上所发射的波长在λ附近的单位波长间隔的辐射功率用e(λ，T)表示，物体的吸收系数，即物体在波长λ和λ+dλ范围内吸收的能量与入射能量的比率用 a(λ，T)表示，则当物体处在辐射平衡时有

$$\frac{e(\lambda,T)}{a(\lambda,T)}=\phi(\lambda,T)$$

当物体的吸收系数 a=1 时，φ(λ、T)就是该物体的辐射本领。在 1860 年，基尔霍夫把 a=1 的理想物体定义为“绝对黑体”，这种黑体在任何情况下能够吸收射在它上面的一切热辐射，所以对绝对黑体的研究成为寻找基尔霍夫函数φ(λ，T)的关键。[3]

1864 年，英国物理学家丁铎尔用加热空腔充作黑体测定了单位表面积、单位时间内黑体辐射的总能量与黑体温度的关系。1879 年德国物理学家斯特藩(Josephstefan，1835—1893)从丁铎尔和法国物理学家所作的测量中导出，黑体单位表面积在单位时间内发出的热辐射总能量 W，与它的绝对温度 T 的四次方成正比，即

$$W=\sigma T^4$$ [3]

式中σ为“斯特藩-玻耳兹曼常量”。但是，斯特藩-玻耳兹曼定律只反映了总的辐射能与温度的关系，不能反映辐射能随波长的分布。1881 年，美国物理学家兰利(S.P.Langley，1834—1906)发明了测辐射仪，用极细薄的铂丝作为惠斯通电桥的两臂，用灵敏电流计检测，可测出 1×10^{-3}

℃的温度变化，大大提高了热辐射能量的测量精度。[2]他虽然没有得到精确的分布定律，却已发现分布曲线并不对称，而且最大能量随温度升高而向短波方向移动。 1893 年，德国物理学家维恩(Wilhelm Wien， 1864—1928)由电磁理论和热力学理论得到了维恩位移定律

$$\lambda_m T=常量$$

此式表明辐射中能量最强的波长λ_m与黑体的温度成反比。

19 世纪末叶，人们对热辐射的规律性，尤其是对黑体辐射能量按波长分布的函数的研究产生了浓厚的兴趣。这是因为那时城市照明提到日程上来了，人们探求新的光源，寻找最有效的发光方式。由于对星体表面测温和工业上高温测量的需要，有必要对辐射能量按波长分布的函数曲线与温度的关系进行详尽的研究。在承认光是电磁波后，人们开始系统地探索这些波的全部频谱，发现完全新的辐射形式。由于欧洲和美国日益增长的工业发展的需要，促进了测量热辐射技术的发展，一些特殊的国家研究机构和实验室也应运而生。在这些研究机构中首推柏林的物理技术研究所，它在 1887 年成立时得到了西门子电力公司的创立者的资助，在 19 世纪快结束的时候，为了提供涉及黑体辐射能量分布的基本数据，在这个研究所里发展起来了各种各样精确测量热辐射的实验方法。[1]

1895 年，德国物理学家卢默尔(Otto Lummer，1860—1925)和维恩指出，由不透射任何辐射的器壁围住的带有一个小孔的空腔，它的热辐射性能等同于黑体，辐射空腔的实现为研究黑体辐射提供了重要手段。

在这期间，德国物理学家帕邢(Paschen，1865—1947)进一步发展了热辐射谱和光谱的测量技术，他测量了各种不同种类物质的热辐射谱，在他的领导下，杜宾根(Tubingen)大学成为了光谱学研究的重要基地。[1]在这些实验的基础上，1896 年，帕邢确立了黑体辐射能量按波长分布的经验公式，即基尔霍夫函数的具体形式为

$$\phi(,T\lambda)=c_1\lambda^{-a}\exp\left(-\frac{c_2}{\lambda T}\right)$$

在这里 c_1 和 c_2 是常量，α 为待定常量。在帕邢建立他的经验定律前不久，维恩在理论论证的基础上已达到同一结果。[1]

在探求辐射空腔中能量密度分布函数$\rho(\lambda,T)$的过程中，维恩作出了杰出的贡献。他从纯热力学理论出发建立了一个辐射能量随波长λ和温度 T 分布的维恩公式。它是由研究“平衡辐射的绝热膨胀”而获得的。首先考虑一个具有完全反射壁的球壳，其中放置一块黑体，在温度 T 达到平衡后将黑体取出，此时球壳中充满黑体辐射。然后设想辐射作绝热膨胀，即设想球壳以缓慢的匀速向外胀大，其温度自然也要发生变化，不过辐射的本质并不因此而改变，仍属黑体辐射。由于球壳壁运动必有多普勒效应产生，因而引起辐射的频率ν或波长λ的变化。通过简单的计算可知，波长与半径成正比；由热力学还可以证明λ与绝对温度 T 的乘积为一常量。由于发生了绝热膨胀，辐射能密度也要改变，即球壳中每单位体积的能量也要相应地改变。[4]可以证明对应于波长λ的辐射能密度ρ与波长的五次幂成反比。因此

$$\rho(\lambda,T)=A\lambda^{-5}\cdot f(\lambda,T)$$

或者以频率表示，可得

$$\rho(v,T) = Bv_3 \cdot \varphi\left(\frac{v}{t}\right)$$

式中 A 和 B 是常量。

这就是维恩公式。1896 年，维恩利用上述公式推得了明晰的分布函数ρ(λ、T)。在推导时，他假设黑体辐射的能量按频率分布，和同温度的理想气体分子的能量按麦克斯韦速度分布律的分布相类同。[4]于是推出

$$\rho(\lambda,T) = c_1\lambda^{-5}\exp\left(-\frac{c_2}{\lambda T}\right)$$

或者以频率表示，可得

$$\rho(v、T) = av^3\exp\left(-\frac{bv}{T}\right) \qquad (10.1)$$

由此看来，维恩定律与帕邢的经验定律是一致的，只要人们使ρ(λ、T)与基尔霍夫函数φ(λ，T)相等，对于帕邢的幂指数值取 5，于是它就精确地重现了观察到的数据。

1900 年 6 月，英国物理学家瑞利(Rayleigh，1842—1919)发表了黑体辐射理论的研究结果，他假定辐射空腔内的电磁辐射形成一切可能的驻波，而根据经典的能量均分定理，每一驻波平均具有能量 kT，由此导出

$$\rho(v,T) = \frac{8\pi v^2}{c_3}kT$$

式中$\frac{8\pi v^2}{c^3}$为驻波数，kT是按照古典的能量均分定理空腔中的每个驻波分配到的平均能量。[5]瑞利的推导中错了一个因数，后来年轻的英国天文学家金斯(T.H.Jeans，1877—1946)投书《自然》杂志作出纠正，故称为瑞利-金斯公式。

这个公式虽然在低频部分与实验符合，但由于辐射的能量与频率的平方成正比，所以随频率增大而单调增加，在高频部分出现趋于无限大，即在紫端发散，后来这个失败被埃伦菲斯特(P.Ehrenfest，1880—1933)称为“紫外灾难”。[3]这个灾难正是经典物理学的灾难。所以开尔文在本世纪的开始，1900 年 4 月 27 日，在英国皇家学会作的题为《在热和光的动力理论的上空的 19 世纪的乌云》的讲演(1901 年 7 月发表在《哲学杂志》和《科学杂志》合刊上)中，开尔文把与“紫外灾难”相联系的能量均分学说比做经典物理学晴空中的第二朵乌云。在这篇文章的最后，他提到瑞利在 1900 年 6 月发表的《论能量均分定律》的论文，在这篇论文里瑞利已经看到了能量均分定理的困难，他力图寻求一条不破坏能量均分定理又能解决这一困难的途径。开尔文与瑞利的做法针锋相对，他明确宣布：“达到所期望结果的最简单的途径就是否认这一结论。”他满怀信心地预言：“对于在 19 世纪最后四分之一时期内遮蔽了热和光分子论亮光的这朵乌云，人们在 20 世纪就可以使其消散。”[6]历史发展表明，这朵乌云终于由量子论的诞生而拨开了。

(二)普朗克前期的物理思想

德国杰出的物理学家普朗克(Max Planck，1858－1947)出生于德国的基尔城，他的父亲是法学教授。1874 年 10 月，普朗克进入慕尼黑大学，

最初决定主攻数学，但很快又被物理学所吸引。他的老师约利(P.Jolly)曾极力劝说他不要研究物理。约利告诉普朗克：“在这一学术领域里，已经没有什么本质上新的东西有待发现了。”但是普朗克还是坚持抛弃纯数学，因为他对宇宙本质问题有浓厚兴趣。[7]他在《科学自传》中谈到了他从青年时期起就爱好科学的原因。他说：“引导我从事科学研究和从青年时期就爱好它的原因，是一个不十分自明的事实，这就是我们的思维规律和我们从外界接受到的自然过程的规律是符合一致的，因而使人们有可能通过纯粹思维对这种规律作出解释。对此具有重要意义的是，外部世界是我们所面对的、独立于我们而存在的绝对所在，而探索这种绝对存在所适用的规律，我认为就是最崇高的科学研究任务。[8]这些话表明了一个科学家朴素的唯物主义思想，他坚信外部世界是独立于我们而存在的客观存在，他坚信真理的客观性，他认为探索这种绝对存在所适用的规律是科学研究的崇高任务。

普朗克早期的物理思想受到克劳修斯的深刻影响。当他在慕尼黑大学学习物理和数学时，就以极大的热情自学了克劳修斯的名著《热力学》。他在《科学自传》中，在谈了课堂教学的不足之处后紧接着说：“在这种情况下，我只能通过自学我所感兴趣的书刊来满足我的求知欲望，这时我所学习的自然都和能量定律有关。我在偶然中得到了克劳穆斯的一本著作，它的明白易懂的语言和深入浅出的叙述给我留下了非常深刻的印象，我钻研它们的兴趣越来越高。”[8]克劳修斯的一些主要热力学观念，如不可逆性、熵增加原理等给他留下了极其深刻印象。他把克劳修斯称为的‘不可逆过程”命名为“自然过程”，认为“在一个不可逆过程中，终态在某种意义上更优越于始态，自然界似乎对它有更大的‘偏爱’。克劳修斯的熵给出了这种偏爱的量度，而且下列叙述具有第二定律的意义，在每个自然过程中，所有参加该过程物体的熵的总和在增加。”[8]普朗克在 1879 年完成的慕尼黑大学的博士论文中，对热力学第二定律作了详细的论述。但这篇论文并没有引起人们的兴趣。

但是，这种淡漠的态度并没有阻止他对熵的研究。他在《科学自传》中说：“由于深刻体会到这个问题的重大意义，这些遭受并没有阻止我继续对熵进行研究，因为我把熵看作和能量一样，也是物理过程最重要的特性。”[8]1880 年 6 月 14 日，普朗克写了《不同温度条件下物体的平衡熵》一文，获得慕尼黑大学授于的特别奖状。[7]所以，我们可以看到在普朗克的热力学思想中，熵增加原理和非平衡态向平衡态发展过程的不可逆性，始终占有最根本的地位。他把熵增加原理看成和能量原理一样是物理学中一条不可缺少的独立定律。这个在他青年时期形成的中心思想，一直主导着他的科学工作。当然也就决定性地影响了他处理黑体辐射问题的方式。[9]

1887 年，哥廷根大学哲学系举办关于能量本质的有奖竞赛。普朗克完成了获奖论文《论能量守恒定律》之后，再一次转向“第二定律”，并在三篇论文中对该定律加以概括，力图使它的应用扩大到电学等领域中去。[7]他在 1891 年就更为明白地肯定，“熵的增长定理必须扩大到所有的自然力……不仅包括热过程、化学过程，还应包括电及其它过程。”[7]

从 1894 年起，普朗克把注意力转向黑体辐射问题，当时他已知道按照基尔霍夫定律辐射本领与吸收率的比值与物体的性质无关，只决定于温度

和波长。他在谈到他的探索的动机时说：“这个所谓的正常的能量分布表明了某种绝对东西的存在，而探索这种绝对东西的存在，我认为就是科学研究的最崇高的任务，因此我一直刻苦勤奋地探索这个问题。”[8]

普朗克从 1895 年起开始持续进行了五年之久的关于空腔共振子体系“不可逆辐射过程”的系统研究，其根本目的就是试图论证在这样一个封闭体系内部热辐射过程在严格意义上是不可逆的。最初，他希望避免明显的原子分子论假定和统计方法，单单利用电动力学来做到这一点。[9]

普朗克分别于 1895 年和 1896 年发表了两篇论述“共振子”(resonator)与辐射场相互作用的论文。他研究了封闭在一个具有理想反射壁的空腔里的电磁辐射。他假定空腔是由最简单的“共振子”、即线性赫兹振子集聚而成。每个振子各有其频率，其作用如共振器，可以吸收周围辐射中相同频率的能量而受到激发，同时又发射能量而减弱，一般来说，发射的能量与吸收的能量不同，通过共振子与辐射场的相互作用而建立起平衡态。[1]普朗克假定振子线度极小，即使在一个不大的空腔中，也可以忽略它的具体结构而把它视为点偶极辐射中心。[9]

1897 年 2 月普朗克在柏林普鲁士科学院的会议上宣读了他《关于不可逆辐射过程》研究的第一篇正式报告，论证空腔共振子体系电磁过程的不可逆性，普朗克认为只要电动力学就够了，不必考虑统计方法，因此立即招来了玻耳兹曼的批评。玻耳兹曼指出普朗克肯定达不到他的目的，因为按照严格的麦克斯韦方程，空腔内的电磁过程完全是可逆的。尽管普朗克用电动力学观点对玻耳兹曼作了答复。他毕竟还是在这年 12 月在他的《关于不可逆辐射过程》的第三篇报告中，修改了他最初宣布的目标，不得不引进他一向感到厌恶的统计方法，特别是他被迫从他的论敌玻耳兹曼那里借用了一个关键性的统计概念“无序性”，证明空腔共振子体系电磁过程的不可逆性。[9]

1898 年 7 月普朗克发表了《关于不可逆辐射过程》的第四篇报告，论证空腔共振子体系电磁过程存在一个随时间单调变化的量，他引入了一个统计假定——自然辐射假定，即不可逆辐射过程的假定。在这项假定的基础上，普朗克最后证明了，存在一个被他称为“熵”的函数，这个函数将随时间单调上升。[9]

1899 年 5 月普朗克发表了《关于不可逆辐射过程》的第五篇报告。根据经典电动力学理论，使谐振子平衡时的发射率与吸收率相等，[10]1899 年普朗克就得到振子的平均能量 U 和具有同一频率的入射辐射能密度之间的关系

$$\rho(v,T)=\frac{8\pi v^2}{c^3} \qquad (10.2)$$

由于当时普朗克不仅不熟悉，而且拒绝热的统计理论，所以他没有使用由能量均分定理导出的公式 U=kT，不然他就会在瑞利之前得到瑞利-金斯辐射公式。[8]

普朗克是带着对不可逆问题的浓厚兴趣转向黑体辐射研究的，他把熵增加原理置于考虑的首位是很自然的。这就体现出普朗克处理问题的特色。普朗克认为热力学第二定律适用于自然界一切过程，一切物体。所以他寻求“振子的能量和熵之间的合适关系，而不是把振子的能量和温度联系起来”。[8]他从热力学来看，在体积一定的条件下

$$dS = \frac{dU}{T}$$

由维恩公式(10.1)和他导出的辐射公式(10.2)，得出

$$U=a'\nu e^{-a\nu/T}$$

解出 $\frac{1}{T}$ 后可得

$$\frac{dS}{dU} = -\frac{1}{a\nu}\ln\frac{U}{a'\nu}$$

$$\frac{d^2S}{dU^2} = \frac{\text{常量}}{U} \qquad (10.3)$$

普朗克将 $\frac{d_2S}{dU^2}$ 的倒数用R表示，则得出

$$R=U\times\text{常量}$$

由于这个关系十分简单，所以普朗克认为它是一个有特征性的普适关系。后来他说："由于整个问题是在研究一条普遍的物理定律，而且由于我以前同现在一样，相信物理定律越带普遍性就越是简单，……所以我一度认为，我从量 R 与能量成正比这一点看到了它是整个能量分布定律的基础所在。"[3]正是由于这种认识，对于一个振子的熵，普朗克写出了下列定义

$$S = -\frac{U}{a\nu}\ln\frac{U}{ea'\nu} \qquad (10.4)$$

他进而继续证明，总熵是时间的单调增函数。他还证明从式(10.3)出发结合式(10.2)可以很容易得出维恩公式。普朗克的上述论证步骤，看起来似乎是经由维恩定律的某种循环推论，但是普朗克所做的是把维恩公式关于一个函数的经验性猜测变为由系统熵变出发的一个假定。如果有办法从更基本的物理考虑给出熵的定义，那么通过这条途径就可以对辐射公式给出一种严格的理论论证。普朗克意识到他选定熵的特殊表示式作为能量的函数决定于最终的能量分布律，因而对能量分布律的实验验证就显得十分重要了。他在 1899 年 5 月向科学院提交的一篇论文中表述了如下的想法："我认为，这必然会使我得出这样的结论，即辐射熵的定义，因而还有维恩的能量分布定律，两者必定都是通过将熵增加原理应用于电磁辐射理论而得出的，因而这条定律有效性的限度，如果它存在着这种限度的话，将会与热力学第二定律所受到的完全相同。显然，这使对这条定律再做一番实验研究更显得极端重要了。"[3]

（三）普朗克黑体辐射定律的建立

1899 年底，普朗克得知德国实验物理学家卢梅尔和普林斯海姆等人在 1899 年 9 月发表的实验报告中指出了维恩定律仅在短波段内与实验相符，而在长波范围内则有明显的偏离。这说明维恩公式并不是一个真正符合客观实际的辐射公式，需要作进一步的修正。[11]1900 年 10 月 7 日德国实验物理学家鲁本斯(H.Rubens，1865—1922)夫妇访问了普朗克，鲁本斯告诉他在长波段，分布函数趋于一个完全不同的形式，变成正比于绝对温度 T 了。这使普朗克受到启发，他立即尝试用"内插法"去寻求新的辐射公式。他当天就得到了所要求的辐射公式。

普朗克的辐射公式是依据熵对能量二阶导数的两个极限值进行内推而

得到的，其中一个极限值对应于热辐射谱短波段，由维恩公式确定，熵对能量的二阶导数与能量成反比。对于热辐射谱的长波段，是根据鲁本斯等人的测量结果而得到。他们发现对于长波领域黑体辐射强度与温度成线性关系。普朗克在他的计算中吸收了这一实验结果，从而确定了熵对能量的二阶导数与能量的平方成反比。据瑞利-金斯公式，振子能量 U=kT，则

$$\frac{dS}{dU}=\frac{\text{常量}}{U}$$

$$\frac{d^2S}{dU^2}=\frac{\text{常量}}{U^2} \qquad (10．5)$$

或 $$R=U^2\times\text{常量}$$

于是，为了能够得到普适的情况，可以令 R 等于能量 U 的一个一次项与一个二次项之和，以便对小能量、对短波，一次项起决定作用；对大能量、对长波，二次项起决定作用。[8]由此可假定

$$\frac{d^2S}{dU^2}=\frac{a}{U(\beta+U)} \qquad (10.6)$$

通过对上式积分可得 $\frac{dS}{dU}$ 的表示式，令它等于绝对温度的倒数，于是得

$$\frac{dS}{dU}=\frac{\alpha}{\beta}\ln\frac{U}{\beta+U}=\frac{1}{T}$$

于是普朗克得到振子的能量方程为

$$U=\frac{\beta}{e^{-\beta/aT}-1} \qquad (10.7)$$

上式虽是关键性的，但是它并没有把主要参量频率 v 引入能量表示式中，为了解决这一问题普朗克把公式(10.2)与维恩公式(10．1)来比较，可知 U 必须满足

$$U=v\varphi\left(\frac{v}{T}\right)$$

所以式(10.8)的形式必须是

$$U=\frac{c_1 v}{e_2^{c'v/T}-1} \qquad (10.8)$$

则

$$\rho(v,T)=\frac{c'_2 v^3}{e_2^{c'v/T}-1} \qquad (10.9)$$

或

$$\rho(\lambda,T)=\frac{c_1\lambda^{-5}}{e^{c_2/\lambda T}-1} \qquad (10.10)$$

当天晚上普朗克就把他推导出的这个辐射公式写在明信片上寄给鲁本斯。鲁本斯接到普朗克的信后，把这一公式和他的测量数据作了认真比较，发现令人满意地符合一致。两天后鲁本斯又来到普朗克家，给他带来了新的公式与实验观察完全一致的消息。[1]普朗克后来说：“以后的测量，也一而再，再而三地证实了我的辐射公式——所用的方法越精密，越能看出我的公式更为正确。”[10]1900 年 10 月 19 日在德国物理学会会议上，普

朗克作了《维恩辐射定律的改进》的报告，提出了他的这一新的辐射公式。

(四)能量量子化假设的提出以及普朗克常量的发现

虽然实验证明了普朗克公式对于全部波长对于所有温度的正确性，然而它是一个侥幸猜中的内推公式，并不具备明确的理论基础。对于普朗克来说他的最重要的任务就是为新的公式提供理论基础。后来，普朗克回忆说：“即使这个新的辐射公式竟然能被证明是绝对精确的，但是如果把它仅仅看作是一个侥幸揣测出来的内插公式，那么它的价值也是有限的。由于这个缘故，从它于10月19日被提出之日起，我即致力于找出这个等式的真正物理意义。这个问题使我直接去考虑熵和概率之间的关系，也就是说，把我引到了玻耳兹曼思想。”[3]

在此之前由于普朗克喜欢毫无例外的普适性，对玻耳兹曼有关统计性的概率概念格格不入。他后来回忆说：“每一条概率规律都有例外情形，而我们那时认为热力学第二定律是毫无例外地有效的。”普朗克在1897年5月的信里对玻耳兹曼的概率提出过尖锐的批评：“诚然，如果事先一无所知的话，那么概率可用于找出最概然的态。但是如果取定一个已知态，那就不能用它来计算在此之后的态，因为后者不再是由概率而是由力学决定的，并且，自然界中的变化总是沿着由较低概率到较高概率的方向进行的假定，是完全没有根据的。”[9]普朗克认为概率的出现是由于我们对体系状态认识不完全，在我们对体系的真实情况缺乏足够信息的情况下，这种统计推断就可能发生错误。

现在为了解释他在用类推法导出辐射公式时所使用的熵的二阶导数公式，他不得不转向玻耳兹曼思想，接受熵的概率解释，这就引起了他思想上的激烈斗争。然而，普朗克知道：“这个问题(指辐射能的分布)对于物理学是至关重要的……因此一个理论上的解释必须不惜任何代价非把它找出来不可。”这个代价就是：除了维护热力学的两条定律之外，普朗克“准备牺牲我以前对物理定律所抱的任何一个信念”。[11]后来，他把这一转变过程称作“孤注一掷”的行动。以后普朗克在获得诺贝尔奖金的演讲中说：“这个问题导致我自动地去考虑熵和概率之间的联系，即导致玻耳兹曼的思想，直到我的生活中最艰苦的几个星期过去以后，光明冲破了黑暗，至今一个新的梦想不到的远景在我眼前出现了。”[1]

早在1877年，玻耳兹曼就发表了题为《论热力学第二定律和概率理论之间的关系》的论文。他在那里提出了熵的统计解释，指出一个系统的任意状态的熵和该状态的热力学概率的对数成正比，即 $S \propto \ln W$。这个概率被认为是对应于宏观态的微观状态数。每种微观态表示能量在分子中的一种分配方式，这种方式称为配容(Complexions)，每种配容出现的概率相等。而宏观态指的是能量份额的一种分布，即只要知道包含能量为 ε、$2\varepsilon \cdots M\varepsilon$ 的分子相应的分子数就可以了，并不要求具体知道是哪个分子。所以每一宏观态包含若干个微观态，而热力学的平衡态是最概然的状态，是包含微观态数量多的宏观态。然而值得注意的是，在计算的适当阶段，玻耳兹曼使 ε 趋于零，这样分子的能量值就是连续的，而不是分立的了，对于玻耳兹曼来说 ε 只不过是使计算成为可能的技巧而已。[10]

普朗克当时的任务是寻找确定一组谐振子配容数目 W 的方法，很显然不能让 ε 变为零，这正是普朗克工作的主旨。这样，他就很自然地引入了

能量不连续的假定，只有把能量分成一份一份地才能计算出确定的微观态数目。这样论及微观态的确定数目就与经典理论完全不同了，因为从经典理论看来，能量是连续的，不能有分立的微观态个数。所以，普朗克实际上已作了革命性的假定，所有可能微观态的总组合是分立的组合，一个系统的每一个宏观态对应有完全确定数目的微观态。然后他根据熵的可加性和热力学概率的可乘性得出下述关系

$$S=k\ln W[4] \qquad (10.11)$$

式中 k 为玻耳兹曼常量。

普朗克设想空腔内有数目很多的振子，它们是辐射着的物质中心，能够与周围电磁场交换能量。在这个系统中频率为 ν 的振子数为 N，总能量为 E，每个振子的能量只能取一些分立的值，它是有限能量元的整数倍。因为微观态数是分立的、有限的，这就要求 E 只能作有限的划分，例如只能划分为 P 个相等的小份额，即

$$E=P\varepsilon \qquad (10.12)$$

这些能量元 ε 在 N 个振子中可以按不同比例 nε (n=0，1，2，…P)分配给单个振子。每个振子的能量只能取一定的分立值，它们是能量 ε 的整数倍，ε 是能量的最小单元，它就是能量子。根据排列组合法则，配容的数目应当是

$$W=\frac{(N+P-1)!}{(N-1)!\,P!}$$

由于 N 和 P 都是很大的数，可以去掉 1，并利用斯特林公式

$$\text{in}x!=x\ln x-x$$

于是得

$$\ln W=(N+P)\ln(N+P)-N\ln N-P\ln P$$

因为 N 个振子系统的熵 S_M 是一个振子的熵 S 的 N 倍，

$$S_N=NS$$

利用公式 $S_N=k\ln W$

以及对N个振子求平均能量 $U=\dfrac{p\varepsilon}{N}$

则可得出

$$S=k\left[\left(1+\frac{U}{\varepsilon}\right)\ln\left(1+\frac{U}{\varepsilon}\right)-\left(\frac{U}{\varepsilon}\ln\frac{U}{\varepsilon}\right)\right]$$

因而

$$\frac{1}{T}=\frac{dS}{dU}=\frac{k}{\varepsilon}\left[\ln\left(1+\frac{U}{\varepsilon}\right)-\ln\frac{U}{\varepsilon}\right]$$

于是得出 $U=\dfrac{\varepsilon}{e^{8/kT}-1}$

黑体辐射公式则为

$$\rho(v,T)=\frac{8\pi v^2}{c^3}\frac{\varepsilon}{e^{8/kT}-1} \qquad (10.13)$$

现在这个通过熵与概率相联系而推得的理论性辐射公式应该与普朗克以前得到的符合实验的辐射定律相一致。必须令

$$\varepsilon=hv(10.14)$$

这个 h 是以比例常量姿态出现的，它是一个与振子特性无关的常量，数值为 6.55×10^{-27} 尔格·秒。它的意义十分重大，它说明不同频率的振子具有不同的能量子，与其频率之比总要等于恒量 h，h 的普适性在此已可见一斑，人们称 h 为普朗克常量。由于它的量纲等于能量与时间的乘积，也就是作用的量纲，所以普朗克把它称为作用量子，而 hν 则称为能量子。将 ε=hν 代入式(10.13)得黑体辐射公式为

$$\rho(v,T)=\frac{8\pi hv^3}{c^3}\frac{1}{e^{hv/kT}-1} \qquad (10.15)$$

或

$$\rho(\lambda,T)=\frac{8\pi ch}{\lambda^5}\frac{1}{e^{ch/\lambda kT}-1} \qquad (10.16)[4]$$

这就是 1900 年 12 月 14 日普朗克在德国物理学会上宣读的论文《关于正常光谱的能量分布定律的理论》中所提出的普朗克辐射公式。从以上分析看来，辐射公式是在能量子假定下建立起来的。普朗克在演讲中强调指出：

“我们采取这种看法——并且这是整个计算中最重要的一点——认为 E 是由一些完全确定的，有限而又相等的部分组成的，而对于这个有限而又相等的部分，我们应用了自然常量 $h=6.55\times10^{-27}$ 尔格·秒。”[3]

所以，这一天就被人们看作是量子论的诞生日。作用量子 h 是最基本的自然常量之一，体现了微观世界的基本特征，它既是支配电磁波和物质相互作用的基本量，又是表征原子结构的重要参数，是物理世界中的一个基本角色，随着这个角色的登台，物理科学就出现了一个新时代。

量子论的诞生冲击了经典物理学长期信奉的一切自然过程都是连续的这条原理。这个理论与经典理论如此格格不入，所以当时物理学界对它的反映是极为冷淡的，从 1900 年到 1904 年的文献中，几乎找不到有关普朗克工作的论文，普朗克的工作几乎被人们遗忘了。人们承认那个与实验相符的辐射公式，却不接受普朗克的量子假说。绝大多数物理学家都没有意识到普朗克的假说已揭开了物理学革命的序幕。

普朗克本人也受着经典物理传统观念的束缚，他发现了能量子，但对能量子并不理解。普朗克本人也对他的这一行动长期惴惴不安，总想回到经典理论的立场。1909 年后，当他重新投入量子问题的研究时，他告诫自己和其他人尽可能避免采取冒险步骤：“在将作用量子 h 引入理论时，应当尽可能保守从事；这就是说，除非业已表明绝对必要，否则不要改变现有理论。”在量子论已为广大物理学家所接受，成了国际物理学界的主要思潮后，普朗克却从自己原来的观点中退却下来。1911 年，普朗克认为只是在发射过程中才是量子化的，而吸收完全是连续进行的。到了 1914 年，又退了一步，干脆撤销了量子假说，而认为发射过程是连续发生的。[3]就在他多年的徘徊中，量子理论却已飞速地向前发展了，所以普朗克的这两个理论都受到了批评，这使普朗克最终放弃了倒退的立场。他后来在《科学自传》中回忆说：

“我一直试图把作用量子纳入经典物理学理论，这种尝试我曾持续了许多年，并且消耗了我许多精力，但仍然归于失败。我的许多同事把它视为一种悲剧，我则持不同意见，因为对我来说，通过这种最基本的研究而得到的益处是最有价值的。现在我已真正了解到，作用量子在物理学中所起的作用要比当初我所设想的广泛得多，重要得多，并且由此对研究原子问题必须引进完全新的方法和新的计算方法提供了全面认识。”[8]在经过

种种失败的教训之后。他得出要变革旧的物理思想，采纳新的物理思维的结论。他指出：“因为作用量子揭示了某种至今未知的东西，所以要求我们彻底变革我们旧的物理思想，而这种物理思想是从莱布尼兹和牛顿创立微积分以来一直建立在因果关系的连续性假设上。”[8]

爱因斯坦对普朗克的工作给以高度的评价。1948 年 4 月，在悼念普朗克的会上，爱因斯坦充分肯定了作用量子发现的重大意义。他说：“这一发现成为 20 世纪整个物理学研究的基础,从那个时候起几乎完全决定了物理学的发展。要是没有这一发现，那就不可能建立起分子、原子以及支配它们变化的能量过程有用的理论。而且，它还粉碎了古典力学和电动力学的整个框架，并给科学提出了一项新任务：为全部物理学找出一个新的概念基础。”[12]

爱因斯坦特别赞赏普朗克为追求真理的崇高的探索动机，他称颂普朗克“追求真理的理想在他身上体现得非常完满”。1918 年 4 月，在普朗克六十岁生日庆祝会上的讲话中他指出：在科学的殿堂里有各种各样的人：有人爱好科学是为了满足智力上的快感；有的人是为了纯粹功利的目的。而普朗克热爱科学是为了得到“现象世界那些普遍的基本定律。”他认为这是普朗克“无穷的毅力和耐心的源泉”。他强调普朗克追求真理的“激情”。从前面的叙述中我们看到在普朗克的科学生涯中，他既注意从前辈科学家那里吸取丰富的科学思想营养，又坚持独创的科学研究道路；他严谨、审慎、执著，一旦发现自己认识上的错误后，又能勇于抛弃自己的偏见，接受新的物理思想。正因为如此，他成为了“一个以伟大的创造性观念造福于世界的人”。[12]

参考文献

[1]Jagdish Mehra，The Historical Development of QuantumTheory，Vol.1， Springer-Verlag， New York，1982，26－27，35，44，48

[2] 郭奕玲等编著，《物理实验史话》，北京，科学出版社，1988 年，181

[3] 申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1986 年 1 月，660—666，709—713

[4] 杨朝潢，“能量子和作用量子的缘起”，《物理通报》，1964 在 2 月

[5]Robert Eisberg，Quantum Physics，Second Edition，TohnWiley and Sons， New York，1985， 6－12

[6] 李醒民，“开尔文勋爵的‘两朵乌云’”，《物理》，1984 年 11 月

[7] H.坎格罗，“普朗克”，《科学与哲学》，1980 年第 1、2 辑

[8] 普朗克，《科学自传》摘译，《科学史译丛》，1988 年 1 月参照 Planck， A Scientific Autobiography，作了校正

[9] 曾心愉，“不可逆性、无序性与自然辐射理论”，复旦大学物理系《量子物理史讲习班参考资料》

[10]〔美〕M.J.Klein，“普朗克和量子论的开端”，《科学与哲学》，1986 年 6 月

[11] 钱时惕，“普朗克与作用量子”，《重大科学发现个例研究》，北京，科学出版社，1987 年 8 月
[12] 许良英、范岱年编译，《爱因斯坦文集》，第一卷，第 1 版，北京，商务印书馆，1976 年 1 月，100—103，445—446

十一、狭义相对论的建立

爱因斯坦创立的狭义相对论是本世纪物理学最伟大的成就之一。它从根本上改变了传统的时间、空间观念，建立了时间、空间的新观念，揭示了质量和能量的内在联系，给出了高速运动物体的运动规律。这个理论不仅由大量实验所证实，而且已经成为近代科学技术不可缺少的理论基础。

相对论是在研究运动物体的光学和动体电动力学的过程中产生的；是在旧理论出现了严重而深刻的矛盾中产生的；是爱因斯坦在前人工作的基础上，经过 10 年酝酿和探索而完成的。研究相对论的起源及其发展的历史对于我们的学习是有重要意义的，我们可以从中受到非常丰富的科学方法论的教益和启迪。

本文要介绍的内容是：相对论建立前地球运动对光学现象影响的研究；洛伦兹变换的发展与他的“对应态理论”；鼓加勒的“相对性原理”的基本内容；爱因斯坦狭义相对论的思想起源及其发展过程。

(一)地球运动对光学现象的影响

1．光行差研究与菲涅耳理论

1728 年英国天文学家布拉德雷(James Bradley，1693—1762)报告，通过对天龙座γ恒星的长期观察，发现恒星的视位置起了变化，以一年为周期在天球上画出一个小椭圆，这种现象就叫光行差。他认为这是由地球的轨道运动和光的有限速度产生的。他依据光的微粒说，用光的速度和地球轨道运动速度的合成来说明。设来自恒星的光对太阳的速度为 c，地球绕太阳运动的速度为 v，则从地球上观测到的从恒星来的光速应为

$$c'=c-v \quad (11.1.1)$$

布拉德雷计算出光行差角δ为

$$\sin\delta=\frac{v}{c}\sin\theta \qquad (11.1.2)$$

式中θ是 c'和-$\vec{v}$之间的夹角。望远镜筒必须指向-c'方向，这是恒星视位置(ap- parent position)所在的方向。恒星的真实位置沿-c 方向。

从对几个恒星的光行差的观察，布拉德雷得出结论，光速与恒星到地球的距离无关，对各个恒星来说光速是相同的，并测量出光速 $c=3.04\times10^8$m/s。

布拉德雷基于速度合成说明了光行差，他是以光速和地球运动相互独立为前提的，这一前提极其自然地为光的微粒说所接受。在 18 世纪期间，人们专门从微粒说的观点出发处理光行差。在波动说看来，这一前提未必成立，说明光行差是一个困难的问题。托马斯·杨在 1804 年的论文《关于物理光学的实验和计算》中说，很难想象地球运动不影响以太(echer)。另一方面，他又说，如果以太和地球一起运动，那么就不会产生光行差。

对于由布拉德雷理论得出的一些结果，人们议论纷纷。其一是当来自恒星的光通过折射率大于 1 的介质时，由于该介质中的光速和空气中的不同，理应显示出不同的光行差。另一个问题是由恒星射来的光和地上的光应该以不同的方式进行折射。1810 年，阿拉果进行了一项实验，他用一块棱镜遮住望远镜的一半，他发现光行差角与通过棱镜的任何一条光线无关，表明来自恒星的光线不受地球运动的影响。

1818 年，菲涅耳在解释阿拉果星光折射行为时，提出透明物质中以太密度和该物质折射率 n 的平方成正比的见解。他假定当一个物体相对于以太参考系运动时，其内部的以太只是超过真空中以太的密度的那部分，即$\left(1-\frac{1}{n^2}\right)$的以太才被物体所带动，这就是“部分曳引假设”。菲涅耳由此得出在以太参考系中，以速度 v 运动的透明物体中以太的速度为

$$\left(1-\frac{1}{n^2}\right)v \qquad (11.1.3)$$

所以，在以太参考系中运动物体内的光速为

$$c'=\frac{c}{n}+Kv \qquad (11.1.4)$$

式中$K=1-\frac{1}{n^2}$称为菲涅耳的曳引系数。曳引系数与折射率n有关，对于折射率 n=1 的物质，应不拖动以太。进而言之，既然大气的折射率接近于 1，它就不能或者只能极微弱地拖动以太。这样，对于地球上的观察系统来说，将有以太风存在。假若如此，光行差就可以说明。

可是，菲涅耳的理论并未马上被人们所承认。1843 年，多普勒发表了《光行差说明的总论》。他说菲涅耳理论建立在一切物体对以太都是透明的假设之上，与地球和其它地上物体的不透光性相抵触，因而是不能允许的。1845 年，英国人斯托克斯(G.Stokes，1819—1903)提出了他的光行差理论，他强烈地反对地球通过以太而毫不影响以太的观念。他的理论以两个假设为基础，其一是，在地球表面上以太和地球表面具有相同的速度，随着与地球距离的增加而减少到零。其二是，以太的速度是无旋的，即具有速度势。根据这两个假设，恰好导出观测到的光行差。

1851 年，法国人菲佐(A.H.L.Fizeau，1819—1896)设计了一个实验对菲涅耳的曳引系数获得了第一个实验证明。[1]如图 11-2 所示，让一束来自静止光源 S 的单色光被玻璃片 G 反射,然后经透镜 L_1 成为平行光。缝 S_1、S_2 选择了两条光线通过 U 形玻璃管，管中的水在如图所示的方向上流动。位于透镜 L_2 焦点

图 11-2

的平面镜 M 改变两条光线的方向，使得其中的一条光线总是沿水流方向传播，而另一条光线沿反方向传播。在通过 U 形管后，两条光线经透镜 L_1 在其焦点 S'点会聚并产生干涉条纹，当从空管到充满静止的水时，菲佐没有测量出任何条纹移动。当水以速率 v 相对于实验室运动时，他考虑了三种可能性：(1)以太完全被水拖动，即 K=1；(2)以太完全没有被水拖动，即$K=0$；(3)按照菲涅耳部分以太被拖动的假设，即$K=1-\frac{1}{n^2}$。

菲佐根据方程(11.1.4)，计算了通过充满流水的 U 形管后，两束光的光程差为

$$\Delta=2Lc\left[\frac{1}{c/n-Kv}-\frac{1}{c/n+Kv}\right] \qquad (11.1.5)$$

式中 L 是 U 形管每个臂的长度，精确到 v/c，预期的条纹移动量δ=△/λ

为

$$\delta_{计算}=\frac{4n^2Lv}{c\lambda}K \qquad (11.1.6)$$

菲佐使用波长λ=5.26×10^{-7}m的黄色光，L=1.487m，n= 1.33，v=7.059m/s。他测量了一个平均的条纹移动值δ$_{观察}$=0.23。

用$K=1-\frac{1}{n^2}$代入，则预期的条纹移动值δ$_{计算}=0.2022$.由此菲佐得出结论“这两个值几乎是相等的”。这就从实验上证明了菲涅耳的理论。

图 11-3

1871年，艾里(G.B.Airy)进行了一项类似于1818年阿拉果所作的实验。[1]如图11-3所示，在地心参考系中，考虑一个充满水的望远镜，它对准恒星的视线垂直于恒星相对于地球的速度方向。按照菲涅耳假设，水中以太的漂移速度为$-v/n^2$，由正弦定律得$\sin\delta'=\frac{v}{cn}$。因为在进入水后星光被折射，所以δ'不是光行差角。据折射定律sinδ=nsinδ'，我们得到$\sin\delta=\frac{v}{c}$。艾里用这一实验装置观察了天龙座的γ星。他做了两次观察，一次在春天，另一次在六个月之后的秋天。当时，他从观察结果计算出观察地点的纬度。如果望远镜筒内的水改变了光行差常数，那么计算得到的纬度会出现矛盾，然而未能找到任何矛盾。

2．以太漂移一级效应零结果与静止以太说

菲佐实验大大提高了菲涅耳的威信，按照该理论以太在宇宙空间是静止的，地球相对于它运动着。对于地球上的观察者来说就有以太风存在，从而人们开始了探索以太存在的实验。1868年奥克(Martinus Hoek)进行了一项实验。[1]如图11-4所示，来自S光源的单色光被半镀银镜P分成在相反方向上通过闭合路径的两束光，M_1、M_2和M_3是平面镜，-V是以太相对于实验室的速度。在M_1、M_2路径上包含了一段充满水的玻璃管W。在通过闭合路径后，被P片所反射的射线1与被P片所折射的射线2会聚到一起，在望远镜T内产生干涉条纹。把仪器转动180°，使射线2在与地球运动相反的方向上通过以太，奥克发现没有干涉条纹移动。

奥克按照菲涅耳的理论解释了这一结果。在这一装置中，射线1和射线2之间的光程差只能由闭合路径中长为l的那部分引起，据方程(11.1.4)，对于先通过长为l的水段，后通过长为l的空气段的射线1所需的时间t_1为

$$t_1=\frac{l}{c+v}+\frac{l}{c/n+Kv-v} \qquad (11.1.7)$$

对于在相反的方向上通过闭合路径的射线2所需的时间t_2为

$$t_2=\frac{L}{c-v}+\frac{L}{c/n-Kv+v} \qquad (11.1.8)$$

式中曳引系数$K = 1 - \frac{1}{n^2}$代入式(11.1.7)式(11.1.8)，精确到$\frac{v}{c}$阶得

$t_1 = t_2 = \frac{l}{c} + \frac{nl}{c}$，$\triangle t = 0$。观察不到干涉条纹，这就解释了实验结果。

图 11-4

1870 年，维耳特曼(Wilhelm Veltmann)证明，如果不涉及高于一阶的效应，人们便不能检测到天体运动对光学现象的任何影响。[2]他考虑了在一个具有公共平动速度 v 的透明物体中光的传播问题。如果我们用 s_i 表示多边形每边的长度，用 w_i 表示光在该边相对于物体的速度，那么光线走完全程所需要的时间将是 $\Sigma s_i/w_i$，若以太相对于物体的速度是 U_i，则

$$w_i=c_i+U_i\cos\varphi_i \qquad (1.1.9)$$

在这里，φ_i是i边和系统平动方向之间的夹角。依照菲涅耳假设$U_i = \upsilon / n_i^2$ n_i 是物体的折射率，略去高于一阶的项，我们得到

$$\frac{s_i}{w_i} \approx \frac{s_i}{c_i} - \frac{s_i}{c_i^2}\frac{v}{n_i^2}\cos\varphi_i = \frac{s_i}{c_i} - \frac{v}{c^2}s_i\cos\varphi_i$$

$$\sum\frac{s_i}{w_i} = \sum\frac{s_i}{c_i} - \frac{v}{c^2}\sum s_i\cos\varphi_i \qquad (11.1.10)$$

这里 c 表示光在真空中的速度。后一个方程右边第一项表示光线经过全程的时间间隔，第二项的求和是全体路程在系统平动方向的投影，对于一个闭合光路它变为零。因此，对于把相同起点和终点连接起来的两束光来说，两条光路长度之差依然是相同的，而与系统运动无关。现在，按照菲涅耳理论，光的传播一般能描述为干涉的结果，并且干涉唯一地由光路长度决定。因此，在一级近似下，光学现象不受地球运动的影响，干涉仪实验不能察觉到地球相对于以太的运动，以太漂移的零结果是不可避免的结果。

以太漂移一阶效应的零结果，在菲涅耳理论的基础上得到了充分的证明和解释。而该理论主张静止以太的存在，地球相对于以太运动。所以这一结果并不象有的文章所说是对静止以太论的否定。从历史实际来看一阶效应实验是个肯定的实验，它的零结果巩固并加强了人们的静止以太观念。

3．以太漂移二级效应零结果与收缩假设

麦克斯韦在 1867 年指出，在地上做测量光速的实验，因为让光在同一路径上往返，地球相对于以太运动的影响仅仅表现在二次效应上，要用实验测出来是极其困难的。迈克耳孙(A.A.Michelson，1852—1931)力图克服这个困难，他想检测出地球相对以太的运动，进一步确证菲涅耳的静止以太说。

图 11—5

1881 年，他用自己发明的干涉仪作了这个实验。以后又在 1887 年与莫雷(E.W.Moley，1838—1923)合作进行了更精确的测量。这一装置如图 11-5 所示。[1]以光束 1 从 M 到 M_1 的往返时间为

$$t_1 = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2L}{c(1-v^2/c^2)} \qquad (11.1.11)$$

光束 2 从 M 到 M_2 的往返路径如图所示，所需时间为 t_2。因为

$$l^2 = \left(\frac{ct_2}{2}\right)^2 - \left(\frac{vt_2}{2}\right)^2$$

所以 $$t_2 = \frac{2L}{c\sqrt{1-v^2/c^2}} \qquad (11.1.12)$$

仪器转动 90°所引起的两光束的时间差的变化为

$$\Delta t = 2(t_1 - t_2) = \frac{4L}{c}\left(\frac{1}{1-v^2/c^2} - \frac{1}{\sqrt{1-v^2/c^2}}\right)$$

$$\approx \frac{2Lv^2}{c^3} \qquad (11.1.13)$$

预期的干涉条纹的移动量为

$$\delta = \frac{c\Delta t}{\lambda} = \frac{2Lv^2}{\lambda c^2} \qquad (11.1.14)$$

将以上数据代入，估计可以测得 δ=0.4 个条纹移动。但是，迈克耳孙和莫雷进行观察时却看不到这样的移动。在这个实验中，他们把仪器安装在很重的石台上以维持稳定，并把这一装置浮在水银上使它可绕水平轴平稳地转动，使他们的仪器能观察到 0.01 个条纹的移动。所以他们得出结论：“实际观测到的条纹移动肯定小于预期值的 $\frac{1}{20}$，或许小于 $\frac{1}{40}$。”

迈克耳孙对这一实验结果感到十分失望，但并没有动摇他对以太说的信赖。他用拖动假设来解释，认为地球运动时，地球表面的以太并不保持静止，而是随地球一起运动。因为这一解释与光行差现象有矛盾，所以不为世人所接受。由于这一实验与静止以太论不相容，所以引起了物理学界的震惊。但是有的书把这一实验说成是相对论赖以产生的判决实验就不妥了。从当时的历史实际来说，人们对菲涅耳的理论是极为信赖的，不少人认为迈克耳孙-莫雷实验失败了，其理由是看不到预想的结果。总之，谁也没有据此判断菲涅耳的静止以太被否定了。

为了解释迈克耳孙-莫雷实验的零结果，1889 年，爱尔兰的物理学家菲兹杰惹(G.Fitzgerald，1851—1901)提出了收缩假设。[3]他指出，如果物质是由带电荷的粒子组成，一根相对于以太静止的量杆的长度，将完全由量杆粒子间取得的静电平衡决定，而量杆相对于以太运动时，组成量杆的带电粒子将会产生磁场，从而改变这些粒子之间的平衡间隔，量杆就会缩短，缩短的程度“取决于物体的运动速度对光速的比率的平方”，即 $\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}$。这样，迈克耳孙-莫雷实验所使用的仪器，当它指向地球运动的方向时就会缩短，而缩短的程度恰好使我们无法用迈克耳孙实验探测到以太漂移。

1892 年，荷兰物理学家洛伦兹(H.A.Lorentz，1853—1928)独立地提

出了他的收缩假设。[4]他说：这个实验(迈克耳孙-莫雷实验)已经使我困惑了一个很长的时间，最后我想到了一种使它的结果和菲涅耳理论

相协调的一种形式.”他依据迈克耳孙实验中两光束的时间差 $\frac{Lv^2}{c^3}$ 提

出处于地球运动方向的那条臂缩短的假设.他说如果这条臂缩短到L

$\left(1-\frac{v^2}{2c^2}\right)$,那么,精确到$\left(\frac{v}{c}\right)^2$,则$\triangle t=0$,迈克耳孙实验的结果就可以圆满

地解释了.

他认为这一收缩是由于分子力引起的，并假定分子力也象电力和磁力那样通过以太而传递，由于物体平移可能影响分子力，从而使物体大小发生变化。在后来写的电动力学论文中，洛伦兹依据他所导出的力的变换式，证明了他的收缩假设。

(二)洛伦兹变换与“对应态理论”

1890 年，赫兹把麦克斯韦电磁场方程写成了四个简单的具有对称形式的微分方程，导出了在静止以太参考系中的波动方程，得出波速不依赖于波源运动的结论，并且在原则上作出了在相互作匀速运动的参考系中，电磁场方程在形式上不变性的论断。但是，他没有进一步探讨在两个参考系中的电磁场量之间的变换方程以及电磁场方程协变性的问题。

本节以洛伦兹关于动体电动力学的三篇论文为依据，试图探讨他的变换理论的发展以及对应态理论的形成。

1．1892 年洛伦兹的电磁理论

1892 年，洛伦兹发表了《麦克斯韦的电磁理论及其对运动物体的应用》的论文，着手研究运动物体的电动力学问题。在赫兹发现电磁波后，洛伦兹明确地接受了麦克斯韦关于场的概念。他首先把以太和有重量的物质鲜明地区别开来，认为以太是电磁场的载体。在他的理论中以太和物质粒子之间仅仅存在下述关系：物质粒子所带的电荷使以太的电磁状态发生变化，以太的电磁状态使带电粒子受到力的作用。就洛伦兹的理论而言，电磁场被认为是由以太承担着，以太和物质之间只有电磁相互作用。这与麦克斯韦和赫兹所说的电磁场被介质所承担的意义截然不同，麦克斯韦的电磁理论是通过具有力学特征的介质状态的变化来理解电磁作用，他认为以太及介质中的力学应变产生电磁现象。洛伦兹认为承担电磁场的以太和通常的物质完全是独立的，他所说的以太就是作为独立实体的电磁场而存在。[2]

洛伦兹在这篇文章中写道，他的目的是为了解决光在运动介质中的传播问题，推导出菲涅耳的曳引系数 K。在这项研究中，他提出了以下基本假设：①静止以太充满整个宇宙空间；②物体中大量微小的带电粒子是电磁现象的源泉；③空间每一点以太的状态是由麦克斯韦方程

$$\nabla \times E - \frac{1}{c}\frac{\partial B}{\partial t} \qquad (11.2.1)$$

$$\nabla \times B = \frac{1}{c}\frac{\partial E}{\partial t} + \frac{4\pi}{c}\rho W \qquad (11.2.2)$$

$$\nabla \cdot E = 4\pi\rho \qquad (11.2.3)$$

$$\nabla \cdot B = 0 \qquad (11.2.4)$$

来描述的。[1]上式中的 E 和 B 是电场强度和磁感应强度。ρ是电荷密度，W 是带电粒子相对于以太的速度。洛伦兹得以太施加在带电粒子上的力密度为

$$f = \rho E + \rho \frac{W}{c} \times B \qquad (11.2.5)$$

他认为以上五个方程是基本方程，这些方程是相对于固定在以太中的参考系 S 写的。

在这些方程的基础上，洛伦兹在这篇论文的最后一章《光在有重量的运动介质中的传播》中，讨论了在惯性参考系中运动物体的光学过程。[1]他处理运动物体光学问题的第一步是把带有波源的波动方程，从静止以太参考系变换到相对于静止以太参考系以速度 v=vi 运动的惯性系 S_r 中。在 S 系中，这些有波源的波动方程是

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)E = 4\pi\nabla\rho + \frac{4\pi}{c^2}\frac{\partial}{\partial t}(\rho W) \qquad (11.2.6)$$

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)B = \frac{4\pi}{c}\nabla \times (\rho W) \qquad (11.2.7)$$

式中 W=V+u，u 是波源相对于 S_r 的速度。为了简单起见，方程(11．2．6)和(11．2．7)可以写为

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t}\right)\alpha_1 = \alpha_2 \qquad (11.2.8)$$

波动方程(11．2．6)和(11．2．7)描述了通过以太以速度 c 传播的电磁波，而波速是与波源的运动无关的。

假设介质中的每个分子以惯性参考系 S_r 的速度(即 W_x=vi，W_y=W_z=0)运动，洛伦兹根据伽利略变换式

$x_r x-v_t$，y_e=y， z_r=z，t_r=t(11．2．9)

和运流导数(convective derivative)，即矢量场对时间的导数在两个惯性系中的关系式

$$\left(\frac{\partial}{\partial t}\right)Sr = \left(\frac{\partial}{\partial t}\right)S + v \cdot \nabla_r \qquad (11.2.10)$$

把方程(11．2．8)从 S 系变换到 S_r 系，因此在 S_r 系中的波动方程就变为

$$\left[\nabla_r^2 - \frac{1}{c^2}\left(\frac{\partial}{\partial t_r} - v\frac{\partial}{\partial x_r}\right)^2\right]K = j \qquad (11.2.11)$$

在这里K和j是（x_r，y_r，z_r，$t_r = t$）的函数，而$\nabla = \nabla_r$。但是由于方程(11．2．11)没有描述波动运动的通常的形式。洛伦兹提出了一个从 S_r 系

到 Q'系的附加的坐标变换式

$x' = \gamma x_r \qquad y' = y_r \qquad z' = z_r$

$t' = t - (v/c^2)\gamma^2 x_r \qquad (11.2.12)$

式中$\gamma = (1/\sqrt{1-v^2/c^2})$。洛伦兹认为从$S_r$系到$Q'$系的变换是纯粹的数学坐标变换，他引入的 x'和 t'是新的独立变量。在 Q'系中，方程(11.2.11)变为

$$\left[\nabla'^2 - \frac{1}{(c^2 - v^2)}\frac{\partial^2}{\partial t^2}\right]K' = j' \qquad (11.2.13)$$

在这里 K'和 J'是(x'，y'，z'，t')的函数。方程(11.2.13)描述了一个以速度$c\sqrt{1-v^2/c^2}$传播的电磁波。因此从S系到S_r的伽利略变换未能产生正确的波动方程,而从 S_r 系到 Q'系的进一步变换产生了取决于波源速度的波动方程，从而违背了建立在以太基础上的光的波动理论。但是精确到 v/c 阶，变换方程(11．2．12)式就变为

$$x' = x_r, y' = y_r, z' = z_r, t' = t - (v/c^2)x_r \qquad (11.2.14)$$

我们把在 v/c 限度内的 Q'系称为 R'系。在 R'系中波动方程就与 S 系中的波动方程具有了相同的形式，在这个方程中波速是 c。因此，到 v/c 阶，数学坐标系统 Q'变成了它的空间坐标和伽利略空间坐标相同，而时间坐标是伽利略的绝对时间 t_r(=t)和伽利略的空间坐标 x_r 的混合。

最后，洛伦兹从 S_r 系中的波动方程推导出了菲涅耳的曳引系数，他的推导取决于光在其中传播的物质中作谐振动的带电粒子所受的恢复力和每单位体积中的带电粒子数。他认为菲涅耳曳引系数的起因是由于光和被束缚的、谐振的带电粒子相互作用的结果。菲涅耳曳引系数的成功导出是洛伦兹电子论的巨大成果。

2．1895 年洛伦兹的—阶交换理论

1895 年，出版了洛伦兹的一本著作《运动物体中电现象和光现象的理论研究》。在这一研究中，洛伦兹系统地、简洁地推导出了 1892 年论文中的全部结果，而且，继续使它们普遍化。在“应用到静电”这章中，他提出了把电动力学问题变成静电问题的步骤。[1]在构成大块物质的离子是静止的 S_r 系中，它们的电磁性质必定不变，电磁场量对时间的导数一定为零。方程(11．2．10)变为

$$\left(\frac{\partial}{\partial t}\right)_s = -v\cdot\nabla_r \qquad (11.2.15)$$

洛伦兹通过利用方程(11.2.15)和伽利略变换式(11.2.9)把方程(11.2.6)和(11．2．7)变换到 S_r系

$$\left[\nabla_r^2 - \left(\frac{v\cdot\nabla_r}{c}\right)^2\right]E = 4\pi\nabla_r\rho - \frac{4\pi}{c^2}v\cdot\nabla r(\rho v) \qquad (11.2.16)$$

$$\left[\nabla_r^2 - \left(\frac{v}{c}\cdot\nabla_r\right)^2\right]B = \frac{4\pi}{c}\nabla_r\times(\rho v) \qquad (11.2.17)$$

在这里，洛伦兹计算的是 S 系中的场量 E 和 B，尽管所有其它的量都是在惯性参考系 S_r 中表示的。其次为了便于求解方程(11.2.16)和(11.2.17)，

洛伦兹通过下式引入标势φ

$$\left[\nabla_r^2-\left(\frac{v}{c}\cdot\nabla r\right)^2\right]\phi=-4\pi\rho \qquad (11.2.18)$$

在此

$$E=-\nabla r\phi+\frac{v}{c}\left(\frac{v}{c}\cdot\nabla_r\phi\right) \qquad (11.2.19)$$

$$B=\frac{v}{c}\times E \qquad (11.2.20)$$

取 v=vi 洛伦兹可以把方程(11．2．18)重新写为

$$\left[\left(1-\frac{v^2}{c^2}\right)\frac{\partial^2}{\partial x_r^2}+\frac{\partial^2}{\partial y_r{}^2}+\frac{\partial^2}{\partial z_r{}^2}\right]\phi=-4\pi\rho \qquad (11.2.21)$$

洛伦兹注意到方程(11．2．21)几乎是一个泊松方程，他接着引入一个从 S_r 系到 Q'系的变换方程

$$x_r=x''\sqrt{1-v^2/c^2},y_r=y'',z_r=z'',t_r=t'' \qquad (11.2.22)$$

把方程(11．2．22)、(11．2．21)变为泊松方程

$$\nabla''^2\phi''=-4\pi\rho'' \qquad (11.2.23)$$

在此，φ"和ρ"是 x"、y"、z"的函数，从 S"系中电荷守恒的要求，洛伦兹推导出

$$\rho''=\rho\sqrt{1-v^2/c^2} \qquad (11.2.24)$$

因此

$$\phi''=\phi\sqrt{1-v^2/c^2} \qquad (11.2.25)$$

洛伦兹关于在 S"系中离子的场特征能够从泊松方程获得的证明，完成了电动力学问题到静电问题的推导。对洛伦兹来说 S"系是纯粹的数学坐标系统，变换过程是一种纯数学过程。

因为在 S"系中

$$E''=-\nabla\text{“}\phi\text{”} \qquad (11.2.26)$$

洛伦兹从方程(11．2．19)、(11．2．20)、(11．2．22)和(11．2．25)导出了在 S''系和 S"系中力的变换式(因为 S 与 S’’系是由伽利略变换联系起来的，所以 S 与 S’’系中的力是相等的)

$$F_{xr}=F_x'',F_{yr}=F_y''\sqrt{1-v^2/c^2},$$

$$F_{2r}=F_2''\sqrt{1-v^2/c^2} \qquad (11.2.27)$$

(11．2．27)式描述了离子在静止以太参考系 S"中受的力 F"和在运动参考系 S_r 中受的力 F_r 的关系。由此可见，只要离子系统在 S"系中处在平衡状态，则在 S_r 系中也是处于平衡状态。上述方程也表明静电现象，即用静止在 S_r 系中的仪器所进行的实验，只有到二阶 v/c 才受地球运动的影响。

在这本书中，为了解决运动物体的光学问题，洛伦兹在静止以太参考系 S 和运动参考系 S_r 之间引入一套新的变换方程

$$x_r=x-vt,y_r=y,z_r=z,$$

$$t_L=t-\frac{v}{c^2}x \qquad (11.2.28)$$

他把引入的“新的独立变量” t_L 称为“当地时间”(Local time)坐标，

以便于与他的“普遍时间”(Universal time)t 相区别。对 t_L 的这种解释是与他毕生的信念一致的。他认为真实的、普遍的空间和时间坐标是相对于静止以太参考系的坐标。在忽略二阶项的情形下，根据式(11．2．28)他得到了在 S_r 系中波速 c 不依赖于波源运动的波动方程

$$\left(\nabla_r^2-\frac{1}{c^2}\frac{\partial^2}{\partial t_L^2}\right)K''=j'' \qquad (11.2.29)$$

式中 K"和 j"是(X_r，y_r，z_r，t_r)的函数，在自由以太空间该波动方程与 S 系中的波动方程具有相同的形式。他假设在 S_r 和 S 系中的麦克斯韦方程具有相同的形式，在改变了的伽利略变换下这些方程是协变的(Covariant)

$$\nabla_r\times E_r=-\frac{1}{c}\frac{\partial B_r}{\partial t},\quad \nabla_r\cdot E_r=0$$

$$\nabla_r\times B_r=\frac{1}{c}\frac{\partial E_r}{\partial t}\qquad \nabla_r\cdot B_r=0 \qquad (11.2.30)$$

因为精确到 $\frac{v}{c}$ 阶，在 S_r 系和 S 系中的麦克斯韦方程具有相同的形式，所以在 S_r 系和 S 系中的光学规律是相同的。他把这一结果称之为“对应态原理”(theorem of corresponding states)。如果在 S 系中存在一个以 E 和 B 为特征的系统的状态，E 和 B 是(x、y、z、t)的函数，那么在 S 系中就存在一个以 E_r 和 B_r 为特征的对应状态，E_r 和 B_r 是(x_r，y_r，z_r，t_L)的函数。该函数与上述函数具有相同的形式。洛伦兹认为对应态理论普遍地回答了，在精确到 $\frac{v}{c}$ 阶的条件下，地球运动对光学现象没有影响。

3．1904 年洛伦兹的二阶变换理论

1904 年 5 月洛伦兹发表了他的二阶理论论文《在一个速度小于光速的运动系统中的电磁现象》。[4]这篇文章是在 1900 年彭加勒对“收缩假设”提出批评后写的。洛伦兹接受了彭加勒的批评。认为“对每一个新的实验结果提出一种特殊假设的作法是不自然的，假使能利用某些基本假定，并且不用忽略各阶 $\frac{v}{c}$ 的量来证明任何电磁作用都与各惯性系的运动速度无关就更好了”。洛伦兹试图用一些“基本假设”来建立新的电动力学理论。

但是，洛伦兹这篇文章并没有实现这一目标。他声称要以基本假设而不是以“特殊假设”为基础。事实上却包含了一系列特殊假设；限于小的速度 v 对光速 c 的比值；先验地假设变换方程；假设有静止以太；假设静止电子是球形的；假设所有力的变换家电磁力一样；……正是由于这一情况，所以后来爱因斯坦认为这是“作为一个对于过去一直是相互独立的种种假设的令人吃惊地简单的总结和概括”。

象他在 1892 年所作的那样,洛伦兹首先用运流导数和伽利略变换把电磁场方程从 S 系变到 S_r 系，然后引入一组新的独立坐标把 S_r 系变到由 X'，y'，z'，和 t'表示的 S'系

$$x'=\gamma Lx_r, y'=Ly_r, z'=Lz_r \tag{11.2.33}$$
$$t'=\frac{L}{\gamma}t-\gamma L\frac{v}{c^2}xr$$

式中$\gamma=1/\sqrt{1-v^2/c^2}$, v是运动参考系的速度.这就是洛伦兹变换式的原始形式。洛伦兹认为当$\frac{v}{x}$的比值不大时, $L=1$.他仍然把t'称为“地方时”。

1906年，彭加勒在《电子动力学》中，消除了S_r系，直接把S'系和静止以太参考系联系起来，写出了他称为的“洛伦兹变换”(Lorentz transformation)式

$$x'=\gamma L(x-vt), y'=Ly, z'=Lz$$
$$t'=\gamma L\left(t-\frac{v}{c^2}x\right) \tag{11.2.34}$$

当L=1时，就成为了现代使用的洛伦兹变换式。洛伦兹用下列各式定义在S'系中的电位移D'和磁场强度H'

$$D_x'=\frac{1}{L^2}Dx, D_y'=\frac{\gamma}{L^2}\left(D_y-\frac{v}{c}H_z\right)$$
$$D_z'=\frac{\gamma}{L^2}\left(D_2+\frac{v}{c}H_y\right)$$
$$H_x'=\frac{1}{L^2}H_x, H_y'=\frac{\gamma}{L^2}\left(H_y+\frac{v}{c}D_2\right)$$
$$H_z'=\frac{\gamma}{L^2}\left(H_z-\frac{v}{c}D_y\right) \tag{11.2.35}$$

这样一来，如果利用上面定义的D'，H'，x'，y'，z'，t'改写运动系中(例如，在地球上)的电磁现象的表达式(以D，H，x， y，z书写)，那么就可以得到和静止系中同一对象的麦克斯韦方程式相同的形式。因此，运动系中的电磁现象也可以通过解静止系中的麦克斯韦方程来处理，只不过把不带撇的变量变为相应的带撇的变量而已。所以前面的对应态原理，不是近似的，而是严格的。因此，这就可以保证不论多么高次的效应，也决不可能发现相对于以太的运动。

洛伦兹把他的收缩假设和新的坐标变换式结合起来说明两个系统中对应状态的存在。他假定：“电子，当它们处于静止状态时我认为是半径为R的球状的，但由于平移的影响，它们的大小就发生了变化，沿着运动方向的长度变小到原来长度的1/γL”“这就是说，在一个以速度v运动的静电系统中，所有电子都是扁平椭球，其短轴沿运动方向。按新的变换式，两个系统的关系是：如果平行X轴的尺寸乘上γL，Y方向或Z方向的尺寸乘上L，S系就变到S_r系。这种变形可以用符号(γL， L，L，)表示。这种变形恰好抵消了或补偿了长度收缩引起的效应，所以重新得到半径为 R的球形电子。

洛伦兹还依据电磁场的变换式导出了力的变换式。对于静电场，他得出了在运动参考系S'和以太静止参考系S中的电力的变换式

$$F_x' = L^2 F_x, F_y' = \frac{L^2}{\gamma} F_y, F_x' = \frac{L^2}{\gamma} F_z \qquad (11.2.36)$$

或写为 $$F(S') = \left(L^2, \frac{L_2}{\gamma}, \frac{L_2}{\gamma}\right) F(S) \qquad (11.2.37)$$

因此，如果在 S 系中一个电子所受的电力为零，则在 S'系中该粒子所受的电力也为零。他又假定分子之间的力具有和电力一样的性质。“不带电粒子之间的力，以及这种粒子和电子之间的力，当系统平移时所受的影响，和静电系统中电力所受的影响完全一样”。由于物体的形状由分子力的平衡来决定，因此，静止时具有某一形状大小的物体，平移时在其运动方向上的长度就会按 $\sqrt{1-v^2/c^2}$ 收缩，这就证明了他的收缩假设。

虽然，洛伦兹借助于对应态原理分析问题，但是他的主导思想仍然是：运动产生了效应，效应之所以不能检测出来，乃是由于它们相互抵消了。洛伦兹本人的话表明了这一点。1909 年，他在《电子论》一本中写道：他的理论和爱因斯坦理论的“主要差别”在于后者使我们在诸如迈克耳孙、瑞利和布雷斯实验的否定结果中看到的不是相反效应的偶然补偿，而是一般的基本原理的体现。

虽然洛伦兹提出了变换方程，但是他是先验地提出这组方程，而且他把变换方程作为一种纯粹的数学手段，把变换式中的时间 t'称为“地方时间”。1915 年，他在总结自己没有提出相对论的原因时说：“如果我现在必须写最后一章，我一定给爱因斯坦的狭义相对论一个更为重要的地位。根据这一理论，运动系统电磁现 237 象的理论达到了我没有获得的简明性。我失败的主要原因是我墨守这一概念，只是把时间 t 认为是真实的时间，而把我的地方时间 t'充其量不过看作是一个数学辅助量而已”[5]

爱因斯坦对洛伦兹的贡献作了高度评价。1907 年，他说：“狭义相对论是洛伦兹理论和相对性原理的结合。” 1955 年，他在致 C.Seelig 的一封信中写道：“毫无疑问，如果我们回顾狭义相对论的发展，那么它在 1905 年发现的时机已经成熟。洛伦兹已经看到为了分析麦克斯韦方程，那些后来以他的名字命名的变换方程是必不可少的，彭加勒甚至已经识破这些联系。”

(三)彭加勒的“相对性原理”

法国数学家物理学家彭加勒(Henri Poincare，1854—1912)1895 年在研究拉摩(J.Larmor，1857—1942)电磁理论的论文中，首次出现了反对绝对运动的提法，他在那里说：“从各种经验事实得出的结论能够概括为下述断言：要证明物质的绝对运动，或者更确切地讲要证明可称量物质相对于以太运动是不可能的。”[2]1899 年，在巴黎大学的讲演中，当他描述了当时所作的一系列测量地球相对于以太运动的一阶、二阶实验没有产生任何效应后，他说：“我认为光学现象很可能只取决于物体、光源和所涉及的仪器的相对运动，不只是到二阶的量，而且到更加精确的量也是对的。”[6]这就是说，在 1899 年彭加勒相信在原则上绝对运动是不存在的，只有相对运动才有意义。

1900 年，在巴黎举行的物理学国际会议上，他坚持了同样的原则。他

说："我们的以太它真的存在吗？我不相信更精确的观察能够揭示出比相对位移更多的东西。"他继续说："对于考虑到二阶项所获得的负结果，寻找一个共同的解释是必要的。有理由假定这种共同的解释适用于更高阶的项，这些项的相消是严格的和绝对的，于是某种新的原理必定引入物理学。这一新原理很像热力学第二定律，它断言某些事情是不可能的，在这里，就是确定地球相对于以太的速度是不可能的。"彭加勒在这里明确指出对各种二阶效应实验要寻找一个共同的解释，他认为对每一个新的实验结果提出一种特殊假设的做法是不自然的。他预期着引入一个新的原理，这一新原理在方法论上类似于热力学第二定律。即确定某些事物是不可能存在的。

迄今，彭加勒论述的只是相对运动原理而不是相对性原理。1902 年，他在《科学与假设》(Science and Hypothesis)中进一步阐述了这一思想。在该书第五章"实验和几何学"中，他断言："在任何时刻物体的状态只取决于初始时物体的状态和物体间的相对位置，而与系统在初始时的绝对位置和绝对速度无关。简而言之，我称它为'相对性定律'(law of relativity)。"[7]他说："例如宇宙系统，实验势必不能告诉我们它在空间有什么绝对的位置和方向，无论我们的仪器何等精巧，我们所能知道的，只是宇宙的各部分的情状及其相对位置。"[7]所以，彭加勒的相对性定律仍然局限在对绝对运动的否定上。

1904 年 9 月 24 日，在美国圣路易斯(St.Louis)城举行的科学和艺术会议上，彭加勒在一篇讲演中首次使用了"相对性原理"(the principle of relativity)一词。他说："按照相对性原理，物理现象的规律对一个固定的观察者象对于一个相对于他作匀速平移运动的观察者一样是相同的。所以我们没有也不可能有任何方法来辨别我们是不是处于这样一个匀速运动系统中。"[7]在这篇讲话中，彭加勒还敏锐地预感到一种新的力学即将出现。他断言："也许我们应该建立一门新的力学，对这门力学我们只能窥见它的一鳞半爪，在这门力学中，惯性随着速度增加，光速将会成为一个不可逾越的界限。"[8]这段话既显示了彭加勒直觉的力量，也显示出他对新力学的见解具有定性的性质。

1906 年，彭加勒在《论电子的动力学》中写道："从实验上证明地球的绝对运动的不可能性是自然界的普遍的规律。相对性假设(postulate of relativity)是成立的而且是普遍有效的。无论这个假设迄今是与实验相一致，还是以后被更精确的实验进一步证实或被推翻，在任何情况下看到这一假设所得出的结果都是有趣的。"[1]

1908 年在另一篇文章《电子动力学》中，他仍然表述的是相对运动原理。"无论使用什么方法，除了相对速度以外，我们将永远不能揭示出任何其他东西；我所说的某些物体的速度是相对于另外一些物体而言的。"[2]1909 年，在法国科学促进协会的一次讲演中，他对洛伦兹的补偿理论的精神实质仍是满怀信心的。在这次讲演中，彭加勒详细地阐明了，以当地时间和收缩假设为基础的洛伦兹理论解释了"在所有光学实验中所观察到的完全补偿"，并且"在电学中也存在着同样的补偿作用"，他"最终坚信，相对性原理是完全正确的"。[2]

由此看来彭加勒的相对性原理虽然类似于爱因斯坦的叙述，但是在内容上的差别是很明显的，关键的差别是彭加勒原理承认静止以太的存在，

认为只有在静止以太中测量的光速才严格地为 c。他的相对性原理的内容主要是断言绝对运动不存在，实质上讲的是相对运动原理。

惠特克(Edmund Whittaker)在他的著作《以太和电磁理论的历史》中，基本上否定了爱因斯坦 1905 年关于狭义相对论的论文的意义，认为这篇论文只是“陈述了彭加勒和洛伦兹的相对论，并加了一些补充，而且引起了很大注意。”[8]他在 1955 年为爱因斯坦写的悼念传记中，也没有改变这一评价。他说：“彭加勒 1904 年 9 月在圣路易斯的讲话中创造了‘相对性原理’这个词组。”“爱因斯坦采用彭加勒的相对性原理作为物理学的一个新基础，并且指出洛伦兹变换群对于彼此在相对运动的物体的物理学提供了一种新的分析。”[8]前面我们对彭加勒的“相对性原理”已经作了详实的介绍，对于惠特克这一不公正的评价也就不攻自破了，何况爱因斯坦并未读过 1904 年洛伦兹的论文，也不知道彭加勒的这个讲话。

爱因斯坦一直坚持工作的连续性这一观点，他曾经说过：“至于相对论，它根本不是一个革命行动的问题，而是一条可以追溯到很多世纪的路线的一种自然发展的问题。”[8]他对他的前辈和同时代人的工作都是充分肯定的。英费尔德(L.Infeld)在他写的《相对论的发展史》中叙述了他与爱因斯坦在普林斯顿的一次谈话：“我对爱因斯坦说：‘在我看来，即使您没有建立它，狭义相对论的出现也不会再等多久。因为彭加勒已经很接近构成狭义相对论的那些东西了。’爱因斯坦回答道：‘是的，这说得对。’”[8]

(四)爱因斯坦的狭义相对论

爱因斯坦(Albert Einstein， 1879—1955)出身于德国符腾堡的乌尔姆镇，父母都是犹太人。1896 年 10 月考入苏黎士工业大学攻读物理学。1900 年 8 月毕业后一度失业。1902 年 6 月到伯尔尼瑞士专利局任技术员。

1905 年发表了阐述狭义相对论、光量子理论和布朗运动理论等的四篇重要论文，推动了物理理论的变革。同年以论文《分子大小的新测定法》取得苏黎士大学博士学位。1908 年秋兼任伯尔尼大学编外讲师。1909 年离开专利局任苏黎世大学理论物理学副教授。1914 年回德国任威廉皇家学会物理研究所所长兼柏林大学教授。1916 年发表了《广义相对论基础》。爱因斯坦由于在光量子论方面的贡献，荣获 1921 年诺贝尔物理学奖。1933 年，希特勒执政后，爱因斯坦成为纳粹的迫害对象，幸而他在美国讲学才未遭毒手。同年 10 月在美国定居，任普林斯顿高级研究所研究员。1940 年取得美国国籍。1950 年发表新的统一场论论文。1955 年 4 月 18 日逝世。1905 年建立狭义相对论的历史条件已经成熟，走到狭义相对论门前的人中，有洛伦兹、彭加勒、爱因斯坦。无论就学识的渊博和数学水平来讲，洛伦兹，彭加勒都比爱因斯坦强，但为何不是由他们建立狭义相对论，而是由‘初出茅庐’的爱因斯坦建立狭义相对论呢？他受了哪些因素的推动？受了哪些思想的启发？经历了哪些曲折、克服了哪些困难而建立狭义相对论的？本文主要依据爱因斯坦写的《自述》、《相对论发展简述》和 1922 年 11 月 14 日他在京都大学的讲话《我是如何创立相对论的？》以及他的亲属和朋友的回忆录来回答上述问题，力图把爱因斯坦相对论思想的起源及其思想发展线索揭示出来，以便使我们能从中受到丰富的思想方法论的教育和启迪。

1．“ 追光 ” 问题引起的沉思

少年时代的爱因斯坦就萌发了相对论的思想。1895 年，当他 16 岁在瑞士阿劳中学念书时，曾在无意中想到一个追光的假想实验。他想：“ 如果我以速度 c 追随一条光线运动，那末我就应当看到，这样一条光线就好象一个在空间里振荡着而停滞不前的电磁场。可是无论是依据经验，还是按照麦克斯韦方程，看来都不会有这样的事情。从一开始我凭直觉就很清楚，从这样一个观察者来判断，一切都应当象一个相对于地球是静止的观察者所看到的那样按照同样一些定律进行。 ” [9]

这是与狭义相对论有关的 “ 第一个朴素的理想实验 ”。这个假想实验提出了个佯谬：一个以光速 c 追随光线运动的人应看见电磁驻波存在，而按照麦克斯韦方程没有这样的驻波存在。从他说的 “ 依据经验 ” 这句话来看，可能在那时他知道一些著名的以太漂移实验；从 “ 按照麦克斯韦方程 ” 这句话来看，可能他已熟悉麦克斯韦方程，读过赫兹的著作。

这个佯谬的提出体现了少年时代的爱因斯坦具有非凡的洞察问题的本领。他后来说，这个问题一直使他思考了许多年，经过 10 年的沉思，他才透彻地解决了这一难题。

2．以太漂移实验零结果的启示

爱因斯坦在京都大学的讲话中谈到他的相对论思想的发展是在大学第二年(1897 年)，从考虑以太漂移实验开始的。他力图从物理学文献中找到以太运动的实验证明，但毫无结果，于是他想通过实验来进行。他说：

“ 那时，我想要检验以太相对于地球的运动……当我最早想到这个问题时，我没有怀疑以太的存在，即不怀疑地球在以太中的运动。我想用两个热电偶作如下实验，放置两个平面镜使来自单光源的光在两个不同方向上反射，一束光平行于地球运动的方向，另一束光与之相反。如果这两束光之间有一能量差，就能用两个热电偶测量出它们所产生的热量之差。以证实以太的存在，但是没有进行这个实验。 ” [10]

这段话可以从1899年9月他给未婚妻米列娃 ·马丽奇(Mileva Maric)的信中得到证实：“ 我也给 W · 维恩教授写了信，研究光以太与有重物体的相对运动，因为 “ 上司 ” (H · 韦伯，爱因斯坦在工业大学的物理教授)对此完全像是位继母。 ” “ 由于他的导师们的怀疑是那样的大，因此没有建造这种仪器的机会。 ” [11]

他接着谈到迈克耳孙实验零结果对他的影响。他说：“ 还在学生时代我就在想这个问题了，当时我知道迈克耳孙实验的奇怪的结果。不久我得到这个结论：如果我们承认迈克耳孙的零结果是事实，那么地球相对以太运动的想法就是错的，这是引导我走向狭义相对论的最早想法。自那以后，我认识到虽然地球在环绕太阳运动，但地球运动不能由任何光学实验检验出来。 ” [10] 当迈克耳孙实验的零结果使一些物理学家感到震惊和迷惘，忙于修补以太论的时候，爱因斯坦却得出 “ 地球相对于以太运动的想法是错误的 ” 结论。这与他在 “ 追光 ” 中的思想是一脉相承的。因为在那时他凭直觉相信在相互作匀速运动的参考系中所看到的物理现象 “ 都按同样的定律进行 ”。

为什么又说这是引导他走向狭义相对论的最早的想法呢？这是因为在追光中他是 “ 凭直觉 ”，而这时是从观察实验中得到的结论，比过去对相对性原理的认识就更加具体、更加巩固了。但迈克耳孙实验在他的观念上

并没有引起很大的变化。所以他在晚年与香克兰的几次谈话以及致美国克利夫兰物理协会的信中，都谈到他在 1905 年前“通过洛伦兹的著作知道迈克耳孙的工作”，但是迈克耳孙实验的零结果对他“并没有产生很大的影响”。从爱因斯坦思想发展的历史过程来分析，对这句话就易于理解了，与他在 1922 年的讲话也就没有什么矛盾了。

3．因处于两大矛盾之中

爱因斯坦在《自述》中说，在大学时代，最使他着迷的课题是麦克斯韦理论。他在大学头一年就自学了基尔霍夫、亥姆霍兹、赫兹等人的著作，在后两年又学习了洛仑兹的著作。从这些著作中他首先吸取了光速不变性以及电磁场方程协变性的概念。

早在 1865 年，在麦克斯韦的《电磁场的动力理论》中，就从波动方程中得出电磁波的传播速度$v = 1/\sqrt{\varepsilon\mu}$，$\varepsilon$和$\mu$分别是介质的介电常量和磁导率。对于在真空中传播的电磁波，其波速等于光速c，$c = \dfrac{1}{\sqrt{\varepsilon_0\mu_0}}$为一常量。1890 年，赫兹在写出静止以太参考系中的波动方程时，明确指出波速 c 与波源的运动速度无关。洛伦兹在探讨运动参考系中的波动方程时，始终把光速不变性作为一个限制条件，作为他写出的变换式是否正确的准绳。所以，爱因斯坦在《相对论的基本思想》中指出：“从麦克斯韦-洛伦兹电动力学出发，这个理论采纳了光速不变定律(光速不变原理)。”[9]

洛伦兹在确立他的变换方程时，除了要满足光速不变性的要求外，还要满足协变性的要求，即用洛伦兹变换式来计算从一个惯性系转移到另一个惯性系的变化时，必须要求物理学的一切方程都不改变它们的形式。爱因斯坦在《相对论发展简述》中说：“洛伦兹理论的重大价值在于使静止物体和运动物体的全部电动力学回到了空虚空间的麦克斯韦方程。”[9] 这句话的意思就是说运动参考系中的麦克斯韦方程与静止参考系中的麦克斯韦方程具有相同的形式。

但是洛伦兹理论在一个“有根本性重要意义的地方”使爱因斯坦感到不满，“比起别的运动参考系来，它给以太静止参考系以突出的地位，在这一点上，这个理论与相对性原理相对立。”而爱因斯坦对相对性原理是确信不疑的。他认识到一切经验，包括电动力学领域里的一切经验(特别是迈克耳孙实验)都支持一切惯性系的等效性这个概念,也就是说都是支持相对性原理的。

当时，在确信伽利略变换是正确的前提下，爱因斯坦遇到了一个带根本性的问题：狭义相对性原理同麦克斯韦方程相矛盾。如果按照伽利略变换，麦克斯韦方程在不同参考系中就具有不同的形式，这就违背了相对性原理，由于爱因斯坦对相对性原理是确信的，所以他对麦克斯韦方程产生了怀疑，花了一些时间试图修正麦克斯韦方程。“如果麦克斯韦方程对一个系统是正确的，那么它们在另一个系统中就不正确了。它们应当加以变化。那些年，爱因斯坦用研究和尝试改变麦克斯韦方程的途径试图弄清楚这个问题。他没有取得成功……”[2]，“白白用了近一年时间”。

这时，他还遇到了另一对矛盾：光速不变与速度合成法则相矛盾。他试图用麦克斯韦-洛伦兹方程处理菲佐关于曳引系数的实验，他相信“这些方程是正确的，它们恰当地描述了事实，它们在运动坐标系的正确性表明

了所谓光速不变的关系……可是，这与从力学中所了解的速度合成法则却格格不入”，“他尝试用某种方法将力学方程和电磁现象统一起来，他遇到了困难”。[2]他一度产生了对光速不变性的怀疑：“必须克服的困难是我们早已想到又不得不放弃真空中光速的不变性。只是在探索若干年后，我才注意到这个困难是建立在运动基本概念的任意性上[大概是同时性(simultaneity)这类概念]。”[1]

4．对电磁场对称性的考虑

这几年，爱因斯坦还对电磁场的对称性进行了考虑，他在《论动体电动力学》一开始就指出：“大家知道，麦克斯韦电动力学——象现在通常为人们所理解的那样——应用到运动的物体上时，就要引起一些不对称，而这种不对称似乎不是现象所固有的。比如设想一个磁体同一个导体之间的电动力的相互作用。在这里，可观察到的现象只同导体和磁体的相对运动有关，可是按照通常的看法，这两个物体之中，究竟是这个在运动，还是那个在运动，却是截然不同的两回事。如果是磁铁在运动，导体静止着，那末在磁体附近就会出现一个具有一定能量的电场，它在导体各部分所在的地方产生一股电流。但是如果磁体是静止的，而导体在运动，那末磁体附近就没有电场，可是在导体中却有一电动势，这种电动势本身虽然并不相当于能量，但是它——假定这里所考虑的两种情况中的相对运动是相等的——却会引起电流，这种电流的大小和路线都同前一情况中由电力所产生的一样。”[9]

爱因斯坦想得到一种只有相对运动才具有物理意义的电磁理论，换句话说他一直寻求电磁学关于运动相对性的统一。1973年霍耳顿(Holton)发现了一篇爱因斯坦在1919年写的一篇文章《相对论发展中的基本思想和方法》。在这篇文章中，有爱因斯坦把法拉第定律应用到磁铁和导体环情形时引起的不对称性的清晰的叙述。他写道：

“在狭义相对论建立的过程中，对法拉第电磁感应实验的思考，对我起着导向的作用。按照法拉第定律，在磁铁相对于导体环路的运动中，在导体环路中有感应电流出现。无论是磁铁在运动或者是导体在运动，这一结果是完全相同的，按照麦克斯韦-洛伦兹理论只与相对运动有关。然而对这一现象的理论解释在两种情形下是完全不同的……”[1]

“我不能接受这是两种基本不同情形的思想。我深信这两种情形之间的差别不是实质性的差别，宁可说只不过是参考点选择的差别，从磁铁参考系来看，一定没有电场，而从导体参考系来看，一定有电场。因此电场的存在是相对的，依赖于所选用的坐标系统的运动状态。唯一可以接受的客观事实是完全与观察者或坐标系统的相对运动状态无关的联合在一起的电磁场。这个电磁感应现象迫使我假设狭义相对论原理。”[1]

1950年，爱因斯坦在给美国克利夫兰物理学会的信中谈到了对电磁场对称性与统一性的考虑与建立狭义相对论的关系。他写道：“直接引导我提出狭义相对论的是由于我深信，物体在磁场中所感受到的电动力不过是一种电场力罢了。”

5．普朗克著作的启发

1900年以后，爱因斯坦从阅读普朗克的著作中受到方法论的启发，认识到解决电磁场的相对性与麦克斯韦方程的矛盾走改变方程的途径是行不通的。他在《自述》中写道：

“辐射必须在能量上具有一种分子结构，当然这种结构与麦克斯韦理论是相矛盾的。……早在 1900 年以后，即在普朗克的首创性工作以后不久，这类思考已使我清楚地看到，不论是力学还是热力学(除非在极限情况下)都不能要求严格有效。渐渐地我对那种根据已知事实用构造性的努力去发现真实定律的可能性感到绝望了。我努力得愈久，就愈加绝望，也就愈加确信，只有发现一个普遍形式的原理，才能使我们得到可靠的结果。我认为热力学就是放在我面前的一个范例。在那里，普遍原理是用这样一条定理来说明的：自然规律是这样的，它们使(第一类和第二类)永动机的建造成为不可能。”[1]

爱因斯坦所依据的认识是狭义相对性原理是普遍正确的，物理定理对于一切惯性系是等价的，这是对自然界定律的一条限制性原理，它可以同不存在永动机这样一条作为热力学基础的限制性原理相比拟。

6．光速不变性概念与速度合成之间的矛盾的解决

经过了一段时间的深入思考后，爱因斯坦坚定了对光速不变的信念，从而反过来对速度合成法则所赖以成立的伽利略变换产生了怀疑。但是后者是建立在牛顿的绝对时间、绝对空间概念的基础上的，这就不得不导致时空观念上的彻底变革。攀登这样一个思想高峰，绝非一件轻而易举的事，从各方面提供的材料分析，从哲学批判中吸取精神营养，是他获得观念上变革的重要因素。他在《自述》中说道：

“今天，虽然谁都知道，只要时间的绝对性或同时性的绝对性这条公理不知不觉地留在潜意识里,那末任何想要令人满意地澄清这个悖论(指追光问题)的尝试都是注定要失败的。清楚地认识这条公理以及它的任意性，实际上就意味着问题的解决，对于发现这个中心点所需要的批判思想，就我的情况来说特别是由于阅读了戴维·休谟和恩斯特·马赫的哲学著作而得到决定性的进展。”[1]

我们知道，爱因斯坦在少年时代就开始对哲学发生兴趣，13 岁的时候，读了康德(Kant，1724—1804)的《纯粹理性批判》。1897 年，在大学二年级时读了奥地利物理学家马赫(Ernst Mach，1838—1916)的《力学及其发展的批判历史概论》(以下简称《力学史》)这本书对牛顿力学的绝对性的批判，给爱因斯坦以深刻的影响。

马赫在《力学史》中对牛顿的绝对时间和绝对空间概念进行了分析，他认为时间与空间的量度是与物质的运动分不开的，时空概念是通过经验形成的，是从比较两个事物的快慢中产生的。可是绝对时空无论依据什么经验也不能把握，它只不过是一种无根据的、先验的概念而已。他由此得出结论在力学中具有意义的只是相对运动，绝对运动是毫无意义的。

马赫的批判使爱因斯坦受到很大启发。他对马赫的论述感到“格外引人入胜”。他说“马赫卓越地表述了那些当时还没有成为物理学家公共财富的思想。”[9]

1902—1905 年，爱因斯坦在伯尔尼结识了索洛文(M.So-lovine)与哈比希特 (C.Habicht)，三个青年组成了“奥林比亚科学院”，他们每晚聚集在一起以极大的兴趣和热情研读了斯宾诺莎(Spinoza，1632—1677)、马赫、休谟(D.Humo，1711—1766)、亥姆霍兹、黎曼、彭加勒等人的科学与哲学著作。从这些大师们的经典著作中，爱因斯坦吸取了许多人类思想的精华。斯宾诺莎关于自然界的统一性观念对爱因斯坦有深刻影响。他曾

经声称：“我信仰斯宾诺莎的那个在存在事物的有秩序的和谐中显示出来的上帝，而不信仰那个同人类的命运和行为有牵累的上帝。”[9]即他相信世界的统一。正是因为他对自然界的统一性具有强烈的执着的信念。所以他在 1905 年发表的几篇文章，都具有同一风格，在文章的起始都提出了不对称性问题，即统一性遭到破坏的问题。

这时，休谟的空间和时间观念对爱因斯坦思想的发展具有比较明显的直接的影响。休谟说：“空间或广延的概念不是别的，而是按一定次序分布的可见的或可感知的点的观念。”他还说：“如果我们没有用可觉察的对象充满空间，我们就不会有任何真实的空间观念。”至于时间，它“总是由能够变化的对象的可觉察的变化而发现的”,“没有任何可变的存在”，我们也就不会有“时间观念”。爱因斯坦在他的第一篇相对论论文中，以借助于量尺和时钟定义空间和时间为开端。这种对空间和时间的研究使我们直接回想起休谟的如下论断：空间概念是建立在可触知的对象的排列基础上，而时间观念是建立在能够变化的对象的可觉察的变化基础上。”[2]

马赫对牛顿的绝对时空观提出的批评以及休谟的时空观点，为爱因斯坦在潜意识中放弃时间的绝对性或同时性的绝对性概念打下了基础。有了这样一个基础，相对性原理与伽利略变换之间的矛盾就有了解决的途径。

爱因斯坦 1922 年在京都大学的讲演中,生动地描述了这一矛盾的解决过程：“1905 年 4 月的一天，天气很好，我带着这个问题(光速不变性概念与速度合成法则的矛盾)去访问我在伯尔尼的一位朋友贝索(Michele Besso)。开始，我告诉他：‘近来，我遇到一个难题，今天到这儿来，请你和我一块攻攻它。’我们讨论了这个问题的各个方面。后来我突然找到了问题的关键。第二天，我又来访问他。甚至没有问候一声就对他说：‘谢谢你，我已经完全解决了这个问题。’我的解决办法是，分析时间这个概念不能绝对定义，时间与信号速度之间有不可分的联系，使用这个新概念，我第一次完满地解决了整个问题。”[11]

爱因斯坦经过 10 年沉思，“放弃了许多无效的尝试”，“终于醒悟到时间是可疑的”。[9]在对空间和时间的物理意义作了较深入的分析后，他认识到“伽利略变换是建筑在任意假定的基础上，特别是建筑在同时性的陈述与参考系无关的基础上”。在时空观上产生了一个升华，形成了同时性的相对性概念。突破了这一难关，就完满地解决了整个困难，五个星期后，他就完成了狭义相对论论文。

7．1905 年两篇狭义相对论论文的发表

1905 年 6 月，爱因斯坦完成了他的第一篇狭义相对论论文《论动体的电动力学》。论文一开始就提出了把麦克斯韦理论应用到运动物体引起的不对称问题。然后叙述了相对性原理的来源：“诸如此类的例子，以及企图证实地球相对于‘光介质’运动的实验的失败，引起了这样一种猜想：绝对静止这个概念，不仅在力学中，而且在电动力学中也不符合现象的特性。倒是应当认为，凡是对力学方程适用的一切坐标系对于上述电动力学和光学的定律也一样适用。”[9]他把这猜想提升为公设。并把另一条在表面上看来同它不相容的公设——光的传播速度不变——同发射体的运动状态无关引入，共同作为狭义相对论的两条基本原理，他把这两条原理定义如下：

(1)物理体系的状态据以变化的定律,同描述这些状态变化时所参考的

坐标系究竟是用两个在互相匀速移动着的坐标系中的哪一个并无关系。

(2)任何光线在“静止的”的坐标系中都是以确定的速度 v 运动着，不管这道光线是由静止的还是运动的物体发射出来的。

文章分为两部分：运动学部分和电动力学部分。运动学部分一开始，作者从定义同时性出发，指出了时间和空间与物质运动的联系，分析了长度和时间的相对性。

图 11—6

首先，为了测量时间，必须要有相互校准的时钟，为此他用一个假想的实验确定了在不同地点两钟同步的意义。设在一惯性系中两点 A 和 B，各放有一个时钟。假设一条光线在 t_A 时刻从 A 射向 B，在 t_B 时刻又从 B 反射向 A，在 t'_A 时刻光回到 A。他依据光速不变原理得到光速从 A 到 B 所需的时间等于光从 B 到 A 所需的时间，从而定义两钟同步的条件是：

$$t_B-t_A=t'_A-t_B$$

根据经验

$$\frac{2AB}{t'_A-t_A}=c$$

c 为真空中的光速，是个普适常数。这样，借助于这个假设的物理实验，他解决了如何理解位于不同地点的同步静止钟的问题，并且得到了“同时”或“同步”以及时间的定义。事件的“时间”就是位于事件所在处的静止钟在事件发生的同时给出的时间。

接着，他依据上述同步钟的定义和相对论的两条基本原理论证了长度和时间的相对性，并得出同时性具有相对性的结论。他说：“我们绝不能给同时性概念以绝对的意义。相反，两个事件若在同一坐标系看来是同时的，从另一与这坐标系作相对运动的系统来观察，就不能看作是同时的。”[9]

在讨论两个参考系的坐标和时间的变换理论时，在洛伦兹认为是“纯数学手段”的地方，爱因斯坦揭示了变换方程的实际意义。所谓变换方程指的是：“对于一个完全确定的事件在静系统中的一组空间时间坐标(x，y，z，t)与同一事件在运动系统中的一组空间时间坐标(x'，y'，z'，t')之间的联系。”[9]更为根本的差别是在洛伦兹那里一切变换方程都是先验地提出来的，而爱因斯坦是依据狭义相对论两条基本原理严格地推导出来的。他还把变换方程应用于分析运动球体，得到“在静系统看来，就具有旋转椭球的形状”的结论。应用于分析运动时钟得到“在静系统看来，时钟变慢的效应”。这就是我们现在所说的“时间膨胀”效应。他说：“由此就产生了以下异乎寻常的结果，假定在 S 系中 A、B 两点各有一只静止钟，在静系统看来，这两钟是同步的；如果 A 点的钟以速度 v 沿 AB 线向 B 运动，那末当它到达 B 点时，这两只钟就不再同步了，从 A 到 B 的钟比留在 B 点的钟慢了。”[9]爱因斯坦认为时间膨胀和同时性的相对性是真实的，在所有的惯性系中测量的时间都是等价的。这与把 t'称为地方时间，认为只有静止参考系中的时间 t 才是真实时间的洛伦兹、彭加勒比较起来，爱因斯坦在认识上比他们高出很多。最后，爱因斯坦依据他的时间和空间的坐标变换式导出了速度变换式。在这篇文章的第二部分，爱因斯坦把他导出的洛伦兹变换方程应用于电动力学。他根据相对性原理，在两个相互作匀速

运动的参考系中的麦克斯韦方程应该具有相同的形式得出了电磁场量的变换方程。但是爱因斯坦对变换方程物理意义的解释是不同的。在静系统中作用在单位静止电荷上的力被定义为 E；作用在单位磁极上的力被定义为H。当同一电荷运动时，按旧的理论则作用在这电荷上的力除了E这种电力外，还有一个$v\times\frac{H}{c}$的“电动力”。而在新的理论中“如果一个单位电荷在电磁场中运动，则作用在它上面的力就等于在电荷所在处的电力，这个电力是我们把这电磁场变换到同这单位电荷相对静止的一个坐标系上去时所得出的”。这样，他就解决了，他在文章开始时所提出的把麦克斯韦方程应用到运动物体所引起的不对称问题。

在论文的其余部分，爱因斯坦应用电磁场变换方程和洛伦兹变换说明了多普勒效应和光行差；得出了光的能量变换和光压公式，应用于缓慢加速的电子导出了运动电子的质量和速度的关系式和电子运动方程式，以及电子动能表示式

$$W=m_0c^2(1/\sqrt{1-v^2/c^2-1})$$

同年 9 月，爱因斯坦发表了狭义相对论的第二篇论文《物体的惯性是否与它所含的能量有关？》他依据在前篇文章中导出的能量变换方程，得出辐射能量为L的物体，由于辐射的结果，物体的动能变化为$\frac{1}{2}$。$\frac{L}{c^2}v^2$

由此得出结论，放出辐射能量 L，相当于物体质量减少 L/c^2。他认为从物体中取出的能量变成了辐射能这件事显然是无关紧要的，于是他得出更一般的结论：这个关系不仅对于辐射成立，而且对于所有的能量形式都成立，“物体的质量是它所包含的能量的量度”，“如果能量改变 L，则质量也要同样改变$\frac{L}{c^2}$”。[9]用现代符号表示：如果ΔE表示能量的改变，Δm表示质量的改变，则得$\Delta E=\Delta mc^2$，这就是著名的爱因斯坦质能关系式。他还指出，如果用镭盐那样的能量显著变化的物体来做实验，可以验证这个理论的正确性。

这两篇文章发表后，起初几乎一点影响也没有。尽管如此，在那个时候有些物理学家非常仔细地阅读了爱因斯坦的文章，并且在其中看出了具有无可限量的洞察力的思想。英费耳德在《相对论发展史》中提到一位波兰物理学家维特科夫斯基对他的同行说：“去读一读爱因斯坦的文章吧。一位新的哥白尼诞生了！”

普朗克很快注意到爱因斯坦的理论。他在 1906 年讲了下面的话：“与洛伦兹相比，爱因斯坦则用更为一般的形式导出了“相对性原理。但是普朗克没有注意到时间、空间概念的变更。爱因斯坦本人当然意识到他的理论核心是时间空间概念的变革。1905 年 5 月他在致朋友的信中说他自己正在热衷于研究“利用空间时间概念变革的电动力学”。除爱因斯坦外，最早注意到空间时间概念变革的是数学家闵可夫斯基(H.Minkowski，1864—1909)。1907 年 11 月 5 日，闵可夫斯基在哥廷根数学会议上作了题为“相对性原理”的讲演，他在讲演的开头就说：“以光的电磁理论为开端，在我们的时空观念中，一个彻底的变革似乎发生了。”[2]1908 年 9 月闵可夫斯基在德国自然科学家大会上发表了“空间和时间”的讲演，预言了爱因斯坦的思想对近代思维所要起的深远影响，他以下面的话开始了他的讲

演：“现在我要向你们提出的空间和时间观念是在实验物理学的土壤中产生的，其力量就在这里，这些观点是根本性的。从现在起，孤立的空间和孤立的时间，注定要销声匿迹，只有两者的统一才能保持独立存在。”[2]然而，要使时空概念的变革引起人们的广泛注意，尚需一段时间。随着历史的发展，相对论的结论为越来越多的实验所证实。1909 年布歇勒(Bucherer)用电磁偏转法测量了镭源放出的β射线(快速电子束)的荷质比 e/m 与电子速度的关系，首次证明了爱因斯坦的质量与速度的关系式。1915 年，盖依(Guye)和拉万曲(Lavanchy)测量了用静电加速器加速的电子在电磁场中的偏转，以更高的精度巩固地确立了这一公式的正确性。在后来的核物理实验中，对大量高速粒子运动的观察更加证实了这一点。

既然相对论能用许多实验方法来证实，那末这也间接地表明它已成为许多富有成果的新理论的出发点和物理学家从事研究的不可缺少的工具。1916 年索末菲(Sommerfeld)用相对论的动量和能量表达式解释了氢原子光谱线的精细结构。 1922 年康普顿用相对论解释了 X 射线经物质散射波长改变的现象。1923 年德布罗意就是根据“时钟频率的相对论性变化及波的频率之间的差异”引入物质波概念的。在以后进行的核反应实验中处处都证实着爱因斯坦质能关系的正确性。事实表明这一理论已成为现代科学技术的基础。

参考文献

[1]Arthur I. Miller， AlbertEinstein’s Speeial Theory of Relativity Emergence(1905) and Early Interpretation(1905—1911)，Addison-Wesley， 1981 11—36， 40—41，70—72，79—81，145

[2][日]广重彻著，李醒民译，《物理学史》，第 1 版，北京，求实出版社，1988 年，610，620，626，629—630，632

[3]申先甲等编著《物理学史简编》。第 1 版，济南、山东教育出版社，1985 年 1 月，673

[4]A.爱因斯坦等著，赵志田，刘一贯译，《相对论原理》(狭义相对论和广义相对论经典论文集)，科学出版社，1980 年，1—5，6—26

[5]Stanley Goldberg，The Lorentz Theory of elec trons andEinstein’s theory of Relativity American Journal of physics，October 1969，Vol37，No10，pp982—994

[6] E.T.Whittaker， AHistoryoftheTheoriesofAetherand Electricity，Edinburgh 1953，pp30—31

[7][法]彭加勒著，叶蕴理译，《科学与假设》，商务印书馆，1989 年，55—65

[8]赵中立、许良英编译，《纪念爱因斯坦译文集》，上海科学技术出版社， 1979 年， 288—289， 317—332

[9]许良英，范岱年，赵中立编译，《爱因斯坦文集》，商务印书馆，1976 年，第一卷， 1—51，83—90， 149—155， 181—190，第二卷 83—115，116—118，第四卷 490

[10] A.Einstein，How I Createdthetheoryofrelativity，Physics Today，August，1982， Vol 35， No8， 45

[11]Johnstachel，“爱因斯坦和以太漂移实验”，《物理通报》，1988

年 5 月，44—46

十二、玻尔原子的起源

尼耳斯·玻尔是物质结构的量子理论的创立者。1913 年他以《原子和分子的结构》为题发表了后人称为“伟大的三部曲”的三篇论文。他把光谱现象、普朗克和爱因斯坦提出的光量子说以及卢瑟福提出的原子核模型结合起来，建立了原子结构的量子理论，为人类认识微观世界打开了大门。

本文的主要任务是探讨玻尔是在什么背景下，经过哪些途径、受到哪些启发、克服了哪些困难把上述三者结合在一起的。

(一)世纪之交原子结构模型的研究

19 世纪末至 20 世纪初，电子、X 射线、放射性元素的发现，突破了经典物理学的框框，冲破了原子绝对不可分、元素绝对不变的传统观念，揭示了原子有复杂的内部结构，标志着人类对物质结构的认识进入到一个新阶段。

1．J.J.汤姆孙的原子模型

英国物理学家 J.J.汤姆孙(JosephJohnThomson，1856—1940)出生于英国曼彻斯特，毕业于剑桥大学，1884 年任卡文迪什实验室教授。他通过阴极射线的研究发现了电子的存在，认识到电子是一切原子的基本组成部分之一。他测定了电子的荷质比为氢离子荷质比的一千多倍，从而得出电子的质量小于氢原子质量千分之一的结论。

1904 年，J.J.汤姆孙提出了第一个比较有影响的原子模型。他设想原子中的正电荷以均匀的密度连续地分布在整个原子球中，许多细小的电子均匀地分布在其内，正象面包里的葡萄干一样，原子的这个“plum pudding”模型如图 12－1 所示。原子中的

图 12—1 汤姆孙的原子模型

电子静止在一定的平衡位置上，这一位置由电子与电子之间的排斥力和电子与正电荷之间的吸引力的平衡来确定。当原子从外界得到一定的能量时，原子内的电子就在其平衡位置附近振动，同时发出各种不同波长的电磁波。在汤姆孙模型的基础上可以定性地理解激发态原子的辐射，但没有得到计算值与观测到的光谱线的实验值的一致性。[1]

应该指出第一个试图将量子假设应用于原子结构的人是维也纳大学的博士生阿图尔·埃里希·哈斯(ArthurErichHass)，他在 1910 年 2 月写的论文里，试图结合原子的组成来考虑作用量子 hv 的本性。他探讨了一个质量为 m、电荷为 e 的单个电子在半径为 a 的中性的汤姆孙原子内的振荡，根据牛顿力学所作的简单计算表明，电子在作简谐运动，振动的频率

$v^2 = \dfrac{e^2}{4\pi^2 ma^3}$，哈斯根据普朗克的能量$E = hv$的公式和古典力学的振幅与能量的关系，使振幅等于半径 a 就得到 $hv=e^2/a$ 从上述方程中消去 v。哈斯得到普朗克作用量子的电动力学解释 $h=2\pi e(ma)^{1/2}$。由此可见

$a = \dfrac{h^2}{4\pi^2 e^2 m}$。哈斯未写出这个等式,严格地说这个值是以后玻尔获得的。他们从完全不同的前提出发得到了一致的结果。哈斯的计算是建立在汤姆孙模型基础上，而玻尔的计算是建立在卢瑟福模型的基础上。[2]

2．卢瑟福的原子模型

物理学家卢瑟福(ErnestRutherford,1871—1937)出生于新西兰。1895年他来到剑桥大学在汤姆孙领导的卡文迪什实验室当研究生。1907年卢瑟福主持了曼彻斯特大学实验室。自从天然放射性现象发现后，经过多年的努力，他终于证实α粒子是带两个正电荷的氦离子。卢瑟福决定用这种新粒子当炮弹来轰击原子，以探索原子的内部结构。他和年轻的德国物理学家盖革(Hans Geiger,1882—1945)、青年学生马斯顿(ErnestMarsden,1889—)一起作α粒子散射实验。他们用镭作放射源，进行α粒子穿射金属箔的实验，发现入射束中多数粒子仍保持其原来的方向，但也有不少粒子偏转了很大角度。他们精心测量了极少的大角度散射的粒子，结果发现约有八千分之一的α粒子偏转角度超过90°，甚至有反弹回来的。通过对这些实验结果的思考，使卢瑟福得到这样的印象：“它是如此难以令人置信，正好象你用十五英寸的枪射击一张薄纸，而子弹居然反弹回来把你打中了一样。”[3]这个结果与汤姆孙模型的预言完全不符。按照汤姆孙模型，原子的质量和正电荷几乎是均匀地分布在整个原子中，在这种情形下，入射粒子的电荷与原子内部的电荷之间的相互作用绝不会强到使α粒子离开其原来的运动方向发生大角度偏折。卢瑟福意识到这种大角散射也不可能是由于很多小偏离的累积效应。1910年底，卢瑟福开始把散射实验事实与新的原子模型联系起来。他认为α粒子是在同靶原子的一次碰撞中改变其方向的，因而静电斥力必须集中在一个极小的范围内，即原子中有一个体积很小、质量很大，对正电荷有很强偏转能力的核。核外则是一个很大的空间，核的体积很小，直径约为 10^{-12}—10^{-13} 厘米，约为原子直径的万分之一到十万分之一，但却几乎集中了原子的全部质量，带负电的轻得多的电子则在很大的空间里绕核运动，它看起来就象行星绕太阳的运动。一定元素的原子核上的正电荷数目等于核外电子数。原子结构的现代模型就这样问世了。[3]

为了检验这一模型是否符合观察到的散射结果，必须根据电学定律导出一个公式能算出α粒子在离排斥中心不同距离处通过时偏转的大小。这个公式指出偏离原来运动方向角θ的α粒子数，应当与$\sin\frac{\theta}{2}$的四次方成反比。这个结论与观察到的散射曲

图12—2 原子的核型模型

线非常相符，这就证明了原子的核型模型的正确性。[1]

不过，卢瑟福模型一开始就由于同经典电动力学理论尖锐矛盾而遇到了困难。因为按照经典理论，电子绕核运动是加速的。会自动向外辐射电磁波，从而丧失能量而逐渐落向原子核。在这一过程中，电子绕核转动频

率连续改变，应向外发射连续光谱。这些推论同原子的稳定性以及原子发射线状光谱的事实相矛盾。于是人们就遇到一个十分棘手的问题，如何说明实在原子的稳定性？这个稳定性问题就成为 20 世纪初期摆在物理学家面前的几个困难问题之一。

(二)玻尔早期的物理思想

尼耳斯·玻尔(NielsBohr，1885—1962)是丹麦杰出的物理学家，近代

量子物理学的奠基者之一。他出生于哥本哈根大学一个生理学教授的家庭，自幼就受到家庭浓厚的科学文化思想的熏陶。1903 年秋季，玻尔进入哥本哈根大学的“ 数学和自然科学系 ”，主修物理学。他在那里作为一个能力非凡的研究者显得与众不同。他的第一件研究工作完成于 1906 年, 当时他还是一个学生。他通过观察水柱的规则振动来测定水的表面张力，这项工作使他获得了丹麦皇家科学文学院的金奖章。1909 年 6 月他完成了科学硕士论文《试论述电子论在解释金属的物理性质方面的应用》。1911 年 5 月 13 日，他答辩了他的博士论文《金属电子论的研究》，获得了哥本哈根大学哲学博士学位。[4]

建立在 J.J.汤姆孙关于电子存在的实验和洛伦兹关于电子行为的理论基础上的金属电子论，试图通过金属内部存在着自由电子来解释金属的奇特性质。玻尔的研究在经典理论的基础上对以往许多人的工作进行了全面的分析和检查。他在作为博士论文的开场白的底稿中写道：“ 这些科学家所发展的理论，在许多基本问题上和实验符合得惊人地好……但是，在许多本质性问题上，经验和理论结果却完全不相符合。 ”[4]玻尔认为这种不符合是由于经典理论本身不完善所致。

在讨论外磁场对金属中自由电子的影响时，当时人们对这一问题的看法是很不一致的。有的人认为金属中自由电子的存在对金属的抗磁性有贡献，而另一些人则认为它们对它的顺磁性有贡献。玻尔运用统计观点讨论了这个问题，证明了自由电子对金属的磁性没有任何影响。他在科学硕士论文中写道：“ 在磁场的影响下，电子的路径将发生弯曲，而且电子将沿一条螺旋线运动，螺旋线的轴线平行于场的轴线。J.J.汤姆孙相信路径的这种弯曲将沿着和外场相反的方向引起磁场，从而包含着自由电子的那块金属就象抗磁体一样地起作用。……但是这种似乎得到普遍承认的观点，在我看来却是错误的。 ”接着他用统计观点对自由电子不引起任何磁效应加以解释。“ 现在，考虑受到一个恒定磁场作用的金属内部的一个小体积元，由于对称性，电子将沿着相对于场轴线的一切方向运动，因此这样一个体积元将不引起任何磁效应，从而整块金属也不会引起磁效应。 ”[4]他在博士论文中讨论了磁场对金属表面处的自由电子的影响后，得出结论：“ 一块处于电平衡和温度平衡的金属，不会因为自由电子的存在而具有任何磁学性质。 ”在他的博士论文的摘要的最后一条中，也就是全篇论文的最后的一句话中，这样写着：“ 在电子论的目前发展阶段，根据这一理论来解释物质的磁学性质看来还是不可能的。 ”[4]在玻尔看来，经典电子理论没有解释物质的磁学性质的可能，要做到这一点必须对经典理论加以改造。

在讨论金属对热射线的吸收和反射的问题时，玻尔在他的博士论文中指出：“ 如果人们坚持保留电磁理论所依据的那些假设，那似乎就不可能解释热辐射定律。这或许是由于这样一种情况：电磁理论和实际情况并不一致。……从而不能用它来考察单独一个电子在短的时间间隔中的运动，因为那将使我们远远超出本论文所处理的那些问题的范围，所以我们将不再接下去讨论在电磁理论中引入根本性变动的那些努力，那些努力似乎正在导致很有兴趣的结果呢！ ”在这段话中，玻尔更加明确地肯定了经典理论的局限性。他提到的“ 在电磁理论中引入根本性变动的那些努力 ”。显然是指普朗克和爱因斯坦在量子理论方面的工作，玻尔认为这些工作“ 似

乎正在导致很有兴趣的结果”。1911 年 10 月 23 日，他给他的弟弟若哈德写的信中谈到汤姆孙在热射线吸收计算中的问题时说：“我唯一确信的是，他认为可以找到将会在普通电磁学定律的基础上解释热辐射定律的一种机械模型，而正如我已经间接证明并且后来又由麦克拉润直接证明了的，这种事情是显然不可能的。”[4]玻尔在此表明经典的电磁理论不能解决热辐射的问题。由此我们可以看到在玻尔早期的物理思想中已经孕育着新的物理思维，包含着量子概念的萌芽。

(三)接受卢瑟福原子模型

1911 年夏季，玻尔从哥本哈根大学毕业。这年 10 月到英国剑桥大学卡文迪什实验室参加研究工作。起初，他希望在 J.J.汤姆孙指导下继续进行他在金属电子论方面的工作。但是，汤姆孙已对这一课题失去了兴趣，而且不肯与玻尔密切合作和经常会见，这种会见对于玻尔发展其物理思想是很必要的。玻尔想发表《金属电子论》英译本的努力也没有成功。正当玻尔处于彷徨歧途，无所适从的心境时，当时在曼彻斯特工作得很有成绩的卢瑟福出现在他的面前了。1911 年 10 月间，卢瑟福到剑桥来参加卡文迪什实验室年度聚餐会。他在集会上发表了长篇演讲，论述了自己的新发现。[5]玻尔后来回忆说：他对“卢瑟福性格上的魅力和魄力得到了深刻的印象。”他有转到曼彻斯特工作去的愿望，为此玻尔请曼彻斯特的生理学家史密斯(J.L.Smith,玻尔父亲生前的学生和朋友)安排了他与卢瑟福的会见。卢瑟福向他介绍了他刚参加第一次索尔维物理会议的情况(这次会议的议题是辐射和量子理论),也提到普朗克和爱因斯坦的看法以及自己对物理学发展前景的看法。这次谈话对玻尔以后的科学生涯产生了十分重要的影响，使他了解到量子论进一步发展的情况，种下了他不久将量子论和原子核模型结合起来的种子。[6]

玻尔于 1912 年 3 月间来到曼彻斯特工作,这在他的终身事业中是一个最有决定意义的转折。当时曼彻斯特大学物理实验室已经形成了一个研究放射性问题的学术中心，在那里集中了一批最有才能的物理学家。玻尔刚刚到达时，就按照卢瑟福的安排作了一些实验。他在这个实验室经受了英国研究风格的训练。后来玻尔时常提到他经历的幸运环境：“生于一个小国，没有民族自负感，在青年时代就接受了两个世界的优良传统，即‘大陆的理论传统和英国的经验主义’。”[6]在卢瑟福实验室，玻尔的这个特点是独树一帜的。玻尔到来时，卢瑟福等人已经确证了原子核的存在，追索原子核所将引起的各个方面的后果，成为整个曼彻斯特科学集体的兴趣中心。

玻尔热心地接受了新的原子模型，而且很快就认识到它的深远含义。他指出元素的化学、物理性质，由原子的核外电子所决定。原子所含的电子数目决定原子在周期表中的位置。也就是说在决定原子的化学性质上，原子序数比原子量更根本。当玻尔听说周期表在原子量顺序方面显示了一两处反常，已经鉴定下来的稳定元素和衰变元素的数目超过了著名的门捷列夫周期表中可以利用的位置时，他一下子就想到这些在化学上无法分离的物质具有相同的核电荷，其不同只在核的质量和内在结构了。玻尔证明了如果人们承认电荷相同而质量不同的原子核的存在，从而有不止一个原子品种占据周期表中的同一位置，则所有这些反常性都可以消除。后来把

这种具有不同质量而在化学上不可区分的原子命名为“同位素”。[7]

玻尔还指出元素的放射性质直接与原子核有关，由放射性元素放出的α粒子及电子来自原子核。早在1902年卢瑟福就提出了放射性元素的嬗变理论。他指出放射性物质是不稳定的，它不断放射出某种射线进行衰变，即由母元素变为子元素直至变成稳定元素为止。玻尔按照核型模型、放射性衰变必须被设想为原子核的实际嬗变。于是玻尔就论证说，通过α射线的发射，原子核就失去两个单位电荷而变成周期表上退后两位处的那一元素的同位素。另一方面，在β衰变中，一个负电子的发射导致一个单位电荷的获得，从而嬗变后的元素就占据了周期表上前一位的位置。[7]尽管看起来很简单，导致这些放射元素的“位移定律”的那种猜想在当时却不是显而易见的。

当时，曼彻斯特集体中的许多科学家，包括卢瑟福本人在内，他们起初对于发现原子核所引起的深远后果，并不是一下子就清楚地意识到了的。卢瑟福作为一个实验物理学家，他习惯于脚踏实地地工作，对于从已有的事实进行“外推”的作法，则采取一种谨慎的态度。玻尔在回忆中说：“当我找到卢瑟福去听听他对这些见解的反应时，他照例对任何有希望的见解表示了敏锐的兴趣，但他以特有的慎重提出告诫说，不要过度强调原子核型模型的意义，并从比较贫乏的实验资料进行外推。”[7]这时，实验室中对玻尔的见解颇感兴趣，并有所理解的人是匈牙利青年化学家G.希维思，他利用放射性同位素作为示踪原子进行的化学研究，支持了玻尔的观点，使玻尔在与他的讨论中受到很大的鼓励。[6]

(四)定态假设

玻尔对于卢瑟福原子模型的含意的敏锐考察，并没有停留在认识原子序数和原子中电子数目之间的关系上。他断然进攻了一个更加困难得多的问题，那就是确定这一关系的确切本性的问题，而这就要对核型模型所代表的原子结构进行动力学分析。

在曼彻斯特的几个月中，玻尔通过参加α射线的吸收实验，对这个问题有了更深入的了解。物质对α射线的吸收提供了原子结构的重要信息，他日以继夜地以惊人的速度完成了这一工作。1912年9月，玻尔在哥本哈根物理学会上发表的演讲中指出：“可以说，α粒子和β粒子通过它们在物质中经过时所遭受的变化，告诉我们它们在物质中经过的途中所遇到的事情……如果我竟然试图……描述我们通过研究α射线的散射和吸收而得出的两种信息的区别，我也许就可以说，第一种现象通过指示原子内部那些力场的强度和本性，即存在着的粒子的数目和电荷，来告诉我们原子的静力学性能，而第二种现象则告诉我们原子的动力学性能，因为吸收(它是依赖于电子的运动的)将提供关于电子运动频率的信息。”[5]玻尔在这里强调α粒子的吸收实验，提供了原子的动力学性能的信息。α粒子在穿透物质的过程中，能使它们击中的原子发生电离，而它们本身将不断地损失能量。其损失的能量是与原子电离时所需的电离能有关的，而电离能也就是电子与原子结合时的结合能，从古典力学知道，这种结合能与频率之间有一定的关系，所以α射线的吸收实验提供了关于电子频率的信息。

1912年7月6日，玻尔在离开曼彻斯特前，把一份论文提纲交给了卢瑟福，这份提纲主要讨论了原子和分子在正常状态时的稳定性问题，后来

海耳布朗和库恩曾把这份提纲称为《卢瑟福备忘录》。在这份提纲中，玻尔对原子结构进行了动力学分析。他假设原子中核外电子按相等的间隔排列在一个圆形轨道上绕原子核转动。这就是玻尔最初采用的原子模型，这种原子模型是一个很薄很薄的圆片，而不是球体。玻尔面临着怎样解释在圆环上运动的电子的稳定性问题。[5]

为了描述电子的运动，他当然只能利用当时唯一的经典力学。例如，当只有一个电子时，该电子受到来自原子核的电吸引力，它起着向心力的作用。即

$$m(4\pi^2 v^2 r) = X\frac{e^2}{r^2}$$

式中 m 是电子质量，v 是电子的绕转频率，r 是轨道半径，而 X 是原子序数。困难在于，只有这样一个方程，显然不能确定 r 和 v。玻尔根据他对金属电子行为的研究，已经确信经典电动力学的适用性在原子范围内将受到根本性的限制，而且他毫不怀疑这种限制将在某种方式下受到普朗克作用量子的支配。玻尔在回忆中说：“1912 年春天，在我停留于曼彻斯特的早期阶段，我已经确信卢瑟福原子内电子的运动是彻头彻尾由作用量子支配着了。这种观点得到不只一件事实的支持。”按照普朗克的能量子假设，每个谐振子的能量只能取一些分立值，它们是有限能量元 ε 的整数倍，ε 是能量的最小单元。ε=hν，h 是普朗克常量，ν 是振子频率。为了解决上述问题，玻尔把普朗克的能量子假设引入他自己的量子化条件，要求各电子的动能和它们的绕转频率成正比，即

$$E=Kv$$

式中 K 是一个和普朗克常量同数量级的比例常量。[5]

到此为止，玻尔还只是在寻找足够的方程，并没有明确涉及稳定性问题。为了解决原子的稳定性问题，玻尔引入了一个定态假设：能量满足上述量子化条件的运动才是一种稳定的运动。这就是最初的定态概念。这种定态假设并不是对原子稳定性的一种“解释”，而只是对事实的一种认可，它是由一些经验事实归纳出来的结论。

玻尔指出，他的这一假设似乎能够从实验事实那里得到有力的支持，而且似乎可以对某些现象和规律作出某种意义下的说明。他举出了四点可以说明的事物，即元素的原子体积随原子序数的周期变化规律等等。在《卢瑟福备忘录》中，玻尔还进行了一些计算，试图判断什么样的电子环可以有相对的稳定性，这时他不但考虑了原子结构模型，而且考虑了几种分子结构模型。到此为止，玻尔还只考虑了原子和分子的“正常态”，没有引入跃迁概念，即有关定态之间的过渡问题。[5]

(五)量子假设

英国的天体物理学家尼科耳孙(J.W.Nicholson,1881—1955)，从 1911 年 11 月到 1912 年 6 月，在英国的《皇家天文学会月报》上发表了一系列有关星云光谱和日冕光谱起源的论文。当时，他并不知道卢瑟福的工作，但是却在论文中提出了关于原子核的思想。更妙的是，尼科耳孙当时也不知道哈斯的工作，但是却独立地想到了利用量子概念来确保“电子环的稳定性”。他假设各电子绕核运动的角动量只能等于普朗克常量的整数倍。这一假设和后来玻尔提出的量子公设的一种表述相同。但是尼科耳孙并没

有引入量子轨道(定态运动)的概念，他只计算了谱线频率。他所设想的发光机制是不对的，他认为原子发光是由于核外电子环的振动所引起的。[5]

尼科耳孙的工作对玻尔思想的转变起了一定的促进作用。1912 年 12 月 23 日,玻尔在给他的弟弟的信中指出,他自己的理论只考虑到原子的‘最终态”，而尼科耳孙的理论考虑了原子正在发光时的状态，涉及到光谱线系频率的计算问题。

当时，光谱学已经积累了浩如烟海的资料。人们已经知道原子所发出的光谱反映了原子的化学种类的特征。各特征谱线的波长包含着发光原子结构的很准确的信息，但是，原子光谱是由包括上千条谱线的线系构成的，看来要破译这样复杂密码的希望甚微。然而,使玻尔出乎意料的是， 1913 年 2 月在他和哥本哈根的光谱学家汉森(H.M.Hansen)的一次闲谈中，得悉光谱学家们已经设法在这一团乱丝后面找出了一些规律。特别谈到里德伯已经发现了表示若干“系列”光谱线频率的一个结构上很简单、很引人注意的公式。里德伯公式的惊人特点是各个波数都可用两个“谱项”之差来表示，即 $v=T'-T''$。其中每一谱项都以一种简单的方式依赖于一个数，它可以取一系列整数值，每个线系都对应于保持一项不变，而令另一项改变的系列。例如氢光谱线的波数 v_{nm} 可以用巴耳末公式表示成

$$v_{nm}=R\left(\frac{1}{n^2}-\frac{1}{m^2}\right)$$

在这里 R 是里德伯常量。这一公式的简洁性使玻尔很受启发。他以后不只一次地说过：“我一看见巴耳末公式，整个问题对我来说就全都清楚了。”玻尔在上大学时，从他的老师那里学到了一些光谱学方面的知识，但当他对原子结构进行了一番研究之后，光谱学的规律对他才又有了新的意义。玻尔后来回忆道：

“我在 1912 年秋季写给卢瑟福的那些信件,谈到了追寻作用量子在卢瑟福原子的电子结构中所起的作用的持续努力，其中也包括分子键的问题和辐射效应及磁效应的问题。但是，稳定性问题在所有这些考虑中引起了纠缠不清的困难，这些困难刺激着我们去寻求更坚实的立脚点。然而，经过应用量子概念的各式各样尝试以后，我在 1913 年的初春才认识到，直接适用于卢瑟福原子的稳定性问题的一个线索，是由支配着元素光谱的惊人地简单的定律提供出来的。”[7]

玻尔把光谱学的规律和卢瑟福的原子模型联系起来，他觉察到光谱学规律给出了所缺少的线索，可以找到必须将作用量子引入原子体系描述中所遵循的正确方式。正如玻尔后来所说：“事实上，接受了爱因斯坦关于能量为 hv(h 是普朗克常量)的光量子或光子概念，人们不免就要假设，原子对辐射的每一发射或吸收，都伴随着传递的能量为 $h(T'-T'')$的一个过程，并将 hcT 解释为原子的某种稳定状态中或所谓定态中的电子结合能。这一假设给线系谱中各发射谱线及吸收谱线的表观特性提供了直截了当的解释。例如，在发射过程中，我们看到的是原子从高能级到低能级的跃迁，而在吸收过程中，我们遇到的则一般是原子从具有最低能量的基态到它的一个受激态的跃迁。”[7]他又说：“在氢光谱这一最简单的情况中，各谱项可以很精确地表示为 $T_n=R/n^2$。”当 n 逐次增加时，就导致氢原子中电子结合能一系列递减的值，相当于电子处于更高的能级。离核很远的电子将

通过辐射跃迁而进入越来越低的 n 值，表示进入结合得越来越牢固的定态，能量越来越低的状态，直到 n=1 的基态为止。

在新的认识基础上，玻尔重新建立了他的理论，并把它详细阐述为一本论著。1913 年 7 月、9 月和 11 月玻尔在英国的《哲学杂志》上先后发表了著名的原子结构和氢光谱理论的三部曲《论原子和分子的构成》。这本论著成了原子物理学的划时代文献。他在第一篇文章中，按照前述的思路提出了三条基本假设(prin-cipal assumption)：

1．在定态中体系的动力学平衡可以借助普通的力学来加以讨论，而体系在不同定态之间的过渡，则不能在这样的基础上进行处理。

2．后一种过程伴随着单色辐射(homogeneous radiation)的发射，对于这种辐射来说，频率和所发射的能量之间的关系式就是由普朗克理论所给出的那一关系式

$$h\nu=W_2-W_1$$

式中 h 是普朗克常量。[8]

3．对任一原子体系的持久态(permanent state)每一电子绕其轨道中心的角动量等于的整数倍，即

$$M=\tau\frac{h}{2\pi}\qquad \tau=1，2，3\cdots$$

THE

LONDON,EDINBURGH,AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE

[SIXTH SERIES.]

JULY 1913.

I.On the Constitution of Aioms and Alolecules.

By N. BoHR,Dr. plail. Copenhagen.

Introduction.

IN order to explain the resalts of experiments on seattering of a rnys by matter Prof. Rutherford has given a theory of the strueture of atonis. According to this theory, the atoms consist of a positively charged nncleus surrommded by a system of electrons kept together by attractive forees from the mucleus; the totnl negative charge of the electrons is equal to the positive charge of the mucleus.Further, the nucleus is assumed to bo the sent of the essntinl part of the mass of the ntom, and to have linear dimensions exceedingly small compared with the linear dimensions of the whole atom. The number of oleetrons in an ntom is dednced th be approximately equal to half the atomie weight. Great interest is to be attributed to this atom-model;for, as Rutherford has shown, the assumplion of the existence of muclei,as those in question,seems to be necessary in order to aeeount for the resnlts of the experiments on large angle senttering of the a raysl

In an attempt to explain some of the properties of matter on the basis of this atom-model we meet, however, with Jifficulties of a sorikous. naturo arising from the apparent

*Communicated by Prof. E. Rutherford,F.R.S.

E. Rutherford,Phil. Mag. xxi,p.660(1911).

See also Geiger and marsden, Phil. Mag. April 1013.

Phil. Mag. S.6.Vol.151.July1913.

图 12－31913 年 7 月玻尔在英国的《哲学杂志》上发表的著名论文《论原子和分子的构成》

玻尔证明了后一假设是与他称为的特殊假设(special assump-tion)等价的。该假设指出：在电子从无限远(此时电子相对于核的速度为零)与核结合的过程中(during the bindingoftheelectron)(结合后电子绕核作圆运动)，所发射单色光的频率等于电子在它最后轨道上转动频率的一半，即

$$v=\frac{1}{2}\omega\text{，}\omega\text{为转动频率}$$

据普朗克理论，上述过程中发射的能量

$$W=\tau h\frac{\omega}{2}\qquad \omega=1\text{，}2\text{，}3\text{，}\cdots$$

玻尔根据他提出的三条假设导出了电子绕核运动的频率ω，椭圆轨道的主轴 2α，以及要把电子从轨道上迁移到无限远的距离必须给系统输入的能量值 W

$$W=\frac{2\pi^2me^2E^2}{\tau^2h^2}$$

$$\omega=\frac{4\pi^2me^2E^2}{\tau^3h^3}$$

$$2a=\frac{\tau^2h^2}{2\pi^2meE}$$

式中 m 是电子质量，e 和 E 分别表示电子和核的电荷。玻尔写道，在上述表示式中如果给τ以不同的值，就得到一系列的 W、ω和α的值，相应于系统的一系列组态，这些组态相应于系统处于稳定的状态。如果τ=1，则 W 的值最大，这种情形相应于系统处于最稳定的状态。也就是相应于把与核结合的电子分裂开(forthebreakingup)所需的最大的能量。

玻尔接着写道：“在上述表示式中，如果令τ=1 和 E=e 并引入实验值 $e=4.7\times10^{-10}$，$\frac{e}{m}=5.31\times10^{11}$，$h=6.5\times10^{-27}$就得到$2\alpha=1.1\times10^{-8}\text{cm}$，$\omega=6.2\times10^{15}S^{-1}$，$\frac{W}{e}=13V$. 由此看到这些值与原子的线度、光的频率和电离势具有相同的数量级。”

在“三部曲”第一部的第二节中，玻尔还讨论了谱线的发射。对于氢原子因为 E=e，从而就得到氢原子的能量公式为

$$W_\tau=\frac{2\pi^2me^4}{h^2\tau^2}$$

因此，这个系统从$\tau=\tau_1$的状态过渡到$\tau=\tau_2$的状态，发射的能量是

$$W_{\tau2}-W_{\tau1}=\frac{2\pi^2me^4}{h^2}\left(\frac{1}{\tau_2^2}-\frac{1}{\tau_1^2}\right)$$

如果，我们假设上述辐射是单色的，发射的能量就等于 hv，其中 v 是辐射

频率，那么，我们就得到

$$W\tau_2-W\tau_1=hv$$

由此得 $$v=\frac{2\pi^2 me^4}{h^3}\left(\frac{1}{\tau_2^2}-\frac{1}{\tau_1^2}\right)$$

由此我们看到这个表示式说明了氢谱线系的规律。结果我们令 $\tau_2=2$，令 τ_1 改变，我们就得到通常的巴尔末线系。如果我们令 $\tau_2=3$，我们就得到由帕邢观察到的红外线系。如果我们令 $\tau_2=1$ 和 $\tau_2=4$，5…，我们就可以分别得到一系列看不到的，但实际上存在的紫外线系和红外线系。

这种一致性不仅是定性的，而且是定量的。我们把

$$\varepsilon=4.7\times10^{-10}，\ \frac{e}{m}=5.31\times10^{11}\text{和}\eta=6.5\times10^{-27}$$

代入上面公式中括号外的因子，我们就得到

$$\frac{2\pi^2 me^4}{h^3}=3.1\times10^{15}$$

它的观察值为

$$3.290\times10^{15}$$

这个理论值与观察值之间的差别是在实验误差所允许的范围内。

玻尔理论提出了一个动态原子结构的轮廓，揭示了光谱线与原子结构的内在联系，指出了分析光谱是研究原子内部结构的重要途径，从而推动了物质结构理论的发展。由于这一开拓性的贡献，玻尔获得了 1922 年的诺贝尔物理学奖。

(六)玻尔理论的确证及其引起的反响

1896 年到 1897 年,美国天文学家匹克林(E.C.Pickering,1846—1919)在船舻座ζ星的光谱中发现了一个很象巴耳末系的线系，这个线系中每隔一条谱线和巴耳末系的谱线差不多重合。[3]因此，这种谱线一般被认为属于氢谱线，甚至里德伯也这样认为。杰出的光谱学家佛勒(A.Fowler)也持有这一观点。当时，他刚刚在用充有氢氦混合气体的放电管作的实验中，观察到了匹克林线及新的有关线系。[7]这一线系曾用 n 的半整数值归结到氢　的公式中

$$\tilde{v}=R_H\left(\frac{1}{2^2}-\frac{1}{n^2}\right)\quad n=2.5,3,3.5,4,\cdots$$

按照玻尔的看法，这些数必须有整数值，而这种情况只可能意味着这一线系的里德伯常量对应于带两倍电荷的核，是氢的里德伯常量的四倍。于是，很自然地可以认为匹克林线系是氦离子产生的。

由于 He 核的电荷 E=2e，所以

$$\tilde{v}=4R_H\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right)$$

上式中如果令 $n_1=4$，就得到匹克林线系

$$\tilde{v}=4R_H\left(\frac{1}{4^2}-\frac{1}{n^2}\right)\quad n=5,6,7,\cdots$$

将此式改写为

$$\tilde{v} = R_H \left(\frac{1}{2^2} - \frac{1}{\left(\frac{n}{2}\right)^2} \right)$$

就得到与经验公式一致的结果。这就论证了匹克林谱线是氦离子的光谱线。玻尔写信给卢瑟福请他在他的实验室中给予验证。

“我已经试着把这些谱线归因于氦。这一点是可以用实验来验证的。在一次讨论中，我提出我的观点，在佛勒实验中氢的存在可能并不是这些谱线的直接原因(正如我在论文中所说的那样)。化学家布杰隆博士向我建议说，如果我的观点是正确的，这些谱线就会在充有氦氯混合气的管子里出现；而且他还说，这样的谱线会有更大的强度。不过，我们在哥本哈根没有条件做出满意的实验来。因此我冒昧请求你在你的实验室里检验一下。”[3]应卢瑟福的要求，伊万斯(E.J.Evans)在一个玻璃管内充入极纯净的氦气得到了匹克林谱线。然而佛勒仍然不完全信服，他尖锐地指出，就这些有争议的谱线而言，里德伯公式中的系数并不精确地等于 4R。玻尔有力地作出答辩，这个微小的差别产生于氦核的不可忽略的运动。在推导公式时假定了核的质量很大，因而可以忽略它的运动。但实际情况是，电子和核都绕公共质心运动。[3]因此里德伯常量表示式中的质量 m 不应被看成自由电子的质量，而应被看成所谓的约化质量 $\frac{mM}{m+M}$，此处M是核的质量。事实上，将这一改正考虑在内，所预言的氢光谱和氦离子光谱之间的关系，就和一切测量结果完全一致了，从而消除了佛勒的怀疑。

玻尔的“三部曲”发表后，引起了物理学界的强烈反响。有赞同的，也有反对的。卢瑟福看了“三部曲”的第一部的初稿，承认玻尔关于光谱发射的看法是“很巧妙的，而且看来是很好用的”。但是他接着就提出了问题：

“……但是，普朗克概念和旧力学的混合，却很难使人对什么是它的基础形成一个物理概念。在我看来，你的假设中有一个严重的困难，这个困难我毫不怀疑地认为你也充分意识到了。那就是，当电子从一个定态过渡到另一个定态时，它怎么决定将以什么频率运动呢？在我看来，你似乎必须假设电子在事先就知道将在什么地方停下来。”[5]

卢瑟福一针见血的批判，击中了玻尔理论的要害，提出了普朗克理论和经典力学的结合，很难形成统一的理论基础，在理解上造成极大的困难，在逻辑上难于自洽一致。

在“三部曲”第一部问世不久，有人在瑞士苏黎世的一次物理讨论会上简略地介绍了它的内容。当讨论即将结束时，劳厄曾经抗议说：“这完全是胡扯，麦克斯韦方程在任何情况下都是成立的。”这时爱因斯坦站起来声明：“这是很值得注意的！在它背后一定有点什么玩意儿。我不相信里德伯常量的导出是全靠运气。”[5]1913 年 9 月 7 日，在伯明翰召开的英国科学促进会会议上，金斯(James Jeans)综述了量子理论应用于原子结

构的进展，他说：“玻尔博士对光谱线规律做出了极富天才的解释。我觉得还应加上‘令人信服的’这几个字。”

当海维塞(Joseph George von Hevesy，1885—1966)在维也纳把玻尔的发现告诉爱因斯坦时，爱因斯坦说：“这可是个重大的成就，玻尔的理论一定是正确的。”[3]

1914年，弗兰克(James Franck，1882—1964)和赫兹通过电子碰撞吸收实验，对玻尔的量子能态假设作出了最直接的实验验证。

玻尔的贡献不仅体现在科学上的非凡成就，而且还以他的科学思想、科学作风、高尚情操影响了一代青年科学家。特别是他的爱国主义精神感人至深。1918 年，玻尔的导师和挚友卢瑟福写出“私人信件，本人亲启”的邀请信，以“把曼彻斯特办成现代物理研究中心”，“年薪两百英磅(相当于在丹麦收入的两倍)”为前提，两次请玻尔去英国任职。“你知道，我们是多么喜欢有你和我们一起工作啊！我想，你我两人能努力使物理学蓬勃发展起来。好好考虑，一有决定，就立刻告诉我。”

卢瑟福提供的美好前景简直无法抗拒，但是，玻尔回信道：

“……我一直强烈地希望在您身旁工作，以分享您对大家施予的，使人受益匪浅的热情和想像力。……”

“我非常喜欢再次到曼彻斯特去。我知道，这对我的科学研究会有极大的重要性。但我觉得不能接受您提到的这一职务，因为哥本哈根大学已经尽全力来支持我的工作，虽则它在财力上，在人员能力上，以及在实验室的管理上，都远远达不到英国的水平。”

“我立志尽力帮助丹麦发展自己的物理学研究工作……我的职责是在这里尽我的全部力量。”

玻尔对自己祖国的热爱促使他一再婉言拒绝外来的高薪聘请，决心在人口不到五百万的小国建立起物理学的国际中心。确实，他成功了，哥本哈根很快成了物理学家“朝拜的圣地”！

1939年，当法西斯的铁蹄正在践踏欧洲各国的时候，玻尔毅然决然地从美国回到丹麦。他不仅要回去维持研究所的工作，而且还要继续帮助逃离纳粹占领区的科学家。这就是玻尔爱国主义最生动的写照。玻尔一直记在心里的，并经常加以引用的是丹麦童话家安徒生的名言：“丹麦是我出生的地方，是我的家乡，这里就是我心中的世界开始的地方。”[9]

参考文献

[1] 〔美〕乔治·盖莫夫，高士圻译，《物理学发展史》，北京，商务印书馆，204， 210—214

[2]John L.Heilbron,Bohr'sfirsttheoriesoftheatom,Physics Today, October 1985, 30—34

[3]申先甲等编著，《物理学史简编》，第1版，济南，山东教育出版社，1985年1月，737—739，743—744

[4] L．罗森菲耳德主编，戈革译，《尼尔斯·玻尔集》，第一卷，北京，商务印书馆，1986年7月，15—21，105，111，127，165，278，412

[5] 戈革，《尼耳斯·玻尔——他的生平、学术和思想》，上海，上海人民出版社，1985年10月，32—34，72—79，90—98，109，121—123

[6] 阎康年，“卢瑟福和尼尔斯·玻尔的友谊与合作及其对现代物理学的影响”，《物理》，1986年7月
[7]玻尔，《1958 年度卢瑟福纪念讲座——关于原子核科学的奠基人以及以他的工作为基础的若干发展的一些回忆》
[8] N.BOHR, On the Constitution of Atoms and Molecules, PHILOSOPHICAL MAGAZINE and JOURNAL of SCIENCE,July 1913, 2－7
[9] 杨福家，“尼尔斯·玻尔的魅力——纪念玻尔诞生一百周年”，王福山主编，《近代物理学史研究》(二)，第一版，上海，复旦大学出版社，1986年7月

十三、物质波理论的创立

路易·德布罗意(Louis de Broglie)是物质波理论的创立者，他提出了电子和原子中的其它粒子具有波动性的假说，创造性地构思了物质波的概念，为开创现代量子力学的新时代奠定了基础。

本文所要介绍的内容是：(一)物质波思想的起源；(二)物质波理论；(三)物质波理论的应用；(四)物质波思想的影响及实验验证。

路易·德布罗意，1892 年 8 月 15 日生于法国塞纳河畔迪埃普(Dieppe)，出身于贵族。中学毕业后，他进入巴黎大学学习，最初学习历史、法律，1910 年获学士学位。在学习期间受到他的哥哥莫里斯·德布罗意研究 X 射线和庞加莱(H. Poincare，1854—1912)著作的影响，对科学感兴趣，1911 年改学物理学，1913 年获得“科学证书”。第一次世界大战期间，在埃菲尔铁塔上的军用无线电报站服役六年(1913—1919)。战后，他重新钻研物理并在他哥哥的实验室参加实验研究工作。1922—1924 年间逐渐形成他的物质波思想。1924 年获巴黎大学科学博士学位。1928 年任巴黎大学理论物理教授。1929 年获得了诺贝尔物理学奖。1933 年被选为法国科学院院士。

(一)物质波思想的起源

德布罗意在谈到关于波粒二象性思想的起源时，这样写道：“在我年轻时代，也就是在 1911—1919 年间，我满腔热情地钻研了那个时期理论物理的一切最新成果。我了解庞加莱、洛伦兹、郎之万……的著作，也了解玻耳兹曼和吉布斯关于统计力学方面的著作。但是，特别引起我注意的是普朗克、爱因斯坦、玻尔论述量子的著作。我注意到爱因斯坦 1905 年在光量子理论中提出的辐射中波和粒子共存是自然界的一个本质现象。在随我哥哥莫里斯作了 X 射线谱的研究后，我觉察到电磁辐射的这种二重性具有十分重要的意义。在研究了力学中的哈密顿-雅可比理论后，我进一步在其中发现了一种波粒统一的初期理论。最后，在深入地研究了相对论后，我深信它一定是一切新的假设的基础。”[1]他在其它地方还谈到布里渊的原子核周围的以太波的思想对他的影响。本节将从德布罗意提出的这几个方面，阐述物质波思想的起源。

1．X 射线的研究

自从 1895 年伦琴(Röntgen，1845—1923)发现 X 射线以后，对 X 射线的研究就成了一个非常热门的课题。据 X 射线与 α 射线及 β 射线从放射性原子中射出，以及 X 射线通过气体时具有电离气体的能力等性质来判断，X 射线显然是一种粒子流。1911 年，布拉格(W. H. Bragg，1862—1942)在一篇题为《γ 射线、X 射线粒子假设的结论和 β 射线的种类》的论文中写道：“一束 X 射线为一束 β 射线提供能量；同样，在 X 射线管中，一束 β 射线激发一束 X 射线……次级 β 射线的速率是与 X 射线通过的距离无关的，因此 X 射线的能量在运动中不会扩散，这就是说，这是一种粒子。”[2] 1912—1914 年间，莫塞莱(Moseley)利用 X 射线波谱验证了玻尔模型，又发现了 X 射线谱的频率正比于原子序数 Z 的平方，再一次揭示了 X 射线的粒子性。1912 年劳厄(Laue 1879—1960)提出如果 X 射线是波长极短的电磁波，它通过晶体会产生衍射现象。这一设想被在实验室工作的弗里德

里希(Friedrich)和克尼平(Knipping)所证实。当 X 射线通过晶体时，发现了明显的衍射现象，有力地显示了 X 射线的波动性。

X 射线时而象波时而象粒子的奇特性质使一些实验物理学家感到困惑不解。1912 年末，在回顾这种情形时，布拉格作了这段深刻的、有远见的叙述：“对我来说，问题似乎并不在于判定 X 光的两种理论何者更为正确，而在于去寻求一种理论，它能同时把握这两个方面。”[2]

莫里斯·德布罗意和布拉格都出席了 1911 年在布鲁塞尔举行的第一次索尔维(Solvay)国际物理会议。布拉格关于 X 光的研究给了莫里斯很大影响，稍后他的主要研究方向即转为对 X 射线本性的探索。他具有与布拉格类似的观点，但却不想从理论上探讨这一问题。路易·德布罗意说：“我的哥哥把 X 射线看作波和粒子的一种结合，但他不是一个理论物理学家，对这个问题没有特别清晰的看法。”[2]莫里斯让他的弟弟阅读索尔维会议的记录，在他的装置优良的私人物理实验室里进行 X 射线的研究。莫里斯在 1913 年 10 月和 1921 年 4 月在布鲁塞尔举行的第二和第三次国际物理会议任科学秘书时，曾将劳厄在两次会议上和布拉格在第二次会议上分别作的关于 X 射线晶体衍射和反射强度的专题报告，及卢瑟福在第三次会议上报告他的《原子结构》时提到莫塞莱的元素 X 光谱实验的文件，带回家整理，使路易详细了解到会议的文件和讨论情况，对 X 射线的波动性和粒子性等问题以及当时的研究动向有了深入的认识。[3]

1919 年战争结束后，路易·德布罗意又回到他兄长的实验室进行 X 射线研究。他的头两篇论文就是用玻尔原子、理论分析 X 射线的吸收。随后又与他兄长合作研究光电效应产生的光谱的测量结果，它还与人合作研究 X 射线光谱问题。这些研究不但使路易了解到原子结构的许多知识，更重要的是使他形成了 X 射线具有波粒二象性的思想。[2]他在一篇自传性的短文中写道：“我曾长期同我的哥哥讨论如何解释关于光电效应和粒子谱的漂亮实验。……同他进行有关 X 射线性质的长时间的讨论……使我陷入波和粒子必定总是结合在一起的沉思中。”[2]

2．接受相对论和光量子学说

由 X 射线的研究所引起的对波粒二象性的思索，并不是决定德布罗意思想方向的唯一因素，普朗克的量子论、爱因斯坦的相对论和光量子学说对德布罗意物质波概念的形成起着极大的作用。他后来回忆道：“我怀着年轻人特有的热情对这些问题发生了浓厚的兴趣，我决心致力于探究普朗克早在 10 年前就已引入理论物理的、但还不理解其深刻意义的奇异的量子。”[4]

战后，他继续在巴黎大学攻读他的博士学位。他对于涉及到时间、空间、物质结构和光的本性等的基本概念特别感兴趣。例如郎之万的相对论讲演和对时间概念的分析(郎之万是第一个详细处理所谓‘时钟佯谬’的人)给德布罗意留下了深刻的印象。事实上，他的相位波概念是与他对时间问题的思考紧密相连的。他曾写道：“时钟频率的相对论性变化及波的频率之间的差异是基本的，它极大地引起了我的注意，仔细地考虑这个差异，决定了我的整个研究方向。”[2]

德布罗意很早就读过关于光量子假说的文章。这些文章和对 X 射线的研究使他接受了光的波粒二象性思想。1922 年 1 月 26 日，他发表了题为《黑体辐射和光量子》[4]的论文。他在引言中谈到他试图‘只利用热力学，

气体动理论和量子理论，而不借助于电磁学推出辐射理论中一系列已知的结果。”他采用光量子假说，把光子当作具有能量 hv 和动量 hv/v 的光量子来处理。在他的计算中，把一定温度下的黑体辐射认为是由光原子组成的气体。他引进了光原子静止质量 m_0 的概念并定义为

$$W = hv = \frac{m_0 c^2}{\sqrt{1 - v^2 / c^2}}$$

在这里 v 是光量子的速度，他认为 v 低于但又无限接近于真空中的光速 c。然后，他按照气体动理论讨论了这些有质量的光原子组成的系统，推导出能量在 W—W＋dW 范围内单位体积中的光原子数为

$$dn_W = \frac{n}{2k^3T^3} e^{-\frac{W}{kT}} W^2 dW$$

在这里 n 是单位体积中的光原子数，k 是玻耳兹曼常量，T 是绝对温度。每个光原子具有 3kT 的平均能量，是气体分子具有非相对论性能量的两倍。接着，德布罗意指出：如果考虑一种由单原子、双原子、三原子……组成的“光的混合气体”，则可得到普朗克的黑体辐射公式。

1922 年 11 月，德布罗意发表了《干涉与光量子》[4] 一文，在此文中，他考虑了干涉问题，认为必须用光量子假设来解释干涉现象。他谈到爱因斯坦用能量涨落对热平衡时的黑体辐射所作的分析，并写出了爱因斯坦的表示式

$$\overline{(E - \overline{E})^2} = \sum_{n=1}^{\infty} nhvE_n$$

在此 $E_n = \frac{8\pi h v^3 \upsilon}{c^3} e^{-\frac{\pi h v}{kT}}$ $n = 1，2，3\cdots$

E表示在体积V中的实际辐射能量，而$\overline{E} = \sum_{n=1}^{\infty}$是在热平衡时它的平均值。

德布罗意认为，上述公式“具有如下的重要性：从光量子的观点来看，干涉现象与光原子的集合有关。这些光原子的运动不仅不是独立的，而且是相干的。因此，如果有一天光量子理论能够解释干涉现象，那么采用这种量子集合的假定是十分自然的。”[4]这就是说利用这种集合效应可以描述辐射所具有的包括干涉现象在内的象波一样的行为。1963 年，在巴黎跟库恩会晤时，德布罗意指出，这篇论文是他后来工作的出发点。“我开始有了那种想法，不过它尚未诞生，我可能不敢讲出来，但我心中已经开始孕育它。”[5]这种对量子论的兴趣促使他去探索把物质的波动方面和粒子方面统一起来的理论。

1923 年 4 月 21 日，康普顿公布了光经散射后波长变长的结果，并用光量子理论作了正确解释，这使德布罗意更加坚定了自已的信念。在德布罗意看来，“最近几年累积的实验证据对光量子的实际存在是一个有力的支持。光电效应看来正越来越倾向于由爱因斯坦的光子假说所统治。不采用光量子概念，要解释康普顿关于散射 X 射线波长变长的最终结果是十分困难的。”[6]但他又不同意一般人把康普顿实验作为粒子判决性实验的看法。德布罗意认为，在一个粒子理论中能量项由 hv 来定义是不能令人满意的。于是他宣称，“仅此一端就提供了一个必要性，即在光的理论中应该

同时引进粒子概念和周期性概念。”[7]

德布罗意后来谈到他是如何把光的波粒二象性推广到实物粒子上来的。他在 1964 年春天致 Fritz Kubli 的信中写道：“在 1922－1923 年期间，当我开始获得波动力学的基本想法时，我的意图是把爱因斯坦发现的光的波粒共存现象推广到所有粒子，因此我开始用爱因斯坦建立的光量子公式：W= hv 和 p=hv/c=h/λ。我把这些公式应用到光子以外的其它粒子上。对于这些粒子有下列公式

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} = hv$$

和

$$p = \frac{m_0 \upsilon}{\sqrt{1-\beta^2}} = \frac{h}{\lambda} \qquad \left(\beta = \frac{\upsilon}{c}\right)$$

不过从我看来，这和爱因斯坦理论还对不上，除非人们赋予光子以适当的质量 μ_0，这样允许人们以下列形式写出爱因斯坦方程

$$W = \frac{\mu_0 c^2}{\sqrt{1-\beta^2}} = hv$$

和

$$p = \frac{\mu^0 \upsilon}{\sqrt{1-\beta^2}} = \frac{h}{\lambda}[4]$$

德布罗意根据相对论力学和光量子假说，从考虑对称性出发，把光看成是由质量极小的，但不为零的，速度接近于但又略低于真空中光速 c 的光原子所组成。他认为光子具有质量的想法对他提出实物粒子具有波动性的想法起了很大的作用。

3．哈密顿-雅可比理论的启示[2]

最早试图揭开光的本性的神秘面纱，把牛顿的粒子说和惠更斯的波动说统一起来的人是哈密顿(W.R.Hamilton1805—1865)，他想使光学成为具有像拉格朗日所赋予力学那样的“审美性、权威性与和谐性”的一门正规科学，并通过寻求同时支配光的传播和粒子运动的一种单一的自然规律来达到这点。借助于他的作用函数 S=∫Ldt(式中拉格朗日函数 L=T-U，T 为质点系的动能，U 为势能。)，他发现在一个力场中质点的运动与光射线的传播，受同一形式的规律所支配。根据哈密顿的变分原理，对于真实运动来讲作用函数的变分为零，即 $\delta\int L dt = 0$。在能量为常数时，可以导出莫伯丢(Maupertu)的最小作用原理，这个动力学的古老的变分原理

$$\delta\int 2T dt = 0$$

或

$$\delta\int [2m(E-U)]^{1/2} ds = 0$$

而几何光学的费马(Fermat)最小时间原理，这一几何光学的基础为

$$\delta\int \frac{n}{c} ds = 0$$

即在光线的实际路程上光程的变分等于零。把以上二式加以比较，我们看到在力学中的表示式 $[2m(E-U)]^{\frac{1}{2}}$ 与光学中的相速度 $u = c/n$ 起着相同的作用，可以用符号表示为

$$u = \frac{c}{n} \leftrightarrow [2m(E-U)]^{\frac{1}{2}}$$

这一关系称为“哈密顿的光学力学类比”。然而，这一类比还可以进一步加以推广。作用函数 $\int_{t_0}^{t} Ldt$ 在位形空间（configura - tion space）确定了一个“作用面”（actionsurface）。众所周知，$p = \nabla S$，而且 $\frac{\partial S}{\partial t} = -E$(E 是力学体系的总能量 E=T+U)。另一方面对平面单色波，波矢 k 和频率 $\omega = 2\pi v$ 与相位 $\varphi = -\omega t + k \cdot r$ 具有如下关系 $k = \nabla\varphi$，及 $\frac{\partial \varphi}{\partial t} = -\omega$。因此，在传播中粒子系统的常量作用面完全类似于光学中的同相位面。波失相应于动量，而频率相应于粒子的能量。这些结果是哈密顿在 1828 年到 1837 年发表的。

哈密顿-雅可比理论给了德布罗意很大的启发。40 年后他在追忆波粒二象性思想是如何产生时说：“突然间，我萌发了把光的二象性推广到物质粒子，尤其是电子的想法。我不能给出确切的日子，但一定是在 1923 年夏天。我认识到，一方面，哈密顿在某种程度上指出了这点，因为它能适用于粒子，同时又表示一种几何光学；另方面，在量子现象中得到的量子数很少在力学中被找到，但却常常出现在波动现象以及与波动运动有关的所有问题中。于是，我自信有一种与量子现象相联系的波。”[4]

4．布里渊的驻波[7]

早在 1913 年，玻尔就提出了原子中电子运动的量子化条件。成功地解释了氢原子光谱等许多当时物理学不能解决的问题。但是玻尔关于电子绕核运动的角动量等于 $\frac{h}{2\pi}$ 的整数倍是人为规定的.没有理论基础,不能使人信服。因此，许多人都构造了各种各样的理论来导出玻尔这一量子化条件。法国的布里渊(M.Bri-llouin)在 1919—1922 年发表了一系列文章,提出了一个解释玻尔量子化条件的理论。布里渊设想原子核周围存在着一层以太。电子就在该以太层中运动，这运动在以太层内掀起波，这些波相互干涉而在原子核周围形成环形驻波。布里渊认为这一情况可以作为对玻尔神秘的量子化条件的一种物理解释。布里渊是第一个把电子与波作为一个整体研究的人，但他借助于当时已被物理学家抛弃的以太介质，使得物理学家难于接受。布里渊曾把自己的上述工作告诉了德布罗意，从而使后者认真地把光的二象性假说同对玻尔模型的研究结合在一起考虑，并把布里渊尊称为“波动力学的真正先驱”。布里渊的儿子证实了这一点：“德布罗意采用了我父亲 1919—1922 年间提出的一些相当不成熟的想法。”但是德布罗意抛弃了以太的错误看法。把属于以太的周期性运动给了电子本身，实现了向波粒二象性思想的过渡。

(二)物质波理论

1923 年 9 月和 10 月，德布罗意在《法国科学院导报》上发表了三篇关于物质波理论的短篇论文。这些短文加起来不过十来页却包含了他的新理论的全部要点，标志着物质波的发现。在写完上述文章后，德布罗意马上投入博士论文的写作。1924 年夏天，他完成了他的题为《量子理论的研

究》的博士论文，11 月 25 日在巴黎大学举行了论文答辩，这篇论文后来发表在 1925 年 1—2 月号的物理杂志上。他的论文对前一年获得的结果提出了一个系统的有逻辑性的报告，完整地阐述了他的物质波理论及其应用。

1．相位波概念的提出

1923 年 9 月 10 日，德布罗意发表了第一篇关于物质波的论文。题为《辐射——波和量子》。[8] 在这篇文章中，他提出实物粒子也有波粒二象性，引入了与运动粒子相缔合的波的概念。德布罗意考虑相对于观察者以速度 $\upsilon=\beta c(\beta<1)$运动，其静质量为 m_0 的一个质点。根据相对论质能关系，该质点具有 m_0c^2 的静能。德布罗意把这一能量同量子现象联系起来，他认为 $hv_0=m_0c^2$，即可把与这一静能相联系的运动视作频率为 v_0 的简单的周期性现象。

接着，德布罗意考察了两个频率：一方面，从静止的观察者看来，对应于动点的能量有一频率 v。他认为 $hv=mc^2$，即

$$v=\frac{m_0c^2}{h\sqrt{1-\beta^2}}=\frac{v_0}{\sqrt{1-\beta^2}}$$

另一方面，按照相对论给出的运动时钟变慢效应，当那位静止的观察者观察动点的内在周期性现象时，他就会认为这一现象变缓慢了，即将它看成频率$v_1=v_0\sqrt{1-\beta^2}$的周期性现象,也就是说这一内在周期性现象按$\sin 2\pi v_1$ t 的形式变化。

这两个频率 v 和 v_1 的差别引起了德布罗意很大的注意，正如他自己所说：“仔细地考虑这个差别决定了我研究工作的整个方向。”为了解决这一矛盾，德布罗意引入了一个“与质点运动相缔合的波”。这波以频率 v，速度 c/β，沿着与质点相同的运动方向传播。因为速度超过真空中的光束 c，这个波不可能传递任何能量。德布罗意把它视为一种与质点运动相缔合的假设的波。然后，他证明了如果在开始的时候，动点的内在周期性现象在相位上与波是一致的，那么，在任何时刻，这种相位的一致性将保持下去。

如果在时刻 t，动点与原点的距离为 $x=\upsilon t$，其内在周期性运动按$\sin 2\pi v_1\frac{x}{\upsilon}$变化，而在同一点上与动点相联系的波可以表示为

$$\sin 2\pi v\left(t-\frac{x}{c/\beta}\right)=\sin 2\pi vx\left(\frac{1}{v}-\frac{\beta}{c}\right)$$

由前面定义的 v_1 和 v 可知

$$v_1=v(1-\beta^2)$$

所以

$$\sin 2\pi v_2\frac{x}{\upsilon}=\sin 2\pi vx\left(\frac{1}{\upsilon}-\frac{\beta}{c}\right)$$

上式说明表示动点内在运动的正弦函数与表示波的正弦函数相等，所以两者在相位上永远保持一致。德布罗意指出：“这一重要结果的论证完全是以狭义相对论原理和严格的量子关系式为基础的。”

两个星期后，也就是同年 9 月 24 日，德布罗意发表了第二篇关于物质

波的论文《光量子、衍射和干涉》[9]。在文章开始德布罗意明确提出了相波概念。他说："为了描述一个速度为 βc 的动点的运动，观察者必须将这一动点与一个非物质的，以速度 $c/\beta=c^2/\upsilon$，在同一方向上传播的正弦波联系起来。在这个观察者看来，这一波的频率等于上述动点的总能量除以普朗克常量 h。"由于"这动点同位于同一点上的波具有相同的相位"，因此德布罗意称此波为"相波"(phasewave)。在博士论文中，德布罗意把这一波称为"相位波"，认为它是与相位的传播相联系的。

2．波与粒子的对应关系

(1) 波射线与粒子路径的一致性

1923 年 10 月 8 日，德布罗意发表了关于物质波的第三篇论文《量子、气体运动理论以及费马原理》[11]。在这篇文章中，德布罗意更加明确地阐述了他的物质波思想。他假定与任何粒子相联系的相波，在空间任何点与粒子同相位。相波的频率与速度由粒子的能量和速度所决定。他强调这些相波特别应当具有这种性质："相波的射线应当与动力学上粒子的可能轨迹相一致。"他声称相波的射线应当用光学上的费马原理来描述。即

$$\delta\int n\mathrm{d}s=0$$

因为 $n=\lambda_0/\lambda$，λ_0 为光在真空中的波长，λ 为光在介质中的波长，因此有

$$\delta\int\frac{\mathrm{d}s}{\lambda}=\delta\int\frac{v\mathrm{d}s}{c^2/\upsilon}=\delta\int\frac{m_0\beta c}{h\sqrt{1-\upsilon^2/c^2}}\mathrm{d}s=0$$

另一方面，粒子的运动轨迹可以由莫伯丢最小作用原理来描述。在空间两点之间粒子的实际路径满足动量的变分为零，即 $\delta\int m\upsilon\mathrm{d}s=0$，因此有

$$\delta\int\frac{m_0(\beta c)^2}{\sqrt{1-\beta^2}}\mathrm{d}t=\delta\int\frac{m_0\beta c}{\sqrt{1-\beta^2}}\mathrm{d}s=0$$

两者在形式上完全一致。这就表明空间两点之间粒子的实际路径，与其相波射线的实际路径是完全一致的。至此"联接几何光学和动力学两大原理的基本关系完全明朗"。[10]

德布罗意在他的博士论文第二章《莫伯丢原理和费马原理》中，引入了相对论的四维动量和相波的四维矢量，证明了这两个原理中的任一个能用来描述量子。他得出结论："适用于相波的费马原理与适用于运动物体的莫伯丢原理是等同的；运动物体的动力学可能轨道与波的可能射线是等同的。"

(2) 群速度与粒子速度的等同性

在博士论文第一章[11]中，德布罗意把相波的两个传播速度、相速度和群速度联系起来，并证明了波的群速度等于粒子速度。

他考虑在同一方向传播的、频率相近且速度随频率变化的两个波。其频率分别为 v 和 $v'=v+\mathrm{d}v$，速度为 υ 和 $\upsilon'=\upsilon+\frac{\mathrm{d}\upsilon}{\mathrm{d}v}\mathrm{d}v$，这一速度称为相速度，即相位的传播速度。将两个波叠加，在略去了 δv 的二次项的条件下可得

$$\sin 2\pi\left(\nu t-\frac{\nu x}{\upsilon}+\varphi\right)+\sin 2\pi\left(\nu' t-\frac{\nu' x}{\upsilon'}+\varphi'\right)$$

$$=2\cos 2\pi\left(\frac{\delta\nu}{2}t-x\frac{d\frac{\nu}{\upsilon}}{d\nu}\frac{d\nu}{2}+\varphi'\right)\sin 2\pi\left(\nu t-\frac{\nu x}{\upsilon}+\varphi\right)$$

式中余弦部分表示一个缓变的“振幅”，它代表合成波的整体轮廓，表示一条调制曲线，上式就是一个振幅受频率 δν 调制的正弦合成波，这种调制波就称之为波包，包迹的移动速度是波包的整体移动速度，称之为群速度，用 υ_g 表示，由上式的余弦部分可以得

$$\frac{1}{\upsilon_g}=\frac{d\left(\frac{\nu}{\upsilon}\right)}{d\nu}$$

现在来证明相波的群速度等于运动物体的速度。如果给予运动物体一个速度 U=βc，β 的值没有完全确定，但仅要求这个速度在 β 和 β+dβ 之间；相应的频率在一个很小的间隔 ν 和 ν+dν 之间。与该粒子运动相缔合的相波的速度 υ 和 ν 可以看成是 β 的函数，因为有

$$\upsilon=\frac{c}{\beta}\qquad \nu=\frac{1}{h}\frac{m_0c^2}{\sqrt{1-\beta^2}}$$

相波波群的群速度可以写成

$$\upsilon_g=\frac{d\nu}{d\beta}\Bigg/\frac{d\left(\frac{\nu}{\upsilon}\right)}{d\beta}$$

而

$$\frac{d\nu}{d\beta}=\frac{m_0c^2}{h}\frac{\beta}{(1-\beta^2)^{3/2}}$$

$$\frac{d\left(\frac{\nu}{\upsilon}\right)}{d\beta}=\frac{m_0c}{h}\frac{d\left(\frac{\beta}{\sqrt{1-\beta^2}}\right)}{d\beta}=\frac{m_0c}{h}\frac{1}{(1-\beta^2)^{3/2}}$$

因此 $\upsilon_g=\beta c=U$

相波的群速度正好等于运动物体的速度。“这个关系对理论的发展是很重要的。”

(3)波长与粒子动量的关系

在博士论文第七章《统计力学和量子理论》中，德布罗意明确指出，对于速度较小的非相对论气体分子，相波波长为

$$\lambda=\frac{c^2/\upsilon}{m_0c^2/h}=\frac{h}{m_0\upsilon}$$

对运动质点同样有 $\lambda=\frac{h}{m\upsilon}$. 这就是著名的德布罗意波长与动量的关系式。

值得注意的是这一公式以如此明晰的形式在他的论文中出现，仅仅只有这

一次。这个式子和 E=hν 一起，后来称为爱因斯坦-德布罗意关系。

(三)物质波理论的应用

1．玻尔量子化条件的物理解释[8]

德布罗意在他的第一篇物质波论文中，把他的相波概念应用到以闭合轨道绕核运动的电子，从而推出了玻尔的量子化条件，这是德布罗意波能够成立的一个有力证据。德布罗意假定在 t=0 时刻，电子位于 O 点，并以速度 v=βc 运动。它所缔合的波也在 O 点并以速度 c/β沿电子轨道运行。假定在 t 时刻波与电子再次重合在 O′点。设电子的运动周期为 T，由于电子与波的速度不同，则二者重合必须满足下式

$$t\frac{c^2}{\upsilon}=\upsilon(t+T)$$

或
$$t=\frac{\beta}{1-\beta^2}T$$

因此在从 O 点到 O′点时，电子内在运动的相位变化为

$$2\pi v_1 t=2\pi\frac{m_0c^2}{h}\frac{T\beta^2}{\sqrt{1-\beta^2}}$$

德布罗意认为如果电子轨道是稳定的，那么经过 O′点的波必须在相位上与电子保持一致，也就是说与电子缔合的波必须在电子轨道上形成驻波，于是就有

$$2\pi\frac{m_0c^2}{h}\frac{T\beta}{\sqrt{1-\beta^2}}=2\pi n \qquad (n\text{为整数})$$

或
$$\frac{m_0\beta^2c^2}{\sqrt{1-\beta^2}}T=nh$$

这一条件与玻尔-索末菲(Bohr-Sommerfeld)条件

$$\int_0^T(p_x dx+p_y dy+p_z dz)=nh$$ 完全一致。

因为
$$\int_0^T(p_x dx+p_y dy+p_z dz)$$
$$=\int_0^T\frac{m_0}{\sqrt{1-\beta^2}}(\upsilon_x^2+\upsilon_y^2+\upsilon_z^2)dt=\frac{m_0\beta^2c^2}{\sqrt{1-\beta^2}}T$$
$$=nh$$

设电子的角速度为ω。在半径为 R 的圆轨道上绕核旋转，则在电子速度相当小的情形下，可再次得到玻尔的角动量量子化条件

$$m\omega R^2=n\frac{h}{2\pi}$$

他在第三篇物质波论文中提出“只有满足相波谐振的那条轨道才是稳定的轨道”。电子的轨道是由 $\frac{v}{\upsilon}$ 对闭合路径的积分取整数值来确定，在博士论文第三章《轨道稳定性的量子条件》中写出下式

$$\oint\frac{v}{\upsilon}dl=n$$

谐振的条件是 l=nλ，即电子轨道的周长是相波波长的整数倍。

2．麦克斯韦分布的证明[10]

在第三篇物质波论文中，德布罗意把相波假设应用于气体系统，用他的统计平衡的新概念证明了麦克斯韦分布。他认为气体原子象电子一样也具有波粒二象性。每一具有速度βc 的原子可以视之与一波包相联系，这一波包的相速度$\upsilon = c/\beta$，频率为$\frac{1}{h}\frac{m_0c^2}{\sqrt{1-\beta^2}}$，群速度$u=\beta c$。除非所有与原子对应的波构成驻波系统，气体的状态不可能是稳定的。他采用金斯的计算方法，得出在单位体积内，在 v 到 v+dv 的频率间隔内连续的驻波数为

$$n_v dv = \frac{4\pi}{u\upsilon^2}v^2dv = \frac{4\pi}{c^3}\beta v^2 dv$$

气体原子的频率 v 和动能 W 的关系式由下式确定

$$hv = \frac{m_0c^2}{\sqrt{1-\beta}} = W - m_0c^2 = m_0c^2(1+a)$$

式中$\alpha = W/m_0c^2$。将以上联立可得

$$n_v dv = \frac{4\pi}{h^3}m_0^2c(1+a)\sqrt{a(a+2)}dW$$

这是在单位体积内，在 W 到 W+dW 间隔内的驻波数。因为每个波可以传输一个、二个或若干个原子，与这些原子相应的能量为 hv ，2hv…nhv。按玻耳兹曼分布，每个波传输的能量为 hv 的原子数为

$$\frac{\sum_1^\infty ne^{-\frac{nhv}{kT}}}{\sum_1^\infty e^{-\frac{nhv}{kT}}}$$

所以在体积元中，动能在 W 到 W+ dW 间隔内能量为 hv 的原子数为

$$dnw = c\frac{4\pi}{h^3}m_0^2c(1+a)\sqrt{a(a+2)}dWdxdydz$$

$$\times\frac{\sum_1^\infty ne^{-\frac{nhv}{kT}}}{\sum_1^\infty e^{-\frac{nhv}{kT}}}$$

式中 c 为常量。对于质量较大因而速度较小的原子所构成的气体，上述级数中除第一项以外，其它各项可以忽略不计，即有 1+α=1，则动能在 W 到 W+ dW 内的原子数为

$$dnw = c\frac{4\pi}{h^3}m_0{}^{3/2}\sqrt{2W}^{-\frac{w}{kT}}Wdxdydz$$

这就得到了麦克斯韦分子按动能分布的规律。

3．电子衍射的预言和干涉现象的解释[9]

在第二篇物质波论文中，德布罗意提出如下假设：自由动点的轨迹是与其相波波射线一致的，即沿着垂直于等相面的方向运动。如果动点从一个线度相当于相波波长的小孔通过，它的轨迹就会象衍射波的射线那样发生弯曲。从而他提出如下预言：“从很小的孔穿过的电子束能够呈现衍射

现象，这或许就是人们能借以寻找关于我们的想法的实验证据的方向。”在这里，德布罗意提示了用电子衍射实验来证实物质波存在的可能性。

德布罗意在此进一步发展了他的相波概念。他把相波看作是导引能量转移的导引波。它的确实存在，允许他进行波与量子的综合。他还指出“自由质点的新动力学与古典力学(也包括爱因斯坦的动力学)的关系，犹如波动光学与几向光学的关系。通过深思熟虑可以看出，我们提出的综合，似乎是从 17 世纪以来动力学和光学的发展比较中得出的一个逻辑结果”。

接着，德布罗意用量子观点解释了干涉现象。他认为“原子具有一种吸引或辐射光原子的概率，这一概率由一个与之交织在一起的相波波矢来决定。自然，只有当原子处于受激状态，才有可能辐射，只有当光原子处于其邻近位置，吸收才有可能发生”。由处于激发态的某一原子辐射的光量子，“其相波在通过邻近原子时会引起其它的量子辐射，假定这种量子辐射的内在振动在相位上与波本身一致，所有被辐射出来的光原子具有与先前一样的相波”，“这些辐射与波耦合，与单一的相波一起传输着”。在博士论文第五章第五节“干涉和相干性”中，德布罗意指出，由于这些波是一样的，这样一个波就能输运很多小的光原子，每个光原子是一个能量凝聚中心，这些能量中心轻轻地在波面上滑动，形成了小块的能量，正如他在《干涉与光量子》一文中所讲的那样，干涉现象与光原子的集合有关。

最后，德布罗意在波动理论的基础上对杨氏双狭缝实验作了量子解释。他说：“有些光原子穿过小孔，沿着环绕这些光原子的那部分相波的射线发生衍射。在屏后面的空间里，随着穿过二个孔而发生衍射的这两股相波，在每一点上所产生的不同的干涉，能得到不同的光电效应。不论射入光的强度小到如何程度，正如波动理论所预见的那样，将出现明暗的条纹。”

(四)物质波思想的影响及实验验证

1．答辩会上提出的问题

由佩兰(Perrin)、郎之万，卡坦(Cartan)和莫格温(Maugain)组成的论文答辩委员会对德布罗意的工作印象极佳。郎之万审阅了德布罗意的论文，他在报告中强调“除了思想的独创性外，德布罗意以非凡的技巧作出努力来克服阻碍物理学家的困难”。[5]他赞赏了德布罗意理论的连贯性、真实性。1972 年莫格温在回忆这次答辩会时说：“今天，我很难理解我在 1924 年的心理状态。当时，我并不相信实际上存在提供这种解释的实体就接受了对这种事实(玻尔-索末菲量子化规则)的解释。”由于当时物质波理论还没有任何实验证据的支持，所以多数考试委员对物质波的真实性存在疑虑。当答辩委员会主席佩兰问道：“这些波怎样用实验来证明？”时，德布罗意回答说：“用晶体对电子的衍射实验是可以做到的。”[2]这是他早已考虑过的方案，并曾向他哥哥的同事道维耶(Dauvilier)(他后来是著名的天体物理学家)提出做这个实验，但因后者忙于其它实验而搁置下来。1973 年 9 月 17 日，德布罗意给美国物理学教授梅迪卡斯的信中说：“在我写论文时，我正在我哥哥的实验室里研究 X 射线，我肯定向道维耶先生提议过，请他用电子进行实验以获得衍射或干涉现象，可是道维耶忙于其它工作，没有按照我的建议去做。”道维耶在纪念德布罗意六十诞辰的一

篇短文中写道：“用以证实电子的这些性质的第一个实验结果是否定的。在这个实验里使用的阴极射线(电子束)太软(即能量太低)。因此高真空中的云母晶体拾取了寄生电荷。”[5]道维耶当时对物质波的存在并无信心，他在 1973 年给梅迪卡斯的信中说：“M．德布罗意、郎之万、佩兰都不关心在他们的实验室里做这种实验，根本没有相信这种实验。”[5]

2．爱因斯坦的有力支持

德布罗意的物质波理论，最初并未受到物理学界的重视。但在他的导师郎之万将论文的复印本寄给了爱因斯坦后，事情起了戏剧性的变化。爱因斯坦向来欣赏物理学中的对称性，而德布罗意理论正是建立了光子和物质微粒之间的这种对称性，所以对德布罗意的想法很感兴趣，非常赞同并竭力支持。1924 年 12 月，爱因斯坦在给郎之万的复信中对德布罗意的工作给以很高的评价，声称这是“揭开了大幕的一角”[4](Hehasliftedacornerof the great veil)。同年 12 月 26 日，爱因斯坦写信给洛伦兹，非常详细地谈到德布罗意的工作：“我们熟知的 M．德布罗意的弟弟已经对于解释玻尔和索末菲的量子规则作了非常有趣的解释。我相信这对于揭露我们物理学中最难以捉摸的谜，开始露出了一线微弱的光芒，我还发现了支持他的解释的一些东西。”[4]次年 1 月，爱因斯坦在柏林科学院的会议周报上发表的一篇论文中写道：“我将很仔细地探讨这个解释，因为我相信它包含了比类推更多的东西。”[2]他在同年 2 月发表的关于爱因斯坦-玻色统计的论文中指出，德布罗意在“非常值得注意的文章中”，把一个粒子归结为一个波场，并认为他的工作不仅是爱因斯坦关于波的二象性的简单类比，而且包含了玻尔和索末菲量子规则的非常卓越的几何解释。在这篇论文的后一部分，爱因斯坦写道：“看来粒子的每一运动都伴随着一个波场，这个场应该是能观察到的。”然后，他谈到了分子穿过小孔的衍射实验，不过他认为在热运动下，小孔必须比分子还要小，因而这种衍射实验实际上是做不到的。[5]爱因斯坦认识到德布罗意工作的重大意义，他对玻恩说：“您一定要读它，虽然看起来有点荒唐，但很可能是有道理的。”[3] 玻恩在 1925 年 7 月 15 日给爱因斯坦的信中写道：“随后我读了 L．德布罗意的论文，并逐渐明白他们搞的是什么名堂，我现在相信物质波理论可能是非常重要的。”

3．对薛定谔的影响

德布罗意的论文经爱因斯坦推荐后，引起物理学界的广泛重视，特别是对奥地利的薛定谔和在哥廷根大学物理系学习的 21 岁大学生埃尔萨塞(Elsasser)产生了深刻的影响。薛定谔也是在郎之万的促使下阅读了德布罗意的论文，他在 1926 年 4 月 23 日给爱因斯坦的信中说：“如果不是你的关于气体简并的第二篇论文硬是把德布罗意的想法的重要性摆到了我的鼻子底下，整个波动力学根本就建立不起来，并且恐怕永远也搞不出来(我指的是光靠我自己)。”[5]

在发表波动力学的一系列论文以前，薛定谔写了一篇关于气体理论的文章，清楚地说明了他对德布罗意假设的看法。他在绪言中说：“这意味着除了认真地考虑德布罗意-爱因斯坦的运动粒子波动理论之外,再无别的途径可言。”该文章详述了爱因斯坦和德布罗意的思想，并且假定处于封闭体积 V 里的分子不能静止不动，否则与该状态对应的波长就会是无限

大，从而最低能态应该是波长为$V^{\frac{1}{3}}$数量级的状态，这是在量子力学建立的过程中第一次讨论零点能。

在 1926 年初，薛定谔发表他的第一篇波动力学论文时，曾清楚地表示他的“这些考虑的灵感主要得自德布罗意先生独创性的论文”。[5]这就说明波动力学是在德布罗意物质波思想的直接影响下所获得的一个丰硕成果。

4．埃尔萨塞的贡献

早在 1921 年戴维孙(Davisson)和孔斯曼(Kunsman)在用电子轰击镍靶时，发现从镍靶上反射回来的二次电子有奇异的角度分布，在不同的角度上出现了两个极大值，而最大值的散射角与电子能量之间有一种关系。他们认为这是原子中出现电子壳层的象征，电子从原子的不同壳层偏转时，其偏转方式是不同的，还没有认识到这是电子衍射现象。这件事引起了哥廷根大学物理教授玻恩的注意。在一次玻恩组织召开的物质结构讨论会上，他的研究生洪德(F.Hund)汇报了自己按戴维孙的电子壳层假设重新计算电子散射曲线的极大值。埃尔萨塞参加了这次讨论会，给他留下了深刻的印象。[12]

受爱因斯坦文章的启发，埃尔萨塞对玻色统计的背景很感兴趣，就到图书馆去借阅德布罗意的论文。德布罗意的物质波假说，大大启发了埃尔萨塞。正好这时玻恩收到了德布罗意寄来的复印本。埃尔萨塞读后马上想到，戴维孙和孔斯曼的极大值和极小值会不会就是电子衍射现象。1925 年 8 月 14 日，埃尔萨塞向《自然科学》杂志提交了一篇短文。他在短文中指出：只要认真运用物质波的概念便能解释戴维孙和孔斯曼的奇怪结果。物质波的波长正好与产生这种反应的大小相当。埃尔萨塞假定电子的透入深度是非常浅的，以致可以假设反射晶面是二维光栅，其光栅常数等于铂的晶格常数，从而计算了反射角的近似值，发现结果正确。爱因斯坦在受托审阅这篇文章时，他签了如下的意见：“当我进行有关玻色气体理论的计算时，在运用理论时并不是那么死板的，不过我感到无论如何要让埃尔萨塞的论文发表。”当埃尔萨塞来访时，爱因斯坦对他说：“年轻人，你正坐在一座金矿上。”[5]

戴维孙在 1927 年 12 月的物理评论上发表的一篇文章中，对埃尔萨塞的工作作了如下的评论：“从电子束对单晶体的散射实验中，可以找到粒子的波动特性的证据，埃尔萨塞在两年以前就预言过了……事实上，埃尔萨塞认为，在那些实验室(于 1923 年)发表的曲线图里，就已经有了这一类证据……我们倒是愿意同意埃尔萨塞对于这些曲线图的解释，但是却不能这么做……散射图上的最大值……我们认为与晶体的结构无关。”戴维孙对于埃尔萨塞的理论并不怎么重视，因此它没有影响他的实验过程。[5]

5．戴维孙和革末的散射实验

最早从实验上证实电子衍射现象是美国的戴维孙和他的合作者革末(Germer)。1925 年，一次偶然事故，使他们的工作呈现新面貌。“正当镍靶处于高温之际，液态空气瓶爆裂了，试验用的管子打破了，靶被进入的空气严重氧化。在采取高温加热的办法净化镍表面后。……氧化物终于减少了，脱了一层靶。但是当他们继续进行实验时，发现散射电子的角分布已经完全改变了。”实验曲线上出现了好几处尖锐的峰值。他们查出这种

变化的原因在于加热过程中，多晶靶重新结晶成少数几块较大的单晶体。但是他们仍然不知道这一现象的本质就是电子衍射。1926 年 8 月 10 日，在牛津召开的一次科学会上，戴维孙听到玻恩讲这样的话：“戴维孙从金属表面散射电子的实验，可能就是德布罗意波的干涉图样。”会议休息时，戴维孙把新近得到的镍单晶实验曲线展示给玻恩等与会的物理学家看，和他们进行了热烈的讨论。他们建议戴维孙研究不久前薛定谔发表的几篇波动力学著作。从此，戴维孙就转入自觉地寻找电子波实验的探索活动之中。他们用的实验装置与现在教科书中所示的一样，研究了散射电流与轰击电压和散射角的关系，肯定了这是电子衍射的结果经过定量计算，证明了德布罗意波长公式的正确性。1927 年 4 月，他们把实验研究结果公布在《自然》杂志上。

6．汤姆孙的衍射实验

同一年，英国物理学家 J.J.汤姆孙之子 G.P.汤姆孙(GeorgePaget Thomson， 1892—1975)也完成了电子衍射实验。他是在德布罗意著作的启发下自觉进行的。他在阿伯丁大学的时候，曾从《哲学杂志》上看到过德布罗意 1924 年写的“光量子试论”一文。他也参加了 1926 年的牛津会议，进一步了解了德布罗意的理论，于是便开始考虑衍射实验。和戴维孙、革末不同，他决定用一种简单、明确和直接的方法，即利用金属薄层做透射实验。这样他便能使用比较容易处理的、能量在 3.9—16.5keV 之间的高能量电子。他所用的第一批薄膜是厚约 3×10^{-6} 厘米的赛璐珞。他和 F.里德很快就观察到了衍射环。根据这些圆环的半径可以计算出电子波的波长。后来他们又研究了电子通过金、铝和铂等金属产生的衍射，从而令人信服地证明了德布罗意波长公式，证实了电子的波粒二象性。戴维孙和汤姆孙同时获得 1937 年的诺贝尔物理奖。[5]

后来，接连发现物质波的衍射现象。1929 年，伊斯特曼和斯特恩成功地研究了原子和分子的衍射。他们把氢分子束和氦原子束对准氟化锂晶体的解理面，发现实验结果定量地与德布罗意公式相符。1936 年，在约里奥实验室工作的冯哈尔巴恩和普赖斯沃克用来自氡(Rn)—铍(Be)源的散射中子在铁上的衍射，获得了中子衍射的实验证据。几十年来，德布罗意的物质波思想不仅成了量子理论的起点，而且还促使人们去进一步做一系列新的实验。

参考文献

[1] Louis de Broglie, The Beginnings of Wave Mechanics, William C.,Wave Mechanics, The First fifty Years, But ter worths, London, 1973 年,12—13

[2] Max Jammer, The Conceptual Development of Quantum Mechanics, MCGRAW-Hill, New York,1966,236—237,239－240,247,249

[3]阎康年，“物质波理论的奠基者——L.V.德布罗意”，《物理》，1982 年 12 月，

[4]Jagdish Mehra, Helmut Rechenberg, The Historical Development of Quantum Theory,VoL. I part2,Springer-Verlag,New york,1982,582-587,592-594,602,604

[5]〔美〕梅迪卡斯，闻人军译，“物质波五十年”，《科学史译丛》，

1982 年，第一辑
[6]胡小敏，“ 德布罗意和波粒二象性思想 ”，复旦大学物理系量子物理史讲习班参考资料 II
[7] 王福山，“ 祝贺德布罗意、狄拉克诞辰 ”，《近代物理学史研究》(一)，第二版，上海，复旦大学出版社，1984 年 4 月
[8]德布罗意，《辐射——波和量子》
[9] 德布罗意，《光学——光量子、衍射和干涉》
[10]德布罗意，《物理学——量子、气体运动理论及费马原理》[11]德布罗意，《量子理论的研究》(博士论文)以上四篇来自《量子力学原始论文选编》复旦大学物理系编[12]郭奕玲等编著，《物理实验史话》，北京，科学出版社，1988，219—220

十四、薛定谔方程的提出

薛定谔方程是量子力学的基本方程，它揭示了微观物理世界物质运动的基本规律，就象牛顿定律在经典力学中所起的作用一样，它是原子物理学中处理一切非相对论问题的有力工具，在原子、分子、固体物理、核物理、化学等领域中被广泛应用。本文主要依据薛定谔的《量子化是本征值问题》的前两篇论文及其它有关著作，探讨以下几个问题：波动力学概念是怎样产生的？薛定谔方程是怎样建立的？薛定谔的研究方法有什么特色？波函数的概率解释是怎样形成的？

薛定谔(ErwinSchrödinger，1887—1961)，1887 年 8 月 12 日出生于奥地利首都维也纳。他的父亲是油毡厂主，又是自然科学爱好者，发表过一系列植物遗传学论文，并长期担任维也纳动植物学会副会长。他的母亲出生于一个化学教授家庭。他从小受到良好的家庭教育。11 岁时，就读于维也纳大学一所预科学校，该校偏重于希腊语和拉丁语，薛定谔爱好语言文学，他曾把荷马史诗译成英文。

1906 年至 1910 年，他就学于维也纳大学物理系。在此期间他曾深入研究过连续介质物理中的本征值问题，这对他以后创立波动力学有重大影响。1910 年获得博士学位。毕业后，在维也纳大学第二物理研究所从事实验物理的工作。第一次世界大战期间，他应征服役于一个偏僻的炮兵要塞，利用闲暇时间研究理论物理。战后他仍回到第二物理研究所。1920 年他到耶拿大学协助维恩工作。1921 年薛定谔受聘到瑞士的苏黎世大学任数学物理教授，在那里工作了六年，薛定谔方程就是这一期间提出的。1927 年薛定谔接替普朗克到柏林大学担任理论物理教授。1933 年希特勒上台后，薛定谔对于纳粹政权迫害爱因斯坦等杰出科学家的法西斯行为深为愤慨，移居牛津，在马达伦学院任访问教授。同年他与狄拉克共同获得诺贝尔物理学奖。

1936 年他回到奥地利任格拉茨大学理论物理教授，不到两年，奥地利被纳粹并吞后，他又陷入了逆境。1939 年 10 月流亡到爱尔兰首府都柏林，就任都柏林高级研究所所长，从事理论物理研究。在此期间还进行了科学哲学，生物物理研究，颇有建树，出版了《生命是什么》一书，试图用量子物理阐明遗传结构的稳定性。1956 年薛定谔回到了奥地利，被聘为维也纳大学理论物理教授。奥地利政府给了他极大的荣誉，设定了以薛定谔命名的国家奖金，由奥地利科学院授给。[1]

(一)物理思想背景

薛定谔把原子系统中的能级理论，德布罗意的物质波思想以及哈密顿的经典力学与几何光学的数学相似性思想结合起来，形成了波动力学概念，建立了薛定谔方程。

1．接受原子能级理论

狄拉克(P.A.M.Dirac)在谈到薛定谔方程产生的背景时曾经说“他(薛定谔)曾经告诉我他是如何作出他的伟大发现的。无疑，在他关于光谱的研究中要用到玻尔的轨道理论，但他总是感到这个理论中的量子化条件是不能令人满意的，原子光谱实际上应当由某种本征值问题所决定。”[2]这就是说薛定谔接受玻尔关于原子系统中电子的能量是分立的概念，认为正确

的原子力学理论必须能够给出分立的能量谱值，但是他对玻尔的量子条件假设始终感到难以接受，他在《量子化是本征值问题Ⅱ》的论文中，认为玻尔的这些“量子条件具有完全奇怪的和不可理解的性质”。应该取代为在满足波函数单值、有限、连续的条件下以及一定的边界条件下求解波动方程的本征值，从而求得一组分立的本征频率。“整数性是这一方法的自然结果”，而不是人为规定的。

2．接受德布罗意物质波思想

薛定谔是怎样接受德布罗意的物质波思想建立其波动力学的呢？薛定谔于 1926 年 4 月 23 日给爱因斯坦的信中这样写道：“假如不是因为你的气体简并论文的第二篇文章把德布罗意思想置于我的面前，如果单靠我个人，很难想象波动力学会建立起来，甚至有可能永远出不来。”[3]这句话给我们提供了薛定谔接受德布罗意思想的一个重要线索，他是通过量子统计的研究接受德布罗意思想的。

薛定谔是在玻耳兹曼毕生工作过的维也纳大学物理系渡过他的大学生活的。他入学的那年 9 月，坚持原子论的玻耳兹曼由于长期受反对原子论的马赫、奥斯特瓦尔德等的攻击，神经有点失常而自杀去世。维也纳大学是玻耳兹曼思想占统治的地方。薛定谔虽然没有直接受学于玻耳兹曼，但也深受玻耳兹曼思想的影响。他曾回忆说：“这里所发生的许多事情都使我直接感知到这位伟人(指玻耳兹曼)所创立的思想。它激发了我早年对科学的热爱，与此相比再也没有其他人的思想能使我这样欣喜若狂。”他在《自然科学的精神》一书中，把玻耳兹曼统计力学和达尔文进化论并列为 19 世纪两大科学发现。正是由于玻耳兹曼的影响，薛定谔早年从事研究的是气体动理论和统计力学问题。

1924 年前后，薛定谔和普朗克、爱因斯坦等人就量子统计问题展开了深入的研究。这年夏天，爱因斯坦收到印度物理学家玻色寄给他的包含着一种新的统计方法的文章。爱因斯坦肯定了他的工作，并着手把这一方法推广到单原子气体上去。正在这时，德布罗意有关物质波的论文通过朗之万送到爱因斯坦手中。德布罗意提出，应该把光量子的波粒二象性假设推广到更大范围中去，即物质微粒运动时都有一定的波相伴随。在那时爱因斯坦正在研究玻色的统计方法，自然对德布罗意的论文很感兴趣，并在 1924 年 12 月发表的关于理想气体量子统计的第二篇论文中引用了它，提出了一种简并气体的新的统计方法，即现在所说的玻色-爱因斯坦统计法。[4]

这时，薛定谔已写过一篇有关气体简并和平均自由程的文章，详细评述了理想气体熵的计算问题。他在读了爱因斯坦的文章后，于 1925 年 2 月 5 日到 1926 年间，同爱因斯坦通信共有 13 封，讨论统计理论中的有关问题。前 9 封写于薛定谔发表波动力学论文以前，对于了解薛定谔建立波动力学的思想过程很有参考价值。薛定谔在 1925 年 2 月 5 日给爱因斯坦的信中谈到不理解爱因斯坦-玻色统计方法,认为量子理论中玻耳兹曼的方法也应是完全有效的。根据这个方法爱因斯坦的计算有误。同年 2 月 28 日爱因斯坦给薛定谔回了信，详细说明了自己的计算过程并没有错，只是所用的方法和玻耳兹曼的方法不同，但薛定谔仍难于接受，迟迟没有给他回信。

1925 年 7 月，普朗克代薛定谔向普鲁士科学院递交了一篇论文《理想气体统计熵定义的评述》。在评述中他认真考虑了爱因斯坦的理论，认为

如果气体分子与光量子在统计学上的行为一致,就势必导致玻色-爱因斯坦统计法。但他并没有接受这种观点。爱因斯坦在久未收到薛定谔复信的情况下，于 1925 年 9 月 28 日主动地给薛定谔写了一封信，肯定了薛定谔对于他的理想气体熵的批评。11 月 3 日薛定谔在回信中有这样一段话：

“几天前，我怀着极大的兴趣拜读了德布罗意的独创性论文，并且终于掌握了它。有了这篇论文，我就第一次清楚理解了您的论文中第八节的内容。……按我个人的意见，德布罗意在数学技巧上的处理和我过去的工作也差不多，只是稍为正规些，却并不那么优美，更没有从普遍性上加以说明。德布罗意能从一个巨大理论的框架上全面地思考问题，这一点确实比我高明，那是我过去所不知道的。”

这封信清楚地表明，薛定谔已经认识到只有德布罗意的理论才为理解光的波粒二象性、理解原子理论所遇到的困难，提供了新的“理论框架”。从信中还可以看出薛定谔对德布罗意的工作并不完全满意，认为还没有从普遍性上说明问题，表明了自己要从这方面进一步发展这一理论的意向。也许这就是薛定谔创建波动力学的起点。

1925 年 12 月 15 日，薛定谔发表了《论爱因斯坦的气体理论》一文。文章一开头薛定谔就指出如果不考虑德布罗意-爱因斯坦的运动粒子波动理论，解决问题的其它途径是没有的。在这篇论文中，薛定谔第一次运用德布罗意的相波思想，把理想气体作为相波谐波系统来处理。他提出不仅把玻色方法应用于光量子可得普朗克公式，而且把玻耳兹曼方法应用于德布罗意相波，同样可得普朗克公式。关键是对物质本性的理解。这篇文章表明薛定谔不仅已经接受了德布罗意的理论，而且已经将这一理论用来研究自由粒子的运动。[3]

不仅如此，在这一年中薛定谔还力图把这一理论用于研究束缚粒子的运动，寻找物质波所满足的普遍方程式，结果得出一个满足相对论要求的波动方程(即克莱因、戈登方程)。他把它应用于氢原子中的电子，发现所得结果与实验不一致。薛定谔极为失望，以为自己的方法错了，因而放弃了它。后来克莱因、戈登也得出了这一方程，并在薛定谔之前发表了它。现在我们知道薛定谔的方法并没有错，结果与实验不一致的原因在于没有考虑电子自旋，而自旋在当时还是一个极新的概念，薛定谔可能还没有听到过。[5]

当时在苏黎世除了苏黎世大学外，还有一所更有名的苏黎世理工学院，德拜(Debye)在那里做教授。德拜并主持一个两校人员都参加的学术讨论会。据德拜当时的学生后来成为著名的物理学家的布洛赫(Bloch)回忆：一次讨论会后，德拜要薛定谔报告当时深受注意的德布罗意的论文。在下次讨论会上，薛定谔对德布罗意如何把波与粒子联系起来，以及他又如何根据在稳定轨道上传播的波其轨道的圆周长应是波长整数倍的要求，从而获得玻尔和索末菲的量子化规则，作了清晰而漂亮的说明。在薛定谔报告后德拜评论说，讨论波动而没有一个波动方程，太幼稚了。几个星期后，薛定谔又作了一次报告，他开头就说：“我的同事德拜提议要有一个波动方程，好，我已经找到了一个。”[6]

3．哈密顿光学力学类比的启示

早在 19 世纪 30—40 年代，哈密顿通过光学力学类比发现了经典力学与几何光学的数学相似性。一百多年来这一重要概念几乎完全被忽略了。

一个值得注意的例外是 19 世纪末克莱因(FelixKlein)反复强调哈密顿思想的重要性，讨论了力学与光学的对应关系。薛定谔在他的论文《量子化是本征值问题》(第二部分)中，一开始就指出：“哈密顿理论和波传播过程之间的内在联系完全不是一个新的概念，这个概念不仅哈密顿本人是熟知的，而且还是他的力学理论的出发点。”他强调“这是一个十分有效的重要概念”。是形成他的波动力学概念的渊源之一。我们可用图 14-1 来表示波动力学产生的物理思想背景。

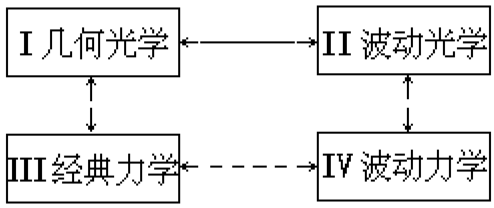

图 14-1

I 和 II 之间的过渡关系早已十分清楚。I 和 III 之间的对应关系，即费马原理与莫培督的最小作用量原理之间的密切相似性早在哈密顿时代已有了完善的理论。薛定谔进一步发展了哈密顿的几何光学与经典力学的类比，提出了微观力学过程是波动过程的论断，形成了波动力学的概念。他从波动光学与波动力学的物理相似性出发，实现了由波动光学向波动力学的过渡。

COLLECTED PAPERS
ON
WAVE MECHANICS
BY
E.SCIIRÖDINGER

TRANSLATEDFROM
THE SECOND GERMAN EDITION

BLACKIE & SON LIMITED
LONDON AND GLASGOW
1928

图 14-21928 年出版的薛定谔的《波动力学论文集》一书的封面

WAVE MECHANICS
Quantisation as a Problcm of
Proper Values(Part I)
(Annalcn der I’hysik(4),vol.79,1926)

§I In this paper I wish to consider, first, the simple case of the hydrogen atom (non-relativeistieand unperturbed), and show that the customary quantum conditions can be replaced by another postulate, in which the notion of “whole numbers”, merely as such, is not introduced. Rather when integralness dose appear, it arises in the same naturnl way as it does in the ease of the node-numbers of a vibrating string. The new conception is capable of generalissation, and strikes, I believe, very deeply at the true nature of the quantum rules.

The usul form of the true nature of the latter is conneeted with the Hamilton-Jacobi difierential equation,

$$\text{(I)} \qquad H(q, \frac{\partial S}{\partial q}) = E.$$

A solution of this equation is sought as can be represented as the sum of functions, each being a runction of one only of the independent varinbles q.

Here we now put for S a new unknown Ψ such that it will appear as a produd of related functions of the single co-ordinates, i.e. we put

(2)

The constant K must be introduced from considerations of dimensions; it has those of aclion. Hence we get

$$\text{(I')} \qquad H\left(q, \frac{K}{\psi}\frac{\partial \psi}{\partial q}\right) = E$$

Now we do not look for a solution of equation(I'), but proceed as follows. If we negleet the relativistie variation of mass, equation(I') can always be transformed so as to become a quadratic form (of Ψand its first derivatives) equated to zero.(For the onc-clectron problem)

(D 804) B

图 14-3 1926 年薛定谔写的《量子化是本征值问题》(第一部分)的第一页

(二)氢原子定态薛定谔方程的提出

1926 年 1 月薛定谔在德国的《物理学杂志》上发表了他的论文《量子化是本征值问题》(第一部分)。[1]他由经典力学的哈密顿-雅可比方程和变分方法建立了氢原子的定态薛定谔方程，得出量子化是本征值问题的结论。在哈密顿函数 H 不显含时间 t 的情形下，H 等于力学体系的总能量 E，即

$$H\left(q, \frac{\partial S}{\partial q}\right) = E \qquad (14.2.1)$$

式中q是广义坐标，$\frac{\partial S}{\partial q}$等于广义动量p，作用函数S是q的函数。他想为方程寻求一个解，而这个解可以表示为一组函数的总和，其中每个都是单个独立变量的函数。为此，他对作用函数 S 引进一个新的未知函数ψ，ψ表示单个独立变量 q 的函数的乘积，令

S=klogψ(14.2.2)

式中 k 具有作用量的量纲，因此得

$$H\left(q，\frac{k}{\varphi}\frac{\partial\varphi}{\partial q}\right)=E \qquad (14.2.1')$$

可以把上式变成ψ和它的一阶导数的二次式等于零的形式。选取直角坐标系，在电子绕核运动的情形下，动能

$$\begin{aligned}T &= \frac{1}{2m}\left(p_x^2+p_v^2+p_z^2\right)\\ &= \frac{1}{2m}\frac{k^2}{\varphi^2}\left[\left(\frac{\partial\varphi}{\partial x}\right)^2+\left(\frac{\partial\varphi}{\partial y}\right)^2+\left(\frac{\partial\varphi}{\partial z}\right)^2\right]\end{aligned}$$

势能 $$V=-\frac{e^2}{r} \qquad (r=x^2+y^2+z^2)$$

哈密顿函数

$$\begin{aligned}H &= T+V\\ &= \frac{1}{2m}\frac{k^2}{\varphi^2}\left[\left(\frac{\partial\varphi}{\partial x}\right)^2+\left(\frac{\partial\varphi}{\partial y}\right)^2+\left(\frac{\partial\varphi}{\partial z}\right)^2\right]-\frac{e^2}{r}\\ &= E\end{aligned}$$

由此得

$$\left(\frac{\partial\varphi}{\partial x}\right)^2+\left(\frac{\partial\varphi}{\partial y}\right)^2+\left(\frac{\partial\varphi}{\partial z}\right)^2-\frac{2m}{k^2}\left(E+\frac{e^2}{r}\right)\varphi^2=0 \qquad (14.2.1'')$$

现在我们来找这样一个实数的，在整个位形空间中，单值的、有限的和两次连续可微的函数ψ，它使上式的左端部分对整个位形空间的积分变为一个极值。薛定谔就用这个变分

$$\begin{aligned}\delta J &= \delta\iiint dxdydz\left[\left(\frac{\partial\varphi}{\partial x}\right)^2+\left(\frac{\partial\varphi}{\partial y}\right)^2+\left(\frac{\partial\varphi}{\partial z}\right)^2-\frac{2m}{k^2}\left(E+\frac{e^2}{r}\right)\varphi^2\right]\\ &= 0 \qquad (14.2.3)\end{aligned}$$

来代替量子化条件。他用通常的方法得到

$$\frac{1}{2}\delta J=\int df\delta\varphi\frac{\partial\varphi}{\partial n}-\iiint dxdydz\delta\varphi\left[\nabla^2\varphi+\frac{2m}{k^2}\left(E+\frac{e^2}{r}\right)\varphi\right]=0 \qquad (14.2.4)$$

因此，第一必须使

$$\left[\nabla^2\varphi+\frac{2m}{k^2}\left(E+\frac{e^2}{r}\right)\varphi\right]=0 \qquad (14.2.5)$$

ψ只是坐标的函数，方程不显含时间，薛定谔把这一方程称为与哈密顿-雅可比方程相对应的波动方程。这就是现在所说的氢原子的定态薛定谔方程。第二必须使沿着无限远封闭曲面进行的积分满足

$$\int df\delta\varphi\frac{\partial\varphi}{\partial n}=0 \qquad (14.2.6)$$

式中 df 是闭合曲面上的面积元。

为了求解方程(14.2.5)，薛定谔引入了球坐标(r、θ、ψ)，写出$\psi = x(r)u(\theta,\varphi)$，求出u是球面谐函数$P_l^m(\cos\theta)\cos m\varphi$或$P_l^m(\cos\theta)\sin m\varphi$。在这里 0≤m≤l，而 l=0，1，2…。m 值的限制条件是根据ψ对极角的单值关系而得出的。他据式(14.2.5)得出 x(r)满足的方程

$$\frac{d^2x}{dr^2}+\frac{2}{r}\frac{dx}{dr}+\left(\frac{2mE}{k^2}+\frac{2me^2}{k^2 r}-\frac{n(n+1)}{r^2}\right)x=0 \qquad (14.2.7)$$

式中 n=0，1，2，3…。对于每个正的 E 值，方程(14.2.7)都有解，这些解在整个空间是单值、有限和连续的，而且在无穷远处它们在不断振荡，以$\frac{1}{r}$的方式趋向于零；然而对于负的E值，只有当E满足

$$\frac{me^2}{k\sqrt{-2mE}}=n$$

时，这个解才唯一的存在。这个分离的本征值谱结果是

$$E=\frac{-me^4}{2k^2n^2} \qquad n=1,2\cdots \qquad (14.2.8)$$

如果给具有作用量量纲的常数k以数值$k=\frac{h}{2\pi}$就得到大家熟悉的氢原子的玻尔能谱

$$E=\frac{-2\pi^2me^4}{h^2n^2} \qquad (14.2.8'')$$

接着，薛定谔谈到函数ψ的意义，得出量子化是本征值问题的结论，从而取代了原来的玻尔-索末菲量子化条件。他写道：“当然，函数ψ可以和原子中一个振动过程联系起来，而这种振动过程比之今天时常对之怀疑的电子轨道要更接近于真实并且清楚得多了。我原先有这个意图，想用这种更直观的方式来建立量子法则的新见解，但最后还是宁愿采取上述不偏不倚的数学形式，因为它能把本质的东西更清晰地揭露出来。在我看来，本质的东西似乎在于量子法则中不再出现神秘莫测的“整数性要求”，而是把这个要求向前更推进一步，它的根源在于某个空间函数的有限性和单值性。”

最后，薛定谔对玻尔的频率条件作了新的解释。根据玻尔的频率条件，原子发射的频率正比于E的差值，即$v=\frac{E_1}{h}-\frac{E_2}{h}$。他把这一频率条件设想为两个频率产生的差频。他说：“至于能量从一个简正振动转移到另一个简正振动时会出现某些东西——我指的是光波，而且作为它所属的频率应是那两个频率之差，是很可以理解的。”他强调指出：“在量子跃迁中能量从一种振动方式转移到另一种振动方式的这种想法，比起电子跳跃的想法不知要合情合理多少。振动方式的改变可以在空间上和时间上连续进行，发射过程持续多长时间可似随它所欲，就象实验中观测到的那样。”

(三)力学与光学的类比

1926 年 2 月，薛定谔发表了论文《量子化是本征值问题》(第二部分)。[8]他通过力学与光学的类比，形成了波动力学的概念。

1．哈密顿-雅可比方程与惠更斯原理的类似性

首先，他考虑含时的哈密顿-雅可比方程

$$\frac{\partial W}{\partial t} \quad (14.3.1)$$

式中W是作用函数，即体系的拉格朗日函数对路径的时间积分$W=\int_{t_C}^{t}(T-V)dt$，它是始点和终点位置和从始点到终点的时间 t 的函数；q_k代表位置坐标；T 是动能，是坐标和动量的函数，对动量来说是二次式，动量表示为 W 对 q_k的偏导数；V 是势能。为解这个方程，令

$$W=-Et+S(q_k) \qquad (14.3.2)$$

$$2T\left(q_k，\frac{\partial W}{\partial q_k}\right)=2（E-V） \qquad (14.3.1')$$

式中 E 是积分常数，即体系的能量。$S(q_k)$是与时间无关的坐标的函数。薛定谔为了把式(14.3.1')简单地表示出来,他在位形空间中利用体系的动能引入一个非欧几里德度规。令动能 T 是速度 q_k的函数，并令线元

$$(ds)^2=2\overline{T}(qk,qk)(dt)^2 \qquad (14.3.3)$$

q 空间中“线段”的量纲不是长度，而是长度乘以质量的开方。

$\frac{\partial W}{\partial q_k}$是作为变量进入式(14.3.1')中去的，它是矢量gradW的分量，

于是方程(14.3.1')等价于这样一种简单的表述

$$(\mathrm{grad}W)^2=2(E-V) \qquad (14.3.1'')$$

或 $|\mathrm{grad}w|=\sqrt{2(E-V)} \qquad (14.3.1''')$

这一要求是容易剖析的，假定由式(14.3.2)所表示的函数 W 已经找到，它满足这一要求，那么这个函数在每一个确定的时刻 t 的值都可清楚地表示出来，只要把 W=常数的曲面族在 q 空间中表示出来，且其中每个曲面都有一个对应的 W 值即可。如果已经知道这一族曲面中任意的某个曲面及其 W 的值，那么方程(14.3.1''')将给出构造这一曲面族中所有其它曲面，并给出一个获得 W 值的确切的规则。

图 14-4

他考虑这一构造规则。令 W_0是图 14-4 中任一曲面的 W 值，取曲面任一侧为正方向。为了找出曲面 W_0+dW_0，作曲面上每一点的垂线，在 dW_0 的正方向截取线段

$$ds=\frac{dW_0}{\sqrt{2(E-V)}} \qquad (14.3.4)$$

这些线段终点的轨迹就是曲面 W_0+dW_0，整个曲面族可按类似的方法在两侧依次构造出来。

现在让我们考虑简单的随时间变化的情况，对这种情况，式(14.3.2)表明，在任一较迟时刻 t+t’，标明 W 分布的仍是相同的一群曲面，尽管每个曲面都有不同的 W 值，但它们都比 t 时刻的值减少 Et’。W 的值好象按一

个确定的简单的规律，从一个曲面到另一个曲面在移动着。当 E 为正时，W 向增加方向移。不过我们也可作另一种设想，即不是 W 的值在不同曲面上移动，而是每一个曲面带着它自己的 W 值在移动，连续地占据下一个曲面的位置和形状。这种移动的规律，可根据这样的事实给出：曲面 W_0 在 t+dt 时刻必须运动到在 t 时刻的曲面 W_0+Edt 所占据的位置。这规律可由(14.3.4)式得到，只要曲面 W_0 上的每个点在垂直于此曲面的正法线方向在 dt 时间内运动的距离为

$$ds = \frac{Edt}{\sqrt{2(E - V)}} \qquad (14.3.5)$$

这就是说，曲面运动的法向速度为

$$u = \frac{ds}{dt} = \frac{E}{\sqrt{2(E - V)}} \qquad (14.3.6)$$

当常数 E 给定时，u 只是位置的函数。

接着，薛定谔作了力学和几何光学之间的类比，指出这两者之间的类似性。他写道 W=常数的曲面族可看作是 q 空间中稳定行波的波面族，这种行波在每一点的相速度由(14.3.6)式给出。法线的构造显然可以用惠更斯子波(具有半径 ds)及其包络面的构造来取代。“折射率 ’正比于式(14.3.6)的倒数，和空间位置有失，但和方向无关，所以 q 空间在光学上是非均匀的，但是是各向同性的，子波是球面波，作用函数扮演了这一波系中相位的角色，哈密顿-雅可比方程是惠更斯原理的表示式。

2．哈密顿变分原理与费马原理的一致性

薛定谔把费马原理表示为

$$O = \delta\int_{p_1}^{p_2}\frac{ds}{u} = \delta\int_{p_1}^{p_2}\frac{ds\sqrt{2(E-V)}}{E} = \delta\int_{t_1}^{t_2}\frac{2T}{E}dt = \frac{1}{E}\delta\int_{t_1}^{t_i}2Tdt \qquad (14.3.7)$$

就得到莫培督给出的那种形式的哈密顿原理。因此“光线”即垂直波面的射线是能量为 E 的体系的运动路径，符合人们熟知的方程组

$$p_k = \frac{\partial W}{\partial q_k} \qquad (14.3.8)$$

由上式可断定动量 p_k 是作用函数的梯度，它说明体系的一组运动路径能够从每个特定的作用函数推导出来。

3．力学体系运动点与波包的等同性

薛定谔指出力学体系形象点(imagepoint)的运动不是以波速 u 沿波面法向运动，相反它的速度对恒定的E是正比于 $\frac{1}{u}$，由式（14.3.3）可以得出

$$\upsilon\frac{ds}{dt} = \sqrt{2T} = \sqrt{2(E - V)} \qquad (14.3.9)$$

u 和 v 之间的不一致是可以理解的。由式(14.3.8)可知在梯度 W 最大的地方，即 W 曲面密集在一起，也就是 u 小的地方，力学体系点的速度大。其次从 W 是拉格朗日函数对时间的积分这一定义可知，当粒子运动时，W 就在改变，因此粒子不能持续地停留在同一个 W 面上。

薛定谔假定 q 空间中的 W 波是正弦波，它的宗量是 W 的线性函数，因为正弦函数的相位的量纲为零，而 W 的量纲是作用量，因此 W 的系数必须具有作用量倒数的量纲，这样波函数ψ的时间因子即为

$$\sin\left(\frac{2\pi W}{h}+\text{常数}\right)=\sigma\iota\nu\left(\frac{-2\pi E\tau}{\eta}+\frac{2\pi S(q_k)}{h}+\text{常数}\right) \qquad (14.3.10)$$

因此波的频率$v=\dfrac{E}{h}$ (14.3.11)

这样我们就自然得到了在 q 空间中波频率正比于体系能量这个结论。由式(14.3.6)和(14.3.11)得波长为

$$\lambda=\frac{u}{v}=\frac{\eta}{\sqrt{2(E-\varsigma)}} \qquad (14.3.12)$$

由式(14.3.6)、(14.3.9)、(14.3.11)可立即证明粒子的速度为

$$\upsilon=\frac{dv}{d(\frac{v}{u})} \qquad (14.3.13)$$

即粒子的速度是在一个小的频率范围内所包含的波的群速度。这与德布罗意在他的相波理论中根据相对论导出的结果是一致的。

他据此阐述了波包与力学体系运动点之间的联系。他写道：“我们能够利用这一事实，在波的传播和代表点运动之间建立一种比以前可能得到的更加自然的关系。我们打算建立一个在各个方向上尺度都相当小的波包。假定这个波包所服从的运动规律和代表力学体系的一个形象点的运动规律相同。只要我们能把波包看作是近似地局限在一个点上，即只要和体系轨道的尺度相比较，我们就能忽略波包的任何扩散，那末，可以说波包就和代表力学体系运动的点是等价的，这只有当轨道的尺度，特别是轨道的曲率半径比波长大得多时才行。”

(四)波动力学概念的形成

薛定谔还把经典力学的限度和几何光学的限度作了类比，从而大胆地提出了波动力学存在的猜测。他写道：“事实上我们现在知道，经典力学在轨道线度很小和曲率很大的情况下是不适用的，也许力学的这种失效和几何光学，即‘波长为无穷小的光学’的失效是十分相似的。几何光学的失效在障碍物或孔径不再比真实的有限大小的波长大时，显得特别明显，也许我们的经典力学和几何光学完全类似，象几何光学一样出错，不符合真实情况；一旦轨道的曲率半径和线度比某个波长小时，经典力学就不正确了。在 q 空间中，这种波长是具有实际意义的，于是提出了寻找波动力学的问题。”

接着他在讨论波包的限制条件时，提出了微观力学过程是波动过程的重要论断，他写道：“在 q 空间中，真正的力学过程是通过一种适当的波动过程，而不是由形象点的运动来实现或表示出来的。作为经典力学的对象，形象点运动的研究只是一种近似处理，正如几何光学或‘射线’光学比之于真实光学过程一样。和轨道的几何结构相比，上面所描述的那类波信号能足够近似地被认为是局限于一点的，这种波信号就可描述宏观力学过程。我们已经看到，这样一种信号或波包的运动规律，和经典力学形象

点的运动规律是严格相同的。然而当轨道的形状和波长比较不再是很大或实际上差不多时，这种处理方式失去了全部意义，那时我们就必须严格地按波动理论来处理，即为了形成各种各样可能过程的一个图象，我们必须按波动方程，而不是按力学的基本方程来处理，后者对解释力学过程的微观结构是无效的，正如几何光学在解释衍射现象时的情况一样。”

最后，薛定谔指出在微观领域中粒子路径没有意义。他写道：“让我们考虑一个具有上面所述的那种性质的波包，这个波包以某种方法进入一个小的其尺度是波长数量级的闭合的“路径”(Path)中去，因此小到可和波包本身的尺度作比较，那么很清楚，在经典力学意义上的‘体系的路径”，即等相位点的路径，将完全失去其路径的特性，因为在这个等相点的前后及其附近区域，有一片连续的点，其相位几乎是完全相等的，然而它们表现完全不同的“路径”，换句话说，波包不仅立即充满整个路径区域，而且也要向四面八方伸展出去。”

在这个意义上，他进一步说明电子轨道没有意义。他写道：“根据德布罗意的理论，这种相波是和电子轨道伴随在一起的。因此电子路径本身不具有空间的意义(至少在原子内部是如此)，电子在路径上的位置更是没有意义的。在我阐明的这个意义上，如今更加明显确信：首先，原子中电子运动的相位的真实意义必须放弃；其次，我们不能由量子条件断言在某一定的时刻，可在某个确定的量子轨道上发现电子；第三，正确的量子力学定律，不是由单一路径的明确的定则所组成的，而是通过一些方程式，把一个体系的各种各样的路径都结合起来，因此在不同的路径之间，明显地存在某种相互制约的作用。”由此他得出结论：“所有这些断言一致要求我们放弃‘电子的位置’和‘电子的路径’的概念。”

(五)含时的薛定谔方程的提出

既然如此，在微观领域中，对于必须用波动来表示的情形，我们现在应遵循哪一条途径来描述力学问题呢？薛定谔明确指出不能从力学基本方程出发，而必须从 q 空间中的波动方程出发，他指出波函数必须满足普遍的波动方程

$$\nabla^2 \boldsymbol{\psi} - \frac{1}{u^2}\boldsymbol{\psi}'' = 0 \qquad (14.5.1)$$

因此，由式(14.3.6)、(14.3.10)、(14.3.11)可得

$$\nabla^2\varphi + \frac{8\pi^2}{h^2}(hv - V)\varphi \qquad (14.5.1')$$

或 $$\nabla^2\varphi + \frac{8\pi^2}{h^2}(E - \varsigma)\varphi = 0 \qquad (14.5.1\forall)$$

写成现在形式 $$-\frac{\eta^2}{8\pi^2\mu}\nabla^2\varphi + \varsigma\varphi = E\varphi \qquad (14.5.1''')$$

这个方程和第一篇论文中的结果是一致的。

最后，薛定谔作出结论再一次阐述了量子化就是本征值问题的含义。他写道：“关于把方程(14.5.1)作为原子力学的基础所表现出来的担忧，我不想明确地断定它不需有进一步的附加限制。但是它们也许不再象以前‘量子条件’那样，具有完全奇怪的和不可理解的性质，而是成为我们所

熟悉的在含有偏微分方程的物理问题中寻找初条件或边值条件的问题，它们决不和量子条件类似——因为在我迄今所已经研究的经典动力学的全部情况中，证明方程(14.5.1)本身内部蕴藏着量子条件。在某些情况下，以及在实验所确实要求的地方，除了十分明显的要求，即作为一个物理量，函数ψ在整个位形空间中必须单值、有限和连续外，方程(14.5.1)不需要任何进一步假定，可以自行分辨出某种频率或能级，这种频率或能级在定态过程中是唯一可能的。”他强调指出：“量子能级可立即由方程(14.5.1)本身所具有的固有边界条件所决定的本征值来确定。”

文章的后一部分，薛定谔把他的理论应用到线性谐振子、定轴转子、绕自由转轴的刚性转子以及非刚性的振动转子(双原子分子)上，举了一系列实例进行分析。

1926 年 5 月，薛定谔发表了论文《量子化是本征值问题》(第三部分)，阐述了与时间无关的微扰论。第一次提出波动力学(wavemechanics)一词。

1926 年 6 月，薛定谔发表了论文《量子化是本征值问题》(第四部分)[9]，阐述了含时的微扰论。他认为不含时的方程(14.5.1)不具有普遍性。确切地说不能把它称为波动方程，而把它称为振动方程或振幅方程更恰当些。他认识到在方程(14.5.1)中的本征值 E 在从一个稳定状态到另一个稳定状态的过程中是变化的，但这一变化情况没有出现在上述的波动方程中。为了体现这一变化情况，他把ψ函数中反映空间分布的空间因子和反映时间分布的时间因子分开，使

$$\psi=\psi(q)\exp[2\pi i(E/h)t]$$

对时间求导得 $\frac{\partial\varphi}{\partial t}=\frac{2\pi iE}{h}=\psi$

所以 $E\psi=\frac{h}{2\pi i}\frac{\partial\varphi}{\partial t}$

代入方程(14.5.1'')就得

$$-\frac{\eta^2}{8\pi^2 m}\nabla^2\varphi+\varsigma\varphi=\frac{\eta}{2\pi}\frac{\partial\varphi}{\partial t} \qquad (14.5.2)$$

这就是含时的薛定谔方程。

(六)波函数的物理解释

薛定谔理论中的波函数ψ究竟代表什么？它的物理意义是什么？薛定谔本人还不太清楚。在他的第一篇论文中认为波函数ψ是实在的，反映原子中的振动过程，比电子轨道更接近于真实情况。在他的第二篇论文中，他认为波函数表示的是德布罗意“相波”，他试图以彻底的波动概念来建立一个直观的图象，把粒子看作是波群中的“波包”，但是波包在传播过程中必然会发散，这不符合实际例子的稳定性。在他阐述ψ函数的另一篇文章中，他把$\psi\psi^*$解释为位形空间的权函数(weightfunction)，表示电荷在空间出现的密度，即$\rho=e\psi\psi^*$[4]，但是薛定谔的波函数所表征的是位形空间，即多维相空间中的波，而不是三维空间中的波，无法直接代表物理实在。

1926 年 6 月，玻恩在题为《散射过程的量子力学》的论文中，提出了

波函数的概率解释。他的依据是微观粒子的波粒二象性。他说：“他在普朗克关于原子和分子碰撞的卓越实验中，每天都目睹粒子概念的丰硕成果，因而确信，粒子不能简单地取消。”不能像薛定谔那样放弃粒子概念，要坚持物质的波粒二象性，就要设法寻找物质粒子与波统一的新途径。玻恩认为$|\psi|^2$表示粒子出现的概率。他在文中写道：“更精确的考虑表明，概率正比于ψ的平方。”即当测量以波函数ψ描述的量子态下电子的位置时，它在一点(x，y，z)附近体积元 dV 中发现的概率与$\psi^*\psi dV$成正比。这就是波函数的概率解释[10]。玻恩的这一解释多少有点源出爱因斯坦的“鬼场”(phantomfield)概念，它导引光量子沿某一轨道运动，其波幅的平方为确定发现光量子的密度。毫无疑问，玻恩也接受了薛定谔的权函数$\psi\psi^*$这一数学形式而换上了新的内容。正如玻恩所说：“必须彻底放弃薛定谔旨在恢复经典连续理论的物理图象，只保留这一形式，并充实以新的物理内容。”[11]玻恩的概率解释找到了统一粒子与波的新线索，不久得到了全世界绝大多数物理家的公认。

概率概念第一次以具有原则的意义引入了物理学，玻恩认为波函数是描述单个粒子行为的，他认为在微观领域中概率是第一性的。他说：“我本人倾向于在原子世界中放弃决定论，但这是个哲学问题，对此有判断权威的不光是物理学上的论据。”[11]

爱因斯坦坚决反对量子力学的概率解释。1926 年 12 月 4 日他给玻恩的信中说：“量子力学固然是堂皇的，可是有一种内在的声音告诉我，它还不是那真实的东西。这理论说得很多，但是一点也没有真正使我更加接近‘上帝’的秘密。我无论如何深信上帝不是在掷骰子。”[12]这里所说的上帝的秘密就是自然界中“客观存在的规律和秩序”。爱因斯坦认为波函数描述的不是单个体系而是体系的系综，量子力学是一种统计的理论。单个粒子的运动状态必须是决定性的，不能是统计性的。

薛定谔同样反对概率解释。他认为玻恩的解释是对他的理论的误解。认为电子像跳蚤一样跳来跳去，真令人毛骨悚然。从这时起量子理论沿着一条与薛定谔歧离的途径发展。薛定谔“关注且失望”的是这一“超经验的，几乎是超自然的关于波现象的解释”成为“几乎普遍接受的信条”。[13]

1926 年 9 月，玻尔邀请薛定谔到哥本哈根讲学，在报告的最后薛定谔提出应该放弃量子跃迁观念，而坚持微观世界无所不在的连续性。他说：“如果我们改变以上图景，不把电子看作粒子，而看作电子波。此时，定态轨道概念及其量子跃迁概念就自然消失了，被激发的电子的拍频可以转化为光。”“光发射就象无线电波通过天线发射一样容易解释了。”玻尔驳斥了薛定谔对波函数的物理解释。认为 25 年的历史已证明，普朗克辐射公式不引入能量的分立值及跃迁概念是推导不出来的。荧光屏上观察到的突现的亮点或观察到电子突然通过云室，都使人直接看到原子现象的瞬间跃迁。我们不能无视这些事实。[10]最后，薛定谔懊丧地说：“如果我们一定要坚持这个该死的量子跳跃，那么，我就只能因为曾对量子理论做了一点工作而感到遗憾。”玻尔却回答说：“但是我们却因为你为澄清量子理论作出了这么多的贡献而感谢你。”[13]

(七)薛定谔的科学思想及其理论的影响

薛定谔在自己一生的科学研究中形成了独特的风格，对科学哲学有浓厚的兴趣。他对科学与哲学的关系有很多精辟的论述。在《我的世界观》一文中，他认为哲学是“ 建成知识大厦所必不可少的脚手架。 ”[14]要排除它就意味着“ 抽去艺术和科学的灵魂，将它们置于裹足不前的境地 ”。他在《科学是时代的风尚吗？》一文中强调一个时代的文化背景或时代精神对物理框架、物理思想的影响，他指出科学“ 依赖于它构成某一部分的那个时代风行的精神框架 ”，“ 我们都是我们的文化环境的成员 ”。[15]

薛定谔在科学观上坚持科学理论是对客观实在的描述。他强调科学真理的客观性，“ 一旦这种真理最终得以阐明，即能为世界上的任何人用实验加以检验，并总得到同样的结果 ”。由于精确的实验是物理知识的泉源，他声称物理学是“ 绝对客观真理的载体 ”。

从薛定谔方程建立的过程中我们可以看到他的科学研究方法的一个突出特点是善于运用物理类比。他通过力学与光学类比形成波动力学概念，通过几何光学与经典力学的相似性得出波动光学与波动力学的相似性，从而根据普遍的波动方程建立起了薛定谔方程。他对类比方法在科学创造中的作用有着深刻的看法，他说：“ 我们这些现代知识分子不习惯于把一个形象化的比拟当作哲学洞见，我们坚持要有逻辑推演。但对这种要求，逻辑思维也许只能向我们揭示这么多，要通过逻辑思维来掌握现象的基础，很可能根本做不到，因为逻辑思维本身就是现象的一个部分，和现象完全牵连在一起，既然如此，我们也就不妨问，我们是否仅仅因为一个形象化的比拟不能被严格证明，就逼得不能够运用它呢？ ”[15]

薛定谔的论文发表后，欧洲物理学界为之一震，爱因斯坦和普朗克都支持波动力学，而对矩阵力学持有怀疑态度。1926 年 4 月普朗克在收到波动力学第一篇文章后给薛定谔写信说：“ 我像一个好奇的儿童听人讲解他久久苦思的谜语那样聚精会神地拜读您的论文，并为我眼前展现的美丽而感到高兴。 ”爱因斯坦也认为“ 薛定谔著作的构思证实着真正的独创性。 ”玻恩赞扬薛定谔的工作说：“ 在理论物理学中，还有什么比他在波动力学方面的最初几篇论文更出色的呢？ ”[17]

在 1926 年以前，矩阵力学与波动力学是平行地发展起来的，只是在证明两者的联系以后，才统一为量子力学，一种力学的思想发展对另一种是有影响的。在未统一以前，两门力学的创建者对彼此的工作缺乏理解，互有责难。海森堡曾说：“ 我越是思考薛定谔理论的物理内容，我就对它越讨厌。 ”薛定谔也对矩阵力学提出了批评，认为“ 这种超越代数的方法简直无法想像，它如果不使我拒绝的话，至少也使我气馁。 ”但是，薛定谔并没有拒绝海森堡的论文，而是钻研它。在他的《量子化是本征值问题》(第二部分)中他写道：“ 这里我愿意提及海森堡、玻恩、约当和其他一些著名的学者目前正在深入进行的一项排除量子困难的研究工作。这些研究已经取得如此值得注目的成就。因此它不容置疑地至少含有一部分真理。从这一理论的发展趋向看，海森堡的意图十分接近于现在我们已经论述的这一工作，他们的方法是完全不同的，我还没有找到二种方法之间的联系，我抱有明确的希望，这二个进展将不会互相冲突。相反，正由于它们的出发点和方法截然相异，它们可互相补充，取长补短。海森堡方案的力量在于它能给出谱线强度，这个问题我们还没有接触到。 ”在 1926 年 4 月发表了《关于海森堡-玻恩-约当的量子力学与我的波动力学之间的关系》的论

文，在这篇文章中，薛定谔证实了矩阵力学和波动力学的等价性。指出这两种力学在数学上则是完全等价的，可以通过数学变换从一理论转换到另一理论。

薛定谔理论刚出现时，狄拉克的反应也是否定的，海森堡与他通信中曾问到他对薛定谔理论的看法。他当时认为海森堡的工作已经为量子力学提供了很满意的基础，我们只要继续发展它就很好了，没有任何必要对这个基础作进一步修订。1929 年 5 月下旬，海森堡给狄拉克写信，详细说明薛定谔理论和矩阵力学的关系，才逐渐使狄拉克的看法转变过来，从此他热情地学习薛定谔理论。[5]

今天，薛定谔方程已经成为在原子物理学的世界文献中应用最广泛的公式。MaxJammer 在他写的《量子力学概念的发展》一书中作了如下的评述：“薛定谔的卓越的论文无疑地是科学史上最有影响的论文之一，它加深了我们对原子现象的理解，是求解原子物理问题，固体物理问题，在某种程度上也是求解核物理问题的合适的基础，而最终开辟了新的思路。事实上非相对论量子理论后来的发展只不过是对薛定谔工作的详尽阐述和应用。”[4]

参考文献

[1]周奇，“薛定谔和他的科学成就”，《大学物理》，1987 年 8 月

[2]F.A.M.狄拉克著，胡新和译，“纪念英国皇家学会外籍会员埃尔温、薛定谔教授”，《自然科学哲学问题》，1987 年 4 月

[3]金尚年，“薛定谔和他所建立的波动力学”，复旦大学物理系《量子物理史讲习班参考资料 II》

[4]Max Jammer，The Conceptual Development of QuantumMechanics，MCGraw-Hill，NewYork，1966，248，266-267

[5]周世勋，“量子力学的诞生”，《大学物理》，1982 年 3 月

[6]Felix Bloch，Reminis cences of Heisenberg and theearly days of guantum mechanics，PhysicsToday，Dec. 1976，23

[7]薛定谔，金尚年译，《量子化是本征值问题》(第一部分)，复旦大学物理系《量子力学原始论文选编》

[8]薛定谔，《量子化是本征值问题》(第二部分)，复旦大学物理系，《量子力学原始论文选编》(续)，

[9] SCHRÖdinger，Quantisationas aProblemofProper Values(part ④)，Collectedpapers on WaveMechanics，102—104[10]朱荣华，《物理学基本概念的历史发展》北京，冶金工业出版社，1987 年 11 月，346—347

[11]宋德生，“波动力学和薛定谔”，《自然杂志》1987 年 10 卷 8 期

[12]许良英、范岱年编译，《爱因斯坦文集》，第一卷，第 1 版，北京，商务印书馆，1976 年 1 月 221

[13]申先甲等编著，《物理学史简编》，第 1 版，济南，山东教育出版社，1985 年 1 月，777

[14]薛定谔，“我的世界观”，《自然科学哲学问题》，1987 年 4 月

[15]薛定谔，“科学是时代的风尚吗？”同[14]

[16]黄元生，钱时惕，“薛定谔的科学贡献及思想方法”，《物理通

报》，1987 年 11 月

[17] 史天一，“ 量子力学发展概况 ”，《物理学史专题讲座汇编》，北京物理学会